KIT

Karlsruher Institut für Technologie

Constraints on anomalous couplings in the hadronic decay channel of Vector Boson Scattering at $\sqrt{s} = 13$ TeV with the CMS Experiment

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN (Dr. rer. nat.)

von der KIT-Fakultät für Physik des

Karlsruher Instituts für Technologie (KIT)

angenommene

DISSERTATION

von

M.Sc. Max Neukum
aus Bad Friedrichshall

Tag der mündlichen Prüfung:	28.07.2023
Referent:	Prof. Dr. Thomas Müller
Korreferent:	Prof. Dr. Markus Klute

Introduction

Curiosity and the urge to understand our surroundings is a key factor of humanity's development. This includes physics, the natural science involving studies of matter with its fundamental constituents and their interaction. The physical world we live in comes in a myriad of facets with the rules at smallest accessible length scales currently best described by the Standard Model (SM) of particle physics [1–3].

It is the culmination of foregoing developments from the idea of atoms being indivisible building blocks of matter, to a multitude of subatomic particles, and finally to the fundamental fermions and bosons described in the SM. This requires to incorporate both quantum mechanics, which revolutionized how particles are perceived, and Einstein's theory of special relativity, which changed the understsanding of space and time. When it was first formulated in 1967, not all included particles had been experimentally confirmed but their observation followed in the upcoming years. The Higgs boson, the last missing piece of the SM to be experimentally observed, was discovered in 2012 [4, 5] at the Large Hadron Collider (LHC) [6] at CERN.

The LHC and accompanying detector experiments [7–10] are dedicated to the investigation of the SM, possible shortcomings, and extensions. They are an international endeavor and the results of great efforts, both financial and personal. The first data taking period lead to the discovery of the Higgs boson, followed by precise measurements of its properties. At the same time, direct searches for extensions explaining phenomena of physics beyond the SM (BSM) are carried out. Looking ahead, the research program at the LHC has not yet reached its midway point and results of direct searches have been found to be compatible with the SM. This motivates the interpretation of the SM as an effective field theory (EFT) [11–14], enabling to investigate the impact of BSM physics present at energies currently not reachable in a model independent way.

In an EFT approach, the theoretical framework of the SM is extended by adding further terms to its Lagrangian, the function which encodes its content. Comparison to measured data allows to draw limits on coefficients of additionally introduced terms depending on an energy scale Λ.

The analysis described in this thesis sets limits on 23 different coefficients additionally introduced in the EFT approach by investigating the hadronic decay channel of vector boson scattering (VBS) in data taken by the CMS experiment corresponding to an integrated luminosity of $137\,\mathrm{fb}^{-1}$. For a subset of these coefficients, no public results are available from VBS and it is the first time limits are presented, while for the others, the sensitivity is of comparable order as existing limits [15–17] but utilizes a different subset of the measured data. Therefore, a combination is possible to further improve the derived limits.

Besides direct searches, the EFT approach is one of the possibilities to extract hints to BSM physics and can be used to steer the development of particle physics in a productive direction. Since it appears in many signal topologies simultaneously, a combination is a natural further step. This analysis sets limits on a so far unique set of additional contributions from the EFT approach with very competitive limits and is thus a possible cornerstone for the legacy results of the LHC.

This thesis is organized as follows:
In chapter 1, the theoretical foundation is described, where the focus lies on VBS processes and EFT contributions. This is followed by an overview of the experimental setup in chapter 2, where the LHC and then the CMS experiment are briefly described. Chapter 3 is about the comparison of simulated events to measured data. First, individual steps from the theoretical description to the simulation of measured signals are laid out, and then, reconstruction algorithms leading to objects entering the statistical analysis are introduced. The comparison of data to simulation is then wrapped up by statistical inference in the last part of this chapter. Following the introductory chapters, the analysis is described in chapter 4 starting with a general composition of signal and background processes and their modeling. Uncertainty treatment and statistical inference then lead to the results described in the last part of the chapter. Finally, chapter 5 summarizes the results and brings them into context with other public results.

Contents

1. Theoretical Introduction

Experimental particle physics, as any experimental scientific effort, is connected to predictions of a theoretical model to interpret and guide complex measurements. In modern particle physics, this theory is given by the Standard Model (SM) of particle physics providing descriptions of fundamental particles and their interactions with one another.

It was developed in the 1960s and 1970s, before which the three forces in the SM were explained separately by quantum electrodynamics (QED) [18], Fermi's four-fermion theory of weak interactions [19, 20], and theories of the strong interaction such as the original Yukawa theory between nucleons [21]. While QED was very successful, the four-fermion theory explains beta decay by direct coupling of a neutron, proton, electron, and a neutrino but is not valid anymore at higher energies. Furthermore, an increasing number of strongly interacting particles was found and following different organizational schemes, the modern model of quarks as fundamental particles was developed [22]. Additional work on the underlying symmetries [23–25] facilitated the unification of QED and the theory of weak interactions accompanied by the introduction of the Higgs boson [25]. These developments solved the problems of validity at higher energies, reduced the number of fundamental particles and finally culminated in the SM. It was first formulated in 1967 [1–3] and was shown to be renormalizable in 1971 [26] after which it began to gain recognition. A more detailed historical review of the history of the SM can be found in Refs. [27, 28].

The SM had many successes and correctly predicted fundamental particles including the top quark and the tau neutrino, which were discovered in 1995 [29, 30] and 2000 [31], respectively. The last missing piece of the SM, the Higgs boson, was postulated in 1964 [25, 32, 33] and observed by the CMS and ATLAS Collaborations in 2012 [4, 5] at the Large Hadron Collider (LHC). Despite its remarkable precision, it is widely believed that the SM might be a low energy approximation of a more fundamental theory that solves the remaining open questions.

The first section of this chapter gives an overview of the SM content followed by an outline of its mathematical structure. Then vector boson scattering (VBS), the process investigated in this thesis, is described in section 1.2. In the last part of this chapter, section 1.3, effective field theory (EFT) as a model-independent approach to describe physics beyond the SM (BSM) is introduced with its application in the VBS topology.

1.1. The Standard Model of Particle Physics

The SM is a quantum field theory including both principles of quantum mechanics and special relativity and describes the fundamental building blocks of nature: quarks and leptons. It includes three of the four fundamental interactions, the electromagnetic, weak, and the strong force, while gravity is separately described by the Theory of General Relativity [34]. On length scales that are typically under consideration in particle physics, gravitation is much weaker than the other interactions and not considered in the SM. The mathematical formulation of the SM is based on the principle of least action, meaning that the variation of the action S vanishes:

$$\delta S = \int \mathcal{L}(\phi, \partial_\mu \phi) \, \mathrm{d}^4 \mathrm{x} = 0 \,, \tag{1.1}$$

where $\partial_\mu = \left(\partial_t, \vec{\nabla} \right)$ is the derivative operator in spacetime, $\mathcal{L}$ the Lagrangian density, which is taken as a starting point in many theories, and ϕ is a placeholder for the fields described by the SM. From this principle, Noether's Theorem [35] can be derived, which states that for a continuous symmetry, there exists a conserved quantity. The concept of symmetries plays a central role in the SM.

In order to ensure that the SM follows the ideas of special relativity, $\mathcal{L}_{SM}$ is invariant under transformations of the Poincaré group, which includes translation in time and space, rotations, and Lorentz boosts. This results in conservation of momentum, energy, and angular momentum. The interaction of the SM are included via additional local gauge symmetries: $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$, where the $\mathrm{SU}(2)_L$ symmetry only affects left-handed particles. This results in the gauge bosons of the SM and corresponding charges. If a particle carries such a charge, it is subject to the corresponding interaction. The local gauge symmetries and interactions are described in more detail in sections 1.1.2 through 1.1.5.

1.1.1. Particle Content

The particle content of the SM consists of six quarks, three charged leptons and three corresponding neutrinos, as well as four gauge bosons and the Higgs boson. The quarks, charged leptons and neutrinos are fermions, and thus carry spin-$1/2$ and follow Fermi-Dirac statistics. The bosons carry integral spin and follow Bose-Einstein statistics.

Fermions

The fermions of the SM are listed in table 1.1 and are grouped into three generations based on their electric charge and weak isospin. Each generation contains one up-type and one down-type quark, as well as a charged lepton and the corresponding neutrino. The first generation contains the up (u) and down (d) quark, the electron (e) and

Table 1.1.: The fermions of the SM grouped into three generations together with their respective electric charge, weak isospin in the case of left-handed fermions, and their possible color charges for quarks. Right-handed fermions have weak isospin of zero. For each fermion, a corresponding antiparticle exists with the same mass and spin but opposite charges, which are not listed here for simplicity. In particular, there are no left-handed antineutrinos in the SM and antiquarks carry anticolor: $\bar{r}, \bar{g}, \bar{b}$.

	Generation			Electric	Weak	Color
	1	2	3	charge	isospin (l.h.)	charge
Quarks	up (u)	charm (c)	top (t)	$+2/3$	$+1/2$	r,g,b
	down (d)	strange (s)	bottom (b)	$-1/3$	$-1/2$	r,g,b
Leptons	ν_e	ν_μ	ν_τ	0	$+1/2$	-
	electron (e)	muon (μ)	tau (τ)	-1	$-1/2$	-

electron neutrino (ν_e). The second generation consists of the charm (c) and strange (s) quark, the muon (μ) and muon neutrino (ν_μ), while the third generation includes the top (t) and bottom (b) quark, as well as the tau (τ) and tau neutrino (ν_τ). Electric charge and isospin of the second and third generation are analogous to the first generation, whereas the mass increases with the number of generation. In this sense, particles of the second and third generation can be thought of as heavy copies of the first generation particles. Furthermore, particles of the second and third generation are not stable and decay to lighter particles, such that neutrons and protons and thus atoms and matter in everyday life consists of particles of the first generation.

The interactions of all particles is determined by their charges: quarks and charged leptons are electrically charged and interact via the electromagnetic force mediated by photons. Additionally, quarks carry one color charge, red (r), green (g), or blue (b), and interact with gluons via the strong interactions. Lastly, the weak interaction is determined by the third component of weak isospin, T_3, and only affects left-handed fermions, while right-handed fermions have weak isospin of zero and do not undergo charged weak interactions. The concept of handedness, or chirality, is a consequence of the Dirac equation, which governs the kinematics of free fermions:

$$\left(i\gamma^\mu \partial_\mu - m_f \right) \psi_f = 0, \tag{1.2}$$

with the field ψ_f of fermion f with mass m_f, and the gamma matrices γ^μ. The solutions of this equation are four-component spinors, which can be represented in a basis directly showing its chirality. For massless particles, chirality is the same as helicity, i.e., the sign of the projection of its spin onto its momentum. Left-handed fermions form isospin doublets with the third component of weak isospin $T_3 = \pm\frac{1}{2}$:

$$\begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \quad \text{and} \quad \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L. \tag{1.3}$$

Right-handed fermions have weak isospin of zero, $T_3 = 0$, form singlets and do not undergo charged weak interactions:

$$\left(u \right)_R, \left(d \right)_R, \left(c \right)_R, \left(s \right)_R, \left(t \right)_R, \left(b \right)_R \quad \text{and} \quad \left(e \right)_R, \left(\mu \right)_R, \left(\tau \right)_R. \tag{1.4}$$

The SM does not contain right-handed neutrinos, since they are massless, colorless, have no electric charge, weak isospin of zero and would not interact. Conversely, left-handed neutrinos only interact via the weak interaction and are very hard to detect. Although neutrinos are assumed to be massless in the SM, experimental results show that neutrinos can change their flavor, i.e., move from one generation to another. So-called neutrino oscillations have been experimentally observed and can only be explained with massive neutrinos. This is an active field of research and so far only upper limits on the neutrino masses of $m_\nu < 0.8\,\text{eV}$ are known [36].

The mass terms for the charged leptons and quarks cannot simply be included in the SM Lagrangian, but they are generated by an additional Yukawa term, which is explained in section 1.1.4 in more detail. Fermion masses are free parameters of the SM and range from $m_u = 2.16^{+0.49}_{-0.26}\,\text{MeV}$ for up quarks to $m_t = 172.69 \pm 0.30\,\text{GeV}$ for top quarks. The masses of charged leptons are given by $m_e = 0.511\,\text{MeV}$ for electrons, $m_\mu = 105.66\,\text{MeV}$ for muons, and $m_\tau = 1.78\,\text{GeV}$ for taus. Up-to-date values of these measured quantities are taken from Ref. [37].

Lastly, there are antiparticles for each fermion, i.e., particles with opposite charge and chirality but same mass and spin. In particular, the SM contains right-handed antineutrinos but no left-handed ones, and antiquarks carry anticolor: $\bar{r}, \bar{g}, \bar{b}$.

Bosons

The four gauge bosons in the SM are spin-1 particles, their existence follows from the local gauge symmetries of the SM Lagrangian and are the mediator particles of the interactions. They are listed in table 1.2 together with the Higgs boson.

The electromagnetic interaction is mediated by the photon and couples to electric charge. The photon itself is massless and carries neither electric nor color charge meaning that there is no interaction between two photons. $W^\pm$ and Z bosons are the gauge bosons of the weak interaction and couple to particles with weak isospin, i.e., fermions. They are the only massive gauge bosons in the SM and interact directly with the Higgs boson and in case of the $W^\pm$ bosons with photons. The weak interaction is unique in the sense that it affects neutrinos, which are electrically neutral and colorless. In the SM, the electromagnetic and weak interaction are included together as electroweak

Table 1.2.: The bosons described in the SM with their corresponding charge, masses and in the case of the gauge bosons, their interaction and underlying symmetry. The electromagnetic and weak forces are based on spontaneously broken $SU(2)_L \times U(1)_Y$ symmetry as described later in section 1.1.2, whereas the strong interaction follow from $SU(3)_C$ symmetry of the Lagrangian. The masses are taken from Ref. [37].

Boson	Force	Coupling to	Mass	Symmetry
photon (γ)	electromagnetic	electric charge	0	
Z			91.188 GeV	$SU(2)_L \times U(1)_Y$
	weak	weak isospin		
$W^\pm$			80.379 GeV	
8 gluons (g)	strong	color charge	0	$SU(3)_C$
Higgs (H)		mass	125.25 GeV	

interaction from $SU(2)_L \times U(1)_Y$ symmetry.

The mediator boson of the strong interaction is the gluon with the corresponding charge being the color charge. Gluons are massless, electrically neutral and do not interact weakly. However, they themselves carry one color and one anticolor. This results in eight gluons from all possible combinations of r,g,b and $\bar{r}, \bar{g}, \bar{b}$, where one of the nine possibilities is a linear combination of the other eight. Theoretically, the strong interaction is included in the SM via the $SU(3)_C$ symmetry, where the eight gluons directly correspond to the eight generators of the group. Since gluons carry color charge, they interact with each other.

As the only scalar particle in the SM, the Higgs boson takes a special role. It couples to mass and thus interacts with each particle of the SM except for photons, gluons, and presumably neutrinos. The mass of bosons is given by the Higgs mechanism through spontaneous symmetry breaking resulting in the Higgs boson.

1.1.2. Electroweak Interaction

In the SM, the weak and electromagnetic interaction are combined in the electroweak theory [1–3] including quantum electrodynamics (QED). Its corresponding gauge symmetry is given by $SU(2)_L \times U(1)_Y$, which is spontaneously broken via the Higgs mechanism in order to include the mass terms of gauge bosons without breaking gauge invariance. The subscript L is added to emphasize that only left-handed particles are affected, whereas the subscript Y stands for the weak hypercharge.

The Lagrangian of the electroweak sector of the SM can be written as:

$$\mathcal{L}^{EW} = \mathcal{L}^{EW}_{Dirac} + \mathcal{L}^{EW}_{Gauge} + \mathcal{L}^{EW}_{Higgs} + \mathcal{L}^{EW}_{Yukawa} \, . \tag{1.5}$$

The first and second terms describe the fermion fields and their interactions ($\mathcal{L}^{EW}_{Dirac}$) and the gauge fields with their self interactions ($\mathcal{L}^{EW}_{Gauge}$), whereas the third and fourth term introduce the Higgs boson and give mass terms to gauge bosons ($\mathcal{L}^{EW}_{Higgs}$) and fermions ($\mathcal{L}^{EW}_{Yukawa}$). The last two terms are described in more detail in sections 1.1.3 and 1.1.4. The Dirac and gauge terms of the Lagrangian are sufficient to describe a massless theory and are given by

$$\mathcal{L}^{EW}_{Dirac} = \bar{L} i \gamma^\mu D^L_\mu L + \bar{R} i \gamma^\mu D^R_\mu R \, , \tag{1.6}$$

where L and R are the left- and right-handed fermions and the covariant derivatives acting on them are given by

$$\begin{aligned} D^R_\mu &= \partial_\mu + i g \frac{Y}{2} B_\mu \, , \\ D^L_\mu &= \partial_\mu + i g \frac{Y}{2} B_\mu + i g_W \frac{\sigma_i}{2} W^i_\mu \end{aligned} \tag{1.7}$$

and

$$\mathcal{L}^{EW}_{Gauge} = -\frac{1}{4} W^i_{\mu\nu} W^{i,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \, , \tag{1.8}$$

where $B^i_{\mu\nu}$ and $W^i_{\mu\nu}$ are the field strength tensors of $U(1)_Y$ and $SU(2)_L$ respectively:

$$\begin{aligned} B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \, , \\ W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - i g_W \epsilon^{ijk} W^j_\mu W^k_\nu. \end{aligned} \tag{1.9}$$

B_μ is the $U(1)_Y$ gauge field with $\frac{Y}{2}$ as the generator and coupling constant g and W^i_μ is the $SU(2)_L$ gauge field with generators $\frac{\sigma_i}{2}$ and the weak coupling constant g_W.

The last term in Eq. 1.9 is especially interesting for the work in this thesis. It arises from the non-abelian structure of $SU(2)_L$, meaning that in general the sequence of two rotations cannot be interchanged. Expanding Eq. 1.8, this term leads to contributions

Table 1.3.: The weak isospin T_3 together with the weak hypercharge $Y_{L/R}$ and electric charge Q of leptons and quarks in the SM shown for the first generation. The three quantities are connected via the Gell-Mann–Nishijima relation $Q = T_3 + \frac{Y}{2}$ for every fermion. Right-handed neutrinos and left-handed antineutrinos are not included in the SM.

	left-handed		right-handed		
	Y_L	T_3	Y_R	T_3	Q
ν_e	-1	$+\frac{1}{2}$	$-$	$-$	0
e^-	-1	$-\frac{1}{2}$	-2	0	-1
u	$+\frac{1}{3}$	$+\frac{1}{2}$	$+\frac{4}{3}$	0	$+\frac{2}{3}$
d	$+\frac{1}{3}$	$-\frac{1}{2}$	$-\frac{2}{3}$	0	$-\frac{1}{3}$

with three or four $SU(2)_L$ gauge fields resulting in trilinear and quartic couplings in the electroweak sector of the SM. These couplings are fundamental to VBS processes and strongly motivate the choice of EFT operators studied in this thesis.

The two terms of the electroweak Lagrangian expanded above follow directly from the requirement of local gauge symmetry and so far introduce massless bosons. Linear combination of B_μ and $W_\mu^{1,2,3}$ are constructed to recover the three massive bosons, corresponding to the $W^\pm$ and Z boson, and one massless boson, the photon γ. The two fields without electric charge are constructed by mixing B_μ and W_μ^3:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}, \tag{1.10}$$

where θ_W is the weak mixing angle given by $\cos\theta_W = \frac{g_W}{\sqrt{g_W^2 + g^2}}$, A_μ the photon field, and Z_μ the field for the Z boson. Linear combination of the two remaining fields then correspond to the weak gauge bosons $W^\pm$ with electric charge $\pm e$:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2). \tag{1.11}$$

Finally, the weak hypercharge Y of the $U(1)_Y$ symmetry is connected to the electric charge and the third component of the weak isospin by the Gell-Mann–Nishijima relation:

$$Q = T_3 + \frac{Y}{2}. \tag{1.12}$$

Because of the $SU(2)_L$ symmetry, this is different for left-handed and right-handed particles as summarized in table 1.3 for the first generation of fermions.
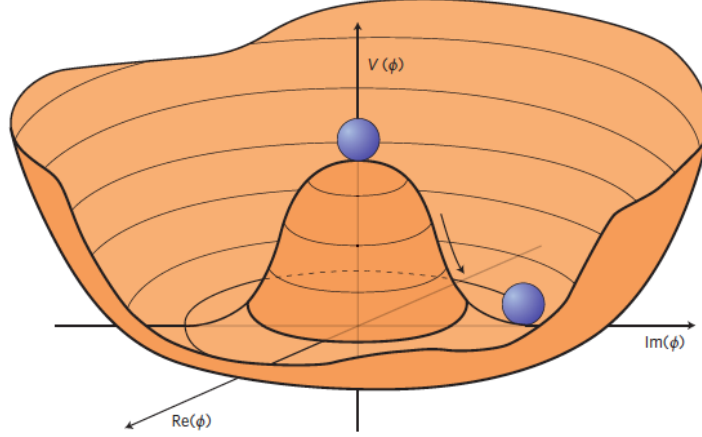
Figure 1.1.: A schematic display of the Higgs potential when $\mu^2 < 0$ as defined in Eq. 1.13. The ground state of the Higgs field is located at nonzero values of the scalar field ϕ resulting in an infinite number of minima. Taken from Ref. [38].

1.1.3. The Higgs Mechanism

One of the major steps that led to the structure of the SM in its current form was the idea of spontaneous symmetry breaking and how it fits perfectly with the electroweak symmetry described above. It leads to masses for weak gauge bosons and also opens a way to solve the challenge that mass terms in the Dirac equation break gauge invariance.

For the Higgs mechanism, a new complex scalar field $\phi = \left(\phi^+, \phi^0\right)^{\mathrm{T}}$ is introduced with hypercharge $Y_\phi = 1$ under $\mathrm{U}(1)_Y$ transformation, weak isospin $T = \frac{1}{2}$ and electric charge of the individual components of $+1$ and 0, which follows from the Gell-Mann–Nishijima relation (Eq. 1.12). The additional term in the Lagrangian with the Higgs potential $V(\phi)$ is given by

$$\begin{aligned}
\mathcal{L}_{\mathrm{Higgs}}^{\mathrm{EW}} &= \left(D_\mu \phi\right)^\dagger \left(D^\mu \phi\right) - V(\phi) \\
&= \left(D_\mu \phi\right)^\dagger \left(D^\mu \phi\right) - \left(\mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2\right) ,
\end{aligned}$$ (1.13)

where

$$D_\mu = \partial_\mu + \mathrm{i} g_{\mathrm{W}} \frac{\sigma_i}{2} W_\mu^i + \mathrm{i} g \frac{Y}{2} B_\mu$$ (1.14)

is the covariant derivative of the Higgs field. The Higgs potential $V(\phi)$, shown in Fig. 1.1, is given by the last two terms in brackets in the Lagrangian. With $\mu^2 < 0$, the characteristic shape of the potential, i.e., the minima are at nonzero values of ϕ is given. Choosing the unitary gauge and expanding the ϕ around a minimum in the unitary gauge results in

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$$ (1.15)

with the vacuum expectation value (VEV) of the Higgs field $v = \sqrt{\frac{-\mu^2}{\lambda}}$ and the scalar component $h(x)$.

The expansion around v for ϕ can then be substituted in Eq. 1.15 and all terms expanded. With the same definitions of the photon field A_μ and the fields for the Z and $W^\pm$ bosons defined in Eq. 1.10 and 1.11, mass terms for the three gauge fields Z and $W^\pm$ appear as well as a new scalar field h, which is identified as the Higgs field. The masses of these fields can be directly taken from the resulting Lagrangian and are

$$M_{W^\pm} = \frac{g_W v}{2}, \qquad M_Z = \frac{M_W}{\cos\theta_W}, \qquad M_H = \sqrt{2\lambda}v, \qquad (1.16)$$

while the photon field A_μ remains massless.

The value of v can be measured via the Fermi coupling constant G_F, which is connected to the mass of the W boson and is precisely measured. The value for the VEV is found to be $v = 246.22\,\text{GeV}$ [37].

1.1.4. Fermion Masses

The mechanism described above gives mass terms to bosons but not to fermion. To include this in a gauge invariant way, a Yukawa term is added to the Lagrangian introducing the interaction between the scalar field ϕ introduced in the Higgs mechanism and fermion fields. This is the last term in Eq. 1.5:

$$\mathcal{L}_{\text{Yukawa}}^{\text{EW}} = -\bar{R}_e C_e \phi L_e + \bar{R}_u C_u \phi L_q - \bar{R}_{d'} C_d \phi L_q + \text{h.c.}, \qquad (1.17)$$

with the left-handed doublet L of $\text{SU}(2)_L$ consisting of one up- and one down-type quark (L_q) or one charged lepton with corresponding neutrino (L_e), $R_{e,u,d}$ the right-handed singlet either for charged leptons, up- or down-type particles, ϕ the Higgs doublet, C_e, C_u and C_d are complex 3×3 matrices for the three generations, and h.c. is the hermitian conjugate of the first three terms. There are two terms for quarks and one term for leptons because there are no right-handed neutrinos.

After spontaneous symmetry breaking and expanding the scalar field ϕ around its minimum, this term generates mass terms for fermions: $m_f = y_f \frac{v}{\sqrt{2}}$ with the Yukawa couplings y_f. Since eigenstates of the weak interaction are not the same as mass eigenstates, the Cabibbo-Kobayashi-Maskawa (CKM) [39, 40] matrix, an additional 3×3 matrix, is introduced relating the matrices $C_{u,d}$ to the Yukawa couplings for quarks. The CKM matrix has to fulfill theoretical constraints such as unitarity and has four free parameters that have to be determined from experiment. Because weak and mass eigenstates are not the same, non-diagonal elements of the CKM matrix allow transition of one generation of quarks to another via the weak interaction. In the case of massive neutrinos, a similar matrix, the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) [41, 42] matrix is introduced to explain for example neutrino oscillations.

1.1.5. Strong Interaction

Besides the electromagnetic and weak interaction, the SM also described the strong interaction described by quantum chromodynamics (QCD) and follows from local $SU(3)_C$ gauge symmetry with the index C emphasizing that it couples to color charge. As described above, quarks and gluons carry color charge and are subject to the strong interaction.

The terms in the SM Lagrangian for QCD are given by

$$\mathcal{L}^{\text{QCD}} = \bar{\psi}(\text{i}\gamma^\mu)(D_\mu)\psi - m\bar{\psi}\psi - \frac{1}{4}F^a_{\mu\nu}F^{a\mu\nu} \,, \tag{1.18}$$

with the covariant derivative in QCD given by

$$D_\mu = \partial_\mu - \text{i}g_\text{s}t^a A^a_\mu \,, \tag{1.19}$$

quark fields ψ, Dirac matrices γ, the gluon field strength tensor $F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g_\text{s}f^{abc}A^b_\mu A^c_\nu$, gluon fields A^a_μ, and coupling strength g_s. The structure constant of $SU(3)_C$ is given by f^{abc} with the generators t^a.

The theoretical structure is very similar to the $SU(2)$ symmetry for the weak interaction with the difference that there are eight massless gluons corresponding to the generators t^a. In particular, the non-abelian structure of both $SU(2)$ and $SU(3)_C$ results in trilinear or quartic vertices in the case of electroweak interaction and in self interactions for gluons. Another very important difference is that the coupling strength for the strong interaction g_s is larger than for the weak interaction. This leads to correction terms due to gluon and quark loops, effectively introducing a running coupling strength α_s depending on the energy scale under consideration, Q^2:

$$\alpha_s\left(Q^2\right) = \frac{\alpha_\text{s}(\mu^2)}{1 + \alpha_\text{s}(\mu^2)\frac{1}{12\pi}(11n_\text{c} - 2n_\text{f})\ln\frac{Q^2}{\mu^2}} \,, \tag{1.20}$$

where n_f is the number of quark generations, n_c the number of color charges, and μ a reference energy scale, which in measurements is often chosen to be the mass of the Z boson: $\mu = m_\text{Z}$.

Large energies and correspondingly small distances lead to vanishing values of α_s, meaning that the strong interaction becomes weaker, called asymptotic freedom. Conversely, the value of α_s increases for larger distances such that color-charged particles cannot be observed isolated and color-neutral composite particles are formed, called hadrons. They can be classified into mesons, which are composed of a quark and an antiquark with matching color and anticolor, and baryons, which are composed of three quarks with all three colors or correspondingly antiquarks with anticolors.

Lastly, perturbation theory can only be applied for values of $\alpha_\text{s} \ll 1$ above an energy scale Λ_QCD.

1.2. Vector Boson Scattering

When the LHC was built, one of the main motivations was the search for the Higgs boson and to investigate the electroweak sector to which it is closely related. A peculiarity of the electroweak sector is the self-coupling of massive bosons, namely triple gauge couplings (TGC) and quartic gauge couplings (QGC). Additionally, the mass of the W and Z bosons give rise to couplings to the Higgs boson. Relevant Feynman diagrams for the scattering of vector bosons resulting in two opposite sign W bosons are shown in Fig. 1.2. They are of the order of $\mathcal{O}(\alpha^2)$, since QGC carry two orders and TGC one order of α. Diagrams with the Higgs boson include it either in the s channel or the t channel as shown in subfigures 1.2(d) and 1.2(e), respectively. If the electric charge of the two vector bosons after scattering sums up to zero, contributions with the Higgs boson via the s channel are allowed, whereas this is not the case, e.g., for one $W^\pm$ and one Z boson. Electroweak symmetry breaking gives rise to the masses of W and Z bosons, as well as their longitudinal polarization. Furthermore, diagrams with the Higgs boson are necessary to keep the scattering cross section finite due to the negative sign of the interference term. Therefore, VBS is a key process when studying the EW sector and properties of the Higgs boson.
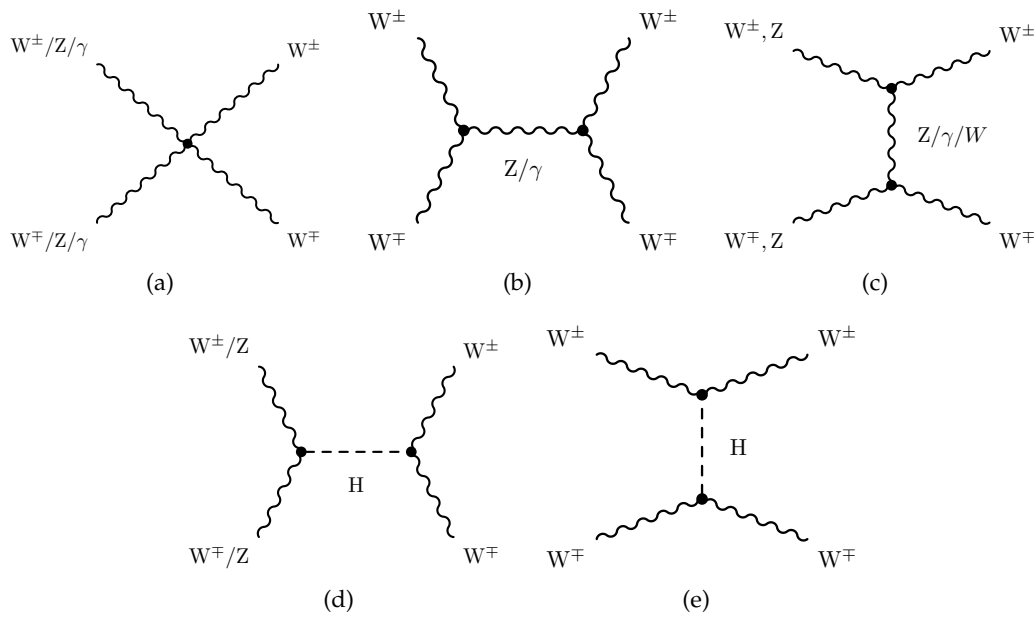

Figure 1.2.: Tree-level Feynman diagrams realized in VBS in the case of two opposite-sign $W^\pm$ bosons. Diagram (a) shows the quartic gauge coupling and TGC can be included via s-channel ((b)) or t-channel exchange (c). Finally, contributions including the Higgs boson are also possible in the s channel (d) and t channel (e).
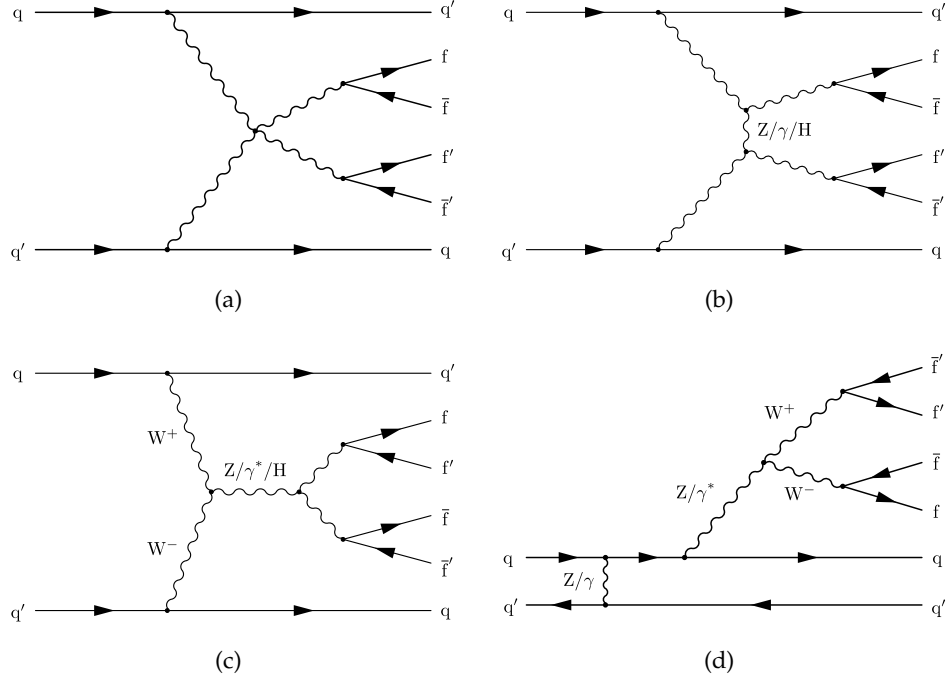
Figure 1.3.: Typical Feynman diagrams that contribute to VBS. Shown are the quartic interaction in (a) and *t*-channel and *s*-channel contributions in (b) and (c). Since electric charge is conserved, *s* channel diagrams are not contributing to the same-sign WW channel. Further diagrams have to be included to form a gauge invariant set such as the one shown in (d).

Since the LHC is a proton-proton collider, the initial state consists of gluons or quarks. In the case of VBS, two quarks each radiate off a vector boson which then scatter and subsequently decay. Heavy vector bosons, $W^\pm$ or Z bosons, decay into two fermions either hadronically into quark-antiquark pairs or leptonically into two charged leptons or one charged lepton and one neutrino. Accordingly, the final state after decay can be split into three categories: *fully-leptonic* with two quarks and four leptons, *semi-leptonic* with four quarks and two leptons, and *fully-hadronic* with six quarks and without leptons. From an simulation point of view, the vector bosons as an intermediate state are required and the decay is simulated afterwards. VBS is then defined as the electroweak production of two vector bosons together with two quarks and subsequent decay of the vector bosons. In perturbation theory, this process is of order $\mathcal{O}(\alpha_{\mathrm{ew}}^6)$ including the decay. More specifically, in the case of two opposite-sign W bosons decaying leptonically:

$$p\,p \to W^+ W^- q\,q \to \ell^+ \nu\, \ell^- \nu\, q\,q, \tag{1.21}$$

where p stands for proton and q for any quark. Many possible Feynman diagrams contribute to the electroweak production in order to form a gauge-invariant set, some of which are exemplary shown in Fig. 1.3.
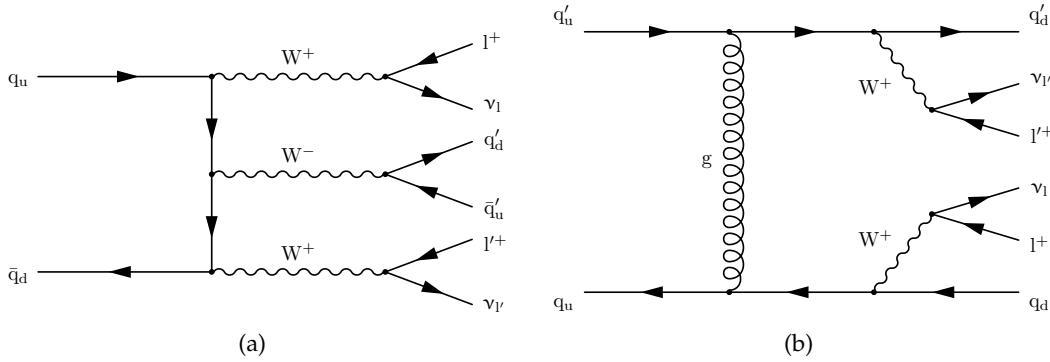
Figure 1.4.: Example of Feynman diagrams for triboson production (a) and QCD-induced VBS production (b) are shown. They contribute to the same final state as the EW VBS production and can be separated by selection criteria on the two outgoing quarks.

In addition to the purely electroweak contribution, there are also QCD-induced diagrams of order $\mathcal{O}(\alpha_{\text{ew}}^4 \, \alpha_{\text{s}}^2)$, where a gluon exchange is included. One of these diagrams is shown in Fig. 1.4(b) and is treated as background in VBS analyses. Following the rules of quantum mechanics, interference of the EWK and QCD-induced contributions of the order $\mathcal{O}(\alpha_{\text{ew}}^5 \, \alpha_{\text{s}})$ exists. Looking at each particle in the final state individually, they cannot be distinguished and a single event cannot be unambiguously attributed to a single contribution. Nevertheless, the colorless diagrams of the electroweak contribution lead to location of two quarks in the forward region of the detector [43], whereas the vector bosons and their decay products tend to be in the central region. This signature is used to separate the contributions and analyses therefore define a signal-enriched phase space, where two quarks are reconstructed as so-called *tagging jets* in the forward region and reconstructed objects in the central region are compatible with the targeted vector bosons. This is schematically shown in Fig. 1.5.

In order to target the EWK VBS production, analysis use cuts targeting both the tagging jets and the decay products of vector bosons. Since the tagging jets of EWK VBS production are in opposite direction with respect to the beam axis, a large separation in pseudorapidity[1] $\Delta\eta_{\text{jj}}$ is required and consequently a large invariant mass $m_{\text{jj}} = \sqrt{(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2}$. This reduces the contribution of QCD-induced diagrams significantly. Additionally, decay products of vector bosons in the central region of the detector are required to be compatible with the vector boson mass, which reduces contributions from non-resonant diagrams. Furthermore, EWK VBS shows a reduced additional jet activity in the central region. In the case of leptonically decaying vector bosons, this leads to a central jet veto (CJV) [44], which in simulation strongly

---

[1]The pseudorapidity is defined in Eq. 2.3

depends on the choice of algorithms [45] and the impact of higher orders. Since this thesis focuses on the hadronic decay channel, a CJV is not further investigated.

Diboson production via VBS has first been studied in the LHC Run 1 at $\sqrt{s} = 8\,\text{TeV}$, where leptonically decaying vector bosons are considered [46–52]. With the higher center-of-mass energy of the LHC Run 2, VBS production of two same-sign W bosons or a W together with a Z boson and subsequent leptonic decay have been observed during the first year of the LHC Run 2 [53–56]. With the full recorded dataset at $\sqrt{s} = 13\,\text{TeV}$, VBS ZZ production with leptonic decays was observed by the ATLAS Collaboration [57] and the CMS Collaboration exceeded the threshold for evidence but not yet for observation [58]. In the case of two opposite-sign W bosons, a Z or a W boson together with a photon, observations have also been made [59–63].
Analyses in the semi-leptonic channel of VBS are expected to reach the threshold for observation but so far only reported evidence [64, 65]. Lastly, there is no public measurement of VBS in the hadronic decay channel up to today. A more detailed overview of the status of VBS can be found in Refs. [66–69].

With the SM production of VBS observed in above-mentioned decay channels and with more likely to follow, the focus of VBS analyses shifted towards precision measurements. Higher number of events and even larger center-of-mass energies in the following runs of the LHC make it possible to separate the final states further by polarization of the vector bosons. In particular, the longitudinal polarization is of interest since it is enabled by the Higgs mechanism and unitarity is achieved by intricate cancellation of contributions from vector bosons, Higgs boson, and interference terms. Furthermore, VBS allows setting constraints in terms of an EFT, especially when TGC and QGC are affected. This is subject of the following section.
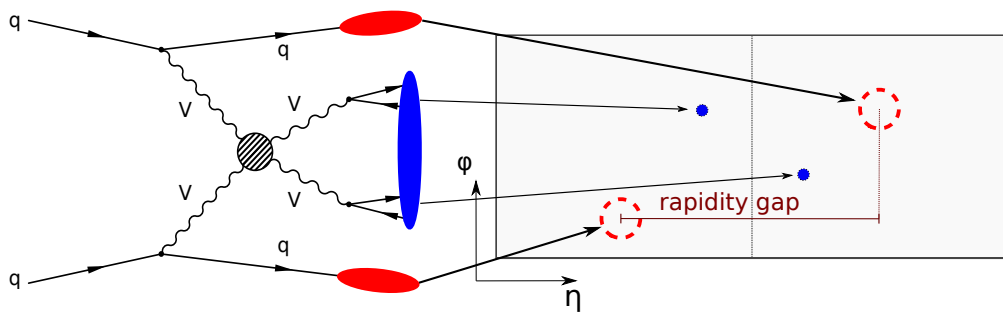

Figure 1.5.: A schematic view of the event topology of VBS at the LHC. On the left, a typical Feynman diagram is shown and on the right a rolled out representation of the detector. The two market quarks are located in the detector in opposite directions with a separation in η, called rapidity gap.

1.3. Effective Field Theory

The SM had many successes as outlined before which are hard to underestimate, but it is nevertheless not a final theory. Most evident is the fact that it does not include all known fundamental forces and gravity is described by the Theory of General Relativity [34]. This leads to the idea of Grand Unification [70], where all fundamental forces are unified at sufficiently high energies. Other challenges are experimentally found phenomena without established theories: neutrino oscillations are well-observed and not possible in the SM, where they are assumed to be massless. Also dark matter [71] and dark energy [72] are not explained in the SM, since it does not contain a particle that could account for their effects.

There are many theories that could extend the SM in order to explain and solve these problems. These searches for physics beyond the SM are a large part of the agenda at the LHC and can be performed by directly looking for new particles via searches for resonances in the measured data. If the new particle is too heavy and not reachable at center-of-mass energies currently available, their impact can potentially still be seen.

A framework to describe the low energy effect of new resonances that are not yet reachable is the effective field theory approach [73, 74]. The underlying idea is to describe in a bottom-up approach the low-energy effect of BSM physics by changing the behavior of already known interactions. The SM is then the low energy approximation and the EFT, which is valid up to an energy scale Λ_{valid}, gives insights to the deviation from the SM. By analyzing the footprints of specific models in EFT parameter space, hints and constraints on them can be deduced.

1.3.1. Fermi Theory

Fermi's theory of beta decay is a theory that was published in 1934 [19, 20] before the SM was formulated as a quantitative theory to explain beta decay. It postulated the existence of neutrinos, which would be emitted together with an electron from a nucleus under beta decay. This explained the continuous energy spectrum of the emitted electrons and can be depicted as a Feynman diagram as shown in Fig. 1.6(a). The contact interaction is given by the coupling strength G_{F} and is accurate for energies below the energy scale given by the mass of the W boson: $m_{\mathrm{W}} \approx 80\,\mathrm{GeV}$. The cross-section for this process increases with energy and violates unitarity at high enough energies. In the SM, this process is then described by the exchange of a W boson as shown in Fig. 1.6(b).

This theory is a famous and historically relevant example of an EFT for the weak interaction and worked very well such that G_{F} is still often considered as one of the fundamental parameters of the SM.
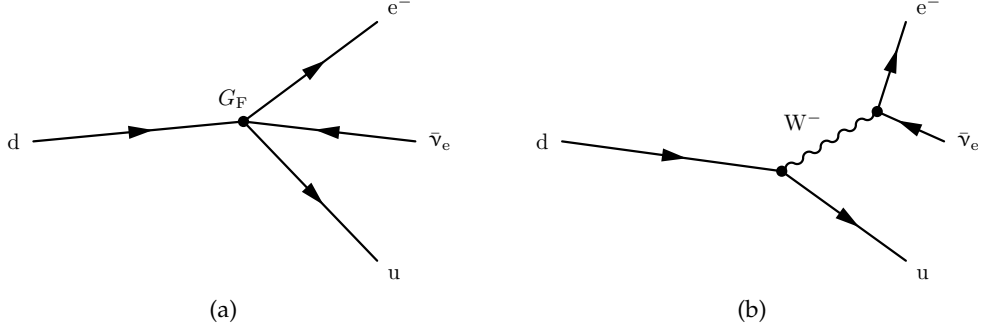
Figure 1.6.: Leading order Feynman diagrams for β^- decay as described in Fermi's theory for β^- decay (a) and in the SM (b). In Fermi's theory, a contact interaction of all four participating fermions is determined by the Fermi coupling constant G_F, which is measured via the muon lifetime.

1.3.2. The SM as an EFT

Because the SM is not the final theory yet as described above, it is often regarded as an incomplete theory and as a low-energy approximation of a more complete theory. In the EFT framework, new physics is assumed to be at an energy scale Λ larger as currently accessible energies. Additional terms are then included in the Lagrangian with higher mass dimensions as terms in the SM contribution, $d > 4$:

$$\mathcal{L}_{\text{EFT}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i^{(6)}}{\Lambda^2} \mathcal{Q}_i^{(6)} + \sum_i \frac{f_i^{(8)}}{\Lambda^4} \mathcal{O}_i^{(8)} + \cdots , \tag{1.22}$$

where Λ is the energy scale of new physics, $c_i^{(6)}$ and $f_i^{(8)}$ are Wilson coefficients for dimension-6 and dimension-8, respectively, i.e., dimensionless coefficients describing the strength of the coupling, and $\mathcal{Q}_i^{(6)}$ ($\mathcal{O}_i^{(8)}$) are operators with mass dimension-6 (8) preserving SM symmetries. Throughout this thesis, c_i will refer to Wilson coefficients of dimension-6 operators and f_i to dimension-8 operators. Adding these operators means that at higher energies, unitarity is violated resulting in divergent cross sections and the invalidity of the probability interpretation.

Operators with odd mass-dimension lead to violation of the lepton or baryon number and are not considered here resulting in the lowest order term of dimension-6. To recover the overall dimension of four, operators of dimension d are suppressed by factors of Λ^{d-4}. Analyses of VBS often focus on dimension-8 operators with the argument that dimension-6 operators are better constraint in diboson production. However, recently the focus also moved to dimension-6 operators to investigate this phase space and with the prospect of a possible combination of dimension-6 and dimension-8 operators, since the sensitivity has been shown to be compatible for some operators [75].

Dimension 6 operators

Since the only dimension-5 operator violates lepton number and may explain neutrino masses via heavy Majorana neutrinos [11], the operators of lowest dimensionality considered here are dimension-6 operators. The number of possible operators is given by all possible combinations of SM fields with reducing all redundant combination and heavily depends on the dimensionality. In total, there are 2499 possible dimension-6 operators [76], although additional symmetries are often considered. Assuming baryon conservation and only a single flavor, the number is reduced to 76 [13].

A systematic way of constructing these operators lead to the Warsaw basis [12], which became more relevant in recent times at the LHC [77]. The SMEFTsim package [13, 78] explicitly implements the Warsaw basis for use in simulation. It offers a choice of EW input schemes, i.e., which parameters are taken from measurements, and a set for different symmetry assumptions regarding the flavor structure. Since in the analyses in this thesis, there is no sensitivity to different quark flavors, the $U(3)^5$ flavor symmetry is chosen. A set of three operators is investigated, since they only affect the direct coupling of vector bosons and do not affect their decay:

$$
\begin{aligned}
Q_W &= \epsilon^{ijk} W_\mu^{i\nu} W_\nu^{j\rho} W_\rho^{k\mu} \\
Q_{\phi\Box} &= (\phi^\dagger \phi)\Box(\phi^\dagger \phi) \\
Q_{\phi W} &= (\phi^\dagger \phi) W_{\mu\nu}^i W^{i\mu\nu} ,
\end{aligned}
\tag{1.23}
$$

where ϕ is the Higgs doublet, $W_\mu^{i\nu}$ the field of $SU(2)_L$, and $\Box$ is the d'Alembert operator. The effect of EFT typically shows up in the high-energy tail of distributions. Sensitive variables depend on the operator and process of choice and were studied for a subset of dimension-6 operators for VBS in Ref. [75].

Limits for dimension-6 operators affecting TGC can be set in VBS and also in diboson production, with the latter typically being the used in the past. While working towards an LHC-wide combination of different processes, VBS also became of interest. Fig. 1.7 shows a schematic representation of Wilson coefficients and affected processes, which are candidates for a possible combination [77].

Dimension 8 operators

When considering dimension-8 operators as a possible extension to the SM, one is confronted with an overwhelming number reaching well above ten thousand [79]. However, only a subset is of interest considering VBS: for purely modification of QGCs, dimension-8 operators are required [14], which are therefore of the most interesting to study in VBS. They are sorted by their theoretical way of construction resulting in operators $\mathcal{O}_{S,i}$ with two covariant derivatives, $\mathcal{O}_{M,i}$ with exactly one, and $\mathcal{O}_{T,i}$ without
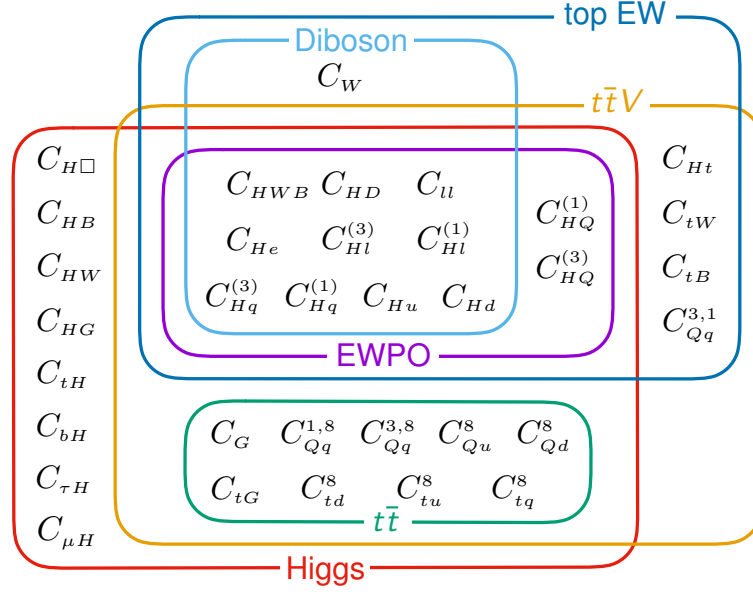
Figure 1.7.: A schematic representation of a subset of Wilson coefficients in the Warsaw basis and sensitive processes. This figure is taken from Ref. [77].

covariant derivatives. Following Ref. [80], the following 18 genuine QGC operators are considered:

$$
\begin{aligned}
\mathcal{O}_{S,0} &= [(D_\mu\phi)^\dagger D_\nu\phi] \times [(D^\mu\phi)^\dagger D^\nu\phi] & \mathcal{O}_{T,0} &= \mathrm{Tr}\left[\widehat{W}_{\mu\nu}\widehat{W}^{\mu\nu}\right] \times \mathrm{Tr}\left[\widehat{W}_{\alpha\beta}\widehat{W}^{\alpha\beta}\right] \\
\mathcal{O}_{S,1} &= [(D_\mu\phi)^\dagger D^\mu\phi] \times [(D_\nu\phi)^\dagger D^\nu\phi] & \mathcal{O}_{T,1} &= \mathrm{Tr}\left[\widehat{W}_{\alpha\nu}\widehat{W}^{\mu\beta}\right] \times \mathrm{Tr}\left[\widehat{W}_{\mu\beta}\widehat{W}^{\nu\alpha}\right] \\
\mathcal{O}_{S,2} &= [(D_\mu\phi)^\dagger D_\nu\phi] \times [(D^\nu\phi)^\dagger D^\mu\phi] & \mathcal{O}_{T,2} &= \mathrm{Tr}\left[\widehat{W}_{\alpha\mu}\widehat{W}^{\mu\beta}\right] \times \mathrm{Tr}\left[\widehat{W}_{\beta\nu}\widehat{W}^{\nu\alpha}\right] \\
\mathcal{O}_{M,0} &= \mathrm{Tr}\left[\widehat{W}_{\mu\nu}\widehat{W}^{\mu\nu}\right] \times [(D_\beta\phi)^\dagger D^\beta\phi] & \mathcal{O}_{T,3} &= \mathrm{Tr}\left[\widehat{W}_{\mu\nu}\widehat{W}_{\alpha\beta}\right] \times \mathrm{Tr}\left[\widehat{W}^{\alpha\nu}\widehat{W}^{\mu\beta}\right] \\
\mathcal{O}_{M,1} &= \mathrm{Tr}\left[\widehat{W}_{\mu\nu}\widehat{W}^{\nu\beta}\right] \times [(D_\beta\phi)^\dagger D^\mu\phi] & \mathcal{O}_{T,4} &= \mathrm{Tr}\left[\widehat{W}_{\mu\nu}\widehat{W}_{\alpha\beta}\right] \times B^{\alpha\nu}B^{\mu\beta} \\
\mathcal{O}_{M,2} &= [B_{\mu\nu}B^{\mu\nu}] \times \left[(D_\beta\phi)^\dagger D^\beta\phi\right] & \mathcal{O}_{T,5} &= \mathrm{Tr}\left[\widehat{W}_{\mu\nu}\widehat{W}^{\mu\nu}\right] \times B_{\alpha\beta}B^{\alpha\beta} \\
\mathcal{O}_{M,3} &= [B_{\mu\nu}B^{\nu\beta}] \times \left[(D_\beta\phi)^\dagger D^\mu\phi\right] & \mathcal{O}_{T,6} &= \mathrm{Tr}\left[\widehat{W}_{\alpha\nu}\widehat{W}^{\mu\beta}\right] \times B_{\mu\beta}B^{\alpha\nu} \\
\mathcal{O}_{M,4} &= \left[(D_\mu\phi)^\dagger \widehat{W}_{\beta\nu}D^\mu\phi\right] \times B^{\beta\nu} & \mathcal{O}_{T,7} &= \mathrm{Tr}\left[\widehat{W}_{\alpha\mu}\widehat{W}^{\mu\beta}\right] \times B_{\beta\nu}B^{\nu\alpha} \\
\mathcal{O}_{M,5} &= \left[(D_\mu\phi)^\dagger \widehat{W}_{\beta\nu}D^\nu\phi\right] \times B^{\beta\mu} + \text{h.c.} & \mathcal{O}_{T,8} &= B_{\mu\nu}B^{\mu\nu}B_{\alpha\beta}B^{\alpha\beta} \\
\mathcal{O}_{M,7} &= \left[(D_\mu\phi)^\dagger \widehat{W}_{\beta\nu}\widehat{W}^{\beta\mu}D^\nu\phi\right] & \mathcal{O}_{T,9} &= B_{\alpha\mu}B^{\mu\beta}B_{\beta\nu}B^{\nu\alpha}
\end{aligned}
$$

$$(1.24)$$

where ϕ is the Higgs doublet, D_μ the covariant derivative as defined in Eq. 1.14 with $Y_\phi = 1$, $B_{\mu\nu}$ the field strength of the $U(1)_Y$ group and $\widehat{W}_{\mu\nu} = W^j_{\mu\nu} \frac{\sigma^j}{2}$ of SU(2)$_L$. These genuine QGC operators conserve electric charge, as well as charge C and parity P both separately and combined.

In the Lagrangian, each operator comes with a separate Wilson coefficient: $f_{S,i}$, $f_{M,i}$, or $f_{T,i}$. Dimensionless limits are then set on f_i/Λ^4. Table 1.4 shows the set of operators and which quartic vertex is affected by the respective operator.

Table 1.4.: Set of dimension-8 operators as described in Ref. [14]. "X" denotes a vertex that is affected by the respective operators. The hadronic decay channel of VBS is sensitive to all listed operators.

	$\mathcal{O}_{S,0}$, $\mathcal{O}_{S,1}$, $\mathcal{O}_{S,2}$	$\mathcal{O}_{M,0}$, $\mathcal{O}_{M,1}$, $\mathcal{O}_{M,7}$	$\mathcal{O}_{M,2}$, $\mathcal{O}_{M,3}$, $\mathcal{O}_{M,4}$, $\mathcal{O}_{M,5}$	$\mathcal{O}_{T,0}$, $\mathcal{O}_{T,1}$, $\mathcal{O}_{T,2}$	$\mathcal{O}_{T,5}$, $\mathcal{O}_{T,6}$, $\mathcal{O}_{T,7}$	$\mathcal{O}_{T,8}$, $\mathcal{O}_{T,9}$
WWWW	X	X		X		
WWZZ	X	X	X	X	X	
ZZZZ	X	X	X	X	X	X
WWZγ		X	X	X	X	
WW$\gamma\gamma$		X	X	X	X	
ZZZγ		X	X	X	X	X
ZZ$\gamma\gamma$		X	X	X	X	X
Z$\gamma\gamma\gamma$				X	X	X
$\gamma\gamma\gamma\gamma$				X	X	X

2. The LHC and the CMS experiment

Since physics gives a quantitative description of nature, it is necessary to connect the theoretical models and calculations to experiments and the behavior of the world around us. Experiments in modern physics are designed to prepare specific situations in order to probe complex phenomena of interest, which are then measured with tools tailored to be as sensitive as possible to possible outcomes.

In the case of particle physics, the SM describes the fundamental building blocks of the universe and how they behave at length scales as small as possible. For this reason, scattering experiments with energies are built, where particles are accelerated and collided at high center-of-mass energies. The largest such accelerator is the Large Hadron Collider (LHC) at CERN colliding protons at center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ during the LHC Run 2. At and around the collision points, detectors are built such as the Compact Muon Solenoid (CMS) detector in order to gather the information about the collision.

The first part of this chapter describes the LHC, the experimental setup necessary for the acceleration and collision of the particles under consideration. In the second part, the setup and function of the CMS detector is explained.

2.1. The Large Hadron Collider at CERN

The LHC [6, 81] at CERN is located near Geneva, Switzerland, and as of today is the most powerful and largest collider. Protons are accelerated in two beams crossing each other and colliding at four collision points. The machine is designed to bring each proton beam to an energy of $7\,\text{TeV}$ resulting in a center-of-mass energy of $\sqrt{s}_\text{design} = 14\,\text{TeV}$ in head-on collisions and deliver a luminosity of $L_\text{design} = 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$. The LHC was first mentioned in 1977 [82] as a possible proton-proton collider, presented to the CERN Council in 1993 [83], and approved in 1994 [84]. The LHC is built inside the $27\,\text{km}$-long tunnel previously housing the Large Electron Positron Collider (LEP) and collide protons with unprecedented energies.

The motivations for building the LHC are detailed in the CERN Resolution from 1993 and include the search and investigation of the top quark and Higgs boson, since they were missing particles of the SM at that time and it was clear that either they are

discovered at the LHC or other phenomena show up. Furthermore, CP violation in the mass matrix of fermions should be investigated and whether there is an underlying substructure in the quark sector explaining the number of generations. Since the Higgs boson is an important part of the electroweak sector of the SM, the motivation included the search and study of it and whether there is ultimately a unified force of the strong and electroweak interactions. Searches for supersymmetry and other theories beyond the SM were also a part of the physics program from the very beginning. From an economical point of view, it was also well motivated to reuse the pre-existing tunnel of LEP, especially due to the political turmoil in some European member states at that time [83].

Many of the initial ambitions have been achieved or boundaries of knowledge have been pushed further, where the most famous achieved goal is the discovery of the Higgs boson in 2012 by the CMS and ATLAS collaborations [4, 5]. Since the top quark was already discovered 1995 at the Tevatron [85], this was the last missing piece of the SM, on which the theory of electroweak symmetry breaking heavily relies. Besides the increase in physical understanding of our world, bringing together people from many different countries is also an enormous achievement.

As a circular collider, the LHC uses the same beam paths multiple times allowing it to reach higher collision energies than a linear collider. In order to keep the beams on the paths predefined by the tunnel, a finely adjusted setup of superconducting magnets preserves their trajectory. Additionally, a chain of pre-accelerators prepares the protons before they can be injected in the final tunnel of the LHC. An overview of the accelerator complex is shown in Fig. 2.1.

Preaccelerators

Before 2020, the protons originated from hydrogen atoms, which are stripped off their electrons with an electrical field and are then accelerated and focussed by a linear accelerator, the Linear Accelerator 2 (LINAC2), using radio frequency quadrupoles tuned to the proton mass and charge. Since 2020, the LINAC2 is replaced by the Linear Accelerator 4 (LINAC4) [86], which accelerates negatively charged hydrogen ions up to an energy of 160 MeV. At the subsequent injection point, the electrons are knocked off, which then results in a beam of protons. The next accelerators in the chain, the BOOSTER [87] and Proton Snychrotron (PS), are circular particle accelerators bringing the speed of protons to almost the speed of light and increase the beam intensity. The PS was the world's leading accelerator when it started in 1959 [88] and is now part of the preaccelerator chain feeding protons with energies of 25 GeV into the Super Proton Synchrotron (SPS), the final accelerator before the LHC ring. Historically, the SPS is the accelerator that provided protons and antiprotons for collisions that allowed the discovery of the W and Z bosons by the UA1 and UA2 collaborations [89, 90]. Finally, the protons are injected into the LHC ring at an energy of 450 GeV.
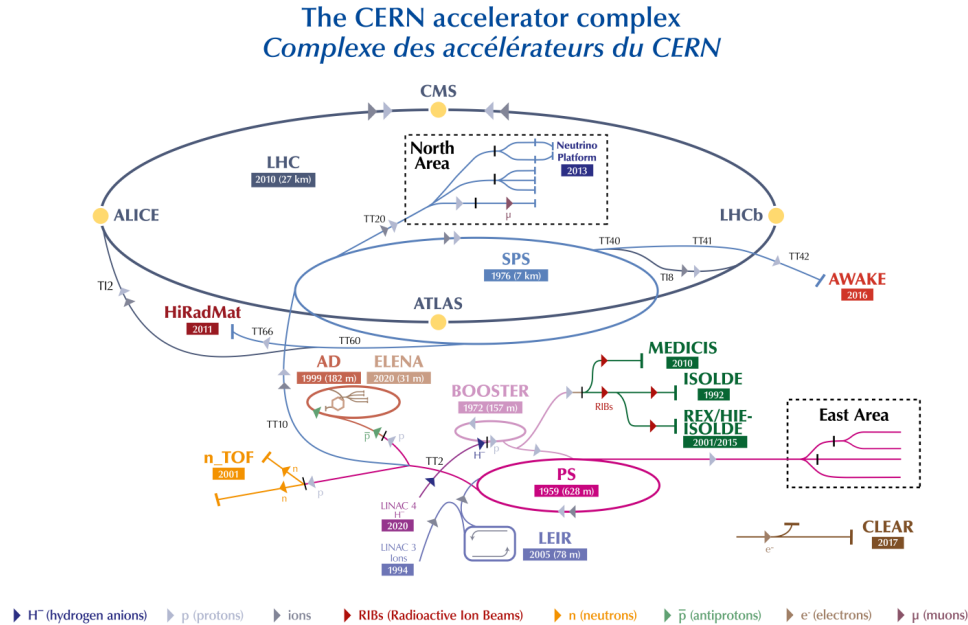
Figure 2.1.: The CERN accelerator complex [91] together with the four experiments CMS, ATLAS, LHCb, and ALICE. Further experiments and preaccelerators are also shown. The LINAC2 accelerator was superseded by the LINAC4 in 2020.

The LHC Ring

The 26.7 km long ring of the LHC houses two vacuum beam pipes, one for each beam of protons travelling in opposite directions. Each beam consists of up to 2808 bunches, which themselves contain in the order of 10^{11} protons. During the 2016 to 2018 data-taking period, the so-called Run 2 of the LHC, the spacing between bunches was 25 ns and the beam energy was 6.5 TeV resulting in a center-of-mass energy of 13 TeV. The beams are deflected by 1232 superconducting main dipoles in order to keep them on their circular trajectory. Additionally, quadrupoles and higher-order multipoles focus the bunches within the beam pipe by alternatingly squeezing their width and height. In order to reach the necessary high magnetic fields of 8.33 T in case of the dipole magnets, superconductivity is achieved by cooling the magnets to a temperature of 1.9 K using liquid helium.

Collisions of beams are produced at four points, where the two beams cross each other. At each of these crossing points, one of the four main detectors is located to measure the resulting product as shown in Fig. 2.1. The CMS (Compact Muon Solenoid) detector [7] and the ATLAS detector [8] are multipurpose detectors located at the northern and southern part of the ring, respectively. ALICE (A Large Ion Collider Experiment) [9] is

a detector specialized in recording lead-lead collisions, and the LHCb (Large Hadron Collider beauty) [10] experiment is an asymmetric detector focusing on measurements involving rare decays of B hadrons and CP violation.

In addition to the type of accelerator and the center-of-mass energy, a very important characteristic of a collider is the instantaneous luminosity L, which together with the cross section of a process p gives the number of produced events per second:

$$\dot{N} = L \cdot \sigma_{\mathrm{p}} \, . \tag{2.1}$$

The cross section σ_{p} depends on both the process and center-of-mass energy, while the instantaneous luminosity L is solely dependent on beam parameters and following Ref. [92] can be written for a Gaussian distribution as

$$L = \frac{N_1 N_2 N_{\mathrm{b}} f_{\mathrm{rev}}}{4\pi \sigma_x \sigma_y} \, , \tag{2.2}$$

where N_1 and N_2 are the number of particles per bunch, N_{b} the number of bunches, $\sigma_{x,y}$ the width of the bunch in the respective direction, and f_{rev} is the revolution frequency given by the length of the LHC tunnel and the speed of light: $f_{\mathrm{rev}} \approx 11245\,\mathrm{Hz}$.

Integrating the instantaneous luminosity over time gives the integrated luminosity, which is used to describe the amount of collisions delivered by the LHC:

$$L_{\mathrm{int}} = \int L \, \mathrm{d}t \, . $$

Figure 2.2 shows the integrated luminosity collected by the CMS detector as a function of time. The analysis described in this thesis uses proton-proton collisions collected by the CMS detector during the Run2, i.e., taken in the years 2016, 2017, and 2018 at a center-of-mass energy of 13 TeV.
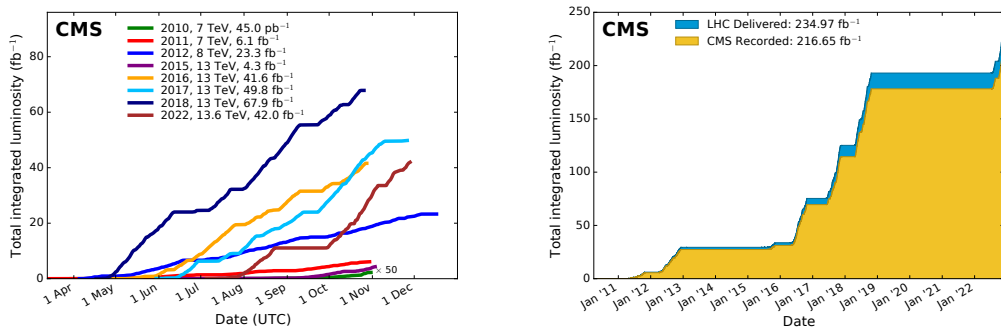

Figure 2.2.: The integrated luminosity delivered to the CMS detector during proton-proton collisions with stable beams. The cumulative luminosity for each year is shown in the left and the overall luminosity for all years in the right. Taken from Ref. [93]

2.2. The Compact Muon Solenoid Experiment

The CMS detector [7] is one of the two multi-purpose detectors at the LHC and is located near Cessy on the border of France and Switzerland. The detector is 100 m underground and built cylindrical around the interaction point. It has a length of 21.6 m, a diameter of 14.6 m, and a total weight of 14 000 t [94]. A design choice when building the detector was the powerful magnetic field of 3.8 T in order to study muons resulting from the subsequent decay of a Higgs boson. The tracker and calorimeters had to fit inside the superconducting solenoid giving it its name.

A slice of the detector with its subsystems and some exemplary particles is shown in Fig. 2.3. The individual parts starting from the inside to the outside are described in the following sections: close to the interaction point is a silicon based tracker containing pixels closest to the collision and strips farther away in order to measure the trajectory of particles, which is bend if they are electrically charged. Then the electromagnetic and hadron calorimeter follow, which measure the energy of the incoming particles by absorbing them. This is followed by the superconducting solenoid, which creates the magnetic field and finally an iron return yoke to stabilize it, inlayered by muon chambers.
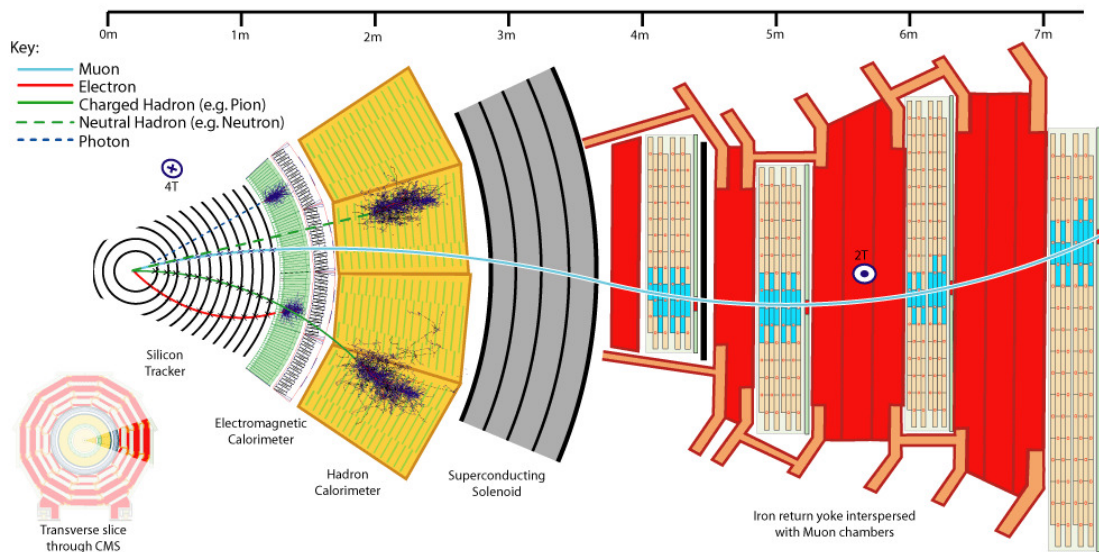

Figure 2.3.: A slice of the CMS detector with its subsystems and exemplary particles. The subsystems are described in sections 2.2.1 – 2.2.4. The inner silicon tracker measures trajectories of charged particles, whereas the electromagnetic and hadron calorimeters measure the energy of particles by ideally absorbing their complete shower. Muons tracks are measured outside the superconducting solenoid in the muon system. Taken from Ref. [95].

In order to describe the location within the CMS detector, a suitable coordinate system is chosen with its origin in the interaction point. With the *x*-axis pointing towards the center of the LHC ring and the *y*-axis pointing upwards, the *z*-axis points counterclockwise along the beam axis in order to form a right-handed coordinate system. Due to the symmetric shape of the proton beams and pipes, the collision products are also expected to be symmetrical in the azimuthal angle *ϕ*, which is measured in the *x*-*y* plane starting from the *x*-axis. A cylindrical coordinate system is then usually used consisting of *ϕ*, the polar angle *θ* to the *z*-axis and the distance to the origin *r*.

Instead of the polar angle *θ*, the pseudorapidity

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) = \frac{1}{2}\ln\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right) \tag{2.3}$$

is used, where the first equation is the definition of the pseudorapidity and explicitly shows the dependence on *θ*, and the second equation uses the introduced *z*-axis along the beam axis and the three-momentum vector $\vec{p}$. This definition is convenient, since differences in pseudorapidity are invariant under Lorentz boosts along the *z*-axis and for massless particles it converges to the rapidity *y* known from special relativity. The pseudorapidity ranges from $-\infty$ to $+\infty$ with values of $\eta = 0$ being perpendicular to the beam axis.

The behavior under Lorentz boosts is exploited to define the angular separation of two particles by

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.4}$$

Due to the symmetry in *ϕ*, the three-momentum of collision results is assumed to be zero and the transverse momentum p_{T} is introduced as the component of the three-momentum perpendicular to the beam axis. The deviation of the vector sum of all collision products from zero is called missing transverse momentum and results from undetected particles such as neutrinos, which mostly transverse the detector undetected:

$$p_{\mathrm{T}}^{\mathrm{miss}} = |-\sum_{\substack{\text{detected} \\ \text{particles } i}} \vec{p}_{\mathrm{T},i}|. \tag{2.5}$$

To conclude the definition of used variables, the scalar sum of transverse momentum is introduced as

$$H_{\mathrm{T}} = \sum_{\substack{\text{detected} \\ \text{particles } i}} |\vec{p}_{\mathrm{T},i}|. \tag{2.6}$$

2.2.1. Silicon Tracker

The innermost part of the CMS detector is occupied by the silicon tracker [96] detecting hits of charged particles. While traversing the detector, charged particles create electron-hole pairs, which induce an electric signal that is then read out by special read-out chips. The track of charged particles is bent in the 3.8 T magnetic field of the solenoid magnet. This allows to determine the magnitude of the momentum and charge of the particle.
Due to the high flux of particles close to the interaction, a dedicated cooling system keeps the temperature at $-20°C$ in order to mitigate damage from radiation. Although the increase in lifetime, radiation is still a challenge in the high luminosity environment of the CMS detector, such that the pixel detector has been upgraded in the year-end technical stop in 2016/2017 [97].

The silicon tracker covers a length of 5.8 m and has a diameter of 2.6 m. A schematic layout in Fig. 2.4 shows the inner layers of the tracker consisting of the pixel detector and outer layers with silicon strips. After the upgrade in 2017, the silicon pixel detector consists of four rings placed around the beam axis with radii from 2.9 to 16.0 cm and three endcaps at each side. It contains a total of 66 million pixels with a size of $100\,\mu$m $\times\,150\,\mu$m reaching a spatial resolution of $15 - 20\,\mu$m and coverage of $\eta = \pm 2.5$. The silicon strip detector is situated at the outer part of the tracker and is made out of four subsystems: the Tracker Inner Barrel (TIB), Tracker Inner Disks (TID), Tracker Outer Barrel (TOB), and two Tracker End Caps (TEC).
The momentum resolution of the tracker is around $1 - 2\,\%$ for high momentum tracks up to $|\eta| \approx 1.6$ after which it degrades up to around 7 % at $|\eta| = 2.5$. Lower momentum tracks show a resolution of 0.5 % up to 2 %.
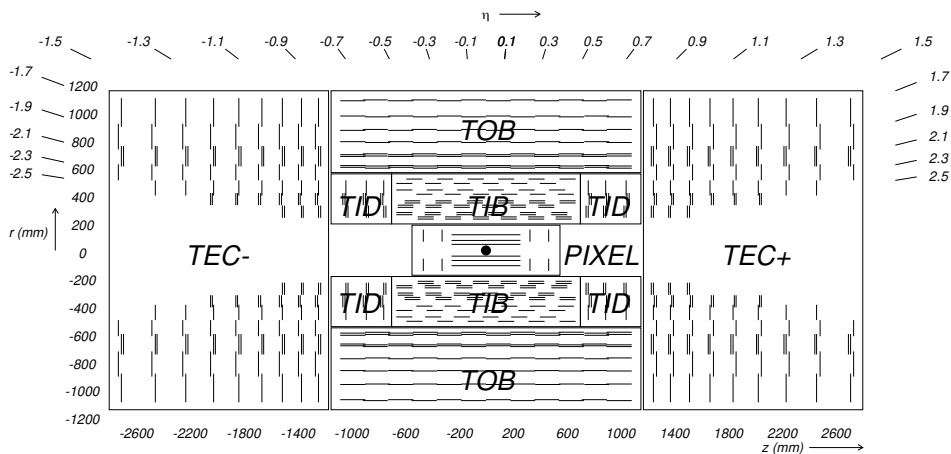

Figure 2.4.: Schematic cross section through the CMS tracker showing the pixel detector near the interaction point and the silicon strip detector with its subsystems. After the upgrade in 2017, the pixel detector consists of four instead of three layers shown here. Taken from Ref. [7]

2.2.2. Electromagnetic Calorimeter

The purpose of the electromagnetic calorimeter (ECAL) [98, 99] is to measure the energy of electromagnetically interacting particles. Other than the tracker, this is also possible for electrically neutral particles and is done by ideally absorbing the particle and all its decay products.

Depending on their energy, electrons and positrons emit photons via bremsstrahlung for high energies when accelerated and deposit their energy in the detector through ionization and thermal excitation for lower energies. The interaction of photons for high energies is dominated by pair production and for low energies they deposit their energy via Compton scattering and through the photoelectric effect. This leads to a chain reaction resulting in an electromagnetic shower containing photons, electrons and positrons. Ideally, the whole electromagnetic shower is absorbed in the calorimeter, which requires it to have a radial length of multiple radiation lengths, the material-specific distance after which the energy is reduced to $1/e$ of its original energy.

The layout of the CMS ECAL is shown in Fig. 2.5. It consists of $76\,000$ lead-tungstate (PbWO$_4$) crystals with a radiation length of $X_0 = 0.89$ cm and a Molière radius of $R_M = 2.2$ cm describing the horizontal spread. Crystals in the ECAL are arranged next to each other with a length of 23 cm corresponding to 25.8 X_0. The width of the crystals is 22×22 mm^2 at the inner part of the calorimeter up to 26×26 mm^2 at the outer part. They are used as absorber and scintillator with a fast response time and read out via avalanche photodiodes in the barrel region ($|\eta| < 1.4$) and vacuum photodiodes in the endcap ($1.4 < |\eta| < 3.0$) regions.

The energy resolution of the ECAL is described by the following equation

$$\frac{\sigma(E)}{E} = \sqrt{\left(\frac{S}{\sqrt{E/\text{GeV}}}\right)^2 + \left(\frac{N}{E/\text{GeV}}\right)^2 + C^2}\,. \tag{2.7}$$

where S describes the stochastic term, N the noise term, and C the constant term. The stochastic term includes statistical fluctuations in the width of the shower and energy deposited in front of the ECAL. The constant term originates from non-uniformity in longitudinal direction, calibration errors, and leaked energy at the back of the crystal. Finally, the noise term consists of contributions from electronic, digitization noise, and from multiple simultaneous interactions.

Using electron test beams, the numerical values have been measured to be $S = 2.8\,\%$, $N = 12\,\%$ and $C = 0.3\,\%$, but ultimately depend on the exact conditions during data taking.
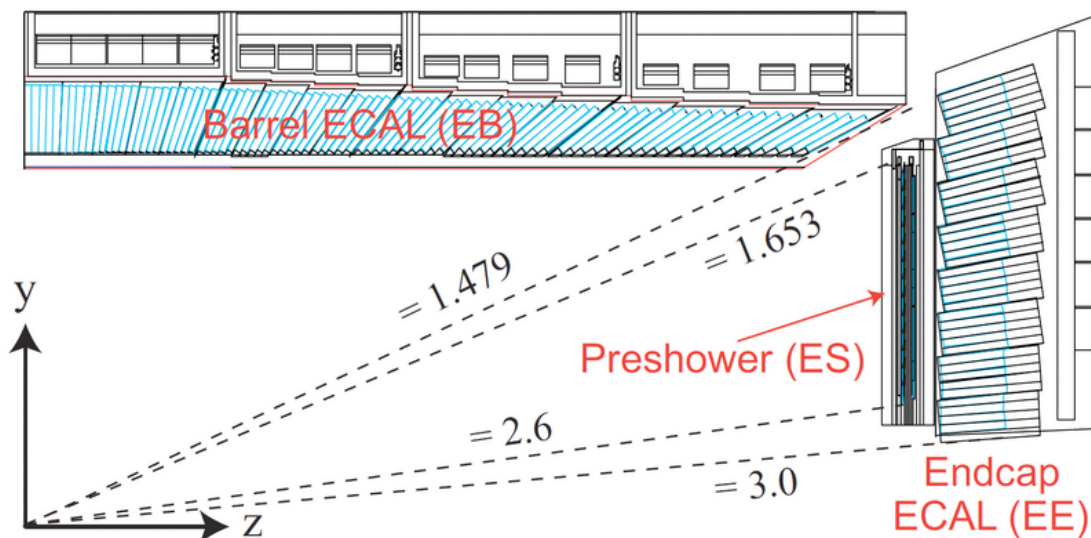
Figure 2.5.: The layout of the CMS ECAL showing the ECAL Barrel (EB) and ECAL Endcap (EE) with their respective ranges in η. Taken from Ref. [99].

2.2.3. Hadron Calorimeter

The hadron calorimeter (HCAL) [100–102] measures similarly to the ECAL the energy of particles and is specialized on particles not stopped previously but also catches tails of electromagnetic shower. Particles entering the HCAL produce a hadron shower through a cascade of inelastic collisions via strong interaction with the detector material. In case of charged hadrons, this measurement is complementary to track measurements, whereas this is the only way to measure the energy of neutral hadrons. Additionally, an electromagnetic shower can appear through sequential decay, e.g., a neutral pion in the hadronic decay chain produces two photons. Consequently, a hadron shower contains a hadronic and an electromagnetic shower fraction and is in general longer. The layout of the CMS HCAL is shown in Fig. 2.6. Unlike the ECAL, the HCAL is a sampling calorimeter consisting of alternating layers of brass absorber plates and plastic scintillators as sampling layers. The brass absorber layers have a density of $8.83\,\mathrm{g/cm^3}$, an electromagnetic radiation length of $X_0 = 1.5\,\mathrm{cm}$, and a hadronic interaction length of $\lambda_I = 16.4\,\mathrm{cm}$. It encloses the ECAL and is divided into the HCAL Barrel (HB) and HCAL Endcap (HE), as well as the HCAL Outer (HO) outside the superconducting solenoid and the HCAL Forward (HF). The HB covers a range of $|\eta| < 1.39$, which is extended by the HE to $|\eta| < 3.0$. While the HB and HE work in conjunction with the ECAL, the HO measures the energy of high-p_T particles with showers not fully inside the ECAL and HB. The HF is a Cherenkov detector specifically designed for the strong particle flux in the forward direction and covers the region of $3.0 < |\eta| < 5.0$.
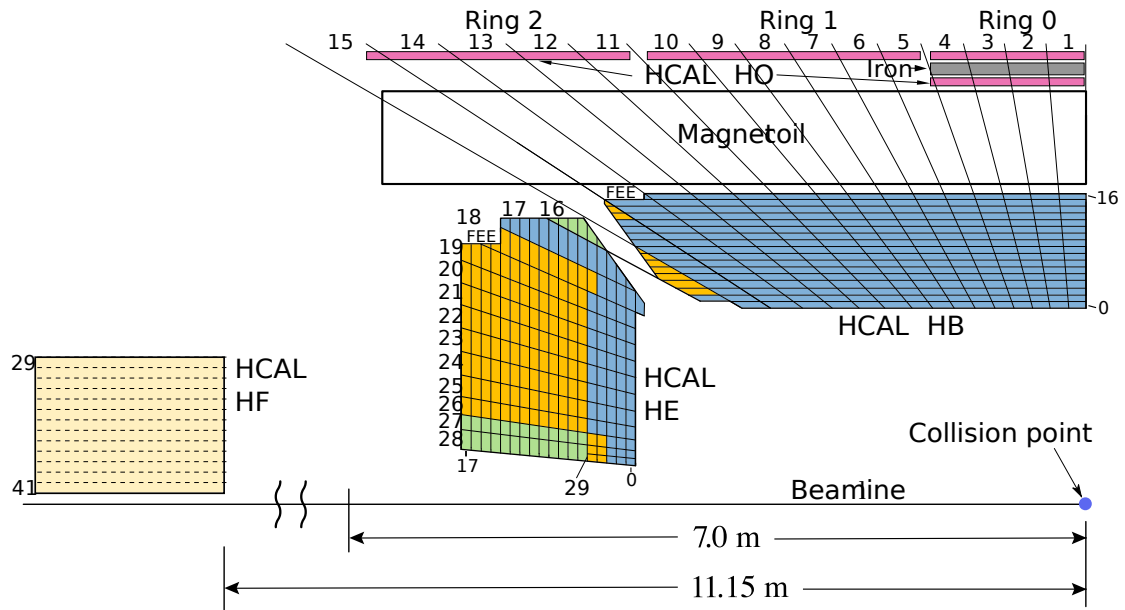
Figure 2.6.: The layout of the CMS HCAL with subsystems. The HCAL Barrel (HB) and HCAL Endcap (HE) work in conjunction with the ECAL, the HCAL Outer (HO) catches tails of the hadronic shower, and the HCAL Forward (HF) extends the range. Taken from Ref. [100].

2.2.4. Muon System

The muon system [103] is the outermost layer of the detector located completely outside the magnetic solenoid. Because the mass of muons is roughly 200 times the mass of electrons, they are less prone to bremsstrahlung, overall interact less with the detector material and can be precisely reconstructed. They play an important role in analyses of the Higgs boson, which can decay via two Z bosons into four muons. Many BSM theories like SUSY can be probed with muons, such that the CMS detector although ultimately a multipurpose detector, is built to have excellent muon reconstruction.

Alternated with the return yoke, there are four muon stations in the muon system giving information that is combined with a Kalman filter algorithm to reconstruct muons. Information from the muon system is also combined with tracks from the inner tracker resulting in so-called global muons. The muon system contains three different kinds of gaseous detectors: drift tubes (DT) with precise spatial resolution, cathode strip chambers (CSC) with both good precision and fast response, and resistive plate chambers (RPC) giving a quick measurement of the muon momentum used for immediate trigger decisions.

In the barrel and overlap region up to $|\eta| < 1.2$, DTs are installed and CSCs are used in the overlap and endcap region covering $0.9 < |\eta| < 2.5$. CSCs are placed in the region with $|\eta| < 1.8$. A schematic overview of the muon system is shown in Fig. 2.7.
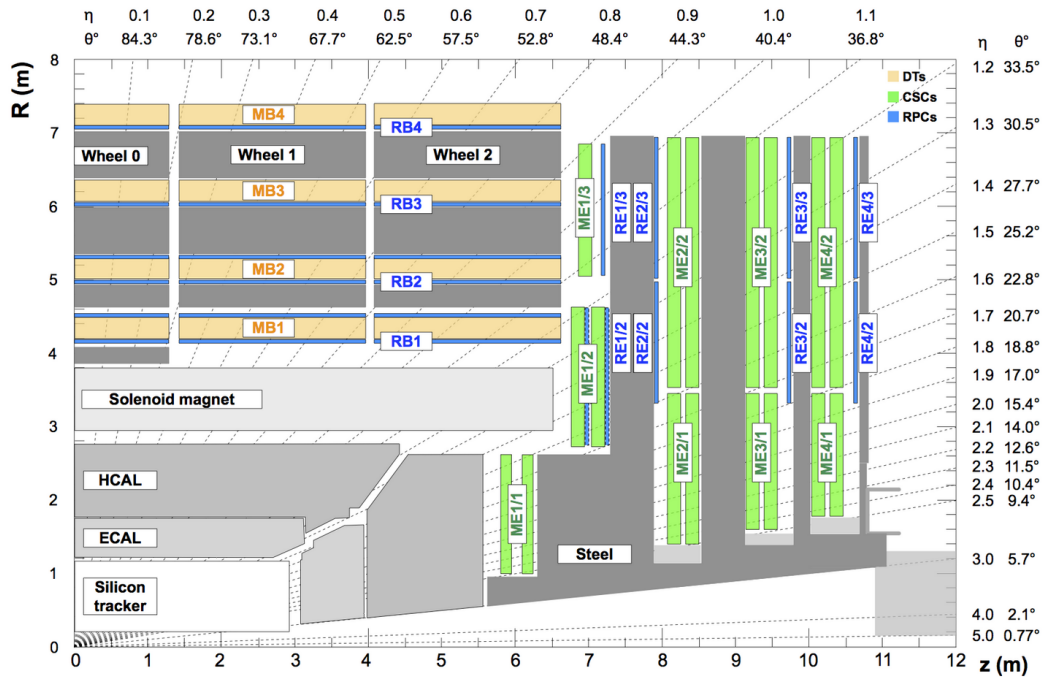
Figure 2.7.: The layout of the CMS muon system. Shown are the locations of the individual subdetectors: Muon Barrel (MB) are the DTs in the barrel and Muon Endcap (ME) the CSCs in the endcap region; RPCs Barrel (RB) and RPCs Endcap (RE) are the locations of RPCs in the respective region. Taken from Ref. [104].

2.2.5. The CMS Trigger System

In order to reduce the amount of data stored from collisions, the CMS detector features a two-level trigger system: the Level 1 (L1) trigger is a hardware based trigger to reduce the rate from bunch crossings every 25 ns or 40 MHz to 100 kHz followed by the High Level Trigger (HLT) reducing it further to 1 kHz during Run 2.

The L1 trigger quickly decides if an event contains physical phenomena interesting for further analysis. It has three major subsystems, using information from the calorimeters, the muon system, and global information, respectively. The L1 calorimeter trigger reads out energies from single calorimeter cells, combines them in a regional calorimeter trigger, which then forwards candidates for electrons, photons, taus, and jets to the global L1 trigger. The L1 muon trigger collects track information from muon subdetectors, combines them and determines the sign of the muon, and finally forwards it to the global L1 trigger. Lastly, the global L1 trigger takes information from the calorimeter and the muon system, synchronizes their arrival time and decides whether to keep or discard an event.

The HLT uses events passing the L1 trigger starting from so-called L1 seeds and then includes more detailed information, e.g., from the inner tracker. Because the HLT is

software-based and programmable, multiple HLT paths are designed and constantly improved. Therefore, complex algorithms with fast application time including artificial neural networks are used.

2.2.6. Computing

Data collected by the CMS detector and other experiments at the LHC amount to a considerable size and require a dedicated computing network, the Worldwide LHC Computig Grid (WLCG) [105, 106]. It is organized in different tiers and is distributed across all continents. Events passing one or more HLT paths are stored in a raw event format at the single Tier-0 data center at CERN in Geneva. Here, raw data is archived and split into datasets based on the passed HLT trigger paths, which are then distributed to thirteen large Tier-1 data centers - one of them at KIT in Karlsruhe - for further data-intensive analysis tasks such as detailed reconstruction, calibration, and further slimming of datasets. Many Tier-2 centers located at research centers and universities provide computing resources for tasks specific to analyses. Finally, Tier-3 centers are local clusters, typically for a small group of users.

Distribution of data is nowadays done automatically by providing certain rules concerning location and redundancy [107]. In general, data used by many analyses and raw data is stored redundantly, whereas files specific to analyses, which can be reproduced within a reasonable time are stored only once at defined locations for faster access.
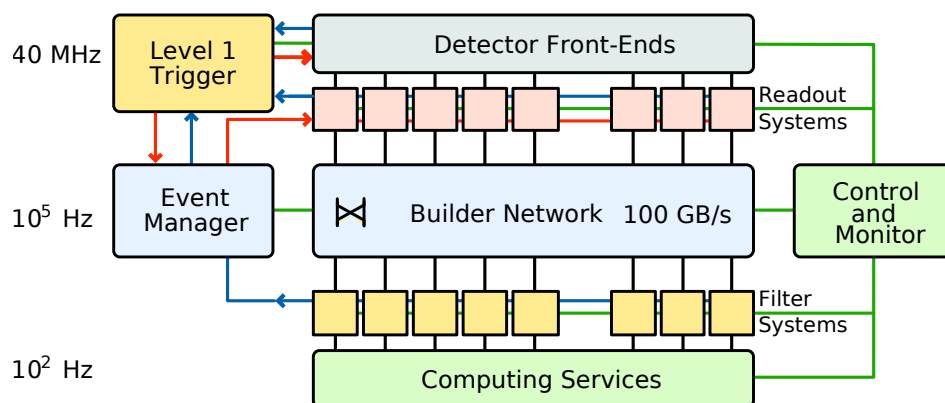

Figure 2.8.: Schematic overview of the CMS data acquisition and trigger system. Data taken at a rate of 40 MHz is reduced to 100 kHz by the L1 trigger and further reduced to 10^3 Hz (10^2 Hz prior to 2016) before writing the data to disk. Taken from Ref. [7].

3. Simulation to Data Comparison

In order to compare theoretical predictions to experimental data, it is necessary to bring predictions given by the theory into a form that is comparable with measured events. In particle physics, this is done using Monte Carlo (MC) event generators simulating events and then reconstructing the output in the same way as is done for data. The general structure for analyses of data taken with the CMS detector is as follows:

- **Data and Simulation:**
  Collisions produced by the LHC are measured in the individual subdetector parts of the CMS detector as described in chapter 2. Simulation follows a similar path starting from the scattering of protons, which ultimately produce stable particles measured by the CMS detector. Specific conditions of the collisions and interactions with the detector are included in the simulation.

- **Object Reconstruction:**
  Starting from hits in the detector, reconstruction algorithms are used to form so-called *physics objects*, such as muons, electrons, or photons. Selection criteria ensure good reconstruction quality and correction factors are applied to account for differences in efficiency, scale, and resolution between data and simulation.

- **Event Selection:**
  Using reconstructed physics objects, a specific phase space is defined to optimize the significance of the analysis. Multiple non-overlapping regions in phase space can be used to derive or cross-check background processes.

- **Background Estimation:**
  Contributions from background processes are estimated either from simulation or data in a control region. Prior to fixing the analysis strategy, data taken with the CMS detector is only used in a region where no signal is expected or which is not used in the limit extraction.

- **Statistical Analysis:**
  Finally, the statistical analysis of signal and background contributions gives a quantitative result of desired quantities. In the simplest case, this can be an analysis of the number of signal and background events. Depending on the analysis, estimation of background and signal extraction are done simultaneously.

Event simulation and event reconstruction are similar for different analyses with the CMS detector and are described in sections 3.1 and 3.2. The third section of this chapter, section 3.3, describes basic principles of the statistical methods applied in the analysis. Since event selection and background estimation are very specific to the analysis under consideration, they are part of the next chapter.

3.1. Event Simulation

Simulation of events at the LHC is done in multiple steps using MC techniques. The single steps are illustrated in Fig. 3.1, starting from two protons similar to collisions at the LHC. Since protons are composite particles, their inner structure has to be taken into account at collision energies reached by the LHC. Then, the hard scattering is simulated followed by the parton shower, hadronization, and decay into particles stable enough to be detected by the CMS detector. Finally, the interaction with the individual detector parts is simulated. Fig. 3.1 shows the individual steps during event simulation starting from Parton Distribution Functions (PDFs) up to and including the decay of color-neutral particles before their interaction with the detector.
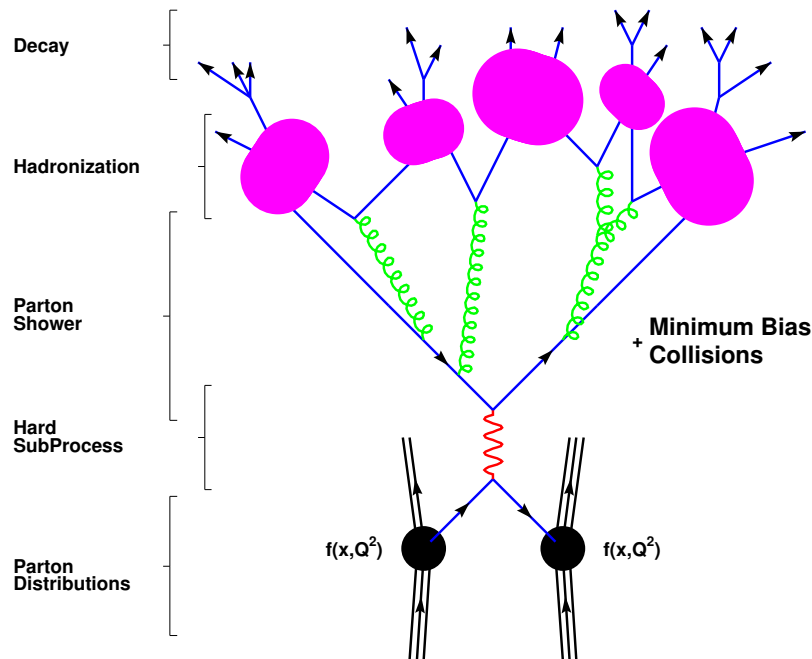
Figure 3.1.: Illustration of the individual steps during event simulation. The structure of incoming protons is described by the PDFs $f(x, Q^2)$ used to sample partons from incoming protons, which initiate the hard scattering process. This is followed by the parton shower, hadronization to color neutral particles, and the subsequent decay to stable particles. Not shown is the simulation of the detector. This figure is taken from Ref. [108].

3.1.1. Parton Distribution Functions

Since the proton is not a fundamental particle but a hadron, i.e., a bound state consisting of three quarks held together by the strong interaction, its structure becomes relevant at energies used in collisions at the LHC. The partonic structure of the proton is described by PDFs giving the probability for a given constituent to carry a specific momentum fraction x.

PDFs cannot be calculated from first principles and must be experimentally extracted from data for a given energy scale μ_F, since perturbation theory cannot be applied in this coupling regime. There are however evolution equations following from splitting functions that describe the radiation of a quark or gluon from another quark or gluon. They are called Dokshitzer–Gribov–Lipatov–Altarelli–Parisi [109–111] (DGLAP) equations and allow to extrapolate PDFs derived at a given energy scale to higher energies. They introduce a systematic uncertainty which is one of the major sources of uncertainty at hadron colliders. Many different collaborations calculate sets of PDFs, such as NNPDF [112], which was used in this study and applies artificial neural networks on data from fixed-target experiments, deep inelastic scattering, and data taken at the LHC and the Tevatron to derive the PDF.

Figure 3.2 shows the NNPDF3.0 set for two values of μ_F^2 with a major part of the momentum carried by the valence quarks.
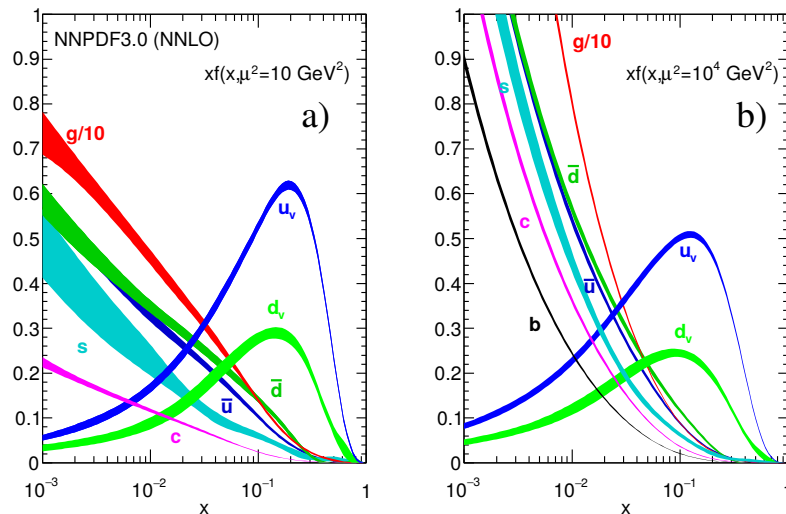

Figure 3.2.: The proton PDFs as given by the NNPDF3.0 set for different values of the factorization scale μ_F. For small momentum fractions x, gluons are dominant, whereas for higher values, the relevance of valence quarks, here up and down quarks, increases. Comparing smaller to higher values of μ_F, up and down quarks carry less momentum and gluons and quark antiquark pairs more. This figure is taken from Ref. [113].

3.1.2. Hard Scattering Process

The hard scattering process is responsible for the main characteristics of an event and is defined as the scattering of partons with the highest momentum transfer in the event. With the probability for a specific parton given by the PDFs, the total proton-proton cross section for the production of a particle X can be written as

$$\sigma_{\text{pp}\to X} = \sum_{i,j \in \{q,g\}} \int \mathrm{d}x_1 \int \mathrm{d}x_2 \cdot f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \cdot \hat{\sigma}_{ij\to X}(x_1 P_1, x_2 P_2), \qquad (3.1)$$

where $f_{i,j}$ are PDFs for the incoming particles, $x_{1,2}$ their momentum fraction of proton momentum $P_{1,2}$, and $\hat{\sigma}_{ij\to X}$ is the differential cross section for the production of particle X from particles i, j. The quantity μ_F is called factorization scale and is the energy scale, below which all constituents are absorbed into the PDFs and the cross section $\hat{\sigma}_{ij\to X}$ is evaluated. It is specific to the generation of events and should be greater than Λ_{QCD} such that perturbation theory can be applied to calculate $\hat{\sigma}_{ij\to X}$.

The differential cross section $\hat{\sigma}_{ij\to X}$ is calculated from the matrix element $|\mathcal{M}_{\text{ij}\to X}|^2$ by

$$\mathrm{d}\hat{\sigma}_{ij\to X} = \frac{|\mathcal{M}_{ij\to X}|^2}{4\sqrt{(p_i \cdot p_j)^2 - m_i^2 m_j^2}} \times \mathrm{d}\Phi, \qquad (3.2)$$

where p_i is the four-momentum of particles i with mass m_i, $\mathrm{d}\Phi$ is the Lorentz-invariant phase-space for the process under consideration, and the matrix element $\mathcal{M}_{\text{ij}\to X}$ is calculated in perturbation theory. A useful way to calculate the matrix element is to write all contributing Feynman diagrams for a given order in α_s and α_{EW} determined by the interaction vertices and their assigned coupling order, which is directly derived from the Lagrangian. The sum of these Feynman diagrams results then in the matrix element, where the interference terms arise from taking the absolute squared of the sum.

In addition to the dependence on the factorization scale μ_F through the PDFs and phase space, the cross section depends on the renormalization scale μ_R in order to deal with ultraviolet divergences due to large momenta in diagrams with loops. The coupling strengths α_s and α_{EW} are then calculated from the bare quantities in the Lagrangian depending on μ_R.

3.1.3. Parton Shower

The hard scattering process is calculated up to a defined order in perturbation theory with the leading order (LO) minimizing both contributions from α_s and α_{EW}. Higher orders in QCD (α_s) or EWK (α_{EW}) have to be explicitly included in the calculation. Furthermore, perturbation theory is only valid below a certain energy scale. In order to include gluon and photon radiation not considered in the hard process, a parton shower (PS) approach is used to add initial state radiation (ISR) and final state radiation (FSR). The evolution of the PS is done with MC techniques, where a random number is chosen and compared to a survival probability [114, 115], which gives the probability for a particle to not undergo radiation. In the case of FSR, the probability is given by the Sudakov form factors [116], whereas for the ISR, PDFs have to be included [117]. In order to include the precision of event simulation, higher-order radiations can be included in the calculation of the hard scattering process, e.g., by generating events with zero, one, and two additional quarks or gluons before the PS. Matching schemes such as MLM matching [118] are then used to pass these events to the PS simulation.

3.1.4. Hadronization

Since color-charged states form bound states due to confinement in QCD, the transformation into colorless hadrons is simulated in the hadronization process. This is not understood from first principle but relies on phenomenological models. One of such models is the Lund string model [119], in which color-charged particles are connected by a field that contains increasing energy for larger distances. If the distance grows too large, the energy allows to iteratively create new color-neutral states.
Finally, short-lived particles, i.e., particles with a lifetime smaller than a given value, decay according to their branching ratio.

3.1.5. Underlying Event

In addition to the hard interaction, further interactions occur during a single collision of two proton beams, which are summed up as underlying event (UE). This includes ISR and FSR as described above, as well as the interaction of beam remnants and multi parton interaction (MPI). Since the hard process is defined as the collision of partons with the highest momentum transfer, MPIs are sub-dominant. However, they can be of similar order, which is then called double parton scattering (DPS) and is also studied at the LHC [120]. In order to include these effects, parameters in the event generator are set to specific values, summarized in event tunes such as the CP5 tune [121] used for simulation in this thesis.
Lastly, an important part of UE is the contribution from pileup (PU), which is included in the simulation but heavily depends on the specific conditions during collisions. The next section will discuss PU in more detail.

3.1.6. Pileup

At the LHC, protons are collided in bunches resulting in a high instantaneous luminosity. This results in an increased probability of seeing an interesting event but also in an increased number of tracks and vertices at the same time. The total number of proton-proton interactions per bunch crossing is called PU. Since this is present in data, it has to be included in the simulation of events as well.

In data, pileup is measured exploiting the formula

$$\mu^i = L^i_{inst} \cdot \frac{\sigma_{inel}}{f_{rev}} \, , \qquad (3.3)$$

where μ^i is the mean number of proton-proton collisions in interval i of approximately 23 seconds [93], L_{inst} the instantaneous luminosity during that time, $f_{rev} = 11246\,\text{Hz}$ the revolution frequency of the LHC, and σ_{inel} the fiducial proton-proton cross section given by

$$\sigma_{inel} = 69.2 \pm 3.2\,\text{mb} \qquad (3.4)$$

in accordance with measurements by the CMS, ATLAS, and TOTEM [122–124] Collaborations. Extrapolation of this cross-section to the total inelastic phase space by the ATLAS and TOTEM Collaborations showed agreement with the theoretical prediction of about 80 mb. In simulation, the number of true interactions is then drawn from a distribution following the measurement. Figure 3.3 shows the measured pileup distribution during the CMS Run 2.
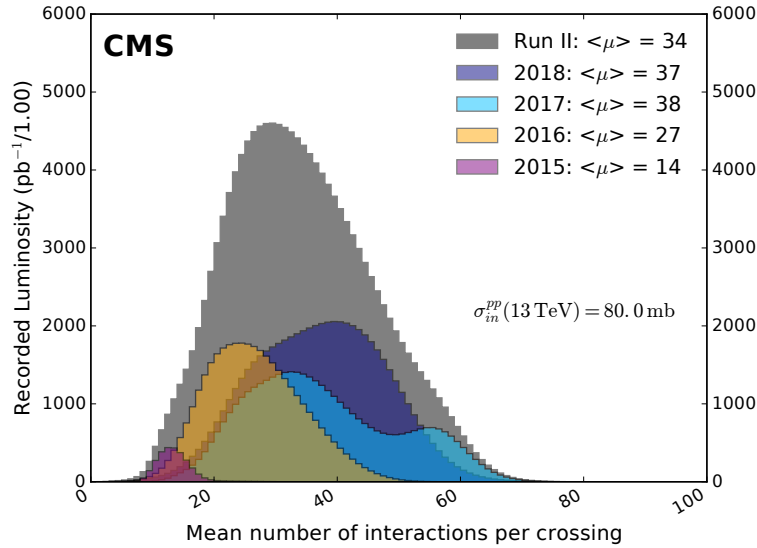
Figure 3.3.: Shown is the number of interactions per bunch crossing (pileup) taken by the CMS Experiment during 2015-2018. This figure is taken from Ref. [93].

3.1.7. Detector Simulation

After all particles have been simulated, the decay chain is followed until resulting particles have decay times long enough to leave signals in the detector. Before reconstruction algorithms are applied to build physics objects from detector signals, the interaction of the specific detector under consideration is included.

A description of the complete detector geometry is included and simulated in the GEANT4 [125] program. The response of each subdetector and electronic readout is simulated as detailed as possible. Since the full simulation of the detector takes up a major part of the computing time, a light-weight simulation of the detector response can be simulated with DELPHES [126], which applies scale and smearing factors to reconstructed quantities before detector simulation. This is useful for studies of possible detector geometries as is the case for studies of the so-called High-Luminosity LHC, the upgraded configuration after the currently running data taking period.

Samples used in the analysis of fully-hadronic VBS undergo the full detector simulation with GEANT4.

3.1.8. Software Tools

Since there is not an unambiguous way to implement the event generation in a software program, there are multiple different event generators. Different event formats to describe the event content at defined stages of event generation are defined as interfaces in order to run multiple programs in one chain.

FeynRules

In order to start from a Lagrangian, FEYNRULES [127] as a MATHEMATICA package generates a specific model in the universal FEYNRULES output [128] (UFO). This output format serves as interface to event generators for hard scattering.

Specifically, in the case of EFT, the desired Lagrangian is implemented in FEYNRULES via the SMEFTsim package [13, 78].

MadGraph5_aMC@NLO

The hard scattering is generated with MADGRAPH5_AMC@NLO [129], a program to generate events at tree-level and next-to-leading order in QCD. It is the combination of MADGRAPH5 [130] and AMC@NLO [131]. As an input, MADGRAPH5_AMC@NLO takes a UFO model and process definitions in the form of textfiles. The SM is implemented in a UFO model in NLO, whereas the SMEFTsim UFO model is in LO. When generating samples with weights for EFT, the desired EFT operators have to be chosen and can be reweighted to different values.

As an interface to store event information from parton-level interactions and to pass to further general-purpose generators, the Les Houches Event Files (LHEF) [132] has been defined as a standard format.

Powheg

Powheg [133–135], for Positive Weight Hardest Emmision Generator, is a MC event generator specialized to generate hard scattering processes at next-to-leading order or higher in QCD with only positive event weights. Processes are explicitly implemented and the output format is given in the LHEF format such that it can be interfaced to general purpose generators handling the parton shower.

Pythia8

Pythia8 [114] is a general purpose event generator able to completely generate events including all steps but the simulation of detector effects. Hard processes can be generated with Pythia8 but this is often done using dedicated tools such as the ones described above. The output of these dedicated tools is then interfaced to Pythia8 via the LHEF format and further simulation steps such as ISR, FSR, followed by the hadronization and decay of short-lived particles. A simulation of the UE is also added. Since not all aspects of these simulation steps can be derived from first principles, phenomenological research is included in the Pythia8 program, such as the Lund string model describing the process of hadronization. A set of parameters can be set to adjust the prediction of the event generator. Fixed sets of such parameters are called tunes, such as the CP5 tune [121], which was used in this thesis.

Geant4

The last step in the chain of event generation is to simulate the detector. Geant4 [125], for **Ge**ometry **An**d **T**racking, is a software package to simulate the interaction of particles and matter. Each component of the detector is encoded separately and the interaction of single particles is simulated such that the response of the detector is as close to the real experiment as possible.

3.2. Event Reconstruction

In collisions at hadron colliders such as the LHC, where two bunches of protons collide with each other, many particles are created. They are observed via their interaction with the detector leading to different kinds of signals. In order to figure out which particles are created, these signals are used to infer kinematic properties.

The CMS Collaboration uses the particle flow (PF) [136] algorithm to gather and combine information of the various subdetectors leading to an optimal identification of stable particles, i.e., electrons and photons, as well as muons and charged and neutral hadrons. It starts with basic *PF elements*: tracks of charged particles in the inner tracker, energy clusters in the ECAL and HCAL, and tracks in the muon system. These are then combined by recursively linking the single elements to *PF blocks*, when they are compatible with each other. Specifically, tracks are linked to entries in the calorimeters if the extrapolated track falls within or close to the energy cluster. Calorimeter entries are linked to other calorimeter entries, if the entry in the calorimeter with higher granularity, Preshower or ECAL, lies within the other energy cluster, ECAL or HCAL. Electrons are linked to photons to account for bremsstrahlung and finally tracks are linked to other tracks by association to the same secondary vertex (SV). In each step, specific quantities are used as distance measure to resolve ambiguity. Furthermore the linking of PF elements is transitive, meaning that they can be connected to the same block. In the next step, PF candidates of physics objects are iteratively derived from the PF blocks and the respective PF elements are removed from the PF block in each step. First, muons are reconstructed, then electrons and isolated photons, and lastly hadrons and non-isolated photons.

3.2.1. Charged Leptons

Muons

The first particles reconstructed and removed from the PF blocks are muons, since they have a relatively clear signature in the detector and their reconstruction is not specific to the PF algorithm. Information from both the inner tracker and the muon chambers are used to reconstruct three different types of muons: *Standalone muons* are based only on information from the muon system by fitting hits in the DT, CSC, and RPC subdetectors. *Global muons* start from standalone-muon tracks, which are then matched to tracks in the inner tracker improving the momentum resolution for high-p_T muons. Lastly, *tracker muons* start from tracks in the inner tracker, which are extrapolated to the muon system and are used for low-p_T muon candidates with insufficient hits in the muon system because of scattering in the return yoke.

Electrons

Reconstruction of electrons uses information from the tracker and calorimeters and takes into account the emission of photons via bremsstrahlung as well as the conversion of photons to pairs of electrons and positrons. Two different approaches are used: Firstly, an ECAL-based approach clustering energy of electrons and photons from bremsstrahlung in the ECAL and secondly, a tracker-based approach, where tracks from the inner tracker and energy clusters in the ECAL are matched requiring a compatible ratio of track momentum and cluster energy. Due to the interaction of electrons with the detector material, track candidates generally are fit with a large χ^2 value in the inner tracker and are therefore fit again with a Gaussian-sum filter [137] as it allows energy losses along the trajectory. Finally, a boosted-decision-tree (BDT) is trained combining the χ^2 values of both fits, lost energy along the track, number of hits in the tracker, and the distance of the extrapolated track to the closest ECAL cluster.

While the ECAL-based approach results in a good efficiency for isolated electrons, including the tracker-based approach improves the efficiency for non-isolated electrons and allows to reconstruct electrons with p_{T} as low as $2\,\mathrm{GeV}$.

3.2.2. Photons and Hadrons

The reconstruction of isolated photons follows after muon and electron reconstruction. Photons deposit energy in the ECAL similarly to electrons but are not detected in the inner tracker. They are reconstructed requiring an ECAL energy cluster that is not linked to a reconstructed track in the tracker.

In a last step, non-isolated photons, and charged and neutral hadrons are reconstructed. Energy clusters in the ECAL and HCAL that are not linked to a track are associated with photons and neutral hadrons, since they leave no tracks in the inner tracker. Within the tracker acceptance, $|\eta| < 2.4$, such entries in the ECAL are identified as photons and entries in the HCAL as neutral hadrons. Outside the tracker acceptance, where charged and neutral hadrons cannot be distinguished anymore, the energy fraction deposited in the ECAL due to hadrons without track is larger in hadronic showers and ECAL entries are not uniquely identified as photons. Instead, if ECAL and HCAL energy clusters are matched, they are identified as a charged or neutral hadron, and unlinked ECAL energy clusters as photons.

For energy clusters linked to at least one track, the track momentum is compared to the calorimeter energy and if found to be smaller, it is interpreted as additional neutral hadrons or photons. If found to be compatible, track and calorimeter measurements are fit with a χ^2-fit and identified as charged hadrons. In special cases, when the track momentum is larger than the calorimeter energy by more than three standard deviations, an additional search for global muons is started with relaxed requirements.

3.2.3. Jets

Analysis of physical processes require quantities that represent the final state particles of the hard scattering. While electrons, muons, and isolated photons are reconstructed as described above, and neutrinos traverse the detector without interaction, quarks and gluons are not detected individually due to the principles of QCD as described in section 1.1.5. They are measured via clusters of secondary particles located in the direction of the initiating particle and form a jet.

Jet Reconstruction

In order to combine the measured particles and group them together, a jet algorithm is used which defines a set of rules to map the measured particles to a set of jets with defined energy and momentum. The resulting jets should approximate the initiating particles and should be both infrared and collinear safe. Infrared safe means that the clustered jet remains the same when adding soft radiation to the initial particles, which is experimentally important due to the finite reconstruction efficiency. Collinear safe refers to the fact that clustered jets are invariant to collinear splitting of the constituents, which is important because of detector resolution.

Jet algorithms can be divided into two classes: cone algorithms looking for stable cones, where the sum of particle momenta falls on the cone axis, and sequential algorithms, where particles are iteratively combined following a distance measure. In this work, the anti-k_T algorithm [138], a sequential algorithm as implemented in the FASTJET package [139, 140] is used.

The anti-k_T, Camebridge-Aachen, and k_T algorithms rely on similar distance measures between pairs of particles:

$$d_{ij} = \min(p_{\mathrm{T}i}^{2n}, p_{\mathrm{T}j}^{2n})\frac{\Delta R_{ij}^2}{R^2}\,,$$

$$d_{i\mathrm{B}} = k_{\mathrm{T}i}\,,$$

(3.5)

with the angular distance ΔR_{ij} defined in Eq. 2.4 and $p_{\mathrm{T}i}$ the transverse momentum of particle i. The radius parameter R is an input of the algorithm defining the size of the jet, and $d_{i\mathrm{B}} = p_{\mathrm{T}i}^2$ is the distance to the beam. Values of $n = -1, 0, 1$ correspond to the anti-k_T, Camebridge-Aachen, and k_T algorithms, respectively.

In the clustering process, two particles i and j are clustered together if $d_{ij} < d_{i\mathrm{B}}$ forming a pseudojet and are subsequently treated as one particle, or if this equation is not fulfilled for at least one particle j, the pseudojet i is removed from the list of clustering candidates. This procedure is repeated until all particles are clustered into a set of jets. The anti-k_T algorithm is by construction collinear and infrared safe. A visual example of jets clustered with the anti-k_T algorithm and $R = 1$ is shown in Fig. 3.4.

Different values of the radius parameter R are used in the CMS Collaboration with the
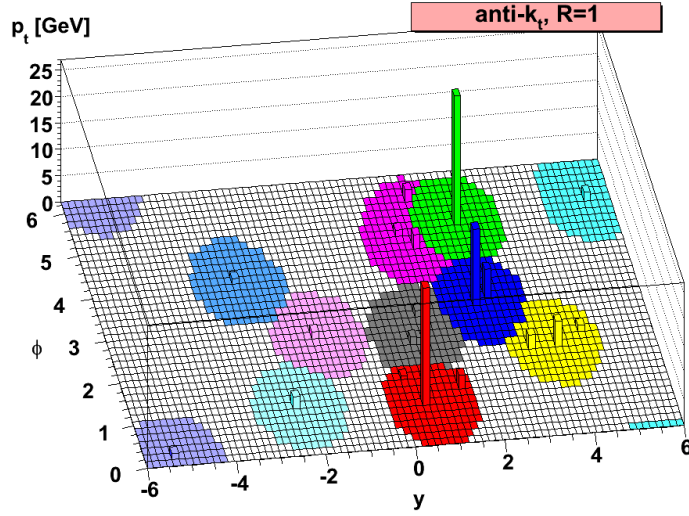
Figure 3.4.: Illustration of jets clustered with the anti-k_T algorithm with input parameter $R = 1$. The transverse momentum p_T of particles and clustered jets is shown in the $y - \phi$ plane, where y is the rapidity and ϕ the azimuthal angle. This figure is taken from Ref. [138].

most prevalent choices of $R = 0.4$ and $R = 0.8$, where the latter is used to reconstruct jets from boosted vector bosons and corresponding jets are called *fatjet*. Jet algorithms can be applied to particles reconstructed with the PF algorithm or on generator level quantities directly, i.e., after parton shower and hadronization but before the simulation of the detector.

Jet Energy Corrections

Before jets can be used in the analysis, they need to be calibrated and their scale and resolution have to be corrected for detector effects and contribution from PU in order to match the true jet energy. The correction is done in multiple steps, where in each step the four-momentum is scaled by a factor. An overview of the applied steps is shown in Fig. 3.5. In a first step, the residual contributions from PU are removed, where the corrections are derived from simulation with and without the overlay of PU. The second step is to correct for the detector response depending on the p_T and η of the jets. These corrections are also derived from simulation and correct the p_T such that the ratio of generated and reconstructed jets is on average one. In the last step, residual effects in data are corrected by comparison with QCD dijet simulation in η and with Z/γ+jets in p_T. This last step is only applied to data, after which the average scale of the jet energy is equal in data and simulation. The procedure is described in detail in Ref. [141].

In addition to the jet energy corrections described above, the jet energy resolution
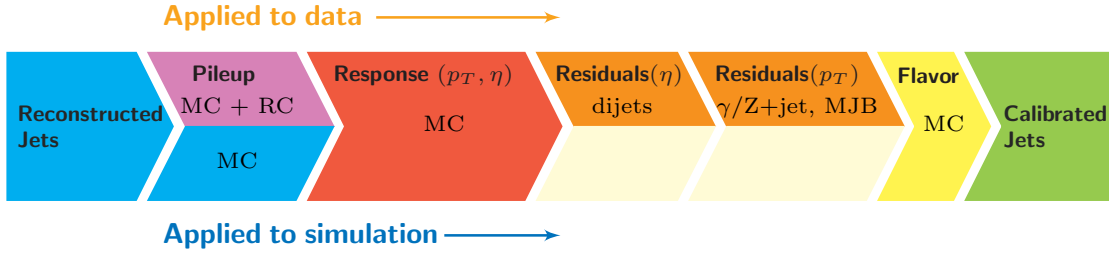
Figure 3.5.: Shown are the consecutive steps applied during the energy correction of jets as described in the text. Residual corrections are only applied to data and the simulation based flavor corrections are optional and not described here. This figure is taken from Ref. [141].

(JER) is adjusted, since it was seen in comparison of data to simulated events that the resolution in simulation is typically lower than in data [142]. For simulated jets with a well-defined matched particle-level jet, this is done by scaling the four-momentum of simulated jets depending on their p_T and a measured data-to-simulation resolution scale factor. If no matched jet is found, a stochastic smearing is applied.

Pileup Subtraction

An important factor of the performance of jet clustering algorithms is the effect of pileup, since additional contamination degrades the accurate reconstruction of jets. Several techniques have been developed to recover correct jet properties. In this work, the *Pileup per Particle Identification* (PUPPI) [143–145] algorithm is used, where a weight is assigned to each particle describing how likely it is to originate from PU contributions. This is done before jets are clustered and the weights are used to scale the four-momentum of particles effectively reducing the effect of pileup. The weight α_i for each particle i is calculated as follows:

$$\alpha_i = \log \sum_{\substack{j \neq i \\ \Delta R_{ij} < R_0}} \left(\frac{p_{\mathrm{T},j} \, / \, \mathrm{GeV}}{\Delta R_{ij}} \right)^2 , \tag{3.6}$$

where j are charged particles from the primary vertex for $|\eta_i| \leq 2.5$ and all particles for $|\eta_i| > 2.5$, and ΔR_{ij} the angular distance of particles i and j required to be smaller than a value R_0.

Jet Identification

Additional quality criteria on reconstructed jets are applied to reject jets originating from calorimetric noise and from mis-reconstructed electrons or muons. Selection criteria are based on PF candidates and result in an acceptance efficiency of real jets of about 98 – 99 %, while more than 98 % of background jets are rejected for $|\eta| < 2.7$. The efficiencies are measured in dijet events using the tag-and-probe technique, where

Table 3.1.: Additional selection criteria applied to jets reconstructed with the anti-k_T algorithm and radius parameter $R = 0.4$ or $R = 0.8$ after the PUPPI algorithm is applied.

2016-2018	$|\eta| \leq 2.4$	$2.4 \leq |\eta| \leq 2.7$	$2.7 \leq |\eta| \leq 3.0$	$3.0 \leq |\eta| \leq 5.0$
Neutral Hadron Fraction	< 0.90	< 0.98	-	-
Neutral EM Fraction	< 0.90	< 0.99	-	< 0.9
Number of Constituents	> 1	-	-	-
Muon Fraction	< 0.80	-	-	-
Charged Hadron Fraction	> 0	-	-	-
Charged Multiplicity	> 0	-	-	-
Charged EM Fraction	< 0.80	-	-	-
Number of Neutral Particles	-	-	>=1	> 2

one jet is required to pass the quality criteria ("tag") and the efficiency is determined by the number of occurrences where the second jet also passes the criteria ("probe"). The selection criteria for data taken in 2016–2018 and after application of the PUPPI algorithm are listed in table 3.1.

3.2.4. Boosted Objects Identification

Decay products of elementary particles with a large hadronic decay branching fraction, such as the W, Z, and Higgs boson, can be reconstructed as jets with an angular separation ΔR depending on the mass and p_T of the decaying particle:

$$\Delta R \sim \frac{2M}{p_\mathrm{T}} \tag{3.7}$$

Therefore, for elementary particles of a given mass M and transverse momentum p_T, jets of decay particles can no longer be resolved and are reconstructed as a single jet with larger radius parameter R. In the case of W or Z bosons and when $R = 0.8$ is chosen for reconstruction, $p_\mathrm{T} > 250\,\mathrm{GeV}$ is a conservative requirement.

Jet Grooming

Analyses that deal with highly boosted particles clustered as a large radius jet, benefit from additional jet grooming techniques. They exploit the substructure of jets from, e.g., hadronically decaying W, Z, or Higgs bosons. In this work, the *soft drop* [146] algorithm is applied to large radius jets to improve the mass-resolution. The algorithm removes soft and wide-angle radiation from effects of QCD. This is done by starting from the anti-k_T jet, which is reclustered with the Camebridge-Aachen algorithm, since

the latter allows to iteratively undo the clustering steps giving two meaningful subjets. The following requirement is imposed on these two subjets

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{cut} \left(\frac{\Delta R_{12}}{R_0} \right)^{\beta} , \qquad (3.8)$$

where z_{cut} and β are two adjustable parameters and R_0 the radius parameter used in the jet clustering algorithm. If the requirement is fulfilled, the final soft drop jet consists of the two considered subjets and otherwise, the subjet with lower p_T is removed and the algorithm further applied to the subjet with higher p_T. Within the CMS Collaboration, $z_{cut} = 0.1$ and $\beta = 0$ is used.

N-subjettiness

When heavy particles decay hadronically and their decay products are reconstructed as a single jets, their substructure can be helpful to distinguish them from QCD initiated jets. For W, Z, or Higgs bosons, the large radius jet tends to have two prongs since the bosons decay into two quarks, and for top quarks three prongs, since top quarks decay into a bottom quark and a W boson, which subsequently decays into two quarks. This tendency is exploited by the N-subjettiness [147] defined as

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min \left(\Delta R_{1,k}, \Delta R_{2,k}, \dots \Delta R_{N,k} \right) , \qquad (3.9)$$

where the sum runs over all constituents, $\Delta R_{i,k}$ is the angular distance between constituent k and subjet candidate i, and $d_0 = \sum_k p_{T,k} R_0$ is used to normalize the expression. The N-subjettiness quantifies how compatible a jet is with the hypothesis of having N prongs, where small values mean a higher degree of compatibility. It has been observed that QCD initiated jets can also have small values of τ_2 but these jets at the same time also tend to have small values of τ_1, such that the ratio $\tau_{21} = \tau_2/\tau_1$ is a good discriminator to, e.g., distinguish boosted W bosons from QCD jets.

ParticleNet Algorithm

With the improvements in the field of artificial neural networks, new techniques have been developed such as the ParticleNet [148] algorithm, which has shown better performance compared to the traditional approach of N-subjettiness [149].

The ParticleNet algorithm starts with the representation of a jet as "particle cloud", i.e., an unordered set of its constituents allowing to include arbitrary information of individual particles. Then multiple "EdgeConv" [150] blocks are used as shown in Fig. 3.6. These blocks take as input particle clouds, connect each particle with its k nearest neighbors, and update the particle cloud with a learnable function taking features of neighbors as input. Repeating these blocks retains the permutation invariant representation inherent to jet constituents.

In order to avoid changes in the spectrum of the jet mass, a mass-decorrelated version of the algorithm, ParticleNet-MD, is trained on a signal sample with a flat mass spectrum from 15 to 250 GeV. The ParticleNet-MD algorithm provides four output scores for the different decay modes of a particle X: $P(X \to b\bar{b})$, $P(X \to c\bar{c})$, $P(X \to q\bar{q})$, and $P(QCD)$.
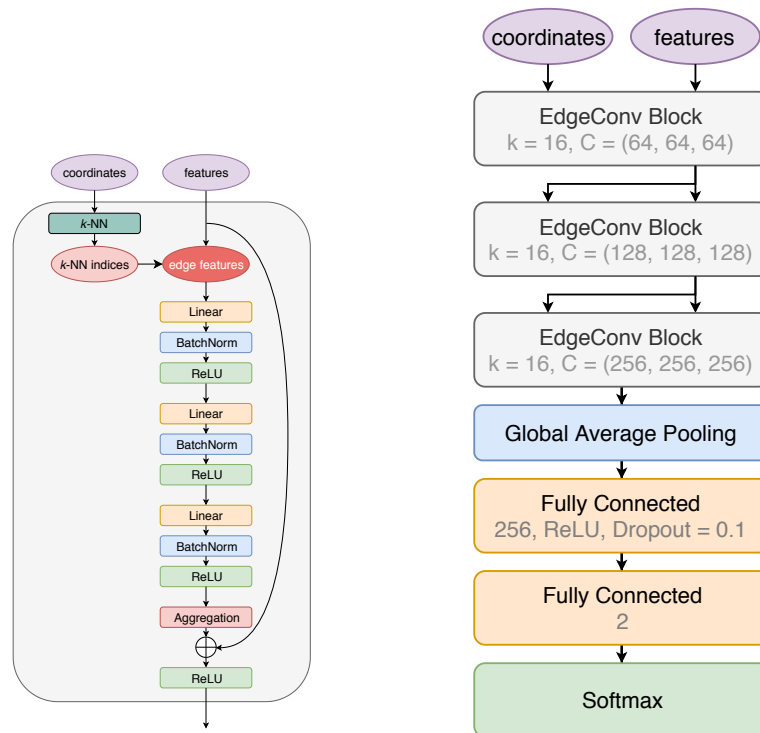
Figure 3.6.: Schematic overview of the architecture used in the ParticleNet algorithm. The EdgeConv block is shown on the left and the complete architecture of ParticleNet on the right. This figure is taken from Ref. [148].

3.3. Statistical Methods

Comparison of predictions and measurements need statistical interpretation in order to make quantitative statements. This is necessary in particle physics for multiple reasons. The SM is based on quantum mechanics and predictions are given in the form of probabilities: the result of two colliding electrons can be different each time, even under same conditions. But even in classical mechanics, where predictions are deterministic, complex systems and incomplete knowledge of the initial state are treated as sources of uncertainty. Additionally, measurements are also subject to random noise from electronics, and lastly, theoretical predictions introduce further sources of uncertainty, e.g., through calculations in perturbation theory.
At the LHC, beams of protons collide in rapid succession and many events are recorded. With the use of statistical methods, this allows to infer values of model parameters.

3.3.1. Parameter Estimation

For the estimation of model parameters, maximum likelihood fits are commonly used in high energy physics. The likelihood function $L(\vec{\theta}|\vec{x})$ is constructed for unknown parameters $\vec{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ given measurements $\vec{x} = (x_1, x_2, \ldots, x_n)$. It is similar to the probability density function f for a measurement x_i given model parameters $\vec{\theta}$. For N independent measurements, the likelihood function L is given by:

$$L(\vec{\theta} \,|\, \vec{x}) = \prod_{i=1}^{N} f(x_i \,|\, \vec{\theta}) \,, \tag{3.10}$$

$$\int L(\vec{\theta} \,|\, \vec{x}) \, \mathrm{d}x_1 \, \mathrm{d}x_2 \ldots \mathrm{d}x_N = 1 \,. \tag{3.11}$$

Estimation of model parameters is then done by maximizing L, which is called maximum likelihood estimation (MLE). At the maximum, the derivatives vanish:

$$\frac{\partial L(\hat{\vec{\theta}} \,|\, \vec{x})}{\partial \theta_j} = 0 \,, \tag{3.12}$$

for each j ranging from 1 to N. Computationally, it is more convenient to minimize the negative logarithm of the likelihood function. In the case of a counting experiment with n_{bins} bins, the likelihood function is given by a product of Poisson distributions:

$$L\left(\vec{\lambda}|\vec{k}\right) = \prod_{i}^{n_{\mathrm{bins}}} \mathrm{Poisson}\left(k_i|\lambda_i\right) = \prod_{i}^{n_{\mathrm{nbins}}} \lambda_i^{k_i} \frac{e^{-\lambda_i}}{k_i!} \,, \tag{3.13}$$

where λ_i are the expected and k_i the observed number of events in bin i.

3.3.2. Systematic Uncertainties

The expected number of events can depend on different parameters, such as introduced by a physical model ($\vec{\mu}$) or due to systematic uncertainties ($\vec{\vartheta}$). Usually, an idea of the strength of systematic uncertainties is known, such that values of the corresponding parameter further away from the nominal value are less likely. This is achieved by introducing constraint terms for nuisance parameters $\vec{\vartheta}$:

$$L\left(\vec{\lambda}|\vec{k}\right) = \prod_i^{n_{\text{bins}}} \text{Poisson}\left(k_i|\lambda_i\right) \prod_i^{n_{\text{nuis}}} p_i(\vartheta_i|\widetilde{\vartheta}_i), \tag{3.14}$$

$$\lambda_i = \lambda_i(\vec{\mu}, \vec{\vartheta}), \tag{3.15}$$

where ϑ_i is the nuisance parameter with nominal value $\widetilde{\vartheta}_i$ and p_i is the function used to constrain the parameter, which is often chosen to be a normal distribution with mean $\widetilde{\vartheta}_i$ and standard deviation of one.

3.3.3. Interval Estimation

In addition to the values maximizing the likelihood, confidence intervals (CIs) for a model parameter at a given confidence level (CL) are interesting. This is done by constructing a test statistic q_μ from the likelihood function in the case of a single model parameter μ:

$$q_\mu = -2 \ln \frac{L(\hat{\hat{\theta}}_\mu, \mu)}{L(\hat{\theta}, \hat{\mu})}, \tag{3.16}$$

where the likelihood function in the numerator is maximized given a value μ, and in the denominator, $\hat{\mu}$ is chosen to be the global maximum of the likelihood function, such that q_μ is always positive expect for numerical instabilities. Since the test statistic ultimately depends on measured data, it itself is a random variable following a distribution $f\left(q_\mu, \mu\right)$. Looking at observed data, a specific value $q_\mu^{\text{obs.}}$ can be calculated. The compatibility of a given value of the model parameter μ can then be expressed via the p-value:

$$p_\mu = \int_{q_\mu^{\text{obs}}}^{\infty} f\left(q_\mu, \mu\right) \, \mathrm{d}q_\mu. \tag{3.17}$$

This value is the probability to measure a value that is at least as incompatible with the assumption of the model parameter μ as the observed data. Often, the p-value is conventionally converted to the corresponding number of standard deviations Z under assumption of a normal distribution, even when $f\left(q_\mu, \mu\right)$ follows a different distribution:

$$Z_\mu = \Phi^{-1}(1 - p_\mu) \tag{3.18}$$

with the inverse cumulative distribution function of the normal distribution Φ^{-1}. Following Wilks' Theorem [151], the distribution of the test statistic follows a χ^2 distribution with k number of freedoms for k floating model parameters:

$$q_\mu \sim \chi_k^2, \tag{3.19}$$

$$\chi_k^2(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}, \tag{3.20}$$

with the gamma function Γ. For a given CL α, the corresponding CI is then given by all free floating model parameters μ that satisfy:

$$q_\mu \leq \int_0^\alpha \chi_k^2(x)\,\mathrm{d}x = I_\alpha. \tag{3.21}$$

In the case of one model parameter, the right-hand side of Eq. 3.21 evaluates to $I_{68\%} = 0.99 \approx 1$ for $\alpha = 68\%$, to $I_{95\%} = 3.84$ for $\alpha = 95\%$, and to $I_{95.45\%} = 4.00$ for $\alpha = 95.45\%$, where the first value of I corresponds to a deviation of one standard deviation and the last value of I to a deviation of two standard deviations compared to a normal distribution.

4. Constraints on Anomalous Couplings in the Hadronic Decay Channel of VBS

The analysis presented here is a search for anomalous gauge couplings in VBS and aims to set limits on the strength of additionally introduced operators.

The Higgs mechanism is deeply connected to the VBS process, where two vector bosons (V = W,Z) are emitted by incoming quarks and interact with each other. The unitarity of this process depends on the cancellation with diagrams involving the Higgs boson and would otherwise be broken at energy scales that are reachable today.

With higher energies, it becomes attractive to analyze VBS processes directly and thereby probe the EWK sector. In recent years, the theory and experimental community became more and more interested in the search for deviations of couplings from the theoretical prediction in a model independent way. Searches for anomalous couplings in the framework of an EFT are motivated by new physical phenomena at high energies: heavy particles that cannot be produced directly at current reachable energies are described by introducing higher order correction terms. Coefficients to additionally introduced operators can be constraint by experimental data resulting in model independent searches. In particular, anomalous triple and quartic gauge couplings (aTGC and aQGC) affecting heavy vector bosons are probed in VBS. While aTGCs are usually better constrained in diboson production due to the higher cross section of the process, aQGC is only accessible in VBS and in triboson production, which is limited by the presence of three heavy vector bosons in the final state. This analysis considers a subset of dimension-6 operators in the EFT approach leading to aTGC and aQGC, as well as a set of dimension-8 operators affecting only quartic interaction leading to aQGC.

The fully hadronic decay channel is targeted, meaning that each of the two vector bosons decays into two quarks. Compared with other channels, the cross section is larger and higher energies can be reached. A drawback of this final state is that it has low missing transverse momentum and no leptons, which results in an overwhelming background due to QCD multijet production. In order to separate signal from background contributions, the boosted topology is considered, where each vector boson is reconstructed as a single jet with a large radius of $R = 0.8$. Therefore, high transverse momentum is required such that the opening angle between the two quarks is sufficiently small. This nicely fits together with the EFT approach, since as seen in previous

analyses [152], the EFT coefficients are mainly constraint in the high energy region.

This analysis is a search for anomalous couplings in the fully hadronic decay channel of VBS using data from pp collisions at $\sqrt{s} = 13\,\text{TeV}$. Data were collected during the full 2016, 2017, and 2018 data taking period by the CMS detector at the CERN LHC and correspond to a total integrated luminosity of $137.2\,\text{fb}^{-1}$.

4.1. Analysis Strategy

The fitting strategy employed in this analysis aims to deal with the major challenge in this decay channel: the large background from QCD multijet production. Therefore, signal and background contributions are fit in three dimensions: the mass of each AK8 jet, m_{J1} and m_{J2}, and the invariant mass of the pair, m_{JJ}. This strategy is based on previous analyses searching for diboson resonances described in Ref. [153, 154] and adopted and optimized for signal contributions from an EFT model. A more detailed description of the fitting procedure is given in section 4.1.4.

4.1.1. Signal and Background Composition

The hadronic decay of both vector bosons in VBS is chosen, since this topology is expected to be highly sensitive to BSM via EFT. However, a major challenge of this decay channel is given by the absence of charged leptons and missing momentum in the final state resulting in a very large background from QCD multijet production. The second-largest background is from top quark pair production with subleading background from QCD-induced VBS, V+jets and diboson production, and the production of single top quarks. In the following, the signal and background processes are briefly described.

Signal and EWK VBS

The signal considered in this analysis is the additional contribution of VBS in the presence of EFT contributions, while the EWK SM production of VBS is considered as background. The topology of EWK VBS in the SM is described in section 1.2, while the additional EFT contribution depends on the operator under consideration. In general, it is a priori not known, which kinematic variables are most sensitive to the EFT contribution under consideration. However, it has been found that in general, the invariant mass of the diboson system, m_{JJ}, shows good if not the best sensitivity [75].

The signal generation is done in MADGRAPH5 using the SMEFTsim package as described in section 3.1.8. In particular, MC events are generated using the version SMEFTsim_U35_MwScheme_UFO_v302, where the name hints to the flavor assumption ($U(35)$), meaning that maximal flavor symmetry is assumed. Since this analysis is
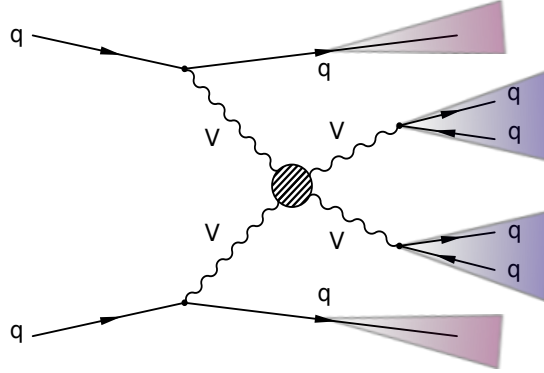
Figure 4.1.: An example Feynman diagram of the EWK VBS process. The hatched circle indicates the presence of anomalous couplings introduced by the considered operators.

not sensitive to different flavor eigenstates, which would be the case if, e.g., $Z \to b\bar{b}$ would be explicitly considered, the assumption is justified. Events are generated at a nominal value of a specific Wilson coefficient and then reweighted.

In total, three dimension-6 operators are considered, as neither of them affects any processes other than EWK VBS production. Neglecting the effect on minor backgrounds or broadening the signal definition, it is possible to expand this analysis to a larger group of operators. The considered dimension-6 operators are the following:

$$\begin{aligned}
Q_{\text{W}} &= \epsilon^{ijk} W_\mu^{i\nu} W_\nu^{j\rho} W_\rho^{k\mu} , \\
Q_{\phi\Box} &= (\phi^\dagger \phi)\Box(\phi^\dagger \phi) , \\
Q_{\phi\text{W}} &= (\phi^\dagger \phi) W_{\mu\nu}^i W^{i\mu\nu} .
\end{aligned} \tag{4.1}$$

As for dimension-8 operators, the full set of 20 operators introduced in section 1.3.2 is considered, one at a time. The operators are not a complete basis but rather a set of operators affecting the quartic vertex of VBS.

MC samples are split by combination of vector bosons to better cover the phase space and allow for potential further studies, e.g., depending on the decay of Z bosons and flavor choice of PDFs. The full list of samples is given in appendix A, where the presence of two heavy vector bosons decaying hadronically is required. The cross section of EWK VBS production in the SM is shown together with QCD-induced VBS production in table 4.1.

Figure 4.1 shows one of many Feynman diagrams of the signal process, where the circle indicates the presence of anomalous couplings through either dimension-6 or dimension-8 operators.

All samples produced with dimension-6 and dimension-8 operators including the ranges used for reweighting are listed in appendix A.1.

QCD-induced VBS

VBS processes can be produced at order $\mathcal{O}(\alpha_{\mathrm{EW}}^6 \alpha_{\mathrm{S}}^0)$ (EWK) or including the exchange of a gluon resulting at order $\mathcal{O}(\alpha_{\mathrm{EW}}^4 \alpha_{\mathrm{S}}^2)$ (QCD-induced VBS). Since the final state is the same, a single event cannot be assigned to one or the other process. However, the topology of these two processes is different and QCD-induced VBS production is treated separately as background. The interference of both processes is neglected as it is small and both processes are further separated by selection criteria on jets in the forward region. Table 4.1 shows the cross section of the EWK and QCD-induced production, as well as the cross section in the case where both processes are considered together. Figure 4.2 shows an example Feynman diagram for this background.

Table 4.1.: Summary of the cross section for EWK and QCD-inducted VBS production as well as their combined production quoted directly from MADGRAPH5. The sum of both processes agrees with the combined production within 1% and the separation is further enhanced due to selection criteria on the tagging jets.

| | cross section in pb | | |
	VBS EWK	VBS QCD	VBS EWK and QCD
$W^\pm W^\pm$	0.126	0.108	0.245
$W^\pm W^\mp$	1.893	160.738	161.005
$ZW^\pm$	0.579	5.985	6.561
ZZ	0.053	1.084	1.143
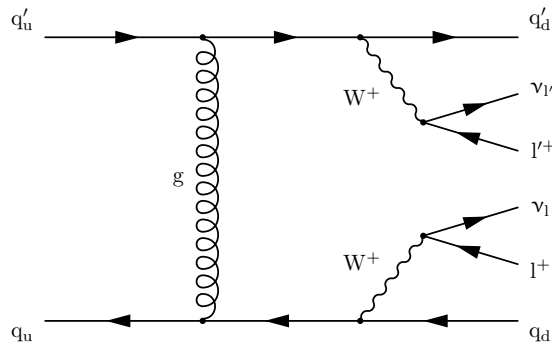
Figure 4.2.: An example diagram of the QCD-induced VBS production. The final state is the same as EWK VBS and is separated to a good degree from the signal contribution by cuts on the tagging jets. The interference of EWK and QCD-induced VBS production is small and neglected.

QCD Multijet Production

The major background in the considered region of phase space is given by QCD multijet production. Two different sets of samples are used in this analysis as listed in appendix A: one set is produced in bins of generator level p_T with PYTHIA8, which includes all hard QCD $2 \rightarrow 2$ processes. The second set of samples is generated in bins of H_T with MADGRAPH5 generating the hard scattering process of two protons to two, three, or four light-flavor quarks or gluons, and is then propagated to PYTHIA8 with MLM matching to simulate the subsequent parton shower. The cross section depends on the specific bin under consideration and is taken from MC simulation as a starting value but ultimately derived from data, which is explained in more detail in the following chapter. Figure 4.3 shows a set of example Feynman diagrams in leading order. The final state of the hard scattering process is given by gluons and light-flavor quarks and the mass of reconstructed jets is not expected to be centered around a specific value, as is the case for, e.g., VBS with a genuine vector boson in the final state. Furthermore, jets reconstructed with a large radius of R = 0.8 tend to have different substructure than jets from a heavy vector boson. These two aspects are exploited in the analysis.

Top Quark Pair Production

The second-largest background is given by top quark pair production, which is the dominant production mode for top quarks at the LHC. It is produced via the strong interaction with the leading order Feynman diagrams shown in Fig. 4.4 and is dominated by gluons in the initial state at a center-of-mass energy of $\sqrt{s} = 13$ TeV. Its cross section is expected [155, 156] to be:

$$\sigma_{t\bar{t}} = 833.9^{+20.5}_{-30.0} \,(\text{scale}) \pm 21.0 \,(\text{pdf} + \alpha_s) \text{ pb.} \tag{4.2}$$

Top quarks decay almost exclusively into a W boson and a bottom quark, which is a consequence of the CKM matrix element V_{tb}. In case of hadronically decaying W bosons, the final state consists of six quarks leading to a similar footprint in the detector as
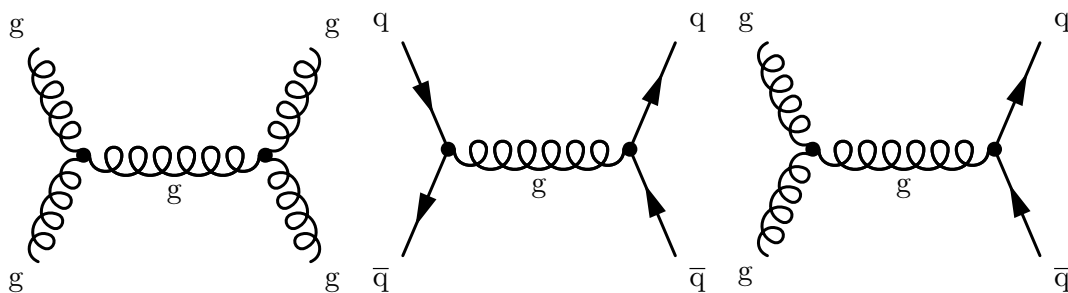
Figure 4.3.: Shown are LO Feynman diagrams for QCD multijet production, the major background in the analysis.
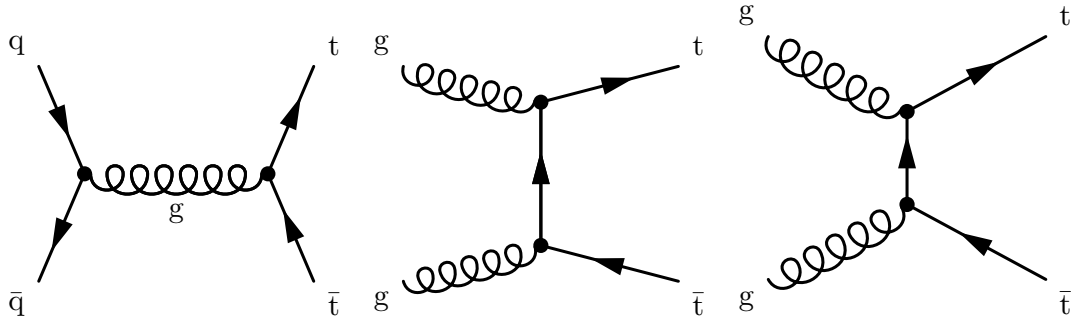
Figure 4.4.: Shown are LO Feynman diagrams for top quark pair production at the LHC. Two hadronically decaying top quarks constitute the second-largest background in the analysis.

EWK VBS production. Keeping Eq. 3.5 in mind, a single top quark can result in one large radius jet for values of $p_T \gtrsim 400\,\text{GeV}$, two jets in case the W boson results in a single jet for $p_T \gtrsim 200\,\text{GeV}$, an in three resolved jets.

V+jets Production

A further contribution to background events is due to the production of a heavy vector boson, W or Z boson, in association with additional jets, called V+jets processes. The production cross section is much smaller than for QCD multijet production but the genuine heavy vector boson leads to a non-vanishing contribution after selection criteria are applied. Both the production of W and Z bosons with additional jets is considered, where the cross section is smaller for additional jets. In order to enhance the number of simulated events entering the signal extraction, the sets of MC samples are generated considering the hadronic decay of the vector bosons and are binned in H_T. The hard scattering is generated without, with one, and with two additional jets in MADGRAPH5 and interfaced with MLM matching to PYTHIA8.

Diboson Production and Single Top Quark Production

Further backgrounds considered in this analysis are the production of two heavy vector bosons, W or Z, and the production of one top quark. Although their cross section is larger than EWK VBS production, imposed selection criteria significantly reduce the number of expected events from these processes. Diboson production is generated with PYTHIA8 and the production of single top quarks with an associated W boson is generated with POWHEG and interfaced to PYTHIA8.

4.1.2. Object Identification

Reconstruction of physics objects is based on the PF algorithm combining information from the various parts of the CMS detector as described in section 2. An overview of used objects is given in table 4.2 and a detailed description is given in the following.

Jets

Jets are reconstructed from all candidate particles using the anti-k_T algorithm with $R = 0.4$ (AK4) and $R = 0.8$ (AK8). As a baseline selection, AK8 jets are required to have $p_T > 200\,\text{GeV}$ and lie within $|\eta| < 2.4$, while the requirement for AK4 jets is $p_T > 30\,\text{GeV}$ and $|\eta| < 5.0$. For both collections of jets, the PUPPI algorithm is applied to reduce the impact of PU and additional identification criteria as summarized in table 3.1 are applied. The soft drop algorithm is used to calculate the mass of AK8 jets. Because the same candidate particles can be clustered in both jet collections, overlapping jets are removed from the list of AK4 jets if their jet axis is within $\Delta R = 1.2$ of the jet axis of an AK8 jet. In order to enhance the fraction of AK8 jets originating from heavy vector bosons and reduce the background from QCD multijet production, the mass-decorrelated version of the ParticleNet algorithm is applied.

Electrons and Muons

Electrons are identified from PF candidates with the HEEP V70 ID [157], which is specialized to select electrons with high p_T. It requires well reconstructed isolated tracks to be matched with energy deposits in the ECAL. In the barrel region, it uses arrays of energy clusters, whereas in the endcap, the lateral extension of the shower in η-direction is used. The electrons are required to have $p_T > 35\,\text{GeV}$ and $|\eta| < 2.5$. Similarly, muons start form PF candidates and are identified with the tight cut-based ID,

Table 4.2.: Overview of physics objects used in the analysis.

Collection	
Jets AK4	anti-k_T algorithm with $R = 0.4$, PUPPI
Jets AK8	anti-k_T algorithm with $R = 0.8$, PUPPI, soft drop
Jet tagger	ParticleNet mass-decorrelated: WvsQCD score
Electrons	PF algorithm, ElectronID_HEEP_V70
Muons	PF algorithm, tight cut-based ID
missing p_T	various filters, PUPPI

which selects a subset of PF muons with additional quality requirements. Additionally, a $p_T > 30\,\text{GeV}$ is required as well as $|\eta| < 2.4$.

Missing transverse momentum and vertex selection

Since protons are collided head-on, the whole system carries zero net transverse momentum. Particles that are not detected result in an imbalance in the transverse momentum and thus in missing transverse momentum (p_T^{miss}), which is calculated as negative vector sum of the p_T of all detected particles.

In order to remove mismodeling due to instrumental noise in the detector, the following noise filters are applied as recommended by the JetMET POG [158]:

- goodVertices

- globalSuperTightHalo2016Filter

- HBHENoiseFilter

- HBHENoiseIsoFilter

- EcalDeadCellTriggerPrimitiveFilter

- BadPFMuonFilter

- BadPFMuonDzFilter

- eeBadScFilter

- ecalBadCalibFilter (only for 2017 and 2018)

In addition, the presence of at least one primary vertex (PV) is requited with a maximum displacement of 24 cm in z direction and 2 cm in radial direction from the nominal proton-proton interaction. Furthermore, at least four degrees of freedom are required. In the case of multiple such vertices, the one with the highest sum of transverse momentum of all associated tracks is chosen.

4.1.3. Event Selection

With the physics objects entering the analysis defined, the next step is the definition of interesting events going into the signal extraction. The event selection aims at refining the data to a signal enriched region.

Trigger Selection

A number of HLT paths is chosen to select signal events. The triggers used in this analysis are listen in table 4.3 and target the fully hadronic final state. Due to the high cut on m_{JJ}, the efficiency is close to one and the analysis operates on the trigger plateau.

In order to study the triggers, their efficiency is estimated using the SingleMuon dataset with "HLT_IsoMu27" and "HLT_Mu50" as reference triggers. The denominator consists of the number of all events passing the reference triggers, whereas events in the numerator pass both the reference and analysis triggers for the respective period. Requirements imposed on the physics objects are looser than in the analysis and are summarized in table 4.4. In particular, there are no selection criteria on the AK4 jets and on m_{JJ} since the efficiency is estimated as a function of this variable.

Table 4.3.: Shown is a complete list of triggers used in the analysis for all years. The year 2016 is split into Run A to H, 2017 into Run A to F, and 2018 into Run A to B, depending on the machine configurations.

Period	Run	trigger name
2016	all	HLT_PFJet{400,450,500}
(pre and postVFP)		HLT_PFHT650_WideJetMJJ{900,950}DEtaJJ1p5
		HLT_AK8PFJet360_TrimMass30
		HLT_AK8PFHT700_TrimR0p1PT0p03Mass50
		HLT_PFHT900
	only B-G	HLT_PFHT800
	only H	HLT_AK8PFJet{400,450,500}
2017	all	HLT_PFHT1050, HLT_AK8PFJet500
	only C-F	HLT_AK8PFJet{360,380,400,420}_TrimMass30
	only C-F	HLT_AK8PFHT{750,800,850,900}_TrimMass50
2018	all	HLT_PFHT1050, HLT_AK8PFJet500
		HLT_AK8PFJet{360,380,400,420}_TrimMass30
		HLT_AK8PFHT{750,800,850,900}_TrimMass50

Figure 4.5 shows the resulting efficiency for all periods as a function of m_{JJ}. The efficiencies split by runs can be found in appendix B.

The analysis cut on m_{JJ} is chosen such that the overall trigger efficiency per period is $> 99\%$.

Table 4.4.: Summary of the preselection used to estimate the trigger efficiency. The applied cuts are looser than the selection criteria used in the analysis to allow for separate optimization.

	VV Selection
AK8 jets	$N \geq 2$
	$p_T \geq 200\,\text{GeV},\ |\eta| \leq 2.5$
	$m_J \geq 55\,\text{GeV}$
	$\Delta\eta < 1.3$
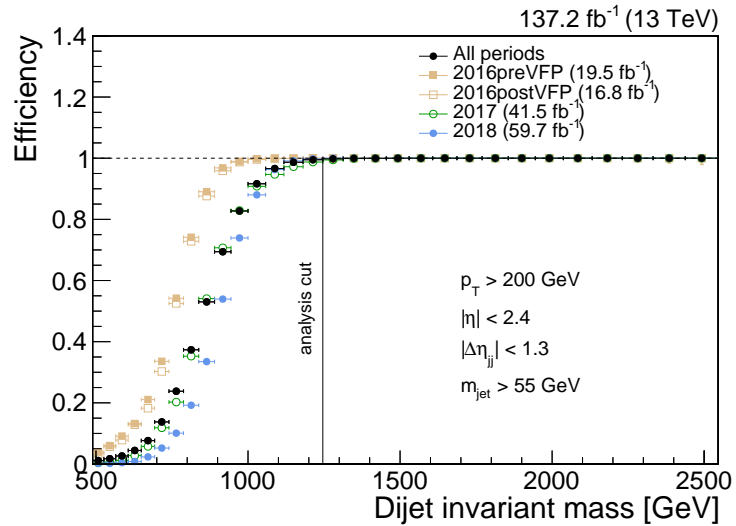
Figure 4.5.: Efficiency of all trigger per period as a function of m_{JJ}. The analysis cut is chosen such that the overall trigger efficiency per period is $> 99\%$.

Heavy object tagging

A multivariate classifier distinguishes AK8 jets originating from a heavy vector boson V from QCD induced jets.

The ParticleNet algorithm, a dynamic graph convolutional neural network, is used, which is described in section 3.2.4. Since the features used in the training and therefore also the output of the algorithm can be highly correlated to the jet masses, the mass-decorrelated version (ParticleNet-MD) is used. This is important, since the shape of the jet mass is heavily relied on in the fitting strategy for limit extraction.

ParticleNet-MD provides four different output nodes: $P(X \to b\bar{b})$ ("Xbb"), $P(X \to c\bar{c})$ ("Xcc"), $P(X \to q\bar{q})$ ("Xqq"), and $P(QCD)$ ("QCD"). To separate against QCD induced jets, they are combined in the following way:

$$\text{WvsQCD} = \frac{\text{Xcc} + \text{Xqq}}{\text{Xcc} + \text{Xqq} + \text{QCD}},$$

where in the final state, jets from b quarks are not explicitly distinguished.

Since the tagging efficiencies in data and MC simulation are different, scale factors have to be applied to simulated events. They are derived by the JetMET POG in Ref. [159] and depend on the p_T of the selected AK8 jet such that each event is assigned a SF depending on the p_T of both selected AK8 jets, the chosen working points, and the run period. Tables 4.5–4.8 list the scale factors applied to this analysis. Working points used in this analysis are defined by their mistag rate as follows: tight (0.5%), medium (1.0%), loose (2.5%). For AK8 jets with $p_T > 800\,\text{GeV}$, the scale factors are linearly extrapolated.

Table 4.5.: ParticleNet-MD scale factors for 2016preVFP.

mistag rate	$200 \leq p_T\,/\,\text{GeV} < 300$	$300 \leq p_T\,/\,\text{GeV} < 400$	$400 \leq p_T\,/\,\text{GeV} < 800$
0.5%	$0.85^{+0.03}_{-0.03}$	$0.86^{+0.04}_{-0.04}$	$0.86^{+0.08}_{-0.08}$
1.0%	$0.90^{+0.03}_{-0.03}$	$0.87^{+0.04}_{-0.04}$	$0.92^{+0.08}_{-0.07}$
2.5%	$0.90^{+0.03}_{-0.03}$	$0.94^{+0.04}_{-0.04}$	$0.94^{+0.07}_{-0.07}$

Table 4.6.: ParticleNet-MD scale factors for 2016postVFP.

mistag rate	$200 \leq p_T\,/\,\text{GeV} < 300$	$300 \leq p_T\,/\,\text{GeV} < 400$	$400 \leq p_T\,/\,\text{GeV} < 800$
0.5%	$0.86^{+0.05}_{-0.04}$	$0.83^{+0.04}_{-0.04}$	$0.69^{+0.07}_{-0.06}$
1.0%	$0.89^{+0.04}_{-0.04}$	$0.86^{+0.04}_{-0.04}$	$0.73^{+0.07}_{-0.07}$
2.5%	$0.95^{+0.05}_{-0.06}$	$0.91^{+0.04}_{-0.04}$	$0.84^{+0.07}_{-0.07}$

Table 4.7.: ParticleNet-MD scale factors for 2017.

mistag rate	$200 \leq p_T / \text{GeV} < 300$	$300 \leq p_T / \text{GeV} < 400$	$400 \leq p_T / \text{GeV} < 800$
0.5%	$0.85^{+0.03}_{-0.03}$	$0.85^{+0.03}_{-0.03}$	$0.86^{+0.05}_{-0.05}$
1.0%	$0.91^{+0.02}_{-0.02}$	$0.90^{+0.03}_{-0.03}$	$0.89^{+0.05}_{-0.04}$
2.5%	$0.96^{+0.03}_{-0.03}$	$0.97^{+0.03}_{-0.02}$	$0.98^{+0.05}_{-0.05}$

Table 4.8.: ParticleNet-MD scale factors for 2018.

mistag rate	$200 \leq p_T / \text{GeV} < 300$	$300 \leq p_T / \text{GeV} < 400$	$400 \leq p_T / \text{GeV} < 800$
0.5%	$0.81^{+0.03}_{-0.03}$	$0.81^{+0.02}_{-0.02}$	$0.77^{+0.04}_{-0.04}$
1.0%	$0.87^{+0.02}_{-0.02}$	$0.86^{+0.02}_{-0.02}$	$0.82^{+0.04}_{-0.04}$
2.5%	$0.92^{+0.03}_{-0.02}$	$0.92^{+0.02}_{-0.02}$	$0.87^{+0.04}_{-0.04}$

Preselection

This section describes the event preselection before applying any jet tagger. In this analysis, the boosted regime is targeted, where the vector bosons have a high $p_T \geq 200\,\text{GeV}$ resulting in a small opening angle and are then reconstructed as a large radius jet with R = 0.8, called AK8 jets. In addition, there are two quarks in the forward region that are typical for VBS resulting in two AK4 jets. Details about jet reconstruction are given in chapter 3.

Table 4.9 summarizes the preselection selection criteria without jet tagger. There are split into the VV selection targeting the AK8 jets and VBS selection targeting the AK4 jets. Using this distinction, two regions are defined: the control region (CR) passes the VV selection but fails the VBS selection, whereas the signal region (SR) passes both selection. They are thus mutually exclusive and discussed in the next section.

Table 4.9.: Summary of the event preselection used in the analysis.

VV Selection: AK8 jets	VBS Selection: AK4 jets
$N \geq 2$	$N \geq 2$
$p_T \geq 200\,\text{GeV},\ |\eta| \leq 2.5$	$p_T \geq 50\,\text{GeV},\ |\eta| \leq 4.7$
$55\,\text{GeV} \leq m_J \leq 215\,\text{GeV}$	
$1246\,\text{GeV} \leq m_{JJ} \leq 6000\,\text{GeV}$	$m_{j_1 j_2} \geq 500\,\text{GeV}$
$\Delta\eta < 1.3,\ \rho < -1.8$	$\Delta\eta > 3.0,\ \eta_1 \cdot \eta_2 < 0$

Here, ρ is defined as $\rho = \log(m_{\text{SD}}^2/p_{\text{T}}^2)$ and this additional selection criterion vetoes events with high jet mass but low p_{T} which affects the tagging efficiently resulting in a better agreement of simulation and data.

Event Categorization

Different selection criteria are applied to define mutually exclusive regions in phase space, which are schematically shown in Fig. 4.6. The SR is optimized to increase the fraction of expected signal events. Events in the SR are required to pass all criteria listed in table 4.9 and one AK8 jet passes the tight working point and another AK8 jet at least the medium working point of the ParticleNet-MD algorithm as they are defined in tables 4.5–4.8. A region enriched in events from QCD multijet production ("CR-QCD") is defined by the same selection criteria as the SR, but both AK8 jets fail the tagger threshold for the SR while passing the loose working point. This region is dominated by QCD multijet production, while the additional requirement of passing the loose working point ensures that it is close enough to the SR to estimate the shape of the QCD background from data. When both AK8 jets fail the loose working point, the event lies in the region called "CR-NoTag", which is used for additional cross checks. Lastly, a control region is defined, where any of the selection criteria listed as "VBS Selection" in table 4.9 is not fulfilled, while the same working points of the ParticleNet-MD as in the SR are required. Backgrounds with a genuine vector boson in the final state where additional AK4 jets do not originate from the hard scattering at leading order are expected to largely fall into this region. Lastly, the two regions marked in red in Fig. 4.6 are not used in the analysis.
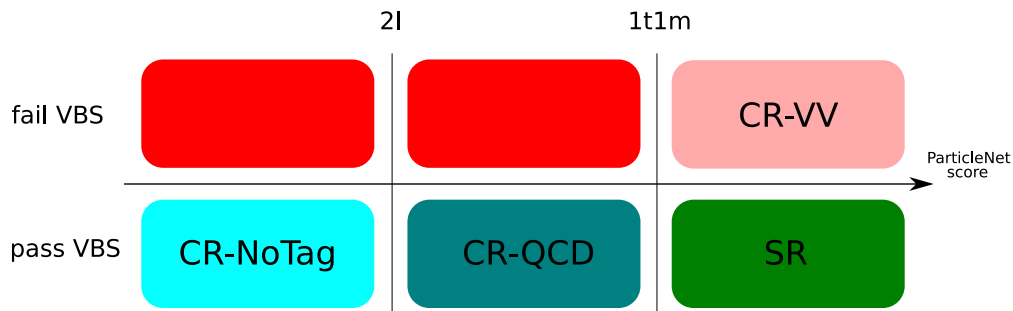
Figure 4.6.: Overview of the defined regions in the analysis. The signal region (in green) requires to pass the full set of selection criteria in table 4.9 as well as one AK8 jet passing the tight, and one AK8 jet passing the medium working point of the ParticleNet-MD algorithm as defined in tables 4.5–4.8.

Data to Simulation Comparison

The following figures show the comparison of Data and MC simulation in the CR and show the MC simulation only in the SR, since this analysis is still blinded. Additionally, four different signal variations are shown: $c_{\mathrm{HBox}} = 7.5$, 15, and 22.5, which are scaled by an arbitrary number. Data corresponds to the full Run2 dataset, which is compared to the MC samples weighted with their respective luminosity. Comparison of data and MC simulation split by run periods can be found in appendix C.

Figure 4.7 shows MC simulation in the combined regions passing all preselection criteria without requirement on the ParticleNet-MD output score, and Fig. 4.8 the corresponding distributions when the selection criteria on AK4 jets are not fulfilled. Scale factors are not applied, which explains the discrepancy of data and MC simulation in the score of the tagging algorithms. Where data is shown, the integral of QCD multijet background is scaled to data to account for normalization that is derived in the fit to data.
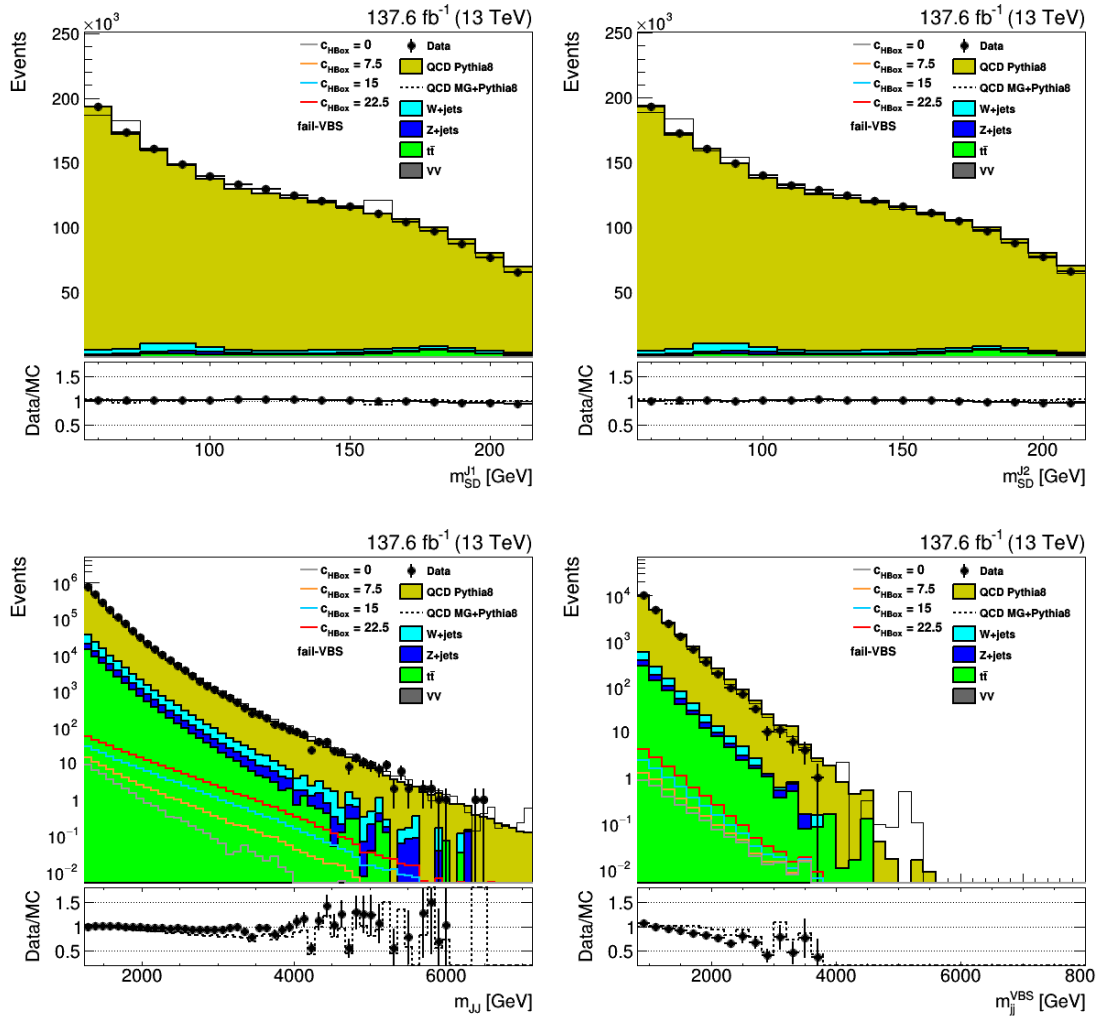
Figure 4.7.: Comparison of simulated events with data in the control region, where selection criteria on AK4 jets are not passed. The QCD contribution is scaled such that the integral of data and simulated events is the same, since the normalization is taken from data. Shown is the soft drop mass of the AK8 jets (top row), the dijet invariant mass of AK8 jets (bottom left), and the dijet invariant mass of the AK4 tagging jets (bottom right). The two AK8 jets leading in p_{T} are chosen, and suitable AK4 jets are not present in every event in the CR.
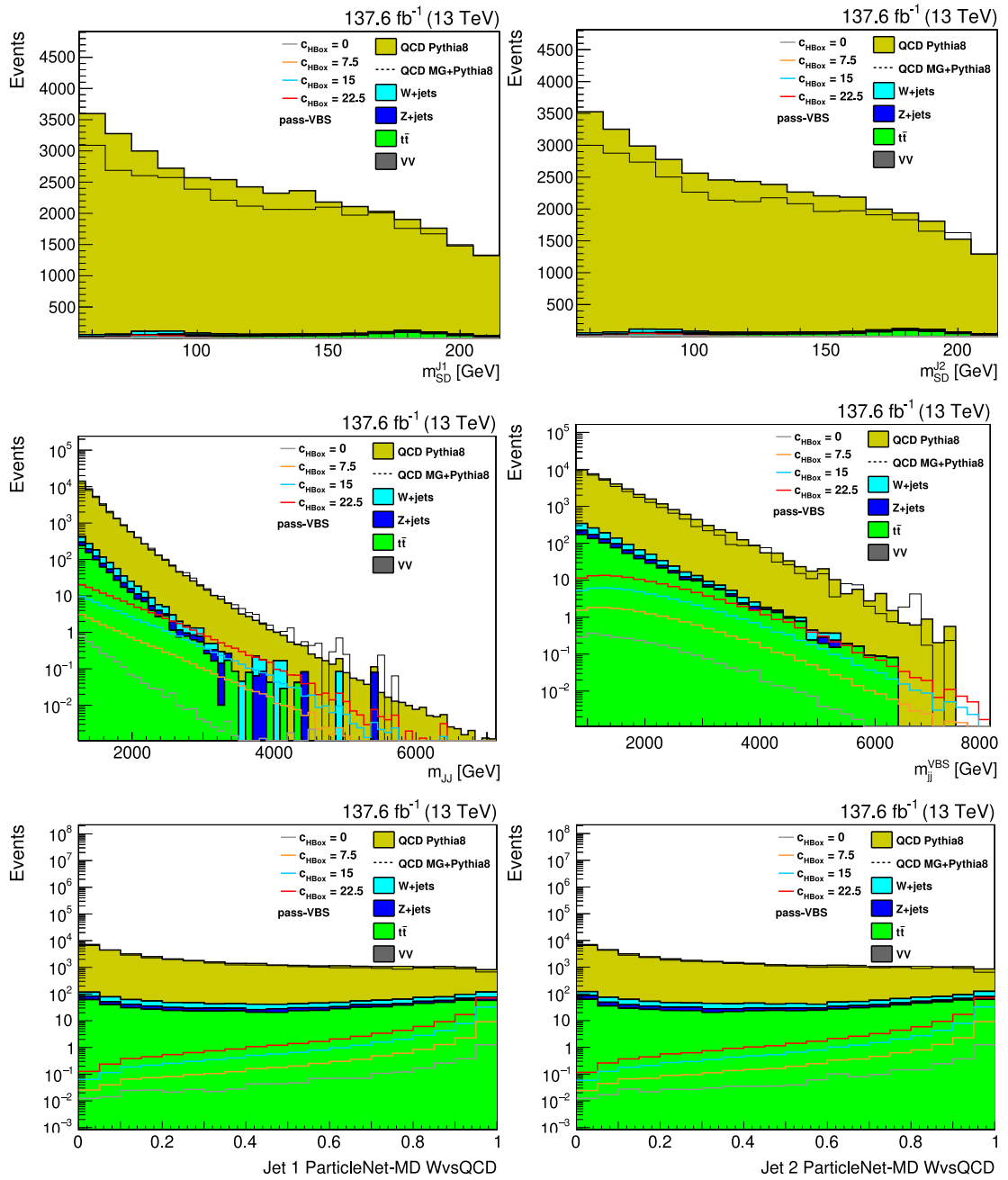
Figure 4.8.: Distributions of simulated events in the signal region, without selection on the ParticleNet-MD score. Shown is the soft drop mass of the AK8 jets (top row), the dijet invariant mass of AK8 jets (middle left), and the dijet invariant mass of the AK4 tagging jets (middle right). The score of the ParticleNet-MD algorithm is shown in the bottom row, where no cuts and scale factors have been applied.

4.1.4. Fitting Strategy

This analysis implements a fit in three dimensions, m_{J1}-m_{J2}-m_{JJ}, to set limits on dimension-6 and dimension-8 operators. Here, m_{J1} and m_{J2} are the masses of the two selected AK8 jets and m_{JJ} is the invariant mass of the AK8 jet pair. The AK4 jets are used to target the SR but are then not explicitly used in the final fit for signal extraction because the shapes of signal and background are non-resonant and falling.

The main background from QCD multijet production is non-resonant in all three dimensions: both AK8 jet masses, m_{J1} and m_{J2}, and the invariant mass of both AK8 jets m_{JJ}. A simultaneous fit in these three dimensions allows to exploit this fact to constrain the large QCD background. The second-largest background from $t\bar{t}$ production shows peaks in m_{J1} and m_{J2} at both the mass of the W boson and top quark after applying the ParticleNet-MD WvsQCD discriminator. These two peaks, together with the fact that an extended jet mass region (55-215 GeV) is fitted, distinguish the $t\bar{t}$ background from signal and further background contributions. Further backgrounds from W+jets, Z+jets, VV-, and QCD-induced VBS production, as well as electroweak VBS production, show the same shapes as the signal but their contributions are much smaller than the above mentioned backgrounds. These contributions peak at the W or Z boson mass in m_{J1} and m_{J2}.

All contributions, signal and backgrounds, show a falling spectrum in m_{JJ}. Reason for this third axis is that it is very sensitive to EFT contributions, the signal considered in this analysis, as has been seen in multiple previous analyses [153, 154]. Figure 4.9 shows a schematic overview of the 3D space with resonant contributions in m_{J1} and m_{J2} at the W boson mass and the change in m_{JJ} due to the EFT contribution.
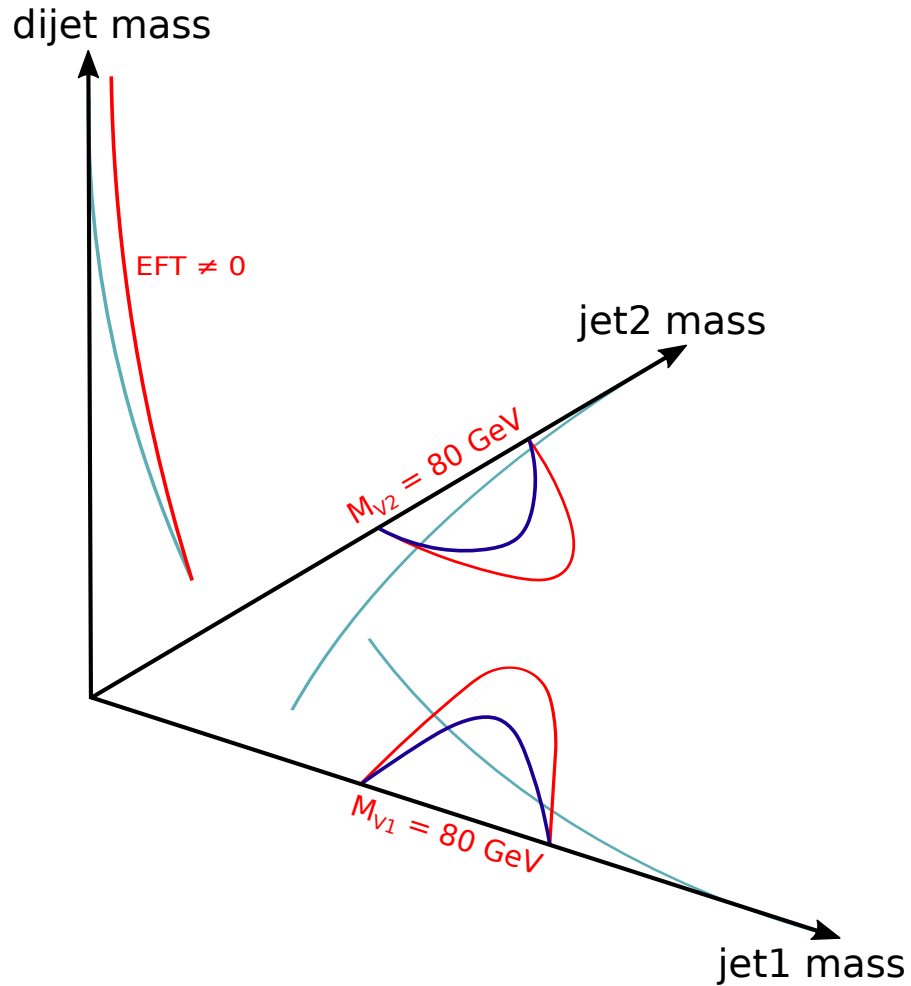
Figure 4.9.: A schematic overview of the fitting strategy in three dimensions: the mass of each individual AK8 jet, and the dijet invariant mass. Background contributions are resonant (blue) or non-resonant (teal) in the jet mass, while the signal contribution (red) is resonant. All contributions show a falling distribution in the dijet invariant mass, where the effect of EFT is seen as a change in the slope.

4.2. Signal Modeling

The signal is defined as the EFT contribution of EWK VBS production, while the EWK production present in the SM is considered as background. Additional terms in the Lagrangian Eq. 1.22 lead to additional terms in the matrix element with linear and quadratic dependence on the Wilson coefficient c_α:

$$N \propto |A|^2 = |A_{SM}|^2 + \sum_\alpha \frac{c_\alpha}{\Lambda^2} \cdot 2\Re(A_{SM}A_{Q_\alpha}^\dagger) + \sum_{\alpha,\beta} \frac{c_\alpha c_\beta}{\Lambda^4} \cdot (A_{Q_\alpha}A_{Q_\beta}^\dagger) \qquad (4.3)$$

The first term, $|A_{SM}|^2$, is the EWK SM production as observed in Ref. [160, 161], while the two additional terms arise from the additional EFT operators. They are either linear in c_α/Λ or quadratic. The case is the same for dimension-6 operators as it is for dimension-8 operators when replacing c_α/Λ with f_i/Λ^2, where c_α is the Wilson coefficient of dimension-6 and f_i for dimension-8 operators.

In order to apply the three-dimensional fit method, the signal has to be parametrized in all three dimensions, which are taken to be uncorrelated:

$$P_{sig}(m_{JJ}, m_{J1}, m_{J2}|\bar{\theta}) = N(\bar{\theta}_{norm}) \times P_{JJ}(m_{JJ}|\bar{\theta}_3) \times P_{J1}(m_{J1}|\bar{\theta}_1) \times P_{J2}(m_{J2}|\bar{\theta}_2). \qquad (4.4)$$

The shapes for m_{JJ}, m_{J1}, m_{J2} are denoted as P_{JJ}, P_{J1} and P_{J2} and depend on additional parameters $\bar{\theta}_i$. The shape of a single AK8 jet mass, P_{J1} and P_{J2}, is given by a DSCB function, whereas the parametrization of the invariant mass of the AK8 jet pair, P_{JJ}, is motivated by Eq. 4.3.
Additionally, the normalization for the signal component is described in the following and is multiplied to the above probability density function (pdf) in three dimensions.

Cross Section

The cross section of the EFT contribution is parametrized with a quadratic function following the consideration in Eq. 4.3. Figure 4.10 shows the reweighted simulated events and the fit function for all three parameters in the SR as defined by the preselection and ParticleNet-MD working point, normalized to the SM value of EWK VBS production. For all three parameters, the quadratic fit functions describes the simulation very well and is used do scale the overall signal normalization in order to interpolate to values between the discrete values used in reweighting the simulated sample.
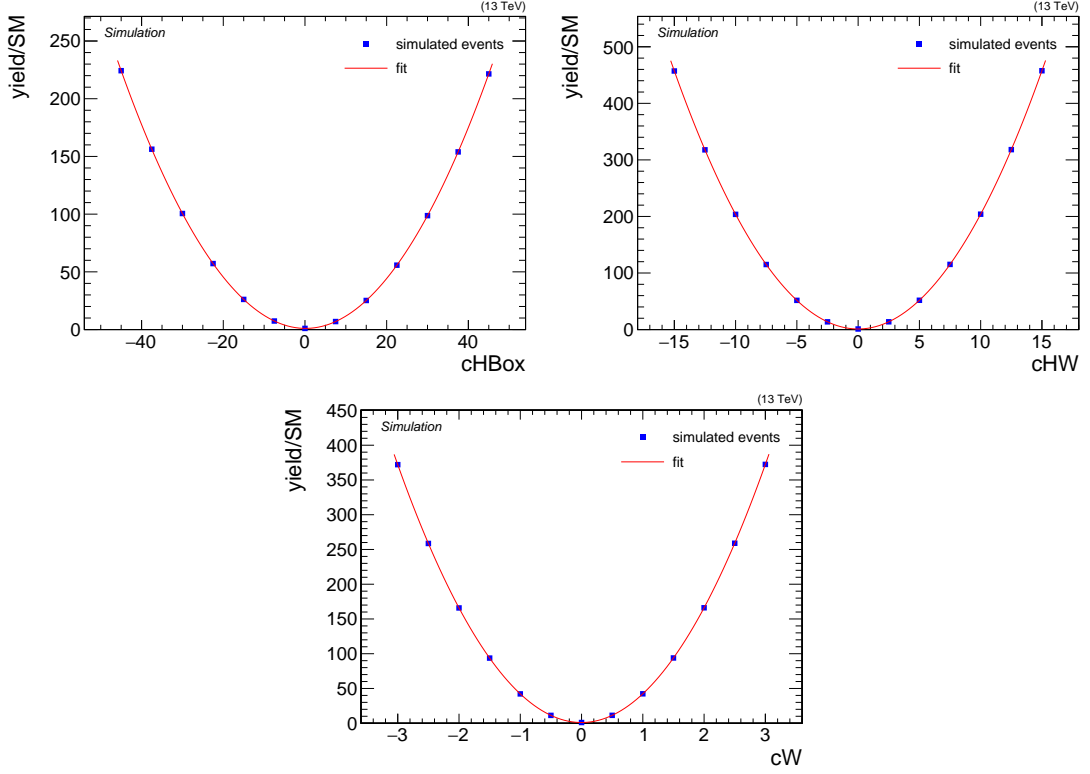
Figure 4.10.: The signal contribution divided by the SM value in the signal region for all three operators: c_{HBox}, c_{HW} and c_{W}.

Jet Mass

The soft drop jet mass m_{SD} for one of the AK8 jets is described by a DSCB as shown in Fig. 4.12:

$$\text{DSCB}(t; \alpha_1, N_1, \alpha_2, N_2) = \begin{cases} \exp\left(-\frac{1}{2}\alpha_1^2\right) \left[1 - \frac{\alpha_1}{N_1}(\alpha_1 + t)\right]^{-\alpha_1} & \text{, if } t \leq -\alpha_1 \\ \exp\left(-\frac{1}{2}t^2\right) & \text{, if } -\alpha_1 < t < \alpha_2 \quad (4.5) \\ \exp\left(-\frac{1}{2}\alpha_2^2\right) \left[1 - \frac{\alpha_2}{N_2}(\alpha_2 - t)\right]^{-\alpha_2} & \text{, if } t \geq \alpha_2 \end{cases}$$

with $t = \frac{x - \text{mean}}{\text{width}}$ describing the Gaussian core, model parameters α_1, α_2 the starting value and N_1, and N_2 the strength of the falling flanks.

The EWK VBS including the EFT contribution are fitted for varying values of the Wilson coefficient under consideration and resulting parameters are then interpolated with a spline as shown in appendix C. Maximum values of the EFT Wilson coefficients are directly given by the endpoints for reweighting in MadGraph5.
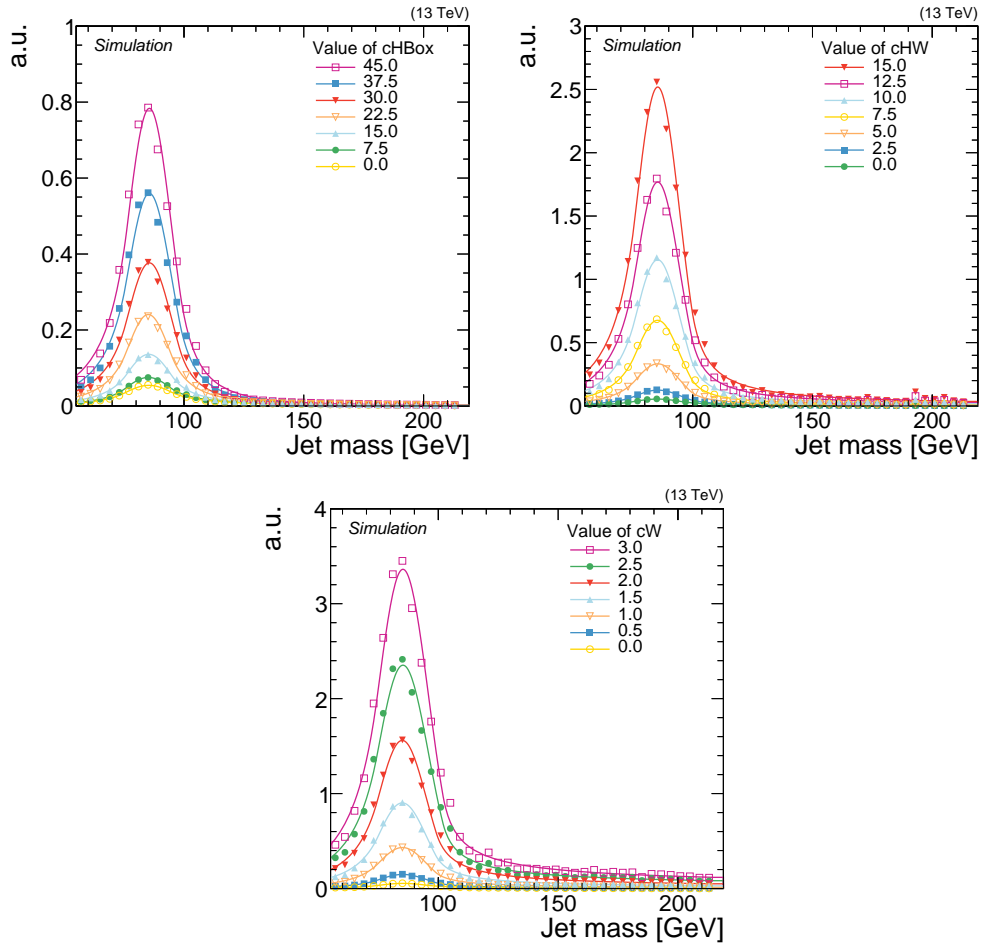
Figure 4.11.: Distribution of the AK8 jet soft drop mass m_{SD} for simulated events and fit of the double-sided Crystal Ball function for different values of c_{HBox}, c_{HW}, and c_{W}. Since the AK8 jets are randomly ordered, the distribution is the same for m_{J1} and m_{J2}.
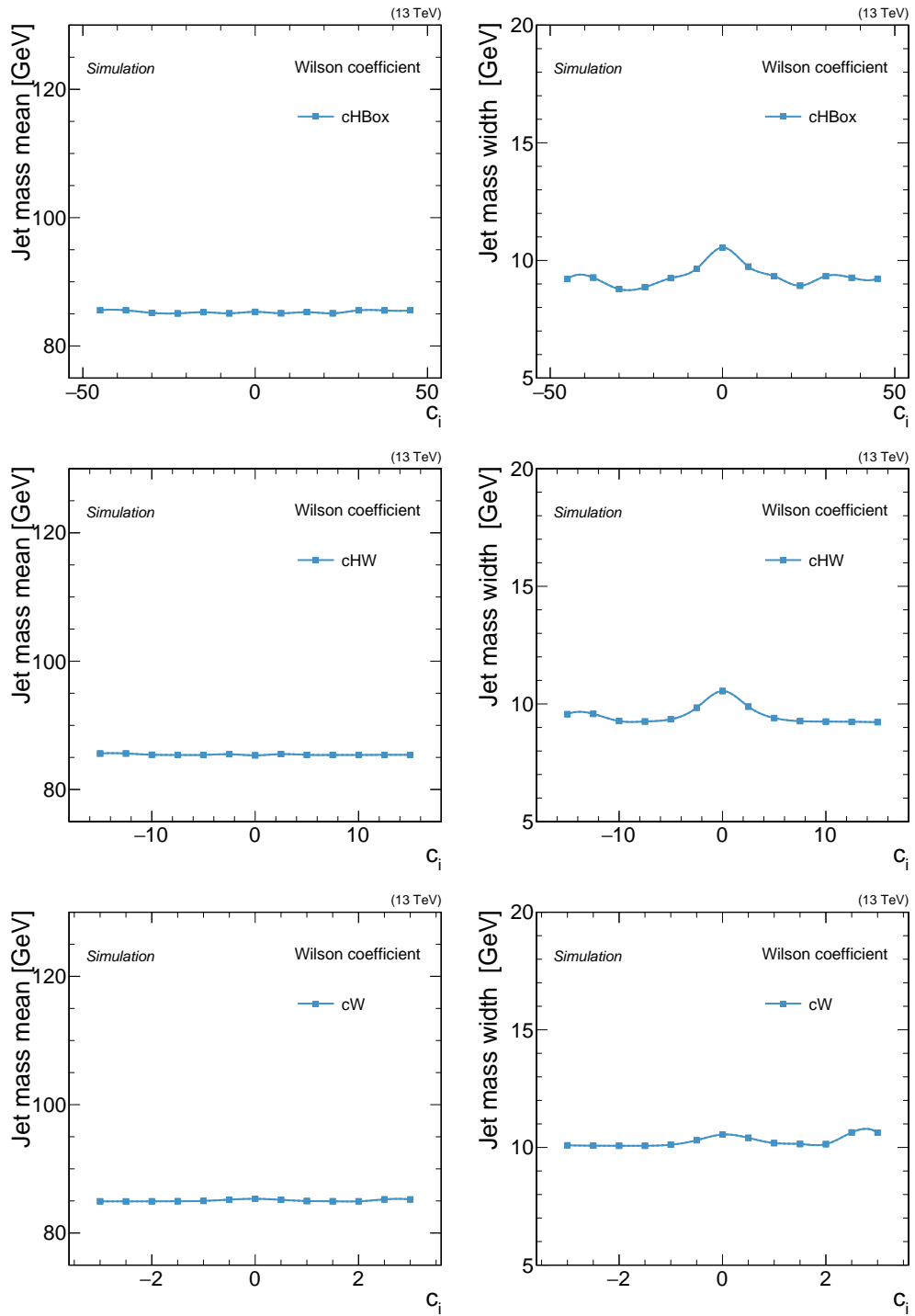
Figure 4.12.: Spline interpolation of the mean (left) and width (right) of the DSCB used to describe the signal distribution in the jet mass for different values of the Wilson coefficients.

Dijet Invariant Mass

Although the shape of m_{SD} does not change significantly for different Wilson coefficients, this is not the case for the dijet invariant mass m_{JJ}.

For both the SM production and the production associated to anomalous couplings, the m_{JJ} distribution is in good approximation exponentially falling. Therefore, a sum of exponential terms is used to model the shape of the SM, pure EFT, and SM-EFT interference contribution. The relative normalization of the EFT associated terms is extracted from the MADGRAPH5 signal samples as described above. The shapes for the SM and EFT contributions are also derived by fits to the signal simulation. Motivated by Eq. 4.3, the m_{JJ} distribution is modelled as:

$$
\begin{aligned}
\text{P}_{\text{JJ}}(m_{\text{JJ}} \mid c_{\text{i}}) = {} & \text{N}_{\text{SM}} \cdot (e^{a_0 m_{\text{JJ}}} + e^{a_{\text{corr}} m_{\text{JJ}}}) \\
& + \text{N}_{\text{int}} \cdot c_{\text{i}} \cdot e^{a_{2,\text{i}} m_{\text{JJ}}} \\
& + \text{N}_{\text{quad}} \cdot c_{\text{i}}^2 \cdot e^{a_{1,\text{i}} m_{\text{JJ}}} \cdot \frac{1 + \text{erf}((m_{\text{JJ}} - a_{0,\text{i}})/a_{\text{w,i}})}{2} \, ,
\end{aligned}
\tag{4.6}
$$

where c_{i} represents the Wilson coefficients, such that there can be three contributions identified: the SM part which is independent of c_{i}, the pure EFT contribution proportional to c_{i}^2 and the EFT-SM interference term proportional to c_{i}. The N_{SM}, N_{int}, N_{quad} represent the normalization of the different contributions and a_0, a_1, a_2 model the exponential slopes of these three contributions. A correction term, a_{corr}, allows for a correction of the SM contribution from a simple exponential at higher values of m_{JJ}. The final parameters $a_{0,\text{i}}$ and $a_{\text{w,i}}$ define the turn-on and steepness of the pure EFT contribution via an Gaussian error function:

$$
\text{erf}(x) = \frac{2}{\pi} \int_0^x \exp^{-t^2} \, \mathrm{d}t \, .
\tag{4.7}
$$

In order to extract the parameter values from simulated MC events, first the SM contribution is fitted on its own to the SM expectation, where the EFT contribution is set to zero. Then the difference between the samples weighted to the same magnitude but opposite sign of c_{i} is used to extract the EFT-SM interference term, since both the SM contribution and pure EFT contribution cancel out. Finally, all three contributions are fit together to the samples weighted to a single non-zero EFT value. This method allows for a stable fit to determine all parameters and can be expanded to consider multiple Wilson coefficients simultaneously by introducing and EFT-EFT interference term to Eq. 4.6.

This fitting procedure is done for a specific choice of the Wilson coefficient as working point given by 7.5, 2.5, and 0.5 for c_{HBox}, c_{HW}, and c_{W}, respectively. Figure 4.13 shows the resulting function for different values of the Wilson coefficient.
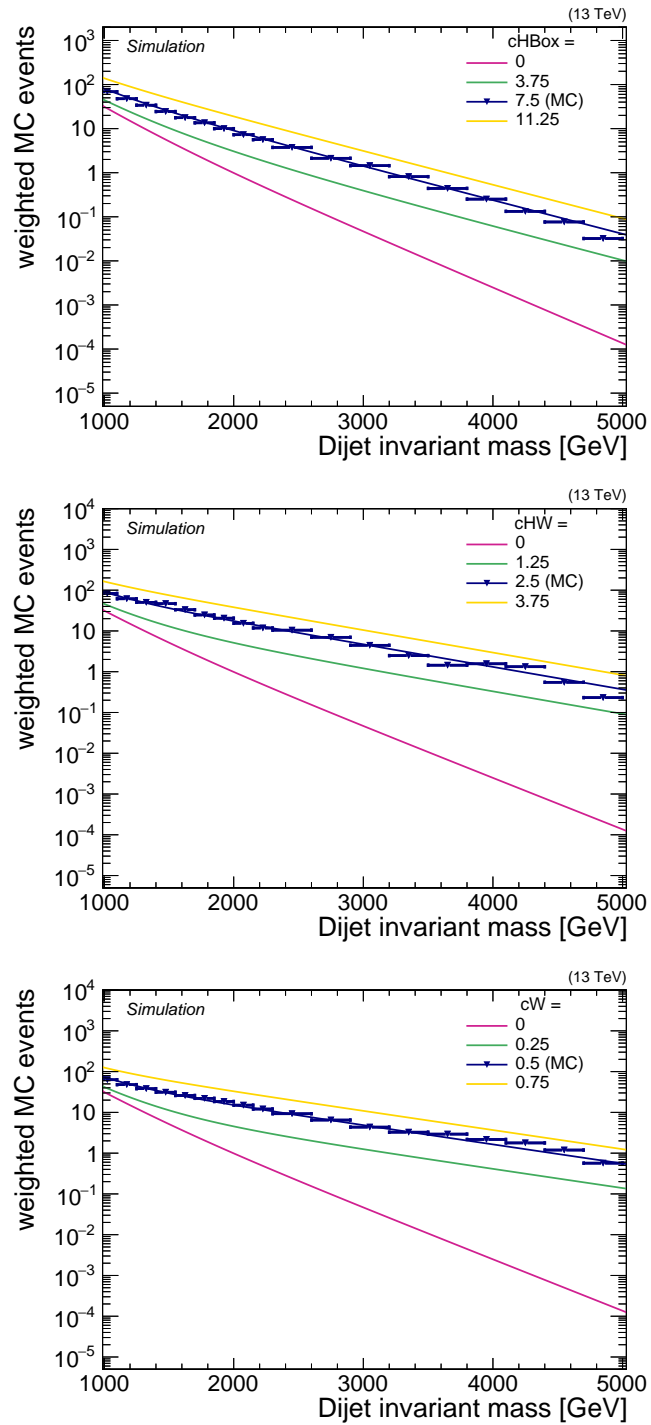
Figure 4.13.: Distribution of the dijet variant mass of the AK8 jet pair m_{JJ} for simulated events and fit of the full parametrization for c_{HBox} (top), c_{HW} (middle), and c_{W} (bottom). The indicated MC events are the nominal events used to derive the template. The other pdfs are derived by setting the Wilson coefficients to the respective value.

4.3. Background Modeling

All background contributions present in the SM are also modelled in three dimensions. The largest background is given by QCD multijet production, which is non-resonant in all three dimensions entering the fit. Top quark pair production is the second-largest contribution with resonant structures in the AK8 jet mass at the mass of the W boson and the top quark. Smaller contributions are given by V+jets, diboson, single top, and QCD-induced VBS production.
The following section describe the modeling of each contribution entering the fit.

4.3.1. QCD Background

The main background in this analysis comes from QCD multijet production. This background has the defining characteristic that all three dimensions in the final fit do not show a resonant structure, which is fully exploited in this fitting strategy. The modeling follows a similar approach as described in Ref. [153].
In order to model the QCD multijets background in the three-dimensional m_{JJ}-m_{J1}-m_{J2} plane, the following conditional product is used:

$$P_{QCD}(m_{JJ}, m_{J1}, m_{J2}) = P_{JJ}(m_{JJ}|\Theta_1) \times P_{cond,1}(m_{J1}|m_{JJ}, \Theta_2) \times P_{cond,2}(m_{J2}|m_{JJ}, \Theta_3), \quad (4.8)$$

where P_{JJ} is the template for the dijet invariant mass, and $P_{cond,1}$, $P_{cond,1}$ are conditional templates for a single AK8 jet mass, practically given by 2D histograms in m_{JJ} and m_{J1} with the same binning in m_{JJ} as P_{JJ}. Since both AK8 jets are randomly sorted, both conditional templates are the same except for statistical fluctuation. The parameters Θ_1, Θ_2, and Θ_3 are nuisance parameters used in the final fit for limit extraction.
This approach allows to include the correlation of m_{JJ}, m_{J1}, and m_{J2} and requires the computation of 2D-conditional pdfs for the AK8 jet masses given m_{JJ}. The fit range for m_{J1}/m_{J2} goes from 55 to 215 GeV with equidistant binning. m_{JJ} is fitted from 1246 to 7600 GeV in order to stay on the trigger plateau and avoid complications due to trigger turn-on effects. The sensitivity to contributions from EFT is almost exclusively in the high energy region such that no relevant regions are excluded. The binning of m_{JJ} is taken to correspond to the actual resolution in the dijet mass, such that a resonant structure would be present in at least three neighbouring bins. The resulting binning in GeV is:
1246,1313,1383,1455,1530,1607,1687,1770,1856,1945,2037,2132,2231,2332,2438,2546,
2659,2775,2895,3019,3147,3279,3416,3558,3704,3854,4010,4171,4337,4509,4686,4869,
5058,5253,5058,5253,5455,5663,5877,6099,6328,6564,6808,7060,7320,7600.

The background model is built starting from MC simulation using the p_T-binned PYTHIA8 sample with a large number of simulated events and the shape is fit to data in the CR-QCD control region with nuisance parameters for each alternative shape variation. For this a forward-folding approach is used: A 2D- or 1D-Gaussian kernel is built from generator level quantities in m_{JJ}-m_J- or m_{JJ}-space, respectively, binned in generator level p_T in order to describe the detector response for the jet and dijet invariant mass. Each simulated event then contributes to the probability density function $P_{\text{cond},1/2}$ or P_{JJ} with the derived kernel effectively smearing the distribution of simulated events. The resulting templates are then fit to data in a defined control region, where the ParticleNet-MD working points as defined in 4.1.3 are failed but the loose ones passed. Alternative shapes allow the template to adapt itself to data, whereas the normalization is constrained in the final fit due to the large range in m_{J1} and m_{J2} and a large normalization uncertainty.

Summarizing the procedure, it can be split in the following steps:

- Deriving Gaussian kernels to describe the detector response

- Smearing jet and dijet invariant mass distributions from simulation

- Building alternative shapes by varying the templates

- Fitting the templates to data in the CR-QCD control region

Detector Response

In order to derive Gaussian kernels from generator level quantities, resolution and scale for m_J and m_{JJ} are derived by fitting a Gaussian function to $m_{\text{jet}}^{\text{reco}}/m_{\text{jet}}^{\text{gen}}$ or $m_{JJ}^{\text{reco}}/m_{JJ}^{\text{gen}}$, respectively, in bins of $p_T^{\text{gen.}}$. Each fit is performed two times to keep the procedure stable: first with an extended range to find starting values for width and mean of the second fit. Figure 4.15 shows the resulting fits for a single bin in $p_T^{\text{gen.}}$ for $m_{\text{jet}}^{\text{reco}}/m_{\text{jet}}^{\text{gen}}$ (left) and $m_{JJ}^{\text{reco}}/m_{JJ}^{\text{gen}}$ (right). From these fits, the Gaussian mean gives the scale and the Gaussian width the mass resolution, i.e., how to scale and smear the generator level quantity. The resulting scales and resolutions depending on $p_T^{\text{gen.}}$ for both m_{jet} and m_{JJ} are shown in Fig. 4.14.

In order to populate the conditional 2D histogram for $P_{\text{cond},1/2}(m_{\text{Jet}1/2}|m_{JJ})$ each generated event i is smeared with the following 2D Gaussian kernel:

$$
k_i(m_{JJ}, m_{\text{jet}}) = \frac{\omega_i}{\sqrt{2\pi} \cdot R_{m_{JJ},i} \cdot R_{m_{\text{Jet}},i}}
$$
$$
\exp\left[-\frac{1}{2}\left(\frac{m_{JJ} - S_{m_{JJ},i} \cdot m_{JJ,i}^{\text{gen}}}{R_{m_{JJ},i} \cdot m_{JJ,i}^{\text{gen}}}\right)^2\right]
$$
$$
\exp\left[-\frac{1}{2}\left(\frac{m_{\text{Jet}} - S_{m_{\text{Jet}},i} \cdot m_{\text{Jet},i}^{\text{gen}}}{R_{m_{\text{Jet}},i} \cdot m_{\text{Jet},i}^{\text{gen}}}\right)^2\right], \tag{4.9}
$$

where R_i is the resolution and S_i the scale derived before depending on $p_{\mathrm{T},i}^{\mathrm{gen}}$ and can be read from the histograms in Fig. 4.14, and ω_i is the event weight resulting from, e.g., PU correction, cross section, and luminosity. All quantities in the equation above with an index i are potentially different for each event such that the weight for m_{Jet} or m_{JJ} is smaller for values further away from the scaled generator-level quantity. In order to build the 1D histogram for m_{JJ}, the same procedure is used where only a one-dimensional Gaussian function is necessary.

Inclusive templates without any requirement on the jet tagger are fitted to increase the statistics and remove residual bias from events at the corner of the phase space. This choice is taken because the categories have limited statistics but the shape is the same as inclusive distributions.
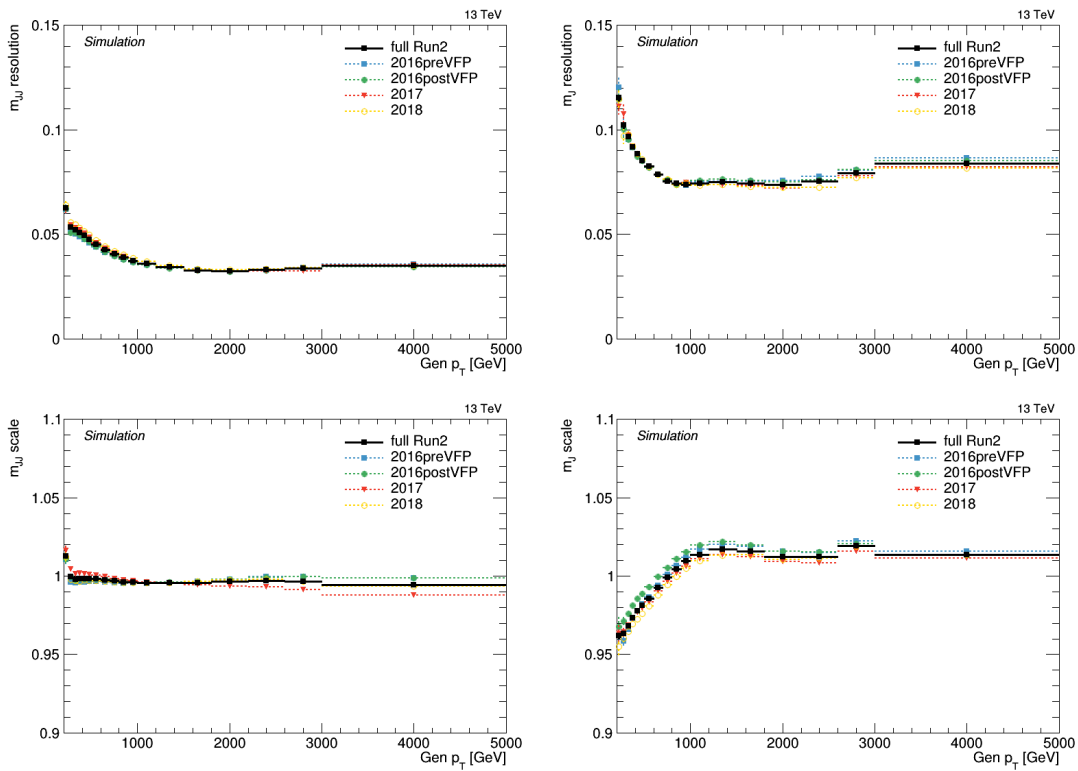
Figure 4.14.: Resolution (top) and scale (bottom) for m_{JJ} (left) and m_{J} (right) as a function of generator-level p_{T}.
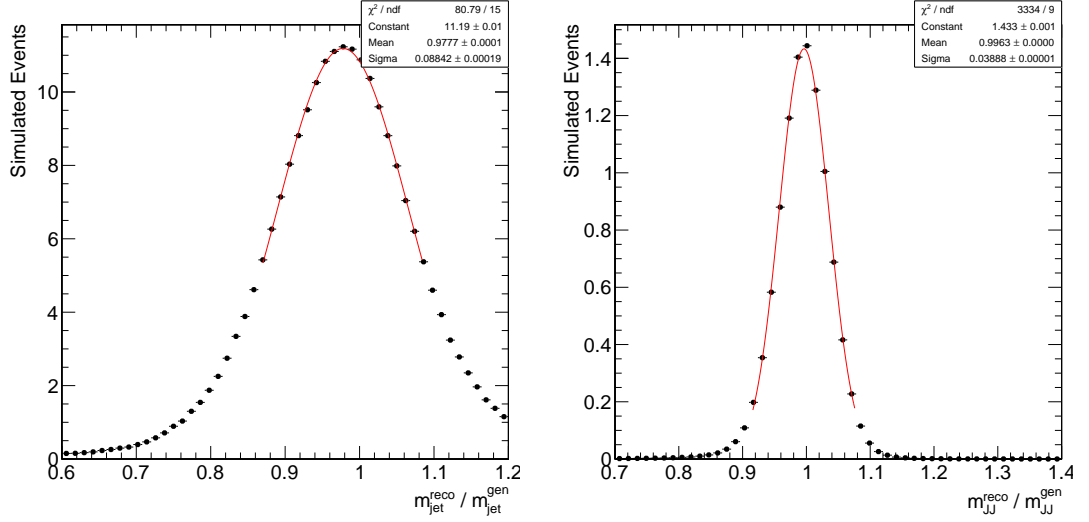
Figure 4.15.: Fit to $m_{\text{jet}}^{\text{reco}}/m_{\text{jet}}^{\text{gen}}$ (left) and $m_{\text{JJ}}^{\text{reco}}/m_{\text{JJ}}^{\text{gen}}$ (right) in one bin of generator-level jet p_{T} for the nominal p_{T}-binned PYTHIA8 sample. The width is taken as resolution and the mean as scale.

Alternative Templates

In order to account for shape mismodeling in the QCD templates, the template derived with the kernel approach and MC simulation are allowed to adapt to data by vertical template morphing, which creates one nuisance parameter for each shape. These shapes simultaneously affect m_{JJ}, m_{J1}, and m_{J2}. In total three alternative shapes are included: The first alternative shape simultaneously varies m_{JJ} and m_{J} by a factor proportional to m_{JJ} and m_{J}, respectively and corresponds to a variation of the underlying p_{T} spectrum. This is achieved by scaling up and down the content of each 1D histogram bin (for m_{JJ}) or 2D histogram (for m_{J1} and m_{J2}) to obtain two new histograms h_u and h_d for each m_{JJ}, m_{J1}, and m_{J2}. The factor with which each bin content is scaled is given by

$$f^i = \left(1 + 1.5 \cdot m^i / m_{\text{max}}^i\right), \tag{4.10}$$

where m^i is either m_{JJ}, m_{J1} or m_{J2}, and m_{max} is the respective maximum of that variable. For the 1D histogram of m_{JJ} with $m_{\text{max}} = 7600\,\text{GeV}$, this results in a maximum variation of 150% and a minimum variation of 25%. The up-variation is obtained by multiplying with this factor, while the down-variation is obtained by dividing by this factor. For the 2D histograms of m_{J1} and m_{J2}, which take the correlations into account, the 2D histogram bin contents are rescaled such that a maximum variation of 150% is achieved, when $m_{\text{J}} = 215\,\text{GeV}$ and a mimimum variation of about 38% for $m_{\text{J}} = 55\,\text{GeV}$, since

$$1.5 \cdot m_{\text{J}}^{\text{min}} / m_{\text{J}}^{\text{max}} = 1.5 \cdot \frac{55}{215} \approx 38\% \,.$$

The final 3D histogram for the up- (down-) variation is then obtained by multiplying the up- (down-) variation of the corresponding 1D and 2D histograms. The resulting alternative shapes are called "$\propto m_{\text{jet}}$".

The second alternative shape is derived similar to the first one with the difference that the factor f is now given by

$$f^i = \left(1 + 1.5 \cdot m^i_{\text{min}} / m^i\right),\tag{4.11}$$

where the dependence on m^i is inverted and the maximum replaced by the minimum value of the quantity corresponding to a variation of the scale. With this change, the procedure is then the same as for the first alternative shape. The resulting alternative shapes are called "$\propto 1/m_{\text{jet}}$".

The final alternative shape considered takes into account the differences in MC generation and modeling of parton shower: these correspond to the template obtained using MADGRAPH5 + PYTHIA8 MC samples and following the same procedure as described in section 4.3.1.

The alternate shapes described above are shown in Fig. 4.16 together with the simulated distributon of QCD multijet production before the fit to data in the control region.
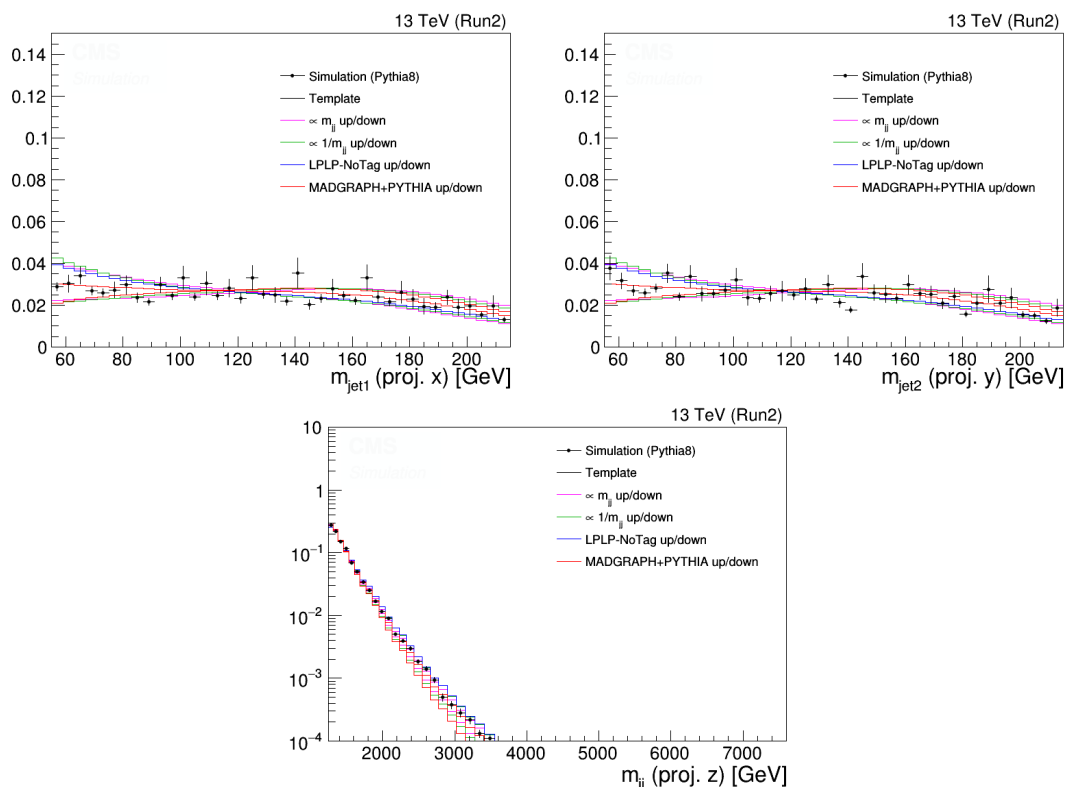
Figure 4.16.: Projections of the QCD template before fit to data in the low purity control region together with simulated events and all considered alternative shapes.

Fit to Data

As the resulting templates are statistically very limited, especially in the SR due to the cuts on the AK4 jets, the shape of the inclusive template is fitted to data in CR-QCD region, which is seen to be similar to the SR and differs in the choice of the tagger working point. This fit strengthens the prediction of the QCD background by looking into data in a QCD-enriched region and also removes residual bias in the template construction due to events at the corner of the phase space.

The fit is performed through vertical morphing among alternative shapes, which are included as shape systematic uncertainties and described in the above section. The shapes resulting of this fit are shown in Fig. 4.18 for the QCD-enriched region. The final templates with shape uncertainties are then shown in Fig. 4.17.
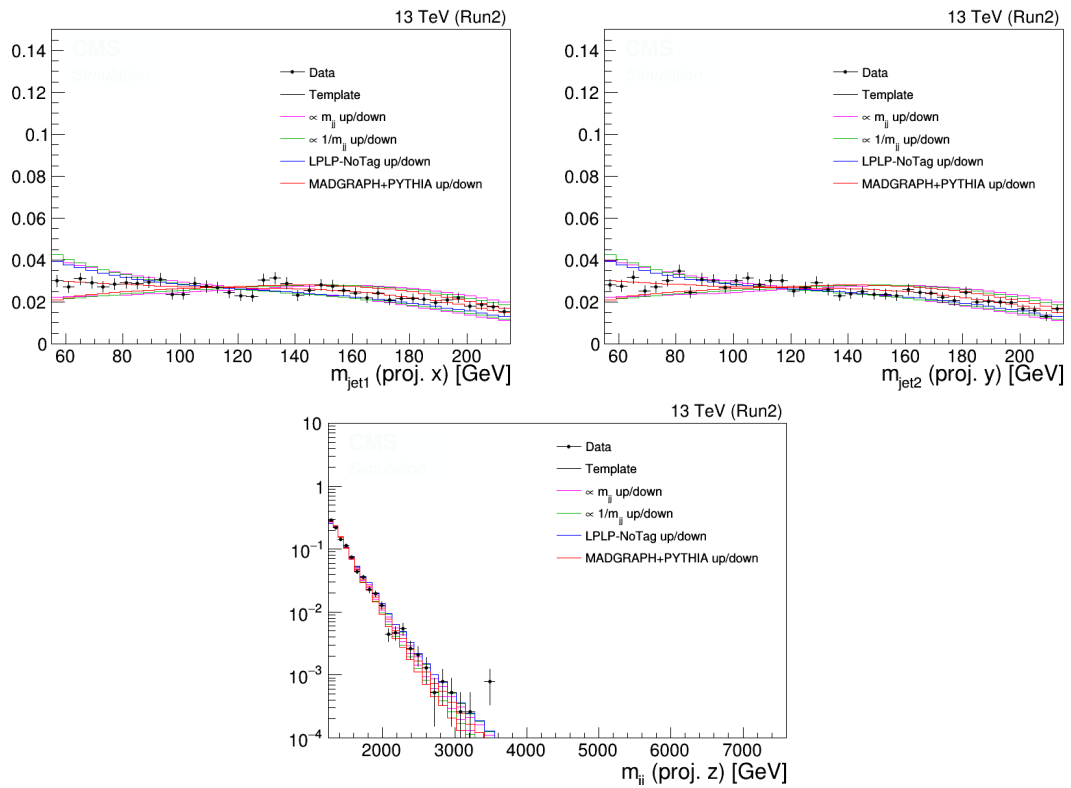
Figure 4.17.: Projection of the QCD template onto all three jet axes in the low purity region after the fit. The alternative shapes described in section 4.3.1 are also shown. The "LPLP-NoTag" is not included in the fit since the shape of the "very low" purity region differs from the shape of the simulated QCD multijet background in the signal region.
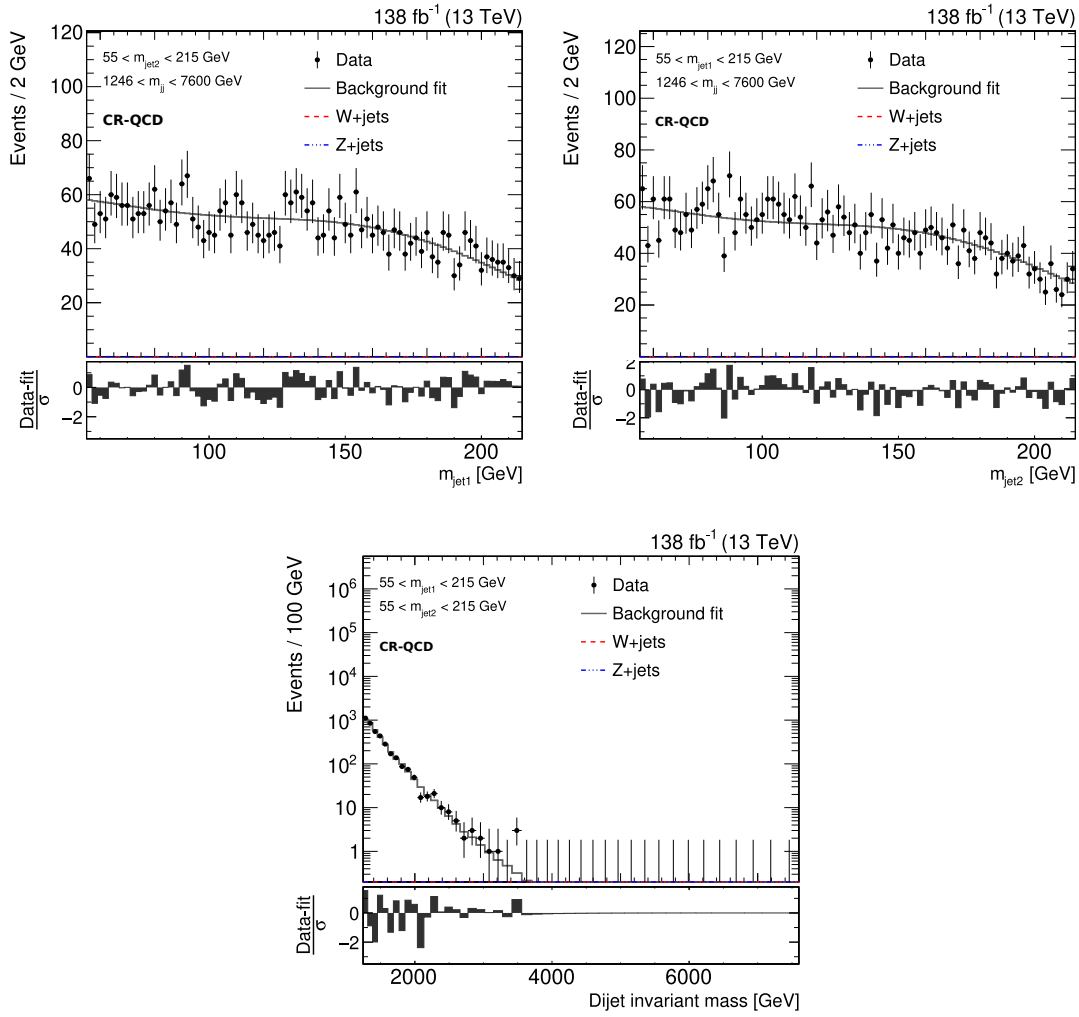
Figure 4.18.: QCD template and data in the low purity region when all other cuts are passed. The AK8 jets are required to fail the medium working point but pass the loose working point as described in 4.1.3.

4.3.2. Modeling of Resonant Backgrounds

Backgrounds with a resonant structure in m_{J1} and m_{J2} are modelled separately. They are categorized into W+jets and Z+jets contribution with a resonance at the W or Z boson mass, respectively, the $t\bar{t}$ contribution with a resonance at both the W boson mass and the top quark mass, and into WZ and ZZ contribution, where both AK8 jets show a resonance in the jet mass spectrum. The peak at the top quark mass is not present when using the n-subjettiness τ_{21}^{DDT} but is due to the more performant ParticleNet-MD algorithm focussing to separate jets from heavy vector bosons from QCD jets.
Since both jet masses, m_{J1} and m_{J2}, are randomly sorted, the resonant structure is present in both quantities. The dijet invariant mass m_{JJ} shows a falling spectrum for all background contributions similar to the QCD non-resonant background described above.

V+jets and VV Backgrounds

Subdominant backgrounds are modelled in groups as follows: W+jets and Z+jets background are modelled separately, since they contain a resonant and non-resonant contribution to the jet mass. Backgrounds containing two Z bosons, namely SM production and QCD-induced VBS production of two Z bosons, and backgrounds containing one W and one Z boson are group accordingly. SM and QCD-induced production of two W bosons are modelled together with the top quark pair production.
For the V+jets background, each jet mass contains two contributions: the resonant part from a real V jet, which shows a clear peak around the V boson, and the non-resonant part, which originates from quark or gluon jets resulting in a very similar distribution to the one for the QCD background. The three-dimensional probability function includes three different terms for each kind of contribution, where the correlation between dijet mass m_{JJ} and the mass of a single AK8 jet m_J was found to be negligible but the correlation between the two single jet masses, m_{J1} and m_{J2}, to be relevant:

$$P_{\mathrm{Vjets}}(m_{JJ}, m_{\mathrm{Jet1}}, m_{\mathrm{Jet2}}) = 0.5 \cdot (P_{JJ}(m_{JJ}|\Theta_1) \times P_{\mathrm{res}}(m_{\mathrm{Jet1}}|, \Theta_2) \times P_{\mathrm{non-res}}(m_{\mathrm{Jet2}}|, \Theta_3))$$
$$+ 0.5 \cdot (P_{JJ}(m_{JJ}|\Theta_1) \times P_{\mathrm{res}}(m_{\mathrm{Jet2}}|, \Theta_2) \times P_{\mathrm{non-res}}(m_{\mathrm{Jet1}}|, \Theta_3)) , \quad (4.12)$$

The contribution from the m_{JJ} spectrum, $(P_{JJ}(m_{JJ}|\Theta_1)$ is modelled with the same kernel approach that is used for the QCD background without fit to data in an enriched control region. Therefore, this background is taken from simulation.
The contribution from m_J is split into two parts: the resonant (P_{res}) and non-resonant ($P_{\mathrm{non-res}}$). For modeling the resonant part, a generated V boson within $\Delta R = 0.8$ around the reconstructed jet is required. A DSCB function, the same function as is used to model the signal contribution, is then fitted to the resonant spectrum for a peak at the W boson and Z boson separately. Uncertainties affecting the resolution and

scale of the m_J distribution are treated as correlated with the signal and $t\bar{t}$ contribution since these uncertainties affect all jets originating from real vector bosons in the same way. The final fits of the DSCB template for the resonant contribution are shown in Fig. 4.19. The non-resonant part of V+jets is modelled with a simple Gaussian to the non-resonant part of the jet mass spectrum. The WZ and ZZ backgrounds are modelled analogously without non-resonant template.
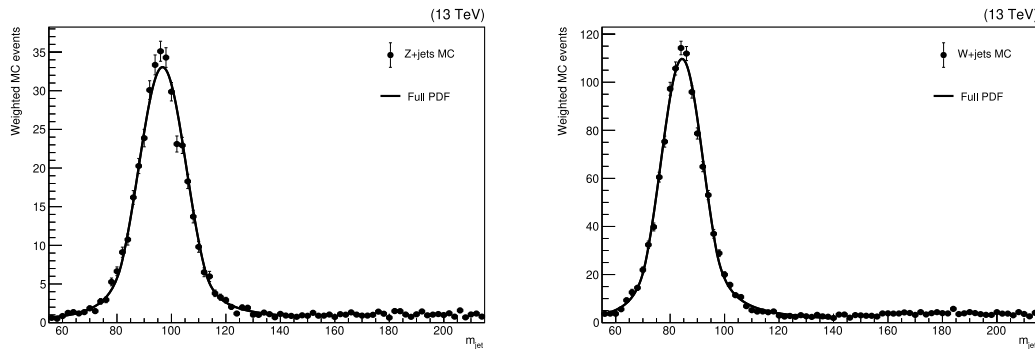
Figure 4.19.: The fit of the DSCB to the resonant part of the Z+jets (left) and the W+jets (right) contributions in the signal region.

Top Quark Pair Production

The SM $t\bar{t}$ background consists of three different components in the jet mass: two resonant components at $m_W \approx 80\,\text{GeV}$ and $m_{\text{top}} \approx 172\,\text{GeV}$, and one non-resonant component. The resonant components arise from merged hadronic W boson decays or merged hadronic top quark decay, respectively. In case the jets are not merged, they are part of the non-resonant contribution. Relative fractions of these components are correlated and related to the mass and p_T of the underlying particle. The opening angle of subjets is given by $\Delta R = 2m/p_T$ resulting in $p_T \approx 200\,\text{GeV}$ for W bosons and $p_T \approx 450\,\text{GeV}$ for top quarks in case of large radius AK8 jets. Above these thresholds, the jets are merged and are expected to be part of the respective contributions. The actual assignment is done by generator level matching: if a W boson or top quark is found in simulated events that is compatible with the reconstructed jet, the jet is assigned to the corresponding contribution, else it is assigned to the non-resonant contribution. In general, this results in a correlation of m_J and m_{JJ}. This is taken into consideration by constructing the m_J template in two different bins of m_{JJ}, a trade-off between the limited statistics after selection criteria on the AK4 tagging jets and capturing the dependence. As a validation, these shapes are also derived in "CR-VV", where the selection criteria on AK4 jets are failed and more simulated events are available.
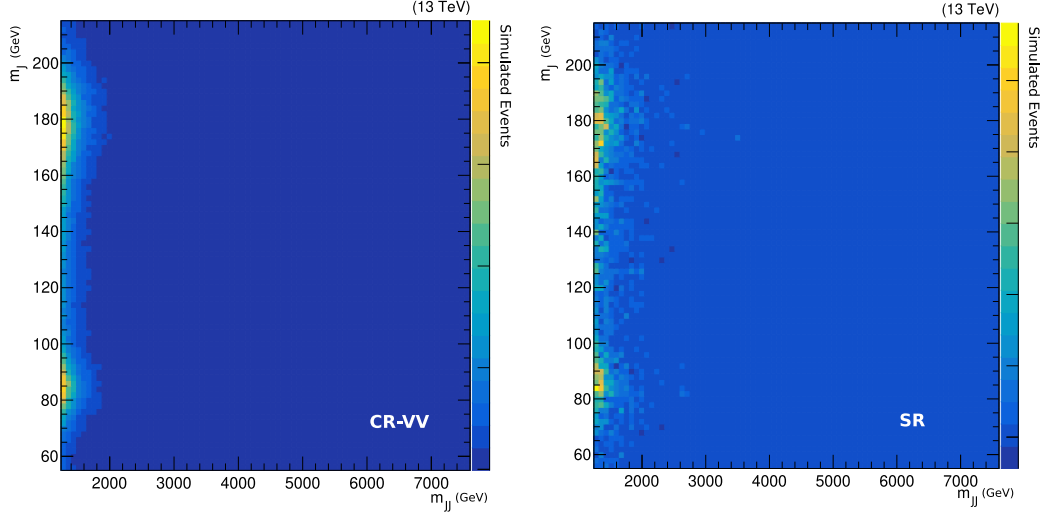
Figure 4.20.: The 2D distribution of the AK8 jet mass versus the dijet invariant mass for the resonant $t\bar{t}$ background, when the selection criteria on AK4 jets are failed (left) and when they are passed (right).

Figure 4.20 shows the distribution of m_J versus m_{JJ} when failing the VBS cuts (left) and when passing them, i.e. the signal region (right), which shows both the correlations of m_J and m_{JJ} as well as the difference in statistics due to the VBS cuts.

First, the m_J spectrum is fitted over the full m_{JJ} range by projecting the 2D histograms in Fig. 4.20 onto the m_J axis. The fit function consists of two Gaussian functions, one describing the peak at the W boson mass and one at the top quark mass, and an Error Function Exponential (Erf-Exp) describing the non-resonant part. The fit results together with simulated events are shown in Fig. 4.21. Then, the same fit is performed in bins of m_{JJ}, chosen such that it captures the dependence of m_J on m_{JJ} and that the statistics allow the fit to converge. The resulting fit parameters of the previous step are chosen as initial values and resulting fit parameters are interpolated.

The final step in modeling the $t\bar{t}$ background is to build the three-dimensional probability density function from the three one-dimensional pdfs. This procedure results in the pdfs for m_{J1} and m_{J2} as described above and the pdf for m_{JJ} follows the same procedure as is done for the V+jets background in the previous chapter. With the one dimensional pdfs given, the final 3D pdf is then built out of 6 components, one for each possible combination of resonance at the top quark mass, resonance at the W boson mass, and non-resonant: P_{TT}, P_{TW}, $P_{T-nonRes}$, P_{WW}, $P_{W-nonRes}$, and $P_{nonRes-nonRes}$:

$$P_{all} = P_{TT} + P_{WT} + P_{T-nonRes} + P_{WW} + P_{W-nonRes} + P_{nonRes-nonRes} \qquad (4.13)$$

This final pdf, $P_{TT,all}$, is then fit to the full $t\bar{t}$ simulation to obtain the prefit values for

each $t\bar{t}$ contribution that will enter the final fit for limit extraction. This is done because the generator-level matching efficiency results in a lower yield than what is expected. The resulting fits are shown in Fig. 4.22 and 4.23.
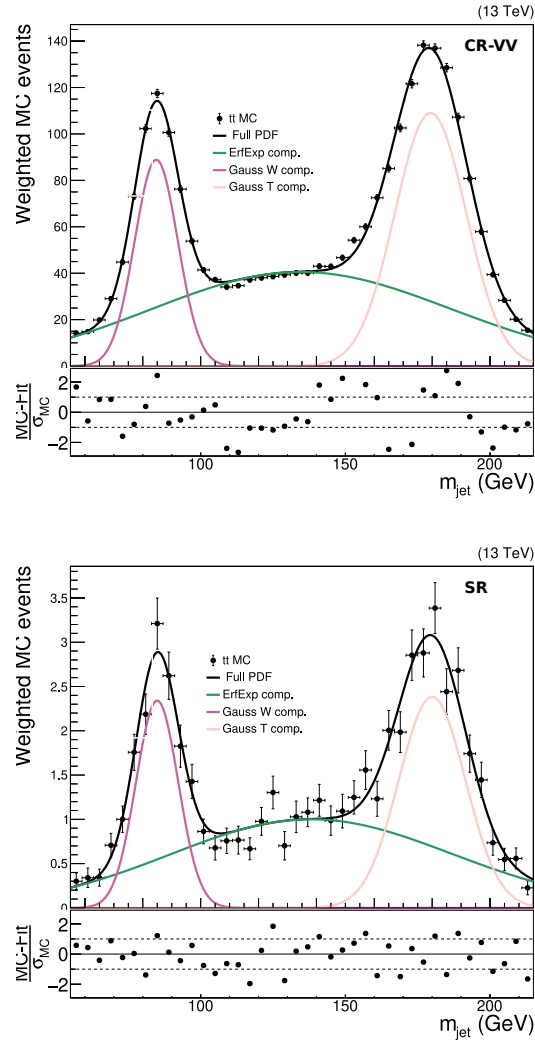
Figure 4.21.: Fit of the $t\bar{t}$ template consisting of all three contributions to the complete m_J distribution, when the selection criteria on AK4 jets are failed (left) and when they are passed (right).
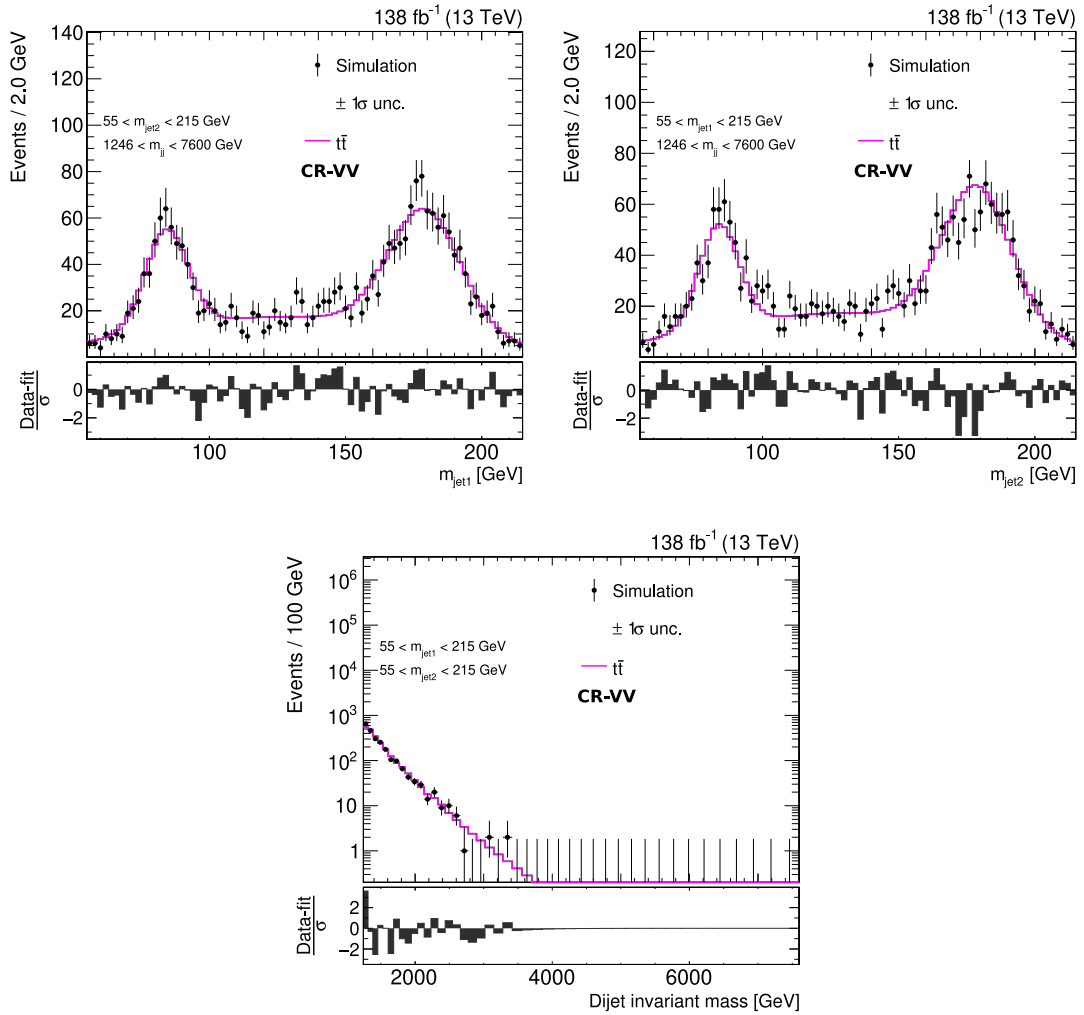
Figure 4.22.: Postfit distributions for the $t\bar{t}$ background parametrisation to simulated event in the CR-VV region to extract the final ratios of all six considered contributions as described in Eq. 4.13. The full three-dimensional template is fitted to simulated events and projected onto its axes one at a time.
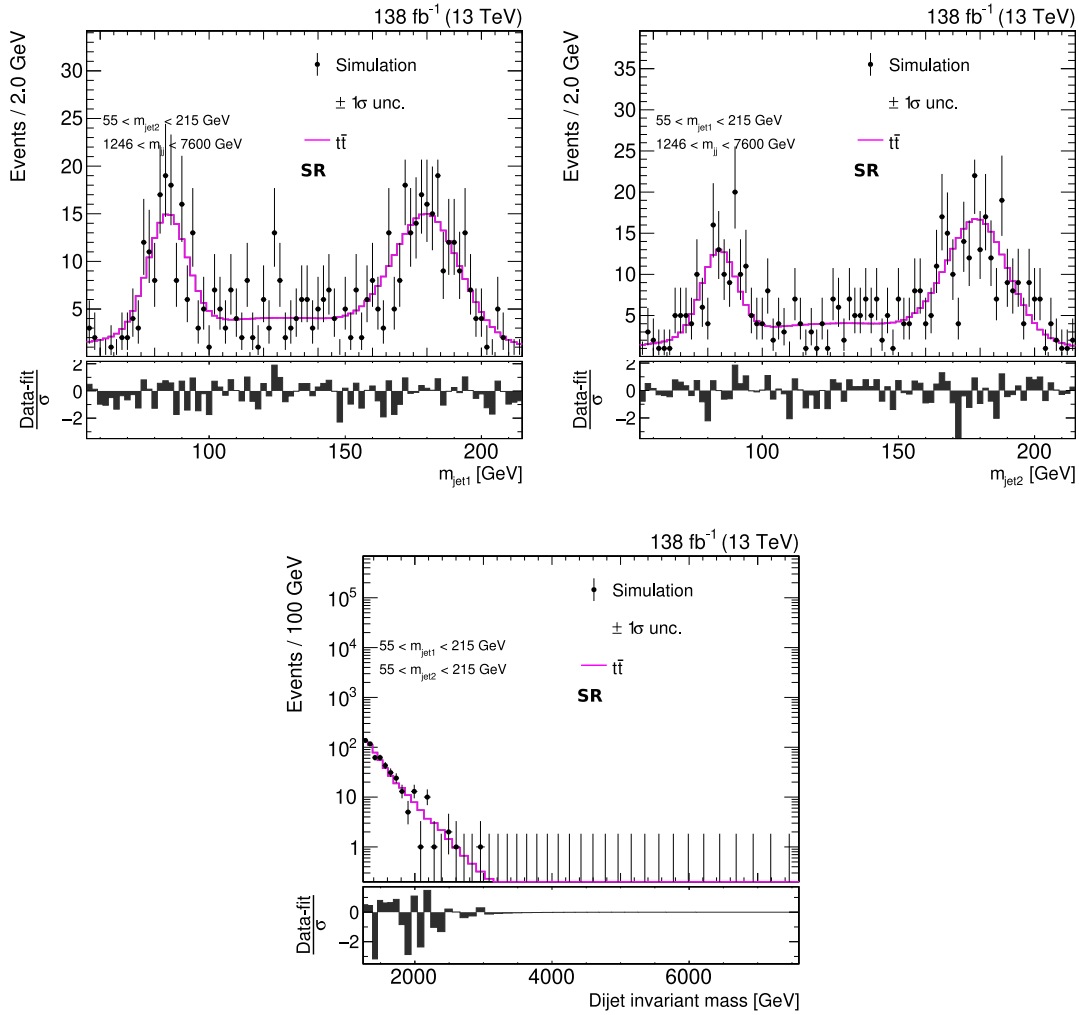
Figure 4.23.: Postfit distributions for the t$\bar{\text{t}}$ background parametrisation to simulated event in the SR to extract the final ratios of all six considered contributions as described in Eq. 4.13. The full three-dimensional template is fitted to simulated events and projected onto its axes one at a time.

4.4. Statistical Inference

With the signal and all background contributions established, the signal extraction is done by simultaneously fitting the three-dimensional distributions. This is done by constructing a likelihood function and following the procedure described in section 3.3. Systematic uncertainties are included for both signal and background contributions.

4.4.1. Systematic Uncertainties

Different sources of uncertainty are taken into account and introduced as nuisance parameters in the fit. They can affect the normalization or shape of signal and background processes.
The contributions to the final fit to data can be split into two categories: background processes estimated with data-driven methods, in this analysis only the QCD multijet background, and contributions estimated from MC simulation, here the signal process and all background processes other than QCD multijet production. Treatment for uncertainties affecting the data-driven QCD background estimation are completely different and described in section 4.3.1. Uncertainties affecting signal and background processes other than QCD multijet production are described in the following section.

Simulation Based Contributions

Multiple sources of uncertainty are considered in the final fit and are listed in the following. They affect all three dimensions in the 3D Fit, where they are estimated for non-resonant contributions by including alternative shapes in the fit and letting the processes float within the respective range. For the resonant contributions, the up- and down- variations affecting the shapes are estimated by fitting the respective distributions each time and including the variation in mean and width of the fit function as shape uncertainties. These are then implemented in the likelihood function with a Gaussian constraint. In general, uncertainties affecting the shape of distributions also change the overall normalization and are implemented taking into account asymmetric values for up- and down-variation of the corresponding nuisance parameter. Sources affecting only the normalization are included with log-normal constraints.
The luminosity-weighted sum of all years is fit together to reduce the impact of statistical fluctuations. All uncertainties are described in the following and listed in tables 4.10.

- **Luminosity**: A normalization uncertainty of 1.8% for the total Run 2 dataset is used, correlated among years [162–164].

- **Pileup**: An uncertainty is applied corresponding to the change in minimum bias cross-section by $\pm 4.6\%$:
$$\sigma_{\mathrm{MB}} = (69.2 \pm 3.2)\,\mathrm{pb}$$

Weights are assigned to simulated events as follows: In simulation, the number of true interactions n_{true} is drawn from a distribution slightly overestimating the distribution in data following the latest recommendations from the CMS Collaboration [165]. The ratio of normalized data to simulated MC histograms then results in an event weight depending on n_{true}.

Up- and down-variations are also propagated through the analysis such that pileup reweighting results in both alternative shapes and a normalization uncertainty. Since the uncertainty originates from above variation of the minimum bias cross-section, this uncertainty is correlated across processes and years.

- **Jet tagging efficiency**: By applying the up- and down- variation of the scale factors measured for ParticleNet-MD WvsQCD discriminator, an uncertainty for the jet tagging is estimated. Since the scale factors are dependent on p_{T}, this uncertainty also affects the shape. All scale factors with uncertainties are listed in tables 4.5–4.8.

- **L1 prefiring**: Timing information from muon stations is propagated to the corresponding trigger primitives, where the limited timing resolutions can lead to an incorrect assignment of L1 muon candidates to bunch crossings. Rules of the CMS L1 trigger system do not allow for the firing in consecutive bunch crossings. In 2016 and 2017, the time alignment in the ECAL endcap was not properly accounted such that a fraction ECAL deposits are reconstructed as belonging to the previous bunch crossing and thus effectively vetoing the event itself. This effect known as "L1 prefiring" generates an inefficiency of the L1 trigger system.

  The effect is measured using "un-prefirable" events, where the previous bunch crossings are forbidden by trigger rules and depends on η as it depends on the muon track finder. In this analysis, mainly events with forward jets are affected, which are used to target the VBS production.

  The L1 prefiring weights and up/down variations for muon prefiring as well as ECAL inefficiencies are applied following the recipe in Ref. [166] resulting in different shapes of background contributions.

- **PDF + α_{s}**: Uncertainties from PDFs are taken into account by reweighting the simulated MC events to the provided PDF sets: the default PDF set, "NNPDF31_nnlo_as_0118_nf_4_mc_hessian" [167], is used and each histogram for final limit extraction is filled 100 times with the corresponding variations. The root mean square in each bin is calculated and summed in quadrature with the variation of α_{s}. The uncertainty on the signal process is estimated from the SM EWK VBS samples.

  The resulting weights affect both shape and normalization, such that both has to be taken into account in the final fit.

- **Factorization and renormalization scale**: The renormalization and factorization scales are shifted up and down by a factor of two resulting in different weights. For each event, the envelope of the six allowed combinations, where both scales are not varied in opposite directions, is taken as uncertainty This also results in different shapes and normalizations, although their impact is sub-dominant.

- **Jet energy scale and resolution**: The largest uncertainties in this analysis come from corrections to the jet energy as described in section 3.2.3. This effect is estimated by varying up and down the jet energy scale (JES) and jet energy resolution (JER) by one standard deviation. This directly impacts the selection of events and requires a rerun of all contributions to the fit. It therefore impacts both shapes and normalization.

- **$t\bar{t}$ normalization:** The cross section for top quark pair production is given by Eq.4.2 and a conservative uncertainty of 6% is assigned to the normalization.

A summary of all assigned uncertainties is given in table 4.10, where in case of asymmetric uncertainties, the average is listed.

Table 4.10.: Summary of systematic uncertainties

Source	Shape	VBS EWK	QCD	W+Jets	Z+Jets	$t\bar{t}$	VV
$t\bar{t}$ normalization		-	-	-	-	1.06	-
Luminosity		1.018	1.018	1.018	1.018	1.018	1.018
L1prefiring		1.01	-	1.01	1.01	1.01	1.01
Pileup	✓	1.02	-	1.015	1.025	1.020	1.08
Jet tagging	✓	1.12	-	1.12	1.12	1.12	1.12
JER	✓	1.05	-	1.07	1.05	1.06	1.20
JES	✓	1.17	-	1.25	1.17	1.20	1.20
PDF + α_s	✓	1.08	-	1.03	1.03	1.06	1.08
Scale ($\mu_\mathrm{R}, \mu_\mathrm{F}$)	✓	1.01	-	1.19	1.20	1.06	1.09
QCD: norm.		-	1.5	-	-	-	-
QCD: $\propto m_\mathrm{JJ}$	✓	-	1	-	-	-	-
QCD: $\propto 1/m_\mathrm{JJ}$	✓	-	1	-	-	-	-
QCD: MG+Pythia8	✓	-	1	-	-	-	-

QCD Estimation from Data

Uncertainties for the QCD multijet background estimated from data are split into shape uncertainties only affecting the shape of the three-dimensional distribution, and into the normalization.

The normalization is derived simultaneously in the final fit for limit extraction. This is done by the enhanced range in the jet mass from 55 GeV to 215 GeV, while a large prefit uncertainty of 50% is assigned. The nominal value for the QCD normalization is taken from simulation and scaled with the ratio as estimated in the CR-QCD region but ultimately is constraint in the fit.

The shape of QCD multijet production is derived in section 4.3.1 together with three alternative shapes, each with a corresponding nuisance parameter.

4.4.2. Likelihood

Limit extraction is then done following the methods described in section 3.3.3, where the likelihood function is given by

$$L\left(\vec{\lambda}|\vec{k}\right) = \prod_{i}^{n_{\text{bins}}} \text{Poisson}\left(k_i|\lambda_i\left(\vec{\mu}, \vec{\vartheta}\right)\right) \prod_{j}^{n_{\text{nuis}}} p_j(\vartheta_j|\widetilde{\vartheta}_j), \tag{4.14}$$

with nuisance parameters $\vec{\vartheta}$ corresponding to uncertainties as introduced above with Gaussian functions as constraint terms p_i, model parameters $\vec{\mu}$ are in this case the Wilson coefficients, expected and observed measurements λ_i and k_i in bin i defined by a range in m_{JJ}, m_{J1}, and m_{J2}. The expected number of events is given by the sum of signal and background contribution:

$$\lambda = N_{\text{signal}}\left(\vec{\mu}, \vec{\vartheta}\right) + \sum_{n} N_{\text{bkg,n}}\left(\vec{\vartheta}\right). \tag{4.15}$$

Normalization uncertainties that do not affect the shapes affect the number of expected events with following a log-normal distribution, implemented in the Higgs combine tool [168] by multiplying with $\kappa_j^{\vartheta_j}$ with the value κ_j describing the strength as given in table 4.10. Shape uncertainties that simultaneously affect the normalization, are interpolated with a spline within one standard deviation and are taken to be linear outside.

The model parameters $\vec{\mu}$ are the Wilson coefficients and affect signal yield given by the quadratic function derived in section 4.10 and change the shape according to the function in Eq. 4.4.

4.5. Results

With the final likelihood set up, the limit extraction is done by a scan of $q_\mu = -2\Delta\mathrm{NLL}$ as defined in Eq. 3.16. This scan is done as a function of one EFT parameter at a time and is shown in Fig. 4.24 for the three dimension-6 operators considered. Horizontal lines at 3.94 and 1 for $-2\Delta\mathrm{NLL}$ correspond to 95 % and 68 % CL, respectively. Resulting expected limits on the Wilson coefficients of dimension-6 operators are listed in table 4.11.

No public analyses exist that set limits on coefficients of dimension-6 operators using VBS processes as signal and dimension-6 operators are often considered in analyses targeting diboson production, since they are expected to be more sensitive due to the larger cross section. However, a recent study [75] suggests that VBS processes can be used to extract limits with a good enough sensitivity that it is beneficial for a combination to utilize data taken by the CMS experiment to the best extent. The study combines leptonic and semi-leptonic decay channels on parton level only considering statistical uncertainties resulting in limits for the coefficient c_{HW} of similar order, and for the coefficints c_{HBox} and c_{W} more stringent by a factor of three. Considering that no systematic uncertainties have been included in the sensitivity study, the here presented results are very competitive.

Looking ahead, the analysis can be extended to a larger set of dimension-6 operators, when the impact on background processes is taken under control. Furthermore, it is possible to investigate correlations by extracting two dimensional limits or profiling of operators.

Table 4.11.: Resulting expected limits on the considered Wilson coefficients for dimension-6 operators.

	Lower limit (TeV^{-2})	Upper limit (TeV^{-2})
$c_{\mathrm{HBox}}/\Lambda^2$	-5.88	6.20
$c_{\mathrm{HW}}/\Lambda^2$	-1.30	1.30
c_{W}/Λ^2	-0.258	0.256
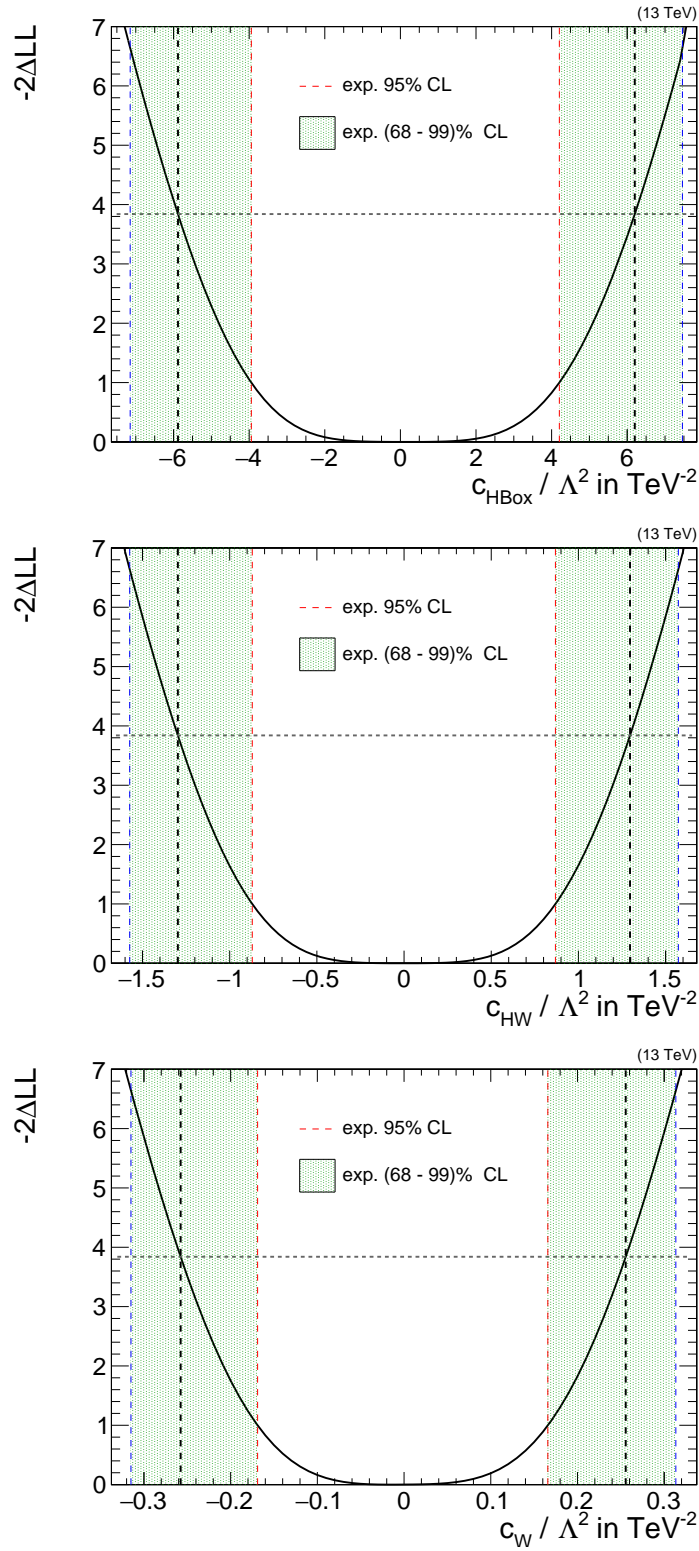
Figure 4.24.: Resulting ΔNLL scans for the three considered dimension-6 operators. Horizontal lines at 3.84 and 1 correspond to 95 % and 68 % CL, respectively.

Besides dimension-6 operators, also limits on dimension-8 operators have been set and can be found in table 4.12. Distributions of q_μ for all dimension-8 operators are shown in appendix C. When considering dimension-8 operators, it is not unambigously clear how to deal with unitarity violation for very high energy values. Since it is an active field of development and focus is on dimension-6 operators, where this problem is less relevant, it is not considered in this analysis.

Comparing the limits to public results of previous analyses and this decay channel promises to set stringent limits: In the semi-leptonic decay channel of VBS [15], the most sensitive limits to the coefficients $f_{S,i}/\Lambda^{-4}$, $f_{T,0,1,2}/\Lambda^{-4}$, and $f_{M,0,1,6,7}/\Lambda^{-4}$ are derived. Best current limits on the coefficiencts $f_{T,5,6,7}/\Lambda^{-4}$ and $f_{M,2-5}/\Lambda^{-4}$ are given in Ref. [16] and for $f_{T,8,9}/\Lambda^{-4}$ in Ref. [17]. The limits expected in this analysis are better for the operators $f_{S,i}/\Lambda^{-4}$ and of comparable size for the others. However, they utilize a separate subset of measured data and no limits are so far set on $f_{T,3}/\Lambda^{-4}$ and $f_{T,4}/\Lambda^{-4}$.

Table 4.12.: Resulting expected limits on the Wilson coefficients for the 20 considered dimension-8 operators.

	Lower limit (TeV^{-4})	Upper limit (TeV^{-4})		Lower limit (TeV^{-4})	Upper limit (TeV^{-4})
f_{S0}/Λ^4	-3.72	3.42	f_{T0}/Λ^4	-0.228	0.225
f_{S1}/Λ^4	-5.31	5.19	f_{T1}/Λ^4	-0.262	0.266
f_{S2}/Λ^4	-7.25	7.17	f_{T2}/Λ^4	-0.568	0.568
f_{M0}/Λ^4	-1.07	1.07	f_{T3}/Λ^4	-0.548	0.548
f_{M1}/Λ^4	-3.52	3.53	f_{T4}/Λ^4	-1.50	1.48
f_{M2}/Λ^4	-1.33	1.33	f_{T5}/Λ^4	-0.516	0.512
f_{M3}/Λ^4	-4.47	4.47	f_{T6}/Λ^4	-0.794	0.790
f_{M4}/Λ^4	-3.09	3.09	f_{T7}/Λ^4	-1.65	1.64
f_{M5}/Λ^4	-8.41	8.47	f_{T8}/Λ^4	-0.388	0.388
f_{M7}/Λ^4	-6.80	6.79	f_{T9}/Λ^4	-0.733	0.735

5. Summary

The last missing piece of the Standard Model (SM) of particle physics, the Higgs boson, was discovered in 2012 [4, 5]. In the following years, measurements of its properties and other parameters of the SM have been the focus of analyses. Additionally, direct searches to new physics beyond the SM (BSM) have shown results compatible with the SM. A separate approach is the interpretation of the SM as a low energy approximation of a more complete theory, to treat it as an effective field theory (EFT) and measure the effect of new phenomena present at an energy scale Λ, which is higher than currently accessible energies in any experiment. This is done by adding terms with higher mass-dimension, dimension-6 and dimension-8, to the SM Lagrangian.

The number of additional terms is large but only a subset affects the process under consideration: vector boson scattering (VBS). In VBS processes, triple and quartic couplings of heavy vector bosons, W or Z bosons, are studied, which reduces the number of interesting terms. In particular, the hadronic decay is investigated, where both vector bosons decay into pairs of quarks, and the branching ratio is largest. This results in a powerful channel to set limits on EFT contributions. A major challenge, however, is the large background from QCD multijet production, which is constrained by requiring the presence of two large-radius jets, each compatible with a W or Z boson, and by a fit in three dimensions: both jet masses of the large-radius jets, and the invariant mass of the jet pair. This fitting approach utilizes the falling distribution in QCD multijet production, while the signal shows a resonance at the mass of the vector boson. Furthermore, the invariant mass of the jet pair is sensitive to additional contributions due to the EFT approach.

In this thesis, sensitivity to limits on dimension-6 and dimension-8 operators is estimated by analyzing VBS in the hadronic decay channel in data collected by the CMS experiment corresponding to an integrated luminosity of $137\,\mathrm{fb}^{-1}$. The sensitivity is estimated considering one operator at a time resulting in the best limits for dimension-6 operators in VBS. For dimension-8 operators, sensitivities are of comparable order as other analyses but utilize a separate subset of data such that a combination that improves the limits is possible. Besides a combination with other analyses, the analysis can be extended by considering multiple operators at a time, simultaneously considering dimension-6 and dimension-8 operators, and increasing the number of investigated dimension-6 operators by handling their impact on background contributions.

A. Appendix: Data and Monte Carlo samples

A.1. List of Signal MC samples

Dim-6 Samples

All used samples of simulated with contributions from dimension-6 operators are listed in table A.1. The considered operators with maximal and minimal values in the reweighting are listed in table A.2.

Table A.1.: List of simulated signal samples: dimension-6 operators

MonteCarlo name	Events			
	2016preVFP	2016postVFP	2017	2018
WPWMjj_EWK_LO_dim6	1288110	1288110	1193108	1028684
WMWMjj_EWK_LO_dim6	1380139	1381045	1241127	1145664
WPWPjj_EWK_LO_dim6	1316421	1115341	1112190	1097577
ZZjj_EWK_LO_dim6	1270943	1270958	1237645	1137366
WPZjj_EWK_LO_dim6	1414169	1237645	1212976	1133218
WMZjj_EWK_LO_dim6	1251591	1234551	1207905	1177412

Table A.2.: Reweighting range of considered dimension-6 operators

Wilson coefficient	minimum	maximum
c_{HBox}	-7.5	7.5
c_{HW}	-2.5	2.5
c_{W}	-0.5	0.5

Dim-8 Samples

All used samples of simulated with contributions from dimension-8 operators are listed in table A.3. The considered operators with maximal and minimal values in the reweighting are listed in table A.4.

Table A.3.: List of simulated signal samples: dimension-8 operators

MonteCarlo name	Events			
	2016preVFP	2016postVFP	2017	2018
WminusTo2JWminusTo2JJJ_(*)	882000	1000000	996000	1000000
WminusTo2JZTo2JJJ_(*)	996000	990000	996000	984000
WplusTo2JWminusTo2JJJ_(*)	1037000	994000	970000	984000
WplusTo2JWplusTo2JJJ_(*)	995000	996000	984000	955000
WplusTo2JZTo2JJJ_(*)	995000	995000	966000	992000
ZTo2JZTo2JJJ_(*)	497468	496257	976900	988574
(*) = EWK_LO_NPle1_aQGC_TuneCP5_13TeV-madgraph-pythia8				

Table A.4.: Reweighting range of considered dimension-8 operators

	minimum	maximum		minimum	maximum
f_{S0}	-30.0	30.0	f_{T0}	-2.0	2.0
f_{S1}	-30.0	30.0	f_{T1}	-2.0	2.0
f_{S2}	-30.0	30.0	f_{T2}	-4.0	4.0
f_{M0}	-36.0	36.0	f_{T3}	-4.0	4.0
f_{M1}	-28.0	28.0	f_{T4}	-4.0	4.0
f_{M2}	-60.0	60.0	f_{T5}	-8.0	8.0
f_{M3}	-80.0	80.0	f_{T6}	-8.0	8.0
f_{M4}	-80.0	80.0	f_{T7}	-16.0	16.0
f_{M5}	-160.0	160.0	f_{T8}	-20.0	20.0
f_{M7}	-80.0	80.0	f_{T9}	-20.0	20.0

A.2. List of Simulated Background Samples

Table A.5.: List of simulated background samples for 2016preVFP.

Type		MonteCarlo name in DAS	Events
VBS EWK	ssWW	WplusminusTo2JWplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v3/MINIAODSIM	2500000
	osWW	WplusTo2JWminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v3/MINIAODSIM	2499000
	WZnob	ZTo0Jho8WplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	2448000
	WZbb	ZTo2BWplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	2499000
	ZZ	ZTo0JZTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	2461558
VBS QCD	ssWW	WplusminusTo2JWplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	5000000
	osWW	WplusTo2JWminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	1000000
	WZnob	ZTo0Jho8WplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	1000000
	WZbb	ZTo2BWplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	1000000
	ZZ	ZTo0JZTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	100000
QCD HT		QCD_HT200to300_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	18273591
		QCD_HT300to500_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	15226670
		QCD_HT500to700_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	56138559
		QCD_HT700to1000_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	15478579
		QCD_HT1000to1500_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	13679903
		QCD_HT1500to2000_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	9830970
		QCD_HT2000toInf_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	4923577
QCD_Pt		QCD_Pt_170to300_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	24691000
		QCD_Pt_300to470_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	54024000
		QCD_Pt_470to600_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	40618000
		QCD_Pt_600to800_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	56314000
		QCD_Pt_800to1000_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	33893000
		QCD_Pt_1000to1400_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	19077000
		QCD_Pt_1400to1800_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	11000000
		QCD_Pt_1800to2400_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	3332000
		QCD_Pt_2400to3200_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	2999000
		QCD_Pt_3200toInf_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	1000000
W+jets		WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	5121341
		WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	7573159
		WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	7488760
Z+jets		ZJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	7638105
		ZJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	6045332
		ZJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	3726992
TTbar		TT_Mtt-700to1000_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	22251771
		TT_Mtt-1000toInf_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	23366615
WW		WW_TuneCP5_13TeV-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	15739000
WZ		WZ_TuneCP5_13TeV-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	7910000
ZZ		ZZ_TuneCP5_13TeV-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	128000
singleTop		/ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	2300000
		/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v1/MINIAODSIM	2300000

Table A.6.: List of simulated background samples for 2016postVFP.

Type	MonteCarlo name in DAS	Events
VBS EWK		
ssWW	/WplusminusTo2JWplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	2498128
osWW	/WplusTo2JWminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	2498120
WZnob	/ZToJJnoBWplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v3/MINIAODSIM	2489638
WZbb	/ZTo2BWplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v3/MINIAODSIM	2491567
ZZ	/ZToJJZTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v3/MINIAODSIM	2458750
VBS QCD		
ssWW	/WplusminusTo2JWplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	1000000
osWW	/WplusTo2JWminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	1000000
WZnob	/ZToJJnoBWplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v3/MINIAODSIM	1000000
WZbb	/ZTo2BWplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v3/MINIAODSIM	1000000
ZZ	/ZToJJZTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v3/MINIAODSIM	1000000
QCD HT	/QCD_HT200to300_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	
	/QCD_HT300to500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	
	/QCD_HT500to700_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	
	/QCD_HT700to1000_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	
	/QCD_HT1000to1500_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	
	/QCD_HT1500to2000_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	
	/QCD_HT2000toInf_TuneCP5_PSWeights_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	
QCD_Pt	/QCD_Pt_170to300_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	29926000
	/QCD_Pt_300to470_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	55264000
	/QCD_Pt_470to600_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	52336000
	/QCD_Pt_600to800_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	64224000
	/QCD_Pt_800to1000_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	37446000
	/QCD_Pt_1000to1400_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	19868000
	/QCD_Pt_1400to1800_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	10722000
	/QCD_Pt_1800to2400_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	5236000
	/QCD_Pt_2400to3200_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	2848000
	/QCD_Pt_3200toInf_TuneCP5_13TeV_pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	996000
W+jets	/WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	4339896
	/WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	6775674
	/WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	6674335
Z+jets	/ZJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	6942718
	/ZJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	5405979
	/ZJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	4287246
WW	/WW_TuneCP5_13TeV-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	15796000
WZ	/WZ_TuneCP5_13TeV-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	7536000
ZZ	/ZZ_TuneCP5_13TeV-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	1151000
TTbar	/TT_Mtt-700to1000_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	33586554
	/TT_Mtt-1000toInf_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v1/MINIAODSIM	24177380
singleTop	/ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	2491000
	/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17-v2/MINIAODSIM	2554000

107

Table A.7.: List of simulated background samples for 2017.

Type	MonteCarlo name in DAS	Events
VBS EWK		
ssWW	/WplusminusTo2WplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	4985000
osWW	/WplusTo2WminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	4994000
WZnob	/ZTo0JnoBWplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	4973000
WZbb	/ZTo2BWplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v3/MINIAODSIM	4919000
ZZ	/ZTo0JZTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	4904000
VBS QCD		
ssWW	/WplusminusTo2WplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	1000000
osWW	/WplusTo2WminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	1000000
WZnob	/ZTo0JnoBWplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v3/MINIAODSIM	1000000
WZbb	/ZTo2BWplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	1000000
ZZ	/ZJJZJjj_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer19UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	1000000
QCD HT	/QCD_HT200to300_TuneCP5_PSWeights_13TeV-madgraph-pythia8/ RunIISummer20UL17MiniAODAPVv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	42433176
	/QCD_HT300to500_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODAPVv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	43100754
	/QCD_HT500to700_TuneCP5_PSWeights_13TeV-madgraph-pythia8/ RunIISummer20UL17MiniAODAPVv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	35816792
	/QCD_HT700to1000_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODAPVv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	33646855
	/QCD_HT1000to1500_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODAPVv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	10136610
	/QCD_HT1500to2000_TuneCP5_PSWeights_13TeV-madgraph-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	7678756
	/QCD_HT2000toInf_TuneCP5_13TeV-madgraph-pythia8/ RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	4089387
QCD_Pt	/QCD_Pt_170to300_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	29739000
	/QCD_Pt_300to470_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	55378000
	/QCD_Pt_470to600_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	50285000
	/QCD_Pt_600to800_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	67268000
	/QCD_Pt_800to1000_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	36458000
	/QCD_Pt_1000to1400_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	19685000
	/QCD_Pt_1400to1800_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	10994000
	/QCD_Pt_1800to2400_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	5488000
	/QCD_Pt_2400to3200_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	2997000
	/QCD_Pt_3200toInf_TuneCP5_13TeV_pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	976000
W+jets	/WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	9881339
	/WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	14338694
	/WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	14414746
Z+jets	/ZJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	14813420
	/ZJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	11442350
	/ZJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	9217675
TTbar	/TT_Mtt-700to1000_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	35862238
	/TT_Mtt-1000toInf_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	22724532
WW	/WW_TuneCP5_13TeV-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	15490000
WZ	/WZ_TuneCP5_13TeV-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v1/MINIAODSIM	7793000
ZZ	/ZZ_TuneCP5_13TeV-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	2706000
singleTop	/ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	5649000
	/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9-v2/MINIAODSIM	5674000

Table A.8.: List of simulated background samples for 2018.

Type	MonteCarlo name in DAS	Events
VBS EWK		
ssWW	/WPMJJWPMJJjj_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer19UL18MiniAODv2-106X_upgrade_realistic_v16_L1v1-v1/MINIAODSIM	993700
osWW	/WplusTo2JWminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer19UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	4991000
WZnob	/ZToJJnoBWplusminusTo2JJJ_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	2448000
WZbb	/ZBBWPMJJjj_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer19UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	987400
ZZ	/ZJJZJJjj_EWK_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer19UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	949600
VBS QCD		
ssWW	/WPMJJWPMJJjj_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer19UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	1000000
osWW	/WplusTo2JWminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	1000000
WZnob	/ZToJJnoBWplusminusTo2JJJ_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11-v2/MINIAODSIM	1000000
WZbb	/ZBBWPMJJjj_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer19UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	1000000
ZZ	/ZJJZJJjj_QCD_LO_TuneCP5_13TeV-madgraph-pythia8/RunIISummer19UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	1000000
QCD HT	/QCD_HT200to300_TuneCP5_PSWeights_13TeV-madgraph-pythia8/ RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	56417004
	/QCD_HT300to500_TuneCP5_PSWeights_13TeV-madgraph-pythia8/ RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	61063399
	/QCD_HT500to700_TuneCP5_PSWeights_13TeV-madgraph-pythia8/ RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	48805955
	/QCD_HT700to1000_TuneCP5_PSWeights_13TeV-madgraph-pythia8/ RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	48090314
	/QCD_HT1000to1500_TuneCP5_PSWeights_13TeV-madgraph-pythia8/ RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	14315177
	/QCD_HT1500to2000_TuneCP5_PSWeights_13TeV-madgraph-pythia8/ RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	10775587
	/QCD_HT2000toInf_TuneCP5_PSWeights_13TeV-madgraph-pythia8/ RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	5278880
QCD_Pt	/QCD_Pt_170to300_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	29292000
	/QCD_Pt_300to470_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	57718500
	/QCD_Pt_470to600_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	52352000
	/QCD_Pt_600to800_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	67484000
	/QCD_Pt_800to1000_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	36872000
	/QCD_Pt_1000to1400_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	19634000
	/QCD_Pt_1400to1800_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	10814000
	/QCD_Pt_1800to2400_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	5491000
	/QCD_Pt_2400to3200_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	2997000
	/QCD_Pt_3200toInf_TuneCP5_13TeV_pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	1000000
W+jets	/WJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	9172656
	/WJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	13469064
	/WJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	13437358
Z+jets	/ZJetsToQQ_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	13787944
	/ZJetsToQQ_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	11912371
	/ZJetsToQQ_HT-800toInf_TuneCP5_13TeV-madgraphMLM-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	9476143
WW	/WW_TuneCP5_13TeV-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	15463000
WZ	/WZ_TuneCP5_13TeV-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	7868000
ZZ	/ZZ_TuneCP5_13TeV-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	3502000
TTbar	/TT_Mtt-700to1000_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	30720345
	/TT_Mtt-1000toInf_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v1/MINIAODSIM	23758200
singleTop	/ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	7956000
	/ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8/RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1-v2/MINIAODSIM	7749000

A.3. List of Data Samples

Table A.9.: Data samples and corresponding luminosity.

Data Set	Integrated Luminosity [fb^{-1}]
/JetHT/Run2016B-ver1_HIPM_UL2016_MiniAODv2-v2	
/JetHT/Run2016B-ver2_HIPM_UL2016_MiniAODv2-v2	
/JetHT/Run2016C-HIPM_UL2016_MiniAODv2-v2	
/JetHT/Run2016D-HIPM_UL2016_MiniAODv2-v2	
/JetHT/Run2016E-HIPM_UL2016_MiniAODv2-v2	
/JetHT/Run2016F-HIPM_UL2016_MiniAODv2-v2	19.5
/JetHT/Run2016F-UL2016_MiniAODv2-v2	
/JetHT/Run2016G-UL2016_MiniAODv2-v2	
/JetHT/Run2016H-UL2016_MiniAODv2-v2	16.8
/JetHT/Run2017B-UL2017_MiniAODv2-v1	
/JetHT/Run2017C-UL2017_MiniAODv2-v1	
/JetHT/Run2017D-UL2017_MiniAODv2-v1	
/JetHT/Run2017E-UL2017_MiniAODv2-v1	
/JetHT/Run2017F-UL2017_MiniAODv2-v1	41.5
/JetHT/Run2018A-UL2018_MiniAODv2-v2	
/JetHT/Run2018B-UL2018_MiniAODv2_GT36-v1	
/JetHT/Run2018C-UL2018_MiniAODv2_GT36-v1	
/JetHT/Run2018D-UL2018_MiniAODv2_GT36-v1	59.7
Total integrated luminosity	137.16 fb^{-1}

B. Appendix: Trigger Plots

The following figures show the trigger efficiencies for all Run periods.
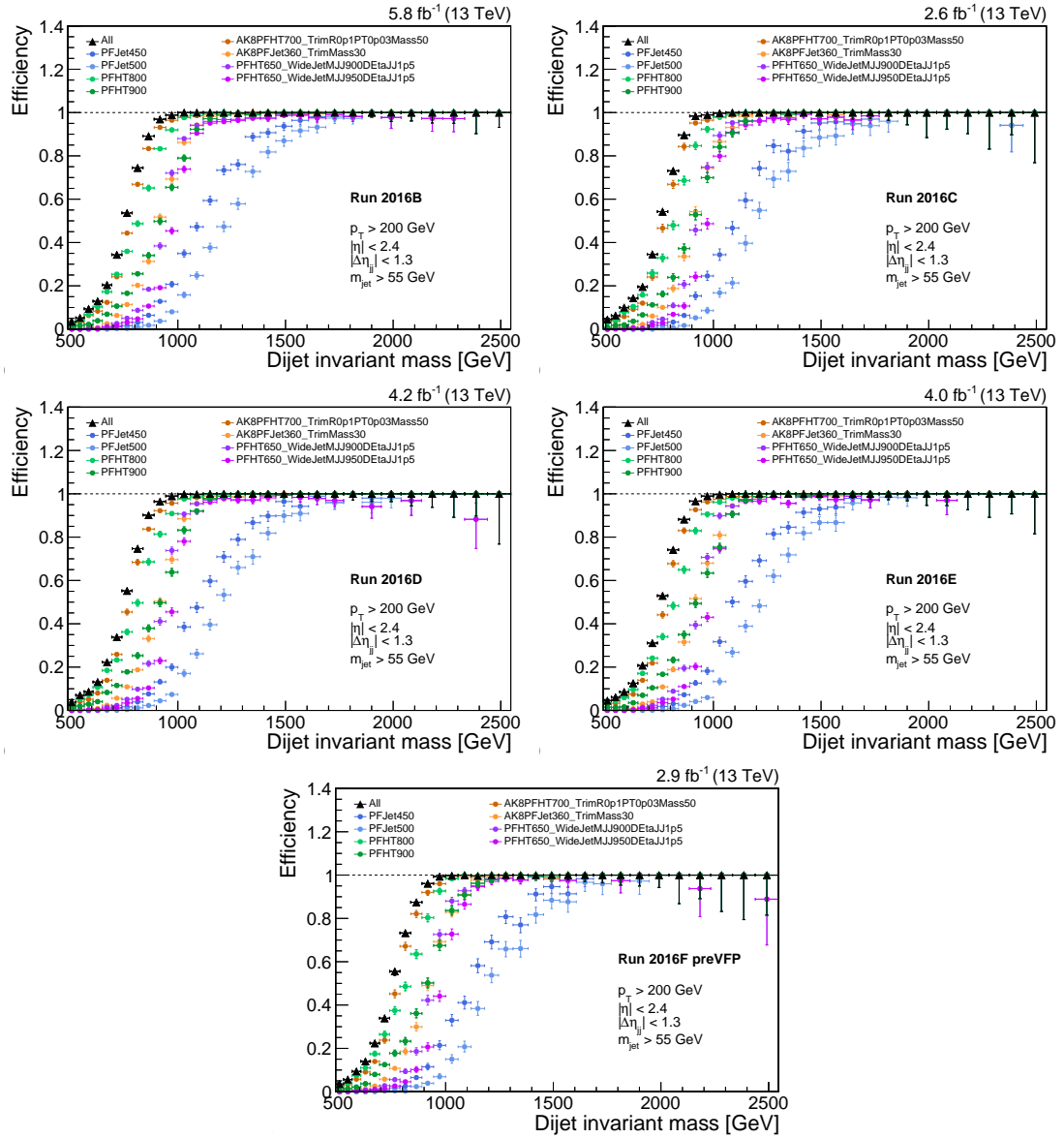
B.1. 2016preVFP

Figure B.1.: Trigger efficiencies for all Runs of 2016preVFP.

B.2. 2016postVFP
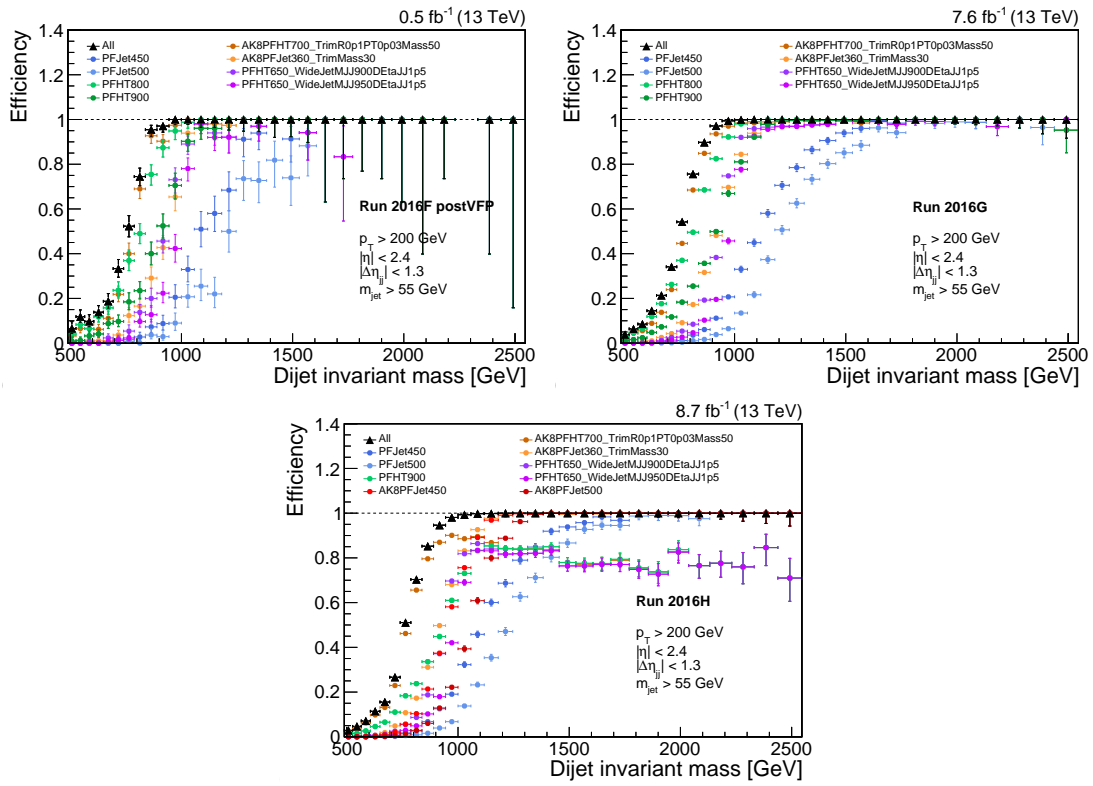
Figure B.2.: Trigger efficiencies for all Runs of 2016postVFP.

B.3. 2017
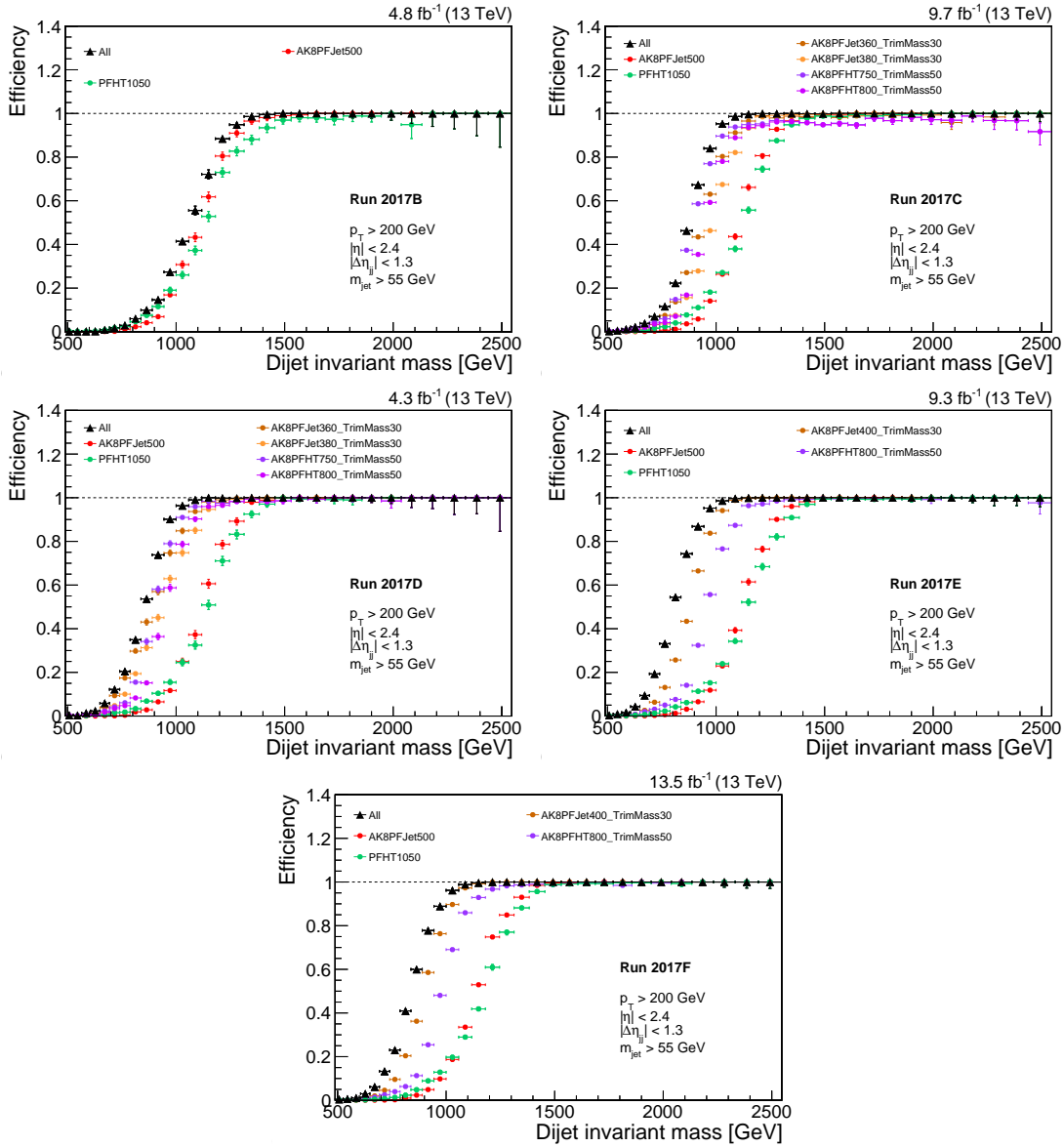
Figure B.3.: Trigger efficiencies for all Runs of 2017.

B.4. 2018
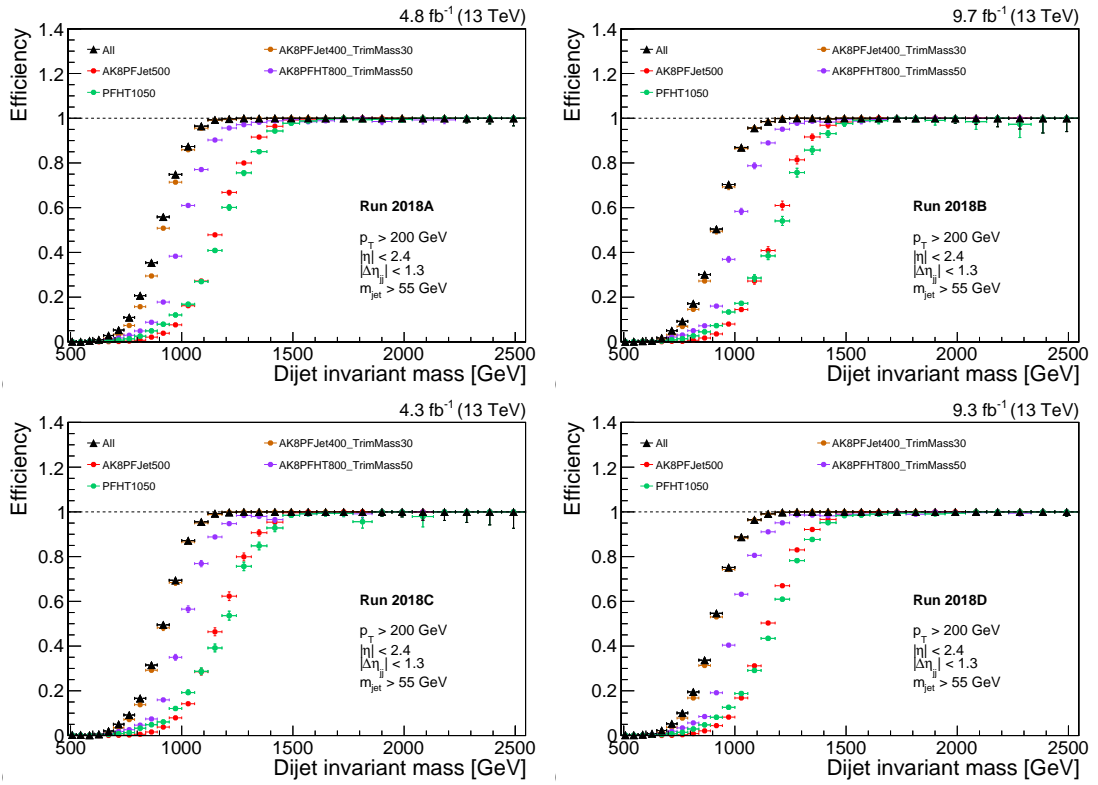
Figure B.4.: Trigger efficiencies for all Runs of 2018.

C. Appendix: Additional Plots for Dimension-8 Operators

This appendix lists additional plots for dimension-8 operators. They are not the main focus and the analysis is optimized for dimension-6 operators. The methods used to derive limits are the same for both kinds of operators.

C.1. Signal Parametrization

Here, additional information about the signal parametrization for all dimension-8 operators is listed. The mean and width of the DSCB function used to describe the AK8 jet mass is interpolated with a spline, which are shown in the following

(a)

(b)

(c)

(d)

Figure C.1.

(a)

(b)

(c)

(d)
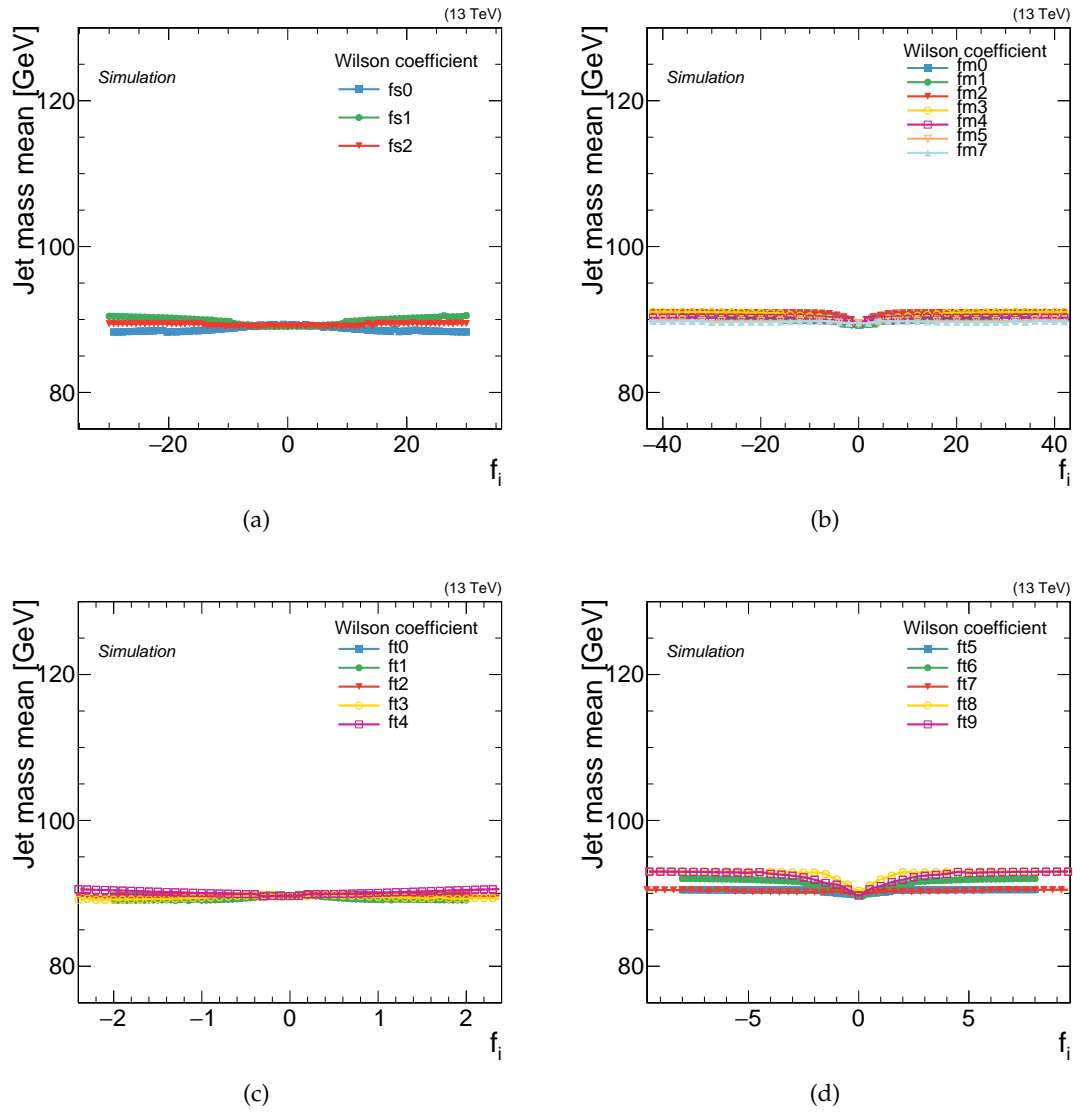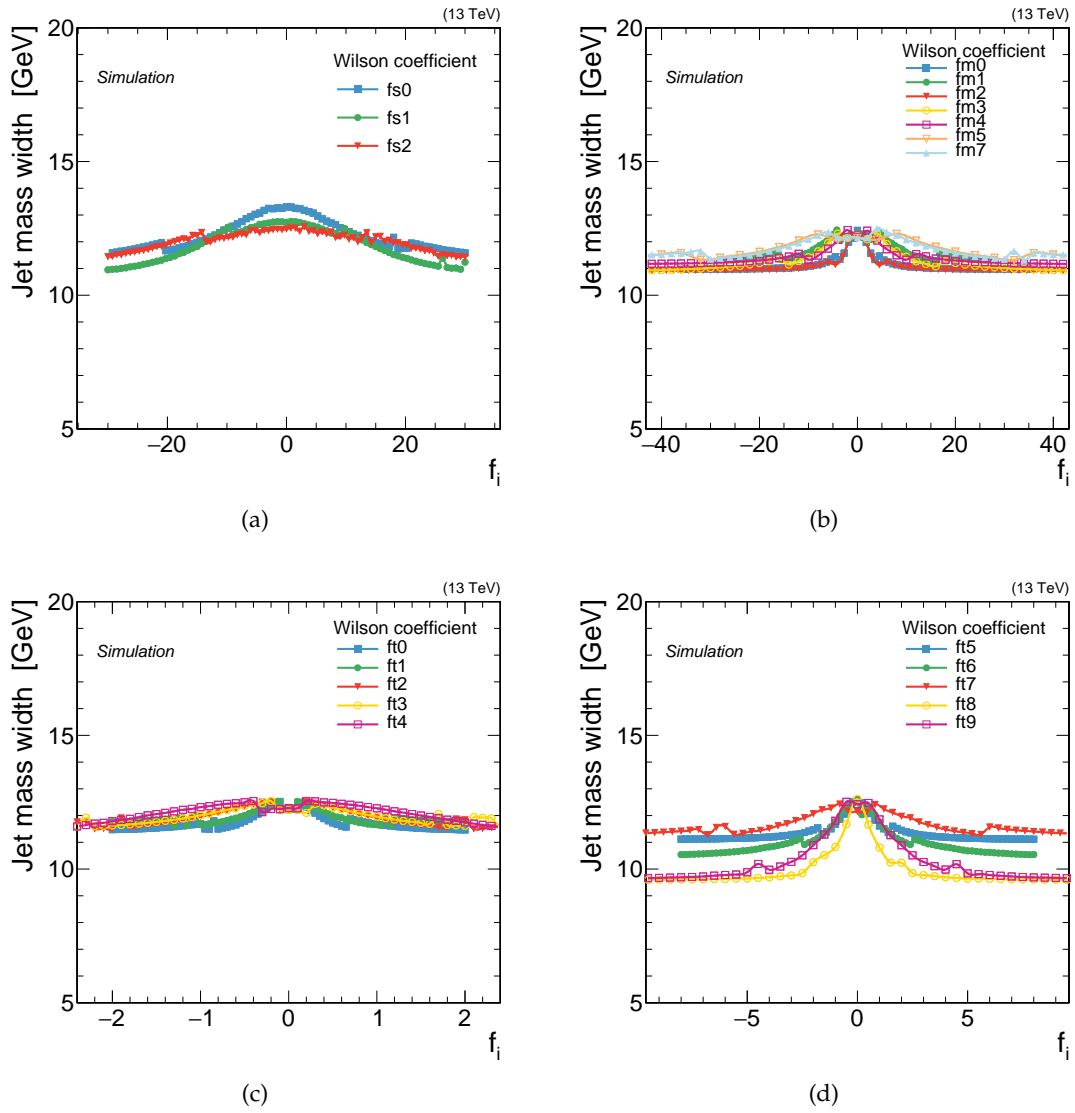
Figure C.2.

C.2. Likelihood Scans

This chapter shows additional likelihood scans for the 20 considered dimension-8 operators.
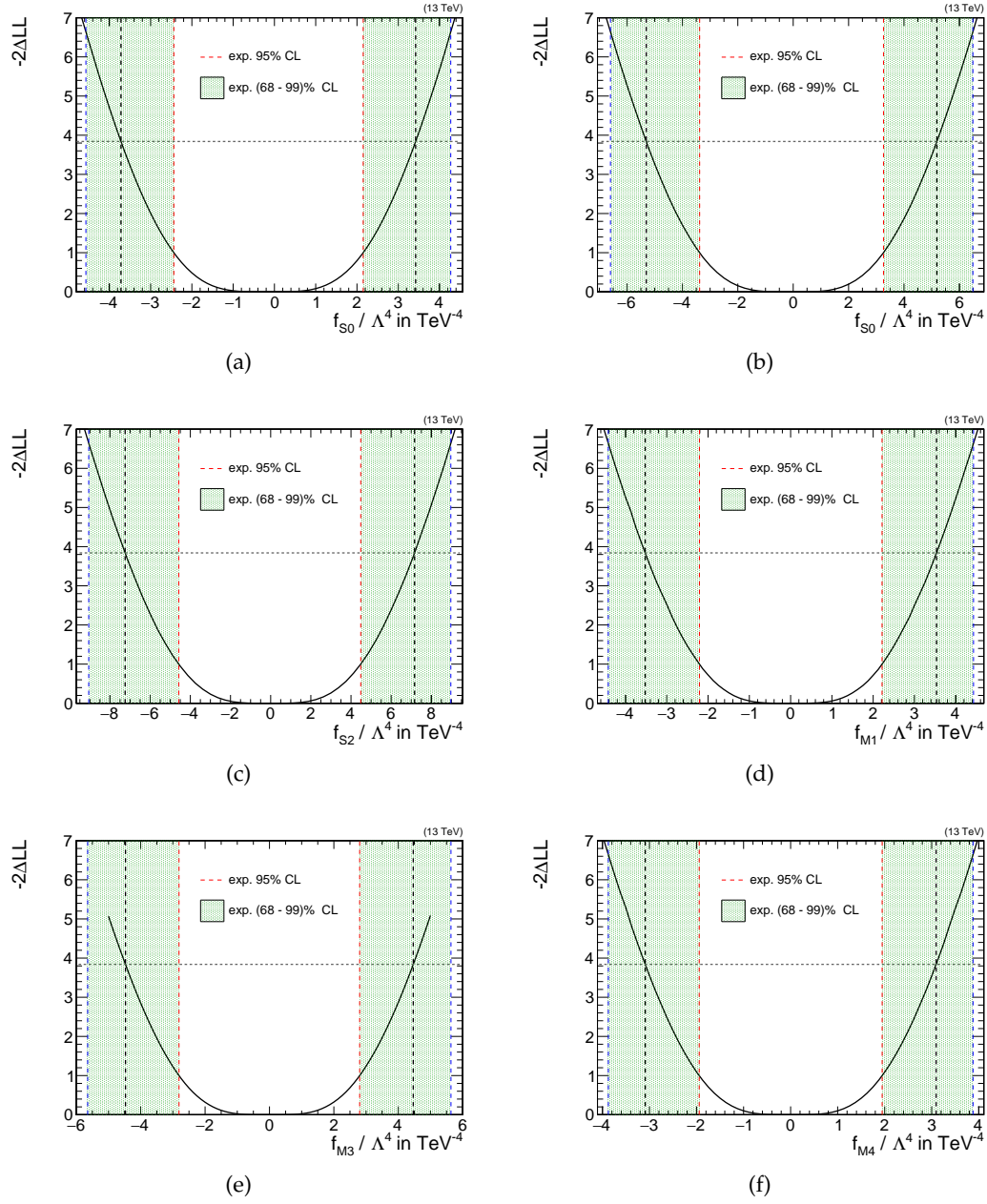
Figure C.3.: Likelihood scans for the fs parameters and fm1–fm4
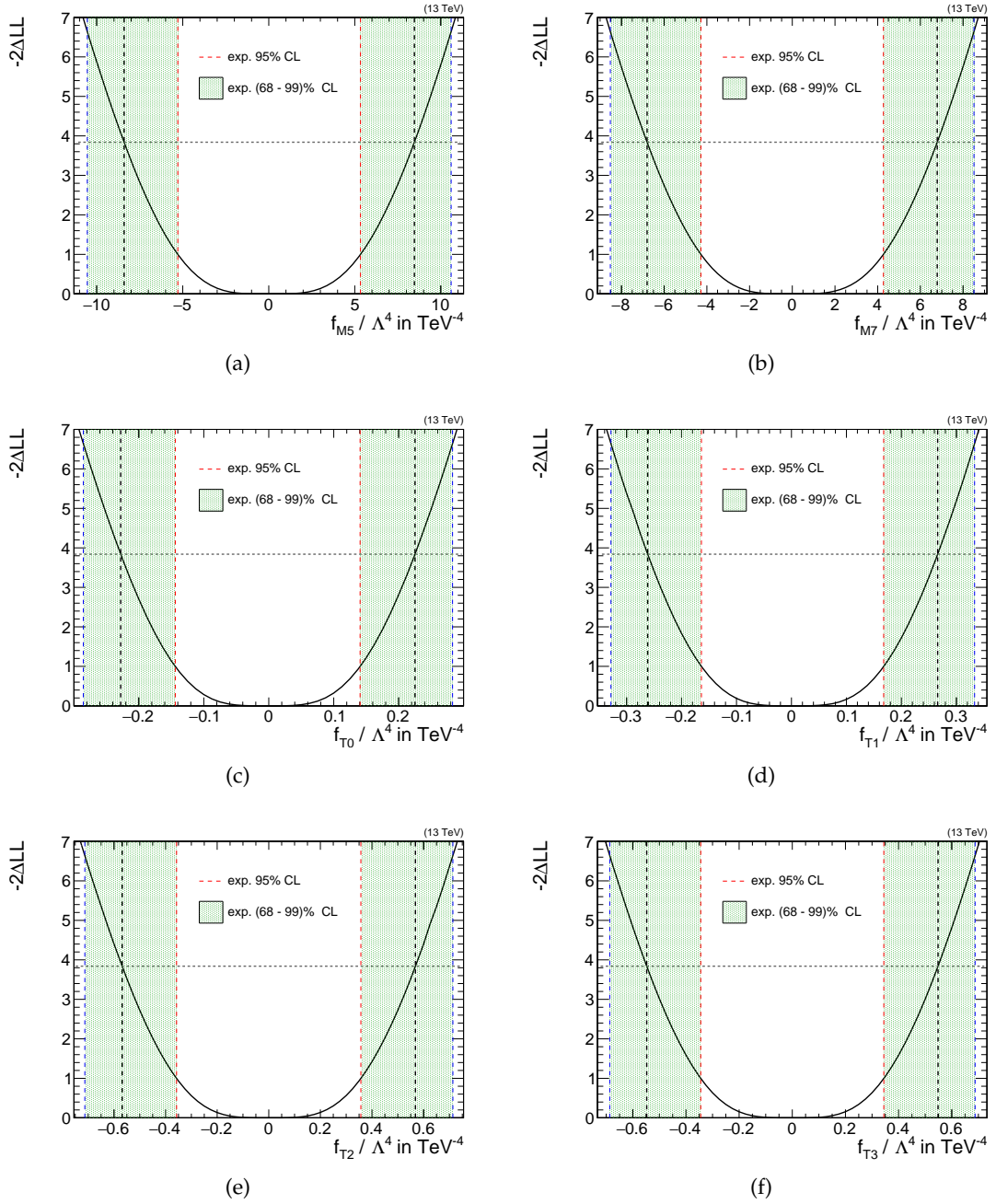
Figure C.4.: Likelihood scans for the fm5, fm7, and ft1–ft3 parameters

(a)

(b)

(c)

(d)

(e)

(f)
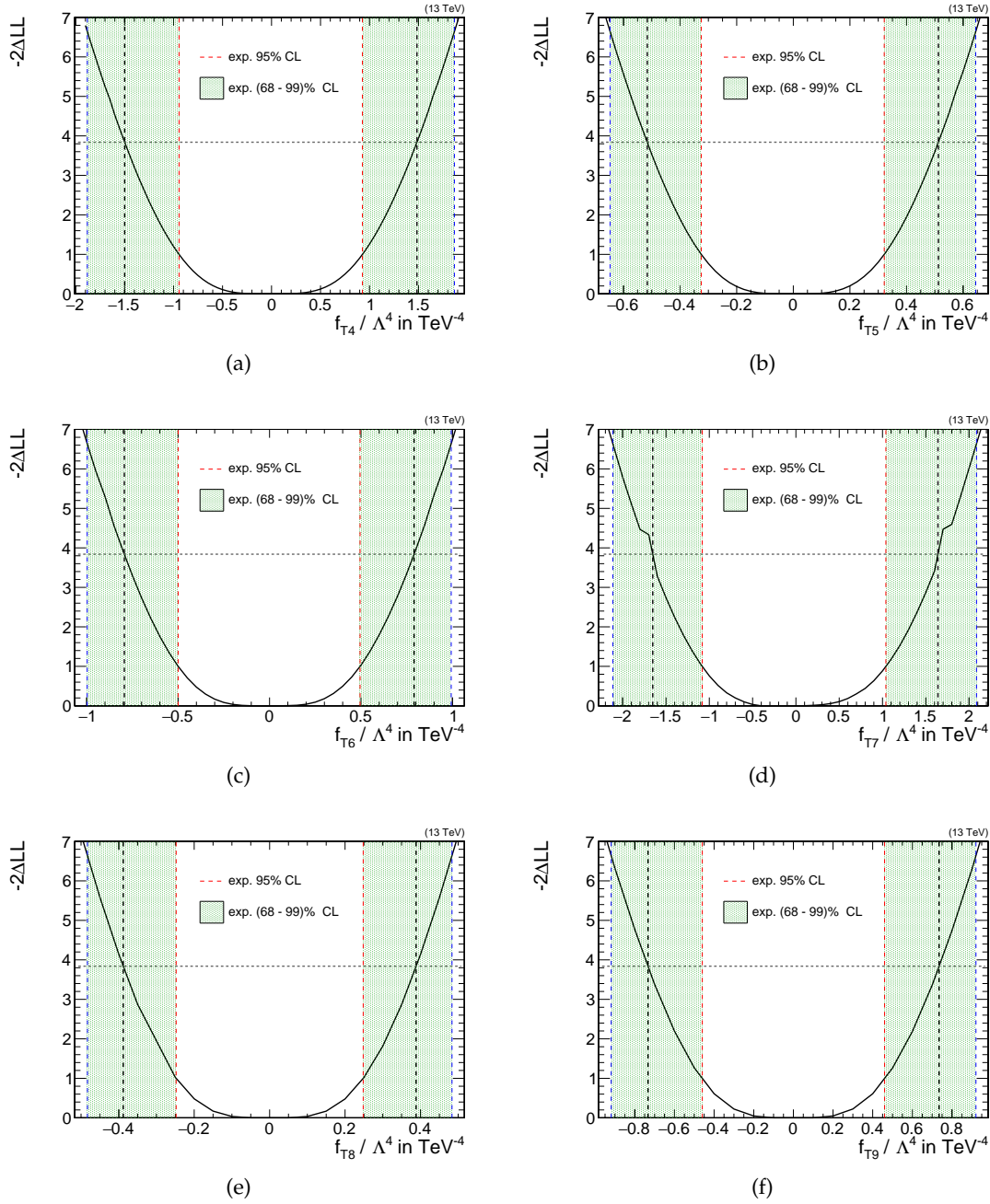
Figure C.5.: Likelihood scans for the ft4–ft9 parameters

List of Tables

List of Figures

Bibliography

[1] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* **19** (1967) 1264–1266, `doi:10.1103/PhysRevLett.19.1264`.

[2] A. Salam, "Weak and Electromagnetic Interactions", *Conf. Proc.* **C680519** (1968) 367–377, `doi:10.1142/9789812795915_0034`.

[3] S. L. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.* **22** (1961) 579–588, `doi:10.1016/0029-5582(61)90469-2`.

[4] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30–61, `doi:10.1016/j.physletb.2012.08.021`.

[5] 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** (2012) 1–29, `doi:10.1016/j.physletb.2012.08.020`.

[6] O. S. Brüning et al., "LHC Design Report". CERN Yellow Reports: Monographs. CERN, Geneva, 2004. CERN-2004-003-V-1. `doi:10.5170/CERN-2004-003-V-1`.

[7] CMS Collaboration, "The CMS Experiment at the CERN LHC", *J. Instrum.* **3** (2008) S08004, `doi:10.1088/1748-0221/3/08/S08004`.

[8] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *J. Instrum.* **3** (2008) S08003, `doi:10.1088/1748-0221/3/08/S08003`.

[9] ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002, `doi:10.1088/1748-0221/3/08/S08002`.

[10] LHCb Collaboration, "The LHCb Detector at the LHC", *J. Instrum.* **3** (2008) S08005, `doi:10.1088/1748-0221/3/08/S08005`.

[11] S. Weinberg, "Baryon- and lepton-nonconserving processes", *Phys. Rev. Lett.* **43** (1979) 1566–1570, `doi:10.1103/PhysRevLett.43.1566`.

[12] B. Grzadkowski, M. Iskrzyński, M. Misiak, and J. Rosiek, "Dimension-six terms in the Standard Model Lagrangian", *J. High Energy Phys.* **2010** (2010), no. 10, `doi:10.1007/jhep10(2010)085`.

[13] I. Brivio, Y. Jiang, and M. Trott, "The SMEFTsim package, theory and tools", *J. High Energy Phys.* **2017** (2017), no. 12, `doi:10.1007/jhep12(2017)070`.

[14] O. J. P. Éboli, M. C. Gonzalez-Garcia, and J. K. Mizukoshi, "$pp \to jje^{\pm}\mu^{\pm}\nu\nu$ and $jje^{\pm}\mu^{\mp}\nu\nu$ at $\mathcal{O}(\alpha_{\mathrm{em}}^6)$ and $\mathcal{O}(\alpha_{\mathrm{em}}^4\alpha_s^2)$ for the study of the quartic electroweak gauge boson vertex at CERN LHC", *Phys. Rev. D* **74** (2006) 073005, `doi:10.1103/PhysRevD.74.073005`.

[15] CMS Collaboration, "Search for anomalous electroweak production of vector boson pairs in association with two jets in proton-proton collisions at 13 TeV", *Phys. Lett. B* **798** (2019) 134985, `doi:10.1016/j.physletb.2019.134985`.

[16] CMS Collaboration, "Measurement of the electroweak production of Wγ in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", 2022. Submitted to: Phys. Rev. D. `doi:10.48550/arXiv.2212.12592`, `http://cds.cern.ch/record/2845242`.

[17] CMS Collaboration, "Measurement of the electroweak production of Zγ and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV and constraints on anomalous quartic gauge couplings", *Phys. Rev. D* **104** (2021) 072001, `doi:10.1103/PhysRevD.104.072001`.

[18] R. P. Feynman, "Space-time approach to quantum electrodynamics", *Phys. Rev.* **76** (1949) 769–789, `doi:10.1103/PhysRev.76.769`.

[19] E. Fermi, "Versuch einer Theorie der β-Strahlen", *Z. Phys.* **88** (1934), no. 3, 161–177, `doi:10.1007/BF01351864`.

[20] C. N. Yang, "Fermi's β-decay theory", *Int. J. Mod. Phys. A* **27** (2012), no. 03n04, 1230005, `doi:10.1142/S0217751X12300050`.

[21] H. Yukawa, "Quantum Theory of Non-Local Fields. Part I. Free Fields", *Phys. Rev.* **77** (1950) 219–226, `doi:10.1103/PhysRev.77.219`.

[22] M. Gell-Mann, "A schematic model of baryons and mesons", *Phys. Lett.* **8** (1964), no. 3, 214–215, `doi:10.1016/S0031-9163(64)92001-3`.

[23] C. N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance", *Phys. Rev.* **96** (1954) 191–195, `doi:10.1103/PhysRev.96.191`.

[24] J. Goldstone, "Field theories with superconductor solutions", *Il Nuovo Cimento* **19** (1961), no. 1, 154–164, `doi:10.1007/BF02812722`.

[25] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys. Lett.* **12** (1964) 132–133, `doi:10.1016/0031-9163(64)91136-9`.

[26] G. 't Hooft, "Renormalizable Lagrangians for Massive Yang-Mills Fields", *Nucl. Phys. B* **35** (1971) 167–188, `doi:10.1016/0550-3213(71)90139-8`.

[27] S. Weinberg, "The Making of the Standard Model", *Eur. Phys. J. C* **34** (2004) 5–13, `doi:10.1142/9789812567147_0005`.

[28] Novaes, S. F., "Standard Model: An Introduction", *Particles and fields. Proceedings, 10th Jorge Andre Swieca Summer School* (1999) 5–102, `doi:10.48550/arXiv.hep-ph/0001283`.

[29] D0 Collaboration, "Observation of the Top Quark", *Phys. Rev. Lett.* **74** (1995) 2632–2637, `doi:10.1103/PhysRevLett.74.2632`.

[30] CDF Collaboration, "Observation of top quark production in $\bar{p}p$ collisions", *Phys. Rev. Lett.* **74** (1995) 2626–2631, `doi:10.1103/PhysRevLett.74.2626`.

[31] K. Kodama et al., "Observation of tau neutrino interactions", *Phys. Lett. B* **504** (2001), no. 3, 218–224, `doi:10.1016/s0370-2693(01)00307-0`.

[32] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", *Phys. Rev. Lett.* **13** (1964) 321–323, `doi:10.1103/PhysRevLett.13.321`.

[33] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles", *Phys. Rev. Lett.* **13** (1964) 585–587, `doi:10.1103/PhysRevLett.13.585`.

[34] A. Einstein, "The Foundation of the General Theory of Relativity", *Ann. Phys.* **49** (1916), no. 7, 769–822, `doi:10.1002/andp.200590044`.

[35] E. Noether, "Invariant variation problems", *J. Comput. Theor. Nanosci.* **1** (1917), no. 3, 186–207, `doi:10.1080/00411457108231446`.

[36] KATRIN Collaboration, "Direct neutrino-mass measurement with sub-electronvolt sensitivity", *Nat. Phys.* **18** (2022), no. 2, 160–166, `doi:10.1038/s41567-021-01463-1`.

[37] Particle Data Group, "Review of Particle Physics", *Prog. Theor. Exp. Phys.* **2022** (2022), no. 8, `doi:10.1093/ptep/ptac097`.

[38] J. Ellis, "Higgs Physics", in *Proceedings, 2013 European School of High-Energy Physics.* 2015, pp. 117–168. `doi:10.5170/CERN-2015-004.117`.

[39] N. Cabibbo, "Unitary Symmetry and Leptonic Decays", *Phys. Rev. Lett.* **10** (1963) 531–533, `doi:10.1103/PhysRevLett.10.531`.

[40] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction", *Prog. Theor. Phys.* **49** (1973) 652–657, `doi:10.1143/PTP.49.652`.

[41] B. Pontecorvo, "Inverse beta processes and nonconservation of lepton charge", *Sov. Phys. JETP* **7** (1958) 172–173.

[42] Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the unified model of elementary particles", *Prog. Theor. Phys.* **28** (1962) 870–880, `doi:10.1143/PTP.28.870`.

[43] M. Rauch, "Vector-Boson Fusion and Vector-Boson Scattering". KA-TP-35-2016. `doi:10.48550/arXiv.1610.08420`.

[44] V. D. Duca, A. Frizzo, and F. Maltoni, "Higgs Boson Production in Association with Three Jets", *J. High Energy Phys.* **2004** (2004), no. 05, 064, `doi:10.1088/1126-6708/2004/05/064`.

[45] A. Ballestrero et al., "Precise predictions for same-sign W-boson scattering at the LHC", *Eur. Phys. J. C* **78** (2018), no. 8, `doi:10.1140/epjc/s10052-018-6136-y`.

[46] ATLAS Collaboration, "Evidence for electroweak production of $W^{\pm}W^{\pm}jj$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Rev. Lett.* **113** (2014) `doi:10.1103/PhysRevLett.113.141803`.

[47] ATLAS Collaboration, "Search for anomalous electroweak production of WW/WZ in association with a high-mass dijet system in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *Phys. Rev. D* **95** (2017) `doi:10.1103/PhysRevD.95.032001`.

[48] ATLAS Collaboration, "Measurement of $W^{\pm}W^{\pm}$ vector-boson scattering and limits on anomalous quartic gauge couplings with the ATLAS detector", *Phys. Rev. D* **96** (2017) `doi:10.1103/PhysRevD.96.012007`.

[49] ATLAS Collaboration, "Studies of $Z\gamma$ production in association with a high-mass dijet system in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector", *J. High Energy Phys.* **2017** (2017), no. 7, `doi:10.1007/jhep07(2017)107`.

[50] CMS Collaboration, "Study of vector boson scattering and search for new physics in events with two same-sign leptons and two jets", *Phys. Rev. Lett.* **114** (2015), no. 5, `doi:10.1103/PhysRevLett.114.051801`.

[51] CMS Collaboration, "Measurement of electroweak-induced production of $W\gamma$ with two jets in pp collisions at $\sqrt{s} = 8$ TeV and constraints on anomalous

quartic gauge couplings", *J. High Energy Phys.* **2017** (2017), no. 6,
doi:10.1007/jhep06(2017)106.

[52] CMS Collaboration, "Measurement of the cross section for electroweak
production of $Z\gamma$ in association with two jets and constraints on anomalous
quartic gauge couplings in proton–proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett.
B* **770** (2017) 380–402, doi:10.1016/j.physletb.2017.04.071.

[53] ATLAS Collaboration, "Observation of electroweak $W^\pm Z$ boson pair production
in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS
detector", *Phys. Lett. B* **793** (2019) 469–492,
doi:10.1016/j.physletb.2019.05.012.

[54] ATLAS Collaboration, "Observation of Electroweak Production of a Same-Sign
W Boson Pair in Association with Two Jets in pp Collisions at $\sqrt{s} = 13$ TeV with
the ATLAS Detector", *Phys. Rev. Lett.* **123** (2019) 161801,
doi:10.1103/PhysRevLett.123.161801.

[55] CMS Collaboration, "Measurement of electroweak WZ boson production and
search for new physics in WZ + two jets events in pp collisions at $\sqrt{s} = 13$ TeV",
Phys. Lett. B **795** (2019) 281–307, doi:10.1016/j.physletb.2019.05.042.

[56] CMS Collaboration, "Observation of Electroweak Production of Same-Sign W
Boson Pairs in the Two Jet and Two Same-Sign Lepton Final State in
Proton-Proton Collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. Lett.* **120** (2018) 081801,
doi:10.1103/PhysRevLett.120.081801.

[57] ATLAS Collaboration, "Observation of electroweak production of two jets and a
Z-boson pair", *Nat. Phys.* **19** (2023) 237–253,
doi:10.1038/s41567−022−01757−y.

[58] CMS Collaboration, "Evidence for electroweak production of four charged
leptons and two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B*
812 (2021) 135992, doi:10.1016/j.physletb.2020.135992.

[59] ATLAS Collaboration, "Observation of electroweak production of two jets in
association with an isolated photon and missing transverse momentum, and
search for a Higgs boson decaying into invisible particles at 13 TeV with the
ATLAS detector", *Eur. Phys. J. C* **82** (2022) 105,
doi:10.1140/epjc/s10052−021−09878−z.

[60] ATLAS Collaboration, "Measurement of the cross-section of the electroweak
production of a $Z\gamma$ pair in association with two jets in pp collisions at $\sqrt{s} = 13$
TeV with the ATLAS detector", 2023. Submitted to: Phys. Lett. B.
doi:10.48550/arXiv.2305.19142,
http://cds.cern.ch/record/2860529.

[61] CMS Collaboration, "Measurement of the electroweak production of $Z\gamma$ and two jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV and constraints on anomalous quartic gauge couplings", *Phys. Rev. D* **104** (2021), no. 7, `doi:10.1103/physrevd.104.072001`.

[62] CMS Collaboration, "Observation of electroweak production of $W\gamma$ with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **811** (2020) 135988, `doi:10.1016/j.physletb.2020.135988`.

[63] CMS Collaboration, "Observation of electroweak W^+W^- pair production in association with two jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *Phys. Lett. B* **841** (2023) 137495, `doi:10.1016/j.physletb.2022.137495`.

[64] ATLAS Collaboration, "Search for electroweak diboson production in association with a high-mass dijet system in semileptonic final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Phys. Rev. D* **100** (2019) 032007, `doi:10.1103/PhysRevD.100.032007`.

[65] CMS Collaboration, "Evidence for WW/WZ vector boson scattering in the decay channel $\ell\nu$qq produced in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **834** (2022) 137438, `doi:10.1016/j.physletb.2022.137438`.

[66] D. Buarque Franzosi et al., "Vector boson scattering processes: Status and prospects", *Reviews in Physics* **8** (2022) 100071, `doi:10.1016/j.revip.2022.100071`.

[67] M. U. Mozer, "Electroweak Physics at the LHC", volume 267 of *Springer Tracts Mod. Phys.* Springer, 2016. `doi:10.1007/978-3-319-30381-9`.

[68] M. Gallinaro et al., "Beyond the Standard Model in Vector Boson Scattering Signatures", 2020. DESY-PROC-2020-002. `doi:10.48550/arXiv.2005.09889`.

[69] R. Bellan et al., "VBSCan Thessaloniki 2018 Workshop Summary", 2019. DESY-19-108. `doi:10.48550/arXiv.1906.11332`.

[70] G. 't Hooft, "The Evolution of Quantum Field Theory, From QED to Grand Unification", *Adv. Ser. Direct. High Energy Phys.* **26** (2016) 1–27, `doi:10.1142/9789814733519_0001`.

[71] V. Trimble, "Existence and Nature of Dark Matter in the Universe", *Annu. Rev. Astron. Astrophys.* **25** (1987), no. 1, 425–472, `doi:10.1146/annurev.aa.25.090187.002233`.

[72] P. J. E. Peebles and B. Ratra, "The cosmological constant and dark energy", *Rev. Mod. Phys.* **75** (2003), no. 2, 559–606, `doi:10.1103/revmodphys.75.559`.

[73] S. Weinberg, "Phenomenological Lagrangians", *Physica* **A96** (1979), no. 1-2, 327–340, `doi:10.1016/0378-4371(79)90223-1`.

[74] C. Degrande et al., "Effective Field Theory: A Modern Approach to Anomalous Couplings", *Ann. Phys. (N. Y.)* **335** (2013) 21–32, `doi:10.1016/j.aop.2013.04.016`.

[75] R. Bellan et al., "A sensitivity study of VBS and diboson WW to dimension-6 EFT operators at the LHC", *J. High Energy Phys.* **2022** (2022) 39, `doi:10.1007/JHEP05(2022)039`.

[76] R. Alonso, E. E. Jenkins, A. V. Manohar, and M. Trott, "Renormalization group evolution of the Standard Model dimension six operators III: gauge coupling dependence and phenomenology", *J. High Energy Phys.* **2014** (2014) 159, `doi:10.1007/jhep04(2014)159`.

[77] N. Castro et al., "LHC EFT WG Report: Experimental Measurements and Observables", 2022. CERN-LHCEFTWG-2022-001. `doi:10.48550/arXiv.2211.08353`.

[78] I. Brivio, "SMEFTsim 3.0 — a practical guide", *J. High Energy Phys.* **2021** (2021), no. 4, `doi:10.1007/jhep04(2021)073`.

[79] C. W. Murphy, "Dimension-8 operators in the Standard Model Effective Field Theory", *J. High Energy Phys.* **2020** (2020), no. 10, `doi:10.1007/jhep10(2020)174`.

[80] E. da Silva Almeida, O.J.P.Éboli, and M.C.Gonzalez-Garcia, "Unitarity constraints on anomalous quartic couplings", *Phys. Rev. D* **101** (2020), no. 11, `doi:10.1103/physrevd.101.113003`.

[81] L. Evans and P. Bryant, "LHC Machine", *J. Instrum.* **3** (2008) S08001, `doi:10.1088/1748-0221/3/08/S08001`.

[82] T. Schörner-Sadenius, "The Large Hadron Collider—Background and History", pp. 1–26. Springer International Publishing, Cham, 2015. `doi:10.1007/978-3-319-15001-7_1`.

[83] CERN, "CERN Council - Ninety-Eighth Session", 1993. `https://indico.cern.ch/event/550126/`, accessed on 2023-04-28.

[84] CERN, "Resolution Approval of the Large Hadron Collider (LHC) Project. Résolution Approbation du projet de grand collisionneur de hadrons (LHC).

100th Session of Council", 1994. `https://cds.cern.ch/record/33618`, accessed on 2023-06-20.

[85] C. Campagnari and M. Franklin, "The Discovery of the top quark", *Rev. Mod. Phys.* **69** (1997) 137–212, `doi:10.1103/RevModPhys.69.137`.

[86] J. Vollaire et al., "Linac4 design report". CERN Yellow Reports: Monographs. CERN, Geneva, 2020. CERN-2020-006. `doi:10.23731/CYRM-2020-006`.

[87] K. Hanke, "Past and present operation of the cern ps booster", *Int. J. Mod.Phys. A* **28** (2013), no. 13, 1330019, `doi:10.1142/S0217751X13300196`.

[88] CERN, "Minutes of the Session. Procès-verbal. 7th Session of the European Council for Nuclear Research", 1953. `https://cds.cern.ch/record/18042`, accessed on 2023-06-20.

[89] UA1 Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540\,\text{GeV}$", *Phys. Lett. B* **122** (1983), no. 1, 103–116, `doi:10.1016/0370-2693(83)91177-2`.

[90] UA2 Collaboration, "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider", *Phys. Lett. B* **122** (1983), no. 5, 476–485, `doi:10.1016/0370-2693(83)91605-2`.

[91] E. Lopienska, "The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022", 2022. `https://cds.cern.ch/record/2800984`, accessed on 2023-06-20.

[92] CERN, "CAS - CERN Accelerator School: Intermediate Accelerator Physics: Zeuthen, Germany 15 - 26 Sep 2003. CAS - CERN Accelerator School: Intermediate Course on Accelerator Physics". CERN, Geneva, (2006). `doi:10.5170/CERN-2006-002`.

[93] CMS Collaboration, "Public CMS Luminosity Information", 2023. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`, accessed on 2023-06-20.

[94] CMS Collaboration, "The CMS: The Compact Muon Solenoid", 2022. `https://home.cern/science/experiments/cms`, accessed on 2023-06-20.

[95] CMS Collaboration, "Interactive Slice of the CMS detector", 2016. `https://cds.cern.ch/record/2205172`.

[96] V. Veszpremi, "Operation and performance of the CMS tracker", *J. Instrum.* **9** (2014), no. 03, C03005, `doi:10.1088/1748-0221/9/03/c03005`.

[97] CMS Collaboration, "The CMS Phase-1 Pixel Detector Upgrade", 2020.
`https://cds.cern.ch/record/2745805`.

[98] CMS Collaboration, "Performance of the CMS electromagnetic calorimeter and its role in the hunt for the Higgs boson in the two-photon channel", *J. Phys.: Conf. Ser.* **455** (2013) 012028, `doi:10.1088/1742-6596/455/1/012028`.

[99] CMS Collaboration, "The CMS Physics: Technical Design Report Volume 1: Detector Performance and Software", 2006. CERN-LHCC-2006-001 `https://cds.cern.ch/record/922757`.

[100] CMS Collaboration, "The CMS hadron calorimeter project: Technical Design Report", 1997. CERN-LHCC-97-031 `http://cds.cern.ch/record/357153`.

[101] CMS Collaboration, "Hadron Calorimeter", 2021. `http://cms.web.cern.ch/news/hadron-calorimeter`, accessed on 2023-06-20.

[102] H. S. Budd, "The CMS Central Hadron Calorimeter", *Nucl. Phys. B Proc. Suppl.* **54** (1997), no. 3, 191–197, `doi:10.1016/S0920-5632(97)00104-7`.

[103] CMS Collaboration, "The CMS muon project". Technical Design Report CMS. CERN, Geneva, 1997. CERN-LHCC-97-032 `https://cds.cern.ch/record/343814`.

[104] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *J. Instrum.* **13** (2018), no. 06, P06015, `doi:10.1088/1748-0221/13/06/P06015`.

[105] CMS Collaboration, "The CMS computing: Technical Design Report", 2005. `http://cds.cern.ch/record/838359`.

[106] WLCG Office, "WLCG (Worldwide LHC Computing Grid) - Tier centres". `wlcg-public.web.cern.ch/tier-centres`, accessed on 2023-06-20.

[107] B. Martin et al., "Rucio: Scientific data management", *Comput. Softw. Big Sci.* **3** (2019), no. 1, 11, `doi:10.1007/s41781-019-0026-3`.

[108] M. A. Dobbs et al., "Les Houches guidebook to Monte Carlo generators for hadron collider physics", in *3rd Les Houches Workshop on Physics at TeV Colliders.* 2004, pp. 411–459. `doi:10.48550/arXiv.hep-ph/0403045`.

[109] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.", *Sov. Phys. JETP* **46** (1977) 641–653.

[110] V. Gribov and L. Lipatov, "Deep inelastic electron scattering in perturbation theory", *Phys. Lett. B* **37** (1971), no. 1, 78–80, `doi:10.1016/0370-2693(71)90576-4`.

[111] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nucl. Phys. B* **126** (1977) 298–318, `doi:10.1016/0550-3213(77)90384-4`.

[112] NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, `doi:10.1007/JHEP04(2015)040`.

[113] Particle Data Group Collaboration, "Review of Particle Physics", *Chin. Phys. C* **40** (2016), no. 10, 100001, `doi:10.1088/1674-1137/40/10/100001`.

[114] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159–177, `doi:10.1016/j.cpc.2015.01.024`.

[115] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *J. High Energy Phys.* **05** (2006) 026, `doi:10.1088/1126-6708/2006/05/026`.

[116] V. V. Sudakov, "Vertex parts at very high-energies in quantum electrodynamics", *Sov. Phys. JETP* **3** (1956) 65–71.

[117] S. Gieseke, "Uncertainties of sudakov form factors", *J. High Energy Phys.* **2005** (2005), no. 01, 058–058, `doi:10.1088/1126-6708/2005/01/058`.

[118] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-pair production in hadronic collision", *J. High Energy Phys.* **2007** (2007), no. 01, 013, `doi:10.1088/1126-6708/2007/01/013`.

[119] B. Andersson et al., "Parton Fragmentation and String Dynamics", *Phys. Rep.* **97** (1983) 31–145, `doi:10.1016/0370-1573(83)90080-7`.

[120] CMS Collaboration, "Observation of same-sign WW production from double parton scattering in proton-proton collisions at $\sqrt{s}$ = 13 TeV", 2022. Submitted to: Phys. Rev. Lett. `doi:10.48550/arXiv.2206.02681`, `https://cds.cern.ch/record/2811478`.

[121] CMS Collaboration, "Extraction and validation of a new set of CMS pythia8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020), no. 1, 4, `doi:10.1140/epjc/s10052-019-7499-4`.

[122] ATLAS Collaboration, "Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s}$ = 13 TeV with the ATLAS Detector at the LHC", *Phys. Rev. Lett.* **117** (2016), no. 18, `doi:10.1103/physrevlett.117.182002`.

[123] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV", *J. High Energy Phys.* **2018** (2018), no. 7, 161, `doi:10.1007/jhep07(2018)161`.

[124] TOTEM Collaboration, "First measurement of elastic, inelastic and total cross-section at $\sqrt{s} = 13$ TeV by TOTEM and overview of cross-section data at LHC energies", *Eur. Phys. J. C* **79** (2019), no. 2, 103, `doi:10.1140/epjc/s10052-019-6567-0`.

[125] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth.* **A506** (2003) 250–303, `doi:10.1016/S0168-9002(03)01368-8`.

[126] DELPHES 3 Collaboration, "DELPHES 3, A modular framework for fast simulation of a generic collider experiment", *J. High Energy Phys.* **02** (2014) 057, `doi:10.1007/JHEP02(2014)057`.

[127] A. Alloul et al., "FeynRules 2.0 - A complete toolbox for tree-level phenomenology", *Comput. Phys. Commun.* **185** (2014), no. 8, 2250–2300, `doi:10.1016/j.cpc.2014.04.012`.

[128] C. Degrande et al., "UFO - The Universal FeynRules Output", *Comput. Phys. Commun.* **183** (2012), no. 6, 1201–1214, `doi:10.1016/j.cpc.2012.01.022`.

[129] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *J. High Energy Phys.* **07** (2014) 079, `doi:10.1007/JHEP07(2014)079`.

[130] J. Alwall et al., "MadGraph 5 : Going Beyond", *J. High Energy Phys.* **2011** (2011) 128, `doi:10.1007/JHEP06(2011)128`.

[131] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations", *J. High Energy Phys.* **06** (2002), no. 06, 029, `doi:10.1088/1126-6708/2002/06/029`.

[132] J. Alwall et al., "A Standard format for Les Houches event files", *Comput. Phys. Commun.* **176** (2007), no. 4, 300–304, `doi:10.1016/j.cpc.2006.11.010`.

[133] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms", *J. High Energy Phys.* **2004** (2004), no. 11, 040, `doi:10.1088/1126-6708/2004/11/040`.

[134] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *J. High Energy Phys.* **2007** (2007), no. 11, 070, `doi:10.1088/1126-6708/2007/11/070`.

[135] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower monte carlo programs: the POWHEG BOX", *J. High Energy Phys.* **2010** (2010), no. 6, 43, `doi:10.1007/jhep06(2010)043`.

[136] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *J. Instrum.* **12** (2017), no. 10, P10003, `doi:10.1088/1748-0221/12/10/p10003`.

[137] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC", *J. Phys. G: Nucl. Part. Phys.* **31** (2005), no. 9, N9–N20, `doi:10.1088/0954-3899/31/9/n01`.

[138] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *J. High Energy Phys.* **04** (2008) 063, `doi:10.1088/1126-6708/2008/04/063`.

[139] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, `doi:10.1140/epjc/s10052-012-1896-2`.

[140] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett. B* **641** (2006) 57–61, `doi:10.1016/j.physletb.2006.08.037`.

[141] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *J. Instrum.* **12** (2017), no. 02, P02014, `doi:10.1088/1748-0221/12/02/p02014`.

[142] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *J. Instrum.* **6** (2011), no. 11, P11002, `doi:10.1088/1748-0221/6/11/p11002`.

[143] D. Bertolini, P. Harris, M. Low, and N. Tran, "Pileup per particle identification", *J. High Energy Phys.* **2014** (2014), no. 10, `doi:10.1007/jhep10(2014)059`.

[144] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *J. Instrum.* **15** (2020), no. 09, P09018, `doi:10.1088/1748-0221/15/09/p09018`.

[145] CMS Collaboration, "Pileup-per-particle identification: optimisation for Run 2 Legacy and beyond", 2021. CMS-DP-2021-001 `https://cds.cern.ch/record/2751563`.

[146] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft drop", *J. High Energy Phys.* **2014** (2014), no. 5, 146, `doi:10.1007/jhep05(2014)146`.

[147] J. Thaler and K. V. Tilburg, "Identifying boosted objects with N-subjettiness", *J. High Energy Phys.* **2011** (2011), no. 3, 15, `doi:10.1007/jhep03(2011)015`.

[148] H. Qu and L. Gouskos, "Jet tagging via particle clouds", *Phys. Rev. D* **101** (2020), no. 5, 056019, `doi:10.1103/physrevd.101.056019`.

[149] CMS Collaboration, "Identification of highly Lorentz-boosted heavy particles using graph neural networks and new mass decorrelation techniques", 2020. CMS-DP-2020-002 `http://cds.cern.ch/record/2707946`.

[150] Y. Wang et al., "Dynamic graph cnn for learning on point clouds", *ACM Trans. Graph.* **38** (2019), no. 5, 1–12, `doi:10.1145/3326362`.

[151] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses", *Annals Math. Statist.* **9** (1938), no. 1, 60–62, `doi:10.1214/aoms/1177732360`.

[152] CMS Collaboration, "Measurements of production cross sections of WZ and same-sign WW boson pairs in association with two jets in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Lett. B* **809** (2020) 135710, `doi:10.1016/j.physletb.2020.135710`.

[153] CMS Collaboration, "A multi-dimensional search for new heavy resonances decaying to boosted WW, WZ, or ZZ boson pairs in the dijet final state at 13 TeV", *Eur. Phys. J. C* **80** (2020), no. 3, 237, `doi:10.1140/epjc/s10052-020-7773-5`.

[154] CMS Collaboration, "Search for new heavy resonances decaying to WW, WZ, ZZ, WH, or ZH boson pairs in the all-jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV", 2022. Submitted to: Phys. Lett. B. `doi:10.48550/arXiv.2210.00043`, `https://cds.cern.ch/record/2835154`.

[155] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders", *Comput. Phys. Commun.* **185** (2014) 2930–2938, `doi:10.1016/j.cpc.2014.06.021`.

[156] CMS Collaboration, "NNLO+NNLL top-quark-pair cross sections". `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO`, accessed on 2023-06-19.

[157] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *J. Instrum.* **16** (2021), no. 05, P05014, `doi:10.1088/1748-0221/16/05/p05014`.

[158] CMS Collaboration, "Jet Energy Resolution", 2023. `https://twiki.cern.ch/twiki/bin/view/CMS/JetResolution`, accessed on 2023-06-20.

[159] CMS Collaboration, "ParticleNet SF for UL", 2022. `https://indico.cern.ch/event/1221195/#4-particlenet-sf-for-ul`, accessed on 2023-06-20.

[160] CMS Collaboration, "Search for anomalous couplings in boosted WW/WZ to $\ell\nu q\bar{q}$ production in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **772** (2017) 21–42, `doi:10.1016/j.physletb.2017.06.009`.

[161] CMS Collaboration, "Search for anomalous triple gauge couplings in WW and WZ production in lepton + jet events in proton-proton collisions at $\sqrt{s} = 13$ TeV", *J. High Energy Phys.* **2019** 62, `doi:10.1007/JHEP12(2019)062`.

[162] CMS Collaboration, "Precision luminosity measurement in proton–proton collisions at $\sqrt{s} = 13$ TeVin 2015 and 2016 at CMS", *Eur. Phys. J. C* **81** (2021), no. 09, 800, `doi:10.1140/epjc/s10052-021-09538-2`.

[163] CMS Collaboration, "CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV", 2018. `https://cds.cern.ch/record/2621960`.

[164] CMS Collaboration, "CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV", 2019. `https://cds.cern.ch/record/2676164`.

[165] CMS Collaboration, "Pileup Scenarios for Run 2 (and Run 3)", 2022. `https://twiki.cern.ch/twiki/bin/view/CMS/PileupScenariosRun2`, accessed on 2023-06-20.

[166] CMS Collaboration, "Reweighting recipe to emulate Level 1 ECAL and Muon prefiring",. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1PrefiringWeightRecipe`, accessed on 2023-06-20.

[167] R. D. Ball et al., "Parton distributions from high-precision collider data", *Eur. Phys. J. C* **77** (Oct, 2017) 663, `doi:10.1140/epjc/s10052-017-5199-5`.

[168] CMS Collaboration, "HiggsAnalysis-CombinedLimit", 2023. `https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/`, accessed on 2023-06-20.

Danksagung

A special thanks goes to the numerous office colleagues during the time in office 9-10 in the physics highrise.

Zu guter Letzt gilt mein Dank auch meinen Eltern, Bruder und Freunden, die stets an mich glaubten, auch wenn manchmal die Arbeit unseren Plänen in die Quere kam.