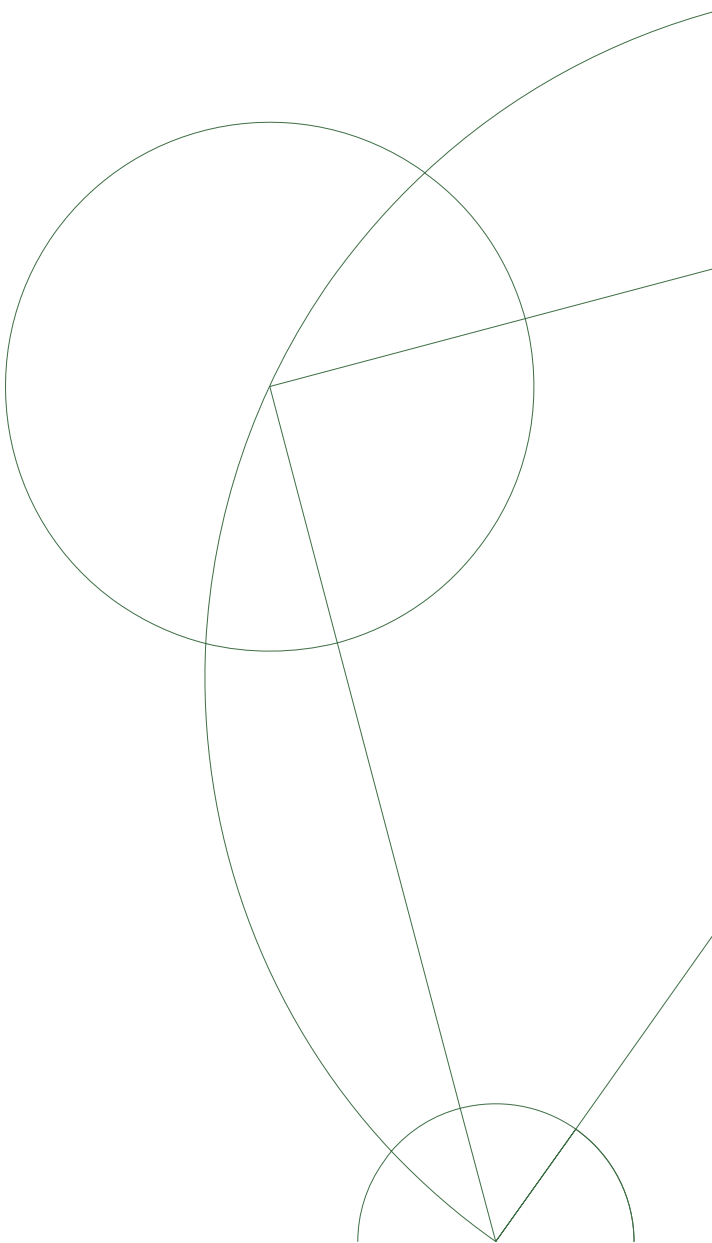
FACULTY OF SCIENCE
UNIVERSITY OF COPENHAGEN

Master's thesis

Limits on triple gauge boson couplings

Limits on the TGCs of the WWZ vertex using 827.87 pb^{-1} of data, recorded with the ATLAS detector at the LHC at $\sqrt{s} = 7$ TeV.

Christian Bierlich



Academic Advisor:
Jørgen Beck Hansen
Niels Bohr Institute, Discovery Center
University of Copenhagen

May 2012

Abstract

This thesis presents limits on the charged triple gauge boson couplings Δg_1^Z, $\Delta\kappa_Z$ and λ_Z using 110 $W^{\pm}Z$ events from pp collisions at a center-of-mass energy of 7 TeV, corresponding to 827.87 pb^{-1} of data recorded with the ATLAS experiment in 2011. From a Monte Carlo study it is shown, that the conventional choice of statistical procedure for extraction of the couplings – a maximum likelihood fit – can cause under coverage due to insufficiencies of the TGC model, in addition to under coverage from poor statistics. Hence a Neyman construction with Feldman-Cousins ordering is used to extract the couplings, and for providing a framework for propagation of systematical errors.

The limits set at a 95 % CL are:
From data corresponding to 827.87 pb^{-1} of integrated luminosity:

$$-0.098 < \lambda_Z < 0.097\ ,\ -0.46 < \Delta\kappa_Z < 0.73 \text{ and } -0.11 < \Delta g_1^Z < 0.20.$$

With Standard Model expectation:

$$-0.056 < \lambda_Z < 0.058\ ,\ -0.46 < \Delta\kappa_Z < 0.53 \text{ and } -0.062 < \Delta g_1^Z < 0.100.$$

All reported limits are compatible with the Standard Model.

Dansk Resume

I dette speciale præsenteres grænser for tre-boson koblingerne Δg_1^Z, $\Delta\kappa_Z$ og λ_Z fra 110 $W^{\pm}Z$ begivenheder i pp kollisioner ved en massemidtpunktsenergi på7 TeV. Der bruges data tilsvarende et integreret luminositet på 827.87 pb^{-1} , opsamlet af ATLAS eksperimentet i 2011. Gennem et Monte Carlo studie vises det, at den konventionelle statistiske procedure til bestemmelse af koblingerne – et maksimum likelihood fit – kan medføre underdækning grundet utilstrækkeligheder i TGC modellen, udover underdækning grundet dårlig statistik. Derfor benyttes en Neyman konstruktion med en Feldman-Cousins sortering til at bestemme koblingerne, og til propagering af systematiske usikkerheder.

De fundne grænser for tre-boson koblingerne er, med 95 % konfidens:
Fra data tilsvarende en integreret luminositet på 827.87 pb^{-1} :

$$-0.098 < \lambda_Z < 0.097\ ,\ -0.46 < \Delta\kappa_Z < 0.73 \text{ og } -0.11 < \Delta g_1^Z < 0.20.$$

Med en forventning givet af Standardmodellen:

$$-0.056 < \lambda_Z < 0.058\ ,\ -0.46 < \Delta\kappa_Z < 0.53 \text{ og } -0.062 < \Delta g_1^Z < 0.100.$$

Alle fundne grænser er kompatible med Standardmodellen.

CONTENTS

CHAPTER 1

INTRODUCTION

The Standard Model is currently being put through a thorough test at the Large Hadron Collider (LHC). Since the Standard Model is conjectured to break down at a finite energy scale, much effort goes into looking beyond the this theory, and into the landscape of New Physics.

The means of looking beyond the Standard Model, can be divided into two categories. A model dependent one and a model independent one.

The model dependent one is by far the most widespread. Although models for new physics are very different, the recipe for such searches always builds on falsification of the predictions made by such a model. In practice, limits are set on masses and cross sections of the particles and processes of such a new theory, making the possible range of these less likely, until they are excluded.

If one has an idea of what kinds of physics to look for, this is indeed the correct way of observing such New Physics. A major problem is, that the number of proposed theories is so vast, and almost every theory contains a number of free parameters that can be tuned in order to make the theory even harder to fully falsify.

Another approach is model independent. The idea is, that since the Standard Model is a very successful theory, it can be taken as a low energy approximation of the theory of New Physics. The deviations from the Standard Model at higher energies are modeled much in the same way as one would make a series expansion of a function in the neighborhood of a minimum. The parametrization that will be considered in this thesis takes as its starting point the vertex where a W particle radiates a Z. This vertex is parametrized in *triple gauge boson couplings*, and the end result of the conducted study is to set limits on these.

As the model considered is a general parametrization of an existing vertex using only Standard Model particle content, limits on triple gauge boson couplings truly are model independent limits on deviations from Standard Model physics.

The search for New Physics is motivated by the desire to *know* how the world works at a most fundamental level. Increasing energy scales corresponds to decreasing length scales, such that the higher the energy, the more fundamental the physics. To set limits

on, New Physics is motivated by a desire to know the boundaries of the yet most precise theory of particle physics.

In experimental high energy physics, these questions and answers are formulated in the language of statistics. Limits on New Physics are quoted as *confidence intervals* of measured parameters, and to understand the final limits presented in this thesis it is necessary to understand how the limit setting procedure influences the intervals themselves.

1.1 Contribution

In this study limits on the three specific triple gauge boson couplings of the WWZ vertex, $\Delta\kappa_Z$, Δg_1^Z and λ_Z, is set at a 95 % CL using 827.87 pb^{-1} of data, taken by the ATLAS detector at LHC in 2011. Much emphasis has been put on the statistical part of this limit setting procedure, and the conclusions reached regarding choice of procedure, is not limited to this specific search with this amount of data. This study employs:

- A theoretical introduction on the approach for including non Standard Model interactions between the gauge bosons.
- The usage of Monte Carlo methods to make predictions of the effects of these interactions, both using direct event generation, and using a Matrix Element Reweighting technique, including a validation of this.
- A discussion of statistical methods for setting limits on couplings, and how the choice of method possibly affects a given limit, based on a Monte Carlo truth study.
- An introduction of the ATLAS and LHC experiments to give the reader a complete understanding of the obtained experimental results, and limitations of the analysis set by the experimental apparatus.
- A cut based event selection procedure of a final state with three leptons and missing transverse energy, supplemented by a multivariate selection procedure.
- An estimation of limits on three triple gauge boson couplings, using 827.87 pb^{-1} of data, including systematic error.

The work conducted in this study is a part of the effort to produce a combined analysis tool, to fit all triple gauge boson couplings from ATLAS data.

CHAPTER 2

THEORETICAL BACKGROUND

In this chapter the theoretical basis of the experimental search, that is the main subject of this thesis, will be laid out in a quite detailed manner. The gauge structure of the Standard Model of Particle Physics (SM) and its background in quantum field theory (QFT) will be explained. The reason for explaining the theory in such detail is to motivate the introduction of triple gauge boson couplings (TGCs) as a natural way of making a model independent parametrization of New Physics (NP), based on the SM. To gain this understanding of the subject it is important to:

- Put the TGCs in a firmly rooted context of the SM and gauge invariance.
- Emphasize the SM rooting in QFT, especially in the context of symmetries.

The four great realms of mechanics can be depicted as in figure 2.1, with the everyday, macroscopic world revolving at low speeds sitting in the lower left corner. When the classical theory of motion is taken to the limits of c and $\hbar$, it breaks down and is replaced by special relativity and quantum mechanics. When applying both at the same time, a theory which allows both for quantum fluctuations according to the Heisenberg Uncertainty Principle, and equivalence of energy and matter via $E = mc^2$ is obtained. This theory is QFT, and is the basis of which the SM resides.

QFT is an active area of research, and especially the attempts of adding another axis to figure 2.1 and make gravity a part of the big picture, is a matter of most interest to researchers. SM physics, however, does not attempt to include this, and even before the theory is built, it is manifest, that it must break down at some scale of energy – at least when gravity comes into play. This is not necessarily a problem, since the SM provides many good predictions at lower energy scales. However, when doing collider physics experiments at the energy frontier, deviations from SM physics is certainly a possibility that must be included.

Besides the inability to describe gravity, the SM suffers from several other challenges pointing in the direction that it must break down and be replaced by a more fundamental

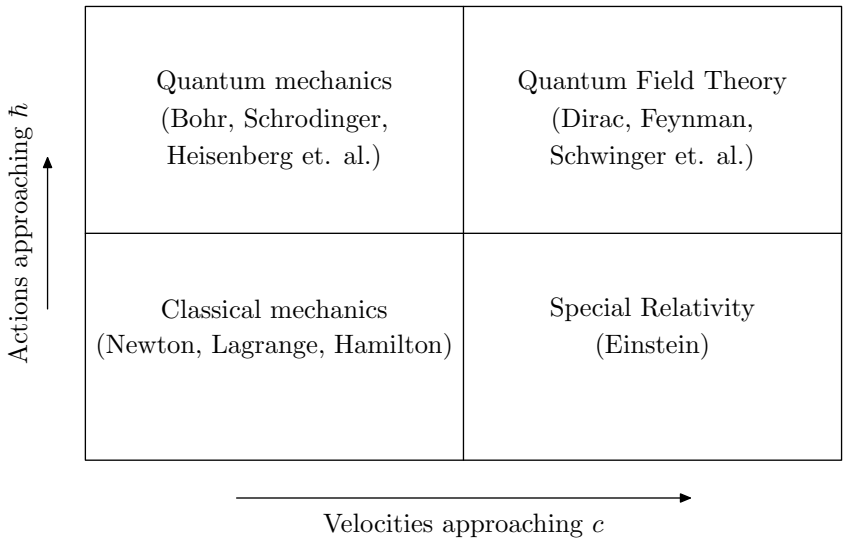


Figure 2.1: The four realms of mechanics, with inspiration from Ref. [1]

theory at a higher energy scale. The lack of Dark Matter candidates, matter anti-matter asymmetry and ad hoc introduction of (experimentally observed) particle masses are all challenges based on experimental observations of which the SM does not adequately explain. Moreover some theoretical aspects about the SM are unsatisfactory, without being a problem per se. This includes the many free parameters of the SM, which must all be measured and put in by hand.

Despite the challenges, the SM have proved extremely successful in providing predictions about particles, their properties and their decays, and is currently regarded as the best theory of fundamental particles and their interactions.

2.1 Particles and symmetries of the Standard Model

This section will serve as a short summary of the SM in tabulated form. In the forthcoming sections the SM and its background in QFT will be more developed in more detail.

The Standard Model is basically a description of the existing fundamental particles and their mutual interactions, governed by underlying principles of symmetry. The matter constituents are the spin-$\frac{1}{2}$ *fermions* and the force carrying *bosons* with integer spin. The fermions are further divided into *quarks* and *leptons*. Quarks always carry a non-zero *color charge* and interacts via the strong force. As color charged objects cannot exist freely, quarks are *confined* in *hadrons* consisting of either a quark and an anti-quark (*meson*) or three quarks (*baryon*).

Leptons are sub-divided into charged leptons and neutrinos, and can exist freely. The charged leptons can interact with quarks through the electromagnetic interaction, as quarks also carry electric charge. All fermions can interact via the weak force. Table 2.1

Quarks			Leptons		
Flavour	Mass	Charge [e]	Flavour	Mass	Charge [e]
up (u)	0.005 GeV	$-\frac{1}{3}$	e	0.511 MeV	-1
down (d)	0.003 GeV	$\frac{2}{3}$	ν_e	$< 2 \cdot 10^{-6}$ MeV	0
charm (c)	1.27 GeV	$-\frac{1}{3}$	μ	105.7 MeV	-1
strange (s)	0.101 GeV	$\frac{2}{3}$	ν_μ	< 0.19 MeV	0
top (t)	172 GeV	$-\frac{1}{3}$	τ	1776.8 MeV	-1
bottom (b)	4.2 GeV	$\frac{2}{3}$	ν_τ	< 18.2 MeV	0

Table 2.1: Table of the fermion content of SM. Masses and charges are from Ref. [3]. For each fermion there exists a corresponding anti-particle partner with the opposite charge. As the SM treats neutrinos as massless, the given neutrino masses are experimental limits.

Force	Gauge boson	Mass	Charge [e]
Electromagnetic	Photon (γ)	$< 1 \cdot 10^{-18}$ eV	0
Weak	$W^\pm$	80.4 GeV	± 1
	Z^0	91.2 GeV	0
Strong	Eight gluons	0	0

Table 2.2: Table of the gauge bosons of SM. Masses and charges are from Ref. [3]. The photon mass is the experimental limit, as photons are treated massless in the SM.

sums up the fermion content of the SM[1]. All the fermions have a so-called *anti-particle* partner carrying the opposite gauge charges and is denoted with a bar, such that $\bar{f}$ is the anti-particle of f.

As mentioned, the forces of the SM are the strong, weak and electromagnetic forces, which are all mediated by corresponding gauge bosons. All interactions are modeled by an exchange of such force carriers. The exchange of gauge bosons are parameterized in strength parameters called *coupling constants*.

Interactions are illustrated with Feynman diagrams. An example of two such diagrams is given in figure 2.2. Intersecting lines are called *vertices*. The allowed vertices in the SM are governed by principles of symmetry, and the aim of this thesis is to investigate an effective theory in which new physics is parameterized in terms of a certain type of coupling constants; so-called TGCs. The gauge bosons are listed in table 2.2.

The last particle of SM is the Higgs boson, which is the quanta of a hypothesized Higgs Field, introduced in order to give mass to the other SM particles. The Higgs boson has yet to be verified experimentally, but is prospected to be either discovered or excluded in the course of 2012. Hints could point in the direction of $m_H \approx 125$ GeV [2].

[1] In this section, and throughout the thesis, units where $c = \hbar = 1$ will be used, as it is customary in particle physics.

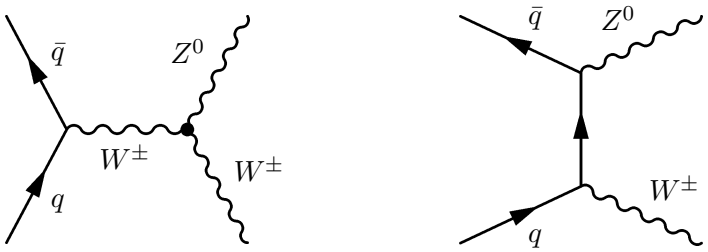


Figure 2.2: The triple gauge vertex of interest (the black dot) is this thesis depicted (left), in a so-called s-channel diagram where two quarks goes to a W, in turn going to a Z and a W, (right) a t-channel diagram which also results in a WZ final state, but does not contain the triple gauge boson vertex.

2.1.1 Symmetries of SM

As mentioned, the SM is built upon the belief that nature obeys a series of symmetries. The most important symmetry in the SM, is the so-called *gauge symmetry*, which will be explained in detail in forthcoming sections. Since the theory is relativistic, it also needs to be symmetric under transformations of the Poincaré group[2]. Besides gauge symmetry and Poincaré symmetry, three other – so-called discrete symmetries – will be of relevance:

- Charge conjugation (C) changes particles to anti-particles and vice versa, by conjugating all internal quantum numbers [3]: $\psi \mapsto \psi' = \bar{\psi}$.
- Parity transformation (P) reverses the spatial part of the space-time coordinates of the system: $\psi(t, \vec{x}) \mapsto \psi' = \psi(t, -\vec{x})$.
- Time reversal (T) reverses the temporal part of the space-time coordinates of the system: $\psi(t, \vec{x}) \mapsto \psi' = \psi(-t, \vec{x})$.

Experimental evidence suggests that the strong and electromagnetic forces are invariant under C, P and T transformations separately (and hence also all combinations of these), whereas the weak force violates both C-symmetry and P-symmetry. The weak force itself preserves CP-symmetry, but in SM where the strong force is also included, CP-symmetry is violated, and the only combination which is not violated, is CPT-symmetry[3]. Violation of CP symmetry is manifest in the SM, in the concept of quark mixing [3].

A way to test if a theory obeys a certain symmetry is to test if the Lagrangian (see section 2.2) is left unaltered when the symmetry is applied.

The importance of symmetries arises from Noether's theorem which states, that for each symmetry exists an associated conservation law [5]. Since conservation laws are good candidates for experimental tests, the language of symmetries turns out to be a suitable choice for building up the SM.

[2]The Poincaré group is the full symmetry group of Special Relativity, and consists of Lorentz boosts and rotations (together the Lorentz group) plus space-time translations. In literature sometimes denoted the *inhomogenous* Lorentz group.

[3]A violation of CPT-symmetry would lead to violation of Lorentz invariance [4].

2.1.2 An effective expansion of the SM

The subject of this thesis is to quantify the level of deviation from the SM in terms of an effective theory involving TGCs. An effective theory can be thought of as a series expansion of SM in energy, with the SM as the zeroth order term, i.e. valid for low energies. As the energy goes up, the expected deviations from the SM becomes more outspoken, and as such the additional effective terms become different from zero.

This thesis will investigate the deviation from SM in the WWZ vertex, depicted in figure 2.2 (left). If only processes that *conserves C and P symmetry separately* are considered, the vertex is fully described by three independent couplings. Two of these couplings already exist in the SM, and is denoted g_1^Z and κ_Z, and have an SM expectation value of one. The third coupling is denoted λ_Z and does not exist in the SM.

2.1.3 Effect of New Physics on TGCs

Many hypothesized models that affect the WWZ vertex exist in literature. Two types of such models can be imagined. Either the NP enters as new particles (fermions and/or bosons) that, created on mass-shell, contributes at tree level to the same final states as the WWZ vertex, else it enters as new, radiative loop corrections. From an experimental point of view, the models that predict contributions at tree level are the most interesting, as the contributions are higher than for the radiative loop corrections, and some even within the sensitivity of the LHC. Noteworthy models contributing at tree level to the WZ final state are (not limited to):

- Supersymmetry [6], where heavy, supersymmetric partner-particles decays to W's and Z's.
- Charged Higgs models, where a charged Higgs plays the role of the W decaying into the WZ pair.
- Technicolor [7], where a techni-R-meson plays the role of the W decaying into the WZ pair.

The TGC model makes the most general expansion of the WWZ vertex as possible, and measurement of TGCs does as such not limit itself to one specific model, but can provide hints of what to look for in present and future scenarios for NP. This is of course both a strength and a drawback of using TGCs, as the ability to do a broad search comes with the cost of not being able to say anything specific, but low energy approximative statements about the NP.

2.2 The Lagrangian formalism and global gauge invariance

As summarized in the previous section, the SM is a QFT made up of symmetries and a particle content. The remarkable thing about the SM is, that by imposing *gauge invariance* under a given local gauge group, the interaction terms of the theory are fixed. It is exactly from these interaction terms the TGCs emerge.

In order to impose gauge invariance, a framework for doing so must be introduced along with sensible objects to serve as the »starting point« on which the symmetry can be imposed. This is done by first introducing the *Lagrangian formalism* and then the Klein-Gordon and Dirac Lagrangians, which are the starting points for building theories of bosonic and fermionic content respectively [5].

The Lagrangian formalism or the *Path Integral Representation* [8] is based on similar notions from classical physics. The basic idea is that a system can be fully characterized by some function of the field (ψ) and its derivative ($\partial_\mu\psi$): the Lagrangian L. The *Principle Of Least Action* states that for a system moving from a configuration t_1 to t_2, the integral over the path:

$$S = \int_{t_1}^{t_2} L dt = \int_{t_1}^{t_2} \mathcal{L}(\psi, \partial_\mu\psi) d^4x, \tag{2.1}$$

will take the least possible value. $\mathcal{L}$ is called the Lagrangian density, and is related to L as indicated. It is customary to work in terms of Lagrangian densities, and $\mathcal{L}$ will simply be referred to as the Lagrangian for the remainder of the thesis. S is denoted the action of the system. Making the variation of the action leads to the equations of motion of the system, called the *Euler-Lagrange equations*:

$$\partial_\mu(\partial_{\partial_\mu\psi}\mathcal{L}) - \partial_\psi\mathcal{L} = 0. \tag{2.2}$$

The equations of motion specified by equation (2.2) are not quantized. The field needs to be replaced with the field operator, and the system needs not follow the classical path described by equation (2.2), but can make fluctuations based on the fact, that the operators do not *a priori* commutate. Taking the simplest example of the free scalar field, the Lagrangian is [9]:

$$\mathcal{L}_{fs} = \partial_\mu\psi^\dagger\partial^\mu\psi - m^2\psi^\dagger\psi. \tag{2.3}$$

As mentioned, the symmetry properties of the theory can be investigated directly in the Lagrangian. In this case, equation (2.3) is for example invariant under *global gauge transformations*, i.e. transformations of the type:

$$\psi \mapsto \psi' = \exp(i\alpha)\psi, \tag{2.4}$$

when α is constant (also referred to as a *global* $U(1)$ symmetry). Later the stronger condition of local gauge invariance will be imposed.

Writing up the Euler-Lagrange equations for this specific Lagrangian (equation (2.3)), yields the *Klein-Gordon equation*:

$$(\Box + m^2)\psi = 0. \tag{2.5}$$

The Klein-Gordon equation is the scalar, free field version of the Schrödinger equation. Since we know that fermions exist, the theory also needs to account for spin. This is done by introducing spinor fields represented by Dirac γ-matrices. The resulting Dirac Lagrangian is [9]:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \tag{2.6}$$

where $\bar{\psi}$ is the Dirac adjoint $\bar{\psi} \equiv \psi^\dagger \gamma^0$. The corresponding Euler-Lagrange equations are the Dirac equation:

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \tag{2.7}$$

which is the spin-$\frac{1}{2}$ free field equivalent of the Schrödinger equation, and serves as the primary »starting point« for the SM.

2.2.1 Feynman diagrams and matrix elements

After writing up a Lagrangian for a system, the next step is to quantize it. This can only be done by approximation. A very successful path is to first consider the classical path, and then introduce quantum corrections through the method of *Feynman diagrams*.

Feynman diagrams is a clever way of distinguishing between possible paths for the transition from a given initial state to a final state. Different paths correspond to topologically different Feynman diagrams. Each diagram contributes to the sum that is called the matrix element $\mathcal{M}$. The matrix element then contains the dynamics of the interaction, and can be translated into a differential cross section.

The calculation of differential cross sections is very central, as these predict the distributions of kinematic variables seen in the detector. Since the differential cross sections depends directly on the matrix element, the effect of TGCs will later on be implemented by *matrix element reweighting*.

Feynman diagrams are basically built up of three ingredients (see also figure 2.3):

Vertices (figure 2.3 (left)) represents an interaction where some particles (lines) go in, and other particles go out. Given a certain theory some vertices are allowed, and some are not. For each type of vertex there exists a *coupling constant*, which describes the strength of this vertex.

Propagators (figure 2.3 (middle)) describes the free movement between vertices, and really corresponds to the momentum of the particle. Only for propagators going into or out of the diagram the particles are required to be on-shell. These are referred to as *external* lines as opposed to *internal* lines.

Internal Loops (figure 2.3 (right)) is a particle/anti-particle pair created and annihilated in two consecutive vertices. Thus the momentum is undetermined, and the problem of divergencies emerges.

How a certain ingredient contributes to the matrix element is governed by the corresponding *Feynman rules*, which are deduced from the Lagrangian. It should be noted that the three ingredients mentioned here, can be many-fold degenerate, such that, for example, a number of different propagators exist, each with their corresponding Feynman rule and contribution to the matrix element.

Using the Feynman rules for a given theory, $\mathcal{M}$ can be calculated for a given process. $\mathcal{M}$ will be a function of the masses (m) and couplings (g) involved, along with phase space information (Ω):

$$\mathcal{M} = \mathcal{M}(m, g, \Omega). \tag{2.8}$$

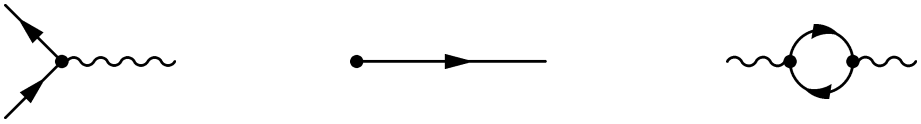


Figure 2.3: Illustration of the three basic ingredients of a Feynman diagram. From the left: A vertex, a(n external) propagator and an internal loop.

In experimental particle physics, the observable of choice is a differential cross section. As $\mathcal{M}$ is interpreted as the amplitude of the transition from initial to final state at a specific phase space point, it can be squared and normalized to the incident flux (Φ), to obtain the differential cross section for a differential phase space element ($d\Omega$):

$$d\sigma = \frac{|\mathcal{M}|^2}{\Phi} d\Omega. \tag{2.9}$$

The direct link from the symmetry content of a theory is now related first through the Lagrangian, then through matrix elements to the differential cross sections, which will depend on the TGCs. As the matrix elements are linear functions of the TGCs (since they enter the Lagrangian linearly), $d\sigma$ will be a quadratic function of the TGCs.

2.2.2 Next to Leading Order

A problem arises as $\mathcal{M}$ is computed according to the Feynman rules of the theory. As the contributions from topologically different Feynman diagrams are added together, it is quickly realized that new diagrams can always be created by adding extra vertices. The matrix element will then appear as a power series in the couplings. Given coupling constants less than one, the higher order terms will be suppressed, and as such, the first couple of terms will be sufficient to make a prediction within experimental uncertainty.

The Leading Order (LO) term is defined to be the contribution from all diagrams containing only the fewest possible number of vertices to go from the desired initial to final state. Then LO+1 is dubbed Next-To Leading Order (NLO), LO+2 is Next-To, Next-To Leading Order (NNLO) and so on.

2.2.3 Renormalization

Since higher order corrections can always be introduced via. loops as seen in figure 2.3, other problems of infinities in the theory emerges [10]. Basically the problem arises because particles in a loop are off-shell, and therefore can have any four-momentum, but particle momentum at the vertices are conserved. When integration over all momentums is performed, unphysical divergences appear.

One specific technique to deal with this, is known as *regularization*, where an upper momentum cutoff scale (Λ_R) is introduced, above which all momenta is ignored (corresponding to a minimal distance). This has the immediate problem, that an extra, arbitrary cut-off parameter is introduced to the theory. This is of course not desirable, and the parameter is »absorbed« into the masses and couplings of the theory with the

process of *renormalization*. This makes m and g functions of Λ_R and the momentum scale at which they are measured (p). This means that:

$$\mathcal{M}(g, m) \mapsto \mathcal{M}(g(\Lambda_R, p), m(\Lambda_R, p)). \tag{2.10}$$

The effect of renormalization is, that the masses (m) and the couplings (g) can no longer be seen as *bare* parameters in experiment. Instead the distinction is made between the bare parameters of the Lagrangian, and the physical parameters that is measured by experiment. Taking the continuum limit, letting $\frac{1}{\Lambda_R} \to 0$, keeping the physical parameters at the observed values, would make the bare parameters diverge. This is not a problem, as long as the physical parameters remain finite. One important problem persists. To be able to make the bare parameters fit the physical parameters, the theory needs to possess enough bare parameters to begin with. For this to be realized, extra parameters may have to be added to the Lagrangian. If this number of extra parameters is finite, the theory is said to be *renormalizable*. Any valid theory of physics is demanded to be renormalizable, in addition to having a sensible particle and symmetry content. The electroweak SM is a renormalizable theory[4], but the λ_Z TGC, which will later be added in an effective Lagrangian framework, makes the theory non-renormalizable, making these considerations about regularization and renormalization very relevant for this study. This point will be addressed further in section 2.5.

2.3 Construction of The Standard Model

In the following, a more thorough treatment of the SM will be given, with emphasis on the *gauge structure* of the SM. The interactions of the SM are introduced by imposing gauge symmetry. Doing this, extra terms are picked up, which represents the gauge bosons. The TGCs emerges when introducing the electroweak force from invariance under the gauge group $U(1) \times SU(2)$, and is fixed by the introduction of masses to the emerging gauge fields using spontaneous symmetry breaking. This is, however, a quite cumbersome process, and in order to preserve readability, the simpler case of Quantum Electrodynamics (QED) shall be introduced first, as an exemplification of how a gauge symmetry is imposed by introducing covariant derivatives, and how this introduces (massless) gauge bosons. By a short explanation of non-Abelian gauge theories, the concepts used to introduce QED is generalized such that the electroweak theory can be considered. From the treatment of electroweak theory, the TGCs present in SM are introduced. This section is largely based on the introductions to the SM given in Ref. [9, 5].

2.3.1 Quantum Electrodynamics

The QFT for electromagnetic interactions, QED, is the simplest to deal with. It starts out by considering the Dirac Lagrangian (equation (2.6)), demanding that it should be

[4] As shown by t'Hooft and Veltman who received a Nobel Prize for this discovery.

invariant under local phase transformations[5]:

$$\psi \mapsto \psi' = \exp[i\alpha(x)]\psi, \tag{2.11}$$

where α is now explicitly a function of x. It is clear that an extra term will carry from the transformation of:

$$\partial_\mu \psi \mapsto \partial_\mu \psi' = \exp[i\alpha(x)]\partial_\mu \psi + \underbrace{i \exp[i\alpha(x)]\psi \partial_\mu \alpha(x)}_{\text{Extra term}}, \tag{2.12}$$

which will only be zero when α is constant. The trick is to replace the ordinary derivative with a covariant derivative (D_μ) carrying a term that cancels the extra term of equation (2.12):

$$\partial_\mu \mapsto D_\mu \equiv \partial_\mu - igA_\mu. \tag{2.13}$$

This implies the introduction of an extra field, A_μ, the gauge field, and g the coupling constant associated with it. In order to obtain local gauge symmetry, A_μ must transform like:

$$A_\mu(x) \mapsto A'_\mu(x) = A_\mu(x) + \frac{1}{g}\partial_\mu \alpha(x). \tag{2.14}$$

Using the commutator of the covariant derivative, the field strength tensor can be constructed, which turns out to be equivalent to the electromagnetic field strength tensor:

$$A_{\mu\nu} = F_{\mu\nu} = -\frac{1}{g}[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{2.15}$$

where $F_{\mu\nu}$ is the (in literature) standard letter for the electromagnetic field strength tensor.

Now the transformation of the problematic term in the Dirac Lagrangian works out nicely, but the Lagrangian also picks up the extra term:

$$\begin{aligned} \mathcal{L} &= i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi \\ &= \underbrace{\bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi}_{\text{Fermionic part}} + \underbrace{e\bar{\psi}\gamma^\mu A_\mu \psi}_{\text{Interaction}}. \end{aligned} \tag{2.16}$$

Since A_μ couples to the fermion field in the same way as the photon field does, the coupling is identified as the electric charge e. The new term is thus an interaction term, as indicated with the braces. We are now just a small step away from QED. The added term is an interaction term between the fermion field and the gauge field, but the gauge field itself is missing from the Lagrangian. One could guess that both a term for kinetic energy and a mass dependent potential energy term should be added, but in fact only the former is true. Adding a term dependent of a gauge boson mass would violate local gauge invariance, and can not be permitted. This does allow for a term involving the

[5] A stronger condition than global phase invariance as introduced in equation (2.4).

field strength tensor of electrodynamics, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, and the kinematic term of the Lagrangian can be added as the field strength tensor squared, such that:

$$\begin{aligned}\mathcal{L}_{QED} &= i\bar{\psi}\gamma^\mu D_\mu \psi - m\bar{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \underbrace{\bar{\psi}(i\gamma^\mu\partial_\mu)\psi}_{\text{Fermion kinematics}} - \underbrace{\bar{\psi}m\psi}_{\text{Fermion mass}} + \underbrace{e\bar{\psi}\gamma^\mu A_\mu\psi}_{\text{Interaction term}} - \underbrace{\frac{1}{4}F_{\mu\nu}F^{\mu\nu}}_{\text{Photon kinematics}} \quad .\end{aligned} \tag{2.17}$$

The absence of a gauge boson mass in the Lagrangian requires that the gauge boson of QED is massless – and the photon indeed is.

The theory of QED is often quoted as a beautiful theory, and the beauty really lies in the fact, that this interacting field theory can be constructed by simply requiring the free fermion to obey local gauge invariance. In the mathematical language of group theory, this is denoted invariance under the local unitary group of degree one, in short $U(1)$. This group can be represented by the complex numbers z and the multiplication operator, $U(1) = (z \in \mathbb{C} : |z| = 1, \otimes)$, where the condition that $|z| = 1$ gives the geometrical interpretation as rotations around a circle. As multiplication with complex numbers commutate, $U(1)$ is an Abelian group. This group theoretical picture becomes a more powerful tool when dealing with the theory of strong interactions and the unification of QED with the weak interactions.

2.3.2 Non-Abelian Gauge theories

As the symmetry of QED can be described in terms of group theory as a local $U(1)$ symmetry, the group theoretical picture is very useful when turning to introduce the strong and weak forces into the SM. The obvious generalization of $U(1) \to U(n)$ is to replace the complex numbers z with all $n \times n$ unitary matrices and the complex multiplication operator with matrix multiplication. An interesting subgroup of this unitary group, turns out to be the *special unitary group* $SU(n)$, where the further constraint is induced, that the matrices must have determinant 1. Matrix multiplication is not commutative, and the gauge theories resulting from imposing invariance under transformations of the $SU(n)$ gauge group are indeed called non-Abelian gauge theories[6].

As the $U(1)$ transformation was represented as exponentials of $i\alpha(x)$, the members of $SU(n)$ are also represented as a generalization of such an exponential. The object of interest then becomes the *generators* of the group, which are matrices M^j satisfying:

$$U = \exp(i\alpha_j M^j) \ , U \in SU(n), \tag{2.18}$$

where the exponential is the usual matrix exponential defined in terms of a power series.

For the special case of $SU(2)$, these generators are the *Pauli matrices* and for $SU(3)$ the *Gell-Mann matrices*[7].

[6] Or *Yang-Mills* theory in honor of Chen Ning Yang and Robert Mills.

[7] The Pauli -and Gell-Mann matrices does not constitute a unique choice for the representation of the generators. All sets of matrices that spans the space of solutions of equation (2.18) could be used.

The $SU(n)$ equivalent of requiring local $U(1)$ invariance of the Dirac Lagrangian, would be to require that the Lagrangian remains invariant as the *multiplet* of fields $\psi = (\psi_1, \psi_2, \ldots, \psi_n)$ undergoes the transformation:

$$\begin{aligned} \psi \mapsto \psi' &= \exp\left(i\alpha_j(x)M^j\right)\psi \\ &= \exp\left(\frac{1}{2}i\alpha_j(x)\sigma^j\right)\psi, \end{aligned} \tag{2.19}$$

where σ^j denotes the Pauli or Gell-Mann matrices (or the proper $n > 3$ generalizations thereof). This specific choice of representation introduces the factor $\frac{1}{2}$. Doing so, the same problem with extra terms, as in equation (2.12) appears. This is dealt with, by introducing a new covariant derivative, which is then reduced to the well known covariant derivative of QED from equation (2.13). The theory of strong interaction – Quantum Chromodynamics – is developed by requiring the Dirac Lagrangian (equation (2.6)) to be invariant under local $SU(3)$ transformations.

2.3.3 Electroweak unification

Following the above »recipe« for constructing a gauge theory by imposing invariance under a certain gauge group, the weak interaction can be constructed by demanding invariance under local $SU(2)$ transformations. Experimental evidence, however, suggests that the weak and electromagnetic forces should be joined into one force, the so-called electroweak force[8], where the relevant gauge group is now the combination of the two, $SU(2) \times U(1)$. This is related to the scale dependence of the couplings, briefly introduced along with renormalization in section 2.2.3, in the sense that when extrapolating the measured scale dependence of the couplings, the electromagnetic and weak couplings are seen to intersect at a high mass scale[9].

The choice of the $SU(2) \times U(1)$ gauge structure, and the following construction of the electroweak interaction is essential to the study of this thesis. As it will be shown, the TGCs of the SM emerges from (a) the choice of gauge group, and (b) the introduction of spontaneous symmetry breaking. In the framework presented, the TGC measurements, are really measurements of deviations from the $SU(2) \times U(1)$ gauge structure of the electroweak theory.

The choices made when constructing the electroweak theory, can seem a bit arbitrary, but are in fact firmly rooted in experimental evidence. The concept of chirality is used to distinguish so-called left-handed from right-handed particles. The concept can be explained most easily for massless particles, where the chirality equals the helicity, which is defined as the spin projection on the momentum axis. The »left« and »right« convention are depicted in figure 2.4. If the spin projection is in the same direction as the momentum,

[8]The unification of the electromagnetic and weak forces into one, was carried out by Steven Weinberg, Abdus Salam and Sheldon Glashow, who shared the 1979 Nobel Prize for the effort.

[9]This gives the idea that the *three* different forces (weak, strong and electromagnetic) could be unified into one superforce, at some mass scale. No current theory which unifies the three forces (so-called Grand Unified Theories) has been validated at an experiment.

the particle is right-handed (denoted with subscript R) and in opposite direction, the particle is left-handed (denoted with subscript L). If the particle is massive, this concept

Figure 2.4: Illustration of the concept of helicity, where the alignment of the spin projection on the momentum defines the helicity of a particle

needs to be generalized. Since it is always possible to Lorentz-boost into a reference frame where left becomes right and vice versa, chirality-handed particles are defined as ones that transforms like the corresponding helicity-handed particles under the transformations of the group of generalized Lorentz transformations, the Poincaré group.

Experimental evidence suggests that left-handed and right-handed particles interacts differently under the electroweak force. That is, the charged weak current couples only to left-handed fermions. Hence this degeneracy needs to be taken into account when constructing the electroweak model, and fields are introduced for both left-handed and right-handed fermions, which should now transform by the local $SU(2)_L \times U(1)_Y$ gauge transformation:

$$\begin{aligned} \psi_L \mapsto \psi'_L &= \exp\left[\frac{i(\alpha^j(x)\sigma_j + \beta(x)\tilde{Y})}{2}\right]\psi_L, \\ \psi_R \mapsto \psi'_R &= \exp\left[\frac{i\beta(x)\tilde{Y}}{2}\right]\psi_R, \end{aligned} \tag{2.20}$$

where σ_j are the Pauli matrices and $\tilde{Y}$ the (one) generator of $U(1)_Y$. Since invariance under a transformation results in a conserved quantity due to Noether's theorem, left-handed fermions carry two such conserved quantum numbers; weak isospin $I_3 = \pm\frac{1}{2}$ and Y the weak hypercharge; the *eigenvalue* of $\tilde{Y}$. Right-handed fermions only carry hypercharge. In equation (2.20) the choice is made to scale both σ_j and $\tilde{Y}$ with a factor of $\frac{1}{2}$, following the choice from Ref. [5]. This will affect the hypercharge quantum number, as it is scaled along with the generator.

The covariant derivative that renders the Dirac Lagrangian (equation (2.6)) invariant under the transformations of equation (2.20) must now carry two gauge fields along with two coupling constants. Following the steps from constructing QED, it is clear that the $SU(2)_L$ gauge field must carry one more index (i) than the $U(1)_Y$, in order to sum over the three generators. The obvious choice is:

$$\partial_\mu \mapsto D_\mu \equiv \partial_\mu + \frac{i}{2}\left[\underbrace{g_1\sigma_i W_\mu^i}_{SU(2)_L \text{ part}} + \underbrace{g_2\tilde{Y}B_\mu}_{U(1)_Y \text{ part}}\right]. \tag{2.21}$$

The field strength tensor of the B-field follows that of QED (equation (2.15)), while the field strength tensor(s) of the W^i-field(s) picks up a self-interaction term:

$$\begin{aligned} B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \\ W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + \underbrace{g_1\epsilon^i_{jk}W^j_\mu W^k_\nu}_{\text{Self-interaction term}} , \end{aligned} \tag{2.22}$$

where the non-Abelian structure is given by the Levi-Civita symbol ϵ_{ijk}. Inserting the kinematic contributions from the W and B fields results in a gauge invariant Lagrangian:

$$\begin{aligned} \mathcal{L}_{EW} = \underbrace{i\bar{\psi}^\lambda_L\gamma^\mu D_\mu\psi_{\lambda L}}_{\text{Left-handed fermions}} + \underbrace{i\bar{\psi}^\lambda_R\gamma^\mu D_\mu\psi_{\lambda L}}_{\text{Right-handed fermions}} \\ - \underbrace{\frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu}}_{\substack{W \text{ kinematics and} \\ \text{self-interaction}}} - \underbrace{\frac{1}{4}B_{\mu\nu}B^{\mu\nu}}_{B \text{ kinematics}} . \end{aligned} \tag{2.23}$$

Here the extra summation index λ has been introduced to sum over the left-handed fermion doublets and the right-handed fermion singlets (see table 2.3 for an overview).

The bilinear term of equation (2.23) is the central piece of this analysis, as it generates the TGCs. This term emerges directly from the choice of constructing a non-Abelian gauge theory, with the choice of symmetry group entering as the Levi-Civita symbol in equation (2.22). The term as it stands cannot, however, stand alone. As when the gauge field of QED were introduced, the gauge bosons cannot carry any mass, as this would violate the gauge symmetry just imposed. The problem is, that the physically observed weak gauge bosons actually do carry masses, as indicated in table 2.2. A possible solution to this problem is to introduce the concept of spontaneous symmetry breaking by the so-called *Higgs Mechanism* (see appendix A for a more thorough introduction). The TGCs will then be dependent of both (a) The imposed $SU(2)_L \times U(1)_Y$ gauge symmetry, and (b) The mechanism responsible for spontaneous symmetry breaking – in the SM this is the Higgs.

A feature of spontaneous symmetry breaking is that the introduced fields does not correspond to the physically observable gauge bosons. The physical fields are obtained as linear combinations of the W and B fields. The $W^\pm$ is thus a mix of the first two W states, and the Z^0 and γ states are a mix of the last W state and the single B state:

$$\begin{aligned} \begin{pmatrix} W^+_\mu \\ W^-_\mu \end{pmatrix} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}\begin{pmatrix} W^1_\mu \\ W^2_\mu \end{pmatrix}, \\ \begin{pmatrix} Z^0_\mu \\ A_\mu \end{pmatrix} &= \begin{pmatrix} \cos(\Theta_W) & -\sin(\Theta_W) \\ \sin(\Theta_W) & \cos(\Theta_W) \end{pmatrix}\begin{pmatrix} W^3_\mu \\ B^1_\mu \end{pmatrix}, \end{aligned} \tag{2.24}$$

			I	I_3	Q	Y
$\begin{pmatrix}\nu_e\\ e^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu^-\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau^-\end{pmatrix}_L$	$^1/_2$ $^1/_2$	$^1/_2$ $-\,^1/_2$	0 -1	$-\,^1/_2$ $-\,^1/_2$
$\begin{pmatrix}u\\ d'\end{pmatrix}_L$	$\begin{pmatrix}c\\ s'\end{pmatrix}_L$	$\begin{pmatrix}t\\ b'\end{pmatrix}_L$	$^1/_2$ $^1/_2$	$^1/_2$ $-\,^1/_2$	$^2/_3$ $-\,^1/_3$	$^1/_6$ $^1/_6$
e^-_R	μ^-_R	τ^-_R	0	0	-1	-1
u_R	c_R	t_R	0	0	$^2/_3$	$^2/_3$
d_R	s_R	b_R	0	0	$-\,^1/_3$	$-\,^1/_3$

Table 2.3: Table of fermion quantum numbers, all quarks and leptons. The d', s' and b' refers to mixed quark states, related to CP violation.

where Θ_W is the Weinberg angle; a shorthand notation for the coupling ratios:

$$\Theta_W = \arccos\left(\frac{g_1}{\sqrt{g_1^2+g_2^2}}\right) = \arcsin\left(\frac{g_2}{\sqrt{g_1^2+g_2^2}}\right). \tag{2.25}$$

The physical and well-known observables of the electromagnetic force, charge quantum number Q and coupling value e, are in the same way combinations of the couplings and quantum numbers of the W and B fields, known as the Gell-Mann-Nishijima relation:

$$e = \frac{g_1 g_2}{\sqrt{g_1^2+g_2^2}}\ ,\ Q = I_3 + Y. \tag{2.26}$$

Note that the hypercharge enters without a scaling of $\frac{1}{2}$, as it is already scaled in equation (2.20), and correspondingly the fermion hypercharges quoted in table 2.3 follows this normalization.

2.3.4 Triple gauge boson couplings in the Standard Model

The electroweak Lagrangian of equation (2.23) can be written out in terms of the physical fields introduced in equation (2.24). The bilinear term becomes [11] (with $W^{\mu\nu\pm} = \partial^\mu W^{\nu\pm} - \partial^\nu W^{\mu\pm}$ and $Z^{\mu\nu} = \partial^\mu Z^\nu - \partial^\nu Z^\mu$):

$$\begin{aligned}\frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu} &= ig_1[W^{\mu\nu-}W^+_\mu - W^{\mu\nu+}W^-_\mu][\cos(\Theta_W)Z_\nu + \sin(\Theta_W)A_\nu] \\ &\quad + \frac{ig_1}{2}[\cos(\Theta_W)Z^{\mu\nu} + \sin(\Theta_W)A^{\mu\nu}][W^-_\mu W^+_\nu - W^+_\mu W^-_\nu] + \mathcal{O}(g^2),\end{aligned} \tag{2.27}$$

expressing the triple gauge boson vertices allowed in the SM. Multiplying the terms together, it is clear, that the only two triple gauge boson vertices allowed, are W^+W^-Z and $W^+W^-\gamma$. It is seen, that the strength of these two vertices are the same as the coupling to the fermions.

The neglected terms of $\mathcal{O}(g^2)$ are the so-called quadruple gauge boson vertices; $\gamma\gamma W^+W^-$, ZZW^+W^-, γZW^+W^-, $W^+W^-W^+W^-$. The quadruple gauge boson vertices are not considered in this thesis, because their strength is of second order.

The triple gauge boson vertex part, however, can be rearranged to an expression where the WWZ and the $WW\gamma$ vertices are separated, but symmetric in the exchange of photons and Zs:

$$\begin{aligned}\mathcal{L}_{TGC} =& \frac{1}{4} W^i_{\mu\nu} W_i^{\mu\nu} - \mathcal{O}(g^2) \\ =& ig_{WW\gamma} \underbrace{\left[g_1^\gamma (W^-_{\mu\nu} W^{+\mu} A^\nu - W^+_{\mu\nu} W^{-\mu} A^\nu) + \kappa_\gamma W^-_\mu W^+_\nu A^{\mu\nu} \right]}_{WW\gamma \text{ part}} \\ &+ ig_{WWZ} \underbrace{\left[g_1^Z (W^-_{\mu\nu} W^{+\mu} Z^\nu - W^+_{\mu\nu} W^{-\mu} Z^\nu) + \kappa_Z W^-_\mu W^+_\nu Z^{\mu\nu} \right]}_{WWZ \text{ part}},\end{aligned} \tag{2.28}$$

where $g_{WW\gamma} = e$ and $g_{WWZ} = e\cot(\Theta_W)$.

As seen, the emergence of the TGCs resides both on the non-Abelian structure of the electroweak theory, but also very much on the Higgs mechanism. The two triple gauge boson vertices described in equation (2.28) is part of the SM, and has been experimentally confirmed both at LEP [12] and TEVATRON [13].

2.4 Effective Lagrangians

The effective Lagrangian formalism is a neat way to introduce New Physics (NP) to an existing model. If elements of NP is to be added to the SM, only knowledge of the low energy behavior of this new physics is necessary, since this will constitute the only observable effects, at present energies. The concept is similar to the familiar notion of a series expansion of a mathematical function, where the expansion is truncated after a couple of terms, but still constitutes a good enough expression to work with in a certain neighborhood of a chosen point. In the same way, effective Lagrangians are ordered in powers of $\frac{1}{\Lambda_{NP}}$, where Λ_{NP} is the energy scale of NP. Perhaps familiar to the reader is the effective four-fermion (Fermi) theory of weak interactions – this is exactly a low energy approximation of the interactions of the weak force.

Since we do not know what characteristics NP will exhibit, the SM triple gauge boson vertex of equation (2.28) is simply expanded into a generalized version, and parameterized into parts that obeys different rules of symmetry. The structure of the effective Lagrangian of the WWZ vertex[10] will be of the structure:

$$\begin{aligned}\mathcal{L}^{total}_{eff} =& \mathcal{L}_1 \text{ (Separately C -and P-conserving)} \\ &+ \mathcal{L}_2 \text{ (CP conserving, C and P violating)} + \mathcal{L}_3 \text{ (CP violating)}.\end{aligned} \tag{2.29}$$

The full expression for this, the most general effective Lagrangian for the WWZ vertex, is readily found in literature [11, 14, 15], and contains seven independent couplings ($g_1^Z, g_4^Z, g_5^Z, \kappa_Z, \tilde{\kappa}_Z, \lambda_Z$ and $\tilde{\lambda}_Z$), here ordered in three separate Lagrangians to maintain

[10] All arguments can be made similarly for the $WW\gamma$ vertex, and indeed literature often just denotes this vertex WWV, $V = \{Z, \gamma\}$.

the structure of equation (2.29):

$$\begin{aligned}
\frac{i\mathcal{L}_1}{g_{WWZ}} &= g_1^Z(W^-_{\mu\nu}W^{+\mu}Z^\nu - W^+_{\mu\nu}W^{-\mu}Z^\nu) + \kappa_Z W^-_\mu W^+_\nu Z^{\mu\nu} + \frac{\lambda_Z}{m_W^2}W^-_{\rho\mu}W^\mu_\nu Z^{\nu\rho},\\
\frac{i\mathcal{L}_2}{g_{WWZ}} &= ig_5^Z\epsilon_{\mu\nu\rho\sigma}\left[(\partial^\rho W^{-\mu})W^{+\nu} - W^{-\mu}(\partial^\rho W^{+\nu})\right]Z^\sigma,\\
\frac{i\mathcal{L}_3}{g_{WWZ}} &= ig_4^Z W^-_\mu W^+_\nu(\partial^\mu Z^\nu + \partial^\nu Z^\mu) - \frac{\tilde{\kappa}_Z}{2}W^-_\mu W^+_\nu\epsilon^{\mu\nu\rho\sigma}Z_{\rho\sigma} - \frac{\tilde{\lambda}_Z}{2m_W^2}W^-_{\rho\mu}W^{+\mu}_\nu\epsilon^{\nu\rho\alpha\beta}Z_{\alpha\beta}.
\end{aligned} \tag{2.30}$$

As seven independent couplings is a vast number of parameters to measure, a common choice is to measure only the couplings associated with operators obeying C and P symmetry separately ($\mathcal{L}_1$). The sub-division of the operators based on symmetry properties indicated in equation (2.29), can be established, by applying the well known C and P transformations of the W and Z fields [14], which were also indicated in section 2.1.1:

$$\begin{aligned}
C\text{-transformations: } & CW_\mu C^{-1} = -W_\mu^\dagger \ , \ CZ_\mu C^{-1} = -Z_\mu.\\
P\text{-transformations: } & PW_\mu(\vec{x},t)P^{-1} = W^\mu(-\vec{x},t) \ , \ PZ_\mu(\vec{x},t)P^{-1} = Z^\mu(-\vec{x},t).
\end{aligned} \tag{2.31}$$

Using these transformations, the symmetry properties of each of the terms can be checked. The C-invariance of the simplest term of $\mathcal{L}_1$; $\kappa_Z W^+_\mu W^-_\nu Z^{\mu\nu} = \kappa_Z W^\dagger_\mu W_\nu Z^{\mu\nu}$, is here shown as an example of such a check:

$$\begin{aligned}
C[W^\dagger_\mu W_\nu Z^{\mu\nu}]C^{-1} & \\
&= CW^\dagger_\mu W_\nu\partial^\mu Z^\nu C^{-1} - CW^\dagger_\mu W_\nu\partial^\nu Z^\mu C^{-1}\\
&= CW^\dagger_\mu C^{-1}CW_\nu C^{-1}C\partial^\mu Z^\nu C^{-1} - CW^\dagger_\mu C^{-1}CW_\nu C^{-1}C\partial^\nu Z^\mu C^{-1}\\
&= -W_\mu W^\dagger_\nu\partial^\mu Z^\nu + W^\dagger_\nu W_\mu\partial^\nu Z^\mu\\
&= W^\dagger_\mu W_\nu Z^{\mu\nu}.
\end{aligned} \tag{2.32}$$

Not only the symmetry properties of the operators associated with the couplings differ, but also their dimensionality. From $\mathcal{L}_1$ in equation (2.30), it is seen that the dimensionality of the operators associated with the couplings differ. As a Lagrangian[11] has dimension $(mass)^4$, the dimensionality of the individual terms of equation (2.30) are fixed. As the dimensionality of the operators associated with the λ_Z and $\tilde{\lambda}_Z$ are dimension $(mass)^6$, the desired dimensionality is restored by dividing with m_W^2.
The dimensionality of the operators associated with all seven couplings is:

$$\begin{aligned}
&\text{Dimension 4: } g_1^Z, g_4^Z, g_5^Z, \kappa_Z, \tilde{\kappa}_Z.\\
&\text{Dimension 6: } \lambda_Z, \tilde{\lambda}_Z.
\end{aligned} \tag{2.33}$$

[11]The action must be dimensionless to preserve Lorentz invariance, forcing the Lagrangian (density) to dimension $(mass)^4$.

With the choice of considering only $\mathcal{L}_1$, with operators obeying C and P symmetry separately, it is seen that the effective Lagrangian that will be the central theoretical object of this study, contains two TGCs associated with dimension 4 operators, both present in the SM with expectation value 1, and one TGC associated with dimension 6 operators, with a SM expectation of 0:

$$\frac{i\mathcal{L}_{eff}}{g_{WWZ}} = g_1^Z(W^-_{\mu\nu}W^{+\mu}Z^\nu - W^+_{\mu\nu}W^{-\mu}Z^\nu) + \kappa_Z W^-_\mu W^+_\nu Z^{\mu\nu} + \frac{\lambda_Z}{m_W^2}W^-_{\rho\mu}W^\mu_\nu Z^{\nu\rho}. \quad (2.34)$$

Since the aim of this thesis is to describe *deviations* from the SM, the derived quantities describing deviations from SM expectation will be used [16]: $\Delta g_1^Z = g_1^Z - 1$ and $\Delta\kappa_Z = \kappa_Z - 1$.

2.4.1 TGC contribution

To give a hint of the effects of TGCs on the differential cross section, the contribution at tree level to the matrix element for the process $q_{h_1}\bar{q}_{h_2} \to Z_{h_1}W^\pm_{h_4}$, where $h_1, ..., h_4$ are the particle helicities in the (see figure 2.4) can be calculated. This is known as the helicity amplitude, and is calculated in the interesting high energy limit with contributions from all seven couplings in Ref. [15]. The helicity amplitudes can be used to make predictions about the sensitivity of the three couplings to measurements on this vertex and the effect on differential distributions.

The relevant helicity amplitudes in the high energy limit are (in notation $\mathcal{M}_{h_3h_4}$) proportional to [15]:

$$\begin{aligned}
\mathcal{M}_{\pm 0} &\propto \frac{\sqrt{\hat{s}}}{m_W}\left(\Delta g_1^Z + \Delta\kappa_Z + \lambda_Z\right)(1 \mp \cos(\theta)),\\
\mathcal{M}_{0\pm} &\propto \frac{\sqrt{\hat{s}}}{m_W}\left(2\Delta g_1^Z + \lambda_Z\right)(1 \pm \cos(\theta)),\\
\mathcal{M}_{\pm\pm} &\propto \frac{\hat{s}}{m_W^2}\lambda_Z \sin(\theta),\\
\mathcal{M}_{00} &\propto \frac{\hat{s}}{m_W^2}\Delta g_1^Z \sin(\theta),
\end{aligned} \quad (2.35)$$

where θ is the center-of-mass scattering angles of the Z, and the $(0,0)$ amplitude is highly suppressed in contribution to the cross section [14].

First of all, equation (2.35) predicts a much stronger $\hat{s}$ sensitivity to variations in λ_Z than the two other couplings. This is already expected from the considerations about dimensionality in equation (2.33), as operators of a higher dimensionality naïvely has a stronger energy dependence. Since the $(0,0)$ contribution is also a factor of $\sqrt{\hat{s}}$ higher than the $(\pm,0)$ and $(0,\pm)$ contributions, the differential cross sections are expected to be relatively more sensitive to Δg_1^Z than κ_Z, but not as much as λ_Z, because of the aforementioned suppression of this contribution [14].

It is therefore expected that the limits extracted using the WWZ vertex will be tightest for λ_Z, and the least tight for $\Delta\kappa_Z$, with Δg_1^Z in the middle [17].

From equation (2.35) it is also possible to point in the direction of what kinematical observable(s) would be the most sensitive to variations in the couplings. As all the matrix elements are dependent on $\hat{s}$ and θ, an observable including this information is desired. The transverse momentum of the Z boson includes both, and does not suffer from the reconstruction inefficiencies that the total invariant mass of the event would, when real data analysis is undertaken.

2.5 Unitarity violation

The addition of the extra term in the Lagrangian in equation (2.34) compared to the electroweak Lagrangian of 2.28 leads to a non-renormalizable theory, in four dimensional space-time, since, for the theory to be renormalizable, only operators of dimension four or less may appear in the Lagrangian [18].

The violation of unitarity is clearly visible from the helicity amplitudes in equation (2.35), as the $(\pm,\pm)$ contribution involving the coupling λ_Z, associated with dimension six operators is allowed to increase unbounded from above. This will, for high energies, result in a violation of unitarity. In many studies of triple gauge couplings, this is dealt with by trying to »salvage« the theory by considering the couplings as form factors, in order to preserve unitarity. The idea is to introduce a regularization function, that dampens the high energy behavior; be it a simple cut-off at some energy scale or a more involved function of $\hat{s}$. A particular popular choice is the so-called n-pole form factor [15, 17, 19, 20]:

$$\Delta g_1^Z(\hat{s}) = \frac{\Delta g_1^Z}{(1+\hat{s}/\Lambda_{FF}^2)^n} \ , \ \Delta\kappa_Z(\hat{s}) = \frac{\Delta\kappa_Z}{(1+\hat{s}/\Lambda_{FF}^2)^n} \ , \ \lambda_Z(\hat{s}) = \frac{\lambda_Z}{(1+\hat{s}/\Lambda_{FF}^2)^n}, \quad (2.36)$$

where the numerators are the couplings with no applied form factor. It is clear that a low Λ_{FF} will suppress more than a high one, and correspondingly high values of n will suppress more than low ones. Using $n = 0$ will restore the couplings without form factors. In figure 2.5 the behavior of the $p_T(Z)$ and the (true) invariant mass of the WZ system ($M_{inv}(WZ)$) (simulated with PYTHIA 8 at $\sqrt{s} = 7$ TeV and rescaled to 5 fb^{-1} of integrated luminosity), are seen for three values of Λ_{FF}, $n = 2$, generated at non-SM TGC value of $\Delta g_1^Z = 0.1$ (colored) compared to SM (black). The Λ_{FF} suppresses the high energy tail, as desired, and for small values of Λ_{FF}, the effect of TGCs diminishes.

As n goes up, the suppression should be more outspoken. In figure 2.6, the same distributions are shown for a fixed $\Lambda_{FF} = 2$ TeV, and $n = \{2, 3, 4\}$. The expected behavior is clearly visible, as a higher value of n, forces the distribution towards SM.

Assuming only one TGC different from zero at a time, the unitarity limits for the

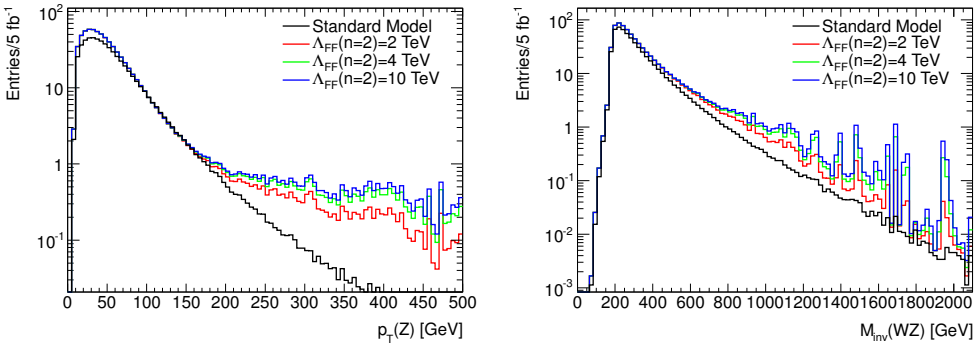


Figure 2.5: Distributions of $p_T(Z)$ (left) and $M_{inv}(WZ)$ (right) generated at $\Delta g_1^Z = 0.1$ for different values of Λ_{FF}, $n = 2$, compared to SM. It is seen that for small Λ_{FF}, the non-SM sample is forced towards the SM.

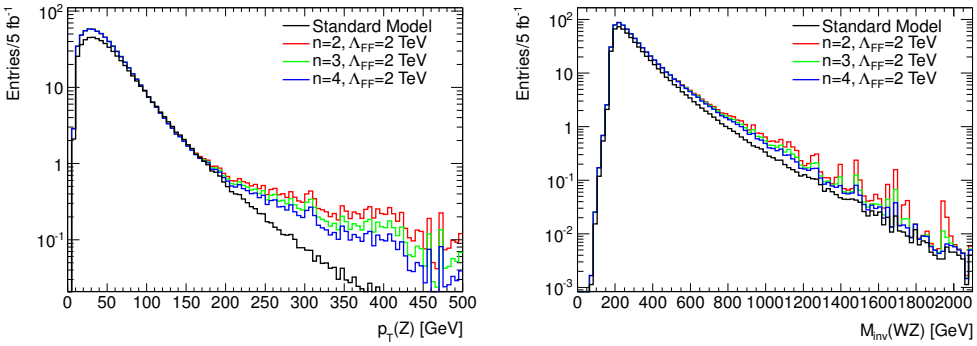


Figure 2.6: Distributions of $p_T(Z)$ (left) and $M_{inv}(WZ)$ (right) generated at $\Delta g_1^Z = 0.1$ for different values of n, Λ_{FF} =2 TeV, compared to SM. It is seen that as n increases, the non-SM sample is forced towards the SM.

n-pole form factors can, for $\Lambda_{FF} >> m_W, m_Z$, be calculated [17] [19]:

$$
\begin{aligned}
|\Delta g_1^Z| &\leq \frac{n^n}{(n-1)^{(n-1)}} \frac{0.84 \text{ TeV}^2}{\Lambda_{FF}^2} \\
|\Delta \kappa_Z| &\leq \frac{n^n}{(n-1)^{(n-1)}} \frac{0.83 \text{ TeV}^2}{\Lambda_{FF}^2} \\
|\lambda_Z| &\leq \frac{n^n}{(n-1)^{(n-1)}} \frac{0.52 \text{ TeV}^2}{\Lambda_{FF}^2}.
\end{aligned}
\tag{2.37}
$$

As the reason to introduce form factors is to avoid breaking of unitarity, Λ_{FF} is as such a regularization scale, related to – but not equal to – the scale of which NP becomes important. This is a neat way to make the effective Lagrangian valid for all energies, and not just for energies much smaller than a new physics scale.

This of course introduces a new parameter (Λ_{FF}) into the model. This is especially

a problem, since Λ_{FF} has to be adjusted every time the sensitivity of the experiment improves. Every time a new value of Λ_{FF} is used, it essentially corresponds to altering the model, with a resulting inability to compare new experimental limits with historic limits.

On a more conceptual note, the concept of applying such a regularization, is contradictory to the idea behind the effective Lagrangian formalism, which is to provide an effective, low energy approximation which *should break down* at the energy scale of NP. Introducing a regularization scheme at this point conceptually »promotes« the effective theory to a real theory. This was neither the intention nor desire in the first place, as the aim was to introduce a model independent framework for searching for deviations from the SM, not a full blown theory of NP.

In the end, the effect of applying a form factor scale is a reduction of sensitivity in the high energy regime, where NP manifesting itself in TGCs would appear if it exists.

The reduction of this sensitivity and the possibility of breaking unitarity, given SM physics, can be directly seen by comparing the limits in equation (2.37) with the estimated sensitivity of TGC fits[12] in pp collisions at $\sqrt{s} = 7$ TeV with data corresponding to 5 fb^{-1} of integrated luminosity.

This is done for $n = 2$ and $n = 4$ in figure 2.7 ($n = 2$ is the usual choice for analyzes of the WWZ vertex [17]). It is clearly seen that a low Λ_{FF} results in a large reduction of sensitivity, as also indicated by figure 2.5. As Λ_{FF} becomes larger than ≈ 3 TeV, no further suppression, and hence no reduction in sensitivity is seen. It is also seen that a larger n means a larger reduction of sensitivity, exactly as indicated in figure 2.6.

From the λ_Z behavior in figure 2.7 (lower), it is seen that variations in λ_Z within the uncertainty attainable with present ATLAS statistics and LHC energies can break with unitarity as expected.

As this thesis was written, the official ATLAS recommendation changed from providing TGC limits using a form factor scale of $\Lambda_{FF} = 2$ TeV with $n = 2$, to no form factor [20]. The decision was based on considerations about the nature of effective Lagrangians and ability to compare with past and future measurements, as summarized above.

[12] This is really a notion which is developed through the thesis. The reader can see this as *the largest allowed value of the TGC parameters given that no NP is observed.* For the reader who have already read the whole thesis, and returns here to get the full picture, a couple of things should be noted: The sensitivities quoted here, are half the asymmetric maximum likelihood CIs from binned fits of an Asimov SM data sample. The samples are made from PYTHIA 8 samples of 10^6 events generated at SM, and reweighted with matrix elements extracted from the BHO generator [21]. The sensitivity quoted here is thus worse than the sensitivities calculated with POWHEG or MC@NLO samples generated at high TGC values.

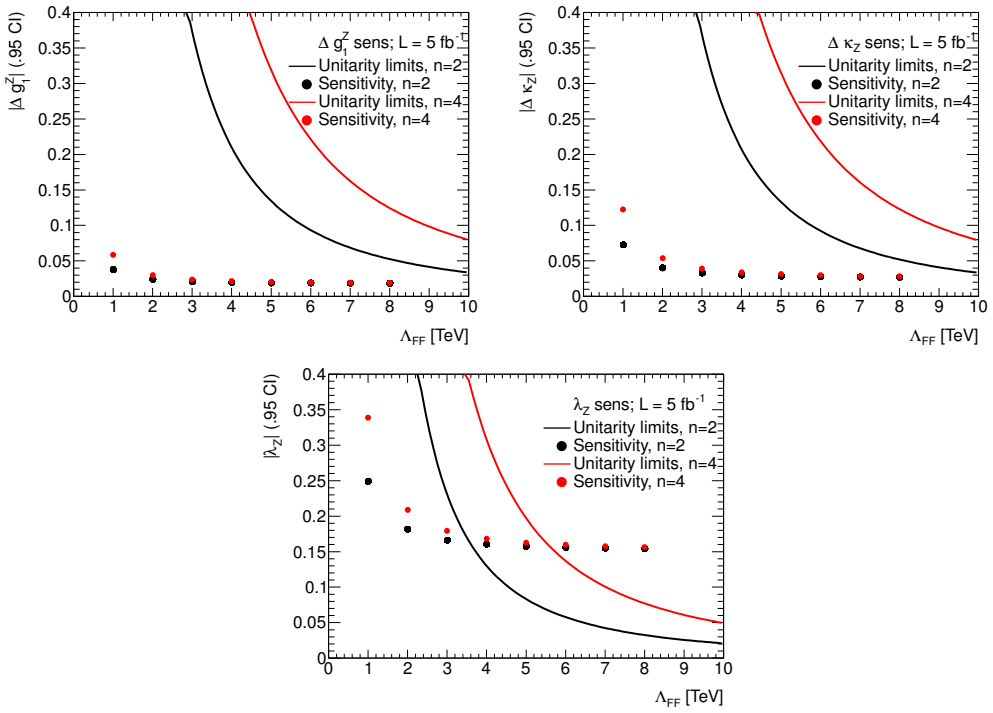


Figure 2.7: Expected sensitivity of TGC fits using simulated pp collisions (at SM) at $\sqrt{s} = 7$ TeV with data equivalent to 5 fb^{-1} of integrated luminosity, as a function of Λ_{FF} for $n = 2, 4$ compared to limits from unitarity. Calculated for Δg_1^Z (left, upper), κ_Z (right, upper) and λ_Z (lower).

CHAPTER 3

PHENOMENOLOGY AT HADRON COLLIDERS

There is a long way from the framework of effective Lagrangians introduced in the former chapter, to the final states observed in a detector. In order for particle physics experimentalists to make precise predictions about kinematics distributions, the use of so-called *event generators* is extremely widespread. The event generator is a software tool that, given $\mathcal{M}$ for the *hard process* and a *parton distribution function*, can compute the final states predicted by the theory. This is often referred to as the *truth*-level. As the detector measuring the final states is not perfect, simulation of the propagation of particles through the detector is also needed. That way the full information about what can actually be measured in a specific detector is obtained. Finally the event needs to be reconstructed, and analysis of the signal of interest can commence, after having properly removed any fake signal from other processes (often referred to as removal of »background«). This process is depicted in figure 3.1.

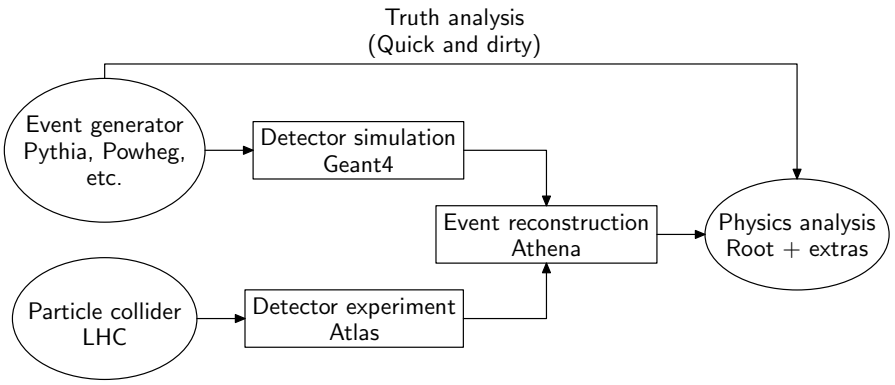


Figure 3.1: The event generation depicted as a flow chart with the various steps of simulation and analysis depicted.

The truth analysis might seem a bit to hasty when depicted as in figure 3.1. While

it is clear that the truth analysis is not sufficient to make direct predictions of what is observed in the detector, a truth analysis can be used to make first approximations, and perhaps most important to demonstrate and investigate a type of analysis on a simulation understood in full detail. Such a truth analysis, relevant to the scope of this thesis, is carried out in chapter 4.

In this chapter the process of producing Monte Carlo simulated events with TGC contributions will be introduced – both in a direct way with POWHEG [22] [23] and indirectly via matrix element reweighting of PYTHIA 8 [24] and POWHEG events. This event generation constitutes the upper left corner of the steps of simulation in figure 3.1.

The structure of the chapter is such, that the general framework of simulation of pp collisions is presented first. This is done by first considering the step from protons to its *parton* constituents; the parton distribution functions. Then the simulation of the hard process is considered, and finally the notions of showers, hadronization and the underlying event is introduced. This is done in order to give the reader an insight in the full event generation process, as this is important to understand how the matrix element reweighting of events to non-SM TGC values, and consequently important to understand the background of the truth analysis carried out in chapter 4.

After this, an introduction to the matrix element reweighting technique used to reweight samples to non-SM values of the TGCs is given. The concept of »reweighting« is explained, and a particularly useful basis for writing the TGC matrix elements is introduced. This technique is finally tested against the POWHEG event generator.

3.1 Proton-proton collisions

A (slightly simplified) view of a high energy physics event can, with inspiration from Ref. [24], be thought of as a time ordered evolution:

1. Two proton beams comes close to each other, each characterized by a parton distribution function, which defines their composition and the energy of their constituents.

2. One ingoing parton from each beam enters the hard process, with the possibility to radiate photons and gluons »on the way« – so-called initial state radiation.

3. The hard process produces a number of outgoing particles, which may radiate final state radiation.

4. Unstable particles are decayed, colored particles are hadronized.

5. Further semi-hard processes might happen between the other partons of the incoming protons.

6. The remnants of the protons may have structure and color charge, and can relate to the rest of the final state particles.

These steps are all carried out in the event generator(s), and the generator provides the user with a transcript of the kinematics of the event constituents (particles) and general event properties.

The items will be explained one by one in the following.

3.1.1 Parton distribution functions

The compositeness of the proton is modelled with so-called *parton distribution functions*. The parton distribution function (pdf) expresses the probability for a quark or a gluon to have a certain *momentum fraction* x of the total energy of the proton [25]. The pdf parametrize the probability as a function of momentum fraction of the i'th parton x_i, and the momentum transfer scale Q of the collision, denoted $f(x_i, Q^2)$. Since the hard process considered in this thesis is the production of WZ pairs, the energy scale of the hard process is the electroweak energy scale, which is much larger than the scale of QCD which is only a couple of hundred MeV. Correspondingly the timescale of the hard process is much smaller than the timescale of the QCD processes. Therefore the two types of interactions can be factorized in accordance with the *QCD factorization theorem* [26]:

$$\sigma(pp \to WZ) = \sum_{\text{quark flavors}} \int\int dx_1 dx_2 f_q(x_1, Q^2) f_{\bar{q}}(x_2, Q^2) \sigma(q\bar{q} \to WZ) \tag{3.1}$$

A key ingredient of equation (3.1) is of course the pdfs. Since they belong to the low-energy regime of QCD which cannot be calculated using perturbation theory (non-perturbative QCD), various pdf models must be fitted to known processes. The fits are extrapolated, and used for other studies. The determinations of pdfs with data is a whole subject in itself. The truth studies of this thesis relies on the CTEQ pdfs [27], which are also used among others when fitting data. The CTEQ6.6 pdf at a $Q^2 = (200)^2\text{GeV}^2$ (roughly the invariant mass scale of the hard process considered in this thesis) is depicted in figure 3.2. At the LHC unprecedented momentum scales will be explored, with values of Q^2 up to $\approx 10^8$ GeV^2.

3.1.2 Hard processes

The key part of a Monte Carlo event generator, is the generation of the *hard process*. The main hard process considered in this thesis, is the production of a WZ system from quarks, having the differential cross section:

$$d\sigma(q\bar{q} \to WZ) = \frac{1}{2\hat{s}} |\mathcal{M}(q\bar{q} \to WZ)|^2 \frac{d\cos\theta d\phi}{8(2\pi)^2}. \tag{3.2}$$

where ϕ and θ are the azimuthal and polar decay angles of the Z.

This constitutes a theoretical prediction of the distribution in question. The prediction is constructed by sampling all off phase space (here (ϕ, θ)). The differential cross section obtained from this procedure, is often referred to as the *event weight*.

A typical Monte Carlo generator technique for producing an event sample from the event weights, is the *hit-and-miss* technique [24]. This allows for a random event selection,

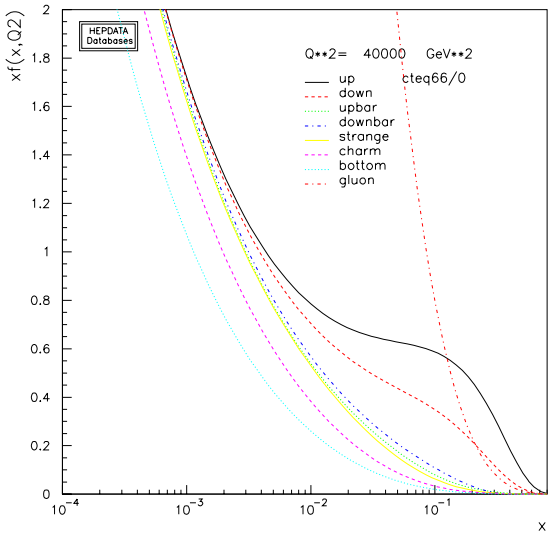


Figure 3.2: The CTEQ6.6 [27] parton distribution function at $Q^2 = (200)^2$ GeV^2 [28]

according to the event weights. The idea behind such sampling techniques are, that selecting a variable according to an n-dimensional distribution is equivalent to making a uniform selection in a corresponding $n+1$-dimensional volume.

The simplified case (which generalizes to higher dimensionality) is the selection of an observable x, according to the distribution $f(x)$. Given that a function $f(x) \leq f_{max}$ is known on the interval $x \in]x_{min}, x_{max}[$, where x is to be sampled from, the $n+1$-dimensional volume is an area. Then x is selected by the algorithm, taking as input a random distributed number $R \in]0, 1[$:

1. Select $x = Rf_{max}$.
2. Select $x_t = x_{min} + R(x_{max} - x_{min})$.
3. While $x > f(x_t)$, cycle to 1.

This algorithm is robust but time consuming, and improvements exist. The most obvious improvement is to introduce further constraints on $f(x)$, eg. $f(x) \leq g(x)$ where $g(x)$ is

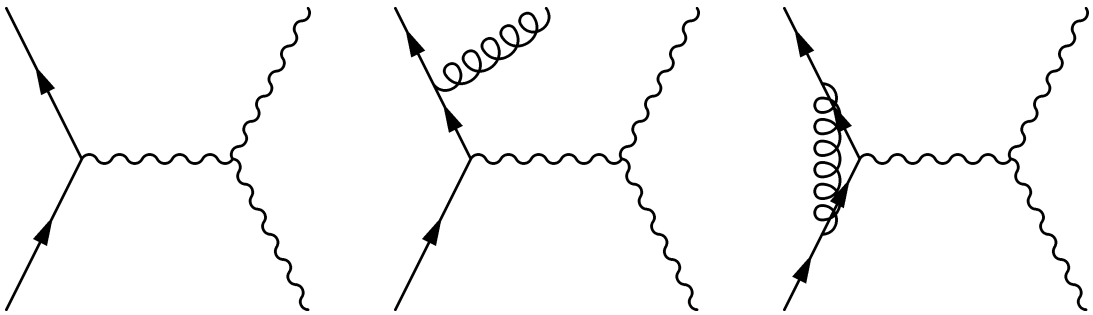


Figure 3.3: The process considered in this study with (left) LO contributions only, (middle) a NLO contribution as an ISR extra »leg« and (right) a contribution in form of a loop correction, which is not covered by ISR and FSR.

some well-known, well behaved function. This is known as *importance sampling*, and is for example used by PYTHIA 8.

The modeling of the hard process is really what is used to test the SM. NP models are all characterized by allowing some new hard process to take place, or modifying existing hard processes.

NLO with k-factor

Recently it is actually possible to calculate full NLO differential cross sections using NLO generators like POWHEG [22] or MC@NLO [29]. For this study, this is only done in full for the most central process considered, and not for all background processes. The reason being that a full NLO calculation is very CPU intensive. A reweighting using k-factors – which is simply the ratio of the NLO cross section to the LO cross section – is applied in such cases. This approach of course comes with its limitations. A differential cross section can be highly phase space dependent so it introduces a source of error to just average over phase space.

3.1.3 Initial -and Final State Radiation

The incoming and outgoing partons of the hard process, will radiate photons and gluons before and after the hard process takes place, as depicted in figure 3.3 (middle). This extra »leg« of the process is really a higher order correction, and should be dealt with as such, but since a full NLO treatment is very cumbersome, the concept of parton showers [26] is introduced to describe this specific NLO contribution.

The probability density functions for the individual parton to branch in the phase of final state radiation, or to have come from a specific parton when in the phase of initial state radiation [26] can be calculated, and in turn the initial -and final state radiation is simulated using a hit-and-miss technique or similar. The force of the showering approach, is that the probability calculated is independent of the hard process considered. Hence the same implementation of the showering can be reused for different processes.

One should emphasize that showering only accounts for the extra »leg« corrections depicted in figure 3.3 (middle), and not virtual corrections (loops) (figure 3.3 (right)).

3.1.4 Hadronization and fragmentation

Since the partons involved in the hard process can only be regarded as free particles on the timescales of the hard process, these are organized into colorless hadrons, obeying the principles of colour confinement. This is also within the non-pertubative QCD regime, and is approached by models. The Lund string model [24] [26] approaches this problem by forming »strings« between the partons, containing potential energy. As the q and $\bar{q}$ move away from each other, the strings may break to form new quark pairs. When the energy becomes too low for the quarks to escape each other, the hadronization and fragmentation stops.

The string model is incorporated in the PYTHIA 8 event generator. Other models of hadronization exist, most notably perhaps the cluster model, which is described in Ref. [26].

3.1.5 The underlying event

The underlying event consists of »the rest« – which is everything not related to the hard interaction, or the associated effects described above. This meaning primarily interactions between beam remnants and multiple parton interaction (MPI). The event generator PYTHIA 8 utilizes an MPI model, described in detail in Ref. [26].

The MPI model, and the other phenomenological models, contains a number of parameters which needs to be determined by an experiment. This is referred to as tuning of the event generator, and is also performed by the ATLAS experiment [30].

3.2 Simulating effects of triple gauge boson couplings

In section 2.4 it was described how the effect of the TGCs can be parametrized as extra terms in the effective Lagrangian formalism. As indicated in section 2.4.1, the matrix elements for the TGC contribution are calculable, and can be included directly in a Monte Carlo generator. This is done in several generators. For this thesis it is relevant to mention POWHEG in the POWHEG-BOX framework [23] and MC@NLO [29]. This method is not computationally efficient, as a full simulation is necessary for every small change in the coupling values. This could be acceptable for a truth analysis (see figure 3.1), but as steps of detector simulation and reconstruction must be done for each sample when analyzing real data, this inefficiency becomes severe. To deal with this inefficiency, a reweighting procedure where a sample is reweighted to other TGC values than it was produced at, is adopted. This will be presented in detail in the following section. The presentation takes the following order:

1. The idea of matrix element reweighting on an event by event basis is introduced.

2. The event weights are written in a basis such that the weights are made explicitly dependent by the TGCs.

3. The reweighting procedure is compared to direct generation using POWHEG.

In the course of introducing these methods, the effects of the TGCs on the relevant distributions will be discussed.

3.2.1 Reweighting procedure

Given a sample of Monte Carlo events generated at the SM – for example using PYTHIA 8, which does not include TGCs – the events can be *reweighted* to non-SM TGC values by applying to each event an event specific weight:

$$w = \frac{d\sigma_{TGC}}{d\sigma_{SM}}. \tag{3.3}$$

The $d\sigma_{TGC}$ can be calculated for an event, given values of the TGCs, full event information and the full matrix elements. Full matrix elements are extracted from the BHO event generator [21] [31], and the distributions $\frac{d\sigma_{TGC}}{dp_T(Z)}$ and $\frac{d\sigma_{TGC}}{dM_{inv}(Z)}$ is calculated varying one TGC at a time, as shown in figure 3.4. The SM sample used is generated with PYTHIA 8, set up to simulate pp collisions at $\sqrt{s} = 7$ TeV using the CTEQ6 LO parton distribution function [27].

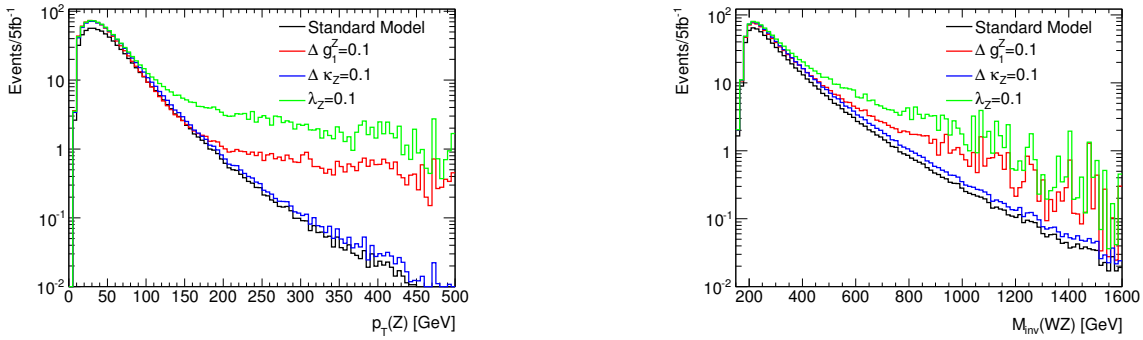


Figure 3.4: Reweighting effect on $p_T(Z)$ (left) and $M_{inv}(WZ)$ (right), where a PYTHIA 8 sample of 10^6 events is reweighted to 5 fb^{-1} and is reweighted to non-zero values of $\Delta g_1^Z = 0.1$, $\Delta\kappa_Z = 0.1$ and $\lambda_Z = 0.1$.

Effect on observables

From figure 3.4 it is seen that the predictions from section 2.4.1 are correct. The distributions are most sensitive to variations in λ_Z, less to Δg_1^Z and even less to $\Delta\kappa_Z$. The characteristic deviations appear for large values of $\hat{s}$ (the high $p_T(Z)$ and $M_{inv}(WZ)$ tail) as expected, and a by-eye assessment asserts that the $p_T(Z)$ distribution is more sensitive than the $M_{inv}(WZ)$. These notions are more visible from the ratios between

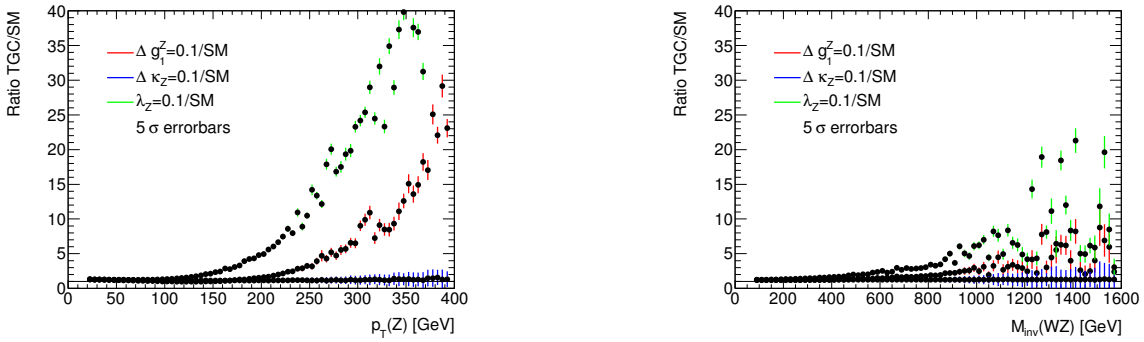


Figure 3.5: Reweighting effect on (left) $p_T(Z)$ and (right) $M_{inv}(WZ)$, where the TGC samples with $\Delta g_1^Z = 0.1$, $\Delta\kappa_Z = 0.1$ and $\lambda_Z = 0.1$ from figure 3.4 is divided by the corresponding SM sample. Error bars are the scaled and propagated bin by bin errors, all multiplied by five for graphical reasons.

Variable	$\Delta g_1^Z = 0.1$	$\Delta\kappa_Z = 0.1$	$\lambda_Z = 0.1$
$p_T(Z)$	729 ± 29	720 ± 25	896 ± 34
$M_{inv}(WZ)$	582 ± 20	559 ± 18	753 ± 28

Table 3.1: Significance of the $p_T(Z)$ and $M_{inv}(WZ)$ variables, as defined in equation (3.4). The quoted errors are the square sum of the weights in each bin, propagated through equation (3.4).

the TGC samples and the SM samples shown in figure 3.5.

This statement can be quantified by considering the cumulated bin-by-bin significance (minus under -and overflow bins):

$$S = \sum_{Bins} \frac{TGC_i - SM_i}{\sqrt{SM_i}}, \tag{3.4}$$

where TGC_i is the i'th bin of the TGC histograms and SM_i the i'th bin of the SM histogram. This value for each coupling, for the two variables is listed in table 3.1. The significance from $p_T(Z)$ is larger than the significance for $M_{inv}(WZ)$, and the significances are largest for deviations in λ_Z, and smallest for deviations in $\Delta\kappa_Z$.

Poorly populated regions of phase space

As seen in figure 3.4 (left) and the error bars of figure 3.5 (left), a feature of reweighting SM samples to non-SM values of the TGCs, is (relatively) poor statistics in the high $p_T(Z)$ tail of the distribution. As this is a poorly populated region of phase space to begin with, the reweighting enhances the effect of poor statistics here. Two techniques will be used to deal with this problem.

A simple way of dealing with poor statistics is to generate a very high number of events, and scale the sample down to the desired luminosity. This is known as an

Asimov data sample [32]. This is a commonly used technique, but deals only with an overall improvement of statistics. The specific problem of relatively poor statistics in the high $p_T(Z)$ region is not dealt with by this. The Asimov weight is given directly from $N_{evt} = \sigma \mathscr{L}_{int}$:

$$w_{asimov} = \frac{N_{gen}}{\sigma} \mathscr{L}_{int,desired}, \tag{3.5}$$

where σ is the cross section, N_{gen} denotes number of generated events and $\mathscr{L}_{int,desired}$ is the target integrated luminosity.

To improve statistics in the high $p_T(Z)$ region, it is necessary to generate relatively more events with high $p_T(Z)$. There are several means of doing so; for this study a method of generating samples at high TGC values and weighting them back to the SM is used. By generating the input events at high TGC values, it is ensured that the parts of phase space which are the most sensitive to variations in TGCs are filled. The idea is, that the weights defined in equation (3.3) can be used to reweight samples between two arbitrary TGC (TGC1 and TGC2) values if written as:

$$w_{TGC1 \to TGC2} = \frac{d\sigma_{TGC2}}{d\sigma_{TGC1}}, \tag{3.6}$$

which equation (3.3) is really a special case of. A sample generated at TGC1, can thus be reweighted to any TGC value – including SM, by letting TGC2 = SM. To quantify the statistical effect of such a reweighting back to the SM, the relative error of the integral of the high $p_T(Z)$ (250GeV $< p_T(Z) <$ 500GeV) tail is compared for three types of samples. All three types are considered at: (1) SM values of the TGCs, (2) $\Delta g_1^Z = 0.1$, (3) $\Delta\kappa_Z = 0.1$ and (4) $\lambda_Z = 0.1$.

The first type is a POWHEG sample, generated at the SM, similar to PYTHIA 8 sample shown in figure 3.4, reweighted to the desired coupling values. The second type is generated at TGC values $\Delta g_1^Z = \Delta\kappa_Z = \lambda_Z = 2.5$ and reweighted to desired values. The third type is samples generated directly at the desired values of the TGCs. The POWHEG setup is described in further detail in section 3.2.4.

The relative error of the four integrals for each type of sample is quoted in table 3.2. This reveals that a by choosing to reweight from a high TGC sample compared to reweighting from a SM sample can decrease this relative error by up to a factor of ≈ 4. The performance of the samples generated at nominal value (middle row) is still somewhat better than for the sample reweighted from high TGC values, for the two sensitive couplings. This is not surprising as the sample generated at high TGC values is generated with all three TGCs at the same value. One would therefore expect that the error is »evened out«, which row three also reveals is the case.

The quoted error on the error in table 3.2 is approximated as the error of the variance of a Gaussian [3]; here approximately the relative error divided by 10.

Examples of this procedure is carried out in section 3.2.4, where it is used to make a sanity test of the used matrix elements [21] by comparing them to the matrix elements included in POWHEG [23], [33].

Type	SM	$\Delta g_1^Z = 0.1$	$\Delta\kappa_Z = 0.1$	$\lambda_Z = 0.1$
Rew. from SM	$4.9 \pm 0.5\%$	$10 \pm 1\%$	$5.3 \pm 0.5\%$	$12 \pm 1\%$
Gen. at TGC	$4.0 \pm 0.4\%$	$1.9 \pm 0.2\%$	$4.0 \pm 0.4\%$	$1.2 \pm 0.1\%$
Rew. from high TGC	$3.5 \pm 0.3\%$	$3.6 \pm 0.4\%$	$3.5 \pm 0.4\%$	$3.2 \pm 0.3\%$

Table 3.2: The relative error of the integral of $p_T(Z)$ distributions given and SM sample and samples with one TGC at a time different from zero. All samples are generated with 100,000 events.

3.2.2 The F_{ij} basis

Using the reweighting technique it is possible to introduce TGCs to Monte Carlo samples from generators that does not support TGCs. Furthermore it is possible to apply the reweighting technique after the full detector simulation (if the necessary truth information is carried along), which is much more efficient than starting the whole chain of figure 3.1 once again. It is, however, still time consuming to generate new event weights from the TGC matrix elements every time a TGC value is to be changed a little bit. Especially since the aim of this thesis is to provide limits on the TGCs through a fitting routine which will need to reweight to many different TGC values, a more efficient approach is highly desirable. Such an approach is developed in the following section, by noting that $d\sigma_{TGC}$ can be written as a quadratic function of the couplings. Once the coefficients of this quadratic function is known, the exact dependence of the TGCs is specified. The simple case of one TGC is explained first, then generalized to more than one. A TGC in general will here, and henceforth, be denoted α.

One coupling

As the TGC terms are added to the effective Lagrangian linearly (equation (2.34)), the TGC matrix element can, with the contribution of one additional TGC, be written as:

$$|\mathcal{M}_{TGC}|^2 = |\mathcal{M}_0 + \alpha\mathcal{M}_1|^2 = \mathcal{M}_0^2 + 2\alpha\mathcal{M}_0\mathcal{M}_1 + \alpha^2\mathcal{M}_1^2, \tag{3.7}$$

The differential cross section for an event can thus be written as a quadratic function function of the couplings:

$$d\sigma_{TGC} = d\sigma_{SM} + \alpha \cdot F_1 + \alpha^2 \cdot F_2 = F_0 + \alpha \cdot F_1 + \alpha^2 \cdot F_2, \tag{3.8}$$

where $d\sigma_{SM} \equiv F_0$ is the contribution from the SM. Note that the conceptual difference between equation (3.7) and equation (3.8) is that the $d\sigma_{TGC}$ in equation (3.8) is really calculated using equation (3.7). This is done for each event, using full truth information of the event. The unknowns of equation (3.8) are the F-coefficients, and the central part of this method is to determine those.

The weight defined in equation (3.3) to reweight a SM sample to non-SM TGC values becomes:

$$w = \frac{d\sigma_{TGC}}{d\sigma_{SM}} = 1 + \frac{\alpha \cdot F_1 + \alpha^2 \cdot F_2}{F_0}. \tag{3.9}$$

The F-coefficients are found from equation (3.8). As this is an equation of three unknowns (F_0, F_1 and F_2), three equations are needed to solve for all three unknowns. Since $d\sigma_{TGC}$ can be computed for any value of α from equation (3.7), using the matrix elements extracted from the BHO event generator [21], choosing $\alpha = \{0, 1, -1\}$ gives a simple system of equations [34]:

$$\begin{aligned} d\sigma_1 &= F_{00} \\ d\sigma_2 &= F_0 + F_1 + F_2 \\ d\sigma_3 &= F_0 - F_1 + F_2. \end{aligned} \tag{3.10}$$

Or equivalently on matrix form:

$$\begin{pmatrix} d\sigma_1 \\ d\sigma_2 \\ d\sigma_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} F_0 \\ F_1 \\ F_2 \end{pmatrix}. \tag{3.11}$$

If the values of α are chosen such that the system of equations is linearly independent (which the equation system (3.10) is), the coefficients can be determined from simple matrix inversion. When the coefficients are determined it is possible to calculate a weight corresponding to any value of α by simply inserting in equation (3.9).

More than one coupling

The generalization of the above procedure to n couplings is conceptually straightforward. Some added complexity appears since cross terms between TGC terms emerge (here only written in entirety for two new TGCs for readability, but the extension to n is trivial):

$$\begin{aligned} |\mathcal{M}_{TGC}|^2 &= |\mathcal{M}_0 + \alpha_1\mathcal{M}_1 + \alpha_2\mathcal{M}_2|^2 \\ &= \mathcal{M}_0^2 + 2\alpha_1\mathcal{M}_0\mathcal{M}_1 + 2\alpha_2\mathcal{M}_0\mathcal{M}_2 + \alpha_1^2\mathcal{M}_1^2 + 2\alpha_1\alpha_2\mathcal{M}_1\mathcal{M}_2 + \alpha_2\mathcal{M}_2^2 \\ &= \begin{pmatrix} 1 & \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} \mathcal{M}_0^2 & \mathcal{M}_0\mathcal{M}_1 & \mathcal{M}_0\mathcal{M}_2 \\ \mathcal{M}_0\mathcal{M}_1 & \mathcal{M}_1^2 & \mathcal{M}_1\mathcal{M}_2 \\ \mathcal{M}_0\mathcal{M}_2 & \mathcal{M}_1\mathcal{M}_2 & \mathcal{M}_2^2 \end{pmatrix} \begin{pmatrix} 1 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} F_{00} & F_{01} & F_{02} \\ 0 & F_{11} & F_{12} \\ 0 & 0 & F_{22} \end{pmatrix} \begin{pmatrix} 1 \\ \alpha_1 \\ \alpha_2 \end{pmatrix}. \end{aligned} \tag{3.12}$$

The matrix of the F_{ij} coefficients is put on an upper triangular form, and the factors of 2 are absorbed into the F_{ij}. Note that two indices are now used for the F_{ij} and $d\sigma_{SM} = F_{00}$. This result reduces to the single coupling result of equation (3.8) by letting F_{ij} be a 2×2 matrix, and setting $F_{00} = F_0$, $F_{01} = F_1$ and $F_{11} = F_2$.

In a compact form (the equivalent of equation (3.8)), the differential cross section can be written as:

$$d\sigma_{TGC} = \sum_{i,j} \alpha_i \alpha_j F_{ij}. \tag{3.13}$$

Note that $\alpha_0 = 1$ is the SM contribution. The weight needed to reweight a sample to a non-SM TGC value is:

$$w = \frac{\sum_{i,j} \alpha_i \alpha_j F_{ij}}{F_{00}}. \tag{3.14}$$

It is clear from the the upper triangular structure of the matrix F_{ij} in equation (3.12), that $N_{coeffs} = \frac{n(n+1)}{2}$ coefficients are needed when considering n TGCs. As it was the case for one coupling, a system of N_{coeffs} linearly independent equations is constructed and put on matrix form; here the equivalent of equation (3.11) is written in a compact form:

$$d\sigma_i = A_{ij} F_j, \tag{3.15}$$

where F_j denotes a vector consisting of all the F_{ij} coefficients, $d\sigma_i$ a vector consisting of the differential cross sections calculated with the TGC input values given in the matrix A_{ij}. These values are – with inspiration from the choice in equation (3.11) – a suitable combination of 0, 1 and -1 that makes the system of equations linearly independent and thus A_{ij} invertible. The coefficients can now be obtained by matrix inversion from equation (3.15). This needs only to be done once for a given sample.

3.2.3 Reweighting on a histogram level

The use of the F_{ij} coefficients to do TGC reweighting is essentially an event by event reweighting, and is as such an unbinned technique, as presented in the former section. As a consequence of equation (3.13), the distribution of the observable X with contributions from TGCs included, can be written as the sum of all contributions:

$$\frac{d\sigma_{TGC}}{dX} = \sum_{i,j} \alpha_i \alpha_j \frac{dF_{ij}}{dX}. \tag{3.16}$$

Since distributions of observables are often subjected to some sort of binning in a histogram, it is worthwhile to note that the TGC reweighting also works on a histogram level. From equation (3.16) the value of the i'th bin of a histogram showing the distribution of the observable X, can be written as the sum of weights of all events in the bin. This makes it interesting to consider the distributions of the F_{ij} coefficients over X, as these histograms can simply be added, with TGC values as pre factors, to obtain a binned version of the $\frac{d\sigma_{TGC}}{dX}$ distribution.

Reducing the discussion from the general level, to the three specific couplings considered in this thesis, the behavior of the F_{ij} wrt. $p_T(Z)$ is interesting to study, as these are indeed the histograms used for reweighting this distribution in the forthcoming chapters. With three couplings, there are 10 coefficients to consider. All projections of the F_{ij} coefficients (apart from F_{00} which is SM, and thus shown on e.g. figure 3.4) on $p_T(Z)$ are shown in figure 3.6, with the naming scheme: $\Delta g_1^Z = \alpha_1$, $\Delta\kappa_Z = \alpha_2$ and $\lambda_Z = \alpha_3$, rescaled to an integrated luminosity of 5 fb^{-1}. It is seen that cross terms with SM (F_{0j}) can give negative values; i.e. destructive interference. This is also the reason why the distributions in figure 3.6 are not shown on a logarithmic scale.

As only one coupling (α_i) will be fitted at a time (with the two others fixed to zero), equation (3.16) describing the distributions of observables simplifies greatly to:

$$\frac{d\sigma_{TGCi}}{dX} = \frac{dF_{00}}{dX} + \alpha_i \frac{dF_{0i}}{dX} + \alpha_i^2 \frac{dF_{ii}}{dX}. \tag{3.17}$$

The non-SM contribution, varying only one TGC at a time, can thus be calculated by adding one histogram from the top row and the corresponding histogram from the bottom row (with appropriate pre factors) in figure 3.6, ignoring the cross terms between TGC terms completely.

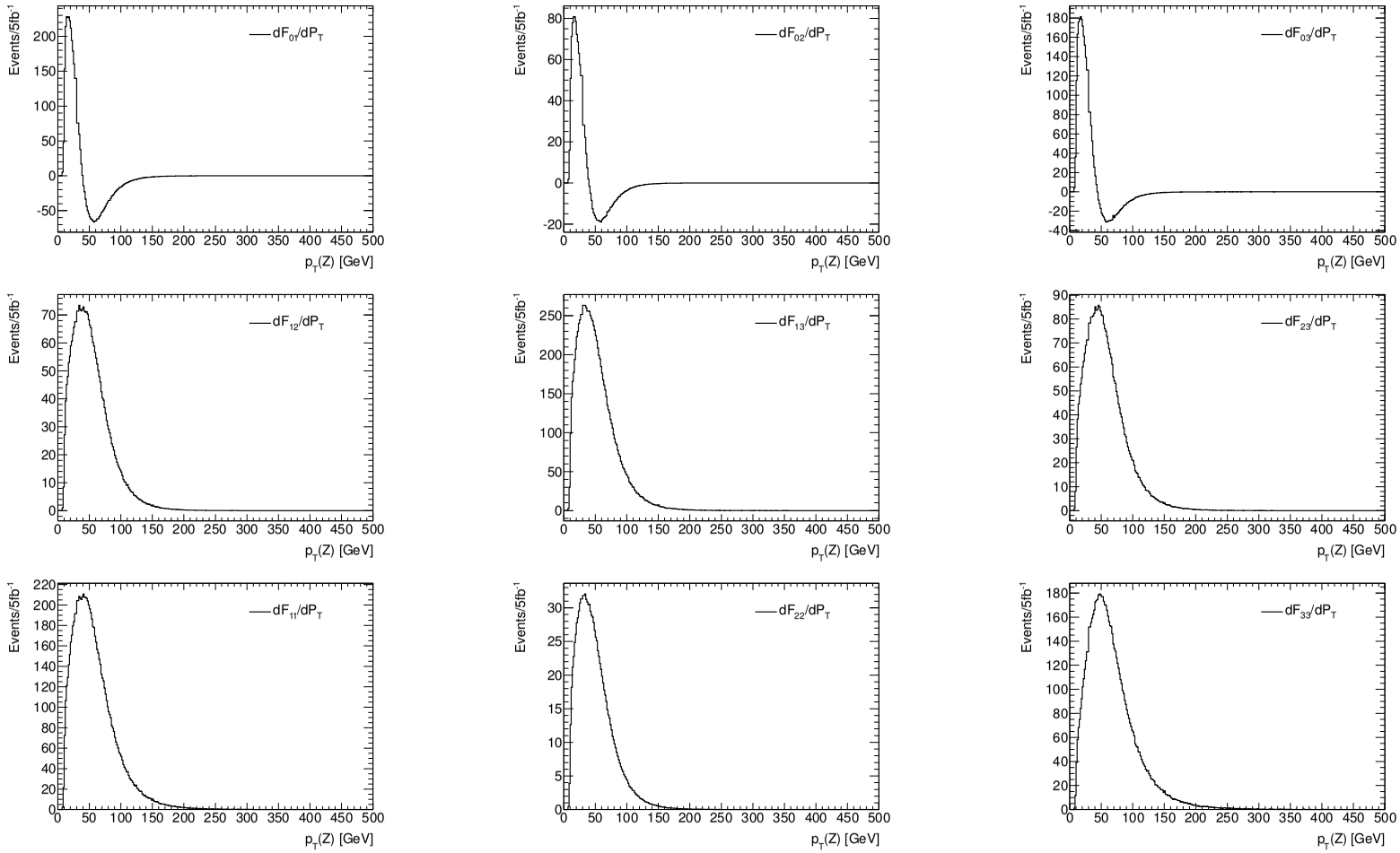


Figure 3.6: Projections of the F_{ij} coefficients ($i + j > 0$) on $p_T(Z)$, generated from a PYTHIA 8 sample, with the naming scheme: $\Delta g_1^Z = \alpha_1$, $\Delta\kappa_Z = \alpha_2$ and $\lambda_Z = \alpha_3$. All are rescaled to an integrated luminosity of 5 fb^{-1}.

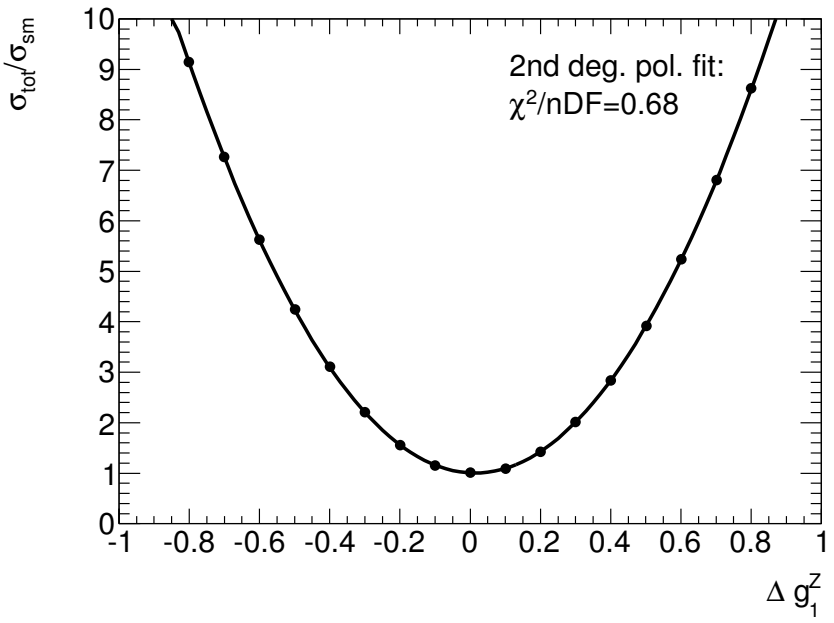


Figure 3.7: The parabolic behavior of the total cross sections normalized to SM cross section as a function of the coupling Δg_1^Z. The simulation is done by reweighting a PYTHIA 8 sample to non-zero values of Δg_1^Z calculating the total cross section for each step. The cross section ratio as a function of Δg_1^Z is fitted with a 2nd degree polynomial, and a $\chi^2/nDF = 0.68$ indicates a good fit.

As a sanity test of the F_{ij} coefficients, the quadratic behavior of the total cross section when varying only Δg_1^Z is shown in figure 3.7. The ratio between the total cross section given a non-zero value of Δg_1^Z and the SM cross section is fitted with a second degree polynomial, and a $\chi^2/nDF = 0.68$ reveals good agreement with the prediction.

3.2.4 Comparison with directly generated TGC events

A sanity check of the idea behind the reweighting method, the use of the F_{ij} coefficients and the used matrix elements [21] is performed by comparison to POWHEG-BOX [23], which uses another set of matrix elements [33], and generates events at a NLO with TGC contributions. POWHEG is set up to run in a mode such that comparison to the used matrix elements can be made. This means that virtual contributions are switched off. NLO contributions are as such limited to an extra leg, in accordance with the used matrix elements. Also only double resonant diagrams are included. As it is not possible to switch off the dipole form factor (Λ_{FF}), the comparison is done using $\Lambda_{FF} = 25$ TeV, $n = 2$, for both the POWHEG events and the reweighting procedure.

In figure 3.8 (left) a sample generated at the enormous TGC values of $\Delta g_1^Z = \Delta\kappa_Z = \lambda_Z = 2.5$ (green) is weighted back to SM value (red) and overlaid with the similar sample generated at SM level (black). Taking note of the different axis of the TGC sample

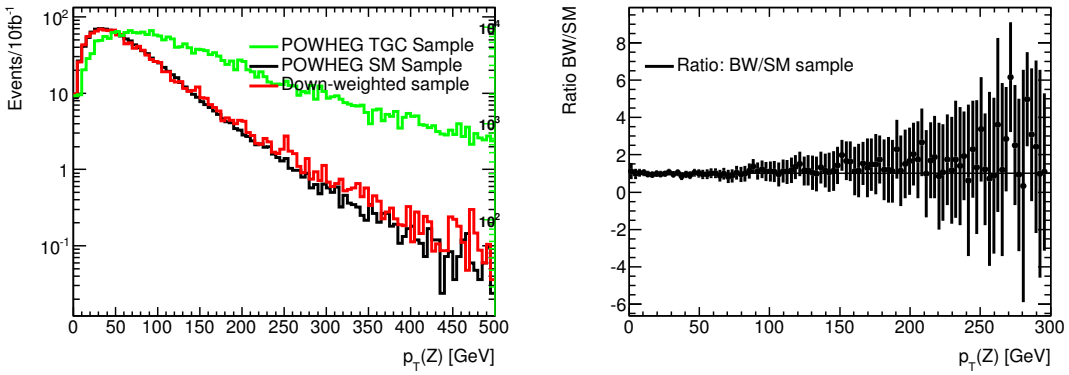


Figure 3.8: Reweighting TGC samples back to SM; (left) a sample, generated with POWHEG (green) at $\Delta g_1^Z = \Delta\kappa_Z = \lambda_Z = 2.5$ is reweighted down to SM level (red) and compared to a sample generated at SM level (black). The (right) bin-by-bin ratio of the back-weighted histogram to the SM histogram is shown, with the error bars indicating the propagated, scaled bin-by-bin error.

histogram in the weighted down histogram, one can appreciate that the sample is weighted down over several orders of magnitude. In figure 3.8 (right) the ratio between the two is shown. The error bars on the ratio represents the propagated bin-by-bin error of the two histograms. The ratio shows very good agreement between the two for $p_T(Z) < 250$GeV, and almost all within the error for 250GeV$< p_T(Z) < 300$GeV.

From the comparison in figure 3.8 the compatibility of the two sets of matrix elements and a proof of principle of the method is made. It is, however, also desirable to see how well the reweighting performs compared to direct generation at the level where at is used in this study; to generate samples at TGC values on the order $\mathcal{O}(0.1)$. Using the aforementioned POWHEG settings, samples at the same values as in figure 3.4 are generated directly. The $p_T(Z)$ distributions are shown in figure 3.9 (left), and the same effects of non-zero coupling values are seen, as also shown in figure 3.9 (left) where the ratios of the TGC samples to the SM sample are shown with errors – the error bars on this figure is multiplied by five for graphical reasons.

The ratio of each of the directly generated POWHEG samples to the corresponding reweighted PYTHIA 8 samples is shown in figure 3.10 with SM (upper left) shown as comparison. Agreement is shown in each of the four cases.

In order to give the reader an idea of the difference in CPU time using the reweighting procedure, a small use case example is here given. Here a SM sample already exists, and a user wishes to know what this sample would look like at non-SM values of the TGCs. Two $\Delta g_1^Z = 0.1$ samples of 100,000 events were created, one using POWHEG, full simulation (see figure 3.9), the other from an existing SM POWHEG sample, from which F_{ij} coefficients and weights were calculated. With a single Intel(R) Xeon(TM) 2.40 GHz processor on a Red Hat Linux system with 2 GB RAM available, the full simulation took more than six hours, whereas the reweighting procedure, including calculation of the

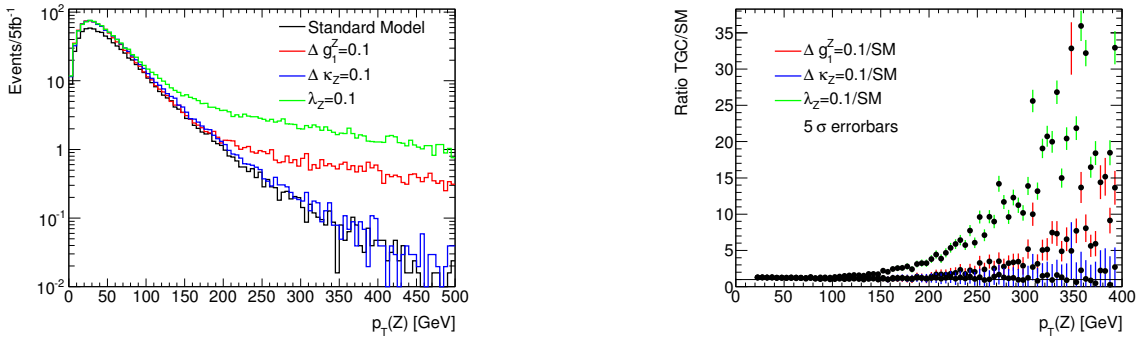


Figure 3.9: TGC effect from direct simulation by POWHEG on the (left) $p_T(Z)$ distributions and (right) the ratio of the TGC samples and SM. Samples are produced with $\Delta g_1^Z = 0.1$, $\Delta\kappa_Z = 0.1$ and $\lambda_Z = 0.1$ and scaled to an integrated luminosity of 5 fb^{-1}. The error bars indicates the propagated, scaled bin-by-bin error, and is multiplied by five for graphical reasons.

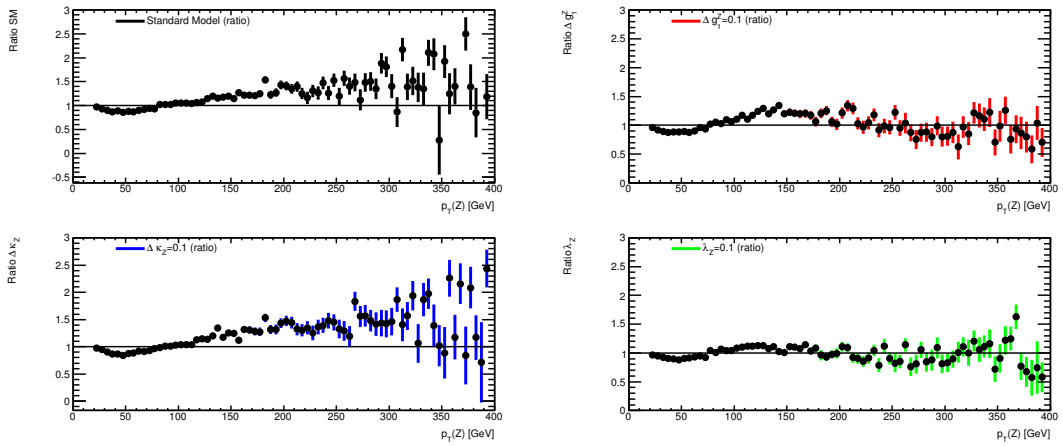


Figure 3.10: Ratio of directly generated POWHEG samples to reweighted PYTHIA 8 samples, (upper left) SM, (upper right) $\Delta g_1^Z = 0.1$, (lower left) $\Delta\kappa_Z = 0.1$ and (lower right) $\lambda_Z = 0.1$. The error bars indicates the propagated, scaled bin-by-bin error.

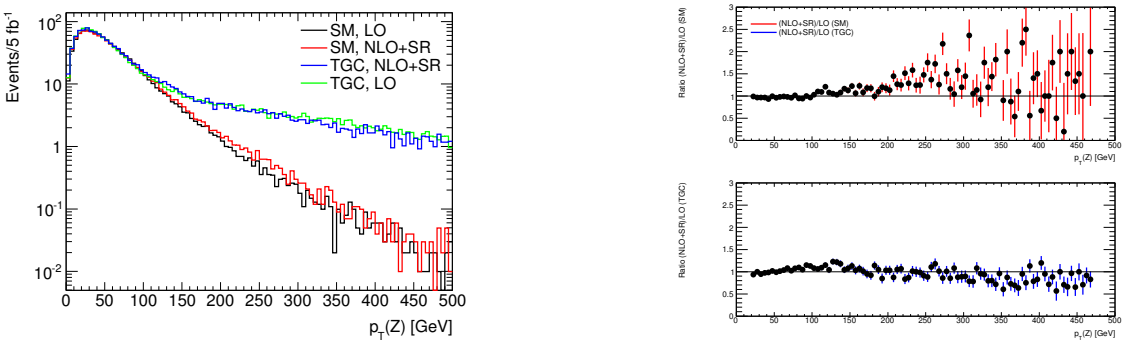


Figure 3.11: The $p_T(Z)$ distributions (left) for directly generated POWHEG samples at SM at LO (black), at SM with NLO+SR contributions (red), at $\Delta g_1^Z = \Delta\kappa_Z = \lambda_Z = 0.1$ at LO (green) and at the same TGC values with NLO+SR contributions (blue) all scaled to an integrated luminosity of 5fb^{-1}. The ratios (right) between the two SM distributions (red) and the two TGC distributions (blue) is also depicted.

weights, took less than ten minutes. The advantage of the reweighting procedure grows even larger, as the user might wish to obtain a sample at yet another value of the TGCs. Using the full simulation, the user can start all over again, but with the reweighting technique presented here, the F_{ij} coefficients are already calculated and saved, and the reweighted sample can be obtained in a matter of seconds if using the binned approach described in section 3.2.3, and a in less than a minute if a full event by event reweighting is used (still on the use case machine).

Next to leading order and single resonant diagrams

As the matrix elements used for the reweighting procedure does not support virtual corrections nor contributions from single resonant (SR) diagrams, these contributions were turned off in the POWHEG generator. Since nature contains both NLO effects and single resonant contributions, it must be tested if these effects contribute significantly – and especially if they resemble effects of TGCs. This effect is tested by comparing POWHEG samples generated at LO to samples generated with NLO+SR contributions (requiring the invariant Z mass $m_Z > 20$ GeV [23]). This comparison is done both at SM level and with directly simulated TGC samples with $\Delta g_1^Z = \Delta\kappa_Z = \lambda_Z = 0.1$. In figure 3.11 (left) the four corresponding $p_T(Z)$ distributions are shown, and a by-eye assessment reveals a small extra contribution from NLO+SR diagrams at SM, at about 250 GeV. The ratio of the distributions (figure 3.11 right) shows that the effect is negligible within the error, and that no effect is visible between the two TGC samples. It should however be noted that the small effect at about 250 GeV is on the same order as a $\Delta\kappa_Z = 0.1$ effect (see figure 3.9), but with the statistics available for this study, this effect is not expected to be visible. The effect of SR contributions by itself was tested by creating samples with NLO contributions and no SR contributions (not shown in the figure), but with the aforementioned cut on the Z mass, no visible effect was noted.

3.2.5 Afterburner vs. event generator

The method described here, using weights calculated *after* the event generation step, is referred to as an *afterburner* method. The obvious alternative to this would be to generate the weights along with the events. The two types of procedures for calculating weights are essentially the same, only the afterburner provides more flexibility, since the two steps are decoupled, and therefore possibly subject for modification. This of course does not come without a price – with added flexibility comes added complexity.

The event generator MC@NLO [29] currently supports generation of weights for selected TGCs. The three couplings considered in this thesis included. Since the MC@NLO samples are already run trough ATLAS detector simulation software, and the samples include virtual corrections, these samples are used when fitting real data in chapters 6 and 7. For the truth study in the next chapter, the POWHEG and PYTHIA 8, reweighted with the afterburner, are used.

The main differences between the two approaches are:

- The afterburner approach provides added flexibility in the sense, that new processes can be considered by adding a new matrix element to the generator. This cannot be done in the closed framework of MC@NLO.
- One is free to use whatever event generator desired when using the afterburner procedure.
- Consistency between F_{ij} coefficients and MC samples is ensured when using MC@NLO.
- Currently only MC@NLO supports virtual corrections on F_{ij} coefficients.

Overall, the choice between the two procedures can be summarized as the choice of flexibility in exchange for simplicity.

CHAPTER 4

LIMITS ON TGCS

In this chapter the statistical framework to estimate the values of the TGCs from data is developed. Different means of calculating a *confidence interval* (CI) at some *confidence level* (CL) is developed and discussed, and it is argued that the estimation of TGC CIs from data, should be done using a so-called *Neyman Construction* instead of an ordinary maximum likelihood fit. This is done through a truth level analysis of the coverage of the to types of limit setting techniques.

4.1 Historic limits

Limits on TGCs have been published dating back to the LEP experiments [35] up to TGC measurements in ATLAS and CMS. Some historic limits on the couplings λ_Z, $\Delta\kappa_Z$ and Δg_1^Z is given in table 4.1. The limits are obtained by varying the couplings one by one. The DØ limits [13] is extracted from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV with an integrated luminosity of 4.1 fb^{-1}. The CMS limits [36] are done using WW decays, with an integrated luminosity of 36 pb^{-1}. The ATLAS limits [37] are extracted with an integrated luminosity of 1 fb^{-1}.

The tightest published limits at the time of writing is set by the CDF collaboration at TeVatron [38], [39]. This is done using 7.1 fb^{-1} of integrated luminosity at $\sqrt{s} = 1.96$ TeV. The analysis is published February 2012, but is not expected to keep its status as the tightest limits, as no more data is recorded at the TeVatron, and the LHC experiments ATLAS and CMS keeps recording data at higher energies than the ones attained by the TeVatron. An ATLAS publication setting limits using 5 fb^{-1} taken at $\sqrt{s} = 7$ TeV is due before the summer conferences 2012.

As this thesis analyzes data equivalent to 827.87 pb^{-1} of integrated luminosity, recorded by the ATLAS detector at $\sqrt{s} = 7$ TeV using no form factor regularization, the ATLAS limits in the bottom row of table 4.1 are the ones relevant to compare to. All Monte Carlo studies – both at truth level in this chapter, and after detector simulation

	λ_Z	$\Delta\kappa_Z$	Δg_1^Z
CDF ($\Lambda_{FF} = 2$TeV)	$[-0.08, 0.10]$	$[-0.39, 0.90]$	$[-0.08, 0.20]$
DØ ($\Lambda_{FF} = 2$TeV)	$[-0.14, 0.15]$	$[-1.01, 1.27]$	$[-0.22, 0.30]$
CMS ($\Lambda_{FF} = \infty$)	$[-0.19, 0.19]$	–	$[-0.29, 0.31]$
ATLAS ($\Lambda_{FF} = 2$TeV)	$[-0.17, 0.17]$	$[-0.9, 1.1]$	$[-0.20, 0.30]$
ATLAS ($\Lambda_{FF} = \infty$)	$[-0.14, 0.14]$	$[-0.8, 1.0]$	$[-0.16, 0.24]$

Table 4.1: Historic limits on WWZ triple gauge couplings, at a 95% CL. Only the most recent limits are quoted. The tightest limit to this date, is set by the CDF experiment at the TeVatron [38] [39].

later – will be done for both 827.87 pb^{-1} and 5 fb^{-1}, such that the method developed in this thesis can be compared to the present ATLAS result, and the official result for 5 fb^{-1} when published. The reason the 5 fb^{-1} analysis is not published at the time of writing – and the reason this study only analyzes 827.87 pb^{-1} of data – is that the run parameters of the machine changed in 2011, making the combination of datasets a non-trivial exercise (see section 6.2 for further explanation).

Regarding comparison of the previously obtained results to this one, the use of a form factor in the CDF case, must be emphasized. Limits with no form factor set forth in this thesis, cannot be directly compared to limits where a form factor scale – with resulting loss of sensitivity – is used.

The limits given in table 4.1 are all obtained using a maximum likelihood method. In this chapter examples of different likelihood based limit setting techniques will be given in the scope of TGC measurements.

4.2 Maximum likelihood fit

The parameter estimation method of choice in this report is the maximum likelihood fit. The basics of parameter estimation techniques (in short – a fit) is assumed known to the reader. Otherwise one is referred to the introductory chapters of Ref. [40], from which a great deal of inspiration to the explanations of this chapter is also taken.

A distinction between two types of fits are made; (1) Fitting total cross section and (2) Fitting the shape of the $p_T(Z)$ distribution. These are two different types of information, and in order to investigate different features of a TGC fit in section 4.5, this turns out to be a sensible distinction. The two types of fits can be combined, as it will be explained in section 4.2.3, to make full use of the available information. The introduction of the different types of fits are supplemented with a demonstration, where pseudo-data (Monte Carlo generated at SM) corresponding to 827.87 pb^{-1} and 5 fb^{-1} of integrated luminosity is put in place of real data. This is done in order to give the reader an idea of the characteristic sizes of such limits, and to enable a comparison of the limit setting techniques based on MC truth results.

4.2.1 Total cross section fits

The simplest way of providing a TGC CI, is by fitting the total number of observed events in a sample. Since total cross section, including contribution from TGCs (σ_{TGC}), integrated luminosity ($\mathscr{L}_{int}$) and number of expected events (N_{exp}) are related by:

$$N_{exp} = \sigma_{TGC}\mathscr{L}_{int}, \tag{4.1}$$

this type of fit can be carried out as a part of a cross section measurement.

The number of events is expected to follow a well known Poisson distribution, where n is now an integer random variable, and $f(n_0|N_{exp})$ denotes the probability of observing n_0 events, given an expectation value (equation (4.1)) of N_{exp}. The Poisson probability density function is:

$$f(n|N_{exp}) = \frac{N_{exp}^n}{n!}\exp(-N_{exp}). \tag{4.2}$$

From the introduction of the F_{ij} coefficients in section 3.2.2, it is clear from equation (3.13) that N_{exp} can be written as a quadratic function of the TGCs with the sum of coefficients[1] as:

$$\begin{aligned}\sigma_{TGC} = \int d\sigma_{TGC} = \int \sum_{i,j}\alpha_i\alpha_j F_{ij} \approx \sum_{events}\sum_{i,j}\alpha_i\alpha_j F_{ij} \Rightarrow \\ N_{exp} = \mathscr{L}_{int}\sum_{i,j}\sum_{events}\alpha_i\alpha_j F_{ij}.\end{aligned} \tag{4.3}$$

The expression in equation (4.3) is the general result including all cross terms used to do multi parameter fits of TGCs. Note that the order of summation is changed in the last line, such that the histograms containing the F_{ij} coefficients can be summed over individually (section 3.2.3) As this study will vary TGCs one by one, the following is explained for one coupling, loosing full generality in exchange for the possibility to write expressions out in a less compact form. The TGCs will therefore carry an index i to illustrate that the equations are really a number of decoupled equations, one for each TGC considered. In that way equation (4.3) becomes:

$$N_{exp}(\alpha_i) = \mathscr{L}_{int}\left[\sum_{events}F_{00} + \alpha_i\sum_{events}F_{0i} + \alpha_i^2\sum_{events}F_{ii}\right]. \tag{4.4}$$

Given this simple polynomial dependence, the TGC parameters can therefore be estimated by constructing the log-likelihood function from equation (4.2), providing an observed value of total number of events ($n = N_{obs}$), and vary $N_{exp}(\alpha_i)$, by varying α_i one at a time. This will give a (number of) log-likelihood function(s)[2]:

$$\begin{aligned}\log(L_i(\alpha_i|N_{obs})) &= \log\left(f(\alpha_i|N_{obs})\right) = \log\left(f(\alpha_i|N_{obs})\right) \\ &= \text{const} + N_{obs}\log\left(N_{exp}(\alpha_i)\right) - N_{exp}(\alpha_i).\end{aligned} \tag{4.5}$$

[1] The sum over events is really an approximation of the integral, which is only valid when Monte Carlo statistics is sufficiently large.

[2] Throughout this thesis $\log(x)$ denotes the *natural logarithm* (base e) and not the base 10 logarithm. As no other logarithms are used, the word *natural* is omitted.

$\mathscr{L}_{int}$	$\Delta\kappa_Z$	Δg_1^Z	λ_Z
828 pb^{-1}	$[-0.59, 0.58]$	$[-0.14, 0.13]$	$[-0.12, 0.11]$
5 fb^{-1}	$[-0.37, 0.36]$	$[-0.08, 0.08]$	$[-0.07, 0.06]$

Table 4.2: Confidence intervals at a 95% CL from expectation fits of *total number of events*, as depicted in figure 4.1. The fits are done with SM expectation as input, having one coupling free at a time.

The maximum likelihood estimators are the value(s) of α_i that maximizes equation (4.5); written $\hat{\alpha}_i$. As equation (4.4) is a second degree polynomial, it is possible to have two values of α_i maximizing equation (4.5). Once the maximum likelihood estimator (the value of the fit) is found, the CI is found by comparing $L(\hat{\alpha}_i)$ to other hypotheses $L(\alpha_i \neq \hat{\alpha}_i)$. The comparison is done by constructing the log-likelihood ratio:

$$\mathfrak{r} = \log\left(\frac{L(\alpha_i)}{L(\hat{\alpha}_i)}\right) = \log(L(\alpha_i)) - \log(L(\hat{\alpha}_i)), \tag{4.6}$$

and allowing values of α_i in the CI, when $\mathfrak{r}$ is close to zero. To quantify this notion, it is assumed that $-2\mathfrak{r}$ goes asymptotically as a χ^2 distribution (which an ordinary Poisson distribution does [40]), and thus compare $\mathfrak{r}$ to a fractile of the χ^2 distribution with one degree of freedom (per free coupling). The 95 % fractile of the χ^2 distribution with one degree of freedom is 3.84, so to allow a certain value of α_i in the 95 % CI, it is demanded that[3]:

$$-\mathfrak{r}(\alpha_i, \hat{\alpha}_i) < 1.92. \tag{4.7}$$

It is noted that a change of sign has happened, and it is indeed the common practice to use the negative log-likelihood ratio, as well as the negative log-likelihood functions. The reason for this is rather pragmatic; the maximization of the log-likelihood is done using a numerical tool, which can only *minimize* a function. To fully demonstrate the procedure, and to provide the reader with a crude estimate of typical limits, the TGCs are fitted to SM »pseudo experiments« corresponding to integrated luminosities of 827.87 pb^{-1} and 5 fb^{-1}. The negative log-likelihood ratios are presented in figure 4.1, and the corresponding limits in table 4.2. The minimization package MINUIT [41] in a ROOT/RooFit framework [42] [43] have been utilized to produce the fits.

From these results it seen that the previous assertions (see section 2.4.1) about the predicted sensitivity are true. The tightest limits given SM are expected on λ_Z, less tight on Δg_1^Z and least tight on κ_Z.

4.2.2 Fitting shape

Fitting the shape of the $p_T(Z)$ distribution can be performed using a binned or an unbinned maximum likelihood fit. As described in section 2.4.1, the high $p_T(Z)$ tail is most sensitive

[3]More free couplings equals more degrees of freedom, and for fits of two TGCs with the last fixed, it should be demanded that $-\mathfrak{r}(\alpha_i, \hat{\alpha}_i) < 3.00$. For fits of all three TGCs at once, it should be demanded that $-\mathfrak{r}(\alpha_i, \hat{\alpha}_i) < 3.91$.

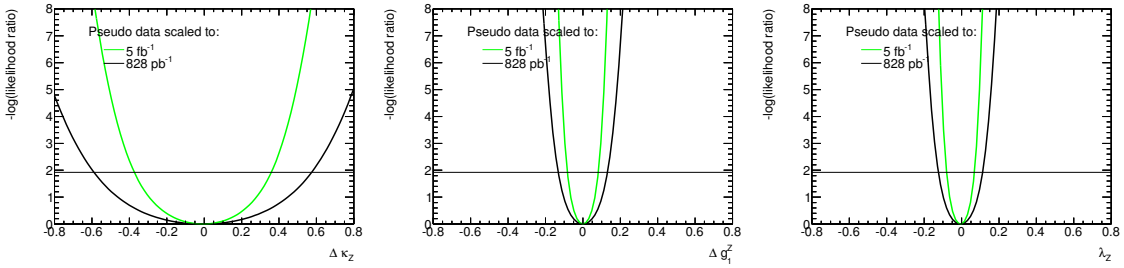


Figure 4.1: The negative log-likelihood ratios ($-\mathfrak{r}$) for the three WZ couplings. $\Delta\kappa_Z$ (left), Δg_1^Z (middle) and λ_Z (right), given SM expectation of *total number of events* as input to the fit. The expectation fits are done with inputs equivalent to integrated luminosities of (black) 827.87 pb^{-1} and (green) 5 fb^{-1}. The 95% CIs are calculated by requiring that $-\mathfrak{r} < 1.92$, as indicated with a horizontal line. The CIs are written out in table 4.2.

to couplings, and it will therefore be the high $p_T(Z)$ behavior that drives the sensitivity of the fits. The study uses POWHEG samples generated at $\Delta g_1^Z = \Delta\kappa_Z = \lambda_Z = 2.5$, to create a template for fitting, as described earlier.

Binned shape fits

The expected number of events N_{exp} is divided into N bins in $p_T(Z)$, such that a histogram with a certain number of entries in each bin is obtained. In other words a vector of N values: $\vec{n} = (n_1, \ldots, n_N)$. The content of each bin is now Poisson distributed [40] with expectation value $\nu_j(\alpha_i)$ dependent on the TGCs:

$$\nu_j(\alpha_i) = N_{exp} \int_{\text{bin j}} f(p_T(Z)|\alpha_i) dp_T(Z). \tag{4.8}$$

We can regard the bin numbers as type labels, and see the whole distribution as a joint experiment of N independent Poisson experiments, each with a Possion probability density function:

$$f_{pois}(n_j|\nu_j(\alpha_i)) = \frac{\nu_j(\alpha_i)^{n_j}}{n_j!} \exp(-\nu_j(\alpha_i)). \tag{4.9}$$

The joint, binned probability density function describing the shape of the $p_T(Z)$ distribution is thus:

$$f_{joint}(\vec{n}|\vec{\nu}) = \prod_{j=1}^{N} \frac{\nu_j^{n_j}}{n_j!} \exp(-\nu_j). \tag{4.10}$$

When constructing the corresponding log-likelihood function, the multiplication becomes a sum of terms each constructed like equation (4.5):

$$\begin{aligned}\log L(\alpha_i|\vec{n}) &= \text{const} + \sum_{j=1}^{N} n_j \log(\nu_j(\alpha_i)) - \nu_j(\alpha_i) \\ &= \text{const} - N_{exp}(\alpha_i) + \sum_{j=1}^{N} n_j \log(\nu_j(\alpha_i)),\end{aligned} \tag{4.11}$$

where it is used that $N_{exp} = \sum_{j=1}^{N} \nu_j$; the total number of expected events is equal to the sum of bin expectation values.

The TGC dependance of ν_j is fully parametrized by the equations (3.16) and (3.17). Varying only one TGC at a time, and using $p_T(Z)$ as the observable of choice, ν_j can be written as:

$$\begin{aligned}\nu_j(\alpha_i) = \int_{\text{bin j}} \frac{d\sigma_{TGC}}{dp_T(Z)} dp_T(Z) &\approx \sum_{\text{bin j}} \left[\frac{dF_{00}}{dp_T(Z)} + \alpha_i \frac{dF_{0i}}{dp_T(Z)} + \alpha_i^2 \frac{dF_{ii}}{dp_T(Z)} \right] \\ &= \sum_{\text{bin j}} \frac{dF_{00}}{dp_T(Z)} + \alpha_i \sum_{\text{bin j}} \frac{dF_{0i}}{dp_T(Z)} + \alpha_i^2 \sum_{\text{bin j}} \frac{dF_{ii}}{dp_T(Z)},\end{aligned} \tag{4.12}$$

where the summation in the last line have been moved, such that each histogram is summed over individually, as described in section 3.2.3. The approximation in the first line is again the one done when going from an integral over a true distribution to a summation over a histogram.

The fit is again done by minimizing the negative log-likelihood function, and limits established by requiring that the negative log-likelihood ratio $-\mathfrak{r}(\alpha_i, \hat{\alpha}_i) < 1.92$.

The generalization of the procedure to accommodate fits of M couplings simultaneously is straightforward, but results in more »dense« expressions in place of equation (4.11) and equation (4.12). Instead of looking at only one coupling α_i, the model is extended to look at the vector containing all M couplings: $\vec{\alpha} = (\alpha_0, \ldots, \alpha_M)$, where $\alpha_0 = 1$; the SM contribution. The log-likelihood function becomes:

$$\log L(\vec{\alpha}|\vec{n}) = \text{const} - N_{exp}(\vec{\alpha}) + \sum_{j=1}^{N} n_j \log(\nu_j(\vec{\alpha})). \tag{4.13}$$

The total expected number of events N_{exp} is given by equation (4.3), and the expression for ν_j in equation (4.12) becomes:

$$\nu_j(\vec{\alpha}) = \int_{\text{bin j}} \sum_{i,k=0}^{M} \alpha_i \alpha_k \frac{dF_{ik}}{dp_T(Z)} \approx \sum_{i,k=0}^{M} \sum_{\text{bin j}} \alpha_i \alpha_k \frac{dF_{ik}}{dp_T(Z)}. \tag{4.14}$$

Fits of the TGCs to SM pseudo experiments corresponding to 827.87 pb^{-1} and 5 fb^{-1} of data, using a binning of $p_T(Z) = \{15, 75, 150, 200, 1500\}$. Negative log-likelihood ratios are shown in figure 4.2 and corresponding limits in table 4.3.

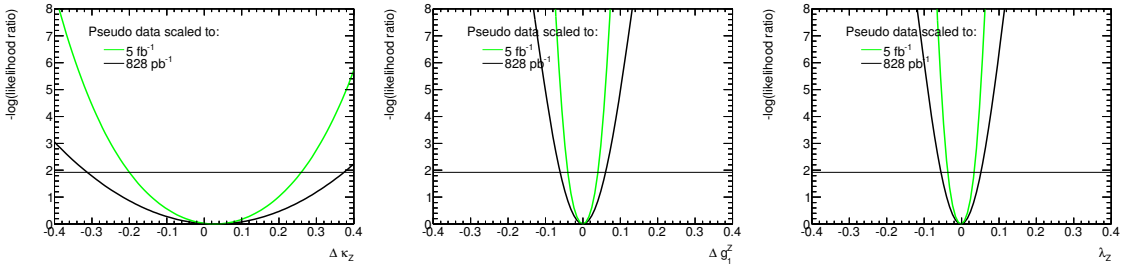


Figure 4.2: The negative log-likelihood ratios ($-\mathfrak{r}$) for the three WZ couplings. $\Delta\kappa_Z$ (left), Δg_1^Z (middle) and λ_Z (right), given SM expectation of the $p_T(Z)$ distribution, using a *binned fit*. The expectation fits are done with inputs equivalent to integrated luminosities of (black) 827.87 pb^{-1} and (green) 5 fb^{-1}. The 95% CIs are calculated by requiring that $-\mathfrak{r} < 1.92$, as indicated with a horizontal line. The CIs are written out in table 4.3.

$\mathscr{L}_{int}$	$\Delta\kappa_Z$	Δg_1^Z	λ_Z
828 pb^{-1}	$[-0.31, 0.37]$	$[-0.06, 0.06]$	$[-0.05, 0.05]$
5 fb^{-1}	$[-0.184, 0.256]$	$[-0.035, 0.035]$	$[-0.031, 0.032]$

Table 4.3: Confidence intervals at a 95% CL from binned shape expectation fits of the $p_T(Z)$ distribution, as depicted in figure 4.2. The fits are done with SM expectation as input, having one coupling free at a time.

Unbinned shape fits

The construction of an unbinned shape fit is, in a way, conceptually more simple than the binned fits. Here it is not necessary to construct a joint distribution, but only to focus on the probability distribution of $p_T(Z)$ given coupling values: $f(p_T|\vec{\alpha})$.

The method behind the construction of an unbinned fit of the TGCs is given here, and MC expectation corresponding to the same luminosities as the binned shape fit and the cross section fit is calculated. The unbinned fit is constructed as a proof of principle of the procedure. To this author's knowledge, unbinned maximum likelihood fits is not used for TGC measurements at any contemporary high energy experiments, even though the TGC analysis is notorious for poor statistics. Unbinned fits have however been employed earlier, notably in ALEPH [11]. Unbinned fits of TGCs could therefore be a possible path to follow for future TGC estimations at hadron colliders1.

The idea is, that for each given value of $p_T(Z)$, a probability density for seeing an event is calculated. Popularly speaking, letting $N \to \infty$ in the binned fit described above. This probability density is simply the differential cross section from equation (3.16):

$$\frac{d\sigma_{TGC}}{dp_T(Z)} = \sum_{i,j=0}^{M} \alpha_i \alpha_j \frac{dF_{ij}}{dp_T(Z)}. \tag{4.15}$$

After adding the histograms containing the relevant F_{ij} projections with pre factors, the

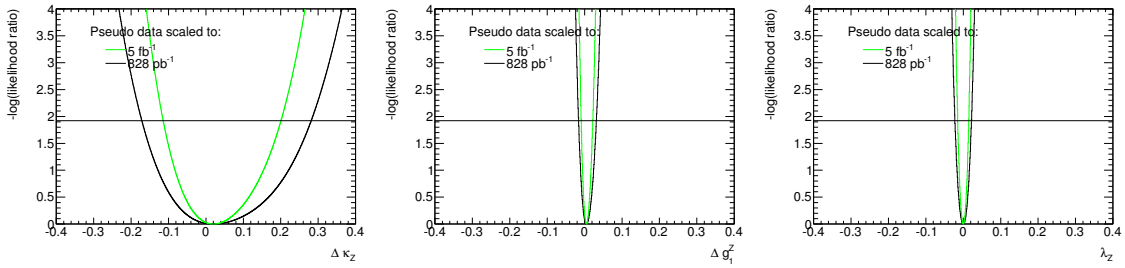


Figure 4.3: The negative log-likelihood ratios ($-\mathfrak{r}$) for the three WZ couplings. $\Delta\kappa_Z$ (left), Δg_1^Z (middle) and λ_Z (right), given SM expectation of the $p_T(Z)$ distribution, using a *unbinned fit*. The expectation fits are done with Asimov samples equivalent to integrated luminosities of (black) 827.87 pb^{-1} and (green) 5 fb^{-1}. The 95% CIs are calculated by requiring that $-\mathfrak{r} < 1.92$, as indicated with a horizontal line. The CIs are written out in table 4.4.

resulting histogram is normalized, and used as a probability density function (one can reconsult figure 3.6 to see all F_{ij} projected on $p_T(Z)$):

$$f(p_T(Z)|\vec{\alpha}) = \frac{1}{N_{exp}(\vec{\alpha})} \sum_{i,j=0}^{M} \alpha_i \alpha_j \frac{dF_{ij}}{dp_T(Z)}. \tag{4.16}$$

The log-likelihood function is now calculated as the sum of all individual log-likelihoods obtained given observed values of $p_T(Z)$.

As this implementation of unbinned fits uses the F_{ij} histograms, it is not fully unbinned in Monte Carlo; only in data, as no smoothing algorithms are used. As the number of generated Monte Carlo events approaches infinity, the unbinned fit is likewise approached.

Fits of the TGCs to SM pseudo experiments corresponding to 827.87 pb^{-1} and 5 fb^{-1} of data, using unbinned shape fits is performed to compare with the other types of fits. Negative log-likelihood ratios are shown in figure 4.3 and corresponding limits in table 4.4.

It is seen that the unbinned fits have an overall better performance when compared to binned fits, which is an indication that the unbinned maximum likelihood fit is a good candidate for a TGC estimation technique for future studies. It should however be stressed that the performance test done here, uses an Asimov data sample with events out to $p_T(Z) = 1500$GeV in place of real data, albeit with a very small Asimov weight. In real data this will not be the case. The last observed event will denote a »cut-off« for the fit, and the information from the *non-observation* of events at a higher $p_T(Z)$ will not be used in the fit. A possible solution to this issue in given in the next section.

4.2.3 Extended fits

The unbinned shape fit and the total number of events fit, can be combined into a so-called extended fit. As this combines the information from both types of fit, this should in

$\mathscr{L}_{int}$	$\Delta\kappa_Z$	Δg_1^Z	λ_Z
828 pb^{-1}	$[-0.18, 0.29]$	$[-0.02, 0.02]$	$[-0.02, 0.02]$
5 fb^{-1}	$[-0.094, 0.198]$	$[-0.010, 0.018]$	$[-0.009, 0.014]$

Table 4.4: Confidence intervals at a 95% CL from unbinned shape expectation fits of the $p_T(Z)$ distribution, as depicted in figure 4.3. The fits are done with SM expectation as input, having one coupling free at a time.

principle provide tighter limits. As the total likelihood function is then the product of the Poisson likelihood used to fit number of events and the likelihood function used for shape fits, the log-likelihood is the sum of the log-likelihood of the two [40]. For this truth study extended information (the expected number of events at SM) was added to the unbinned fits of table 4.4. No improvement of expected sensitivity was observed.

In a similar manner the information from the non-observation of events above the last data entry is used. This is accounted for in the binned fit via. the overflow bin, but by construction not accounted for in an unbinned fit. A constructed »overflow bin« can be employed, understood as a Poisson distribution with an expectation value given as the integral over the expected number of events above a certain threshold; effectively adding a Poisson overflow bin to the fit.

4.3 Systematics using profile likelihood

An advantage of using maximum likelihood estimators is that evaluation of systematic uncertainties is fairly easily incorporated with the use of *profile likelihood* [3, 32, 44]. The formal construction of a profile likelihood and the corresponding profile likelihood ratio is an extension of the construction of the ordinary likelihood ratio in equation (4.6). As the likelihood now depends both on TGCs ($\vec{\alpha}$) and parameters describing systematics ($\vec{\mu}$, often called *nuisance parameters*), the log-likelihood ratio is now $\mathfrak{r} = \mathfrak{r}(\vec{\alpha}, \vec{\mu})$. The explicit $\vec{\mu}$ dependence is removed by the construction of:

$$L_p(\alpha) = L(\vec{\alpha}, \hat{\hat{\vec{\mu}}}(\vec{\alpha})), \tag{4.17}$$

where $\hat{\hat{\vec{\mu}}}(\vec{\alpha})$ is given by the $\vec{\mu}$ that maximizes the likelihood for fixed $\vec{\alpha}$ [3]. The expression in equation (4.17) is the definition of the profile likelihood, and following this, the profile log-likelihood ratio is constructed analogous to equation (4.6):

$$\mathfrak{r}_p = \log\left(\frac{L_p(\vec{\alpha})}{L(\hat{\vec{\alpha}}, \hat{\vec{\mu}})}\right). \tag{4.18}$$

The ratio $\mathfrak{r}_p$ is then used in place of $\mathfrak{r}$.

The practical evaluation of systematic errors using profile likelihood is a twofold procedure [44]:

- Providing a framework for propagation of systematic uncertainties. This is a matter of (a) understanding the influence of each specific systematic uncertainty, (b) propagation of the uncertainty into the measurement of the couplings.
- Turning the measurement of the coupling into a joint measurement of both coupling and systematic uncertainties.

In order to ease the modeling of systematic uncertainties using the profile likelihood approach, some assumptions are made:

1. The distribution of the magnitude of all systematic uncertainties are assumed to be Gaussian.
2. The systematic uncertainties considered are assumed to be uncorrelated. If such a correlation exists, the covariance matrix can be diagonalized to obtain $\vec{\mu}$ in a basis with no correlation.
3. Coherence of multiple systematic uncertainties are assumed, i.e. the order of application is assumed to have no effect on the final result.

The method itself consists of adding to the simulation the effect of the desired systematics, by adding the contribution from the Gaussian distribution of the systematic uncertainty. Adding these effects to the model described, yields a new negative likelihood ratio curve, which is wider compared to the one without added systematic uncertainty. The method is described in further detail in chapter 8.

4.4 Neyman construction

It is known that the method of maximum likelihood causes under coverage in cases with very low amounts of statistics [45]. A known solution to this problem is to use a full frequentist way of estimating CIs, such as The Neyman Construction [3, 46].

As it will be shown in this section, under coverage because of poor statistics is not the only problem of under coverage when fitting TGCs. It is possible to create »problematic« data samples, which are configurations of data which are entirely possible to measure in the detector, given the SM is true, but are not allowed by the TGC model, causing the validity of the CIs to break down.

The Neyman Construction estimates CIs based on *pseudo experiments*. A pseudo experiment is a Monte Carlo sample drawn (constructed) from the expected distribution desired to fit, at a given value of the couplings. In the simple case where a pseudo experiment in the variable of »number of observed events« is desired, the pseudo experiment is drawn from a Poisson: $N_{pseudo} = \text{draw}(\text{Pois}(n \mid N_{exp}(\vec{\alpha}_{in})))$ with $N_{exp}(\vec{\alpha}_{in})$ being the expected mean value at some integrated luminosity, given the couplings $\vec{\alpha}_{in}$. In the case where a shape fit is desired, a sample of $p_T(Z)$ is drawn from the expected distribution. It is of course also possible to combine the two by drawing from the extended model. As this is a computationally very demanding procedure, and the CPU time of all existing

implementations of the procedure grows as the number of fit parameters squared, only fits of one coupling at a time is considered in this type of parameter estimation. The method used for drawing from a distribution is provided in ROOT/RooFit [42], [43], which is a variant of the hit-and-miss method.

The pseudo experiments are used to scan the parameter space by fitting the individual experiment using a maximum likelihood fit, retrieving only information about the central value. By repeating the procedure many times, one obtains a »slice« of coupling values possible to obtain from experiment, given some true value of the coupling.

After constructing a slice, the $1 - a$ fractile which is *least likely* to occur is removed from the slice. If a slice (*after* the removal) covers the central value of a fit to real data, the α_{in} of this particular slice is contained in the confidence interval. A simple two step algorithm where an input interval that surely over covers the CI is used. Step one is obtaining a full confidence belt [3]:

1. Construct a slice at an input value α_{in} beneath the lower boundary of the CI
2. Increase α_{in} by the desired precision of the fit
3. While α_{in} is less than the upper boundary of the CI, cycle to (2)

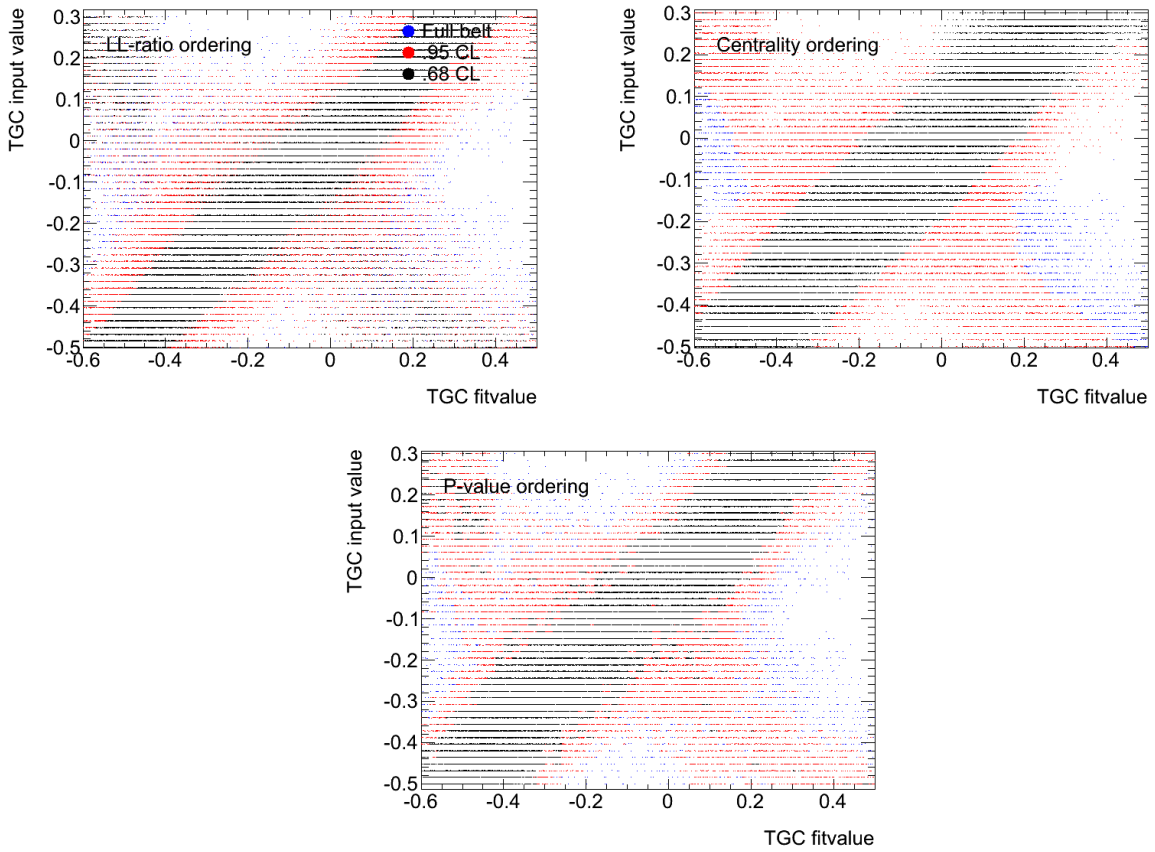


Figure 4.4: Example of full confidence belts, with application of three different »ordering rules«, as explained in section 4.4.1. The considered coupling is $\Delta\kappa_Z$, and the Neyman Constructions are made assuming 1fb^{-1} of integrated luminosity. The input is truth level POWHEG SM samples fitted with a binned maximum likelihood fit.

4.4.1 Ordering

Before proceeding to extract CIs, it is necessary to define the removal of the »least likely« values from the slices in figure 4.4. Since the notion of least likely is not unique, a choice of *ordering priciple* is needed [47]. A particular popular choice stems from Feldman and Cousins [45], and is motivated by the Neyman-Pearson lemma [40]. The choice is to use the likelihood ratio $\mathfrak{r}$ (equation (4.6)), with the null hypothesis being the input TGC value (α_{in}) of a given slice. By doing so, one can extend the aforementioned algorithm by step two[4]:

1. Calculate a central value using data, or use SM expectation.

2. Calculate $\mathfrak{r}$ for each pseudo experiment.

3. Remove the $1 - a$ fractile with the smallest $\mathfrak{r}$ from each slice.

4. If the central value is in the slice, the slice input value is covered by the CI.

Since the ordering principle decides what is meant by »least likely«, it will clearly also drastically alter the size of the fully frequentist CI. In figure 4.4 three different ordering principles are employed on the same sample of pseudo experiments, done only for $\Delta\kappa_Z$. The 68% and 95% CI are shown in black and red, whereas the 5% most unlikely fit values are marked blue. It can be noted that sorting based on likelihood ratio is equivalent to sorting based on log-likelihood ratio (equation (4.6)), denoted LL-ratio in figure 4.4 (left, upper). Here the Feldman Cousins ordering/log-likelihood ordering is used as it will be for the remainder of this study. This is compared to (right, upper) a centrality based ordering, where the distance to the input value is used to ensure a central CI, and (lower) p-value which is given as the probability to obtain a TGC fit value at least as extreme as the one given.

Error on the error

The error on the obtained CI is of course affected by the number of pseudo experiments in each slice. To make an estimation of this error, the error of a binomial CI is used; the so-called Wald interval [48]. For 95 % CIs this is $\sigma_{err} = \frac{0.43}{\sqrt{N_{\text{ps. exp}}}}$ where $N_{\text{ps. exp}}$ is the number of pseudo experiments per slice. With a 1000 pseudo experiments per slice, this number is $\sigma_{err} = 0.01$, which determines the number of significant figures on the TGC limits.

[4] An advantage of using such a two step algorithm is, that it is possible to (a) apply a variety of different ordering rules on the same ensemble of pseudo experiments (b) calculate CIs with different values of a on the same ensemble of pseudo experiments. This is especially useful when generating large samples, and fitting extended functions – or profiling of nuisance parameters, which can be very time consuming.

$\mathscr{L}_{int}$	$\Delta\kappa_Z$	Δg_1^Z	λ_Z
828 pb^{-1}	$[-0.32, 0.37]$	$[-0.07, 0.06]$	$[-0.05, 0.05]$
5 fb^{-1}	$[-0.20, 0.26]$	$[-0.04, 0.04]$	$[-0.03, 0.03]$

Table 4.5: Confidence intervals at a 95% CL from Neyman Construction with Feldman-Cousins ordering, using binned shape fits of the $p_T(Z)$ distribution. The CIs are quoted for a SM fit value.

4.4.2 A point-by-point approach

The described procedure is very useful for reasons of clarity. The plots of the full bands in figure 4.4 provides a transparent explanation of the Neyman Construction. However, calculating such full bands, is computationally very inefficient for the purpose of calculating a CI. The idea behind the point-by-point procedure is to spend more CPU time near the boundary of the CI. This also addresses the point that the »error on the error« is determined by a preset granularity.

The slices are tested one by one. If a slice is close to being included or excluded, by a preset multiple of the binomial error, more pseudo experiments should be added to the slice[5]. In this implementation, the sample amount is increased in a pre-defined interval around the endpoints.

4.4.3 Systematic error

The Neyman Construction using Feldman Cousins sorting incorporates systematic error propagation in a nuisance parameter approach, by construction. Systematics are incorporated as nuisance parameters as described in section 4.17, and instead of sorting based on likelihood ratio, the *profile likelihood ratio* (equation (4.18)) is used.This will be used to estimate systematic errors on the TGC measurements in chapter 8.

4.5 Comparison

The usage of the Neyman Construction to complement the ordinary maximum likelihood is traditionally done to accommodate problems of under coverage in cases of low statistics. In the case of TGCs, the problem is more complicated, as the TGC model itself gives rise to under coverage of the maximum likelihood fits. These points are addressed in the following, by showing that the likelihood intervals undercovers[6] the Neyman intervals, given certain datasets which will be referred to as »problematic«[7]. All arguments regarding CIs are done using Δg_1^Z, but are identical for all three TGCs considered in this thesis.

[5] This is also known as *adaptive sampling*.

[6] And not the other way around, as the Neyman CI can be taken as the *definition* of exact coverage [3, 46].

[7] A problematic dataset is sometimes (in literature) referred to as a »lucky« sample.

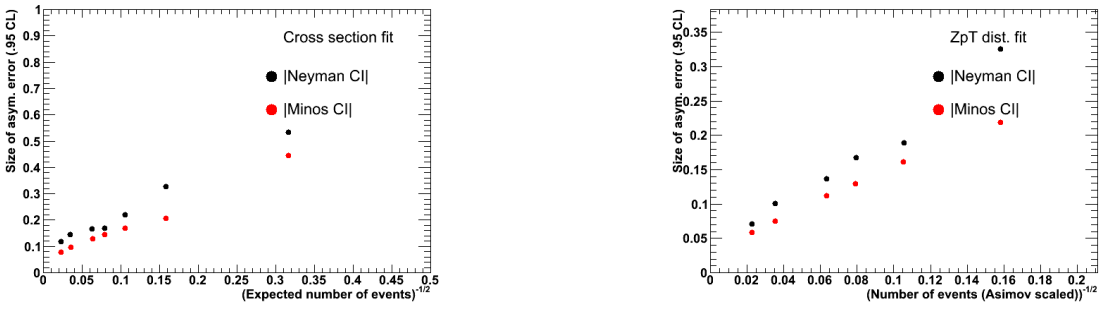


Figure 4.5: The size of the Neyman CI (black) and the Likelihood CI (red, marked MINOS) as a function of number of observed events in a fit of total number of observed events (left) and an unbinned shape fit (right). Mind the dimension of the primary axis, which is (number of events)$^{-1/2}$.

4.5.1 Large luminosity behavior

The comparison between the Neyman interval and the likelihood interval with SM truth MC carried out in the previous sections, is now extended to cover a wider range of luminosities. It is expected that the difference between the Neyman intervals and likelihood intervals when fitting a SM sample, should diminish when the integrated luminosity goes up. As seen in figure 4.5, this is indeed the case.

Each red point in figure 4.5 represents a likelihood interval given an Asimov SM sample, and a black point a Neyman CI given SM. The expected approximate square root dependence of the size of the CIs is evident, and the two types of CI follows each other, with the likelihood under covering, as expected.

4.5.2 Problematic fits of total number of events

Since the total cross section is a quadratic function in the couplings with the SM as the approximate minimum, it is not clear how the fits will behave if the total number of observed events is below the SM expectation. Clearly this behavior is – from a physical point of view – not incorporated in the TGC model, but is anyway a likely experimental scenario, given that the SM is true. Since a total number of events lower than SM expectation is *not allowed* in the framework of the TGC model, this is a problematic sample. No additional information about the couplings is gained by obtaining a data sample not allowed by the model, and the CI of full coverage should not change in size when the observed number of events goes below SM expectation.

In figure 4.6 this behavior of problematic samples is shown. The number of observed events is varied $\pm 3\sigma$ from SM expectation; i.e:

$$N_{obs} = N_{exp}(\vec{\alpha} = 0) \pm n\sqrt{N_{exp}}, \text{ with } n = \{0, 1, 2, 3\}, \tag{4.19}$$

and the likelihood fits clearly under covers when the number of observed events is below SM expectation.

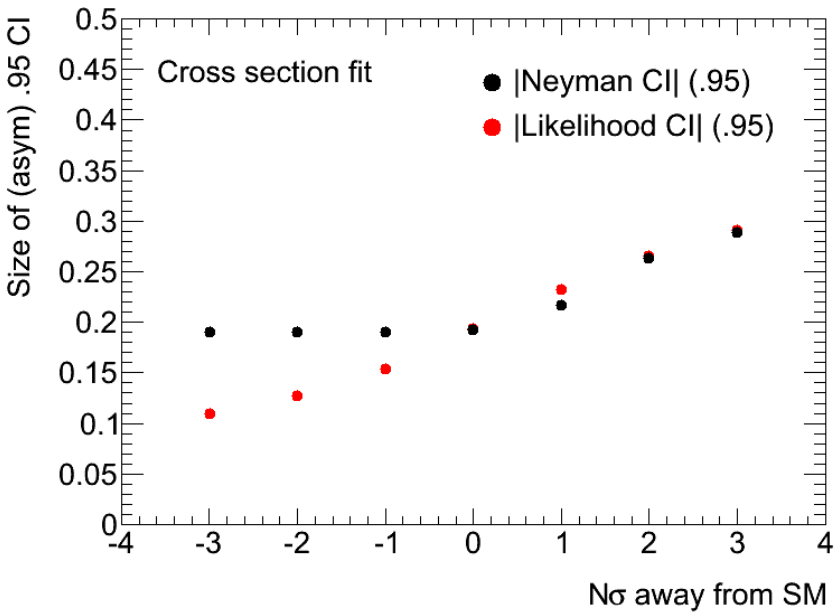


Figure 4.6: Fit of total number of events, size of CI for Neyman and likelihood (MINOS) fits are given. A problematic behavior is observed when observation fluctuates below SM expectation.

The limits from the Neyman Construction on the other hand behaves as expected, and does not decrease as one goes below SM expectation.

4.5.3 Problematic shape fits

Since the notion of above or below SM expectation is not as clear when considering shape fits, two different ways of constructing problematic samples for fitting shape will here be employed. One is designed for constructing problematic samples for binned fits and one designed with unbinned fits in mind. The idea behind both types of problematic samples is, that the shape of the distribution is to be tweaked, such that it lies above SM expectation for low values of $p_T(Z)$ and below for high values of $p_T(Z)$. The reasoning behind this is, that the high $p_T(Z)$ tail is the most sensitive to TGCs and, as previously seen, the TGCs (always) leads to an increase in this tail.

Problematic binned fits

It is desirable to construct a problematic data sample such that it is as transparent as possible, that the problem at hand stems from the incapabilities of the TGC model to accommodate all possible observations. Furthermore it is desirable to have the effect so outspoken as possible, in order to illuminate the problem.

A problematic data sample is constructed by first drawing a sample at SM and a TGC

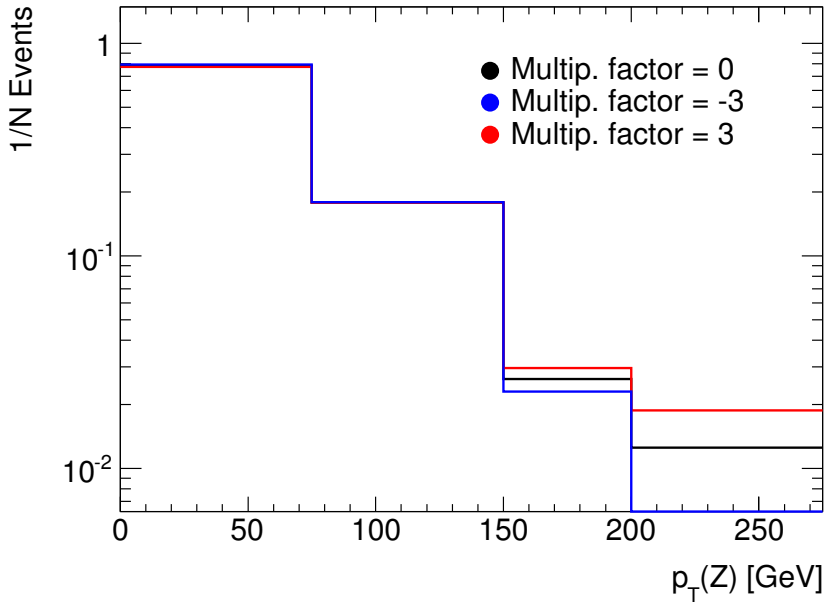


Figure 4.7: Example of construction of a datasets problematic when fitting TGCs in a binned maximum likelihood fit. The SM sample (black) is added to the ratio between a sample constructed with $\Delta g_1^Z = 0.25$ and the SM sample, multiplied with a factor of (blue) -3 and (red) 3. The blue sample is expected to be problematic.

sample, here exemplified with $\Delta g_1^Z = 0.25$. This specific value of Δg_1^Z is chosen, as it (a) Is in the ballpark of the expected sensitivity of a measurement using data corresponding to 827.87 pb^{-1} of luminosity, (b) The features of the resulting figure is easily compared to the ones of figure 4.6. The ratio of the two is taken, and this ratio is added to the SM sample multiplied by a factor, which will be varied between -3 and 3. In figure 4.7 examples of such samples in the chosen binning is shown, with 0 corresponding to SM.

If the problematic behavior corresponds to the one seen in figure 4.6, the blue sample on figure 4.7 should show the most problematic behavior of the three.

As seen in figure 4.8, this is indeed the case. If the observed number of events in the high $p_T(Z)$ tail is less that expected from SM, the likelihood fit under covers the Neyman Construction limits.

This is in some sense more problematic than the behavior shown for the total number of event-fits as the high $p_T(Z)$ tail will always suffer from poor statistics, since, as the amount of data and energy goes up, the criteria for what constitutes as a high $p_T(Z)$ tail will also move.

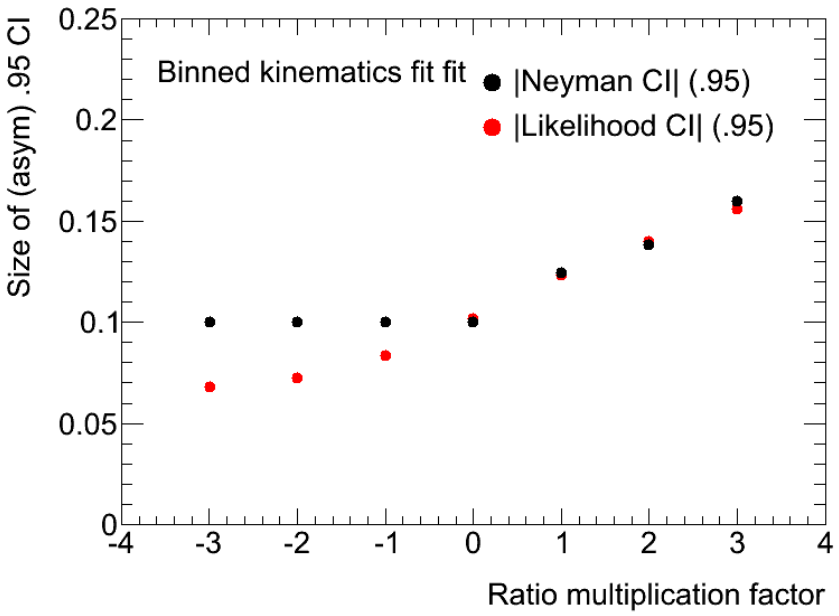


Figure 4.8: Binned shape fit, size of CI for Neyman and likelihood (MINOS) fits are given. A problematic behavior is observed when observation fluctuates below SM expectation.

Problematic unbinned fits

When doing an unbinned fit, the tweaking of a sample must be applied on an event by event basis. Each event is assigned a weight inversely proportional to the $p_T(Z)$:

$$w = \frac{k}{p_T(Z)} \tag{4.20}$$

where the factor k will be varied. One could very well obtain a sample tweaked in this characteristic way, since the p_T resolution is better for low p_T leptons, which are used to reconstruct the Z. An Asimov data sample is created and weighted this way, and the size of the CIs using a maximum likelihood fit and the Neyman Construction is shown in figure 4.9. As seen, samples like this also exhibits a behavior with likelihood fits under covering.

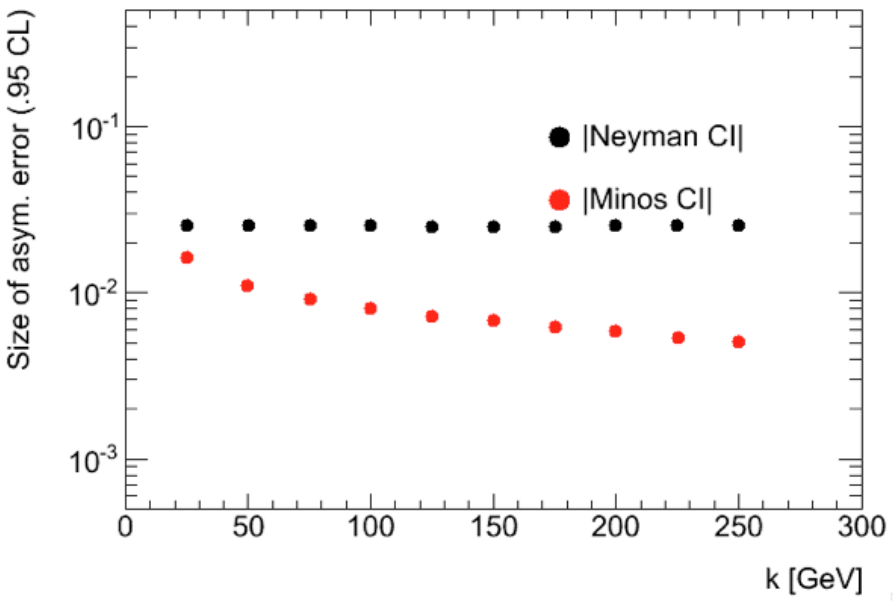


Figure 4.9: Unbinned shape fit, size of CI for Neyman and likelihood (MINOS) fits are given. A problematic behavior is observed when observation fluctuates below SM expectation.

4.6 Summary

From the Monte Carlo study performed here, it is clear that if one simply uses a maximum likelihood fit to determine the limits on the TGCs, one can end up in a situation with tighter limits than data provides justification for. It is possible to overcome this specific problem by using a Neyman Construction for setting limits, and the Neyman Construction will therefore be used when setting limits.

The virtues of maximum likelihood fits using an Asimov MC dataset should be noted. As this type of (pseudo) data set does not (by construction) posses any of the considered problematic features, such a MC dataset is an option to use for various optimization purposes. When doing MC predictions in the later chapters of this thesis, expectations given an Asimov MC dataset is quoted along with the Neyman expectations.

CHAPTER 5

THE EXPERIMENT

The energy frontier of particle physics have been pushed enormously over the last 50 years, with the ring colliders at CERN, Geneva, and Fermilab, Illinois, as some of the principal experiments.

Currently, the only collider experiment operating at the energy frontier is the Large Hadron Collider (LHC) located at CERN. The LHC spans the border between France and Switzerland 100 m underground, with a circumference of about 27 km. It feeds two general purpose experiments with collisions, ATLAS and CMS, along with several other more specifically purposed high energy physics experiments.

The idea behind this chapter is to provide the reader with a basis for understanding:

- How the desire to test the theoretical and phenomenological considerations up until this point in the thesis is actualized in a detector.
- How specific physics goals plays together with knowledge about how particles interact with matter, to decide how to build up the ATLAS detector.
- How the limitations of the detector and following reconstruction algorithms makes certain steps in the following analysis necessary, and limits the accessible physics.

A single chapter in a Master's thesis neither can nor should be a thorough introduction to the LHC and ATLAS machines, and the computing infrastructure that surrounds them. With the above mentioned three points in mind, this chapter will focus on putting the detector into the context of the present analysis – and introduce a reader who is not familiar with the ATLAS experiment. When this is not sufficient, the reader is referred to the many references (also in the text), which describes the complex in much greater detail. Especially the references Ref. [49] and [50] are key to understanding the details of LHC and the ATLAS experiment.

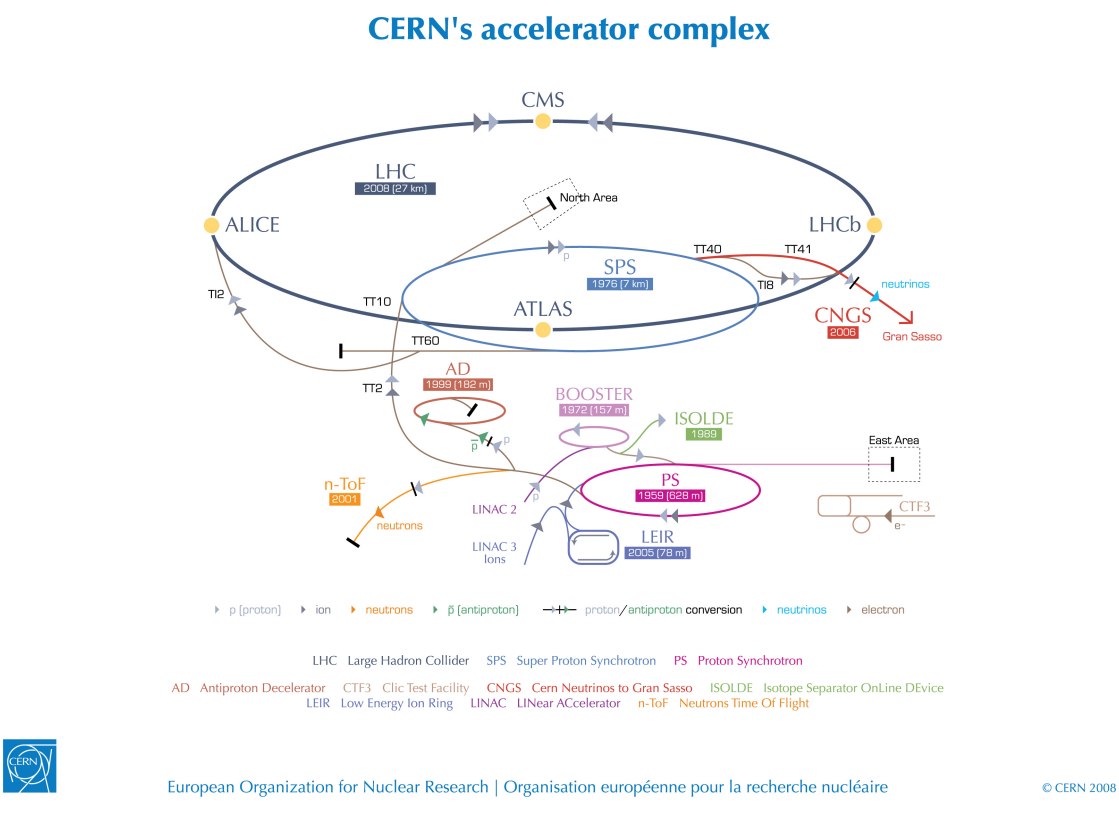


Figure 5.1: Schematic drawing of the accelerator complex, including the main experiments [51].

5.1 The Large Hadron Collider

The LHC is the proton-proton collider [49] which supplies the ATLAS experiment with collisions. It consists of a large complex of proton sources and pre-accelerators, depicted in figure 5.1. The protons are pre-accelerated in the LINAC 2. They are further accelerated in the BOOSTER, and after having circulated in the PS and SPS synchrotrons, they are fed into the LHC, where they reach a beam energy of 3.5 TeV per beam[1]. The proton beams are kept on a circular path by dipole magnets at a nominal strength of 8.33 T, and kept focused by additional magnets. The magnets are superconducting, and cooled by liquid helium, which brings the operating temperature down to 1.9 K.

In the LHC, the protons are bunched together in up to 2808 bunches per beam, with a nominal bunch spacing of 25 ns. The nominal number of protons in a single bunch is $1.15 \cdot 10^{11}$. The amount of data taken is measured in units of luminosity, which can also

[1] As per 2012 the beam energy is increased to 4 TeV per beam. The machine is designed to go up to 7 TeV per beam.

	2010	2011	2012	Design (2014-)
C.O.M. energy ($\sqrt{s}$)	7 TeV	7 TeV	8 TeV	14 TeV
Inst. Lumi. ($\mathscr{L}$ [$\mathrm{cm}^{-2}\mathrm{s}^{-1}$])	$9 \cdot 10^{26} - 2 \cdot 10^{32}$	$10^{30} - 3.6 \cdot 10^{33}$	-	10^{34}
Bunch spacing	50 ns	50 ns	50 ns	25 ns

Table 5.1: Principal LHC collision parameters. Historical from Ref. [52], prospected (2012) from Ref. [51] and design from Ref. [49].

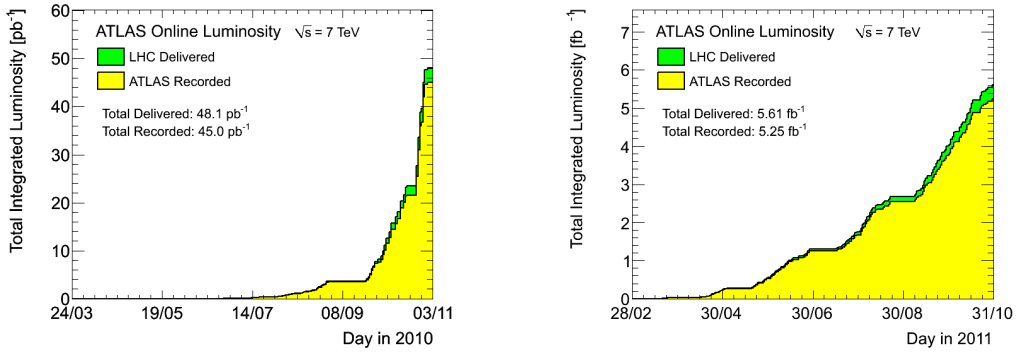


Figure 5.2: Integrated Luminosity (cumulative plots) at $\sqrt{s} = 7$ TeV delivered by LHC and recorded by ATLAS [53] in 2010 (left) and 2011 (right). Note the factor 10^3 difference in units on the secondary axis.

be calculated from the parameters of the collider [3]:

$$\mathscr{L} = fA\frac{n_1 n_2}{4\pi\sigma_x\sigma_y}, \tag{5.1}$$

where f is the collision frequency, n is the number of particles in a bunch and σ is the transverse beam size, assuming Gaussian profiles. Since the beams are brought together at a slight crossing angle, the luminosity is reduced by an efficiency factor A. The nominal instantaneous luminosity is $\mathscr{L} = 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

From the beginning of LHC operations in November 2009 (after a one year shutdown due to a malfunction of the superconducting magnet interconnects), the actual luminosity have been increased, as will the energy in 2012 (see table 5.1 for an overview).

The interesting number regarding the amount of for this analysis, is of course the total integrated luminosity $\mathscr{L}_{int} = \int \mathscr{L} dt$, which in 2011 exceeded 5 fb^{-1}, making the LHC high luminosity experiments the most data-rich experiments at the energy frontier. The integrated luminosity recorded by ATLAS is shown in figure 5.2. The data target for 2012 is 15 fb^{-1} – three times higher than 2011. For this study, 827.87 pb^{-1} collected in the start of 2011 is used.

5.2 ATLAS

The ATLAS experiment (A Toroidal LHC ApparatuS) [50] is a multi-purpose detector designed for receiving the full luminosity from the LHC collider. It is the largest of the LHC detectors with a height of 25 m and a length of 44 m. An overview of the detector is provided in figure 5.3. At a glance, the ATLAS detector consists of:

- Tracking devices in the *Inner Detector*, designed for taking »pictures« of about 1000 particle tracks each time two bunches collides.
- Calorimetry systems designed for stopping electromagnetic and hadronic showers and provide readouts of the energy.
- Muon system providing state-of-the-art measurements of muon momentum.
- Forward detectors (LUCID and ALFA) providing luminosity measurements.

Not quite a detector, but also worth mentioning, is the large magnet system, which bends charged particles using two superconducting magnets.

In this section only the parts relevant for this specific study will be described. For a more detailed description of the ATLAS detector, Ref. [50] can be consulted, where also the information put in this section is taken from, if not stated otherwise.

The detector geometry is cylindrical, meaning that positions are given in usual cylindrical coordinates (R, ϕ, z). To describe the angle of a particle relative to the beam axis, usually the *pseudorapidity* is used[2]:

$$\eta = -\log\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{5.2}$$

and from η and ϕ the derived quantity $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ is obtained to describe a 2-dimensional radial distance between two points in the detector. Of particular relevance to this study, is the ability to observe final states containing muons and electrons. Good resolution on p_T is essential. The neutrino in the final state is not directly observable, but is seen indirectly through *missing transverse energy* E_T^{miss} [55].

5.2.1 Inner Detector

The Inner Detector provides momentum and vertex measurements of charged particles. As shown on figure 5.4, the Inner Detector consists of several subsystems, providing different granularity. The pixel detector and the silicon strip detector both measures particle trajectories at a high precision. The transition radiation tracker provides both particle trajectory measurements for ordinary particles (low threshold), and adds additional coverage for ultra relativistic particles by measuring transition radiation (high threshold). The three measurements are combined in a track reconstruction (see section 5.3). All

[2] A nice feature of pseudorapidity is, that it reduces to ordinary rapidity in the limit of massless particles. Rapidity is in turn additive under Lorentz transformations.

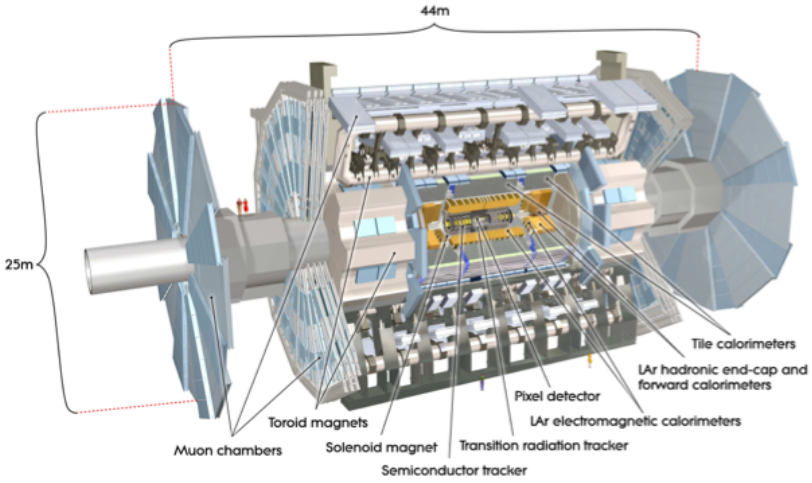


Figure 5.3: Cutaway view of the ATLAS detector, showing its vital parts (without the forward detectors). Picture taken from Ref. [54].

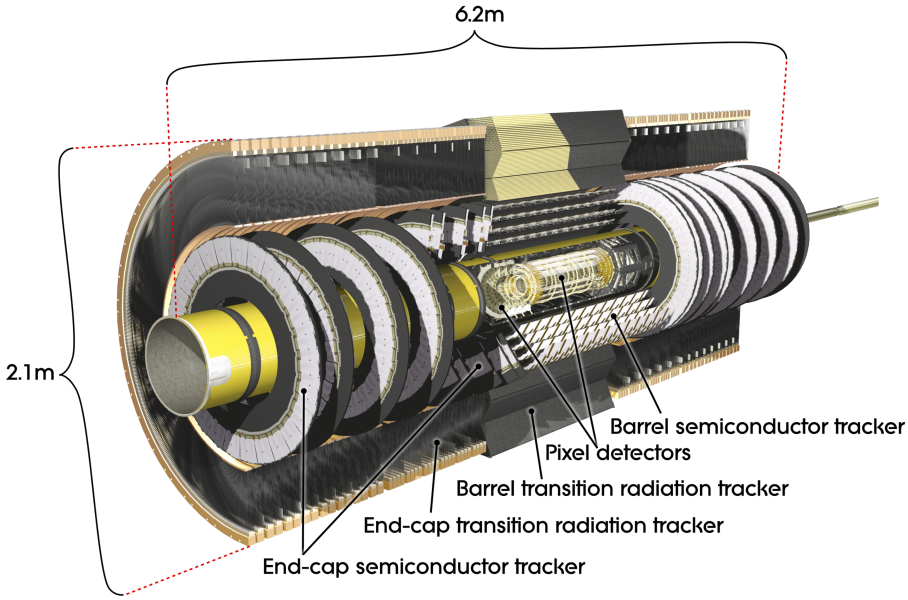


Figure 5.4: Cutaway wiew of the Inner Detector, showing its subsystems. Picture taken from Ref: [54].

Detector	Intrinsic accuracy [μm]	Readout channels [10^6]
Pixel		80.4
Barrel	$10(R-\phi)115(z)$	-
Endcaps	$10(R-\phi)115(R)$	-
Strip		6.3
Barrel	$17(R-\phi)580(z)$	-
Endcaps	$17(R-\phi)580(R)$	-
TRT	130	0.35

Table 5.2: Intrinsic measurement accuracies (expected) and amount of readout of the Inner Detector subsystems. Numbers taken from Ref. [50].

three parts of the Inner Detector relies on the fact, that a charged particle will deposit energy when traversing a material, according to the Bethe-Bloch equation [55], which describes the deposited energy in the material $^{\mathrm{dE}}/_{\mathrm{dX}}$ as a function of particle velocity.

Silicon detectors

The two detectors situated closest to the interaction point are the silicon pixel detector and the silicon strip detectors. They both rely on the semiconductor properties of silicon, which ensures that electron-hole pairs are created proportionally to the deposited energy [56]. By applying an electric field, the electrons and holes drift to the electrodes and measured. This provides both the information that a strip/pixel has been hit and deposited energy.

The pixel detector provides the highest granularity, using silicon pixels with a minimum size of $(R-\phi)\times z = 50\times 400\mu\mathrm{m}^2$, and has approximately $8\cdot 10^7$ readout channels. It is arranged as three concentric cylinders (barrels) around the beam axis, with disks (endcaps) on the side to ensure optimal (ϕ,η) coverage.

The strip detector is also designed with barrels and endcaps. The layers are double sided with a stereo angle of 40 mrad between the sides in order to gain the ability to observe in three spatial dimensions. The silicon strip detector has approximately $6.3\cdot 10^6$ channels.

Transition Radiation Tracker

The TRT consists of 4 mm×144 cm tubes filled with a Xe-based gas mixture, and measures the ionization of particles traversing the detector. The free electrons in the gas, stemming from the ionization, is driven to the electrodes by an electric field, and then measured. The TRT-type detector is much cheaper than the Si types, but also worse off in terms of spatial resolution, as seen from table 5.2.1. However, since the number of hits is so large (36 per track), the lack of resolution is compensated.

The TRT also provides electron identification by measuring transition radiation from

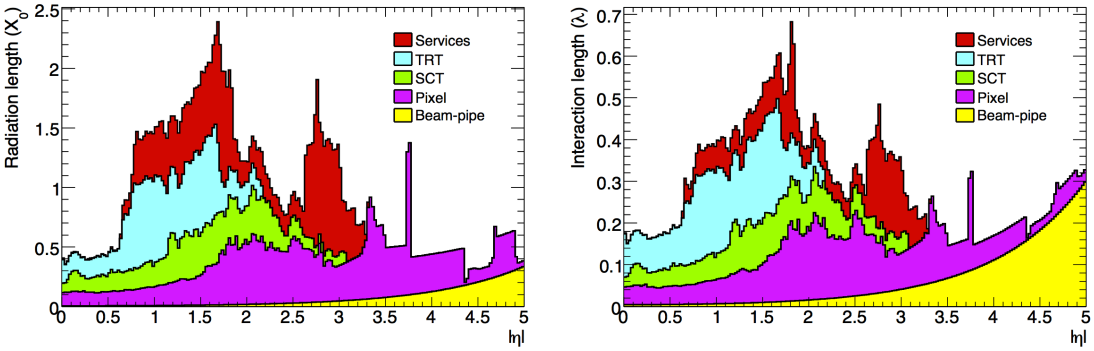


Figure 5.5: Material budget for the Inner Detector (left) by radiation length, (right) by interaction length. Taken from Ref. [54].

the transition between the layers, as the space between the tubes is filled with a special multilayered radiator material.

Material budget

The amount of material used for the Inner Detector is a major concern in the sense, that the traversing particles loses energy before arriving at the calorimeter. The material budget of the inner detector is shown in terms of radiation length (X_0) and interaction length (λ) in figure 5.5.

It is seen that the total material distribution exceeds $2X_0$ for $|\eta| \approx 1.5$, even though it is very low for the central region. This influences the later reconstruction and identification (see section 5.3).

5.2.2 Calorimeters

The calorimeter system in ATLAS is designed to cover the full $|\eta| < 4.9$ range, using techniques suitable for the type of physics expected in a given η-range. In the region matched to the Inner Detector, a granularity adequate for precision measurements of electron and photons is maintained. The rest of the calorimeters are coarsely grained to satisfy the weaker physics requirements for E_T^{miss} measurements and jet reconstruction. A full cutaway view of the calorimeter system is shown in figure 5.6.

Again the amount of material used is an important parameter. One wishes to provide containment for EM -and hadronic showers and limit punch-through into the muon system. Good containment in the calorimeter system is essential for providing good E_T^{miss} measurements, which will be used in this study. The thickness of the EM calorimeters are summed up in table 5.6.

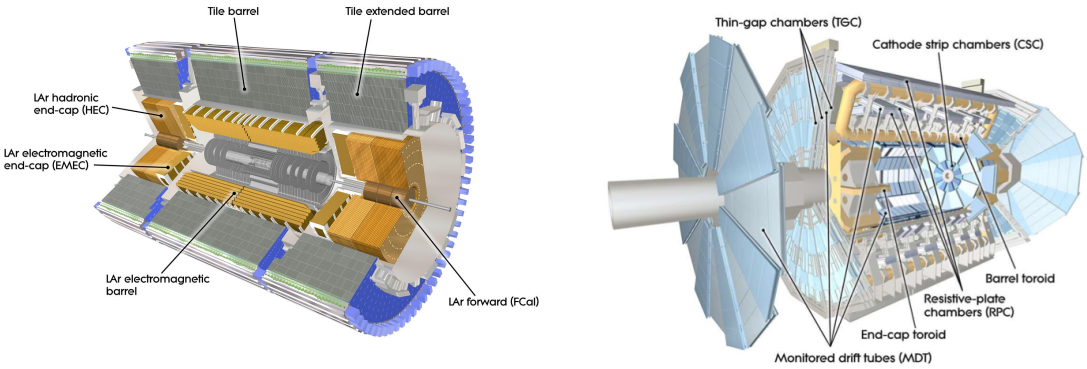


Figure 5.6: Cutaway view of (left) the calorimeter system and (right) the muon spectrometer system. Pictures taken from Ref. [51].

Calorimeter	Thickness $[X_0]$	Coverage	Readout chan.
EM	-	$\|\eta\| < 3.2$	173,312†
Barrel	> 22	$\|\eta\| < 1.475$	101,760‡
End cap	> 24	$1.375 < \|\eta\| < 3.2$	62,208‡

† Including presampler
‡ Not including presampler

Table 5.3: Summary of EM calorimeter parameters. Numbers taken from Ref. [50] and [57].

Electromagnetic calorimeter

The electromagnetic calorimeter is a lead (Pb)-liquid argon (LAr) calorimeter of the sampling type. EM calorimeters in general, are devices which rely on radiation of bremsstrahlung, happening within the characteristic length X_0 called the radiation length, which goes as Z^{-2} [55] (Z denotes the atomic number). An incoming electron will radiate bremsstrahlung, and if the photon energy is high enough, an e^+e^- pair is created, which in turn will also radiate bremsstrahlung. This process continues until there is not enough energy in the system to create new e^+e^- pairs resulting in a shower. In the ATLAS EM calorimeter, the shower is developed in Pb plates (high $Z \rightarrow$ short X_0) and sampled through ionization in the LAr, which in turn have a long X_0. The geometry of the Pb-LAr layer structure is shown in figure 5.7 (left).

In addition to the three layers seen in figure 5.7 (left), the central region ($|\eta| < 1.8$) is also equipped with a presampler, which corrects for energy lost upstream to the calorimeter.

Calorimeters based on liquid noble gasses, can in general obtain a good resolution, as seen in figure 5.7 (right), where a relative energy resolution as a function of electron energy is quoted. The great disadvantage of a LAr calorimeter compared to for example

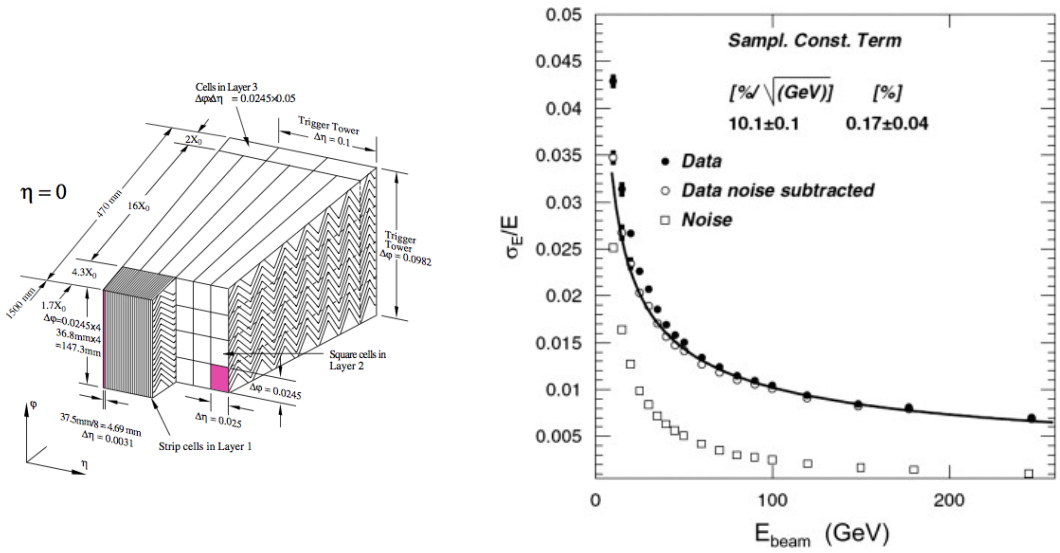


Figure 5.7: The accordion geometry of the EM calorimeter (left) . The energy resolution as a function of the electron energy (right), see section 5.3 for details. Both are from Ref. [50].

the CMS calorimeter using tungsten scintillating crystals [57], is that LAr calorimeters have a comparatively longer collection time for the charge, which limits the counting rate of the calorimeter [56]. In ATLAS this is solved by the accordion design.

Hadronic calorimeters

The hadronic calorimeters are for this study primarily interesting for the E_T^{miss} measurements. When looking at any colored final states, hadronic calorimetry becomes immensely important, as one needs to reconstruct jets from hadronic calorimetry data. This is relevant for this study since jets are used in the event selection.

The hadronic calorimetry system consists of:

Hadronic barrel (Tile) which develops showers in steel plates, and uses scintillating tiles as the active material. In total, it covers the $|\eta| < 1.7$ range. It is divided into three layers, with a more coarse granularity in the third layer

Hadronic end cap (HEC) develops showers in cooper plates, and uses LAr as the active material. It covers the $1.7 < |\eta| < 3.2$ region.

Forward calorimeter (FCAL) makes both hadronic and electromagnetic calorimetry with respectively a cooper and a tungsten module, in the very forward region $3.1 < |\eta| < 4.9$.

The total thickness is 11 interaction lengths (λ) at $\eta = 0$ which is adequate to reduce the punch-through to the muon system [50].

Chamber	Coverage	No. of chambers	No. of channels
Monitored drift tubes	$\|\eta\| < 2.7$	1088	339,000
Cathode strip chambers	$2.0 < \|\eta\| < 2.7$	32	31,000
Resistive plate chambers	$\|\eta\| < 1.05$	544	359,000
Thin gap chambers	$1.05 < \|\eta\| < 2.7$	3588	318,000

Table 5.4: Main parameters of the muon spectrometer. Taken from Ref. [50].

5.2.3 Muon spectrometer

The muon system is designed for providing muon momentum at a good resolution, using tracking chamber techniques. The muon spectrometer is the outermost part of the ATLAS detector, and pictured in a cut away view on figure 5.6 (right).

A large toroid magnet system is used to bend the trajectory of the muons as they pass though three stations of tracking chambers. The tracking chambers provides position measurements, from which the momentum can be calculated. Both parameters are used in the muon reconstruction (see section 5.3).

The tracking stations consists of four types of chambers, using different principles [55]:

Monitored drift tubes which consists of a gas tube with a wire in the center. The gas is ionized by a passing muon. The ionization drifts to an electrode where it is collected, amplified and measured.

Cathode strip chambers where a passing charge is detected by an anode wire, creating an induced charge on a cathode.

Resistive plate chambers where two resistive electrode plates encapsulate a gas gap. Traversing particles causes the gas to ionize, and deformations of an applied electric field is translated to muon position.

Thin gap chambers where closely separated wired are situated in a thin gap between grounded resistive cathode planes. Traversing particles draws a spark between the wires, and a signal is read out.

The main parameters of the four types of chambers in the ATLAS muon spectrometer, is summed up in table 5.4.

5.2.4 Trigger system

Given the physics goals of the LHC experiments, ATLAS must be able to work at a very high input luminosity. As a natural consequence of looking for very rare processes, a lot of very common events must be filtered away. As one event takes 1.6 MB in raw data format (25 MB without zero suppression), the ability to filter away the most uninteresting events away, becomes very urgent. Putting all events on tape, and search through them

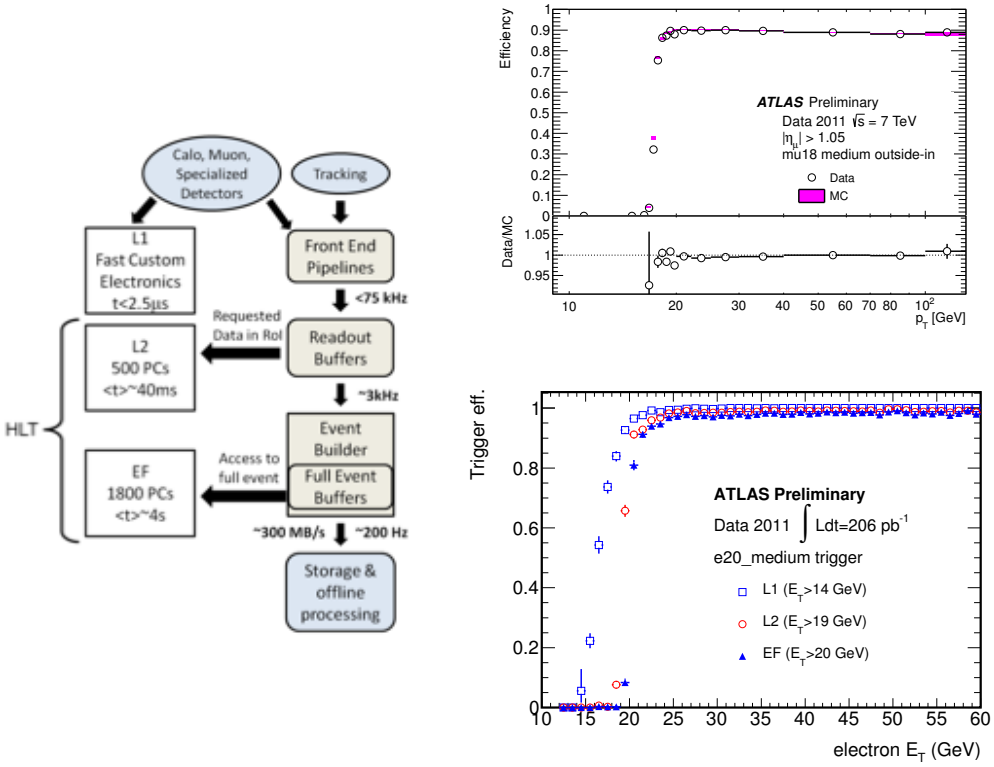


Figure 5.8: Overview of (left) the whole ATLAS triggering proceduce, (right upper) the efficiency of mu18 medium trigger in the endcap regians, terms of offline muon p_T, (right lower) e20 electron trigger efficiency in terms of electron E_T. All adapted from Ref. [54].

is not an option, since 1.6 MB per event would yield an output of about 64 TB/s running at design specifications.

The filtering procedures are known as *triggering*, as the procedure triggers on the events that looks interesting, by rapidly identifying the signatures of muons, electrons, photons, taus, jets and B mesons as well as missing energy [58]. The process is depicted in figure 5.8 (left). As seen in the figure, detector output of up to 40 MHz times 25 interactions per bunch crossing [58] is stored in front-end buffers, whilst the L1 trigger makes a decision whether to keep an event, reducing the event rate to a maximum of 75 kHz. The L1 trigger also points out regions of interest (RoI) as depicted in figure 5.8 (left). The data in the RoI are read in to the high level trigger (HLT) which is software based, running on a massive CPU farm. Here the output is reduced to $\approx$ 300 Hz, or equivalently $\approx$ 300 MB/s, which is stored on tape to be processed offline.

Example efficiencies for some of the used trigger for this study (see next chapter) is shown on figure 5.8 (right). It is seen that the efficiencies are high, but that a low efficiency region exists for both muons and electrons.

For this analysis muons and electrons will be demanded to have a $p_T > 15$ GeV. This

will remove the low efficiency region seen in figure 5.8, without loosing interesting events, as we are primarily interested in high p_T leptons which provide the highest sensitivity to TGCs.

5.3 Reconstruction

The information returned from the detectors needs to be combined into the physical objects for the study of TGCs. The basic objects are *tracks* left by traversing particles and *vertices* from which the physical objects electrons and muons will be reconstructed. The most complicated physical object needed is the E_T^{miss} used to reconstruct the neutrinos.

Since jets are used indirectly for removal of fake signal events, the jet reconstruction procedure is outlined.

5.3.1 Track reconstruction

The data obtained from the Inner Detector is recombined into particle tracks. The technique used in ATLAS is based on the so-called Kalman Filter [59], which is a recursive estimator technique. The basic idea of the process is to provide a guess of where a track is, given a certain initial value and the dynamics of the system. The guess is compared with data, and the initial guess is updated accordingly. The Kalman filter is run in Athena software before data is selected for this analysis. In ATLAS the Kalman filter is run both inside out and outside in. This is especially suitable when electrons are needed – such as for this analysis – which have a high energy loss in the Inner Detector, leading to a wrong inside out extrapolation. By running the algorithm outside in, particles that have been extrapolated wrongly into the SCT and TRT detectors can be catched.

5.3.2 Vertex reconstruction

If two or more tracks points stems from a common point, this point is denoted a vertex. The existence of a common vertex for multiple tracks, is taken as an indication that the tracks originate from the same collision.

The vertex with the highest number of tracks associated to it, is denoted *the primary vertex*. All other vertices are dubbed *secondary vertices*.

The vertex reconstruction algorithm is based on a χ^2 fit, and is described further in Ref. [60].

Since multiple interactions may take place per bunch crossing, multiple primary vertices is constructed in each event. In figure 5.9 a $Z \to \mu\mu$ event with very high pileup (25 reconstructed vertices), taken on April 15th 2012 is shown. This figure is in itself a good illustration of why the computational part of the reconstruction process takes up such a great amount of work in the ATLAS experiment.

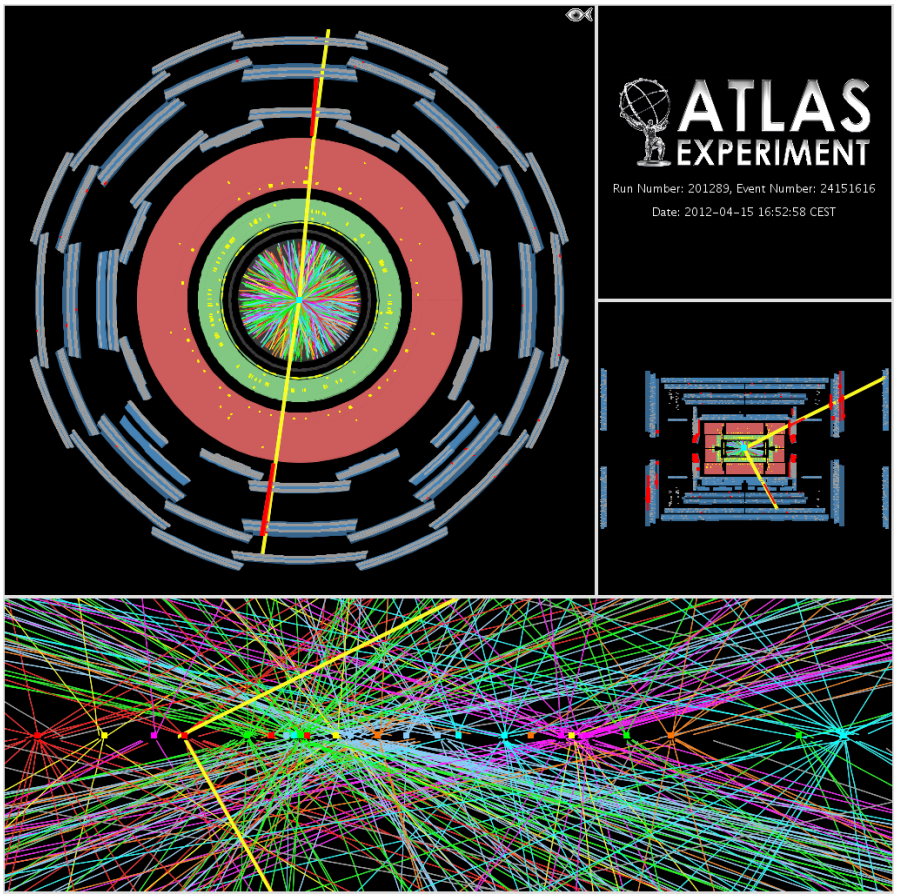


Figure 5.9: An example of a $Z \to \mu\mu$ event with very high pileup (25 reconstructed vertices), taken on April 15th 2012. The vertices are reconstructed using tracks with $p_T > 0.4$GeV, but the constraint on the number of hits on the tracks is tighter than in 2011.

5.3.3 Muon reconstruction

The muons used for this study is so-called *combined muons* which uses a combination of fitted track information from the inner detector as described in section 5.3.1 and a muon spectrometer track [61]. This combination is done to increase the momentum resolution, which is highly desirable, as recombination of WZ pairs needs well-defined high p_T muons.

In figure 5.10 (left) the efficiency of the muon tracking is shown.

There exist two algorithms for combination of inner detector tracks and muon spectrometer tracks. The STACO algorithm makes a statistical recombination of the two, whereas the MuID makes a new global fit of the muon track, using the Moore algorithm [61]. For this study the MuID muons are used.

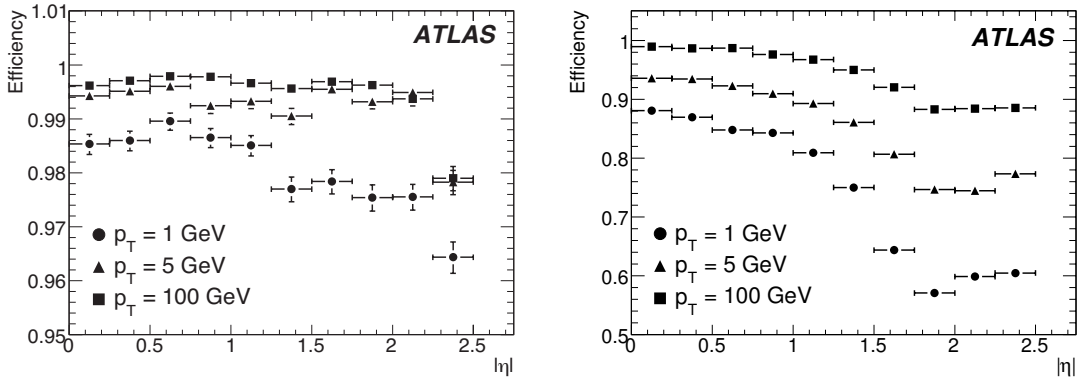


Figure 5.10: Tracking performance of muons of pT of 1, 5 and 100 GeV as a function of η (left) and for electrons (right). Taken from Ref. [54].

5.3.4 Electron reconstruction

The electrons are reconstructed using a combination of the fitted track information from the inner detector and clusters in the calorimeter [62]. The algorithm used starts from the calorimeter cluster, forming seed clusters with energies above 2.5GeV. Then tracks are matched to the cluster seeds to define electrons. If more than one track match the same seed cluster, the track with the smallest ΔR between the impact point on the calorimeter and the seed cluster is considered the best match. To deselect hits which are more likely to stem from photon conversions, silicon hits are given priority over non-silicon hits.

In figure 5.10 (right) the efficiency of the electron tracking is shown.

5.3.5 Lepton overlap removal

The reconstruction software can accidentally reconstruct a single physical object as two. Two such cases are considered by imposing cuts on the ΔR distance between neighboring objects.

As a single electron could be reconstructed as two electrons, it is required that if two electrons lie within a distance of $\Delta R < 0.1$, only the leading electron (with higher p_T) is kept. This cannot be done without care, as two electrons close in ΔR, could stem from a highly boosted Z. Using as an approximation to illustrate this effect, a Z decaying in the $\eta = 0$ plane only, and for electron mass equal to zero, the well known formula for the aberration of light can be used. Given that the electrons are back to back in the Z rest frame, the opening angle (and thus ΔR since $\Delta\eta = 0$) is given as:

$$\phi = 2 \arccos \left(\frac{1}{\sqrt{\frac{m_Z^2}{p_T^2(Z)} + 1}} \right) \tag{5.3}$$

A requirement of a $\Delta\phi$ separation of 0.1 will cut away all Zs decaying to electrons with a $p_T(Z) \gtrapprox 1800\text{GeV}$.

In order to ensure that electrons are not double counted as muons, electrons and muons are demanded to be separated by $\Delta R > 0.1$. In the case they are not, only the muon is kept.

5.3.6 Missing transverse energy

As the E_T^{miss} is not observed directly, but representing everything that is not observed, given energy -and momentum conservation, the reconstruction involves all parts of the detector.

The definition of E_T^{miss} used in this thesis is the so-called MET_REF_Final. This is calculated by applying energy- and momentum conservation in the calorimeter; and correcting for energy loss in the cryostat and energy of the muons. In figure 5.11 an event display of a candidate for a $W \to \mu\nu$ decay, where the W is produced in association with a photon is shown. Here both calorimeter information (the photon) and muon information is used for the reconstruction.

The reconstructed missing transverse energy is clearly a vector, defined as:

$$\vec{E}_T^{miss} = \vec{E}_T^{miss}(\text{calo}) + \vec{E}_T^{miss}(\text{muon}) + \vec{E}_T^{miss}(\text{cryo}), \tag{5.4}$$

where $\vec{E}_T^{miss}(\text{calo})$ is the missing transverse energy from the calorimeter; the negative vector sum of the transverse energy of all reconstructed calorimeter objects plus what is left from other cells. The vector $\vec{E}_T^{miss}$ is usually given as a magnitude and an angle.

Jet reconstruction

Jets are in this thesis used indirectly for vetoing certain types of background events. The jets used are reconstructed using the anti-k_t-four algorithm [63]. The idea is to combine tracks or calorimeter clusters using a metric distance measure with two components to combine tracks in the close vicinity of each other into jets. This results a list of jets, consisting of tracks bunched together. One is then able to combine the kinematics of the tracks, and obtain the jet kinematics used in the selection process.

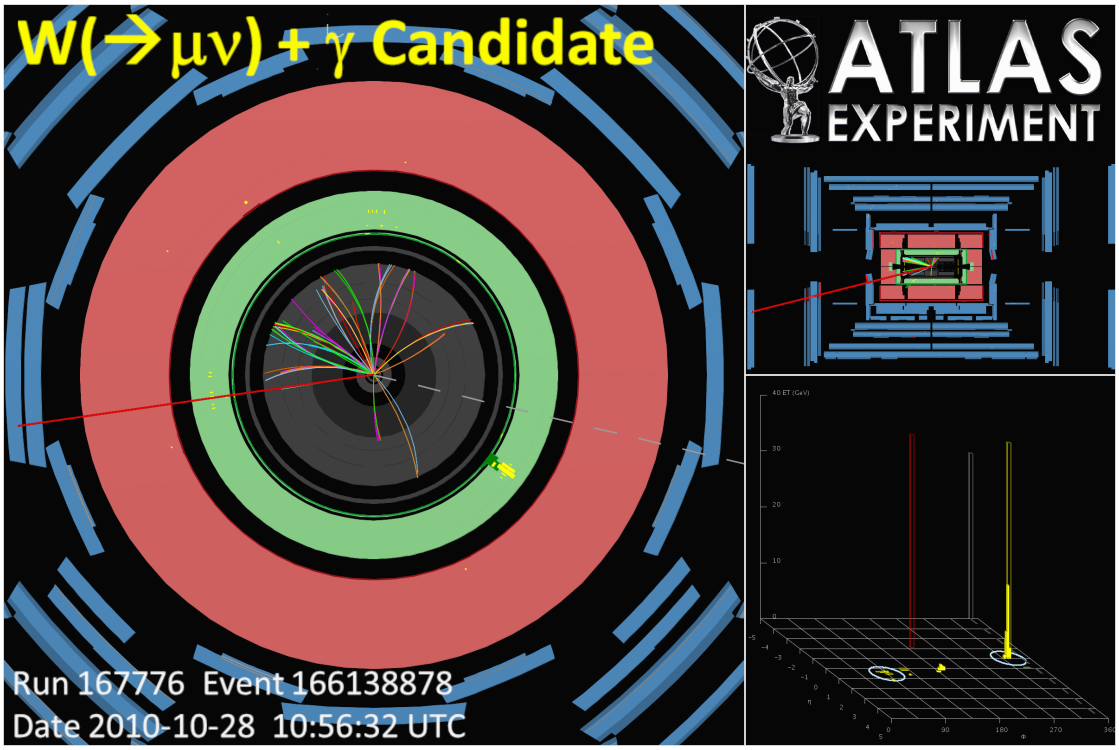


Figure 5.11: Display of an event taken on September 28th 2010, where E_T^{miss} is reconstructed from both calorimeter and muon energy. The event is a candidate for a $W \to \mu\nu$ decay, where the W is produced in association with a photon. The photon energy is measured by the EM calorimeter (the green ring) and the muon continues all the way out to the muon system. The E_T^{miss} (grey) is reconstructed using information from both.

5.4 Detector simulation

The complex interaction of particles with matter taking place in the ATLAS detector needs to be simulated for the Monte Carlo samples in order for these to make predictions about the behavior of the data. The study in this thesis is based on the official ATLAS simulation done in the Athena framework (see figure 3.1), which is based on the GEANT4 [64] program.

The detector simulation is really an umbrella notion for the process of:

- Modeling the physical interactions of particles with the detector material.
- Modeling detector geometry, dead regions, etc.
- Digitizing the signal, i.e. translating from energy deposits in the detector to the voltages and currents which are actually measured.

An output like the one from the ATLAS detector makes it possible to use the same aforementioned Athena reconstruction software as on data, thus treating data and Monte

Carlo in the same way from this point on (see again figure 3.1 and Ref. [65]). As the detector simulation itself is not a part of this study, it shall only be given this short mention. The interested reader is referred to the references [65, 64, 66].

CHAPTER 6

EVENT SELECTION

The event selection is the common description for all the different means used to extract the final states originating from the process the study wants to consider, from final states originating from other processes which produce a similar looking final state. In the jargon of experimental particle physics, the former is dubbed »signal« and the latter »background«.

The signal is chosen as one or more final states that are observable in the detector. For this study, final states containing three leptons and a neutrino is considered.

Although the choice of this particular final state removes a lot of WZ events, one obtains a very pure signal with almost no backgrounds [37].

The goal of this selection is to remove as much as possible of the backgrounds associated with this particular final state, while leaving as much signal as possible. Two different techniques are optimized on Monte Carlo samples, before proceeding to real data analysis. The two techniques are:

Cut based event selection A cut based selection is carried out to ensure (a) well defined leptons and (b) reconstruction of W and Z is possible. This is based largely on Ref. [37], but using somewhat looser cuts.

Boosted decision tree A multivariate signal selection using boosted decision trees is carried out to remove remains of the specific backgrounds described in section 6.1.

6.1 Backgrounds to WZ events

Although the final states considered are very pure, there are still some physics processes that could get reconstructed as this particular final state. Not only processes that actually ends up as this particular final state is considered, but also processes that can be misidentified as such. The following processes have been chosen as important background

contributions for this analysis [37, 67, 68], categorized according to the number of leptons from the hard process present in the final state, originating from the background process:

Backgrounds with four leptons from hard process

ZZ A ZZ with leptonic decay can mimic the WZ signature if one of the leptons escapes detection. The $ZZ \to 4l$ is sensitive to another category of TGCs – the so-called neutral TGCs – which are not the subject of this study, and is taken to be at SM level. The ZZ process is an obvious candidate for an important background, as it is a very similar process to WZ. The kinematics of the two types of events is as such expected to be close to identical. The difficult scenario where one Z decays to electrons or muons and the other to τs with one of the τs decaying further $\tau \to l\nu\nu$ provides the final state of exactly three charged leptons and E_T^{miss}. Fortunately the branching ratio is very small [67]: $BR(ZZ \to ll\tau\tau \to ll\tau l\nu_{tau}\nu_l) = 0.16\%$.

Backgrounds with two leptons from hard process

$t\bar{t}$ A $t\bar{t} \to WbW\bar{b}$ with Ws decaying leptonically and one jet is mis-identified as a lepton. The $b\bar{b}$ is also allowed to decay semileptonically which will allow for more leptons in the final state.

Z+**jets** Zs produced with jets can be mis-identified as a WZ signature if a jet is mis-interpreted as a lepton.

WW The WW final state includes two leptons, but more could emerge from ISR/FSR; especially QED FSR from final state leptons.

Backgrounds with one lepton from hard process

W+**jets** Ws produced with jets can be mis-identified as a WZ signature if two jets are mis-interpreted as leptons.

Backgrounds with no leptons from hard process

Dijet events Dijet signifies the huge background from QCD. It has the signature of multijets, depending on p_T scale.

All backgrounds demanding jets mis-identified as leptons are processes with a large cross section (factor $10^4 - 10^6$ higher than signal) [69], but there is an equally large rejection factor for mis-identified jets [67].

6.2 Data samples

The data set selected for analysis is taken with the ATLAS detector in 2011 at $\sqrt{s} = 7$ TeV (data periods B through G) and corresponds to an integrated luminosity of 827.87 pb^{-1} . As the total integrated luminosity recorded by the ATLAS experiment in 2011 amounts to nearly 5 fb^{-1}, an obvious question is to ask why this study does not include

all of it. Starting summer 2011, the LHC running conditions changed, leading to a significant increase in pileup. This had a drastic impact on the reconstruction of E_T^{miss}. Consequently this study uses data from the initial phase of 2011. To put this decision in perspective, the official ATLAS paper on WZ cross section and TGCs using all 5 fb^{-1} is not published yet, but expected for the summer conferences of 2012.

6.3 Monte Carlo samples

The Monte Carlo samples of WZ events used to test the event selection, and later to compare with data, are officially produced ATLAS Monte Carlo, generated with the MC@NLO [29] event generator, and using the CTEQ6.6 [27] parton distribution function. The $t\bar{t}$ background is also modeled with MC@NLO, and for the Z/W+jets samples Alpgen [70] has been used. The ZZ background is modeled with Herwig [71]. The underlying event is simulated with Jimmy [72], and the decay of τ particles to the lighter leptons are handled with the stand-alone application Tauola [73]. Herwig is used to handle ISR and FSR.

Dijet events are modeled with PythiaB [74], which employs its own model for ISR, FSR and underlying event.

All detector simulation is done with Geant4 [64]. In table 6.1 the signal Monte Carlo samples are summarized along with key parameters.

Corresponding tables for all Monte Carlo samples used for background estimation is put in appendix B. A feature of the used Alpgen Z/W+jets background samples is that the same heavy flavor final states arise in multiple samples. When combining samples together, certain classes of events will be double counted, unless they are properly vetoed. This is done using the hfor algorithm described in Ref. [77].

Signal samples generated at non-SM values of the TGCs are used to enhance statistics in the high $p_T(Z)$ tail, as described in section 3.2.1.

6.3.1 Sample weighting

All Monte Carlo samples must be reweighted to the total integrated luminosity of the data sample of 827.87 pb^{-1} in order to be properly compared with this. This essentially makes the Monte Carlo an Asimov data set as described in section 3.2.1.

Some of the Monte Carlo samples are not produced with an NLO generator, and the samples are reweighted to NLO with the corresponding k-factor, cited in appendix B.

MCID	Process	Generator	Events	Δg_Z^1	$\Delta\kappa_Z$	λ_Z	Cross section [pb]
126071-9	$W^+Z \to l\nu ll$	MC@NLO	50000	0	0	0.13	0.14097
126080-8	$W^-Z \to l\nu ll$	MC@NLO	50000	-0.3	1.0	0.3	0.06772

Table 6.1: The signal Monte Carlo samples used for this study are listed. All samples are obtained from the GRID, and parameters of the samples obtained from Ref. [69, 75, 76].

6.3.2 Pileup reweighting

The Monte Carlo samples used for this study was generated prior to the recording of data. As a consequence, the event pileup (multiple proton collisions in one bunch collision) have not been simulated correctly, and this discrepancy needs to be accounted for. For the data considered in this thesis, the average number of interactions per bunch crossing was below 10, with a grand average of 6 (see figure 6.1). For data collected later in 2011, this number exceeded up to 17.

In figure 6.1 the pileup profiles for data and Monte Carlo is shown. Clearly the Monte Carlo needs to be reweighted.

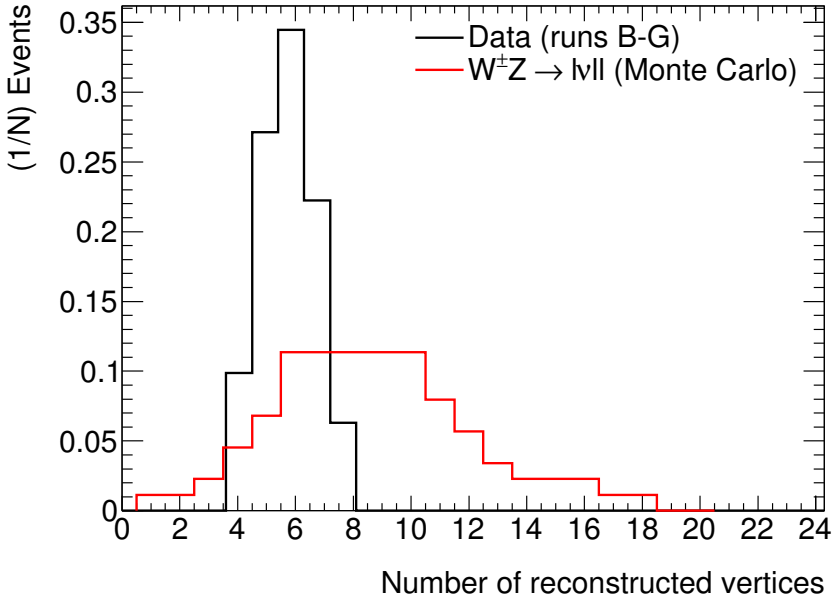


Figure 6.1: The average number of interactions per bunch crossing in data (black) and the $W^{\pm}Z \to ll\nu l$ Monte Carlo used for this study.

The reweighting of Monte Carlo to the data pileup conditions is done using the ATLAS tool for pileup reweighting [78].

6.4 Preselection

The event preselection is split up in three parts. First the events are »cleaned« to ensure that (a) the detector is in a condition such that the event is sensible (b) the event itself fulfills some minimum requirements for further investigation. Next the physical objects muons, electrons and E_T^{miss} are selected according to a closer definition of what specifies »good« objects. Finally it is required that a WZ can be reconstructed from the physics objects. The steps of the preselection will be explained in detail in the following.

6.4.1 Event cleaning

Before even looking at the physical objects in a candidate event, it is tested if the event fulfills a number of minimal criteria.

Luminosity and Good-run-lists

The event is required to be in the so-called Good-run-list (GRL), which is a run list prepared by the ATLAS data quality group, assembled from information from all the detector subsystems. If the event is in the GRL, it is simply deemed »good«, that is, valid for analysis. It is from the GRL, that the total integrated luminosity is calculated [53].

LAr calorimeter error

In a large part of the data taking period, the LAr calorimeter was affected by a partial failure. Resultingly the energy could not be be measured within $0 < \eta < 1.4$ and $-0.8 < \phi < -0.6$. Since the LAr calorimeter is vital to the reconstructed of the physics objects used for the analysis, events with objects within the selected area are removed.

Trigger

Three trigger are used to select events that contains a minimum of one muon or electron. The triggers are EF_e20_medium, EF_mu18_MG or EF_mu40_MSonly_barrel, of which at least one must have fired.

Number of tracks

The primary vertex (the reconstructed vertex with the highest track multiplicity) must have at least two reconstructed tracks associated to it.

6.4.2 Object selection

If the event passes the cleaning criteria, the physical objects are selected according to a series of cuts. The lepton selection are largely based on the selection cuts in Ref. [37].

Both electrons and muons must originate from the primary vertex. The longitudinal impact parameter of the corresponding lepton track with respect to the primary vertex is demanded to be $|z_0| < 10\text{mm}$.

Muons

The reconstructed muons are all Combined muons from MUid, and are required to have $|\eta| < 2.5$ and $p_T > 15\text{GeV}$. To minimize background from jets, muons must be isolated: The scalar sum of the track p_T in a cone of $\Delta R < 0.2$ around the muon is required to be less than 10 % of the muon p_T.

Electrons

Electrons are formed with the standard reconstruction algorithm described in section 5.3.4, and are required to be in the region $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ to select central electrons and avoid the transition region between the barrel and end-cap. The electrons are required to have $p_T > 15$ GeV. The sum of energy deposited in the calorimeter in a cone of $\Delta R < 0.3$ around the electron is required to be less than 4 GeV to ensure isolation.

Missing transverse energy

The E_T^{miss} used for this study is MET_RefFinal [79]. It is built from other reconstructed objects as described in section 5.11. It is required that E_T^{miss} >20 GeV.

6.4.3 WZ reconstruction

Once the leptons and the E_T^{miss} are extracted from the event, they can be combined into a WZ pair, provided that a number of conditions on the reconstruction are fulfilled.

Z reconstruction

The event is scanned for Z candidates, which is defined as every pair of leptons in the event satisfying:

- Opposite electric charge
- Same flavour
- Combined invariant mass compatible with Z mass [3]: $|M_{ll} - 91.1876| < 20$ GeV

If multiple Z candidate exist for an event, the Z candidate most compatible with the $m_Z = 91.1876$ GeV is chosen.

W reconstruction

The W is reconstructed from the remaining lepton after the Z reconstruction, meaning that the leading, leftover lepton is combined with the E_T^{miss} of the the event to reconstruct the W. It is demanded that:

- The lepton used for the W reconstruction must have a $p_T > 20$ GeV, and fulfill the *tight* lepton criteria.
- The reconstructed W must have a transverse mass [3]:

$$m_T(W) = 2E_T^{miss} p_T(\text{lepton})(1 - \cos\theta^T) > 10 \text{ GeV}.$$

With these criteria a maximum of one W is reconstructed per event.

6.4.4 Preselection performance

The event preselection is performed on Monte Carlo and data. In order for the reader to get an idea of where in the preselection most background is removed, the preselection summary table (table 6.2) is split up in the following categories (GRL is required before this cut flow table begins):

1. Event cleaning.
    - No LAr Error.
    - Event trigger.
    - Number of tracks.
2. The event must have $E_T^{miss} > 20\text{GeV}$.
3. Three good leptons
    - At least three leptons must exist, which fulfills the criteria of section 6.4.2.
4. Cut on Z.
    - At least one Z which fulfills the criteria of section 6.4.3 must be reconstructed
5. Cut on W.
    - At least one W which fulfills the criteria of section 6.4.3 must be reconstructed

The performance of the preselection is summarized in table 6.2. It is seen that the preselection removes nearly all background. A total of 100 ± 1.1 events are expected (signal plus background) and 110 ± 10 events are observed after all cuts.

Cut	Data	$W^{\pm}Z$	ZZ	$t\bar{t}$	W/Z+jets	WW+Dijet (All other)
Event cleaning	$886,738 \pm 942$	1446 ± 5.8	1838 ± 4.5	$627,746 \pm 70$	$10,130 \pm 72.1$	$269,756 \pm 1456$
$E_T^{miss} > 20$ GeV	$200,438 \pm 448$	1016 ± 5.4	881 ± 3.1	$54,831 \pm 65.2$	4704 ± 43.2	$104,437 \pm 886$
Three good leptons	3070 ± 55	144 ± 1.5	16.2 ± 0.4	148 ± 3.4	7.8 ± 0.70	155 ± 35.1
Z cut	961 ± 31	128 ± 1.2	14.2 ± 0.4	29 ± 1	6.7 ± 0.61	1 ± 0.3
W cut (**Total**)	110 ± 10	88.0 ± 0.9	6.8 ± 0.3	4.5 ± 0.60	0.3 ± 0.12	0.1 ± 0.09

Table 6.2: Summary of cut flow efficiency for data and Monte Carlo (after pileup reweighting). Total expected after cuts (signal plus background) is 100 ± 1.1, with 110 ± 10 observed in data. Additional information for signal MC is in table 6.1. Additional information for all of the used MC samples are put in appendix B.

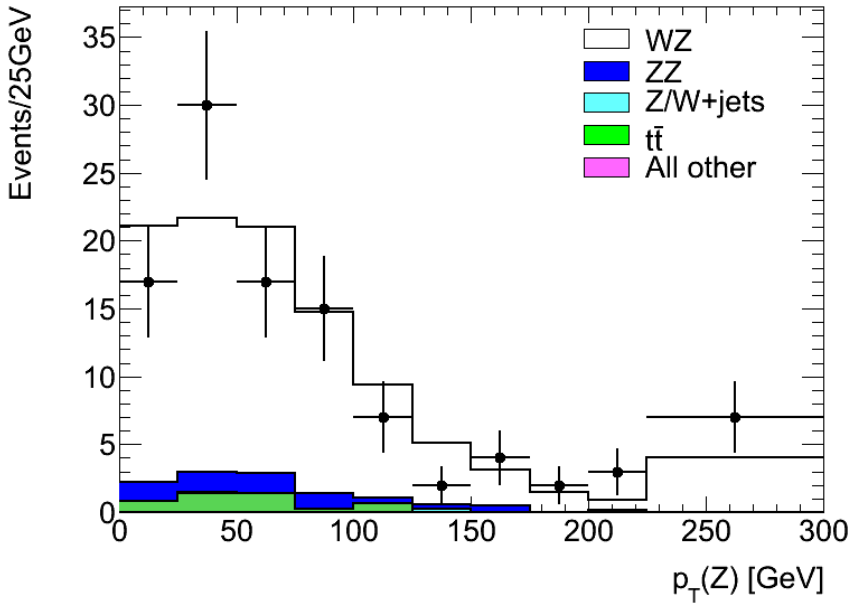


Figure 6.2: The $p_T(Z)$ distributions for all Monte Carlo samples combined and compared to data after cuts.The last bin includes overflow. The »All other« category covers Dijet and WW, which are seen to be practically excluded, as also seen from table 6.2.

In figure 6.2 the central variable $p_T(Z)$ is shown, after the cut based selection.

6.5 Multivariate selection

In table 6.2 and from figure 6.2 it is seen that a cut based analysis does pretty well in discriminating between the signal and background for this analysis. As the reconstruction has provided well defined physical objects, and a sensible estimation of the reduced background exists, one can in principle extract the TGCs at this point. It is, however, desirable to remove as much background as possible before extracting the TGCs, as this can provide tighter limits. This is done using a multivariate technique known as boosted decision trees.

The scope of this multivariate analysis is not limited to tightening the limits of the TGCs. The three lepton final state is used in other present searches for new physics (for example SUSY [80]), and as multivariate selection methods becomes more and more common in diboson physics studies [81] to distinguish signal from background, it is interesting in itself to see if the signal to noise ratio can improve significantly with the application of multivariate analysis.

It is important to note, that care needs to be taken when using multivariate methods. Since multivariate methods are less transparent than a simple cut based selection, it is therefore essential that one remains completely blind to data, when tuning the multivariate

method of choice to the particular analysis. All tuning of the multivariate selection method is therefore done exclusively on Monte Carlo.

6.5.1 Boosted decision trees

Boosted decision trees (BDT) is a multivariate analysis method designed for discriminating between signal and background using several variables at a time. The key idea is that an algorithm is *trained* to distinguish between signal and background by letting it look at several variables at a time, and then »learn« what a signal looks like in the combination of all the variables. The motivation for doing so is to be able to combine many variables which are weakly sensitive into one powerful [81].

The BDT selection performed in this study in based on the implementation done in Toolkit for Multivariate Data Analysis (TMVA) [82], which is shipped with the ROOT [42] package. TMVA can do many types of multivariate analysis, and every type of analysis provides several sub-options. Since BDT is a rather transparent method, it is selected for the study. A single decision tree is built as depicted in figure 6.3.

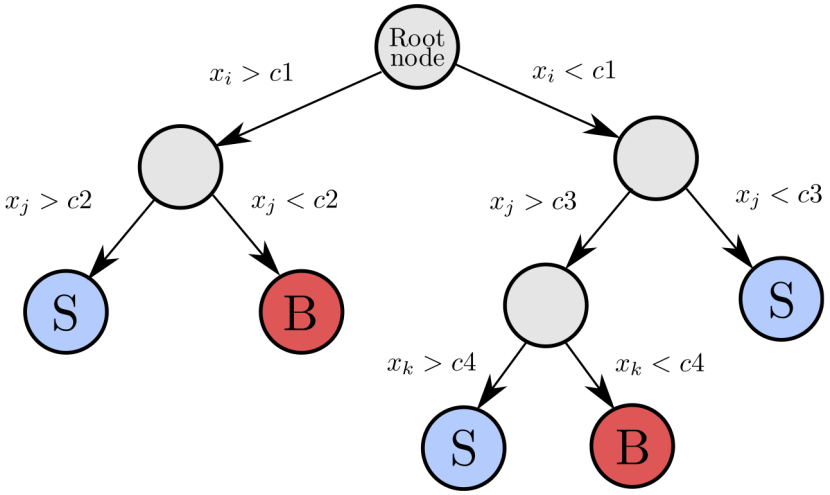


Figure 6.3: Schematic of a decision tree. Starting from the »root« node, a series of splits used on the discriminating variables (x_i) ensures good separation of Signal and Background, by, at each node, using the variable and cut (c_i) that gives the best separation at this specific node. Figure taken from Ref. [34].

An event propagated through the tree will then be assigned a value of $h(\vec{x}) = 1$ if deemed signal or $h(\vec{x}) = -1$ if deemed background, where $\vec{x}$ is a vector of the input variables (classifiers).

Separating signal from background using a single decision tree like figure 6.3 poses an immediate problem. If the *training sample* suffers from some statistical fluctuation making an otherwise favorable variable bad to cut on, one might disfavor it with respect to others, and obtain a suboptimal discrimination routine (or in the converse scenario

a too-good-to-be-true routine, known as overtraining). This is dealt with by growing a large number of trees (a so called forest), and assign a weight to the m'th tree (α_m) as a function of the misclassification rate (e_m) of the previous tree:

$$\alpha_m = \log\left(\frac{1-e_m}{e_m}\right). \tag{6.1}$$

The total BDT response is then the weighted average of the (N) individual responses:

$$y_R = \frac{1}{N}\sum_N \alpha_m h_M(\vec{x}). \tag{6.2}$$

Growing such a multitude of trees is known as *boosting*, and the algorithm indicated here, is one of the TMVA built-in algorithms. It is called AdaBoost [83], and is the one used for this study – again because of transparency[1].

6.5.2 Training and testing

The training of a BDT consists of defining a so-called *splitting criteria* for each node. It is clear from figure 6.3 that one should choose the path through the tree, that provides the best separation, but this separation needs to be defined. The TMVA package provides a number of separation criteria, the default being the Gini index:

$$G = p(1-p), \tag{6.3}$$

where $p = \frac{S}{S+B}$ is called the purity. The splitting criterion is decided for each node, for every single variable by optimizing the increase in Gini index between a node and the sum of the Gini indexes of its two daughters, weighted by their relative fraction of events.

The training uses half of the events passed on from the cut based analysis, denoted the *training sample*. The remaining part will be used to test the performance of the BDT, and is denoted the *testing sample*.

As seen in the previous section, the most important backgrounds to remove, are the Z/W+jets, and $t\bar{t}$. The BDT will therefore be trained to remove these three types.

6.5.3 Selection variables

Six variables are chosen based on an expectation of what will separate signal from background better.

Transverse momentum of leading jet

The p_T of the leading jet (the jet with the highest p_T) is used as a classifier. Since the LO hard process involves no quarks in the final state, it is expected that the signal distribution will peak at a low p_T, whereas the background distribution will be a little more flat.

[1] An easy-to-read proof of the AdaBoost algorithm is provided in Ref. [84]. It contains numerous football references.

Rank	Variable	Separation ($S/\sqrt{S+B}$)
1	E_T^{miss}/E_T^{Total}	0.1656
2	p_T(Leading jet)	0.1275
3	M_T(W)	0.0756
4	ΔR(leptons)	0.0736
5	Number of jets	0.07028

Table 6.3: $S/\sqrt{S+B}$ for each of the BDT selection variables, as calculated by TMVA.

Transverse mass of W

Using the transverse mass of the W reconstructed from the lepton and the E_T^{miss}, one is able to distinguish properly and badly reconstructed Ws. As the $t\bar{t}$ background will have a wider range of E_T^{miss} because of the (min.) two neutrinos in the final state, and as E_T^{miss} is essential in the reconstruction of $m_T(W)$, this variable is expected to be particularly good at eliminating $t\bar{t}$ background. The same goes for the two other variables involving E_T^{miss}.

Distance between W lepton and E_T^{miss}

The $\Delta R(leptons)$ is the ΔR distance between the lepton from the decayed W and the E_T^{miss} of the event (for signal this is the neutrino).

Number of jets (jet multiplicity)

The total number of reconstructed jets is used as a discriminating variable since signal events are characterized by low QCD activity.

Missing transverse energy

The E_T^{miss} normalized to the scalar sum of E_T of all visible objects is used.

Performance

The distributions of selection variables are shown in figures 6.4, and 6.5 where it is clearly seen that some discriminates better than others.

To quantify the notion of which variable separates signal from background best, the separation variable $S/\sqrt{S+B}$, where S and B are integrals of the signal and background distributions, is used. The ranking of classifiers according to this variable is shown in table 6.3.

Figure 6.4: The BDT variables (left) E_T^{miss}/E_T^{Total} and (right) ΔR(leptons)

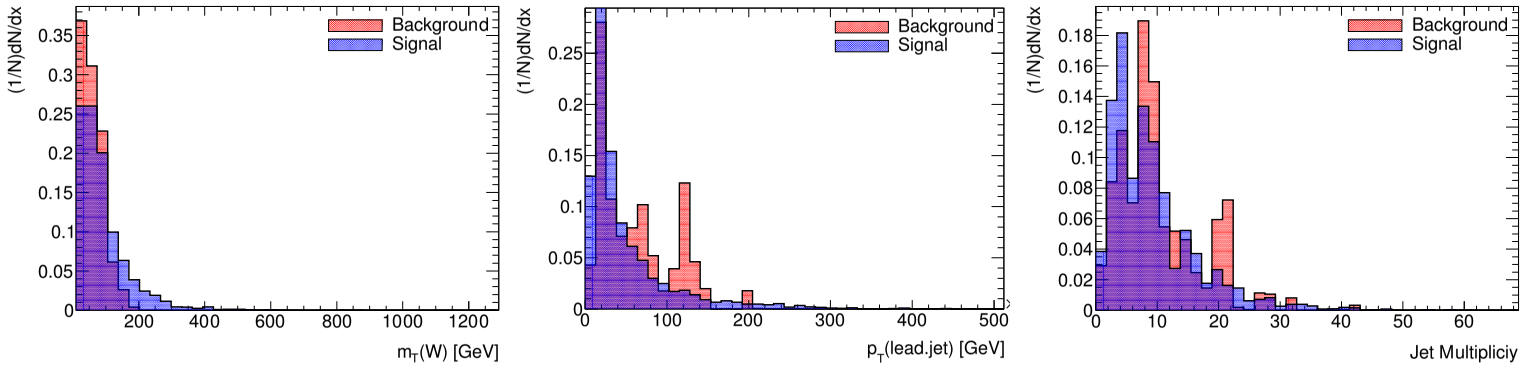


Figure 6.5: The BDT variables (left) M_T(W), (middle) p_T(Leading jet) and (right) p_T(Leading jet). Separation efficiencies are quoted in table 6.3.

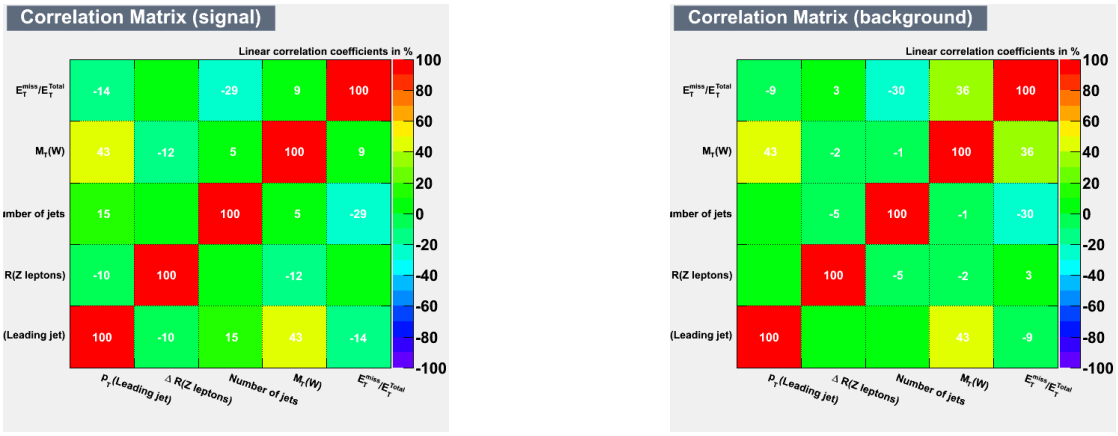


Figure 6.6: Correlation matrix of selection variables given signal (left) and background (right). No correlation coefficients are over 50.

6.5.4 Correlations

It is not a goal in itself to have as many variables as possible in a multivariate analysis. The goal is to obtain a good separation of signal and background, using the most transparent analysis as possible. Correlations between selection variables in signal or background could point towards redundancy, and therefore possible sources for overtraining. Secondly one must also look for correlations between the final BDT classifier and $p_T(Z)$, the variable the TGC extraction will eventually be performed on. If the BDT classifier correlates strongly with the $p_T(Z)$, the usage of this could introduce an undesired bias.

In figure 6.6 the linear correlation coefficients between selection variables is shown in a matrix form, on a scale from -100 to 100, where 0 means no correlation, -100 means large negative correlation and 100 means large positive correlation.

In figure 6.6 it is seen, that there are no correlation coefficients above 50 between the chosen selection variables for neither signal nor background. The possible correlations between the combined BDT classifier and $p_T(Z)$ is dealt with in the next chapter.

6.5.5 BDT response

When combining weak classifiers into a BDT response using TMVA, a number of parameters can be set. As mentioned, the AdaBoost algorithm is chosen, with the use of Gini Index for separation. With this choice in mind, the parameters which could be fine tuned to optimize the BDT (and the chosen values) for this study are:

- Number of trees (1000)
- Minimum number of events at end nodes (50)
- Maximum tree depth (3)

The used values are suggested defaults. No optimization procedure regarding these parameters were undertaken.

The BDT response for this particular choice of values is shown for the testing and training samples overlaid in figure 6.7.

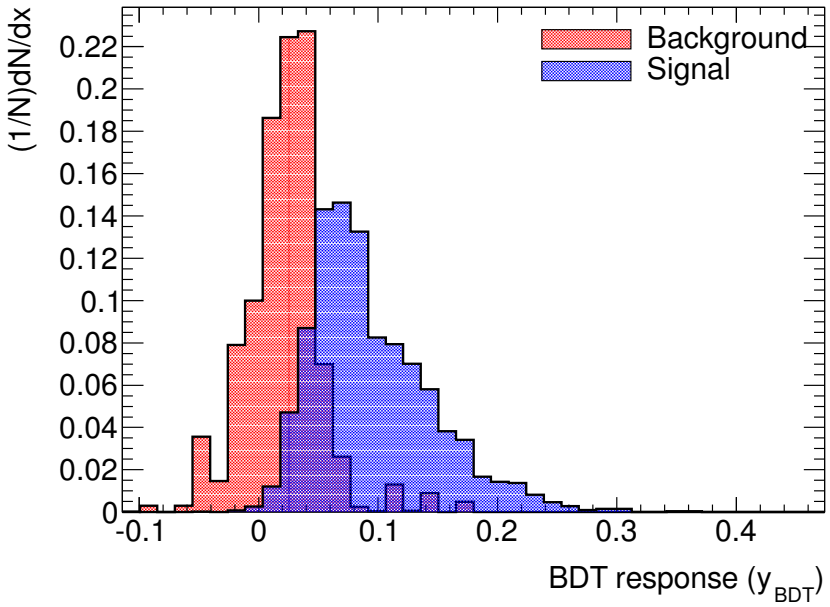


Figure 6.7: The BDT response from (blue) signal and (red) background. It is important to note that the distributions are normalized to one, and does not take into account difference in the expected amount of signal and background events. A by-eye assessment tells that the optimal BDT cut lies somewhere between $y_{BDT} = 0.00$ and $y_{BDT} = 0.05$.

6.5.6 Optimization of selection cut

The decision on where the cut on the BDT variable is made, should be done following a number of criteria, which will be addressed one by one:

- The cut on BDT should improve the significance, $S/\sqrt{S+B}$.
- The cut on BDT should aim to improve the sensitivity for this specific search as much as possible.
- The cut on BDT must not introduce a bias towards any parts of the distribution of $p_T(Z)$, which is the variable this search hinges on.

The signal and background efficiencies are calculated for BDT cut values from $y_{BDT} = -0.1$ to $y_{BDT} = 0.4$, and the significance is calculated as $S/\sqrt{S+B}$, now weighted

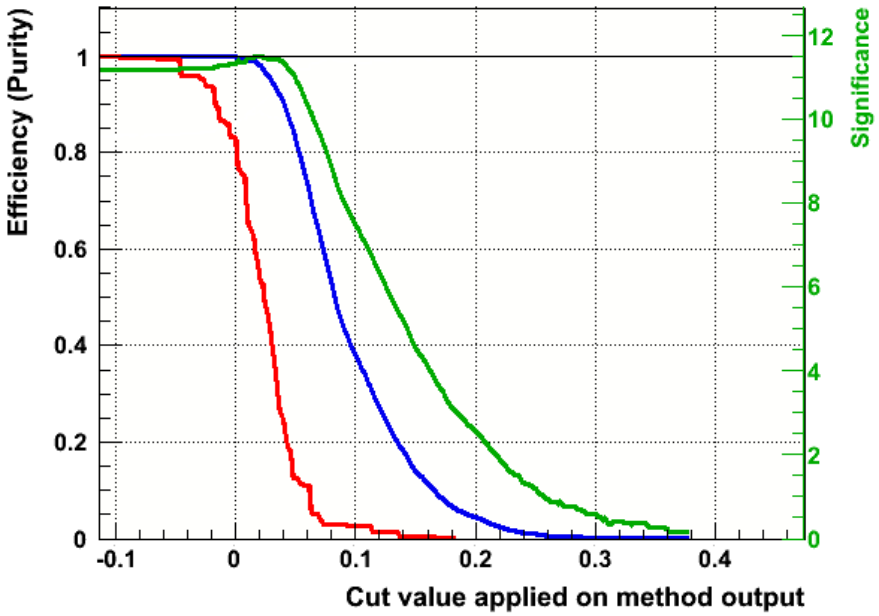


Figure 6.8: Signal (blue) and background (red) efficiencies along with significance (green) as a function of cut applied on y_{BDT}. It is seen that a slight improvement can be made using a cut of $y_{BDT} = 0.02$

	$\lvert\Delta\kappa_Z\rvert$	$\lvert\lambda_Z\rvert$	$\lvert\Delta g_1^Z\rvert$
With background	0.90	0.093	0.14
Without background	0.85	0.090	0.13

Table 6.4: Expected sensitivity before BDT and for a perfect BDT (without background).

according to the expected number of signal and background events after the cut based selection. All three are plotted as a function of cut in y_{BDT} in figure 6.8.

The maximal improvement of significance is at a cut value of $y_{BDT} = 0.02$, as figure 6.8 also indicates.

Addressing the second criteria, the sensitivity for this specific study is taken as the size of the maximum likelihood CI, given an Asimov MC sample at SM. Before optimizing with respect to this criteria, an assessment of the largest possible gain in sensitivity, given SM, is in place. Therefore the sizes of the CIs are calculated (a) without applying any BDT cut, and (b) by removing all simulated background, leaving all signal MC. This last sensitivity can be regarded as the expected sensitivity for a »perfect« BDT which would remove all background before removing any signal. The expected sensitivities are summarized in table 6.4. As seen in table 6.4 the expected sensitivity can be improved by less than 5 % by removing all background without removing signal. Since any BDT cut

y_{BDT}	Background efficiency	Signal efficiency	$\lvert\Delta\kappa_Z\rvert$	$\lvert\lambda_Z\rvert$	$\lvert\Delta g_1^Z\rvert$
-1	100 %	100 %	0.90	0.093	0.14
-0.02	93 %	100 %	0.90	0.093	0.14
-0.01	87 %	99 %	0.90	0.093	0.14
0.00	83 %	99 %	0.89	0.093	0.14
0.01	75 %	99 %	0.89	0.093	0.14
0.02	67 %	98 %	0.88	0.093	0.14
0.03	40 %	95 %	0.88	0.093	0.14
0.04	32 %	84 %	0.89	0.093	0.14

Table 6.5: Expected sensitivity with background removal with BDT.

	Data	$W^{\pm}Z$	ZZ	$t\bar{t}$	W/Z+jets	WW+Dijet
Exp.	110 ± 10	86.2 ± 0.9	4.4 ± 0.2	3.0 ± 0.2	0.3 ± 0.12	0.1 ± 0.09

Table 6.6: BDT efficiency for data and Monte Carlo signal and background. The reader can compare to table 6.2.

of $y_{BDT} > -0.014$ will remove signal (see figures 6.7 and 6.8), the expected sensitivities of the last row in table 6.4 will not be reached. It is thus a sensible choice to calculate the expected sensitivity for a series of BDT cuts $y_{BDT} \geq -0.02$, including $y_{BDT} = 0.02$, which maximizes significance, to see when the sensitivity no longer improves from removal of background, due to removal of signal. Such a calculation of sensitivities for BDT cuts up to $y_{BDT} = 0.04$ is carried out, and shown along with background and signal efficiencies in table 6.5. It is seen that the expected sensitivity is optimal somewhere between $y_{BDT} = 0.00$ and $y_{BDT} = 0.03$, but also that only a small improvement of the $\Delta\kappa_Z$ limit is expected. No improvement is expected on the Δg_1^Z or λ_Z limits. As $y_{BDT} = 0.02$, the cut that maximized significance, is one of the cut candidates giving the optimal sensitivity, this is chosen as the BDT cut. Choosing cuts higher than the ones in table 6.5, will only remove sensitivity, as relatively more signal will be removed.

The bias towards parts of the $p_T(Z)$ distribution is also studied. If a correlation between the BDT variable and $p_T(Z)$ exists, the overlapping area between response values of 0.0 and 0.1 on figure 6.7 could contain events of relatively high $p_T(Z)$. An estimate of this correlation for signal is shown in figure 6.9, where it is seen that no such correlation exists.

The final BDT cut is decided to be at $y_{BDT} = 0.02$, since this is the value that optimizes the $S/\sqrt{S+B}$. Acceptances for signal and background Monte Carlo is computed for this cut, and shown in table 6.6, with the different background contributions separated.

It is seen that the BDT cut does not remove any data events, nor any of the W/Z+jets or WW+Dijet background expectation.

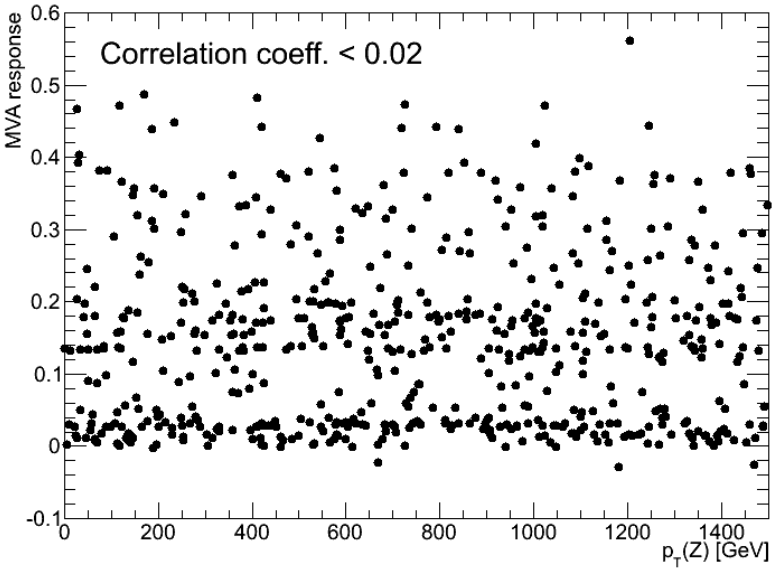


Figure 6.9: Correlation between MVA response and $p_T(Z)$ for signal (testing and training) samples. The linear correlation coefficient is less than 0.02.

CHAPTER 7

TGC EXTRACTION

As described in the previous chapter, background is removed first with a cut based selection, then with a multivariate analysis based on boosted decision trees. In order to evaluate the usage of boosted decision trees, limits are calculated both with and without the use of this method. Limits are calculated using Neyman Construction with Feldman Cousins ordering, and limits from maximum likelihood fits are given for comparison. Predictions for limits on all couplings are calculated using the same procedure, with Monte Carlo instead of the 827.87 pb^{-1} of real data.

7.1 Cut based selection only

The fit is carried out before any BDT selection is done. This is essentially a fit of the data presented in figure 6.2, but the fit uses only four bins: $p_T(Z) = \{25, 75, 150, 200, 1500\}$ (the figures 7.1 and 7.4 shows data and MC all the way down to $p_T(Z) = 0$ GeV to enable comparison with cut flow tables). The data and Monte Carlo for signal and background in the selected binning is shown in figure 7.1.

The fit is carried out both with a Monte Carlo Asimov sample generated at SM expectation, in place of real data, and the real data. Likelihood curves for the maximum likelihood fits to real data are presented in figure 7.2 and the fits using the Monte Carlo Asimov sample in figure 7.3.

The obtained limits at a 95% CL is given in table 7.1. As expected, the limits using the maximum likelihood fit tend to undercover slightly, but in all three couplings the limits given the two methods of fitting is seen to be compatible. The observed limits is seen to differ from the SM prediction. This is due to the fact that more events than expected was observed, and especially in the sensitive high $p_T(Z)$ region (as seen in figure 7.1). As this results in a double minimum structure of the negative log-likelihood ratio, the 95 % CI is larger than expected. All limits on the TGCs are compatible with the SM at a 95 % CL.

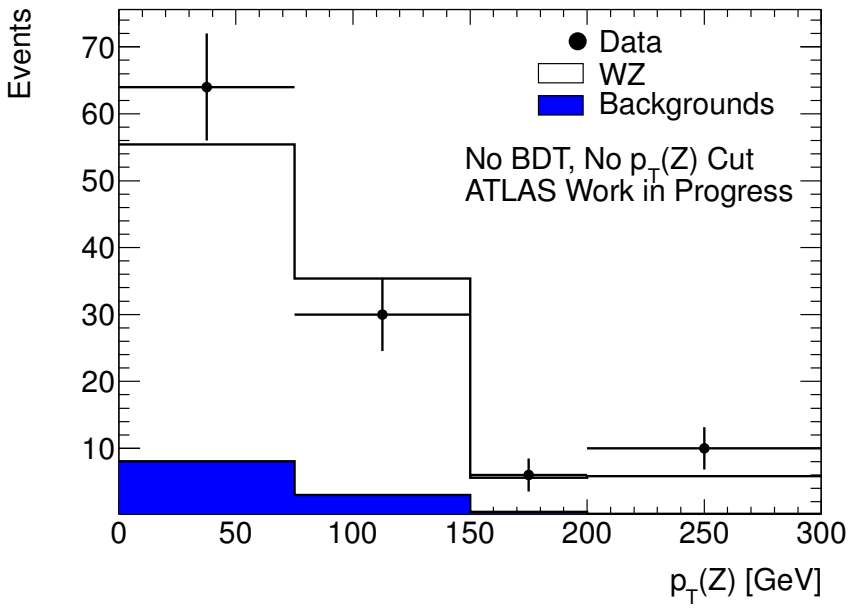


Figure 7.1: Data and Monte Carlo signal and background prior to a cut on BDT output. The primary axis is truncated at $p_T(Z) = 300$ GeV, but the last bin contains overflow. Limits on TGCs are quoted in table 7.1

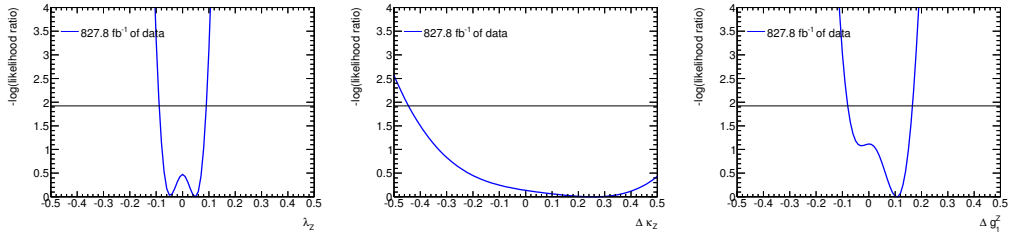


Figure 7.2: Likelihood curves for fits of data corresponding to 827.87 pb^{-1} of integrated luminosity, without BDT cut, for λ_Z (left), $\Delta\kappa_Z$ (middle) and Δg_1^Z (right). Corresponding limits (along with expected limits) is quoted for both likelihood fit and Neyman Construction in table 7.1.

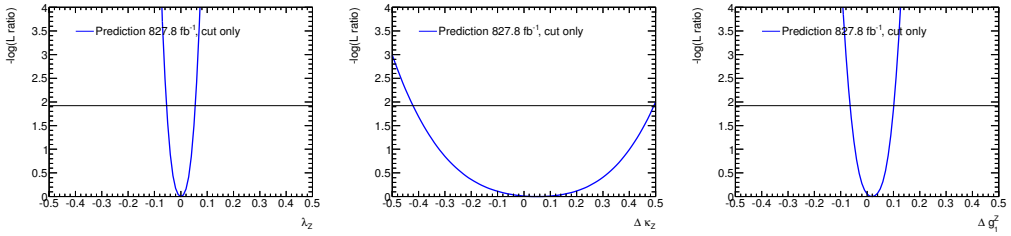


Figure 7.3: Likelihood curves for Monte Carlo expectation, without BDT cut, for λ_Z (left), $\Delta\kappa_Z$ (middle) and Δg_1^Z (right). Corresponding limits (along with limits for fits to 827.87 pb^{-1} of data) is quoted for both likelihood fit and Neyman Construction in table 7.1.

	λ_Z	$\Delta\kappa_Z$	Δg_1^Z
Likelihood prediction	[-0.046 , 0.047]	[-0.42 ,0.48]	[-0.040 , 0.100]
Neyman prediction	[-0.053 , 0.053]	[-0.45 , 0.48]	[-0.062 , 0.100]
Binned Likelihood	[-0.089 , 0.090]	[-0.45 , 0.69]	[-0.10 , 0.18]
Neyman Construction	[-0.092 , 0.093]	[-0.45 , 0.69]	[-0.11 , 0.19]

Table 7.1: Limits on charged triple gauge couplings, for cut-based selection only (see figure 7.1). The two upper rows gives the MC prediction (using the Asimov SM data sample) and the two lower rows using 827.87 pb^{-1} of data recorded with the ATLAS detector. Corresponding likelihood curves are shown in figures 7.2 and 7.3.

7.2 With BDT selection

The BDT selection cut of 0.0203 obtained in section 6.5.6 is applied, and the fit is carried out. Again both for a Monte Carlo Asimov sample generated at SM expectation, in place of real data, and the real data. Monte Carlo signal and background compared to data after applying the BDT cut is shown in figure 7.4.

Likelihood curves for the maximum likelihood fits to real data are presented in figure 7.5 and the fits using the Monte Carlo Asimov sample in figure 7.6.

The obtained 95% CIs is given in table 7.3. In the same way as before, the observed CIs are larger than the expected due to observation of more events than expected. All limits are compatible with the SM.

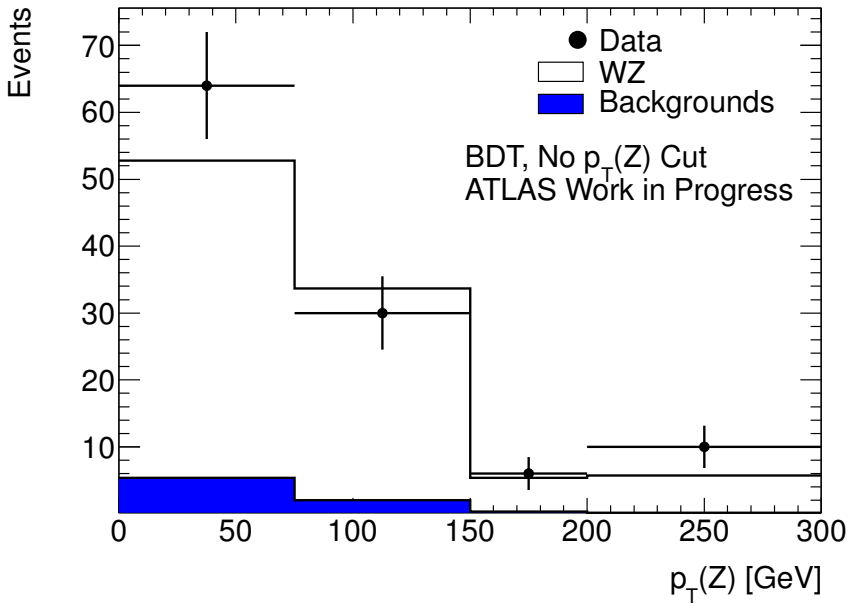


Figure 7.4: Data and Monte Carlo signal and background after imposing a cut on BDT output. The primary axis is truncated at $p_T(Z) = 300$ GeV, but the last bin contains overflow. In this figure, all backgrounds are collected in the blue histogram. Limits on TGCs are quoted in table 7.3.

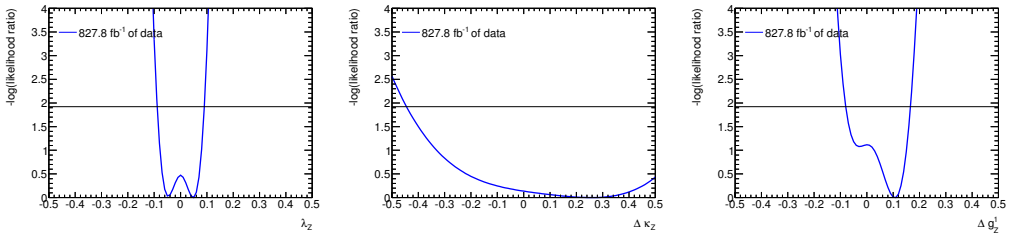


Figure 7.5: Likelihood curves for fits of data corresponding to 827.87 pb^{-1} of integrated luminosity, with additional cut on BDT output, for λ_Z (left), $\Delta\kappa_Z$ (middle) and Δg_1^Z (right). Corresponding limits (along with expected limits) are quoted for both likelihood fit and Neyman Construction in table 7.3.

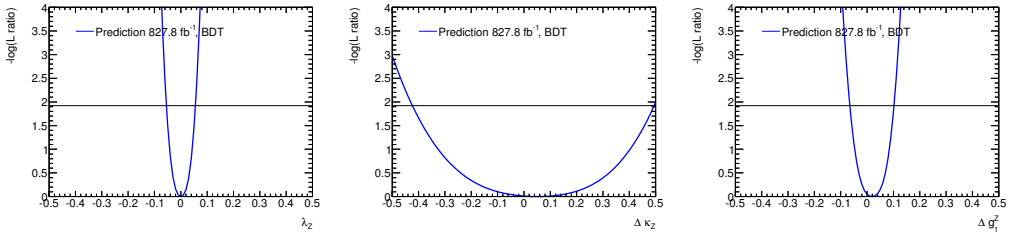


Figure 7.6: Likelihood curves for Monte Carlo expectation, with additional cut on BDT output, for λ_Z (left), $\Delta\kappa_Z$ (middle) and Δg_1^Z (right). Corresponding limits (along with limits for fits to 827.87 pb^{-1} of data) is quoted for both likelihood fit and Neyman Construction in table 7.3.

	λ_Z	$\Delta\kappa_Z$	Δg_1^Z
Likelihood prediction	[-0.046 , 0.047]	[-0.41 , 0.47]	[-0.040 , 0.100]
Neyman prediction	[-0.054 , 0.054]	[-0.46 , 0.47]	[-0.062 , 0.100]
Binned Likelihood	[-0.091 , 0.092]	[-0.43 , 0.70]	[-0.08 , 0.18]
Neyman Construction	[-0.097 , 0.093]	[-0.45 , 0.71]	[-0.11 , 0.19]

Table 7.2: Limits on charged triple gauge couplings with additional cut on BDT output (see figure 7.4). The two upper rows gives the MC prediction (using the SM Asimov data sample) and the two lower rows using 827.87 pb^{-1} of data recorded with the ATLAS detector. Corresponding likelihood curves are given in figures 7.5 and 7.6.

7.3 Prediction for 5 fb^{-1} of integrated luminosity

In order for the reader to compare the method set forth in this thesis to the official ATLAS analysis including 5 fb^{-1}, which is due for the summer conferences of 2012, but not published at the time of writing, a prediction of the limits, given an SM Asimov data sample corresponding to 5 fb^{-1} of data, with a cut on BDT output applied, is made.

The predicted improvement on the limits are approximately on the same scale as predicted by the truth study in chapter 4 (see table 4.3). The corresponding negative likelihood ratio plots are given in next chapter (figure 8.5) along with estimates of expected systematic error with 5 fb^{-1} of integrated luminosity.

	λ_Z	$\Delta\kappa_Z$	Δg_1^Z
Likelihood prediction	[-0.030 , 0.032]	[-0.29 , 0.33]	[-0.052 , 0.055]
Neyman prediction	[-0.030 , 0.033]	[-0.33 , 0.36]	[-0.053 , 0.055]

Table 7.3: Predicted limits for MC corresponding to 5 fb $^{-1}$ of luminosity, given a SM Asimov sample, using cut on BDT output.

7.4 Discussion

Limits at a 95 % CL for the TGCs using 827.87 pb^{-1} of data are reported, both with and without event selection using BDT. Limits are reported using both a likelihood fit, and a Neyman Construction using a likelihood ratio ordering (Feldman Cousins). The reported limits are not as tight as expected from the SM, as more events than expected is observed (see chapter 6). As the TGC sensitivity is the largest for high $p_T(Z)$, observation of more events than expected in the overflow bin, is really the driving factor in the fit. Prior to the BDT cut, ten events is observed in the overflow bin with an expectation of five (see figure 7.1).

Comparing the expected limits with and without BDT cut, the statements about the sensitivity set forth in table 6.5 is confirmed – only a very slight improvement on the size of the $\Delta\kappa_Z$ limit is expected. As no optimization of the cuts should be done on data, the decision on which limits to report is made by comparing the expected sensitivities only.

This discussion will be continued after the estimation of systematic error.

CHAPTER 8

SYSTEMATICS

Until now the calculated error on the TGC extraction have only been of statistical nature. In this chapter sources of systematic error will be identified, and a quantitative analysis of the magnitude of the systematic error on the TGC CIs will be undertaken.

Statistical error arises from the fact, that an experiment only provides a finite data sample. This type of error is expected to diminish as the sample size increases. Systematic error arises from poorly calibrated instruments, theoretical inadequacies and other analysis input encumbered with an uncertainty, and does not *a priori* vanish for large sample sizes.

The systematic error will be treated with a profile likelihood analysis, as introduced in chapter 4. The considered systematic effects will be all parametrized individually with a Gaussian distribution in each of the four bins considered in the fit. The expectation value for each bin will then be a product of the expectation value from equation (4.14) and a number of Gaussians. Systematic effects are all assumed uncorrelated. The individual nuisance parameters are either allowed full correlation between the bins or no correlation, based on a case-by-case assessment. When all nuisance parameters have been parametrized, the fit of chapter 7 is carried out once more with »likelihood ratio« replaced by »profile likelihood ratio«.

The systematic effects considered this way are:

Limited Monte Carlo statistics using the numbers of table 6.2.

Uncertainty from Parton Density Functions estimated using PDF reweighting.

Luminosity uncertainty using the official numbers on the uncertainty from the ATLAS luminosity group.

Shape reconstruction uncertainty using the reconstruction uncertainty averaged over all of phase space, taken from the cross section measurement in Ref. [75].

Uncertainty on determination of TGC weights by using two different samples, generated at non SM TGC values.

8.1 Limited Monte Carlo statistics

As seen in table 6.2, the limited amount of Monte Carlo statistics gives rise to a statistical uncertainty of the predictions of the amount of signal and background events. This uncertainty is much smaller than the corresponding uncertainty in data because of the use of Asimov samples, but they are anyway non-negligible, and must be treated as a source of systematic error. This uncertainty will differ from bin to bin, and is calculated separately for each bin (uncorrelated). This source of systematic error is only considered for signal Monte Carlo. The uncertainty is parametrized as four uncorrelated Gaussians each with a mean value of $\mu = 1.0$ and a standard deviation given as the relative uncertainty in the specific bin; summarized in table 8.1.

	Bin 1	Bin 2	Bin 3	Bin 4
Std. dev. (σ)	0.006	0.009	0.041	0.030

Table 8.1: The standard deviation of the Gaussian parametrization of the error from limited Monte Carlo statistics.

8.2 PDF uncertainty

The usage of parton density functions (PDFs) is a source of systematic uncertainty. The estimation of this uncertainty is done by carrying out the final calculation using different PDFs. Instead of redoing the full simulation with another PDF as input, the samples are reweighted on an event by event basis, by considering the probability to hit the same flavored partons, given the new PDF. This is done by calculating an event weight as the following ratio [85]:

$$w = \frac{f_{PDF2}(f_1, x_1, Q^2)}{f_{PDF1}(f_1, x_1, Q^2)} \cdot \frac{f_{PDF2}(f_2, x_1, Q^2)}{f_{PDF1}(f_2, x_1, Q^2)} \tag{8.1}$$

where f_1 and f_2 denotes the flavor. After applying this weight, MC events generated using PDF1 will be effectively distributed according to PDF2. This method carries some disadvantage to a full simulation, in the sense that it relies heavily on sufficient statistics in the input events. Also tuning to the new PDF with regards to the underlying event (section 3.1.5) is not accounted for. The method is implemented using an interface to the LHAPDF library [86] in particular the CTEQ66 [27], MSTW2008 [87] and NNPDF2.0 [88] PDFs. The procedure follows the official recommendation from PDF4LHC [89].

The PDF uncertainty is only calculated for signal events, but used for both signal and background.

Intra-PDF uncertainty

The fitting of PDFs to data gives rise to an internal PDF-uncertainty. Since a PDF parametrization is a function of a number of uncorrelated parameters, one can estimate this uncertainty by varying these parameters internally in the PDF by $\pm 1\sigma$. The resulting PDF is the so-called Hessian error PDF, which is shipped along with the LHAPDF library. The intra-PDF uncertainty is then calculated by reweighting each event to all error PDFs, and take the mean and standard deviation (68 %) of this procedure as the PDF mean and the intra-PDF error. The mean value is the one calculated with this procedure, and the band around the mean is the standard deviation (68 %). For each bin there is now – given $PDFi$ – a mean value of the differential cross section with a symmetric error $d\sigma = d\sigma_0^{PDFi} \pm \delta(d\sigma^{PDFi})$.

Inter-PDF uncertainty

As we have no *a priori* reason to believe that one specific parametrization of the proton PDF is more correct than the other, an inter-PDF uncertainty is the difference obtained by changing from one PDF to another. On figure 8.1 the $p_T(Z)$ distribution of signal only Monte Carlo is shown, reweighted to the three PDFs CTEQ6.6, MSTW2008 and NNPDF2.0. The bin by bin multiplication factor to get from one histogram to another is simply the mean value of weights of events in the bin.

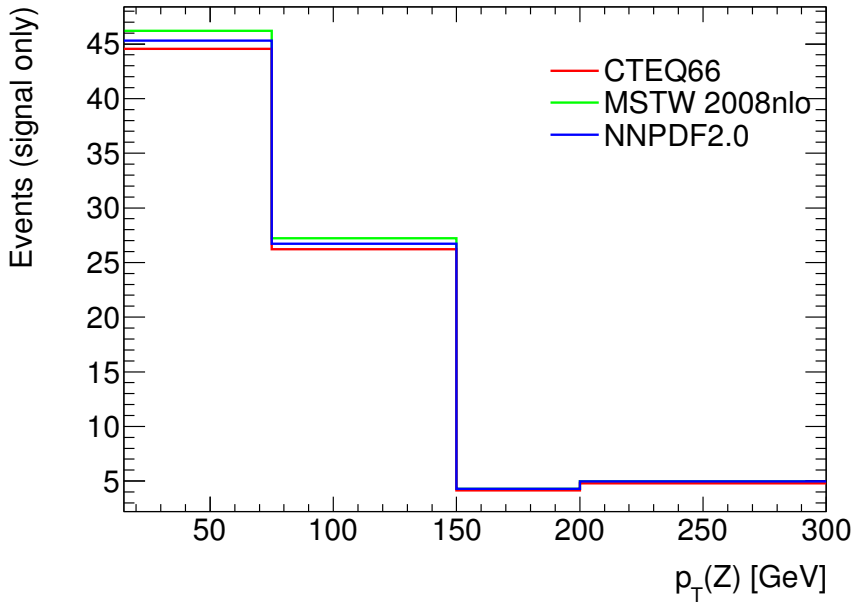


Figure 8.1: PDF reweighting on the $p_T(Z)$ distribution of signal Monte Carlo. Inter-PDF uncertainty is shown, with the means of the three used PDF parametrizations. The envelope is created by also considering internal error on the three PDFs.

	Bin 1	Bin 2	Bin 3	Bin 4
Mean (μ)	1.018	1.022	1.019	1.023
Std. dev. (σ)	0.017	0.021	0.018	0.023

Table 8.2: The bin by bin Gaussian parametrization of the PDF error

Combination

The total PDF error on the $p_T(Z)$ distribution is the envelope [89] of the two uncertainties. This simply states that the limits on the differential cross section in a bin is given by:

- The high limit is the highest of the three $d\sigma_0^{PDFi} + \delta(d\sigma^{PDFi})$
- The lower limit is the lowest of the three $d\sigma_0^{PDFi} - \delta(d\sigma^{PDFi})$

The mean value is the central value between the high and the low limit, and the Gaussian used to parametrize this error is in this way allowed to have a mean value different than one, and can as such shift the central value of the TGC fit.

The mean value and standard deviation of the Gaussian parametrization of the PDF error is calculated, and given bin by bin in table 8.2. The PDF uncertainty is assumed uncorrelated between the bins, and the resulting estimate will be conservative.

8.3 Luminosity uncertainty

The systematic luminosity uncertainty is estimated to be 3.7% [75]. The parametrization of this is straightforward, as a Gaussian of mean $\mu = 1$ and standard deviation $\sigma = 0.037$ is used. This effect is assumed correlated across all bins.

8.4 Shape reconstruction uncertainty

The systematic uncertainty on the $p_T(Z)$ determination from the reconstruction of the physical object electrons, muons and E_T^{miss} stems from several sources, all with different values for each of the final states. The uncertainties are considered here as the relative uncertainty from each of the sources on total cross section. As indicated in chapter 5, the reconstruction uncertainties for the leptons are highly p_T dependent. The approach of using such an average over the whole phase space is therefore an approximation.

The quoted uncertainties are all in percent, and taken from the WZ cross section study in Ref. [75]:

- Muon uncertainty includes: Muon reconstruction efficiency ($^{+1.36}_{-1.33}$) and muon p_T smearing ($^{+0.17}_{-0.17}$). The muon efficiency stems from the fact that the efficiency of the muon reconstruction is introduced in Monte Carlo as a scale factor. The muon p_T smearing covers the fact, that the p_T of the muons are smeared in Monte Carlo to resemble detector resolution. This smearing comes with an uncertainty.

- Electron uncertainty includes: Electron reconstruction and identification efficiency ($^{+2.09}_{-2.03}$), electron energy scale ($^{+0.18}_{-0.18}$) and calorimeter isolation ($^{+1.83}_{-1.78}$).

- E_T^{miss} uncertainty includes uncertainties from the objects it is reconstructed from (see section 5.3.6). This includes: The uncertainty on the topo cluster energy scale ($^{+0.56}_{-0.55}$), a resolution uncertainty ($^{+0.31}_{-0.31}$), uncertainty on jet energy scale ($^{+0.34}_{-0.34}$), a pileup uncertainty ($^{+0.57}_{-0.57}$) and a scale uncertainty ($^{+0.18}_{-0.18}$).

The uncertainties are assumed mutually uncorrelated, and are therefore combined in quadrature to $^{+3.24}_{-3.16}$. The combination is parametrized with a Gaussian of $\mu = 1.04$ and $\sigma = 0.032$, and as mentioned this uncertainty is assumed fully correlated between the bins.

8.5 Uncertainty on determination of TGC weights

The event-by-event TGC weights determined along with the MC@NLO samples are associated with an uncertainty. Currently no official recommendation to deal with this uncertainty exists. An estimation of this uncertainty will be made by comparing samples generated at different TGC values. Signal samples of $W^+Z \to lll\nu$ generated at two different coupling values (see table 8.3) are reweighted to SM prediction. The difference between the two predictions, in each bin, will have a contribution from (a) Uncertainty

MCID	Process	Generator	Events	Δg_1^Z	$\Delta\kappa_Z$	λ_Z	Cross section (pb)
126071-9	$W^+Z \to l\nu ll$	MC@NLO	50000	0	0	0.13	0.06772
126089-97	$W^+Z \to l\nu ll$	MC@NLO	50000	-0.3	1.0	0.3	0.05129

Table 8.3: The signal Monte Carlo samples used for estimating the uncertainty on determination of TGC weights. Parameters of the samples obtained from Ref. [69], [75] and [76]

	Bin 1	Bin 2	Bin 3	Bin 4
Std. dev. (σ_W)	0.02	-	-	0.08

Table 8.4: The standard deviation of the Gaussian parametrization of the error on the TGC weights determination.

from determination of TGC weights, (b) A statistical uncertainty from limited Monte Carlo. As the statistical part is already estimated (section 8.1), this component must be subtracted in order to avoid double counting of this effect. In figure 8.2 the samples generated at the different TGC values are shown with error bars indicating the scaled error from limited MC statistics. As the two errors are assumed uncorrelated, it is assumed that the total error (σ_{tot}; half the difference of the counts in each bin), is the TGC weight error (σ_W) and the limited Monte Carlo error (σ_{MC}) combined in quadrature. The TGC weight error is thus calculated as:

$$\sigma_W = \sqrt{\sigma_{tot}^2 - \sigma_{MC}^2}. \tag{8.2}$$

This error turns out to be non-negligible in two bins (one and four), and is parametrized as two uncorrelated Gaussians, with $\mu = 1$ and standard deviations as indicated in table 8.4.

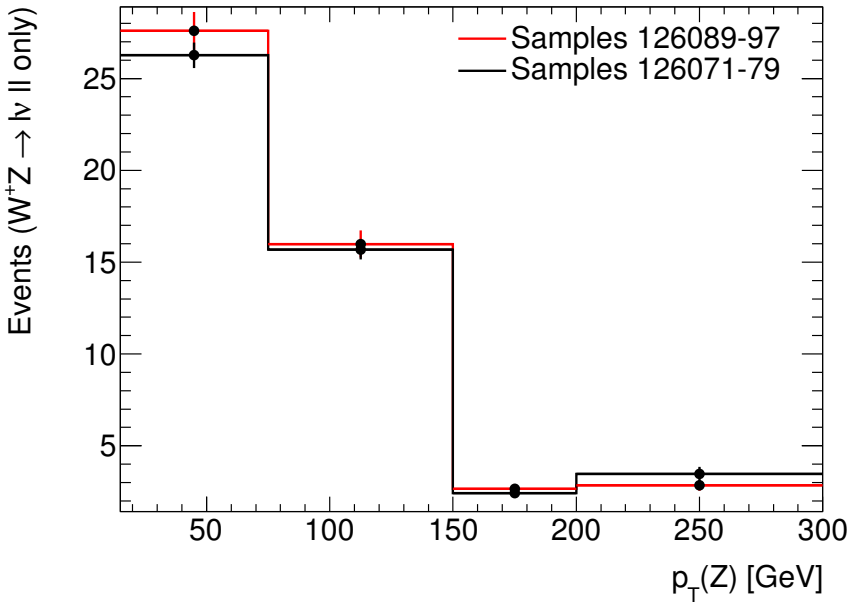


Figure 8.2: Estimation of the systematic error from determination of the TGC weights. The red histogram is for the samples 126089-97 generated at $\lambda_Z = 0.13$ and the black histogram is for the samples 126089-97 generated at $\Delta g_1^Z = -0.3$, $\Delta\kappa_Z = 1.0$ and $\lambda_Z = 0.3$. Both are reweighted to SM level. The error bars represent the statistical error from finite sample size, and the systematic error from determination of the TGC weights is calculated from the residual difference between the histograms.

8.6 Results with combined systematics

All sources of systematic error is combined by multiplying the bin expectation (equation (4.14)) with the Gaussian parametrization of the relevant nuisance parameters. In table 8.5 the distributions entering the construction of the profile likelihood is summarized. If the mean value for a specific error is not quoted, it is $\mu = 1.0$.

The TGC extraction, with the imposed BDT cut, of chapter 7 is repeated with the combined systematics. The profile likelihood ratios for expectation from Asimov samples generated at SM are shown in figure 8.3, with the likelihood ratios of the previous chapter shown for comparison (in some cases overlapping). The corresponding fits of the data sample of 827.87 pb^{-1} are shown in figure 8.4. The limits at a 95 % CL are given in table 8.6, both using a Neyman Construction with Profile Likelihood ratio ordering and ordinary maximum likelihood fit.

	Bin 1	Bin 2	Bin 3	Bin 4
MC stat (Uncorrelated)				
σ	0.006	0.009	0.041	0.030
PDF (Uncorrelated)				
μ	1.018	1.022	1.019	1.023
σ	0.017	0.021	0.018	0.023
Weights (Uncorrelated)				
σ	0.02	0.0	0.0	0.08
Lumi (Correlated)				
σ	0.037	0.037	0.037	0.037
Shape (Correlated)				
μ	1.04	1.04	1.04	1.04
σ	0.032	0.032	0.032	0.032

Table 8.5: All considered systematic effects summarized.

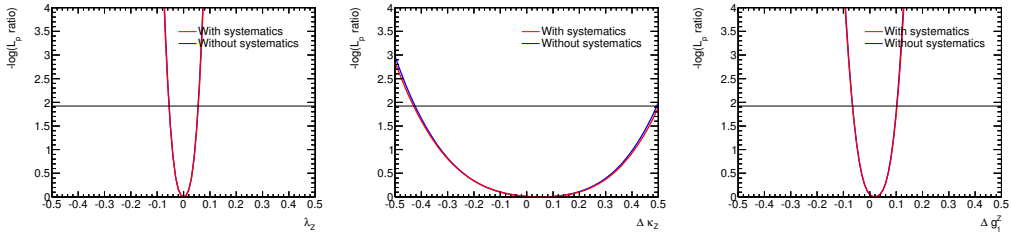


Figure 8.3: (Profile) Likelihood curves for Monte Carlo expectation including systematics, with background removal using BDT, for λ_Z (left), $\Delta\kappa_Z$ (middle) and Δg_1^Z (right). Corresponding limits (along with limits for fits to 827.87 pb^{-1} of data) are quoted for both likelihood fit and Neyman Construction in table 8.6. In the figures profile likelihood is red and likelihood is blue.

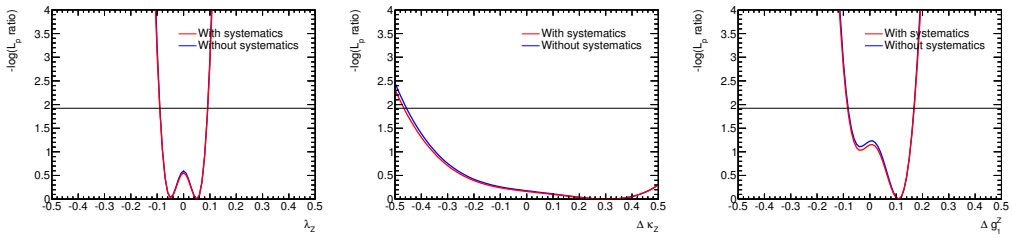


Figure 8.4: (Profile) Likelihood curves for fits of data corresponding to 827.87 pb^{-1} of integrated luminosity, including systematics, with background removal using BDT, for λ_Z (left), $\Delta\kappa_Z$ (middle) and Δg_1^Z (right). Corresponding limits (along with limits for fits to 827.87 pb^{-1} of data) are quoted for both likelihood fit and Neyman Construction in table 8.6. In the figures profile likelihood is red and likelihood is blue.

	λ_Z	$\Delta\kappa_Z$	Δg_1^Z
Likelihood prediction	[-0.046 , 0.047]	[-0.43 , 0.51]	[-0.040 , 0.100]
Neyman prediction	[-0.056 , 0.058]	[-0.46 , 0.53]	[-0.062 , 0.100]
Binned Likelihood	[-0.092 , 0.092]	[-0.44 , 0.71]	[-0.082 , 0.190]
Neyman Construction	[-0.098 , 0.097]	[-0.46 , 0.73]	[-0.11 , 0.20]

Table 8.6: Limits on charged triple gauge boson couplings, with removal of backgrounds using BDT, including systematic effects. The two upper rows gives the MC prediction (using Asimov data sample) and the two lower rows using a 827.87 pb^{-1} of data from the ATLAS detector. Corresponding likelihood curves are given in figures 8.4 and 8.3.

8.7 Prediction for 5 fb^{-1} of integrated luminosity

In order for the reader to compare the method developed in this thesis to the coming official ATLAS analysis (expected for the summer 2012), expected limits for 5 fb^{-1}, including systematics, given SM samples are provided in table 8.7 and figure 8.5. As seen, the systematic error becomes more important as the amount of data goes up, as expected. This is due to the fact that systematic effects does not vanish with large sample sizes.

It is important to note that the expectation given here is not a full 5 fb^{-1} prediction, but simply an extrapolation of the present study. As machine parameters have changed, the systematic effects have also changed, and an extension of the present study to 5 fb^{-1} should keep this in mind.

	λ_Z	$\Delta\kappa_Z$	Δg_1^Z
Likelihood prediction	[-0.032 , 0.035]	[-0.32 , 0.35]	[-0.055 , 0.057]
Neyman prediction	[-0.033 , 0.035]	[-0.35 , 0.37]	[-0.056 , 0.058]

Table 8.7: Predicted limits for MC corresponding to 5 fb $^{-1}$ of luminosity, including systematic effects, given a SM Asimov sample, using cut on BDT output.

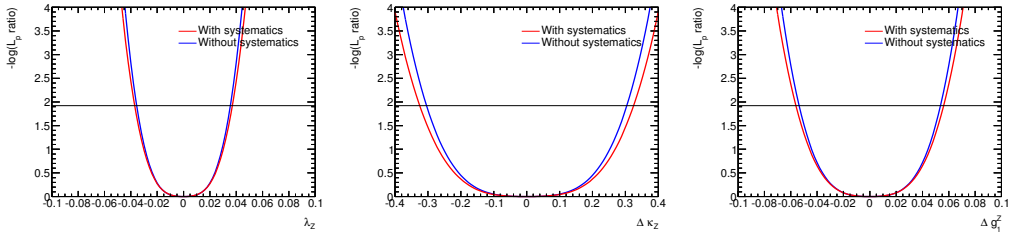


Figure 8.5: (Profile) Likelihood curves for Monte Carlo expectation for 5 fb^{-1} of integrated luminosity, including systematics, with background removal using BDT, for λ_Z (left), $\Delta\kappa_Z$ (middle) and Δg_1^Z (right). Corresponding limits are quoted for both likelihood fit and Neyman Construction in table 8.7. In the figures profile likelihood is red and likelihood is blue.

8.8 Discussion

As it was the case without systematics, the limits for actual data is wider than SM expectation, due to more observed events than expected. Limits are compatible with the SM. It is also seen that systematic error does not contribute much to the set limits with the amount of statistics available. This will change as more data is added. The limits are seen to improve the official ATLAS result, which uses slightly more data [75]. The reason is that the official ATLAS result relies on TGC extraction from total cross section, which is seen to be much less sensitive (chapter 4) than using the $p_T(Z)$ distribution. The official 5 fb^{-1} analysis is expected to use $p_T(Z)$, and in table 8.7 predictions for 5 fb^{-1} is calculated in order for the reader to make a comparison when this analysis is published.

CHAPTER 9

CONCLUSION AND PERSPECTIVE

With the work of this thesis, limits at a 95 % CL using a Neyman Construction with a Feldman-Cousins ordering, are put on the triple gauge boson couplings of the WWZ vertex.
From data corresponding to 827.87 pb^{-1} of integrated luminosity:

$$-0.098 < \lambda_Z < 0.097 \ , \ -0.46 < \Delta\kappa_Z < 0.73 \text{ and } -0.11 < \Delta g_1^Z < 0.20.$$

With Standard Model expectation:

$$-0.056 < \lambda_Z < 0.058 \ , \ -0.46 < \Delta\kappa_Z < 0.53 \text{ and } -0.062 < \Delta g_1^Z < 0.100.$$

The systematic effects included are: limited MC statistics, PDF uncertainty, uncertainty on calculated TGC weights, luminosity and shape reconstruction.

It is seen that the observed limits are less tight than predicted. This is due to the fact that more high $p_T(Z)$ events were observed than expected. Adding more data to the analysis will be essential in order to see if this effect stems from underlying physics, or is merely a statistical fluke. The key considerations leading to these limits, and worth to keep in mind for future analysis are:

- Even with slightly less integrated luminosity than used for the limits published in Ref. [37], tighter expectation and tighter limits were provided. This shows that using the $p_T(Z)$ distribution is a better choice than using total cross section. This point is further backed by an MC truth study which shows that the expected sensitivity increases with about a factor of 2, when using the $p_T(Z)$ distribution rather than total cross section.

- Not much improvement in signal purity is gained by using a BDT after a cut based selection. An obvious extension of this study would be to use a data driven background removal technique, compared to the BDT analysis, using a further relaxed cut based selection.

- The maximum likelihood fits tend to undercover given certain sample configurations, due to the TGC model's lack of ability to describe all types of sample configurations.

The last point is worth dwelling at, as this is something future studies relying on likelihood based estimation techniques needs to address. The solution used for this study – a fully frequentist limit setting technique – is recommendable. Another option would be to extend the TGC model such that it is able to cover the configurations made with problematic data samples in this thesis. This can for example be done by altering the choice of TGCs to include additional (complex valued) TGCs.

The use of the Neyman Construction as a limit setting technique for TGCs, is also a subject that is not exhausted. As described in this thesis, existing implementations makes use of importance sampling to make the procedure computationally less demanding. This is even more essential when fitting more than one TGC at a time, as the CPU time used to make a Neyman Construction grows with the number of parameters squared. As indicated by the likelihood ratios depicted in chapter 7, two disjunct regions can be obtained if the observed sample is far away from the expected one – i.e. a large TGC value is observed. The importance sampling technique cannot currently accommodate such disjunct regions, and it is necessary to supplement with drawing full »confidence belts« as the ones shown in figure 4.4.

A possibility for limit setting not mentioned in this thesis, is to make use of a Bayesian method. As such methods does not claim to be fully frequentist, under coverage with respect to the Neyman Construction would not conceptually be as severe a problem as the problem of under coverage by the maximum likelihood method described in this thesis.

As mentioned, the most efficient way to improve TGC limits at the moment, is to add more data to the analysis. This will not, however, remain to be the case. A long term LHC shutdown is scheduled for 2013, and to get the most out of the data available at that point, it is important that the analyzes striving to cover the largest possible amount of data is supplemented by attempts to improve the methods of fitting the TGCs. In this thesis a proof-of-principle for using an unbinned fit of the TGCs is given, and it is shown that the expected sensitivity improves compared to the binned fits of the $p_T(Z)$ distributions used so far. An obvious continuation of this study would be to:

- Test if the improvement in expected sensitivity remains after detector simulation.
- Use the unbinned technique described in this thesis directly in fits of data.
- Improve the suggested technique by adding a Poisson distribution to accommodate for the non-observation of data at $p_T(Z)$ above the data point of the highest $p_T(Z)$ (section 4.2.3).
- Improve the suggested technique by making it truly unbinned in Monte Carlo by applying a smoothing of the F_{ij} histograms.

To go even further beyond the scope of this thesis, future studies of limits on TGCs could improve the sensitivity by:

- Investigating the possibilities of combining information of several distributions in a single fit. The $p_T(Z)$ distribution was chosen before the $M_{inv}(WZ)$ distribution, but a combination of the two may provide improved sensitivity compared to using just $p_T(Z)$.
- Combining multiple channels. The three TGCs considered in this thesis influence (along with the three corresponding TGCs of the $WW\gamma$ vertex) the W^+W^- final state, allowing for combination studies.

From February 2012 the LHC collision energy went up to $\sqrt{s} = 8$ TeV. The strong $\hat{s}$ dependance on the effect of TGCs, means that larger collision energy will improve the ability to set limits. Moreover, raising the energy opens up a new, unexplored territory which could be home to New Physics, making the triple gauge boson coupling analysis as important as ever.

CHAPTER 10

ACKNOWLEDGEMENTS

It has been a pleasure for me spend a year in the Experimental High Energy Physics group, at the Discovery Center at NBI. The group have made me feel most welcome, and have provided facilities for writing this thesis beyond all expectations. The group have also made it possible for me to make two visits to CERN during this period, which have been both inspiring and fruitful to me.

First I would like to thank my supervisor Jørgen Beck Hansen as the one who introduced me to physics beyond the Standard Model and the world of couplings. The many discussions on particle physics, physics in general and many totally non-related issues, have been both inspiring and motivating.

This analysis would of course not have been possible without the hard work done by the ATLAS collaboration and CERN/LHC to provide data and simulation tools.

I would also like to thank the always friendly ph.d.-students of the group for their helpful attitude, and willingness to discuss a problem of physics or computing. Kristian, Morten, Ask and Peter, please continue that approach.

Finally the fellow students in the group have made this year much more enjoyable indeed. Jesper, Bjørn, Lars, Ingrid, Christine, Martin, Therkel, Gorm, Maria and Thomas – best of luck on your road ahead.

Copenhagen, May 2012.

CHAPTER 11

BIBLIOGRAPHY

[1] David Griffiths. *Introduction to Electrodynamics*. Pearson, 2nd edition, 1999. ISBN 1.

[2] ATLAS Collaboration. Combined search for the Standard Model Higgs boson using up to 4.9 fb^{-1} of pp collision data at $\sqrt{s}$ =7 TeV with the ATLAS detector at the LHC. *arXiv:hep-ex/12021408v3*, 2012.

[3] Particle Data Group, K. Nakamura, et al. Review of Particle Physics. *J. Phys. G*, 37(075021), 2010.

[4] O. W. Greenberg. CPT violation implies violation of Lorentz invariance. *arXiv:hep-ph/0201258v3*, 2002.

[5] W. N. Cottingham and D. A. Greenwood. *An Introduction to the Standard Model of Particle Physics*. Cambridge, 2007. ISBN 987-0-521-85249-4.

[6] E. N. Argyres et al. Trilinear gauge boson couplings in the MSSM. *Phys. Lett.*, B383 (63), 1996.

[7] J. R. Andersen et al. Discovering technicolor. *arXiv:hep-ph/11041255v2*, 2011.

[8] R. P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.*, 20:367–387, 1948.

[9] W. Buchemüller and C. Lüdeling. Field theory and Standard Model. *arXiv:hep-ph/0609174v1*, 2006.

[10] A. Zee. *Quantum Field Theory in a Nutshell*. Princeton Press, 2nd edition, 2003. ISBN 9-780-69101-0199.

[11] Jørgen Beck Hansen. *Triple Gauge-boson Couplings in W Pair Production via e^+e^- Annihilation*. PhD thesis, University of Copenhagen, NBI, 1996.

[12] LEP Collaboration and LEP TGC Working Group. A combination of results of charged triple gauge boson couplings measured by the LEP experiments, accessed 27.03.2012, 2005. URL `http://lepww.web.cern.ch/lepww/tgc/2005/gc_pdg05.ps`.

[13] D0 Collaboration. Measurement of the WZ→ eνee cross section and limits on anomalous triple gauge couplings in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *arXiv:hep-ex/1006.0761v1*, 2010.

[14] K. Hagiwara, R.D. Peccei, and D. Zeppenfeld. Probing the weak boson sector in $e^+e^- \to W^+W^-$. *Nucl. Phys. B*, 282:253–307, 1988.

[15] D. Zeppenfeld and Scott Willenbrock. Probing the three-vector-boson vertex at hadron colliders. *Phys. Rev. D*, 37(7):1775–1786, 1988.

[16] M. Bilenky et al. Trilinear couplings among the electroweak vector bosons and their determination at LEP2. *Nucl. Phys. B*, 408:22–68, 1993.

[17] H. Aihara et al. Anomalous gauge boson interactions. *arXiv:hep-ph/9503425v1*, 1995.

[18] Kristian Gregersen. Limits on anomalous trilinear gauge couplings in $Z\gamma$ production at $\sqrt{s} = 7$ TeV and $L = 300\text{pb}^1$ in the ATLAS experiment. Master's thesis, University of Copenhagen, NBI, 2010.

[19] U. Baur and D. Zeppenfeld. Unitarity constraints on the electroweak three vector boson vertices. *Phys. Lett. B*, 201:383–389, 1987.

[20] Jørgen Beck Hansen. TGC limit determition (ATLAS internal recommendation), accessed 27.03.2012. URL `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TGCLimitDetermination`.

[21] U. Baur, T. Hahn, and J. Ohnemus. QCD corrections and non-standard three vector boson couplings in W^+W^- production at hadron colliders. *arXiv:hep-ph/9507336v1*, 1995.

[22] S. Alioli, P. Nason, C. Oleari, and E. Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 1006(043), 2011.

[23] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi. W^+W, WZ and ZZ production in the POWHEG BOX. *JHEP*, 1111(078), 2011.

[24] T. Sjöstrand, S. Mrenna, and P. Skands. Pythia 6 (8) manual. *JHEP*, 05(026), 2006.

[25] Dan Green. *High p_T physics at hadron colliders*. Cambridge University Press, 1st edition, 2004. ISBN 1.

[26] Andy Buckley et al. General-purpose event generators for LHC physics. *arXiv:hep-ph/1101.2599v1*, 2011.

[27] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *arXiv:hep-ph/0201195v3*, 2002.

[28] Durham University HEPDATA. On-line unpolarized parton distribution calculator with graphical display, accessed 29.2.2012. URL `http://hepdata.cedar.ac.uk/pdfs/`.

[29] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *arXiv:hep-ph/0204.244*, 2002.

[30] ATLAS Collaboration. Measurements of underlying-event properties using neutral and charged particles in pp collisions at $\sqrt{s} = 900$GeV and $\sqrt{s} = 7$TeV with the ATLAS detector at the LHC. *arXiv:hep-ex/1103.1816v2*, 2011.

[31] Gideon Bella. Weighting di-boson Monte Carlo events in hadron colliders. *arXiv:hep-ph/08033307v1*, 2008.

[32] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *arXiv:data-an/1007.1727v2*, 2010.

[33] L. J. Dixon, Z. Kunszt, and A. Signer. Helicity Amplitudes for $\mathcal{O}(\alpha_s)$ Production of $W^+W^-, W^\pm Z, ZZ, W^\pm$, or Z Pairs at Hadron Colliders. *arXiv:hep-ph/9803250*, 1998.

[34] Morten Badensøe. Limits on anomalous triple gauge couplings in WW production with semileptonic final state. Master's thesis, University of Copenhagen, NBI, 2011.

[35] LEP Collaboration. The public WWW page of the LEP Electroweak Gauge-Couplings Group, accessed 30.4.2012, . URL `http://lepww.web.cern.ch/lepww/tgc/`.

[36] CMS Collaboration. Measurement of W^+W^- Production for the Higgs Boson in pp collisions at $\sqrt{s} = 7$ TeV. *arXiv:hep-ex/1002.5429v2*, 2011.

[37] ATLAS Collaboration. Measurement of the WZ production cross section and limits on anomalous triple gauge couplings in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *arXiv:hep-ex/1111.5570v2*, 2011.

[38] The CDF Collaboration. Measurement of WZ cross section and triple gauge couplings in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$TeV. *arXiv:hep-ex/12026629v1*, 2012.

[39] The CDF Collaboration. Measurement of triple gauge couplings in WZ in trilepton plus MET events using 7.1 fb^{-1}. *CDF note 10595*, 2011.

[40] Glen Cowan. *Statistical data analysis*. Oxford Science Publications, 1st edition, 1998. ISBN 0-19- 850156-0.

[41] F. James and M. Roos. Minuit: A system for function minimization and analysis of the parameter errors and correlations. *Comput. Phys. Commun.*, 10:343–367, 1975.

[42] R. Brun and F. Rademakers. ROOT - an object oriented data analysis framework v. 5.32. *Nucl. Instrum. Meth. (see also http://root.cern.ch/)*, A389:81–86, 1997.

[43] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling v. 3.14. *arXiv:physics-data-an/0306116v1*, 2003.

[44] W. Verkerke. Guide to parameterized likelihood analyses (draft). *Unpublished (Not public)*, (Version 1.01), 2012. URL `svn/atlasgrp/browser/Physics/StatForum/FrequentistRecommendation/doc/profiling_recommendations_draft.pdf`.

[45] Gary J. Feldman and Robert D. Cousins. A unified approach to the classical statistical analysis of small signals. *arXiv:physics-data-an/9711021v2*, 1999.

[46] J. Neyman. Outline of a theory of statistical estimation based on the classical theory of probability. *Phil. Trans. of the Royal Society of London. A*, 236:333–380, 1937.

[47] Louis Lyons and Luc Demortier. Importance of ordering rule for Neyman construction. *CDF/MEMO/STATISTICS*, 2003.

[48] Lawrence D. Brown, T. Tony Cai, and Anirban DasGupta. Interval estimation for a binomial proportion. *Statistical Science*, 16(2):101–133, 2001.

[49] Lyndon Evans and Philip Bryant (ed.). LHC machine. *JINST 3*, S08001, 2008.

[50] ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider. *JINST*, 3, 2008.

[51] CERN Collaboration. CERN Homepage, accessed 29.2.2012, . URL `http://www.cern.ch/`.

[52] ATLAS Collaboration. ATLAS Run Query, accessed 29.2.2012, . URL `http://atlas-runquery.cern.ch/`.

[53] ATLAS Collaboration. ATLAS Luminosity, public results, accessed 29.2.2012, . URL `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`.

[54] ATLAS Collaboration. ATLAS Public results, accessed 29.2.2012, . URL `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/`.

[55] Dan Green. *The physics of particle detectors*. Cambridge University Press, 1st edition, 2000. ISBN 1.

[56] Konrad Kleinknecht. *Detectors for Particle Radiation*. Cambridge University Press, 2nd edition, 1992. ISBN 0-521-64854-8.

[57] R.M. Brown and D.J.A. Cockerill. Review: Electromagnetic calorimetry. *Nucl. Inst. Phys. A*, 666:47–79, 2012.

[58] ATLAS Collaboration. Performance of the ATLAS trigger system in 2010. *arXiv:hep-ex/1110.1530v2*, 2011.

[59] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Inst. Phys. A*, 262:444–450, 1987.

[60] P. Billoir and S. Qian. Fast vertex fitting with a local parametrization of tracks. *Nucl. Inst. Meth. A*, 311(1-2):139–150, 1992.

[61] Bernado R V de Melo Xavier. Muon identification algorithms in ATLAS. *European Physical Society Europhysics Conference on High Energy Physics*, EPS-HEP 2009, 2009.

[62] ATLAS Collaboration. Expected electron performance in the ATLAS experiment. *Technical report ATL-PHYS-PUB-2011-006*, 2011.

[63] Matteo Cacciari, Gavin Salam, and Gregory Soyez. The anti-k_t clustering algorithm. *arXiv:hep-ph/0802.1189v2*, 2008.

[64] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth.*, A506 (250), 2003.

[65] ATLAS Collaboration. The ATLAS simulation infrastructure. *arXiv:physics-ins-det/1005.4568v1*, 2010.

[66] J. Allison et al. GEANT4 Developments and applications. *Trans. Nucl. Sci.*, 53 (270-278), 2006.

[67] Matthew Dobbs. *Probing the Three Gauge-boson Couplings in 14 TeV Proton-Proton Collisions*. PhD thesis, University of Victoria, 1997.

[68] Matthew Dobbs. Prospects for probing triple gauge-boson couplings at the LHC. *arXiv:hep-ph/0506174v1*, 2005.

[69] ATLAS Collaboration. Monte Carlo samples luminosity normalization info for p574 mc (ATLAS internal), accessed 01.03.2012, . URL `https://svnweb.cern.ch/trac/atlasgrp/browser/Physics/StandardModel/ElectroWeak/Analyses/Summer2011/Common/mc10b_p574_info.txt`.

[70] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. Polosa. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *arXiv:hep-ph/0206.239v2*, 2003.

[71] G. Corcella et al. HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *arXiv:hep-ph/0011.363*, 2001.

[72] ATLAS Collaboration. First tuning of HERWIG/JIMMY to ATLAS data. *ATLAS Note*, ATL-PHYS-PUB-2010-014, 2010.

[73] R. Decker, S. Jadach, J. Kuhn, and Z. Was. The tau decay library TAUOLA, version 2.4. *Comput. Phys. Commun*, 76(361), 1993.

[74] S. P. Baranov and M. Smizanska. Semihard b quark production at high-energies versus data and other approaches. *Phys. Rev. D*, 62(014012), 2000.

[75] J. Barreiro Guimaraes da Costa et al. A measurement of WZ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *(Unpublished) ATLAS-COM-PHYS-2011-1108*, 2011.

[76] Eve Le Menedeu. *Etude des performances du spectrometre a muons du detecteur Atlas au LHC, des cosmiques aux collisions. Mesure de la section efficace de production des dibosons WZ.* PhD thesis, Universite Pierre et Marie Curie (Paris 6), 2011.

[77] S. Allwood-Spires et al. Monte Carlo samples used for top physics. *ATLAS-PHYS-INST*, 132, 2010.

[78] ATLAS Collaboration. Pileup reweighting, accessed 12.2.2012, . URL `https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/PileupReweighting`.

[79] ATLAS Collaboration. Reconstruction and Calibration of Missing Transverse Energy and Performance in W and Z Events in ATLAS Proton-Proton Collisions at $\sqrt{s} = 7$ TeV. *(Unpublished) ATLAS-CONF-2011-080*, 2011.

[80] ATLAS Collaboration. Search for supersymmetry in events with three leptons and missing transverse momentum in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS detector. *arXiv:hep-ex/12045638*, 2012.

[81] ATLAS Collaboration. Diboson physics studies with the ATLAS detector. *ATLAS Note*, ATL-PHYS-PUB-2008-036, 2008.

[82] Andreas Hoecker et al. TMVA – Toolkit for multivariate data analysis. *arXiv:physics-data-an/0703039v5*, 2009.

[83] Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *J. of Computer and System Science*, 119 (55), 1997.

[84] Raúl Rojas. AdaBoost and the Super Bowl of classifiers, accessed 25.3.2012. *Unpublished, Freie Universität Berlin*, 2009. URL `www.inf.fu-berlin.de/inst/ag-ki/adaboost4.pdf`.

[85] D. Bourilkov et al. LHAPDF: PDF use from the Tevatron to the LHC. *arXiv:hep-ph/0605240v2*, 2006.

[86] M. R. Whalley et al. The Les Houches Accord PDFs (LHAPDF) and Lhaglue. *arXiv:hep-ph/0508110*, 2006.

[87] A. D. Martin et al. Parton distributions for the LHC. *arXiv:hep-ph/0901000 2v3*, 2009.

[88] R. D. Ball et al. A first unbiased global NLO determination of parton distributions and their uncertainties. *arXiv:hep-ph/10024407v2*, 2010.

[89] CERN Collaboration. Recommendation for LHC cross section calculations, accessed 30.4.2012, . URL `http://www.hep.ucl.ac.uk/pdf4lhc/PDF4LHCrecom.pdf`.

[90] Jochen Hartert. *Measurement of the $W \to e\nu$ and $Z/\gamma^* \to WW$ Production Cross-Sections in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Experiment.* CERN Thesis 2011-186, Universitat Freiburg, 2011.

APPENDIX A

HIGGS MECHANISM

The motivation for introducing the Higgs Mechanism is to allow gauge boson masses without loosing gauge invariance. The principle behind the Higgs Mechanism is an elaborate breaking of the local gauge symmetry imposed on the Lagrangian. We need to break only the $SU(2)$ symmetry, and leave the $U(1)$ symmetry intact, since we only want to add mass to the weak gauge bosons, and leave the photon massless, in accordance with the experimental results summarized in table 2.2. An appropriate Higgs field Lagrangian must be constructed and added to equation (2.23). The Higgs field is a complex scalar field doublet ($Y = 1$, $I = \frac{1}{2}$), which in all generality can be written as [5]:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{A.1}$$

The Lagrangian describing a free scalar field is the Klein-Gordon Lagrangian introduced in equation (2.3) making a preliminary Higgs Lagrangian:

$$\mathcal{L}_{Higgs} = \partial_\mu \phi^\dagger \partial^\mu \phi - \mu^2 \phi^\dagger \phi. \tag{A.2}$$

We now wish to alter the potential energy density given by $\mu^2 \phi^\dagger \phi$, such that its minimum is no longer situated at $\phi = 0$. If one simply changes the sign of the term, the potential energy density would be unbounded from below (as μ^2 is always positive). This is resolved by adding another positive valued parameter λ, making the full Higgs potential:

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{A.3}$$

This seems like a somewhat arbitrary choice, but it can be shown to be the most general form of the potential that does the job. The potential is sometimes dubbed »The Mexican Hat«-potential because of its characteristic shape, depicted in figure A.1.

The partial derivative in equation (A.2) is now replaced with the electroweak covariant derivative introduced in equation (2.21), in order to make the Higgs field couple to the

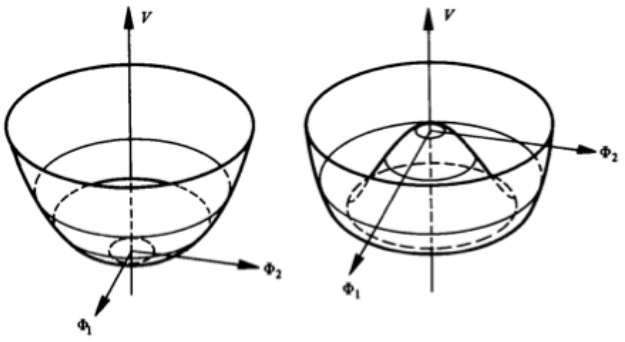


Figure A.1: The Higgs Potential for (left) $\mu^2 < 0$ and (right) $\mu^2 > 0$ (the Mexican Hat). Figure adapted from Ref. [11].

gauge bosons. The full Higgs Lagrangian is then:

$$\begin{aligned} \mathcal{L} &= D_\mu \phi^\dagger D^\mu \phi - V(\phi) \\ &= D_\mu \phi^\dagger D^\mu \phi + \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2. \end{aligned} \tag{A.4}$$

The Higgs Lagrangian is invariant under local $SU(2)_L \times U(1)_Y$ gauge transformations as required.

The behavior of the Higgs potential (equation (A.3)) can now be explored. Since $\mu^2 > 0$ and $\lambda > 0$, the minimum of the potential is fixed at:

$$\phi^\dagger \phi = |\phi|^2 = \frac{\mu^2}{2\lambda} \tag{A.5}$$

This does not fix the values of $\phi_1, ...\phi_4$ in equation (A.1), so we are free to choose the vacuum expectation value, as long as it fulfills equation (A.5):

$$\langle \phi \rangle_{vac} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \ , \ v = \sqrt{\frac{\mu^2}{\lambda}}. \tag{A.6}$$

The full expression for the Higgs field can now be written as some local $SU(2)$ gauge transformation of the vacuum expectation value plus a term dependent of the position [90]:

$$\begin{aligned} \phi(x) = \frac{\exp(i\theta^j(x)\sigma_j/2)}{\sqrt{2}} \left[\begin{pmatrix} 0 \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ H(x) \end{pmatrix} \right] \\ = \frac{\exp(i\theta^j(x)\sigma_j/2)}{\sqrt{2}} \begin{pmatrix} 0 \\ \rho(x) \end{pmatrix}, \end{aligned} \tag{A.7}$$

where $\rho(x)$ is introduced as the total density distribution of the Higgs field. The full Higgs field can then be substituted into the Lagrangian (equation (A.2)), where the electroweak

covariant derivative (equation (2.21)) acts on it. Using line (1) in equation (A.7), the mass term in can be written in terms of the vacuum expectation value of the Higgs Field and (by using 2.24) the physical fields. The masses are now generated, and related to each other and the Weinberg angle through:

$$\frac{M_W}{M_Z} = \cos\Theta_W. \tag{A.8}$$

The Higgs mechanism is a neat way of introducing masses to the gauge bosons, but it does not come without a cost. The quanta of the Higgs field is a charge neutral, massive scalar particle called the Higgs boson. Because of the arbitrary extra parameter (λ) introduced, the Higgs mass cannot be predicted, but is a so-called free parameter of the theory, determining λ:

$$\lambda = \frac{M_H}{2v^2}. \tag{A.9}$$

In february 2012, the LHC experiments ATLAS and CMS made a preliminary claim of seeing the Higgs Boson decaying through $H \to 2Z^0 \to 4l$ and $H \to 2\gamma$ at a mass of $M_H \approx 125GeV$, but conclusive evidence still remains, and it is in fact still uncertain whether a Higgs boson exists or not. Much points in the direction that the Higgs' hypothesis will be accepted or rejected at a 95% CL in the course of 2012 or 2013.

APPENDIX B
BACKGROUND SAMPLES

In this appendix information on the individual Monte Carlo samples used for background estimation is quoted. The information is all taken from Ref. [75] and Ref. [76].

MCID	Process	Generator	Events	k-factor	Filter efficiency	Cross section [pb]
105986	ZZ	Herwig	249906	1.41	0.21152	4.60
105985	WW	Herwig	249915	1.52	0.38863	29.59

Table B.1: MC samples/processes to model ZZ and WW background.

MCID	Process	Generator	Events	k-factor	Filter efficiency	Cross section [pb]
105200	$t\bar{t}$	MC@NLO	1496936	1.0	0.55551	164.57
105200	$t\bar{t}$	MC@NLO	3495579	1.0	0.55551	164.57
105200	$t\bar{t}$	MC@NLO	9853199	1.0	0.55551	164.57
105757	$bbcc\_\mu\mu$	PythiaB	297118	1.0	1.0	2830
105758	$bbcc\_\mu e$	PythiaB	796391	1.0	1.0	4017
105759	$bbcc\_ee$	PythiaB	290326	1.0	1.0	1693

Table B.2: MC samples/processes to model $t\bar{t}$ and dijet (multijet) background.

MCID	Process	Generator	Events	k-factor	Filter efficiency	Cross section [pb]
107650	ZeeNp0	Alpgen	6582268	1.26	1.0	664.100
107651	ZeeNp1	Alpgen	1323746	1.26	1.0	132.462
107652	ZeeNp2	Alpgen	404873	1.26	1.0	40.226
107653	ZeeNp3	Alpgen	109942	1.26	1.0	11.138
107654	ZeeNp4	Alpgen	29992	1.26	1.0	3.12818
107655	ZeeNp5	Alpgen	8992	1.26	1.0	0.753425
107660	Z$\mu\mu$Np0	Alpgen	6619010	1.26	1.0	663.79
107661	Z$\mu\mu$Np1	Alpgen	1334723	1.26	1.0	132.95
107662	Z$\mu\mu$Np2	Alpgen	403886	1.26	1.0	40.375
107663	Z$\mu\mu$Np3	Alpgen	109954	1.26	1.0	11.161
107664	Z$\mu\mu$Np4	Alpgen	29978	1.26	1.0	2.8987
107665	Z$\mu\mu$Np5	Alpgen	9993	1.26	1.0	0.75662

Table B.3: MC samples/processes to model Z+jets. NpX (X=0...5) refers to the number of additional partons in the final state.

MCID	Process	Generator	Events	k-factor	Filter efficiency	Cross section [pb]
107680	W$e\nu$Np0	Alpgen	2072731	1.21	1.0	6870.5
107681	W$e\nu$Np1	Alpgen	641361	1.21	1.0	1293.0
107682	W$e\nu$Np2	Alpgen	4521053	1.21	1.0	376.60
107683	W$e\nu$Np3	Alpgen	1211068	1.21	1.0	101.29
107684	W$e\nu$Np4	Alpgen	301832	1.21	1.0	25.246
107685	W$e\nu$Np5	Alpgen	83447	1.21	1.0	7.1239
107690	W$\mu\nu$Np0	Alpgen	3466523	1.21	1.0	6871.1
107691	W$\mu\nu$Np1	Alpgen	641867	1.21	1.0	1294.7
107692	W$\mu\nu$Np2	Alpgen	4519611	1.21	1.0	376.08
107693	W$\mu\nu$Np3	Alpgen	1213499	1.21	1.0	100.72
107694	W$\mu\nu$Np4	Alpgen	306848	1.21	1.0	25.993
107695	W$\mu\nu$Np5	Alpgen	83452	1.21	1.0	7.1300

Table B.4: MC samples/processes to model W+jets. NpX (X=0...5) refers to the number of additional partons in the final state.

LIST OF FIGURES

LIST OF TABLES