UNIVERSITY
OF OSLO

Master's thesis

Implementation of Beyond the Standard Model Z' models using reweighting tools applied on PowHeg and Sherpa

Martin Føll

Nuclear and Particle Physics

60 ECTS study points

Department of Physics

Faculty of Mathematics and Natural Sciences

Spring 2023

Martin Føll

Implementation of Beyond the Standard Model Z' models using reweighting tools applied on PowHeg and Sherpa

Supervisors:

Farid Ould-Saada

James Catmore

Abstract

The Standard Model (SM) has been successful in describing the particles and interactions that we know of today with great precision. We know however that it is not the ultimate theory since it does not, for example, describe gravity nor accommodate dark matter. Among the various beyond Standard Model (BSM) theories, we concentrate on extensions which predict one or several neutral massive spin-1 gauge bosons, Z'. Monte Carlo simulations are important for making predictions and interpreting measurements based on data from the ATLAS detector at the LHC. A method called signal reweighting is used to produce a Z' signal from an existing Drell-Yan simulation, thus avoiding simulating a large number of Z' events based on various models depending on many parameters. The signal reweighting method is extended to use the next-to-leading order PowhegPythia8 Drell-Yan sample. We obtain promising results when comparing the signal reweighted Z' sample to a fully simulated Pythia8 Z' sample in Run 2 at $\sqrt{s} = 13$ TeV with 58.5 fb^{-1} of data. The signal reweighted Z' samples are also used in Run 3 on top of data and SM backgrounds at $\sqrt{s} = 13.6$ TeV corresponding to 26.1 fb^{-1} of data collected by ATLAS in 2022. An attempt to apply signal reweighting to a Sherpa Z + jets sample is also performed but the information from the event generator is found to be insufficient.

Acknowledgements

I will first and foremost thank my supervisor Farid Ould-Saada for giving me the opportunity to work on such an interesting project and for his good guidance and suggestions. A huge thanks goes to my co-supervisor James Catmore for helping me to navigate in all the different corners of the ATLAS software and for answering my many questions. I am also grateful to Eirik Gramstad for his help with the different samples and analysis codes. Thanks also to my fellow master students for good laughs and always having the time to play fussball, and to my neighbors on the master room, Tobias Opdalshei Baumann and Ruben Guevara for the many interesting conversations about physics and other topics. The rest of the HEP group also needs to be thanked for making a good social environment and for all the countless quizzes. The coffee machine also needs its credits for giving me a good and refreshing start on the day.

For inspiring me to study physics I need to thank my high school physics teacher Geir Arne Bosnes. I will also thank my best friends from Trondheim, Markus Valås Hagen, Magnus Mæhlum and Mikkel Aslak Sokki for all the fun times we have had and for your visits during the master. Last but not least, I will thank my parents for always supporting me and giving me good advice.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

At the Large Hadron Collider at CERN the data collected from the proton-proton (pp) collisions in the ATLAS detector is used to test Standard Model (SM) predictions in particle physics, and to search for new physics beyond the Standard Model (BSM). The SM has been studied with great precision, but we know that is not the final theory as both gravity and dark matter are not accommodated. We are interested in BSM models which predict new neutral gauge bosons, Z', which are present in several dark matter or grand unification models. Monte Carlo (MC) simulations of pp collisions in the ATLAS detector are compared with real data in order to test SM predictions and to search for new physics. The simulations of the SM and BSM processes are therefore important to make accurate predictions and for our understanding of the SM.

In this thesis we will study a method called signal reweighting to produce BSM Z' simulations from SM Drell-Yan simulations in a dilepton final state. Signal reweighting is a method that is much less computationally expensive than if we would simulate the Z' samples every for every mass point, and are therefore useful when studying a wide variety of Z' models depending on many parameters and each of which predict a different Z' mass. An tool used for applying signal reweighting to Pythia8 samples, called the LPXSignalReweightingTool, will be extended to include the signal reweighting of a PowhegPythia8 Drell-Yan sample. In addition an attempt to extend it to reweighting a Sherpa Z + jets sample will be made. The results of the signal reweighting with PowhegPythia8 will be compared to fully simulated Z' samples from Pythia8 in Run 2 and also include a comparison of data and MC in Run 3 with signal reweighted Z' samples.

The structure of the thesis is as follows. In the next chapter we will introduce the SM

and some BSM Z' models, and describe how these hypothetical particles are produced in pp collisions at the LHC. In Chapter 3 we describe how the outcome of a pp collision (called an event) is modeled with MC simulations. We describe the ATLAS detector in Chapter 4 together with samples that are used for the signal reweighting. In Chapter 5 the Z' signal reweighting method is described in detail and extended such that it can be used on the next-to-leading order (NLO) PowhegPythia8 simulation. The attempt to extend the signal reweighting to Sherpa Z + jets, allowing up to 5 jets in the final state, is described in Chapter 6. The results of the signal reweighting of PowhegPythia8 are presented in Chapter 7 where it is compared to fully simulated Z' Pythia8 samples. Finally, a first look at Run 3 data at 13.6 TeV is made. Data are compared to SM simulations, including Z' produced from the reweighting work presented in this thesis.

Chapter 2

The Standard Model and beyond

in pp collisions

2.1 The Standard Model

The Standard Model (SM) [1, 2, 3, 4, 5, 6] describes the particles and the interactions between the particles that we know of today, and is build up from the symmetry group

$$G_{SM} = SU(3)_C \times SU(2)_L \times U(1)_Y, \tag{2.1}$$

where every interaction is described by the different groups. The subscript C stands for color, L for left-handed and Y for hypercharge. It is build up of three forces, the electromagnetic force with Quantum electrodynamics (QED), the strong force with Quantum chromodynamics (QCD) described by the group $SU(3)_C$ and the weak force (W), and where the electromagnetic and weak force unifies to the electroweak force (EW) described by the product of groups $SU(2)_L \times U(1)_Y$. The SM does not describe the gravitational force. The SM has been successful in describing how all the known particles interact with great precision. Before we describe the particles and forces we will explain what are called *Feynman diagrams* that are both useful to visualize a process and to assist in the calculation of observables from the theory. We will use the Feynman diagrams actively when describing the SM.

2.1.1 Feynman diagrams

The SM is mathematically described with Quantum Field Theory (QFT) [7] where every particle has its own field, and the excitations of the fields are the particles. The kinematics of the fields are described by the Lagrangian density $\mathcal{L}$. In perturbation theory the Lagrangian can be expanded such that it can be formulated diagrammatically with what are called *Feynman diagrams*. An example of a Feynman diagram is shown Figure 2.1 where we read the diagram from left to right which is the direction of time[1]. In Figure 2.1 two particles, an electron (e^-) and position (e^+) represented as lines meet in a point called an interaction vertex *annihilation*. From this interaction a photon γ is produced and propagated to another point where an electron and positron are produced. This way of representing a process is both visual and practical since all particles and points have their own mathematical expressions called *Feynman rules* that are used to calculate observables, and which are derived from the Lagrangian $\mathcal{L}$. The Feynman diagrams consist of lines and nodes, where the lines denote the particles and the nodes denote the interactions. Every line and node represent a mathematical expression constructed using *Feynman rules*. The Feynman rules are used to calculate different observables of the theory such as the differential cross section of a process. The advantages of these Feynman diagrams is that the calculations in the theory become more systematic and that the Feynman diagrams visualize the process at the same time. We will therefore make active use of Feynman diagrams to describe the interactions in the SM. Before we go into detail about the interactions we will first review the different particles and interactions in the SM.

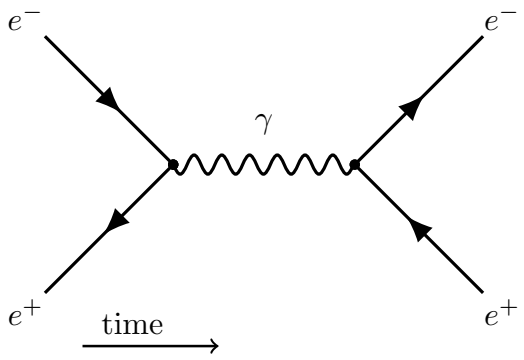


Figure 2.1: The Feynman diagram showing the process $e^-e^+ \to \gamma \to e^-e^+$ with the time going from left to right as indicated by the arrow.

[1]The time direction of every Feynman diagram in this thesis will be from left to right

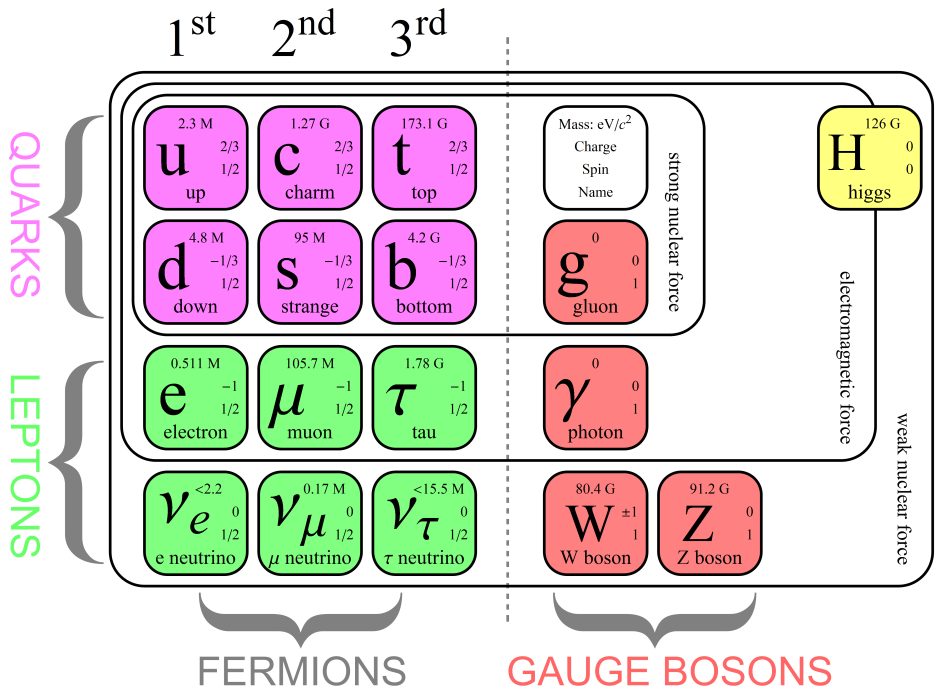


Figure 2.2: Table of the particles in the Standard Model [8].

2.1.2 Interactions and particles

In the SM there are two different kinds of particles called *fermions* and *bosons*. The fermions are particles that make up what we see around us, while the bosons are the mediators of the forces between the different particles. We will first describe the fermions before we go on to the bosons. Figure 2.2 shows all the fundamental particles that we know of today.

The fermions can be grouped together into particles known as *quarks* and *leptons*. There are six quarks in the SM called the up (u), down (d), charm (c), strange (s), top (t) and bottom (b) quark. The quarks comes in three generations as seen in Figure 2.2 and where each generation consists of an up- and down-type quark (not to be confused by the up and down quark). The up-type quarks have electric charge $+2/3$ while the down-type quarks have electrical charge $-1/3$. Quarks are not observed freely, but are grouped together with other quarks to make up what we call *hadrons*. The proton and neutron are examples of the hadrons and the proton consists on average of two up-quarks and a down-quark, while the neutron consists on average of two down-quarks an one up-quarks. The quarks are subject to the strong force, contrary to the leptons, which together with electrons build the atoms.

The leptons also comes in three generations, where each generation has one charged and one neutral particle. The charged particles are the electron (e^-) (1st generation), the muon (μ^-) (2nd generation) and the tau τ^- (third generation) and the corresponding neutral particles are the electron neutrino (ν_e), the muon neutrino (ν_μ) and the tau neutrino (ν_τ).

Every fermion has a corresponding anti-particle that has the opposite electric charge to its particle.

2.1.3 Quantum electrodynamics

Quantum electrodynamics (QED) is the force that acts between charged particles and is mediated by the photon which is the generator in the $U(1)$ group. There are no self interactions with the photon since it has no electric charge and the $U(1)$ group is abelian. The conserved quantity in QED is the electric charge. Figure 2.3 shows the annihilation of a fermion f with an anti-fermion $\overline{f}$ to the photon γ.

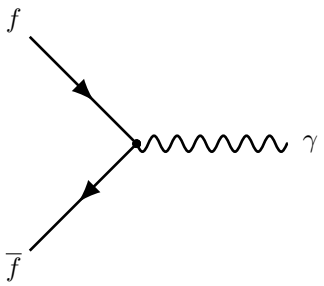
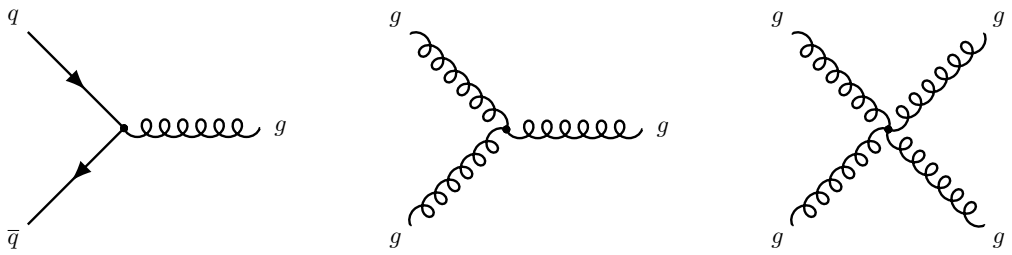


Figure 2.3: The QED interaction vertex

2.1.4 Quantum chromodynamics

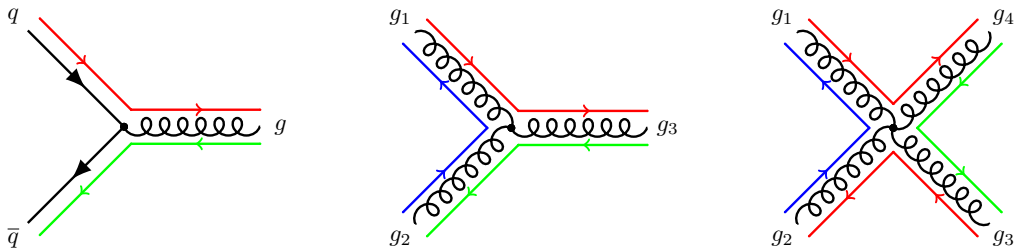
In Quantum chromodynamics (QCD) the force carriers are eight gluons corresponding to the generators of the $SU(3)_C$ group and are spin-1 bosons. The $SU(3)_C$ group describing QCD is a non-abelian, and as a consequence the gluons can interact with each other in so-called self-interactions. The interaction vertices in QCD are shown in Figure 2.4 where 2.4b and 2.4c are self-interactions, and Figure 2.4a is a quark anti-quark annihilation to a gluon.



(a) $q\bar{q} \to g$ (b) $gg \to g$ (c) $gg \to gg$

Figure 2.4: The QCD interaction vertices

In QCD the conserved quantity is the *color charge* where a quark carries a color c, an anti-quark carries an anti-color $\bar{c}$ and a gluon carries in general a combination of color-anticolor $c_1\bar{c}_2$. There are three different color charges called red (r), green (g) and blue (b) and the corresponding anti-color charges anti-red ($\bar{r}$), anti-green ($\bar{g}$) and anti-blue ($\bar{b}$). The color charges can be represented as colored lines on the sides of the quark flavour as seen in Figure 2.5.



(a) $q\bar{q} \to g$ (b) $g_1g_2 \to g_3$ (c) $g_1g_2 \to g_3g_4$

Figure 2.5: Color Feynman vertices

The colors charges for the particles in Figure 2.5a, 2.5b and 2.5c are shown in Eq. (2.2), (2.3) and (2.4) respectively.

$$\text{Figure 2.5}a: \quad q : r \quad \bar{q} : \bar{g} \quad g : r\bar{g} \tag{2.2}$$

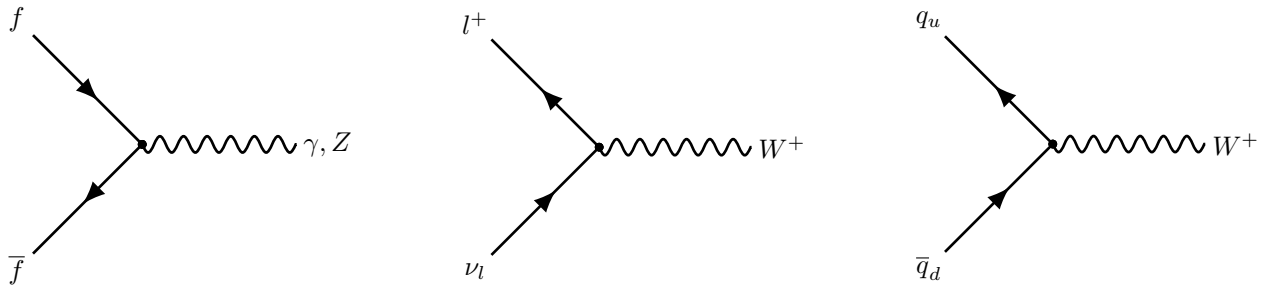
$$\text{Figure 2.5}b: \quad g_1 : r\bar{b} \quad g_2 : b\bar{g} \quad g_3 : r\bar{g} \tag{2.3}$$

$$\text{Figure 2.5}c: \quad g_1 : r\bar{b} \quad g_2 : b\bar{r} \quad g_3 : g\bar{r} \quad g_4 : r\bar{g} \tag{2.4}$$

2.1.5 The Electroweak interaction

In the electroweak theory there are four force carriers which are the neutral photon γ and the Z and the charged $W^{\pm}$ bosons, where the weak interactions are mediated through the Z and $W^{\pm}$ bosons . The conserved quantity in weak theory is the weak isospin. Figure

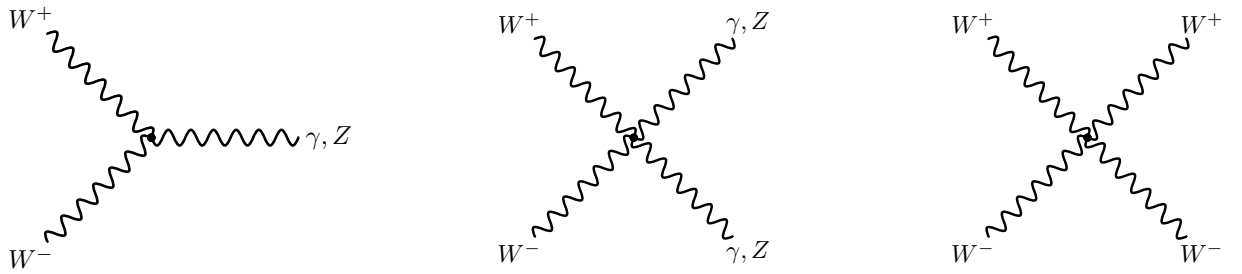
2.6 shows the different interactions in EW theory involving fermions. The photon γ and Z boson can only interact with charged particles as shown in Figure 2.6a. The W^+ and W^- bosons can interact with a neutrino and a lepton as shown in Figure 2.6b, and with quarks as shown in Figure 2.6c, where the subscript u and d on the quarks denote that they are a down- and up-type quark respectively.



(a) $f\overline{f} \to \gamma, Z$ (b) $l^+\nu_l \to W^+$ (c) $q_u\overline{q}_d \to W^+$

Figure 2.6: The interaction vertices involving fermions in EW theory.

The $SU(2)_L \times U(1)_Y$ group describing EW theory consist of the abelian group $U(1)_Y$ and the non-abelian group $SU(2)_L$ and we will consequently get some self-interaction vertices between the bosons in EW theory which are shown in Figure 2.7.



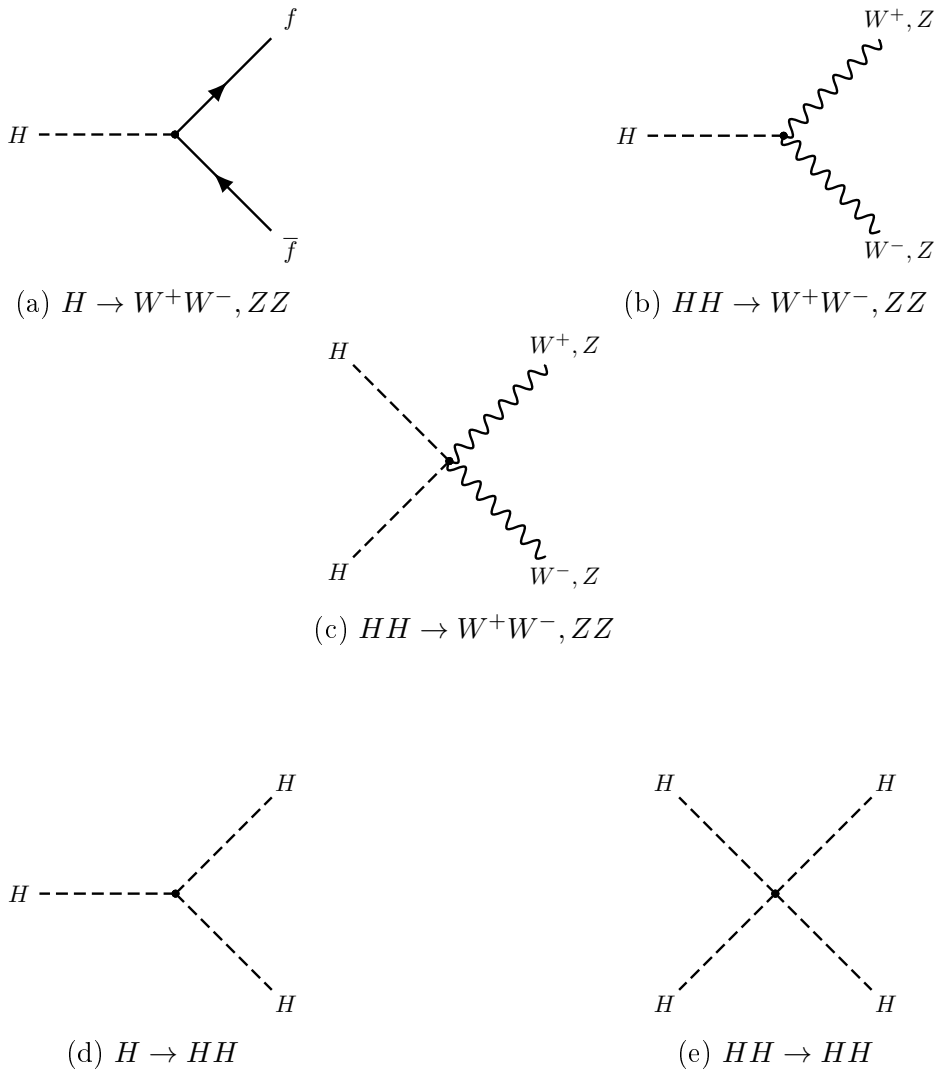
(a) $W^+W^- \to \gamma, Z$ (b) $W^+W^- \to \gamma\gamma, ZZ, \gamma Z$ (c) $W^+W^- \to W^+W^-$

Figure 2.7: The self-interaction vertices in EW theory.

2.1.6 The Higgs mechanism

The $SU(2)_L \times U(1)_Y$ group spontaneously brake down to the $U(1)_{QED}$ group in what is known as the Brout-Englert-Higgs mechanism [5, 6] which is responsible for giving mass to the fermions and bosons in the SM. The Higgs particle was discovered in 2012 by ATLAS [9] and CMS [10] and was the final member of the SM to be experimentally observed. The weak interaction vertices involving the Higgs boson are shown in Figure

2.8.



(a) $H \to W^+W^-, ZZ$

(b) $HH \to W^+W^-, ZZ$

(c) $HH \to W^+W^-, ZZ$

(d) $H \to HH$

(e) $HH \to HH$

Figure 2.8: Interaction vertices involving the Higgs boson

2.2 Beyond Standard Model - Z' models

There are several phenomena that the SM cannot explain. It does not describe the gravitational interaction, and is unable to account for dark matter in a way which is consistent with all current observations. There exist many so-called Beyond the Standard Model (BSM) theories that aim to solve one or several problems with the SM. One BSM extension to the SM adds new neutral massive spin-1 gauge bosons called Z' bosons. Many BSM models have the Z' boson included in their extension such as grand unified theories (GUT) and some dark matter models, where the Z' plays a role of the mediator. We will study the Sequential Standard Model (SSM) Z' model and the E_6' Z' models

[11, 12, 13] in this thesis. There have been several searches [14, 15, 16, 17, 18] and studies [19, 20, 21, 22, 23] of the Z' boson at the Large Hadron Collider in general and ATLAS in particular. Before we study these models we will study how a general Z' model is build up by describing its Lagrangian $\mathcal{L}$. When adding a Z' boson to the SM the G_{SM} symmetry group from Eq. (2.1) is extended by adding a $U(1)$ group, and the symmetry with a Z' extension therefore becomes

$$SU(3)_c \times SU(2)_L \times U(1)_Y \rightarrow SU(3)_c \times SU(2)_L \times U(1)_Y \times U'(1). \tag{2.5}$$

The interactions between a Z' boson and fermions in a neutral current can be summarized in the Lagrangian

$$\mathcal{L}_{Z'} = g_{Z'}\overline{f}\gamma^{\mu}(z_{f_L}P_L + z_{f_R}P_R)fZ'_{\mu} = g_{Z'}J^{\mu}_i Z'_{\mu}, \tag{2.6}$$

which describes how the Z'_{μ} interacts with the fermion fields f and $\overline{f}$. The coupling constant $g_{Z'}$ describes the strength while the operators P_L and P_R project the left- and right-handed part of a field and are called chiral projection operators. The constants z_{f_L} and z_{f_R} are called the left- and right-handed chiral charges of the fermions. If $z_{f_L} \neq z_{f_R}$ we call it a *chiral* theory. The Z' boson interactions with fermions are described by the values of the coupling constant $g_{Z'}$ and the left- and right-handed chiral charges z_{f_L} and z_{f_L} to all the SM fermions.

2.2.1 The Sequential Standard Model

The first model we will describe is the Sequential Standard Model (SSM). Here the Z' is similar to the SM Z boson. In the SSM the coupling constant $g_{Z'}$ and left and right handed chiral charges z_{L_f} and z_{R_f} describing the theory are the same as those of the Z boson in the SM and are given as

$$g_Z = \frac{e}{\sin\theta_W \cos\theta_W} \tag{2.7}$$

$$z_{f_L} = t_f^3 - q_f \sin^2\theta_W \tag{2.8}$$

$$z_{f_R} = -q_f \sin^2\theta_W, \tag{2.9}$$

where θ_W is the weak mixing angle, t_f^3 is the 3rd component of the weak isospin and q_f the charge of the fermion. The only difference between a SSM Z' and the SM Z boson is the mass of the Z' boson, $m_{Z'}$ that can be different from the Z boson mass $m_Z = 91.1876$ GeV. The SSM model is called a toy model since it is not renormalizable when coupling to SM fermions. It can be made a renormalizable theory if it couples to exotic fermions. The width of the Z_{SSM} boson as compared to its mass is assumed to be the same as for the Z boson and is approximately 3% of its mass.

2.2.2 The E_6' model

The E_6-inspired Z' models [11, 12, 13, 24] are motivated from the unification of all the forces and is therefore called a Grand Unified model. The forces are unified at higher scales in the E_6 group which comes from a superstring theory in 10 dimensions formulated in the $E_8 \times E_8$ group. One of the E_8 groups breaks down to the E_6 group that we will schematically study the symmetry breaking with and shown in Figure 2.9. In the first symmetry breaking the E_6 group breaks down to $SO(10) \times U(1)_\psi$, where the $U(1)_\psi$ and gives the Z'_ψ boson.

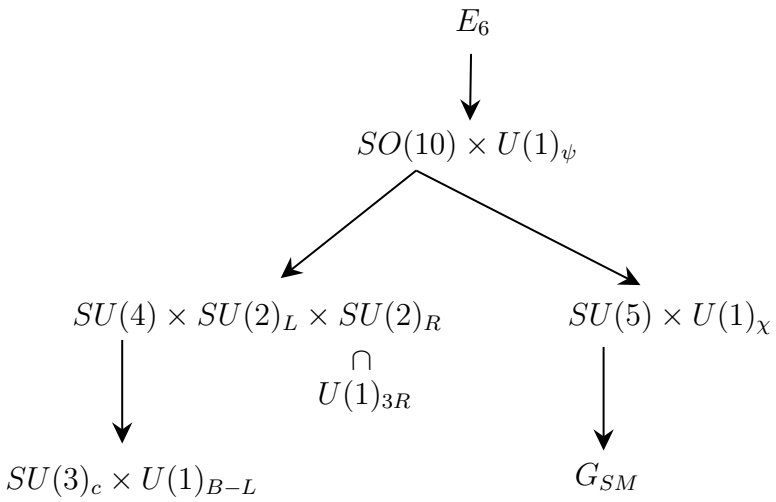


Figure 2.9: Schematic diagram showing the decomposition of the E_6 group, where G_{SM} is the SM group

The $SO(10)$ group can then break down in two different ways. In the first branch to the right in Figure 2.9 it breaks down to $SU(5) \times U(1)_\chi$ and where the $SU(5)$ group breaks down to the SM group G_{SM} in Eq. (2.1). The extension to the SM group in this branch leads to two new Z' bosons, Z'_ψ and Z'_χ. These two Z' bosons can at lower

energies mix via a mixing angle θ_{E_6} and give one $Z'(\theta_{E_6})$ boson given as

$$Z'(\theta_{E_6}) = Z'_\psi \cos\theta_{E_6} + Z'_\chi \sin\theta_{E_6}. \tag{2.10}$$

The parameter describing these models are [24]

$$g_{Z'} = \sqrt{\frac{5}{3}} g_Z \sin\theta_W = \sqrt{\frac{5}{3}} g \tan\theta_W \tag{2.11}$$

$$z_{f_L} = \frac{q^\psi_{f_L}}{2\sqrt{6}} \cos\theta_{E_6} - \frac{q^\chi_{f_L}}{2\sqrt{10}} \sin\theta_{E_6} \tag{2.12}$$

$$z_{f_R} = \frac{q^\psi_{f_R}}{2\sqrt{6}} \cos\theta_{E_6} - \frac{q^\chi_{f_R}}{2\sqrt{10}} \sin\theta_{E_6}, \tag{2.13}$$

where $q^\psi_{f_{L,R}}$ and $q^\chi_{f_{L,R}}$ gives the fermionic charges under $U(1)_\psi$ and $U(1)_\chi$. The mixing angle θ_{E_6} is a free parameter of the model and can therefore be varied together with the Z' boson mass. Table 2.1 the values of θ_{E_6} in the some of the most traditional models together with corresponding values of $\sin(\theta_{E_6})$ and $\cos(\theta_{E_6})$. Z'_ψ and Z'_χ corresponds to $\theta_{E_6} = 0$ and $\pi/2$ respectively.

Model	θ_{E_6}	$\sin\theta_{E_6}$	$\cos\theta_{E_6}$
Z'_S	0.6293π	$\frac{3\sqrt{6}}{8}$	$-\frac{\sqrt{10}}{8}$
Z'_I	0.7098π	$\sqrt{\frac{5}{8}}$	$-\sqrt{\frac{3}{8}}$
Z'_N	-0.0804π	$-\frac{1}{4}$	$\frac{\sqrt{15}}{4}$
Z'_ψ	0	0	1
Z'_η	0.2098π	$\sqrt{\frac{3}{8}}$	$\sqrt{\frac{5}{8}}$
Z'_χ	0.5π	1	0

Table 2.1: Different values for θ_{E_6} for some traditional Z' models, including the special cases of Z'_ψ and Z'_χ from Figure 2.9.

The relative width of the Z'_ψ an Z'_χ bosons are 0.5% and 1.2% of their mass respectively.

2.2.3 Minimal Z' models

We will now describe the breakdown in the other branch to the left of Figure 2.9 that gives what we call *minimal models* and are described by a small number of parameters. In this branch the $SO(10)$ group breaks first down to the $SU(4) \times SU(2)_L \times SU(2)_R$

group, where $U(1)_{3R}$ is a subgroup of $SU(2)_R$ and $3R$ is the third component of the right-handed weak isospin. The Z'_{3R} boson that comes from this symmetry breaking is an example of a model where the left-right symmetry violation in the SM gets restored. Finally the $SU(4)$ group breaks down to $SU(3)_C \times U(1)_{B-L}$ where $B-L$ is the conserved quantum number describing the difference between the baryon (B) and lepton (L). These two different models are examples of minimal models which describe the model with the fewest possible parameters. These parameters are the Z' mass and two effective coupling constants. The two effective coupling constants are g_{BL}, the coupling constant to the Z'_{B-L} boson and g_Y the coupling constant to the the weak hypercharge Y. Instead of using g_{BL} and g_Y we use the ratio of them to the SM Z coupling constant g_Z and we write them as

$$\tilde{g}_{B-L} = \frac{g_{BL}}{g_Z} \tag{2.14}$$

$$\tilde{g}_Y = \frac{g_Y}{g_Z}. \tag{2.15}$$

It is useful to reparameterize the $\tilde{g}_{B-L}$ and $\tilde{g}_Y$ in terms of two independent parameters γ', which is the strength between the Z' and the SM Z boson, and θ_{Min} that is the mixing angle between the generators of the $B-L$ and the weak hypercharge Y gauge group. We can then rewrite the coupling constants in Eq. (2.14) and (2.15) as

$$\tilde{g}_{B-L} = \gamma' \cos\theta_{Min} \tag{2.16}$$

$$\tilde{g}_Y = \gamma' \sin\theta_{Min}. \tag{2.17}$$

Table 2.2 shows the values for these two parameters in some minimal models. We observe that the Z'_χ from the first branch is a minimal model.

Model	γ'	$\sin\theta_{Min}$	$\cos\theta_{Min}$
Z'_{B-L}	$\sqrt{\frac{5}{8}}\sin\theta_W$	1	0
Z'_χ	$\sqrt{\frac{41}{24}}\sin\theta_W$	$\sqrt{\frac{25}{41}}$	$-\sqrt{\frac{16}{41}}$
Z'_{3R}	$\sqrt{\frac{25}{12}}\sin\theta_W$	$\sqrt{\frac{1}{5}}$	$-\sqrt{\frac{4}{5}}$

Table 2.2: Different values for γ' and θ_{Min} for some Minimal Z' models

We have now studied the E_6 and SSM Z' model in some detail. In the next section

we will introduce the kinematics of a collision between protons.

2.3 Z' production in pp collision

In this section we will introduce the needed theory to study pp collisions in the ATLAS detector. We will first describe the relevant variables used to describe pp collisions before describing the kinematics of a particle.

2.3.1 Z' production in pp collisions

Now we will explain important variables when studying pp collision where the protons collide in bunches with many particles. The number of collisions that can be detected over a given time is

$$N = \sigma \int \mathcal{L}(t)dt, \tag{2.18}$$

where σ is the total cross section and $\mathcal{L}(t)$ is the instantaneous luminosity that is given as

$$\mathcal{L} = \frac{n_1 n_2}{4\pi\sigma_x\sigma_y} f \tag{2.19}$$

where $\sigma_{x,y}$ is the beam size, f is the frequency of bunch crossings and $n_{1,2}$ is the number of particles in the two colliding bunches of particles. The cross section describes the probability that an interaction can happen and is given as

$$\sigma = \int \frac{d\sigma}{d\Omega} d\Omega, \tag{2.20}$$

where $\frac{d\sigma}{d\Omega}$ is the differential cross section that is given as

$$\frac{d\sigma}{d\Omega} = \frac{1}{F}\frac{dN}{d\Omega}. \tag{2.21}$$

The differential cross section describes the N number of particles that are scattered in a solid angle $d\Omega$ per unit time per unit flux F.

The picture of the proton is more complex than what we described in Section 2.1.2. In addition to the *valence quarks* which are the two up-quarks and the down-quark, the proton consist of other quarks called *sea quarks* and other gluons, and we call them all *partons*. The interacting partons each carry a fraction of the momentum of the proton.

The probability that a parton of given flavour carries a fraction x of the proton is described by the Parton Distribution Functions (PDFs). The PDFs for two different values of the momentum transfer is shown in Figure 2.10.

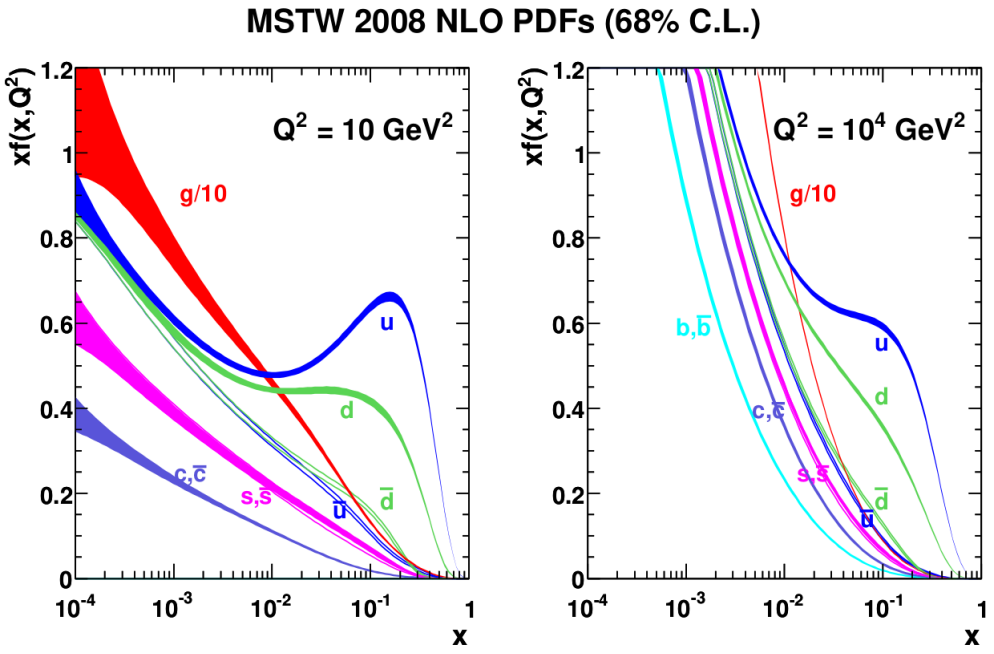


Figure 2.10: The PDFs for the partons in the proton for two values of the momentum transfer Q. The x-axis shows the fraction x while the y-axis shows the product $xf(x, Q^2)$ of the fraction x and the density function for a parton carrying the momentum [25].

The Z' production in a pp collision is shown in the process in Figure 2.11 with two incoming protons with momentum p_1 and p_2 respectively. The Z' boson is produced from the anhilation of a quark and anti-quark from the protons and decays to a di-lepton pair l^+l^-. The hadrons formed from the quarks that do not take part in the Z' production are denoted X.

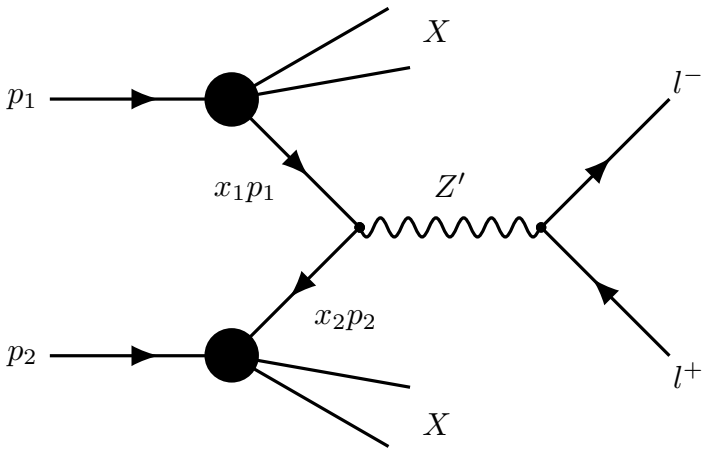


Figure 2.11: The Drell-Yan process of a Z' production decaying to a di-leptpn pair from a quark and anti-quark from the protons p_1 and p_2 respectively. The fraction of the momentum of the quark and anti-quark from the protons are x_1 and x_2 respectively. The quarks not taking part in the Z' production form hadrons that are denoted X.

The differential cross section of the Drell-Yan process with respect to the fractions of momentum x_1 and x_2 of the protons is [7]

$$\frac{d\sigma(p(p_1)+p(p_2)\to Z'\to l^+l^-+X)}{dx_1dx_2}=\sum_i \sigma(q_i\overline{q}_i\to Z'\to l^+l^-+X)q_i(x_1)\overline{q}_i(x_2), \quad (2.22)$$

where sum i is over the quark flavours and σ is the cross section of the s-channel process $q_i\overline{q}_i\to Z'\to l^+l^-+X$. The PDFs of the quark and anti-quark with flavour i are $q_i(x_1)$ and $\overline{q}_i(x_2)$ respectively. We will now shortly describe the kinematics of a particle.

2.3.2 Kinematics of a of particle

The kinematic of a particle is defined by the energy E, momentum $\vec{p}=(p_x,p_y,p_z)$ and the mass of the particle. The energy and momentum of a relativistic particle can in special relativity [26] be written as

$$E=\gamma mc^2 \quad \text{and} \quad \vec{p}=\gamma m\vec{\beta}, \quad (2.23)$$

where c is the speed of light and γ is the Lorentz factor given as

$$\gamma=\frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}}. \quad (2.24)$$

We will use natural units from now on where we set the speed of light c and Plank's constant $\hbar$ equal to one, and use GeV as units for the energy, momentum and the mass of a particle. The energy and momentum can be combined into what is called the four-vector

$$p^{\mu} = \big(E, p_x, p_y, p_z\big), \tag{2.25}$$

which when contracted with itself gives

$$p^2 = p^{\mu}p_{\mu} = E^2 - p^2 = m^2, \tag{2.26}$$

This is a Lorentz invariant quantity and is conserved. The component of the momentum that is in the xy-plane is called the *transverse momentum* and is defined as

$$p_T = \sqrt{p_x^2 + p_y^2} = p\sin\theta, \tag{2.27}$$

where θ is the polar angle that goes around the z-axis as shown in Figure 2.12 and p is the magnitude of the momentum. By using the transverse momentum together with Eq. (2.26) we can define the *transverse energy* as

$$E_T = \sqrt{m^2 + p_T^2}. \tag{2.28}$$

Now that we have described the kinematics we can continue to study the collision of two particles in the next section.

2.3.3 Collision of two particles

We will now describe the collision of two particles in the center of mass (CoM) frame where the total momentum vector of the two particles equals zero. If the two particles travel along the z-axis as shown in Figure 2.12 in opposite directions, their four-momentum in CoM becomes

$$p_1^{\mu} = (E_1, 0, 0, p) \quad \text{and} \quad p_2^{\mu} = (E_2, 0, 0, -p), \tag{2.29}$$

and the sum of their four-momenta becomes the initial four-momentum $P^\mu = (E_1 + E_2, 0, 0, 0)$. The squared of the initial four-momentum becomes

$$P^2 = P^\mu P_\mu = (E_1 + E_2)^2 \equiv s, \tag{2.30}$$

where $s = (p_1 + p_2)^2$ is a Mandelstam variable and is the square of the CoM energy $\sqrt{s}$. In the LHC, since the colliding particles are hadrons, the collision involves interactions amongst the constituent quarks and gluons and so many particles may emerge from each collision. If for example a Z boson is produced in the collision, and it decays to a di-lepton pair, we can reconstruct the mass of the Z boson from the invariant mass of the two leptons l_1 and l_2 it decays to, that is defined as

$$m_{l_1 l_2} = \sqrt{(E_{l_1} + E_{l_2})^2 - (\vec{p}_{l_1} + \vec{p}_{l_2})^2}. \tag{2.31}$$

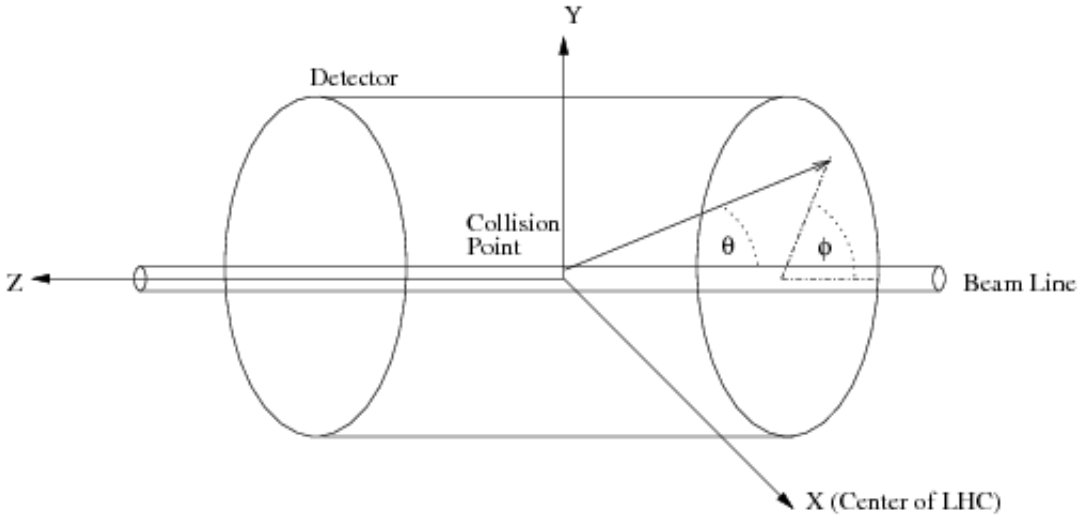


Figure 2.12: View of the geometry of a cylindrical detector where two particles travel in opposite directions and collide along the z-axis [27]

We have now studied how proton collisions happen and will in the next chapter see how pp collisions are studied at the Large Hadron Collider and in the ATLAS detector. Before we study pp collisions in the ATLAS detector we will see how pp collisions are generated in simulations.

Chapter 3

Event generators

Particle collisions in high energy physics are modelled using Monte Carlo event generators. These software packages simulate on a statistical basis the behaviours of particles predicted by theoretical models. The output of event generators can be studied on their own or can further be simulated in a "digital twin" of a detector such as ATLAS. The simulation of the pp collisions for different processes and the different stages of the collision evolution will be explained in Section 3.1. The event generators that will be important for this thesis are Pythia8, Powheg and Sherpa, and will be shortly described in Section 3.2. The outcome of the event generation is stored in what is called an event record and we will study the HepMC event record in Section 3.3. This will be important for the signal reweighting in Section 5.

3.1 Event generation

The generation of an event is often done in a time-ordered way which is equivalent to the transverse momentum ordering. This means that the processes which have the outgoing particles with the highest transverse momentum are modelled first followed by the processes with lower transverse momentum. Another word for this ordering is the *hardness* of the collision which typically is quantified by the transverse momentum of the particles. The time-ordering does not necessary go in the direction from past to future, but follows the hardness of the processes. The time-ordering generally follows how well the physics is understood in the process. We will now take a brief look at the different stages in the time-ordering of the event generation based on the collision in Figure 3.1

from Pythia8 that shows the process $pp \to t\bar{t}$. The hardness scale in Figure 3.1 starts in the middle of the circle where the hardest scattering happens and then the hardness of the processes decreases radially outwards. We will therefore describe the hardest scattering first followed by the other stages in the time-ordered way.

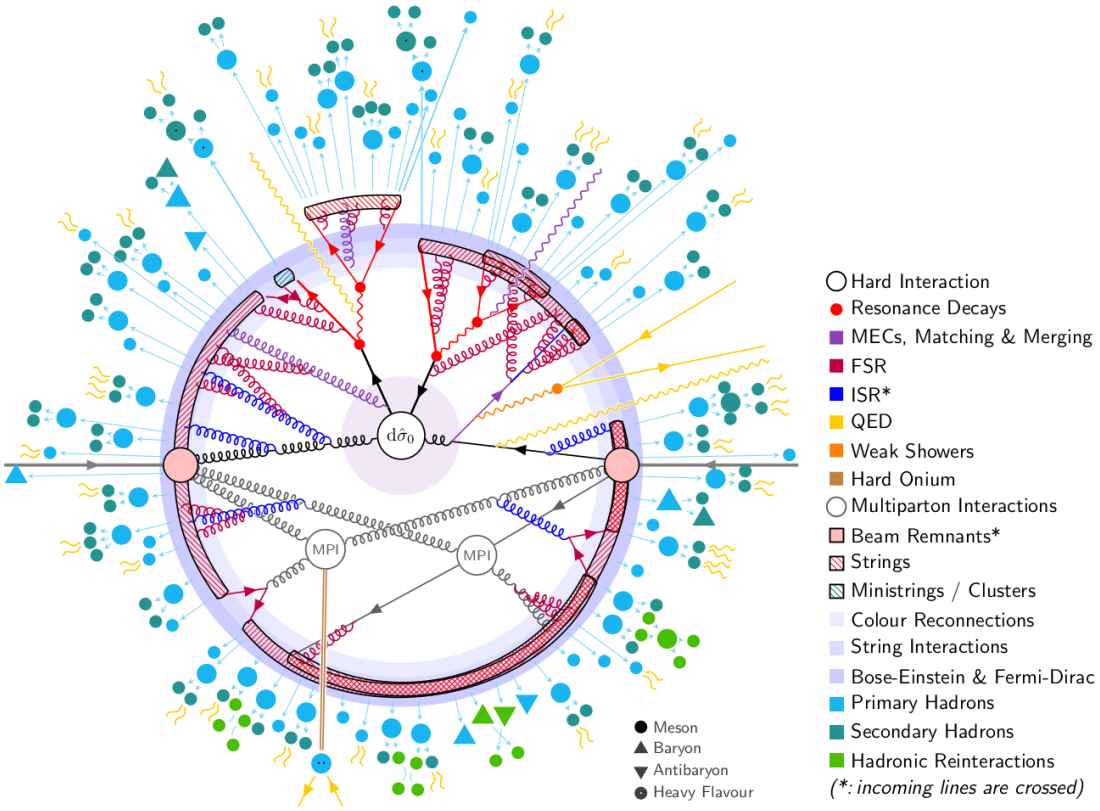


Figure 3.1: Schematic diagram of the time-ordering of a $pp \to t\bar{t}$ process from Pythia8[28].

1. In the *hardest scattering* one parton from each proton collides and produces a few outgoing particles. The incoming partons are chosen based on the parton distribution functions of the protons. The kinematics of the outgoing partons are calculated from the matrix element of the hard scattering in perturbation theory. The hardest scattering process in Figure 3.1 is $pp \to t\bar{t} + X$, where X are particles that coms from the parton showering. The hard scattering is shown in the circle which surrounds $d\hat{o}_0$ in the middle of the big circle. The top and anti-top quarks are outgoing particles from the hard scattering and are shown as thick black lines with arrows and are connected to the hardest scattering.

2. One or several particles from the hardest scattering can be a short-lived resonance like the Z or $W^{\pm}$ boson or top quarks. These resonances will then decay to other

more stable particles. In Figure 3.1 the top and anti-top quark are unstable and their decays are shown as red dots that are connected to the top and anti-top quarks.

3. *Fixed-order corrections* can be included in the matrix-element corrections in QCD or EW in perturbation theory. Most commonly the hard scattering is calculated in leading order (LO) or next-to-leading (NLO) in EW and QCD.

4. The incoming partons to the hardest scattering can radiate other partons and photons and is known as *initial-state radiation* (ISR). In Figure 3.1 the ISR is shown with the blue lines. Due to the Fermi motion of the particles inside the proton, the incoming particles also gets some transverse momentum added because of this and is called *primordial* k_T.

5. The outgoing particles and resonances from the hardest scattering can also radiate other particles and is known as *final-state radiation* (FSR). Figure 3.1 shows the FSI as red lines.

6. In addition to the ISR and FSR, multiple partons from the beam remnants can further scatter in what are known as *Multiple Parton Interactions* (MPI). Figure 3.1 shows the MPI as the circles that surrounds MPI.

7. *Color dipoles* begins to form after the MPI and resonance decays stage. These color dipoles are defined by color connections in the leading-color approximation, i.e the $N_c \to \infty$ limit. The color connections are shown in the innermost shaded circle.

8. The strong interaction confines the QCD partons into color-singlet known as *strings*, or *clusters* in the small-mass limiting cases. The leftover partons from the beams are also combined into beam remnants. In Figure 3.1 the strings are shown in the the striped arcs. The beam remnants are shown as the two pink circles lying on opposite side of the big circle. The partonic and hadronic parts of the schematic diagram in Figure 3.1 are separated by the shaded blue rings before the blue and green circles and triangles.

9. The strings then fragment into hadrons and are shown as the blue circles for the mesons or triangles for the baryons and are called the *primary hadrons* in Figure 3.1.

10. When multiple identical particles are close in phase space this leads to increased Bose-Einstein effects and a suppression of Fermi-Dirac effects that is shown in the outermost circle.

11. The primary hardrons from the fragmentation can decay to *secondary hadrons* and are shown in the dark green circles or triangles in Figure 3.1.

12. The primary and secondary hadrons can in densely populated regions re-scatter, re-annhilate, and/or recombine with each other to produce other hadrons. Figure 3.1 shows these *hadronic reinteractions* in the light green circles.

3.2 Event generators

A number of different event generators are used in High Energy Physics to generate events that can then be simulated in a detector model. Here we will only talk about the most common event generators used in ATLAS.

3.2.1 Pythia8

`PYTHIA8`[28] is a general purpose Monte Carlo event generator that produces high-energy collisions between electrons, protons, photons and nuclei. It models all the steps that were described in Section 3.2 from the hard scattering until the fragmentation and decays of the final state hadrons. The hard scattering can be calculated at both leading-order and next-to-leading order in perturbation theory [29]. However, the NLO matrix element can also be merged into Pythia8 in the hardest scattering with either MC@NLO[30] or `POWHEG`. Pythia8 has over 200 scattering processes from both SM and BSM.

3.2.2 Powheg

The `POWHEG`[31, 32, 33] method is a general framework for implementing NLO QCD and EW calculations into parton showers. POWHEG is short for Positive Weight Hardest Emission Generator and only generates the hardest radiation using the NLO matrix elements. The output form the hardest radiation is then passed on to other Shower Monte Carlo programs where the showering is done. The events are produced with a method

that produces positive event weights in NLO, whereas other NLO event generators (such as Sherpa) produce a large fraction of negatively weighted events.

3.2.3 Sherpa

`Sherpa`[34] is another general purpose Monte Carlo event generator that generates high-energy collision events in hadron-hadron, lepton-hadron and lepton-lepton colliders. Sherpa is organised as different modules that are collected in a framework. Sherpa can simulate multi-jet events in final state, and as rarer and rarer processes are persued, accurately modelling such backgrounds is becoming increasingly important to accurately describe the complexity of events at the LHC. The NLO merging with Powheg fail to describe the multi-jet final state [35]. Sherpa has therefore developed its own technique for multi-jet merging at NLO called MEPS@NLO. In NLO Sherpa generation there is however a large fraction of negative weights that comes from effects when calculating the differential cross section at higher order, but there exist methods to reduce the amount of negative weights[36].

3.3 HepMC - Event record

The output information from the event generation is stored in what is called the event record. Every event generator has its own internal event record. HepMC[37, 38, 39] is a general event record that is independent of any event generator and experiment and is commonly used in the High Energy Physics community. We will describe how the HepMC event record is structured and introduce some of the features that will be important when discussing signal reweighting and color flows in for the Chapter 5 and Chapter 6.

3.3.1 Event record

In particle physics a collision process is typically visualized as a diagram that shows the particles that interact in different parts of the process and is illustrated in the left diagram in Figure 3.2. In HepMC the collision process is represented in an event record which connects event vertices where the interactions happen in a graph structure as shown in the right graph in Figure 3.2. A vertex lists the incoming and outgoing particles from an interaction. In Figure 3.2 the incoming and outgoing particles are represented with edges,

and the interaction is represented as nodes. A vertex will generally not correspond to an interaction vertex in a Feynman diagram, but must rather be thought of as a collection of different processes that transform the incoming particles into the outgoing particles.

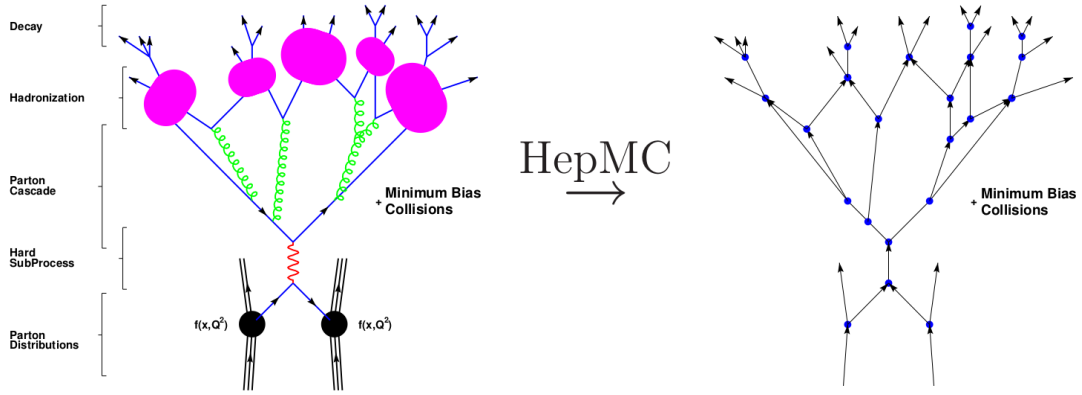


Figure 3.2: Visualization of a collision process (left) and the HepMC representation of it in the event record (right) from [38].

Figure 3.3 shows an example of an event vertex from the HepMC event record in Pythia8. The event vertex shows an incoming proton from a pp collision and a collection of outgoing quarks and gluons. We will study this event vertex in more detail to learn how to read the information in an event vertex.

Barcode	PDG ID	(Px, Py, Pz, E)	Stat	DecayVtx
TruthVertex:	-310	ID: 0 (X,cT)= 0		
I: 1	1	2212 +0.00e+00,+0.00e+00,+6.50e+06,+6.50e+06	4	-310
O: 11	399	21 -1.54e+03,-1.49e+03,+2.04e+04,+2.05e+04	61	-248
	407	21 +3.74e+02,-5.72e+02,+3.49e+03,+3.56e+03	61	-20
	417	21 -2.98e+03,-1.18e+03,+7.02e+04,+7.02e+04	61	-95
	446	2 +2.28e+02,+3.31e+03,+4.40e+05,+4.40e+05	61	-290
	467	-1 +1.16e+03,-5.91e+02,+1.91e+03,+2.31e+03	61	-271
	478	1 +2.07e+01,-9.09e+02,+1.41e+05,+1.41e+05	61	-172
	488	21 +1.39e+03,+4.50e+00,+7.64e+05,+7.64e+05	61	-178
	506	21 +4.36e+02,-5.23e+01,+2.90e+06,+2.90e+06	61	-71
	510	2103 +2.97e+02,+5.91e+02,+2.06e+06,+2.06e+06	63	-437
	511	1 +6.29e+02,+6.14e+02,+1.17e+04,+1.18e+04	63	-494
	512	-1 -1.81e+00,+2.75e+02,+2.08e+03,+2.13e+03	63	-452

Figure 3.3: Example of a Pythia8 event vertex with a unique vertex barcode (purple box) with 1 incoming particle and 11 outgoing particles (blue box). Each row represent a particle and has a unique barcode (orange box) and the particle type (red box) is represented with PDG ID codes (see Table 3.1). The status codes (yellow box) gives more information about which stage the particle is in (see Table 3.2). The kinematics of the particles (green box) is of the form (p_x, p_y, p_z, E) in units of GeV. Finally, the barcode for the event vertex where the particle decays is provided (brown box).

The colored boxes in Figure 3.3 shows the different types of information in the event

vertex. The number of incoming (I) and outgoing (O) is shown in the blue box. Every event vertex has its own unique barcode as shown in the purple box. The rows in Figure 3.3 each represents a particle where the orange box shows the unique particle barcode. The particle type is identified by what is called the PDG ID number that follows the Particle Data Group numbering scheme[40] and is shown in the red box. Table 3.1 shows some examples of the PDG ID codes for some common particles. The kinematics for every particle is given in the green box where the four columns show the four-momentum in the form (p_x, p_y, p_z, E) in units of GeV. The numbers in the yellow box are used to identify which stage of the event generation process the particle is in and is called a status code. Table 3.2 shows some relevant status codes that will be used in Chapter 5. The information about where we can find the particle next in the event record is shown in the brown box which gives the vertex barcode where that particle decays.

Quarks			
d	1	Leptons	
u	2	e^-	11
s	3	μ^-	13
c	4	τ^-	15
b	5		
t	6	Gauge bosons	
Diquarks		g	21
$(ud)_1$	2103	γ	22
Hadrons		Z	23
		W^+	24
p	2212		

Table 3.1: Relevant PDG ID codes for different particles [40].
Anti-particles have negative numbers

As an example of how to use the event record with all this information we can look in the first particle in the event record in Figure 3.3. There we see a proton (PDG ID 2212) which is incoming from the beam (Stat 4) and decays in the event vertex -310 which is the barcode of the vertex itself since it decays there. If we look at the kinematics of the proton we also see that it comes from the beam since it has an energy of 6.5 TeV, which is exactly half of the energy of the LHC CoM energy of $\sqrt{s} = 13$ TeV. It has no momentum in x- or y-direction. We will use the event record actively in Chapter 5 to trace particles in an event through different processes.

1-9 :	stage of event generation
1 :	a final-state particle
4 :	an incoming beam particle
21 - 29 :	particles of the hardest subprocess
21 :	incoming
22 :	intermediate (intended to have preserved mass)
23 :	outgoing
41 - 49 :	particles produced by initial-state-showers
44 :	outgoing shifted by a branching
51 - 59 :	particles produced by final-state-showers
51 :	outgoing produced by parton branching
61 - 69 :	particles produced by beam-remnant treatment
61 :	incoming subprocess particle with primordial k_T included
62 :	outgoing subprocess particle with primordial k_T included
63 :	outgoing beam remnant

Table 3.2: Relevant status codes for different processes[41].

Other kinds of information from the event generators can also be included in the event record like the color flow information which will be further described in the next section.

3.3.2 Color flow

In the event generation the color charges for every quark, anti-quark and gluon is stored in the event record. In HepMC this information is stored in color flows in the Flow class [42]. A color flow contains all the particles that have the same color. The event generators use the leading color approximation, where the number of colors goes towards infinity, $N_c \to \infty$. This means that every color flow in a diagram will have an unique label. We will illustrate how the color flows are stored in HepMC with the s-channel diagram in Figure 3.4.

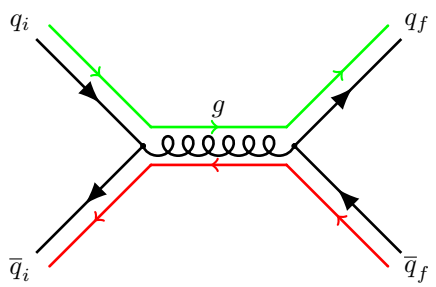


Figure 3.4: Feynman diagram of process $q_i\overline{q}_i \to g \to q_f\overline{q}_f$ with the color charges are denoted on the side of the particle lines.

In this diagram we have two different color flows labeled *red* and *green*, which would be labeled with integers in the HepMC Flow class. The quarks carries color, the anti-quark carries anti-color and the gluon carries both carries color and anti-color as explained in Section 2.1.4. The color flow contains the particles that have the same color label and are ordered in the direction of the color. The color flows red and green are therefore

$$red : \overline{q}_f \rightarrow g \rightarrow \overline{q}_i \tag{3.1}$$

$$green : q_i \rightarrow g \rightarrow q_f. \tag{3.2}$$

We will come back to the color flow in Section 6.1 and Chapter 6 where we will use the color flow information in an attempt to do signal reweighting with Sherpa.

Chapter 4

Z' production in pp collisions

Capturing and reconstructing the fragments emerging from the Large Hadron Collider (LHC) collisions require large and complex detectors which are constructed around each collision point. ATLAS is the largest of the LHC detectors. In this chapter we will describe how it works and how different particles interact with it. We will then briefly describe how the generated events from the last chapter are simulated in the detector. We will also describe the simulated samples that are of particular importance in this thesis: Pythia8 Z' boson production in pp collisions with the Z' decaying to a dilepton pair, $q\bar{q} \to Z' \to l^+l^-$; and the Drell-Yan sample $q\bar{q} \to \gamma, Z \to l^+l^-$ that will be used in the signal reweighting and is simulated with PowhegPythia8. Both these MC samples will be described and corrected with higher order corrections in Section 4.3. The event selection for the electrons and muons are described in Section 4.4.

4.1 The ATLAS detector

The ATLAS detector is a multipurpose particle detector and is one of the LHC detectors at CERN. The LHC accelerates protons to an energy of 6.5 TeV in opposite directions in a 27 km ring located $\sim$ 100 m underground. The two proton beams collide at four points around the ring where the detectors ATLAS, CMS, ALICE and LHCb are located. The ATLAS detector has a cylindrical symmetry in both the forward and backward direction relative to the LHC beam axis. The ATLAS detector uses the right handed coordinate system with origin in the interaction point (IP) and is shown in Figure 2.12. The z-axis is in the direction of the beam while the x-axis is orthogonal to the beam at the

IP and points towards the center of the LHC ring. The y-axis is both orthogonal to the x- and y-axis and points in the upward direction from the IP. The transverse plane with origin in the IP is spanned by the x- and y-axis and is described in cylindrical coordinates with the azimuthal angle ϕ and radius r in addition to the polar angle θ around the z-axis. In particle physics we often use the pseudorapidity η instead of the polar angle θ: as $\eta = -\ln(\tan(\theta/2))$. Finally the angular distance ΔR is defined as $\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$.

Figure 4.1 shows a schematic view of the different subdetectors in the ATLAS detector. The innermost area around the beam pipe is called the inner detector (ID) and consists of tracking detectors which are restricted to $|\eta| < 2.5$. The ID consist of three layers of detectors with the pixel detectors closest to the beam surrounded by semiconductior tracker and the transition radiation tracker as the outermost detector. Outside the ID there is a superconducting solenoid which provides a 2 T axial magnetic field. The electromagnetic (EM) and hadronic calorimeters surround the solenoid and are restricted to $|\eta| < 4.9$. Generally speaking only muons and neutrinos are able to penetrate all the layers of the calorimeter, so an external muon spectrometer (MS) forms the outermost layers of the detector. It covers $|\eta| < 2.7$. The MS consists of detectors that can activate the trigger when a muon is detected, and can also precisely track the trajectories of muons, in addition to the tracking of the muon in the ID. Three toroidal magnets are used to determine the momentum of the muons in their magnetic field. One of the toroidal magnets surrounds the center of the detector and consist of eight coils and is called the barrel toroidal. The other two are at each of the ends and of the detector and are called the end-cap toroids and consist of eight coils each. The triggering in the MS is used to determine which events that are kept for analysis based on some criteria to determine if the event is interesting to keep or not.

Particles that are detected in the ATLAS detector interact differently in the different subdetectors. Figure 4.2 shows the schematic of a portion of the transverse plane of the detector and illustrates how different particles are identified. The inner detector measures the trajectory of electrically charged particles such electrons, muons, protons, charged kaons and pions. They leave bent tracks from their interaction with the magnetic field from the superconducting solenoid. The momentum and the charge of the particles can be determined from the curvature of the tracks. Particles that interact electromagnetically

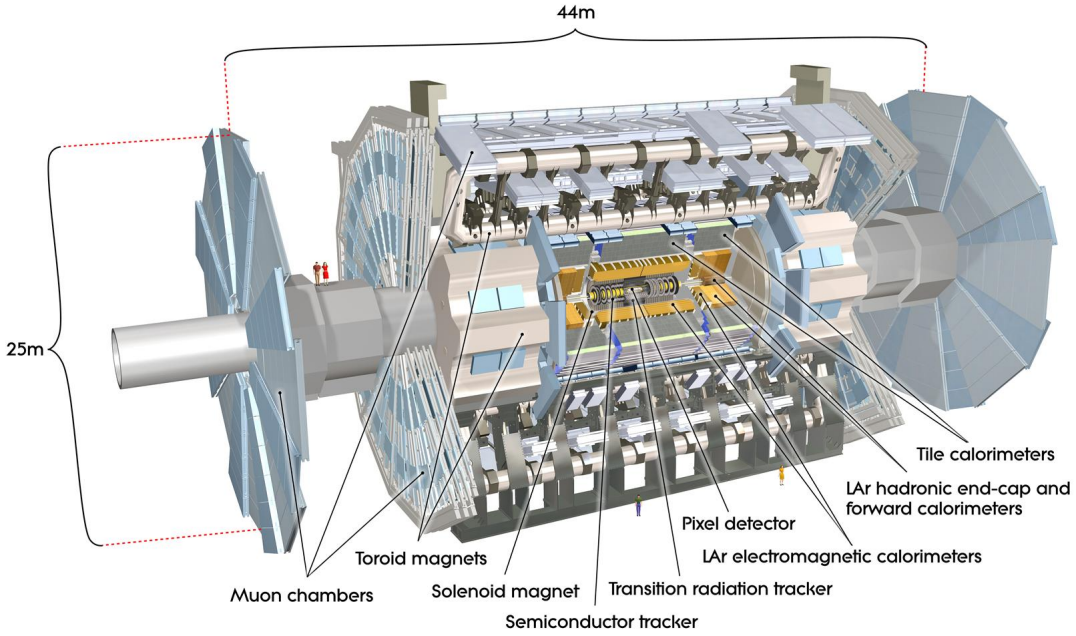


Figure 4.1: Schematic view of the ATLAS detector with the different subdetectors[43]

are detected in the EM calorimeter. The electron and the photon are stopped in the EM calorimeter through EM showers and their energy gets absorbed and measured. Protons and muons loose little energy in the EM calorimeter. The hadronic calorimeter is designed to measure the energy of hadrons such as protons, neutrons, kaons and pions. Such particles produce showers in the hadron calorimeter and their energy is absorbed and measured. It is particularly important for reconstructing particle jets which are associated with high momentum quarks and gluons. Muons and neutrinos are not stopped in the ATLAS detector. The momenta of the muons is however determined from the curvature of the tracks they leave behind in the inner detector and in the muon spectrometer. The neutrinos are not detected in any of the subdetectors in ATLAS since they interact weakly with matter. Since overall each collision event must have net zero transverse momentum, the presence of undetected particles with significant energy can be readily inferred as “missing energy".

The LHC operates in different periods called runs and undergoes upgrades between them. During Run 2 from 2015 to 2018 the LHC produced collisions at the CoM energy $\sqrt{s} = 13$ TeV and the ATLAS detector collected 139 fb^{-1} of data. The protons in the LHC ring are not in a continuous beam, but are bunched together with empty gaps between the bunches. In each bunch there were $1.1 \cdot 10^{11}$ protons in Run 2 and the time

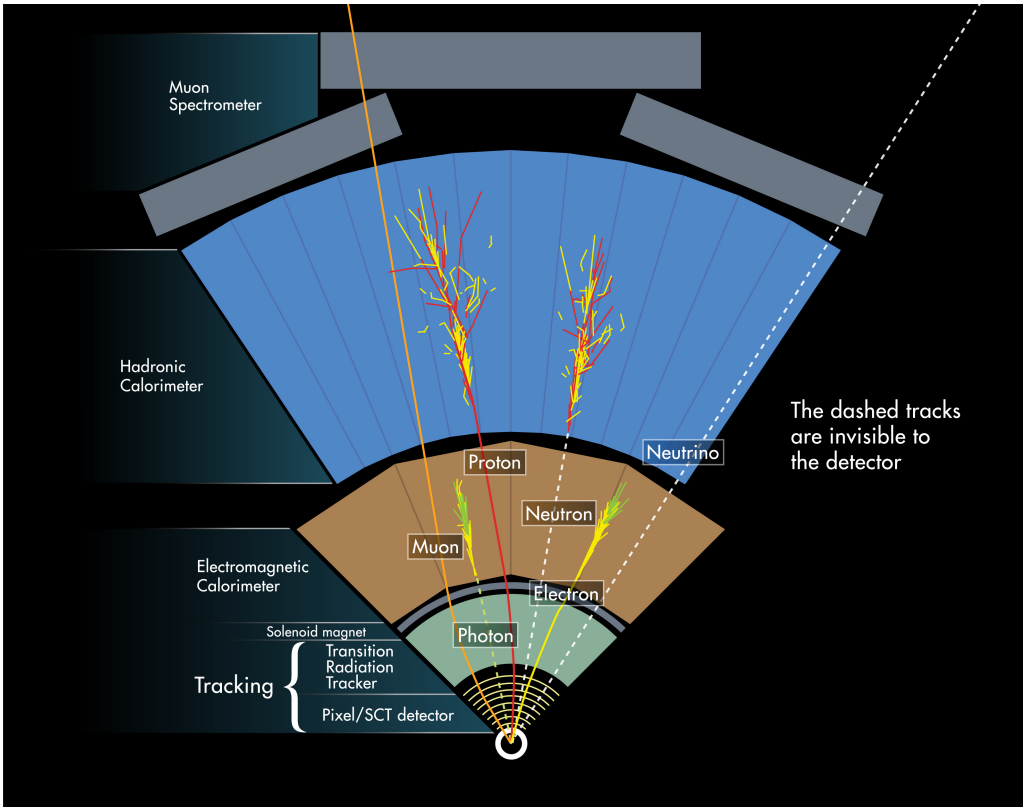


Figure 4.2: Schematic view of the transverse plane of the ATLAS detector that illustrates how different particles interact with the ATLAS detector[43]

between each bunch collision was 25 ns where on average 30 collisions occur in every bunch. In other words there were 40 millions collisions happening each second. It is not possible to store all this data so the triggering described above is needed to filter out the interesting events.

4.2 Event simulation

The interaction of particles generated in the event generators in the detector can be simulated in a “digital twin" of the detector with GEANT4[44]. The simulated energy deposits can then be “digitized" leading to a data format which looks like the digital output of the real detector. These can then be reconstructed in the same way as the real data, producing events that can be analysed in exactly the same way as real data, but which, unlike real data, have known provenance and associated ground truth. This enables Standard Model predictions from the event generators to be compared with collected data.

4.2.0.1 Detector simulation

The generated events are stored in a HepMC file with the truth information of the particles containing the history of the interactions in the generation. Since many generated particles will be outside of the fiducial volume of the detector, some cuts are done to the particle before they are simulated [45]. Not all stable particles are simulated in the detector. Stable particles in the detector are defined as particles that have a life time $c\tau > 10$mm. The generated events are then run through detector simulations in ATLAS. The detector simulation is done with GEANT4 [44] that simulates the geometry of ATLAS. This is the most computationally expensive part of the simulations and take about 80 % of the total CPU time. There does also exist fast simulations [46] that is less computationally expensive, but does not simulate the detector output as accurately. The output from the detector simulation is stored in a “hits" file that is a collection of the truth information and the collection of hits in each subdetector. A hit is an energy deposit in a subdetector that is stored with position and time. The next step is to convert these hits in the detector to something that mimic the output from collected data.

4.2.0.2 Digitization

The hits file from the detector simulation is the input into the ATLAS digitization software that converts the hits into detector responses called *digits*. This procedure is done to mimic what happens in the detector for collected data. A digit is produced when the voltage or current output in a readout channel rises above a predefined threshold over a time window. In some subdetectors the digit is described with the shape over some time-window, while in others it is only recorded if a threshold is exceeded within a time window.

4.3 Data and Monte-Carlo samples

In this section we will describe the data and simulations that will be used for the signal reweighting, and the processes that contribute a dilepton final state. We use only MC simulations from the $mc16e$ period in Run 2 since it is the only period we have fully simulated Z' Pythia8 samples. The MC samples are simulated at $\sqrt{s} = 13$ TeV and are scaled to the integrated luminosity 58.5 fb^{-1} that was collected during the $mc16e$ period.

For Run 3 we use both data and simulations from 2022 where 26.1fb^{-1} of data that is used for analysis was collected at $\sqrt{s} = 13.6$ TeV. The MC samples from Run 3 are therfore scaled to 26.1fb^{-1}.

4.3.1 Simulated Z' samples used in Run 2

In a dilepton final state the Z' production process is shown in leading order in Figure 4.3. It shows the process where a quark q and anti-quark $\overline{q}$ from the protons annihilate to a Z' boson that decays to a lepton pair l^-l^+.

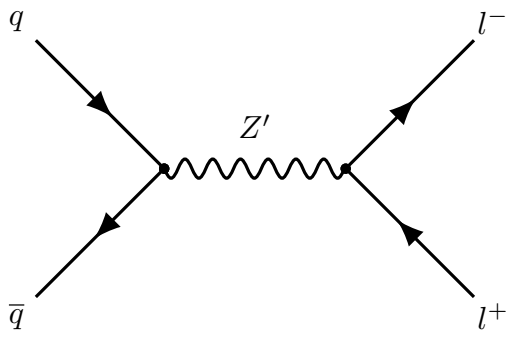


Figure 4.3: Z' production process

The Z' signal simulated samples we will use to compare to the signal reweighted samples obtained using the procedure in Chapter 5 are LO Pythia8[47] Z' samples which has the signal process in Figure 4.3. The only available Z' samples are based on the Z'_χ model where a 3 TeV-Z' leads to e^+e^- and $\mu^+\mu^-$ final states from the 2018 $mc16e$ period. Table A.3 in Appendix A.2 describes the samples in more detail.

4.3.2 Samples used for Z' signal reweighting in Run 2 and 3

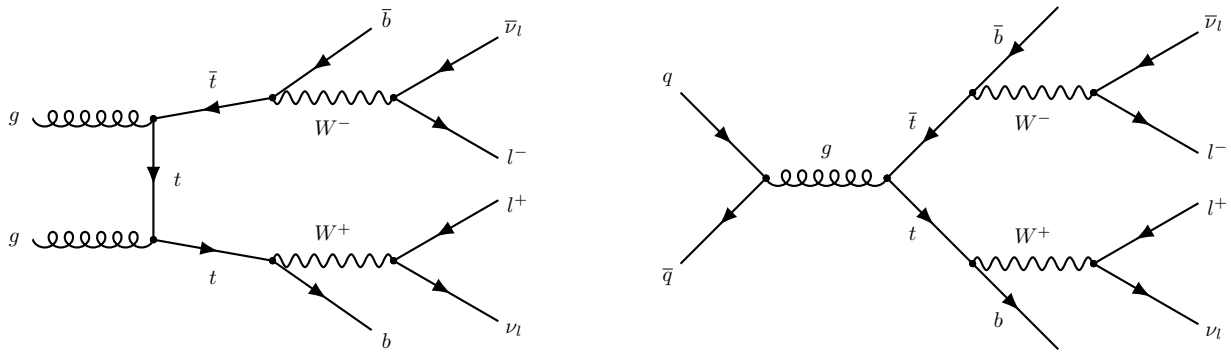
In the next chapter we will do signal reweighting of a Drell-Yan (γ^*, Z^*) background sample to a Z' signal sample. The Drell-Yan sample we will use is a fully PowhegPythia8[48] simulated sample where the matrix element is generated in NLO by Powheg and the parton showering is done by Pythia8. We will study the signal process of the PowhegPythia8 Drell-Yan sample in Section 5.3.1. Tables A.1 and A.2 in Appendix A.1 describe the samples from Run 2 in more detail. A fully simulated PowhegPythia8 Drell-Yan sample will also be used for signal reweighting in Run 3.

4.3.3 Background samples used in Run 3

In addition to the Drell-Yan background sample there are several other background processes in a di-lepton final state that we will shortly describe here. The other background processes are diboson, single top, $t\bar{t}$ and W + jets. The diboson and W + jets samples are simulated with Sherpa while single top and $t\bar{t}$ are simulated with PowhegPythia8.

4.3.3.1 Top

The production of a top-antitop ($t\bar{t}$) pair can produce a di-lepton pair in the final state through a decay from two W bosons as shown in Figure 4.4. The top quark almost always decays to a W boson and bottom quark. The two W bosons then decays leptonically and gives a di-lepton pair in final state. Figure 4.4a shows the $t\bar{t}$ production from two gluons while Figure 4.4b shows a gluon decaying to the $t\bar{t}$.

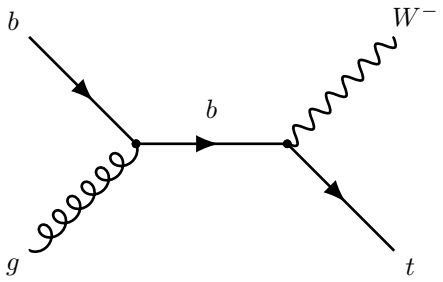


(a) Production of $t\bar{t}$ from two gluons.

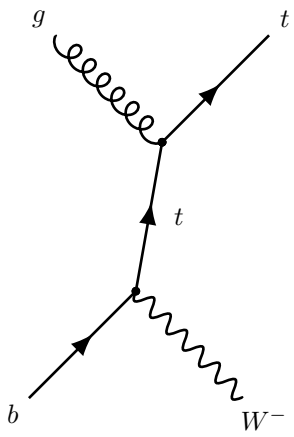
(b) Production of $t\bar{t}$ from the decay of a gluon.

Figure 4.4: Production of $t\bar{t}$ from two gluons (a) and the decay of a gluon (b).

The production of a top/antitop quark together with a W boson can also give a di-leption in the final state as shown in Figure 4.5 and 4.6 and are called single top processes. In Figure 4.5 the top quark is produced in two diagrams involving a gluon, while in Figure 4.6 there are no gluons involved in the top production. The top quark in all these diagrams decay to a W^+ boson which decays leptonically together with the W^- boson that is not shown in the diagrams.

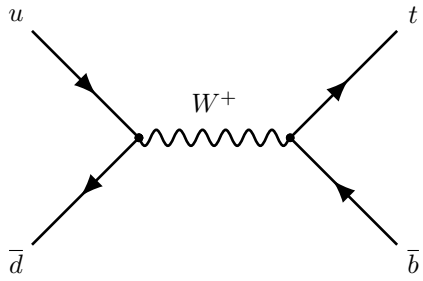


(a) The single top production in a s-channel diagram involving a gluon.

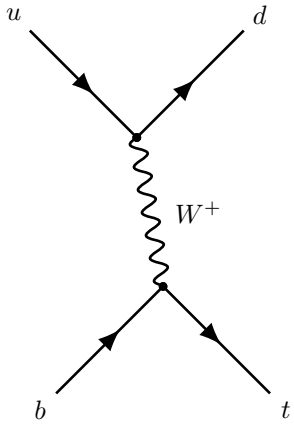


(b) The single top production in a t-channel diagram involving a gluon.

Figure 4.5: The single top production involving a gluon in a s-channel (a) and t-channel (b).



(a) The single top production in a s-channel diagram.

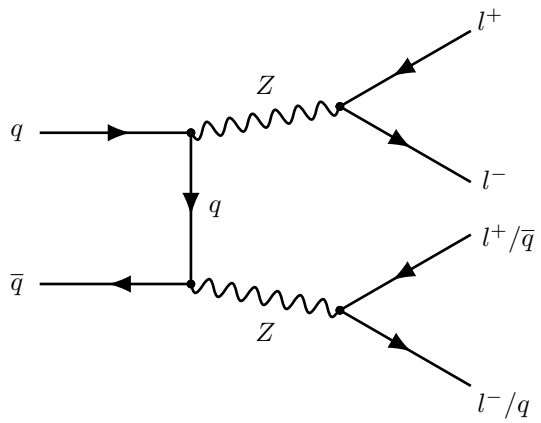


(b) The single top production in a t-channel diagram.

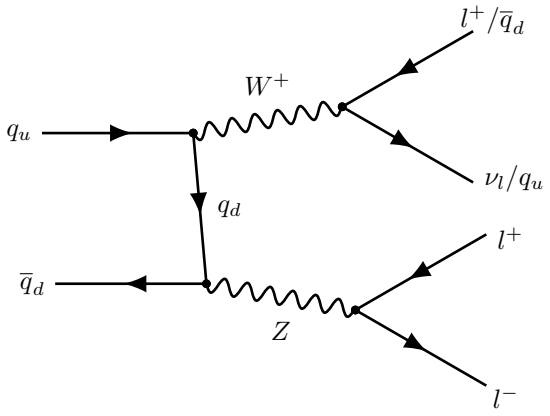
Figure 4.6: The single top production involving in a s-channel (a) and t-channel (b).

4.3.3.2 Diboson

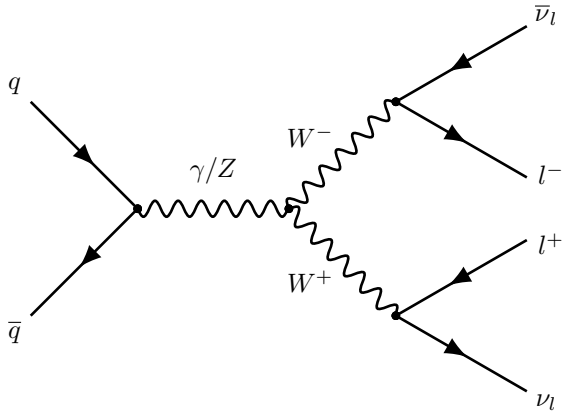
A production of two bosons can decay to a di-lepton pair as shown in Figure 4.7. Figure 4.7a shows the production of two Z bosons, 4.7b the production of a Z boson and W^+ boson and 4.7c the production of a W^+ and W^- boson. At least one Z boson must decay leptonically in Figure 4.7a, while the Z boson in Figure 4.7b must decay leptonically and both W^+ and W^- must decay leptonically in Figure 4.7c. The subscript u and d on the quarks means that the quark/antiquark is an up- or down-type quark.



(a) Production of two Z bosons where at least one of the boson decay leptonically.



(b) Production of a Z boson and W^+ boson where the Z boson must decay leptonically.



(c) Production of a W^+ boson and W^- boson where both bosons decay leptonically.

Figure 4.7: Production of diboson showing ZZ (a), ZW^+ (b) and W^+W^- (c).

4.3.3.3 W + jets

The last background we will look at is W + jets which can be described by the same diagrams in Figure 4.5 and 4.6, but where the top quark does not decay to a W boson. Then there is only one W boson that decays leptonically, giving one lepton. The quark can however be misidentified as a lepton in the detector where it makes a jet. The jet is then called a fake lepton.

We will now describe how higher order corrections can be applied to an already generated sample which we will make use of in Chapter 7.

4.3.4 Higher order corrections

Higher order corrections can also be applied to the MC samples after the event generation. To correct to higher orders than provided by the event generators we apply k-factors that we need for the Run 2 MC samples in a comparison. A k-factor is the ratio of the differential cross section in the signal process calculated in the order we want to correct to, divided by the differential cross section for the signal process in the sample. The k-factor is an event-by-event scale factor. The k-factors we will use are calculated from the LPXKfactorTool[49]. We need to apply k-factors to the LO Pythia8 Z' and the NLO PowhegPythia8 Drell-Yan samples so that we compare them in equal order when we do signal reweighting. We apply a $LO \rightarrow NNLO$ QCD k-factor to the LO Pythia8 Z' samples and a $NLO \rightarrow NNLO$ QCD k-factor to the NLO PowhegPythia8 Drell-Yan sample. There do not exist any EW k-factors for the simulated Z' sample. Since both Pythia8 Z' and the PowhegPythia8 Drell-Yan samples are simulated at LO in EW we will not apply any EW k-factors to the PowhegPythia8 Drell-Yan samples either.

4.4 Event selection in Run 2

Most of the event selection criteria for the electron and muon candidates in Run 2 in the final state are taken from [14] and are described below.

The electron candidates reconstructed from tracks in the ID and must deposit energy in EM calorimeters. They must have $E_T > 30$ GeV and $|\eta| < 2.47$ in order to pass through the fine-granularity region of the EM calorimeters. In addition the electron candidates in the region between the barrel and the end cap of the EM calorimeter corresponding to the region $1.37 < |\eta| < 1.52$ are excluded. The candidates should also pass the 'medium' electron working point which has an identification and reconstruction efficiency for electrons above 92 % for $E_T > 80$ GeV.

The muon candidates are reconstructed by matching tracks in the ID to tracks reconstructed in the MS. The muon candidate must have $p_T > 30$ GeV and $|\eta| < 2.5$. The muon candidates must also pass the 'high p_T' identification working point. It ensures optimal resolution for muons with high p_T and is specified by at least 3 hits in each of the three layers of precision tracking chambers in the MS.

Some cuts are common for both electron and muon candidates. One of them is that the longitudinal impact parameter z_0 is required to satisfy $|z_0 \sin\theta| < 0.5$ mm. In addition the electron (muon) candidate must have a significance of the transverse parameter that satisfies $|d_0/\sigma(d_0)| < 5(3)$. The electrons should pass the 'gradient' isolation working point. For the muons the isolation requirement is that the summed scalar p_T inside a cone or size $\Delta R = 0.3$ should be less than 6 % of the p_T of the muon candidate.

In the final state the candidates in the pair needs to have the same flavor. However an electron pair is not required to have opposite charge because of the probability of charge misidentification for high-E_T electrons. For muons the opposite charge requirement must be satisfied.

The event selection for 3 and 4 leptons in the final state will not be considered in this study, due to the minimal number of such events. The final requirement for the events considered in this study is that the invariant mass of the pair should satisfy the invariant mass criteria, $m_{ll} > 120$ GeV, slightly above the Z-peak.

4.5 Event selection in the electron channel in Run 3

The event selection criteria for the electron candidates in the Run 3 final state are almost the same as for Run 2. The electron candidates reconstructed from tracks in the ID and must deposit energy in EM calorimeters. They must have $p_T > 30$ GeV and $|\eta| < 2.47$ in order to pass through the fine-granularity region of the EM calorimeters. In addition the electron candidates in the region between the barrel and the end cap of the EM calorimeter corresponding to the region $1.37 < |\eta| < 1.52$ are excluded. The longitudinal impact parameter z_0 is required to satisfy $|z_0 \sin\theta| < 0.5$ mm. In addition the electron candidate must have a significance of the transverse parameter that satisfies $|d_0/\sigma(d_0)| < 5$. An electron pair is not required to have opposite charge because of the probability of charge misidentification for high-p_T electrons. The final requirement for the events considered in this study is that the invariant mass of the electron pair should satisfy the invariant mass criteria, $m_{ee} > 180$ GeV, slightly above the Z-peak.

Chapter 5

Z' Signal reweighting with Powheg

In this chapter we will extend the LPXSignalReweightingTool [24] to apply signal reweighting to a PowhegPythia8 Drell-Yan di-lepton invariant mass distribution rather than the Pythia8 Drell-Yan distribution described in Section 5.1. The motivation for doing signal reweighting on PowhegPythia8 is because there are no available ATLAS Pythia8 Drell-Yan samples available for Run 2 and 3. Another reason is that the hard process is calculated in next-to-leading order in QCD for PowhegPythia8 while it is calculated at leading order for Pythia8. The PowhegPythia8 Drell-Yan distribution is therefore more accurately modelled than the Pythia8 Drell-Yan distribution. The truth information from the HepMC event record will be important for extending the LPXSignalReweightingTool [24]. In the two first sections we will explain how the signal reweighting is done for Pythia8 and how we need to change the LPXSignalReweightingTool to apply signal reweighting to a PowhegPythia8 sample.

5.1 LPXSignalReweightingTool

The LPXSignalReweightingTool is a software tool within the ATLAS software that uses the truth information from the HepMC event record to calculate per-event weights to transform a LO Pythia8 Drell-Yan sample to a BSM Z' sample. This is known as signal reweighting and avoids having to produce and fully reconstruct many new MC samples (many models, masses, ...), which is computationally expensive. The scale factor is calculated at truth level but is applied to the reconstructed sample. We do therefore not need to simulate any new signal samples other than the Drell-Yan sample that is

anyway needed for background modelling. The LPXSignalReweigtingTool was originally written to reweight LO Pythia8 Drell-Yan as input sample. In Chapter 5 we will extend the LPXSignalReweightingTool to also reweight NLO PowhegPythia8 Drell-Yan samples, and describe the attempt that was made to use NLO Sherpa Z + jets in the next chapter. We will therefore describe the LPXSignalReweingTool in greater detail there, and only describe the basics of the tool here.

5.1.1 Z' reweighting scale factor

The scale factor that is used to reweight a Drell-Yan distribution to a Drell-Yan + Z' distribution is

$$SF_{SM+BSM} = \frac{\frac{d\sigma}{d\hat{t}}\left(q\overline{q} \rightarrow \gamma, Z, Z' \rightarrow l^{-}l^{+}\right)}{\frac{d\sigma}{d\hat{t}}\left(q\overline{q} \rightarrow \gamma, Z \rightarrow l^{-}l^{+}\right)}, \tag{5.1}$$

and is the ratio of the differential cross section of the s-channel diagram in Figure 5.1, where where $\hat{t}$ is the Mandelstam [7] variable defined as $\hat{t} = (q - l^{-})^2 = (\overline{q} - l^{+})^2$. The denominator in Eq. (5.1) is the differential cross section for the SM process $q\overline{q} \rightarrow \gamma, Z \rightarrow l^{-}l^{+}$, while the numerator is the differential cross section for the SM + BSM process $q\overline{q} \rightarrow \gamma, Z \rightarrow l^{-}l^{+}$. The scale factor SF_{SM+BSM} is applied to every reconstructed event in a histogram like e.g the invariant mass distribution.

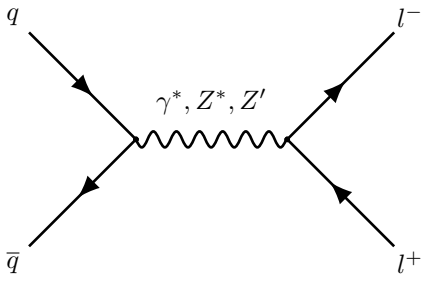


Figure 5.1: The Drell-Yan s-channel diagram that is used to calculate the scale factor per event where either the photon, Z boson or Z' boson is the propagator.

The scale factor that is used to reweight a Drell-Yan distribution to only a Z' distribution is

$$SF_{BSM} = \frac{\frac{d\sigma}{d\hat{t}}\left(q\overline{q} \rightarrow Z' \rightarrow l^{-}l^{+}\right)}{\frac{d\sigma}{d\hat{t}}\left(q\overline{q} \rightarrow \gamma, Z \rightarrow l^{-}l^{+}\right)}, \tag{5.2}$$

where the photon and Z boson are not included in the differential cross section in the numerator. We can use these different modes in Eq. (5.1) and (5.2) depending on what we want to reweight to. The scale factor in Eq. (5.1) and (5.2) needs the kinematics of

the incoming quarks q and $\bar{q}$, the intermediate particles γ, Z and Z' and the final state leptons l^- and l^+. We also need to know the couplings to the Z' model. In the next section we will go throught how this information can be found. We will further work on this in Section 5.1.2 in more detail.

5.1.2 Finding the information for the LPXSignalReweighitng-Tool

We need to specify the Z' model we are reweighting to in the scale factor in Eq. (5.1) or (5.2). The LPXSignalReweighitngTool supports various E_6 inspired and the SSM Z' models for now. It can however be extended to any Z' model when we know the coupling constant $g_{Z'}$ and the left and right handed chiral charges z_{f_L} and z_{f_R} to the fermions from the Lagrangian in Eq. (2.6). Table 2.1 lists the values of the coupling constant $g_{Z'}$ and the chiral charges for the E_6 model.

We also need to specify the kinematics and flavours for the incoming quarks $q\bar{q}$, the intermediate γ or Z boson and the final state leptons l^-l^+ from the s-channel process in Figure 5.1. This information is needed to calculate the scale factor. We get this information from the HepMC event record of the LO Drell-Yan Pythia8 sample. More precisely, we get it by looping over the different event vertices that follow the different stages in the event generation from Section 3.1. In the next chapter we will see how this is done in detail in order to adjust the LPXSignalReweightingTool to do signal reweighting of NLO PowhegPythia8 Drell-Yan samples. First, let us give some details about the LPXKfactorTool that allows taking into account higher order corrections.

5.1.3 LPXKfactorTool

The LPXKfactorTool [49] is used to account for higher order corrections for Monte Carlo simulations with a scale factor. These scale factors are called k-factors and are the ratio between the higher order cross sections compared to the cross section at the order that the sample is generated at. As an example a k-factor can correct a LO Pythia8 sample to a NLO sample. The LPXKfactorTool uses the invariant mass of the final state lepton to calculate a scale factors. The k-factor is therefore not constant, but is a function of the invariant mass. [49]

5.2 Pythia8

In Section 5.1 we described in general how the scale factor (Eq. (5.2)) from the LPXSignalReweightingTool is used to signal reweight a Pythia8 Drell-Yan distribution to a Z' distribution. We will now go in more detail on how the scale factor is calculated from the event record of a Pythia8 sample. This section will therefore help us to extend the LPXSignalReweightingTool to apply Z' signal reweighting to a PowhegPythia8 Drell-Yan distribution in Section 5.3. The structure of this section will closely follow the time-ordering in the event generator described in Section 3.1, starting from the hard scattering. When we analyze the event record we make use of the PDG ID codes and the status codes summarized in Table 3.1 and 3.2, respectively.

5.2.1 Signal process

The hard subprocess, also known as the signal process, is where we find the quarks and the Z boson needed to calculate the scale factor (Eq. (5.2)) in the LPXSignalReweigthingTool. An example of a signal process vertex is shown in Figure 5.2 where a quark-antiquark pair (PDG ID 1 and -1) annihilates into a Z boson (PDG ID 23).

```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:       -3 ID:    0 (X,cT)= 0
 I:  2         3         1 +0.00e+00,+0.00e+00,+1.29e+04,+1.29e+04  21        -3
               4        -1 +0.00e+00,+0.00e+00,-1.49e+05,+1.49e+05  21        -3
 O:  1         5        23 +0.00e+00,+0.00e+00,-1.36e+05,+1.62e+05  22       -19
```

Figure 5.2: Example of signal process vertex $d\bar{d} \to Z$ in the event record in Pythia8.

The quarks and antiquark $d\bar{d}$ are incoming particles into the hard subprocess (Stat 21) and will be used to calculate the scale factor (Eq. (5.2)) from the LPXSignalReweightingTool. The outgoing Z boson is an intermediate particle from the hardest scattering (Stat 22) that will also be used to calculate the scale factor. We have now found both the incoming $d\bar{d}$ pair and the intermediate Z boson in the event record. We need to follow the Z boson further in the event record to find the final-state leptons. Summarized, the signal process in a Pythia8 Drell-Yan event can be represented as

$$q\bar{q} \to Z_q, \tag{5.3}$$

where Z_q denotes that the Z boson is produced from a quark-antiquark annihilation. The process in Eq. (5.3) can also be represented as the Feynman vertex in Figure 5.3, and is the first vertex of the s-channel diagram in Figure 5.1.

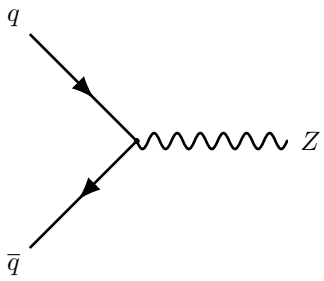


Figure 5.3: The Feynman vertex of the Z boson production process in Eq. (5.3).

5.2.2 Initial state showers

The Z_q boson from the signal process is not unique for calculating the scale factor (Eq. (5.2)) from the LPXSignalReweightingTool. After the signal process the Z_q boson decays several times in the event record which is shown in the event vertices in Figure 5.4, 5.5 and 5.6. These decays are branchings from the initial state showers (Stat 44) and change the kinematics of the Z_q boson. In the last branching in Figure 5.6 the outgoing Z boson has primordial k_T (Stat 62) included as explained in Section 3.1.

```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:      -19 ID:    0 (X,cT)= 0
 I:  1         5       23 +0.00e+00,+0.00e+00,-1.36e+05,+1.62e+05  22       -19
 O:  1        25       23 +1.35e+03,+3.80e+03,-1.38e+05,+1.63e+05  44       -55
```

Figure 5.4: The first initial state branching of the Z_q boson to the Z_{b_1} boson in the event record in Pythia8.

```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:      -55 ID:    0 (X,cT)= 0
 I:  1        25       23 +1.35e+03,+3.80e+03,-1.38e+05,+1.63e+05  44       -55
 O:  1        71       23 +3.09e+03,+5.60e+03,-1.38e+05,+1.64e+05  44      -250
```

Figure 5.5: The second initial state branching of the Z_q boson to the Z_{b_1} boson in the event record in Pythia8.

```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:     -312 ID:    0 (X,cT)= 0
 I:  1       323       23 +3.36e+03,+5.89e+03,-1.38e+05,+1.64e+05  44      -312
 O:  1       401       23 +2.63e+03,+5.04e+03,-1.37e+05,+1.63e+05  62      -407
```

Figure 5.6: The fourth and final state branching of the Z_q boson into the final Z_l boson that decays to a lepton pair in the event record in Pythia8.

The outgoing Z boson from Figure 5.6 is the Z boson that decays to the lepton pair and will be labeled Z_l. Summarized, the process of initial state showering in n branching of the Z_q boson is

$$Z_q \to Z_{b_1} \to ... \to Z_{b_n} \to Z_l, \tag{5.4}$$

where Z_{b_n} is the nth and final branched Z boson and Z_l is the Z boson that decays to the lepton pair.

5.2.3 Decay of Z boson and final state showers

The leptons that are used to calculate the scale factor appear in the event record after the final-state showers from the decay of the Z_l boson. An example of the Z_l boson decaying to an electron and a positron (PDG ID 11 and -11) in the event record is shown in Figure 5.7, while the radiation of a photon (PDG ID 22) from the positron is shown in the event vertex in Figure 5.8.

```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:     -407 ID:    0 (X,cT)= 0
 I:  1       401       23 +2.63e+03,+5.04e+03,-1.37e+05,+1.63e+05  62      -407
 O:  2       515       11 -8.83e+03,-4.01e+04,-5.62e+04,+6.96e+04  23      -409
             516      -11 +1.15e+04,+4.52e+04,-8.12e+04,+9.36e+04  23      -408
```

Figure 5.7: An event vertex showing the Z_l boson decaying to an electron pair in the event record in Pythia8.

```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:     -408 ID:    0 (X,cT)= 0
 I:  1       516      -11 +1.15e+04,+4.52e+04,-8.12e+04,+9.36e+04  23      -408
 O:  2       517      -11 +1.15e+04,+4.51e+04,-8.12e+04,+9.36e+04  51      -411
             518       22 +1.91e+00,+7.64e+00,-1.39e+01,+1.60e+01   1
```

Figure 5.8: An event vertex showing the radiation of a photon from the positron in the event record in Pythia8.

The final state positron that we use to calculate the scale factor is the outgoing positron (Stat 1) from the event record in Figure 5.9. Radiation of photons may also happen to the electron.

```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:     -411 ID:    0 (X,cT)= 0
 I:  1       517      -11 +1.15e+04,+4.51e+04,-8.12e+04,+9.36e+04  51      -411
 O:  1       522      -11 +1.15e+04,+4.51e+04,-8.12e+04,+9.36e+04   1
```

Figure 5.9: The event vertex with the final state positron in the event record in Pythia8.

Generally we can write the decay of the Z_l boson to a lepton pair l^-l^+ with m radiated photons as

$$Z_l \to l^-l^+\gamma_1...\gamma_m, \tag{5.5}$$

where γ_m is the mth radiated photon. The energy/momentum of the photon in Figure 5.8 is negligible compared to that of the positron. This is generally the case, so we can ignore the photon radiation in the process Eq. (5.5) such that it simplifies to the process

$$Z_l \to l^-l^+. \tag{5.6}$$

The process in Eq. (5.6) can be represented as the Feynman vertex in Figure 5.3, and is the second vertex of the s-channel diagram in Figure 5.1.

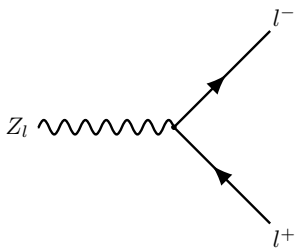


Figure 5.10: The Feynman vertex for the Z boson decay process in Eq. (5.6)

This ends the search to find the particles needed for the LPXSignalReweightingTool. In the next subsection we will put together the pieces in order to calculate the scale factor.

5.2.4 Combining the Z production and Z decay vertices

We can now combine the Z production vertex in Figure 5.3 with the Z decay vertex in Figure 5.7 to make the s-channel diagram in Figure 5.11. This is the Feynman diagram that the LPXSignalReweightingTool uses to calculate the differential cross section in the scale factor. The $Z_{q/l}$ boson in Figure 5.11 denotes that we can either use the Z_q boson from the production vertex, or the Z_l from the decay vertex to calculate the differential cross section in the scale factor. The LPXSignalReweightingTool uses the Z_q boson to calculate the scale factor for Pythia8.

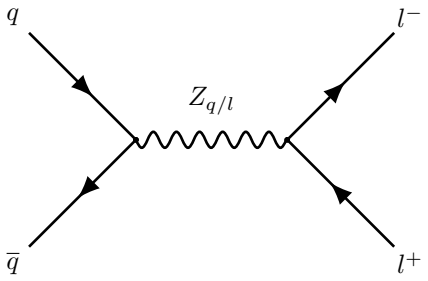


Figure 5.11: The s-channel Feynman diagram obtained from combining the Z production vertex and the Z decay vertex in Figure 5.3 and 5.10 respectively.

5.2.5 Calculating the scale factor

We have now found all the particles that we need to calculate the scale factor for this Pythia8-generated event. These are listed with their kinematics in Table 5.1. The invariant mass of the final state e^+e^- pair in this event is $m_{ee} = 87.92$ GeV and is close to the Z boson mass. The final information we need to calculate the scale factor is the Z' model and the Z' mass that we set ourselves.

Particle	(p_x,	p_y,	p_z,	E) [GeV]
d	(0,	0,	12.94,	12.94)
$\overline{d}$	(0,	0,	-149.37,	149.37)
Z_q	(0,	0,	-136.44,	162.31)
e^-	(-8.83,	-40.11,	-56.16,	69.57)
e^+	(11.46,	45.15,	-81.21,	93.63)

Table 5.1: The kinematics of the particles from the event record used to calculate the scale factor.

Table 5.2 shows the scale factors SF_{BSM} from Eq. (5.2) for this event in the Z'_{SSM}, Z'_{χ} and Z'_{ψ} models at $m_{Z'} = 87.92$ GeV in Table 5.2b and $m_{Z'} = 3000$ GeV in Table 5.2b. The Z' width, $\Gamma_{Z'}$ and the ratios $\Gamma_{Z'}/m_{Z'}$ are also listed in Table 5.2a and 5.2b. We see that the ratios $\Gamma_{Z'}/m_{Z'}$ are in good agreement with the theoretical ratios 3% (Z'_{SSM}), 1.2% (Z'_{χ}) and 0.5% (Z'_{ψ}) between the width and the mass of the Z' boson as given in Section 2.2.1 and 2.2.2.

The scale factors calculated at $m_{Z'} = 87.92$ GeV are larger compared to the scale factors calculated at $m_{Z'} = 3000$ GeV in Table 5.2a and 5.2b, respectively. This is expected since $m_{Z'} = 87.92$ GeV is equal to the invariant mass of the final state electron pair, $m_{ee} = 87.92$ GeV where there are many events in the Z' resonance. While $m_{Z'} = 3000$ GeV is far from the invariant mass of the final state electron pair where there are fewer events. We can generally say that events with invariant mass close to the Z' mass are up-weighted, while events with invariant mass far from the Z' mass are down-weighted.

Model	$\Gamma_{Z'}$[GeV]	$\Gamma_{Z'}/m_{Z'}$[%]	SF_{BSM}
Z'_{SSM}	2.42	2.76	8.80
Z'_{χ}	1.02	1.16	15.29
Z'_{ψ}	0.42	0.48	12.62

(a) $m_{Z'} = 87.92$ GeV

Model	$\Gamma_{Z'}$[GeV]	$\Gamma_{Z'}/m_{Z'}$[%]	SF_{BSM}
Z'_{SSM}	95.22	3.17	$4.95 \cdot 10^{-9}$
Z'_{χ}	37.06	1.24	$1.51 \cdot 10^{-9}$
Z'_{ψ}	16.88	0.56	$2.10 \cdot 10^{-10}$

(b) $m_{Z'} = 3000$ GeV

Table 5.2: Ths scale factors SF_{BSM}, Z' width and the ratio $\Gamma_{Z'}/m_{Z'}$ for different Z' models at $m_{Z'} = 87.92$ GeV (a) and $m_{Z'} = 3000$ GeV (b).

We have now shown how the LPXSignalReweightingTool uses the event record to calculate the scale factor for one event that is applied to the kinematics variables to reweight a Drell-Yan distribution to a Z' distribution. In the next section we will extend the signal reweighting method that we have presented here to signal reweight PowhegPythia8 Drell-Yan distributions and give an example on how the whole signal reweighted distribution might look like.

5.3 PowhegPythia8

In this section we will extend the signal reweighting method used in LPXSignalReweightingTool to a PowhegPythia8 Drell-Yan sample. The PowhegPythia8 Drell-Yan sample

has NLO order QCD signal processes while the Pythia8 Drell-Yan sample has LO signal processes. Consequently there will be a difference in how we find the Z production vertex in PowhegPythia8 compared to Pythia8. We studied the event record in detail for Pythia8 Drell-Yan samples, and will not go in such detail here whenever it is not needed.

5.3.1 Signal process

We need to use another method to find the Z production vertex in the signal process in NLO PowhegPythia8 compared to what we did for Pythia8. We see this in Figure 5.12 which shows an example of a signal process in the PowhegPythia8 Drell-Yan sample with an incoming quark pair $u\overline{u}$ (PDG ID 2 and -2), and a gluon (PDG ID 21) and Z boson (PDG ID 23) as outgoing particles. The signal process in PowhegPythia8 has an extra outgoing gluon compared to the signal process in Pythia8. As a consequence of the extra gluon we cannot any longer represent Figure 5.12 directly as a Z production vertex from the annihilation of two quarks like we did for Pythia8 in Section 5.2.1.

```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:       -3 ID:    0 (X,cT)= 0
 I:  2         3        2 +0.00e+00,+0.00e+00,+2.19e+05,+2.19e+05  21        -3
               4       -2 +0.00e+00,+0.00e+00,-7.48e+04,+7.48e+04  21        -3
 O:  2         5       23 -1.50e+02,-1.16e+04,+9.50e+04,+2.43e+05  22       -58
               6       21 +1.50e+02,+1.16e+04,+4.91e+04,+5.05e+04  23       -59
```

Figure 5.12: The event vertex showing the signal process $u\overline{u} \to gZ$ in the event record in PowhegPythia8.

The signal process in the Figure 5.12 is not the only type of signal processes that exist in a NLO PowhegPythia8 sample. There are in total four different types of signal processes which are

$$q\overline{q} \to Z \tag{5.7}$$

$$q\overline{q} \to gZ \tag{5.8}$$

$$qg \to qZ \tag{5.9}$$

$$\overline{q}g \to \overline{q}Z. \tag{5.10}$$

The signal process in Eq. (5.7) is the same as the Pythia8 signal process, while the

signal processes in Eq. (5.8) - (5.10), involving a Z and an additional parton, are not in LO Pythia8. We will see in the next section how we can extract the Z production vertex from the signal processes in Eq. (5.8) - (5.10).

5.3.2 Finding the Z production vertex

To find the Z production vertex we first need to represent the signal processes in Eq. (5.8) - (5.10) as t-channel Feynman diagrams as shown in Figure 5.13. Figures 5.13a and 5.13b show the t-channel diagrams of the signal process in Eq. (5.8), while Figures 5.13c and 5.13d show the t-channel diagram of Eq. (5.9) and (5.10) respectively. We will come back the signal process in Eq. (5.8) which can be represented in the two different t-channel diagrams in Figures 5.13a and 5.13b.

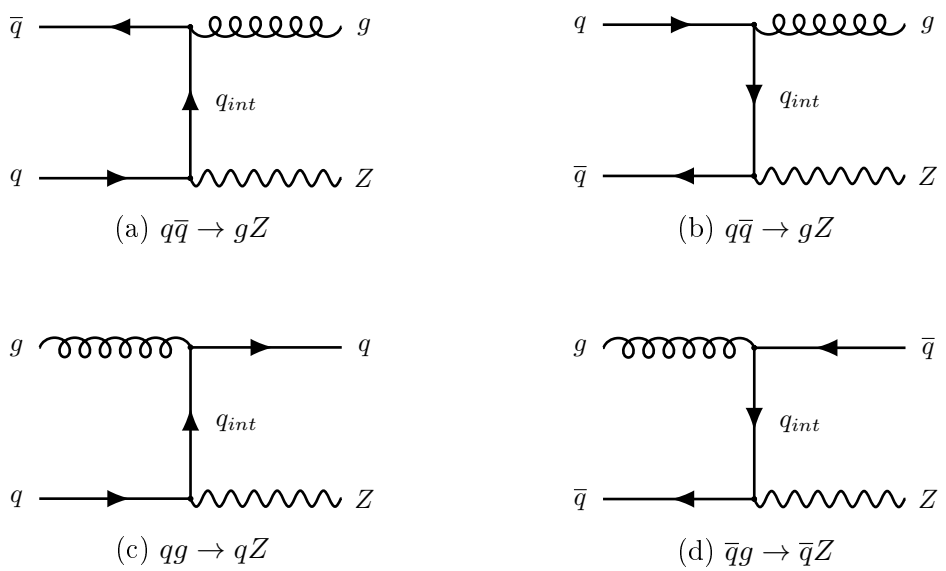
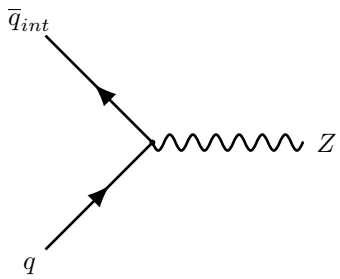
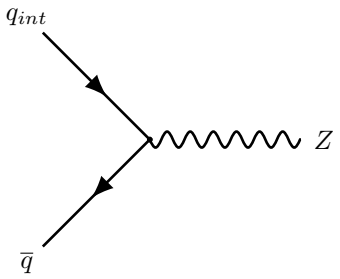


Figure 5.13: Feynman diagrams for the different signal processes in Eq. (5.8) - (5.10).

The intermediate particle in each of the Feynman diagrams in Figures 5.13 is the quark q_{int}. We can now find the Z production vertex $qq_{int}Z$ in the Feynman diagrams in Figure 5.13a and 5.13c and the vertex $\overline{q}q_{int}Z$ in Figure 5.13b and 5.13d. Figure 5.14 shows these Z production vertices where the q_{int} quark is time-ordered such that the Z production vertices are quark-anitquark annihilations.



(a) The time-ordered Z production vertex from the Feynman diagrams in Figure 5.13a and 5.13c.



(b) The time-ordered Z production vertex from the Feynman diagrams in Figure 5.13b and 5.13d.

Figure 5.14: The time-ordered Z production vertices from the Feynman diagrams in Figure 5.13.

The two time-ordered Z production vertices from Figure 5.14 can be represented as the processes

$$q\overline{q}_{int} \rightarrow Z_q \tag{5.11}$$

$$\overline{q}q_{int} \rightarrow Z_q, \tag{5.12}$$

where Z_q denotes that the Z boson is produced from a quark-antiquark annihilation. We have now found the Z production vertices from the signal process in PowhegPythia8. The next step is to find the Z decay vertex. The initial-state branching of the Z_q boson is the same as for Pythia8 in Section 5.2.2. We will therefore jump right to the decay and final state radiation in the next section.

5.3.3 Decay of Z boson and final-state radiation

In PowhegPythia8 the Z boson decay and the final-state radiation take place in one event vertex as shown in Figure 5.15. This is slightly different compared to Pythia8 where the decay and the final-state radiation takes place in multiple event vertices as explained in Section 5.2.3. The Z boson (PDG ID 23) in Figure 5.15 decays to an electron-positron pair (PDG ID 11 and -11 with Stat 1) and two photons (PDG ID 22). There is also another electron pair (Stat 3), but this is only used as a documentation entry. We will therefore use the electron pair with Stat 1 in the LPXSignalReweightingTool since it is the final stare particles.

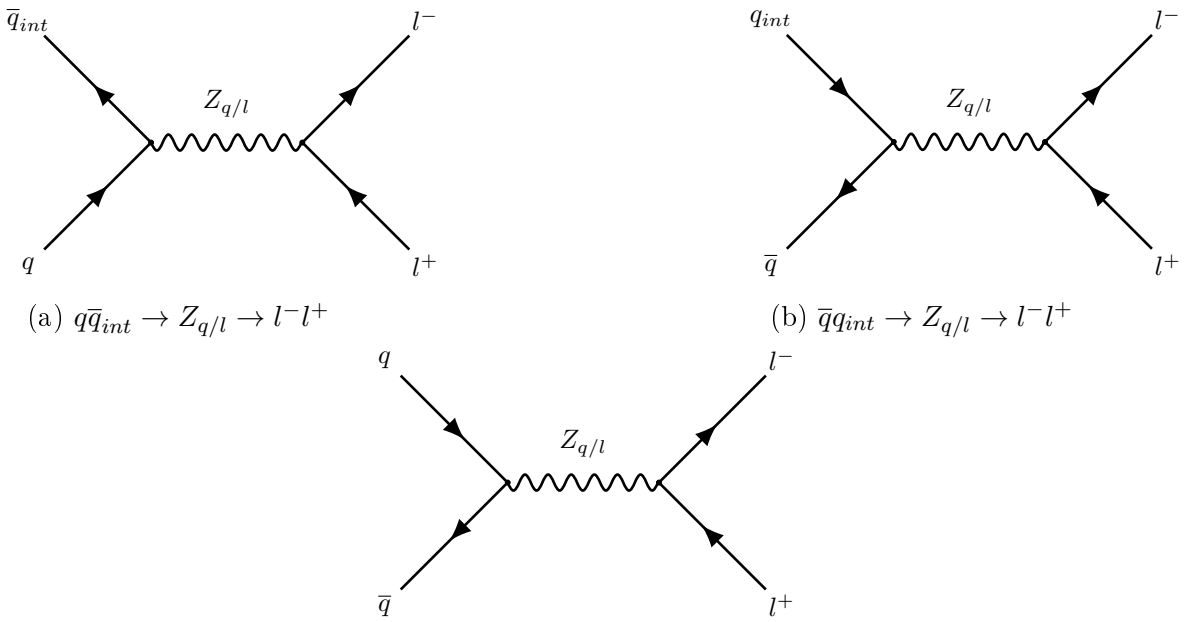
```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:     -471 ID:    0 (X,cT)= 0
 I:  1       464       23 -4.24e+02,-1.44e+04,+9.52e+04,+2.43e+05  62      -471
 O:  6       618       11 +4.67e+04,-1.37e+04,+1.58e+05,+1.65e+05   1
             619      -11 -4.71e+04,-6.48e+02,-6.24e+04,+7.82e+04   1
           10001       22 -2.42e+01,-8.75e+01,-1.65e+02,+1.89e+02   1
           10002       22 -5.96e-01,-8.09e-03,-7.89e-01,+9.89e-01   1
           10003       11 +4.67e+04,-1.37e+04,+1.58e+05,+1.65e+05   3
           10004      -11 -4.71e+04,-7.34e+02,-6.26e+04,+7.83e+04   3
```

Figure 5.15: The event vertex showing the Z boson decaying to an electron pair and two radiated photons in PowhegPythia8.

We can now write the general process for the Z boson decay with final state radiation like we did for Pythia8 in Eq. (5.5). The kinematics of the photons in Figure 5.15 are tiny compared to the electron-positron pair and we can therefore neglect the photons. The simplified process of Eq. (5.5) is then approximated to the Z boson decay process in Eq. (5.6) which is represented as the decay vertex in Figure 5.15.

5.3.4 Combining the Z production and Z decay vertices

The Z boson production and Z decay vertices can now be combined like we did for Pythia8 in Section 5.2.4. We found three different Z production vertices for PowhegPythia8. The Z boson production vertices are $q\overline{q}_{int} \to Z$ and $q_{int}\overline{q} \to Z$ from Figure 5.14a and 5.14b respectively, and the production vertex $q\overline{q} \to Z$ from Figure 5.3. Figure 5.16 shows the three s-channel diagrams we get when we combine these Z boson production vertices with the Z boson decay vertex $Z_l \to l^- l^+$ in Figure 5.10.



(a) $q\overline{q}_{int} \to Z_{q/l} \to l^-l^+$

(b) $\overline{q}q_{int} \to Z_{q/l} \to l^-l^+$

(c) $q\overline{q} \to Z_{q/l} \to l^-l^+$

Figure 5.16: The s-channel Feynman diagrams we get when combining the Z boson decay vertex in Figure 5.10 with the Z boson production vertex in Figure 5.14a (a), 5.14b (b) and 5.3 (c).

The $Z_{q/l}$ boson in Figure 5.16 denotes that we can either use the Z_q boson from the production vertex, or the Z_l boson from the decay vertex to calculate the differential cross section in the scale factor in Eq. (5.2), page 42. The intermediate quark q_{int} and anti-quark $\overline{q}_{int}$ are not listed in the event record, and so their four-momentum are not known. We can however calculate the four-momentum of q_{int} and $\overline{q}_{int}$ from the four-momentum conservation in the Z boson production vertices in Figure 5.16a and 5.16b respectively as

$$p^{\mu}_{\overline{q}_{int}} = p^{\mu}_{Z_{q/l}} - p^{\mu}_{q} \tag{5.13}$$

$$p^{\mu}_{q_{int}} = p^{\mu}_{Z_{q/l}} - p^{\mu}_{\overline{q}}. \tag{5.14}$$

The four-momentum of the intermediate quarks in Eq. (5.13) and (5.14) depend on whether we choose the Z_l boson or the Z_q boson. In Section 7.1 we will investigate how the scale factor changes based on whether we choose the Z_q boson before the initial-state radiation, or the Z_l boson after the initial-state radiation in the LPXSignalReweighting-Tool to calculate the four-momentum in Eq. (5.13) and (5.14). We will also investigate how the scale factor changes based on whether we choose the t-channel in Figure 5.13a

or 5.13b for the events with the signal process $q\bar{q} \to gZ$.

Table 5.3 summarizes the signal processes and their corresponding s-channel processes. The implementation of PowhegPythia8 into the LPXSignalReweightingTool will therefore consist of looping over the events and, based on the signal process, the corresponding s-channel process in Table 5.3 will be used to calculate the scale factor.

Signal process	s-channel process
$q\bar{q} \to Z_q$	$q\bar{q} \to Z_{q/l} \to l^-l^+$
$q\bar{q} \to gZ_q$	$q\bar{q}_{int} \to Z_{q/l} \to l^-l^+$
$q\bar{q} \to gZ_q$	$q\bar{q}_{int} \to Z_{q/l} \to l^-l^+$ $\bar{q}q_{int} \to Z_{q/l} \to l^-l^+$
$qg \to qZ_q$	$q\bar{q}_{int} \to Z_{q/l} \to l^-l^+$
$\bar{q}g \to \bar{q}Z_q$	$\bar{q}q_{int} \to Z_{q/l} \to l^-l^+$

Table 5.3: Overview over the signal processes in PowhegPythia8 and their corresponding s-channel processes.

5.3.5 Calculating the scale factor

We have now found all the particles that we need to calculate the scale factor for this PowhegPythia8 event. The particles are listed with their kinematics in Table 5.4. The steps to calculate the scale factor follow the same procedure as in Section 5.2.5 with the exception of the $\overline{u}_{int}$ quark in Table 5.4. The signal process in this event is $u\overline{u} \to gZ$ from Figure 5.12 and we choose the $\overline{u}_{int}$ quark from the s-channel process in Table 5.3. The $\overline{u}_{int}$ kinematics is calculated from Eq. (5.14) in this example, where the Z_l boson is choosed in this example. The invariant mass of the final state e^+e^- pair in this event is $m_{ee} = 223.30$ GeV. The final information we need to calculate the scale factor is the Z' model, and the corresponding couplings, and the Z' mass that we set ourselves.

Particle	(p_x,	p_y,	p_z,	E) [GeV]
u	(0,	0,	218.90,	218.90)
$\overline{u}_{int}$	(-0.42,	-14.41,	-123.69,	24.53)
Z_l	(-0.42,	-14.41,	95.21,	243.43)
e^-	(46.68,	-13.68,	157.76,	165.08)
e^+	(-47.08,	-0.65,	-62.38,	78.15)

Table 5.4: The kinematics of the particles from the event record used to calculate the scale factor.

Table 5.5 shows the scale factors SF_{BSM} from Eq. (5.2) for this event in the Z'_{SSM}, Z'_{χ} and Z'_{ψ} model at $m_{Z'} = 223.30$ GeV in Table 5.5a and $m_{Z'} = 3000$ GeV in Table 5.5b. The Z' width, $\Gamma_{Z'}$ and the ratios $\Gamma_{Z'}/m_{Z'}$ are also listed in Table 5.5a and 5.5b. We see that the ratios $\Gamma_{Z'}/m_{Z'}$ are in good agreement with the theoretical ratios 3% (Z'_{SSM}), 1.2% (Z'_{χ}) and 0.5% (Z'_{ψ}) in Section 2.2.1 and 2.2.2, page 10 and 11.

The scale factors calculated at $m_{Z'} = 223.30$ GeV are larger compered to the scale factors calculated at $m_{Z'} = 3000$ GeV in Table 5.5a and 5.5b respectively. This is expected since $m_{Z'} = 223.30$ GeV is equal to the invariant mass of the final state e^+e^- pair, $m_{ee} = 223.30$ GeV where there are many events. While $m_{Z'} = 3000$ GeV is far from the invariant mass of the final state e^+e^- pair where there are few events. We can generally say that event with invariant mass close to the Z' mass are up-weighted, while events with invariant mass far from the Z' mass are down-weighted form the discussion in Section 5.2.5.

Model	$\Gamma_{Z'}$[GeV]	$\Gamma_{Z'}/m_{Z'}$[%]	SF_{BSM}
Z'_{SSM}	6.20	2.78	162.14
Z'_{χ}	2.60	1.16	83.59
Z'_{ψ}	1.07	0.48	227.92

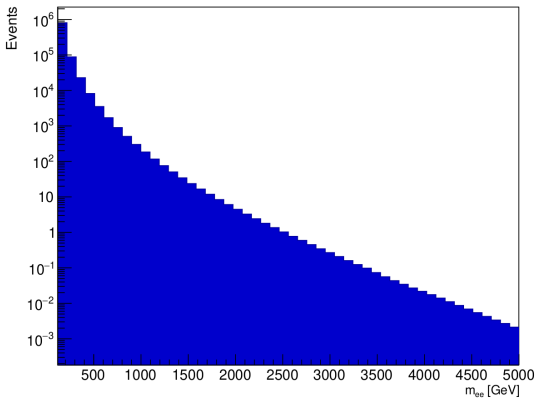
(a) $m_{Z'} = 223.3$ GeV

Model	$\Gamma_{Z'}$[GeV]	$\Gamma_{Z'}/m_{Z'}$[%]	SF_{BSM}
Z'_{SSM}	95.22	3.17	$3.93 \cdot 10^{-6}$
Z'_{χ}	37.06	1.24	$3.68 \cdot 10^{-7}$
Z'_{ψ}	16.88	0.56	$2.04 \cdot 10^{-7}$

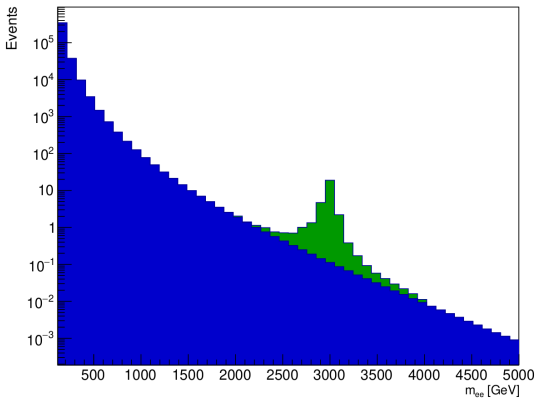
(b) $m_{Z'} = 3000$ GeV

Table 5.5: The scale factors SF_{BSM}, Z' width and the ratio $\Gamma_{Z'}/m_{Z'}$ for different Z' models at $m_{Z'} = 223.3$ GeV (a) and $m_{Z'} = 3000$ GeV (b)

We have now extended the LPXSignalReweightingTool to use the event record for PowhegPythia8 to calculate the scale factor that is applied to the kinematical variables to reweight a Drell-Yan distribution to a Z' distribution. Figure 5.17 shows an example of signal reweighing of the whole invariant mass distribution where Figure 5.17a shows the PowhegPythia8 Drell-Yan distribution that is used to calculate the scale factor and Figure 5.17b shows the signal reweighted Z' distribution on top of the Drell-Yan distribution.



(a) The PowhegPythia8 Drell-Yan invariant mass distribution that is used to calculate the scale factor.



(b) The PowhegPythia8 Drell-Yan invariant mass distribution (blue) with the signal reweighted Z' distribution (green) on top.

Figure 5.17: The PowhegPythia8 Drell-Yan invariant mass distribution (blue) (a) used for signal reweghting and the signal reweighted Z' (green) on top of the Drell-Yan distribution (blue) (b).

In the next section we will try to extend the signal reweighting method that we have presented here to signal reweight a Sherpa Z + jets distribution, and explain why it is harder to do it there.

Chapter 6

Z' signal reweighting with Sherpa - an attempt

In this chapter we will present an attempt to extend the LPXSignalReweightingTool to apply signal reweighting to a Sherpa Z + jets distribution rather than the Pythia8 or PowhegPythia8 Drell-Yan distribution described in Section 5.1 and 5.3 respectively. The motivation for doing signal-reweighting on Sherpa is because it can simulate multi-jet events and therefore model LHC events more accurately. The hard process is also calculated in NLO in QCD as for PowhegPythia8. The color flow information from the HepMC event record will be important for the attempt to extend the LPXSignalReweightingTool. In the first section we will study the signal process in Sherpa and explain why the method we used to reweight PowhegPythia8 is not suitable to apply signal reweighting to a Sherpa Z + jets distribution. The issue is finding the Z production vertex, as is explained in Section 6.2. Significant effort was expended to develop a method for unambiguously locating the Z production vertex in Sherpa events. The method developed for this thesis uses the color flow information from the event generators together with methods from graph theory to find the Z production vertex. In Section 6.3 we will describe the challenges we met when trying to use this method on the color flows from Sherpa. Lets us first look at the signal process in Sherpa.

6.1 Signal process

We need to use another method to find the Z production vertex in the signal process in Sherpa compared to what we did for PowhegPythia8 and Pythia8. We see this in Figure 6.1 which shows an example of a signal process in the Sherpa Z + jets sample with two incoming partons, a down quark (PDG ID 1) and a gluon (PDG ID 21), and five outgoing particles, one down quark, two gluons and an e^-e^+ pair (PDG ID 11 and -11). The Z boson is not shown in the event record in Sherpa, but the e^+e^- decayproducts of the Z boson (Figure 6.1).

```
        Barcode   PDG ID      ( Px,       Py,       Pz,     E ) Stat  DecayVtx
-------------------------------------------------------------------------------
TruthVertex:       -5 ID:    1 (X,cT)= 0
 I:  2     10010       21 +0.00e+00,+0.00e+00,+5.79e+04,+5.79e+04   3        -5
           10011        1 +0.00e+00,+0.00e+00,-1.22e+06,+1.22e+06   3        -5
 O:  5     10012       11 +2.93e+04,-2.93e+04,-7.92e+04,+8.94e+04   3
           10013      -11 -3.71e+03,-1.06e+04,-4.26e+05,+4.27e+05   3
           10014       21 +4.11e+03,-1.47e+04,+3.63e+04,+3.94e+04   3
           10015       21 +1.71e+04,+4.97e+04,-3.54e+04,+6.34e+04   3
           10016        1 -4.69e+04,+4.98e+03,-6.59e+05,+6.61e+05   3
```

Figure 6.1: The event vertex showing the signal process $dg \to dgge^-e^+$ in the event record in Sherpa.

The signal process in Sherpa has two extra outgoing partons, making it a Z + 3jets, compared to the signal process in PowhegPythia8, which is at most Z + 1jet. As a consequence of the two extra partons we can represent Figure 6.1 in 28 different Feynman diagrams as shown in Figure 6.2 using CompHEP [50]. The method we used to find the Z production vertex from the annihilation of a pair quark-antiquark in PowhegPythia8, see Section 5.3.2, will not work here because of the large number Feynman diagrams and therefore the many different Z production vertices. In other words we can not find the Z production vertex unambiguously with the method we used with PowhegPythia8. The signal process in Figure 6.1 is also only one of many different signal processes, initiated by other pairs of partons. We will therefore need a more general way to reconstruct the Feynman diagram from the event vertex than with PowhegPythia8. We will make use of a method based on the color flow information from the event record, first descried in Section 3.3.2, to find a general method to reconstruct the signal process.

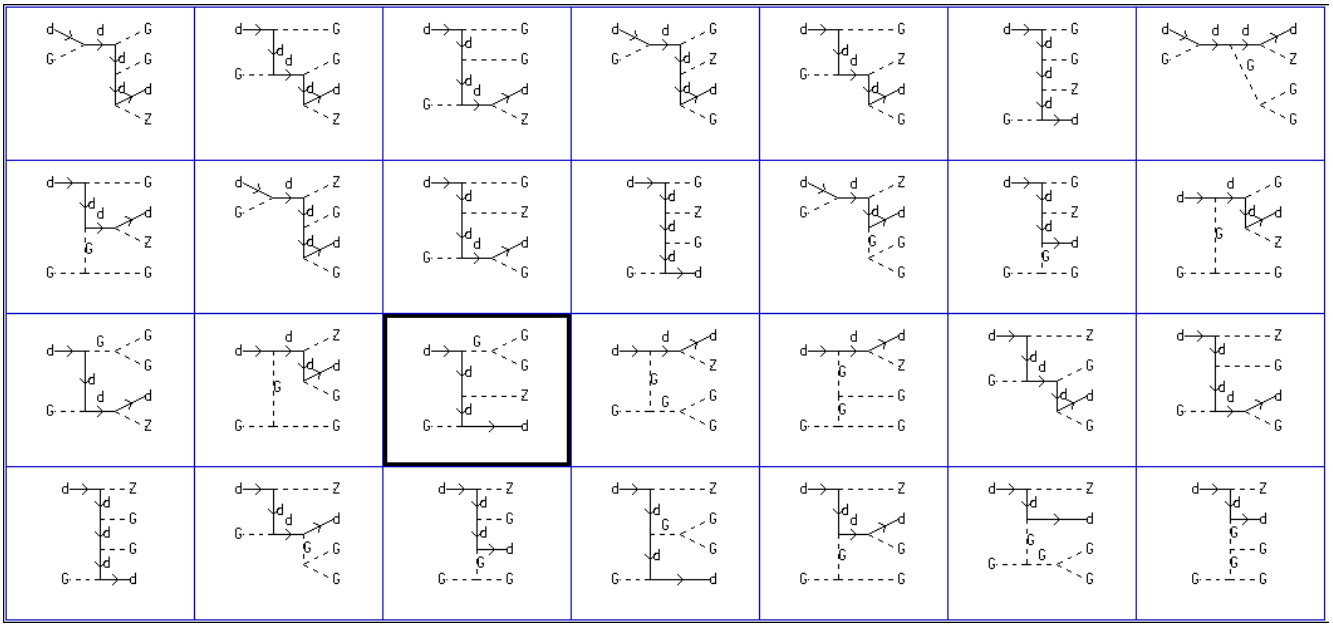


Figure 6.2: The 28 different LO Feynman diagrams of the signal process $dg \to dggZ$ generated by CompHEP [50].

6.2 Finding the Z production vertex

In Section 3.3.2 we described how the event generators store the information about the colors for the different partons as color flows as illustrated in Figure 3.4. We will now describe how we can use this color flow information to reconstruct the Feynman diagram for an event vertex and thus find the Z boson production vertex. This method uses the color flow from QCD to obtain the reconstructed Feynman diagram. We will show in Subsections 6.2.2 and 6.2.3 that both Feynman diagrams and color flows are two different types of graphs that can be related, which is central to the method. To do this it is necessary to introduce some concepts from graph theory that will be used in Section 6.2.1. In Section 6.2.5 we will study two examples where the color flows are used to reconstruct Feynman diagrams, where the highlighted Feynman diagram in Figure 6.2 will be the final example.

6.2.1 Introduction to Graph theory

We will in this section introduce the graph theory needed to understand how Feynman diagrams and color flows are connected to each other and are based on the references [51, 52]. Graphs are used in many different problems to represent and describe how different entities relate to one another, and can be visualized as a set of dots and lines

between the dots. Graphs are used to model relationships in everything from the atoms in molecules to bus stops in a city. In the last few years techniques have been developed to allow neural networks to learn from and process data that is represented as graphs, in so-called graph neural networks (GNNs). In particle physics graphs are used to visualize particle interactions in the form of Feynman graphs and these are the graphs that we will study in this chapter. We will split up the description of graphs into *undirected* and *directed* graphs, concepts that will be explained in Subsections 6.2.1.1 and 6.2.1.4 respectively. The examples that we use in these sections will be relevant for the Feynman diagrams and color flows respectively.

6.2.1.1 Undirected graphs

An *undirected graph* G is defined as a pair $G = (V, E)$ with a set of nodes V and a set of edges E. The edges $e \in E$ is a pair $e = (n_1, n_2)$ of nodes $n_1, n_2 \in V$ where the order of the nodes in the edges does not matter, i.e $e = (n_1, n_2) = (n_2, n_1)$. A graph can be represented as a picture with dots and lines where the dots represent the nodes and the lines that connects the dots are the edges as shown in Figure[1] 6.3. Figure 6.3 shows the star graph S_4[2] with the nodes $V = \{n_1, n_2, n_3, n_4\}$ and edges $E = \{e_1 = (n_1, n_2), e_2 = (n_1, n_3), e_3 = (n_1, n_4)\}$. We will use the star graph in Figure 6.3 to introduce some terminology to describe how the nodes in a graph are connected to each other that we need for later use.

The S_4 graph is called a *tree* since it has no loops in it. If two nodes share an edge we say that those nodes are *adjacent* to each other. In Figure 6.3 the node n_1 is adjacent to n_2, n_3 and n_4. The set of nodes that are adjacent to a node n is called the *neighborhood* of n and denoted $N(n)$.

[1] All the diagrams in this chapter is made using the *TikZ* package in LaTeX

[2] A star graph S_4 has a single common connecting node, and the number indicates the total number of nodes.

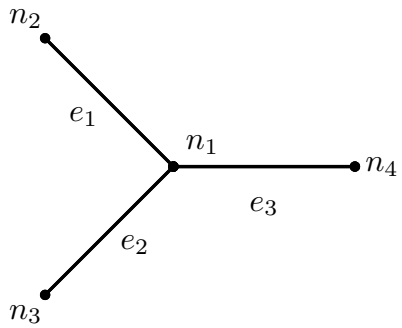
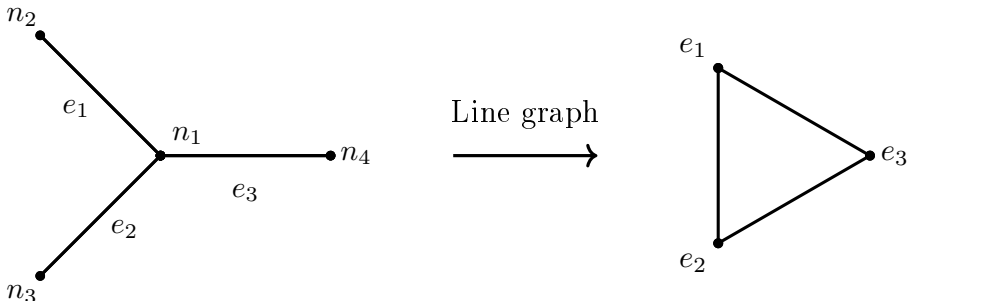


Figure 6.3: The star graph S_4 with 4 nodes and 3 edges.

The neighborhood of n_1 is therefore $N(n_1) = \{n_2, n_3, n_4\}$ and for the other nodes $N(n_2) = \{n_1\}$, $N(n_3) = \{n_1\}$ and $N(n_4) = \{n_1\}$. The number of nodes that are adjacent to a node n is called the *degree* of n and denoted $d(n)$. The degree of a node n is therefore equal to the size of the neighborhood of n, i.e $d(n) = |N(n)|$. In Figure 6.3 the degree of the respective nodes are $d(n_1) = 3$, $d(n_2) = 1$, $d(n_3) = 1$ and $d(n_4) = 1$. We have now introduced the necessary terminology that enables us to describe how the nodes in an undirected graph are connected. Next we will describe some methods that can be used to study the connections in a graph.

6.2.1.1.1 Line graphs To represent the adjacencies between the edges of an undirected graph G, we can construct a *line graph* $L(G)$. This $L(G)$ is built from G by replacing each *edge* in G by a *node*, such that the nodes of $L(G)$ are the *edges* of G. In addition, if two edges in G have a common node these edges become adjacent to each other in the line graph $L(G)$ (where these edges are nodes). It is easiest to see this with the example in Figure 6.4 which shows the star graph S_4 in Figure 6.4a and the corresponding line graph $L(S_4)$ in Figure 6.4b.



(*a*) The star graph S_4

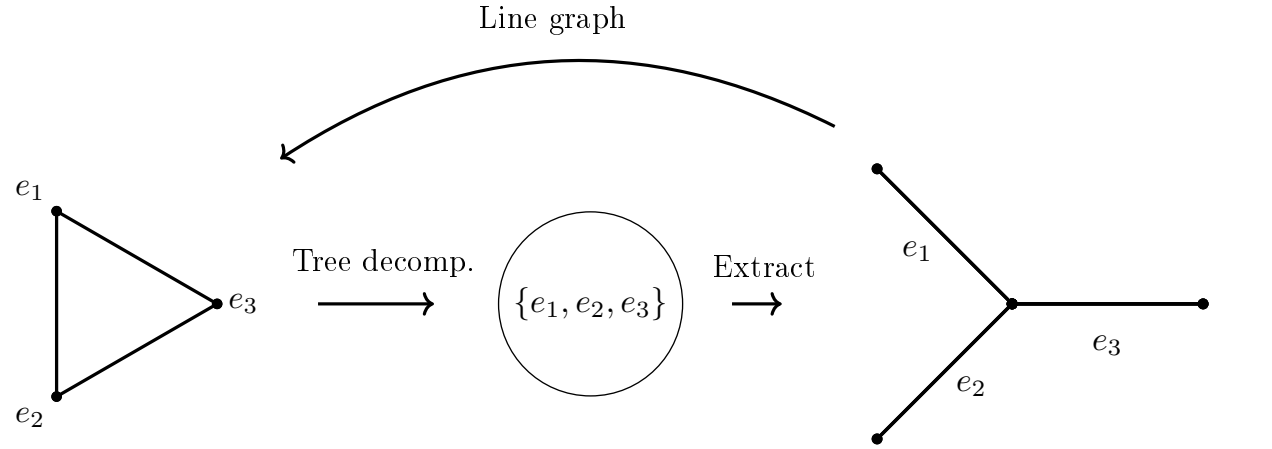
(*b*) The line graph $L(S_4)$ of S_4

Figure 6.4: The line graph $L(S_4)$ (b) of the star graph S_4 (a)

In Figure 6.4a the edges $e_1 = (n_1, n_2)$ and $e_2 = (n_1, n_3)$ have the node n_1 in common. Similarly the edges $e_1 = (n_1, n_2)$ and $e_3 = (n_1, n_4)$ have the node n_1 in common, and the edges $e_2 = (n_1, n_3)$ and $e_3 = (n_1, n_4)$ have the node n_1 in common. The edges e_1, e_2 and e_3 are therefore nodes in the line graph in Figure 6.4b that all are adjacent to each other. The line graph $L(S_4)$ in Figure 6.4b is called the triangle graph C_3 and is not a tree since it has a loop. The next operation we will study describes how *tree-like* a loop graph is and is called *tree decomposition*. This operation will enable us to find a method to go back from the line graph $L(S_4)$ to a star graph S_4 when we know that $L(S_4) = C_3$.

6.2.1.2 Tree decomposition of a triangle graph

The tree decomposition $T(G)$ of a graph G is a tree T with a set of *bags* B as nodes. A bag is a set of nodes in G that becomes a single node in the tree T. Every node in G needs to be contained in at least one bag. If two bags contain the same node they are adjacent to each other in the tree. The bags of a graph can be chosen in many ways and can therefore give many different tree decompositions of a graph. However, for the triangle C_3 in Figure 6.5 there exist only one tree decomposition which is the one where all the nodes $\{n_1, n_2, n_3\}$ are in one bag. Figure 6.5b shows the tree decomposition with this bag that is the only node.



(*a*) Triangle graph C_3 (*b*) Tree decomposition (*c*) Extracted tree, star graph S_4

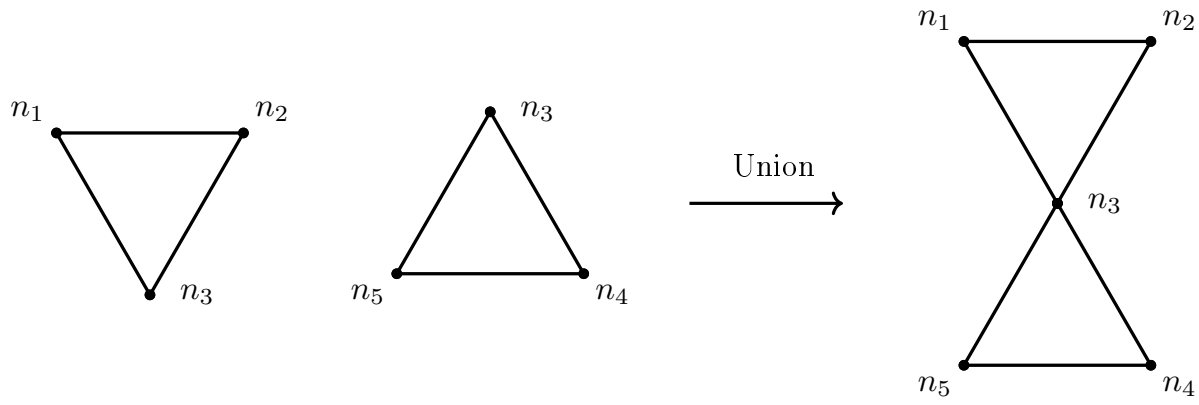
Figure 6.5: The cycle that shows how to get from the triangle graph C_3 (a) to its tree decomposition (b) to the extracted graph S_4 (c) and back to the graph in (a) via the line graph

We can go back to the star graph S_4 in Figure 6.5c by choosing the nodes in the bag

in the tree decomposition to be the edges in the S_4 graph. We will call this process the *extracting* of the nodes in a bag. The line graph of S_4 is again C_3 in Figure 6.5a and we have now shown how to go back to the star graph S_4 from its line graph $C_3 = L(S_4)$. It can be summarized as the extracted tree decomposition of the triangle graph C_3. In the next subsection we will extend the tree decomposition to a graph with multiple triangles.

6.2.1.3 Tree decomposition of a graph with multiple triangles

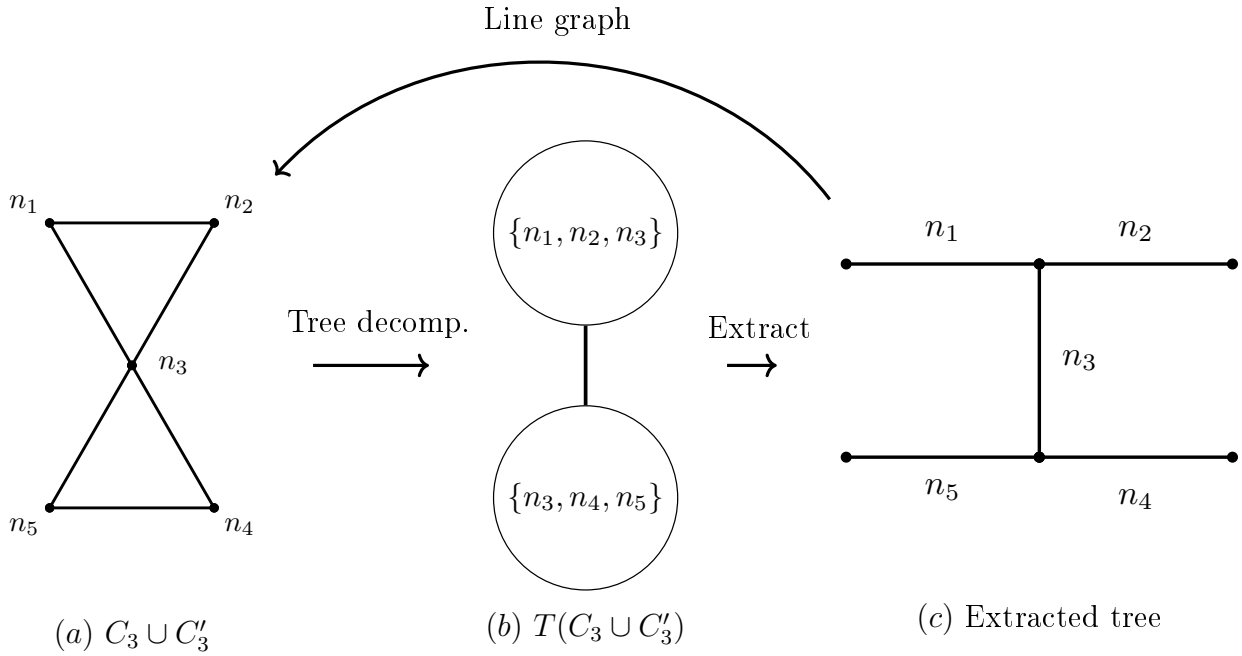
We will first see how we can join multiple triangles to one graph with the *union* operator. The union of two graphs G and G' with edges and nodes E,V and E',V', respectively, is defined as the union of their nodes and edges as $G \cup G' = (V \cup V', E \cup E')$. Figure 6.6a and 6.6b shows the two triangles C_3 and C_3' with the nodes $\{n_1, n_2, n_3\}$ and $\{n_3, n_4, n_5\}$ respectively. The union $C_3 \cup C_3'$ is shown in Figure 6.6c where the node n_3 joins the two triangles together since it is common to both triangles. The tree decomposition of



(a) C_3 (b) C_3' (c) $C_3 \cup C_3'$

Figure 6.6: The union $C_3 \cup C_3'$ (c) for the two triangles C_3 (a) and C_3' (b).

$C_3 \cup C_3'$ in Figure 6.7a is shown in Figure 6.7b where we again have chosen the bags to be the two triangles with the nodes $\{n_1, n_2, n_3\}$ and $\{n_3, n_4, n_5\}$ respectively. The two bags are adjacent to each other since they have the node n_3 in common and these two connected bags makes up the tree structure of the tree decomposition. Figure 6.7c shows the extracted graph of the tree decomposition where the nodes in the bags becomes edges. The line graph of the extracted tree is again the graph in Figure 6.7c and we have now completed the cycle as we did in Figure 6.5. This cycle in Figure 6.7 can be extended to a graph with more than two triangles and will be used later in this chapter. This



(a) $C_3 \cup C_3'$ (b) $T(C_3 \cup C_3')$ (c) Extracted tree

Figure 6.7: The cycle that shows how to get from the graph with two triangles (a) to its tree decomposition (b) to the extracted tree (c) and back to the graph in (a) via the line graph

concludes the introduction to undirected graphs. In the next section we will introduce directed graphs which will be used for the color flows.

6.2.1.4 Directed graphs

A *directed graph* G is defined as a pair $G = (V, E)$ with a set of nodes V and a set of directed edges E. The edge $\vec{e} \in E$ is a pair $\vec{e} = (n_1, n_2)$ of nodes $n_1, n_2 \in V$ where the order of the nodes in the edges does matter, i.e $\vec{e}_1 = (n_1, n_2) \neq (n_2, n_1)$. The order of the nodes in the edge determines the direction of the edge and is denoted with an arrow as shown in Figure 6.8. Figure 6.8 shows the *directed path* $\vec{P}_3$ with the nodes $V = \{n_1, n_2, n_3\}$ and edges $E = \{\vec{e}_1 = (n_1, n_2), \vec{e}_2 = (n_2, n_3)\}$. We will use the directed path in Figure 6.8 to introduce some terminology to describe how the nodes in a directed graph are connected to each other, that we will needed later. The directed path $\vec{P}_3$ can be written as an ordered list with arrows pointing between the nodes, such as $\vec{P}_3 = n_1 \to n_2 \to n_3$.

We say that a node n_1 is *adjacent to* another node n_2 if there is an edge between them and the arrow of the edge is directed from n_1 to n_2. In Figure 6.8 the node n_1 is adjacent to the node n_2, which is itself adjacent to the node n_3.

The set of nodes that a node n is adjacent to is called the *out-neighborhood* of n

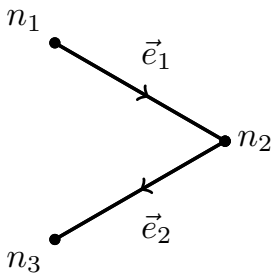
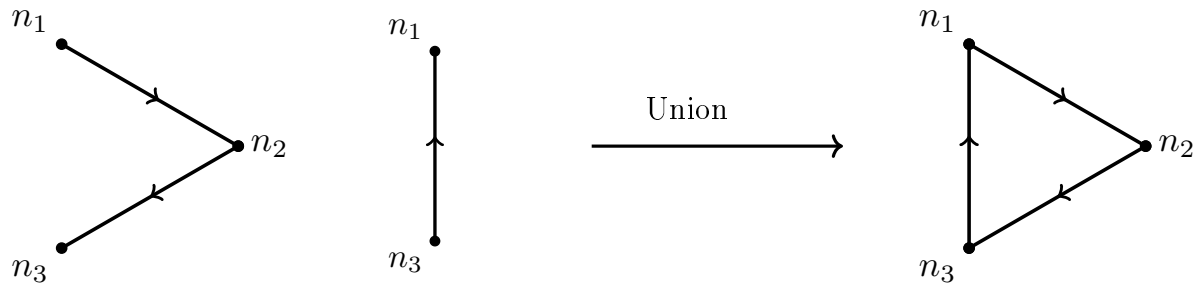


Figure 6.8: The directed path $\vec{P}_3$.

and denoted $N^+(n)$, while the set of nodes that are adjacent to a node n is called the *in-neighborhood* of n and denoted $N^-(n)$. The out- and in-neighborhood of n_2 in Figure 6.8 are $N^+(n_2) = \{n_3\}$ and $N^-(n_2) = \{n_1\}$ respectively, while the node n_1 only has an out-neighborhood $N^+(n_1) = \{n_2\}$ and the node n_3 only has a in-neighborhood $N^-(n_3) = \{n_2\}$. This is the only terminology that we need for a directed graph. In the next section we will look at the union of directed graphs and how we can transform a directed graph into an undirected graph.

6.2.1.4.1 Union of directed paths The union of two directed graphs G and G' with edges and nodes E,V and E',V' respectively is defined as the union of their nodes and edges as $G \cup G' = (V \cup V', E \cup E')$. Figure 6.9a and 6.9b show the two paths $\vec{P}_3 = n_1 \to n_2 \to n_3$ and $\vec{P}_2 = n_3 \to n_1$ with the nodes $\{n_1, n_2, n_3\}$ and $\{n_1, n_3\}$ respectively. The union $\vec{P}_3 \cup \vec{P}_2$ is shown in Figure 6.9c where the nodes n_1 and n_3 join the two paths together since they are common to both paths.

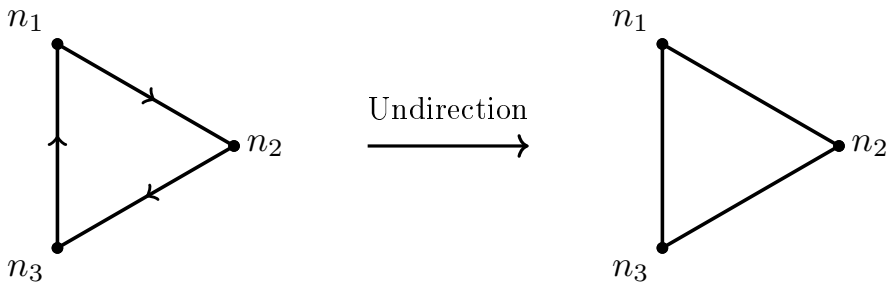


(a) $\vec{P}_3$ (b) $\vec{P}_2$ (c) $\vec{P}_3 \cup \vec{P}_2$

Figure 6.9: The union $\vec{P}_3 \cup \vec{P}_2$ (c) for the two directed paths $\vec{P}_3$ (a) and $\vec{P}_2$ (b).

The $\vec{P}_3 \cup \vec{P}_2$ graph is a directed triangle $\vec{C}_3$ graph, and the name for this process of getting a triangular graph from the union of paths is *triangulation*. We can make the

directed triangle an undirected triangle by removing the arrows on the edges as shown in Figure 6.9.



(a) Directed triangle graph $\vec{C}_3$ (b) Undirected triangle graph C_3

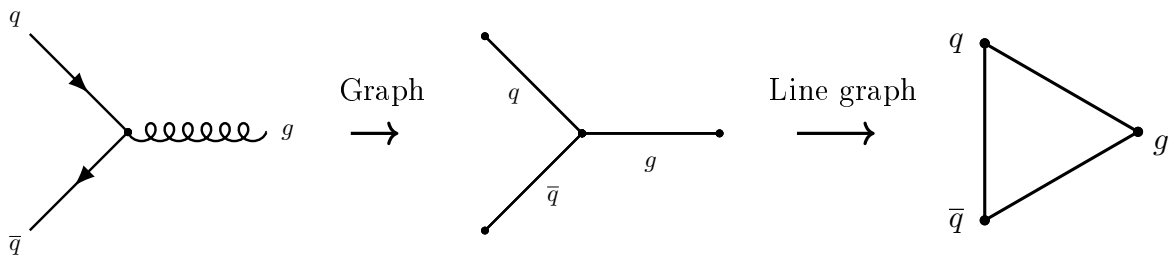
Figure 6.10: The undirected triangle C_3 (b) of the directed triangle $\vec{C}_3$ (a)

We have now come back to the triangle graph that we first studied in Section 6.2.1.1.1 and in Figure 6.10 and this is the last piece of graph theory that we need to study the Feynman diagrams and color flows in the next sections.

6.2.2 Feynman diagrams as graphs

We will first see how a Feynman vertex is a graph. The Feynman vertex in Figure 6.11a has three lines that represent the quark q, anti-quark $\overline{q}$ and gluon g that have a common node where they interact. The particle lines are formatted differently depending whether it is a fermion or a boson as described in Section 2.1.2. The formatting is not necessary for the vertex since the label on the particle line tells which particle it is, but it makes it easier to read the Feynman diagrams. The particle lines in Figure 6.11a only have a node where they interact. We can however put a node on the other ends of the lines where they do not interact with other particles. We can therefore get the graph from the Feynman vertex if we remove the formatting on the lines and add nodes on the end of the particle lines as shown in Figure 6.11b. The particle line is an edge in the graph, and the particles interact if they have a node in common.

The graph in Figure 6.11b is the star graph S_4 that we studied in detail in Section 6.2.1.1. The line graph of S_4 is the triangle graph in Figure 6.11c with the particles q,$\overline{q}$ and g as nodes. The triangle graph will be important for the next section where we will see how the color flow is a graph, and how the color flows in different vertices can be transformed to Feynman diagrams.



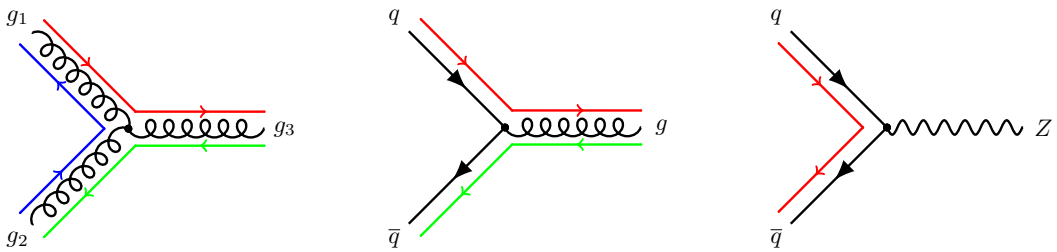
(a) Feynman vertex $q\overline{q} \to g$ (b) Star graph S_4 (c) The line graph C_3 of S_4.

Figure 6.11: The star graph S_4 (b) of the Feynman vertex (a), and the triangle C_3 (c) of the line graph S_4

6.2.3 Color flows as graphs

We showed in the previous section how the Feynman vertex in Figure 6.11 can be transformed to a triangle graph. In this section we will show that color flows are directed paths and how the color flows of a Feynman vertex can be transformed to a triangle graph.

We will only consider the QCD vertices which include three particles from Section 2.1.4 when working with the color flow. Thus we will not consider the four gluon interaction $gg \to gg$ from Figure 2.4c since it adds some complications, although this could be a topic for future work. In Section 2.1.4 the color flow for different vertices was described and are shown for the three particles in Figure 6.12 with the color flows added, where Figure 6.12a shows the $g_1g_2 \to g_3$ vertex, Figure 6.12b shows the $q\overline{q} \to g$ vertex and Figure 6.12c shows the $q\overline{q} \to Z$ vertex with the respective color flows represented with the colored arrows on the side of the particle lines. We listed the colors of the quarks and gluons in Eq. (2.4) and (2.3) corresponding to Figure 6.12a and 6.12b respectively.



(a) $g_1g_2 \to g_3$ (b) $q\overline{q} \to g$ (c) $q\overline{q} \to Z$

Figure 6.12: Color Feynman vertices

In each of the color vertices in Figure 6.12 the individual color flows are directed paths which include the particles with the same colors and follows the direction of the color flows. The different color flows for the vertices in Figure 6.12 can therefore be written as

the paths

$$\text{Figure 6.12}a: \quad red: g_1 \to g_3, \quad green: g_3 \to g_2, \quad blue: g_2 \to g_1 \tag{6.1}$$

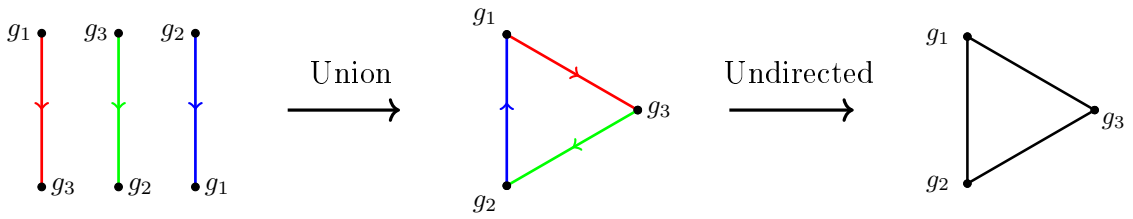
$$\text{Figure 6.12}b: \quad red: q \to g, \quad green: g \to \overline{q} \tag{6.2}$$

$$\text{Figure 6.12}c: \quad red: q \to \overline{q}. \tag{6.3}$$

The color flows represented as ordered lists in Eq. (6.1) - (6.3) will be the starting point for the next three sections where each of the sets of color flows will be represented as graphs which we will also call *color flows*.

6.2.3.1 Color flows in a $g_1 g_2 \to g_3$ vertex

We will first study the color flows of the $g_1 g_2 \to g_3$ vertex from Figure 6.12a with the three color flows red, green and blue given in Eq. (6.1). The color flows in Eq. (6.1) are themselves directed paths and are represented as graphs in Figure 6.13a. The nodes in the graphs are particles from the color flow, and the direction of the edges follows the arrows in Eq. (6.1). We will call the directed paths in Figure 6.13 *color flow* and denote them $\mathcal{C}_{red}$, $\mathcal{C}_{green}$ and $\mathcal{C}_{blue}$.



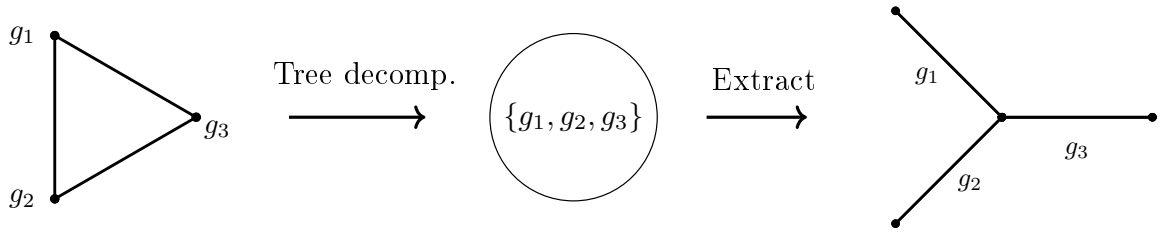
(a) $\mathcal{C}_{red}$, $\mathcal{C}_{green}$ and $\mathcal{C}_{blue}$ (b) $\mathcal{C} = \mathcal{C}_{red} \cup \mathcal{C}_{green} \cup \mathcal{C}_{blue}$ (c) Triangle C_3

Figure 6.13: The individual color flows (a), the union of them (b) and the undirected graph of the union (c).

The color flow graphs $\mathcal{C}_{red}$, $\mathcal{C}_{green}$ and $\mathcal{C}_{blue}$ share some of the particles in pairs. As an example, g_1 is contained in both $\mathcal{C}_{red}$ and $\mathcal{C}_{blue}$. Therefore, if we connect $\mathcal{C}_{red}$, $\mathcal{C}_{green}$ and $\mathcal{C}_{blue}$ together it gives us a new color flow graph $\mathcal{C}$ in Figure 6.13b from the union

$$\mathcal{C} = \mathcal{C}_{red} \cup \mathcal{C}_{green} \cup \mathcal{C}_{blue}. \tag{6.4}$$

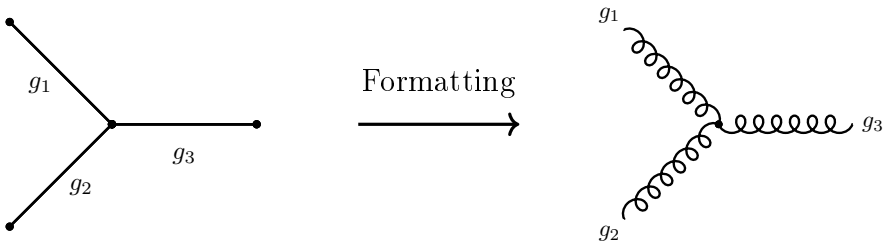
The color flow graph $\mathcal{C}$ is a directed triangle where the gluons g_1, g_2 and g_3 connect the color flow graphs from Figure 6.13a. The edges in $\mathcal{C}$ are directed in the same direction. Removing the direction and the color on the edges in the color flow graph in Figure 6.13b results in the undirected triangle C_3 in Figure 6.13c. In Section 6.2.2 we showed that a Feynman vertex could be represented as an undirected triangle C_3 like the one in Figure 6.13c. We can therefore use the tree decomposition of C_3 (Figure 6.14a) shown in Figure 6.14b and extract it to the star graph in Figure 6.14c.



(*a*) Triangle C_3 (*b*) Tree decomposition (*c*) Extracted tree

Figure 6.14: Tree decomposition (b) of the triangle C_3 (a) and the extracted graph of the tree decomposition (c).

The final step to obtain the Feynman vertex from the extracted tree is to add back the formatting of a Feynman vertex described in Section 6.2.2 and is shown in Figure 6.15.



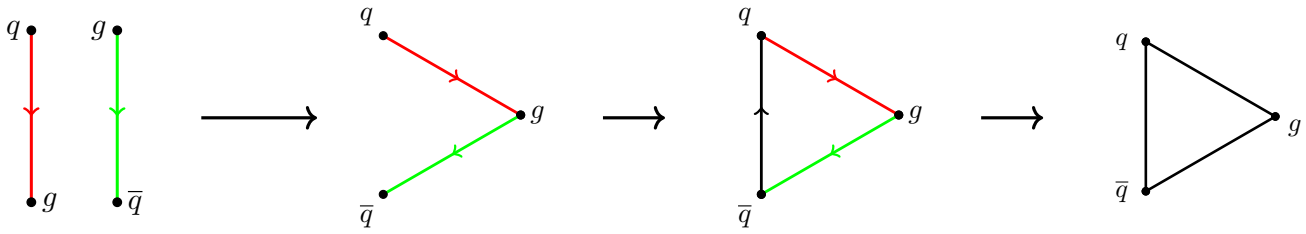
(*a*) Extracted tree (*b*) Feynman vertex $g_1 g_2 \to g_3$

Figure 6.15: The formatting from the extracted tree (a) to the Feynman vertex (b)

We have now shown how the color flows in Figure 6.13a can be combined together to the Feynman vertex in Figure 6.12a by using the methods we described in Section 6.2.1. We will do the same for the other two Feynman vertices in Figure 6.12b and 6.12c respectively, but will skip the details that are the same as in the 3-gluon vertex.

6.2.3.2 Color flows in a $q\overline{q} \to g$ vertex

The next vertex we will study is $q\overline{q} \to g$ in Figure 6.12b with the color flows red and green in Eq. (6.2). The color flow graphs $\mathcal{C}_{red}$ and $\mathcal{C}_{green}$ of the color flows are shown in Figure 6.16a.



(a) $\mathcal{C}_{red}$ and $\mathcal{C}_{green}$ (b) $\mathcal{C}_{red} \cup \mathcal{C}_{green}$ (c) $\mathcal{C} \cup \mathcal{G}_{\overline{q}q}$ (d) C_3

Figure 6.16: The individual color flows (a), the union of them (b), the union with gluon flow (c) and the undirected graph of the union (d).

The gluon g is a common particle in both $\mathcal{C}_{red}$ and $\mathcal{C}_{blue}$. We can therefore connect them to get a new color flow graph $\mathcal{C}$ shown in Figure 6.13b from the union

$$\mathcal{C} = \mathcal{C}_{red} \cup \mathcal{C}_{green}. \tag{6.5}$$

The color flow graph $\mathcal{C}$ has no triangular shape. However, we can make $\mathcal{C}$ triangular if we add a path from $\overline{q}$ to q such that all edges point in the same direction as shown in Figure 6.16c. We call this path a *gluon flow*,

$$\mathcal{G}_{\overline{q}q} = \overline{q} \to q, \tag{6.6}$$

since both quarks have the gluon g as a neighbor. From the color flow graph $\mathcal{C}$ in Figure 6.16c we see that g is the in-neighbor of $\overline{q}$ and the out-neighbor of q. If we use the notation from Section 6.2.1.4, we can be write it more compactly as

$$N_{\mathcal{C}}^{-}(\overline{q}) = g = N_{\mathcal{C}}^{+}(q). \tag{6.7}$$

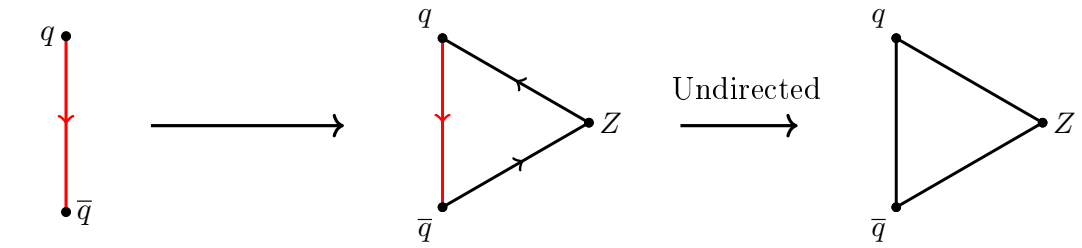
From Figure 6.16b we also see that $\overline{q}$ and q are in different color flows. We can now generalize the gluon flow graph $\mathcal{G}_{q_i q_j}$ between two quarks q_i and q_j in a color flow graph

$\mathcal{C}$ as

$$\mathcal{G}_{q_i q_j} = \begin{cases} q_i \to q_j & \text{, if } N_{\mathcal{C}}^{-}(q_i) = N_{\mathcal{C}}^{+}(q_j) \text{ and } q_i, q_j \notin \text{ same color flow} \\ \emptyset & \text{, otherwise,} \end{cases} \tag{6.8}$$

where $\emptyset$ is the empty set if the conditions are not satisfied. We can now remove the colors and directions of the edges of $\mathcal{C} \cup \mathcal{G}_{\bar{q}q}$ in Figure 6.16c to get the undirected triangle in Figure 6.16d. We have now arrived to the same stage as in Figure 6.13d in the $g_1 g_2 \to g_3$ vertex. The triangle in Figure 6.16d can be transformed to the vertex $q\bar{q} \to g$ in Figure 6.12b when we follow the same steps as in Section 6.2.3.1 in Figure 6.14 and 6.15. We have now showed how to go from the color flow to the Feynman diagram in the Figure 6.12b where we had to introduce the gluon flow in this section to triangulate the color flow graph.

6.2.3.2.1 Color flow in a $q\bar{q} \to Z$ vertex The final vertex $q\bar{q} \to Z$ in Figure 6.12c has only one color flow between the quarks in Eq. (6.3). The color flow $\mathcal{C}_{red}$ is shown in Figure 6.17a. The Z boson is a colorless particle and is therefore not included in any color flow. We can transform the color flow $\mathcal{C}_{red}$ into a triangle if we join the path $\bar{q} \to Z \to q$ and the color flow $\mathcal{C}_{red}$ that is shown in Figure 6.17b.



(a) $\mathcal{C}_{red}$ (b) $\mathcal{C}_{red} \cup \mathcal{B}_{\bar{q}q}$ (c) C_3

Figure 6.17: The individual color flow (a), the union with boson flow (b) and the undirected graph of the union (c).

We call this path a *boson flow*,

$$\mathcal{B}_{\bar{q}q} = \bar{q} \to Z \to q, \tag{6.9}$$

since the Z boson is a neighbor to both quarks. We notice that the in-neighborhood of $\bar{q}$ is q which we formally write

$$N_{\mathcal{C}}^{-}(\bar{q}) = q, \tag{6.10}$$

and that both quarks are in the same color flow. We can now generalize the boson flow graph $\mathcal{B}_{q_iq_j}$ between two quarks/antiquarks q_i and q_j in a color flow graph $\mathcal{C}$ as

$$\mathcal{B}_{q_iq_j} = \begin{cases} q_i \to Z \to q_j & \text{, if } N_{\mathcal{C}}^{-}(q_i) = q_j \text{ and } q_i, q_j \in \text{ same color flow} \\ \emptyset & \text{, otherwise} \end{cases} \tag{6.11}$$

where $\emptyset$ is the empty set if the conditions are not satisfied. We can now remove the colors and directions of the edges of $\mathcal{C}_{red} \cup \mathcal{B}_{\bar{q}q}$ in Figure 6.17b to get the undirected triangle in Figure 6.17c. We have now arrived to the same stage as in Figure 6.13d in the $g_1g_2 \to g_3$ vertex. The triangle in Figure 6.17d can be transformed to the vertex $q\bar{q} \to Z$ in Figure 6.12c when we follow the same steps as in Section 6.2.3.1 in Figure 6.14 and 6.15. We have now shown how to go from the color flow to the Feynman diagram in the Figure 6.12c where we had to introduce the boson flow in this section to triangulate the color flow graph.

6.2.4 Color flow in Feynman diagrams

We will now summarize what we learned from color flow graphs of the vertices and show how it generalizes to Feynman diagrams. We will use all the results obtained so far in the next section, applying them to two examples. In Section 6.2.3.1 we combined the color flows for the different vertices. We can combine the individual color flows $\mathcal{C}_i$ in a Feynman diagram to the total color flow

$$\mathcal{C} = \bigcup_{i \in colors} \mathcal{C}_i, \tag{6.12}$$

where all colors i are distinct since we are working in the leading-color approximation explained in Section 3.3.2. We can further write the individual color flows as

$$\mathcal{C}_i = p_{i_1} \to p_{i_2} ... \to p_{i_n}, \tag{6.13}$$

where all the particles $p_{i_1}, p_{i_2}, ..., p_{i_n}$ have the same color. In Section 6.2.3.2 we introduced the gluon flow graph $\mathcal{G}_{q_i q_j}$ in Eq. (6.8)

$$\mathcal{G}_{q_i q_j} = \begin{cases} q_i \to q_j & \text{, if } N_{\mathcal{C}}^{-}(q_i) = N_{\mathcal{C}}^{+}(q_j) \text{ and } q_i, q_j \notin \text{ same color flow} \\ \emptyset & \text{, otherwise.} \end{cases}$$

If we have a Feynman diagram with multiple gluon flows we can combine them to the total gluon flow

$$\mathcal{G} = \bigcup_{q_i, q_j \in Q} \mathcal{G}_{q_i q_j}, \tag{6.14}$$

where Q is the set of quarks in the Feynman diagram. Finally, we introduced the boson flow graph in the $q\bar{q} \to Z$, $\mathcal{B}_{q_i q_j}$ given in Eq. (6.11)

$$\mathcal{B}_{q_i q_j} = \begin{cases} q_i \to Z \to q_j & \text{, if } N_{\mathcal{C}}^{-}(q_i) = q_j \text{ and } q_i, q_j \in \text{ same color flow} \\ \emptyset & \text{, otherwise.} \end{cases}$$

If we have a Feynman diagram with multiple boson flow graphs we can combine them to the total boson flow

$$\mathcal{B} = \bigcup_{q_i, q_j \in Q} \mathcal{B}_{q_i q_j}, \tag{6.15}$$

where Q is the set of quarks in the Feynman diagram. We have now summarized the three different flows that make up what we will call the *particle flow*, as a union of color, gluon and boson flows

$$\mathcal{P} = \mathcal{C} \cup \mathcal{G} \cup \mathcal{B}. \tag{6.16}$$

The particle flow is a directed graph made up of triangles from the vertices in the Feynman diagram. The last step we need to do is to transform the particle flow graph to an undirected graph, and make a tree decomposition of it in the same way as we did for the vertices. We choose the bags of the tree decomposition to be the triangles in the undirected particle flow graph. The tree graph is then extracted as we did previously and formatted such that we get back to the Feynman diagram that the color flows corresponds. We have now presented the full framework for how we can use the color flows from an event record to reconstruct the Feynman diagrams. We can now apply these tools to some examples.

6.2.5 Examples

We will use two different Feynman diagrams as examples to show how their color flows can be used to reconstruct the Feynman diagram. The first Feynman diagram is a t-channel diagram. In the second example we will look at a more complex Feynman diagram with a Z boson that is taken from Figure 6.2 showing the possible Feynman diagrams of the Sherpa Z + jets signal process in Figure 6.1.

6.2.5.1 t-channel

We will start with a simple example of the t-channel diagram in Figure 6.18 with the color flows red and green.

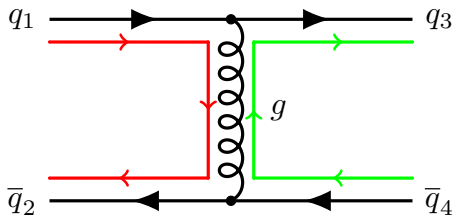
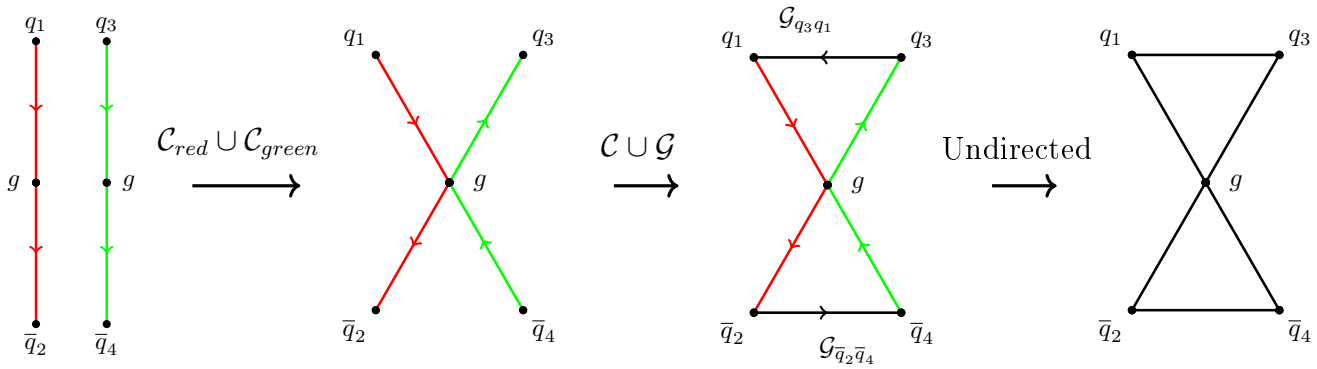


Figure 6.18: $q_1\overline{q}_2 \to q_3\overline{q}_4$

Figure 6.19 shows the individual color flows $\mathcal{C}_{red}$ and $\mathcal{C}_{green}$. We can combine the color flows $\mathcal{C}_{red}$ and $\mathcal{C}_{green}$ to the total color flow $\mathcal{C} = \mathcal{C}_{red} \cup \mathcal{C}_{green}$ and this is shown in Figure 6.19b.



(a) $\mathcal{C}_{red}$ and $\mathcal{C}_{green}$ (b) $\mathcal{C} = \mathcal{C}_{red} \cup \mathcal{C}_{green}$ (c) $\mathcal{C} \cup \mathcal{G}$ (d) Undirected

Figure 6.19: The individual color flows (a), the union of them (b), the union with gluon flow (c) and the undirected graph of the union (d).

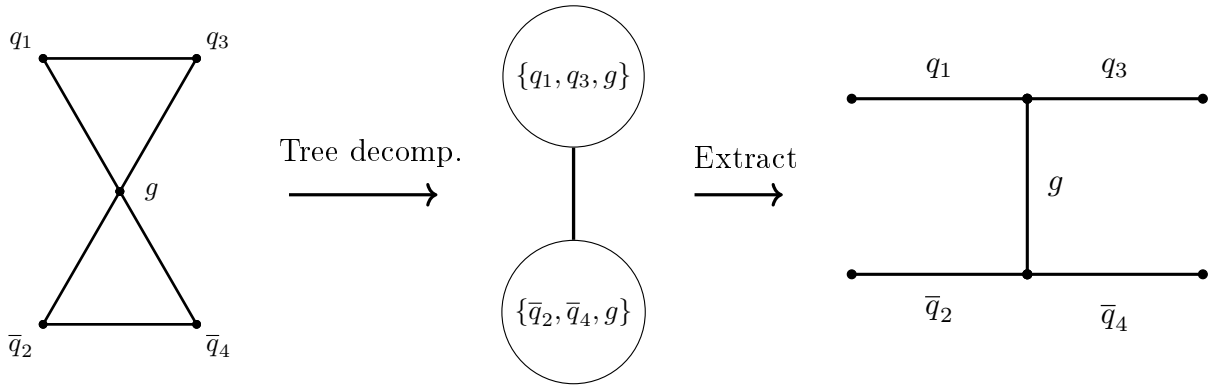
In Figure 6.19b we find the two gluon flows

$$\mathcal{G}_{q_3q_1} = q_3 \to q_1, \text{ since } N_{\mathcal{C}}^{-}(q_3) = N_{\mathcal{C}}^{+}(q_1) \text{ and } q_1 \in \mathcal{C}_{red}, q_3 \in \mathcal{C}_{green}, \tag{6.17}$$

$$\mathcal{G}_{\overline{q}_2\overline{q}_4} = \overline{q}_2 \to \overline{q}_4, \text{ since } N_{\mathcal{C}}^{-}(\overline{q}_2) = N_{\mathcal{C}}^{+}(\overline{q}_4) \text{ and } \overline{q}_2 \in \mathcal{C}_{red}, \overline{q}_4 \in \mathcal{C}_{green}, \tag{6.18}$$

that we can join together to the total gluon flow $\mathcal{G} = \mathcal{G}_{q_3q_1} \cup \mathcal{G}_{\overline{q}_2\overline{q}_4}$.

There are no boson flows in Figure 6.19b so the particle flow becomes $\mathcal{P} = \mathcal{C} \cup \mathcal{G}$ and is shown in Figure 6.19c where the gluon flows $\mathcal{G}_{q_3q_1}$ and $\mathcal{G}_{\overline{q}_2\overline{q}_4}$ are drawn. The particle flow in Figure 6.19c is a directed graph with two triangles. Figure 6.19d shows that it is an undirected graph. We can now do the tree decomposition of the graph in Figure 6.20a (like we did in Figure 6.7) and is shown in 6.20b. The extracted tree of the tree decomposition is shown in Figure 6.20c.

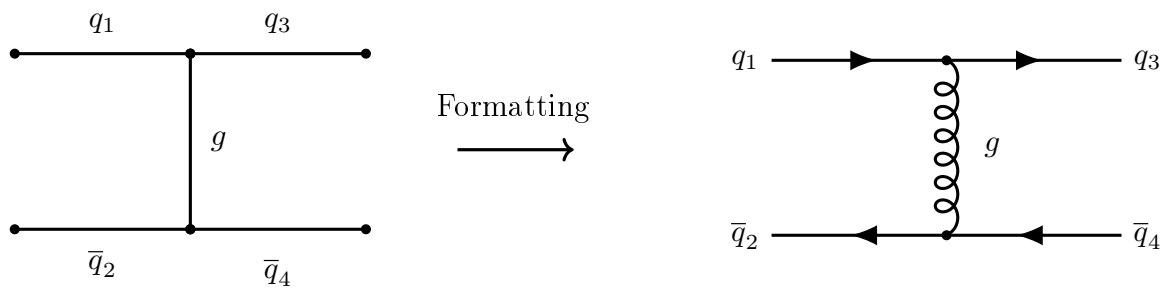


(a) Undirected (b) Tree decomposition (c) Extracted tree

Figure 6.20: The tree decomposition (b) of the undirected (a), and the extracted graph of the tree decomposition (c).

The final step to reconstruct the Feynman diagram from the color flows in 6.19a is to apply the formatting described in Section 6.2.2 of the graph in Figure 6.21a. This gives the Feynman diagram in Figure 6.21b.

The Feynman diagram in Figure 6.21 is exactly the diagram that has the color flows that we started out with in Figure 6.18. We have therefore reconstructed the Feynman diagram from the color flows in Figure 6.19a.



(*a*) Extracted tree

(*b*) Feynman diagram

Figure 6.21: The formatted Feynman diagram (b) from the extracted tree (a).

6.2.5.2 Sherpa Z + jets signal process

We started this chapter with the aim of finding a method to reconstruct the Feynman diagram from its color flows, and thereby finding the Z production vertex in that Feynman diagram. We are now ready to find the Z boson production vertex in the Feynman diagram $dg \to dggZ$ in Figure 6.2 that was one of the signal process for a Sherpa Z + jets sample. In Figure 6.22 this Feynman diagram is drawn with its color flows.

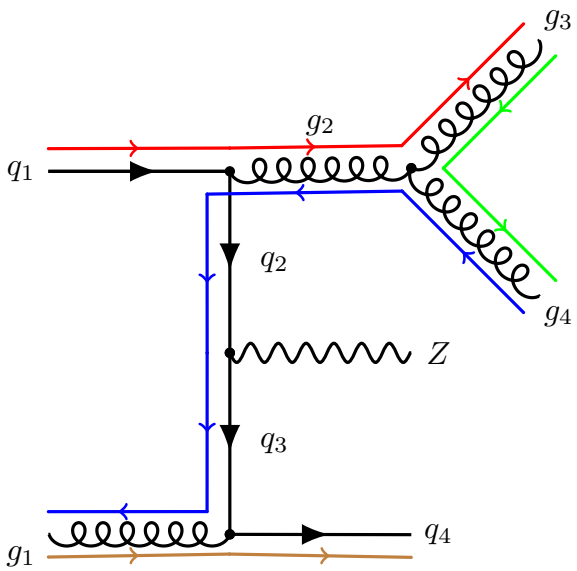
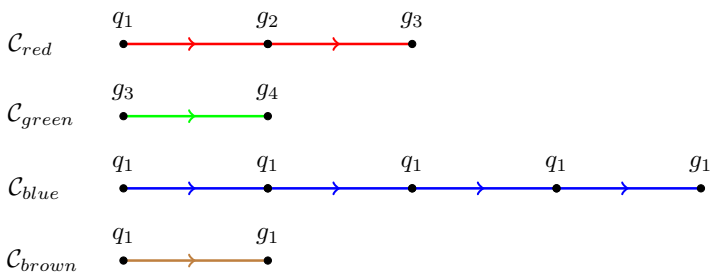
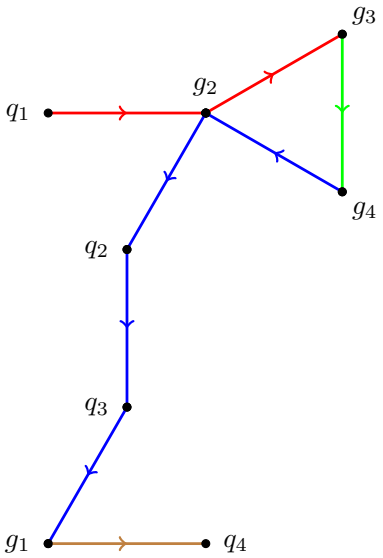


Figure 6.22: $q_1 g_1 \to q_4 g_3 g_4 Z$

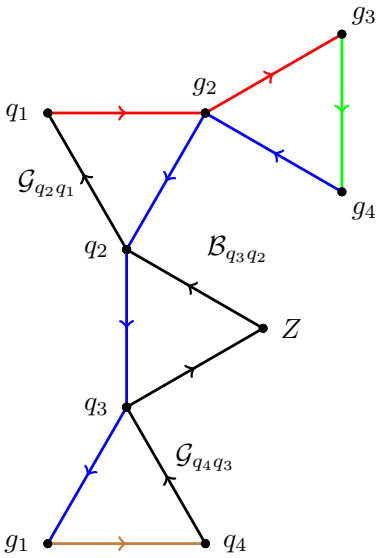
Figure 6.23 shows the individual color flows $\mathcal{C}_{red}$, $\mathcal{C}_{green}$, $\mathcal{C}_{blue}$ and $\mathcal{C}_{brown}$. We use the leading color approximation and the color *brown* is therefore one of these colors since every color becomes unique. We can combine these color flows to the total color flow $\mathcal{C} = \mathcal{C}_{red} \cup \mathcal{C}_{green} \cup \mathcal{C}_{blue} \cup \mathcal{C}_{brown}$ and is shown in Figure 6.23b.



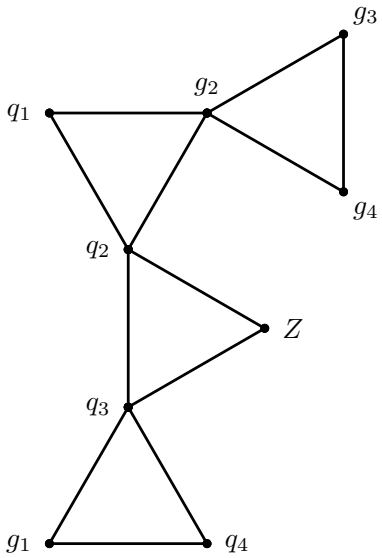
(a) $\mathcal{C}_{red}, \mathcal{C}_{green}, \mathcal{C}_{blue}, \mathcal{C}_{brown}$



(b) $\mathcal{C}$



(c) $\mathcal{C} \cup \mathcal{G} \cup \mathcal{B}$



(d) Unidirected

Figure 6.23: The individual color flows (a), the union of them (b), the union with gluon flow and boson flow (c) and the undirected graph of the union (d).

In Figure 6.23b we find the two gluon flows

$$\mathcal{G}_{q_2q_1} = q_2 \to q_1, \text{ since } N_{\mathcal{C}}^{-}(q_2) = g_2 = N_{\mathcal{C}}^{+}(q_1) \text{ and } q_2 \in \mathcal{C}_{blue}, q_1 \in \mathcal{C}_{red}, \tag{6.19}$$

$$\mathcal{G}_{q_4q_3} = q_4 \to q_3, \text{ since } N_{\mathcal{C}}^{-}(q_4) = g_3 = N_{\mathcal{C}}^{+}(q_3) \text{ and } \overline{q}_4 \in \mathcal{C}_{brown}, \overline{q}_3 \in \mathcal{C}_{blue}, \tag{6.20}$$

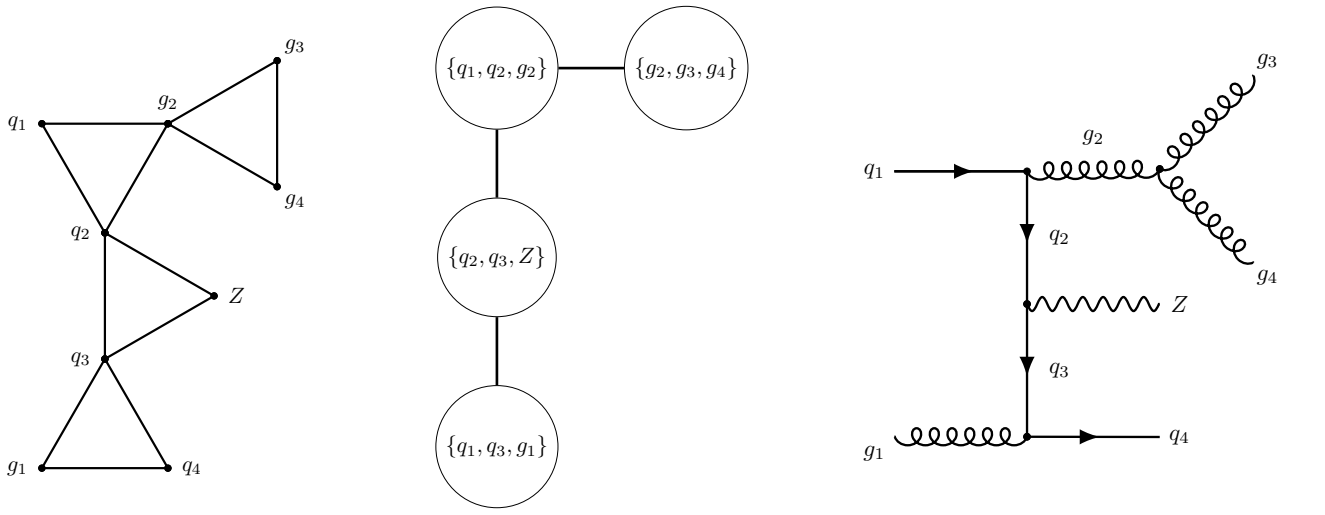
that we can join together to the total gluon flow $\mathcal{G} = \mathcal{G}_{q_3q_1} \cup \mathcal{G}_{\overline{q}_2\overline{q}_4}$. We also find one unique boson flow

$$\mathcal{B}_{q_3q_2} = q_3 \to Z \to q_2, \text{ since } N_{\mathcal{C}}^{-}(q_3) = q_2 \text{ and } q_2, q_3 \in \mathcal{C}_{blue}, \tag{6.21}$$

and the total boson flow is therefore $\mathcal{B} = \mathcal{B}_{q_2q_3}$. The particle flow becomes $\mathcal{P} = \mathcal{C} \cup \mathcal{G}$

where the gluon flows $\mathcal{G}_{q_2q_3}$, $\mathcal{G}_{q_4q_3}$ and boson flow $\mathcal{B}_{q_3q_2}$ are drawn. The particle flow in Figure 6.19c is a directed graph with four triangles. Figure 6.23d shows its undirected graph. We can now do the tree decomposition of the graph in Figure 6.24a and this is shown in 6.24b. The extracted tree of the tree decomposition is shown in Figure 6.24c which also is formatted to a Feynman diagram.

The Feynman diagram in Figure 6.24b is exactly the diagram that has the color flows that we started out with in Figure 6.18. We have therefore reconstructed the Feynman diagram from the color flows in Figure 6.19a, and the Z production vertex is $q_2q_3 \to Z$, and is the only place it can be, which was what we wanted to achieve.



(*a*) Undirected (*b*) Tree decomposition (*c*) Extracted graph with formatting

Figure 6.24: The tree decomposition (b) of the undirected (a), and the extracted graph of the tree decomposition with formatting (c).

6.2.6 Summary

To summarize we have demonstrated a method that uses the QCD color flows to unambiguously reconstruct Feynman diagrams by using different methods from graph theory. The Z production vertex used in the signal reweighting method in Section 5.1.2 can therefore be found once we have the color flows of the signal process.

6.3 Color flows in the Sherpa event record

When investigating the color flows in the event record in Sherpa we discovered that the intermediate particles between the incoming and outgoing particles in the signal process event vertex were not included in the color flow. Instead the color flows in Sherpa contain only the incoming and outgoing particles in the event vertex. Color flows with only the incoming and outgoing particles in the event vertex do not contain sufficient information to reconstruct the Feynman diagram unambiguously. This means that there is insufficient information in the color flows from the Sherpa event record to be able to use the explicitly method described above to reconstruct the Feynman diagram, and thus find the Z production vertex which is needed for the signal reweighting method.

The many possible Feynman diagrams that can represent a Sherpa Z + jets signal process makes it therefore difficult to find the Z production vertex.

Chapter 7

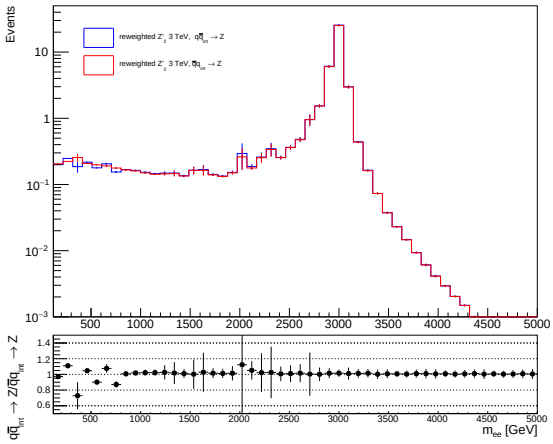
Results

In this chapter we will present and discuss the results following the signal reweighting of PowhegPythia8 Drell-Yan distributions to Z' distributions. In the first section we will investigate the two ambiguities in the PowhegPythia8 signal reweighing method in described in Section 5.3, namely the use of Z_q or Z_l in Eq. (5.13) and (5.14), and the choice of t-channel diagram in Figure 5.13a and 5.13b. After that we will validate the signal reweighting with different kinematical distributions between the signal reweighted PowhegPythia8 Z'_χ and the only available simulated Pythia8 Z'_χ sample described in Section 4.3.1. The QCD k-factors from Section 4.3.4 will be applied to every distribution such that we compare both the signal reweighted and simulated Z' distributions in $NNLO$ order. We will also look at different Z' models that we will compared to each other to see how to differentiate between various Z' models. In the final section Run 3 data will be compared to background simulations with several signal reweighted Z' models on top.

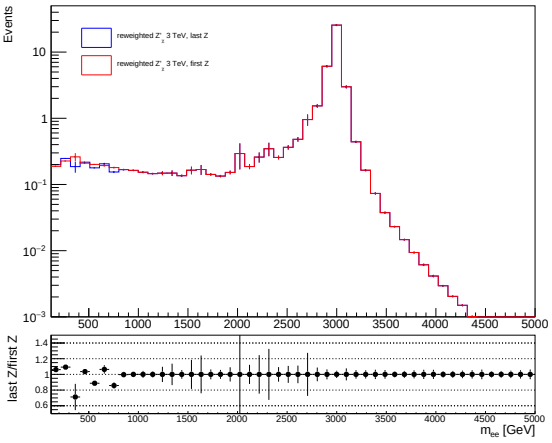
7.1 Comparison between different configurations in the PowhegPythia8 signal reweighting

In Section 5.3.4 we discussed the two possible diagrams that can be used to calculate the scale factor in the t-channel process $q\bar{q} \to gZ$. The two Feynman diagrams are shown in Figure 5.13a and 5.13b, page 51. The Z production vertex for these two diagrams are $q\bar{q}_{int} \to Z$ and $\bar{q}q_{int} \to Z$ respectively. In Figure 7.1a we compare the reweighted Z' invariant mass distributions when we use the diagram in Figure 5.13a or 5.13b to calculate the scale factor given in Eq. (5.2), page 42. We see good agreement when

swapping the quark and antiquark in Figure 5.13a and 5.13b to calculate the scale factor. We also discussed in Section 5.3.4 which Z boson we could use from the event record to calculate the scale factor. Figure 7.1b shows the Z' invariant mass comparison when we choose either the Z_q boson from the $Zq\overline{q}$ vertex or the Z_l boson that decay to the l^+l^- pair in the event record. We see an overall good agreement between the Z' distribution with the only difference being in the lower masses.



(a) Comparison between the reweighted Z' distributions where the diagram in Figure 5.13a (blue) or 5.13b (red) has been used to calculate the scale factor.



(b) Comparison between the reweighted Z' distributions where the Z boson from the signal process (blue) or the Z boson from the decays to leptons (red) has been used to calculate the scale factor.

Figure 7.1: Comparisons of reweighted PowhegPythia8 Z' distributions at 3 TeV where different diagrams (a) and Z bosons (b) has been used to calculate the scale factor

The only visible discrepancy in Figure 7.1 is at low invariant mass where there are some fluctuations between these two histograms. We can conclude based on the comparisons in Figure 7.1a and 7.1b that the choice of the diagram in the signal process $q\overline{q} \to gZ$, and Z boson in the event record has little influence on the results of the signal reweighted Z' distributions. From here on we will calculate the scale factor from the diagram in Figure 5.13a for the events with the signal process $q\overline{q}_{int} \to gZ$, and use the Z boson from the decay to the lepton pair, and has initial-state radiation included.

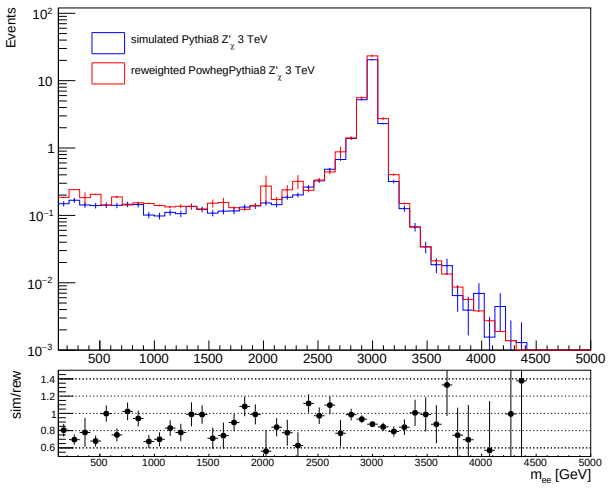
7.2 Validation in the electron channel

We validate the signal reweighted Z' distributions by comparing the the invariant mass, transverse momentum and pseudorapidity distributions for the signal reweighted and the simulated Z' samples. The samples we use for the validation is the Z' signal simulated sample from Pythia8 in Section 4.3.1 and the PowhegPyhia8 Drell-Yan sample in Section 4.3.2 used for the Z' signal reweighting. We do the comparison with samples from the $mc16e$ period with 58.5 fb^{-1} with the Z'_χ model since it is the only available Z' signal simulated Pythia8 sample. The k-factors described in Section 4.3.4 will also be applied. One way we can test the comparison between two histograms is with a Kolmogorov–Smirnov test [53]. The Kolmogorov-Smirnov test is used to determine how likely it is that two histograms have been sampled from the same underlying probability distribution. The Kolmogorov-Smirnov test returns a p-value being uniformly distributed under a true hypothesis and takes therefore values between 0 and 1. Another method we can use to compare two histograms is to fit a function to each of the histograms and compare the fit parameters and results. We fit the invariant histogram using the double-sided Crystal ball function [54].

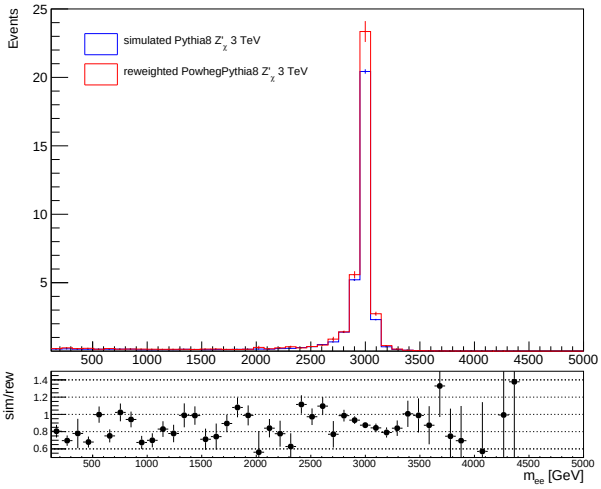
7.2.1 Comparison of kinematical variables

We will first compare the signal reweighted Z' invariant mass distribution from the PowhegPythia8 Drell-Yan sample to a fully simulated Pythia8 Z' sample in the electron channel. We will first see the comparison when no k-factors are applied and is shown in Figure 7.2. The fully simulated Z'_χ distribution is therefore in LO while the signal reweighted Z'_χ distribution is NLO in QCD. In Figure 7.3 the Kolmogorov-Smirnov test gives a p-value of 0.9900, and the two histograms are therefore likely to be compatible since the probability of getting a lower p-value from a pair of histograms sampled from the same distribution is very high. We see however that the signal reweighted Z'_χ distribution is consistently above the fully simulated Z'_χ. We will therefore apply the k-factors from Section 4.3.4 such that both the fully simulated and the signal reweighted Z' distributions have NNLO order in QCD. Figure 7.3 shows this comparison of the invariant mass distributions between the signal reweighted and fully simulated Z'_χ where both are in NNLO order in QCD. The ratio plot in Figure 7.3 shows some fluctuations

between the two histograms, but there is no trend that one distribution is consistently above the other.

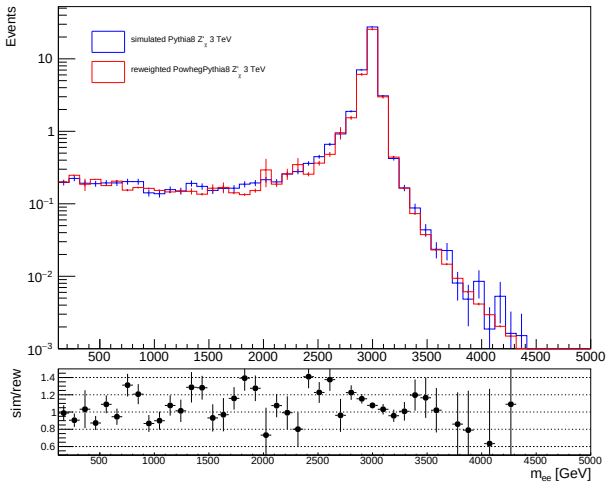


(a) Comparison between simulated (blue) and signal reweighted (red) Z'_χ invariant mass distributions (log)

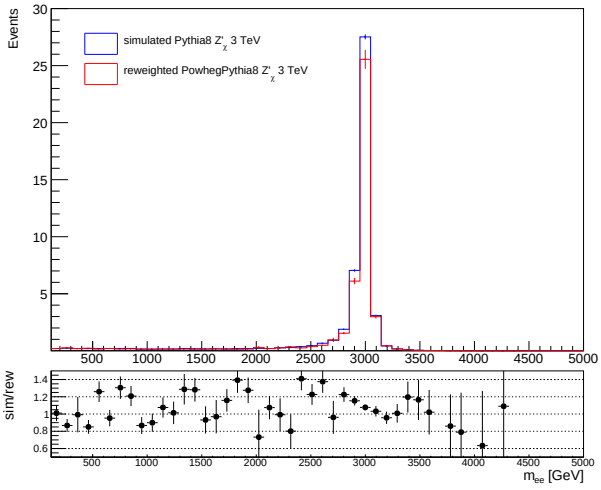


(b) Comparison between simulated (blue) and signal reweighted (red) Z'_χ invariant mass distributions (lin)

Figure 7.2: Comparison between simulated (blue) (LO) and signal reweighted (red) (NLO) Z'_χ invariant mass distributions in linear (a) and logarithmic (b) scale.



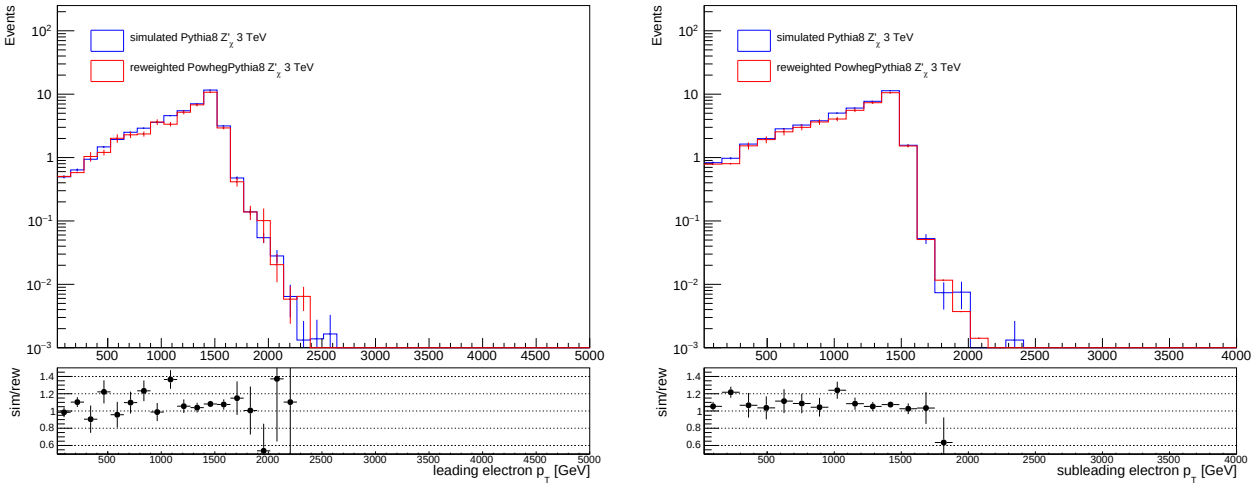
(a) Comparison between simulated (blue) and signal reweighted (red) Z'_χ invariant mass distributions (log)



(b) Comparison between simulated (blue) and signal reweighted (red) Z'_χ invariant mass distributions (lin)

Figure 7.3: Comparison between simulated (blue) (NNLO) and signal reweighted (red) (NNLO)Z'_χ invariant mass distributions in linear (a) and logarithmic (b) scale.

In Figure 7.3 the Kolmogorov-Smirnov test gives a p-value of 0.9808, and the two histograms are therefore likely to be compatible since the probability that they are not compatible is tiny, $1-0.9809 = 0.0191$. From now on we will apply the k-factors on other kinematical variables since it was a better agreement and they are compared at the same order. Figure 7.4a and 7.4b show the electron leading and subleading p_T distributions. The Kolmogorov-Smirnov test for these histograms gives the p-values 0.8841 and 0.9824, respectively.

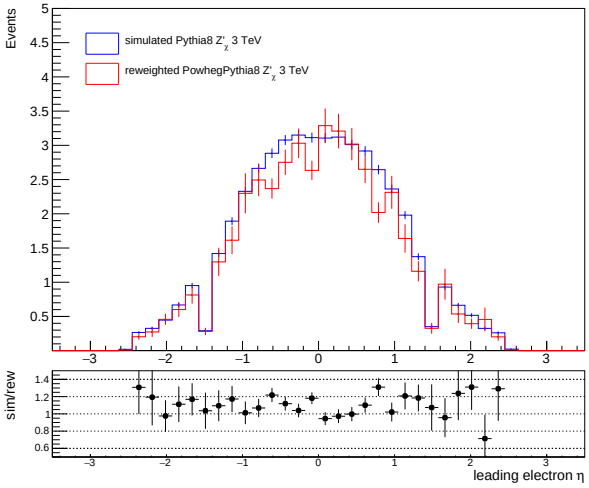


(a) Comparison between simulated (blue) and signal reweighted (red) Z'_χ leading electron p_T distributions.

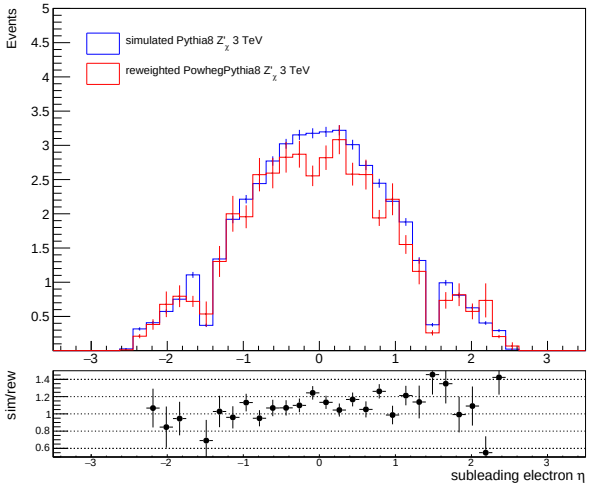
(b) Comparison between simulated (blue) and signal reweighted (red) Z'_χ subleading electron p_T distributions.

Figure 7.4: Comparison of Z'_χ leading (a) and electron subleading (b) p_T distributions.

The final kinematical variable we will test is the pseudo rapidity η. Figure 7.5a and 7.5b shows the leading and subleading η distributions. The Kolmogorov-Smirnov test for these histograms gives the p-values 0.9343 and 0.6703 respectively. In Figure 7.5b there is a discrepancy at $\eta = 0$, that is the muons in the transverse plane.



(a) Comparison between simulated (blue) and signal reweighted (red) Z'_χ leading η distributions



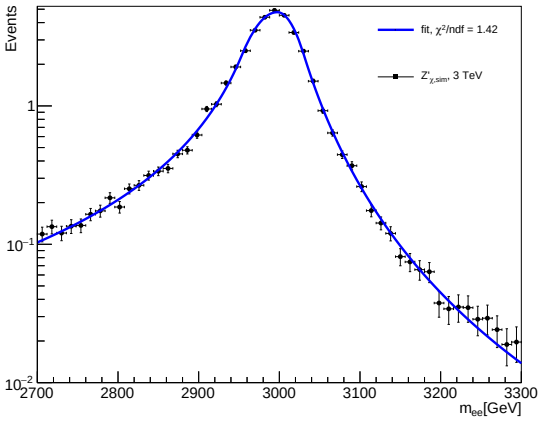
(b) Comparison between simulated (blue) and signal reweighted (red) Z'_χ subleading η distributions

Figure 7.5: Comparison of Z'_χ leading (a) and subleading (b) η distributions

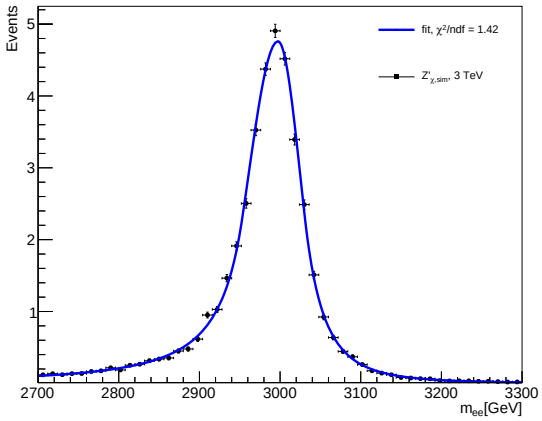
7.2.2 Fits of invariant mass distribution

We will now use the double-sided Crystal Ball function to do a fit of both the simulated and the reweighted Z'_χ invariant mass distributions. Figure 7.6a and 7.6b shows the fit of the simulated Z'_χ distribution in logarithmic and linear scale, respectively, around a Z' mass of 3 TeV. A chi-squared test between the simulated Z'_χ distribution and the fitted function gives $\chi^2/ndf = 1.42$.

We can get the pull distribution of the fit if we calculate the difference between the value of the simulated Z'_χ and the fit in every bin. The normalized pull distribution is called the residual distribution. Figure 7.7a and 7.7b show the pull and residual distributions respectively. From the residual distribution in Figure 7.7a we see that there are some discrepancies around the Z'_χ peak between the fit and the simulated Z'_χ distribution that is also visible in the linear fit in Figure 7.6b. To get a better fit we could have used different combinations of e.g a Crystal Ball and a Gaussian function to take into account the different resolutions in different parts of the ATLAS detector.

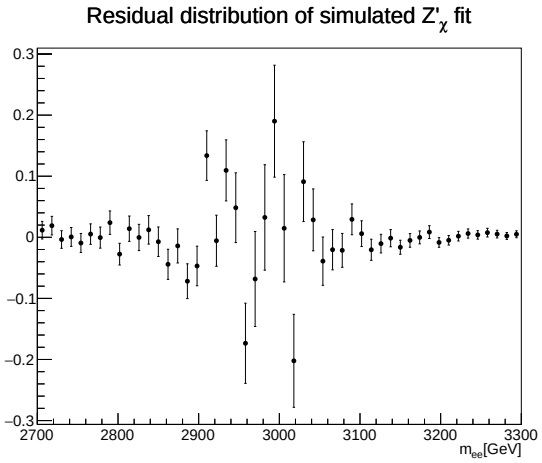


(a) Double-sided Crystal Ball fit of simulated Z'_χ invariant mass distribution (log).

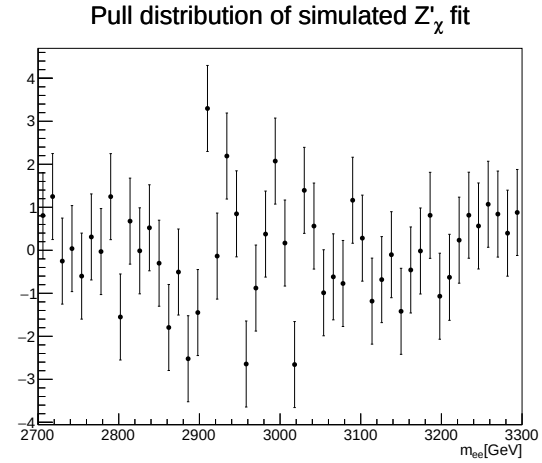


(b) Double-sided Crystal Ball fit of simulated Z'_χ invariant mass distribution (lin).

Figure 7.6: Double-sided Crystal Ball fit of simulated Z'_χ invariant mass distribution in logarithmic (a) and linear (b) scale.



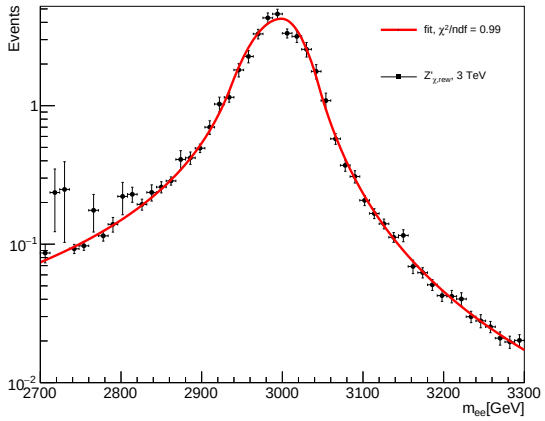
(a) Residual distribution of simulated Z'_χ invariant mass fit.



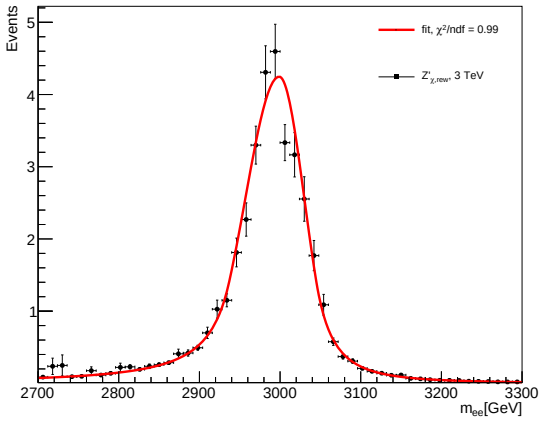
(b) Pull distribution of simulated Z'_χ invariant mass fit.

Figure 7.7: Residual (a) and pull (b) distribution of simulated Z'_χ invariant mass fit compared to the histogram.

We use the Crystal Ball function again to fit the signal reweighted Z'_χ invariant mass distribution. Figure 7.8a and 7.8b shows the fit result in logarithmic and linear scale respectively. The Chi-squared test of the fit gives $\chi^2/ndf = 0.99$. Figure 7.9a and 7.9b shows the residual and pull distribution of the fit. We see again the discrepancy around the Z'_χ peak between the fit and the simulated Z'_χ distribution in the residual distribution that is also visible in the linear scale in Figure 7.8b.

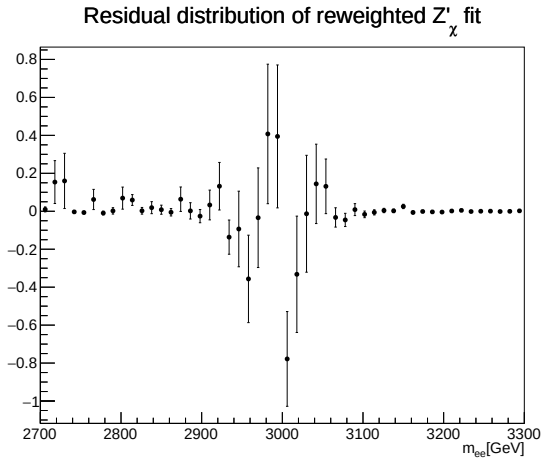


(a) Double-sided Crystal Ball fit of reweighted Z'_χ invariant mass distribution (log)

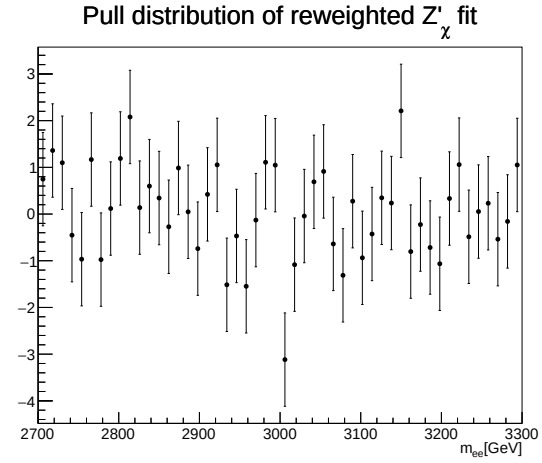


(b) Double-sided Crystal Ball fit of reweighted Z'_χ invariant mass distribution (lin)

Figure 7.8: Double-sided Crystal Ball fit of the reweighted Z'_χ invariant mass distribution in logarithmic (a) and linear (b) scale.



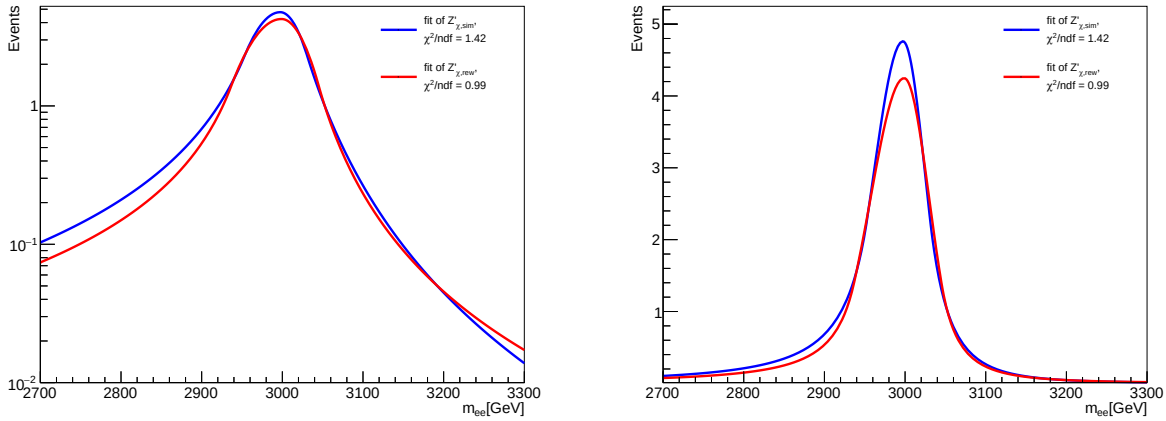
(a) Residual distribution of the reweighted Z'_χ invariant mass fit.



(b) Pull distribution of the reweighted Z'_χ invariant mass fit.

Figure 7.9: Residual (a) and pull (b) distribution of the reweighted Z'_χ invariant mass fit.

The final comparison in the electron channel is between the fit of the simulated and signal reweighted Z'_χ distribution. Figure 7.10a and 7.10b show the comparison in logarithmic and linear scale. We see that the fits do not completely overlap and one reason for this might be the fits fail to describe in Z'_χ peak as discussed for the simulated and signal reweighted fits.



(a) Comparison between fit of the simulated and reweighted Z'_χ invariant mass distribution (log)

(b) Comparison between fit of simulated and reweighted Z'_χ invariant mass distribution

Figure 7.10: Comparison between fit of simulated and reweighted Z'_χ invariant mass distribution

7.3 Validation in the muon channel

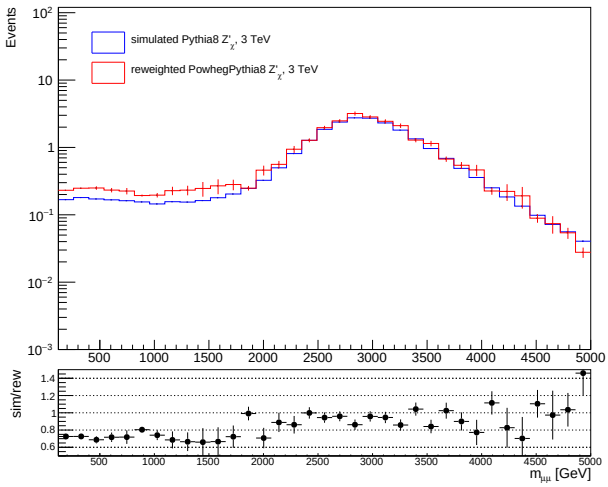
We validate the signal reweighted Z' distributions by comparing the the invariant mass, transverse momentum and pseudorapidity distributions for the signal reweighted and the simulated Z' samples. The samples we use for the validation is the Z' signal simulated sample from Pythia8 in Section 4.3.1 and the PowhegPythia8 Drell-Yan sample in Section 4.3.2 used the Z' signal reweighting. We do the comparison with samples from the $mc16e$ period with 58.5 fb^{-1} with the Z'_χ model since it is the only available Z' signal simulated Pythia8 sample. The k-factors described in Section 4.3.4 will also be applied.

7.3.1 Comparison of kinematical variables

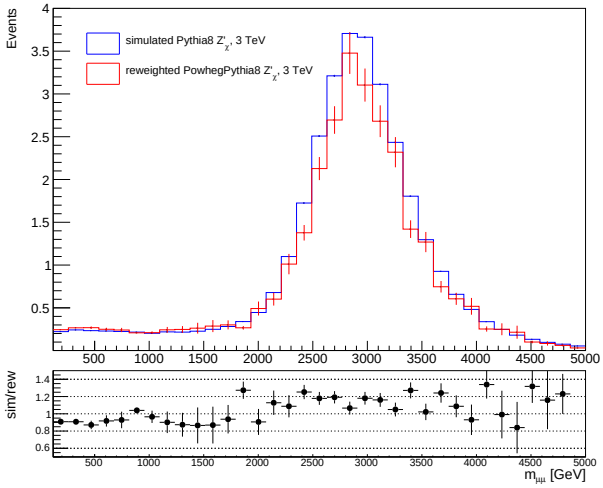
We will now compare the signal reweighted Z' from the PowhegPythia8 Drell-Yan sample to a simulated Pythia8 Z' sample in the muon channel.

We will first see the comparison when no k-factors are applied and is shown in Figure 7.11. The fully simulated Z'_χ distribution is therefore in LO while the signal reweighted Z'_χ distribution is NLO in QCD. In Figure 7.11 the Kolmogorov-Smirnov test gives a p-value of 0.0401, and the two histograms are therefore not likely to be compatible since the probability is tiny. We see that the signal reweighted Z'_χ distribution is consistently

above the fully simulated Z'_χ. We will therefore apply the k-factors from Section 4.3.4 such that both the fully simulated and the signal reweighted Z' distributions have NNLO order in QCD.



(a) Comparison between simulated (blue) and signal reweighted (red) Z'_χ invariant mass distributions (log)

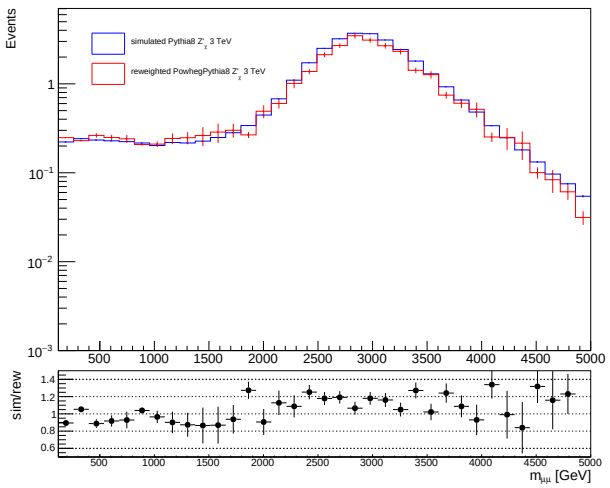


(b) Comparison between simulated (blue) and signal reweighted (red) Z'_χ invariant mass distributions (lin)

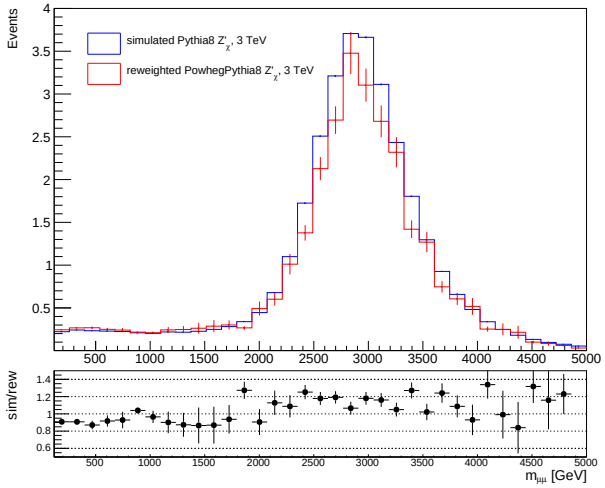
Figure 7.11: Comparison between simulated (blue) (LO) and signal reweighted (red) (NLO) Z'_χ invariant mass distributions in linear (a) and logarithmic (b) scale.

Figure 7.12 shows this comparison of the invariant mass distributions between the signal reweighted and fully simulated Z'_χ where both are in NNLO order in QCD. The ratio plot in Figure 7.12 shows some fluctuations between the two histograms, but there is no trend that one distribution is consistently above the other. In Figure 7.12 the Kolmogorov-Smirnov test gives a p-value of 0.2300, and the two histograms are therefore be compatible since the probability is not too low.

From now on we will apply the k-factors on other kinematical variables since it was a better agreement and they are compared at the same order Figure 7.13a and 7.13b shows the leading and subleading p_T distributions. The Kolmogorov-Smirnov test for these histograms gives the values 0.9035 and 0.2320 respectively. There is some discrepancies at high p_T in both the muon leading and subleading p_T distribution and is biggest for the subleading p_T distribution. KS also gives a lower p-value for the muon subleading compared to the leading p_T distribution.

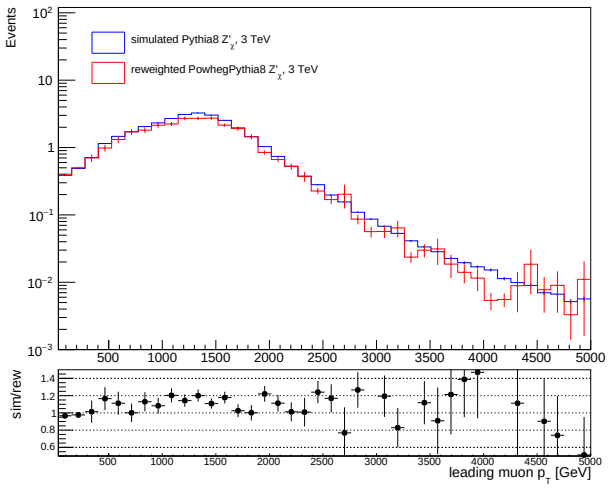


(a) Comparison between simulated (blue) and signal reweighted (red) Z'_χ invariant mass distributions (log)

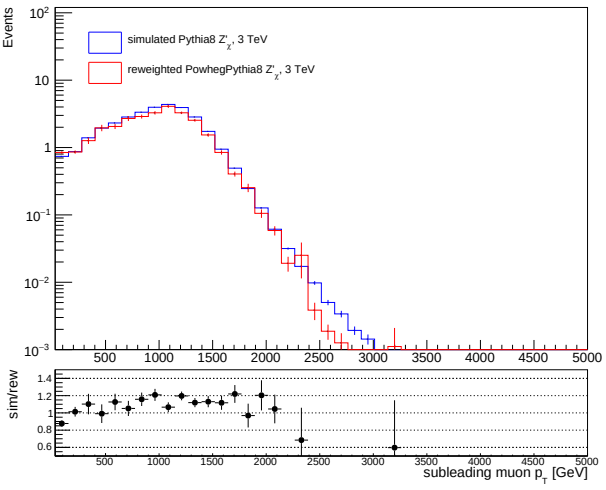


(b) Comparison between simulated (blue) and signal reweighted (red) Z'_χ invariant mass distributions (lin)

Figure 7.12: Comparison between simulated (blue) (NNLO) and signal reweighted (red) (NNLO) Z'_χ invariant mass distributions in linear (a) and logarithmic (b) scale.



(a) Comparison between simulated (blue) and signal reweighted (red) Z'_χ leading p_T distributions.

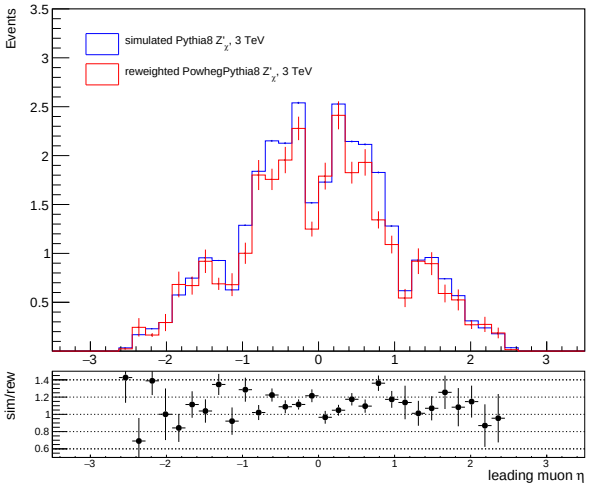


(b) Comparison between simulated (blue) and signal reweighted (red) Z'_χ subleading p_T distributions.

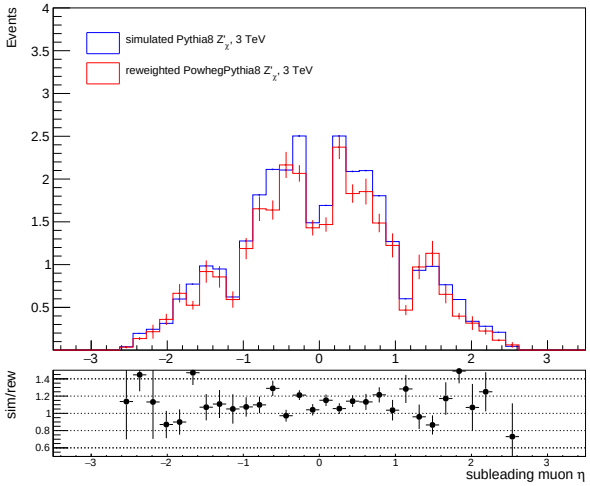
Figure 7.13: Comparison of Z'_χ leading (a) and subleading (b) p_T distributions.

The final kinematics variable we will test is the pseudo rapidity η. Figure 7.14a and 7.14b show the muon leading and subleading η distributions. The Kolmogorov-Smirnov test for these histograms gives the values 0.8346 and 0.9936 respectively, and shows good

agreement overall.



(a) Comparison between simulated (blue) and signal reweighted (red) Z'_χ leading η distributions.



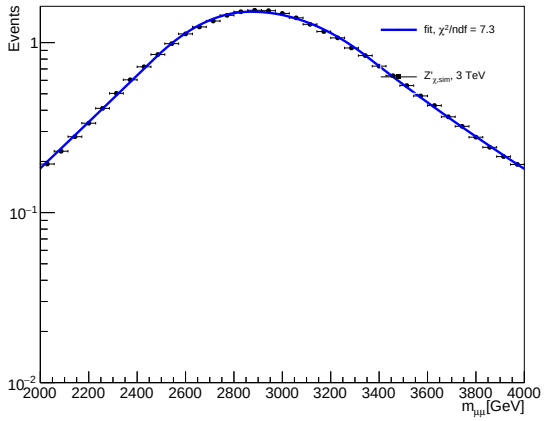
(b) Comparison between simulated (blue) and signal reweighted (red) Z'_χ subleading η distributions.

Figure 7.14: Comparison of Z'_χ leading (a) and subleading (b) η distributions.

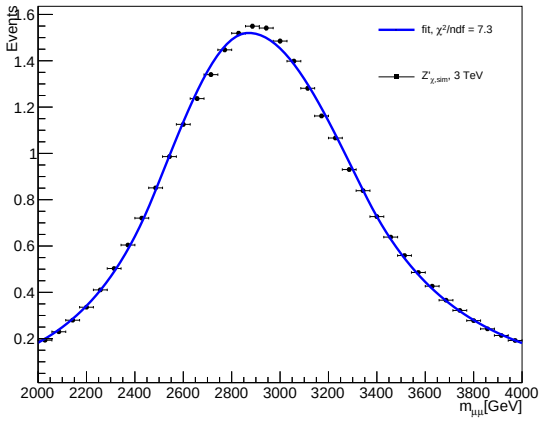
7.3.2 Fits of invariant mass distribution

We fit the histograms with a Crystal Ball function again. Figure 7.15a and 7.15b show the fit in logarithmic and linear scale. The fit is not so good and gives a $\chi^2/ndf = 7.3$. One factor for the large χ^2/ndf comes from the relative small uncertainties that we can explain from the large statistics in the simulated Pythia8 Z'_χ sample shown in Appendix A.1.

The residual and pull distributions in Figure 7.16a and 7.16b show some structure which is also an indication that the fit is not so good. We see that there is a shift between the fit and invariant mass distribution that explains the structure in Figure 7.15b which leads to the structure in the residual and pull distributions in Figure 7.16. We can explain some of the discrepancies of the fit with the difference in muon resolution in the endcap and barrel region of the ATLAS detector.

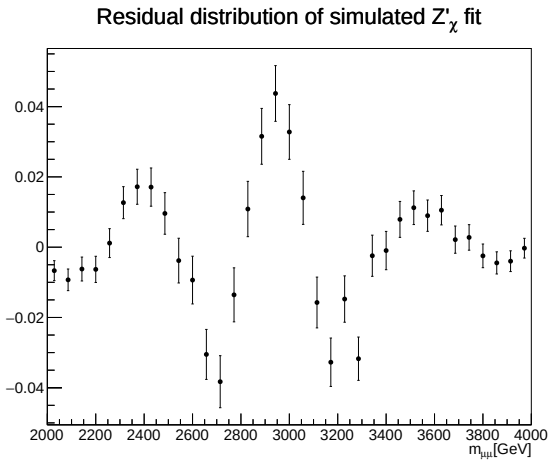


(a) Double-sided Crystal Ball fit of simulated Z'_χ invariant mass distribution (log)

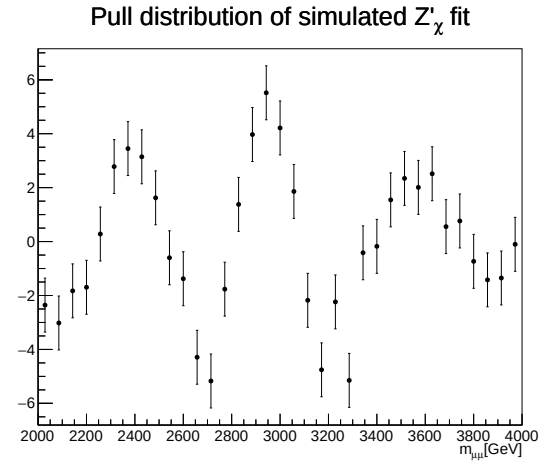


(b) Double-sided Crystal Ball fit of simulated Z'_χ invariant mass distribution (lin)

Figure 7.15: Double-sided Crystal Ball fit of simulated Z'_χ invariant mass distribution in logarithmic (a) and linear (b) scale.



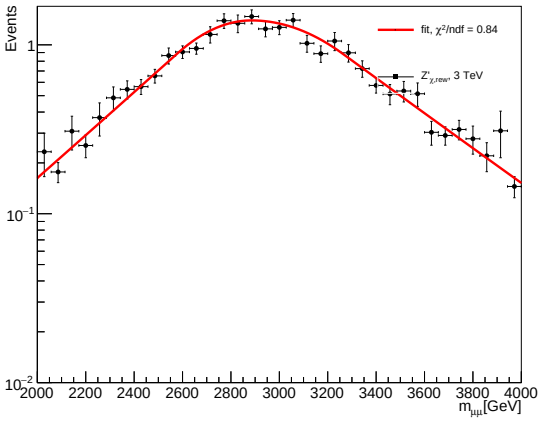
(a) Residual distribution of simulated Z'_χ invariant mass fit.



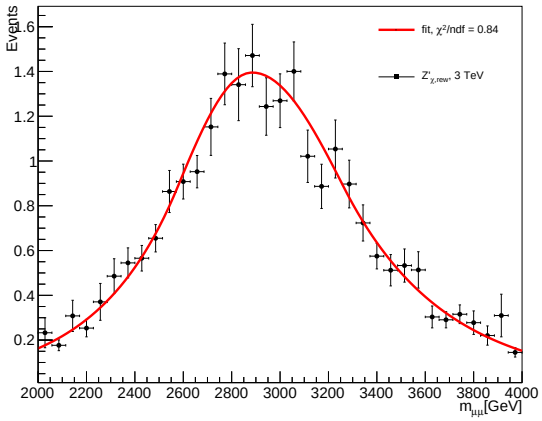
(b) Pull distribution of simulated Z'_χ invariant mass fit.

Figure 7.16: Residual (a) and pull (b) distribution of simulated Z'_χ invariant mass fit.

The Crystal Ball fit of the invariant mass distribution of the signal reweighted Z'_χ is shown in Figure 7.17a and 7.17b in logarithmic and linear scale. The signal reweighted Z'_χ distribution has relative large errors and gives $\chi^2/ndf = 0.84$. Figure 7.18a and 7.18b show the residual and pull distribution of the fit. We see a little discrepancies in the Z'_χ peak in the residual distribution that is also visible in Figure 7.17b.

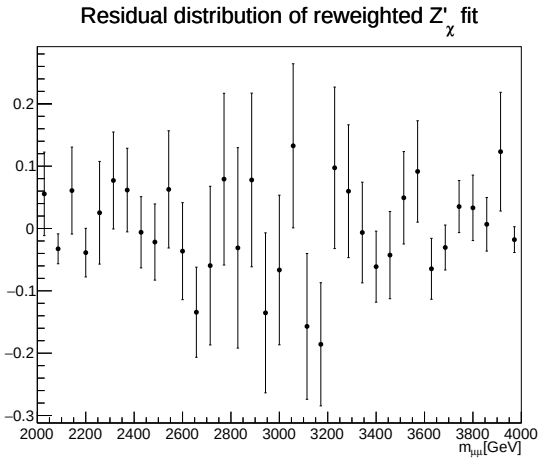


(a) Double-sided Crystal Ball fit of reweighted Z'_χ invariant mass distribution (log).

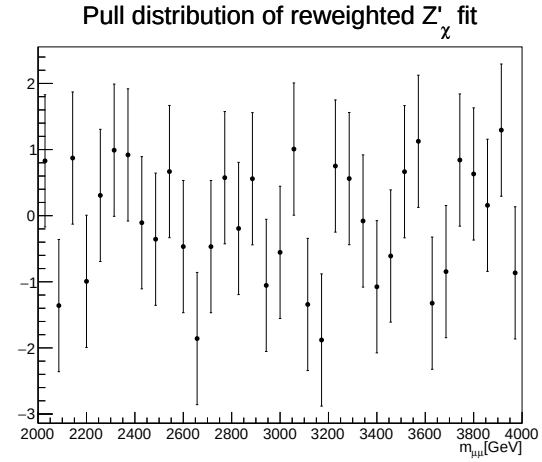


(b) Double-sided Crystal Ball fit of reweighted Z'_χ invariant mass distribution (lin).

Figure 7.17: Double-sided Crystal Ball fit of reweighted Z'_χ invariant mass distribution in logarithmic (a) and linear (b) scale.



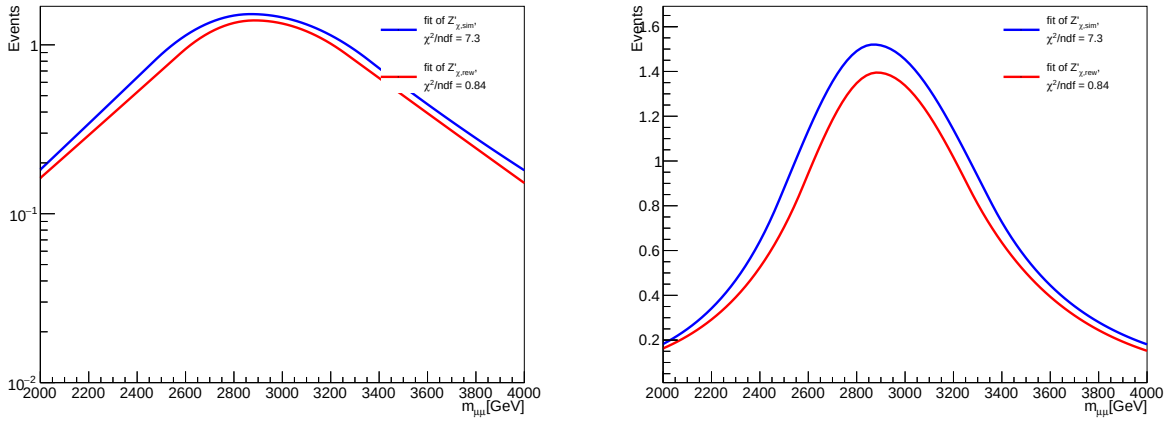
(a) Residual distribution of reweighted Z'_χ invariant mass fit.



(b) Pull distribution of reweighted Z'_χ invariant mass fit.

Figure 7.18: Residual (a) and pull (b) distribution of reweighted Z'_χ invariant mass fit.

We can now compare the Crystal Ball fit of the simulated and signal reweighted Z'_χ invariant mass distributions that is shown in Figure 7.19. The fit of the simulated Z'_χ distribution is consistently below the signal reweighted Z'_χ distribution which we also observed in Figure 7.12.



(a) Comparison between fit of simulated and reweighted Z'_χ invariant mass distribution (log)

(b) Comparison between fit of simulated and reweighted Z'_χ invariant mass distribution

Figure 7.19: Comparison between fit of simulated and reweighted Z'_χ invariant mass distribution

7.4 Comparison of different Z' models

With the LPXSignalReweightingTool we can choose between the E_6 and the SSM Z' model described in Section 2.2, and also choose the mass freely. We can therefore compare different Z' models at the same mass. Figure 7.20 shows the comparison of the Z'_{SSM}, Z'_χ and Z'_ψ model at 3 TeV in the electron channel. The width follows the ordering explained in Section 2.2.1 and 2.2.2, i.e $\Gamma_{Z'_{SSM}} > \Gamma_{Z'_\chi} > \Gamma_{Z'_\psi}$. The asymmetry between different models can also be studied in order to distinguish between different models.

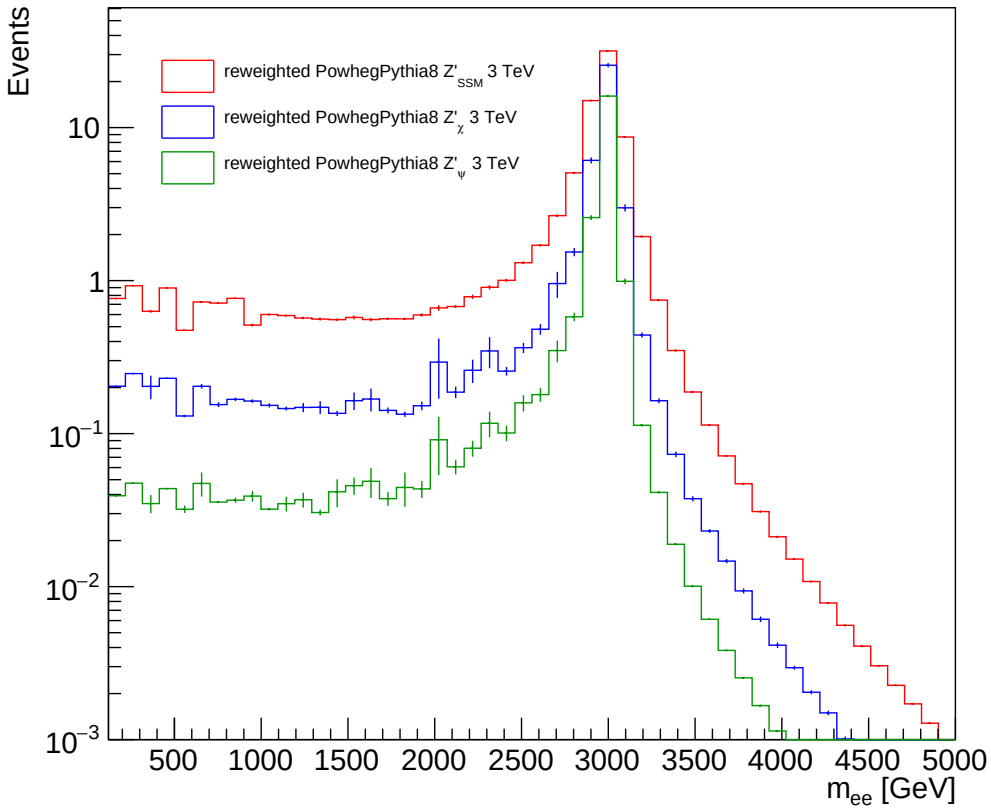


Figure 7.20: Comparison of the signal reweighted Z'_{SSM}, Z'_{χ} and Z'_{ψ} invariant mass distribution at 3 TeV in the electron chennel.

7.5 Z' models in Run 3 data and simulation comparison

The LPXSignalReweightingTool can also be used to apply signal reweighting to the fully simulated PowhegPythia8 Drell-Yan samples produced to match Run 3 data described in Section 4.3.2, to different Z' samples. Figure 7.21 shows a first result with signal reweighted Z' in Run 3 in the invariant mass distribution of the electron channel with the signal reweighted Z'_{SSM}, Z'_{χ} and Z'_{ψ} models at different masses, and are plotted on top of the SM background described in Section 4.3.3 and compared to data. The data is collected in 2022 at $\sqrt{s} = 13.6$ TeV where 26.1 fb^{-1} is used for analysis. In Figure 7.21 the data over 1 TeV is blinded. We see again that the width of the Z'_{SSM}, Z'_{χ} and Z'_{ψ} resonance at 1 TeV follows the ordering $\Gamma_{Z'_{SSM}} > \Gamma_{Z'_{\chi}} > \Gamma_{Z'_{\psi}}$ which is the same as in Figure 7.20. The Z'_{χ} model is plotted with Z' mass at 500 GeV, 750 GeV and 1000 GeV, and the width of the resonance increases as the mass of the Z'_{χ} boson increases.

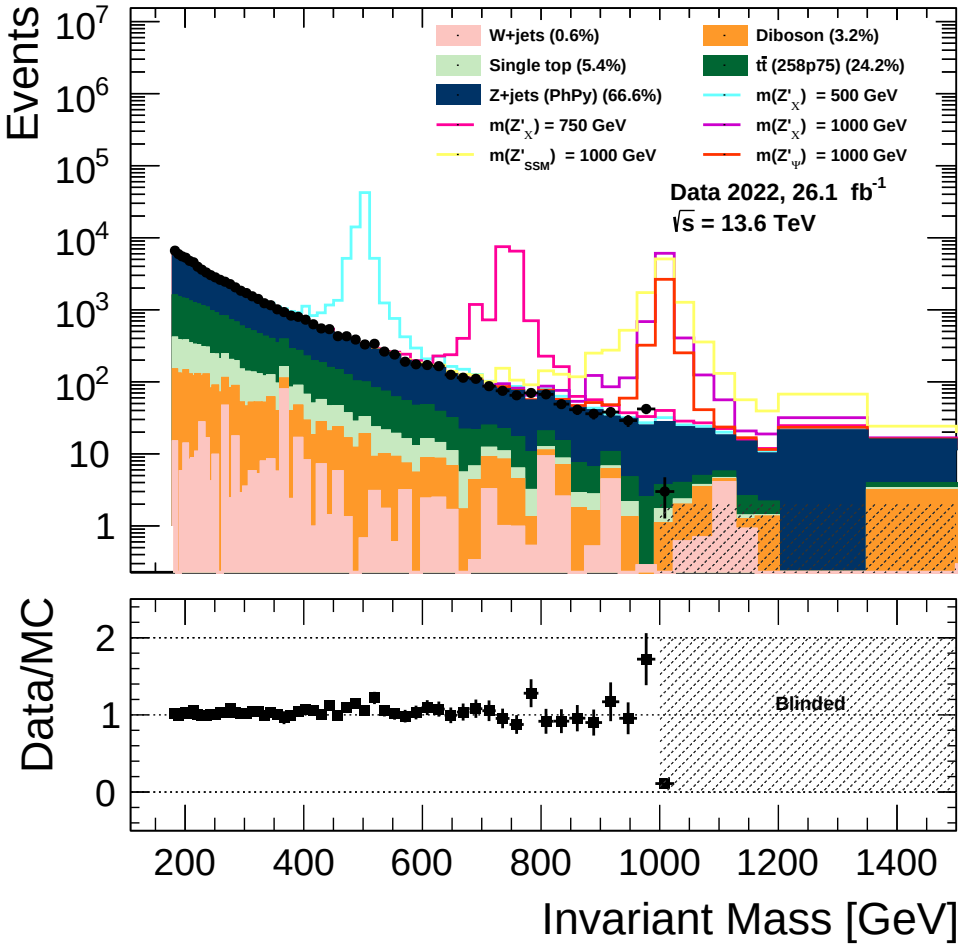


Figure 7.21: Invariant mass distribution in the electron channel in Run 3 with several signal reweighted Z'_{SSM}, Z'_{χ} and Z'_{ψ} models (solid lines) plotted on top of the SM background (shaded histograms) and compared to data.

Chapter 8

Conclusion and outlook

At the LHC the data collected by the ATLAS detector can be compared to MC simulations to test Standard Model predictions. MC simulations can further be used in searches for new physics beyond the Standard Model. We have in this thesis studied a method called signal reweighting to produce BSM Z' samples from a SM Drell-Yan sample. The LPXSignalReweightingTool has been extended to include the signal reweighing of a next-to-leading order PowhegPythia8 Drell-Yan sample that has not been done before. The results of the signal reweighted Z' samples shows some promising results when compared to fully simulated Z' Pythia8 samples produced to match with data taken in 2018 as part of Run 2 at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of 58.5 fb^{-1}. We also used the signal reweighting for Run 3 to produce some Z' samples that were superimposed on the comparison between the data and background simulations at $\sqrt{s} = 13.6$ TeV with 26.1 fb^{-1} of data. This extension to the LPXSignalReweightingTool will be used by the ATLAS collaboration for further Z' studies. An attempt to reweight a Sherpa Z + jets based sample to Z' models was also made using a method involving graphs developed as a attempt to understand the Sherpa Z + jets generator. This was unsuccessful due to insufficient information from the generator, so a statistical or machine learning approach may be needed to find an effective method to perform reweighting with Sherpa Z + jets. The signal reweighting method could be extended to include signal reweighting of new physics processes other than the Z' models discussed. One possibility is to apply signal reweighting to an already simulated Z' + dark matter sample to transform a specific model to another. We also applied the signal reweighting above to MC produced to match Run 3 data at $\sqrt{s} = 13.6$ TeV collected in 2022 corresponding to 26.1 fb^{-1}, to

produce Z' signals on top of data and SM simulated background.

Appendix A

Monte-Carlo samples

A.1 Monte-Carlo Drell-Yan samples used for signal reweighting

Process	Filter	N_{events}^{mc16e}	σ_{gen}[fb]
$Z \to e^-e^+$	120 GeV $< m_{ee} <$ 180 GeV	$2.48 \cdot 10^6$	$1.7478 \cdot 10^{-2}$
$Z \to e^-e^+$	180 GeV $< m_{ee} <$ 250 GeV	$1.50 \cdot 10^6$	$2.9212 \cdot 10^{-3}$
$Z \to e^-e^+$	250 GeV $< m_{ee} <$ 400 GeV	$2.00 \cdot 10^6$	$1.0820 \cdot 10^{-3}$
$Z \to e^-e^+$	400 GeV $< m_{ee} <$ 600 GeV	$1.49 \cdot 10^6$	$1.9550 \cdot 10^{-4}$
$Z \to e^-e^+$	600 GeV $< m_{ee} <$ 800 GeV	$8.50 \cdot 10^5$	$3.7401 \cdot 10^{-5}$
$Z \to e^-e^+$	800 GeV $< m_{ee} <$ 1000 GeV	$4.20 \cdot 10^5$	$1.0607 \cdot 10^{-5}$
$Z \to e^-e^+$	1000 GeV $< m_{ee} <$ 1250 GeV	$1.00 \cdot 10^5$	$4.2582 \cdot 10^{-6}$
$Z \to e^-e^+$	1250 GeV $< m_{ee} <$ 1500 GeV	$9.00 \cdot 10^4$	$1.4219 \cdot 10^{-6}$
$Z \to e^-e^+$	1500 GeV $< m_{ee} <$ 1750 GeV	$9.00 \cdot 10^4$	$5.4521 \cdot 10^{-7}$
$Z \to e^-e^+$	1750 GeV $< m_{ee} <$ 2000 GeV	$1.70 \cdot 10^5$	$2.2991 \cdot 10^{-7}$
$Z \to e^-e^+$	2000 GeV $< m_{ee} <$ 2250 GeV	$9.00 \cdot 10^4$	$1.0387 \cdot 10^{-7}$
$Z \to e^-e^+$	2250 GeV $< m_{ee} <$ 2500 GeV	$9.00 \cdot 10^4$	$4.9400 \cdot 10^{-8}$
$Z \to e^-e^+$	2500 GeV $< m_{ee} <$ 2750 GeV	$9.00 \cdot 10^4$	$2.4452 \cdot 10^{-8}$
$Z \to e^-e^+$	2750 GeV $< m_{ee} <$ 3000 GeV	$9.00 \cdot 10^4$	$1.2487 \cdot 10^{-8}$
$Z \to e^-e^+$	3000 GeV $< m_{ee} <$ 3500 GeV	$9.00 \cdot 10^4$	$1.0029 \cdot 10^{-8}$
$Z \to e^-e^+$	3500 GeV $< m_{ee} <$ 4000 GeV	$9.00 \cdot 10^4$	$2.9342 \cdot 10^{-9}$
$Z \to e^-e^+$	4000 GeV $< m_{ee} <$ 4500 GeV	$9.00 \cdot 10^4$	$8.9764 \cdot 10^{-10}$
$Z \to e^-e^+$	4500 GeV $< m_{ee} <$ 5000 GeV	$9.00 \cdot 10^4$	$2.8071 \cdot 10^{-10}$
$Z \to e^-e^+$	5000 GeV $< m_{ee} <$	$9.00 \cdot 10^4$	$1.2649 \cdot 10^{-10}$

Table A.1: PowhegPythia8 Drell-Yan samples used for signal reweighting in the electron

Process	Filter	N_{events}^{mc16e}	σ_{gen}[fb]
$Z \to \mu^-\mu^+$	120 GeV $< m_{\mu\mu} <$ 180 GeV	$3.32 \cdot 10^6$	$1.7478 \cdot 10^{-2}$
$Z \to \mu^-\mu^+$	180 GeV $< m_{\mu\mu} <$ 250 GeV	$1.90 \cdot 10^6$	$2.9212 \cdot 10^{-3}$
$Z \to \mu^-\mu^+$	250 GeV $< m_{\mu\mu} <$ 400 GeV	$2.25 \cdot 10^6$	$1.0820 \cdot 10^{-3}$
$Z \to \mu^-\mu^+$	400 GeV $< m_{\mu\mu} <$ 600 GeV	$1.66 \cdot 10^6$	$1.9550 \cdot 10^{-4}$
$Z \to \mu^-\mu^+$	600 GeV $< m_{\mu\mu} <$ 800 GeV	$1.00 \cdot 10^6$	$3.7399 \cdot 10^{-5}$
$Z \to \mu^-\mu^+$	800 GeV $< m_{\mu\mu} <$ 1000 GeV	$4.95 \cdot 10^5$	$1.0607 \cdot 10^{-5}$
$Z \to \mu^-\mu^+$	1000 GeV $< m_{\mu\mu} <$ 1250 GeV	$2.50 \cdot 10^5$	$4.2582 \cdot 10^{-6}$
$Z \to \mu^-\mu^+$	1250 GeV $< m_{\mu\mu} <$ 1500 GeV	$1.70 \cdot 10^5$	$1.4219 \cdot 10^{-6}$
$Z \to \mu^-\mu^+$	1500 GeV $< m_{\mu\mu} <$ 1750 GeV	$1.70 \cdot 10^5$	$5.4521 \cdot 10^{-7}$
$Z \to \mu^-\mu^+$	1750 GeV $< m_{\mu\mu} <$ 2000 GeV	$1.70 \cdot 10^5$	$2.2991 \cdot 10^{-7}$
$Z \to \mu^-\mu^+$	2000 GeV $< m_{\mu\mu} <$ 2250 GeV	$1.70 \cdot 10^5$	$1.0387 \cdot 10^{-7}$
$Z \to \mu^-\mu^+$	2250 GeV $< m_{\mu\mu} <$ 2500 GeV	$1.10 \cdot 10^5$	$4.9400 \cdot 10^{-8}$
$Z \to \mu^-\mu^+$	2500 GeV $< m_{\mu\mu} <$ 2750 GeV	$1.70 \cdot 10^5$	$2.4452 \cdot 10^{-8}$
$Z \to \mu^-\mu^+$	2750 GeV $< m_{\mu\mu} <$ 3000 GeV	$1.70 \cdot 10^5$	$1.2487 \cdot 10^{-8}$
$Z \to \mu^-\mu^+$	3000 GeV $< m_{\mu\mu} <$ 3500 GeV	$1.70 \cdot 10^5$	$1.0029 \cdot 10^{-8}$
$Z \to \mu^-\mu^+$	3500 GeV $< m_{\mu\mu} <$ 4000 GeV	$1.70 \cdot 10^5$	$2.9342 \cdot 10^{-9}$
$Z \to \mu^-\mu^+$	4000 GeV $< m_{\mu\mu} <$ 4500 GeV	$1.70 \cdot 10^5$	$8.9764 \cdot 10^{-10}$
$Z \to \mu^-\mu^+$	4500 GeV $< m_{\mu\mu} <$ 5000 GeV	$1.70 \cdot 10^5$	$2.8071 \cdot 10^{-10}$
$Z \to \mu^-\mu^+$	5000 GeV $< m_{\mu\mu} <$	$1.70 \cdot 10^5$	$1.2649 \cdot 10^{-10}$

Table A.2: PowhegPythia8 Drell-Yan samples used for signal reweighting in the muon channel for the $mc16e$ period in Run 2

A.2 Monte-Carlo Z' signal samples

Process	Model	$m_{Z'}$[GeV]	DSID	N_{events}^{mc16e}	σ_{gen}[fb]
$Z' \to e^-e^+$	Z'_χ	3000	301216	$4.00 \cdot 10^4$	$8.102 \cdot 10^{-7}$
$Z' \to \mu^-\mu^+$	Z'_χ	3000	301221	$1.66 \cdot 10^6$	$8.045 \cdot 10^{-7}$

Table A.3: Pythia8 Z' sample for the Z'_χ model in the electron and muon channel for the $mc16e$ period in Run 2.

Bibliography

[1] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

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

[3] A. Salam. Weak and Electromagnetic Interactions. *Conf. Proc. C*, 680519:367–377, 1968. 10.1142/9789812795915_0034.

[4] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev. D*, 2:1285–1292, 1970.

[5] P W Higgs. Spontaneous symmetry breakdown without massless bosons. *Physical Review (U.S.) Superseded in part by Phys. Rev. A, Phys. Rev. B: Solid State, Phys. Rev. C, and Phys. Rev. D.*

[6] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, 1964.

[7] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory.* Westview Press, 1995. Reading, USA: Addison-Wesley (1995) 842 p.

[8] Universität zürich. standard model. https://www.physik.uzh.ch/groups/serra/StandardModel.html. Accessed: 2023-05-27.

[9] ATLAS collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1–29, 2012.

[10] CMS collaboration. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B*, 716:30–61, 2012.

[11] D. London and J. L. Rosner. Extra Gauge Bosons in E(6). *Phys. Rev. D*, 34:1530, 1986.

[12] P. Langacker. The Physics of Heavy Z' Gauge Bosons. *Rev. Mod. Phys.*, 81:1199–1228, 2009.

[13] G. Senjanovic and Rabindra N. Mohapatra. Exact Left-Right Symmetry and Spontaneous Violation of Parity. *Phys. Rev. D*, 12:1502, 1975.

[14] ATLAS collaboration. Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector. *Phys. Lett. B*, 796:68–87, 2019.

[15] ATLAS collaboration. Search for new high-mass phenomena in the dilepton final state using 36 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 10:182, 2017.

[16] ATLAS collaboration. Search for heavy resonances decaying to a Z boson and a photon in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Lett. B*, 764:11–30, 2017.

[17] ATLAS collaboration. Search for high mass dilepton resonances in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS experiment. *Phys. Lett. B*, 700:163–180, 2011.

[18] V. Morisbak. *Searching for new forces of Nature: New neutral gauge bosons in dimuon final states with the ATLAS detector at the LHC.* PhD thesis, Oslo U., Oslo U., 2021.

[19] M. V. Chizhov, V. A. Bednyakov, and J. A. Budagov. On resonance search in dilepton events at the LHC. *Phys. Part. Nucl. Lett.*, 10:144–146, 2013.

[20] E. Salvioni, A. Strumia, G. Villadoro, and F. Zwirner. Non-universal minimal Z' models: present bounds and early LHC reach. *JHEP*, 03:010, 2010.

[21] M. V. Chizhov, V. A. Bednyakov, and J. A. Budagov. Proposal for chiral bosons search at LHC via their unique new signature. *Phys. Atom. Nucl.*, 71:2096–2100, 2008.

[22] M. Dittmar, A. Nicollerat, and A. Djouadi. Z-prime studies at the LHC: An Update. *Phys. Lett. B*, 583:111–120, 2004.

[23] E. Accomando et al. Z' physics with early LHC data. *Phys. Rev. D*, 83:075012, 2011.

[24] C. Willis. The LPXSignalReweightingTool: A Package for Signal Template Reweighting in High-Energy Physics Analysis. Technical report, CERN, Geneva, 2016.

[25] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J. C*, 63:189–285, 2009.

[26] A. Einstein. Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 322(10):891–921, 1905. ...

[27] M. Schott and M. Dunford. Review of single vector boson production in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J. C*, 74:2916, 2014.

[28] C. Bierlich et al. A comprehensive guide to the physics and usage of PYTHIA 8.3. 3 2022. https://arxiv.org/abs/2203.11601.

[29] Torbjörn Sjöstrand. The PYTHIA Event Generator: Past, Present and Future. *Comput. Phys. Commun.*, 246:106910, 2020.

[30] S. Frixione, F. Stoeckli, P. Torrielli, B. R. Webber, and C. D. White. The MCaNLO 4.0 Event Generator. 10 2010.

[31] S. Alioli, P. Nason, C. Oleari, and E. Re. A general framework for implementing NLO calculations in shower Monte Carlo programs the POWHEG BOX. *JHEP*, 06:043, 2010.

[32] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[33] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[34] E. Bothmann et al. Event Generation with Sherpa 2.2. *SciPost Phys.*, 7(3):034, 2019.

[35] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert. QCD matrix elements + parton showers: The NLO case. *JHEP*, 04:027, 2013.

[36] K. Danziger, S. Höche, and F. Siegert. Reducing negative weights in Monte Carlo event generation with Sherpa. *MCNET-21-30, FERMILAB-PUB-21-536-T*, 10 2021.

[37] A. Buckley et al. The HepMC3 event record library for Monte Carlo event generators. *Comput. Phys. Commun.*, 260:107310, 2021.

[38] HepMC 2 user manual, May 17 2010. http://hepmc.web.cern.ch/hepmc/releases/HepMC2_user_manual.pdf. Accessed: 2023-05-13.

[39] M. Dobbs and J. B. Hansen. The HepMC C++ Monte Carlo event record for High Energy Physics. *Comput. Phys. Commun.*, 134:41–46, 2001.

[40] R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022. 10.1093/ptep/ptac097.

[41] Particle properties in the pythia generator. https://pythia.org/latest-manual/ParticleProperties.html. Accessed: 2023-05-19.

[42] The HepMC flow class. https://rivet.hepforge.org/code/hepmc.bak/Flow_8h-source.html. Accessed: 2023-05-19.

[43] ATLAS collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[44] S. Agostinelli et al. GEANT4–a simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250–303, 2003.

[45] ATLAS collaboration. The ATLAS Simulation Infrastructure. *Eur. Phys. J. C*, 70:823–874, 2010.

[46] Wolfgang Lukas. Fast simulation for atlas: Atlfast-ii and isf. *Journal of Physics: Conference Series*, 396(2):022031, dec 2012.

[47] C. Ciobanu et al. Z' generation with PYTHIA. *FERMILAB-FN-0773-E*, 7 2005.

[48] S. Alioli, P. Nason, C. Oleari, and E. Re. NLO vector-boson production matched with shower in POWHEG. *JHEP*, 07:060, 2008.

[49] The LPXKfactorTool. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/LPXKfactorTool. Accessed: 2023-05-19.

[50] A. Pukhov, E. Boos, M. Dubinin, V. Edneral, V. Ilyin, D. Kovalenko, A. Kryukov, V. Savrin, S. Shichanin, and A. Semenov. CompHEP: A Package for evaluation of Feynman diagrams and integration over multiparticle phase space. 8 1999.

[51] D. Reinhard. *Graph Theory*, volume Graduate Texts in Mathematics (GTM, volume 173). Springer Nature, 01 2017.

[52] J. S. Bagga L. W. Beineke. *Line Graphs and Line Digraphs*, volume Developments in Mathematics, 68. Springer Nature, London, 2021. ISBN 978-3-030-81384-0.

[53] Frank J. Massey Jr. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.

[54] M. Oreglia. A Study of the Reactions $\psi' \to \gamma\gamma\psi$. 12 1980. https://www.osti.gov/biblio/5670739.