

# **Probing new physics with boosted $H \rightarrow b\bar{b}$ decays with the *ATLAS* detector at 13 TeV**

Dissertation  
zur  
Erlangung des Doktorgrades (Dr. rer. nat.)  
der  
Mathematisch-Naturwissenschaftlichen Fakultät  
der  
Rheinischen Friedrich-Wilhelms-Universität Bonn

von  
**Ruth Magdalena Jacobs**  
aus  
Bonn

Bonn, Juli 2019

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

1. Gutachter: Prof. Dr. Norbert Wermes  
2. Gutachter: Prof. Dr. Ian Brock

Tag der Promotion: 08.10.2019  
Erscheinungsjahr: 2020

Nicht alle Dinge lassen  
sich einfach in Worte fassen.  
Aber sie stehen bisweilen  
unsichtbar zwischen den Zeilen.  
– *Frantz Wittkamp*



# Acknowledgements

---

I would like to thank Norbert Wermes for the opportunity to work on my PhD in his research group and for his continued support from the very beginning of my studies in physics. My thanks go to Peter Jenni, Fabiola Gianotti and the CERN & Society foundation for their generous funding through the ATLAS PhD grant. Furthermore I would like to thank the Studienstiftung des deutschen Volkes for their funding during the final year of my PhD.

I would like to thank Tatjana Lenz for her daily supervision and advice regarding all ATLAS matters. Also I would like to thank Eckhard von Törne for a critical reading of my thesis. My thanks go to all members and former members of the Bonn ATLAS group, who helped me with countless chats, rounds of comments, tips and tricks and coffees.

A special thank you goes to my parents for their moral support and advice, for many cups of tea, hugs, and everything that is helping me to keep my feet attached to firm ground and walking forwards. To my "support team" at Hagerhof, especially Anneka and Eva: thank you for taking such good care of Soli while I was away somewhere yet again. Finally I would like to thank Artem for bearing with me and for reminding me from time to time that philosophy, art and the world outside are interesting, too.

*Mein besonderer Dank gilt Ingrid Kromik – für Soli, der mein allerbestes Ausgleich zur Physik ist, für Deine Großzügigkeit und Deine langjährige Unterstützung.*



# Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Basic Concepts of High-Energy Physics</b>	<b>5</b>
2.1	The Standard Model of particle physics	5
2.2	Symmetries and Gauge invariance	6
2.3	The Higgs mechanism	7
2.4	Production and decay of the Standard Model Higgs boson	9
2.5	The Higgs boson as a probe for new physics	9
<b>3</b>	<b>Experimental Physics at the Large Hadron Collider</b>	<b>13</b>
3.1	The Large Hadron Collider	13
3.1.1	Luminosity	14
3.1.2	Proton-proton collisions	14
3.2	The <i>ATLAS</i> detector	16
3.2.1	Coordinates and reconstructed quantities	16
3.2.2	Detector components	17
3.2.3	Trigger	20
3.2.4	Monte Carlo simulation	21
3.3	Tracks	22
3.4	Jets	23
3.4.1	The anti- $k_T$ jet algorithm	24
3.4.2	Large- $R$ jets	25
3.4.3	Large- $R$ jet mass	26
3.4.4	Jet substructure	27
3.4.5	Ghost association	29
3.5	Flavour tagging	29
3.5.1	Properties of $b$ -jets	29
3.5.2	Flavour tagging procedure in <i>ATLAS</i>	30
3.5.3	Impact-parameter based $b$ -tagging	32
3.5.4	Secondary-vertex based $b$ -tagging	33
3.5.5	Flavour tagging algorithms in <i>ATLAS</i>	34
3.5.6	Jet truth labelling	35
3.5.7	Performance of flavour tagging	35
3.5.8	Track jet $b$ -tagging	37
3.5.9	Flavour tagging calibration	37
3.6	Higgs boson tagging	38
3.6.1	Properties of boosted $H \rightarrow b\bar{b}$ decays	38
3.6.2	Standard Higgs tagging	39

3.6.3	Higgs tagging performance . . . . .	40
3.6.4	Advanced Higgs tagging . . . . .	41
3.7	Statistical methods . . . . .	44
3.7.1	Basic concepts . . . . .	44
3.7.2	Parameter estimation: The Maximum Likelihood method . . . . .	45
3.7.3	Binned Maximum Likelihood fits . . . . .	47
3.7.4	The Profile Likelihood Ratio . . . . .	47
3.7.5	Hypothesis tests . . . . .	48
3.7.6	Systematic uncertainties: Nuisance parameters . . . . .	50
3.7.7	Approximation of the Test Statistics distribution: Asymptotic formulae . . . . .	51
3.7.8	Ensemble tests . . . . .	52
<b>4</b>	<b>Search for a new resonance <math>V' \rightarrow VH</math> decaying into final states with leptons and b-jets</b>	<b>53</b>
4.1	Introduction to the Heavy Vector Triplet model . . . . .	53
4.2	Object selection . . . . .	57
4.3	Event categorization . . . . .	60
4.4	Background processes . . . . .	62
4.5	Event selection . . . . .	64
4.5.1	Dataset and Trigger strategy . . . . .	64
4.5.2	Resolved selection . . . . .	65
4.5.3	Merged selection . . . . .	67
4.6	Reconstruction of the $VH$ invariant mass . . . . .	69
4.7	Systematic uncertainties . . . . .	70
4.8	Statistical analysis . . . . .	70
4.8.1	Fit inputs, regions and binning . . . . .	72
4.9	Fit diagnostics . . . . .	75
4.9.1	Nuisance parameter pull distributions . . . . .	76
4.9.2	Nuisance parameter ranking . . . . .	78
4.9.3	Nuisance parameter correlations . . . . .	80
4.9.4	Breakdown of statistical and systematic uncertainties . . . . .	80
4.9.5	Validation regions . . . . .	82
4.10	Combined analysis of 0-, 1- and 2-lepton channels . . . . .	82
4.11	Results . . . . .	87
4.12	Interpretation within the HVT simplified model . . . . .	90
4.13	Conclusions and prospects . . . . .	91
<b>5</b>	<b>Double-<math>b</math>-tagging calibration in <math>g \rightarrow b\bar{b}</math> events</b>	<b>93</b>
5.1	Calibration strategy . . . . .	93
5.2	Object selection . . . . .	95
5.3	Event selection . . . . .	96
5.4	Flavour fraction fit . . . . .	98
5.5	Validation of Higgs tagging variables . . . . .	102
5.5.1	Systematic uncertainties in validation study . . . . .	103
5.5.2	Jet kinematics and $b$ -tagging results . . . . .	105
5.5.3	Jet mass and substructure results . . . . .	107
5.6	Scale factor calculation . . . . .	109
5.6.1	Systematic uncertainties in double- $b$ -tagging calibration . . . . .	110

5.7	Extrapolation from semi-leptonic to inclusive $b$ -jet sample	113
5.8	Calibration Results	114
5.9	Conclusions and outlook	117
<b>6</b>	<b>Conclusion</b>	<b>119</b>
<b>A</b>	<b>Additional information</b>	<b>121</b>
A.1	$VH$ resonance additional post-fit spectra	121
A.2	$g \rightarrow b\bar{b}$ calibration additional studies and plots	122
A.2.1	Studies of alternative template variables	122
A.2.2	Template similarity test	124
A.2.3	Template correlation test	124
A.2.4	Small- $R$ jet trigger threshold studies	124
A.2.5	Additional studies for extrapolation from semi-leptonic to inclusive $b$ -jet sample	126
A.2.6	$\langle s_{d_0} \rangle$ post-fit plots	128
	<b>Bibliography</b>	<b>133</b>



# Introduction

The current model of the evolution of the universe postulates the emergence of space and time in a singularity called the *Big Bang* estimated to have taken place 13.7 billion years ago. Following the Big Bang, the universe expanded and the energy density decreased over time. A timeline of the evolution of the universe is shown in figure 1.1.

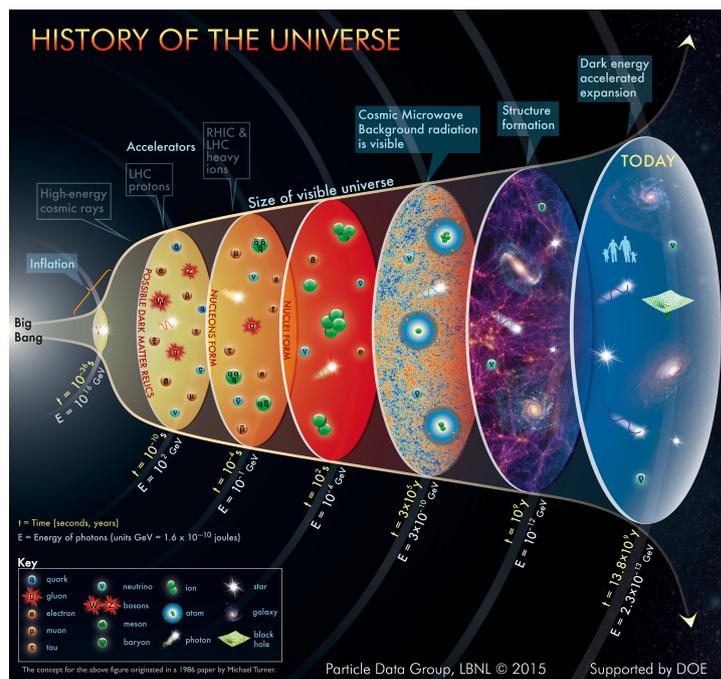


Figure 1.1: Schematic view of the evolution of the universe in the Big Bang model. From [1]

After a period of rapid expansion (*inflation*), the so called *elementary particle era* spans from a time of  $t = 10^{-10}$  s after the Big Bang, at an energy scale of around 100 GeV, until the formation of the first nucleons at  $t = 10^{-4}$  s, corresponding to an energy scale of 0.1 GeV. The physics studied in this thesis takes place at the high-energy frontier of the particle era. To infer information about physics processes in this regime, data from the Large Hadron Collider (LHC), one of the largest particle accelerators in the world today, are analysed. At the LHC, protons are brought to collision at a centre-of-mass energy of 13 TeV, recreating in a laboratory the conditions that prevailed at the very beginning of the particle era shortly after the Big Bang.

The dynamics and interactions of elementary particles are described by the *Standard Model of Particle Physics (SM)*. It is one of the most well-tested theories in physics at present. However, there are known inconsistencies within the SM, which makes physicists believe that it is an incomplete theory. Observations from astronomy suggest that only 4.6% of the matter content of our universe consists of so-called baryonic matter, whose particle composition is described by the SM, while 23% are made of a type of matter which only interacts via gravity, also known as *dark matter*. The remaining 72% are attributed to so-called *dark energy* which is deemed responsible for the observed accelerated expansion of the universe and whose composition is known to an even lesser degree than dark matter. The SM provides no particle candidate for dark matter and also gives no compelling explanation of the nature of dark energy.

With the discovery of the Higgs boson in the year 2012 by the *ATLAS* and *CMS* collaborations at the LHC, one of the last missing pieces of the SM fell into place, by observing the mechanism responsible for giving mass to the known particles. Even though the properties of the observed Higgs boson seem to be consistent with the SM expectation, there are still many open questions connected to it. One of these problems, called the *naturalness problem*, is related to the great scale difference between the SM electroweak scale (250 GeV) and the Planck scale ( $10^{19}$  GeV), where gravity becomes strong. This scale difference leads to extremely large quantum corrections to the Higgs boson mass. To obtain the observed Higgs boson mass of 125 GeV the SM parameters need to be tuned exactly to cancel out the corrections. Such fine-tuning is considered to be un-natural and undesirable in a theory.

Due to these conceptual problems it is believed that the SM cannot be the final answer regarding the description of the universe in the framework of particle physics. There are many proposed extensions of the SM (*Beyond the SM (BSM)* or *New Physics theories*) which attempt to overcome some of the difficulties mentioned above, e.g. by providing a natural candidate for a dark matter particle, or introducing a new interaction at the TeV energy scale which leads to natural cancellation of the large terms correcting the Higgs boson mass.

One of the aims of the LHC programme is to search for evidence of New Physics by producing new particles directly. The *ATLAS* experiment is designed as a general-purpose detector such that it is sensitive to all SM particle types and able to record many different signatures of new particle decays. Many new particles that could be produced at the LHC are predicted to have masses up to several TeV.

In this thesis a search for a new heavy resonance  $V'$  in the decay channel  $V' \rightarrow VH$  with the *ATLAS* detector is presented.  $V$  denotes a SM electroweak boson ( $W$  or  $Z$  boson) and  $H$  is a SM Higgs boson. The  $W$  and  $Z$  bosons are required to decay leptonically, and the Higgs boson decays further into two  $b$ -quarks. Data recorded by the *ATLAS* detector during 13 TeV LHC proton-proton collisions are analyzed and scrutinized for the presence of a  $V'$ . The main challenge in this analysis is the reconstruction of the Higgs boson, which obtains a large Lorentz-boost due to the large resonance mass. As a result, the  $b$ -quarks from the  $H \rightarrow b\bar{b}$  decay are collimated and require dedicated reconstruction techniques. The second part of this thesis is dedicated to the study of boosted Higgs boson identification algorithms and their calibration in data. The data-based calibration was newly developed for the *ATLAS* experiment within this thesis work. It is using gluon splitting ( $g \rightarrow b\bar{b}$ ) events, which are produced abundantly at the LHC, as a proxy for a boosted  $bb$  event topology.

This thesis is structured as follows: Chapter 2 gives an introduction to the SM and the Higgs mechanism and introduces different classes of New Physics models. In chapter 3, an overview of the Large Hadron

---

Collider and the *ATLAS* detector is given, as well as an introduction to reconstruction techniques for different particle signatures in the detector. The search for a new heavy resonance  $V' \rightarrow VH(b\bar{b})$  is presented in chapter 4, introducing the effective model used as New Physics benchmark and outlining in detail the analysis strategy and design. Chapter 5 contains the study of the performance of Higgs boson identification algorithms in data using  $g \rightarrow b\bar{b}$  events as well as the first direct calibration of double  $b$ -tagging for the *ATLAS* experiment. Finally, the conclusion is presented in chapter 6.



---

## Basic Concepts of High-Energy Physics

---

In this chapter an introduction to the physics of elementary particles will be given. The basic principles of the current theoretical framework, called the Standard Model of particle physics, will be introduced, as well as the generation of particle masses via the Higgs mechanism. Finally the role of the Higgs boson in connection with new physics extensions of the Standard Model will be outlined.

### 2.1 The Standard Model of particle physics

The theory of elementary particle physics attempts to unify quantum mechanics with the theory of Special Relativity. Quantum mechanics first established the duality between discrete particles and continuous waves at small scales. When contemplating the Lorentz-invariant formulation of the equations of quantum mechanics, a curious dynamics of creation and annihilation of quanta from the continuum emerged which finally led to the new concept of a quantum field. Particles manifest themselves as quantum excitations of fields and interact via force fields, which in turn also possess particle nature via quantum excitations.

The current theoretical framework to describe elementary particles and their interactions is called the Standard Model of particle physics (SM). Particles are classified according to their behavior under Lorentz transformation. While all matter particles transform as Dirac spinor with half-integer spin (*fermions*), the force carriers transform like vectors and possess integer spin (*bosons*). The exception to these cases is the Higgs boson, which transforms as a Lorentz-scalar and will be discussed in more detail in section 2.3.

Figure 2.1 shows a graphical depiction of the particle content of the SM the corresponding forces to describe particle interactions. Matter particles are subdivided into quarks, which have fractional electric charge of  $+2/3$  (*up-type*) and  $-1/3$  (*down-type*) as well as leptons with an electric charge of  $-1$  and neutrinos, which are electrically neutral. Quarks and leptons can be grouped into three generations which differ only in their mass. In total there are six different types of quarks (*flavours*) with three quark doublets consisting of one up-type and one down-type quark. Of special interest for this thesis is the *b*-quark, which is part of the third quark generation, with a mass of  $4.18^{+0.02}_{-0.03}$  GeV (in  $\overline{MS}$  scheme) [3]. The leptons follow a similar structure as the quarks, with three generations of doublets each containing a charged lepton and the corresponding neutrino. For each matter particle in the SM there is a corresponding anti-matter particle with the same mass but opposite charges with respect to the interactions.

The three interactions of the SM include *strong*, *electromagnetic* and *weak* interaction. Each interaction is connected to a conserved current and a charge, where only particles that possess a nonzero charge

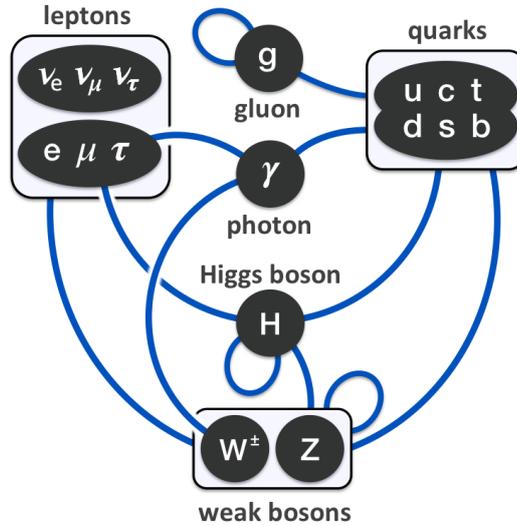


Figure 2.1: Particles and their interactions within the Standard Model of Particle Physics. The blue connecting lines represent the interactions between different SM particles. From:[2]

participate in the corresponding interaction. The strong interaction (*Quantum Chromodynamics, QCD*) is mediated by the *gluon*, which is a massless spin-1 vector boson. It couples to particles which carry *colour charge*, namely quarks and gluons themselves. Colour charge assumes three different values, symbolically denoted as *red, blue* and *green* and three anti-colours, *anti-red, anti-blue* and *anti-green*. The nature of the strong interaction requires that all physical objects are colour-neutral. This leads to the formation of bound states of quarks, called *hadrons*. Gluons themselves carry both colour and anticoulour, which means they can couple to themselves (*self-interaction*). The weak interaction is mediated by three massive vector bosons,  $W^+, W^-$  and  $Z^0$ . The charges associated with the weak interaction are the *hypercharge* and *isospin*. The large mass of the mediator bosons leads to smaller couplings of the weak interaction and a short interaction range. Finally the electromagnetic interaction is mediated by the massless spin-1 photon, often denoted as  $\gamma$ . In the SM, the neutral fields of the electromagnetic and weak interaction mix, leading to the unification of the electromagnetic and weak interactions into the so-called *electroweak* interaction.

## 2.2 Symmetries and Gauge invariance

In the previous section the structure of the SM interactions was introduced in an ad hoc way. However the behaviour of the different interactions can be motivated by underlying symmetry principles in a field theoretical formulation of the SM. The theoretical basis of this formulation of the SM is the Lagrangian density, which in its shortest form can be written as:

$$\begin{aligned}
 \mathcal{L}_{\text{SM}} = & -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\
 & + i\bar{\psi}\not{D}\psi \\
 & + \mathcal{L}_{\text{H,bosons}} \\
 & + \mathcal{L}_{\text{H,fermions}}
 \end{aligned}
 \tag{2.1}$$

The first line of equation 2.1 contains the energy-momentum tensor  $F_{\mu\nu}$  of the electroweak fields and the gluon fields. The first line therefore encodes the kinetic energies of the electroweak bosons and gluons and their self-interactions. The second line contains the fermion field  $\psi$  (for leptons and quarks) and the so-called covariant derivative  $D$ , which contains both the lepton and quark kinetic energies, as well as their interactions with the electroweak fields. Furthermore this term contains the interaction of quarks with gluons. The third line contains the Higgs potential as well as the interactions of the electroweak fields with the Higgs doublet. Finally the last line contains the interaction of leptons and quarks with the Higgs doublet. The last two lines are commonly referred to as the *Higgs sector* of the Lagrangian and will be discussed in more detail in section 2.3.

The terms in the Lagrangian density must fulfill two criteria: Firstly, each term must be a Lorentz-scalar, since the action (the integral of the Lagrangian density over time) must be a scalar following the postulates of special relativity. This requirement dictates the behaviour of the fields under Lorentz transformations. Secondly, each term must fulfill the principle of *local gauge invariance*, stating that the terms must be invariant under spacetime-dependent transformations of a certain symmetry group (*gauge group*). The gauge groups are specific for each interaction. The strong interaction requires invariance under  $SU(3)$  transformations whereas terms related to the electroweak interaction must satisfy invariance under  $SU(2)_L \times U(1)$  transformation. (The subscript  $L$  denotes that the weak interaction couples only to particles of left-handed chirality and thereby maximally violates parity. For a more detailed discussion, see e.g. [4].) These symmetry requirements dictate which terms are possible in the Lagrangian density and which terms are forbidden. This governs the structure of all interactions in the SM. In principle, all allowed terms in the SM Lagrangian density should take place in nature, or there must be an underlying mechanism for their suppression.

## 2.3 The Higgs mechanism

The observed particles of the SM have a wide range of different masses. However, including a mass term directly in the Lagrangian density violates local gauge invariance. To solve this problem, a complex scalar doublet Higgs field together with a Higgs potential is introduced in the Lagrangian density eq. 2.2:

$$\begin{aligned}\mathcal{L}_{\text{H,bosons}} &= D_\mu \phi^\dagger D^\mu \phi - V(\phi) \\ V(\phi) &= \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad \text{with } \mu^2 < 0, \lambda > 0\end{aligned}\tag{2.2}$$

Figure 2.2 shows an illustration of the Higgs potential, often referred to as the *Mexican-hat potential*. The potential itself is invariant under  $SU(2) \times U(1)$  transformations, however the minimum is moved away from zero, therefore the potential has a nonzero vacuum expectation value  $v$ , also called (*VEV*). Due to quantum fluctuations the vacuum will spontaneously be moved into the minimum of the potential and will no longer be invariant under transformations of the original symmetry group. The original symmetry is said to be *broken spontaneously*. From the form of the potential it is visible that fluctuations along the minimum result in no gradient in the potential, therefore states corresponding to these fluctuations remain massless. Each linearly independent direction in which these fluctuations can occur without a potential gradient gives rise to one such degree of freedom. In case of the Higgs potential there are three massless degrees of freedom which are absorbed by the  $W$  and  $Z$  bosons. They generate the  $W$  and  $Z$  boson masses by providing a degree of freedom for the longitudinal polarization direction needed for a massive state. This is referred to as the *Higgs mechanism*. The radial quantum fluctuations around the

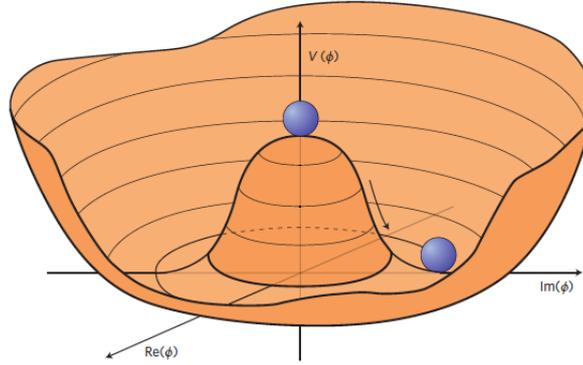


Figure 2.2: Graphical representation of the Higgs potential. From: [5]

vacuum expectation value correspond to a massive scalar field, the *Higgs field*  $h(x)$ , excitations of which give rise to the *Higgs boson*. In the vacuum state one can therefore substitute:

$$\phi = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (2.3)$$

The energy scale at which electroweak symmetry breaking (EWSB) occurs is called the *electroweak scale* which corresponds to an energy of  $E_{\text{EWSB}} = 250 \text{ GeV}$ . From the SM theory point of view, it makes sense only after EWSB to speak of the massive  $W$  and  $Z$  bosons instead of the massless fields of the unified electroweak interaction. It should be noted that, after electroweak symmetry breaking, the vacuum state still fulfills the  $U(1)_{em}$  symmetry of the electromagnetic interaction, therefore the photon remains massless.

The *spontaneous symmetry breaking* (SSB) gives rise to the masses of the weak vector bosons. Fermion masses still need to be introduced, via their interaction with the complex Higgs doublet:

$$\mathcal{L}_{\text{H,fermions}} = \bar{\psi}_L \hat{Y} \psi_R \phi + h.c. \quad (2.4)$$

If the Higgs doublet field is substituted by eq. 2.3, it becomes immediately visible that mass terms arise, as well as terms describing the interaction of the Higgs field with fermions:

$$\begin{aligned} \mathcal{L}_{\text{H,fermions}} &= \frac{\hat{Y}v}{\sqrt{2}} \bar{\psi} \psi + \hat{Y} \bar{\psi} \psi h \\ m_{\text{fermions}} &= \frac{\hat{Y}v}{\sqrt{2}} \end{aligned} \quad (2.5)$$

In the above equation,  $\hat{Y}$  is called *Yukawa-coupling*. This coupling is a free parameter of the SM which must be measured experimentally. The coupling strength between fermions and the Higgs field is proportional to the fermion mass.

## 2.4 Production and decay of the Standard Model Higgs boson

The production modes of a SM Higgs boson at the LHC are shown in figure 2.3 (a). The dominant production mode at the LHC is the so-called *gluon-gluon fusion* ( $ggF$ ), where the Higgs boson is created via a top quark loop. In the *vector boson fusion* ( $VBF$ ) mode, the Higgs is created in association with two forward quark jets. The *vector boson associated* production mode is more rare (see figure 2.3 (b)), however the decays of the associated vector bosons provide a good separation power with respect to hadronic backgrounds. The rarest production mode among the four shown in figure 2.3 (a) is the *top-quark associated* production mode, which is interesting because it directly probes the coupling among the two heaviest known particles in the SM.

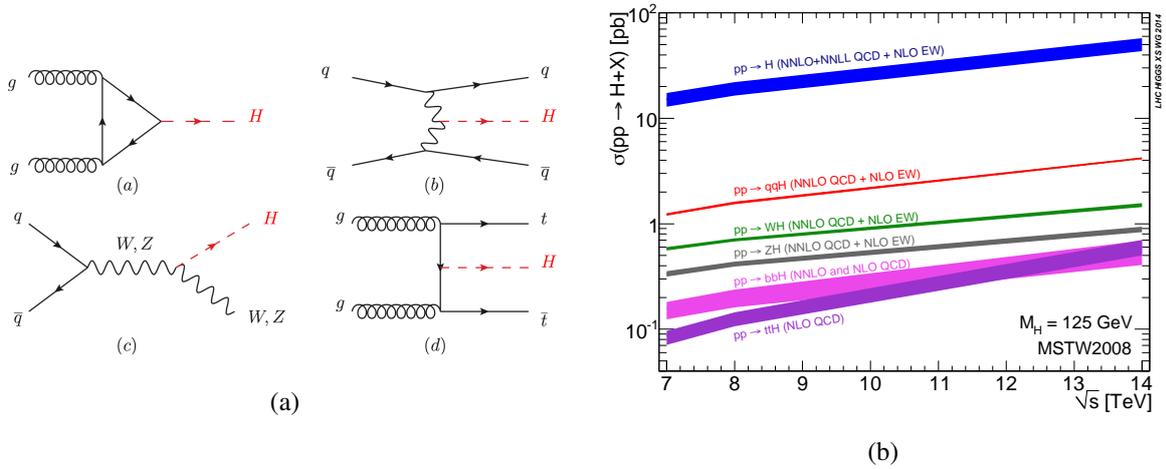


Figure 2.3: Left: Feynman diagrams for SM Higgs boson production at the LHC. (a) Gluon-gluon fusion, (b) vector boson fusion, (c) vector boson associated production and (d) top associated production. From: [3]. Right: Cross-sections for different Higgs production modes in 13 TeV  $pp$ -collisions at the LHC [6].

The branching fractions for a SM Higgs boson with a mass of 125 GeV decaying into SM particles are shown in figure 2.4. The most probable decay mode is to two  $b$ -quarks (the decay into top quarks is kinematically forbidden) with a branching fraction of 58%.

The Higgs boson was discovered in the year 2012 by the *ATLAS* and *CMS* collaborations at CERN [7, 8]. Since then, further studies of the properties of the Higgs boson have found it to be consistent with the SM expectation [9–12]. The bosonic decay modes of  $ZZ, \gamma\gamma$  [7] and  $WW$  [13] have been observed in *ATLAS*, as well as the fermionic decay into  $b\bar{b}$  [14] and  $\tau\tau$  [15].

## 2.5 The Higgs boson as a probe for new physics

Even though the experimental results obtained at the LHC so far suggest that the observed Higgs boson is consistent with the SM expectation, there are still several open questions related to the origin and nature of spontaneous symmetry breaking and the Higgs field. One prominent problem, that was already mentioned in the introduction is the *naturalness problem* [16–18]. It is related to the mass of the Higgs boson, which, in next-to-leading order (NLO) perturbative calculation receives radiative corrections with extremely large values. To obtain the observed Higgs boson mass of 125 GeV precise fine-tuning and can-

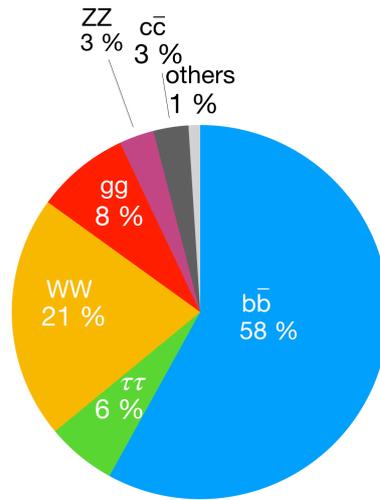


Figure 2.4: Relative branching fractions of the SM Higgs boson with a mass of 125 GeV [6]

cellation between the correction terms is required for which there is no intrinsic explanation within the SM.

There are several extensions of the SM that can protect the Higgs boson mass against large radiative corrections [19]. One broad range of models assume a new symmetry between fermions and bosons, called *Supersymmetry (SUSY)*. Assuming SUSY, every particle participating in the loop corrections to the Higgs mass would acquire a supersymmetric partner contributing an additional loop diagram which, due to the super partner's difference in spin would cancel out each SM loop correction at every order of perturbation theory. Figure 2.5 shows the top loop Feynman diagram correcting the Higgs mass, together with the diagram contributed by the super partner of the top, called *Stop*. Supersymmetric models predict an extended Higgs sector with at least three neutral Higgs bosons and two charged Higgs bosons. For many SUSY models the lightest supersymmetric particles are expected to be produced at LHC energy scales, but none have been observed so far.

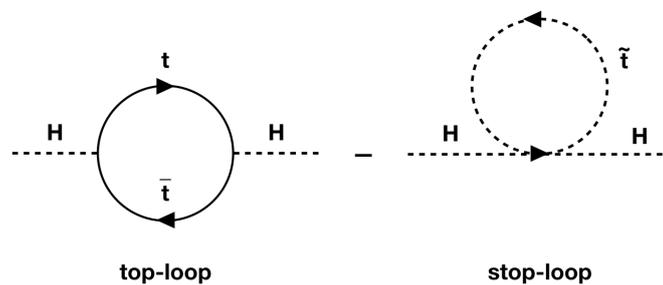


Figure 2.5: The top loop contributing to NLO corrections to the Higgs boson mass. Also shown is the stop loop which appears in supersymmetric extensions of the SM and cancels the top-loop corrections.

Another class of extensions aiming to resolve the naturalness problem assumes a new strong interaction

regime at the TeV energy scale or above. The new interaction leads to dynamical breaking of the electroweak symmetry. In dynamical symmetry breaking the role of the Higgs field is played by bound states of the original strong interaction which, similar to light mesons in Chiral symmetry breaking of QCD, break the electroweak symmetry. This would imply that the Higgs boson is not an elementary particle, but a composite bound state. The introduction of a new intermediate energy scale in the TeV range resolves the naturalness problem, because the magnitude of the NLO Higgs boson mass correction is proportional to the scale difference between the electroweak scale and the next higher scale where New Physics enters. If there is New Physics at the TeV scale, the corrections are much smaller than if one assumes that the SM holds up to the Planck scale of  $10^{19}$  GeV. For a graphical illustration of the influence of an intermediate scale to the Higgs mass correction, see figure 2.6. There are several sub-types of the TeV-scale strong interaction models, including *Minimal Walking Technicolor* [20–22], *Little Higgs* [23] and *Composite Higgs models* [24, 25].

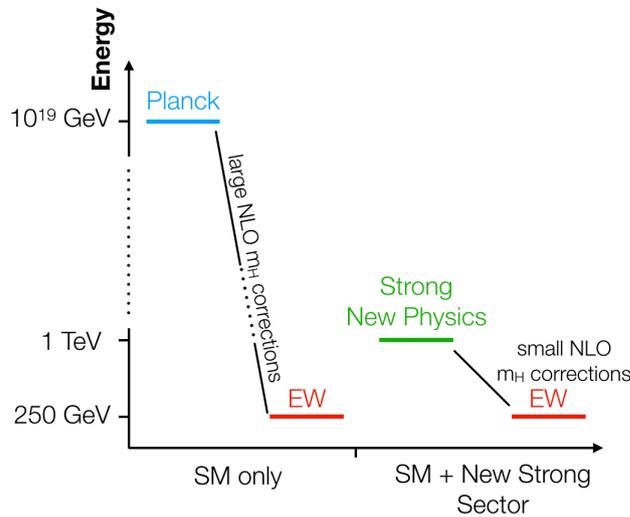


Figure 2.6: Illustration of different energy scales with and without the introduction of a new strong interaction regime at the TeV scale. The corrections to the Higgs boson mass are proportional to the scale difference between the electroweak (EW) scale and the next higher scale of New Physics.

All of the above new physics models predict the existence of heavy vector resonances decaying into SM particles. It is expected that resonances with masses up to several TeV can be produced directly at the LHC. The decay of heavy vector resonances into an electroweak boson and a SM Higgs boson is a natural prediction by many new physics models. In many cases it is an enhanced decay mode. In searches for heavy resonances, the SM Higgs boson therefore serves as a probe for new physics associated to the SM Higgs sector. Many new physics searches are not sensitive to all parameters of the underlying theoretical model, therefore simplified effective models are widely used to conduct searches that are less model-dependent [26]. An example of an effective model is the so-called *heavy vector triplet* (HVT) model [27, 28]. It is introduced in detail in section 4.1.



---

## Experimental Physics at the Large Hadron Collider

---

This chapter will serve as an introduction to the Large Hadron Collider, one of the world's largest particle accelerators, as well as the *ATLAS* detector, one of the two multi-purpose experiments at the LHC. Experimental methods and mathematical tools to extract information about physics processes from the data recorded by *ATLAS* will be described in detail.

### 3.1 The Large Hadron Collider

The *Large Hadron Collider (LHC)* is the world's largest particle accelerator. It is located at the European Centre for Particle Physics CERN on the Franco-Swiss border close to Geneva. The LHC tunnel has a circumference of 27km. Within the tunnel, two proton beams circulate contrarotatingly and are brought to collision at four different points where particle physics detectors are located. The protons, being charged particles are forced to circular orbits within the LHC tunnel by powerful superconducting dipole magnets. Particles are accelerated in several pre-accelerators, but also in a single location within the LHC using electric fields varying at radio frequencies. The beams within the LHC are not continuous but they consist of many discrete packets of protons called bunches. The maximum number of bunches within the LHC in Run-2 were 2556.

One of the main goals of the Large Hadron Collider physics programme is the discovery of new particles by producing them directly and detecting them via their decay products. To produce heavy new particles directly, a large center-of-mass collision energy  $\sqrt{s}$  is needed. For two beams of energy  $E_1$  and  $E_2$  that collide head-on, the centre-of-mass energy is defined as:

$$\sqrt{s} = \sqrt{(p_1 + p_2)^2} = \sqrt{4E_1E_2}. \quad (3.1)$$

To motivate the use of proton beams at the LHC it is useful to compare electron-positron colliders and proton-proton colliders. For a circular electron-positron collider the maximum beam energy that is reachable is always constrained by the losses through synchrotron radiation, which is emitted by charged particles being deflected on a circular orbit by a magnetic field. The total power radiated via synchrotron radiation scales with the particle energy to the power of four, decreases with the radius squared and with the mass of the particle to the power of four. This means that for protons which have 2000 times the mass of an electron, much higher energies are reachable without considerable losses through synchrotron radiation. Therefore a proton-proton collider is able to reach higher energies necessary to produce new heavy particles. The limiting aspect of the center-of-mass energy at a hadron collider is the magnetic field

strength necessary to force protons onto a circular orbit at high energies. In order to reach a design beam energy of 7 TeV in the 27 km circumference LHC tunnel, magnetic dipole fields of 8.3 T are necessary.

### 3.1.1 Luminosity

One of the most important defining parameters of a storage ring is the instantaneous luminosity, defined as:

$$L = \frac{nN_1N_2f}{4\pi\sigma_x\sigma_y}, \quad (3.2)$$

where  $n$  is the number of particle bunches in the beams,  $N_1$  and  $N_2$  are the numbers of particles per bunch in beams 1 and 2,  $f$  is the collision frequency and  $\sigma_x$  and  $\sigma_y$  describe the standard deviation of the beam in the transverse plane, assuming a Gaussian beam profile. The luminosity quantifies the number of particle interactions per area and time. Together with the cross section of a particle reaction process it determines the number of times a particle reaction is expected to occur in an accelerator experiment:

$$N(pp \rightarrow X) = \int L \sigma(pp \rightarrow X) dt. \quad (3.3)$$

In the above equation,  $N$  denotes the number of events in which the process  $pp \rightarrow X$  takes place,  $L$  is the instantaneous luminosity and  $\sigma(pp \rightarrow X)$  is the cross section for the process  $pp \rightarrow X$ . Equation 3.3 shows that the instantaneous Luminosity has the units  $\text{cm}^{-2}\text{s}^{-1}$ .

The peak instantaneous luminosity achieved at the LHC in Run-2 is  $2.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  [29]. To maximize the luminosity of the LHC the beam parameters need to be tuned. Typical tuning consists of increasing the number of particle bunches and to decrease the diameter of the bunches in the transverse plane. For this purpose focusing magnets and corrector magnets at higher multipole order are used.

### 3.1.2 Proton-proton collisions

Protons are composite objects that consist of quarks and gluons. In general there is an inverse relation between the energy scale of a probing process and the distance scale that can be resolved, thus, the higher the energy transfer in a process, the finer the structures that can be resolved. At the energy scale of the LHC, the sub-structure of the protons can be resolved and the collisions need to be described in the picture of the high-energy limit of QCD, where partons are asymptotically free. In this picture, not the protons themselves collide, but the partons within them. The term parton in this case does not refer only to the known valence quarks (u,u,d) within the protons but include gluons, and the so called *sea-quarks*, which are created by gluon splitting within the protons. The fraction of different parton types and their momentum distribution within the proton are described by the *parton-distribution functions* or PDFs. A PDF is usually shown as a dependence on the parton momentum fraction:

$$x = \frac{p_{\text{parton}}}{p_{\text{proton}}}. \quad (3.4)$$

For a proton made of three bound valence quarks the PDF would be a smeared function with a peak at  $x = 1/3$ . However the results from deep-inelastic scattering show that there is a significant contribution at low  $x$  from sea quarks and gluons. Figure 3.1 shows the PDF of the proton at different energy scales  $\mu^2$ . PDFs cannot be predicted from first principle theoretically, but they need to be measured. The evolution of the PDFs as a function of the energy scale can be calculated from theory, therefore PDFs that are

measured at a lower energy scale can be extrapolated to the energy scale of the LHC. At the LHC energy scale (fig. 3.1 (b)) the interaction of two gluons is the most likely process.

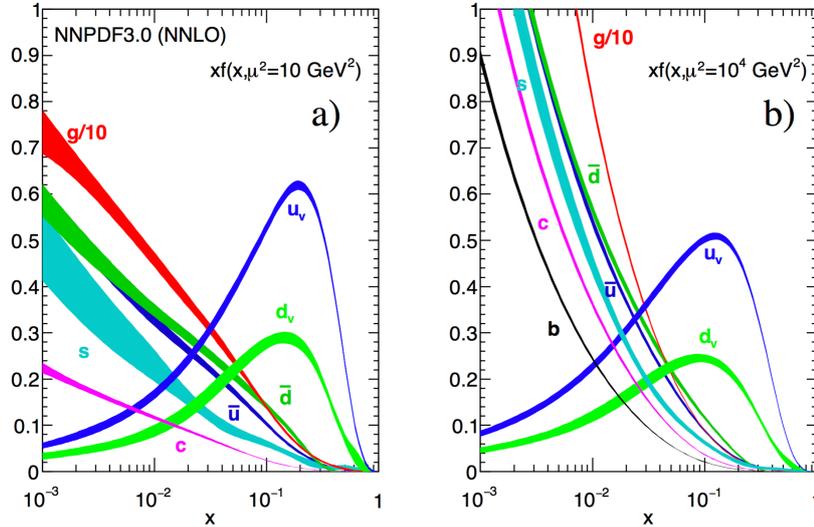


Figure 3.1: Parton distributions functions inside the proton for different energy scales  $\mu^2$ . The bands show  $x$  times the unpolarized PDF from the NNLO NNPDF3.0 global analysis [30]. From [3].

At a proton-proton collider the true center-of-mass (CM) energy does not correspond to the collider CM energy (13 TeV in Run-2), but it is the CM energy of the interacting partons, which carry a certain fraction of the initial proton momentum, depending on the PDF. The momentum fraction and the CM energy of the colliding partons is not known a priori. However, since one of the main goals of the LHC programme is the search for new particles with unknown mass, the different parton CM energies ensure that a broad range of masses is covered for the direct production of new particles.

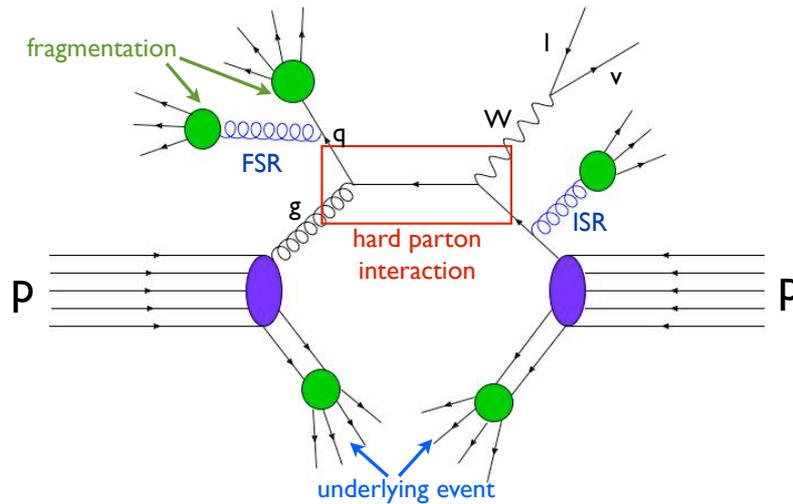


Figure 3.2: Schematic view of a proton-proton collision producing a  $W$  boson in association with jets.

A schematic view of a proton-proton collision is shown in figure 3.2. Only a small fraction of the

proton takes part in the "hard" (high momentum transfer) scattering interaction, which in the end may produce an interesting physics event. The proton remnants travel onwards along the beam axis, or are scattered at very small angles. This leads to a high particle density close to the beam which decreases for larger angles with respect to the beam axis. In the transverse plane the momentum spread of the protons is well-constrained. In the longitudinal plane the momentum of the bunches is unknown. Therefore it is not possible to make a statement about the longitudinal momentum of the rest system of the hard scattering interaction.

Due to the large number of protons in the beam, several collisions occur per bunch crossing, which manifest themselves as additional interaction vertices within the detector. These additional interaction vertices are called *pile-up*. The average number of pile-up vertices in the years 2015 and 2016 was  $\langle\mu\rangle = 13.4$  and 25.1, increasing up to  $\langle\mu\rangle = 36.1$  in 2018 [29]. In general, a higher instantaneous luminosity also corresponds to more pile-up interactions in the detector.

## 3.2 The ATLAS detector

The ATLAS detector is one of four particle detectors at the LHC. It is a multi-purpose detector of cylindrical shape (see figure 3.3) that consists of several layers of sub-systems dedicated to the detection of different particle types [31]. One of the main principles of particle detection in ATLAS is hermeticity of the detector, meaning that no interacting particle should escape detection. This is important because (transverse) momentum balance can be used as a constraint for indirect detection of particles that interact only weakly or not at all with the detector material. To achieve hermeticity the detector covers almost the full solid angle of  $4\pi$ .

### 3.2.1 Coordinates and reconstructed quantities

The ATLAS detector aims not only to determine the species of particles produced in a hard scattering reaction, but also to determine their kinematic properties. This is necessary since short-lived particles cannot be measured directly in the detector, but their properties have to be inferred from their decay products using relativistic four-momentum conservation. To determine the four-momentum of a particle, the energy, transverse momentum and two angles must be measured.

Figure 3.3 shows an overview of the ATLAS detector together with the axes of its coordinate system. The ATLAS coordinate system is a right-handed spherical coordinate system, where the z-axis lies in the direction of the beam and the x-axis points to the center of the LHC. The origin lies at the nominal interaction point in the middle of the detector. The *azimuthal angle*  $\phi$  is the angle measured in the transverse plane spanned by the x- and y-axis. Instead of the polar angle  $\theta$ , which is the ascension with respect to the beam axis, the *pseudo-rapidity*  $\eta$  is commonly used in coordinate systems at hadron colliders. The pseudo-rapidity is defined as:

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right). \quad (3.5)$$

The use of the pseudo-rapidity  $\eta$  is motivated by the fact that differences in  $\eta$  are invariant under Lorentz-boost. At a hadron collider this property is advantageous, because the particle multiplicity within intervals of  $\eta$  is approximately constant.

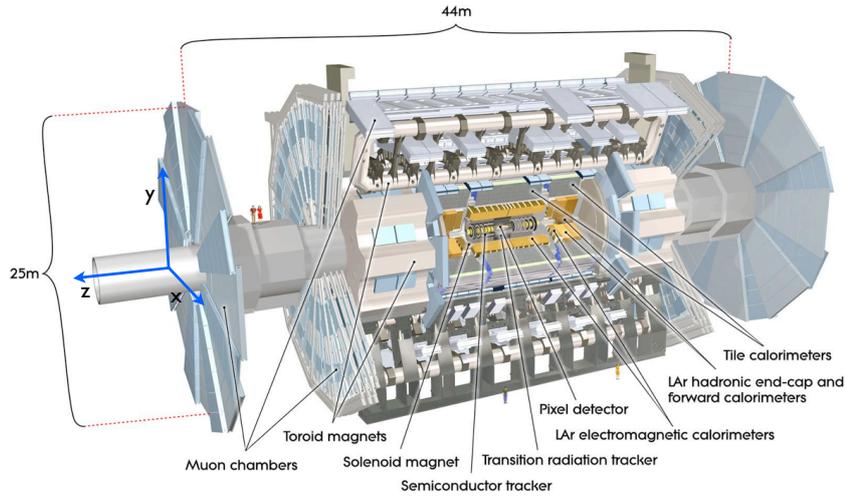


Figure 3.3: Overview of the *ATLAS* detector with the detector coordinate system. Two humans are added for scale. [32]

Having obtained both information about  $\phi$  and  $\eta$  one can define a distance measure in the solid angle between particle 1 and 2:

$$\Delta R_{12} = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2}. \quad (3.6)$$

One further important quantity besides angles, transverse momentum and energy is the missing transverse energy,  $E_T^{\text{miss}}$ . If a particle such as a neutrino leaves the detector without detection, its transverse momentum can be determined by reconstructing all visible particle signatures  $k$  in the detector and calculate the component that is missing to achieve transverse momentum conservation:

$$\mathbf{E}_T^{\text{miss}} = \sum_k \mathbf{p}_{T,k}. \quad (3.7)$$

As mentioned before, for a meaningful determination of  $E_T^{\text{miss}}$  it is necessary that all visible particles produced at the hard scattering interaction are reconstructed as precisely as possible within the detector.

### 3.2.2 Detector components

Figure 3.4 presents a schematic view of a segment of the *ATLAS* detector showing the different detector components which will briefly be introduced in this section. Since the most important physics objects studied in this thesis are jets and tracks for jet flavour identification, the focus of this chapter is to discuss the subsystems contributing to their reconstruction, namely the Inner Tracker and the Calorimeter.

#### Inner detector

The innermost detector layer [31], also called *Inner Detector* or *Inner Tracker* is dedicated to the precise reconstruction of charged particle tracks. The innermost part consists of four layers of silicon pixel detectors. The sensors of these detector layers are divided into two-dimensional segments, so-called *pixels*, where each pixel constitutes one separate channel that is read out to record if a charged particle traversed it.

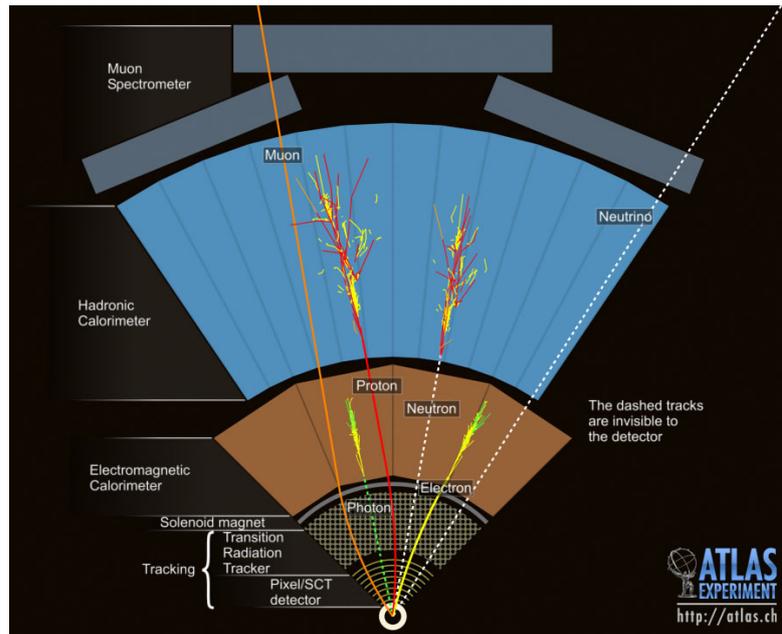


Figure 3.4: Schematic slice of a segment in the barrel region of the *ATLAS* detector showing signatures of different particle species in the sub-detector layers. Relative size of the sub-detectors are not to scale. [33]

Of the four pixel detector layers the innermost *insertable B-layer (IBL)* is situated closest to the beam pipe, at 33 mm from the interaction point. The second layer is called the *B-layer*, because historically it was the layer closest to the distance in which *B*-hadrons produced at the primary vertex decay. The following two layers are named Pixel Layer 1 and 2 <sup>1</sup>. Each layer of the pixel detector provides a three-dimensional space-point on the particle trajectory. Due to its fine segmentation, the pixel detector can tolerate extremely high radiation doses close to the interaction point [34]. The Pixel detector is followed by four layers of silicon strip detectors, called the *semi-conductor tracker (SCT)*: A single layer of strip detectors only provides a 2-dimensional space-point measurement, however three-dimensional hit information can be determined by using a double layer of strip sensors overlaid at an angle. Outside of the SCT a *transition radiation tracker (TRT)* is installed which consists of gas-filled straw tubes, that add complementary hit information up to a distance of 1 182 mm from the beam line. The extremely high granularity of the silicon pixel and strip detectors ensures a reconstruction of tracks with excellent spatial resolution necessary to reconstruct the primary vertex where the hard scattering takes place. Furthermore, the decay vertices of long-lived particles, such as *B*-hadrons can be reconstructed precisely thanks to the excellent track resolution due to hit information provided by the IBL and B-layer located very close to beam line.

As a whole, the inner detector functions as a magnetic spectrometer: it is inserted in a superconducting solenoid producing a magnetic field along the beam direction, which forces charged particles on helical trajectories. The transverse momentum of the charged particles can be deduced from the bending radius of their reconstructed tracks. For more detail on track reconstruction and parametrization, see section 3.3.

<sup>1</sup> The naming scheme of Pixel layers is historical. Pixel Layers 1 and 2 were devised before the B-layer and IBL.

## Calorimeters

The ATLAS Calorimeters [31] are used to measure the energy of particles. This is done via a destructive measurement, which means that the particles are stopped within the calorimeter material and their deposited energy is measured. There are two different types of calorimeters in ATLAS, namely the *electromagnetic calorimeter* and the *hadronic calorimeter*. The *electromagnetic calorimeter* is dedicated mainly to the reconstruction of electrons and photons. In the hadronic calorimeter, strongly interacting particles such as pions, kaons or protons are detected. If a particle enters into the calorimeter and interacts with the material it creates a cascade of secondary particles, called a *shower*. In the electromagnetic calorimeter, the dominant processes at high energies are bremsstrahlung for electrons and electron-positron pair production for photons. In a simplified model the shower generation is pictured as an avalanche, where each of the branchings corresponds to either a bremsstrahlung or pair-production process. The shower development stops, when the photon energy is below the threshold for pair production. The characteristic length of a calorimeter is called *radiation length* and is defined as the length where the energy deposited in the shower is  $1 - 1/e$  times the energy of the incoming particle. For the hadronic calorimeter, the shower development is dominated by the strong interaction and nuclear reactions, such as spallation, fission or nuclear evaporation. Some electromagnetically decaying hadrons, such as  $\pi^0$  or  $\eta$  also initiate an electromagnetic shower. Hadronic showers are in principle similar to electromagnetic showers, but their development is much more complex than electromagnetic interactions. Since nuclear interactions have a smaller cross section than electromagnetic interactions, the hadronic interaction length  $\lambda_{\text{int}}$  is larger than the electromagnetic radiation length. This is the reason why the hadronic calorimeter in the ATLAS detector is thicker (1.6 m thickness at  $\eta = 0$ ) than the electromagnetic calorimeter (0.53 m thickness at  $\eta = 0$ ).

Both ATLAS calorimeters are so-called *sampling calorimeters*, which consist of alternating layers of absorber material and active material to record the shower shape. In the electromagnetic calorimeter the absorber material consists of lead plates with layers of liquid argon scintillator in between. To reduce the amount of dead regions and to enable a fast read-out, an accordion-like geometry is chosen for the lead absorber plates. In the ATLAS hadronic calorimeter, iron plates are used alternatingly with scintillating tiles.

If a particle deposits energy in a certain segment of the calorimeter it leaves a signal in this segment, or *cell*. The individual cell signals are clustered together by the so-called *topological cell clustering algorithm* based on their signal significance with respect to the noise [35]. The topological cell clusters contain spatial and shape information and calibration and corrections depending on the cluster location and shape are applied to them. The calibrated and corrected topological cell cluster form the basic input to jet reconstruction algorithms described in 3.4.1.

For charged particles calorimeters provide a complementary measurement to the track measurement. Neutral particles, such as photons and neutrons, can only be measured in the calorimeters. The energy resolution of an electromagnetic calorimeter is typically given as:

$$\frac{\sigma_E}{E} = a \cdot \frac{1}{\sqrt{E}} \oplus b \oplus c \cdot \frac{1}{E} \quad (3.8)$$

where the first term is called the *stochastic term*, which is due to the statistical fluctuations of the number of particles  $N$  counted in the shower, fluctuations due to the sampling and quantum fluctuations that scale with  $\sqrt{N} \propto \sqrt{E}$ . The second term is the *constant term*, which comes about due to inhomogeneities and imperfections of the calorimeter material or calibration. The third term is called *noise term* which

is contributed e.g. by noise of the readout electronics. For the hadron calorimeter there is an energy-dependent modification of the above behavior due to the effect of *non-compensation*, meaning a different response for electromagnetically and hadronically induced shower parts. However, to good approximation the relation:

$$\frac{\sigma_E}{E} \propto \frac{10\%}{\sqrt{E}} \quad (3.9)$$

holds for the electromagnetic calorimeters in *ATLAS*, while

$$\frac{\sigma_E}{E} \propto \frac{42\%}{\sqrt{E}} \quad (3.10)$$

holds for the hadronic calorimeter. The increase in energy resolution at high energies and the almost constant behaviour within the full  $4\pi$  solid angle as well as neutral particle reconstruction and fast readout are among the main motivations to include calorimeters in addition to the Inner Detector in a particle detector such as *ATLAS*.

### Muon Spectrometer

Muons are the only SM particles besides neutrinos that traverse the detector material and reach the outermost detector layers. The muon spectrometer (MS) [31] provides a complementary measurement of muons in the outermost layer of the detectors. It consists of gaseous tracking detectors and large toroid magnets which bend the muon trajectories. Since the volume that the muons travel within the toroid field is much larger than in the inner detector, the deflection is high enough to retain good momentum resolution. Muons are reconstructed by matching their MS track to a track in the inner detector. If the matching fails, muons can also be reconstructed using a stand-alone measurement from the MS. For more details on muon reconstruction and quality definitions, see 5.2.

### 3.2.3 Trigger

One of the main challenges at the LHC are the enormous event rates and large backgrounds. The cross section for an inelastic proton-proton scattering is of the order of 100 mb. In comparison, the production cross section for a Higgs boson in association with a W or Z boson is ten orders of magnitude smaller, of the order of 1 pb. Most of the events produced at the LHC are soft QCD events that are not of particular interest to most current physics analyses and are far too numerous to be recorded and stored. Therefore it is of vital importance to have an automated procedure that comes to a very fast decision whether or not an event contains an interesting signature and should be kept or whether it should be discarded. This automated procedure is called *event trigger*.

The trigger used in the *ATLAS* detector consists of two steps called the *Level-1 (L1)* trigger and the *high-level trigger (HLT)* [36]. The L1 trigger is purely hardware-based and makes a first decision based on geometric or kinematic properties of simplified trigger objects reconstructed from one of the detector systems. The L1 trigger reduces the event rate from about 4 GHz to 100 kHz. The HLT is purely based on software and receives information about so-called *Regions of Interest (RoI)* from the L1 Trigger. The HLT trigger algorithms perform a local reconstruction of objects within these regions and come to a decision to keep or reject the event on the basis of their properties. There are different trigger algorithms for different objects and kinematic properties available in the *ATLAS* experiment. For some triggers the

event rate would still be too high despite a kinematic requirement on a physics object. For these triggers an additional mechanism is introduced called *prescale*, which ensures that only a certain fraction of the events passing the trigger requirements are recorded. In total, the HLT reduces the event rate further to 1 kHz. The trigger algorithms used to pre-select events in the analyses presented in this thesis include requirements on the presence of high-transverse-momentum leptons, high- $p_T$  jets or missing transverse energy.

### 3.2.4 Monte Carlo simulation

A hadronic collision event consists of several different processes that take place at a wide range of energy scales. The processes of main interest for physics take place only in the hard scattering interaction. The transition matrix elements for the hard scattering can be calculated using perturbative methods. However processes such as fragmentation and hadronization that are governed by soft QCD as well as the interaction of particles with the detector material cannot be calculated from first principle, therefore they rely on simulations based on models with tunable parameters. Simulations in particle physics make use of pseudo-random number generators and are commonly referred to as *Monte Carlo (MC) simulation*. This section is sub-divided into generation of particle-level events and the detector simulation.

#### Event generation

Event generation refers to the process of simulating a set of particles that can be fed into the detector simulation. A variety of general purpose event generators are available, which generate simulated events starting from a proton-proton-collision, step-wise adding more and more processes to the event until a complex final states with many particles is reached [37]. One of the default generators used within the ATLAS experiment, also employed in this thesis, is PYTHIA [38, 39]. The event generation in PYTHIA begins by simulating the hard scattering interaction computed to lowest order of corrections in QCD. In the next step higher order QCD and QED corrections are added to the event in a parton shower approximation, which assumes that radiation is emitted at small angle. The PYTHIA generator simulated both hard radiation from the hard scattering interaction as well as soft components to simulate the underlying event. At the end of the parton shower, the quarks and gluons from the shower are hadronized using a phenomenological model with many parameters which have to be tuned to match the observed data. To simulate specific decay modes, such as decays of  $B$ -hadrons, data-based tables of branching fractions can be used within the EVTGEN [40] package. To add the contribution from pile-up, so called *minimum-bias* events are simulated and overlaid with the previously generated MC events in the correct admixture of hard and soft components. Other background processes such as beam gas and beam halo background can be added by studying data events using the same trigger as is used to study detector noise.

As mentioned above, one of the default event generator in ATLAS is PYTHIA. However for specific processes the hard scattering simulation may be done with an alternative generator (e.g. POWHEG [41–43]) and its output is fed into PYTHIA. Samples which are generated in this way are usually denoted as e.g. POWHEG + PYTHIA. To assess the size of systematic uncertainties due to MC modelling of a particular process, events generated with an alternative shower model (e.g. with the HERWIG++ [44] generator) are used for comparison with PYTHIA (see section 5.5.1).

### Detector simulation

The detector simulation serves as a model for the interaction of the generated particles with the detector material. For the *ATLAS* experiment, the package *GEANT4* [45] is used. It contains a very finely detailed description of the *ATLAS* detector geometry and material composition and is able to simulate signatures that different particles induce in the detector material including the response of the readout electronics. The material description contains both active detector material (sensors) and passive material (support structures, readout electronics, cooling pipes etc.).

The first step of the detector simulation describes interaction processes with the material, such as ionization, bremsstrahlung, multiple scatter, photon conversion or nuclear scattering for hadrons. One of the most complex and time-consuming steps of the detector simulation is the description of electromagnetic and hadronic showers in the calorimeters. To reduce the computing time needed per simulated event, simplified models of the shower development can be used to describe the shower shapes. The second step of the detector simulation, called *digitization* contains a model description of the response of the readout electronics to the processes mentioned above. After the detector simulation an event has the exact same data format as a real proton-proton collision event from the LHC.

A realistic detector simulation is of vital importance both in the conception stage of an experiment, to optimize the expected performance of a setup, as well as in final data taking, to enable precise calibrations and good understanding of the measured data.

## 3.3 Tracks

Under the luminosity conditions of the LHC in Run-2 up to a thousand charged particle tracks are present in the *ATLAS* detector in each event. Tracks are used in almost all reconstruction processes of physics objects in the detector, be it the reconstruction of leptons, jets, missing transverse momentum or the removal of pile-up. It is therefore crucial to find a robust way to reconstruct tracks. Tracks are reconstructed from hits of charged particles in the inner detector (see 3.2.2) by the *ATLAS* track reconstruction algorithm [46]. The track reconstruction algorithm employs a step-wise procedure, which starts with the clustering of hits in the different inner detector layers, and is followed by a first loose combinatorial track candidate finding sequence. After all possible track candidates are found, ambiguities are resolved based on a scoring system which ranks track candidates higher that are more likely to correctly represent a charged particle track. After solving the ambiguities, a track fit is run for a precise determination of the track parameters using information from all inner detector systems.

Due to the solenoid field, tracks in the inner detector are bent onto a helical trajectory. The parametrization of the track is based on five parameters:

$$\mathbf{q}_{\text{track}} = (q/p, d_0, z_0, \theta, \phi). \quad (3.11)$$

In the above parametrization,  $q/p$  denotes the charge-signed inverse of the track momentum and  $d_0$  and  $z_0$  are the transverse and longitudinal impact parameters (IP). The IP denote the distance between a reference point and the track at the point of closest approach (*perigee*) in the transverse or longitudinal direction. A common choice of reference point is either the primary vertex (PV), or the beam spot. In this thesis the transverse IP is defined with respect to the beam spot, because this definition provides the tightest constraint on  $d_0$ . The longitudinal impact parameter is defined with respect to the PV. A sketch of the

impact parameter determination is shown in figure 3.5. Sometimes the *signed impact parameters* are used, which for tracks in general are defined by the direction of the track curvature. The sign is a measure for the particle charge. To characterize secondary vertices in flavour tagging it is common to use a different definition of the impact parameter sign based on the lifetime of the particle whose decay produced the track (see 3.5.3). The angles  $\theta$  and  $\phi$  are the polar and azimuthal angles in the *ATLAS* coordinate system (3.2.1). The uncertainties of the track parameter measurements are correlated and are encoded in a  $5 \times 5$  covariance matrix  $\sigma_{\text{track}}$ . The diagonal elements correspond to the squared errors on the five track parameters.

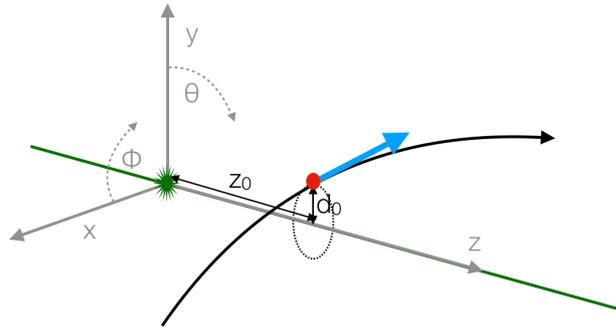


Figure 3.5: Sketch of the perigee track parametrization.

The reconstruction of tracks in *ATLAS* is especially challenging in dense environments that arise for example in boosted jet topologies. The track reconstruction efficiency in these environments is limited, resulting in degradation of reconstruction efficiencies e.g. for tracks from long-lived *B*-hadron decays from 95% at a *B*-hadron  $p_T$  of 100 GeV to 82% at 1 TeV [46]. At the beginning of Run-2 the *ATLAS* track reconstruction algorithm was specifically optimized in view of the challenges to be met in dense environments [46].

### 3.4 Jets

If strongly-interacting objects such as quarks or gluons are produced in a proton-proton collision, they form a shower of hadrons due to the confinement existing in QCD. This process is subdivided into the so-called *fragmentation*, where the initial parton splits into multiple partons, and *hadronization*, where the partons form bound states. Both fragmentation and hadronization take place instantaneously, leading to the production of a collimated shower of particles in the detector, called a *jet*. To determine the properties of hadronically decaying objects, for example the Higgs boson invariant mass in the  $H \rightarrow b\bar{b}$  decay, one needs to reconstruct the four-vectors of the initial partons. Since the only physical object in the detector is the hadronic jet, the aim of the jet reconstruction algorithm is to obtain a single four-vector from the energy depositions in the detector that is related as closely as possible to the four-vector of the jet of partons after showering.

### 3.4.1 The anti- $k_T$ jet algorithm

A jet clustering algorithm is intended to reduce the complexity of a final state with many hadrons to more simple objects to be able to make quantitative predictions for hadronic events. Jet clustering means the mapping of the four-momenta of the hadrons to the momenta of reconstructed jets. A jet algorithm must fulfill two criteria in order to avoid divergences when considering perturbative calculations. The first condition is that a good jet algorithm should be *infrared safe*, so the reconstructed jet axis should be insensitive to the radiation of a soft gluon. The second criterion is *collinear safety*, stating that the jet algorithm should be insensitive to the collinear splitting of a "hard" (high  $p_T$ ) particle. Also it is important that a jet stay as insensitive as possible to the non-perturbative effects like the hadronization, which cannot be calculated from first principle, or contributions from the underlying event.

In addition to the above criteria, which are mandatory for calculations one needs to consider the sensitivity of the jet boundaries to soft radiation. For a *soft-adaptable* jet algorithm radiation can cause irregularities at the jet boundaries. For a *soft-resilient* algorithm the jet shape is insensitive to soft radiation. In principle it is not clear which degree of sensitivity to soft radiation is desirable, however experience showed that it is easier to calibrate the energy of a jet if it is more soft-resilient. Additionally, for a soft-resilient algorithm it is simpler to mitigate contributions from the underlying event.

The default jet algorithm used in the *ATLAS* experiment is the *anti- $k_T$*  algorithm [47]. The *anti- $k_T$*  algorithm is used to cluster together four-vectors into proto-jets in a step-wise process, based on the distance measure:

$$d_{ij} = \min(k_{Ti}^{2p}, k_{Tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \quad (3.12)$$

$$d_{iB} = k_{Ti}^{2p}$$

where  $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  is the angular distance between two objects  $i$  and  $j$  and  $k_{Ti}$  is the transverse momentum. The variables  $y_i$  and  $\phi_i$  are the rapidity and azimuthal angle of particle  $i$  respectively. The parameter  $R$  is a free parameter of the jet algorithm and defines the jet cone radius. The exponent parameter  $p$  is equal to  $-1$  for the anti- $k_T$  jet algorithm. The measure  $d_{ij}$  denotes the distance between objects  $i$  and  $j$ , and  $d_{iB}$ . Also two other algorithms exist for different values of  $p$ , namely  $p = 0$ , called *Cambridge-Aachen algorithm* and  $p = 1$ , called  *$k_T$  algorithm*. Both differ from the anti- $k_T$  algorithm regarding the clustering of soft components [47].

The jet reconstruction is performed by finding the minimal distance parameter among a set of objects (initial particles or proto-jets). If the minimal distance is  $d_{ij}$ , objects  $i$  and  $j$  are grouped together to form a proto-jet. If the minimal distance is  $d_{iB}$ , the object  $i$  is called a jet and it is removed from the set of input objects. The clustering procedure is repeated until no objects are left in the set of inputs.

For an event with a small number of hard objects and many soft ones, the anti- $k_T$  algorithm starts by clustering the hard objects with their surrounding soft objects. An example event is shown in figure 3.6. If there is a hard particle, surrounded by soft particles, with no other hard particles in a vicinity of  $\Delta_{12} < 2R$ , it will form a conical jet of radius  $R$  (see turquoise jet in figure 3.6). If there is another hard particle within  $R < \Delta_{12} < 2R$ , there will be two jets, both of which are not perfectly conical. Their shape depends on the relative size of their momenta  $k_{T1}$  and  $k_{T2}$ . If  $k_{T1} \gg k_{T2}$  (e.g. green and magenta

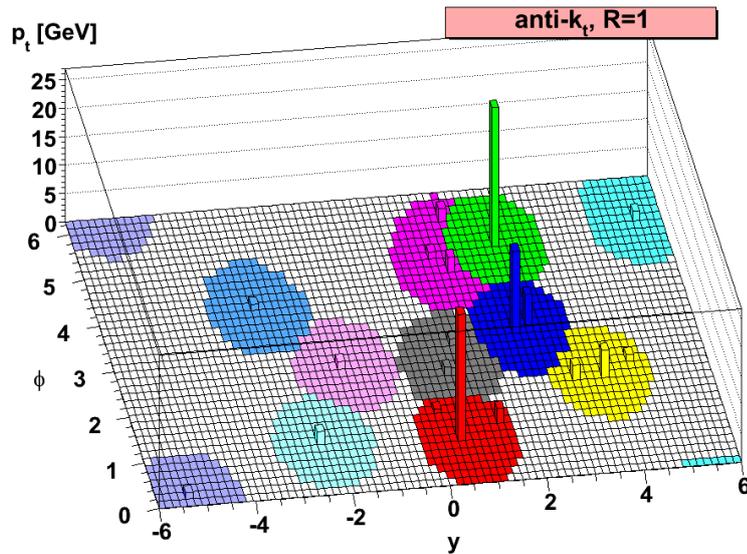


Figure 3.6: Example event at parton-level, clustered with the anti- $k_T$  jet algorithm. [47]

jets), jet 1 will be conical and jet 2 will be semi-conical, missing the section overlapping with jet 1. For  $k_{T1} = k_{T2}$ , both jets will not be conical but separated by a straight line between their boundaries (e.g. dark blue and yellow jet in figure 3.6).

In principle, any set of four-vectors can be used as inputs to a jet clustering algorithm. The standard choice in the *ATLAS* collaboration is to use clusters of energy depositions in the hadronic calorimeter. For use in the jet flavour identification also jets reconstructed from inner detector tracks (*track jets*) are used. For MC-based studies also the four-vectors of MC truth particles are used as inputs to the jet clustering to create so-called *truth jets*. For the clustering of truth jets, only stable truth particles ( $c\tau_0 > 10$  mm) are considered, also particles which do not leave a significant energy deposit in the calorimeter, such as muons and neutrinos are excluded.

Finally, a jet reconstruction algorithm should be fast, because of the large number of hadrons in an average jet at the LHC energy scale. Within the *ATLAS* collaboration the anti- $k_T$  algorithm from the FASTJET package is used for fast and efficient jet reconstruction [48]

### 3.4.2 Large- $R$ jets

The jet cone radius  $R$  is a free parameter of the jet reconstruction algorithm and determines the angular scale to which the decay products in hadronic final states can be resolved in separate jets. The default jet radius used by the *ATLAS* collaboration in Run-2 is  $R = 0.4$ . In boosted topologies with very collimated hadronic final states (see 3.6) a radius of 0.4 is often too wide to resolve the final state in two or more separate jets, therefore one follows the strategy to reconstruct all decay products in the final state (e.g.  $H \rightarrow b\bar{b}$ ) within a single *large- $R$*  jet. The default jet radius for a large- $R$  jet used by the *ATLAS* collaboration is  $R = 1.0$ .

Because of their large catchment area, large- $R$  jets are more likely to have energy contributions from

pile-up events in the detector than  $R = 0.4$  jets. To mitigate this, a removal of pile-up contributions is applied, called *jet grooming*. Several different algorithms for jet grooming exist. The default jet grooming algorithm used by the *ATLAS* collaboration is called *jet trimming* [49, 50]: For trimming, small- $R$   $k_t$  subjets are reconstructed within the ungroomed large- $R$  jet (*parent jet*). The  $k_t$  algorithm is chosen, because it clusters neighbouring soft components before clustering them with hard components, which is desirable for the identification of pile-up contributions. After the subjet reconstruction is completed, all subjets whose transverse momentum fraction is below a certain fraction of the large- $R$  jet  $p_T$  are removed. Default trimming parameters used by the *ATLAS* collaboration in Run-2 are a subjet radius of  $R = 0.2$  and a momentum fraction threshold of  $f = 5\%$ . The remaining jet is called the *trimmed jet*. Trimming leads to a more stable behaviour of jet-related quantities with respect to the average number of pile-up interactions per event  $\langle \mu \rangle$ .

Large- $R$  calorimeter jets are calibrated in the same way as their equivalents with smaller radius, by studying the ratio between properties of the calorimeter jets and their equivalent truth jet. This calibration is purely based on simulation and includes a correction of jet mass as well, since the energy scale correction is applied as a multiplicative factor to the jet four-vector [51]. In a second step called the *in situ calibration* [52], energy and mass scale uncertainties are derived based on comparisons between data and simulation. The energy scale in situ calibration methods make use of the momentum balance in events with either  $\gamma + \text{jet}$ ,  $Z(l\ell) + \text{jet}$  or multijet events where the momenta of the calibrated photon, Z-boson or, for multi-jet events, a small- $R$  jet can be used to deduce the momentum of the recoiling large- $R$  jet. For the mass scale in situ calibration, two methods are combined. The first method is using semi-leptonic  $t\bar{t}$  events to measure the jet mass response using the mass spectra of the hadronically decaying high- $p_T$  W boson or fully hadronic top quark. The second method uses the so-called  $R_{\text{trk}}$  method, where the ratio between jet properties derived using calorimeter information only and properties derived using inner detector tracks only is studied, to have two independent measurements from different detector systems [52]. It should be noted that the in-situ calibration is only applied to jets in data.

### 3.4.3 Large- $R$ jet mass

The jet mass can be, at first glance, a surprising quantity. If jets originating from a quark (with the exception of the top quark) are produced at LHC energies the assumption that the jet invariant mass is negligible compared to the transverse momentum is justified in almost all cases. For gluon jets it may appear surprising at first that some jets have a non-zero mass, even though the gluon is massless in the SM. This can be explained by gluons being produced as virtual, off-shell particles. For large- $R$  jets that contain the full hadronic decay of a heavy object (e.g.  $H \rightarrow b\bar{b}$  or  $t \rightarrow bq\bar{q}$ ) the assumption that the invariant mass is negligible no longer holds. In fact, the invariant mass of the large- $R$  jet is a very important discriminating variable for the identification of boosted heavy object decays.

Even though the four-vectors of the energy clusters for calorimeter jet reconstruction are assumed to be massless, the resulting jet four-vector, reconstructed by sequentially adding up the constituent four-vectors, can have a non-zero mass. The mass resolution depends on the angular resolution of the constituents. At high transverse momentum, the jet constituents become more collimated, therefore the angular distance between the energy clusters in the calorimeter decreases on average. As a result the jet mass resolution drops at high transverse momentum. This loss in resolution can be mitigated by the so-called *track assisted mass*,  $m^{\text{TA}}$  [52], which uses inner detector tracks for the mass reconstruction due to their higher angular resolution compared to calorimeter clusters. A weighted combination of the calorimeter-based ( $m^{\text{calo}}$ ) and track-assisted mass definition, called the *combined mass*  $m^{\text{comb}}$  is used in many analyses with

large- $R$  jets [52]. For a  $H \rightarrow b\bar{b}$  jet,  $m^{\text{comb}}$  has a similar resolution as  $m^{\text{calo}}$  in the transverse momentum range  $p_T < 700$  GeV, and improves at high transverse momentum [53].

### 3.4.4 Jet substructure

Jet substructure variables encode the distribution of energy depositions within a large- $R$  jet. The substructure is very distinct for jets produced by different hadronically decaying objects (see also 3.6.2) therefore it can be used to identify which object decay produced the jet. A multitude of substructure variables are available, which are often based on energy correlation functions or geometrical properties of the jet. A short description of these variables can be found in [54, 55]. In this thesis two substructure variables are used, namely  $\tau_{21}$  and  $D_2^{\beta=1}$ , which will be explained in more detail in the following.

The variable  $\tau_{21}$  is defined as the ratio between the 2-subjettiness  $\tau_2$  and the 1-subjettiness  $\tau_1$  [56, 57] :

$$\tau_{21} = \frac{\tau_2}{\tau_1} \quad (3.13)$$

The  $n$ -subjettiness  $\tau_n$  is determined by reconstructing exactly  $n$  subjets using the exclusive  $k_T$  algorithm (see 3.6.4) and calculating the  $p_T$ -weighted sum of the minimal angular distance  $\Delta R$  (see definition in 3.4.1) between a subjet and an object  $k$  that has not been clustered into one of the subjets:

$$\begin{aligned} \tau_n &= \frac{1}{c_0} \sum_k p_{T,k} \min \{ \Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{n,k} \}, \\ c_0 &= \sum_k p_{T,k} R_0, \end{aligned} \quad (3.14)$$

where  $R_0$  is the jet radius used to reconstruct the original jet. Note that in the special case  $n = 1$ , the 1-subjettiness encodes the distribution of energy depositions around the original jet axis. A modification of the  $n$ -subjettiness with additional separation power is called  $\tau_n^{\text{wta}}$ , where *wta* stands for *winner-takes-all*, where the angular distance  $\Delta R$ , is calculated with respect to the hardest subjet instead of the nearest subjet. In this thesis, the  $\tau_n^{\text{wta}}$  is used to calculate subjettiness ratios.

For a large- $R$  jet with two separated energy depositions (e.g. a  $H \rightarrow b\bar{b}$  jet) the 2-subjettiness has a small value, whereas the 1-subjettiness has a large value (see figure 3.7). Therefore the ratio  $\tau_{21}$  tends towards small values for 2-prong jets. For a jet from QCD, with just one single collimated energy deposition, the 1-subjettiness would tend to small values, therefore  $\tau_{21}$  tends to large values (again figure 3.7). Thus, the variable  $\tau_{21}$  is a good discriminator between Higgs boson decays and QCD. Following a similar argument, the ratio  $\tau_{32}$  can be used to discriminate between jets from hadronically decaying top quarks and Higgs boson jets.

The second important substructure variable studied in this thesis is called  $D_2^{(\beta)}$  [58, 59]. It is a variable encoding the shape of the jet, defined as the ratio of energy correlation functions:

$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}, \quad (3.15)$$

where  $e_2^{(\beta)}$  and  $e_3^{(\beta)}$  are the 2- and 3-point energy correlation functions. The energy correlation functions

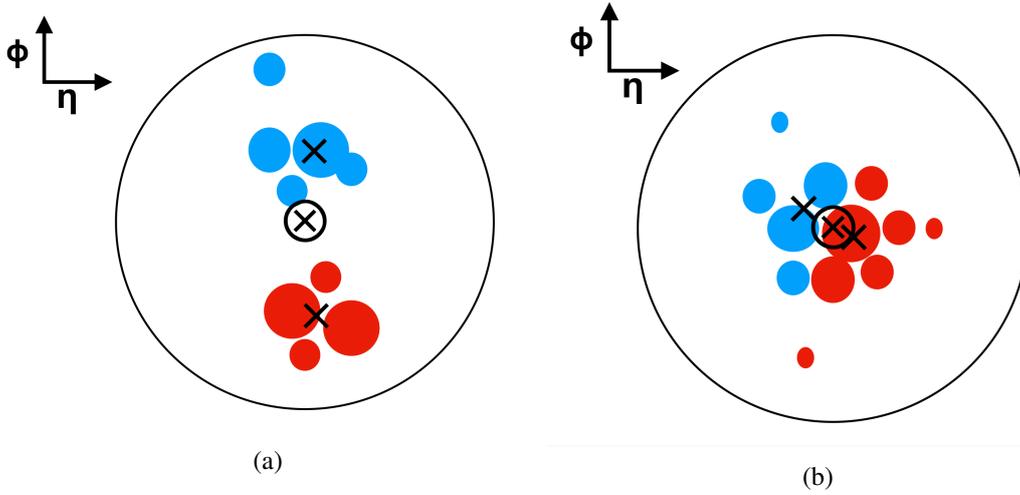


Figure 3.7: Sketch of the determination of the  $\tau_{21}$  subjeettiness ratios for (a) a 2-prong jet (e.g.  $H \rightarrow b\bar{b}$ ) with small  $\tau_{21}$  and (b) a 1-prong jet (e.g. QCD) with large  $\tau_{21}$ . The colours indicate the subject splitting with the  $k_T$  algorithm. The  $\times$  symbols denote the subject axes and  $\otimes$  denotes the original jet axis.

are defined as:

$$\begin{aligned}
 e_2^\beta &= \frac{1}{p_{T,J}^2} \sum_{ij} p_{T,i} p_{T,j} \Delta R_{ij}^\beta, \\
 e_3^\beta &= \frac{1}{p_{T,J}^3} \sum_{ijk} p_{T,i} p_{T,j} p_{T,k} \Delta R_{ij}^\beta \Delta R_{ik}^\beta \Delta R_{jk}^\beta,
 \end{aligned} \tag{3.16}$$

where  $p_{T,J}$  is the transverse momentum of the jet and  $p_{T,i}$  is the transverse momentum of object  $i$  and  $\Delta R_{ij}^\beta$  is the angular distance between particles  $i$  and  $j$ .

Figure 3.8 illustrates the structure of radiation within a jet with a single energy deposition (*1-prong*, similar to QCD jets) and with two correlated energy depositions (*2-prong*, similar to  $H \rightarrow b\bar{b}$  jets). While the 1-prong jet is dominated by a hard radiation component surrounded by soft radiation, the 2-prong jet has two hard, collinear components, soft radiation as well as soft collinear radiation that is emitted by the dipole created by the two hard components. Considerations about the angular and momentum scaling of different combinations of objects from the hard, soft, or soft-collinear components yield the inequalities (for the case  $\beta = 1$ ):

$$\begin{aligned}
 (e_2)^3 &\lesssim e_3 \lesssim (e_2)^2 && \text{(1-prong)} \\
 0 < e_3 &\ll (e_2)^3 && \text{(2-prong)}
 \end{aligned} \tag{3.17}$$

It follows from equation 3.17 that the ratio of the energy correlation functions  $D_2^{\beta=1}$  is small for 2-prong jets and large for 1-prong jets.

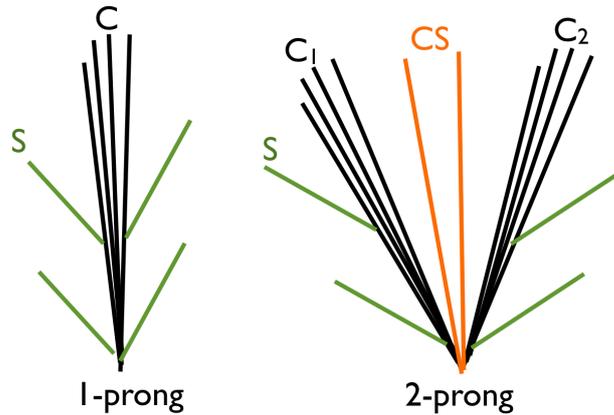


Figure 3.8: Sketch of the radiative structure for a 1-prong jet (left) and a 2-prong jet (right). For the 1-prong jet the collinear (C) component and soft (S) components are dominant, for the 2-prong jet there is additional soft collinear (CS) radiation from the hard colour dipole.

### 3.4.5 Ghost association

In some cases it is necessary to match objects to a jet, for example to match small- $R$  track jets to a large- $R$  calorimeter jet for flavour tagging. Especially in boosted topologies, simple  $\Delta R$ -based matching procedures often result in ambiguities. *Ghost association* provides a robust ambiguity-free matching prescription in boosted jet topologies. For ghost-association, the four-vectors of the objects to be associated to a jet are added among the list of constituents for the reconstruction of the original jet. The same jet reconstruction algorithm is run for the ghost-association as was used for the original jet reconstruction (see 3.4.1). In the example of associating track jets to a large- $R$  calorimeter jet, the four-vectors corresponding to the track jet axes are added to the list of four-vectors of calorimeter clusters for the large- $R$  jet reconstruction algorithm. If the track jet four-vectors are clustered within one of the large- $R$  jets, they are said to be *ghost-associated* to this jet. To ensure that the four-vectors of the objects to be associated (track jets) do not distort the large- $R$  jet axis, their transverse momenta are set to zero. The ghost-association method has the advantage that it can account for irregular jet shapes, or, in the case of groomed jets, that objects can be associated to an area of a jet that was removed by the grooming.

## 3.5 Flavour tagging

The identification of the jet flavour is paramount at a hadron collider where the backgrounds from light-flavour jets from QCD are large and many interesting final states such as the  $H \rightarrow b\bar{b}$  decay crucially rely on good separation between light-flavour and heavy-flavour jets.

### 3.5.1 Properties of $b$ -jets

The most important property of a  $b$ -jet is the presence of a secondary vertex (SV) from the decay of a  $B$ -hadron. The long lifetime of the  $B$ -hadron, of the order of pico-seconds, as well as its high momentum fraction can lead to a measurable distance between primary vertex (PV) and the SV. The long lifetime of the  $B$ -hadron can be explained by the fact that  $B$ -hadrons decay primarily via the weak interaction which

has a much smaller transition probability than e.g. strong-interaction processes. The high momentum transfer is due to the fragmentation function of the  $b$ -quark which is shown in figure 3.9. As can be seen the momentum transfer from the original  $b$ -quark to the  $B$ -hadron peaks at a value of about 85% of the original  $b$ -quark momentum. For  $b$ -jets created at the LHC the distance between PV and SV can be of the order of millimetres to centimetres.

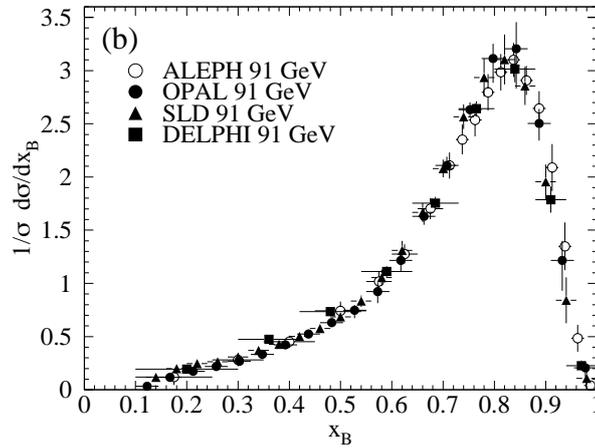


Figure 3.9: Fragmentation function of a  $b$ -quark fragmenting into a  $b$ -hadron measured in  $e+e$  collisions at  $\sqrt{s} = 91$  GeV [3].  $x_B = E_B/E_{\text{beam}}$  denotes the scaled energy of the  $B$ -hadron.

Figure 3.10 shows the topology of a  $B$ -hadron decay within a  $b$ -jet. The  $b$ -quark (e.g. from a  $H \rightarrow b\bar{b}$  decay) is produced at the primary vertex and immediately hadronizes, forming a  $B$ -hadron, which, after a finite flight distance decays either directly into light hadrons, or, as shown in figure 3.10 decays via a charmed hadron, such as the  $D$  or  $D^*$ , producing a secondary decay vertex in both cases. Since charmed hadrons themselves have a significant lifetime, decays via charm hadrons can produce an additional tertiary vertex, which is not as far removed as the secondary from the primary vertex, but, with sufficient momentum transfer, the flight distance between secondary and tertiary vertex can be measured in the inner detector.

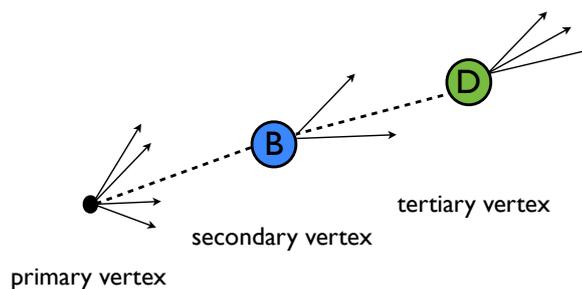


Figure 3.10: Topology of the production and decay of a  $B$ -hadron in a  $b$ -jet. [60]

### 3.5.2 Flavour tagging procedure in ATLAS

The jet flavour identification (*flavour tagging*) used in the *ATLAS* experiment [61] consists of several steps that transform the input jet into a jet with a decision on its flavour content (*tagged jet*). This procedure is

depicted in figure 3.11.

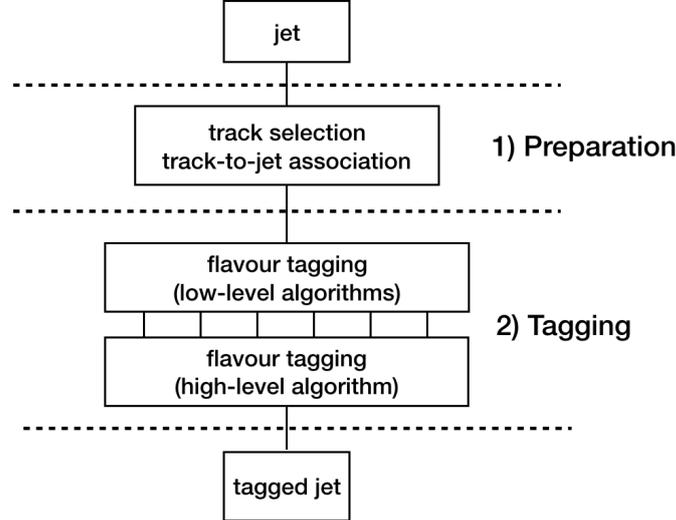


Figure 3.11: Procedure of jet flavour tagging in *ATLAS*.

Most quantities that are used in the flavour tagging algorithms are related to inner detector tracks, which are associated to the input jet in the first preparatory step. Tracks are associated to jets using a shrinking-cone association method, where tracks are associated to a jet if they are within a cone of  $\Delta R < R_{\text{assoc}}$  where  $R_{\text{assoc}}$  follows the functional form [61]:

$$R_{\text{assoc}} = 0.239 + e^{-1.22 - 0.0164 \cdot p_T / \text{GeV}} \quad (3.18)$$

This means that  $R_{\text{assoc}}$  decreases from a value of  $R_{\text{max}} = R_{\text{assoc}}(20 \text{ GeV}) = 0.45$  down to a minimal value of  $R_{\text{min}} = 0.239$  with increasing jet transverse momentum. The track association cone size as a function of  $p_T$  is shown in figure 3.12. The track-to-jet association is exclusive, meaning that a track that is within the shrinking cone around more than one jet is associated to the closest jet.

The track-association step takes place regardless of whether the input jet four-vector has been reconstructed from calorimeter clusters or from inner detector tracks. Upon first consideration it may surprise the reader that the additional step of associating tracks to a track-jets is necessary. However, the selection criteria applied to tracks used for the reconstruction of the track jet axis can be different than for tracks used for  $b$ -tagging.

The second step in the flavour tagging procedure in figure 3.11 is the application of a  $b$ -tagging algorithm. Flavour tagging algorithms can be subdivided into low-level algorithms, such as impact-parameter based algorithms or secondary-vertex-based algorithms, and high-level algorithms which make use of multivariate analysis methods and use variables calculated by the low-level algorithms as inputs. The final discriminant for flavour tagging is the output variable of one high-level flavour tagging algorithm. Typically a threshold on the discriminant at a given  $b$ -jet efficiency is used. If the flavour tagging discriminant determined for a certain jet lies above that threshold, the jet is denoted as *tagged*.

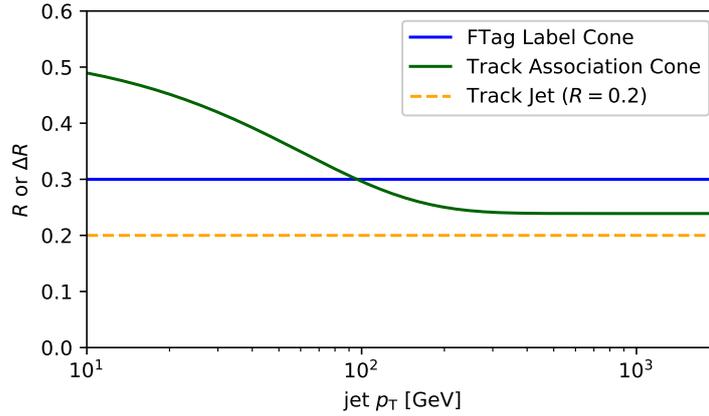


Figure 3.12: Size of the  $\Delta R$  cone used to associate tracks to jets as well as MC truth particles to jets as a function of jet  $p_T$ . The green curve shows the shrinking size of the track association cone. The blue curve shows the fixed cone for jet labelling in MC based on the association of truth hadrons. The dashed orange line is added for orientation and shows the fixed  $R=0.2$  cone used for the reconstruction of track jets [61, 62].

For more detail on the definition of thresholds on the flavour tagging discriminant, see section 3.5.7.

### 3.5.3 Impact-parameter based $b$ -tagging

The impact parameter (IP) of a track is defined as the distance of closest approach between the track and the PV in the transverse ( $d_0$ ) or longitudinal ( $z_0$ ) direction. Impact parameter based  $b$ -tagging algorithms make use of the fact that tracks from  $B$ -hadrons decaying at a SV typically have a large transverse IP. Furthermore it is exploited that secondary vertices from decays of long-lived  $B$ -hadrons have a positive flight-distance from the PV along the jet axis. Therefore one can define the so-called *lifetime-signed* impact parameter:

$$\text{sign}(d_0) = \begin{cases} -1 & \vec{v}_{\text{PVtrk}} \cdot \vec{v}_{\text{jet}} < 0 \\ +1 & \vec{v}_{\text{PVtrk}} \cdot \vec{v}_{\text{jet}} \geq 0 \end{cases}$$

Here,  $\vec{v}_{\text{PVtrk}}$  denotes the vector from the primary vertex to the track on the point of closest approach, and  $\vec{v}_{\text{jet}}$  denotes a vector pointing in the direction of the jet axis. Figure 3.13 shows an illustration of the lifetime-sign determination. Prompt tracks with negative transverse impact parameter are caused by the finite impact parameter resolution. Also other effects, such as photon conversion or hadronic material interactions can cause negative  $d_0$ . For a more detailed description of causes for negative  $d_0$ , see 5.4.

For  $b$ -tagging, not only the magnitude of the transverse impact parameter is of interest, but also its uncertainty. One therefore defines the transverse IP significance  $s_{d_0}$ :

$$s_{d_0} = \frac{d_0}{\sigma_{d_0}}, \quad (3.19)$$

where  $\sigma_{d_0}$  denotes the measurement uncertainty of  $d_0$ . Tracks from a  $B$ -hadron decay typically have a well-measured large transverse IP, therefore a large significance  $s_{d_0}$ .

The IP based  $b$ -tagging approach crucially relies on good-quality tracks, therefore the track selection is tight compared to the SV based approach. The IP track selection requires at least 7 precision layer hits (in

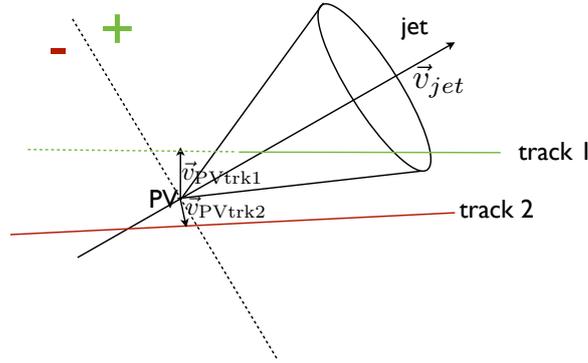


Figure 3.13: Illustration of the lifetime-signed transverse IP. From [60].

the pixel or strip detectors) of which 2 hits must be in the pixel detector,  $p_T > 1$  GeV, as well as tight cuts on the IP,  $|d_0| < 1$  mm and  $|z_0| \sin(\theta) < 1.5$  mm which are intended to reduce the contribution from pileup tracks [61].

#### 3.5.4 Secondary-vertex based $b$ -tagging

Since the decay of  $B$ -hadrons can produce a SV, a natural approach is to reconstruct these vertices using a vertex finding tool. The presence of a SV within a jet is a good indicator for a  $b$ -jet. However, it is not sufficient to discriminate between  $b$ -jets,  $c$ -jets and light jets, since charmed hadrons, or light-long-lived hadrons such as kaons or  $\Lambda$ -particles can also produce a SV. Therefore different properties of the reconstructed SV are studied, such as the three-dimensional flight distance between PV and SV (depicted in figure 3.14). Other important properties are the invariant mass calculated from the sum of track four-vectors at the SV, which is close to the  $B$ -hadron mass, as well as the track multiplicity at the SV.

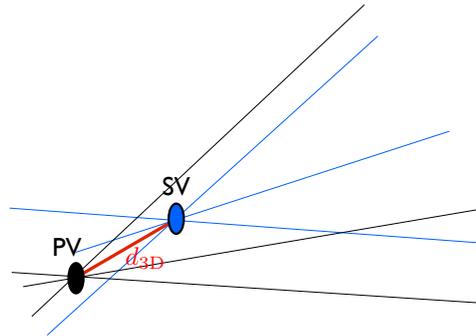


Figure 3.14: Reconstruction of the secondary vertex and definition of the three-dimensional flight distance. The black tracks stem from the primary vertex and the blue tracks stem from the SV. The red line corresponds to the three-dimensional flight distance between the secondary and primary vertex.

To retain efficiency for the reconstruction of the SV, the track preselection for SV-based  $b$ -tagging is looser than for IP-based  $b$ -tagging. Typical selection requirements are [61]:  $p_T > 500$  MeV,  $|d_0| < 3.5$  mm, as well as at least one hit in the Pixel layer. Due to the looser requirement, also tracks from material interaction or decays of light long-lived particles enter, however these are removed by

imposing certain criteria on the position of the reconstructed SV, or the invariant mass calculated from the track four-momenta associated to the vertex.

### 3.5.5 Flavour tagging algorithms in ATLAS

Different algorithms are used for identifying the flavour content of jets in the ATLAS detector. In general, quantities calculated by the low-level  $b$ -tagging algorithms are used as input variables to the high-level algorithms. In the end, a single  $b$ -tagging discriminant is obtained by a high-level algorithm.

#### Impact-parameter based: IP2D/IP3D

The IP2D and IP3D algorithm are the default impact-parameter based  $b$ -tagging algorithms used in ATLAS [63, 64]. IP2D makes use of the transverse IP significance and IP3D simultaneously uses the transverse and longitudinal IP significances in two-dimensional templates to account for correlations between the two. Both algorithms construct templates for different jet flavours and use a Log-Likelihood-ratio to discriminate between tracks from  $b$ -jets and jets of other flavour origin. To increase the separation power the tracks are sorted into categories depending on their hit pattern in the inner detector.

#### Secondary-vertex based: SV

The secondary vertex finder (SV) aims to explicitly reconstruct a SV within a jet [63, 64]. All possible track pairs within the jet are tested for a two-track vertex hypothesis. After removing candidates likely to originate e.g. from conversions or light long-lived particles, a new vertex fit is performed using only the tracks from the preceding two-track vertex candidates, removing outlier tracks iteratively until the  $\chi^2$  of the fit is optimal. Properties of the found SV that are used for  $b$ -tagging are the invariant mass at the SV, the energy fraction as well as the distance between the primary and SV positions.

#### Topological $b$ -tagging: JetFitter

The JetFitter algorithm [65] makes use of a so-called *Kalman filter* to fit the entire topology of the  $B - C$ -hadron decay cascade within the jet. The first step is to assume that both the  $B$ -hadron and  $C$ -hadron decay vertices lie on a single line in the direction of the jet axis, which is assumed as seed for the  $B$ -hadron flight direction. All tracks crossing the axis are considered as possible vertex candidates. These vertex candidates are merged step-by-step until a minimization criterion for the vertex fit is reached. The axis connecting the vertex candidates is updated in each step. It is assumed to be a straight line and interpreted as the  $B$ -hadron flight direction. The quantities used for  $b$ -tagging are similar to the SV algorithm, however the JetFitter algorithm has the advantage that it can find vertices with only one track, due to the assumptions about the  $B$ -hadron decay topology and flight direction.

#### High-level $b$ -tagging: MV2

The MV2 algorithm is a high-level  $b$ -tagging algorithm based on multivariate Boosted Decision Trees (BDT) [64]. It uses a set of 24 input variables from the low-level algorithms IP2D, IP3D, SV and JetFitter [66]. The algorithm is trained on  $b$ -,  $c$ - and light jets from a simulated sample of  $t\bar{t}$  events as well as on a simulated sample of a hypothetical  $Z'$  decaying to two top quarks, to improve the performance for  $b$ -tagging of high- $p_T$  jets. A certain fraction of charm jets is introduced in the training sample to improve the separation power between  $b$ -jets and  $c$ -jets. Different versions of MV2 which have been trained originally with a  $c$ -jet-fraction of  $f_c = 0\%$ ,  $10\%$ ,  $20\%$  are named MV2c00, MC2c10 or MV2c20 [66]. In

a later refinement of the MV2 algorithm the fractions were modified to enhance the charm jet rejection while keeping the light flavour jet rejection similar to the previous version [64]. The refined version of MV2 was trained with  $f_c = 0\%, 7\%, 15\%$ , whereas the naming scheme (MV2c00, MC2c10, MV2c20) remained the same. The default algorithm used in *ATLAS* in Run-2 is the MV2c10 algorithm. The MV2 algorithm is trained for  $R = 0.4$  calorimeter jets. A common approach in *ATLAS* Run-2, also in this thesis (see 3.5.8), is to use track jets for flavour tagging. It should be noted that the MV2 algorithm has not been retrained specifically for track jets, so the MV2 version optimized for  $R=0.4$  calorimeter jets is used.

### 3.5.6 Jet truth labelling

In order to define flavour tagging efficiencies in MC simulation, it is necessary to define the true flavour of a jet. As explained in section 3.2.4, for every simulated event a record of the types of simulated particles as well as their four-vectors is kept, before any detector effects are taken into account. This record of so called *truth particles* can be used to determine the true flavour content of a jet. In principle there are two ways to determine the flavour of a jet from the truth particles. The first possibility is to use the flavour of the parton that initiated a jet to determine its flavour, the other is to use the flavour of truth hadrons within the jet. Since partons are non-physical objects, the standard procedure in *ATLAS* is to use a hadron-based truth label. A  $\Delta R$  cone-based association is used, where a truth hadron with  $p_T > 5$  GeV and  $|\eta| < 2.5$  is associated to a jet if its four vector lies within  $\Delta R < 0.3$  around the jet axis (cf. figure 3.12). If a hadron is within that cone for several jets, which is commonly the case in boosted topologies, it is associated to the closest jet. Therefore, also the flavour labelling scheme is exclusive. The association sequence is as follows:

1. If a  $B$ -hadron fulfills the association criteria, the jet is labelled a " $b$ -jet".
2. If 1.) is not fulfilled and a  $C$ -hadron fulfills the association criteria, the jet is labelled a " $c$ -jet".
3. If neither 1.) nor 2.) are fulfilled and a  $\tau$  fulfills the association criteria, the jet is labelled a " $\tau$ -jet".
4. If neither 1.) nor 2.), nor 3.) are fulfilled, the jet is labelled a "light-jet".

### 3.5.7 Performance of flavour tagging

The performance of flavour tagging is expressed in terms of the efficiency to correctly identify true  $b$ -jets as well as the rejection of charm or light jets. The former is called the  *$b$ -tagging efficiency* defined as:

$$\epsilon_b = \frac{N(\text{true } b\text{-jet} \mid \text{identified as } b\text{-jet})}{N(\text{true } b\text{-jet})}. \quad (3.20)$$

The fraction of light, charm or  $\tau$  jets that mistakenly is identified as a  $b$ -jet by the algorithm is called the *mis-tag efficiency*, which, depending on the jet flavour  $j$  is defined as:

$$\epsilon_j = \frac{N(\text{true } j\text{-jet} \mid \text{identified as } b\text{-jet})}{N(\text{true } j\text{-jet})} \quad (3.21)$$

For most  $b$ -tagging algorithms the mis-tag rates are very small, especially for light jets, therefore the inverse of the mis-tag efficiency, called *rejection*, is commonly used as a performance indicator:

$$r_j = \frac{1}{\epsilon_j}. \quad (3.22)$$

To compare the discrimination power of two  $b$ -tagging algorithms it is common to study the light- or charm jet rejection curves as a function of the  $b$ -tagging efficiency. This curve is also called *Receiver operation characteristic*, or *ROC*. A  $b$ -tagging algorithm that yields a higher rejection at a given  $b$ -tagging efficiency is the more performant one. In general it is a matter of choice for each analysis whether high efficiency or a high purity are more important. Depending on this consideration, different operating points of the  $b$ -tagging algorithm are chosen. For the *ATLAS* experiment, operating points are defined by their  $b$ -tagging efficiency in a standard model  $t\bar{t}$  sample. The recommended operating points, also called *working points (WP)* correspond to 60%, 70%, 77% and 80%  $b$ -tagging efficiency.

The  $b$ -tagging performance has a strong dependence on the jet transverse momentum. Since the pseudo-rapidity is correlated with the transverse momentum, the performance also varies as a function of  $\eta$ . Figure 3.15 shows the behaviour of the  $b$ -tagging efficiency for varying jet  $p_T$  for the MV2c20 algorithm:

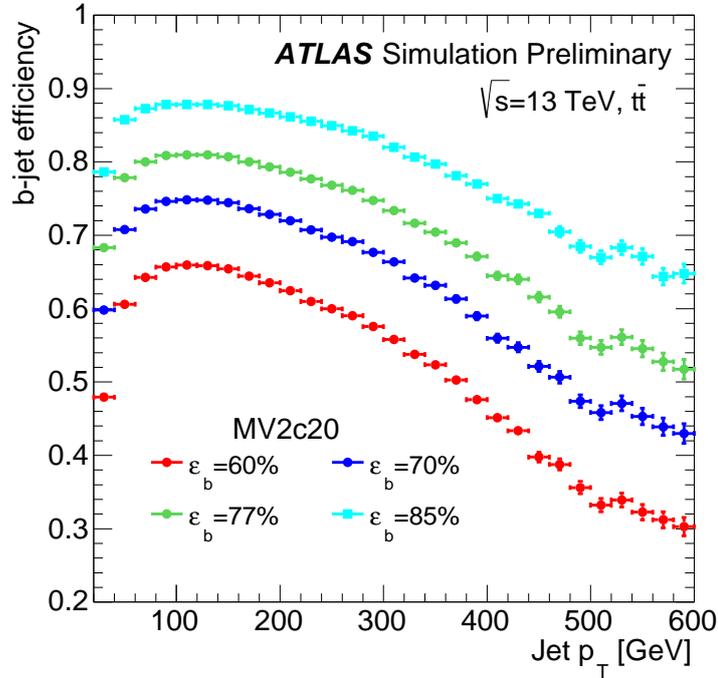


Figure 3.15: Dependence of the  $b$ -tagging efficiency of the MV2c20  $b$ -tagging algorithm on the jet  $p_T$  for different working points. [66]

Figure 3.15 shows that for all working points the  $b$ -tagging efficiency decreases at low jet  $p_T$  as well as at high jet  $p_T$ . At low  $p_T$  the cross-section for multiple scattering and hadronic material interactions is enhanced. Multiple scattering can lead to a less precise measurement of the track parameters. Tracks from material interactions can create fake SV. Furthermore the distance between PV and SV is lower, making it more difficult to discriminate between  $b$ -jets,  $c$ -jets and light jets. At high jet  $p_T$  the track multiplicity from fragmentation is increased therefore the fraction of tracks from the SV with respect to all tracks in the jet decreases. This makes the reconstruction of a SV more difficult and leads to a decrease in the SV resolution. At high  $p_T$ , tracks become very collimated which can lead to tracks sharing several hits in the inner detector. This decreases the precision on the track parameter measurements. The same decrease in measurement precision happens if a  $B$ -hadron decays beyond the first inner detector layers.

All effects lead to a lower SV reconstruction efficiency at high jet  $p_T$  as well as a higher rate of fake SV, and therefore a lower  $b$ -tagging efficiency [61].

### 3.5.8 Track jet $b$ -tagging

A close correspondence between the jet axis and the  $B$ -hadron flight direction is important for  $b$ -tagging, since the jet axis is used as a seed direction for the SV finding as well as a reference axis for the determination of the lifetime sign of the track impact parameter. It was shown that for jets reconstructed from inner detector tracks, instead of the standard calorimeter jets, the jet axis is closer to the  $B$ -hadron flight direction [67]. It was found that using track jets for  $b$ -tagging leads to a good  $b$ -tagging performance especially for high- $p_T$  jets and in dense environments. This is due to the excellent angular resolution of tracks in the inner detector. Track jets with small radius ( $R = 0.2$ ) can also be used to resolve the sub-structure of jets in boosted topologies, where  $R = 0.4$  calorimeter jets already begin to merge. The procedure where  $b$ -tagging algorithms are applied to two track jets within a jet is called *double- $b$ -tagging*, and will be discussed in more detail in section 3.6.2.

### 3.5.9 Flavour tagging calibration

All flavour tagging algorithms used by the *ATLAS* collaboration are optimized and their performance is evaluated using simulated events. Differences exist between data and simulation due to imperfect modelling of properties of tracks in jets that are used by the  $b$ -tagging algorithms. This leads to differences in the performance of the  $b$ -tagging algorithms in data and MC. This difference must be understood precisely and corrected by measuring efficiencies and rejections of a  $b$ -tagging algorithm in data. Based on these measurements, a correction factor is applied to all simulated samples in analyses using  $b$ -tagging. This correction procedure is called  *$b$ -tagging calibration*. The correction factors, also called  *$b$ -tagging scale factors (SF)*, are defined as:

$$\kappa = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}}. \quad (3.23)$$

Scale factors are determined not only for  $b$ -tagging efficiencies but also for  $c$ -jet and light jet mis-identification efficiencies. The standard approach for calibrating the  $b$ -tagging efficiency in *ATLAS* in Run-2 is using a combinatorial likelihood approach in di-leptonic  $t\bar{t}$  events [68]. For the calibration of  $c$ -jets, events containing a  $W$  boson in association with a single charm jet are selected, using a soft muon to identify a semi-leptonic charm decay [61]. The light jet mis-identification efficiency, or *mis-tag rate* is determined in di-jet events using the so-called *negative tag* method, which exploits the fact that for tracks within light jets the distribution of the transverse IP significance, or 3-dimensional flight direction is approximately symmetric around zero [61]. Each of the SFs is quoted with statistical and systematic uncertainties whereas these uncertainties are decomposed using an eigenvector decomposition into 3 components for  $b$ -jets, 3 for  $c$ -jets, and 5 for light jets. For the *ATLAS*  $b$ -tagging calibration for individual jets, the SFs for the  $b$ -tagging efficiency are mostly compatible with one, whereas the SFs for the light mis-tag rate can reach values up to a factor of two. The standard  $b$ -tagging calibration methods for isolated jets within the *ATLAS* experiment provide data-based  $b$ -tagging efficiency SFs as a function of jet  $p_T$  up to values of  $p_T = 250$  GeV for track jets <sup>2</sup>. For high- $p_T$  jets above this threshold the SF is assumed to have the same value as the SF in the highest bin for which a data-based calibration is available, with an additional extrapolation uncertainty derived from simulation.

<sup>2</sup> For calorimeter jets, data-based  $b$ -tagging efficiency SFs are available up to  $p_T = 300$  GeV.

Calibration SFs in *ATLAS* are derived separately for track jets. However, for track jets only the  $b$ -jet efficiency SFs were determined in data. For the  $c$ - and light-jet mis-identification efficiency calibration, the track jet SFs were determined for calorimeter jets and extrapolated to track jets using simulated events.

## 3.6 Higgs boson tagging

The study of boosted Higgs bosons is of interest for many physics analyses in high-energy physics. The production of high  $p_T$  Higgs bosons can happen in the decays of hypothetical new heavy resonances, but also the SM Higgs production at high transverse momentum is interesting, because the relative contribution of other SM backgrounds decreases at high  $p_T$ . Since the Higgs boson decay into two  $b$ -quarks has the highest SM branching fraction it is useful to consider this final state in boosted analyses. In a boosted  $H \rightarrow b\bar{b}$  decay, the fragmentation products of the two  $b$ -quarks become very collimated, making this final state challenging to reconstruct. To meet this challenge, special techniques to identify boosted  $H \rightarrow b\bar{b}$  decays have been developed within the *ATLAS* collaboration, also called *Higgs tagging* [53].

### 3.6.1 Properties of boosted $H \rightarrow b\bar{b}$ decays

If a Higgs boson is produced with a transverse momentum  $p_T^H$  and decays via the  $H \rightarrow b\bar{b}$  process, the angular separation between the  $b$ -quarks approximately follows the relation:

$$\Delta R(b, b) \approx \frac{2m_H}{p_T^H}. \quad (3.24)$$

This means that for a Higgs boson transverse momentum of  $p_T^H > 320$  GeV the angular separation would be smaller than 0.8. The consequence would be that, when reconstructing the two  $b$ -jets with standard  $R = 0.4$  calorimeter jets, there would be regions where the jet cones start to overlap and the association of decay products to a jet becomes ambiguous. Among other effects this can lead to distortions of the jet axis, which in turn causes the degradation of the  $b$ -tagging performance and a loss of signal efficiency. To mitigate this, a large- $R$  calorimeter jet with  $R = 1.0$  is used to reconstruct the entire  $H \rightarrow b\bar{b}$  decay within a single jet.

Because of the larger jet area, large- $R$  jets are more likely to pick up unwanted contributions from pile-up processes. To counteract this, low-energy contributions are removed using jet grooming algorithms. The standard grooming algorithm in *ATLAS* is jet trimming. For more detail on jet trimming, see 3.4.2.

A large- $R$  jet from a  $H \rightarrow b\bar{b}$  decay contains two correlated energy depositions. Common backgrounds include jets from strong interaction quark or gluon production or hadronic top decays. Figure 3.16 shows an illustration of the substructure of jets produced by the different processes.

Jets from QCD production of a quark or a gluon mostly have a single collimated energy deposition within the jet, whereas jets from boosted hadronic top quark decays have a three-prong substructure with one  $b$ -jet and two light-flavour jets from the  $W$  boson decay. The aim of a Higgs tagging algorithm is to discriminate between jets from  $H \rightarrow b\bar{b}$  and jets from QCD and top decays using the known mass of the Higgs boson, as well as flavour tagging information and the substructure of the jet.

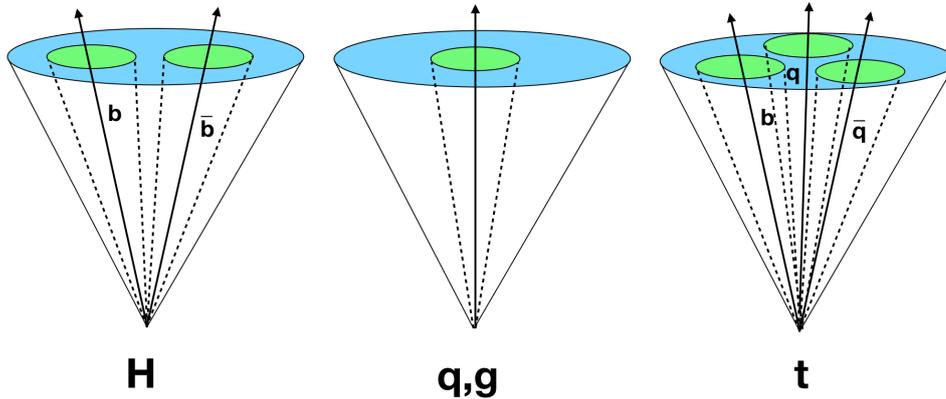


Figure 3.16: Illustration of the substructure of jets produced by decays of different boosted SM particles.

### 3.6.2 Standard Higgs tagging

The standard Higgs tagging procedure in *ATLAS* refers to an algorithm where the discrimination between jets from  $H \rightarrow b\bar{b}$  decays and jets of other origin is made based on a series of cut selections [53]. The input object for Higgs tagging in *ATLAS* is always a trimmed large- $R$  jet candidate, with  $p_T > 250$  GeV and  $|\eta| < 2.0$ . The variables used for the discrimination are the combined jet mass, information about the flavour content as well as substructure variables. Different sequences of cuts can be combined into a tagging algorithm with certain characteristics.

#### Jet mass

The mass of the Higgs boson has been measured to be  $m_H = (125.18 \pm 0.16)$  GeV [3]. Therefore the jet mass is a good discrimination variable between jets from  $H \rightarrow b\bar{b}$  and jets from QCD or top decays. The combined jet mass (see 3.4.3) is used, because it provides optimal mass resolution over the full mass range. A mass window around the Higgs boson mass is defined, in which 95% (*loose*) or 68% (*tight*) of the signal events are contained. Since the mass resolution varies with large- $R$   $p_T$ , the mass window boundaries are derived as a function of the jet  $p_T$ .

#### Track jet $b$ -tagging

large- $R$  jet candidates from the  $H \rightarrow b\bar{b}$  decay contain the fragmentation products of two  $b$ -quarks. To determine the flavour content of the large- $R$  jet, small- $R$  track jets are associated to the large- $R$  jet via ghost-association (see 3.4.5). It is required that the large- $R$  jet has at least two such jets associated. The MV2c10 algorithm is applied to the track jets with different configurations:

- **Double- $b$ -tagging:** Both track jets associated to a jet are required to pass a threshold on the MV2c10 discriminant. The threshold value can be varied according to the desired Higgs efficiency working point. The default choice is to use the MV2c10 70%  $b$ -tagging efficiency working point.
- **Single  $b$ -tagging:** At least one jet among the leading- and subleading- $p_T$  track jets associated to the large- $R$  jet must pass the threshold on the MV2c10 discriminant.
- **Leading track jet  $b$ -tagging:** The leading-  $p_T$  track jets associated to the large- $R$  jet must pass the threshold on the MV2c10 discriminant.

- **Asymmetric track jet  $b$ -tagging:** One jet among the leading- and subleading- $p_T$  track jets must pass a threshold on the MV2c10 discriminant corresponding to the 70%  $b$ -tagging working point, the other must pass a configurable threshold, which can be different from the 70% working point.

### Jet substructure

A great variety of jet variables is available for discrimination of jets with different substructure patterns. Detailed studies regarding the performance of different combinations were conducted in *ATLAS* [53]. Among the most promising variables for Higgs tagging are the variables  $\tau_{21}^{wta}$  and  $D_2^{\beta=1}$ . For a more detailed description of jet substructure and  $\tau_{21}^{wta}$  and  $D_2^{\beta=1}$ , see 3.4.4 .

### 3.6.3 Higgs tagging performance

Similar to  $b$ -tagging, the Higgs tagging performance is assessed by considering both the ability to identify true Higgs jets as well as the ability to reject jets produced by top quarks or from QCD multijet processes. In the same way as for flavour tagging, the true origin of a jet must be defined for simulated samples, using the MC truth record. A true Higgs boson jet (*Higgs jet*) is defined as a jet with a truth Higgs boson and two truth  $B$ -hadrons from the Higgs boson decay within a cone of  $\Delta R = 1.0$  around the large- $R$  jet axis. True top jets are required to have a top quark from the truth record within  $\Delta R = 1.0$  of the large- $R$  jet axis. For QCD jets there are no truth requirements.

The Higgs efficiency is defined as:

$$\epsilon_H = \frac{N(\text{true Higgs - jet} \mid \text{identified as Higgs-jet})}{N(\text{true Higgs jet})}. \quad (3.25)$$

The background mis-identification efficiency is defined analogously to flavour tagging:

$$\epsilon_{\text{top,QCD}} = \frac{N(\text{true top/QCD - jet} \mid \text{identified as Higgs jet})}{N(\text{true top/QCD-jet})}. \quad (3.26)$$

The top and QCD jet rejection is defined as:

$$r_{\text{top,QCD}} = \frac{1}{\epsilon_{\text{top,QCD}}}. \quad (3.27)$$

Figure 3.17 shows the Higgs jet efficiency for double- $b$ -tagging using the 70% working point as a function of the Higgs jet  $p_T$ .

As can be seen in figure 3.17 the Higgs jet efficiency drops at high jet  $p_T$ . This is due to the decrease in  $b$ -tagging efficiency for high- $p_T$  jets. Finally, for very high  $p_T$  jets the Higgs jet reconstruction efficiency drops due to merging of the  $R = 0.2$  track jets. This feature is studied in detail in [69].

Figure 3.18 shows the multijet and top jet rejection for a loose mass window cut in combination with double- $b$ -tagging at the 70% working point as a function of jet  $p_T$ . In general the multijet rejection of the Higgs tagger is about a factor of two larger than the top-rejection. This is mainly because the top jet contains one real  $b$ -jet which makes the distinction between boosted  $H \rightarrow b\bar{b}$  and a boosted hadronically decaying top quark more difficult than for a QCD jet. Also shown is the impact of systematic uncertainties on the QCD and top rejections. For the multijet rejection, the  $b$ -tagging calibration uncertainty is the

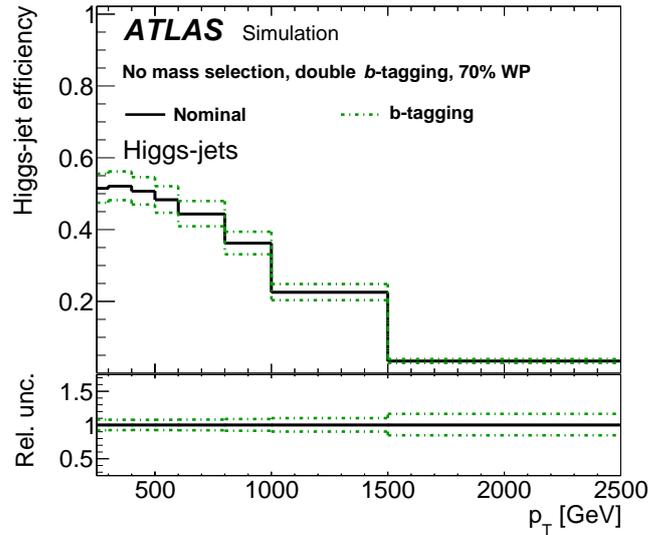


Figure 3.17: Higgs jet efficiency as a function of jet  $p_T$  for double- $b$ -tagging with MV2c10 at 70% working point. From [53].

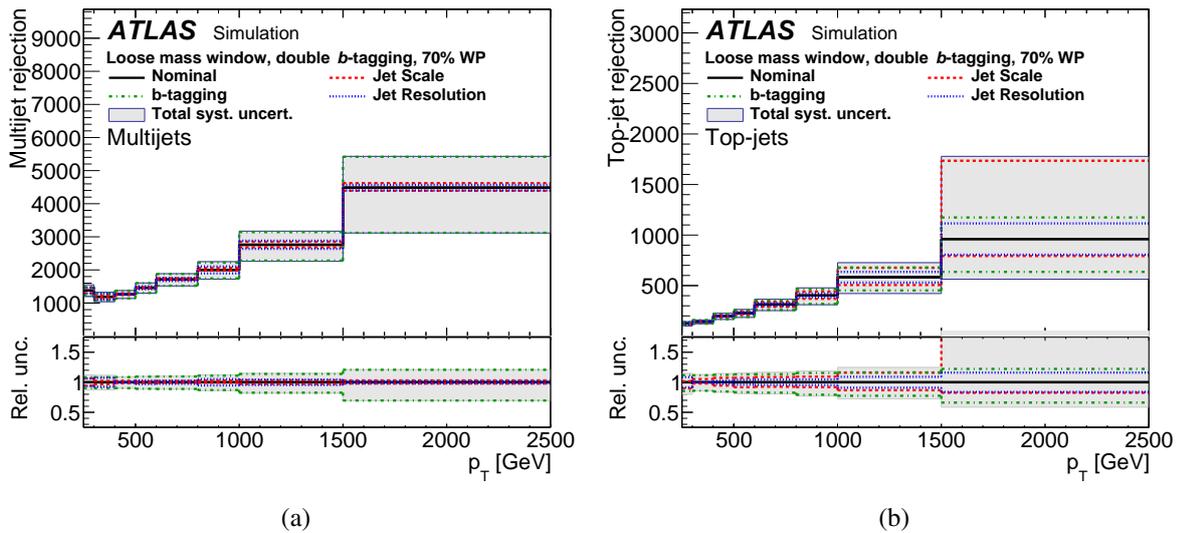


Figure 3.18: Multi-jet rejection (a) and top jet rejection (b) for a loose mass window cut in combination with double- $b$ -tagging at the 70% efficiency working point as a function of jet  $p_T$ . Also shown is the impact of systematic uncertainties from  $b$ -tagging as well as large- $R$  jet energy and mass scale and resolution uncertainties. From [53].

dominant one. For the top rejection, the  $b$ -tagging, large- $R$  jet energy and mass scale and resolution uncertainties each have a similar impact on the rejection.

### 3.6.4 Advanced Higgs tagging

To counteract the efficiency loss due to merging of fixed-radius track jets (see figure 3.17), more advanced Higgs tagging methods include the use of three different sub-jet reconstruction techniques which aim to

be more flexible at high track jet  $p_T$  [69]. They are briefly explained in the following:

### Variable-radius track jets

The reconstruction cone radius of variable-radius (VR) track jet shrinks as a function of track jet  $p_T$ , following the functional form:

$$R_{\text{eff}}(p_T) = \frac{\rho}{p_T}. \quad (3.28)$$

In addition to the parameter  $\rho$ , which determines how quickly the effective radius  $R_{\text{eff}}$  decreases as a function of  $p_T$ , two other parameters are needed, namely  $R_{\text{min}}$  and  $R_{\text{max}}$  which define a minimal and maximal radius, to prevent the VR-track jets from getting too large or too small. A typical parameter set for VR track jets used for  $b$ -tagging in the *ATLAS* experiment is  $R_{\text{max}} = 0.4$ ,  $R_{\text{min}} = 0.02$  and  $\rho = 30$  GeV.

### Exclusive- $k_T$ track jets

The exclusive  $k_T$  algorithm is a modification of the standard inclusive  $k_T$  algorithm (see section 3.4.1). The difference to the standard algorithm is that the exclusive  $k_T$  algorithm will not stop clustering until either all  $d_{ij}$  (defined in eq. 3.12) are above a certain value or a fixed number of sub-jets is achieved. For Higgs tagging the exclusive  $k_T$  algorithm is used to reconstruct exactly two jets from tracks ghost-associated to the large- $R$  jet. The advantage of the exclusive  $k_T$  algorithm is that it has no dependence on an angular scale (such as the jet-radius), therefore it is expected to be more robust at high jet  $p_T$ . Exclusive  $k_T$  with the requirement of exactly two sub-jets is equivalent to reverting the last clustering step of an existing  $k_T$  jet into two sub-regions which each are expected to contain remnants of one  $B$ -hadron (keeping in mind that the  $k_T$  algorithm clusters the most highly energetic object last), which makes the approach promising for Higgs tagging.

### Exclusive center-of-mass subjets

As the name says, center-of-mass (CoM) subjet reconstruction relies on information about the boosted jet system in its center-of-mass system. In the first step, the Higgs candidate jet, including the tracks associated to the jet are transformed via a Lorentz-boost back to the Higgs candidate rest frame. In the rest frame, an angular measure  $y_{ij}$  is defined as

$$y_{ij} = 2 \cdot (1 - \cos(\theta_{ij})), \quad (3.29)$$

where  $\theta_{ij}$  is the angle between the momenta of four-vectors  $i$  and  $j$ . Firstly, in the rest frame, subjets within the Higgs jet candidate are constructed using the EECambridge jet algorithm [70] with the angular measure  $y_{ij}$ . The algorithm is stopped, when exactly two subjets are found. Afterwards tracks  $i$  are associated to a subjet when they are within an angular distance  $y_{i,\text{sub}} < y_{\text{cut}}$  to one of the subjets. The track is always associated to the nearest subjet if the cut criterion results in an ambiguity. After the subjet reconstruction and track-to-subjet association are finished, the  $H \rightarrow b\bar{b}$  system is transformed back to the laboratory frame. The tracks associated to the subjets clustered in the CM frame are used for  $b$ -tagging.

Figure 3.19 shows the QCD- and top-jet rejection for a fixed Higgs-jet efficiency of 50% for the three different advanced subjet reconstruction techniques described above, comparing to the performance of the standard Higgs tagging algorithm using fixed-radius  $R = 0.2$  track jets. All three advanced subjet reconstruction techniques outperform the standard fixed-radius jets at high jet  $p_T$ . In the  $p_T$  range from

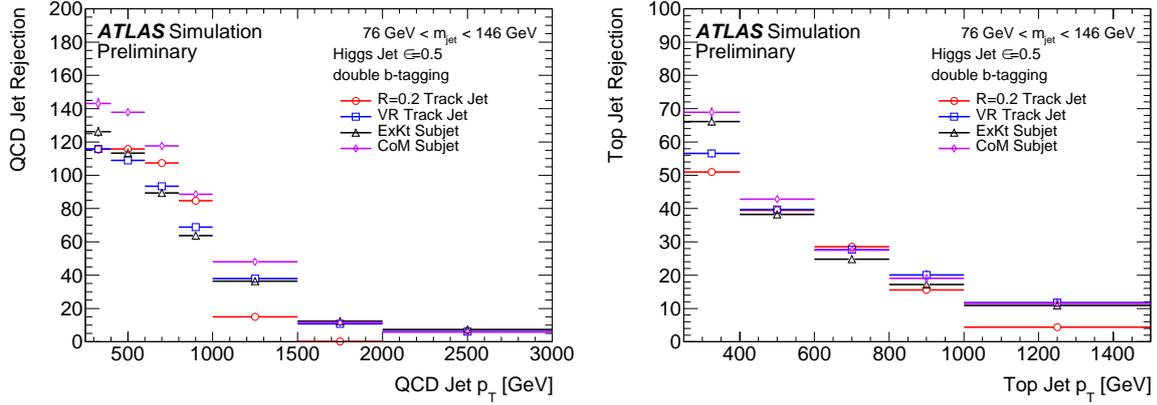


Figure 3.19: QCD jet (left) and Top jet (right) rejection at fixed Higgs efficiency of 50% as a function of  $p_T$  for different subjet reconstruction techniques. [69]

250 GeV to 1.5 TeV the best QCD rejection is achieved by the CoM subjet tagging, as well as the highest Top jet rejection for the range  $250 \text{ GeV} < p_T < 600 \text{ GeV}$ . Despite it having the best performance, the CoM tagging as well as the exclusive  $k_T$  explicitly rely on the double- $b$ -jet topology, and therefore cannot be calibrated in the standard approach using SM  $t\bar{t}$  events. Therefore the most recent recommendation (for full Run-2 analyses) in ATLAS is to use VR track jets for  $b$ -tagging. Chapter 5 of this thesis shows a prototype for a  $R = 0.2$  track jet double- $b$ -tagging calibration analysis using  $g \rightarrow b\bar{b}$  events. The calibration strategy presented in 5 is universally applicable for all three subjet reconstruction techniques presented in this section.

## 3.7 Statistical methods

At the heart of every physics experiment is the process of inscription, meaning the creation of human-digestible information from observation. At a collider experiment in particle physics, observations typically consist of high rates of data events which are selected following certain criteria, recorded and transcribed into various distributions of count rates. However, the process of inscription of experimental data does not end at this stage. Higher-level physical quantities, such as production cross sections can be derived from the count-rate distributions, or they can even be used to formally create scientific language constructs such as the observation or the exclusion of a previously formulated hypothesis. Statistical methods provide a powerful mathematical framework commonly used for this process of data inscription in physics. In the following chapter, the basic concepts of statistical data analysis methods used in this thesis will be introduced, namely the construction of Likelihood models, incorporation and treatment of systematic uncertainties and, finally, the role of hypothesis tests in searches for new physics. For a comprehensive introduction to statistical methods in high energy physics, see e.g. [71], upon which the introductory sections of this chapter are based.

### 3.7.1 Basic concepts

#### Inference

Given a theoretical model, with model parameters  $\theta = (\theta_0, \dots, \theta_n)$  and a measurement result  $x$ , the process of inference is to obtain information from the observed measurement result  $x$  about the model parameters  $\theta$ , or, in more simple words, to get from data to the theory. An example from the context of this thesis is the inference of the rate of events with a certain jet flavour ( $\theta = (\theta_0)$ ), given a measured transverse impact parameter significance of tracks within the jet ( $x$ ) in these events.

#### The Likelihood

The probability to obtain an outcome  $x$  in a measurement, given the model parameters  $\theta$  is commonly denoted as the Likelihood  $L(x; \theta)$ . For practical reasons one usually considers the logarithm of the Likelihood, because of in case of a set of independent measurement outcomes  $\mathbf{x} = (x_0, \dots, x_n)$  it has an additive behavior:  $\ln L(\mathbf{x}; \theta) = \sum_i \ln l(x_i; \theta)$ , where  $l(x_i; \theta)$  denotes the probability to obtain the outcome of an individual measurement  $x_i$ , given the model parameters  $\theta$ .

#### Measurements as a Bayesian or Frequentist?

Given a measurement outcome  $x$  that deviates from the true value  $\mu$ , which is considered to be Gaussian-distributed with a known standard deviation  $\sigma$ , the result is commonly quoted as:

$$x \pm \sigma.$$

A good example is the current value quoted for the Higgs boson mass  $m_H = (125.18 \pm 0.16)$  GeV. As a *frequentist*, the interpretation of this value is the following: "The statement that  $m_H$  is in the interval [125.02, 125.34] GeV has a probability of 68% to be true", or, to reformulate: "The statement that  $m_H$  is in the interval [125.02, 125.34] GeV is true at a *level of confidence* of 68%". The frequentist interpretation always requires a compromise between the accuracy of the statement and the level of confidence in it (e.g. for the Higgs boson mass a more conservative option would have been to say that at 95% confidence level

$m_H$  is in the interval [124.86, 125.5] GeV, which would be less accurate). In other cases of problems, one-sided upper or lower limits on a parameter can be a more appropriate choice, as described in 3.7.5. For the *frequentist* the above statements only make sense within the context of the hypothetical ensemble of all independent experiments and their inferred statements, among which 68% (or whichever level of confidence is chosen) are true. In practice however this ensemble is often hard to achieve, or even to imagine, especially at the LHC, since it is impossible to e.g. repeat LHC Run-2.

For the *Bayesian* statistician the interpretation is somewhat more simple. Bayes' theorem states that after a measurement  $x$  is obtained the *probability density function (PDF)* describing the observers prior belief in the value of  $\theta$ , denoted as  $\pi(\theta)$  is updated according to:

$$f(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{f(x)}, \quad (3.30)$$

where  $f(\theta|x)$  is called the *posterior belief*, and  $f(x) = \int f(x|\theta)\pi(\theta) d\theta$ . Assuming a uniform prior it follows that  $f(\theta|x) \propto f(x|\theta)$ .

Returning to the example of the Higgs boson mass measurement, the Bayesian interpretation states that, assuming a prior that is flat in a given quantity  $\theta$ , the Likelihood is equivalent to the posterior probability and follows a Gaussian PDF:  $f(\mu|x) = L(x; \mu) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$ . For the Bayesian statistician, quoting the Higgs boson mass as  $m_H = (125.18 \pm 0.16)$  GeV therefore means that the possible outcomes of a mass measurement are believed to be Gaussian-distributed, with a width of  $\sigma = 0.16$  GeV around the most probable measurement outcome of  $\mu = 125.18$  GeV. With the Bayesian forward definition of updating the observers belief, it is simple to incorporate new results, simply replacing the prior of the new measurement by the posterior PDF resulting from the preceding experiment.

It is important to note that even though *frequentist* and *Bayesian* approach to statistical analysis tend to be presented as opposing and incommensurable, in practical high energy physics both are used alternately and complementarily. The choice of approach is made depending on which one of the two approaches is the most appropriate for a particular problem at hand. The statistical tools used by the *ATLAS* collaboration mostly tend towards frequentist methods, which are also used in this thesis.

### 3.7.2 Parameter estimation: The Maximum Likelihood method

The task of estimating parameters from observed data (*fitting*) is ubiquitous in experimental high-energy physics, e.g. for the reconstruction of tracks, for calibration analyses (see chapter 5 of this thesis) up to searching for new particles. As already mentioned, the estimation of a parameter consists of the determination of the so called *point estimate*, the best approximation of a parameter to its true value, and its *uncertainty*, in a frequentist approach typically given as confidence intervals (see 3.7.1). The maximum likelihood (ML) method is a common frequentist approach to parameter estimation. For a set of  $N$  statistically independent quantities measured in data  $\mathbf{x} = (x_1, \dots, x_N)$ , which each are described by a PDF  $f(x; \theta)$  with parameters  $\theta = (\theta_1, \dots, \theta_m)$ , the joint PDF is described by the Likelihood:

$$L(\mathbf{x}; \theta) = \prod_{i=1}^N f(x_i; \theta). \quad (3.31)$$

The maximum-likelihood estimate of the parameters  $\theta$ , denoted as  $\hat{\theta}$ , is defined as the set of values for which the Likelihood reaches a global maximum. For practical purposes most often the negative

logarithm of the Likelihood (NLL) is used:

$$-\ln L(\mathbf{x}; \boldsymbol{\theta}) = -\sum_{i=1}^N \ln f(x_i; \boldsymbol{\theta}). \quad (3.32)$$

The use of the NLL is more convenient because it is often easier to search for a global minimum rather than a global maximum, and calculate using sums instead of products of (sometimes rapidly varying)  $f(x_i; \boldsymbol{\theta})$ . The minimization criterion is therefore:

$$-\frac{\partial \ln L(\mathbf{x}; \hat{\boldsymbol{\theta}})}{\partial \theta_j} = 0 \quad \text{for } j = 1, \dots, m. \quad (3.33)$$

The uncertainties of the parameters  $\boldsymbol{\theta}$  are estimated using the contour given by parameter values  $\boldsymbol{\theta}'$  where

$$\Delta \ln L \equiv \ln L(\boldsymbol{\theta}') - \ln L_{\max} = -\frac{s^2}{2}. \quad (3.34)$$

The outer boundaries of this contour for the parameter  $\theta_j$  correspond to the  $s \cdot \sigma$  confidence interval for  $\theta_j$ . Figure 3.20 depicts the 1-dimensional case.

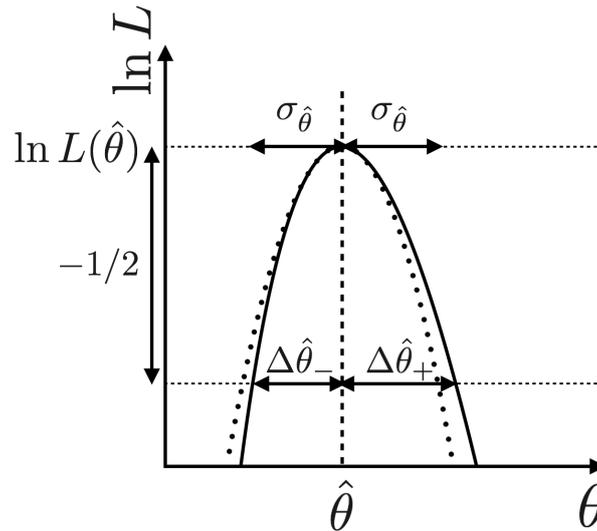


Figure 3.20: Sketch of the logarithm of a likelihood function. The plot shows the likelihood function (solid) and the Gaussian approximation around the ML value (dotted line).

For any choice of PDF  $f(x; \boldsymbol{\theta})$  it can be shown that the Likelihood approaches a multi-dimensional Gaussian distribution for an increasing sample size, indicated for the one-dimensional case as the dotted parabola on the logarithmic scale of figure 3.20. Moreover,  $-2\ln L$  will follow a  $\chi^2$  distribution in this approximation. Already for a finite number of events, the Gaussian approximation, obtained by Taylor-expanding the likelihood function around the maximum, is a good approximation of the true likelihood function. The variance  $\sigma_{\hat{\theta}}$  can be obtained by the inverse of the second derivative of the approximated Gaussian. This is easily visible for the one-dimensional case, because  $\ln L \propto -\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}$  for a Gaussian approximation of  $L$ . For the  $n$ -dimensional case the covariance matrix  $\mathbf{V}(\hat{\boldsymbol{\theta}})$  of the parameter vector is determined:

$$\mathbf{V}(\hat{\boldsymbol{\theta}}) = \left[ \frac{-\partial^2 \ln L(x; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \right]^{-1} \equiv \mathbf{H}^{-1}, \quad (3.35)$$

where  $\mathbf{H}$  is called the *Hesse matrix* which contains the second derivatives of the NLL at its global minimum.

### 3.7.3 Binned Maximum Likelihood fits

When considering statistical problems in high-energy physics at collider experiments usually the considered data samples are very large. It is therefore useful to bin the data to make the calculation of the likelihood function more efficient. Binned maximum likelihood fits are appropriate for problems where the parameters affect both the normalization and the shape of the fit function, and where one is interested in the total rate of a process. In chapter 5 of this thesis a calibration analysis is presented, where the rates of different jet flavour production modes are measured in data using a shape- and normalization-sensitive binned maximum-likelihood fit. A similar binned likelihood fit is performed for the VH resonance search presented in chapter 4.

If one considers a fit problem where the total number of data events  $N$  is fixed, and the number of bins is  $B$ , the Likelihood is:

$$L = \prod_{i=1}^B e^{-\nu_i} \frac{\nu_i^{n_i}}{n_i!}. \quad (3.36)$$

It is clearly visible that the Likelihood in eq. 3.36 corresponds to the product of the Poisson probabilities per bin, to observe  $n_i$  events in bin  $i$ , when  $\nu_i =: \nu_i(\nu, \boldsymbol{\theta})$  events are expected. The variable  $\nu = \sum_i \nu_i$  is the total normalization of the expected events. The NLL is then given as:

$$-\ln L(\mathbf{n}; \nu, \boldsymbol{\theta}) = \sum_{i=1}^B n_i \ln \nu_i(\nu, \boldsymbol{\theta}) - \nu + \text{const.} \quad (3.37)$$

### 3.7.4 The Profile Likelihood Ratio

A common situation arising in high-energy physics is that a likelihood function depends on many parameters, but one is interested only in one specific parameter and its uncertainty. An example for this is a search for a new particle, where a measurement can depend on multiple parameters (e.g. background normalizations and shapes that are known only with a certain precision, energy calibrations etc.), but in the end one is interested only in the signal strength parameter  $\mu$  of the new particle production and its uncertainty. In a new physics search the signal strength is defined as:

$$\mu = \frac{\sigma_{\text{obs}}}{\sigma_{\text{benchmark}}}, \quad (3.38)$$

where  $\sigma_{\text{obs}}$  is the observed cross section and  $\sigma_{\text{benchmark}}$  is the cross section predicted by a certain benchmark model. As can be seen from equation 3.38, the case where  $\mu = 0$  corresponds to the SM-only hypothesis, whereas  $\mu = 1$  corresponds to a scenario where the new particle is produced with the exact cross section as predicted by the benchmark model.

For a likelihood function with a dependence on many parameters  $\theta$ , and one *parameter of interest* (POI)  $\mu$ , one can use the so-called *Profile-Likelihood-Ratio* defined as:

$$\lambda(\mu) = \frac{L(\mathbf{x}; \mu, \hat{\theta})}{L(\mathbf{x}; \hat{\mu}, \hat{\theta})}. \quad (3.39)$$

In the numerator of eq. 3.39 the parameters  $\theta$  are fitted to their maximum-likelihood values  $\hat{\theta}$ , for a given value  $\mu$  of the POI. In practise the parameter  $\mu$  is stepped through and  $\hat{\theta}$  is determined for each value of  $\mu$ . This procedure is commonly called *profiling* of the likelihood function. In the denominator both the parameters  $\theta$  and the POI  $\mu$  are fitted to their maximum likelihood values,  $\hat{\theta}$  and  $\hat{\mu}$  which constitute the global maximum of the likelihood.

Approximating the profile likelihood ratio in the large-event limit works in the exact same way as described in paragraph 3.7.2, and the confidence intervals of the POI  $\mu$  can be evaluated according to equation 3.34, however the evaluation of the profile likelihood ratio is much more complicated, since at every step of  $\mu$  a new maximization of the numerator of eq. 3.39 with respect to the parameters  $\theta$  is required. For more detail on how to find an approximation for the profile-likelihood ratio, see section 3.7.7 and [72].

A new physics search which makes use of the profile likelihood ratio is presented in chapter 4 of this thesis.

### 3.7.5 Hypothesis tests

One important application of statistical inference is the process of decision making on hypotheses, based on a set of measurements. The aim is to investigate how consistent the data is with a certain hypothesis, thereby drawing conclusions regarding the validity of the hypothesis, or finally, to discard it. Hypothesis tests need a clearly-defined, unambiguous hypothesis that is undergoing the test, called the *null hypothesis*  $H_0$ . Often also a complementary *alternative hypothesis*  $H_1$  is defined, that is used to optimize the hypothesis test.

A classical example of a hypothesis test in particle physics is a search for a new particle. In this case the null hypothesis would be that there is no production of new particle, only previously known background processes from the Standard Model. This hypothesis is also called *background-only* hypothesis. The alternative hypothesis used for limit setting is the *signal-plus-background* hypothesis, which states that there is a signal from the production of a new particle on top of the background from the Standard Model.

In frequentist hypothesis testing, the quantity that is used to determine the level of agreement of the hypothesis with the observed data  $\mathbf{x}$  is called the *test statistic*  $t(\mathbf{x})$ . There is some freedom of choice regarding which test statistic to use. A common choice, which is also used in this thesis, is the profile likelihood ratio (eq. 3.39).

To perform a hypothesis test, one needs to define criteria when to reject the null hypothesis based on the test statistic . One therefore defines a *critical region* such that if the test statistic  $t$  lies in this region,  $H_0$  is rejected. If the alternative hypothesis  $H_1$  leads to larger values of  $t$  the critical region can be defined by a cut  $t \geq t_{\text{crit}}$  (see fig 3.21 (a)). With the choice of critical region a compromise has to be

made between Type-I errors ( $H_0$  rejected even though it is true) or Type-II errors ( $H_0$  not rejected even though it is not true). In physics, claiming a false discovery is the more risky behaviour, therefore there are stringent restrictions on Type-I errors.

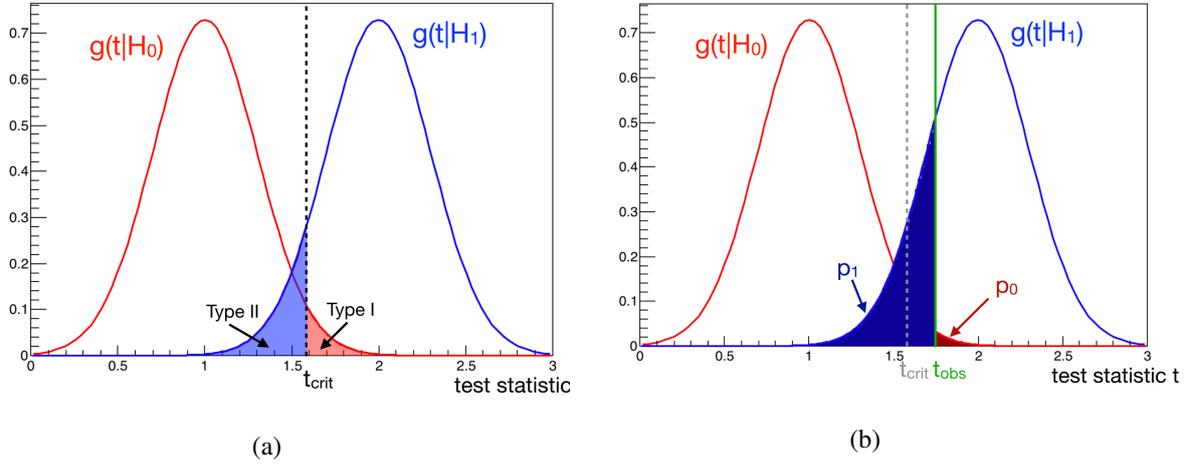


Figure 3.21: Left: Sketch of the statistic probability functions for the null hypothesis (red) and the alternative hypothesis (blue), with the critical region, as well as Type I and II errors. Right: Same as left, but an observed value of the test statistic  $t_{\text{obs}}$  is indicated by the green line, as well as the  $p$ -values  $p_0$  and  $p_1$  by the dark red and dark blue shaded areas.

If finally the test statistic is calculated for observed data ( $t_{\text{obs}}$ ), the  $p$ -value  $p_0$  quantifies how often, if the determination of  $t$  was repeated very often, one would obtain a result as far (or further) removed from the null hypothesis as the result obtained from data, assuming the null hypothesis to be true (see fig 3.21 (b)). One defines  $p$ -values for both  $H_0$  and  $H_1$ :

$$p_0 \equiv \int_{t_{\text{obs}}}^{+\text{inf}} g(t|H_0) dt, \quad (3.40)$$

$$p_1 \equiv \int_{-\text{inf}}^{t_{\text{obs}}} g(t|H_1) dt, \quad (3.41)$$

where  $g(t|H_i)$  is the PDF of the test statistic given the hypothesis  $H_i$ . Often, the  $p$ -value is expressed as the observed significance level  $Z$ , defined as:

$$Z = \Phi^{-1}(1 - p), \quad (3.42)$$

where  $\Phi$  denotes the cumulative distribution of the unit Gaussian. For  $p = 2.87 \cdot 10^{-7}$  we get  $Z = 5$  or  $5\sigma$  significance level. In particle physics the convention is to require a  $5\sigma$  significance level with respect to the null hypothesis to claim a discovery and  $3\sigma$  to announce evidence for new physics.

When there is no excess significant enough to claim evidence, a search can be used instead to constrain parameters of the alternative hypothesis, such as the signal strength  $\mu$  of the new physics process with respect to the Standard Model expectation. This is done by inverting the hypothesis test, namely switching the role of the alternative hypothesis  $H_1$  to the one to be subjected to the test. The parameter of interest

(in the example the signal strength  $\mu$ ) is varied and the corresponding hypothesis is tested, until a value of  $\mu$  is found that results in a value  $CL_s = 0.05$ , where  $CL_s$  is defined as:

$$CL_s = \frac{P_1}{1 - p_0}. \quad (3.43)$$

This means that all  $\mu > \mu_{CL_s=0.05}$  are excluded at the confidence level of 95%. The value of  $\mu_{CL_s=0.05}$  is also called the 95% *CL limit* on the parameter  $\mu$ . To investigate whether the observed data is consistent or in tension with the background-only hypothesis, the observed limit on a parameter is commonly shown in comparison to the *expected limit*. The expected limit is the median limit obtained testing simulation-based datasets generated for the signal-plus-background hypothesis against the background-only hypothesis. The median expected limit is usually shown with green and yellow shaded bands which indicate the range in which 68% (95%) of the results are expected if the limit determination was repeated many times, under the assumption that the background-only hypothesis is true. An illustration of the upper limit distribution from pseudo-experiments and the determination of the median expected limit and the  $1\sigma$  and  $2\sigma$  shaded bands is shown in figure 3.22.

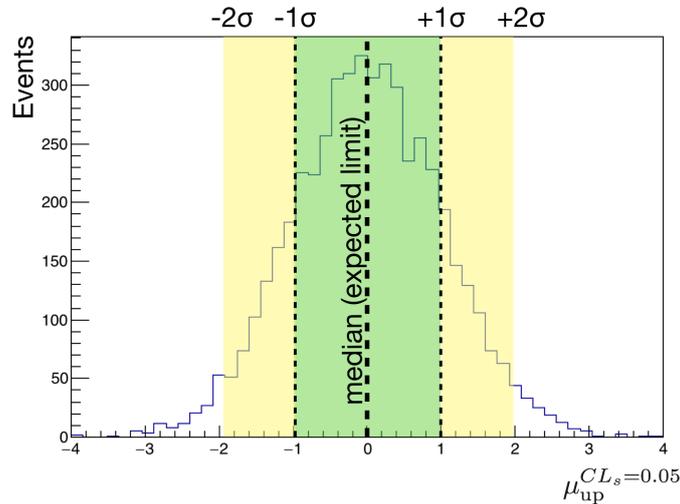


Figure 3.22: Example distribution of the upper limit at  $CL_s = 0.05$  on the signal strength  $\mu$  using ensemble tests. The median indicates the expected limit. The yellow and green bands indicate the  $1\sigma$  and  $2\sigma$  uncertainty bands.

### 3.7.6 Systematic uncertainties: Nuisance parameters

In hypothesis testing the separation between the null hypothesis and an alternative hypothesis depends on the level of overlap between the test statistic of the two hypotheses. If there is less separation between the test statistic distributions for the two hypotheses, the ability to distinguish between them is reduced. Systematic effects in an analysis need to be taken into account in the test statistic, and often cause a larger overlap between hypotheses and reduce the sensitivity of the test. In the following this will be illustrated for the profile likelihood approach, which is applied in chapter 4 of this thesis.

A common approach to include systematic uncertainties in the test statistic is via *nuisance parameters*. As already mentioned in 3.7.4 the profile likelihood ratio contains nuisance parameters  $\theta$  which are taken into account but one is not interested in measuring their values. The parameters  $\theta$  are treated as free when

profiling the likelihood, but they need to be constrained externally e.g. by another measurement. An example is the rate of a background process  $\nu_b$ , which is predicted with insufficient precision by MC, but can be determined by measuring the number of events  $N'$  in a sideband region (*auxiliary measurement*). In this case another Poisson term is multiplied to the likelihood:

$$L(N'; \nu_b) = \frac{(\tau \nu_b)^{N'}}{N'!} e^{-\tau \nu_b}, \quad (3.44)$$

where  $\tau$  is a background scale factor that is assumed to be known. The likelihood ratio including the additional term is profiled with respect to the nuisance parameter  $\nu_b$ .

Similarly it may happen that the scale factor  $\tau$  introduced before is measured with a Gaussian uncertainty  $\sigma_\tau$ . This uncertainty can be taken into account in the test statistic by also profiling  $\tau$  while multiplying the likelihood with a Gaussian constraint term (assuming a measured value  $t_0$ ) of  $1/(\sqrt{2\pi}\sigma_\tau) e^{-(\tau-t_0)^2/(2\sigma_\tau^2)}$ . Often also a logarithmic normal (also called *log-normal*) distribution is used, because it resembles more closely the multiplicative product of many positive random numbers, which is desirable since bin scale factors are not expected to become negative.

### 3.7.7 Approximation of the Test Statistics distribution: Asymptotic formulae

In the general case the distribution of the test statistic  $g(t|H)$  can become very complicated. This is the case especially for the profile likelihood ratio since at every step the minimization procedure has to be performed for the numerator. One possibility is to use numerical methods such as Monte Carlo simulations (see 3.7.8), however with a large number of nuisance parameters this becomes computationally very expensive. An alternative approach is to approximate the distribution of the profile likelihood ratio test statistic using asymptotic formulae [72]. It was already mentioned in 3.7.2 that the approximation of the Likelihood with a Gaussian distribution in the limit of large datasets holds also for the profile likelihood ratio. For a test of the signal strength  $\mu$  in a dataset that is distributed according to  $\mu'$  this is expressed as:

$$-2 \ln \lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + O\left(\frac{1}{\sqrt{N}}\right), \quad (3.45)$$

where  $N$  is the size of the data sample. The above relation is also referred to as *Wilks' theorem* or *Wald approximation* [73, 74]. The best-fit  $\mu$  is Gaussian-distributed with a mean  $\mu'$  and a standard deviation  $\sigma$ . The standard deviation can be determined from the covariance matrix of the estimators of the fit parameters  $\hat{\theta}_i$ , (in the following we set  $\mu = \theta_0$  for ease of notation) which in the large sample limit can be written as the expectation value of the second derivative of the Likelihood, similar to the Hesse matrix in section 3.7.2:

$$V_{ij}^{-1} = -E \left[ \frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j} \right]. \quad (3.46)$$

For a fit with only one degree of freedom the test statistic follows a *non-central  $\chi^2$ -distribution*. For multiple parameters, as is the case for most analyses in particle physics, the determination of the standard deviation is more complicated and is estimated using an artificial dataset, called the *Asimov dataset*.

The Asimov dataset is defined such that if it is used for evaluating the estimators of the parameters  $\theta_i$ , it yields their true values. In case of a typical particle physics analysis this would mean that the calculation

of an estimator for any nuisance parameter, for example the jet energy scale, would yield their nominal value. The Asimov dataset can be used to evaluate the *Asimov likelihood*  $L_A$  and the profile-likelihood ratio  $\lambda_A$ :

$$\lambda_A(\mu) = \frac{L_A(\mathbf{x}; \mu, \hat{\boldsymbol{\theta}})}{L_A(\mathbf{x}; \hat{\mu}, \hat{\boldsymbol{\theta}})} = \frac{L_A(\mathbf{x}; \mu, \hat{\boldsymbol{\theta}})}{L_A(\mathbf{x}; \mu', \boldsymbol{\theta})}. \quad (3.47)$$

The replacement of  $\hat{\mu} = \mu'$  is justified by the fact that in the Asimov dataset, all parameters are equal to their estimator values.

The standard method to determine the standard deviation  $\sigma$  in eq. 3.45 is to determine the second derivatives of  $\ln L_A$  using numerical methods, then inverting the Hesse matrix and extracting  $V_{00}$  which is the standard deviation  $\sigma$  of  $\theta_0 = \mu$ . This means that all free parameters of the approximate distribution of the profile likelihood ratio (eq. 3.45) have been determined. Based on the approximate distribution, the test statistic, its pdf and the cumulative function needed for the determination of the p- and Z-value can be determined, without the need for running many pseudo experiments.

As already stated the asymptotic approximation is only valid on a dataset of sufficient size. The recommendation within the *ATLAS* collaboration is not to use asymptotic formulae when there are zero observed events. The 95% CL observed upper limit should in this case not be smaller than 3 events. The methods explained in this section are available in the RooFit package [75, 76], which is used in the analysis presented in chapter 4

### 3.7.8 Ensemble tests

It is possible to estimate certain properties of a statistical model  $M$  using so-called *ensemble tests* or *pseudo-experiments*. This means repeating a measurement of a certain parameter many times under the same conditions and analyzing the result of this measurement. Ensemble tests are used for the analysis presented in chapter 5 of this thesis.

The ensemble tests begin with the definition of a statistical model. The second step consists of the generation of pseudo-data, e.g. using a random generator to generate numbers according to a Poisson distribution as pseudo-data. The statistical model is evaluated for each of the newly generated pseudo-dataset and, for a large number of pseudo-events, the probability distribution of the parameter in question can be estimated.

For ensemble tests, it is important to show that the estimator of the parameter is consistent and unbiased. One important test quantity is the *linearity* of the estimator: If a pseudo-dataset is generated using the parameter value  $\theta_{gen}$ , the estimator  $\hat{\theta}$  should yield the same value. One can define the so-called *pull*, defined as the difference between  $\theta_{gen}$  and  $\hat{\theta}$ , divided by the standard deviation of  $\hat{\theta}$ . If the mean of the pull distribution deviates from zero, or if the standard deviation is not equal to one, this points to a problem with the estimator.

---

## Search for a new resonance $V' \rightarrow VH$ decaying into final states with leptons and b-jets

---

In this chapter a search for a new heavy resonance in the decay channel  $V' \rightarrow VH \rightarrow ll, l\nu, \nu\nu + b\bar{b}$  is presented. Many scenarios for physics beyond the Standard Model (BSM) predict the existence of new heavy particles at the TeV energy scale, which could be produced at the LHC. Some of the predicted resonances decay preferably into an electroweak vector boson W or Z and a SM Higgs boson. The presence of a new heavy resonance would manifest itself as a localized excess of data events above the expected SM background in the  $m_{VH}$  invariant mass spectrum, therefore the search is performed by scrutinizing for such a localized excess. In this search, a simplified model is used as a benchmark, namely the Heavy Vector Triplet (HVT) model [27, 28]. This model predicts a set of three vector resonances with different bosonic and leptonic coupling strengths.

The presented search analysis is performed blindly, meaning that the optimization of the event selection and the development of the statistical analysis strategy are performed based on simulated data without looking at the observed data in analysis regions where signal events are expected. Only after the full analysis strategy is optimized and defined, the final results using observed data are obtained. The  $VH$  resonance search analysis documented here is published in [77]. In this chapter, special attention will be paid to sections of the analysis where the author of this thesis made a significant contribution, namely the  $\nu\nu b\bar{b}$  channel optimization as well as the combined statistical analysis of all three lepton channels ( $\nu\nu b\bar{b}$ ,  $\ell\nu b\bar{b}$ ,  $\ell\ell b\bar{b}$ ).

### 4.1 Introduction to the Heavy Vector Triplet model

As of today there is no complete and compelling theory of physics beyond the SM that is able to make precise predictions about experimental observables without major assumptions about underlying BSM model parameters. Several general theoretical frameworks (such as Supersymmetry) exist, that are able to make predictions for observable phenomena, however the quantitative interpretation of observed effects can only be made within an explicit implementation of the framework and the translation of experimental results to other theoretical frameworks or other models is difficult.

The *Heavy Vector Triplet (HVT)* model [27, 28] attempts to overcome this difficulty by introducing a simplified, phenomenological approach that makes use of the fact that many searches for heavy new particles are only sensitive to a very limited set of BSM model parameters governing the resonance mass

and its couplings to the SM particles in the decay final state. This insensitivity of the searches is due to the fact that for most analyses it is not possible to resolve the line shape of the resonance peak, but the measured width is dominated by the detector resolution. The simplified model therefore focuses on the prediction of on-shell narrow-width new particles, which, depending on the model parameters, can either be weakly coupled, similarly to theories with an extended Gauge sector [78] or strongly coupled, similar to composite-Higgs models [79]. A brief overview of different classes of New Physics models is given in section 2.5.

The simplified model is implemented using a phenomenological Lagrangian that explicitly contains a narrow-width heavy electroweak-charged spin-1 resonance  $V_\mu^a$ ,  $a = 1, 2, 3$ , which manifests itself as one charged and one neutral vector resonance charge eigenstate, similar to the W and Z bosons in the SM:

$$V^\pm = \frac{V_\mu^1 \mp iV_\mu^2}{\sqrt{2}} \quad V_\mu^0 = V_\mu^3. \quad (4.1)$$

The phenomenological Lagrangian describing the dynamics and interactions of the field  $V^a$  can be written as:

$$\begin{aligned} \mathcal{L}_V = & -\frac{1}{4}[D_\mu V_\nu^a][D^\mu V^{\nu a}] + \frac{m_V}{2}V_\mu^a V^{\mu a} \\ & + ig_V c_H V_\mu^a (H^\dagger \tau^a D^\mu H - D^\mu H^\dagger \tau^a H) + \frac{g^2}{g_V} c_F V_\mu^a \sum_f \bar{f}_L \gamma^\mu \tau^a f_L \\ & + \frac{g_V}{2} c_{VVV} \epsilon_{abc} V_\mu^a V_\nu^b [D^\mu, V^{\nu c}] + g_V^2 c_{VVHH} V_\mu^a V^{\mu a} H^\dagger H - \frac{g}{2} c_{VW} \epsilon_{abc} W^{\mu\nu a} V_\mu^b V_\nu^c. \end{aligned} \quad (4.2)$$

The first line of equation 4.2 contains the dynamics of the resonance as well as the mass term. It should be noted however that due to the mixing with the SM  $W$  field, the charge eigenstates in the Lagrangian are not equivalent to the mass eigenstate, therefore  $m_V$  does not correspond to the physical resonance mass. The coupling  $g_V$  governs the strength of the interactions of the resonance  $V_\mu^a$ , where as the parameters denoted as  $c$  describe coupling modifiers for the individual interaction operators. The coupling  $g$  is the SM  $SU(2)_L$  gauge coupling.

The first term in the second line contains the interaction of the resonance with the Higgs field, with  $\tau^a = \sigma^a/2$  being the Pauli matrices, and the covariant derivative:

$$D_\mu V_\nu^a = \partial_\mu V_\nu^a + g \epsilon^{abc} W_\mu^b V_\nu^c, \quad (4.3)$$

where  $W_\mu^a$  are the SM weak isospin triplet fields with their coupling  $g$ . The second line, after electroweak symmetry breaking, gives rise to the interaction of the new resonance  $V_\mu^a$  with the SM W and Z boson and the physical Higgs field due to the second term of the covariant derivative. The coupling modifier  $c_H$  therefore controls the strength of the interaction of the new V resonance with the SM vector fields and the Higgs field. This coupling strength governs the decays  $W' \rightarrow WH$  and  $Z' \rightarrow ZH$ <sup>1</sup> studied in this search analysis.

The second term in the second line of equation 4.2 contains the interaction of the  $V'$  with the SM left-handed fermions  $f_L$ . This term is important for the analysis presented here, because it describes the tree-level production mechanism of the heavy new resonance via the Drell-Yan (DY) process. The

<sup>1</sup>  $W'$  and  $Z'$  are the mass eigenstates of the HVT resonances, commonly summarized as  $V'$ .

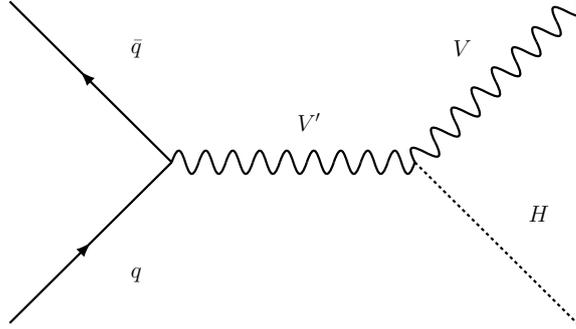


Figure 4.1: Feynman diagram of the Drell-Yan production of a new heavy resonance  $V'$  and its decay into  $V = W, Z$  and a SM Higgs boson.

Feynman diagram for the production and decay of a  $V'$  is shown in figure 4.1. The fermionic couplings are controlled by the parameter  $c_F$ .

The third line of equation 4.2 contains three terms and coupling parameters  $c_{VVV}$ ,  $c_{VVHH}$ ,  $c_{VW}$ . The interaction terms describe triple couplings among the  $V'$  themselves, the coupling of two  $V'$  with a pair of Higgs bosons or a weak  $W$  field. However, since the simplified model is only intended for the description of the direct interaction of one  $V'$  with SM particles these terms are considered to be not relevant for the phenomenological description of physics accessible at the LHC. Specifically, they only affect the interactions described above via the mixing of the  $V'_\mu$  with the  $W_\mu^a$ , but the mixing is a small effect, such that it can be neglected.

In summary, the coupling  $g_V c_H$  is governing the probability of the  $V'$  resonance to decay into SM bosons and the coupling  $g^2 c_F / g_V$  is governing the production and decay via SM fermions. Two different benchmark model types, A and B, for different values of  $g_V$  are used in the presented search analysis. For the benchmark Model A the assumption for the parameters is the following:

$$g_V c_H \simeq g^2 c_F / g_V \simeq g^2 / g_V. \quad (4.4)$$

This means that the decay branching fractions into fermions and bosons are comparable for Model A. This assumption is used in certain models predicting an extended gauge sector [78]. Also, the total decay width of the new resonance decreases for increasing  $g_V$ , because of the suppression due to the factor  $g^2 / g_V$ .

For Model B, the assumption is:

$$g_V c_H \simeq -g_V, \quad g^2 c_F / g_V \simeq g^2 / g_V, \quad (4.5)$$

which means that the coupling term proportional to  $c_H$  is not suppressed for increasing  $g_V$ . This means that for increasing  $g_V$  the decay to di-boson final states becomes dominant. This is common to some strongly coupled models, such as the minimal composite Higgs model [79]. However, also the total width increases with  $g_V$  such that for very large  $g_V$  the assumption of a narrow resonance does no longer hold. Figure 4.2(a) and 4.2(b) show the decay branching fractions for Model A and  $g_V = 1$  and Model B and  $g_V = 3$ . The suppression of fermionic decays for Model B is clearly visible. Figure 4.2(c) shows the ratio of the cross section for DY resonance production mechanism with respect to that of vector boson fusion (VBF) production for different resonance masses and collider energies, as a function of the

coupling modifiers  $c_F$  and  $c_H$ . As expected, for low  $c_F/c_H$  the VBF process dominates, whereas, for higher  $c_F/c_H$  the DY process is the dominant production mode.

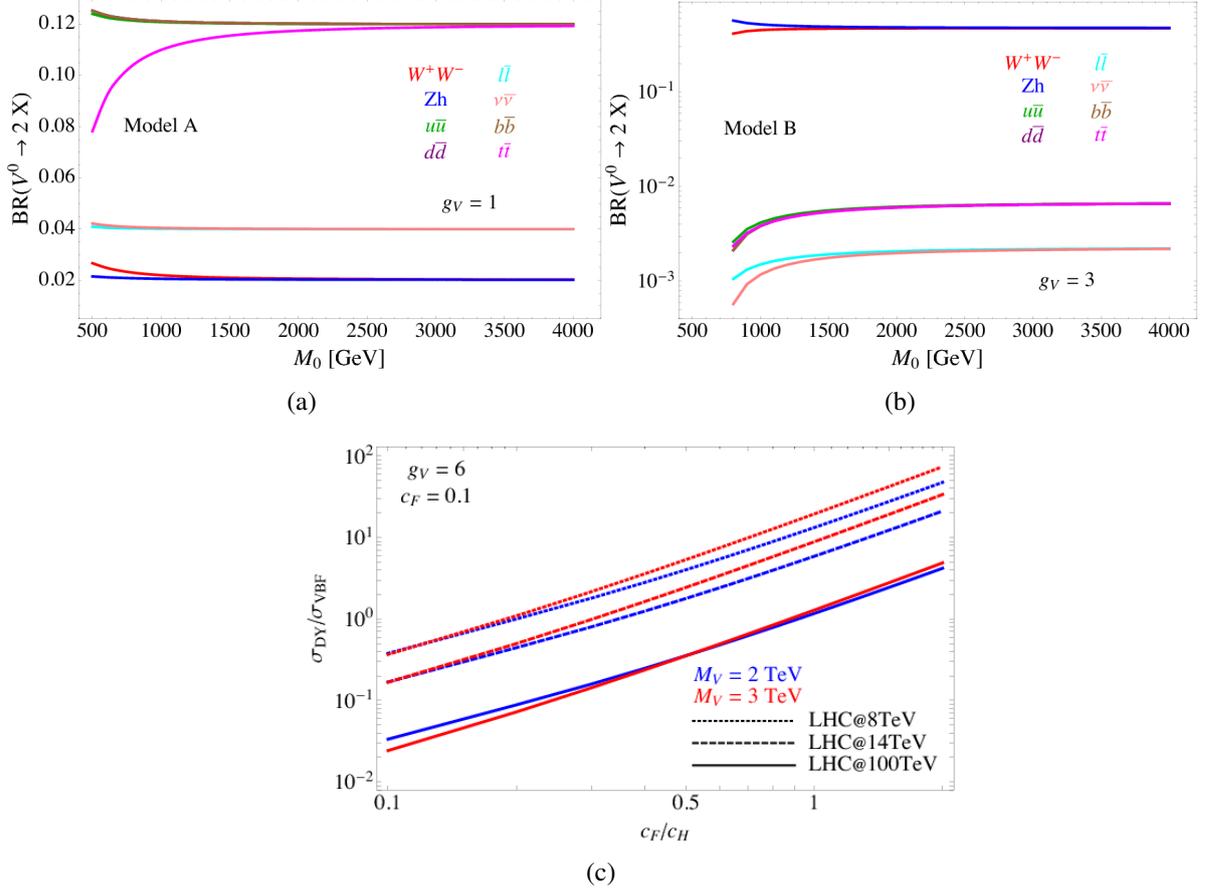


Figure 4.2: (a),(b): Fermionic and bosonic branching fractions of a neutral heavy new vector resonance for the HVT Model  $A_{g_V=1}$  (left) and  $B_{g_V=3}$  (right) as a function of the resonance mass. (c): Cross section ratio for DY and VBF production of a heavy new vector resonance as a function of the ratio of the fermionic and bosonic couplings modifiers, for different resonance masses and accelerator center-of-mass energies. From [27].

In the presented search analysis, Model A with  $g_V = 1$  is used as a benchmark model. In addition the presented results are interpreted in terms of Model B with  $g_V = 3$ . Finally the obtained results are interpreted in the full two-dimensional HVT parameter space  $(g_V c_H, g^2 c_F/g_V)$  in the phase space region where the narrow-width approximation holds.

The presented analysis is designed to search for the HVT  $V'$  in the  $V' \rightarrow VH$  final state with  $H \rightarrow b\bar{b}$  and leptonic  $V$  boson decays. The Higgs boson is assumed to have purely SM-like properties including production cross sections and decay branching fractions. The advantage of studying leptonic vector boson decays is that the presence of high- $p_T$  leptons or large missing transverse momentum allows for efficient event triggering. On the other hand,  $H \rightarrow b\bar{b}$  is the most probable decay of the SM Higgs boson, therefore a good signal acceptance is expected in this particular channel. To make the search analysis as versatile and sensitive as possible in view of an a priori unknown resonance mass, a wide resonance

mass range is covered. The total width of the signal invariant mass spectrum is dominated by the detector resolution, therefore the use of the simplified HVT model is justified. The covered resonance mass regime spans from  $m_{V'} = 500$  GeV to  $m_{V'} = 5$  TeV. The lower mass boundary is chosen due to inability of the HVT model to reproduce the SM parameters for lower resonance mass hypotheses. For resonances above  $m_{V'} = 5$  TeV the expected reconstruction efficiency is expected to be very low due to the extremely high Lorentz-boost of the resonance decay products.

## 4.2 Object selection

### Lepton selection type definition

Electrons are reconstructed by matching inner detector tracks to energy clusters in the electromagnetic calorimeter. Additional identification requirements are imposed on the electron candidates regarding the track quality, the shape of the electromagnetic shower, the matching between track and energy clusters and the isolation. A Likelihood-approach is used to combine the information and the Loose and Tight working points are used in this analysis as defined in [80].

Muons are identified using tracks in the inner detector matched to tracks or track segments from the muon spectrometer. To extend the detector acceptance beyond that of the inner detector, also muons with only a *stand-alone* track in the muon spectrometer are used. Additional selection criteria are imposed on the quality of the inner detector track and the track or track segment from the muon spectrometer. The Loose and Medium working points as defined in [81] are used.

Both the electron and muon candidates must fulfill certain kinematic requirements and are sorted into three different quality categories, called *VH loose*, *ZH signal* and *WH signal*. The *VH loose* category is defined to ensure orthogonality (meaning that there is no overlap of events) between the lepton channels ( $\nu\nu b\bar{b}$ ,  $\ell\nu b\bar{b}$ ,  $\ell\ell b\bar{b}$ ), based on a looser selection of leptons. The *WH signal* and *ZH signal* leptons have more stringent quality requirements to reduce the background contribution from fake leptons. The criteria are summarized in table 4.1.

Lepton flavour	Selection type	$p_T$ [GeV]	$ \eta $	ID	$s_{d_0}$ [mm]	$ z_0 \sin \theta $ [mm]	Isolation	
Electron	VH loose	>7	<2.47	Loose	<5	<0.5	LooseTrackOnly	
	ZH signal	>27	same as VH loose					
	WH signal	>7	<2.47	Tight	same as ZH signal		Tight	
Muon	VH loose	>7	<2.7	Loose	<3	<0.5	LooseTrackOnly	
	ZH signal	>27	<2.5	same as VH loose				
	WH signal	>27	<2.5	Medium	same as ZH signal		Tight	

Table 4.1: Selection criteria for leptons used in the VH resonance analysis. Here,  $s_{d_0}$  denotes the transverse track impact parameter significance defined as  $|d_0|/\sigma_{d_0}$ .  $z_0$  represents the longitudinal track impact parameter and  $\theta$  is the polar angle.

### Tau leptons

A veto on events containing tau leptons is used in the 0-lepton channel to suppress backgrounds from hadronically decaying tau lepton candidates ( $\tau_{\text{had}}$ ). The  $\tau_{\text{had}}$  candidates are reconstructed using small- $R$

calorimeter jets with  $p_T > 20$  GeV that need to lie within the tracking acceptance region of  $|\eta| < 2.5$  and outside of the gap between calorimeter barrel and end-cap region  $1.37 < |\eta| < 1.52$ . The candidates are required to have either one or three associated tracks and must comply with the Medium tau identification criteria. [82]

## Jets

Jets are reconstructed from energy clusters in the calorimeter or from tracks in the inner detector, using the anti- $k_T$  reconstruction algorithm from the FASTJET [48] package. For a brief introduction to jet reconstruction, see also 3.4. Large- $R$  jets are used to reconstruct Higgs candidate in the merged category, and track jets are used to  $b$ -tag the large- $R$  jets. Small- $R$  calorimeter jets are used for the reconstruction of the Higgs candidate in the resolved category as well as for  $b$ -tagging. The different jet types used in this analysis and their selection criteria are summarized in table 4.2.

Jet selection type	Radius	Constituents	Selection criteria
Forward jets	0.4	Calorimeter cluster	$p_T > 30$ GeV $2.5 \leq  \eta  < 4.5$
Signal jets	0.4	Calorimeter cluster	$p_T > 20$ GeV $ \eta  < 2.5$ $JVT \geq 0.59$ if $ \eta  < 2.4$ and $p_T < 60$ GeV
Large- $R$ jets	1.0	Calorimeter cluster	$p_T > 250$ GeV $ \eta  < 2.0$ at least one track jet
Small- $R$ track jets	0.2	Tracks	$p_T > 10$ GeV $ \eta  < 2.5$ at least 2 track constituents

Table 4.2: Selection criteria for Jets used in the VH resonance analysis.

## Muon-in-jet correction

Muons inside the  $b$ -jets stem from semi-leptonic decays of  $B$ -hadrons. Since a muon does not deposit a significant portion of its energy in the calorimeter, this component does not contribute to the reconstructed jet energy. To correct the jet energy, the reconstructed muon four-vector is added to the jet four-vector [53]. For small- $R$  calorimeter jets the closest muon within a cone of  $\Delta R < 0.4$  around the jet axis is added. Muons with  $p_T > 5$  GeV and without requirements on the isolation are considered for the correction. For large- $R$  jets the closest muon within a cone of  $\Delta R < 0.2$  to the axis of one of the associated track jets is used. This association scheme is chosen because the track jets serve as proxies for the  $B$ -hadron flight directions within the Higgs candidate jet. For small- $R$  calorimeter jets an additional correction is used to account for biases due to the presence of neutrinos and jet resolution effects. This so-called *PtReco* correction is based on a jet- $p_T$ -dependent scaling of the reconstructed jet  $p_T$  to truth-level jet  $p_T$ . [14] The corrected jet four-vectors only enter the calculation of the invariant mass  $m_{VH}$  of the reconstructed final state. No *PtReco*-like correction is available for large- $R$  jets, therefore none is applied.

### Missing transverse energy

As described in section 3.2 the missing transverse energy is calculated by imposing energy conservation in the x-y-plane of the detector and calculating the missing component to the energy balance based on all reconstructed physics objects. In addition to so-called *hard terms* in the  $E_T^{\text{miss}}$  calculation from muons, electrons, photons,  $\tau$ -leptons and jets, a track-based soft term is introduced, using tracks from the inner detector that are not associated with the hard terms. Tight quality requirements and association to the primary vertex ensures robustness of the soft term against pile-up tracks. [83]

An additional quantity that is related to the missing transverse energy and which is used in the event selection of the VH resonance analysis is  $H_T^{\text{miss}}$ , defined as:

$$H_T^{\text{miss}} = \frac{E_T^{\text{miss}}}{\sqrt{\sum(p_T^e)^2 + \sum(p_T^\mu)^2 + \sum(p_T^\tau)^2 + \sum(p_T^{\text{jet}})^2}} \quad (4.6)$$

where the denominator is the square root of the scalar sum of the transverse momenta of the hard terms. The scalar  $p_T$  sum is proportional to the  $E_T^{\text{miss}}$  resolution, therefore  $H_T^{\text{miss}}$  is related to the  $E_T^{\text{miss}}$  significance. Placing a cut on the significance  $H_T^{\text{miss}}$  corresponds to applying a tighter selection in events with well-measured  $E_T^{\text{miss}}$  as opposed to a looser selection in events with a larger uncertainty on the  $E_T^{\text{miss}}$ . In the presented analysis, a cut on  $E_T^{\text{miss}}/\sqrt{H_T^{\text{miss}}}$  is used to veto missing transverse energy in the 2-lepton channel.

### Overlap removal

Different physics objects can have similar signatures in the detector, so they can in principle pass reconstruction criteria for more than one type of object. For example, an electron in an environment with additional activity in the detector could also be interpreted as part of a hadronic jet. To avoid duplicate objects, a procedure to resolve these reconstruction ambiguities is imposed, which is called *overlap removal*. The standard procedure recommended in *ATLAS* is as follows:

- **Electrons and Muons:** If an electron and a muon candidate share the same track, the electron candidate is removed, and the muon candidate is kept.
- **Electrons and Jets:** If there is a calorimeter jet candidate within  $\Delta R = 0.2$  of a surviving electron candidate, the jet is removed and the electron is kept. For a surviving jet that fulfils the criteria  $JVT > 0.64$ ,  $p_T < 50$  GeV and  $|\eta| < 2.4$ , electron candidates are removed if they are within  $\Delta R = \min(0.4, 0.04 + 10 \text{ GeV}/p_T^e)$  with respect to the jet axis.
- **Muons and Jets:** If there is a muon candidate within  $\Delta R = 0.2$  of a jet candidate with a low number of tracks carrying a significant portion of the jet momentum <sup>2</sup> the jet is removed. For surviving jets that pass the following criteria:  $JVT > 0.64$ ,  $p_T < 50$  GeV and  $|\eta| < 2.4$  a concurring muon candidate is removed if it is within  $\Delta R = \min(0.4, 0.04 + 10 \text{ GeV}/p_T^\mu)$  with respect to the jet axis.
- **Electrons and large- $R$  Jets:** If there is a large- $R$  jet candidate within  $\Delta R = 1.2$  of an electron candidate, the large- $R$  jet is removed and the electron is kept. No overlap removal between muons and large- $R$  jets is applied.

<sup>2</sup>  $\text{NumTrkPt500PV}^{\text{jet}} < 3$ , or  $(p_T^\mu/\text{SumPtTrkPt500PV}^{\text{jet}} > 0.7$  and  $p_T^{\text{jet}}/p_T^\mu < 2$ ), where  $\text{NumTrkPt500PV}$  and  $\text{SumPtTrkPt500PV}$  are the number and  $p_T$  sum of tracks from the primary vertex with a track  $p_T$  of at least 500 MeV associated to a jet.

### 4.3 Event categorization

As mentioned in the introduction to this chapter, events from three different decay channels are considered in this analysis, which differ in the number of leptons produced in the electroweak vector boson decay. The events are categorized into three different lepton channels (0-, 1-, and 2-lepton channel, for  $\nu\nu b\bar{b}$ ,  $\ell\nu b\bar{b}$ ,  $\ell\ell b\bar{b}$  respectively) based on the number of VH-loose leptons (see 4.2) in the event. The lepton channels are strictly disjoint, by requiring that the 0-lepton channel contains no VH-loose leptons, the 1-lepton channel contains exactly one, and the 2-lepton channel contains exactly two same-flavour VH-loose leptons. A graphical summary of the regions after the full categorization sequence within a lepton channel is shown figure 4.3. The details of the categorization will be discussed in the following.

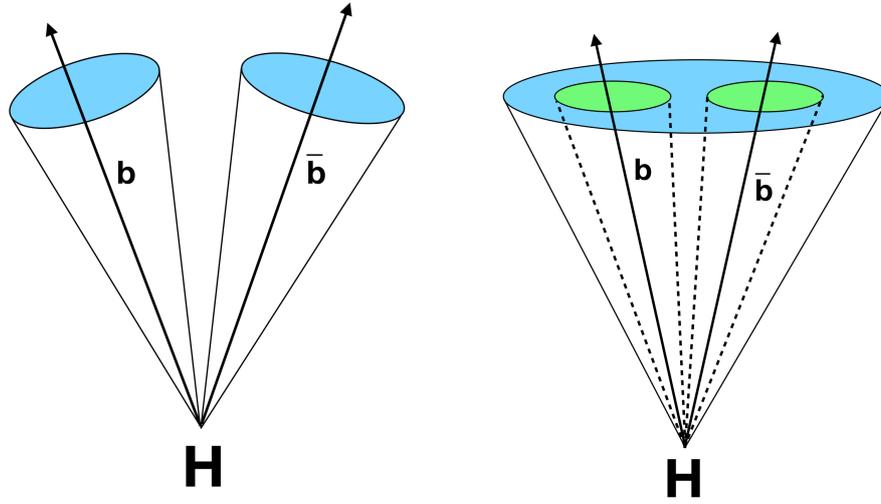
H $\rightarrow$ bb topology	number of b-tagged (Track-) jets	Higgs candidate mass	Top control region?
resolved	1-tag	signal	opposite flavour leptons (2-lepton only)
		sideband	
	2-tag	signal	
		sideband	
merged	1-tag	signal	additional b-tags
		sideband	
	2-tag	signal	
		sideband	

Figure 4.3: Sketch of the analysis regions used in the VH resonance analysis.

Since the  $V'$  search analysis covers a wide range of hypothetical resonance masses, the  $p_T$  range of the resonance decay products is very large. This has an impact especially on the reconstruction of the  $H \rightarrow b\bar{b}$  decay. At low resonance masses the  $p_T$  of the Higgs boson is low, therefore its decay products can be reconstructed as a pair of two  $R = 0.4$  calorimeter jets. This event topology is referred to in the following as the resolved category. In case of a large resonance mass the Higgs boson is produced with a large transverse momentum, therefore the  $H \rightarrow b\bar{b}$  decay products are very collimated. The complete  $H \rightarrow b\bar{b}$  decay is reconstructed within a single large- $R$  jet. This event topology is called the merged category. Figure 4.4 shows a sketch of the resolved and merged  $H \rightarrow b\bar{b}$  decay topologies.

To categorize the events into resolved or merged topology while achieving an optimal analysis sensitivity, the so called *Priority Resolved Signal Region (PRSR)* categorization scheme is adopted, which classifies events according to the classification sequence shown below. The sequence is followed from top to bottom, whereas the next step is invoked only if none of the previous attributes apply to the event. If none of the listed attributes apply, the event is discarded.

1. The event belongs to the resolved signal region.
2. The event belongs to a resolved side-band, or top control region.



(a)  $H \rightarrow b\bar{b}$  decay in resolved topology      (b)  $H \rightarrow b\bar{b}$  decay in merged/boosted topology

Figure 4.4

3. The event belongs the merged signal region.
4. The event belongs to a merged side-band, or top control region.

The categorization in signal regions and background control regions is performed based on the invariant mass of the  $H \rightarrow b\bar{b}$  candidate. In the resolved category this corresponds to the invariant mass of the di-jet system formed by the two signal small- $R$  jets. In the merged category it is the invariant mass of the large- $R$  jet, using the so-called combined mass that uses not only calorimeter information, but also tracking information to improve the mass resolution [84] (see also 3.4.3). A two-sided mass window requirement around the Higgs boson mass  $m_H = 125$  GeV defines the signal region. Events that lie in the sideband-regions outside this range are sorted into the background control region. The lower and upper bound of the signal mass region depend on the topology category and the lepton channel, and are summarized in Table 4.3. In the resolved category, a tighter mass window is chosen, because the mass resolution is better compared to the merged category.

topology	0-lepton	1-lepton	2-lepton
resolved	$110 < m_{jj} < 140 \text{ GeV}/c^2$	$110 < m_{jj} < 140 \text{ GeV}/c^2$	$100 < m_{jj} < 145 \text{ GeV}/c^2$
merged	$75 < m_j < 145 \text{ GeV}/c^2$	$75 < m_j < 145 \text{ GeV}/c^2$	$75 < m_j < 145 \text{ GeV}/c^2$

Table 4.3: Definition of the  $H \rightarrow b\bar{b}$  candidate mass range for signal regions in the different lepton- and topology-based categories.

The Higgs candidate mass windows in the resolved analysis are defined such that an optimal sensitivity is obtained. In the merged regime the cut corresponds to 90% signal efficiency, when using the calorimeter-based mass [85]. The window was not re-optimized for the combined calorimeter mass definition that is used in this analysis, but the difference is expected to be small.

Furthermore, the categorization within each lepton channel and topological region is based on the number of jets identified as  $b$ -jets ( $b$ -tagged jets) in the event. In this analysis the MV2c10 algorithm is used to identify  $b$ -jets at the 70% efficiency working point [61, 64]. In the resolved category, the events are classified, based on the number of  $b$ -tagged small- $R$  jets. If there are two  $b$ -tagged small- $R$  jets the event belongs to the 2-tag category. If only one of the small- $R$  jets is  $b$ -tagged the event is sorted into the 1-tag category. In the merged category the 1-tag and 2-tag categories are defined depending on the number of  $b$ -tagged track jets associated to the large- $R$  jet.

In addition, a categorization is performed in the merged category regarding the number of additional  $b$ -tagged track jets not associated to the large- $R$  jet. Events with a so-called *additional  $b$ -tag* can be used to constrain the background from top-quark pair production. Figure 4.5 shows a sketch of the event topology with an additional  $b$ -tagged track jet. In the resolved 2-lepton analysis another top background control region is introduced, where two opposite-flavour leptons are required. For signal events one would expect two leptons with the same flavour.

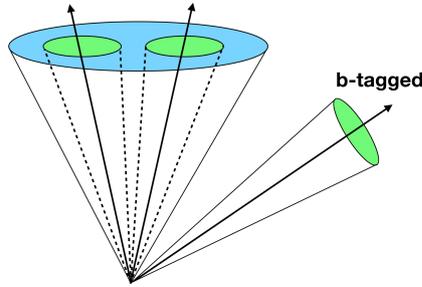


Figure 4.5: Sketch of the top-background-sensitive merged region with an additional  $b$ -tagged track jet not associated to the large- $R$  jet candidate.

## 4.4 Background processes

There are several SM processes with a similar final state as the decay  $V' \rightarrow VH \rightarrow \ell\ell, \ell\nu, \nu\nu + b\bar{b}$ . The Feynman diagrams of the most important background processes are shown in figure 4.6. The top quark pair production process ( $t\bar{t}$ ) with one or both  $W$  boson decaying leptonically is one of the main backgrounds, because it contains real  $b$ -jets from the top quark decay. However, also fully hadronic top-quark pairs with mis-identified non-prompt leptons or mis-measured jet energies leading to non-zero missing transverse momentum can contribute. Another main background is the production of leptonically decaying SM electroweak bosons ( $W$  or  $Z$ ) in association with jets ( $V$ +jets). The  $W$  or  $Z$  boson can be produced together with real  $b$ -jets, or with charm or light jets that may be mis-identified by the  $b$ -tagging algorithm. Smaller backgrounds originate from the production of SM electroweak boson pairs (diboson), one of which decays hadronically, the other leptonically, as well as events with a single top quark with real leptons and  $b$ -jets in the final state. The contribution from QCD multijets, as discussed in 4.5.2 stems mainly from mis-measured jets faking missing transverse momentum, or contributing a non-prompt lepton. Finally, the non-resonant production of a vector-boson associated SM Higgs boson is considered as an irreducible background in this analysis. The yields of the individual SM background processes in the signal regions of the 0-lepton channel are presented in table 4.4.

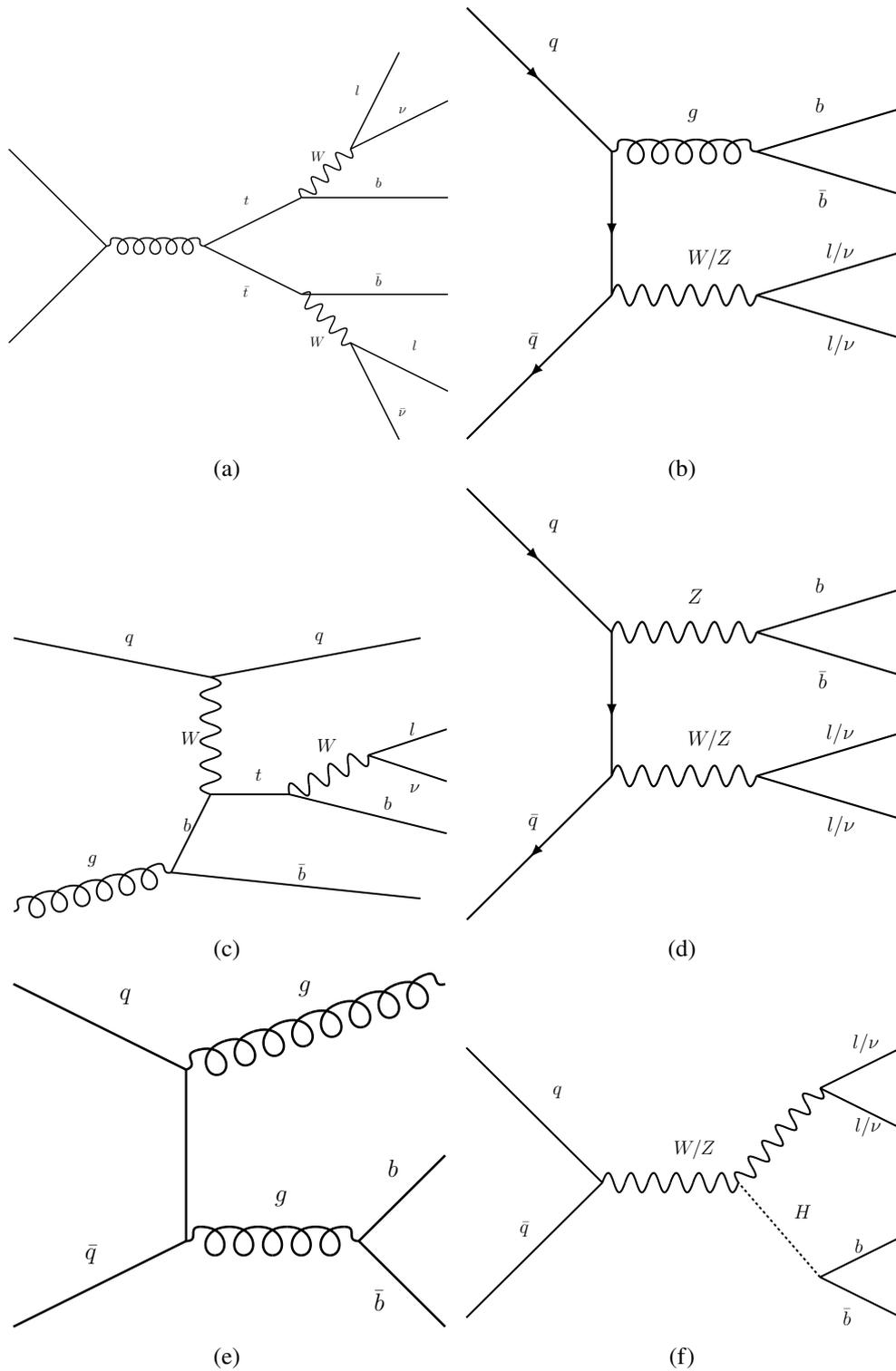


Figure 4.6: Background processes in the  $VH$  resonance analysis. Top: Main backgrounds (a) top quark pair production (example: dileptonic  $t\bar{t}$ ), (b) SM vector boson  $V$ + jets production, Middle and Bottom: processes with smaller background yields (c) single top quark production (example: t-channel), (d) SM vector boson pair production (example:  $WZ$ ) and (e) QCD multijet production (f) SM vector-boson associated Higgs production.

Process	Resolved		Merged	
	1-tag	2-tag	1-tag	2-tag
$t\bar{t}$	$22900 \pm 890$	$6640 \pm 180$	$1650 \pm 160$	$68 \pm 12$
Single top	$2440 \pm 330$	$552 \pm 76$	$217 \pm 52$	$15.4 \pm 4.1$
Diboson	$317 \pm 41$	$41.2 \pm 5.8$	$188 \pm 30$	$34.8 \pm 4.8$
$Z + l$	$580 \pm 210$	$1.3 \pm 1.3$	$310 \pm 130$	$0.38 \pm 0.29$
$Z + (bl, cl)$	$8240 \pm 840$	$50 \pm 17$	$910 \pm 160$	$10.1 \pm 3.7$
$Z + (bb, bc, cc)$	$1280 \pm 170$	$1270 \pm 150$	$238 \pm 45$	$101 \pm 16$
$W + l$	$960 \pm 300$	$3 \pm 2$	$227 \pm 95$	$1.0 \pm 0.6$
$W + (bl, cl)$	$5960 \pm 1100$	$56 \pm 17$	$770 \pm 230$	$6.6 \pm 3.2$
$W + (bb, bc, cc)$	$530 \pm 150$	$470 \pm 130$	$112 \pm 44$	$40 \pm 16$
$SM VH$	$55 \pm 21$	$102 \pm 39$	$7.4 \pm 2.9$	$4.7 \pm 1.8$
$t\bar{t}h$	$10.4 \pm 5.3$	$7.8 \pm 3.9$	$1.4 \pm 0.7$	$0.2 \pm 0.1$
$t\bar{t}V$	$102 \pm 54$	$41 \pm 22$	$17.7 \pm 9.5$	$1.4 \pm 0.8$
<b>Total</b>	$43400 \pm 200$	$9240 \pm 95$	$4650 \pm 79$	$282 \pm 14$

Table 4.4: Summary of postfit background yields in the 0-lepton channel of the VH resonance analysis. From [77].

The backgrounds in this analysis are estimated based on the prediction from Monte Carlo Simulation, except for the QCD multijet background in the 1-lepton channel which is estimated using a data-based method [77]. Table 4.5 summarizes the Monte Carlo generators used for each of the physics processes that constitute the background. All of the listed simulated samples include the effect of pile-up (see section 3.1.2) by overlaying every generated signal and background event with simulated minimum-bias events. Every event is passed through the GEANT4 ATLAS detector simulation [45] (see also section 3.2). The same standard ATLAS reconstruction software is applied in simulated events and in data. More information about the generated samples is given in [77].

## 4.5 Event selection

In order to improve the ratio of signal events to background a series of selection requirements are applied in the three different lepton channels. The focus will be on the 0-lepton channel event selection, but also 1-lepton and 2-lepton will be outlined. The selection requirements are different for the resolved and merged analysis categories.

### 4.5.1 Dataset and Trigger strategy

The data used in this analysis is proton-proton collision data from the LHC at a center-of-mass energy of  $\sqrt{s} = 13$  TeV recorded with the ATLAS detector in the years 2015 and 2016 corresponding to an integrated luminosity of  $(36.1 \pm 1.2) \text{ fb}^{-1}$ . Only data events are used, where all subsystems of the detector were fully operational.

In the zero-lepton channel events are selected using the unrescaled  $E_T^{\text{miss}}$  triggers with the lowest possible threshold. The triggers used in the 0-lepton channel of the VH resonance analysis for different periods of data-taking are listed in table 4.6.

Physics Process	MC Generator
Top-quark events	
$t\bar{t}$	POWHEG +PYTHIA 6
single top s-channel	POWHEG +PYTHIA 6
single top t-channel	POWHEG +PYTHIA 6
single top Wt-channel	POWHEG +PYTHIA 6
V + jets	
$W \rightarrow \ell\nu$	SHERPA 2.2.1
$Z/\gamma^* \rightarrow \ell\ell$	SHERPA 2.2.1
$Z/\gamma^* \rightarrow \nu\nu$	SHERPA 2.2.1
Diboson	
WW	SHERPA 2.2.1
WZ	SHERPA 2.2.1
ZZ	SHERPA 2.2.1
Standard Model Higgs	
$qq \rightarrow WH \rightarrow \ell\nu bb$	PYTHIA 8
$qq \rightarrow ZH \rightarrow \ell\ell bb$	PYTHIA 8
$qq \rightarrow ZH \rightarrow \nu\nu bb$	PYTHIA 8
$qq \rightarrow ZH \rightarrow \ell\ell bb$	POWHEG + PYTHIA 8
$gg \rightarrow ZH \rightarrow \nu\nu bb$	POWHEG + PYTHIA 8

Table 4.5: List of MC generators used for the SM background simulation for the VH resonance analysis.

Dataset/Period	Trigger
2015	HLT_xe70
2016 - Period A - D3	HLT_xe90_mht_L1XE50
2016 - Period $\geq$ D4	HLT_xe110_mht_L1XE50

Table 4.6: Summary of  $E_T^{\text{miss}}$  triggers used in the 0-lepton channel of the VH resonance analysis.

The nomenclature of  $E_T^{\text{miss}}$  triggers in table 4.6 is as follows: *HLT\_xe70* denotes an  $E_T^{\text{miss}}$  trigger with a threshold on the trigger-level missing transverse energy of  $E_T^{\text{miss}} > 70$  GeV, whereas *L1XE50* corresponds to a Level-1  $E_T^{\text{miss}}$  trigger threshold of 50 GeV and *mht* refers to an algorithm that calculated  $E_T^{\text{miss}}$  as the negative transverse momentum vector sum of jets reconstructed with the anti- $k_T$  algorithm from topological calorimeter cells in the detector [36]. For more detail on the different trigger levels, see also 3.2.3. The trigger strategy in the 1- and 2-lepton channels is documented in tables 4.8 and 4.7. For more details on the 1-lepton and 2-lepton channel strategy, see [77].

#### 4.5.2 Resolved selection

The events in the resolved category are required to have at least two signal jets (see table 4.2). Also events with a larger number of signal jets or forward jets are included.

In the 0-lepton channel the missing transverse energy  $E_T^{\text{miss}}$  is required to be larger than 150 GeV. The

cut is chosen such that the efficiency of the  $E_T^{\text{miss}}$  trigger (HLT\_xe70) has reached the maximum-efficiency plateau of 85% [36]. An additional cut on  $p_T^{\text{miss}} > 30$  GeV is imposed in the region with 0 and 1  $b$ -tagged jets to suppress non-collision backgrounds.

Additionally a cut is placed on the scalar sum of the  $p_T$  of the two leading signal jets (2-jet events) and the third leading  $p_T$  signal jet or forward signal jet (in events with three or more jets). The scalar sum  $S_T$  is required to be larger than 120 GeV for 2-jet events and larger than 150 GeV for events with three or more jets. This requirement is intended to remove a region that is mis-modelled in simulation, due to a dependency of the trigger efficiency on the jet activity in the event.

In order to suppress the multijet background contribution, a series of requirements is placed in the 0-lepton channel. The main source of multijet background in this channel is a mis-measurement of the energy of a jet. As a result, the missing transverse energy is aligned with the respective jet for multijet background events. The cuts for multijet-suppression therefore mainly make use of the angles between  $E_T^{\text{miss}}$  and the jets:

- $\Delta\phi(\mathbf{E}_T^{\text{miss}}, \mathbf{E}_{T,\text{trk}}^{\text{miss}}) < 90^\circ$
- $\Delta\phi(\text{jet}_1, \text{jet}_2) < 140^\circ$
- $\Delta\phi(\mathbf{E}_T^{\text{miss}}, h) > 120^\circ$
- $\min[\Delta\phi(\mathbf{E}_T^{\text{miss}}, \text{presel. jets})] > 20^\circ$

In the above list,  $\phi$  refers to the azimuthal angle,  $\text{jet}_1$  and  $\text{jet}_2$  are the two signal jets that form the Higgs boson candidate  $h$ . The above cuts are designed such that they reduce the multijet background in the 0-lepton channel below 1% such that it can be neglected. In the region with 4 or more jets, some multijet remains, which is why the cut on  $\min[\Delta\phi(\mathbf{E}_T^{\text{miss}}, \text{presel. jets})]$  was tightened to  $30^\circ$ .

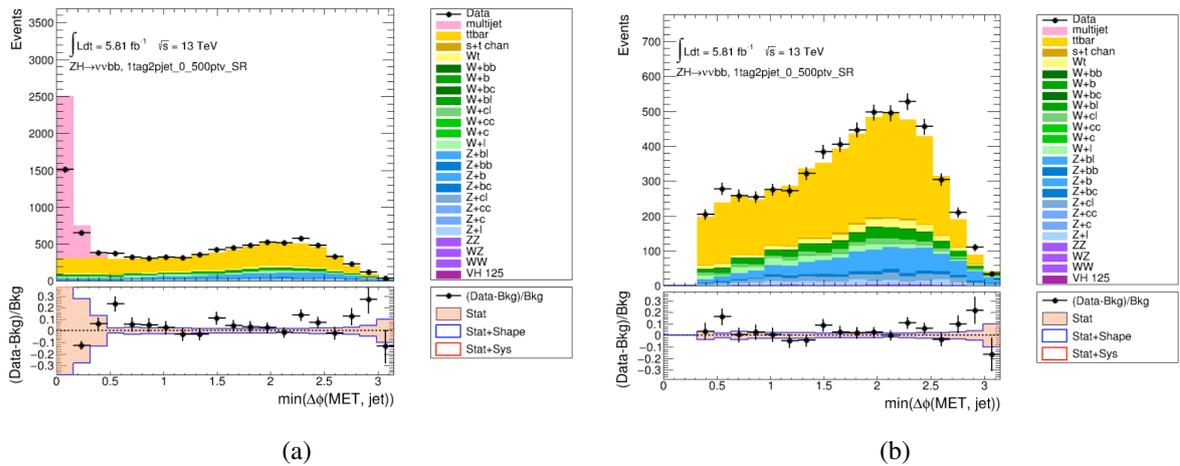


Figure 4.7: Pre-fit distribution of the angular distance  $\Delta\phi$  between the  $E_T^{\text{miss}}$  vector and the closest jet. Left: Without any Anti-QCD cuts, Right: With all Anti-QCD cuts applied.

Figure 4.7<sup>3</sup> shows the effect of the anti-QCD cuts for one region in the 0-lepton analysis. Before the anti-QCD cuts listed above are applied, it is clearly visible that the multijet contribution is cumulated at angles close to the jet axis, due to the previously described mis-measurement of the jet energy. After all anti-QCD cuts including a cut on  $\min(\Delta\phi(E_T^{\text{miss}}, \text{jet}))$  is applied, no more background from QCD multijets is present in this region.

A veto on events containing hadronically decaying tau leptons is applied, to reduce the background contribution from  $t\bar{t}$  and  $W$  + jets background. No tau veto is applied in the final HVT  $V'$  combination when combining 0-, 1- and 2-lepton channels, in order to enhance the signal efficiency for the  $W'$  signal for the decay  $W' \rightarrow WH \rightarrow \tau\nu b\bar{b}$ .

Additionally the invariant mass range of the Higgs boson candidate is reduced to  $50 \text{ GeV} < m_H < 200 \text{ GeV}$  to reduce the mis-modelling in MC. A summary of the selection in the 0-lepton channel can be found in table 4.7. The selection requirements for the 1- and 2-lepton channels in the resolved category are summarized in table 4.8 and 4.9.

Selection	0-lepton resolved	0-lepton merged
Trigger	$E_T^{\text{miss}}$ trigger	
0 lepton	veto $VH$ loose electrons and muons	
	veto medium taus	
$E_T^{\text{miss}}$	$150 \text{ GeV} < E_T^{\text{miss}}$	$200 \text{ GeV} < E_T^{\text{miss}}$
$p_T^{\text{miss}}$	$> 30 \text{ GeV}$	
$\sum p_T$	$> 120$ (2 jets) or $150 \text{ GeV}$ (at least 3 jets)	-
Jets	$\geq 2$ signal jets	$\geq 1$ large- $R$ jet $\geq 1$ track jet associated to leading large- $R$ jet
Leading jet $p_T$	$> 45 \text{ GeV}$	-
$m_H$	$50 \text{ GeV} < m_H < 200 \text{ GeV}$	
$ \min\Delta\phi(E_T^{\text{miss}}, \text{jet}) $	$> 20^\circ$ or $> 30^\circ$ (at least 4 jets)	
$ \Delta\phi(E_T^{\text{miss}}, h) $	$> 120^\circ$	
$ \Delta\phi(\text{jet1}, \text{jet2}) $	$< 140^\circ$	-
$ \Delta\phi(E_T^{\text{miss}}, p_T^{\text{miss}}) $	$< 90^\circ$	

Table 4.7: Event selection requirements in the 0-lepton resolved and merged categories in the VH resonance analysis.

### 4.5.3 Merged selection

Events in the merged category are required to have at least one large- $R$  jet. The leading large- $R$  jet in the event, which is the candidate for the  $H \rightarrow b\bar{b}$  decay is required to have at least one ghost-associated track jet. The same requirement on  $p_T^{\text{miss}} > 30 \text{ GeV}$  is required as in the resolved case.  $E_T^{\text{miss}}$  is required to be larger than 200 GeV. The multijet-suppression is achieved by a similar set of cuts as in the resolved 0-lepton analysis:

<sup>3</sup> It is important to note that this figure shows the distribution before the fit, as well as for multijet from MC, therefore discrepancies between data and MC are not unexpected.

Selection	1-lepton resolved	1-lepton merged
Trigger	Electron Event: Lowest unprescaled electron trigger, Muon Event: Lowest unprescaled $E_T^{\text{miss}}$ trigger	
1 lepton	exactly 1 WHLoose electron/muon which should also pass the WHSignal criteria	
$E_T^{\text{miss}}$	$E_T^{\text{miss}} > 80$ GeV in electron events $E_T^{\text{miss}} > 40$ GeV in muon events	$E_T^{\text{miss}} > 100$ GeV
$p_T(W)$	$p_T(W) > (-3.26 \times 10^5) \cdot \frac{1}{m_{\text{VH}}[\text{GeV}]} + 709.60$ lower cap of $p_T(W) > 150$ GeV	$p_T(W) > 394 \cdot \ln(m_{\text{VH}} [\text{GeV}]) - 2350$
$m_H$	$50 \text{ GeV} < m_H < 200 \text{ GeV}$	
Jets	2 or 3 signal jets $\geq 0$ forward jets	$\geq 1$ large- $R$ jet $\geq 1$ track jet associated to leading large- $R$ jet
Leading jet $p_T$	$> 45$ GeV	-
	Veto on events with 3 or more $b$ -jets	-

Table 4.8: Event selection requirements in the 1-lepton resolved and merged categories in the VH resonance analysis.

Selection	2-lepton resolved	2-lepton merged
Trigger	OR of lowest un-prescaled single lepton triggers	
2 lepton	1 $VH$ loose +1 $ZH$ signal	2 $ZH$ signal
	same flavour ( $ee$ or $\mu\mu$ ) oppositely charged leptons for $\mu\mu$ events in resolved regime	
$m_{\ell\ell}$	$\max(40\text{GeV}; -0.03 \cdot m(ZH) + 87 \text{ GeV}) < m(\ell\ell)[\text{GeV}] < 0.013 \cdot m(ZH) + 97\text{GeV}$	
$E_T^{\text{miss}}/\sqrt{H_T}$	$E_T^{\text{miss}}/\sqrt{H_T} < \frac{0.008 \cdot m(ZH)}{\sqrt{\text{GeV}}} + 1.15\sqrt{\text{GeV}}$	
Jets	$\geq 2$ signal jets	$\geq 1$ fat jet $\geq 1$ track-jet associated to leading fat jet
$m_H$	$50 \text{ GeV} < m_H < 200 \text{ GeV}$	
Leading jet $p_T$	$p_T > 45$ GeV	-
$p_T(Z)$	$m_{\text{VH}} > 320 \text{ GeV}: p_T(Z) > 20 \text{ GeV} + 9\sqrt{\text{GeV}} \cdot \sqrt{m(ZH)[\text{GeV}]} - 320 \text{ GeV}$	

Table 4.9: Event selection requirements in the 2-lepton resolved and merged categories in the VH resonance analysis.

- $\Delta\phi(\text{jet}, E_T^{\text{miss}}) > 20^\circ$  where jet denotes every accepted small- $R$  jet
- $\Delta\phi(\mathbf{E}_T^{\text{miss}}, \mathbf{p}_T^{\text{miss}}) < 90^\circ$
- $\Delta\phi(\mathbf{E}_T^{\text{miss}}, \text{large-}R \text{ jet}) > 120^\circ$

Also in the merged category the tau veto is applied, but not for the HVT  $V'$  combination. Also the mass range is restricted to  $50 \text{ GeV} < m_J < 200 \text{ GeV}$ , where  $m_J$  denotes the Higgs boson candidate large- $R$  jet invariant mass. A summary of the selection requirements in the 0-lepton channel can be found in table 4.7. The selection requirements for the 1- and 2-lepton channels in the merged category are summarized in table 4.8 and 4.9.

## 4.6 Reconstruction of the VH invariant mass

The final discriminant in this analysis is the invariant mass of the resonance candidate,  $m_{VH}$ . To obtain  $m_{VH}$ , the four-vectors of the Higgs boson candidate and the electroweak V boson candidate are added. In the merged category the Higgs boson candidate is the large- $R$  jet. In the resolved category the Higgs boson is reconstructed by adding the four-vectors of the two signal jets. The electroweak boson candidate is reconstructed from  $E_T^{\text{miss}}$  and leptons, depending on the lepton channel:

- **0-lepton channel ( $Z \rightarrow \nu\bar{\nu}$ ):**

In the 0-lepton channel, only the transverse momentum of the  $Z \rightarrow \nu\bar{\nu}$  candidate can be measured. Therefore the full invariant mass of the  $ZH$  system is not accessible, and the transverse mass  $m_T(ZH)$  is used instead:

$$m_T(ZH) = \sqrt{(|\mathbf{p}_T^h| + |\mathbf{E}_T^{\text{miss}}|)^2 - (\mathbf{p}_T^h + \mathbf{E}_T^{\text{miss}})^2} \quad (4.7)$$

In the above equation,  $\mathbf{p}_T^h$  denotes the transverse momentum vector of the Higgs boson candidate and  $\mathbf{E}_T^{\text{miss}}$  the missing transverse energy vector.

- **1-lepton channel ( $W \rightarrow \ell\nu$ ):**

In the 1-lepton channel the W boson candidate is reconstructed from a lepton and a neutrino. To reconstruct the full invariant mass of the W boson, an additional step is needed, since only the transverse momentum of the neutrino is accessible in the detector. A mass constraint is applied, since the W boson mass of  $m_{\ell\nu} = m_{W,\text{PDG}} = 80.385 \text{ GeV}$  [3] is known. The mass constraint for  $m_{\ell\nu}$  leads to a quadratic equation, which can be solved for the z-component of the neutrino momentum,  $p_z$ . If there is a complex solution to the quadratic equation, the real component is taken as  $p_z$ . If there are two real solutions, the smaller one is used for  $p_z$ . With  $p_z$  known, the full invariant mass of the WH system,  $m_{WH}$  is calculated.

- **2-lepton channel ( $Z \rightarrow \ell\ell$ ):**

In the 2-lepton channel all lepton four-vectors can be fully reconstructed in the detector. The Z boson candidate is reconstructed by adding the two lepton four-vectors and  $m_{ZH}$  is constructed by adding the four-vectors of the Z boson and Higgs boson candidates.

Since the mass of the SM Higgs boson is known to be  $m_{H,\text{PDG}} = (125.18 \pm 0.16) \text{ GeV}$  [3] the resolution of the invariant mass  $m_{VH}$  in the resolved analysis can be further improved by scaling both four-vectors of the two b-jet candidates by the factor:

$$f_H = \frac{m_{H,\text{PDG}}}{m_{bb}} \quad (4.8)$$

where  $m_{H,\text{PDG}}$  is the known mass of the SM Higgs boson and  $m_{bb}$  is the mass of the di-b-jet system. This operation effectively constrains the mass of the di-b-jet system to exactly the known Higgs boson mass. It should be noted that the mass-dependent scaling of the di-b-jet system is performed only after the  $m_{bb}$  mass window selection cut, therefore the b-jet energy scale corrections described in section 4.2 are very important for the event selection, despite the mass constraint being applied afterwards. In the merged category no mass-dependent scaling is applied to the large- $R$  jet four vector.

To further improve the mass resolution in the 2-lepton channel, the four-vectors of the leptons from the Z candidate are scaled by:

$$f_Z = \frac{m_{Z,\text{PDG}}}{m_{\ell\ell}} \quad (4.9)$$

where  $m_{Z,\text{PDG}}$  is the known mass of the Z-Boson of  $m_{Z,\text{PDG}} = (91.1876 \pm 0.0021)$  GeV [3] and  $m_{\ell\ell}$  is the invariant mass of the dilepton system. Note that in both cases the rounded values of  $m_H = 125$  GeV and  $m_Z = 91$  GeV have been used for the scaling, since the invariant mass resolution achievable in this analysis is much lower than the world-average mass measurement uncertainties.

## 4.7 Systematic uncertainties

One can categorize the systematic uncertainties applied in this analysis into experimental and Monte-Carlo-modelling-related uncertainties. The experimental uncertainties are related to the reconstruction of the different physics objects in the detector, whereas the modelling uncertainties parametrize the difference in the final discriminating variable as observed in different MC simulations. A summary of the considered experimental uncertainties related to different objects can be found in Table 4.10.

The modelling uncertainties are derived for the (transverse) invariant mass of the  $VH$  system. They are determined for the main background processes  $t\bar{t}$ ,  $V$ +jets and single top (for 1-lepton channel only) by comparing the distribution of  $m_{VH}$  and  $m_{T,VH}$  (in the 0-lepton channel) in the nominal MC samples to various alternative samples with different generators and varied simulation parameters. The modelling uncertainties are split in two components namely the normalization, or acceptance uncertainty and the shape uncertainty.

## 4.8 Statistical analysis

The aim of this search analysis is to scrutinize the  $VH$  final state invariant mass spectrum for the presence of an excess of data events over the SM background prediction. Using the terminology of statistics, this analysis is a hypothesis test of the SM background-only hypothesis against a New Physics signal plus background hypothesis.

The parameter one is interested in measuring is the New Physics signal strength  $\mu$ . However, many of the cross sections of the SM background processes are not known with perfect precision and also experimental uncertainties as well as MC modelling uncertainties of quantities used for the event selection play an important role. As mentioned in section 3.7.4, for a statistical problem with many parameters but only a single parameter of interest (in this case the signal strength  $\mu$ ), the profile likelihood ratio is a useful concept. In this analysis, the profile likelihood ratio is therefore used as a test statistic for the hypothesis test.

A complex likelihood fit model is used, with  $\mu$  as a parameter of interest and all other parameters, such as background normalizations and systematic uncertainties included as nuisance parameters (NP) (see 3.7.6). The fit to data yields information about the signal strength of a possible New Physics contribution, while simultaneously providing auxiliary measurements to constrain the NP describing normalizations of MC templates as well as systematic uncertainties.

In the 1- and 2-lepton channel, the full invariant mass  $m_{VH}$  is used in the fit, whereas the transverse mass  $m_{T,VH}$  is used in the 0-lepton channel. The fit is performed individually in each of the channels. In the end the statistical combination of the different channels is performed. In total six different fits

Systematic Uncertainty	Brief Description
<b>Event-related</b> Lumi PRW_SF	uncertainty on total integrated luminosity uncertainty on pile-up reweighting
<b>Electron-related</b> EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR EG_SCALE_ALL EG_RESOLUTION_ALL	trigger efficiency uncertainty reconstruction efficiency uncertainty identification efficiency uncertainty isolation efficiency uncertainty energy scale uncertainty energy scale uncertainty
<b>Muon-related</b> MUON_EFF_TrigStatUncertainty MUON_EFF_TrigSysUncertainty MUON_EFF_STAT MUON_EFF_SYS MUON_EFF_STAT_LOWPT MUON_EFF_SYS_LOWPT MUON_ISO_STAT MUON_ISO_SYS MUON_TTVA_STAT MUON_TTVA_SYS MUON_SCALE MUON_ID MUON_MS	trigger efficiency uncertainty    reconstruction and identification efficiency uncertainty for muons with $p_t > 15$ GeV    reconstruction and identification efficiency uncertainty for muons with $p_t < 15$ GeV    isolation efficiency uncertainty    track-to-vertex association efficiency uncertainty    energy scale uncertainty energy resolution uncertainty from inner detector energy resolution uncertainty from muon spectrometer
<b>R=0.4 Jet-related</b> JET JET_JER_SINGLE_NP JvtEfficiency FT_EFF_Eigen_B FT_EFF_Eigen_C FT_EFF_Eigen_L FT_EFF_Eigen_extrapolation FT_EFF_Eigen_extrapolation_from_charm	energy scale uncertainty in 21 components energy resolution uncertainty Jet Vertex Tagger efficiency uncertainty flavour tagging efficiency uncertainty, 3 components for b-jets 4 components for c-jets 5 components for light jets efficiency uncertainty from extrapolation to high- $p_T$ jets efficiency uncertainty for tau-jets
<b>R=1.0 Jet-related</b> SysJET_Comb_Baseline_Kin SysJET_Comb_Modelling_Kin SysJET_Comb_Tracking_Kin SysJET_Comb_TotalStat_Kin FATJET_JER FATJET_JMR	large- $R$ jet energy scale ( $p_T$ and mass scale fully correlated)          large- $R$ energy resolution large- $R$ mass resolution
<b>R=0.2 track Jet-related</b> FT_EFF_Eigen_B FT_EFF_Eigen_C FT_EFF_Eigen_L FT_EFF_Eigen_extrapolation FT_EFF_Eigen_extrapolation_from_charm	flavour tagging efficiency uncertainty, 3 components for b-jets 4 components for c-jets 5 components for light jets efficiency uncertainty from extrapolation to high- $p_T$ jets efficiency uncertainty for tau-jets
<b>MET-related</b> METTrigStat METTrigTopZ MET_SoftTrk_ResoPara MET_SoftTrk_ResoPerp MET_SoftTrk_Scale MET_SoftTrk_Scale MET_JetTrk_Scale	trigger efficiency uncertainty    track-based MET soft-term longitudinal resolution uncertainty track-based MET soft-term transverse resolution uncertainty track-based MET soft-term longitudinal scale uncertainty track-based MET soft-term longitudinal scale uncertainty track-based MET scale uncertainty due to tracks in jets

Table 4.10: Summary of the experimental systematic uncertainties used in the  $VH$  resonance analysis with their explanations.

are performed for this analysis: the 0-,1- and 2-lepton channels, as well as the combination for the  $Z'$  (0+2-lepton channels),  $W'$  (0+1-lepton channels) and  $V'$  (0+1+2-lepton channels) each have an individual fit. The focus here will be laid on the 0-lepton channel as well as the  $V'$  combination.

To study the  $m_{VH}$  spectrum and to search for an excess, a binned maximum-likelihood fit using the RooStats package [72, 75, 76, 86, 87] is performed to observed data using templates describing the signal and background distributions. For a more detailed description of the profile likelihood method, hypothesis tests and the limit setting procedure, see section 3.7. The result of the statistical analysis is the 95%CL limit set on the HVT resonance production cross section times branching ratio.

#### 4.8.1 Fit inputs, regions and binning

The main backgrounds in the  $VH$  resonance analysis are  $W$  or  $Z$  bosons events produced in association with two or more jets, as well as  $t\bar{t}$  events. All background processes that are considered in the fit are listed in table 4.11. The jet flavour labels ( $ll, cl, bl, bc, cc, bb$ ) of the  $V$ +jets backgrounds refer to the truth hadron labels (described in 3.5.6) of the two signal jets in the resolved category and the two track jets associated to the large- $R$  jet in the merged category. Several background templates with small event yields were merged before fitting, indicated by summation signs, because the fit is not sensitive to the individual components.

Sample nomenclature in $VH$ fit model
$W + l = (W + ll)$
$W + clbl = (W + cl) + (W + bl)$
$W + hf = (W + bb) + (W + bc) + (W + cc)$
$Z + l = (Z + ll)$
$Z + clbl = (Z + cl) + (Z + bl)$
$Z + hf = (Z + bb) + (Z + bc) + (Z + cc)$
$t\bar{t}$
$t\bar{t} + V$
$t\bar{t} + h$
diboson= $ZZ + WW + WZ$
single-top= $s$ -channel+ $t$ -channel+ $Wt$
multijet (only in 1-lepton channel)
SM $VH$ production
HVT $W'$ and $Z'$ (separate)

Table 4.11: List of processes and merging of MC templates for the  $VH$  resonance fit.

The fit regions used in the individual lepton channels are summarized in table 4.12. In some categories with low event yields, regions can be combined across different number of  $b$ -tags or different  $m_{bb}$  mass ranges. This helps to decrease the influence of statistical fluctuations on the fit parameters in these regions. For example, the low and high  $m_{bb}$  sideband regions are combined in the 1-lepton channel.

Figure 4.8 shows the contribution from the merged and the resolved category to the expected limit to illustrate which categories contribute the largest sensitivity to the analysis in different regions of phase space. It can be observed that the resolved signal region contributes the most sensitivity for a resonance mass up to 1.2 GeV. The sensitivity for the resolved category increases with increasing resonance mass in

topology	region	0-lepton	1-lepton	2-lepton
2-tag resolved signal region	2 $b$ -tags, $\geq 2$ jets, SR	✓	✓	✓
1-tag resolved signal region	1 $b$ -tag, $\geq 2$ jets, SR	✓	✓	✓
2-tag merged signal region	2 $b$ -tags, $\geq 1$ large- $R$ jets, 0 add. $b$ -tags	✓	✓	✓
1-tag merged signal region	1 $b$ -tag, $\geq 1$ large- $R$ jets, 0 add. $b$ -tags	✓	✓	✓
resolved $m_{bb}$ -sideband control region	1 or 2 $b$ -tags, $\geq 2$ jets, mBBCR	✗	✓	✗
resolved top $e\mu$ control region	1 or 2 $b$ -tags, $\geq 2$ jets, opp. flav. leptons	✗	✗	✓

Table 4.12: Definition of the  $H \rightarrow b\bar{b}$  candidate mass range for signal regions in the different lepton- and topology-based categories.

this regime, because the SM background contributions get smaller, as can also be seen in the  $m_{VH}$  spectra in figure 4.20. For higher masses the sensitivity in the resolved category deteriorates because of lower signal efficiency. This is due to the merging of  $R = 0.4$  calorimeter jets as well as the known decrease in  $b$ -tagging efficiency at high jet  $p_T$ . This effect has been studied extensively for the optimization of Higgs tagging [53]. The merged category is not contributing to the sensitivity at low mass, mainly due to the fact that the PRSR categorization scheme ( cf. 4.3) employed here, preferably sorts events at low resonance mass into the resolved category. However, at low mass the merged category also suffers from a low signal efficiency. This is caused by the fact that for low resonance masses the separation between the  $b$ -jets from the  $H \rightarrow b\bar{b}$  decay is too large to be reconstructed within the large- $R$  jet. At high resonance masses the merged category contributes the most sensitivity. The plateau in the expected limit is reached due to the fact that at very high resonance masses there are almost no events from SM background any more. In addition, the signal reconstruction efficiency drops due to  $b$ -tagging efficiency loss at high  $p_T$  and, at very high resonance masses, the merging of the  $R = 0.2$  track jets. The sensitivity in the intermediate mass region from 1.2 TeV to 1.8 TeV both the contribution from the resolved and the merged region are important. To ensure a smooth transition when combining the two categories the PRSR scheme is used, because it provides optimal sensitivity in the intermediate resonance mass regime.

To constrain the normalization of a certain SM background process in the fit, control regions that are enhanced in events from this process can be used. For the two-lepton channel a region where the two leptons are of opposite flavour is used as a control region for the  $t\bar{t}$  background contribution. Figure 4.9 shows the expected limit for the 0-lepton channel when considering only the signal regions in the fit and when adding additional control regions in the  $m_{bb}$  sidebands or a top control region. As can be seen, there is no significant improvement in the expected limit when adding additional control regions in the  $m_{bb}$  sidebands or a top control region, except for an improvement up to 15% when adding the  $m_{bb}$  sideband regions for very low and very high resonance masses. Since the signal regions themselves provide enough constraint on the normalizations of the other background contributions, all other regions introduced in section 4.3 are used as fit validation regions instead. Since the contribution of the 0-lepton channel compared to the other lepton channels expected to be small for very low resonance masses (see figure 4.19) and in order to benefit from studying the post-fit agreement between data and MC in an additional validation region (see 4.9.5) only the resolved and merged signal regions were used in the 0-lepton fit.

The fits to the  $m_{VH}$  and  $m_{T,VH}$  spectra are performed using a binning of the distributions that is chosen such that the sensitivity of the analysis is maximised, while not being subject to statistical fluctuations. The choice of binning has a significant impact on the fit stability and sensitivity.

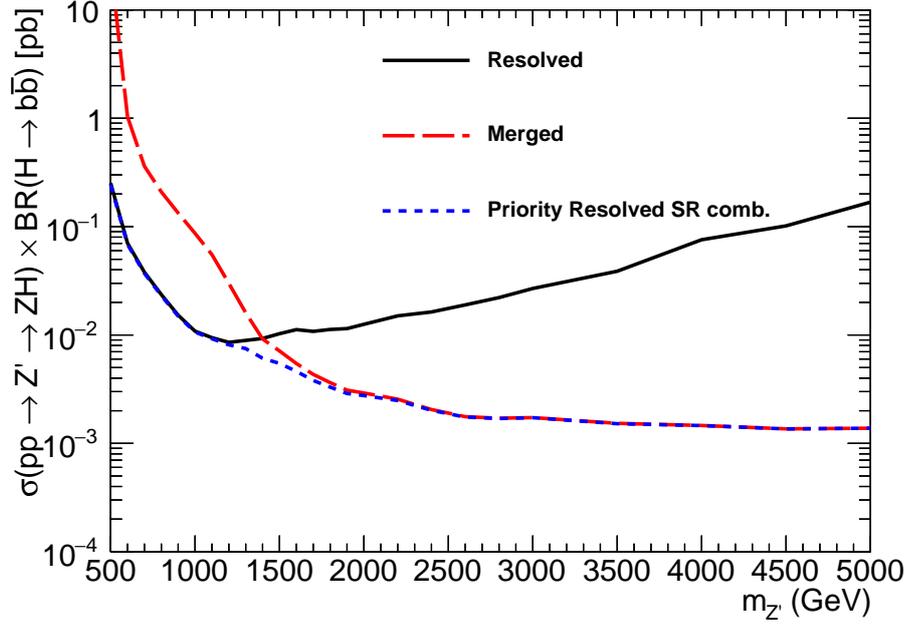


Figure 4.8: Comparison of the expected limits for the 0-lepton channel with different regions being considered in the fit. The solid black line shows the limit using the resolved signal region only. The red curve represents the limit using the merged signal region only. The dashed blue line shows the two regions combined.

Figure 4.10 shows the expected limit in the 0-lepton channel for different binning strategies. The standard binning uses non-equidistant bins which are summarized in table 4.13. This binning strategy originates from the search for a pseudo-scalar boson  $A \rightarrow ZH$ <sup>4</sup>, and was optimized by hand to yield the best sensitivity.

$m_{T,VH}$ range [TeV]	bin width [GeV]
0-1.0	100
1-1.2	200
1.2-1.5	300
1.5-2.2	700
>2.2	1000
<b>Statistics Requirement</b>	$\Delta N_{\text{bin}}^{MC} / N_{\text{bin}}^{MC} < 0.75$
<b>Bin Content</b>	$N_{\text{bin}}^{MC} > 0$

Table 4.13: Standard non-equidistant bins used for the 0-lepton channel  $m_{VH}$  distribution in the resolved and merged categories.  $N_{\text{bin}}^{MC}$  denotes the content per bin, and  $\Delta N_{\text{bin}}^{MC}$  is the statistical error on the bin content, both for the total sum of the background Monte Carlo.

An alternative binning strategy shown in figure 4.10 uses the following parametrization of twice the standard deviation of the  $W'$  resonance peak in simulation:

$$\frac{w_{\text{bin}}}{\text{GeV}} = 40 + \frac{155 - 40}{2800 - 500} \cdot \left( \frac{m_{VH}}{\text{GeV}} - 500 \right) \quad (4.10)$$

<sup>4</sup> The  $A \rightarrow ZH$  search analysis was published alongside the presented HVT resonance search in [77]

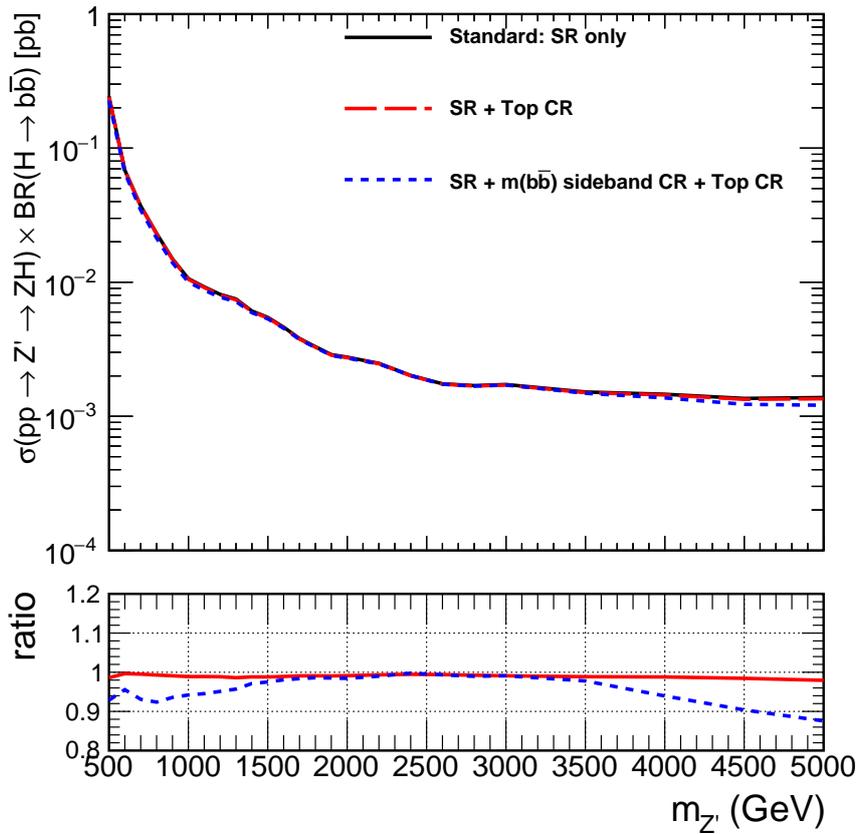


Figure 4.9: Comparison of the expected limits for the 0-lepton channel with different regions being considered in the fit

Eq. 4.10 is determined by studying the standard deviation of simulated HVT  $W' m_{\text{VH}}$  signal distribution, assuming a Gaussian signal distribution, for different resonance mass points and fitting a linear function. For the expected limit curves shown in figure 4.10 an additional requirement on the bin statistical error to be smaller than 70% (15%) was made. For all three binning strategies bins with zero predicted SM background content are merged into the next adjoining bin, to avoid having bins with zero background prediction, for which the binned likelihood makes no sense.

As can be seen in figure 4.10, the standard binning from table 4.13 and the resolution-based binning with loose bin statistics requirement yield similar sensitivity, whereas the resolution-based binning with a tighter requirement on the statistical uncertainty yields up to 50% worse expected limit, especially at high masses, where a tight requirement on the bin statistics leads to a coarser binning. Because of the higher sensitivity and more simplicity with respect to the resolution-based binning, the standard binning scheme as described in table 4.13 is used in the 0-lepton channel.

## 4.9 Fit diagnostics

In the previous section the main tool to study the sensitivity of the resonance search was the expected limit. To study the performance of the profile likelihood fit in more detail it is important to define additional

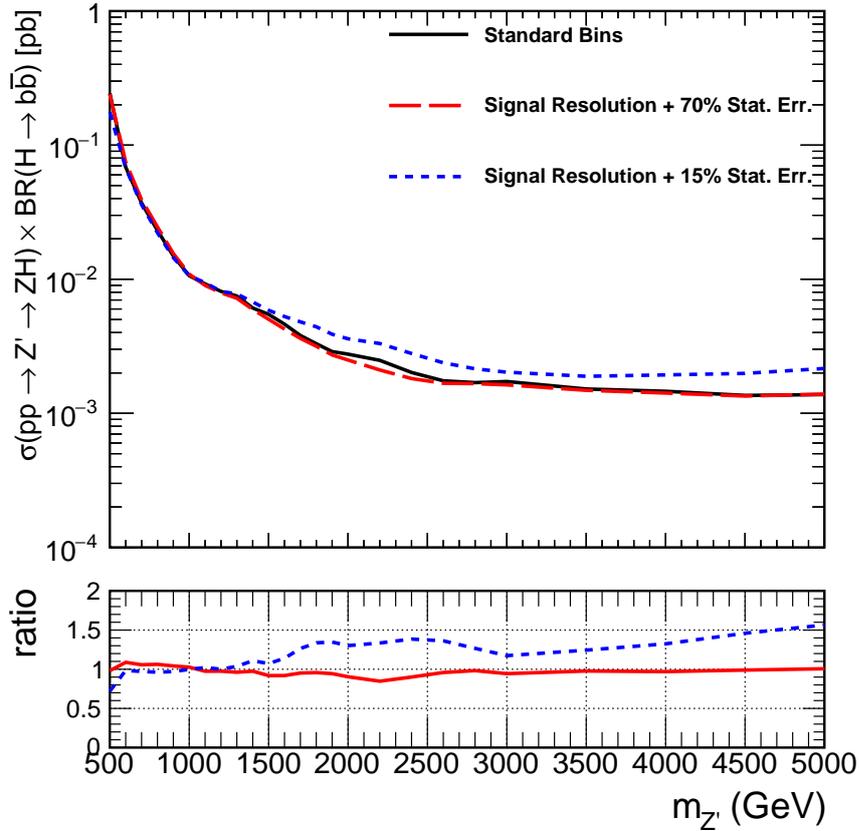


Figure 4.10: Comparison of the expected limits for the 0-lepton channel with different binning options being considered in the fit

figures of merit to estimate and compare the fit performance as well as methods to diagnose possible problems. In the following section the diagnostic tools used to validate and optimize the statistical model will be introduced and results will be discussed:

#### 4.9.1 Nuisance parameter pull distributions

As discussed in section 3.7.6 systematic uncertainties are incorporated in the Profile Likelihood fit via the introduction of Nuisance Parameters (NP). One distinguishes between so-called *floating* nuisance parameters, which are allowed to vary freely in the fit, and *constrained* nuisance parameters which are associated with a Gaussian- or log-normal penalty function. Floating nuisance parameters are used to adjust the normalization of processes whose cross-sections are not well measured and which can be sufficiently well constrained in the data. Constrained nuisance parameters are used in all other cases, where a value for a cross-section or for a systematic shift like the Jet energy scale is measured or calibrated with a certain uncertainty.

In the  $VH$  resonance 0-lepton analysis the  $t\bar{t}$  background normalizations are treated as floating nuisance parameters, as well as the normalizations of the  $Z + (bb, cc)$  and  $Z + (bl, cl)$  backgrounds. In figure 4.12 these nuisance parameters are denoted by the prefix *norm*.

Figure 4.11 illustrates how a pull distribution is determined. A pull distribution is created by releasing one nuisance parameter from the fit, scanning it, and studying the effect on the profile likelihood ratio (PLR). According to Wilk's theorem (section 3.7.7) the scan of a nuisance parameter should yield a  $\chi^2$  distribution of the PLR. The pull distribution shows the best-fit value of the nuisance parameter as well as the uncertainties based on the  $1\sigma$  band around the minimal negative logarithm of the likelihood. It therefore allows to study if the PLR indeed follows the expectation from Wilk's theorem and whether the best-fit value and width of the nuisance parameter is consistent with the expectation.

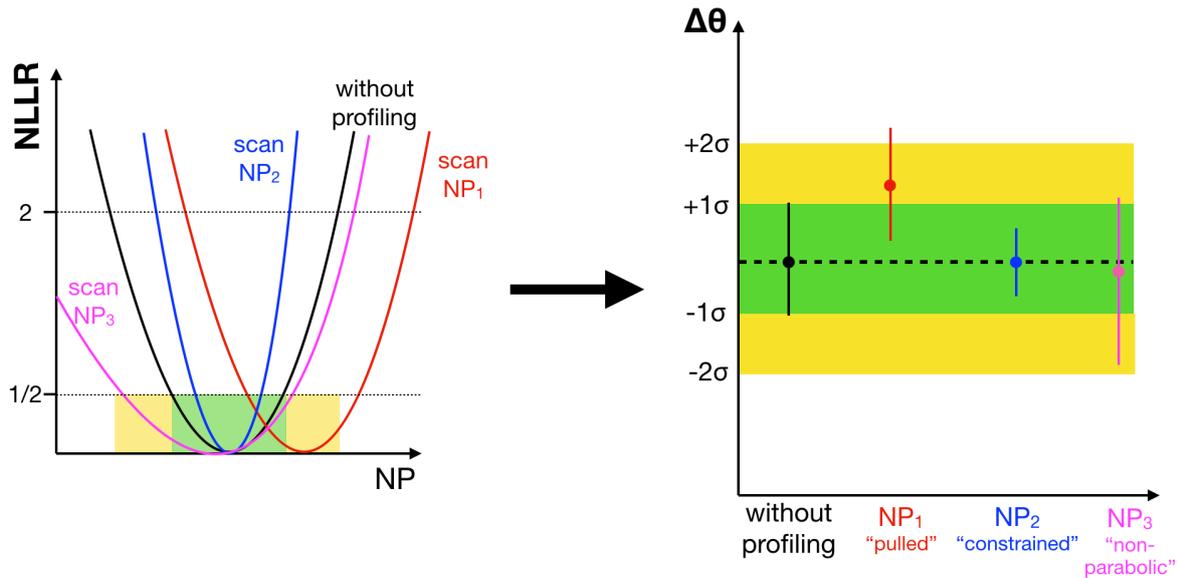


Figure 4.11: Sketch explaining the determination of the nuisance parameter (NP) pull distribution. The left figure illustrates the behaviour of the negative log likelihood ratio (NLLR) when profiling different nuisance parameters. The right figure shows the corresponding NP pull distribution.

When speaking about pull distributions often the term *constraint* is used, not to be confused with the aforementioned log-normal or Gaussian constraint of a NP. This term describes the narrowing of the parabola of the negative log-likelihood ratio (NLLR) while profiling the NP compared to the NLLR without profiling the NP. In the pull distribution this narrowing can be seen for nuisance parameters whose uncertainties are smaller than the yellow and green  $2\sigma$  and  $1\sigma$  error bands.

For nuisance parameters that are associated with a penalty function it is unlikely that the maximum likelihood is reached for a parameter configuration where a nuisance parameter is shifted (*pulled*) significantly from its central value at zero (see figure 4.11). If several nuisance parameters show such behaviour, this can point to an issue with the fit convergence or an insufficient agreement between data and simulation. If the fit is able to constrain a nuisance parameter in data much more than in simulation, this can indicate that the prefit uncertainty was overestimated. An occurrence of many nuisance parameters that are significantly more - or less - constrained in a fit to data compared to the fit to the Asimov dataset (defined in 3.7.7) can serve as an indicator for problems with non-parabolic PLR behaviour, fit convergence or modelling problems in simulation. To diagnose such problems, figure 4.12 is used, which shows the nuisance parameter pull distribution for the fit in the VH resonance 0-lepton channel. Each point corresponds to a nuisance parameter. For an explanation of the

NP names, see table 4.10. The coloured boxes indicate groups of different types of systematic uncertainties.

The error bars of the points shown in 4.12 encode how well the fit is able to constrain the respective nuisance parameters. Figure 4.12 shows a comparison between nuisance parameters of the fit to the Asimov dataset (red points) and the fit to data (black points) for a resonance mass of 1.5 TeV. For a good fit, the black points should not be removed beyond the one-sigma-band of the red point. One exception in 4.12 is the NNLORW nuisance parameter, which encodes a sequential reweighting of the top-quark and  $t\bar{t}$  pair  $p_T$  spectra in simulation to agree with next-to-next-to-leading-order parton level prediction. This reweighting is intended to improve the agreement between data and MC. If the NNLO reweighted spectrum matches the observed data better, a pull of the corresponding nuisance parameter is expected and no sign of a fit problem. Furthermore it should be noted that for a NP that describes a freely floating NP (such as  $t\bar{t}$  and  $Z$ +jets normalizations), a strong constraint and a pull are expected.

Another criterion for a good fit behaviour is that the error bars on the black nuisance parameters from the fit to Data should not be much bigger or smaller than for the red points from the fit to the Asimov dataset. Figure 4.12 gives an indication that the 0-lepton channel fit converges and behaves as expected. For each step of the fit model optimization, new pull distributions such as 4.12 are created and a good fit behaviour is verified.

#### 4.9.2 Nuisance parameter ranking

To understand the behaviour of the profile likelihood fit it is useful to understand which of the systematic uncertainties has the most impact on the determination of the signal strength  $\mu$ . To estimate the impact of a certain nuisance parameter, it is scanned within the  $\pm 1\sigma$  band around their nominal value during the profile likelihood fit while all other nuisance parameters are allowed to vary as usual and are minimized to their best-fit-value. The best-fit signal strength  $\hat{\mu}$  is recalculated for each value of the nuisance parameter under study. The nuisance parameter for which  $\hat{\mu}$  varies the most is ranked the highest.

Figure 4.13 shows the nuisance parameter ranking plots for the 0-lepton channel fit for different mass hypotheses for the  $Z'$  resonance. For both mass points the floating normalization nuisance parameters for the dominant backgrounds have a significant impact on the estimation of  $\hat{\mu}$ . For the lowest mass point  $m_{Z'} = 500$  GeV the most influential nuisance parameters are the uncertainties on the flavour tagging of  $c$ -jets for  $R = 0.4$  calorimeter jets. In general one can see that many of the flavour tagging uncertainties are among the ten highest-ranked nuisance parameters. This is because  $V$ +jets and  $t\bar{t}$  processes contribute a significant background fraction in the 0-lepton channel and the flavour admixture plays an important role for the shape of the  $m_{T,VH}$  distribution. The  $c$ -jet-related uncertainty is of particular importance since the  $c$ -jet contamination is high and uncertainties on the  $c$ -jet mis-identification efficiency are large. The importance of the calorimeter jet flavour tagging uncertainty for low resonance masses can be explained by the fact that at low mass mainly the resolved categories contribute to the analysis sensitivity (cf. figure 4.8), which make use of  $R = 0.4$  calorimeter jets for flavour tagging.

As an example for the nuisance parameter ranking at high resonance masses, the mass point  $m_{Z'} = 2$  TeV is shown in figure 4.13. Furthermore, at high resonance masses masses, the merged Higgs boson decay category becomes the most important contribution as can be seen by the appearance of the track-jet related  $b$ -tagging uncertainties in the nuisance parameter ranking. As can be seen also in figure 4.15 at this mass point the statistical uncertainties are dominant.

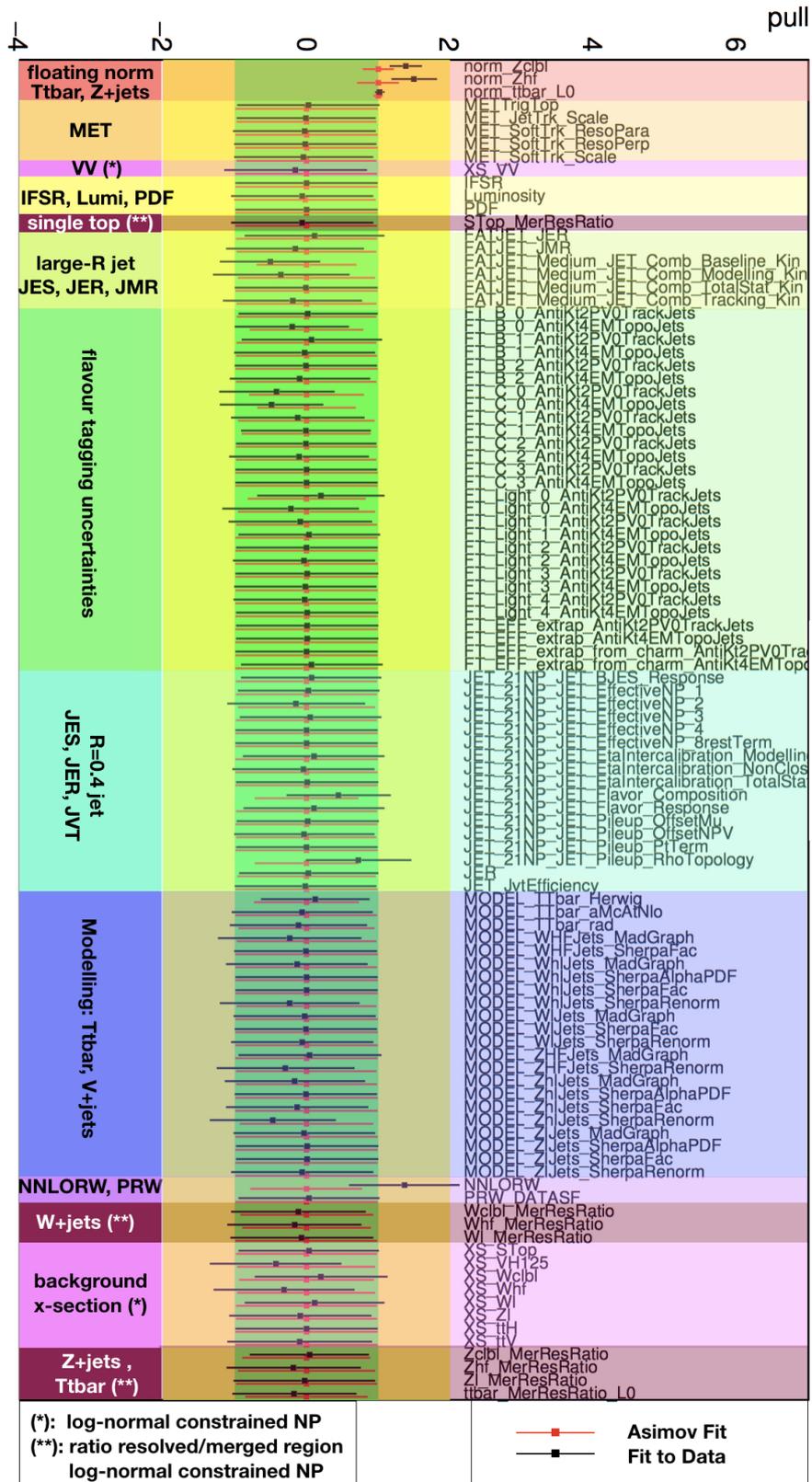


Figure 4.12: Nuisance parameter pull distribution for the 0-lepton channel in the final fit configuration at the 1.5 TeV  $Z'$  mass point.

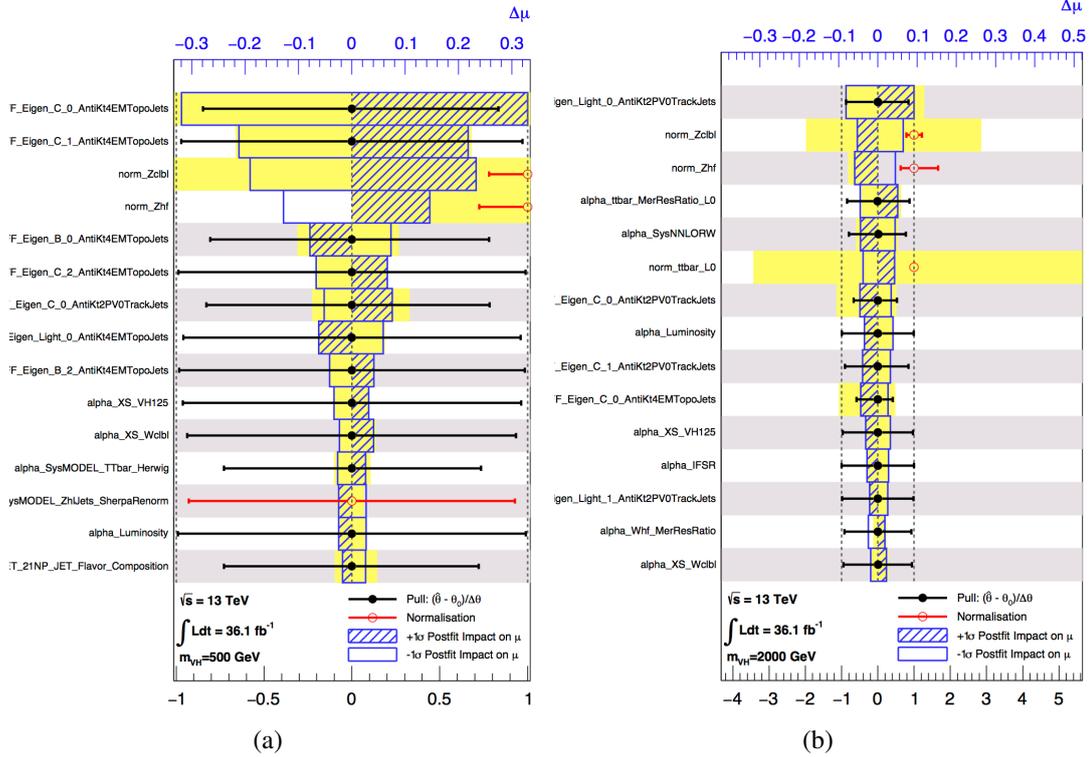


Figure 4.13: Nuisance parameter ranking for the 0-lepton channel fit. Figure (a) shows the ranking for a  $VH$  resonance mass hypothesis of 500 GeV, (b) shows the ranking for the 2000 GeV mass hypothesis. The yellow bands show the influence of the respective NP on  $\mu$  before the fit.

### 4.9.3 Nuisance parameter correlations

To understand the behaviour of the nuisance parameters in the fit it can be helpful to consider correlations between them and study if they behave as expected.

Figure 4.14 shows the correlation matrix for the fit in the 0-lepton channel. Only nuisance parameters that have a correlation of at least 35% are included in the matrix. Comparatively large correlations include e.g. the uncertainty related to the  $b$ -tagging efficiency and the normalization of the  $t\bar{t}$  background. This is expected since there are real  $b$ -jets among the signal jet candidates from  $t\bar{t}$  event. Therefore, if the  $b$ -tagging efficiency increases within its uncertainty, so does the  $t\bar{t}$  background normalization. A smaller correlation of the  $b$ -tagging efficiency uncertainty with the  $Z$ +heavy-flavour-jets background can be observed for the same reason, also between the  $c$ -tagging uncertainty and the  $Z + c/b$  background. The largest anti-correlation is observed between nuisance parameters describing the jet pile-up and flavour composition. Also visible is the anti-correlation between the large- $R$  jet energy and mass scale (JET\_Rtrk\_comb\_Kin) with the ratio of the merged and resolved category. This is expected since the change of large- $R$  jet energy scale only affects the merged category while the resolved category is only indirectly affected because of the PRSR categorization scheme.

### 4.9.4 Breakdown of statistical and systematic uncertainties

It was already mentioned that one of the challenges of the  $VH$  resonance is the fact that it covers a very wide mass range. Especially in the high-mass regime it is expected that the dominating uncertainty will

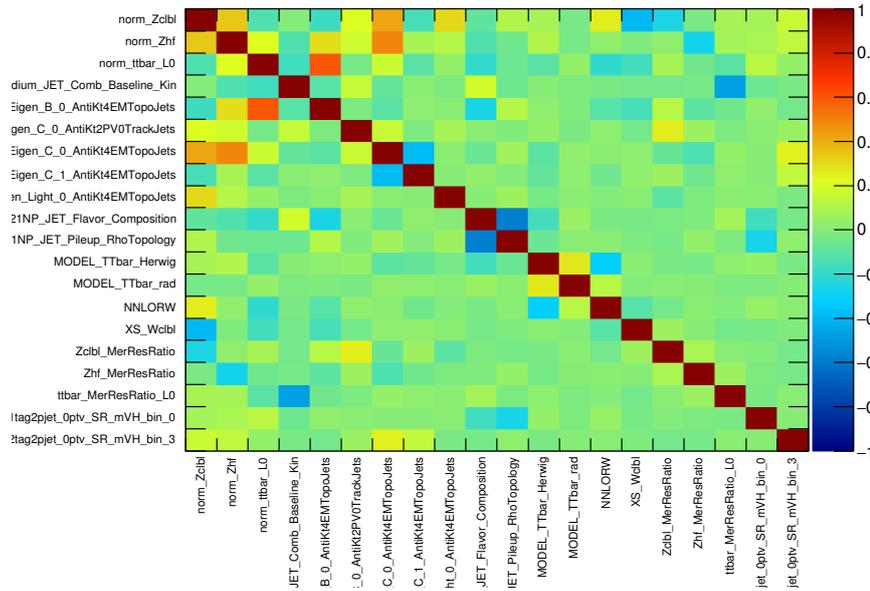


Figure 4.14: Nuisance parameter correlations for the 0-lepton channel in the final fit configuration. Only correlations above 35% are shown.

be the statistical uncertainty, since not many background events are expected at high mass.

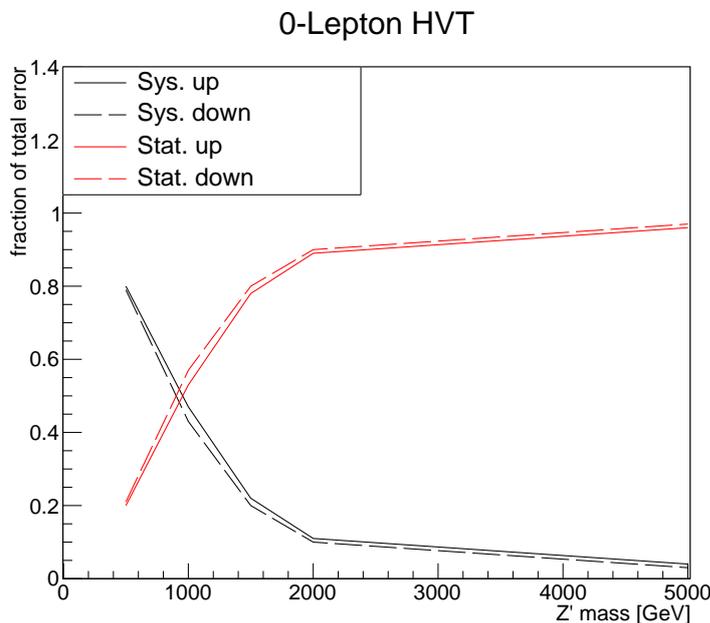


Figure 4.15: Expected fractional impact of the statistical and systematic uncertainties on the total uncertainty on the best-fit signal strength  $\hat{\mu}$

Figure 4.15 shows a breakdown of the fractional impact of the statistical and systematic uncertainties on the total uncertainty on the best-fit signal strength  $\hat{\mu}$  for different  $Z'$  resonance mass hypotheses. It can

be seen that below a resonance mass of  $m'_{Z'} = 1$  TeV the systematic uncertainties dominate. For resonance masses above the crossing point at 1 TeV the statistical uncertainties are dominating the total uncertainty. Therefore one can conclude that the  $VH$  resonance search analysis can still benefit, especially in the high-mass regime, from studying a larger dataset.

#### 4.9.5 Validation regions

As mentioned in section 4.8.1 the  $m_{bb}$  sideband regions as well as the top-background-sensitive regions with an additional  $b$ -jet outside of the large- $R$  jet are not used as control regions in the 0-lepton fit, but can be used to validate the post-fit agreement between data and simulation. For this validation study, the background correction factors determined in the fit to the signal regions are extrapolated to the side-band and additional jet regions. If the post-fit agreement between data and simulation in the validation regions is good, this serves as an indicator that the fit is able to describe the SM background expectation.

As can be seen from figures 4.16 and the agreement between data and SM background prediction in the validation regions is good. Some deviations are visible for the low-mass regime of merged 1- $b$ -tag side-band region. This is due to the fact that the yield in the low-mass region is dominated by the resolved categories, so the constraint on the background normalizations are driven by these categories.

### 4.10 Combined analysis of 0-, 1- and 2-lepton channels

As already mentioned in 4.8, several scenarios for the production of a new heavy vector resonance decaying to  $VH$  are studied within this analysis. Among these scenarios is the one where the  $Z'$  and  $W'$  are mass-degenerate. To study this case, all three lepton channels are combined in order to obtain the optimal sensitivity. In addition, the case where  $Z'$  and  $W'$  have different masses is studied. For this scenario a combination of the 0- and 2-lepton channels is performed for the  $Z'$  signal hypothesis, as well as a combination of the 0- and 1-lepton channels for the  $W'$  signal hypothesis. In the following chapters the focus will lie on the  $Z'$  combination as well as on the full HVT  $V'$  combination, because they were developed and optimized in the course of this thesis work. More details on the  $W'$  combination can be found in [77].

Whereas the 2-lepton channel has only a contribution from the  $Z'$  signal, both the 0-lepton and the 1-lepton channel have contributions from both the  $W'$  signal as well as the  $Z'$  signal. For the 0-lepton channel the contribution from  $W' \rightarrow WH$  contains mainly events where the  $W$ -boson decays further into a hadronically decaying  $\tau$  lepton. For the 1-lepton channel, the contribution from the  $Z' \rightarrow ZH$  mostly are 2-lepton events where one lepton fails the selection criteria. For the  $Z'$  combination the contribution from  $\tau$  leptons is removed by applying a veto on hadronically decaying tau leptons (cf. 4.2). The input channels and signal contributions for the three different combined fits are summarized in table 4.14.

Figure 4.17 shows the contribution of the  $W'$  and  $Z'$  signals to the 0-lepton channel expected limit, with and without the veto on tau leptons applied. As can be seen the contribution from the  $W'$  signal is small compared to the contribution from  $Z' \rightarrow ZH \rightarrow \nu\nu + b\bar{b}$ .

The combination is performed after the fit model optimization for the individual lepton channels. This means that the probability density functions (PDF) for signal and background templates of the individual channels are combined. It is very important that the individual channels are orthogonal (no overlap of events between lepton channels, ensured by the selection criterion on the number of loose leptons) such

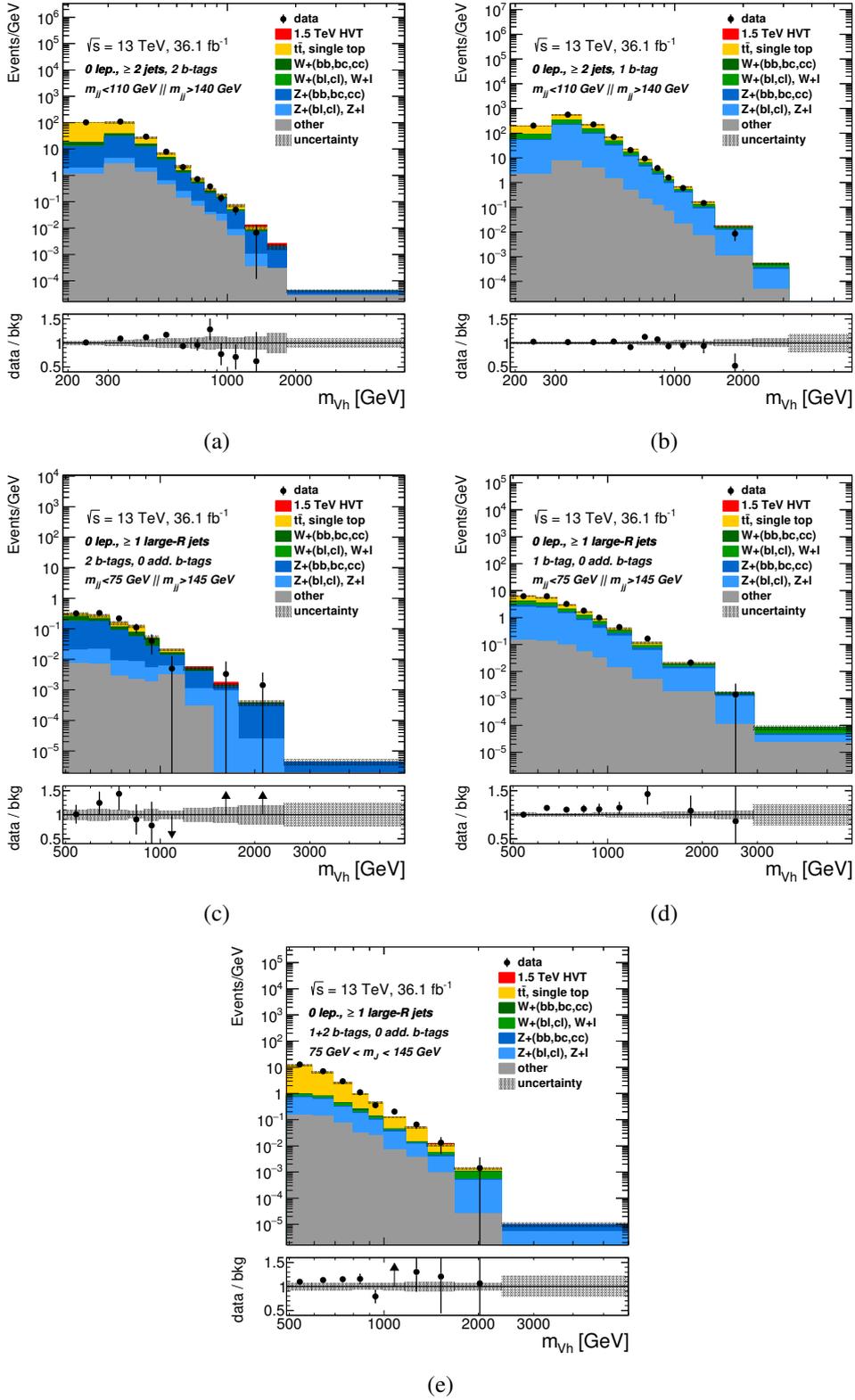
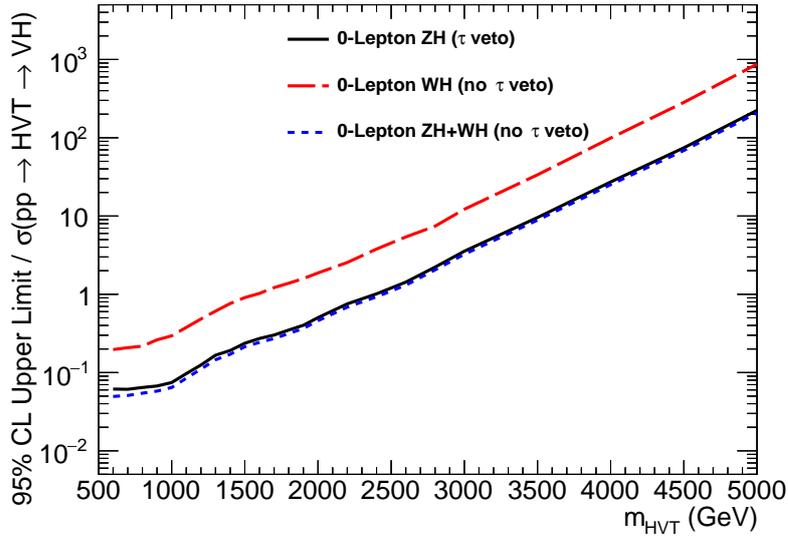


Figure 4.16: Post-fit distribution of the mass of the reconstructed  $VH$  system in the 0-lepton channel in the resolved ((a) with 2  $b$ -tags, (b) with 1  $b$ -tag)  $m_{bb}$  side-band regions as well as the top-background-sensitive region with an additional jet ((c) with 2  $b$ -tags, (d) with 1  $b$ -tag)  $m_{bb}$  side-band regions as well as the top-background-sensitive region with an additional jet ((e) merged signal mass region with  $\leq 1$   $b$ -tags).

	0-lepton channel	1-lepton channel	2-lepton channel
<b>Z' combination</b>	Z' signal with $\tau$ -veto	-	Z' signal
<b>W' combination</b>	W' signal without $\tau$ -veto	W' signal	-
<b>HVT combination</b>	W' & Z' signals without $\tau$ -veto	W' & Z' signals	Z' signal

Table 4.14: Summary of the channels and signals used in the combinations.


 Figure 4.17: Comparison of the expected limits for the 0-lepton channel with or without  $\tau$ -veto as well as with or without contribution from the  $W' \rightarrow WH$  decay.

that the combined PDF can be determined from the product of the individual channel PDFs. The main development task for the combined analysis is then how to treat nuisance parameters that are correlated between the input channels.

All NP corresponding to experimental systematic uncertainties, such as the jet energy scale uncertainty, are correlated between channels. For the NP regulating the background normalizations, the correlation scheme has to be decided based on physics arguments. In some cases, the NP governing the normalizations are treated fully decorrelated between the lepton channels. In other cases, additional degrees of freedom are introduced via constrained NP to allow the fit to adjust the background normalizations slightly differently in the individual channels. Figure 4.18 shows the NP pull distribution for the NP responsible for adjusting the dominant background normalizations ( $t\bar{t}$  and  $V$ +jets) for the HVT combination and the individual lepton channels.

In the following, the correlation scheme for the nuisance parameters responsible for adjusting the dominant background normalizations will be explained:

- $t\bar{t}$ : The  $t\bar{t}$  cross-section is treated using three separate floating nuisance parameters in the three

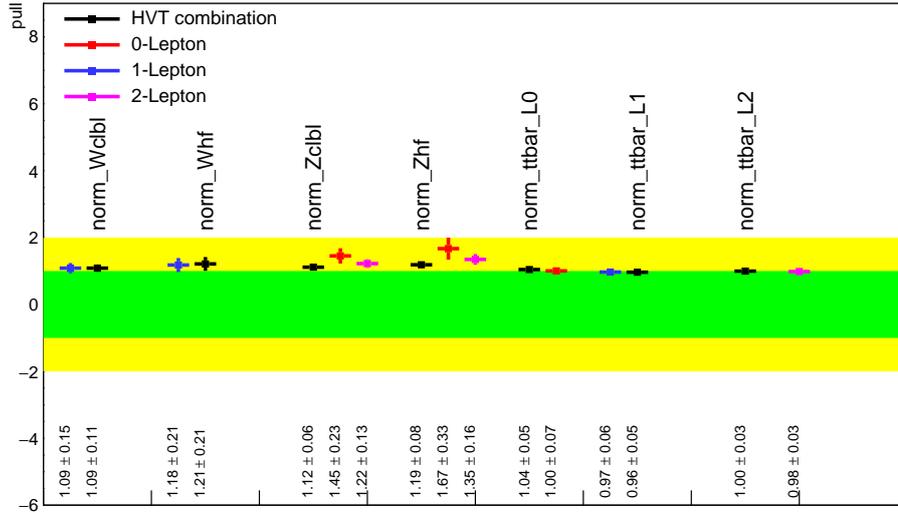


Figure 4.18: Comparison of the floating normalization Nuisance Parameter pulls for the 0-, 1- and 2-lepton individual fit and the HVT combination.

lepton channels ( $\text{norm\_ttbar\_L0}, *_L1, *_L2$ ). This choice can be motivated by the fact that the  $t\bar{t}$  normalization can be well constrained in all three lepton channels, indicated by the small error bars in figure 4.18. A further motivation for keeping the normalizations separate is the fact that the  $t\bar{t}$  contributions to the 0-, 1- or 2-lepton channels come from different processes (dileptonic  $t\bar{t}$ , semi-leptonic  $t\bar{t}$ , hadronic  $t\bar{t}$ ) and different regions of phase space. Additionally all nuisance parameter regulating the ratio of the  $t\bar{t}$  background between merged and resolved channel are decorrelated for all three channels.

- $Z + (bb, cc)$  and  $Z + (bl, cl)$ : The SM background of a jet associated either with two heavy-flavour jets or one heavy-flavour and one light jet is well constrained in the 0- and 2-lepton channels, particularly in the 2-lepton channels, as can be seen in figure 4.18 ( $\text{norm\_Zc1b1}, *_Zhf$ ). The 1-lepton channel does not provide a significant constraint on both normalizations. The process and phase space from which the  $Z + (bb, cc)$  and  $Z + (bl, cl)$  events originate are similar. Therefore a single common floating nuisance parameter is used across all three channels to describe the normalization of the two processes. To allow for additional freedom to adjust the normalization in the less-sensitive 0-lepton channel a constrained acceptance ratio between 0- and 2-leptons is introduced. All acceptance ratio nuisance parameters, such as the acceptance ratio between the merged and resolved region are treated fully correlated between the two channels.
- $Z + l$ : The contribution from  $Z$ +jets events with two light jets is small in all three lepton channels due to the  $b$ -tagging requirements, as can be seen in figures 4.20 and table 4.4. Therefore a common constrained nuisance parameter is used in all three channels to describe the normalization of this background process in the fit. Additionally, a constrained acceptance ratio parameter is introduced in the 0-lepton channel to allow the  $Z + l$  normalization to differ between 0-, and 2-lepton channel.
- $W + (bb, cc)$  and  $W + (bl, cl)$ : For the  $W$ +jets process with either two heavy-flavour or one heavy-flavour and one light-flavour jet the constraint is mainly driven by the 1-lepton channel, as can be seen in figure 4.18 ( $\text{norm\_Wc1b1}, *_Whf$ ). Therefore the normalization of this process is described by a common floating normalization parameter in all three lepton channels. A constrained nuisance parameter describing the acceptance ratio between 0- and 1-lepton as well as between

0- and 2-lepton channel is added in the 0-lepton channel. All other acceptance ratio nuisance parameters are correlated between channels in the same way as for  $Z + (bb, cc)$  and  $Z + (bl, cl)$ .

- $W + l$ : The SM production of a  $W$  with two light-flavour jets yields a small contribution in all three channels. Therefore the normalization nuisance parameter treatment is similar as for  $Z + l$ , with an acceptance ratio parameter that allows for freedom for the normalization to differ between 0- and 1-lepton channel.

All other background normalizations are small in comparison to  $t\bar{t}$  and  $V$ +jets. They are described by log-normal constrained nuisance parameters which are treated as fully correlated between the different lepton channels.

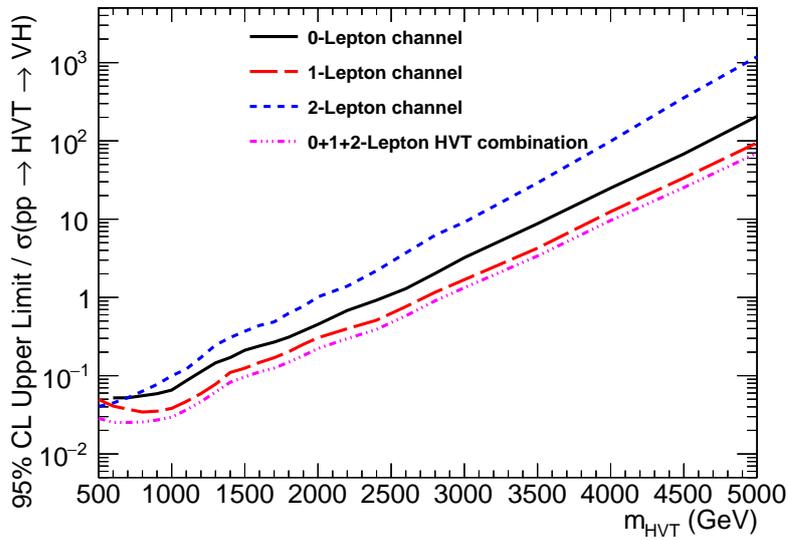


Figure 4.19: Comparison of the ratio of the expected HVT production cross section limits with respect to the theoretical prediction for the benchmark Model A, for the 0-, 1- and 2-lepton channel individual analyses as well as the combined HVT  $V'$  fit.

Figure 4.19 shows the ratio of the expected 95% CL upper limit on the HVT  $V'$  production cross section with respect to the theoretical prediction of the HVT Model A for the three individual lepton channels as well as their combination. It is visible that the 2-lepton channel dominates the sensitivity a low resonance mass hypotheses ( $m_{V'} < 600$  GeV). This is due to the excellent mass resolution of the 2-lepton channel, especially in the resolved category, which allows for a good discrimination between signal and SM background in the low-mass regime. At high masses the 2-lepton channel is the least sensitive of the three, because the efficiency to reconstruct two extremely high- $p_T$  leptons drops, which leads to a small signal acceptance for high resonance masses. In addition, because for high- $p_T$  leptons the tracks in the inner detector become almost straight, therefore the momentum resolution deteriorates, which leads also to a worse resolution of the invariant mass  $m_{VH}$ . Furthermore the branching ratios for the 2-lepton decay mode decrease for high mass. Both the 0-lepton channel and the 1-lepton channel retain a larger signal acceptance up to high masses. The 1-lepton channel is the most sensitive at high mass, because due to the presence of one lepton and the use of the  $W$  boson mass constraint the resolution of the invariant mass  $m_{VH}$  is higher than in the 0-lepton channel, which, due to the presence of only missing transverse energy to reconstruct  $m_{T,VH}$ , has a lower resolution. At low mass both the 0- and 1-lepton

channels are less sensitive than the 2-lepton channel, because both rely on the reconstruction of missing transverse energy which has a lower resolution and therefore also leads to a lower resolution of  $m_{VH}$ .

## 4.11 Results

After the statistical model is validated and well understood, the final fit model configuration is decided upon and the post-fit  $m_{VH}$  distributions in the signal regions can be studied in data. Figure 4.20 shows the transverse mass  $m_{T,VH}$  in the 0-lepton channel resolved and merged signal regions. The observed data agrees well with the simulated SM background prediction within the quoted uncertainties. More post-fit  $m_{VH}$  distributions in the 1- and 2-lepton channels can be found in figures A.1 and A.2.

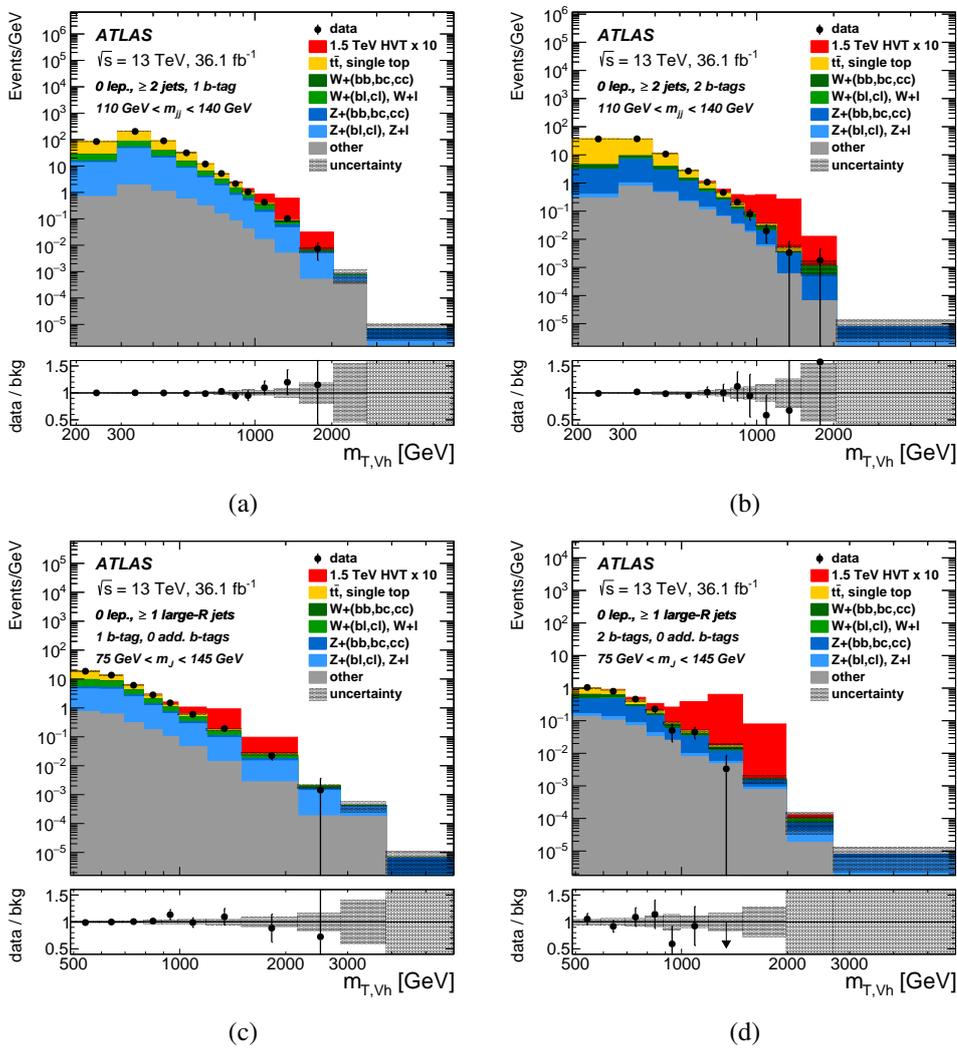


Figure 4.20: Post-fit distribution of the transverse mass of the reconstructed  $VH$  system in the 0-lepton channel in the merged and resolved signal regions (SR). (a): resolved SR with 1  $b$ -tag, (b): resolved SR with 2  $b$ -tags, (c): merged SR with 1  $b$ -tag, (d): merged SR with 2  $b$ -tags. From [77].

Since no significant excess of the data with respect to the estimated SM background is observed, limits at 95% confidence level on the Heavy vector triplet production cross-section times the branching ratio of the decay  $H \rightarrow b\bar{b} + c\bar{c}$  are set using a modified frequentist  $CL_s$  method with asymptotic formulae (see 3.7.5 ff.). The reason why the branching ratio for the decay of the Higgs boson both to  $b$ -quarks and  $c$ -quarks is considered is the significant fraction of charm quarks mis-identified as  $b$ -quarks, between 8.3% for  $R=0.4$  calorimeter jets and 14.2% for small- $R$  track jets [61, 64]. The HVT signal MC sample used in this analysis contains resonance decays with both  $H \rightarrow b\bar{b}$  and  $H \rightarrow c\bar{c}$  final states, and, even though the analysis is optimized for the  $H \rightarrow b\bar{b}$  case, there is a non-negligible fraction of  $H \rightarrow c\bar{c}$  final states.

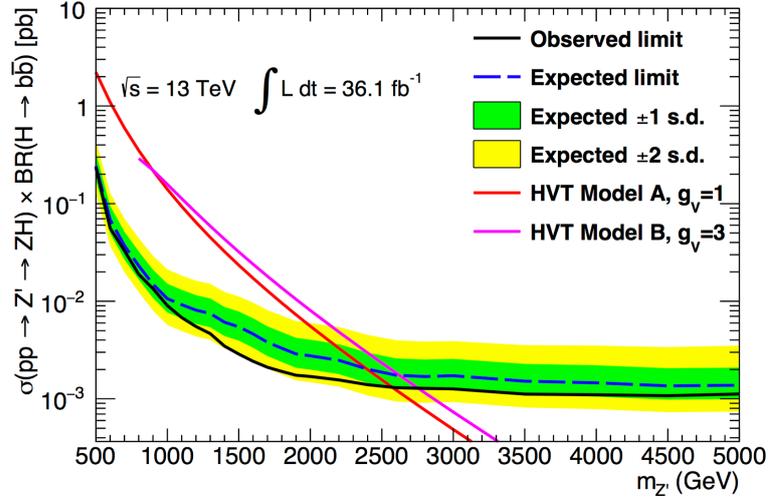


Figure 4.21: Expected and observed limit for  $Z'$  production for the 0-lepton channel individual fit. The dashed blue line shows the expected limit with the green and yellow bands representing the  $1\sigma$  and  $2\sigma$  uncertainty intervals. The solid black curve describes the observed limit. The red dashed and magenta solid line describe the theoretical prediction from the HVT model A and B respectively.

Figure 4.21 shows the expected and observed limit on the  $Z'$  production cross section times branching ratio based on the 0-lepton channel only. The observed limit is lower by  $1 - 2\sigma$  compared to the expected limit for resonance masses above 700 GeV. This is due to a downwards fluctuation of the observed data with respect to the SM background prediction. When comparing to the post-fit  $m_{T,VH}$  distributions in figure 4.20 there are bins with downward fluctuations visible in the data in the resolved 2-tag region as well as the merged region. Especially at high mass the low observed limit is caused by bins with no observed data events but with a small predicted SM background content.

To study the scenario where the  $Z'$  and  $W'$  resonances have different masses, limits are set on the production cross section of the  $Z'$  and  $W'$  resonances separately. The  $Z'$  and  $W'$  combinations are performed according to the description in 4.10. The limits for  $Z'$  and  $W'$  production are shown in figures 4.22 and 4.23 respectively. For the  $Z'$ , in the mass regime above 1 TeV the observed limit is lower than the expected limit, due to a statistical fluctuation of the data in the 0-lepton channel. From the intersection of the theoretical prediction with the observed limit curve one can deduce that the existence of a  $Z'$  resonance is excluded for  $m_{Z'} < 2.65$  TeV for the HVT Model A, and for  $m_{Z'} < 2.83$  TeV for Model B. For the  $W'$ , the mass hypotheses  $m_{W'} < 2.67$  TeV for HVT Model A and  $m_{W'} < 2.82$  TeV for Model B.

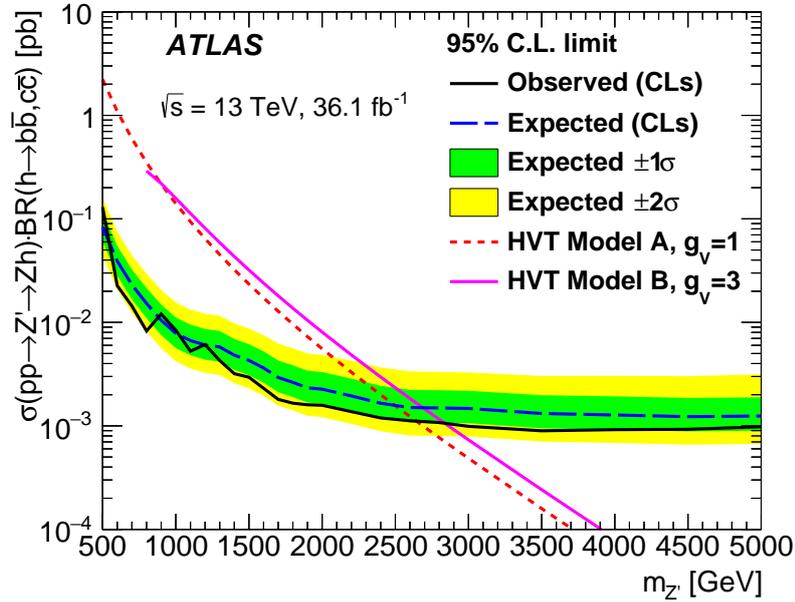


Figure 4.22: Expected and observed limit for the  $Z'$  combined fit. The dashed blue line shows the expected limit with the green and yellow bands representing the  $1\sigma$  and  $2\sigma$  uncertainty intervals. The solid black curve describes the observed limit. The red dashed and magenta solid lines describe the theoretical prediction from the HVT model A and B, respectively. From [77].

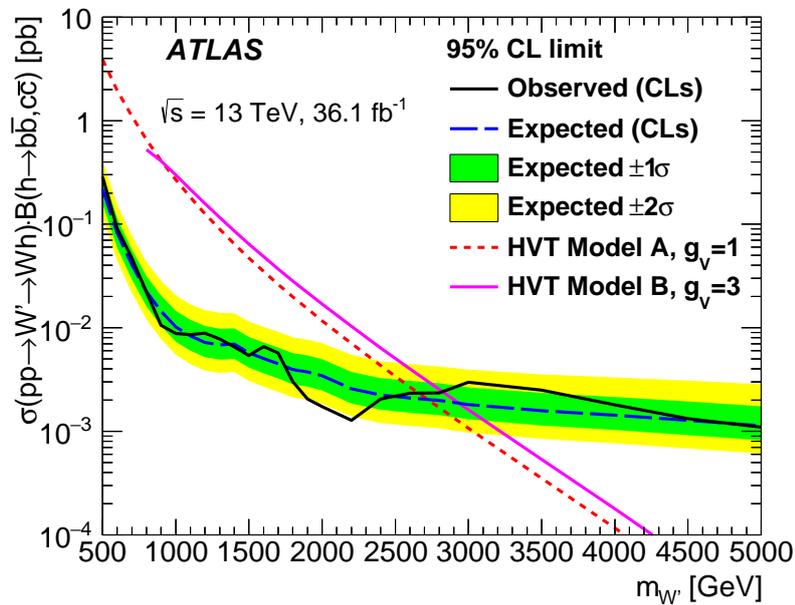


Figure 4.23: Expected and observed limit for the  $W'$  combined fit. The dashed blue line shows the expected limit with the green and yellow bands representing the  $1\sigma$  and  $2\sigma$  uncertainty intervals. The solid black curve describes the observed limit. The red dashed and magenta solid lines describe the theoretical prediction from the HVT model A and B, respectively. From [77].

Finally, the scenario is studied in which the  $W'$  and  $Z'$  resonances are mass-degenerate (figure 4.24). All three lepton channels are combined according to the procedure described in 4.10 under the assumption  $m_{Z'} = m_{W'}$ . The existence of a HVT resonance  $V'$  is excluded for a mass  $m_{V'} < 2.80$  TeV for the benchmark Model A and  $m_{V'} < 2.93$  TeV for the benchmark Model B.

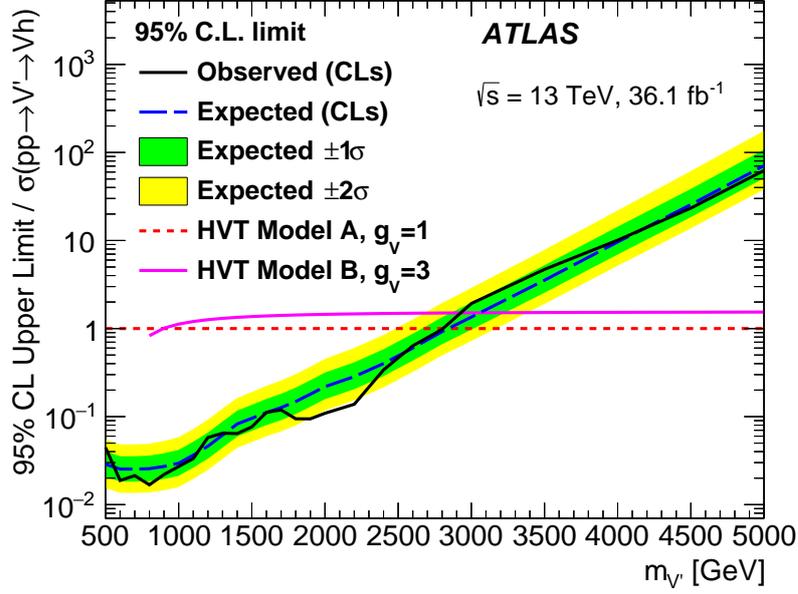


Figure 4.24: Expected and observed limit for the HVT  $V'$  combined fit. The dashed blue line shows the expected limit with the green and yellow bands representing the  $1\sigma$  and  $2\sigma$  uncertainty intervals. The solid black curve describes the observed limit. The red dashed and magenta solid lines describe the theoretical prediction from the HVT model A and B, respectively. From [77].

## 4.12 Interpretation within the HVT simplified model

The observed limits on the  $V'$  production cross section times branching ratio as shown in figure 4.24 can be translated into exclusion contours within the parameter space of the HVT model. Figure 4.25 shows the exclusion limit at 95% confidence level in the 2-dimensional HVT parameter plane  $\{g^2 c_F / g_V, g_V, c_H\}$  (see section 4.1) for  $V'$  resonance mass hypotheses of 1.2 TeV, 2.0 TeV and 3.0 TeV. All HVT models whose parameters lie outside of the exclusion contours, are excluded for the specific mass hypothesis. The violet points denote the parameter combinations that constitute the aforementioned benchmark model A and B, for different values of the coupling  $g_V$ . Figure 4.25 assumes a mass-degeneracy between the  $W'$  and  $Z'$ . Also it is assumed that the fermion coupling  $c_F$  is universal for all quark and lepton generations and that other coupling parameters concerning more than one new heavy vector boson, such as  $g_V c_{VVV}$ ,  $g_V^2 c_{VVHH}$  and  $c_{VW}$  do not contribute to the  $V'$  cross section in the considered decay channel (see also 4.1). This assumption can be motivated by the fact that the only way these couplings enter the considered decay scenario is via the mixing of the new heavy vector boson with the  $W$ , which can be shown to be a small contribution. [27]

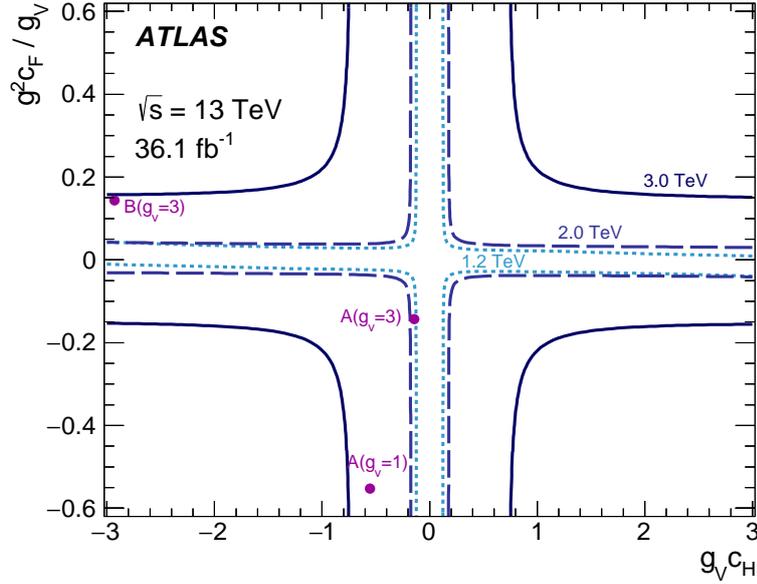


Figure 4.25: Observed exclusion contours within the HVT model parameter space for the HVT  $V'$  combined fit. The blue solid and dashed lines show the boundaries of regions in the parameter space that are excluded for different  $V'$  resonance mass hypotheses. The magenta points indicate the location of specific model parameter configurations corresponding to Model A and B with  $g_V = 1$  and  $g_V = 3$ . From [77].

### 4.13 Conclusions and prospects

In this chapter a search for a new heavy resonance  $V'$  in the decay channel  $V' \rightarrow VH \rightarrow \ell\ell, \ell\nu, \nu\nu + b\bar{b}$  with the ATLAS experiment in  $36.1 \text{ fb}^{-1}$  of LHC pp-collision data was discussed. This decay channel is interesting because it makes use of the SM Higgs boson as a probe for BSM physics. The final state with leptons and b-jets considered in this analysis is advantageous because of efficient triggering of high- $p_T$  leptons and the large Higgs boson branching fraction into b-quarks. The presented analysis covers a wide Higgs boson  $p_T$  range and, especially in the high  $p_T$  regime, makes extensive use of boosted Higgs boson identification methods, such as  $R=0.2$  track jet double- $b$ -tagging as well as the large- $R$  jet mass.

No significant excess of the observed data with respect to the Standard Model background was found, therefore limits at 95% C.L. are set, excluding the resonance hypothesis for various parameter configurations in framework of the HVT model for resonance masses  $m_{V'} \lesssim 2 - 3 \text{ TeV}$ .

For resonance mass hypotheses above 1 TeV the uncertainty of the presented analysis is still dominated by the data statistical uncertainty, therefore it would be desirable to perform a similar search analysis on a larger dataset. Figure 4.26 shows a projection of the expected limit for the full Run-2 dataset of  $140 \text{ fb}^{-1}$ , neglecting possible analysis improvements, updated reconstruction and higher simulation statistics. Figure 4.26 shows an expected improvement of about 50%, which is expected considering that the Full Run-2 dataset is about 3.8 times bigger than the dataset used in this analysis.

Further improvements to the presented analysis could include the use of more advanced Higgs boson identification algorithms (introduced in section 3.6.4) that retain a higher signal efficiency for very high- $p_T$   $H \rightarrow b\bar{b}$  decays. One promising method is the use of so-called *variable-R* (VR) track jets, whose

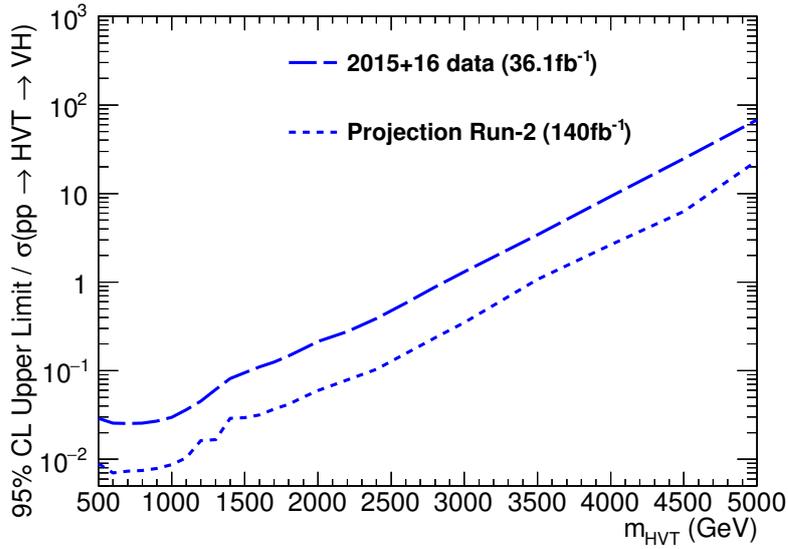


Figure 4.26: Expected limit for the HVT  $V'$  combined fit. The widely dashed blue line shows the expected limit with a luminosity of  $36.1 \text{ fb}^{-1}$ . The narrowly dashed line shows the projection of the expected limit with a luminosity scaled to  $140 \text{ fb}^{-1}$ . No new data or analysis improvements are taken into account.

jet cone radius shrinks as a function of the jet  $p_T$  and is therefore more resilient against the merging of fixed-R track jets at very high Higgs boson  $p_T$  [69]. At a resonance mass around 3 TeV, the use of VR track jets in the merged 2-tag category<sup>5</sup> is expected to gain a factor of two in signal efficiency while retaining a similar top-jet rejection.

Furthermore it became clear that in this analysis the flavour-tagging-related uncertainties are among the dominant systematic uncertainties. At high jet  $p_T$  the uncertainties are derived using MC-based extrapolation methods with large uncertainties and no data-based estimate is yet firmly established within *ATLAS*. A possible improvement to the presented resonance search analysis would be to improve the understanding of flavour tagging performance in data for high- $p_T$  and boosted close-by  $b$ -jet topologies. In the next chapter of this thesis the first direct data-based measurement of track-jet double- $b$ -tagging efficiencies in gluon splitting ( $g \rightarrow b\bar{b}$ ) events with the *ATLAS* experiment will be presented.

<sup>5</sup> The leading background contribution in the 2-tag merged category is from  $t\bar{t}$  events.

---

## Double- $b$ -tagging calibration in $g \rightarrow b\bar{b}$ events

---

The identification of boosted Higgs bosons decaying to  $b$ -quarks (Higgs tagging) is an important method for analyses searching for the decay of hypothetical heavy resonances into final states with Higgs bosons, or for studying single or double Higgs boson production at high transverse momenta. An example for a search analysis making heavy use of Higgs tagging techniques is the search for a heavy resonance  $V'$  in the decay channel  $VH(b\bar{b})$ , as presented in this thesis.

In order to apply a Higgs tagging algorithm in an analysis, it needs to be calibrated using real data, however no direct calibration method for Higgs tagging existed within the *ATLAS* experiment so far. The work presented in this chapter pioneered Higgs tagging calibration techniques in *ATLAS* by using events with gluons splitting into  $b$ -quark pairs ( $g \rightarrow b\bar{b}$ ) to study the close-by double- $b$ -jet topology in data.

One of the simplest ways to construct a Higgs tagging algorithm is to study the flavour content of the Higgs boson candidate jet. Double- $b$ -tagging with small- $R$  track jets is used to investigate whether the candidate jet contains the hadronization products of two  $b$ -quarks. A detailed introduction to Higgs boson identification and double- $b$ -tagging in *ATLAS* is given in 3.6.

In the following chapter, a method to calibrate double- $b$ -tagging directly using  $g \rightarrow b\bar{b}$  events will be introduced. A useful feature of this calibration analysis is that it provides a test bench to validate the agreement between data and simulation for jet properties used for Higgs tagging in an environment with two close-by  $b$ -jets. The validation of jet properties related to Higgs tagging was adapted for LHC collision data at 13 TeV by the author of this thesis and is documented in [53]. The development of the  $g \rightarrow b\bar{b}$  analysis into a full double- $b$ -tagging calibration is a novel approach for the *ATLAS* experiment and is documented for the first time in this thesis.

### 5.1 Calibration strategy

Double- $b$ -tagging relies on a specific event topology of a large- $R$  jet with two subjets, therefore a direct calibration of double- $b$ -tagging can only be performed using events with this topology. The aim of the study presented in this chapter is a direct data-based calibration of the *double- $b$ -tagging efficiency*, therefore one needs to select a data sample that contains boosted jets with two  $B$ -hadrons inside. A sample of such events that are produced with high rate at the Large Hadron Collider contain the decay of a gluon into two  $b$ -quarks, also referred to as *gluon splitting* or  $g \rightarrow b\bar{b}$ .

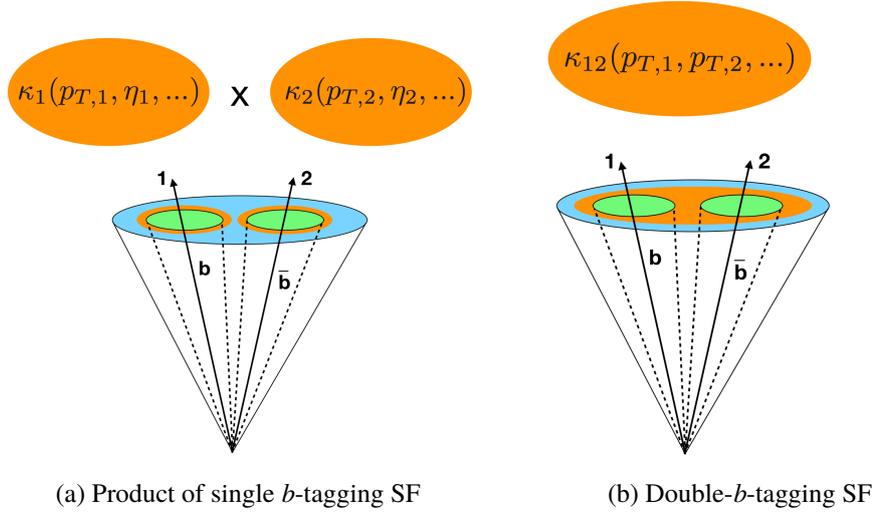


Figure 5.1: Different strategies to calibrate double- $b$ -tagging. Figure (a) shows the standard approach where two single- $b$ -jet scale factors are multiplied, (b) shows the new approach to determine a direct double- $b$ -tagging scale factor.

Figure 5.1 illustrates the difference between the standard *ATLAS* approach for calibrating double- $b$ -tagging of track jets and a direct double- $b$ -tagging calibration. The standard approach in *ATLAS* is to apply two single track jet scale factors<sup>1</sup>, derived for isolated jets, via an event weight calculated by multiplying the two track jet  $b$ -tagging scale factors and treating them as uncorrelated. In the direct calibration one double- $b$ -tagging scale factor is applied which depends on the correlated kinematic properties of both track jets.

The advantage of calibrating double- $b$ -tagging directly on a sample of boosted di- $b$ -jets is that all correlations between the two  $b$ -jets and topology-specific properties of the events are taken into account, which is not the case in the standard approach. Also the new method allows to perform the calibration using real data in the regime of extremely high momenta beyond a track jet transverse momentum of  $p_{T,\text{track jet}} \geq 250$  GeV. The standard  $b$ -tagging calibration analyses only give MC-based extrapolation uncertainties of the scale factors in this high- $p_T$  regime. In addition, the new calibration method provides a natural approach to determine scale factors for more complicated  $b$ -tagging approaches used in boosted Higgs boson identification, for example using a so-called *asymmetric  $b$ -tagging* approach where different  $b$ -tagging operating points are used for the two sub-jets (see section 3.6.2). When using the standard single- $b$ -tagging scale factors in this case, the treatment of the uncertainties becomes very difficult, because they are correlated between the looser and tighter  $b$ -tagging working points. The correlation between the uncertainties comes about because the calibration of the tighter working point would use a subset of the same data and MC as the calibration for the looser working point. The calibration of a Higgs boson identification algorithm using asymmetric  $b$ -tagging in  $g \rightarrow b\bar{b}$  events provides scale factors and their uncertainties in a simple, monolithic way.

<sup>1</sup> For an introduction to *ATLAS* flavour tagging calibrations, see section 3.5.9.

## 5.2 Object selection

In this section the object selection for the  $g \rightarrow b\bar{b}$  calibration analysis is described. There is a large overlap with the selection for the VH resonance analysis described in section 4.2. References are added where applicable.

### Jets

Large- $R$  and small- $R$  calorimeter jets as well as small- $R$  track jets for the  $g \rightarrow b\bar{b}$  calibration analysis are reconstructed in the same way as for the VH resonance search analysis (see section 4.2 and table 4.2). The large- $R$  jet is required to have at least two ghost-associated (see section 3.4.5) small- $R$  track jets.

### Flavour tagging

The double- $b$ -tagging strategy chosen for the  $g \rightarrow b\bar{b}$  calibration is the same as for the merged category of the VH resonance search analysis presented in chapter 4, section 4.5.3. The 70%  $b$ -jet efficiency working point of the MV2c10 algorithm [61, 64] is used on the small- $R$  track jets associated to the large- $R$  jet.

### Muons

Muons are reconstructed in the same way as for the VH resonance analysis, however the selection requirements are different in this analysis, because in the  $g \rightarrow b\bar{b}$  calibration analysis muons are used to identify semi-leptonic  $B$ -hadron decays within jets. A series of requirements on the muon ID track are applied, as well as at least Loose quality requirements, defined in [81]. The quality requirements for muons are summarized in table 5.1, as recommended by the ATLAS collaboration [81]. The quality requirements are inclusive, meaning that all medium and tight muons are also loose muons. The muon inner detector track selection is summarized in table 5.2. These requirements are intended to remove contributions from pile-up events and are applied to all muon types and qualities [81]. Furthermore, muons are required to pass a selection on the transverse momentum  $p_T > 5$  GeV and the pseudorapidity  $|\eta| < 2.5$ .

### Track-to-jet association and flavour labelling

A central part of the presented analysis is a template fit of a track-based flavour sensitive variable  $\langle s_{d_0} \rangle$  (see 5.4), therefore tracks are associated to the small- $R$  track-jets. Since the analysis is intended for a flavour tagging efficiency calibration the track selection is kept similar to tracks used as input to the  $b$ -tagging algorithms. As a result the same track association scheme as for flavour tagging is adopted, which is described in more detail in 3.5.2.

It is necessary to define the true flavour of a jet to define flavour tagging efficiencies in MC. In the  $g \rightarrow b\bar{b}$  analysis the exclusive  $\Delta R$  cone-based association of hadrons from the MC truth record, described in 3.5.6 is used for truth-labelling jets. Only the truth flavour labels of the small- $R$  track-jets are considered.

### Track selection

In addition to the  $\Delta R$  requirements for the track-to jet association, tracks are required to pass the loose track selection requirement defined in [88]:

Quality	Muon Type	Requirements
<b>Tight</b>	Combined	# precision layer hits >1 combination $\chi^2/NDF < 8$ $p_T$ - & $\eta$ -dependent cuts on $q/p$ -significance $p_T$ - & $\eta$ -dependent cuts on ID/ME/CB momentum imbalance
<b>Medium</b>	Combined	$q/p$ -significance < 7 # precision layer hits >1, <b>or</b> (# precision layer hits = 1 and # holes in precision layer < 2 and $ \eta  < 0.1$ )
	StandAlone	$ \eta  > 2.5$ #precision layer hits >2
<b>Loose</b>	CaloTagged	$ \eta  < 0.1$ CaloLRLikelihood > 0.9 <b>or</b> CaloMuonIDTag > 10
	StandAlone	$ \eta  < 0.1$

 Table 5.1: Summary of muon requirements used for the  $g \rightarrow b\bar{b}$  analysis [81].

ID Track Requirements for Muons
# pixel hits+ # crossed dead pixel sensors > 0
# SCT hits+ # crossed dead SCT sensors > 4
# pixel holes + # SCT holes < 3
successful TRT extension where expected: 1.1 < $ \eta $ < 1.9: $n > 5$ and $n_{\text{TRT}}^{\text{outliers}} < 0.9n$ where $n = n_{\text{TRT}}^{\text{hits}} + n_{\text{TRT}}^{\text{outliers}}$
longitudinal impact parameter $z_0 \sin(\theta) < 2$ mm
transverse impact parameter of ID track: $d_0 < 2$ mm

 Table 5.2: Summary of Muon ID track requirements used for the  $g \rightarrow b\bar{b}$  analysis [81].

Loose Track Requirements
$p_T > 400$ MeV
$ \eta  < 2.5$
# silicon hits (Pixel + SCT) $\geq 7$
# shared modules $\leq 1$
# silicon holes $\leq 2$
# pixel holes $\leq 1$

 Table 5.3: Summary of LOOSE ID track requirements used for the  $g \rightarrow b\bar{b}$  analysis. [88]

### 5.3 Event selection

Events considered in this analysis are required to have a hard-scatter primary vertex with at least two tracks with  $p_T > 500$  MeV [89]. In case there is more than one primary vertex candidate, the one with the highest  $\sum p_T^2$  of the associated tracks is used. To reject non-collision backgrounds such as calorimeter noise, beam halo interactions or cosmic rays, a standard jet cleaning procedure is used [90].

In order to select boosted  $g \rightarrow b\bar{b}$  topologies, events are required to contain at least two jets, one of

which is the  $g \rightarrow b\bar{b}$  candidate and the other is a recoil jet that ensures a sufficiently high momentum of the  $g \rightarrow b\bar{b}$  candidate jet and also is used to trigger the events when recording data with the *ATLAS* detector. The  $g \rightarrow b\bar{b}$  candidate jet is a trimmed large- $R$  calorimeter jet, with the selection criteria described in 5.2. The recoil jet is a small- $R$  calorimeter jet. The reason for using single small- $R$  jet triggers and therefore small- $R$  calorimeter jets as proxies for the recoil jet is mainly historical, because no large- $R$  triggers were available in the reduced datasets used for the first implementation of the calibration. A sketch of the  $g \rightarrow b\bar{b}$  event topology is shown in figure 5.2.

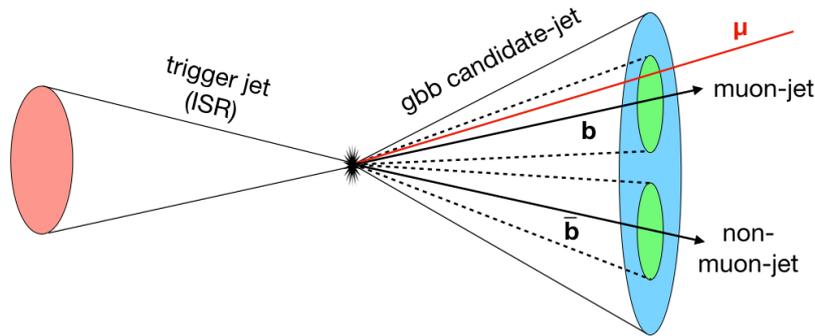


Figure 5.2: Sketch of the  $g \rightarrow b\bar{b}$  event topology.

To select di-jet events in data, a single jet trigger with an online  $E_{T,\text{jet}}$  threshold of 380 GeV is used [91]. This was chosen, because it is the lowest un-prescaled single jet trigger in the *ATLAS* data-taking periods of the years 2015 and 2016. It is required that the recoil offline small- $R$  calorimeter jet is matched to the online jet by requiring that the online trigger jet axis lies within  $\Delta R < 0.4$  of the leading offline small- $R$  calorimeter jet axis. To ensure that the plateau of the trigger efficiency is reached, a cut on the transverse momentum of the offline recoil jet  $p_T > 500$  GeV is imposed. The plateau of the trigger efficiency should be reached to avoid including events from the trigger-turn-on region for which there are known discrepancies between data and simulation, which would require a special  $p_T$ -dependent jet trigger calibration. In the plateau region just one jet trigger calibration scale factor is needed to adjust the total normalization. A cut on the large- $R$  jet transverse momentum in the range of  $p_T > 450$  GeV to  $p_T > 500$  GeV is also commonly used in other analyses using boosted Higgs or W boson tagging techniques [92, 93].

The  $g \rightarrow b\bar{b}$  candidate jet corresponds to the leading large- $R$  jet with two small- $R$  track jets associated to it as described in section 5.2. In addition it is required that at least one of the small- $R$  track jets has a muon within a  $\Delta R < 0.2$  of the track jet axis. The presence of a muon within the track jet is required to enrich the selected multijet sample in events with b-jets, because semi-leptonic  $B$ -hadron decays produce muons within jets. In principle, also electrons in jets are an indicator for semi-leptonic  $B$ -hadron decays, however it is difficult to distinguish between non-isolated electrons and hadronic jets, therefore electrons are not used in this analysis. The leading track jet with an associated muon is in the following referred to as *muon-jet* ( $\mu$ -jet). The leading jet among the list of remaining track jets associated to the large- $R$  jet is referred to as *non-muon-jet* ( $!\mu$ -jet), regardless of whether it contains a muon or not.

Only events with a  $g \rightarrow b\bar{b}$  candidate jet  $p_T > 500$  GeV are considered, in order to avoid a bias due to the mis-match of the jet areas between the online jet used for triggering ( $R = 0.4$ ) and the large- $R$  jet ( $R = 1.0$ ). For a more detailed description on this requirement, see section A.2.4.

The main focus of this analysis lies on the flavour composition of the  $g \rightarrow b\bar{b}$  candidate large- $R$  jet. The flavour composition is defined as the flavour of the partons that initiated the muon- and non-muon subjets. In simulations this information can be accessed from the MC truth record, where the hadron- and cone-based exclusive truth labelling strategy as described in section 3.5.6 is applied. In the following, the flavour content of the large- $R$  jet will be encoded in a double-label of the structure  $F_\mu F_{\bar{\mu}}$ , where  $F_\mu$  is the flavour truth label of the muon-track-jet and  $F_{\bar{\mu}}$  is the flavour truth label of the non-muon-track-jet. Both  $F_\mu$  and  $F_{\bar{\mu}}$  are elements of the truth label set (B,C,L), resulting in a total of nine flavour combinations. It was found that the contribution from jets produced by hadronic  $\tau$  decays can be neglected.

## 5.4 Flavour fraction fit

Since the double- $b$ -tagging scale factor should encode only the difference in double- $b$ -tagging efficiency between data and simulation, all other sources of differences should be corrected beforehand. It is a known fact that the currently available multijet simulations do not provide adequate modelling of the relative fractions of the large- $R$  jet flavour content. To correct for this mis-modelling, a data-based normalization correction is applied, based on a template fit of the flavour-sensitive variable  $\langle s_{d_0} \rangle$ . In the following, this procedure is referred to as the *flavour fraction correction*.

The variable  $\langle s_{d_0} \rangle$  is defined as follows: From the tracks that are associated to the small- $R$  track jets within the large- $R$  jet, as described in section 5.2, the three tracks with the highest track transverse momenta are selected and the average of their transverse impact parameter (IP) significance (defined in equation 3.19) is calculated. For more detail on the calculation of the transverse IP significance, see section 3.5.3. The distribution of the variable  $\langle s_{d_0} \rangle$  for different jet flavour pairs is shown in figure 5.3.

The transverse IP significance is large for well-measured tracks with a large IP. This is typically the case for tracks originating from the decay of long-lived  $B$ -hadrons. In other related analyses [94, 95] the  $s_{d_0}$  of the most significant, or second most significant track among those associated to the track jet is chosen as flavour-sensitive variable, however the average among the three leading tracks is better modelled in simulation.<sup>2</sup> The three tracks leading in  $p_T$  are chosen, since tracks from the  $B$ -hadron decay typically carry a large fraction of the initial  $b$ -quark momentum.

Different types of events contribute in different regions of the signed transverse impact parameter significance distribution. To illustrate the different components, figure 5.4 shows the  $s_{d_0}$  distribution for tracks from  $B$ - and  $C$ -hadrons as well as primary and secondary tracks from light-flavour jets. One can discern three main components in 5.4: Firstly, the  $s_{d_0}$  of prompt tracks in light-flavour jets follows an approximately Gaussian distribution. Since prompt tracks stem from the primary vertex, the distribution is symmetric around zero, with positive and negative tails due to the limited resolution of the transverse impact parameter measurement. Tracks from real  $C$ - and  $B$ - hadrons have a long tail towards positive significances, because of the long lifetime of heavy-flavour hadrons. Since the  $B$ - and  $C$ - hadrons possess a large fraction of the initial parton momentum, their flight direction and the resulting secondary vertex position are correlated with the jet axis, leading to predominantly positive lifetime signs. This

<sup>2</sup> For more detail on alternative variables see A.2.1

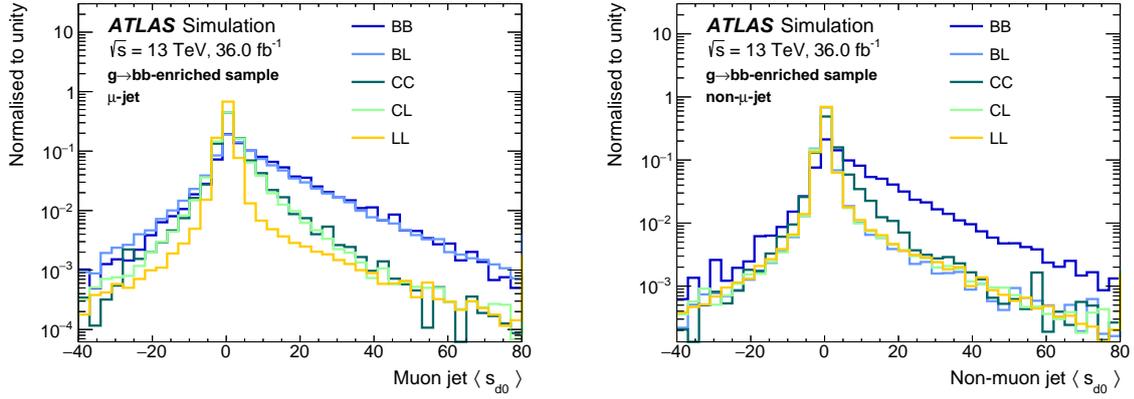


Figure 5.3: Averaged transverse impact parameter significance,  $\langle s_{d_0} \rangle$ , distributions for the muon (left) and non-muon jets (right). The double flavour labels denote the flavour of the jet pair, with the muon-jet stated first. From [53].

situation is different for light long-lived particles and, especially, photon conversions or hadronic material interactions occurring in light-flavour jets. Their flight direction is not correlated, or, in case of long-lived light hadrons, much less correlated with the jet axis. Therefore even though there are still tails towards large impact parameter significances, the probability to obtain a positive or negative lifetime-sign is approximately equal, leading to a more symmetric distribution with large tails both in the positive and negative regime. The total light jet  $s_{d_0}$  distribution is an admixture of prompt tracks, tracks from light long-lived hadrons and photon conversion, leading to a narrower  $s_{d_0}$  distribution than for heavy-flavour jets, but with enhanced tails.

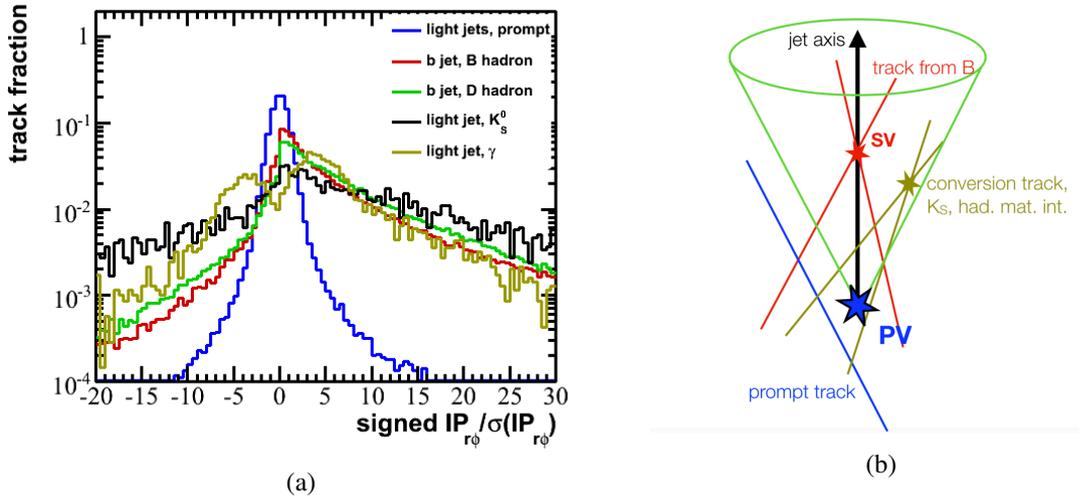


Figure 5.4: a) Transverse signed impact parameter significance distributions for tracks from  $B$ - and  $D$ -hadrons as well as primary and secondary tracks in light jets. From [96]. (b) illustration of distributions of tracks within the jet for track categories used in (a).

Since the flavour composition varies as a function of the track jet transverse momenta, the template fit is performed in twelve two-dimensional bins of muon-jet  $p_T$  (denoted as  $p_T(j_\mu)$ ) and non-muon jet  $p_T$

(denoted as  $p_T(j_\mu)$ ). The bin boundaries are summarized in table 5.4.

fit category	
muon-jet $p_T/\text{GeV}$	non-muon-jet $p_T/\text{GeV}$
$p_T(j_\mu) < 100$	$p_T(j_{\bar{\mu}}) < 100$
	$100 < p_T(j_{\bar{\mu}}) < 200$
	$200 < p_T(j_{\bar{\mu}}) < 300$
	$p_T(j_{\bar{\mu}}) > 300$
$100 < p_T(j_\mu) < 200$	$p_T(j_{\bar{\mu}}) < 100$
	$100 < p_T(j_{\bar{\mu}}) < 200$
	$200 < p_T(j_{\bar{\mu}}) < 300$
	$p_T(j_{\bar{\mu}}) > 300$
$p_T(j_\mu) > 200$	$p_T(j_{\bar{\mu}}) < 100$
	$100 < p_T(j_{\bar{\mu}}) < 200$
	$200 < p_T(j_{\bar{\mu}}) < 300$
	$p_T(j_{\bar{\mu}}) > 300$

Table 5.4: Two-dimensional track jet  $p_T$  bins used for the flavour fraction fit.

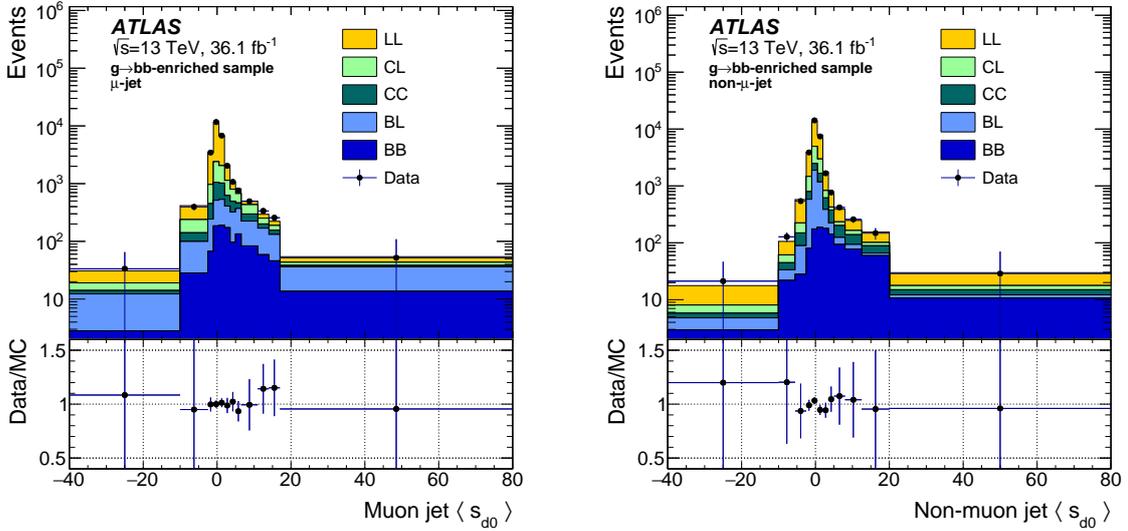


Figure 5.5: Distributions of the average transverse IP significance  $\langle s_{d_0} \rangle$  for the muon and non-muon jet in (100, 200) GeV bin of the muon jet and non-muon jet  $p_T$ 's. From [53].

As mentioned before, each flavour component is named after the truth label of the respective muon- and non-muon track jet among ( $B$ ,  $C$ ,  $L$ ). This allows for a total of 9 different components, however the components  $LC$ ,  $BC$ ,  $CB$  and  $LB$  are very small ( $< 1\%$ ), therefore they are merged into one of the larger components, namely the one with most similar  $\langle s_{d_0} \rangle$  template shape. The shape-similarity between two templates is estimated using the integral measure:

$$S = \frac{1}{2} \int \frac{(p_1(x) - p_2(x))^2}{p_1(x) + p_2(x)} dx. \quad (5.1)$$

A small value of  $S$  is obtained for very similar distributions  $p_1(x)$  and  $p_2(x)$ . For each of the small-yield components, the most similar component with larger yield is determined based on the minimal value of  $S$  (tables in A.2.2). As a result, the merging scheme  $LB, LC \rightarrow LL$ ,  $BC \rightarrow BL$ ,  $CB \rightarrow CL$  is chosen. A total of 5 flavour components remains, namely  $BB$ ,  $BL$ ,  $CC$ ,  $CL$  and  $LL$ .

The normalization of each of the five flavour component is adjusted based on the simultaneous fit of templates for the  $\langle s_{d_0} \rangle$  of the muon- and non-muon track jet to data <sup>3</sup>. It is possible to fit the muon-jet and non-muon-jet distributions simultaneously because the  $\langle s_{d_0} \rangle$  are uncorrelated between the two track jets (see section A.2.3). The largest correlation between muon-jet and non-muon jet  $\langle s_{d_0} \rangle$  is 4.5%, for the  $BB$  flavour component.

The template fit is performed as a binned Maximum-Likelihood (ML) fit. For an introduction to the ML method, see 3.7.3. The likelihood in each two-dimensional track jet  $p_T$  category is defined as follows:

$$\mathcal{L} = \prod_{i=0}^{N_{\text{bins}}} \frac{(N_i^{\text{MC,total}})^{N_i^{\text{data}}}}{N_i^{\text{data}}!} \cdot \exp(-N_i^{\text{MC,total}}) \quad (5.2)$$

where

$$N_i^{\text{MC,total}} = \sum_{j=0}^{N_{\text{flav}}=5} f_j \cdot n_{ji} \quad (5.3)$$

Here,  $n_{ji}$  is the content of each flavour component  $j$  in bin  $i$  and  $f_j$  is the fitted normalization correction factor for each of the five aforementioned flavour components.  $N_i^{\text{data}}$  is the content of bin  $i$  in data. In total one obtains  $12 \times 5 = 60$  fitted flavour fractions and normalization correction factors.

A non-equidistant binning is used for the  $\langle s_{d_0} \rangle$  distributions in the fit. The binning is chosen such that, going from left to right, a new bin is created only if the relative MC statistical error on the total bin content is smaller than 25%. If the rightmost bin has an error larger than that it is merged with its adjoining neighbour. In the region with lowest statistics this binning scheme results in 12-56 bins for the muon- and non-muon jet  $\langle s_{d_0} \rangle$  combined, which means that the fit has enough degrees of freedom to constrain the five flavour fractions in each fit category. The  $\langle s_{d_0} \rangle$  distributions of the muon- and non-muon-jet are shown in fig. A.7 to A.10 for all fit categories.

Table 5.5 shows the fitted flavour fractions in the two-dimensional bins of track jet  $p_T$ . The  $BB$  fraction in the fitted bins varies between 3.2% and 8.9%.

Table 5.6 shows the fitted normalization correction factors that determine how much the prefit-fractions were corrected to match the data. The correction factors quoted in this table were adjusted such that the total normalization difference between data and MC is not included.

<sup>3</sup> It should be noted that the data is "double-counted" because the muon- and non-muon jet distributions are fitted simultaneously. The fit uncertainty is therefore in principle underestimated by a factor  $\sqrt{2}$ , however the fit uncertainties due to data statistical errors are negligible with respect to the other uncertainties.

fit category $p_T/\text{GeV}$		BB	BL	CC	CL	LL
$p_T(j_\mu) < 100$	$p_T(j_{i\mu}) < 100$	$0.044 \pm 0.007$	$0.117 \pm 0.012$	$0.091 \pm 0.021$	$0.204 \pm 0.024$	$0.544 \pm 0.014$
	$100 < p_T(j_{i\mu}) < 200$	$0.087 \pm 0.004$	$0.139 \pm 0.006$	$0.052 \pm 0.009$	$0.238 \pm 0.013$	$0.484 \pm 0.006$
	$200 < p_T(j_{i\mu}) < 300$	$0.066 \pm 0.004$	$0.164 \pm 0.007$	$0.110 \pm 0.011$	$0.140 \pm 0.013$	$0.520 \pm 0.006$
	$p_T(j_{i\mu}) > 300$	$0.085 \pm 0.004$	$0.121 \pm 0.008$	$0.090 \pm 0.013$	$0.176 \pm 0.018$	$0.528 \pm 0.006$
$100 < p_T(j_\mu) < 200$	$p_T(j_{i\mu}) < 100$	$0.056 \pm 0.004$	$0.101 \pm 0.006$	$0.019 \pm 0.010$	$0.177 \pm 0.014$	$0.647 \pm 0.009$
	$100 < p_T(j_{i\mu}) < 200$	$0.063 \pm 0.004$	$0.117 \pm 0.006$	$0.071 \pm 0.009$	$0.159 \pm 0.013$	$0.590 \pm 0.007$
	$200 < p_T(j_{i\mu}) < 300$	$0.063 \pm 0.005$	$0.117 \pm 0.008$	$0.065 \pm 0.012$	$0.206 \pm 0.017$	$0.548 \pm 0.009$
	$p_T(j_{i\mu}) > 300$	$0.071 \pm 0.007$	$0.113 \pm 0.012$	$0.029 \pm 0.018$	$0.206 \pm 0.030$	$0.581 \pm 0.014$
$p_T(j_\mu) > 200$	$p_T(j_{i\mu}) < 100$	$0.034 \pm 0.002$	$0.113 \pm 0.003$	$0.028 \pm 0.005$	$0.116 \pm 0.008$	$0.709 \pm 0.005$
	$100 < p_T(j_{i\mu}) < 200$	$0.052 \pm 0.003$	$0.107 \pm 0.005$	$0.054 \pm 0.007$	$0.123 \pm 0.013$	$0.664 \pm 0.009$
	$200 < p_T(j_{i\mu}) < 300$	$0.053 \pm 0.005$	$0.076 \pm 0.011$	$0.081 \pm 0.013$	$0.154 \pm 0.025$	$0.635 \pm 0.015$
	$p_T(j_{i\mu}) > 300$	$0.034 \pm 0.010$	$0.138 \pm 0.022$	$0.050 \pm 0.031$	$0.052 \pm 0.052$	$0.726 \pm 0.022$

 Table 5.5: Fitted flavour fractions in the  $g \rightarrow b\bar{b}$ -enriched sample. The errors are statistical errors from the fit only.

fit category $p_T/\text{GeV}$		BB	BL	CC	CL	LL
$p_T(j_\mu) < 100$	$p_T(j_{i\mu}) < 100$	$0.769 \pm 0.125$	$1.014 \pm 0.102$	$2.161 \pm 0.492$	$1.572 \pm 0.183$	$0.830 \pm 0.023$
	$100 < p_T(j_{i\mu}) < 200$	$1.420 \pm 0.062$	$0.845 \pm 0.039$	$1.014 \pm 0.185$	$1.442 \pm 0.077$	$0.867 \pm 0.011$
	$200 < p_T(j_{i\mu}) < 300$	$1.080 \pm 0.069$	$0.997 \pm 0.041$	$1.976 \pm 0.196$	$0.863 \pm 0.083$	$0.935 \pm 0.012$
	$p_T(j_{i\mu}) > 300$	$1.475 \pm 0.078$	$0.744 \pm 0.049$	$1.744 \pm 0.246$	$1.157 \pm 0.115$	$0.917 \pm 0.011$
$100 < p_T(j_\mu) < 200$	$p_T(j_{i\mu}) < 100$	$1.698 \pm 0.130$	$1.126 \pm 0.072$	$0.705 \pm 0.371$	$1.527 \pm 0.124$	$0.881 \pm 0.012$
	$100 < p_T(j_{i\mu}) < 200$	$1.266 \pm 0.080$	$1.000 \pm 0.053$	$1.855 \pm 0.228$	$1.223 \pm 0.103$	$0.887 \pm 0.011$
	$200 < p_T(j_{i\mu}) < 300$	$1.100 \pm 0.084$	$0.922 \pm 0.064$	$1.531 \pm 0.290$	$1.602 \pm 0.134$	$0.851 \pm 0.014$
	$p_T(j_{i\mu}) > 300$	$1.302 \pm 0.125$	$0.796 \pm 0.084$	$0.659 \pm 0.411$	$1.668 \pm 0.246$	$0.913 \pm 0.023$
$p_T(j_\mu) > 200$	$p_T(j_{i\mu}) < 100$	$1.691 \pm 0.096$	$1.581 \pm 0.041$	$1.473 \pm 0.236$	$1.290 \pm 0.085$	$0.886 \pm 0.006$
	$100 < p_T(j_{i\mu}) < 200$	$1.378 \pm 0.074$	$1.183 \pm 0.057$	$1.913 \pm 0.239$	$1.239 \pm 0.126$	$0.892 \pm 0.012$
	$200 < p_T(j_{i\mu}) < 300$	$1.184 \pm 0.113$	$0.855 \pm 0.121$	$2.660 \pm 0.430$	$1.583 \pm 0.254$	$0.860 \pm 0.021$
	$p_T(j_{i\mu}) > 300$	$0.891 \pm 0.257$	$1.437 \pm 0.227$	$1.581 \pm 0.984$	$0.520 \pm 0.521$	$0.989 \pm 0.031$

 Table 5.6: Fitted relative flavour fraction correction factors in the  $g \rightarrow b\bar{b}$ -enriched sample. The errors are statistical errors from the fit only. The correction factors were adjusted, so that the correction of the total normalization difference between data and MC is not included.

## 5.5 Validation of Higgs tagging variables

As mentioned in the introduction, one of the goals of the studies in the  $g \rightarrow b\bar{b}$  final state is the validation of Higgs-tagging-related variables, such as large- $R$  jet properties and substructure variables, as well as distributions related to  $b$ -tagging. The aim is to investigate the agreement between data and simulation in a sample that contains two close-by  $b$ -jets. A sufficiently pure  $g \rightarrow b\bar{b}$  sample is obtained by applying double- $b$ -tagging using the MV2c10 algorithm at a 70% working point after the template fit.

The reason why it is important to validate the modelling of jet properties related to Higgs tagging in a sample of close-by  $b$ -jets is, firstly, that the standard jet energy scale calibration is performed in a flavour-inclusive multijet sample [52] and, secondly, that the standard  $b$ -tagging calibration is using  $t\bar{t}$ -events [68, 97]. Both of these samples contain few events with a close-by  $b$ -jet topology or none at all. Therefore it is not a priori clear whether the default uncertainties derived in these samples are sufficient to cover possible differences between data and simulation in close-by  $b$ -jet environments. Studying the modelling of jet properties and flavour tagging in data in  $g \rightarrow b\bar{b}$  events provides an ideal setup to test this.

In this subsection large- $R$  jet properties, such as the transverse momentum, the large- $R$  jet mass

and the substructure variables  $D_2$  and  $\tau_{21}$  (introduced and defined in section 3.4.4) will be studied in the  $g \rightarrow b\bar{b}$ -enriched sample after the flavour fraction correction and applying double- $b$ -tagging. The variables  $D_2$  and  $\tau_{21}$  were selected for the validation study because they were among the most powerful jet substructure variables for Higgs tagging [53]. In addition, the double- $b$ -tagging rate and kinematic properties of the muon- and non-muon-jet are studied.

It is important to note that, for the validation study, the  $\langle s_{d_0} \rangle$  fit for the flavour fraction correction is performed in the sample before applying the double- $b$ -tagging (*pre-tag*), and the pre-tag flavour correction is applied in the sample after double- $b$ -tagging (*post-tag*).

### 5.5.1 Systematic uncertainties in validation study

For the validation, the impact of systematic variations on the Higgs-tagging-related variable distributions are studied directly, meaning that the effect of a systematic variation on the variable shape and not only on the flavour fraction normalizations are considered. This way it is possible to investigate whether the default systematic uncertainties are sufficient to cover possible discrepancies between data and simulation also in topologies with close-by  $b$ -jets.

Systematic uncertainties on the variable distributions shown in section 5.5 are determined bin-by-bin, by considering the difference between the nominal and the systematically varied variable distribution. In both cases the flavour fraction correction factors from the same  $\langle s_{d_0} \rangle$  fit, performed in the nominal sample, are applied. The upward and downward differences between the nominal and the varied distributions in each bin are added up in quadrature. In the following, the systematic variations considered in the validation study will be discussed briefly. A summary of the impact of the systematic uncertainties on the BB and non-BB event yields is given in figure 5.7.

#### Large- $R$ jet energy, mass and substructure scale

The large- $R$  jet energy mass and substructure scale uncertainties are determined in a sample of flavour-inclusive dijet events, by comparing track-only-based quantities with calorimeter-only-based quantities [52]. The scale uncertainties on the energy and the mass as well as the substructure are treated as fully correlated.

#### Large- $R$ jet energy and mass resolution

The uncertainties on the large- $R$  jet energy and mass resolution are parametrized as a smearing degrading the large- $R$  jet energy resolution by 2% and the mass resolution by 20% [49, 98]. The resolution uncertainties are one-sided and are symmetrized in this analysis.

#### Flavour tagging scale factor uncertainties:

The standard single track-jet flavour tagging efficiency scale factors and their uncertainties are applied in the post-tag distributions. The scale factors are applied using the standard multiplicative-event-weight approach as described in section 5.1. For more information about flavour tagging calibration scale factors and their uncertainties, see section 3.5.9.

### Generator modelling uncertainties

To investigate the influence of a different parton shower model on the Higgs-tagging-related jet properties, simulated events generated with the HERWIG++ MC generator were compared to the nominal events generated with PYTHIA8. The difference between the distributions is a one-sided systematic uncertainty which is smoothed and symmetrized.

### Fit uncertainties

The uncertainties on the flavour fractions from the  $\langle s_{d_0} \rangle$  template fit are correlated among each other, because the flavour template normalizations are varied simultaneously in the fit to minimize the negative log likelihood. The impact of the fit uncertainties on the large- $R$  jet properties are therefore estimated using fully correlated error propagation for a variable  $f$ :

$$(\Delta f)^2 = \sum_{i,j=0}^4 \frac{\partial f}{\partial N_i} \frac{\partial f}{\partial N_j} \Sigma_{ij} \quad (5.4)$$

where  $i, j$  are indices of the five flavour components (BB, BL, CC, CL, LL) and  $\Sigma_{ij}$  is the covariance matrix obtained for each fit. Especially, we obtain for the total normalization of all flavour components:

$$(\Delta N_{\text{total}})^2 = \mathbf{v} \Sigma \mathbf{v}^T, \quad (5.5)$$

where  $N_{\text{total}} = \sum_{i=0}^4 N_i$  is the total normalization,  $\mathbf{v} = (N_{BB}, N_{BL}, N_{CC}, N_{CL}, N_{LL})$  is a row vector of the normalizations of each flavour component and  $\Sigma$  is the covariance matrix.

Since the fit is performed in two-dimensional categories of the muon jet and non-muon jet  $p_T$  and a covariance matrix is obtained for each category, calculating the impact of the statistical fit uncertainty on the global<sup>4</sup> distributions of the large- $R$  jet properties becomes too complex for analytical evaluation of the error propagation. This is because the relative contribution of each fit category in a single bin of the large- $R$  jet distribution varies. Therefore an approach using pseudo-experiments is used in combination with a principle component decomposition of each of the covariance matrices from the flavour fraction fits. The covariance matrix is diagonalized and five eigenvalues  $\lambda_i$  and eigenvectors  $\mathbf{e}_i$  are obtained. A new set of fit parameters  $\mathbf{c}_{\text{pseudo}}$  is created by allowing the parameters obtained from the nominal fit  $\mathbf{c}_{\text{nom}}$  to randomly vary within the fit parameter space as spanned by the eigenvectors:

$$\mathbf{c}_{\text{pseudo}} = \mathbf{c}_{\text{nom}} + \sum_i r_i \mathbf{e}_i \quad (5.6)$$

where  $r_i$  is a Gaussian-distributed random variable whereas the Gaussian probability density function has a standard deviation of  $\sqrt{\lambda_i}$  of the eigenvalues of the covariance matrix. This random-walk in the fit parameter space is repeated two thousand times for each fit category and a new post-fit distribution for the large- $R$  jet related variable is obtained. For each bin of the large- $R$  jet related variable, a probability distribution of the bin content is obtained, which is fit using a Gaussian function. The standard deviation of the fitted Gaussian is used as an estimate for the total fit uncertainty of the considered bin. As an example, figure 5.6 shows the distribution of the bin content in bin 7 of the large- $R$  jet  $D_2^{\beta=1}$ -distribution, demonstrating the Gaussian behaviour.

---

<sup>4</sup> Global here means summed over all (muon-jet  $p_T$ , non-muon-jet  $p_T$ ) fit categories.

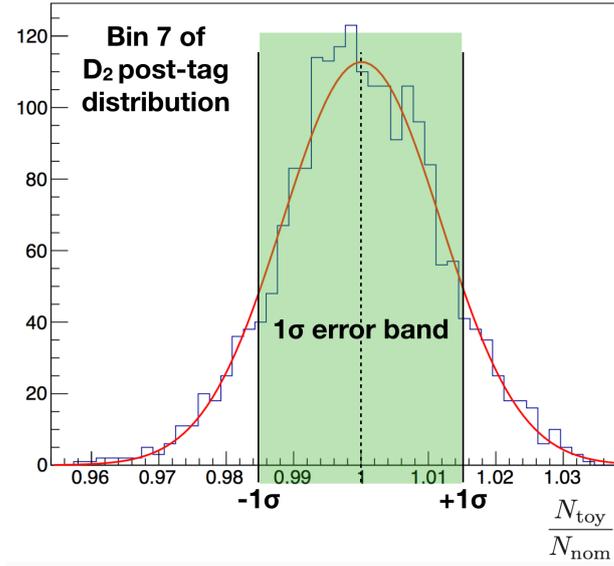


Figure 5.6: Determination of the fit uncertainty in bin seven of the  $D_2^{\beta=1}$  distribution using pseudo-experiments.

Source	Signal (B,B) [%]	Background (non-B,non-B) [%]
Jet scales	9.0	7.3
Jet energy resolution	1.0	1.8
Jet mass resolution	0.1	0.2
JVT	0.01	0.01
$b$ -tagging related	9.5	27
Modelling	23	13
Fit statistics	1.0	0.1

Figure 5.7: Impact of the systematic uncertainties on the BB and non-BB yields in the  $g \rightarrow b\bar{b}$  validation study. From [53].

### 5.5.2 Jet kinematics and $b$ -tagging results

Figure 5.8 shows the transverse momentum distribution of the muon and non-muon track-jets as well as the large- $R$  jet  $p_T$  distribution, before and after double- $b$ -tagging. It is visible how the double- $b$ -tagging algorithm changes the flavour composition such that the main contribution are now events from the BB category and the contribution from other flavour categories are small. The muon-jet  $p_T$  spectrum is harder than that of the non-muon jet, because of the selection criterion of the muon-jet being the leading  $p_T$  track jet with an associated muon. The agreement between data and simulation is good both before and after  $b$ -tagging. The dominant uncertainties in the post-tag sample are the  $b$ -tagging SF uncertainties.

Figure 5.9 shows the double- $b$ -tagging rate, defined as:

$$r_{b\text{-tag}} = \frac{N_{\text{jets, double-}b\text{-tagged}}}{N_{\text{jets, total}}} \quad (5.7)$$

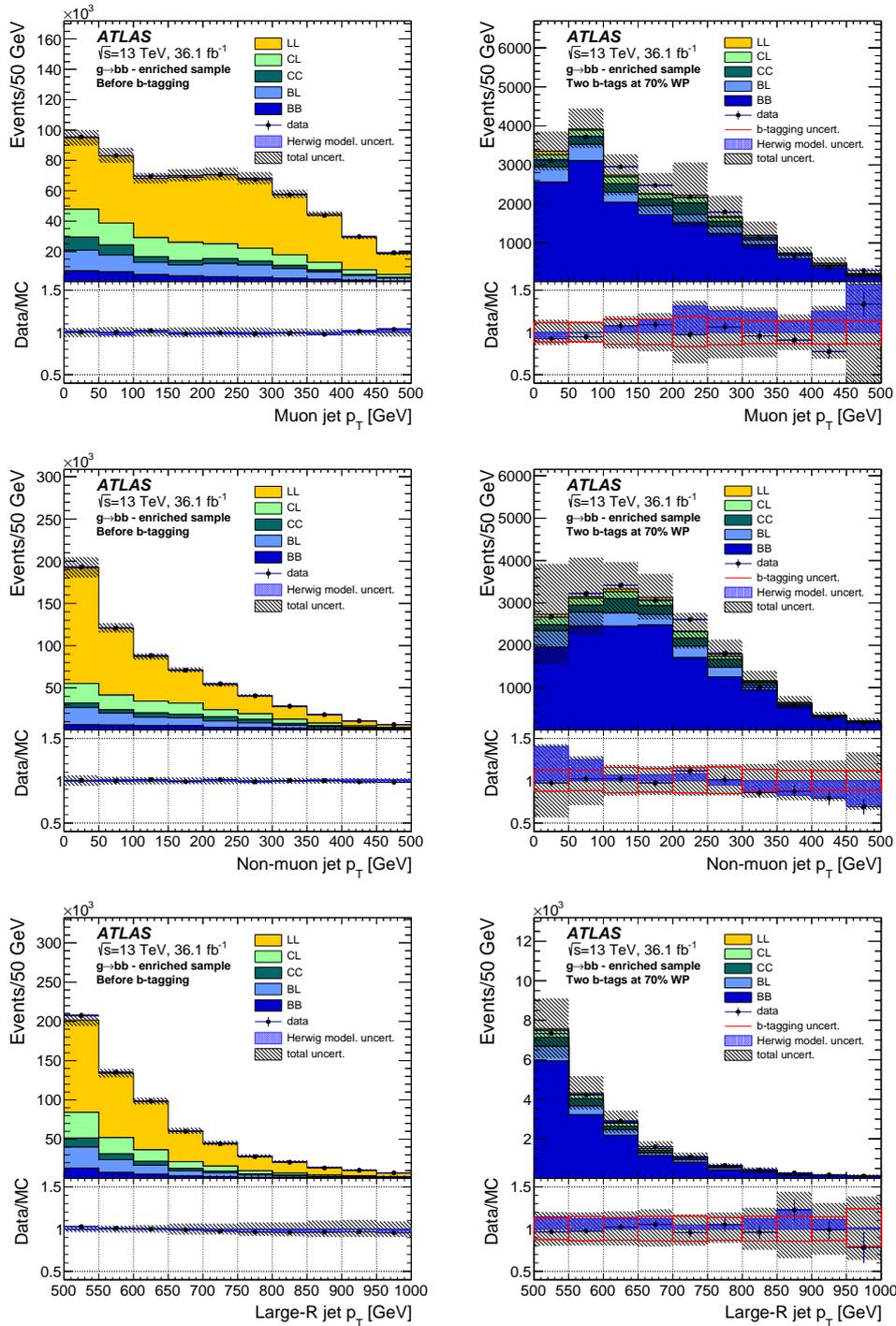


Figure 5.8: Transverse momentum distributions of the muon jet (top), non-muon jet (middle) and large- $R$  jet (bottom) before (left) and after (right)  $b$ -tagging. The  $b$ -tagging scale factors and the flavour fraction correction have been applied. The light shaded band indicates the total uncertainty. The red band shows the  $b$ -tagging uncertainty. The blue shaded band in the ratio panel indicates the modelling uncertainty given by the difference between the distributions predicted by PYTHIA and HERWIG. From [53].

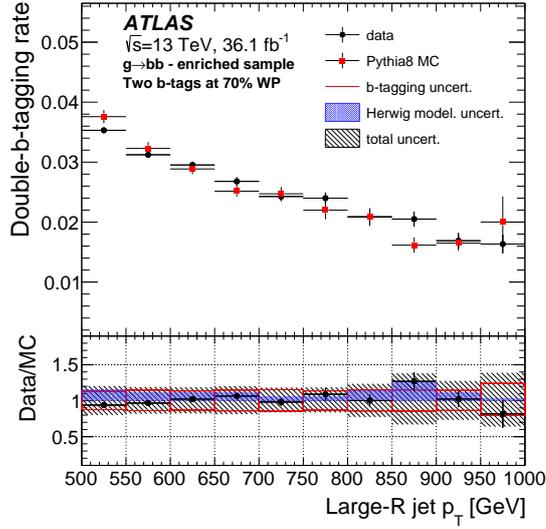


Figure 5.9: Comparison of the double- $b$ -tagging rate as a function of the large- $R$  jet  $p_T$  in data and MC simulation. The light shaded band indicates the total uncertainty. The red band shows the  $b$ -tagging uncertainty. The blue shaded band in the ratio panel indicates the modelling uncertainty given by the difference between the distributions predicted by PYTHIA and HERWIG. From [53].

where  $N_{\text{jets, double-}b\text{-tagged}}$  denotes the number of large- $R$  jet  $g \rightarrow b\bar{b}$  candidates, independent of their flavour content that have a positive double- $b$ -tagging decision and  $N_{\text{jets, total}}$  denotes the number of all  $g \rightarrow b\bar{b}$  candidates after the preselection described in 5.2. The  $b$ -tagging efficiency of MV2c10 drops for higher jet  $p_T$ , which leads to a lower double- $b$ -tagging rate at high  $p_T$ . For the double- $b$ -tagging rate, good agreement is observed between data and simulation.

### 5.5.3 Jet mass and substructure results

Since jet substructure variables can also be used for the identification of boosted Higgs bosons, two example variables,  $\tau_{21}$  and  $D_2^{\beta=1}$  that showed promising separation power between large- $R$  jets from Higgs boson decays and jets from top or QCD background [53] were studied in the  $g \rightarrow b\bar{b}$  environment in terms of their agreement between data and simulation. For an introduction to jet substructure, including a detailed description of the variables  $\tau_{21}$  and  $D_2^{\beta=1}$ , see 3.4.4. In addition, the large- $R$  jet mass is studied, which provides also good discrimination power for Higgs boson identification.

Figure 5.10 shows the large- $R$  jet mass distribution as well as the large- $R$  jet  $D_2^{\beta=1}$  and  $\tau_{21}$  distributions before and after double- $b$ -tagging. The mass distribution is nonzero because firstly the gluon in the splitting process can be virtual therefore having a nonvanishing mass. Secondly, since the fraction of different QCD processes is not well modelled in simulation, there is no explicit requirement on the truth events for the jet be to caused by a single gluon splitting process. Therefore, in principle also other QCD processes with a similar topology could contribute, which leads to non-zero mass.

Both the  $D_2^{\beta=1}$  and  $\tau_{21}$  substructure variables should peak towards small values for a two-prong Higgs jet. Since the events considered in this study are produced by the decay of a colour-octet gluon, instead

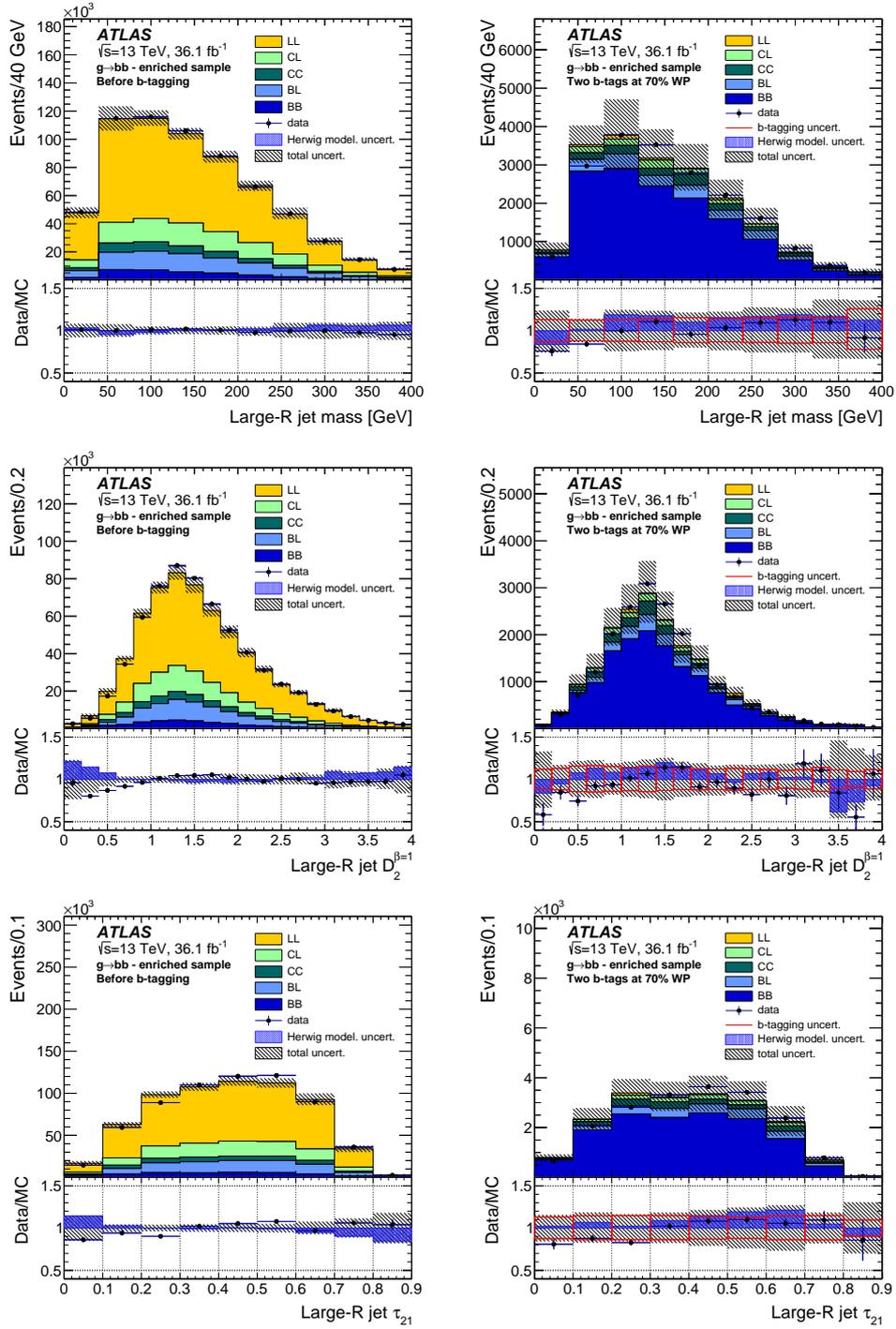


Figure 5.10: Distributions of large- $R$  jet mass (top),  $D_2^{\beta=1}$  (middle) and  $\tau_{21}$  (bottom) before (left) and after (right)  $b$ -tagging. The light shaded band indicates the total uncertainty. The red band shows the  $b$ -tagging uncertainty. The blue shaded band in the ratio panel indicates the modelling uncertainty given by the difference between the distributions predicted by PYTHIA and HERWIG. From [53].

of a colour singlet Higgs boson, they are expected to have different radiation characteristics and energy correlations between the two hard energy deposits. This is why the substructure distributions in  $g \rightarrow b\bar{b}$  can deviate from the expectation for a real  $H \rightarrow b\bar{b}$  jet. For both variables, the ratio between data and simulation shows a deviation from unity, however the total magnitude of the deviation is covered by the total uncertainty. Since the correct modelling of parton showers is one of the most difficult tasks in a simulation (see section 3.2.4) and the substructure variables strongly depend on the parton shower, deviations are not unexpected. As mentioned, the difference between two different shower generators (PYTHIA and HERWIG) is included as a systematic uncertainty and shown also as a separate uncertainty band. However, based on figures 5.8 and 5.10, no clear statement is possible on whether one shower model shows better agreement with data than the other for  $D_2^{\beta=1}$  and  $\tau_{21}$ .

The successful validation of jet substructure variables concludes the use of the  $g \rightarrow b\bar{b}$ -enriched sample as a test bench for studying the modelling of Higgs tagging variables in a close-by b-jet environment. In the following, the next steps are taken to use the  $g \rightarrow b\bar{b}$  sample to calibrate double- $b$ -tagging using real data.

## 5.6 Scale factor calculation

The main goal of the  $g \rightarrow b\bar{b}$  calibration is to measure the double- $b$ -tagging efficiency in data. The template fit of the variable  $\langle s_{d_0} \rangle$  as described in section 5.4 can be used to measure the flavour composition of the two track jets within the large- $R$  jet in data. As a result, the double- $b$ -tagging efficiency calibration scale factors can be determined using  $g \rightarrow b\bar{b}$  events.

The scale factor  $\kappa$  encodes the difference between the double- $b$ -tagging efficiency in data and MC and is defined as:

$$\kappa = \frac{\epsilon_{\text{data}}^{\text{BB}}}{\epsilon_{\text{MC}}^{\text{BB}}} \quad (5.8)$$

where  $\epsilon_{\text{data}}^{\text{BB}}$   $\epsilon_{\text{MC}}^{\text{BB}}$  are the double- $b$ -tagging efficiencies in data and simulation respectively. The efficiency is defined in the following way:

$$\epsilon^{\text{BB}} = \frac{N_{\text{double-b-tagged}}^{\text{BB}}}{N_{\text{total}}^{\text{BB}}}, \quad (5.9)$$

where  $N_{\text{double-b-tagged}}^{\text{BB}}$  is the number of BB jets, which pass the double- $b$ -tagging requirement and  $N_{\text{total}}^{\text{BB}}$  is the total number of BB jets. The efficiency from simulation can easily be extracted by using the truth flavours of the track jets associated to the large- $R$  jets. In data, it is not possible to determine the double flavour content of the large- $R$  jet on a jet-by-jet basis. However it is possible to estimate the fraction of true BB jets in data via the  $\langle s_{d_0} \rangle$  template fit as is described in section 5.4.

Figure 5.11 demonstrates the extraction of the fraction of BB jets in data for a simplified case where there are only two bins in the  $\langle s_{d_0} \rangle$  distribution and only three different flavour pairs<sup>5</sup>. The fit to data is performed using the  $\langle s_{d_0} \rangle$  templates in the sample where no double- $b$ -tagging is yet applied (①  $\rightarrow$  ②). Let this sample be called the *post-fit pre-tag sample*. The flavour fraction correction factors that are

<sup>5</sup> In the real analysis the binning is based on the bin statistical error, which leads to at least 12 bins in  $\langle s_{d_0} \rangle$  and five flavour pairs

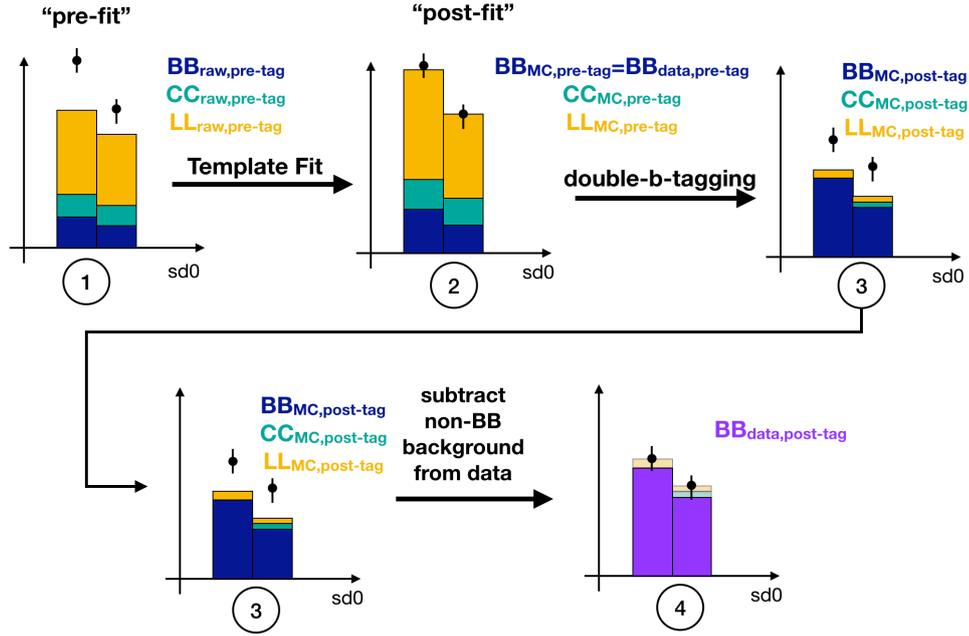


Figure 5.11: Sketch of the determination of the BB fraction in data from the  $\langle s_{d_0} \rangle$  flavour fraction fit.

determined from the  $\langle s_{d_0} \rangle$  fit are then also applied in the sample where the double- $b$ -tagging is applied. It is important to note that there is no new fit performed in the sample that passes the double- $b$ -tagging requirement. The sample that passes the double- $b$ -tagging cut and has the flavour fraction correction factors applied is called the *post-fit post-tag* sample (3).

The denominator of equation 5.9 can be determined from the post-fit pre-tag sample. Since the template normalizations are fit to match the data in the pre-tag sample, the integral of the corrected BB template can be interpreted as the total number of BB-jets before double- $b$ -tagging. The numerator of equation 5.9 must be determined in the post-fit post-tag sample. Since the scale factors encode the difference between data and simulation, we assume that the difference that we observe in the post-fit post-tag sample (3) is entirely due to a difference between the double- $b$ -tagging efficiency in data and MC. Therefore the numerator of equation 5.9 is determined by subtracting the integral of the post-fit non-BB background templates from the integral of the data distribution (4).

### 5.6.1 Systematic uncertainties in double- $b$ -tagging calibration

In the calibration analysis the influence of the systematic variations on the flavour fractions and double- $b$ -tagging efficiencies are of main interest, to determine the total uncertainty on the double- $b$ -tagging scale factors.

For the double- $b$ -tagging calibration the treatment of the systematic uncertainties is focusing mainly on the  $\langle s_{d_0} \rangle$  fit. One therefore discerns between systematic uncertainties that affect the shape of the  $\langle s_{d_0} \rangle$  distribution and those which only affect the normalization. In general, for each systematic variation, the

double- $b$ -tagging scale factors are rederived and the differences of the upward and downward variations to the nominal value are added up in quadrature. Table 5.7 presents a summary of the size of the systematic uncertainties described below.

### Systematics not affecting $\langle s_{d_0} \rangle$ shape

The systematic uncertainties that do not affect the  $\langle s_{d_0} \rangle$  shape include large- $R$  jet energy scale, energy resolution and mass resolution. The main effect of varying these systematic uncertainties up and down is a change in the number of jets that pass or fail the kinematic thresholds for the large- $R$  jet (see section 5.3). No change in the distribution of  $\langle s_{d_0} \rangle$  is expected, since the small- $R$  track jets are unaffected by changes in the large- $R$  jet energy and mass. Therefore the template fit is not performed again in the varied sample, but the nominal fit correction is applied to the systematically varied sample to recalculate the scale factors.

### Systematics affecting $\langle s_{d_0} \rangle$ shape

The systematics listed below affect the  $\langle s_{d_0} \rangle$  shape. New simulated templates are derived from MC for each systematic variation in the following way:

- **Transverse impact parameter resolution:**

Since the transverse impact parameter  $d_0$  is one of the main ingredients of the template variable  $\langle s_{d_0} \rangle$ , the quality of its description in simulation is crucial for this analysis. It has been observed that there is a disagreement between simulation and data for the transverse impact parameter resolution. Therefore a smearing procedure has been derived by measuring the transverse impact parameter resolution in dijet,  $Z$  + jets and minimum bias events [61] in fine categories of  $p_T$  and  $\eta$ . In this analysis the smeared transverse impact parameter is used as nominal. The difference between the smeared and un-smeared distributions are assumed as a one-sided systematic uncertainty, which is symmetrized.

- **Light long lived hadrons:**

Light hadrons such as Kaons or Lambda particles can have a significant lifetime. In connection with a high momentum transfer these light flavour hadrons can create a secondary vertex which can lead to large  $\langle s_{d_0} \rangle$  values even though no  $B$ -hadron is present in the jet. This mainly affects the  $\langle s_{d_0} \rangle$  template shape of the CL, BL, and LL templates, as was described in 5.4. Since the fraction of these light hadrons is not well described in simulation, a systematic uncertainty is associated to this. It is derived by searching the event truth record for the presence of a track from light long-lived hadrons among the tracks used for the  $b$ -tagging of one of the small- $R$  track jets and varying their event weight up and down by 10%, following the procedure described in [61].

- **Photon conversions:**

Photons in the ATLAS detector can convert into an electron-positron-pair, spontaneously or triggered by interactions with the detector material. If a conversion occurs within the inner detector the two electron tracks can mimic a secondary vertex which can lead to large  $\langle s_{d_0} \rangle$  values despite no  $B$ -hadron being present in the jet (see also 5.4). Since the conversion rate is not precisely known, the truth record is searched for the presence of a conversion occurring within a small- $R$  jet (with the same association requirements of truth particles to tracks as for the light long lived hadrons)

and the corresponding events are weighted up and down by 10% [61].

- **Hadronic material interactions:**

The interaction of hadrons from the primary vertex with material in the inner detector can lead to secondary tracks with a large impact parameter and therefore a large  $\langle s_{d_0} \rangle$  which can distort the templates as described in section 5.4. Since the rate of this process is not well known, a systematic uncertainty is assigned to it by searching the truth record for the occurrence of a hadronic material interaction and weighting the corresponding events up and down by 10% [61]. A hadronic material interaction is searched for by checking all truth particles with a specific barcode from the detector simulation that indicates a secondary interaction. If the track does not stem from a  $B$ -hadron and if none of the above conditions for conversions or light long-lived particles is met, a hadronic material interaction is assumed.

- **Muon Fakes:**

In this analysis context a fake muon refers to a reconstructed muon where no muon is present in the truth record within a cone of  $R = 0.2$  around the reconstructed muon. Reasons for this can be calorimeter punch-through effects resembling a muon signature, or in-flight decays. An associated systematic uncertainty is derived by weighting the fraction of events with a fake muon up by a factor of three [61].

- **Generator dependence:**

Differences between the event generators and the shower model can affect the jet kinematics and also the shapes of the  $\langle s_{d_0} \rangle$  distributions. To estimate the impact the fractions are re-fitted using events generated with HERWIG++. Since the number of events in the alternative HERWIG++ MC sample is much smaller than in the PYTHIA8 sample, the fit is performed without splitting in bins of muon-jet- $p_T$  and non-muon-jet- $p_T$  and one global double- $b$ -tagging scale factor is calculated for both generators. The difference between the two scale factors of 11% is assigned as a global modelling uncertainty on the double- $b$ -tagging scale factors.

For each of these systematic variations, a new set of  $\langle s_{d_0} \rangle$  templates is derived and the flavour fraction correction fit is newly performed for each variation. With the re-fitted flavour fraction, the systematically varied scale factors are calculated.

### Data statistical uncertainty

The data statistical uncertainty concerns the limited statistics on the  $\langle s_{d_0} \rangle$  distributions from measured LHC data. To estimate the impact of this uncertainty, a method based on pseudo-experiments is used. A set of ten thousand new  $\langle s_{d_0} \rangle$  distributions per fit category is created by setting each histogram bin content to a Poisson-distributed random integer, where the mean of the Poisson-distribution is the bin content in the original data histogram. The error in the new histogram's bin is set to the square root of the bin content. With each of the pseudo-datasets, the flavour fraction correction fit is newly performed and the scale factors are obtained for each set. The scale factors are histogrammed and a Gaussian distribution is fit to the histogram similar to section 5.5.1 (cf. also figure 5.6). The standard deviation of the fitted Gaussian function is taken as the data statistical error on the scale factor.

### MC statistical uncertainties

The uncertainties that stem from MC statistical uncertainties have to be subdivided into two parts:

1. The first part concerns the MC uncertainty associated with the template distributions. To estimate the impact of the template statistical uncertainties on the scale factor, pseudo-experiments are used once more. A set of ten thousand new  $\langle s_{d_0} \rangle$  templates (in each double flavour and track jet  $p_T$  category) is created, where in each template bin the bin content is replaced by a Gaussian-distributed random number, where the mean of the Gaussian corresponds to the bin content in the original MC template and the standard deviation of the Gaussian distribution corresponds to the MC statistical error in this bin. The template fits are newly performed for each set of these pseudo-templates and a new scale factor is determined for each. The scale factor is histogrammed for each of the ten thousand pseudo-experiments, and a Gaussian distribution is fit to it. The standard deviation of the fitted Gaussian distribution is assumed as the scale factor uncertainty due to the template MC statistical uncertainty.
2. The second component of the MC statistical uncertainty concerns the count rates  $N_{\text{double-}b\text{-tagged}}^{\text{BB}}$  and  $N_{\text{total}}^{\text{BB}}$  in the numerator and denominator of the double- $b$ -tagging efficiency in MC (eq. 5.9). It is calculated by propagating the statistical error on the count rates through the efficiency calculation using Gaussian error propagation. The efficiency measured in data is unaffected by this uncertainty.

### Non-BB flavour tagging uncertainties

After double- $b$ -tagging, the non-BB flavour templates are subtracted from data to determine the number of double- $b$ -tagged BB jets in data (see section 5.6). However it is known, that also the mis-identification of non- $b$ -jets, especially for light-jets can be subject to large scale factors. The presented calibration is intended only for a double- $b$ -tagging BB efficiency calibration, which should only correct for disagreement between data and MC in the BB components. Therefore the standard flavour tagging scale factors derived for  $b$ -tagging of isolated jets are applied for all non-BB flavour fractions (BL, CC, CL and LL) and their respective uncertainties are taken into account in the same way as for the validation study (see 5.5.1). Since the flavour tagging efficiency scale factor uncertainties only affect the total flavour normalizations, the fit is not reperformed for this systematic variation.

## 5.7 Extrapolation from semi-leptonic to inclusive $b$ -jet sample

While selecting  $g \rightarrow b\bar{b}$  events (5.3), the requirement that one of the two track jets associated to the large- $R$  jet has an associated muon from a semi-leptonic  $B$ -hadron decay serves the purpose to enrich the sample in BB jets. However, this selection requirement also biases the  $B$ -hadron composition of the  $g \rightarrow b\bar{b}$  sample, because it increases the fraction of semi-leptonic  $B$ -hadron decays. This also causes a bias in the admixture of  $B$ -hadron species in the  $g \rightarrow b\bar{b}$  sample. Different  $B$ -hadron types decay with different track multiplicities. In addition, different  $B$ -hadrons have different lifetimes, therefore the separation between primary and secondary vertex is affected by the hadron composition in the sample. Since the  $b$ -tagging algorithm MV2c10 is sensitive to these properties, it is expected that the  $b$ -tagging efficiency is different for semi-leptonic and inclusive  $B$ -hadron decays.

The aim of the  $g \rightarrow b\bar{b}$  double- $b$ -tagging calibration is to provide scale factors that are independent of a specific sample selection. However, the scale factors are derived in a sample that is biased in the

fraction of semi-leptonic  $B$ -hadron decays, which may also bias the double- $b$ -tagging scale factor. As a result, an extrapolation procedure is introduced to estimate the impact of this bias. Ideally, one would remove the muon selection requirement to obtain an inclusive unbiased  $B$ -hadron sample, recalculate the double- $b$ -tagging scale factors for the inclusive sample in data and MC and calculate a multiplicative correction to the original scale factor. However, since the  $\langle s_{d_0} \rangle$  fit and the determination of the efficiency in data is not possible without the muon requirement, one cannot determine the efficiency in data for the inclusive sample. As a result it was decided to make a purely MC-based estimate of the extrapolation uncertainty, by studying the difference between the double- $b$ -tagging efficiencies in simulation with and without the muon selection requirement. This is possible, because the efficiency in simulation is independent of the  $\langle s_{d_0} \rangle$  fit. It is then assumed that the efficiencies in data and simulation vary in the same way, which is equivalent to the statement that the scale factor as defined in equation 5.8 does not change when extrapolating from the semi-leptonic to the inclusive sample. The MC statistical error on the estimated efficiency ratio between semi-leptonic- and inclusive sample is then used as an estimate for the extrapolation uncertainty of the scale factor.

The double- $b$ -tagging efficiency (for the 70%  $\times$  70% WP) in the inclusive sample without the muon requirement was found to be  $\epsilon_{MC, \text{incl}}^{BB} = (45.9 \pm 2.4(\text{stat.}))\%$ . For the sample with the muon selection the double- $b$ -tagging efficiency is  $\epsilon_{MC, \mu}^{BB} = (46.6 \pm 1.4(\text{stat.}))\%$ . For the efficiency ratio

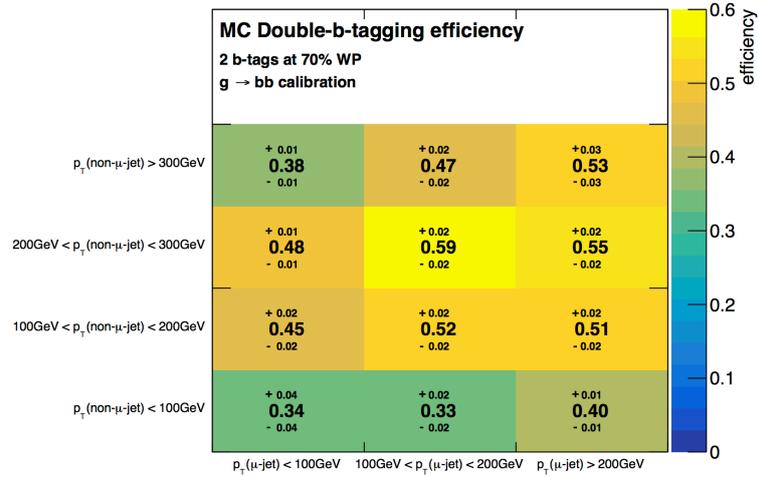
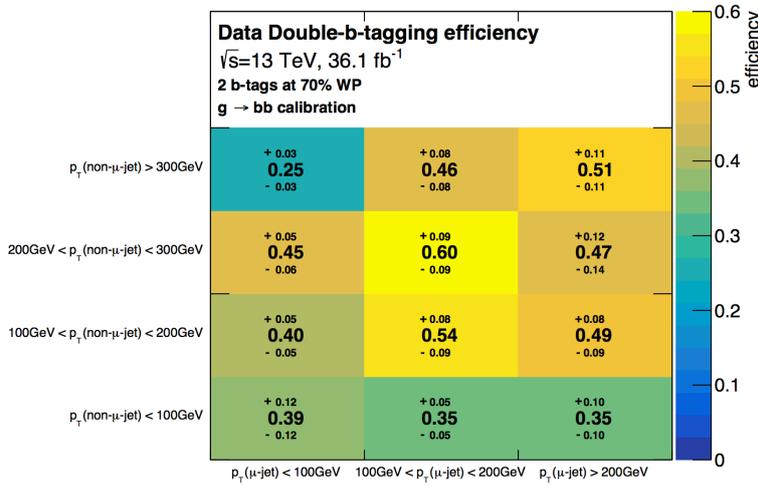
$$c_{\epsilon}^{MC} = \frac{\epsilon_{MC, \text{incl}}^{BB}}{\epsilon_{MC, \mu}^{BB}} \quad (5.10)$$

the value  $c_{\epsilon}^{MC} = 0.98 \pm 0.06(\text{stat.})$  is obtained. A summary of further studies regarding the dependence of the ratio  $c_{\epsilon}^{MC}$  on the semi-leptonic branching fractions,  $B$ -hadron  $p_T$  spectrum, as well as the generator can be found in the appendix A.2.5. No clear dependence on these parameters was found. In the following, the uncertainty  $\Delta c_{\epsilon}^{MC} = 0.06$  is assumed as extrapolation uncertainty on the scale factor.

## 5.8 Calibration Results

Figure 5.12 shows the double- $b$ -tagging efficiencies in data,  $\epsilon_{\text{data}}^{BB}$ , and in MC,  $\epsilon_{MC}^{BB}$ , as defined in equations 5.8 and 5.9. One would expect that the efficiency of double- $b$ -tagging using twice the 70% working point of the MV2c10 algorithm is  $\epsilon^{BB} \approx 0.7^2 \approx 0.5$ . There are several reasons for deviations from this expectation: Firstly, the cut requirement on the MV2c10 discriminant for the 70% working point corresponds to 70%  $b$ -tagging efficiency inclusive in jet  $p_T$  only in a  $t\bar{t}$  sample, and may deviate in the considered  $g \rightarrow b\bar{b}$  sample. Secondly, the efficiency drops for very low and high  $p_T$  jets (see also section 3.5.7).

Figure 5.13 (a) shows the double- $b$ -tagging scale factors determined in the  $g \rightarrow b\bar{b}$ -enriched sample. For the first time, scale factors have been determined in a dataset with two close-by  $b$ -jets for jet  $p_T$  above 250 GeV, where the standard *ATLAS*  $b$ -jet efficiency calibration fully relies on extrapolation uncertainties derived in simulation only. It can be seen that all scale factors are compatible with a value of one, with the exception of the scale factor in the bin of  $p_T(\mu\text{-jet}) < 100$  GeV,  $p_T(\text{non-}\mu\text{-jet}) > 300$  GeV which is shifted from one by 2.4 standard deviations (top left bin in figure 5.13 (a)). For comparison with the standard *ATLAS* approach (see also 5.1), figure 5.13 (b) shows the double- $b$ -tagging scale factors calculated by multiplying the standard single track-jet  $b$ -tagging scale factors (see section 3.5.9) for the two track jets associated to the large- $R$  jet, assuming the  $p_T$  spectrum of the simulated  $g \rightarrow b\bar{b}$ -enriched sample. Comparing figure 5.13 (a) and (b) it is clearly visible that all scale factors from the  $g \rightarrow b\bar{b}$ -based

(a) Double- $b$ -tagging efficiency in MC(b) Double- $b$ -tagging efficiency in dataFigure 5.12: Double- $b$ -tagging efficiencies in MC (top) and data (bottom) in a  $g \rightarrow b\bar{b}$ -enriched sample.

calibration agree within their uncertainties with the values determined by the standard *ATLAS* calibration. The uncertainties on the  $g \rightarrow b\bar{b}$  scale factors are larger than the uncertainties assumed for the standard *ATLAS* calibration, however, especially at high track jet momenta the  $g \rightarrow b\bar{b}$  scale factors access a new regime for data-based calibrations for the first time.

Table 5.7 shows an exemplary breakdown of the statistical and systematic uncertainties considered in the calibration analysis (as described in 5.6.1) for one bin of (muon jet  $p_T > 200\text{ GeV}$ , non-muon jet  $p_T > 300\text{ GeV}$ ) and their fractional impact on the calculated double- $b$ -tagging scale factors. It can be seen that the  $g \rightarrow b\bar{b}$  calibration analysis is dominated by systematic uncertainties, with the transverse impact parameter smearing uncertainty having the largest impact. Further important systematic uncertainties include flavour tagging uncertainties for the non-BB flavour components as well as the shower and hadronization modelling uncertainties and the MC statistical uncertainty on the  $\langle s_{d_0} \rangle$  templates. One

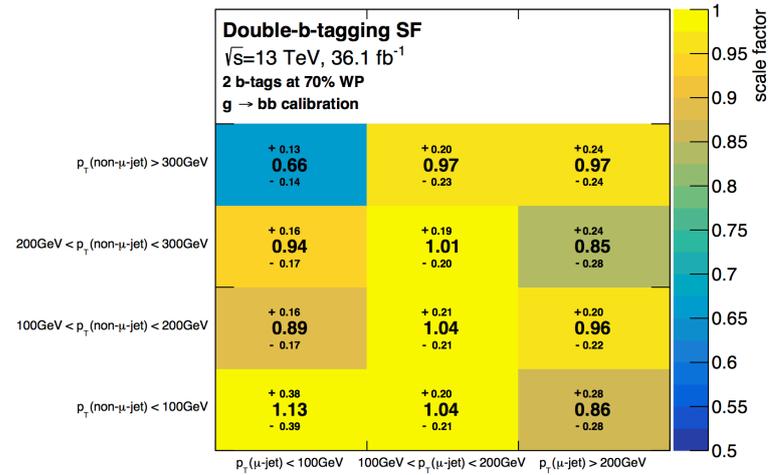
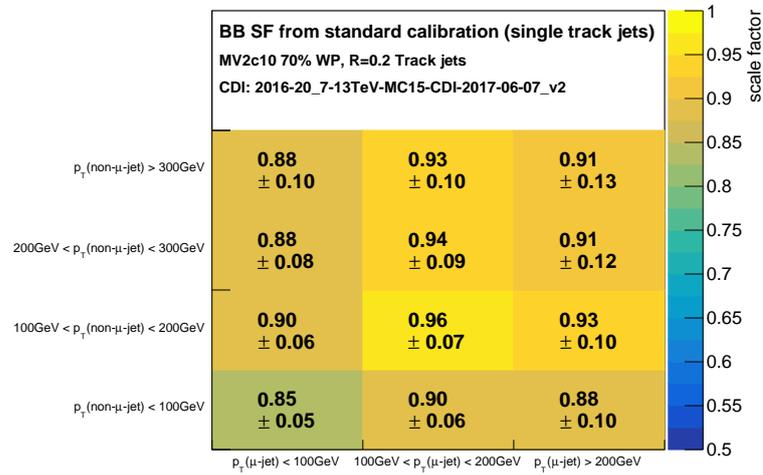

 (a) Double- $b$ -tagging scale factors determined in  $g \rightarrow b\bar{b}$  calibration

 (b) SF from ATLAS standard single  $b$ -jet calibration.

Figure 5.13: Top: Double- $b$ -tagging scale factors determined via the template fit in the  $g \rightarrow b\bar{b}$  calibration. Bottom: BB scale factors as calculated from the ATLAS standard calibration for single track jets, by multiplying the scale factors of the individual track jets. The assumed (muon-jet, non-muon jet)  $p_T$  distribution is taken from the dijet MC as is used for the  $g \rightarrow b\bar{b}$  scale factor calculation.

can conclude from table 5.7, that, in order to improve the  $g \rightarrow b\bar{b}$  calibration analysis and to reduce the scale factor uncertainties, better understanding of the transverse impact parameter modelling in the ATLAS inner detector is essential. Furthermore improved flavour tagging calibrations of the light-jet and  $c$ -jet mis-tag rates could yield a substantial reduction of uncertainty. Furthermore, additional simulation statistics would be desirable.

Systematic Variation	Impact on Scale Factor
Jet energy & mass scale	up -1.3%, down 4%
Jet energy resolution	0.3%
Jet mass resolution	0.9%
Jet JVT efficiency	up -0.3%, down -0.3%
Conversion	up -0.9%, down 1.2%
Hadronic material interactions	up 0.7%, down -0.6%
Light long-lived particles	up 0.4%, down -0.3%
Fake muons	0.13%
transverse IP smearing	12%
Non-BB flavour tagging	up -10%, down 8%
Modelling (PYTHIA vs. HERWIG)	10%
Semi-leptonic-inclusive extrapolation	4%
Template statistical uncertainty	10%
Data statistical uncertainty	8%
<b>total</b>	<b>24 %</b>

Table 5.7: Exemplary breakdown of the fractional impact of the statistical and systematic uncertainties on the double- $b$ -tagging scale factors in the category ( $\mu$ -jet  $p_T > 200$  GeV, non- $\mu$ -jet  $p_T > 300$  GeV ).

## 5.9 Conclusions and outlook

A direct calibration of track jet double- $b$ -tagging has the advantage that it can take into account correlations between track jets in a close-by  $b$ -jet topology. Furthermore, it can be implemented in a straightforward way to be used for more complex  $X \rightarrow b\bar{b}$  identification algorithms, where the standard approach of multiplicative event weights fails, e.g. when several  $b$ -tagging working points or even multivariate  $X \rightarrow b\bar{b}$  tagging approaches which make use of the full MV2c10 discriminant and not a fixed working point. The results documented in this chapter show that it is possible to measure the double- $b$ -tagging efficiency in  $g \rightarrow b\bar{b}$  events and calculate calibration scale factors. It was shown that the associated uncertainties are of the order of twenty percent. The most important uncertainties in this calibration analysis stem from the understanding of the modelling of  $\langle s_{d_0} \rangle$  and flavour tagging uncertainties of the non-BB flavour components, as well as from the data and simulation statistical uncertainties. To improve the  $g \rightarrow b\bar{b}$ -based calibration in the future, studies of the track impact parameters in data and precise modelling of the inner detector are crucial. Furthermore it is important to advance the calibrations of light- and  $c$ -jet mis-tag efficiencies in data and to reduce the uncertainties of the corresponding scale factors. Studies aiming towards an advancement and optimization of the  $g \rightarrow b\bar{b}$ -based direct double- $b$ -tagging calibration have already begun within the *ATLAS* collaboration.

To apply the  $g \rightarrow b\bar{b}$ -based calibration in a physics analysis with boosted  $X \rightarrow b\bar{b}$  final states one further step is needed, to remove the dependence on the muon-in-jet selection criterion, namely to change the jet bins in which the scale factors are calculated from  $\mu$ -jet  $p_T$  and non- $\mu$ -jet  $p_T$  to leading and sub-leading track jet  $p_T$ . This way, the definition of the scale factors is more universal, such that analyses applying direct double- $b$ -tagging scale factors do not explicitly have to require a muon in one of the track jets. No significant change in the central value and uncertainties of the scale factors is expected. Also other variables may be used for the differential scale factor measurement, such as the  $\Delta R$  between the two

track jets associated to the large- $R$  jet.

To extend the presented  $g \rightarrow b\bar{b}$ -based double- $b$ -tagging calibration to a full calibration of an  $X \rightarrow b\bar{b}$  identification algorithm (which is using additional information to  $b$ -tagging, such as the jet mass or substructure) some careful considerations are necessary. In this context it is important to note that  $g \rightarrow b\bar{b}$  events, due to their origin from the decay of a color octet gluon, have different radiative properties as the decay of a color-singlet object into two  $b$ -jets such as a Higgs boson, in addition to the different gluon mass. Therefore a  $g \rightarrow b\bar{b}$ -based calibration so far can only serve as a background calibration for  $X \rightarrow b\bar{b}$  tagging, similar to a light-jet mis-identification-efficiency calibration for  $b$ -tagging in *ATLAS* (see 3.5.9). A calibration in events containing  $Z \rightarrow b\bar{b}$  decays can serve as a more signal-like calibration for  $X \rightarrow b\bar{b}$  tagging, however current studies showed that the signal-to-background ratio in  $Z \rightarrow b\bar{b}$  events is still small. Therefore  $g \rightarrow b\bar{b}$ -based calibrations are an important test bench to explore how  $X \rightarrow b\bar{b}$  identification algorithms can be calibrated directly in the future. The identification of boosted  $X \rightarrow b\bar{b}$  final states becomes more important, e.g. in searches for Mono-Higgs bosons, as well as for di-Higgs-boson final states extending to higher-mass ranges, therefore there is a growing need for a direct calibration method. The work performed in the course of this thesis has laid the foundation for such direct calibration approaches with the *ATLAS* detector.

---

## Conclusion

---

This thesis summarizes work performed on the search for new physics in final states with boosted Higgs bosons decaying to  $b$ -quarks with the *ATLAS* detector. The presented analysis makes use of the effective theory framework of Heavy Vector Triplets (HVT) that predicts a set of three vector resonances  $V'$  (two charged  $W'^{\pm}$  and one neutral  $Z'$ ) with masses in the TeV range. The aim is to search for evidence for a heavy resonance decaying via a SM  $W$  or  $Z$  boson and a SM Higgs boson into final states with leptons and  $b$ -jets. No significant excess of the data over the SM prediction are found, therefore limits are set on the production cross section of the HVT resonances. The presence of a  $W'^{\pm}$  is excluded for a mass  $m_{W'} < 2.67(2.82)$  TeV for the HVT Model A (Model B) and a  $Z'$  is excluded for masses up to  $m_{Z'} < 2.65(2.83)$  TeV. Assuming a mass degeneracy between  $W'^{\pm}$  and  $Z'$ , the HVT  $V'$  can be excluded for masses  $m_{V'} < 2.80(2.93)$  TeV for Model A (Model B).

One of the main challenges in the search for new physics in connection with the SM Higgs boson is the reconstruction of boosted Higgs boson decays. As shown in the VH resonance search, if the Higgs boson is produced in the decay of a heavy new particle it can obtain a large Lorentz-boost. In this thesis methods for reconstructing boosted Higgs bosons (*Higgs tagging*) in the decay mode  $H \rightarrow b\bar{b}$  are introduced. The main focus is on the study of jet properties and flavour tagging in close-by  $b$ -jet environments in data. Gluon splitting events in the mode  $g \rightarrow b\bar{b}$  are used as a proxy for collimated  $b$ -jet topologies, because gluons are produced much more abundantly in  $pp$ -collisions at the LHC than Higgs bosons. Finally, a new method to calibrate double  $b$ -tagging in  $g \rightarrow b\bar{b}$  events is presented. This new approach allows to derive efficiency scale factors for double  $b$ -tagging directly, which take into account correlations between the two small- $R$  track-jets in the boosted  $bb$ -topology. The derived double  $b$ -tagging scale factors are compatible with the scale factors determined for isolated  $b$ -jets in *ATLAS* and have a total uncertainty of the order of 20%.

The calibration of double- $b$ -tagging in  $g \rightarrow b\bar{b}$  events is a pioneering study for background mis-identification efficiency calibrations for Higgs tagging algorithms in general, e.g. using multivariate analysis techniques. Several important search analyses with boosted  $H \rightarrow b\bar{b}$  or  $X \rightarrow b\bar{b}$  decays, such as the search for a boosted Higgs boson in association with a dark matter particle, the study of boosted SM Higgs bosons, or low-mass resonances decaying to two  $b$ -quarks can benefit from the new calibration approach described in this thesis. Similar calibration approaches in more signal-like events such as  $Z \rightarrow b\bar{b}$  are currently under study.

It has been shown in this thesis, that, after the discovery of the Higgs boson and establishment of its

SM-like properties, the research connected to the Higgs sector still provides a wide scope for testing the SM and constraining new physics. The study of the boosted  $H \rightarrow b\bar{b}$  decay mode sparked the development of many new reconstruction techniques to meet the challenges posed by this unique hadronic final state. The development of a prototype for the calibration of boosted  $H \rightarrow b\bar{b}$  identification algorithms in *ATLAS* opens up new possibilities to employ more complicated algorithms in the future that can help improve the sensitivity of physics analyses in this final state.

## Additional information

### A.1 $VH$ resonance additional post-fit spectra

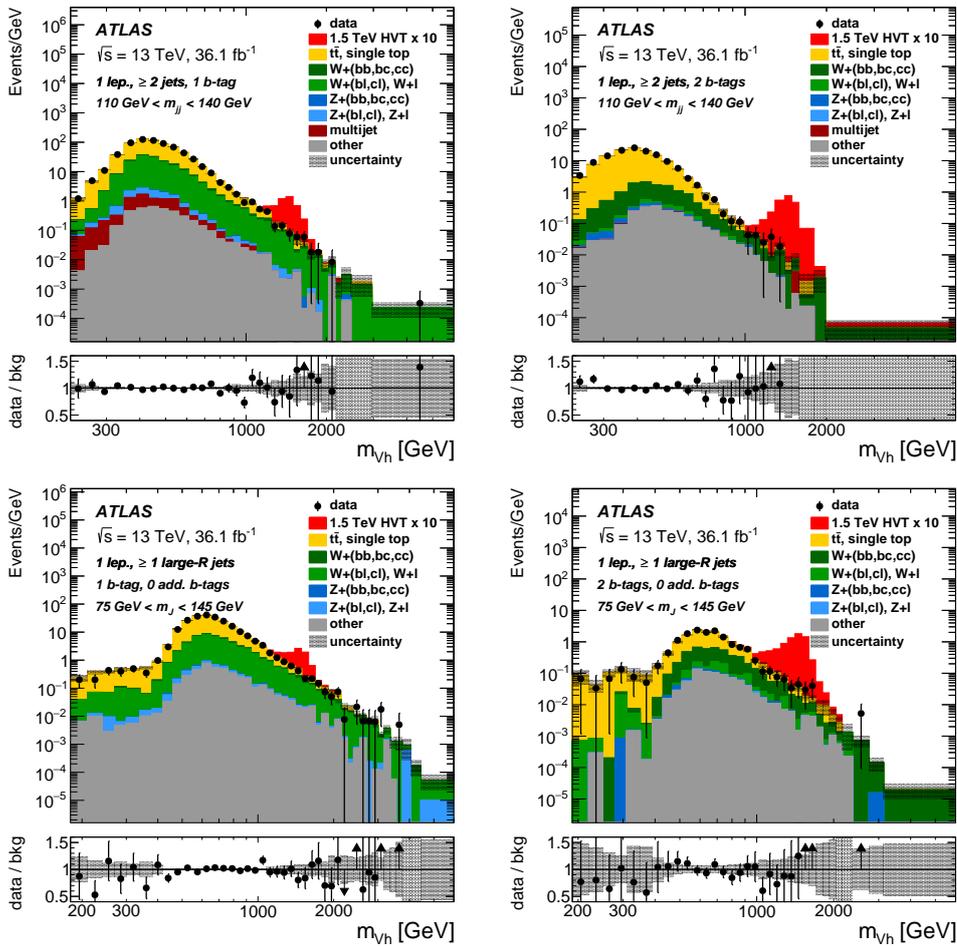


Figure A.1: Post-fit distribution of the transverse mass of the reconstructed  $VH$  system in the 1-lepton channel in the merged and resolved signal regions. From [77].

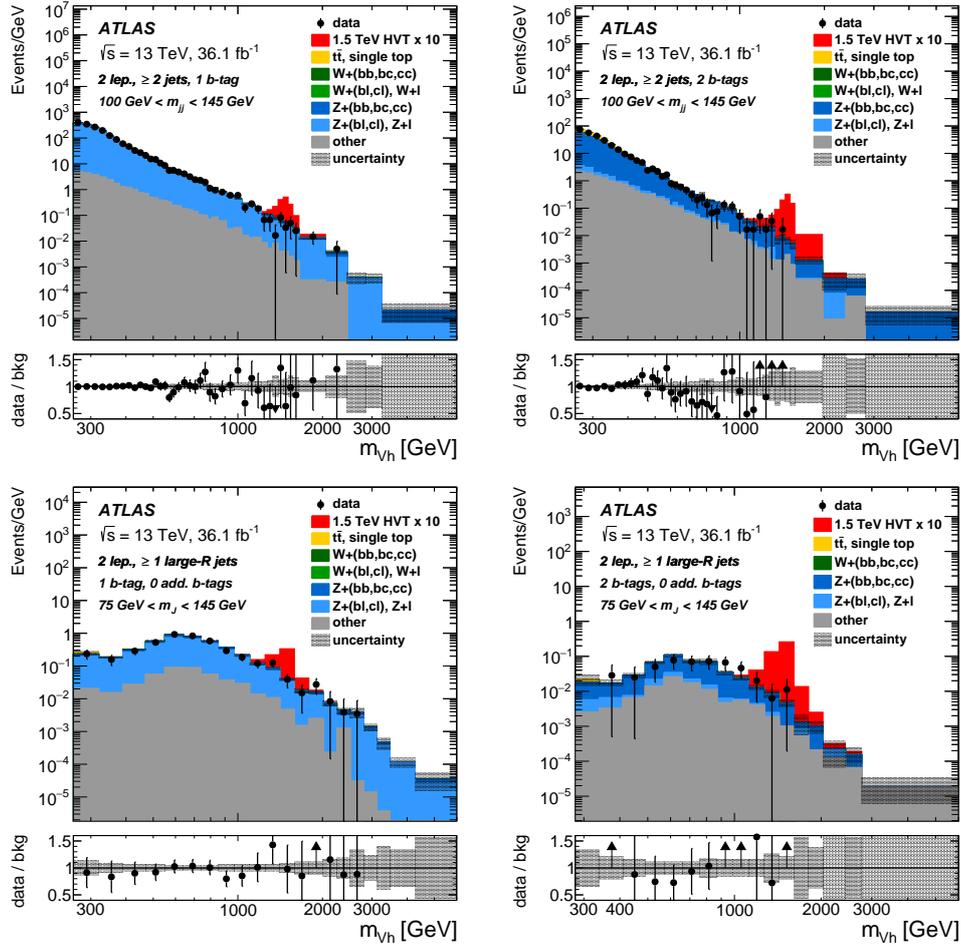


Figure A.2: Post-fit distribution of the transverse mass of the reconstructed  $VH$  system in the 2-lepton channel in the merged and resolved signal regions. From [77].

## A.2 $g \rightarrow b\bar{b}$ calibration additional studies and plots

### A.2.1 Studies of alternative template variables

An alternative flavour-sensitive variable for the flavour fraction template fit is the transverse impact parameter significance  $s_{d_0}$  (defined in equation 3.19) of the most significant track (with the largest absolute value  $|s_{d_0}|$ ), denoted forthwith as  $s_{d_0}^{\max}$ . Figure A.3 show the distributions of  $s_{d_0}^{\max}$  for the different flavour components of the  $g \rightarrow b\bar{b}$  candidate jet.

In comparison with the average transverse IP significance  $\langle s_{d_0} \rangle$  (see figure 5.3), the variable  $s_{d_0}^{\max}$  has more separation power between the different flavour fractions. However there is a shape in the post-fit Data/MC ratio for this variable, as can be seen in figure A.4.

Among the reasons for the mis-modelling observed in figure A.4 could be that the  $s_{d_0}^{\max}$  could be more sensitive to track outliers, which are difficult to describe correctly in simulation. As a result, the more

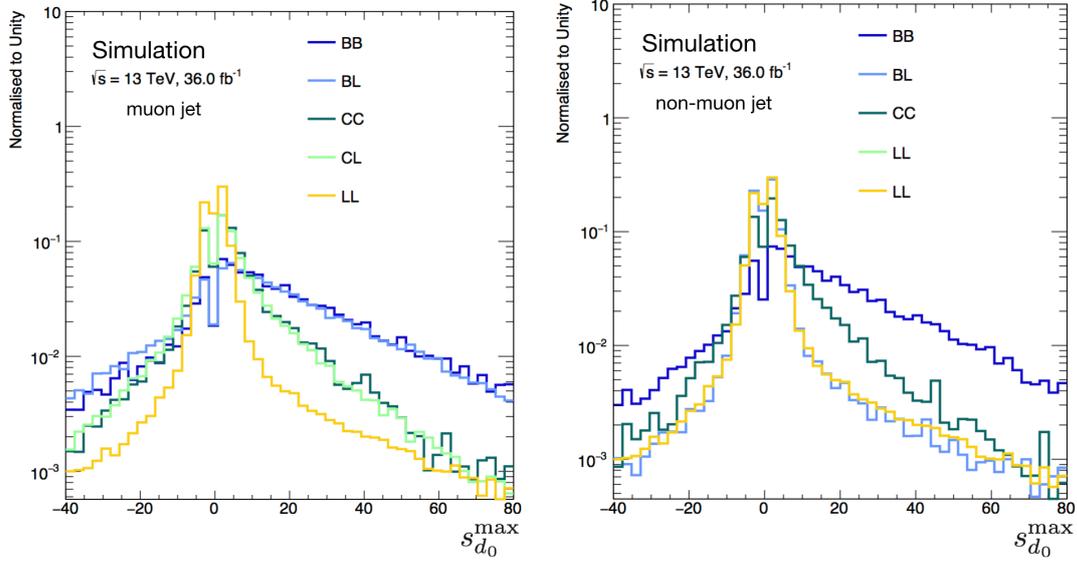


Figure A.3: Maximum impact parameter significance  $s_{d_0}^{max}$  distributions for the muon (left) and non-muon jets (right). The double flavour labels denote the flavour of the jet pair, with the muon-jet stated first.

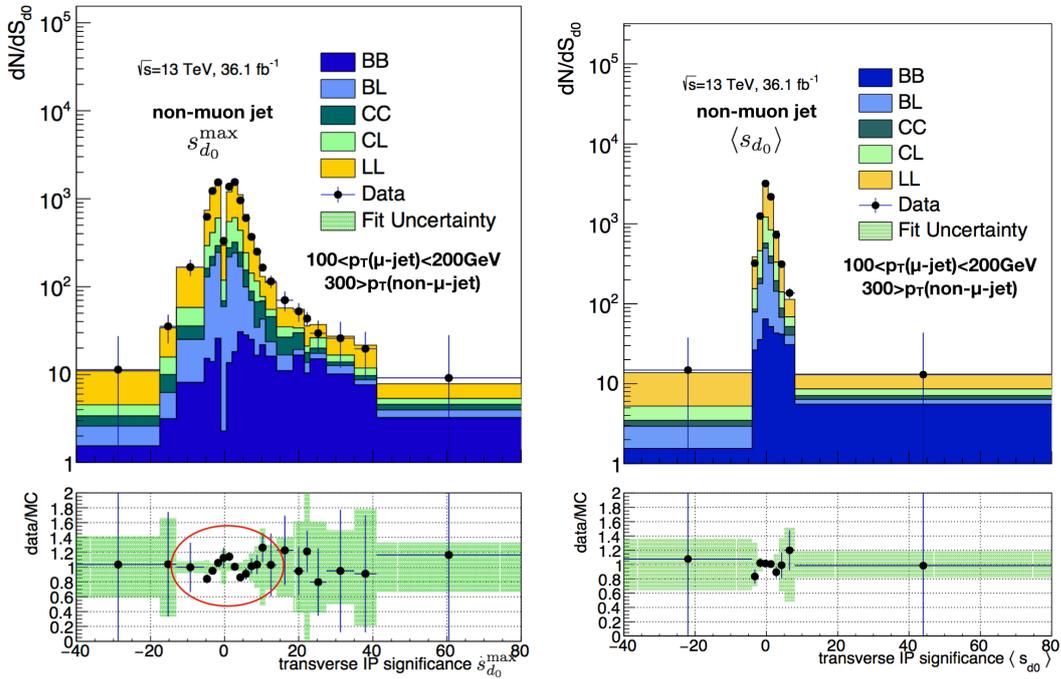


Figure A.4: Maximum impact parameter significance  $s_{d_0}^{max}$  (left) and average IP significance  $\langle s_{d_0} \rangle$  (right) distributions of the non-muon jet in the bin  $100 < p_T(\mu - \text{jet}) < 200 \text{ GeV}$ ,  $300 > p_T(!\mu - \text{jet})$ . The double flavour labels denote the flavour of the jet pair, with the muon-jet stated first.

robust  $\langle s_{d_0} \rangle$  was chosen as a fit variable for the  $g \rightarrow b\bar{b}$  calibration.

### A.2.2 Template similarity test

The pre-fit yields of all possible di-trackjet flavour components in the  $g \rightarrow b\bar{b}$  candidate jet are shown in table A.1<sup>1</sup>.

Flavour	Pre-fit fraction
BB	6.5%
BC	1.0%
BL	13.1%
CB	0.46%
CC	4.9%
CL	15.2%
LB	1.6%
LC	2.9%
LL	54.1%

Table A.1: Pre-fit flavour fractions before  $b$ -tagging. No split in 2-dimensional track-jet  $p_T$  bins is applied.

Templates with small yields ( $< 3\%$ ) should be merged into template with larger yields ( $> 13\%$ ). As described in section 5.4, the integral measure

$$S = \frac{1}{2} \int \frac{(p_1(x) - p_2(x))^2}{p_1(x) + p_2(x)} dx \quad (\text{A.1})$$

is used to determine the similarity between the template shapes, to decide which templates to merge. The results of the similarity check are shown in tables A.2:

From the highlighted numbers in table A.2 (c) it is visible that the merging prescription should be LB $\rightarrow$ LL, CB $\rightarrow$ CL, BC $\rightarrow$ BL, and LC $\rightarrow$ LL. This merging prescription agrees with the one found for a similar 8 TeV analysis [94].

### A.2.3 Template correlation test

For the correlations between the muon-jet  $Sd_0$  and the non-muon jet  $Sd_0$ , the following values were found:

The errors quoted are only the fit statistical errors. As can be seen the correlations between muon-jet and non-muon jet  $Sd_0$  are small.

### A.2.4 Small- $R$ jet trigger threshold studies

As described in section 5.3 data events for the  $g \rightarrow b\bar{b}$  calibration analysis are selected using a single small- $R$  jet trigger. This is due to large- $R$  jet triggers not being available yet in the considered dataset.

---

<sup>1</sup> This study was made for a previous *ATLAS* software version (R20.1) but the result is not expected to change

	BL	LL	CL
LB	0.5	0.02	0.16
CB	0.32	0.25	0.12
BC	0.09	0.44	0.21
LC	0.47	0.02	0.13

a)

	BL	LL	CL
LB	0.41	0.43	0.43
CB	0.50	0.52	0.52
BC	0.12	0.13	0.13
LC	0.14	0.16	0.16

b)

	BL	LL	CL
LB	0.46	0.23	0.29
CB	0.41	0.38	0.32
BC	0.10	0.28	0.17
LC	0.31	0.09	0.15

c)

Table A.2: Similarity measure  $S$  (see eq. A.1) for the muon-track jet (a), and non-muon-track jet (b) and the average  $S$  between muon- and non-muon track jet (c). No split in 2-dimensional track-jet  $p_T$  bins is applied. The highlighted numbers show the most similar large- and small-yield templates.

Flavour	Corr(muon-jet $S_{d0}$ , non-muon jet $S_{d0}$ ) [%]
BB	$4.5 \pm 0.9$
BL	$1.7 \pm 0.6$
CC	$1.2 \pm 0.9$
CL	$1.3 \pm 0.5$
LL	$0.1 \pm 0.3$

Table A.3: Correlations between the  $S_{d0}$  of the muon-jet and non-muon jet.

To avoid the trigger turn-on region, which is not well described in simulation, a cut is placed on the trigger-matched small- $R$  jet in the event to ensure that only events are considered for which the jet trigger is fully efficient. The  $g \rightarrow b\bar{b}$  candidate jet is a large- $R$  jet which is back-to-back to the small- $R$  trigger jet. Due to the difference in jet area, a bias is introduced, as can be seen as a discrepancy between data/MC in the  $g \rightarrow b\bar{b}$  candidate large- $R$  jet  $p_T$  in figure A.6. This bias is present, even if the probed jet is not used for triggering, due to the dijet balance in the event, which means that a cut on the trigger jet also affects the back-to-back  $g \rightarrow b\bar{b}$  candidate jet.

Figure A.6 shows a schematic view of how the bias due to the mis-match of the jet areas comes about. There is a mapping between small- $R$  and large- $R$  jets which is different for different types of jets. This leads to a shift in scale as well as a dilution of the trigger turn-on curve for the large- $R$  jet. This means that for a too low cut on the small- $R$  trigger jet, the trigger is not fully efficient regarding the large- $R$  jet.

To resolve this mis-match of the jet areas, either large- $R$  jet triggers must be used, which was not possible due to their absence in the dataset, but will be implemented in future versions of the  $g \rightarrow b\bar{b}$  calibration analysis. For the time being, a tighter requirement on the transverse momentum of the  $g \rightarrow b\bar{b}$  candidate jet  $p_T > 500$  GeV was introduced to remove the region where mis-modelling was observed.

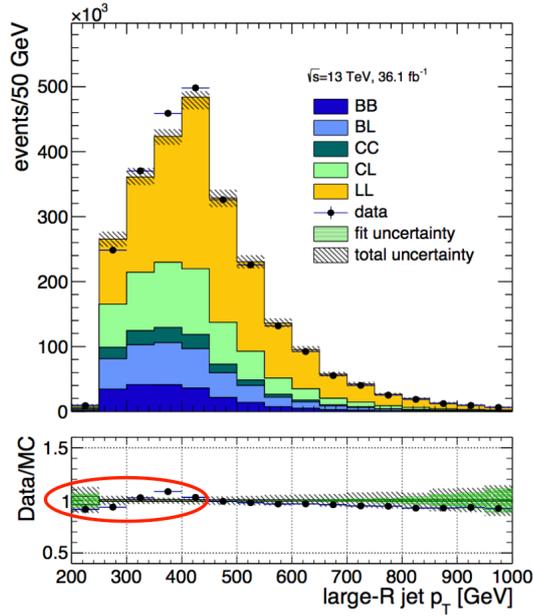


Figure A.5: Large- $R$  jet  $p_T$  of the  $g \rightarrow b\bar{b}$  candidate jet. A discrepancy between data and MC is observed in the low- $p_T$  regime, which is attributed to a bias due to the mis-match between the small- $R$  trigger-jet area and the large- $R$  probe jet.

### A.2.5 Additional studies for extrapolation from semi-leptonic to inclusive $b$ -jet sample

As described in section 5.7 the extrapolation uncertainty is estimated based on MC simulation only, assuming that the efficiencies change in the same way for data and MC. To validate this assumption, the influence of possible differences between data and simulation on the efficiency in the inclusive sample,  $\epsilon_{MC, \text{incl}}^{BB}$  is estimated. The efficiency is determined without splitting in categories of muon-jet and non-muon jet  $p_T$ . Among the effects studied the difference in the fraction of semi-leptonic  $B$ -hadron decays, the  $B$ -hadron  $p_T$  and the difference between two different MC generators. This study serves to estimate how much the extrapolation correction factor in data  $c_\epsilon^{\text{data}}$  is expected to vary with respect to  $c_\epsilon^{MC}$  and to test if the assumption that  $c_\epsilon^{\text{data}} \sim c_\epsilon^{MC}$  is justified.

#### $B$ -hadron decay fractions

The truth  $B$ -hadron fractions in the sample were reweighted to the fragmentation fractions and branching ratio values currently measured in data, based on [3]. The difference between the efficiencies observed between the nominal and the reweighted sample is below 0.1%. This is expected, since the MC samples used for these studies make use of the EVTGEN [99] package to correct the  $b$ - and  $c$ -hadron decay fractions in the generated sample to match the values measured in data.

#### $B$ -hadron $p_T$ spectrum

Differences in the modelling of the  $B$ -hadron fragmentation could manifest themselves in a difference of the  $B$ -hadron  $p_T$  spectrum, which in turn could affect the  $b$ -tagging performance. To study the impact of this, the sample was reweighted based on the truth  $B$ -hadron  $p_T$  spectrum, weighing events such that the

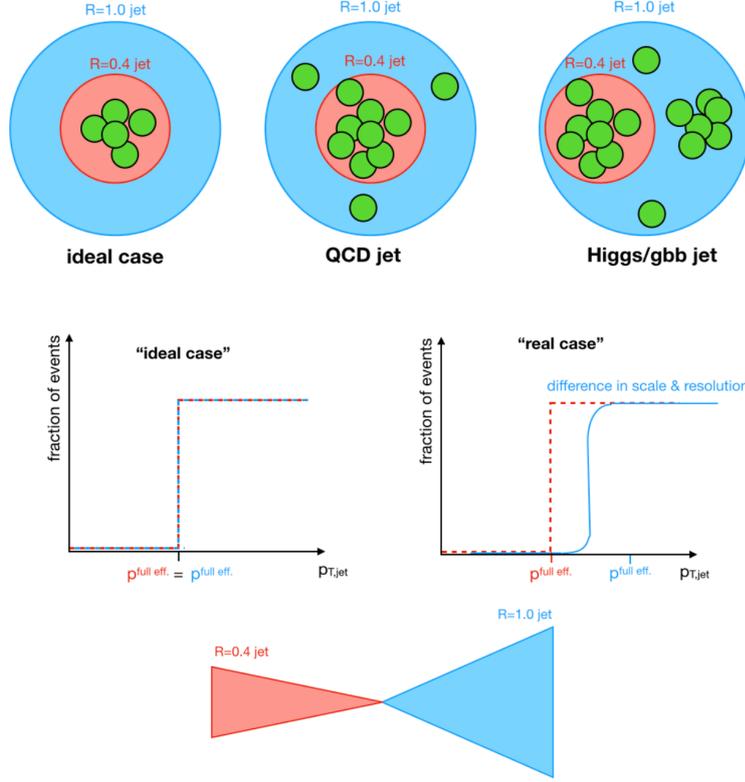


Figure A.6: Schematic explanation of the bias introduced by the mis-match between the small- $R$  trigger-jet area and the large- $R$  probe jet in the  $g \rightarrow b\bar{b}$  calibration analysis. The mis-match due to different jet catchment areas (top row, for different cases) leads to a shift and dilution of the trigger turn on curve for the large- $R$  jet (middle row). The dijet balance (bottom row) in the event leads to the bias taking effect even though the probe jet is required to be back-to-back to the trigger jet.

$B$ -hadron  $p_T$  scale is shifted by 10% up and down. The efficiencies  $\epsilon_{MC, \text{incl}}^{BB}$  measured for the nominal and reweighted samples differ by 1.2%.

### Generator difference PYTHIA/HERWIG

Differences due to the parameters for event generation or introduced by the shower model could affect the double  $b$ -tagging efficiencies. To estimate the size of this effect, the efficiency  $\epsilon_{MC, \text{incl}}^{BB}$  was calculated once in the sample generated with PYTHIA and once in the sample generated with HERWIG. No difference was observed.

In summary, none of the variations studied above showed a significant impact on the behavior of the MC double- $b$ -tagging efficiency in the inclusive sample. The conclusion is therefore that the estimation of the extrapolation uncertainties based on MC only is supportable in this study.

### A.2.6 $\langle s_{d_0} \rangle$ post-fit plots

This section documents the postfit  $\langle s_{d_0} \rangle$  distributions of all fit categories split in  $p_T(\mu - \text{jet})$ ,  $p_T(!\mu - \text{jet})$ . The distributions are shown separately for the muon-jet (fig. [A.7](#) and [A.8](#)) and the non-muon jet (fig. [A.9](#) and [A.10](#)).

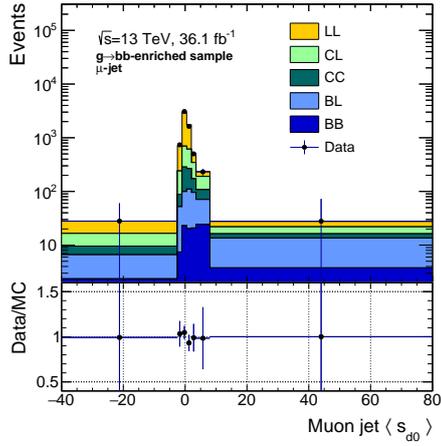
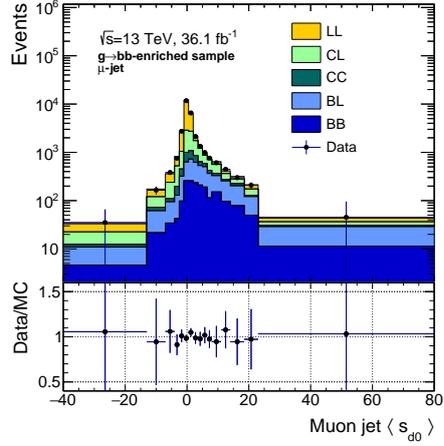
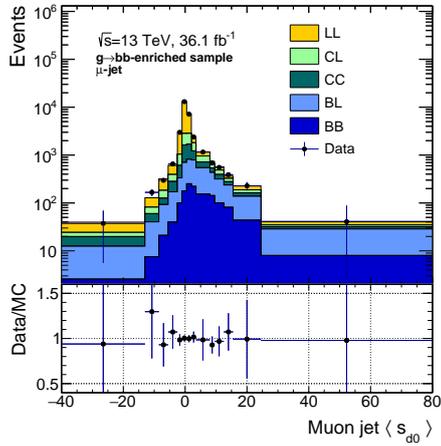
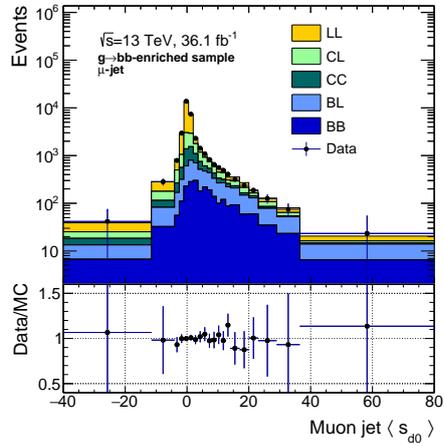
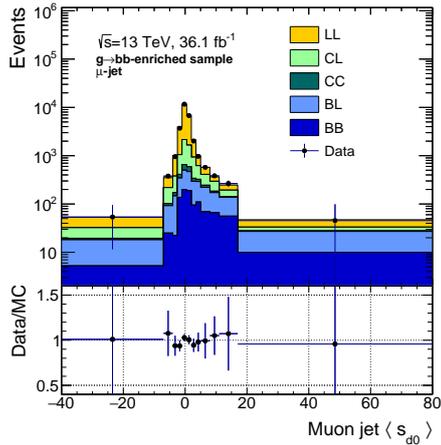
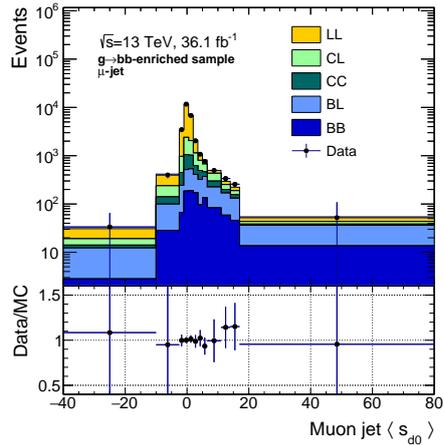

(a)  $100 > p_T^{\mu-jet}$ ,  $100 > p_T^{\mu-jet}$ 

(b)  $100 > p_T^{\mu-jet}$ ,  $100 < p_T^{\mu-jet} < 200$ 

(c)  $100 > p_T^{\mu-jet}$ ,  $200 < p_T^{\mu-jet} < 300$ 

(d)  $100 > p_T^{\mu-jet}$ ,  $300 < p_T^{\mu-jet}$ 

(e)  $100 < p_T^{\mu-jet} < 200$ ,  $100 > p_T^{\mu-jet}$ 

(f)  $100 < p_T^{\mu-jet} < 200$ ,  $100 < p_T^{\mu-jet} < 200$ 

Figure A.7: Post-fit distributions of the flavour-sensitive variable  $\langle s_{d_0} \rangle$  of the  $\mu$ -jet in two-dimensional bins of  $\mu$ -jet and non- $\mu$ -jet  $p_T$ . The bin boundaries in the individual captions are given in GeV.

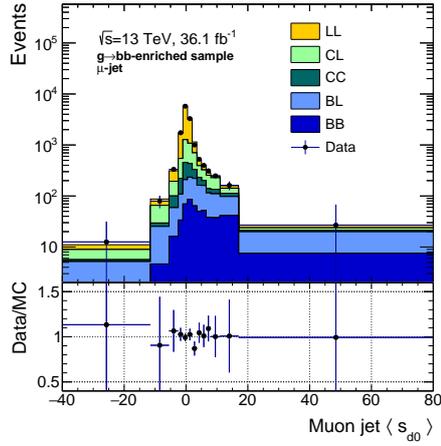
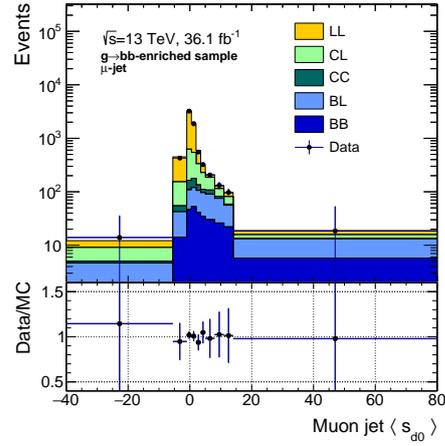
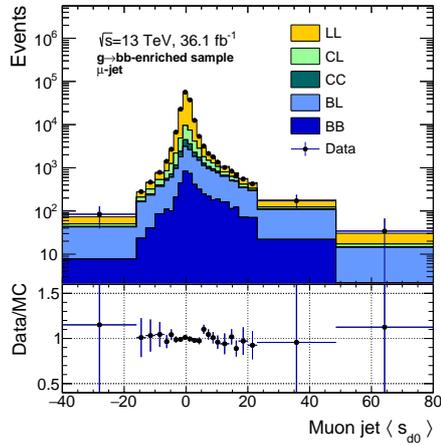
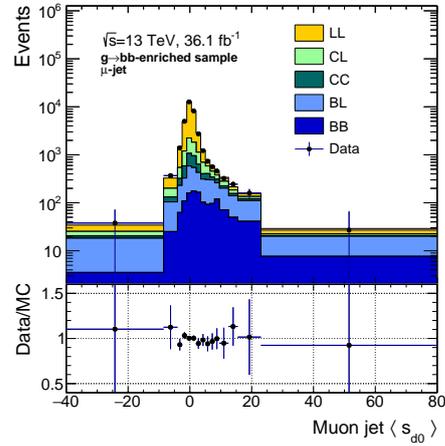
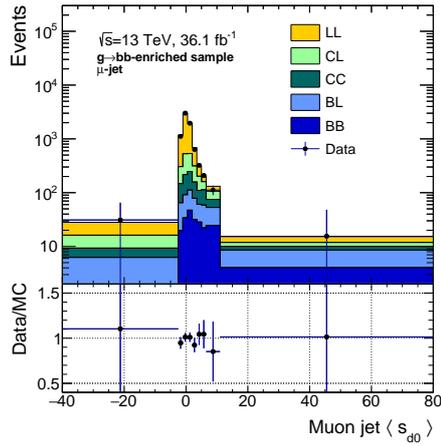
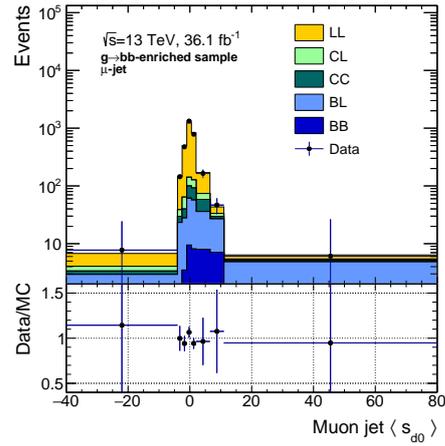

 (a)  $100 < p_T^{\mu-jet} < 200, 200 < p_T^{\mu-jet} < 300$ 

 (b)  $100 < p_T^{\mu-jet} < 200, 300 < p_T^{\mu-jet}$ 

 (c)  $200 < p_T^{\mu-jet}, 100 > p_T^{\mu-jet}$ 

 (d)  $200 < p_T^{\mu-jet}, 100 < p_T^{\mu-jet} < 200$ 

 (e)  $200 < p_T^{\mu-jet}, 200 < p_T^{\mu-jet} < 300$ 

 (f)  $200 < p_T^{\mu-jet}, 300 < p_T^{\mu-jet}$ 

 Figure A.8: Post-fit distributions of the flavour-sensitive variable  $\langle s_{d_0} \rangle$  of the  $\mu$ -jet in two-dimensional bins of  $\mu$ -jet and non- $\mu$ -jet  $p_T$ . The bin boundaries in the individual captions are given in GeV.

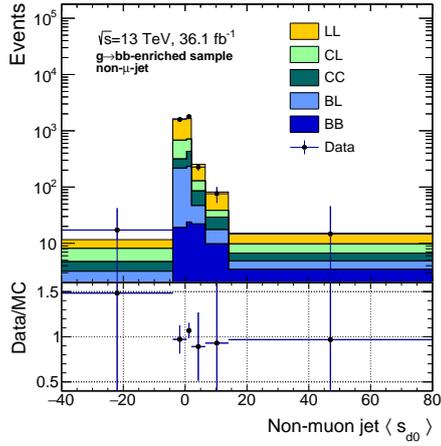
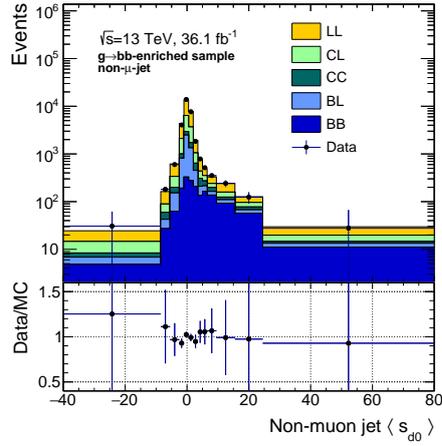
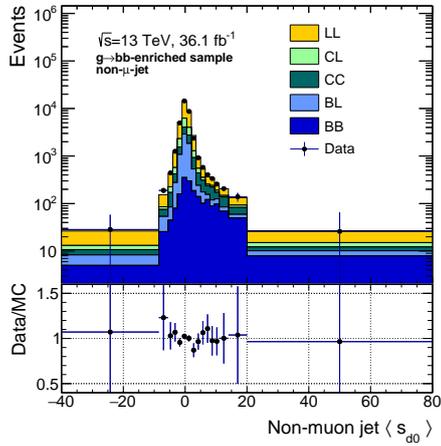
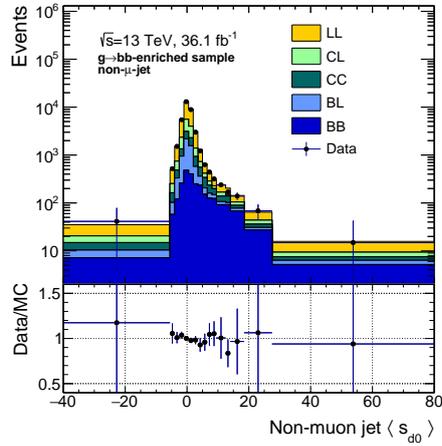
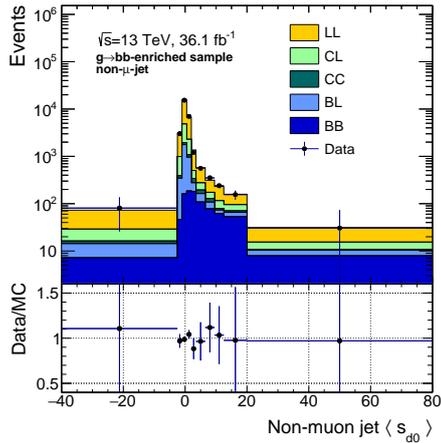
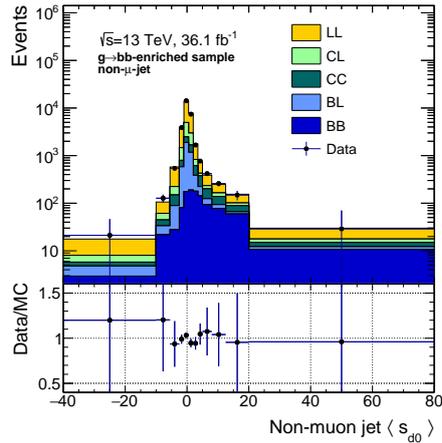

(a)  $100 > p_T^{\mu-jet}, 100 > p_T^{\mu-jet}$ 

(b)  $100 > p_T^{\mu-jet}, 100 < p_T^{\mu-jet} < 200$ 

(c)  $100 > p_T^{\mu-jet}, 200 < p_T^{\mu-jet} < 300$ 

(d)  $100 > p_T^{\mu-jet}, 300 < p_T^{\mu-jet}$ 

(e)  $100 < p_T^{\mu-jet} < 200, 100 > p_T^{\mu-jet}$ 

(f)  $100 < p_T^{\mu-jet} < 200, 100 < p_T^{\mu-jet} < 200$ 

Figure A.9: Post-fit distributions of the flavour-sensitive variable  $\langle s_{d_0} \rangle$  of the non- $\mu$ -jet in two-dimensional bins of  $\mu$ -jet and non- $\mu$ -jet  $p_T$ . The bin boundaries in the individual captions are given in GeV.

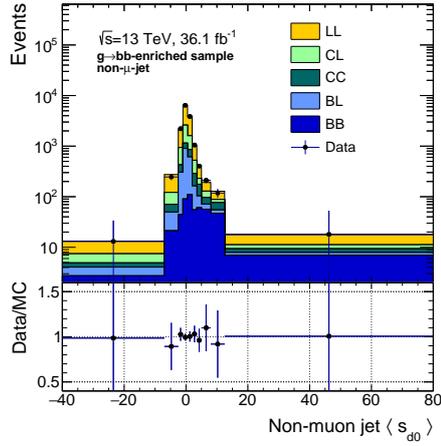
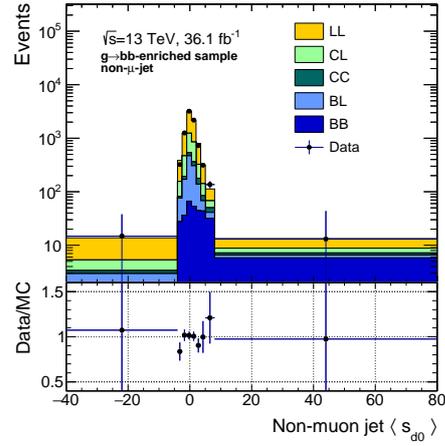
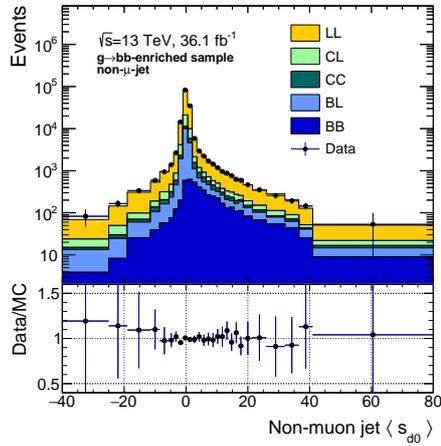
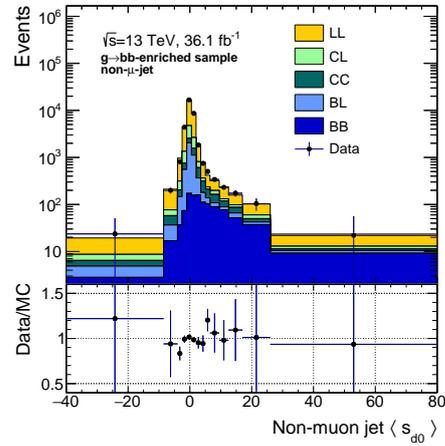
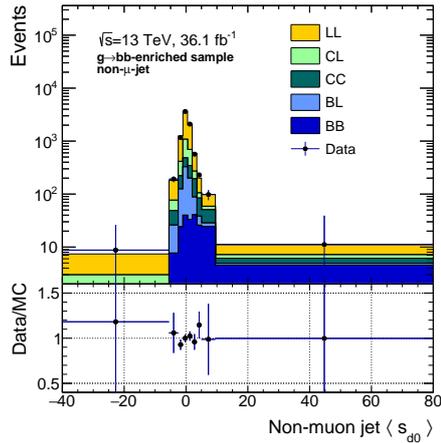
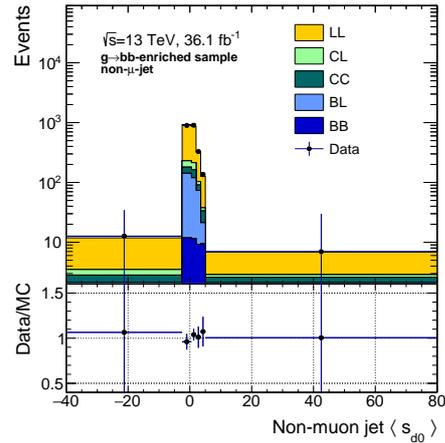

 (a)  $100 < p_T^{\mu-jet} < 200, 200 < p_T^{\mu-jet} < 300$ 

 (b)  $100 < p_T^{\mu-jet} < 200, 300 < p_T^{\mu-jet} < 300$ 

 (c)  $200 < p_T^{\mu-jet}, 100 > p_T^{\mu-jet}$ 

 (d)  $200 < p_T^{\mu-jet}, 100 < p_T^{\mu-jet} < 200$ 

 (e)  $200 < p_T^{\mu-jet}, 200 < p_T^{\mu-jet} < 300$ 

 (f)  $200 < p_T^{\mu-jet}, 300 < p_T^{\mu-jet}$ 

 Figure A.10: Post-fit distributions of the flavour-sensitive variable  $\langle s_{d0} \rangle$  of the non- $\mu$ -jet in two-dimensional bins of  $\mu$ -jet and non- $\mu$ -jet  $p_T$ . The bin boundaries in the individual captions are given in GeV.

# Bibliography

---

- [1] Particle Data Group, LBNL, 2015,  
URL: <http://www.particleadventure.org/history-universe.html> (cit. on p. 1).
- [2] E. Drexler, *WikiMedia commons: Summary of interactions between particles described by the Standard Model*, [Online; accessed 13-June-2019], 2014, URL:  
[https://en.wikipedia.org/wiki/Standard%5C\\_Model#/media/File:Elementary%5C\\_particle%5C\\_interactions%5C\\_in%5C\\_the%5C\\_Standard%5C\\_Model.png](https://en.wikipedia.org/wiki/Standard%5C_Model#/media/File:Elementary%5C_particle%5C_interactions%5C_in%5C_the%5C_Standard%5C_Model.png)  
(cit. on p. 6).
- [3] M. Tanabashi et al., *Review of Particle Physics*, *Phys. Rev. D* **98** (3 2018) 030001,  
URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>  
(cit. on pp. 5, 9, 15, 30, 39, 69, 70, 126).
- [4] F. Halzen and A. Martin, *Quarks & Leptons: An introductory course in modern particle physics*, John Wiley & Sons, 1984 (cit. on p. 7).
- [5] J. Ellis, *Higgs Physics*, (2013) 117, 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report,  
URL: <https://cds.cern.ch/record/1638469> (cit. on p. 8).
- [6] LHC Higgs Cross Section Working Group,  
URL: <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>  
(cit. on pp. 9, 10).
- [7] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett. B* **716** (2012) 1,  
arXiv: 1207.7214 [hep-ex] (cit. on p. 9).
- [8] CMS Collaboration,  
*Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*,  
*Phys. Lett. B* **716** (2012) 30, arXiv: 1207.7235 [hep-ex] (cit. on p. 9).
- [9] ATLAS Collaboration,  
*Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*,  
*Eur. Phys. J. C* **75** (2015) 476, arXiv: 1506.05669 [hep-ex] (cit. on p. 9),  
Erratum: *Eur. Phys. J. C* **76** (2016) 152.
- [10] V. Khachatryan et al., *Constraints on the spin-parity and anomalous HVV couplings of the Higgs boson in proton collisions at 7 and 8 TeV*, *Phys. Rev. D* **92** (2015) 012004,  
arXiv: 1411.3441 [hep-ex] (cit. on p. 9).
- [11] 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*,  
*Phys. Rev. Lett.* **114** (2015) 191803, arXiv: 1503.07589 [hep-ex] (cit. on p. 9).

- [12] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, *Phys. Lett. B* **784** (2018) 173, arXiv: [1806.00425 \[hep-ex\]](#) (cit. on p. 9).
- [13] ATLAS Collaboration, *Observation and measurement of Higgs boson decays to  $WW^*$  with the ATLAS detector*, *Phys. Rev. D* **92** (2015) 012006, arXiv: [1412.2641 \[hep-ex\]](#) (cit. on p. 9).
- [14] ATLAS Collaboration, *Observation of  $H \rightarrow b\bar{b}$  decays and  $VH$  production with the ATLAS detector*, *Phys. Lett. B* **786** (2018) 59, arXiv: [1808.08238 \[hep-ex\]](#) (cit. on pp. 9, 58).
- [15] ATLAS Collaboration, *Cross-section measurements of the Higgs boson decaying into a pair of  $\tau$ -leptons in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Rev.* (2018), arXiv: [1811.08856 \[hep-ex\]](#) (cit. on p. 9).
- [16] S. Weinberg, *Gauge hierarchies*, *Phys. Lett. B* **82** (1979) 387 (cit. on p. 9).
- [17] M. J. G. Veltman, *The infrared - ultraviolet connection*, *Acta Phys. Polon. B* **12** (1981) 437 (cit. on p. 9).
- [18] C. H. Llewellyn Smith and G. G. Ross, *The real gauge hierarchy problem*, *Phys. Lett. B* **105** (1981) 38 (cit. on p. 9).
- [19] *Review of Particle Physics: 11. Status of Higgs Boson physics*, *Phys. Rev. D* **98** (3 2018) 030001, URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001> (cit. on p. 10).
- [20] F. Sannino and K. Tuominen, *Orientifold theory dynamics and symmetry breaking*, *Phys. Rev. D* **71** (2005) 051901, arXiv: [0405209 \[hep-ph\]](#) (cit. on p. 11).
- [21] R. Foadi et al., *Minimal walking technicolor: set up for collider physics*, *Phys. Rev. D* **76** (2007) 055005, arXiv: [0706.1696 \[hep-ph\]](#) (cit. on p. 11).
- [22] A. Belyaev et al., *Technicolor walks at the LHC*, *Phys. Rev. D* **79** (2009) 035006, arXiv: [0809.0793 \[hep-ph\]](#) (cit. on p. 11).
- [23] M. Schmaltz and D. Tucker-Smith, *Little Higgs review*, *Ann. Rev. Nucl. Part. Sci.* **55** (2005) 229, arXiv: [0502182 \[hep-ph\]](#) (cit. on p. 11).
- [24] M. J. Dugan, H. Georgi and D. B. Kaplan, *Anatomy of a composite Higgs model*, *Nucl. Phys. B* **254** (1985) 299 (cit. on p. 11).
- [25] K. Agashe, R. Contino and A. Pomarol, *The minimal composite Higgs model*, *Nucl. Phys. B* **719** (2005) 165, arXiv: [0412089 \[hep-ph\]](#) (cit. on p. 11).
- [26] D. Alves et al., *Simplified models for LHC new physics searches*, *J. Phys. G* **39** (2012) 105005, arXiv: [1105.2838 \[hep-ph\]](#) (cit. on p. 11).
- [27] D. Pappadopulo et al., *Heavy vector triplets: bridging theory and data*, *JHEP* **09** (2014) 060, arXiv: [1402.4431 \[hep-ph\]](#) (cit. on pp. 11, 53, 56, 90).
- [28] J. de Blas, J. M. Lizana and M. Perez-Victoria, *Combining searches of  $Z'$  and  $W'$  bosons*, *JHEP* **01** (2013) 166, arXiv: [1211.2229 \[hep-ph\]](#) (cit. on pp. 11, 53).
- [29] ATLAS Luminosity Working Group (Public Results), URL: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2> (cit. on pp. 14, 16).

- [30] J. Rojo et al., *The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II*, *J. Phys.* **G42** (2015) 103103, arXiv: [1507.00556 \[hep-ph\]](#) (cit. on p. 15).
- [31] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1*, Technical Design Report ATLAS, CERN, 1999, URL: <https://cds.cern.ch/record/391176> (cit. on pp. 16, 17, 19, 20).
- [32] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003 (cit. on p. 17).
- [33] The ATLAS Collaboration, ATLAS Photos, ATLAS Experiment, ©2014 CERN, URL: <http://www.atlas.ch/photos/> (cit. on p. 18).
- [34] H. Kolanoski and N. Wermes, *Teilchendetektoren*, Springer, 2016, ISBN: 9783662453490, 9783662453506 (cit. on p. 18).
- [35] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J. C* **77** (2017) 490, arXiv: [1603.02934 \[hep-ex\]](#) (cit. on p. 19).
- [36] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, *Eur. Phys. J. C* **77** (2017) 317, arXiv: [1611.09661 \[hep-ex\]](#) (cit. on pp. 20, 65, 66).
- [37] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, arXiv: [1005.4568 \[physics.ins-det\]](#) (cit. on p. 21).
- [38] T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **0605** (2006) 026, arXiv: [0603175 \[hep-ph\]](#) (cit. on p. 21).
- [39] T. Sjöstrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852, arXiv: [0710.3820 \[hep-ph\]](#) (cit. on p. 21).
- [40] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152 (cit. on p. 21).
- [41] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, arXiv: [0409146 \[hep-ph\]](#) (cit. on p. 21).
- [42] 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 \[hep-ph\]](#) (cit. on p. 21).
- [43] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, arXiv: [1002.2581 \[hep-ph\]](#) (cit. on p. 21).
- [44] M. Bahr et al., *Herwig++ Physics and Manual*, *Eur. Phys. J. C* **58** (2008) 639, arXiv: [0803.0883 \[hep-ph\]](#) (cit. on p. 21).
- [45] S. Agostinelli et al., *GEANT4 - a simulation toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250 (cit. on pp. 22, 64).
- [46] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, *Eur. Phys. J. C* **77** (2017) 673, arXiv: [1704.07983 \[hep-ex\]](#) (cit. on pp. 22, 23).
- [47] M. Cacciari, G. P. Salam and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, *JHEP* **04** (2008) 063, arXiv: [0802.1189 \[hep-ph\]](#) (cit. on pp. 24, 25).

- [48] M. Cacciari, G. P. Salam and G. Soyez, *FastJet User Manual*, *Eur. Phys. J. C* **72** (2012) 1896, arXiv: [1111.6097 \[hep-ph\]](#) (cit. on pp. 25, 58).
- [49] ATLAS Collaboration, *Performance of jet substructure techniques for large- $R$  jets in proton–proton collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector*, *JHEP* **09** (2013) 076, arXiv: [1306.4945 \[hep-ex\]](#) (cit. on pp. 26, 103).
- [50] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, *JHEP* **02** (2010) 084, arXiv: [0912.1342 \[hep-ph\]](#) (cit. on p. 26).
- [51] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Rev. D* **96** (2017) 072002, arXiv: [1703.09665 \[hep-ex\]](#) (cit. on p. 26).
- [52] ATLAS Collaboration, *In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 135, arXiv: [1807.09477 \[hep-ex\]](#) (cit. on pp. 26, 27, 102, 103).
- [53] ATLAS Collaboration, *Identification of boosted Higgs bosons decaying into  $b$ -quark pairs with the ATLAS detector at 13 TeV*, (2019), arXiv: [1906.11005 \[hep-ex\]](#) (cit. on pp. 27, 38–41, 58, 73, 93, 99, 100, 103, 105–108).
- [54] ATLAS Collaboration, *Identification of boosted, hadronically decaying  $W$  bosons and comparisons with ATLAS data taken at  $\sqrt{s} = 8$  TeV*, *Eur. Phys. J. C* **76** (2016) 154, arXiv: [1510.05821 \[hep-ex\]](#) (cit. on p. 27).
- [55] ATLAS Collaboration, *Identification of high transverse momentum top quarks in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, *JHEP* **06** (2016) 093, arXiv: [1603.03127 \[hep-ex\]](#) (cit. on p. 27).
- [56] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with  $N$ -subjettiness*, *JHEP* **03** (2011) 015, arXiv: [1011.2268 \[hep-ph\]](#) (cit. on p. 27).
- [57] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing  $N$ -subjettiness*, *JHEP* **02** (2012) 093, arXiv: [1108.2701 \[hep-ph\]](#) (cit. on p. 27).
- [58] A. J. Larkoski and others, *Power Counting to Better Jet Observables*, *JHEP* **12** (2014) 009, arXiv: [1409.6298 \[hep-ph\]](#) (cit. on p. 27).
- [59] A. J. Larkoski and others, *Energy Correlation Functions for Jet Substructure*, *JHEP* **06** (2013) 108, arXiv: [1305.0007 \[hep-ph\]](#) (cit. on p. 27).
- [60] R. Jacobs, *Identification of  $b$ -jets in boosted  $H \rightarrow b\bar{b}$  decays and a study of the decay  $V'^{\pm} \rightarrow WH \rightarrow l\nu b\bar{b}$  in  $\sqrt{s} = 13$  TeV  $pp$ -collisions with the ATLAS detector*, Presented 09/2015, 2015, URL: <http://www.hep1.physik.uni-bonn.de/results/data/internal/jacobs.pdf> (cit. on pp. 30, 33).
- [61] ATLAS Collaboration, *Performance of  $b$ -jet identification in the ATLAS experiment*, *JINST* **11** (2016) P04008, arXiv: [1512.01094 \[hep-ex\]](#) (cit. on pp. 30–33, 37, 62, 88, 95, 111, 112).
- [62] D. Guest, *Private communication*, 2019 (cit. on p. 32).
- [63] ATLAS Collaboration, *Commissioning of the ATLAS high performance  $b$ -tagging algorithms in the 7 TeV collision data*, ATLAS-CONF-2011-102, 2011, URL: <https://cds.cern.ch/record/1369219> (cit. on p. 34).

- [64] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, 2016, URL: <https://cds.cern.ch/record/2160731> (cit. on pp. 34, 35, 62, 88, 95).
- [65] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, (2009), arXiv: [0901.0512](https://arxiv.org/abs/0901.0512) [hep-ex] (cit. on p. 34).
- [66] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATL-PHYS-PUB-2015-022, 2015, URL: <https://cds.cern.ch/record/2037697> (cit. on pp. 34, 36).
- [67] ATLAS Collaboration, *Flavor Tagging with Track-Jets in Boosted Topologies with the ATLAS Detector*, ATL-PHYS-PUB-2014-013, 2014, URL: <https://cds.cern.ch/record/1750681> (cit. on p. 37).
- [68] ATLAS Collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using  $t\bar{t}$  events at  $\sqrt{s} = 13$  TeV*, *JHEP* **08** (2018) 089, arXiv: [1805.01845](https://arxiv.org/abs/1805.01845) [hep-ex] (cit. on pp. 37, 102).
- [69] ATLAS Collaboration, *Variable Radius, Exclusive- $k_T$ , and Center-of-Mass Subjet Reconstruction for Higgs( $\rightarrow b\bar{b}$ ) Tagging in ATLAS*, ATL-PHYS-PUB-2017-010, 2017, URL: <https://cds.cern.ch/record/2268678> (cit. on pp. 40, 42, 43, 92).
- [70] Y. L. Dokshitzer et al., *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, arXiv: [hep-ph/9707323](https://arxiv.org/abs/hep-ph/9707323) [hep-ph] (cit. on p. 42).
- [71] O. Behnke et al., *Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods*, 1st, Wiley-VCH, 2013, ISBN: 3527410589, 9783527410583 (cit. on p. 44).
- [72] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) 1554, arXiv: [1007.1727](https://arxiv.org/abs/1007.1727) [physics.data-an], Erratum: *Eur. Phys. J. C* **73** (2013) 2501 (cit. on pp. 48, 51, 72).
- [73] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, *Ann. Math. Statist.* **9** (1938) 60, URL: <https://doi.org/10.1214/aoms/1177732360> (cit. on p. 51).
- [74] A. Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, *Transactions of the American Mathematical Society* **54** (1943) 426, ISSN: 00029947, URL: <http://www.jstor.org/stable/1990256> (cit. on p. 51).
- [75] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003, arXiv: [physics/0306116](https://arxiv.org/abs/physics/0306116) [physics.data-an] (cit. on pp. 52, 72).
- [76] L. Moneta et al., *The RooStats project*, 2010, arXiv: [1009.1003](https://arxiv.org/abs/1009.1003) [hep-ph] (cit. on pp. 52, 72).
- [77] ATLAS Collaboration, *Search for heavy resonances decaying into a W or Z boson and a Higgs boson in final states with leptons and b-jets in  $36\text{fb}^{-1}$  of  $\sqrt{s} = 13$  TeV pp collisions with the ATLAS detector*, *JHEP* **03** (2018) 174, arXiv: [1712.06518](https://arxiv.org/abs/1712.06518) [hep-ex] (cit. on pp. 53, 64, 65, 74, 82, 87, 89–91, 121, 122).
- [78] V. D. Barger, W.-Y. Keung and E. Ma, *A gauge model with light W and Z bosons*, *Phys. Rev. D* **22** (1980) 727 (cit. on pp. 54, 55).

- [79] R. Contino et al., *On the effect of resonances in composite Higgs phenomenology*, *JHEP* **10** (2011) 081, arXiv: [1109.1570 \[hep-ph\]](#) (cit. on pp. 54, 55).
- [80] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton–proton collision data*, *Eur. Phys. J. C* **74** (2014) 2941, arXiv: [1404.2240 \[hep-ex\]](#) (cit. on p. 57).
- [81] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C* **76** (2016) 292, arXiv: [1603.05598 \[hep-ex\]](#) (cit. on pp. 57, 95, 96).
- [82] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at  $\sqrt{s} = 8$  TeV*, *Eur. Phys. J. C* **75** (2015) 303, arXiv: [1412.7086 \[hep-ex\]](#) (cit. on p. 58).
- [83] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at  $\sqrt{s} = 13$  TeV*, tech. rep. ATL-PHYS-PUB-2015-027, CERN, 2015, URL: <https://cds.cern.ch/record/2037904> (cit. on p. 59).
- [84] ATLAS Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, ATLAS-CONF-2016-035, 2016, URL: <https://cds.cern.ch/record/2200211> (cit. on p. 61).
- [85] ATLAS Collaboration, *Boosted Higgs ( $\rightarrow b\bar{b}$ ) Boson Identification with the ATLAS Detector at  $\sqrt{s} = 13$  TeV*, ATLAS-CONF-2016-039, 2016, URL: <https://cds.cern.ch/record/2206038> (cit. on p. 61).
- [86] M. Baak et al., *HistFitter software framework for statistical data analysis*, *Eur. Phys. J. C* **75** (2015) 153, arXiv: [1410.1280 \[hep-ex\]](#) (cit. on p. 72).
- [87] A. L. Read, *Presentation of search results: The  $CL_s$  technique*, *J. Phys. G* **28** (2002) 2693 (cit. on p. 72).
- [88] ATLAS Collaboration, *Track Reconstruction Performance of the ATLAS Inner Detector at  $\sqrt{s} = 13$  TeV*, ATL-PHYS-PUB-2015-018, 2015, URL: <https://cds.cern.ch/record/2037683> (cit. on pp. 95, 96).
- [89] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at  $\sqrt{s} = 13$  TeV*, ATL-PHYS-PUB-2015-026, 2015, URL: <https://cds.cern.ch/record/2037717> (cit. on p. 96).
- [90] ATLAS Collaboration, *Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015, URL: <https://cds.cern.ch/record/2037702> (cit. on p. 96).
- [91] ATLAS Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, ATL-DAQ-PUB-2016-001, 2016, URL: <https://cds.cern.ch/record/2136007> (cit. on p. 97).
- [92] ATLAS Collaboration, *Search for boosted resonances decaying to two b-quarks and produced in association with a jet at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, ATLAS-CONF-2018-052, 2018, URL: <https://cds.cern.ch/record/2649081> (cit. on p. 97).

- 
- [93] ATLAS Collaboration, *Search for heavy resonances decaying to a W or Z boson and a Higgs boson in the  $q\bar{q}^{(\prime)}b\bar{b}$  final state in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Lett. B* **774** (2017) 494, arXiv: [1707.06958](https://arxiv.org/abs/1707.06958) [[hep-ex](#)] (cit. on p. 97).
- [94] ATLAS Collaboration, *Studies of b-tagging performance and jet substructure in a high  $p_T$   $g \rightarrow b\bar{b}$  rich sample of large-R jets from pp collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, ATLAS-CONF-2016-002, 2016, URL: <https://cds.cern.ch/record/2135187> (cit. on pp. 98, 124).
- [95] ATLAS Collaboration, *Properties of  $g \rightarrow b\bar{b}$  at small opening angles in pp collisions with the ATLAS detector at  $\sqrt{s} = 13$  TeV*, (2018), arXiv: [1812.09283](https://arxiv.org/abs/1812.09283) [[hep-ex](#)] (cit. on p. 98).
- [96] T. Lenz and P. Maettig, *Prospects for  $t\bar{t}$  resonance searches at ATLAS*, Presented 21 Dec 2010, 2010, URL: <https://cds.cern.ch/record/1366357> (cit. on p. 99).
- [97] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in  $t\bar{t}$  events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001, 2018, URL: <https://cds.cern.ch/record/2306649> (cit. on p. 102).
- [98] ATLAS Collaboration, *Identification of Boosted, Hadronically-Decaying W and Z Bosons in  $\sqrt{s} = 13$  TeV Monte Carlo Simulations for ATLAS*, ATL-PHYS-PUB-2015-033, 2015, URL: <https://cds.cern.ch/record/2041461> (cit. on p. 103).
- [99] D.J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instr. Meth. A* **462** (2001) 152 (cit. on p. 126).