Department of Physics and Astronomy

University of Heidelberg

Master thesis

in Physics

submitted by

Josua Göcking

born in Tübingen (Germany)

2018

Using track observables to improve the sensitivity to weak boson fusion processes

This Master thesis has been carried out by Josua Göcking

at the

Institute for Theoretical Physics

under the supervision of

Prof. Dr. Tilman Plehn

Abstract

Higgs to invisible decays are an important signature in the search for physics beyond the Standard Model. Due to its tagging properties the weak boson fusion (WBF) Higgs production is a promising channel for such searches. Hence, it is of great importance to improve the separation of signal and background of this process. In this study we first start by comparing the Monte Carlo event generators SHERPA, MADGRAPH+PYTHIA and HERWIG to each other with the aim of finding the one best suited to describe track propterties in the respective signal and background processes. Afterwards, we apply these findings to examine if analyzing tracks instead of jets can improve the sensitivity. We find that SHERPA is best suited describing the track activity. Furthermore, using tracks instead of jets for the analysis of events increases the sensitivity. Moreover, the jet properties are complementary to track properties. Finally, we show that using tracks to discriminate Higgs WBF from the Z WBF is possible and preferable for jets with minimum jet transverse momenta of 30 GeV and higher.

Zusammenfassung

Higgs Zerfälle in unsichtbare Teilchen sind ein wichtiges Signal auf der Suche nach Physik jenseits des Standardmodells. Aufgrund seiner Tagging-Eigenschaften ist die Analyse der Higgs Produktion durch die Fusion schwacher Eichbosonen (WBF) vielversprechend. Daher ist es von großer Bedeutung die Trennung von Signal und Untergrund in diesem Prozess zu optimieren. Wir beginnen diese Analyse damit die Monte Carlo Event Generatoren SHERPA, MADGRAPH+PYTHIA und HERWIG miteinander zu vergleichen. Ziel ist es den Generator zu finden, der für die Beschreibung von geladenen Teilchen (“Tracks”) in den jeweiligen Signal und Untergrundprozessen am besten geeignet ist. Danach benutzen wir diese Ergebnisse um zu untersuchen, ob die Studie von Tracks anstatt von Jets die Sensitivität erhöhen kann. Es stellt sich heraus, dass SHERPA am besten geeignet ist die Aktivität von Tracks zu beschreiben. Außerdem kann wenn Tracks an Stelle von Jets für die Analyse verwendet werden eine höhere Sensitivität erreicht werden. Darüber hinaus, zeigt sich, dass Jets und Tracks einander ergänzen. Abschließend zeigen wir, dass Tracks auch zur Trennung des Higgs WBF Prozesses vom Z WBF Prozess verwendet werden können. Für Jets mit minimalem transversalem Impuls über 30 GeV ist eine Track Analyse der von Jets vorzuziehen.

Contents

1 Introduction

The Standard Model is one of the greatest achievements in particle physics of the last decades. Its description of the strong and electroweak interactions predicts particles that could be successfully observed at colliders.

However, there are still some missing pieces for a full description of our universe. Firstly, the Standard Model does not describe the gravitational force at all. A complete theory, however, would need to include these interactions.

Secondly, by now it is experimentally established that neutrinos do oscillate in flavor, which means that in contrast to the Standard Model prediction, they are not massless. Furthermore, cosmological findings suggest that there is matter in our universe, which is not described by the Standard Model and seems to interact only gravitationally and possibly weakly. This matter is referred to as dark matter.

These are only some of the open questions in particle physics that show that the Standard Model is not the final answer.

Therefore, after the Higgs boson was found at the Large Hadron Collider (LHC), and with it the last missing piece in the Standard Model, experimentalists are now searching for these so-called “new physics”.

In these endeavors the Higgs boson, in particular, becomes interesting, since it could serve as a portal to new particles. Such new particles would be invisible to current detectors at colliders and could only be detected by accounting for the missing energy in the event as for neutrinos. Such a signal could show up, for example, in an enhancement of the Higgs to invisible branching ratio.

Since the weak boson fusion channel for Higgs production has some characteristic features that allow for a neat separation of signal and background, this channel is favored above the dominating gluon fusion production process in searches for Higgs invisible decays at the LHC.

Hence, it is worthwhile to find ways to improve the separation of weak boson fusion processes from the background. One way to do this might be using the track information, i.e. the tracks charged particles leave in the detector, instead of the jet information. One reason for this is that such tracks would be sensitive much to lower energies than jets are. Whether this is really helpful can be simulated using Monte Carlo event generators. However, to make accurate predictions, the event generator must describe the physics well.

The aim of this study is to first find the best Monte Carlo event generator to describe the track activity in weak boson fusion and second analyzing if using tracks instead of jets can in fact improve the separation of signal and background.

We begin by introducing the theoretical background needed for our study. This can be found in section 2. In section 3 the physics of colliders and some technicalities of our study are described. This is followed by discussing Higgs invisible decays and weak boson fusion in section 4. A comparison of the event generators SHERPA, MADGRAPH combined with PYTHIA and HERWIG is presented in section 5. Finally, in section 6 we study if the discrimination of signal and background can be improved by using tracks instead of jets. Furthermore, a summary of what we have done and how one could proceed from these results can be found in section 7.

2 The Standard Model

The Standard Model (SM) [1–3] reflects our current understanding of nature. It describes three of the four fundamental forces, the electromagnetic force and the strong and weak nuclear forces. Gravity has not yet been included in the SM. The SM is established using the concepts provided by Quantum Field Theory (QFT). In the following we will briefly introduce the basic concepts of this theory. For a deeper insight we refer to the literature [9, 10].

Matter and their interactions are described in the SM by two kinds of particles, fermions and bosons. Fermions carry half integer spin (for all known elementary particles we have spin-$\frac{1}{2}$) and make up the matter part of our universe. Bosons carry integer spin. Spin-1 bosons like the photon and the gluons are force mediators, while the spin-0 Higgs boson is responsible for the masses of the weak gauge bosons and the fermions.

We begin by studying the fermions. One distinguishes two kinds of fermions: Quarks, which take part in the strong interaction and leptons, which do not. The lepton sector consists of electron-like particles and respective neutral partner, the neutrino. Similarly the quark sector is made up of up-type quarks with electrical charge $\frac{2}{3}$ and down-type quarks with electric charge $-\frac{1}{3}$. There are three generations of such pairs. The full particle content of the SM can be seen in Fig. 1.

In QFT fermions are described using Dirac spinor fields ψ. With the mass m_ψ of the fermion field the Lagrangian becomes

$$\mathcal{L} = i\overline{\psi}\gamma^\mu\partial_\mu\psi - m_\psi\overline{\psi}\psi \ , \tag{2.1}$$

We are now considering fields that are invariant under a continuous $SU(N)$ symmetry,

$$\psi(x) \to U\psi \ , \tag{2.2}$$

where $U \in SU(N)$, i.e. U is a $N \times N$ special unitary matrix and can be expressed by,

$$U(\theta) = \exp(i\theta_a T^a) \ , \tag{2.3}$$

with the real coefficients θ_a and the hermitian and traceless generator matrices T^a. Index a runs from 1 to $N^2 - 1$, i.e. the $SU(N)$ is described by $N^2 - 1$ generators.

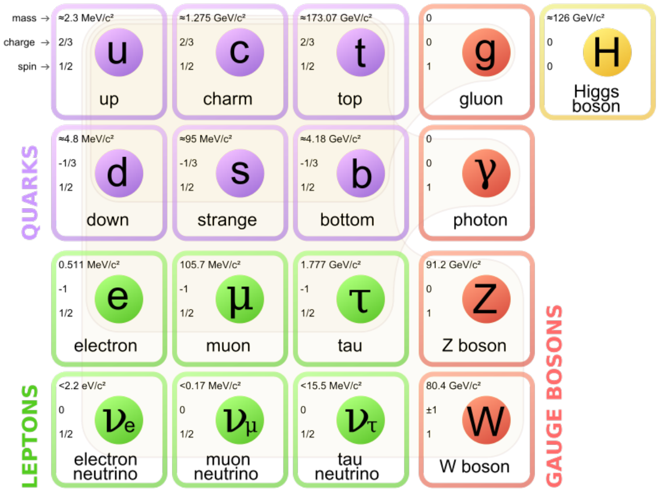

Figure 1: The elementary particles of the Standard Model. Taken from [8].

Furthermore, the generators fulfill the commutation relation

$$[T^a, T^b] = if^{abc}T_c \,, \tag{2.4}$$

where the f^{abc} are the structure constants. For an abelian theory it is $f^{abc} = 0$.

In general the introduced symmetry transformation can be local, i.e. depending on spacetime. This is obtained by setting $\theta_a = \theta_a(x)$.

It can be easily seen that such a local symmetry transformation would not be symmetric under the Lagrangian in Eq. (2.1).

Therefore, we define the $SU(N)$ gauge fields with the gauge transformation property

$$A_\mu(x) = U(x)A_\mu(x)U^\dagger(x) + \frac{i}{g}U(x)\partial_\mu U^\dagger(x) \,. \tag{2.5}$$

This allows us to construct the covariant derivative

$$D_\mu = \partial_\mu - igA_{\mu a}T^a \,. \tag{2.6}$$

with $A_\mu = A_{\mu a}T^a$. The covariant derivative transforms as $D^\mu \to UD^\mu U^\dagger$ and therefore, restores the gauge invariance of the Lagrangian.

To complete the Lagrangian we need to include a kinetic term for the gauge fields, which is found by introducing the field strength tensor

$$F_{\mu\nu} = \frac{i}{g}[D_\mu, D_\nu] \ . \tag{2.7}$$

Apparently it has the same gauge transformation behavior as the covariant derivative.

However, since this term is not gauge invariant by itself, the kinetic term is obtained by taking the trace of the square of this tensor. Multiplying this by a factor of $-1/2$ the final Lagrangian becomes

$$\mathcal{L} = \overline{\psi}(i\gamma^\mu D_\mu - m_\psi)\psi - \frac{1}{2}\,\mathrm{Tr}(F_{\mu\nu}F^{\mu\nu}) \ . \tag{2.8}$$

The full symmetry group of the SM is $SU(3)_C \times SU(2)_L \times U(1)_Y$. In the following we will briefly study these symmetry groups and the resulting interactions.

2.1 Quantum Chromodynamics

The interactions based on the $SU(3)_C$ group are the strong interactions of quantum chromodynamics (QCD). Leptons are singlets under this group, i.e. are not taking part in the strong force. Quarks, on the other hand, come in three colors, which are the values of the $SU(3)$ index. That is, quarks are $SU(3)$ triplets

$$q = \begin{pmatrix} q_R \\ q_G \\ q_B \end{pmatrix} \ , \tag{2.9}$$

where $R,\ G,\ B$ denote the color charges red, green and blue, respectively.

There are eight $SU(3)_C$ generators t^a. The covariant derivative is given by

$$D_\mu = \partial_\mu - ig_s G_{\mu a} t^a \ , \tag{2.10}$$

with the strong coupling g_s and the eight gluon fields G^a_μ

The QCD Lagrangian then becomes

$$\mathcal{L}_{\mathrm{QCD}} = \overline{q}\left(i\gamma^\mu\partial_\mu - m_u\right)q + g_s G_{\mu a} t^a \overline{q}\gamma^\mu q - \frac{1}{2}\,\mathrm{Tr}\left(F_{\mu\nu}F^{\mu\nu}\right) \ , \tag{2.11}$$

where q describes the quark fields. Note that gauge fixing terms and Fadeev-Popov ghost fields where neglected in this expression.

2.2 Electroweak Interactions

The weak interactions are peculiar in acting only on left handed fermions or right handed antifermions, as experiments have shown [15]. From a Dirac field ψ we obtain the left handed part L and the right handed part R using the projection operators:

$$L = P_L\psi = \frac{1}{2}(1-\gamma_5)\psi\ , \qquad R = P_R\psi = \frac{1}{2}(1+\gamma_5)\psi\ , \tag{2.12}$$

with $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$.

The electroweak interactions are based on the $SU(2)_L \times U(1)_Y$ group. Left-handed fields are doublets and right-handed fields are singlets under the $SU(2)_L$.

Therefore, the covariant derivatives for the left- and right-handed fermions are given by

$$\begin{aligned} D_{L\mu} &= \partial_\mu - ig_2 W_{\mu a}\tau^a - ig_1 Y B_\mu \\ D_{R\mu} &= \partial_\mu - ig_1 B_\mu\ , \end{aligned} \tag{2.13}$$

with $\tau_a = \frac{1}{2}\sigma_a$, σ_a being the Pauli matrices and the couplings g_1 and g_2 of the gauge field B_μ corresponding to the $U(1)_Y$ symmetry and the gauge fields W^a_μ corresponding to the $SU(2)_L$ symmetry, respectively.

For left- and right-handed fermions denoted as L and R, respectively, the Lagrangian becomes

$$\mathcal{L}_{\text{fermions}} = \sum_L \overline{L} i\gamma^\mu D^L_\mu L + \sum_R \overline{R} i\gamma^\mu D^R_\mu R\ , \tag{2.14}$$

Together with the Lagrangian of the electroweak gauge fields given by

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}B^{\mu\nu}B_{\mu\nu} - \frac{1}{2}\operatorname{Tr}\left(W^{\mu\nu}_a W^a_{\mu\nu}\right)\ , \tag{2.15}$$

we obtain almost the complete electroweak Lagrangian. What's still missing are the mass terms for the fermion fields. However, since we split the fields into left- and righ-handed ones, we cannot simply add a mass term for each field to the Lagrangian, since such a term would not be invariant under the $SU(2)_L$ group. To solve this problem we first need to introduce the mechanism of electroweak symmetry breaking.

2.3 Electroweak Symmetry Breaking

In the SM interactions are included by imposing local gauge invariance. As we have seen this introduces gauge fields to the Lagrangian. One can easily show that introducing a mass term to a gauge boson in the Lagrangian, i.e. the term $m^2 A_a^\mu A_\mu^a$ breaks local gauge invariance.

Therefore, gauge bosons are supposed to be massless. However, from experiment we know that the gauge bosons of the weak interaction are massive.

To save local gauge invariance Peter Higgs [4, 5], Francois Englert and Robert Brout [6] came up with the mechanism of electroweak symmetry breaking (EWSB). In the following, we will describe this mechanism.

We begin by introducing the scalar Higgs field ϕ with the covariant derivative

$$(D_\mu \varphi)_i = \partial_\mu \varphi_i - i \left[g_2 W_\mu^a \tau^a + g_1 B_\mu Y \right]_i^j \varphi_j \ , \tag{2.16}$$

where the φ_i are the components of ϕ and $Y = -\frac{1}{2} I$ is the hypercharge of the scalar field.

Writing the term in the brackets in matrix form yields

$$g_2 W_\mu^a \tau^a + g_1 B_\mu Y = \frac{1}{2} \begin{pmatrix} g_2 W_\mu^3 - g_1 B_\mu & g_2(W_\mu^1 - i W_\mu^2) \\ g_2(W_\mu^1 + i W_\mu^2) & -g_2 W_\mu^3 - g_1 B_\mu \end{pmatrix} \ . \tag{2.17}$$

For the potential of ϕ we then take

$$V(\varphi) = \frac{1}{4} \lambda (\varphi^\dagger \varphi - \frac{1}{2} v^2)^2 \ , \tag{2.18}$$

which will give us a nonzero vacuum expectation value (vev) as is shown in Fig. 2. Applying a gauge transformation that brings this vev to the first component and makes it real (unitary gauge) will give us

$$\langle 0 | \varphi(x) | 0 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} v \\ 0 \end{pmatrix} \ . \tag{2.19}$$

If we now compute the kinetic term $-(D^\mu \varphi)^\dagger D_\mu \varphi$ and replace ϕ by its vev we will get

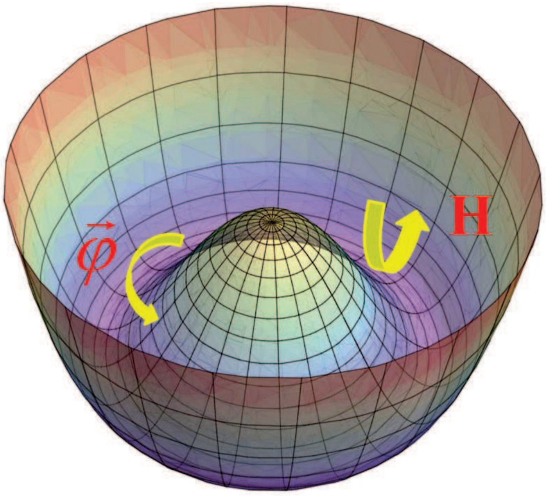

Figure 2: Scalar SM potential, due to its shape commonly referred to as the Mexican hat potential. The Goldstone ($\vec{\varphi}$) and Higgs (H) field describe excitations in the direction indicated by their respective arrow. The Goldstone field corresponds to a phase shift, which using the unitary gauge is set to unity. Taken from [14].

a mass term for the gauge fields,

$$\begin{aligned}\mathcal{L}_{\text{mass}} &= -\frac{1}{8}v^2(1,\ 0)\begin{pmatrix} g_2W^3_\mu - g_1B_\mu & g_2(W^1_\mu - iW^2_\mu) \\ g_2(W^1_\mu + iW^2_\mu) & -g_2W^3_\mu - g_1B_\mu \end{pmatrix}^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= -\frac{1}{8}v^2\left[(g_2W^3_\mu - g_1B^2_\mu)^2 + g_2^2(W^1_\mu + iW^2_\mu)(W^1_\mu - iW^2_\mu)\right] \ .\end{aligned} \tag{2.20}$$

Introducing the Weinberg angle $\theta_W = \tan^{-1}(g_1/g_2)$ we define the fields

$$\begin{aligned} W^\pm_\mu &\equiv \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu) \ , \\ \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} &\equiv \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix} \ . \end{aligned} \tag{2.21}$$

Substituting these into Eq. (2.20) this will become

$$\begin{aligned}\mathcal{L}_{\text{mass}} &= -\left(\frac{g_2v}{2}\right)^2 W^{+\mu}W^-_\mu - \frac{1}{2}\left(\frac{g_2v}{2\cos\theta_W}\right)^2 Z^\mu Z_\mu \\ &= -m_W^2W^{+\mu}W^-_\mu - \frac{1}{2}m_Z^2Z^\mu Z_\mu \ ,\end{aligned} \tag{2.22}$$

where we have introduced the mass terms

$$m_W = \frac{g_2 v}{2} = \cos\theta_W m_Z \ . \tag{2.23}$$

From experiment we know these masses to be $m_W = 80.4$ GeV and $m_Z = 91.2$ GeV [11], from which we get $\sin^2\theta_W = 0.223$.

Apparently, by this symmetry breaking the A_μ field remains massless. Therefore, the $SU(2)_L \times U(1)_Y$ symmetry is broken to a $U(1)$ symmetry. This unbroken subgroup is the gauge group of electromagnetism.

Since our complex scalar field ϕ consists of two components we will altogether have four real scalar fields. Each weak gauge boson will "eat up" one of these scalar fields, which then becomes its longitudinal component, leaving us with one remaining scalar field, responsible for the shifts in the overall scale of ϕ.

In the unitary gauge, we then have

$$\varphi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} v + H(x) \\ 0 \end{pmatrix} \ , \tag{2.24}$$

with the real scalar field H, whose corresponding particle is the Higgs boson. The potential of this field is given by

$$V(\varphi) = \frac{1}{4}\lambda v^2 H^2 + \frac{1}{4}\lambda v H^3 + \frac{1}{16}\lambda H^4 \ . \tag{2.25}$$

From this we can read off the Higgs mass term $m_H^2 = \frac{1}{2}\lambda v^2$. On July 4, 2012 the ATLAS and CMS collaborations at CERN announced the discovery of this particle with a mass $m_H^2 = 125$ GeV [12, 13].

Now, we want to see how the couplings g_1 and g_2 are related to the electrical charge e. For this we consider the covariant derivative of the field W_μ^3

$$D_\mu = \partial_\mu - i g_2 W_\mu^3 = \partial_\mu - i g_2 (\sin\theta_W A_\mu + \cos\theta_W Z_\mu) \ . \tag{2.26}$$

Since the coupling of A_μ has to be equal to e we can conclude that

$$e = g_2 \sin\theta_W = g_1 \cos\theta_W \ , \tag{2.27}$$

where in the second part we used that $\sin\theta_W = g_1 \Big/ \sqrt{g_1^2 + g_2^2}$ and $\cos\theta_W = g_2 \Big/ \sqrt{g_1^2 + g_2^2}$.

Finally, we note that the masses for the fermions are also obtained by the Higgs mechanism. This is shown in Appendix A. Here we only state the results for electron-like leptons e with Yukawa coupling y, up-type quarks u with Yukawa coupling y_1 and down-type quarks d with Yukawa coupling y_2:

$$m_e = \frac{yv}{\sqrt{2}}\,, \qquad m_d = \frac{y_1 v}{\sqrt{2}}\,, \qquad m_u = \frac{y_2 v}{\sqrt{2}}\,. \tag{2.28}$$

3 Collider phenomenology and Technicalities

Many tests of the SM and theories beyond it (BSM) are done by using colliders. To relate the experimental results to the theoretical prediction we have to keep in mind, that the outcoming particles are rarely what we are interested in, since the particles we want to observe typically decay before reaching the detector. Furthermore, due to color confinement, quarks are never observed freely and can only be resolved using jet algorithms. Methods to describe these behaviors are what is usually called phenomenology.

There are certain colliders that one can build. The main differences are in shape (linear or circular), the colliding particles (leptons or hadrons) and the beam parameters like the center-of-mass energy and the luminosity. The advantage of colliders using leptons as colliding particles is, that the initial momentum of the particles is exactly known, while for hadrons one only knows the momentum of the full hadron instead of the interacting individual quarks. However, since quarks are color charged some processes are easier obtained using hadron colliders. Also high energy losses in lepton colliders due to synchrotron radiation favor the use of hadron colliders.

The Large Hadron Collider (LHC) is a proton-proton collider that currently works at a center-of-mass energy of $\sqrt{s} = 13$ TeV.

To describe the kinematics in a hadron collider, we need variables that suit the geometry of the detector and are invariant under longitudinal Lorentz boosts. We will consider the z-axis to be the beam axis. Apparently the momentum transverse to the beam axis, defined as

$$p_T = \sqrt{p_x^2 + p_y^2} = p \sin\theta \tag{3.1}$$

fulfills these requirements. If one assumes that the transverse momentum of the beam particles vanishes, which is justified for high energies, momentum conservation tells us, that for the final state particles we have $\sum \vec{p_T} = 0$. This allows to resolve particles which are invisible to the detector, like neutrinos, by computing the missing transverse momentum. The azimuthal angle ϕ of the plane is apparently also invariant under longitudinal Lorentz boost. The polar angle θ however does not have this property. Therefore, we use the rapidity y, which is for a particle with momentum p^μ and energy E, defined as

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

In the limit of massless particles this becomes

$$\eta = \lim_{m \to 0} y = \frac{1}{2} \ln \left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} \right) = \operatorname{artanh} \left(\frac{p_z}{|\vec{p}|} \right) = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] . \tag{3.3}$$

This new quantity η is called the pseudorapidity. As one can see it only depends on the polar angle θ and is like the rapidity additive under longitudinal Lorentz boosts (as long as $m \sim 0$).

Together with the azimuthal angle ϕ, the pseudorapidity spans the $\eta - \phi$-plane. This allows the definition of a distance measure ΔR_{ij}, defined as

$$\Delta R_{ij} = \sqrt{\Delta\phi_{ij}^2 + \Delta\eta_{ij}^2} = \sqrt{(\phi_j - \phi_i)^2 + (\eta_j - \eta_i)^2} . \tag{3.4}$$

When colliding particles, the resulting events will consist of signal events, i.e. the processes we are interested in, and background events. To accurately test a theory we, therefore, have to distinguish signal from background. To do this we first have to know how many events N are expected for the respective process with cross section σ and collider luminosity $\mathcal{L}$, i.e. the number of particles crossing a unit area. This is given by

$$N = \sigma \mathcal{L} . \tag{3.5}$$

When analyzing the events we get the signal events $S = \sigma_{\text{sig}} \mathcal{L}$ and the background events $B = \sigma_{\text{bg}} \mathcal{L}$ with the signal and background cross sections σ_{sig} and σ_{bg}, respectively. Therefore, for N events have $N = S + B$. The number of events obeys Poisson statistics, which means for an expected number of events N the standard deviation will be $\sqrt{N} = \sqrt{S+B} \sim \sqrt{B}$, where in the last step we used that usually we have $S \ll B$. To get the significance of the signal we compute the number of signal events relative to one standard deviation: $\frac{S}{\sqrt{B}}$.

However, usually it is not possible to perfectly resolve signal and background. Therefore, we have to include the efficiency of signal and background in our analysis. Effectively we will measure $\varepsilon_{\text{sig}} S$ signal events with the signal efficiency ε_{sig} and $\varepsilon_{\text{bg}} B$ background events with the background efficiency ε_{bg}. Substituting S and B by these values will give us the actual significance.

Therefore, we get

$$\frac{S}{\sqrt{S+B}} = \frac{\varepsilon_{\text{sig}}\sigma_{\text{sig}}}{\sqrt{\varepsilon_{\text{sig}}\sigma_{\text{sig}} + \varepsilon_{\text{bg}}\sigma_{\text{bg}}}} \cdot \sqrt{\mathcal{L}} \overset{S \ll B}{\approx} \frac{\varepsilon_{\text{sig}}\sigma_{\text{sig}}}{\sqrt{\varepsilon_{\text{bg}}\sigma_{\text{bg}}}} \cdot \sqrt{\mathcal{L}} = \frac{S}{\sqrt{B}} \; . \tag{3.6}$$

To claim a discovery the signal needs to exceed the background by at least 5σ, which is equivalent to

$$\frac{S}{\sqrt{B}} > 5 \; . \tag{3.7}$$

assuming $S \ll B$.

3.1 Jet Algorithms

Observing free quarks or gluons is impossible. Such final state particles will emit final state radiation (FSR), hadronize and decay in several stages. Therefore for each such particle in the final state a whole bunch of hadrons will arrive at the detector. To get the kinematic properties of the quark we have to combine these hadrons into so-called jets, containing all decay products of the final state particle. This is done using jet algorithms. One distinguishes two kinds of jet algorithms: Cone algorithms and recombination algorithms. The former are by now obsolete since it turned out they are not infrared save. Therefore, we will only consider recombination algorithms. They use a certain distance measure to decide if the hadron still belongs to the jet or not. This measure is given by

$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2} \min\left(p_{T,i}^{2n}, p_{T,i}^{2n}\right) \, , \; d_{iB} = p_{T,i}^{2n} \; , \tag{3.8}$$

where d_{iB} is the beam distance and R is the size of the jet. Usually one chooses $R \in [0.4, \; 0.7]$. The ATLAS collaboration uses $R = 0.4$ and $R = 0.6$ for their analyses [21]. To recombine heavier particles like the top quark to jets one uses so-called fat jets with $R \sim 1.5$.

There are three kinds of recombination algorithms, which differ in the choice of the parameter n: For $n = 0$ we have the purely geometric Cambridge/Aachen (C/A) algorithm [18] and for $n = \pm 1$ we have the (anti-) k_T algorithms, respectively [19, 20]. The algorithm now proceeds in the following steps:

1. Find combinations of two jets with the minimal distance $d_{\min} = \min_{ij}(d_{ij}, d_{iB})$ for all jets.

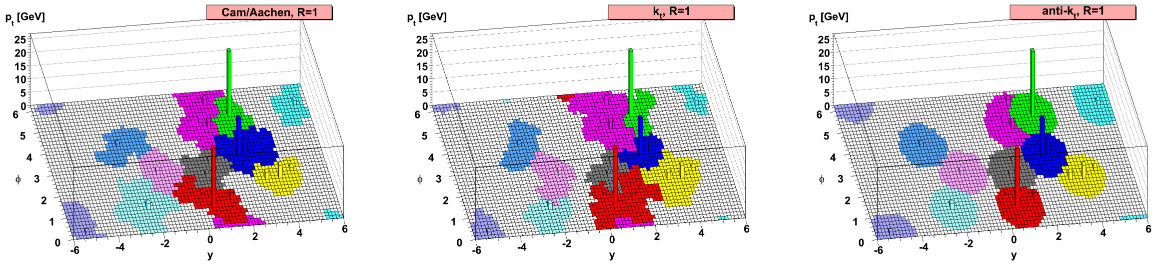

Figure 3: Jet shapes for the different recombination algorithms in the $y - \phi$ plane. Taken from [20].

2. (a) If $d_{\text{min}} = d_{ij} < d_{\text{cut}}$ join jets i and j together (usually by summing their four momenta, which is called E-scheme) and return to step 1.
   (b) If $d_{\text{min}} = d_{iB} < d_{\text{cut}}$ consider i to be beam radiation, disregard it and return to step 1.
   (c) If $d_{\text{min}} > d_{\text{cut}}$ keep all jets and stop.

Here the variable d_{cut} was introduced, which serves as a cut of the jet separation. In exclusive jet algorithms this number has to be put in, while in inclusive jet algorithms it is $d_{\text{cut}} = d_{iB}$, i.e. the jets in step 2 (b) are kept as jets.

Obviously, the shapes of these jets differ in their different distant measure. The k_T algorithm should yield the most accurate results, considering the physical aspects, since it combines the soft constituents first. Conversely, the anti-k_T algorithms starts with combining the hard constituents and is physically ill-motivated. However, as one can see in Fig. 3 the anti-k_T algorithm achieves circular jet shapes, which is why the experimentalists like to use it to easily remove the effects of underlying event. To be able to compare our results with experimental data we will, therefore, use the anti-k_T algorithm throughout this study.

The jets might not only consist of the decay products but also other QCD effect contribute to the jet's quantities:

Initial and final state radiation

Each particle can emit radiation. Radiation emitted by the initial (final) particles is called initial (final) state radiation, in short ISR (FSR). While we would like to remove ISR in the clustering we want to keep FSR, since it contributes to the particle's momentum.

Underlying Event (UE)

In hadron-hadron collisions it is possible that there are interactions of the particles not contributing in the hard process. This is called underlying event and contributes to the jet mass in the following manner

$$\langle \delta m_j^2 \rangle \simeq \Lambda_{UE} p_{T,j} \left(\frac{R^4}{4} + \frac{R^8}{4608} + \mathcal{O}\left(R^{12}\right) \right) , \tag{3.9}$$

where we introduced the parameter Λ_{UE}. This equation shows why it is important to keep the jet size as small as possible.

Pile-up

A proton beam at the LHC is made out of bunches of protons instead of single ones. This allows for more than one collision per bunch crossing and makes it harder to distinguish the individual events. This piling up of events is called pile-up and can be removed using filtering methods.

3.2 Event Generation using Monte Carlo methods

To make predictions about the experimental results of colliders, Monte Carlo event generators are used. They use Monte Carlo methods, i.e. techniques that use (pseudo-) random numbers to compute numerical results, to simulate events at a collider. This is possible because the event can be separated into several stages, which are simulated separately. There is the hard process, which one is usually interested in, the radiation of the incoming and outgoing particles, which is covered by the parton shower, the hadronization of those particles and the decays of the unstable hadrons. Additionally, there is as already mentioned the underlying event. A schematic of such an event can be seen in Fig. 4. In this section, we will briefly describe these subprocesses in more detail.

3.2.1 Hard Process

Most interesting processes include the production of heavy particles and therefore large momentum transfers. This is called the hard process. Thanks to the asymptotic freedom of those particles the amplitude of this process can be computed perturbatively. The most prominent method of doing this is using Feynman diagrams. The important quantity is of course the cross section. In a hadron collider, the cross section for a scattering process $ab \to X$ is computed using collinear factorization. For a partonic cross section $\hat{\sigma}_{ab\to X}$ this

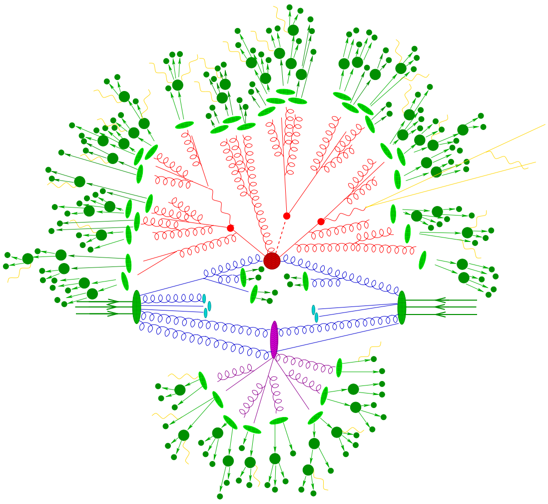

Figure 4: Schematic of an event at a collider. The hard process denoted by the big red blob is apparently a $t\bar{t}H$ process as the decay products show (small red blobs). The underlying event is shown in pink, the parton shower takes place in the outgoing particles in pink and red. The light green blobs signify the hadronization and the dark green plots the decays of these into stable hadrons. Taken from [23].

gives

$$\sigma = \sum_{a,b} \int_0^1 \mathrm{d}x_a \mathrm{d}x_b f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \hat{\sigma}_{ab\to X}(\mu_F, \mu_R) \;, \tag{3.10}$$

where $f_a^{h_1}(x_a, \mu_F)$, $f_b^{h_2}(x_b, \mu_F)$ are the parton distribution functions (PDFs). $f_a^{h_1}(x_a, \mu_F)$ for example gives the probability that for a hadron h_1 with momentum k, the parton a in this hadron has the momentum fraction $x_a k$. μ_F and μ_R are the factorization and renormalization scale, respectively. The partonic cross section for a hadronic center of mass energy s is given by

$$\sigma_{ab\to X} = \frac{1}{2\hat{s}} \int \mathrm{d}\Phi_X |\overline{\mathcal{M}}_{ab\to X}|^2 \;, \tag{3.11}$$

with the parton flux $1/(2\hat{s}) = 1/(2x_a x_b s))$ and the matrix element squared $|\overline{\mathcal{M}}_{ab\to X}|^2$, averaged over initial spins and color degrees of freedom.

The differential phase space element $\mathrm{d}\Phi_X$ of the final state X, consisting of n particles is given by

$$\mathrm{d}\Phi_X = \prod_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 2E_i} \cdot (2\pi)^4 \delta^{(4)}(p_a + p_b - \sum_{i=1}^{n} p_i) \ , \tag{3.12}$$

with the initial state momenta p_a and p_b.

Since any summation over quantum numbers can be pulled outside the squared matrix element this allows not only to sample the phase space using Monte Carlo methods, but also the helicity and color states.

In most event generators a comprehensive list of leading order (LO) matrix elements for $2 \to 1,\ 2,\ 3$ production channels for the SM framework and in some cases even for some BSM frameworks are included. For processes with more final states matrix element and phase space generators are used. SHERPA [23], for example, has built in two such generators: AMEGIC++ [24] and COMIX [25].

To fully specify the cross section in Eq. (3.10) we need to choose values for the renormalization and factorization scale as well as a parametrization of the PDFs. There is not one correct value for these scales, but there is certainly a limited range of reasonable values. For $2 \to 2$ and $2 \to 1$ processes one usually chooses the scale such that $Q^2 = \mu_F = \mu_R$. If one considers the production of a resonance of mass m one chooses $Q^2 = m^2$, where for the production of two massless particles one takes their transverse momentum squared as a scale, $Q^2 = p_T^2$. In event generators this scale Q^2 plays also the role of a starting scale for the parton shower. The PDFs can be chosen arbitrarily, as long as the parametrization allows for an accurate calculation of the cross section. The “Les Houches Accord PDF” (LHAPDF) [27] delivers an interface for commonly used PDF sets, that can be accessed by event generators, where each generator has one default PDF set. Changing the default PDF set might affect the predictions of parton shower, hadronization and underlying event.

Since most generators only include leading-order matrix elements one runs into trouble, when comparing these predictions with experimental results. While the shape of the distributions is accurately described by only considering leading-order (LO) effects, the normalization will be affected by higher-order corrections. This is taken care of by multiplying the LO cross section with a so-called K-factor, obtained by the ratio of the NLO to the LO cross section.

3.2.2 Parton Shower

The hard process computation is well suited to describe the momenta of the outgoing jet, however does not suffice to describe the internal structure of the jets and the distributions of accompanying particles well, since it is calculated at a fixed order. Therefore, one uses parton shower algorithms to simulate the behavior at higher orders. They are formulated as an evolution on momentum transfer from the high scales of the hard process to the low scales of order 1 GeV, where quark confinement begins to play a role.

To further describe how the parton shower works, let us illustrate this using the electroweak $e^+e^- \to q\bar{q}$ process. The leading-order cross section $\sigma_{q\bar{q}}$ is finite and can be computed perturbatively. We are now interested in the NLO process $e^+e^- \to q\bar{q}g$, where one of the outgoing quarks emits a gluon. If we parameterize the phase space with the opening angle between quark and gluon θ, the energy fraction z of the gluon we get

$$\frac{\mathrm{d}\sigma_{q\bar{q}g}}{\mathrm{d}\cos\theta \mathrm{d}z} \approx \sigma_{q\bar{q}} C_F \frac{\alpha_s}{2\pi} \frac{2}{\sin^2\theta} \frac{1+(1-z)^2}{z} \, , \tag{3.13}$$

with $C_F = \frac{N_c^2-1}{2N_c}$ and neglecting non-divergent terms. Since $\sigma_{q\bar{q}}$ appears in this equation we can consider the remaining factors to be the probability of gluon emission. Apparently this equation diverges in the collinear ($\theta \to 0$ and $\theta \to \pi$) and soft ($z \to 0$) limit.

Even though this divergence can be handled and does not matter for the physical description, we need to consider these regions, since there the emission distribution will be large.

First we want to study the collinear limit. There we can approximate $\sin^2\theta \sim \theta^2$ and $\mathrm{d}\cos\theta \sim \mathrm{d}(1+\theta^2) = \mathrm{d}\theta^2$ and therefore obtain

$$\mathrm{d}\sigma_{q\bar{q}g} \approx \sigma_{q\bar{q}} \sum_{\text{partons}} C_F \frac{\alpha_s}{2\pi} \frac{\mathrm{d}\theta^2}{\theta^2} \mathrm{d}z \frac{1+(1-z)^2}{z} \, . \tag{3.14}$$

This equation is completely general. Therefore, for any hard process with cross section σ_0 that has outgoing partons of flavor i the change in the cross section when emitting a parton j with momentum fraction z is given by

$$\mathrm{d}\sigma \approx \sigma_0 \sum_{\text{partons},i} \frac{\alpha_s}{2\pi} \frac{\mathrm{d}\theta^2}{\theta^2} \mathrm{d}z P_{ji}(z,\phi) \mathrm{d}\phi \, , \tag{3.15}$$

where $P_{ji}(z,\phi)$ is a set of universal but flavor dependent functions and ϕ the azimuth of j. As already stated, the factors beside the cross section describe the probability of gluon emission. Therefore, we can compute the distribution of the total probability of all branchings of a parton i between q^2 and $q^2 + \mathrm{d}q^2$ with the virtuality q^2 since as one can show we have $\frac{\mathrm{d}\theta^2}{\theta^2} = \frac{\mathrm{d}q^2}{q^2}$. This distribution is given by

$$\mathrm{d}\mathcal{P}_i = \frac{\alpha_s}{2\pi} \frac{\mathrm{d}q^2}{q^2} \int_{Q_0^2/q^2}^{1-Q_0^2/q^2} \mathrm{d}z P_{ji}(z) \ , \tag{3.16}$$

where the limits on the integral are required to make sure the partons are resolvable and hence get rid of the divergence. Q_0 therefore describes the point below which we cannot distinguish the emitted parton from the emitting parton anymore.

We are now interested in the probability that there are no branchings for virtualities greater than q^2 given a maximum virtuality Q^2. This function we define to be $\Delta_i(Q^2, q^2)$ given by the differential equation

$$\frac{\mathrm{d}\Delta_i(Q^2, q^2)}{\mathrm{d}q^2} = \Delta_i(Q^2, q^2) \frac{\mathrm{d}\mathcal{P}_i}{\mathrm{d}q^2} \ . \tag{3.17}$$

Apparently, the result of this differential equation is an exponential

$$\begin{aligned} \Delta_i(Q^2, q^2) &= \exp\left[- \int_{q^2}^{Q^2} \frac{\mathrm{d}k^2}{k^2} \frac{\alpha_s}{2\pi} \int_{Q_0^2/k^2}^{1-Q_0^2/k^2} \mathrm{d}z P_j i(z) \right] \\ &\overset{q^2=Q_0^2}{\approx} \exp\left[-C_F \frac{\alpha_s}{2\pi} \log^2 \frac{Q^2}{Q_0^2} \right] \ , \end{aligned} \tag{3.18}$$

where in the second line we set $q^2 = Q_0^2$, i.e. calculated the probability of no branching happening at all. This is the so-called Sudakov form factor. Similar to the case of radioactive decay $\Delta_i(Q^2, q^2)$ it describes the probability of non-branching as e to the minus the total inclusive branching probability over that region.

With these results we can now iteratively attach additional partons to the hard process one after another. With these additional partons we then have to proceed in the same way until no more branchings need to be applied. However, one needs to be careful to avoid double-counting by requiring the q^2 values of the child partons to be smaller than the ones generated for this splitting.

The Monte Carlo implementation usually looks like this: Choose a random number $\rho \in [0,\ 1]$ and solve $\Delta_i(Q^2, q^2) = \rho$ for q^2. If the solution is above Q_0^2 generate a branching at q^2 otherwise there is no resolvable branching and the evolution terminates. In principle Q_0^2 can be chosen arbitrarily and is a lower cutoff at which two partons cannot be distinguished anymore. However, since we have a running coupling one has to make sure to choose $Q_0^2 \gg \Lambda_{\mathrm{QCD}}$ to avoid the coupling to become of order 1.

Next to the collinear region the soft emission limit leads to dominant contributions to the emission distribution. We will not go into the specifics of this case, however, one can show that these soft gluon effects can be accurately described by a collinear parton shower algorithm as described above. However, this only works if one uses the opening angle θ instead of the virtuality q^2 as evolution scale. Therefore, many parton showers are angular-ordered as the one implemented in HERWIG [48, 49].

3.2.3 Multi-jet Merging

Before we can start to hadronize the outgoing jets of the parton shower, we first have to find a way to combine the hard matrix element calculation with the results of the parton shower. Since both approaches have their advantages and shortcomings we would like to find a way to combine them in a way, which maximizes the merits of both while keeping the influence of their shortcomings as small as possible.

Fixed-order matrix elements work well in describing separated hard partons, however they run into problems when describing collinear and soft partons. For the parton shower it's the other way around. Combining these two approaches is, however, not as trivial as one might think. One reason for this is, that the matrix element calculation is inclusive in giving the probability of at least n outgoing particles at lowest order while the parton shower is exclusive in computing the probability of having exactly n outgoing particles to all orders. Also if one would simply add the matrix element and the parton shower together, it is possible that double counting or under counting will take place.

There are two main schemes to achieve a proper merging: the CKKW and the MLM scheme. In the following we will briefly describe these two algorithms.

The CKKW scheme was developed by Catani, Kuhn, Krauss and Webber [28] and is used as default in SHERPA. First we introduce a matching parameter y_{match}. Two jets i and j are considered to be best described by the parton shower if $y_{ij} < y_{\mathrm{match}}$, while for $y_{ij} > y_{\mathrm{match}}$ the hard matrix element might be the better fit. This can be translated into a virtuality scale by assigning a typical energy/virtuality t_{hard} to the hard process yielding $t_{\mathrm{match}} = y_{\mathrm{match}}^2 t_{\mathrm{hard}}$. This parameter will serve as a lower jet radiation cut-off.

The CKKW scheme now aims at assigning a weight to given phase space points to make sure that the processes will be chosen statistically in the correct way. We start by generating events with n final states and compute the corresponding cross section $\sigma_{n,i}$, where the subscript i describes the different possible final states. After this, the algorithm proceeds in the following way:

1. Compute the relative probability for each final state $P_{n,i} = \sigma_{n,i} \Big/ \sum_{k,j} \sigma_{k,j}$ and select one according to its probability.
2. assign the momenta from a phase space generator to the hard external particles and calculate $|\mathcal{M}|^2$ adding the parton shower below t_{match}.
3. Recover the shower history using a jet algorithm and check if the splittings are in line with possible Feynman diagrams and do not disobey any symmetries.
4. Compute the Sudakov factors, giving the non-splitting probability, of each external and internal line down to t_{match}.
5. Modify the α_s values of each splitting taking into account the k_T scale from the shower history.
6. Use the obtained matrix element, Sudakov factors and α_s values to compute a weight for the event.

This scheme achieves using t_{match} a clear separation of the hard matrix element computation and the parton shower algorithm.

Another method of for the merging of parton shower and hard matrix element computations is the MLM-scheme [29]. In this scheme also n-jet events are simulated including hard and parton shower jets. Then we apply a jet algorithm on these events. For MADGRAPH the k_T-algorithm is used. Now we compare the showered events with the unshowered ones and try to identify each parton with a parton shower jet. If we can identify a jet to each parton without any jets remaining we know that the parton shower did not change the structure of the event. However, if we observe significant differences between the parton level event and the showered event it has to be dropped.

3.2.4 Hadronization

Usually, by hadronization we mean the model used to describe the transition from the partonic final state to the actually measured hadronic final state. Since quarks carry color charge, due to color confinement one will only be able to measure the hadrons containing the quarks. For this there up to now exist only models, which are, however, inspired by QCD. The reason for this is, that at the energy scale where hadronization occurs QCD gets non-perturbative and can only be described by lattice QCD, which since it is formulated in Euclidian space-time is not suited for the Minkowskian time evolution of the partons.

There are two main kinds of hadronization models: string and clustering ones. In Appendix B a brief introduction into the basics of these models is found. Let us just mention here, that SHERPA and HERWIG use the clustering model, while PYTHIA uses the string model.

3.3 Boosted Decision Trees

The purpose of colliders like the LHC is to search for new physical processes, which are considered signals. Each time such a process is successfully observed it is moved from being a signal to being background. Apparently, over time this will lead to a large amount of background making it harder to distinguish the desired signals from the huge backgrounds.

Therefore, sophisticated methods for separating the signal from the background have been developed. Restricting the analysis on only one variable is only rarely sufficient and multivariate analyses that take into account sets of variables at once have become common practice by now.

One of these methods are boosted decision trees (BDT). We will first describe what a decision tree is and after that consider the method of boosting.

A decision tree serves in our purposes to classify if the considered event is signal or background using phase space variables. This is done by repeatedly applying binary splits, where at each node the variable is used, which allows for the best discrimination between signal and background at this stage. Therefore, some variables might be used several times, while others aren't considered at all. This is done until a minimum cut on the events in one node is achieved. The resulting nodes at the bottom are called leaf nodes. A schematic of such a tree is shown in Fig. 5.

In contrast to rectangular cuts, which select one hypercube in the phase space as a separation of signal and background, the decision tree can split the phase space into many hypercubes of arbitrary size each corresponding to signal or background. This will in general improve the performance of decision trees compared to rectangular cuts. However, one needs to be careful, since one in principle could split the phase space such that each leave node consists of only one event. Such a tree would describe the data perfectly, but since it describes also statistical fluctuations it would have less predictive power. This

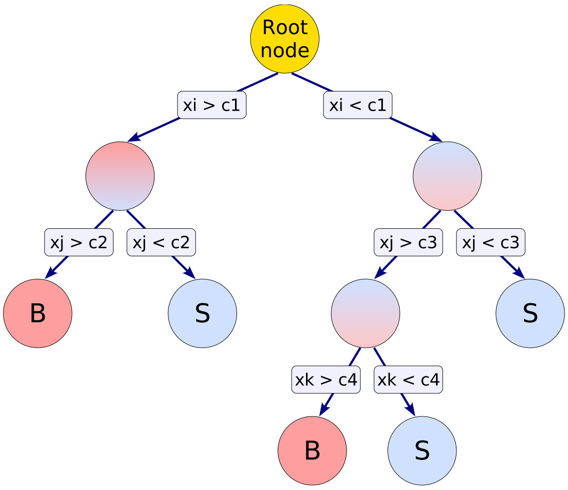

Figure 5: Schematic of a decision tree. Taken from [54].

is the problem of overtraining and is circumvented by imposing a minimum cut on the number of events in the resulting leaves.

Training (building) a tree is done by starting at the root node and determine the variable and cut, which achieves the best separation on the whole sample. Now, one proceeds to checking the subsets at the splittings for the optimal variable and cut to split those on. This is continued until the minimum amount of events is reached. A leaf is considered signal (background) if the majority of the events in the leaf are signal (background) events.

This procedure makes it possible that the splitting goes on until each leaf node only consists of signal or background. While this might seem worthwhile it is probable that in such a case the tree is highly overtrained. Therefore, a so-called pruning that simplifies the tree has to be applied.

Pruning is done by removing nodes that only achieve a small separation of signal and background and, hence, are less significant. To first train the tree and apply the pruning afterwards is better than prohibiting such nodes in the first place, since these nodes still affect the splittings of the remaining nodes in the tree.

As mentioned, some variables might be used more than once and others not at all. This allows for a measure in the importance of a variable. This importance is a function of the number of times it is used in the tree, the number of events in each node it's used and the separation improvement it achieves in each node.

However, decision trees do have a weakness. If we have two variables which are similar in the amount of separation they can achieve, but due to a fluctuation in one variable it is preferred towards the other one, this will possibly change the whole structure below this node.

This can be solved by constructing several decision trees, so-called forests. The classification of an event is now given by a majority vote of the classification of each tree in the forest instead of one single tree. The trees in the forest are generated using the same training sample, but each time weighting the events differently. This is called boosting. Through boosting the classifier becomes stable under statistical fluctuations and improves in performance.

There are several boosting methods that have been developed over time. In our boosted decision tree analyses we will use an algorithm called AdaBoost [55]. This adaptive boosting algorithm is able to take a weak learning algorithm, i.e. one that achieves results that are not significantly better than random guessing, and “boosts” it into a strong algorithm. In Appendix C we show how the AdaBoost algorithm works.

Finally, to make sure that a BDT analysis is not suffering from overtraining one usually applies an overtraining check. This check is done by splitting the sample into a training and a test sample. That is, instead of using the whole sample for training we also use one part to test the trained hypothesis. If no overtraining occurs both samples should yield the same prediction except for statistical fluctuations. Most times overtraining hints that we use too few statistics. Therefore, if the analysis suffers from overtraining increasing the sample size will solve it most of the time.

There is no clear rule to how much of the sample should be used for training and how much for testing. Both samples should be large enough to effectively rule out overtraining. Throughout this study we will use 50% of the sample for testing and training, respectively.

4 Higgs invisible decays in weak boson fusion

We know by now that the current description of particle physics, the SM, is incomplete. This is due to cosmological findings that suggest that there is matter, which cannot be described by any particles in the SM. This matter is at most weakly and gravitationally interacting, and hence, invisible to todays detectors. Therefore, this matter is referred to as dark matter (DM).

After the Higgs boson has been found, the main goal of the LHC is to search for new physics, including dark matter. As we will see in Section 4.2, searches for invisible Higgs decays in weak boson fusion Higgs production are particularly promising.

To describe dark matter properly the proposed dark matter particle has to recover the relic density. The density parameter is given by the ratio of the observed density to the critical density, where the critical density signifies the junction between an expanding and a collapsing universe for a vanishing cosmological constant. The relic density Ω_χ was measured by the Planck satellite to be $\Omega_\chi h^2 = 0.1198 \pm 0.0015$ [36], where h is the dimensionless Hubble parameter.

In the following, we will outline why the Higgs might be a portal to observing dark matter particles and how the properties of the weak boson fusion process can be used to allow for a good separation from the backgrounds.

4.1 Higgs Portal

One possible candidate for a dark matter can be obtained by simply extending the SM introducing a scalar field S, which will then be coupled to the Higgs. This is, particularly, interesting, since the so-called Higgs portal interactions are renormalizeable, i.e. the coupling constant two of these new scalars and two Higgs bosons will be of mass unit zero.

Before we will couple our new scalar to the SM, let us first discuss, which properties the new scalar should have.

First, we need to clarify if we want the scalar to be real or complex, that is, if we want to introduce two particles with a positive and negative charge or not. We decide for the latter, giving us a real scalar field.

The second property we need to determine is, if we want the scalar, like the Higgs, to have a nonzero vacuum expectation value (vev) or not. Since such a vev would contribute to the masses of gauge bosons and fermions in the SM it would only overcomplicate the underlying physics. Therefore, we will assume the new scalar to have no nonzero vev.

If we were including the terms for the S in the potential of Eq. (2.25) there would be terms $\propto H^2S$. This would allow for decays of the S particle into two Higgs bosons on-shell, $S \to HH$ and off-shell, $S \to H^*H^* \to 4b$. This is not what we want, since dark matter should be stable or at least with a decay time comparable to the age of the universe.

This can be solved by imposing a so-called global $\mathbb{Z}_2$ symmetry

$$S \to -S \qquad H \to +H \; . \tag{4.1}$$

Apparently, this gets rid of any terms of uneven powers in S.

Therefore, with the S the scalar potential of Eq. (2.25) becomes

$$V = \frac{m_H^2}{2}H^2 + \frac{m_H^2}{2v_H}H^3 + \frac{m_H^2}{8v_H^2}H^4 - (\mu_S^2 - \lambda_3\frac{v_H}{2})S^2 + \lambda_S S^4 + \lambda_3 v_H H S^2 + \frac{\lambda_3}{2}H^2S^2 \; , \tag{4.2}$$

where we expressed the λ of Eq. (2.25) in terms of the Higgs mass using $m_H^2 = \frac{1}{2}\lambda v$. Furthermore, we added the subscript H to distinguish the Higgs from the S terms.

From this equation, we can read of the mass term and the couplings of the S:

$$m_S = \sqrt{2\mu_S^2 - \lambda_3 v_H^2} \; , \qquad g_{SSH} = -2\lambda_3 v_H \; , \qquad g_{SSHH} = -2\lambda_3 \; . \tag{4.3}$$

To detect the S particle at the LHC we want to observe the process $pp \to H^* \to SS$. The decay width of $H \to SS$ for $m_S < m_H$ is given by

$$\begin{aligned} \Gamma(H \to SS) =& \frac{\lambda_3^2 v_H^2}{32\pi m_H}\sqrt{1 - \frac{4m_S^2}{m_H^2}} \\ \approx& \frac{m_H \lambda_3}{8\pi}\sqrt{1 - \frac{4m_S^2}{m_H^2}} < \frac{\lambda_3}{8\pi} \; , \end{aligned} \tag{4.4}$$

where we used that $m_H \approx \frac{v_H}{2}$. The SM prediction for the total width of the Higgs boson is $\Gamma_H = 4 \cdot 10^{-5} m_H$.

So if we would have a invisible branching ratio BR($H \to SS$) of 10% for light scalars $m_S \ll m_H/2$ the portal coupling λ_3 has to be

$$\frac{\lambda_3}{8\pi} = 4 \cdot 10^{-6} \qquad \Rightarrow \qquad \lambda_3 = \sqrt{32\pi}10^{-3} \approx 10^{-2} \; . \tag{4.5}$$

As is shown in Ref. [34] to be in agreement with the relic density, the light scalar dark matter has to have a portal coupling of $\lambda_3 = 0.3$ with $m_S \lesssim 50$ GeV.

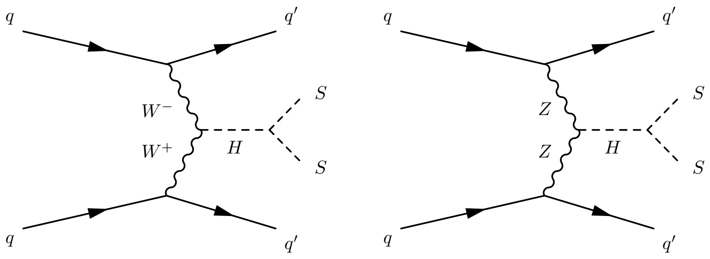

Figure 6: Feynman diagrams for the weak boson fusion Higgs production with the Higgs decaying to the scalar S and a mediating W (left) and Z boson (right), respectively.

Currently, ATLAS and CMS searches at the LHC are, in particular, considering the weak boson fusion Higgs production process to search for invisible Higgs decays [39, 40]. The LHC running at high luminosities (around 3000 fb^{-1}) will be sensitive to invisible branching ratios around

$$\text{BR}(H \to \text{invisible}) = (2 \ldots 3)\% \,, \tag{4.6}$$

if we assume that the SM Higgs production rate is essentially unchanged [37].

With the results from Eq. (4.5) and the bound on the portal coupling from the relic density, we see that a branching ratio of 10% would be too small for the Higgs portal to account for dark matter. Since the LHC at high luminosities will be sensitive to even smaller branching ratios it will be possible to test this model for light scalar dark matter in the near future.

4.2 Weak Boson Fusion and Tagging Jets

The Higgs boson can be produced in three main production channels: gluon fusion, weak boson fusion and associated production with W and Z bosons or top quarks. While gluon fusion is the dominating process at the LHC, the weak boson fusion (WBF) channel is promising in the search for invisible Higgs decays.

The Feynman diagrams for this process are shown in Fig. 6. The WBF Higgs production is interesting, because its kinematics allow for a efficient separation from QCD backgrounds. This is due to the two forward jets in the hard process. These tend to be back-to-back with large longitudinal momentum. This means that the invariant mass of these jets m_{jj} will be of order $\mathcal{O}(\text{TeV})$ at the LHC. This is in stark contrast to the QCD backgrounds.

Therefore, these jets can be used for tagging the signal and are hence called tagging jets. This is a prime example of how the features of a specific production process can help to reduce the background. This distinctive signature, however, comes at a prize: at the LHC WBF production rate is roughly a factor of (1/10) below gluon fusion.

In Fig. 6, we can see that there is no color flow between the two quarks. This means that non-collinear gluon radiation will not be possible in this process. That is, additional jet activity will be restricted to collinear radiation along the two tagging jets. Therefore, we expect the hadronic activity in general to be small between those jets. Since such a behavior will not occur in most QCD background processes this is a further feature of the WBF process that can be used to separate signal and background by looking at the central jets. By central jets we mean jets that are between the two tagging jets, i.e. jets with $\eta_{j_1} < \eta_j < \eta_{j_2}$, where η_{j_1} and η_{j_2} correspond to the pseudorapidities of the tagging jets. This will be later referred to as the pseudorapidity- or η-gap. It can be shown that most backgrounds can be reduced by vetoing central jets with transverse momenta above $p_{T,j} > 30$ GeV. This is what is usually called the central jet or minimal jet veto.

As we have seen using the WBF Higgs production process, allows for invisible Higgs decay searches. However, there are still challenging backgrounds that have to be treated. One of them is the $Z \to \nu\bar{\nu}$ QCD background. Since the Higgs boson decays invisibly, it is only possible to use the tagging jet information for the reconstruction. Usually by distinguishing the jet kinematics and multiplicities of these processes, a separation can be achieved. In such an analysis one uses that in the QCD background there is usually more jet activity.

Relying only on the jet properties of the process might not always be sufficient. Therefore, it is interesting to see if using the information of tracks might improve the discrimination of signal and background. By tracks we mean charged particles, that would be detected in the tracker of a detector. Due to the higher multiplicity and being sensitive to softer p_T regions using track information might achieve higher significances than using jets.

Furthermore, this might even allow to distinguish between the Higgs WBF production and the Z WBF production. With the Z decaying to neutrinos these processes would consist of almost the exact same properties, including the tagging jets feature. This makes a discrimination of these processes a challenging task. Due to the different masses of the Z and the Higgs, there might be some differences in the radiation pattern. Hence, using the jet activity might still be useful to separate those processes. But, since we expect the radiation differences to be small using tracks instead of jets might work even better.

The aim of this study will be to examine if using tracks instead of jets for the separation of the WBF processes to each other and to the Z QCD background will improve the sensitivity.

5 Comparison of Monte Carlo generators

Before we start our analysis of tracks and jets we want to find the Monte Carlo event generator that serves our purposes best. As we have seen in Section 3.2 there are different models simulating events used in event generators. Also there are different ways to tune the parameters.

Therefore, to avoid predictions that depend rather on the subtleties of an event generator than actual physics we need to compare several generators and choose the one which is best suited for our analysis. Event generators usually are in good agreement to with data at high-p_T. Therefore, especially the modelled physics of the generators at low-p_T is interesting, since this regime is highly model dependent.

This comparison will be divided into two parts: In the first part we will compare the generators to a data using jets constructed from charged tracks. This serves to examine the performance of the generators in simulating the tracks accurately. The second part will focus on processes which is are important for the study of track properties in Section 6. The used processes of Z production will later serve as signal and background. In both parts we will compare our results to data. By checking which generator works best in generating accurate tracks in the respective processes we should be able to determine the ideal event generator for our purposes.

We will compare the three event generators Sherpa [23], Pythia [47] and Herwig [48, 49]. However, to circumvent the limited number of processes that Pythia provides we will generate the matrix elements using MadGraph [50] and use Pythia for the parton shower and hadronization of these events. This event generation method we will denote as MadGraph+Pythia. In appendix D it is shown that both methods (using Pythia directly and using MadGraph for the matrix elements) achieve the same results.

5.1 Jets from charged tracks

To see which generator serves our purposes the best a comparison to actual data is preferable. Thanks to Rivet such a comparison is possible without further ado. The Rivet toolkit serves to validate Monte Carlo generators by comparing their results to experimental data. The set of analyses contained in Rivet [51], however, is limited.

Since in our later study we will consider the properties of charged tracks, the generator we will use should achieve a simulation of charged tracks that is in close agreement with the data. Hence, the main goal of this analysis is to see which generator describes the charged tracks the best.

RIVET contains a $pp \to$ jets analysis provided by ATLAS [44], which we use in the first part of this comparison analysis. In this analysis jets are constructed using charged tracks with $p_T > 300$ MeV with the anti-k_T algorithm and cone size $R = 0.4$, at a c.o.m. energy of $\sqrt{s} = 7$ TeV.

5.1.1 Analysis

We now use SHERPA, MADGRAPH+PYTHIA and HERWIG to generate $pp \to$ jets events and compare them to data using RIVET. Except for HERWIG, where we generate $pp \to jj$, we generate a merged sample of two or three jets in the final state. We expect that each of these generators to be in close agreement to data for high transverse momenta. However, at low transverse momenta differences in the generators will become apparent since every generator models the parton shower and hadronization differently. In all generators the c.o.m. energy $\sqrt{s}$ and minimum transverse momentum cut on the jets p_T^{jet} is set to be

$$\sqrt{s} = 7 \text{ TeV} , \qquad p_T^{\text{jet}} > 4 \text{ GeV} . \tag{5.1}$$

In the following we will give further details on the setup chosen for each generator.

Sherpa

For the Sherpa analysis we generate $pp \to jj(j)$, that is a merged sample of at least two and up to three jets in the final state.

MadGraph+Pythia

To improve the statistics we split the phase space into three regions:

$$\text{A: } p_T \in [4,\ 10] \text{ GeV} , \qquad \text{B: } p_T \in [10\ , 40] \text{ GeV} , \qquad \text{C: } p_T > 40 \text{ GeV} . \tag{5.2}$$

Since we apply these cuts on the two hardest jets this gives altogether six phase space regions, which we divide into three parts that are computed separately. Those are: The region where the harder jet is not restricted (except for the lower cut of $p_T > 4$ GeV) and the softer jet is restricted to A, the region where both jets are restricted to B and the region where the harder jet has a lower cut of $p_T > 40$ GeV and the softer jet has a lower cut of $p_T > 10$ GeV. This covers the whole phase space region we are considering. Additionally, we use the cuts put in MADGRAPH by default. Those are:

$$
\begin{aligned}
p_T^{\text{jet}} > 20\text{ GeV},\ p_T^{\ell} > 10\text{ GeV},\ |\eta_{\text{jet}}| < 5,\ |\eta_\ell| < 2.5,&\\
\Delta R_{jj} > 0.4,\ \Delta R_{\ell\ell} > 0.4,\ \Delta R_{j\ell} > 0.4\ ,&
\end{aligned}
\tag{5.3}
$$

where in this analysis the default cut on p_T^{jet} is changed according to the phase space region considered. In MadGraph we generated dijet and three jet events separately and combined them in the final analysis, by merging them according to the ratio of their respective cross sections to each other. This was done for each phase space region separately.

Herwig

For Herwig, we only generate $pp \to jj$ without a possible third jet in the matrix element. Again, we compute certain phase space regions separately. Not only to improve the statistical accuracy but rather to deal with the much higher computation time Herwig claims. With Herwig we part the phase space computation into three parts different from those we use in MadGraph+Pythia: The region were both jets are restricted to B and C, respectively and the region where the harder jet is not restricted but the softer one is restricted to A. This way the region, where the harder jet is not restricted and the softer jet is restricted to B is not included, since including it results in negative cross sections. Since, as we will see later, Herwig with these settings already highly overestimates the data, a complete computation of these phase space regions would only have worsened these results even more.

5.1.2 Results

In Fig. 7 we show the cross sections of charged particle jets as a function of p_T in four rapidity intervals.

Apparently, Herwig highly overestimates the data. In its highest deviation Herwig's prediction more than a factor of two above the data. Agreements with data, appear to be rather coincidental. The smallest deviation that doesn't agree with data is about 40% above the corresponding data point. On average Herwig deviates more than 50% from data. And this despite the fact that in this analysis not the whole phase space region has been covered. Therefore, we conclude that Herwig is not suited for our analysis.

MadGraph+Pythia, only twice deviates more than 80% from data.. For the most part, especially in the intermediate transverse momentum region (40 GeV $\leq p_T \leq$ 80 GeV)

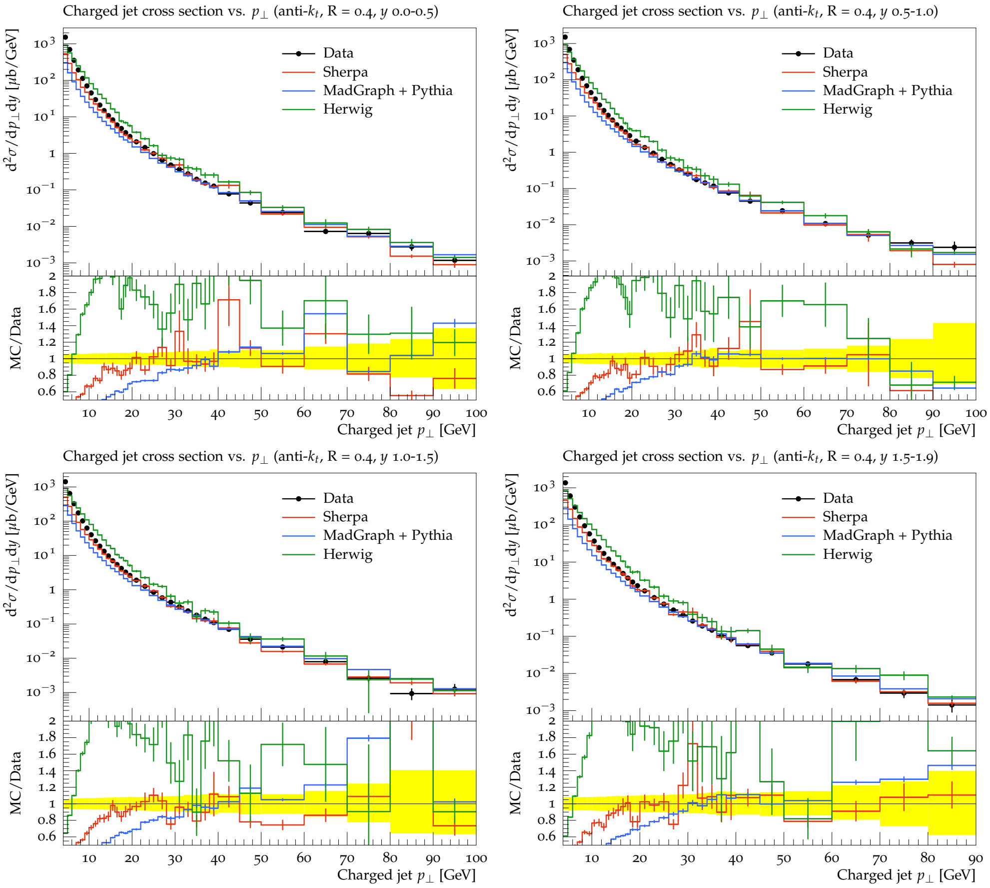

Figure 7: Cross section of charged particle jets as a function of p_T in different rapidity y intervals at $R = 0.4$ using the anti-k_T algorithm.

MadGraph+Pythia is in close agreement to data. In this region it is closer to data than any other of the tested generators.

Sherpa seems to achieve the most reasonable results. Except for one data point it never deviates more than 60% from data. Also the description of the low-p_T regime is closest to data using Sherpa. It is able to predict the data points accurately down to $p_T = 15$ GeV.

In conclusion we can say that, Sherpa and MadGraph+Pythia achieve reasonable results, which agree with data for the most part. Herwig, however, is not at all suited for our purposes. Despite describing most of the data points accurately both Sherpa and MadGraph+Pythia suffer from deflexions at certain points in the p_T spectrum. While this occurs predominantly at low-p_T it also happens at higher p_T, sometimes even yielding a large deviation. The generators failing to describe the process well at low-p_T is due to the $p_T > 4$ GeV cut and the fact that this regime is described in a highly model-dependent way. Apparently none of the generators is suited to describe the behavior in this regime, however, SHERPA seems to achieve the best description. If one looks at the individual deflexions at high-p_T and compares the data points to the neigboring ones, especially for the excess at $p_T = 80$ GeV for $1.0 \leq y \leq 1.5$ it becomes apparent that those data points deviate from the behavior one would expect by extrapolating the other data points to this point. Therefore, we assume that the simulation deviating from these points is due to statistical fluctuations in the data.

In summary, we state that SHERPA and MADGRAPH+PYTHIA seem to be viable candidates for the analysis of tracks. Therefore, in the next part of our comparison we will focus on comparing these two generators.

5.2 Z production

In the last section we ruled HERWIG out as a candidate to use for our analysis. In this section we want to see, which of the two remaining generators, SHERPA or MADGRAPH+PYTHIA, serves our purposes the best.

For this we will use processes similar to the signal and background process, which we will use later in our analysis of tracks. Apparently, the best suited generator for our purposes will be the one who is able to describe these processes the best.

The processes we intend to use are two kinds of Z production with two jets in the final state. For the signal process we use the weak boson fusion Z production denoted as "$Z +$ jets WBF". The background process is any other process with the same final state of a Z and two jets. These processes we will refer to as "$Z +$ jets QCD". Additionally, to compare our results to data, we will let the Z decay leptonically in this part of our analysis. Corresponding Feynman diagrams are shown in Fig. 8.

As before, we expect the generators to achieve basically the same results in the matrix element calculation, however, there will most likely be differences in the parton shower and the hadronization. Therefore, besides the p_T and η spectrum of the Z boson, especially the (parton shower) jet variables will be interesting. Therefore, we will consider the jet multiplicity as well as the p_T distribution of the hardest parton shower jet, which will be the third hardest one overall.

The comparison to data is done using again an analysis by ATLAS [45] where the production cross section of a Z bosons decaying leptonically (electrons and muons) accompanied by up to seven jets. However, before we do this we want to study both processes

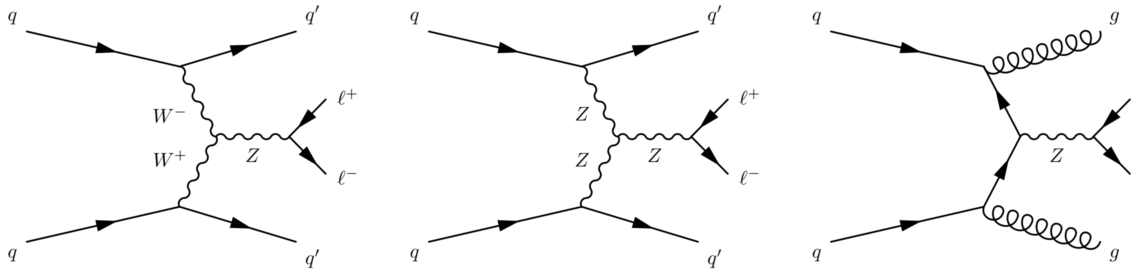

Figure 8: From left to right: Feynman graphs for the WBF signal with mediating W and Z boson and the QCD background. In each process the Z decays leptonically.

individually.

5.2.1 Z + jets WBF

We first start by considering the signal process Z + jets WBF. For both generators we generate $pp \to \ell^+\ell^- jj$ with $\ell = e,\ \mu$, where the Z is produced via weak boson fusion. For the center-of-mass energy we take $\sqrt{s} = 13$ TeV as it is the case for the LHC run II. To make sure that the mediator is a Z we impose $m_{\ell\ell} \in [80,\ 100]$ GeV. Furthermore, we use the default cuts of MadGraph, Eq. (5.3), in both generators to allow for a proper comparison between those two.

The events are analyzed as follows: First we read in the events from a HepMC [52] file and stored the electrons, positrons, muons, antimuons as well as the neutrinos and hadrons separately. Then we took the two leptons of the same flavor with the invariant mass closest to the Z mass $m_Z = 91$ GeV and used them for the Z boson reconstruction. In the next step we took the rest of the leptons, photons and hadrons and clustered them into jets with FastJet using the anti-k_T algorithm with cone size $R = 0.4$ and a transverse momentum cut of $p_T^{\text{jets}} > 10$ GeV. The anti-k_T algorithm with these setting was used following the default ATLAS analyses.

In Fig. 9 one can see the distribution of the p_T and η of the Z boson.

We see that in the plots normalized to the cross section the prediction of Sherpa is above the one of MadGraph+Pythia for the most part. However, in the plots normalized to one this is not the case. This is due to differences in the approach the generators take on the computation. In contrast to Sherpa, MadGraph uses a fixed renormalization and factorization scale at the Z mass. Furthermore, in Sherpa the simulation of the underlying event is included and neglected in MadGraph+Pythia. Further differences

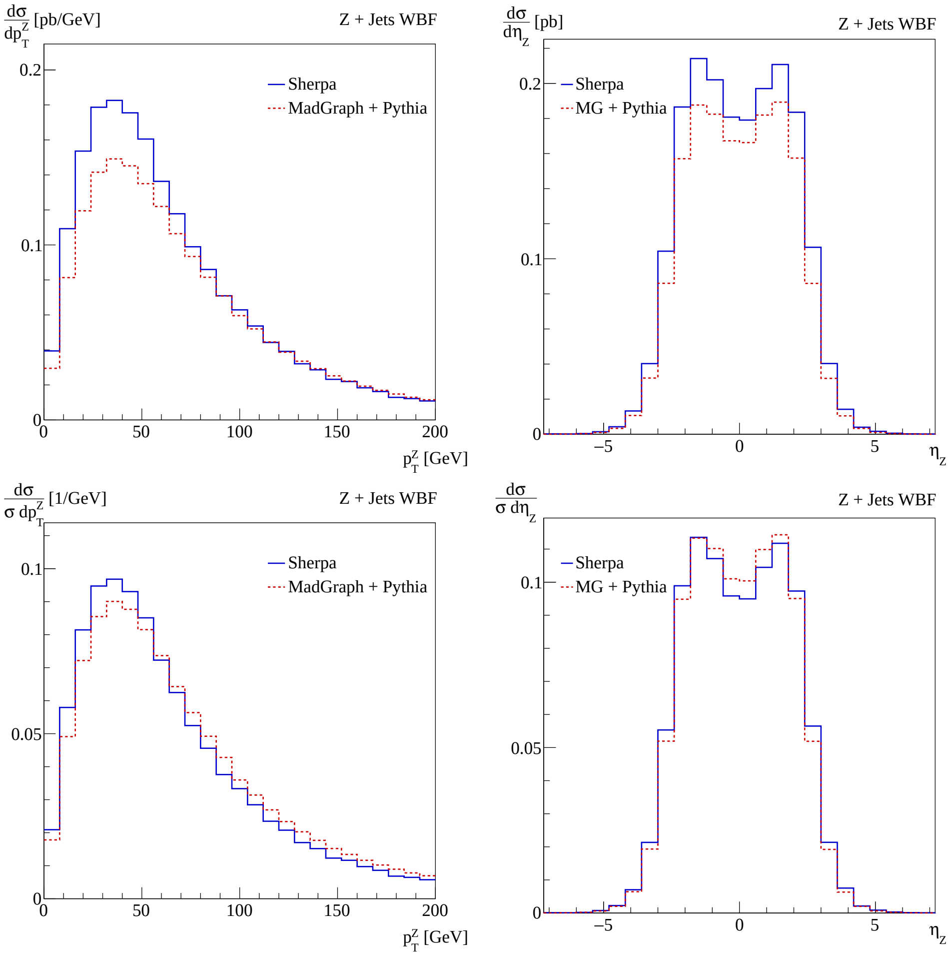

Figure 9: Plots of the p_T (left) and η (right) spectrum of the reconstructed Z boson, normalized to the cross section (upper row) and normalized to one (lower row) in the weak boson fusion process.

occur by the different modelling of the parton shower.

To see that these effects are actually responsible for these changes we generate the events again, using a fixed scale at the Z mass also for Sherpa and neglecting the underlying event in both samples. Additionally we forgo the parton shower in both cases.

In Fig. 10 the p_T distribution of the Z boson can be seen after applying these changes. This increases the agreement of the to generators significantly. Normalized to one the to curves are almost in perfect agreement. However, when normalized to the cross section, there are still large differences, especially at the maximum. This can be explained by scale uncertainties, which are approximately 20%. Furthermore, the remaining differences might be due to different PDFs being used by the respective generators. This, however, was not

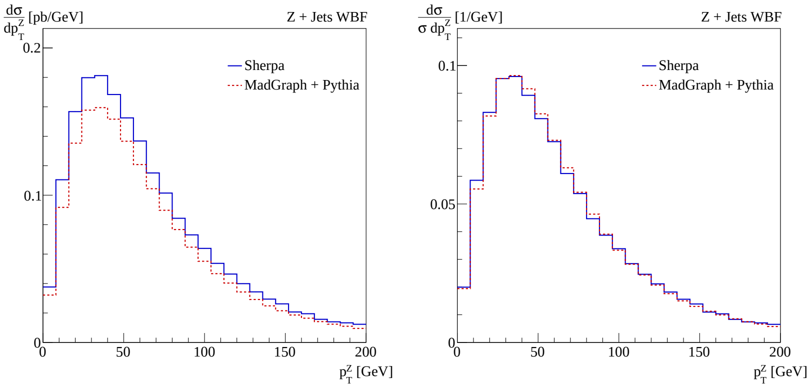

Figure 10: Plot of the p_T spectrum of the Z boson in the weak boson fusion process, where both generators use a fixed scale and neither simulate the underlying event nor the parton shower nomalized to the cross section (left) and normalized to one (right).

investigated further.

Since our default minimum jet p_T cut of $p_T^{\text{jets}} > 10$ GeV is pretty generous we want to see how this behavior changes if we increase this cut. Therefore, we plotted the jet multiplicity and the p_T spectrum of the hardest parton shower jet for different minimum jet p_T cuts as can be seen in Fig. 11. We see in that for $p_T^{\text{jets}} > 10$ GeV SHERPA seems to have more events with 10 jets and more than MADGRAPH+PYTHIA. However, the jets of MADGRAPH+PYTHIA tend to be the harder. Furthermore, increasing the p_T^{jets} cut achieves a better agreement between those generators. Since the cut is imposed in the analysis this cut does not change the behavior of the generator, but rather gets rid of light jets that are due to radiation. Since, in the SHERPA sample the underlying event is included it is understandable why more and softer jets would be present.

5.2.2 Z + jets QCD

For the background process $Z +$ jets QCD we apply the same settings as in the case of the Z WBF process. Again we generate $pp \to \ell^+\ell^- jj$ with $\ell = e,\ \mu$. We only include those cases where the Z gets not produced by weak boson fusion. For the center-of-mass energy we again use $\sqrt{s} = 13$ TeV. To make sure that the mediator is a Z we impose $m_{\ell\ell} \in [80,\ 100]$ GeV. Also, we use the default cuts of MADGRAPH, Eq. (5.3), in both generators.

In Fig. 12 one can see the distributions for the p_T and η of the Z. In these plots we observe a similar behavior as before. The differences between the plots normalized the cross

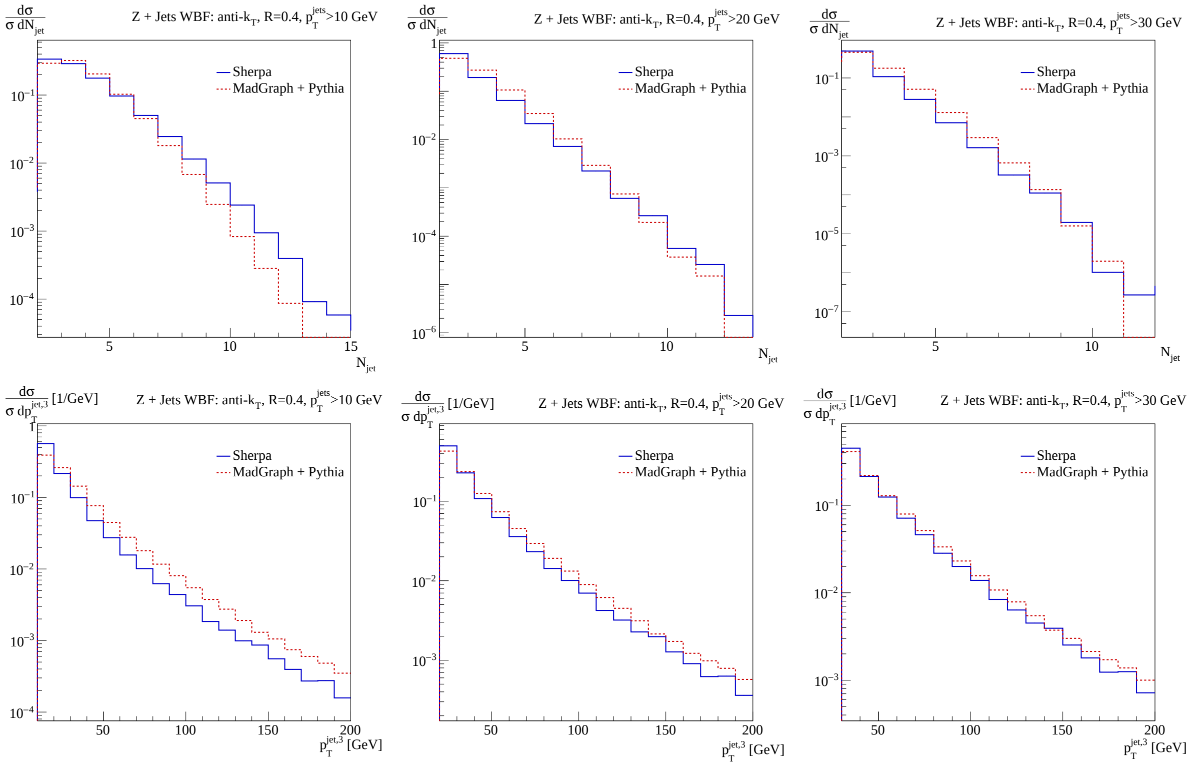

Figure 11: Plots of the jet multiplicity and the p_T spectrum of the hardest parton shower jet normalized to one for three different minimal jet p_T cuts: $p_T^{\text{jets}} > 10$ GeV (left), $p_T^{\text{jets}} > 20$ GeV (center) and $p_T^{\text{jets}} > 30$ GeV (right).

section and those normalized to one can be explained by the same reasons as in the weak boson fusion case. However, there seem to be not only differences in the normalization, but also in the shape of the p_T spectrum of the Z boson. The transverse momentum spectrum of the Z boson tends to be harder for MadGraph+Pythia. To a small extend this can be also observed in the Z + jets WBF process, however is mostly removed after the underlying event and the parton shower are turned off. Therefore, this effect might also be explained by the different treatment of the parton shower. Furthermore, the already mentioned different PDFs the generators use might have contributed to these differences.

As before, we would like to see how increasing the minimum jet p_T cut changes the distributions of the jet multiplicity and the hardest parton shower jet.

This can be seen in Fig. 13. Similar to the former case increasing this cut improves the agreement of both generators with each other, however, in this case the best agreement seems to be obtained already by a cut of $p_T^{\text{jets}} > 20$ GeV. Apparently, for higher cuts Sherpa tends to drop more jets than MadGraph+Pythia, which implies that the Sherpa parton

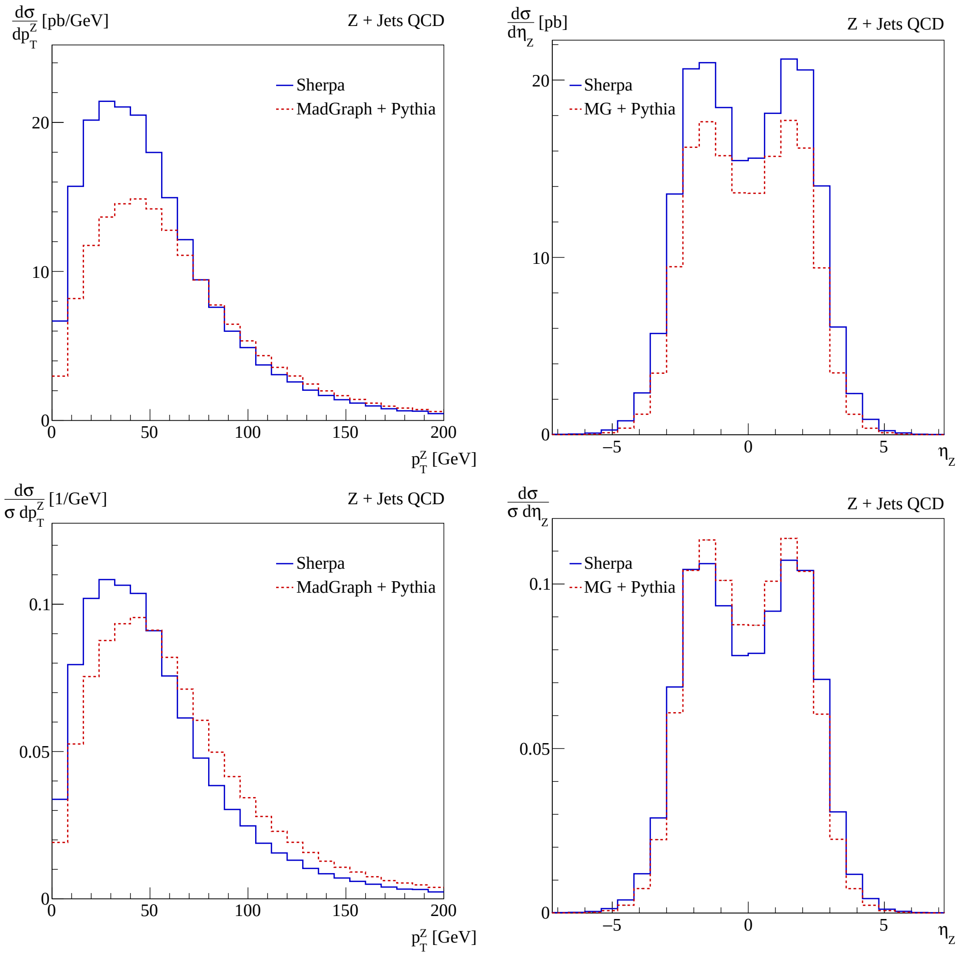

Figure 12: Plots of the p_T (left) and η (right) spectrum of the reconstructed Z boson, normalized to the cross section (upper row) and normalized to one (lower row) in the $Z+$jets QCD process.

shower generates softer jets than the one from MadGraph+Pythia.

We will keep the effects of this cut in mind in our later analysis and consider the changes it has on our results.

5.2.3 Comparison to data

After studying both of these processes in detail we want to see, which of the generators is better suited to describe the physics of those processes. For this we will use a study done by ATLAS [45], where the production of a Z boson accompanied by up to seven jets at $\sqrt{s} = 13$ TeV was studied. The data used corresponds to an integrated luminosity of 3.16 fb^{-1} collected in 2015. The decay of the Z boson was restricted to electrons and

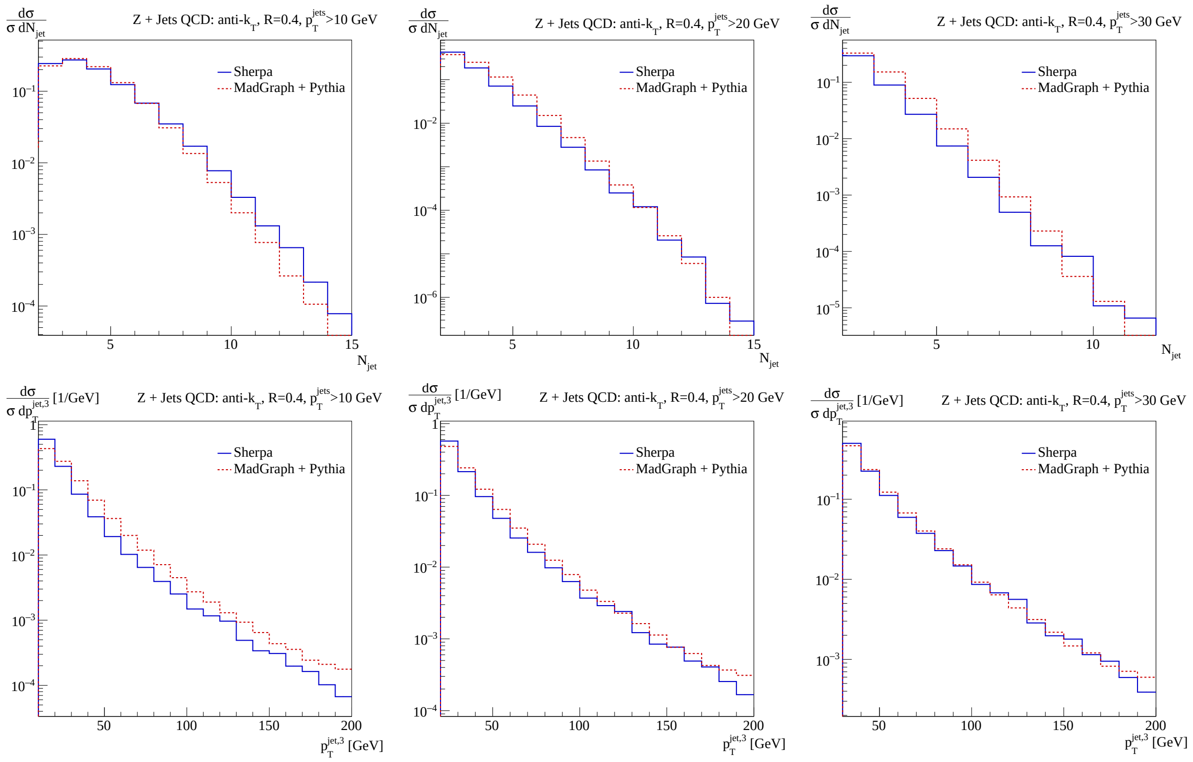

Figure 13: Plots of the jet multiplicity and the p_T spectrum of the hardest parton shower jet normalized to one for three different minimal jet p_T cuts: $p_T^{\text{jets}} > 10$ GeV (left), $p_T^{\text{jets}} > 20$ GeV (center) and $p_T^{\text{jets}} > 30$ GeV (right).

muons and a minimum jet p_T cut of $p_T^{\text{jets}} > 30$ GeV and an rapidity cut of $|y_{\text{jets}}| < 2.5$ was imposed.

This study is well suited to compare to our generated samples. Since both processes contribute to the result of the study we need to combine them and compare them to the data.

In Fig. 14 the results of this comparison are shown. The plots on the upper left and center, respectively, show the jet multiplicity and the distribution of events that are above a lower jet number threshold. At $N_{\text{jet}} < 2$ both generators fail to describe the data accurately, which is to be expected, since such events were not part of our analysis. That there is still a small number of events that fall in that category, is probably due to one or both jets not being resolved by the jet algorithm. However, the rest of the spectrum is in close agreement to data for SHERPA, which has only one deflexion at $N_{\text{jet}} = 6$. MADGRAPH+PYTHIA on the other hand over- or underestimates the data for the most part and only at $N_{\text{jet}} \geq 7$ agrees with the data.

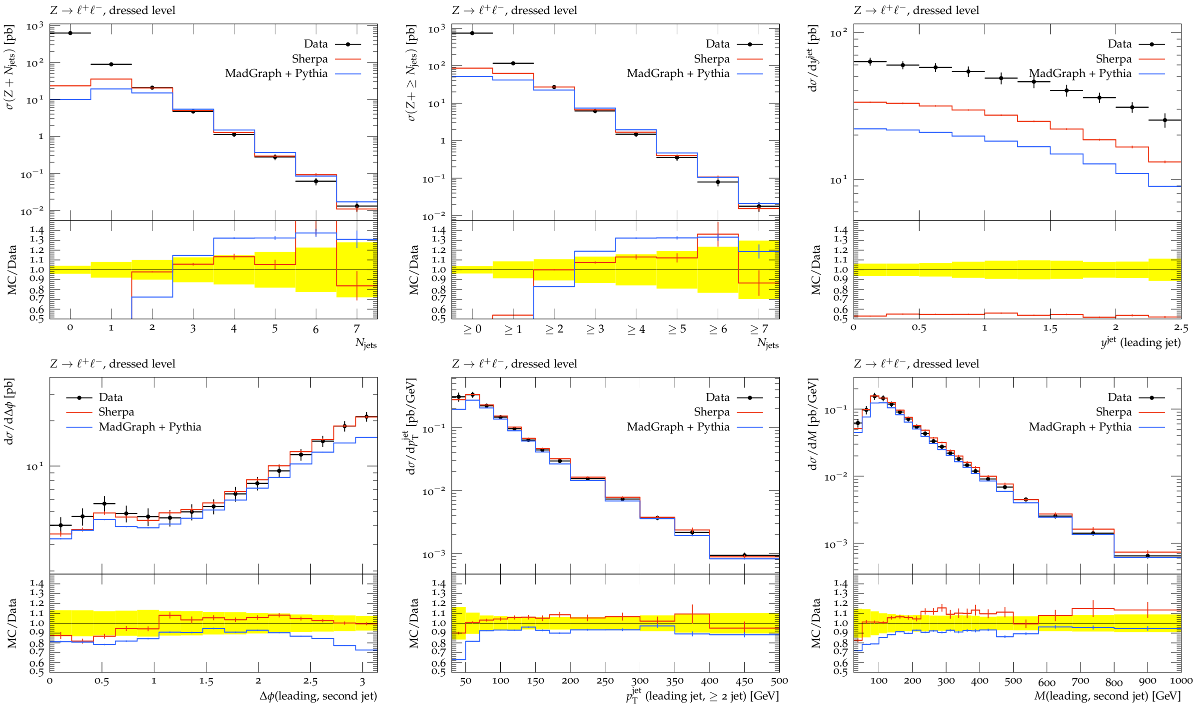

Figure 14: Upper row: Jet multiplicity (left), cross section in dependence of the minimal number of jet in the event (center), rapidity of the hardest jet (right). Lower row: Difference in ϕ between the two hardest jets(left), p_T distributions of the hardest jet if more than two jets are in the event (center) and the distribution of the invariant mass of the two leading jets (right).

The plot in the upper right shows the rapidity of the hardest jet. For the rapidity both generators fail do describe the data well. While the shape is described correctly for both generators, they fail to obtain the right normalization. This is because our samples do not describe the whole cross section due to the absence of events with less than two jets. The fact that SHERPA achieves a better description than MadGraph+Pythia for both $N_{\text{jet}} < 2$ and the rapidity curve, confirms this reasoning.

In the lower left we show the distribution of the difference in ϕ of the two hardest jets. We see that SHERPA is in agreement to data except for one deflexion, while MADGRAPH+PYTHIA achieves reasonable results only in the region of $1 \leq \Delta\phi_{jj} \leq 2.25$.

The plot in the lower center shows the p_T distributions of the hardest jet if at least two jets are present. SHERPA is in total agreement with data, but MADGRAPH+PYTHIA fails to describe the data at low transverse momentum.

The last plot on the lower right shows the invariant mass of the two hardest jets. Both generators are in close agreement to data, however, both suffer from deflexions. While

SHERPA tends to overestimate the data MADGRAPH+PYTHIA tends to underestimate it. Still, SHERPA is in the closest agreement to the data points, which is also due to its error bars being larger than those of MADGRAPH+PYTHIA. This might again indicate that the errors in MADGRAPH+PYTHIA are underestimated.

Finally, we see that SHERPA does a better job at achieving a close agreement with the data than MADGRAPH+PYTHIA. Especially at low-p_T SHERPA seems to achieve a better description.

5.3 Conclusions

This part of our study was aimed at finding out which generator to use in our study of track properties. We considered the three generators SHERPA, MADGRAPH+PYTHIA and HERWIG.

To be a good fit for our purposes those generators need to deliver a good description of the processes we intend to simulate and especially the track and jet properties we will focus our study on. Therefore, in the first part we compared the results of the generators to data, where jets were constructed out of charged tracks. In the second part we focussed on the $Z+$jets WBF and $Z+$jets QCD processes, which we will use in the next part as signal and background in our study of track properties. In this part we also examined some of the differences in the modelling of the generators further. Also the variation of certain cuts was put to a test.

The first part showed that SHERPA and MADGRAPH+PYTHIA both describe the track properties well. However, SHERPA achieves more accurate results. On the other hand we could rule out HERWIG as a candidate already in this part of the analysis, since it is not suited to describe these properties well. Therefore, we spared to include HERWIG in the remaining part of this analysis.

In the second part it was apparent that SHERPA is the better fit for our purposes. While MADGRAPH+PYTHIA still achieves reasonable results, in particular the low-p_T regime is better described using SHERPA.

Also a tuning of the PYTHIA parameters to the p_T spectrum of the Z/γ boson was put to a test. It was not included, since it achieved essentially the same results that the default settings do.

We conclude from this study that since SHERPA well suited to describe tracks as well as the Z production processes. Therefore, we intend to use it for the rest of our study.

Using MadGraph+Pythia would also have been an adequate option, Herwig, on the other hand, is not at all be suited for this study.

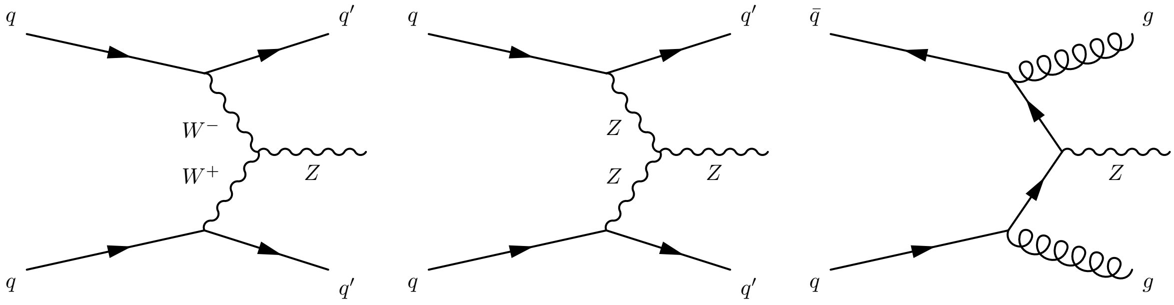

Figure 15: From left to right: Feynman graphs of the WBF Z signal with mediating W and Z boson and the QCD background process.

6 Using tracks to separate signal and background

In the following analysis we want to see if using track information for the separation of signal and background can achieve better results than using jets.

We will start by considering the Z production processes as before, to make general observations to see how far the track observables can distinguish between signal and background.

We will then continue by considering the Higgs to invisible process and also applying our findings to this process. As was shown in Section 5, Sherpa is best suited for our purposes. Therefore, we will use Sherpa for the generation of events throughout the following study.

6.1 Track variables in Z production

The signal process $Z+$jets WBF, is defined as the Z boson being produced via weak boson fusion and is expected to have less jet and track activity than $Z+$ jets QCD. Feynman graphs of both processes are shown in Fig. 15. Therefore, we expect the signal process to have fewer tracks than the background process.

6.1.1 Analysis

In contrast to before we consider the Z to be stable, i.e. we set the branching ratio of Z to invisible to be 100%. Furthermore, in the hard process we allow up to one additional jet. We impose the invariant mass of the tagging jets to be $m_{jj} > 500$ GeV and the difference in pseudorapidity to be $\Delta\eta_{jj} > 2.4$. Again we use a c.o.m. energy of $\sqrt{s} = 13$ TeV as in the current LHC run II.

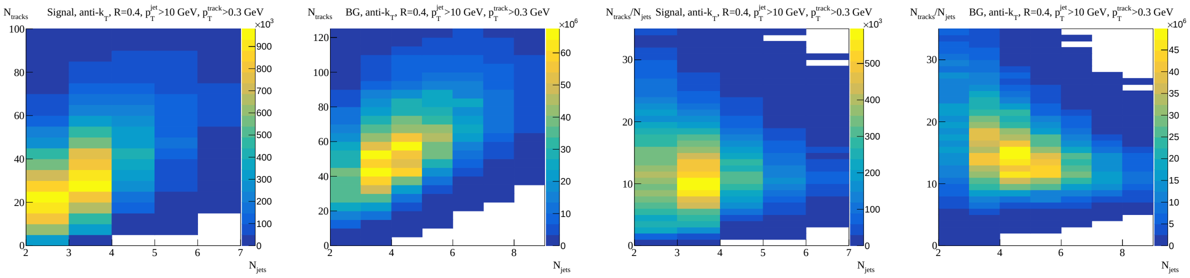

Figure 16: Plots of the number of tracks vs. the number of tracks for signal (left) and background (center left) and the number of tracks per jet vs the number of jets for signal (center right) and background (right).

We cluster the jets using the anti-k_T algorithm with $R = 0.4$ as it is default for ATLAS analyses. Furthermore, we use a minimum cut on the jets of $p_T^{\text{jets}} > 10$ GeV. We restrict our analysis on the central activity in the pseudorapidity gap spanned by the tagging jets.

Since, we're only interested in the activity in this η-gap we only consider the tracks and jets in this region. We impose a transverse momentum cut of $p_T^{\text{tracks}} > 300$ MeV on the tracks as in Ref. [44].

In Fig. 16 the two plots on the left show the number of tracks against the number of jets for signal and background and the two plots on the right show the same but with the number of tracks per jet. We can see that, for the signal, there are about 10 tracks per jet, while in the background there are about 15 tracks per jet. Furthermore, the signal tends to have 2-3 jets, while the background tends to have 3-4 jets in the η-gap. However, in contrast to the signal, the number of tracks per jet in the background sample decreases with an increasing number of jets.

This is a hint that we could use tracks to discriminate between signal and background. Let us now consider the distributions of these variables.

In Fig. 17, one can see the track multiplicities for all tracks as well as just the negatively and positively charged ones. As one can clearly see in all these distributions there are, as expected, more tracks in the background QCD process. Cutting at 50 tracks for all charged tracks, or at about 25 for the negatively and positively charged ones, seems to achieve the best separation between signal and background.

Differences in the transverse momentum and pseudorapidity could also allow for a separation between signal and background. Since we know that the background has more jets with, on average, more tracks than the signal, we expect the background to have harder jets and tracks.

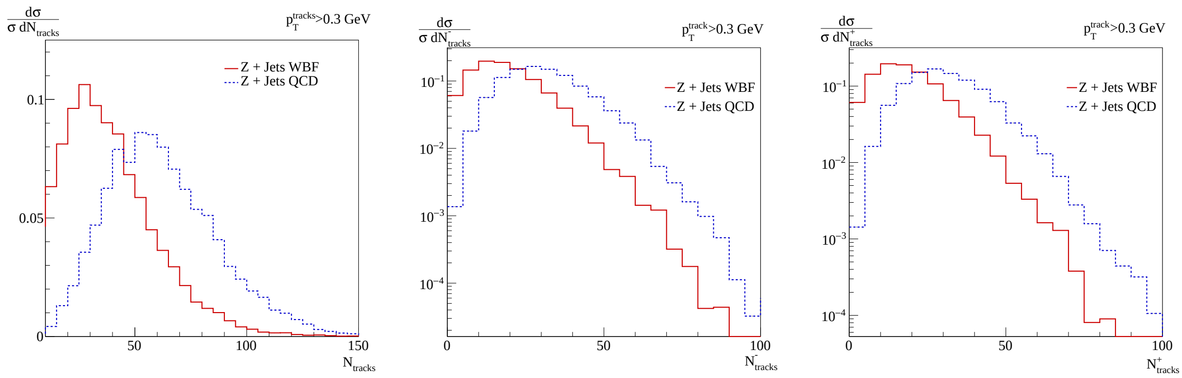

Figure 17: Track multiplicity distributions for all charged tracks (left), negatively charged tracks (center) and positively charged tracks (right).

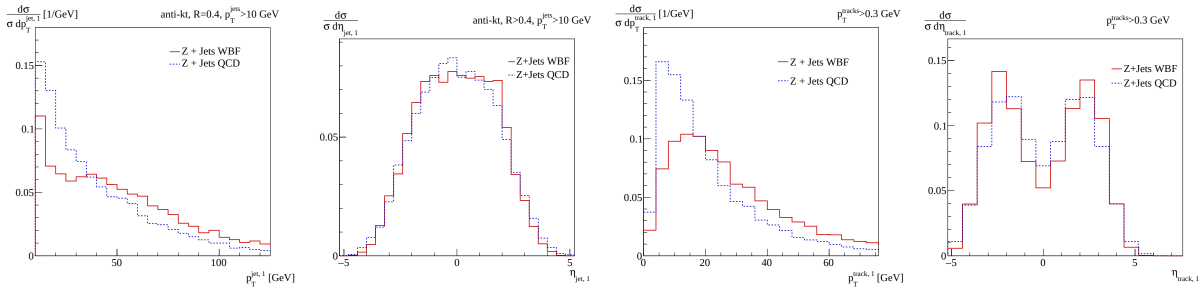

Figure 18: Distributions of the transverse momentum and pseudorapidity of the jets (left) and the tracks (right).

In Fig. 18 we plot the distributions of the transverse momentum and the pseudorapidity of the hardest track and the hardest jet in the η-gap. As expected the transverse momentum of the jets and the tracks is higher for the background process. The pseudorapidity distributions, however, are pretty similar. Therefore, it seems that as an additional variable to cut on we can use the the transverse momentum.

To examine if one can further use these variables to improve the significances we ran the analysis again, each time applying a veto on the transverse momentum of the hardest central jet and the number of tracks and jets, respectively. For N events we each time compute the number of events v that survive the veto. From this we calculate the fraction of the cross section, σ that corresponds to these events $\frac{v}{N}\sigma$. Now, we are able to get the respective significances using Eq. (3.6) with $\frac{v}{N}$ as efficiency.

In Fig. 19 the results of this analysis are shown. On the x-axis the veto on the transverse momentum of the hardest central jet is shown, while the y-axis shows the veto on

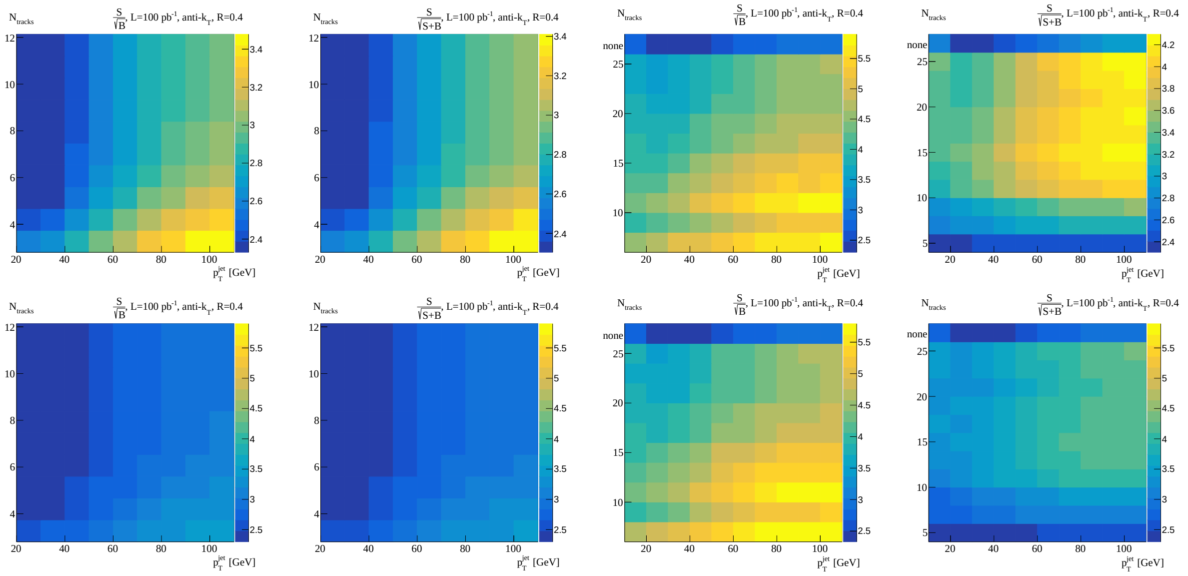

Figure 19: Upper row: Plots of the significances $\frac{S}{\sqrt{B}}$ and $\frac{S}{\sqrt{S+B}}$ for the jet number veto (left) and the track number veto (right) vs. the veto of the transverse momentum of the hardest central jet, respectively. Lower row: Same plots, however, with a fixed scale. The label "none" corresponds to the case, where no veto was imposed.

the number of jets and tracks, respectively. The z-axis indicated by the color shows the significance. The plots are shown, twice, once with an arbitrary scale for each plot and once fixed to allow for a comparison between the performance of vetoing the track number vs. vetoing the jet number.

All plots show that vetoing on the transverse momentum of the hardest central jet doesn't improve the performance at all. This was to be expected, since the signal jets are harder than the background jets for $p_T^{\text{jet, }1} > 40$ GeV as can be seen in Fig. 18. Placing a minimum transverse momentum cut of 40 GeV might achieve an improvement.

However, placing a cut on the number of jets and tracks improves of the sensitivity. When cutting on the number of jets it seems to be the best to force no jets in the gap, while for the tracks achieve high significances also for higher vetoes.

Comparing the significances with and without the signal in the denominator we see that for vetoes of 10 tracks and lower only few events survive and those are predominantly signal events. Both signal and background will be of comparable size and, therefore, $S \ll B$, which $S/\sqrt{B}$ assumes, is not valid anymore. Hence, including the signal in the denominator of the significances is the better choice.

Despite this, we still see that using the number tracks for vetoing events is a good way to improve the significance, as can also be seen in the $S/\sqrt{S+B}$ significance distribution, where no veto is significantly worse than applying a veto on the tracks. Also, the improvement achieved by using tracks surpasses the one using jets as one can see, when considering the plots on a fixed scale (lower row of 19).

Therefore, we conclude that while vetoing on the transverse momentum does not improve the significances, vetoing on the number of tracks and jets surely does. Furthermore, using tracks instead of jets allows for a better separation of signal and background and, hence, it is promising to use the track information for such analyses.

In this analysis we use only two variables to improve the significances. In the following section we want to analyse this more quantitatively and see how much improvement can be actually achieved using the track information.

6.1.2 Boosted Decision Tree analysis

In the last section we have seen that it is possible to improve the significance of the signal by applying vetoes on the events. However, this was only done by considering the jet transverse momentum and the number of tracks.

To investigate a possible improvement of the signal to background ratios we will in the following apply a boosted decision tree (BDT) analysis including further track variables. As before we will also do an analysis using only jet variables to see if the sensitivity of the track analysis is in fact higher than the one using only jets. Additionally, we do an analysis where both the track and jet variables are combined to see if the tracks behave as a complement or rather a substitute, that is, if using the variables combined can improve the performance even further.

For this analysis we generate events in the same manner as in the section before. Again, we only consider the activity in the pseudorapidity gap spanned by the tagging jets. For the jets we use the jet multiplicity as well as the p_T and η of the hardest jet in the η-gap. Since, in this sample, the third hardest jet is usually from the matrix element and the softer ones from the parton shower, we do not consider softer jets in the η-gap, because the parton shower simulates soft and collinear jets, while the matrix element jets are hard. The pseudorapidity and the transverse momentum of the jet are highly correlated and both of these variables are also expected to be correlated to the number of jets. This makes it even harder to improve the signal only using these variables. Therefore, one advantage of the track analysis is having variables that are less correlated.

In the track analysis we use the number of positively and negatively charged tracks,

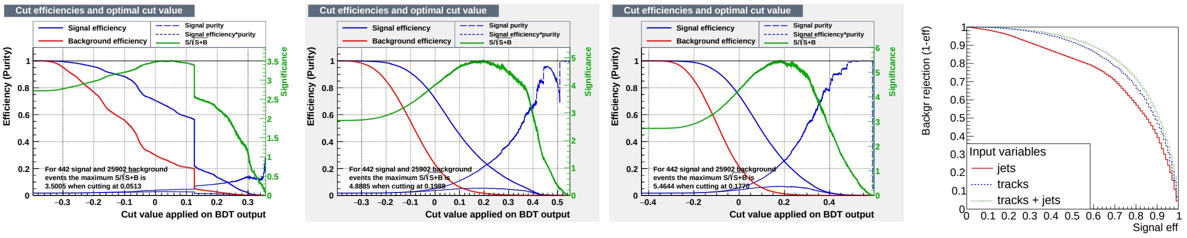

Figure 20: Plots of the signal efficiencies depending on the BDT cut variable for jets (left) and tracks (center-left) and both combined (center-right) and the ROC curves for the three analyses (right). For the jets the transverse momentum and pseudorapidity of the third hardest jet and the number of jets were used as input variables. Similarly we used the transverse momentum and pseudorapidity of the two hardest tracks and the number of positively and negatively charged tracks as well as the angular differences $\Delta\eta, \Delta\phi$ and ΔR of the two hardest tracks as input variables for the track analysis.

background acceptance	1%	10%	30%
jet analysis	0.067 (0.066)	0.332 (0.329)	0.706 (0.703)
track analysis	0.202 (0.209)	0.537 (0.535)	0.780 (0.779)
combined analysis	0.228 (0.237)	0.589 (0.585)	0.811 (0.809)

Table 1: Overtraining check: Signal efficiency from the test sample (from the training sample) for different background acceptances with Z WBF signal and Z QCD background.

respectively, the p_T and η of the two hardest tracks as well as the difference in η, ϕ and R between those two hardest tracks. While the number of positively and negatively charged tracks are expected to be highly correlated, as well as the angular differences η, ϕ and R we do not expect such a behavior for the other variables. Since the number of tracks is much higher than the number of jets, the properties of individual tracks will be far less correlated to the track multiplicities than it is the case for the jets.

The analysis is done using the TMVA [54] toolkit provided by ROOT [53].

In Fig. 20 the results of this analyses are shown. The three plots on the left show the significance as a function of the BDT cut variable for the three analyses of the jets, tracks and both combined. The plot on the right shows the ROC-curves of these analyses.

The discontinuous behavior in the curve of the jet analysis is due to the strong discontinuity of the input variables, since when there was no jet in the gap the transverse momentum and pseudorapidity got assigned a negative value.

In Table 1 we show the overtraining check. In this test the signal efficiency from the test sample is compared to the one from the training sample for different background acceptances. The higher the differences between those to efficiencies is, the higher the

analysis	$\frac{S}{\sqrt{S+B}}$	$\frac{S}{\sqrt{B}}$
jet analysis	3.502	3.58
track analysis	4.892	5.523
combined analysis	5.467	6.211

Table 2: Significances for the jet, track and combined analysis for the Z WBF signal and the Z QCD background with a luminosity of $\mathcal{L} = 100$ pb^{-1}.

overtraining is expected to be. As we can see in this analysis both efficiencies are small and can be explained by statistical fluctuations.

Furthermore, the analysis output tells us that the the variables of the hardest jet (p_T and η) have a correlation of about 91% to each other and of about 68% to the jet multiplicity for the signal and about 87% to each other and about 63% and 54%, respectively, to the jet multiplicity for the background. These high correlations were expected. Hence, the analysis considered them of a similar importance, however, the jet multiplicity was ranked the highest.

In the track analysis, the correlation between the positively and negatively charged tracks iss about 93% for signal and background. This high correlation was, also expected. Additionally, as one would expect the angular differences are correlated to each other with a correlation of about 70% for the signal. Lastly, the transverse momentum and pseudorapidities of the two hardest tracks were correlated to each other and the pseudorapidities also to the angular differences. However, altogether, the variables in the track analysis were less correlated than the ones in the jet analysis, as we expected. The track multiplicities were ranked to be of greatest importance, each by a similar factor. The remaining variables were ranked in a similarly, however, the angular differences (except for the difference in η) and the pseudorapidity of the tracks were ranked higher than the transverse momentum of the tracks.

From the significance curves in Fig. 20 we can read off the efficiencies of the signal and the background at the maximum. Using these we can compute the sensitivity possible in this analysis. The results of this calculation are shown in Table 2. The signal can be observed at a 5σ excess. For a luminosity of $\mathcal{L} = 100$ pb^{-1} this is the case in the track analysis. The luminosity is, for our purposes, just an arbitrary conversion factor and not LHC-relevant, since we do not include statistical and systematic uncertainties in the $S/\sqrt{B}$ limits. However, we will stick with a luminosity of $\mathcal{L} = 100$ pb^{-1} throughout this study, to compare these significances to the ones we will yield in the following Higgs invisible analysis.

This study, is aimed at motivating the analysis of track observables as a means to distinguish WBF processes from a QCD background. Therefore, we use these significances only for a technical comparison between tracks and jets. Table 2 clearly shows that using tracks for the separation of signal and background achieves better results than using only jets. Furthermore, using both jet and track variables for the analysis achieves the best result. Hence, we see that the track and jet variables serve as a complement. This can be already seen in the ROC curves of Fig. 20.

Therefore, we can conclude that it is in fact possible to improve the significance of the $Z+$jets WBF process by using tracks for the analysis. Furthermore, the results achieved using tracks alone are better than when using jets alone. Both combined yield the best results. Hence, it is recommended to use track variables in the separation of signal and background for these processes alongside traditional jet observables.

To see if this is also the case for the Higgs invisible decay, we will apply the same analysis for the Higgs invisible process in the following section.

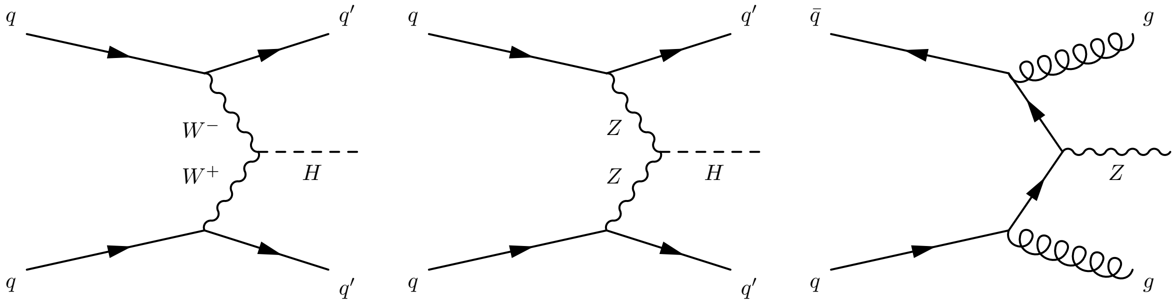

Figure 21: From left to right: Higgs WBF signal with mediating W and Z boson and the Z QCD background.

6.2 Track variables in Higgs invisible decays

The last section showed that it is possible to use the track information to separate signal from background in the $Z+$jets WBF against the $Z+$jets QCD process. In this section we want to investigate if the same can be done for the Higgs invisible process. This process is, very similar to the $Z+$jets WBF process. In Fig. 21 we show Feynman graphs for signal and background.

Since the Higgs and the Z boson are, both weakly interacting we expect a similar behavior to occur. Due to the higher mass of the Higgs the cross section of this process is smaller than for the Z. Therefore, we expect the significances to be lower due to this smaller cross section.

Furthermore, since the Higgs is heavier more energy will be required its production, i.e. we expect more radiation, due to a higher partonic c.o.m. energy.

The differences between these two processes will examined further in the last part of this analysis, where we use the $Z+$jets WBF process as background and try to separate it from the Higgs invisible process. But first we will investigate if it is possible to separate the Higgs invisible signal from the $Z+$jets QCD background. When trying to extract the Higgs invisible from the background this process is the dominating one and, therefore, we will neglect other possible backgrounds.

6.2.1 Analysis

We generate events in the Higgs invisible process, where we use the same settings as before and just replace the Z in the $Z+$jets WBF process with a Higgs boson. As for the Z we consider the Higgs to be stable and not decaying. This is equivalent to assuming a 100%

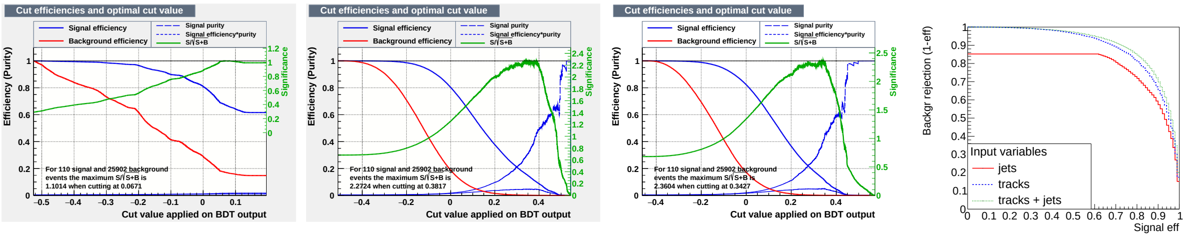

Figure 22: Plots of the signal efficiencies depending on the BDT cut variable for jets (left) and tracks (center-left) and both combined (center-right) and the ROC curves for the three analyses (right). For the jets the transverse momentum and pseudorapidity of the third hardest jet and the number of jets were used as input variables. Similarly we used the transverse momentum and pseudorapidity of the two hardest tracks and the number of positively and negatively charged tracks as well as the angular differences $\Delta\eta, \Delta\phi$ and ΔR of the two hardest tracks as input variables for the track analysis.

background acceptance	1%	10%	30%
jet analysis	-	-	0.819 (0.819)
track analysis	0.324 (0.327)	0.688 (0.691)	0.882 (0.885)
combined analysis	0.339 (0.344)	0.728 (0.732)	0.904 (0.905)

Table 3: Overtraining check: Signal efficiency from the test sample (from the training sample) for different background acceptances with Higgs WBF signal and Z QCD background.

branching ratio for Higgs to invisible decays. Since we want to find out, which fraction of the Higgs could be invisibly decaying this is the starting point in such an analysis.

As before, we use $\sqrt{s} = 13$ TeV and $m_{jj} > 500$ GeV and $\Delta\eta_{jj} > 2.4$ for the tagging jets. For the analysis of the events, we again use the anti-k_T algorithm for the jet clustering with $R = 0.4$ and $p_T^{\text{jets}} > 10$ GeV and use the tagging jets to span the η-gap and will in the following only consider the activity in this gap. The minimum transverse momentum cut on the tracks is once again $p_T^{\text{tracks}} > 300$ MeV.

We directly start with the BDT analysis. For the jet analysis we again use the transverse momentum and pseudorapidity of the hardest central jet as well as the jet multiplicity. In the track analysis once again the number of positively and negatively charged tracks, the transverse momentum and pseudorapidity of the two hardest tracks in the gap as well as the differences in η, ϕ and R between those tracks is used.

In Fig. 22, the result of these analyses are shown.

In Table 3 the overtraining check of these analyses is shown. Since, the jet analysis could at most achieve a background acceptance of about 15% the first two columns could not be filled. However, from the remaining results we can conclude that overtraining again

analysis	$\frac{S}{\sqrt{S+B}}$	$\frac{S}{\sqrt{B}}$
jet analysis	1.103	1.112
track analysis	2.274	3.026
combined analysis	2.363	3.08

Table 4: Significances for the jet, track and combined analysis for the Higgs signal and the Z QCD background with a luminosity of $\mathcal{L} = 100$ pb^{-1}.

seems to not have been an issue in these analyses.

The reason why the jet analysis is not able to reject more than 85% of the background might be that the H WBF process consists of fewer jets than the Z WBF process, which decreases the number of variables to improve on in many events from three to one. While we expected more activity in the H WBF process, we see in Fig. 23 that it is in fact such that the Z WBF gap activity is slightly higher.

The ROC curves clearly show that the combined analysis achieves the best results, followed closely by the track analysis.

Compared to the Z + jets process, the correlations in the jet analysis between the transverse momentum and the pseudorapidity of the hardest jet are even higher at about 97% and the correlation to the jet multiplicity is at about 83%. for the signal. The background is less correlated and the transverse momentum and pseudorapidity have a correlation of about 89% and are correlated to the jet multiplicity with 63% and 55%, respectively. However, this did not change anything to the ranking of the variables in the BDT analysis and, while again the jet multiplicity is ranked the highest all variables are ranked by a similar factor.

For the track analysis we observe a similar behavior. Compared to the $Z+$jets process the variables are correlated similarly, however, the percentages (especially in the signal) are altogether higher than before. Hence, also in this track analysis the track multiplicities are ranked the highest, followed by the angular differences and pseudorapidities and finally the transverse momenta.

In Table 4 we show the significances each of the analyses has achieved. As can be seen the track analysis achieves a significance that is more than double that of the jet analysis. The combined analysis, as the ROC curve already indicated, improves the significance compared to the track analysis only slightly. Furthermore, we see that when we neglect the signal in the denominator we get significantly better result in the track and combined analysis than it is the case for the jet analysis. This is because the track analysis allows for a high background rejection, as one can also see in the steep slope of the respective ROC

	$Z+\text{jets}$		$H+\text{jets}$	
analysis	$\frac{S}{\sqrt{S+B}}$	$\frac{S}{\sqrt{B}}$	$\frac{S}{\sqrt{S+B}}$	$\frac{S}{\sqrt{B}}$
jet analysis	3.502	3.58	3.622	3.721
track analysis	4.892	5.523	5.352	10.13
combined analysis	5.467	6.211	5.641	10.311

Table 5: Significances for the jet, track and combined analysis for the $Z+\text{jets}$ and the $H+\text{jets}$ process, where we used $\sigma_{Hjj}=\sigma_{Zjj}$ with a luminosity of $\mathcal{L}=100\ \text{pb}^{-1}$.

curves.

The luminosity of $\mathcal{L}=100\ \text{pb}^{-1}$ was used again to be able to compare these results to the ones from the $Z+\text{jets}$ process. However, to see which effects are due to the different cross sections of these processes and which one are technical we will in the following recompute these significances setting $\sigma_{Hjj}=\sigma_{Zjj}$.

The results of this calculation compared to the results obtained in the analysis of the $Z+\text{jets}$ process can be seen in Table 5. Even the jet analysis achieves better results in the $H+\text{jets}$ process. However, the greatest improvement compared to the former analysis is achieved by the track analysis. Since the cross section of the signal was increased for the $H+\text{jets}$ process, the results for $S\Big/\sqrt{B}$ are extremely well in the track and combined analysis. The jet analysis shows no such behaviour due to the lower background rejection.

Therefore, we can conclude that using the track information for the separation of signal and background works even better in the $H+\text{jets}$ process than in the $Z+\text{jets}$ process. The rejection of the background especially, is improved compared to before. To further investigate these differences in the two processes and to see if this might make it possible to differentiate both of them using the same techniques as before, we will, in the following section, compare the signal of both processes to each other.

6.2.2 Separating the $H+$jets WBF from the $Z+$jets WBF process

In the last section we showed that the sensitivity using track variables for the separation of signal and background is higher for the $H+\text{jets}$ process than for $Z+\text{jets}$.

In this section we want to investigate how this comes about and check if this might allow us to discriminate between the two weak boson fusion processes.

In Fig. 23 the multiplicities for the jets and the tracks as well as the transverse momentum and pseudorapidity distributions of the hardest jet and hardest track in the pseudorapidity gap are shown.

Although the track multiplicities are pretty similar the curve for the Higgs is shifted

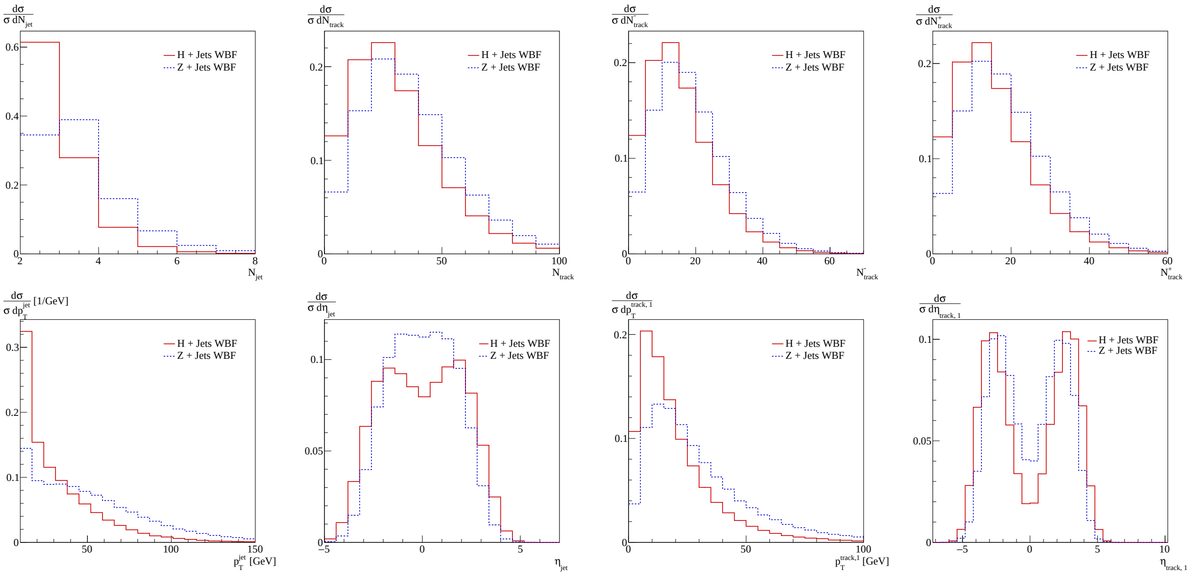

Figure 23: Upper row: Plots of the jet and track multiplicity (left) as well as the multiplicities of the positively and negatively charged tracks (right). Lower row: Plots of the transverse momentum and pseudorapidity of the hardest jet (left) and the hardest track (right) in the pseudorapidity gap.

slightly to the left, i.e. there are fewer tracks in the Higgs weak boson fusion process. This is also confirmed by the jet multiplicity, which clearly shows that there are fewer jets in the Higgs WBF process than in the Z WBF process. Furthermore, the transverse momentum distributions of the hardest jet and hardest track in the gap show that the activity in the Z WBF process tends to be harder than for the Higgs WBF process.

The pseudorapidity distributions show that the hardest jet is more central in the Z WBF process. The same holds for the pseudorapidity of the hardest track. This might improve the discrimination of signal and background using the pseudorapidity variable compared to before.

To see if these differences in the distribution allow for a discrimination between those two processes, we performed a BDT analysis using the Higgs WBF process as the signal and the Z WBF process as background. We examined the same variables as before.

The results of this BDT analyses are shown in Fig. 24. Table 6 demonstrates that these results are not suffering from overtraining. The differences between the test and training sample are small and can be explained by statistical fluctuations.

The ROC curves show again that the combined analysis works best, followed closely by the track analysis, while the jet analysis achieves similar results at high signal efficiencies,

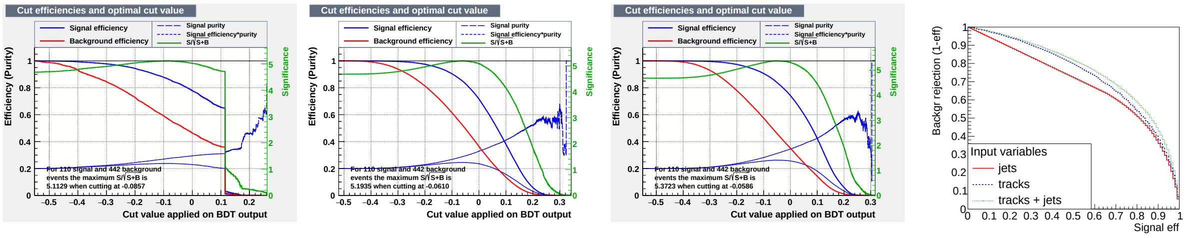

Figure 24: Plots of the signal efficiencies depending on the BDT cut variable for jets (left) and tracks (center-left) and both combined (center-right) and the ROC curves for the three analyses (right). For the jets the transverse momentum and pseudorapidity of the third hardest jet and the number of jets were used as input variables. Similarly we used the transverse momentum and pseudorapidity of the two hardest tracks and the number of positively and negatively charged tracks as well as the angular differences $\Delta\eta, \Delta\phi$ and ΔR of the two hardest tracks as input variables for the track analysis.

background acceptance	1%	10%	30%
jet analysis	0.020 (0.020)	0.182 (0.181)	0.542 (0.535)
track analysis	0.046 (0.049)	0.305 (0.305)	0.640 (0.636)
combined analysis	0.053 (0.053)	0.328 (0.328)	0.684 (0.679)

Table 6: Overtraining check: Signal efficiency from the test sample (from the training sample) for different background acceptances with Higgs WBF signal and Z WBF background.

however, turns linearly for higher background rejection rates.

In the jet analysis, the highest ranked variable is the pseudorapidity of the hardest jet followed by the transverse momentum of the hardest jet and the lowest ranked variable is the jet multiplicity. This behavior differs the results of the former analyses, however is to be expected, considering Fig. 23. As was outlined in the discussion of these plots, the jet multiplicities of the signal and background are very similar. However, in the pseudorapidity distribution of the jets for the background are more central than for the signal. Also the transverse momentum distributions, with the background jets being harder, is a good variable for a separation of signal and background. Hence, the variables were ranked in such a way in this analysis.

For the track analysis the variables were ranked in a similar way as before, putting the track multiplicities on top. However, the transverse momentum of the second hardest track was ranked on second place. As we can see in the distribution of the transverse momentum of the hardest tracks, the signal tracks are softer than the ones from the background. We also expect this to be the case for the second hardest track, probably even more so, and, therefore, this could explain the high ranking of this variable. For the

analysis	$\frac{S}{\sqrt{S+B}}$	$\frac{S}{\sqrt{B}}$
jet analysis	5.116	6.006
track analysis	5.196	6.161
combined analysis	5.376	6.458

Table 7: Significances for the jet, track and combined analysis for the Higgs WBF signal and the Z WBF background with a luminosity of $\mathcal{L} = 100$ pb^{-1}.

hardest track, however, the differences are not yet significant enough, which is why the transverse momentum of the hardest track is, as before, ranked in last place

To see how this changes the results, we once more computed the significances. These are shown in Table 7. As one can see, while the combined analysis achieves the best results followed by the track analysis, all analyses achieve similar results. Since the maximum sensitivity is for all analyses, at high signal efficiencies and high background acceptances the differences in the ROC curves, which only occur at lower signal efficiencies, do not play a role in this regime, resulting in similar significances. We see that the background rejection for the analysis with the Z WBF background is worse than in the ones with the Z QCD background. This is perfectly reasonable, since the signal and background are almost equivalent and only change in the type of boson being produced.

In this case the tracks do not improve the separation between signal and background significantly compared to the jets. However, the cut on the transverse momentum of the jets of $p_T^{\text{jets}} > 10$ GeV might be to small in an actual experiment. A higher transverse momentum cut on the jets might might affect the significance of the jet analysis and, therefore, in such a region the track analysis might have an even larger advantage.

To check this more quantitatively, we proceeded in the following way: We analyzed the events once again. This time we took the events that only had two jets or if they had an additional jet in the gap, we demanded this jet to have a transverse momentum of $p_T^{\text{jet, 3}} > p_T^{\text{cut}}$, where p_T^{cut} was varied from 10 to 100. Otherwise we removed the event. For each of these ten analyses we calculated the significance and plotted the significance in dependence of p_T^{cut}.

The results of this analysis can be seen in Fig. 25. The significance curves on the left clearly show that for an increasing cut on the transverse momentum in the gap the performance of the jet analysis steadily decreases. The track analysis, however, is also decreasing in performance up to $p_T^{\text{cut}} \sim 40(50)$ GeV for the $S/\sqrt{B}$ ($S/\sqrt{S+B}$) significance but afterwards it stabilizes and remains constant for higher cuts on the transverse momentum. This indicates that at high transverse momentum cuts the performance of the track

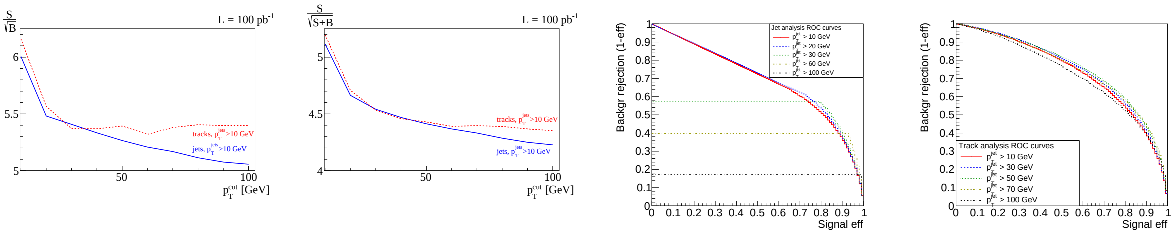

Figure 25: Plots of the significances $S/\sqrt{B}$ (left) and $S/\sqrt{S+B}$ (center left) in dependence of the jet cut variable p_T^{cut} and the corresponding ROC curves for the analysis at several cuts for the jets (center right) and the tracks (right).

analysis still achieves proper results, while the jet analysis struggles to do so. This can also be seen in the ROC curves, where for the jet analysis with an increasing cut the amount of background that cannot be rejected increases. For the track analysis, on the other hand, the signal acceptance and background rejection rates are similar for each cut.

This is due to the tracks being sensitive to the softer particles, which the jets after such cuts can no longer resolve. Therefore, especially in cases where a lot of soft activity is to be expected, the track analysis is a better fit for the separation of signal and background.

Increasing the minimum transverse momentum cut on all the jets, would show a similar behavior. But, in such a case events with central jets softer than the cut would be considered two jet events and included in the analysis. However, in our case such events would be removed. Still the track analysis would be the better choice, since for the jet analysis this event only contributes the jet multiplicity, while the track analysis is able to describe the softer activity and will have more variables.

Therefore, to simulate this behavior instead of dropping events, whose jets in the gap do not pass the cut on the transverse momentum, they could be considered as two jet events. Including this was beyond the scope of this study.

Finally, we see that for the Higgs to invisible process, using tracks for the separation of signal and background is also promising, even more than for $Z+$jets. Also, to investigate the behavior of soft activity, track observables are very useful and can separate the $H+$ jets WBF process from the $Z+$jets WBF process.

6.3 Conclusions

This main part of our study was aimed at determining whether using the track information of events can improve the separation of signal and background more than using jets alone. After showing, by vetoing on the number of tracks and jets and the transverse momentum of the hardest central jet that using tracks is promising for our purposes we proceeded with a boosted decision tree analysis done, altogether, in three ways:

First we separated the $Z+$ jets WBF from the $Z+$ jets QCD process. It could be seen that using the tracks in this process allows for a better separation of signal and background than the jets. Furthermore, a combination of jet and track variables leads to the best results.

The same was the case in our second analysis, where we considered the $H+$ jets WBF process as the signal and the $Z+$ jets QCD process as the background. Compared to the analysis of the Z WBF signal and the Z QCD background this achieved even better results for each analysis. Particularly, the track analysis allowed for a large background rejection. This can be explained by the observation that in the Higgs WBF process there is even less radiation than in the Z WBF case.

In the third part we tried to use the differences in the activity of the Higgs WBF process and the Z WBF process to differentiate them from each other. This is a lot harder than in the cases before, since these processes are topologically similar as the distributions of the jet and track variables have shown. The results we get indicate that using tracks or jets for the separation of signal and background achieves similar results. Despite the ROC curves showing that the track analysis performs better altogether, the maximum of the significance lies in the region where both analyses work equally well. However, as we have seen using a less generous cut on the minimum transverse momentum of the jets seems to worsen the performance of the jet analysis, while the track analysis remains rather constant. Therefore, for higher p_T using tracks could provide a more significant separation of signal and background.

The results we achieved are in this form not predictions for LHC searches, since we ignore uncertainties on the significances. However, the treatment of jets and tracks is consistent, which allows for a meaningful comparison of the different analyses.

It is possible to calculate the 95% CLs limit on the branching ratio of Higgs invisible decays. Then one can use this limit as an prediction of what this analysis might achieve when applied to experimental data. However, such an examination is beyond the scope of this study.

7 Conclusion and Outlook

In this thesis we compare Monte Carlo event generators to find the one best suited to describe track properties in the weak boson fusion process as well as the corresponding Z+jets QCD background. This was done comparing the results to data from ATLAS using two studies. One studying the jets created from tracks, as a measure for the performance in describing the track properties and the other studying the $Z+$jets process (WBF and QCD) with the Z decaying leptonically. The latter was a good measure to analyze the performance in generating the processes of interest. We compared the generators, Sherpa, MadGraph+Pythia and Herwig.

Sherpa was best suited for our purposes. MadGraph+Pythia achieves reasonable results and would have been also a good fit. Herwig, however, did not describe the data well.

The main purpose of this was to determine the best generator and not analyze how these differences come about. However, such a comparison study would be an interesting separate study. Finding the event generators that describe the respective data well would not only turn out to be useful in future analyses, but also improve our current understanding of the fundamental physics.

We proceeded analyzing track properties and checking if using tracks instead of jets for the separation of the Z and Higgs weak boson fusion signal from the Z QCD background could improve the significance.

For the $Z+$jets WBF signal with $Z+$jets QCD background the significance increased from 3.502 using jets to 4.892 using tracks. This shows that using tracks clearly achieves an improvement compared to using jets. The combined analysis gave a significance 5.467, i.e. both analyses complement each other.

The separation of the Higgs WBF signal from the $Z+$jets QCD background, confirms this observation. From a significance of 1.112 in the jet analysis using tracks improved the significance to 2.274. Compared to the former results, i.e. using the same cross sections, the significance improved from 3.622 for jets to 5.352.

The next step was to see if it is also possible to distinguish both weak boson fusion processes, $Z+$jets WBF as background and Higgs WBF as signal, from each other.

In this case the significance for the jets was 5.116 and for the track we got 5.196. This shows that both analyses allow a discrimination of signal and background. However, the tracks are only slightly more efficient than jets. The reason for this could be that the minimum transverse momentum cut of $p_T^{\text{jets}} > 10$ GeV is too generous. As was qualitatively shown, using a harder cut will decrease the significance for the jet analysis. This analysis also hinted that the track study might be unaltered by such a cut for and preferable to jets for p_T^{jets}.

It would be interesting to use these results to calculate the 95% CLs limits on the branching ratio of Higgs invisible decays. This would allow a prediction of results that can be achieved when this analysis is applied to experimental data. But such considerations were beyond the scope of this study.

Using the tracks instead of jets in the separation of the Higgs WBF signal and the $Z+$ jets QCD background improved the significance by a factor of 2, assuming a 100% Higgs to invisible branching ratio. While this is no prediction for analyses at the LHC it shows the potential that using tracks instead of jets would have.

Appendix

A Fermion Masses

In Section 2.2 we saw that, since the electroweak interactions only affect left-handed fields we cannot simply add a mass term to the Lagrangian.

However, this can be solved by obtaining the mass of the fermions via Yukawa couplings to the Higgs scalar field. In the following, we consider this for leptons and quarks.

For the leptons with the left-handed $SU(2)_L$ doublet ℓ and the right-handed singlet $\overline{e}$ we introduce the Yukawa coupling term

$$\mathcal{L}_{\text{Yukawa}} = -y\varepsilon^{ij}\varphi_i \ell_i \overline{e} + \text{h.c.} \ , \tag{A.1}$$

with the Yukawa coupling constant y and the Levi-Civita tensor ε. Using unitary gauge only the φ_1 component of the complex scalar field is nonzero and we obtain

$$\mathcal{L}_{\text{Yukawa}} = -\frac{1}{\sqrt{2}} y(v+H)(\ell_2 \overline{e} + \text{h.c.}) \ . \tag{A.2}$$

Using

$$\ell = \begin{pmatrix} \nu \\ e \end{pmatrix} \ . \tag{A.3}$$

we finally get

$$\mathcal{L}_{\text{Yukawa}} = -\frac{1}{\sqrt{2}} y(v+H)(e\overline{e} + \overline{e}^\dagger e^\dagger) = -\frac{1}{\sqrt{2}} y(v+H)\overline{E}E \ , \tag{A.4}$$

where we have introduced a Dirac field for the electron $E = (e \ , \overline{e}^\dagger)^T$. Now we can read off the mass term for the electron

$$m_e = \frac{yv}{\sqrt{2}} \ , \tag{A.5}$$

As we can see there is no mass term for the neutrinos, implying that they are massless. However, experiments have shown that neutrinos oscillate in flavor [7], which is, as can be proven mathematically, equivalent to neutrinos being massive.

For the quark sector, we have the left-handed quark doublet $q = (u, \ d)^T$ and the two

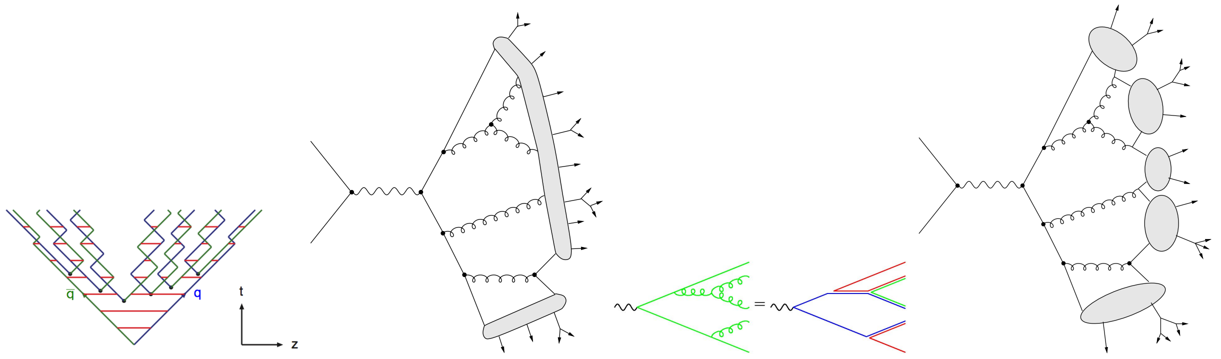

Figure 26: Left: Time evolution and fragmentation in the string model. Right: Color structure of a parton shower and fragmentation in the cluster model.Taken from [22, 42]

right handed doublets $\overline{u}$ and $\overline{d}$, where in each case the u corresponds to the up-type quarks and the d to the down-type quarks. Proceeding in the same way as for the leptons, however, with an additional coupling term that is missing in the lepton sector due to the absence of right-handed neutrinos, will give us the mass terms for the quarks. With the coupling constants y_1 for down-type quarks and y_2 for up-type quarks the quark mass terms m_d and m_u are given as

$$m_d = \frac{y_1 v}{\sqrt{2}} \, , \qquad m_u = \frac{y_2 v}{\sqrt{2}} \, . \tag{A.6}$$

B Hadronization

B.1 String Model

In the string model one uses the fact that in QCD at large distances the confinement is expected to be linear. That is, it can be described through an linearly rising potential $V(r) = \kappa r$ with the string constant $\kappa \approx 1$ GeV/fm. Additional to this there should be also an Coloumb term, which following the Lund model [43] we will neglect. We now describe these dynamics as massless, relativistic strings with no transversal degrees of freedom. If we now consider a $q\bar{q}$ pair moving apart along the z-axis the potential energy stored in the string will increase until it breaks, creating an quark-antiquark pair at the breaking vertex. This way at the end of the process the string will have broken into a set of new $q_i\bar{q}_i$ pairs with $i \in [1, \ n-1]$ for a fragmentation into n hadrons. This evolution of the string can be seen in Fig. 26 on the left.

To describe this behavior kinematically we can make use of the fact that

$$\left|\frac{\mathrm{d}E}{\mathrm{d}z}\right| = \left|\frac{\mathrm{d}p_z}{\mathrm{d}z}\right| = \kappa \ , \tag{B.1}$$

and, therefore, $E = \kappa \Delta z$. Taking into account that two adjacent breaks are constrained by the string piece having to be on the mass shell for the hadron being produced one can compute the total probability of an event as the product of $n-1$ breakup vertex probabilities and n delta functions for the hadron masses. However, this can be simplified using that, since the break vertices are causally disconnected they can be considered in arbitrary order ("left-right symmetry"). From this we can derive the fragmentation function $f(x)$ with x being the momentum fraction of remaining momentum that the produced hadrons receives.

$$f(x) \propto \frac{1}{x}(1-x)^a \exp\left(-\frac{bm_T}{x}\right) \ , \tag{B.2}$$

with free parameters a and b and the transverse mass $m_T = \sqrt{m^2 + p_T^2}$. In the derivation of this equation one also gets the probability distribution for the breakup vertices. With this at hand one can simulate the hadronization following the string model, however there are still some subtleties that have to be taken care of, which we will not discuss here. A schematic of the fragmentation in the string model is shown in Fig. 26. The advantage of the string model is that it offers a predictive framework of the space-time motion and the translation into an energy-momentum distribution of the hadrons. However, there are many parameters in this model, which have to be tuned by data. The string model is, for example, implemented in PYTHIA.

B.2 Clustering Model

In the cluster model the preconfinement property of parton showers is exploited. It states that the color structure of the shower is such that color singlet combinations of partons, i.e. clusters, can be built irrespective of the evolution scale Q_0 (see Fig. 26). The invariant mass distribution of these clusters is universal and asymptotically that is it depends only on Q_0 and Λ_{QCD} (universal) and we have $Q \gg Q_0$ (asymptotically). Furthermore, this distribution is suppressed for large masses.

In the cluster model one splits gluons non-perturbatively into quark-antiquark pairs at the cutoff scale Q_0. This way using the preconfinement property adjacent color lines will become quark-antiquark pairs and form clusters and will then decay into pairs of hadrons. A schematic of this fragmentation is shown in Fig. 26. This model suggests that due to the $g \to q\bar{q}$ being enhanced at low scales the running of the coupling should be reduced at these scales. This behavior is confirmed by hadronization studies suggesting a finite value of α_s at low scales (see e.g. Ref. [32]). Additionally this mechanism gives a possible solution to the high yield of soft photons in hadronic Z decays, since it predicts many charged particles created before hadronization.

Therefore, the cluster model seems to be a good description of the physics at low scales, at least qualitatively. Although the description of data the cluster model delivers is usually slightly less accurate than the string model its main advantage lies in having fewer parameters. HERWIG and SHERPA are using this model for the hadronization, albeit differing in their detailed description.

After having constructed the primary hadrons to finish the event generation one needs to let them decay into stable hadrons. By stable one usually means able to reach the detector. One might think that this can be simply done by implementing the decays using the information of the Review of Particle Physics by the Particle Data Group (PDG) [41], however, since the information in the PDG is often insufficient one also has to make theoretic assumptions to fill in the gaps. For this up to now sophisticated models have been developed, which will not be discussed here.

C AdaBoost

AdaBoost is a boosting algorithm that takes a weak learning algorithm and boosts it into a strong algorithm. This is done by changing the weights of the test sample adaptively, assigning higher weights to data points misclassified by previous classifiers. For a binary classification it proceeds in the following way:

1. Take a sequence of N labeled samples $[(x_1, y_1), \ldots, (x_N, y_N)]$, a distribution D over the N samples, a weak learning algorithm and the number T of iterations as input.

2. Initialize the weights for each event to be $w_i = D(i)$ with $i \in [1, N]$.

3. For each $t \in [1, T]$

   (a) Set $p^t = \frac{w^t}{\sum_{i=1}^{N} w_i^t}$

(b) Execute the weak learning algorithm with the events weighted according to p^t and get the hypothesis $h_t(\vec{x})$

(c) Calculate the error of this hypothesis: $\varepsilon_t = \sum_{i=1}^{N} p_i^t |h_t(x_i) - y_i|$.

(d) Set $\beta_t = \frac{\varepsilon_t}{1-\varepsilon_t}$ and set the weights for the next iteration to $w_i^{t+1} = w_i^t \beta_t^{1-|h_t(x_i)-y_i|}$

4. Return the output hypothesis $h_f(\vec{x}) = \begin{cases} 1 & \text{if } \sum_{t=1}^{T} (\log \frac{1}{\beta_t}) h_t(\vec{x}) \geq \frac{1}{2} \sum_{t=1}^{T} \log \frac{1}{\beta_t} \\ 0 & \text{otherwise} \end{cases}$

D Pythia vs. MadGraph+Pythia

The analyses in Section 5 use MadGraph instead of Pythia for the computation of the matrix elements. Since they are computed perturbatively and are not dependent on a specific model these computations should achieve the same results in each generator. The only differences in event generators are expected to be due to differences in modelling the parton shower and the hadronization. Therefore, we expect the methods of generating events with Pythia directly and generating them by using MadGraph for the matrix elements and Pythia for the parton shower and hadronization (denoted as MadGraph+Pythia) to achieve the same results.

To confirm this we generate the Drell-Yan process $q\bar{q} \to Z \to \ell^+\ell^-$ with $\ell = e,\ \mu$ for both tool chains and compare the results. We generate events with $\sqrt{s} = 13$ TeV and impose the invariant mass of the leptons to be $m_{\ell\ell} \in [80,\ 100]$ GeV to make sure that the propagator is a Z. Except for those the MadGraph default cuts as given in Eq. (5.3) have been used. We compare the kinematical distributions in Fig. 27.

As one can see, there are only minor differences in for Z of low p_T and small pseudorapidities, which are probably due to scale uncertainties as can be seen in the plots where the distributions are normalized to one. Still, there are minor differences, which are due to statistics. Clearly, both methods of event generation are in close agreement to each other. It is justified to interface MadGraph and Pythia.

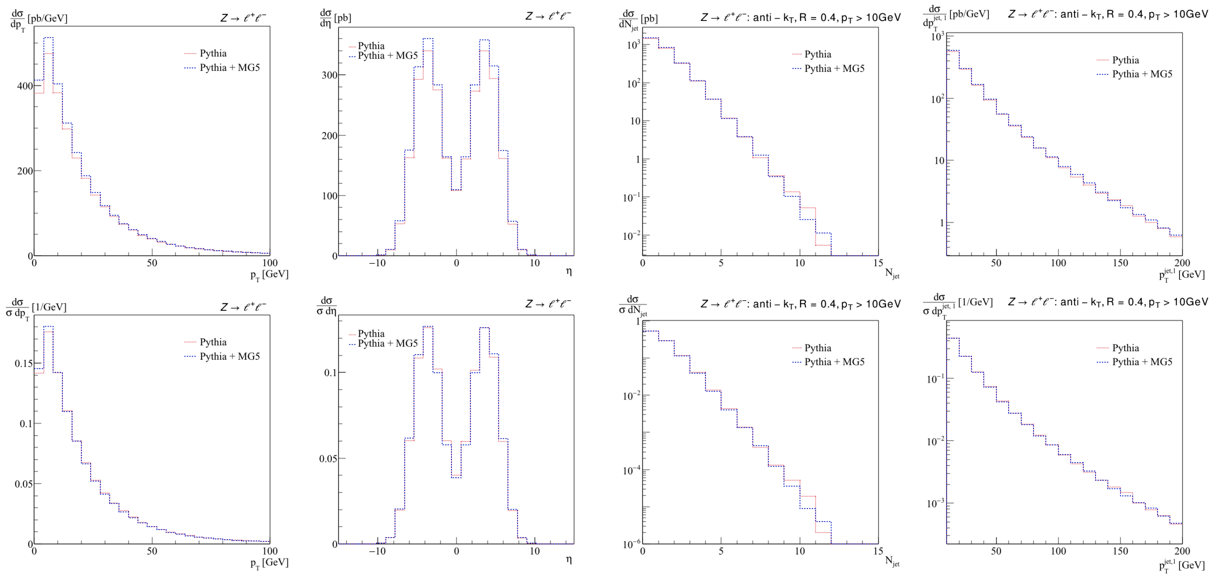

Figure 27: From left to right: Plots of the p_T and pseudorapidity of the Z and the number of jets and the p_T distribution of the hardest jet in the $Z \to \ell^+\ell^-$ process using Pythia directly and MadGraph+Pythia, where MadGraph is used for the matrix element computation and Pythia for the parton shower. The upper row shows the distributions normalized to the cross section, while the lower row shows them normalized to one.

References

[1] S. L. Glashow, "Partial Symmetries of Weak Interactions," Nucl. Phys. **22**, 579 (1961).

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

[3] A. Salam, "Weak and electromagnetic interactions," in Elementary particle theory: relativistic groups and analyticity, N. Svartholm, ed., p.367. Almqvist & Wiskell, 1968. Proceedings of the eighth Nobel symposium.

[4] P. W. Higgs, "Broken symmetries, massless particles and gauge fields," Phys. Lett. **12**, 132 (1964).

[5] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," Phys. Rev. Lett. **13**, 508 (1964).

[6] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," Phys. Rev. Lett. **13**, 321 (1964).

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

[8] Wikipedia Commons, `https://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg` (August 30, 2017)

[9] Mark Srednicki, *Quantum Field Theory*, Cambridge University Press, 1st edition, 2007.

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

[11] J. Beringer *et al.* [Particle Data Group], "Review of Particle Physics (RPP)," Phys. Rev. D **86**, 010001 (2012). doi:10.1103/PhysRevD.86.010001

[12] G. Aad *et al.* [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 (2013) [arXiv:1207.7214 [hep-ex]].

[13] S. Chatrchyan *et al.* [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 (2012) [arXiv:1207.7235 [hep-ex]].

[14] A. Pich, "Electroweak Symmetry Breaking and the Higgs Boson," Acta Phys. Polon. B **47**, 151 (2016) doi:10.5506/APhysPolB.47.151 [arXiv:1512.08749 [hep-ph]].

[15] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, “Experimental Test of Parity Conservation in Beta Decay,” Phys. Rev. **105**, 1413 (1957). doi:10.1103/PhysRev.105.1413

[16] T. Han, “Collider phenomenology: Basic knowledge and techniques,” hep-ph/0508097.

[17] L. Lista, “Practical Statistics for Particle Physicists,” doi:10.23730/CYRSP-2017-005.213 arXiv:1609.04150 [physics.data-an].

[18] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, “Better Jet Clustering Algorithms,” JHEP **9708**, 001 (1997); M. Wobisch and T. Wengler, “Hadronization corrections to jet cross sections in deep-inelastic scattering,” arXiv:hep-ph/9907280.

[19] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, “Longitudinally invariant K_t clustering algorithms for hadron hadron collisions,” Nucl. Phys. B **406**, 187 (1993).

[20] M. Cacciari, G. P. Salam and G. Soyez, “The Anti-k(t) jet clustering algorithm,” JHEP **0804**, 063 (2008) [arXiv:0802.1189 [hep-ph]].

[21] G. Aad *et al.* [ATLAS Collaboration], “Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector,” Phys. Rev. D **86**, 014022 (2012) doi:10.1103/PhysRevD.86.014022 [arXiv:1112.6297 [hep-ex]].

[22] A. Buckley *et al.*, “General-purpose event generators for LHC physics,” Phys. Rept. **504**, 145 (2011) doi:10.1016/j.physrep.2011.03.005 [arXiv:1101.2599 [hep-ph]].

[23] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert and J. Winter, “Event generation with SHERPA 1.1,” JHEP **0902**, 007 (2009) doi:10.1088/1126-6708/2009/02/007 [arXiv:0811.4622 [hep-ph]].

[24] F. Krauss, R. Kuhn and G. Soff, “AMEGIC++ 1.0: A Matrix element generator in C++,” JHEP **0202**, 044 (2002) doi:10.1088/1126-6708/2002/02/044 [hep-ph/0109036].

[25] T. Gleisberg and S. Hoeche, “Comix, a new matrix element generator,” JHEP **0812**, 039 (2008) doi:10.1088/1126-6708/2008/12/039 [arXiv:0808.3674 [hep-ph]].

[26] T. Plehn, “LHC Phenomenology for Physics Hunters,” doi:10.1142/9789812838360_0003 arXiv:0810.2281 [hep-ph].

[27] M. R. Whalley, D. Bourilkov and R. C. Group, “The Les Houches accord PDFs (LHAPDF) and LHAGLUE,” hep-ph/0508110.

[28] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, “QCD matrix elements + parton showers,” JHEP **0111**, 063 (2001) doi:10.1088/1126-6708/2001/11/063 [hep-ph/0109231].

[29] M. L. Mangano, M. Moretti, F. Piccinini and M. Treccani, “Matching matrix elements and shower evolution for top-quark production in hadronic collisions,” JHEP **0701**, 013 (2007) doi:10.1088/1126-6708/2007/01/013 [hep-ph/0611129].

[30] S. Frixione and B. R. Webber, “The MC@NLO 3.3 Event Generator,” hep-ph/0612272.

[31] P. Nason, “A New method for combining NLO QCD with shower Monte Carlo algorithms,” JHEP **0411**, 040 (2004) doi:10.1088/1126-6708/2004/11/040 [hep-ph/0409146].

[32] Y. L. Dokshitzer and B. R. Webber, “Calculation of power corrections to hadronic event shapes,” Phys. Lett. B **352**, 451 (1995) doi:10.1016/0370-2693(95)00548-Y [hep-ph/9504219].

[33] T. Ohl, “Drawing Feynman diagrams with Latex and Metafont,” Comput. Phys. Commun. **90**, 340 (1995) [hep-ph/9505351].

[34] T. Plehn, “Lectures on LHC Physics,” Lect. Notes Phys. **844**, 1 (2012) doi:10.1007/978-3-642-24040-9 [arXiv:0910.4182 [hep-ph]].

[35] T. Plehn, “Yet Another Introduction to Dark Matter,” arXiv:1705.01987 [hep-ph].

[36] J. McDonald, “Gauge singlet scalars as cold dark matter,” Phys. Rev. D **50**, 3637 (1994) doi:10.1103/PhysRevD.50.3637 [hep-ph/0702143 [HEP-PH]].

[37] N. Aghanim *et al.* [Planck Collaboration], “Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters,” Astron. Astrophys. **594**, A11 (2016) doi:10.1051/0004-6361/201526926 [arXiv:1507.02704 [astro-ph.CO]].

[38] C. Bernaciak, T. Plehn, P. Schichtel and J. Tattersall, “Spying an invisible Higgs boson,” Phys. Rev. D **91**, 035024 (2015) doi:10.1103/PhysRevD.91.035024 [arXiv:1411.7699 [hep-ph]].

[39] A. Biekötter, F. Keilbach, R. Moutafis, T. Plehn and J. Thompson, “Tagging Jets in Invisible Higgs Searches,” arXiv:1712.03973 [hep-ph].

[40] CMS Collaboration [CMS Collaboration], “Search for invisible decays of a Higgs boson produced via vector boson fusion at $\sqrt{s} = 13$ TeV.,” CMS-PAS-HIG-16-009.

[41] G. Aad *et al.* [ATLAS Collaboration], “Search for invisible decays of a Higgs boson using vector-boson fusion in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector,” JHEP **1601**, 172 (2016) doi:10.1007/JHEP01(2016)172 [arXiv:1508.07869 [hep-ex]].

[42] C. Patrignani *et al.* [Particle Data Group], “Review of Particle Physics,” Chin. Phys. C **40**, no. 10, 100001 (2016). doi:10.1088/1674-1137/40/10/100001

[43] B. R. Webber, “Fragmentation and hadronization,” Int. J. Mod. Phys. A **15S1**, 577 (2000) [eConf C **990809**, 577 (2000)] doi:10.1142/S0217751X00005334 [hep-ph/9912292].

[44] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, “Parton Fragmentation and String Dynamics,” Phys. Rept. **97**, 31 (1983). doi:10.1016/0370-1573(83)90080-7

[45] G. Aad *et al.* [ATLAS Collaboration], “Properties of jets measured from tracks in proton-proton collisions at center-of-mass energy $\sqrt{s} = 7$ TeV with the ATLAS detector,” Phys. Rev. D **84**, 054001 (2011) doi:10.1103/PhysRevD.84.054001 [arXiv:1107.3311 [hep-ex]].

[46] M. Aaboud *et al.* [ATLAS Collaboration], “Measurements of the production cross section of a Z boson in association with jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” Eur. Phys. J. C **77**, no. 6, 361 (2017) doi:10.1140/epjc/s10052-017-4900-z [arXiv:1702.05725 [hep-ex]].

[47] T. Ohl, “Drawing Feynman diagrams with Latex and Metafont,” Comput. Phys. Commun. **90**, 340 (1995) doi:10.1016/0010-4655(95)90137-S [hep-ph/9505351].

[48] T. Sjöstrand *et al.*, “An Introduction to PYTHIA 8.2,” Comput. Phys. Commun. **191**, 159 (2015) doi:10.1016/j.cpc.2015.01.024 [arXiv:1410.3012 [hep-ph]].

[49] M. Bahr *et al.*, “Herwig++ Physics and Manual,” Eur. Phys. J. C **58**, 639 (2008) doi:10.1140/epjc/s10052-008-0798-9 [arXiv:0803.0883 [hep-ph]].

[50] J. Bellm *et al.*, “Herwig 7.0/Herwig++ 3.0 release note,” Eur. Phys. J. C **76**, no. 4, 196 (2016) doi:10.1140/epjc/s10052-016-4018-8 [arXiv:1512.01178 [hep-ph]].

[51] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, “MadGraph 5 : Going Beyond,” JHEP **1106**, 128 (2011) doi:10.1007/JHEP06(2011)128 [arXiv:1106.0522 [hep-ph]].

[52] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz and F. Siegert, “Rivet user manual,” Comput. Phys. Commun. **184**, 2803 (2013) doi:10.1016/j.cpc.2013.05.021 [arXiv:1003.0694 [hep-ph]].

[53] M. Dobbs and J. B. Hansen, “The HepMC C++ Monte Carlo event record for High Energy Physics,” Comput. Phys. Commun. **134**, 41 (2001). doi:10.1016/S0010-4655(00)00189-2

[54] I. Antcheva *et al.*, “ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization,” Comput. Phys. Commun. **182**, 1384 (2011). doi:10.1016/j.cpc.2011.02.008

[55] A. Hocker *et al.*, “TMVA - Toolkit for Multivariate Data Analysis,” PoS ACAT , 040 (2007) [physics/0703039 [PHYSICS]].

[56] Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", JCSS 55, 119 (1997).

Acknowledgements

First of all, I would like to thank my supervisor Tilman Plehn for giving me the opportunity to work on this project and hand out helpful advise whenever necessary. Furthermore, I have to thank Jennifer Thompson and Johann Brehmer for guiding me through this project and being a great source of help with any kind of problem. The readability of this thesis was ensured through the kind help of Anke Biekötter and Peter Reimitz. In addition I want to thank the whole group for always being fun to be around and many delicious cakes. Last but definitely not least I thank my family and friends for supporting me, wherever needed, throughout my studies.

Erklärung

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 30.01.2018,

Josua Göcking