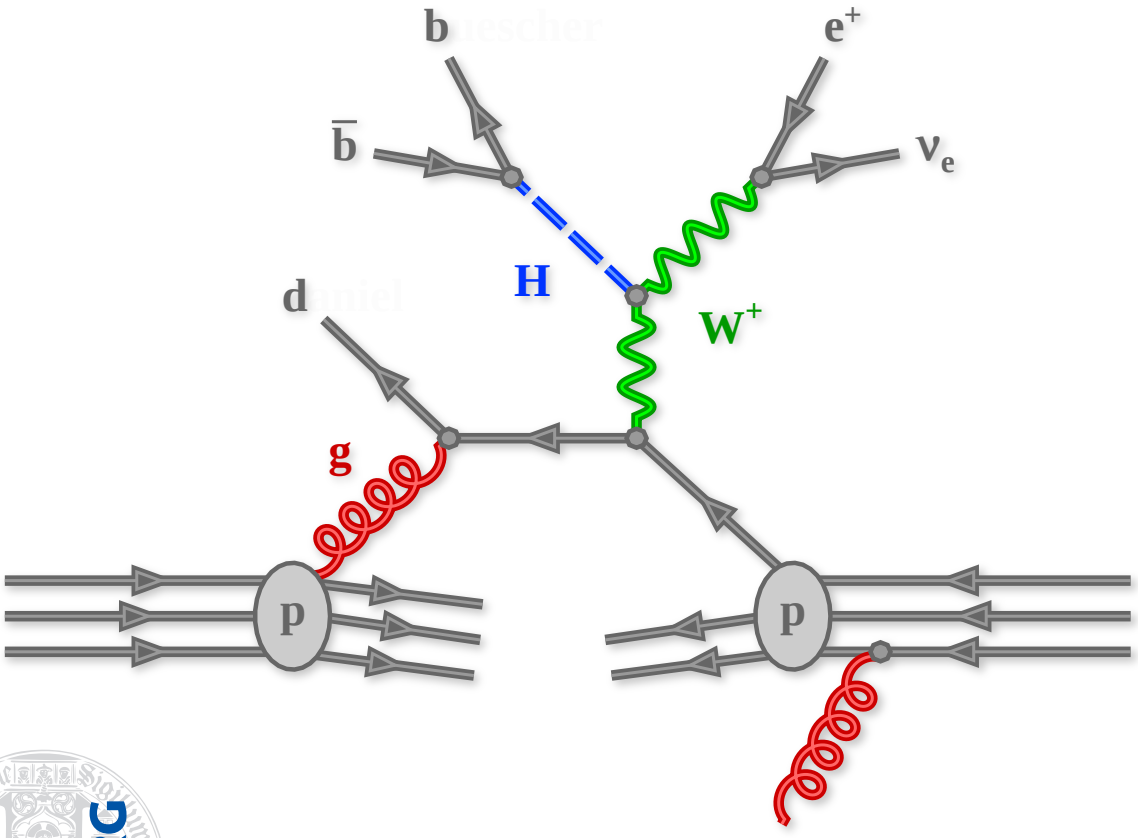
Dissertation

Search for Higgs bosons with b-jets in the final state in proton-proton collisions with the ATLAS experiment



Daniel Büscher

Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität Freiburg

Search for Higgs bosons with b-jets in the final state in proton-proton collisions with the ATLAS experiment

Dissertation

vorgelegt von

Daniel Büscher

August 2016

Prof. Dr. Karl Jakobs
Fakultät für Mathematik und Physik
Albert-Ludwigs-Universität
Freiburg im Breisgau

Dekan:	Prof. Dr. Dietmar Kröner
Betreuer der Arbeit:	Prof. Dr. Karl Jakobs
Referent:	Prof. Dr. Karl Jakobs
Koreferent:	Prof. Dr. Markus Schumacher
Prüfer:	Prof. Dr. Gregor Herten
	J. Prof. Dr. Harald Ita
	Prof. Dr. Karl Jakobs

Datum der mündlichen Prüfung: 26. Oktober 2016

The Standard Model of particle physics seems to be complete with the discovery of the Higgs boson in 2012 by the ATLAS and CMS experiments at the LHC. However, many open questions remain. In particular, no firm evidence for the decay of the Higgs boson to a b-quark pair has been found so far, although this decay is expected to be the most frequent one.

Within this dissertation I present the most sensitive search for this decay to date, which has been carried out with the ATLAS experiment. It uses sophisticated analysis techniques, relying on multivariate methods and on the most advanced calibrations available. The search makes use of many features of the ATLAS detector, from the innermost tracking devices for efficient b-tagging, over the calorimeters for precise measurements of the electron and jet energies, to the outermost muon system for an optimal signal acceptance.

Additionally, a novel search strategy is investigated, aiming for the most abundant production mode of the Higgs boson, the gluon fusion, in combination with the decay to b quarks. This process is usually deemed to be not measurable at the LHC, but the results presented here suggest that it might not be that hopeless.

Further studies are carried out, searching for possible Beyond Standard Model signatures, which can be mediated through the Higgs boson and can have b-jets in the final state as well. A specific jet reconstruction method is applied for reaching an optimal acceptance of hypothetical new particles with very large masses.

Altogether, a wide variety of analysis techniques and searches with the ATLAS experiment is presented.

Daniel Büscher

Contents

1. Introduction **1**

2. Theoretical background **5**
2.1. The Standard Model . . . 5
2.1.1. Particle content . . . 5
2.1.2. Fundamental interactions . . . 7
2.1.3. Spontaneous symmetry breaking . . . 10
2.2. Phenomenology of proton-proton collisions . . . 12
2.3. Phenomenology of the Higgs boson . . . 17
2.4. Current knowledge of the Higgs boson . . . 19
2.5. Higgs boson portal models . . . 23
2.6. Heavy Vector Triplet models . . . 24

3. The ATLAS experiment **27**
3.1. The Large Hadron Collider . . . 27
3.2. The ATLAS detector . . . 28
3.2.1. Inner detector . . . 30
3.2.2. Calorimeters . . . 31
3.2.3. Muon spectrometer . . . 32
3.2.4. Trigger system . . . 33

4. Reconstruction of physical objects **35**
4.1. Tracks and vertices . . . 35
4.2. Leptons . . . 37
4.3. Jets . . . 39
4.4. Flavor tagging . . . 41
4.5. Overlap removal . . . 44
4.6. Missing transverse energy . . . 45
4.7. List of experimental uncertainties . . . 46

5. Search for the VH(bb) process **49**
5.1. Signal and background processes . . . 49
5.2. Event selection . . . 51
5.2.1. Selection requirements . . . 53
5.2.2. Signal acceptance . . . 56
5.2.3. Multijet estimate . . . 57
5.2.4. Dijet mass correction . . . 58

5.3. Multivariate analysis . . . 59
5.3.1. Boosted decision trees . . . 59
5.3.2. Input variables . . . 61
5.3.3. Configuration . . . 66
5.3.4. Signal sample size . . . 68
5.4. Statistical treatment . . . 69
5.4.1. Profile likelihood fit . . . 69
5.4.2. Fit input distributions . . . 71
5.4.3. Smoothing and pruning . . . 71
5.4.4. Inter- and extrapolation . . . 73
5.5. Systematic uncertainties . . . 74
5.5.1. Experimental uncertainties . . . 75
5.5.2. Signal . . . 75
5.5.3. W bosons + jets . . . 75
5.5.4. Z bosons + jets . . . 78
5.5.5. Top-quark pairs . . . 79
5.5.6. Single top quarks . . . 80
5.5.7. Vector-boson pairs . . . 81
5.5.8. Multijet . . . 81
5.6. Binning strategies . . . 82
5.6.1. Numerical optimization . . . 83
5.6.2. Analytical algorithms . . . 83
5.6.3. Performance comparison . . . 86
5.6.4. Optimization of transformation D . . . 88
5.6.5. Optimization of transformation F . . . 92
5.6.6. Further observations . . . 93
5.6.7. Final binning . . . 93
5.7. Fit model validation . . . 94
5.7.1. Post-fit plots . . . 94
5.7.2. Nuisance parameter pulls . . . 95
5.7.3. Correlations . . . 104
5.7.4. Ranking . . . 106
5.7.5. Uncertainty breakdown . . . 107
5.7.6. Likelihood scans . . . 109
5.7.7. Toy experiments . . . 114
5.7.8. Compatibility of regions . . . 114
5.8. Results . . . 115
5.8.1. Most sensitive distributions . . . 115
5.8.2. Signal strength . . . 118
5.8.3. Significance and limits . . . 118
5.8.4. Summary distributions . . . 121

6. Sensitivity to H(bb) in gluon fusion **127**
6.1. Simulated samples . . . 127
6.2. Event selection . . . 128

6.3. Sensitivity estimate . . . 129
6.4. Conclusion . . . 137

7. Search for the VH(invisible) process 139
7.1. Signal and background processes . . . 139
7.2. Event selection . . . 140
7.3. Statistical treatment . . . 145
7.4. Systematic uncertainties . . . 145
7.4.1. Signal . . . 146
7.4.2. W boson + jets . . . 147
7.5. Binning strategies . . . 147
7.6. Fit model validation . . . 149
7.6.1. Post-fit plots . . . 149
7.6.2. Nuisance parameter pulls . . . 153
7.6.3. Correlations . . . 157
7.6.4. Ranking . . . 159
7.6.5. Uncertainty breakdown . . . 159
7.6.6. Compatibility of regions . . . 161
7.7. Results . . . 162

8. Search for VH resonances with H→bb 165
8.1. The CxAOD framework . . . 165
8.2. Signal and background processes . . . 167
8.3. Reconstruction of physical objects . . . 168
8.3.1. Leptons . . . 168
8.3.2. Small jets . . . 169
8.3.3. Large jets . . . 169
8.3.4. Track jets and b-tagging . . . 170
8.3.5. List of experimental uncertainties . . . 171
8.4. Event selection . . . 171
8.4.1. Selection requirements . . . 171
8.4.2. Signal acceptance . . . 175
8.4.3. Neutrino reconstruction . . . 175
8.5. Statistical treatment . . . 180
8.6. Systematic uncertainties . . . 181
8.6.1. Experimental . . . 181
8.6.2. Signal and backgrounds . . . 181
8.7. Binning strategies . . . 182
8.7.1. Regions and transformation D . . . 182
8.7.2. Final binning . . . 183
8.8. Fit model validation . . . 184
8.8.1. Post-fit plots . . . 185
8.8.2. Nuisance parameter pulls . . . 185
8.8.3. Correlations . . . 191
8.9. Results . . . 193

9. Summary **199**

A. VH(bb) appendix **203**
A.1. Pull comparisons . . . 203
A.2. Unfolded b-tagging scale factors . . . 203
A.3. Correlation matrices . . . 211
A.4. Post-fit yields . . . 211

B. Gluon fusion H(bb) appendix **217**
B.1. Signal sample validation . . . 217

C. VH(invisible) appendix **219**
C.1. Equal bin-width plots . . . 219
C.2. Correlation matrices . . . 219

D. VH resonances appendix **227**
D.1. Correlation matrices . . . 227

Bibliography **231**

Acknowledgements **242**

1. Introduction

The Standard Model of particle physics describes all known elementary particles and their interactions with great precision. It states the existence of fermions, spin-1/2 particles, which form the known matter. The interactions among the fermions are mediated by bosons, spin-1 particles. Several of these particles were predicted by the Standard Model[1] and all of them were experimentally observed until the turn of the century, except for one: the Higgs boson.

The Standard Model (SM) symmetry, depicted in Quantum Field Theory (QFT), does not allow for mass terms for both bosons and fermions. Instead, the masses result in the Standard Model from a scalar field, causing a spontaneous symmetry breaking. This *Brout-Englert-Higgs mechanism*[2] predicts observable properties of the Standard Model, e.g. the relation between the masses of the W and the Z bosons, which was experimentally verified[3]. In addition, the Higgs mechanism postulates the existence of a new boson, the *Higgs boson* with spin 0. The direct search for the Higgs boson, however, turned out to be one of the largest scientific ventures of the last decades.

In 2009 a new particle accelerator started its operation: the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN[4]) in Geneva. The LHC is a circular proton-proton collider with a design center-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$. During its *Run 1*, started in 2010, the LHC delivered a large amount of integrated luminosity, corresponding to $\mathsf{L} = 5.5\,\text{fb}^{-1}$ in 2011 at $\sqrt{s} = 7\,\text{TeV}$ and to $\mathsf{L} = 22.8\,\text{fb}^{-1}$ in 2012 at $\sqrt{s} = 8\,\text{TeV}$. After a technical shutdown its operation started again for *Run 2* in 2015 with an increased center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. The protons are collided at four points of the ring where detectors are located. Two of them, the ATLAS and CMS experiments, are multi-purpose detectors designed to address searches for new physics as well as precise tests of the Standard Model.

In 2012 the ATLAS and CMS collaborations announced the discovery of a boson with a mass of about 125 GeV, consistent with the Higgs boson, H, predicted in the Standard Model. This observation was driven by three specific decay modes: $H \to \gamma\gamma$, $H \to ZZ^* \to llll$ and $H \to WW^* \to l\nu l\nu$. These channels provide clean signatures in the detector, due to photons or leptons in the final state, and can be well separated from the background processes. Other important decay modes, such as $H \to \tau\tau$ and $H \to b\bar{b}$, are more difficult to detect. In particular, the $H \to b\bar{b}$ decay, which produces only jets in the detector, is

[1] The Standard Model allowed to infer the existence of some particles before their discovery from measurements of known particles and from theoretical considerations. However, the SM gives no explanation of its particle content as a whole.

[2] The mechanism of spontaneous symmetry breaking in field theories was independently published by Anderson, Brout and Englert, Higgs, as well as by Hagen, Kibble and Guralnik.

[3] The verification with electroweak precision measurements, performed by the experiments at the Large Electron-Positron collider (LEP) in the 1990s, provides indirect evidence for the Higgs mechanism.

[4] Conseil Européen pour la Recherche Nucléaire.

hard to separate from the so-called multijet background: jets are copiously produced by QCD processes in a hadron collider. Nonetheless, these two channels are very important since they directly probe the decay of the Higgs boson to fermions, while the three former channels only provide evidence for the decay to other bosons[5].

By now the observation of the Higgs boson is well established and even confirmed in the $H \to \tau\tau$ decay mode. However, no firm evidence for the $H \to b\bar{b}$ decay has been found so far. This decay has the largest partial decay-width (the largest branching ratio) among all possible decays for the discovered Higgs boson in the Standard Model. Therefore, its measurement is particularly important for determining the total decay width and for indirect constraints of the coupling to possible Beyond Standard Model (BSM) particles.

The search for the $H \to b\bar{b}$ decay is possible using a specific production mode of the Higgs boson, the associated production with a vector boson ($V = W$ or Z boson). The decay products of the vector boson, in particular if these are leptons, allow for a good rejection of background processes and for an efficient triggering of the events. Therefore, this process can provide a better sensitivity than the most abundant production mode, the gluon fusion, although it occurs less frequently.

A major part of this thesis discusses the search for the SM Higgs boson in associated production with a leptonically decaying vector boson and the decay to b quarks, $VH(\to b\bar{b})$ [1]. The analysis is carried out with the data recorded in 2012 with the ATLAS detector at $\sqrt{s} = 8\,\text{TeV}$, corresponding to an integrated luminosity of $\mathsf{L} = 20.3\,\text{fb}^{-1}$. Even though the leptons from the vector-boson decay are used to suppress the background processes, they are still very large. For an optimal sensitivity a *multivariate analysis* (MVA) is performed. The MVA exploits various event properties and their correlations to separate the signal from the background processes. Several types of MVAs exist with different capabilities. *Boosted Decision Trees* (BDTs) are used in this analysis due to their easy configuration and robust separation power.

The output distributions of various BDTs are used to measure the Higgs boson production cross section by fitting the signal and the background processes to the data. The fit, employing a profile likelihood approach, is quite complex due to the use of many distributions and a large number of parameters for the systematic uncertainties. The validation of this fit is discussed in some detail.

As a second study, a sensitivity estimate for the direct detection of the SM Higgs boson via the gluon-fusion production and the decay to b quarks, $gg \to H \to b\bar{b}$, is performed. This processes is often deemed not to be measurable at the LHC due to the large multijet background. However, using dedicated topologies and multivariate analysis techniques a reasonable sensitivity might be achieved, in particular in regard of the Run 2 of the LHC.

Another analysis presented in this thesis is the direct search for decays of the Higgs boson to Beyond Standard Model particles, χ, that are invisible to the detector [2]. This search is performed in the associated production with a vector boson as well, but this time using its hadronic decays: $VH \to q\bar{q}'\chi\chi$. The search is done in the context of a *Higgs boson portal model*, where a hidden sector is postulated, interacting with the SM only through the Higgs boson. Such models can potentially provide candidates for dark matter

[5] The $H \to \gamma\gamma$ decay is mediated in the SM, besides the predominant virtual W-boson loop, by a virtual top-quark loop, which are fermions as well.

particles. The analysis is also performed with the data recorded in 2012 with $\sqrt{s} = 8\,\text{TeV}$ and $\mathsf{L} = 20.3\,\text{fb}^{-1}$.

A second BSM analysis is discussed within this thesis: the search for new and very heavy vector bosons ($V' = W'$ or Z' boson), decaying via $V' \to VH(\to b\bar{b})$ [3]. These decays are postulated in the *Heavy Vector Triplet* model, which is a simplified model used as benchmark for this search. The decay products of the hypothesized V' resonance have large momenta in the detector and cannot be reconstructed efficiently using traditional methods. A specific *boosted* jet reconstruction is applied to recover a good signal acceptance. The analysis is carried out with the first data recorded during Run 2 of the LHC in 2015 with an increased center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$, and an integrated luminosity of $\mathsf{L} = 3.2\,\text{fb}^{-1}$.

Between Run 1 and Run 2 not only the LHC has been upgraded, but also the ATLAS detector has been improved. An innermost silicon pixel layer has been added for the tracking detector, which benefits in particular analyses with b jets in the final state. Another improvement was made in the software system: a new data format for physics analysis, aiming for more efficient storage of systematic uncertainties, has been introduced. To make use of this, a new analysis framework is developed for Run 2.

In summary, four studies are discussed in this thesis: three searches in the VH final state and one sensitivity estimate for the $gg \to H \to b\bar{b}$ process. Although the physics motivations for the searches are quite different, they all have similar final states. Hence, they share reconstruction techniques and systematic uncertainties, which is a good motivation from the experimental point of view to discuss them together.

This thesis is organized as follows. In Chapter 2 a brief introduction to the Standard Model is given, the phenomenologies of proton-proton collisions and of Higgs boson decays are discussed and recent results in the searches for the Higgs boson are presented. Further, possible extensions of the SM are sketched. The ATLAS experiment is described in Chapter 3 and the reconstruction of physical objects in Chapter 4. The SM $VH(\to b\bar{b})$ analysis is presented in Chapter 5 and the estimate of the sensitivity to the SM $gg \to H \to b\bar{b}$ process in Chapter 6. The search for the BSM $VH(\to \text{inv.})$ process is discussed in Chapter 7 and for the BSM $V' \to VH(\to b\bar{b})$ process in Chapter 8. A summary is given in Chapter 9.

2. Theoretical background

An overview of the most important aspects of the theoretical background for this thesis is given in the following. A brief introduction to the Standard Model is given in Section 2.1, the phenomenologies of proton-proton collisions are discussed in Section 2.2 and of Higgs boson decays in Section 2.3. Recent results in the searches for the Higgs boson are presented in Section 2.4. Possible extensions of the SM in terms of Higgs boson portal models are sketched in Section 2.5 and in terms of the Heavy Vector Triplet models in Section 2.6.

2.1. The Standard Model

The *Standard Model* (SM) of particle physics represents a description of the elementary particles and their interactions. It was developed in the 1960s and 1970s and proved very successful in explaining the results of existing experiments and predicting the outcome of later ones. The most important predictions that were experimentally confirmed are the existence of the $W^{\pm}$ and Z bosons, found at CERN [4], the top-quark, found at the Tevatron [5], and most recently the Higgs boson, found at CERN [6, 7].

The Standard Model postulates the existence of elementary particles that form the known matter, the *fermions*, which are classified into *leptons* and *quarks*. Their interactions are described by gauge fields mediated by the exchange of *bosons*. With these ingredients the model is able to depict three of the four known forces: the electromagnetic, the weak and the strong force. Only the gravitational force is not included.

The theoretical description is given in the framework of quantum field theory, where the key assumption is an invariance under local phase transformations, also called "gauge invariance". The underlying symmetry is $SU(3)_C \times SU(2)_L \times U(1)_Y$ and will be motivated in the following. The description follows to a large extend the textbooks of Refs. [8–10].

2.1.1. Particle content

Fermions Experiments show that the elementary spin-1/2 particles of the SM, the fermions, exist in three generations ("flavors"). The generations differ only by the particle masses, increasing from the first to third generation. A summary of all known elementary fermions is given in Table 2.1. Listed is the classification into the three generations, in quarks and leptons and their electromagnetic charge.

All ordinary matter consists only of the first-generation particles: up- and down-quarks, forming protons and neutrons, and the electron. The other fermions can occur in high energy processes, e.g. in cosmic rays or in collider experiments, such as the LHC.

Bosons Interactions of the fermions are mediated by bosons, which are integer-spin particles. The vector bosons (spin 1) are shown in Table 2.2. The best-known boson is the

Fermions	1st generation		2nd generation		3rd generation		Charge [e]
Quarks	up	u	charm	c	top	t	$+\frac{2}{3}$
	down	d	strange	s	bottom	b	$-\frac{1}{3}$
Leptons	e-neutrino	ν_e	μ-neutrino	ν_μ	τ-neutrino	ν_τ	0
	electron	e	muon	μ	tau	τ	-1

Table 2.1.: Elementary fermions categorized into quarks and leptons and the three generations with their electric charge. For each charged particle an anti-particle exists with opposite charge.

Interaction	Bosons	Charge [e]	Mass [GeV]
Electromagnetic	Photon γ	0	0
Strong	Gluons g	0	0
Weak	$W^\pm$ bosons	± 1	80.4
	Z boson	0	91.2

Table 2.2.: Bosons of the Standard Model with their electric charges and masses. The gluons exist in eight colored states.

photon, γ, mediator of the electromagnetic force. It is known as visible light for a narrow energy range. The photon is massless and stable, therefore the electromagnetic force has an infinite interaction range. It interacts with all particles carrying electric charge.

The gluons, g, are the carriers of the strong force. Gluons interact with colored particles, the quarks and the gluons themselves, as described in the next section. They exist in eight different colored states. Also gluons are massless. However, the strong interaction is short-ranged, since the SM forbids the existence of unbound colored particles.

The $W^\pm$ and Z bosons are the mediators of the so-called weak force. They are massive particles, therefore the weak force has only a small range. They interact with all particles carrying weak charges, as described in the next section. The $W^\pm$ bosons are the only bosons known to be able to change the flavor of the fermions, but only for the quarks.

The masses of the bosons (and of the fermions via the Yukawa coupling, as described later) are generated by the Brout-Englert-Higgs (BEH) mechanism or Higgs mechanism for short. It predicts the only elementary spin-0 particle, the Higgs boson. It is the only particle of the Standard Model that has eluded experimental proof up to recent measurements at the LHC. The current knowledge of the Higgs boson is further discussed in Section 2.4.

All quarks and charged leptons are massive, but neutrinos are assumed to be massless in the following. This agrees with direct mass measurements to a very good level. However, it should be mentioned, that oscillations in the neutrino sector are observed [11]. This provides evidence for neutrinos having small, but non-zero masses.

2.1.2. Fundamental interactions

The Standard Model consists of several components, each describing specific interactions of the elementary particles. Historically, the first part is the Quantum Electro Dynamics (QED). The strong force is described by the Quantum Chromo Dynamics (QCD) and the weak force is unified with QED in the Electroweak (EW) model [12–14]. The latter is also called "GSW model", named after Glashow, Salam and Weinberg.

Quantum Electro Dynamics The QED is mathematically described by a Lagrangian density, $\mathcal{L}$, denoted by simply *Lagrangian* in the following. It is derived using symmetry arguments, especially the invariance under local phase transformations. A free fermion (spin 1/2) is described by the Lagrangian

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

where ψ is the Dirac spinor of the fermion field, $\overline{\psi} = \psi^\dagger\gamma^0$ is the adjoint, γ^μ are the Dirac gamma matrices and $\partial_\mu = \frac{\partial}{\partial x_\mu}$ are the partial derivatives. The Dirac equation,

$$(i\gamma^\mu\partial_\mu - m)\psi = 0, \tag{2.2}$$

is the corresponding equation of motion.

The Lagrangian is requested to be invariant under local phase transformations of the form

$$\psi(x) \to \psi'(x) = e^{i\alpha(x)}\psi(x), \tag{2.3}$$

where the local phase is given by $\alpha(x)$, depending on space and time. These transformations form the abelian unitary group $U(1)$, since $e^{i\alpha(x)}$ can be written as a 1×1-matrix U with $U^\dagger U = 1$. One finds that the first term of the Lagrangian (Equation 2.1) is not invariant under this transformation, since

$$\partial_\mu\psi \to \partial_\mu\psi' = e^{i\alpha(x)}\partial_\mu\psi + ie^{i\alpha(x)}\psi\partial_\mu\alpha(x). \tag{2.4}$$

The invariance is established by introducing an additional field A_μ, transforming as

$$A_\mu \to A'_\mu = A_\mu + \frac{1}{e}\partial_\mu\alpha(x), \tag{2.5}$$

and replacing ∂_μ with the covariant derivative D_μ:

$$D_\mu = \partial_\mu - ieA_\mu. \tag{2.6}$$

One can show that

$$D_\mu\psi \to D'_\mu\psi' = e^{i\alpha(x)}D_\mu\psi \tag{2.7}$$

holds and the Lagrangian becomes

$$\begin{aligned}\mathcal{L} &= i\overline{\psi}\gamma^\mu D_\mu\psi - m\overline{\psi}\psi \\ &= \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi + e\overline{\psi}\gamma^\mu\psi A_\mu.\end{aligned} \tag{2.8}$$

Thus, by introducing the *gauge field* A_μ, the local phase invariance is restored. This field couples to the fermion field with the coupling constant e, the elementary electric charge. By introducing a kinetic term, using the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, the gauge field A_μ is identified as the physical photon field and the Lagrangian of QED is complete:

$$\mathcal{L}_{\text{QED}} = i\overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi + e\overline{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{2.9}$$

The individual terms are the kinetic energy and mass of the fermion, the coupling between fermion and photon fields and the kinetic energy of the photon. The local phase invariance forbids the introduction of a mass term of the form $\frac{1}{2}m^2 A_\mu A^\mu$ for the photon field. Hence this formalism requests the photon to be massless, which is in agreement with all experiments.

Quantum Chromo Dynamics The QCD describes the interaction of quarks and gluons. Quarks are fermions as the leptons, but an additional degree of freedom is introduced: the *color charge*, which exists in the states red (r), green (g) and blue (b). The simple Dirac spinors are replaced by vectors of three spinors denoting the quarks:

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix} \tag{2.10}$$

Together with the eight gauge fields G^a_μ, $a = 1, ..., 8$, representing the gluons, the Lagrangian of QCD is derived:

$$\mathcal{L}_{\text{QCD}} = i\overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi - g_s\left(\overline{\psi}\gamma^\mu\frac{\lambda_a}{2}\psi\right)G^a_\mu - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a. \tag{2.11}$$

It is invariant under transformations of the non-abelian unitary group $SU(3)_C$, whose generators are $T_a = \lambda_a/2$. Here, g_s denotes the strong coupling constant, λ_a are the eight Gell-Mann-matrices and $G^a_{\mu\nu}$ is the field strength tensor, written as

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu. \tag{2.12}$$

The structure constants f_{abc} fulfill $\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = i f_{abc}\frac{\lambda_c}{2}$. The last term of the field strength tensor is due to the $SU(3)_C$ being non-abelian and does not have an equivalent in QED. It enables the gluons to interact with themselves. As for QED the local phase invariance requests the force carriers, here gluons, to be massless.

Electroweak model All fermions are subject to the weak force, manifested e.g. in nuclear β-decays. The interaction is carried out by the exchange of the W^+ and W^- bosons, which is called *charged current* interaction, and of the Z boson, called *neutral current* interaction. While the exchange of the $W^\pm$ bosons allows to change the flavor of quarks, there are no flavor-changing neutral currents (FCNC) observed at tree level.

Experiments show that the charged weak current only couples to left-handed fermions. In the electroweak model, conserving a $SU(2)_L \times U(1)_Y$ symmetry, the left handed fermions

	Generation			Quantum numbers			
	1st	2nd	3rd	I	I_3	Y	Q [e]
Quarks	$\begin{pmatrix} u \\ d' \end{pmatrix}_L$	$\begin{pmatrix} c \\ s' \end{pmatrix}_L$	$\begin{pmatrix} t \\ b' \end{pmatrix}_L$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{2}{3}$
				$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{3}$
	u_R	c_R	t_R	0	0	$\frac{4}{3}$	$\frac{2}{3}$
	d_R	s_R	b_R	0	0	$-\frac{2}{3}$	$-\frac{1}{3}$
Leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	$\frac{1}{2}$	$\frac{1}{2}$	-1	0
				$\frac{1}{2}$	$-\frac{1}{2}$	-1	-1
	e_R^-	μ_R^-	τ_R^-	0	0	-2	-1

Table 2.3.: The fermions in the electroweak model with their quantum numbers.

are assigned to $SU(2)_L$ doublets with corresponding *isospin* of $I = 1/2$ and gauge fields W_μ^a. The right-handed fermions are described by $U(1)_Y$ singlets with $I = 0$ and gauge field B_μ. Here, the hypercharge Y is introduced, which is related to the electric charge via $Q = I_3 + \frac{Y}{2}$, as derived below.

An overview of the fermions and their quantum numbers in the electroweak model is given in Table 2.3. The weak eigenstates of the down-type quarks, denoted by d', s' and b', are a mixture of the mass eigenstates d, s and b. The mixing is described by the CKM matrix [15].

The left-handed isospin doublets χ_L and right-handed singlets ψ_R behave under local phase transformations as

$$\chi_L(x) \to \chi_L'(x) = e^{i\alpha_a(x)\tau_a} e^{i\beta(x)Y} \chi_L, \tag{2.13}$$

$$\psi_R(x) \to \psi_R'(x) = e^{i\beta(x)Y} \psi_R \tag{2.14}$$

and form the $SU(2)_L \times U(1)_Y$ symmetry. Here, $\alpha(x)$ and $\beta(x)$ are the local phases, τ_a with $a = 1, 2, 3$ are the generators of $SU(2)_L$, and Y is the weak hypercharge operator, generating $U(1)_Y$. The covariant derivative is given by

$$D_\mu = \partial_\mu + igW_\mu^a \frac{\tau_a}{2} + ig' B_\mu \frac{Y}{2}, \tag{2.15}$$

where g is the coupling constant of the $SU(2)_L$ gauge fields, W_μ^a, and g' is the coupling constant of the $U(1)_Y$ gauge field, B_μ.

The resulting electroweak Lagrangian is

$$\mathcal{L}_{\mathrm{EW}} = i\overline{\chi_L^i}\gamma^\mu D_\mu \chi_L^i + i\overline{\psi_R^i}\gamma^\mu D_\mu \psi_R^i - \frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{2.16}$$

where a summation over the three isospin doublets and six singlets is done (index i). The field tensors are given by

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g\epsilon_{abc} W_\mu^b W_\nu^c \tag{2.17}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{2.18}$$

where the completely antisymmetric tensor, ϵ_{abc}, denotes the structure constants of $SU(2)_L$, while the structure constant of $U(1)_Y$ is zero. This enables the W_μ^a fields to interact with themselves, while the B_μ field has no self-interaction.

The physical fields are given by linear combinations of W_μ^a and B_μ,

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \tag{2.19}$$

$$Z_\mu = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \tag{2.20}$$

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu, \tag{2.21}$$

where the weak mixing angle, θ_W, is introduced. It relates the coupling constants via $\cos\theta_W = g/\sqrt{g^2+g'^2}$ and $\sin\theta_W = g'/\sqrt{g^2+g'^2}$. By rewriting the Lagrangian in terms of the physical fields and comparing the A_μ components to the photon field of QED (Equation 2.9) one obtains the relations

$$e = g\sin\theta_W = g'\cos\theta_W \tag{2.22}$$

$$\text{and} \qquad Q = I_3 + \frac{Y}{2}. \tag{2.23}$$

As for QED and QCD, the local phase invariance forbids the introduction of mass terms for the bosons. In addition, mass terms of the form $-m\overline{\psi}\psi$ for the fermions are forbidden in the EW sector. This is in conflict with experiments, where the Z and $W^\pm$ bosons and the fermions are found to be massive, as shown for the bosons in Table 2.2. This problem is solved by the mechanism described in the next section.

2.1.3. Spontaneous symmetry breaking

The local phase invariance in the electroweak model requests the fermions and bosons to be massless particles. However, experiments show that the vector bosons $W^\pm$ and Z^0 are very massive, as shown in Table 2.2. This contradiction can be solved by a spontaneous symmetry breaking, induced by the Brout-Englert-Higgs (BEH) mechanism [16–18] or "Higgs mechanism" for short.

The scalar Higgs field, Φ, a weak isospin doublet, containing complex scalar fields with hypercharge $Y = 1$, is postulated as

$$\Phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2\\ \phi_3 + i\phi_4\end{pmatrix}. \tag{2.24}$$

The corresponding Lagrangian,

$$\mathcal{L}_{\text{Higgs}} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - \mu^2\Phi^\dagger\Phi - \lambda(\Phi^\dagger\Phi)^2, \tag{2.25}$$

is invariant under $SU(2)_L \times U(1)_Y$ phase transformations. The potential, parameterized as

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2, \tag{2.26}$$

has its minimum at $\Phi \neq 0$ for $\mu^2 < 0$ and $\lambda > 0$. This is illustrated for a single complex scalar field, ϕ, in Figure 2.1.

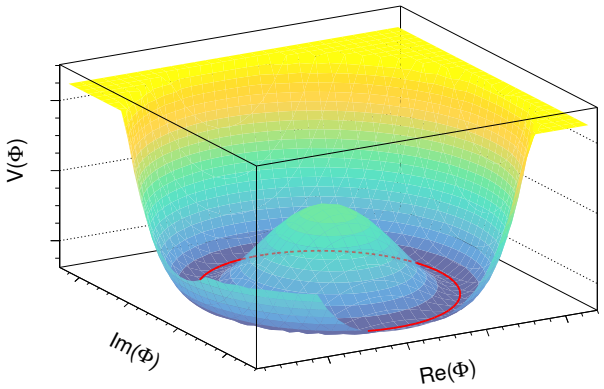


Figure 2.1.: Illustration of the Higgs potential in the complex plane for a single complex scalar field. The minimum of the potential is at $\phi \neq 0$ (red circle), causing a spontaneous symmetry breaking.

One point of the minimum,

$$\Phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{2.27}$$

is chosen as ground state, where $v = \sqrt{-\mu^2/\lambda}$ depicts the vacuum expectation value (*v.e.v.*). Since $v > 0$, the symmetry of the group is spontaneously broken. The field can be parameterized as

$$\Phi(x) = \frac{e^{i\tau_a \theta_a(x)/v}}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{2.28}$$

where $\theta_a(x)$ $(a = 1, 2, 3)$ and $h(x)$ are real fields. The particle corresponding to the $h(x)$ field is called Higgs boson and has spin 0. The exponential containing the θ_a fields (*Goldstone bosons*) is eliminated in the Lagrangian due to the local phase invariance and does not have a physical meaning.

Substituting this parameterization in the Lagrangian of Equation 2.25 the term

$$\begin{aligned} \left| \left(ig\frac{\tau_a}{2} W_\mu^a + ig'\frac{Y}{2} B_\mu \right) \Phi_0 \right| \; = \; & \left(\frac{1}{2}vg\right)^2 W_\mu^+ W^{\mu-} \\ & + \left(\frac{1}{2}vg\right)^2 \frac{1}{2\cos^2\theta_W} Z_\mu Z^\mu \\ & + 0 \cdot A_\mu A^\mu \end{aligned} \tag{2.29}$$

is found, where the notation $|...|^2 = (...)^\dagger(...)$ was used and the physical bosons occur in the final expression. The obtained terms can be identified as mass terms of the vector bosons and the relations

$$m_W = \frac{1}{2}vg, \qquad m_Z = \frac{m_W}{\cos\theta_W} \qquad \text{and} \qquad m_\gamma = 0 \tag{2.30}$$

are found. This represents a strong prediction of the SM for the relation between m_W and m_Z, which was experimentally verified by the experiments at the LEP collider [19].

The vacuum expectation value is determined from measurements of the Fermi coupling G_F in muon decays to be $v = (\sqrt{2}G_F)^{-1/2} = 246\,\text{GeV}$. Further, a mass term for the Higgs boson can be found, which leads to $m_H = \sqrt{-2\mu^2}$. This relation does not have predictive power, as the parameter μ can be freely chosen in the SM.

Similarly, the fermion masses are generated using their coupling to the Higgs boson, also called *Yukawa coupling*. It is described by an additional $SU(2)_L \times U(1)_Y$ invariant component to the Lagrangian,

$$\mathcal{L}_{\text{Yukawa}} = -G_l^{ij}\overline{L}_L^i \Phi l_R^j - G_d^{ij}\overline{Q}_L^i \Phi d_R^j - G_u^{ij}\overline{Q}_L^i \Phi_C u_R^j + \text{h.c.}, \tag{2.31}$$

where $\overline{L}_L^i$ ($\overline{Q}_L^i$) are the lepton (quark) isospin doublets, l_R^j (d_R^j, u_R^j) are the lepton (down, up-type quark) singlets and $\Phi_C(x) = \sqrt{1/2}\,(v+h(x),0)^\top$. The coupling constants are given by G_l^{ij}, G_d^{ij} and G_u^{ij}. In case of the quarks the eigenstates of the weak interaction are a mixture of the mass eigenstates, while no mixing is present for the leptons. For example, the Yukawa term for the electron is given by

$$\begin{aligned}\mathcal{L}_{\text{Yukawa}}^e &= -\frac{G_e}{\sqrt{2}} v\,(\overline{e}_L e_R + \overline{e}_R e_L) - \frac{G_e}{\sqrt{2}}\,(\overline{e}_L e_R + \overline{e}_R e_L)\,h \\ &= -m_e \overline{e} e - \frac{m_e}{v}\overline{e} e h\end{aligned} \tag{2.32}$$

In summary, the Lagrangian of the Standard Model is composed as

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \tag{2.33}$$

This Lagrangian is invariant under local phase transformations of the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group. The Higgs boson interacts with all massive weakly-interacting particles, including itself. One can see from Equation 2.32 that the coupling of the Higgs boson to the fermions is proportional to their mass. Similarly, its couplings to the bosons are found to be quadratic in their masses.

2.2. Phenomenology of proton-proton collisions

The predictions of the Standard Model, and models beyond that, can be tested with scattering experiments such as the ATLAS experiment. It is located at the Large Hadron Collider (LHC), which is a proton-proton collider. The advantage of using protons instead of electrons, the most common elementary particle for collider experiments, is the higher mass of the proton. It enables the collider to reach higher center-of-mass energies due to a drastically reduced energy loss from synchrotron radiation, which is proportional to m^{-4}.

However, protons have, in contrast to electrons, a substructure. Since the Standard Model only describes the interaction of elementary particles, the substructure has to be modeled for a complete description of the proton-proton scattering process. This cannot be done with perturbative QCD. Instead, phenomenological methods are used, as described in the following.

The *cross section*, σ, can be related to the probability of a scattering process. The occurrence of such a process is also called *event*. From the experimental point of view the

cross section is defined as

$$\sigma = \frac{R}{L}, \tag{2.34}$$

where $R = dN/dt$ is the rate of events and L is the *instantaneous luminosity*. The cross section can be rewritten as

$$\sigma = \frac{N}{\int L\, dt} = \frac{N}{\mathsf{L}}, \tag{2.35}$$

where $\mathsf{L} = \int L\, dt$ is the *integrated luminosity* and the N total number of events. The unit of σ is usually taken as $1\,\mathrm{barn} = 1\,\mathrm{b} = 10^{-28}\,\mathrm{m}^2$.

The luminosity depends on the parameters of the particle accelerator. For a circular accelerator with a rotation frequency f_r and two colliding beams with n_b bunches of particles each, the luminosity can be written as

$$L = f_r \frac{n_b N_1 N_2}{A}, \tag{2.36}$$

where A is the transverse area of the beam and $N_{1,2}$ are the numbers of particles inside each bunch of beam 1 and 2, respectively. Assuming $N_1 = N_2 = N_p$ and a Gaussian shape of the beam with widths σ_x and σ_y in the transverse plane the luminosity becomes

$$L = f_r \frac{n_b N_p^2}{4\pi\sigma_x\sigma_y}. \tag{2.37}$$

If the total cross section for inelastic processes, σ_{inel}, is known, the luminosity can be determined as

$$L = \frac{R_{\mathrm{inel}}}{\sigma_{\mathrm{inel}}} = \frac{\mu n_b f_r}{\sigma_{\mathrm{inel}}}, \tag{2.38}$$

where μ is the average number of interactions per bunch crossing.

From Equation 2.35 the number of events for a specific process can be predicted as

$$N = \sigma\, \mathsf{L}, \tag{2.39}$$

where a theoretical computation of the cross section is needed. The computation is complicated by the fact that protons are composite particles. They consist of three valence quarks (uud), sea quarks of any flavor and gluons. This structure is described by Parton Distribution Functions (PDFs), which have to be determined experimentally. The PDFs, parameterized as $f_{q_i}(x_i, Q^2)$, give the probability to find a parton q_i with a momentum fraction x_i of the total proton momentum in a collision with a momentum transfer Q^2 in the Breit frame.

The PDFs, which are measured at a specific Q_0^2, can be transfered to a different Q^2 by using the DGLAP equations [21–23]. In Figure 2.2 the proton PDFs for $Q^2 = 10\,\mathrm{GeV}$ and $Q^2 = 10^4\,\mathrm{GeV}$ are shown. For large x the up and down valence quarks of the proton dominate, while for low x the gluons dominate and the sea quarks have similar contributions as the valence quarks. For further increased Q^2 the fraction of sea quarks and gluons rises in comparison to the valence quarks.

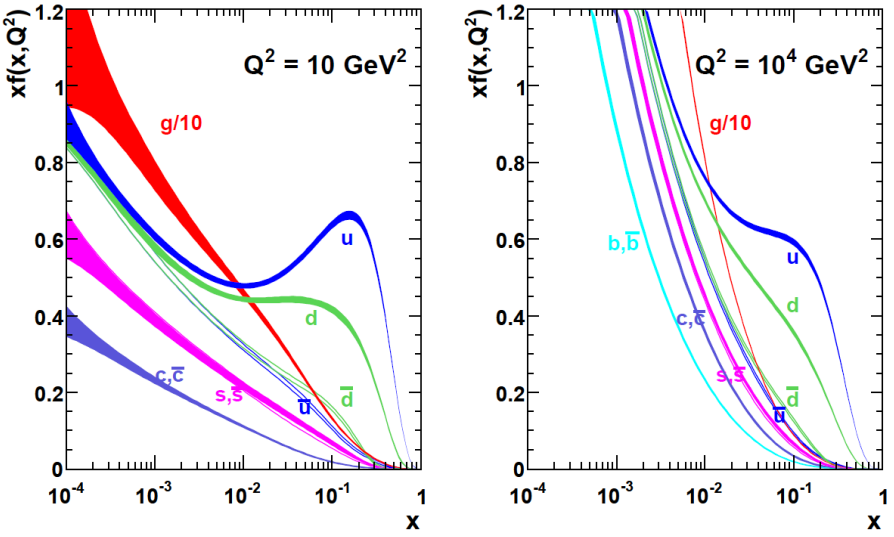


Figure 2.2.: Parton distribution functions for $Q^2 = 10\,\mathrm{GeV}^2$ (left) and $Q^2 = 10^4\,\mathrm{GeV}^2$ (right) calculated at NNLO [20]. Shown are the functions for the valence quarks, sea quarks and gluons inside the proton. The widths of the bands represent the uncertainties.

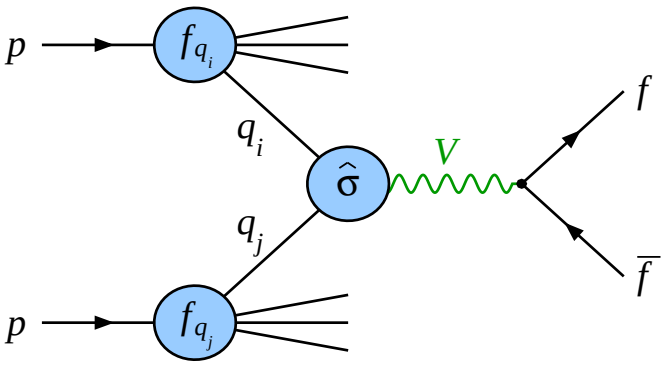


Figure 2.3.: The production (and decay) of a vector boson V (W or Z) in a proton-proton collision, as described by the factorization theorem. The time axis points from left to right.

The total cross section for the process $pp \to X$ is calculated from the partonic cross section, $\hat{\sigma}$, multiplied with the PDFs, integrating over the parton momenta and summing over all parton combinations leading to the final state X [24]:

$$\sigma(pp \to X) = \sum_{i,j} \int dx_i dx_j f_{q_i}\left(x_i, Q^2\right) f_{q_j}\left(x_j, Q^2\right) \hat{\sigma}(q_i q_j \to X). \tag{2.40}$$

where the partons are denoted by q_i and q_j. Here, the assumption is made that the partonic interaction is independent of the PDFs. This *factorization theorem* [25] requires the introduction of the factorization scale, μ_F. It separates between the soft radiation, which is included in the PDFs, and the hard interactions, which are accounted for in $\hat{\sigma}$. The factorization scheme is depicted in Figure 2.3 for the production of a vector boson V (W or Z boson).

The partonic cross section is written as

$$\hat{\sigma} = \int \frac{|\mathcal{M}|^2}{F} dQ, \tag{2.41}$$

where $\mathcal{M}$ is the matrix element for the transition from initial to final state, F is the incident particle flux and dQ is the phase space factor for the given kinematics. The probability amplitude is given by $|\mathcal{M}|^2$.

The matrix element $\mathcal{M}$ can be calculated using the Lagrangian densities and perturbation theory. The hadronic cross section for strong interactions can be written as a power expansion series of the strong coupling constant, α_s:

$$\hat{\sigma} = \hat{\sigma_0}\left(1 + c_1\alpha_s + c_2\alpha_s^2 + ... + c_n\alpha_s^n\right), \tag{2.42}$$

where $\hat{\sigma_0}$ denotes the leading order cross section. Similarly, additional corrections from electroweak processes can be expressed in orders of the electroweak coupling constant. The numerical calculation of such a series can only take a limited number of orders into account, which is denoted by LO (leading order), NLO (next-to leading oder), NNLO (next-to-next-to leading order) and so on.

The calculated cross sections for various processes are shown in Figure 2.4 as a function of the center-of-mass energy, $\sqrt{s}$. One can see that the electroweak productions of W and Z bosons have cross sections that are several orders of magnitude below the total inelastic cross section of $\sigma_{\text{tot}} \approx 100\,\text{mb}$ for $\sqrt{s} = 8\,\text{TeV}$. One of the processes of interest in this thesis, the Higgs boson production, is again several orders below with $\sigma_{\text{Higgs}} \approx 20\,\text{pb}$ for $m_H = 125\,\text{GeV}$ and $\sqrt{s} = 8\,\text{TeV}$. Therefore, about one event out of five billion is expected to contain a Higgs boson, making its detection a great challenge.

The remaining partons of the protons that did not take part in the hard scattering are not in a stable compound anymore. They hadronize and cause additional objects in the detector, which is called *underlying event*. The fragmentation of the partons is described by fragmentation functions. These functions cannot be calculated by perturbation theory, but have to be modeled based on experimental results. Common models are the cluster [27] and the string fragmentation [28].

Another effect, making the measurements at the LHC even harder, is called *pile-up*. Each bunch contains a large number of protons and the total cross section for hard scattering is very large. Therefore, typically several tens of proton-proton interactions occur

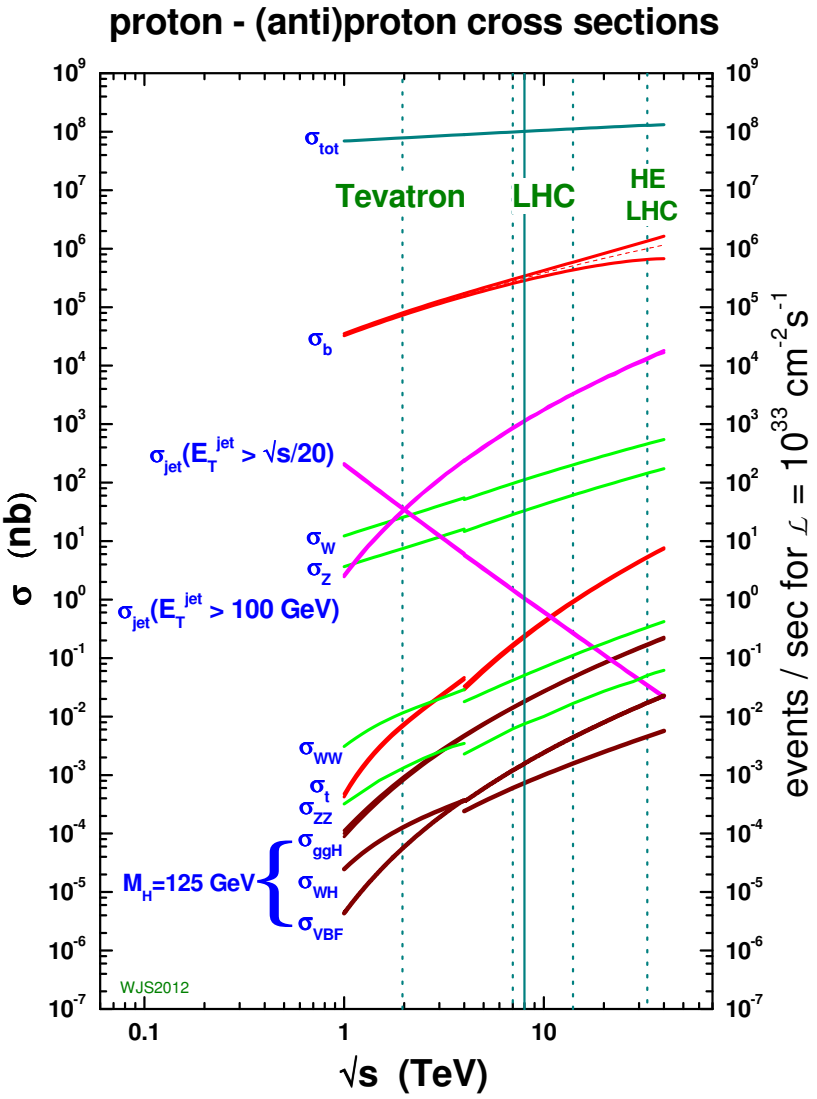


Figure 2.4.: Predicted cross sections (left axis) and expected number of events per second for $L = 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (right axis) for various processes occurring in proton-(anti)proton collisions as a function of $\sqrt{s}$. The vertical dotted lines represent the center-of-mass energy of 1.96 GeV at the Tevatron and 7, 8 (solid) and 14 GeV at the LHC. A dotted line for 33 GeV is shown for a possible high energy (HE) LHC [26].

in a single bunch crossing in the ATLAS experiment [29], which is called *in-time* pile-up. Additionally, *out-of-time* pile-up can occur from the scattering products of a bunch crossing before or after the one under consideration. Specific reconstruction techniques exist to alleviate the resolution loss from pile-up, which are described in Chapter 4.

In conclusion, the phenomenology of proton-proton collisions is far more complex than for elementary particles. All the effects discussed above (in addition to many others like detector response/resolution) have to be modeled by a Monte Carlo simulation, which is used to compare the data to the expectation from theory. This modeling cannot be perfect, such that corrections and uncertainties have to be applied, which are discussed throughout this thesis.

2.3. Phenomenology of the Higgs boson

Production modes The production of the Higgs boson at the LHC can be classified into several modes. The four most important ones, depicted in Figure 2.5, are discussed here. The production modes differ in their cross sections, which are shown in Figure 2.6 (left) for $m_H = 125\,\text{GeV}$ as a function of $\sqrt{s}$, and their phenomenology.

The gluon fusion (ggF) is by far the dominant production mode for this Higgs boson mass. Its cross section is over ten times larger than for the next one, which is due to the large fraction of gluons for small values of x inside the protons [20]. Two gluons fuse through a quark loop to a Higgs boson. Since the coupling of the Higgs boson is proportional to the fermion mass, a top-quark loop is preferred, but the contribution from bottom quarks is sizable as well [30].

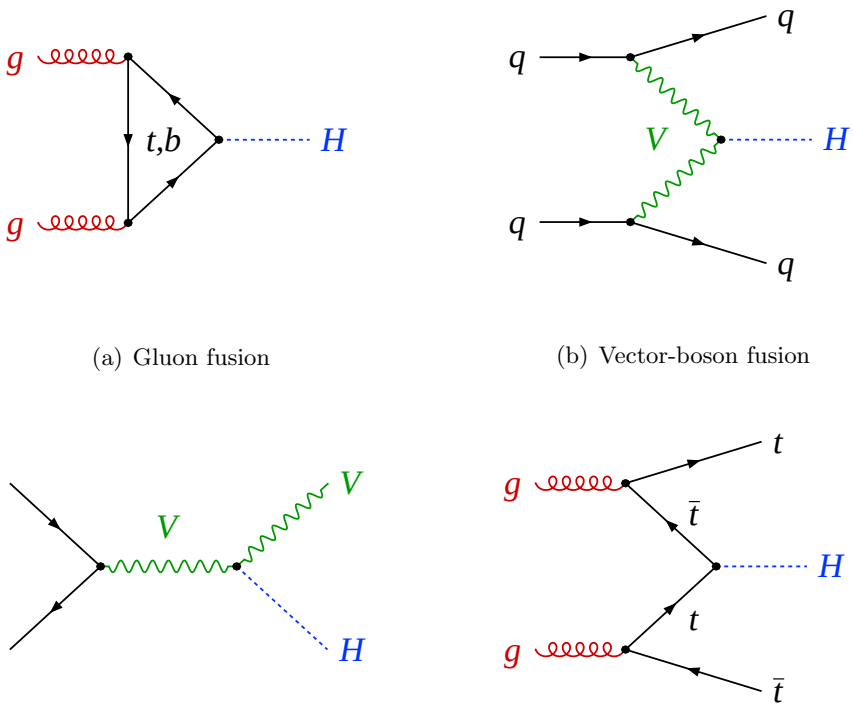
The vector boson fusion (VBF) is the next important production mode. Here, two W or Z bosons are radiated from quarks of the protons and fuse to a Higgs boson. The two quarks produce jets inside the detector in opposite directions along the beam. These can be reconstructed and used to reject backgrounds.

The production in association with a vector boson (VH) has an even smaller cross section. Here, two quarks fuse to a W or Z boson, which radiates the Higgs boson. The vector boson can decay into leptons. They constitute additional objects in the detector, which can be reconstructed with high efficiencies. This allows for a very good background suppression.

The associated production with two top quarks (ttH) has the fourth largest cross section among the processes discussed. The decays of the top quarks produce various objects in the detector, therefore this production mode is the most complex one regarding the reconstruction.

Branching ratios The branching ratios of the Higgs boson decay are dictated by (besides the phase space) the masses of the decay products, since the coupling of the Higgs boson is proportional to the fermion masses and quadratic in the boson masses. The branching ratios are shown as a function of the Higgs boson mass, m_H, in Figure 2.6 (right).

The decay to b quarks, $H \to b\bar{b}$, has the largest branching ratio for $m_H < 135\,\text{GeV}$. In this region b quarks are the heaviest elementary particles that are kinematically accessible. Pairs of heavier decay products, namely W and Z boson and top-quark pairs, are



(a) Gluon fusion

(b) Vector-boson fusion

(c) Vector-boson associated

(d) Top-quark associated

Figure 2.5.: Feynman diagrams of the most relevant production modes of the Higgs boson at the LHC. The process with the largest cross section is (a) the gluon fusion, followed by (b) the vector-boson fusion, (c) the associated production with a vector boson and (d) with a top-quark pair.

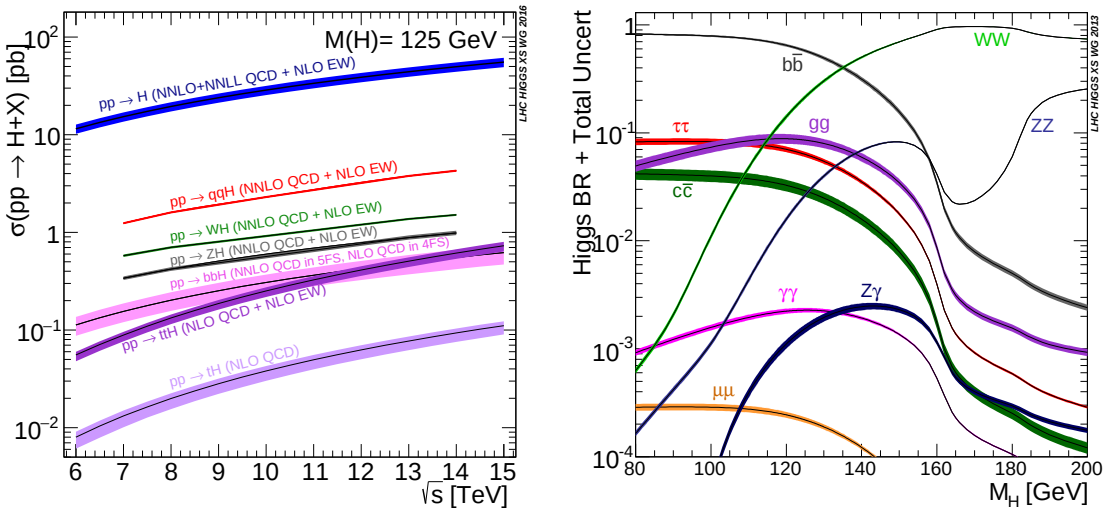


Figure 2.6.: Left: cross sections for various Higgs boson production-modes for $m_H = 125\,\text{GeV}$ as a function of $\sqrt{s}$. $pp \to H$ denotes the gluon fusion, $pp \to qqH$ the vector boson fusion, $pp \to WH$ and $pp \to ZH$ the associated production with a vector boson and $pp \to ttH$ with a top-quark pair. Right: branching ratios of the Higgs boson decay as a function of m_H [31].

suppressed, since they are in sum heavier than the Higgs boson itself. At least one of them has to be virtual.

For $m_H > 135\,\text{GeV}$ the process $H \to WW^*$ has the largest branching ratio. This threshold is significantly smaller than twice the W mass, since the production of virtual vector bosons is allowed due to their intrinsic width.

Even though the decay to b quarks is preferred in the low-mass region, this channel is not the most sensitive one. The decays of the Higgs boson into photons or vector bosons produce cleaner signatures in the detector and therefore can be separated from the backgrounds more efficiently.

The mass resolution for the various decay channels differs greatly. The best resolution is reached with a few GeV in the decay channels $H \to \gamma\gamma$ and $H \to ZZ^* \to llll$, since the energies of the photons and leptons can be measured well and the invariant mass can be fully reconstructed. The processes $H \to b\bar{b}$ has a worse resolution due to larger uncertainties on the reconstructed energies of the jets originating from the b-quarks. Decay channels that contain neutrinos in the final state, e.g. $H \to WW^* \to l\nu l\nu$, do not allow for a complete reconstruction of the mass and therefore have an even worse resolution.

2.4. Current knowledge of the Higgs boson

The Higgs boson has eluded experimental proof for a long time. Its mass is a free parameter in the Standard Model. However, an upper bound is provided by perturbation theory: unitarity in the elastic scattering of vector bosons is only preserved for Higgs boson masses below 1 TeV [32]. For these masses the LHC was expected to either exclude the existence

of the Higgs boson or find evidence for it.

Before the operation of the LHC the mass of the Higgs boson was already constrained by the experiments at the LEP collider. The Higgs boson was excluded for $m_H < 114.4\,\text{GeV}$ at a 95 % confidence level by direct searches [19]. Additionally, it was possible to constrain the mass by the results of electroweak precision measurements. The theoretical predictions for these results involve loop corrections, which are sensitive to m_H. From these measurements, low masses just above the excluded range of the LEP experiments were preferred with a best-fit value of $m_H = 89^{+35}_{-26}\,\text{GeV}$ [19].

The Tevatron was able to constrain the Higgs boson mass with direct searches as well. Masses in the ranges of 100 - 108 GeV and 156 - 177 GeV were excluded [33] before the recent findings with the data from the LHC.

In July 2012 the two large experiments at the LHC, ATLAS and CMS, announced the independent observation of a new particle with a mass of around 126 GeV, which was consistent with the Standard Model Higgs boson. For this claim, the requirement of a significance of over $5\,\sigma$ was satisfied for each experiment. The results were published by the ATLAS and CMS collaborations shortly after [6,7]. These results were driven by the bosonic decay channels $H \to \gamma\gamma$, $H \to WW^* \to l\nu l\nu$ and $H \to ZZ^* \to llll$.

Later, in July 2013, the Tevatron found an excess of events in the mass range of 115 - 140 GeV in the search for the SM Higgs boson as well. The observed (expected) local significance at $m_H = 125\,\text{GeV}$ was determined to be $3.0\,\sigma$ ($1.9\,\sigma$) and was mainly driven by the search for the $H \to b\bar{b}$ decay [34].

The LHC measurements have been updated since then. The latest invariant mass spectra of the $H \to \gamma\gamma$ [35] and $H \to ZZ$ [36] searches from the ATLAS collaboration are shown in Figure 2.7. The significance of the signal is now larger than $5\,\sigma$ for each analysis alone and clear peaks are visible in the data. These analyses are combined with the corresponding measurements from the CMS collaboration to determine the mass of the Higgs boson [37]. The result is

$$m_H = 125.09 \pm 0.21\,(\text{stat.}) \pm 0.11\,(\text{syst.})\,\text{GeV}. \tag{2.43}$$

The quantum numbers of the observed resonance are determined as well. Due to the observed decay to two photons (two Z bosons, a W^+W^- pair) the resonance has to be electrically neutral. Further, according to the Landau-Yang theorem, the spin cannot be one, leaving only the spin-0 and spin-2 hypotheses possible.

The spin-2 option has been experimentally excluded by a spin and parity analysis by the ATLAS collaboration [39]. The SM Higgs boson hypothesis, corresponding to the quantum numbers $J^P = 0^+$, is tested against several alternative spin/parity scenarios, including non-SM spin-0 and spin-2 models. All alternative models are excluded at more than 99.9% confidence level in favor of the SM hypothesis. Similar studies by the CMS collaboration have come to the same conclusion [40,41]. Additionally, the spin-2 hypothesis is excluded by the observation of the Higgs boson decay to two taus, as mentioned below.

The searches for the SM Higgs boson from the ATLAS and CMS collaborations have also been combined in a measurement of the Higgs boson couplings [42]. This combination uses the results of five search channels: $H \to \gamma\gamma$, $H \to WW$, $H \to ZZ$, $H \to \tau\tau$ and $H \to b\bar{b}$. A combined fit to the data is performed to determine the signal-strength parameters in

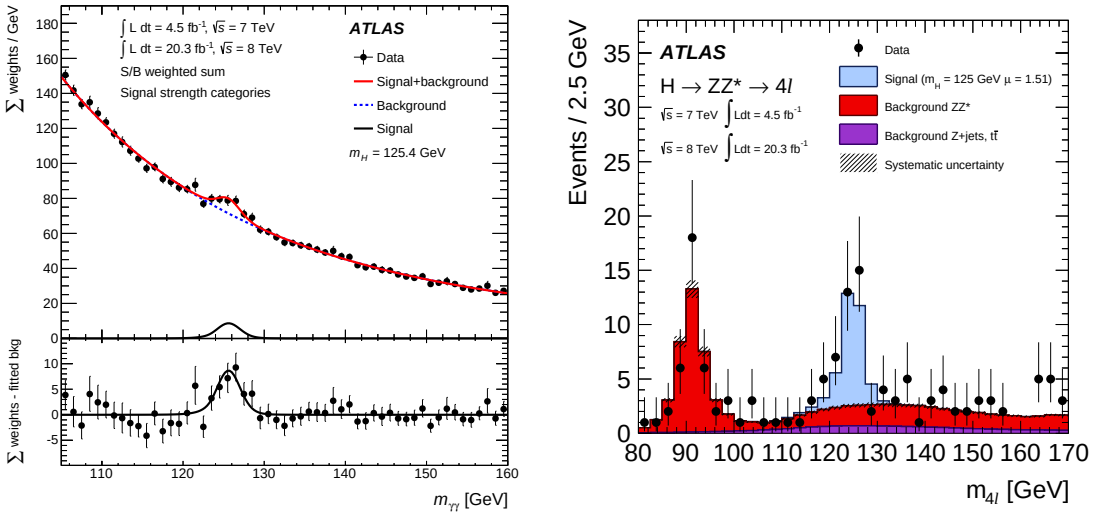


Figure 2.7.: Left (right): diphoton (four-lepton) invariant mass spectrum observed in the $H \to \gamma\gamma$ [35] ($H \to ZZ^* \to llll$ [36]) search. The data are compared to the background plus signal models with $m_H = 125.4\,\text{GeV}$ [38] ($m_H = 125\,\text{GeV}$), as determined from the measurements. The data events in the $H \to \gamma\gamma$ search are weighted by the signal-to-background ratio in the various categories of the analysis.

the various channels. They are defined as the observed production or decay strength, divided by the SM expectation: μ_i or μ^f, where i is the production mode and f is the decay channel, $i \to H \to f$.

The measured decay strengths for the five channels are shown in Figure 2.8. They are well compatible with the SM prediction. The sensitivity of each analysis can be estimated from the error bars. Although the bosonic decay channels have driven the discovery of the Higgs boson, the measurements of the fermionic channels, $H \to \tau\tau$ and $H \to b\bar{b}$, have uncertainties not much larger by now. In this combination the measurement of the $H \to \tau\tau$ decay even reaches an observed signal strength with more than $5\,\sigma$ significance.

The measurement of the $H \to b\bar{b}$ signal strength in the combination from the ATLAS and CMS collaborations has not reached a significance of $3\,\sigma$, which would be needed to claim evidence for this decay. However, this is not due to the sensitivity of the analyses (a combined significance of $3.7\,\sigma$ is expected), but due to the rather low signal strength of $\mu^{bb} = 0.69^{+0.29}_{-0.27}$ measured in the data.

Also the production strengths for the dominant modes, ggF, VBF, VH and ttH, are investigated. A measurement of the bosonic production strength, $\mu_\text{V} = \mu_\text{VBF+VH}$, versus the fermionic production via heavy quarks, $\mu_\text{F} = \mu_\text{ggF+ttH}$, is shown in Figure 2.8 (right). The results, measured separately for the five decay channels, are well compatible with the SM prediction. The $H \to b\bar{b}$ channel shows the best constraint of the bosonic production modes, while the other channels are more sensitive to the fermionic modes.

A different parameterization, the κ-framework [43], is employed to measure the Higgs boson couplings in the combined measurement from the ATLAS and CMS collaborations [42].

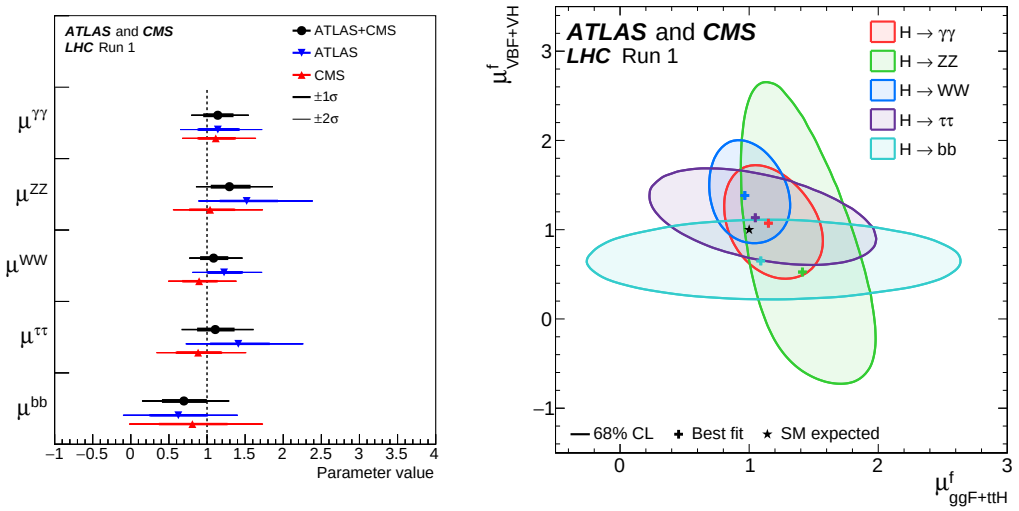


Figure 2.8.: Results for various Higgs boson signal-strength parameters from a combined ATLAS and CMS measurement [42]. Left: signal strengths for the five decay channels under consideration. Right: likelihood contours in the plane of bosonic versus fermionic production-strength for the five decay channels, where contours at the 68 % confidence level (CL) are shown.

The measurement is performed in terms of the coupling modifiers,

$$\kappa_j^2 = \sigma_j/\sigma_j^{\mathrm{SM}} \quad \text{or} \quad \kappa_j^2 = \Gamma_j/\Gamma_j^{\mathrm{SM}}, \tag{2.44}$$

which equal to unity in the Standard Model. The event yield of an individual channel $i \to H \to f$ is parameterized as

$$\sigma_i \cdot B^f = \sigma_i(\vec{\kappa})\frac{\Gamma^f(\vec{\kappa})}{\Gamma_H(\vec{\kappa})}, \tag{2.45}$$

where Γ_H is the total decay width of the Higgs boson, defined as

$$\Gamma_H(\vec{\kappa}) = \frac{\kappa_H^2 \cdot \Gamma_H^{\mathrm{SM}}}{1 - B_{\mathrm{BSM}}}. \tag{2.46}$$

Here, the branching ratio for possible BSM decays, B_{BSM}, may be fixed to zero in the fit to the data. The coupling modifier κ_H^2 is a sum over the individual couplings for the various Higgs boson decays multiplied by the corresponding SM branching ratios:

$$\begin{aligned}\kappa_H^2 = \quad & 0.57\,\kappa_b^2 + 0.22\,\kappa_W^2 + 0.09\,\kappa_g^2 + 0.06\,\kappa_\tau^2 \\ & + 0.03\,\kappa_Z^2 + 0.03\,\kappa_c^2 + 0.002\,\kappa_\gamma^2 + ...\end{aligned} \tag{2.47}$$

Clearly, this sum is dominated by the first term involving κ_b, which is parameterizing the coupling to b quarks. Other terms, in particular the ones involving κ_Z and κ_γ, provide only little to the total decay width.

The low signal strength for the decay of the Higgs boson to b quarks causes the measured $\Gamma_H/\Gamma_H^{\mathrm{SM}}$ to be smaller than unity. As a consequence, all other couplings are pulled to lower

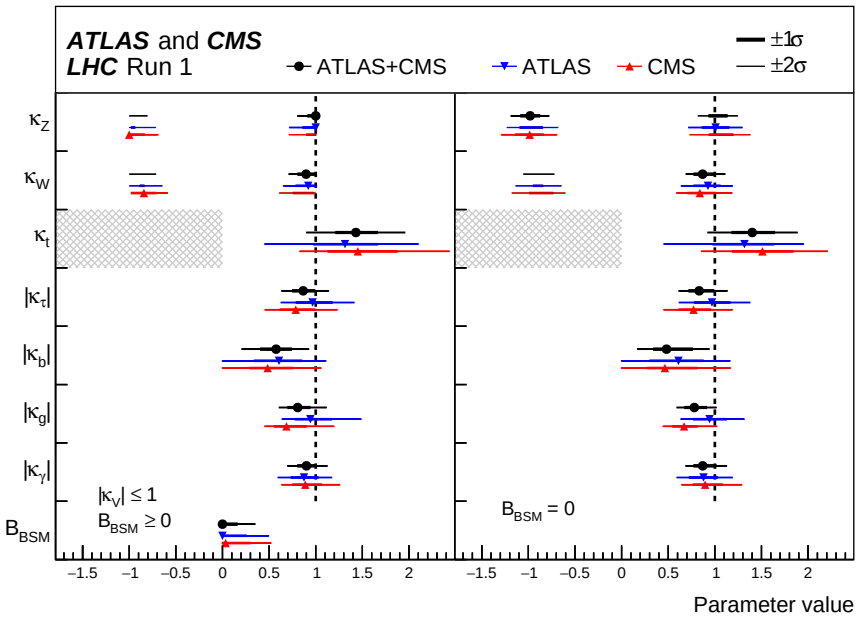


Figure 2.9.: Results for various Higgs boson coupling modifiers from a combined ATLAS and CMS measurement [42]. The branching ratio for possible BSM decays is required to be $B_{\text{BSM}} \geq 0$ on the left and $B_{\text{BSM}} = 0$ on the right.

values in the fit to compensate for the small denominator in Equation 2.45. The result is shown in Figure 2.9.

For example, the absolute value of the coupling modifier for Z bosons, $|\kappa_Z|$, is measured close to unity (for both requirements of $B_{\text{BSM}} \geq 0$ and $B_{\text{BSM}} = 0$). Instead, the signal strength parameter for the decay to Z bosons, μ^{ZZ}, is measured about $1.2\,\sigma$ above unity, as shown in Figure 2.8. This apparent discrepancy is caused by the low signal strength for $H \to b\bar{b}$, which has an impact on κ_Z, but not on μ^{ZZ}.

Hence, the $H \to b\bar{b}$ decay is particularly interesting due to the large branching ratio of 57 % predicted in the Standard Model. Its measurement is crucial to validate the total decay width of the Higgs boson, as discussed above. The search for the associated production, $VH(\to b\bar{b})$, is presented in Chapter 5. It is presently the most sensitive search for the $H \to b\bar{b}$ decay and contributed to the combined measurement of the Higgs boson couplings from the ATLAS and CMS collaborations discussed above.

2.5. Higgs boson portal models

One of the unsolved problems of the Standard Model is the apparent abundance of *dark matter* (DM) in the universe. Extensions of the SM, often called Beyond Standard Model (BSM) theories, can provide candidates for particles possibly constituting the dark matter. For example, the *Supersymmetry* (SUSY) contains a lightest supersymmetric particle (LSP), which is stable under the hypothesis of a conserved R-parity. The LSP is a good candidate for a DM particle. However, SUSY predicts many more particles, which have

not been found so far.

A simpler group of extensions of the Standard Model are Higgs boson portal models [44–47]. They predict the existence of a *hidden sector*, which does only interact with the SM through the Higgs boson. The hidden sector could contain candidates for DM particles. Possibly the simplest approach is the addition of a new scalar S, respecting a $U(1)_S$ global symmetry, to the SM Lagrangian [47]:

$$\Delta\mathcal{L}_S = -\frac{1}{2}m_S^2 S^2 - \frac{1}{4}\lambda_S S^4 - \frac{1}{4}\lambda_{HSS} H^\dagger H S^2 \tag{2.48}$$

Here, the last term yields the interaction of S with the SM Higgs boson, H. Within this approach, the relic abundance of dark matter particles is obtained through the s-channel annihilation of energetic SM particles via the exchange of the Higgs boson.

The possible decay of the Higgs boson to such scalars would be invisible to the ATLAS detector, hence it is denoted by $H \to \text{inv.}$ or $H \to \chi\chi$ in the following. Its detection, however, would be possible by reconstructing an excess of events with large missing transverse momentum, $E_{\mathrm{T}}^{\mathrm{miss}}$, over the expected background yield from SM processes. The large $E_{\mathrm{T}}^{\mathrm{miss}}$ in such events would be caused by the recoil of the Higgs boson against the initial state particles.

A limit on the invisible branching ratio of the Higgs boson has been set as $\mathrm{BR}(H \to \text{inv.}) < 0.25$ by a combination of several direct searches performed by the ATLAS collaboration [48]. One of these searches, using the associated production of the Higgs boson with a hadronically decaying vector boson, $VH \to q\bar{q}'\chi\chi$, is presented in Chapter 7.

2.6. Heavy Vector Triplet models

The Higgs boson mass in the Standard Model undergoes radiative corrections from loops of virtual particles, most importantly top-quark loops. These corrections are quadratically divergent with the momentum in the loop and make it unnatural for the SM to be valid beyond a scale of a few TeV. This is known as one of the hierarchy problems of the SM.

Several extensions of the SM try to solve this problem by assuming a new strong interaction at a higher scale. Such models often predict the existence of new heavy resonances, which can decay to SM particles, e.g. to a vector boson and a Higgs boson. Examples for such models are the Minimal Walking Technicolor [49–51], Little Higgs [52], or composite Higgs boson models [53, 54].

These models are quite involved. However, simpler benchmark models serve the purpose of an experimental search better. The Heavy Vector Triplet (HVT) model [55] predicts the existence of new heavy vector bosons ($V' = W'$ or Z' boson), which can decay to SM bosons and fermions.

A simple phenomenological Lagrangian is employed, $\mathcal{L}_V$, which describes the interaction with the SM particles. It contains the terms [55]

$$\mathcal{L}_V \supset i g_V c_H V_\mu^a H^\dagger \tau^a \widehat{D}^\mu H + \frac{g^2}{g_V} c_F V_\mu^a J_F^{\mu a}, \tag{2.49}$$

where g_V, c_H and c_F are new coupling constants and $2\tau^a = \sigma^a$ are the Pauli matrices. The generalized covariant derivative in the first term, $\widehat{D}$, with $H^\dagger \widehat{D}^\mu = H^\dagger D^\mu + D^\mu H^\dagger$, is

defined by

$$D_\mu V_\nu^a = \partial_\mu V_\nu^a + g\epsilon^{abc} W_\mu^b V_\nu^c, \tag{2.50}$$

and allows for the decay $V' \to VH$. The second term describes the coupling to the left-handed fermion current of the SM, $J_F^{\mu a}$. Further terms allow decays such as $V' \to VV$ with other coupling constants.

Two specializations are further discussed: the HVT Models A and B. The Model A represents an extended gauge symmetry with an symmetry breaking pattern of $SU(2)_1 \times SU(2)_2 \times U(1)_Y \to SU(2)_L \times U(1)_Y$ [56], which leads to the vector triplet. The Model B represents a minimal composite Higgs boson model and the vector triplet emerges from the breaking of a global $SO(5)$ symmetry to an $SO(4)$ subgroup [57]. These two models lead to specific constraints on the coupling constants [55]:

$$\text{Model A:} \quad c_H \sim -g^2/g_V^2 \quad \text{and} \quad c_F \sim 1 \tag{2.51}$$

$$\text{Model B:} \quad c_H \sim c_F \sim 1 \tag{2.52}$$

In a search for W' resonances, performed during Run 1 of the LHC, the ATLAS collaboration has found an interesting excess in the invariant mass spectrum around $m_{W'} = 2.0\,\text{TeV}$, as shown in Figure 2.10 (left). The analysis uses two jets to reconstruct hadronically decaying W or Z bosons, each as one jet, from the hypothetical W' decay. The local excess corresponds to a deviation from the background spectrum with a significance of $3.4\,\sigma$. However, this excess is not confirmed by W' searches in the leptonic decay channels. The significance is reduced to $2.5\,\sigma$ in the combination [58].

Interestingly, the CMS collaboration has also found an excess in the search for W', but in the decay $W' \to WH(\to bb)$. The excess is seen for an invariant mass close to the one above with about $m_{W'} = 1.9\,\text{TeV}$, as shown in Figure 2.10 (right). It corresponds to a local significance of $2.9\,\sigma$ [60].

These excesses are to be confirmed or excluded with the data to be taken during Run 2 of the LHC. With the increased center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ the parton luminosities for producing heavy resonances with masses above 1 TeV increase dramatically, as shown in Figure 2.11. For example the parton luminosity for a quark-induced resonance with a mass of $M_X = 2\,\text{TeV}$ increases by a factor of about eight. Hence, a large increase of sensitivity is to be expected with a similar amount of luminosity. The search for V' resonances in the channel $V' \to VH(\to b\bar{b})$ with the first data taken during Run 2 in 2015 with the ATLAS experiment is presented in Chapter 8.

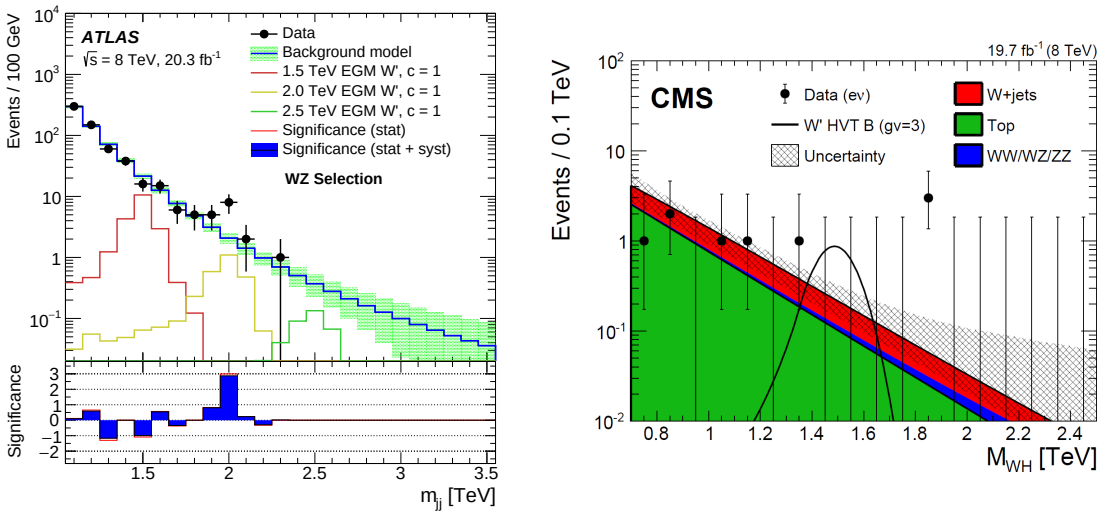


Figure 2.10.: Invariant mass spectra in W' searches performed during Run 1 of the LHC. Left: $W' \to WZ \to qqqq$ search from the ATLAS collaboration [59]. Each vector boson is reconstructed within one large jet, such that the invariant mass is defined by the dijet mass, m_{jj}. Right: $W' \to WH \to e\nu bb$ from the CMS collaboration [60]. Both spectra show excesses for a reconstructed mass of about 2 TeV.

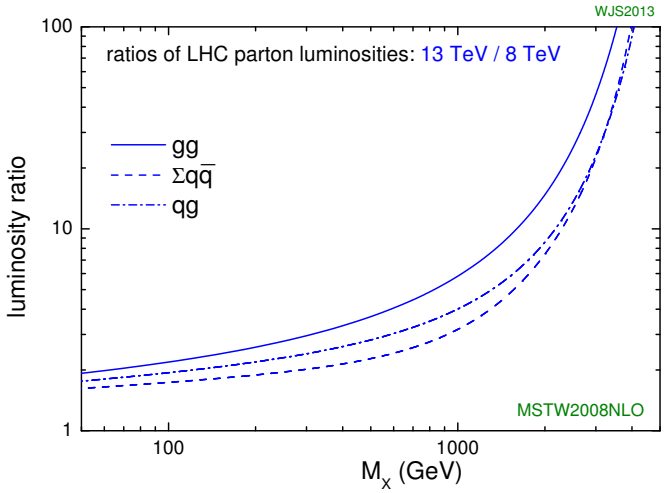


Figure 2.11.: Ratio of the parton luminosities for $\sqrt{s} = 13\,\text{TeV}$ relative to $\sqrt{s} = 8\,\text{TeV}$ as a function of the invariant mass of the two interacting partons, M_X. The ratio is given for the gg, $q\bar{q}$ and qg initial states [20].

3. The ATLAS experiment

The ATLAS[1] experiment is a multi-purpose detector operating at the Large Hadron Collider (LHC), located at CERN near Geneva, Switzerland. It is designed to address searches for new physics as well as precise tests of the Standard Model (SM). Its first discovery was the SM Higgs boson in the year 2012 together with the CMS[2] experiment [6, 7]. The following description of the ATLAS experiment is based on Refs. [61, 62].

3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) is a circular collider, mainly used for proton-proton collisions. It was built inside a tunnel with 27 km circumference and about 100 m below ground-level, formerly housing the Large Electron-Positron Collider (LEP). It was planned in the 1990s and installed after the shutdown of LEP in the year 2000.

The LHC is able to accelerate protons and heavy ions in two beams running in opposite directions. The beams are bent by dipole magnets, generating fields up to 8.3 T. These magnets are super-conducting and cooled down to 1.9 K by liquid helium. In 2011 a maximum beam energy of 3.5 TeV for protons was reached, thus a maximum center-of-mass energy of $\sqrt{s} = 7$ TeV. The energy was increased to $\sqrt{s} = 8$ TeV in 2012 and further to $\sqrt{s} = 13$ TeV in 2015, which is close to the design energy of $\sqrt{s} = 14$ TeV. The data taking in 2011 and 2012 is referred to as *Run 1* and the 2015 period as *Run 2* in the following.

Protons are fed into the LHC with an energy of 450 GeV, pre-accelerated by a chain of linear and circular accelerators, as shown in Figure 3.1. The accelerator chain consists of a linear accelerator (LINAC) and three circular accelerators, the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS).

The particle beams in the LHC are circulating in two evacuated beam pipes and are bent by the fields of 1232 super-conducting dipole magnets, operating at a temperature of 1.9 K. The maximum beam energy of the LHC is limited by their field strength of up to 8.33 T. Eight superconducting cavities, operating at a frequency of 400 MHz, are generating the electric fields for the particle acceleration. In addition, 392 quadrupole magnets are used to focus the beams.

During the acceleration phase, the particles gain 485 keV energy per turn. Once accelerated, the life time of the beams is limited to about $\tau = 15$ h, mainly by the luminosity loss due to collisions at the interaction points. At design conditions, 2808 bunches per beam are circulating in the LHC, each of them containing about 10^{11} protons.

[1] **A** **T**oroidal **L**HC **A**peratu**S**
[2] **C**ompact **M**uon **S**olenoid

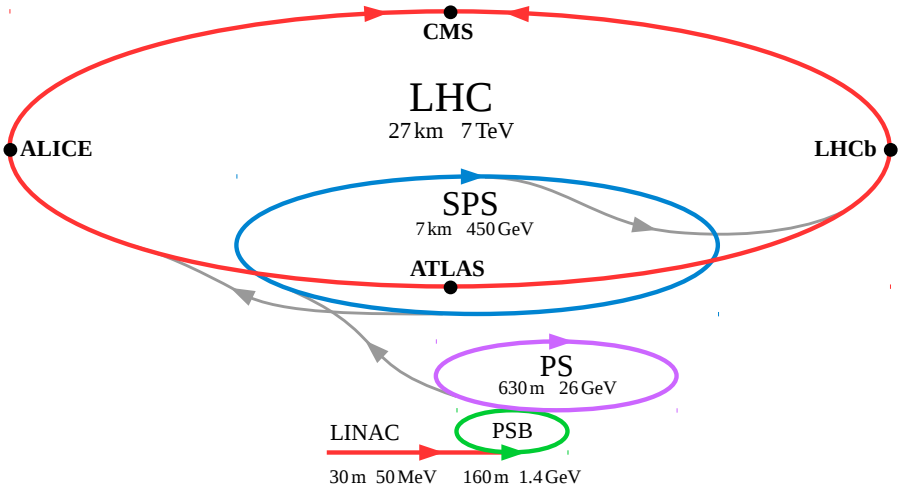


Figure 3.1.: Schematic view of the LHC accelerator chain for protons. Shown are the individual accelerators with their names, lengths and design proton energies, and the location of the four experiments at the LHC, as described in the text.

The beams of the LHC are brought to collision at four linear sections of the tunnel. Two of these interaction points are designed with a crossing angle of 285 μrad for an instantaneous luminosity of $L = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. This is where the two multi-purpose detectors ATLAS and CMS are located. At the other two interaction points the LHCb[3] and ALICE[4] experiments are located. This thesis is based on data taken with the ATLAS experiment, which is described in the following.

3.2. The ATLAS detector

The ATLAS detector is a multi-purpose detector, aiming at the discovery of new particles, such as the Higgs boson and physics beyond the Standard Model in proton-proton collisions. Its cylindrical 4π-design with high density of detection material up to low scattering angles allows for near-complete reconstruction of hard scattering processes.

The common coordinate system used for measurements is centered in the beam crossing, with the z-axis along the beam, the y-axis pointing upwards and the x-axis pointing towards the center of the LHC ring. The azimuthal angle ϕ is measured in the x-y-plane with respect to the x-axis and the polar angle θ is the angle to the z-axis. A commonly used coordinate, the pseudo-rapidity η, is defined as $\eta = -\ln(\tan(\theta/2))$.

The detector consists of several sub-detectors: the inner detector with a coverage of $|\eta| < 2.5$, the calorimeters with $|\eta| < 4.9$ and the muon system with $|\eta| < 2.7$, as shown in Figure 3.2. The ATLAS detector has a multi-stage triggering system to cope with the high interaction rates occurring at the LHC. The subsystems of the detector are described in more detail in the following.

[3] **L**arge **H**adron **C**ollider **b**eauty
[4] **A** **L**arge **I**on **C**ollider **E**xperiment

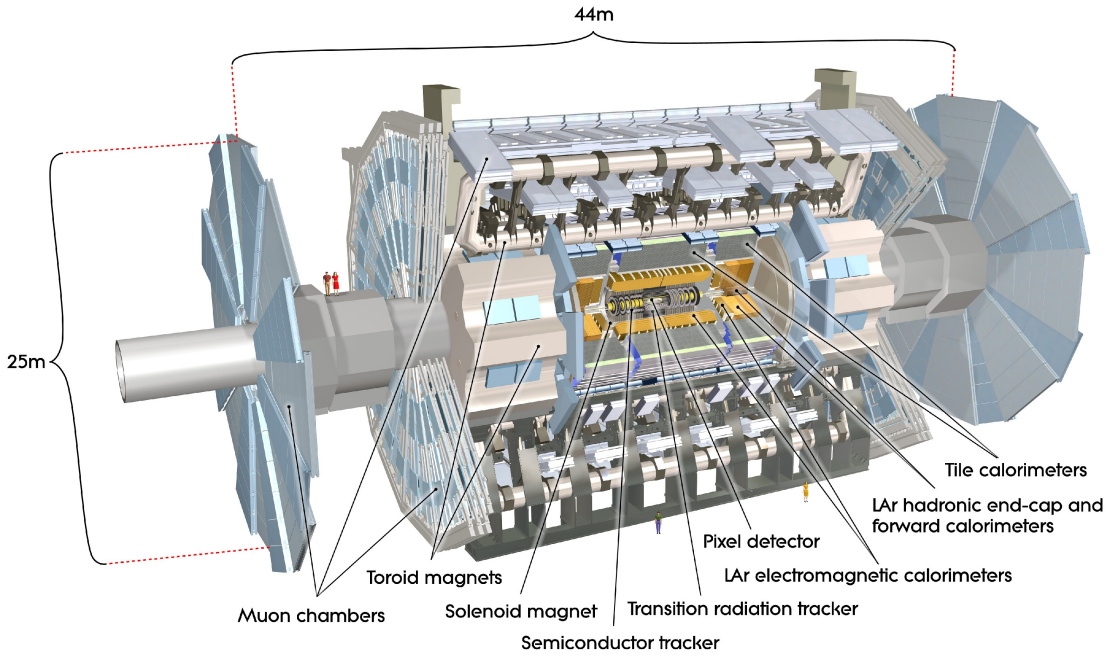


Figure 3.2.: Schematic view of the ATLAS detector [62]. The visible components are the muon system, the toroidal and solenoid magnets, the calorimeters and the inner detector.

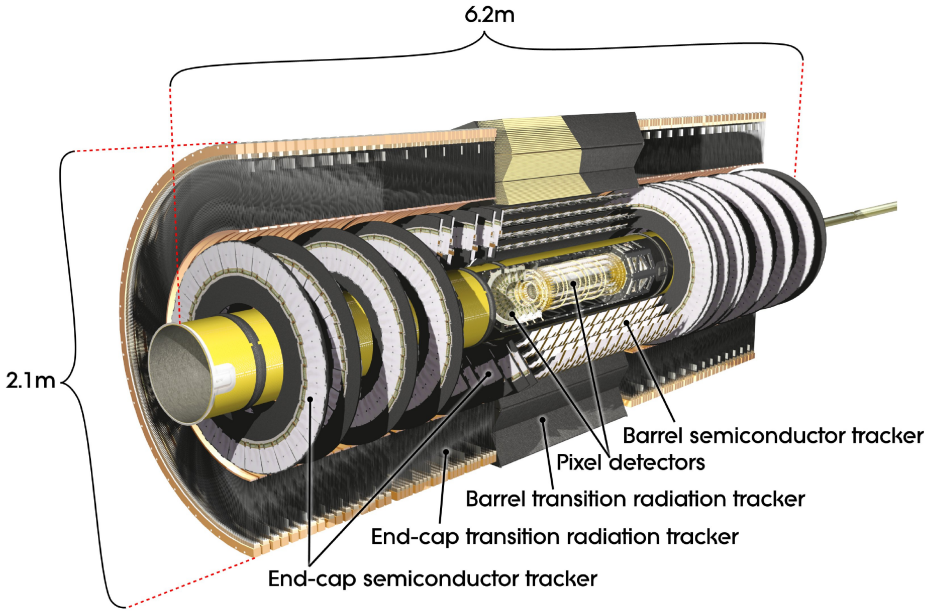


Figure 3.3.: Sketch of the inner detector with its pixel detectors, SCT and TRT [62].

3.2.1. Inner detector

The inner detector consists of pixel detectors, semiconductor trackers (SCTs) and transition radiation trackers (TRTs), as shown in Figure 3.3. The pixel detector allows precise measurements of tracks from charged particles close to the interaction point. This enables reliable reconstruction of the primary interaction point and of secondary vertices. Together with the surrounding tracking detector and a solenoid, generating a magnetic field of 2 T, the track momenta can be determined with a design resolution of[5] $\sigma_{p_T}/p_T = 0.05\%\, p_T/\text{GeV} \oplus 1\%$ [62].

Insertable B-Layer The *Insertable B-Layer* (IBL) was installed during the the first shutdown of the LHC between its Runs 1 and 2. It is a pixel detector and represents the closest tracking detector to the beam pipe with a mean radius of $R = 33\,\text{mm}$ [63] and a granularity of 250x50 μm [64]. Its main purpose is the improvement of the reconstruction of secondary vertices, such as from b-hadron decays. Due to its vicinity to the interaction point the IBL has to cope with high radiation doses. This led to the development of specific read-out electronics and the use of so-called 3D sensors, which are specifically radiation hard [64].

Pixel detector The pixel detector is the next closest detector to the interaction point, the innermost layer being located at $R = 50.5\,\text{mm}$. It consists of cylindrical layers ordered around the beam in the central region (barrel) and of disks arranged radially in the end caps. Particles typically pass three of these layers. The pixels have a minimum size of

[5] Using the notation $a \oplus b = \sqrt{a^2 + b^2}$.

R-$\phi \times z = 50 \times 400\,\mu\text{m}^2$ and a total of 80 million read-out channels is reached. The track resolution in the barrel (end caps) is about $10\,\mu$m in the R-ϕ-plane (along ϕ) and about $115\,\mu$m along the z-axis (R-axis) [62].

Semiconductor tracker The SCT surrounds the pixel detector. It consist of four layers of silicon strip detector modules in the barrel and nine layers in the end caps. Particles typically pass four of these layers. The modules in the barrel have two layers of silicon that are sightly rotated against each other, which allows for the determination of the position along the strips. The resolution in the barrel (end caps) is about $17\,\mu$m in the R-ϕ-plane (along ϕ) and about $580\,\mu$m along the z-axis (R-axis) [62].

Transition radiation tracker The TRT surrounds the SCT and is the outermost part of the inner detector. It consists of gas-filled drift tubes along the z-axis in the barrel (144 cm long) and along the radial direction in the end caps (37 cm long). The tubes have a diameter of 4 mm and a thin wall out of two $35\,\mu$m thick multi-layer films [62]. The TRT contains a large number of polypropylene fibers with a diameter of $9\,\mu$m, which serve as transition radiation material.

In the barrel the TRT allows for the position measurement in the R-ϕ-plane with a resolution of about $130\,\mu$m. The TRT provides on average 36 coordinate measurements over the radial distance from 55 cm to 108 cm [62]. By having a larger radius than the other trackers it improves significantly the resolution of the momentum measurement. In addition it provides potential for particle identification. The transition radiation, emitted by particles crossing the interface of two materials with different dielectric constants, is proportional to $\gamma = E/mc^2$ [65]. This is especially important for electrons, since they are by far the lightest stable charged particles and therefore emit the most transition radiation.

Since the components of the inner detector are very close to the interaction point, they have to cope with high radiation which damages the detector material. To reduce this damage the pixel detector and the SCT are cooled down to about $-7°$C, while the TRT operates at room temperature. At the design luminosity of the LHC a total of about 85 kW of heat has to be removed from the inner detector, which is done with an elaborate cooling system [62].

3.2.2. Calorimeters

The calorimeter system of the ATLAS detector consists of an electromagnetic ($|\eta| < 3.2$) and a hadronic calorimeter ($|\eta| < 4.9$), covering the full ϕ-range, as shown in Figure 3.4. Both are sampling calorimeters, which means they use layers of active detector material and absorber material. Their main purpose is the determination of particle energies. The calorimeters are *non-compensating*, meaning they have a different response to electromagnetic and hadronic showers.

Electromagnetic calorimeter The electromagnetic (EM) calorimeter is a liquid argon (LAr) sampling-calorimeter with lead absorbers and kapton electrodes. They are arranged in an accordion geometry, enabling a full ϕ-coverage without gaps. The calorimeter is

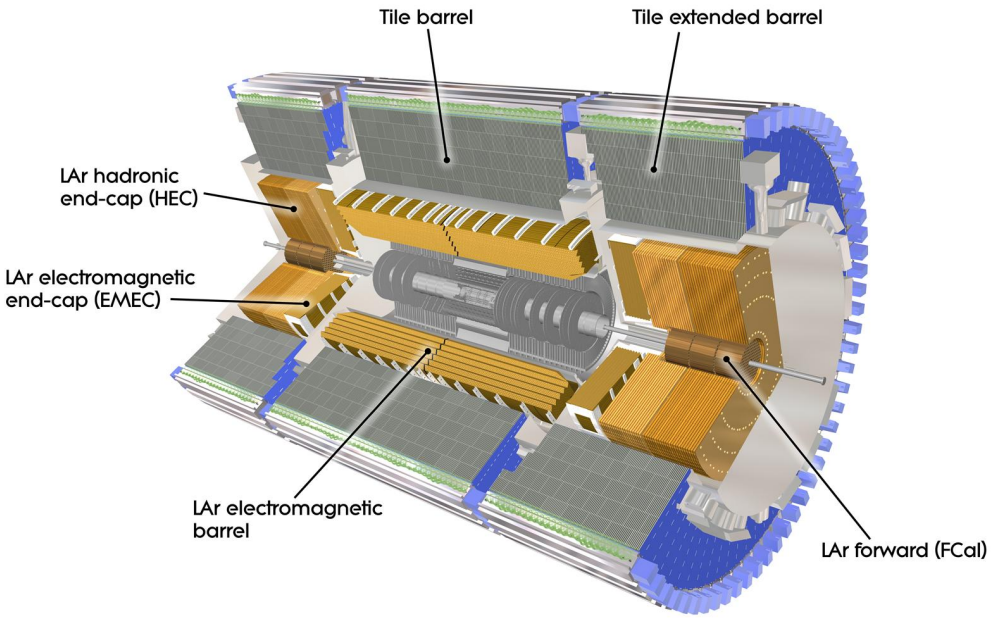


Figure 3.4.: Sketch of the ATLAS calorimeters with their electromagnetic and hadronic subcomponents [62].

divided into a barrel region ($|\eta| < 1.475$) and end caps (EMEC, $1.375 < |\eta| < 3.2$). The end caps are further divided into a central region ($1.375 < |\eta| < 2.5$) with finer granularity and a forward region ($2.5 < |\eta| < 3.2$) with coarser granularity. It has a design energy resolution of $\sigma_E/E = 10\%/\sqrt{E/\text{GeV}} \oplus 0.7\%$ [62].

Hadronic calorimeter The hadronic calorimeter has a sampling structure as well. In the central region ($|\eta| < 1.7$) it is called tile calorimeter and uses scintillating tiles and steel as absorber. The hadronic end caps (HEC) at the outer region ($1.5 < |\eta| < 3.2$) use liquid argon (LAr) as detector material and copper as absorber. The design resolution of the hadronic barrel and end caps for jets is $\sigma_E/E = 50\%/\sqrt{E/\text{GeV}} \oplus 3\%$ [62].

Forward calorimeter The forward calorimeter (FCal) is dedicated to the very forward region ($3.1 < |\eta| < 4.9$) and has only three layers of absorber material. The first one is copper, optimized for electromagnetic measurements and the other two are made of tungsten for hadronic measurements. The design resolution of the forward calorimeter for jets is $\sigma_E/E = 100\,\%/\sqrt{E/\text{GeV}} \oplus 10\,\%$ [62].

3.2.3. Muon spectrometer

The muon spectrometer is the outermost and largest sub-system of the ATLAS detector. Its toroidal magnet system does not only give ATLAS its name, but also is responsible for

the distinct look of the detector, shown in Figure 3.2.

It is divided into three regions: barrel ($|\eta| < 1.4$), end caps ($1.6 < |\eta| < 2.7$) and transition region ($1.4 < |\eta| < 1.6$). It provides triggering capabilities for muons up to $|\eta| < 2.4$. Three separate magnet systems are used for generating the bending fields: one in the barrel and the two end caps. In the transition region their overlapping fields are used. Each magnet system consist of eight coils with air-cores, arranged in a toroidal geometry. They produce a magnetic field integral of $\int B_{\perp} dl = 1.5$ to $5.5\,\mathrm{Tm}$ in the barrel, 1.0 to $7.5\,\mathrm{Tm}$ in the end caps and less in the transition region. Due to the toroidal geometry, charged particles are bend orthogonal to the bending direction in the inner solenoid.

Monitored drift tube (MDT) chambers are installed in the full η-range up to $|\eta| = 2.7$ for track detection. For $2.0 < |\eta| < 2.7$ cathode-strip chambers provide better spatial resolution and can cope with higher signal rates. The trigger system consists of resistive-plate chambers in the barrel and thin-gap chambers in the end caps. The muon spectrometer allows for precise measurement of large muon momenta. For a track crossing three MDT chambers a resolution of $\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 9\,\% \, p_\mathrm{T}/\mathrm{TeV}$ is expected, satisfying the design goal of $\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 10\,\%$ for $p_\mathrm{T} = 1\,\mathrm{TeV}$ [62].

3.2.4. Trigger system

The design event-rate provided by the LHC is $40\,\mathrm{MHz}$. This rate has to be reduced drastically to a few hundred Hertz to be able to record the events to disk or tape. The decision, which events are recorded, is taken by a multi-level triggering system [62].

The level-1 (L1) trigger uses the trigger chambers of the muon system and the full calorimeter system with reduced granularity to search for objects with high transverse energy, e.g. electrons, photons, muons, jets or missing transverse energy. The L1 trigger is hardware-based and a decision is taken within $2.5\,\mu\mathrm{s}$. It reduces the event rate to $75 - 100\,\mathrm{kHz}$. The remaining events are passed to the level-2 trigger system (L2) together with η and ϕ information about so-called Regions of Interest (RoIs).

The level-2 (L2) trigger is software-based and uses the full detector information inside the RoIs and brings the event rate down to about $3.5\,\mathrm{kHz}$. The event filter (EF) is the final triggering stage, reducing the rate to about $400\,\mathrm{Hz}$. It uses the full detector information and needs about $4\,\mathrm{s}$ for a decision. All events passing the EF are recorded for further offline analysis. A typical event has a size of 1 to $2\,\mathrm{MB}$ of raw data.

This three-level triggering system has been optimized for Run 2 of the LHC, resulting in a two-level system. The output from the hardware-based L1 trigger is now directly passed to the final algorithm, the high-level trigger (HLT). The HLT is software-based and can make use of the full event information.

4. Reconstruction of physical objects

The reconstruction of physical objects is presented in this chapter. For each type of object several sets of selection criteria are applied to fit the needs of the analyses presented in this thesis. The following description focuses on the procedure of the analyses carried out with the dataset recorded at $\sqrt{s} = 8\,\text{TeV}$. Minor differences to the analyses with $\sqrt{s} = 7$ and 13 TeV exist and are pointed out where appropriate.

The reconstruction of tracks and vertices is described in Section 4.1, of leptons in Section 4.2 and of jets in Section 4.3. An important aspect of the reconstruction for this thesis is the flavor tagging of jets, described in Section 4.4. Further, the overlap removal is described in Section 4.5 and the reconstruction of missing transverse energy in Section 4.6. A summary of the systematic uncertainties, assigned to the reconstruction of all objects, is given in Section 4.7.

4.1. Tracks and vertices

The reconstruction of charged particles, their trajectories and momenta, is done by reconstructing their tracks in the detector. Due to the high resolution of the inner detector tracks provide precise spatial information. They are vital for identifying electrons and muons (Section 4.2). The space points where charged particles merge can be reconstructed as vertices, which play an important role in the jet selection (Section 4.3) and for the flavor tagging (Section 4.4).

The reconstruction of tracks starts from three-dimensional space points, reconstructed from energy deposits ("hits") of charged particles in the inner detector components, in the pixel, SCT and TRT detectors [66]. So-called track seeds are generated using three or more hits in the pixel detector and the first layer of the SCT. These are extended to the full SCT and a first track fit is performed. Quality requirements on the χ^2 of the fit and the number of hits are applied to reject tracks that are not caused by a real charged particle, but by accidental alignment of some hits ("fake tracks"). The remaining tracks are associated to drift circles in the TRT and, finally, the fit is redone using the full information of the three detector systems. This is called the inside-out reconstruction.

Another algorithm, the outside-in reconstruction, starts the track reconstruction from the TRT. This approach does not require hits in the innermost detector layers and therefore has sensitivity to particles originating from secondary vertices, which are very distant to the primary interaction point. These can e.g. originate from interactions in the detector material, like electrons from photon conversions, hadronic interactions, or decays of long-lived particles, such as the the K_L^0 meson or the Λ^0 baryon.

Vertices are reconstructed from the tracks, which are extrapolated to find intersections between them. For primary vertices, the tracks are required to fulfill $p_\text{T} > 150\,\text{MeV}$,

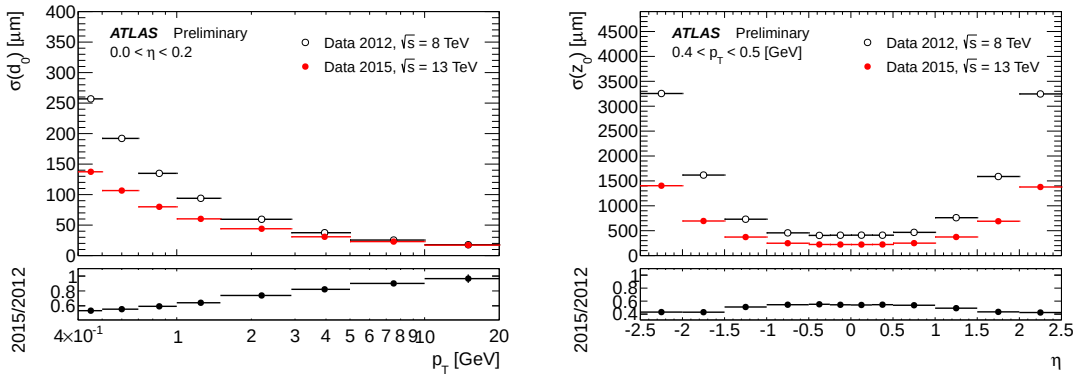


Figure 4.1.: Transverse impact parameter resolution as a function of the track p_{T} (left) and longitudinal impact parameter resolution as a function of the track η (right). The measurement performed with the data collected in 2012 with $\sqrt{s} = 8\,\mathrm{TeV}$ (black) is compared to the one with the data collected in 2015 with $\sqrt{s} = 13\,\mathrm{TeV}$ (red), including the newly-installed IBL (from Ref. [70]).

$|d_0| < 4\,\mathrm{mm}$, $\sigma(d_0) < 5\,\mathrm{mm}$ and $\sigma(z_0) < 5\,\mathrm{mm}$ [67]. Here, d_0 is the transverse *impact parameter*: the distance of the track to the beam spot in the point of closest approach, and z_0 is the longitudinal impact parameter, measured at the same point of the track.

In an iterative procedure, a first vertex seed is found as the maximum in the longitudinal impact parameter distribution. The position of the vertex is then determined from a χ^2-fitting algorithm [68, 69], which uses the vertex seed and the tracks as input. Tracks that are incompatible with the vertex with more than $7\,\sigma$ are removed are removed from the vertex. This procedure is repeated until all track are associated to a vertex, or no additional vertex can be found. For the reconstruction of secondary vertices less stringent requirements on the track positions are applied for recovering the decay of particles with larger lifetimes.

The impact parameter resolution is an important quantity for the accuracy of the vertex reconstruction, and therefore for the flavor tagging (Section 4.4). The resolution in both, the transverse and longitudinal impact parameters, was significantly improved in 2015 with the installment of the new IBL detector (Section 3.2.1). It is compared with and without the IBL in Figure 4.1. The $\sigma(d_0)$ is reduced in particular for tracks with low transverse momenta up to about 50 %, and the $\sigma(z_0)$ is even reduced up to about 60 %.

Since the total interaction cross section at the LHC is very large, multiple primary vertices for each event are expected, which is called pile-up. The vertex that corresponds to the hardest interaction is identified as *signal vertex*. It is required to have the largest sum of squared transverse momenta,

$$\sum_{i=1}^{N_{\mathrm{track}}} p_{\mathrm{T,i}}^2, \tag{4.1}$$

where the index i runs over all outgoing tracks of the vertex. Here, a requirement of a minimal number of tracks can be applied, such that vertices with lower N_{track} are ignored

for determining the signal vertex. Important track properties are the radial distance, d_0, and the longitudinal distance, z_0, of the closest approach to the signal vertex.

4.2. Leptons

The reconstruction of charged leptons, electrons and muons, is described in the following. Tau leptons are not considered in this thesis. Leptons are categorized with increasing purity into *loose*, *medium* and *tight*, based on various quality requirements.

In general, a minimal transverse momentum of $p_{\mathrm{T}} > 7\,\mathrm{TeV}$ is required for leptons. This rather low threshold was chosen to maximize the acceptance of the signal involving $Z \to ll$ and the rejection of the $t\bar{t}$ background. All leptons must pass the overlap procedure as described in Section 4.5, while the electron and muon-specific criteria are described in the following.

Electrons Electron candidates are reconstructed from energy deposits (clusters) in the EM calorimeter. An associated tracks in the inner detector is required, which aims to distinguish electrons from photons. Therefore, the electron reconstruction is only available up to $|\eta| < 2.47$.

using the standard ATLAS sliding window algorithm [71]. A so-called *likelihood identification* (LH ID) is employed, which combines various identification variables [72]. It reaches factors of two (low p_{T}) to three (high p_{T}) better rejection of fake electrons induced by jets, compared to the previous cut-based approach [73].

The product of the efficiencies for reconstructing and identifying an electron are about 85 % (65 %) at low transverse energy, E_{T}, and go up to 95 % (90 %) at $E_{\mathrm{T}} = 95\,\mathrm{GeV}$ for the so-called LooseLH (VeryTightLH) ID criteria. The efficiencies are measured using a *tag-and-probe* method in $Z \to ee$, $Z \to ee\gamma$ and $J/\psi \to ee$ events, reconstruted in the data [73]. Clean samples of theses processes are selected with tight criteria for one of the electrons ("tag") and looser criteria for the other electron ("probe"). The latter is then used to estimate the efficiency for a tighter selection.

The efficiencies are shown as a function of E_{T} in Figure 4.2 (left). The discrepancy between the data and the MC is corrected for using event-weight scale-factors, parameterized in the E_{T} and η of the electrons. The associated systematic uncertainty is of order 1-2 % for electrons with $E_{\mathrm{T}} > 25\,\mathrm{GeV}$. The electron energy is calibrated with an scale uncertainty of below 1 % for $E_{\mathrm{T}} > 10\,\mathrm{GeV}$. The energy resolution is about 2 % at $E_{\mathrm{T}} = 25\,\mathrm{GeV}$ and 1 % at $E_{\mathrm{T}} = 200\,\mathrm{GeV}$ [71].

In the analyses presented in this thesis, the VeryLooseLH ID is applied for loose electrons with $E_{\mathrm{T}} > 7\,\mathrm{GeV}$, while $E_{\mathrm{T}} > 25\,\mathrm{GeV}$ is required for medium electrons. The VeryTightLH ID is additionally applied for tight electrons.

Muons Several algorithms are used to reconstruct muons: *combined* (CB) muons require independent tracks in the inner detector (ID) and the muon system (MS). CB muons are of highest quality, but least acceptance. *Segment-tagged* (ST) muons require a track in the ID and only one hit in the MS, which allows for a lower p_{T} of the muon. *Stand-alone* muons on the other hand require only a track in the MS, which increases the acceptance

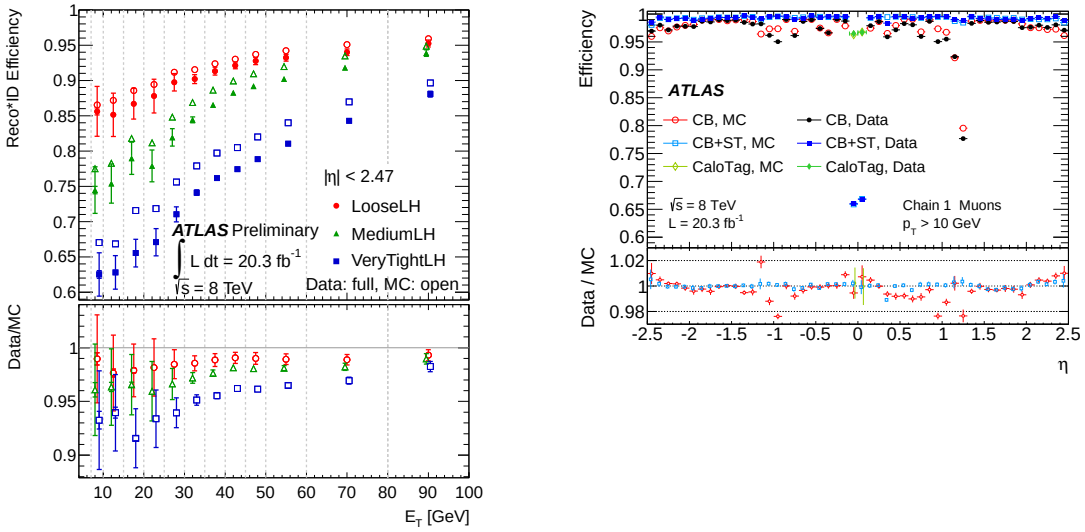


Figure 4.2.: Left: electron reconstruction times identification efficiency as a function of the electron E_{T} [73]. Right: muon reconstruction efficiency as a function of the muon η [74].

from $|\eta| < 2.5$ to $|\eta| < 2.7$. These algorithms, which rely on the MS, are inefficient for $|\eta| < 0.1$, due to the supporting structure of the ATLAS detector. *Calorimeter-tagged* (CaloTag) muons are used to increase the acceptance in this region. They require only a track in the ID and in addition energy deposits in the calorimeter, which are compatible with a minimum ionizing particle.

The reconstruction efficiency by requiring either of these four algorithms is above 99 % for most of the covered phase space of $|\eta| < 2.7$ and $5 \leq p_{\mathrm{T}} \leq 100\,\mathrm{GeV}$ [74]. It is measured similarly as for electrons with a tag-and-probe method in $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events. The muon reconstruction efficiency as a function of η is shown in Figure 4.2 (right). It is measured with per-mille precision, such that the related uncertainty is of minor importance in this thesis.

The momentum resolution uncertainties are measured in $Z \to \mu\mu$, $J/\psi \to \mu\mu$ and $\Upsilon \to \mu\mu$ events [74]. They range from 1.7 % in the central region and for transverse momenta of $p_{\mathrm{T}} \simeq 10\,\mathrm{GeV}$, to 4 % at large rapidity and $p_{\mathrm{T}} \simeq 100\,\mathrm{GeV}$. The momentum scale uncertainty is below 0.2% and is neglected in this thesis.

The loose muon selection allows CB and ST muons in $|\eta| < 2.7$ and SA muons in $2.5 < |\eta| < 2.7$, both with $p_{\mathrm{T}} > 7\,\mathrm{GeV}$. Their tracks are required to fulfill $|d_0| < 0.1\,\mathrm{mm}$ and $|z_0| < 10\,\mathrm{mm}$. In addition, CaloTag muons are allowed for $|\eta| < 0.1$ and $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ without further requirements on the track. The medium selection allows only CB and ST muons with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$. Stricter isolation cuts are applied to tight muons, as described above.

Isolation Important lepton selection criteria are the track-based isolation, i_{trk}, and the calorimeter-based isolation, i_{calo}, defined as:

$$i_{\text{trk}}(\ell) = \frac{1}{p_T^{\ell}} \sum_{\text{trk}} p_T^{trk} \quad \text{for} \quad \Delta R(\ell, trk) < 0.2 \tag{4.2}$$

$$i_{\text{calo}}(\ell) = \frac{1}{p_T^{\ell}} \sum_{clus} E_T^{clus} \quad \text{for} \quad \Delta R(\ell, clus) < 0.3. \tag{4.3}$$

Here, *trk* (*clus*) are the tracks (clusters) around the lepton, excluding the track (clusters) used to reconstruct the lepton itself. Loose leptons must pass $i_{\text{trk}} < 0.1$, while for tight leptons $i_{\text{trk}} < 0.04$ and $i_{\text{calo}} < 0.04$ are required.

4.3. Jets

Jets are reconstructed by grouping energy deposits in the calorimeter into clusters. These are then combined by the anti-k_t algorithm [75] into jets with a radius of $R = 0.4$ (small jets) or $R = 1.0$ (large jets). The anti-k_t algorithm defines a measure of distance d_{ij} between clusters in the calorimeter,

$$d_{ij} = \min \left\{ \frac{1}{k_{\text{T},i}^2}, \frac{1}{k_{\text{T},j}^2} \right\} \times \frac{(\Delta R_{ij})^2}{R^2}, \tag{4.4}$$

where $k_{\text{T},i}$ is the transverse momentum of cluster i and $\Delta R_{ij} = \sqrt{(\Delta y_{ij})^2 + (\Delta \phi_{ij})^2}$ is the geometrical distance between clusters i and j with y being the rapidity. The algorithm calculates a minimal distance d_{min} in the event from the list of all clusters i and j:

$$d_{\text{min}} = \min\{d_{ij}, d_{iB}\}, \tag{4.5}$$

where the stopping criterion, d_{iB}, is defined as the distance of cluster i to the beam axis B:

$$d_{iB} = \frac{1}{k_{\text{T},i}^2}. \tag{4.6}$$

If $d_{\text{min}} = d_{iB}$ the cluster i is regarded as jet and is removed from the list. Otherwise ($d_{\text{min}} = d_{ij}$) the clusters i and j are grouped together. Then d_{min} is recalculated. This procedure is repeated until all clusters are grouped into jets. The algorithm can not only be applied to cluster, but to tracks or any other collection four-vector objects.

The jet energy is calibrated in several steps. First, the expected energy contributions from pile-up jets are removed. Then, the jet energy scale (JES) is applied, which tries to correct the reconstructed jet energy to its true value, as derived from the MC truth. This scale can be a factor 2 for jets with $E = 30\,\text{GeV}$, depending on η, and goes down to about 1.2 for $E = 1.2\,\text{TeV}$ [76]. A large contribution to this scale arises from the cluster energy being measured at the electromagnetic (EM) scale, while most of the deposited energy is of hadronic origin. Further, the *global sequential calibration* (GSC) is applied, mainly to account for a different response to gluon- and quark-initiated jets [76]. The GSC includes

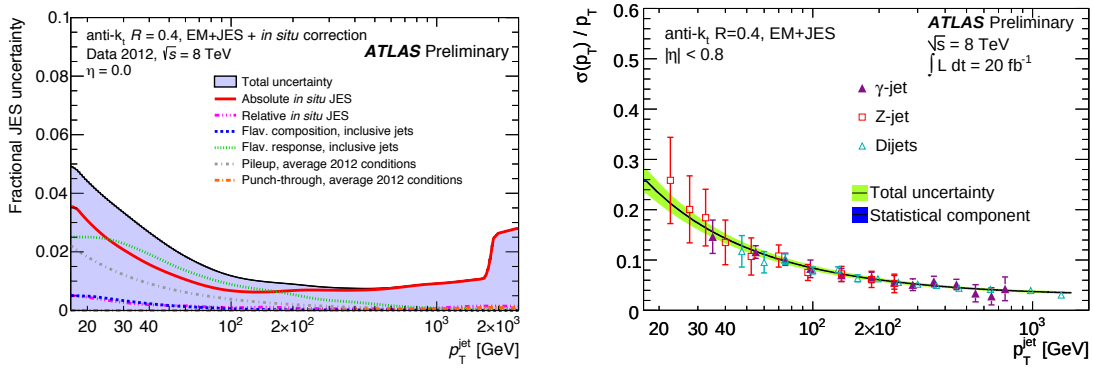


Figure 4.3.: Jet energy scale (left) and resolution uncertainties (right) for EM jets with in-situ correction [76].

a "punch-through" correction for jets with large p_{T} whose energy is not fully contained in the calorimeter.

Further, *in-situ* calibrations are applied. Jets in the forward region ($0.8 < |\eta| < 4.5$) are balanced with jets in the central region ($|\eta| < 0.8$). This "η-intercalibration" gives correction factors for the jet energy of below 2 % [76]. The central jets ($|\eta| < 0.8$) are corrected from measurements of jets recoiling against well-measured objects in specific ranges of p_{T} i.e. photons ($20 \leq p_{\mathrm{T}} \leq 200\,\mathrm{GeV}$ and Z bosons ($30 \leq p_{\mathrm{T}} \leq 800\,\mathrm{GeV}$. High p_{T} jets ($300 \leq p_{\mathrm{T}} \leq 1700$) are corrected from events, where a single jet recoils against a system of lower p_{T} jets. Finally, a correction is applied for MC samples using the Atlfast-II detector simulation [77], and a specific non-closure uncertainty is assigned. It is at most ±1% at 20 GeV and rapidly falls with increasing jet p_{T} [76].

Uncertainties for the various JES corrections are taken into account, as listed in Section 4.7. The total JES uncertainty is shown in Figure 4.3 (left) and belongs to the most important experimental uncertainties within this thesis. It ranges from ±4 % at $p_{\mathrm{T}} = 25\,\mathrm{GeV}$ to ±1 % at $p_{\mathrm{T}} = 500\,\mathrm{GeV}$ and is rising above.

The jet energy resolution (JER) varies between $20\,\% \pm 3\,\%$ at $p_{\mathrm{T}} = 25\,\mathrm{GeV}$ to $6\,\% \pm 1\,\%$ at $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ [76]. It can be parameterized as

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{N}{p_{\mathrm{T}}} \oplus \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus C, \tag{4.7}$$

where N is caused by the noise from electronics and pile-up, S is stochastic noise from the sampling calorimeters and C is a p_{T} independent term. A fit of this function to the measurement is shown in Figure 4.3 (right).

Jets are selected for the analyses in two categories: *signal* and *forward*. Signal jets are required to satisfy $|\eta| < 2.5$, $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\mathrm{JVF}| > 0.5$. Forward jets are selected within $2.5 < |\eta| < 4.5$ and $p_{\mathrm{T}} > 30\,\mathrm{GeV}$. Here, JVF is the *jet vertex fraction*, defined as

$$\mathrm{JVF}(jet) = \frac{\sum_i p_{\mathrm{T}}^i}{\sum_j p_{\mathrm{T}}^j}, \tag{4.8}$$

where i (j) runs over all tracks matched to the jet and the signal (any) vertex. Hence, the

JVF is used to reject jets not originating from the signal vertex, but from pile-up. The corresponding uncertainty is evaluated by varying the cut on JVF by 0.03.

4.4. Flavor tagging

The identification of b jets, jets initiated by b quarks, plays a central role for this thesis. So-called b-tagging algorithms are employed to distinct b jets from c jets (initiated by c quarks) or light jets (initiated by gluons, u, d or s quarks). They use the fact that b quarks form hadrons with relatively long lifetimes and large masses, and the hard fragmentation of the b quark.

The life time of b hadrons are typically around 1.5 ps. This is long enough for the the hadrons to travel several mm from the primary interaction before they decay and create secondary vertices. These can be reconstructed using the tracks from the inner detector, as described in Section 4.1. The pixel detector with its innermost "b-layer", and for Run 2 the additional insertable b-layer, deliver the highest precision for tracks close to the interaction point.

Tagging algorithms The b-tagging algorithms rely either on reconstruction of secondary vertices or on the impact parameters of the tracks. The impact parameter d_0 is the minimal distance of a track to the signal vertex in the transverse plane and z_0 in the longitudinal direction. A track originating from a secondary vertex has predominantly large impact parameters. Commonly used is the impact parameter significance, $S_i = \{d_0/\sigma_{d_0}, z_0/\sigma_{z_0}\}$, where the impact parameters are divided by their uncertainties. The *IP3D* algorithm uses a two dimensional log-likelihood ratio,

$$W_{\mathrm{IP3D}} = \sum_{i=1}^{N_{tr}} \ln \frac{b(S_i)}{u(S_i)}, \tag{4.9}$$

with $u(S_i)$ and $b(S_i)$ being the probability density functions (PDFs) for light and b jets, respectively.

The *SV1* algorithm tries to reconstruct one secondary vertex for each jet, for which it combines all of its tracks with impact parameters above a certain threshold. Several of the vertex parameters are combined in a likelihood ratio, such as the its invariant mass and the transverse momentum fraction of its tracks over all tracks in the jet. A more advanced algorithm, *JetFitter*, is trying to fit tracks into secondary vertices using the decay topologies of b and c hadrons in the jet. It is not relying on a single geometrical vertex and even secondary vertices with one track can be reconstructed [78].

Several tagging algorithms can be combined to achieve better performance. *JetFitterCombNN* uses a neural network to combine IP3D and JetFitter. *MV1* finally combines IP3D, SV1 and JetFitterCombNN in another neural network [79]. The output distribution of this algorithm is shown in Figure 4.4 (left). While the distribution is concentrated at values close to zero for light jets, it is mostly close to one for b jets. The output for c jets is distributed in between. The performance of the various algorithms is compared in Figure 4.4 (right). Clearly, the MV1 algorithm has the largest light-jet rejection over the

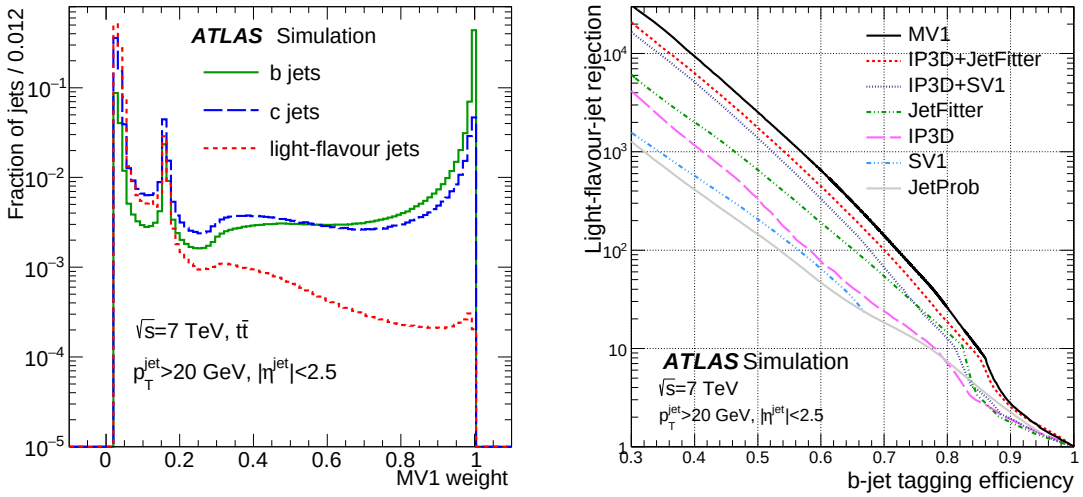


Figure 4.4.: Left: distribution obtained from the MV1 algorithm, for b, c and light jets, evaluated on a sample of simulated $t\bar{t}$ events. Right: light-jet rejection versus b-jet tagging efficiency, for various tagging algorithms [80].

full range of the b jet tagging efficiency. Another variant of the MV1 algorithm, *MV1c*, which is used here, is specifically optimized to reject c jets.

Another algorithm, *MV2c20*, was introduced for Run 2 [81]. It uses *Boosted Decision Trees* (BDTs) instead of a neural network and exploits the measurements from the newly installed IBL (Section 3.2.1) efficiently. The algorithm is trained with 20 % of c jets in the background sample, hence the "20" in its name. The expected rejection factors for light and c jets as a function of the b-tagging efficiency are compared between Run 1 and Run 2 in Figure 4.5. For a typical working point with 70 % b-tagging efficiency, the rejection is increased by a factor of 4 (1.7) for light (c) jets, which can mostly be attributed to the use of the IBL.

Truth labeling The jets in simulated events need to be labeled by their *truth flavor* for the training of the b-tagging algorithms and for the b-tagging calibration. This is accomplished using *truth matching*: if a b quark (b hadron for Run 2) with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ is found in the MC truth record, which is within $\Delta R < 0.3$ of a jet, the jet is labeled as b jet. Otherwise, if a c quark (c hadron) is found, it is labeled as c jet and otherwise, if a τ is found, it is labeled as τ jet. If none of the above applies, the jet is labeled as light jet.

Calibration Several methods exist to calibrate the b-tagging efficiency in simulation with respect to the data. Jets containing muons can be used to obtain a b-jet enriched sample. Other methods rely on the large cross section for $t\bar{t}$ production at the LHC. For example, a clean sample of $t\bar{t}$ can be selected by requiring at least four jets in the event and two lepton with opposite charges and different flavors. The number of b tags counted in the simulation after this selection is compared to the data in Figure 4.6 (left).

Scale factors are derived from a fit to the data to correct the observed discrepancy

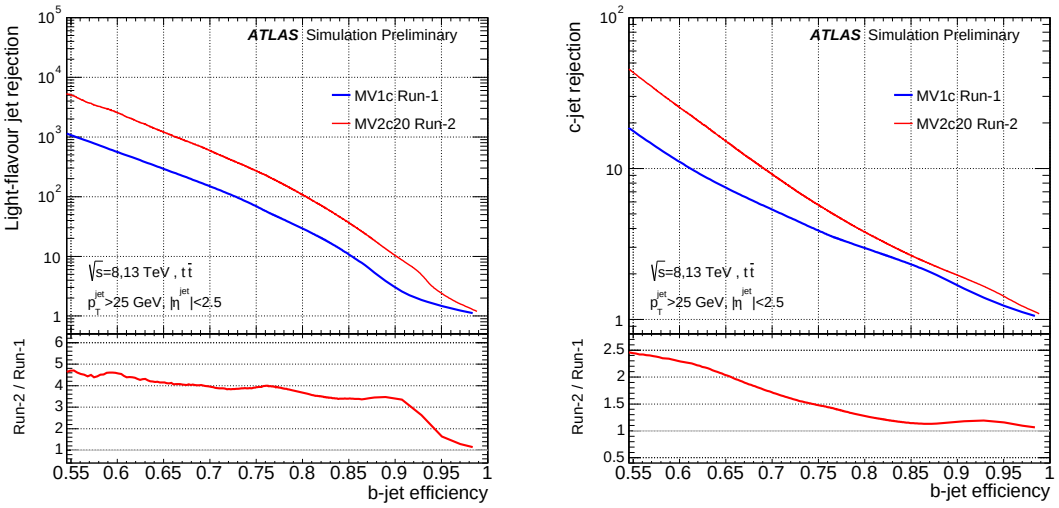


Figure 4.5.: Expected rejection factor of light jets (left) and c jets (right) as a function of the b-tagging efficiency, compared between the MV1c algorithm in Run 1 (blue) and MV2c20 in Run 2 (red). The efficiencies are evaluated on a sample of simulated $t\bar{t}$ events [81].

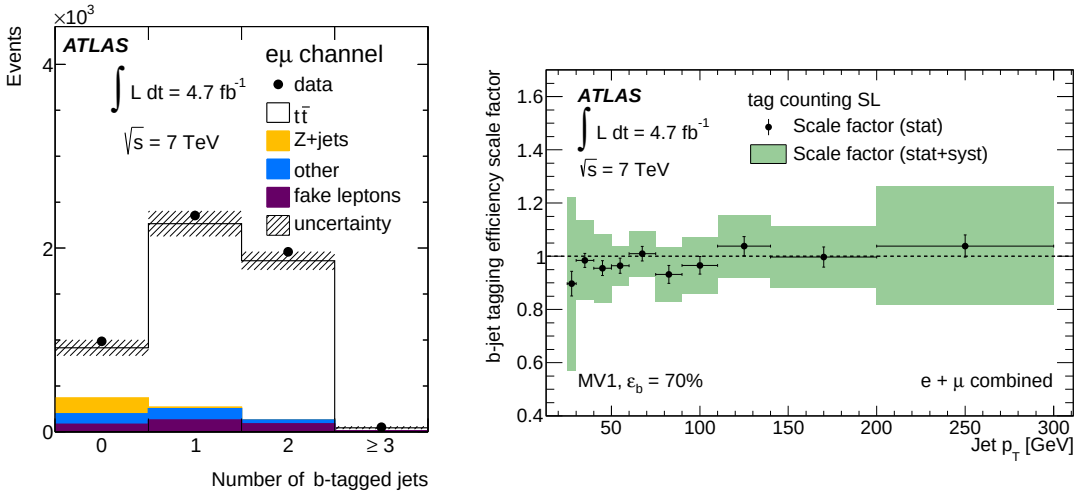


Figure 4.6.: Number of b tags, as measured in a $t\bar{t}$ enriched sample (left) and the corresponding b-tagging efficiency scale-factors obtained from a fit to the data (right) [80].

in the b-tagging efficiency. This is done for every b-tagging algorithm at various *working points*, each corresponding to a cut on the output of the algorithm with a specific b-tagging efficiency. Further, this is done separately for the different jet flavors and as a function of the jet p_{T} (and jet η for light jets). An example is shown for b jets with the MV1 algorithm and an b-tagging efficiency of 70 % in Figure 4.6 (right). The corresponding systematic uncertainties are taken into account in the analyses presented in this thesis.

Pseudo-continuous calibration The calibration described above is only valid for a specific cut on the b-tagging output. However, this might not always be the most efficient use of the algorithm. By using the full distribution, e.g. as input to a MVA, often a better separation of a signal from the backgrounds could be achieved. This would require the calibration of the whole output distribution of the b tagger, which is not feasible. Instead, a *pseudo-continuous* calibration can be employed: the b-tagging output is calibrated in coarse bins.

This technique is used in the SM $VH(\rightarrow b\bar{b})$ analysis (Section 5.3.2), for which a calibration of the MV1c tagger in six bins was made available, for which the bin boundaries correspond to b-tagging efficiencies of 100 %, 80 %, 70 %, 60 %, 50 % and 0 %.

The number of systematic uncertainties resulting from this method is very large: number of tagging bins $\times$ jet flavors $\times$ jet-p_{T} bins ($\times$ jet-η bins). However, strong correlations among these uncertainties exist. An example is shown in Appendix A.2. The correlations are exploited in an eigenvector (EV) decomposition, one for each jet flavor. A large number of eigenvectors have very small effect (small eigenvalues) and can be neglected. The remaining numbers of uncertainties are 10 for b jets, 15 for c jets and 10 for light jets.

Truth tagging The b-tagging algorithms are usually used in analyses by applying a cut on the b-tagging output distribution. However, the number of events for some simulated samples might be very small after the cut, e.g. for samples containing only light jets. A different method can be applied to increase the number of events after b tagging for such samples. Instead of applying a cut, the events are weighted by the probability of passing the b-tag cut.

The b-tagging probability is obtained from efficiency maps, which contain the b-tagging efficiency, $\epsilon(j)$, for single jets, j. They measured for a specific tagging algorithm, for each jet flavor and in bins of p_{T} and η of the jet. The tagging probabilities of the jets are combined to obtain the event weight. For example, if the first jet in the event is requested to be b tagged and a b-tag veto is applied on a second jet in the event ("1-tag region"), the corresponding event weight would be $w = \epsilon(j_1) \cdot (1 - \epsilon(j_2))$. This *truth-tagging* method is applied in the SM $VH(\rightarrow b\bar{b})$ analysis (Section 5.2).

4.5. Overlap removal

The reconstruction of physical objects, such as leptons and jets, is not exclusive. For example, an energy deposit in the calorimeter, which is reconstructed as electron, is usually reconstructed as jet as well. The *overlap removal* (OR) is applied to resolve such ambiguities. It is applied to the loose electron, muon and jet collections.

First, all loose electrons that share a track with a muon are removed. This is a rather rare case. Then the most important step is applied: jets that are within $\Delta R < 0.2$ of an electron are removed. This is the case for most of the electrons, since their energy deposits in the calorimeters are used as input to the jet reconstruction, regardless of the outcome of the electron reconstruction. Without this step, almost each electron would be double-counted as jet in the analysis. Further, electrons are removed if they are within $\Delta R < 0.4$ of a jet to remove any remaining ambiguity.

Jets and muons are defined as overlapping if they are within $\Delta R < 0.4$ of each other. The jet is then removed if it has ≤ 2 tracks. In this case it is likely to be a real prompt muon, which has caused - or accidentally coincides - with some energy deposit in the calorimeter. Otherwise the muon is removed. The latter case is particularly important for semileptonic decays of b hadrons, where a real, but non-prompt muon can be emitted. Such muons can be used later to correct the energy of the jet, as described in Section 5.2.4.

4.6. Missing transverse energy

Neutrinos cannot be detected in the ATLAS detector, since they interact only weakly and the cross sections are extremely small. However, they can be reconstructed indirectly through the transverse energy imbalance of the event. In an ideal event the total transverse energy vectors of all particles should sum up to zero, but if undetectable neutrinos are involved this is generally not the case. The absolute value of this energy is called $E_{\mathrm{T}}^{\mathrm{miss}}$, the missing transverse energy. The z component cannot be reconstructed, since the longitudinal momenta of the colliding partons are a-priori unknown.

The missing transverse energy is defined by its x and y components as

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2}, \tag{4.10}$$

where $E_{x(y)}^{\mathrm{miss}}$ are defined by the negative sum over the energies of all objects in the event:

$$E_{x(y)}^{\mathrm{miss}} = -E_{x(y)}^{\mathrm{SoftTerm}} - \sum_{obj} E_{x(y)}^{obj}. \tag{4.11}$$

Here, *obj* are the reconstructed and calibrated physical objects, *obj* = jet, e, μ, τ or γ, and *SoftTerm* are calorimeter clusters not assigned to any such object. The energy components of each object can be calculated from their energy deposits E_i and their angular position:

$$E_x^{\mathrm{obj}} = \sum_i E_i \sin\theta_i \cos\phi_i \tag{4.12}$$

$$E_y^{\mathrm{obj}} = \sum_i E_i \sin\theta_i \sin\phi_i. \tag{4.13}$$

An example distribution for the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ in simulated events, compared to the data, is shown in Figure 4.7 (left). The apparent discrepancy between the simulation and the data is covered by the systematic uncertainties discussed below. The $E_{\mathrm{T}}^{\mathrm{SoftTerm}}$ is scaled with the soft term vertex fraction (STVF) to suppress pile-up. It is defined similarly

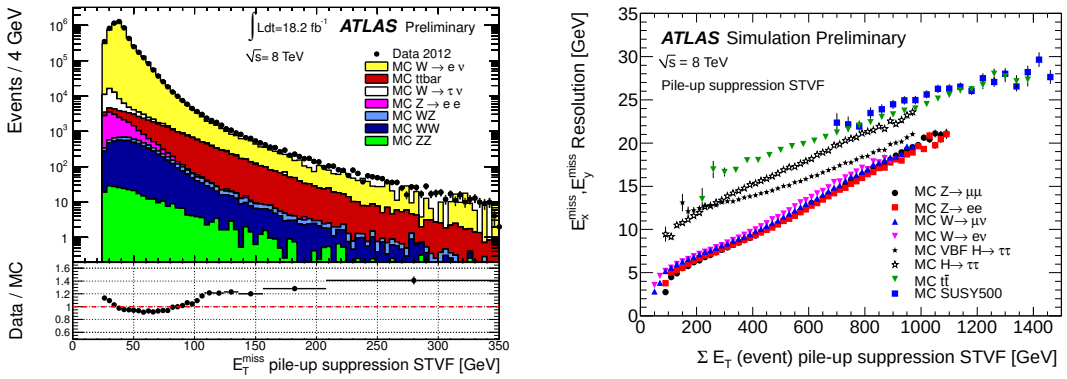


Figure 4.7.: Left: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, as measured in a $W \to e\nu$ enriched sample. The simulation is normalized to the data. Right: $E_{x(y)}^{\mathrm{miss}}$ resolution as a function of H_{T} for various simulated processes. Both quantities are shown after pile-up suppression using the STVF scaling [82].

to the JVF (Section 4.3), but the sums run over the tracks associated to the SoftTerms instead of the jets.

The scalar sum of the transverse energy in the event, H_{T}, is an important quantity to parameterize and understand the $E_{\mathrm{T}}^{\mathrm{miss}}$ performance. Within this thesis, it is defined as

$$H_{\mathrm{T}} = E_{\mathrm{T}}^{\mathrm{SoftTerm}} + \sum_{obj} E_T^{obj}. \tag{4.14}$$

The resolution of $E_{x(y)}^{\mathrm{miss}}$ in simulated $W \to e\nu$ events is estimated be about 7 GeV for H_{T} = 200 GeV and about 20 GeV for H_{T} = 1 TeV, as shown in Figure 4.7 (right).

All energy scale and resolution uncertainties of the reconstructed objects are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$. E.g. a variation of the jet energy affects the jet term of Equation 4.11 and therefore the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ as well. Additionally, scale and resolution uncertainties arising from the SoftTerms are taken into account and are listed in the next section.

The vectorial missing transverse momentum, $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$, and its magnitude, $p_{\mathrm{T}}^{\mathrm{miss}}$, can be defined in a similar fashion as $E_{\mathrm{T}}^{\mathrm{miss}}$, but using tracks:

$$\boldsymbol{p}_{x,y}^{\mathrm{miss}} = -\sum_{trk} \boldsymbol{p}_{x,y}^{trk}. \tag{4.15}$$

Here, the tracks, trk, are categorized as soft, jet, electron and muon tracks. The former two are provided by the specific electron or muon reconstruction algorithms, while this is not done for tracks associated to jets. Soft tracks are not associated to any of the former objects.

4.7. List of experimental uncertainties

The full list of experimental uncertainties, which are taken into account in the analyses performed with the data taken at $\sqrt{s}$ = 8 TeV, is given in Table 4.1. The uncertainties are related to the reconstruction of physical objects, as described above. A name is

assigned to each uncertainty for its identification in the statistical treatment, as described in Section 5.4.

Most of the experimental uncertainties are kinematic variations, such they affect the four-vectors of reconstructed objects. They are propagated by performing an analysis (applying cuts, evaluating multivariate algorithms, ...) once for each variation, for obtaining systematically varied histograms. Hence their evaluation is computing intense. Some variations, such as b-tagging uncertainties, affect only the event weights. They can be applied in the very last step of an analysis, before filling histograms, and are computationally inexpensive.

Name	Source	Count
Luminosity (2)		
Lumi	Total integrated luminosity	1
MuScale	Pile-up profile	1
Leptons (8)		
ElecEffic	Trigger, reco., and ID efficiencies	1
ElecE	Energy scale	1
ElecEResol	Energy resolution	1
MuonEffic	Trigger, reco., and ID efficiencies	1
MuonEResolID	Energy resolution from inner detector	1
MuonEResolMS	Energy resolution from muon system	1
LepIso	Isolation scale factors	1
LepVeto	Lepton veto efficiency in 0-lepton selection	1
Jet Energy Scale (23)		
JetNPx	EV decomposition of in-situ calib. (x = 1-6)	6
JetEtaModel	η inter-calibration model	1
JetEtaStat	Statistical error of η inter-calibration	1
JetNonClos	Calibration non-closure	1
JetMu	Average pile-up correction	1
JetNPV	N_{PV} correction	1
JetPile(Pt/Rho)	Pile-up in jet area correction	2
JetFlavB†	Energy scale for b jets	1
JetBE†	Energy scale for semileptonic decays	1
JetFlavComp_X‡	Knowledge of light quark vs. gluon fraction	4
JetFlavResp_X‡	Response to light quarks vs. gluon jets	4
Jet Energy Resolution (2)		
JetEResol	Energy resolution of all jets	1
BJetReso†	Energy resolution of b jets	1
Jet Quality (1)		
JetJVF	Jet vertex fraction efficiency	1
$E_{\mathrm{T}}^{\mathrm{miss}}$ (2)		
METResoSoftTerms	Resolution of soft component	1
METScaleSoftTerms	Scale of soft component	1
$E_{\mathrm{T}}^{\mathrm{miss}}$ Trigger (2)		
METTrigZ	W efficiency curve versus Z curve	1
METTrigStat	Statistical uncertainty of efficiency curve fit	1
Flavor Tagging (40)		
BTagBxEffic	EV decomp. of b-jet tagging efficiency (x = 0-9)	10
BTagCxEffic	EV decomp. c-jet tagging efficiency (x = 0-14)	15
BTagLxEffic	EV decomp. light-jet tagging efficiency (x = 0-9)	10
BTagTruthTagDR	Correction to ΔR(cc) bias from truth-tagging	1
BTag(B/C)Sherpa	Sherpa specific tagging efficiency	2
BTag(B/C)Pythia8	Pythia8 specific tagging efficiency	2
Total	80 with priors, 0 floating	

Table 4.1.: Names and sources of experimental systematic uncertainties. The last column states the number of parameters in combined fit, as described in Section 5.4. † (‡) Applied only to (non) truth-matched b jets.

5. Search for the VH(bb) process

The search for the decay of the SM Higgs boson into a b-quark pair, $H \to b\bar{b}$, is presented in this chapter. It is performed in the associated production with a vector boson, $VH(\to b\bar{b})$ ($V = W$ or Z boson), which is decaying leptonically. The analysis is published as Ref. [1] and yields the most sensitive search for this process so far[1].

The analysis is carried out with the data recorded in 2012 with the ATLAS detector at a center-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ and corresponding to an integrated luminosity of $\mathsf{L} = 20.3\,\text{fb}^{-1}$. A combination with the corresponding analysis using the data recorded in 2011 with $\sqrt{s} = 7\,\text{TeV}$ and $\mathsf{L} = 4.7\,\text{fb}^{-1}$ [83] is performed. The result contributes to a measurement of the Higgs boson couplings using seven of its decay modes, published by the ATLAS collaboration [84], and to a combined measurement by the ATLAS and CMS collaborations [42].

The signal and background processes are discussed in Section 5.1 and the event selection in Section 5.2. The multivariate analysis (Section 5.3) is an important topic for this thesis. Further, the statistical treatment (Section 5.4) and the systematic uncertainties (Section 5.5) are discussed. Particular emphasis is put on the strategies for defining optimal bin boundaries in the histograms for the fit to the data (Section 5.6) and on the fit model validation (Section 5.7). The results are presented in Section 5.8.

5.1. Signal and background processes

The signal and background processes are simulated using various Monte Carlo (MC) generators. Their direct output, containing the generated particles and denoted as *truth* events in the following, are propagated through the ATLFAST-II simulation [85]. It is based on a detailed simulation of the ATLAS detector using the GEANT4 program [86], except for the response of the calorimeters for which a parameterized simulation is used.

The signal process for this analysis, $VH(\to b\bar{b})$, is categorized by the decay of the vector boson into $ZH \to \nu\nu b\bar{b}$, $WH \to \ell\nu b\bar{b}$ and $ZH \to \ell\ell b\bar{b}$, referred to as the 0-, 1- and 2-lepton *channels* in the following. The corresponding leading order (LO) Feynman diagrams are shown in Figure 5.1. The MC generator used for quark-initiated $VH(\to b\bar{b})$ production is PYTHIA8 [87] with the CTEQ6L1 [88] parton distribution functions (PDFs). The AU2 tune [89,90] is used for the parton shower, hadronization, and multiple parton interactions. QED final-state radiation is simulated with the PHOTOS program [91].

The additional gluon-induced ZH process, as depicted in Figure 5.2, is generated at LO in QCD with the POWHEG generator [92–94]. It is used within the MiNLO approach [95]

[1] Although the gluon-fusion production of the Higgs boson has a significantly larger cross section, as shown in Figure 2.6, it is hard to measure at the LHC in the decay to a b-quark pair, which is discussed in Appendix 6.

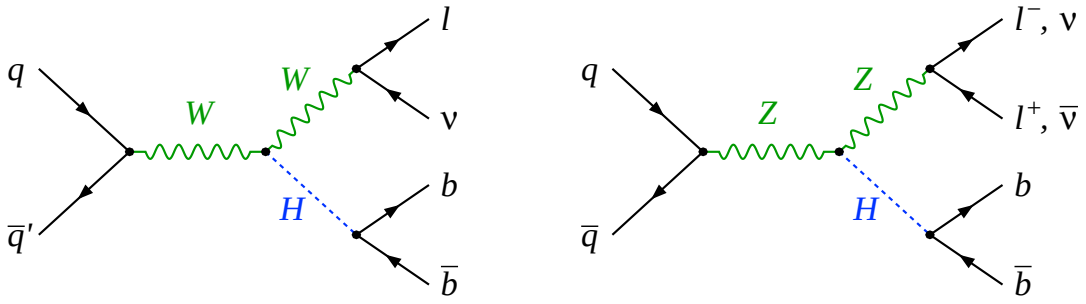


Figure 5.1.: Leading order Feynman diagrams of the $WH \to \ell\nu b\bar{b}$ (left) and $ZH \to \ell\ell b\bar{b}$ and $ZH \to \nu\nu b\bar{b}$ (right) signal processes.

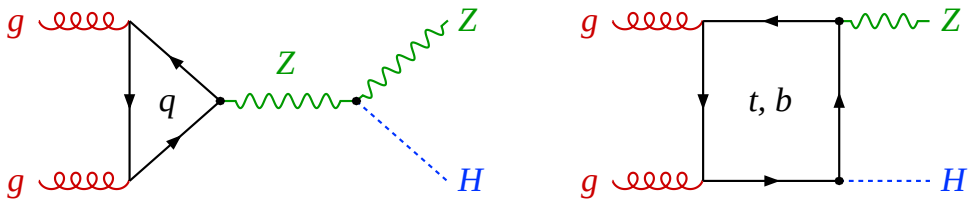


Figure 5.2.: Leading order Feynman diagrams of the gluon-induced ZH signal process. Left: triangle, right: box diagram. The Z and Higgs boson decays are omitted.

with the CT10 PDFs [96], interfaced to Pythia8 with the AU2 tune. A cross check is made with an independent calculation [97]. The same Powheg generator is used as a cross check for the quark induced $VH(\to b\bar{b})$ process and to evaluate systematic uncertainties on the signal acceptance and kinematic properties (Section 5.5.2).

Important backgrounds to the $VH(\to b\bar{b})$ signal are the production of W and Z bosons with additional jets: W+*jets* and Z+*jets*. The corresponding LO Feynman diagrams are shown in Figure 5.3. The V+*jets* samples are produced with the Sherpa 1.4.1 generator [98] with massive b/c quarks and interfaced with the CT10 PDFs.

The V+*jets* backgrounds consists of the same number of leptons and quarks in the final state as the signal. However, the additional jets are more likely to be caused by light flavor quarks than by b quarks. Hence, this background can be reduced efficiently by applying b tagging. The specific background components with two additional b quarks, W+$b\bar{b}$ and Z+$b\bar{b}$, are *irreducible*: they have exactly the same final state particles as the signal. They can, however, be distinguished by different kinematic distributions. This is exploited by the multivariate analysis (MVA, Section 5.3).

Another important background to the $VH(\to b\bar{b})$ signal is the production of top-quark pairs: $t\bar{t}$. Almost each top quark decays into a b quark and a W boson, the latter decaying hadronically or leptonically. The LO Feynman diagrams for the semileptonic and dileptonic $t\bar{t}$ processes are shown in Figure 5.4. The dominant production mode at the LHC is the gluon fusion. The fully hadronic $t\bar{t}$ decay can be neglected in this analysis due

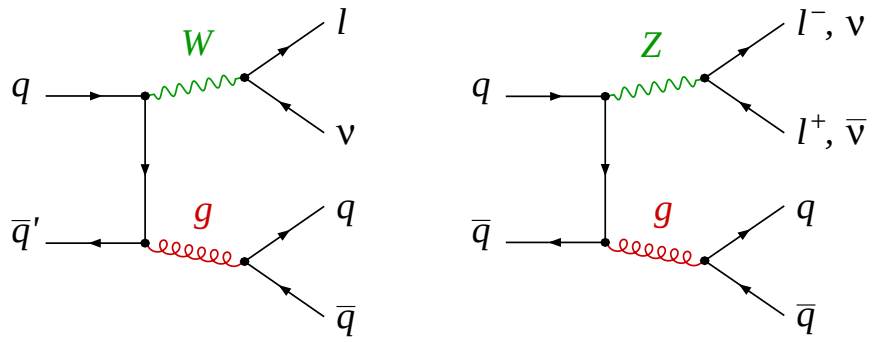


Figure 5.3.: Leading order Feynman diagrams of the W+*jets* (left) and Z+*jets* (right) background processes with at least two jets in the final state.

to the lack of leptons in the final state. The $t\bar{t}$ process is simulated with the POWHEG generator with the CT10 PDFs, interfaced with PYTHIA6 [99]. The CTEQ6L1 PDFs and the Perugia2011C tune [89, 90] are used.

The $t\bar{t}$ decays can produce the same particles in the final state as the signal processes. However, this background is reducible, as it always has additional leptons or quarks. Vetoes on low-p_{T} leptons and jets are employed to reduce this background before applying the multivariate analysis.

The production of single top quarks yields another background. The s-channel exchange process and Wt production are simulated with POWHEG, as for $t\bar{t}$. The t-channel process is simulated with the ACERMC generator [100] interfaced with PYTHIA6, using the CTEQ6L1 PDFs and the Perugia2011C tune. The corresponding LO Feynman diagrams are shown in Figure 5.5.

Further, the production of vector-boson pairs, WW, WZ and ZZ, often called *diboson* background, is taken into account. It is generated using POWHEG with the CT10 PDFs, interfaced to PYTHIA8 with the AU2 tune [101]. The corresponding LO Feynman diagrams are shown in Figure 5.6.

Finally, the multijet production from QCD processes is considered as a background. Due to the large production cross section and low selection efficiency, this process cannot be generated by MC efficiently. Instead, it is estimated from the data: in the 1- and 2-lepton channels a multijet template is selected by inverting the lepton track isolation cuts. For the 0-lepton channel the so-called ABCD method is employed. This is explained in more detail in Section 5.2.3.

5.2. Event selection

The event selection is based on an earlier study performed by the ATLAS collaboration [102], which used a cut-based approach and the dijet mass as final discriminant. However, the kinematic selection for the present analysis has been optimized for the multivariate analysis.

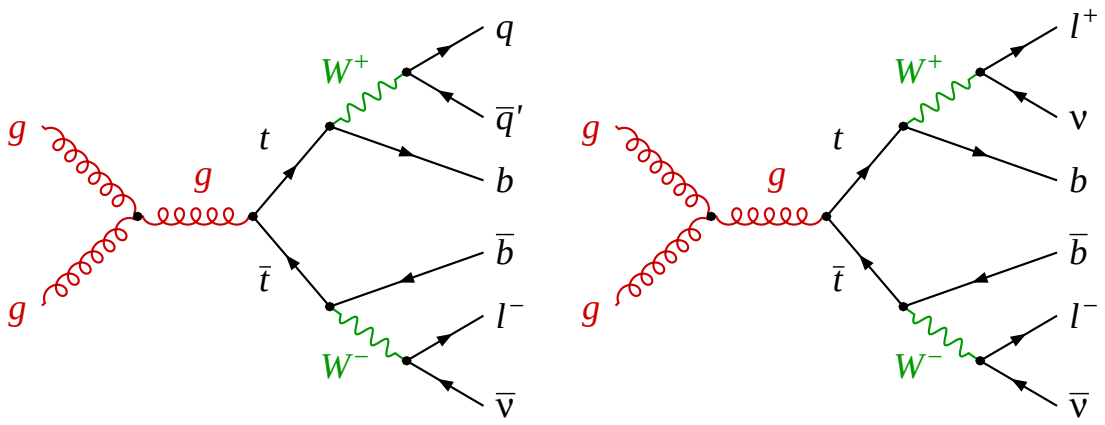


Figure 5.4.: Leading order Feynman diagrams of the semileptonic (left) and dileptonic $t\bar{t}$ (right) background processes.

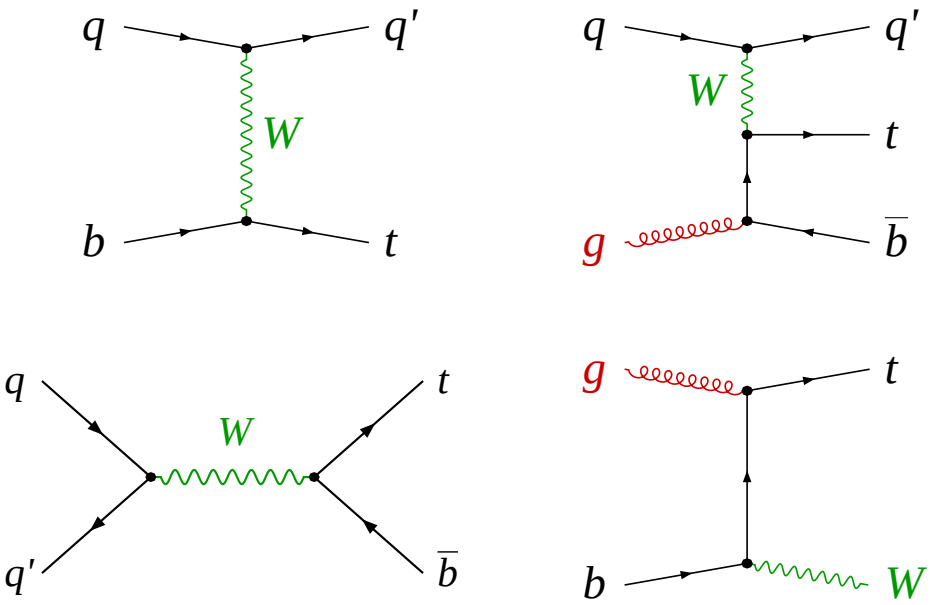


Figure 5.5.: Leading order Feynman diagrams of the single top-quark production in the t-channel (top row), s-channel (lower left) and associated Wt production (lower right). The top-quark decays are omitted.

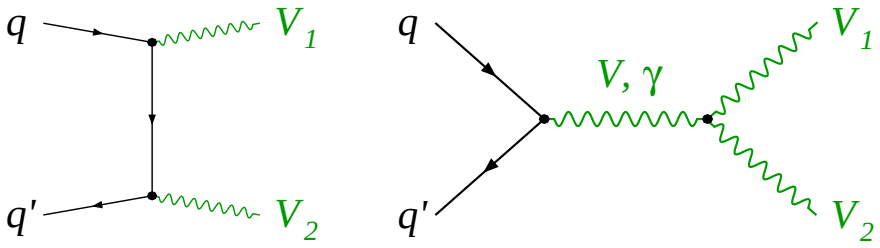


Figure 5.6.: Leading order Feynman diagrams of the diboson production in the t-channel (left) and s-channel (right). The u-channel is obtained from the t-channel by exchanging V_1 and V_2. In the s-channel only the charged couplings $WW\gamma$ and WWZ are allowed by the SM. The vector-boson decays are omitted.

5.2.1. Selection requirements

Event cleaning The first part of the event selection is devoted to remove events with deficiencies during data taking. The data are required to satisfy the *Good Run List* (GRL), ensuring all essential elements of the ATLAS detector were operational with good efficiency during data-taking. A series of standard cleaning cuts are applied to avoid sporadic problems in the detector, during reconstruction, or due to activity in the detector from non-collision background. The cuts are based on error flags provided by the ATLAS reconstruction software.

Further, each event is required to contain a signal vertex with at least three tracks, as defined in Section 4.1. Events with so-called *bad jets* are vetoed for $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 4.5$ to prevent a faulty $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. These jets correspond to energy deposits from other sources than the beam spot, e.g. cosmic ray showers, LHC beam conditions or spikes in the noise from the calorimeter electronics. Simulated events without truth record are removed.

Triggers Events in the ATLAS detector are only recorded if they activate one of the various triggers. To select as many events as possible, efficient triggers have been chosen for the three lepton channels.

The 0-lepton analysis uses a trigger based on $E_{\mathrm{T}}^{\mathrm{miss}}$ with a threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 80\,\mathrm{GeV}$. The trigger efficiency has a wide turn-on range of $50\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} < 150\,\mathrm{GeV}$, above which it reaches an efficiency close to 100 %. A similar trigger with a threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 70\,\mathrm{GeV}$ was used for the analysis carried out with the data taken at $\sqrt{s} = 7\,\mathrm{TeV}$. The corresponding trigger efficiency for the data and for simulated $W \to \mu\nu$ events is shown in Figure 5.7.

Differences in the trigger efficiency between the data and the simulated events are corrected for: the efficiencies for the data and for the MC as a function of the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ are fitted with an error function. The resulting ratio of data/MC is then applied as event weights to the simulated events.

The 1-lepton analysis uses single electron and muon triggers. One trigger for each lepton flavor is used, both with p_{T} thresholds of 24 GeV and isolation requirements, and

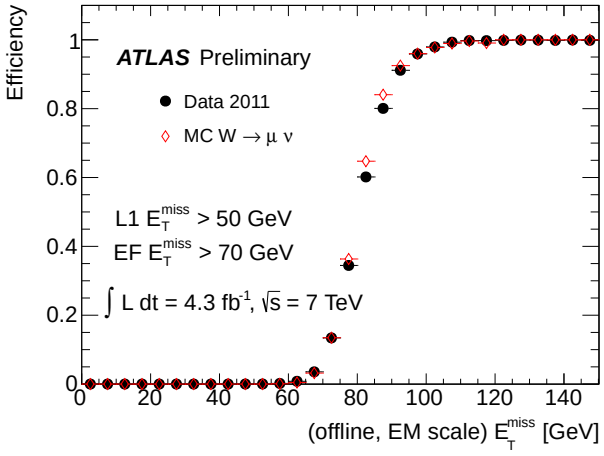


Figure 5.7.: Efficiency of an $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger with a threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 70\,\mathrm{GeV}$ as a function of the reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ for the data (black) and for simulated $W \to \mu\nu$ events (red) with $\sqrt{s} = 7\,\mathrm{TeV}$. Taken from Ref. [103].

two other triggers with higher p_{T} thresholds of 60 GeV (electrons) and 36 GeV (muons) without isolation requirements are used. In events with $p_{\mathrm{T}}^{W} > 120\,\mathrm{GeV}$ an additional $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger (same as for the 0-lepton analysis) is used to supplement inefficiencies in the muon trigger coverage for low $|\eta|$.

The 2-lepton analysis uses the single-lepton triggers as above and additional dilepton triggers with low p_{T} thresholds of 12 and 13 GeV, for muons and electrons, respectively.

Leptons Events containing any loose leptons, as defined in Section 4.2, are vetoed for the 0-lepton channel. The 1-lepton selection requires one tight lepton (electron or muon), while one medium and one loose lepton are required for the the 2-lepton channel. Any additional leptons are vetoed.

Jets Events with two or three signal jets, as defined in Section 4.3, are selected. The 2-jet category provides the best signal-to-background ratio, while the 3-jet category (allowing for some hard hadronic radiation in the signal process) provides additional sensitivity. Any forward jets are vetoed in order to reduce the contribution from the $t\bar{t}$ and single-top backgrounds.

Flavor tagging The analysis is split into 0-, 1- and 2-tag regions, based on the number of b tags on the two leading jets, defined by the MV1c discriminant with an b-tagging efficiency of 80 % (denoted as *loose* working point). The 0-tag region is only used for validation purposes, while the 1-tag region is used to control some backgrounds in the combined fit, as described in Section 5.4.2.

The 2-tag signal regions is further divided into loose-loose (LL), medium-medium (MM) and tight-tight (TT) b-tagging categories, defined by cuts on the MV1c discriminant at the 50 % (tight) and 70 % (medium) b-tagging efficiency points, as depicted in Figure 5.8.

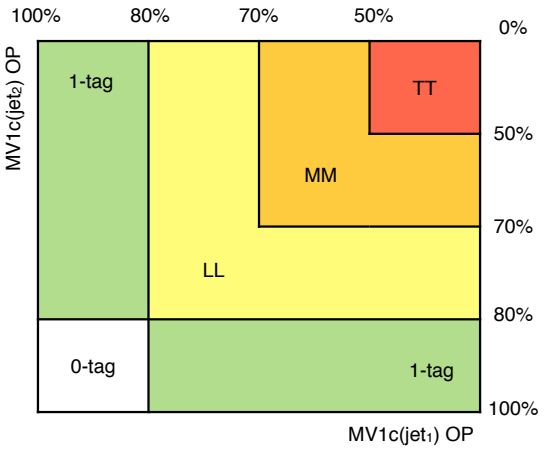


Figure 5.8.: Flavor tagging regions based on the MV1c discriminant of the two leading p_{T} jets (from Ref. [1]).

If both jets satisfy the tight requirement the event is classified as TT, else if both jets require the medium requirement as MM and otherwise as LL.

Specifically for the simulated V+jets events, the signal jets are used to categorize the V+jets background into the flavor components bb, bc, bl, cc, cl and l. In the presence of a b hadron (c hadron) with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ within $\Delta R = 0.4$ of a jet, the jet is labeled as b (c), otherwise as l for light flavor. The event is then labeled based on the flavor of the two leading jets.

The *truth-tagging* method, described in Section 4.4, is applied to the V+cc, V+cl, V+l and WW backgrounds in the 2-tag regions to obtain a reduced statistical uncertainty.

Event kinematics The analysis is split into a low- and a high-p_{T}^{V} region at a value of $p_{\mathrm{T}}^{V} = 120\,\mathrm{GeV}$. Here, V is the vector boson candidate constructed from the $E_{\mathrm{T}}^{\mathrm{miss}}$, the vectorial sum of $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_{T}^{ℓ} or the vectorial sum of $p_{\mathrm{T}}^{\ell_1}$ and $p_{\mathrm{T}}^{\ell_2}$ for the 0, 1, and 2-lepton analyses, respectively. The two leading jets are required to fulfill $\Delta R(j_1, j_2) > 0.7$ for $p_{\mathrm{T}}^{V} < 200\,\mathrm{GeV}$.

The 0-lepton channel requires $E_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$, which eliminates a large fraction of the multijet background. A number of cuts is applied to suppress it further: $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$, $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}}) < \pi/2$ and $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jets}) > 1.5$. These angular requirements are motivated by the following observations: for events with real $E_{\mathrm{T}}^{\mathrm{miss}}$ the directions of the calorimeter-based $E_{\mathrm{T}}^{\mathrm{miss}}$ and the track-based $p_{\mathrm{T}}^{\mathrm{miss}}$ are similar, hence $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})$ is small. In events with fake $E_{\mathrm{T}}^{\mathrm{miss}}$, arising from a jet energy fluctuation, the direction of $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected to be close to the direction of the poorly measured jet, hence $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jets})_{\mathrm{min}}$ is small.

Further, the scalar sum of the transverse momenta of the jets is required to be larger than 120 GeV (150 GeV) for the 2-jet (3-jet) regions. The MVA in the 0-lepton channel is only applied for the high-p_{T}^{V} region, while a specific cut-based approach is used for $100\,\mathrm{GeV} < p_{\mathrm{T}}^{V} < 120\,\mathrm{GeV}$ [1].

Process	$\sigma \times$ BR [fb]	Acceptance [%]		
		0-lepton	1-lepton	2-lepton
$q\bar{q} \to ZH \to \ell\ell b\bar{b}$	14.9	–	1.3	13.4
$gg \to ZH \to \ell\ell b\bar{b}$	1.3	–	0.9	10.5
$q\bar{q} \to WH \to \ell\nu b\bar{b}$	131.7	0.3	4.2	–
$q\bar{q} \to ZH \to \nu\nu b\bar{b}$	44.2	4.0	–	–
$gg \to ZH \to \nu\nu b\bar{b}$	3.8	5.5	–	–

Table 5.1.: Cross section times branching ratio ($\sigma \times$ BR) and acceptance for the three lepton channels and production modes. The acceptance is given for the inclusive 2-tag event selection.

For the 1-lepton channel, $M_{eff} > 180\,\text{GeV}$ is required for low-p_{T}^V and $E_{\text{T}}^{\text{miss}} > 20\,\text{GeV}$ for high-p_{T}^V. Here, M_{eff} is the scalar sum over the p_{T} of all the objects in the event, i.e. the jets, leptons and $E_{\text{T}}^{\text{miss}}$. Only muon events are used for the low-p_{T}^V region in the 1-lepton channel, which is discussed in Section 5.7.8. For the 2-lepton channel a cut is applied on the dilepton mass of $71\,\text{GeV} < m_{ll} < 121\,\text{GeV}$ to select $Z \to ll$ events.

Event categories Several categories are defined for the statistical analysis (Section 5.4) based on the selection requirements mentioned above. In summary, the events are categorized into

- 0, 1 and 2 leptons,
- 2 and 3 jets,
- 0, 1, LL, MM and TT tags and
- low and high-p_{T}^V.

5.2.2. Signal acceptance

The signal acceptance after the inclusive 2-tag event selection is shown in Table 5.1. Only muons and electrons are considered in the BR for $Z \to \ell\ell$ (due to a very low acceptance for $Z \to \tau\tau$), while all three lepton flavors are considered for $W \to \ell\nu$ (some acceptance for $W \to \tau_{\text{lep}}\nu_\tau$) and $Z \to \nu\nu$ (same acceptance for $Z \to \nu_\tau\nu_\tau$ as for the other flavors).

The acceptance is with more than 10 % significantly larger for $ZH \to \ell\ell b\bar{b}$ than for the other processes, which is due to the good lepton reconstruction efficiency and less stringent kinematic requirements. However, this channel does show the smallest cross section times branching ratio. For $ZH \to \nu\nu b\bar{b}$ and $WH \to \ell\nu b\bar{b}$ acceptances of about 4 % are observed.

Although the three lepton channels are designed to select one specific decay of the vector boson, some cross-contamination exist. For example, the lepton from the $WH \to \ell\nu b\bar{b}$ decay can be lost in an inefficient region of the detector and the event is reconstructed in the 0-lepton channel. Similarly, 2-lepton events can migrate to the 1-lepton selection. The corresponding acceptances are suppressed by a factor of about ten compared to the native selection.

Region	A	B	C	D
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})$	$< \pi/2$	$< \pi/2$	$> \pi/2$	$> \pi/2$
$\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jets})_{\mathrm{min}}$	> 1.5	< 0.4	> 1.5	< 0.4

Table 5.2.: Definition of the regions A, B, C and D. The multijet background in the signal region A is estimated using the other regions, as described in the text.

5.2.3. Multijet estimate

The multijet background is estimated from the data using different methods for the three lepton channels, as described in the following. In the course of this thesis, an estimate of the multijet background has been made for the 1-lepton channel. However, the final recipe has been developed within the ATLAS collaboration.

0-lepton The multijet (MJ) estimate in the 0-lepton channel employs the so-called ABCD method. Four regions, A, B, C and D are defined based on $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}})$ and $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jets})_{\mathrm{min}}$, as listed in Table 5.2. These two variables, motivated in Section 5.2.1, do not show significant correlations, which is necessary to apply the ABCD method. Region A is the signal region of the nominal event selection, while regions B, C and D are control regions dominated by the MJ background.

A template for the MJ background is taken from region C after subtracting the electroweak (EW) backgrounds, taken from simulation. The template is normalized by the ratio of the number of events in the regions B and D, again after subtracting the EW backgrounds:

$$N_A = \frac{N_B}{N_D} N_C. \tag{5.1}$$

No b-tagging requirement is applied in the regions B, C and D to reduce the statistical uncertainty. Instead, an additional normalization factor is applied, taken as the fraction of 2-tag events in region D. The MJ background in the signal region A is found to be about 1 % of the total background and is taken into account in the combined fit.

1-lepton The multijet background in the 1-lepton channel is determined separately for the electron and muon channels. Templates are extracted from MJ dominated regions defined by dropping the cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ adjusting the requirements on the leptons.

The medium selection criteria, as defined in Section 4.2, are applied (instead of tight) and the isolation criteria are approximately inverted. Instead of requiring the track-based isolation, i_{trk}, to be smaller than 0.04 for both lepton flavors, the cut $0.05 < i_{\mathrm{trk}} < 0.12$ is applied for electrons and $0.07 < i_{\mathrm{trk}} < 0.50$ for muons. The calorimeter-based isolation is loosened from $i_{\mathrm{calo}} < 0.04$ to $i_{\mathrm{calo}} < 0.07$.

One template for each of the regions in the nominal analysis is extracted. The small remaining EW backgrounds in the template, scaled by the results of an preliminary global fit, are subtracted from the template. The normalization of the template is then estimated from a fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the nominal analysis (without cut on $E_{\mathrm{T}}^{\mathrm{miss}}$). Two floating parameters are determined: one for the MJ background normalization and one

for the EW backgrounds. This is done separately for the 2- and 3-jet and 1- and 2-tag regions. The MJ background in the 1-lepton channel constitutes about 11 % (6 %) of the total background in the 2-jet, LL-tag (TT-tag), low-p_{T}^{W} category. It is significantly smaller for the high-p_{T}^{W} regions.

A specific procedure is applied to decrease the statistical uncertainties of the templates in the 2-tag regions. The 1-tag events, which show similar kinematic properties, are used to enrich the 2-tag regions. However, the MV1c distribution, which is used as input to the MVA, is obviously not valid for the untagged jet. It is emulated by randomly drawing values from the MV1c distribution measured in 2-tag events. The distribution is measured separately for the untagged jet being the leading or sub-leading jet in p_{T} and in bins of the MV1c value of the tagged jet.

Residual differences are observed in some distributions between these pseudo-2-tag MJ events and the actual 2-tag MJ events. A reweighing is applied as a function of the MV1c distribution of the tagged jet and, for the electron channel, as a function of $\Delta R(j_1, j_2)$ and p_{T}^{W}. This procedure is applied before estimating the normalization in the template fit mentioned above. Systematic uncertainties are assigned, as listed in Section 5.5.8.

2-lepton The MJ background estimation for the 2-lepton channels is similar to the 1-lepton channel. A template for the 2-electron channel is obtained by loosening the identification and isolation requirements. The normalization is determined from a fit to the $m_{\ell\ell}$ distribution, where the normalization of the Z+jets and MJ backgrounds are free parameters. The other backgrounds, mostly $t\bar{t}$, are fixed to the MC prediction. Consistent scale factors are found for the 0-, 1- and 2-tag regions.

A similar procedure for improving the statistical uncertainty of the 2-tag template, by reweighing 1-tag events, is applied. For the 2-muon channel the MJ background is found to be negligible. Altogether, the MJ background amounts to less than 1 % of the total background in the 2-lepton channel.

5.2.4. Dijet mass correction

The dijet mass is the most important kinematic quantity to discriminate the signal from the backgrounds. Hence, a good mass resolution is critical for the sensitivity. The resolution is measured to be 16.4 GeV for the signal process after applying the jet-energy calibration, as shown in Figure 5.9.

Part of the rather long tail toward low m_{jj} is caused by semileptonic b-hadron decays. The resulting jets can contain a muon, which looses only little energy in the calorimeter, while most of its energy is lost for the jet reconstruction. Additional energy of the jet is lost due to neutrinos. This is partially corrected by adding the four-vector of muons, which are reconstructed within the jet, to the four-vector of the jet. With this *muon-in-jet* correction the resolution is improved to 14.4 GeV.

Another small gain is observed from applying a resolution correction to each jet. It is derived from the signal MC by comparing the reconstructed p_{T} of the jet to its truth p_{T}. A scale factor is applied to the jet energy as a function of the reconstructed jet p_{T}.

A different correction is applied in the 2-lepton channel, consisting of a kinematic fit. It uses $m_{\ell\ell} \approx m_Z$ and $E_{\mathrm{T}}^{\mathrm{miss}} \approx 0$ as constraints to adjust the energy of the jets. Adding this

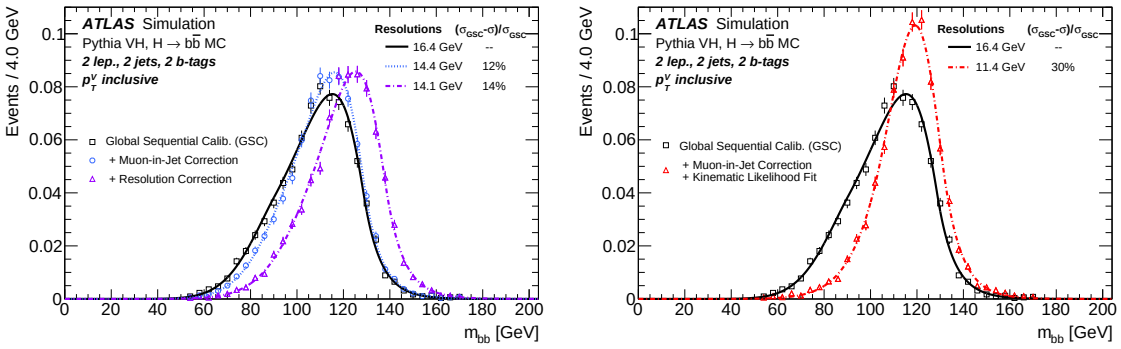


Figure 5.9.: Dijet mass resolution for signal events with various corrections. The correction on the left are used for the 0- and 1-lepton channel, while the kinematic fit on the right is used for the 2-lepton channel (from Ref. [1]).

correction on top of the muon-in-jet correction, a resolution in the dijet mass of 11.4 GeV is achieved.

5.3. Multivariate analysis

In order to maximize the sensitivity a multivariate analysis (MVA) is performed. *Boosted Decision Trees* (BDTs) are chosen among several MVA methods. They have proven to provide good performance, while being robust and easy to configure. The following choices are based on earlier studies performed by the author of this thesis in Ref. [104]. The implementation of BDTs in the TMVA toolkit [105] is used and is described in the following.

5.3.1. Boosted decision trees

A Boosted Decision Tree (BDT) is a collection of decision trees, which are combined using *boost weights*. The construction of a BDT is called *training* and is performed using simulated background and signal samples. For each event a number of variables is provided as input to the BDT.

Each decision tree is constructed by first defining the *root node*, containing the full amount of signal, S, and background events, B. A separation index is defined for each input variable, x_i, as

$$g(c_i) = p(c_i) \cdot (p(c_i) - 1), \tag{5.2}$$

where $p(c_i) = S(c_i)/B(c_i)$ is the purity of the resulting sample, if the cut $x_i < c_i$ is applied. This specific index is called *Gini index*. Other choices are possible, but usually do not affect the performance of the BDT significantly.

The index is evaluated for each variable at a number of equidistant points ("nCuts" in the TMVA configuration) and the maximum, $g(c_i^{\max})$, is determined. The variable with the largest maximum is chosen and the sample is split at $c_i^{\max}$ into a background- and a

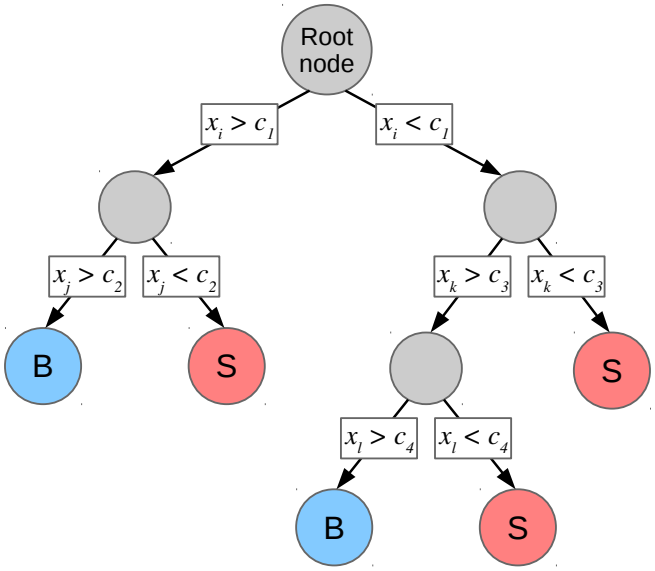


Figure 5.10.: Sketch of a decision tree with a depth of three. The root node contains the first decision, based on variable x_i. The sample is split at the cut value c_1 into a signal- and a background-like sub-sample. Each sub-sample is divided again, possibly using different variables, until each event is assigned to one of the leaf nodes.

signal-like sub-sample. This procedure is repeated recursively for each sub-sample, adding decision nodes to the tree. Each decision can be based on a different variable.

One could let such a tree grow until each sub-sample consists of only one event. However, the resulting BDT would be extremely *over-trained*, meaning it would respond to the statistical fluctuations in the training sample. This can be tested for by comparing the separation power of the tree on the training sample to an independent test sample. If the former is significantly better, the BDT is over-trained.

To reduce over-training the minimum number of events in each sub-sample can be limited ("nEventsMin"). Further, the maximum depth of the tree can be limited as well ("MaxDepth"). The depth is defined as the maximum number of decision nodes, including the root node, an event can pass in sequence. A decision tree with a depth of three is sketched in Figure 5.10. The final sub-samples are called leafs and are labeled as "B" for background-like and as "S" for signal-like.

Over-training can also be reduced by reducing the tree size after the training, which is called *pruning*. This technique is not used here ("PruneMethod" is set to "NoPruning").

The above requirements put strong limitations on the size of each tree. Such a small tree alone cannot have great separation power, but this can be solved with *boosting*. Boosting does not only increase the performance compared to a single tree, but also makes them more robust against over-training. The specific method described in the following is called *adaptive boosting* ("AdaBoost").

The boosting is done by training a first tree and then applying a *boost weight*, α, to

the misclassified events. These are the signal (background) events that end up in the background-like (signal-like) sub-samples. The weight for each tree, i, is calculated as

$$\alpha_i = \frac{1 - \text{err}_i}{\text{err}_i}, \tag{5.3}$$

where err_i is the misclassification rate of the tree. By construction is err ≤ 0.5, and therefore $\alpha \geq 1$. A second tree is trained using the reweighed events. This is repeated until a saturation in performance is reached, where the corresponding number of trees has to be chosen by the user ("NTrees").

The individual trees are combined in one output variable, y, which is calculated using the boost weights as

$$y(\boldsymbol{x}) = \frac{1}{N_{\text{trees}}} \sum_{i=1}^{N_{\text{trees}}} \ln(\alpha_i) \cdot h_i(\boldsymbol{x}), \tag{5.4}$$

where $\boldsymbol{x} = (x_1, ..., x_n)$ represents the input variables for a specific event and $h_i(\boldsymbol{x})$ is defined as $+1$ (-1) for the event ending up in an signal-like (background-like) leaf of tree i.

For the present analysis, the training of the BDTs is done by splitting all available events into two samples, A and B. First, a BDT is trained on sample A and evaluated on sample B. Another BDT is trained on sample B and evaluated on sample A. This *cross-evaluation* method ensures to avoid a possible bias from over-training, by separating the test and training events. Additionally, by training the second BDT, all available events are used in the evaluation, which benefits the statistical analysis (Section 5.4).

5.3.2. Input variables

The input variables for the MVA are chosen by a *forward selection* from a set of physical quantities. Starting from a basic set of variables the performance gain by adding any of the remaining variables is evaluated. The variable with the largest gain is chosen and the procedure is repeated until no significant gain is observed anymore.

The chosen variables are listed in Table 5.3 together with short descriptions. Here, H is the Higgs boson candidate constructed from the vectorial sum of the two leading p_{T} jets, m_{jj} is the corresponding dijet mass and m_{T}^W is the transverse mass of the W, defined as

$$m_{\text{T}}^W = \sqrt{2p_{\text{T}}^{\ell} E_{\text{T}}^{\text{miss}} (1 - \cos\Delta\phi(\ell, E_{\text{T}}^{\text{miss}}))}. \tag{5.5}$$

The most important kinematic quantities, m_{jj}, $\Delta R(j_1, j_2)$ and p_{T}^V, yield a large portion of the discrimination power. For the 3-jet region the $p_{\text{T}}^{jet_3}$ and m_{jjj} provide additional sensitivity. By using the MV1c b-tagging outputs (four calibrated working points, as described in Section 4.4) as input to the MVA another significant gain is observed. The other quantities provide additional percent-level improvements.

The distributions of the input variables are shown in Figures 5.11, 5.12 and 5.13 for the 0-, 1- and 2-lepton channels, respectively. The additional quantities defined using the third jet are shown in Figure 5.14 for all channels. The distributions are shown after the fit

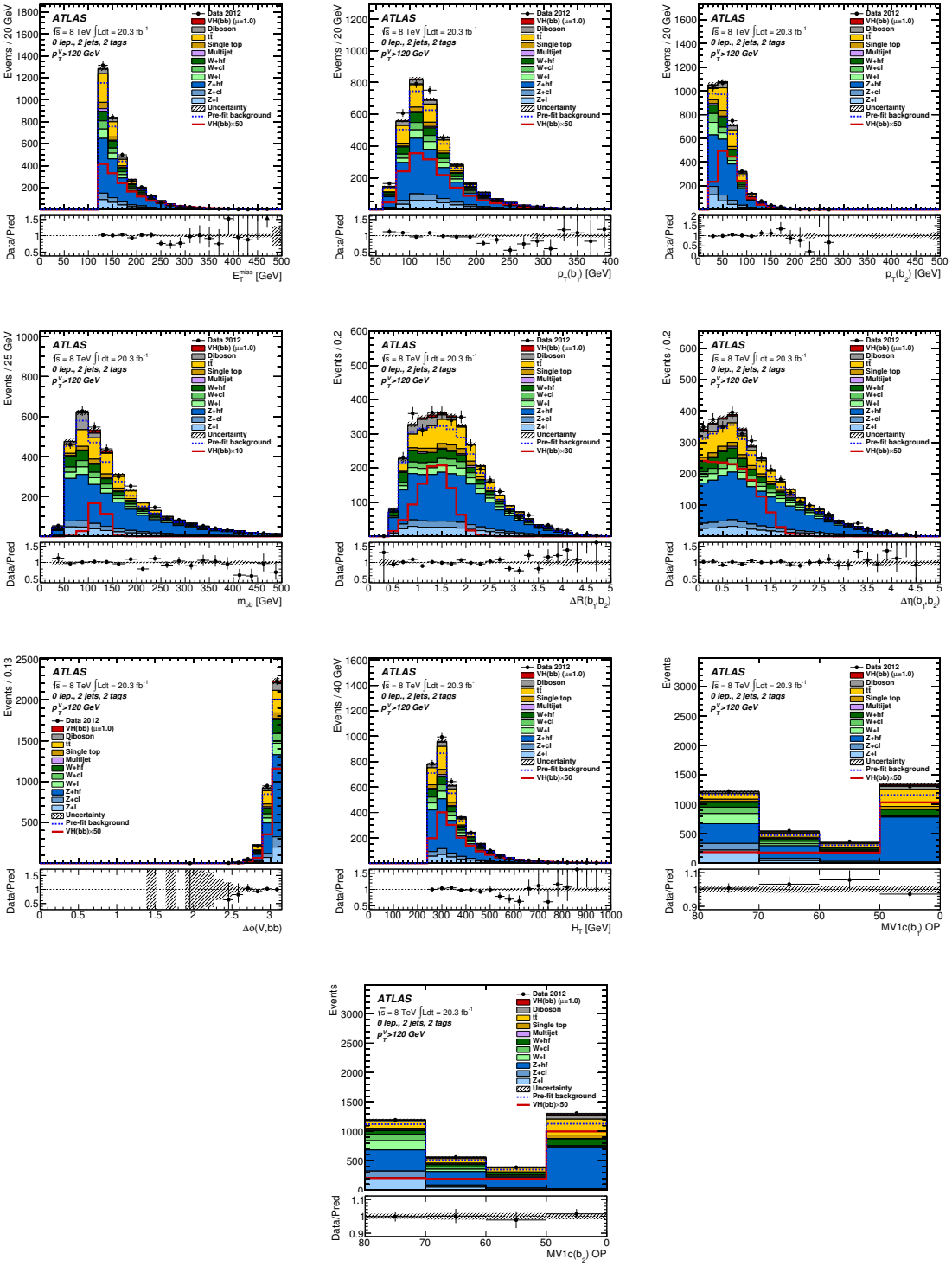


Figure 5.11.: BDT input variables for the 0-lepton channel, 2-jet, 2-tag, $p_{\mathrm{T}}^{V} > 120\,\mathrm{GeV}$. The distributions are shown after the fit to the data, as described in Section 5.4.1. A formal description of the plot elements is given in Section 5.7.1 (from Ref. [1]).

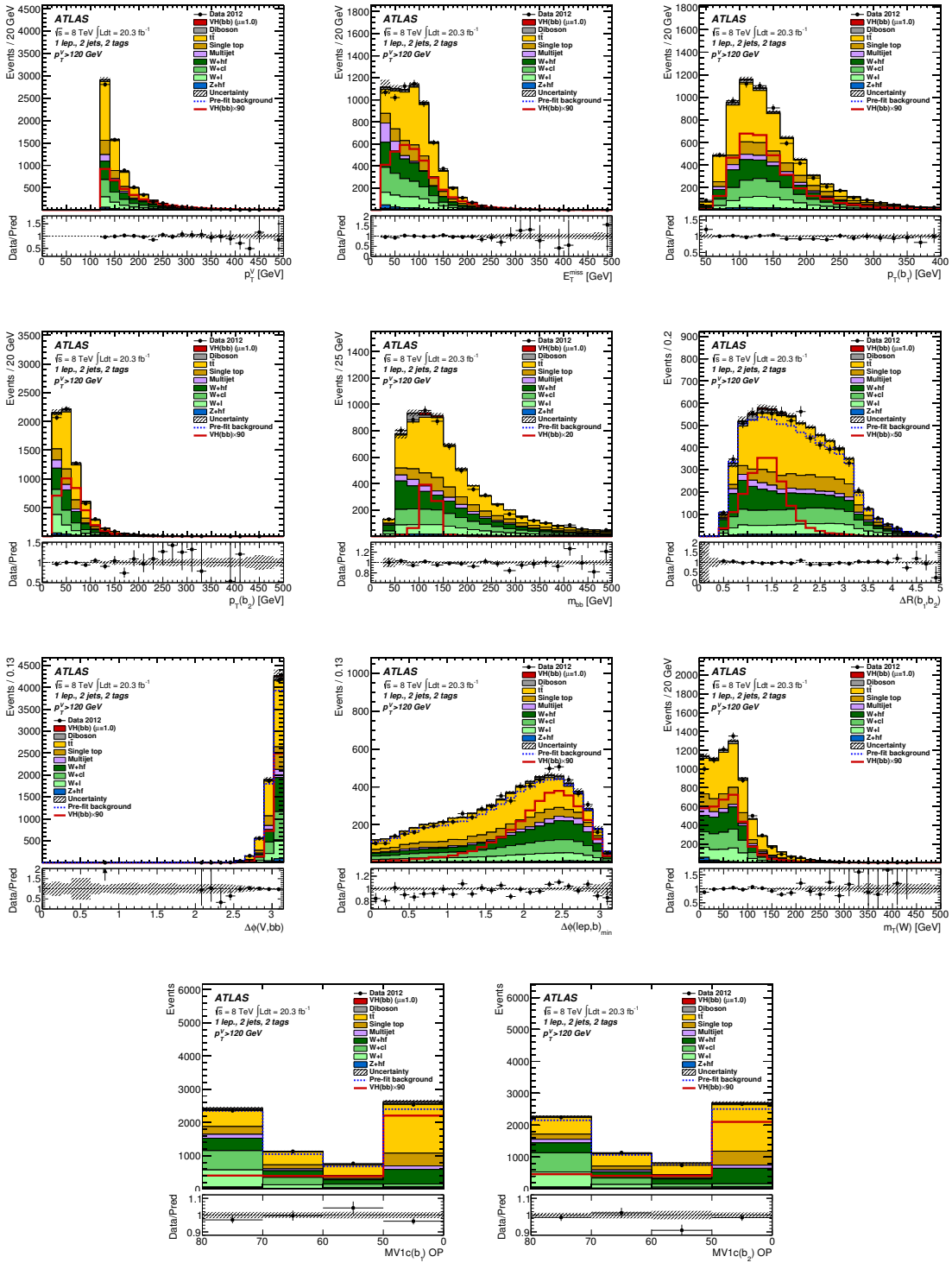


Figure 5.12.: BDT input variables for the 1-lepton channel, 2-jet, 2-tag, $p_{\mathrm{T}}^{V} > 120\,\mathrm{GeV}$. The distributions are shown after the fit to the data, as described in Section 5.4.1. A formal description of the plot elements is given in Section 5.7.1 (from Ref. [1]).

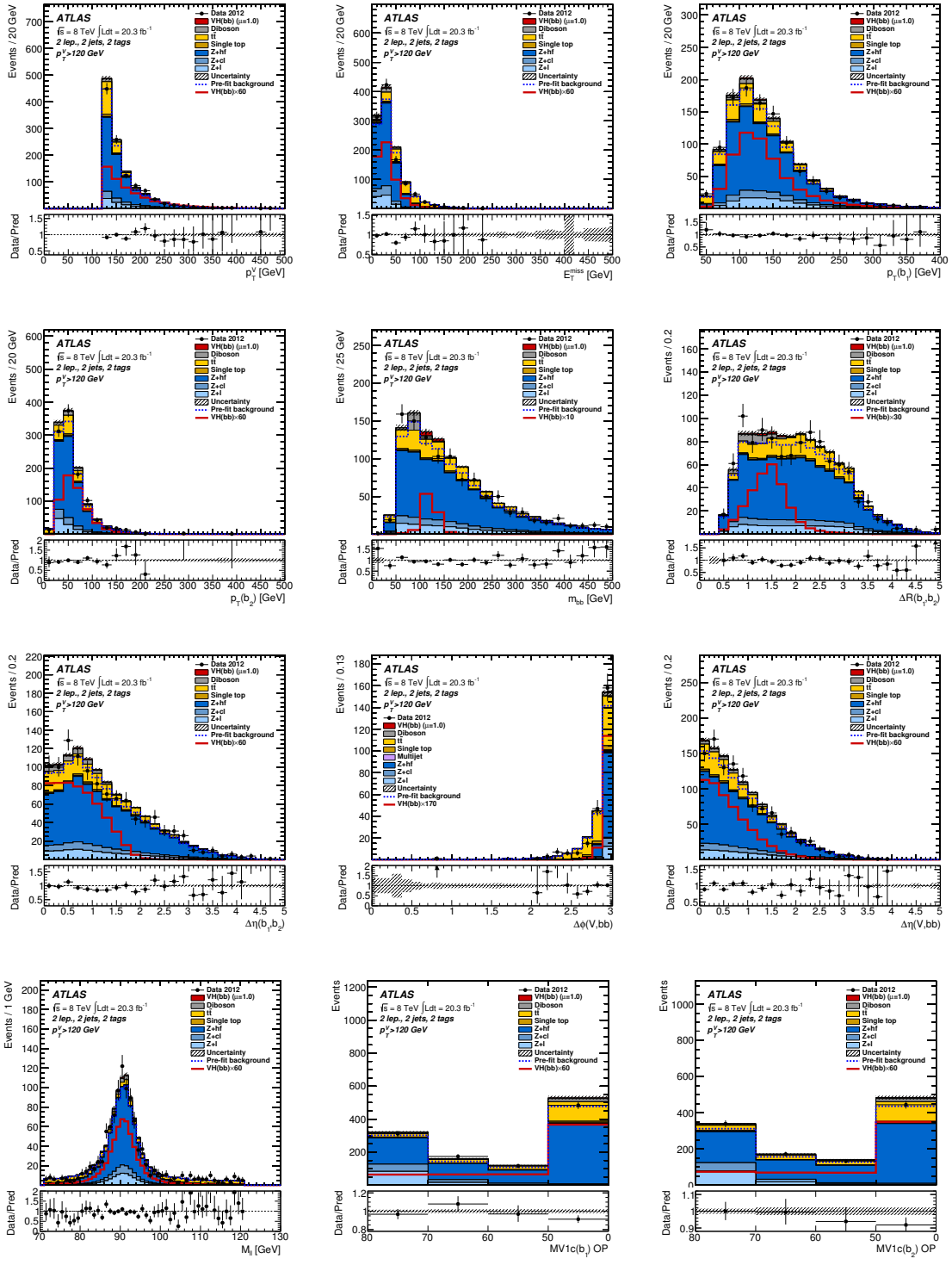


Figure 5.13.: BDT input variables for the 2-lepton channel, 2-jet, 2-tag, $p_{\mathrm{T}}^{V} > 120\,\mathrm{GeV}$. The distributions are shown after the fit to the data, as described in Section 5.4.1. A formal description of the plot elements is given in Section 5.7.1 (from Ref. [1]).

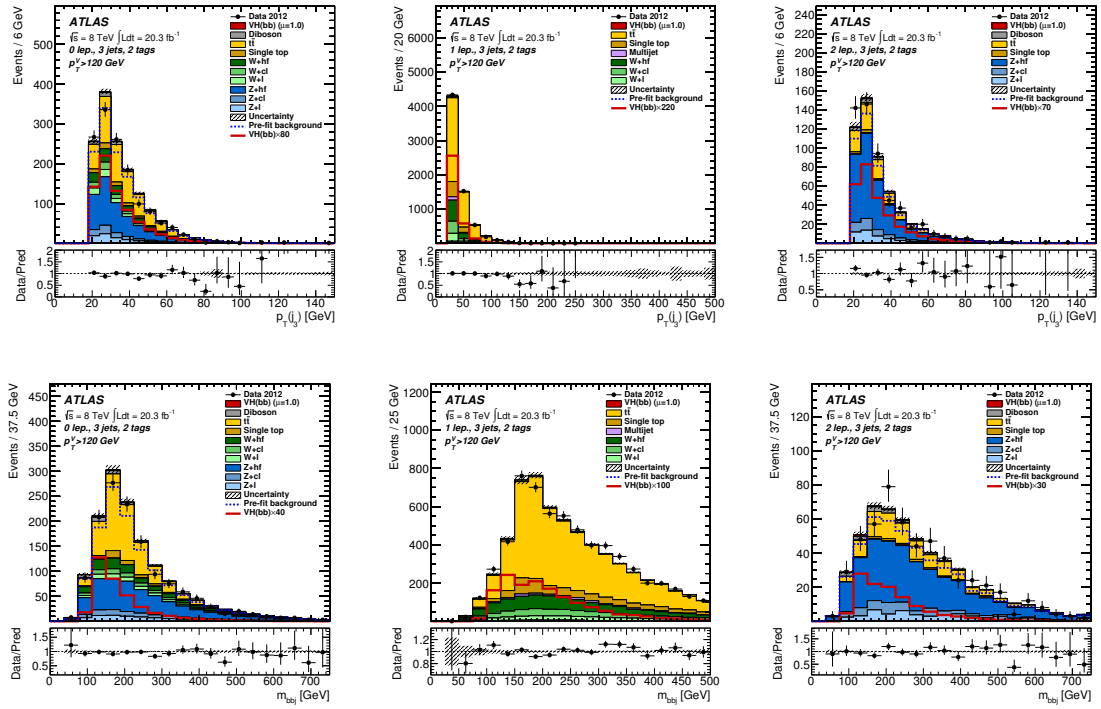


Figure 5.14.: BDT input variables for the (from left to right) 0-, 1- and 2-lepton channels and the 3-jet, 2-tag, $p_{\mathrm{T}}^{V} > 120\,\mathrm{GeV}$ regions. The distributions are shown after the fit to the data, as described in Section 5.4.1. A formal description of the plot elements is given in Section 5.7.1 (from Ref. [1]).

Variable	Name	0-lep.	1-lep.	2-lep.	Description
p_{T}^{V}	pTV		✓	✓	Vector boson p_{T}
$E_{\mathrm{T}}^{\mathrm{miss}}$	MET	✓	✓	✓	Missing transverse energy
$p_{\mathrm{T}}^{j_1}$	pTB1	✓	✓	✓	Leading jet p_{T}
$p_{\mathrm{T}}^{j_2}$	pTB2	✓	✓	✓	Subleading jet p_{T}
m_{jj}	mBB	✓	✓	✓	Dijet mass
$\Delta R(j_1, j_2)$	dRBB	✓	✓	✓	Dijet angular separation
$\lvert\Delta\eta(j_1, j_2)\rvert$	dEtaBB	✓		✓	Dijet longitudinal separation
$\lvert\Delta\phi(V, H)\rvert$	dPhiVBB	✓	✓	✓	Vector, Higgs boson azim. separation
$\lvert\Delta\eta(V, H)\rvert$	dEtaVBB			✓	Vector, Higgs boson long. separation
M_{eff}	HT	✓			Scalar sum of all p_{T}
$\lvert\Delta\phi(\ell, j)\rvert_{\mathrm{min}}$	dPhiLBmin		✓		Lepton, closest jet azim. separation
m_{T}^{W}	mTW		✓		Transverse mass of W boson
m_{ll}	mLL			✓	Dilepton mass
$MV1c(j_1)$	MV1cB1	✓	✓	✓	Leading jet MV1c value
$MV1c(j_2)$	MV1cB2	✓	✓	✓	Subleading jet MV1c value
Only for 3-jet events					
$p_{\mathrm{T}}^{j_3}$	pTJ3	✓	✓	✓	Third jet p_{T}
m_{jjj}	mBBJ	✓	✓	✓	Trijet mass

Table 5.3.: Variables used to train the multivariate discriminant for the three lepton channels.

to the data, as described in Section 5.4.1. Good agreement of the data with the adjusted prediction is observed for all variables.

The MVA not only exploits each variable alone, but also the correlations among them. Therefore, good modeling of the correlations in the simulation is required. Two-dimensional plots of each two input variables are used to investigate this. Figure 5.15 shows the example of m_{jj} versus $\Delta R(j_1, j_2)$. The two-dimensional plot, as well as the one-dimensional projections, show good agreement between the data and the simulation.

The response of the MVA can be investigated by plotting the BDT output against the input variables. An example for the dijet mass is shown in Figure 5.16. A clear behavior is observed: events with masses around the expected signal mass are given a high BDT output value, while the output is lower at both sides in the m_{jj} spectrum. Good agreement between the data and the simulation is observed here as well.

5.3.3. Configuration

The configuration of the BDTs, the choice of *meta-parameters*, is determined from an optimization, performed for the 1-lepton channel. The figure-of-merit is the *best-cut* statistical significance, which is determined by maximizing the significance as a function of a cut on the BDT-output distribution. Here, the significance is defined as $S/\sqrt{B}$, where S (B) is number of expected signal (background) events after the cut.

A first working point is chosen as the best-performing BDT from a coarse 5-dimensional scan of the number-type BDT meta-parameters. Further, one-dimensional scans are used

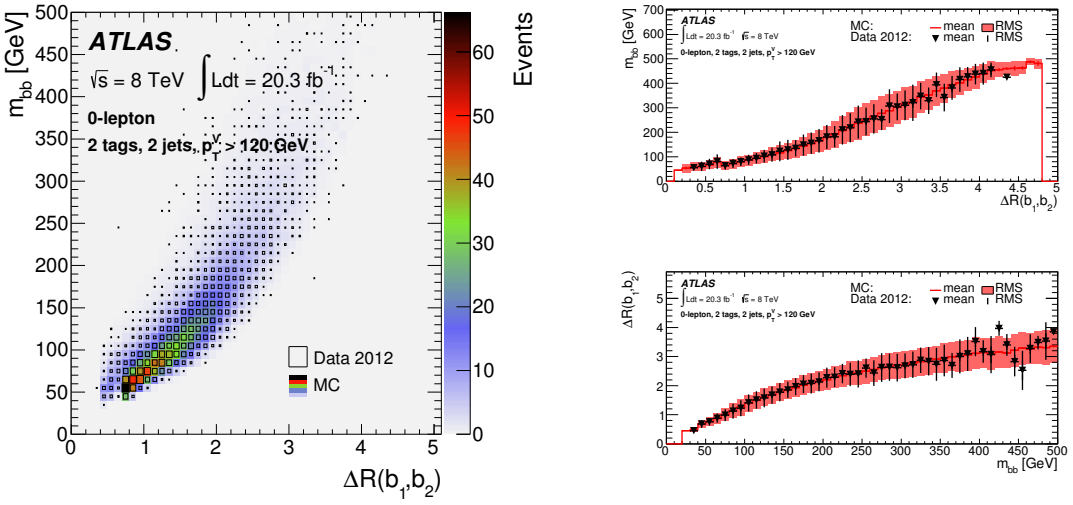


Figure 5.15.: Plot of m_{jj} versus $\Delta R(j_1, j_2)$ for the 0-lepton channel, 2-jet, 2-tag, $p_{\mathrm{T}}^{V} > 120\,\mathrm{GeV}$ (left). One-dimensional projections are shown on the right. The error bars indidicate the RMS of the projected distribution (from Ref. [1]).

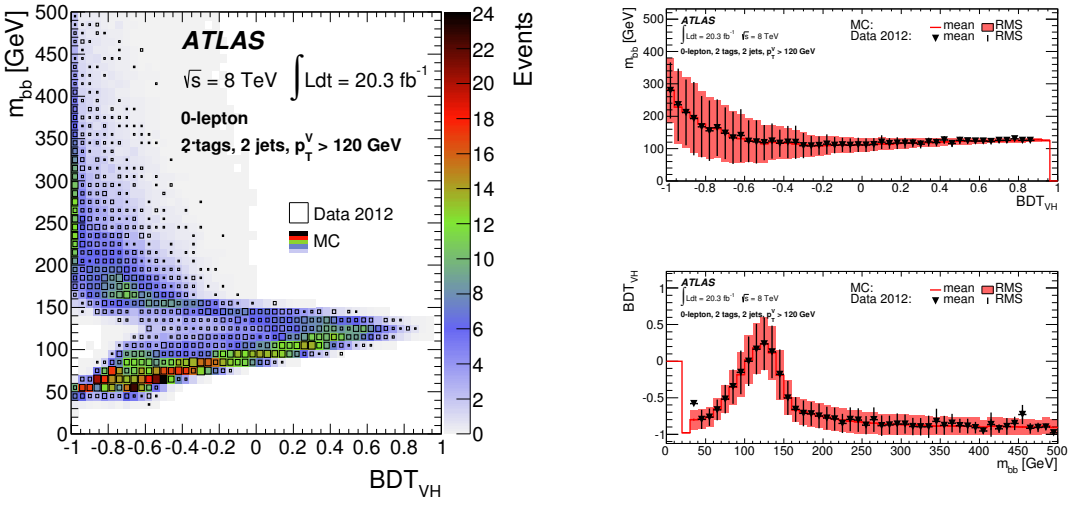


Figure 5.16.: Plot of m_{jj} versus the BDT output for the 0-lepton channel, 2-jet, 2-tag, $p_{\mathrm{T}}^{V} > 120\,\mathrm{GeV}$ (left). One-dimensional projections are shown on the right (from Ref. [1]).

TMVA Setting	Value	Description
SeparationType	GiniIndex	Node separation measure
PruneMethod	NoPruning	Pruning method
BoostType	AdaBoost	Boost procedure
AdaBoostBeta	0.15	Learning rate
NTrees	200	Number of trees
MaxDepth	4	Maximum tree depth
nCuts	100	Number of cuts tested per variable
nEventsMin	100	Minimum number of events in a node

Table 5.4.: BDT meta-parameters after optimization.

to refine each parameter setting, while the other parameters are fixed. These scans show only small slopes in the significance around the initial working point, such that one iteration of these scans is deemed to be sufficient. The result of the optimization is shown in Table 5.4. They were found to be optimal for the 0- and 2-lepton channels as well.

5.3.4. Signal sample size

The performance of the BDTs as a function of the number of signal and background events has been evaluated in the course of this thesis. A rather strong dependence on the number of training events is found. In particular, the number of signal events is critical, as many more events for the backgrounds are available.

The cross-evaluation method, as described in Section 5.3, can be generalized to *k-folds*: the events are split into k samples. In total, k BDTs are constructed, each trained on a different set of $k-1$ sub-samples and evaluated on the remaining one. This way, each BDT is trained on a larger fraction ($(k-1)/k$) of the total events than before ($1/2$).

Although each BDT is tested only on a fraction of $1/k$ of the total events, the output distributions can be combined, such that all events are used in the evaluation. This is not done for the following performance study. Instead, the best-cut statistical significance, as defined in the previous section, is evaluated for each of the k BDTs and the average is calculated. The corresponding standard deviation allows to estimate the statistical uncertainty arising from the training and test events.

The signal sample for the 1-lepton channel, 2-jets, 2-tags and $p_{\mathrm{T}}^{V} > 120\,\mathrm{GeV}$ provided originally about 7500 events, resulting in 3750 events for the training of each BDT. By varying the number of events in the training (randomly removing events) the impact on the BDT performance is assessed. The results are shown in Figure 5.17 (left). A 5-fold is used to increase the number of training events to 6000. A clear upwards trend is observed in the performance beyond the available number of events.

Following this study the number of signal events were increased by a factor of ten. The re-evaluation of the performance (Figure 5.17, right) shows a saturation above 20 k events in the training. Here, a 5-fold was used as before. The number of training events for the baseline analysis (using a 2-fold) is about 38 k, which is well in the saturation region.

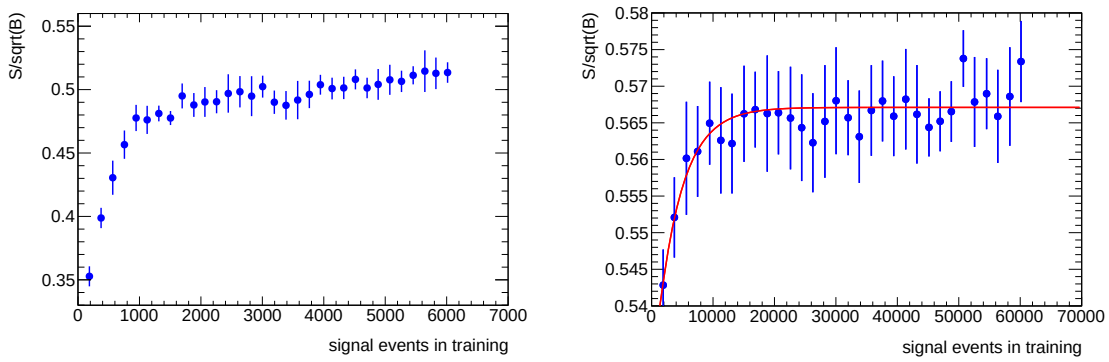


Figure 5.17.: BDT performance as a function of the number of signal events in the training with a small sample (left) and an enlarged sample (right). Each point represents the average best-cut statistical significance of five BDTs and the corresponding standard deviation. An exponential fit (red) is shown to guide the eye.

5.4. Statistical treatment

The measurement of the signal strength in the data is performed using a *profile likelihood fit*. Its description in the following is adapted from Ref. [106] and the implementation is described in Refs. [107, 108].

5.4.1. Profile likelihood fit

The binned likelihood function is a product of Poisson probability terms,

$$\mathcal{L}_{\text{Meas}}(\mu, \theta) = \prod_{i \in \text{bins}} \text{Pois}(N_i | \mu s_i + b_i) = \prod_{i \in \text{bins}} \frac{(\mu s_i + b_i)^{N_i}}{N_i!} e^{-(\mu s_i + b_i)}, \tag{5.6}$$

where μ is the signal-strength parameter, s_i (b_i) is the expected signal (background) yield and N_i is the observed number of events from the data in histogram bin i. The μ parameter is also called the *parameter of interest* (POI). For the present analysis, the signal is normalized to the SM Higgs boson production cross section so that $\mu = \sigma_{\text{meas}}/\sigma_{\text{SM}}$.

Systematic uncertainties are parameterized by a set of nuisance parameters (NPs), θ, which affect the expected signal and background yields: $s_i = s_i(\theta)$ and $b_i = b_i(\theta)$. The nominal values of most NPs have been determined from *auxiliary measurements*. To enforce this knowledge, a so-called penalty term,

$$\mathcal{L}_{\text{Aux}}(\theta) = \text{Gauss}(0|\theta, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\theta^2} \tag{5.7}$$

is multiplied to the likelihood. Here, the nominal value of θ is zero and the parameter is assumed to follow a Gaussian distribution with a standard deviation of $\sigma \equiv 1$. Some parameters, typically background normalization factors, are completely determined from

the data and do not have a penalty term. They are called *freely floating* parameters of the fit.

The statistical uncertainties of the total background are taken into account as γ-parameters. The initial $\mathcal{L}_{\text{Meas}}(\mu, \theta)$ is modified to multiply the background yield by γ_i for each bin i:

$$\text{Pois}(N_i|\mu s_i + b_i) \rightarrow \text{Pois}(N_i|\mu s_i + \gamma_i b_i) \tag{5.8}$$

This allows to modify the background expectation in each bin about the nominal value of $\gamma_i = 1$. Further, another likelihood term is defined to represent the statistical uncertainty, δ_i, as auxiliary measurement:

$$\mathcal{L}_{\text{BkgStat}}(\gamma_i) = \prod_{i \in \text{bins}} \text{Pois}(n_i|\gamma_i n_i) \tag{5.9}$$

The number of generated background events, n_i, is not directly available. However, it can be estimated from $\delta_i/b_i = 1/\sqrt{n_i}$, where δ_i/b_i is the relative statistical uncertainty of the background in bin i. This number n_i would fluctuate around $\gamma_i n_i$ for a newly generated sample, which justifies the Poisson term above. The full likelihood is then given as the product of the individual terms:

$$\mathcal{L}(\mu, \theta) = \mathcal{L}_{\text{Meas}}(\mu, \theta, \gamma_i) \cdot \mathcal{L}_{\text{Aux}}(\theta) \cdot \mathcal{L}_{\text{BkgStat}}(\gamma_i). \tag{5.10}$$

The nominal fit result in terms of μ is obtained by maximizing the likelihood function with respect to all parameters. The *test statistic* is constructed as

$$q_\mu = 2 \log \left(\frac{\mathcal{L}(\mu, \hat{\hat{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\theta})} \right), \tag{5.11}$$

where $\hat{\mu}$ and $\hat{\theta}$ are the parameters that maximize the likelihood and $\hat{\hat{\theta}}_\mu$ are the NP values that maximize the likelihood for a given value of μ.

The test statistic is used to estimate the level of disagreement of a specific hypothesis μ with the data, quantified by the p-value:

$$p_\mu = \int_{q_{\mu,\text{obs}}}^{\infty} f(q_\mu|\mu) dq_\mu. \tag{5.12}$$

Here, $q_{\mu,\text{obs}}$ is the observed value of the test statistic from the data and f denotes the probability density function (p.d.f.) of q_μ. An important special case is the p_0-value for rejecting the background only hypothesis ($\mu = 0$) in favor of the signal hypothesis.

The p-value is often given as a significance, Z, in terms of standard deviations, σ, of a Gaussian distribution,

$$Z = \Phi^{-1}(1 - p), \tag{5.13}$$

where Φ^{-1} is the inverse of the cumulative Gaussian distribution. Typical values are $Z_0 = 5\,\sigma$ ($p_0 = 2.87 \cdot 10^{-7}$) for claiming a discovery or $p_1 = 0.05$ (95 % confidence level, $Z_1 = 1.64\,\sigma$) for rejecting a signal hypothesis.

The p.d.f. of the test statistic, $f(q_\mu|\mu)$, is required for obtaining the p-values. This can be achieved by sampling the distribution with the Monte-Carlo method [109]. However, this method is computationally expensive. Instead, an approximation for f is employed, using a non-central χ^2 distribution [106].

Upper limits on the signal-strength parameter, μ, are derived with the CL_s method [110,111]. The CL_s value is defined by the ratio of two confidences, CL_{s+b} and CL_b, given by the p-values for the background and the background plus signal hypothesis:

$$\mathrm{CL}_s = \frac{\mathrm{CL}_{s+b}}{\mathrm{CL}_b} = \frac{1-p_\mu}{1-p_0} \tag{5.14}$$

The CL_s value is required to deviate from unity less than a chosen value, $1-\mathrm{CL}_s \leq \mathrm{CL}$, usually CL is set to 95 %. The corresponding values of μ are determined by varying this parameter in the estimation of p_μ.

The validation of the statistical model, encoded in the likelihood, is an important part of the analysis. A useful tool is the so-called *Asimov dataset*. It is not measured in the experiment, but is taken from the nominal background plus signal prediction. The validation techniques are discussed in Section 5.7.

5.4.2. Fit input distributions

The likelihood function is built based on the expected distributions of the BDT discriminants in the 2-tag regions, separately for the three lepton channels. Electron events with $p_\mathrm{T}^V < 120\,\mathrm{GeV}$ are not included for the 1-lepton channel. They were found to be not well modeled, as described in Section 5.7.8.

No BDT is available for the 0-lepton channel for the low-p_T^V region, due to uncertainties in the modeling of the multijet background for the full phase space. Instead, the m_{jj} distribution, which is found to be well modeled, is used in the fit, separately for the LL, MM and TT-tag regions.

In addition, the MV1c distribution of the tagged jet is used in the 1-tag regions, separately for the leading and sub-leading jet. The various tagging regions and the MV1c discriminant help to control the flavor fractions of the W/Z+jet backgrounds in the fit.

All regions and distributions that are used in the fit are summarized in Table 5.5. The LL, MM and TT-tag regions are used separately for the 1-lepton channel, while MM and TT are merged for the 0- and 2-lepton channels.

The BDT distributions are binned according to a specific algorithm, which is described in Section 5.6. Four bins are used for each MV1c distribution, corresponding to the calibrated working with 80 %, 70 %, 60 % and 50 % b tagging efficiency.

5.4.3. Smoothing and pruning

The likelihood function is built from one-dimensional histograms of the nominal distributions, as listed in the previous section. A set of alternative one-dimensional histograms is used to determine the impact from systematic uncertainties, one histogram for each variation and region.

Most variations consist of an up and of a down variation. If this is not the case, e.g. for jet energy resolution uncertainties, the existing variation is symmetrized with respect

N	Comment	Distr.	Leptons	Jets	b tags	p_{T}^{V}
4	SRs	BDT	0	2, 3	LL, MM+TT	high
6	SRs, muon only		1		LL, MM, TT	low
6	SRs		1		LL, MM, TT	high
8	SRs		2		LL, MM+TT	low, high
3	SRs, low $E_{\mathrm{T}}^{\mathrm{miss}}$	m_{jj}	0	2	LL, MM, TT	low†
1	Flavor CRs	MV1c	0	2	1	low†
2			0	2, 3	1	high
8			1, 2	2, 3	1	low, high

Table 5.5.: List of the 38 regions entering into the combined fit. In total 27 signal regions (SRs) and 11 control regions (CRs) are used. Low (high) p_{T}^{V} refers to events with $p_{\mathrm{T}}^{V} < 120\,\mathrm{GeV}$ ($p_{\mathrm{T}}^{V} \geq 120\,\mathrm{GeV}$). †: low p_{T}^{V} in the 0-lepton channel refers to the region with $100\,\mathrm{GeV} < E_{\mathrm{T}}^{\mathrm{miss}} \leq 120\,\mathrm{GeV}$.

to the nominal distribution to obtain the corresponding variation in the other direction. Some analyses apply a generic symmetrization to all variations. This is not done for the present analysis, since asymmetric responses can be of physical origin.

A specific issue arises for kinematic variations, such as the jet energy uncertainties, opposed to event weight variations, such as b tagging uncertainties. The latter do not change anything in the event, except its weight. A kinematic variation, instead, can for example change the energy of a jet, which will affect also the values of m_{jj} and of the BDT output. Hence, such a variation can cause migrations of the predicted events between the bins of the nominal histogram to the systematically varied histogram. In this case, the bin-wise difference between the nominal distribution and the variation can have a large statistical uncertainty.

This uncertainty is not taken into account in the fit, since it has usually no significant impact on the result. In addition, it is not available from the two one-dimensional histograms as they do not carry information about the statistical correlation between them. A proper estimate of the uncertainty would be possible from a two-dimensional histogram of the nominal distribution versus the variation. However, such histograms are not trivially constructed due to the separation between the nominal event record ("N-tuple") and the corresponding record of the variation[2].

Statistical fluctuations in the difference between the nominal distribution and the variations can lead to issues in the fit, which are further discussed in Section 5.7. To alleviate these issues, *smoothing* and *pruning* techniques are applied. The following prescriptions have been developed within the ATLAS collaboration

Smoothing The smoothing of a systematic variation assumes a correlation between bins in the ratio of a systematically varied distribution over the nominal one. Under this assumption, an algorithm can be applied to wash out ("smooth") statistical fluctuations between neighboring bins.

The specific algorithm devised for the present analysis assumes a monotonic behavior

[2]The CxAOD data format, devised for Run 2, is well suited for this task, as described in Section 8.1.

of the ratio as a function of the BDT output, i.e. exactly two maxima are allowed in the ratio. This is justified by the one-sided behavior of the BDT distribution: the expected signal is concentrated at large values, while the background events have predominantly lower values. Instead for the dijet mass, where the expected signal peaks on top of a broad background distribution, an (inverted) u-shape is likely to occur in the ratio. Consequently, three maxima are allowed in the ratio in this case by the algorithm.

Bins are merged until the desired number of (local) extrema remain in the ratio. This is performed iteratively, each time merging neighboring bins, one of which constitutes a local extremum. If there are more than two extrema, the merging is done where the difference before and after is smallest.

In a second step, the resulting bins are further merged until the statistical uncertainty in each of the merged bins of the nominal distribution is smaller than 5 %. No smoothing is applied to the MV1c distributions.

The merging of bins is only applied to the ratio of a systematically varied distribution over the nominal one. It is not intended to reduce the number of bins in the fit. Instead, the obtained smoothed ratio is ported back to the original distributions: each bin of the nominal distributions is multiplied by the corresponding value of the smoothed ratio to obtain the systematically varied and smoothed value.

Pruning Systematic variations that have a very small effect and are negligible for the measurement can be removed ("pruned"). This helps to improve the performance and stability of the fit. For the present analysis, the pruning is applied after the smoothing described above.

The pruning consists of removing a specific variation for a specific region and sample (signal or one of the backgrounds) if its effect on the normalization is less than $\pm 0.5\,\%$, or if its up and down variations go into the same direction ("one-sided"). Shape variations are removed if they do not show any bin with a shift above 0.5 % or if this is true for the corresponding variation in the other direction. Further pruning steps are applied to very small backgrounds if they are less than 2 % of the total background.

5.4.4. Inter- and extrapolation

The effect of each systematic variation is described by three values for each bin of the analysis: the nominal yield of the prediction, $I^0 = I(\theta = 0)$, and the corresponding values for the up and down variations, $I^{\pm} = I(\theta = \pm 1)$, where θ is the corresponding nuisance parameter. However, the fit requires a continuous parameterization of the likelihood as a function of θ. This is achieved using inter- and extrapolation functions, which are determined from the three values above. The following description uses $\alpha \equiv \theta$ and is adapted from the HistFactory documentation [112].

The HistFactory package provides several options for the functional form of the inter- and extrapolation. The present analysis uses the polynomial interpolation and exponential extrapolation for normalization uncertainties that have a prior. A typical example is depicted in Figure 5.18 (left, green line). Here, the yields are normalized to the nominal

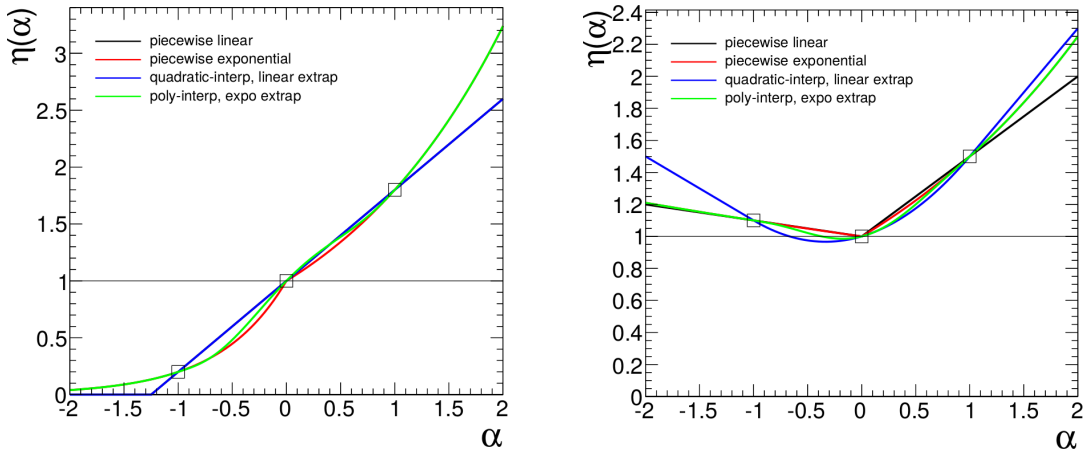


Figure 5.18.: Examples for inter- and extrapolation of a systematic variations in a single bin of the analysis, using various functions. Left: $\eta^- = 0.2$, $\eta^+ = 1.8$, right: $\eta^- = 1.1$, $\eta^+ = 1.5$ (from Ref. [112]).

one as

$$\eta^{\pm} = \frac{I^{\pm}}{I^0} \quad \text{and} \quad \eta^0 = \frac{I^0}{I^0} = 1 \tag{5.15}$$

and the continuous parameterization is given by $\eta(\alpha)$.

An important feature of the exponential extrapolation is to prevent the yield to become negative, which would be unphysical. Instead, for a linear extrapolation of a large negative prior the yield could become negative (black and blue lines).

For shape uncertainties the quadratic interpolation and linear extrapolation is used. An example for an asymmetric, even one-sided, variation is shown in Figure 5.18 (right, blue line). Such cases can create specific features in the fit result, in particular since the slope of $\eta(\alpha)$ can depend strongly on α. This is further discussed in Section 5.7.

5.5. Systematic uncertainties

As stated above, systematic uncertainties are parameterized as nuisance parameters (NPs) in the fit. Some parameters, mostly normalization scale-factors for the backgrounds and the one for the signal (μ), are freely floating, while others have a-priori uncertainties from auxiliary measurements ("priors" in the following). Further, some parameters affect only the normalization of some sample (the corresponding priors are given in tables in the following), while others affect the shape of distributions as well (abbreviated as "S" in the tables) and some explicitly preserve the normalization for each distribution and affect the shape only ("SO").

An important choice in the parameterization of the NPs is the correlation among regions, which has been evaluated for the present analysis in the course of this thesis. Some uncertainties, such as experimental effects, have one specific source and should be correlated among all regions. Others, such as multijet normalization parameters, are derived

independently for specific regions and should be treated uncorrelated. The chosen correlation schemes are indicated by appending tags to the NP names in the following. The naming convention is:

- `LX` with `X` = `0,1,2` for the lepton channel,
- `JX` with `X` = `2,3` for the jet multiplicity,
- `BX` with `X` = `0,2` for the p_{T}^{V} regions (bins) and
- `TTypeX` with `X` = `ll,mm,tt,xx` for the b tagging regions.

Here, `xx` denotes the merged MM+TT region of the 0- and 2-lepton channels. The two p_{T}^{V} regions are specific to the MVA analysis discussed in this thesis. The corresponding cut-based analysis, described in Ref. [1], employs five p_{T}^{V} regions (`BX` with `X` = `0` to `4`). Sometimes the tag `Y2012` is appended to specify a parameter is only valid for the MVA or cut-based analyses based on the data taken during 2012 at $\sqrt{s} = 8\,\mathrm{TeV}$.

5.5.1. Experimental uncertainties

All experimental systematic uncertainties, as listed in Section 4.7, are taken into account. They are propagated to the BDT and MV1c distributions and are parameterized as shape NPs, correlated among all regions and samples.

An exception are the `JetFlavComp/Resp_X` parameters, which parameterize the different calorimeter response to gluon and quark jets. They are process dependent, such that they are taken into account uncorrelated between the W+jets, Z+jets, $t\bar{t}$ plus single-top, and VH plus VV processes: `X` = `Wjets`, `Zjets`, `Top` or `VHVV`.

5.5.2. Signal

The signal-specific nuisance parameters are listed in Table 5.6. The total cross section is taken from Ref. [30] and the $H \to b\bar{b}$ decay branching ratio uncertainty (`TheoryBRbb`) is taken from Refs. [30, 113]. Uncertainties on NNLO QCD corrections (`TheoryQCDscale` and `TheoryVPtQCD`) are taken from Ref. [114] and on NLO EW corrections (`TheoryVHPt`) from Ref. [115]. Further, uncertainties on the acceptance (`TheoryAcc`), the parton shower (`TheoryAccPS`) and the PDF set (`TheoryPDF` and `TheoryAccPDF`) are taken into account.

All signal NPs are correlated throughout all regions. They are treated as uncorrelated between the $qq \to VH$ (`qqVH`) and $gg \to ZH$ (`ggZH`) processes (hence two parameters per uncertainty), except for the branching ratio and parton shower uncertainties (one parameter per uncertainty).

5.5.3. W bosons + jets

The W+jets process is a major background for the 0- and 1-lepton channels. It has been studied in comparisons of the simulation to the data in various kinematic distributions for the 1-lepton channel in the course of this thesis. The background is categorized into W+bb, bc, bl, cc, cl and l, based on the truth flavor of the two leading jets, as explained in Section 5.2.1. Uncertainties for the W+l and W+cl components are derived from

Name	Regions	qqWH	qqZH	ggZH	Count
μ	all		float		1
`TheoryBRbb`	all	3.3%	3.3%	3.3%	1
`TheoryQCDscale`	all	1%	1%	50%	2
`TheoryVPtQCD`	all	S	S	S	2
`TheoryVHPt`	all	S	S	–	1
`TheoryAcc_J2`	2-jet	3.0%	3.4%	1.5%	2
`TheoryAcc_J3`	2-jet	-1.1%	-0.9%	-1.9%	2
	3-jet	4.1%	3.5%	3.3%	
`TheoryAccPS`	all		7% - 13%		1
`TheoryPDF`	all	2.4%	2.4%	17%	2
`TheoryAccPDF`	2-jet	3.5%	3.0%	2.1%	2
	3-jet	2.8%	5.0%	3.4%	
Total	15 with priors, 1 floating				

Table 5.6.: Nuisance parameters, as defined in the text, and their priors for the signal-specific systematic uncertainties.

the 0- and 1-tag control regions, respectively. However, for the heavy flavor components ($hf = bb, bc, bl, cc$) no clean control regions are available due to the large $t\bar{t}$ background.

The W+jets specific nuisance parameters are listed in Table 5.7. Normalization uncertainties are taken into account separately for the 2- and 3-jet regions and for the flavor components. Floating parameters are used for the W+hf (and not only the W+bb components, as the name of `norm_Wbb` for the parameter suggests) and W+cl (named `norm_Wcl`) components, while the W+l normalization uncertainty has a prior (`WlNorm`).

Additional normalization uncertainties are taken into account for the 3-jet regions, which are defined relative to the normalization in the 2-jet regions (`WlNorm_J3`, `WclNorm_J3` and `WbbNorm_J3`). They are derived from comparing the default Sherpa generator to the Powheg + Pythia8, aMC@NLO + Herwig++ and Algen + Herwig generators. Uncertainties on the ratio of the heavy flavor components are derived from the same generator comparisons and are defined relative to the W+bb component (`WblWbbRatio`, `WbcWbbRatio` and `WccWbbRatio`).

The agreement of the MC with the data in the 0-tag regions, which constitute very clean W+jets control regions, was found to be rather poor initially, as shown in Figure 5.19 (left). Hence, a correction has been derived, which is based on the $\Delta\phi(j_1, j_2)$ distribution. A mostly linear fit to the ratio of data/MC is performed (second-order polynomials are fitted to the edges for small and large values of $\Delta\phi(j_1, j_2)$). The resulting function is used to derive weights for correcting the MC events. This is done separately for the 2- and 3-jet regions.

The agreement in the $\Delta\phi(j_1, j_2)$ distribution is clearly better after the correction as expected, but also other variables improve, such as p_{T}^{V} (Figure 5.19, right). Corrections derived from other variables than $\Delta\phi(j_1, j_2)$ were investigated, but found to perform worse. The correction, derived in the 0-tag region, where W+l dominates, is applied to W+cl as well and improves the modeling in the 1-tag region. Half of the correction is applied as

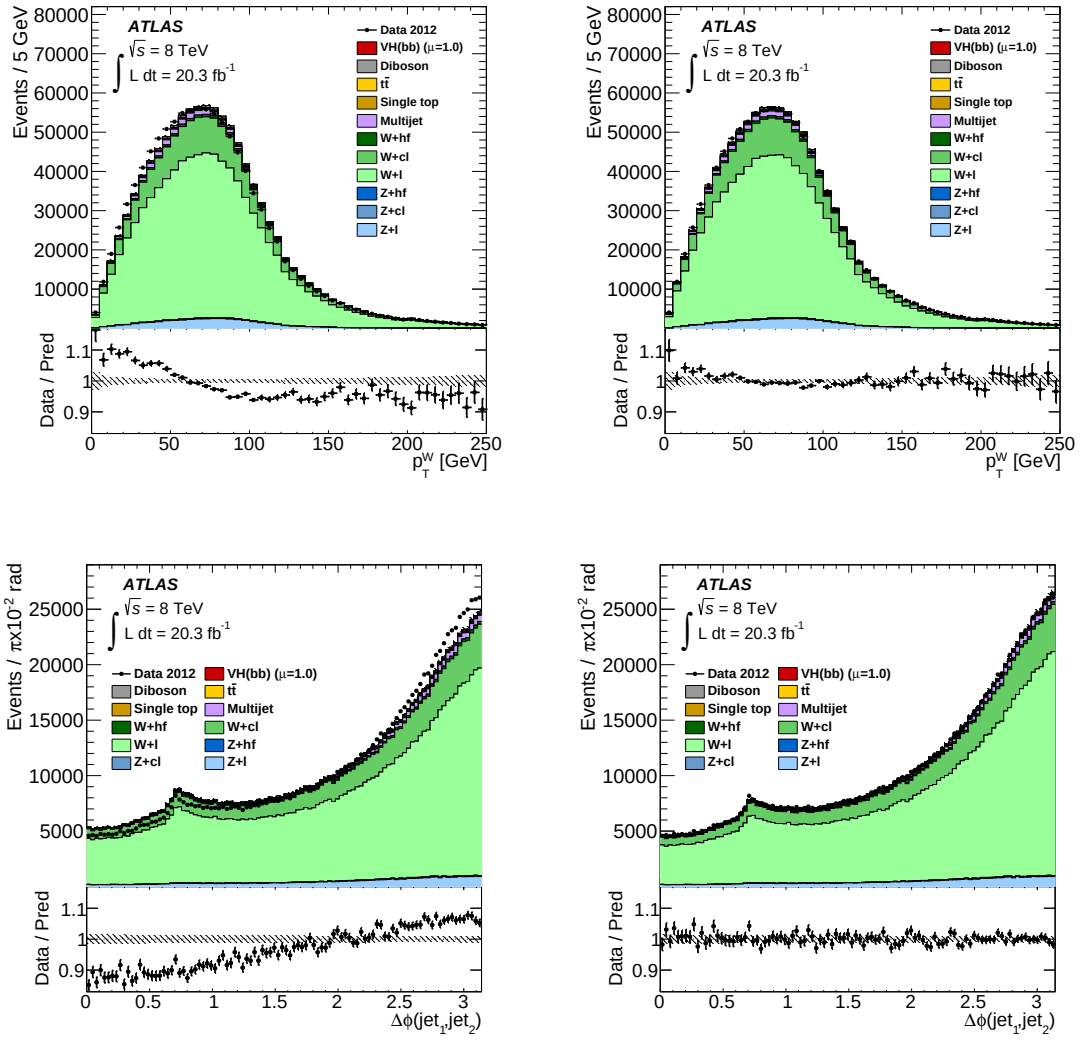


Figure 5.19.: Distributions of p_{T}^{W} (top) and $\Delta\phi(j_1, j_2)$ (bottom) for the 1-lepton channel, 0-tag region. The plots on the left are shown before the $\Delta\phi(j_1, j_2)$ correction, while is it applied on the right. The data are shown as dots with error bars and the expected background yields are shown as stacked colored histograms. The lower panels show the ratio of the data over the background yield (from Ref. [1]).

Name	Samples	Regions	Value	Count
WlNorm	W+l	all	10%	1
WlNorm_J3	W+l	3-jet	10%	1
norm_Wcl	W+cl	all	float	1
WclNorm_J3	W+cl	3-jet	10%	1
norm_Wbb	W+hf	all	float	1
WbbNorm_J3	W+hf	3-jet	10%	1
WblWbbRatio	W+bl	p_{T}^{V}-bins	35%	3
WbcWbbRatio	W+bc	all	12%	1
WccWbbRatio	W+cc	all	12%	1
WDPhi	W+hf, cl, l	2-, 3-jets	S	6
WMbb	W+bb/cc, bc/bl, cl, l	all	SO	6
WPtV	W+hf	2-, 3-jets	S	2
Total	23 with priors, 2 floating			

Table 5.7.: Nuisance parameters, as defined in the text, and their priors for the systematic uncertainties on the W+jets background.

a systematic uncertainty (WDPhi). However, it is not clear whether the correction should be applied to the heavy flavor components as well, since different processes are involved (mostly gluon splitting for W+bb and W+cc, which is not the case for W+cl and W+l). This is not done. Instead, the full impact of the correction is applied as uncertainty.

The uncertainty on $\Delta\phi(j_1, j_2)$ is treated as uncorrelated for the W+hf, W+cl and W+l components in the fit. This is indicated by commas in Table 5.7 and similar for other cases in the following. Additional uncertainties on the m_{jj} (WMbb) and p_{T}^{W} shape (WPtV) are derived from comparing the default Sherpa generator to Powheg + Pythia8, aMC@NLO + Herwig++ and Algen + Herwig.

5.5.4. Z bosons + jets

The Z+jets process is a major background for the 0- and 2-lepton channels. It is categorized into the flavor components, based on the two leading jets, as W+jets. However, for Z+jets control regions can be derived even for the heavy flavor components, since other SM processes can be suppressed efficiently in the 2-lepton selection. The $ZH \to \ell\ell b\bar{b}$ signal contribution is largely reduced by vetoing events having a value of m_{jj} close to the expected value for the signal of $m_H = 125\,\mathrm{GeV}$.

The Z+jets specific nuisance parameters are listed in Table 5.8. Normalization uncertainties are taken into account separately for the 2- and 3-jet regions and the flavor components, following the same procedure and naming schemes as for the W+jets background.

Similar to W+jets, a bad modeling of the $\Delta\phi(j_1, j_2)$ has been observed initially for Z+jets in the 2-lepton, 0-tag selection. A correction is derived and applied to the Z+l component. This improves the agreement of the data with the MC prediction in the 1-tag region as well. Half of the correction is applied as a systematic uncertainty (ZDPhi). However, it was found that the correction is not necessary for the non-light flavor components

Name	Samples	Regions	Value	Count
ZlNorm	Z+l	all	5%	1
ZlNorm_J3	Z+l	3-jet	5%	1
norm_Zcl	Z+cl	all	float	1
ZclNorm_J3	Z+cl	3-jet	26%	1
norm_Zbb	Z+hf	all	float	1
ZbbNorm_J3	Z+hf	3-jet	20%	1
ZblZbbRatio	Z+bl	2, 3-jet	12%	2
ZbcZbbRatio	Z+bc	all	12%	1
ZccZbbRatio	Z+cc	all	12%	1
ZDPhi	Z+b/c, Z+l	2, 3-jet	S	4
ZPtV	Z+b/c, Z+l	all	S	2
ZMbb	Z+b/c, Z+l	all	SO	2
Total	16 with priors, 2 floating			

Table 5.8.: Nuisance parameters, as defined in the text, and their priors for the systematic uncertainties on the Z+jets background.

(Z+b/c). Instead, the full correction is applied as uncertainty.

For Z+jets in the 2-lepton, 2-tags selection a residual disagreement between the data and the MC prediction in the p_{T}^{V} distribution is observed. A logarithmic fit to the ratio of data/MC is performed and applied as correction to Z+b/c. Half of the correction is applied as a systematic uncertainty uncorrelated between Z+b/c and Z+l (ZPtV). An additional uncertainty on the m_{jj} distribution (ZMbb) is derived from a linear fit to the data/MC ratio in the 0-, 1- and 2-tag regions separately.

5.5.5. Top-quark pairs

The list of all $t\bar{t}$ specific systematic uncertainties is given in Table 5.9. The freely floating normalization scale-factor of the $t\bar{t}$ background (norm_ttbar) was initially correlated for all regions in the fit. However, it was found to be inconsistent between the three lepton channels from independent fits, which is probably related to the different phase spaces selected for the three lepton channels, as discussed in Section 5.7.2. Therefore, the corresponding nuisance parameter was de-correlated: three independent parameters are used in the fit, one for each channel, instead of one parameter for all three channels.

The $t\bar{t}$ background prediction from the simulation does not well model the p_{T}^{V} distribution observed in the data. This mis-modeling is consistent with the top p_{T} being too hard in the simulation. A dedicated differential cross section measurement has found a similar behavior [116]. Correction scale-factors as a function of the average top-p_{T} are derived from that measurement and half of the correction is applied as a systematic uncertainty (TopPt).

Further, the default Powheg+Pythia generator is compared to the Powheg+Herwig, AcerMC, aMC@NLO and Algen+Pythia generators in various distributions. The largest discrepancies are found for Algen+Pythia in m_{jj} and $E_{\mathrm{T}}^{\mathrm{miss}}$. Linear fits to the ratios in those distributions are done and taken as systematic uncertainties (TtbarMBBCont

Name	Regions	Value	Count
norm_ttbar	lepton channels	float	3
ttbarNorm_J3	3-jet in 2, 0+1 lepton	20%	2
TopPt	all	S	1
TtbarMBBCont	all	SO	1
TtbarMetCont	1-lepton	S	1
ttbarHighPtV	$p_{\mathrm{T}}^{V} > 120$ GeV	7.5%	1
Total	6 with priors, 3 floating		

Table 5.9.: Nuisance parameters, as defined in the text, and their priors for the systematic uncertainties on the $t\bar{t}$ background.

Name	2-jet		3-jet		Count
	low p_{T}^{V}	high p_{T}^{V}	low p_{T}^{V}	high p_{T}^{V}	
s-channel					
stopsNorm	4%				1
SChanAcerMC	13%	22%	18%	30%	1
SChanAcerMCPS	6%	8%	4%	4%	1
t-channel					
stoptNorm	4%				1
TChanPtB2	52%	25%	12%	-18%	1
Wt-channel					
stopWtNorm	7%				1
WtChanAcerMC	1%	S/-2%	S/4%	S/-15%	1
WtChanPythiaHerwig	S/5%	3%	5%	-3%	1
Total	8 with priors, 0 floating				

Table 5.10.: Nuisance parameters, as defined in the text, and their priors for the systematic uncertainties on the single-top background.

and TtbarMetCont).

The propagation of these systematic uncertainties to the BDT output show that the differences between any of the generators are covered. Only a remaining normalization discrepancy between low and high-p_{T}^{V} was found and taken as an additional uncertainty (ttbarHighPtV).

5.5.6. Single top quarks

For the single-top background no clean control region is available. Hence, all systematic uncertainties, as listed in Table 5.10, are derived from generator comparisons. The uncertainties on the cross section from the calculation are taken into account seperately for the three processes (stopsNorm, stoptNorm and stopWtNorm).

The default generators are compared to aMC@NLO, Herwig and AcerMC with various factors (0.5, 1 and 2) for the renormalization scale. The largest deviations in terms of normalization are used as systematic uncertainties (SChanAcerMC(PS), WtChanAcerMC and

Name	Regions	WW	WZ	ZZ	Count
VVJetScalePtST1	all	S	S	S	1
VVJetScalePtST2	all	S	S	S	1
VVJetPDFAlphaPt	2-jet	2%	2%	3%	1
	3-jet	3%	4%	3%	
VVMbb_WW	all	SO	–	–	1
VVMbb_WZ	all	–	SO	–	1
VVMbb_ZZ	all	–	–	SO	1
Total	6 with priors, 0 floating				

Table 5.11.: Nuisance parameters, as defined in the text, and their priors for the systematic uncertainties on the diboson backgrounds.

WtChanPythiaHerwig). Additional disagreement in kinematic distributions is observed for the t-channel comparing the default Powheg+Pythia generator to AcerMC+Pythia in the p_{T}^{j2} distribution and to Powheg+Herwig in m_{jj}. The corresponding systematic uncertainties (TChanPtB2) are parameterized as linear fits to the ratios and propagated to the fit.

5.5.7. Vector-boson pairs

For the diboson backgrounds, similar to single-top, no clean control region is available and the systematic uncertainties, as listed in Table 5.11, are derived from generator comparisons.

A systematic uncertainty from scale variations is derived differentially in p_{T}^{V}. The Stewart-Tackmann method [117, 118] is employed to evaluate the uncertainty for the 2- and 3-jet regions separately (VVJetScalePtST1,2). The uncertainty scales from a few percent at low p_{T}^{V} to about 28 % for $p_{\mathrm{T}}^{V} > 200\,\mathrm{GeV}$ in the 2-jet region.

The uncertainty from the PDF set and the strong scale are evaluated as well (VVJetPDFAlphaPt). It is derived taking the envelope of the error bands provided by the CT10 and MSTW2008 PDF sets, evaluated for two values of 0.117 and 0.119 for α_S.

Additionally, uncertainties on the m_{jj} distribution (VVMbb_WW, WZ and ZZ) are derived by comparing the default Powheg+Pythia8 generator to Herwig. It ranges from 10 % to about 20 % for ZZ, from 10% to about 45 % for WZ and up to about 20 % for WW.

5.5.8. Multijet

The multijet (MJ) background in the 0- and 2-lepton analyses constitutes about or less than 1 % of the total background, so that conservative uncertainties of ±100 % on the normalization (MJ_L0 and MJ_L2) are sufficient for a reliable measurement of the signal.

In the 1-lepton channel, the MJ background is larger and a more sophisticated estimate of the systematic effects is done. The normalization uncertainties are taken from the statistical uncertainties of the templates fits, which are described in Section 5.2.3. The uncertainties are determined for the electron and muon channels separately and parameterized as MJEl/MuNorm, as listed in Table 5.5.8.

Name	Regions	Value	Count
	0- and 2-lepton		
MJ_L0	2, 3-jet; 1, 2-tag; low, high p_{T}^{V}	100%	6
MJ_L2	all	100%	1
	1-lepton		
MJMuNorm	2-jet; 1, 2L, 2M, 2T-tag	12, 28, 42, 60%	4
	3-jet; 1, 2-tag	11, 14%	2
MJElNorm	2-jet; 1, 2L, 2M, 2T-tag	3, 11, 14, 22%	4
	3-jet; 1, 2-tag	4, 6%	2
MJMuTrkIso	2, 3-jet; 1, 2-tag	S	4
MJElTrkIso	2, 3-jet; 1, 2-tag	S	4
MJElCaloIso	2-jet; 1, 2-tag	S	2
MJDR	electron; 2-jet; 2-tag	S	1
MJPtV	electron; 2-jet; 2-tag	S	1
Total	31 with priors, 0 floating		

Table 5.12.: Nuisance parameters, as defined in the text, and their priors for the systematic uncertainties of the multijet background normalization.

Shape and normalization uncertainties are derived by varying the requirements on the track and calorimeter isolation. For the electron channel, an alternative template with $0.12 < i_{\mathrm{trk}} < 0.50$ and another with $0 < i_{\mathrm{calo}} < 0.04$ is defined. For the muon channel the alternative criteria are $0.07 < i_{\mathrm{trk}} < 0.095$ and $0.095 < i_{\mathrm{trk}} < 0.5$. The effects on the multijet estimate are parameterized as MJElCaloIso and MJMu/ElTrkIso. Further, half of the corrections on the $\Delta R(j_1, j_2)$ and p_{T}^{W} distributions for the electron channel, described in Section 5.2.3, are taken as systematic uncertainties (MJDR and MJPtV).

5.6. Binning strategies

A variable-size binning has been developed within the course of this thesis for the distributions entering into the statistical model. It is designed to optimize the search sensitivity, while improving the stability of the fit. The two figures of merit used to optimize the binning are the expected sensitivity and the reduction of the number of bins.

For example, the total number of bins in the analysis, if each BDT distribution has 20 bins, is over 500. The fit would need to handle all of them, all the uncertainties (including the bin-wise statistical uncertainties of the background) and the correlations across them. This is hardly feasible.

Several binning algorithms are defined and optimized in the following. Each algorithm works with a finely binned input histogram and generates an output histogram with a coarser binning. Each bin in the output histogram is defined by an interval, $I(k, l) = [k, l]$, of the bins from k to l in the input histogram.

The statistical significance of the signal over the total background expectation is a useful figure of merit. It is defined as a function of the bin contents of a histogram as $\sqrt{llr_s}$,

where llr_s is an approximate likelihood ratio [106],

$$llr_s(k,l) = \sum_{i=k}^{l} s_i \ln\left(1 + s_i/b_i\right), \tag{5.16}$$

and s_i (b_i) are the number of signal (background) events in bin i.

5.6.1. Numerical optimization

A numerical optimization algorithm is defined to search for the optimal binning. The goal is to find the maximal square of statistical significance, $LLR_s = llr_s(1, N_{\text{bin}})$, for a given number of bins, N_{bin}, in the output histogram.

A function $f(\boldsymbol{k})$ is defined as the negative LLR_s,

$$f(\boldsymbol{k}) = -LLR_s(k_0, k_1, ..., k_{N_{\text{bin}}}), \tag{5.17}$$

where the vector $\boldsymbol{k}$ contains the bin boundaries that map the input to the output histogram: bin i in the output histogram is defined by the interval $I(k_{i-1}, k_i - 1)$. The MINUIT package [119] is used to minimize f with respect to all k_i.

This generic approach is certainly suited to find the optimal binning. However, the global minimum is not easily found, since $f(k_i)$ can be a high-dimensional (N_{bin}) and non-continuous function.

An iterative approach is used to obtain a reliable minimization. Starting from $N_{\text{bin}} = 2$ the minimization is performed with ease. The result is used as starting point for a larger N_{bin}. This is repeated for increasing N_{bin} until the minimization breaks down. Two different strategies are used to increase the number of bins. At each step

a) N_{bin} is doubled, dividing each bin into two bins of half the size or

b) N_{bin} is increased by 1, adding a bin on the right.

The added bin in the second approach has a width of $\Delta x / N_{\text{bin}}$, where Δx is the x-axis range of the histogram. After adding the bin, the width of all bins is scaled down to fit the histogram. The results from these numerical optimization algorithms are discussed in Section 5.6.3.

5.6.2. Analytical algorithms

Several analytical binning algorithms ("transformations") are defined in the following to obtain intervals, I, of bins in an input histogram. The bins of each interval are merged to obtain one bin of the corresponding output histogram.

For each algorithm, a function Z is defined as

$$Z(I) = Z(z_s, N_s, n_s(I), z_b, N_b, n_b(I)), \tag{5.18}$$

where $z_{s,b}$ are free parameters, $N_{s,b}$ are the total number of signal (background) events in the input histogram and $n_{s,b}(I)$ are the number of signal (background) events in the interval $I(k,l) = [k,l]$, which contains the bin numbers from k to l of the input histogram.

The *transformations D, E* and *F*, which are discussed in the following, correspond to specific choices of the Z function[3]:

$$Z_{\mathrm{D}} = z_s n_s/N_s + z_b n_b/N_b \tag{5.19}$$

$$Z_{\mathrm{E}} = \sqrt{z_s llr_s/LLR_s} + \sqrt{z_b llr_b/LLR_b} \tag{5.20}$$

$$Z_{\mathrm{F}} = \sqrt{z_s llr_s} + \sqrt{z_b n_b/N_b} \tag{5.21}$$

Here, $\sqrt{llr_s(k,l)}$ is the statistical significance of the signal over the total background in the interval $I(k,l)$ of the input histogram and $\sqrt{llr_b(k,l)}$ is the statistical significance of the total background over the signal, defined by exchanging s and b in Equation 5.16. The total statistical significance of the signal (background) in the input histogram is defined by the square root of $LLR_{s(b)} = llr_{s(b)}(1, N_{\mathrm{bin}}^{\mathrm{input}})$, where $N_{\mathrm{bin}}^{\mathrm{input}}$ is the number of bins in the input histogram.

Each of these binning algorithms employs the following steps. A first boundary for the output histogram, k_0, is defined as $k_0 = N_{\mathrm{bin}}^{\mathrm{input}}+1$. Then the following steps are performed interatively. Starting from the rightmost bin of the input histogram, $k = l = k_0 - 1$, the range of the interval $I(k,l)$ is increased by adding one bin after the other (decreasing k by one each time). The value of Z is calculated at each step. Once the condition $Z(I(k,l)) > 1$ is fulfilled the second bin boundary, $k_1 = k$, is defined. All bins in the interval $I(k_1,l)$ are merged into a single bin for the output histogram. The next bin boundary, k_2, is searched for by redoing the steps above, starting now from $k = l = k_1 - 1$. This is repeated until all input bins are remapped and $k_{N_{\mathrm{bin}}}$ is found, where N_{bin} is the number of bins in the output histogram.

In transformation F the intervals I are requested, in addition to the condition on Z, to satisfy $\sigma_b < S$, where σ_b is the relative statistical uncertainty on the number of expected background events and S can be chosen freely, typically as $S = 5\,\%$ or $10\,\%$.

Examples for the effect of the transformations D and F in the m_{jj} and BDT distributions are given in Figure 5.20.

In order to allow the algorithms to work efficiently, the input histogram needs to provide a large number of bins. Here, $N_{\mathrm{bin}}^{\mathrm{input}} = 1000$ is chosen. However, for transformations E and F it should be avoided to have many bins without background expectation, but with some signal. For such bins the llr_s cannot be calculated and is taken as zero. A possible improvement would be to merge each bin without background expectation with one of its neighboring bins in a step before applying the transformation.

Transformation D is introduced for being a simple algorithm, which produces a well-defined shape, and has some useful features. First, it can be noted that the number of bins in the output histogram, N_{bin}, is directly connected to the free z parameters: $N_{\mathrm{bin}} = z_s + z_b$.

Further, the parameters allow to adjust the shape of the output distribution. For $z_s = 0$, $z_b > 0$ the background is equally distributed among the bins with $N_{\mathrm{bin}} = z_b$. Accordingly, for $z_s > 0$, $z_b = 0$ the output results in a flat signal distribution. It is possible to continuously interpolate between these cases, while keeping the number of bins constant, by choosing $z_s > 0$, $z_b > 0$ and fixing $z_s + z_b$.

[3]The naming of the binning algorithms has historical reasons and is kept here for consistency.

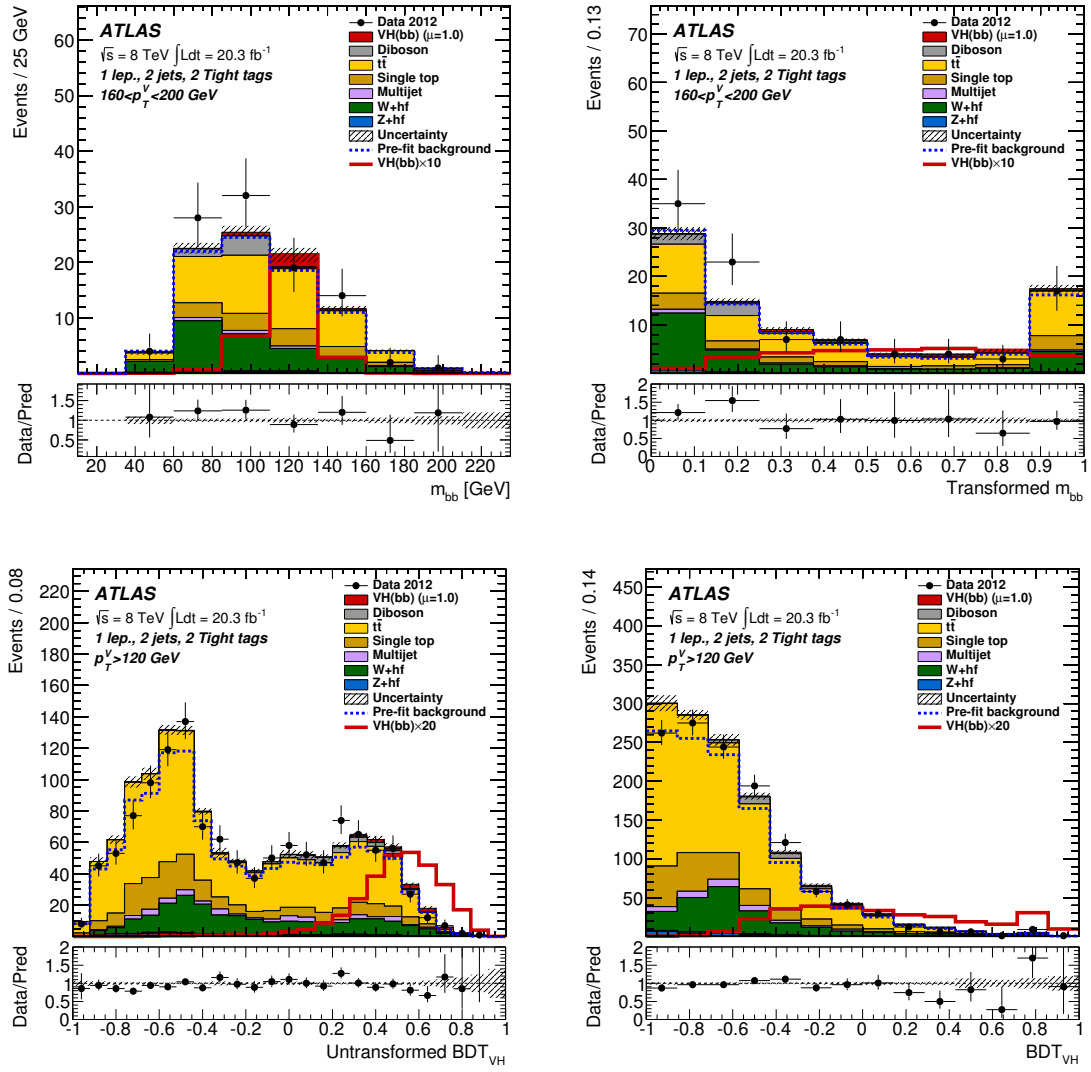


Figure 5.20.: Examples for the analytical binning algorithms. The top row shows the m_{bb} distribution with equal-width binning (left) and after transformation D with $z_s = 6$ and $z_b = 2$ (right). The bottom row shows the BDT distribution with equal-width binning (left) and after transformation F with $z_s = z_b = 4.5$ and $\sigma_b < 10\,\%$ (right). The transformation parameters have been determined from dedicated optimizations, as described in Section 5.6.7. The distributions are shown after the fit to the data, as described in Section 5.4.1. A formal description of the plot elements is given in Section 5.7.1 (from Ref. [1]).

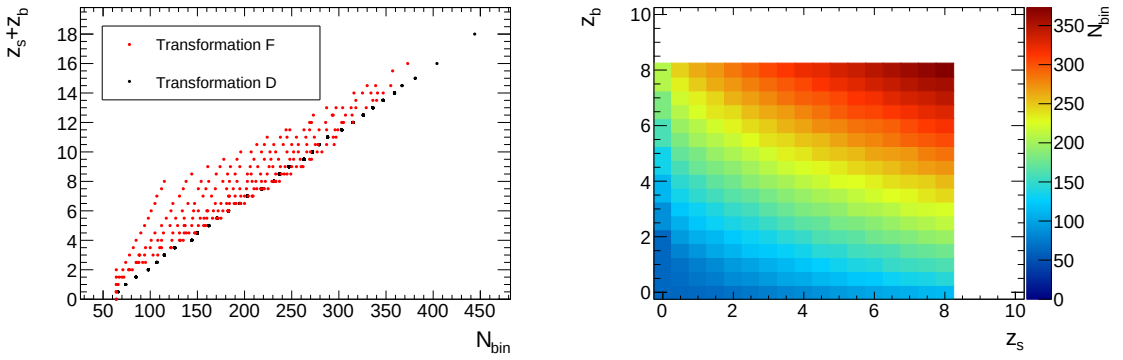


Figure 5.21.: Correlation between the total number of bins in the fit and $z_s + z_b$ for transformation D and F (left). Transformation D shows a clear linear dependence, while it is more complex for transformation F, as shown on the right.

Transformation F shows a similar dependence of N_{bin} on the z parameters, but it is not a simple linear one, as shown in Figure 5.21. Instead, there is another dependence: due to the llr_s term the number of bins in the output is positively correlated with the statistical significance of the signal in a given input histogram.

Transformations E and F were derived from investigating the results of the numerical optimization (Section 5.6.1), which is discussed in the following.

5.6.3. Performance comparison

The results from the numerical optimization and the analytical algorithms are discussed in the following. The statistical significance of the signal in the output histogram, LLR_s, of each algorithm is shown as a function of N_{bin} in Figure 5.22. Histograms with 1000 bins of the BDT output for the 2-jet, 2-tag region for the 0 (left) and 2-lepton (right) channel were used as input for the algorithms.

The results of the numerical optimization (minimization) algorithms (a) and (b) are compared to a equal-size binning ("rebinning") of the input histogram and to the *transformations D, E* and *F*. Clear trends are visible in the performance of the various binning algorithms in Figure 5.22 (left), where input distributions with relatively small statistical uncertainties were used.

As expected, all of the algorithms achieve a better sensitivity with a larger number of bins, N_{bin}, in the output. Further, some plateau seems to be reached for larger N_{bin}. This is expected as well, since it is not possible to separate the signal perfectly from the backgrounds in any (finite) phase space. Hence, some background is expected to be present in all of the bins that contain some signal yield, and the statistical significance is limited[4].

It can also be noted that the simple rebinning performs better than transformation D in terms of the statistical significance. This, however, depends on the shape of the BDT

[4]This is also true for the continuous limit with an infinite number of bins, given an infinite amount of predicted events and a BDT output with a monotonic behavior in the signal over background ratio. In this limit, the maximal, but still finite, statistical significance of the signal would be achieved.

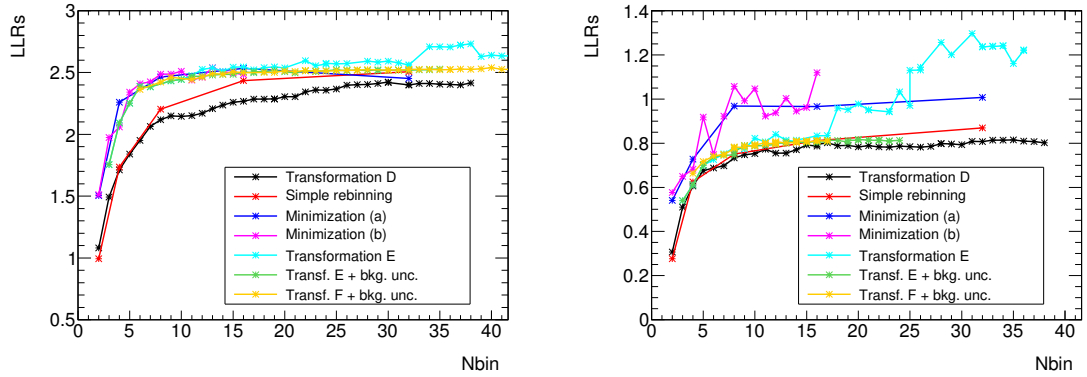


Figure 5.22.: Squared statistical significance of the signal, LLR_s, as a function of the number of bins in the output, N_{bin}, for various binning algorithms as described in the text. The BDT distributions of the 2-jet, 2-tag region from the 0-lepton (left) and 2-lepton channel (right) are used as input. The 2-lepton channel has significantly larger statistical uncertainties on the number of expected background events.

output and may vary.

Further, the two minimization algorithms increase the LLR_s significantly w.r.t the simple rebinning and transformation D, in particular for lower numbers of bins. Both minimization algorithms, though quite different, show very similar performance. This indicates that they indeed found the optimal binning.

The minimization algorithms are computing intense and can fail. Therefore, they are not used to define the final binning for the analysis. Instead, the results were further investigated to derive an analytical algorithm.

In particular, the binned statistical significance shows an interesting behavior, as shown in Figure 5.23 (upper left and middle). For both minimization algorithms the significance is almost equally distributed among the bins in the higher BDT output. This is quite different to transformation D, where the significance peaks at high BDT output. Here, most of the sensitivity is obtained from one single bin, resulting in worse performance than for the minimization algorithms.

This observation is translated into transformation E, as defined by Equation 5.20. It is designed to mimic the behavior of the minimization algorithms by directly using the statistical significance to define the bin boundaries. This is achieved approximately, as shown in Figure 5.23 (lower left). A rough optimization of the z parameters has led to a fixed ratio of $z_s/z_b = 4$. The number of bins in the output is varied by scaling z_s and z_b with the same factor. This algorithm is compared with the others in Figure 5.22 (left).

It is found that transformation E performs as well as the minimization algorithms. However, there is a worrying feature for very large number of bins. The significance shows upward fluctuations, which can be caused by bins with a large relative statistical uncertainty on the number of expected background events, σ_b. Indeed, this effect is enhanced for a BDT output with larger statistical uncertainties (Figure 5.22, right). The minimization

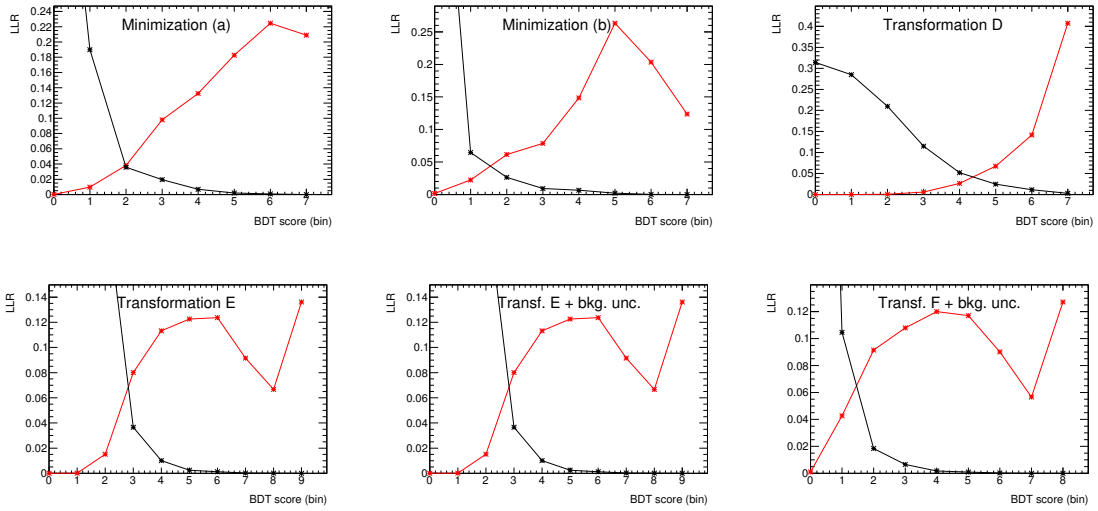


Figure 5.23.: Bin-wise squared statistical significance, LLR, for various binning algorithms as described in the text and for $N_{\text{bin}} = 8$ to 10, depending on the algorithm. The significance of the signal is shown in red and that of the background in black. The BDT distribution of the 0-lepton, 2-jet, 2-tag region is used as input for the algorithms.

algorithms show this behavior as well, even for lower numbers of bins in the output. This feature is prevented by introducing an explicit requirement of $\sigma_b < 5\,\%$ on the output bins (denoted by "+ bkg. unc." in the legend).

Transformation F, as defined by Equation 5.20, is derived from transformation E. The fractional statistical significance of the background, llr_b/LLR_b is replaced by the fraction of background events, n_b/N_b. This allows for a more predictable background shape and does not visibly affect the sensitivity. The fractional significance of the signal, llr_s/LLR_s, is replaced by the absolute value, llr_s. This causes the dependence of N_{bin} on the expected significance, as mentioned above, and is a good feature for the combined fit, as described in Section 5.6.5.

The requirement of $\sigma_b < 5\,\%$ is introduced to transformation F as well. The z parameters were chosen from a rough optimization as $z_b = 1$ and the number of bins in the output is varied by adjusting z_s only. The performance of transformation F is comparable to that of transformation E and the minimization algorithms, as shown in Figure 5.22.

5.6.4. Optimization of transformation D

The parameters of transformation D, which has been defined in Section 5.6.2, are optimized for the $VH(\to b\bar{b})$ analysis in the following. The two figures of merit are the number of bins in the output, N_{bin}, and the expected upper limit on the $VH(\to b\bar{b})$ signal cross section, derived with the combined profile likelihood fit (Section 5.4).

A first working point for transformation D was chosen as $z_s = z_b = 6$ for further optimization. To allow for more flexibility the values for z_s and z_b are redefined for each

region of the analysis (defined in Section 5.2):

$$z_{s,b} = f_{\text{tags}} \times f_{\text{jets}} \times f_{p_{\text{T}}^V} \times f_{\text{leptons}} \times \zeta_{s,b}. \tag{5.22}$$

Each parameter f and ζ can be freely chosen and is sequentially optimized in the following. The f parameters are designed to take different values for the various regions. For example, the f_{jets} parameter can have different values of $f_{\text{2-jet}}$ and $f_{\text{3-jet}}$ for the 2- and 3-jet regions, respectively.

First, the impact from varying the number of bins in the various b-tag categories (LL versus MM and TT) is studied by varying f_{tags}. All other f-parameters are set to 1, and $\zeta_s = \zeta_b = 6$ is chosen. The expected limit is then evaluated as a function of f_{LL}. The total number of bins is kept constant by setting

$$f_{\text{MM+TT}} = 2 - f_{\text{LL}} \qquad \text{(0- and 2-lepton)}, \tag{5.23}$$

$$f_{\text{MM}} = f_{\text{TT}} = (2 - f_{\text{LL}})/2 \qquad \text{(1-lepton)}. \tag{5.24}$$

The different parameterization for the lepton channels results from the choice of regions in the combined fit, as listed in Section 5.4.2. The result from scanning f_{LL} is shown in Figure 5.24 (first row). The optimal working point is chosen as $f_{\text{LL}} = 0.6$ for the 0- and 2-lepton channels and $f_{\text{LL}} = 0.2$ for 1-lepton.

Next, the fraction of bins in the 2-jet category versus 3-jet (f_{jets}) is optimized, then the fraction of bins in the low p_{T}^V versus the high p_{T}^V region ($f_{p_{\text{T}}^V}$), and finally to the fraction of bins in the 3 lepton categories (f_{leptons}). In each step the optimized value of the previous steps are used. In general, it can be noted that a lower number of bins for regions with low sensitivity is favored, while a higher number of bins for regions with higher sensitivity is preferred.

Finally, the optimal values for ζ_s and ζ_b are evaluated. The expected limit as a function of the parameter of interest for each step are shown in Figure 5.24. The overall results are summarized in Table 5.13. This output transformation reduces the total number of bins in the fit to about 300 with a degradation of less than 1% of the expected limit. This has to be compared to the case of 20 equal-size bins in each BDT distribution, resulting in over 500 bins in the fit.

A second iteration of the full optimization procedure is performed. The results are largely consistent with the first iteration. Figure 5.25 shows the final two steps of scanning f_{leptons} and $\zeta_{s,b}$.

The expected limit improves by either increasing ζ_s or ζ_b. This is equivalent to increasing the number of bins in the fit, since N_{bin} is proportional to $\zeta_s + \zeta_b$. This dependence is shown in Figure 5.26 (left). There is a strong dependence for low N_{bin}, while it is weaker for a larger N_{bin}.

Increasing the number of bins is certainly expected to improve the sensitivity. However, it has to be noted that no plateau is reached, which would have been expected due to the limited separation between the signal and the backgrounds. This might point to the fact that the sensitivity is overestimated for a very large number of bins and that the fit result cannot be trusted anymore. A possible cause might be the apparent increase of separation of the signal with the backgrounds that have increasingly large statistical fluctuations.

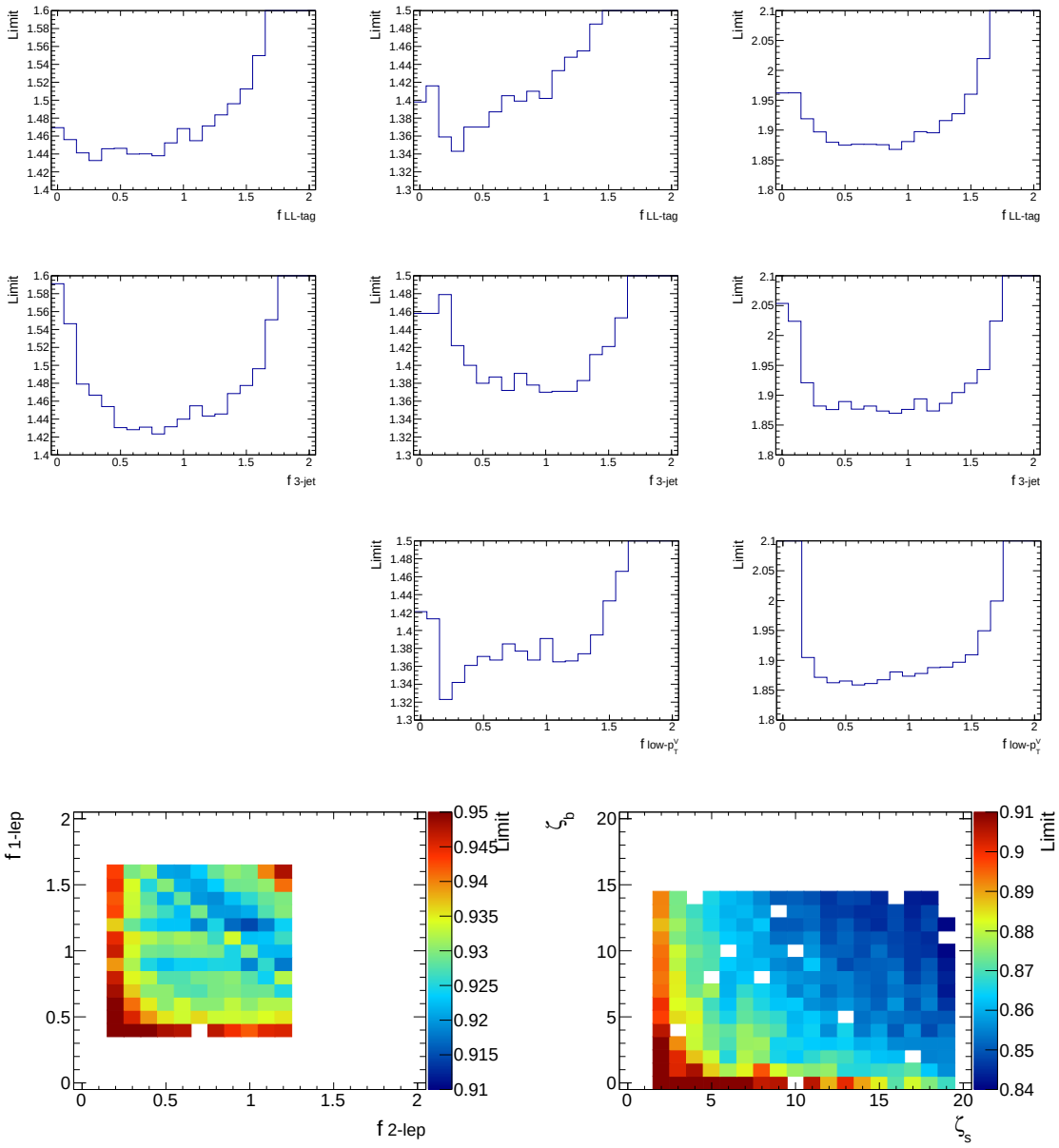


Figure 5.24.: Dependence of the expected limit on the transformation parameters during the first optimization iteration of transformation D. Rows 1 to 3 show the dependence on f_{tags}, f_{jets} and $f_{p_{\text{T}}^V}$ for the 0 (left), 1 (middle) and 2 lepton (right) channels. Row 4 shows the dependence on f_{leptons} (left) and $\zeta_{s,b}$ (right). Not all results in the scan region are available (white spots).

Parameter	Description	Region	Value
f_{tags}	fractions in the tagging regions (0- and 2-lepton)	LL	0.6
		MM+TT	1.4
	fractions in the tagging regions (1-lepton)	LL	0.2
		MM	0.9
		TT	0.9
f_{jets}	fractions in the jet categories	2-jet	1.2
		3-jet	0.8
$f_{p_{\text{T}}^V}$	fractions in the low and high p_{T}^V region	low p_{T}^V	0.6
		high p_{T}^V	1.4
f_{leptons}	fractions in the lepton channels	0-lepton	0.9
		1-lepton	1.2
		2-lepton	0.9
ζ_s	weight for the signal	any	6
ζ_b	weight for the background	any	4

Table 5.13.: Parameters for transformation D from the binning optimization, as described in the text.

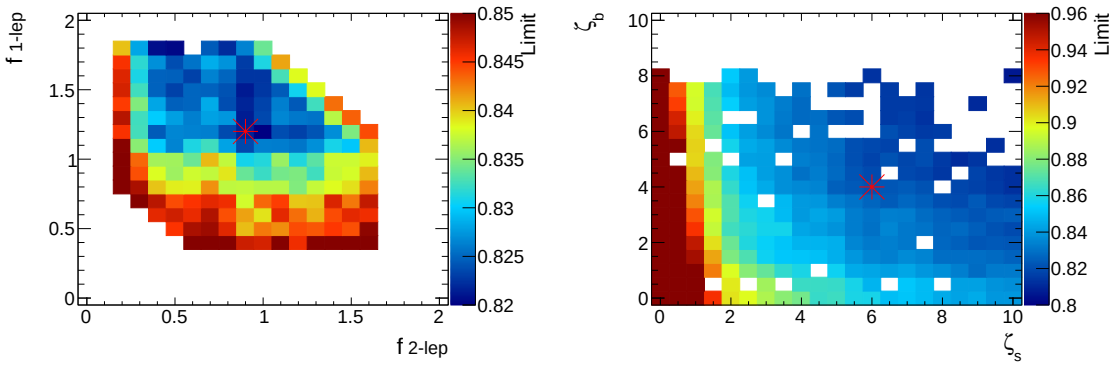


Figure 5.25.: Expected limit as a function of f_{leptons} (left) and $\zeta_{s,b}$ (right) during the second optimization iteration of transformation D. The chosen values are marked with red stars. Not all results in the scan region are available (white spots).

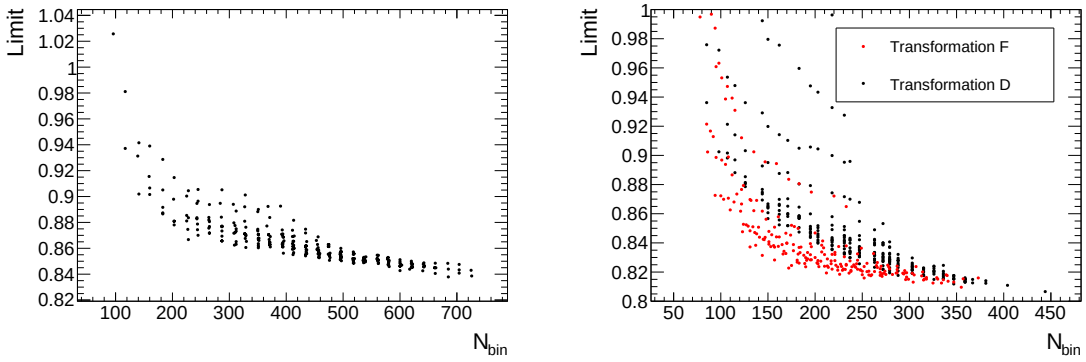


Figure 5.26.: Expected limit as a function of the total number of bins in the fit. Left: the points correspond to the two-dimensional scan of $\zeta_{s,b}$ during the first optimization of transformation D (Section 5.6.4). Right: second optimization of transformation D and scanning of $z_{s,b}$ for transformation F (Section 5.6.5). The relative statistical uncertainty on the number of expected background events, σ_b, is required to be less than 10 % for transformation F. No such requirement is applied for transformation D.

5.6.5. Optimization of transformation F

The parameters of transformation F, which has been defined in Section 5.6.2, are optimized for the $VH(\to b\bar{b})$ analysis in the following. The two figures of merit are the number of bins in the output, N_{bin}, and the expected upper limit on the $VH(\to b\bar{b})$ signal cross section, derived with the combined profile likelihood fit (Section 5.4).

As shown in the previous section, it is beneficial to allow for more bins in regions with higher sensitivity than for regions with lower sensitivity. This was implemented manually for transformation D. Instead, for transformation F, this behavior is directly contained in the algorithm by using the llr_s term.

Hence, the optimization of transformation F is simple: just the z_s and z_b parameters are varied. The relative statistical uncertainty on the number of expected background events, σ_b, is required to be less than 10 %. It was found that the previous value of 5 % degrades the sensitivity significantly and the larger value is still deemed sufficient for a reliable measurement.

The resulting expected limits are shown in Figure 5.27 for the full likelihood including systematic uncertainties (left) and for using only the statistical uncertainty of the Asimov dataset (right). The limit without systematic uncertainties shows less fluctuations as a function of $z_{s,b}$ than for the full likelihood.

As expected, the limit improves by increasing each of the z parameters, which increases the number of bins in the fit. The dependence of the limit on N_{bin} is shown in Figure 5.26 (right). It is similar to transformation D. However, the limit is superior already for lower N_{bin} and shows a saturation for larger N_{bin}. The latter is due to the requirement on σ_b in transformation F.

The working point for the $VH(\to b\bar{b})$ analysis is chosen as $z_s = z_b = 4.5$, which corre-

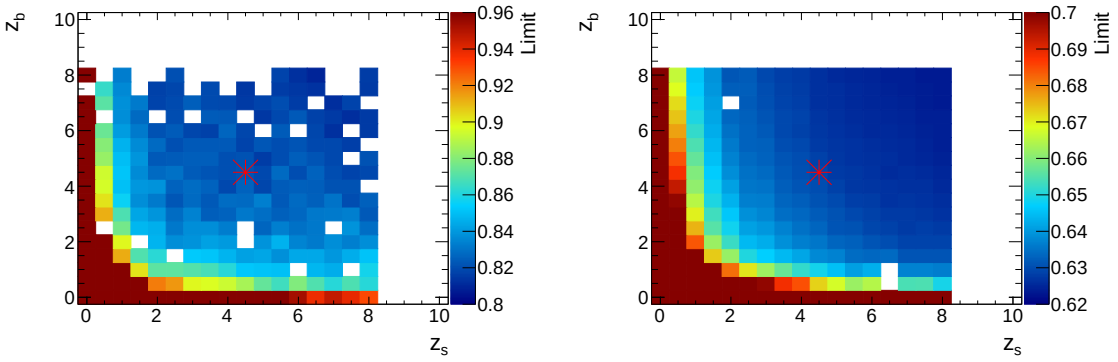


Figure 5.27.: Expected limit as a function of z_s and z_b in transformation F. The left plot shows the result using the full likelihood, while only the statistical uncertainty of the Asimov dataset is used on the right. The chosen values are marked with red stars. Not all results in the scan region are available (white spots).

sponds to about 250 bins in the fit in total. The expected limit is not changed compared to the optimized transformation D with about 300 bins.

5.6.6. Further observations

The chosen parameters for transformation F in the previous section are likely to be suboptimal for other analyses. In particular, since the number of bins in the output depends on the expected significance, the z_s parameter might have to be adjusted. A starting point could be $z_s = 10/\sqrt{LLR_s}$, where $\sqrt{LLR_s}$ is the expected statistical significance of the signal, to get sensible output distributions.

As mentioned, the binning algorithms allow for a continuous variation of the shape in the output histogram. A shape parameter can be defined as $S = z_s/(z_s + z_b)$, which is the fraction of "signal bins" in transformation D and similar for F. The expected limit as a function of S, while keeping $z_s + z_b$ constant, is shown in Figure 5.28.

For a lower number of bins a preference towards $S = 1$ (flat signal distribution) is observed in transformation D. For larger numbers of bins values around $S \simeq 0.5$ seem to be preferred in both algorithms.

5.6.7. Final binning

The final binning for the BDT distributions of the $VH(\to b\bar{b})$ analysis is obtained from transformation F with $z_s = z_b = 4.5$ and $\sigma_b < 10\,\%$, which corresponds to about 250 bins in the fit in total.

The binning for the dijet mass distributions of the cut-based analysis, which is described in Ref. [1], is obtained from transformation D and was optimized within the ATLAS collaboration. The z parameters are chosen independently for the three lepton channels and the 2- and 3-jet regions, as listed in Table 5.14. This binning reduces the number of

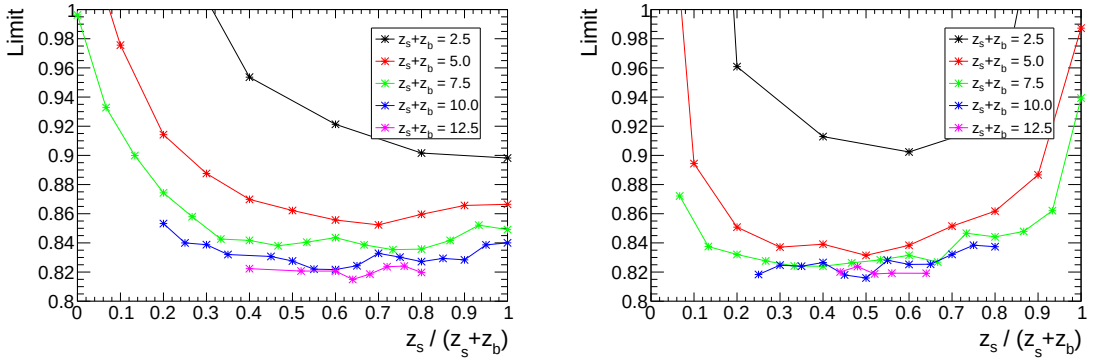


Figure 5.28.: Expected limit as a function of the shape parameter, $S = z_s/(z_s + z_b)$, for constant $z_s + z_b$ in transformation D (left) and transformation F (right).

Region	2-jet		3-jet	
	z_s	z_b	z_s	z_b
0-lepton	6	2	4	2
1-lepton	6	2	4	2
2-lepton	4	4	2	2

Table 5.14.: Parameters of transformation D used for the binning of the dijet mass distribution of the cut-based analysis, as described in Ref. [1].

bins by almost 45 % and improves the expected limit by 2 %, compared to 20 GeV bins in the dijet mass distributions (which was used for the preliminary result in Ref. [102]).

5.7. Fit model validation

The fit model, as described above, is quite complex, due to the use of many regions and many parameters in the fit. The validation of this model with respect to the data is an important part of the analysis and has been studied extensively in the course of this thesis.

5.7.1. Post-fit plots

A first impression of the result from the fit to the data is provided by the *post-fit plots*. They show the distributions of the data compared to the prediction, which has been adjusted according to the result of the fit. The VH signal with $m_H = 125\,\text{GeV}$ is taken into account and its normalization parameter, μ, is floating in the fit. The result can be found in Section 5.8.2.

All post-fit distributions of the regions that are used as input to the fit are shown in Figure 5.29 (BDTs for 0- and 2-lepton), Figure 5.30 (BDTs for 1-lepton), Figure 5.31 (MV1c for 1-tag) and Figure 5.32 (m_{jj} for 0-lepton, low-p_{T}^V region). The most sensitive distributions can be found in Section 5.8.1 in logarithmic scale for better visibility of the high BDT-output tail.

The various background components are shown stacked as colored filled histograms. The signal prediction with $\mu = 1$ is shown on top of the backgrounds as red filled histogram. It is also shown for better visibility as red hollow histogram with varying scale factors, as indicated in the legend. The data are shown as points with error bars. They are compared as ratio to the post-fit prediction in the lower panels. The combined statistical and systematic uncertainty on the prediction is indicated by the hatched band. The pre-fit background prediction is shown in the plots as dotted blue line.

In general, good agreement of the prediction with the data is observed after the fit, while the pre-fit prediction shows larger discrepancies. This difference is attributed the adjustment of the model parameters to the data. This is discussed further below.

5.7.2. Nuisance parameter pulls

Nuisance parameter *pulls* represent the result for each parameter of the fit, $\hat{\theta}$, which has been adjusted according to the dataset. Each pull is given relative to its pre-fit value of the parameter, θ_0, and normalized to its pre-fit uncertainty, σ_{θ_0}:

$$\text{pull} = \frac{\hat{\theta} - \theta_0}{\sigma_{\theta_0}} = \frac{\hat{\theta}}{\sigma_{\theta_0}}, \tag{5.25}$$

where $\theta_0 \equiv 0$ by construction of the likelihood function. Thus, if the data agrees with the nominal value of a specific NP, its post-fit pull is expected to be zero.

The post-fit uncertainties of the NPs are taken from the diagonal of the covariance matrix, $\sigma_i^2 = \mathbf{cov}(i,i)$, which is estimated from the inverse of the Hessian matrix, $\mathbf{H}_{i,j}$, around the maximum of the likelihood [106]:

$$\mathbf{cov}(i,j) = \mathbf{H}_{i,j}^{-1} = \left[-\frac{\partial^2 \log \mathcal{L}(\theta)}{\partial\theta_i \partial\theta_j} \right]^{-1} \tag{5.26}$$

The uncertainty of each NP results from the combination of the prior and the measurement from the data. If the fit model is not able to extract information about a specific NP from the data, its post-fit uncertainty is expected to be $\pm 1\,\sigma_{\theta_0}$, where σ_{θ_0} is the prior. If the model is sensitive to this NP, its post-fit uncertainty is reduced.

The floating normalization parameters are special cases since they do not have a prior. Hence, their fit result cannot be given relative to a pre-fit uncertainty. Instead, the result is given in absolute terms, where 1 represents the nominal value from the pre-fit expectation.

The full set of NP pulls from the fit to the data is shown in Figures 5.33, 5.34 and 5.35. They are compared to the pulls from a fit to the Asimov dataset. The parameter names are introduced in Section 5.5. The pulls from the Asimov dataset are all centered around zero as expected, since this dataset corresponds to the nominal MC expectation. The floating parameters from the fit to the data are summarized in Table 5.15.

In the pulls from the data several features can be observed. First, the three $t\bar{t}$ normalization parameters (`norm_ttbar_Lx`) are inconsistent between the lepton channels (Figure 5.33, top left and Table 5.15). This is likely to be caused by selecting quite different phase-space regions for the three channels. The $t\bar{t}$ normalization parameters were correlated in the initial fit model as one parameter. This, however, led to tensions between

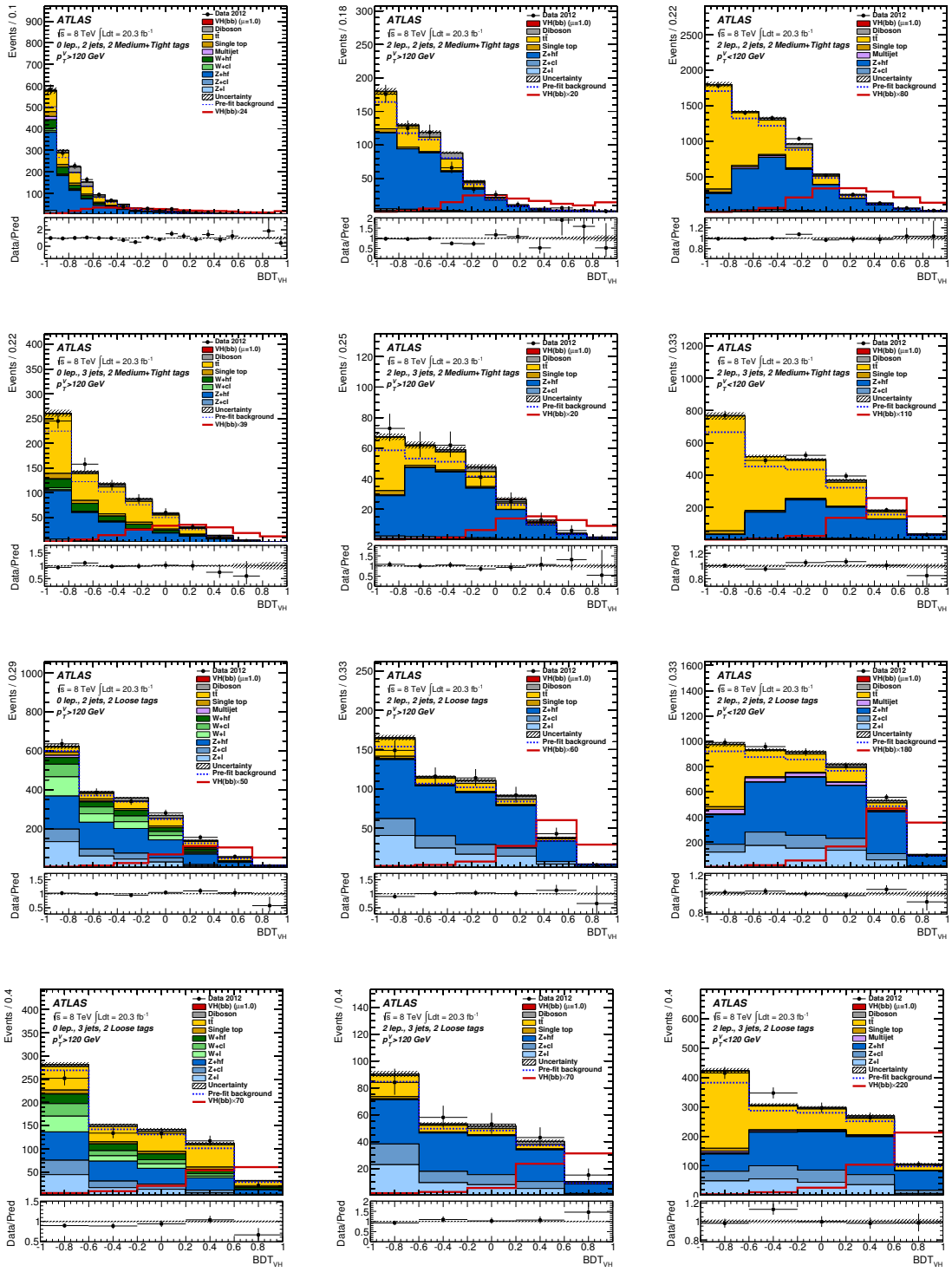


Figure 5.29.: Distributions of the BDT output for the 0-lepton channel (left) and for the 2-lepton channel with high p_{T}^{V} (middle) and low p_{T}^{V} (right) after the combined fit to the data, ordered by decreasing sensitivity. Shown are the regions used in the fit: 2- and 3-jet with MM+TT-tags (two top rows) and LL-tag (two bottom rows). A formal description of the plot elements is given in the text (from Ref. [1]).

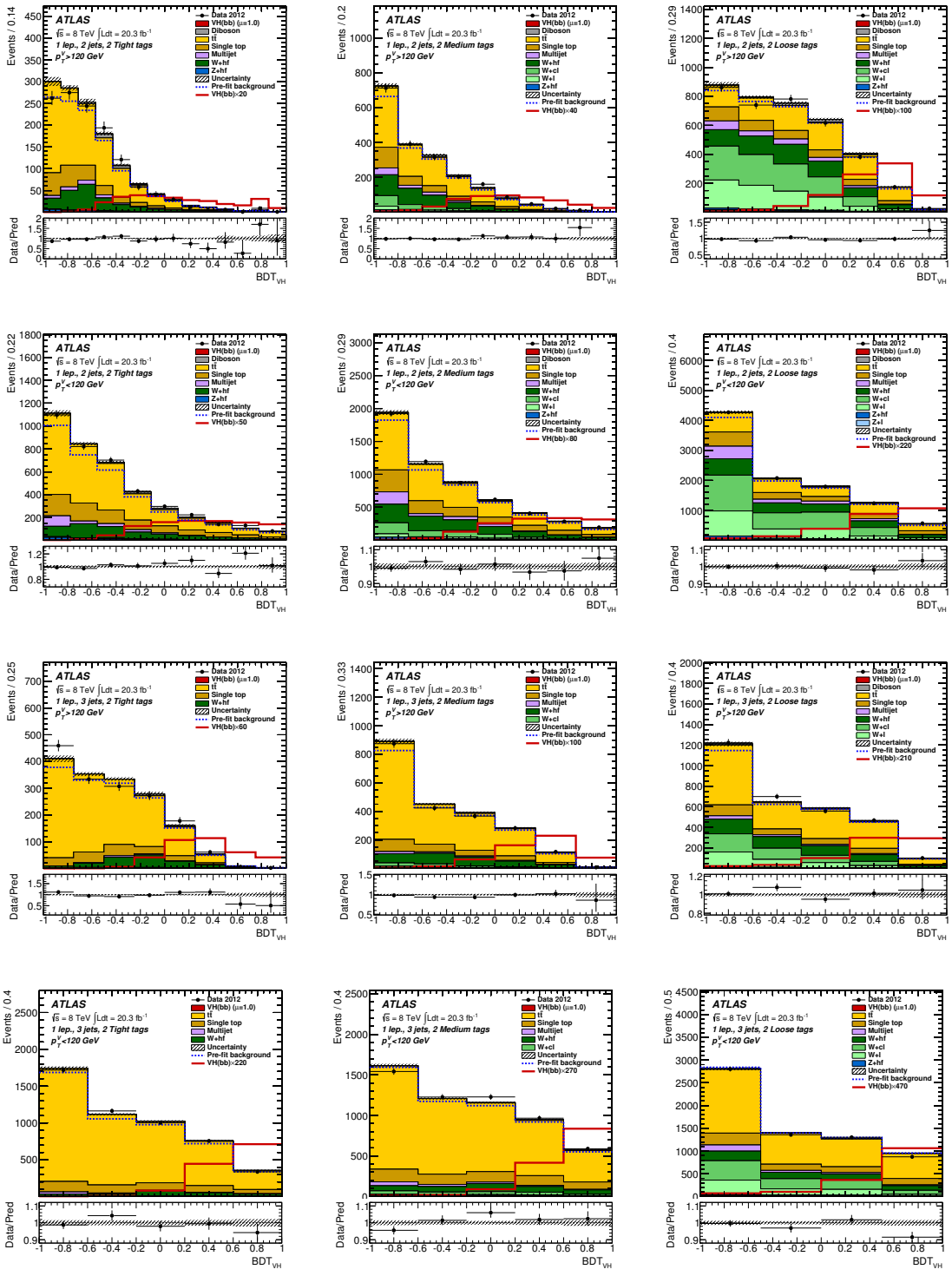


Figure 5.30.: Distributions of the BDT output for the 1-lepton channel after the combined fit to the data, ordered by decreasing sensitivity. Shown are the regions used in the fit: 2-jet (two top rows) and 3-jet (two bottom rows); TT-tag (left), MM-tag (middle) and LL-tag (right) with low and high-p_{T}^{V}. A formal description of the plot elements is given in the text (from Ref. [1]).

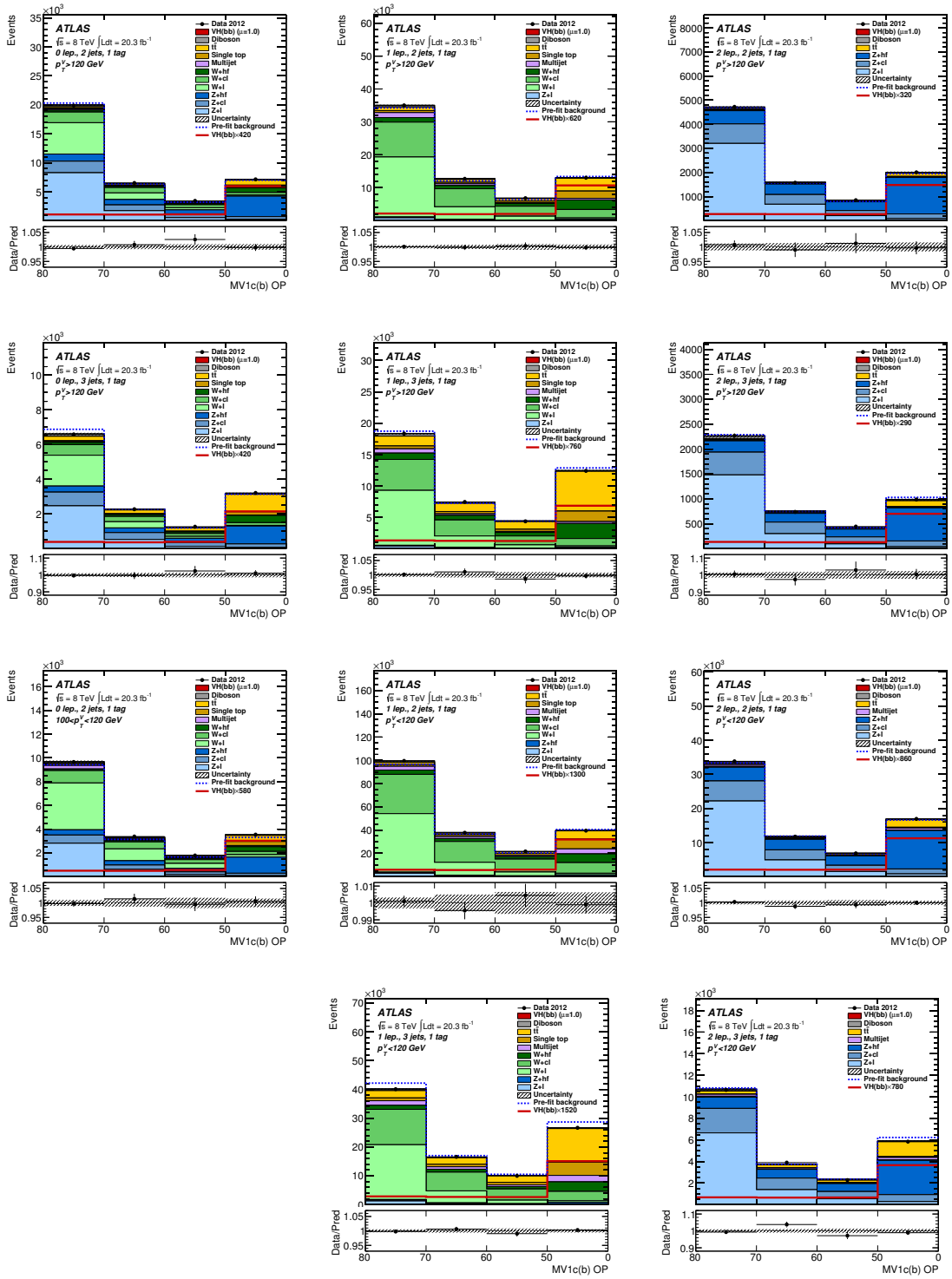


Figure 5.31.: MV1c distributions for the 0 (left), 1 (middle) and 2-lepton (right) channels after the combined fit to the data. Shown are the regions used in the fit: 2- and 3-jet, 1-tag with high p_{T}^{V} (two top rows) and low p_{T}^{V} (two bottom rows). A formal description of the plot elements is given in the text (from Ref. [1]).

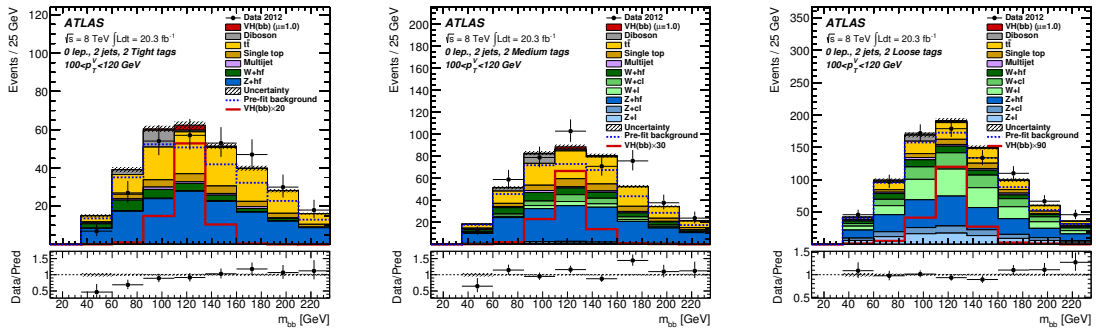


Figure 5.32.: Distributions of m_{jj} for the 0-lepton channel after the fit to the data, ordered by decreasing sensitivity. Shown are the regions used in the fit: 2-jet; TT-tag (left), MM-tag (middle) and LL-tag (right) with medium p_{T}^{V}. A formal description of the plot elements is given in the text (from Ref. [1]).

Process	Scale factor
$t\bar{t}$ (0-lepton)	1.36 ± 0.14
$t\bar{t}$ (1-lepton)	1.12 ± 0.09
$t\bar{t}$ (2-lepton)	0.99 ± 0.04
W+bb	0.83 ± 0.15
W+cl	1.14 ± 0.10
Z+bb	1.09 ± 0.05
Z+cl	0.88 ± 0.12

Table 5.15.: Scale factors for the largest backgrounds from the fit to the data. They are, together with the signal strength, the only floating parameters in the fit.

the combined fit and independent fits of the three channels. The corresponding NP pulls are shown in Appendix A.1. Hence, the decision was taken to treat them as independent parameters.

A similar feature can be observed for some W+jets modeling parameters (Figure 5.33, bottom right). Here, the pull on `WPtV` is inconsistent between the 2- and 3-jet regions. Again, this was not expected and a de-correlation was introduced a-posteriori. Instead, the slightly inconsistent behavior of `WDPhi` between jet regions and flavor components is not surprising. Its correlation scheme was chosen based on the derivation of the corresponding systematic uncertainty (Section 5.5.3).

The issues mentioned above were discovered using a technique that is generic to any analysis. A specific nuisance parameter, which is initially correlated among several regions, is de-correlated for some regions. Then the fit to the data is performed. By comparing the pulls on the de-correlated parameter inconsistencies can be found. This was done for all parameters and all regions of the fit in the present analysis. Examples are shown in Appendix A.1.

Although this technique is quite computing intense, it turned out to be a valuable tool. However, if an inconsistency is found it is not necessarily a good approach to de-correlate

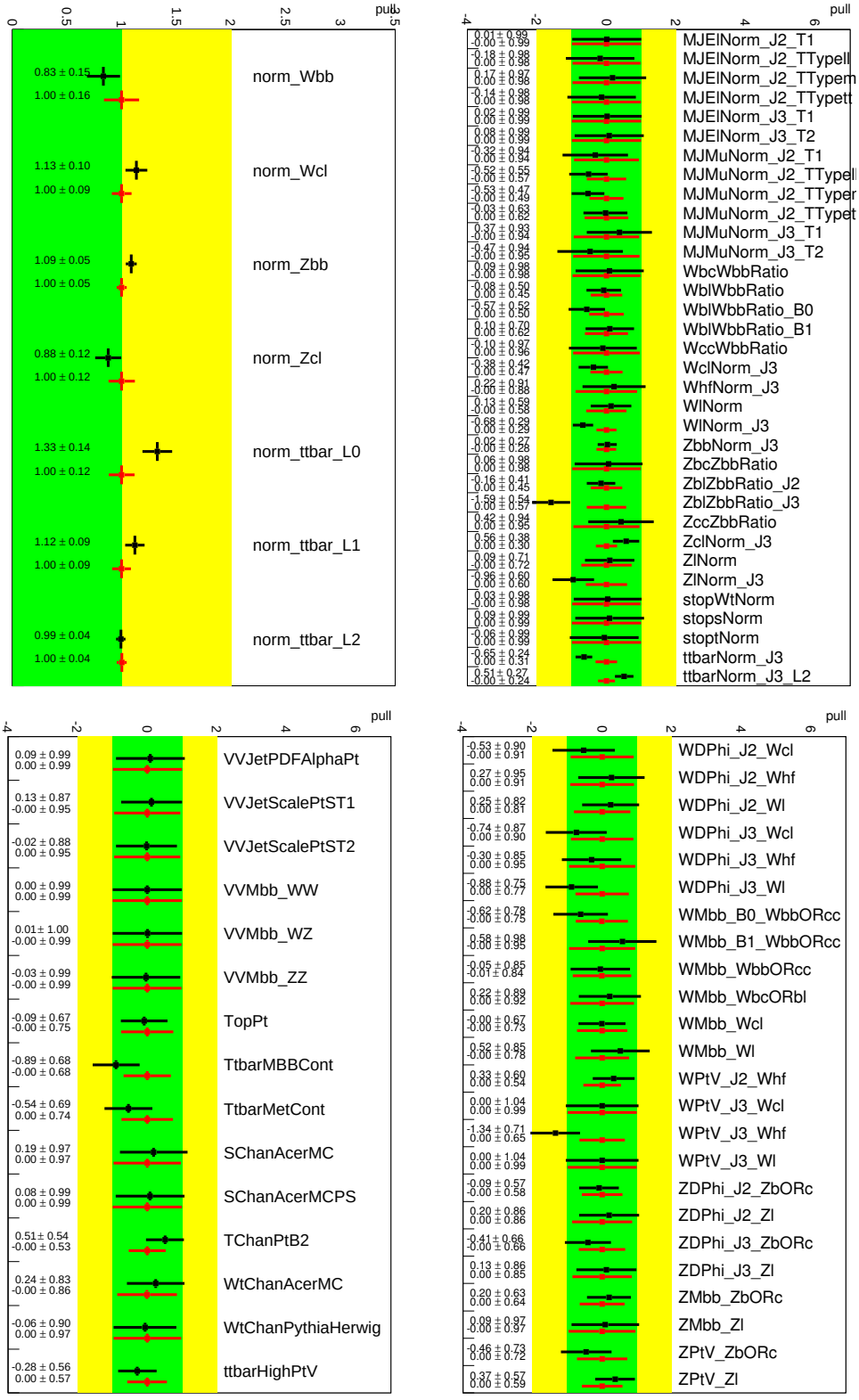


Figure 5.33.: Nuisance parameter pulls from the combined fit to the data (black) and to the Asimov dataset (red). Shown are the NPs for the floating normalization parameters (upper left), the normalization with priors (upper right), for the top quark modeling (lower left) and the V+jets modeling (lower right).

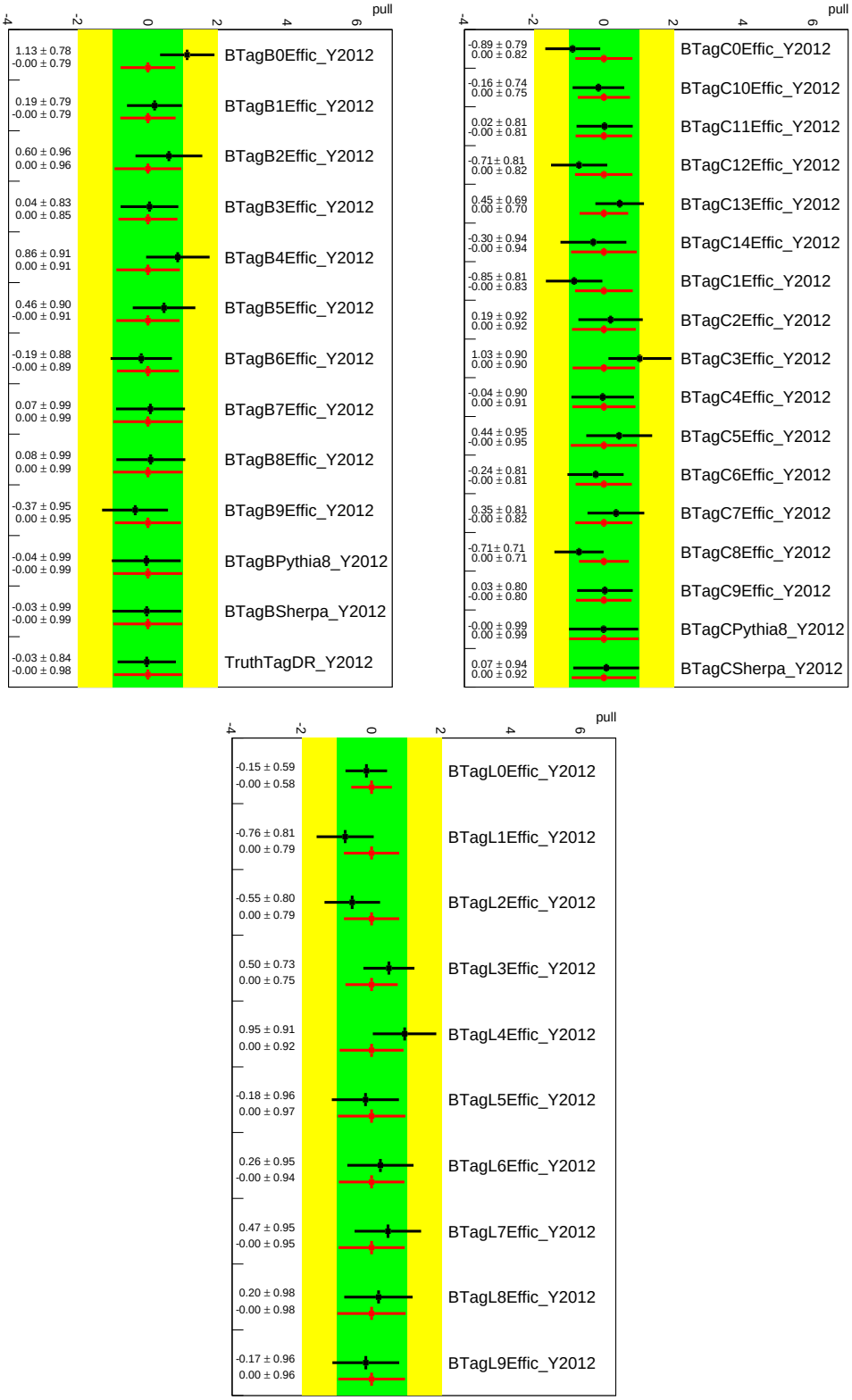


Figure 5.34.: Nuisance parameter pulls from the combined fit to the data (black) and to the Asimov dataset (red). Shown are the b tagging NPs for b jets (upper left), c jets (upper right) and light jets (bottom).

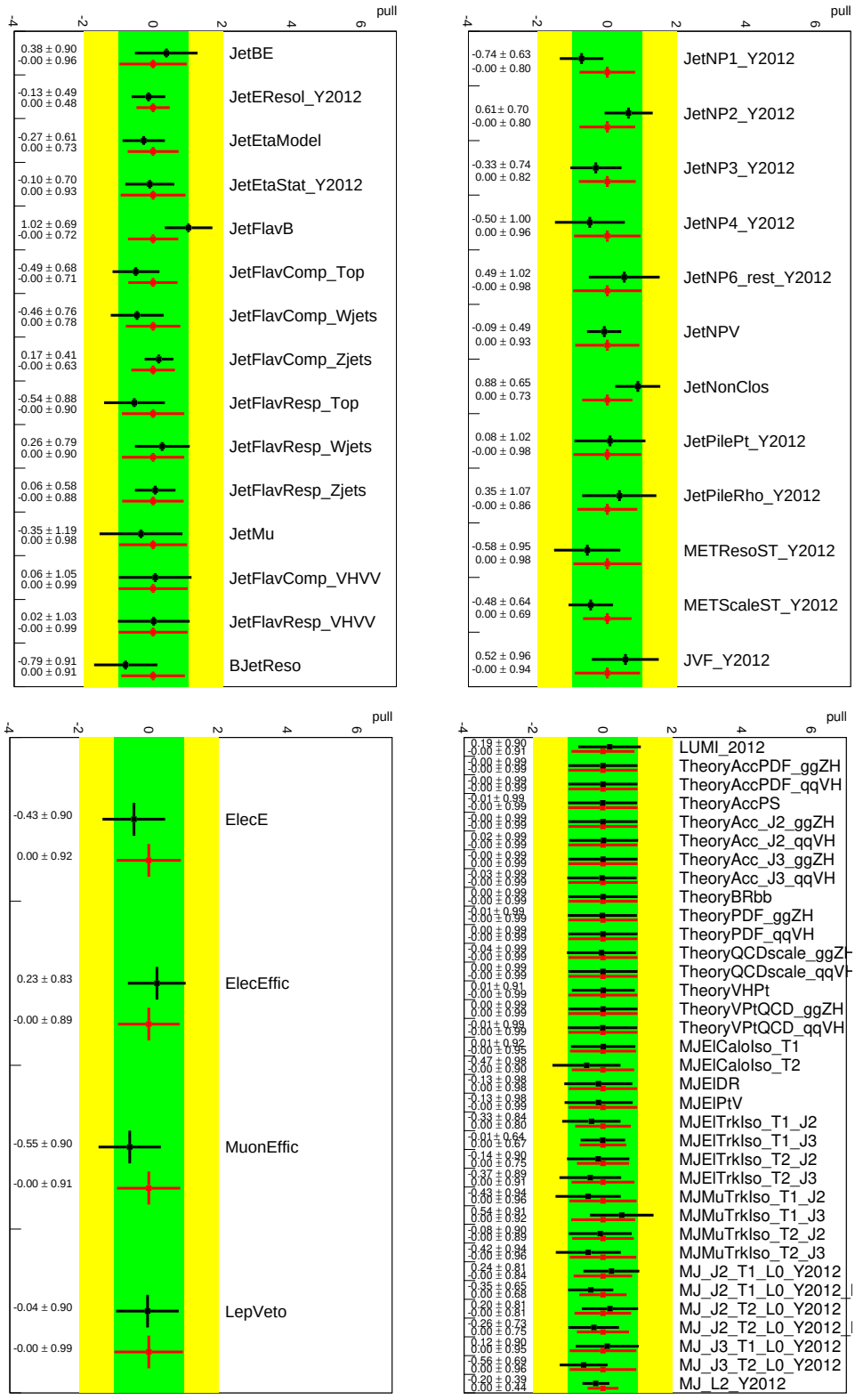


Figure 5.35.: Nuisance parameter pulls from the combined fit to the data (black) and to the Asimov dataset (red). Shown are the jet related NPs (top), the lepton related NPs (lower left) and the remaining NPs (lower right).

the corresponding parameter for the baseline analysis. Instead, a more profound correction might be needed, which has to be judged case by case.

The fit to the Asimov dataset is useful to compare the resulting post-fit uncertainties of the nuisance parameters to the fit to the data. They should be very similar, assuming the total numbers of events in the two datasets are roughly the same. This is the case for the analyses presented in this thesis. Large discrepancies, however, can point to issues in the likelihood. For example, a very noisy and / or one-sided parameterization of a specific NP could cause different post-fit uncertainties in the two cases.

An *over-constraint* is observed for `JetNPV` (Figure 5.35, top right). Its post-fit uncertainty from Asimov is $\pm 0.9\,\sigma$, while it is $\pm 0.5\,\sigma$ for the data. This parameter is affecting the jet kinematics, which causes bin migrations in the MC histograms. In this case, large statistical uncertainties can be present on the differences between the nominal and the systematically varied distributions, as discussed in Section 5.4.3.

As a consequence, bins with asymmetric or even one-sided variations exist. This can, together with the specific inter- and extrapolation technique discussed in Section 5.4.4, cause a strong dependence of the slope in the nuisance parameter response, $\eta(\alpha)$, on its pull, $\alpha \equiv \theta$. Typically, a stronger constraint of the parameter is observed once it is pulled away from the nominal to a value with a steeper slope (e.g. corresponding to the region with $\alpha > 0$ in Figure 5.18, right, blue line). A steeper response means a faster change in the likelihood, hence a narrower maximum and a reduced post-fit uncertainty.

Statistical fluctuations in the differences between the nominal and the systematically varied distributions can not only lead to over-constraints, but also to unreasonable pulls. For example, the `JetEResol` parameter (Figure 5.35, top left) was found very inconsistent between the lepton channels at first. This was traced back to a strong pull in the 2-lepton channel from a single bin of the MV1c distribution in the 3-jet, 1-tag, low p_{T}^{V} region in a minor background. It was regarded as a statistical fluctuation, since no other MV1c distribution or sample showed this behavior. As a consequence, the parameter was removed from this specific region and the consistency between the channels was restored.

A rarer case, the slight increase of a NP uncertainty is observed for `JetMu` (Figure 5.35, top left). Its post-fit value is $\pm 1.0\,\sigma$ for the Asimov dataset, while it is $\pm 1.2\,\sigma$ for the data. Part of the cause can be a similar effect as for over-constraints, just that the parameter is pulled to a value where the response is flatter than for the nominal value (e.g. corresponding to the minimum at $\alpha \approx -0.3$ in Figure 5.18, right, blue line). This could explain a weaker constraint in the fit to the data compared to the Asimov dataset.

However, it cannot explain a post-fit uncertainty larger than the one before the fit. Even if the response is completely flat, the prior should limit the uncertainty to $\pm 1\,\sigma$. Instead, a one-sided effect from the uncertainty can be the cause. In this case, pulls to both sides of the nominal value can lead to a larger value of the measurement likelihood (Equation 5.6), creating two degenerate maxima. This broadens the maximum in the full likelihood (including the prior, Equation 5.10) and leads to the increased post-fit uncertainty, as sketched in Figure 5.36. No significant impact on the signal measurement is expected, as discussed further below.

Moderate pulls up to about $1\,\sigma$ are observed on the b-tagging NPs. This is not unexpected, since the MV1c distributions are directly used as input to the fit. These pulls on the

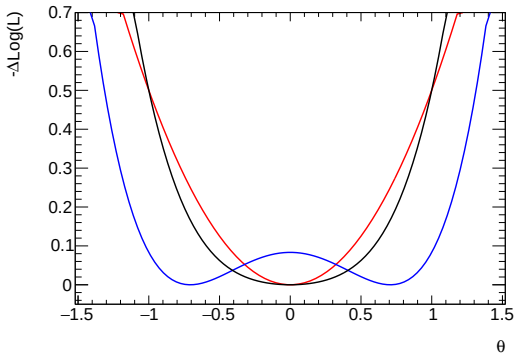


Figure 5.36.: Sketch of the inflation of a systematic uncertainty. Two degenerate minima are present in the negative logarithm of the measurement likelihood (blue). In combination with the prior (red), a broader minimum is found in the full likelihood (black) and the post-fit uncertainty is larger than $\pm 1\,\sigma$.

eigenvector variations are translated back to the b-tagging scale factors in Appendix A.2.

In conclusion, some tensions in the pulls for fits with the initial model are observed and are addressed. The remaining NP pulls from the fit to the data are mostly within $\pm 1\,\sigma$ and do not raise concerns.

5.7.3. Correlations

Another important tool for investigating the fit result is the correlation matrix. However, the full matrix with all parameters is very large and contains many small correlations. A reduced version from the fit to the data is shown in Figure 5.37. Here, only parameters with at least one correlation with any other parameter above 20 % are shown. The full matrix can be found in Appendix A.3. The parameter names are introduced in Section 5.5.

Similar to the NP pulls, the correlations can be compared between the fit to the data and to the Asimov dataset. Both should be very similar, considering only mild pulls occur in the fit to the data[5]. This is the case for the present analysis (and the other analyses presented in this thesis). The correlation matrix from the fit to the Asimov dataset can be found in Appendix A.3.

The largest correlations observed in the fit are listed in Table 5.16. They are discussed in the following with focus on the leading effects, while further contributions from higher order correlations are neglected.

Some correlations are easily understood, since they are generated by design of the fit model. For example, the large negative correlation between `norm_Zbb` and `ZblZbbRatio_J2` is due to the specific parameterization of the flavor composition uncertainties. The `ZblZbbRatio_J2` parameter affects only the Z+bl component (in the 2-jet region). The

[5]The correlations are encoded in the likelihood function. In case strong pulls are observed in a fit to a dataset that deviates from the nominal prediction, the correlations are evaluated in a different point of the parameter space of the likelihood. In this case they can differ to those from a fit to the Asimov dataset.

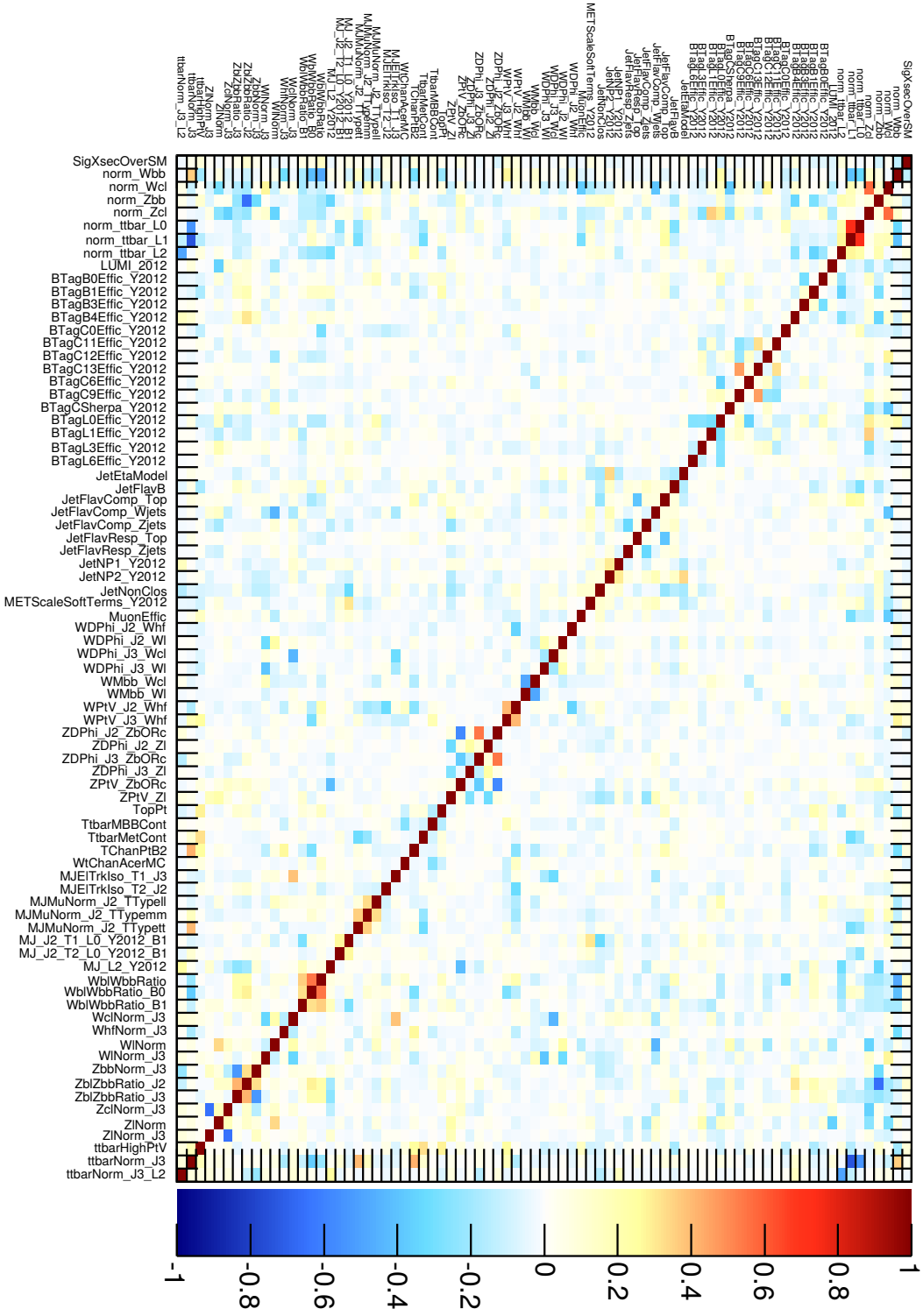


Figure 5.37.: Correlation matrix from the fit to the data. Only parameters with at least one correlation with any other parameter larger than 20 % are shown.

Parameter 1	Parameter 2	Correlation
SigXsecOverSM	norm_Wbb	−15 %
	WMbb_WbbORcc	−15 %
	WblWbbRatio	+15 %
norm_Zbb	ZblZbbRatio_J2	−65 %
norm_Wcl	norm_Zcl	+60 %
norm_ttbar_L0	norm_ttbar_L1	+75 %
norm_ttbar_L0	norm_ttbar_J3	−55 %
	norm_Wbb	−30 %
norm_ttbar_L1	norm_ttbar_J3	−75 %
	norm_Wbb	−45 %

Table 5.16.: Largest correlations observed in the fit to the data for pairs of any two parameters and for the signal-strength parameter, SigXsecOverSM, with any other parameter.

floating normalization parameter, norm_Zbb, however, affects all Z+jets heavy flavor components, including bl. Therefore, the two parameters have a competing effect and the negative correlation is created.

Positive correlations in normalization parameters are usually more complex, since they require at least a third component. The one observed between the floating normalization parameters norm_ttbar_L0 and norm_ttbar_L1 is induced by two effects. A first contribution is from the fact that both parameters have negative correlations with norm_ttbar_J3. This is again by design: The 3-jet parameter affects only the 3-jet region, while the floating parameters are applied to both, the 2- and 3-jet regions. Further, there is a physical component. Both parameters are negatively correlated with norm_Wbb. Both processes, $t\bar{t}$ and W+bb, are major backgrounds in the 0- and 1-lepton signal regions. Since they are not well separated, the negative correlations in the normalization parameters are expected.

The largest correlation for the signal strength parameter, SigXsecOverSM or μ, are observed with norm_Wbb, WMbb_WbbORcc and WblWbbRatio. They are not very large with about 15 %. However, any correlation with μ can possibly affect the sensitivity. In the end, the main point of the analyses is to separate, or de-correlate, the signal from the backgrounds. How severe the correlations with μ are is discussed in the next section.

5.7.4. Ranking

The *nuisance parameter ranking* is an important tool to estimate the impact of individual NPs on the sensitivity. The ranking is compiled by performing scans of the likelihood function. First, the nominal fit is performed to find the global maximum in the likelihood and the corresponding $\hat{\mu}$. Then, the likelihood is scanned as a function of one specific nuisance parameter, while all other parameters are re-fitted. This N-dimensional scan is described in more detail in Section 5.7.6.

The scan stops once the logarithm of the likelihood decreases by 1/2 with respect to the global maximum. This corresponds to the $\pm 1\,\sigma$ uncertainty interval. At this point the change of the signal strength, $\Delta\hat{\mu}$, is evaluated. This procedure is done in both directions

of the NP and is repeated for all other NPs of the fit.

Once all parameters have been scanned, they are ordered by decreasing $\Delta\hat{\mu}$. The result is shown in Figure 5.38 for the parameters with the largest impact on $\hat{\mu}$.

It is apparent from the ranking that the W+jets background plays an important role in the signal measurement. The top-four parameters are all related to this background, in particular the heavy flavor components. The floating W+bb normalization, which shows the largest correlation with μ in the reduced correlation matrix (Figure 5.37), turned up at the third rank. Varying this parameter by $\pm 1\,\sigma$ changes the signal strength by $\Delta\hat{\mu} \simeq 0.05$.

The nuisance parameter with the largest impact on the signal strength of $\Delta\hat{\mu} \simeq 0.06$ is `WMbb_WbbORcc`. This NP parameterizes the uncertainty on the m_{jj} shape for the W+bb and W+cc background in the high p_{T}^{V} region. It does not show up in the reduced correlation matrix, but a correlation of about $-15\,\%$ with μ is observed in the full matrix (Figure A.8).

The experimental parameters with the largest impact can be found on rank seven and nine: `BJetReso` and `JetEResol`. This is not surprising, since they parameterize the jet energy resolution and therefore affect the resolution of the dijet mass. The large impact of `JetEResol` shows that the issue with the statistical fluctuation in a minor background, mentioned in Section 5.7.2, was indeed severe. A bias from an otherwise insignificant control region could have occurred if this issue would have went unnoticed.

Further, it can be noted that the parameters `JetMu` and `JetNPV` do not show up in the ranking plot. They can be found at rank 90 and 103, respectively, with $\Delta\hat{\mu} \lesssim 0.01$. Hence, their slightly problematic pulls, discussed in Section 5.7.2, are not expected to impact the signal measurement.

The nuisance parameter ranking is easier to read than the correlation matrix, but it has a caveat. Parameters that have an impact on the signal might not show up in the ranking due to strong correlations with other parameters. For example, two NPs that are completely degenerate, i.e. have a correlation of $\pm 100\,\%$, can compensate each other in the likelihood scan. The $\Delta\hat{\mu}$ would be zero. Still, they can have an impact on the signal measurement if they have a correlation with μ.

5.7.5. Uncertainty breakdown

Another approach for looking at the most important parameters in the fit is the *uncertainty breakdown*. Here, the nuisance parameters are grouped by similar sources of uncertainties. For example, the W+jets modeling NPs build one group.

The contributions of such groups to the signal strength uncertainty are estimated by first performing the nominal fit and extracting $\sigma_{\hat{\mu}}$. Another fit is done, where a specific group of NPs, g, is fixed to their values of the nominal fit. Thus, they do not contribute to the corresponding uncertainty of the signal strength, $\sigma'_{\hat{\mu}}$. The contribution of the group to the nominal signal strength uncertainty is given as

$$\sigma_{\hat{\mu}}^{\mathrm{g}} = \sqrt{(\sigma_{\hat{\mu}})^2 - (\sigma'_{\hat{\mu}})^2}. \tag{5.27}$$

The uncertainties of the signal strength are estimated from scans of the likelihood, as described in the next section. Although the scans give asymmetric uncertainties, they are symmetrized in the following for better readability.

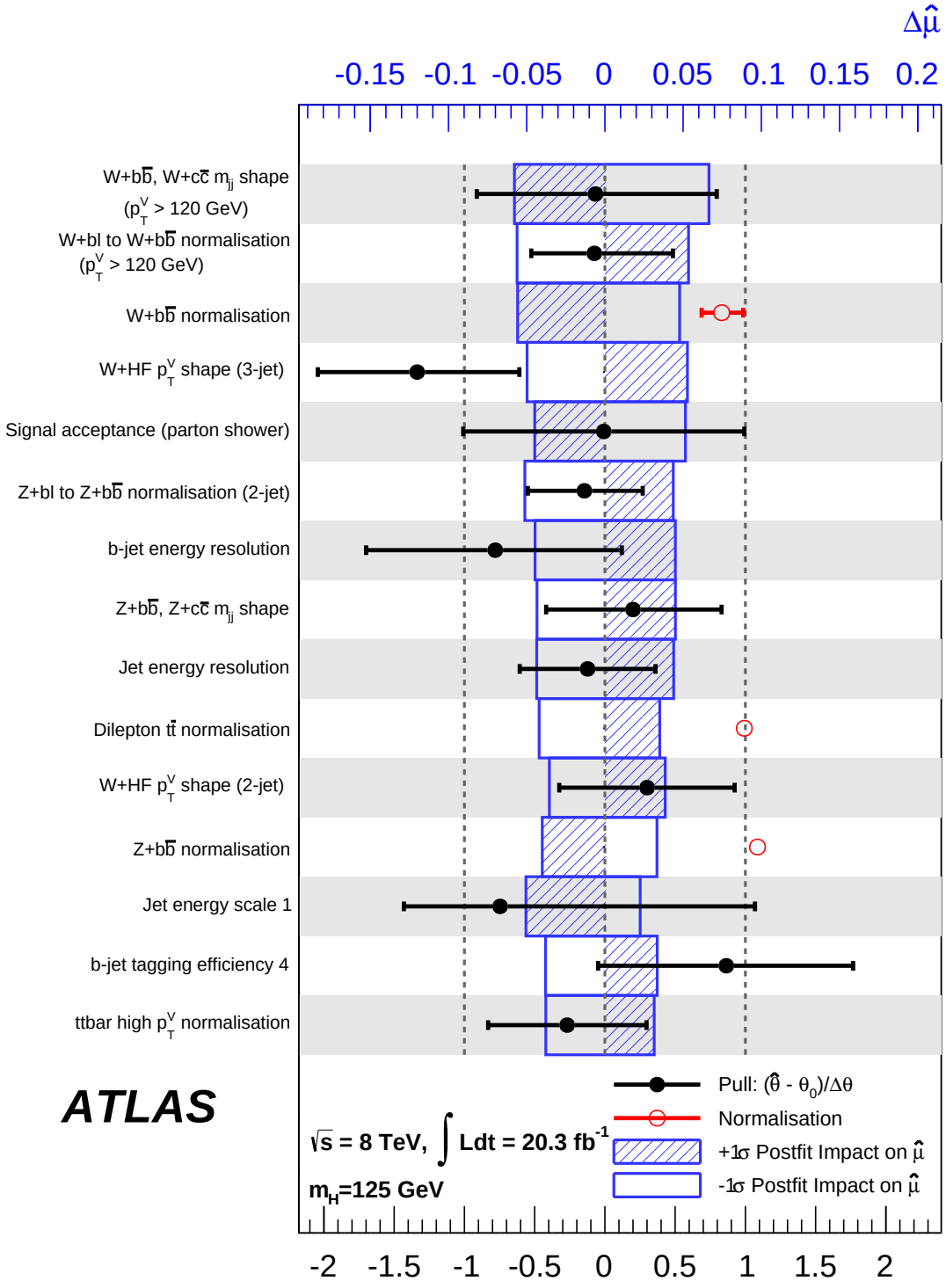


Figure 5.38.: Ranking of the post-fit impact from nuisance-parameter uncertainties on the observed signal strength $\hat{\mu}$ (blue areas). The pulls of the parameters from the fit to the data are shown as well (markers). (from Ref. [1]).

Experimental	
Source of unc.	$\sigma_{\hat{\mu}}$
Jets	0.08
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.03
Leptons	0.01
b tag: b jets	0.07
c jets	0.04
light jets	0.04
Luminosity	0.03

Theory and modeling	
Source of unc.	$\sigma_{\hat{\mu}}$
W+jets	0.11
Z+jets	0.08
$t\bar{t}$	0.05
Single-top	0.04
Diboson	0.02
Multijet	0.06
Signal	0.07

Floating norm.	
Source of unc.	$\sigma_{\hat{\mu}}$
W+jets	0.06
Z+jets	0.03
$t\bar{t}$	0.04

Combined	
Systematic	0.26
Statistical	0.32
Total	0.41

Table 5.17.: Breakdown of $\sigma_{\hat{\mu}}$ into groups of uncertainty sources in the fit to the data. Various systematic components are given, as well as the total statistical and total uncertainties.

The results are shown in Table 5.17. Similarly, the combined impact from all systematic sources, $\sigma_{\hat{\mu}}^{\mathrm{sys}}$, is estimated using $\sigma'_{\hat{\mu}} = \sigma_{\hat{\mu}}^{\mathrm{stat}}$ from a fit where all nuisance parameters are fixed. This total systematic uncertainty includes a component arising from the limited number of events in the background prediction (Equation 5.9). This component is not listed separately, as it amounts to less than 1 %.

The largest contribution to the total uncertainty on $\hat{\mu}$ is the statistical uncertainty arising from the limited number of events in the data. However, the combined systematic uncertainty is not much smaller. Therefore, if the measurement shall be improved in the future using more data, work has to be done to reduce the systematic uncertainties as well.

In general, the sensitivity of this analysis, opposed to others, is affected by many different sources of uncertainties. Hence, work has to be done on various aspects, on the experimental and on the theoretical side. In particular the W+jets modeling, which is the largest contribution, has to be improved. Here, an upgrade of the LO Sherpa 1.4.1 generator to an NLO generator would be desirable.

However, the $t\bar{t}$ modeling should be of interest as well. In particular, since the cross section for this background rises stronger with increasing center-of-mass energy than for any other process considered in the analysis, as shown in Figure 2.4.

5.7.6. Likelihood scans

A direct look at the likelihood around the minimum is provided by *likelihood scans*. The one-dimensional scan is the simplest one. Here, the likelihood is evaluated as a function of one parameter, θ_i, while all other parameters are fixed to the result of the minimization:

$$\mathcal{L}_1(\theta_i) = \mathcal{L}(\hat{\mu}, \hat{\theta}_1, ..., \theta_i, ..., \hat{\theta}_N) \tag{5.28}$$

The result is given for convenience as the negative logarithm of the likelihood, subtracted by the global maximum,

$$-\Delta \log \mathcal{L}_1(\theta_i) = -\left(\log \mathcal{L}_1(\theta_i) - \log \mathcal{L}(\hat{\mu}, \hat{\theta})\right), \tag{5.29}$$

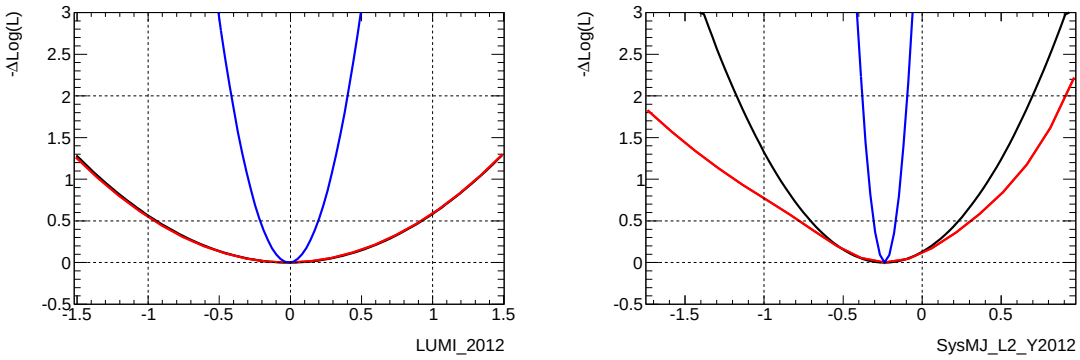


Figure 5.39.: Examples for scans of the likelihood along the LUMI and MJ_L2 parameters in an earlier stage of the analysis for the 2-lepton channel in a fit to the data. Black: parabola from Hesse-matrix inversion, blue: one-dimensional scan, red: full scan.

so that $-\Delta \log \mathcal{L}_1(0) = 0$ is the global minimum.

Examples for one-dimensional scans are shown as blue curves in Figure 5.39 for the LUMI and MJ_L2 parameters. The parabolic shapes indicate Gaussian uncertainties from the combination of the priors and the measurements from the data.

Each scan is compared to a simple parabola, $-\Delta \log \mathcal{L}_p$ (black curve), whose minimum is at $\theta_i = \hat{\theta}_i$ and whose width at $-\Delta \log \mathcal{L}_p = 1/2$ is set to two times the post-fit uncertainty,

$$-\Delta \log \mathcal{L}_p(\theta_i) = \frac{1}{2} \left(\frac{\theta_i - \hat{\theta}_i}{\sigma_{\theta_i}} \right)^2 , \tag{5.30}$$

where σ_{θ_i} is estimated from the inverted Hessian matrix, as described in Section 5.7.2.

The proper N-dimensional scan is done by evaluating the likelihood as a function of θ_i, while all other parameters are re-adjusted to find the corresponding local minimum:

$$\mathcal{L}_N(\theta_i) = \mathcal{L}(\hat{\hat{\mu}}, \hat{\hat{\theta}}_1, ..., \theta_i, ..., \hat{\hat{\theta}}_N) \tag{5.31}$$

The results, again as negative logarithms and subtracted by the global minima, are shown as red curves in the same Figure.

The example scans show some specific features. First, the LUMI parameter, shown in Figure 5.39 (left), is very well behaved. The N-dimensional scan agrees perfectly with the parabola from the Hesse matrix inversion. Hence, the latter gives is a good approximation of the uncertainty.

The one-dimensional scan gives a narrower curve. This is the case for NPs that have correlations with other parameters, which can compensate for the change of θ_i during the N-dimensional scan. In the case of LUMI, the compensating parameters are mostly background normalization parameters, as seen from the correlation matrix in Figure 5.37. Many parameters of the fit are as well behaved as the LUMI parameter.

Some other parameters show asymmetric responses, for example the MJ_L2 parameter, shown in Figure 5.39 (right). In this case, the behavior is well understood. The MJ_L2

is the normalization scale parameter for the multijet background in the 2-lepton channel, which has a very large prior of 100 %, as listed in Table 5.12.

For normalization parameters a polynomial interpolation and exponential extrapolation function is used, as discussed in Section 5.4.4. The exponential extrapolation prevents the yield of the background to become negative. Hence, the response of the likelihood flattens for large negative pulls on θ_i, which is seen in the plot. Also the slight disagreement for positive values of the `MJ_L2` parameter is explained by the changing slope in the interpolation.

For such asymmetric responses the uncertainty from the Hesse matrix inversion can by definition not be a good estimate, since it is symmetric. Instead the intersect of $-\Delta \log \mathcal{L}_N = 1/2$ from the N-dimensional scan can be used to define asymmetric uncertainties. This is done for the signal-strength parameter, as listed in Section 5.8.2.

More alarming examples of likelihood scans are shown in Figure 5.40. These scans were derived in an earlier stage of the analysis and do not reflect the final result. The `JetEResol` parameter shows a pathological behavior in certain configurations of the fit. For the 0-lepton channel alone good behavior is observed for the fit to the data as well as in the fit to the Asimov dataset, shown in Figure 5.40 (top left and top right, respectively). This is true for the 2-lepton channel as well, although a strong pull is observed. This pull is artificially caused by one bin, as discussed in Section 5.7.2.

However, the scan of `JetEResol` for the 1-lepton channel reveals fluctuations in the likelihood. These can be oscillations between two local minima. Even more severe, the fit has not found the global minimum in the likelihood. The nominal result gives a value of about -0.55 for `JetEResol`, while a better minimum is visible around -0.35. The oscillations are amplified in the combined fit with all three channels and the global minimum is not found here as well.

These issues were resolved by a fix in the symmetrization of the `JetEResol` parameter, together with the optimization of the binning, as described in Section 5.6, and by introducing the smoothing and pruning techniques, as described in Section 5.4.3. All these changes aim to alleviate statistical fluctuations and one-sided effects from experimental systematic uncertainties and have improved the fit stability significantly.

The final fit does not show oscillations in likelihood scans and only one minimum is found. A scan of the signal-strength parameter in the combined fit is shown in Figure 5.41 (upper left).

The response of other parameters during the N-dimensional scan can be evaluated as well. Three examples are shown in the same figure. The `JetEResol` parameter shows a simple linear correlation with the signal strength. The `WMbb_B1_WbbORcc`, which is the highest-ranked NP (Section 5.7.4), shows a similarly simple behavior. Only on the edges of the scan some artifacts occur. They are well beyond the $1\,\sigma$ interval of the signal-strength parameter and therefore do not affect the measurement.

Some parameters, such as `MJElCaloIso_T1`, show more complex behavior. Clearly, the linear correlation coefficients, given in Section 5.7.3, are not sufficient to describe such effects. However, this does not raise concerns as it occurs only for parameters that are loosely correlated with the signal strength parameter.

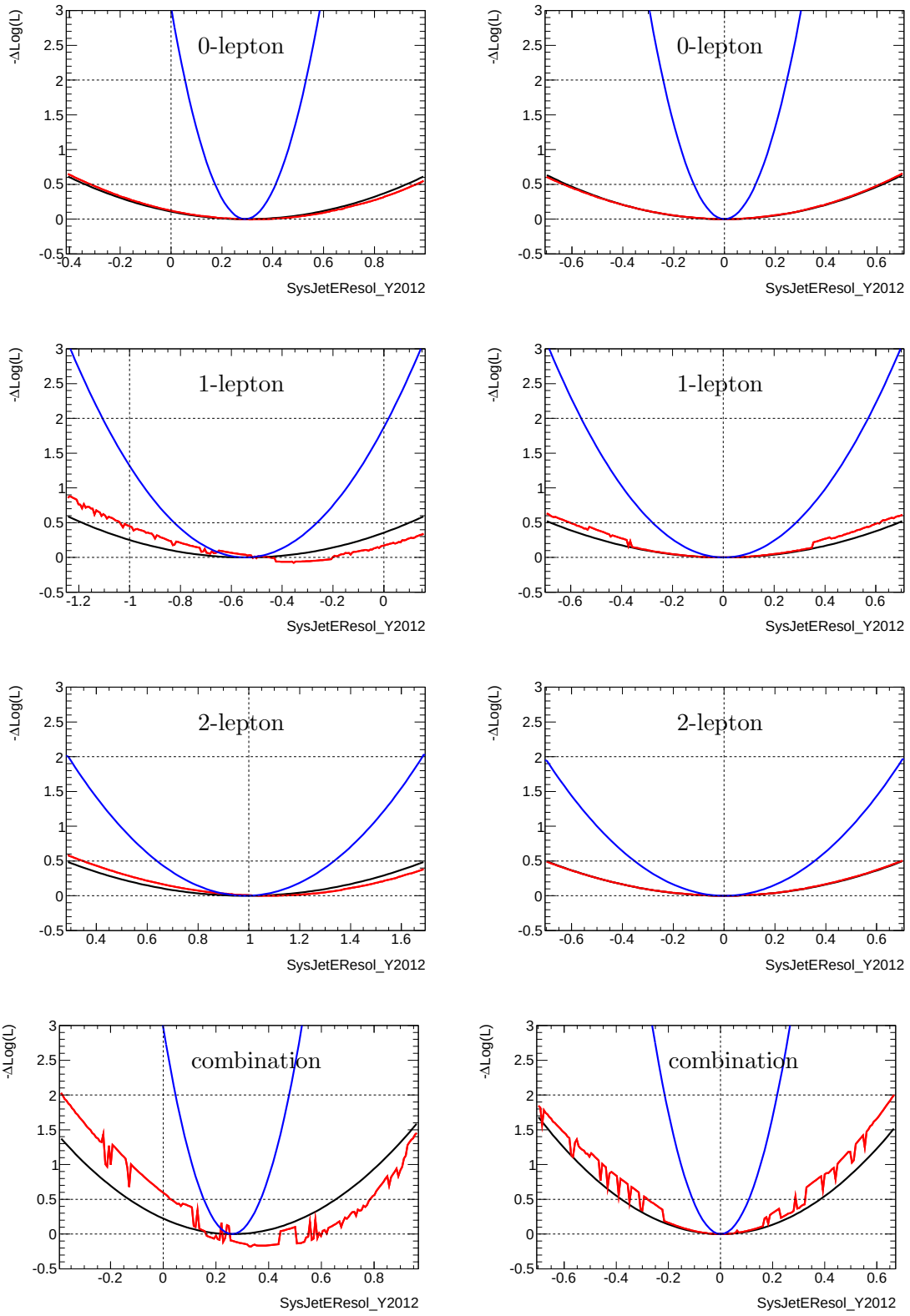


Figure 5.40.: Examples for scans of the likelihood along the `JetEResol` parameter in an earlier stage of the analysis for the (from top to bottom) 0-, 1- and 2-lepton channels and the combination. Fits to the data are shown on the left and to Asimov datasets on the right. Black: parabola from Hesse-matrix inversion, blue: one-dimensional scan, red: full scan.

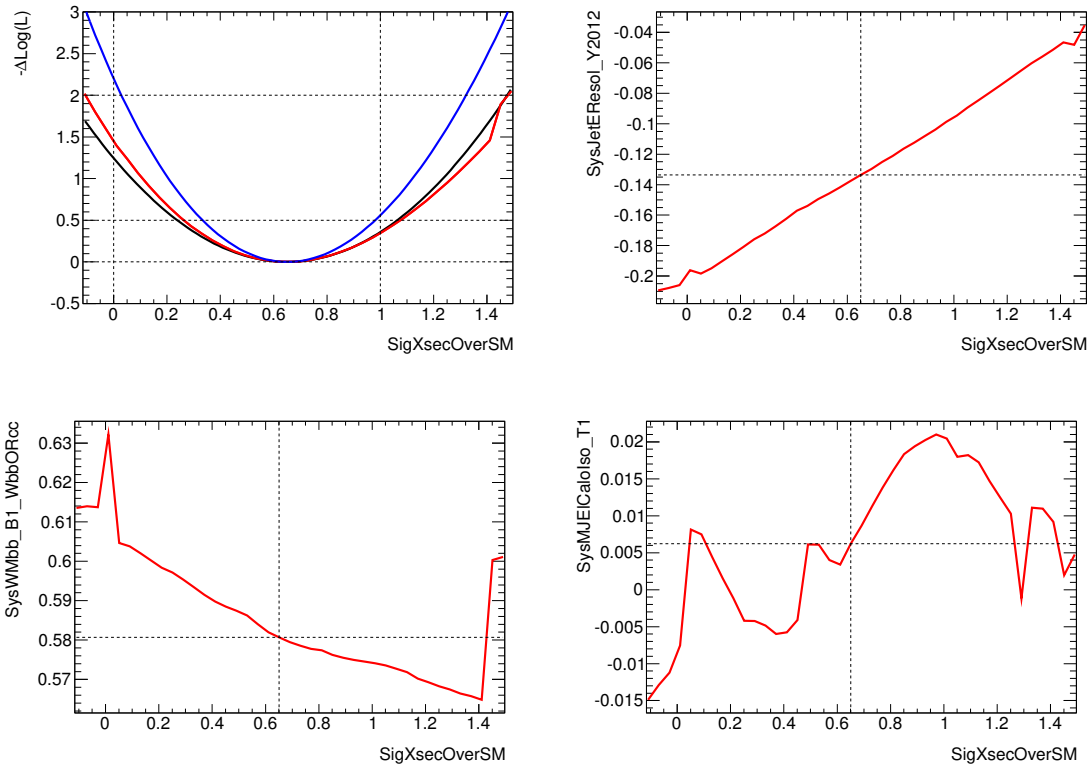


Figure 5.41.: Scans of the likelihood along the signal-strength parameter, `SigXsecOverSM`, for the combined fit to the data (upper left). Black: parabola from Hesse-matrix inversion, blue: one-dimensional scan, red: full scan. The responses of the `JetEResol`, `WMbb_B1_WbbORcc` and `MJElCaloIso_T1` parameters during the scan are shown as well.

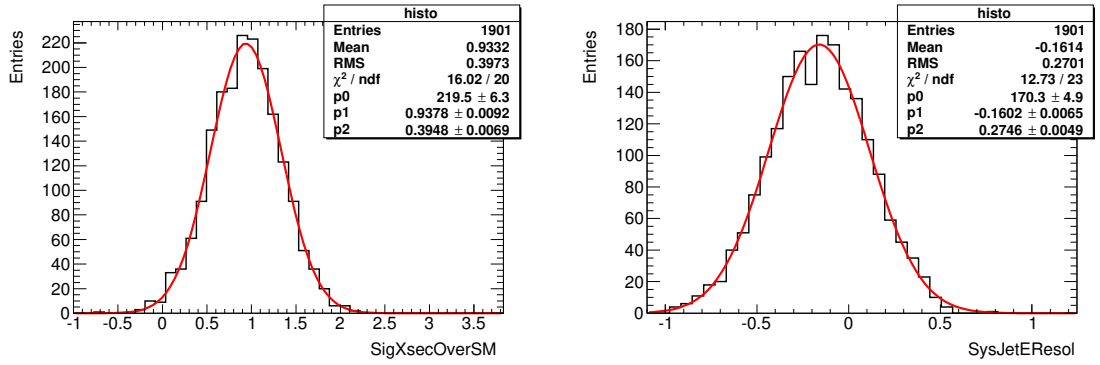


Figure 5.42.: Examples for the result of 1901 toy experiments for the `SigXsecOverSM` and `JetEResol` parameters in the combined fit in an earlier state of the analysis. Gaussian fits (red) are used to estimate the mean and its standard deviation.

5.7.7. Toy experiments

Another tool to investigate the fit model are *toy experiments*. Each experiment consists of generating ("throwing") a toy dataset from the the likelihood function, $\mathcal{L}$, using the Monte-Carlo method [109]. The number of events in each bin, N_i of Equation 5.6, are randomized according to their Poissonian priors. Similarly, the auxiliary measurements, m_i of Equation 5.7, are randomized as well. A fit is performed to each toy dataset to extract the results in terms of the fit parameters.

Examples for the `SigXsecOverSM` and `JetEResol` parameters are shown in Figure 5.42. These studies were done in an earlier state of the analysis and do not reflect the final result. The signal-strength parameter shows a bias of the expected fit result. Although the datasets were generated with an average of $\overline{\mu} = 1$, the result from the fits is $\overline{\mu} = 0.94 \pm 0.01$, where the statistical uncertainty is estimated using a Gaussian fit (red lines in Figure 5.42) to the distribution obtained from the fits to the toy datasets.

Further investigation reveals the strongest bias for the `JetEResol` parameter, which is on average $\overline{\theta} = -0.160 \pm 0.006$, with zero being the expected value. This issue was resolved by the same changes described in the previous section, the fix in the symmetrization of `JetEResol` and by the smoothing and pruning techniques. A residual bias on the signal strength is observed: $\overline{\mu} = 0.97 \pm 0.01$. However, considering of the expected uncertainty on $\hat{\mu}$ of about 0.4, this is deemed small enough for a reliable measurement.

The toy experiments can also be used to calculate the test statistic and to extract limits on the signal strength. However, as described in Section 5.4.1, asymptotic formulae are used for a computationally faster calculation.

5.7.8. Compatibility of regions

An important test of the the fit model is the compatibility of the signal strength, $\hat{\mu}$, between regions. It allows to discover potential inconsistencies for different phase spaces.

This test can be done *blinded*, i.e. without looking at the actual value of $\hat{\mu}$[6].

This is accomplished by first evaluating the maximum likelihood, $\mathcal{L}(\hat{\mu}, \hat{\theta})$, in the combined fit to the data. Then, $\hat{\mu}$ is de-correlated for a chosen set of N regions into $\hat{\mu}_1, ..., \hat{\mu}_N$ and $\mathcal{L}(\hat{\mu}_1, ..., \hat{\mu}_N, \hat{\theta})$ is evaluated. The difference between the logarithms of the two likelihood values is expected to follow the χ^2 distribution,

$$\chi^2 = \log\left(\mathcal{L}(\hat{\mu}, \hat{\theta})\right) - \log\left(\mathcal{L}(\hat{\mu}_1, ..., \hat{\mu}_N, \hat{\theta})\right), \tag{5.32}$$

with $k = N - 1$ degrees of freedom. The compatibility is then given by the cumulative probability density function of the χ^2 distribution,

$$P(\chi^2, k) = \frac{1}{\Gamma(\frac{k}{2})2^{k/2}} \int_0^{\chi^2} x^{k/2-1} e^{-x/2} dx, \tag{5.33}$$

where $\Gamma(t)$ is the gamma function. The values of P are expected to be distributed uniformly between zero and one. The implementation of P is given by `TMath::Prob(Double_t chi2, Int_t ndf)` in the ROOT package [120].

The compatibility between the three lepton channels in the combined fits was initially observed to be quite low with 4.3 %. However, when removing all low p_{T}^V regions from the fit, the compatibility was increased to 40 %. This was further investigated and modeling issues in the 1-lepton, low p_{T}^V region for the electron channel were found, while the muon channel behaved well. The compatibility between the low and high p_{T}^V regions for the electron (muon) channel was estimated to be 0.005 % (66 %).

It was concluded that the multijet background was not properly modeled for the 1-lepton, electron channel, low p_{T}^V region and those events were removed from the fit. The compatibility of the lepton channels in the final combined fit is 51 %.

5.8. Results

The results from the fit to the data are presented in the following. The post-fit plots of the most sensitive distributions are shown in Section 5.8.1, the numerical results of the fit are given in Sections 5.8.2 and 5.8.3, and summary distributions are given in Section 5.8.4.

5.8.1. Most sensitive distributions

The BDT outputs for the three lepton channels in the 2-jet, MM and TT-tag, high p_{T}^V regions are the most sensitive distributions in the analysis. The corresponding 3-jet or low p_{T}^V regions are the next most sensitive distributions. All of them are shown after the fit to the data in Figures 5.43 (0- and 2-lepton) and 5.44 (1-lepton) in logarithmic scale. The corresponding linear-scale plots can be found in Section 5.7.1 and the event yields in Appendix A.4.

[6] Blinding is an important technique to obtain results without bias from human intervention during the development of the analysis. For the present analysis, the data in the sensitive regions of the BDT and m_{jj} distributions have been removed in comparisons to the simulation. Fits have been performed to the full distribution, but the result for signal-strength parameter has been removed from the output.

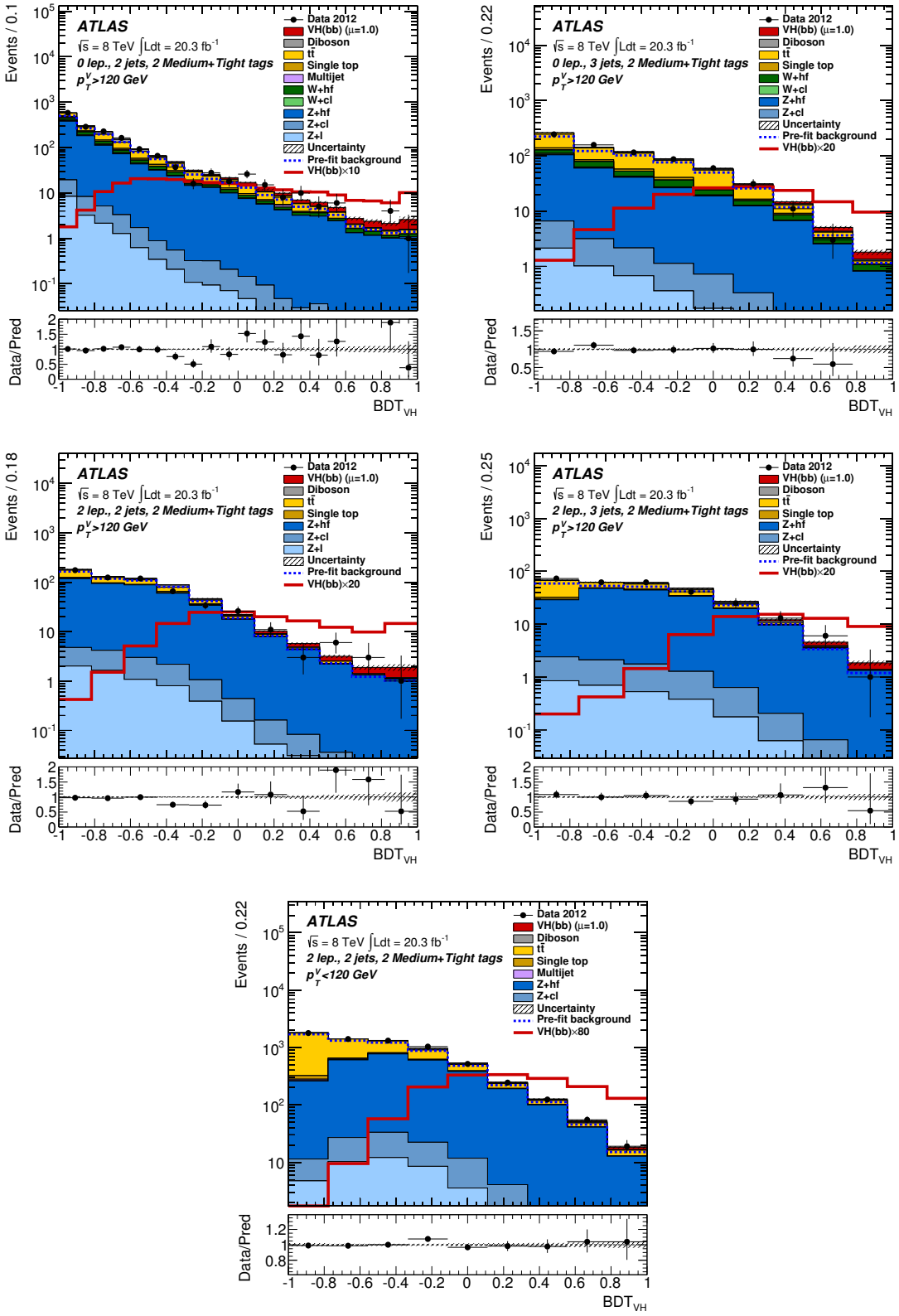


Figure 5.43.: Most sensitive BDT distributions for the 0-lepton (top) and 2-lepton (middle and bottom) channels after the fit to the data in logarithmic scale. A formal description of the plot elements is given in Section 5.7.1 (from Ref. [1]).

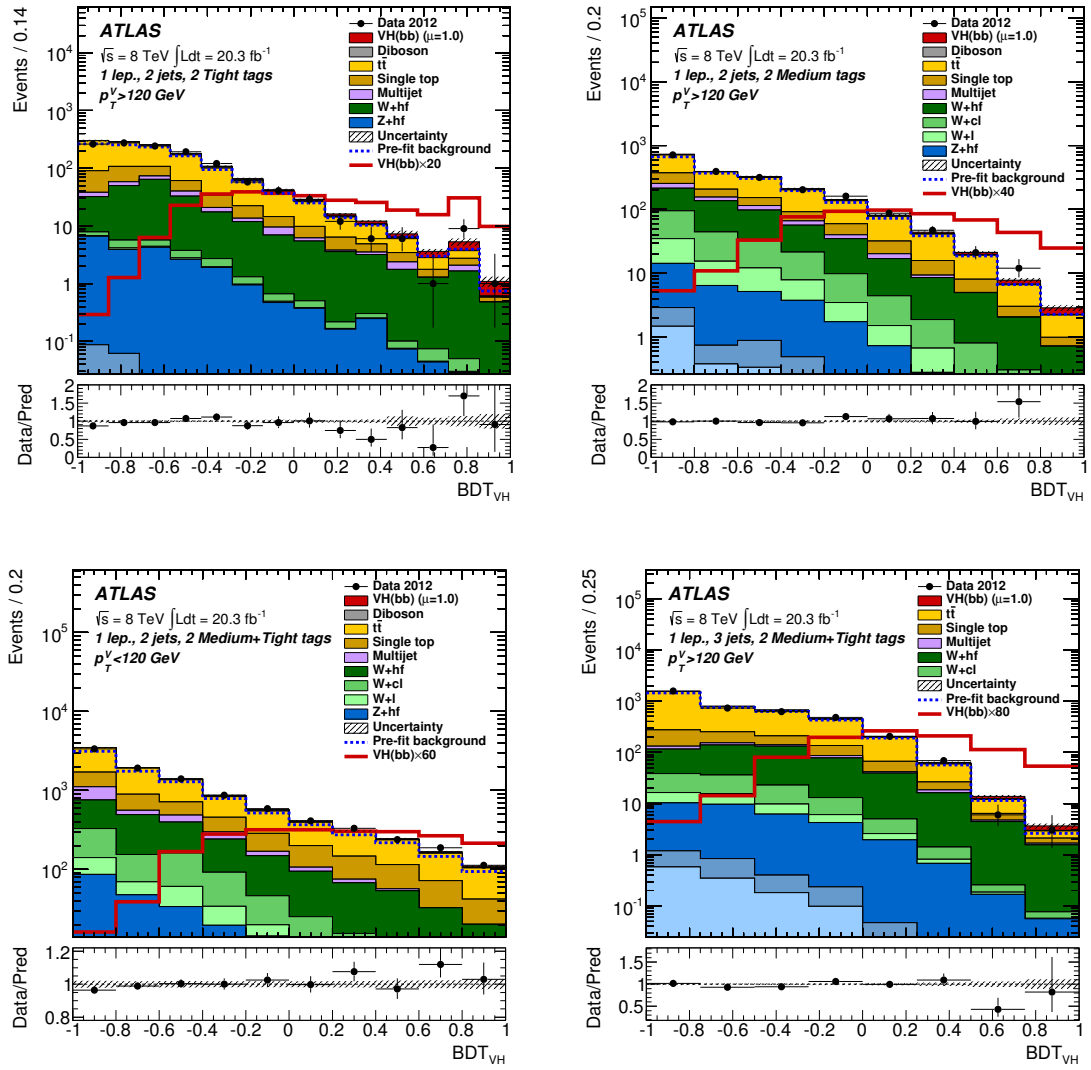


Figure 5.44.: Most sensitive BDT distributions for the 1-lepton channel after the fit to the data in logarithmic scale. A formal description of the plot elements is given in Section 5.7.1 (from Ref. [1]).

Due to the use of transformation F, the number of bins in each distribution is positively correlated with its sensitivity to the signal, as described in Section 5.6.2.

The signal is drawn on top of the backgrounds with a signal strength of $\mu = 1$. It is apparent that the data in the 1- and 2-lepton channels is preferring the background plus signal hypothesis, while some deficit is visible in the 0-lepton distributions.

5.8.2. Signal strength

The observed signal strength for the SM Higgs boson with $m_H = 125\,\text{GeV}$ in the fit of all three lepton channels to the data taken at $\sqrt{s} = 8\,\text{TeV}$ is

$$\hat{\mu}_{8\,\text{TeV}} = \sigma/\sigma_{\text{SM}} = 0.65^{+0.43}_{-0.40}$$

This is combined with the corresponding result from the analysis performed with the data taken at $\sqrt{s} = 7\,\text{TeV}$ [121], which is $\hat{\mu} = -1.61^{+1.50}_{-1.46}$. The combination is performed by multiplying the likelihoods of the two analyses, leaving all parameters uncorrelated, expect for the signal-strength parameters. The result from maximizing the likelihood for the $7 + 8\,\text{TeV}$ combination is

$$\hat{\mu}_{7+8\,\text{TeV}} = 0.51^{+0.40}_{-0.37},$$

The breakdown of the uncertainty into statistical and systematic components is shown in Figure 5.45 (upper left) for the combined result and with separate signal-strength parameters for the two analyses. For both analyses and for the combination the statistical component of the uncertainty constitute the largest part. However, the systematical component is not much smaller.

The combined result is further investigated by having independent signal-strength parameters for the ZH and WH processes (Figure 5.45, top right) and for the lepton channels (Figure 5.45, bottom). In particular, the separate measurements of the ZH and WH signal strengths are interesting, since they provide information on the couplings of the Higgs boson to the W and Z bosons, which are expected to differ in the SM due to the different masses.

As already apparent in the post-fit BDT distributions, the data in the 1- and 2-lepton channels is preferring the background plus signal hypothesis, while the result for the 0-lepton is closer to the background only hypothesis. Since most of the sensitivity for the WH process is coming from the 1-lepton channel, the WH signal strength is close to unity as well. Instead, the ZH signal strength is is closer to zero, since most sensitivity is coming from the 0-lepton channel.

5.8.3. Significance and limits

The probability of the background only hypothesis, p_0, is evaluated as a function of the hypothetical Higgs boson mass from $m_H = 110$ to $140\,\text{GeV}$ in $5\,\text{GeV}$ steps. The results are shown in Figure 5.46. The observed significance is below the $2\,\sigma$ level for all mass points and it is $1.4\,\sigma$ for $m_H = 125\,\text{GeV}$. The corresponding expected significance assuming a signal strength of $\mu = 1$ is $2.6\,\sigma$. The discrepancy between the observed and expected

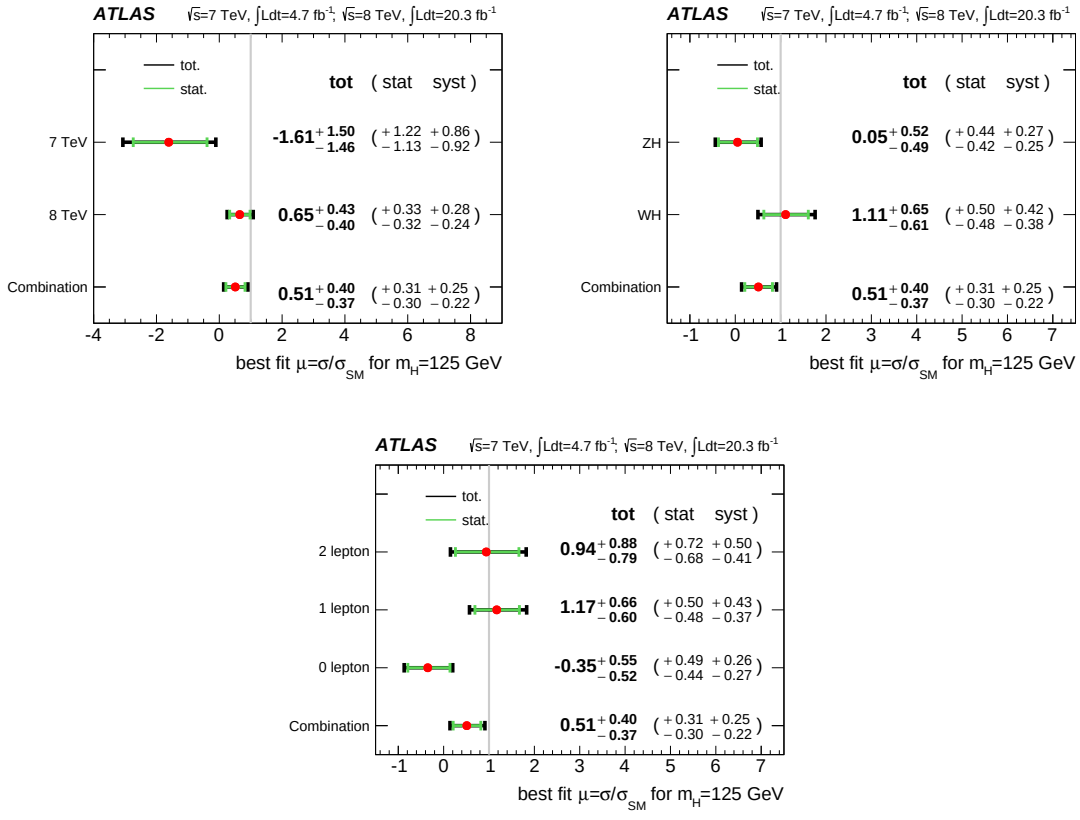


Figure 5.45.: Signal strength, μ, from fits to the data with breakdown of the uncertainty into statistical and systematic components. Fits with independent μ parameters for the 7 and 8 TeV analyses (top left), ZH and WH processes (top right) and lepton channels (bottom) are compared to the combined result (from Ref. [1]).

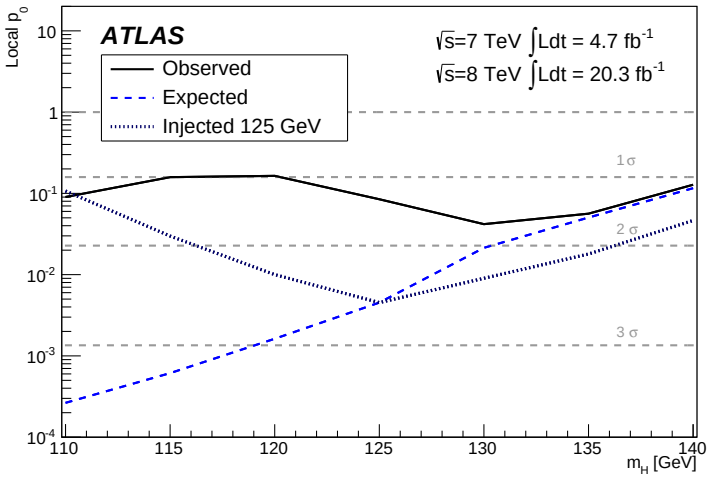


Figure 5.46.: Probability for the background only hypothesis, p_0, from the combined 7+8 TeV fit as a function of the hypothesized Higgs boson mass. The observed p_0, as measured in the data, is compared to the expected one, when assuming the existence of the Higgs boson with $\mu = 1$. The hypothesis of the Higgs boson with any mass is also tested against an injection of the signal with $m_H = 125\,\text{GeV}$ into the background model. The 1, 2 and $3\,\sigma$ significance levels are indicated by horizontal dashed lines (from Ref. [1]).

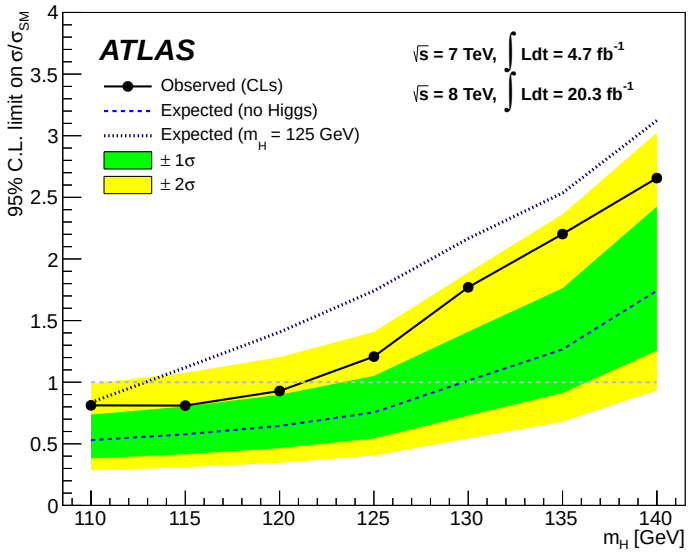


Figure 5.47.: 95 % CL upper limits on the Higgs boson production cross section times branching ratio relative to the SM expectation from the combined 7+8 TeV fit as a function of the hypothesized Higgs boson mass. The observed limit, as measured in the data, is compared to the expected one, when assuming the absence of the Higgs boson. The hypothesis of the Higgs boson with any mass is also tested against an injection of the signal with $m_H = 125\,\text{GeV}$ into the background model (from Ref. [1]).

values is due to the observed signal strength, which is below the SM expectation. The observed (expected) significance for the 8 TeV analysis alone is $1.7\,\sigma$ $(2.5\,\sigma)$.

The p_0 values are also given for a fit to the background hypothesis with an injection of the signal with $m_H = 125\,\text{GeV}$. This corresponds to the expected value at $m_H = 125\,\text{GeV}$, while the other signal mass hypothesis are less compatible with the injected signal and give smaller significances. This allows to estimate the sensitivity of the analysis to the Higgs boson mass.

Further, 95 % confidence level (CL) upper limits on the cross section of the signal, relative to the SM expectation, are shown in Figure 5.47. A slight excess above the expected value for the background-only hypothesis is observed for all mass points. However, it is withing the $2\,\sigma$ level and smaller than expected from the injection of the signal with $m_H = 125\,\text{GeV}$ into the background model. The SM Higgs boson with $m_H < 121\,\text{GeV}$ is excluded at the 95 % confidence level.

5.8.4. Summary distributions

Bins ordered by S/B Summary distributions are created by ordering all bins of the analyses by the signal-to-background ratio, S_i/B_i. Here, S_i is the expected signal yield

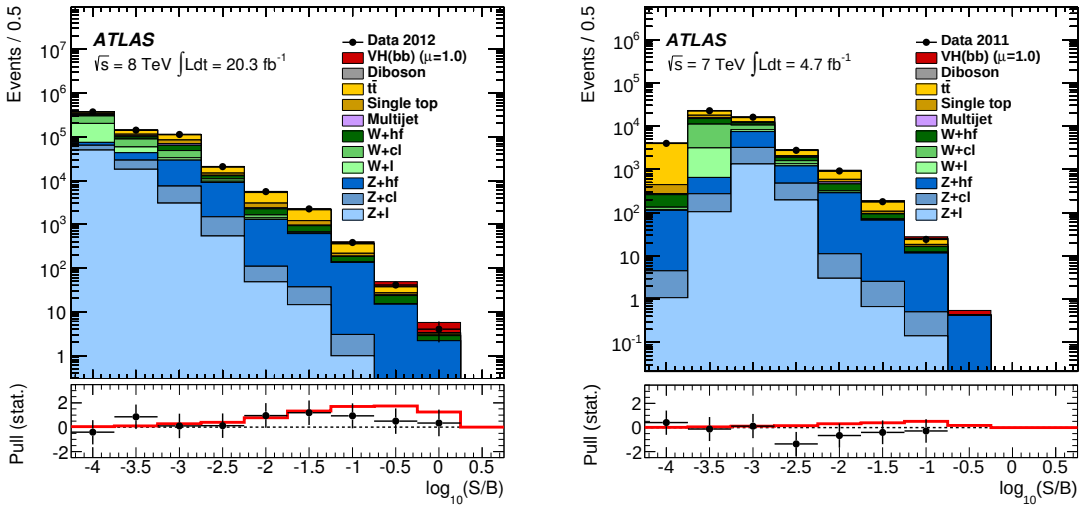


Figure 5.48.: Summary distributions of all bins in the 8 TeV (left) and 7 TeV (right) analyses ordered by S/B (from Ref. [1]).

in bin i for $\mu = 1$ and B_i is the corresponding background yield according to the post-fit expectation. The bins are merged in fixed intervals of $\Delta log_{10}(S/B) = 0.5$. The results are shown in Figure 5.48 separately for the 7 and 8 TeV analyses.

The pull of the data with respect to the background yield in each bin, $(N_i - B_i)/\sqrt{N_i}$, where N_i is the number of observed data events, is shown with its statistical uncertainty of ± 1. It is compared to the expected signal. As before, a slight excess in the data of the 8 TeV analysis over the background is observed, but not as large as expected from the SM Higgs boson. The plot for the 7 TeV analysis shows a deficit of the data compared to the background expectation, resulting in the negative signal strength quoted above.

Similar plots are shown separately for the three lepton channels of the 8 TeV analysis in Figure 5.49. Again, the data of the 0-lepton channel agrees more with the background only hypothesis, while the 1- and 2-lepton channels prefer the background plus signal expectation.

The post-fit yields corresponding to the bins of the 8 TeV analysis ordered by the signal-to-background ratio (Figure 5.48, left) are shown in Table 5.18.

Dijet mass The fit result of the dijet-mass analysis [1] is used to visualize the expected signal in the m_{jj} distributions. That analysis uses the dijet mass in five bins of p_{T}^{V} as final discriminant in the fit to the data. Each of the m_{jj} distributions is weighted by its ratio of S/B. Then they are combined and all backgrounds, except the diboson process, are subtracted from the data.

The result is shown in Figure 5.50 for the 7 and 8 TeV analyses and for the three lepton channels of the 8 TeV analysis in Figure 5.51. The combined statistical and systematic uncertainty on the background is indicated by the hatched band. Similar trends for the signal strengths are observed in these plots as for the fit results of the MVA analysis.

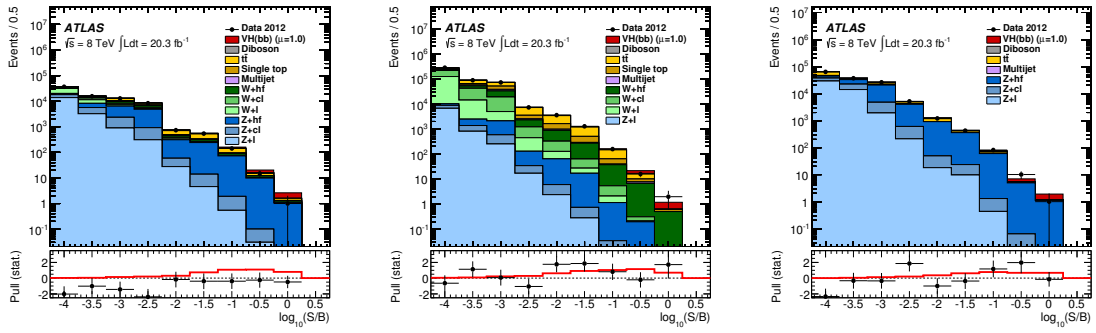


Figure 5.49.: Summary distributions of all bins in the 8 TeV analysis for the (left to right) 0-, 1- and 2-lepton channels ordered by S/B (from Ref. [1]).

Process	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6	Bin 7	Bin 8	Bin 9
Data	368550	141166	111865	20740	5538	2245	382	41	4
Signal	29	43	96	57	58	62	32	10.7	2.3
Background	368802	140846	111831	20722	5467	2189	364	37.9	3.4
S/B	8×10^{-5}	0.0003	0.0009	0.003	0.01	0.03	0.09	0.3	0.7
W+hf	14584	10626	15297	1948	618	250	45	8.2	0.7
Wcl	96282	30184	15227	1286	239	47	4.2	0.2	0.005
Wl	125676	14961	3722	588	107	16	1.3	0.03	0.001
Z+hf	10758	14167	21684	7458	1178	577	130	14.8	2.2
Zcl	13876	11048	4419	941	61	22	2.1	0.1	0.008
Zl	49750	18061	3044	537	48	15	1	0.05	0.004
$t\bar{t}$	30539	24824	26729	5595	2238	922	137	10	0.3
Single top	10356	9492	14279	1494	688	252	31	2.7	0.1
Diboson	4378	1831	1247	474	186	62	9.7	1	0.2
Multijet	12603	5650	6184	400	103	26	3	0.9	0

Table 5.18.: Predicted and observed numbers of events for the bins ordered by the signal-to-background ratio. The yields of the backgrounds are shown after the combined fit to the data. The signal process is listed for $m_H = 125\,\text{GeV}$ and $\mu = 1$ (from Ref. [1]).

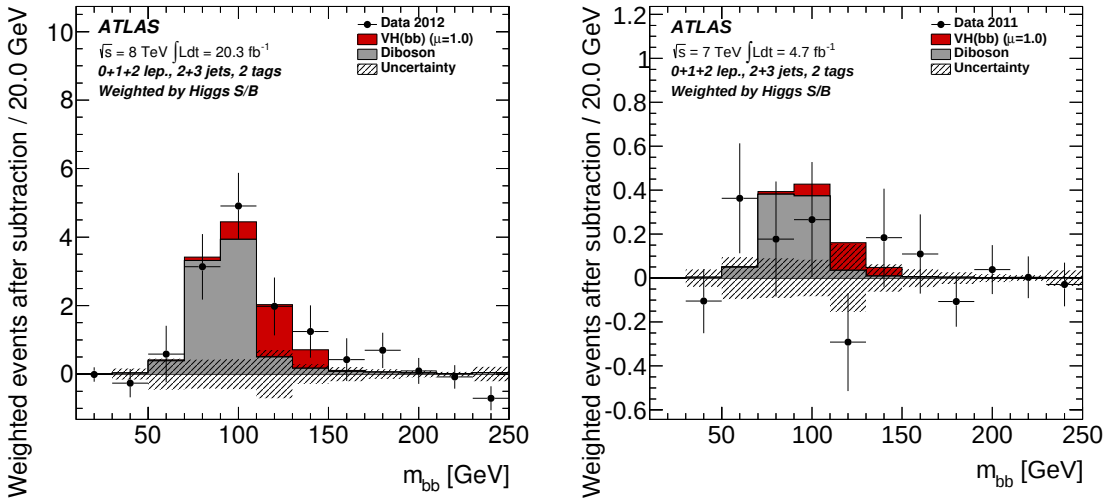


Figure 5.50.: Distributions of m_{jj} after weighting the regions in the dijet-mass analysis by S/B and subtracting the backgrounds for the 8 TeV (left) and 7 TeV (right) analyses (from Ref. [1]).

However, the two analyses do not directly correspond to each other, since they employ different selection criteria and are sensitive to somewhat different phase spaces.

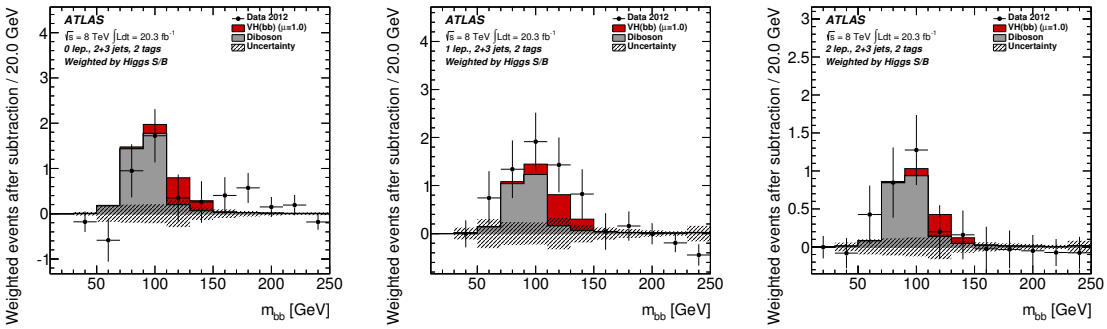


Figure 5.51.: Distributions of m_{jj} after weighting the regions in the dijet-mass analysis by S/B and subtracting the backgrounds for the 8 TeV analysis in the (left to right) 0-, 1- and 2-lepton channels (from Ref. [1]).

6. Sensitivity to H(bb) in gluon fusion

The search for the Higgs boson in the associated production with a vector boson, presented in the previous chapter, is the most sensitive search for the $H \to b\bar{b}$ decay. Searches in the ttH [122, 123] and VBF production modes [124] were performed by the ATLAS collaboration as well, resulting in somewhat less sensitive measurements.

The gluon-fusion production of the Higgs boson with the decay to b quarks, $gg \to H \to b\bar{b}$, is usually deemed not to be measurable at the LHC at all. This is mostly attributed to the production of b-quark pairs via QCD processes, which is an irreducible background and has an over seven orders of magnitudes larger cross section than the Higgs boson production, as shown in Figure 2.4.

However, the signal-to-background ratio improves for large transverse momenta of the Higgs boson. This can be exploited in a cut-based or multivariate analysis. Based on this, a feasibility study for measuring the $gg \to H \to b\bar{b}$ process in the data taken with the ATLAS experiment at $\sqrt{s} = 8\,\text{TeV}$, corresponding an integrated luminosity of about $20\,\text{fb}^{-1}$ has been performed in the course of this thesis and is presented in the following.

The analysis is designed close to the cross section measurement of $Z \to b\bar{b}$ decays with large transverse momenta of the Z boson [125]. The $Z \to b\bar{b}$ process has the same final state as the $gg \to H \to b\bar{b}$ signal and is the largest background from electroweak processes for the present analysis.

The dijet mass spectrum from the $Z \to b\bar{b}$ analysis is shown in Figure 6.1. Already from this spectrum one might guess that there is some sensitivity to the roughly 120 events from the $H \to b\bar{b}$ process (3 % of $Z \to b\bar{b}$ [125]). Further optimization for the $gg \to H \to b\bar{b}$ signal can be done, as presented in the following.

6.1. Simulated samples

Specific signal samples have been generated in the course of this thesis using the Powheg generator [92–94]. It is used with the CT10 PDFs [96] and interfaced to Pythia8 [87] with the AU2 tune [89, 90]. The signal samples are generated in slices of truth-level p_{T}^{H}. This allows for more events to be generated in the region with larger p_{T}^{H}, where most of the sensitivity is expected. The chosen ranges of p_{T}^{H} with the corresponding filter efficiencies are 70 - 120 GeV, 120 - 200 GeV and $> 200\,\text{GeV}$. An inclusive sample is generated as well for validation purposes, as discussed in Appendix B.1.

The $Z \to b\bar{b}$ background is simulated using the Sherpa 1.4.1 generator [98] with the CT10 PDF set. A truth level filter is applied, which requires at least one jet from the Cambridge-Aachen algorithm [126], using the truth-level particles as input, with $R = 1.2$ and $p_{\text{T}} > 160\,\text{GeV}$.

A possible bias from this truth-level filter is possible, since the event selection in the

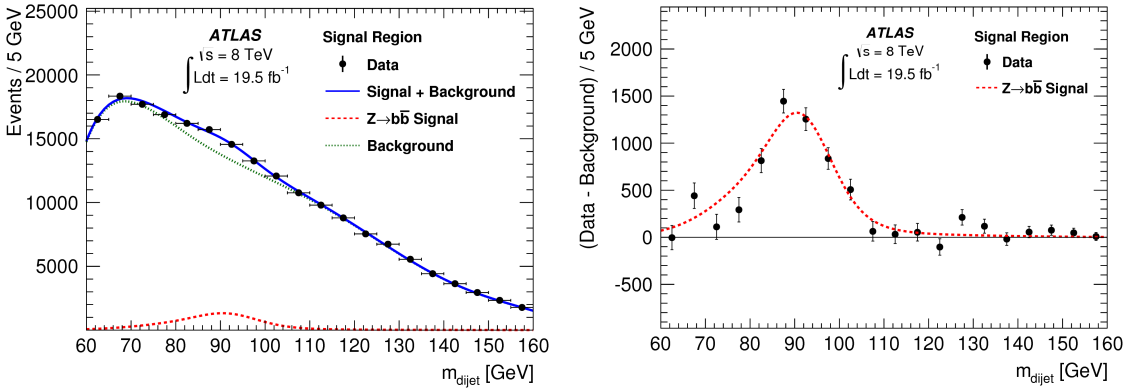


Figure 6.1.: Dijet mass spectrum from the $Z \to b\bar{b}$ cross section measurement (left) in the data (dots with error bars) together with background estimate (dotted line), the $Z \to b\bar{b}$ signal (dashed line) and the sum of both (solid line). The difference between the data and the background estimate is shown on the right (from Ref. [125]).

Name	Selected jets	b tags
EF_2b35_loose_j145_j35_a4tchad	$p_{\mathrm{T}}^{j1} > 145\,\mathrm{GeV}$, $p_{\mathrm{T}}^{j2} > 35\,\mathrm{GeV}$	≥ 2
EF_b45_medium_j145_j45_a4tchad_ht500	$p_{\mathrm{T}}^{j1} > 145\,\mathrm{GeV}$, $p_{\mathrm{T}}^{j2} > 45\,\mathrm{GeV}$	≥ 1
EF_b45_medium_4j45_a4tchad_L2FS	$p_{\mathrm{T}}^{j1-4} > 45\,\mathrm{GeV}$	≥ 1
EF_b145_medium_j145_a4tchad_ht400	$p_{\mathrm{T}}^{j1} > 145\,\mathrm{GeV}$	≥ 1
EF_j360_a4tchad	$p_{\mathrm{T}}^{j1} > 360\,\mathrm{GeV}$	≥ 0
EF_4j80_a4tchad_L2FS	$p_{\mathrm{T}}^{j1-4} > 80\,\mathrm{GeV}$	≥ 0

Table 6.1.: List of triggers with their requirements on the transverse momenta of the jets and on the number of b tags.

following applies requirements only on jets from the anti-k_t algorithm [75] with a radius of $R = 0.4$. However, the $Z \to b\bar{b}$ background constitutes less than 1 % of the total background. For the sensitivity estimate, using statistical uncertainties only, this possible bias can be neglected.

6.2. Event selection

The event selection requires one of the following six jet-based triggers, as listed in Table 6.1. Other triggers, such as muon triggers (for selecting muons from semi-leptonic b hadron decays) or other jet triggers, were investigated. Only percent-level gain in acceptance seem possible.

The object selection is very similar to the one used in the SM $VH(\to b\bar{b})$ analysis, as described in Chapter 4. At least three signal jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ are required. Two jets are expected from the $H \to b\bar{b}$ decay. Since Higgs bosons with large transverse momenta are to be selected, a third jet is expected, serving as recoil against the Higgs boson.

The b-tagging algorithm MV1 [79] is applied at a looser working point with 85 % tagging efficiency for the pre-selection and a tighter working point with 70 % efficiency for the signal region. Exactly two b-tagged jets are required at each step. The signal-region selection is applied in the following if not stated otherwise.

Cuts on the angular separation, $\Delta R(j_1, j_2) < 1.8$, and their combined transverse momentum, $p_{\mathrm{T}}^{jj} \geq 200\,\mathrm{GeV}$, are applied to the b-tagged jets. The cut on $\Delta R(j_1, j_2)$ is looser than for the $Z \to b\bar{b}$ analysis [125], which accounts for the larger mass of the Higgs boson and the resulting larger angular separation of the jets. The third jet is defined as the one that minimizes the transverse momentum of the vectorial sum of three jet four-vectors, p_{T}^{jjj}, in case of more than three jets in the event. This aims to select the jet recoiling against the Higgs boson.

It should be noted that the requirement of exactly two b-tagged jets with two different working points of the b-tagging algorithm is quite inefficient for the signal. A gain of about 10 % in acceptance is observed when cutting on a working point with 80 % tagging efficiency instead of 85 % in the pre-selection. This effectively loosens the veto on additional b-tagged jets and could be improved further.

The trigger efficiency for the signal after the signal-region selection is evaluated. It is shown independently for the six jet triggers as a function of p_{T}^{jj} in Figure 6.2. The combined trigger efficiency as a function of various kinematic distributions is shown in Figure 6.3. In combination the triggers reach an efficiency of about 60 % for $p_{\mathrm{T}}^{jj} = 200\,\mathrm{GeV}$ and close to 100 % for $p_{\mathrm{T}}^{jj} > 500\,\mathrm{GeV}$.

Muons that are within a jet are used to correct the jet four-momenta, as described for the SM $VH(\to b\bar{b})$ analysis in Section 5.2.4, to improve the dijet mass resolution. The resulting mass spectrum is shown in Figure 6.4. The resolution is estimated from a Gaussian fit to be about $\sigma \approx 11.7\,\mathrm{GeV}$. For a loosened cut on the dijet momentum of $p_{\mathrm{T}}^{jj} > 160\,\mathrm{GeV}$ the resolution is slightly worse with $\sigma \approx 12.3\,\mathrm{GeV}$. These fit results are only approximate, as a simple Gaussian distribution does not well describe the reconstructed mass peak.

6.3. Sensitivity estimate

The sensitivity to the $gg \to H \to b\bar{b}$ signal is estimated in the signal region using the dijet mass, m_{jj}, as discriminant. It is shown for the signal, the $Z \to b\bar{b}$ background and the data in Figure 6.5. The data is blinded around the expected signal peak for $110\,\mathrm{GeV} < m_{jj} < 140\,\mathrm{GeV}$. A third-order polynomial fit to the side-bands in the range $80\,\mathrm{GeV} < m_{jj} < 200\,\mathrm{GeV}$ is performed. It is used to generate pseudo data, estimating the background in the blinded region. The signal to background ratio in this region is 0.1 %. The statistical significance of the signal is estimated from a binned likelihood ratio,

$$LLR = \sum_{i=1}^{N} s_i \ln\left(1 + s_i/b_i\right), \tag{6.1}$$

where the index i runs over all bins of the signal, s_i, and background, b_i, histograms. The result for the significance is $\sqrt{\mathrm{LLR}} = 0.58\,\sigma$.

A multivariate discriminant is trained to estimate the possible gain from using more kinematic properties of the event. The method of choice are Boosted Decision Trees

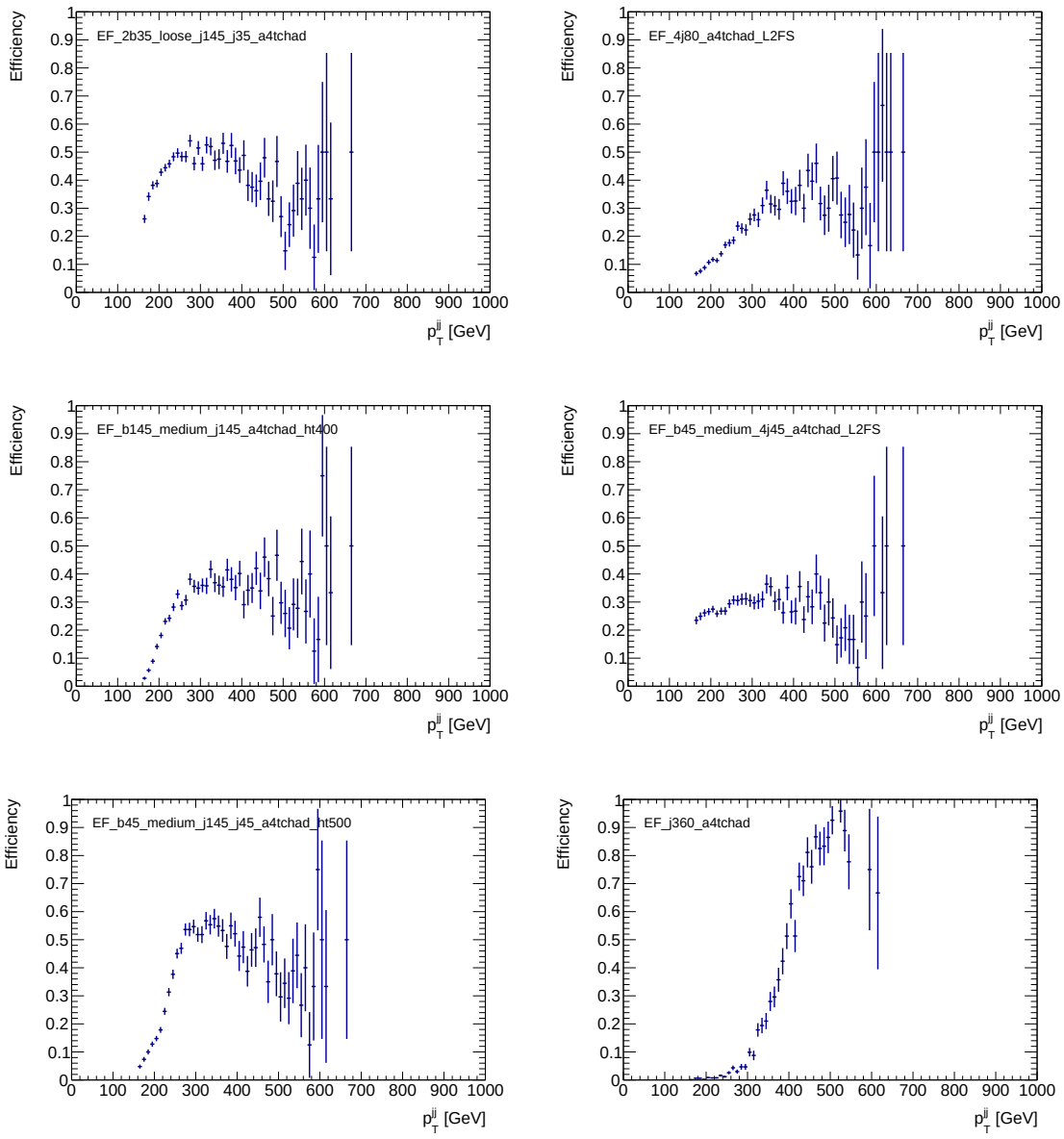


Figure 6.2.: Trigger efficiency for the $gg \to H \to b\bar{b}$ signal with respect to the signal-region selection. It is shown independently for the six jet triggers as a function of p_{T}^{bb}.

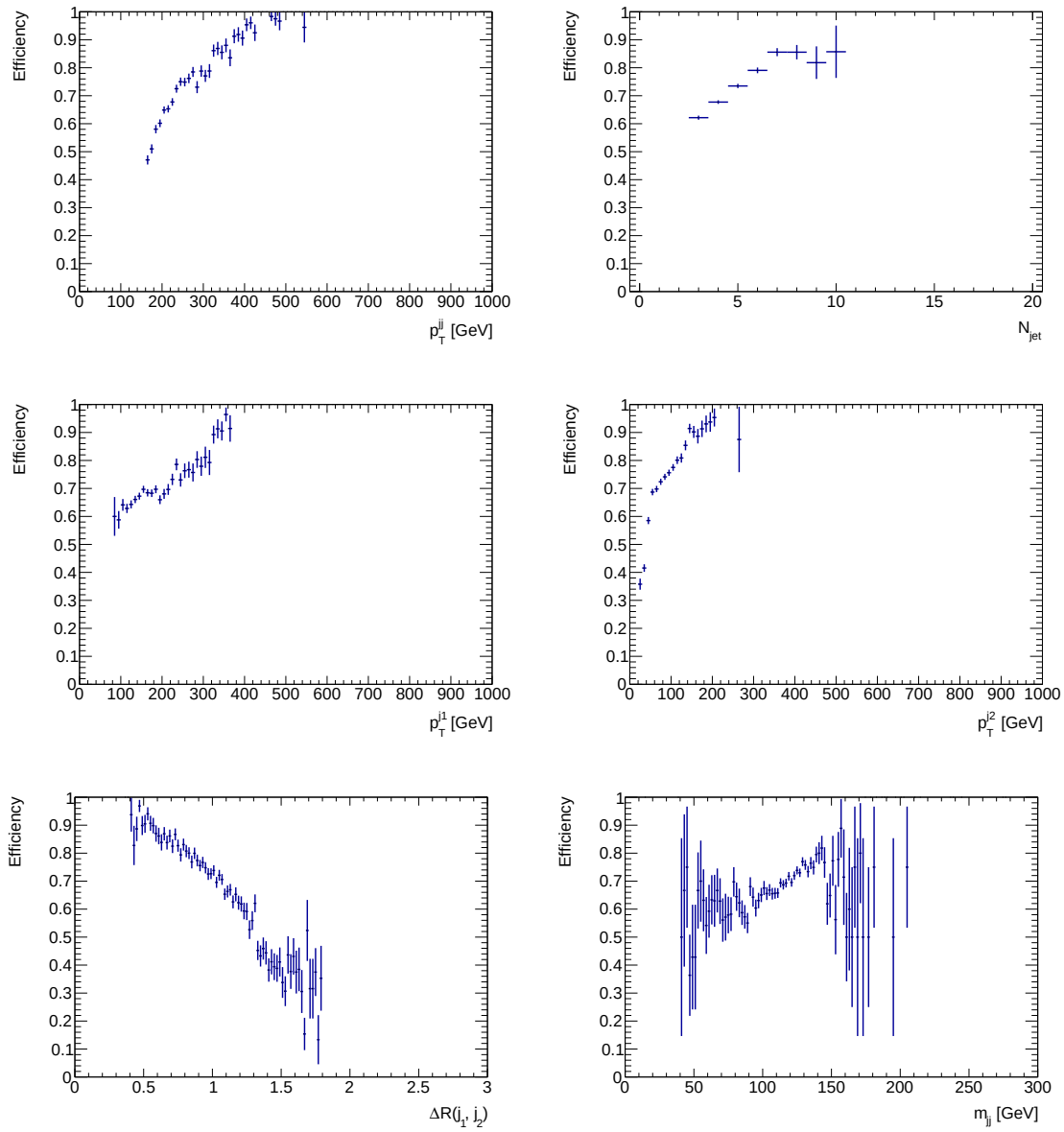


Figure 6.3.: Trigger efficiency for the $gg \to H \to b\bar{b}$ signal with respect to the signal-region selection. It is shown for the combination of the six jet triggers (at least one is triggered) as a function of various kinematic distributions.

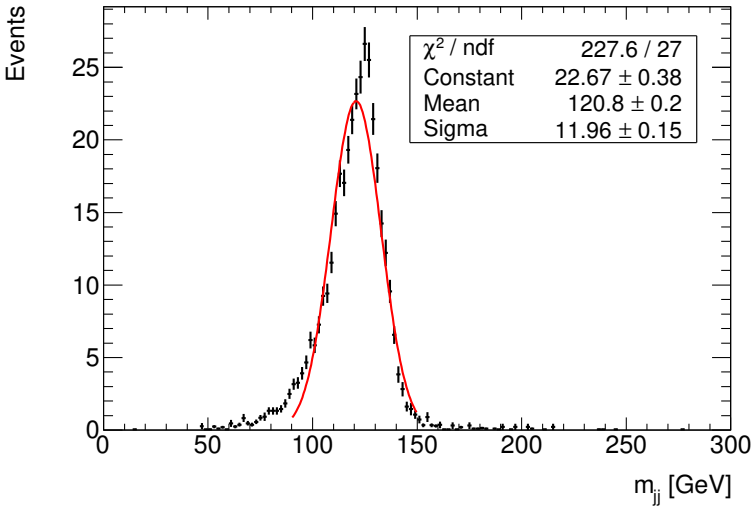


Figure 6.4.: Dijet mass distribution for the $gg \to H \to b\bar{b}$ signal with a Gaussian fit. The signal-region selection is applied.

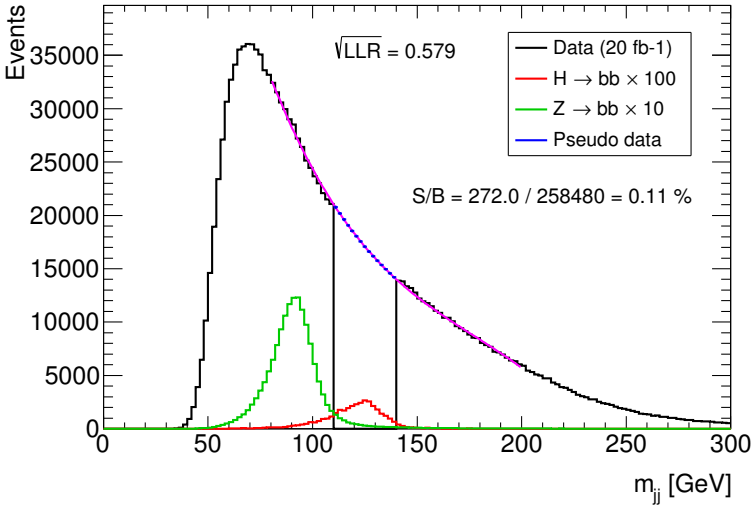


Figure 6.5.: Dijet mass distribution for the $gg \to H \to b\bar{b}$ signal, the $Z \to b\bar{b}$ background and the blinded data. A third-order polynomial fit (magenta) and the corresponding pseudo data are shown as well. The signal-region selection is applied. The statistical significance of the signal is estimated by $\sqrt{LLR}$ in the full distribution (binned) and the signal-to-background ratio, S/B is estimated in the blinded region (integral).

Variable	Name	Comment
N_{jet}	nJ	Number of jets
MV1($j_{1,2}$)	MV1B1,2	MV1 b-tagging output for the leading, sub-leading jet
p_{T}^{jj}	pTBB	p_{T} of the dijet system
η_{jj}	etaBB	η of the dijet system
$\Delta\eta(jj, j_3)$	dEtaBBJ	$\Delta\eta$ between the dijet system and the third jet
p_{T}^{jjj}	pTBBJ	p_{T} of the three-jet system
$\Delta R(jj, j_3)\|_{R=\text{min}}$	dRBBJm	ΔR between the dijet system and the closest other jet
$m_{jjj}\|_{R=\text{min}}$	mBBJm	Mass of the dijet system together with closest other jet

Table 6.2.: Explanation of the variables that are used to train the multivariate discriminant in addition to those defined for the SM $VH(\to b\bar{b})$ analysis.

(BDTs) with the same configuration as for the SM $VH(\to b\bar{b})$ analysis, as given in Section 5.3. The jet-based variables from that analysis, listed in Table 5.3, are used here as well. Additional variables are defined based on the third jet. They are listed in Table 6.2 together with a short explanation. The distributions of all variables used in the MVA, besides the dijet mass, are shown in Figures 6.6, 6.7 and 6.8.

The MVA discriminant is trained with the $H \to b\bar{b}$ signal sample against the data as background sample. The expected amount of $H \to b\bar{b}$ events in the data is small enough to have no noticeable effect on the training. The training is performed for various selection criteria and sets of variables, as listed in Table 6.3. Two trainings are performed for each configuration by swapping the training and test samples.

The performance of each BDT is evaluated on the corresponding test sample as the best-cut value of $N_{\text{sig}}/\sqrt{N_{\text{data}}}$. This corresponds, assuming the agreement of the data with the Standard Model, to $N_{\text{sig}}/\sqrt{N_{\text{sig}} + N_{\text{bkg}}}$. The results from the two trainings are averaged. The improvements for each configuration, relative to the simplest one, are given in Table 6.3. They range from a few percent to about 20 % for the most complex configuration.

It can be noted, that the variables η^{jj} and $\Delta\eta(jj, j_3)$, which were used in the $Z \to b\bar{b}$ analysis [125], do not contribute significantly to the separation of the Higgs boson signal from the multijet background.

The mass of the dijet system together with the closest other jet, $m_{jjj}|_{R=\text{min}}$, shown in Figure 6.8 (middle right), aims to reconstruct the Higgs boson mass in case of additional final-state radiation. In fact, a peak at the expected value of 125 GeV is visible in the distribution for the simulated signal. This feature is exploited by the MVA, but also a cut-based approach might make use of it, possibly with a separate 4-jet category.

One caveat of the present event reconstruction is apparent in the p_{T}^{jj} distribution in Figure 6.7 (upper right). The acceptance for the $H \to b\bar{b}$ signal seems to drop for $p_{\text{T}}^{jj} > 600$ GeV. This is likely due to the merging of the two jets for large Higgs boson transverse momenta and could be improved using boosted techniques, as discussed in Section 8.3.3.

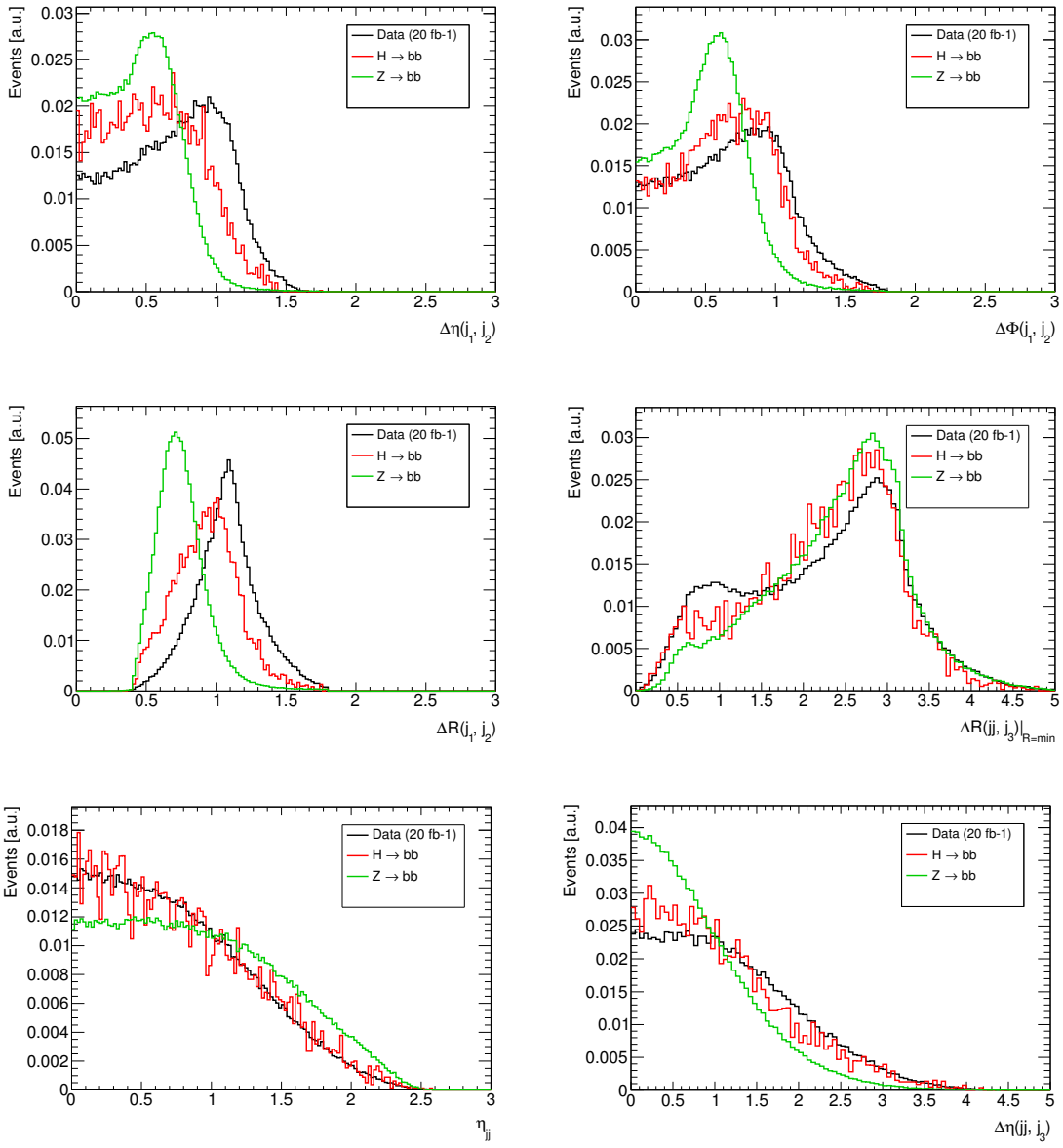


Figure 6.6.: Various angular distributions for the $gg \to H \to b\bar{b}$ signal, the $Z \to b\bar{b}$ background and the data, normalized to unity. The signal-region selection is applied.

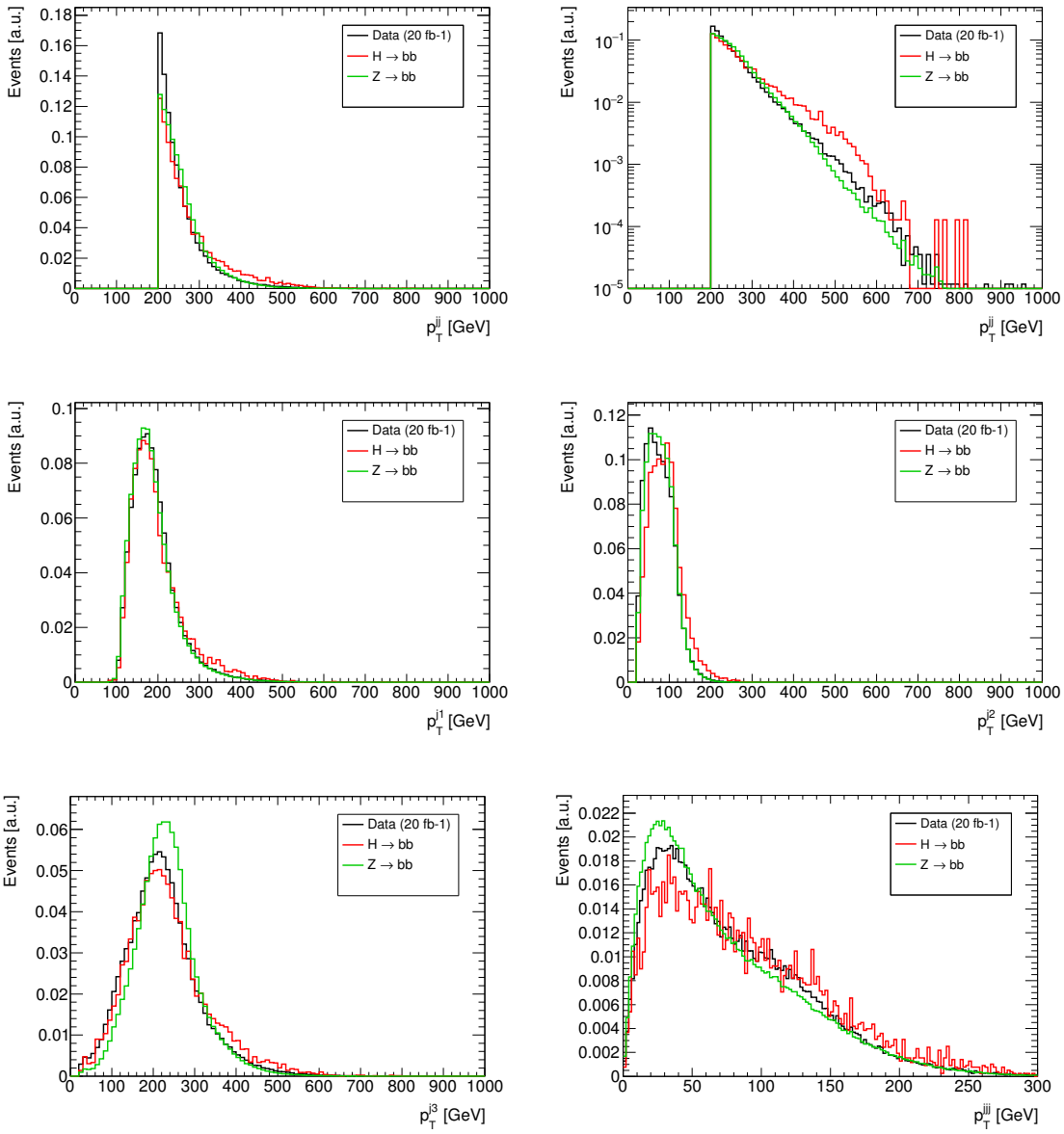


Figure 6.7.: Various transverse momenta distributions for the $gg \to H \to b\bar{b}$ signal, the $Z \to b\bar{b}$ background and the data, normalized to unity. The signal-region selection is applied.

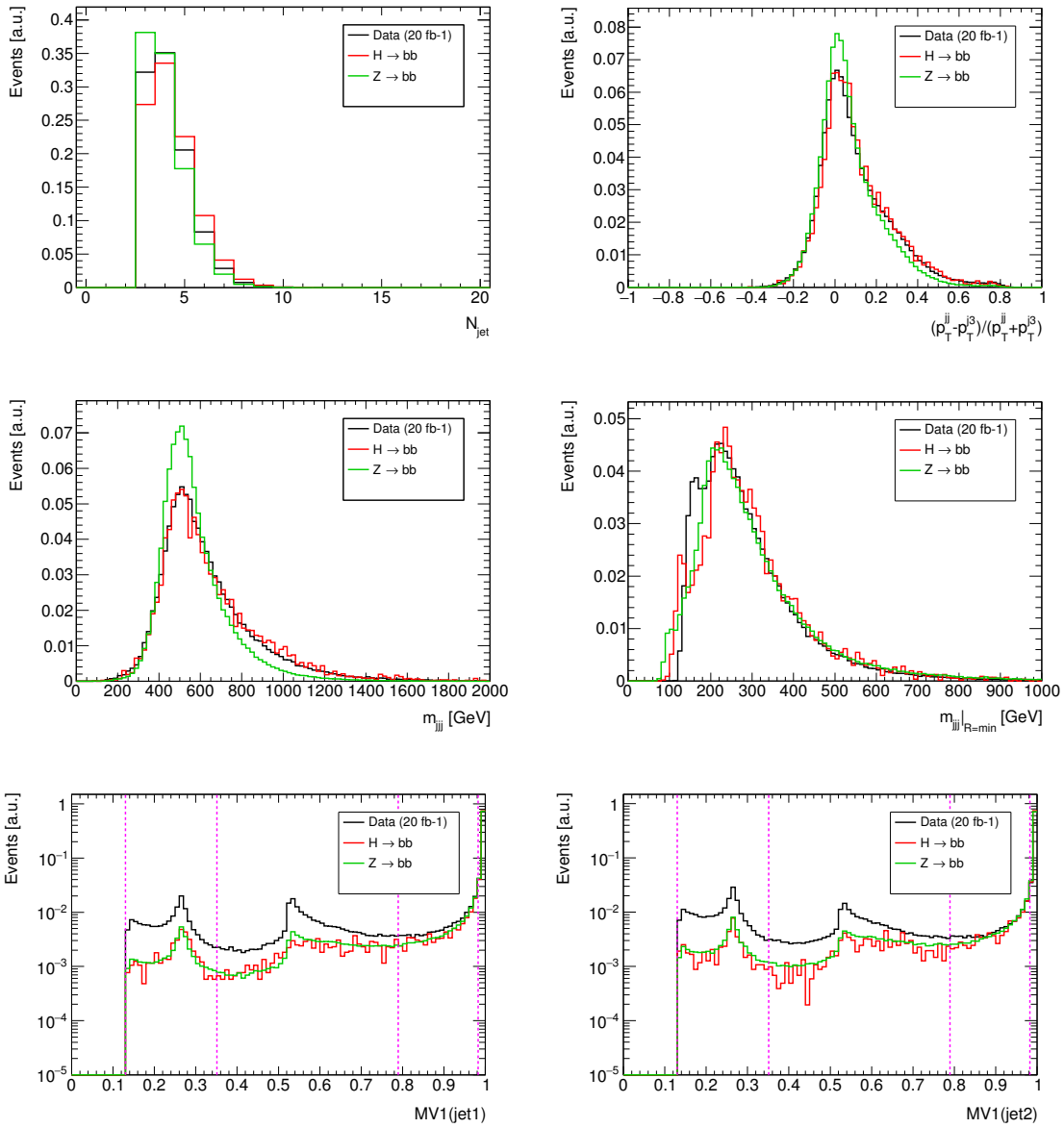


Figure 6.8.: Various kinematic and MV1 distributions for the $gg \to H \to b\bar{b}$ signal, the $Z \to b\bar{b}$ background and the data, normalized to unity. The signal-region selection is applied, except for the MV1 distribution, where the pre-selection (looser b-tagging requirement) is applied. The dashed magenta lines indicate the cut values corresponding to b-tagging efficiencies of (left to right) 85 %, 80 %, 70 % and 50 %.

Selection		Variables							Impr.
p_{T}^{jj}/GeV	MV1	m_{jj}	p_{T}^{jj}	$\Delta R(j_1, j_2)$	eta†	MV1($j_{1,2}$)	$p_{\mathrm{T}}^{j_{1,2}}$	add‡	
> 200	70 %	✓							0.0 %
		✓	✓						2.5 %
> 160	70 %	✓							7.3 %
		✓	✓						9.0 %
		✓	✓	✓					11.8 %
		✓	✓		✓				9.1 %
> 160	80 %	✓	✓			✓			13.7 %
		✓	✓	✓		✓			17.0 %
		✓	✓	✓	✓	✓	✓		18.9 %
		✓	✓	✓	✓	✓	✓	✓	20.3 %

Table 6.3.: MVA performance results for various sets of selection criteria and training variables. † eta = η_{jj}, $\Delta\eta(jj, j_3)$ ‡ add = N_{jet}, $\Delta\Phi(j_1, j_2)$, $\Delta\eta(j_1, j_2)$, $p_{\mathrm{T}}^{j_3}$, p_{T}^{jjj}, m_{jjj}, $\Delta R(jj, j_3)|_{R=\mathrm{min}}$, $m_{jjj}|_{R=\mathrm{min}}$

Channel	$\hat{\mu}$	Total unc.	Stat. unc.
$ttH(\to b\bar{b})$ hadronic	1.6	± 2.6	± 0.8
$ttH(\to b\bar{b})$ lepton+jets	1.2	± 1.3	± 0.8
$ttH(\to b\bar{b})$ dilepton	2.8	± 2.0	± 1.4
VBF $H \to b\bar{b}$	-0.8	± 2.3	± 1.3
$gg \to H \to b\bar{b}$	Expected:		± 1.4

Table 6.4.: List of the observed signal strengths with their total and statistical uncertainties from various searches for the $H \to b\bar{b}$ decay in the ttH [122, 123] and VBF production modes [124]. The uncertainties are compared to the expected statistical uncertainty for the $gg \to H \to b\bar{b}$ process from the present study. All of these analysis are performed with the data taken at $\sqrt{s} = 8\,\mathrm{TeV}$, corresponding to $\mathsf{L} = 20\,\mathrm{fb}^{-1}$.

6.4. Conclusion

The statistical significance for a measurement of the $gg \to H \to b\bar{b}$ signal in the data taken with the ATLAS experiment at $\sqrt{s} = 8\,\mathrm{TeV}$, corresponding to an integrated luminosity of $\mathsf{L} = 20\,\mathrm{fb}^{-1}$, is estimated to be about $0.7\,\sigma$. This corresponds to an expected statistical uncertainty of 1.4 on the signal strength and includes the possible improvements from the analysis methods described above.

The expected statistical uncertainty is compared to the observed ones from various searches for $H \to b\bar{b}$ in the ttH and VBF production modes in Table 6.4. It is apparent that a search for the $gg \to H \to b\bar{b}$ process might be competitive with these other production modes.

Further steps have to be taken for a full analysis, as precise estimates of the background processes will be needed. In particular, the modeling of the high-mass tail of the $Z \to b\bar{b}$ background will be critical. Here, the mass side-bands can be of use, but also the η^{jj} and

$\Delta\eta(jj, j_3)$ variables, which are mostly uncorrelated with the dijet mass [125], could be used to define a $Z \to b\bar{b}$ control region.

Another control region, defined by a looser b-tagging requirement, might help to estimate the multijet background. The $t\bar{t}$ background, which was not studied here, will have to be estimated as well. Defining bins in the jet multiplicity can probably be used for defining a corresponding control region. Further work has to be done on the jet triggers, which might have different efficiencies in the simulation than in the data.

For Run 2 of the LHC a measurement of the $gg \to H \to b\bar{b}$ process is becoming more attractive. The increased amount of data will help, but also the increased center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ is beneficial: the cross section for Higgs boson production rises stronger than for the multijet and $Z \to b\bar{b}$ backgrounds, as shown in Figure 2.4. Further, boosted techniques for reconstructing the $H \to b\bar{b}$ decay with large-R jets, as discussed in Section 8.3.3, will become more important. These techniques might improve the sensitivity to $gg \to H \to b\bar{b}$ significantly.

7. Search for the VH(invisible) process

In this chapter the search for the Higgs boson in the associated production with a hadronically decaying vector boson, $V \to qq$, and the decay to invisible Beyond Standard Model (BSM) particles, $H \to \chi\chi$, is presented. This process is referred to as $VH(\to \text{inv.})$ in the following. The analysis was published as Ref. [2] and contributed to the combination of several searches for invisible Higgs boson decays performed by ATLAS [48].

The search for invisible decays of the Higgs boson is complementary to the search for $H \to b\bar{b}$ decays, as presented in Chapter 5. The latter is particularly important for validating the total decay width of the Higgs boson predicted in the SM, and thus for indirect constraints on invisible BSM decays. Instead, the present analysis represents a direct search for such decays.

The signal and background samples are discussed in Section 7.1 and the event selection in Section 7.2. The statistical treatment is presented in Section 7.3, the systematic uncertainties in Section 7.4 and the binning strategies in Section 7.5. Finally, the fit model validation is discussed in Section 7.6 and the results in Section 7.7.

7.1. Signal and background processes

The signal process, $VH(\to \text{inv.})$, is categorized into $WH \to q\bar{q}'\chi\chi$ and $ZH \to q\bar{q}\chi\chi$. The corresponding LO Feynman diagrams are shown in Figure 7.1. The quark-induced signal events are generated using the NLO Powheg method as implemented in the Herwig++ generator [127] using the CTEQ6L1_CT10ME PDFs [128] with the UEEE3CTEQ6L1-_CT10ME tune [129].

The additional gluon-induced ZH production contributes approximately 5% to the total ZH cross section. The corresponding LO Feynman diagrams are shown in Figure 5.2. These events are not simulated, but are taken into account by increasing the quark-induced cross section as a function of the Higgs boson p_{T} by the appropriate amount. This correction is estimated on truth level using events from the Powheg generator [92–94]. It is used within the MiNLO approach [95] with the CT10 PDFs [96], interfaced to Pythia8 for parton showering and hadronization with the AU2 tune.

While the $VH(\to \text{inv.})$ yields a major part of the the sensitivity, the $gg \to H \to \text{inv.}$ process with additional jets provides a sizable contribution as well. It is generated using Powheg at NLO with the CT10 PDFs, interfaced to Pythia8 with the AU2 tune [101].

The contributions to the sensitivity from the $q\bar{q}' \to q\bar{q}'H$ production via vector-boson fusion (VBF) and from $qq/gg \to t\bar{t}H$ (ttH) production are estimated to be approximately 1 % and 0.2 %, respectively. The VBF contribution is strongly suppressed by requirements on the dijet mass and both are suppressed by the forward jet veto used to reduce the $t\bar{t}$ background. These signal processes are neglected in the following.

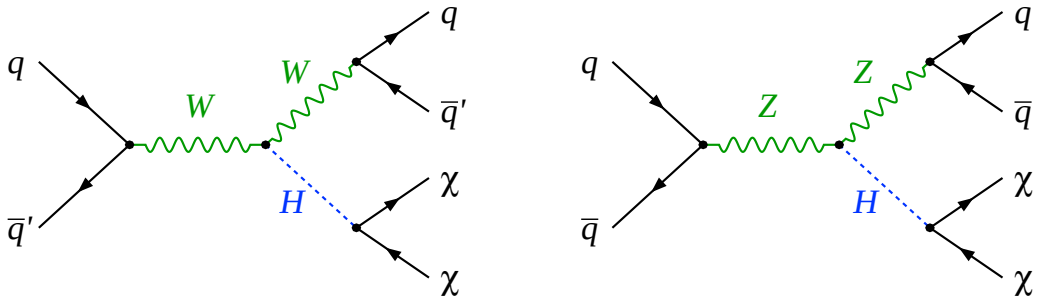


Figure 7.1.: Leading order Feynman diagrams of the $WH \to q\bar{q}'\chi\chi$ (left) and $ZH \to q\bar{q}\chi\chi$ (right) signal processes.

The background processes for this search are the same as for the SM $VH(\to b\bar{b})$ analysis, presented in Chapter 5: V+$jets$, $t\bar{t}$, single-top, diboson and multijet. In addition, the SM $VH(\to b\bar{b})$ process is considered as background. The samples are generated using the same techniques as in the SM $VH(\to b\bar{b})$ analysis (Section 5.1). The SM Higgs boson decay $H \to ZZ^* \to 4\nu$ leads to the same signature as the signal, since the neutrinos escape undetected. However, its BR is with about 0.1 % [43] far below the sensitivity of this analysis and is neglected.

The largest background is the production of Z bosons decaying to neutrinos with additional jets: $Z(\to \nu\nu)$+jets (Figure 5.3, right). This process has the same signature of $E_{\mathrm{T}}^{\mathrm{miss}}$ with jets as the signal. However, it can be reduced with cuts on the dijet mass. Further, b-tagging can be employed to extract the $ZH \to b\bar{b}\chi\chi$ component of the signal and suppress the Z+jets background, which has a smaller branching ratio for producing heavy flavor jets.

Similar to the 0-lepton channel of the SM $VH(\to b\bar{b})$ analysis, also the production of W+jets and $t\bar{t}$ are important background processes. Ideally, these processes can be rejected in the event selection by using vetoes on the additional leptons or jets in the final. However, their large cross sections together with inefficiencies in the reconstruction lead to sizable contributions in the end.

The only irreducible backgrounds to this search are the diboson processes $WZ \to q\bar{q}'\nu\nu$ and $ZZ \to q\bar{q}\nu\nu$. They can only be distinguished from the signal by kinematic distributions, such as $E_{\mathrm{T}}^{\mathrm{miss}}$, which is the final discriminant for this analysis.

7.2. Event selection

The event selection for the $VH(\to \mathrm{inv.})$ analysis employs many aspects of the selection for the SM $VH(\to b\bar{b})$ analysis (Section 5.2). The same selection criteria for the event cleaning, triggers, leptons and jets are applied.

The signal is expected in the 0-lepton channel, while the 1- and 2-lepton channels are used as control regions for the backgrounds. Similarly, most sensitivity is coming from the 2-jet selection, while the 3-jet region provides some additional one. The MV1c algorithm

Variable		Selection			
$E_{\mathrm{T}}^{\mathrm{miss}}$	[GeV]	120–160	160–200	200–300	> 300
$\Delta R(j_1, j_2)$		0.7–2.0	0.7–1.5	< 1.0	< 0.9
m_{jj} (2-jet)	[GeV]	70–100	70–100	70–100	75–100
m_{jj} (3-jet)	[GeV]	50–100	55–100	60–100	70–100

Table 7.1.: Event selection for the 0-lepton signal region. The cuts are optimized in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$ for the 2- and 3-jet selection separately.

is applied to the two leading jets with an b-tagging efficiency working-point of 70 % to define 0-, 1- and 2-tag regions.

The gain of the expected sensitivity due to the splitting into the three b-tagging categories was found to be about 28 % considering only the $VH(\rightarrow \mathrm{inv.})$ signal and about 17 % considering the $gg \rightarrow H \rightarrow \mathrm{inv.}$ process in addition. Not only the improved signal separation contributes to this gain, but also the better control of the various flavor components of the backgrounds.

Any loose lepton is vetoed for the 0-lepton selection. The kinematic selection requires $E_{\mathrm{T}}^{\mathrm{miss}} > 120\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$, which suppresses a large fraction of the multijet background. Further, the scalar sum of the transverse momenta of the jets is required to be larger than 120 GeV (150 GeV) for the 2-jet (3-jet) region. Similar to the SM $VH(\rightarrow b\bar{b})$ analysis, the cuts $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}}) < \pi/2$, $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{jets}) > 1.5$ are applied. The same ABCD method, based on these two variables, as described in Section 5.2.3, is applied to estimate the multijet background.

Besides the 0-lepton selection, also 1- and 2-lepton selections are used to define control regions for the background processes, mainly targeting the W/Z+jets and $t\bar{t}$ backgrounds. These control regions are used in the final fit, as described in Section 7.3.

The 1-lepton events are selected by requiring exactly one tight lepton, $E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{W} > 100\,\mathrm{GeV}$ ($E_{\mathrm{T}}^{\mathrm{miss}} > 50\,\mathrm{GeV}$ for $p_{\mathrm{T}}^{W} > 200\,\mathrm{GeV}$) and $m_{\mathrm{T}}^{W} < 120\,\mathrm{GeV}$. The 2-lepton events are selected by requiring two loose leptons of the same flavor, opposite charges and with invariant mass $83\,\mathrm{GeV} < m_{\ell\ell} < 99\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{Z} > 100\,\mathrm{GeV}$. The cuts on p_{T}^{W} and p_{T}^{Z} are applied in order to approximate the phase space of the signal region.

At least two signal jets are required for the 0- and 2-lepton channels and exactly two for 1-lepton. In the 0-lepton channel cuts on the $\Delta R(j_1, j_2)$ and m_{jj} are applied in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$ to select the hadronic W- and Z-boson decays. The corresponding values, optimized for the statistical significance of the $VH(\rightarrow \mathrm{inv.})$ signal, are given in Table 7.1. The dijet mass is calculated from the two leading jets. The distributions before the cut on m_{jj} and after the fit to the data are shown in Figures 7.2, 7.3 and 7.4 for the 0-, 1- and 2-tag regions, respectively. The selected interval around the W and Z boson masses define the 0-lepton signal region. The low and high side bands are merged and used as control regions.

The V+jets backgrounds are categorized based on the two leading signal jets into the flavor components bb, bc, bl, cc, cl and l, as described in Section 5.2. The flavor components are handled separately in the combined fit (described in the next section), however, they are shown inclusively in the post-fit plots.

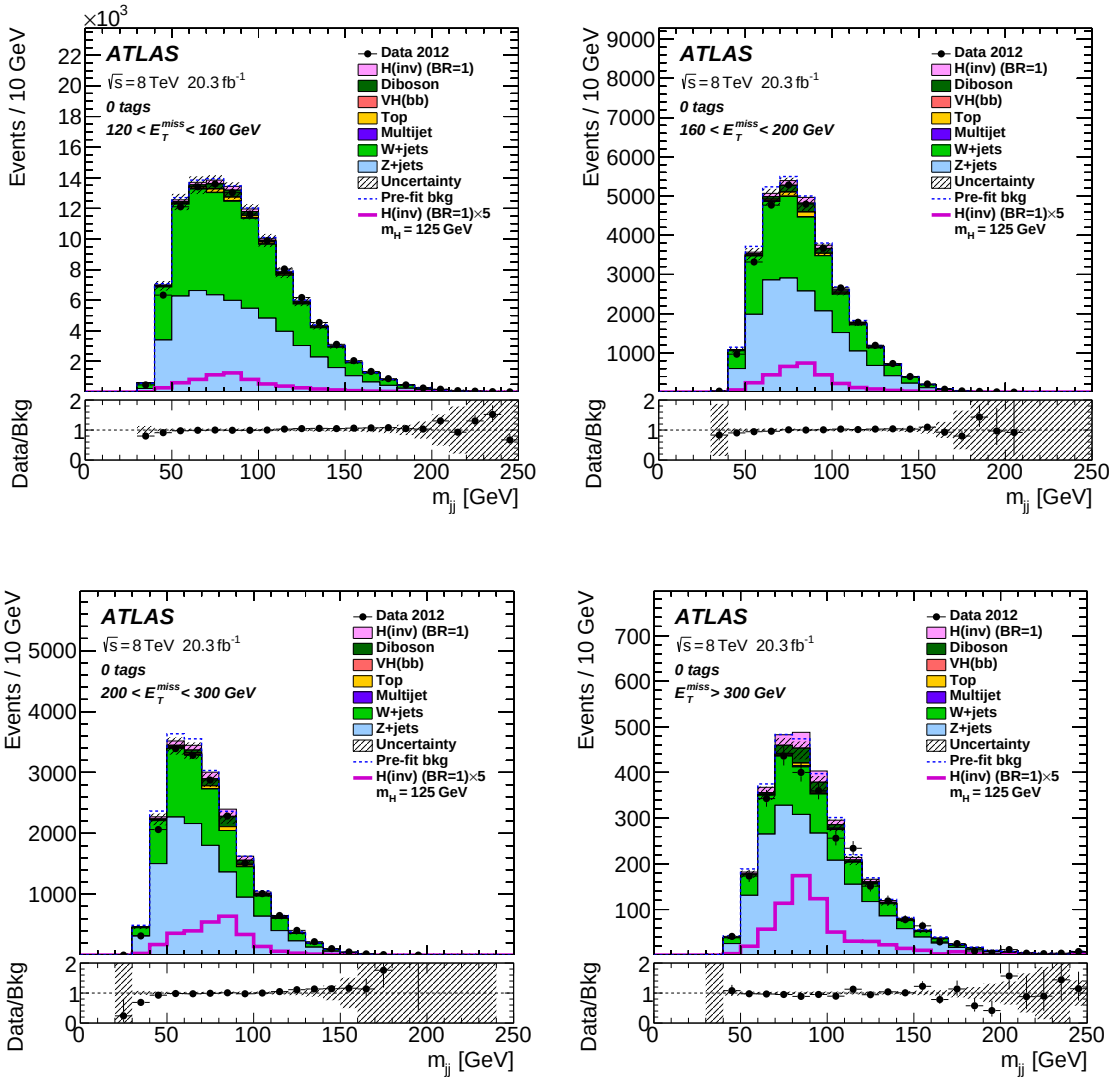


Figure 7.2.: Dijet mass distributions for the 0-tag region of the 0-lepton channel after the fit to the data in four bins of $E_{\mathrm{T}}^{\mathrm{miss}}$. A formal description of the plot elements is given in Section 7.6.1 (from Ref. [2]).

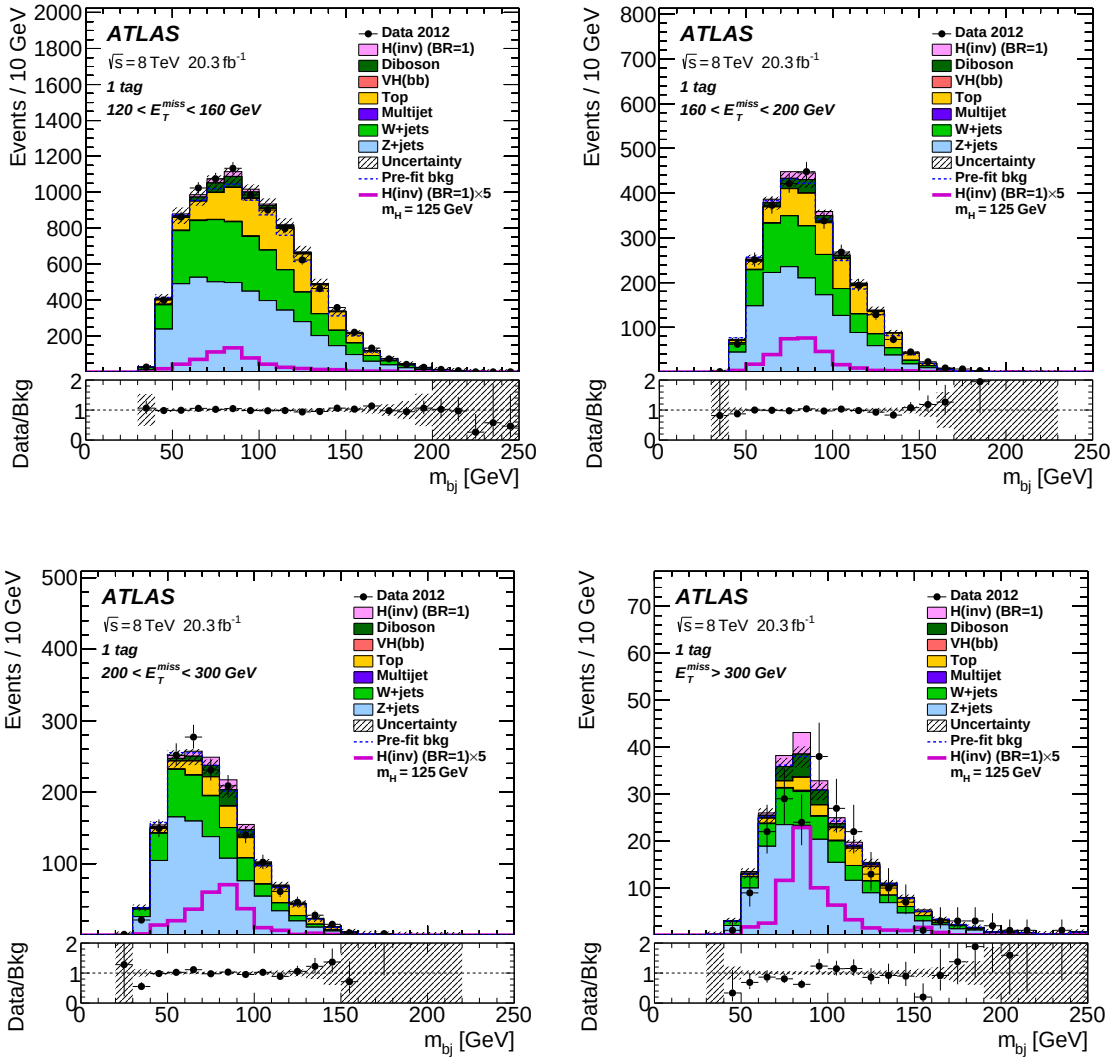


Figure 7.3.: Dijet mass distributions for the 1-tag region of the 0-lepton channel after the fit to the data in four bins of $E_{\mathrm{T}}^{\mathrm{miss}}$. A formal description of the plot elements is given in Section 7.6.1 (from Ref. [2]).

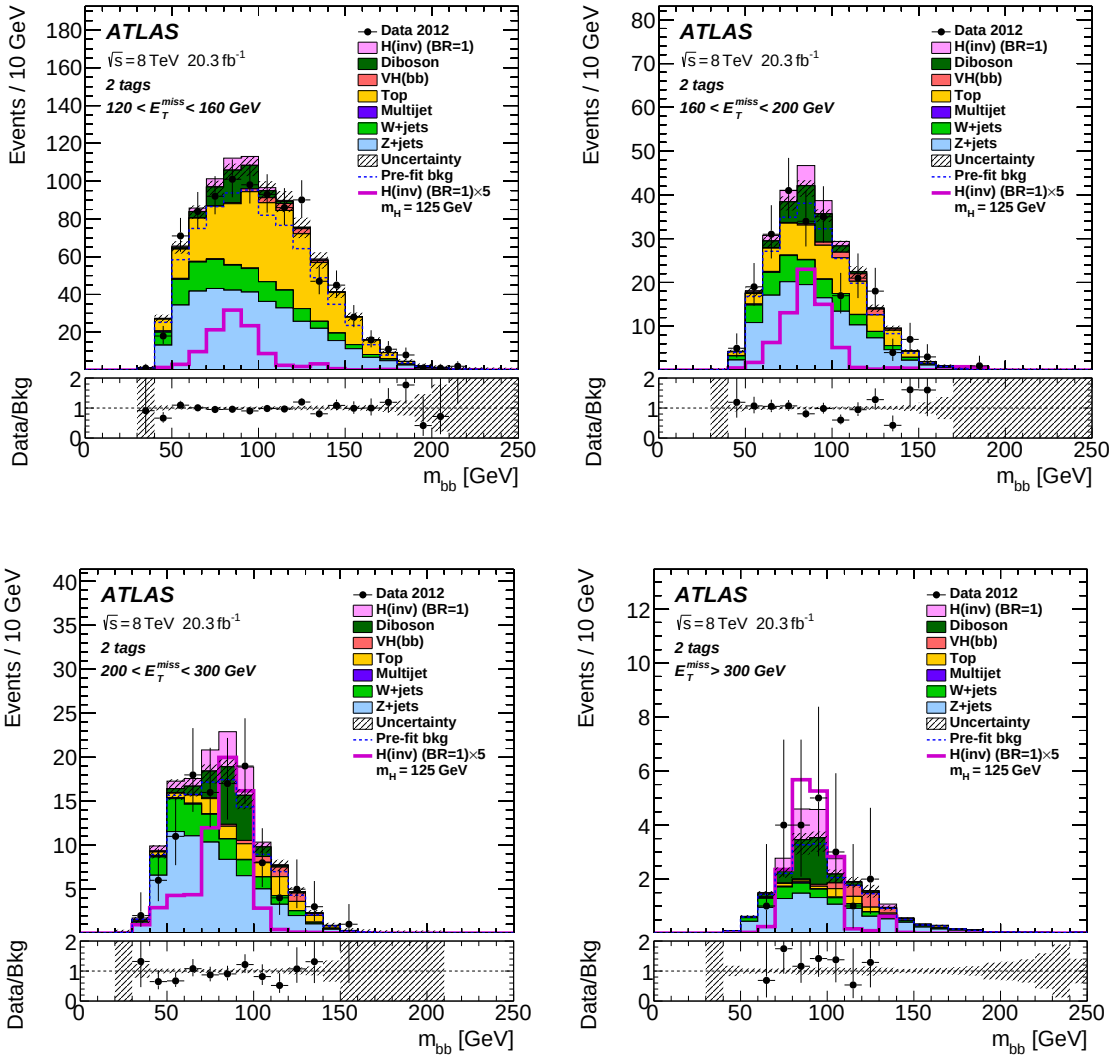


Figure 7.4.: Dijet mass distributions for the 2-tag region of the 0-lepton channel after the fit to the data in four bins of $E_{\mathrm{T}}^{\mathrm{miss}}$. A formal description of the plot elements is given in Section 7.6.1 (from Ref. [2]).

N	Comment	Leptons	Jets	b tags	m_{jj}
6	Signal regions	0	2, 3	0, 1, 2	SR
6	Side bands				SB
6	V+jets CRs	1, 2	2	0, 1, 2	incl.
1	$t\bar{t}$ $e\mu$ CR	2	2	2	incl.

Table 7.2.: List of the 19 regions entering into the combined fit. In total, 6 signal regions (SRs) and 13 control regions (CRs), including dijet mass side-bands (SBs), are defined for the $VH(\to \text{inv.})$ signal.

7.3. Statistical treatment

The statistical analysis of the data is performed using a binned profile-likelihood fit, as described in Section 5.4. The largest differences are the choice of regions and the final discriminant, which have been evaluated in the course of this thesis. In the present analysis the final discriminant is $E_{\mathrm{T}}^{\mathrm{miss}}$, opposed to BDT outputs as was the case for the SM $VH(\to b\bar{b})$ analysis. Minor differences exist in the parameterization of the systematic uncertainties, as described in the next section.

The parameter of interest is defined as the signal strength relative to the SM Higgs boson production cross section with $\mathrm{BR}(H \to \text{inv.}) = 100\,\%$. It is a freely floating parameter in the fit and is named `SigXsecOverSM`.

Six signal regions are used in the final fit: 0-lepton with 2 or 3 jets and 0, 1 or 2 b tags with signal region selection on the dijet mass. The six corresponding side-bands (low and high mass merged) are used as well. They serve as control regions for the background processes of the $VH(\to \text{inv.})$ signal. However, for the $gg \to H \to \text{inv.}$ signal the mass side-bands serve as signal regions as well, since the mass is no strong discriminant for this process with respect to the background processes.

Further, seven control regions are used in the fit, as listed in Table 7.2. The 1-lepton (2-lepton) regions with 0, 1 and 2 b-tags provide control of the various flavor components of the W+jets (Z+jets) backgrounds. The additional $e\mu$ control region helps to constrain the $t\bar{t}$ background for the 2-lepton channel.

Limits on the production cross section times $\mathrm{BR}(H \to \text{inv.})$ are derived for several values of m_H in the range of 115 - 300 GeV, considering only the VH production mode. For the mass of $m_H = 125\,\mathrm{GeV}$ limits on the $\mathrm{BR}(H \to \text{inv.})$ are derived considering all relevant production modes and their cross sections as predicted by the SM.

7.4. Systematic uncertainties

The experimental systematic uncertainties that are applied in the SM $VH(\to b\bar{b})$ analysis, listed in Section 5.5, are taken into account for the present analysis as well. Also the background modeling uncertainties are adapted with only minor adjustments, which have been developed in the course of this thesis and will be discussed in the following.

Floating scale parameters are assigned to all of the major background components, including W+l and Z+l. No prior is applied to the light components, since the 0-tag

regions provide sufficient constraint for a reliable measurement from the data.

Systematic uncertainties can be parameterized in the fit as shape (corresponding to "S" for the tables of Section 5.5), or as shape only ("SO"), neglecting an possible normalization effect. In the first case, the normalization effect of the uncertainty is taken into account for each region of the fit in a correlated way. In the latter case, the normalization effect is removed from the parameterization for each region.

These two choices correspond to two different scenarios: either an uncertainty is derived for some distribution that is used to define the regions of the fit or it is derived particularly for the distribution that is used as discriminant in the fit, and not for defining the regions. In the first case, the shape parameterization is more fitting to allow for changing the normalization differently in the various regions, hence providing an "extrapolation uncertainty". In the latter case, the shape-only parameterization might be the better choice to avoid possible degeneracies with generic normalization parameters.

Opposed to the SM $VH(\to b\bar{b})$ analysis, the normalization effects of the systematic uncertainties on the m_{jj} distribution are taken into account for each region in this analysis. This provides the necessary extrapolation uncertainty for the categorization in signal region and mass side-bands. The following NPs are parameterized as shape uncertainty, including a normalization effect (corresponding to "S" for the tables of Section 5.5): `W/ZMbb`, `TtbarMBBCont` and `VVMbb_WW/WZ/ZZ`. However, the normalization effects of the $E_{\mathrm{T}}^{\mathrm{miss}}$-related uncertainties, `WPtV` and `ZPtV`, are neglected (shape only, "SO"), as this distribution is directly used in the fit and not for defining regions.

The remaining differences with respect to the SM $VH(\to b\bar{b})$ analysis are discussed for specific samples in the following.

7.4.1. Signal

The systematic uncertainties on the signal include scale and PDF variations. The parton shower uncertainty is evaluated by multiplying the renormalization scale, μ_r, with factors of 0.5 and 2 for the $VH(\to \text{inv.})$ signal. The resulting yield variation is $\pm 5\,\%$ in 2-jet events and $\mp 8\,\%$ in 3-jet events (anti-correlated as indicated by the inverted $\pm$ sign). The factorization scale, μ_f, and PDF uncertainties for the $VH(\to \text{inv.})$ signal are estimated to be $\pm 1\,\%$ and $\pm 2\,\%$, respectively.

The renormalization-scale variations for the $gg \to H \to \text{inv.}$ signal have larger effects on the normalization. They are estimated in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$ to be $\pm 4\,\%$ below 200 GeV, $\pm 7\,\%$ up to 300 GeV and $\pm 15\,\%$ above. The PDF uncertainty is estimated to be $\pm 5\,\%$.

The $gg \to H \to \text{inv.}$ sample is generated with Powheg at NLO for zero or more additional jets. However, since two or three jets are required by the selection, an additional correction from an higher-order calculation is applied to achieve a better accuracy. The default sample is compared to the one from MiNLO HJJ [130], which generates $H + 2$ jet events at NLO. A correction scale factor is derived as a function of the p_{T}^{H} spectrum. It ranges from zero to $-10\,\%$ for $p_{\mathrm{T}}^{H} < 300$ GeV and up to $+50\,\%$ for $p_{\mathrm{T}}^{H} > 450$ GeV.

The same correction is used in the $H \to \gamma\gamma$ search published by the ATLAS collaboration [35] and is derived for a phase space with two or more jets. The corresponding systematic uncertainty is $\pm 15\,\%$ of the correction, which is applied for the 2-jet region, while a conservative uncertainty of $\pm 100\,\%$ is chosen for the 3-jet region.

Channel	Region	2-jet		3-jet	
		z_s	z_b	z_s	z_b
0-lepton	signal region	10	0	5	0
0-lepton	side bands	5	0	3	0
1- and 2-lepton	0-tag	0	1	–	
1- and 2-lepton	1- and 2-tag	0	5	–	
2-lepton	$t\bar{t}$ control region	0	1	–	

Table 7.3.: Parameters of transformation D after optimization.

7.4.2. W boson + jets

The systematic uncertainties derived in the SM $VH(\to b\bar{b})$ analysis are mostly sufficient to cover any systematic disagreement between the data and the simulation in the control regions of the present analysis. Only for the 0-tag region a residual mis-modeling of the p_{T}^{V} distribution is observed for the 1-lepton control region.

A correction is derived by fitting a linear function to the ratio of data/MC. The resulting function is applied as correction factor to the $W+l$ background in all regions. It ranges from $-5\,\%$ at $p_{\mathrm{T}}^{V} = 300\,\mathrm{GeV}$ to $-15\,\%$ at $p_{\mathrm{T}}^{V} = 500\,\mathrm{GeV}$ and is consistent between the electron and muon channels. No such disagreement is observed for the $Z+l$ background.

7.5. Binning strategies

A binning optimization is performed in order to increase the number of bins where signal is expected, while avoiding the proliferation of bins in regions that are dominated by the backgrounds. This optimization has been performed in the course of this thesis. The following strategy is based on the binning studies for the SM $VH(\to b\bar{b})$ analysis 5.6. Two options for the binning are evaluated: transformation D and F.

For transformation D the z parameters are chosen based on the sensitivity of the various regions. For the 0-lepton regions z_b is set to zero, resulting in a flat signal distribution. The z_s parameter ranges from three for regions with lower sensitivity to ten for the most sensitive distributions. All values are given in Table 7.3.

For the 1- and 2-lepton control regions, where no signal is expected, a flat background distribution ($z_b = 0$) is chosen. The number of bins ranges from one for regions constraining the normalization of some background to five for regions that yield some information about shape uncertainties as well.

The chosen parameters for transformation D result in a total of 92 bins in the fit. The expected limit is evaluated for varying the z_s parameter in the 0-lepton channel. The result is shown in Figure 7.5 (top, black markers). The limit with statistical uncertainties only (left plot) shows a clear improvement for increasing the number of bins, with an approximate saturation around 100 bins. A similar trend is observed when including systematic uncertainties (right plot).

The expected limit with transformation D is also evaluated as a function of the shape parameter $S = z_s/(z_s+z_b)$, introduced in Section 5.6.6. The results are shown in Figure 7.5

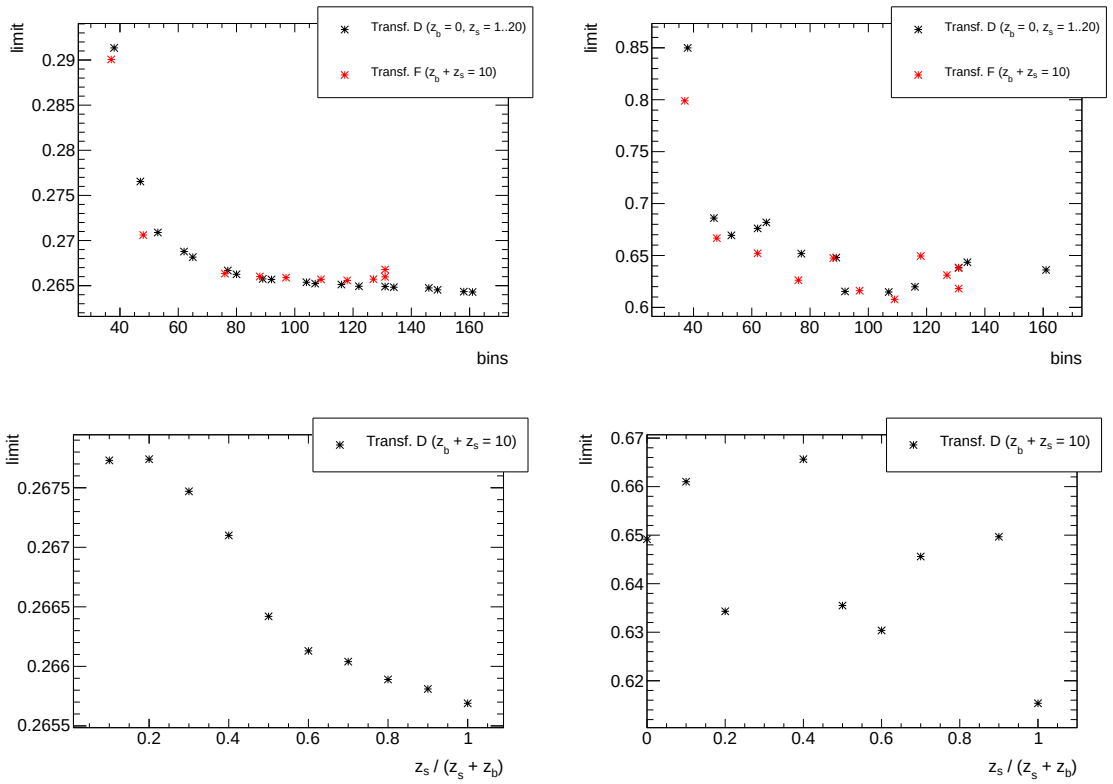


Figure 7.5.: Expected limit on BR($H \to$ inv.) as a function of the number of bins (top) and shape parameter $S = z_s/(z_s + z_b)$ (bottom) for transformations D (black) and F (red). The limit is shown on the left calculated with statistical uncertainties only and including systematics on the right.

(bottom). Here, the number of bins, $N_{\text{bin}} = z_s + z_b$, in the various regions are fixed to the values chosen above, i.e. ten for the most sensitive regions. A clear preference for $z_b = 0$ is visible in the limit with statistical uncertainties only, although the limit with $z_s = 0$ differs only by about 1 %. This trend is washed out with systematic uncertainties, which cause fluctuations of the order of a few percent in the limit.

Transformation F is evaluated as a function of the S parameter for $z_s + z_b = 10$ as well. The results are shown in Figure 7.5 (top, red markers). The number of bins is not constant in this case. Here, in contrast to the SM $VH(\to b\bar{b})$ analysis, transformation D and F show very similar performance. This might be attributed to the fact that the sensitivity of this analysis is driven by the systematic uncertainties, while transformation F was designed to minimize the statistical uncertainty only.

In conclusion, transformation D is chosen for its simplicity and good performance for this analysis. The parameters of the transformation are chosen as listed in Table 7.3. The expected limit is improved by about 4 % and the number of bins is reduced by about 50 % compared to a 50 GeV fixed-width binning of the $E_{\text{T}}^{\text{miss}}$ and p_{T}^{V} distributions.

7.6. Fit model validation

The fit model is validated in the following. Similar techniques as for the SM $VH(\to b\bar{b})$ analysis are applied. Motivational comments are given in Section 5.7. The validation has been performed in the course of this thesis.

7.6.1. Post-fit plots

All post-fit distributions of the regions that are used as input to the fit are shown in Figure 7.6 ($E_{\text{T}}^{\text{miss}}$ for the 0-lepton signal region) and Figure 7.7 (side bands), Figure 7.8 (p_{T}^{V} for the 1- and 2-lepton control regions) and Figure 7.9 (normalization in the top control region). The normalization parameter of the signal, μ, is floating in the fit and the result can be found in Section 7.7.

The various background components are shown stacked as colored filled histograms. The expected signal is shown for BR($H \to$ inv.) = 100 % (μ set to unity) on top of the backgrounds as pink filled histogram. It is also shown for better visibility as magenta hollow histogram with varying scale factors, as indicated in the legend. The data are shown as points with error bars. They are compared as ratio to the post-fit background prediction in the lower panels. The combined statistical and systematic uncertainty on the prediction is indicated by the hatched band. The pre-fit background prediction is shown in the plots as dotted blue line.

In general, good agreement of the background prediction with the data is observed after the fit. The chosen binning algorithm causes smaller bins for lower $E_{\text{T}}^{\text{miss}}$ and larger bins in the higher $E_{\text{T}}^{\text{miss}}$ tail. The bins of the 0-lepton distributions each contain about the same amount of expected signal events, which is not directly apparent in the plots due to the normalization of the bins to their width. The same is true for the backgrounds for the distributions of the 1- and 2-lepton channels. Alternatively, these plots can be drawn with equal bin widths, which is shown in Appendix C.1.

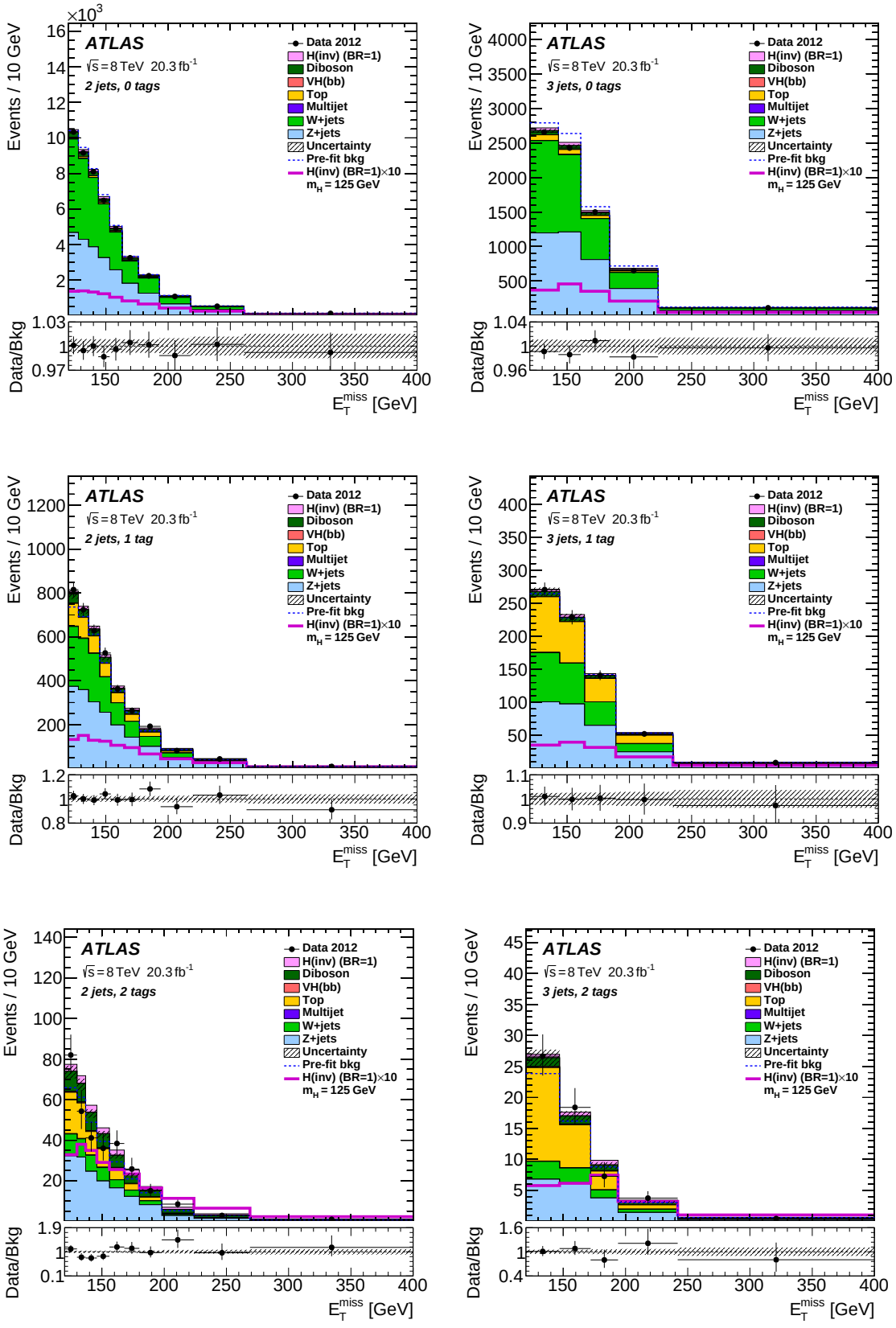


Figure 7.6.: Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions of the 0-lepton signal region for the 2-jet (left) and 3-jet (right) and (top to bottom) 0-, 1- and 2-tag regions. A formal description of the plot elements is given in the text (from Ref. [2]).

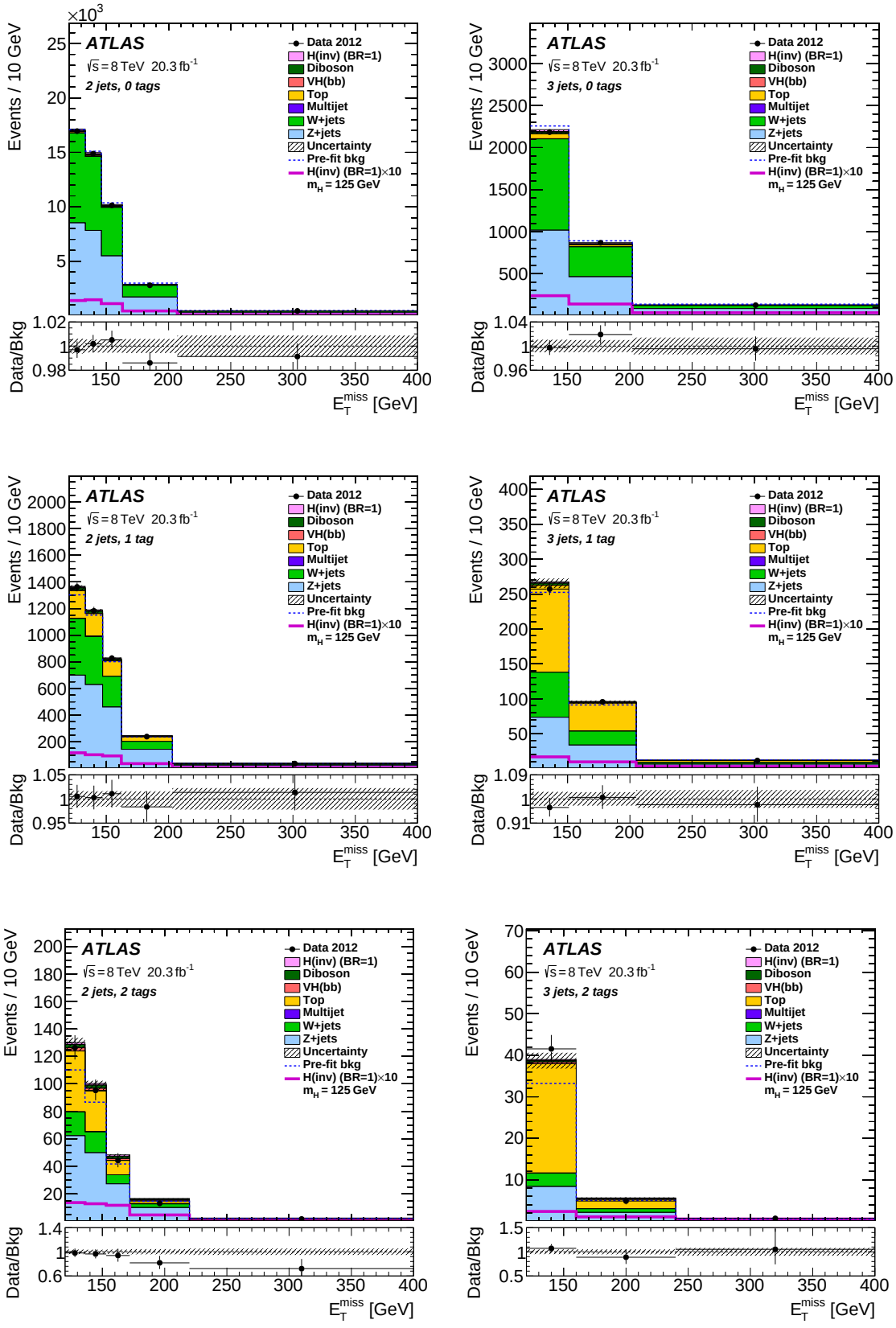


Figure 7.7.: Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions of the 0-lepton side-bands for the 2-jet (left) and 3-jet (right) and (top to bottom) 0-, 1- and 2-tag regions. A formal description of the plot elements is given in the text (from Ref. [2]).

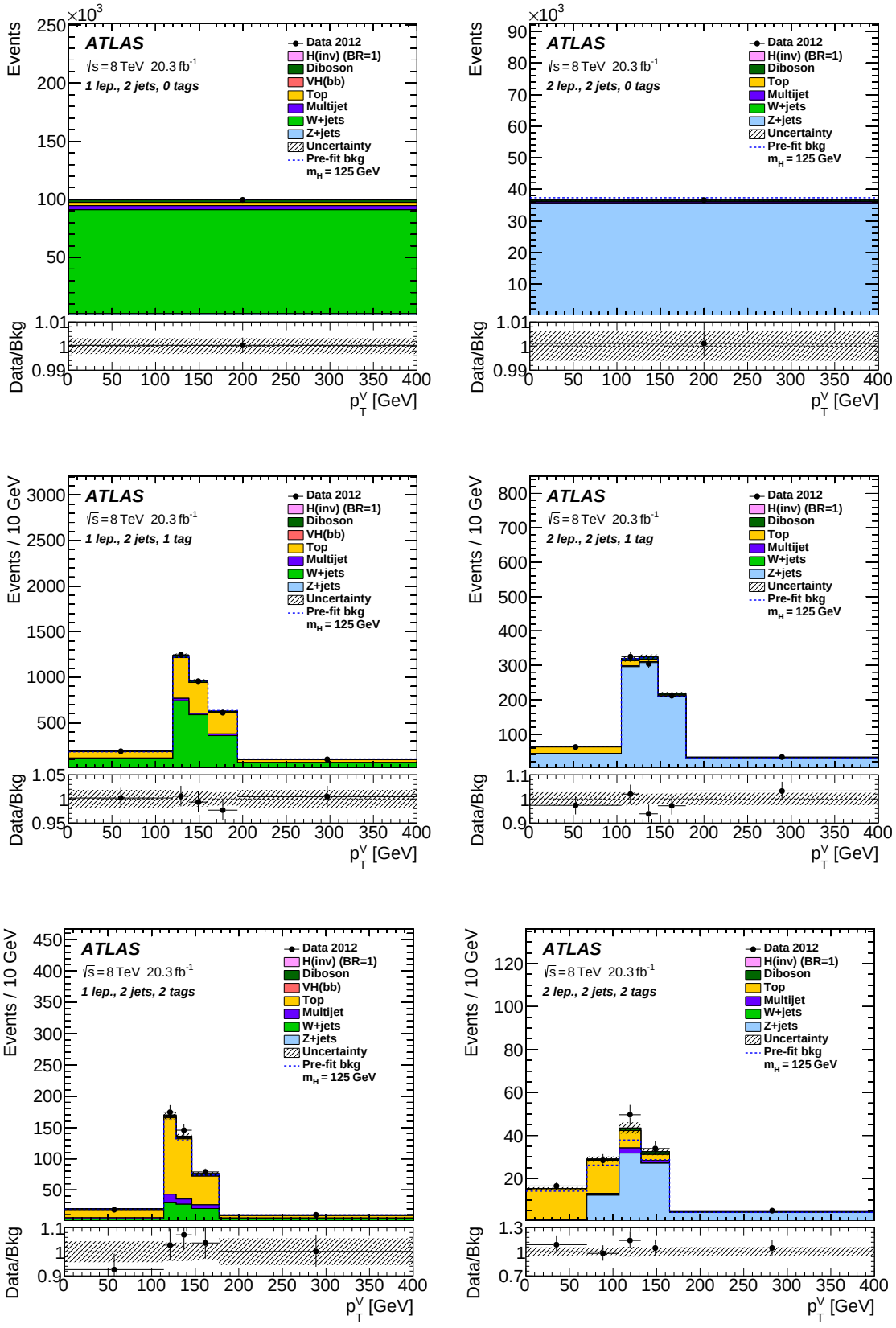


Figure 7.8.: Post-fit p_{T}^{V} distributions of the 1-lepton (left) and 2-lepton (right) channels for the (top to bottom) 0-, 1- and 2-tag regions. A formal description of the plot elements is given in the text (from Ref. [2]).

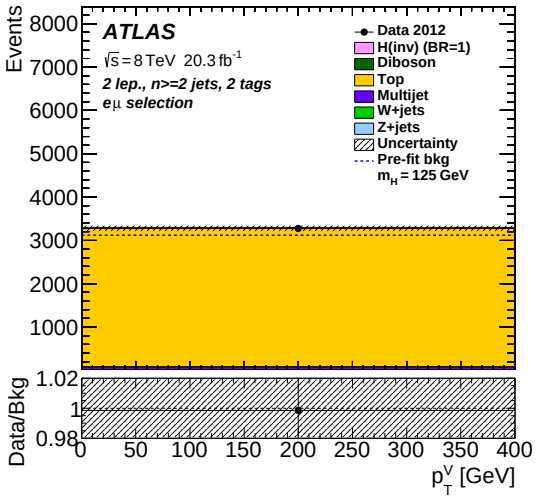


Figure 7.9.: Post-fit normalization of the 2-lepton top e-μ control region. A formal description of the plot elements is given in the text (from Ref. [2]).

7.6.2. Nuisance parameter pulls

The full set of nuisance-parameter pulls from the fit to the data is shown in Figures 7.10, 7.11 and 7.12. They are compared to the pulls from a fit to the Asimov dataset. The definitions of the parameters can be found in Sections 7.4 and 5.5.

The pulls in general are well behaved and do show similar trends as for the SM $VH(\to b\bar{b})$ analysis. This is expected, as similar phase spaces are used. For example, the $t\bar{t}$ normalization parameter for the 0-lepton channel, `norm_ttbar_L0` (Figure 7.10, upper left), is with 1.3 ± 0.2 above nominal. Instead, the parameters for the 1- and 2-lepton channels are close to unity as for the SM analysis.

The additional normalization parameters for the light components of the W+jets and Z+jets backgrounds, `norm_Wl` and `norm_Zl`, are consistent with unity.

An interesting feature is the `HiggsNorm` parameter (Figure 7.10, upper right), which parameterizes the 50 % prior on the $VH(\to b\bar{b})$ background. The pull of -0.54 ± 0.85 corresponds to a scale factor of 0.7 for this background. This is, with a large uncertainty, very close to the result of the dedicated SM $VH(\to b\bar{b})$ analysis.

Further some mild pulls are observed, which are all within $1\,\sigma$ of the nominal value. Some of the largest pulls ($> 0.5\,\sigma$) are observed on `WlNorm_J3`, `VVJetScalePTST1/2`, `ZDphi_J2_Zl`, `ZPtV_Zl`, `JetNP1/2` and `JetJVF`. All these pulls are consistent with the ones of the SM $VH(\to b\bar{b})$ analysis, discussed in Section 5.7.2.

The largest discrepancy is observed for `JetEResol`, which is at 0.70 ± 0.66, while it is -0.13 ± 0.49 for the SM $VH(\to b\bar{b})$ analysis. This might be attributed to different event categorization, using the dijet-mass cuts for the present analysis. Statistical fluctuations might also play a role for this experimental uncertainty. As the pull is within $1\,\sigma$, it was regarded as tolerable and was not further investigated.

Most other pulls are weaker than for the SM $VH(\to b\bar{b})$ analysis, which can be attributed

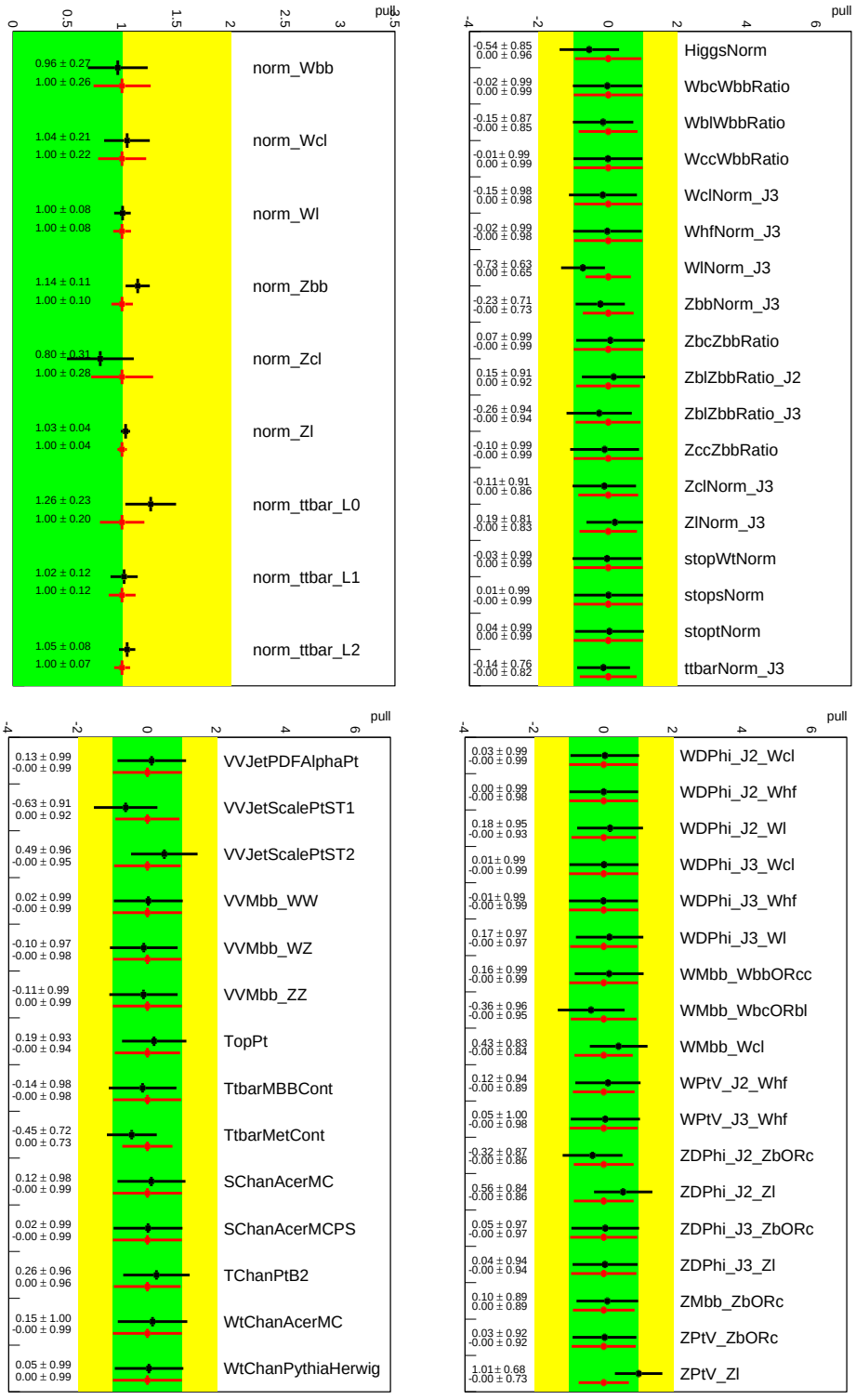


Figure 7.10.: Nuisance parameter pulls from the combined fit to the data (black) and to the Asimov data (red). Shown are the NPs for the floating normalizations (upper left), normalizations with priors (upper right), top-quark modeling (lower left) and V+jets modeling (lower right).

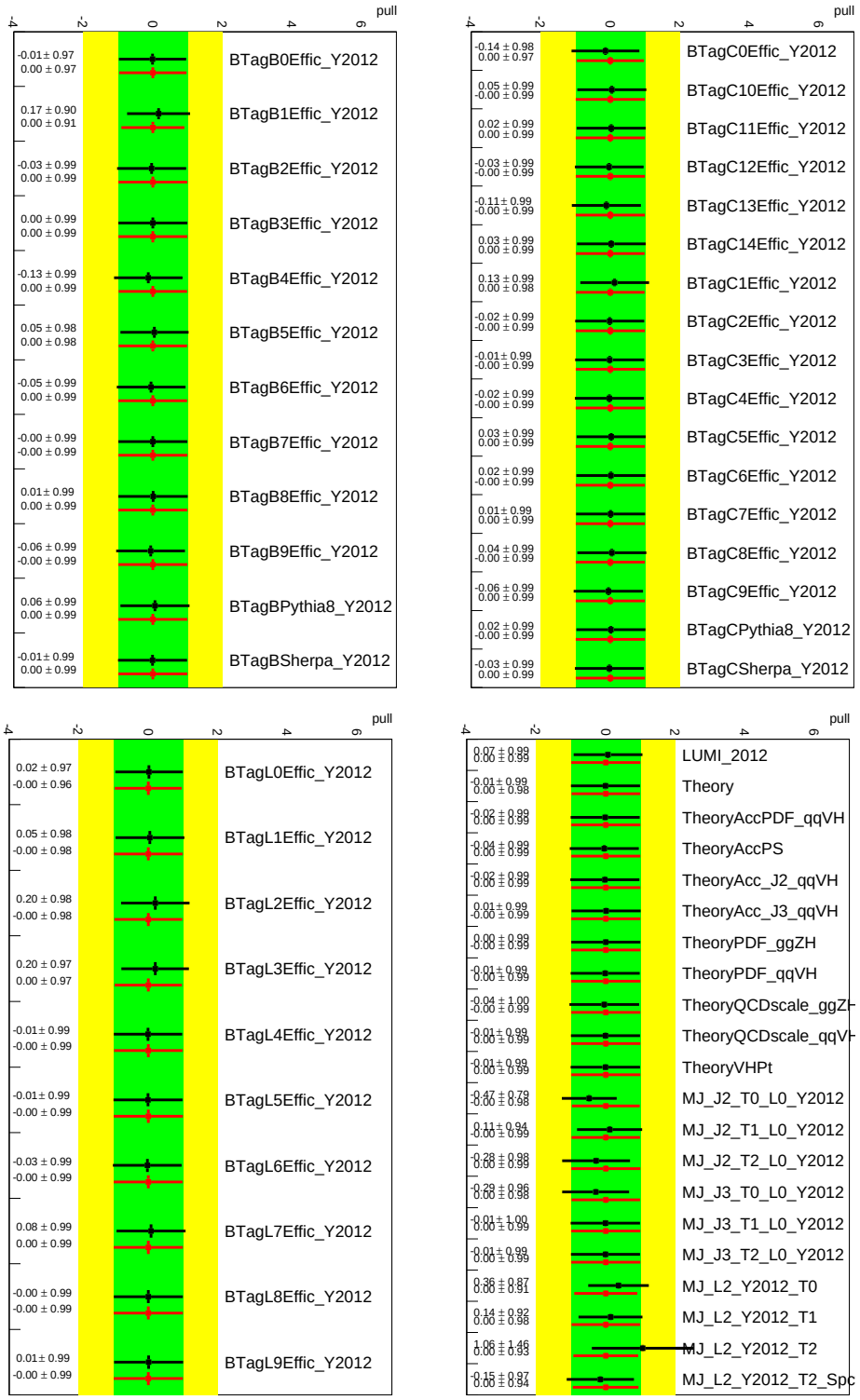


Figure 7.11.: Nuisance parameter pulls from the combined fit to the data (black) and to the Asimov data (red). Shown are the b-tagging NPs for b jets (upper left), c jets (upper right), light jets (lower left) and the remaining NPs (lower right).

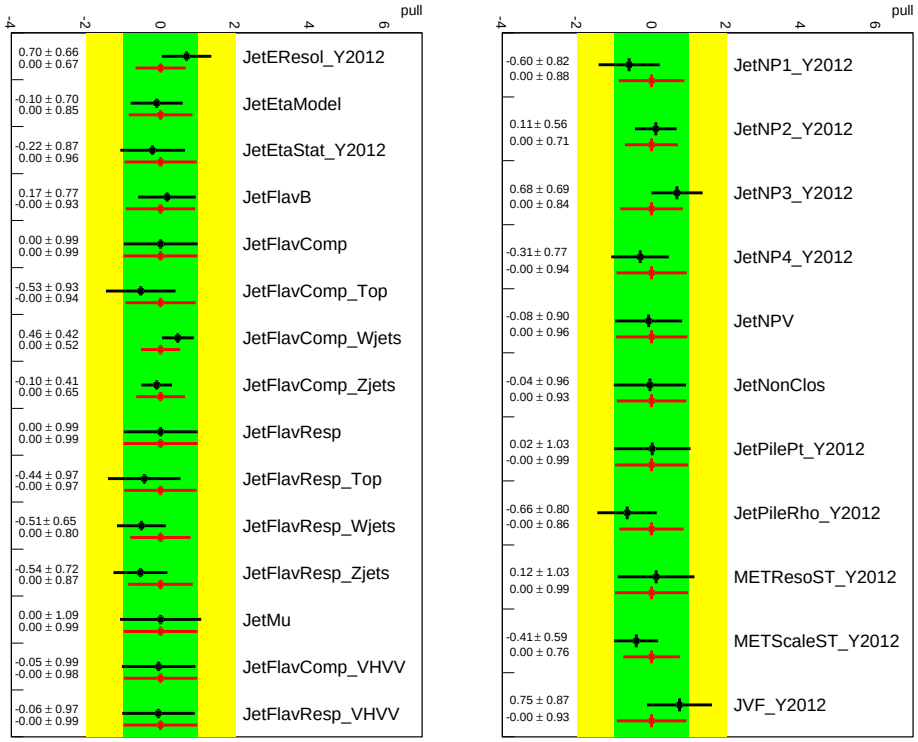


Figure 7.12.: Nuisance parameter pulls from the combined fit to the data (black) and to the Asimov data (red). Shown are the jet related NPs.

Parameter 1	Parameter 2	Correlation
SigXsecOverSM	VVJetScalePTST1	+30 %
	VVJetScalePTST2	−45 %
	ZPtV_Zl	+30 %
norm_Zl	norm_Zcl	−65 %
	ZblZbbRatio_J2	+45 %
norm_Zcl	ZblZbbRatio_J2	−60 %
norm_Wl	norm_Wcl	−30 %
	WDPhi_J2_Wl	+65 %
	WblWbbRatio	+35 %
norm_Wcl	WblWbbRatio	−55 %
norm_ttbar_L0	norm_ttbar_L1	+50 %
	ttbarNorm_J3	−80 %

Table 7.4.: Largest correlations observed in the fit to the data for any parameter and for the signal strength parameter, SigXsecOverSM.

to the simpler fit model. In particular, the b-tagging NPs (7.11) are very close to their nominal values in this analysis, while they are mildly pulled for the SM $VH(\to b\bar{b})$ analysis. This is explained by the use of the MV1c distribution directly in the fit for the SM $VH(\to b\bar{b})$ analysis, which is not the case here.

Overall, the pulls observed in the fit to the data are under control and do not raise concerns.

7.6.3. Correlations

The reduced correlation matrix (containing only parameters with at least one correlation with any other parameter above 20 %) as observed in the fit to the data is shown in Figure 7.13. The full correlation matrix and the one from the fit to the Asimov dataset can be found in Appendix C.2.

Due to the use of less control region compared to the SM $VH(\to b\bar{b})$ analysis some degeneracies are created and larger correlations occur. The largest ones are shown in Table 7.4.

In particular, strong correlations appear among the parameters for the Z+light and W+light backgrounds. This is due to the use of the floating normalization parameters for the light components and due to the lack of specific flavor-sensitive distributions, like MV1c, in the fit.

Similar to the SM $VH(\to b\bar{b})$ analysis, the $t\bar{t}$ normalization parameters show strong correlations between them. However, no strong correlation with the signal-strength parameter is observed. The positive correlation of norm_ttbar_L0 with norm_ttbar_L1 is mediated by negative correlations, such as with the W+jets normalization and modeling parameters. The negative correlation with ttbarNorm_J3 is due to its specific parameterization, as discussed in Section 5.7.2. These correlation could be reduced for the present analysis by using 3-jet regions for 1- and 2-lepton channels in the fit.

In both cases, for the V+light and $t\bar{t}$ normalization parameters, the use of additional

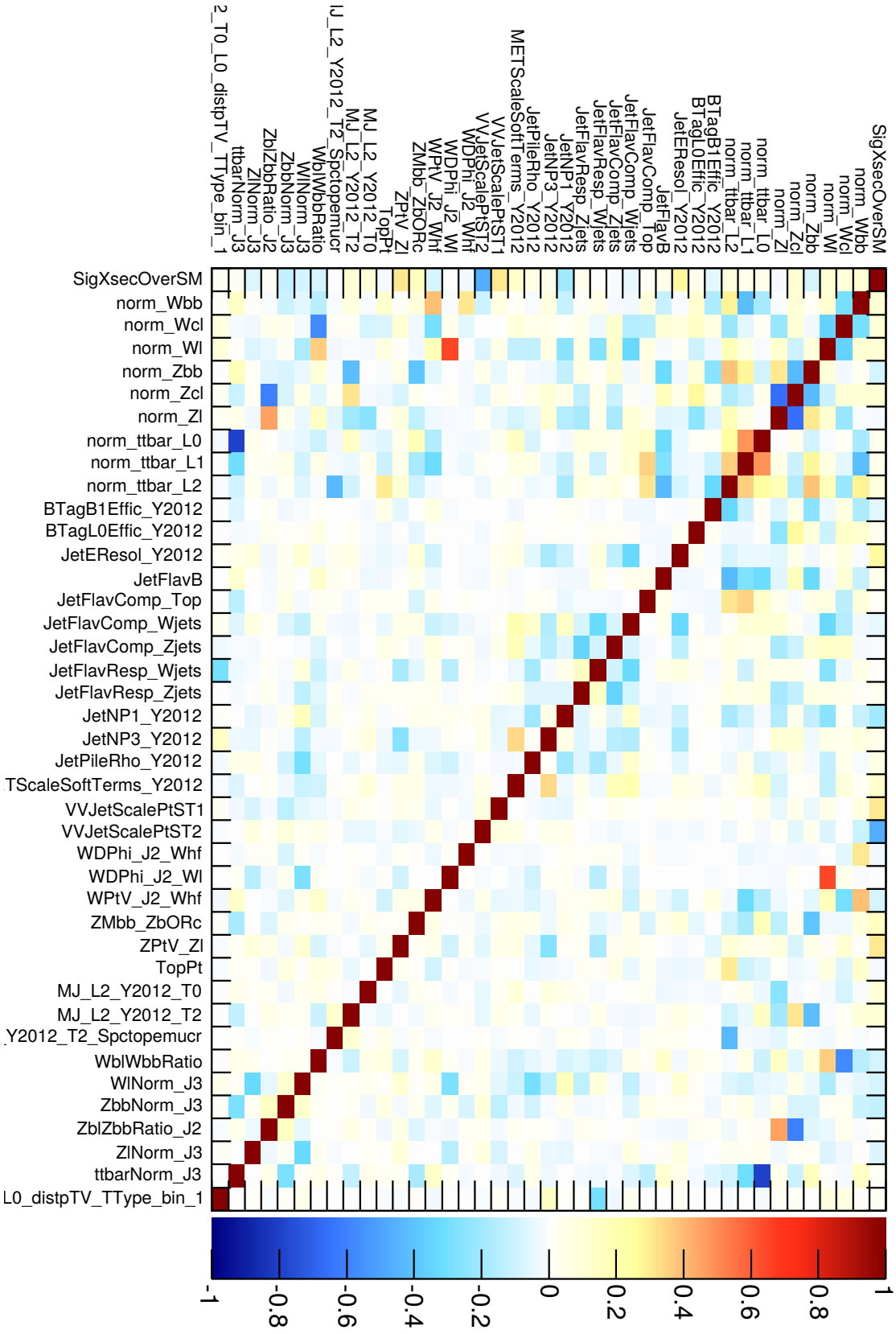


Figure 7.13.: Correlation matrix from the fit to the data. Only parameters with at least one correlation larger than 20 % are shown.

control regions was investigated to reduce the degeneracies. The correlations were indeed reduced, however only percent-level improvements to the sensitivity were found together with increased nuisance parameter pulls. This means creating sensitivity to parameters that are not of interest for the measurement. So the decision was taken to stay with the current fit model.

The signal strength parameter, `SigXsecOverSM`, shows a different behavior than for the SM $VH(\to b\bar{b})$ analysis. Its strongest correlations appear with the diboson normalization parameters, `VVJetScalePTST1/2`, and with the Z+light p_{T}^{V} shape parameter, `ZPtV_Zl`, as defined in Sections 5.5.7 and 5.5.4, respectively. Both cases are not surprising: a large portion of the sensitivity arises from the $E_{\mathrm{T}}^{\mathrm{miss}}$ ($\equiv p_{\mathrm{T}}^{V}$) distribution of the 0-lepton channel, 0-tag region. Here, the Z+light is the largest background component. The diboson normalization parameters are expected to play an important role as well, since the WZ and ZZ components are the only irreducible backgrounds to this search.

7.6.4. Ranking

The nuisance parameter ranking is shown in Figure 7.14. The highest ranked parameters are `VVJetScalePTST1/2` and `ZPtV_Zl`, which also showed the strongest correlations with the signal strength parameter in the previous section. The largest impact on $\hat{\mu}$ of about 0.20 is observed from `VVJetScalePTST2`.

The jet energy resolution and scale uncertainties show up on the following ranks. These are expected to play an important role as well, since they affect the definition of the jet categories. Further, the jet energy directly impacts the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, which is the discriminant chosen to extract the signal.

The dijet mass modeling parameter for the WZ background, `VVMbb_WZ`, is on the next rank, followed by Z+jets and W+jets modeling and normalization parameters.

The `METScaleSoftTerms` parameter, which is the uncertainty of the soft components in the $E_{\mathrm{T}}^{\mathrm{miss}}$, is ranked at place 14 with an impact of about 0.05 on $\hat{\mu}$. Apparently, the jet energy uncertainties, which have impacts around 0.10, are more important to this analysis. This is not surprising, since most of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is reconstructed from the jets of the hadronically decaying vector boson, which recoils against the invisibly decaying Higgs boson of the signal hypothesis.

7.6.5. Uncertainty breakdown

The breakdown of the uncertainties of $\hat{\mu}$ in the fit to the data, grouped by similar sources of uncertainties, is shown in Table 7.5. It is apparent that this analysis, opposed to the SM $VH(\to b\bar{b})$ analysis, is affected mostly by the systematic uncertainties, while the statistical uncertainty plays a minor role.

Further, the systematic uncertainty breaks down into a few dominant contributions. In order of decreasing importance: the modeling of the diboson background, the jet-energy uncertainties, and the modeling of the Z+jets and W+jets backgrounds. The other components contribute only little to the overall uncertainty.

A more detailed view at the major components of uncertainty is shown in Table 7.6. Here, the modeling uncertainties of the W+jets, Z+jets and $t\bar{t}$ backgrounds are each

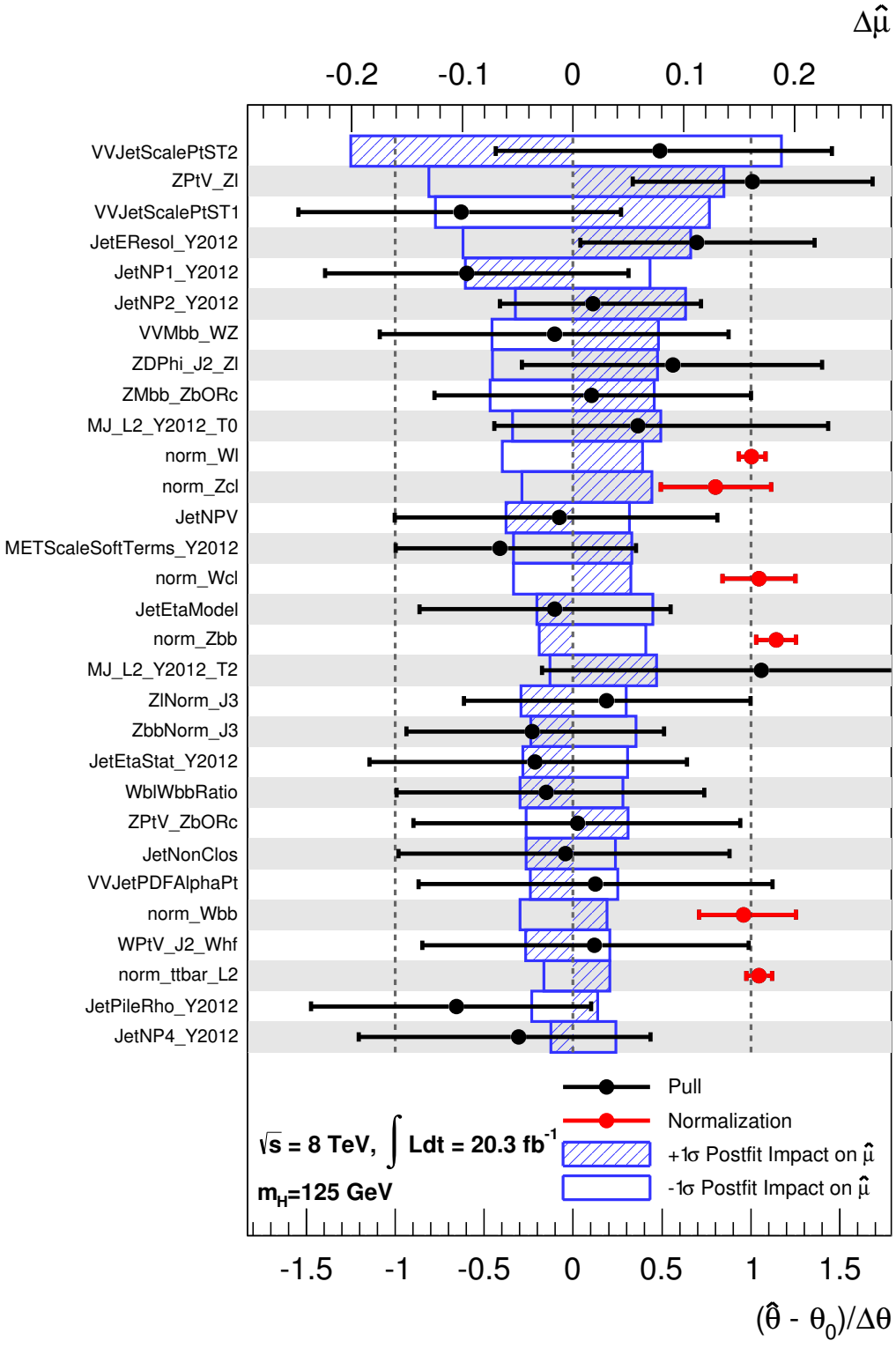


Figure 7.14.: Ranking of the post-fit impact from nuisance-parameter uncertainties on the observed signal strength $\hat{\mu}$ (blue areas). The pulls of the parameters from the fit to the data are shown as well (markers).

Experimental	
Source of unc.	$\sigma_{\hat{\mu}}$
Jets	0.22
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.06
b tag: b jets	0.03
c jets	0.02
light jets	0.02
Luminosity	0.04

Theory and modeling	
Source of unc.	$\sigma_{\hat{\mu}}$
W+jets	0.07
Z+jets	0.20
$t\bar{t}$	0.02
Single-top	0.02
Diboson	0.27
Multijet	0.07

Floating norm.	
Source of unc.	$\sigma_{\hat{\mu}}$
W+jets	0.10
Z+jets	0.07
$t\bar{t}$	0.04

Combined	
Systematic	0.42
Statistical	0.12
Total	0.44

Table 7.5.: Breakdown of $\sigma_{\hat{\mu}}$ into groups of uncertainty sources in the fit to the data. Various systematic components are given, as well as the total statistical and total uncertainties.

combined with the corresponding floating normalization uncertainty. This way their total impacts can be directly compared to total impact from the diboson background.

The uncertainties in the fit to the data are compared to the ones from a fit to the Asimov dataset. They are in good agreement, which shows that the data is not that far from the nominal MC expectation.

Several fits to the Asimov dataset are performed using various de-correlation schemes for $\hat{\mu}$. This allows to investigate the contributions of the uncertainty components in the various regions.

As mentioned, in the combined fit the uncertainty is dominated by the systematic component. This is not true for the 2-tag region, where the statistical uncertainty is of similar size. Similarly, in the 1-tag region the Z+jets background is, instead of diboson, the largest contributor to the uncertainty. This is also the case for the mass side-bands. The $t\bar{t}$ background plays a more important role in the 1- and 2-tag regions compared to the 0-tag region.

On the experimental side the jet energy uncertainties are the dominant component throughout all regions, except for 1-tag, where the b-tagging uncertainties are of similar size.

The contributions of the various regions to the sensitivity can be evaluated from the total uncertainty. As expected, the signal region is the main contributor, while the side bands are less sensitive (being mostly sensitive to the $gg \to H \to$ inv. signal). A similar behavior is observed comparing the 2-jet to the 3-jet regions.

Among the b-tagging regions the 0-tag region is the most sensitive, closely followed by 2-tag, while 1-tag contributes less. This is true when considering all signal processes in combination. For the VH signal process alone, in particular the ZH production, the 2-tag region is expected to be of more importance.

7.6.6. Compatibility of regions

The compatibility of $\hat{\mu}$ between the lepton channels, jet and b-tagging categories and the signal region and the dijet-mass side-bands is measured in the data, as explained in

	Data	Asimov	SR	SB	2-jet	3-jet	0-tag	1-tag	2-tag
Experimental									
Jets	0.22	0.22	0.25	0.46	0.19	0.47	0.29	0.23	0.17
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.06	0.07	0.01	0.19	0.05	0.17	0.09	0.03	0.05
b tagging	0.04	0.06	0.08	0.14	0.06	0.09	0.03	0.22	0.12
Modeling and floating norm.									
W+jets	0.16	0.13	0.16	0.31	0.12	0.26	0.16	0.20	0.10
Z+jets	0.21	0.21	0.26	0.54	0.21	0.62	0.26	0.40	0.19
$t\bar{t}$	0.05	0.07	0.10	0.20	0.06	0.21	0.05	0.20	0.15
Diboson	0.27	0.24	0.22	0.19	0.27	0.18	0.25	0.27	0.21
Combined									
Systematic	0.42	0.43	0.46	0.89	0.45	0.92	0.51	0.70	0.45
Statistical	0.12	0.12	0.14	0.25	0.14	0.25	0.13	0.38	0.48
Total	0.44	0.44	0.48	0.92	0.47	0.95	0.53	0.80	0.66

Table 7.6.: Breakdown of $\sigma_{\hat{\mu}}$ into groups of uncertainty sources in the fit to the data (first column) and to the Asimov dataset (remaining columns). The main systematic components are given, as well as the statistical and total uncertainty. Various de-correlation schemes for μ are shown: signal region (SR), side bands (SB); 2, 3-jet and 0, 1, 2-tag.

Regions	N_μ	Compatibility
0, 1, 2-lepton	3	91 %
0, 1, 2-tag	3	73 %
2, 3-jet	2	75 %
SR, SB	2	43 %

Table 7.7.: Compatibility of $\hat{\mu}$ between regions of the combined fit. Various de-correlation schemes for μ are shown: 0, 1, 2-lepton; 2, 3-jet; 0, 1, 2-tag and signal region (SR), side bands (SB).

Section 5.7.8. The results are shown in Table 7.7.

A good compatibility is observed for all of the de-correlation schemes. The 1- and 2-lepton channels have very little sensitivity to the signal, such that a low compatibility would be worrying, but this is not the case. This study was done blinded, meaning the actual $\hat{\mu}$-values measured in the various regions were not recorded.

7.7. Results

The post-fit numbers of predicted background events are compared to the observed number of events in the data for the six signal regions in Table 7.8. The predicted number of signal events are listed for for $m_H = 125\,\mathrm{GeV}$ and $\mathrm{BR}(H \to \mathrm{inv.}) = 100\,\%$. No significant excess of events over the expected SM background processes is observed in the data.

The combined fit gives a signal strength of $\hat{\mu} = -0.13^{+0.43}_{-0.44}$, which is consistent with zero.

	b-tag category	0-tag	1-tag	2-tag
	Process	2-jet events		
Background	Z+jets	24400 ± 1100	1960 ± 200	164 ± 13
	W+jets	20900 ± 770	1160 ± 130	47 ± 7
	$t\bar{t}$	403 ± 74	343 ± 65	57 ± 10
	Single top	149 ± 16	107 ± 14	11 ± 2
	Diboson	1670 ± 180	227 ± 25	64 ± 7
	SM VH($b\bar{b}$)	1.5 ± 0.5	6 ± 2	3 ± 1
	Multijet	26 ± 43	8 ± 7	0.7 ± 0.9
Total		47560 ± 490	3804 ± 64	347 ± 15
Signal	$gg \to H$	403 ± 95	25 ± 6	2.1 ± 0.5
	$W(\to jj)H$	425 ± 45	44 ± 6	0.6 ± 0.1
	$Z(\to jj)H$	217 ± 19	42 ± 4	26 ± 2
Data		47404	3831	344
		3-jet events		
Background	Z+jets	9610 ± 580	795 ± 93	53 ± 7
	W+jets	7940 ± 510	479 ± 70	21 ± 4
	$t\bar{t}$	443 ± 53	437 ± 53	63 ± 7
	Single top	97 ± 14	66 ± 9	6.4 ± 0.9
	Diboson	473 ± 54	55 ± 6	13 ± 2
	SM VH($b\bar{b}$)	0.8 ± 0.3	2.6 ± 0.9	1.4 ± 0.5
	Multijet	22 ± 29	4 ± 4	0.6 ± 0.6
Total		18580 ± 200	1840 ± 40	158 ± 7
Signal	$gg \to H$	224 ± 55	15 ± 4	1.2 ± 0.5
	$W(\to jj)H$	110 ± 16	11 ± 1	0.14 ± 0.03
	$Z(\to jj)H$	65 ± 7	12 ± 1	6.1 ± 0.7
Data		18442	1842	159

Table 7.8.: Predicted and observed numbers of events for the six signal regions. The yields and the combined statistical and systematic uncertainties of the backgrounds are shown after the combined fit to the data. The signal processes are listed for $m_H = 125\,\text{GeV}$ and $\text{BR}(H \to \text{inv.}) = 100\,\%$ (from Ref. [2]).

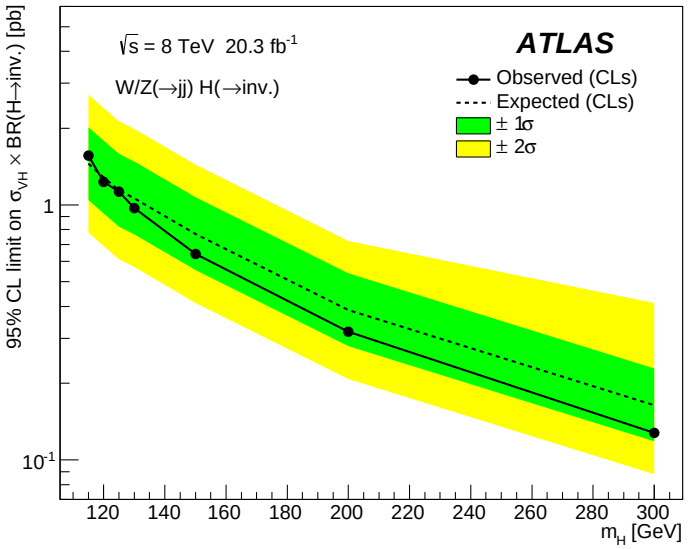


Figure 7.15.: 95 % CL upper limits on the VH cross section × BR($H \to$ inv.) as a function of the hypothesized Higgs boson mass (from Ref. [2]).

This result is obtained by considering all relevant SM production modes of the Higgs boson, VH and ggH, and a mass of $m_H = 125\,\text{GeV}$. It is translated into an observed (expected) upper limit of 78 % (86 %) at 95 % CL on the branching ratio of the Higgs boson to invisible particles, assuming the Higgs boson production as predicted in the SM. The gluon-fusion production mode contributes about 39 % (29 %) to the observed (expected) sensitivity, defined as unity divided by the limit squared.

Limits on the VH production cross section times the branching ratio for invisible decays are set as a function of m_H in the range $115 < m_H < 300\,\text{GeV}$. The results are shown in Figure 7.15. The observed 95 % CL upper limit on $\sigma_{VH} \times \text{BR}(H \to \text{inv.})$ varies from 1.6 pb at 115 GeV to 0.13 pb at 300 GeV.

At $m_H = 125\,\text{GeV}$ a limit of 1.1 pb is observed compared with 1.1 pb expected for the VH production mode. These results assume the proportions of the WH and ZH production cross sections as predicted in the SM. Observed (expected) limits are derived on the cross section times branching ratio for the two contributions separately as well: 1.2 pb (1.3 pb) for WH and 0.72 pb (0.59 pb) for ZH production.

This independent result is comparable to that of a similar search carried out with the ATLAS detector in the channel $ZH \to \ell\ell\chi\chi$. That analysis has set an observed (expected) 95 % CL upper limits on BR($H \to$ inv.) of 75 % (62 %) for $m_H = 125.5\,\text{GeV}$ [131].

The results are combined with other searches carried out with the ATLAS experiment, of which the vector-boson fusion production of the Higgs boson provides the largest part of the sensitivity. In the combination, an observed (expected) upper limit at the 95 % CL on BR($H \to$ inv.) of 25 % (27 %) is obtained in the direct search [48].

8. Search for VH resonances with H→bb

With the increased center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ for Run 2 of the LHC, physics beyond the standard model (BSM) might be in reach. New resonances, such as the resonant production of Standard Model (SM) boson-pairs, would be a clear hint of BSM effects. In the following, searches for VH resonances are discussed. The analysis is published as Ref. [3] and is one of a number of diboson-resonance searches carried out with the early Run 2 data [132].

A simplified model is used as benchmark for this search. The Heavy Vector Triplet (HVT) model, discussed in Section 2.6, predicts new heavy vector bosons W' and Z', each decaying into a SM vector boson ($V = W$ or Z boson) and a Higgs boson (H). Other decays are possible as well, but are not considered here. The final states where the vector boson decays leptonically and the Higgs boson decays to a b-quark pair are investigated: $W' \to WH \to \ell\nu b\bar{b}$, $Z' \to ZH \to \nu\nu b\bar{b}$ and $Z' \to ZH \to \ell\ell b\bar{b}$.

The search for W' with masses around $m_{W'} = 2\,\text{TeV}$ is of particular interest. For this signature a local excess corresponding to a significance of $3.4\,\sigma$ (standard deviations) was observed with data taken during the LHC Run 1 at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS experiment [58], and also the CMS experiment has reported an excess of $2.9\,\sigma$ [60]. Although these excesses were only seen in particular decay channels, as discussed in Section 2.6, such resonances are not ruled out completely.

The present study is carried out with data taken with the ATLAS detector during the year 2015 at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$, corresponding to an integrated luminosity of $\mathsf{L} = 3.2\,\text{fb}^{-1}$. The analysis strategy is common to many resonance searches. The decay products of a hypothesized resonance are reconstructed and the invariant mass is calculated. The signal is then expected as a peak on top of the continuous SM backgrounds.

8.1. The CxAOD framework

The present analysis is implemented using a novel software framework, the *CxAOD framework*, which has been developed in the course of this thesis. It is based on ROOT [120] and on the *xAOD* data format [133], developed within ATLAS for the Run 2 of the LHC. The xAOD format combines features from the old *Analysis Object Data* (AOD) and *Derived Physics Data* (DPD) formats. The AOD format was used in the ATLAS reconstruction software Athena [134]. Physics analyses, however, often preferred the DPD format, which were directly readable using ROOT.

The basic xAOD format contains a number of uncalibrated objects, such as electrons or jets, for each event. Further, a large amount of reconstruction-related information is stored, which is needed for the calibration of the objects.

The CxAOD format is the central piece of the framework. It is derived from the xAOD

CxAOD

Object	Nominal	JES
Jet 1	p_{T}, η, ϕ, ...	p'_{T}
Jet 2	p_{T}, η, ϕ, ...	p'_{T}
Muon 1	p_{T}, η, ϕ, ...	

N-tuple

Object	Nominal	JES
Jet 1	p_{T}, η, ϕ, ...	p'_{T}, η, ϕ, ...
Jet 2	p_{T}, η, ϕ, ...	p'_{T}, η, ϕ, ...
Muon 1	p_{T}, η, ϕ, ...	p_{T}, η, ϕ, ...

Table 8.1.: Comparison of the CxAOD format (left) with the N-tuple format (right) for the example of one jet energy scale (JES) variation in one event with two jets and one muon. The variation is stored as shallow copy in the CxAOD, overwriting only the p_{T} of the jets with new values p'_{T} (green). Other properties are referenced back to the nominal values (blue). This reduces the file size typically by a factor of eight compared to the N-tuple. The N-tuple uses more space by storing the nominal values several times (red).

format, but it contains calibrated objects (hence the "C"). This allows to drop the information needed in the calibration step and to store only information needed for physics analyses. This reduces the file size on disk significantly.

An important feature of the xAOD format are *shallow copies* (SCs). They allow to store copies of objects that are mostly empty. Only properties selected by the user are written, while other properties are referenced back to the original object. This is very useful for efficient storage of systematic variations.

For example, a jet energy scale (JES) variation affects only the p_{T} of jets, while all other jet properties (η, ϕ, m, ...) and other objects are unaffected. In the CxAOD a collection of jets with the nominal calibration and their associated properties is stored. The JES variation is stored as a shallow copy, overwriting only the p_{T} of jets. This is illustrated in Table 8.1. In the traditional *N-tuple* format, instead, full copies for each object and variation are stored, resulting in a larger file size.

A typical example consists of a sample of simulated $t\bar{t}$ events with 40 kinematic variations for various objects. When using N-tuples one would expect a factor of 41 in disk size comparing all variations to nominal only. In the CxAOD format the size increases only by a factor of five. Additionally, less computing time is needed for creating and reading the CxAOD. This yields a significant advantage for the analysis setup in Run 2.

Part of the framework is the *CxAODMaker* package. It contains the core software, which processes xAOD input files, applies the calibrations and writes CxAOD files. As mentioned above, unnecessary information (defined by the user) is removed from the objects. Additionally, objects and even events are removed completely if they do not pass a user-defined selection.

Due to the use of shallow copies the selection step is more complicated than with N-tuples. An object has to be written if it passes the selection for any of the variations, including the nominal one. The same is true for events. Due to this logic, more events have to be written when processing systematic variations than for nominal only. This overhead constitutes about 20 % of the disk size estimated for the $t\bar{t}$ example above.

The framework contains a *CxAODReader* package, which allows to read the CxAODs for further analysis in an efficient way. For each event a loop over all variations is performed,

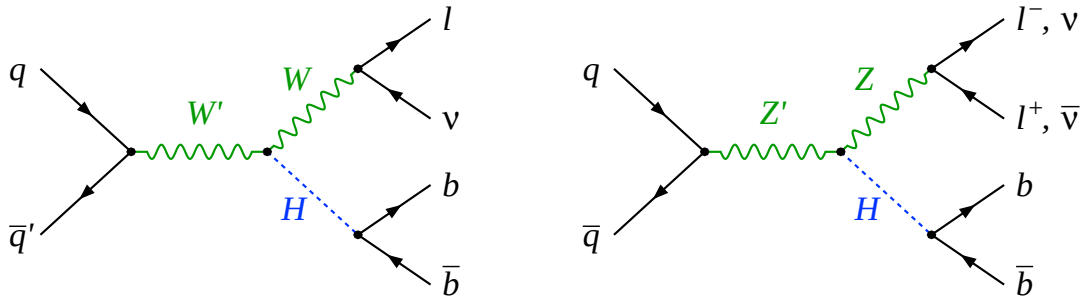


Figure 8.1.: Leading order Feynman diagrams of the $W' \to WH \to \ell\nu b\bar{b}$ (left) and $Z' \to ZH \to \ell\ell b\bar{b}$ and $Z' \to ZH \to \nu\nu b\bar{b}$ (right) signal processes.

during which histograms or other output for each variation is generated. Then the next event is processed. This is in contrast to most Run 1 software and makes best use of the CxAOD format by reducing the computing time.

Another advantage of the CxAOD format is the direct availability of systematic variations for any object in any event, which passed the nominal selection. This allows to easily fill two-dimensional histograms with some variation of any distribution against the nominal one. This can be used for validation, but more importantly, it would allow for a more sophisticated evaluation of the statistical uncertainty on systematic variations, as sketched in Section 5.4.3.

The CxAOD framework is the only framework in ATLAS to implement such an efficient data format in ROOT, according to an internal review in early 2016 by the *ATLAS Software Group* (ASG). It is not only used for the present analysis, but more than ten official derivations for other analyses exist. One other framework, which is based on Athena, implements a very similar format (PxAOD, "P" for physics) and was developed in collaboration with the CxAOD framework.

8.2. Signal and background processes

The signal process for this analysis, $V' \to VH(\to b\bar{b})$, is categorized by the decay of the V boson into the 0-, 1- and 2-lepton channels: $Z' \to ZH \to \nu\nu b\bar{b}$, $W' \to WH \to \ell\nu b\bar{b}$ and $Z' \to ZH \to \ell\ell b\bar{b}$. The corresponding leading order (LO) Feynman diagrams are shown in Figure 8.1.

Although the final state particles of the signal are the same as for the SM $VH(\to b\bar{b})$ analysis, kinematic distributions are quite different. The mass of the V' causes the final-state particles to be more boosted and is itself the most important discriminant in the analysis. It is reconstructed from the decay products of the V- and H-boson candidates, as described in Section 8.4.1.

The HVT signal is simulated with the MadGraph5_aMC@NLO v2.2.2 generator [135] using the NNPDF2.3LO PDFs [135], interfaced to Pythia 8.186 [87] using the A14 tune [136] for parton showering and hadronization. The Higgs boson in the signal process

is simulated using $m_H = 125\,\text{GeV}$ and allowing the $H \to b\bar{b}$ and $H \to c\bar{c}$ decays with $\text{BR}(H \to c\bar{c})/\text{BR}(H \to b\bar{b}\ ) = 5\,\%$, as predicted from SM calculations [43]. The signal samples are generated for resonance masses in the range of $0.7 \leq m_{V'} \leq 5\,\text{TeV}$.

The production of W or Z bosons with additional jets, W+jets or Z+jets, are among the major contributions to the SM background processes for this search. The corresponding matrix elements are calculated for up to two partons at NLO and four partons at LO using the Comix [137] and OpenLoops [138] generators and are merged with the Sherpa 2.1.1 [98] parton shower using the ME+PS@NLO prescription [139]. The CT10 PDF set [96] is used.

The diboson background processes, WW, WZ and ZZ, where one of the bosons decays hadronically and the other leptonically, are simulated using the same procedure as for the V+jets backgrounds. The matrix elements are calculated for up to one (ZZ) or no (WW, WZ) additional parton at NLO and up to three additional partons at LO.

The production of $t\bar{t}$ constitutes another major background process. It is simulated using the Powheg-BOX v2 [93, 94, 140] generator with the CT10 PDF set. The parton shower, fragmentation and the underlying event are simulated using Pythia 6.428 [99] with the CTEQ6L1 PDF set [141] and the Perugia P2012 tune [142]. The single-top Wt and s-channel processes are simulated using the same techniques as for the $t\bar{t}$ background. This is also true for the t-channel process, except that the Powheg-BOX v1 generator is used instead of v2.

The SM $VH(\to b\bar{b})$ process is considered as background for this search as well. It is simulated using Pythia 8.186 for the quark-induced production and using Powheg showered with Pythia 8.186 for the gluon-induced production. The CT10 PDFs and the AZNLO tune [143] are used in both cases.

8.3. Reconstruction of physical objects

The reconstruction of physical objects for the present analysis differs partially from the reconstruction for the Run 1 analyses, described in Chapter 4. Some of the changes are designed to cope with the specific kinematics of the signal process and others are due updates in the ATLAS software for Run 2. The differences, which affect mostly the jet reconstruction, are elaborated in the following.

8.3.1. Leptons

Leptons are categorized based on identification (ID) and isolation criteria as *loose* or *tight*, which are similar to the ones described in Section 4.2. The loose electron selection requires the LooseLH ID, while the TightLH ID is required for tight electrons. Both ID criteria are described in Ref. [72]. Similarly, the loose (tight) muon selection requires the Loose (Tight) ID, as described in Ref. [144]. As before, the minimum p_{T} for lepton candidates is 7 GeV.

Two different requirements on the lepton isolation are applied: LooseTrackOnly and Tight, as defined in the ATLAS reconstruction software. The Tight requirement applies p_{T} and η-dependent cuts on the track and calorimeter isolation: i_{trk} and i_{calo}, as defined

in Section 4.2. They are designed to accept 95 % of $Z \to \ell\ell$ events, while maximizing the rejection of non-prompt leptons.

The LooseTrackOnly requirement is based on the track isolation and employs a variable cone size, ΔR, as a function of the transverse momentum of the lepton as

$$\Delta R(\ell, trk) = \min\left\{R_{\text{min}}, \frac{10\,\text{GeV}}{p_{\text{T}}^{\ell}}\right\}, \tag{8.1}$$

where $R_{\text{min}} = 0.2$ (0.3) is chosen for electrons (muons). The p_{T}-dependency effectively tightens the isolation requirement for lower momenta, where more non-prompt leptons are expected, while loosening it for larger momenta.

For larger momenta high energetic photons from the leptons are more likely to be radiated. These photons can convert and create tracks close to the lepton and thus reduce the acceptance for a fixed-radius isolation requirement. Using the variable cone size for the definition of i_{trk} these events can be recovered. The cut on i_{trk} is then designed as a function of p_{T} and η to accept 99 % of $Z \to \ell\ell$ events, while maximizing the rejection of non-prompt leptons.

Systematic uncertainties for the leptons are taken into account on their trigger, reconstruction, identification and isolation-requirement efficiencies. Further, energy scale and resolution uncertainties are taken into account. All of them, listed in Section 8.3.5, have very little impact on the search results, as shown in Section 8.8.3.

8.3.2. Small jets

Calorimeter jets that are reconstructed using the anti-k_t algorithm with $R = 0.4$, as described in Section 4.3, are referred to as *small jets* in this analysis. They are used for the overlap removal with electron and muons, as described in Section 4.5, and in the calculation of the missing transverse energy, as described in Section 4.6. For the present analysis, calorimeter jets are reconstructed from noise-suppressed topological clusters [145].

As for the Run 1 analysis, systematic uncertainties on the jet-energy scale and resolution are taken into account. A reduced set of three eigenvector variations is employed for the scale uncertainties and one parameter for the resolution, as listed in Section 8.3.5. These uncertainties do not have a large impact on the search results, as they have only an indirect effect through the propagation to the overlap removal and to the $E_{\text{T}}^{\text{miss}}$ calculation.

8.3.3. Large jets

The use of small jets for reconstructing the Higgs boson candidate is not efficient for the HVT signal process. The average distance ΔR between the b quarks from the Higgs boson decay of the signal decreases with an increasing resonance mass. Hence, the b quarks cannot be reconstructed in separate small jets with $R = 0.4$ for very large $m_{V'}$. This effect starts to be important around $m_{V'} = 1\,\text{TeV}$ and is illustrated in Figure 8.2.

It is not feasible to decrease the R parameter of the jet algorithm due to the granularity of the calorimeters. Instead, the anti-k_t algorithm with $R = 1.0$ is used to reconstruct the Higgs boson candidate within one large jet. Such large jets typically contain significant energy depositions from pile-up interactions or other radiation in the event. To reduce

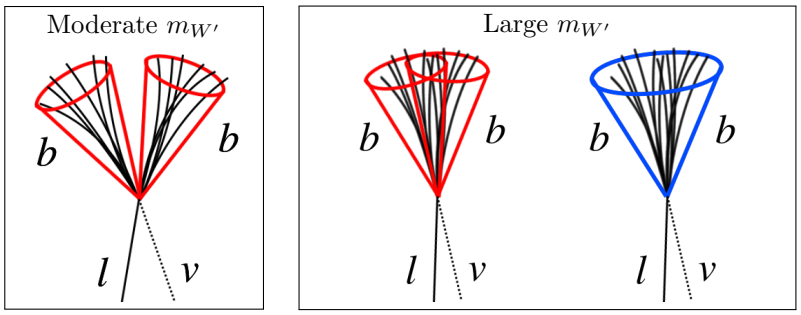


Figure 8.2.: Illustration of the b-jet reconstruction for the $W' \to WH \to \ell\nu b\bar{b}$ signal. Left: the b jets are well separated for moderate $m_{W'}$ and can be reconstructed in two separate jets (red). Right: the b jets merge for large $m_{W'}$ and need to be reconstructed in one large jet (blue).

the noise from such effects the jets are *groomed*. The chosen grooming technique is the *trimming* algorithm [146].

Trimming takes the original constituents of the large jet and reclusters them using the k_t algorithm [147] with a smaller distance parameter, R_{sub}, to produce a collection of subjets. Subjets that carry less than a specific fraction, f_{cut}, of the original jet p_{T}, are removed. Finally, the large-jet four-momenta are recomputed from the selected subjets. The trimming parameters, optimized for identifying hadronic W or Z boson decays [148], are $R_{\text{sub}} = 0.2$ and $f_{\text{cut}} = 5\,\%$ and are used in the present analysis as well.

The energies and masses of the large jets are corrected for energy losses in passive material, for the non-compensating response of the calorimeter, and for any additional energy due to multiple pp interactions [149].

The scale and resolution uncertainties on the jet energy and mass are evaluated by comparing the calorimeter-based to track-based measurements in multijet events selected in the data to the simulation [150]. These uncertainties are among the most important ones for the present analysis. Their impact is typically between 2 to 20 % on the normalization and on differential distributions of the background processes.

The groomed and calibrated large jets are required to satisfy $p_{\text{T}} > 250\,\text{GeV}$ and their rapidity range is limited to $|\eta| < 2.0$. Typical jet substructure variables, such as energy correlation functions, e.g. the D_2 variable [151, 152], did not show significant separation after applying b tagging (described in the next section) and are therefore not used.

8.3.4. Track jets and b-tagging

Track jets are reconstructed using the anti-k_t algorithm with $R = 0.2$. They are used to apply b tagging for identifying the $H \to b\bar{b}$ candidate. Track jets can be reconstructed with a smaller R parameter than calorimeter jets due to the high granularity of the tracking detector. Tracks with $p_{\text{T}} > 400\,\text{MeV}$ associated with the primary vertex are used as input to the jet algorithm.

The track jets are required to satisfy $p_{\text{T}} > 10\,\text{GeV}$ and $|\eta| < 2.5$. Further, the MV2c20 b-tagging algorithm [81, 153] with an tagging efficiency of 70% is applied. It has a rejection factor of about 5.6 (180) for jets containing c hadrons (only light-quark hadrons) in a

simulated sample of $t\bar{t}$ events.

The track jets are associated to the large jets via ghost-association [75]. The ghost-association is performed by rerunning the algorithm for the calorimeter jets, but using the track jets as input in addition. However, the p_{T} of the tracks are set to zero, such that they do not affect the shape of the reconstructed jet. This results in the same calorimeter jet as before, but this time each track jet is uniquely associated to one of them.

The calibration of the b-tagging efficiency is taken from Run 1, as described in Section 4.4, due to insufficient data from Run 2 for an updated calibration at the time of the analysis. A generic extrapolation uncertainty, and a specific one for c jets, is applied to cover possible inconsistencies. The b-tagging uncertainties have a typical impact of 5 to 15 % on the normalization of the various background processes.

8.3.5. List of experimental uncertainties

The full list of experimental uncertainties, taken into account in the present analysis, is given in Table 8.2. The uncertainties are related to the reconstruction of physical objects, as described above. A name is assigned to each uncertainty, which is used for its identification in the statistical treatment, as described in Section 8.5.

8.4. Event selection

The event selection for the search for the $V' \to VH(\to b\bar{b})$ process is described in the following. It is partially adapted from an earlier study carried out by the ATLAS collaboration with the data taken at $\sqrt{s} = 8\,\mathrm{TeV}$ [154].

However, an important different exists: the earlier search employed small jets for reconstructing the Higgs boson candidate. Instead, for the present analysis the Higgs boson candidate is reconstructed as one large jet, as described in Section 8.3.3, resulting in a better signal acceptance for large resonance masses.

8.4.1. Selection requirements

Event cleaning The first step of the event selection consists of standard cleaning cuts, which are similar to the ones in the SM $VH(\to b\bar{b})$ analysis (Section 5.2). An updated Good Run List for the data taken during the year 2015 is used. The requirement on the primary vertex has been loosened to contain at least two tracks (instead of three).

As for the previous analyses, the event selection is split into 0-, 1- and 2-lepton channels, targeting the leptonic decay modes of the W and Z bosons. However, the 0-lepton selection has a sizable acceptance for the $W' \to WH \to \ell\nu b\bar{b}$ signal of about 10 % compared to the 1-lepton selection. It is used in addition when searching for the W' signal alone, assuming zero abundance of the Z' signal. This is further discussed in Section 8.5.

Triggers The triggers have changed partially with respect to the Run 1 analyses. The 0-lepton analysis still uses a trigger based on $E_{\mathrm{T}}^{\mathrm{miss}}$ with the same threshold of $E_{\mathrm{T}}^{\mathrm{miss}} > 80\,\mathrm{GeV}$. Both the 1- and 2-lepton analyses employ single electron and muon triggers. Three electron triggers with increasing p_{T} thresholds of 24, 60 and 120 GeV and decreasing

Name	Source	Count
Luminosity (1)		
Luminosity	Total integrated luminosity	1
Electrons (3)		
EL_EFF_ID_TotalCorrUnc.	Trigger, reco., and ID efficiencies	1
EG_SCALE_ALL	Energy scale	1
EG_RESOLUTION_ALL	Energy resolution	1
Muons (7)		
MUON_EFF_STAT/SYS	Trigger, reco., and ID efficiencies	2
MUONS_SCALE	Energy scale	1
MUON_ID	Energy resolution from inner detector	1
MUON_MS	Energy resolution from muon system	1
MUON_ISO_STAT/SYS	Isolation scale factors	2
Small jets (4)		
JET_GroupedNP_x	EV decomposition of energy scale (x = 1-3)	3
JET_JER_SINGLE_NP	Energy resolution	1
Large jets (5)		
JET_Rtrk_Baseline	Ratio of calo. to track meas.: baseline	1
JET_Rtrk_Modelling	Ratio of calo. to track meas.: modeling	1
JET_Rtrk_Tracking	Ratio of calo. to track meas.: tracking	1
FATJET_JER	Energy resolution	1
FATJET_JMR	Mass resolution	1
$E_{\mathrm{T}}^{\mathrm{miss}}$ (3)		
MET_SoftTrk_ResoPara/Perp	Resolution of soft component	2
MET_SoftTrk_Scale	Scale of soft component	1
Flavor Tagging (13)		
EF_EFF_Eigen_B_x	EV decomp. of b-jet tagging eff. (x = 0-2)	3
EF_EFF_Eigen_C_x	EV decomp. c-jet tagging eff. (x = 0-3)	4
EF_EFF_Eigen_L_x	EV decomp. light-jet tagging eff. (x = 0-3)	4
EF_EFF_extrapolation	Extrapolation Run 1 to Run 2	1
EF_EFF_extr._from_charm	Charm quark specific extrapolation	1
Total	36 with priors, 0 floating	

Table 8.2.: Names and sources of experimental systematic uncertainties. The last column states for each uncertainty the number of corresponding parameters in the combined fit.

identification criteria are used. For the muon selection two triggers with p_{T} thresholds of 20 and 50 GeV are used. The electron and muon triggers with the lowest p_{T} threshold apply requirements on the lepton isolation.

The 1-lepton channel additionally accepts $E_{\mathrm{T}}^{\mathrm{miss}}$ triggered events to recover some of the efficiency loss in the central η region for muons. A requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV is applied for these events, such that the scale factors for the trigger efficiency are close to one.

Jets Each channel requires at least one large jet with $p_{\mathrm{T}} > 250\,\mathrm{GeV}$ and $|\eta| < 2.0$. If more than one large jet is present in the event, the leading jet in p_{T} is used as Higgs boson candidate. It is referred to as "the large jet" in the following.

The track jets that are within the large jet are used to define the number of b tags (0, 1 or 2 tags) in the event, based on the MV2c20 algorithm. These track jets are also used to categorize the V+jets background into flavor components. In the presence of a b hadron (c hadron), ghost associated to a track jet, the event is labeled as V+b (V+c), otherwise as V+light. Additional b tags (0 or at least 1) are counted for the track jets that are outside of the large jet.

The mass of the large jet, m_{jet}, is an important discriminant, as it is expected around the SM-Higgs mass of 125 GeV for the signal. The distributions of m_{jet} after the fit to the data are shown in Figure 8.3. Cuts on the jet mass are applied to define the signal region for jet masses of $75\,\mathrm{GeV} < m_{\mathrm{jet}} < 145\,\mathrm{GeV}$. The low- and high-mass side-bands are used as control regions in the fit.

Lepton channels For the 0-lepton channel any loose lepton in the event is vetoed and $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\mathrm{miss}} > 30\,\mathrm{GeV}$ are required. Further, angular cuts are applied to reject the multijet background: $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, p_{\mathrm{T}}^{\mathrm{miss}}) < \pi/2$, $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \text{small jets}) > 20°$ and $\Delta\phi(E_{\mathrm{T}}^{\mathrm{miss}}, \text{large jet}) > 120°$. The invariant mass of the $Z' \to ZH \to \nu\nu b\bar{b}$ system can only be partially reconstructed for the 0-lepton channel and is defined as

$$m_{\mathrm{T},VH} = E_{\mathrm{T}}^{\mathrm{miss}} + \sqrt{p_{\mathrm{T},J}^2 + m_{\mathrm{jet}}^2} \tag{8.2}$$

The 1-lepton selection requires one tight electron or one tight muon with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and vetoes any additional loose lepton. In addition, a cut of $E_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$ is applied to reject the remaining multijet background to a negligible level.

To fully reconstruct the invariant mass of the $W' \to WH \to \ell\nu b\bar{b}$ system the momentum of the neutrino in the z-direction, $p_{z,\nu}$, is required. The transverse components are indirectly measured and combined in the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$. However, this cannot be done for the longitudinal component, since the initial-state momenta of the partons are a-priory unknown at a hadron collider. However, it can be obtained from imposing the known W-boson mass as constraint on the lepton-neutrino system, which is described in Section 8.4.3. The mass of the WH system, m_{WH}, is then reconstructed from the four-vector sum of the lepton, the neutrino and the Higgs boson candidate jet.

For the 2-lepton selection exactly two loose electrons or two loose muons with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and opposite charge are required. One of them must pass the tight identification with $p_{\mathrm{T}} > 60\,\mathrm{GeV}$. The invariant mass of the two leptons is required to be within $70\,\mathrm{GeV} < m_{ee} < 110\,\mathrm{GeV}$ for electrons and $55\,\mathrm{GeV} < m_{\mu\mu} < 125\,\mathrm{GeV}$ for muons.

The muon momentum resolution deteriorates significantly for large p_{T}. In order to improve the resolution of large m_{VH}, the four-momentum of the di-muon system is scaled by $m_Z/m_{\mu\mu}$ with $m_Z = 91.2\,\mathrm{GeV}$. The ZH system mass, m_{ZH}, is then the mass of the four-momentum sum of the dilepton system and the $H \to b\bar{b}$ candidate jet.

A $t\bar{t}$ control region is defined for the 2-lepton channel by requiring leptons of different flavor, i.e. one electron and one muon, with opposite charges.

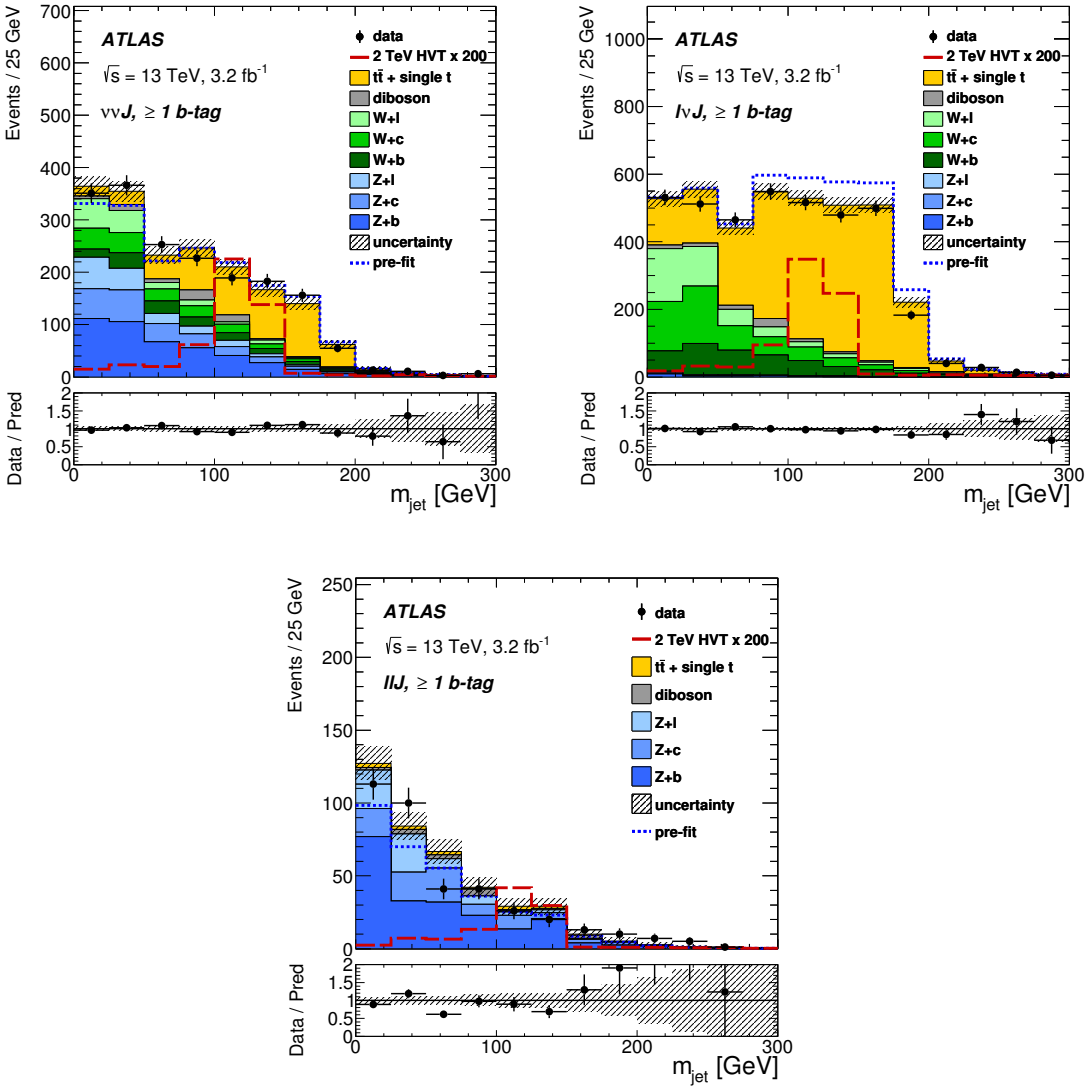


Figure 8.3.: Distributions of m_{jet} for the 0-lepton (upper left), 1-lepton (upper right) and 2-lepton channels (bottom) with at least one b tag. The distributions are shown after the fit to the data. The expected HVT signal with $m_{V'} = 2\,\text{TeV}$ is shown as dashed red histogram and is normalized to 200 times the corresponding expected 95 % CL upper cross section limit. A formal description of the plot elements is given in Section 8.8.1 (from Ref. [3]).

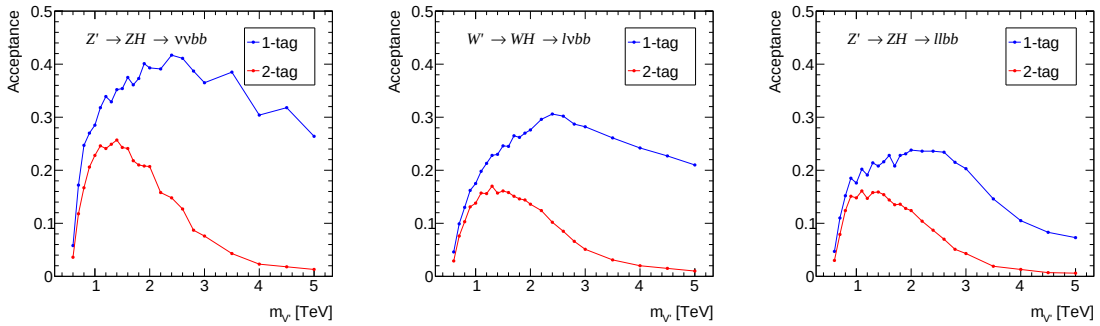


Figure 8.4.: Acceptance for the (left to right) $Z' \to ZH \to \nu\nu b\bar{b}$, $W' \to WH \to \ell\nu b\bar{b}$ and $Z' \to ZH \to \ell\ell b\bar{b}$ signal processes with the 0-, 1- and 2-lepton selections, respectively, as a function of the resonance mass, $m_{V'}$.

8.4.2. Signal acceptance

The selection efficiency for the signal in the three lepton channels after b tagging and without cut on m_{jet} as a function of the resonance mass, $m_{V'}$, is shown in Figure 8.4

For smaller masses of $m_{V'} \lesssim 1\,\text{TeV}$ the acceptance is low, since the two b quarks from the Higgs boson decay are too distant to be reconstructed in one large jet. On the other hand, the acceptance with two b tags drops for larger masses of $m_{V'} \gtrsim 2\,\text{TeV}$, since the two b quarks come too close to be reconstructed with two separate track jets. The 1-tag region helps to recover the acceptance for very large masses.

The cut on the large-jet mass of $75\,\text{GeV} < m_{\text{jet}} < 145\,\text{GeV}$ has an signal efficiency of about 90 % for smaller resonance masses of $m_{V'} \leq 2\,\text{TeV}$ and about 80 % for $m_{V'} = 5\,\text{TeV}$, counting the 1- and 2-tag events inclusively.

8.4.3. Neutrino reconstruction

The reconstruction of the neutrino momentum in the 1-lepton channel, $W' \to WH \to \ell\nu b\bar{b}$, and its optimization are described in the following. These studies have been carried out in the course of this thesis.

It is not possible to determine the invariant mass of the $W' \to WH \to \ell\nu b\bar{b}$ system solely from the objects reconstructed in the detector, due to the neutrino in the final state. The transverse components of the neutrino momentum are indirectly measured and combined in the missing transverse energy, $E_{\text{T}}^{\text{miss}}$. However, this cannot be done for the longitudinal component, since the initial-state momenta of the partons are a-priory unknown at a hadron collider.

Instead, the momentum of the neutrino in the z-direction, $p_{z,\nu}$, is obtained by imposing the known W-boson mass as constraint on the lepton-neutrino system. The four-momentum conservation gives:

$$p_W^2 = m_W^2 = (p_\nu + p_l)^2. \tag{8.3}$$

Squaring this equation and solving for $p_{z,\nu}$ with $m_l = m_\nu = 0$ one obtains:

$$p_{z,\nu} = \frac{\mu p_{z,l}}{p_{\mathrm{T},l}^2} \pm \sqrt{\frac{\mu^2 p_{z,l}^2}{p_{\mathrm{T},l}^4} - \frac{p_l^2 p_{\mathrm{T},\nu}^2 - \mu^2}{p_{\mathrm{T},l}^2}} \tag{8.4}$$

$$\text{with} \qquad \mu = \frac{1}{2} m_W^2 + p_{\mathrm{T},l} p_{\mathrm{T},\nu} \cos\Delta\phi(p_l, p_\nu). \tag{8.5}$$

Here, the transverse momentum of the neutrino is taken as the missing transverse momentum, $p_{\mathrm{T},\nu} = E_{\mathrm{T}}^{\mathrm{miss}}$, and its direction is taken as the direction of $\boldsymbol{E}_{\mathrm{T}}^{\mathrm{miss}}$. Further, the known mass of the W boson is used, $m_W = 80.4\,\mathrm{GeV}$. The remaining quantities are derived from the measured lepton four-momentum.

As consequence of the quadratic equation, either two real or two complex solutions for $p_{z,\nu}$ are possible. In case of complex solutions $p_{z,\nu}$ is taken as its real component and the imaginary part is set to zero. In case of real solutions $p_{z,\nu}$ is taken as the one that minimizes $|p_{z,\nu}|$. The mass of the WH system, m_{WH}, is then reconstructed as the mass of the four-momentum sum of the lepton, the neutrino and the $H \to b\bar{b}$ candidate jet.

Longitudinal momentum optimization The choice of using the minimal $|p_{z,\nu}|$ proofed to be better than several other options, which are discussed in the following. These studies have been carried out in an earlier stage of the analysis. Hence, the numbers for the resolution in the reconstructed mass spectra do not reflect the final result, but the outcome of the optimization is used in the final analysis.

The resolution in the reconstructed m_{WH} spectra of the W' signal with $m_{W'} = 1\,\mathrm{TeV}$ and $\sigma_{W'} \times \mathrm{BR}(W' \to WH \to \ell\nu b\bar{b}) = 1\,\mathrm{pb}$ are used as benchmark. Additionally, the W/Z+jets, $t\bar{t}$, single-top and SM $VH(\to b\bar{b})$ background processes are taken into account for sensitivity estimates.

The selection for this study deviates slightly from the final one. It requires exactly one signal electron or muon and exactly zero loose leptons. Furthermore, exactly one large jet with $p_T > 250\,\mathrm{GeV}$ is required and, due to technical reasons, at least one small jet. The large and small jets are allowed to overlap. The b tagging is applied to the small jets that are within $\Delta R = 1.0$ of the large jet and the MV2c00 algorithm with a tagging efficiency of 70 % is used. The requirements on these jets are $N_{\mathrm{jets}} = 1$ and $N_{\mathrm{tags}} = 1$ or $N_{\mathrm{jets}} \geq 2$ and $N_{\mathrm{tags}} = 2$. Further selection cuts are $p_T^W > 120\,\mathrm{GeV}$ and $95\,\mathrm{GeV} < m_J < 140\,\mathrm{GeV}$.

The large jet is corrected using a simple Higgs boson mass constraint: its four-vector is scaled by m_H/m_J with $m_H = 125\,\mathrm{GeV}$. This correction improved the resolution of $m_{W'}$ for this study, as shown in the following. However, it did not show an improvement in the final analysis, as a more evolved jet-energy scale correction is applied, and is not used for the final results.

The momentum of the neutrino in the z-direction, $p_{z,\nu}$, is obtained by imposing the W-boson mass constraint on the lepton and neutrino system, as described above. In case of complex solutions $p_{z,\nu}$ is taken as its real component and the imaginary part is set to zero. In case of two real solutions various choices for the neutrino $p_{z,\nu}$ are defined:

- Set $p_{z,\nu}$ to zero. Only the transverse components are used.
- Use $p_{z,\nu}$ from the W-mass constraint with the smaller $|p_{z,\nu}|$.

- Use $p_{z,\nu}$ from the W-mass constraint with the larger $|p_{z,\nu}|$.
- Using the output of a regression BDT, which uses the two solutions from the W-mass constraint as input and the truth $p_{z,\nu}$ as target.
- The $p_{z,\nu}$ from the W-mass constraint that is closer to the truth $p_{z,\nu}$.
- Set $p_{z,\nu}$ to the truth $p_{z,\nu}$.

Here, the truth $p_{z,\nu}$ refers to the longitudinal momentum of the neutrino given by the Monte Carlo generator for the signal events. This quantity can clearly not be used in the analysis of the data, but it provides a benchmark for the best solution possible.

As before, the mass of the WH system, m_{WH}, is reconstructed from the lepton, the neutrino and the Higgs boson candidate jet. The m_{WH} spectra of the signal, using the various solutions for the neutrino $p_{z,\nu}$, are shown in Figure 8.5.

The peak width is estimated by the RMS of the histogram and by the σ of a Gaussian fit. The fit uses a two step procedure with different fit ranges: in the first step the fit is performed in the range of mean $\pm$ RMS of the histogram. In the second step the fit is done again, this time in the range of mean $\pm\, 1.5\,\sigma$ of the first fit.

The signal is compared to the SM backgrounds in Figure 8.6. These spectra are used to estimate the statistical significance from a binned log-likelihood ratio as $\sqrt{LLR}$ with

$$LLR = \sum_{i=1}^{N} s_i \ln\left(1 + s_i/b_i\right), \tag{8.6}$$

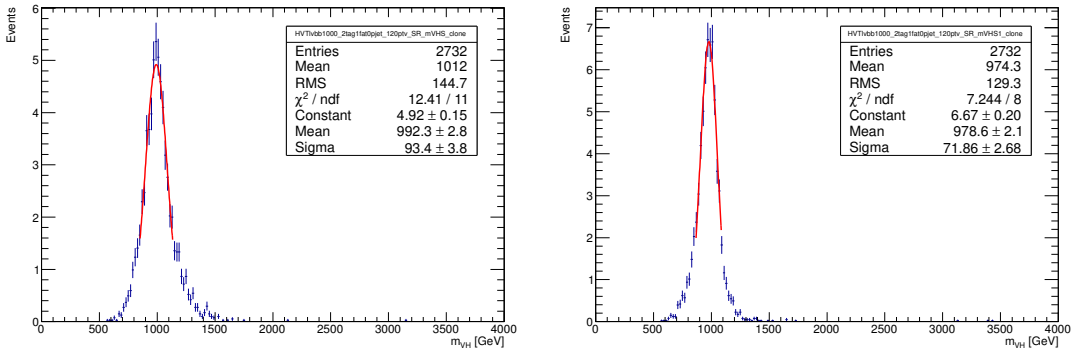
where the index i runs over all bins of the signal, s_i, and background, b_i, histograms. It can be noted that all choices for the neutrino $p_{z,\nu}$ broaden the background spectrum compared to setting it to zero.

The results are summarized in Table 8.3. It is found that the Higgs boson mass constraint improves the resolution, as well as the statistical significance in all cases (which is not the case for the final analysis).

The best choice for the neutrino $p_{z,\nu}$ is the one with smaller $|p_{z,\nu}|$. It shows the smallest σ and largest significance, which are improved by about 23 % and 6 %, respectively, compared to setting $p_{z,\nu}$ to zero. This result can be of statistical origin: the distribution of the truth $|p_{z,\nu}|$ is clustered towards low values, such that choosing a smaller value is more likely to be correct than a larger value.

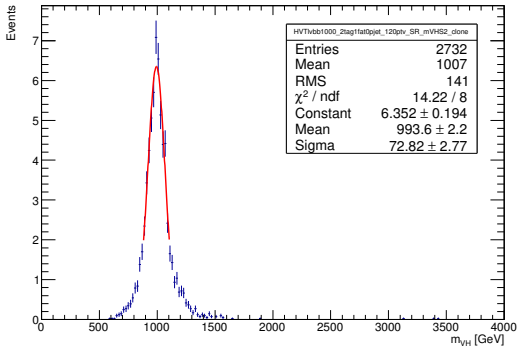
It can be noted that the regression BDT achieves a smaller RMS than the solution with smaller $|p_{z,\nu}|$. However, the resulting σ is larger and the significance is smaller. A different configuration of the BDT with more input variables (e.g. $\Delta\Phi(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$) in addition to the two solution for the neutrino $p_{z,\nu}$ has been tried. However, no gain in performance has been found. The training of the BDT employs gradient boosting with a specific loss function, which might not be optimal in noisy settings [105]. Possibly, a different definition of the loss function could lead to a better result.

Choosing the neutrino $p_{z,\nu}$ that is closer to the truth $p_{z,\nu}$ or the truth $p_{z,\nu}$ itself improve the resolution further by about 10 % and 25 %, respectively, compared to the solution with the smaller reconstructed $|p_{z,\nu}|$. These numbers provide an estimate for the maximal gain from a further improved neutrino $p_{z,\nu}$ reconstruction.

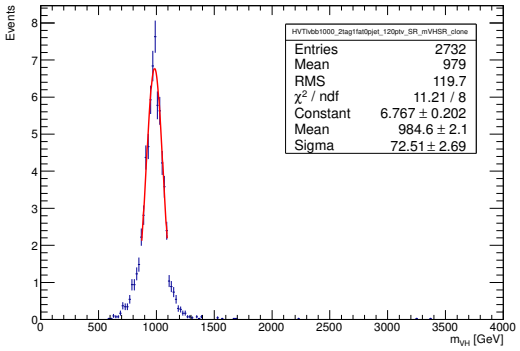


(a) Zero $p_{z,\nu}$

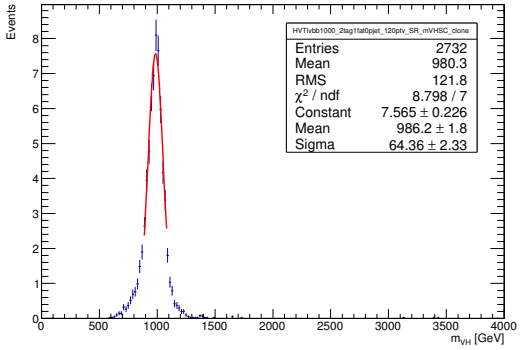
(b) Smaller $|p_{z,\nu}|$



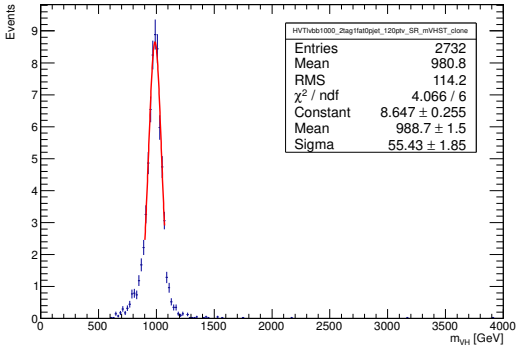
(c) Larger $|p_{z,\nu}|$



(d) Regression $p_{z,\nu}$



(e) Closer $p_{z,\nu}$ to truth



(f) Truth $p_{z,\nu}$

Figure 8.5.: Reconstructed m_{VH} spectra for the W' signal with $m_{W'} = 1\,\text{TeV}$ using various solutions for the neutrino $p_{z,\nu}$. The large jet four-vector is scaled by m_H/m_J. The peak width is estimated by the RMS of the histogram and by the σ of a Gaussian fit (red) with varying fit ranges, as described in the text. The fit parameters are given in the legend.

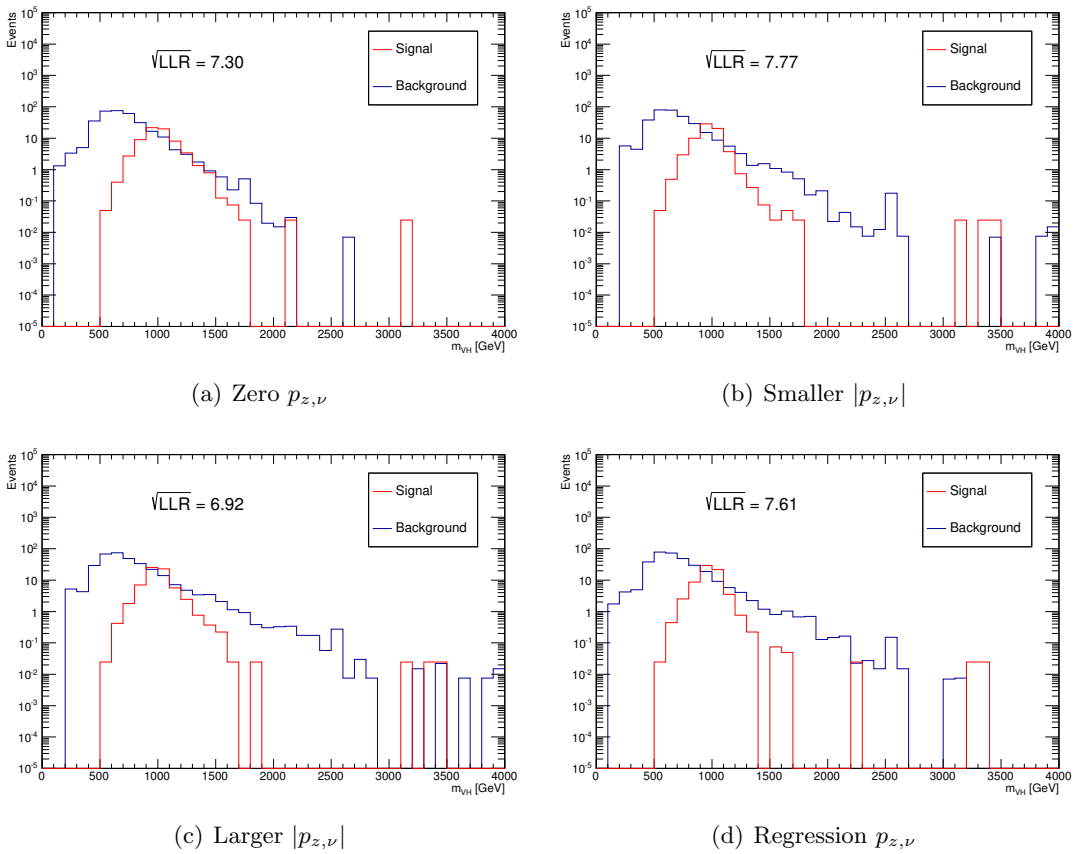


Figure 8.6.: Reconstructed m_{VH} spectra for the W' signal with $m_{W'} = 1\,\text{TeV}$ (red) and the sum of backgrounds (blue) using various solutions for the neutrino $p_{z,\nu}$. The large jet four-vector is scaled by m_H/m_J. The statistical significance is estimated by a binned log likelihood ratio as $\sqrt{LLR}$.

m_H	$p_{z,\nu}$	Mean ±	σ	RMS	$\sqrt{LLR}$
No	Zero	953.0 ±	109.0	146	7.06
No	Smaller	937.1 ±	87.1	130	7.39
No	Larger	958.9 ±	81.7	141	6.66
Yes	Zero	992.2 ±	93.4	145	7.30
Yes	Smaller	978.5 ±	71.9	129	7.77
Yes	Larger	993.5 ±	72.8	141	6.92
Yes	Regression	984.6 ±	72.5	120	7.61
Yes	Closer to truth	986.2 ±	64.4	122	—
Yes	Truth	988.7 ±	55.4	114	—

Table 8.3.: Benchmark numbers of various choices of neutrino $p_{z,\nu}$ for the W' signal with $m_{W'} = 1\,\text{TeV}$. The results are shown with ("Yes") and without ("No") Higgs boson mass constraint. Listed are the mean and width, σ, of Gaussian fits to the reconstructed m_{VH} spectra, the RMS of the spectra and the significance estimated as $\sqrt{LLR}$ considering the SM backgrounds.

N	Comment	Leptons	b tags	Add. tags	m_{jet}
6	Signal regions	0, 1, 2	1, 2	0	medium
3		0, 1, 2	2	0	low+high
4	V+jets CRs	0, 1	1	0	low, high
1		2	1	0	low+high
2	$t\bar{t}$ CRs	0, 1	1	1	low+high
2	$t\bar{t}$ $e\mu$ CRs	2	1, 2	incl.	medium

Table 8.4.: List of the 18 regions entering into the combined V' fit. In total 6 signal regions (SRs) and 12 control regions (CRs) are used. Low, medium and high m_{jet} refers to events with $m_{\text{jet}} < 75\,\text{GeV}$, 75 - 145 GeV and $> 145\,\text{GeV}$, respectively.

In conclusion, choosing the solution with smaller $|p_{z,\nu}|$ proofed to be the best option out of the investigated ones. The regression BDT performs worse, but could possibly be improved with a different definition of the loss function. The unphysical cases with imaginary solutions could also be looked at. For example a kinematic fit could be performed for each event, correcting the jet momenta with propagation to $E_{\text{T}}^{\text{miss}}$, to obtain a real solution.

8.5. Statistical treatment

The statistical analysis of the data is performed using a binned profile-likelihood fit, as described in Section 5.4. The largest differences are the choice of prior uncertainties ("priors" in the following) for the V+jets and $t\bar{t}$ backgrounds and the updated systematic uncertainties for Run 2, as described in the next section. The final discriminant is the invariant mass of the VH system, m_{VH}.

The parameter of interest is defined as the signal strength relative to the production cross section times branching ratio as predicted. It is the only freely floating parameter in the fit and is named `mu`.

Six signal regions are used in the final fit: 0-, 1- and 2-lepton with 1 or 2 b tags and medium jet mass, corresponding to $75\,\text{GeV} < m_{\text{jet}} < 145\,\text{GeV}$. The low and high m_{jet} side bands are used as eight control regions for the V+jets backgrounds, as listed in Table 8.4. The four $t\bar{t}$ control regions are defined with additional b tags or the different-flavor lepton selection. The choice of regions is the outcome of an optimization, as described in Section 8.7.

Three configurations are employed for extracting limits on the HVT signal cross sections. The W' and Z' hypotheses are tested separately by assuming no production of the corresponding other process. The limits on W' production are derived using the 1-lepton regions, where the largest acceptance is expected, and the 0-lepton regions for recovering events where the lepton from the W boson decay is not reconstructed. The 0- and 2-lepton regions are used in the fit for testing the Z' hypothesis. The 0-lepton region is sensitive to the $Z \to \nu\nu$ decays and the 2-lepton region to the $Z \to \ell\ell$ decays. Finally, limits on the HVT model cross sections are derived for both, the W' and Z' production modes, in a combined V' fit using all three lepton channels. The relative contributions from W' and

Name	Process	Value	Count
`XS_Wb/c/l`	W+jets	30/30/10 %	3
`XS_Zb/c/l`	Z+jets	30/30/10 %	3
`XS_ttbar`	$t\bar{t}$	30 %	1
`XS_st`	Single-top	30 %	1
`XS_diboson`	Diboson	11 %	1
`XS_vh`	SM $VH(\to b\bar{b})$	40 %	1
Total	9 with priors, 0 floating		

Table 8.5.: Nuisance parameters and their priors for the normalization uncertainties of the backgrounds.

Z' are taken from the HVT model.

8.6. Systematic uncertainties

The list of systematic uncertainties, which are taken into account in the fit as nuisance parameters, is given in the following.

8.6.1. Experimental

All experimental systematic uncertainties as listed in Section 8.3.5 are taken into account. They are propagated to the m_{VH} distributions and parameterized as shape NPs, correlated in all regions.

8.6.2. Signal and backgrounds

The normalization uncertainty of each background is taken into account as a constrained nuisance parameter in the fit. Measurements performed in ATLAS during Run 1 are employed to derive the priors for the $t\bar{t}$ and single-top [155], V+light [156], V+b, V+c and SM $VH(\to b\bar{b})$ backgrounds [1]. The uncertainty for the diboson cross section are derived from MC [157]. The corresponding priors are listed in Table 8.5.

The prior uncertainty of 30 % for the $t\bar{t}$ background is somewhat conservative. However, the normalization of this background is well constrained in the fit to the data and even a floating parameter would be suited for its parameterization, as discussed in Section 8.8.2.

The shape uncertainty on the m_{VH} distribution is evaluated for the $t\bar{t}$ and V+jets backgrounds, as described in the following. The corresponding nuisance parameters for the fit are summarized in Table 8.6.

The distribution obtained from the default Powheg sample of $t\bar{t}$ is compared to the one from the aMC@NLO 2.2.2 generator. The full difference is parameterized with a logarithmic function and taken into account symmetrized as a systematic shape uncertainty (`MODEL_TTbar_aMcAtNlo`).

The same procedure is applied to an alternative sample of $t\bar{t}$ with the default generator, but showered with Herwig++ 2.7.1 [158] instead of Pythia (`MODEL_TTbar_Herwig`). Additionally, samples of $t\bar{t}$ with variations of the factorization and renormalization scales

Name	Source	Value	Count
MODEL_Vjets_MadGraph	V+jets MadGraph	S	1
MODEL_TTbar_aMcAtNlo	$t\bar{t}$ generator	S	1
MODEL_TTbar_Herwig	$t\bar{t}$ showering	S	1
MODEL_TTbar_rad	$t\bar{t}$ radiation	S	1
IFSR	Signal radiation	3 %	1
PDF	Signal PDF	1 %	1
Total	6 with priors, 0 floating		

Table 8.6.: Nuisance parameters and their priors for the systematic uncertainties on the background and the signal processes.

(factors of 2 and 0.5) are used to derive a shape uncertainty on additional radiation in the events (MODEL_TTbar_rad).

For the V+jets backgrounds, the m_{VH} distribution obtained from the default Sherpa generator is compared to the one from MadGraph5_aMC@NLO 2.2.2, showered with Pythia 8.186 using the AZNLO tune. The difference is parameterized and taken into account as a systematic shape uncertainty in the fit (MODEL_Vjets_MadGraph).

The initial- and final-state radiation (IFSR) and PDF uncertainties (PDF) for the signal are taken into account as well. They vary the normalization of the signal by at most 3 %, as listed in Table 8.6, such that their impact on the search results is insignificant.

8.7. Binning strategies

The fit model, as designed initially, was very complex. The 1-lepton channel alone consisted of 18 regions: 0-, 1- and 2-tag split up in 0 and 1 additional tag and in low, medium and high m_{jet}. Each region consisted of 20 bins with a width of 250 GeV from 0 to 5 TeV, resulting in 360 bins in the fit in total.

This fit model is simplified in the next section by reducing the number of regions and by using transformation D, which was developed for the SM $VH(\to b\bar{b})$ analysis. The outcome of this study is not used directly for the final results, but it is a prime example for a fit-model optimization. The final binning strategy for the analysis, which was motivated by the optimization, is described in Section 8.7.2.

8.7.1. Regions and transformation D

This initial fit model is simplified in the following using transformation D, as defined in Section 5.6.4. The expected limit for the signal with $m_{VH} = 1$ TeV serves as benchmark for the sensitivity. The optimization has been performed in the course of this thesis for the 1-lepton channel.

The first step of the simplification consists of dropping control regions (CRs). Only those regions are kept that are expected to contribute to the sensitivity from a physical perspective:

- 0-, 1- and 2-tag, 0 add. tag, low m_{jet}: three CRs for W+l, W+c and W+b

- 1-tag, 1 add. tag, high m_{jet}: one CR for $t\bar{t}$
- 1- and 2-tag, 0 add. tag, medium m_{jet}: two signal regions

From this configuration the $W+l$ (0-tag) control region is dropped as well, since a negligible change in sensitivity was observed. Altogether, this drastically reduces the number of regions from 18 to five, while the limit degrades mildly by about 1 %.

Of the remaining regions, the $W+c$ (1-tag) and $W+b$ (2-tag) control regions contribute slightly to the sensitivity ($\lesssim 1\,\%$). However, if the priors on the W+jets background components are dropped, their contribution rises to about 5 %. The $t\bar{t}$ control region provides about 1 % of the sensitivity. These observations depend on the signal mass, since the background composition changes with m_{VH}.

The number of bins is reduced further significantly by using only the normalization (one bin) in each of the three remaining control regions. Another slight degradation of the limit by about 0.5 % is observed.

The usefulness of this last change is debatable. By using only one bin, statistical fluctuations in the experimental uncertainties are reduced and artificial constrains of the corresponding nuisance parameters can be alleviated, as discussed in Section 5.7.2. However, the shape of the distributions in the control regions might provide additional information about some modeling uncertainties. This information, possibly resulting in a pull of some nuisance parameter(s), might be useful or harmful (e.g. due to a missing extrapolation uncertainty to the signal region). This can only be judged from further studies. In most cases using less shape information, resulting in weaker constraints of nuisance parameters, seems to be the more conservative approach.

Finally, the fixed-width binning in the signal regions is replaced with transformation D. This allows for more bins in the sensitive region of m_{VH} (around the signal mass), while broader bins in the tails are created. The parameters of the algorithm are chosen as $z_s = 1$ and $z_b = 5$ without further optimization. This results in six bins in each of the two signal regions. Due to the choice of $z_s > 0$ the bin boundaries depend on the signal sample and vary with the mass under consideration. An example of the effect of transformation D on the m_{VH} distribution is shown in Figure 8.7.

Changing the binning in the signal region the expected limit improves by about 10 %, which compensates well the percent-level degradation from the simplifications above. The total number of bins in the fit is reduced from 360 to 15 with respect to the initial fit model. By reducing the number of control regions and bins severe over-constraints of some nuisance parameters are resolved.

This very encouraging result is not directly used in the fit, but is partially translated to the final binning choice, as described in the next section.

8.7.2. Final binning

The binning scheme that is employed to extract the results is a compromise between a fixed-width binning with all regions and the aggressive reduction described in the previous section.

The control regions are chosen based on the number of expected background events. If this yield is very low the region is dropped or merged with a region of different m_{jet},

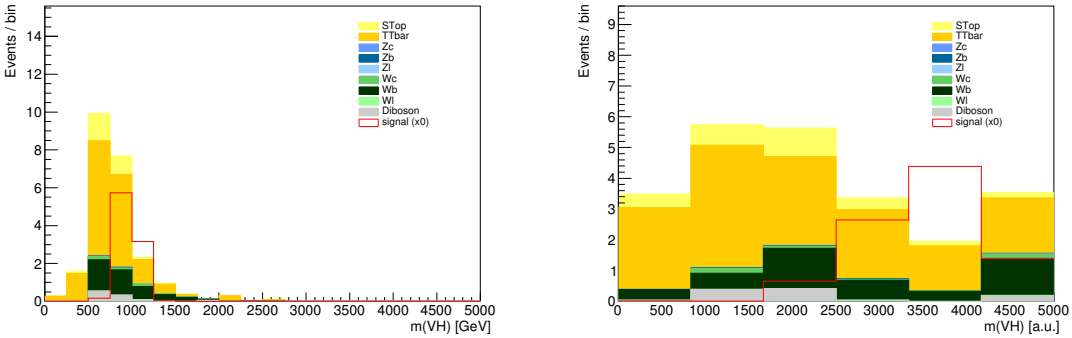


Figure 8.7.: Distributions of m_{VH} for the 1-lepton channel with the 2-tag signal region selection before the fit to the data. The fixed-width (250 GeV) binning is shown on the left and the same distribution after transformation D on the right. The normalization of the total background and of the signal is arbitrary, but consistent for both distributions.

but only within the same lepton and b-tagging selection. The resulting configuration is summarized in Table 8.4.

Broad bins with a width of 500 GeV are used for the control regions of the 0- and 2-lepton channels and with 1 TeV in the 1-lepton channel. This prevents large statistical fluctuations and keeps some shape information. No severe over-constraints due this choice are observed in the final fit.

The signal region binning is based on the experimental resolution, σ, of the m_{VH} distributions. The resolution is parameterized by a linear function as

$$\sigma(m_{VH}) = 15\,\text{GeV} + 5\,\% \times m_{VH}. \tag{8.7}$$

For the 0-lepton channel m_{VH} is replaced by $m_{\text{T},VH}$. The bin width is chosen as $2\,\sigma$ ($3\,\sigma$) for the 0- and 2-lepton (1-lepton) channels. For both the signal and control regions, bins that have zero background expectation are merged with the neighboring bin on the right.

This binning scheme produces more bins for the fit than transformation D. However, it has the advantage of being independent of the signal distribution, resulting in same binning for the various signal masses. Hence, it was preferred in the analysis group and was chosen as final binning.

8.8. Fit model validation

The fit model is validated in the following. Similar techniques as for the SM $VH(\to b\bar{b})$ analysis are applied. Motivational comments are given in Section 5.7. The validation has been performed in the course of this thesis.

8.8.1. Post-fit plots

All post-fit $m_{(\mathrm{T},)VH}$ distributions of the regions used in the fit are shown in Figure 8.8 for the signal regions and in Figures 8.9, 8.10 and 8.11 for the 0-, 1- and 2-lepton control regions, respectively.

The various background components are shown stacked as colored filled histograms. The expected HVT signal with $m_{V'} = 2\,\mathrm{TeV}$ is shown as dashed red histogram and is normalized to 50 times the corresponding expected 95 % CL upper cross section limit. The data are shown as points with error bars. The combined statistical and systematic uncertainty on the prediction is indicated by the hatched band. The pre-fit background prediction is shown in the plots as dotted blue line.

In general, good agreement of the background prediction with the data is observed after the fit. The six signal regions plots are shown with logarithmic scale and with the ratio of the data to the post-fit prediction in Section 8.9.

The chosen binning strategy causes bins of varying width as intended. However, in the control regions many bins with very low event yields are present. As discussed in Section 8.7.1, they do probably not contribute to the sensitivity and could be further reduced for a future analysis.

8.8.2. Nuisance parameter pulls

The full set of nuisance-parameter pulls from the fit to the data is shown in Figure 8.12 for the V' combination using the signal with $m_{V'} = 2\,\mathrm{TeV}$ and all three lepton channels. The pulls are compared to the ones from a fit to the Asimov dataset.

The parameters are well behaved and most pulls are withing the $1\,\sigma$ level. In particular, the background normalization parameters are all compatible with the prior expectation. This is important, since the choice of priors instead of floating parameters might be questioned otherwise.

Most normalization parameters are only mildly constrained with the exception of `XS_ttbar`, since clean control regions for the $t\bar{t}$ background are used in the fit. Its pull is -0.48 ± 0.23, which corresponds, with the prior of 30 %, to a scale factor for $t\bar{t}$ of 0.86 ± 0.07, which is still reasonable. Opposed to the Run 1 analyses, no significant deviation between the three lepton channels is observed for this scale factor. Due to the strong constraint from the data, the prior provides very little additional information and could be dropped.

Mild pulls and constraints are observed in the fit to the data for the $t\bar{t}$ modeling parameter `MODEL_TTbar_Herwig` of 0.38 ± 0.78 and for `MODEL_TTbar_aMcAtNlo` of -0.38 ± 0.52. These parameters affect the shape of the $m_{\mathrm{T},VH}$ distributions and a constraint from the clean $t\bar{t}$ control regions is not unexpected. This confirmed by the constraints in the fit to the Asimov dataset. These are not expected to affect the search result, since no significant correlation with the signal strength parameter is observed, as shown in the next section.

The strongest pulls are observed in `FATJET_JER` with $1.17\pm0.66\,\sigma$ and on `FT_EFF_Eigen_B0` with $-0.85 \pm 0.51\,\sigma$. The constraints are consistent with the expectation from the fit to the Asimov dataset.

Some of the nuisance parameters for experimental uncertainties show over-constraints. For example, the post-fit uncertainties on `JET_Rtrk_Baseline` and `JET_Rtrk_Modelling`

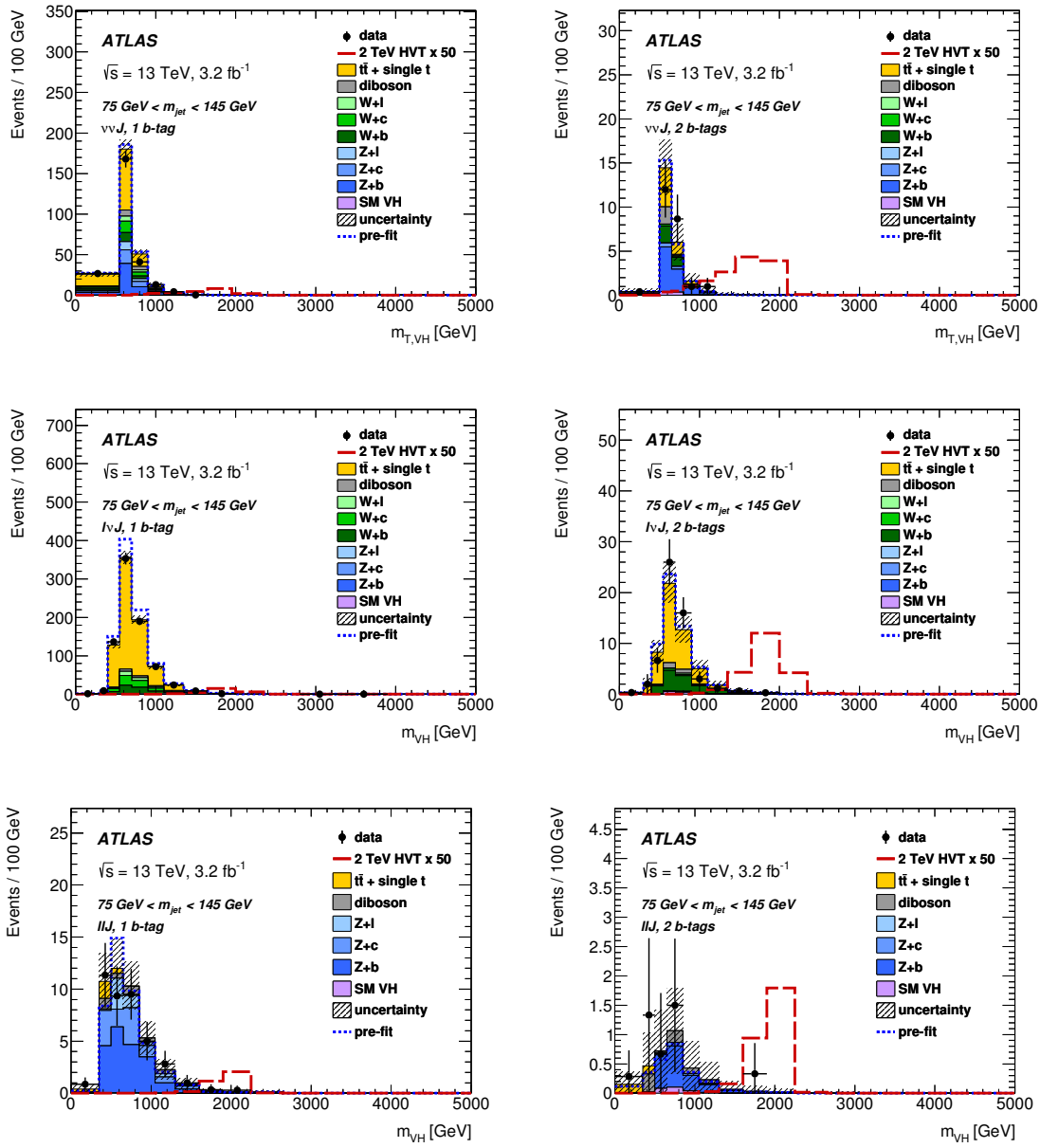


Figure 8.8.: Distributions of $m_{(\mathrm{T},)VH}$ in linear scale for the (from top to bottom) 0-, 1- and 2-lepton signal regions with 1-tag (left) and 2-tag selections (right) after the fit to the data. The expected HVT signal with $m_{V'} = 2\,\mathrm{TeV}$ is shown as dashed red histogram and is normalized to 50 times the corresponding expected 95 % CL upper cross section limit. A formal description of the plot elements is given in the text (from Ref. [3]).

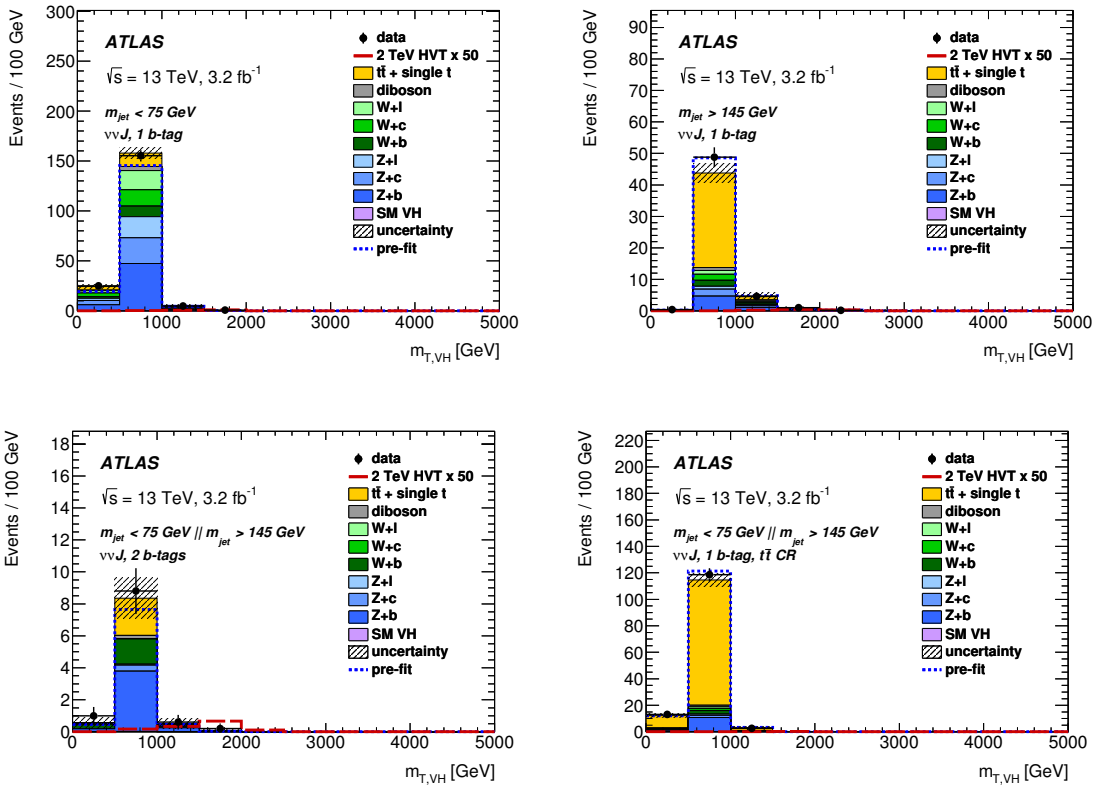


Figure 8.9.: Distributions of $m_{\mathrm{T},VH}$ in the 0-lepton control regions after the fit to the data: 1-tag, low m_{jet} (upper left), 1-tag, high m_{jet} (upper right), 2-tag, low+high m_{jet} (lower left) and the 1-tag $t\bar{t}$ CR (lower right). The expected HVT signal with $m_{V'} = 2\,\mathrm{TeV}$ is shown as dashed red histogram and is normalized to 50 times the corresponding expected 95 % CL upper cross section limit. A formal description of the plot elements is given in the text (from Ref. [3]).

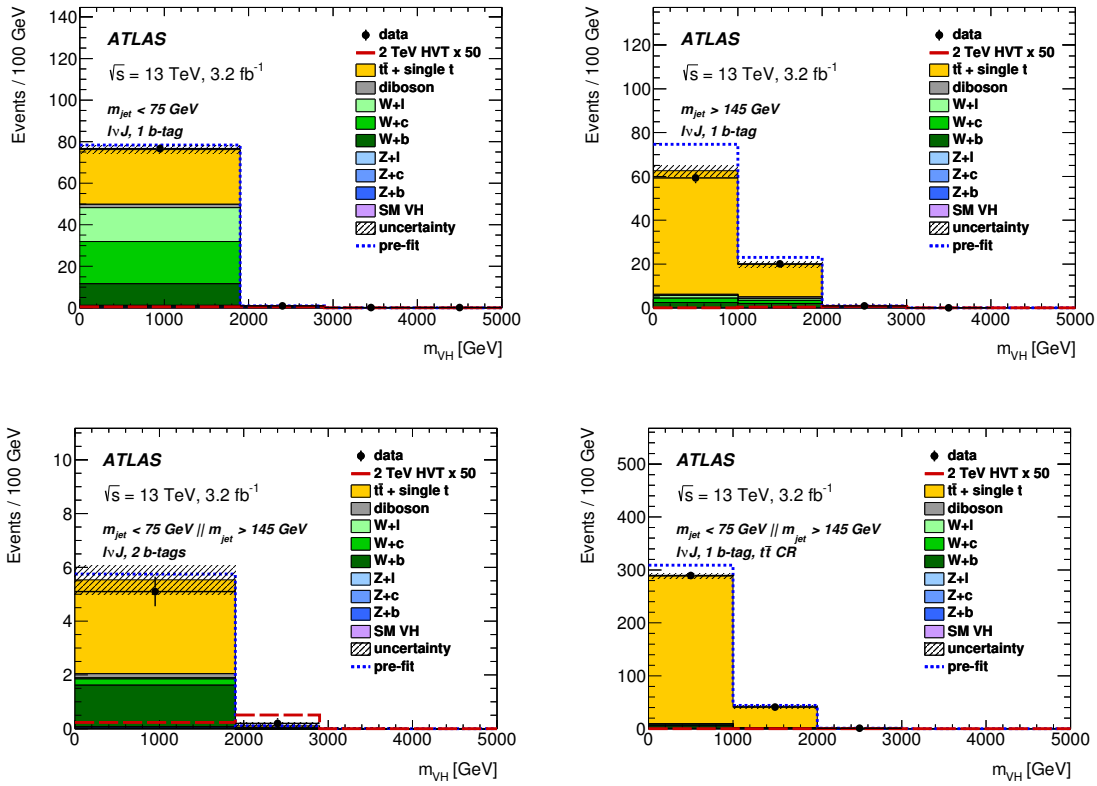


Figure 8.10.: Distributions of m_{VH} in the 1-lepton control regions after the fit to the data: 1-tag, low m_{jet} (upper left), 1-tag, high m_{jet} (upper right), 2-tag, low+high m_{jet} (lower left) and the 1-tag $t\bar{t}$ CR (lower right). The expected HVT signal with $m_{V'} = 2\,\text{TeV}$ is shown as dashed red histogram and is normalized to 50 times the corresponding expected 95 % CL upper cross section limit. A formal description of the plot elements is given in the text (from Ref. [3]).

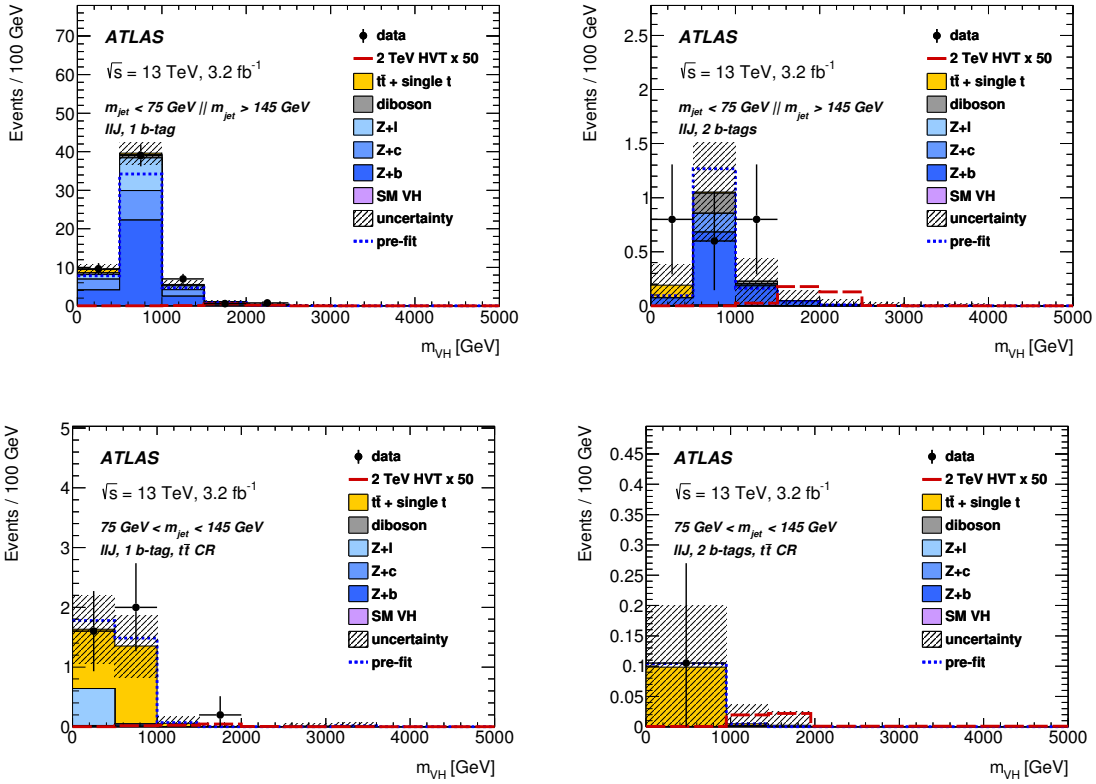


Figure 8.11.: Distributions of m_{VH} in the 2-lepton control regions after the fit to the data. 1-tag, low+high m_{jet} (upper left), 2-tag, low+high m_{jet} (upper right), the 1-tag $t\bar{t}$ CR (lower left) and the 2-tag $t\bar{t}$ CR (lower right). The expected HVT signal with $m_{V'} = 2\,\text{TeV}$ is shown as dashed red histogram and is normalized to 50 times the corresponding expected 95 % CL upper cross section limit. A formal description of the plot elements is given in the text (from Ref. [3]).

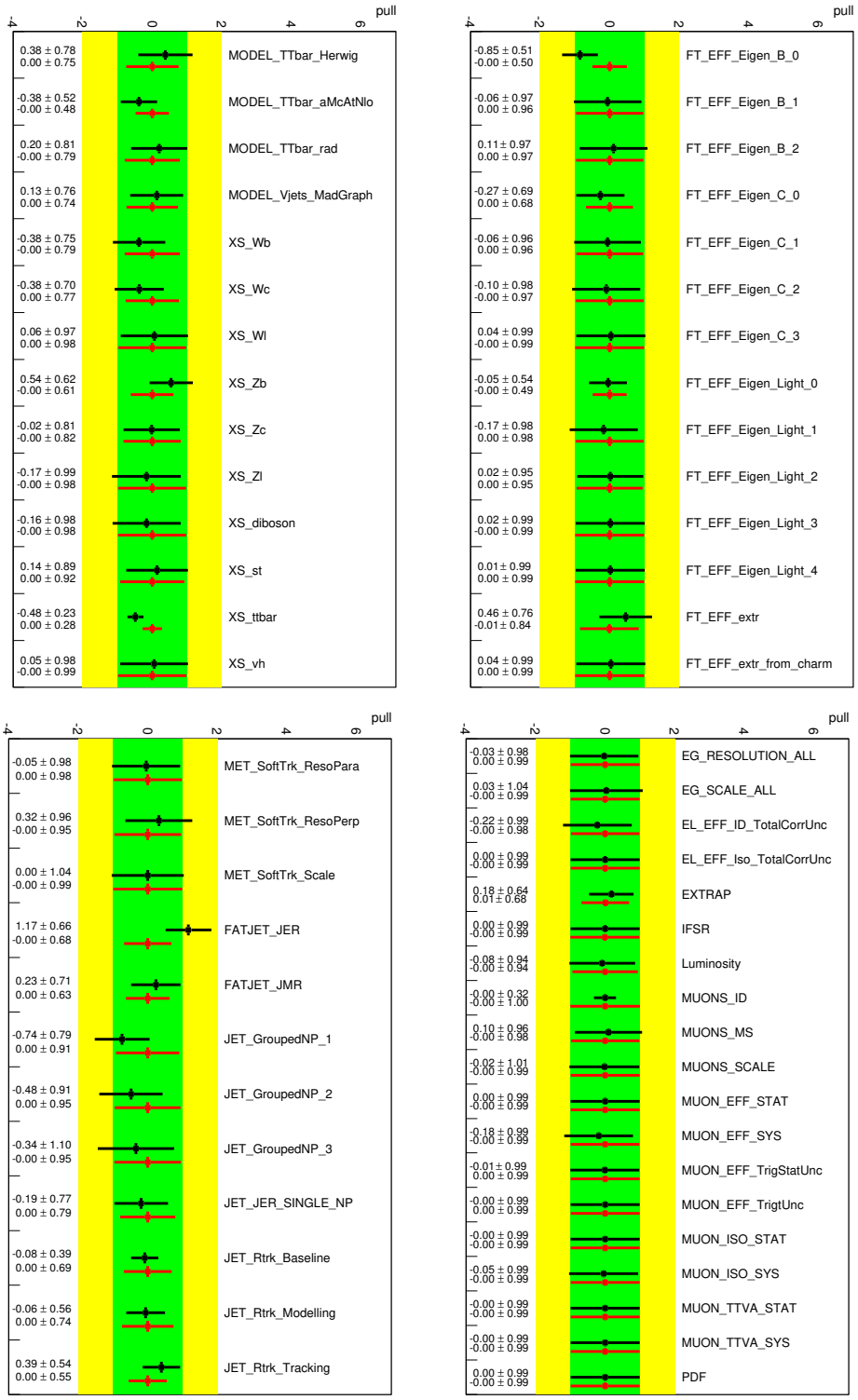


Figure 8.12.: Nuisance parameter pulls from the V' combination fit with $m_{V'} = 2\,\text{TeV}$ to the data (black) and to the Asimov data (red). Shown are the background normalization and modeling NPs (upper left), b-tagging NPs (upper right), jets and $E_{\text{T}}^{\text{miss}}$ NPs (lower left) and the remaining NPs (lower right).

Parameter 1	Parameter 2	Correlation
mu	FATJET_JER	−20 %
	FATJET_JMR	+15 %
JET_Rtrk_Baseline	JET_Rtrk_Modelling	−45 %
XS_Wc	FT_EFF_Eigen_Light_0	+50 %
XS_ttbar	FT_EFF_Eigen_B_0	+55 %
	Luminosity	−50 %

Table 8.7.: Largest correlations observed in the combined V' fit to the data for any parameter and for the signal strength parameter, mu.

observed in the fit to the data are only about 2/3 compared to the fit to the Asimov dataset. A strong over-constraint is observed on MUONS_ID, which is constraint to ±0.3 in the fit to the data, while no constraint is expected from the Asimov dataset. However, the parameter is not pulled and has no impact on the signal measurement, as discussed in the next section.

The over-constraints are likely caused by the large statistical uncertainty on the backgrounds in the tails of the m_{VH} distributions, as discussed in Section 5.7.2. For the control regions, this could be easily circumvented with an optimized binning. One could use only the normalization in these regions, which would not affect the sensitivity, as shown in Section 8.7.1.

8.8.3. Correlations

The correlation matrix, as observed in the fit to the data for the V' combination including all parameters of the fit, except for the bin-wise statistical uncertainties of the simulated backgrounds, is shown in Figure 8.13. The corresponding matrix from the fit to the Asimov dataset and the full matrices are shown in Appendix D.1. The largest correlations from the fit to the data are summarized in Table 8.7.

The strongest correlation of +55 % shows up between the $t\bar{t}$ normalization parameter, XS_ttbar, and the leading b-tagging uncertainty, FT_EFF_Eigen_B_0. This is well understood, since the construction of the $t\bar{t}$ control regions relies on additional b tags in the event. A similar behavior is observed for the $W+c$ normalization parameter, XS_Wc, which has a +50 % correlation with FT_EFF_Eigen_Light_0.

Further, a negative correlation of −50 % of XS_ttbar with Luminosity is observed. This is attributed to the fact that the largest contribution of the selected events is expected from the $t\bar{t}$ background.

The next largest correlation with −45 % is observed between the large-jet uncertainty-parameters JET_Rtrk_Baseline and JET_Rtrk_Modelling. A possible explanation is that the present analysis cannot resolve the differences between the these parameters.

Further, the energy scale parameters show correlations of about −20 % to −30 % among them and with the background normalization parameters. The cut on the large-jet momentum of $p_{\mathrm{T}} > 250\,\mathrm{GeV}$, together with variations in the jet energy, cause changes in the acceptance of the backgrounds, which can lead to the observed correlations.

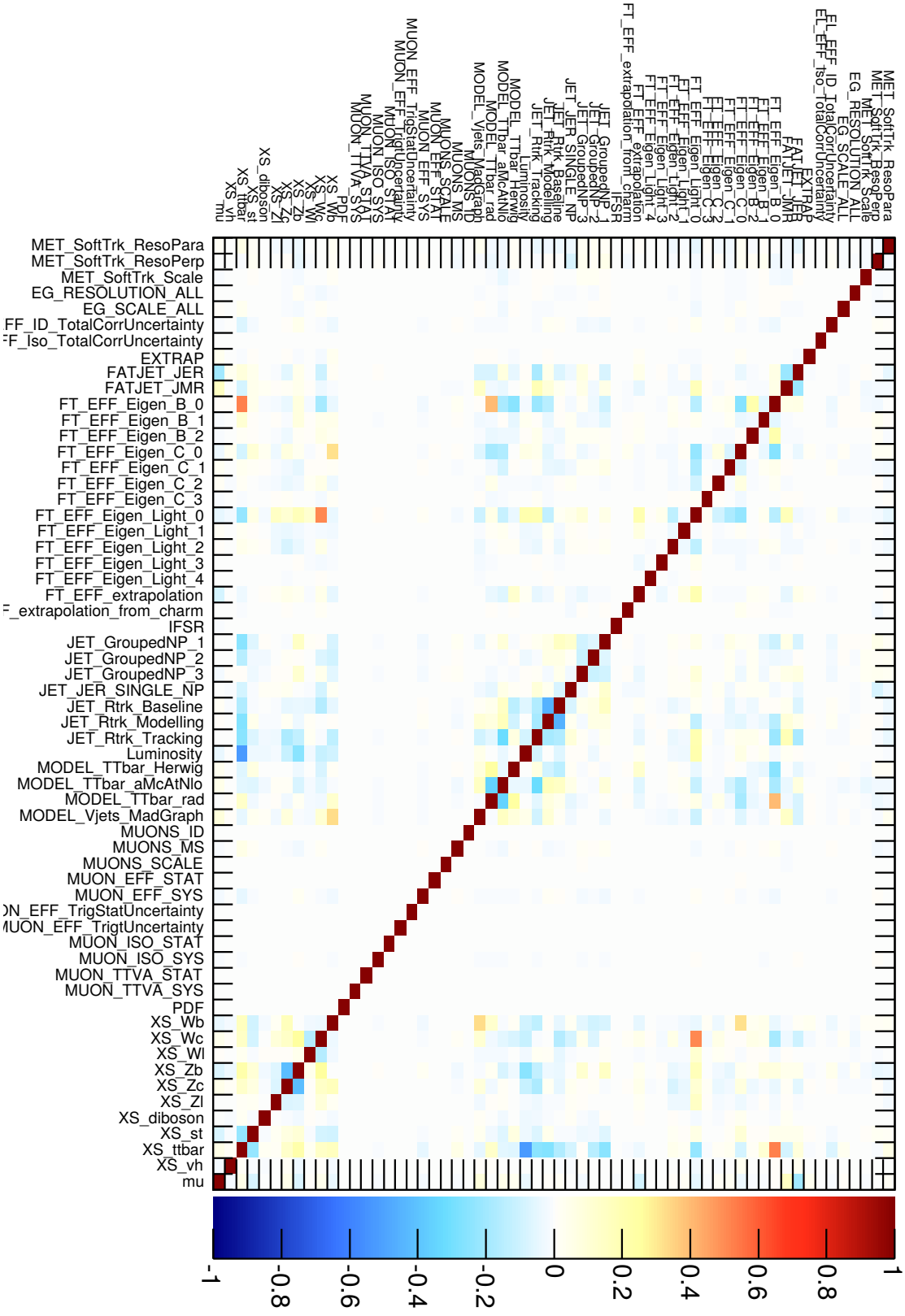


Figure 8.13.: Correlation matrix from the combined V' fit to the data. All parameters of the fit are shown, except for the bin-wise statistical uncertainties of the simulated backgrounds.

For the signal-strength parameter, `mu`, the largest correlation of $-20\,\%$ is observed with the large-jet energy-resolution parameter, `FATJET_JER`. This is not surprising, as this parameter affects the resolution of the reconstructed invariant mass, m_{VH}, which is the final discriminant. This moderate correlation is somewhat critical, as the `FATJET_JER` is also the parameter with the strongest pull in the fit, as discussed in the previous section. However, no other sign of mis-modeling is observed, such that this effect is deemed to be of physical origin and no artificial bias on the signal measurement is expected.

The large-jet mass resolution, `FATJET_JMR`, has a similar correlation of $+15\,\%$ with the signal strength. This correlation is caused by the selected jet-mass window around the Higgs boson mass of $75\,\text{GeV} < m_{\text{jet}} < 145\,\text{GeV}$. Together with a variation of the mass resolution this selection cut causes acceptance changes in the signal regions.

Other parameters show correlation of at most $\pm 10\,\%$ with the signal strength. The large-jet energy scale parameters and `MUONS_ID` are with less than $\pm 5\,\%$ very weakly correlated with the signal. As discussed in the previous section, these parameters show some over-constraints. If the correlation with the signal would be larger, it might be questioned if the uncertainty on the signal strength is well measured. However, this is not the case.

8.9. Results

The most sensitive distributions after the fit to the data are shown in Figure 8.14 in logarithmic scale. The expected HVT signal with $m_{V'} = 2\,\text{TeV}$ is shown as dashed red histogram and is normalized to 50 times the corresponding expected 95 % CL upper cross section limit. The corresponding post-fit event yields are listed in Table 8.8. No significant excess of events over the expected SM backgrounds is observed in the data.

The data are used to set 95 % CL upper limits separately on the W' or Z' production cross section times the branching fraction into WH or ZH times $\text{BR}(H \to b\bar{b}/c\bar{c})$ as a function of the resonance mass, as shown in Figure 8.15. For the limit on W' production zero abundance of the Z' signal is assumed and vice versa.

They limits range from about 100 fb for masses around $m_{V'} = 1\,\text{TeV}$ to about 10 to 30 fb for the most sensitive region between $m_{V'} = 2$ to 5 TeV for both hypotheses. The HVT Model A with $g_V = 1$ is excluded for $m_{W'} < 1.75\,\text{TeV}$ and $m_{Z'} < 1.49\,\text{TeV}$ and Model B with $g_V = 3$ for $m_{W'} < 2.22\,\text{TeV}$ and $m_{Z'} < 1.58\,\text{TeV}$.

A combined limit on the V' production cross section times branching ratio is derived assuming the W' and Z' masses to be degenerate and taking the WH/ZH branching ratios from the HVT model. The result is shown in Figure 8.16 relative to the HVT Model A prediction. The corresponding values from the HVT Model B, predicting the same relative amount of W' and Z' production, are shown as well. The HVT Model A with $g_V = 1$ is excluded for $m_{V'} < 1.9\,\text{TeV}$ and Model B with $g_V = 3$ for $m_{V'} < 2.1\,\text{TeV}$. Corresponding exclusion contours in the HVT parameter space of $\{g_V c_H, (g_2/g_V)c_F\}$ for resonances of masses of 1.2 TeV, 2.0 TeV and 3.0 TeV are shown in Figure 8.17.

These results exceed the sensitivity of earlier $V' \to VH$ searches performed by the ATLAS and the CMS collaborations using data taken at $\sqrt{s} = 8\,\text{TeV}$. The earlier search from the ATLAS collaboration is performed in the same final as the present one and excludes the HVT Model A with $g_V = 1$ for $m_{W'} < 1.47\,\text{TeV}$ and $m_{Z'} < 1.36\,\text{TeV}$ [154].

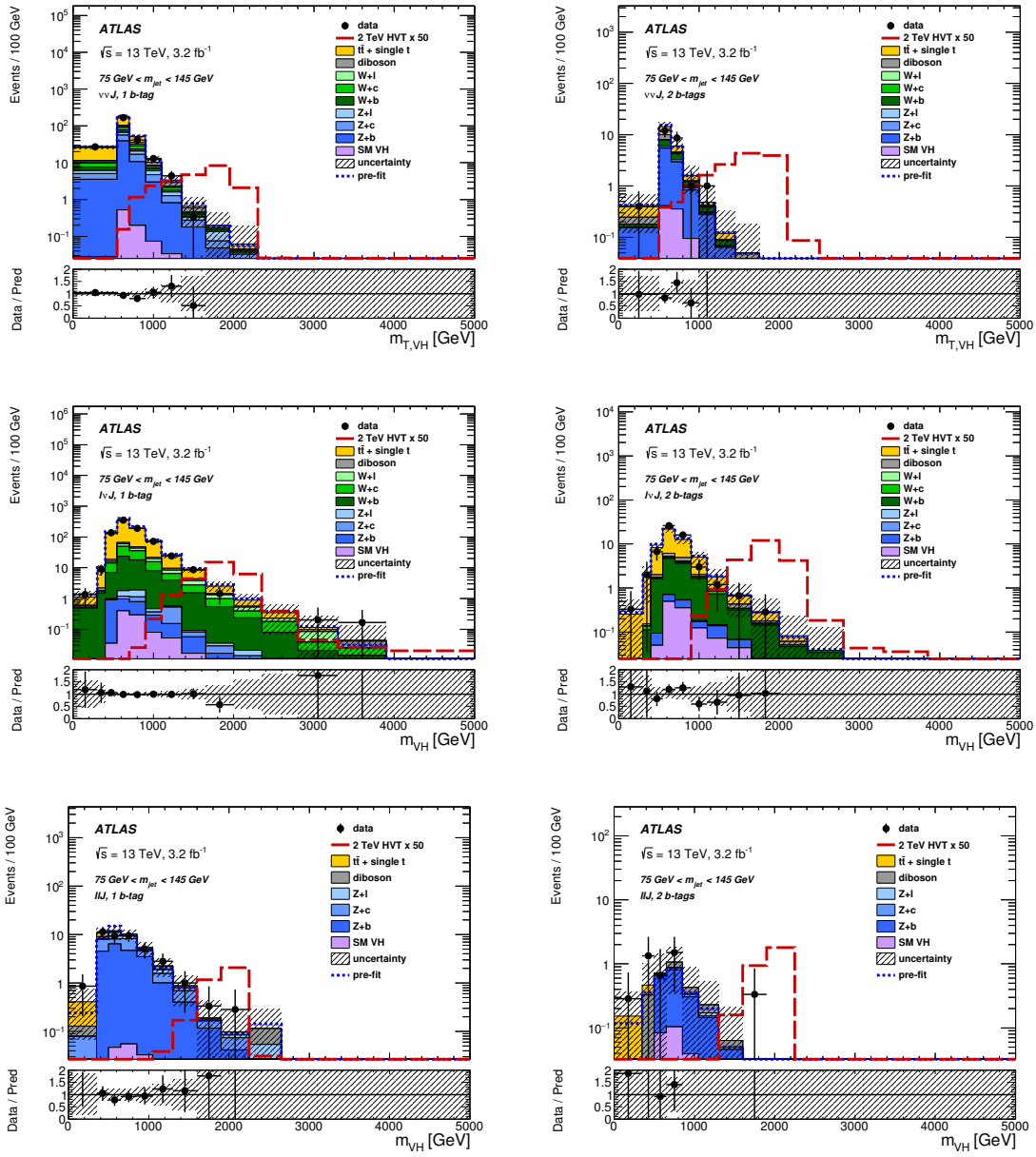


Figure 8.14.: Distributions of m_{VH} in logarithmic scale for the (from top to bottom) 0-, 1- and 2-lepton signal regions with 1-tag (left) and 2-tag selection (right) after the fit to the data. The expected HVT signal with $m_{V'} = 2\,\text{TeV}$ is shown as dashed red histogram and is normalized to 50 times the corresponding expected 95 % CL upper cross section limit. A formal description of the plot elements is given in Section 8.8.1 (from Ref. [3]).

	Two b-tags		
	$\nu\nu b\bar{b}$	$\ell\nu b\bar{b}$	$\ell\ell b\bar{b}$
$t\bar{t}$	9.6 ± 1.4	50 ± 7	0.54 ± 0.36
Single top	2.0 ± 0.6	11.4 ± 3.0	0.20 ± 0.10
$W + b$	5.2 ± 1.3	18 ± 5	
$W + c$	0.64 ± 0.18	2.0 ± 0.7	
$W + q$	0.06 ± 0.03	2.0 ± 0.8	
Diboson	4.2 ± 1.8	4.6 ± 0.8	1.28 ± 0.27
SM VH	1.43 ± 0.57	0.03 ± 0.01	0.45 ± 0.19
$Z + b$	12.3 ± 2.4	1.0 ± 0.4	3.4 ± 0.8
$Z + c$	1.46 ± 0.43	0.05 ± 0.02	0.31 ± 0.10
$Z + q$	0.13 ± 0.05		0.04 ± 0.04
Backgrounds	36.9 ± 3.5	90 ± 6	6.2 ± 1.0
Data	37	96	8

	One b-tag		
	$\nu\nu b\bar{b}$	$\ell\nu b\bar{b}$	$\ell\ell b\bar{b}$
$t\bar{t}$	216 ± 17	969 ± 50	3.8 ± 0.8
Single top	26 ± 7	112 ± 30	0.58 ± 0.19
$W + b$	33 ± 8	100 ± 24	
$W + c$	41 ± 10	109 ± 31	
$W + q$	20 ± 5	1.5 ± 0.6	
Diboson	28 ± 5	32 ± 5	6.4 ± 1.0
SM VH	1.6 ± 0.6	0.04 ± 0.01	0.30 ± 0.12
$Z + b$	99 ± 17	3.8 ± 1.0	36 ± 6
$Z + c$	51 ± 13	2.7 ± 1.6	19 ± 5
$Z + q$	32 ± 8	3.0 ± 1.0	9 ± 4
Backgrounds	548 ± 16	1385 ± 30	75 ± 7
Data	520	1364	75

Table 8.8.: Predicted and observed numbers of events for the six signal regions. The yields and the combined statistical and systematic uncertainties of the backgrounds are shown after the combined fit to the data (from Ref. [3]).

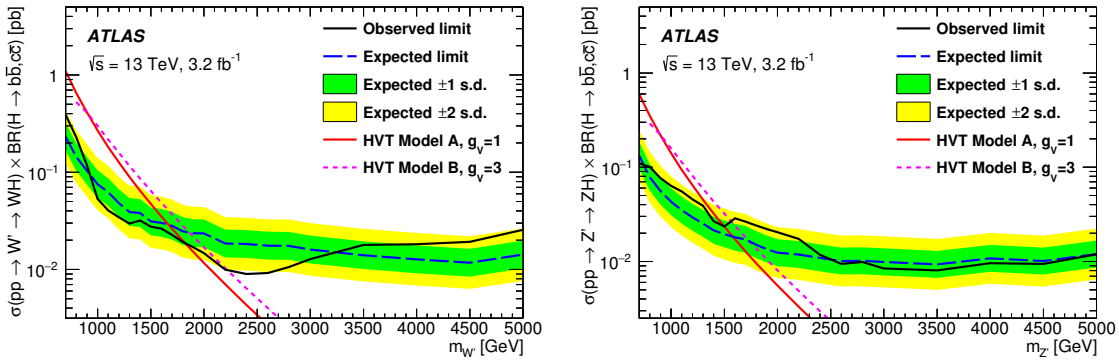


Figure 8.15.: 95 % CL upper limits on the W' (left) and Z' (right) production cross section times branching ratio from the combined fits as a function of the hypothesized resonance mass. The observed limit, as measured in the data, is compared to the expected one, when assuming the absence of the signal. The predicted values from the HVT Models A and B are shown as well (from Ref. [3]).

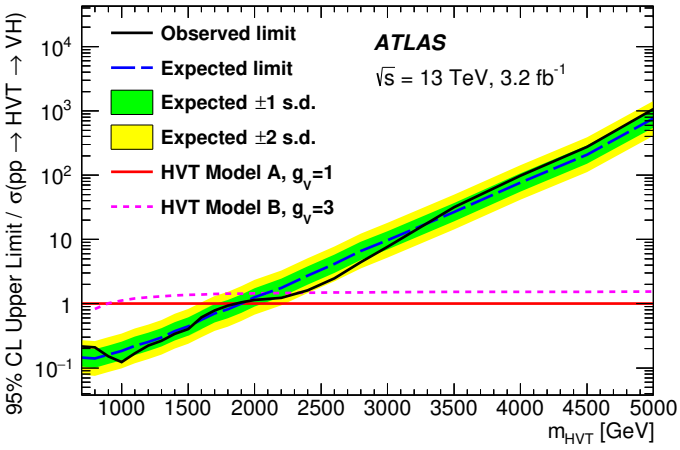


Figure 8.16.: 95 % CL upper limits on the V' production cross section times branching ratio from the combined V' fit relative to the HVT Model A prediction as a function of the hypothesized resonance mass. The observed limit, as measured in the data, is compared to the expected one, when assuming the absence of the signal. The predicted values from the HVT Model B are shown as well (from Ref. [3]).

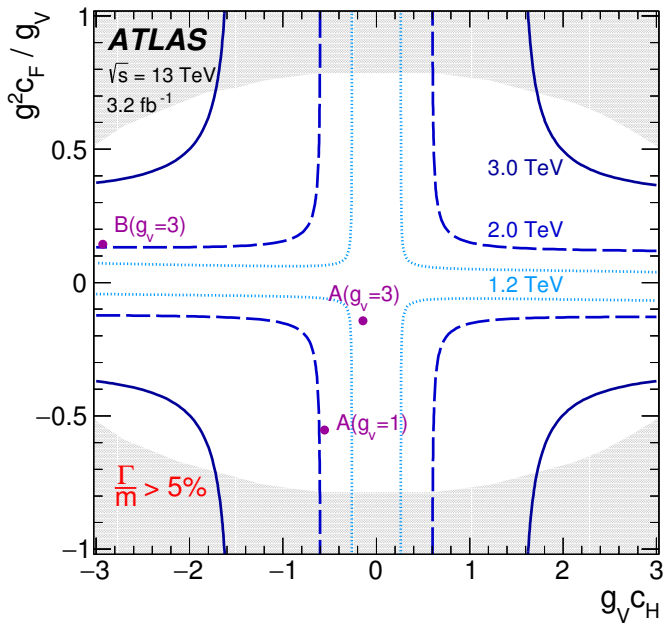


Figure 8.17.: 95 % CL exclusion contours for the HVT signal with $m_{V'} = 1.2$, 2.0 and 3.0 TeV from the combined V' fit in the plane of $g_V c_H$ and $g^2 c_F / g_V$. The areas outside the contours are excluded. The values for the HVT Model A with $g_V = 1$ and $g_V = 3$ and HVT Model B with $g_V = 3$ are indicated by dots. For the shaded regions the total width of the resonance, Γ, is larger than 5 % of its mass. In this case, it is not negligible compared to the experimental resolution and the results are not reliable (from Ref. [3]).

The earlier search from the CMS collaboration is performed in the fully hadronic decay mode and excludes the HVT Model B with $g_V = 3$ for $m_{W'} < 1.6\,\text{TeV}$, $m_{Z'} < 1.1\,\text{TeV}$ and in the range of $1.3\,\text{TeV} < m_{Z'} < 1.5\,\text{TeV}$, and in the combination for $m_{V'} < 1.7\,\text{TeV}$ [159].

In conclusion, no significant deviation from the Standard Model prediction has been found. Upper limits on the production cross section of W' and Z' resonance times the branching ratios for the decays $W' \to WH \to \ell\nu b\bar{b}$, $Z' \to ZH \to \nu\nu b\bar{b}$ and $Z' \to ZH \to \ell\ell b\bar{b}$ have been set. The results exceed the sensitivity of earlier searches and are to be improved further with more data from Run 2 of the LHC.

9. Summary

Three searches and one sensitivity estimate for Higgs boson decays with b jets in the final state have been presented in this thesis. The decay of the Higgs boson to b quarks, $H \to b\bar{b}$, is of particular interest, since it has the largest branching ratio in the Standard Model (SM) for a mass of $m_H = 125\,\text{GeV}$.

First, the search for the SM Higgs boson in associated production with a vector boson and the decay to b quarks, $VH(\to b\bar{b})$, is discussed. It is based on the data taken in 2012 at a center-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$. The analysis yields an important component for the measurements of the Higgs boson couplings and has contributed to a combined measurement by the ATLAS and CMS collaborations.

A sophisticated multivariate analysis (MVA) has been employed, using boosted decision trees (BDTs). The configuration of the BDTs was optimized and the performance was evaluated to be saturated with respect to the number of events in the training. The MVA has been optimally utilized together with a pseudo-continuous b-tagging calibration. The calibration not only allows to apply a cut on the b-tagging output as usual, but to use its distribution as input to the MVA. As a consequence, the separation of the signal from the background processes was improved.

A novel binning strategy has been developed by using numerical optimization algorithms, which were translated to analytical ones. These algorithms allowed to alleviate statistical fluctuations in the combined likelihood fit, while preserving a good sensitivity. The results of the fit were thoroughly investigated using various techniques, such as nuisance parameter de-correlations, uncertainty breakdowns, likelihood scans, toy experiments or the compatibility of the signal strength between regions.

As a final result, the signal strength of the Standard Model (SM) Higgs boson with $m_H = 125\,\text{GeV}$ relative to the SM expectation, $\mu = \sigma/\sigma_{\text{SM}}$, has been measured to be

$$\hat{\mu}_{8\,\text{TeV}} = 0.65^{+0.43}_{-0.40}$$

with the data taken in 2012 at $\sqrt{s} = 8\,\text{TeV}$. This corresponds to an observed (expected) significance of $1.7\,\sigma$ ($2.5\,\sigma$). In combination with the corresponding analysis carried out with the data taken in 2011 at $\sqrt{s} = 7\,\text{TeV}$ the signal strength is decreased to

$$\hat{\mu}_{7+8\,\text{TeV}} = 0.51^{+0.40}_{-0.37},$$

corresponding to an observed (expected) significance of $1.4\,\sigma$ ($2.6\,\sigma$).

The deviation from the SM expectation is compatible with a statistical downward fluctuation. However, there is also a chance that it is a sign of new physics, since a lower signal strength is expected in some BSM scenarios. The uncertainties are too large to draw any conclusions yet. This has to be clarified with further measurements using the data from Run 2 of the LHC, which has started in 2015.

As the uncertainty for the present measurement is composed about equally from statistical and systematic components, the possible gain in sensitivity from just increasing the amount of data would be limited. Further work has to be done on the understanding of the background processes to reduce the associated uncertainties. In particular the irreducible $W+bb$ background, which has the largest impact on the present measurement, will be of interest. Further, the $t\bar{t}$ background is likely to become more important, since its production cross section rises the most with the increased center-of-mass energy for Run 2.

In addition, a sensitivity estimate for the SM Higgs boson in gluon-fusion production and the decay to two b quarks, $gg \to H \to b\bar{b}$, is presented for the first time. This feasibility study is based on the data taken in 2012 at $\sqrt{s} = 8\,\text{TeV}$. The expected statistical significance has been estimated to be $0.7\,\sigma$ from a likelihood ratio binned in the dijet mass. This includes possible improvements estimated using multivariate methods.

Further steps have to be taken for a full analysis, in particular concerning precise estimates of the backgrounds. The jet multiplicity, angular jet quantities and b-tagging working points are suggested for the construction of control regions. Boosted jet reconstruction techniques are expected to improve the sensitivity significantly.

The second analysis presented in this thesis is the search for decays of the Higgs boson to Beyond Standard Model particles, χ, that are invisible to the ATLAS detector. The search has been performed in the associated production with a hadronically decaying vector boson, $VH \to q\bar{q}'\chi\chi$. Also this analysis is based on the data taken in 2012 at $\sqrt{s} = 8\,\text{TeV}$.

The analysis shares many aspects with the SM $VH(\to b\bar{b})$ analysis. Similar backgrounds are present, such that it was possible to reuse the definition of control regions and also the well-established statistical treatment could be extended and adapted. The binning algorithms, developed for the SM $VH(\to b\bar{b})$ analysis, were evaluated here as well and allowed to reduce the number of bins in the fit significantly, while increasing the sensitivity.

No excess of events over the expected SM background processes is observed. The results have been used to set 95 % CL upper limits on $\sigma(VH) \times BR(H \to \text{inv.})$. They range from 1.6 pb to 0.13 pb for Higgs boson masses from $m_H = 115\,\text{GeV}$ to $300\,\text{GeV}$. Assuming a production of the Higgs boson as predicted in the SM, including the $gg \to H$ contribution, an observed (expected) upper limit of 78 % (86 %) on $BR(H \to \text{inv.})$ has been derived for $m_H = 125\,\text{GeV}$.

The third analysis is the search for new and heavy vector bosons, V', decaying via $V' \to VH(\to b\bar{b})$. The analysis was carried out with the first data taken during Run 2 of the LHC in 2015 with $\sqrt{s} = 13\,\text{TeV}$ and an integrated luminosity of $\mathsf{L} = 3.2\,\text{fb}^{-1}$.

A novel analysis software, the CxAOD framework, has been developed for the analysis. It is based on the xAOD data format, introduced within the ATLAS software for Run 2, and allows for an efficient storage of systematic variations. The size of simulated samples has been typically reduced by a factor of eight compared to the traditional N-tuple data format, which is a substantial improvement for the analysis setup for Run 2. This framework has been adapted by other analysis groups and is now widely used in the ATLAS collaboration.

Boosted jet-reconstruction techniques have been employed in the search for V' reso-

nances to preserve the signal acceptance for large $m_{V'}$ above 1 TeV. Jets with large radii in the calorimeter were used in combination with small-radius track jets to reconstruct the Higgs boson candidates. A good acceptance of the signal process is observed up to $m_{V'} = 5\,\text{TeV}$.

The reconstruction of the invariant mass in the 1-lepton channel, assuming W' production, required the estimation of the longitudinal momentum of the neutrino. This was accomplished using the known W-boson mass as constraint, which leads to two solutions. The resolution in $m_{W'}$ was optimized with respect to several combinations of these solutions.

The binning algorithms developed during Run 1 have shown very good performance for this analysis as well. The number of bins in the fit has been reduced by a factor of 24 compared to the initial binning choice, while the sensitivity has been improved by 10 %.

No significant excess of events over the expected SM background processes is observed. Limits on the production cross section times branching ratio have been set independently for the W' and Z' hypotheses. They range from about 100 fb for masses around $m_{V'} = 1\,\text{TeV}$ to about 10 to 30 fb for the most sensitive region between $m_{V'} = 2$ to 5 TeV for both hypotheses.

In conclusion, three searches and one sensitivity estimate for yet unobserved Higgs boson decays with b jets in the final state have been presented. Similar reconstruction techniques were employed and the statistical evaluation was optimized in many aspects. No significant deviation from the SM prediction was found. All these analyses have great potential to be consolidated during Run 2 of the LHC with more data and the increased center-of-mass energy.

Very recently a new search for the SM $VH(\to b\bar{b})$ process has been performed with the data taken at $\sqrt{s} = 13\,\text{TeV}$, corresponding to an integrated luminosity of $\mathsf{L} = 13.2\,\text{fb}^{-1}$ with the ATLAS experiment [160]. The analysis is designed close to the one presented in this thesis, but with an expected significance of $1.9\,\sigma$ it has not yet reached the same sensitivity. The observed signal strength is $\hat{\mu}_{13\,\text{TeV}} = 0.21^{+0.36}_{-0.35}(\text{stat.})^{+0.36}_{-0.36}(\text{syst.})$, again smaller than the SM expectation, resulting in an observed significance of $0.42\,\sigma$.

This result is to be improved with more data from the LHC, as are the searches for possible BSM signatures. Exciting times lie ahead for Higgs boson physics, in particular with b jets in the final state!

A. VH(bb) appendix

This appendix contains additional information for the search for the SM $VH(\to b\bar{b})$ process, discussed in Chapter 5.

A.1. Pull comparisons

Lepton channel comparison The full set of NP pulls from fits for the individual lepton channels and from the combined fit to the data is shown in Figures A.1, A.2 and A.3. In particular the `norm_ttbar` parameter is inconsistent between the channels, as discussed in Section 5.7.2.

Parameter de-correlations Examples for de-correlations of specific nuisance parameters in the combined fit to the data are shown in Figure A.4. These were done in an earlier stage of the analysis and do not reflect the final result. Shown are de-correlations of the `norm_ttbar`, `WPtV` and `JetEResol` parameters. The various de-correlations schemes consist of having independent parameters for an uncertainty for specific regions, instead of one for all regions. In the examples, fits are shown with independent paramaters for either the bins in p_{T}^{V}, the jet multiplicity bins, the lepton channels or the b-tagging regions.

The behaviour discussed in Section 5.7.2 is observed: `norm_ttbar` and `JetEResol` are inconsistent between lepton channels and `WPtV` is inconsistent between the 2- and 3-jet regions. The response of the other parameters of such de-correlations can be investigated as well. For example, the impact on the floating normalization parameters by de-correlating the `JetEResol` parameter is shown in Figure A.4 (lower right). No significant change is observed.

A.2. Unfolded b-tagging scale factors

The pulls of the b-tagging nuisance parameters, listed in Section 5.7.2, correspond the pulls on the eigenvectors (EV) from the decomposition of the uncertainties, as described in Section 4.4.

The pulls on the EV variations can be translated back to the original uncertainties as function of the jet p_{T} (and η for light jets). The resulting SF before and after the combined fit to the data in the SM $VH(\to b\bar{b})$ analysis are shown in Figure A.5 (b and c jets) and Figure A.6 (light jets).

The correlations between the original uncertainties, binned in jet flavor and p_{T} (and η), are quite strong. An example is shown for b jets in Figure A.7. The correlations are shown before the fit of the SM $VH(\to b\bar{b})$ analysis to the data. Slight differences exist after the fit, but the overall picture is the same.

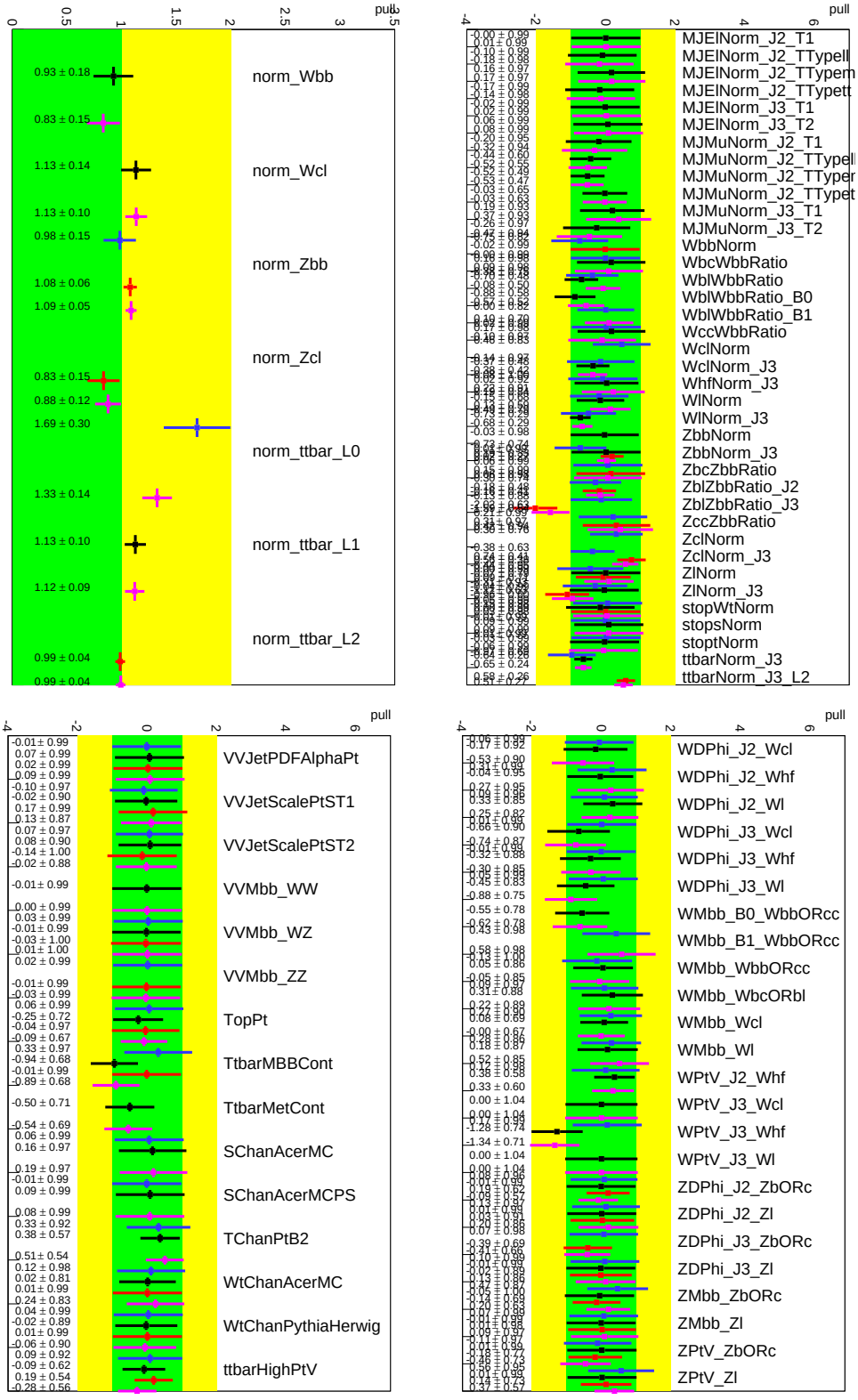


Figure A.1.: Nuisance parameter pulls from the fit to the data separately for the 0-lepton (blue), 1-lepton (black) and 2-lepton (red) channels and for the combined fit (magenta). Shown are the NPs for floating normalization (upper left), normalization with priors (upper right), top quark modeling (lower left) and V+jets modeling (lower right).

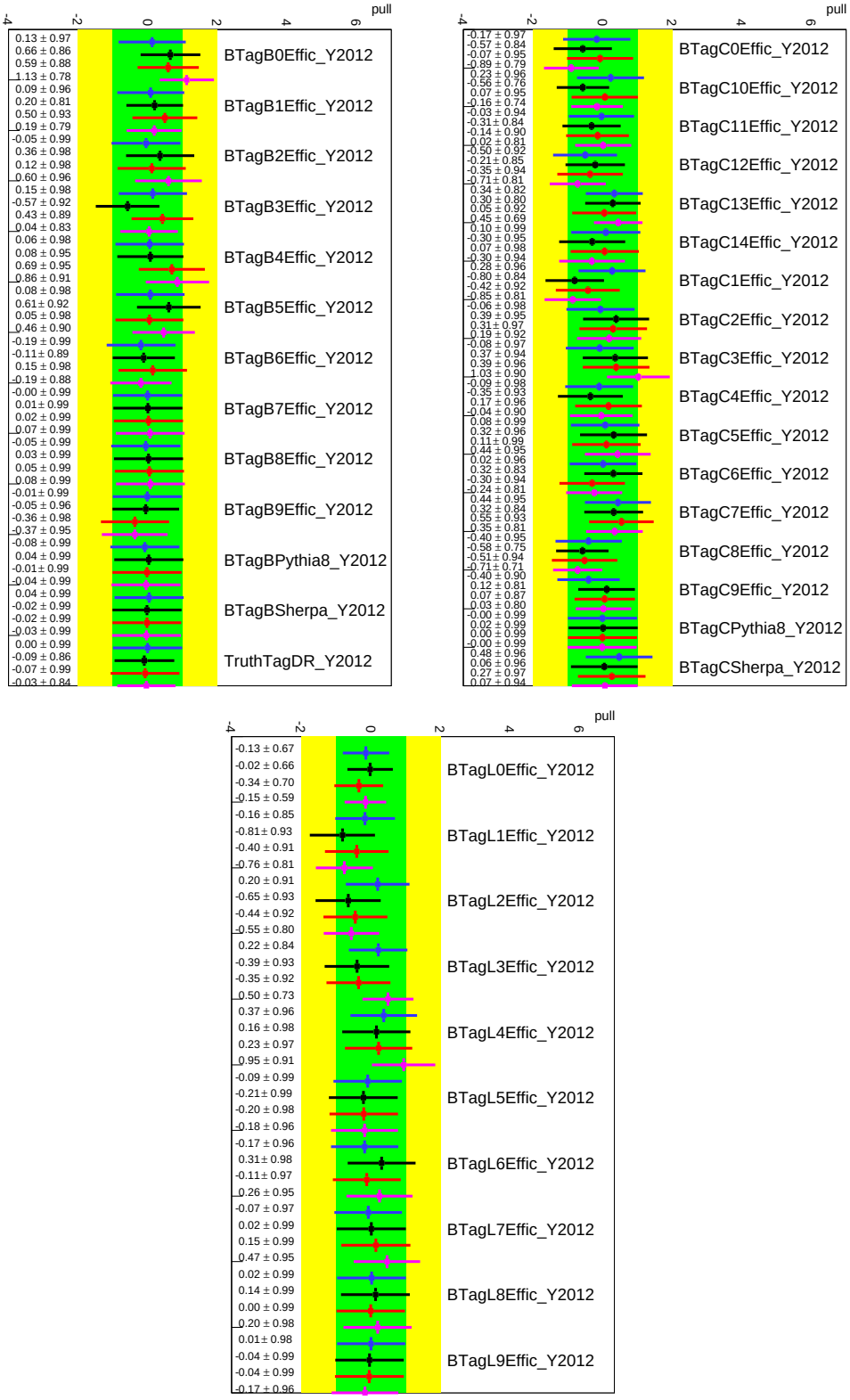


Figure A.2.: Nuisance parameter pulls from the combined fit to the data separately for the 0-lepton (blue), 1-lepton (black) and 2-lepton (red) channels and for the combined fit (magenta). Shown are the b tagging NPs for b jets (upper left), c jets (upper right) and light jets (bottom).

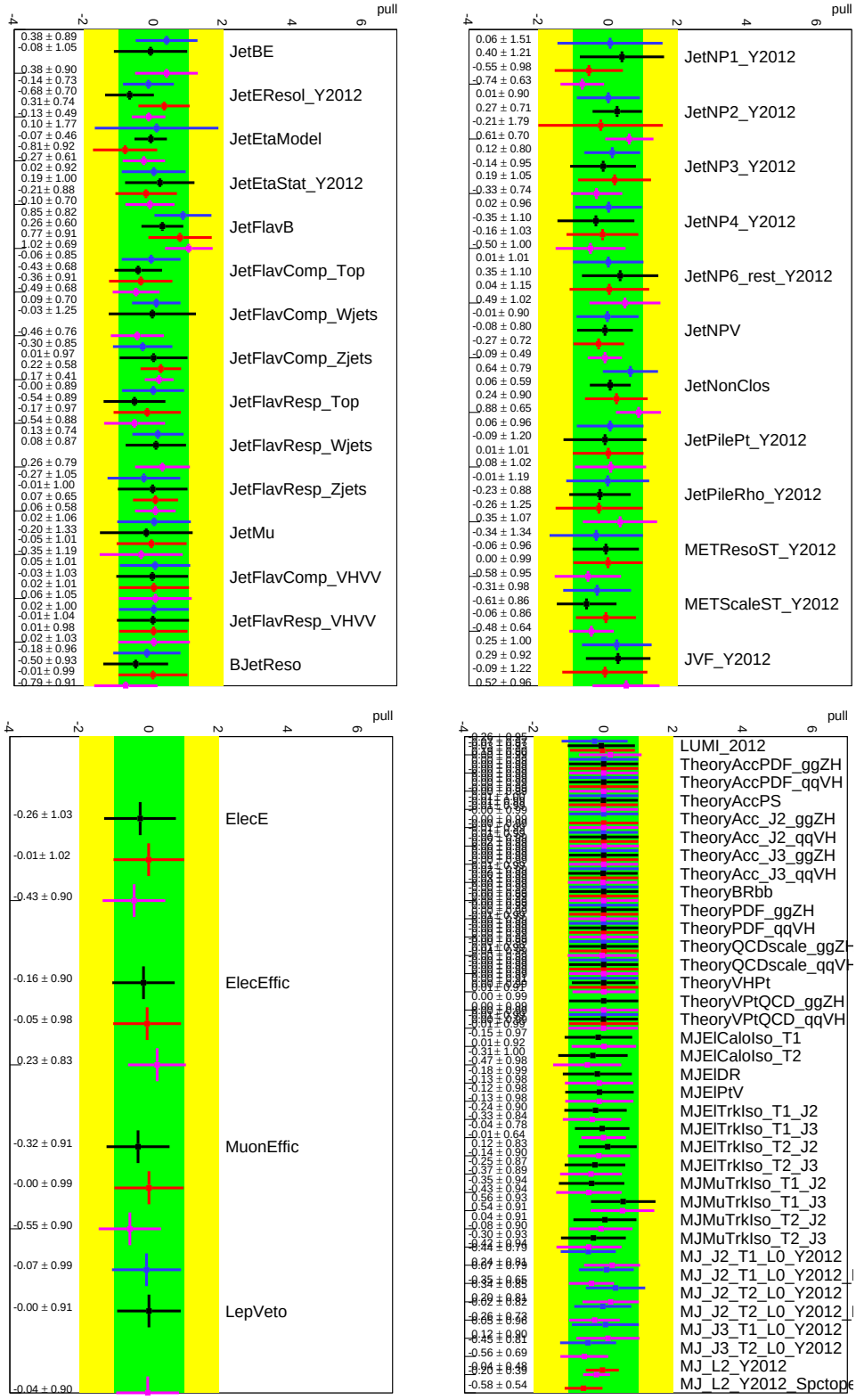


Figure A.3.: Nuisance parameter pulls from the combined fit to the data separately for the 0-lepton (blue), 1-lepton (black) and 2-lepton (red) channels and for the combined fit (magenta). Shown are the jet related NPs (top), lepton related NPs (lower left) and the remaining NPs (lower right).

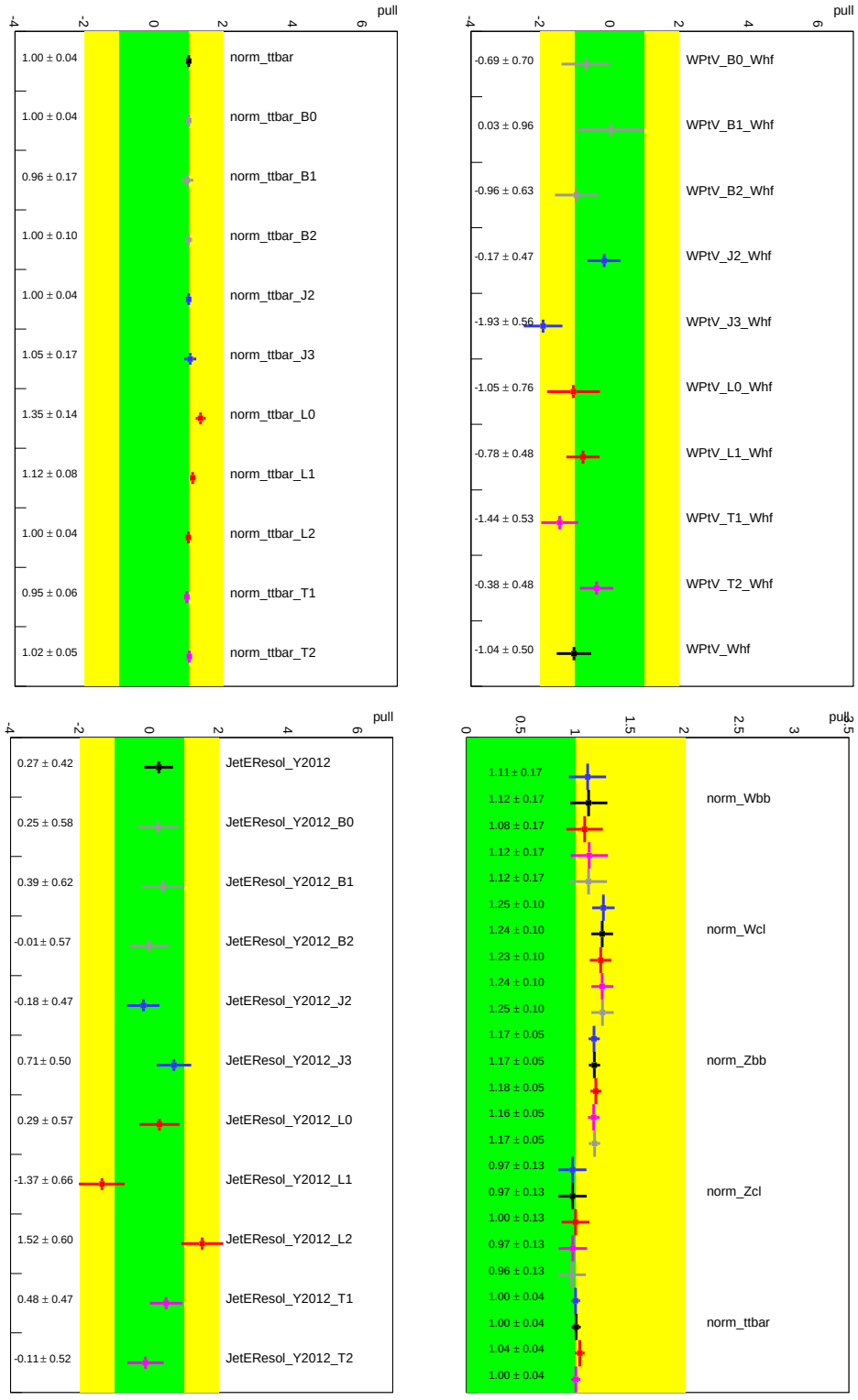


Figure A.4.: Nuisance parameter pulls from combined fits to the data with various de-correlation schemes of specific parameters. Shown are the de-correlation of `norm_ttbar` (upper left), `WPtV` (upper right) and `JetEResol` (lower left) with the corresponding change in the floating normalization parameters (lower right). Black: no de-correlation, gray: bins of p_{T}^{V}, blue: jet multiplicity, red: lepton channels, magenta: b-tagging regions.

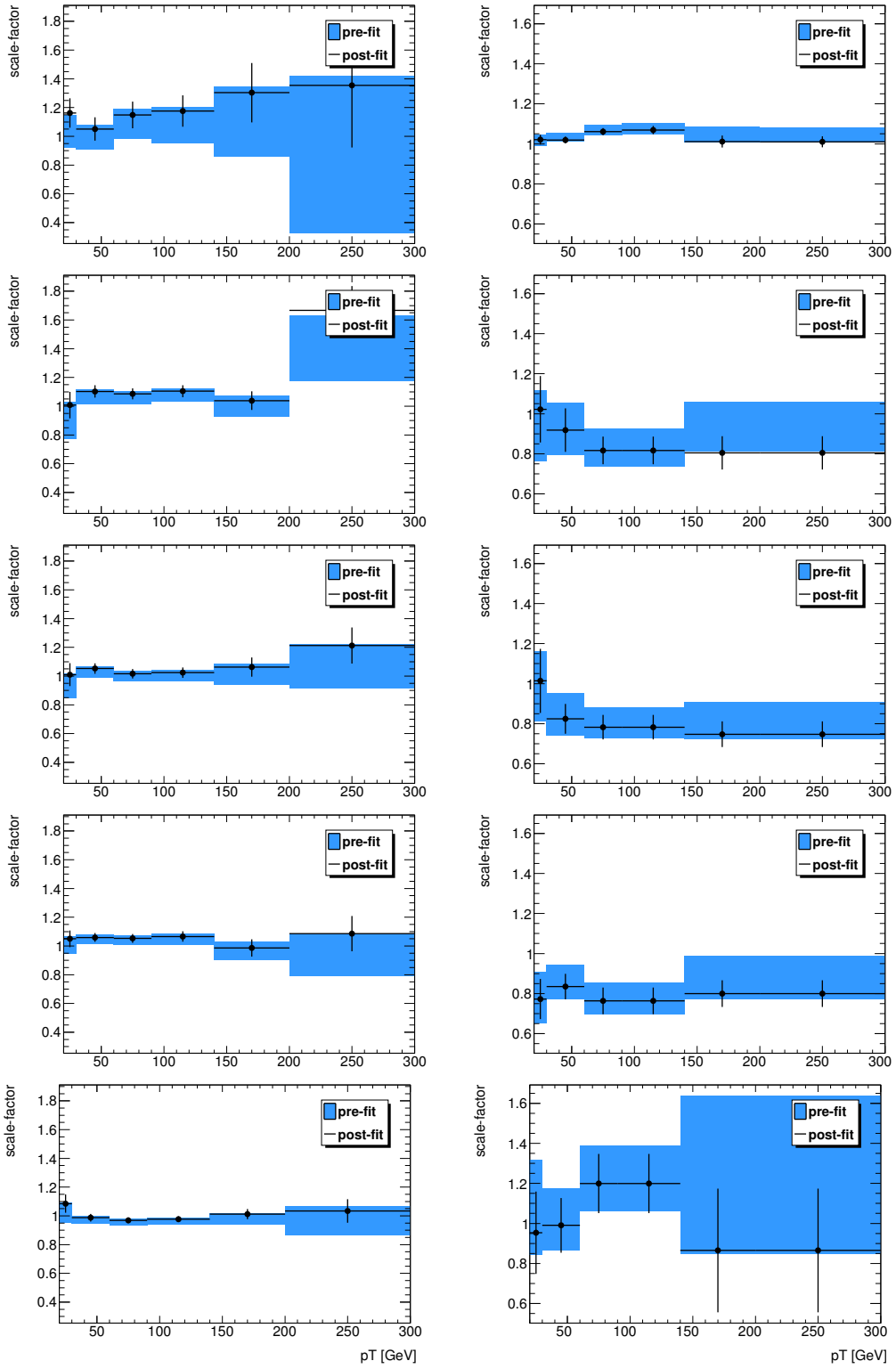


Figure A.5.: Calibration scale-factors for the MV1c algorithm before and after the fit to the data as a function of the jet p_{T} Ḷeft: b jets, right: c-jets. From top to bottom: five bins in MV1c with boundaries corresponding to 100 %, 80 %, 70 %, 60 %, 50 % and 0 % b-tagging efficiency.

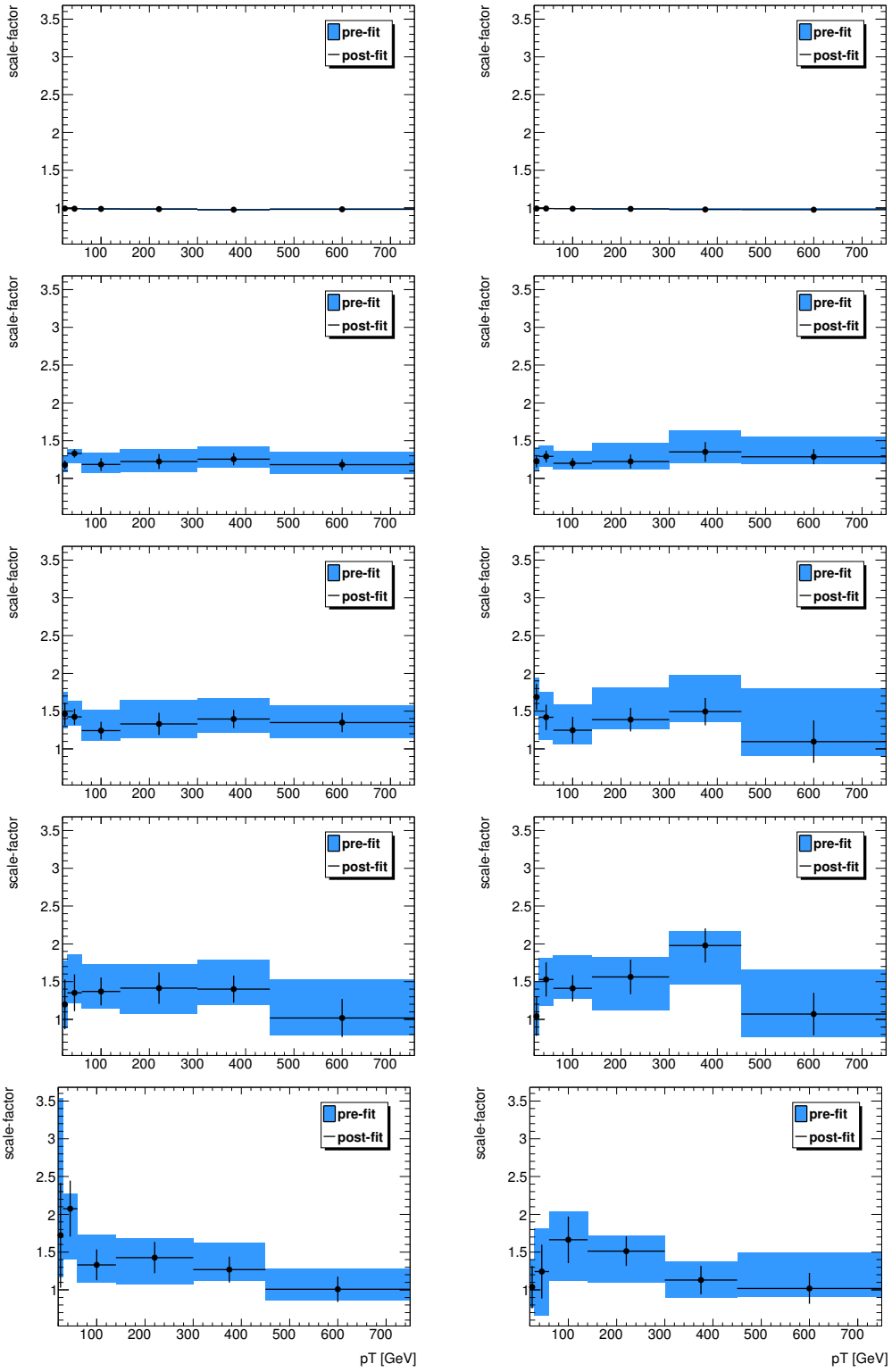


Figure A.6.: Calibration scale-factors for the MV1c algorithm before and after the fit to the data as a function of the jet p_{T} for light jets. Left: $|\eta| < 1.2$, right: $1.2 < |\eta| < 2.5$. From top to bottom: five bins in MV1c with boundaries corresponding to 100 %, 80 %, 70 %, 60 %, 50 % and 0 % b-tagging efficiency.

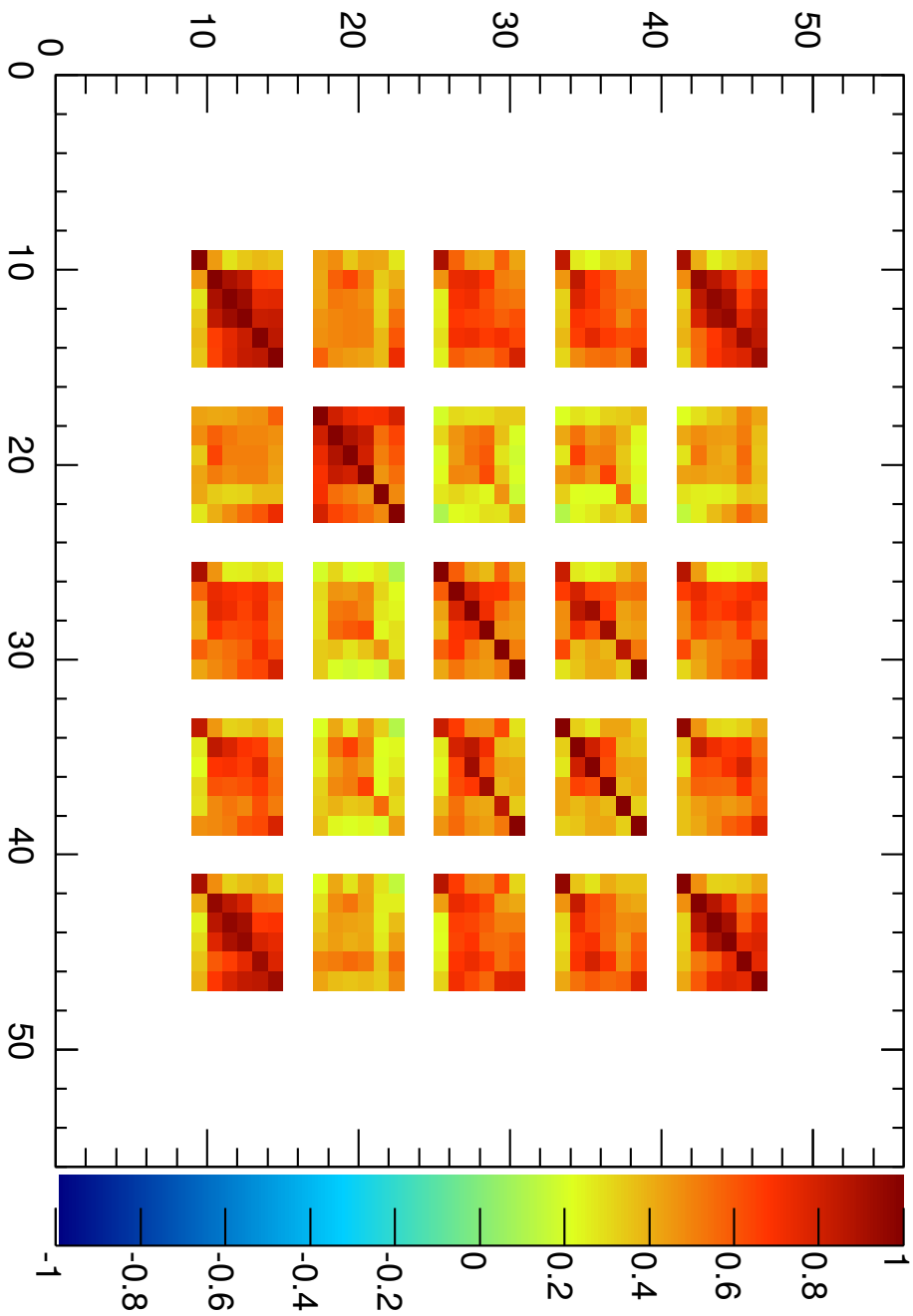


Figure A.7.: Example for the correlations in the b-tagging scale factors for b jets with the MV1c algorithm and pseudo-continuous calibration. The sub-matrices correspond to the five bins in MV1c and their entries correspond to six bins in the jet p_{T}.

Sample	1-tag		LL-tag		MM-tag	MM+TT-tag	TT-tag
p_{T}^{V}	$100-120\,\mathrm{GeV}$	$>120\,\mathrm{GeV}$	$100-120\,\mathrm{GeV}$	$>120\,\mathrm{GeV}$	$100-120\,\mathrm{GeV}$	$>120\,\mathrm{GeV}$	$100-120\,\mathrm{GeV}$
	2-jet						
VH	7.9	23	2.2	6.9	3.5	23	4.0
WZ, ZZ, WW	235	635	18	49	14.7	81	13.3
$t\bar{t}$	840	1520	114	183	129	332	116
Single top	531	704	40	56	32.6	66	22.8
Wl	5470	7100	159	206	14.4	16	0.2
Wcl	2230	3710	106	159	23.4	27	1.1
W+hf	762	1520	54	124	33.6	128	21.8
Zl	3890	10750	96	272	6.1	17	0.1
Zcl	1590	3990	59	162	9.4	24	0.3
Z+hf	2550	6510	225	607	186	876	151
Total	18340 ± 150	36890 ± 200	886 ± 17	1841 ± 25	458.5 ± 9.8	1599 ± 28	333.2 ± 8.9
Data	18343	36903	887	1860	477	1592	306
	3-jet						
VH	–	8	–	2	–	7	–
WZ, ZZ, WW	–	260	–	17	–	20	–
$t\bar{t}$	–	1670	–	186	–	315	–
Single top	–	318	–	25	–	30	–
Wl	–	2280	–	59	–	4.3	–
Wcl	–	1240	–	53	–	8.9	–
W+hf	–	750	–	60	–	62	–
Zl	–	3190	–	79	–	4.5	–
Zcl	–	1620	–	65	–	9.8	–
Z+hf	–	1890	–	170	–	259	–
Total	–	13310 ± 100	–	718 ± 12	–	719 ± 17	–
Data	–	13344	–	657	–	710	–

Table A.1.: Predicted and observed numbers of events for all regions of the 0-lepton channel. The yields and the combined statistical and systematic uncertainties of the backgrounds are shown after the combined fit to the data. The signal process is listed for $m_H = 125\,\mathrm{GeV}$ and $\mu = 1$ (from Ref. [1]).

A.3. Correlation matrices

The full correlation matrices from the fit to the data and to the Asimov dataset are shown in Figures A.8 and A.9, respectively.

The reduced correlation matrix (containing only parameters with at least one correlation with any other parameter above 20 %) as observed in the fit to the Asimov dataset is shown in Figure A.10.

No significant deviation between the correlations from the fit to the data and to the Asimov dataset is observed.

A.4. Post-fit yields

The numbers of events for all regions after the fit to the data for the 0-, 1- and 2-lepton channels are shown in Tables A.1, A.2 and A.3, respectively.

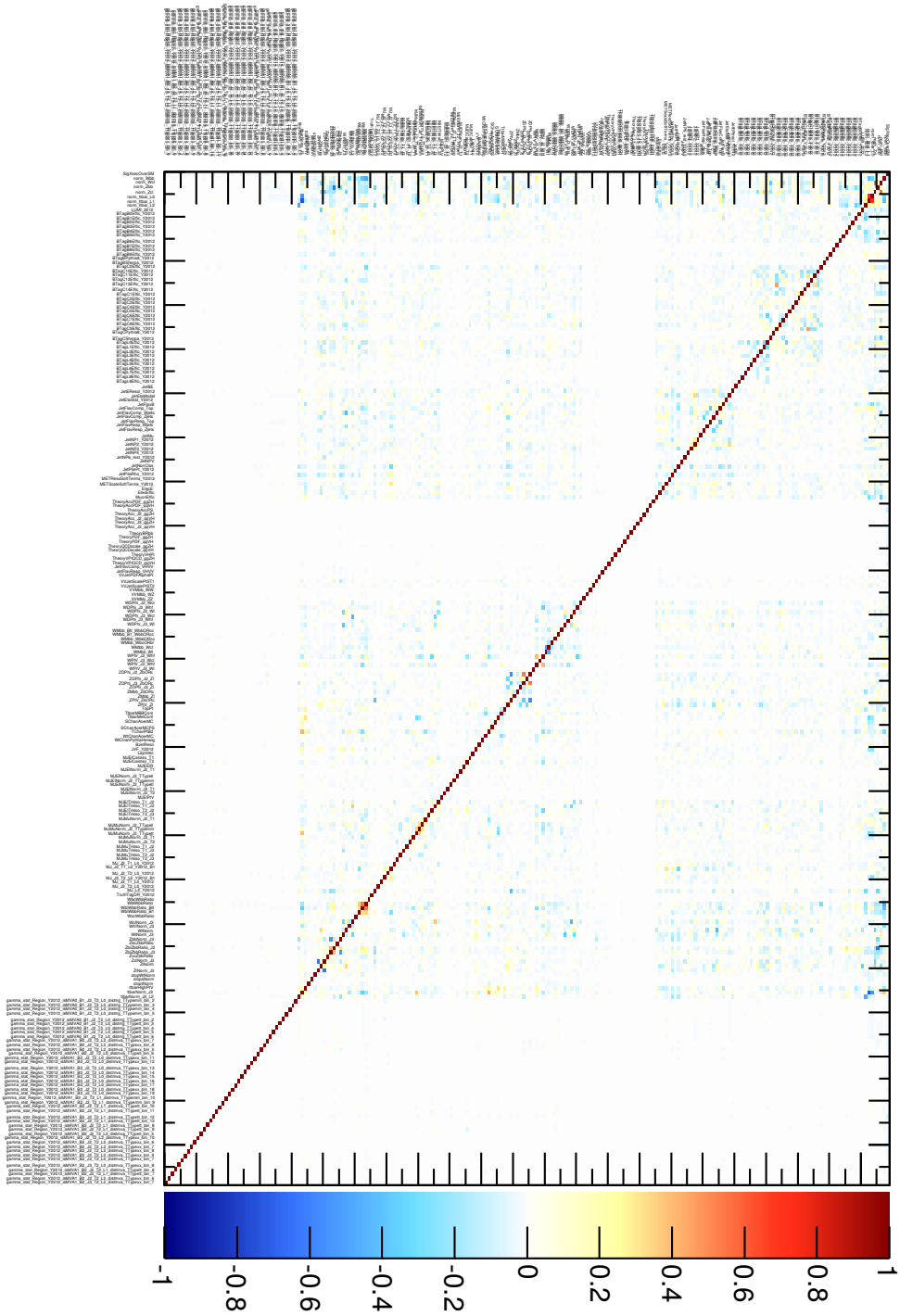


Figure A.8.: Correlation matrix from the fit to the data. All parameters of the fit are shown.

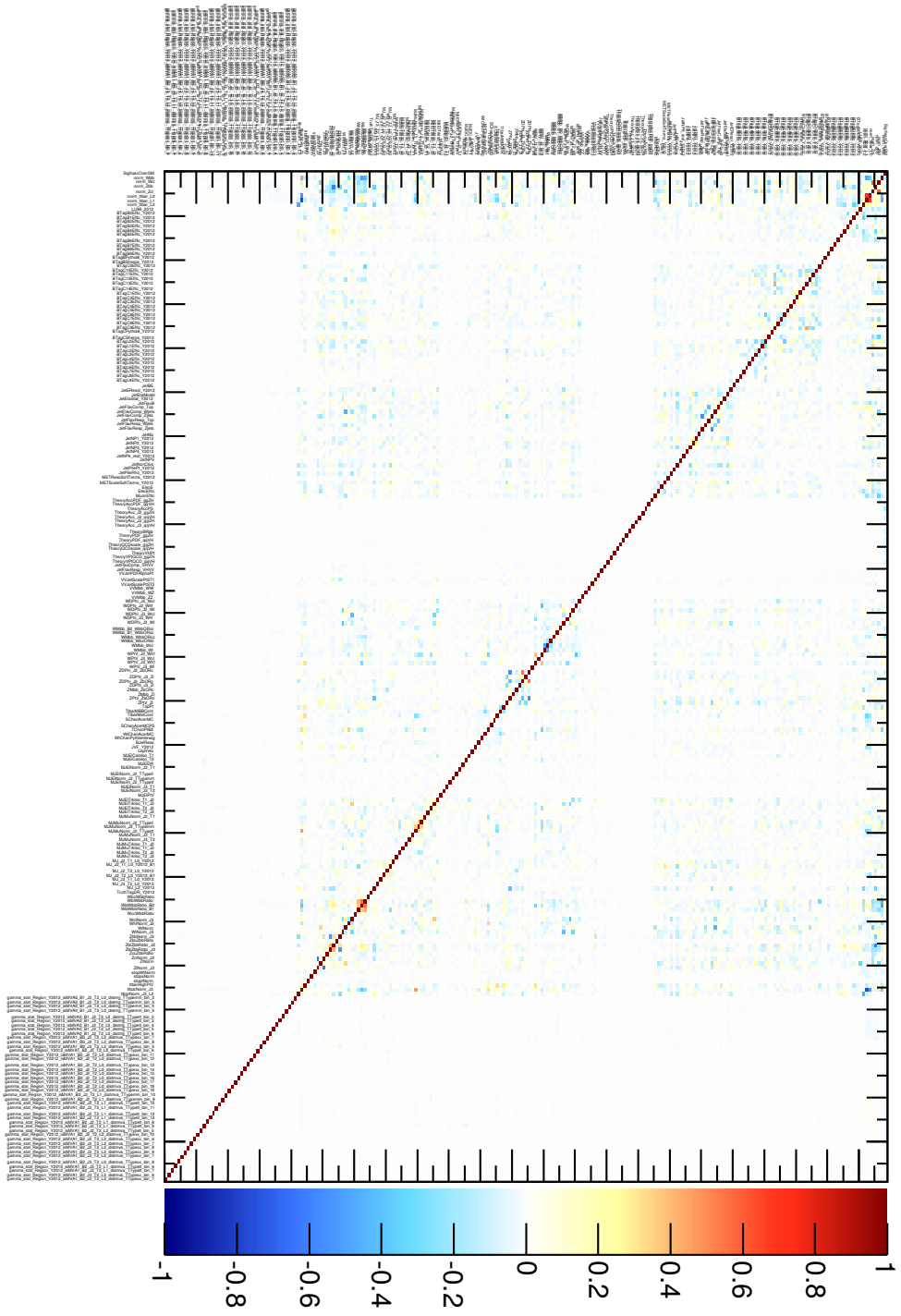


Figure A.9.: Correlation matrix from the fit to the Asimov dataset. All parameters of the fit are shown.

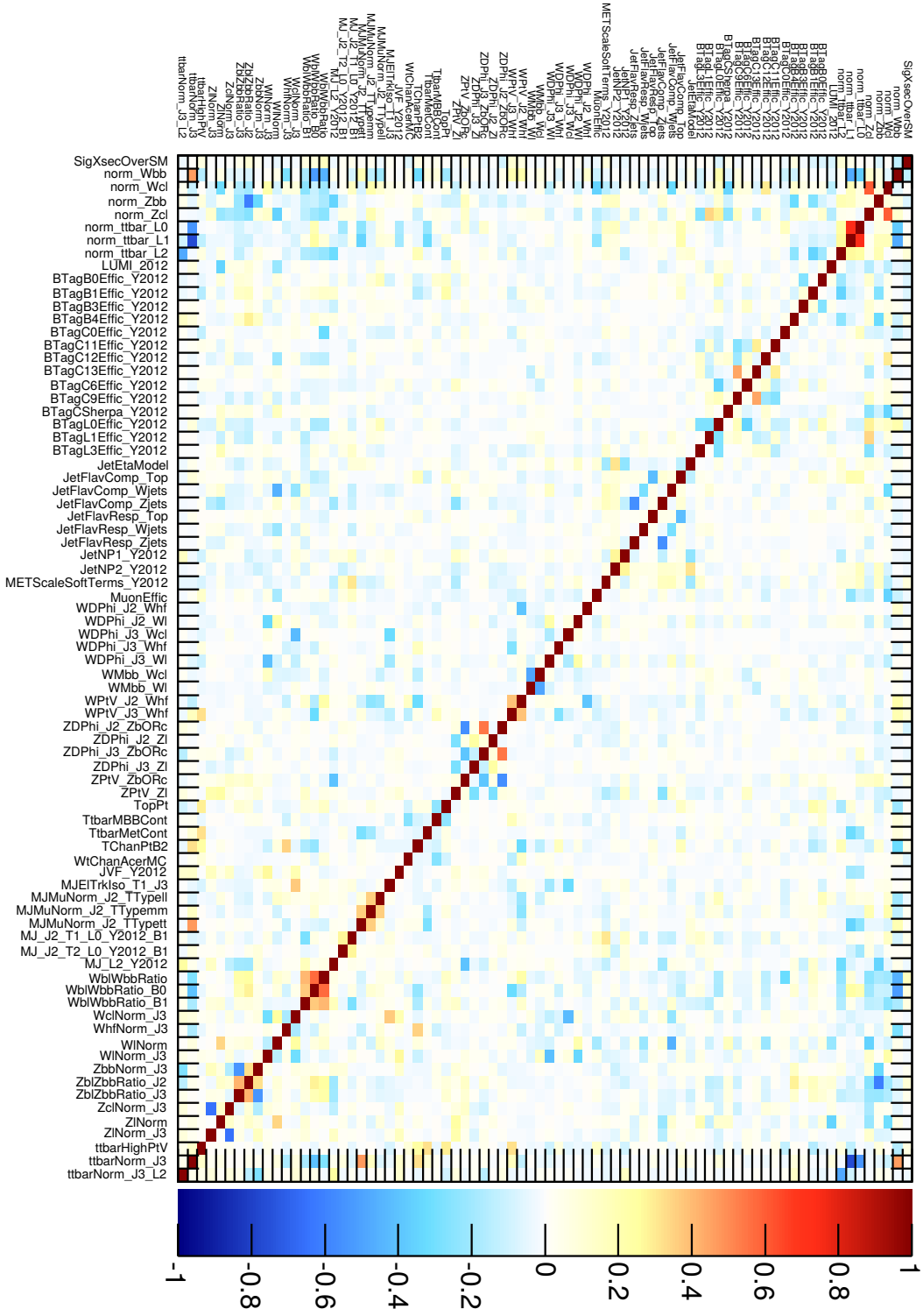


Figure A.10.: Correlation matrix from the fit to the Asimov dataset. Only parameters with at least one correlation with any other parameter larger than 20 % are shown.

Sample	1-tag		LL-tag		MM-tag		TT-tag	
p_{T}^{V}	< 120 GeV	> 120	< 120 GeV	> 120 GeV	< 120 GeV	> 120 GeV	< 120 GeV	> 120 GeV
	2-jet							
VH	39	28	11	9.3	17	14	19	16
WZ, ZZ, WW	1950	927	103.6	62	64	36	52	29.5
$t\bar{t}$	11380	6641	1954	1051	2426	1080	2290	890
Single top	13680	3730	1150	398	975	307	739	219
Wl	65980	23702	1603	697	124	48	3.1	0.9
Wcl	71930	21650	2630	966	465	153	23	6.5
W+hf	16030	6112	1470	648	954	402	506	227
Zl	3940	1223	101	37	7.4	2.8	0.1	0.0
Zcl	1350	333	53	18	10	3.1	0.5	0.2
Z+hf	2080	475	161	45	126	30	85	24
MJ (e)	–	2618	–	162	–	89	–	40
MJ (μ)	10230	164	721	16.0	329	4.8	178	1.3
Total	198540 ± 500	67600 ± 290	9953 ± 91	4106 ± 50	5492 ± 66	2161 ± 33	3889 ± 55	1448 ± 27
Data	198544	67603	9941	4072	5499	2199	3923	1405
	3-jet							
VH	15	14	3.2	3.8	4.8	5.8	5.4	6.5
WZ,ZZ,WW	1100	689	50	39.6	22.6	18	14	14
$t\bar{t}$	18660	10490	3240	1622	4119	1670	4181	1388
Single top	7390	2815	66	318	619	261	503	188
W+l	24980	11320	588	322	42	20	1.1	0.4
W+cl	25900	10080	952	454	164	72	7.7	3.2
W+hf	6530	4740	576	490	353	297	187	168
Z+l	1780	572	43	18.1	2.8	1.4	0.0	0.0
Z+cl	690	193	27	9.8	4.5	1.6	0.2	0.1
Z+hf	1024	272	77	25.9	54	18.8	40	14
MJ (e)	–	1290	–	68.6	–	36	–	15
MJ (μ)	5300	91	227	4.9	117	3.2	58	0.8
Total	93350 ± 320	42570 ± 200	6447 ± 57	3376 ± 43	5501 ± 50	2405 ± 33	4995 ± 55	1796 ± 30
Data	93359	42557	6336	3472	5551	2356	4977	1838

Table A.2.: Predicted and observed numbers of events for all regions of the 1-lepton channel. The yields and the combined statistical and systematic uncertainties of the backgrounds are shown after the combined fit to the data. The signal process is listed for $m_H = 125\,\mathrm{GeV}$ and $\mu = 1$ (from Ref. [1]).

Sample	1-tag		LL-tag		MM-tag		TT-tag	
p_{T}^{V}	$< 120\,\mathrm{GeV}$	> 120	$< 120\,\mathrm{GeV}$	$> 120\,\mathrm{GeV}$	$< 120\,\mathrm{GeV}$	$> 120\,\mathrm{GeV}$	$< 120\,\mathrm{GeV}$	$> 120\,\mathrm{GeV}$
	2-jet							
VH	39	28	11	9.3	17	14	19	16
WZ, ZZ, WW	1950	927	103.6	62	64	36	52	29.5
$t\bar{t}$	11380	6641	1954	1051	2426	1080	2290	890
Single top	13680	3730	1150	398	975	307	739	219
Wl	65980	23702	1603	697	124	48	3.1	0.9
Wcl	71930	21650	2630	966	465	153	23	6.5
W+hf	16030	6112	1470	648	954	402	506	227
Zl	3940	1223	101	37	7.4	2.8	0.1	0.0
Zcl	1350	333	53	18	10	3.1	0.5	0.2
Z+hf	2080	475	161	45	126	30	85	24
MJ (e)	–	2618	–	162	–	89	–	40
MJ (μ)	10230	164	721	16.0	329	4.8	178	1.3
Total	198540 ± 500	67600 ± 290	9953 ± 91	4106 ± 50	5492 ± 66	2161 ± 33	3889 ± 55	1448 ± 27
Data	198544	67603	9941	4072	5499	2199	3923	1405
	3-jet							
VH	15	14	3.2	3.8	4.8	5.8	5.4	6.5
WZ,ZZ,WW	1100	689	50	39.6	22.6	18	14	14
$t\bar{t}$	18660	10490	3240	1622	4119	1670	4181	1388
Single top	7390	2815	66	318	619	261	503	188
W+l	24980	11320	588	322	42	20	1.1	0.4
W+cl	25900	10080	952	454	164	72	7.7	3.2
W+hf	6530	4740	576	490	353	297	187	168
Z+l	1780	572	43	18.1	2.8	1.4	0.0	0.0
Z+cl	690	193	27	9.8	4.5	1.6	0.2	0.1
Z+hf	1024	272	77	25.9	54	18.8	40	14
MJ (e)	–	1290	–	68.6	–	36	–	15
MJ (μ)	5300	91	227	4.9	117	3.2	58	0.8
Total	93350 ± 320	42570 ± 200	6447 ± 57	3376 ± 43	5501 ± 50	2405 ± 33	4995 ± 55	1796 ± 30
Data	93359	42557	6336	3472	5551	2356	4977	1838

Table A.3.: Predicted and observed numbers of events for all regions of the 2-lepton channel. The yields and the combined statistical and systematic uncertainties of the backgrounds are shown after the combined fit to the data. The signal process is listed for $m_H = 125\,\mathrm{GeV}$ and $\mu = 1$ (from Ref. [1]).

B. Gluon fusion H(bb) appendix

This appendix contains additional information for the sensitivity estimate for the SM $gg \to H \to b\bar{b}$ process, discussed in Chapter 6.

B.1. Signal sample validation

The signal samples have been generated and validated in the course of this thesis, as described in the following. The samples are generated in slices of truth-level p_{T}^{H}. The chosen ranges of p_{T}^{H} with the corresponding filter efficiencies are shown in Table B.1.

An inclusive sample is generated as well. The slices are validated by comparing the p_{T}^{H} spectra on truth and reconstruction level with the inclusive sample, as shown in Figure B.1. Good agreement is observed as expected.

After the validation, larger samples with 50 k events for each slice are generated and propagated through the full ATLAS simulation. The corresponding ATLAS MC channel numbers are listed in Table B.1 as well.

p_{T}^{H} [GeV]	total	selected	efficiency	yield	MC channel
inclusive	5000	5000	1.000	154530	206728
30 – 70	15416	5000	0.324	50120	(not requested)
70 – 120	30000	3380	0.113	17410	206729
120 – 200	30000	1128	0.038	5810	206730
> 200	30000	295	0.010	1520	206731

Table B.1.: Number of total and selected events for a filter using the p_{T}^{H} on truth level. The selection efficiency and the expected event yield, scaled to $20.3\,\mathrm{fb}^{-1}$, are shown as well. The numbers are evaluated on a small test sample. The last column lists the ATLAS MC channel numbers for the final samples.

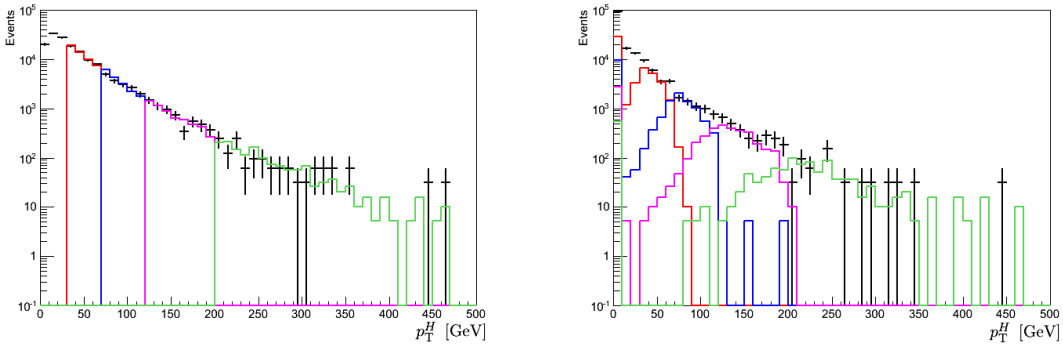


Figure B.1.: Comparison of the p_{T}^{H} spectra of the sliced samples (colored) with the inclusive sample (black) on truth level (left) and on reconstruction level (right).

C. VH(invisible) appendix

This appendix contains additional information for the search for the BSM $VH(\rightarrow \text{inv.})$ process, discussed in Chapter 7.

C.1. Equal bin-width plots

All post-fit distributions that are used as input to the fit are shown in Figure C.1 ($E_{\mathrm{T}}^{\mathrm{miss}}$ for the 0-lepton signal region) and Figure C.2 (side bands), Figure C.3 (p_{T}^{V} for the 1- and 2-lepton control regions). The normalization in the top control region is omitted.

These plots correspond to the ones shown in Section 7.6.1, but all flavor components of the backgrounds are shown separately here. Further, the bins are drawn with equal widths on the x-axis. This way the distributions are less physical, but correspond to what the fit actually "sees".

C.2. Correlation matrices

The reduced correlation matrix (containing only parameters with at least one correlation with any other parameter above 20 %) as observed in the fit to the Asimov dataset is shown in Figure C.4.

The full correlation matrices from the fit to the data and to the Asimov dataset are shown in Figures C.5 and C.6, respectively.

No significant deviation between the correlations from the fit to the data and to the Asimov dataset is observed.

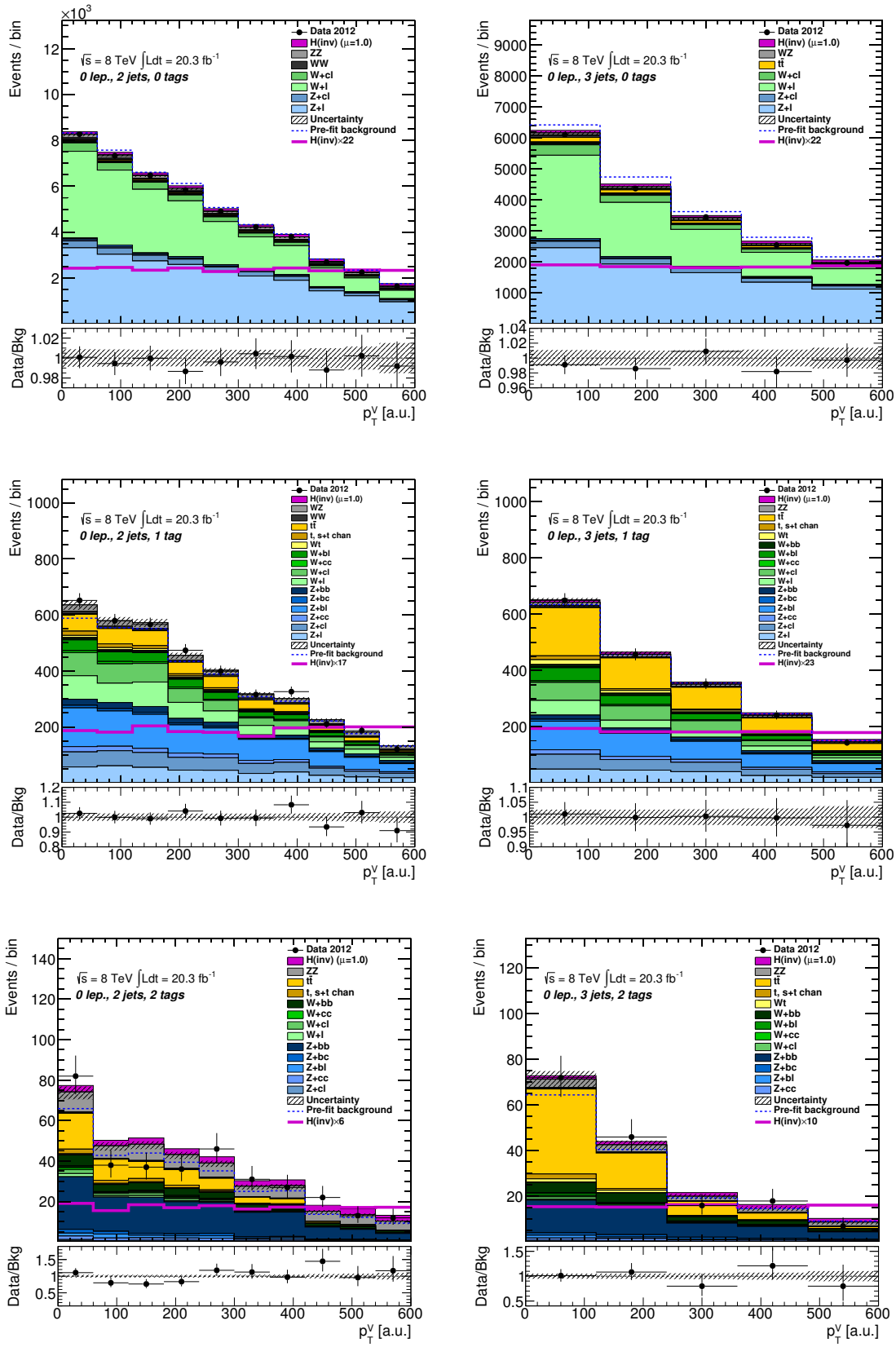


Figure C.1.: Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions with equal bin-width of the 0-lepton signal region for the 2-jet (left) and 3-jet (right) and (top to bottom) 0-, 1- and 2-tag regions (from Ref. [2]).

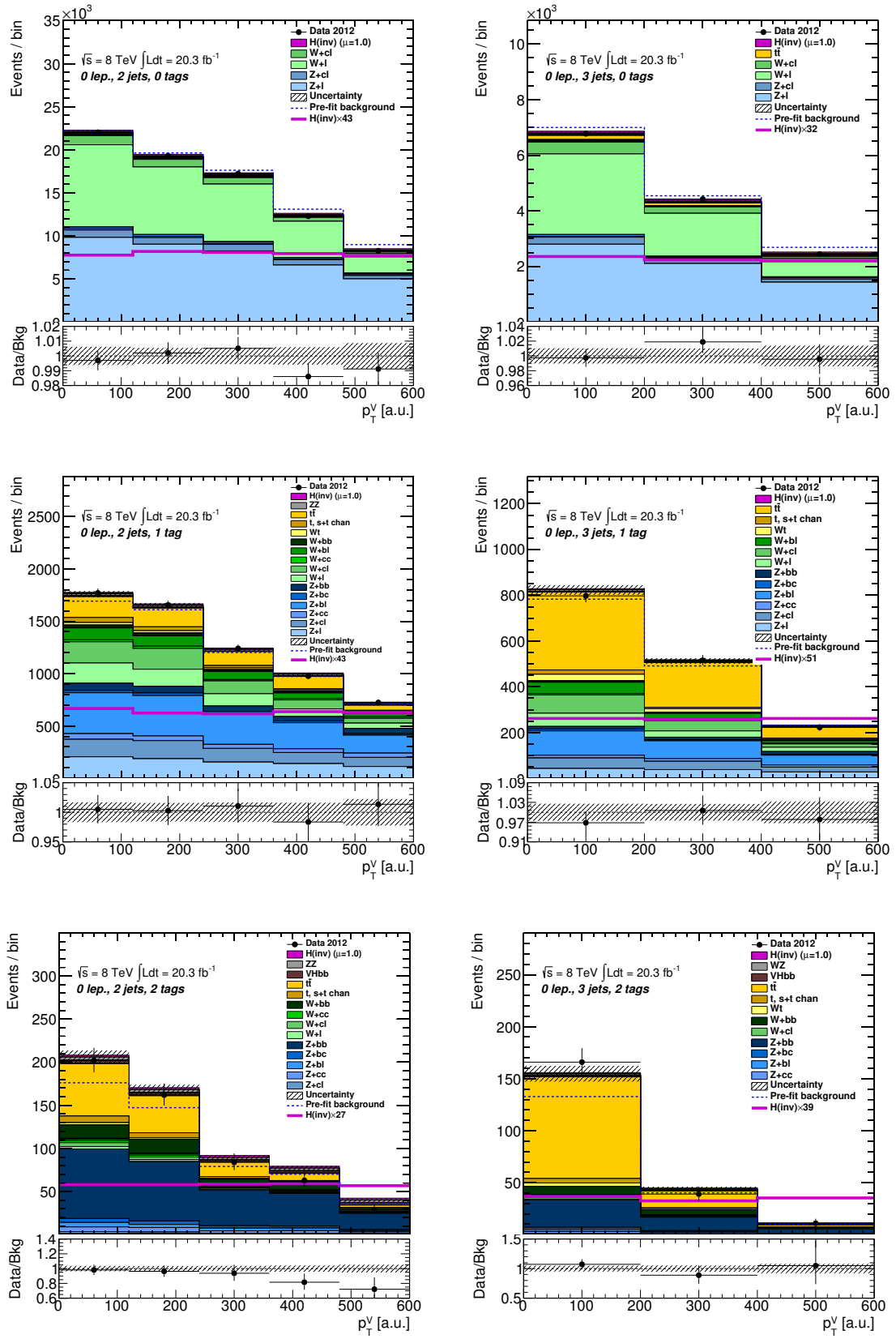


Figure C.2.: Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions with equal bin-width of the 0-lepton side-bands for the 2-jet (left) and 3-jet (right) and (top to bottom) 0-, 1- and 2-tag regions (from Ref. [2]).

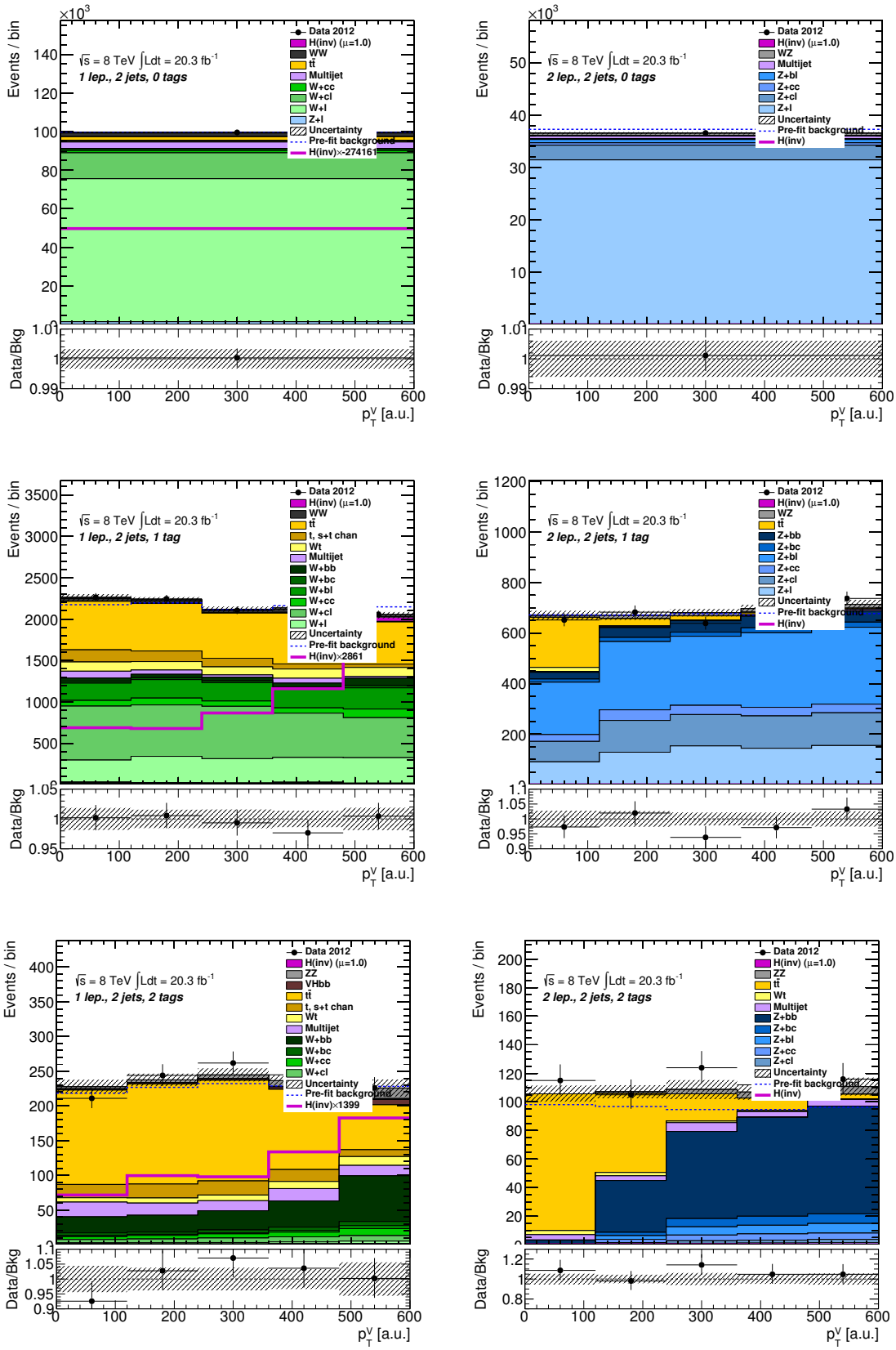


Figure C.3.: Post-fit p_{T}^{V} distributions with equal bin-width of the 1-lepton (left) and 2-lepton (right) channels for the (top to bottom) 0-, 1- and 2-tag regions (from Ref. [2]).

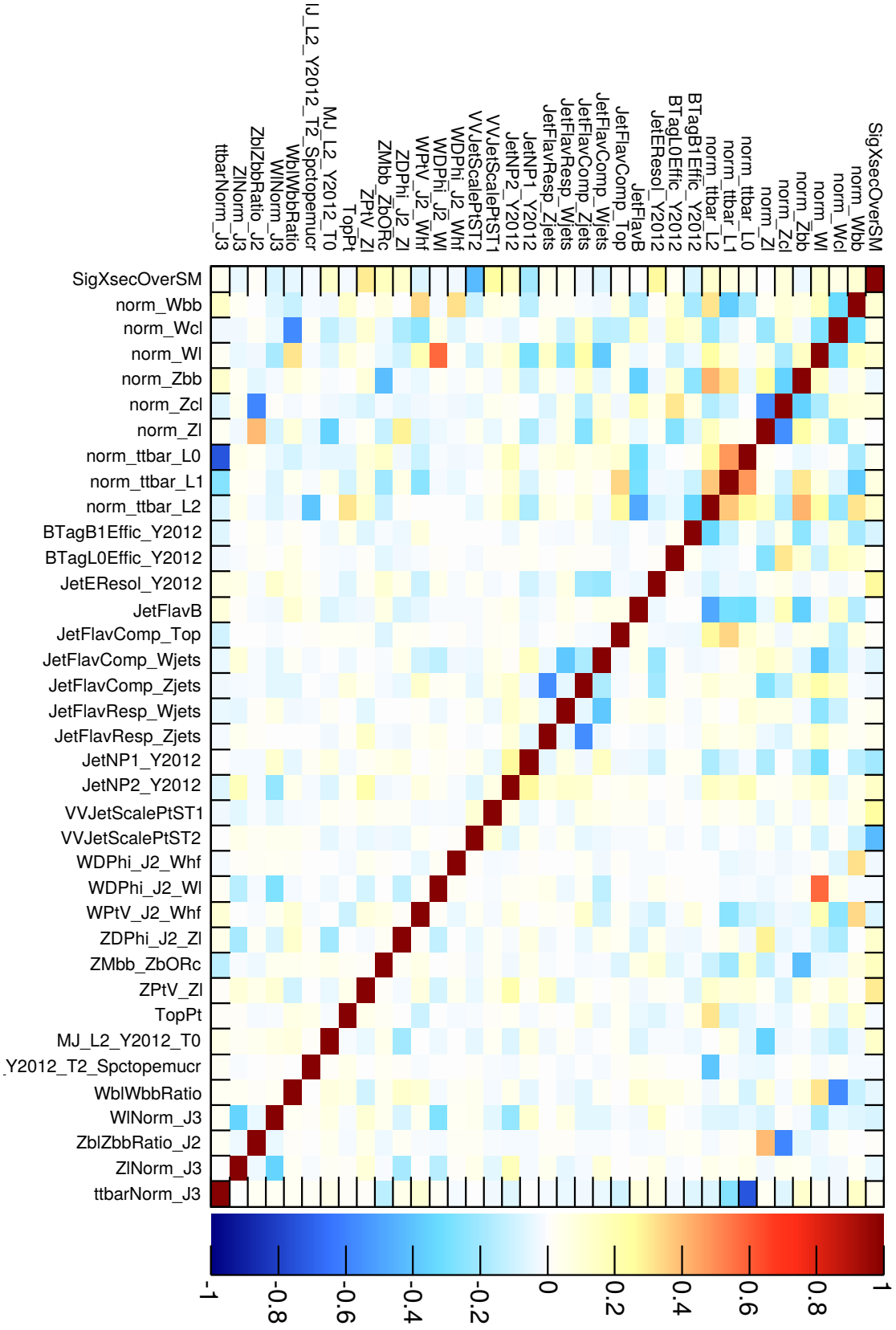


Figure C.4.: Correlation matrix from the fit to the Asimov dataset. Only parameters with at least one correlation with any other parameter larger than 20 % are shown.

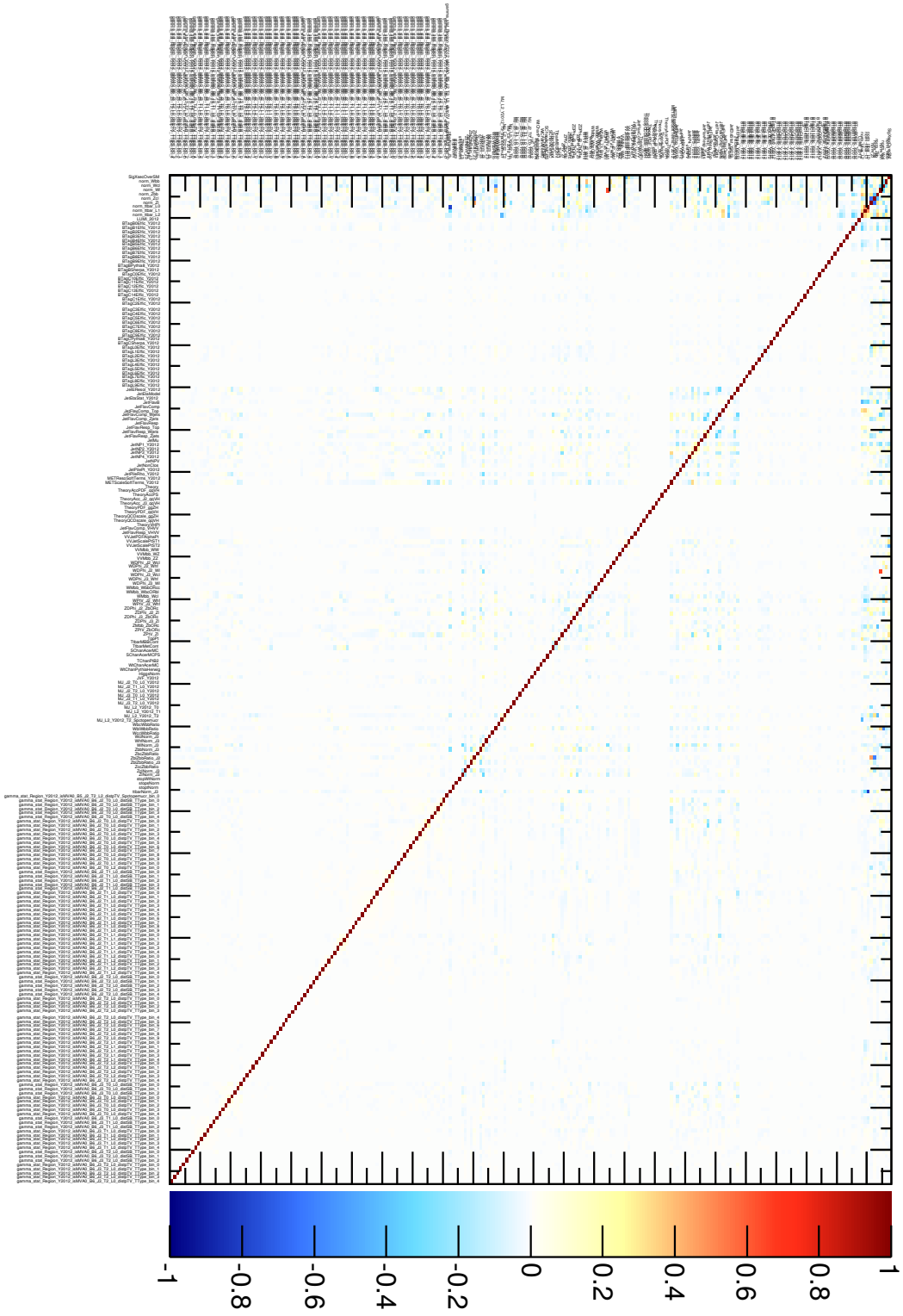


Figure C.5.: Correlation matrix from the fit to the data. All parameters of the fit are shown.

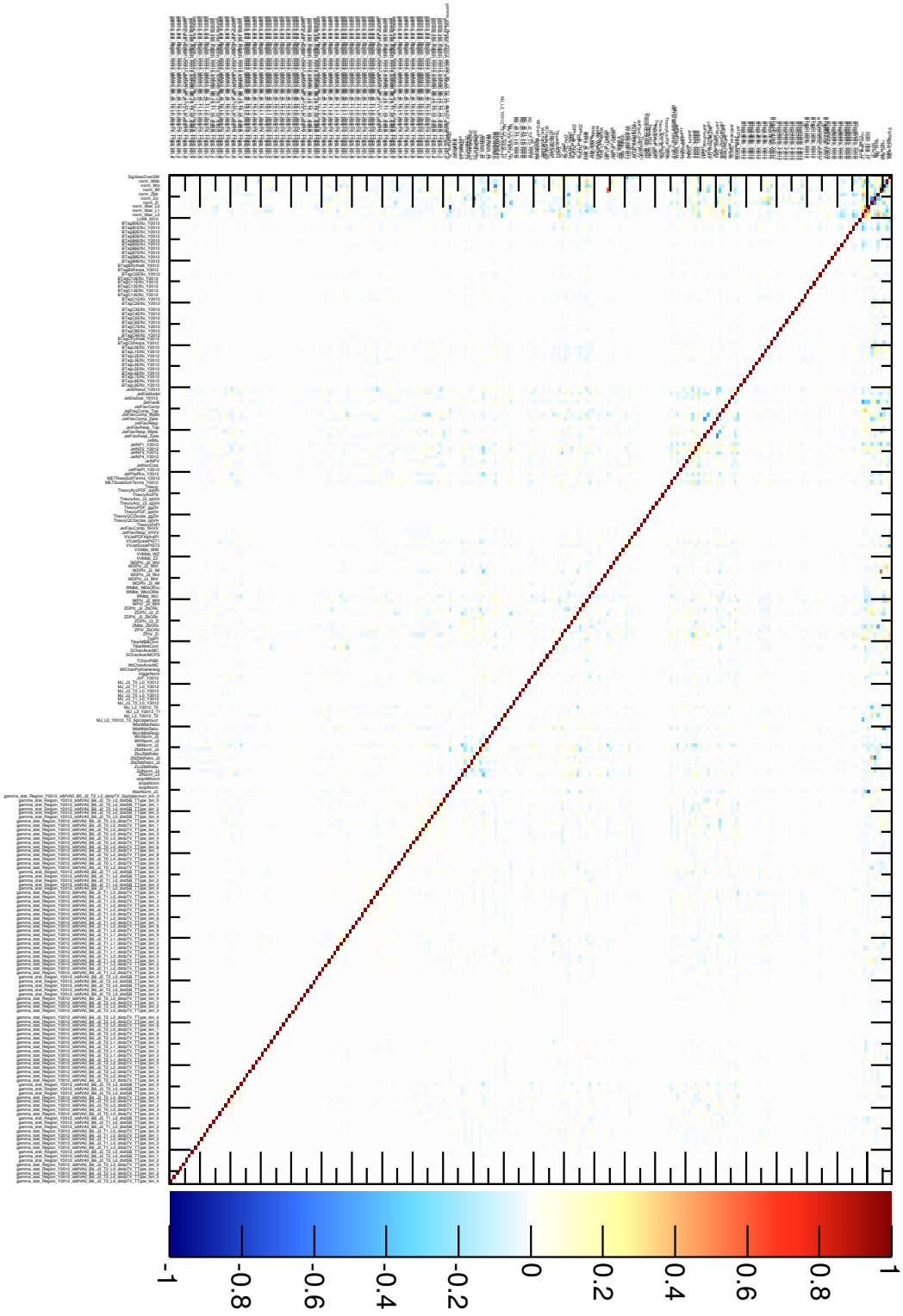


Figure C.6.: Correlation matrix from the fit to the Asimov dataset. All parameters of the fit are shown.

D. VH resonances appendix

This appendix contains additional information for the search for BSM VH resonances with $H \to b\bar{b}$, discussed in Chapter 8.

D.1. Correlation matrices

The full correlation matrices from the fit to the data and to the Asimov dataset for the V' combination are shown in Figures D.1 and D.2, respectively.

The correlation matrix from the fit to the Asimov dataset for the V' combination including all parameters of the fit, except for the bin-wise statistical uncertainties of the simulated backgrounds, is shown in Figures D.3.

No significant deviation between the correlations from the fit to the data and to the Asimov dataset is observed.

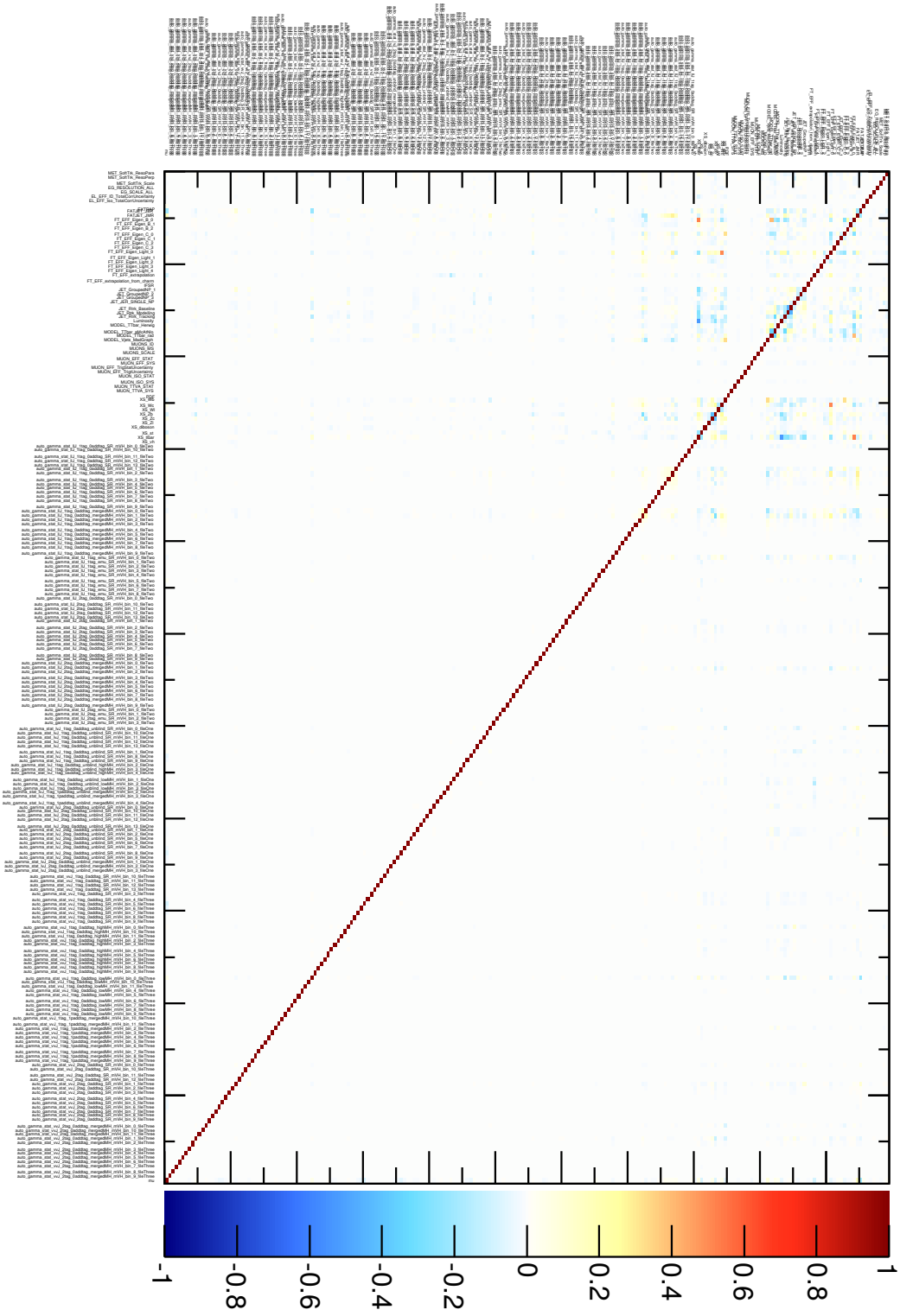


Figure D.1.: Correlation matrix from the fit to the data. All parameters of the fit are shown, including the bin-wise statistical uncertainties of the simulated backgrounds ("gamma" parameters).

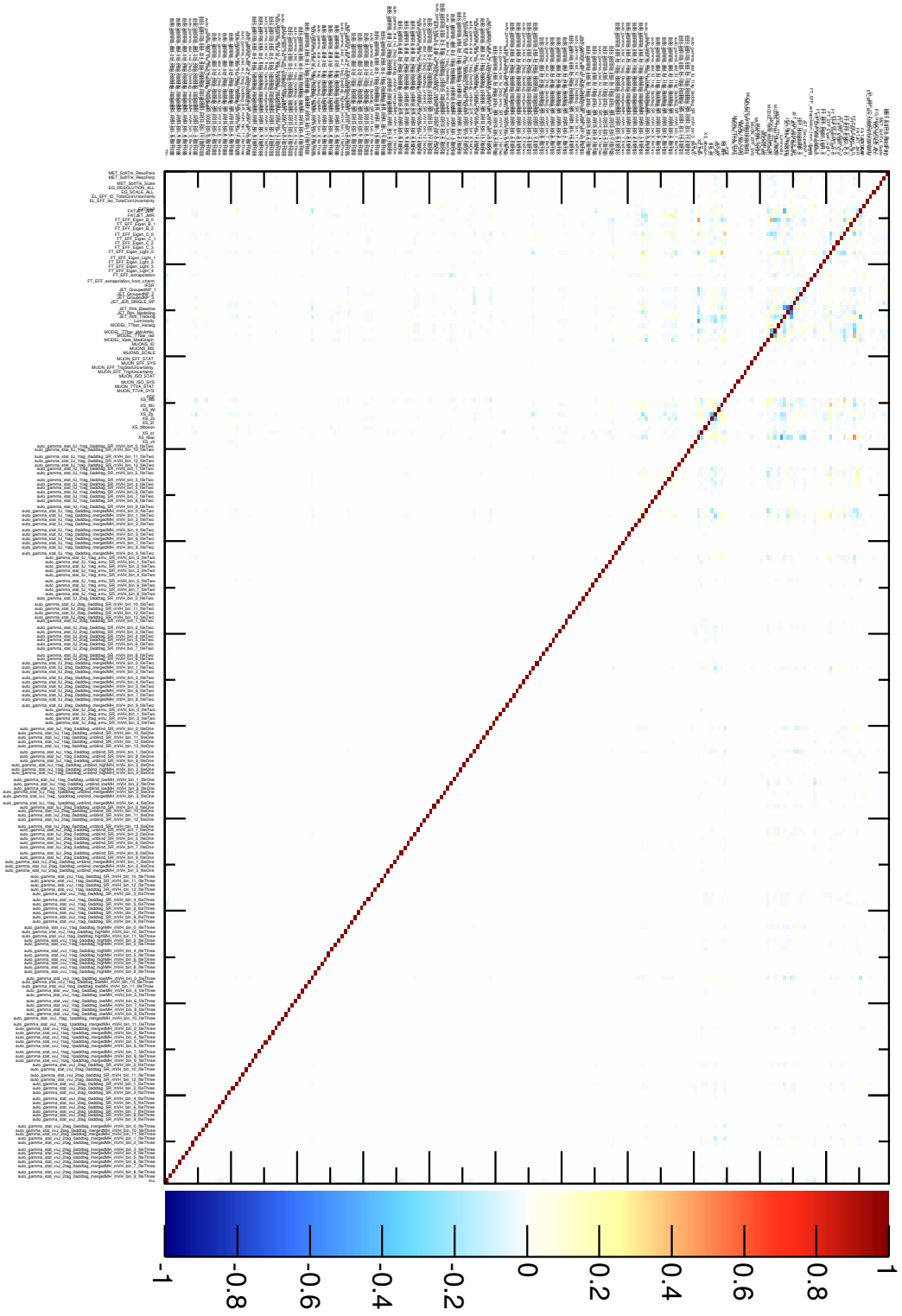


Figure D.2.: Correlation matrix from the fit to the Asimov dataset. All parameters of the fit are shown, including the bin-wise statistical uncertainties of the simulated backgrounds ("gamma" parameters).

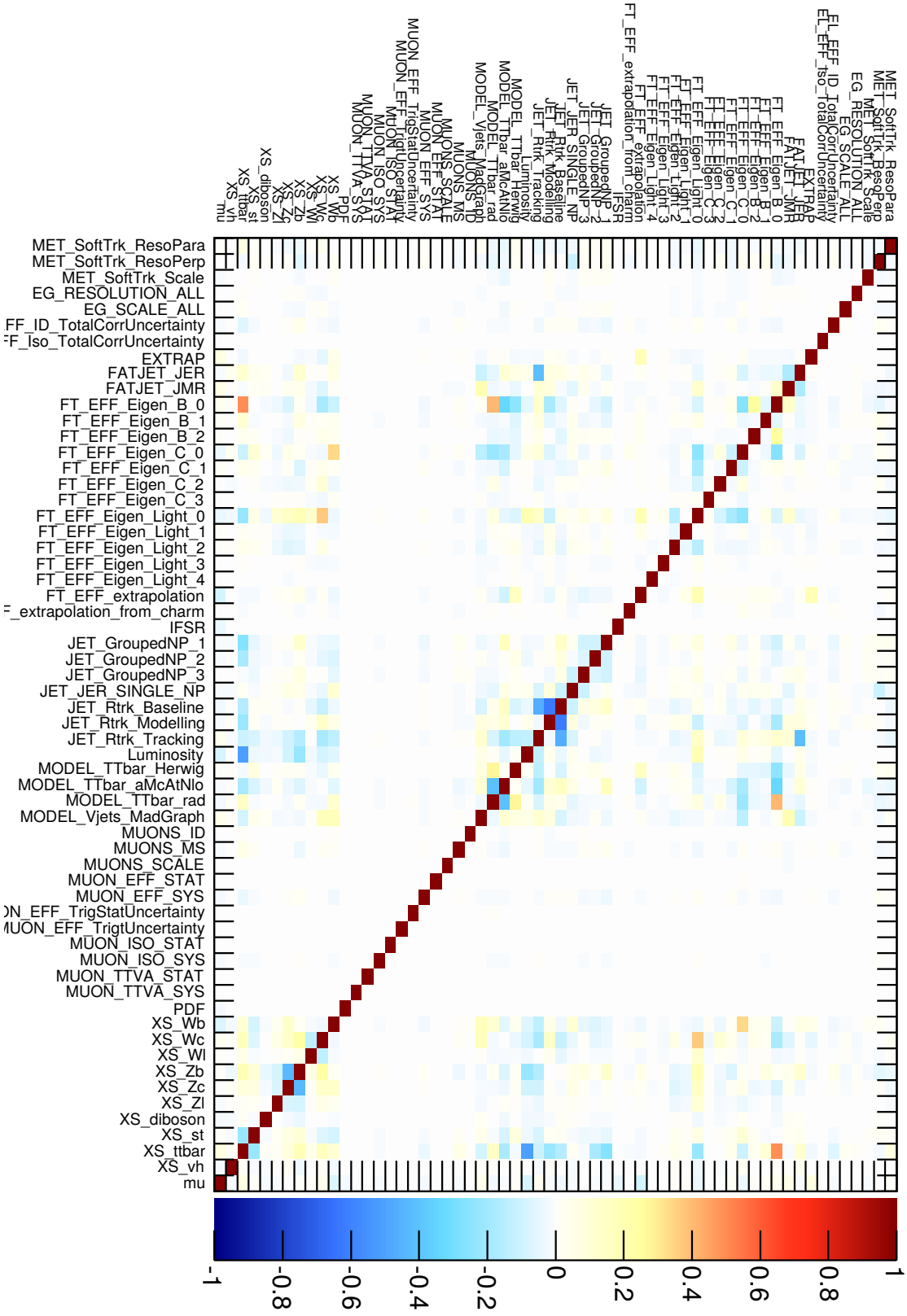


Figure D.3.: Correlation matrix from the fit to the Asimov dataset. All parameters of the fit are shown, except for the bin-wise statistical uncertainties of the simulated backgrounds.

Bibliography

[1] ATLAS Collaboration, *Search for the decay of the Standard Model Higgs boson in associated (W/Z) H production with the ATLAS detector*, Journal of High Energy Physics **1** (Jan., 2015) 69, arXiv:1409.6212 [hep-ex].

[2] ATLAS Collaboration, *Search for invisible decays of the Higgs boson produced in association with a hadronically decaying vector boson in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, European Physical Journal C **75** (July, 2015) 337, arXiv:1504.04324 [hep-ex].

[3] ATLAS Collaboration, *Search for new resonances decaying to a W or Z boson and a Higgs boson in the $\ell^+\ell^- b\bar{b}$, $\ell\nu b\bar{b}$, and $\nu\bar{\nu}b\bar{b}$ channels with pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ArXiv e-prints (July, 2016) , arXiv:1607.05621 [hep-ex].

[4] G. Thubes, *Carlo Rubbia and the discovery of the W and the Z*, Phys. World **16** (2003) 23–28.

[5] C. Collaboration, *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, Physical Review Letters **74** (Apr., 1995) 2626–2631, hep-ex/9503002.

[6] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (Sept., 2012) 1–29, arXiv:1207.7214 [hep-ex].

[7] C. Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (Sept., 2012) 30–61, arXiv:1207.7235 [hep-ex].

[8] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984.

[9] D. Griffith, *Introduction to Elementary Particles*. Wiley-VCH, 2. ed., 2008.

[10] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[11] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Phys.Rev.Lett. **81** (1998) 1562, arXiv:hep-ex/9807003 [hep-ex].

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

[13] A. Salam and J. Ward, *Electromagnetic and weak interactions*, Phys.Lett. **13** (1964) .

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

[15] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

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

[17] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys.Rev.Lett. **13** (1964) 321.

[18] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys.Rev.Lett. **13** (1964) 585.

[19] ALEPH, CDF, D0, DELPHI, L3, OPAL and SLD Collaborations and LEP, Tevatron and SLD Electroweak Working Groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, `arXiv:1012.2367 [hep-ex]`, CERN-PH-EP-2010-095.

[20] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189, `arXiv:0901.0002 [hep-ph]`.

[21] Y. L. Dokshitzer, *Calculation of structure functions of deep-inelastic scattering and e+e- annihilation by perturbation theory in quantum chromodynamics*, Sov.Phys. JETP **46** (1977) 641.

[22] V. Gribov and L. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov.J.Nucl.Phys **15** (1972) 438.

[23] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. **B126** (1977) 298.

[24] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard interactions of quarks and gluons: a primer for LHC physics*, Reports on Progress in Physics **70** (Jan., 2007) 89–193, `hep-ph/0611148`.

[25] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization of Hard Processes in QCD*, ArXiv High Energy Physics - Phenomenology e-prints (Sept., 2004) , `hep-ph/0409313`.

[26] W. J. Stirling, *Private communication*.

[27] B. R. Webber, *A QCD model for jet fragmentation including soft gluon interference*, Nucl. Phys. B **238** (1984) 492.

[28] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys.Rept. **97** (1983) 31.

[29] ATLAS Collaboration, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`.

[30] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), *Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables*, CERN-2011-002 (CERN, Geneva, 2011) , `arXiv:1101.0593 [hep-ph]`.

[31] LHC Higgs Cross Section Working Group, `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSections`.

[32] J. Ellis, G. Ridolfi, and F. Zwirner, *Higgs boson properties in the Standard Model and its supersymmetric extensions*, Comptes Rendus Physique **8** (Nov., 2007) 999–1012, `hep-ph/0702114`.

[33] CDF and D0 Collaborations, *Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb-1 of Data*, ArXiv e-prints (July, 2011) , `arXiv:1107.5518 [hep-ex]`.

[34] CDF and D0 Collaborations, *Higgs Boson Studies at the Tevatron*, ArXiv e-prints (Mar., 2013) , `arXiv:1303.6346 [hep-ex]`.

[35] ATLAS Collaboration, *Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. **D90** no. 11, (2014) 112015, `arXiv:1408.7084 [hep-ex]`.

[36] ATLAS Collaboration, *Measurements of Higgs boson production and couplings in the four-lepton channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector*, Phys. Rev. **D91** no. 1, (2015) 012006, `arXiv:1408.5191 [hep-ex]`.

[37] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, ArXiv e-prints (Mar., 2015) , `arXiv:1503.07589 [hep-ex]`.

[38] ATLAS Collaboration, *Measurement of the Higgs boson mass from the $H \to \gamma\gamma$ and $H \to ZZ^* \to 4\ell$ channels with the ATLAS detector using 25 fb^{-1} of pp collision data*, ArXiv e-prints (June, 2014) , `arXiv:1406.3827 [hep-ex]`.

[39] ATLAS Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, ArXiv e-prints (June, 2015) , `arXiv:1506.05669 [hep-ex]`.

[40] CMS Collaboration, *Constraints on Anomalous HWW Interactions using Higgs boson decays to W+W- in the fully leptonic final state* Tech. Rep. CMS-PAS-HIG-14-012, CERN, Geneva, 2014. `http://cds.cern.ch/record/1728250`.

[41] CMS Collaboration, *Constraints on anomalous HVV interactions using H to 4l decays* Tech. Rep. CMS-PAS-HIG-14-014, CERN, Geneva, 2014. `http://cds.cern.ch/record/1728251`.

[42] ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, ArXiv e-prints (June, 2016) , arXiv:1606.02266 [hep-ex].

[43] The LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, ArXiv e-prints (July, 2013) , arXiv:1307.1347 [hep-ph].

[44] A. Djouadi, P. Janot, J. Kalinowski, and P. M. Zerwas, *SUSY decays of Higgs particles*, Physics Letters B **376** (Feb., 1996) 220–226, hep-ph/9603368.

[45] S. P. Martin and J. D. Wells, *Motivation and detectability of an invisibly decaying Higgs boson at the Fermilab Tevatron*, Phys. Rev. **D60** no. 3, (Aug., 1999) 035006, hep-ph/9903259.

[46] C. Englert, T. Plehn, D. Zerwas, and P. M. Zerwas, *Exploring the Higgs portal*, Physics Letters B **703** (Sept., 2011) 298–305, arXiv:1106.3097 [hep-ph].

[47] A. Djouadi, O. Lebedev, Y. Mambrini, and J. Quevillon, *Implications of LHC searches for Higgs-portal dark matter*, Physics Letters B **709** (Mar., 2012) 65–69, arXiv:1112.3299 [hep-ph].

[48] ATLAS Collaboration, *Constraints on new phenomena via Higgs boson couplings and invisible decays with the ATLAS detector*, arXiv:1509.00672 [hep-ex].

[49] F. Sannino and K. Tuominen, *Orientifold theory dynamics and symmetry breaking*, Phys. Rev. **D71** no. 5, (Mar., 2005) 051901, hep-ph/0405209.

[50] R. Foadi, M. T. Frandsen, T. A. Ryttov, and F. Sannino, *Minimal walking technicolor: Setup for collider physics*, Phys. Rev. **D76** no. 5, (Sept., 2007) 055005, arXiv:0706.1696 [hep-ph].

[51] A. Belyaev, R. Foadi, M. T. Frandsen, M. Järvinen, F. Sannino, and A. Pukhov, *Technicolor walks at the LHC*, Phys. Rev. **D79** no. 3, (Feb., 2009) 035006, arXiv:0809.0793 [hep-ph].

[52] M. Schmaltz and D. Tucker-Smith, *Little Higgs Theories*, Annual Review of Nuclear and Particle Science **55** (Dec., 2005) 229–270, hep-ph/0502182.

[53] M. J. Dugan, H. Georgi, and D. B. Kaplan, *Anatomy of a Composite Higgs Model*, Nucl. Phys. **B254** (1985) 299–326.

[54] K. Agashe, R. Contino, and A. Pomarol, *The minimal composite Higgs model*, Nuclear Physics B **719** (July, 2005) 165–187, hep-ph/0412089.

[55] D. Pappadopulo, A. Thamm, R. Torre, and A. Wulzer, *Heavy vector triplets: bridging theory and data*, Journal of High Energy Physics **9** (Sept., 2014) 60, arXiv:1402.4431 [hep-ph].

[56] V. D. Barger, W.-Y. Keung, and E. Ma, *A Gauge Model With Light W and Z Bosons*, Phys. Rev. **D22** (1980) 727.

[57] R. Contino, D. Marzocca, D. Pappadopulo, and R. Rattazzi, *On the effect of resonances in composite Higgs phenomenology*, JHEP **10** (2011) 081, arXiv:1109.1570 [hep-ph].

[58] ATLAS Collaboration, *Combination of searches for WW, WZ, and ZZ resonances in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, Physics Letters B **755** (Apr., 2016) 285–305, arXiv:1512.05099.

[59] ATLAS Collaboration, *Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at $\sqrt{s}$=8 TeV with the ATLAS detector*, Journal of High Energy Physics **12** (Dec., 2015) 55, arXiv:1506.00962 [hep-ex].

[60] CMS Collaboration, *Search for massive WH resonances decaying into the l nu b anti-b final state at $\sqrt{(s)} = 8$ TeV*, arXiv:1601.06431.

[61] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[62] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[63] A. La Rosa and Atlas Ibl Collaboration, *ATLAS IBL Pixel Upgrade*, Nuclear Physics B Proceedings Supplements **215** (June, 2011) 147–150, arXiv:1104.1980 [physics.ins-det].

[64] ATLAS Collaboration, *Overview of the ATLAS Insertable B-Layer (IBL) Project*, PoS **RD13** (2013) 012.

[65] B. Dolgoshein, *Transition radiation detectors*, Nuclear Instruments and Methods in Physics Research **A326** (1993) 434–469.

[66] ATLAS Collaboration, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)* Tech. Rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar, 2007. https://cds.cern.ch/record/1020106.

[67] ATLAS Collaboration, *Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment* Tech. Rep. ATLAS-CONF-2010-069, CERN, Geneva, Jul, 2010. https://cds.cern.ch/record/1281344.

[68] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444.

[69] R. Frühwirth, W. Waltenberger, and P. Vanlaer, *Adaptive vertex fitting*, J. Phys. **G34** (2007) N343.

[70] ATLAS Collaboration, *Impact Parameter Resolution*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2015-007/.

[71] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, European Physical Journal C **74** (Oct., 2014) 3071, arXiv:1407.5063 [hep-ex].

[72] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, ArXiv e-prints (Apr., 2014) , arXiv:1404.2240 [hep-ex].

[73] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data* Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014. https://cds.cern.ch/record/1706245.

[74] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data*, European Physical Journal C **74** (July, 2014) 3130, arXiv:1407.3935 [hep-ex].

[75] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, Journal of High Energy Physics **4** (Apr., 2008) 063, arXiv:0802.1189 [hep-ph].

[76] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS* Tech. Rep. ATLAS-CONF-2015-037, CERN, Geneva, Aug, 2015. https://cds.cern.ch/record/2044941.

[77] ATLAS Collaboration, M. Beckingham, M. Duehrssen, E. Schmidt, M. Shapiro, M. Venturi, J. Virzi, I. Vivarelli, M. Werner, S. Yamamoto, and T. Yamanaka, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim* Tech. Rep. ATL-PHYS-PUB-2010-013, CERN, Geneva, Oct, 2010. https://cds.cern.ch/record/1300517.

[78] ATLAS Collaboration, *Commissioning of high performance b-tagging algorithms with the ATLAS detector*, ATLAS-CONF-2011-102, July, 2011 .

[79] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, PERF-2012-04, 2012 .

[80] ATLAS, ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** no. 04, (2016) P04008, arXiv:1512.01094 [hep-ex].

[81] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2* Tech. Rep. ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul, 2015. http://cds.cern.ch/record/2037697.

[82] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV* Tech. Rep. ATLAS-CONF-2013-082, CERN, Geneva, Aug, 2013. https://cds.cern.ch/record/1570993.

[83] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a b-quark pair with the ATLAS detector*, Physics Letters B **718** (Dec., 2012) 369–390, arXiv:1207.0210 [hep-ex].

[84] ATLAS Collaboration, *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s}$ = 7 and 8 TeV in the ATLAS experiment*, European Physical Journal C **76** (Jan., 2016) 6, arXiv:1507.04548 [hep-ex].

[85] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J **C 70** (2010) 823, arXiv:1005.4568 [physics.ins-det].

[86] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A 506** (2003) 250–303.

[87] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[88] J. Pumplin *et al.*, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:hep-ph/0201195 [hep-ph].

[89] ATLAS Collaboration, *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11* Tech. Rep. ATL-PHYS-PUB-2011-009, CERN, Geneva, Jul, 2011. http://cds.cern.ch/record/1363300.

[90] ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data* Tech. Rep. ATL-PHYS-PUB-2011-008, CERN, Geneva, Apr, 2011. http://cds.cern.ch/record/1345343.

[91] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C 45** (2006) 97–107, arXiv:hep-ph/0506026 [hep-ph].

[92] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv:hep-ph/0409146 [hep-ph].

[93] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[94] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv:1002.2581 [hep-ph].

[95] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, *$HW^{\pm}$/HZ + 0 and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO*, JHEP **10** (2013) 083, arXiv:1306.2542 [hep-ph].

[96] H.-L. Lai *et al.*, *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph].

[97] C. Englert, M. McCullough, and M. Spannowsky, *Gluon-initiated associated production boosts Higgs physics*, Phys. Rev. **D89** no. 1, (2014) 013013, arXiv:1310.4828 [hep-ph].

[98] T. Gleisberg *et al.*, *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv:0811.4622 [hep-ph].

[99] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv:hep-ph/0603175.

[100] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5*, arXiv:hep-ph/0405247 [hep-ph].

[101] P. Nason and G. Zanderighi, W^+W^- , WZ *and* ZZ *production in the POWHEG-BOX-V2*, Eur. Phys. J. **C 74** (2014) 2702, arXiv:1311.1365 [hep-ph].

[102] ATLAS Collaboration, *Search for the Standard Model Higgs boson in produced in association with a vector boson and decaying to bottom quarks with the ATLAS detector* Tech. Rep. ATLAS-CONF-2012-161, CERN, Geneva, Nov, 2012. https://cds.cern.ch/record/1493625.

[103] ATLAS Collaboration, *The ATLAS transverse-momentum trigger performance at the LHC in 2011* Tech. Rep. ATLAS-CONF-2014-002, CERN, Geneva, Feb, 2014. https://cds.cern.ch/record/1647616.

[104] D. Büscher, *Search for the Higgs Boson in the Channel* $WH \to evbb$ *using Boosted Decision Trees in Proton-Proton Collisions at* $\sqrt(s) = 7\,TeV$ *with the ATLAS Experiment*, Diploma thesis http://portal.uni-freiburg.de/jakobs/dateien/abschluss/diplombuescher, 2012.

[105] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[106] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, European Physical Journal C **71** (Feb., 2011) 1554, arXiv:1007.1727 [physics.data-an].

[107] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, ArXiv Physics e-prints (June, 2003) , physics/0306116.

[108] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke, *The RooStats project* in *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research.*, p. 57. 2010. arXiv:1009.1003 [physics.data-an].

[109] N. Metropolis and S. Ulam, *The Monte Carlo Method*, Journal of the American Statistical Association **44** no. 247, (1949) 335–341.

[110] A. L. Read, *Modified frequentist analysis of search results (the CL_s method).*

[111] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693–2704.

[112] ROOT Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats* Tech. Rep. CERN-OPEN-2012-016, New York U., New York, Jan, 2012. `https://cds.cern.ch/record/1456844`.

[113] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: A program for Higgs boson decays in the Standard Model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56–74.

[114] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys.Lett. **B579** (2004) 149–156, `arXiv:hep-ph/0307206 [hep-ph]`.

[115] M. Ciccolini, S. Dittmaier, and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys.Rev. **D68** (2003) 073003, `arXiv:hep-ph/0306234 [hep-ph]`.

[116] ATLAS Collaboration, *Measurement of top-quark pair differential cross-sections in the l+jets channel in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector* Tech. Rep. ATLAS-CONF-2013-099, CERN, Geneva, Sep, 2013. `https://cds.cern.ch/record/1600778`.

[117] I. W. Stewart and F. J. Tackmann, *Theory uncertainties for Higgs mass and other searches using jet bins*, Phys. Rev. D **85** (Feb, 2012) 034011.

[118] S. Gangal and F. J. Tackmann, *Next-to-leading-order uncertainties in Higgs + jets from gluon fusion*, Phys. Rev. D **87** (May, 2013) 093008.

[119] F. James and M. Roos, *Minuit: Function Minimization and Error Analysis*, Comput. Phys. Commun. **10** (Feb, 1975) 343–367.

[120] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81–86.

[121] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a b-quark pair with the ATLAS detector*, Phys. Lett. **B718** (2012) 369–390, `arXiv:1207.0210 [hep-ex]`.

[122] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into in collisions at with the ATLAS detector*, European Physical Journal C **75** (July, 2015) 349, `arXiv:1503.05066 [hep-ex]`.

[123] ATLAS Collaboration, *Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{s}$=8 TeV with the ATLAS detector*, ArXiv e-prints (Apr., 2016) , `arXiv:1604.03812 [hep-ex]`.

[124] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced by vector-boson fusion in 8 TeV pp collisions and decaying to bottom quarks with the ATLAS detector*, ArXiv e-prints (June, 2016) , arXiv:1606.02181 [hep-ex].

[125] ATLAS Collaboration, *Measurement of the cross section of high transverse momentum production in protonproton collisions at with the ATLAS detector*, Physics Letters B **738** (2014) 25 – 43.

[126] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better jet clustering algorithms*, Journal of High Energy Physics **8** (Aug., 1997) 001, hep-ph/9707323.

[127] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, *et al.*, *Herwig++ Physics and Manual*, Eur. Phys. J. C **58** (2008) 639, arXiv:0803.0883 [hep-ph].

[128] J. Pumplin, D. R. Stump, J. Huston, H.-L. Lai, P. Nadolsky, and W.-K. Tung, *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, Journal of High Energy Physics **7** (July, 2002) 012, hep-ph/0201195.

[129] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes* Tech. Rep. ATL-PHYS-PUB-2012-003, CERN, Geneva, Aug, 2012. https://cds.cern.ch/record/1474107.

[130] J. M. Campbell, R. K. Ellis, and G. Zanderighi, *Next-to-leading order Higgs + 2 jet production via gluon fusion*, Journal of High Energy Physics **10** (Oct., 2006) 028, hep-ph/0608194.

[131] ATLAS Collaboration, *Search for Invisible Decays of a Higgs Boson Produced in Association with a Z Boson in ATLAS*, Physical Review Letters **112** no. 20, (May, 2014) 201802, arXiv:1402.3244 [hep-ex].

[132] D. Büscher, *Searches for diboson resonances with the ATLAS detector at the LHC*, Proceedings of XXIV International Workshop on Deep-Inelastic Scattering and Related Subjects , PoS(DIS2016)101.

[133] ATLAS Collaboration, *Implementation of the ATLAS Run 2 event data model* Tech. Rep. ATL-SOFT-PROC-2015-003. 7, CERN, Geneva, May, 2015. https://cds.cern.ch/record/2014150.

[134] P. Calafiura, W. Lavrijsen, C. Leggett, M. Marino, and D. Quarrie, *The Athena Control Framework in Production, New Developments and Lessons Learned*.

[135] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, Journal of High Energy Physics **7** (July, 2014) 79, arXiv:1405.0301 [hep-ph].

[136] ATLAS Collaboration, *ATLAS Run 1 Pythia8 tunes* Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov, 2014. http://cds.cern.ch/record/1966419.

[137] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, Journal of High Energy Physics **12** (Dec., 2008) 039, arXiv:0808.3674 [hep-ph].

[138] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Physical Review Letters **108** no. 11, (Mar., 2012) 111601, arXiv:1111.5206 [hep-ph].

[139] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, Journal of High Energy Physics **4** (Apr., 2013) 27, arXiv:1207.5030 [hep-ph].

[140] P. Nason, *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*, Journal of High Energy Physics **11** (Nov., 2004) 040, hep-ph/0409146.

[141] J. Huston *et al.*, *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **07** (2002) 012, arXiv:0201195 [hep-ph].

[142] P. Z. Skands, *Tuning Monte Carlo generators: The Perugia tunes*, Phys. Rev. **D82** no. 7, (Oct., 2010) 074018, arXiv:1005.3457 [hep-ph].

[143] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, ArXiv e-prints (June, 2014) , arXiv:1406.3660 [hep-ex].

[144] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$=13 TeV*, ArXiv e-prints (Mar., 2016) , arXiv:1603.05598 [hep-ex].

[145] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, ArXiv e-prints (Mar., 2016) , arXiv:1603.02934 [hep-ex].

[146] D. Krohn, J. Thaler, and L.-T. Wang, *Jet trimming*, Journal of High Energy Physics **2** (Feb., 2010) 84, arXiv:0912.1342 [hep-ph].

[147] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (Oct., 1993) 3160–3166, hep-ph/9305266.

[148] ATLAS Collaboration, *Identification of boosted, hadronically-decaying W and Z bosons in $\sqrt{s}$ = 13 TeV Monte Carlo Simulations for ATLAS* Tech. Rep. ATL-PHYS-PUB-2015-033, CERN, Geneva, Aug, 2015. https://cds.cern.ch/record/2041461.

[149] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s}$ = 7 TeV*, European Physical Journal C **73** (Mar., 2013) 2304, arXiv:1112.6426 [hep-ex].

[150] ATLAS Collaboration, *Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s}$ = 8 TeV*, ArXiv e-prints (Oct., 2015) , arXiv:1510.05821 [hep-ex].

[151] A. J. Larkoski, I. Moult, and D. Neill, *Power counting to better jet observables*, Journal of High Energy Physics **12** (Dec., 2014) 9, arXiv:1409.6298 [hep-ph].

[152] A. J. Larkoski, I. Moult, and D. Neill, *Analytic Boosted Boson Discrimination*, arXiv:1507.03018 [hep-ph].

[153] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, Journal of Instrumentation **11** (Apr., 2016) P04008, arXiv:1512.01094 [hep-ex].

[154] ATLAS Collaboration, *Search for a new resonance decaying to a W or Z boson and a Higgs boson in the final states with the ATLAS detector*, European Physical Journal C **75** (June, 2015) 263, arXiv:1503.08089 [hep-ex].

[155] ATLAS Collaboration, *Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in* $\sqrt{s}$ *= 8 TeV proton-proton collisions using the ATLAS detector*, ArXiv e-prints (Oct., 2015) , arXiv:1510.03818 [hep-ex].

[156] ATLAS Collaboration, *Measurements of the production cross sections in association with jets with the ATLAS detector*, European Physical Journal C **75** (Feb., 2015) 82, arXiv:1409.8639 [hep-ex].

[157] J. M. Campbell and R. K. Ellis, *MCFM for the Tevatron and the LHC*, Nuclear Physics B Proceedings Supplements **205** (Aug., 2010) 10–15, arXiv:1007.3492 [hep-ph].

[158] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, *HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, Journal of High Energy Physics **1** (Jan., 2001) 010, hep-ph/0011363.

[159] CMS Collaboration, *Search for a massive resonance decaying into a Higgs boson and a W or Z boson in hadronic final states in proton-proton collisions at sqrt(s) = 8 TeV*, ArXiv e-prints (June, 2015) , arXiv:1506.01443 [hep-ex].

[160] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with a vector boson and decaying to a* $b\bar{b}$ *pair in pp collisions at 13 TeV using the ATLAS detector*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2016-091/, Aug., 2016.

Acknowledgements

BIG thanks (elephant-sized) go to Karl Jakobs for his support and for the leadership of his Freiburg group. This thesis would not have been possible without his dedication for the scientific cause. Also I would like to thank Christian "Super" Weiser for the many fruitful discussions we had and for his numerous comments on this thesis.

During my PhD I have gained many insights on physics and on statistical methods from the ATLAS $H \to b\bar{b}$ group. In particular, I would like to thank the conveners during my time in the group: Heather Gray, Andrew Mehta and Giacinto Piacquadio (in particular for the motivation to pursue the binning studies). Also the discussions with Nicolas Morange, Gabriel Facini, Remi Zaidan and Garabed Halladjian have helped me tremendously. Thanks!

I would like to thank Georges Aad. This is for his tutoring during my diploma thesis. I have missed to write acknowledgements back then. Georges does not like German Bandwurmsätze. He likes short sentences.

From the group in Freiburg I would further like to thank Felix Bührer for making his WWEventSelector available, which allowed me to implement the $gg \to H \to b\bar{b}$ study with ease. Also thanks to Philip "der Normale" Sommer for all the amusement (and some useful discussions as well!).

Ich möchte meiner Familie danken, insbesondere Reinhold, welcher in mir schon früh das Interesse an der Mathematik und den Naturwissenschaften geweckt hat. Auch möchte ich Ilse danken für ihre allzeit gutgelaunte Motivation! Ich danke Ira und Andreas für ihre Unterstützung während meines Studiums. Zu guter Letzt möchte ich Alena danken. Nur durch Dich konnte ich meine Zeit in Freiburg wirklich genießen!