Study of Vector Boson Scattering with Pile-up with the ATLAS Detector at the LHC

Diplomarbeit
zur Erlangung des wissenschaftlichen Grades
Diplom-Physiker

vorgelegt von

Philipp Anger
geboren am 13.12.1985 in Annaberg-Buchholz

Technische Universität Dresden

Institut für Kern- und Teilchenphysik
Fachrichtung Physik
Fakultät für Mathematik und Naturwissenschaften

2010

CERN-THESIS-2010-141
2010

Eingereicht am 12.11.2010

1. Gutachter: Prof. Dr. Michael Kobel
2. Gutachter: Prof. Dr. Kai Zuber

Abstract

Effects of physics beyond the Higgs-less Standard Model are expected in the vector boson scattering channel. The chiral electroweak Lagrangian extended by additional resonances and unitarized with the K-matrix model can parametrize these effects.
An analysis of the sensitivity of the ATLAS experiment at the Large Hadron Collider at CERN to these resonances is presented. Especially the effects of additional proton-proton interactions per bunch crossing (pile-up) are investigated. A training and evaluation of a multivariate analyzer is performed and used to determine ranges of sensitivity for couplings of the resonances. Systematic uncertainties are taken into account.

Kurzdarstellung

Effekte neuer Physik jenseits des Higgs-freien Standardmodells werden bei der Streuung schwacher Eichbosonen erwartet. Der um zusätzliche Resonanzen erweiterte und mittels K-matrix Model unitarisierte chirale elektro-schwache Lagrangian kann diese Effekte parametrisieren.
Diese Arbeit ist eine Analyse der Sensitivität des ATLAS Experiments am Large Hadron Collider (CERN) für diese Resonanzen. Besonders die Effekte von zusätzlichen Proton-Proton Wechselwirkungen (pile-up) werden untersucht. Eine multivariate Analyse wird angewendet, um die erwarteten Entdeckungssignifikanzen zu bestimmen und experimentell zugängliche Bereiche für die Kopplungen der Resonanzen zu erhalten. Zudem wird der Einfluss systematischer Unsicherheiten betrachtet.

Contents

1 Introduction **1**

2 Theory **3**
2.1 Standard Model . . . 3
2.1.1 Electroweak Theory . . . 5
2.1.2 Electroweak Symmetry Breaking . . . 9
2.2 Effective Theories of Electroweak Interactions . . . 11
2.3 Resonances and Unitarity . . . 13

3 Experiment **17**
3.1 Large Hadron Collider . . . 17
3.2 The ATLAS Detector . . . 18
3.2.1 Detector Description . . . 18
3.2.2 Trigger and DAQ System . . . 22
3.2.3 ATLAS Event Data Model . . . 22
3.3 Pile-up . . . 23
3.3.1 Handling Pile-up . . . 25
3.3.2 Pile-up Simulation . . . 27
3.4 Experimental Signature of Vector Boson Scattering . . . 28

4 Monte Carlo Samples and Discriminating Variables **31**
4.1 ATLAS Monte Carlo Production Chain . . . 31
4.2 Event Generator WHIZARD . . . 32
4.3 Monte Carlo Samples . . . 33
4.3.1 Signal Samples . . . 33
4.3.2 Background Samples . . . 35
4.4 Discriminating Variables . . . 41

5 Multivariate Analysis **45**
5.1 Pre-cuts . . . 45
5.2 Multivariate Methods . . . 46
5.2.1 Toolkit for Multivariate Analysis . . . 46
5.2.2 Boosted Decision Tree . . . 47
5.3 Application of a Boosted Decision Tree . . . 48
5.3.1 Input Samples . . . 49
5.3.2 Input Variables . . . 51
5.3.3 Overtraining Detection . . . 52
5.3.4 Evaluation of the Training Output of a Boosted Decision Tree . 53

5.3.5 Influence of a Boosted Decision Tree on the Angular Separation of the Leptons . . . 56

6 Statistical Analysis 59
6.1 Profile Likelihood . . . 59
6.1.1 p-Values and Equivalent Gaussian Significance . . . 59
6.1.2 Profile Likelihood . . . 59
6.1.3 Connection of Profile Likelihood Ratio to Significance . . . 61
6.2 Application of an Approximate Significance Formula . . . 62
6.3 Application of the Profile Likelihood Method . . . 62
6.4 Effects of Pile-up . . . 65
6.5 Luminosity Studies . . . 66
6.6 Training Crosschecks and Statistical Error Estimation . . . 68
6.7 Systematic Uncertainties . . . 74

7 Conclusions and Outlook 77

A Shape of the Boosted Decision Tree Input Variables 79

B Boosted Decision Tree Cutflow for Different Resonances 85

C Discovery Significance as a Function of Boosted Decision Tree Cut and Coupling for all Resonances 87

List of Figures 91

List of Tables 93

Bibliography 95

1 Introduction

The current theoretical fundament of particle physics is the Standard Model (see Section 2.1). All aspects of this theory are in good agreement with the current experimental data.

Despite the elegance and power of the Standard Model it is not the final theory of matter. Several open questions exist that are impossible to be answered within this model. The most discussed and perhaps one of the most important is how particles acquire mass. The leading idea to answer the question is the Higgs mechanism (see Section 2.1.2).

In the unbroken Standard Model without Higgs mechanism all particles have to be massless. However, we know from experiment that most of the particles found in nature have non-zero masses. This is highly correlated to the mechanism of electroweak symmetry breaking (see Section 2.1.2). Without the mechanism of electroweak symmetry breaking, the scattering of vector bosons violates perturbative unitarity for longitudinally polarized W bosons at about 1.2 TeV center-of-mass energy.

In the Higgs mechanism particles acquire mass by interaction with a scalar Higgs field whose excitations introduce a new particle, the Higgs boson. Electroweak symmetry is spontaneously broken by this new scalar field and unitarity is conserved in the vector boson scattering channel.

One problem remains: Despite great effort the Higgs boson has escaped detection and continues to do so. A conclusive decision about the existence of this particle has been one of the main reasons for building the Large Hadron Collider at CERN. This machine is expected to give the final answer to the question of the existence of the Higgs boson.

If the Higgs boson does not exist, physics beyond the Standard Model is required. This is where this work enters the scene. New physics can be described by an effective approach (see Section 2.2). In the electroweak chiral Lagrangian considered in this thesis, the Standard Model is treated as a low-energy effective theory. Higher orders in perturbation theory enforce additional anomalous couplings as new degrees of freedom. Also resonances that might exist at higher WW center-of-mass energies are introduced. Because unitarity is not assured for all resonances and couplings, the theory is properly unitarized with the K-matrix unitarization scheme (see Chapter 2.3). The search for these kind of resonances present in the vector boson scattering channel is the goal of this work.

An integrated luminosity of $45\,\mathrm{pb}^{-1}$ of data has been taken by the ATLAS detector by the end of this work. However, an integrated luminosity on the order of $100\,\mathrm{fb}^{-1}$ is required to achieve the discovery limits for sufficiently large ranges of the couplings of the resonances. Therefore, this study only deals with simulated Monte Carlo events (see Chapter 4.3) scaled to the required amount of real data.

The double-leptonic final state is investigated in this work. This channel takes

advantage of the very clean signature and reconstruction of leptons by ATLAS. All simulated signal Monte Carlo samples are generated by WHIZARD (see Chapter 4.2). This event-generator has been selected because angular correlations expected in the double-leptonic final state are preserved. Furthermore, WHIZARD is the only generator that implements the K-matrix unitarization with additional resonances.

This study is based on the work of Jan Schumacher [1] where no additional proton-proton interactions per bunch crossing (pile-up) are considered (see Section 3.3). The additional tracks in the detector due to this effect are already present at current instantaneous luminosities and their effect will rise with increasing luminosities while data taking continues. The aim of this work is to take a closer look at the effects of pile-up in the context of resonances in the vector boson scattering channel.

2 Theory

2.1 Standard Model

The subjects of elementary particle physics are the fundamental constituents of matter and the interactions they experience - the basic questions of nature. As far as nature is understood at the moment, there are four basic types of forces interacting with several different types of elementary particles. During the last century many theories have been developed to acquire an understanding of the main concepts and origins of these interactions. The present status of these efforts is the *Standard Model* of particle physics [2,3]. All aspects of this theory are in good agreement with the current reach of experimental data.

According to the Standard Model the elementary particles are classified in fermions and bosons. Fermions are particles with half-integer spin obeying Fermi-Dirac statistics. There are two types of elementary fermions with spin one-half: leptons and quarks. Electrically charged leptons are the electron e^-, the muon μ^- and tau τ^- together with their corresponding antiparticles e^+, μ^+ and τ^+. The associated electrically neutral leptons are the neutrinos: electron-neutrino ν_e, muon-neutrino ν_μ and tau-neutrino ν_τ also with their corresponding anti-neutrinos $\overline{\nu}_e$, $\overline{\nu}_\mu$ and $\overline{\nu}_\tau$. Quarks are non-integer electrically charged particles. Twelve types are part of the Standard Model: up u, down d, strange s, charm c, bottom b and top t with corresponding antiparticles $\overline{u}$, $\overline{d}$, $\overline{s}$, $\overline{c}$, $\overline{b}$ and $\overline{t}$. These fundamental fermions can be grouped into three generations with same properties but different masses and are listed in Table 2.1.

Bosons (see Table 2.2) have an integer spin and obey Bose-Einstein statistics. Because interactions between fermions are described in terms of local gauge theory (see Section 2.1.1), the bosons in the Standard Model are called gauge bosons. They are carriers of the forces and have an spin of one. A special boson is the hypothetical Higgs boson necessary for electroweak symmetry breaking (see Chapter 2.1.2). It would be the only fundamental boson in nature not associated to a force and the only fundamental particle with a spin of zero.

Electrically charged particles are affected by the *electromagnetic force*. The theory of this force, *quantum electrodynamics* [4–6], identifies the photon γ as the force mediating gauge boson. The weak force acts on all fermions via three massive gauge bosons $W^\pm$ and Z^0. No consistent gauge theory describing this force on its own is known, but it was shown by Glashow, Salam and Weinberg [2,3,7] that the electromagnetic and weak theory can be incorporated into one theory, the *electroweak theory*. The third force of the Standard Model, the *strong force*, is mediated by eight gauge bosons called gluons g and described by *quantum chromodynamics* [8–13].

Generation		Fermion	Electric charge	Mass [MeV][1]
1st	e^-	electron	-1	0.511
	ν_e	electron-neutrino	0	$< 2 \times 10^{-6}$ (95% CL) [2] [3]
	u	up	2/3	$2.49^{+0.81}_{-0.79}$ ($\overline{\text{MS}}$, $\mu = 2$ GeV) [4]
	d	down	-1/3	$5.05^{+0.75}_{-0.95}$ ($\overline{\text{MS}}$, $\mu = 2$ GeV)
2nd	μ^-	muon	-1	105.658
	ν_μ	muon-neutrino	0	< 0.19 (90% CL)
	c	charm	2/3	1270^{+70}_{-90} ($\overline{\text{MS}}$, $\mu = \overline{m}_c$)
	s	strange	-1/3	101^{+29}_{-21} ($\overline{\text{MS}}$, $\mu = 2$ GeV)
3rd	τ^-	tau	-1	1776.84 ± 0.17
	ν_τ	tau-neutrino	0	< 18.2 (95% CL)
	t	top	2/3	$(172.0 \pm 2.3) \times 10^3$
	b	bottom	-1/3	4190^{+180}_{-60} ($\overline{\text{MS}}$, $\mu = \overline{m}_b$)

Table 2.1: Fermions of the Standard Model with their electrical charges and experimental masses [14].

Boson		Interaction	Mass [GeV]	Electric charge	Spin
photon	γ	electromagnetic	$< 10^{-18}$	0	1
W bosons	$W^\pm$	weak	80.399 ± 0.023	± 1	1
Z boson	Z^0	weak	91.188 ± 0.002	0	1
gluons	g	strong	0 [5]	0	1
Higgs	H		> 114.4 (95% CL)	0	0

Table 2.2: Bosons of the Standard Model, the interactions they are mediating and experimental masses [14]. The special role of the Higgs boson is explained in Chapter 2.1.2.

[1] In particle physics it is commonly used to set $\hbar = c = 1$ ("natural units"). This convention is used throughout this work.

[2] In the neutrino sector, a mixing between mass and flavor eigenstates similar to the quark sector occurs [15]. If the neutrino mass eigenstates are written as ν_1, ν_2 and ν_3, a mass of the flavor eigenstates ν_e, ν_μ and ν_τ can be defined with the following equation: $m^2_{\nu_\ell} = \sum_i |U_{\ell i}|^2 m^2_{\nu_i}$, $\ell \in \{e, \mu, \tau\}$. Upper limits on the masses of the flavor states are given in the table.

[3] The confidence level (CL) is a qualification of the likelihood of a true parameter to be included inside a given confidence interval [16].

[4] Quark masses are dependent on the renormalization scheme and the renormalization scale μ [17]. The values of the quark masses are given for the modified minimal subtraction scheme $\overline{\text{MS}}$ at the quoted renormalization scale μ.

[5] Theoretical value. A mass of as large as a few MeV may not be precluded [18].

Until now it was not possible to include the fourth force, *gravitation* [19], into this context. The theory of gravitation, Einstein's general relativity [20], describes the gravitational interactions sufficiently precise to understand the current accessible experimental range. Many efforts have been made to unify the Standard Model and general relativity and many new ideas are on the way [21, 22].

The fundamental forces are characterized by three quantities [23]: The strong force by the proton mass ($m_p \approx 1$ GeV), the electroweak force by the Fermi constant ($G_\mathrm{F} \approx 10^{-5}$ GeV^{-2}, see Equation 2.2) and the gravitation by Newton's constant ($G_\mathrm{N} \approx 10^{-38}$ GeV^{-2}). Out of these quantities, three different fundamental scales are derived:

$$\begin{aligned} \text{Proton mass:} \quad & m_p & & \approx 1 \text{ GeV} \\ \text{Fermi scale:} \quad & v = \left(\sqrt{2} G_\mathrm{F}\right)^{-1/2} & & \approx 246 \text{ GeV} \\ \text{Planck scale:} \quad & M_\mathrm{Planck} = G_\mathrm{N}^{-1/2} & & \approx 1.22 \times 10^{19} \text{ GeV}. \end{aligned}$$

These scales set the ranges of the energy to explore them. Currently, with the first data of the LHC, energies are leaving the Fermi scale and dunk into the TeV region. This work is about new insights when entering this experimental hidden new energy ranges.

All the ingredients of the Standard Model can be explained via relativistic quantum field theory [24] incorporating a number of symmetries. The most significant symmetries, gauge symmetries, are explained and applied in the following chapters.

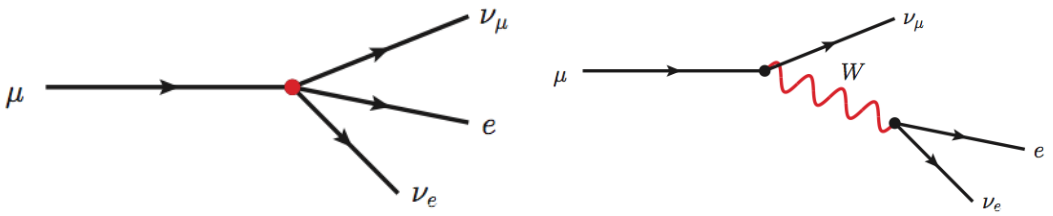
Although the Standard Model is an elegant and successful theory, there are also lots of open questions that cannot be explained in the context of the Standard Model:

- Are there really three generations of particles? Why?
- Why is the mass of the neutrinos so tiny?
- How to explain and deduce free parameters?
- How to explain the hierarchy problem [25]?
- What are the constituents of dark matter [26]?
- How can baryon asymmetry be described [27]?

Current as well as future theories and experiments try to find answers to these and related questions.

2.1.1 Electroweak Theory

Because vector boson scattering is a feature of the electroweak interaction, a closer look at the electroweak theory is taken in this work. This includes the breaking of electroweak symmetry. Both aspects of the Standard Model are explained in the following sections.



(a) Muon decay in Fermi's theory.

(b) Muon decay in the Standard Model in leading order.

Figure 2.1: Comparison of the description of the muon decay in Fermi's original theory (left) and the Standard Model Feynman diagram in leading order (right).

Fermi Theory

The first attempt to find a theory of the weak interaction was proposed by Enrico Fermi's theory of beta decay [28]. Because no W gauge boson was known to the time of this theory (1933), the fermions are described to interact directly with one and another at one vertex. The original vertex of the Fermi interaction and the Standard Model description with a Feynman diagram can be compared in Figure 2.1 for the muon decay.

Fermi's original calculation uses *Fermi's golden rule* [29] and postulates a theoretical muon lifetime τ of:

$$\tau = \frac{192\pi^3}{G_\mathrm{F}^2 m_\mu^5}. \tag{2.1}$$

Here, m_μ is the mass of the muon and G_F is Fermi's constant. The value of this constant was not known at the time of Fermi's calculation.

However, when considering the muon lifetime in leading order perturbative quantum field theory, the calculation[6] results in exactly the same equation. But here, the value of Fermi's constant is a function of the mass M_W of the mediating W boson and the weak coupling constant g_W (see explanation of Equation 2.6):

$$G_\mathrm{F} = \frac{\sqrt{2}}{8}\left(\frac{g_\mathrm{W}}{M_W}\right)^2. \tag{2.2}$$

Fermi's theory is astonishingly precise for such an early state of knowledge about the true fundamental processes.

This is the first *effective theory* of the weak interaction, introduced because the energy was too low to recognize the heavy W boson in experiments. In Chapter 2.2 an effective theory for energies above the next great leap of energy, the TeV range, is introduced.

[6]The momentum p in the W boson propagator is neglected because it can be assumed that $p^2 \ll M_W^2$. Therefore, the propagator is not effected by the phase space integration and the calculation leads to Fermi's simple result. Also the mass of the electron is set to zero.

Leptons				Y	Q	T_3
L_L^j	$\begin{pmatrix}\nu_{e,L}\\ e_L\end{pmatrix}$	$\begin{pmatrix}\nu_{\mu,L}\\ \mu_L\end{pmatrix}$	$\begin{pmatrix}\nu_{\tau,L}\\ \tau_L\end{pmatrix}$	$-1/2$ $-1/2$	0 -1	$+1/2$ $-1/2$
l_R^j	e_R	μ_R	τ_R	-1	-1	0

Quarks				Y	Q	T_3
Q_L^j	$\begin{pmatrix}u_L\\ d'_L\end{pmatrix}$	$\begin{pmatrix}c_L\\ s'_L\end{pmatrix}$	$\begin{pmatrix}t_L\\ b'_L\end{pmatrix}$	$+1/6$ $+1/6$	$+2/3$ $-1/3$	$+1/2$ $-1/2$
u_R^j	u_R	c_R	t_R	$+2/3$	$+2/3$	0
d'^j_R	d'_R	s'_R	b'_R	$-1/3$	$-1/3$	0

Table 2.3: Fermion fields of the Standard Model. The left-handed fields are expressed in doublet notation, whereas the right-handed counterparts are SU(2) singlets. The quantum numbers Y, Q and T_3 are introduced in Equation 2.6 and 2.13.

Gauge Theory of Electroweak Interactions

In the electroweak theory of Glashow, Salam and Weinberg [2, 3, 7], the main input is the Lagrangian and the chiral gauge invariance of the theory. The gauge group is chosen as $\mathrm{SU}(2) \otimes \mathrm{U}(1)$ and the constituents of the the theory are described by left- and right-handed four-component Dirac fermion fields:

$$\psi_{L/R} = \frac{1}{2}\left(1 \mp \frac{\gamma_5}{2}\right)\psi. \tag{2.3}$$

The full Lagrangian of the theory has the following form[7]:

$$\begin{aligned}\mathscr{L} = {} & \mathrm{i}\,\overline{L}_L^j \gamma^\mu D_\mu L_L^j + \mathrm{i}\,\overline{l}_R^j \gamma^\mu D_\mu l_R^j \\ & + \mathrm{i}\,\overline{Q}_L^j \gamma^\mu D_\mu Q_L^j + \mathrm{i}\,\overline{u}_R^j \gamma^\mu D_\mu u_R^j + \mathrm{i}\,\overline{d}'^j_R \gamma^\mu D_\mu d'^j_R \\ & - \frac{1}{4} W_{a,\mu\nu} W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}.\end{aligned} \tag{2.4}$$

The doublet notation is used for the left-handed fields. A complete list of all fermion fields can be seen in Table 2.3.

The bar notation represents the adjoint spinor[8]

$$\overline{\psi} \equiv \psi^\dagger \gamma^0. \tag{2.5}$$

The four Dirac matrices γ^μ [30] and the covariant derivative,

$$D_\mu = \partial_\mu + \mathrm{i}\, g_\mathrm{W} T_a W_\mu^a + \mathrm{i}\, g_Y Y B_\mu, \tag{2.6}$$

[7] A sum taken over the three generations j of leptons and quarks is suppressed to preserve readability.
[8] Not to be confused with the hermitian adjoint $\psi^\dagger$.

are summed over μ according to the Einstein summing convention. The covariant derivative is introduced to ensure local gauge invariance under $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$:

$$\psi \to \mathrm{e}^{\mathrm{i}\,(\alpha(x)Y + \beta_a(x)T_a)}\psi. \tag{2.7}$$

In this equation, ψ is either a two-component left-handed $\mathrm{SU}(2)_L$ doublet or a right-handed singlet under the same symmetry group.

This postulation therefore introduces the gauge fields W_μ^a ($a = 1, 2, 3$) and B_μ with the corresponding gauge couplings g_W and g_Y. The generators are the weak isospin $T_a \equiv \frac{1}{2}\sigma_a$ (with the Pauli matrices σ_a) and the weak hypercharge Y (see Equation 2.13 and Table 2.3).

Field strength tensors in the last part of the Lagrangian ensure kinetic terms for the gauge bosons and can be expressed in terms of the gauge fields:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.8}$$

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g_\mathrm{W}\epsilon^{abc}W_\mu^b W_\nu^c. \tag{2.9}$$

However, the gauge fields need to be identified with the physical gauge bosons. Because the photon does not couple to neutrinos, the real physical gauge fields of photon (A_μ) and Z boson (Z_μ) are linear combinations of two of the gauge fields of the covariant derivative in Equation 2.6:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_\mathrm{W} & \sin\theta_\mathrm{W} \\ -\sin\theta_\mathrm{W} & \cos\theta_\mathrm{W} \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}. \tag{2.10}$$

Mixing between these two fields is therefore a rotation by the *electroweak mixing angle* θ_W. The value of this angle slightly depends on the energy scale. For a renormalization scale at the Z boson mass and with the $\overline{\mathrm{MS}}$ renormalization scheme[3], the value of this angle is found to be

$$\theta_\mathrm{W}(M_Z) = 0.502. \tag{2.11}$$

The fields of the physical charged W bosons ($W_\mu^\pm$) are a simple superposition:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp \mathrm{i}\,W_\mu^2\right). \tag{2.12}$$

To ensure the exact coupling of the photon A^μ to the charged particles, a relation between the weak hypercharge, the electrical charge and the third component of the weak isospin can be derived:

$$Q = T_3 + Y. \tag{2.13}$$

In all three families the *weak eigenstates* d', s', b' of the down-type quarks[9] appear in the full Lagrangian of Equation 2.4. These fields are a linear combination of the *mass eigenstates* d, s and b:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.14}$$

[9] These are the quarks with the same quantum numbers as the down quark: down, strange, bottom.

The unitary 3×3 matrix in the above equation is called *CKM-Matrix*[10] V_{CKM} and contains the information of the strength of flavor-changing charged weak currents. When rewriting in canonical form [31], the matrix contains three independent Euler angles and one CP-violating phase. A fit to current measurements leads to the following absolute values [14]:

$$V_{\mathrm{CKM}} = \begin{pmatrix} 0.9743 & 0.2253 & 0.0035 \\ 0.2252 & 0.9735 & 0.0410 \\ 0.0086 & 0.0403 & 0.9992 \end{pmatrix}. \tag{2.15}$$

An equivalent mixing between mass eigenstates and weak eigenstates occurs in the neutrino sector [15].

With this complete Lagrangian of electroweak theory and with the help of principles of quantum field theory, all Feynman rules can be derived to calculate *transition matrix elements*[11]. Experimentally accessible values like cross sections and lifetimes are now reachable on the theoretical side via these matrix elements and phase space integrations [32].

This generalization of the Fermi theory shows the transition from an effective theory to a more complete description. The effective theory of electroweak interaction of Chapter 2.2 is again a attempt to reach new regions of energy by an effective ansatz.

2.1.2 Electroweak Symmetry Breaking

In the theory described in the last section all particles are assumed to be massless. But the experiment [14] clearly shows non-zero masses of all particles - only the photon remains massless[12]. Mass terms of weak gauge bosons and electrically charged fermions would break the fundamental local gauge invariance. For instance, a mass term of the Z boson,

$$\mathscr{L}_{M_Z} = \frac{1}{2} M_Z^2 Z^\mu Z_\mu, \tag{2.16}$$

is forbidden by the gauge invariance of Equation 2.7. The symmetry must therefore be broken to describe nature. Mass terms of gauge bosons can be introduced by spontaneous breaking of this symmetry[13]. A system in a symmetrical state ends up in a non-symmetrical state after such a breaking.

Higgs Mechanism

The mechanism of spontaneous symmetry breaking of electroweak theory in the Standard Model is the Higgs mechanism [33, 34]. At energies high enough that the symmetry is unbroken all elementary particles are massless. At a critical energy, the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry is broken when the Higgs field leaves the instable point of maximum energy, moves into a local stable minimum and receives a non-zero vacuum

[10] Introduced by Cabibbo, Kobayashi and Maskawa.
[11] The transition matrix elements are also called *amplitudes* or simply *matrix elements*.
[12] Albeit QCD is not considered here, the gluons as gauge bosons of QCD are also massless.
[13] The name is a little bit misleading, because the symmetry is hidden and not really broken.

expectation value v. Through this mechanism gauge bosons acquire mass. Fermions acquire their mass as a result of the interaction with the Higgs field.

The potential of the Higgs mechanism has the following form:

$$V(\Phi) = -\mu^2 |\Phi|^2 + \lambda |\Phi|^4 . \tag{2.17}$$

The complex weak isospin doublet field Φ transforms under local $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ gauge symmetry in the following way:

$$\Phi \to \mathrm{e}^{\mathrm{i}\,\alpha(x) Y_\Phi} U(x) \Phi. \tag{2.18}$$

A transformation matrix of the $\mathrm{SU}(2)_L$ gauge symmetry is $U(x)$. $Y_\Phi = \frac{1}{2}$ is the generator of the $\mathrm{U}(1)_Y$ transformation and $\alpha(x)$ is a continuous transformation parameter.

The Higgs doublet field can be written in the following general form:

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} = \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}} (v + H + \mathrm{i}\, G) \end{pmatrix} . \tag{2.19}$$

When using the electrical charge operator $Q = T_3 + Y_\Phi$ (equivalent to Equation 2.13), the electrical charge can be calculated for every component as the value indicated by the upper indices. G^+ is a complex field, H and G are scalar fields and v is the vacuum expectation value and represents one arbitrary vacuum ground state of the infinite number of ground states:

$$\Phi_0 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} v \end{pmatrix} . \tag{2.20}$$

Because a special choice of gauge fixing is possible, *unitary gauge* ($G^+ = G = 0$) can be used to identify the physics and simplify the problem:

$$\Phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} (v + H) \end{pmatrix} . \tag{2.21}$$

The Standard Model Lagrangian of Equation 2.4 can now be extended by adding the Higgs Lagrangian

$$\mathscr{L}_{\text{Higgs}} = (D^\mu \Phi)^\dagger (D_\mu \Phi) + \mu^2 |\Phi|^2 - \lambda |\Phi|^4 . \tag{2.22}$$

When written out, this equation clearly shows the vertices of the Higgs field coupling to the gauge bosons and the mass terms of gauge bosons and the Higgs:

$$m_W = v \cdot \frac{g_\mathrm{W}}{2} \tag{2.23}$$

$$m_Z = v \cdot \frac{g_\mathrm{W}}{2 \cos \theta_\mathrm{W}} \tag{2.24}$$

$$m_\gamma = 0 \tag{2.25}$$

$$m_H = v \cdot \sqrt{2\lambda} \tag{2.26}$$

$$m_f = v \cdot \frac{y_f}{\sqrt{2}}. \tag{2.27}$$

The last line is a result of adding the Yukawa terms to the Lagrangian of Equation 2.22 to allocate mass to the electrically charged fermions:

$$\mathscr{L}_{\text{Yukawa}} = y_{L^j}\overline{L}_L^j \Phi l_R^j + y_{u^j}\overline{Q}_L^j \Phi^{\text{c}} u_R^j + y_{d'^j}\overline{Q}_L^j \Phi d_R'^j + \text{h.c.} \tag{2.28}$$

Charge conjugated fields Φ^{c} are transformations of the original field with the help of the second Pauli matrix σ^2:

$$\Phi^{\text{c}} = \text{i}\,\sigma^2 \Phi. \tag{2.29}$$

How neutrino masses are represented in the Lagrangian is not yet known.

2.2 Effective Theories of Electroweak Interactions

As continuation of the effective theory of Fermi described in Chapter 2.1.1 to the next step of energy at the TeV scale, an effective theory of the electroweak interaction [23] is considered here. The formalism of effective Lagrangians provides a general framework for the description of electroweak interactions in the absence of a completely renormalizable field theory. Historically, the formalism of effective Lagrangians has been developed for the strong interaction [35, 36].

When parametrizing not yet accessed regions of energy, an infinite set of new parameters has to be introduced, while preserving all aspects of low-energy physics. These new infinite number of free parameters in the model mark the infinite set of possible new theories, but destroy the renormalizability. The goal is to do experiments revealing a finite subset of these parameters chosen by nature. This yields a finite theory.

The complete effective Lagrangian can be written as a sum of Lagrangians of operator dimension[14] two to four and a self-interaction Lagrangian of a new field $\Sigma(x)$:

$$\mathscr{L} = \mathscr{L}_\Sigma + \mathscr{L}_2 + \mathscr{L}_3 + \mathscr{L}_4. \tag{2.30}$$

The additional field Σ is an isospin 2×2 matrix valued field introduced to parametrize our ignorance about the symmetry breaking sector. This field *can* be associated to the Standard Model Higgs boson. A general local gauge invariant potential term of the field Σ can be introduced:

$$\mathscr{L}_\Sigma = -\mu^2 \left(\frac{v}{2}\right)^2 \text{tr}\left[\Sigma^\dagger\Sigma\right] + \lambda \left(\frac{v}{2}\right)^4 \text{tr}\left[\Sigma^\dagger\Sigma\right]^2. \tag{2.31}$$

This term can be seen as an generalization of the Higgs potential term in Equation 2.17. Equivalent to the Standard Model Higgs, the electroweak scale v and additional free parameters μ and λ are introduced.

The operator dimension two term is a generalization of the kinetic term of the Standard Model Higgs field and the gauge boson mass terms and includes the Σ field to preserve local gauge invariance:

$$\mathscr{L}_2 = -\left(\frac{v}{2}\right)^2 \text{tr}\left[V_\mu V^\mu\right] - \frac{1}{2}\beta' \left(\frac{v}{2}\right)^2 \text{tr}\left[TV_\mu\right] \text{tr}\left[TV^\mu\right]. \tag{2.32}$$

[14] The Operator dimension is a classical scaling dimension. For an operator A, the operator dimension is the power of energy and determined by dimensional analysis of the Lagrangian. An elementary bosonic field contributes 1, an elementary fermionic field 3/2 and the contribution of a vector field is also 1.

The trace is taken in isospin space and an additional free parameter β' is introduced. A non-zero value of β' parametrizes a breaking of custodial symmetry [37]. V_μ and T are isospin matrix fields and can be parametrized with the following formulas:

$$V_\mu = \Sigma(D_\mu\Sigma)^\dagger \tag{2.33}$$

$$T = 2\,\Sigma T_3 \Sigma^\dagger, \tag{2.34}$$

with the covariant derivative $D_\mu\Sigma = \partial_\mu\Sigma + \mathrm{i}\, g_\mathrm{W} T_a W_\mu^a \Sigma + \mathrm{i}\, g_Y Y \Sigma B_\mu$ similar to Equation 2.6.

Modified mass terms of the fermions are introduced by dimension three operators. Also all possible terms for neutrino masses that are forbidden for charged particles are included by the last two terms:

$$\begin{aligned}\mathscr{L}_3 = &-\left(\bar{Q}_L^i \Sigma M_Q^{ij} Q_R^j + \bar{L}_L^i \Sigma M_L^{ij} L_R^j + \mathrm{h.c.}\right) \\ &- \bar{L}_L^{i\,\mathrm{c}} \Sigma^* M_{N_L}^{ij} \frac{1+\tau^3}{2} \Sigma L_L^{j\,\mathrm{c}} - \bar{L}_R^{i\,\mathrm{c}} M_{N_R}^{ij} \frac{1+\tau^3}{2} L_R^j.\end{aligned} \tag{2.35}$$

Here, M_Q^{ij}, M_L^{ij}, $M_{N_L}^{ij}$, $M_{N_R}^{ij}$ are mass matrizes in the generation space. Furthermore, the isospin doublet notation is used for left- *and* right-handed fields:

$$Q_{L/R}^i = \begin{pmatrix} U_{L/R}^i \\ D_{L/R}^i \end{pmatrix} \qquad L_{L/R}^i = \begin{pmatrix} N_{L/R}^i \\ E_{L/R}^i \end{pmatrix}. \tag{2.36}$$

The fermion fields $U_{L/R}^i$, $D_{L/R}^i$, $L_{L/R}^i$, $N_{L/R}^i$ and $E_{L/R}^i$ are understood as two-component Weyl spinors (compared to the four-component Dirac fields used in Chapter 2.1.1) and carry additional generation indices ($i = 1, 2, 3$).

Operator dimension four parts are characterizing the kinetic terms of bosons and fermions and the bosonic interaction terms:

$$\begin{aligned}\mathscr{L}_4 = &\bar{Q}_L^j \mathrm{i}\,\gamma^\mu D_\mu Q_L^j + \bar{Q}_R^j \mathrm{i}\,\gamma^\mu D_\mu Q_R^j + \bar{L}_L^j \mathrm{i}\,\gamma^\mu D_\mu L_L^j + \bar{L}_R^j \mathrm{i}\,\gamma^\mu D_\mu L_R^j \\ &- \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{a,\mu\nu} W_a^{\mu\nu}.\end{aligned} \tag{2.37}$$

Field strength tensors $B_{\mu\nu}$ and $W_a^{\mu\nu}$ are defined according to Equation 2.6 and the covariant derivative D_μ is equivalent to Equation 2.8.

Up to now, Equation 2.30 is a leading order expansion in E/Λ with a energy scale Λ where perturbation theory breaks down. In the absence of Higgs-like states, a more precise calculation [38] sets this scale down to

$$\Lambda = 4\pi v \approx 3 \text{ TeV}. \tag{2.38}$$

To make the theory finit at next-to-leading order in E/Λ, additional operator dimension four terms have to be introduced as counterterms [39]. Only five terms are CP invariant and conserve custodial symmetry [37]. Two of them are observable *only* in weak-boson scattering and are thus unconstrained by experiments so far:

$$\mathscr{L}_{\alpha_4} = \alpha_4 \,\mathrm{tr}\left[V_\mu V_\nu\right] \mathrm{tr}\left[V^\mu V^\nu\right] \tag{2.39}$$

$$\mathscr{L}_{\alpha_5} = \alpha_5 \left(\mathrm{tr}\left[V_\mu V^\mu\right]\right)^2. \tag{2.40}$$

The additional parameters α_4 and α_5 introduced within these Lagrangians are referred to as *anomalous couplings*.

But what happens to the higher operator dimension terms in Equation 2.30? Every new order receives a factor of the order of $1/16\pi^2$. As long as the energy is small enough, one can truncate the perturbative series to obtain an adequate approximation of the true Lagrangian.

2.3 Resonances and Unitarity

A comprehensive list of phenomenological models for electroweak symmetry breaking includes all types of resonances that can emerge in quasi-elastic vector boson scattering. Resonances could be elementary particles such as a Higgs boson or bound states of more fundamental objects yet to be discovered.

The effective chiral Lagrangian of Section 2.2 describes no such resonances. To include resonances to the theory, new degrees of freedom in terms of additional fields are added to the effective chiral Lagrangian [39]. A list of all additional fields with their properties considered in this work is shown in Table 2.4. For each resonance r, a kinetic term including a mass m_r and a linear coupling term to longitudinally polarized vector bosons parametrized by an coupling g_r is introduced. For example, the Lagrangian of the σ resonance has the following form:

$$\mathscr{L}_\sigma = -\frac{1}{2}\sigma(m_\sigma^2 + \partial^2)\sigma + \frac{1}{2}\sigma g_\sigma v \,\mathrm{tr}[V_\mu V^\mu]. \tag{2.41}$$

Here, the additional field is called $\sigma = \sigma(x)$.

Couplings to the photon can be ignored because the electroweak symmetry breaking is a property of the longitudinal degrees of freedom [23] which are absent for the photon. For the same reason couplings to transversally polarized W and Z bosons are not considered here.

To connect the effective theory with the additional vector boson resonances to actual experiments, one has to look at scattering amplitudes. If these amplitudes are expanded in the electroweak coupling g_W, the leading term corresponds to the scattering of longitudinally polarized particles. This term asymptotically rises with energy. This is a *violation of unitarity*. On the other hand, the scattering amplitude of transversally polarized vector bosons does not rise with energy.

When assuming custodial symmetry [37] and using the Goldstone boson equivalence theorem [40], the quasi-elastic scattering amplitude for longitudinal W boson scattering can be written in terms of a master amplitude[15] $A(s,t,u)$:

$$A(W_L^+W_L^- \to W_L^+W_L^-) = A(s,t,u) + A(t,s,u). \tag{2.42}$$

A motivation of this formula and equivalent scattering amplitudes for the scattering of the other vector bosons can be found in [39].

[15] This is actually the Goldstone boson scattering amplitude.

Type	Spin J	Isospin I	Width Γ/Γ_0	Electric charges
σ	0	0	6	0
φ	0	2	1	$0, +, -, ++, --$
ρ	1	1	$4/3 \cdot v^2/m^2$	$0, +, -$
f	2	0	1/5	0
t	2	2	1/30	$0, +, -, ++, --$

Table 2.4: All resonances considered in this work with their properties: spins J, isospins I, widths Γ as multiple of a unit width $\Gamma_0 = g^2m^3/64\pi v^2$ and electric charges.

At tree level but to next-to-leading order in the energy expansion, the master amplitude for energies around the TeV scale can be approximated with:

$$A^{\text{tree}}(s,t,u) = \frac{s}{v^2} + A_{\text{r}}^{\text{tree}}(s,t,u) \tag{2.43}$$

$$A^{\text{tree}}(t,s,u) = \frac{t}{v^2} + A_{\text{r}}^{\text{tree}}(t,s,u). \tag{2.44}$$

Each possible resonance r contributes a term A_r^{tree} to the master amplitude. In the following analysis it is assumed that only one resonance is present in the vector boson scattering channel. This can easily be extended to more than one resonance. For instance, the calculation [39] for the σ resonance yields:

$$A_\sigma^{\text{tree}}(s,t,u) = -\frac{g_\sigma^2}{v^2}\frac{s^2}{s-M_\sigma^2} \tag{2.45}$$

$$A_\sigma^{\text{tree}}(t,s,u) = -\frac{g_\sigma^2}{v^2}\frac{t^2}{s-M_\sigma^2}. \tag{2.46}$$

Setting $g_\sigma = 1$, this amplitude cancels the rise of the first term of Equation 2.43. This is the case of the Standard Model Higgs boson. All other amplitudes are rising with energy, leading to unitarity violation of the quasi-elastic scattering amplitude for longitudinal vector boson scattering of Equation 2.42.

Several methods are available to unitarize this theory [41, 42]. In this thesis, the K-matrix formalism [43,44] is chosen. When applying this method, the master amplitude $A(s,t,u)$ has to be expressed in terms of isospin eigenamplitudes $A_I(s,t,u)$ [39]. These eigenamplitudes can be decomposed in spin-isospin eigenamplitudes $a_{IJ}(s)$ using Legendre polynomials.

The point of the K-matrix formalism is, that every unitary eigenamplitude $a_{IJ}(s)$ has to respect the Argand-circle condition:

$$\left| a_{IJ}(s) - \frac{\mathrm{i}}{2} \right| = \frac{1}{2}. \tag{2.47}$$

If the eigenamplitude $a_{IJ}(s)$ of an arbitrary scattering amplitude fails this requirement, the following eigenamplitude fulfilling the Argand-circle condition is chosen to

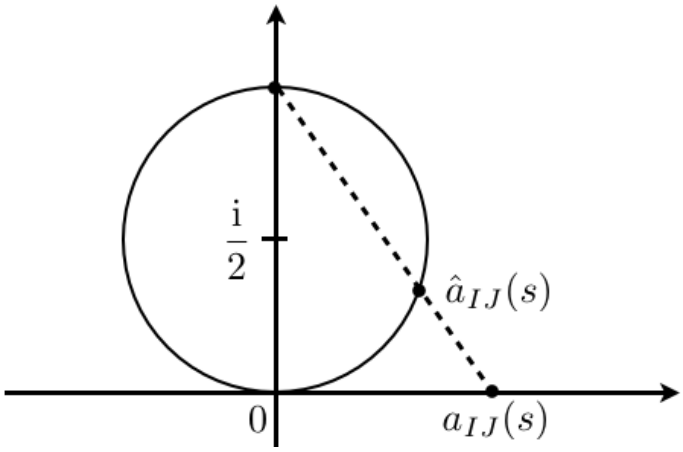


Figure 2.2: Illustration of the K-matrix unitarization formalism of a real scattering eigenamplitude $a_{IJ}(s)$. An unitarized eigenamplitude $\hat{a}_{IJ}(s) = 1/\left[\mathrm{Re}\left(1/a_{IJ}(s)\right) - \mathrm{i}\right]$ can be geometrically interpreted as a projection (dashed line) on the Argand-circle (solid lined circle).

be calculated:

$$\hat{a}_{IJ}(s) = \frac{1}{\mathrm{Re}\left(1/a_{IJ}(s)\right) - \mathrm{i}}. \tag{2.48}$$

This equation can be interpreted as a projection on the Argand-circle if the original eigenamplitude $a_{IJ}(s)$ is real (see Figure 2.2). For an arbitrary complex scattering eigenamplitude $a_{IJ}(s)$ the real eigenamplitude,

$$a_{IJ}^{\mathrm{real}}(s) \equiv \frac{1}{\mathrm{Re}\left(1/a_{IJ}(s)\right)}, \tag{2.49}$$

is projected on the Argand-circle with the same Equation 2.48.

After application of the K-matrix formalism, the original scattering amplitude $A(W_L^+W_L^- \to W_L^+W_L^-)$ is therefore replaced by an unitarized amplitude $\hat{A}(W_L^+W_L^- \to W_L^+W_L^-)$ with asymptotically saturation.

3 Experiment

3.1 Large Hadron Collider

The *Large Hadron Collider* (LHC) [45] is a proton-proton[1] collider at the European Organisation for Nuclear Research (CERN[2]) near Geneva, Switzerland. The circular ring is located 70 to 100 meters under ground and has a circumference of about 26.7 kilometers. It uses the tunnel of the previous Large Electron-Positron Collider (LEP) which was closed down in November 2000. The still existing PS/SPS accelerator complexes are used to pre-accelerate the protons before they are injected into the LHC. In total, 1 232 superconducting dipole and several thousand multipole magnets are installed.

After 7 years of construction, the LHC was completed in 2008 and first proton beams were successfully circulated in the LHC ring for the first time in September 2008. About one week later the operation was disrupted by a faulty electrical connection between two superconducting magnets. A total of 53 magnets were damaged by the pressure wave of the vaporizing helium and repaired during a nearly one year shutdown. Since the end of march 2010 the LHC is successfully operating in continuous operation.

Designed for an energy of 7 TeV per beam resulting in 14 TeV nominal center-of-mass energy, the LHC runs with lowered energy of 3.5 TeV per beam (7 TeV in the center-of-mass frame) in the first years. After a scheduled shutdown, energies of 14 TeV should be possible.

The design luminosity of the LHC is $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. To achieve this luminosity, the protons will be concentrated in up to 2 808 bunches each consisting of up to 10^{11} protons. A peak luminosity of more than $10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ has been reached up to the end of this work. More than $1\,\mathrm{fb}^{-1}$ of data is planned to be collected in the 2010/11 run and an integrated luminosity of 50 to $70\,\mathrm{fb}^{-1}$ is being aimed for until 2016. To finally reach the planned $3\,\mathrm{ab}^{-1}$ by 2030, a big upgrade is scheduled [46].

Six experiments are running at the LHC: ALICE [47], ATLAS [48], CMS [49], LHCb [50], LHCf [51] and TOTEM [52]. ALICE (A Large Ion Collider Experiment) is constructed for studying collisions of heavy ions. One aim is the exploration of the quark-gluon plasma in lead-lead nuclei collisions at 2.76 TeV nucleon energy of the lead beam. The LHCb ('b' stands for the bottom quark) is a specialized b-physics experiment aimed at measuring CP violation and the interactions of bottom quarks. The smallest experiment, LHCf (Large Hadron Collider forward) is designed for the study of the forward region of the interaction and is located close to the beam

[1] The LHC is also designed to accelerate heavy ions at a center-of-mass energy of 574 TeV.

[2] Original in French: Conseil Européen pour la Recherche Nucléaire. Official name: European Organization for Nuclear Research.

axis to measure the energy and numbers of neutral pions. TOTEM (TOTal Elastic and diffractive cross section Measurement) aims at measurement of total elastic scattering, diffractive processes and cross section. The CMS (Compact Muon Solenoid) and ATLAS (A Toroidal LHC Apparatus) experiments are multi-purpose detectors. Their main tasks are the search for the Higgs boson, the exploration of physics at the TeV scale and the validation and measurement of parameters of the Standard Model.

A seventh experiment, MoEDAL (Monopole and Exotics Detector At the LHC) [53], is approved by the CERN Research Board and will be installed during the next long shutdown in 2011. It will search for massive stable or pseudo-stable particles such as magnetic monopoles or dyons.

3.2 The ATLAS Detector

3.2.1 Detector Description

The ATLAS detector [48] (see Figure 3.1) is a multipurpose detector at point 1 of the LHC designed to cover a wide range of physics measurements [54]. With a nominally forward-backward symmetric cylindrical geometry it covers almost the whole solid angle (4π) around the interaction point.

The ATLAS coordinate system is a Cartesian right-handed coordinate system. Origin is the nominal collision point, the positive x-axis points to the centre of the LHC ring, the positive y-axis points upwards and the positive z-direction completes the right-handed coordinate system by pointing in anti-clockwise beam direction. The polar angle θ is defined with respect to the z-axis and φ, the azimuthal angle, is defined in the transverse x-y plane around the beam axis.

Transverse variables are defined in the x-y plane. The transverse momentum p_{T} is defined as

$$p_{\mathrm{T}}^2 = p_x^2 + p_y^2. \tag{3.1}$$

Instead of the polar angle θ the *pseudorapidity* is widely used in particle physics. Unlike velocities v_i, rapidities $y_i = \tanh^{-1}(v_i/c)$ can simply be added on parallel velocities in special relativity. The differential cross section is invariant under Lorentz boosts and therefore naturally expressed in terms of rapidity. Therefore, the flux of particles produced in collisions is constant over a fixed interval of rapidity. The prefix *pseudo* is used because the exact formula for calculating the rapidity,

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{3.2}$$

can be simplified to

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \tag{3.3}$$

for high energies making it easy to calculate the pseudorapidity out of the polar angle θ. E is the energy and p_z is the momentum component along the beam direction.

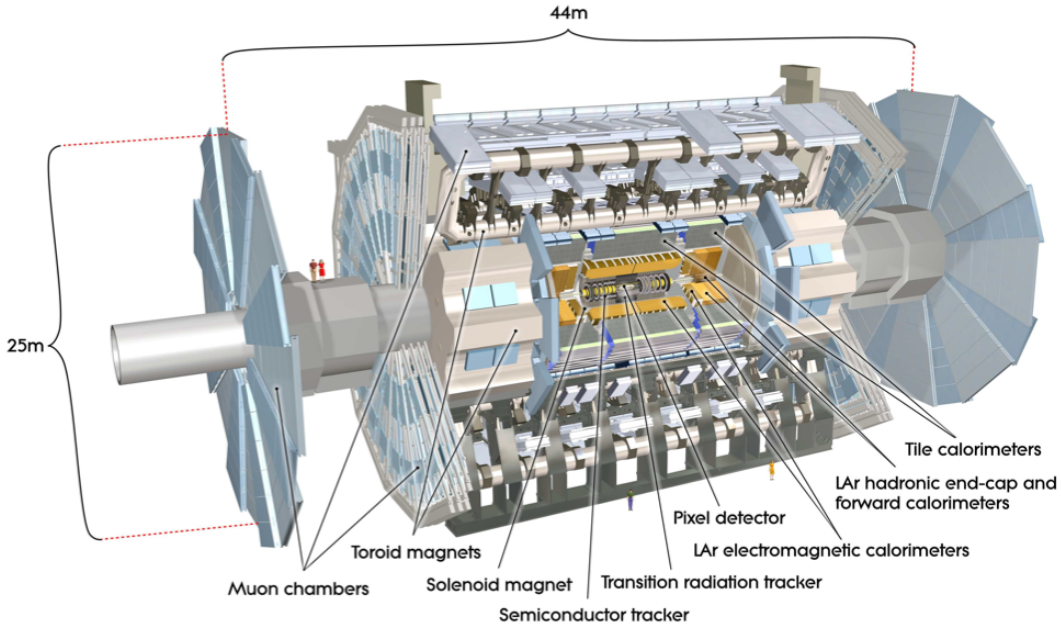


Figure 3.1: Overview of the ATLAS detector at the LHC [48].

The ATLAS detector consists of three overall detector systems: tracking detectors, calorimeters and muon chambers; together with a four-magnet system: a thin superconducting solenoid surrounding the inner cavity and three large superconducting toroids (one barrel and two endcaps) surrounding the calorimeters.

Inner Detector

The inner detector (ID) tracking system [55] is designed to reconstruct primary and secondary vertices as well as vertices from additional proton-proton interactions (pile-up) using charged tracks. It is therefore an important part of ATLAS for this analysis. Technical details about vertex finding and pile-up classification can be found in Chapter 3.3.

The ID operates inside an axial solenoid magnetic field of 2 T covering a pseudorapidity range of $|\eta| < 2.5$. Three main types of tracking devices are part of the inner detector. All detector sensors are subject to large integrated radiation doses due to their position close to the beam axis and the interaction point.

At the innermost part, the B-layer of the *silicon pixel detectors* [56], the 1 MeV neutron equivalent fluence F_{neq} rises to more than 10^{14} cm^{-2} at the design luminosity of $10^{34}\,\text{cm}^{-1}\text{s}^{-1}$ [57]. To ensure efficient semiconductor performance in this environment the sensors are operating in the temperatur range $-5\,°\text{C}$ to $-10\,°\text{C}$. The intrinsic resolution of 115 μm in the beam direction and 10 μm in the R-φ plane is reached by 1 744 pixel sensors with 47 232 pixels on each sensor. The nominal size of a pixel is $50 \times 400\,\mu\text{m}^2$ for about 90% of the pixels and $50 \times 600\,\mu\text{m}^2$ for the remaining pixels.

Located outside this innermost part of the ATLAS detector is the *Semiconductor Tracker* (SCT). It covers $|\eta| < 1.4$ in the barrel and $|\eta| < 2.5$ in the endcaps. Intrinsic accuracy of 17 μm in the R-ϕ plane and 580 μm in the beam direction is reached by this silicon strip detectors. Also the SCT has to be cooled down up to $-10\,°\text{C}$ to

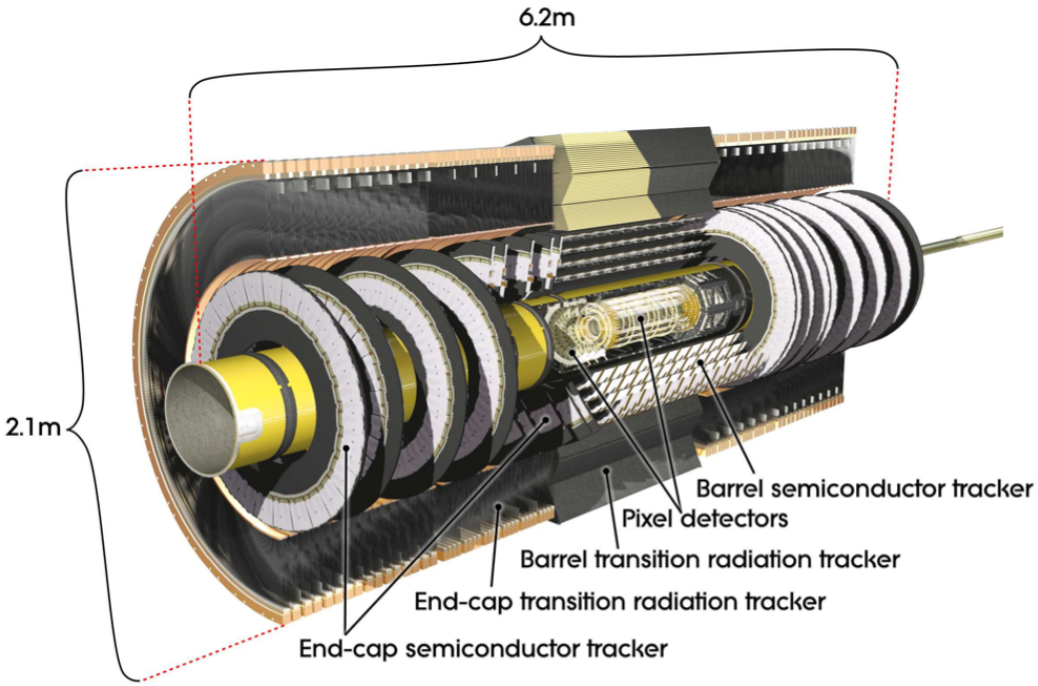


Figure 3.2: Cut-away view of the ATLAS inner detector [48].

maintain an adequate noise performance. It provides up to eight hits per track with 768 active strips of 12 cm length per sensor. The total number of sensors is 15 912.

The *Transition Radiation Tracker* (TRT) [58] is a combination of a straw tracker and a transition radiation detector. 298 000 straw tubes are used as detecting elements. They are filled with a $Xe/CO_2/O_2$ gas mixture that becomes ionized when passed by charged particles. A wire with a positive potential in the centre of the straw collects the negative ions and produces a current signal to be read out. Material between the straws causes particles to produce transition radiation and are an efficient identifier for light charged particles like electrons.

A cut-away view of the complete ATLAS inner detector can be seen in Figure 3.2.

Hybrid Calorimeter System

The calorimeter system of the ATLAS detector provides the measurement of energies of jets, electrons and photons and the identification of these particles. It is divided into several parts with different properties and tasks.

The barrel part of the detector contains the *Electromagnetic Calorimeter* (EM) and the *Hadronic Calorimeter* (HC). The EM is a sampling liquid argon[3] calorimeter with a high granularity covering at a pseudorapidity range of $|\eta| < 3.2$. It uses accordion geometry to ensure uniform response of the particles and shows excellent performance in terms of energy and position resolution. The energy resolution is found to be

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E/\ \mathrm{GeV}}} \oplus 0.7\%. \tag{3.4}$$

[3]Parts with liquid argon are often called *LAr calorimeter*.

Here and in the following, $\oplus$ is denoting addition in quadrature.

To correct for energy losses in the material of the inner detector, a pre-sampler detector is located in front of the EM calorimeter at $|\eta| < 1.8$.

Installed outside the EM and with a pseudorapidity range of $|\eta| < 1.7$ is the HC. It is a sampling calorimeter made of steel (used as absorber material) and scintillating tiles[4] (active medium). The HC is separated in a large barrel and two smaller extended barrel cylinders on each side of the central barrel. In the region $|\eta| < 3.2$ the energy resolution is found to be

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E/\text{ GeV}}} \oplus 3\%. \tag{3.5}$$

The endcap parts of the ATLAS calorimeters are also divided into hadronic (*Hadronic Endcap Calorimeter*, HEC), and electromagnetic (*Electromagnetic Endcap Calorimeter*), parts. In contrast to the HC, the HEC is liquid argon based and covers $|\eta| > 1.5$.

In the very forward region, the liquid argon *Forward Calorimeters* (FCal) provides both, electromagnetic and hadronic energy measurements for pseudorapidities up to $|\eta| = 4.9$.

Muon System

Because of their high mass in comparison to electrons, muons with a momentum below about 200 GeV do not initiate a shower in the calorimeters. Therefore, a large tracking and trigger system is installed around the calorimeters to accurately measure the properties of muons.

Four main detector types are installed: *Monitored Drift Tubes* (MDT) measure tracks for most of the barrel pseudorapidity range whereas *Cathode Strip Chambers* (CSC) deliver tracking information in the high pseudorapidity region. Trigger information is given by *Resistive Plate Chambers* (RPC) in the barrel and *Thin Gap Chambers* (TGC) in the endcaps.

The whole muon system is located inside an air-core toroid generated magnetic field of 0.5 T in the barrel and 1.0 T in endcaps.

Specialized Detectors

Special detectors are installed in the forward region around the beam axis [48].

To monitor the timing of the beam, the *BPTX* [59] system is installed 175 m away from the interaction point along the LHC beam pipe on both sides of ATLAS. These detectors are electrostatic beam pick-up detectors and installed as a part of the LHC beam instrumentation. They are used for timing purposes by ATLAS.

Forward photons and neutrons with $|\eta| > 8.3$ are detected by the *ZDCs* (Zero-Degree Calorimeters). They are located 140 m away from the interaction point where the combined beam pipe is divided back into two independent beam pipes. These detectors play a key role in heavy-ion collision by determining the centrality of such collisions, which is strongly correlated to the number of very forward (spectator) neutrons.

[4] This is the reason for the often used name *TileCal* or short *Tile*.

The *LUCID* (LUminosity measurement using a Cherenkov Integrating Detector) detectors are based on a Cherenkov luminosity counter and are primarily dedicated to online luminosity monitoring. Consisting of twenty aluminum tubes and directly surrounding the beam pipe in each endcap region of ATLAS these detectors have to be very resistent against radiation.

To trigger on minimum bias events during initial data-taking at low luminosities the *MBTS* (Minimum Bias Trigger Scintillators) are mounted in front of the electromagnetic endcaps. They consist of big wheels at $|z| = 3.56\,\mathrm{m}$ with installed scintillator counters. Due to radiation damage, the scintillators will get irreversibly damaged after the first months of higher-luminosity operation.

3.2.2 Trigger and DAQ System

The trigger and data acquisition system (TDAQ) of the ATLAS detector is a multi-level system. There are three trigger levels where the trigger decision is based on. Directly attached to the detector is the hardware based *first level trigger* (L1). It is designed to select events at a maximum rate of 75 kHz out of an input rate of 40 MHz and identify regions of interest (RoIs). The *High Level Trigger* (HLT) system is software based and runs on processor farms. It consists of the *second level trigger* (L2) and a last trigger in the chain, *Event Filter* (EF). The L2 trigger uses more detector information and better and more complex algorithms. The EF revisits the events passing the L2 trigger looking in more detail at the measurements in the RoIs. The rate is thereby reduced to a maximum of 300 Hz representing the final trigger output.

A main part of the TDAQ system is the buffering at all trigger levels. While waiting for decisions of lower levels the data is stored until the decision is made.

In the initial phase of data taking trigger decisions are also based on hits in one or both of the MBTS scintillator wheels in coincidence with the BPTX. This system is also part of the L1 trigger.

The trigger algorithms are in permanent development and improvement and will change over the data taking period. Exact triggers for 14 TeV data taking are therefore not yet defined. Trigger efficiencies will also change when measured directly from data.

3.2.3 ATLAS Event Data Model

The data collected by the ATLAS detector or generated with Monte Carlo generators (see Chapter 4.1) is the main source for the physics research done with ATLAS. This data is structured according to the ATLAS event data model. Because of the huge amount of real or simulated data, different data types exist for the various kinds of physics analyses. Their size and structure depend on the stage of reconstruction. The most important data formats for these analyses are AOD (Analysis Object Data) and ROOT files, but there is a wide range of other types:

Byte-stream data is a persistent presentation of the event data flowing from the trigger system. This data is converted to a C++ object representation, the RDO (*Raw Data Object*). This primary type of data is the base for the detector reconstruction, available with the ESD (*Event Summary Data*). The AOD is a summary

of the reconstructed events and contains sufficient information for common analyses. C++ objects are used to store the data. For reconstructed data, these objects are electrons, jets, reconstructed tracks, etc. and contain all the information and parameters required for physics analyses (e.g. pseudorapidity η, transverse momentum p_{T}). The TAG files are a short summary and allow to search for events of interest. The user can also do analyses on a format derived from the AOD, for example a ROOT file. There are many kinds of derived formats, such as standard ROOT Ntuples or D3PDs.

The ATLAS event data model also allows for back navigation. This means, if a user is doing his analysis on AOD level and might need to look at more details for a given event, it is possible for the user to navigate back from the AOD to the ESD (and, if necessary, to the RDO).

3.3 Pile-up

During the initial years of the LHC, the accelerator will operate at the luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-1}\,\mathrm{s}^{-1}$ (low-luminosity mode), reaching $10^{34}\,\mathrm{cm}^{-1}\,\mathrm{s}^{-1}$ (high-luminosity mode) at the later stages. In this environment, more than one proton-proton interaction per bunch crossing is expected. This is often referred to as *pile-up*, but is more precisely *in-time pile-up*. This means, that in general the event with interesting physics will trigger the readout but additional traces in the detector from uninteresting additional proton-proton collisions not related to the main physics event will occur. The interesting event will therefore contain a superposition of tracks coming from different events. The depositions of these tracks are a serious background.

There is another type of pile-up, *out-of-time pile-up*, which is due to the integration time of detector components. This means, that the ATLAS sub-detectors are also sensitive to tracks produced in bunch crossings before and after the main interaction. The sensitivity of the sub-detectors to such pile-up varies with the detector type and position within ATLAS. With the main bunch crossing labeled as 0, crossings after the physics event are counted with positive numbers and the one before with negativ numbers. Particularly the liquid argon calorimeter is sensitive to 38 bunch crossings for 25 ns bunch spacing.

Pile-up therefore correctly refers to both the in-time minimum bias and cavern background interactions as well out-of-time effects.

The number of proton-proton interactions per bunch crossing is Poisson distributed with averages of 23 (high-luminosity case) and 4.6 (low-luminosity case) respectively in the case of 25 ns bunch spacing. A mean number of additional pile-up collisions can be calculated from the non-diffractive cross section σ[mb], the instantaneous luminosity L[$\mathrm{mb}^{-1}\mathrm{s}^{-1}$], the number of bunches in the LHC ring N_b and the revolution frequency of protons in the LHC[5]:

$$\overline{N} = \frac{L \cdot \sigma}{f \cdot N_b}. \tag{3.6}$$

[5]If one expects protons circulating at the speed of light c and using the circumference of the LHC ring u, the revolution frequency f can be estimated to $f = c/u = 11\,245\,\mathrm{s}^{-1}$.

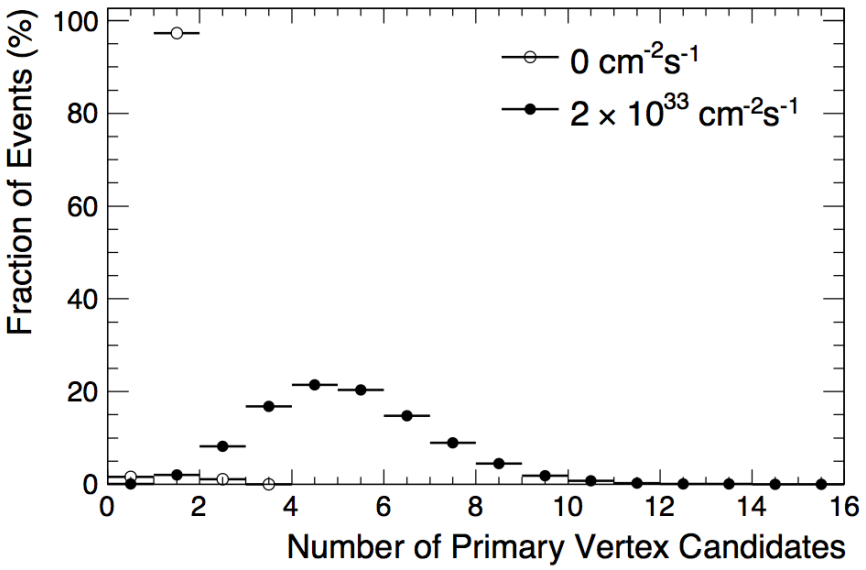


Figure 3.3: Expected number of reconstructed primary vertex candidates in $H \to \tau^+\tau^- \to lh$ [60]. $0\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is defined here as no pile-up. The pile-up case is shown for a luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, a bunch spacing of 25 ns and a corresponding average number of interactions per bunch crossing of 4.6.

In a Monte Carlo study [60] of $H \to \tau^+\tau^- \to lh$ events, a closer look at the distribution of the number of expected reconstructed primary vertex candidates was taken. The result in Figure 3.3 shows the distribution without additional pile-up events and low-luminosity pile-up (luminosity: $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, bunch spacing: 25 ns, average number of interactions per bunch crossing: 4.6). Already at this low-luminosity stage of the detector, events with more than five additional interactions are expected to take place in the detector.

Other types of background leading to additional particles to be reconstructed must be taken into account:

One important additional background is the *cavern background*: Neutrons appear in the ATLAS cavern for some seconds until they are thermalized. This leads to a permanent background of neutrons and photons. A constant rate of spallation protons and Compton electrons due to this neutrons and photons in the cavern are observed. This parasitic particles are referred to as cavern background. The result is a quality fall-off of the muon system resulting in lower resolution, decreased efficiency and a higher fake rate.

Because the vacuum in the LHC beam pipe is not ideal, elastic and inelastic scattering of the beam on the residual gas is expected. These *beam gas* background leads to spurious secondary particle fluxes in the ATLAS detector [61]. The main expected gases present in the beam pipe are H_2, CH_4, CO and CO_2 with residual densities of around $10^{13}\,\mathrm{molecules/m}^3$ [62].

Because the beam itself is spread out (a Gaussian distribution of the beam shape is a good approximation), particles of the beam can interact with the beam pipe. This leads to machine-induced secondary particles. These *beam halo* particles can cross the detector and leave signals predominantly in the endcap regions [61].

3.3.1 Handling Pile-up

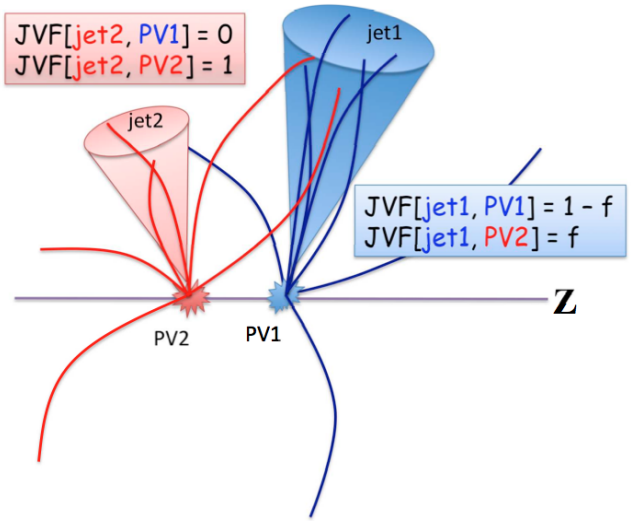


Figure 3.4: Schematic explanation of the jet-vertex fraction $JVF(i,j)$ of a single jet i matched to a considered vertex j [63]. The JVF represents the fraction of each jet's constituent transverse track momentum contributed by each vertex. All tracks of the left (red) jet are originating in the left vertex (JVF[jet2,PV2] = 1) and no track originates in the right vertex (JVF[jet2,PV1] = 0). A fraction f of the tracks transverse momentum of the right (blue) jet are originating in the left vertex (JVF[jet1,PV2] = f) and a fraction $1-f$ in the right vertex (JVF[jet1,PV1] $= 1-f$). If the primary interaction vertex is the right vertex (PV1), a cut on the jet-vertex fraction related to this vertex of $1 \geq JVF$[jet,PV1] $>$ cut > 0 would exclude the left (red) pile-up jet.

There are several methods to suppress the influence of pile-up. Nevertheless, no pile-up suppressing has been applied in the analysis of this work due to very low effects of pile-up in the final results (see Chapter 6.4).

The vertices of the primary interactions of the event and additional in-time pile-up interactions are reconstructed by the ATLAS software using an adaptive multi-vertex fitter [64]. The multiplicity of the vertex candidates is shown in Figure 3.3 for the low-luminosity mode ($2 \times 10^{33}\,\mathrm{cm^{-1}s^{-1}}$) and a bunch spacing of 25 ns. From the list of primary vertex candidates, the primary interaction vertex is the one of the hard process. This vertex is identified by the fitter as the vertex with the highest $\sqrt{N}\sum p_{\mathrm{T}}^2$. In this formula, the sum is taken over all N tracks with transverse momentum p_{T} assigned to the considered vertex. If the signature of the event is special (if for example a lepton has been detected), the primary interaction vertex can be reconstructed using these additional information.

On possibility to suppress the influence of the extra jets traversing from these additional vertices is a *jet-vertex association*. The jets are associated to the vertex they are originating (with the help of the tracks associated to the jet) and suppressed if this vertex was not selected as the primary interaction vertex. A *jet-vertex fraction* JVF can be calculated to parametrize this belonging [60, 63]. This discriminating variable represents the fraction of each jet's constituent transverse track momentum

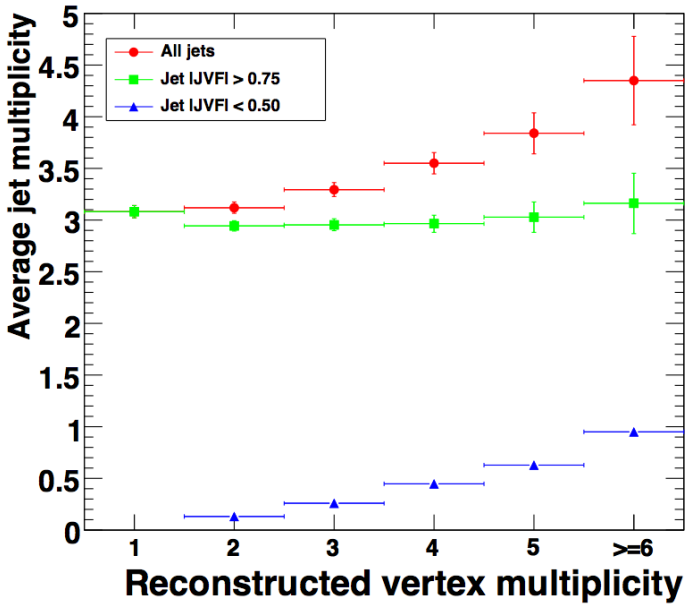


Figure 3.5: Jet multiplicity as a function of reconstructed vertex multiplicity with and without application of a jet-vertex fraction [63]. Results for a top pair sample are shown with pile-up for a luminosity of $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and a bunch spacing of 25 ns. "All jets" refers to all jets passing the selection criteria as described in [63]. This means no jet-vertex association is applied. After a cut of $JVF > 0.75$ associated to the selected primary interaction vertex, a flat distribution is observed. This means that 75% of each jet's track transverse momentum have to originate in the selected primary interaction vertex. Pile-up jets are selected by a cut of $JVF < 0.5$.

contributed by each vertex. It can be calculated with the following formula for a single jet i and a considered vertex j:

$$JVF(\mathrm{jet}_i, \mathrm{vtx}_j) = \frac{\sum_k p_\mathrm{T}(\mathrm{trk}_k^{\mathrm{jet}_i}, \mathrm{vtx}_j)}{\sum_{l,m} p_\mathrm{T}(\mathrm{trk}_l^{\mathrm{jet}_i}, \mathrm{vtx}_m)}. \tag{3.7}$$

Jet i has a fraction $JVF(i,j)$ of it's total track momentum originating in vertex j. The association is shown schematically in Figure 3.4.

The performance of jet-vertex association can be probed by the average jet multiplicity distribution as a function of the reconstructed vertex multiplicity. A flat distribution is a sign of well suppressed pile-up because no additional jets from additional vertices not associated to the primary interaction vertex are selected.

Results of an analysis [63] of the top pair process with pile-up for a luminosity of $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and a bunch spacing of 25 ns can be seen in Figure 3.5. Without a cut on JVF (red circles), an increasing number of the average jet multiplicity due to additional pile-up jets is observed. When using the jet-vertex association with a suitable cut ($JVF > 0.75$ associated to the selected primary interaction vertex, green squares), a flat distribution is recovered. For comparison, pile-up jets are selected by a cut on the jet-vertex fraction of $JVF < 0.5$ associated to the selected primary interaction vertex (blue triangles). This result clearly highlights the importance of a jet-vertex selection algorithm.

3.3.2 Pile-up Simulation

from SimuJobTransforms.DigitConfig import digitConfig	
digitConfig.pileupCollisions	= 6.9
digitConfig.bunchSpacing	= 75
digitConfig.cavernEvents	= 5
digitConfig.pileupInitialBunch	= -32
digitConfig.pileupFinalBunch	= 32
digitConfig.numberOfBeamHalo	= 0.05
digitConfig.numberOfBeamGas	= 0.0003

Table 3.1: The pile-up options file *Lumi010DigitConfig_75ns.py* of the samples in this work. It assumes an luminosity of $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and a bunch spacing of 75 ns.

To study the influence of pile-up, simulated events including these extra tracks to be reconstructed have to be produced. The general Monte Carlo production chain is explained in Chapter 4.1. Including pile-up is performed at the digitization step (see Figure 4.1). One possibility is to use ATLAS digitization *job transformations*. This is the same way as it is done in the central production system and ensures to use the latest options and bugfixes. Job transformations take input parameters and input files and convert this information into one or more output files.

A *minimum bias hits file* and a *cavern hits file* have to be added to the input to overlay minimum bias and cavern background from which the pile-up is formed. When taking into account pile-up of beam halo or beam gas events (see Chapter 3.3), an input file containing beam halo and beam gas Monte Carlo is necessary.

Additionally, a file containing the pile-up options has to be specified. The content for a luminosity of $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and a bunch spacing of 75 ns used within this work (see Chapter 4.3) can be seen in Table 3.1. The mean number of the Poisson-distributed number of pile-up collisions can be determined directly from luminosity [$\mathrm{cm}^{-2}\mathrm{s}^{-1}$] and bunch spacing [ns] with the following formula:

$$\text{mean number of pile-up collisions} = 2.3 \times \frac{\text{luminosity}}{10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}} \times \frac{\text{bunch spacing}}{25\,\mathrm{ns}}. \qquad (3.8)$$

This formula scales the number of pile-up interactions expected at a luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and a bunch spacing of 25 ns to some other luminosity given a certain bunch spacing. For the scenario of this work, the mean number of pile-up collisions is calculated to 6.9 (see Table 3.1).

On average five additional pile-up events per bunch crossing due to cavern interactions are chosen for the samples of this work. The earliest and latest filled bunch crossing contributing to a hypothetical pile-up event at zero due to out-of-time pile-up are set to ± 32. These values are determined by the sub-detector element with the slowest integration time (MDT in the muon chambers $(-32 \ldots + 32)$ and the LAr calorimeter $(-32 \ldots + 6)$. Arguments for beam halo and beam gas pile-up contributions are set in the last lines of Table 3.1. They specify the mean number of the Poisson-fluctuating beam halo and beam gas events per bunch crossing.

3.4 Experimental Signature of Vector Boson Scattering

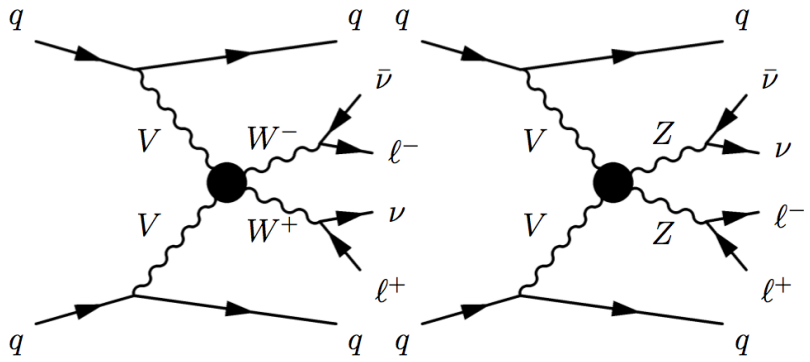


Figure 3.6: Feynman diagrams of vector boson scattering $WW/ZZ \rightarrow l\nu l\nu$. Proton remnants are not shown. Same-sign WW is also possible, but not shown. The diagrams are taken from [1].

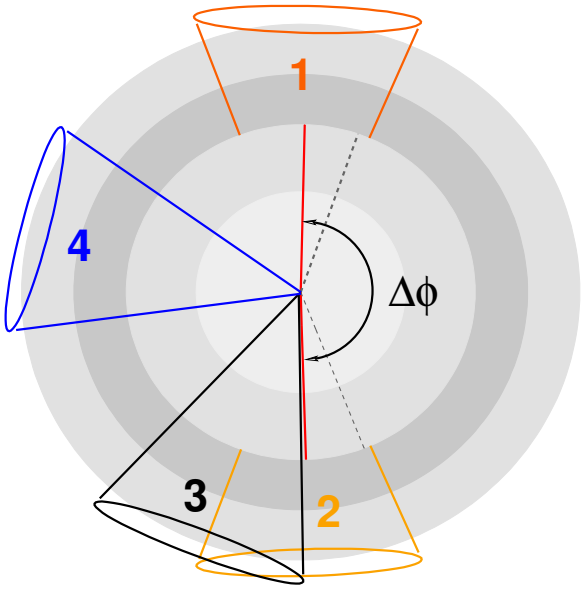
The term *Vector Boson Scattering* (VBS) refers to the interaction of two vector bosons resulting in two vector bosons. There are various channels to be studied due to the different three and four boson vertices.

In hadron colliders, two distinctive jets from the final state quarks that radiated the vector bosons at high mass will occur ("3" and "4" in Figure 3.6). These *tagging jets* [65] are expected to be well separated in pseudorapidity and carry large transverse momentum. They can be defined as the jets with the highest transverse momentum in every pseudorapidity hemisphere ($\pm\eta$) of the detector and are a strongly discriminating variable in VBS.

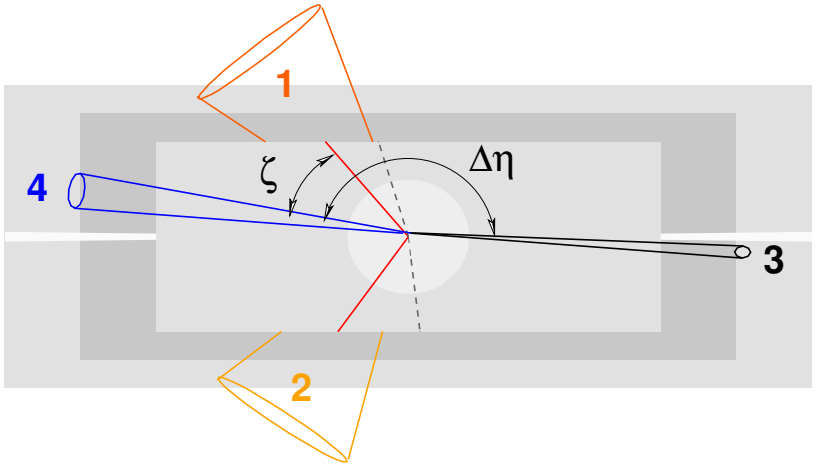
Additional jets in the pseudorapidity range between the tagging jets are rare due to lack of color flow between the initial quarks in this channel. These additional jets are called *minijets*. A veto on *minijets* can be used to separate the signal from the background but is also expected to be affected by additional jets due to pile-up.

Because each of the final bosons can decay either leptonically or hadronically, different numbers of leptons can occur in the final state. The channel without leptons is challenging in hadron colliders like the LHC due to the large background from QCD multijet final states. The di-lepton channel chosen in this work has a clean signature but a lower branching ratio compared to the single-leptonic case. However, this channel is selected because of the excellent lepton reconstruction of the ATLAS detector and a measurable angular correlation of the leptons. The two leptons ("1" and "2" in Figure 3.7) are expected to be located inside (in terms of pseudorapidity) the tagging jets. This can be parametrized with the lepton centrality ζ (see Chapter 4.4). For massive resonances, the outgoing vector bosons are highly boosted. The angular distributions of their decay products have a similar angular correlation. This leads to a separation of the leptons in φ with a peak at π.

Because of the two neutrinos from the decay of the final state bosons, missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is also an important observable in this analysis.



(a) R-φ plane



(b) R-z plane

Figure 3.7: Schematic view of a vector boson scattering event at the ATLAS detector. "1" and "2" are charged leptons from vector boson decays. "3" and "4" are tagging jets (see text). The angular correlations ζ (lepton centrality), $\Delta\varphi^{\text{leplep}}$ and $\Delta\eta^{\text{tagjets}}$ are explained in Chapter 4.4. The graphs are taken from [1].

4 Monte Carlo Samples and Discriminating Variables

4.1 ATLAS Monte Carlo Production Chain

Before real LHC data was available, all analyses were done with simulated events rather than real experimental output. These simulated events are called *Monte Carlo events*[1]. To estimate background, to compare the detector performance with simulated predictions, to estimate analyses results at higher assumed integrated luminosities and to predict detector responses for physics processes, Monte Carlo analyses are an indispensable tool also while data taking. Because data taking only started at the beginning of this analysis and because the currently collected data is far from what is needed, the complete work was done with simulated Monte Carlo events.

To produce Monte Carlo samples, multiple steps are needed. Adapted to this analysis, these steps from generation to real physics analysis are shown in Figure 4.1 and explained below for the signal processes studied in this work (see Chapter 4.3).

In the *generation* step the particle four vectors are produced from specified physics processes. The signal samples of this work are generated by the event generator Whizard [67] as described in Section 4.2. For a hadron collider like the LHC, the generation is done at parton level using PDFs (Parton Distribution Functions) [68]. Additional soft radiation is simulated with a *parton shower*. To finally arrive at unconfined particles, *hadronization* is applied to the partons in the last generation step. For the signal samples, Pythia [69] is used for parton shower and hadronization.

The next step, *simulation*, produces a record of where each particle traversed the detector and how much energy is deposited. Because this process of complete detector simulation takes a huge amount of computational power, two simulations are possible: An expensive and precise full simulation with GEANT [70, 71] and a fast simulation with Atlfast-II [72].

The output from the full simulation is subjected to *digitization* where the response of the detector is simulated. How pile-up is added in this step is described in Chapter 3.3.2.

Monte Carlo events after the digitization step are now a complete simulation of real data detector output. Therefore, simulated events and real data events from a physical process are both reconstructed in the same way. *Reconstruction* is the process where digits are merged to tracks and physics object candidates corresponding to jets and electrons for example. Finally, these are the actual objects for an analysis.

[1] The name refers to the applied *Monte Carlo method* [66] where random methods of computation and stochastic simulation are applied.

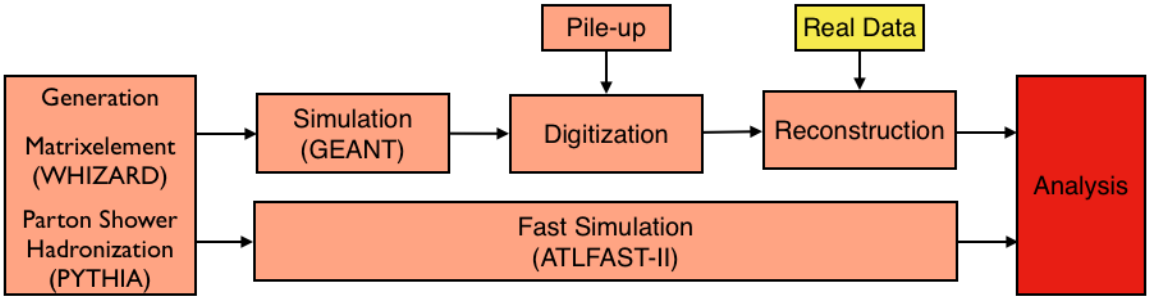


Figure 4.1: ATLAS full Monte Carlo production chain for the signal samples of this analysis. Although a variety of generators, simulation software and reconstruction algorithms are available, the general intermediate steps are found in every ATLAS Monte Carlo production.

4.2 Event Generator Whizard

To generate the signal events of this work, the event generator WHIZARD is used. It allows the calculation of tree-level matrix elements in leading order perturbation theory with the physics of the Standard Model (SM), Minimal Supersymmetric Standard Model (MSSM) and also anomalous couplings for partonic processes via optimized helicity amplitudes. Currently, WHIZARD is the only generator available that is able to simulate additional resonances with the K-matrix model (see Chapter 2.3).

Its optimized matrix elements combined with a very efficient phase space setup allow the simulation of the full six-fermion final state considered in this work. Based on this, cross sections, scattering processes and decay channels of particles can be computed. Other traditional Monte Carlo tools such as PYTHIA or HERWIG [73] contain hard-coded libraries of leading order on-shell matrix elements. Instead, WHIZARD computes the matrix element for every process that is possible in the selected physics model using the program O'MEGA [74].

When using WHIZARD, one of the physics models has to be selected and a list of processes has to be specified. The user can also add hadronization, parton showering or decay of particles in the initial state. The WHIZARD infrastructure also allows to handle structured beams. After this initial options, it is possible to set variables - beam polarization, center-of-mass energy and WHIZARD process run options. If the event generation is activated, WHIZARD allows for creation of weighted and unweighted events. Weighted events have the advantage of probing regions of phase space where most of the real event will lie.

The Vector Boson Scattering process with two leptonically decaying vector bosons analyzed in this work, $qq \to qql\nu l\nu$, is a "two-to-six" process. WHIZARD by default generates all tree-level matrix elements for a given initial and final state. Therefore, all sub-processes are included in the final output. Nevertheless, processes containing QCD vertices can be switched off by setting the strong coupling constant α_s to zero.

4.3 Monte Carlo Samples

The signal and main contributing background samples used in this analysis are in parts central and private production. Centrally produced samples are provided by the ATLAS production team and are approved within the ATLAS collaboration. Private samples are generated due to special properties not available for the central samples. Equivalent tools as in the central production are used. Specific properties of the samples can be found in Table 4.1 and 4.3.

4.3.1 Signal Samples

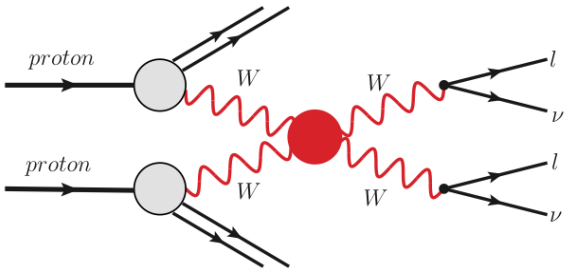


Figure 4.2: A schematic Feynman diagram of a signal process. A possible resonance is marked with a circle at the vector boson four-vertex.

This is a work on the discovery setting potential of the ATLAS detector to additional resonances in vector bosson scattering as theoretically introduced in Chapter 2.3. These resonances are therefore defined as signal. The signal samples are generated with the event generator WHIZARD for a center-of-mass energy of 14 TeV. A schematic Feynman diagram can be seen in Figure 4.2 where a possible resonance is marked with a circle at the vector boson four-vertex.

Five possible resonances at LHC energies (see Table 2.4) are taken into account: scalar isoscalar σ, scalar isotensor φ, vector isovector ρ, tensor isoscalar f and tensor isotensor t. For each resonance there are two independent and free parameters: mass m and coupling g. The mass is set to $m = 850$ GeV within this analysis. Results for a higher mass of $m = 1150$ GeV are shown in [1]. The couplings, cross sections and number of events generated are listed in Table 4.1.

To use computer time economically, only one coupling per resonance is generated. Other couplings are derived by reweighting [1]. The main reason for this decision is the slow and expensive detector simulation.

Reweighting is done downwards. This means, the highest coupling per resonance is selected and used to generate signal samples. Smaller couplings are derived from these values by assigning event weights. The highest couplings as written in Table 4.1 have been chosen to reach equivalent estimated discovery potential for all resonances.

At the detector simulation step, 6 000 of 30 000 generated events per resonance[2] are fully reconstructed with GEANT, the remaining events are reconstructed with fast detector simulation ATLFAST-II. For the pile-up samples, all 30 000 events are fully

[2]There are only 6 000 events available for the σ resonance.

Resonance	Coupling	Cross section [fb]	Events	Pile-up	Simulation
σ	2.0	78.0	6 000	no	full
			6 000	yes	full
φ	2.0	78.4	6 000	no	full
			24 000	no	fast
			30 000	yes	full
ρ	4.0	74.0	6 000	no	full
			24 000	no	fast
			30 000	yes	full
f	3.0	88.3	6 000	no	full
			24 000	no	fast
			30 000	yes	full
t	4.0	89.7	6 000	no	full
			24 000	no	fast
			30 000	yes	full

Table 4.1: Monte Carlo signal samples used in this analysis. Only masses of $m = 850$ GeV are taken into account. The coupling is the coupling generated by WHIZARD, but is reweighted to smaller values at analysis level. "full" in the simulation column refers to full simulation with GEANT and "fast" to fast simulation with ATLFAST-II. All samples are produced for a center-of-mass energy of 14 TeV.

reconstructed with GEANT. To avoid comparing full simulation with fast simulation and to have the same amount of events for every case, only 6 000 full simulated pile-up events are compared to the 6 000 full simulated events without pile-up in the following analysis. Because the pile-up and no pile-up samples match at matrix element level, the matching events with identical matrix element are identified and used for analysis. The difference between the cases with and without pile-up is therefore to be not influenced by generator effects. The remaining events are used for crosschecks in Chapter 6.6.

Pile-up is digitized for a luminosity of $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and 75 ns bunch spacing (see Chapter 3.3.2). This leads to a mean number of pile-up collisions per bunch crossing of 6.9. All samples are produced within the ATLAS framework ATHENA[3] for a center-of-mass energy of 14 TeV.

As described in Chapter 4.2, WHIZARD generates all sub-processes for a given initial and final state. Therefore, many other processes not considered as signal are included in the signal samples. Because the strong coupling constant α_s can be set to zero, all QCD processes may be switched off. Signal can be switched off when setting the coupling g to zero. Nevertheless, electroweak background processes (EW) like the Standard Model vector boson quartic vertex and other non-QCD processes resulting in two vector bosons are still part of the signal samples (see Section 4.3.2). All combinations that are possible to be generated by WHIZARD are shown in Table 4.2.

Signal is identified with additional resonances in the vector boson scattering channel

[3]Release 14.2.25

WHIZARD sample		$\alpha_s = 0$	$g = 0$
EW	B^{EW}	$\surd$	$\surd$
EW + Resonance r	$S^{\mathrm{EW}+r}(m, g)$	$\surd$	$\times$
EW + QCD	$B^{\mathrm{EW+QCD}}$	$\times$	$\surd$

Table 4.2: Samples that are possible to be generated by WHIZARD and the options to execute the generation. For the electroweak (EW) sample, the strong coupling constant α_s is set to zero and no resonance is generated by setting the coupling g of the resonance to vector bosons to zero. If the coupling g is switched on within WHIZARD, the EW and signal process are generated together. EW+QCD is generated by switching off additional resonances with the option $g = 0$ and restoring the strong coupling constant to its original value. The notation for the samples introduced here is used in the text for the expected number of signal (S) and background (B) events for the given process.

and depends on mass m and coupling g. If the coupling g is set to zero, the resonance does not couple to the vector bosons and the expected number of signal events (S) for a given resonance r is zero:

$$S := S^r(m, g) \text{ with } S^r(m, g = 0) = 0. \tag{4.1}$$

As described above and can be seen in Table 4.2, signal samples are always generated by WHIZARD including EW background. To extract the pure expected number of signal events, the EW background has to be subtracted:

$$S \equiv S^{\mathrm{EW}+r}(m, g) - B^{\mathrm{EW}}. \tag{4.2}$$

An estimate of the EW background is the WHIZARD generated signal sample (including EW as explained in Table 4.2) if the coupling g of the resonances is set to zero. To avoid statistical fluctuations at low couplings, it is therefore allowed to subtract these expected number of events, $S^{\mathrm{EW}+r}(m, g = 0)$, *instead* of the expected number of EW events B^{EW} as in Equation 4.2:

$$S \equiv S^{\mathrm{EW}+r}(m, g) - S^{\mathrm{EW}+r}(m, g = 0). \tag{4.3}$$

When doing counting experiments, this signal definition is used throughout this work. For example, "φ resonance with coupling g" means: "WHIZARD generated φ resonance with coupling g (including EW background) minus the same with coupling $g = 0$".

4.3.2 Background Processes

Because the signal is defined as new physics in form of additional resonances in the vector boson scattering channel, all other processes allowed in the Standard Model are defined as background.

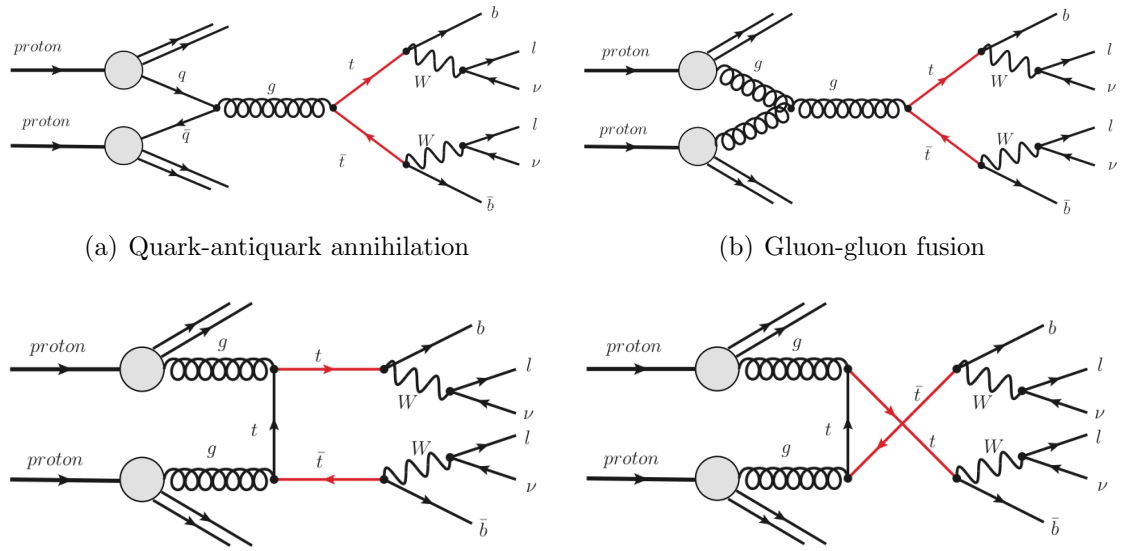
Two leptons, two neutrinos and quarks are expected in the signal final state. Charged leptons can be measured well at hadron colliders, but neutrinos will not

be directly visible in the detector. Their presence can be deduced from missing energy. Backgrounds must have these visible final states or fake the products due to detector effects like pile-up in order to mimic the signal process.

The main contributing background samples are listed in Table 4.3. They are in parts private production [1], but most of them are produced centrally. Cross sections are computed at next-to-leading order.

Only at higher orders of perturbation theory the individual backgrounds will interfere with each other or the signal process. Therefore they are treated as independent at tree level and can be added up.

Top pair Background



(a) Quark-antiquark annihilation

(b) Gluon-gluon fusion

(c) Gluon-gluon fusion

(d) Gluon-gluon fusion

Figure 4.3: Tree-level Feynman diagrams for the top quark pair production background. At LHC energies, the dominant production mode is via gluon-gluon fusion ($\approx 85\%$).

Top quarks decay almost exclusively into a W boson and a bottom quark[4]. Leptonically decaying W bosons mimic the W bosons from the signal process. The bottom quarks will form jets equivalent to the signal jets. Because of their long lifetime, b-tagging [75] can be applied to identify the bottom quarks and suppress this background. Feynman diagrams of the leading order top pair production with leptonically decaying W bosons,

$$t\bar{t} \to b\bar{b}\, W^+W^- \to b\bar{b}\, l^+\nu_l\, l'^-\overline{\nu}_{l'}, \tag{4.4}$$

are shown in Figure 4.3.

The calculated top pair production cross section for 14 TeV is [76]:

$$\sigma_{t\bar{t}} = 833^{+52}_{-39}\,\text{pb}. \tag{4.5}$$

[4]The Standard Model prediction is: $t \to Wb$: 99.82%, $t \to Ws$: 0.17%, $t \to Wd$: 0.01%.

The calculation is done at next-to-leading order including next-to-leading logarithms. Verifying this value by measuring the top pair cross section with early data [77, 78] was not completed by the time of this analysis[5]. The values of the cross section in Table 4.3 are after the filters described below.

Three different samples of this background are part of this analysis. Central MC@NLO [80] produced samples with 997 340 (877 364) events were available without (with) pile-up and are used to study the influence of pile-up. They are called $t\bar{t}$ in this analysis.

However, to reach statistical goals in the training steps of the multivariate analysis (see Section 5.3), a locally produced sample with 32.5 million events is used. The sample was centrally generated with MC@NLO by the ATLAS top group and privately simulated, digitized and reconstructed with ATLFAST-II [1]. In order to make the amount of simulation manageable, special filters and cuts are applied:

- Overlap removal between jets and leptons/photons: A jet is removed if a lepton or photon with $p_{\mathrm{T}} > 15$ GeV and $|\eta| < 2.5$ is found within $\Delta R < 0.05$ ($\Delta R \equiv \sqrt{\Delta\varphi^2 + \Delta\eta^2}$) of the truth jet [81]
- Two-lepton filter: Only events with two leptons pass
- Cut on the transverse momentum of the leptons: $p_{\mathrm{T}} \geq 10$ GeV
- Cut on jet observables: Two jets have to be found after overlap removal with $p_{\mathrm{T}} > 15$ GeV, $|\eta| < 5.0$, invariant mass $m^{jj} \geq 300$ GeV and $\Delta\eta^{jj} > 2.0$.

After application of these filters, a total of 1 832 994 events is left for analysis. No pile-up is available for this sample. In this work, this sample is called $T\bar{T}$.

MC@NLO is a parton shower Monte Carlo tool including the full next-to-leading order QCD corrections. To include full QCD corrections (real and virtual) at next-to-leading order, negative event weights are generated. These weights have to be taken into account as negative events when counting the number of events. After this recalculation, the centrally produced $t\bar{t}$ sample has a total number of events of 729 534 (641 504) for the no pile-up (pile-up) case, while the total number of events for the privately produced $T\bar{T}$ sample (without pile-up) is 1 373 788.

Single Top Background

Like the signal process, the Wt process of single top production also has two isolated leptons and two neutrinos in the final state (see Figure 4.4). Also an s-channel and t-channel process are part of the single top background, but only Wt is expected to have two isolated leptons in the final state. One b quark occurs but one jet is missing with respect to the number of jets in the signal process. The additional jet may be present due to higher order effects. The Wt sample was generated with ACERMC [82] and parton shower and hadronization are done with PYTHIA. Cross section calculations at next-to-leading order give:

$$\sigma_{Wt} = 66^{+2}_{-2}\,\mathrm{pb}. \tag{4.6}$$

[5] Data-driven results for a center-of-mass energy of 7 TeV are available within [79].

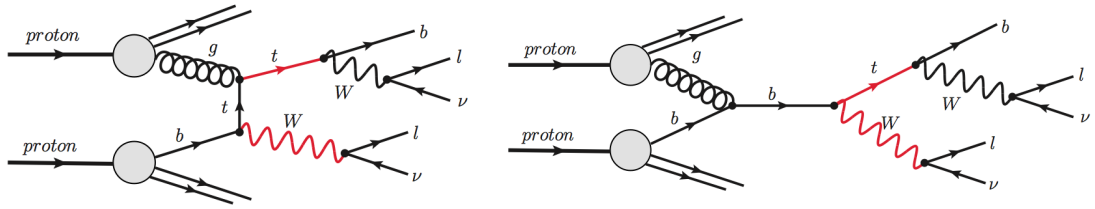


Figure 4.4: Tree-level Feynman diagram for the Wt background contribution to single top production.

After the same cuts and filters as in the privately produced $T\bar{T}$ sample the number of events after the filter is 59 482. No pile-up samples are available for this process.

W+jets and Z+jets Background

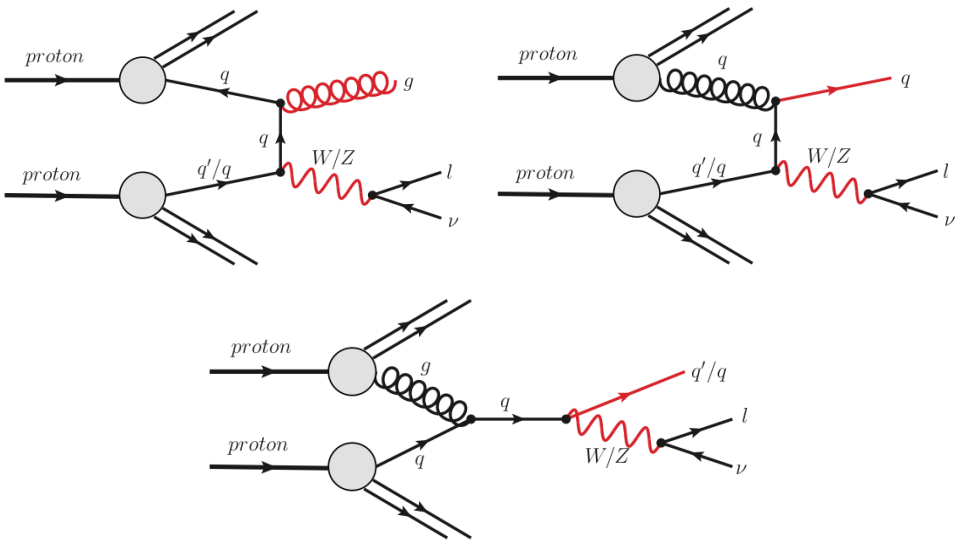
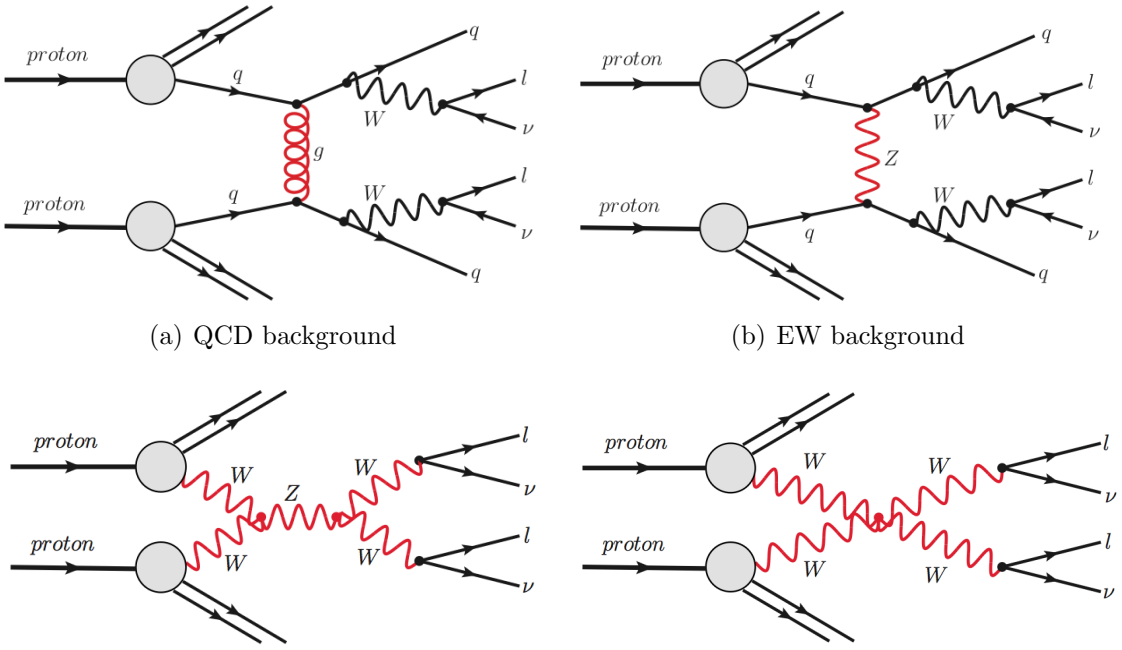


Figure 4.5: Tree-level example Feynman diagrams of the W+jets and Z+jets production backgrounds.

Another background is the production of W and Z bosons with jets. While W+jets events only contain one real lepton, Z+jets events are more similar to the signal due to two isolated leptons from the leptonically decaying Z boson. The samples are centrally produced with the leading-order generator ALPGEN [83]. HERWIG [73] was used for parton shower and hadronization and TAUOLA [84] and PHOTOS [85] for tau decays and final state radiation of the weak bosons. Two leptons with transverse momentum $p_{\mathrm{T}} > 10$ GeV are required and a cut on the invariant mass of the leptons, $m^{\mathrm{leplep}} < 200$ GeV, is applied.

All the samples are split into jet multiplicities (up to five jets) and lepton flavors. Leading-order cross sections used for this analysis are shown in Table 4.3 and example tree-level Feynman diagrams of W+jets and Z+jets can be seen in Figure 4.5.

Irreducible Backgrounds



(a) QCD background

(b) EW background

(c) EW background with triple-boson vertex

(d) EW background with quartic-boson vertex

Figure 4.6: Typical Feynman diagrams of the irreducible QCD and EW backgrounds. Due to WHIZARD properties the QCD background is always entangled with the EW background (EW+QCD).

As described in Chapter 4.3.1 and can be seen in Table 4.2, electroweak background (EW) is included within all WHIZARD samples. When switching off additional resonances and setting α_s to zero, it is possible to generate this irreducible background on its own. Included in this background are the important vector boson triple and quartic vertices with the same final states as the signal process.

The other irreducible background of this analysis is QCD background. It can only be generated in combination with the EW background by restoring α_s to its original value. When doing counting experiments, the expected number of QCD background events, B^{QCD}, can be extracted from the WHIZARD generated combined expected number of EW and QCD background events, $B^{\mathrm{EW+QCD}}$, by a subtraction of the expected number of EW events:

$$B^{\mathrm{QCD}} \equiv B^{\mathrm{EW+QCD}} - B^{\mathrm{EW}}. \tag{4.7}$$

All other parameters and production mechanisms for these backgrounds are exactly the same as in the signal samples generated by WHIZARD. The cross sections are:

$$\sigma_{\mathrm{EW}} = \quad 63.8\,\mathrm{fb} \tag{4.8}$$

$$\sigma_{\mathrm{EW+QCD}} = 382.7\,\mathrm{fb}. \tag{4.9}$$

A total of 6 000 events for EW and 29 000 for EW+QCD are available. Typical Feynman graphs can be seen in Figure 4.6.

Sample	Number of Events		σ [pb]
	no pile-up	pile-up	
$t\bar{t}$	729 534	641 504	467.3
$T\bar{T}$	1 373 788	–	60.5
Wt	59 482	–	3.7
EW	6 000	6 000	0.1
EW+QCD	29 000	29 000	0.4
$W \to e\nu + 0j$	19 950	19 926	147.1
$+ 1j$	29 998	29 998	96.2
$+ 2j$	99 999	99 590	96.1
$+ 3j$	149 997	146 031	79.4
$+ 4j$	69 956	69 956	43.4
$+ 5j$	29 975	29 975	25.0
$W \to \mu\nu + 0j$	19 999	19 999	109.8
$+ 1j$	29 949	29 949	84.2
$+ 2j$	98 000	99 750	87.9
$+ 3j$	149 748	149 248	75.0
$+ 4j$	69 999	69 999	42.7
$+ 5j$	29 749	29 999	24.7
$W \to \tau\nu + 0j$	10 000	10 000	20.3
$+ 1j$	15 000	15 000	19.5
$+ 2j$	50 000	50 000	20.3
$+ 3j$	72 700	72 104	16.2
$+ 4j$	35 000	34 750	10.0
$+ 5j$	15 000	20 000	6.2
$Z \to ee + 0j$	39 000	39 000	12.0
$+ 1j$	75 000	74 750	8.1
$+ 2j$	104 996	104 746	8.8
$+ 3j$	99 998	99 748	7.2
$+ 4j$	50 000	49 750	4.1
$+ 5j$	29 954	29 954	2.2
$Z \to \mu\mu + 0j$	39 748	39 748	7.4
$+ 1j$	98 899	79 478	6.8
$+ 2j$	98 000	98 492	7.6
$+ 3j$	99 750	99 500	7.1
$+ 4j$	48 850	48 861	3.8
$+ 5j$	30 000	29 750	2.1
$Z \to \tau\tau \to ll + 0j$	59 383	62 879	0.3
$+ 1j$	45 500	45 250	0.3
$+ 2j$	338 748	348 498	0.3
$+ 3j$	217 011	216 761	0.3
$+ 4j$	32 499	33 198	0.2
$+ 5j$	16 000	16 000	0.1

Table 4.3: Background samples considered in this work. All samples are generated for a center-of-mass energy of 14 TeV. Because negative event weights occur when generating top pair samples with MC@NLO, "Number of Events" refers to the recalculated number of events including these weights in the case of $t\bar{t}$ (central production) and $T\bar{T}$ (private production).

4.4 Discriminating Variables

Variable	Description	Separation $\langle S^2 \rangle$ no pile-up	pile-up
b_1	Largest flavor tag weight	0.621	0.528
b_2	Second to largest flavor tag weight	0.411	0.381
m^{tagjet}	Invariant mass of tagging jets	0.296	0.296
$p_{\text{T},1}^{\text{minijet}}$	Largest minijet transverse momentum	0.188	0.160
ζ	Lepton centrality	0.182	0.180
$p_{\text{T},1}^{\text{lepton}}$	Largest lepton transverse momentum	0.183	0.177
$\Delta\eta^{\text{tagjets}}$	Pseudorapidity separation of tagging jets	0.161	0.143
$p_{\text{T}}^{\text{bal}}$	Transverse momentum balance	0.144	0.125
$p_{\text{T},2}^{\text{minijet}}$	Second to largest minijet transverse momentum	0.120	0.058
$p_{\text{T},2}^{\text{lepton}}$	Second to largest lepton transverse momentum	0.098	0.106
$E_{\text{T}}^{\text{miss}}$	Missing transverse energy	0.073	0.081
$p_{\text{T},+}^{\text{tagjet}}$	Transverse momentum of tagging jet with $\eta \geq 0$	0.064	0.068
$m_{\text{T}}^{\text{VBS}}$	Transverse mass	0.062	0.077
$p_{\text{T},-}^{\text{tagjet}}$	Transverse momentum of tagging jet with $\eta < 0$	0.056	0.075
x_1	Lepton momentum fraction 1	0.025	0.028
x_2	Lepton momentum fraction 2	0.025	0.022

Table 4.4: Variables used for the training of the multivariate method, their meaning and the separation $\langle S^2 \rangle$ (see Equation 4.21) between signal (φ resonance with $m = 850$ GeV, $g = 2.0$) and background ($T\bar{T}$, Wt, EW+QCD, Z+jets, W+jets) for no pile-up and pile-up samples. The order of the variables is chosen to represent the degree of separation of the samples without pile-up.

The properties of the reconstructed physics objects are represented by various variables like energy or momentum. Other variables connected to the VBS process of this analysis are derived from these variables to result in discrimination between signal and background. Variables used within this work and their separation between signal and background are listed in Table 4.4 and are explained in the text below.

Two *tagging jets* are characteristic for VBS (see Chapter 3.4). Tagging jets are produced directly from valence quarks and are therefore rather energetic. They also tend to be well separated in pseudorapidity η. They are therefore defined as the jets with the largest transverse momentum p_{T} per hemisphere of pseudorapidity. An important variable is the separation in pseudorapidity between the two tagging jets, $\Delta\eta^{\text{tagjets}}$. This variable is expected to be large on average for the signal process. The peculiarity of tagging jets also leads to increased di-jet invariant mass of the tagging jets, m^{tagjet}, compared to background.

Minijets are defined as those jets with pseudorapidity between the two tagging jet candidates:

$$\eta_{-}^{\text{tagjet}} < \eta^{\text{minijet}} < \eta_{+}^{\text{tagjet}}. \tag{4.10}$$

Due to lack of color flow between the decay products of the initial protons and because in leading order only two jets are in the final state for the signal process, minijets are expected to be well suppressed for the signal process. Properties of the minijets are therefore expected to result in good separation between signal and background. On the other side, minijet variables are susceptible to be influenced by effects of pile-up.

The *lepton centrality* ζ is introduced because the leptons of the di-leptonic VBS of this work are expected to be located inside the pseudorapidity range between the tagging jets. It is defined as:

$$\zeta = \min\{\Delta\eta_{-}, \Delta\eta_{+}\} \tag{4.11}$$

with:

$$\Delta\eta_{-} = \min\{\eta_1^{\text{lepton}}, \eta_2^{\text{lepton}}\} - \min\{\eta_1^{\text{tagjet}}, \eta_2^{\text{tagjet}}\} \tag{4.12}$$

$$\Delta\eta_{+} = \max\{\eta_1^{\text{tagjet}}, \eta_2^{\text{tagjet}}\} - \max\{\eta_1^{\text{lepton}}, \eta_2^{\text{lepton}}\}. \tag{4.13}$$

This represents the closest pseudorapidity distance between the tagging jets and the two leptons with the largest p_{T}. If at least one lepton is at a larger pseudorapidity η than the tagging jets, the lepton centrality becomes negative. Leptons in VBS tend to have larger transverse momenta $p_{\text{T},1}^{\text{lepton}}$, $p_{\text{T},2}^{\text{lepton}}$ compared to most of the backgrounds. Therefore, these variables are expected to result in good separation between signal and background.

Flavor tag weight is the output of a b-tagger [86] and should separate the signal from the backgrounds with bottom quarks in the final state. Signal events will usually contain no real bottom quarks. A combined b-tagger that uses the information of the impact parameters of the tracks and the secondary vertex[6] was used in this analysis. Other flavor tagging algorithms are also available within the ATLAS software framework [86]. The two highest flavor tag weights b_1 and b_2 are chosen as variables here.

The *transverse momentum balance* $p_{\text{T}}^{\text{bal}}$ represents the completeness in the reconstruction of tagging jets, leptons and missing energy transverse momentum. For a true $qql\nu l\nu$ final state like the signal, this value is expected to be close to zero because the event is completely reconstructed. The transverse momentum balance can be calculated by:

$$p_{\text{T}}^{\text{bal}} = \sqrt{\left(p_{\text{T,x}}^{\text{bal}}\right)^2 + \left(p_{\text{T,y}}^{\text{bal}}\right)^2} \tag{4.14}$$

$$p_{\text{T,x}}^{\text{bal}} = p_{1,\text{x}}^{\text{tagjet}} + p_{2,\text{x}}^{\text{tagjet}} + p_{1,\text{x}}^{\text{lepton}} + p_{2,\text{x}}^{\text{lepton}} + p_{\text{x}}^{\text{miss}} \tag{4.15}$$

$$p_{\text{T,y}}^{\text{bal}} = p_{1,\text{y}}^{\text{tagjet}} + p_{2,\text{y}}^{\text{tagjet}} + p_{1,\text{y}}^{\text{lepton}} + p_{2,\text{y}}^{\text{lepton}} + p_{\text{y}}^{\text{miss}}. \tag{4.16}$$

[6] IP3D+SV1 tagger [86]

The *transverse mass* $m_{\mathrm{T}}^{\mathrm{VBS}}$ is a very rough approximation of the invariant mass of the boson pair when identifying p^{miss} with the neutrinos. The real invariant mass of the boson pair is not direct accesible in the di-leptonic final state. The transverse mass is defined as:

$$m_{\mathrm{T}}^2 \equiv \begin{pmatrix} E_{\mathrm{T}}^{\mathrm{leplep}} + E_{\mathrm{T}}^{\mathrm{miss}} \\ \mathbf{p}_{\mathrm{T}}^{\mathrm{leplep}} + \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}} \end{pmatrix}^2 \tag{4.17}$$

$$= 2\mathbf{p}_{\mathrm{T}}^{\mathrm{leplep}}\mathbf{p}_{\mathrm{T}}^{\mathrm{miss}} \left(1 - \cos\Delta\varphi(\mathbf{p}_{\mathrm{T}}^{\mathrm{leplep}}, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})\right). \tag{4.18}$$

Neutrinos cannot be detected directly but they are visible in the detector as missing transverse momentum and missing energy. Therefore, and because two neutrinos appear in the final state, the properties of each particular neutrino are not direct accesible in the experiment. The momenta of the vector bosons from the decaying resonance can therefore not be completely reconstructed. Nevertheless, these values can be estimated using *collinear approximation*.

Because there is a large mass difference between the resonance and the vector bosons, the vector bosons are highly Lorentz boosted. For the leptonic decay channel of the W bosons, the opening angle between the lepton and the neutrino of each decaying W boson should therefore be small in the lab frame and is assumed here to vanish. Also, the decay products of the vector bosons are expected to be produced in the same direction as the parent vector boson.

Therefore, the transverse momentum of each particular W boson can be expressed as a scaled transverse momentum of the corresponding final state lepton:

$$\mathbf{p}_{\mathrm{T}}^{W1/W2} \equiv \frac{\mathbf{p}_{\mathrm{T}}^{\mathrm{lep1/lep2}}}{x_{1/2}}. \tag{4.19}$$

When assuming that the missing momentum p^{miss} is equal to the sum of the neutrino momenta, the momentum fractions x_1 and x_2 can be calculated [87]:

$$x_{1/2} = \frac{p_{1,\mathrm{x}}^{\mathrm{lepton}} p_{2,\mathrm{y}}^{\mathrm{lepton}} - p_{1,\mathrm{y}}^{\mathrm{lepton}} p_{2,\mathrm{x}}^{\mathrm{lepton}}}{\pm p_{\mathrm{x}}^{\mathrm{miss}} p_{2/1,\mathrm{y}}^{\mathrm{lepton}} \mp p_{\mathrm{y}}^{\mathrm{miss}} p_{2/1,\mathrm{x}}^{\mathrm{lepton}} + p_{1,\mathrm{x}}^{\mathrm{lepton}} p_{2,\mathrm{y}}^{\mathrm{lepton}} - p_{1,\mathrm{y}}^{\mathrm{lepton}} p_{2,\mathrm{x}}^{\mathrm{lepton}}}. \tag{4.20}$$

With these results and Equation 4.19, the momentum of each W boson can be approximately reconstructed.

The separation $\langle S^2 \rangle$ between signal and background of a variable x is defined by [88]:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{\left(f\left(x|H_0\right) - f\left(x|H_1\right)\right)^2}{f\left(x|H_0\right) + f\left(x|H_1\right)} \mathrm{d}y, \tag{4.21}$$

where $f\left(x|H_0\right)$ is the probability density function (p.d.f.) of the variable x given the signal hypothesis and $f\left(x|H_1\right)$ the p.d.f. of the variable for the background hypothesis. The separation is one if there is no overlap between the two functions and zero if the two shapes are identical.

A full list of all variables with expected discrimination power between signal and background and their separation between signal (φ resonance with $m = 850$ GeV

and $g = 2.0$) and background ($T\bar{T}$, Wt, EW+QCD, Z+jets, W+jets) for no pile-up and pile-up samples can be found in Table 4.4. Similar results are observed for other resonances. Distributions of the variables for no pile-up and pile-up samples can be found in Appendix A. All the variables of this section are used as input for the multivariate analysis to separate the background from the signal. This is explained in detail Chapter 5.

The largest separation is observed for the two flavor tag weights b_1 and b_2. This is found to be caused by the top pair background due to two real b quarks. When leaving out this background, the separation of the flavor tag weights drops down to about 0.1. Top pair background is the largest reducible background of this analysis (see for example Figure 5.6) and the large separation of the flavor tag weights is an important discriminant.

Very low separation power is observed for the lepton momentum fractions x_1 and x_2. Nevertheless, a slight worsening of the final results (see Chapter 6.3) has been observed when dropping these variables, so they have been restored for the multivariate analysis.

In general, only small differences between variables of pile-up samples and variables of samples without pile-up are observed. The difference in terms of separation between both cases (see Table 5.6) is found to be of the same order as the difference when changing the resonance or the coupling g. Without comparison algorithms like a Kolmogorov-Smirnov test [89], no significant difference between the shape of the variables for pile-up and no pile-up (see Appendix A) is observable. An application of the Kolmogorov-Smirnov test results in less than 5% deviation between both cases for all variables.

5 Multivariate Analysis

5.1 Pre-cuts

Trigger

To take advantage of the leptons occurring in the signal process, a combination of electron and muon triggers [1] is used. An event passes the trigger requirements if one of the two triggers is fired. The electron trigger demands an isolated electron candidate of at least 25 GeV whereas the muon trigger demands an isolated muon candidate of at least 20 GeV.

The averaged trigger efficiency of the irreducible QCD background sample (see Section 4.3.2) in the transverse momentum range between 30 GeV and 100 GeV is found to be essentially constant at 0.9816 ± 0.0014 [1]. An trigger efficiency of 0.98 is multiplicatively taken into account for the final significance calculation (see Section 6.3) after all other cuts by assuming a factorization in the cut efficiencies.

The following fiducial transverse momentum cuts of the leptons make sure to be in the plateau region of the trigger efficiency:

$$p_{\mathrm{T},1}^{\mathrm{lepton}} \geq 30 \text{ GeV} \tag{5.1}$$

$$p_{\mathrm{T},2}^{\mathrm{lepton}} \geq 30 \text{ GeV}. \tag{5.2}$$

Overlap Removal

Reconstruction output data consists of lists of reconstructed physics object *candidates*. Because the reconstruction algorithms for the physics objects use the complete event as input, one physical object in the detector (e.g. electron, muon) can be identified by more than one algorithm. It is possible for example, that one electron in the detector is identified as electron *and* jet candidate at the same time and is represented in the list of electron *and* jet candidates. This ambiguity has to be resolved by choosing one of these hypotheses. This is called *overlap removal*.

In this analysis, jets found within a radius $\Delta R \equiv \sqrt{\Delta\eta^2 + \Delta\varphi^2} < 0.3$ in the η-φ plane of an electron or muon candidate are removed.

Fiducial Pre-cuts

Furthermore, quality cuts on reconstructed physics objects are applied to remove any possible effect of generator level cuts.

Electrons have to fulfill the ATLAS tight electron identification [90], must be found within $|\eta| \leq 2.5$ (because of the gap region between barrel and endcaps, electrons between $1.35 < |\eta| < 1.57$ are excluded) and must have a transverse momentum of $p_{\mathrm{T}} \geq 15$ GeV. To ensure isolation, only electron candidates with $E_{\mathrm{T}} < 10$ GeV

(including energy deposition in the hadronic calorimeter) within a cone of $\Delta R < 0.2$ around the candidate are accepted.

Because muons are traversing the whole detector before they reach the muon chambers, matching tracks from the inner detector to the muon system have to be found to improve the quality of the muon reconstruction. Isolation is ensured by rejecting muon candidates with additional energy of $E \geq 10$ GeV within $\Delta R < 0.2$ around the muon candidate. The candidate also has to have $|\eta| \leq 2.7$ and $p_{\mathrm{T}} \geq 15$ GeV.

Jet candidates are identified by a cone algorithm with $\Delta R = 0.4$ and have to fulfill $|\eta| \leq 4.2$ and $p_{\mathrm{T}} \geq 15$ GeV. In the event selection, the last cut is tightened to:

$$p_{\mathrm{T},1}^{\mathrm{jet}} \geq 30 \text{ GeV} \tag{5.3}$$

$$p_{\mathrm{T},2}^{\mathrm{jet}} \geq 30 \text{ GeV}. \tag{5.4}$$

Another fiducial cut on the angle between the tagging jets is applied equivalent to the one on generator level:

$$|\Delta\eta^{\mathrm{tagjets}}| \geq 3.0. \tag{5.5}$$

A cut on the invariant mass of the lepton pair is applied to reduce the Z+jets background:

$$m^{\mathrm{leplep}} \geq 150 \text{ GeV}. \tag{5.6}$$

5.2 Multivariate Methods

The event topology of the di-leptonic VBS signal process with additional resonances considered within this work (see Chapter 3.4) differs from background and allows for cuts on discriminating variables (see Table 4.4) to separate signal and background. Because a large suppression of background events over four orders of magnitude while conserving a relatively low number (~ 100) of signal events is necessary in order to reach the goals for discovery (see Chapter 6.1.1), a simple cut analysis is found to be too inefficient. Therefore, different multivariate methods are applied in this work.

A *Boosted Decision Tree* [91] (BDT) is finally found to be the most efficient tool for the event topology of di-leptonic VBS with additional resonances. Furthermore, a *Fisher Linear Discriminant* [92] (LD) and a *Multi-Layer Perceptron* [93] (MLP) are applied for comparison but have not been applied in the final analysis due to lower separating power between signal and background.

5.2.1 Toolkit for Multivariate Analysis

A tool to realize the multivariate methods described above is the Toolkit for Multivariate Analysis, TMVA [91]. It is part of the ROOT package frequently used in high energy physics. All multivariate methods of this work are implemented in TMVA.

The methods used within TMVA are *supervised learning* algorithms. They use training events for which the output is known and test events for evaluation. This means that the input events are internally split into one *training* and one *testing* collection to guarantee statistical independence. In theory of multivariante methods it can be shown that statistical independent classification can only be warranted with three independent sub-samples: One sample for parameter optimization, another

sample for performance validation and the last one for overtraining (see Section 5.2.2) detection. The merging of the last two by TMVA is done to increase statistics. Bias introduced due to this optimization does not affect the analysis as long as classification cut efficiencies are independently validated with data. In this study, independent validation with real data was not possible because of the low integrated luminosity collected by ATLAS up to now.

5.2.2 Boosted Decision Tree

The main classifier of this thesis is a Boosted Decision Tree (BDT). Similar to a living tree, growth starts with a single node using the variable with the highest *separation* between signal and background. This node represents a single cut on this variable.

Different figures of merit for separation are available, in this work the *Gini index* is selected. It calculates an estimate of the separation S out of the purity p of the signal sample:

$$S = p(1-p). \tag{5.7}$$

Purity is the quotient of the number of signal events passing a cut divided by the total (signal + background) number of events passing this cut.

Two sub-nodes are generated following this node. One sub-node if the event passes the cut and another one if the event does not pass. For every new node this method is repeated until the number of events in this node is below a certain minimal value. If this happens, the node becomes a leaf node and the growing stops at this point.

It is important to take care of statistical fluctuations when applying a BDT because small fluctuations at the root node can lead to big changes later in the tree. This is taken into account by using not only one tree but a forest of single trees. All trees in the forest are derived from the same training samples but special weights are allocated to the events of a tree. An event gets *boosted* by receiving a larger weight in a particular tree if the same event has been misclassified in the previous tree.

The decision r_{ij} of each tree j for a given event i is defined as $r_{ij} = +1$ for signal decision and $r_i = -1$ for background decision. Also, the classification performance of each tree j is parametrized by a weight β_j [91].

With these informations, the response of a forest of N trees to a particular event i is defined as:

$$r_i = \frac{1}{N} \sum_{j=1}^{N} \beta_j r_{ij}. \tag{5.8}$$

The *gradient boost* option is used for the training in this thesis because it is typically less susceptible to overtraining [91]. Its robustness is enhanced by reducing the learning rate of the algorithm through the *shrinkage* parameter which controls the weights of the individual trees. A small shrinkage $(0.1 \ldots 0.3)$ demands more trees to be grown but improves the accuracy of the boosting. A shrinkage of 0.3 is selected for all trainings of this work.

Additional stabilization can be achieved by using random sub-samples of the training events for growing the trees. This is called *bagging* and was enabled within this work with the fraction of events for each sub-sample of 0.6.

A common challenge when applying a BDT is to deal with *overtraining* (see Section 5.3.3). To minimize this effect, statistically insignificant branches can be removed by *pruning*. They get cut back from the bottom up after the growing of the tree is finished. A pruned tree has less degrees of freedom and is therefore less susceptible to overtraining.

The *cost complexity* pruning method [94] is used in this work. It calculates a gain in terms of misclassified events of a sub-tree below a node compared to the node itself. The node with the smallest cost complexity ρ [91] in the tree is recursively pruned away as long as ρ is less than the *pruning strength*. In this work, a pruning strength of 50 is selected.

5.3 Application of a Boosted Decision Tree

For the training of a BDT with TMVA, the input samples of Chapter 4.3 and the input variables of Chapter 4.4 are selected, all fiducial pre-cuts (Chapter 5.1) are applied and the options as explained in Chapter 5.2.2 are chosen.

Compared to a previous study [1], where a different signal definition for the training has been used[1], the signal definition of Section 4.3.1 is completely taken into account here. Also, negative event weights are fully incorporated for training and evaluation of the BDT which was not done in [1]. This is important because top pair background is generated with MC@NLO introducing negative weights. Also samples with a negative number of events due to statistical fluctuations when subtracting EW background as explained in detail in Chapter 4.3.1 occur. They are now treated as samples with a positive number of events but negative weights.

As described in Section 5.2.1, the input samples are internally split into one training and one testing sub-sample by TMVA. This splitting is performed randomly. Other splitting methods like alternate splitting (events are chosen in alternating turns for training and testing) and block splitting (forces to use the events of signal and background as they occur in the source files) are checked with similar results in the final statistical analysis. Nevertheless, random splitting is selected to remove every possible bias from creation of the trees.

After the training is finished, overtraining is checked and the results are evaluated. The exact procedure and the results are explained in the following sections. The φ resonance is selected as an example to present the results. This resonance is chosen because it shows the largest yields of all resonances and therefore the best expected separation between signal and background. In addition, 24 000 fast-simulated ATLFAST-II events without pile-up and additional 30 000 fully-simulated GEANT events with pile-up are available for this resonance (see Figure 5.1). Training with other resonances is done equivalently. The results are analyzed in Chapter 6.4 and are summed in Appendix C.

[1] With the notation of Section 4.3.1, the expected number of signal events for the training has been defined there as $S \equiv S^{\mathrm{EW}+r}(m, g)$.

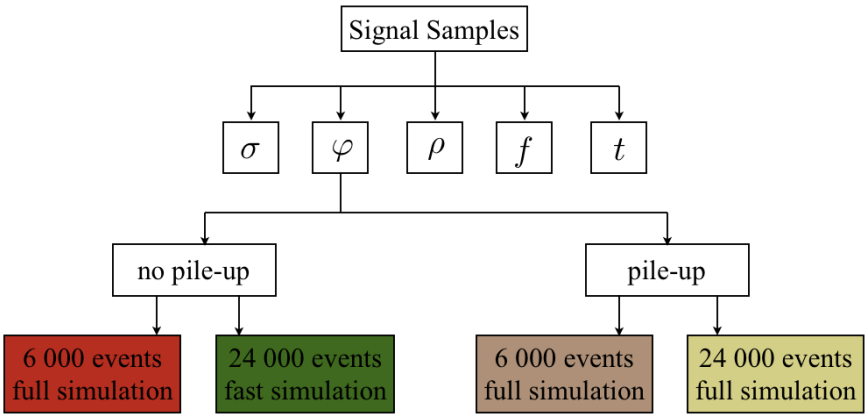


Figure 5.1: Visualization of the signal samples used for the training of the BDT. As an example, the φ resonance is shown in detail. All other samples and other properties are listed in Table 4.1. The colors of the signal sample definitions introduced within this graph are used throughout this section. "full simulation" refers to fully-simulated signal samples with GEANT and "fast simulation" to fast-simulated signal samples with ATLFAST-II. In the case of pile-up, only fully-simulated samples are available and splitting in a small and big sample is done to adapt the number of events to the case without pile-up as explained in the text.

5.3.1 Input Samples

Signal

Various possible signal processes in terms of different resonances are investigated within this work. The mass of all resonances is fixed ($m = 850$ GeV) but different couplings and resonances are possible. Furthermore there are samples with and without pile-up that have to be compared. All signal samples of this work are listed in Table 4.1. The samples of the φ resonance are visualized as an example in Figure 5.1 and the colors of the signal sample definitions introduced within this graph are used throughout this section.

In order to reduce the number of trainings, *only one coupling g per resonance* is selected in the training phase of this analysis. A study of the influence of the effect of different couplings in the training but evaluation with one fixed coupling resulted in deviations comparable to the statistical error of the training of about 2.8% (see Section 6.6). This behavior is expected because the Lagrangian depends linearly on the couplings (see Equation 2.41) and the cross section is therefore also scaled by a constant factor. On the other side, the results of the training are not influenced by an additional constant factor in the signal cross section.

The following samples are selected for training and evaluation of the BDT in the main analysis as explained in detail in the following text:

	Training		Evaluation
no pile-up:	6 000 events — full simulation	→	6 000 events — full simulation
pile-up:	6 000 events — full simulation	→	6 000 events — full simulation

In the main part of the study, all GEANT fully-simulated signal events without pile-up (red) are chosen as training events for the analysis without pile-up. Because there are only 6 000 events available for this sample, the same events are also used for evaluation. This introduces a bias because training events and evaluation events are not statistically independent. In Chapter 6.6 this is avoided by a training with ATLFAST-II fast-simulated signal events without pile-up (green). It is shown there, that the deviation between both training sample selections is of the same order as the statistical error of the training.

A larger amount of fully-simulated signal events exists for the pile-up case and independent samples for training and evaluation would be possible. Nevertheless, a sample definition equivalent to the one without pile-up is used to stay on the safe side when comparing results with and without pile-up. Therefore, a signal sample with only 6 000 fully-simulated signal events (brown) has been extracted from the main sample and used for training and evaluation for the analysis with pile-up.

No significant improvement is found when additionally using the pile-up sample with more events (yellow) in training and/or evaluation as done in Section 6.6. There, the larger amount of events is used for the estimation of the statistical error of the training. This error is found to be small and the sample definition of this chapter can therefore be seen as statistically safe.

Background

Only the important background processes are selected for the training in order to save computation time.

Significance studies as explained in Section 6.3 have revealed that the results are very sensitive to the choice of the top pair background sample in the training phase. When training with the centrally produced top pair sample $t\bar{t}$ (see Chapter 4.3.2), almost all discovery sensitivity is gone while a training with the locally produced lepton-filtered top pair sample $T\bar{T}$ results in higher sensitivity. This effect is expected due to the large difference in the number of events after cuts between these samples and was the main reason for privately producing the additional sample $T\bar{T}$. Because no pile-up sample is available for $T\bar{T}$, this top pair background sample without pile-up is also used to execute the training with pile-up.

Also the Wt background sample (see Chapter 4.3.2) does not have pile-up included. So the sample without pile-up is used for trainings with and without pile-up. The EW+QCD sample is the only background process with pile-up used for the TMVA training of the multivariate method.

W+jets background samples are left out of the TMVA training because this background vanishes already at small cuts on the BDT output (see Figure 5.6). Z+jets background events are suppressed by a fiducial pre-cut on the invariant mass of the lepton pair (see Equation 5.6) and a cut on the angle between the leptons with the largest p_{T} after the evaluation of the BDT. Another reason for leaving W+jets and Z+jets background samples out of the training is that this type of background is split into various jet multiplicities and final states making it computational expensive when including it into the training.

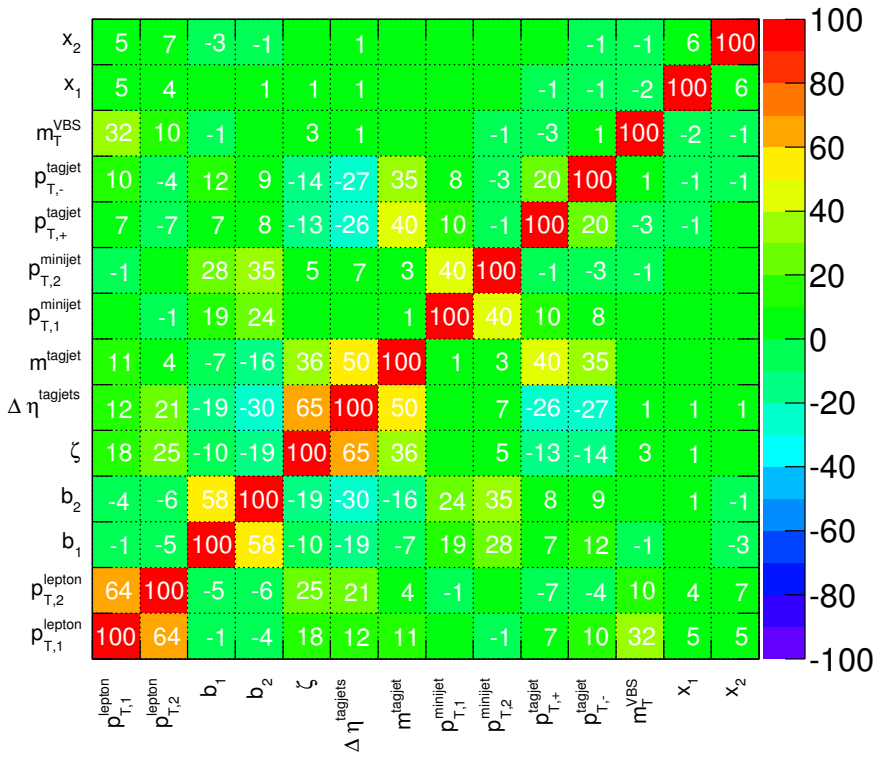


Figure 5.2: Linear correlations between the BDT training input variables for the φ resonance ($m = 850$ GeV) signal sample with pile-up.

5.3.2 Input Variables

All the variables explained in detail in Chapter 4.4 are chosen as input variables for the training of the BDT. Several studies with fewer variables are performed. Variables with a separation of larger than approximately 0.1 are found to be indispensable (see Table 4.4). The remaining variables result in smaller enhancements for the final result in terms of discovery significance (see Section 6.3). A worsening due to too large numbers of input variables is not observed for a BDT whereas a MLP classifier shows a huge sensitivity on the number of input variables.

To characterize the correlations between the different variables, the linear correlation between all variables used as input for the training is shown in Figure 5.2. The results are presented for the variables of the φ resonance ($m = 850$ GeV) signal samples with pile-up.

Some positive, relatively large linear correlations are observed between some of the variables, but most of the variables are uncorrelated. Background samples and samples of other resonances show a similar behavior in these variables. Similarly, the correlations between the variables for samples with and without pile-up are essentially the same.

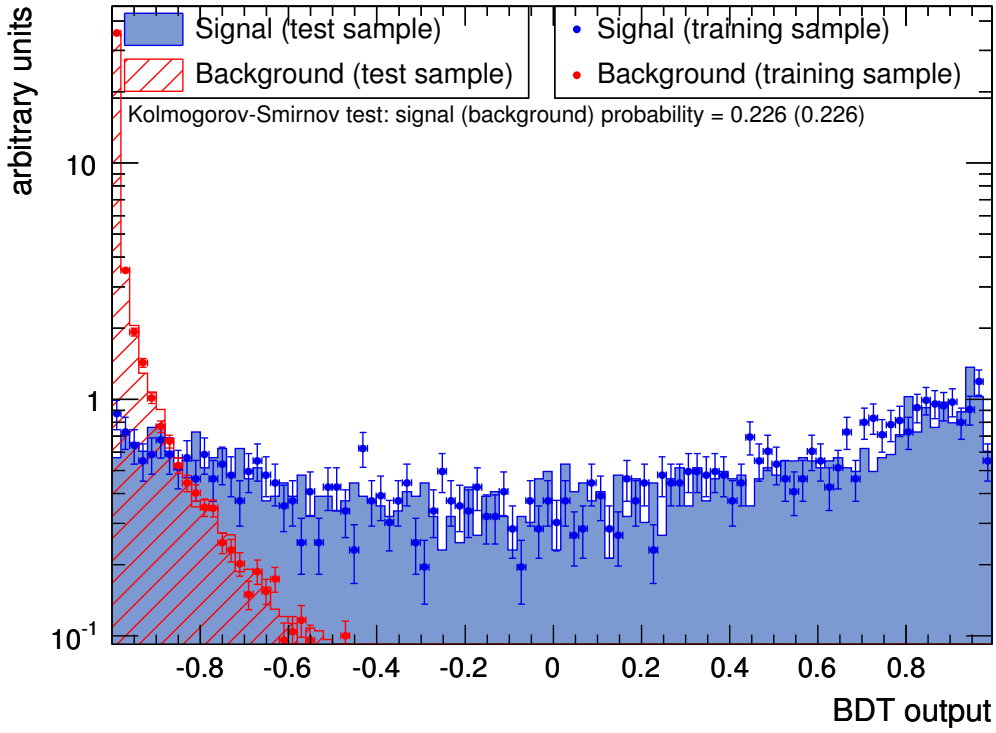


Figure 5.3: Overlay of the test sample and the training sample for a training (including pruning) of a BDT. The shapes are compared applying a Kolmogorov-Smirnov test. As signal, the φ resonance with pile-up is chosen.

5.3.3 Overtraining Detection

A multivariate method is said to be overtrained if the training is too specifically adapted to the training data. This means, the method does not find the underlying trend in the sample but rather learns the current sample by heart. It is therefore not reasonable to apply the training output to any independent data sample.

Overtraining can occur if a data sample has too many degrees of freedom and the phase space is therefore too sparsely populated. Too few data points are then adjusted to too many parameters. Overtraining may be detected by comparing the shapes of the test sample and the training sample. This was done with a Kolmogorov-Smirnov test [89], a non-parametric test of equality.

In Figure 5.3 the results are shown for the training of the φ resonance without pile-up. A probability of 0.262 for the signal and the background samples indicates agreement between test and training sample and thus no overtraining.

BDTs are susceptible to overtraining because of their large number of nodes, whereas a Linear Discriminant can hardly be overtrained. To deal with overtraining when using a BDT, *pruning* of the trees as described in Chapter 5.2.2 is applied in the complete analysis.

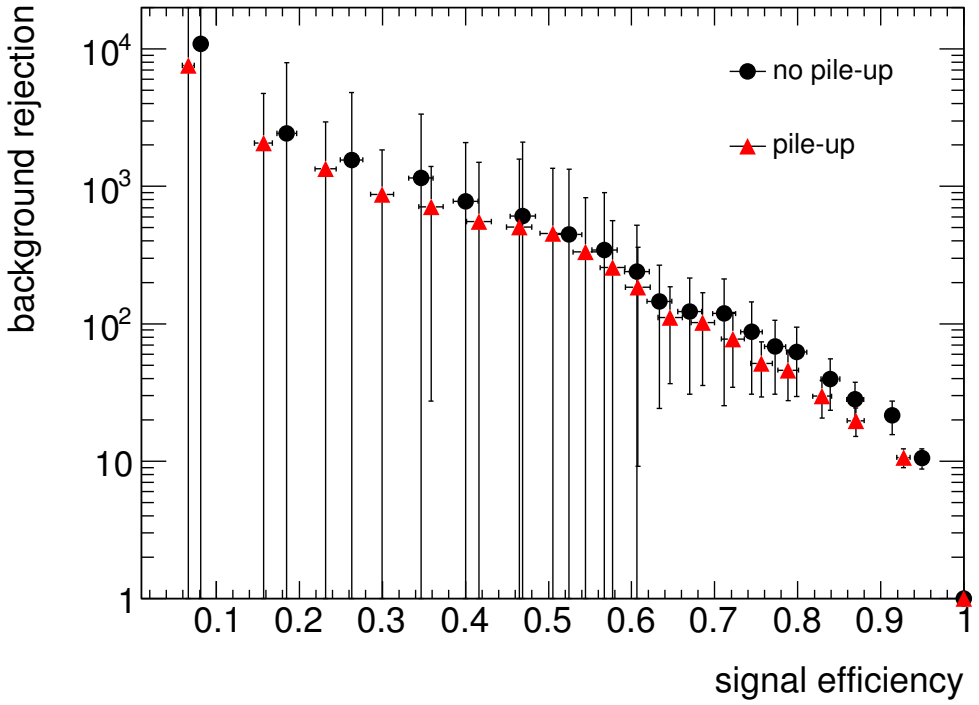


Figure 5.4: Background rejection versus signal efficiency after application of a BDT for the φ resonance with pile-up. All backgrounds are taken into account and an expected integrated luminosity of $100\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed.

5.3.4 Evaluation of the Training Output of a Boosted Decision Tree

The result of the training of a BDT is an abstract *Boosted Decision Tree output variable* r (see Equation 5.8) with $-1 \leq r < +1$. This is the actual value to cut on. No rejection arises for a cut at $r \geq -1$ (this means, that all events are kept) and all events are cut away when cutting at $r > +1$.

Results in terms of signal efficiency versus background rejection can be seen in Figure 5.4 for samples with and without pile-up trained and evaluated as explained in Section 5.3.1. Large suppression of background over more than three orders of magnitude is achieved for the analyses with and without pile-up. The difference between both cases is consistent with statistical fluctuations.

The signal efficiency is calculated as the quotient of the number of signal events after a cut on the BDT output divided by the total number of signal events without any cut on the BDT output:

$$\text{signal efficiency} = \frac{\#\text{ signal events } \; r \geq r_{\text{cut}}}{\#\text{ signal events } \; r \geq -1}. \tag{5.9}$$

Background rejection is defined as the inverse of the background efficiency:

$$\text{background rejection} = \frac{\#\text{ background events } \; r \geq -1}{\#\text{ background events } \; r \geq r_{\text{cut}}}. \tag{5.10}$$

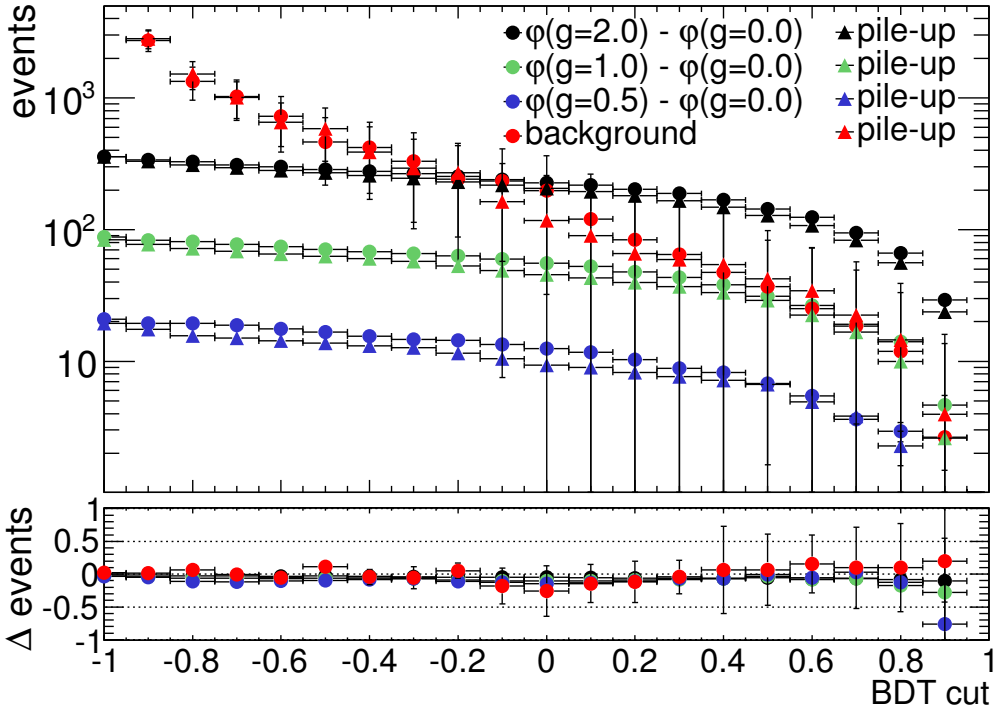


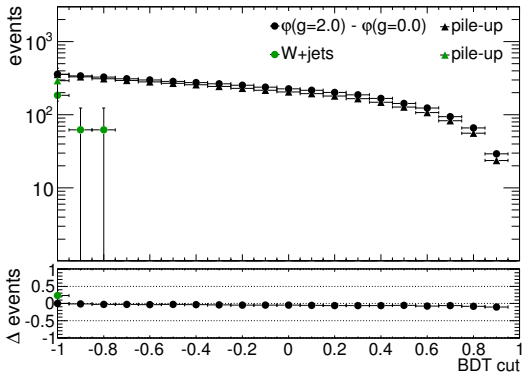
Figure 5.5: Number of events left after a cut on the BDT output for the φ resonance with mass $m = 850$ GeV for three different couplings. Expected events are scaled to 100 fb^{-1} integrated luminosity at $\sqrt{s} = 14$ TeV. A comparison between pile-up (pu) and no pile-up ($\overline{\text{pu}}$) is given in the lower part of the plot: $\Delta\text{events} \equiv (\text{pu} - \overline{\text{pu}})/(\text{pu} + \overline{\text{pu}})$. The angle between the leptons with the largest p_{T} is required to fulfill $\Delta\varphi^{\text{leplep}} > 0.8\pi$. An integrated luminosity of 100 fb^{-1} at $\sqrt{s} = 14$ TeV is assumed and the integrated background Monte Carlo luminosity is scaled to 100 fb^{-1}

The number of events left after a cut on the BDT output of -1 $\leq r_{\text{cut}} <$ +1 for the φ resonance is shown in Figure 5.5. For the pile-up analysis, the training and evaluation are done with pile-up samples and samples without pile-up are used for training and evaluation of the analysis without pile-up. A good global agreement between the results with and without pile-up is observed. This is especially the case when taking the statistical errors into account.

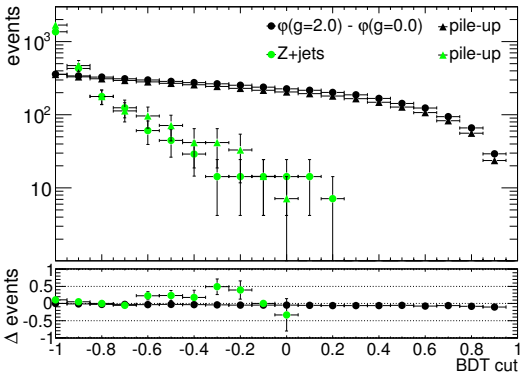
In terms of the coupling g of the resonance it can be seen that a lowering of the coupling value leads to a lowered total number of signal events after the application of the BDT. This is expected because the coupling enters linearly in the Lagrangian of Equation 2.41. Therefore, a constant variation of the coupling strength leads to a constant factor in the cross section and a constant factor for the total number of events. The BDT is trained with only one coupling as described in Section 5.3.1.

A plot of the same cutflow for different signal resonances can be found in Appendix B. The stable behavior under pile-up is similar for other resonances.

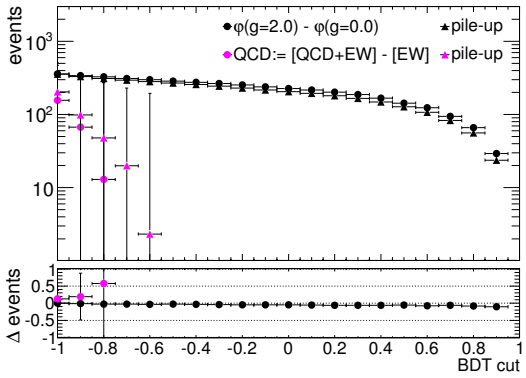
In Figure 5.6, results for the φ resonance are shown for the individual backgrounds. Within the estimated region of best sensitivity for the BDT cut ($r_{\text{cut}} \geq 0.2$, see Chapter 6.3), all backgrounds apart from EW vanish.



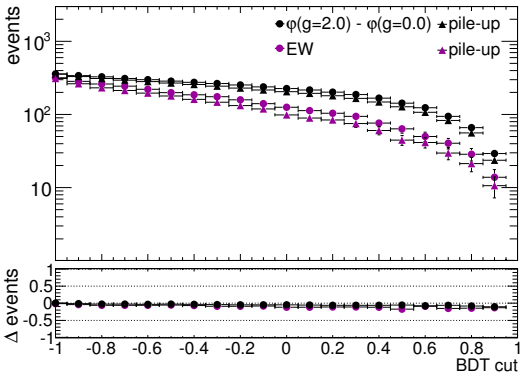
(a) W+jets background



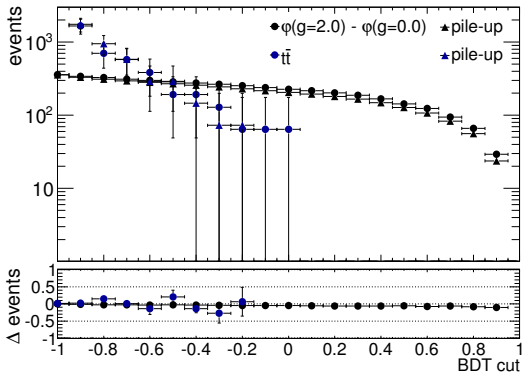
(b) Z+jets background



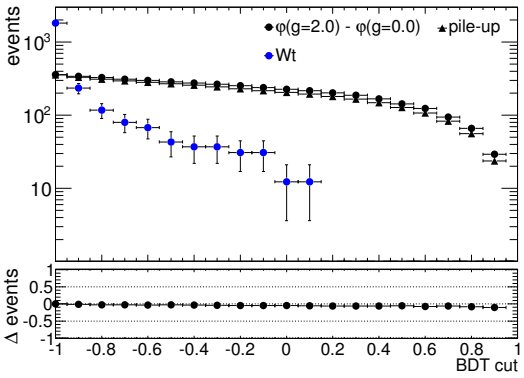
(c) QCD background



(d) EW background

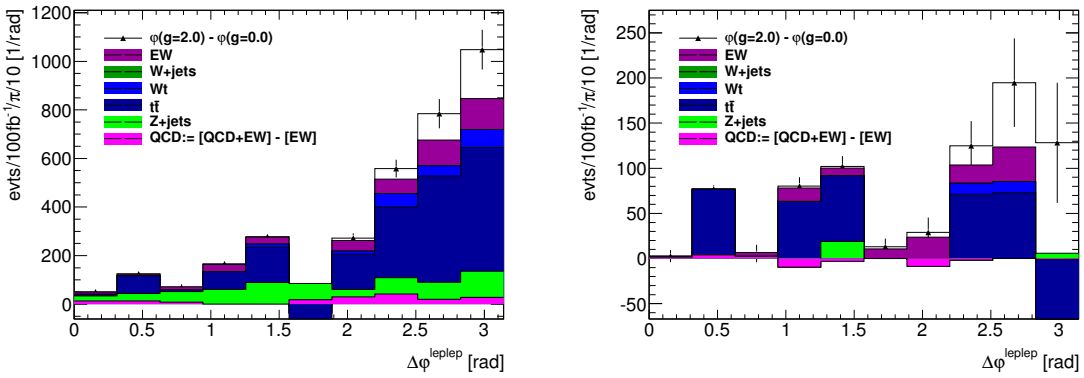


(e) $t\bar{t}$ background



(f) Wt background

Figure 5.6: Number of events left after a cut on the BDT output for signal (φ resonance with mass $m = 850$ GeV and coupling $g = 2.0$) and background. An integrated luminosity of $100\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and the integrated background Monte Carlo luminosity is scaled to $100\,\mathrm{fb}^{-1}$. Pile-up samples are marked with triangles and samples without pile-up are marked with circles. A comparison between results with pile-up (pu) and without pile-up ($\overline{\mathrm{pu}}$) is given in the lower part of the plots: $\Delta\text{events} \equiv (\mathrm{pu} - \overline{\mathrm{pu}})/(\mathrm{pu} + \overline{\mathrm{pu}})$. The angle between the leptons with the largest p_T is required to fulfill $\Delta\varphi^{\mathrm{leplep}} > 0.8\pi$.



(a) Cut on the BDT output: $r \geq -0.8$

(b) Cut on the BDT output: $r \geq 0.0$

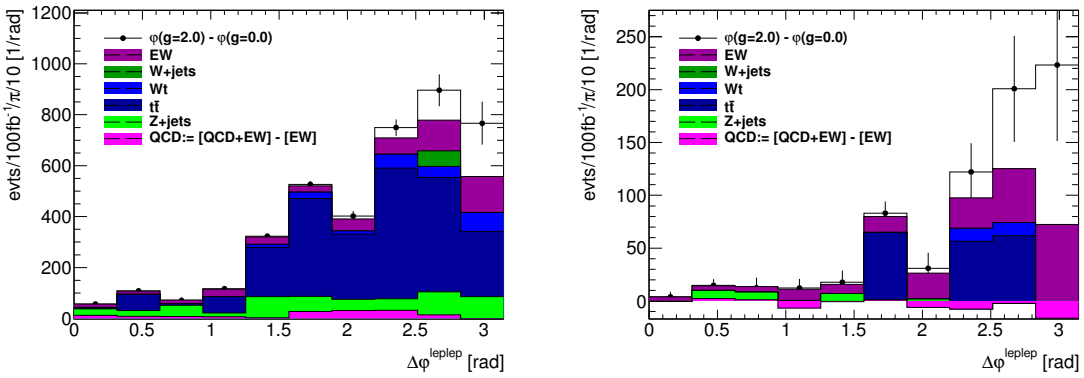
Figure 5.7: Angular separation $\Delta\varphi^{\text{leplep}}$ of signal and background samples *with pile-up* after the cut in the BDT training output. The BDT is trained with the same samples as used for evaluation. The number of events is shown for an expected integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV. A φ resonance with a mass of $m = 850$ GeV and a coupling of $g = 2$ is selected as signal.

5.3.5 Influence of a Boosted Decision Tree on the Angular Separation of the Leptons

To recognize the actual influence of the BDT on the shape of the variables, a look is taken at the angular separation $\Delta\varphi^{\text{leplep}}$ between the two leptons with the largest p_{T}.

For the signal, this variable is expected to peak at π. This is due to kinematics. Because of the high mass ($m = 850$ GeV) of the resonances considered here, they tend to be produced with low transverse momentum and the vector bosons are therefore emitted in opposite direction in the lab frame and are expected to be highly boosted. This boost is passed to the leptons and neutrinos and leads to well separated leptons with high transverse momentum in the detector for the signal process. The angular separation was originally thought to represent a control region for background estimation, because no signal is expected at low values while background is present there (see Figure 5.7(a)). However, after the evaluation of the BDT, the background also peaks at high values [1] although this variable is not part of the set of BDT input variables. The BDT seems to have used the angular separation via correlations to other variables. No significant control region is available after the evaluation of the BDT.

Nevertheless, the angular separation between the leptons with the largest transverse momentum is used here to show the influence of a cut on the BDT output on the shape of this variable and to compare samples with and without pile-up. Shapes of the angular correlation after a very soft cut on the BDT output of $r \geq -0.8$ and moderate cut of $r \geq 0.0$ can be seen in Figure 5.7. There, results of the evaluation of a BDT for the φ resonance with pile-up are shown. The BDT is trained with pile-up samples as described in Section 5.3. An integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed.



(a) Cut on the BDT output: $r \geq -0.8$

(b) Cut on the BDT output: $r \geq 0.0$

Figure 5.8: Angular separation $\Delta\varphi^{\text{leplep}}$ of signal and background samples *without pile-up* after the cut in the BDT training output. The BDT is trained with the same samples as used for evaluation. The number of events is shown for an expected integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV. A φ resonance with a mass of $m = 850$ GeV and a coupling of $g = 2$ is selected as signal.

After a moderate cut of $r \geq 0.0$, Wt[2] and Z+jets background are well suppressed. The irreducible EW and QCD backgrounds are still present in an amount comparable to the number of signal events. Only six top pair background ($t\bar{t}$) events survive the cut, but they receive a relatively large weight. Their uncertainty could be reduced with a larger sample. Negative events are an artifact of the generation process (MC@NLO, see Section 4.3.2), one such event survives the cut.

Negative events for the QCD background arise due to the definition of this background (see Section 4.3.2). Because the EW sample is not affected by the BDT in the same way as EW+QCD, the subtraction of EW can lead to negative events at tighter cuts on the BDT output. This can be seen as a consequence of the sample definition in combination with statistics and should be no general problem of the analysis. With the help of more Monte Carlo statistics this effect should disappear.

The results for the same analysis without pile-up (see Section 5.3.1) are shown in Figure 5.8. The BDT is trained with samples that do not include pile-up and the evaluation of the BDT is applied to the same samples.

To compare the results for pile-up and no pile-up, the shapes in Figure 5.7 (pile-up) can be compared to the shapes in Figure 5.8 (no pile-up). A general similar behavior within statistical uncertainties between the results with and without pile-up can be deduced from these graphs.

Different responses of the BDT for both cases are expected because the training samples are different. The BDT output variable r should correctly be separated in $r_{\text{pile-up}}$ for the training with pile-up samples and $r_{\text{no pile-up}}$ for the training without pile-up. This would prevent a direct comparison and is therefore not done here. The problem is avoided in the analysis of Chapter 6.4 by comparing pile-up and no pile-up results both for a training of the BDT with samples including pile-up.

[2]This is actually the sample without pile-up because no pile-up sample for Wt is available (see Chapter 4.3.2).

6 Statistical Analysis

6.1 Profile Likelihood

6.1.1 p-Values and Equivalent Gaussian Significance

The usual way of quantifying the consistency between data and a given hypothesis is to compute the p-value. In a goodness-of-fit test, the p-value is the probability of obtaining data less or equal compatible as the one actually observed, assuming that the hypothesis in question is true.

Since this is a work on discovery, the hypothesis of "background-only" is to be tested and ruled out. A small p-value indicates low agreement between the observed data and the background-only hypothesis. The observed data is therefore unlikely to be caused by fluctuating background. This is a hint of observed signal.

Related to the p-value is another value often used: The equivalent Gaussian significance[1]. It represents the equivalent fluctuation Z of a standard Gaussian variable and is therefore defined as the number of the standard deviation at which a Gaussian random variable of zero mean would give a one-sided tail area equal to the p-value:

$$p = \int_Z^\infty \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} \mathrm{d}x = 1 - \Phi(Z). \tag{6.1}$$

The cumulative distribution Φ of the standard (zero mean, unit variance) Gaussian is defined as:

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} \mathrm{d}x. \tag{6.2}$$

One can now write the significance Z as a function of the p-value:

$$Z = \Phi^{-1}(1 - p). \tag{6.3}$$

In high energy physics, two values of the significance are quoted frequently:

- 3σ ($Z = 3$, $p \approx 1.35 \times 10^{-3}$) is considered as *evidence* and
- 5σ ($Z = 5$, $p \approx 2.87 \times 10^{-7}$) is considered as *discovery*.

6.1.2 Profile Likelihood

The likelihood of observing n events in an experiment with s^{MC} simulated signal events and b_i^{MC} simulated background events can be written [1, 16] as a product of two Poisson distributions and a Gaussian distribution,

$$L(g, b_i) \equiv L_{n,\, s^{\mathrm{MC}},\, b_i^{\mathrm{MC}}}(g, b_i) = L_n^{(1)}(g, b_i) \times L_{b_i^{\mathrm{MC}}}^{(2)}(b_i) \times L_{s^{\mathrm{MC}}}^{(3)}(g), \tag{6.4}$$

[1] In the literature often only called *significance*. From now on the short name *significance* is also used within this work.

with:

$$
\begin{aligned}
L_n^{(1)}(g, b_i) &= \frac{\left(s(g) + \sum_i b_i\right)^n}{n!} \exp\left[-\left(s(g) + \sum_i b_i\right)\right] \\
L_{b_i^{\mathrm{MC}}}^{(2)}(b_i) &= \prod_i \frac{(\tau_i b_i)^{b_i^{\mathrm{MC}}}}{b_i^{\mathrm{MC}}!} \exp\left[-\tau_i b_i\right] \\
L_{s^{\mathrm{MC}}}^{(3)}(g) &= \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left[-\frac{\left(s^{\mathrm{MC}}(g) - s(g)\right)^2}{2\sigma_s^2}\right].
\end{aligned}
\tag{6.5}
$$

The first Poissonian represents the likelihood of observing n events for a signal model predicting an *expected number of signal events* s and a background model predicting an *expected number of background events* b_i of type i. In this work, the signal model parametrizes additional resonances in VBS with K-matrix unitarization and the background model is representing the Standard Model without Higgs. A conversion from expected signal yields s to the coupling of the resonances g is possible after a fit on Monte Carlo yields and results in a functional dependence $s(g)$. The expectation value of the Poisson distributed observed number of events n can therefore be written as:

$$
E[n] = s + \sum_{i=1}^{N} b_i. \tag{6.6}
$$

The second term in the above likelihood encompasses the Monte Carlo generation of the background processes. Because Monte Carlo samples are usually generated for a different integrated luminosity than the experiment is expected to collect, the simulated background events have to be scaled by a weight factor τ_i:

$$
\tau_i = \frac{L_i^{\mathrm{MC}}}{L^{\mathrm{data}}}. \tag{6.7}
$$

The equivalent integrated luminosity of the Monte Carlo background samples L_i^{MC}, can be calculated from the number of events generated for a given background i after cuts b_i^{MC}, the generator cross section σ_i^{MC} and the cut efficiency ϵ_i:

$$
L_i^{\mathrm{MC}} = \frac{b_i^{\mathrm{MC}}}{\epsilon_i \, \sigma_i^{\mathrm{MC}}}. \tag{6.8}
$$

Expectation values of the Poisson distributed number of generated background events can therefore be written as:

$$
E[b_i^{\mathrm{MC}}] = \tau_i b_i. \tag{6.9}
$$

The third term in the likelihood of Equation 6.4 takes into account the uncertainties on $s(g)$ if the generated yield of signal Monte Carlo $s^{\mathrm{MC}}(g)$ is treated to be Gaussian distributed with the expectation value in experiment $s(g)$ and width σ_s.

Because no measurement on n is available yet, an *Asimov dataset* [95] is assumed:

$$
n = s^{\mathrm{MC}}(g) + \sum_i \frac{b_i^{\mathrm{MC}}}{\tau_i}. \tag{6.10}
$$

The goal is to calculate discovery ranges for certain couplings g of the different resonances considered as signal. A *profile likelihood ratio* [16] can be calculated to test a hypothesized value of g and get rid of additional nuisance parameters b_i in the likelihood of Equation 6.4:

$$\lambda(g) = \frac{L(g, \hat{\hat{b}}_i)}{L(\hat{g}, \hat{b}_i)}. \tag{6.11}$$

The hat notation refers to simultaneous maximum-likelihood estimators and the double-hat refers to conditional maximum-likelihood estimators. In the case of this ratio, $\hat{\hat{b}}_i$ represents the value b_i that maximizes likelihood $L(g, b_i)$ given a fixed g.

For later applications the following quantity is defined:

$$q_g = -2 \ln \lambda(g). \tag{6.12}$$

6.1.3 Connection of Profile Likelihood Ratio to Significance

To claim a discovery, the background-only hypothesis H_0 has to be rejected. Therefore, for every assumed experimental result the level of agreement between the data and the hypothesis $s = 0$ has to be determined. This is given by the p-value, representing equal or less compatibility with H_0:

$$p = P(q_0 \geq q_{0,\text{obs}} | g = 0) = \int_{q_{0,\text{obs}}}^{\infty} f(q_0 | g = 0) \mathrm{d}q_0. \tag{6.13}$$

In this formula, the definition $s(g = 0) = 0$ is applied. The background-only hypothesis is thus equivalent to zero coupling of the resonances, $g = 0$.

The sampling distribution $f(q_0|g = 0)$ of q_0 given the background only hypothesis $g = 0$ can be calculated when assuming a set of regular conditions and a sufficiently large data sample [16] and the integral can be simplified to:

$$p = 1 - \Phi(\sqrt{q_0}). \tag{6.14}$$

The cumulative distribution Φ of the standard Gaussian is defined in Equation 6.2. Combining the result with Equation 6.3 leads to the final result for calculating the discovery significance with the profile likelihood ratio,

$$Z(g) = \sqrt{q_0}, \tag{6.15}$$

with q_g defined in Equation 6.12. The dependence of the significance Z to the coupling g enters with the Asimov dataset (see Equation 6.10). When taking the limit $s \ll b$ and taking into account only terms linear in s, one finds a short formula for the median significance:

$$\tilde{Z}(g) \approx \frac{s}{\sqrt{b}}. \tag{6.16}$$

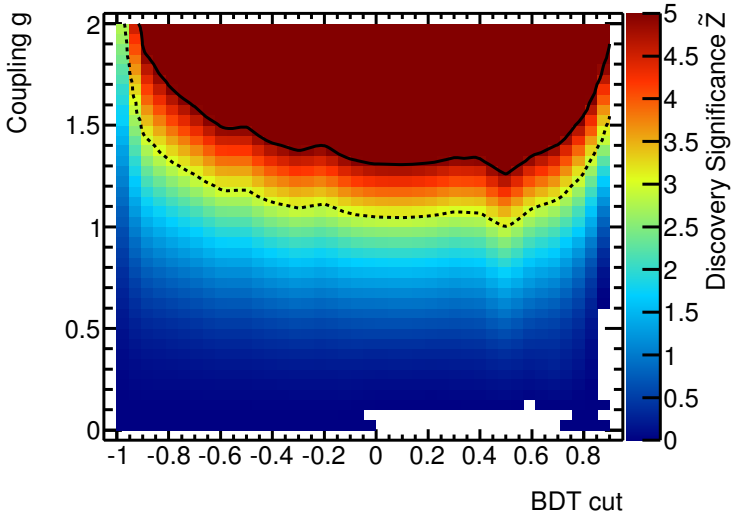


Figure 6.1: $\tilde{Z} = s/\sqrt{b}$ as a first approximation of the discovery significance as a function of the coupling g and the cut on the BDT output. An integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and a cut on the angular separation $\Delta\varphi$ between the leptons with the largest transverse momentum is applied. The plot is done for the φ resonance and includes pile-up.

6.2 Application of an Approximate Significance Formula

As a first estimate, the approximate formula 6.16 can be used to study discovery significances. This equation is often used in particle physics due to the interpretation of $\tilde{Z}$ as a guess of how many units of background uncertainty the signal yields are away when assuming a Poisson distributed background with mean b and error $\sqrt{b}$.

Since studying the dependence of the discovery significance on different couplings g for various resonances is one goal of this study, the approximate discovery significance $\tilde{Z}$ as a function of coupling g and the cut on the BDT output is shown in Figure 6.1. An expected integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and a cut on the angle between the leptons with the largest transverse momentum of $\Delta\varphi > 0.8\pi$ is applied. The result is obtained for a φ resonance including pile-up and the training of the BDT is done as explained in Chapter 5.3.1.

It can be seen that a discovery significance of 5σ can be pushed down to coupling $g = 1.2$ when using a BDT cut of around 0.5. This has to be checked and compared with the full formalism of profile likelihood.

6.3 Application of the Profile Likelihood Method

A more realistic significance calculation is possible when applying the full formalism of profile likelihood as described in Section 6.1.2.

As in Section 6.2, a cut on the angle between the leptons with the largest transverse momentum of $\Delta\phi > 0.8\pi$ is applied and an expected integrated luminosity of $100\,\text{fb}^{-1}$

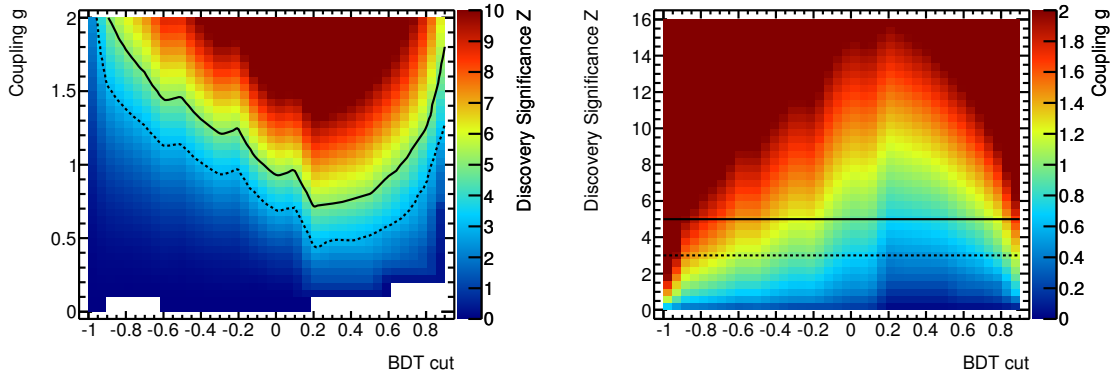


Figure 6.2: Discovery significance Z calculated with the profile likelihood method as a function of the coupling g and a cut on the BDT output (left plot). The minimal coupling g discoverable for a given discovery significance Z and a cut on the BDT output can be read off the right plot. An integrated luminosity of $100\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and a cut on the angle between the leptons with the largest transverse momentum of $\Delta\varphi > 0.8\pi$ is applied. The equivalent integrated Monte Carlo luminosity of the background samples is scaled to $100\,\mathrm{fb}^{-1}$. The plots show the φ resonance with pile-up.

at $\sqrt{s} = 14$ TeV is assumed. The equivalent integrated Monte Carlo luminosity of the background samples L_i^{MC} (see Equation 6.7) is scaled to $100\,\mathrm{fb}^{-1}$ for each background separately. This assumes that when the experiment will have taken the amount of data supposed in this analysis, a lot more Monte Carlo events than today will be available.

The results for the φ resonance with pile-up can be seen in Figure 6.2. Training of the BDT is performed as explained in Chapter 5.3.1. At a cut on the BDT output of $r \geq 0.2$ for example, discovery is expected for couplings $g \geq 0.7$. Results of all other resonances and samples without pile-up can be found in Appendix C.

When comparing to the approximate result in Figure 6.1, the overall shape is approximately retained. The same training is used for both calculations. Values of the couplings expected to be discovered are shifted to much better discovery significance values in the interesting regions of the BDT output. The reason for this advancement can be seen in the assumed higher background Monte Carlo luminosities and the removal of the nuisance background parameters with the help of the profile likelihood method.

The lowest values of the couplings g for discovery with a given significance Z at the optimal cuts on the BDT output can be read off Figure 6.3 for various resonances.

An optimal cut on the BDT output is selected by minimizing the coupling g for a discovery significance of 5σ in the results presented in Figure 6.2 and the equivalent figures of the other resonances in Appendix C. This approach is chosen because at the minimum of a contour the widest range of couplings can be discovered for a given discovery significance. Results of the optimal cuts on the BDT outputs can be seen in Table 6.1. Because a binning of 0.1 is used for the BDT output, these values are not rounded.

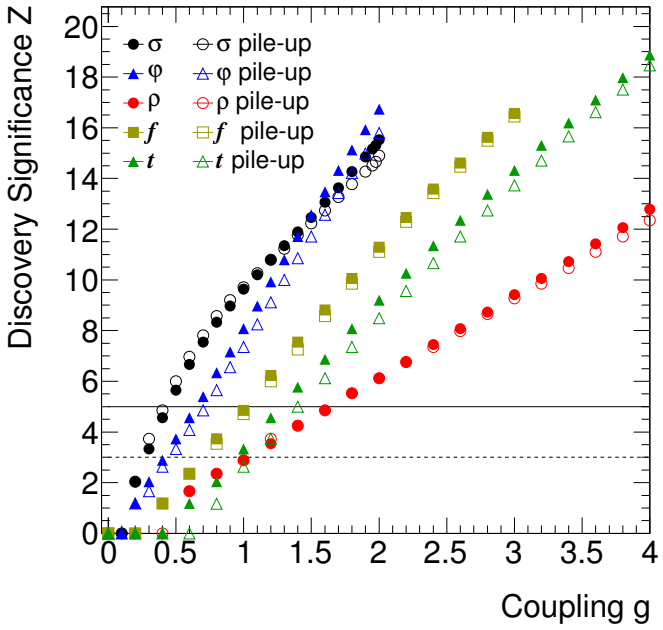


Figure 6.3: Discovery significance Z as a function of coupling g for various resonances. Shown is the distribution for a mass of $m = 850$ GeV and an expected integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV. The equivalent integrated background Monte Carlo luminosity is scaled to $100\,\text{fb}^{-1}$.

Resonance	σ		φ		ρ		f		t	
Pile-up	no	yes	no	yes	no	yes	no	yes	no	yes
Optimal BDT cut	0.1	0.3	0.2	0.2	0.2	0.2	0.4	0.3	0.4	0.3
Coupling g_{min}	0.44	0.41	0.65	0.71	1.64	1.64	1.02	1.04	1.27	1.40
Pile-up effect	-6.3%		10.1%		0.0%		2.3%		10.1%	

Table 6.1: Optimal cuts on the BDT output, minimal coupling g discoverable with 5σ per resonance ($m = 850$ GeV) and the effect of pile-up (see Equation 6.19). The method of obtaining these values is described in the text. All couplings between the minimal value and $g \approx 2.5$ (see text) are expected to be discovered at ATLAS at 14 TeV with $100\,\text{fb}^{-1}$ of LHC data.

The main results are the couplings g for each resonance that can be discovered with 5σ discovery significance Z. They are obtained by interpolating the results of Figure 6.3 to the desired discovery significance by the TSpline method implemented in ROOT. These results must be seen with the constraint that only couplings up to $g \approx 2.5$ are reasonable [1]. A Higgs boson (σ resonance with $g = 1.0$) is contained in the discoverable range of couplings. No systematic uncertainties are taken into account. Results with systematic uncertainties can be found in Section 6.7. 6.4

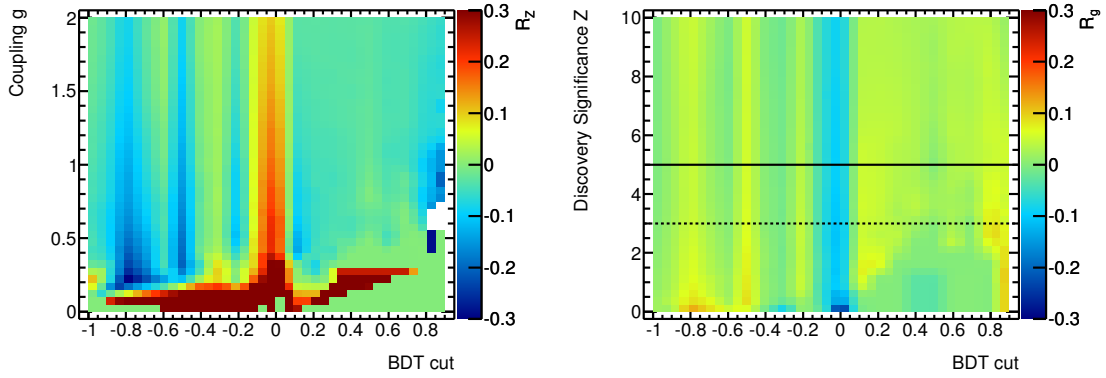


Figure 6.4: Comparison between analyses with and without pile-up. Shown is the quotient (pile-up - no pile-up)/(pile-up + no pile-up) for the discovery significance Z calculated with the profile likelihood method (left, see Equation 6.17) and the coupling g (right, see Equation 6.18). The results are obtained for the φ resonance and the BDT is trained with samples including pile-up. An integrated luminosity of $100\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and a cut on the angle between the leptons with the largest transverse momentum of $\Delta\phi > 0.8\pi$ is applied. The equivalent integrated Monte Carlo luminosity of the background samples is scaled to $100\,\mathrm{fb}^{-1}$.

6.4 Effects of Pile-up

Samples with and without pile-up are compared by taking the quotient of the difference and the sum of the results with and without pile-up as shown in the equations below.

Different responses of the BDT for the analyses with and without pile-up are expected because the training samples are different. The BDT output variable r should correctly be separated in $r_{\text{pile-up}}$ for the training with pile-up samples and $r_{\text{no pile-up}}$ for the training with samples that do not include pile-up. To avoid this, the results for an analysis with and without pile-up are compared here by a training with either pile-up samples or samples without pile-up. In the following, the training with the samples of the φ resonances that include pile-up is selected. Similar results are obtained by selecting the training without pile-up and with other resonances.

To ensure a global normalization between -1 and $+1$, the following quotient is calculated as a measure of comparison between results with and without pile-up:

$$R_Z = \frac{Z(\text{pile-up}) - Z(\text{no pile-up})}{Z(\text{pile-up}) + Z(\text{no pile-up})}. \tag{6.17}$$

The application of this formula to obtain the left result of Figure 6.4 provides a very good agreement between the analysis with and without pile-up. Samples without pile-up show a slightly better discovery significance Z in the region of the optimal cut on the BDT output of BDT cut ≥ 0.2. The difference is relatively conform for different couplings and fixed cuts on the BDT output. A worse agreement at lower couplings is supposed to be caused by statistical limitations because of the subtraction of the signal sample with coupling $g = 0$ (see signal definition in Chapter 4.3.1). At BDT cuts around zero, a conspicuous structure appears. This effect is found to be induced

by the top pair sample $t\bar{t}$. If this sample is left out for the evaluation of the BDT ($T\bar{T}$ remains included in the training[2]), the structure disappears. The exact reason has not been determined, but the relatively low statistics of the top pair sample in combination with the disappearance of top pair events at BDT cut values of larger than zero (see Figure 5.6) is assumed to be the main reason of this effect.

A relation between the analysis with and without pile-up can also be found when comparing the minimal discoverable couplings for both cases. In Figure 6.4 (right), the quotient

$$R_g = \frac{g_{\min}(\text{pile-up}) - g_{\min}(\text{no pile-up})}{g_{\min}(\text{pile-up}) + g_{\min}(\text{no pile-up})} \tag{6.18}$$

has been calculated. A very good agreement between the couplings of the analysis with and without pile-up at soft and tight cuts on the BDT output is observed. The reason for the structure at BDT cuts around zero should be equivalent to the assumption above.

With the results for the minimal coupling g discoverable for a significance of 5σ (see Table 6.1), the final effect of pile-up is calculated with the following equation:

$$\text{Pile-up effect} = \frac{g_{\min}(\text{pile-up})}{g_{\min}(\text{no pile-up})} - 1. \tag{6.19}$$

The largest effect of pile-up of about 10% is observed for the φ and t resonances. This value can be interpreted as the general result for the influence of pile-up. When including systematic uncertainties (see Chapter 6.7), the effect of pile-up is found to be larger.

6.5 Luminosity Studies

In the previous chapters an expected integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV has been assumed. But in the first years ATLAS will presumably collect only a few inverse femtobarn of data at $\sqrt{s} = 7$ TeV.

Simulation of a different center-of-mass energy is very extensive and is not performed here. This is also not a main parameter to change because $\sqrt{s} = 14$ TeV are planned to be reached in the next years at the LHC and most of the data necessary for this analysis will be collected at this larger center-of-mass energy.

When assuming a different integrated luminosity, the training is not repeated with the adjusted luminosity values. Because luminosity enters the global weight for signal and background equally (see Equation 6.7), all samples are scaled in the same way. Therefore, the behavior of the BDT and the selected cut on the angle between the leptons with the largest transverse momentum should not be affected. A re-training for $10\,\text{fb}^{-1}$ and $1\,\text{fb}^{-1}$ and a comparison to the scaled yields trained for $100\,\text{fb}^{-1}$ showed a perfect agreement and the same resulting cuts on the BDT output.

The profile likelihood method is repeated for all considered integrated luminosities. The expected equivalent background Monte Carlo luminosity L_i^{MC} is scaled according

[2]For the definitions of the different top pair samples $t\bar{t}$ and $T\bar{T}$ see Section 4.3.2.

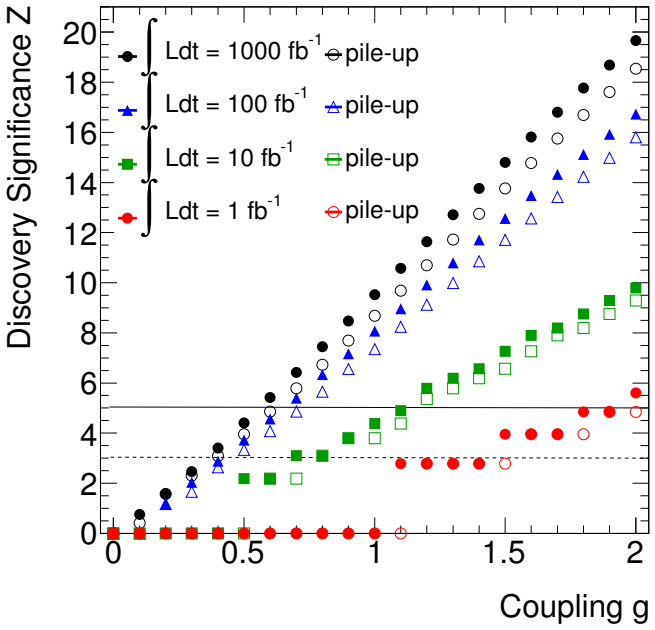


Figure 6.5: Discovery significance Z versus coupling g for different assumed integrated luminosity scenarios at $\sqrt{s} = 14$ TeV. The cut on the BDT output is optimized separately for each case. The results are shown here for the φ resonance with a mass of $m = 850$ GeV.

to the integrated luminosity in the experiment to get comparable results and the cut on the BDT output is optimized for each assumed integrated luminosity.

Results for the φ resonance and four benchmark luminosities are shown in Figure 6.5. For each luminosity case the optimal cut on the BDT output is 0.2. Only at $1\,\text{fb}^{-1}$ an optimal cut of 0.3 is found.

In Figure 6.6, a scanning of a complete range of integrated luminosities between $10\,\text{fb}^{-1}$ and $1\,000\,\text{fb}^{-1}$ is shown. The optimal cut on the BDT output is optimized for all considered luminosity bins.

For an integrated luminosity of $1\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV a discovery at 5σ is only possible for very high couplings close to the reasonable limit around $g = 2.5$. Because an integrated luminosity of this order is expected to be collected by ATLAS at the end of 2011, *but* only at $\sqrt{s} = 7$ TeV, a discovery of the φ resonance with a mass of $m = 850$ GeV is not possible with this first amount of data. The same result is obtained for other resonances considered in this work with a mass of $m = 850$ GeV.

An integrated luminosity of at least $10\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV could lead to a discovery at the 5σ level above a coupling of $g = 1.2$. A huge gain in discovery significance is achieved at integrated luminosities reaching $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV. Higher values of the collected integrated luminosity will lead only to small enhancements of the sensitivity. Systematic uncertainties are expected (see Chapter 6.7) to result in an enlargement of the minimal coupling of about 50% (see Section 6.7).

A general very good agreement between the results with and without pile-up is observed. However, this approach and the general study for an expected integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV must be seen as a very rough estimation. The

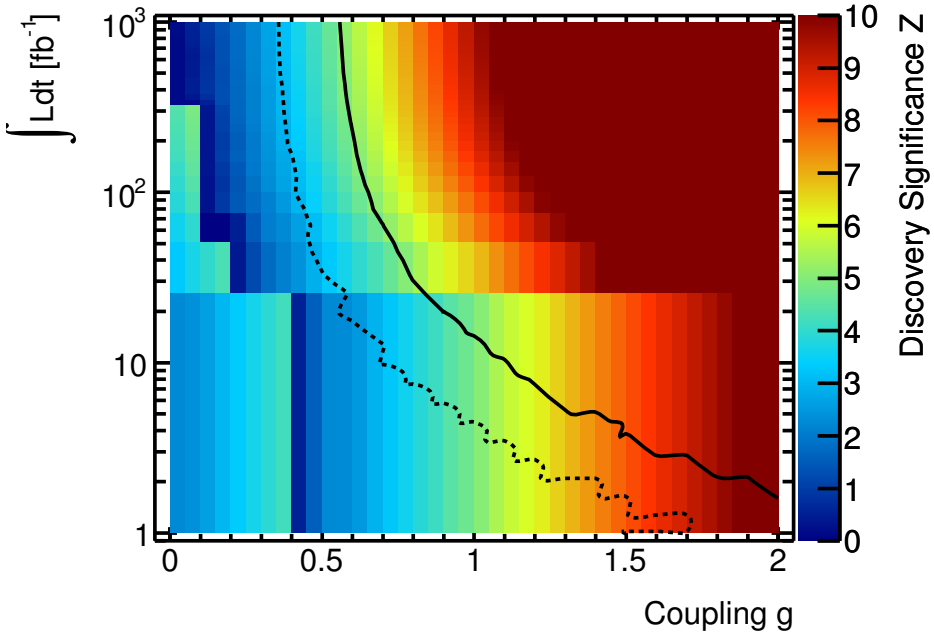


Figure 6.6: Discovery significance Z versus coupling g and integrated luminosity $\int Ldt$ at $\sqrt{s} = 14$ TeV for the φ resonance with a mass of $m = 850$ GeV and included pile-up.

ATLAS software will experience some serious adjustments due to better understanding of the detector until even $1\,\text{fb}^{-1}$ of data is reached and the final extent and the global behavior of pile-up is not known for higher luminosity cases. When increasing the luminosity, the amount of pile-up will change rapidly (see Equation 3.6). The relatively low rate of additional pile-up events used in this study will be exceeded by higher rates with increasing luminosity. More pile-up leads to more tracks and jets to be reconstructed and to be identified as not part of the main physics event. The challenge is to suppress these effects to preserve the sensitivity estimated in this study.

6.6 Training Crosschecks and Statistical Error Estimation

To test the behavior and the stability of the training of the BDT against statistical fluctuations, several crosschecks are performed. It will also be possible to estimate statistical errors of the training.

As described in Section 4.3.1 and is visualized in Figure 5.1, only 6 000 fully-simulated events (red and brown in Figure 5.1) are used for training *and* evaluation of the BDT in the last chapters. The justification for this choice is also given here.

Because additional 24 000 fully-simulated signal events with pile-up[3] (yellow in Figure 5.1) are available, pile-up samples are used for these crosschecks. Also a result of the training with additional 24 000 fast-simulated signal events (green in Figure 5.1) without pile-up is obtained.

[3]For the σ resonance, only 6 000 fully-simulated events are available as can be seen in Table 4.1.

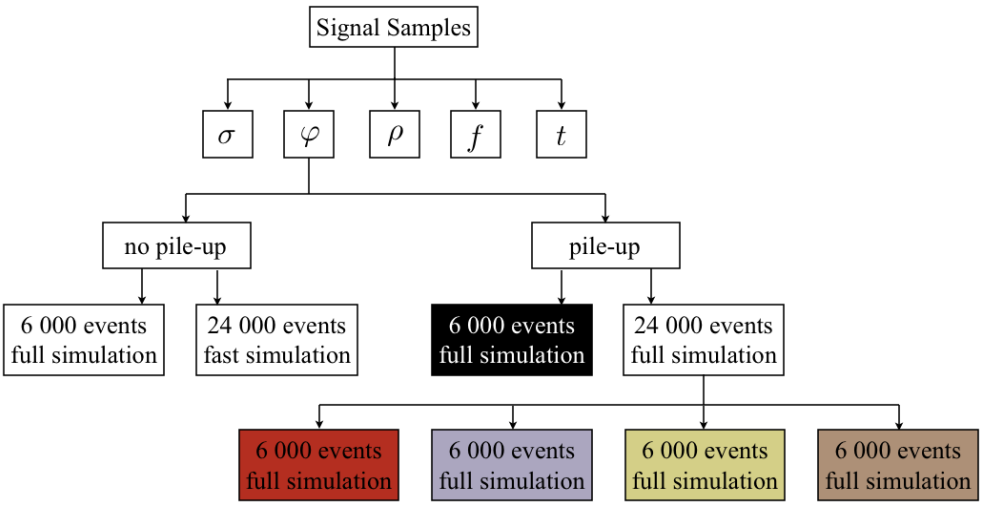


Figure 6.7: Visualization of the signal samples with pile-up used in this section for trainings with independent samples of the same resonance. In this section the φ resonance has been selected.

Training with independent samples

A training with independent samples (see Figure 6.7) under identical conditions is performed here. To get a meaningful comparison, the training results are applied to *one fixed sample* in the evaluation. The same sample used for training and evaluation in the last chapters (black in Figure 6.7) is selected as this one sample and therefore called *original sample* or *sample 0* in the following. For each evaluation, the optimal cut on the BDT output is identified independently.

Training		Evaluation
sample 0 — 6 000 events — full simulation	→	6 000 events — full simulation
sample 1 — 6 000 events — full simulation	→	6 000 events — full simulation
sample 2 — 6 000 events — full simulation	→	6 000 events — full simulation
sample 3 — 6 000 events — full simulation	→	6 000 events — full simulation
sample 4 — 6 000 events — full simulation	→	6 000 events — full simulation

The results for the φ resonance ($m = 850$ GeV) with pile-up given in Figure 6.8 (left) show no significant dependence on the selected training sample. The slightly better performance of the original sample (black in Figure 6.7 and Figure 6.8 (left)) can be an effect of overtraining and statistical dependence of training and evaluation events because this sample has also been used for evaluation of *all* trainings.

Arithmetic mean and spread for the minimal coupling at a discovery significance of 5σ is calculated to:

$$g_{\text{min}} = 0.74 \pm 0.02. \tag{6.20}$$

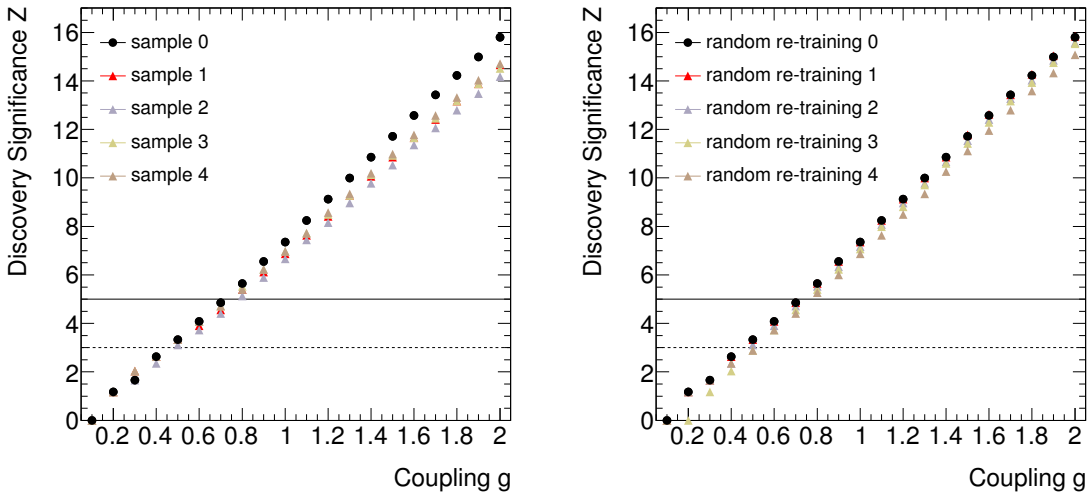


Figure 6.8: Retraining of the BDT under identical conditions with five independent, fully-simulated signal samples with 6 000 events each (left) and five different seeds of the random number generator to pick the training sub-samples within TMVA (right). All results of the trainings are applied to one fixed sample (definition in the text). The optimal cut on the BDT output is selected for each evaluation independently. The results are shown for the φ resonance with a mass of $m = 850$ GeV and with pile-up. An integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed.

This corresponds to a relative spread of 2.8%. For the calculation, the training with the original sample is dropped because of expected statistical dependence of training and evaluation events. The result for the training with the original sample is $g_{\text{min}} = 0.71$. The bias caused by using the same sample for training and evaluation as done in the last chapters is therefore found to be small.

Random re-training

Within TMVA, it is selected to internally split the samples in training and testing sub-samples randomly (see Chapter 5.2.2). The results of a re-training with four different seeds for the random number generator is shown in Figure 6.8 (right) for a training with the original sample (black in Figure 6.7). All results of the trainings are applied to the original sample.

	Training		Evaluation
5×	6 000 events — full simulation	→	6 000 events — full simulation

Again, a good stability and a small error of the training is observed. The arithmetic mean and spread for the minimal coupling at a discovery significance of 5σ is calculated to:

$$g_{\text{min}} = 0.74 \pm 0.02 \tag{6.21}$$

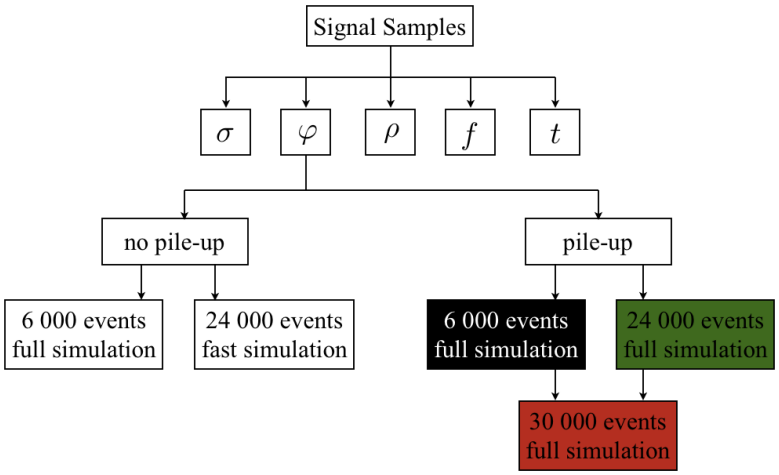


Figure 6.9: Visualization of the signal samples with pile-up used in this section to compare the trainings with different signal sample sizes for the same resonance. In this section the φ resonance has been selected.

with a relative spread of 2.7%. This value is very similar to the result of the last chapter obtained by choosing independent training samples and its relative spread can be interpreted as the statistical error of the training of the BDT.

Training with samples of different sizes

A dependence of the discovery significance estimated with a BDT and the profile likelihood method on the number of events of the training input samples is also possible. To analyze this effect, the BDT is trained with two additional samples of the φ resonance with pile-up as visualized in Figure 6.9. One sample with 24 000 signal events (green in Figures 6.9 and 6.10) and one sample with all 30 000 signal events (red) are chosen in addition to the original sample (black). All trainings are evaluated with the original sample (black). The events of the original sample (black) are included in the largest sample (red) but are absent in the smaller one (green) as can be seen in Figure 6.9.

Training		Evaluation
6 000 events — full simulation	→	6 000 events — full simulation
24 000 events — full simulation	→	6 000 events — full simulation
30 000 events — full simulation	→	6 000 events — full simulation

The discovery significance Z as a function of the coupling g and the cut on the BDT output for the evaluation of a training with the largest sample (red) is shown in the right plot of Figure 6.10. A similar result is obtained for the training with the medium sized sample (green). For the multiple re-trainings of the last chapters, no modification of the cut on the BDT output (0.2) was necessary. Here, the size of the training sample affects the global shape of the output and the optimal BDT cut. Its value was found to be 0.7 for the two larger samples (green and red).

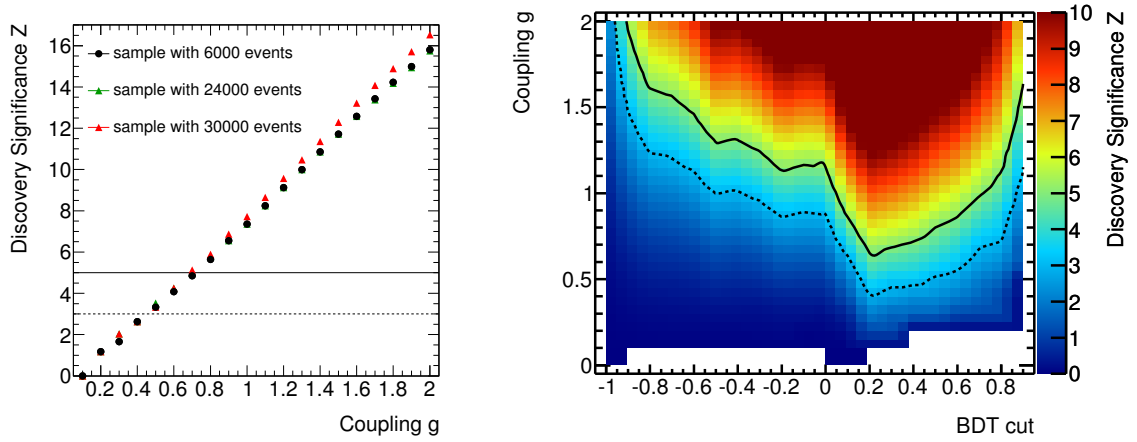


Figure 6.10: Retraining of a BDT under identical conditions but with different signal sample sizes. The results of the trainings are applied to one fixed sample (the original sample, defined in the text). The results of the statistical analysis are shown in terms of the discovery significance Z as a function of the coupling g for the optimal cut on the BDT output for three different signal sample sizes (left plot). The right plot shows the discovery significance Z as a function of the coupling g and the cut on the BDT output for a BDT training with 30 000 fully-simulated signal events. For all trainings, the φ resonance ($m = 850$ GeV) with pile-up is chosen and an integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed. For all evaluations, the optimal cut on the BDT output is selected independently.

Nevertheless, a good agreement in terms of discovery significance between all three training results is observed (see left plot of Figure 6.10). When comparing to the re-trainings in the last sections, the larger size of the training sample is found to have the same effect as the simultaneous usage of the same sample for training *and* evaluation. This can be seen at the largest sample (red) with 30 000 training events, where both situations are combined due to the original sample (black) being part of this sample as seen in Figure 6.9.

Comparison between full and fast simulation

As a last test, the BDT training with the GEANT fully-simulated signal sample without pile-up (black in Figure 6.11) is compared to a training with the ATLFAST-II fast-simulated signal sample without pile-up (red). As signal, the φ resonance without pile-up is selected. Both results of the trainings are applied to the fully-simulated signal sample without pile-up (black).

Training		Evaluation
6 000 events — full simulation	→	6 000 events — full simulation
24 000 events — fast simulation	→	6 000 events — full simulation

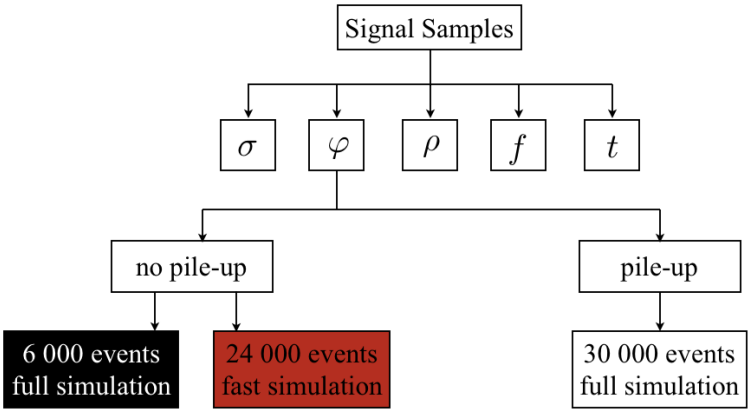


Figure 6.11: Visualization of the signal samples without pile-up used in this section to compare the training with fast-simulated events to the training with 24 000 fully-simulated events of the same resonance. In this section the φ resonance has been selected and all results of the trainings are applied to the full-simulated signal sample with 6 000 events and without pile-up.

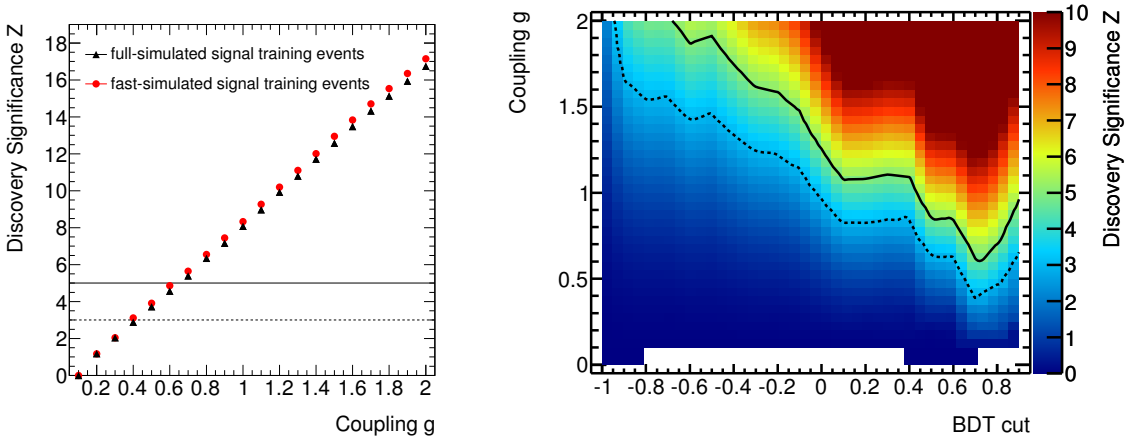


Figure 6.12: Training of a BDT with 24 000 fast-simulated signal events without pile-up (red) compared to a training with 6 000 fully-simulated signal events without pile-up (black). For both cases, the results of the trainings are applied to the fully-simulated sample (black) without pile-up. Identical variables and options for the trainings are used. The optimal cut on the BDT output is selected for each evaluation independently. The right plot shows the discovery significance Z as a function of the coupling g and the cut on the BDT output for a training with the fast-simulated signal events (red). The φ resonance ($m = 850$ GeV) without pile-up is selected as signal and an integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed.

The training output in terms of a BDT cut value r obtained by an with the fast-simulated signal events (red) as shown in the right plot of Figure 6.12 differs from the training output obtained with fully-simulated training events (black). Equivalent to the trainings with the larger samples for the pile-up crosscheck analysis (see Figure 6.9), the optimal cut on the BDT output is shifted to higher values. A selection

of the optimal cut resulted in 0.7. Results for the statistical evaluation of the BDT training outputs are compared in Figure 6.12 (left). A good agreement in terms of discovery significance is observed.

Here, a combination of the effects of the larger number of events in a sample (training with fast-simulated (red) signal sample) and the re-using of the same events for training and evaluation (training and evaluation with fully-simulated (black) signal sample) leads to the better results in terms of discovery significance compared to the last sections. These effects are found to be approximately equal (see Section 6.6). Therefore, the slightly different results between fully-simulated and fast-simulated training events are expected to be caused by differences between fast and full simulation or statistical fluctuations.

6.7 Systematic Uncertainties

The reconstruction performance of the ATLAS detector will be affected by statistical and systematic uncertainties. While statistical effects were analyzed in the previous chapters, systematic uncertainties have not been taken into account up to now.

When considering systematic uncertainties, the likelihood of Equation 6.4 has to be extended. One way [1] is the modification of the signal and background expectation values s and b_i:

$$s(g) \to s'(g) = s(g) - \sum_j c_j e_j \quad \text{and} \quad b_i \to b'_i = b_i - \sum_j d_{ij} e_j. \tag{6.22}$$

Here, the transformation is expected to be linear with the linear transfer constants c_j and d_{ij}. Deviations e_j from the expected values for background i and systematic variable numbered j are expected to be Gaussian distributed around zero with the standard deviation σ_j. This results in a complete likelihood including systematic uncertainties:

$$L(g, b_i, e_j) = L_n^{(1)}(g, b_i, e_j) \times L_{b_i^{\mathrm{MC}}}^{(2)}(b_i, e_j) \times L_{s^{\mathrm{MC}}}^{(3)}(g, e_j) \times L^{(4)}(e_j), \tag{6.23}$$

with:

$$\begin{aligned}
L_n^{(1)}(g, b_i, e_j) =& \frac{(s'(g) + \sum_i b'_i)^n}{n!} \exp\left[-(s'(g) + \sum_i b'_i)\right] \\
L_{b_i^{\mathrm{MC}}}^{(2)}(b_i, e_j) =& \prod_i \frac{(\tau_i b'_i)^{b_i^{\mathrm{MC}}}}{b_i^{\mathrm{MC}}!} \exp\left[-\tau_i b'_i\right] \\
L_{s^{\mathrm{MC}}}^{(3)}(g, e_j) =& \frac{1}{\sqrt{2\pi}\sigma_s} \exp\left[-\frac{\left(s^{\mathrm{MC}}(g) - s'(g)\right)^2}{2\sigma_s^2}\right] \\
L^{(4)}(e_j) =& \prod_j \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{e_j^2}{2\sigma_j^2}\right].
\end{aligned}$$

The likelihood is therefore extended by additional nuisance parameters e_j. An Asimov dataset equivalent to Equation 6.10 is used:

$$n = s^{\mathrm{MC}}(g) + \sum_i \frac{b_i^{\mathrm{MC}}}{\tau_i}. \tag{6.24}$$

Nuisance parameters are removed by a modified profile likelihood ratio similar to Equation 6.11:

$$\lambda(g) = \frac{L(g, \hat{\hat{b}}_i, \hat{\hat{e}}_j)}{L(\hat{g}, \hat{b}_i, \hat{e}_j)}. \tag{6.25}$$

Finally, the discovery significance can be calculated with Equation 6.15,

$$Z(g) = \sqrt{-2 \ln \lambda(g = 0)}, \tag{6.26}$$

where the dependence on the coupling g enters with the Asimov dataset.

The following systematic uncertainties are considered [1]:

- *Jet-energy scale*: $E'_{\text{jet}} = (1 + e_1)E_{\text{jet}}$ with $\sigma_1 = 3.5\%$ ($|\eta| \leq 3.5$) or $\sigma_1 = 7.5\%$ ($|\eta| > 3.5$)
- *Jet-energy resolution*: $E' = E + \Delta E$ with $\sigma_2 = 1$ and ΔE randomly drawn from a Gaussian with $\sigma(E) = \kappa\, e_2 \sqrt{E \times 1\ \text{GeV}}$ and $\kappa = 0.45$ ($|\eta| \leq 3.5$) or $\kappa = 0.63$ ($|\eta| > 3.5$)
- *Electron-energy scale*: $E'_{\text{e}} = (1 + e_3)E_{\text{e}}$ with $\sigma_3 = 0.5\%$
- *Electron-energy resolution*: $E'_{\text{e}} = \left(1 + \frac{\Delta E_{\text{T}}}{E_{\text{T}}}\right) E_{\text{e}}$ with $\sigma_4 = 1$ and ΔE_{T} randomly drawn from a Gaussian with $\sigma(E_{\text{T}}) = 0.0073\, e_4\, E_{\text{T}}$
- *Muon-energy scale*: $E'_{\mu} = (1 + e_5)E_{\mu}$ with $\sigma_5 = 1\%$
- *Muon-reconstruction resolution*: $E'_{\mu} = \left(1 + p_{\text{T}}\, \Delta\left(\frac{1}{p_{\text{T}}}\right)\right)^{-1} E_{\mu}$ with $\sigma_6 = 1\%$ and $\Delta\left(\frac{1}{p_{\text{T}}}\right)$ randomly drawn from a Gaussian with $\sigma(\frac{1}{p_{\text{T}}}) = e_6 \sqrt{\left(\frac{0.011}{p_{\text{T}}}\right)^2 + \left(\frac{0.00017}{\text{GeV}}\right)^2}$
- *b-tag efficiency*: $b'_i = (1 + e_7)b_i$ with $\sigma_7 = 10\%$
- *Luminosity*: $\sigma_8 = 3\%$

Same options for training and evaluation of the BDT as in Section 5.3 are applied here and all systematic uncertainties mentioned above are taken into account. An integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and the background Monte Carlo luminosity is scaled to $100\,\text{fb}^{-1}$.

The results for the couplings g in terms of discovery significance Z can be seen in Figure 6.13. A big effect of systematic uncertainties compared to the results of Figure 6.3 with no systematic uncertainties is observed. Also the difference between the results of the analyses with and without pile-up is larger. This indicates a larger influence of systematic uncertainties to pile-up samples compared to samples without pile-up.

When considering the minimal couplings g_{min} discoverable at a significance of 5σ (see Table 6.1), the effect of systematic uncertainties for samples with pile-up is up to 70% for the σ resonance. By switching on only one systematic uncertainty, the largest systematic uncertainties for the φ resonance with a mass of 850 GeV are found to be

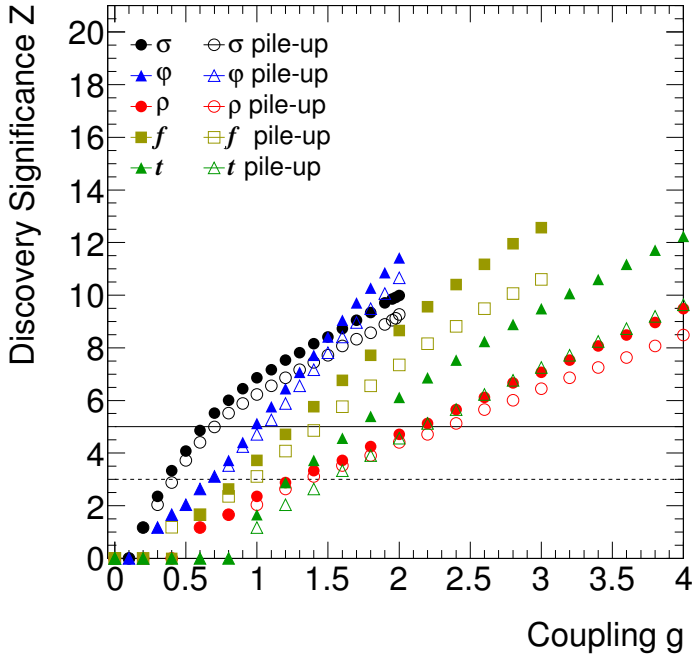


Figure 6.13: Discovery significance Z as a function of coupling g including systematic uncertainties for various resonances. Shown is the distribution for a mass of $m = 850$ GeV and an expected integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV. The equivalent integrated background Monte Carlo luminosity is scaled to $100\,\text{fb}^{-1}$.

muon-reconstruction resolution, *b-tag efficiency* and *jet-energy scale* with an effect of 13.1%, 9.8% and 7.7%.

The effect of pile-up is found to be up to 26.3% for the t resonance and about 10% for the other resonances and increases compared to the results without systematic uncertainties (see Table 6.1).

Resonance	σ	φ	ρ	f	t
Coupling $g_{\text{min}}^{\text{systematics}}$	0.70	1.05	2.34	1.43	2.15
Effect of systematics	69.9%	46.5%	42.2%	37.1%	53.7%
Effect of pile-up	12.7%	7.2%	9.4%	14.1%	26.3%

Table 6.2: Minimal couplings g discoverable with a significance of 5σ per resonance ($m = 850$ GeV) for samples with pile-up and systematic uncertainties included. The effects of systematic uncertainties and the effects of pile-up for results including systematic uncertainties are shown as percentage deviations. All couplings above the minimal value are expected to be discoverable with $100\,\text{fb}^{-1}$ of LHC data at $\sqrt{s} = 14$ TeV.

7 Conclusions and Outlook

In this thesis, a Monte Carlo simulation analysis of the vector boson scattering in the di-leptonic final state at ATLAS was presented. In particular, the effects of pile-up were investigated for the first time.

It is the main result that ATLAS has discovery potential in this channel for additional resonances of mass in the TeV-range and in the relevant coupling range up to $g \approx 2.5$ with $L = 100\,\text{fb}^{-1}$ of data at $\sqrt{s} = 14$ TeV.

The scalar resonances σ and φ with a mass of $m = 850$ GeV are detectable at the 5σ discovery significance level for the range of couplings larger than $g_{\text{min}} = 0.70$ and $g_{\text{min}} = 1.05$. The Higgs boson (σ resonance with $g = 1$) is included in this range. The tensor resonances f and t with a mass of $m = 850$ GeV are more difficult to detect. They are discoverable for couplings larger than $g_{\text{min}} = 1.43$ and $g_{\text{min}} = 2.15$ at the 5σ discovery significance level. Hardest to discover is the vector resonance ρ. Only in the small range of $2.34 < g < 2.5$ a 5σ discovery can be claimed for a mass of $m = 850$ GeV.

Pile-up is one effect amongst many. The effect of low-luminosity pile-up including systematic uncertainties is around 10%, only the t resonance channel shows a large influence of additional pile-up events (26%).

All of the results are highly dependent on systematic uncertainties. An effect compared to an analysis without systematic uncertainties of around 50% is observed. Systematic uncertainties are dominating the results compared to effects of pile-up.

Although pile-up is included in the results for the couplings quoted above, an improvement is achieved compared to a similar study without pile-up [1]. This is found to be mainly caused by a more sophisticated training of the multivariate analysis. An updated current version of TMVA is applied, negative events are fully taken into account and modified samples are used.

The approach of this work must be seen as a first basic step of including pile-up. While data taking continues, the amount of additional proton-proton interactions per bunch crossing will increase rapidly. For the collected data required within this work, a high instantaneous luminosity leading to high rates of pile-up is necessary. Therefore, the rate of low-luminosity pile-up considered here will be exceeded by much higher rates. In a future Monte Carlo study, these rates of pile-up have to be included. Also mechanisms to suppress the influence of additional pile-up events have to be applied and improved.

Also the latest version of the ATLAS software framework Athena has to be applied to be sure of using the current software for reconstruction and trigger algorithms. Furthermore, switching to 7 TeV center-of-mass energy seems reasonable when looking at the first inverse femtobarn of data collected by ATLAS.

A Shape of the Boosted Decision Tree Input Variables

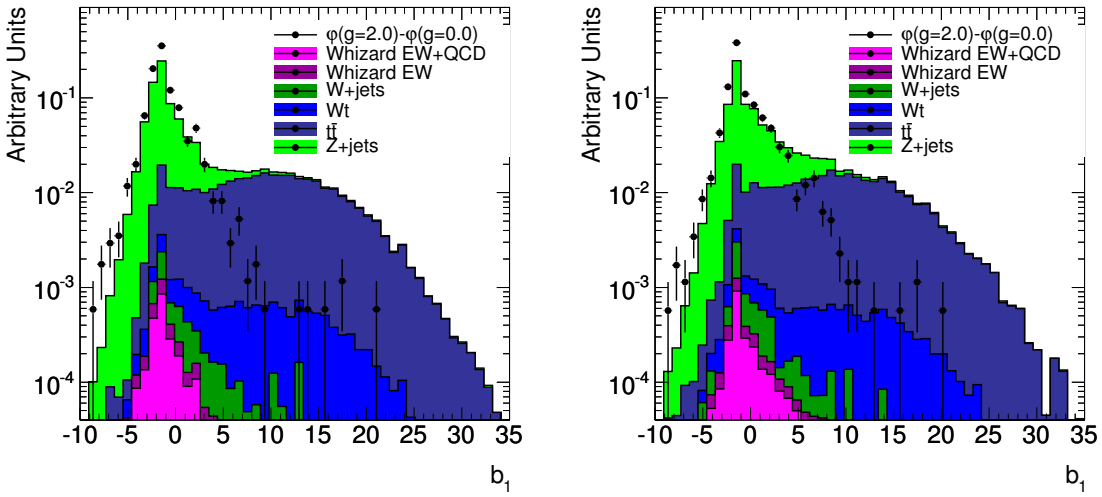


Figure A.1: Distribution of largest flavor tag weight b_1 for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

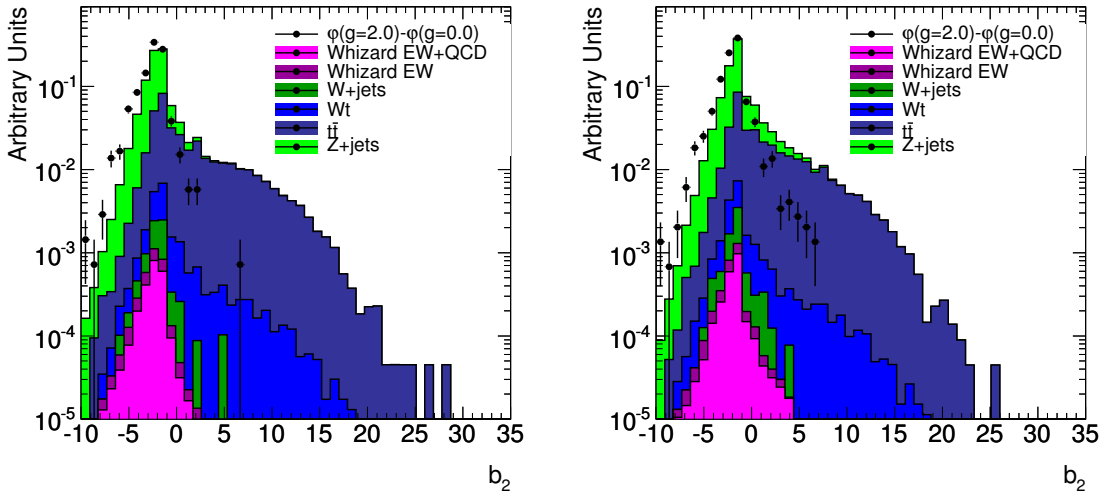


Figure A.2: Distribution of second to largest flavor tag weight b_2 for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

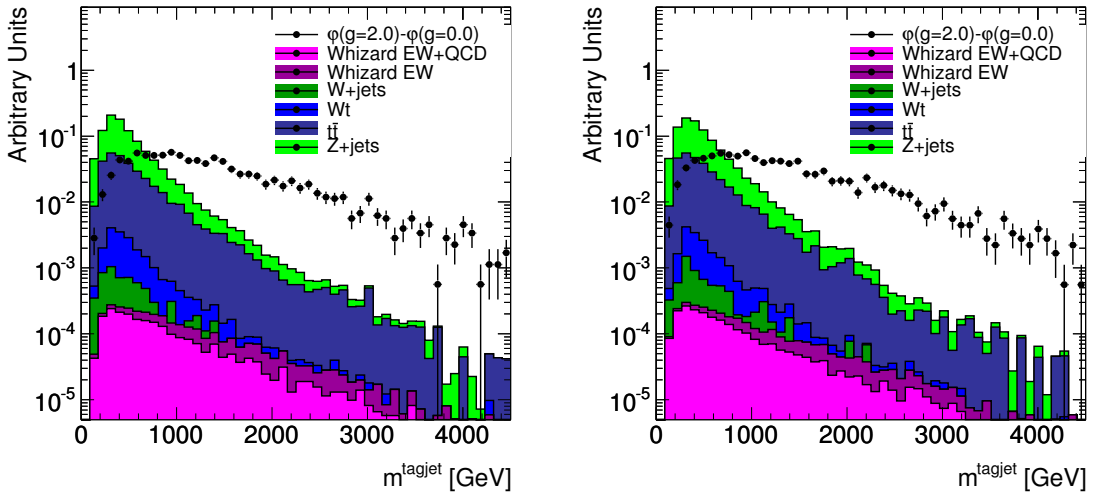


Figure A.3: Distribution of invariant mass of tagging jets m^{tagjet} for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

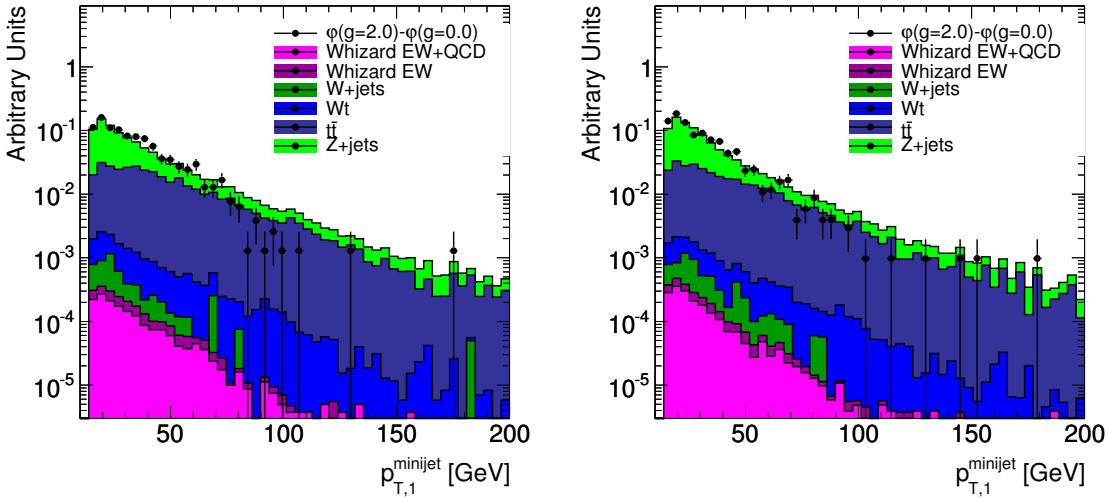


Figure A.4: Distribution of largest minijet transverse momentum $p_{\mathrm{T},1}^{\mathrm{minijet}}$ for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

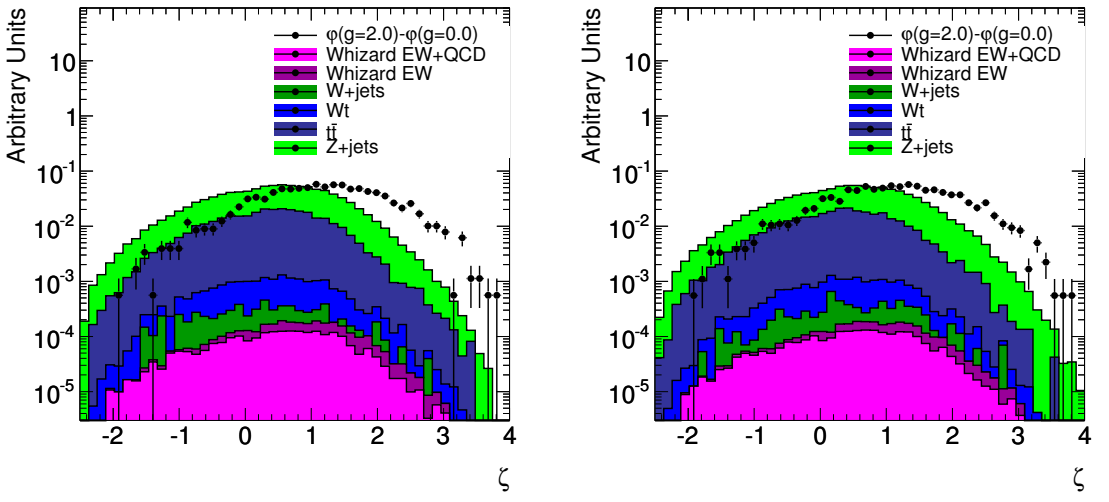


Figure A.5: Distribution of the lepton centrality ζ for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

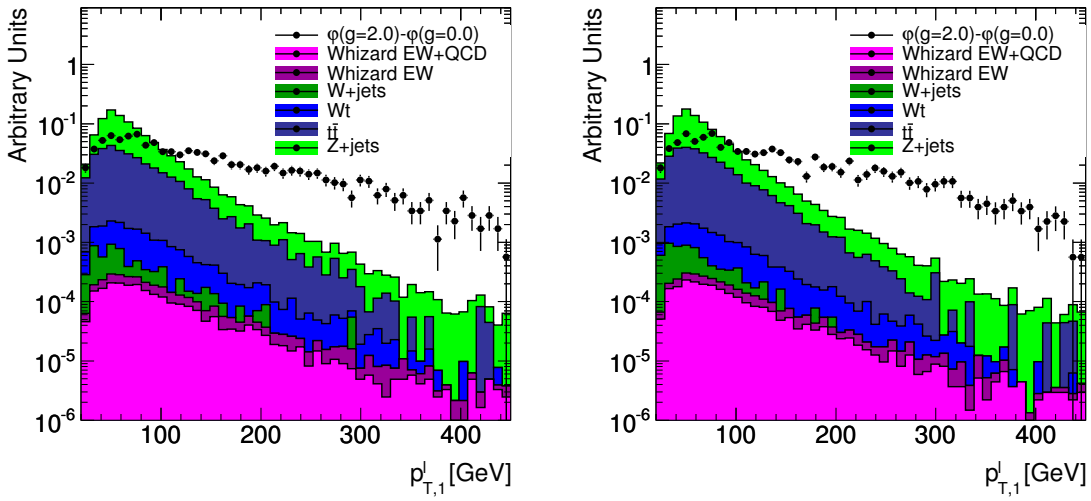


Figure A.6: Distribution of the largest lepton transverse momentum $p_{\mathrm{T},1}^{\mathrm{lepton}}$ for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

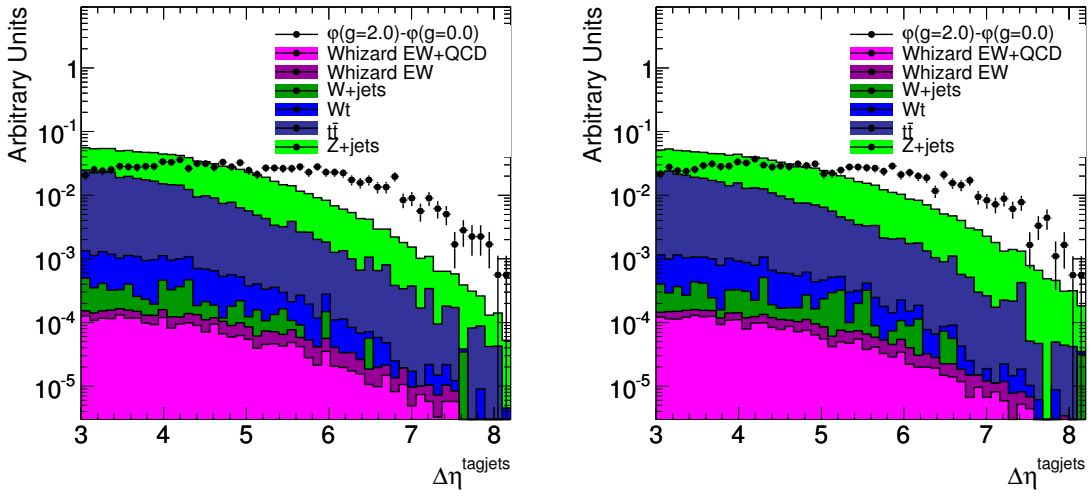


Figure A.7: Distribution of the pseudorapidity separation of the tagging jets $\Delta\eta^{\mathrm{tagjets}}$ for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

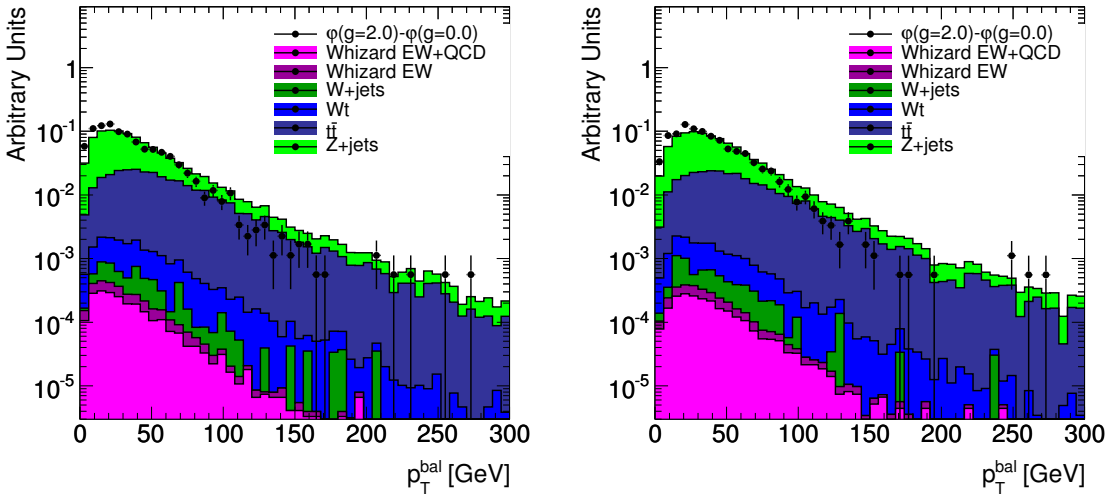


Figure A.8: Distribution of the transverse momentum balance $p_{\mathrm{T}}^{\mathrm{bal}}$ for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

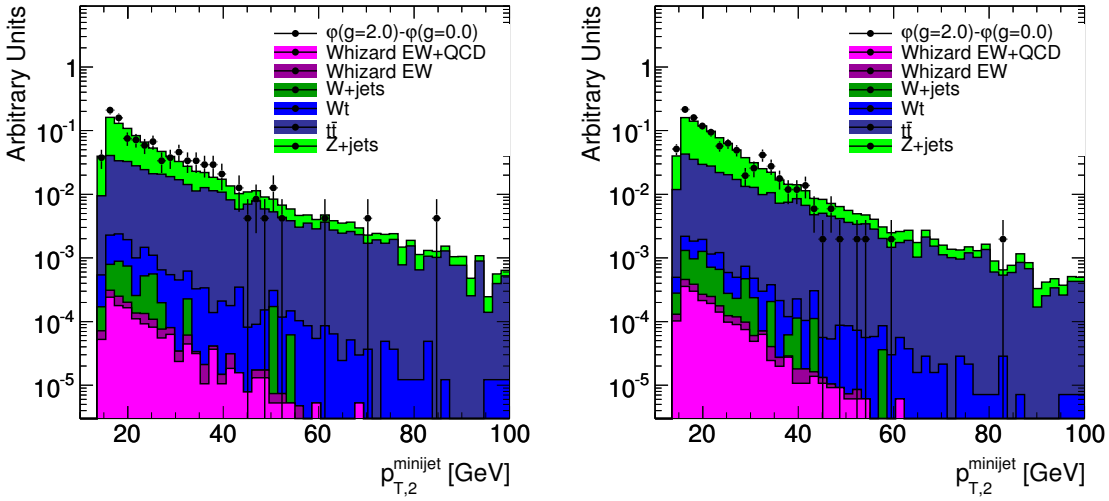


Figure A.9: Distribution of the second to largest minijet transverse momentum $p_{\mathrm{T},2}^{\mathrm{minijet}}$ for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

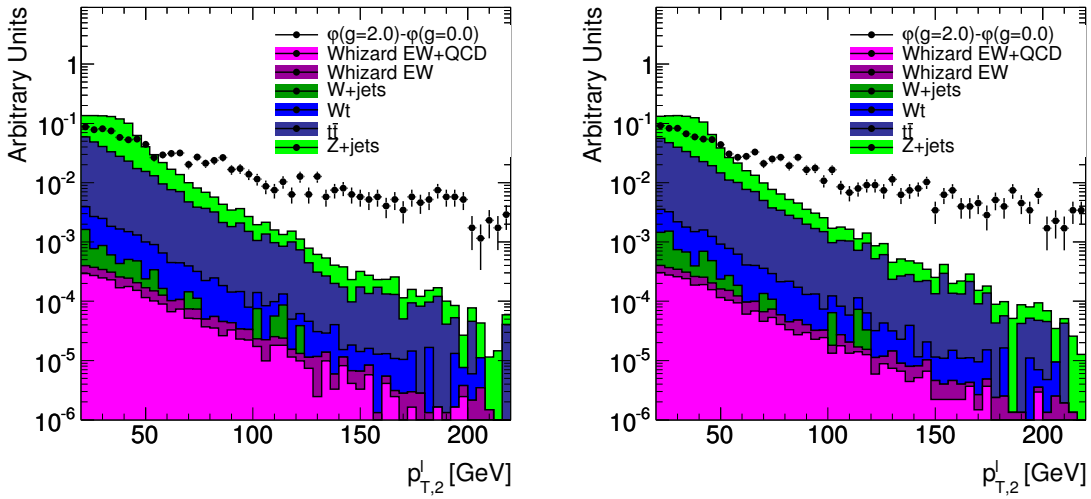
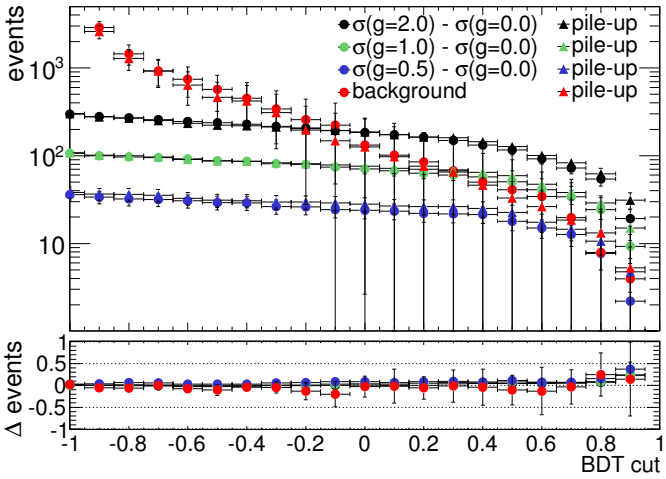
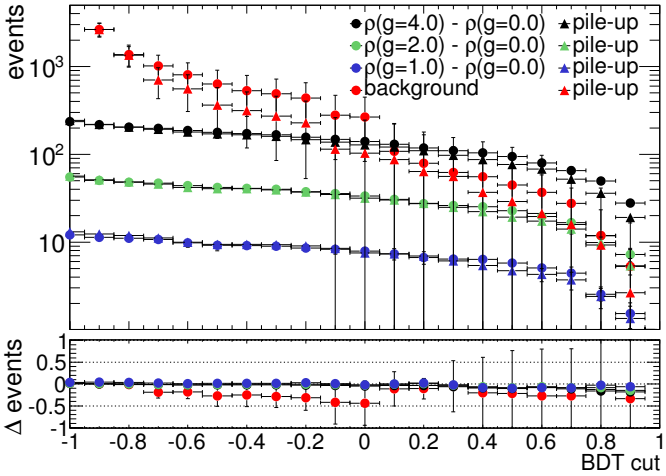


Figure A.10: Distribution of the second to largest lepton transverse momentum $p_{\mathrm{T},2}^{\mathrm{lepton}}$ for input variables without (left) and with (right) pile-up for the φ resonance with $m = 850$ GeV and $g = 2$. Signal and background are independent normalized to unity.

B Boosted Decision Tree Cutflow for Different Resonances

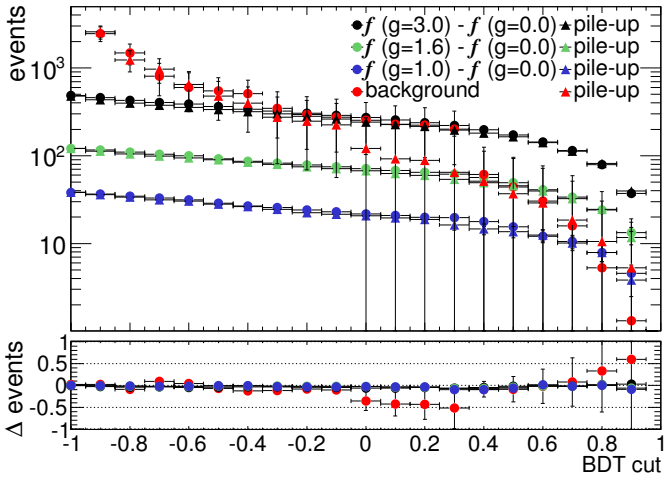


(a) σ resonance with $m = 850$ GeV

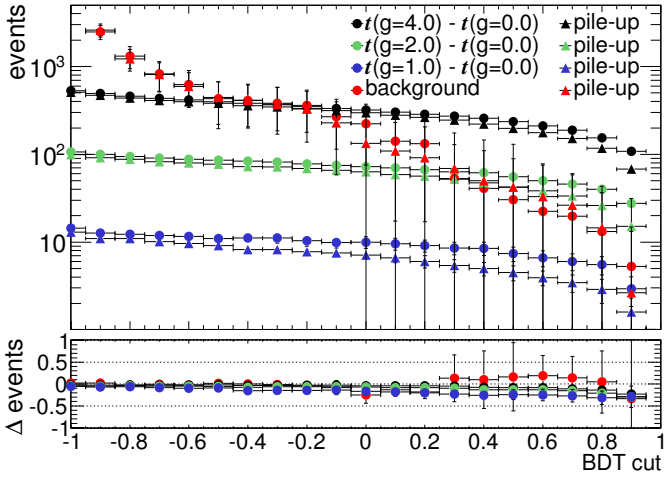


(b) ρ resonance with $m = 850$ GeV

Figure B.1: Number of events left after a cut on the BDT output for background and different resonances with different couplings as signal. An integrated luminosity of $100\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and the integrated background Monte Carlo luminosity is scaled to $100\,\mathrm{fb}^{-1}$. Pile-up samples are marked with triangles and samples without pile-up are marked with circles. A comparison between results with pile-up (pu) and without pile-up ($\overline{\mathrm{pu}}$) is given in the lower part of the plots: $\Delta\mathrm{events} \equiv (\mathrm{pu} - \overline{\mathrm{pu}})/(\mathrm{pu} + \overline{\mathrm{pu}})$.



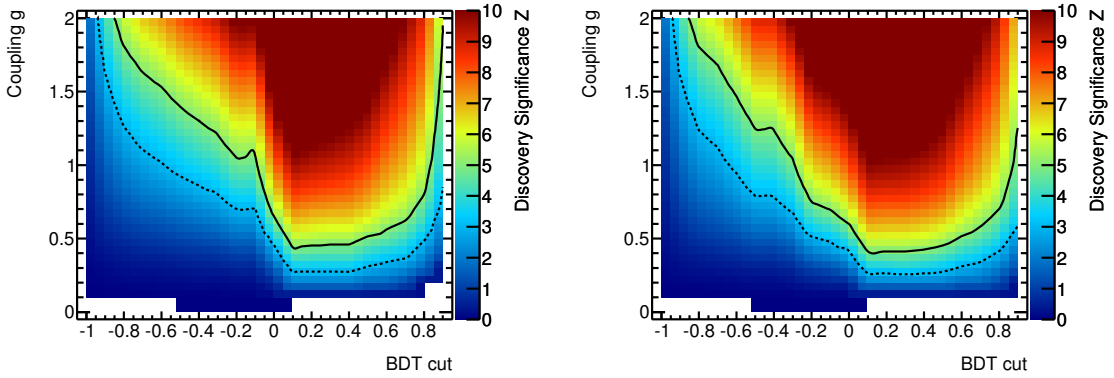
(a) f resonance with $m = 850$ GeV



(b) t resonance with $m = 850$ GeV

Figure B.2: Number of events left after a cut on the BDT output for background and different resonances with different couplings as signal. An integrated luminosity of $100\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and the integrated background Monte Carlo luminosity is scaled to $100\,\mathrm{fb}^{-1}$. Pile-up samples are marked with triangles and samples without pile-up are marked with circles. A comparison between results with pile-up (pu) and without pile-up ($\overline{\mathrm{pu}}$) is given in the lower part of the plots: $\Delta\text{events} \equiv (\mathrm{pu} - \overline{\mathrm{pu}})/(\mathrm{pu} + \overline{\mathrm{pu}})$.

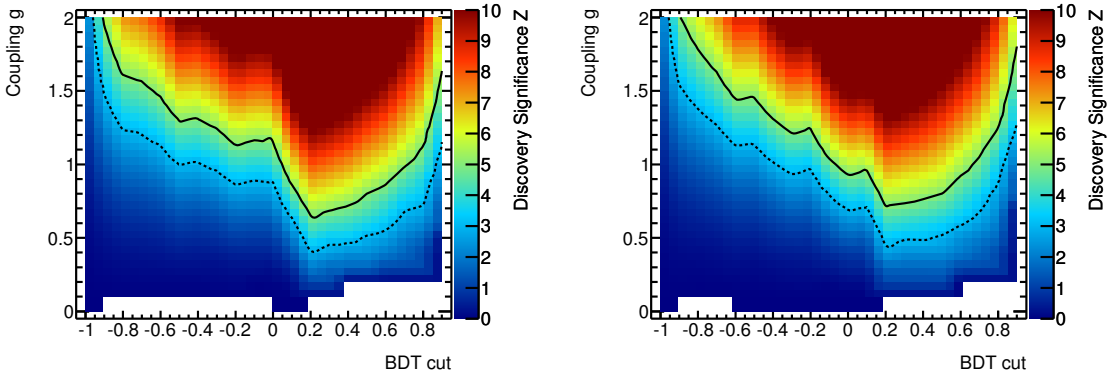
C Discovery Significance as a Function of Boosted Decision Tree Cut and Coupling for all Resonances



(a) σ resonance without pile-up.

(b) σ resonance with pile-up.

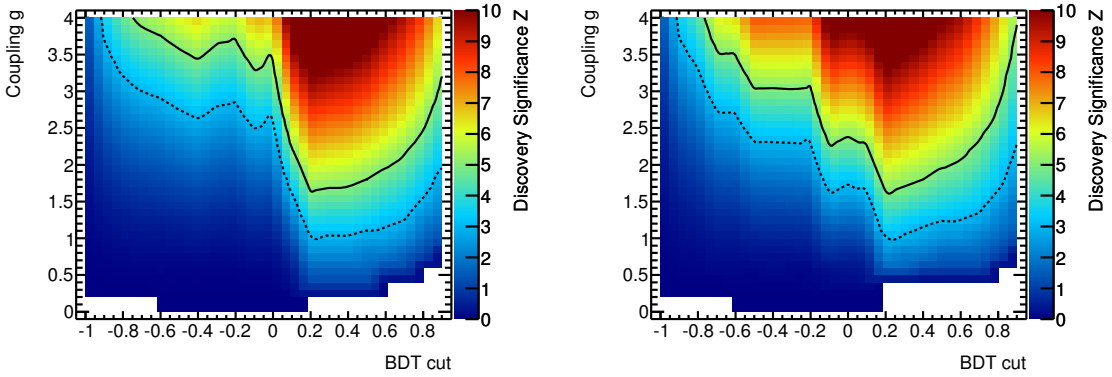
Figure C.1: Significance calculated with the profile likelihood method for the σ ($m = 850$ GeV) resonance without and with pile-up. An integrated luminosity of $100\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and a cut on the angle between the leptons with the largest p_T is applied. The equivalent integrated Monte Carlo luminosity of the backgrounds is scaled to $100\,\mathrm{fb}^{-1}$.



(a) φ resonance without pile-up.

(b) φ resonance with pile-up.

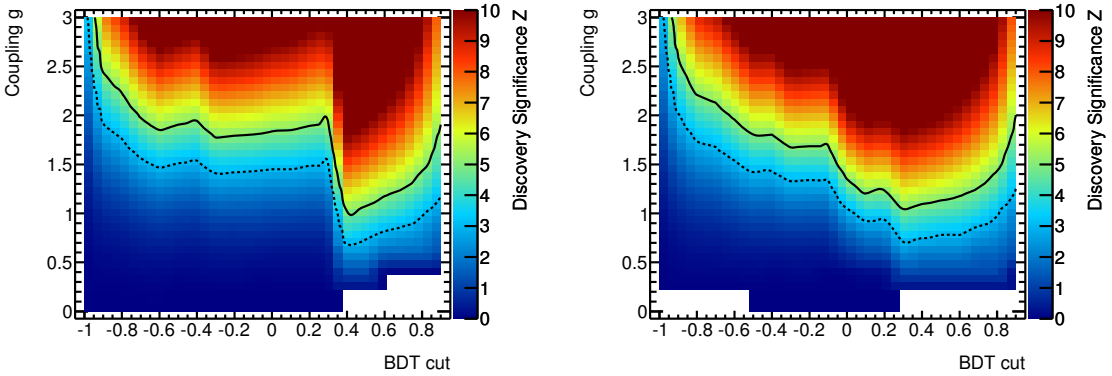
Figure C.2: Significance calculated with the profile likelihood method for the ϕ ($m = 850$ GeV) resonance without and with pile-up. An integrated luminosity of $100\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and a cut on the angle between the leptons with the largest p_T is applied. The equivalent integrated Monte Carlo luminosity of the backgrounds is scaled to $100\,\mathrm{fb}^{-1}$.



(a) ρ resonance without pile-up.

(b) ρ resonance with pile-up.

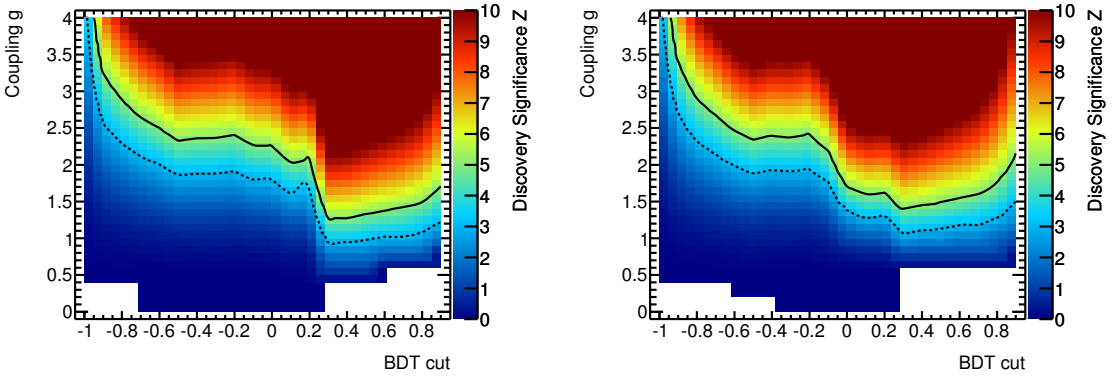
Figure C.3: Significance calculated with the profile likelihood method for the ρ ($m = 850$ GeV) resonance without and with pile-up. An integrated luminosity of $100\,\mathrm{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and a cut on the angle between the leptons with the largest p_T is applied. The equivalent integrated Monte Carlo luminosity of the backgrounds is scaled to $100\,\mathrm{fb}^{-1}$.



(a) f resonance without pile-up.

(b) f resonance with pile-up.

Figure C.4: Significance calculated with the profile likelihood method for the f ($m = 850$ GeV) resonance without and with pile-up. An integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and a cut on the angle between the leptons with the largest p_T is applied. The equivalent integrated Monte Carlo luminosity of the backgrounds is scaled to $100\,\text{fb}^{-1}$.



(a) t resonance without pile-up.

(b) t resonance with pile-up.

Figure C.5: Significance calculated with the profile likelihood method for the t ($m = 850$ GeV) resonance without and with pile-up. An integrated luminosity of $100\,\text{fb}^{-1}$ at $\sqrt{s} = 14$ TeV is assumed and a cut on the angle between the leptons with the largest p_T is applied. The equivalent integrated Monte Carlo luminosity of the backgrounds is scaled to $100\,\text{fb}^{-1}$.

List of Figures

2.1 Muon decay in Fermi Theory and Standard Model 6
2.2 K-matrix unitarization formalism 15

3.1 Overview of the ATLAS detector at the LHC 19
3.2 Cut-away view of the ATLAS Inner Detector 20
3.3 Expected number of primary vertex candidates 24
3.4 Schematic explanation of the jet-vertex fraction 25
3.5 Jet multiplicity as a function of reconstructed vertex multiplicity 26
3.6 Feynman diagrams of Vector Boson Scattering 28
3.7 Schematic signature of Vector Boson Scattering 29

4.1 ATLAS Monte Carlo production chain 32
4.2 Feynman diagram: Schematic signal process 33
4.3 Feynman diagrams: Top pair background 36
4.4 Feynman diagram: Wt background 38
4.5 Feynman diagrams: W+jets and Z+jets background 38
4.6 Feynman diagrams: WHIZARD irreducible backgrounds 39

5.1 Visualization of the available signal samples used for training 49
5.2 Signal linear correlations of the training input variables 51
5.3 Overtraining test for a training of a Boosted Decision Tree 52
5.4 Background rejection versus signal efficiency 53
5.5 Boosted Decision Tree cutflow for summed backgrounds and φ resonance with different couplings 54
5.6 Boosted Decision Tree cutflow for all backgrounds 55
5.7 Angular separation between leptons with largest p_{T} - pile-up samples . 56
5.8 Angular separation between leptons with largest p_{T} - pile-up samples . 57

6.1 Application of approximate significance formula $\tilde{Z} = S/\sqrt{B}$ 62
6.2 Discovery significance as a function of coupling and cut on the Boosted Decision Tree output 63
6.3 Discovery significance for all resonances and couplings 64
6.4 Comparison between analyses with and without pile-up 65
6.5 Discovery significance as a function of the coupling for different luminosities 67
6.6 Discovery significance as a function of coupling and luminosity 68
6.7 Visualization of the signal samples for a training with independent samples 69
6.8 Retraining of a Boosted Decision Tree with independent samples and random training events 70

6.9 Visualization of the signal samples for the training with different sample size . . . 71
6.10 Retraining of a Boosted Decision Tree with different signal sample sizes 72
6.11 Visualization of the signal samples for a training comparison with fast- and fully-simulated events . . . 73
6.12 Comparison between fast- and fully-simulated signal samples . . . 73
6.13 Discovery significance for all resonances and couplings including systematic uncertainties . . . 76

A.1 Distribution of largest flavor tag weight b_1 . . . 79
A.2 Distribution of second to largest flavor tag weight b_2 . . . 80
A.3 Distribution of invariant mass of tagging jets m^{tagjet} . . . 80
A.4 Distribution of largest minijet transverse momentum $p_{\text{T},1}^{\text{minijet}}$ . . . 81
A.5 Distribution of the lepton centrality ζ . . . 81
A.6 Distribution of the largest lepton transverse momentum $p_{\text{T},1}^{\text{lepton}}$ . . . 82
A.7 Distribution of the pseudorapidity separation of the tagging jets $\Delta\eta^{\text{tagjets}}$ 82
A.8 Distribution of the transverse momentum balance $p_{\text{T}}^{\text{bal}}$ . . . 83
A.9 Distribution of the second to largest minijet transverse momentum $p_{\text{T},2}^{\text{minijet}}$ 83
A.10 Distribution of the second to largest lepton transverse momentum $p_{\text{T},2}^{\text{lepton}}$ 84

B.1 Boosted Decision Tree cutflow for background and different signal samples 85
B.2 Boosted Decision Tree cutflow for background and different signal samples 86

C.1 Profile likelihood discovery significance for σ resonance . . . 87
C.2 Profile likelihood discovery significance for φ resonance . . . 88
C.3 Profile likelihood discovery significance for ρ resonance . . . 88
C.4 Profile likelihood discovery significance for f resonance . . . 89
C.5 Profile likelihood discovery significance for t resonance . . . 89

List of Tables

2.1 Fermions of the Standard Model 4
2.2 Bosons of the Standard Model 4
2.3 Fermion fields of the Standard Model and their quantum numbers . . . 7
2.4 Resonances considered in this work with their properties 14

3.1 Pile-up options file of the samples considered within this work 27

4.1 Signal samples, couplings, cross sections and events 34
4.2 Samples generated by WHIZARD and generation options 35
4.3 Background samples, events and cross sections 40
4.4 Boosted Decision Tree input variables, separation 41

6.1 Optimal cuts on the Boosted Decision Tree output and minimal coupling per resonance for discovery 64
6.2 Minimal coupling per resonance for discovery and effects of systematic uncertainties and pile-up 76

Bibliography

[1] J. Schumacher, *Sensitivity of ATLAS to Alternative Mechanisms of Electroweak Symmetry Breaking in Vector Boson Scattering $qq \to qql\nu l\nu$*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2010.

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264–1266.

[3] A. Salam, *Elementary Particle Theory*. Almqvist and Wiksell, Stockholm, 1968.

[4] S. Tomonaga, *On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields*, Progress of Theoretical Physics **1** (1946) no. 2, 27–42.

[5] J. Schwinger, *On Quantum-Electrodynamics and the Magnetic Moment of the Electron*, Phys. Rev. **73** (1948) no. 4, 416–417.

[6] R. P. Feynman, *Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction*, Phys. Rev. **80** (1950) no. 3, 440–457.

[7] S. L. Glashow, *Partial-Symmetries of Weak Interactions*, Nuclear Physics **22** (1961) no. 4, 579–588.

[8] K. Nakamura et al., *Quantum Chromodynamics*, Journal of Physics G **37** (2010) 075021.

[9] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) no. 26, 1346–1349.

[10] H. D. Politzer, *Asymptotic Freedom: An Approach to Strong Interactions*, Physics Reports **14** (1974) no. 4, 129 – 180.

[11] D. J. Gross and F. Wilczek, *Asymptotically Free Gauge Theories. I*, Phys. Rev. D **8** (1973) no. 10, 3633–3652.

[12] G. Zweig, *An SU_3 Model for Strong Interaction Symmetry and its Breaking; Part II*, CERN-TH-412, 1964.

[13] M. G. H. Fritzsch and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Physics Letters B (1973) no. 47, 365.

[14] K. Nakamura et al., *The Review of Particle Physics*, Journal of Physics G **37** (2010) 075021.

[15] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Progress of Theoretical Physics **28** (1962) no. 5, 870–880.

[16] G. Cowan and E. Gross, *Discovery Significance with Statistical Uncertainty in the Background Estimate*, ATLAS Statistics Forum (2008) ATL–PHYS–PUB–2009–086.

[17] S. Weinberg, *New Approach to the Renormalization Group*, Phys. Rev. D **8** (1973) no. 10, 3497–3509.

[18] F. J. Yndurain, *Limits on the Mass of the Gluon*, Phys. Lett. **B345** (1995) 524–526.

[19] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. W. H. Freeman, 1973.

[20] A. Einstein, *Die Feldgleichungen der Gravitation (The Field Equations of Gravitation)*, Königlich Preußische Akademie der Wissenschaften, 1915.

[21] G. Ross, *Grand Unified Theories*. Westview Press, 1984.

[22] K. Nakamura et al., *Grand Unified Theories*, Journal of Physics G **37** (2010) 075021.

[23] W. Kilian, *Electroweak Symmetry Breaking: The Bottom-Up Approach*. Springer, 2003.

[24] S. Weinberg, *The Quantum Theory of Fields*. Cambridge University Press, Vol. I-III, 1995.

[25] M. C. Brak, *The Hierarchy Problem in the Standard Model and Little Higgs Theories*, Master's thesis, NIKHEF, 2004.

[26] G. Bertone, D. Hooper, and J. Silk, *Particle Dark Matter: Evidence, Candidates and Constraints*, Physics Reports **405** (2005) no. 5–6, 279–390.

[27] U. Sarkar, *Particle and Astroparticle Physics*. CRC Press. pp. 429, 2007.

[28] F. L. Wilson, *Fermi's Theory of Beta Decay*, American Journal of Physics **36** (1968) no. 12, 1150–1160.

[29] E. Fermi, *Golden Rule*, Nuclear Physics, 1950.

[30] F. Halzen and M. Alan, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley and Sons, 1984.

[31] L.-L. Chau and W.-Y. Keung, *Comments on the Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **53** (1984) no. 19, 1802–1805.

[32] M. E. Peskin and D. V. Schroeder, *An Introduction To Quantum Field Theory*. Westview Press, 1995.

[33] P. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) no. 16, 508–509.

[34] P. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (1966) no. 4, 1156–1163.

[35] S. Coleman, J. Wess, and B. Zumino, *Structure of Phenomenological Lagrangians. I*, Phys. Rev. **177** (1969) no. 5, 2239–2247.

[36] C. G. Callan, S. Coleman, J. Wess, and B. Zumino, *Structure of Phenomenological Lagrangians. II*, Phys. Rev. **177** (1969) no. 5, 2247–2250.

[37] P. Sikivie, L. Susskind, M. Voloshin, and V. Zakharov, *Isospin Breaking in Technicolor Models*, Nuclear Physics B **173** (1980) no. 2, 189 – 207.

[38] S. Weinberg, *Physica*, 96A, 327, 1979.

[39] A. Alboteanu, W. Kilian, and J. Reuter, *Resonances and Unitarity in Weak Boson Scattering at the LHC*, JHEP **11** (2008) 10, `arXiv:hep-ph/0806.4145`.

[40] G. Valencia and S. Willenbrock, *Goldstone-Boson Equivalence Theorem and the Higgs Resonance*, Phys. Rev. D **42** (1990) no. 3, 853–859.

[41] A. Dobado and J. R. Pelaez, *The Inverse Amplitude Method in Chiral Perturbation Theory*, Phys. Rev. **D56** (1997) 3057–3073, `arXiv:hep-ph/9604416`.

[42] P. Masjuan, J. J. Sanz-Cillero, and J. Virto, *Some Remarks on the Pade Unitarization of Low-Energy Amplitudes*, Phys. Lett. **B668** (2008) 14–19, `arXiv:hep-ph/0805.3291`.

[43] S. N. Gupta, *Quantum Electrodynamics*. Gordon and Breach, 1981.

[44] M. S. Chanowitz, *K-Matrix*, Phys. Rept. **320** (1999) 139.

[45] L. Evans and P. Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) no. 08, S08001.

[46] A. Straessner, *LHC State of the Art and News*, ATL-GEN-PROC-2010-005, 2010.

[47] The ALICE Collaboration, K. Aamodt et al., *The ALICE Experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) no. 08, S08002.

[48] The ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) no. 08, S08003.

[49] The CMS Collaboration, S. Chatrchyan et al., *The CMS Experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) no. 08, S08004.

[50] A. Augusto et al., *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008) no. 08, S08005.

[51] The LHCf Collaboration, O. Adriani et al., *The LHCf Detector at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) no. 08, S08006.

[52] The TOTEM Collaboration, G. Anelli et al., *The TOTEM Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) no. 08, S08007.

[53] J. Pinfold, *Technical Design Report of the MoEDAL Experiment*, Tech. Rep. CERN-LHC-2009-006. MoEDAL-TDR-001, CERN, Geneva, 2009.

[54] The ATLAS Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, `arXiv:hep-ex/0901.0512`.

[55] The Inner Detector Collaboration, G. Aad et al., *ATLAS Inner Detector: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1997.

[56] The ATLAS Collaboration, G. Aad et al., *ATLAS Pixel Detector Electronics and Sensors*, Journal of Instrumentation **3** (2008) no. 07, P07007.

[57] G. Lindström, *Radiation damage in silicon detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **512** (2003) no. 1-2, 30–43. Proceedings of the 9th European Symposium on Semiconductor Detectors.

[58] The ATLAS TRT Collaboration, E. Abat et al., *The ATLAS Transition Radiation Tracker (TRT) Proportional Drift Tube: Design and Performance*, Journal of Instrumentation **3** (2008) no. 02, P02013.

[59] C. Ohm and T. Pauly, *The Software of the ATLAS Beam Pick-up Based LHC Monitoring System*, Journal of Physics: Conference Series **219** (2010) no. 2, 022040.

[60] C. Ruwiedel, G. Gaycken, and J. Kroseberg, *Study of the Use of Jet-Vertex Association for the Suppression of Pileup Jets in the Central Jet Veto in VBF* $H \to \tau^+\tau^- \to lh$, Tech. Rep. ATL-COM-PHYS-2009-237, CERN, Geneva, 2009.

[61] M. Boonekamp, F. Gianotti, R. A. McPherson, M. Nessi, and P. Nevksi, *Cosmic Ray, Beam-Halo and Beam-Gas Rate Studies for ATLAS Commissioning*, Tech. Rep. ATL-GEN-2004-001, CERN, Geneva, 2004.

[62] A. Rossi and N. Hilleret, *Residual Gas Density Estimations in the LHC Experimental Interaction Regions*, Prepared for 2nd national conference of the german vacuum society, berlin, germany, 13-26 june 2003.

[63] D. W. Miller, A. Schwartzman, and D. Su, *Pile-up Jet Energy Scale Corrections Using the Jet-Vertex fraction method*, Tech. Rep. ATL-COM-PHYS-2009-180, CERN, Geneva, 2009.

[64] G. Piacquadio, K. Prokofiev, and A. Wildauer, *Primary Vertex Reconstruction in the ATLAS Experiment at LHC*, Journal of Physics: Conference Series **119** (2008) no. 3, 032033.

[65] D. L. Rainwater and D. Zeppenfeld, *Observing $H \to W^{(*)}W^{(*)} \to e^{\pm}\mu^{\mp}\not{p}_T$ in Weak Boson Fusion with Dual Forward Jet Tagging at the CERN LHC*, Phys. Rev. **D60** (1999) 113004, arXiv:hep-ph/9906218.

[66] N. Metropolis and S. Ulam, *The Monte Carlo Method*, Journal of the American Statistical Association **44** (1949) no. 247, 335–341.

[67] W. Kilian, T. Ohl, and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, arXiv (2007) , arXiv:hep-ph/0708.4233v1.

[68] D. E. Soper, *Parton Distribution Functions*, Nucl. Phys. Proc. Suppl. **53** (1997) 69–80, arXiv:hep-lat/9609018v1.

[69] T. Sjöstrand, *PYTHIA 8 Status Report*, arXiv:hep-ph/0809.0303v1.

[70] S. Agostinelli et al., *Geant4 — A Simulation Toolkit*, Nuclear Instruments and Methods in Physics Research A **506** (2003) 250–303.

[71] J. Allison, K. Amako, J. Apostolakis, et al., *Geant4 — Developments and Applications*, IEEE Transaction on Nuclear Science **53, No. 1** (2006) 270–278.

[72] E. Richter-Was and D. Froidevaux, *ATLFAST 2.0 — A Fast Simulation Package for ATLAS*, ATL-PHYS-98-131, 1998.

[73] T. Sjöstrand and M. H. Seymour, *PYTHIA and HERWIG for Linear Collider Physics*, arXiv:hep-ph/9909346v1.

[74] M. Moretti, T. Ohl, and J. Reuter, *O'Mega: An Optimizing Matrix Element Generator*, arXiv:hep-ph/0102195v1.

[75] M. Lehmacher, *b-Tagging Algorithms and their Performance at ATLAS*, arXiv:hep-ex/0809.4896v3.

[76] R. Bonciani, S. Catani, M. L. Mangano, and P. Nason, *NLL Resummation of the Heavy-Quark Hadroproduction Cross Section*, arXiv:hep-ph/9801375v2.

[77] The ATLAS Collaboration, G. Aad et al., *Prospects for Measuring Top Pair Production in the Dilepton Channel with early ATLAS Data at $\sqrt{s} = 10$ TeV*, ATLAS note (2009) no. ATL-PHYS-PUB-2009-086, 20.

[78] The ATLAS Collaboration, G. Aad et al., *Prospects for the Top Pair Production Cross-section at $\sqrt{s} = 10$ TeV in the Single Lepton Channel in ATLAS*, Tech. Rep. ATL-PHYS-PUB-2009-087, CERN, 2009.

[79] A. Collaboration, *Observation of Top Quark Pair Production in ATLAS at $\sqrt{s} = 7\,TeV$*, Tech. Rep. ATL-COM-PHYS-2010-854, CERN, Geneva, 2010.

[80] S. Frixione and B. R. Webber, *The MC@NLO 3.4 Event Generator*, arXiv:hep-ph/0812.0770v1.

[81] T. A. C. Group, *ATLAS Computing: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2005. ATLAS TDR–017.

[82] B. P. Kersevan and E. Richter-Was, *The Monte Carlo Event Generator AcerMC 2.0 with Interfaces to PYTHIA 6.2 and HERWIG 6.5*, arXiv:hep-ph/0405247v1.

[83] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions*, arXiv:hep-ph/0206293v2.

[84] Z. Was, *Precision Simulations with TAUOLA and PHOTOS*, arXiv:hep-ph/0610386v2.

[85] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision Tool for QED Corrections in Z and W Decays*, arXiv:hep-ph/0506026v2.

[86] M. Lehmacher, *b-Tagging Algorithms and their Performance at ATLAS*, arXiv:hep-ph (2008) ATL–COM–PHYS–2009–180.

[87] J. Schaarschmidt, *A Study of b-Quark Associated Higgs Production in the Decay Mode $b\bar{b}h/A/H \to \tau\tau \to 2\ell + 4\nu$ with ATLAS at LHC*, Institut für Kern- und Teilchenphysik, TU Dresden (2007) Diploma Thesis.

[88] The BABAR Collaboration, P. F. Harrison, H. Quinn, et al., *The* BABAR *Physics Book*, SLAC-R (1998) 0504.

[89] M. A. Stephens, *EDF Statistics for Goodness of Fit and Some Comparisons*, Journal of the American Statistical Association **69** (1974) no. 347, 730–737.

[90] H. J. Kim, *Electron and Photon Identification Performance in ATLAS*, arXiv:hep-ex/0810.3415v1.

[91] A. Hoecker, P. Speckmayer, J. Stelzner, J. Therhaag, E. von Toerne, H. Voss, et al., *TMVA - Toolkit for Multivariate Data Analysis*, arXiv:physics/0703039.

[92] R. A. Fisher, *The Use of Multiple Measurements in Taxonomic Problems*, Ann. Eugen. **7** (1936) 179.

[93] C. M. Bishop, *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.

[94] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Wadsworth, 1984.

[95] I. Asimov, *Isaac Asimov: The Complete Stories, Vol. 1*. Broadway Books, Franchise, 1990.

Acknowledgment

An dieser Stelle möchte ich mich bei all denen bedanken, die zum Gelingen dieser Arbeit und meines Studiums beigetragen haben.

Besonders danke ich Prof. Michael Kobel, der es mir ermöglichte ein spannendes und vor allem aktuelles Thema aus dem Bereich der Teilchenphysik zu bearbeiten. Es ist wie erhofft ein guter Kompromiss aus experimenteller und theoretischer Physik geworden.

Am meisten möchte ich Dr. Jan Schumacher danken. Er war immer bereit meine Fragen zu beantworten und seine Vorarbeiten waren unverzichtbar. Die Arbeit mit seinen Programmen war eine Freude und hat meine Programmierkenntnisse sehr erweitert.

Ausserdem danke ich der gesamten ATLAS Gruppe der TU Dresden für Antworten auf zahlreiche Fragen, spannende Diskussionen, den täglichen Mensagang und außeruniversitäre Aktivitäten. Besonders möchte ich Thomas Göpfert, Andreas Ludwig, Dr. Wolfgang Mader, Marcus Morgenstern, Dr. Jana Schaarschmidt, Peter Steinbach, Jun. Prof. Arno Straessner, Felix Thomas, meiner Betreuerin Dr. Anja Vest und Klaus Zenker danken.

Dem Graduiertenkolleg 1504 “Masse-Spektrum-Symmetrie” danke ich für die interessanten Vorlesungen und Blockkurse.

Hiermit möchte ich mich auch bei Prof. Dominik Stöckinger bedanken, der mein Interesse für Theorie durch seine besondere Art gestärkt hat. Die Entscheidung für die experimentelle Teilchenphysik hat er mir sehr erschwert und lässt mich noch heute zweifeln. Eine Freude und große HIlfe waren außerdem die spannenden und teilweise Kopfschmerzen bereitenden Diskussionen mit den theoretischen Doktoranden und Diplomanden - auch euch ein großer Dank!

I also want to thank all CERN summerstudents ’08 and my supervisor Giulio Usai for the great experience at CERN and the wonderful NYEs. Without you this work would probably not be related to particle physics.

Meinen Freunden danke ich für die offenen Ohren und Türen. Für unvergessliche und vor allem motivierende Wochen in und außerhalb von Dresden möchte ich Doro, allen Erzgebirgern, den Lenbachs, den Mohns, Monsters und meiner ehemaligen Wohnheim-WG danken. Friedi danke ich für das leidteilende Aufschreiben in der finalen Phase dieser Arbeit.

Ein besonderes Dank geht hiermit an meine Eltern und Großeltern, da nur sie mir mein Studium ermöglicht haben und immer für mich da waren.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Dresden, den 12.11.2010

Philipp Anger