Limits on anomalous trilinear gauge couplings in $Z\gamma$ production at $\sqrt{s} = 7\mathrm{TeV}$ and $\mathcal{L} = 300\mathrm{pb}^{-1}$ in the ATLAS experiment

Thesis submitted for the degree of
Candidatus Scientiarum in Physics

by

Kristian A. Gregersen
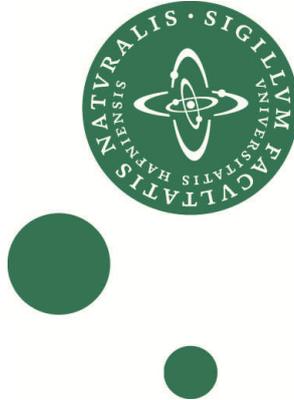
27.01.2010

Experimental High Energy Physics Group
Niels Bohr Institute
University of Copenhagen

Contents

Chapter 1

Introduction

High energy physics, or particle physics, is a branch of physics that studies the properties and interactions between elementary particles. A theory describing phenomena belonging to this area should include all fundamental forces of nature. Indeed, it has been the prime objective to find such a theory ever since the field emerged, however, it has not yet been accomplished. The current theoretical framework, known as the Standard Model (SM), is able to describe three out of the four known forces in nature, electromagnetism, the weak and the strong forces, leaving out only gravity. It has succeeded in describing nearly all high energy phenomena in great detail which has made it the only widely acknowledged model in particle physics.

The Standard Model is a quantum field theory based on symmetry principles. It describes the interactions between fermions (the matter particles) and bosons (the force carriers). A special feature which originates from the symmetries of the model is the unification of some of the forces. The unification of seemingly different forces indicates that at a fundamental level they are not different, but merely different manifestations of the same force. This idea is intriguing since it suggests that perhaps all forces in nature (including gravity) are unified at a fundamental level and can be described by a single unified theory.

The unification seen in the Standard Model started with the unification of the electric and magnetic forces in the theory called electromagnetism in the 19th century by Maxwell. This path was continued with the unification of the weak and electromagnetic forces in the 1970's by Glashow, Weinberg and Salam. This was done in a scheme that utilizes a special kind of symmetry known as gauge symmetry. At that time (the 1950-70's), several different ways of constructing theories were explored, but it became the gauge principle that survived and is the foundation of the Standard Model. As will be demonstrated later, the most noticeable consequence of imposing gauge symmetry on the Lagrangian is the natural introduction of gauge fields which can be interpreted as distinct particles—the gauge bosons. When promoting these from being mere auxiliary fields to having kinematic degrees of freedom, they are able to propagate and interact. These interactions are tightly constrained by the symmetry principles, so experimental deviations from the allowed interactions suggest that the symmetries contained in the theory are incomplete.

Though no such deviations have been discovered yet, it turns out that the Standard Model is unsuitable as a fundamental theory for other reasons. One obvious reason is its exclusion of gravity as mentioned above—but this is only one of many problems. Some of these will be addressed in subsequent chapters. This has lead to numerous speculations of more or less exotic theories which try to solve some of the problems by making extensions to the Standard

Model or by proposing entirely new ways of thinking of the fundamental constituents or even space and time.

When it comes to the experimental part of high energy particle physics it seems there are two main roads to take when searching for physics beyond the Standard Model.

One way is to search for particles or interactions predicted by a specific theory or model. This is a so called model-dependent search for new physics. The most prominent example of this kind is perhaps a search for supersymmetric particles. In short, supersymmetry is an extension to the Standard Model which proposes a symmetry between SM fermions and bosons which manifests itself in a range of new particles. A positive outcome of a search for these particles will make supersymmetry more attractable as an extension to the Standard Model. It will also make other theories, such as string theory, look more promising.

Another way to search for new physics is to probe different sectors of the SM, e.g. the bosonic sector for self interactions. If any non-SM interactions are discovered it means that the symmetries in the Standard Model are incomplete. It can be seen as a subtle hint that something needs to be changed in the very foundation of the Standard Model. It can also mean that an entirely new way of constructing the model will have to be adapted. In either case, a positive outcome of this method will not point towards a specific theory or model. Thus, a search of this kind is called a model-independent search for new physics.

This thesis takes the second approach. The work presented here is dedicated to the study of the self interactions in the bosonic sector—more precisely the anomalous neutral trilinear gauge vertex $ZZ\gamma$. The term "anomalous" means that this vertex is absent in the Standard Model. The process considered is $q\bar{q} \to Z\gamma$, also known as a di-boson $Z\gamma$ production. If the $ZZ\gamma$ couplings are set to non-zero values (due to effects originating from physics beyond the Standard Model) the cross section of $Z\gamma$ production will be altered. The effects of this will show up in the kinematics of the final states. In this thesis, the only final state considered is $Z\gamma \to e^+e^-\gamma$. There is no specific reason for choosing this particular final state, $Z\gamma \to \mu^+\mu^-\gamma$ for instance, could have been chosen equally well. The different decay channels of the Z poses different challenges and advantages. In any case, a full fledged study should include all channels, but this is not the aim of the work presented here.

The Monte Carlo dataset used in this study corresponds to the data taken during the first year of running at the LHC (i.e. 2010). It is expected that the center of mass energy in the collisions will be 7 TeV and that the integrated luminosity will be $\sim 300\text{pb}^{-1}$. A collision energy upgrade to 10 TeV after approximately 6 months may occur, but this has not been taken into account here. The numbers are taken from internal CERN talks at the 27th LHC Machine Committee meeting in September 2009. Considering only this low amount of data, the present study may not be able to put tighter constraints on the triple gauge vertex $ZZ\gamma$ than what has already been achieved at LEP and the Tevatron. However, it still is exciting and interesting to see what the first year of running can bring.

As will become clear during the first part of this thesis, a study of triple gauge couplings is a study of the underlying gauge group $SU_L(2) \times U_Y(1)$ describing the unification of electromagnetic and weak forces in the Standard Model. Before one really can appreciate this motivation, one has to understand the basic ideas behind the Standard Model. Thus, the first part of this thesis is dedicated to the Standard Model and how the anomalous neutral triple gauge vertex $ZZ\gamma$ can be introduced. The next part focuses on the experimental setup. The last part describes the analysis.

Theory

Chapter 2

The Standard Model of Particle Physics

All of our current understanding of elementary particle physics is represented by the Standard Model of Particle Physics. It contains a description of the elementary particles and forces through a mathematical framework called *quantum field theory* in which the particles exist as excitations of fundamental fields.

The particles are divided into two classes, "matter particles" and "force carriers". The matter particles consist of three generations of leptons and three generations of quarks which are all spin-1/2 particles, i.e. fermions. To each of these matter particles there is a corresponding anti-particle which has the same mass, but the opposite electromagnetic charge. The force carriers are the particles that mediates the forces included in SM, i.e. the electromagnetic force (the photon), the weak force (the Z and the W's) and the strong force (the gluons). These are all spin-1 particles, i.e. bosons. Last, but not least, SM also contains a spin-0 boson called the Higgs which is introduced to provide masses for all the other particles. It is the only particle predicted by SM which has not yet been verified experimentally. A summary of the particles and some of their properties is given in table 2.1 [1].

	Particles			mass/MeV			EM charge	weak charge	color charge
fermions	e	μ	τ	0.511	106	1777	-1	yes	no
	ν_e	ν_μ	ν_τ	$< 2 \cdot 10^{-6}$	<0.19	<18.2	0	yes	no
	u	c	t	2	1.27	$171 \cdot 10^3$	2/3	yes	yes
	d	s	b	5	104	$4.2 \cdot 10^3$	-1/3	yes	yes
bosons	photon			0			0	no	no
	Z			$91.187 \cdot 10^3$			0	no	no
	W's			$80.39 \cdot 10^3$			±1	yes	no
	gluons			0			0	no	yes
	Higgs			$> 114 \cdot 10^3$			0	yes	no

Table 2.1: This table shows the matter particles and the force carriers along with some of their properties as described in the text. The fermions are organized vertically in generations. In SM the neutrinos are considered massless, the neutrino masses listed here represent the experimental bounds.

---

[1]Throughout this thesis the units used are the so called *natural units* in which $\hbar = c = 1$.
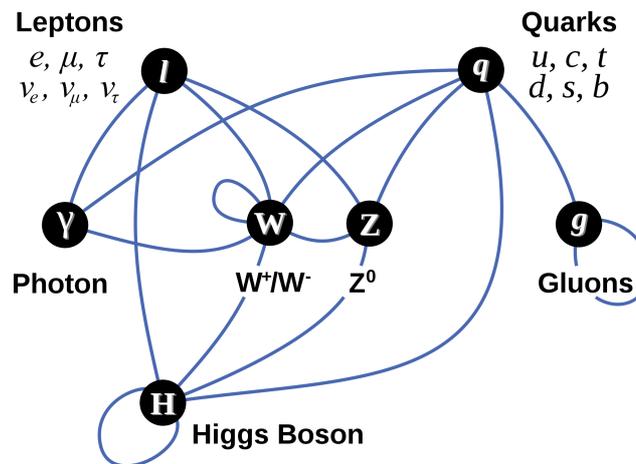
Figure 2.1: This picture shows how the SM particles interact which is a consequence of the quantum numbers displayed in table 2.1. For instance, the Z boson couples to particles with weak and/or electromagnetic charges and hence to leptons, quarks, W's and the Higgs. Likewise, the photon only couple to charged leptons, quarks and the W's since these particles carry electric charge. Note that the line connecting the photon with the leptons only is valid for the charged leptons. The figure is from www.wikipedia.org/wiki/Standard_Model.

The fermions of the Standard Model are classified according to how they interact (or equivalently, by what charges they carry). There are six quarks (up, down, charm, strange, top, bottom), and six leptons (electron, electron neutrino, muon, muon neutrino, tauon, tauon neutrino). Pairs from each classification are grouped together to form a generation, with corresponding particles exhibiting similar physical behavior.

The defining property of the quarks is that they carry color charge, and hence, interact via the strong interaction. The low energy confining behavior of the strong force results in quarks being bound to one another, forming color-neutral composite particles (hadrons) containing either a quark and an antiquark (mesons) or three quarks (baryons). The proton and the neutron are the two baryons having the smallest mass. Quarks also carry electric charge and weak isospin. Hence they interact with other fermions both electromagnetically and via the weak force.

The remaining six fermions do not carry color charge and are called leptons. The three neutrinos do not carry electric charge either, so they only participate in reactions involving the weak force. The remaining leptons (e, μ and τ) carry both weak and electric charges and thus interact electromagnetically and via the weak force.

Each member of a generation has greater mass than the corresponding particles of lower generations. The first generation charged particles do not decay; hence all ordinary (baryonic) matter is made of such particles. Second and third generations charged particles, on the other hand, decay immediately after being formed and are observed only in very high-energy environments. Neutrinos of all generations also do not decay and pervade the universe, only rarely interacting with baryonic matter. A summary of the interactions among SM particles is given in figure 2.1.

The SM particles are considered to be point-like, but, as already revealed, they contain an internal 'spin' (angular momentum) degree of freedom which is quantized and can have values

of 0, 1/2 or 1. Spin-1/2 particles obey Fermi statistics (Pauli's exclusion principle), which have as a consequence that 2 electrons cannot be in the same quantum state. This feature is necessary for forming atoms more complex than hydrogen. Spin-1 and spin-0 particles obey Bose-Einstein statistics, which prefer to have many particles in the lowest energy or ground state. As we shall see in a later section, the existence of the bosons in SM is a consequence of imposing a local (or gauge) symmetry on the fermions.

2.1 Short review of quantum field theories

This section will give a short introduction to quantum field theories which supply the basic framework for the Standard Model. It is only meant as a cursory review and hence will not provide detailed mathematical formulae. A more thorough introduction may be found in e.g. [1], [2] or [3].

2.1.1 Lagrangian formalism

The fundamental quantity in classical mechanics is the action, S. The action is the time integral of the Lagrangian which contains all one needs to know in order to determine the dynamics and kinematics of the particles in a physical system. In a local field theory particles only exist as excitations of the fields. In this case the Lagrangian can be written as the spatial integral of a Lagrangian density, denoted by $\mathcal{L}$, which is defined locally for each point as a function of one or more fields entering the theory and their derivatives[2].

Classically, when a physical system evolves from one configuration to another between two points in time, it is done according to the principle of least action. This principle states that the system is evolved along the path in configuration space for which the action is an extremum (usually a minimum). This path is called the classical path. The extremum is found by *varying* the action with respect to one of the fields, say $\phi(x)$, and demand that this variation vanishes for a given set of boundary conditions which reflect the physics of the situation. The equations which follow from this procedure are called the Euler-Lagrange equations of motion for the field $\phi(x)$. By varying the action with respect to the other fields entering the theory, their equations of motion can be determined as well.

In quantum physics the situation is slightly changed. It is still possible to employ the above machinery and find the Euler Lagrange equations, but contrary to a classical system the quantum system is not constrained to follow the classical path—in principle, all paths are allowed. In the vast majority of cases, however, the classical path is the dominant one along with small *quantum fluctuations* around it. This way of thinking of quantum physics is beautifully captured in the Feynman path integral formalism. The path integral formalism generalizes the action principle of classical mechanics by replacing the classical notion of a single, unique trajectory for a system with a sum, or functional integral, over an infinity of possible trajectories to compute a quantum amplitude.

2.1.2 Feynman diagrams

The actual calculation of transition amplitudes from the Lagrangians which seem to describe nature (e.g. Quantum Electro Dynamics, QED) is more complicated than what might have

---

[2]It is custom in field theory to simply refer to the Lagrangian density as the Lagrangian. This approach will be adopted here.
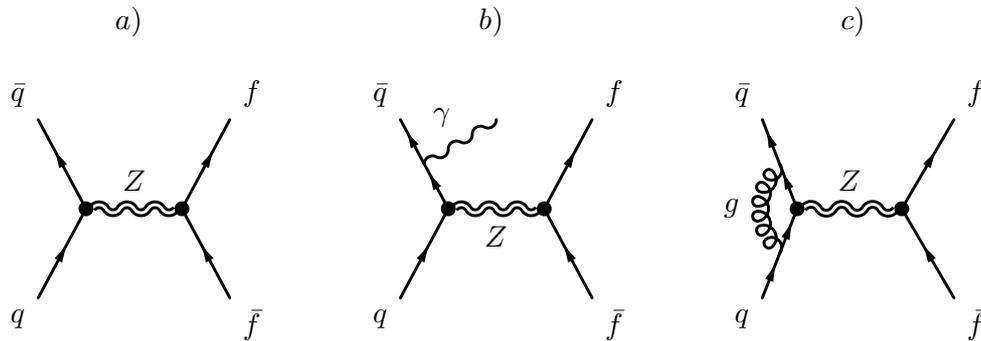
Figure 2.2: Examples of three different feynman diagrams. a) Drell Yan process with subsequent decay of the Z boson to a fermion-antifermion pair. b) Same as the first diagram, but with initial state radiation (ISR) of a photon. c) Same as the first diagram, but with loop including a gluon.

been insinuated above. It turns out that it is efficient to split up the action into a sum of two separate parts, a *free* part and an *interaction* part. The interaction part describes the non-trivial transition from one state to another, i.e. a transition which involves some kind of interaction between the fields. This part of the amplitude is often denoted by a matrix element from the scattering matrix, $\mathcal{M}$. This matrix appears in expressions such as cross sections and decay probabilities as $|\mathcal{M}|^2$, and thus it is the main objective to determine it.

In order to determine the interaction part, one often relies on a basic assumption in quantum field theory: interactions are weak. This means that the interaction part can be perturbed, i.e. we can make a systematic expansion which (hopefully) converges. This is perturbation theory. Perturbation theory may be seen as a way of allowing, in a systematic way, for field configurations in the neighborhood of the classical solution. Hence, even though we are concerned with *quantum* field theory, the classical solutions will play a crucial role.

The actual expansion is carried out through the use of *Wick's theorem* and the parameter in which the perturbation is done is the coupling constant of the force under consideration. This approach involves rather complicated calculations. In order to ease these calculations, a set of rules, called the Feynman rules, have been developed. By using these rules it is possible to describe each term in the expansion by a Feynman diagram from which the contents of the term can easily be written down. Thus, to write down the scattering amplitude for a given process up to some order in the expansion, one only has to draw the corresponding diagrams. Figure (2.2) shows different types of feynman diagrams. In a Feynman diagram particles are represented by lines. Initial state particles are drawn as lines in the left side of the diagram while final state particles are represented by lines furthest to the right of the diagram. In the middle the lines may interact with each other and create new intermediate lines (corresponding to new particles) through different kinds of vertices. As indicated, time progresses from left to right in the diagram. The lines can be squiggly or straight, with an arrow or without, depending on the type of particle. A point where lines connect to other lines is an interaction vertex, and this is where the particles meet and interact by emitting or absorbing new particles, deflecting one another, or changing type.

Lines entering or leaving the diagram represent real particles, i.e. they satisfy the equations

of motion ($E^2 = p^2 + m^2$). Such particles are called "on mass-shell" or simply "on shell" particles. Lines in intermediate stages of the diagram are called virtual particles because they cannot be observed in an experiment. It is not even necessary for them to have the right relationship between E, p, and m. Particles which do not satisfy the equations of motion are said to be "off-shell". The types of interaction vertices that can be drawn are dictated by the Lagrangian. For instance, in QED a vertex where two electrons and a photon meet is possible, whereas a vertex where three photons meet is not, since the Lagrangian does not contain a term for this.

Diagrams in which the given initial and final states are achieved in the very simplest manner, meaning as few vertices as possible, are called leading order diagrams. Diagrams containing more vertices than the leading order diagrams are called higher order diagrams, since they correspond to higher order terms in the perturbative expansion. The more complicated a diagram is, the more suppressed it is, since each vertex in a given diagram gives a factor proportional to the coupling constant of the force describing the interaction. For instance, the second diagram in figure (2.2) is next-to-leading order (NLO) in QED with respect to the first diagram, since there is one extra vertex where a photon is emitted. The third diagram in figure (2.2) is next-to-next-to-leading order in QCD[3], since it contains a loop with a gluon and quarks and thus has two additional vertices with respect to the first diagram. For practical uses it is seldom necessary to go beyond the first few orders to get a reasonable approximation to $\mathcal{M}$.

Naturally, there are important cases where perturbation theory is believed to be insufficient, the most prominent example being the phenomenon of quark confinement. This non-perturbative behavior has to do with the running of coupling constants as a function of energy scale, which is a phenomenon connected to renormalization. In short, it turns out that the forces vary in strength depending on the specific energy scale on which they are considered. The strong coupling constant increases as the energy decreases. This makes it unsuitable in a perturbative expansion. However, at high energies the strong coupling becomes small, a phenomenon also known as asymptotic freedom, which allows perturbation theory to be used accurately in experiments performed at energies much larger than $\Lambda_{QCD} \approx 200$ MeV. Although limited in scope, this approach has resulted in the most precise tests of QCD to date.

In the case of electroweak interactions, perturbation theory works fine since the coupling constants increase with energy and are small for all experimental purposes.

2.1.3 Renormalization

The feynman approach mentioned above works fine at tree-level, i.e. for diagrams without loops. However, when including loop diagrams there is a chance that the scattering amplitude goes to $\pm$ infinity. The difficulties come from the fact that the internal momentum flow in loops is not constrained by the external momentum conservation. Since the momentum around a loop can take on any value and all values contribute to the diagram we have to perform an integration over these momenta—this usually leads to infinities as the momentum goes to zero (infrared divergences) or infinity (ultraviolate divergences). Since all observable quantities

---

[3]QCD, or Quantum Chromo Dynamics, is the gauge theory describing the strong interactions. The principle of gauge invariance will be explained in section 2.2.2.

must be finite the unphysical divergences have to be cancelled somehow[4].

One way to solve the problem is to impose a momentum cutoff on all momenta, i.e. simply ignoring all momenta higher than some fixed value[5]. The momentum cutoff can also be thought of as a distance cutoff. In other words, it amounts to ignoring effects that involve fields varying on distance scales shorter than some distance Λ_D. By doing this, our theory acquires one more parameter, namely the parameter Λ_D. However, since space-time is considered a continuum in quantum field theories, a shortest distance is not desirable. Thus, we wish to let Λ_D go to zero. This continuum limit can be achieved through a procedure known as "renomalization".

Renormalization is deeply connected to the coupling constants in a theory. It turns out that even though the Lagrangian contains coupling constants, it is hard to measure them directly. The parameters named "mass" and "charge" do not correspond to the parameters that are directly measured by colliding particles in an accelerator. In fact, if one tries to measure the charge of the electron (say) by smashing two electrons into each other in an accelerator, seeing how much they repel each other, the answer will depend on their momenta in the center-of-mass frame. The same is true for the electron mass and any other coupling constants there are in the Lagrangian. They have a "bare" value, the value that appears in the Lagrangian, and a "physical" value, the value that is measured by doing an experiment. The physical values depend on the bare values, the cutoff Λ_D, and a momentum scale p.

The idea is that we can measure the physical coupling constants at some momentum scale p, and then figure out which bare coupling constants would give these physical values, assuming some cutoff Λ_D. Then we can try to take a limit as $\Lambda_D \to 0$, adjusting the bare coupling constants as we take the limit, in order to keep the predicted physical coupling constants at their experimentally determined values. We redefine the bare coupling constants with expressions that themselves depend on momenta and have the exactly opposite diverging behavior as the one found in the loop integrals. This "continuum limit", if it exists, will be a theory without any shortest distance scale in it. We also notice that this procedure will make the bare coupling constants diverge. However, this does not matter since they will never be observed in an experiment. As long as the *physical* coupling constants are finite, there is no problem.

Sometimes it is not possible to renormalize a Lagrangian. The main thing to worry about is this: Even if certain bare coupling constants are zero, the corresponding physical coupling constants may be nonzero. For example, if the Lagrangian does not contain a mass term for some particle, the particle may still have a physical mass (measured at some momentum scale)—it acquires a mass through its interactions with other particles. This means that in order to succeed in adjusting the bare coupling constants to fit the experimentally observed physical coupling constants, we need to start with a Lagrangian that has enough bare coupling constants to begin with (since we cannot expect to fit N numbers with fewer than N numbers). So, for a given Lagrangian, we may have to add some extra terms with some extra bare coupling constants before doing the renormalization. If this is possible with only finitely many extra terms, the theory is said to be renormalizable. If we need infinitely many terms, we say the theory is nonrenormalizable.

There is an easy way to check whether a Lagrangian is renormalizable or not. Or, to

---

[4]It should be mentioned that these problems are directly related to the inclusion of quantum physics. By only including tree-level diagrams to describe a process, quantum physics is neglected and the situation is entirely classical.

[5]This procedure is known as "regularization".

be more precise, whether it is superficially renormalizable[6]. A term in the Lagrangian is superficially renormalizable if its coupling constant has dimension of length^d with d being less or equal to zero, but not if d is greater than zero [3]. This leads to the fact that if space-time is four dimensional, only operators of dimensions less than or equal to four can appear in the Lagrangian, if this is to be renormalizable. As we shall see in chapter 3, the $ZZ\gamma$ interaction term can only be introduced by operators of dimension 6 or higher, which means that they will not be renormalizable.

2.1.4   Running coupling constants

As already mentioned above, the physical coupling constants depend on the momentum scale at which they are measured. This momentum scale can also be thought of as a distance which we denote Λ'_D. If we denote the bare couplings as g, we can write the physical coupling constants, g', as some function f: $g' = f(g, \Lambda_D, \Lambda'_D)$. This is also called the renormalization group equation.

We now fix the cuttoff scale Λ_D and the bare couplings g and see how the physical couplings change, or "run", as we vary the distance scale Λ'_D (i.e. the momentum scale at which the physical couplings are measured). Seeing how the physical constants change is called the renormalization group flow or simply the running of the coupling constants. Once a physical coupling constant is known at a given Λ'_D the renomalization group equation can be used to extrapolate to other Λ'_D. In theories like QED and the electroweak theory, the physical coupling constants start out fairly low and increase slowly as Λ'_D decreases. This means that for all experimentally relevant energies the perturbative expansion previously discussed is valid, i.e. the feynman approach will give meaningful results.

To give a more intuitive explanation of running coupling constants let us consider QED. As mentioned, the electron charge increases as we look at smaller distances. This can be understood in physical terms as due to the existence of virtual electron-positron pairs in the vacuum around the real electron. These virtual pairs come from the decays of virtual photons which are constantly emitted and reabsorbed by the real electron. The virtual charged pairs will organize themselves with the virtual positron closest to the real electron. The total effect is that this cloud of virtual electron-positron pairs effectively acts like a polarized vacuum which will lower the value of the electron charge measured in an experiment just like a charge in a dielectric medium is partially cancelled out by the polarization it causes in the medium. This is also called charge screening. As we go to lower Λ'_D we are able to probe inside the cloud and the charge of the real electron will appear larger than at the higher distance scale—just like the renormalization group equations predict.

QCD is quite different. It turns out that the couplings constants decrease with momentum scale. Measurements well above ~ 200 MeV show that the strong coupling constant is significantly below 1. Thus, the perturbative expansion is valid and will even converge faster as the momentum scale is increased. This effect is known as asymptotic freedom since the quarks will behave more and more as free particles. However, when moving to lower momentum scales the strong coupling increases and becomes larger than 1. At this point the perturbative expansion breaks down and the feynman approach will give meaningless results since higher order terms in the expansion will become more significant than lower orders. Physically, this

---

[6]Superficial renormalizability is a weaker statement than renormalizability, since it can be spoiled by other deeper problems such as "anomalies" (which briefly will be addressed in a later section), but it turns out that it gives a very good hint as to whether a given term in a Lagrangian is renormalizable or not.

effect is interpreted as the point where quarks no longer behave as free particles and the effect of color confinement becomes important. Color confinement is the phenomenon that color charged particles (such as quarks) cannot be isolated, and therefore cannot be directly observed. For instance, when two quarks become separated, as it happens in particle accelerator collisions, at some point it is more energetically favorable for a new quark-antiquark pair to spontaneously appear out of the vacuum, than to allow the quarks to separate any further.

As a final point, it is worth noticing that the renormalization group flow can be considered regardless of the Lagrangian being renormalizable or not. This, of course, comes from the fact that we do not take the continuum limit $\Lambda_D \to 0$ when considering the running of the coupling constants. If the Lagrangian contains non-renormalizable terms, it turns out that these will actually get smaller as the distance gets larger. At large distances non-renormalizable interactions become irrelevant. As revealed in the last section, the terms for the $ZZ\gamma$ interaction is non-renormalizable, and thus we will expect the coupling constants to increase with energy. We will see that the running behavior actually has to be controlled somehow, since it breaks with unitarity—but, more on this later.

2.2 Symmetries in the Standard Model

Symmetries play a rather special role in physics. First of all, at a philosophical level, symmetry principles are believed to be an expression of the simplicity of nature. Secondly, and at a more technical level, continuous symmetries of a Lagrangian are linked to conserved currents (i.e. conservation laws) in a one-to-one correspondence through *Noether's theorem*[7] [1].

In fact, symmetries can be seen as the focal point when constructing models such as the Standard Model. As explained in the previous section, quantum field theory can provide the mathematical framework in which a Lagrangian controls the dynamics and kinematics of a physical system. The modern approach of constructing most field theories starts by first describing each kind of particle in terms of a dynamical field that pervades space-time. Then a set of symmetries of the system is postulated, and finally the most general renormalizable Lagrangian from its particle (field) content that obeys these symmetries is written down.

In the case of the Standard Model, global Poincare symmetry is postulated[8]. It consists of the familiar translational symmetry and rotational symmetry as well as the inertial reference frame invariance central to the theory of special relativity. Additionally, local $SU(3) \times SU(2) \times U(1)$ *gauge* symmetry is also postulated. This symmetry is an internal symmetry that in some sense defines the Standard Model. Roughly, the three factors of the gauge symmetry give rise to the three fundamental interactions contained in SM; $SU(3)$ is accountable for the strong interaction, while combinations of $SU(2) \times U(1)$ take care of the electromagnetic and the weak forces. The electroweak theory which contains the electromagnetic and the weak forces will be described in detail in section 2.2.2. For a description of the strong force, I refer to [1].

---

[7]The typical examples are translational symmetry leading to four-momentum conservation and rotational symmetry leading to angular momentum conservation. One more thing is important, however. Strictly speaking, Noether's theorem only applies to classical symmetries, i.e. for fields obeying the equations of motion. In quantum field theory, where we integrate over all fields most of which do not satisfy the equations of motion, there is a similar result known as the Ward Identities. These say that under certain circumstances the *expectation value* of the Noether current is conserved.

[8]A *global* symmetry is a symmetry that does not depend on space-time coordinates. *Local*, on the other hand, means that the symmetry transformation depends on the space-time coordinates.

2.2.1 Parity and Charge conjugation

The Standard Model also contains a set of discrete symmetries. Among these we find the *parity* and *charge conjugation* operations.

The parity operator performs a spatial inversion or reflection through the origin, i.e. it changes the sign of the spatial coordinates. While, e.g. translation is a continuous operation, reflections are discrete—it is impossible to reflect by an infinitesimal amount. The effect is different when operating on different quantities. If the quantity is an eigenfunction of the parity operator, $\mathbf{P}\psi(x) = P\psi(x)$, we can assign an parity-eigenvalue (intrinsic parity) to the quantity. Since two successive reflections brings the quantity back to what it originally was, $\mathbf{P}^2 = \hat{\mathbf{1}}$ and thus $P = \pm 1$. Scalars (e.g. time and energy) and pseudovectors (e.g. angular momentum and the magnetic field) have $P = 1$ and are referred to as having even parity. Vectors (e.g. position, momentum and the electric field) and pseudoscalars (e.g. helicity, i.e. the projection of a spin pseudovector onto the momentum vector, and the magnetic flux, i.e. the dot product between the surface normal vector and the magnetic field pseudovector) have $P = -1$ and are referred to as having odd parity. In quantum field theory, particles and anti-particles have the same parity if they are bosons, and opposite parity if they are fermions. All gauge bosons have odd intrinsic parity, $P_\gamma = P_Z = P_{W^+} = P_{W^-} = P_{gluon} = -1$.

Charge conjugation, sometimes also called C parity, turns a particle into its antiparticle. That is, it reverses all quantic charges, e.g. electric charge, baryon number, lepton number and flavor charges (e.g. strangeness). Thus, we have $\mathbf{C}\psi = \bar{\psi}$. Mass, position, energy, momentum, angular momentum and spin are not affected by the C operator. Particles that are their own anti-particles, such as the photon, are eigenstates of the C operator, i.e. $\mathbf{C}\psi = C\psi$. As with the parity operator, $\mathbf{C}^2$ equals the identity operator and hence the eigenvalues can take on the values $C = \pm 1$. The vector bosons have odd C parity, $C = -1$.

Parity and charge conjugation are said to be preserved if the Lagrangian remains the same after having applied these two operations. Some parts of the SM Lagrangian are unchanged while others are affected. The strong and the electromagnetic interactions preserve parity and charge conjugation, while the weak force violates both. During the next sections we will see how this is realized using gauge groups to represent the forces. We will also see that the weaker combination CP is preserved in the SM.

2.2.2 Gauge symmetries

In physics, the mathematical description of any physical situation usually contains excess degrees of freedom; the same physical situation is equally well described by many equivalent mathematical configurations. Sometimes it is possible to relate these configurations by some kind of transformation, and if this is the case, these transformations form a group of "symmetries" of the theory. This means that a physical situation does not correspond to an individual mathematical configuration, but to a class of configurations related to one another by this symmetry group.

A gauge theory is a mathematical model that has symmetries of this kind. More specifically, a gauge transformation is a "phase" transformation, which means that the Lagrangian describing the theory is invariant under the transformation

$$\psi(x) \to e^{i\alpha}\psi(x) \tag{2.1}$$

where $\psi(x)$ describes the particle fields. This implies that the phase α is immeasurable and can be chosen arbitrarily. The phase in (2.1) is a constant, thus, once it is fixed, the value

is specified for all points in space-time. In this case the transformation is called a *global* gauge transformation. The more general case where the parameter α depends on space-time, i.e. $\alpha(x)$, is called a *local* gauge transformation. Clearly, invariance under local gauge transformations is a stronger demand than invariance under global gauge transformations.

At present, it is believed that local gauge theories may describe all fundamental interactions in nature. The reason why gauge theories seem to be so important is not clear, however. The only thing that supports this point of view is the fact that particle physics experiments through the last 50 years or so indicate that nature prefers gauge theories. This can most easily be seen by the great success of the Standard Model in describing the wide variety of phenomena occurring in particle physics. In fact, the most successful theory in physics till this day is a gauge theory, namely Quantum Electro Dynamics.

Quantum Electro Dynamics and $U(1)$ gauge invariance

Quantum Electro Dynamics (QED) is a prime example of a successful gauge quantum field theory and has served as a general template for constructing many other gauge field theories. The starting point is the non-interacting electrons and positrons described by a common spinor field $\psi(x)$. The dynamics and kinematics of this field is captured by the Dirac Lagrangian

$$\mathcal{L}_{Dirac} = \bar{\psi}(x)(i\gamma^{\mu}\partial_{\mu} - m)\psi(x) \tag{2.2}$$

We now consider how this Lagrangian behaves under gauge transformations. If the gauge transformation is global, as in (2.1), is is clear that the Lagrangian is invariant. However, if the transformation is local, something interesting happens. The mass term is still invariant, but the term containing the derivative is not. Actually, we should make clear what we mean by ∂_{μ} since this is not obvious. The derivative is supposed to be defined as the difference between the field evaluated at two nearby space-time points in the limit where the distance between these points are taken to zero. However, since the field in general transforms differently at these two points, the derivative has no simple transformation law, and no useful geometrical interpretation. In order to subtract the values of the field at neighboring points in a meaningful way, we must introduce a factor that compensates for the difference in phase transformations from one point to the next. This can be done by introducing another field, $A_{\mu}(x)$, and replacing the ordinary derivative with the covariant derivative according to the principle of minimal substitution

$$\partial_{\mu} \rightarrow D_{\mu} = \partial_{\mu} - ieA_{\mu} \tag{2.3}$$

and demand that the gauge field $A_{\mu}(x)$ transforms together with $\psi(x)$

$$\begin{aligned} \psi(x) &\rightarrow e^{i\alpha(x)}\psi(x) \\ A_{\mu}(x) &\rightarrow A_{\mu}(x) - \frac{1}{e}\partial_{\mu}\alpha(x) \end{aligned} \tag{2.4}$$

where e is the positron charge. We see that with this prescription, the Lagrangian is invariant under the local gauge transformation. Moreover, this invariance has led to a theory that is no longer free, but describes the interactions of fermions with the vector field A_{μ}

$$\mathcal{L} = \bar{\psi}(x)(i\gamma^{\mu}\partial_{\mu} - m)\psi(x) - \mathcal{L}_{int} \tag{2.5}$$

where the interaction term is

$$\mathcal{L}_{int} = e\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu(x) \tag{2.6}$$

Of course, the gauge field A_μ is interpreted as the photon field. At this point, it has no kinematic degrees of freedom and thus cannot propagate. In order to give the photon field this property, we have to introduce terms containing first order time derivatives in our Lagrangian[9]. If, besides requiring gauge invariance, also symmetry with respect to parity is imposed, only one term is allowed

$$\mathcal{L}_{Gauge} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{2.7}$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.8}$$

is the electromagnetic field strength tensor [1]. We notice that the addition of a mass term for the photon field, $1/2m_A^2 A_\mu A^\mu$, is prohibited by gauge invariance which implies that the photon is massless. Regarding the fermion mass term inherited from the Dirac Lagrangian, we notice that although it obeys the imposed gauge symmetry, it has been put in by hand. The origin of this mass term is thus not explained. The situation becomes even worse when we try to explain the weak interactions in terms of the gauge group $SU(2)_L \times U(1)_Y$. Here, the mass term will not be gauge invariant, since the fermion fields split into two components which transform differently. We shall see in a later section, how we may explain the presence of mass terms anyway through the notion of spontaneous symmetry breaking.

For the moment, we are able to write down the QED Lagrangian

$$\mathcal{L}_{QED} = \bar{\psi}(x)(i\gamma^\mu D_\mu - m)\psi(x) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.9}$$

$SU(2)_L \times U(1)_Y$ gauge invariance and Electroweak theory

The idea of local $U(1)$ gauge invariance can be generalized to other symmetry groups. The electroweak theory is based on the gauge group $SU(2)_L \times U(1)_Y$. For the moment, we will not discuss the meaning of the subscripts L and Y. Actually, we will pretend they are not there, since, for now, we are merely interested in the implications of choosing the more complex gauge group $SU(2) \times U(1)$. We will also take advantage of the fact that it is a product of gauge groups. Therefore, we can study the effects of the two groups separately, and the first of them was explained in the previous section. The second one, $SU(2)$, represents a generalized version of the phase rotation discussed in section 2.2.2. Consider a doublet of Dirac fields which transforms as

$$\psi \to \exp\left(i\alpha^i(x)\frac{\sigma^i}{2}\right)\psi \quad , \quad \psi = \begin{pmatrix} \psi_1(x) \\ \psi_2(x) \end{pmatrix} \tag{2.10}$$

where σ^i are the Pauli matrices and the repeated index implies a summation. This is a local transformation of the Dirac field under $SU(2)$. The Pauli matrices do not commute and thus if

---

[9]As with the Dirac field entering the theory, we consider infinitesimal displacements when the kinematic term is defined. Thus, we only go to first order in the time derivatives.

they are applied in sequence, the order of them will be important. This is contrary to the $U(1)$ symmetry, and the implications show up when we impose the symmetry on the Lagrangian. Taking once again the free field Dirac Lagrangian (2.2) as starting point, we investigate one term at the time. The mass term is still invariant, remembering that $(\sigma^i)^\dagger = \sigma^i$, this is easily realized

$$m\bar{\psi}\psi \to m\bar{\psi} \exp\left(-i\alpha^i(x)\frac{(\sigma^i)^\dagger}{2}\right) \exp\left(i\alpha^j(x)\frac{\sigma^j}{2}\right)\psi = m\bar{\psi}\psi \tag{2.11}$$

We know the procedure to make sense of the derivative. As in the $U(1)$ case, we replace the ordinary derivative with the covariant derivative with the introduction of gauge fields. Employing this we get

$$\partial_\mu \to D_\mu = \partial_\mu - igA_\mu^i \frac{\sigma^i}{2} \tag{2.12}$$

The gauge fields transform along with the Dirac doublet, however, finding the transformation law for for $A_\mu^i \frac{\sigma^i}{2}$ is less trivial due to the non-commutative nature of the Pauli matrices. It can be shown that to first order in α, the transformation law is

$$A_\mu^i \frac{\sigma^i}{2} \to A_\mu^i \frac{\sigma^i}{2} + \frac{1}{g}(\partial_\mu \alpha^i)\frac{\sigma^i}{2} + i[\alpha^i \frac{\sigma^i}{2}, A_\mu^j \frac{\sigma^j}{2}] \tag{2.13}$$

where $[\cdot, \cdot]$ is the commutator [1]. With this transformation, the Dirac Lagrangian is invariant under the $SU(2)$ transformation. Once again, we see that this required the introduction of gauge fields, one for each generator of the symmetry group, summing up to a total of 3. To give these fields kinematic degrees of freedom to propagate, we introduce the field strength tensors

$$F_{\mu\nu}^i = \partial_\mu A_\nu^i - \partial_\nu A_\mu^i + +g\epsilon^{ijk}A_\mu^j A_\mu^k \tag{2.14}$$

where g is a coupling constant and ϵ^{ijk} is the anti-symmetric pseudo-tensor which is naturally introduced through the commutation relations between the Pauli matrices [1]. We see that the main difference between this field strength tensor and the one displayed in (2.8) is that (2.14) contains self-interactions between the gauge fields. This is a direct consequence of the non-commuting generators of the group. For this reason, theories with non-commuting local symmetries are referred to as *non-abelian* theories, whereas theories with commuting local symmetries, such QED, are called *abelian* theories. The field strength tensor itself is no longer gauge invariant, but it is easy to form a gauge invariant quantity from the field strength tensor

$$-\frac{1}{2}\text{tr}\left[(F_{\mu\nu}^i \frac{\sigma^i}{2})^2\right] = \frac{1}{4}(F_{\mu\nu}^i)^2 \tag{2.15}$$

This term contains both cubic and quartic terms in A_μ^i, again reflecting the non-abelian nature of the symmetry. The so-called *Yang-Mills* Lagrangian is now obtained by adding the kinematic term for the self-interacting gauge fields to the Dirac Lagrangian

$$\mathcal{L}_{Yang-Mills} = \bar{\psi}(x)(i\gamma^\mu D_\mu - m)\psi(x) - \frac{1}{4}(F_{\mu\nu}^i)^2 \tag{2.16}$$

This is the simplest example of Yang-Mills vector fields interacting with fermions. It is possible to construct other Yang-Mills theories using other non-abelian symmetry groups.

The electroweak theory (GWS theory) developed by Glashow, Weinberg and Salam unifies electromagnetism and the weak force, by imposing $SU(2) \times U(1)$ gauge invariance on the fermionic fields. Since the weak force is known to act differently on the left-handed and right-handed particles, the fermions are split into a left-handed ψ_L and a right-handed part ψ_R. The gauge group is then given the appropriate subscripts L and Y such that the symmetry group is $SU(2)_L \times U(1)_Y$. Y is called the hypercharge, L denotes the handed-ness and is sometimes replaced by T which denotes the weak isospin. The transformation laws for the fermion fields become

$$
\begin{array}{rclcl}
\psi_L &=& \frac{1}{2}(1 - \gamma_5)\psi &\rightarrow& e^{i\alpha(x)\mathcal{T}^i + i\beta(x)\mathcal{Y}}\psi_L \\
\psi_R &=& \frac{1}{2}(1 + \gamma_5)\psi &\rightarrow& e^{i\beta(x)\mathcal{Y}}\psi_R
\end{array}
\tag{2.17}
$$

where $\mathcal{T}^i$ with $i = 1, 2, 3$ are the generators of $SU(2)_L$ and $\mathcal{Y}$ is the one generator of $U(1)_Y$. The conservation of weak isospin and hypercharge is insured by gauge invariance, and for a given fermion, these quantum numbers uniquely determine how the fermion interacts with other particles in the theory.

Knowing the generators of the gauge group, it is possible to write down the covariant derivative and the field strength tensor as we did for QED and the Yang-Mills theory discussed above. Denoting the three gauge fields associated with $SU(2)_L$ symmetry $\mathcal{W}^i_\mu$ with $i = 1, 2, 3$, and the one gauge field associated with $U(1)_Y$ symmetry $\mathcal{B}_\mu$, we get

$$
D_\mu = \partial_\mu + ig\mathcal{T}^i\mathcal{W}^i_\mu + ig'\frac{\mathcal{Y}}{2}\mathcal{B}_\mu
\tag{2.18}
$$

where g and g' are the couplings between the fermions and the gauge fields. The field strength tensors are

$$
\begin{array}{rcl}
\mathcal{W}^i_{\mu\nu} &=& \partial_\mu\mathcal{W}^i_\nu - \partial_\nu\mathcal{W}^i_\mu - g\epsilon_{ijk}\mathcal{W}^j_\mu\mathcal{W}^k_\nu \\
\mathcal{B}^i_{\mu\nu} &=& \partial_\mu\mathcal{B}_\nu - \partial_\nu\mathcal{B}_\mu
\end{array}
\tag{2.19}
$$

We would like to identify the 4 gauge fields with the physical fields, $W^\pm$, Z and γ, however, the gauge fields in (2.19) are all massless by construction (and do not either possess the proper characteristics, such as the correct couplings to the fermions). The physical content of the gauge fields becomes clear after a linear transformation determined by the Higgs mechanism (see the next section)

$$
\begin{array}{rcl}
\left(\begin{array}{c} W^+_\mu \\ W^-_\mu \end{array} \right) &=& \left(\begin{array}{cc} 1/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{array} \right) \cdot \left(\begin{array}{c} \mathcal{W}^1_\mu \\ \mathcal{W}^2_\mu \end{array} \right) \\
\left(\begin{array}{c} Z_\mu \\ A_\mu \end{array} \right) &=& \left(\begin{array}{cc} \cos\theta_W & -\sin\theta_W \\ \cos\theta_W & \sin\theta_W \end{array} \right) \cdot \left(\begin{array}{c} \mathcal{W}^3_\mu \\ \mathcal{B}_\mu \end{array} \right)
\end{array}
\tag{2.20}
$$

where the weak mixing angle θ_W, also called the Weinberg angle, is defined by

$$
\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}
\tag{2.21}
$$

writing the covariant derivative in terms of the eigenstates defined in (2.20), we get

$$
\begin{array}{rcl}
D_\mu &=& \partial_\mu - i\frac{g}{\sqrt{2}}(W^+_\mu T^+ + W^-_\mu T^-) \\
&& - i\frac{1}{\sqrt{g^2 + g'^2}}Z_\mu(g^2 T^3 - g'^2 Y) - i\frac{gg'}{\sqrt{g^2 + g'^2}}A_\mu(T^3 + Y)
\end{array}
\tag{2.22}
$$

where $T^{\pm} = (T^1 \pm iT^2) = \frac{1}{2}(\sigma^1 \pm i\sigma^2)$. In this form, we identify the dimensionless coefficient in front of A_{μ} with ie, where e is the electric charge, and the quantity behind A_{μ} with the gauge generator (i.e the electric charge quantum number)

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}} \qquad , \qquad Q = T^3 + \frac{Y}{2} \tag{2.23}$$

Using this, we can write the covariant derivative in an even more transparent form

$$D_{\mu} = \partial_{\mu} - i\frac{g}{\sqrt{2}}(W_{\mu}^+ T^+ + W_{\mu}^- T^-) - i\frac{g}{\cos\theta_W}Z_{\mu}(T^3 - \sin^2\theta_W Q) - ieA_{\mu}Q \tag{2.24}$$

where $g = e/\sin\theta_W$. We see that the couplings of all the weak bosons are described by two parameters: the well-measured electron charge e, and the Weinberg angle θ_W.

The W's couple only to left-handed components of fermions, while the Z couples to both left- and right-handed components of fermions, but with different strengths. This is insured by assigning the proper values of isospin and hypercharge to the fermions. The left-handed fermions are assigned to doublets of $SU(2)_L$, while the right-handed are singlets under this group. By specifying the value of T^3 for each field, the value of Y must be given in accordance with (2.23). This means that the hypercharge will be different for left-handed and right-handed components of the quarks and leptons. The right-handed fields have $T^3 = 0$ and thus the hypercharges are twice the electric charges

$$\begin{aligned}
T^3 &= 0 \quad \text{and} \quad Y = -2 \quad \text{for} \quad e_R^-, \mu_R^-, \tau_R^- \\
T^3 &= 0 \quad \text{and} \quad Y = 4/3 \quad \text{for} \quad u_R, c_R, t_R \\
T^3 &= 0 \quad \text{and} \quad Y = -2/3 \quad \text{for} \quad d_R, s_R, b_R
\end{aligned} \tag{2.25}$$

$$\tag{2.26}$$

For the left-handed fields we have

$$\begin{aligned}
T^3 &= \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix} \quad \text{and} \quad Y = -1 \quad \text{for} \quad \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \\
T^3 &= \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix} \quad \text{and} \quad Y = 1/3 \quad \text{for} \quad \begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L
\end{aligned} \tag{2.27}$$

From the covariant derivative in (2.24), it is now clear that A_{μ} (the photon field) does not distinguish between the left- and right-handed particles. Z_{μ} (the Z boson) couples to electric charge, but also to weak isospin (or flavor). $W_{\mu}^{\pm}$ (the $W^{\pm}$ bosons) only couples to left-handed fermions and is responsible for the flavor changing decays, such as $d' \to u$, which take place inside the doublets. However, since the fermion mass eigenstates differ slightly from the gauge eigenstates (the weak isospin doublets) displayed in (2.27), the $W^{\pm}$ can cause decays between all flavors. The relation between the mass eigenstates and the gauge eigenstates is given by the Cabbibo-Kobayashi-Maskawa matrix, V_{CKM}

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.28}$$

where the u, c and t quarks by convention are unmixed. An element V_{xy}, where $x \in \{u, c, t\}$ and $y \in \{d, s, b\}$ in the V_{CKM} matrix describes the probability of transition between the different quark eigenstates.

The kinematic part of the gauge fields is found in the field strength tensors which contains the couplings between the gauge fields themselves. It can be written as

$$
\begin{aligned}
-\frac{1}{4}\mathcal{W}_{i\mu\nu}\mathcal{W}_i^{\mu\nu} &= ig\big[(\partial_\mu W_\nu^+ - \partial_\nu W_\mu^+)W^{\mu-}A^{\nu3} - (\partial_\mu W_\nu^- - \partial_\nu W_\mu^-)W^{\mu+}A^{\nu3} \\
&\quad + \tfrac{1}{2}(\partial_\mu A_\nu^3 - \partial_\nu A_\mu^3)(W^{\mu+}W^{\nu-} - W^{\mu-}W^{\nu+})\big]
\end{aligned}
\tag{2.29}
$$

where $A_\mu^3 = \cos\theta_W Z_\mu + \sin\theta_W A_\mu$ and $g = e/\sin\theta_W$. These couplings reflect the nature of the gauge bosons. While $W^\pm$ carries both electric charge and weak isospin, the Z and the photon carry neither. Thus, the W's couple to themselves and the Z and the photon couple to the W's in cubic and quartic couplings. However, couplings between the neutral gauge bosons are not present. Thus, the vertex $ZZ\gamma$ of which the present study concerns is not present in the Standard Model Lagrangian.

Recalling, that a fermion mass term in the general Yang-Mills Lagrangian was gauge invariant (see (2.11)), we see that this is no-longer the case since a mass term would mix the left-handed and right-handed fields which transform differently

$$
m\bar{\psi}\psi = m\bar{\psi}\left[\tfrac{1}{2}(1 - \gamma_5) + \tfrac{1}{2}(1 + \gamma_5)\right]\psi = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)
\tag{2.30}
$$

The gauge bosons are also massless by construction, which is in contradiction to experiment where the "weakness" of the weak force is thought to be explained by a short range behavior of the particles carrying the force, thus suggesting that they are massive. Although the gauge principle unifies the electromagnetic and the weak forces, the theory, as it stands, lacks the ability to introduce gauge invariant mass terms for both the gauge bosons and the fermions. In the next section we shall see how mass terms can be introduced without spoiling gauge invariance through the notion of spontaneous symmetry breaking.

2.3 Spontaneous symmetry breaking

As explained in the previous section, mass terms for the gauge bosons are prohibited by gauge invariance. This is in contradiction to experimental evidence of the weak decays (beta decays) which suggest massive gauge bosons. However, the problem can be resolved by coupling the SM to some new sector, known as the symmetry breaking sector, which drives the so called spontaneous symmetry breaking of $SU(2)_L \times U(1)_Y$. Though this procedure makes the gauge bosons acquire mass, as we shall see, the mechanism responsible for such breaking is not known. The most popular version of the symmetry breaking sector is build from a self interacting complex scalar called the Higgs scalar.

First, we will show how the mechanism works in a simpler case. Consider a complex scalar field which couples to itself and to an electromagnetic field. The corresponding Lagrangian is

$$
\mathcal{L} = -\tfrac{1}{4}(F_{\mu\nu})^2 + |D_\mu\phi|^2 - V(\phi)
\tag{2.31}
$$

where $D_\mu = \partial_\mu + ieA_\mu$. This Lagrangian is invariant under the local $U(1)$ transformation

$$
\phi(x) \rightarrow e^{i\alpha(x)}\phi(x), \qquad A_\mu \rightarrow A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x)
\tag{2.32}
$$

If the potential is chosen to be of the form

$$
V(\phi) = -\mu^2\phi^*\phi + \frac{\lambda}{2}(\phi^*\phi)^2
\tag{2.33}
$$
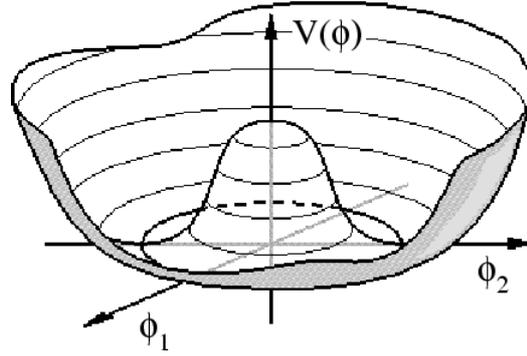
Figure 2.3: This figure shows the Higgs potential, also known as the Mexican hat potential.

with $\mu^2 > 0$, the field will acquire a vacuum expectation value different from zero and the $U(1)$ global symmetry is said to be spontaneously broken (see figure 2.3). The minimum of this potential occurs at

$$\langle \phi \rangle = \phi_0 = \left(\frac{\mu^2}{\lambda} \right)^{1/2} \tag{2.34}$$

or any other value related by the $U(1)$ symmetry. To see the effects of the complex field on the gauge field, it is expanded around its minima (2.34)

$$\phi(x) = \phi_0 + \frac{1}{\sqrt{2}} \left(\phi_1(x) + i\phi_2(x) \right) \tag{2.35}$$

By inserting this expansion in the expression for the kinetic term for ϕ, it becomes clear how the gauge field A_μ will acquire a mass

$$|D_\mu \phi|^2 = \tfrac{1}{2}(\partial_\mu \phi_1)^2 + \tfrac{1}{2}(\partial_\mu \phi_2)^2 + \sqrt{2}e\phi_0 \cdot A_\mu \partial^\mu \phi_2 + e^2 \phi_0^2 A_\mu A^\mu + \cdots \tag{2.36}$$

where terms which are cubic or quartic in the fields have been omitted. The last term in (2.36) is recognized as the mass term for the A_μ field since $\mathcal{L}_{\mathrm{mass}} = \tfrac{1}{2} m_A^2 A_\mu A^\mu$. Thus, the mass is

$$m_A^2 = 2e^2 \phi_0^2 \tag{2.37}$$

and is direct consequence of a non-zero vacuum expectation value of the scalar field ϕ.

The same mechanism can be employed on the $SU(2)_L \times U(1)_Y$ gauge theory described in the previous section. This is done by introducing four new scalar fields which spontaneously break $SU(2)_L \times U(1)_Y$ while leaving the gauge symmetry associated with electric charge unbroken. This will make the three weak bosons acquire mass while the photon remains massless. The scalar fields are arranged in a complex gauge invariant isospin doublet ($T = 1/2$) with hypercharge $Y = 1/2$

$$\Phi \rightarrow e^{i\alpha^a \tau^a} e^{i\beta} \Phi \tag{2.38}$$

where $\tau^a = \sigma^a/2$. If the scalar doublet acquires a vacuum expectation value of the form

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.39}$$

then a gauge transformation with $\alpha^1 = \alpha^2 = 0$ and $\alpha^3 = 2\beta$ will leave the vacuum expectation value invariant. This provides us with a theory that contains one massless gauge boson, the photon, and three massive gauge bosons, the $W^\pm$ and Z, which acquire their masses through the Higgs mechanism. By expanding Φ around its minimum, the covariant derivative acting on Φ contains mass terms for these three bosons.

$$D_\mu \Phi = \left(\partial_\mu - ig\mathcal{W}_\mu^a \tau^a - i\tfrac{1}{2}g'\mathcal{B}_\mu \right) \Phi \tag{2.40}$$

The relevant terms for producing the masses are

$$\mathcal{L}_{\text{mass}} = \frac{1}{2} \begin{pmatrix} 0 & v \end{pmatrix} \left(g\mathcal{W}_\mu^a \tau^a + \frac{1}{2}g'\mathcal{B}_\mu \right) \left(g\mathcal{W}^{b\mu}\tau^b + \frac{1}{2}g'\mathcal{B}^\mu \right) \frac{1}{2} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.41}$$

and evaluating this explicitly using the Pauli matrices, we get

$$\mathcal{L}_{\text{mass}} = \frac{1}{2}\frac{v^2}{4} \left[g^2(\mathcal{W}_\mu^1)^2 + g^2(\mathcal{W}_\mu^2)^2 + (-g\mathcal{W}_\mu^3 + g'\mathcal{B}_\mu)^2 \right] \tag{2.42}$$

By performing the rotation into the physical fields, as explained in section 2.2.2 (see equation (2.20)), we arrive at the physical gauge fields

$$
\begin{array}{llll}
W_\mu^\pm & = \frac{1}{\sqrt{2}} \left(\mathcal{W}_\mu^1 \mp i\mathcal{W}_\mu^2 \right) & , & \text{with mass } m_W & = g\frac{v}{2} \\
Z_\mu & = \frac{1}{\sqrt{g^2+g'^2}} \left(g\mathcal{W}_\mu^3 - g'\mathcal{B}_\mu \right) & , & \text{with mass } m_Z & = \sqrt{g^2+g'^2}\,\frac{v}{2} \\
A_\mu & = \frac{1}{\sqrt{g^2+g'^2}} \left(g\mathcal{W}_\mu^3 + g'\mathcal{B}_\mu \right) & , & \text{with mass } m_A & = 0
\end{array}
$$

The masses of the $W^\pm$ and the Z are not independent, but related through the Weinberg angle

$$m_W = m_Z \cos\theta_W \tag{2.43}$$

The fermions are given masses by coupling them to the Higgs doublet through Yukawa couplings. This introduces a new coupling for each fermion and the masses are therefore not predicted, but has to be measured experimentally. The Higgs fields will also introduce a new scalar particle, the Higgs boson. This boson will acquire a mass

$$m_H = 2v^2\lambda \tag{2.44}$$

By expanding the scalar Higgs field in equation (2.41) to first order, such that it reads $v + h(x)$ where $h(x)$ is the Higgs boson, couplings between the gauge bosons and the Higgs boson are introduced. In the case where the Higgs boson has a very large mass, $m_H \sim 1TeV$, it will not easily be found as a resonant peak. However, its couplings to the gauge bosons will allow measurable effects that can be modelled in an effective Lagrangian. These will mimic trilinear gauge coupling which shows that a study of diboson production is important in discovering new physics. Effective Lagrangians and trilinear gauge couplings will be explained in the next chapter.

The Higgs boson is the only particle predicted by the SM not yet discovered. The Higgs mass is a parameter to be measured experimentally. Upper theoretical bounds exist which are of the order 1TeV, and previous searches at LEP and at the Tevatron has excluded the lower mass range up to about 114GeV. Despite the attractive feature that the Higgs give masses to the gauge bosons and the fermions, the Higgs sector is considered to be the least satisfactory part of the Standard Model. The reason is that none of the masses are predicted and that a scalar particle, the Higgs boson, appeared in the process.

2.4 Symmetries in a quantum theory

Until now, the description of the electroweak Standard Model in terms of the non-abelian gauge group $SU(2)_L U(1)_Y$ has been entirely classical. However, the Standard Model is a *quantum* field theory and thus the gauge theory presented above needs quantization. There are several challenges connected to the quantization of the Standard Model—this section will briefly address one of them, the *anomalies*. In the next chapter, we will explore the possibility of including terms in the Lagrangian that generate the $ZZ\gamma$ vertex. In the previous section, it was shown that the existence of the gauge bosons is directly connected with the symmetries of the theory. The present section will show that there are deep subtleties involved with the symmetries of the Standard Model indicating that the inclusion of additional interactions will introduce additional difficulties. We shall return to these matters in the next chapter.

Naively, one would think that if a classical theory has a specific symmetry implemented, this symmetry will also be present after the theory is quantized. This was what most theoreticians thought in the 1960s at the time where gauge theories were explored in the hope of finding a way to describe the weak and strong forces [2].

However, as they discovered, a classical symmetry does not necessarily survive the process of quantization. This can be understood in terms of transition amplitudes. Consider a particle that travels between two points in space. Classically, the particle will follow the path that minimizes the action, i.e. the path that satisfies the classical equations of motion. In quantum mechanics, the particle is allowed to follow *all* paths between the two points, however with different probabilities for each path. In the path integral formulation, the amplitude for the transition between the two points is given by $\int \mathcal{D}q \exp[iS(q)]$ where the integral includes all paths[10]. From this we see the implications regarding the symmetry property. Classically, a symmetry is recognized if the action is invariant under the corresponding transformation. In quantum mechanics, the measure $\mathcal{D}q$ has to experience the same invariance—which it not necessarily does. Thus, at first sight it is not at all obvious whether the classical symmetry is still present after quantization. If the symmetry does not survive the process of quantization it is called an *anomaly* [2]. Often the term 'anomaly' is used for the current that corresponds to the symmetry transformation (through Noether's theorem)—or rather the divergence of this current, since this will be non-zero if the symmetry is absent.

The Standard Model offers several examples of such anomalies. Some of them originate from the non-conservation, or breaking, of global symmetries, others from breaking of local symmetries. In turns out that some of these anomalies are problematic to the theory, while others can be allowed. Gauge anomalies lead to an inconsistency, since gauge symmetry is required in order to cancel unphysical degrees of freedom with a negative norm (such as a photon polarized in the time direction) [4]. Thus, it is necessary to get rid of the gauge anomalies—they must cancel. The possible appearance of gauge anomalies comes from the fact that the GWS bosons couples to fermions in a chiral fashion. Within GWS theory, the requirements from experiment that the weak interaction currents are left-handed has forced us to choose a chiral gauge coupling.

Fortunately, it turns out that the gauge anomalies cancel out in the Standard Model as required [1]. However, only through a remarkable interplay between the quantum numbers of

---

[10]This amplitude should also include some kind of normalization constant, which however is of no importance in this context. In fact, the expression really should be explained in more detail—e.g. what does the path integral mean? Here, we shall simply think of the integral as a sum. I refer to the commonly used textbooks [1], [2] and the lecture notes [3] for a detailed description.

quarks and leptons. Given the assignments of weak isospin and hypercharge to the left- and right-handed fields, the cancellation of the gauge anomalies requires that leptons and quarks appear in complete multiplets with the structure of $(E_L, e_R, Q_L, u_R, d_R)$, where E_L and Q_L denotes the isospin doublets and e_R, u_R and d_R denotes the singlets (see section 2.2.2). This set of fields is recognized as the first generation of fermions in SM. Thus, the consistency of the theory requires that quarks and leptons appear in nature in equal numbers, organizing themselves into successive generations in this way.

For a thorough and excellent description of anomalies, see [4]. References [2] and [1] also present details on some anomalies.

2.5   Status of the Standard Model

Even though the Standard Model has been very successful in describing almost every phenomena in particle physics within reach of the former and currently operating particle accelerators, it suffers from several flaws which make it unsuitable as a fundamental theory. These problems are both of theoretic as well as observational origin. However, the very same problems also leave clues as to where new physics may be found in the search for a more fundamental theory. This section will briefly describe some the successes of the Standard Model and give an overview of some of the open questions and proposed solutions.

Successes of the Standard Model

The perhaps most precise verification of the electromagnetic part of the Standard Model is the agreement between calculation and measurement of the anomalous magnetic moment of the electron. This has been done with higher order loop calculations and an extremely sensitive experimental setup which has revealed an agreement to 10 significant digits [5]. This agreement is among the best known in physics today. The Standard model also has predicted the existence of the heavy vector bosons in the electroweak sector. The direct observations of the W [6], [7] and Z [8], [9] bosons at CERN around 1983 signified a major milestone in particle physics as it put the Standard Model on solid ground. The exact shape and height of the Z resonance has furthermore been used to test the picture of exactly 3 fermionic generations provided by the Standard Model[11]. The success of the Standard Model continues into physics governed by the strong force. QCD has been able to explain the wide variety of hadronic states using the notion of quarks and colour.

Problems in the Standard Model

First of all, for purely esthetical reasons, the Standard Model seems unpleasing. It is unsatisfying that there are so many free parameters to be determined by experiment: Gauge couplings, CKM parameters, masses and the parameters associated with the electroweak symmetry breaking. This raise questions such as why the fermion masses are so widely spread ranging from the top quark at ~ 175 GeV to the electron at ~ 0.5 MeV, and why there are exactly 3 generations of matter particles.

At a more technical level, the gauge group structure $U_Y(1) \times SU_L(2) \times SU_C(3)$ seems like a rather arbitrary choice. It is also a bit surprising that the gauge group $SU_L(2)$ has a chiral

---

[11]This was one of the great achievements at LEP [10].

structure in the Standard Model—why should an internal gauge symmetry know about the chirality of fermions, which is a space-time property? In 1967, Coleman and Mandula [11] proved a theorem which puts severe restrictions on the groups that can serve as physical symmetry groups. The theorem states that it is impossible to combine space-time and internal symmetries in any but a trivial way. The fact that this happens anyway in the Standard Model is only because the gauge symmetry is spontaneously broken—in this case the Coleman-Mandula theorem does not hold.

There has been several attempts to implement the Standard Model in a larger symmetry group which (perhaps) would make it look less arbitrary and more fundamental. This procedure also include a unification of the electroweak and the strong forces in a Grand Unified Theory (GUT). However, none of these attempts have been successful so far.

Aside from these 'non-devastating' issues, there are also some problems in the Standard Model which need to be solved in a fundamental theory.

The biggest question is how to include gravity in the Standard Model. The problem with gravity when trying to describe it in quantum field theory is its non-renormalizable action, the Einstein-Hilbert action, which is hard to get around. However, gravity is so much weaker than the other forces that it can be neglected in almost all cases. This, on the other hand, leaves us with the question of *why* gravity is so much weaker. From the electroweak scale $M_{EW} \sim v = 246 \text{GeV}$ to the Planck scale $M_{Planck} \sim 10^{19} \text{GeV}$ which is where gravity is expected to be equally strong to the other forces, there is a huge gap of many orders of magnitude. If the SM is valid over this vast range, the Higgs boson will receive enormous radiative corrections from the vacuum polarization. In order to have a light Higgs (i.e. $m_H \sim m_{EW}$) there has to be a large amount of fine-tuning cancellations between the quantum corrections. While this is not prohibited by any fundamental principle, it is believed to be highly unnatural. The problem is known as the Hierarchy problem.

Another problem is the masses of the neutrinos. In the Standard Model the neutrinos are massless. However, there is evidence proving otherwise through neutrino oscillations [12]. Even though the neutrino masses are very small, it is possible to measure the mass difference between the 3 flavors through their oscillations into one another over large distances. This is actually proof of physics beyond the Standard Model.

As we see, many problems in the SM could be connected to the origin of mass. This issue will undoubtedly be related to gravity. Whatever theory which describes gravity at a quantum level, it will affect the couplings between the gauge bosons, especially the neutral ones which are usually not affected in models that do not include gravity. This provides additional motivation for studying the gauge boson vertex $ZZ\gamma$.

Chapter 3

Anomalous Triple Gauge Couplings

In this chapter we introduce the anomalous triple gauge vertex $ZZ\gamma$. The framework is constituted by the topics covered in the previous chapter, starting with considering the feynman rule for the vertex and then uncovering the terms that is added to the SM Lagrangian to generate the vertex. The goal is to determine which observable effects the inclusion of the $ZZ\gamma$ triple vertex introduces.

In the Standard Model, the electroweak gauge bosons interact not only with matter particles, but also with each other because of the non-Abelian nature of the gauge group. These interactions manifest themselves as vertices between three (or four) gauge bosons, such as W^+W^-Z and $W^+W^-\gamma$ which we call triple (quartic) gauge boson vertices, or in short TGV's. The Lagrangian contains the corresponding terms multiplied by each their coupling constant which tells us the strength of a particular vertex. In general, more than one term will contribute to the same vertex. The coupling constants representing the terms which contribute to triple gauge vertices are referred to as triple gauge couplings (TGC's). If a specific coupling is absent within the theory or model considered, the coupling is labeled *anomalous*. If all the couplings contributing to the same vertex are zero, the vertex itself is completely absent within the theory and it is also said to be anomalous.

One of the most important reasons for studying triple gauge couplings is the fact that they are tightly connected with the symmetry properties of the Standard Model. They reflect the full mathematical gauge group structure of the fundamental interactions. By now, it should be clear that it is far from easy to write down a theory based on a gauge group which includes TGC's that are absent in the Standard Model. Whatever is modified at the group structure level will affect not only the interactions between the gauge fields, but in fact the gauge fields themselves. This means that one cannot simply make a small adjustment to the Standard Model Lagrangian to include anomalous couplings.

Recognizing the fact that a change in the fundamental group structure is a very sensitive operation unavoidably leads to the conclusion that the presence of anomalous TGC's imply very different physics from what the SM offers. Experimental attempts to measure anomalous TGC parameters can, as will be explained below, be seen as a way to probe the low energy remnants of new physics which is operating at a much higher energy scale. Measurements of this type would be most interesting in the scenario where direct searches for new particles which affect the gauge boson interactions fail to observe any substantial deviation from the Standard Model. Physics beyond the Standard Model (such as compositeness models[1] [13],

---

[1]The composite Z boson theory states that the Z boson is not it self an elementary particle, but consists

excited Z bosons [14], Minimal Supersymmetric SM[2] [15], or Non Commutative Geometry [16] could alter the cross sections and the production kinematics of the $Z\gamma$ final state. A general study of all anomalous triple gauge boson vertices could give constraints on the possible structure of the new physics and perhaps show which part of the Standard Model the new physics affects and the symmetry properties it preserves.

As shown in the previous sections not all triple gauge couplings in the electroweak part of the Standard Model are zero. For instance, the presence of W^+W^-Z and $W^+W^-\gamma$, also called the charged triple gauge vertices, have already been verified at LEP [17]—these are the only triple gauge vertices allowed in the Standard Model at tree-level. Other gauge boson self interactions such as ZZZ, $ZZ\gamma$, $Z\gamma\gamma$ and $\gamma\gamma\gamma$ (referred to as neutral TGV's) are not allowed in the Standard Model, because neither the Z nor the γ carries electric charge or weak isospin which are the quantum numbers to which these gauge bosons couple. Vertices containing an odd number of W bosons ($W^\pm ZZ$, $W^\pm\gamma\gamma$, $W^\pm Z\gamma$ and $W^\pm W^\pm W^\pm$) are excluded by charge conservation. Among the tree-level self interactions we also have the quartic couplings which represent the meeting of four gauge bosons at the same vertex. In the Standard Model the following four interactions are allowed: $W^+W^-\gamma\gamma$, $W^+W^-\gamma Z$, W^+W^-ZZ and $W^+W^-W^+W^-$ which can be read off from the terms in (2.29). Since our focus will be on the anomalous $ZZ\gamma$ vertex the quartic couplings will not be discussed further.

When speaking of $Z\gamma$ production at tree-level within the Standard Model, one implicitly means t- and u-channel diagrams (see figure 3.1). The triple gauge vertices belong to the s-channel diagram. Even though the s-channel diagram is absent at tree-level in the Standard Model, it receives contributions from loop diagrams through the so called "fermion-triangle-diagram" or higher order box diagrams. As shown is figure 3.1, this diagram contains a loop of fermions instead of the tree-level trilinear vertex. In principle, this contribution should also be accounted for, since the size of it is roughly of the same order as the contributions from the anomalous couplings [18]. However, the loop diagram will be neglected in this study for reasons that will be explained in section 3.4.

In general, the neutral triple gauge vertices have received less attention than the charged ones. The reason is that the charged boson vertices already gets tree-level SM contributions, whereas the neutral ones do not, so that they may be considered as purely anomalous. In the next sections, we will show the conventional way of including anomalous contributions to the production of the diboson final state $Z\gamma$. This is done in the effective Lagrangians approach by introducing higher dimensional operators in a generic operator expansion which generates the $ZZ\gamma$ vertex. First, however, we consider the most general forms of vertex functions (i.e. feynman rules) that contribute to the anomalous s-channel production. Then, we will try to understand which kind of operators in the Lagrangian that will produce these vertex functions.

3.1 Vertex functions for $ZZ\gamma$

The vertex functions for $ZZ\gamma$ have been studied extensively in the literature [15, 19–23]. The most general studies have all three bosons off-shell, e.g. [19]. Others consider the case where only the intermediate boson is off-shell while the two final state bosons are on-shell. In the

---

of the hypothetical particles: *preon* and *anti-preon*. Preons couple to photons, and hence the trilinear $ZZ\gamma$ vertex is a consequence of the radiation of a photon by the preon.

[2]In MSSM, heavy SUSY particles in one-loop diagrams can enhance the strength of the couplings between the Z boson and the photon.—see e.g. figure 3.1 with the fermion triangle replaced by a triangle charginos. The expected values of the anomalous couplings in the MSSM model are of the order of $\mathcal{O}(10^{-3})$ [15].
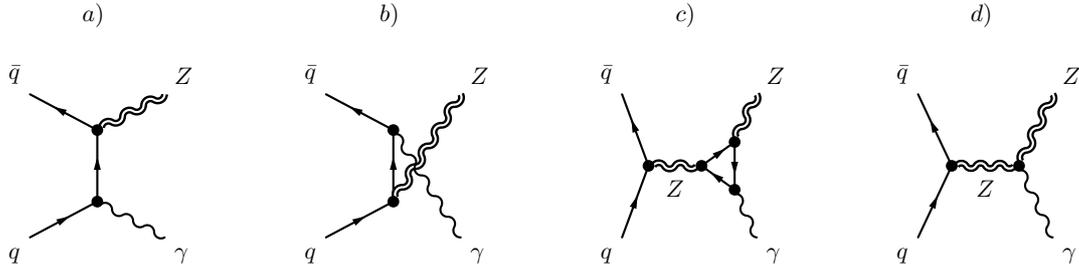
Figure 3.1: a) SM tree-level t-channel $Z\gamma$ diagram. b) SM tree-level u-channel $Z\gamma$ diagram. c) SM one-loop s-channel diagram, also known as the fermion triangle diagram. d) Anomalous tree-level s-channel $Z\gamma$ diagram.

present study, the situation in which the two final state bosons are on-shell is addressed. This is expected to be a good starting point in a new experiment such as ATLAS, since the off-shell cases will be more challenging, partly due to the fact that an off-shell photon will decay, but also since it is more difficult to identify the decay products of off-shell bosons.

The form of the vertex function is determined by imposing symmetries. Given the good agreement between present data and SM predictions, it is natural to expect that any signal of new physics from a large scale will respect the SM symmetries at lower energies.

In the present study, we use the so-called "$U(1)_{em}$ framework" which consists of the following constraints on the vertex functions: Lorentz invariance, $U(1)_{em}$ gauge symmetry and Bose symmetry. Bose symmetry is special to neutral TGV's and does not apply to the charged sector. This is due to the fact that in all neutral TGV's, there will always be at least two (potentially) identical bosons involved. In the case of $ZZ\gamma$, Bose symmetry means that the two Z's can be interchanged at the vertex, if they have the same mass. Therefore, unless the vertex function is symmetric when interchanging the two Z's, it must vanish when they have the same mass.

It is also relevant to consider $SU(2)_L \times U(1)_Y$ gauge invariance, however, most studies [15, 19, 20, 22, 23] choose not to include this symmetry. The reason is that by doing so, one has to decide which specific realization of the electroweak symmetry breaking sector to use. Since this sector is one the big questions to be answered in the near future, it is convenient not to assume anything about it.

Given the constraints of the $U(1)_{em}$ framework, the most general vertex function (shown schematically in figure 3.2) is

$$\Gamma^{\alpha\beta\mu}_{Z\gamma Z}(q_1, q_2, P) = \frac{i(P^2 - m_Z^2)}{m_Z^2}\left[h_1^Z(q_2^\mu g^{\alpha\beta} + q_2^\alpha g^{\mu\beta}) + \frac{h_2^Z}{m_Z^2}P^\alpha\left(P_\rho q_2^\rho g^{\mu\beta} - q_2^\mu P^\beta\right)\right.$$

$$\left. - h_3^Z \epsilon^{\mu\alpha\beta\rho}q_{2\rho} - \frac{h_4^Z}{m_Z^2}P^\alpha \epsilon^{\mu\beta\rho\sigma}P_\rho q_{2\sigma}\right] \tag{3.1}$$

where P, q_1, q_2 refer to the four momenta of the participating bosons, $g^{\alpha\beta}$ is the metric tensor, $\epsilon^{\alpha\beta\rho\sigma}$ is the anti-symmetric pseudo-tensor, and the h_i^Z are four coupling constants, the size of which give the strength of that term to the vertex [20]. The overall factor $P^2 - m_Z^2$ in
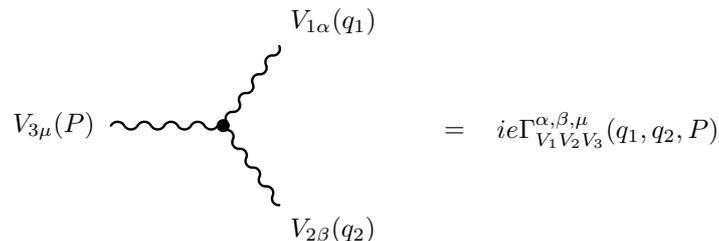
$$= \quad ie\Gamma^{\alpha,\beta,\mu}_{V_1 V_2 V_3}(q_1, q_2, P)$$

Figure 3.2: Vertex function for triple gauge couplings between the vector bosons V_1, V_2 and V_3. P, q_1 and q_2 indicate the four momenta of the intermediate and the two final state bosons, respectively.

(3.1) means that the vertex vanishes when the two Z's have the same mass and is a direct consequence of Bose symmetry, as explained above.

The reason why there are exactly four couplings can be seen from a helicity argument. Helicity is defined as spin projection onto the momentum vector. The gauge bosons are spin-1 particles which imply that their spin projections can take on the values $-1, 0, 1$ in units of $\hbar$. However, massless particles move with the speed of light which means that their spins projections look the same regardless of Lorentz reference frame. In other words, it is not possible to boost to the rest frame and thus the massless particle will always appear to have its spin fixed with respect to its direction of motion.

Applying this to the current situation, we see that the massless photon in the final state can have two helicities, -1 and +1. The final state Z boson can in addition have helicity 0, since it is massive. This restricts the total number of final state helicities to $2 \times 3 = 6$. However, the two combinations where the final state boson spins are pointing in the same direction ($J = \pm 2$) cannot be reached by the virtual Z vector boson in the s-channel exchange, since it only has helicity states -1, 0 and 1. Thus only four out of six possible helicity amplitudes are allowed, and they generally correspond to four different couplings[3].

3.2 Generic operator expansion for the $ZZ\gamma$ vertex

In the absence of a specific model of new physics which can introduce the anomalous couplings, effective Lagrangians are extremely useful. An effective Lagrangian parametrizes in a model independent way the low-energy effects of phenomena to be found at higher energies. It is only necessary to specify the particle content and the symmetries of the low-energy approximation.

The effective Lagrangian will contain an infinite number of terms, which, however, are organized in powers of Λ^{-1}, where Λ is the energy-scale at which the new physics can be directly seen. Thus, at energies much smaller than this scale, only the first few terms of the expansion are important. This approach makes it possible to include effects of new physics on the SM particles without knowing the precise origin of the new physics.

The perhaps best-known effective Lagrangian in particle physics comes from Fermi theory, which is an approximation to the weak interactions at energies way below the mass of the vector gauge bosons in the Standard Model, i.e. $\Lambda = m_W$. Within the Standard Model,

---

[3]If the final state Z and γ are allowed to be off-shell, we get $3 \times 3 - 2 = 7$ final state helicities since the photon then also can have helicity 0.

the charged-current weak interaction between two fermions is described by the exchange of a W-boson

$$\frac{g^2}{8}\bar{\psi}\gamma_\mu(1-\gamma_5)\psi\frac{1}{q^2-m_W^2}\bar{\psi}\gamma^\mu(1-\gamma_5)\psi \tag{3.2}$$

where $1/(q^2-m_W^2)$ is the W propagator and q^2 is the momentum transfer (energy scale) of the interaction [24]. The propagator can be expanded in powers of q^2/m_W^2

$$\frac{1}{q^2-m_W^2} = -\frac{1}{m_W^2}\left[1+\frac{q^2}{m_W^2}+\left(\frac{q^2}{m_W^2}\right)^2+\cdots\right] \tag{3.3}$$

Thus, we see that (3.2) can be written as a sum of an infinite number of terms. In cases where the energy transfer is significantly below the W mass, it is clear that only the contribution from the first term in the expansion will be important while the rest of the terms can be neglected. Writing this out explicitly, we see that this is nothing but the four-fermion interaction of Fermi theory

$$-\frac{G_F}{\sqrt{2}}\bar{\psi}\gamma_\mu(1-\gamma_5)\psi\bar{\psi}(1-\gamma_5)\psi \tag{3.4}$$

The fact that higher dimensional operators will be suppressed by powers of the new physics scale has to do with the so-called "mass-dimension" of the operators, which has briefly been discussed in the end of section 2.1.3 in connection with renormalizability. Ultimately, we want to consider the action, S, which is a dimensionless quantity. It is linked to the Lagrangian (or more correctly, the Lagrangian density) through the four dimensional integral $\int dx^\mu$ which has mass dimension -4. Thus, the operators of the Lagrangian must have dimension +4 in order for S to be dimensionless. This is true for all the operators in the SM Lagrangian.

However, when doing a generic operator expansion, we may have operators in the Lagrangian with a dimension larger than 4, so these operators must be multiplied by a quantity which can bring down the dimension to 4. This quantity is the scale of the new physics, Λ, which will enter the appropriate terms as $1/\Lambda^n$, where n equals the "excess" mass dimensions of the operators.

When writing down a generic operator expansion for the $ZZ\gamma$ vertex, we need to identify which kind of operators can be included. As mentioned above, we first have to specify the fields involved and then impose symmetries as appropriate.

In principle, the combination of operators which can give rise to the neutral anomalous trilinear gauge couplings may involve the four vectors A_μ and Z_μ, together with the antisymmetric tensors $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $Z_{\mu\nu} = \partial_\mu Z_\nu - \partial_\nu Z_\mu$. Since the higher dimension operators will be suppressed, it is natural to consider the lowest dimension operators first.

We see that due to $U(1)_{em}$ gauge symmetry, the electromagnetic field cannot appear as A_μ. Therefore, the photon must appear in any term through the field tensor $F_{\mu\nu}$.

Considering first dimension 4 operators, this means that any such operator will necessarily be of the form $Z_\mu Z_\nu F^{\mu\nu}$. However, due to the antisymmetry of $F^{\mu\nu}$, no 4 dimensional operators can be allowed (since by changing summation indices we get: $Z_\nu Z_\mu F^{\nu\mu}$, while the antisymmetry suggests: $Z_\nu Z_\mu(-F^{\nu\mu})$, which are only compatible if the operator is zero). Going one dimension up, the 5 dimensional operators are trivially excluded since they will contain an index which is not summed over. The same goes for any other odd dimensional operator.

Dimension 6 and 8 operators will be suppressed by the the factors $1/\Lambda^n$, where $n = 2, 4$, respectively. Higher dimensional operators will be even more suppressed and therefore we will not consider operators with dimension higher than 8. This is the conventional approach when working within the $U(1)_{em}$ framework [20, 22].

It can be shown that it is impossible to have structures involving only the field strength tensors $Z_{\mu\nu}$ and $F_{\mu\nu}$ due to their anti-symmetric properties [21]. Thus, operators like $Z_{\mu\nu}Z^{\nu\sigma}F_\sigma{}^\mu$ cannot exist in the expansion. To generate the $ZZ\gamma$ vertex, we must use at least one Z boson in the form Z_μ (possibly with one or more derivatives acting on it).

In general, the operator expansion can be written on the form

$$\mathcal{L} = e\left(\sum_i l_i \mathcal{O}_i + \sum_i l' \mathcal{O}'_i\right) \tag{3.5}$$

where the operators have been split into CP conserving operators ($\mathcal{O}_i$) and CP violating operators ($\mathcal{O}'_i$) with corresponding coupling constants l and l'.

In the previous section, we saw that the $ZZ\gamma$ vertex function contained four couplings h_i^Z, and therefore the Lagrangian generating the vertex function will also have these four couplings.

The effective Lagrangian generating the $ZZ\gamma$ vertex is

$$\begin{aligned}\mathcal{L}_{ZZ\gamma} &= \frac{-e}{m_Z^2}\Big[h_1^Z(\partial^\sigma Z_{\sigma\mu})Z_\beta F^{\mu\beta} + h_3^Z(\partial_\sigma Z^{\sigma\rho})Z^\alpha \tilde{F}_{\rho\alpha} \\ &\quad + \frac{h_2^Z}{m_Z^2}\left[\partial_\alpha\partial_\beta(\Box + m_Z^2)Z_\mu\right]Z^\alpha F^{\mu\beta} + \frac{h_4^Z}{2m_Z^2}\left[(\Box + m_Z^2)\partial^\sigma Z^{\rho\alpha}\right]Z_\sigma \tilde{F}_{\rho\alpha}\Big]\end{aligned} \tag{3.6}$$

where $\tilde{F}_{\mu\nu} = 1/2\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}$ and $\Box = \partial_\mu\partial^\mu$ [20]. By counting mass dimensions (i.e. counting indices) we recognize the first two terms as dimension 6 operators and the last two terms as dimension 8 operators. The operators can be divided into the two classes suggested in equation (3.5) by recalling that the photon and the Z boson are C and P odd ($C = P = -1$) and that $P_{\partial_\mu} = -1$. Performing a C-transformation we see that h_1^Z violates C-parity:

$$h_1^Z(\partial^\sigma Z_{\sigma\mu})Z_\beta F^{\mu\beta} \quad \xrightarrow{C} \quad h_1^Z(\partial^\sigma(-Z_{\sigma\mu}))(-Z_\beta)(-F^{\mu\beta}) \tag{3.7}$$

while

$$h_1^Z(\partial^\sigma Z_{\sigma\mu})Z_\beta F^{\mu\beta} \quad \xrightarrow{P} \quad h_1^Z(-\partial^\sigma(+Z_{\sigma\mu}))(-Z_\beta)(+F^{\mu\beta}) \tag{3.8}$$

shows that h_1^Z conserves P-parity, and thus in total is CP violating. The same is true for h_2^Z. The remaining two terms in (3.6) can be investigated in the same way. Remembering, that $\epsilon_{\alpha\beta\gamma\delta}$ is a pseudotensor and thus have $P_\epsilon = -1$, we see that h_3^Z and h_4^Z are both C and P violating and so in total are CP conserving.

3.3 Unitarity and Form Factors

As with most effective theories, the predictions of the effective Lagrangian (3.6) are in conflict with unitarity at high energies, because the effective couplings will increase with energy due to their non-renormalizable nature (as explained in section 2.1.4). It is important to constrain

this behavior, especially at hadron colliders where the center of mass energy of a given reaction is not known due to the parton distribution functions of the proton. Otherwise, the result will be unphysically large cross sections at high energies leading to an overestimate of experimental sensitivities.

Unitarity would be recovered if the heavy modes of the underlying theory were included in the model again, but this is only possible if we actually know this underlying theory. Hence, we need to address the problem of unitarity by other means.

One solution to the unitarity problem is to invoke a *form factor* behavior on the coupling constants.

Conventionally, the form factor is assumed to have a power law behavior expressed by the generalized dipole form factor

$$h_i^Z(s) = \frac{h_{i0}^Z}{(1 + s/\Lambda^2)^n} \tag{3.9}$$

where $\sqrt{s}$ is the center of mass energy of the interaction and Λ is the energy scale of the new physics [18]. An appropriate choice of n and Λ will constrain the high-energy behavior.

By considering the helicity amplitudes $F(q(\sigma_1)\bar{q}(\sigma_2) \to Z(\tau_1)\gamma(\tau_2))$, where σ_i and τ_i denotes the helicities of the incoming quarks and final state bosons, it is possible to find the exact behavior of the anomalous contributions to the cross section. The behavior is given by

$$
\begin{aligned}
F_i(q(\sigma_1)\bar{q}(\sigma_2) \to Z(\tau_1)\gamma(\tau_2)) &\propto \left(\frac{\sqrt{s}}{m_Z}\right)^3 &, \; i = 1, 3 \\
F_i(q(\sigma_1)\bar{q}(\sigma_2) \to Z(\tau_1)\gamma(\tau_2)) &\propto \left(\frac{\sqrt{s}}{m_Z}\right)^5 &, \; i = 2, 4
\end{aligned}
\tag{3.10}
$$

where F_i is the contribution to the helicity amplitude for the anomalous coupling h_i^Z [18]. The helicity amplitudes grow with center of mass energy $\sqrt{s}$, as we expect.

Equation (3.10) show exactly the energy dependence we need to constrain in order to preserve unitarity. By investigating the energy dependence in more detail than presented here, the helicity amplitudes can be subjected to constraints that make them preserve unitarity, and by using the form factor (3.9) these constraints can be translated into limits on the couplings h_{i0}^Z [18]. Assuming that only one anomalous coupling is non-zero at the time and that $\Lambda \gg m_Z$ the following is obtained

$$
\begin{aligned}
|h_{10}^Z|, |h_{30}^Z| &< \frac{(\frac{2}{3}n)^n}{(\frac{2}{3}n - 1)^{n-3/2}} \cdot \frac{0.126\text{TeV}^3}{\Lambda^3} \\
|h_{20}^Z|, |h_{40}^Z| &< \frac{(\frac{2}{5}n)^n}{(\frac{2}{5}n - 1)^{n-5/2}} \cdot \frac{2.1 \cdot 10^{-3}\text{TeV}^5}{\Lambda^5}
\end{aligned}
\tag{3.11}
$$

An investigation of these inequalities (or by comparing the powers of s in (3.9) and (3.10)) shows that unitarity will be satisfied for

$$
\begin{aligned}
n &> \frac{3}{2} &, \text{ for } h_1^Z \text{ and } h_3^Z \\
n &> \frac{5}{2} &, \text{ for } h_2^Z \text{ and } h_4^Z
\end{aligned}
\tag{3.12}
$$

In the present study, we use the traditional values (see e.g. [22]) $n = 3$ for $h_{1,3}^Z$ and $n = 4$ for $h_{2,4}^Z$. These choices guarantee that unitarity is preserved and that terms proportional to $h_{20,40}^Z$ have the same high-energy behavior as those proportional to $h_{10,30}^Z$, and also makes a direct comparison with other studies possible.

The unitarity limits (3.11) constitute the reference when limits are extracted from experiment. It is clear, that experimental constraints on the anomalous couplings only are useful, if they are more stringent than the unitarity bounds for a given Λ.

It should also be mentioned, that the appearance of form factors actually implies that the effective Lagrangian description breaks down, i.e. we are probing the effects of the anomalous couplings at the scale of the new physics. Remembering, that the effective Lagrangian approach assumed that the new physics is operating at a much higher energy scale, this appears to be in contradiction with the use of form factors. However, the cross sections related to direct observations of the new physics might be too small to be immediately observed or the experimental signatures might be hidden by backgrounds. Thus, form factors should be seen as a way of extending the use of effective Lagrangians to the entire energy range which is accessible at hadron colliders.

3.3.1 Current limits on anomalous couplings

In the past, studies of the $Z\gamma$ production have been performed by the CDF and DØ collaborations at the Tevatron collider [14,25], as well as the LEP [10] collider by the ALEPH, DELPHI, L3 and OPAL collaborations. These experiments have put limits on the anomalous couplings at the 95% confidence level (CL). A 95% CL is used when the data is consistent with the SM prediction. It puts a lower and an upper limit on the anomalous couplings, meaning that if the experiment was repeated 100 times, 95 of these would suggest values of the anomalous couplings that are within this interval. Anomalous couplings which lies beyond the range of the 95% CL are considered excluded.

It is common to only consider one coupling to be non-zero at the time. If more than one couplings is non-zero at the same time, cancellations may occur and the effect of the anomalous couplings is smaller [18].

The current limits from LEP and the Tevatron on anomalous triple gauge couplings contributing to the $ZZ\gamma$ vertex are summarized in Table 3.1 (obtained from [25]). The limits on h_1^Z and h_2^Z are not displayed. It is conventional only to give the limits on h_3^Z and h_4^Z since the other two have nearly the same limits, respectively. The limits on $h_{2,4}^Z$ are smaller than on $h_{1,3}^Z$ reflecting that the first two are dimension 8 whereas the latter are dimension 6. This makes $h_{2,4}^Z$ depend stronger on energy which provides better limits.

The limits from the LEP experiments can be set directly on h_i^Z, since the center of mass energy is fixed and therefore the use of form factors is redundant. For a comparison with the Tevatron limits, the LEP result should be translated into limits on h_{i0}^Z by multiplying with the conversion coefficient $(1 + s/\Lambda^2)^n$. Since this coefficient always is larger than 1 for all possible values of s, n and Λ, we see that the Tevatron provides the best limits.

3.4 Monte Carlo event generators

In order to see the observational effects of anomalous trilinear gauge couplings, we now turn to a discussion on Monte Carlo event generators. It seems convenient to finish this chapter

Experiment	95% CL
LEP	$-0.20 < h_3^Z < 0.07$, $-0.05 < h_4^Z < 0.12$
Tevatron	$|h_{30}^Z| < 0.033$, $|h_{40}^Z| < 0.0017$

Table 3.1: Limits on anomalous couplings at 95% confidence level from previous studies. For the Tevatron limits, a formfactor scale of $\Lambda = 1.5$TeV and $n = 3, 4$ for $h_{3,40}^2$ has been used. The Tevatron limits is obtained from an integrated luminosity of 1fb^{-1}. Notice that the limits from the LEP experiments can be set directly on h_i^Z.

by showing how to actually extract observables from theory.

Monte Carlo event generators allow us to study the properties of a given process by simulating events containing the particles of interest. This can be done in many different ways and thus there are numerous event generators available. Each one of them includes important aspects of di-boson production, but none of them cover *all* the important aspects. This means that we will have to combine some of the generators to get the results we are interested in.

In the present study, the event generator PYTHIA [26] has been used to produce $Z\gamma$ events. The effects of the anomalous couplings are included in a subsequent stage where these events have been reweighted by using another generator, called BHO [27], which includes some of the aspects that are missing in PYTHIA. More precisely, PYTHIA does not include the anomalous couplings that we are interested in, while BHO does. On the other hand, the BHO generator only simulates the so-called *hard process*, while neglecting the surrounding environment, including fragmentation and hadronization (to be defined below), all of which are crucial if the generator output is to be used in a detector simulation. PYTHIA, being a so-called general purpose generator, has all these features and is therefore ideal as a starting point.

PYTHIA uses leading order matrix element calculations, while the BHO generator is next-to leading order. At hadron colliders, the QCD next-to leading order contributions from gluon ISR and loops containing quarks and gluons (virtual corrections) are larger than the electroweak loop corrections mentioned in the beginning of this chapter (the "fermion triangle") and thus should be included before considering these. However, only the virtual part of the QCD corrections are included in this study, since the reweighting method will not create the additional particles which are introduced by gluon ISR. The electroweak corrections (including the "fermion triangle") are neglected.

Before embarking on the reweighting method mentioned above, we first look at the basic principles behind Monte Carlo generators. A schematic overview is shown in figure 3.3 and the individual subprocesses are listed below. The first point applies to both PYTHIA and BHO, while the rest applies to PYTHIA only.

Hard process: Monte Carlo generators starts by simulating the so-called the hard process. The hard process is based on a theoretical description using the cross-section of the process. For $Z\gamma$ we have

$$\mathrm{d}\sigma(q\bar{q} \to Z\gamma) = \frac{1}{2s}|\mathcal{M}(q\bar{q} \to Z\gamma)|^2 \frac{\mathrm{d}\cos\theta \mathrm{d}\phi}{8(2\pi)^2} \tag{3.13}$$

where $\mathcal{M}$ is the matrix element for the process, s is the center of mass energy squared and ϕ and η are the polar and azimuthal decay angles, respectively [28].

The first step in generating events is to sample the phase space. In general, the phase space is a multi-dimensional hypercube which spans all of the degrees of freedom. In our case, it is the two-dimensional space $-1 < \cos\theta < 1$, $0 < \phi < 2\pi$. The sampling is done by randomly choosing the $\cos\theta$, ϕ from a uniform distribution. This procedure is said to define the candidate events. Each of these are given an event weight, $d\sigma$, according (3.13) and thus the average of the candidate event weights will converge towards the total cross-section in the limit where the size of the sample goes to infinity.

However, at this point, there is no physical information in the distribution of the candidate events in phase space. To get a sample with the correct distribution, a method is applied which *unweigh* the events such that they all have a weight equal to 1, but are distributed according to the theoretical prediction (3.13).

This is usually done by using the so-called *hit-and-miss* technique. To apply this technique, the maximum event weight of the sample will have to be known (which can usually be approximated by randomly scanning the parameter space). Then, for each candidate event, the ratio of the event weight over the maximum event weight, $d\sigma/d\sigma_{\max}$, is compared to a number g generated from a uniform distribution in the interval $[0, 1]$. If the $d\sigma/d\sigma_{\max} > g$ the event is kept, otherwise it is rejected. In this way, a physical sample is generated in which the events are distributed according to (3.13).

In order to take into account that the partons entering the hard process each contains a certain amount of the proton's momentum, the parton distribution functions (pdf) are also considered in the hard process. The pdf's express the probability for the quarks and gluons in the proton to have the fraction x of the total proton energy evaluated at the energy scale Q of the collision. The pdf's cannot be determined theoretically, but have to be found experimentally[4]. Using this, the cross-section for $pp \to Z\gamma$ is given by

$$\sigma(pp \to Z\gamma) = \sum_q \iint dx_1 dx_2 f_q(x_1, Q^2) f_{\bar{q}}(x_2, Q^2) \sigma(q\bar{q} \to Z\gamma) \qquad (3.14)$$

where $\sum_q$ is the sum over all quark flavors.

Parton showers: When colliding protons with each other, the incoming partons in the protons will radiate photons and gluons before the hard process takes place. The same is the case for the outgoing particles. This is called initial and final state radiation (ISR/FSR), respectively, and is typically simulated through *parton showers.*

Ideally, we would like the radiation to be described through the matrix element as higher order contributions. However, this approach is not easy since the calculations involves the handling of divergences introduced by e.g. soft or collinear gluon emissions[5]. These complications are avoided by introducing parton showers instead of doing higher order calculations. Parton showers are based on the fact that in most processes, a factorization of the full process into individual subprocesses is possible. These subprocesses are then calculated separately, and the probabilistic branching between them are performed using Monte Carlo methods (such as the hit-and-miss technique mentioned above).

The advantage of parton showers is that since they are assumed to be independent of the hard process, they can be implemented for any process we wish to simulate.

---

[4]We will elaborate on this point in section 4.1.

[5]Many NLO generators exist which deal with these divergences. For instance, BHO is a NLO generator.

Hadronization and decays: Due to the phenomenon of colour confinement, no colored particles can exist long enough to be detected (see section 2.1.4). Hence, all quarks and gluons emerging from the hard process or the parton shower have to be merged into colorless objects. This process is called hadronization. In PYTHIA, it is done using the Lund string model [26] which is based on a picture with linear confinement. As two quarks move apart a gluon string is build up in the gap between them. As the distance between the quarks gets larger, the potential energy increases (linearly), and it becomes energetically favorable for a new quark and antiquark to emerge out of the vacuum. At this point, the gluon string breaks and each of the new quark and anti-quark pair up with the exiting quark and antiquark to form colour neutral mesons.

The whole framework is very constrained in terms of energy-momentum structure, but the flavor selection involves a multitude of parameters.

If unstable particles occur, the decays of these are simulated in the end.

Underlying event: The incoming protons are color neutral objects. When one of the partons of the proton interacts with a parton from another proton, the system of partons left behind is no longer color neutral. Since all objects in nature must be color neutral, the partons left behind will pick up partons from the vacuum to make them color neutral. This will result in additional jets of particles leaving the interaction point. The composite nature of protons also allows for several partons from each of the incoming protons to undergo scattering. The interactions are described by perturbation theory and are approximated by separated $2 \rightarrow 2$ scatterings which are colour-connected with each other and with the beam remnants. A p_T cutoff in the order of 2GeV is applied on the participating partons. For lower p_T, it is assumed that color screening in the incoming protons will strongly dampen the perturbative interaction rate.

3.5   Reweighting procedure

One way of including the anomalous triple gauge couplings in $Z\gamma$ events is to invoke a reweighting routine which rescales the cross section of SM $Z\gamma$ events to include the anomalous contributions. This procedure, of course, removes the notion of an individual event. The reason for using a reweighting routine as opposed to a Monte Carlo generator that includes all effects of the anomalous couplings is twofold. First of all, there *is* no suitable monte carlo generator that has this feature, as explained in the previous section. More importantly, even if there existed a Monte Carlo generator that had these properties, it would not be very attractive. The reason is that all relevant points in phase space would have to be simulated, i.e for each combination of values of the coupling constants, h_i^Z, a full Monte Carlo sample would have to be simulated, including the time consuming detector simulation. This is a highly inefficient approach. For comparison, the method of reweighting only requires one sample no matter how many different combinations of the anomalous couplings we wish to examine[6].

---

[6]There is, however, one disadvantage by using the reweighting method. As we shall see in chapter 6, the approach of reweighting SM $Z\gamma$ events makes the event selection a bit tricky, since here "signal events" are the ordinary SM $Z\gamma$ events and not the reweighted ones. Since the kinematics are different in the two cases, extra care must be applied when doing the event selection in order not to disfavor events with kinematics that are highly sensitive to the anomalous couplings. This should become more clear in chapter 6.
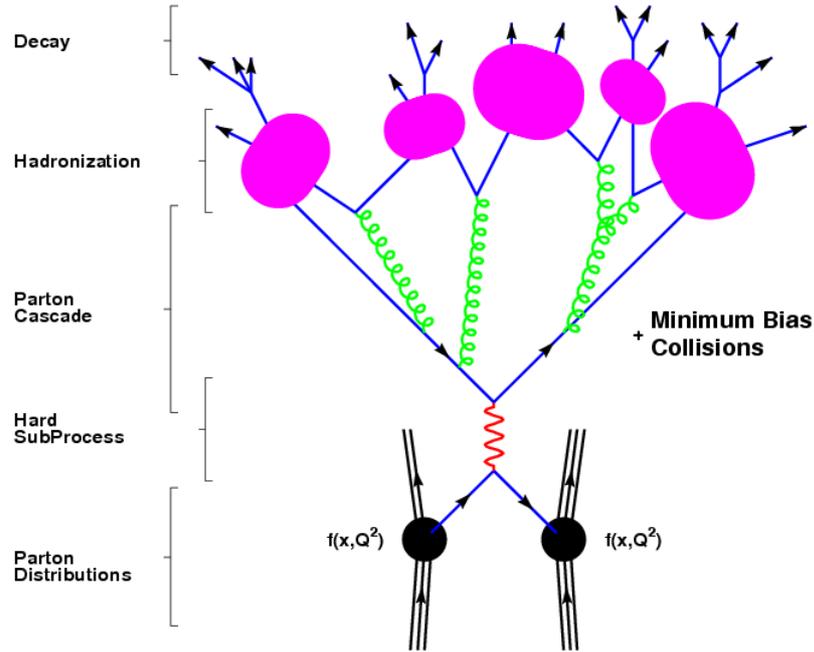
Figure 3.3: This figure shows the different subprocesses that occurs during event generation with PYTHIA. The figure is from [29].

The reweighting is done on the SM t- and u-channel diagrams of $q\bar{q} \to Z\gamma$. This procedure will rescale the SM cross section to include the contribution from the (anomalous) s-channel diagram (see figure 3.1).

In practise, this is done by first generating SM $Z\gamma$ events with Pythia and then using pieces of code extracted from the BHO generator to reweigh each event [30]. The reweigh program is given the relevant four vectors from Pythia (the incoming quarks, the photon and the Z decay products) and then calculates the differential cross-section (the helicity amplitude) for the process when the s-channel diagram is included.

In monte carlo event generators, the differential cross-section of a hard process is calculated using the matrix element, |M|, of the process. This differential cross-section, $d\sigma$, is also called the absolute weight of the event. When including anomalous triple gauge couplings in the Lagrangian, the cross-section will be affected through the matrix element. Since the anomalous coupling constants enter linearly in the Lagrangian and the cross-section depends on the matrix element squared, the inclusion of a single coupling will result in the following contributions to the absolute weight[7]

$$d\sigma = d\sigma_{\text{SM}} + h_i^Z \cdot F_1 + (h_i^Z)^2 \cdot F_2 \tag{3.15}$$

where $d\sigma_{\text{SM}}$ is the usual standard model contribution and h_i^Z is one the anomalous coupling constants ($i = 1, ..., 4$). The F_i's are the coefficients and consist of SM operator contributions as well as operator contributions associated with the anomalous couplings. For each event,

---

[7]One has to include the relevant terms in equation 3.6 to the Standard Model Lagrangian and square the result.

these coefficients are completely specified by the parton level kinematics of the particles and does not depend on the anomalous coupling constants. The linear term $h_i^Z F_1$ is the interference term between the Standard Model and the TGC amplitudes.

By considering more than one anomalous coupling constant at the same time, the absolute weight will also include cross terms which each contain a factor $h_i^Z \cdot h_j^Z$. The expression (3.15) then becomes

$$\mathrm{d}\sigma = \sum_{i,j=0}^{4} h_i^Z h_j^Z \cdot F_{ij} \tag{3.16}$$

where the SM contribution enters the sum via $h_0^Z = 1$ and $\mathrm{d}\sigma_{SM} = F_{00}$. By normalizing the weight with respect to the SM contribution it is possible to reproduce the effects of the anomalous couplings on the SM generated events by simply giving each event the weight

$$\mathrm{weight} = \frac{\sum_{i,j=0}^{4} h_i^Z h_j^Z \cdot F_{ij}}{\sigma_{\mathrm{SM}}} \tag{3.17}$$

3.6   Observables

The effects of anomalous couplings are found by studying the helicity amplitudes. Our investigation of the helicity amplitudes presented in section 3.3 does not reveal many features. However, we do see that the amplitudes increase with center of mass energy. Clearly, this will affect the energies of the Z and the photon and therefore effects will be seen in the distributions of the transverse energies of the Z and the photon. If the full helicity amplitudes are considered, more subtle features are revealed. We will not enter a detailed discussion on this matter, but merely mention that it can be shown that also the decay angles of the leptons are affected [18, 19, 22]. This will, however, not be pursued any further here.

In the present work, we have studied the effects of anomalous couplings on the p_T distribution of the photon. The method described in the previous section has been used to reweigh $Z\gamma$ events produced with PYTHIA. We have followed the convention described in section 3.3.1 and only included one non-SM value among the four couplings at the time. The histograms in figure 3.4 show the resulting effect when including $h_3^Z = 0.085$ and $h_3^Z = 0.015$ with $\Lambda = 2\mathrm{TeV}$. We see the expected behavior. The effect of the anomalous couplings is more prominent at higher transverse energies. Figure 3.5 shows the same as figure 3.4, but with a form factor scale of $\Lambda = 10\mathrm{TeV}$. When the form factor scale is raised, the strong suppression on the cross section comes at higher energies. This effect is seen by comparing figure 3.5 and 3.4.

The exact values for the h_3^Z coupling are chosen only to illustrate the expected behavior. Comparing them to table 3.1, we see that they are already excluded.

It is also possible to illustrate the effects of the anomalous couplings when they are varied in smaller steps than what is displayed in the histograms in figures 3.4 and 3.5. Figure 3.6 shows how the transverse energy of the photon is distributed as function of the coupling constant h_4^Z. We see a clear parabola shaped behavior in each p_T-bin as the coupling is varied, as we expect from the reweighting function (3.17). The ratio between the number of SM events and the number of events in the reweighted sample increases with energy. The sample used in this plot has not been normalized to $300\mathrm{pb}^{-1}$.
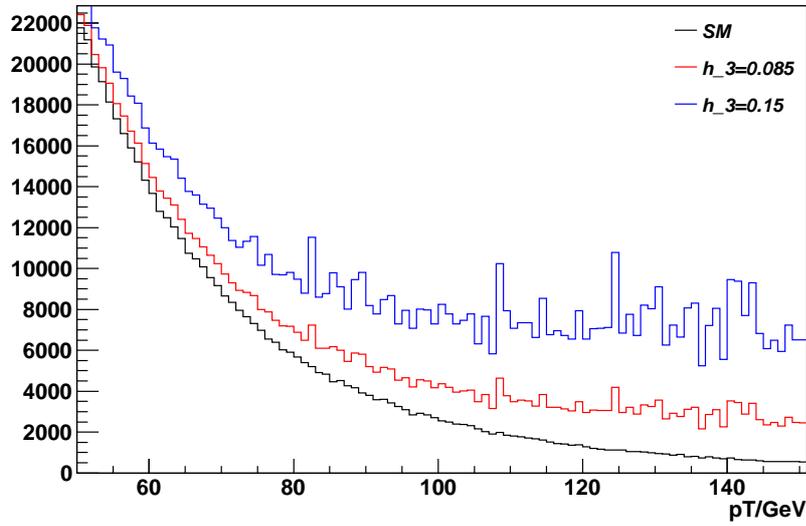
Figure 3.4: This figure shows the distributions of transverse momentum for the photon for the SM and when including $h_3^Z = 0.085$ and $h_3^Z = 0.15$ with a scale of $\Lambda = 2$TeV. The vertical axis shows the number of events out of a sample of 0.6M events.
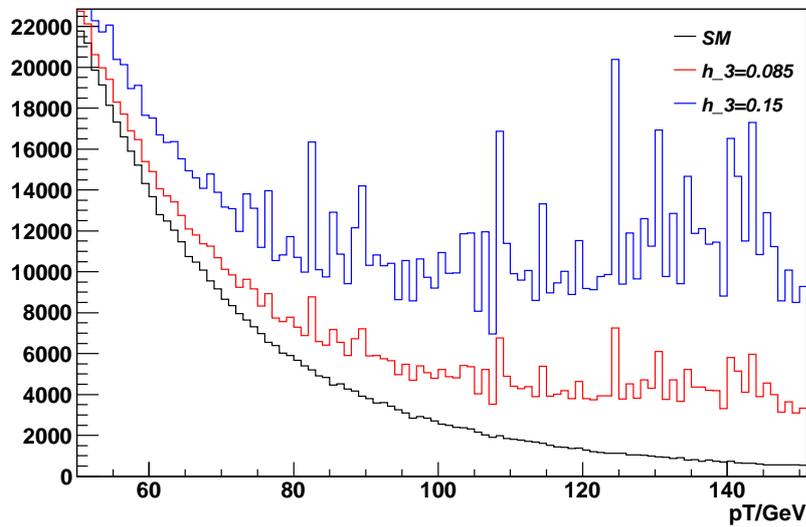
Figure 3.5: This figure shows the distributions of transverse momentum for the photon for the SM and when including $h_3^Z = 0.085$ and $h_3^Z = 0.15$ with a scale of $\Lambda = 10$TeV. The vertical axis shows the number of events out of a sample of 0.6M events.
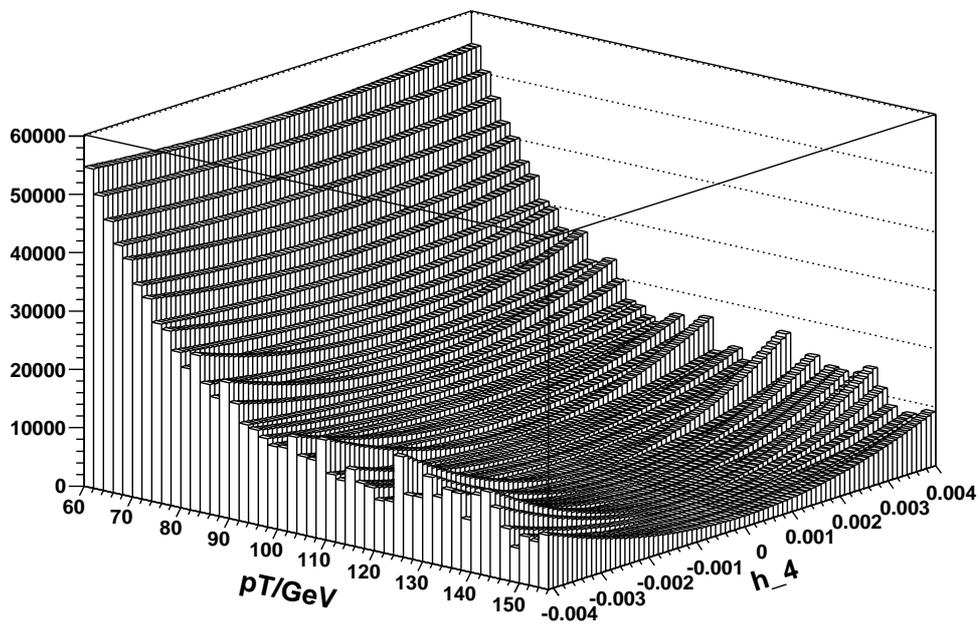
Figure 3.6: This figure shows how the transverse energy of the photon is distributed as function of the coupling constant h_4^Z. The vertical axis shows the number of events.

Experiment

Chapter 4

The Large Hadron Collider

The Large Hadron Collider (LHC) is the world's largest particle accelerator. It is located at the European centre for particle physics, CERN[1], near Geneva, Switzerland.

The LHC is designed to accelerate and collide protons or heavy ions (Pb). The accelerator ring has a circumference of 27km on which four experiments are situated, ATLAS, ALICE, CMS and LHCb. ATLAS and CMS are multipurpose detectors and their designs make them able to discover any new physics that appears at the high energies at LHC. LHCb is designed to study CP symmetry violation in the b-quark sector and ALICE will study quark gluon plasma (this will mainly be done in special runs with Pb-Pb collisions). Figure 4.1 shows an overview of the accelerator including some of the earlier stages of accelerating the proton beams.

First collisions at an injection energy of 450GeV per beam was recorded on the 23th of November 2009. In December, the beam energy was pushed to 1.18 TeV providing the highest center of mass energy collisions achieved by any collider so far.

The data taking will begin in February 2010 at a center of mass energy of 7TeV. The luminosity will be lower by 4-5 orders of magnitude than the design specifications since the bunch spacing will be much larger and there will be fewer protons in each bunch. At design luminosity the proton bunches in each beam are separated by 25ns (corresponding to approximately 7.5m) and contain up to 10^{11} protons. It is expected that the bunch spacing will go from 450ns to 50ns during 2010. The bunches will contain from $3 \cdot 10^{11}$ to $10 \cdot 10^{11}$ protons. The resulting peak luminosity will be $8.6 \cdot 10^{29} \mathrm{cm}^{-2} \mathrm{s}^{-1}$ in February and is expected to slowly increase to $2.1 \cdot 10^{32} \mathrm{cm}^{-2} \mathrm{s}^{-1}$ in November 2010. This will amount to an integrated luminosity in 2010 of approximately 300pb^{-1}.

At design specifications, it is expected that LHC will collide bunches of protons at a rate of 40 million times per second generating collisions with a center of mass energy of 14 TeV and a peak luminosity of $10^{34} \mathrm{cm}^{-2} \mathrm{s}^{-1}$. The integrated luminosity in one year of running is then expected to be 100fb^{-1}.

It is clear, that the upgrade from the starting conditions in February 2010 to the design specifications will take more than a year.

Besides from the increase in luminosity during 2010, it is also expected, that the center of mass energy will be increased to 4-5 TeV per beam in June or July. All the numbers presented here are preliminary[2].

---

[1]The abbreviation comes from the french name "Centre Européen pour la Recherche Nucléaire"

[2]They can be found on http://lhc-commissioning.web.cern.ch/lhc-commissioning/luminosity/09-10-lumi-
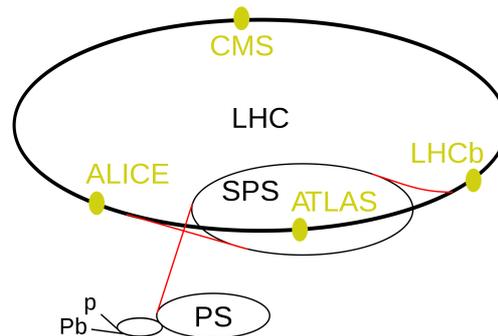
Figure 4.1: This figure shows the layout of the accelerator stages at the LHC along with the positions of the four experiments. The Proton Synchrotron Booster (not marked in the picture) is the first in the accelerator chain and takes 50MeV protons from the linear accelerator, Linac2, and accelerates them up to an energy of 1.4 GeV. From here, the protons are passed on to the Proton Synchrotron (PS) which will increase the energy to 26GeV before injecting the protons into the Super Proton Synchrotron. This is the last stage before the LHC ring and the protons will be accelerated to 450GeV before entering the LHC where the protons will be accelerated to a staggering energy of 7TeV. The figure is borrowed from http://en.wikipedia.org/wiki/Large_Hadron_Collider.

4.1   Collider Physics

The LHC will predominantly collide protons with protons, in contrast to its predecessor, LEP, which was an electron positron-collider. The differences between such two colliders are crucial. The reason for this lies in the fact that electrons and positrons are elementary particles, whereas protons are composite systems.

When colliding two leptons, the energy entering the hard collision will be exactly the invariant mass of the colliding leptons. This makes lepton colliders well suited for precision experiments where a very small energy range is considered at the time.

In hadron colliders, the energy in the hard collision is not that of the incoming protons, but rather of the participating proton constituents. These constituents are called partons. For the proton, they will be gluons, valence quarks and sea quarks. At low energies, proton consists of three quarks, two "up" and one "down". However, when dealing with energies such as those at the LHC, this simple description is insufficient. This has to do with the fact that we are probing well inside the cloud of virtual particles (see section 2.1.4) that surrounds each of the three quarks from the low energy approximation. These quarks will constantly emit gluons which subsequently can split into a quark anti-quark pair. At low energies, the chance of seeing these virtual particles is vanishing, while at higher energies they will become active participants. The three quarks from the low energy approximation are called valence quarks, while the virtual quarks are called sea quarks. The sea quarks are not restricted to be of the first generation only.

The fraction of momentum carried by a parton is denoted x and is described by a so-called parton distribution function (pdf). Given that the process occurs at a certain momentum squared, Q^2, a pdf gives the probability that a given parton inside the proton carries the mo-
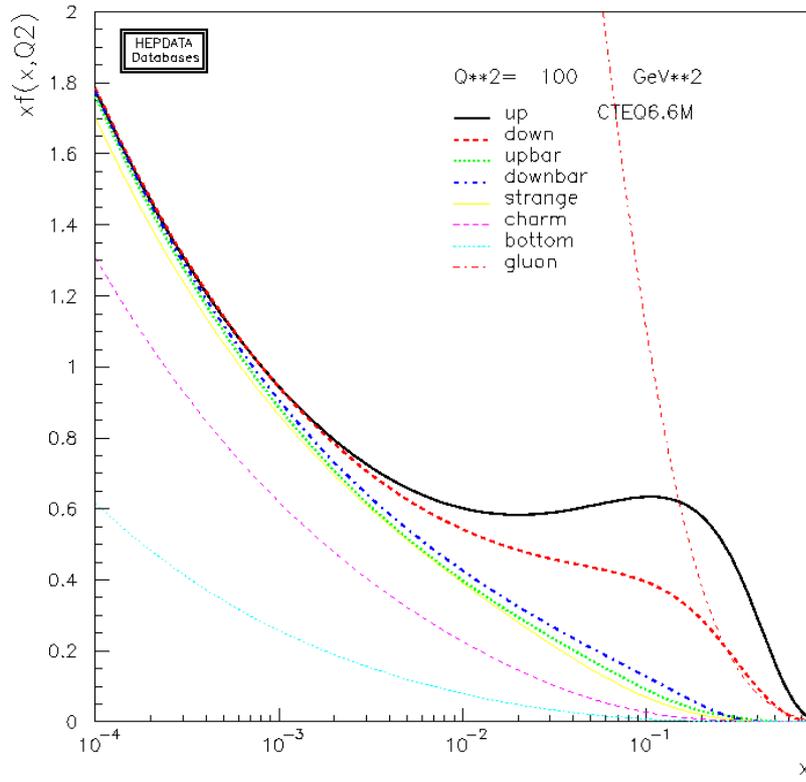
Figure 4.2: This figure shows the CTEQ6.6M parton distribution functions at $Q = 100$GeV. The figure was generated by using an on-line plotting tool available at http://durpdg.dur.ac.uk/hepdata/pdf3.html.

mentum fraction x of the total momentum of the proton. The momentum Q is the momentum that is exchanged in the collision and is also called the scale of the probing. This momentum determines the shapes of the pdf's. As an example, figure 4.2 shows the pdf's at a momentum scale of 100GeV. The peak of the valence quarks at roughly $1/3$ of the momentum is seen. This will be even more clear if we consider lower values of Q^2. The values for low momentum fractions x show the importance of radiative processes quarks and gluons undergo. If the momentum scale is increased the radiative processes become even more important which will enhance all the distributions at lower x and deplete them at higher x.

Because of the non-perturbative nature of QCD at low energies, the pdf's cannot be calculated, but have to be measured from known processes. The pdf's provide important information for Monte Carlo generators, and therefore collaborations, such as CTEQ, frequently publish a fit to existing data.

The energy range available for the hard collision in hadron colliders is quite broad. It is also clear that the partons involved in hadron collisions are able to produce a wider spectrum of particles than at lepton colliders, since the partons carry additional charges. For these

reasons, hadron colliders are often referred to as *discovery* machines. This kind of collider fits the present needs in high energy particle physics, since we have no specific predictions, such as a Higgs mass or SUSY masses, to make a confirmation of. The LHC will explore the energy spectrum from the Z mass at 90GeV and up to approximately 1-2 TeV for new particles and interactions.

Chapter 5

The ATLAS detector

The ATLAS detector (A Toroidal LHC ApparatuS) is a multipurpose detector intended to investigate many different types of physics that might become detectable at the high energies achievable at the LHC. Some will be confirmations or improvement of measurements of the Standard Model, while others are searches for new physics.

The ATLAS detector is a large and very complex detector consisting of many subdetectors, each designed for a specific purpose, which combined enable a full reconstruction of the particles forming an event. This chapter gives an overview of the ATLAS detector. Each of the subsystems will be described separately with emphasis on the parts most relevant for the present study. First, however, a brief summary of the major physics goals of the ATLAS experiment is given.

5.1   Physics goals

One of the most important goals of ATLAS is to investigate a missing piece of the Standard Model, the Higgs boson. If the Higgs boson is not discovered by ATLAS, it is expected that another mechanism of electroweak symmetry breaking that explains the same phenomena, such as Technicolor, will be discovered.

The asymmetry between the behavior of matter and antimatter, known as CP violation, will also be investigated. Current CP-violation experiments, such as BaBar and Belle, have not yet detected sufficient CP violation in the Standard Model to explain the lack of detectable antimatter in the universe. It is possible that new models of physics will introduce additional CP violation, shedding light on this problem; these models might either be detected directly by the production of new particles, or indirectly by measurements made of the properties of B-mesons. LHCb is likely to be better suited to the latter.

The top quark, discovered at Fermilab in 1995, has thus far had its properties measured only approximately. With much greater energy and greater collision rates, LHC will produce a tremendous amount of top quarks, allowing ATLAS to make much more precise measurements of its mass and interactions with other particles. These measurements will provide indirect information on the details of the Standard Model, perhaps revealing inconsistencies that point to new physics. Similar precision measurements will be made of other known particles; for example, ATLAS may eventually measure the mass of the W boson twice as accurately as has previously been achieved.

Other exciting lines of investigation are those directly involved with new physics. One

theory that is the subject of much current research is broken supersymmetry. The theory is popular because it could potentially solve a number of problems in theoretical physics. Models of supersymmetry involve new, highly massive particles; in many cases these decay into stable heavy particles that are very unlikely to interact with ordinary matter. These would escape the detector, leaving a large amount of "missing" momentum.

Other hypothetical massive particles, like those in Kaluza-Klein theory, might also leave a similar signature.

5.2 Coordinate system

Before describing each of the subdetectors it is worth taking a look a the coordinate system used to describe the ATLAS detector. The z axis is defined along the beam direction and the $x - y$ plane is the transverse plane with the x axis pointing from the interaction point towards the center of the LHC ring. The polar angle θ is the angle from the z axis, and the azimuthal angle ϕ is defined as the angle measured in the transverse plane from the top of the detector. It is usual to consider the pseudorapidity, defined as $\eta = -\ln(\tan\theta/2)$, instead of the polar angle, since the particles from collisions are roughly evenly distributed in this variable (referred to as the pseudorapidity plateau). In fact, the subdetectors are built with a constant cell size in pseudorapidity.

We will often use the "transverse" component of a variable, which is defined as the component of the given variable in the $x - y$ plane.

Finally, distances in the ATLAS detector are measured in the ϕ, η plane and we will often use the measure $\triangle R$.

5.3 The Magnet system

The ATLAS magnet system is divided into two separate systems, a solenoidal and a toroidal system (see figure 5.1). The solenoidal magnet system surrounds the inner detector and generates a magnetic field of approximately 2 Tesla parallel to the beam axis. This allows the measurement of the momentum of charged particles in the inner detector from the curvature of their tracks. The field also prevents very low energetic particles from reaching the calorimeters thus reducing the "noise". Choosing a solenoid magnet system in this part of the detector ensures a small field value in the calorimeters which is important for measuring the shower shapes in the calorimeters.

The toroidal magnet system consists of a barrel toroid and two end-cap toroids (see figure 5.1). The end-cap toroids consists of 8 coils which are housed inside a common cryostat and are inserted at the ends of the solenoid and inside the barrel toroid. The barrel toroid consists of 8 coils arranged radially around the hadronic calorimeter. The toroidal magnet system has a field strength of approximately 4 Tesla which provide the bending power necessary to measure high energetic muons up to about 1TeV.

5.4 The Inner Detector

The main purpose of the inner detector is to provide high precision tracking of charged particles. It consists of 3 subsystems, the pixel detector, the silicon strip detector (SCT) and the transition radiation tracker (TRT). All of them consist of a barrel part and two end-cap parts.
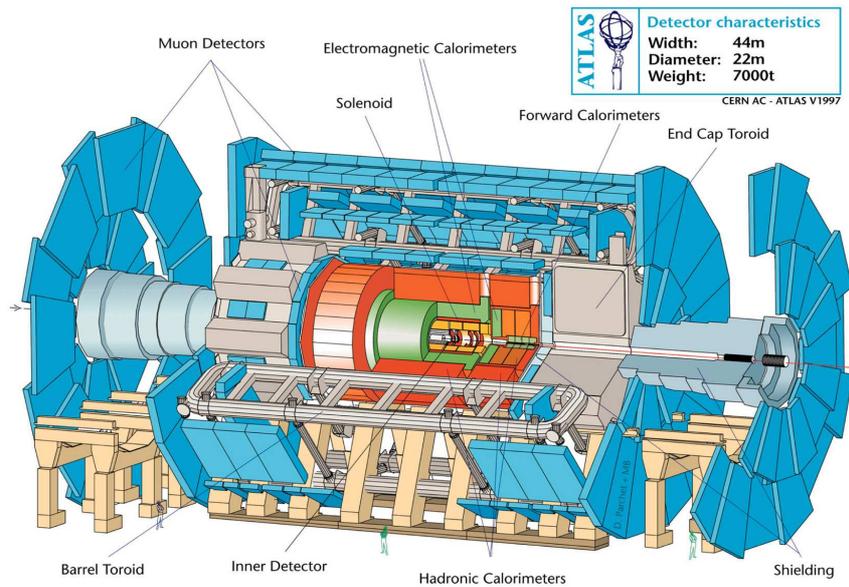
Figure 5.1: This figure shows a cutaway of the ATLAS detector. The figure can be found at http://atlas.ch

Combining the particle hits in each of these subdetectors allows a reconstruction of the paths of the charged particles and the determination of the interaction vertex. The arrangement of the subdetectors can be seen in figure 5.2. The inner detector is important for electron identification, thus, we will spend time on some of the details.

5.4.1 The pixel detector

The innermost part of the inner detector is the pixel detector. It consists of 3 concentric layers of semiconductor pixel sensors at radii of 5cm, 9cm and 12cm, where each gives a two-dimensional coordinate. They have an η coverage up to $|\eta| = 1.05$. The first layer is called the B-layer since is important for detecting secondary vertices used in B-tagging. In each end region there are 3 disks placed at radii between 9cm and 15cm which extends the η coverage up to $\eta = 1.7$. The detecting material is 250μm thick and the pixels have sizes down to $50{\times}400\mu$m closest to the interaction point.

5.4.2 The Semi-Conductor Tracker

The Semi-Conductor Tracker (SCT) is the middle component of the inner detector. It is similar in concept and function to the pixel detector, but with long, narrow strips rather than small pixels. It is designed to make 8 precision measurements per track on average. The silicon sensors are 6.3cm×6.4cm and are arranged in 4 double sided layers in the barrel at radii 29.9, 37.3, 44.3 and 51.4cm. Each side of the sensors is subdivided into strips of 80μm which are rotated slightly with respect to each other in order to get the z-component of the track. Ideally, the strips on each side of a layer would be perpendicular to each other (since this gives the best resolution in z), but due to lack of space for the readout cables they are
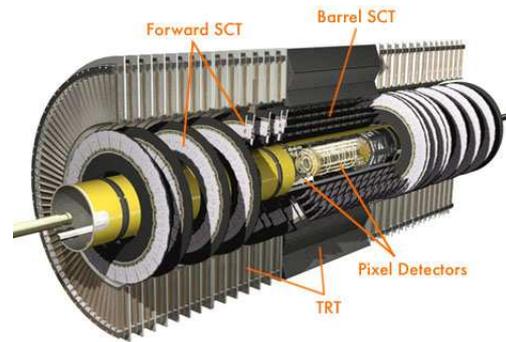
Figure 5.2: This figure shows a cutaway of the ATLAS inner detector. The figure can be found at http://atlas.ch

only slightly rotated. The barrel of the SCT extends to $|\eta| = 1.4$. Together with the end-caps the SCT covers the region $|\eta| < 2.5$.

5.4.3 The Transition Radiation Tracker

The last of the 3 subsystems in the inner detector is the Transition Radiation Tracker (TRT). This subdetector enables an almost continuous tracking in a large part of the inner detector at the expense of having lower resolution than the two semiconductor trackers. Another important feature is its ability to recognize electrons through transition radiation.

The detecting elements are drift tubes (straws), each 2mm in radius and up to 144cm long, which are filled with a mixture of gasses consisting of Xenon (70%), CO_2 (30%) and a small amount of O_2. In the centre of the tubes, a thin wire is suspended and held at a potential of approximately 1.5kV. When charged particle cross a straw, the atoms of the gas inside will be ionized. The liberated electrons from this ionization will drift towards the wire in the center while the leftover ions will drift towards the wall of the tube. The drifting electrons will cause an avalanche of electrons by secondary ionization of the gas along their way. This will generate a puls which can be read out. The straws with signal will provide a pattern in the TRT, and by measuring the different drift times in the straws, it is possible to reconstruct a track with a resolution up to 150μm.

Besides from measuring the track of a charged particle, the TRT also uses the effect of transition radiation to identify electrons. Between the straws, materials with different dielectric constants cause ultra-relativistic charged particles (with a γ-factor of $\sim$1000 or more) to produce transition radiation and leave a much stronger signal in some of the straws. The emitted photons are in the X-ray region and since Xenon is particularly good at absorbing such photons, they will cause a lot of extra ionizations providing the stronger signal. An electron will cause transition radiation at energies as low as 0.5GeV and since the electrons coming from a collision and passing through the straws typically will be much more energetic than this, it is expected that they will trigger 8 so-called High Threshold hits (HT) on average. The hits are defined by the height of the generated pulse in the wire. In HT hits, lots of transition radiation have been generated and thus the generated electrical pulse is high, while low threshold hits (HL) are hits where little transition radiation has been generated resulting in a small electrical pulse. The HT hits are a special feature of electrons. For comparison, a
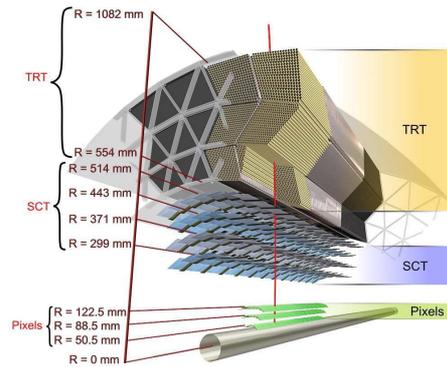
Figure 5.3: This figure shows a close up of the three subsystems in the inner detector. The figure can be found at http://atlas.ch

charged pion with an energy of 140GeV will not generate much transition radiation, and thus it is possible to identify electrons based on their signature in the TRT.

The TRT consists of a barrel part and two end-caps. The straws in the barrel are placed horizontally and cover the region up to $|\eta| = 0.8$ while the straws in the end-caps are placed radially and cover the region $0.8 < |\eta| < 2.5$. The modules in the barrel extend radially between $r = 55$cm and $r = 108$cm while the end-caps consists of 3 disks at $z = 49.5$, $z = 58.0$ and $z = 65.0$cm. In all, there are about 370.000 straws in the TRT. A particle transverses 20-36 straws on average.

Another view of the TRT along with the 2 other subsystems of the inner detector is shown in figure 5.3.

5.5 Calorimeters

The calorimeters are situated outside the solenoidal magnet system that surrounds the inner detector. There are two basic calorimeter systems: an inner electromagnetic calorimeter and an outer hadronic calorimeter. The purpose of both of them is the measure the energy from particles by absorbing them. This is done by stopping the particles in layers of high density metal and measure the shape and energy deposit of the resulting particle shower. If the particle is completely absorbed in the calorimeter, it is possible to give an estimate of its original energy. An overview of the calorimeters is seen in figure 5.4.

5.5.1 EM calorimeter

The electromagnetic calorimeter (ECal) measures the energy of particles that interact electro-magnetically, which includes electrons and photons. Hadrons will typically not loose much energy here, because they do not interact strongly with the material used in the ECal. It consists of a barrel part covering $|\eta| < 1.475$ and two end-cap parts each covering $1.375 < |\eta| < 2.5$. The energy-absorbing materials are lead, with liquid argon (LAr) as the sampling material. These are placed in alternating layers in an accordion shape structure as shown in figure 5.5. The depth of the calorimeters are often measured in radiation lengths X_0, where X_0 is defined as the mean distance over which an electron has $1/e$ of its original energy left.
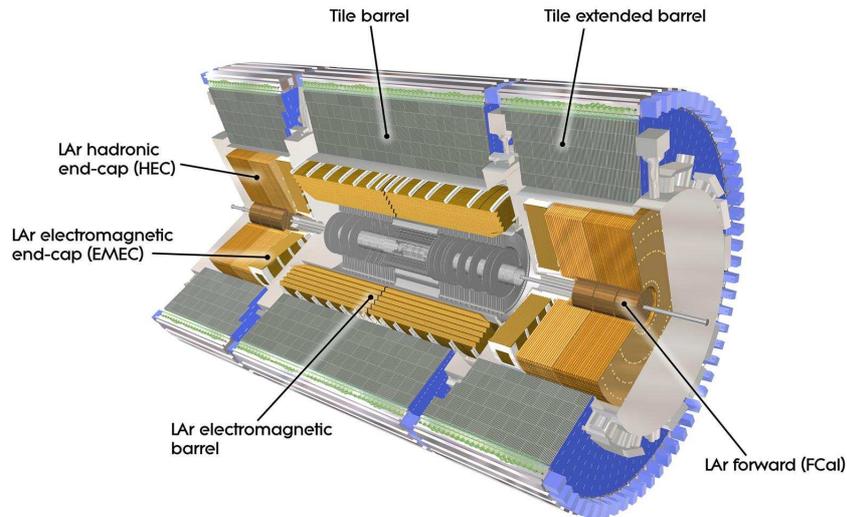
Figure 5.4: This figure shows the calorimeters in the ATLAS detector. The figure can be found at http://atlas.ch

The electromagnetic calorimeter is important for the present study since we wish to identify electrons, positrons and photons. Since the photon is only seen in the EM calorimeter, a precise measurement of its energy and shower shape is crucial to its identification.

The ECal consists of 3 layers in depth (and a presampler) with different dimensions of calorimeter cells. This structure allows for an precise determination of the lateral and longitudinal shape of a shower. The regions are called *samplings* and are described in the following list:

- *The presampler*: A certain loss of energy is expected before the transversing particle reaches the calorimeter due to the presence of material in front of it (the inner detector and the solenoid), and therefore a single layer of LAr has been put in front of the solenoid magnet. This layer is called the presampler, covering the region $|\eta| < 1.8$, and is used to correct the energy measurement in the ECal for this energy loss. The presampler is not shown in figure 5.5.

- *First sampling*: Since this is the first layer of the calorimeter that the particles encounters, it has the finest granularity in η in order to get a good resolution. This is needed for separating single photons from the two spatially close photons coming from neutral pion decays. The depth of the first sampling is $4.3X_0$. Since the charged particles will have curved tracks in the ϕ-direction, a fine granularity is not efficient in this direction.

- *Second sampling*: The second sampling is the largest of the three with a depth of $16X_0$. This is where most of the energy will be deposited. Electrons with an energy below 50GeV will be completely absorbed in this sampling.

- *Third sampling*: The purpose of the third sampling is mainly to measure the energy leakage from the second sampling generated by high energetic particles. For this reason,
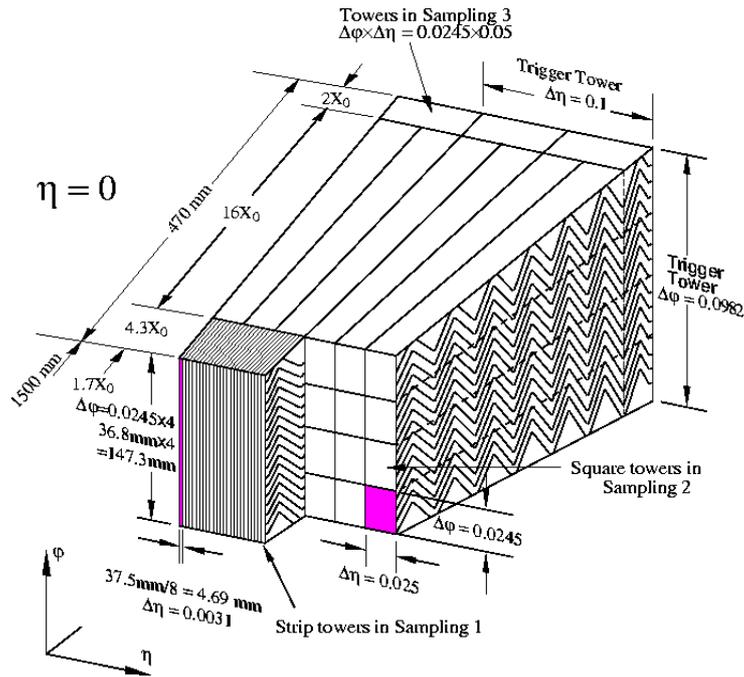
Figure 5.5: This figure shows the structure of the ECal. The individual parts are described in the text. The figure can be found at http://atlas.ch

the granularity of the third sampling is a bit worse than the others. The depth is $2X_0$.

5.5.2 Hadronic calorimeter

The hadronic calorimeter is used to measure the energy of hadrons. Precise measurements are important to the present study, since we want to be able to exclude events containing a lot of jets (these suggest that the event is not a $Z\gamma$ event). It will also help us to establish whether a given electron/positron/photon candidate from the EM calorimeter is isolated from other activity in the detector or if it perhaps is part of a QCD jet.

The hadronic calorimeter is divided into a barrel part covering $|\eta| < 1.0$, an extended barrel covering $0.8 < |\eta| < 1.7$ and two end-caps covering $1.7 < |\eta| < 3.2$. These parts are collective referred to as HCal. In addition, there is the forward calorimeters (FCal) which covers the region closest to the beam pipe, $3.1 < |\eta| < 4.9$ (see figure 5.4).

The barrel consists of alternating layers of plastic scintillator plates and iron absorbers. The scintillators are 3mm thick and the iron plates 14mm. The end-caps and the forward calorimeters are more exposed to radiation and thus cannot be constructed by the same materials as the barrel part. Instead, they are LAr based calorimeters as the ECal, however, using copper as the absorbing material (tungsten for the FCal).

The HCal has a depth of 11 hadronic interaction lengths which is sufficient to stop nearly all hadrons from punching through to the muon detector. The calorimeters covers almost the entire η-range all the way down to the beam pipe which is important in order to get a precise energy measurement of all particles that entered the interactions. Equally important, it constitutes a necessary tool in estimating the size and direction of missing energy in interactions

where neutrinos are involved.

5.6 Muon spectrometer

The last line of measurements in the ATLAS detector is done by the muon chambers. They consist of three layers separated by the toroidal magnet system which makes it possible to measure the momentum from the curvature of the tracks. At the energies relevant at the LHC, muons will penetrate the calorimeters without leaving significant amounts of energy— they are so-called "minimum ionizing" particles at these energies. No other particles have this feature, and thus hits in the muon chambers almost for sure means that a muon was involved in the process. However, very energetic mesons will also have a non-vanishing chance of punching through the hadronic calorimeter and leave a track in the muon chambers. Another background comes from mesons that decay into muons before leaving the detector.

The barrel part of the muon chambers are Monitoring Drift Chambers (MDT) which are aluminum tubes with a diameter of 3cm filled with a mixture of 93% Ar and 7% CO_2. When a charged particle passes, it will ionize the gas and electrons will drift towards a wire in the middle of the tube. The functional principle is thus very alike that of the TRT. Furthest out on each side, a MDT is situated (see figure 5.1).

The end-cap drift chambers are Cathode Strip Chambers (CSC). The change of technology is due to the intense radiation that detector parts closer to the beam pipe experience.

A third kind of muon chamber is also implemented. The so-called Resistive Plate Chambers (RPC) are positioned in the barrel region and Thin Gap Chambers in the end-cap regions. The fast response time from these modules make them ideal as trigger detectors.

The muon detectors cover the region $|\eta| < 2.7$.

5.7 Triggering

One of the major tasks at the ATLAS experiment is the selection of events to be recorded on disk. When operating at high luminosity with bunch crossings every 25th ns, 800.000.000 collisions will occur every second (~ 20 per event). However, recording full event data at this rate is far beyond the capabilities of the storage elements at ATLAS, which can only handle about 100 events per second. Most of the events will be soft collisions without interest for the physics we want to study. These events are known as minimum-bias events. To make use of the high event rate anyway, a complex trigger system has been developed which selects events that may be of interest during data taking.

The ATLAS trigger is organized in three trigger levels (LVL1, LVL2, LVL3), as shown in figure 5.6. The LVL1 trigger accepts data at the full LHC bunch-crossing rate of 40 MHz. The latency (time taken to form and distribute the LVL1 trigger decision) is about $2\mu s$, and the maximum output rate is limited to 100 kHz by the capabilities of the subdetector readout systems and the LVL2 trigger. The LVL1 trigger is implemented in hardware, since the response time needs to be very fast. Its decision is based on reduced granularity calorimetry information and the output of the RPC of the muon spectrometer. In the calorimeters, high p_T objects, such as electrons, photons, jets and τ's will trigger the LVL1 as well as large missing or total transverse energies. The information from the LVL1 trigger system is used to identify the regions-of-interest (RoI) of the detector, i.e. regions containing interesting features such
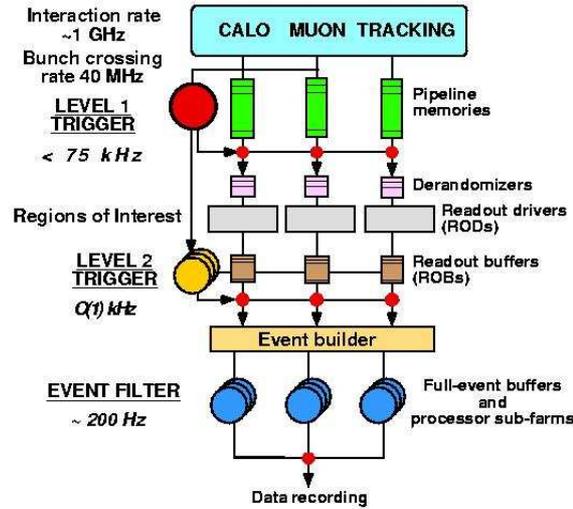
Figure 5.6: This figure shows a layout of the trigger system at the ATLAS experiment. The figure can be found at http://atlas.ch

as high-energy electromagnetic clusters (electrons or photons), jets and muons. The LVL1 will reduce the event rate to $\sim$ 100kHz.

The LVL2 trigger is software based. It has access to the full event information, but primarily investigates the RoI identified by LVL1. Normally, only a small fraction of the detector is selected as RoI with corresponding advantages in terms of the required processing power and data-movement capacity. For electrons the LVL2 searches for a track in the inner detector as well as particle identification from the TRT. For photons the inner detector information cannot be used due to the possible conversions of photons. The total LVL2 latency is variable, up to about 10 ms. The LVL2 trigger reduces the rate from up to 100 kHz after LVL1 to about 1 kHz.

After an event is accepted by the LVL2 trigger, the full data are sent to the LVL3 processors (also called the event filter). Complete event reconstruction is possible at LVL3, with decision times up to about 1 s. The LVL3 system must achieve a data-storage rate of 10–100 MB/s by reducing the event rate and/or the event size. For some triggers, for example Higgs boson candidates, the full event data will be recorded with an event size of up to about 6 MB, corresponding to a maximum event rate of about 100 Hz.

5.8   Electron and photon reconstruction

In order to perform a physics analysis, the detector output has to be processed and reconstructed into physical objects such as electrons and photons. The exact form of the reconstruction routines is an ongoing process constantly improving as the understanding of the detector increases.

Efficient electron and photon reconstruction will be a task of unprecedented difficulty at the LHC. There are several reasons for this, one being the ratios of inclusive electrons and photons to jets from QCD processes which are expected to be between one and two orders of magnitude
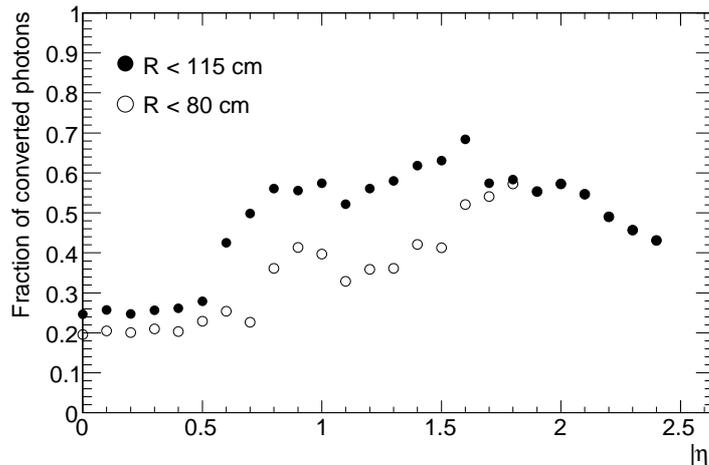
Figure 5.7: This figure shows the fraction of photons that convert at a radius below 80cm and 115cm as open and full circles, respectively. The figure is borrowed from [31].

worse than at the Tevatron. For instance, for electrons with $p_T = 40\text{GeV}$, the electron to jet ratio is expected to be $\sim 10^{-5}$ [31]. In addition, the large amount of material situated in front of the calorimeters provides not only a serious challenge to the energy measurement, but also adds the possibility of photon conversions.

Converted photons appear when photons that pass through material (e.g. the beam-pipe, the subdetectors, etc.) interact with the nuclei on their way. This will result in a momentum transfer between the two which effectively makes the photon gain mass (it goes off-shell). A massive photon has the possibility of decaying into an electron/positron pair if $m_{\gamma*} > 2m_e$ which means that a lot of photons are lost before reaching the calorimeter. In general, the effect is larger the more material the photon has to pass through, so the effect is more apparent at large η, see figure 5.7. We see that the effect is quite large and for that reason the electron and photon reconstruction algorithms have to deal with this issue.

There are two reconstruction algorithms for electrons and photons. The first starts by scanning the calorimeter cells for "seeds" around which to build a fixed sized cluster consisting of typically 50 calorimeter cells (the "sliding window" approach). Then a loose matching for a track is searched for among all reconstructed tracks in the inner detector. Additionally, it is checked whether the candidate matches a reconstructed photon conversion in the inner detector. Electrons and photon are thus separated by requiring the electrons to have an associated track, but no associated conversion. The photons are defined as having no matched track, or as having been matched to a reconstructed conversion.

The second reconstruction algorithm applies only to electrons and uses a good quality track from the inner detector as seed and constructs a cluster around an extrapolated impact point in the calorimeter.

The first algorithm is the standard choice in the reconstruction software in ATLAS. However, the second has shown efficiency improvements for low p_T electrons as well as for electrons close to jets.

The objects collected from the first reconstruction algorithm are collectively called egamma-
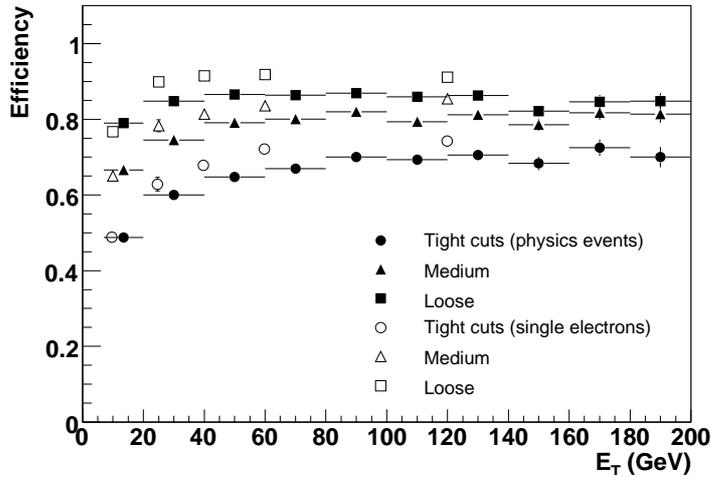
Figure 5.8: This figure shows the overall reconstruction and identification efficiency of various levels of electron cuts: loose, medium and tight isolations, as a function of the transverse energy for single electrons (open symbols) and for isolated electrons in a sample of physics events with a busy environment (full symbols). Figure borrowed from [31].

objects. Each object that is reconstructed as an egamma object carries a tag called IsEM, which is a bit flag telling how much the object "looks" like an electron or photon based on information from all subdetectors. There are three pre-defined IsEM bit patterns which can be used to make a tighter selection of electrons.

Electrons:

- *Loose cuts*: These cuts do not add many requirements to the electron candidates. It consists of simple shower-shape cuts (small longitudinal leakage into th hadronic calorimeter, matching of shower shape in second sampling of the EM calorimeter) and very loose matching between reconstructed track and calorimeter cluster.

- *Medium cuts*: In addition to the "loose" requirements, the fine granularity from the first sampling is used to discriminate candidates which have several entrance points in the calorimeter. These could be photons from a decayed π^0. The track is required to have pixel and SCT hits.

- *Tight cuts*: The track is required to have hits in the B-layer of the pixel detector in order to reject photon conversions, and must also contain a certain amount of HT hits in the TRT. The energy measurement from the calorimeter must match the momentum measurement from the inner detector and the cluster must be isolated from other clusters.

Clearly, tighter cuts lead to lower reconstruction efficiency. However, the jet rejection power also increases. For transverse energies above 17GeV the jet rejection efficiency is 570, 2200 and 10^5 for loose, medium and tight electrons, respectively. The electron reconstruction efficiencies are shown in figure 5.8.

For photons, there is only one predefined setting for the isEM bit-flag since measurements are provided only by the EM calorimeter. This setting represents cuts analogous to the "tight" cuts for electrons.

Photons:

- *Tight cuts:* The cuts on the shower shape are optimized with special emphasis on separating single π^0 from photons using the fine granularity in η of the first sampling of the EM calorimeter. In addition, a standard track-isolation criterion has been added to further improve the rejection while preserving the vast majority of converted photons.

Using these criteria, an efficiency of about 84% is expected for photons with an energy spectrum as expected from one of the discovery channels of a light Higgs, $H \to \gamma\gamma$, with $m_H = 120\text{GeV}$ [31]. This efficiency is quite uniform over the entire η-range and comes with an jet rejection efficiency of ~ 5000.

5.9 Detector simulation

In order to study the detector response for a wide range of physics processes and scenarios, a software package has been constructed that carries events from the event generation (described in section 3.4) through a detailed detector response simulation with an output format identical to that of the true detector. The simulation program is integrated into the ATLAS software framework, Athena.

The simulation software chain is generally divided into 3 parts. The first is the event generation which has been explained in section 3.4. The next is simulation of the detector and physics interactions. The last is the digitization of the energy deposits in the detector into voltages and currents for comparison to the readout of the ATLAS detector. The output of the simulation chain can be presented in an object-based format or in an format identical to the output of the ATLAS data acquisition system (DAQ). In this way, both simulated and real data from the detector can be run through the same trigger and reconstruction algorithms. For detailed descriptions of the ATLAS simulation project, see [32].

5.9.1 Full simulation

A full simulation refers to a simulation where all physics processes that can occur in the detector has been considered and where all details of the detector has been taken into account. The software package implemented in Athena that is responsible for this is Geant4 [33] which is often referred to as the core simulation.

Geant4 takes as input the events from the event generator. The generated particles are then propagated through the detector and the energy deposited in the detector-elements is saved as "hits". Geant4 considers all the physics processes that occur, such as photon conversions, ionizations and the implications of the magnetic fields which bends the path of the particles. All this relies on a detailed map of the entire detector where the different types of material used in the detector elements and their exact positions are taken into account.

The ATLAS digitization software converts the hits produced by the core simulation into detector responses: "digits". Typically, digits are produced when the voltage or current on a particular readout channel rises above some predefined threshold within a particular time-window. The peculiarities of each subdetector's response, such as electronic noise and channel

dependent variations, have been modelled in subdetector-specific digitization software. The general properties of the algorithms have been tuned to reproduce the detector response seen in lab tests, test beam data and cosmic ray running.

Effectively, the digitization algorithms transform the hits produced by the core simulation into Raw Data Objects (RDO).

The process up to this point models the data acquisition phase under real experimental conditions. From here on the simulated data is treated as if it was real data and handled by the reconstruction algorithms presented in the previous section.

5.9.2   AtlfastII

The complicated detector geometry and detailed physics description used in the ATLAS Geant4 simulation make it impossible to achieve the required statistics needed for many physics analysis, including the present one. For this reason, various "fast simulations" have been developed to complement the standard Geant4 simulation. None of these are meant to replace the Geant4 simulation, but can work as a supplement. In the present study, time did not allow for a full simulation, and thus the fast simulation AtlfastII was used. Detailed descriptions of this and other fast simulations can be found in [32].

All fast simulations make approximations in order to reduce time consumption. When considering the time spent in the different parts of a full simulation, it turns out that simulating the particles as they transverse the calorimeters take up almost 80% of the time [32]. Hence, a fast simulation of the calorimeters will provide a substantial time reduction.

The aim of AtlfastII is to simulate events as fast as possible while still being able to run the standard ATLAS reconstruction. AtlfastII uses a special parametrized algorithm (FastCaloSim) to simulate the calorimeters while pertaining Geant4 simulation of the inner detector and the muon chambers. More recently, AtlfastII has been been further supplemented with a fast simulation of the inner detector and the muon chambers called Fatras (Fast ATLAS Tracking Simulation). The current standard configuration includes Fatras, but optionally, any subdetector can be simulated with Geant4 to provide the higher level of accuracy without the same computing time consumption as full simulation of the entire detector. In this study, the older version of AtlfastII which did not include Fatras was used. Running in this mode, the time consumption is reduced by a factor of 10 for most physics processes. Including Fatras will reduce the time consumption even further, reaching a factor of 100 with respect to the full simulation [32]. Below is a brief description of the algorithm FastCaloSim used in AtlfastII to simulate the calorimeters.

FastCaloSim: Instead of simulating the particle interaction in detail, the energy of a single particle shower is deposited by FastCaloSim directly using parametrizations of the longitudinal and lateral energy profile. The distribution of active and inactive material in the calorimeter needs to be respected by the parametrization, so a fine binning of the parametrization in the particle energy and pseudorapidity is used. Furthermore, the energy deposition depends strongly on the origin of the shower in the calorimeter, and therefore this has been taken into account as well.

The parametrizations are based on a 30 million event sample of fully simulated (i.e. using Geant4) single photons and charged pions in an energy range between 200MeV and 500GeV, evenly distributed in $|\eta| < 5.0$ and $-\pi < \phi < \pi$ [32]. All electron and

photon showers are approximated by the photon parametrization, while hadronic showers are approximated by the charged pion parametrization.

Detailed studies of the physics performance of FastCaloSim has been carried out [32], and it has been found that FastCaloSim differs at the 5% level from full simulation after reconstruction. Especially, the properties that are sensitive to the shape of hadronic showers are affected. Since all particles are simulated using an average lateral shape function, visible effects such electromagnetic sub-showers in charged pion showers are not described by FastCaloSim. For electrons, the difference between FastCaloSim and full simulation can be seen in electron identification efficiencies for which the two differs by about 5% [32].

The performance of AtlasfastII was also studied in [34]. Here, the conclusion was that although AtlfastII is an improvement with respect to the older Atlfast[1] (which was the first fast simulation of the ATLAS detector) there are still noticeable discrepancies.

Figures 5.9 and 5.10 show distributions of the reconstructed objects from $Z\gamma$ events obtained from both full simulation, Atlfast and AtlfastII. The differences between the three simulations are quite clear, for instance we see that the efficiency of Atlfast is much higher than both full simulation and AtlfastII due the fact that Atlfast reconstructs all particles.

---

[1] Atlfast is extremely fast, several thousand times faster than full simulation, but does very coarse simulation of the physics processes. It performs a "smearing" of momenta and energies of the particles from the event generator output using parametrization functions obtained from fitting to full simulations. This also means that no particles are lost in the process. Details on Atlfast can be found in [32]

Invariant mass of Z candidat
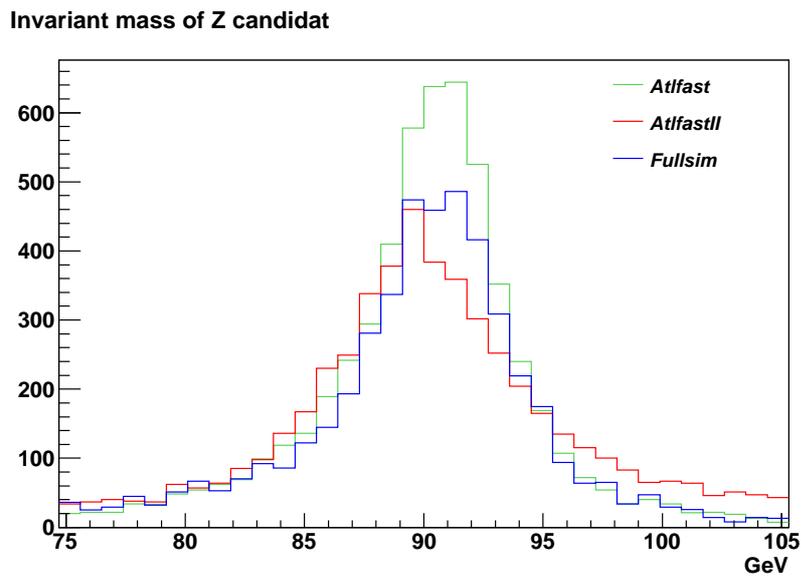
Figure 5.9: This figure shows the distribution of events as function of the invariant mass in GeV of the reconstructed Z in $Z\gamma$ events for full simulation, Atlfast and AtlfastII. The Z is reconstructed from egamma objects recognized as electrons and positrons with $p_T > 10$GeV using the e^+e^- candidate whose invariant mass is closest to 91.187GeV. As indicated, the selection criteria are very loose which means that the sample is somewhat contaminated. Nonetheless, the plot provides a test of the two fast simulations. The simulated events are $Z\gamma$ events produced with PYTHIA. Figure taken from [34].

Electron detection efficiency
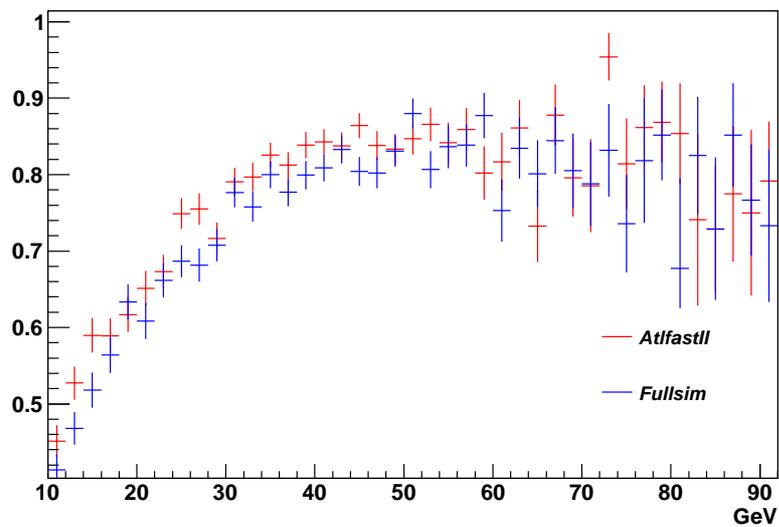
Figure 5.10: This figure shows the reconstruction efficiencies as function of transverse energy of egamma objects recognized as electrons with the same selection criteria as mentioned in figure 5.9. The efficiencies are calculated by matching the reconstructed electrons to the event generator output. As indicated, the selection criteria are very loose (in fact, looser than "loose") which means that not only electrons enter the plot, but also e.g. pions which constitute the main contamination. Thus it cannot be compared directly to figure 5.8. Nonetheless, it provides a test of AtlfastII since the objects exposed to the two simulations are the same. The simulated events are $Z\gamma$ events produced with PYTHIA. Figure taken from [34].

Analysis

Chapter 6

Event selection

An event selection represents a certain way of retrieving a (usually very small) signal from a (usually very large) background. With "background" we mean events that resemble the signal events in one or more ways, but are not part of the signal. As previously discussed, "signal" events are SM $pp \to Z\gamma$ events where the Z decays into an electron/positron pair. In this thesis, the effects of anomalous couplings will be included by reweighting the SM $Z\gamma$ events.

In the ATLAS experiment, most of the events produced will be of no interest to this thesis or to most other studies for that matter. As discussed in section 5.7, the first selection of events happens on-line in the trigger system which only lets a small fraction of the produced events pass for further analysis.

Since the cross-section of SM $pp \to Z\gamma$ is relatively low (as we will see), these events will only constitute a small fraction of the events that are triggered on and therefore will have to be selected by an offline selection. For this purpose, advanced techniques of optimizing the selection can be put into practice. Recently, the method of Boosted Decision Trees (BDT) has been used in high energy particle physics [16, 35], and this will also be the main selection method presented here.

However, it should be emphasized that a complex selection method such as BDT should be used with care. First of all, it should be preceded by a pre-selection to make sure a series of minimum requirements are met. In our case, a pre-selection should select 1 electron, 1 positron and 1 photon from each event and reject events where this is not possible. Preferably, the pre-selection should also make a severe cut on background events while maintaining as large a portion of the signal as possible and hence bring the signal/background ratio closer to unity. When this is achieved, the more advanced selection method can be applied to remove some of the background events that more closely resemble the signal.

Since this method is less transparent due to its complexity, it should be used in a simple way in order to have at least some control of its output. This means that only a few, but "good" variables should be used in the selection. Variables on which the anomalous triple gauge couplings in some way depend should be avoided if possible. The reason for this is that the selection is done on SM $Z\gamma$ events—hence, if the BDT classifier somehow disfavors signal events which later on will receive a relatively large weight in the reweighting procedure, the overall analysis will be less sensitive to the anomalous couplings. A variable such as the invariant mass of the $e^-e^+\gamma$ is a typical example. As explained in chapter 3, the weight increases with center of mass energy, i.e. with invariant mass of $e^-e^+\gamma$. Thus, if the BDT at some point finds it suitable to cut away events with high invariant mass of $e^-e^+\gamma$, the overall

analysis will be less sensitive to the anomalous couplings.

When two backgrounds have significantly different features, it can be hard to train one BDT to remove both at the same time. In the present study, this turned out to be the case.

Initially, the idea was to tune the pre-selection to get rid of the most obvious background source to $Z\gamma$ which, as we shall see, is radiative Z events which have final state radiation (FSR) of a photon from one of the Z's decay products (electron/positron). With this background reduced by the pre-selection, only one other large background is left ("fakes" from $t\bar{t}$) which the BDT was to take care of[1].

However, due to the low amount of data considered, it turned out to be difficult to increase the signal to background ratio while maintaining a significant amount of signal events (more than ~ 10). Thus, a slightly different approach was considered. The cuts in the pre-selection was made looser and an extra BDT was included to compensate for this. This first BDT was then trained to remove the FSR background which previously was handled by the pre-selection, while the "fakes" were discriminated in the second BDT. This approach turned out to be superior to the one with a more tight pre-selection and only one BDT.

6.1 Backgrounds to $Z\gamma$ events

This thesis will concentrate on the process $pp \rightarrow Z\gamma$ with the Z decaying into an electron/positron pair. The final state particles $e^-e^+\gamma$ are objects that are well reconstructed in ATLAS. As described in chapter 5 the four momentum of these objects can be fully reconstructed from ID and calorimeter information, although the photon resolution is a bit worse since it leaves no tracks in the ID and thus only relies on the calorimeter information. The knowledge of the lepton four momenta will enable a reconstruction of the Z, which then in principle, if background events can be excluded, makes it possible to uniquely identify $Z\gamma$ events in ATLAS.

Roughly, the background can be put into two categories: Events that actually contain the final state $e^-e^+\gamma$ and get reconstructed as such; and events which only have some or none of the final state particles $e^-e^+\gamma$, but where other particles in the event are (falsely) reconstructed as this final state.

The last category is hard to get a handle on. In fact, this will be one of the major challenges to all LHC experiments, due to the presence of the large number of additional soft pp collisions, known as minimum bias, that occur simultaneously with any hard interaction in high energy collisions.

Here follows a list of the major part of the background important for this thesis.

- Radiative Z: Since the production rate for $q\bar{q} \rightarrow Z$ is much higher than $q\bar{q} \rightarrow Z\gamma$ (see table 6.2), this process will have a large contribution to the background, due to the possibility of final state radiation (FSR) of a photon from one of the Z's decay products, the electron or positron. The possibility of a jet faking the photon also constitutes a large background. The fake photons could be neutral mesons (π^0, η) which decay into two spatially close photons, or charged mesons or electrons missing tracks. They typically originate from underlying event remnants.

- $t\bar{t}$: Another large background comes from $t\bar{t}$ events, with $t \rightarrow Wb \rightarrow e\bar{\nu}_e b$ (and likewise for the $\bar{t}$) and the photon produced either by FSR of the electron/positron or a jet

---

[1]Both of these backgrounds will be explained in more detail below.

faking the photon. The branching ratio of $t \to Wb$ with $W \to e^- \bar{\nu}_e$ is $\sim 9\%$ [36]. The production rate of $t\bar{t}$ is approximately a factor 30 larger than for $Z\gamma$.

- ZW, ZZ, W^+W^-, $W\gamma$: These other diboson events can also have the three final state particles or at least some of them while faking the remaining. However, taking into account the cross-sections and branching ratios (and the mistag probabilities in case of fakes), they all become negligible. None of these backgrounds have been included in this study.

- $Z\gamma \to \tau^+\tau^-\gamma$: Since we only consider $Z\gamma$ events where the Z decays into an electron/positron pair, the tau channel will constitute a background through $\tau^- \to e^- \bar{\nu}_e \nu_\tau$ and likewise for τ^+. The branching ratio for taus to electrons is 17.8% and thus the contribution from this background is more than 20 times smaller than the signal. When reconstructed, it will not either look like a signal event since the neutrinos carry away much of the energy that should have gone into the reconstruction of the Z. This background has not been included in the present study.

Besides these backgrounds, the so-called *dijets* may also be a serious background. Dijets refer to $2 \to 2$ QCD processes where the final state particles from the hard process are quarks or gluons. As explained in section 3.4, these quarks and gluons will "hadronize" and create several composite particles which are reconstructed as jets in the detector. Dijets can be labeled according to the transverse momentum of the two partons exiting the hard process as shown in table 6.1. Low p_T dijets have huge cross-sections at hadron colliders. This makes it practically impossible to simulate a representable amount of these events no matter which detector simulation is put into use.

Dijet label	p_T range (GeV)	σ (pb)
J8	$p_T > 2240$	$5 \cdot 10^{-6}$
J7	$1120 < p_T < 2240$	0.12
J6	$560 < p_T < 1120$	281
J5	$280 < p_T < 560$	$1.810 \cdot 10^3$
J4	$140 < p_T < 280$	$64.8 \cdot 10^3$
J3	$70 < p_T < 140$	$1.6 \cdot 10^6$
J2	$35 < p_T < 70$	$30.5 \cdot 10^9$
J1	$17 < p_T < 35$	$519 \cdot 10^9$

Table 6.1: This table shows the dijet cross-sections in different p_T ranges of the jets. All cross-sections are found using PYTHIA.

In general, the dijets in the higher p_T-range will more easily fake a $Z\gamma$ event than dijets in the lower p_T-range, since more of the particles that are produced in the jets will have significant sizes of transverse momentum and therefore pass the basic selection cuts (which will be explained in section 6.3). Fortunately, we see from table 6.1 that the cross-sections are relatively low for the most energetic dijet events. It is clear that J8 cannot contribute considerably to the background due to the low cross-section. However, the dijets in some of the remaining p_T range could virtually ruin the selection in worst case scenarios. This can be realized by comparing the dijet cross-sections to the ones showed in table 6.2.

In the present study, we simulated a small sample consisting of a few hundred thousands of dijet events of the types: J7-J4[2]. It was checked that none of these made it completely through the selection. This is, of course, not sufficient to conclude that these dijets are rejected by the selection, but it does, however, serve as a nice check that the selection is on the right track. It is clear, that even if only a few of these dijet events had survived the selection, the dijets, when scaled to the correct luminosity, would completely dominate the dataset.

6.2 Dataset

The simulated dataset for this analysis will represent the amount of data that is expected to be collected within the first year of running at the LHC (year 2010). This means that for both signal and background a number of events corresponding to 300pb[1] at a beam energy of 3,5 TeV is produced[3]. Taking the relatively small cross-section of the signal process into account, this will not amount to a lot of events. In order to have a well determined prediction, substantially more events have been produced (which will, of course, be rescaled to the correct luminosity). These extra events are needed in order to get small statistical errors and to get a proper training of the BDT. Table 6.2 summarizes the dataset used.

	Dataset	CSC-ID	Events (sim)	$\sigma(pb)$	$\sigma_{\mathrm{eff}}(pb)$	Events $(300\mathrm{pb}^{-1})$
	$Z\gamma \to e^+e^-\gamma$	5899	0.6M	2.47	1.56	468
	$Z \to e^+e^-$	5144	2.680M	$0.733 \cdot 10^3$	$0.666 \cdot 10^3$	199860
	$t\bar{t}$	5568	1.285M	81.7	81.6	24470
Dijets	J7	5016	0.1M	0.12	0.12	36
	J6	5015	0.1M	281	281	84300
	J5	5014	0.1M	$1.810 \cdot 10^3$	$1.810 \cdot 10^3$	543000
	J4	5013	0.1M	$64.8 \cdot 10^3$	$64.8 \cdot 10^3$	19.4M

Table 6.2: This table shows the dataset for the event selection and analysis. Indicated are the number of simulated events, the cross-sections, the effective cross-sections after applying event filter in the Mont Carlo generator and the number of events corresponding to a luminosity of 300pb^{-1}. The CSC-ID refers to the official ATLAS software joboptions used in the Monte Carlo generator (PYTHIA) which has been used to retrieve the cross-sections. They are all leading order (LO).

The samples are simulated using the ATLAS software default values in the event filter of the MC generator. This means that for signal events, the event filter consists of a p_T-cut of 10GeV and a η-cut of 2.7 on the photon and the electrons/positrons, while the Z is required to have a mass of at least 60GeV. The same applies to radiative Z events, except for the photon filter. For $t\bar{t}$ events, the event filter is a cut on the transverse momentum of the scattering process of 18GeV. The cuts on transverse energy mentioned here are all chosen a bit conservatively, at least when comparing to other studies, e.g. [22] in which the p_T cut of the

---

[2]Preferably, the entire p_T range would have been sampled, however, due to lack of time, this could not be achieved.

[3]These numbers are, of course, preliminary. It has been discussed to take a step up in energy from 3,5TeV to 5TeV per beam during the summer 2010. The luminosity will slowly increase during 2010. An integrated luminosity of 300pb^{-1} at 3,5TeV per beam in 11 months of data taking in 2010 seems reasonable. All numbers are taken from internal cern talks at the 27$^{\mathrm{th}}$ LHC Machine Committee meeting in September 2009.

photon in the signal sample is set to 100GeV. However, the study presented in [22] considers the prospects of measuring the anomalous couplings at the LHC design specifications in one year of collecting data, i.e. $\sqrt{s} = 14$TeV and $\int \mathcal{L} = 100$fb^{-1}. Only a few signal events have a photon with $p_T > 100$ GeV. Therefore, we choose a more conservative generator cut.

6.3 Pre-selection

As previously mentioned, the relatively low amount of data at the given luminosity and center-of-mass energy constitutes a challenge to the selection and analysis of the present study. In other analysis in the ATLAS collaboration, e.g. [22], the dataset typically corresponds to a luminosity of 100fb^{-1} and with a center-of-mass collision energy of 14TeV. At this energy, the cross-sections are enhanced. The cross-section of $Z\gamma$ is $\sigma = 2.9$pb (using PYTHIA), which means that approximately 3M $Z\gamma$ event will be produced at $\mathcal{L} = 100$fb^{-1}. This allows a much tighter selection. In [22] they consist of strong cuts on transverse energy of the photon ($p_T > 100$GeV), the separation of the photon from nearby electrons/positrons ($\triangle R_{\gamma l} > 0.7$), a jet-veto that rejects events which contain jets with $p_T > 30$GeV. These cuts reduce the background significantly. The amount of signal events, however, are also strongly reduced, leaving less than 0.1% of the events for further analysis [22].

Clearly, this procedure will not be applicable to the present analysis, since we only have 468 $Z\gamma$ event to begin with. Therefore, the pre-selection presented here will consist of looser cuts than what is normally applied. The BDT may then be able to improve the signal/background ratio while maintaining a sizeable amount of signal events.

The pre-selection consists of several layers. First, electrons, positrons and photons are selected. The criteria for electrons and positrons are that they are egamma objects that pass an IsEM bit flag corresponding to "loose" and have a transverse momentum larger than 20 GeV. It is also required that there should be one of each so that it is possible to reconstruct a Z candidate. An electron/positron candidate is labeled "electron" if the curvature of the track in the inner detector is that of a negatively charged particle (and vice versa for a "positron"). The choice of which electron/positron pair to use for the Z candidate can be made it different ways. The photon is required to be an egamma object with an transverse momentum above 20 GeV and no associated track.

Then all possible combinations of the electron, positron and photon candidates are made to construct $Z\gamma$ candidates. In order to remove the most improbable candidates, we make two additional cuts. We demand that the invariant mass of the Z candidates are in the range $70 < m_Z < 110$ GeV, and make a cut on the $Z\gamma$ invariant mass, $m_{Z\gamma} > 95$ GeV. Though these cuts are quite conservative, they will get rid of some of the candidates with the photon being final state radiation from Z events, since these in general will have lower invariant masses of both the Z and the $Z\gamma$. In fact, the invariant mass of the $Z\gamma$ is expected to be centered around 91.187GeV in the case of FSR in radiative Z events.

In most cases ($\sim 97\%$ for signal events), we only have one $Z\gamma$ candidate. In the remaining cases ($\sim 3\%$ for signal events), we have more than one candidate and thus we have to make a set of criteria that identifies the "best" candidate.

The usual practise is to select the Z candidate whose invariant mass is closest to 91.187GeV which is the peak of the Breit Wigner distribution for the Z boson, and pair it with the photon with highest p_T. This choice seems very reasonable, since we expect the reconstructed Z mass to be close to the peak of the Breit Wigner, and since the photon from the hard process is

likely to have higher p_T than a photon from further down the decay chain, e.g. a FSR photon.

To test this way of selecting, we have also considered some other possibilities which we quantified by introducing the following weights:

- Weight according to the probability of the photon candidate being a FSR photon (w1): This weight represents a measure of how close the photon is to a given reconstructed electron/positron measured in η, ϕ space with the distance given by $\triangle R = \sqrt{(\triangle\eta)^2 + (\triangle\phi)^2}$. If the photon is close to an electron or positron, it is likely to be FSR. We define w1 $\propto \triangle R_{Min}$, where $\triangle R_{Min}$ is the distance to the closest electron/positron. The weights are normalized such that $\sum$ w1 $= 1$, where the sum goes over all photon candidates in the event.

- Weight according to the probability of the photon candidate being a "fake" (w2): If there are a lot of particle tracks in the vicinity of the reconstructed photon (in η, ϕ-space), it may actually be part of a jet. However, simply counting the number of tracks may not be a good discriminator, since in many cases, there will be hundreds of low p_T tracks distributed over the entire inner detector. What we are really interested in is "how much" p_T these tracks have. Thus, we take the sum of p_T of all tracks in a cone of $\triangle R < 0.45$ centered on the photon candidate and define the ratio: w2 $\propto p_T^\gamma/(p_T^\gamma + \sum p_T^{\text{tracksC45}})$. The weights were normalized so that $\sum$ w2 $= 1$.

- Weight according to transverse momentum of photon (w3): This weight represents the usual way of selecting the photon. w3 $\propto p_T^\gamma$ with $\sum$ w3 $= 1$.

- Weight according to the probability of the electron/positron candidate being a "fake" (w4): Defined the same way as w2. w4 $\propto p_T^{e^\pm}/(p_T^{e^\pm} + \sum p_T^{\text{tracksC45}})$. The weights were normalized so that $\sum$ w4 $= 1$ where the sum goes over all electron/positron candidates in the event.

- Weight according to the distance of the invariant mass of the Z candidate to 91.187 GeV (w5): This weight is defined by w5 $\propto |m_Z^{\text{reco}} - 91.187|$ with $\sum$ w5 $= 1$ where the sum goes over all Z candidates in the event.

It is now possible to investigate different combinations of these weights. In each case, we compare to "truth", i.e. the MC generator output, to see how often we choose the true electron, positron and photon in the reconstructed signal events[4].

The final weight is then composed by using one or more of the weights listed. The final weight is constructed by multiplying the individual weights (in the case of w1 and w5, we multiply by 1-w1 and 1-w5), and the candidate with the highest weight is chosen. If, for instance, we want to include the first, the second and the third weight, the final weight will be: $w(1,2,3) = (1 - w1) \cdot w2 \cdot w3$. 6 different combinations of weights were tried. w5 is included in all combinations, since this weight provides the best check of the Z candidates. The combinations are listed in table 6.3.

---

[4]The matching from truth to reconstructed data was done by trying to match each reconstructed egamma object to the true electron, positron and photon and make an identification if these were within $\triangle R = \sqrt{(\triangle\eta)^2 + (\triangle\phi)^2} < 0.15$ and $\triangle p_T < 15$GeV. It was checked that at most one matching was possible for each of the 3 truth particles.

Combination	w(1,2,3,4,5)	w(1,2,4,5)	w(1,2,5)	w(3,5)	w(1,2,3,5)	w(2,3,4,5)
Purity (%)	54	53	53	57	54	57

Table 6.3: This table shows the performance of 6 selection methods in signal events in which more than one reconstructed $Z\gamma$ candidate was found. The first row shows the combination of weights in the different methods (with, hopefully, an obvious notation). The second line shows how often the true candidate was chosen, given by the "purity" defined as the no. of times the true candidate was chosen divided by the sum of events in which the true or a false candidate was chosen. The errors on all purities displayed in the table are $\sim 2-3\%$.

We see that the performance of the different selection methods does not vary much. It seems, that simply choosing the photon with the highest p_T is preferable, so this will be the method we use[5].

We also checked the truth matching when applied to all signal events (i.e. with one or more $Z\gamma$ candidates). When doing a fairly loose pre-selection, it seems vital to do this check, since remnants from the underlying event or pile-up have greater chance of making the cuts and thus contamine the signal sample. Actually, it was more or less this check that defined the preselection. The pre-selection cuts earlier mentioned result in a purity of 80.6%. This could not be tuned any higher even if the pre-selection was made relatively tight.

We will comment on selection efficiencies for signal and background in a later section when we also have the results for the main selection method, the boosted decision trees.

6.4 Boosted Decision Trees

The pre-selection was a so called rectangular cut optimization method. In such an optimization, a cut on one variable does not depend on the value of another variable, nor can a, say, background-like value of one variable be counterweighted by signal-like values of the other variables.

The opposite is true when dealing with so-called multivariate analyzers (MVAs), such as boosted decision trees (BDT). BDT combines the technique of boosting with the well-known selection method of decision trees. A decision tree is a way of classifying events by successively dividing the sample by cutting on one variable at the time until a stop criterion is reached. They are in this respect similar to rectangular cuts. However, whereas a cut-based analysis is able to select only *one* hypercube as region in phase space, the decision tree is able to split the phase space into a large number of hypercubes, each of which is identified as either "signal-like" or "background-like". This can be visualized by a simple two-dimensional tree structure as shown in figure 6.1.

In order to build (train) the decision tree, we have to select an estimator to optimize in each node so that a possible split can be determined. For this thesis an analysis package called Toolkit for MultiVariate Analysis (TMVA) is used. The estimators available for optimization in TMVA are

- *Gini Index*, defined by $p \cdot (1-p)$

---

[5]From looking at the number in the table, we could have chosen the last method equally well. In fact, taking the errors into account they are all comparable. However, the other methods are more complicated and thus opens up the possibility of introducing additional systematic errors.
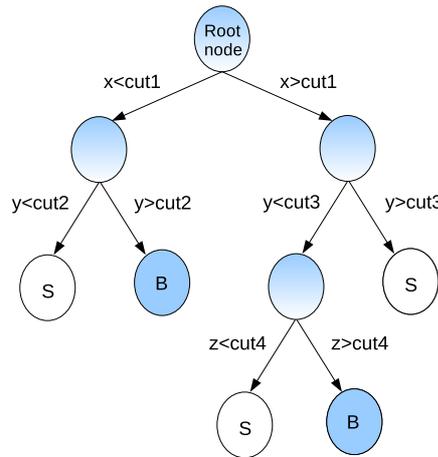
Figure 6.1: This figure shows the basic principle behind decision trees. Starting from the root node a sequence of binary splits is performed. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may be used several times. The leaf nodes at the bottom of the tree are labeled "signal" or "background" depending on the majority of events that end up in the respective nodes.

- *Cross Entropy*, defined by $-p \cdot \ln(p) - (1-p) \cdot \ln(1-p)$

- *Misclassification error*, defined by $1 - \max(p, 1-p)$

- *Statistical significance*, defined by $S/\sqrt{S+B}$

where S and B are the numbers of selected signal and background events, respectively, and p is the purity, defined as $p = S/(S+B)$.

All separation criteria have a maximum where the samples are fully mixed, i.e. $p = 0.5$, and fall off to zero when the sample consists of only signal or background. A cut that selects predominantly background is as valuable as one that selects signal. This is reflected in the fact that all estimators are symmetric with respect to the two event classes. The four separation criteria have no significant performance disparity [37]. In our tests, we could confirm this for our samples and thus we used the default setting which is the Gini Index.

6.4.1   Boosting

A shortcoming with decision trees is their instability with respect to statistical fluctuations in the training sample from which the tree structure is derived. Consider, for instance, the case where two variables exhibit similar separation power at a given node. A small fluctuation in the training sample may cause the tree growing algorithm to decide to use one of the variables, say, the first, at this node. However, the same algorithm would perhaps have chosen the second variable if the fluctuation had been different. The consequences of such a situation is that the whole tree structure below the node would be altered, possibly resulting in a substantially different classifier response.

Ideally, we would like to choose both variables and somehow take the average of the resulting classifier response. Then, we would be less sensitive to statistical fluctuations.

This can be achieved by constructing a forest of decision trees and classifying an event according to a majority vote of the classifications done by each tree in the forest. Since all trees are derived from the same training sample, it is clear, that if the sample is not somehow altered between growing each tree, they will all be exactly the same and nothing is gained.

This situation is avoided by subjecting the events to a so-called "boosting" routine between growing each decision tree. The boosting alters the sample by modifying the event weights.

In TMVA there are several types of boosting, the default one is called AdaBoost (adaptive boosting). It represents a way of taking into account the shape of the distribution of the variables in the regions where the cut is applied. In AdaBoost, the events that are misclassified in one tree are given a higher weight in the following tree. This means that the condition for determining which variable to cut on and where to cut is changed. The misclassified events will, in general, have a higher probability of getting the correct classification in the following tree.

The weight is the same for all misclassified events in a tree and is derived from the misclassification rate, err, of the previous tree

$$\alpha = \frac{1 - \text{err}}{\text{err}} \tag{6.1}$$

The entire event sample is then renormalized to keep the total number of events (i.e. the sum of weights) in the trees constant. We notice that the weight α can be seen as a quantity belonging to the individual *tree* rather than the misclassified events themselves.

The classification from a single tree for an event $\mathbf{x}$ is given by $h(\mathbf{x}) = \pm 1$ corresponding to either signal (+1) or background (-1). The total classifier response $y_{\text{BDT}}(\mathbf{x})$ is then a weighted sum of all the classifications from the individual trees. In TMVA, the default way to do this average is

$$y_{\text{BDT}}(\mathbf{x}) = \frac{1}{N} \sum_{i}^{N} \ln(\alpha_i) \cdot h_i(\mathbf{x}) \tag{6.2}$$

where N is the number of trees in the forest. A large value of $y_{\text{BDT}}(\mathbf{x})$ indicates that the event is signal-like, while a small value suggests that the event is background-like. It is also possible to define the final classification (6.2) without including the weights or by using the purity ($p = S/(S + B)$) in each final node in the decision trees to define the weights instead of relying on the binary decision. However, this last procedure is potentially sensitive to overtraining and neither does show any significant performance increase [37]. Therefore, we used the procedure described in (6.2) to train our classifier.

6.4.2   Training and testing

The training or *growing* of a decision tree is the process that defines the exact structure of the tree, i.e. the splitting criteria for each node.

In order to test the performance of the BDT, the event sample is divided into two equal parts; a training and a test sample. The training is then done by applying the techniques described in the previous sections to the training sample. The performance is tested by applying the trained classifier on the test sample. Preferably, the classifier response should be (approximately) the same for the two samples.

We intend to train two BDT classifiers (the first one to recognize FSR and the second to recognize "fakes"). We use radiative Z events (CSC-ID 5144) for background in the first classifier training and $t\bar{t}$ events (CSC-ID 5568) for the second.

In the final selection, they will be applied in sequence, and thus we train and test the first classifier until it is optimal, before proceeding with the second, because we will need to apply the first classifier on the samples to be used in the second classifier training in order to optimize this classifier.

6.4.3 Variables

The variables that is used in a selection should be variables that well describe the different topologies of the signal and the background. It is not always straight forward to pick these variables and some will be better than others. In general, BDTs are not very sensitive to poorly chosen variables, since these will not be used very often by the decision algorithm if they have negligible separation power. However, the variables used should really come out of a physics consideration. By this, we mean that all variables used should somehow be justified and not just used because they are available[6]. Poorly chosen variables could result in an inefficient classifier.

The variables chosen for this selection are:

1. m_Z: The invariant mass of the reconstructed Z. Since the two selected leptons should be the decay products of a Z boson, the invariant mass of the two leptons is expected to be a good variable. The radiative Z events are expected to have a lower reconstructed Z mass, since the photon carries away some of the energy. The events where the leptons are not coming from a Z (mostly $t\bar{t}$ events) will not have a well defined peak around 91.187GeV. This variable is used in both classifiers.

2. $\triangle R(\gamma, e^{\pm})$: The distance from each photon candidate to the nearest electron/positron candidate. The radiative Z background events will differ from signal events because the direction of the FSR photons will be close to the leptons from which they radiated as oppose to the ISR photons which are more evenly distributed. This variable is also important for discriminating $t\bar{t}$ events, since if the photon originates from a jet it will seldom come alone. The jet may also contain electrons (both real and fakes) in which case the photon will be close to the electron. If the electron is real it may radiate and then the photon would be FSR.

3. $p_T^{(\gamma, e^{\pm})}/(p_T^{(\gamma, e^{\pm})} + \sum p_T^{\text{tracksC45}})$: To have the possibility of discriminating jets reconstructed as electrons, positrons or photons, the sum of transverse momentum of all tracks in a cone of 0.45 around the electron, positron and photon are used as a variable. This is the same variable as we used in the pre-selection (see 6.3). The argument for using it here is the same. In signal events we expect the three reconstructed particles to be fairly isolated. In $t\bar{t}$ events all three particles can be faked in a jet which typically leaves multiple tracks. In radiative Z events only $p_T^{\gamma}/(p_T^{\gamma} + \sum p_T^{\text{tracksC45}})$ is considered, since the electron and positron predominantly will come from the Z. The photon be-

---

[6]To give an example, there are approximately 100 variables related to the measurements of an electron or photon in a standard Ntuple. In principle, we could use all of them, but this is not desireable, since we would loose any feeling we may have for how the cuts are made by the decision algorithm.

ing FSR, however, will be accompanied by an electron or positron which leaves a track within the cone.

4. $E_{\text{EM calo}}/E_{\text{All calos}}$: The ratio of energy measured in the electromagnetic calorimeter vs all calorimeters. This variables can be used to discriminate events with jets, since our signal does not leave significant amounts of energy in the hadronic calorimeter, while jets from a $t\bar{t}$ event leave lots of energy here.

5. $m_{Z\gamma}/(m_{Z\gamma}+m_{jets})$: The ratio of the invariant mass of the $Z\gamma$ candidate vs the invariant mass of the $Z\gamma$ candidate plus the sum of the invariant masses of all jets with $p_T > 30$GeV. Our signal events are expected to be very clean (i.e. no significant jets), whereas $t\bar{t}$ events typically will have some jets.

6. p_T^γ/p_T^Z: The ratio of the transverse energy of the photon candidate and the Z candidate is expected to be unity for signal events, since the incoming colliding quarks have very low transverse momentum. For both types of background this is not expected.

7. E_T^{miss}: Missing transverse energy usually appears when neutrinos are involved. This is is not expected for signal events, while $t\bar{t}$ events with the decay chain noted in section 6.1 is expected to have two neutrinos.

We did not use all the variables for both BDT classifiers, but only the ones that were relevant in the specific BDT. For the first BDT we used the variables 1,2,3,6 from the list. For the second BDT we used all the variables listed. In figure 6.2 and 6.3 the distributions for the input variables are shown for both signal and background.

It is also possible to rank the variables according to how often they are used, and their separation power when used. This ranking tells us whether we can exclude some of the variables. The ranking for the variables used in the two BDTs are presented in table 6.4. We see that it could be considered to remove some of the variables in the second classifier, whereas the first seems optimal. The variables were checked for correlations, and it was found that $\triangle R(\gamma, e^{\pm})$ and $p_T^\gamma/(p_T^\gamma + \sum p_T^{\text{tracksC45}})$ was slightly correlated, however, only very little. We chose to keep all variables listed.

Variable	BDT 1	BDT 2
m_Z	40.5	8.6
p_T^γ/p_T^Z	23.1	3.3
$\triangle R(\gamma, (e^+, e^-))$	16.3	1.5
$p_T^\gamma/(p_T^\gamma + \sum p_T^{\text{tracksC45}})$	23.1	10.9
$p_T^{e^\pm}/(p_T^{e^\pm} + \sum p_T^{\text{tracksC45}})$	.	12.7
$m_{Z\gamma}/(m_{Z\gamma} + m_{jets})$	.	39.4
E_T^{miss}	.	17.3
$E_{\text{EM calo}}/E_{\text{All calos}}$	.	6.3

Table 6.4: Importance in percent of input variables in the two BDT classifiers.

6.4.4 Pruning

Pruning is the process of cutting back a tree from the bottom up after it has been grown to its maximum size. Its purpose is to remove statistically insignificant nodes and thus reduce
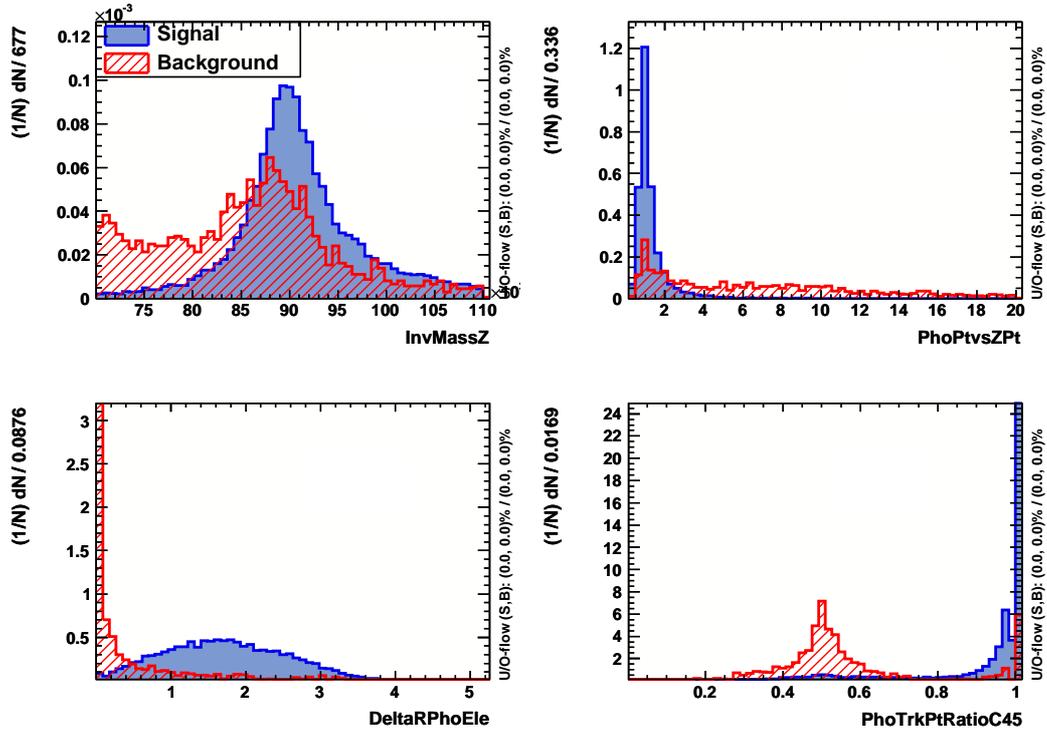
Figure 6.2: This figure shows the variables used in the first BDT which is trained to discriminate FSR. The variables displayed are: m_Z/GeV (top left corner), p_T^γ/p_T^Z (top right corner), $\triangle R(\gamma, e^\pm)$ (bottom left corner), $p_T^\gamma/\sum p_T^{\text{tracksC45}}$ (bottom right corner). We see that the features explained in the list of variables are present in the distributions. For instance, the distribution for $\triangle R(\gamma, e^\pm)$ is strongly peaked towards zero for the background while being more evenly distributed for the signal. The reconstructed Z mass is shifted and broadened for the background, indicating that the FSR photon carries away some of the energy. Notice also the small peak around 0.5 in $p_T^\gamma/(p_T^\gamma + \sum p_T^{\text{tracksC45}})$ for the background. The reason for this peak is not clear since FSR does not suggest that the energy is split evenly between the electron and the radiated photon. Further investigation is needed to clarify this matter. The number of events has been normalized to the total number of events and the bin width.
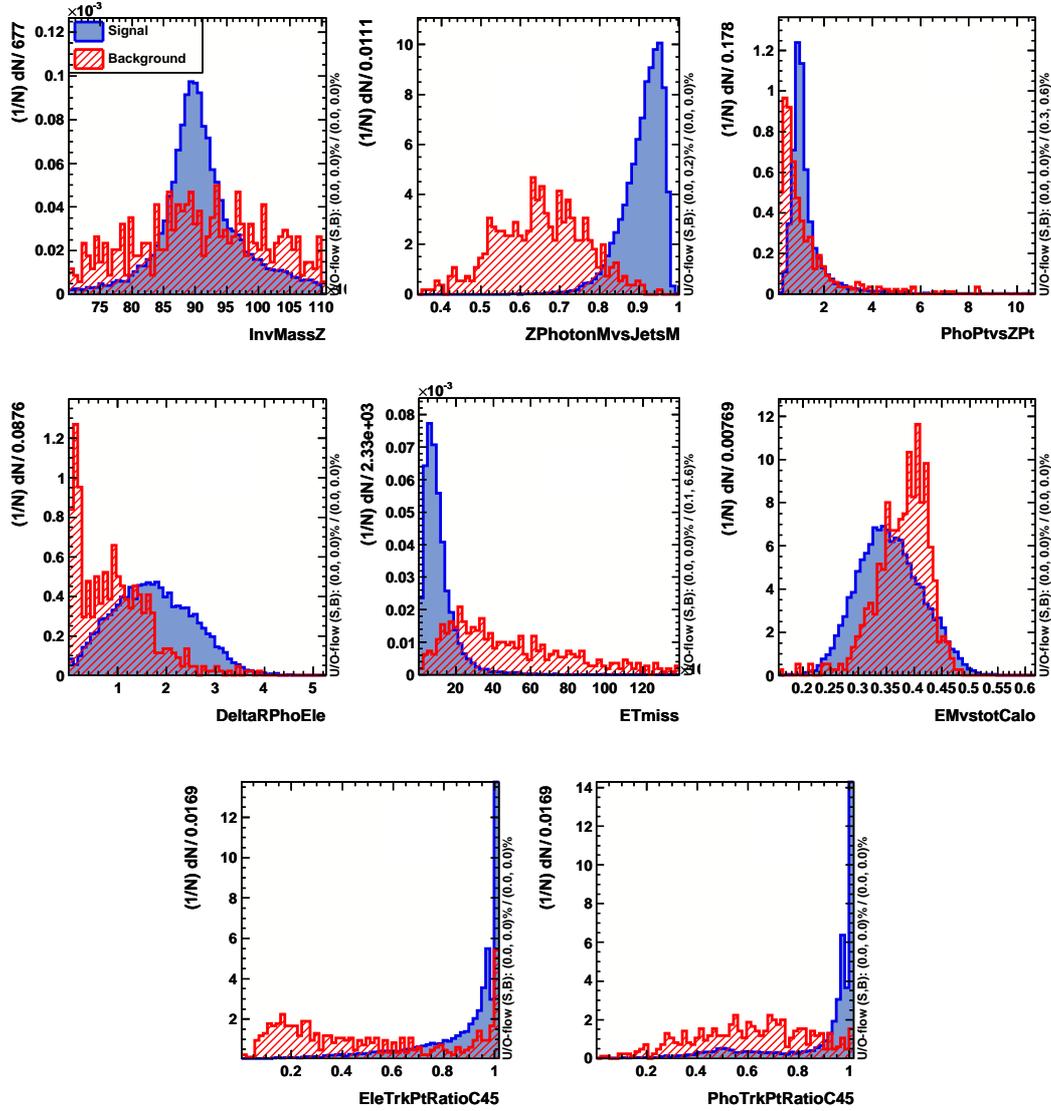
Figure 6.3: This figure shows the variables used in the second BDT which is trained to discriminate "fakes". All events entering these histograms have been subject to a selection cut made in the first BDT. The variables displayed are: m_Z/GeV (top left corner), $m_{Z\gamma}/(m_{Z\gamma} + m_{jets})$ (top middle), p_T^γ/p_T^Z (top right corner), $\triangle R(\gamma, e^\pm)$ (middle left), E_T^{miss}/GeV (middle), $E_{\text{EM calo}}/E_{\text{All calos}}$ (middle right), $p_T^{e^\pm}/(p_T^{e^\pm} + \sum p_T^{\text{tracksC45}})$ (bottom left corner), $p_T^\gamma/(p_T^\gamma + \sum p_T^{\text{tracksC45}})$ (bottom right corner). We see that the features explained in the list of variables are present in the distributions. For instance, the background distribution for m_Z does not have the characteristic peak around 91.187 GeV as we expect for signal events. The number of events has been normalized to the total number of events and the bin width.

the *overtraining* of the tree. To explain this phenomenon, consider the situation where the algorithm has been trained in an unconstrained fashion, such that the trees are grown to the extend that each final node only contains one event. This would yield a perfect separation. However, it is also an artificially effective separation, called overtraining, since the algorithm now is only suitable for the specific sample it has been trained on.

The situation does not have to be this extreme for overtraining to occur. In general, an overtrained algorithm should be thought of as an algorithm which has been tweaked and fine-tuned to separate signal from background in the training sample only, taking advantage of small fluctuations which are not general features.

An overtrained algorithm is identified when the performance of the classifier on the test sample is worse than on the training sample.

To avoid overtraining, the individual trees are pruned. This is done from the bottom up, since apparently insignificant nodes in the top of the tree can lead to good separation in the following nodes. For the same reason, the tree is first grown to its maximum size, which however is constrained by the stop criterion previously mentioned (the default criterion in TMVA is that the end nodes should contain at least 400 events).

The pruning method used in this study was the so-called *Cost-Complexity* pruning. There are also other methods in TMVA, but this is the default one, and since they do not differ significantly in performance [37] we simply used this one. The Cost-Complexity pruning removes nodes that do not provide enough gain in terms of misclassified events. This is done by calculating the cost complexity ρ of each node, defined by

$$\rho = \frac{R(\text{node}) - R(\text{subtree below that node})}{\#\text{nodes}(\text{subtree below that node}) - 1} \tag{6.3}$$

where the cost estimator R is defined as $1 - \max(p, 1-p)$ and p is the purity.

The nodes with a cost complexity value lower than some value, called the prune strength, are then removed. As implied, this is done after the boosting has taken place and thus the weights α_i are not affected.

In our case, we used different approaches for each of our classifiers to get optimal performance. Actually, the first classifier was not pruned at all. Instead, we set a maximum *tree depth*, meaning the maximum number of cuts a given event is subjected to. We found that the optimal performance was achieved when this value was 3. The second classifier was build using the pruning method explained above.

6.4.5   BDT response

The two BDT classifiers used in this selection was build in different ways. As already mentioned in the previous section, the first did not receive pruning whereas the second one did. There are also a number of other parameters to be set when training the BDTs. Among these, one has to choose the number of trees in the forest. We used 400 and 1000 in the two classifiers, respectively. The reason for using more trees in the last classifier, was that we had lowered the stop criterion from being a demand of a minimum of 400 events on the end nodes to 100 events. This was done because we had fewer background events available for training the second classifier. By increasing the number of trees, it seemed we were able to "smooth" out the fluctuations in the classifier response from the individual trees that came when the stop criterion was lowered. We tried also to use a maximum tree depth of 3-4 and not use pruning (which is actually the recommended setting [37]), however, this did not work very

well. Instead, we set the maximum depth relatively high (=10) and employed the pruning method as explained above.

The first classifier performed best when most of the settings were as their default values. The only parameter that we changed was the number of steps that is available to the algorithm when it tries to find the best place to make a cut on a variable. The default value is 20. By setting it to -1, the algorithm tries all possible step sizes and uses the best one. This choice turned out to improve the performance.

Figure 6.4 shows the performances of the two classifiers on both the training and the test samples. It is seen that neither of them seems to suffer significantly from overtraining, since in both cases, the test samples triggers the same responses as the training samples.

6.4.6 Comparison with other MVA's

Within the TMVA package it is quite easy to run other MVA's besides the BDT algorithm. In the present study, we investigated the performance of two other MVA's, Fisher's Linear Discriminant and Artificial Neural Networks, and compared their performances with the BDT.

First, let us briefly give a description of the other two MVA's. We will not provide any details on the specific calculation methods, but refer to [37] for a more thorough exposition.

Fisher's Linear Discriminant

Fisher's linear discriminant is a classification method that projects high-dimensional data onto a line and performs classification in this one-dimensional space.

For simplicity, we consider a two-dimensional example. Imagine we have two classes (say, signal and background) which we describe by two variables. 'Normally', i.e. when using a traditional cut-based method (or a BDT), we will make cuts on these variables in order to separate signal and background.

The Fisher Discriminant method takes a different approach. Here, we construct one single variable from a linear combination of the two original variables. This can, of course, be done in many different ways, see figure 6.5. The trick is to find the "direction" or projection in the phase space of variables which gives the optimal separation.

The Fisher discriminant approaches this problem by assuming that the variable distributions for both signal and background are normally distributed. The, the optimal projection maximizes the distance between the means of the two classes while minimizing the variance within each class.

Artificial Neural Network (ANN)

An artificial neural network is most generally speaking any collection of interconnected nodes, with each node producing a certain response at a given set of input signals. The nodes are most often referred to as neurons due to the correspondence with biological neural networks.

In general, all neurons are connected. For the present application, however, we have a so-called multi-layer perceptron (MLP), in which the neurons are divided into layers and a neuron from one layer only is connected to the neurons in the layers next to it. The first layer is the input layer, the last one, consisting of only one neuron in our case (but in general can have more neurons), is the output layer, while all layers in between are referred to as hidden layers (see figure 6.6). The neuron in the output layer will contain the final response which is used to classify the input.
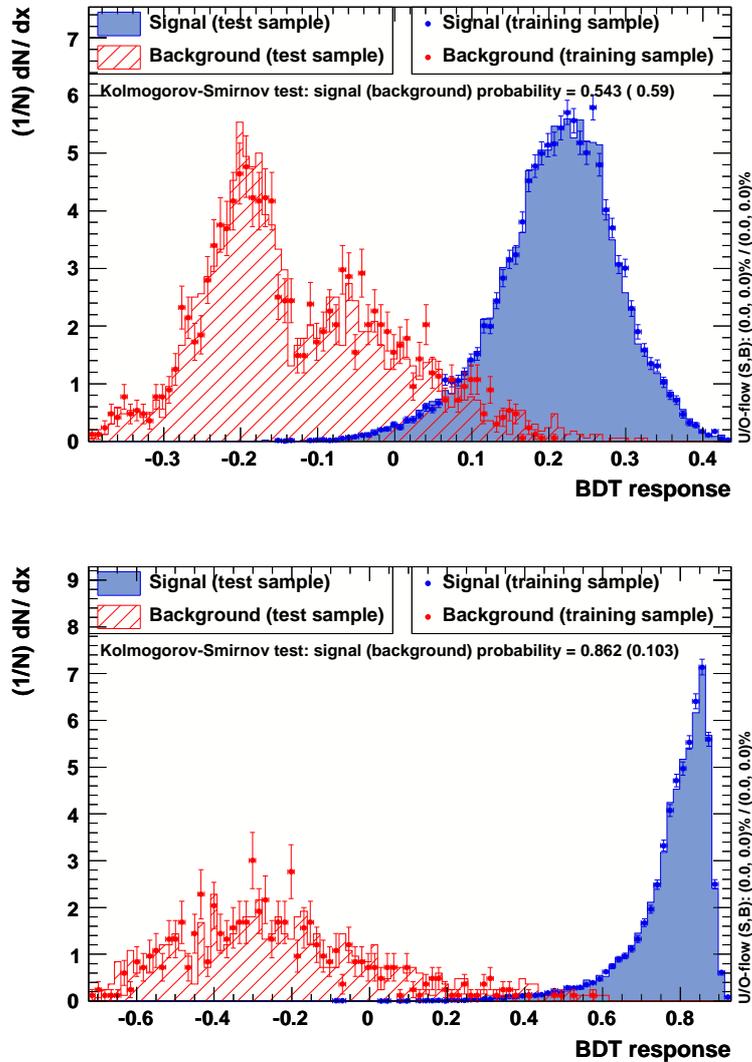
Figure 6.4: This figure shows the classifier responses from the two BDT classifiers. The top figure is for the first BDT that was trained to remove FSR. The bottom figure shows the second classifier which was trained to remove "fakes". We see that in both cases we are able to get a good discrimination. Neither of them seems to suffer significantly from overtraining, since for both classifiers, the test samples triggers the same responses as the training samples. This is also seen from the Kolmogorov-Smirnov test result shown in each histogram.
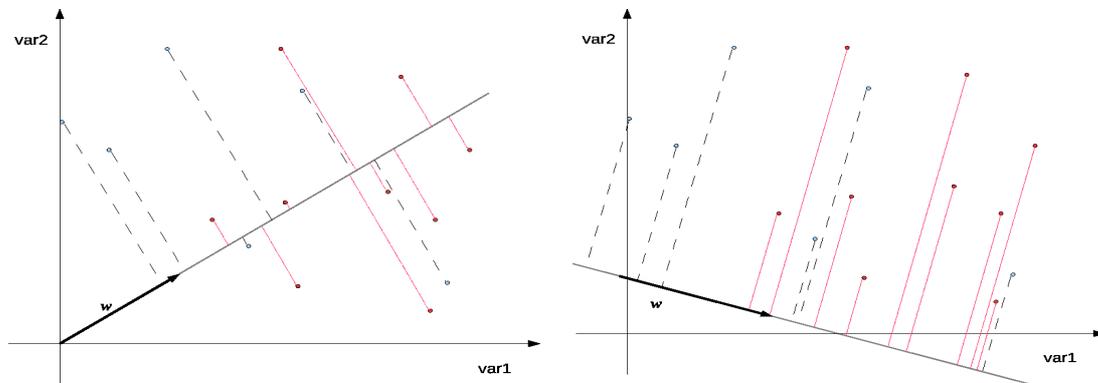
Figure 6.5: This figure shows the basic principle behind the Fisher Discriminant implemented on a fictitious sample of two classes (blue and red points) and two variables (var1 and var2). The two lines given by the vector w represent two different projection lines of the sample onto a linear combination of the two original variables. We notice that the choice of projection in the figure to the left seems to be a bad choice since it mixes the two classes instead of separating them. The choice of projection to the right looks more promising.
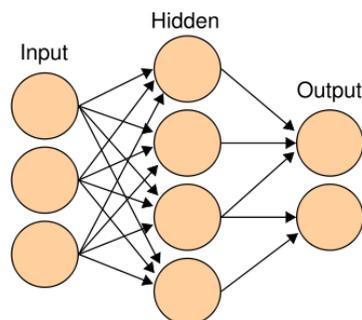
Figure 6.6: This figure shows the basic principle behind a Artificial Neural Network. It shows the input layer consisting of 3 neurons, one hidden layer consisting of 4 neurons and the output layer consisting of 2 neurons. The figure is borrowed from http://en.wikipedia.org

The input to be processed by a given neuron in one of the layers (except for the input layer) is given by the output from the neurons in the preceding layer where each of these outputs is given a weight. As with the BDT, it is necessary to train the ANN, i.e. adjust these weights, until the performance is optimal. This is done by presenting a known input (e.g. 'signal' events) to the neural network and setting a corresponding desired or target response at the output. An error is composed from the difference between the desired response and the system output. This error information is fed back to the system and the weights are adjusted in a systematic fashion. The process is repeated until the performance is acceptable.

The neural network can be viewed as a mapping from a space of input variables onto an one-dimensional (e.g. in the case of a signal-versus-background discrimination problem) or multi-dimensional space of output variables. The mapping is non-linear if at least one neuron has a non-linear response to its input.

The performance depends heavily on the data. If one does not have data that cover a significant portion of the operating conditions or if they are noisy, then neural network technology is probably not the right solution. On the other hand, if there is plenty of data and the problem is poorly understood to derive an approximate model, then neural network technology is a good choice.

Background rejection vs. signal efficiency

The performances of the multivariate analyzers can be compared to each other by plotting the background rejection against signal efficiency when the discrimination threshold is varied. Background rejection is defined as (1-efficiency). Figure 6.7 shows this for the Fisher discriminant, the artificial neural network (labeled MLP in the plot) and the BDT for the two applications of multivariate classification estimators used in this selection.

6.5 Overall selection performance

Ideally, each part of the selection should be such that the error on the final result, i.e. the error on the anomalous couplings, is minimized. However, this would mean that each step taken from the pre-selection and to the determination of the couplings should be done multiple times, which would lead to an multi-dimensional optimization, which is clearly out of the scope for this thesis.

Another obvious way of optimizing the selection is to optimize the statistical significance ($S/\sqrt{S+B}$) for each BDT. However, this might not be such a good choice after all, since it would not yield information of which *kind* of events the selection actually ended up with. Even though we have tried to avoid TGC dependent variables, we have not checked how well this has been achieved. Maybe the "optimal" one would be highly efficient at selecting events with low invariant mass of $Z\gamma$, while discarding events with high invariant mass of $Z\gamma$. This would obviously be a bad choice, since the anomalous couplings grow with center of mass energy. The dilemma comes from the fact that what is defined as "signal" events in the selection, is not the "real" signal before the reweighting procedure described in chapter 3 has taken place.

We chose to take a slightly different approach. The cut values on the classifier responses were made a bit conservative in order not to loose a significant amount of the signal. In general, the more classifiers that are used in a sequence the more signal events will be cut away. The situation is made more complex by the fact that the background used to train the first tree also contains events that will more easily be removed by the second classifier (i.e.
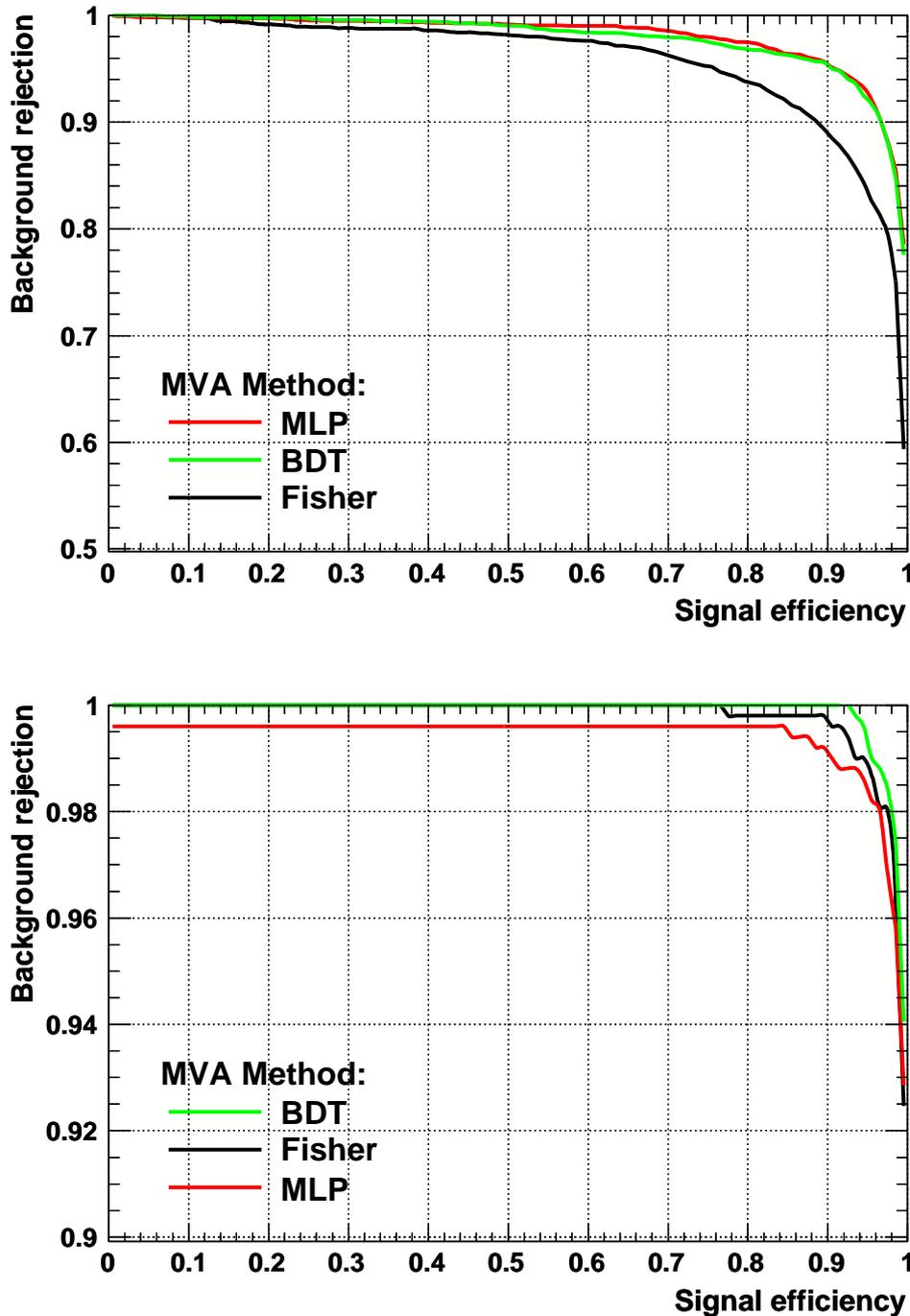
Figure 6.7: This figure shows the the background rejection (defined as 1-efficiency) against signal efficiency (defined as the number of events that passes the selection divided by the initial number of events) when the discrimination threshold is varied for the Fisher discriminant, the artificial neural network (MLP) and the BDT. The top plot is for the discrimination of FSR, the bottom plot is for discrimination of "fakes". In the first case the BDT and MLP are very similar in performance, perhaps with the MLP being slightly better. However, for the chosen signal efficiency (see next section) the difference is zero. In the second case, the BDT clearly outperforms the other two selection methods. It should also be mentioned that in each plot, the "background" consists only of the particular type of background that the selection methods were trained to discriminate.

the part of the Z events that contains "fakes"). A complicating factor is that every time we make a change to first classifier, we have to build the second one again. We tried different cuts, and ended up with $y_{\text{BDT1}}^{\text{cut}} = 0.1$ and $y_{\text{BDT2}}^{\text{cut}} = 0.6$. The first cut value is chosen such that the second cut value can be chosen such that it removes the $t\bar{t}$ background completely without loosing much signal. If the first cut value is raised further to remove more of the first background, a lot of signal signal is be lost here, while the separation power in the second BDT is approximately unchanged. If the first cut value is lowered, the second background cannot be removed without cutting significantly on the signal. With the chosen values, the signal efficiency and background rejection (only for the background against which the specific BDT was trained) for the two BDT classifiers are (sig. eff. , bg reject.): (0.916 , 0.951) and (0.958 , 99.9), respectively. These numbers ("coordinates") can be found on the curves shown in figure 6.7. The results from the entire selection are summarized in table 6.5. The purity of the final sample is 77.6%.

Dataset	Sample	Events (300pb^{-1})	Pre-sel.	BDT1	BDT2	Eff. (%)
$Z\gamma \rightarrow e^+e^-\gamma$	0.27M	468	59.2±0.3	54.2±0.3	51.9±0.3	11.1
$Z \rightarrow e^+e^-$	1.081M	199860	347±8	17±2	15±0.2	0.008
$t\bar{t}$	0.688M	24770	27±1	9±1	0	0
J7	0.1M	36	0	0	0	0
J6	0.1M	84300	6±2	3±2	0	0
J5	0.1M	543000	43±15	5±5	0	0
J4	0.1M	19.4M	1358±513	582±336	0	0

Table 6.5: This table shows the performance of the event selection. The number of events in the columns 3-6 are scaled to a luminosity of 300pb^{-1}. The three selection parts (Pre-sel., BDT1 and BDT2) are applied in sequence. Thus, the final number of selected events can be read off from the BDT2 column. The statistical errors are not simply given by the square root of the number, since the number of produced events (the sample size) differs from the number of selected events. The error is given by $w\sqrt{N_{sample}} = \sqrt{wN_{sel}}$, where N_{sel} is the number of pre-selected events, N_{sample} is the sample size and the event weight w is the ratio between the number of events scaled to 300pb^{-1} and the sample size.

The contamination within the signal sample (as mentioned in section 6.3) is also checked again. The BDT's are not designed to improve the purity in the signal sample, on the contrary, it is trained to exactly recognize events labeled as signal events, including the particles in signal events that are falsely identified. However, since these events constitute a minority of the sample, the BDT's are primarily trained on the "pure" part of the signal, and therefore we may see an increase of the purity within the signal sample after having applied the BDT selection. The purity in the signal is improved to $p_{\text{sig}} = 84.3\%$ (before the BDT-selection, it was 80.6%). Possibly, this number can be improved by making even tighter cuts on the BDT classifiers.

The samples of signal and background described above have all been generated with Pythia with relatively conservative generator cuts as explained in section 6.2. The effects of the anomalous gauge couplings show up at high transverse energies. In order to get a representative sample at higher energies, we have chosen also to simulate a signal sample with a generator cut on the transverse energy of the photon of 40GeV. Table 6.6 shows the selection on this sample. The preselection has been modified with a p_T cut on the photon of 50GeV to take

into account the generator cut. The radiative Z background has also been subjected to this additional cut from which only 2.8 events survived (when scaled to 300pb^{-1}).

Dataset	Sample	Events (300pb^{-1})	Pre-sel.	BDT1	BDT2	Eff. (%)
$Z\gamma \to e^+e^-\gamma$	0.787M	51.1	9.9	9.4	8.8	17

Table 6.6: This table shows the performance of the event selection on the supplementary signal sample which has an additional generator cut on the transverse energy of the photon. The number of events in the columns 3-6 are scaled to a luminosity of 300pb^{-1}. Statistical errors on columns 3-6 are < 0.1. The effective cross section is $\sigma_{\text{eff}} = 0.17025\text{pb}$. For details on definitions, see text in table 6.5

Chapter 7

Analysis

This chapter presents the analysis on the selected $Z\gamma$ candidate events to arrive at the final results on limits on the anomalous trilinear gauge couplings contributing to the $ZZ\gamma$ vertex at a 95% confidence level. As described in section 3.6, we expect to see an effect of the anomalous couplings in the p_T distribution for the photon and therefore we will study deviations from the SM predictions in this observable.

   The naive way of doing this is to make a fit (e.g. χ^2 or likelihood) of the distribution given by a specific value of h_i^Z to the distribution predicted by the SM. However, since the contributions from the anomalous couplings always are positive, no values of the couplings will correspond to a situation where less events than the SM prediction is observed. In such cases, only $h_i^Z = 0$ (the SM values) will fit the distribution and a there is no reasonable way to estimate the error of the fit. However, if we somehow could run the experiment again, a new data set with a slightly different distribution would emerge. Repeating this many times would give us an estimate of how often the distribution would be fitted to non-zero values of h_i^Z and hence an error estimate. Of course, we cannot physically repeat the experiment, but there is way of doing it artificially by introducing so-called pseudo-experiments. We will return to the exact procedure in section 7.2.

   First, we take a look at the reweighting method described in section 3.5. In this chapter, we have omitted the "0" in h_{i0}^Z for practical reasons.

7.1 Reweighting procedure – revisited

This section will explain how the reweighting technique presented in section 3.5 has been employed. It is one of the core features in the analysis program used in the present study, and only through the specific implementation described below, has it been possible to conduct the analysis in a time efficient way. It comprises a boost in analysis time and methodology compared to previous studies.

   The effects of the anomalous contributions are included by reweighting SM $Z\gamma$ events. This is done by giving each event a weight according to the expression (equation (3.16))

$$\text{weight} = \frac{\sum_{i,j=0}^{4} h_i^Z h_j^Z \cdot F_{ij}}{\sigma_{\text{SM}}} \tag{7.1}$$

where σ_{SM} is the SM cross section, the h_i^Z's are the anomalous couplings and the F_{ij}'s are coefficients which consist of SM operator contributions as well as operator contributions as-

sociated with the anomalous couplings. We recall, that for each event, these coefficients are completely specified by the kinematics of the participating particles and does not depend on the anomalous coupling constants.

For a specific set of h_i^Z's, (7.1) is used to obtain the reweighted cross section. If we want to examine the effects for a different set of h_i^Z's, the weight is calculated again using these new couplings. This means that if many different configurations of anomalous couplings are investigated, the reweigh program (see section 3.5) is applied multiple times. Since this is a very time consuming task, a different approach is adopted.

By inspection of (7.1) it is recognized that for a specific sample of events, the only thing that differs in the calculation of two different weights, are the couplings. The F_{ij}'s are unaffected by a change of anomalous coupling constants, since they are determined entirely by the kinematics contained in the operator contributions. Thus, the re-calculation of these coefficients are redundant, and it would be desirable if they could be determined once and for all for each event.

To see how this is possible, we first notice from (7.1) that the F_{ij}'s are symmetric with respect to interchanging i and j

$$F_{ij} = F_{ji} \tag{7.2}$$

A priori, there are 25 different coefficients (remember that (i, j) runs from 0 to 4 since the SM contribution is absorbed in the expression (7.1)). However, using the symmetry property (7.2), we see that there are $25 - 10 = 15$ *independent* coefficients. By using (7.1) it is possible to write down 15 equations that uniquely will determine the F_{ij}'s. This is done with 15 different combinations of h_i^Z and by using the reweigh program to calculate the resulting cross sections for each combination. Clearly, the h_i^Z's must be chosen in a way such that the resulting equations are independent of each other.

To illustrate the procedure, we consider the simplified situation where only one coupling constant exists. In this case, there are 3 coefficients to determine which can be seen from the expression for the cross section

$$d\sigma = F_0 + h^Z \cdot F_1 + (h^Z)^2 \cdot F_2 \tag{7.3}$$

where h^Z is one of the four couplings and $F_0 = d\sigma_{\text{SM}}$.

By using (7.3) three times with $h^Z = \{0, 1, -1\}$ the following matrix equation is constructed

$$\begin{pmatrix} d\sigma_1 \\ d\sigma_2 \\ d\sigma_3 \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{pmatrix} F_0 \\ F_1 \\ F_2 \end{pmatrix} \tag{7.4}$$

If the matrix containing the coupling values are denoted $\hat{A}$, the cross sections by $d\vec{\sigma}$ and the coefficients by $\vec{F}$, the equation is easily manipulated to give the coefficients

$$d\vec{\sigma} = \hat{A}\vec{F} \qquad \Rightarrow \qquad \vec{F} = \hat{A}^{-1}d\vec{\sigma} \tag{7.5}$$

Clearly, $\hat{A}$ must be invertible. This is the case if the couplings used to derive (7.5) are chosen such that the three equations in (7.4) are independent. When considering all four couplings, the matrix $\hat{A}$ is 15×15 and $d\vec{\sigma}$ and $\vec{F}$ are 15-dimensional vectors, but otherwise the situation is the same.
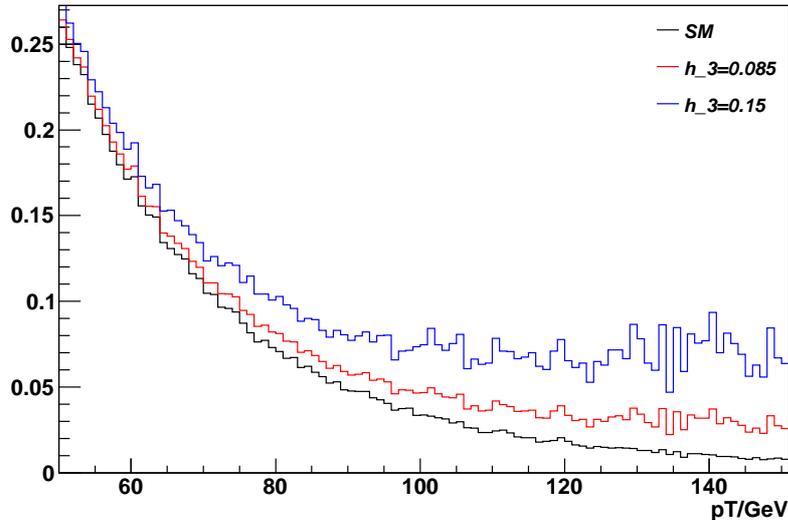
Figure 7.1: This figure shows the distributions of transverse momentum for the reconstructed photons for the SM and when including $h_3^Z = 0.085$ and $h_3^Z = 0.15$ with a form factor scale of $\Lambda = 2\text{TeV}$. The distribution is normalized to 300pb^{-1}, it corresponds to approximately 10 events.

Once $\hat{A}^{-1}$ is determined it can be encoded into the reweigh program. Then, the reweigh program will be initiated 15 times for each event to determine the coefficients $\vec{F}$. Since we will investigate the effects from the anomalous couplings on the p_T distribution of the photon, it is convenient, with regards to consumption of computing time, to take the following approach: 15 histograms of the p_T of the photon are populated. However, with different weights. The first is populated using an event weight of 1 (the SM contribution). The rest are populated using weights equal to F_i, where $i = 1, \ldots, 14$. To get the p_T distribution for a given set of couplings, the individual histograms are simply multiplied with the appropriate couplings and then added together in accordance with (7.1). In this way, it is extremely fast to compute the effects of a given combination of anomalous couplings.

7.2   Fit procedure

As discussed in section 3.6, we expect to see the effects of anomalous couplings in the p_T distribution of the photon. Figure 7.1 shows the distributions of transverse momentum for the reconstructed photons from the selection described in chapter 6 for the case of the SM and when including $h_3^Z = 0.085$ and $h_3^Z = 0.15$ with a form factor scale of $\Lambda = 2\text{TeV}$.

To find the optimal value of h_i^Z for a given data sample, a binned maximum likelihood fit is used. In the following, a brief introduction to this method is given[1].

For a data sample with n_{tot} events, the distribution of some variable, say, p_T^γ is given by a probability density function $f(p_T^\gamma; h_i^Z)$, $i = 1, \ldots, 4$, for which we wish to estimate the

---

[1]For an introduction to maximum likelihood methods and other statistical methods, we refer to [38] from which the following description can be found in a more general form and in much more detail.

unknown parameters h_i^Z. The events from the data sample are divided into N bins and the expectation value $\nu_j(h_i^Z)$ for the number of events in bin i is given by

$$\nu_j(h_i^Z) = n_{\text{tot}} \int_{j^{\text{th}} \text{ bin}} f(p_T^\gamma; h_i^Z) \mathrm{d}p_T^\gamma \tag{7.6}$$

Since the number of events in a given bin is independent of the other bins, the probability $f_{n_{\text{tot}}}(n_1, \ldots, n_N; \nu_1, \cdots, \nu_N)$ of having a certain occupation of events in the bins (i.e. a certain configuration) is given by a multinomial distribution

$$f_{n_{\text{tot}}}(n_1, \ldots, n_N; \nu_1, \ldots, \nu_N) = \frac{n_{\text{tot}}!}{n_1! \cdots n_N!} \left(\frac{\nu_1}{n_{\text{tot}}}\right)^{n_1} \cdots \left(\frac{\nu_N}{n_{\text{tot}}}\right)^{n_N} \tag{7.7}$$

where the probability for having n_i events in bin i has been expressed as the expectation value $\nu_j(h_i^Z)$ divided by the total number of events n_{tot}.

As a further step, the total number of entries in the data sample n_{tot} is regarded as a random variable from a Poisson distribution with mean $\nu_{\text{tot}} = \sum_{i=1}^N \nu_i$. Therefore, the joint probability density function for observing n_{tot} events distributed in N bins with $n_1, \ldots, n_N$ events in each bin is the product of a Poisson distribution and a multinomial distribution similar to the one given in (7.7), however, with the probability for having n_i events in bin i expressed as the expectation value $\nu_j(h_i^Z)$ divided by the *expectation value* of the total number of events ν_{tot}:

$$f_{\text{joint}}(n_1, \ldots, n_N; \nu_1, \ldots, \nu_N) = \frac{\nu_{\text{tot}}^{n_{\text{tot}}} e^{-\nu_{\text{tot}}}}{n_{\text{tot}}!} \frac{n_{\text{tot}}!}{n_1! \cdots n_N!} \left(\frac{\nu_1}{\nu_{\text{tot}}}\right)^{n_1} \cdots \left(\frac{\nu_N}{\nu_{\text{tot}}}\right)^{n_N} \tag{7.8}$$

where the expected number of entries in each bin $\nu_j(h_i^Z)$ now depends on ν_{tot} instead of n_{tot}:

$$\nu_j(\nu_{\text{tot}}, h_i^Z) = \nu_{\text{tot}} \int_{j^{\text{th}} \text{ bin}} f(p_T^\gamma; h_i^Z) \mathrm{d}p_T^\gamma \tag{7.9}$$

Using the constraints $\nu_{\text{tot}} = \sum_{i=1}^N \nu_i$ and $n_{\text{tot}} = \sum_{i=1}^N n_i$, it is possible to write (7.8) in a more compact form [38]. We get

$$f_{\text{joint}}(n_1, \ldots, n_N; \nu_1, \ldots, \nu_N) = \prod_{i=1}^N \frac{\nu_i^{n_i}}{n_i!} e^{-\nu_i} \tag{7.10}$$

This is called the likelihood function. We see that the distribution of p_T^γ is described by the product of N individual Poisson distributions (one for each bin) with mean values $\nu_j(\nu_{\text{tot}}, h_i^Z)$. The likelihood function (7.10) is the function we would like to maximize to find the value of h_i^Z that best describes the distribution. The logarithm of the likelihood function has the same maximum as the likelihood function. Therefore, the log-likelihood function is maximized instead, since the logarithm reduces the product in the likelihood function to a sum, which is easier to compute. It also allows us to drop all terms that does not depend on the parameters we wish to determine. The log-likelihood function is defined as

$$\log L(\nu_{\text{tot}}, h_i^Z) = -\nu_{\text{tot}} + \sum_{j=1}^N n_j \log \nu_j(\nu_{\text{tot}}, h_i^Z) \tag{7.11}$$
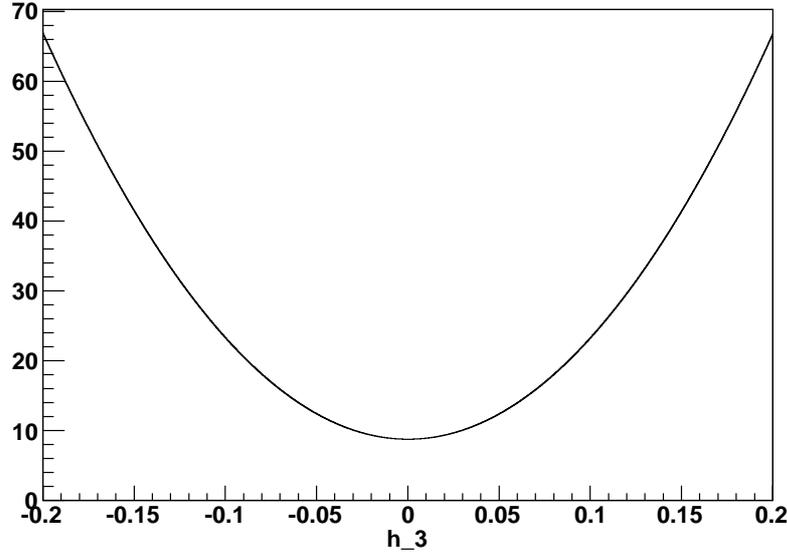
Figure 7.2: This figure shows the total number of events (or equivalently, the total cross section) as a function of h_3^Z. The plot is essentially the same as the plot in figure 3.6, however, only showing the projection onto the plane perpendicular to the p_T-axis, and the entries being generated by 10000 pseudo-experiments. The form factor scale is $\Lambda = 2\text{TeV}$.

The values of h_i^Z for which the log-likelihood function is maximized are called $h_{i,\text{fit}}^Z$. In the following only one h_i^Z is considered at the time. The case of multiple simultaneous couplings is predominantly of interest in the event that deviations are observed in a single parameter fit, in order to understand the nature of the model.

As already mentioned, there are problems connected with the determination of the h_i^Z from distributions with less events than the SM prediction. Since no values of the anomalous couplings will generate a cross section lower than the SM cross section, the theoretical sample space is only a subspace of the experimental sample space. This is illustrated in figure 7.2, where the total number of events (reflecting the total cross section) is plotted as a function of the coupling h_3^Z. We see that for all values of $|h_3^Z| > 0$, the predicted number of events is above the SM prediction ($h_3^Z = 0$). If the data sample contains fewer events than what the SM predicts, the best fit will be at the minimum of the curve, i.e. the SM prediction. However, it is not possible to estimate the error correctly, since samples with fewer events than the SM prediction are indistinguishable from each other and the estimation of the error is strongly biased. It is also evident that for data samples which contain more events than the SM predicts, there are two values of the coupling that give this amount of events. This suggests that we will not be able to distinguish between $-h_i^Z$ and $+h_i^Z$.

In order to get the correct error estimate, it is necessary to consider a larger ensemble of SM samples, which naturally will fluctuate around the SM prediction and thus provide samples in the full region to create th non-trivial transformation from experiment space to theory space, reflecting the union of the theoretical sample space and the experimental sample space.
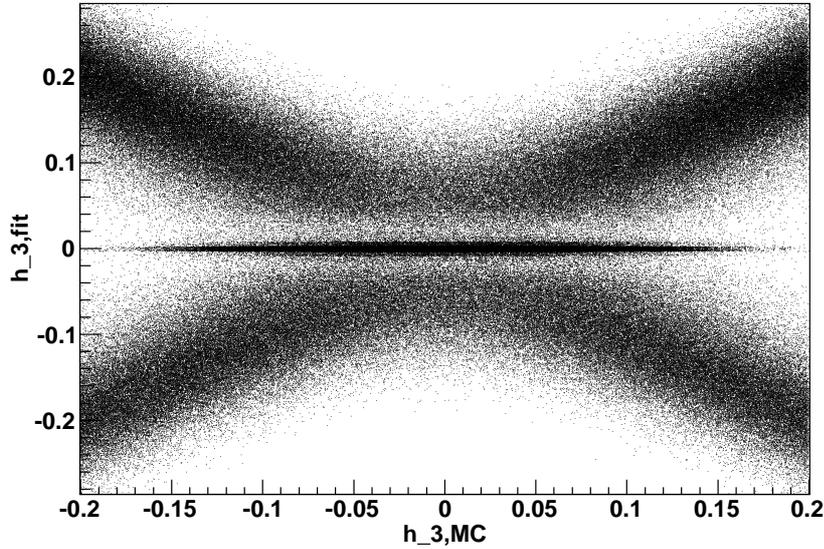
Figure 7.3: This figure shows the distribution of $h^Z_{i,\mathrm{fit}}$ as a function of $h^Z_{i,\mathrm{MC}}$ for the p_T distribution for the photon. The form factor scale is $\Lambda = 2\,\mathrm{TeV}$.

Practically, this is not possible, of course, since we would have to build a large ensemble of "LHC's". Instead, we generate an ensemble of so-called Monte Carlo pseudo-experiments which represent possible outcomes of repeated experiments. A pseudo-experiment for a given value of $h^Z_{i,\mathrm{MC}}$ is generated by fluctuating the number of events in each bin according to a Poisson distribution. In each bin, the mean of the Poisson distribution is set by the predicted amount of events for the given choice of $h^Z_{i,\mathrm{MC}}$.

Each pseudo-experiment is then subjected to the likelihood fit to yield the value $h^Z_{i,\mathrm{fit}}$. From this, we are able to get the distribution of $h^Z_{i,\mathrm{fit}}$ as function of $h^Z_{i,\mathrm{MC}}$. Figure 7.3 shows a plot of this dependence. We notice that for large positive or negative values of $h^Z_{i,\mathrm{MC}}$, the fit reproduces the input value $|h^Z_{i,\mathrm{MC}}|$ well. This is expected, since large couplings easily will be detected. Also, the lack of sensitivity to the sign of the coupling is clearly visible. As $h^Z_{i,\mathrm{MC}}$ comes closer to zero, we see that an increasingly large portion of $h^Z_{i,\mathrm{fit}}$ is situated around zero, i.e. the SM prediction. This is what we expect, since an increasingly large portion of the pseudo-experiments will be samples with numbers of events below the SM prediction that cannot be reached by the theory including couplings. It is seen in the plot as a thin horizontal line which gets slightly broader towards (0,0) and is symmetrical about this point. At $h^Z_{i,\mathrm{MC}} \sim 0$ most of the pseudo-experiments are fitted with $h^Z_{i,\mathrm{fit}} \sim 0$.

7.3   Results

We are now ready to estimate the limits that can be obtained from the present study. We have included the backgrounds discussed in the previous chapter, thus, the data sample consists of all events selected in chapter 6 which are 8.8 signal events (from table 6.6) and 2.8 radiative Z

events. We only consider the p_T range going from 50GeV to 150GeV. The reason for removing the high-p_T region is that the reweighting procedure gives large fluctuations in the weights (differences of a factor $\sim 10^6$) distributed on few events in this region. One of the underlying assumptions when using the reweighting technique is that all kinematic configurations are well populated. This is not the case in the present sample if including the p_T-region. It *is* possible to include the region and derive the corresponding sensitivity limits following the procedure described above, however, the error on the limits would have to be re-evaluated to take into account that only a few events ($\sim 2-3$), yielding large statistical uncertainties, completely determine the limits. It is, of course, not desirable to remove the high-p_T tail of the distribution, since this is where we expect the effects of the couplings to show up. A common feature to investigate is the so-called "over-flow" bin, which is located at the point where only a few events are collected in the individual bins. The over-flow bin contains the sum of all contributions in the very high-p_T region where the sensitivity is maximum. Consequently, the limits obtained here are conservative estimates of what can be achieved within the first year of running.

The limits are estimated by using an ensemble of 10000 pseudo-experiments generated with $h_{i,\text{MC}}^Z = 0$ which are representable for the SM. For each pseudo-experiment, the likelihood fit described in the previous section is performed. The 95% confidence level is then defined as by an upper and a lower limit on $h_{i,\text{fit}}^Z$ such that 2.5% of the pseudo-experiment outcome are the regions above and below these limits, respectively. Figure 7.4 shows the distribution of $h_{i,\text{fit}}^Z$. As expected, there is a large peak at zero and two smeared peaks on each side. The peak at $h_{i,\text{fit}} \sim 0$ represents $\sim 50\%$ of the pseudo-experiments, i.e. the cases with less events than the SM prediction. Naively, we would expect the remaining cases to more or less follow the tail of a Poisson distribution, and this does not suggest two small peaks on each side of the central peak. The reason that these smeared peaks appear anyway, is that even pseudo-experiments that contain more events than the SM prediction can have the p_T distribution in a way that forces $h_{i,\text{fit}}^Z$ towards zero. Imagine, that for a given pseudo-experiment, the events are distributed such that there is an overpopulation for low p_T with respect to the SM prediction, but a small underpopulation at high p_T, and that the total number of events are slightly larger than the SM prediction. Since the effects of the couplings are most sensitive in the high-p_T region, the fit will be forced towards $h_{i,\text{fit}}^Z = 0$, even though we would expect the fittet value to be above zero based on the total number of events. In this way the pseudo-experiments that we, naively, would expect to contribute to the regions between the central peak and the two smeared peaks, are "vacuumed" by the central peak. The small asymmetry around the central peak reflects the effects of the linear term in h_i^Z that interferes with the Standard Model in the weight (7.1).

The final results for all four couplings are displayed in table 7.1. The first column shows the limits on the couplings when not including the background in the fit, the second column shows the result when the background is included. Also displayed are the bounds from unitarity, as discussed in section 3.3.

The limits on $h_{2,4}^Z$ are smaller than on $h_{1,3}^Z$ reflecting that the first two are dimension 8 whereas the latter are dimension 6 (see chapter 3). This makes $h_{2,4}^Z$ depend stronger on energy which provides better limits.

Not all of the limits are within the unitarity bounds, only h_{10}^Z and h_{20}^Z are within the bounds. Comparing the limits with those obtained at the Tevatron (see section 3.3.1), we see that the prospects are that the ATLAS experiment will not improve the current limits
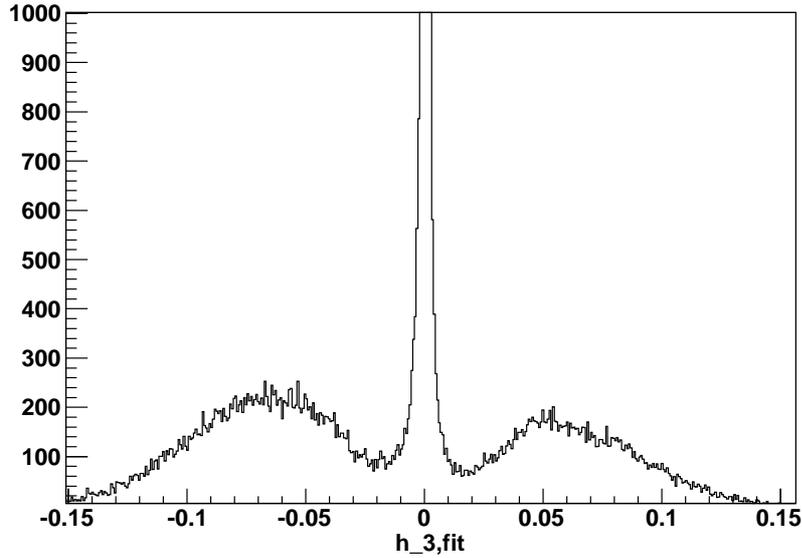
Figure 7.4: This figure shows the likelihood fit to the SM.

95% CL (without bg)	95% CL (with bg)	Unitarity bound
$-0.1085 < h_{10}^Z < 0.0975$	$-0.1096 < h_{10}^Z < 0.0952$	$|h_{10}^Z| < 0.126$
$-0.00592 < h_{20}^Z < 0.00611$	$-0.00577 < h_{20}^Z < 0.00621$	$|h_{20}^Z| < 0.0022$
$-0.1002 < h_{30}^Z < 0.1060$	$-0.1040 < h_{30}^Z < 0.1040$	$|h_{30}^Z| < 0.126$
$-0.00617 < h_{40}^Z < 0.00590$	$-0.00613 < h_{40}^Z < 0.00588$	$|h_{40}^Z| < 0.0022$

Table 7.1: Limits on anomalous couplings at 95% confidence level estimated without background. Unitarity bounds have been calculated using the inequalities (3.11) with $\Lambda = 2$TeV.

within the first year of running. This is, perhaps, a bit surprising. The loss in statistics due to lower luminosity (we consider 300pb^{-1} whereas the Tevatron limits are for 1fb^{-1}) are more than compensated for by the increase in collision energy (this study considers $\sqrt{s} = 7$TeV whereas the Tevatron has $\sqrt{s} = 1.96$TeV). Higher luminosity provides better statistics since more events are collected. The relative error decreases as $\sqrt{N}$ where N is total number of events and since the integrated luminosity is proportional to N, the sensitivity will increase as $\sqrt{\mathcal{L}}$. Recalling, that the helicity amplitudes depend on the center of mass energy of the interaction through $\sqrt{s}^3$ for $h_{1,3}^Z$ and $\sqrt{s}^5$ for $h_{2,4}^Z$ (see equation 3.10), we see that as energy is increased, the amplitudes grows with powers of 3 and 5. Hence, a simple calculation with $\mathcal{L}_{\text{LHC}}/\mathcal{L}_{\text{Tevatron}} \sim 1/3$ and $\sqrt{s}_{\text{LHC}}/\sqrt{s}_{Tevatron} \sim 3$, suggests an expected increase in sensitivity of a factor of $\sim 3^3 \cdot \sqrt{\frac{1}{3}} \sim 16$ for $h_{1,3}^Z$ and $\sim 3^5 \cdot \sqrt{\frac{1}{3}} = 140$ for $h_{2,4}^Z$. These estimates are clearly too optimistic as we have just seen from comparing the limits displayed in tables 3.1 and 7.1. The explanation is to be found in the different environments under which the data have been/will be taken. LHC is a proton-proton collider which means that Z bosons are primarily created in the collision of a *valence u* quark and a $\bar{u}$ *sea-quark* (for definitions of valence and sea quarks, see section 4.1). The sea quark will, on average, not carry as large a

momentum fraction as the valence quark which will have about 20% of the proton momentum (see figure 4.2). The Tevatron is a proton-antiproton collider which means that the $\bar{u}$ will not be a sea-quark, but a valence quark. Thus the fraction of energy entering the hard process with respect to the energy available (the proton/antiproton energy) is larger at the Tevatron than at LHC, on average. This difference will even out the increase in collision energy and thus partly explain the discrepancy between a scaling of the Tevatron results and the limits presented in table 7.1.

Including the background should, in principle, further decreases the sensitivity. This is obvious, since, by including backgrounds, the "data" sample constituting the SM prediction will be a sum of the $Z\gamma$ sample and the background samples. Thus, the fluctuations represented by the pseudo-experiments will be larger, while only the $Z\gamma$ sample is reweighted to fit these larger fluctuations. This will result in numerically larger fitted values of $h^Z_{i,\text{fit}} = 0$, and hence numerically larger limits. However, from table 7.1 we that the background does not have much effect on the final result. The background sample comprises 15 events which are rescaled to 2.8 to correspond to 300pb^{-1}. A larger background sample would provide a better estimate of the effects. However, 7Tev samples are not available in ATLAS, and therefore all samples used here have been simulated specifically as part of this thesis.

There are numerous systematic uncertainties, not included in this study, that also affect the cross section and hence the sensitivity limits on the transverse energy of the photon. They can be divided into experimental and theoretical uncertainties. The experimental uncertainties mainly comes from the measurement of the luminosity and the reconstruction and identification of electrons and photons. There are different experiments that measure the luminosity at ATLAS. The ALPHA experiment [39] measures scattered particles 240 meters from the interaction point and is able to estimate the luminosity from this. It is estimated that the error on the luminosity measurement can be brought down to 2-3%. The particle identification relies on the energy measurements in the detector. According to [35], the uncertainty on electron and photon identification efficiencies are $\sim 3\%$. The theoretical uncertainties mainly come from the uncertainties related to the parton distribution functions (pdf). In order to estimate this uncertainty, new samples of signal and background with different pdf's should be simulated and analyzed as in the present and the previous chapters. However, this is too time consuming and thus out of reach for our study. Instead, we refer to the standard ATLAS estimate of 3-5%.

All numbers cited above are estimated for the situation where the experiment operates at the design specifications, i.e. at a collision energy of 14TeV and an integrated luminosity of approximately 100fb^{-1}. The uncertainties are expected to decrease with time. It is difficult to estimate how large the uncertainties will be under the conditions of the first year of running (as considered in the present study). In general, the errors will be larger, since experimental conditions are better determined the longer the experiment has been running. However, the parton distribution functions are better determined at lower energies of the LHC, because the environment is closer in energy to previous or currently operating experiments from which the pdf's are determined.

Chapter 8

Conclusion and outlook

As the LHC experiments start taking data at unprecedented collision energies from February 2010, the world of particle physics will enter a new era. There are strong indications that new physics will appear at the energies accessible to the LHC and different approaches are taken to investigate the observational consequences. The main goal of the ATLAS experiment is to investigate the missing pieces of the Standard Model and discover the mechanism of electroweak symmetry breaking that explains how the particles achieve their masses. The new physics to be discovered may very well be associated with gravity, the only known force in nature not described through quantum field theory and therefore not included in the Standard Model framework.

The new physics could be operating at too high an energy scale to allow direct detection. In this case, it may be possible to observe the indirect consequences instead, since any new phenomena undoubtedly will affect the couplings between all other particles in the model, especially between the gauge bosons. The tight constraints on these due to the symmetries imposed on the SM Lagrangian make an investigation of the gauge bosons self-interactions a sensitive probe in discovering the effects of physics operating at a higher energy scale.

The present study has been concerned with a Monte Carlo study of the observational effects of non-SM self-interactions in the ATLAS experiment by introducing the anomalous trilinear gauge vertex $ZZ\gamma$. The work has involved both: a study of the modern theoretical approach using effective Lagrangians; extensive Monte Carlo simulations using GRID facilities; advanced statistical multivariate data analysis techniques including Boosted Decision Trees and log-likelihood fitting techniques combined with statistical methods employing so-called pseudoexperiments for setting limits on the non-SM couplings. The amount of data used in the analysis corresponds to the expected amount of data to be collected within the first year of running, $\mathcal{L} = 300\text{pb}^{-1}$ at a center of mass energy of $\sqrt{s} = 7\text{TeV}$.

Using the effective Lagrangian approach, it was shown that four independent couplings h_i^Z contribute to the $ZZ\gamma$ vertex in the $U(1)_{\text{em}}$ framework in which Lorentz invariance, $U(1)_{\text{em}}$ gauge symmetry and Bose symmetry are imposed. The observational effects turn up in the p_T distributions of the bosons because of the strong center of mass energy dependence on the couplings which is directly related to their non-renormalizability. We studied the p_T distribution of the photon. The study was carried out by reweighting SM diboson $Z\gamma$ events simulated with Pythia using calculations extracted from the BHO event generator which includes the non-SM couplings relevant for our study (whereas Pythia itself does not).

The primary backgrounds to $Z\gamma$ events are radiative Z events and $t\bar{t}$ events. The use

of Boosted decision trees (BDT) proved to be extremely efficient in separating signal and background. By using two BDT's in sequence, nearly all background was removed and a purity of 77.6% was achieved.

The results from the analysis on the selected $Z\gamma$ candidate events consisted of limits at a 95% confidence level on the anomalous trilinear gauge couplings contributing to the $ZZ\gamma$ vertex. The expected limits were established using pseudo-experiments without which a proper estimate on the statistical error could not be made. It was found that the prospects of improving the limits on the couplings relevant for the $ZZ\gamma$ vertex at $\sqrt{s} = 7$TeV and $\mathcal{L} = 300$pb^{-1} are not promising. The current limits on the anomalous couplings are set by the experiments at the Tevatron ($|h_{30}^Z| < 0.033$, $|h_{40}^Z| < 0.0017$) and the current study does not show improvements. The limits obtained in the present study are $-0.1040 < h_{30}^Z < 0.1040$ and $-0.00613 < h_{40}^Z < 0.00588$ (citing only $h_{3,4}^Z$ since $h_{1,2}^Z$ are similar). The limits on $h_{2,4}^Z$ in this study are already excluded by unitarity constraints.

The perspectives of improving the limits using the full potential of the ATLAS experiment look more promising. According to [22] it is possible to put strong constraints on the anomalous couplings in the ATLAS experiment. Studies carried out for a center of mass energy of 14TeV and an integrated luminosity of 100fb^{-1} show that ATLAS may bring down the limits to $|h_3^Z| < 5.2 \cdot 10^{-4}$ and $|h_4^Z| < 1.2 \cdot 10^{-6}$. Part of this improvement is due to the use of additional observables. It is possible that new physics will appear before reaching these limits either as direct signal or as increases in the tails of the p_T distributions of the bosons. In the latter case, the form factor scale gives the energy scale at which the new physics can be seen directly.

Acknowledgements

Bibliography

[1] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory.* Westview Press, 1995.

[2] A Zee. *Quantum Field Theory in a nutshell.* Princeton University Press, 2003.

[3] Jan Ambjørn and Jens Lyng Petersen. *Quantum Field Theory.* Lecture Notes, 1999.

[4] A. Dobado, A. Gomez-Nicola, A.L. Maroto, and J.R. Pelaez. *Effective Lagrangians for the Standard Model.* Springer-Verlag Berlin Heidelberg New York, 1997.

[5] T. Kinoshita and M. Nio. Improved alpha**4 term of the electron anomalous magnetic moment. *Phys. Rev. D73, 013003 (2006), hep-ph/0507249*, 2006.

[6] G. et al. (UA1 Collaboration) Arnison. Experimental observation of isolated large transverse energy electrons with associated missing energy at s**(1/2)=540-GeV. *Phys. Lett. B122, 103-116 (1983)*, 1983.

[7] M. et al. (UA2 Collaboration) Banner. Observation of single isolated electrons of high transverse momentum in events with missing transverse momentum at the CERN anti-p p collider. *Phys. Lett. B122 476-485 (1983)*, 1983.

[8] G. et al. (UA1 Collaboration) Arnison. Experimental observation of lepton pairs of invarint mass around 95-GeV/c**2 at the CERN SPS collider. *Phys. Lett. B126, 398-410 (1983)*, 1983.

[9] P. et al. (UA2 Collaboration) Bagnaia. Evidence for Z0 -> e+e- at the CERN anti-p p collider. *Phys. Lett. B129, 130-140 (1983)*, 1983.

[10] J. et al Abdallah. Study of Triple-Gauge-Boson Couplings ZZZ, $ZZ\gamma$ and $Z\gamma\gamma$ at LEP. *arXiv:0706.2741v1*, 2006.

[11] Sidney Coleman and Jeffrey Mandula. All Possible Symmetries of the S Matrix. *Phys. Rev. 159 (1967) 1251*, 1967.

[12] Y. et al. (Super-Kamiokande Collaboration) Fukuda. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett. 81, 1562-1567 (1998), hep-ex/9807003*, 1998.

[13] H. Tardioli P. Mery, M. Perrottet. Z boson pair production and compositeness at large hadron colliders. *Z. Phys. C - Particles and Fields 55, 251-255 (1992)*, 1992.

[14] D. et al. Acosta. Measurement of $W\gamma$ and $Z\gamma$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *arXiv:hep-ex/0410008v2*, 2005.

[15] D. Choudhury. Trilinear Neutral Gauge Couplings. *arXiv:hep-ph/0011205v2*, 2001.

[16] P. Rosendahl. Search for Non Commutative Geometry in ATLAS using dilepton events. *Master thesis, Experimental High Energy Physics Group, Niels Bohr Institute, University of Copenhagen*, 2008.

[17] R. et al. Barate. *Phys. Lett., B453, 107-120*, 1999.

[18] U. Baur and E. L. Berger. Probing the weak boson sector in $Z\gamma$ production at hadron colliders. *Physical Review, D47, 4889*, 1993.

[19] G. J. Gounaris, J. Layssac, and F. M. Renard. Off-shell structure of the anomalous Z and γ self-couplings. *Physical Review D, volume 61, 073012*, 2000.

[20] G. J. Gounaris, J. Layssac, and F. M. Renard. Signatures of the anomalous $Z\gamma$ and ZZ production at lepton and hadron colliders. *Physical Review D, volume 61, 073013*, 2000.

[21] F. et al Larios. Trilinear Neutral Gauge Boson Couplings in Effective Theories. *arXiv:hep-ph/0012180v1*, 2000.

[22] S. Hassani. Prospects for measuring neutral gauge boson couplings in $Z\gamma$ production with the ATLAS detector. *ATL-PHYS-2003-023*, 2003.

[23] J. Alcaraz. On the experimental effects of the off-shell structure in anomalous neutral triple gauge vertices. *arXiv:hep-ph/0111283v4*, 2002.

[24] T. Barklow, S. Dawson, H. Haber, and J. Siegrist. *Advanced Series on Directions in High Energy Physics—Vol. 16: Electroweak symmetry breaking and new physics at the TeV scale.* World Scientific Publishing Co. Pte. Ltd., 1996.

[25] Alexey V. Ferapontov. Measurements and searches for new physics in diboson processes with the DØ detector. *Phd theses, available at http://lss.fnal.gov/archive/thesis/fermilab-thesis-2009-10.shtml*, 2009.

[26] T. Sjöstrand, Edén, L Friberg, L. Lönblad, G. Miu, S. Mrenna, and E. Norrbin. PYTHIA. *Comput. Phys. Commun. 135 (2001) 238. arXiv:hep-ph/0010017*, 2001.

[27] Han T. Ohnemus J. Bauer, U. QCD corrections and anomalous couplings in $Z\gamma$ production at hadron colliders. *Phys. Rev. D, volume 57 number 5*, 1998.

[28] M. A. et al Dobbs. Les Houches guidebook to monte carlo generators for hadron colliders. *arXiv:hep-ph/0403045v2*, 2004.

[29] M. A. Dobbs and J. B. Hansen. The HepMC++ Monte Carlo Event Record for high enrgy physics. *CERN/ATL-SOFT-200-001*, 2004.

[30] Gideon Bella. Weighting Di-Boson Monte Carlo Events in Hadron Colliders. *arXiv:0803.3307v1 [hep-ph] 23 Mar 2008*, 2008.

[31] ATLAS collaboration. The ATLAS experiment at the Large Hadron Collider. *JINST, 3 S08003 (available at https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasTechnicalPaper)*, 2008.

[32] K. et al Assamagan. The ATLAS simulation project. *preprint to be submitted to JINST*, 2009.

[33] GEANT4 collaboration. GEANT4 — a simulation toolkit. *preprint submitted to Elsevier Science*, 2002.

[34] K. Gregersen. Project on the ATLAS detector simulation software. *Internal project at NBI*, 2009.

[35] The ATLAS collaboraton. Diboson physics studies with the ATLAS detector. *ATLAS-PHYS-PUB-2008*, 2008.

[36] Particle Data Group Collaboration. Review of particle physics. *J. Phys. G., 33*, 2006.

[37] A. Hoecker, J Speckmayer, J Stelzer, E Therhaag, E. von Torne, and H Voss. TMVA 4 - Toolkit for Multivariate Data Analysis with ROOT. *arXiv:physics/0703039v5*, 2009.

[38] G. Cowan. *Statistical Data Analysis*. Oxford science publications, 1998.

[39] Matthieu Heller. ATLAS CNI Experiment - Total cross sections and luminosity. *Preliminary paper. http://cdsweb.cern.ch/record/1210688/files/ATL-COM-PHYS-2009-557.pdf*, 2009.