

**Masterarbeit**  
zur Erlangung des Hochschulgrades  
Master of Science

**Boosted Decision Trees in the  
analysis of  $A/H \rightarrow \tau^+ \tau^-$  with  
ATLAS**

vorgelegt von

Max Märker  
geboren am 22.04.1992

TECHNISCHE UNIVERSITÄT DRESDEN  
INSTITUT FÜR KERN- UND TEILCHENPHYSIK  
FACHRICHTUNG PHYSIK  
FAKULTÄT FÜR MATHEMATIK UND NATURWISSENSCHAFTEN

April 18, 2017

**Datum der Abgabe:** 19.04.2017  
**Betreuer:** Dr. Wolfgang Mader, Dipl. Phys. Lorenz Hauswald  
**Erstprüfer:** Prof. Dr. Arno Straessner  
**Zweitprüfer:** Prof. Dr. Michael Kobel

## Abstract

The Minimal Supersymmetric Standard Model (MSSM) is an extension of the Standard Model (SM) of particle physics, that is capable of resolving many problems, like the fine tuning problem of the Higgs mass corrections, that are unanswered in the SM. The MSSM postulates the existence of two additional heavy neutral Higgs bosons, the CP-even H boson and the CP-odd A boson, next to the SM like Higgs particle. The search for these heavy scalar bosons might reveal new physics beyond the Standard Model.

An analysis is presented searching for heavy neutral Higgs resonances in the channel  $A/H \rightarrow \tau\tau$  where both tau leptons decay into hadrons using  $13.2 \text{ fb}^{-1}$  of data recorded by the ATLAS detector. To enhance the sensitivity of the analysis Boosted Decision Trees (BDT) are used. Since no excess of data over the SM expectancy is observed, upper limits in the  $m_A\text{-tan}(\beta)$  parameter plane in the  $m_h^{\text{mod}+}$  scenario of the MSSM are set to exclude certain parameter regions. The BDT approach improved the limits when compared to previous results.

## Kurzfassung

Das Minimal Supersymmetrische Standardmodell (MSSM) ist eine Erweiterung des Standardmodells (SM) der Teilchenphysik, welches in der Lage ist viele Fragen zu beantworten die sich nicht im SM erklären lassen, wie z.B. das Problem der Feinabstimmung der Korrekturen der Higgsmasse. Das MSSM sagt die Existenz zusätzlicher schwerer neutraler Higgs Bosonen vorher, namentlich das CP-gerade H Boson und das CP-ungerade A Boson, welche neben dem SM ähnlichen Higgs Teilchen existieren. Die Suche nach solchen schweren Higgs Bosonen wäre ein Tor zu neuer Physik jenseits des Standardmodells.

Vorgestellt wird eine Analyse auf der Suche nach schweren neutralen Higgs Resonanzen im Zerfallskanal  $A/H \rightarrow \tau\tau$ , indem beide Tau Leptonen in Hadronen zerfallen. Analysiert werden  $13.2 \text{ fb}^{-1}$  Daten aufgenommen mit dem ATLAS Detektor. Um die Sensitivität der Analyse zu verbessern werden Boosted Decision Trees (BDT) verwendet. Da kein Überschuss an Daten über der SM Erwartung beobachtet wurde, werden obere Limits im  $m_A\text{-tan}(\beta)$  Parameterraum im  $m_h^{\text{mod}+}$  Szenario des MSSM berechnet, die gewisse Parameterkonfigurationen ausschließen. Der BDT basierte Ansatz zeigt dabei eine deutliche Verbesserung der Limits im Vergleich zu den vorherigen Resultaten.





# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theoretical Foundations</b>	<b>3</b>
2.1	The Standard Model of particle physics . . . . .	3
2.2	Higgs mechanism in the Standard Model . . . . .	4
2.3	Supersymmetric extensions and the MSSM . . . . .	8
2.4	Physics with tau leptons . . . . .	12
<b>3</b>	<b>The ATLAS Experiment at the LHC</b>	<b>15</b>
3.1	The Large Hadron Collider . . . . .	15
3.2	The ATLAS experiment . . . . .	15
3.2.1	Coordinate system of the ATLAS detector . . . . .	16
3.2.2	Inner detector . . . . .	17
3.2.3	Calorimeters . . . . .	19
3.2.4	Muon spectrometer . . . . .	21
3.2.5	Event collecting and the trigger system . . . . .	22
<b>4</b>	<b>Monte Carlo generation and data samples</b>	<b>25</b>
<b>5</b>	<b>Analysis Selection</b>	<b>29</b>
5.1	Physics object reconstruction and identification . . . . .	29
5.2	Event selection . . . . .	31
5.3	Background estimation . . . . .	34
5.3.1	Data driven QCD background estimation . . . . .	34
5.3.2	Monte Carlo background estimation . . . . .	35
<b>6</b>	<b>Multivariate analysis</b>	<b>37</b>
6.1	Boosted Decision Trees in the TMVA framework . . . . .	37
6.2	Training setup for TMVA . . . . .	39
6.2.1	The classic splitting model . . . . .	43
6.2.2	The K-Fold cross validation model . . . . .	46
6.3	Training results and validation . . . . .	46
6.3.1	The classic splitting model . . . . .	47
6.3.2	The K-Fold cross validation model . . . . .	52
<b>7</b>	<b>Limit setting</b>	<b>57</b>
7.1	Theory of limit setting . . . . .	57
7.2	Limit setting results and uncertainties . . . . .	61
7.2.1	Limit results of the BDTs trained with the classic splitting model	62
7.2.2	Limit results of the BDTs trained with K-Fold model . . . . .	65
<b>8</b>	<b>Summary</b>	<b>67</b>

<b>A</b>	<b>BDT configuration and validation</b>	<b>69</b>
A.1	Validation of the training variables . . . . .	69
A.2	Signal region variable distributions . . . . .	71
A.3	Configuration of the BDTs . . . . .	74
A.4	Event weights of the b-associated signal production Monte Carlo . . .	78
A.5	Validation of the BDTs . . . . .	80
A.6	Signal region distributions of the BDTs . . . . .	85
<b>B</b>	<b>Limit</b>	<b>91</b>
B.1	Uncertainties . . . . .	91
B.2	Limit result tables . . . . .	95

# List of Tables

2.1.1	List of bosons in the SM . . . . .	4
2.1.2	List of fermions in the SM . . . . .	5
2.3.1	Parameter configuration for the $m_h^{\text{mod+}}$ scenario . . . . .	12
2.4.1	List of dominant decays of the tau lepton . . . . .	13
4.0.1	List of Monte Carlo generators used for signal and background simulation	27
5.2.1	Event yields after event selection . . . . .	33
6.2.1	List of variables used for BDT training . . . . .	43
A.3.1	Common configuration for BDT training . . . . .	75
A.3.2	Configuration for BDTs trained with classic splitting model . . . . .	76
A.3.3	Configuration for BDTs trained with K-Fold model . . . . .	77
B.1.1	Table of systematic uncertainties for B-TAG region . . . . .	93
B.1.2	Table of systematic uncertainties for B-VETO region . . . . .	94
B.2.1	Table of the upper limit values using BDTs trained on the classic splitting method . . . . .	96
B.2.2	Table of the upper limit values using BDTs trained on the K-Fold method . . . . .	97



# List of Figures

2.2.1	Higgs potential before and after symmetry breaking . . . . .	6
2.3.1	Particle spectrum of the MSSM . . . . .	9
2.3.2	Feynman diagrams for neutral MSSM Higgs production . . . . .	10
2.3.3	Higgs mass distributions with loop corections . . . . .	11
2.4.1	Decay of Higgs bosons into a pair of tau leptons . . . . .	12
2.4.2	Branching ratios for heavy neutral MSSM Higgs bosons . . . . .	14
3.1.1	Particle accelerators at CERN . . . . .	16
3.2.1	Overview of the ATLAS detector components . . . . .	17
3.2.2	Azimuthal angle dependence of the pseudorapidity . . . . .	18
3.2.3	Cut-away view of the inner detector . . . . .	18
3.2.4	Cut-away view of the calorimeter system . . . . .	20
3.2.5	Cut-away view of the muon spectrometer . . . . .	21
3.2.6	Schematic representation of the ATLAS trigger system . . . . .	23
4.0.1	Integrated luminosity recorded by ATLAS . . . . .	25
6.0.1	Total transverse mass distributions used in the previously published analysis . . . . .	37
6.1.1	Schematic representation of a decision tree . . . . .	38
6.2.1	Background and signal configuration of the BDTs . . . . .	40
6.2.2	Signal region distribution of the main separating training variables . .	42
6.2.3	Same-sign distribuntions of the main separating training variables . .	44
6.2.4	Shape comparrison between negative and positive MC weights for 1000 GeV signal mass . . . . .	45
6.3.1	Normalized response distributions of 500 GeV and 1000 GeV signal mass BDTs . . . . .	47
6.3.2	Signal correlation for 1000 GeV and 500 GeV signal mass BDT trained using classic splitting . . . . .	48
6.3.3	Background correlation of training variables . . . . .	49
6.3.4	ROC-Curve for BDTs trained on the classic splitting model . . . . .	49
6.3.5	BDT response in the same-sign validation region for the 500 and 1000 GeV BDTs using classic splitting . . . . .	50
6.3.6	BDT response in the signal region for the 500 and 1000 GeV BDTs using classic splitting . . . . .	51
6.3.7	Normalized response distributions of the 500 GeV signal mass BDTs trained with K-Fold method . . . . .	52
6.3.8	Normalized response distributions of the 1000 GeV signal mass BDTs train with K-Fold method . . . . .	53
6.3.9	BDT response in the same-sign validation region for the 500 and 1000 GeV BDTs using the K-Fold method . . . . .	54
6.3.10	BDT response in the signal region for the 500 and 1000 GeV BDTs using the K-Fold method . . . . .	55

7.1.1	Propability density function of the test statistic . . . . .	60
7.2.1	B-VETO and B-TAG expected limits for the classic splitting method	63
7.2.2	Combined expected and observed limits for the classic splitting method	64
7.2.3	B-VETO and B-TAG expected limits for the K-Fold method . . . . .	65
7.2.4	Combined expected and observed limits for the K-Fold method . . . . .	66
A.1.1	Same-sign distribuntions of variables used for training in B-TAG . . . .	69
A.1.2	Same-sign distribuntions of variables used for training in B-TAG . . . .	70
A.1.3	Same-sign distribuntions of variables used for training in B-VETO . . .	70
A.1.4	Same-sign distribuntions of variables used for training in B-VETO . . .	71
A.2.1	Signal region distribution of the training variables in B-TAG . . . . .	72
A.2.2	Signal region distribution of the training variables in B-TAG . . . . .	73
A.2.3	Signal region distribution of the training variables in B-VETO . . . . .	73
A.2.4	Signal region distribution of the training variables in B-VETO . . . . .	74
A.4.1	Shape comparrison between negative and positive MC weights for 1000 GeV signal mass in B-VETO . . . . .	78
A.4.2	Shape comparrison between negative and positive MC weights for 1000 GeV signal mass in B-TAG . . . . .	79
A.5.1	Same-sign validation of BDTs trained with classic splitting for 300, 350, 400 GeV signal mass . . . . .	80
A.5.2	Same-sign validation of BDTs trained with classic splitting for 600, 700, 800 GeV signal mass . . . . .	81
A.5.3	Same-sign validation of BDTs trained with classic splitting for 1200 GeV signal mass . . . . .	82
A.5.4	Same-sign validation of BDTs trained using K-Fold method for 300, 350, 400 GeV signal mass . . . . .	82
A.5.5	Same-sign validation of BDTs trained using K-Fold method for 600, 700, 800 GeV signal mass . . . . .	83
A.5.6	Same-sign validation of BDTs trained using K-Fold method for 1200 GeV signal mass . . . . .	84
A.6.1	Signal region response distributions of BDTs trained with classic split- ting for 300, 350, 400 GeV signal mass . . . . .	85
A.6.2	Signal region response distributions of BDTs trained with classic split- ting for 600, 700, 800 GeV signal mass . . . . .	86
A.6.3	Signal region response distributions of BDTs trained with classic split- ting for 1200 GeV signal mass . . . . .	87
A.6.4	Signal region response distributions of BDTs trained using K-Fold method for 300, 350, 400 GeV signal mass . . . . .	87
A.6.5	Signal region response distributions of BDTs trained using K-Fold method for 600, 700, 800 GeV signal mass . . . . .	88
A.6.6	Signal region response distributions of BDTs trained using K-Fold method for 1200 GeV signal mass . . . . .	89

# 1 Introduction

The discovery of the electron by J. J. Thomson in 1897 [1] marks the birth of fundamental particle physics, one of youngest fields of physics that tries to explain the mechanics of particles at very small scales. Since then this new field of physics experienced a rapid development. With the formulation of Quantum Electrodynamics (QED) by S. I. Tomonaga, J. Schwinger and R. Feynman in the late 40s, the first fully covariant formulation of a quantum theory of electromagnetism was available [2–4]. Based on the formalism used in QED, a theory of the weak nuclear interaction was developed by A. Salam, S. Weinberg and S. L. Glashow [5–7]. Shortly after P. Higgs, F. Englert and R. Brout independently postulated the existence of the Higgs mechanism that is responsible for generating the mass of the known fundamental particles and introduced a unified theory of electromagnetism and the weak force [8–10]. Almost at the same time M. Gell-Mann and G. Zweig constructed the theory of Quantum Chromo Dynamics that describes the interaction between quarks and their formation to hadrons. These theoretical models, most of which are formulated in the 1960s, are combined in a model that explains the fundamental nature of particles, the so called Standard Model (SM).

The SM has passed vigorous tests of its predictions using experiments with ever increasing complexity. Today the SM is one of the most accurately tested theories. All particles that are predicted by the SM were observed and measured in the last few decades by numerous experiments, the last one being the Higgs boson discovered in 2012 [11,12]. However, some of the predictions made by the SM seem to be unnatural and a quantum field theory that describes gravity and unify all known fundamental forces into one theory has yet to be found. One of the possible modifications to the SM that are able to resolve many of the issues are supersymmetric (SUSY) models that postulate an additional symmetry transformation between bosons and fermions. One extension of the SM, the Minimal Supersymmetric Standard Model (MSSM), predicts the existence of additional heavy Higgs bosons, which decay to a significant fraction into tau leptons.

To search these particles, enormous accelerators that collide particles are constructed with complex detectors built around the points where the particles collide. The biggest hadron collider build by man is the Large Hadron Collider where the ATLAS detector, a multi-purpose particle detector, is located.

None of the additional Higgs bosons has been found yet, so certain regions of the parameter space of the SUSY theories are excluded.

To enhance the sensitivity of the analysis in order to eventually find the signature of the theoretical Higgs bosons, sophisticated algorithms and data analysis techniques have to be used.

Outside of high energy physics fields multivariate analysis techniques (MVA) are developed to automatically evaluate vast amounts of data while achieving remarkable performance in identifying characteristics in the data to separate or categorize them. These algorithms rely on supervised learning to learn characteristic features. In the last decade modern MVA methods gained more attention in high energy physics. Especially Boosted Decision Trees (BDT) are found to be very useful to identify particles

by their signature left behind in the detector.

In this thesis a BDT based analysis searching for heavy neutral MSSM Higgs bosons decaying into tau leptons is presented. Here the multivariate method is applied to enhance the separation between hypothetical signal and background using only basic observables of the decay topologies.



## 2 Theoretical Foundations

### 2.1 The Standard Model of particle physics

The Standard Model (SM) of particle physics is currently the most successful theory that describes all known particles and their interactions through the fundamental forces, which are the electromagnetic force, the weak and the strong nuclear force [13]. Gravity is not included within the SM and a theory that describes all four fundamental forces is yet to be discovered. However, because the magnitude of gravity is negligible at subatomic scales, this does not impact the accuracy of predictions made by the SM.

The particle content of the SM consists of 12 fermionic particles with a spin of  $\frac{1}{2}$  and 5 bosons which have integer spins. All particles are thought to be elementary, that means they are point like and have no further substructure. The SM is a relativistic quantum field theory that can be described by a single Lagrangian density function with a local gauge symmetry of  $SU(3)_c \times SU(2)_L \times U(1)_Y$ . The  $SU(3)_c$  gauge group corresponds to the strong nuclear force with gluons as gauge bosons that carry color charge. On the other hand the symmetry group  $SU(2)_L \times U(1)_Y$  is the combined gauge symmetry of the electroweak interaction. This symmetry however is spontaneously broken by the Higgs mechanism [8–10] leaving behind  $U(1)_Q$  gauge group of the electromagnetic interaction with photons as gauge bosons and heavy  $W^\pm$  and  $Z$  bosons of the weak nuclear interactions, both were predicted by S. L. Glashow, S. Weinberg and A. Salam [5–7]. In addition to the gauge particles, a heavy Higgs boson with spin 0 is created by the electroweak symmetry breaking. Gauge bosons like photons and gluons are the transmitter particles of the fundamental interactions. The boson content of the SM is listed in Table 2.1.1.

Fermions, the elementary particles that make up matter, are separated in leptons and quarks grouped into three generations. Each generation contains a pair of "up" and "down" type fermions. Particles of different generations only differ in their invariant mass while all other properties, like charge and spin, stay the same. This creates a variety of different fermion particles which are parametrized by a quantum number called flavor. The fermionic particle content of the SM is listed in 2.1.2.

Quarks are elementary particles that have color charge and therefore take part in strong interactions which is described in the theory of Quantum Chromo Dynamics (QCD) [16, 17]. Because of the high magnitude of the strong nuclear force on small scales, free quarks quickly bond together to form color-neutral hadrons so that no free color charge can be observed. This effect is known as color confinement [14]. Since the interaction strength between color charged particles does not diminish fast enough when the particles are pulled apart, it is favorable to create quark-antiquark pairs from vacuum that quickly bond to the free charged particles, creating new color-neutral hadrons. This process of recombination is called hadronization or fragmentation.

Left handed quarks have a weak isospin of  $T = \frac{1}{2}$  and form doublets under weak interactions, whereas right handed particles do not interact with  $W$  bosons in charged flavor changing currents.

Leptons on the other hand have no color charge and thus are not subject to strong

Name		Spin	Mass [GeV]	Electric charge	$T_3$	Color charge
Photon	$\boldsymbol{\gamma}$	1	0	0	0	-
W-Boson	$\boldsymbol{W}^\pm$	1	80.4	$\pm 1$	$\pm 1$	-
Z-Boson	$\boldsymbol{Z}$	1	91.2	0	0	-
Gluon	$\boldsymbol{g}$	1	0	0	0	8 combinations
H-Boson	$\boldsymbol{h}_0$	0	125.4	0	$-\frac{1}{2}$	-

Table 2.1.1: List of the bosons that are predicted by the SM [15]. If no charge is assigned, the particle is a singlet under this interaction.

interactions. Under the weak interaction, left handed leptons also generate doublets of a charged lepton and a corresponding electric neutral neutrino. Because neutrinos only have a weak isospin charge, they only interact over weak interaction so that right handed neutrinos are effectively sterile particles [18]. The SM assumes that neutrinos are massless but the observation of neutrino oscillation suggest none zero masses [19, 20]. While it can not be excluded that the lightest neutrino has no mass, upper limits for the sum of the neutrino masses can be set to  $\sum m_\nu \leq 0.3 \text{ eV}$  [21] through cosmic observations and thus neutrinos are several orders of magnitudes lighter than their charged counterparts. The fermionic content of the Standard Model and their properties are summarized in Table 2.1.2.

## 2.2 Higgs mechanism in the Standard Model

The electroweak sector is described by an  $SU(2)_L \times U(1)_Y$  gauge theory, where the field  $W_\mu^a$  corresponds to  $SU(2)_L$  with the generators  $T_a$  ( $a = 1, 2, 3$ ) that are proportional to the  $2 \times 2$  Pauli matrices [13]:

$$T_a = \frac{1}{2}\sigma_a \text{ with } \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and  $B_\mu$  corresponds to  $U(1)_Y$  with the hypercharge  $Y$  as generator of the group,

$$Y = Q - T_3.$$

The kinematic part of Lagrangian function that describes the electroweak sector is then defined as [22]:

$$\mathcal{L}_{\text{ew}} = -\frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}$$

with the field strength tensor:

$$\begin{aligned} W_{\mu\nu}^a &= \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g_W \epsilon^{abc} W_\mu^b W_\nu^c, \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \end{aligned}$$

where  $\epsilon^{abc}$  is the 3-dimensional Levi-Civita symbol. The Lagrangian function contains no mass terms for the vector fields and thus the particles are massless. However,

Quarks					
Name		Mass	Electric charge	$T_3$	Color charge
1. Generation	<b><i>u</i></b>	2.3 MeV	$+\frac{2}{3}$	$+\frac{1}{2}$	r,g,b
	<b><i>d</i></b>	4.8 GeV	$-\frac{1}{2}$	$-\frac{1}{2}$	
2. Generation	<b><i>c</i></b>	1.2 GeV	$+\frac{2}{3}$	$+\frac{1}{2}$	r,g,b
	<b><i>s</i></b>	95 MeV	$-\frac{1}{2}$	$-\frac{1}{2}$	
3. Generation	<b><i>t</i></b>	173.1 GeV	$+\frac{2}{3}$	$+\frac{1}{2}$	r,g,b
	<b><i>b</i></b>	4.2 GeV	$-\frac{1}{2}$	$-\frac{1}{2}$	

Leptons					
Name		Mass	Electric charge	$T_3$	Color charge
1. Generation	<b><i>e</i></b>	511 keV	$+\frac{2}{3}$	$+\frac{1}{2}$	-
	<b><math>\nu_e</math></b>	$\leq 0.3 \text{ eV}^*$	$-\frac{1}{2}$	$-\frac{1}{2}$	
2. Generation	<b><math>\mu</math></b>	105.7 MeV	$+\frac{2}{3}$	$+\frac{1}{2}$	-
	<b><math>\nu_\mu</math></b>	$\leq 0.3 \text{ eV}^*$	$-\frac{1}{2}$	$-\frac{1}{2}$	
3. Generation	<b><math>\tau</math></b>	1.78 GeV	$+\frac{2}{3}$	$+\frac{1}{2}$	-
	<b><math>\nu_\tau</math></b>	$\leq 0.3 \text{ eV}^*$	$-\frac{1}{2}$	$-\frac{1}{2}$	

Table 2.1.2: List of fermions predicted by the SM [15]. \* Upper limit is set for the sum of all neutrino masses.

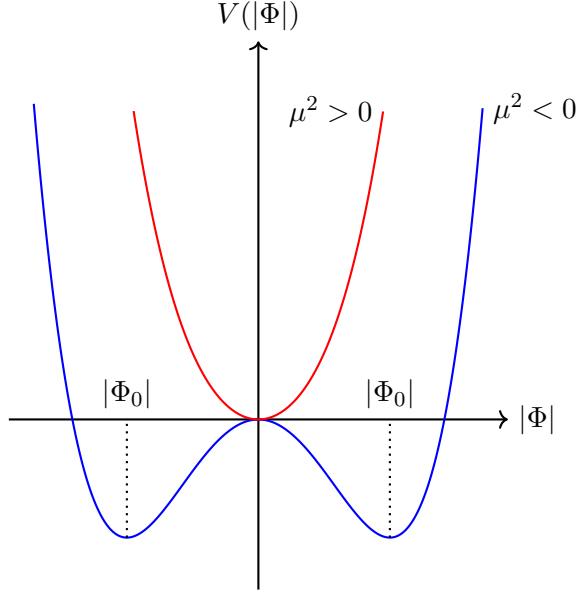


Figure 2.2.1: Form of the Higgs potential before (red) and after (blue) spontaneous electroweak symmetry breaking [22].

measurements have shown that the electroweak  $Z$  and  $W$  gauge bosons as well as the quarks and charged leptons have quite substantial masses [15]. Simple mass terms of the structure  $m\phi\phi^*$  would break the gauge symmetry of the Lagrangian function and are therefore not possible. To let  $W$  and  $Z$  bosons acquire mass while conserving gauge invariance, one has to add an additional complex scalar Higgs field isodoublet  $\Phi$  with a potential  $V(\Phi)$  that enables symmetry breaking:

$$\Phi = \begin{pmatrix} \Phi_+ \\ \Phi_0 \end{pmatrix}, \quad (2.1)$$

with hypercharge  $Y_\Phi = +1$ , which results in a positively charged field component  $\Phi_+$  and a neutral one  $\Phi_0$ . The corresponding Lagrange density function is defined by:

$$\mathcal{L}_\Phi = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi), \quad (2.2)$$

where  $V(\Phi)$  is an appropriately chosen Higgs potential:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \quad (2.3)$$

This Lagrangian function is invariant under  $SU(2)_L \times U(1)_Y$ . The vacuum state however can break this symmetry. As shown in Figure 2.2.1, the minimum of the potential depends on the sign of  $\mu^2$ . If  $\mu^2 > 0$ , only one minimum  $\Phi_+ = \Phi_0 = 0$  exists, but for negative  $\mu^2 < 0$  the minimum is found to be a circle in the  $|\Phi|$  plane. Because it is not expected that the vacuum state of the Higgs field is electrically charged,  $\Phi_+$  can be set to zero. This leads to a vacuum expectation value for the neutral Higgs field of

$$\begin{aligned} \Phi_0 &= \frac{v}{\sqrt{2}}, \\ v &= \sqrt{-\frac{\mu^2}{\lambda}}. \end{aligned} \quad (2.4)$$

The Higgs field can be expanded around the vacuum expectation value by a general first order expansion resulting in

$$\Phi^0 = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + h(x)) \end{pmatrix}. \quad (2.5)$$

If (2.5) is put in Lagrangian density (2.2) with the gauge covariant derivation

$$D_\mu = (\partial_\mu - ig_w T^a W_\mu^a - ig_Y Y B_\mu), \quad (2.6)$$

the Lagrange function will become

$$\begin{aligned} \mathcal{L}_{\Phi^0} &= (\partial_\mu h(x))^2 + \frac{v^2 g_W^2}{4} W_\mu^+ W_\mu^- \left(1 + \frac{h(x)}{v^2}\right) \\ &+ \frac{v^2}{4} (g_Y^2 + g_W^2) Z_\mu Z^\mu \left(1 + \frac{h(x)}{v^2}\right) + V(\Phi^0), \end{aligned} \quad (2.7)$$

where  $Z_\mu$  and  $W_\mu^\pm$  fields are mixtures of the  $B_\mu$  and  $W_\mu^a$  that diagonalize the mass matrix:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \\ Z_\mu &= \frac{g_W W_\mu^3 - g_Y B_\mu}{\sqrt{g_W^2 + g_Y^2}}. \end{aligned} \quad (2.8)$$

After the Higgs field acquires its vacuum state in the case  $\mu^2 < 0$ , the Lagrangian is no longer invariant under electroweak gauge transformations, instead mass terms for the  $Z_\mu$  and  $W_\mu^\pm$  fields are generated. Also the scalar field  $h$  becomes massive. However, the  $U(1)_Q$  gauge symmetry, the symmetry of electromagnetism with the electric charge  $Q$  as generator, remains. The corresponding gauge boson  $A_\mu$ , the photon, is left untouched by electroweak symmetry breaking and therefore remains massless. The Higgs field can also couple with fermions over Yukawa couplings to generate their mass:

$$\begin{aligned} \mathcal{L}_{\text{Yukawa}} &= -\lambda_e \bar{L} \Phi e_R - \lambda_d \bar{Q} \Phi d_R - \lambda_u \bar{Q} \tilde{\Phi} u_R + h.c. \\ &= -\frac{1}{\sqrt{2}} \lambda_e (\bar{\nu}_e, \bar{e}_L) \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} e_R + \dots \\ &= -\frac{1}{\sqrt{2}} \lambda_e (v + h(x)) \bar{e}_L e_R + \dots \end{aligned} \quad (2.9)$$

with  $L$  and  $Q$  the isodoublets of left handed and  $e_R, d_R, u_R$  for the right handed singlet fields of leptons and quarks. To generate the mass of the up-type quarks, a charge conjugated Higgs doublet  $\tilde{\Phi} = i\tau_2 \Phi^*$  is required, introducing terms with complex conjugated fields in the potential. The Yukawa coupling constants  $\lambda_l$  for all lepton species are free parameters of the theory and are connected with the invariant lepton mass:

$$m_l = \frac{v \lambda_l}{\sqrt{2}}. \quad (2.10)$$

The massive scalar boson discovered in 2012 by the ATLAS and CMS experiments is associated with the Higgs Boson predicted by the SM, as many properties are in alignment with the expected theoretical values [11, 12, 23, 24].

## 2.3 Supersymmetric extensions and the MSSM

Although predictions made by the SM are in very good agreement with results of experiments, there are still questions within particle- and astrophysics that are left unanswered. For one, a theory that unifies the electroweak and strong nuclear force, a so called Grand Unified Theory, is yet to be found [25]. Also not resolved in the SM is the so called fine-tuning problem. Because the magnitude of gravity is very small at subatomic scales, gravitational effects would play a role at the Planck energy scale which is about 16 orders of magnitudes higher than the electroweak scale. A quantum field theory that includes gravity would add very large loop corrections to the Higgs mass. These loop corrections are quadratically divergent with the highest energy scale of the theory. Normalizing the mass of the Higgs boson to its measured value, which is near the electroweak scale, would require an unnatural fine tuning of free parameters [22]. Another problem that can not be satisfyingly understood within the SM is the existence of dark matter. Despite that dark matter does not seem to interact with photons, it can be measured via its gravitational impact on the movement of galaxies and by measuring gravitational lensing effects [26,27]. Currently there is no well-defined answer in the SM.

However, there are extensions to the Standard Model that are able to answer these questions. Very popular models are supersymmetric (SUSY) theories where the existence of an additional symmetry transformation is postulated that transforms bosons into corresponding fermions and vice versa, thus expanding the particle content of the model by the supersymmetric partners of the already known particles [28]. The model with the smallest particle content that is still compatible with experimental results is the Minimal Supersymmetric Standard Model (MSSM). It predicts the existence of a superpartner for each known particle in the SM. A SUSY transformation in the MSSM would transform gauge bosons and Higgs bosons to half integer spin gauginos and Higgsinos, whereas leptons and quarks would become so called sleptons and squarks with spin zero. The entire particle spectrum predicted by the MSSM is summarized in Figure 2.3.1.

The MSSM also requires the existence of an additional Higgs doublet [29], so that the Higgs sector is described by the two doublets  $H_1$  with vacuum expectation value  $v_1$  and  $H_2$  with  $v_2$  which have the structure:

$$H_1 = \begin{pmatrix} H_1^0 \\ H_1^- \end{pmatrix} \text{ with } Y = -1 \text{ and } H_2 = \begin{pmatrix} H_2^+ \\ H_2^0 \end{pmatrix} \text{ with } Y = +1. \quad (2.11)$$

At least two Higgs doublets of opposite hypercharge are required to generate the masses for fermions via Yukawa couplings. Terms proportional to  $\tilde{\Phi} = i\tau_2 \Phi^*$  that are used in the SM to generate the mass of up-type fermions are not allowed within the MSSM since the superpotential should not contain the complex conjugate of the fields. Otherwise the superpotential would not be invariant under SUSY transformation. The Higgs potential is then a function of both doublets

$$V_H = (\mu^2 + m_{H_1}) |H_1|^2 + (\mu^2 + m_{H_2}) |H_2|^2 - \mu B \epsilon_{ij} (H_1^i H_2^j + h.c.) + \frac{g_2^2 + g_1^2}{8} (|H_1|^2 - |H_2|^2)^2 + \frac{1}{2} g_2^2 |H_1^\dagger H_1|^2. \quad (2.12)$$

The vacuum expectation values,  $v_1$  for  $H_1^0$  and  $v_2$  for  $H_2^0$ , are acquired after spontaneous symmetry breaking analog to the procedure outlined in chapter 2.2. Here also  $SU(2)_L \times U(1)_Y$  gauge symmetry breaks and  $U(1)_Q$  symmetry remains. Other

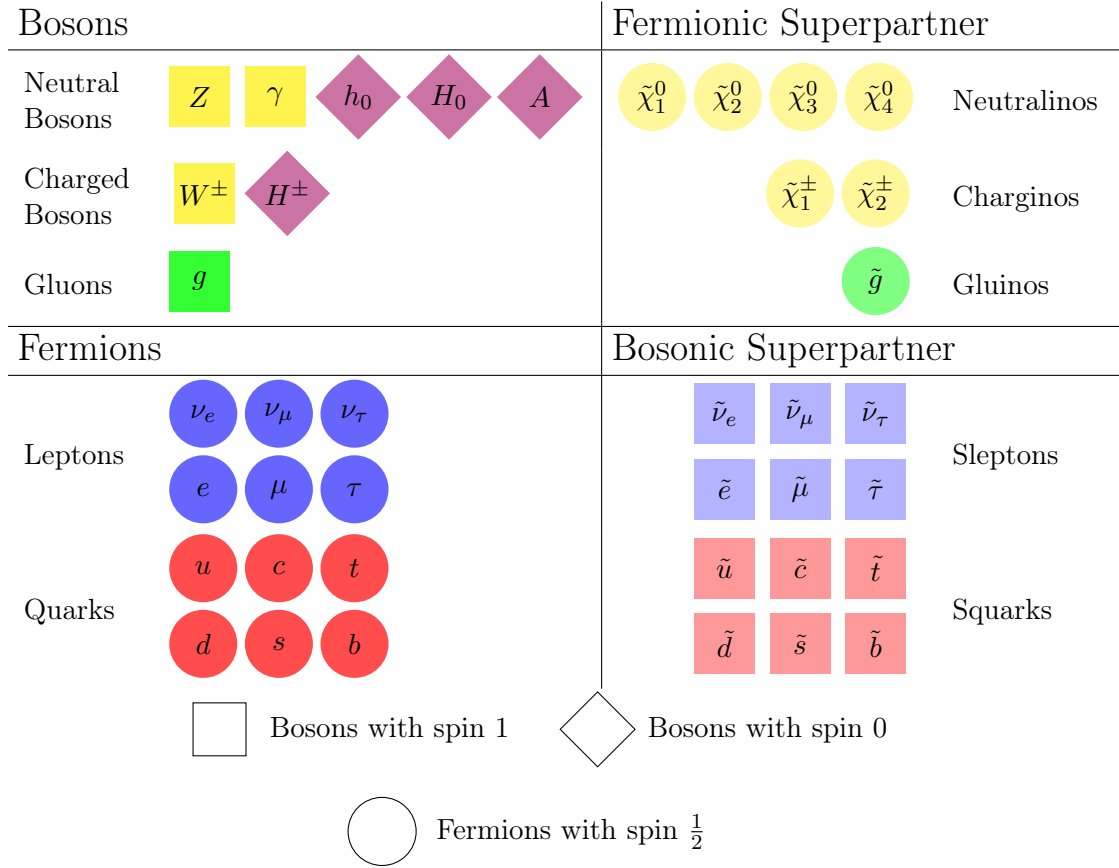


Figure 2.3.1: Particle spectrum of the MSSM [28]. Each SM like particle has a supersymmetric partner. In addition, two Higgs doublets are predicted leading to the existence of two charged and three neutral Higgs bosons. Particles in their mass eigenstate are listed.

than in the Standard Model the two Higgs doublets have 8 degrees of freedom where 3 are gauged away by the  $W^\pm$  and  $Z$  boson. This implies the existence of 5 Higgs bosons. One light and one heavy CP-even boson  $h_0$  and  $H_0$ , a CP-odd boson  $A$  and two charged bosons  $H^\pm$ . The entire Higgs sector is described at tree level by two parameters which are often chosen to be the mass of the CP-odd boson  $m_A$  and the ratio of the vacuum expectation values  $\tan(\beta)$ :

$$\tan(\beta) = \frac{v_2}{v_1}. \quad (2.13)$$

The Higgs fields of  $H_1$  and  $H_2$  can interact via Yukawa couplings with fermions where  $H_1$  couples to up-type fermions with coupling strength  $\lambda_u$  and  $H_2$  to down-type fermions with  $\lambda_d$ :

$$\lambda_u = \sqrt{2} \frac{m_u}{v_1}, \quad \lambda_d = \sqrt{2} \frac{m_d}{v_2}, \quad (2.14)$$

so heavier fermions couple stronger to the Higgs field. By constructing the ratio of coupling constants from (2.14), one finds:

$$\frac{\lambda_d}{\lambda_u} = \frac{m_d}{m_u} \tan(\beta). \quad (2.15)$$

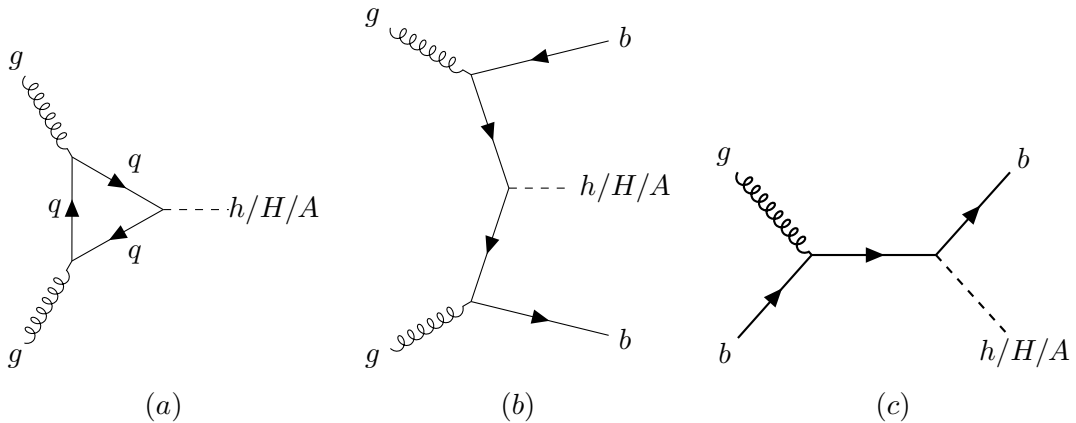


Figure 2.3.2: Feynman diagrams for gluon fusion production (a) and b-associated production in four (b) and five (c) flavor scheme for neutral MSSM Higgs bosons at the LHC [29].

This shows, that for increasing  $\tan(\beta)$  the coupling to down type fermions with  $T_3 = -\frac{1}{2}$  like bottom quarks and tau leptons are strongly enhanced, whereas couplings to up type fermions become suppressed.

Currently no SUSY particle has been directly observed yet, but because most supersymmetric models like the MSSM require one additional Higgs doublet, it is therefore a viable strategy to look for new heavy Higgs bosons.

At hadron colliders like the LHC, Higgs bosons are produced by proton-proton collisions. There are two major production types that involve QCD processes, namely gluon fusion production, where two gluons fuse together over a quark loop emitting a Higgs boson, and the b-associated production with additional bottom quarks in the final state as shown in figure 2.3.2.

With the two free parameters of the theory, the mass of the additional Higgs bosons can be calculated at tree level:

$$m_{H^\pm}^2 = m_A^2 + m_{W^\pm}^2, \quad (2.16)$$

$$m_{h,H}^2 = \frac{1}{2} \left( m_A^2 + m_Z^2 \mp \sqrt{(m_A^2 + m_Z^2)^2 - 4m_A^2 m_Z^2 \cos^2(2\beta)} \right).$$

From the last equation in (2.16) the mass of the lightest Higgs boson can be constrained by an upper limit:

$$m_h \leq \min(m_A, m_Z) |\cos(2\beta)| \leq m_Z. \quad (2.17)$$

If the light  $h$  boson is set to be the Higgs boson discovered by the ATLAS and CMS experiments with mass  $m_h = 125.09 \pm 0.24$  GeV [30], equation (2.17) leads to a contradiction because the mass of the lightest MSSM Higgs boson would be greater than the  $Z$  mass. This contradiction can be solved through loop corrections to the Higgs self-energy that allows a significantly higher lowest Higgs mass within additional tuning parameters of the higher order perturbation calculations.

Driving loop corrections are caused by top/stop loops and can be calculated to:

$$\Delta m_h^2 \sim \frac{3}{2\pi^2} \frac{\bar{m}_t^4}{v^2 \sin^2 \beta} \left( \log \frac{M_S^2}{\bar{m}_t^2} + \frac{X_t^2}{M_S^2} \left[ 1 - \frac{X_t^2}{12M_S^2} \right] \right), \quad (2.18)$$



with the vacuum expectation value  $v = \sqrt{v_1^2 + v_2^2}$ , running top mass  $\overline{m}_t^2$  and the geometric average of the stop mass  $M_S = \sqrt{\overline{m}_{\tilde{t}_1} \cdot \overline{m}_{\tilde{t}_2}}$ . The variable  $X_t$  denotes the mixing of the stop mass eigenstates  $\tilde{t}_1$  and  $\tilde{t}_2$  and is defined:

$$X_t = A_t - \frac{\mu}{\tan(\beta)}, \quad (2.19)$$

with  $A_t$  as the trilinear stop-Higgs coupling and  $\mu$  the Higgsino mass parameter. To raise the light Higgs mass above the  $Z$  mass to the measured value, either the stop masses have to be significantly higher as their contribution grows logarithmically or stop mixing parameter has to be adjusted. Former is regarded as "unnatural" because the SUSY mass scale is expected to be not far off from the electroweak symmetry breaking scale, so the light Higgs mass can only be corrected using mainly stop mixing. The dependency of the h boson mass as a function of the stop mixing parameter is shown in Figure 2.3.3.

As the light Higgs mass can be tuned by two possible  $X_t$  inside the local maxima, this leads to the definition of two possible benchmark scenarios, one with positive and one with negative sign of the mixing parameter.

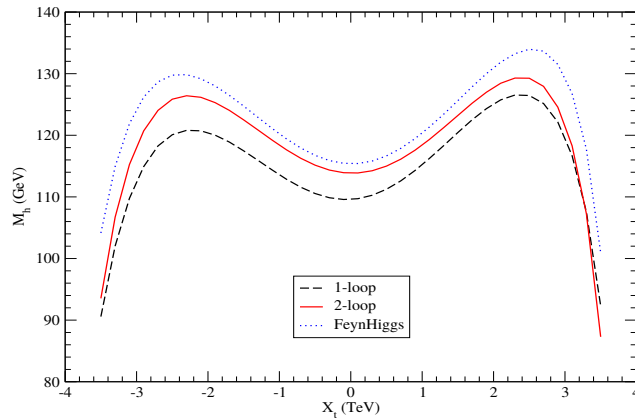


Figure 2.3.3: Higgs mass as function of  $X_t$  with loop corrections calculated by SuSpect (dashed and red line) and 2-loop corrections using FeynHiggs (dotted line) from Ref. [29]. The loop corrections to the Higgs self-energy allows a significantly higher lighter Higgs mass.

The scenario with positive signed  $X_t$ , called  $m_h^{\text{mod}+}$  and the scenario with negative signed stop mixing  $m_h^{\text{mod}-}$ . Both scenarios yield similar results, but the  $m_h^{\text{mod}+}$  scenario describes better the  $(g-2)_\mu$  measurement of the muon while  $m_h^{\text{mod}-}$  is in better conformity with measurements of  $BR(b \rightarrow s\gamma)$ . Furthermore, the definition of a  $m_h^{\text{max}}$  scenario is possible where loop corrections maximize the mass of the lightest Higgs. However, depending on the benchmark configuration the  $m_h^{\text{max}}$  scenario allows Higgs masses to be over 130 GeV which is not in agreement of the measurement.

In this thesis, the  $m_h^{\text{mod}+}$  scenario is used to compare the sensitivity of different analysis results. The setup for the parameters of the scenario is described in Ref [31] and listed in Table 2.3.1.

Variable	Description
$M_{\text{SUSY}} = 1000 \text{ GeV}$	SUSY mass scale
$\mu = 200 \text{ GeV}$	Higgsino mass parameter
$M_2 = 200 \text{ GeV}$	SU(2) gaugino mass parameter
$M_{\tilde{t}_3} = 1000 \text{ GeV}$	Stau sector mass parameter
$M_{\tilde{l}_{1,2}} = 500 \text{ GeV}$	Mass parameter of the 1. and 2. slepton generation
$M_{\tilde{q}_{1,2}} = 1500 \text{ GeV}$	Mass parameter of the 1. and 2. squark generation
$m_t = 173.2 \text{ GeV}$	Top mass
$m_{\tilde{g}} = 1500 \text{ GeV}$	Gluino mass
$X_t^{\text{OS}} = 1.5 \times M_{\text{SUSY}}$	Stop mixing in OS scheme
$X_t^{\overline{\text{MS}}} = 1.6 \times M_{\text{SUSY}}$	Stop mixing in $\overline{\text{MS}}$ scheme
$A_f = 0$	Trilinear Higgs to fermion couplings ( $f = u, d, c, s, \mu, e$ )
$A_b = A_\tau = A_t$	Trilinear Higgs to sbottom, stau and stop coupling

Table 2.3.1: Parameter configuration for the  $m_h^{\text{mod}+}$  scenario of the MSSM [31].

## 2.4 Physics with tau leptons

The tau lepton is the heaviest known lepton having an invariant mass of  $1776.82 \pm 0.16 \text{ MeV}$  [15] and thus being almost 17 times heavier than the muon. Due to the resulting large phase space, the mean lifetime of a tau lepton is very small ( $290.3 \times 10^{-15} \text{ s} \pm 0.5 \times 10^{-15} \text{ s}$  [15]). Because it decays so fast, the tau lepton can not be detected directly, thus only the decay products are visible and used to analyze scattering processes where tau leptons are involved in the final states. Tau leptons decay almost 58 % of the time to light hadrons like pions and only to 35 % in lighter charged leptons and lepton neutrinos [15]. Table 2.4.1 summarizes the dominant decay channels of the tau decay (Figure 2.4.1).

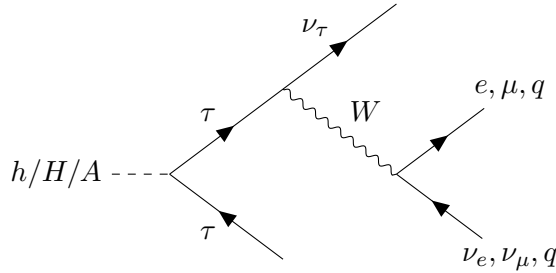


Figure 2.4.1: Decay of Higgs bosons into a pair of tau leptons. The tau lepton decays further into leptons and hadrons.

In all decay channels at least one tau neutrino is present in the final state. Because neutrinos can only interact by the weak interaction, most of the time they leave the detector undetected. Although there is a lepton involved when a tau decays into pions, it is still convenient to call this decay the hadronic decay channel since only the hadrons can be detected. The visible hadronic decay products are labeled  $\tau_{\text{had, vis}}$ . In the hadronic decay a categorization in charge multiplicity is made. Since electric charge is conserved, a tau lepton can decay in odd amounts of charged pions. However, tau leptons mostly decay hadronically with charge multiplicity of 1 or 3 because

### Leptonic decay

Decay products	Branching fraction $\frac{\Gamma_i}{\Gamma}$
$\tau^+ \rightarrow \mu^+ + \nu_\mu + \nu_\tau$	17.4 %
$\tau^+ \rightarrow e^+ + \nu_e + \nu_\tau$	17.8 %
Total	35.2 %

### Hadronic decay

1 charged hadron	
$\tau^+ \rightarrow \pi^+ + \nu_\tau$	10.8 %
$\tau^+ \rightarrow \pi^+ + \pi^0 + \nu_\tau$	25.5 %
$\tau^+ \rightarrow \pi^+ + 2 \times \pi^0 + \nu_\tau$	9.3 %
$\tau^+ \rightarrow \pi^+ + 3 \times \pi^0 + \nu_\tau$	1.1 %
Total	46.7 %
3 charged hadrons	
$\tau^+ \rightarrow 2 \times \pi^+ + \pi^- + \nu_\tau$	9.0 %
$\tau^+ \rightarrow 2 \times \pi^+ + \pi^- + \pi^0 + \nu_\tau$	2.7 %
Total	11.7 %

Table 2.4.1: List of dominant decays of the tau lepton [15].

decays involving more charged pions are increasingly kinematically suppressed. In the leptonic decay one additional lepton neutrino is present in the final state. Since neutrinos are not detected and reconstructed, this loss of decay information subsequently leads to potentially worse tau reconstruction resolutions.

As shown in (2.15) the coupling of Higgs bosons in the MSSM is strongly enhanced for high  $\tan(\beta)$ . This results in high branching ratios for Higgs bosons decaying into tau leptons as shown in Figure 2.4.2, while the coupling to top quarks and other up-type fermions gets smaller. In this thesis the full hadronic decay of neutral Higgs bosons is analyzed where both tau leptons decay hadronically:

$$h/H/A \rightarrow \tau^+ \tau^- . \quad (2.20)$$

This decay channel is particularly sensitive for the search of MSSM Higgs bosons at high values of  $\tan(\beta)$ .

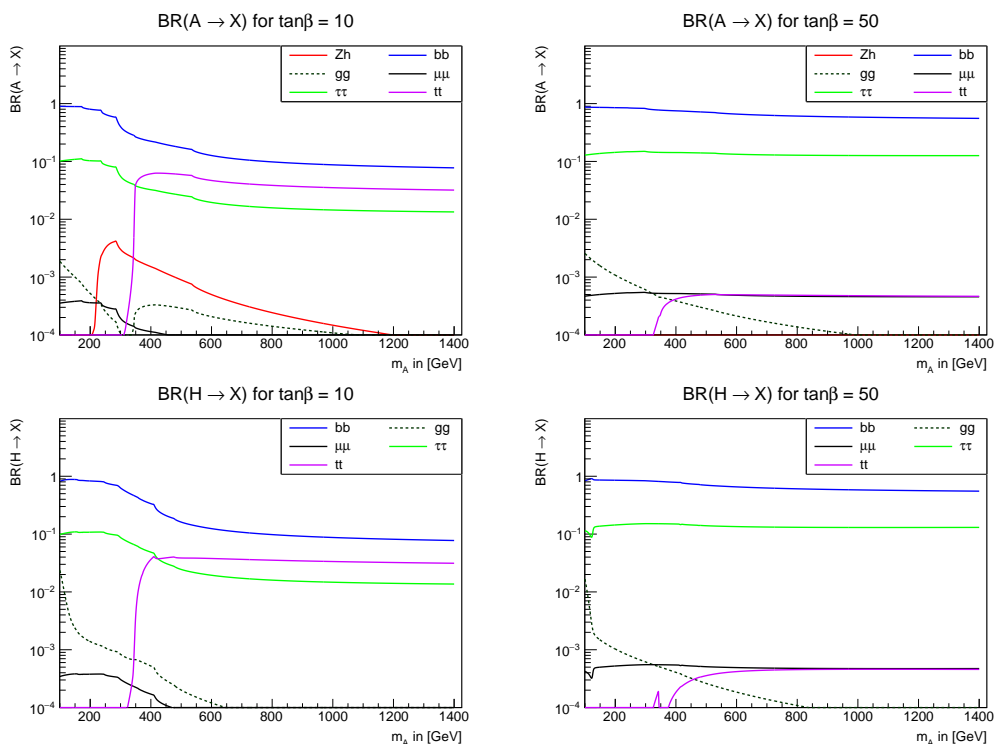


Figure 2.4.2: Branching ratios for heavy neutral MSSM Higgs bosons in dependence of  $m_A$  for  $\tan \beta = 10$  and  $50$ . The branching ratios are calculated using FeynHiggs and SusHi [31–38]

# 3 The ATLAS Experiment at the LHC

## 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [39] is the world largest particle collider, located at the CERN (European Organization for Nuclear Research) laboratories in Geneva, Switzerland. The LHC is a synchrotron for accelerating proton or lead ion beams in opposite direction. The ring like structure of the beam pipes measures a circumference of 27 km and lies in the tunnel of the precursor experiment LEP (Large Electron Positron collider) [40] up to 175 m beneath the surface. To keep the charged particle beams on their circular path, the beam pipe includes 1232 superconducting dipole magnets, which have to be cooled down to 1.9 K using liquid helium. The dipole magnets create a magnetic field of up to 8.3 T, bending the flight path of the particles. Additional 392 quadrupole magnets and several higher order multi-pole magnets are used to focus and counteract dispersion of the particle beams. The LHC is capable of reaching a beam energy in proton-proton mode of 6.5 TeV resulting into a center of mass energy of currently  $\sqrt{s} = 13$  TeV. To achieve the high particle energy, protons are pre-accelerated in several smaller accelerators. The first is the linear particle accelerator LINAC 2 which accelerates protons from hydrogen gas stripped of their electrons, up to an energy of 50 MeV. The protons are then fed into the Proton Synchrotron Booster (PSB) raising the energy to 1.4 GeV. The Proton Synchrotron (PS) accelerates the particles further up to 26 GeV before they are entering the last stage, the Super Proton Synchrotron (SPS), which increases the energy to 450 GeV and injects the proton beams into the LHC. The entire acceleration chain is visualized in Figure 3.1.1. The proton in the beams are packed in bunches. These bunches cross and collide with each other at 4 interaction points, where the 4 big detectors ATLAS [42], CMS [43], ALICE [44], LHCb [45] and the smaller ones LHCf [46], TOTEM [47] and MoEDAL [48] are located. After the first Long Shutdown from February 2013 to April 2015 (LS1), the number of bunches was steadily increased to a maximum of 2076 bunches crossing with a minimal bunch space of 25 ns [49]. In June 2016 the LHC was capable to reach the designed luminosity of  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  [50].

## 3.2 The ATLAS experiment

The ATLAS experiment (A Toroidal LHC ApparatuS) is one of the four bigger experiments placed at the interaction points of the LHC. It is a multi-purpose particle detector, measuring 25 m in height and 44 m in length [42]. Its main purpose is to measure particles created by high energy bunch crossings, recording and reconstructing them for each scattering event. The ATLAS has to fulfill high standards as it is used as a tool for precision measurements of physical quantities predicted by the SM, especially the properties of the Higgs boson in which discovery in 2012 the ATLAS experiment was involved, as well as for looking for new physics beyond the SM like supersymmetry [51].

To accomplish the ambitious goals, the detector is itself made out of several sub-

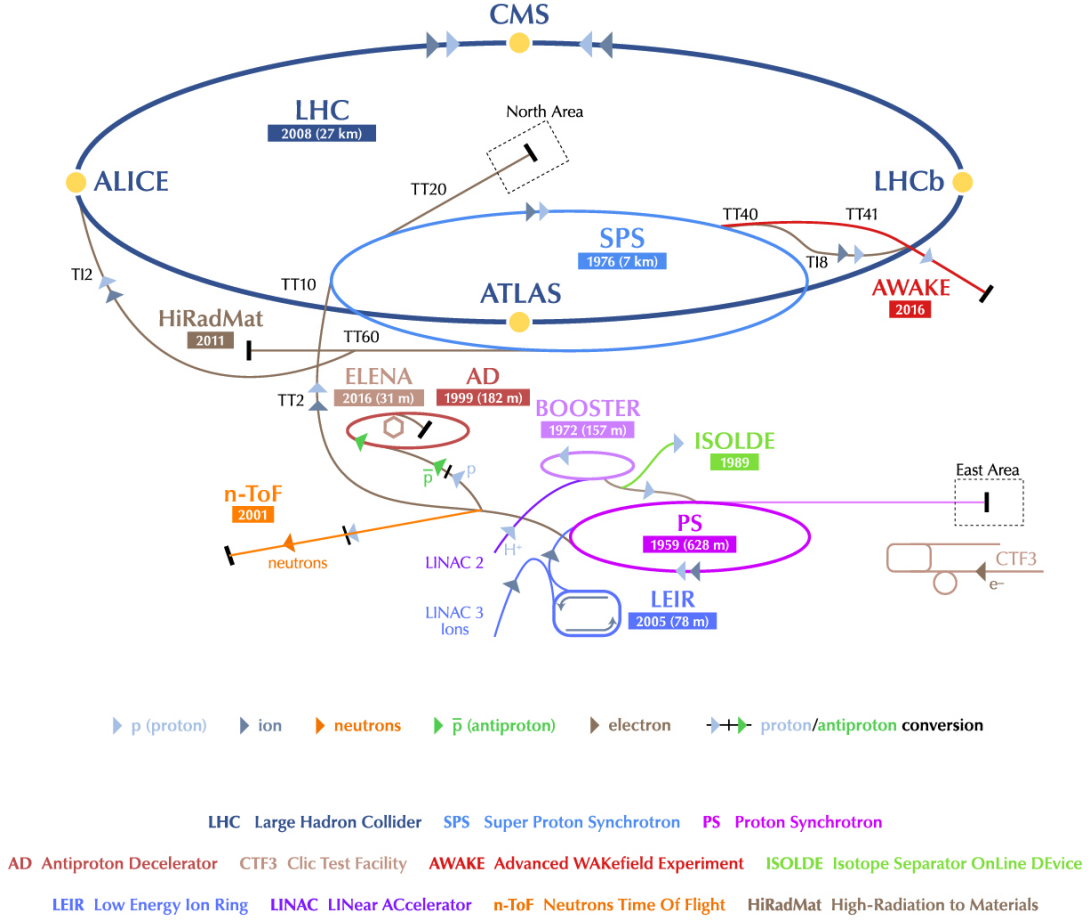


Figure 3.1.1: Schematic representation of the particle accelerators [41]. The LINAC2, PSB, PS and SPS are involved in the proton acceleration chain of the LHC.

detectors concentrically layered around the interaction point of the beam pipe. The cylindrical shape allows to cover most of the spatial directions. An overview of all detector components is given in Figure 3.2.1.

### 3.2.1 Coordinate system of the ATLAS detector

To parametrize the parts of the detector and subsequently the particle tracks reconstructed from energy deposits within the detector, a right handed Cartesian coordinate system can be introduced with its origin at the point of interaction, the longitudinal z-axis defined along the beam pipe. The x-y-plane is further referred as the transversal plain as it is orthogonal to the flight direction of the protons in the beam pipe. The transversal component of a momentum vector is then defined as:

$$p_T = \sqrt{p_x^2 + p_y^2}. \quad (3.1)$$

The transverse plane is of particular interest since momentum in this plane is conserved so that the sum of all transverse momenta should be zero. However, some particles can escape the detector, like neutrinos which only interact weakly and thus mostly leave the detector without interacting with the material at all. Exploiting transversal

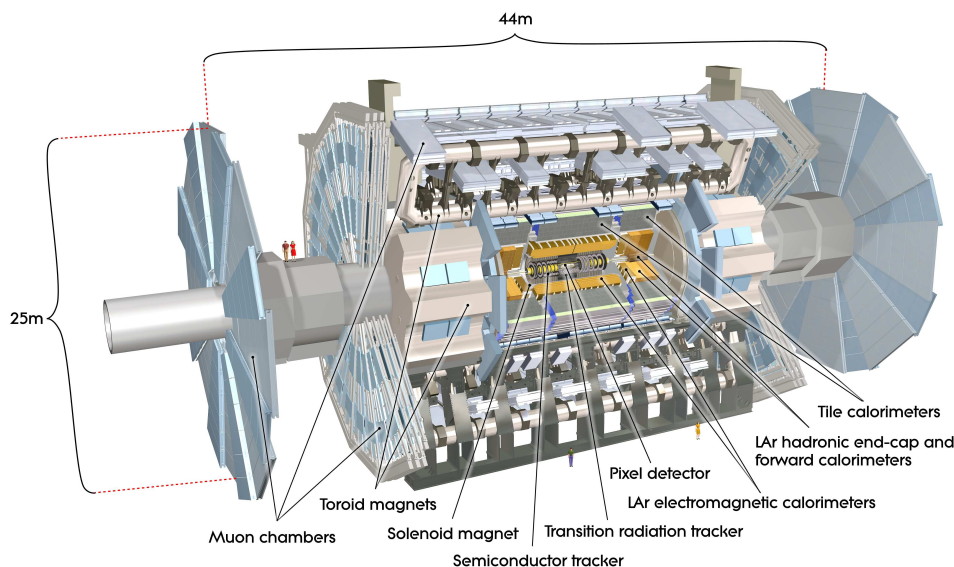


Figure 3.2.1: Schematic overview of the ATLAS detector components [42].

momentum conservation, the sum of the energy carried away by these particles in the transversal plane can be calculated as the negative vectorial sum of the visible transverse momenta [52]. The derived quantity is called missing transverse energy  $E_{T, \text{mis}}$ :

$$E_{T, \text{mis}} = \left| - \sum \vec{p}_T \right|, \quad (3.2)$$

Also spherical coordinates  $(r, \phi, \theta)$  are used with the radius  $r = \sqrt{x^2 + y^2 + z^2}$ , the azimuthal angle  $\phi$  defined by  $\tan \phi = \frac{y}{x}$  and the polar angle defined by  $\cos \theta = \frac{z}{r}$  measuring the spacial angle to the beam pipe. In hadron collider physics, instead of the polar angle, usually the pseudorapidity  $\eta$  is used:

$$\eta = - \ln \tan \left( \frac{\theta}{2} \right). \quad (3.3)$$

The advantage of using the pseudorapidity instead of the polar angle is, that for particles, which masses are negligible compared to their energy, the pseudorapidity is invariant under Lorentz transformations along the beam axis. This allows to take the unknown longitudinal momentum of the partons inside the proton into account. Vectors perpendicular to the  $z$ -axis have a pseudorapidity of  $\eta = 0$  while for vectors parallel to the beam pipe  $\eta$  approaches  $\pm\infty$ . The azimuthal angle dependency of  $\eta$  is illustrated in Figure 3.2.2. Using the pseudorapidity the solid angle element  $R$  is defined:

$$R = \sqrt{\phi^2 + \eta^2}, \quad (3.4)$$

so that solid angle difference between two vectors is calculated as:

$$\Delta R = \sqrt{(\phi_1 - \phi_2)^2 + (\eta_1 - \eta_2)^2}. \quad (3.5)$$

### 3.2.2 Inner detector

The inner detector (ID) of the ATLAS experiment is composed of three different sub-detectors, the pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT) layered around the beam pipe. Its task is to precisely

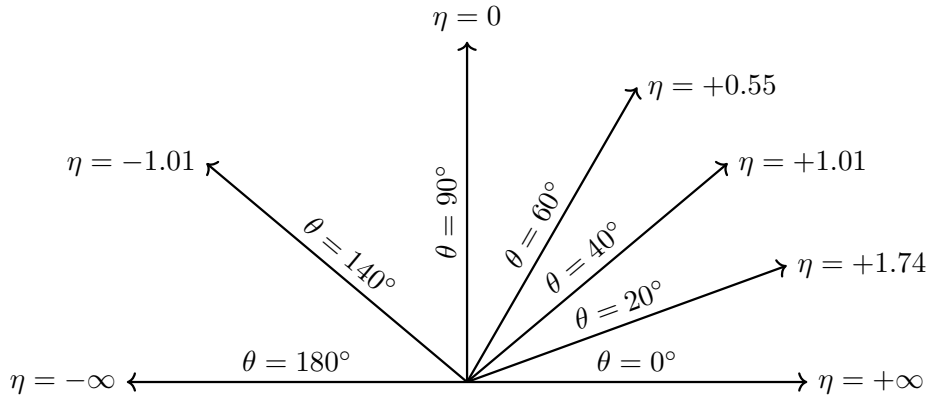


Figure 3.2.2: Azimuthal angle dependence of the pseudorapidity visualized for certain angles.

measure the tracks of charged particles flying through the detectors. All of the sub-detectors are segmented into a cylindrical barrel region and end-caps, covering a total pseudorapidity range of  $|\eta| < 2.5$ . The cross sectional image of the ID is presented in Figure 3.2.3. The entire inner detector is surrounded by a superconducting magnet

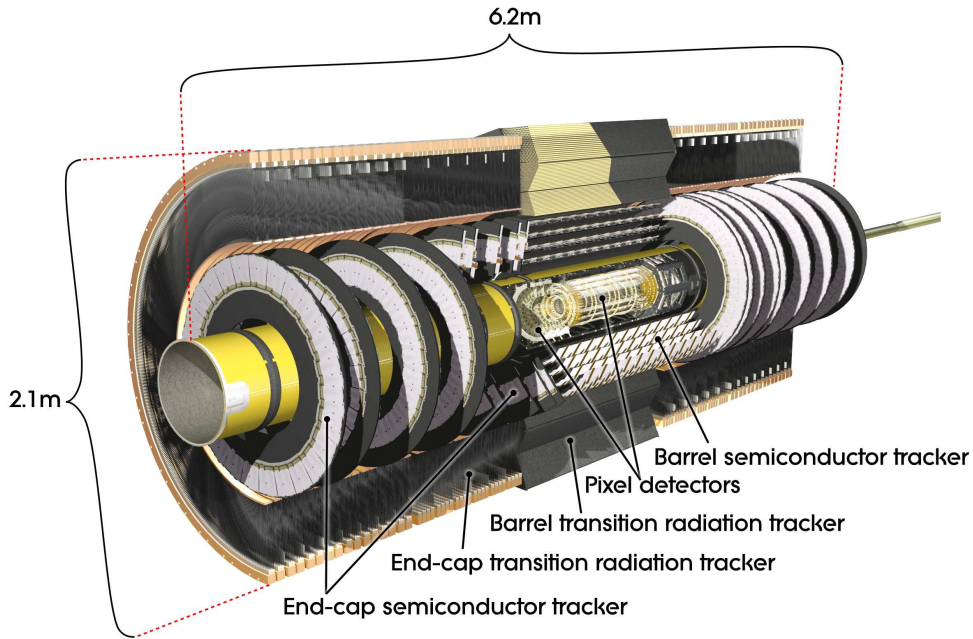


Figure 3.2.3: Cut-away view of the inner detector [42].

solenoid, creating a homogeneous magnetic field with field strength of 2 T. Charged particles going through the magnetic field experience a Lorentz force, bending the particle trajectory. From the curvature of the tracks the momentum of the particle can be calculated. Also track measurements are used to reconstruct the decay vertex, thus track measurement plays a critical role for distinguishing particles originating from different vertices.



### The pixel detector

The pixel detector is the innermost part of ATLAS, being the detector closest to the interaction point. It consists of four pixel layers concentrically arranged around the beam pipe in the barrel region. The innermost pixel layer, the so called Insertable B-Layer (IBL), was implemented during the LS1 phase, enhancing track resolution near the interaction point and significantly improving particle reconstruction [53]. The radial distance between the beam and the IBL is only 3.3 cm, so that the detector has to withstand enormous radiation throughout its planned lifetime. In the end-cap region the pixels are arranged in discs perpendicular to the beam. The pixels are doped silicon semiconductor detectors of size  $(R - \phi) \times z$  of  $50 \times 400 \mu\text{m}^2$ , resulting in an intrinsic accuracy in the barrel of  $10 \mu\text{m}$  in  $(R - \phi)$  and  $115 \mu\text{m}$  along  $z$ . The pixel detector covers a range of  $|\eta| < 2.5$  where layers are aligned such, that on average three layers are hit by a particle flying through.

### The semiconductor tracker

The pixel detector is surrounded by the SCT, made out of doped silicon microstrips as active material. There are 4 cylindrical layers in the barrel region covering  $|\eta| < 1.4$  and 9 disk layers for each end-cap extending the range up to  $|\eta| < 2.5$ . In total around 6.3 million silicon sensors are chained together in microstrips, where two strips are combined in stereo pairs. The SCT provides a slightly worse but still very high granularity of  $17 \mu\text{m}$  in  $(R - \phi)$  and  $580 \mu\text{m}$  in  $z$  ( $R$ ) direction in the barrel (end-cap) region compared to the pixel detector.

### The transition radiation tracker

The outermost detector of the ID is the transition radiation tracker. It consists of gas filled drift tubes with the length of 144 cm in the barrel region and 37 cm in the end-caps, covering the total range of  $|\eta| < 2$ . The tubes are 4 mm in diameter and stacked in layers parallel to the beam pipe, providing only  $R - \phi$  information. On average a particle crosses 36 straw tubes, adding a significant amount of additional track information for reconstruction. As fill gas a mixture of 27% carbon dioxide, 3% oxygen and 70% xenon is used. Due to gas leaks, some of the tubes are filled with argon instead of xenon for cost reasons. The gold plated tungsten wires going through the tubes are  $31 \mu\text{m}$  in diameter and operate with an electric potential difference between anode wire and cathode wall of 1.53 kV. As particles go through the tubes and collide with the gas atoms, these atoms are ionized causing drift currents between tube wall and anode wire where they are collected, amplified and read out. In addition, foils and polymer fibers are placed in between the tubes to create transition radiation if a charged particle goes through. The produced soft X-rays are absorbed by the xenon atoms creating an increased signal yield in the tubes. Since relativistic electrons produce the most amount of transition radiation, leaving behind a unique signature in the TRT, this adds additional electron identification capabilities.

### 3.2.3 Calorimeters

To measure the energy of particles, calorimeter systems are arranged around the solenoid magnet. The calorimeter system is divided into a electromagnetic and a hadronic calorimeter delivering the granularity required for energy reconstruction. Energy measurements are done by completely absorbing the particles and measuring the energy deposit in the active detector material. A cut-away view of the calorimeter system is shown in Figure 3.2.4.

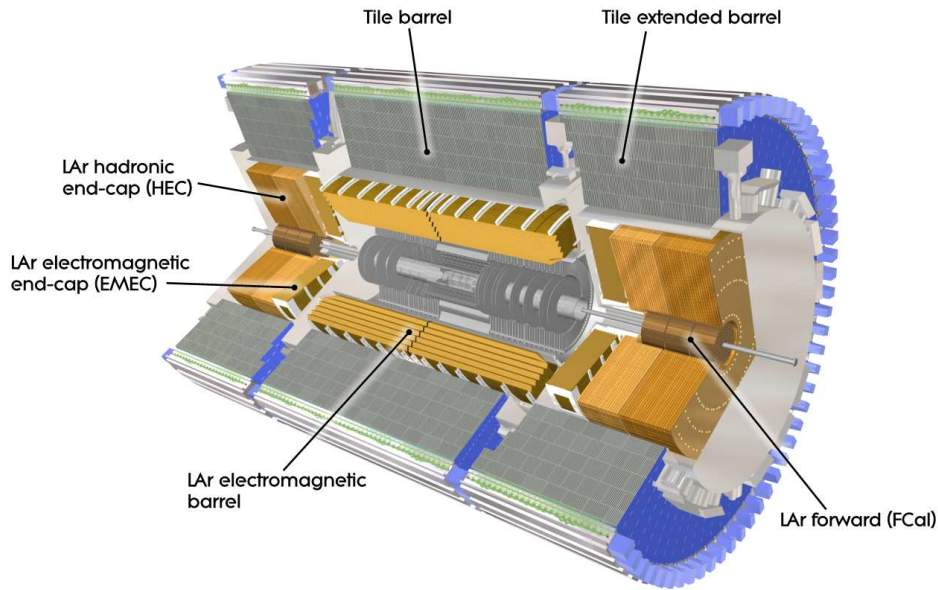


Figure 3.2.4: Cut-away view of the calorimeter system [42].

### The electromagnetic calorimeter

Following the ID, the electromagnetic calorimeter (ECAL) surrounds the solenoid magnet. The ECAL is divided into the barrel calorimeter covering the  $|\eta| < 1.475$  region and two end-caps for  $1.375 < |\eta| < 3.2$ , with the region of highest granularity at  $|\eta| < 2.5$ . However, the coarser-grained end-cap regions are sufficiently fine for the electromagnetic jet reconstruction. The ECAL is based on sampling detector technology with liquid argon (LAr) as active material combined with lead absorbers. When electrons or photons traverse the absorber materials, they will lose energy due to bremsstrahlung and pair production. The produced radiation ionizes argon atoms in the active materials creating free charges which are collected by electrodes. The resulting signal yields are proportional to the energy deposited. To minimize punch-through of electrons and photons into the hadronic calorimeter, the three Pb-LAr layers in the barrel region have a minimum total thickness of 22 times radiation length ( $X_0$ ) and at least 24 times  $X_0$  in the end-caps.

### The hadronic calorimeter

The ECAL is followed by the hadronic calorimeter (HCAL) which itself consists of different sub-detector parts. The part directly encompassing the ECAL is the tile calorimeter. The tile barrel covers the  $|\eta| < 1$  region and is extended on both sides with additional barrels covering  $0.8 < |\eta| < 1.7$ . It is a sampling calorimeter composed of steel plates as absorber material with scintillating tiles as active material in between. A total of 64 modules per barrel are segmented in three layers, reaching a thickness of 9.7 times the hadronic interaction length, thus minimizing punch-through of non-muon particles to the outer muon detectors. Particles that interact via strong interactions create showers in the absorber material. Secondary electromagnetic components of the showering stimulate the scintillator tiles which optical signals are captured and amplified by photomultipliers.

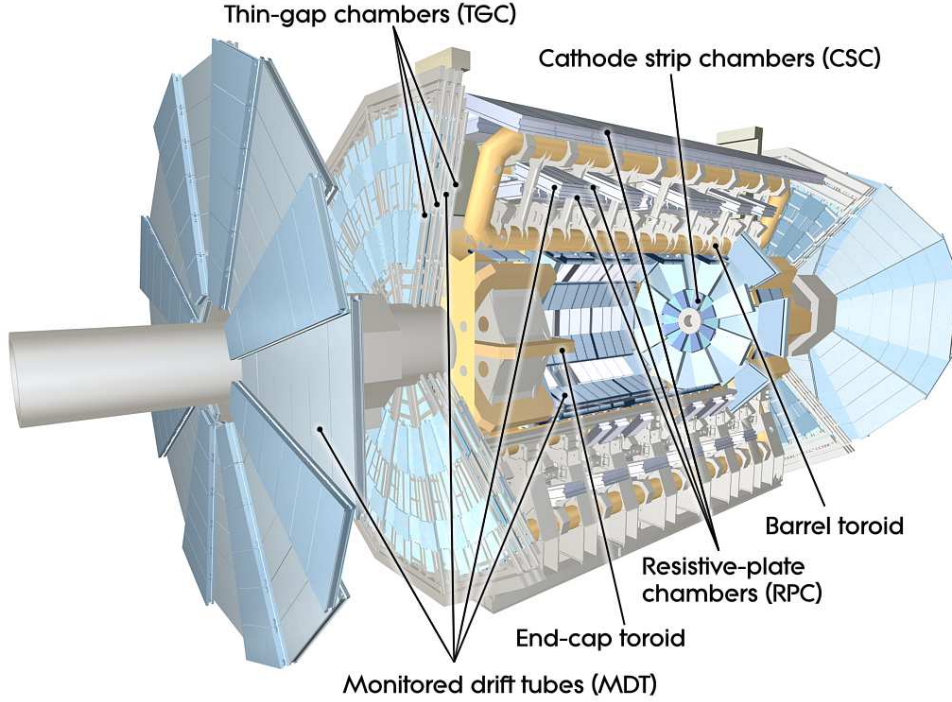


Figure 3.2.5: Cut-away view of the muon spectrometer [42].

In addition to the tile calorimeters, a LAr hadronic end-cap calorimeter (HEC) is placed behind the ECAL end-caps. The HEC covers  $1.5 < |\eta| < 3.2$ , overlapping the tile calorimeters. Similar to the ECAL, the HEC is made out of copper plates as absorber alternated by LAr as active material.

To further extend reconstruction at high pseudorapidity range, additional LAr forward calorimeters (FCal) are placed within the HEC near the beam pipe, covering  $3.1 < |\eta| < 4.9$ . The FCal is separated into two parts, one containing copper plates as absorber material, optimized for measuring electromagnetic interactions, and the other using tungsten plates measuring the energy of hadronic jets.

### 3.2.4 Muon spectrometer

Since muons are low ionizing particles at energies above several GeV, they mostly pass the calorimeters without depositing a lot of energy. To precisely measure the tracks of muons, dedicated muon detectors are placed around the calorimeter, building the outermost part of the ATLAS detector. In the barrel and end-cap regions, most parts are located around or inside the toroid magnet, creating an almost circular 0.5 T magnetic field that bends the trajectories of muons passing through. The toroid magnets are divided into a large barrel magnet ( $|\eta| < 1.4$ ) and end-cap magnets ( $1.6 < |\eta| < 2.4$ ). The end-cap detectors are arranged in large wheels perpendicular to the  $z$ -axis in distances of  $|z| \approx 7.4$  m, 10.8 m, 14 m and 21.5 m. To track the muons, four different types of detectors are used (see Figure 3.2.5).

Monitored drift tubes (MDTs) are cylindrical drift chambers filled with a gas mixture of argon and carbon dioxide operating at a pressure of 3 bar. Muons crossing the MDT ionize gas atoms creating free electrons that are collected by a central gold plated tungsten-rhenium wire anode with a potential difference of 3080 V. However,

due to the drift time of electrons from the chamber wall to the wire of about 700 ns, the readout speed is limited. The tubes are stacked in 3 to 8 layers on a rectangular support structure creating MDT chambers. These chambers achieve a resolution of 35  $\mu\text{m}$  and are used throughout the entire spectrometer in the end-cap wheel and in the barrel region, enabling the precise measure of muon tracks.

In the forward region,  $2.7 > |\eta| > 2$ , cathode strip chambers (CSC) are implemented to match the higher readout speed requirement in regions with higher particle flow densities. Similar to the MDT, the CSC are drift chambers filled with an argon-carbon dioxide gas mixture. The CSCs are constructed as flat chambers holding parallel aligned anode wires operating at 1900 V. Cathode strips are placed in orthogonal direction to the wires with a cathode-anode gap of 0.25 mm, allowing for  $\eta$  and  $\phi$  measurement. Position measurements of the tracks are done by analyzing and interpolating the signals on the cathode strips, while the anode wires are not read out. Four CSC elements are packed together in planes and are placed at the inner side of the toroid end-cap magnets. The resolution of the CSC planes vary between 80  $\mu\text{m}$  and 5 mm. The CSCs have fast time resolution compared to MDTs of 7 ns per plate. In the barrel region, additional to the MDTs, resistive plate chambers (RPC) are implemented. The RPC consists of 2 parallel superimposed resistive plates with a gas gap in between. On the outer faces of resistive plates, graphite electrodes are mounted, generating a electric field 9.8 kV. On top of both electrodes are readout strips running orthogonal to each other, allowing for  $\phi$  and  $\eta$  resolution of the muon tracks. The 2 mm wide gas gap is filled with a mixture of tetrafluoroethane, methylpropane and sulfur hexafluoride. Muons passing through the gap ionize atoms and create free electrons. Under the high electrical potential differences, these free electrons ionize the gas themselves, causing avalanches of free charges as they travel to the anode. The cascading charges induce a signal in the readout strips. RPCs are capable to operate at a readout speed of  $\sim 1 \text{ kHz}/\text{cm}^2$ , thus can be used to trigger events where muons are present. Thin gap chambers (TGC) are placed in the inner three wheels in the  $1.05 < |\eta| < 1.92$  range, to provide trigger capabilities in the end-cap region and additional azimuthal measurements. Like CSCs, TGCs are multi-wired proportional chambers. The fill gas used is based on carbon dioxide and n-pentane, which has good quenching properties and thus allows for a low drift time of the free electrons of  $\sim 25 \text{ ns}$ .

### 3.2.5 Event collecting and the trigger system

Since the LHC is colliding particles at a rate of 40 MHz, the enormous amount of data can not be recorded for every single event. Also most components, like the ID, are due to technical limitations unable to be read out at this frequencies. Therefore a two layered trigger system is introduced to select events based on predefined cuts using informations from fast detector components [54]. The first trigger level reduce the data rate significantly, enabling the next trigger to make further selections.

The first trigger is the Level-1 trigger (L1) which is directly implemented on hardware components. The L1 trigger system is separated into a L1Calo, using deposit information of all calorimeters, an L1Muon trigger that rely on the RPC in the barrel and TGC in the end-cap region, and a dedicated L1Topo trigger that looks for specific topological structures in the event by analyzing the signatures in all trigger systems. Using calorimeter information, the L1Calo trigger is capable of selecting events with high  $E_T$ , such as events involving jets and hadronically decaying tau leptons as well as events with high  $E_{T,\text{mis}}$ . The L1Muon triggers perform a threshold selection on

the  $p_T$  of muons. In total, the L1 trigger reduces the event rate to 100 kHz. All of the trigger informations are forwarded to the Central Trigger Processor (CTP) that decides whether an event is triggered or not. Upon accepting an event, the information of the Region of Interest (RoI) is send to the High Level Triggers (HLT). During the time it takes for the L1 trigger to form a decision, the data collected by the detector parts are stored temporally in data-pipelines often directly located near the detector. The HLT is a software based trigger which works on a large distributed computer farm. From the L1 trigger, the HLT gets the information about the RoI and send a request to gather the data from the detector. Using the full detector information, the HLT is able to apply more dedicated selections and consequently lowering the event yield to values between 600 and 1000 Hz. Events that pass the HLT are collected and written to the storage system. Different event filters of a HLT can be applied on the events, optimized to select specific physics objects like tau lepton decays. Since the total write out bandwidth is limited, for each of the HLT a prescale factor is calculated according to the expected rate a HLT will fire in dependence of the luminosity. If the trigger rate of a HLT is expected to exceed the data stream the system can handle under certain run conditions, only every other event is written to the storage. The actual amount of the events can be recalculated by applying the prescale factors. An overview over the ATLAS trigger system is given in Figure 3.2.6.

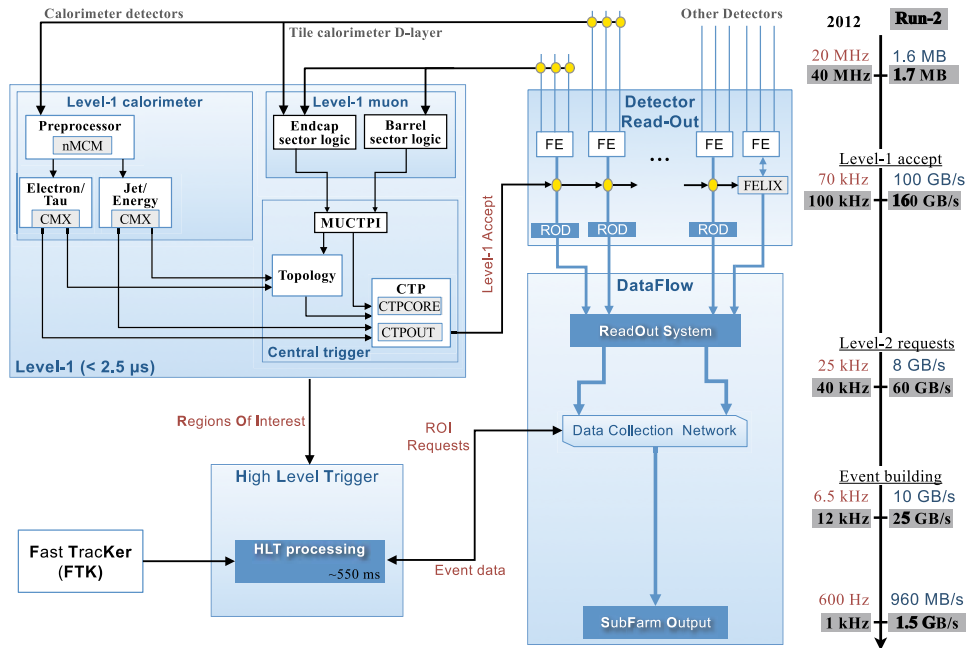


Figure 3.2.6: Schematic representation of the ATLAS trigger system [54].



## 4 Monte Carlo generation and data samples

Data recorded by the ATLAS detector that is used for this analysis was gathered in the data taking periods beginning from August 2015 to July 2016, corresponding to a total integrated luminosity of  $13.2 \text{ fb}^{-1}$  [50]. The luminosity history for this and previous periods is shown in Figure 4.0.1. For the data taking period, the ATLAS

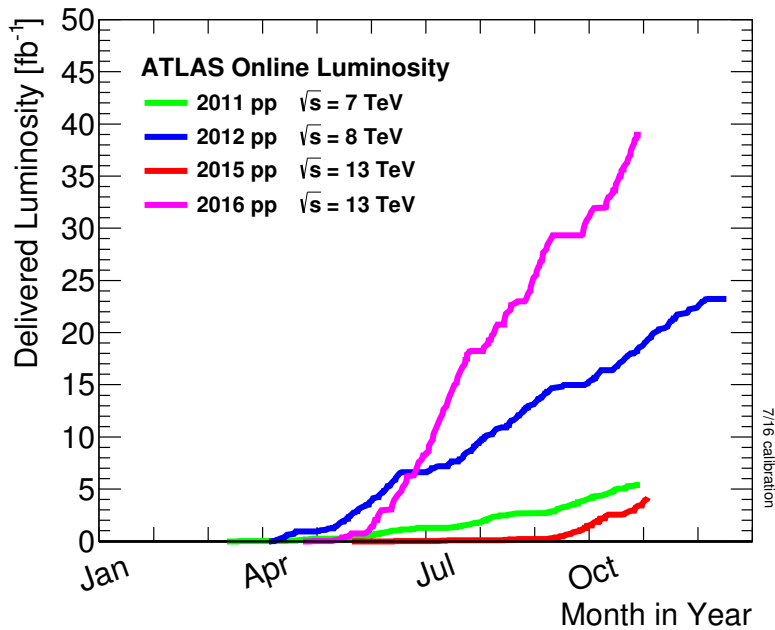


Figure 4.0.1: Integrated luminosity recorded by ATLAS [50]. The red line corresponds to luminosity taken in 2015 and the purple line in 2016 for proton-proton collisions. Data from these periods are used for this analysis.

detector recorded data from proton-proton collisions of the LHC with a center of mass energy of  $\sqrt{s} = 13 \text{ TeV}$ .

The generation of simulated events plays a crucial role in the search for physics beyond the SM, as it enables the simulation of SM processes, which are used for background analysis and the simulation of signal processes according to the prediction of the theory tested. For simulation Monte Carlo generators (MC) are used like PYTHIA [55,56] or Sherpa [57]. The MC generators calculate the matrix element for a given process using Feynman diagrams and the Lagrangian function of the underlying theory, such as electroweak quantum field theory or perturbative QCD. Since protons are collided and thus QCD effects play a major role, the perturbative Matrix Element (ME) approach works only for hard scattering processes, because QCD contributions at lower energy scales can not be described by perturbative theories due to confinement of QCD color charge. Examples of soft scattering events that have to be considered are

initial and final state radiation. These arise when particles emit other particles before or after the main hard scatter event. For example a quark created in a hard scatter process can radiate gluons that decay into quarks and further hadronizes, forming cone like propagating clusters of particles called QCD jets. Contributions from non-perturbative QCD processes can be calculated using Parton Shower (PS) models. The PS allows to model the regime between hard scattered and soft QCD processes where hadronization of quarks plays a major role. As collisions of protons are examined, simulations have to take into account the Parton Distribution Functions (PDF) of the quarks and gluons in the proton. The PDFs used for simulation have to be measured by experiments as they can not be predicted by theory.

To test the hypothesis of heavy neutral Higgs bosons predicted by the MSSM, signal event samples for gluon-gluon fusion (ggH) and b-associated production (bbH) described in Figure 2.3.2 are simulated. To test a wide range of the  $m_A$ - $\tan(\beta)$  parameter space of the MSSM, different event samples for mass hypotheses from 300 GeV up to 1200 GeV are produced. The cross section of the signal event samples are calculated using SusHi [38, 58] and Higgs couplings by FeynHiggs [32–37].

For background estimation  $Z$ ,  $W$ , di-boson ( $ZZ$ ,  $WW$ ,  $ZW$ ),  $t\bar{t}$  and single top events are simulated. The generators used to generate these event samples can be found in Table 4.0.1. Different MC generators are used for generating hard scattered events with matrix element approximation and for parton showering, except for event samples generated with Sherpa that can do both.

To improve event population at high transverse momenta, the event samples of  $W$  production processes are sliced in  $p_T$  and the Drell-Yan process event samples in resonance mass at matrix element level. The cross sections for  $Z$  and  $W$  production are calculated using the FEWZ simulation framework at next-to-next-to-leading order [59–61].

To model the background arising from top-quark production and decay,  $t\bar{t}$  and single top events are generated. For generating the single top background, s- and t-channels as well as the so called  $Wt$ - production channel [62], with a top-quark and an additional  $W$  boson in the final state, are considered. The top quark decay in all event samples are simulated by MadSpin [63] while the decay of lighter hadrons, including b- and c-quarks, are corrected with the EvtGen v1.2.0 [64] decay simulator.

Since the readout speed of the detector is limited and the readout cycle of inner detector parts is multiple times longer than the average time between bunch-crossings, the overlay of multiple scatter processes, the so called pile-up, has to be modeled in the Monte Carlo simulations. There are two types of pile-up. In the so called in-time pile-up additional inelastic scatter events are present next to the main hard scatter vertex, while out-of-time pile-up describes multiple overlaying hard scatter events from different bunch crossings due to the readout inertia of the detector. The pile-up is modeled using events generated with Pythia 8.186.

For all Monte Carlo generated background event samples and the ggH signal events the detector response is simulated with the GEANT4 [65] simulation software combined with a detailed description of all parts of the ATLAS detector. Large amounts of negative events weights in the bbH event sample due to redundancy in the shower modeling of the ME and PS Monte Carlo generators are observed. Therefore more signal events have to be generated to match the required event yields for modeling the signal. Because the full detector simulation with GEANT4 is very processing expensive, the fast ATLAS simulation ATLFAST-II [66, 67] is used for the MSSM bbH signal. After detector simulation all MC event samples are processed with the



same software used for data processing. However, particle information from the MC generator are also kept as so called truth information in the simulated event samples. The processed data is stored in a container format called Derived Analysis Object Data (DxAOD) [78] which allows a standardized access to event information.

Process	MC Generator	
	ME-Generator	PS-Generator
<b>Signal</b>		
ggH	POWHEG-BOX v2 [68–71]	PYTHIA 8.210 [56]
bbH	MADGRAPH5_aMC@NLO 2.1.2 [72, 73]	PYTHIA 8.210
<b>Background</b>		
$Z$	POWHEG-BOX v2 [74, 75]	PYTHIA 8.168 [76]
$W$	Sherpa 2.2.0 [57]	
Di-boson	Sherpa 2.2.0	
$t\bar{t}$ , single top s- $W$ t-channel	POWHEG-BOX v2 [62, 77]	PYTHIA 6.428 [55]
single top t-channel	POWHEG-BOX v1 [77]	PYTHIA 6.428

Table 4.0.1: List of Monte Carlo generators used for signal and background simulation



# 5 Analysis Selection

## 5.1 Physics object reconstruction and identification

In this thesis the decay of heavy neutral Higgs bosons predicted by the MSSM into a pair of hadronically decaying taus is studied. A major task is to reconstruct and identify these tau leptons in the data collected by the ATLAS experiment. Furthermore, it is also important to reconstruct electrons, muons and quark-gluon jets as well as possible because they can leave similar detector signatures, thus contributing to the process background. Hereinafter is listed how these objects are reconstructed from data recorded by the ATLAS detector.

### Electron reconstruction and identification

Electrons are reconstructed from the signature left in the electromagnetic calorimeter as well as pixel, SCT and TRT hits in the inner detector which can be used to calculate the track parameters of the particles [79]. Clusters of the calorimeter information are built using a seeded cluster reconstruction. The seed is generated by a sliding window technique where the calorimeter layers are divided into grids which granularity correspond to the energy resolution of the calorimeter and the energy deposits of the grid layers are summed and stacked. The clusters are then seeded by stacks with a total transverse energy deposit of more than 2.5 GeV.

The track reconstruction is done via pattern recognition [80]. The pixel hits in the inner detectors are compared with the expected bremsstrahlung energy loss of the particle for each crossing with the detector material. The track candidates are then matched with the cluster barycenter by a fitting algorithm that take bremsstrahlung effects into account.

For the identification of electrons a likelihood method is used. This multivariate method assigns a identification class to electron candidates that discriminates electrons from background. Variables used for the identification arise from calorimeter and track entries, track-cluster matching and bremsstrahlung effects.

There are three predefined working points for the electron identification loose, medium and tight corresponding to increasing background rejection efficiency.

### Muon reconstruction and identification

Muons are reconstructed independently using information from the muon spectrometers and the inner detectors [81]. The spectrometer reconstruction utilizes information from hits in the muon drift tubes that have to be aligned with features observed in other spectrometer layers.

The muon candidate is then matched to hits in the inner detector, taking into account the distortion of the path caused by magnetic fields.

The identification of muons is achieved by cuts on variables optimized to separate muon signal from background. Like for the electron identification there are three inclusive working points available, loose, medium and tight and an additional high- $p_T$  working point.

### Jet reconstruction

QCD jets, caused by quark and gluon showering, are reconstructed from energy deposits in the calorimeter using an anti- $k_t$  algorithm to build clusters [82, 83]. The algorithm calculates the weighted distance between two calorimeter deposits  $d_{ij}$  and the beam axis  $B$ :

$$d_{ij} = \min \left( k_{T,i}^{2p}, k_{T,j}^{2p} \right) \frac{\Delta_{ij}^2}{R^2},$$
$$d_{iB} = k_{T,i}^{2p}, \tag{5.1}$$

$$\text{with } \Delta_{ij}^2 = (y_i - y_j)(\phi_i - \phi_j),$$

where  $k_{T,i}$  is the transverse momentum,  $\phi_i$  the azimuthal angle and  $y_i$  the rapidity of the  $i$ -th entity. The parameter  $R$  is the radius parameter and  $p$  a factor that governs the behavior of the algorithm. This clustering algorithm is called  $k_t$ -algorithm when  $p > 0$  and anti- $k_t$ -algorithm when  $p < 0$ . In each iteration the two entries with the shortest distance from one another are merged to form a cluster. If  $d_{iB}$  is the shortest distance, the corresponding cluster is removed from further iterations and called jet. This is done until all deposits are clustered to jets.

In this analysis anti- $k_t$  jets with  $R = 0.4$  and  $p = -1$  are used.

### Tau reconstruction and identification

The baseline of reconstructing hadronically decaying tau leptons are jets formed on calorimeter clusters by an anti- $k_t$  algorithm with  $R = 0.4$  [84]. The reconstruction is restricted to the  $|\eta| < 2.5$  region, excluding the barrel–end-cap overlap region.

The tau clusters are matched with the primary vertex of the event. If more than one vertex is present that matches with the tau jet, the one with the highest squared transverse momentum sum of the associated particles is used as tau vertex. All variables further used for analysis are calculated with respect to the tau vertex.

In addition to calorimeter deposits, also track information is used for reconstruction. All tracks in the inner detectors that fall inside the core region of  $\Delta R < 0.2$  around the tau vertex corrected jet axis and exceed a  $p_T$  threshold of 1 GeV are associated with the tau candidate.

However, the tau reconstruction yields not sufficient rejection power against QCD jets. Therefore an identification algorithm based on Boosted Decision Trees (BDT) [90,92], a multivariate analysis technique, is used to separate jets originated from visible tau lepton decays and quark-gluon jets. To improve separation power, different BDTs are trained for taus with charge multiplicity of 1 and 3. The BDTs are trained on Monte Carlo simulated events of  $Z/\gamma^* \rightarrow \tau\tau$  as signal and QCD di-jet samples as background. Configuration and training of the BDTs used for tau lepton identification is described in Ref. [84].

Likewise to the electron identification, three working points are defined for the BDT based identification, loose, medium and tight corresponding to 0.6 (0.5), 0.55 (0.4), 0.45 (0.3) signal efficiency for a charge multiplicity of 1 (3).

### B-Tagging

Jets are also tested whether they originated from decays of bottom-flavored hadrons to exploit the presence of b-quarks in the final state of b-associated Higgs production. B-mesons have relatively long lifetimes of  $\tau_b \sim 10^{-12}$  s, thus the decay of the

b-hadrons leave behind a distinctive signature where the decay vertex is slightly displaced from the primary vertex. The B-Tagger delivers a multivariate analysis tool to discriminate bottom decay jets from QCD background [85, 86]. It uses a BDT classifier with variables derived from fitting particle tracks to secondary b-decay vertices. Different working points on the BDT output are given at 60 %, 70 %, 77 % and 80 % identification efficiency.

## 5.2 Event selection

The event selection plays a crucial role in the final analysis results, so kinematic cuts and selection criteria have to be applied carefully. The selection is chosen to suppress background contributions while enriching the expected signal that is searched for and is described in detail in Ref. [87].

In this analysis at least two  $\tau_{\text{had,vis}}$  are required, where the one with the highest  $p_{\text{T}}$  is called leading tau candidate and the one with the second highest subleading tau.

The leading tau candidate has to match the object reconstructed by a tau HLT that triggered the event in a cone around  $\Delta R \leq 0.2$ . Because of the high instantaneous luminosity, two inclusive high level triggers with a  $p_{\text{T}}$  threshold of 80 GeV and 125 GeV are used. These HLT trigger when at least one jet originating from a possible tau decay is observed in the event. For each event the unrescaled trigger with the lowest  $p_{\text{T}}$  threshold is used for matching.

Muons that are selected have to pass a medium identification requirement. In addition they have to exceed a  $p_{\text{T}}$  threshold of 7 GeV and be measured within a pseudorapidity range of  $|\eta| < 2.5$  to be considered as muons.

In the electron selection, electrons have to pass a loose identification requirement and satisfy  $p_{\text{T}} > 15$  GeV. Due to constraints in the reconstruction, only electrons in the  $|\eta| \leq 2.47$  region, excluding the barrel–end-cap overlap region  $1.37 \leq |\eta| \leq 1.52$ , are taken into account. To suppress background originating from light leptons that are misidentified as the decay of tau leptons, a veto on muons and electrons is applied, so that an event gets discarded if a muon or electron fulfill the requirements presented above.

Reconstructed jets have to pass a  $p_{\text{T}} > 20$  GeV threshold and be located in the  $|\eta| \leq 2.5$  region. As pointed out in section 5.1, hadronically decaying tau leptons are reconstructed from jets. To avoid double counting of jet objects as QCD jets and  $\tau_{\text{had,vis}}$  candidates, an overlap removal is applied that removes jets that are within a  $\Delta R \leq 0.2$  cone of a  $\tau_{\text{had,vis}}$  candidate. To suppress effects caused by pile-up for jets with  $p_{\text{T}} \leq 60$  GeV and  $|\eta| < 2.4$ , a jet vertex tagger algorithm is used [88]. It delivers a multivariate technique to calculate the likelihood that a jet originated from the primary vertex and relies on information of the jet tracks, energy and vertex matching. Only jets accepted by the tagging algorithm are considered.

In the following the main selection criteria on the  $\tau_{\text{had,vis}}$  candidates are listed:

- Leading  $\tau_{\text{had,vis}}$  passes medium BDT identification requirement
- Leading tau  $p_{\text{T}} \geq 110$  GeV or 140 GeV depending on the  $p_{\text{T}}$  threshold of the lowest unrescaled tau trigger
- Subleading tau passes loose BDT identification requirement
- Subleading tau has  $p_{\text{T}} \geq 55$  GeV

- Both taus have a reconstructed charge multiplicity of 1 or 3
- $\Delta\phi > 2.7$  between leading and subleading tau
- Oppositely charged tau leptons with absolute charge of  $1e$

Since the trigger efficiency is not flat from the beginning of the trigger  $p_T$  threshold, but has a slope that reaches a plateau above higher  $p_T$  values, the  $p_T$  cuts are chosen to ensure a stable trigger efficiency. The cut on  $\Delta\phi$  is motivated from the expectation, that heavy Higgs bosons are produced almost at rest in the transverse direction and thus decay products would radiate outwards back-to-back in the  $\phi$ -plane.

In addition to kinematic cuts and identification requirements, the data blocks analyzed have to be listed in the so called good-runs-list that includes all data tanking periods where the detector was fully functional and operated normally [89].

To further enhance the sensitivity of the analysis, topological differences of the Higgs production channels can be exploited. As shown in Figure 2.3.2 the b-associated production of Higgs bosons, additional b-quarks are involved in the final state. To enrich the b-associated signal, a dedicated B-Tagging algorithm with a identification efficiency of 70 % is used to identify jets originating from b-hadron decay.

Two orthogonal categories are defined depending on whether a b-jet is present or not:

- B-TAG category: At least one b-tagged jet
- B-VETO category: No b-tagged jets.

In the B-TAG category, the  $p_T$  threshold for the subleading tau candidate is increased to 65 GeV to improve background modeling.

The selection outlined here is further referred to as signal region. The requirement of oppositely charged tau leptons allows for the definition of a same-sign validation region where both taus have the same charge. This same-sign region is orthogonal to the signal region and serves for validation of the background modeling in the variable distributions since the signal expectation in this region is very small.

The event yields in the signal region is summarized in Table 5.2.1. No significant deviation of the total event yields between the expected SM background and the data in the signal region are found. How the differed backgrounds are modeled and estimated is covered in section 5.3.

### Event yields in B-VETO

Process	Nominal	Statistical Unc.	Systematic Unc.
Multi-jet	1510	$\pm 12$	+190 -190
$Z/\gamma^* \rightarrow \tau\tau$	340	$\pm 6$	+90 -80
$W \rightarrow \tau\nu + \text{jets}$	79	$\pm 3.6$	+13 -12
$t\bar{t}$ , single top	15	$\pm 1.4$	+17 -17
Others	12.9	$\pm 1.1$	+2.8 -2.5
SM expected	1960	$\pm 13$	+210 -210
Data	2010	$\pm 40$	
Signal $m_A = 500 \text{ GeV}$ $\tan(\beta) = 20$	69	$\pm 1.3$	+19 -18

### Event yields in B-TAG

Process	Nominal	Statistical Unc.	Systematic Unc.
Multi-jet	47	$\pm 2.1$	+12 -12
$t\bar{t}$ , single top	29	$\pm 2.0$	+11 -10
$Z/\gamma^* \rightarrow \tau\tau$	4.1	$\pm 0.9$	+1.4 -1.4
$W \rightarrow \tau\nu + \text{jets}$	1.39	$\pm 0.17$	+0.27 -0.27
Others	0.7	$\pm 0.5$	+0.27 -0.29
SM expected	83	$\pm 3.1$	+16 -16
Data	63	$\pm 8$	
Signal $m_A = 500 \text{ GeV}$ $\tan(\beta) = 20$	29	$\pm 0.9$	+8 -8

Table 5.2.1: Event yields after event selection in B-VETO (top) and B-TAG (bottom). No significant excess of data over the SM expectation is observed.

## 5.3 Background estimation

Although the event selection suppresses background while minimizing signal loss, the remaining background contributions have to be well modeled to see if any significant deviation from the SM expectations in the data of the signal region are observed. Estimation of background contributions can be either done by Monte Carlo simulation or by data driven methods, which require a definition of a control region to measure and estimate the background contributions, that have to be extrapolated to the signal region.

Most of the background contribution arises from jets and leptons faking  $\tau_{\text{had, vis}}$ . However, there is also irreducible background from processes with real  $\tau_{\text{had}}$  in the final state. The procedure of background estimation is fully described in Ref. [87].

### 5.3.1 Data driven QCD background estimation

The leading background contribution in the B-TAG and B-VETO category arises from multi-jet events, where jets from hadronizing gluons or quarks are misidentified as  $\tau_{\text{had, vis}}$ . The background is estimated using a data driven fake factor technique. Fake factors are measured in a fake factor control region which is enriched with multi-jet background, while having a similar but orthogonal phase space to the signal region and containing almost no signal contamination. To estimate the background, a tag and probe technique is applied, where the leading tau selection is orthogonal to the signal region selection (tag) whereas the subleading tau is used to measure the fraction of jets faking tau decays (probe). The background estimation is done independently for B-TAG and B-VETO in the following fake factor control region:

- Event is triggered by single jet triggers
- Leading  $\tau_{\text{had, vis}}$   $p_T > 100$  GeV
- Leading  $\tau_{\text{had, vis}}$  fails medium BDT identification requirement
- $\frac{p_T^{\text{subleading } \tau_{\text{had, vis}}}}{p_T^{\text{leading } \tau_{\text{had, vis}}}} > 0.3$
- B-VETO:
  - Leading  $\tau_{\text{had, vis}}$  candidate has charge multiplicity of 1 or 3
  - No b-tagged jet
- B-TAG:
  - No more than 7 charged tracks for leading  $\tau_{\text{had, vis}}$
  - No charge requirement on leading  $\tau_{\text{had, vis}}$
  - At least 1 b-tagged jet
- Signal region fake factors:
  - Opposite sign charged leading and subleading  $\tau_{\text{had, vis}}$
- Same-sign validation region fake factors:
  - Same sign charged leading and subleading  $\tau_{\text{had, vis}}$



Separate fake factor measurements are done for signal region and same-sign validation region. The cut on the  $p_T$ -ratio is applied to match the ratio between jets originating from gluons and quarks to the ratio expected in the signal region, therefore creating a more similar phase space occupancy. In order to enrich the number of multi-jet events in B-TAG, the cut on the charge multiplicity of the leading  $\tau_{\text{had,vis}}$  is loosened.

The fake factors  $f_{mj}$  are then defined as the ratio of the number of subleading  $\tau_{\text{had,vis}}$  that pass the loose BDT identification requirement and those that fail the requirement:

$$f_{mj}(p_T, N_{\text{Track}}) = \frac{N_{\text{pass BDT-ID}}(p_T, N_{\text{Track}})}{N_{\text{fail BDT-ID}}(p_T, N_{\text{Track}})}. \quad (5.2)$$

To model the transverse momentum dependence,  $f_{mj}$  is parameterized in  $p_T$  and the charge multiplicity  $N_{\text{Track}}$  of the subleading tau decay.

The shape of the multi-jet background in the signal region is estimated with data from a fail-ID control region by applying the fake factors estimated in the fake factor control region. This fail-ID region is similar to the signal region, except that the subleading  $\tau_{\text{had,vis}}$  fails the loose BDT identification requirement. Like the fake factor control region, the fail-ID region contains mostly multi-jet contributions. Background contaminations of real tau leptons in the fail-ID region are estimated using Monte Carlo simulation and are subtracted from the multi-jet estimations. For a given kinematic variable  $x$ , the multi-jet contribution in the signal region is then calculated to:

$$N_{\text{multi-jet}}^{\text{signal region}}(x, p_T, N_{\text{Track}}) = N_{\text{data}}^{\text{fail-ID region}}(x, p_T, N_{\text{Track}}) \times f_{mj}(p_T, N_{\text{Track}}). \quad (5.3)$$

The modeling of the multi-jet background can be validated in the same-sign validation region which is especially enriched with QCD events.

### 5.3.2 Monte Carlo background estimation

The second leading background contribution in the B-VETO region comes from off-shell  $Z$  boson decays into pairs of tau leptons. Because there are true taus in the final state, the  $Z \rightarrow \tau\tau$  decay causes an irreducible background in the signal region. As described in Chapter 4,  $Z \rightarrow \tau\tau$  background is simulated using Monte Carlo generators.

Another important background is the decay of a  $W$  boson into a tau lepton and a neutrino, where an additional QCD jet is misidentified as second  $\tau_{\text{had,vis}}$ , thus this background is further referred as  $W \rightarrow \tau\nu + \text{jets}$ .

In the B-TAG region, the second leading background arises from top-quark decays into a bottom-quark and a  $W$  boson, which can decay further into a lepton and lepton-neutrino. Especially the decay of a  $t\bar{t}$  system can look very similar to the b-associated Higgs production with two true tau leptons and two b-quarks in the final state. Background contributions from  $Z$  decays are strongly suppressed in this category due to the requirement of at least 1 b-tagged jet.

Minor backgrounds estimated with Monte Carlo generators are di-boson production and the decay of  $Z$  and  $W$  bosons into muons and electrons.

In most of the Monte Carlo simulated backgrounds one or more QCD jets are misidentified as  $\tau_{\text{had,vis}}$ . A significant mismatch between the rate of QCD jets faking hadronic tau decays in data and Monte Carlo generated events is observed. To correct the mis-modeling, fake rates are measured from data and applied to  $\tau_{\text{had,vis}}$  candidates that do not come from true tau decays in Monte Carlo generated events.

Fake rate measuring is done separately for B-TAG and B-VETO in a fake rate control region defined by:

- Event is triggered by a muon HLT with  $p_T > 50$  GeV threshold
- HLT muon  $p_T > 55$  GeV
- $\Delta\phi(\mu, \tau_{\text{had, vis}}) > 2.4$
- $m_T(\mu, E_{T, \text{mis}}) > 40$  GeV
- $N_{\text{b-jet}} = 0$  for B-VETO and  $N_{\text{b-jet}} \geq 1$  for B-TAG

This region is enriched with  $t\bar{t}$  and single top events in B-TAG and  $W \rightarrow \tau\nu + \text{jets}$  in B-VETO.

## 6 Multivariate analysis

After event selection and background estimation a final distribution has to be chosen that yields good separation power between signal and background. In the previously published search for heavy neutral MSSM Higgs bosons decaying into tau leptons at ATLAS in Ref. [87], the total transverse mass  $m_T^{\text{tot}}$  of the di-tau system is used for the statistical analysis. It is defined as:

$$m_T^{\text{tot}} = \sqrt{m_T^2(E_{T,\text{mis}}, \tau_1) + m_T^2(E_{T,\text{mis}}, \tau_2) + m_T^2(\tau_1, \tau_2)}, \quad (6.1)$$

with:

$$m_T(a, b) = \sqrt{2p_T(a)p_T(b)[1 - \cos(\Delta\phi(a, b))]}.$$
 (6.2)

The distribution of  $m_T^{\text{tot}}$  in B-TAG and B-VETO in the signal region can be found in Figure 6.0.1.

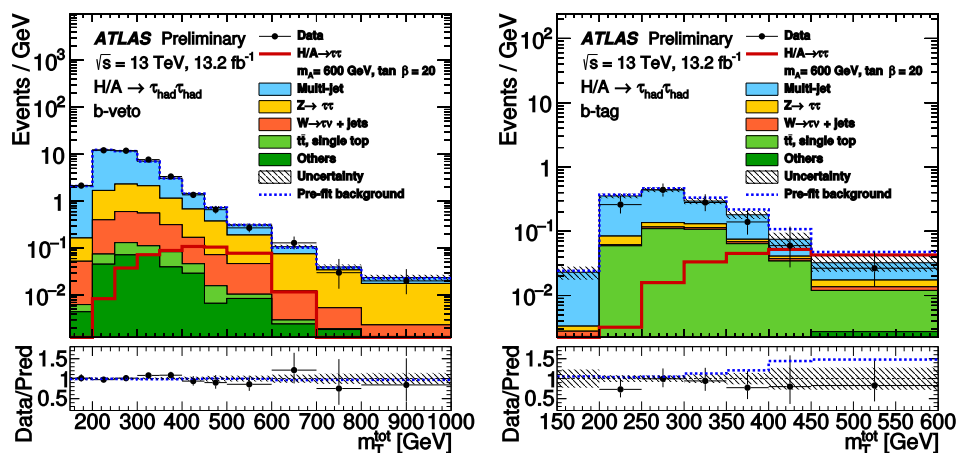


Figure 6.0.1: Total transverse mass distributions used in the previously published analysis [87].

In this thesis a new approach is presented using a multivariate analysis technique, the Boosted Decision Tree (BDT), to improve separation of signal and background. The BDT is trained and applied at event level, thus using basic kinematic variables from the reconstructed detector objects. A BDT uses information from multiple variables to decide if an event is more background or signal like. This decision is reflected in a so called response distribution which can be used for statistical analysis instead of  $m_T^{\text{tot}}$ .

First the theory of BDTs is described. After that, the implementation and training procedure is explained in detail, followed by the training results.

### 6.1 Boosted Decision Trees in the TMVA framework

Boosted Decision Trees are multivariate learning algorithms that are based on many simple decision trees as weak learner and can be used for classification problems [90].

The trees of a BDT have to be trained on a training data set that provides event information of a given set of variables  $\mathbf{X}$  and the desired outcome. Because the final distribution for this analysis can be simplified to signal and background like events, the BDT is applied to a classification problem and thus also referred to as classifier. An individual tree is built out of leaves and nodes in a tree structure originating from the root node as shown in Figure 6.1.1. For each node, the algorithm takes the best

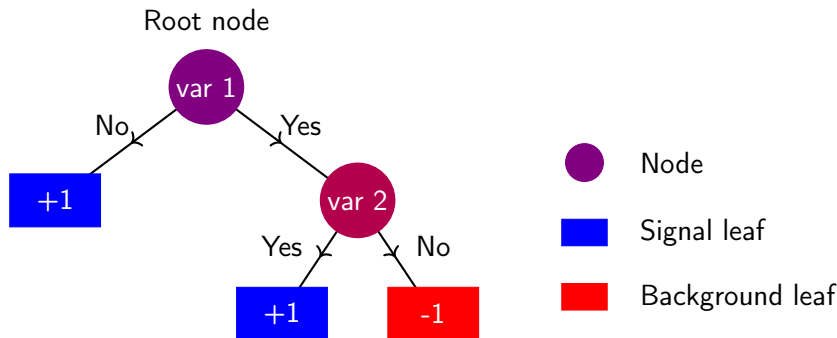


Figure 6.1.1: Schematic representation of a decision tree. Each node sets a criterion on a variable that can be passed or failed. Depending on the result, events go down further nodes in the tree until a leaf is reached that assigns a value to the event. The color of the nodes represents the mixture of background (blue) and signal (red) that goes into the node.

separating variable  $x \in \mathbf{X}$  and finds a cut that best separates signal and background events by maximizing the gain of the Gini-Index [91].

The Gini-Index  $G$  is a statistical measure that represents the inequality between signal and background. It is defined as:

$$G = p(1 - p), \quad (6.3)$$

where purity  $p$  is the ratio of the number of signal events over the number of all events:

$$p = \frac{S}{S + B}. \quad (6.4)$$

The best separating variable and cut value is defined to be the cut that maximizes the difference of the Gini-Index between the parent node and the sum of the two daughter nodes weighted by their event fractions.

After finding the optimal cut, the event samples are split in two subsamples that either pass or fail the cut criteria of the node and proceed down into the daughter nodes. The next nodes then find the best separating variables to split the subsamples further. This splitting is done until a stop criteria is reached, like the maximal depth of a tree or the minimum number of events that should pass a node. At the end of the node structure are leaves that assign events that pass the leaf either -1, if the leaf contains mostly background, or +1, if the leaf is dominated by signal.

Decision trees trained on a training data set are able to learn general features of the data which can be used to discriminate signal and background. However, it is possible that the decision trees pay attention to features that only exist in the training data. The focus of non-general features can lead to a lower separation performance on new data which was not included in the training. This loss in generality of classifiers is called overtraining and can be disadvantageous and diminish the separation power.

Because single decision trees often have low separation power and are prone to overtraining, multiple trees are trained and combined with a boosting algorithm called Adaptive Boosting (AdaBoost) to form a single BDT classifier [92]. This procedure allows to combine shallow decision trees with a depth of 2 or 3 nodes, that perform often only slightly better than random guessing, into one strong classifier with highly enhanced separation power. Given very small trees as weak learners, the overtraining of the combined BDT is strongly suppressed leading to a more generalized classifier. The AdaBoost algorithm takes the initial decision tree classifier  $h_0$  and calculates the boosting weight  $\alpha$  defined as:

$$\alpha = \frac{1 - err}{err}, \quad (6.5)$$

where  $err$  is the misidentification rate of  $h_0$  in the training set. The next tree  $h_1$  is then grown on the training data set, where the misidentified events are weighted by the common boost weight  $\alpha$ . The entire training set is renormalized so that the sum of weights remains constant. This reweighting procedure allows the next tree to focus on events that are harder to distinguish between signal and background. This is done successively for hundreds of trees, forming a decision forest. The responses of the grown trees are finally combined into a single classifier response  $H_{\text{BDT}}$  using the boosting equation:

$$H_{\text{BDT}}(y) = \frac{1}{N_{\text{trees}}} \times \sum_{i=0}^{N_{\text{trees}}} \ln(\alpha_i) h_i(y), \quad (6.6)$$

where  $N_{\text{trees}}$  denotes the number of trees grown and  $h_i(y)$  the response of the  $i$ -th tree given the event  $y$ . For each event the BDT returns a response between -1 and +1 where -1 corresponds to background like and +1 to signal like events. To further enhance training results, a learning rate parameter  $\beta$  is introduced that modifies the boosting weight:

$$\alpha \rightarrow \alpha^\beta. \quad (6.7)$$

Using the parameter  $\beta$ , the speed in which the classifier adopts to the training set can be adjusted.

To decide if one BDT performs better on the test sample than another, the Receiver-Operating-Characteristic-Curve (ROC-Curve) is used as a measure. The ROC-Curves are constructed by plotting the inverse of the background efficiency against the signal efficiency. For the BDT distributions  $\mathcal{D}(x)_B$  and  $\mathcal{D}(x)_S$  for background and signal which are bound between  $x \in [-1, 1]$ , the efficiency  $Eff$  in dependence of the BDT response score threshold  $x$  is defined by:

$$Eff_{B,S} = \frac{\int_x^1 \mathcal{D}(x')_{B,S} dx'}{\int_{-1}^1 \mathcal{D}(x')_{B,S} dx'}. \quad (6.8)$$

The software framework used to configure and train BDTs in this analysis is delivered by TMVA (Toolkit for Multivariate Analysis) [93].

## 6.2 Training setup for TMVA

To improve the performance of the BDTs that are finally used to separate background from signal, the training and application strategy has to be chosen carefully. For each mass point of the signal masses 400 GeV, 500 GeV, 600 GeV, 700 GeV, 800 GeV,

1000 GeV and 1200 GeV separate BDTs are trained to let each BDT optimize to individual features in the signal event samples and thus enhancing separation power. The only exception are the 300 GeV and 350 GeV signal mass samples which are covered by one BDT to enhance the statistics of the training sets. Additionally individual BDTs are trained for the B-TAG and B-VETO categories to benefit from the topological differences of the dominant signal and background processes. This results in a total amount of 16 BDTs which have to be individually set up and validated.

For all BDTs the two leading background contributions, that are multi-jet for both categories and  $Z \rightarrow \tau\tau$  (top) for B-VETO (B-TAG), are used as background to train against. The signal samples ggH and bbH are both used in B-VETO BDTs while in B-TAG only the dominating bbH signal is fed into the training. The configuration of the BDTs is illustrated in Figure 6.2.1.

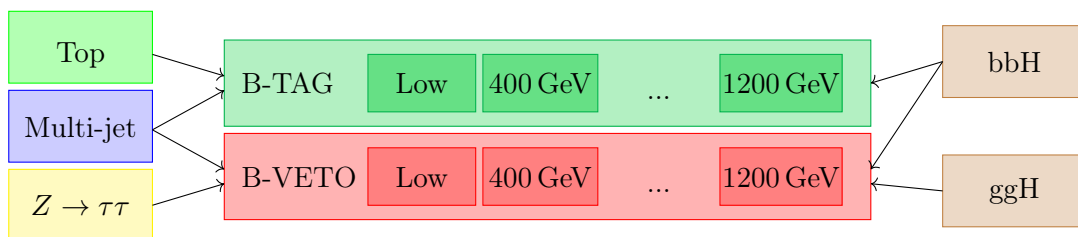


Figure 6.2.1: Background and signal configuration of the BDTs. For each mass point in both categories BDTs are trained (middle) using major contributing signal samples (right) and the two most important backgrounds of the regions (left). The "Low" labeled BDT is trained on the combined 300 GeV and 350 GeV signal event samples.

Variables used for the training are mostly simple kinematic quantities. The background and signal distributions of the variables in the signal region are compared in Figure 6.2.2 and in Appendix A.2.1, A.2.2, A.2.3, A.2.4 and are described hereinafter:

**Total transverse mass:**  $m_T^{\text{tot}}$

The total transverse mass as defined in (6.1) of the di-tau system is one of the best separating variables especially for high signal masses. It is strongly correlated with the Higgs boson mass and gets shifted to higher values for heavier Higgs particles, thus separating background and signal. It is the main discriminant variable used in Ref. [87] whose results are compared to the results of this thesis.

**Absolute value of the pseudorapidity of the leading tau:**  $|\eta_{\text{leading } \tau}|$

The pseudorapidity distribution of the leading  $\tau_{\text{had, vis}}$  is a major discriminating variable. It is especially powerful to separate multi-jet from signal, as the background shows a uniform  $|\eta_{\text{leading } \tau}|$  distribution while bbH and ggH signals peak at 0.

**Missing transverse energy:**  $E_{T, \text{mis}}$

As mentioned in Chapter 3.2.1,  $E_{T, \text{mis}}$  is calculated by the vectorial sum of  $p_T$  from reconstructed electrons, muons, tau leptons and jets using informations from calorimeter deposits and tracks [52].

Although this variable is strongly correlated with  $m_{\text{T}}^{\text{tot}}$  at high signal masses since it is used in the calculation, it delivers extended background separation power for lower mass signals.

**Sum of visible tau masses:**  $\sum m_{\tau, \text{vis}}$

This variable is defined as the sum of the invariant masses calculated from the visible hadronic decay products of the leading and subleading  $\tau_{\text{had, vis}}$ . Like  $E_{\text{T, mis}}$  it is strongly correlated with  $m_{\text{T}}^{\text{tot}}$  as similar variables are used for its calculation. However, it has been shown to increase the performance for low mass signals.

**Transverse momentum balance:**  $p_{\text{T, balance}}$

This variable is defined as the  $p_{\text{T}}$  ratio of the two tau lepton candidates:

$$p_{\text{T, balance}} = \frac{p_{\text{T}}^{\text{leading } \tau}}{p_{\text{T}}^{\text{subleading } \tau}}. \quad (6.9)$$

Using the ratio allows the BDT to select different  $p_{\text{T}}$  regions and subsequently increasing the performance.

**Angular separation in the transverse plane:**  $\Delta\phi$

A cut on  $\Delta\phi$  between the tau leptons is already applied in the event selection. This limits the range of this variable to  $\pi \geq \Delta\phi > 2.7$ . However, the additional degree of freedom improves the performance of the BDT.

**Number of charged tracks:**  $N_{\text{track, leading } \tau}$

The number of tracks that are associated with the leading tau candidate is used to enable the BDT to see differences between tau decays with different charge multiplicities.

**Logarithm of jet  $p_{\text{T}}$ :**  $\ln(p_{\text{T, jet}})$

The BDT performance is further boosted by enabling access to jet  $p_{\text{T}}$  information of jets that are potentially present. The variable is defined by:

$$\ln(p_{\text{T, jet}}) = \begin{cases} \ln(p_{\text{T, leading jet}}) & , \text{ if at least one jet is present} \\ 0 & , \text{ if no jet is present} \end{cases}. \quad (6.10)$$

**Number of b-jets:**  $N_{\text{b-jet}}$

The information about the number of jets that are tagged as b-jets significantly increases the performance of the BDT.

**Transverse mass of di-jet system:**  $m_{\text{T, di-jet}}$

This quantity is used to discriminate the b-associated Higgs production signal from multi-jet background by exploiting the special topology of the additional b-jets in the signal process. The transverse mass is only calculated if at least two jets are reconstructed, otherwise it is set to zero. The numeric value is calculated in analogy to equation (6.2):

$$m_{\text{T, di-jet}} = \sqrt{2 \cdot p_{\text{T, jet}_1} \cdot p_{\text{T, jet}_2} \cdot [1 - \cos(\Delta\phi(\text{jet}_1, \text{jet}_2))]}, \quad (6.11)$$

where  $\text{jet}_1$  is the leading jet with the highest and  $\text{jet}_2$  the subleading jet with the second highest  $p_T$ .

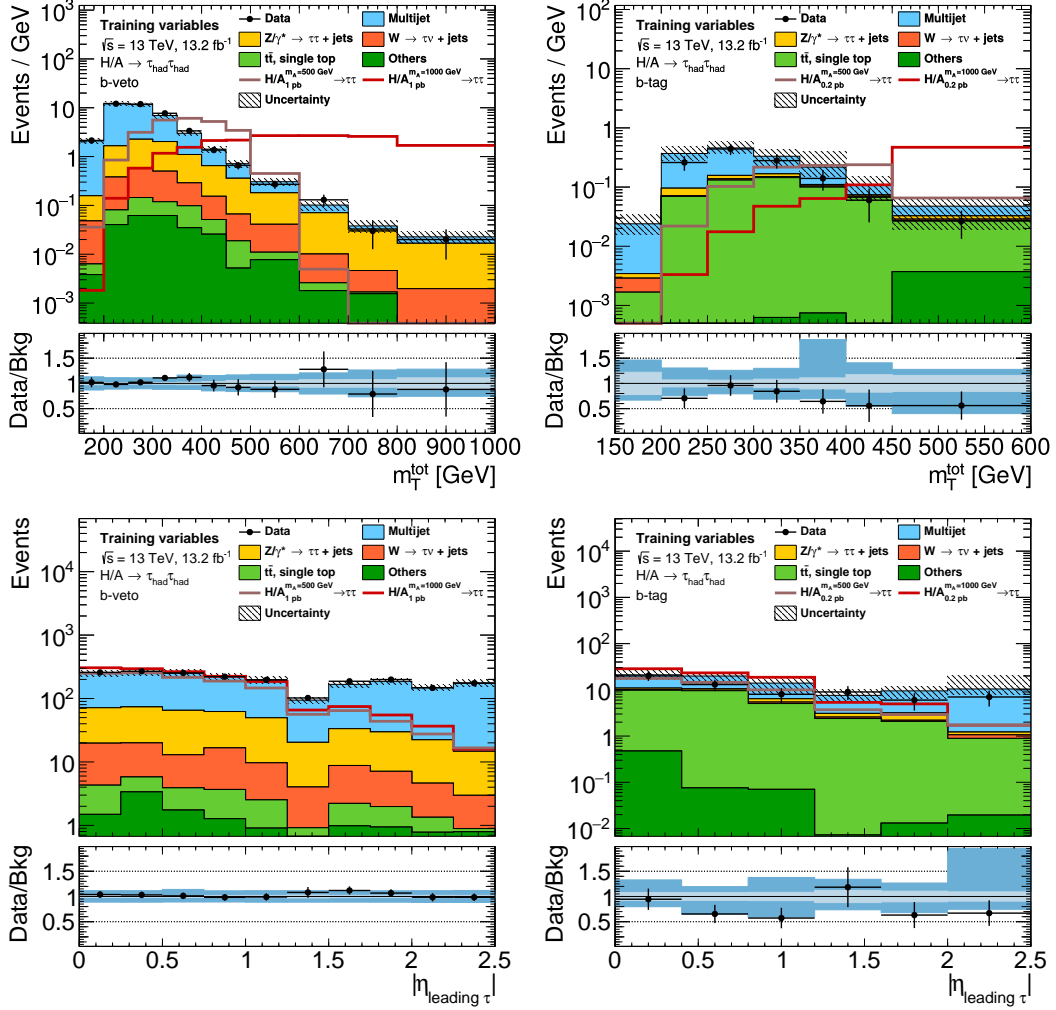


Figure 6.2.2: Signal region distribution of the main separating variables  $m_T^{\text{tot}}$  and  $|\eta_{\text{leading } \tau}|$ . The 1000 and 500 GeV Higgs boson cross section is scaled to 1 pb in B-VETO (left) and 0.2 pb in B-TAG (right) to allow better comparison between the signal and background distributions.

Most variables are used for all BDTs with exception of  $N_{\text{b-jet}}$  and  $m_{T,\text{di-jet}}$  that are not implemented for B-VETO type BDTs, since they are either trivial or do not hold significant separation power. A summary of all variables and when they are used can be found in Table 6.2.1.

The background modeling of all variables used for the training are validated in the same-sign validation region in B-TAG and B-VETO with data from the ATLAS detector. This is done to ensure that the background is modeled correctly and does not bias the signal region results. As seen in Figure 6.2.3 for the main separating variables and A.1.1, A.1.2, A.1.3 as well as A.1.4 for the other variables, the background modeling of the variable distributions are in agreement with the measured data in the same-sign control region, thus no background mismodeling is observed.



Variable	B-VETO	B-TAG
$m_{\text{T}}^{\text{tot}}$	✓	✓
$ \eta_{\text{leading } \tau} $	✓	✓
$E_{\text{T, mis}}$	✓	✓
$\sum m_{\tau, \text{vis}}$	✓	✓
$p_{\text{T, balance}}$	✓	✓
$\Delta\phi$	✓	✓
$N_{\text{track, leading } \tau}$	✓	✓
$\ln(p_{\text{T, jet}})$	✓	✓
$N_{\text{b-jet}}$		✓
$m_{\text{T, di-jet}}$		✓

Table 6.2.1: List of variables used for the training of B-TAG and B-VETO type BDTs.

As already mentioned, some of the variables are directly correlated by the way they are defined. However, as the decision trees are trained by selecting only the best separating variables at each training iteration, the resulting boosted classifier is very resilient against correlation and redundant information. Instead, the BDT quite often profits if more variables are used.

To minimize bias through overtraining in the BDT discriminant, events that are used for training can not be used by the same BDT in the subsequent analysis. This has to be taken into account when choosing a strategy to split the event samples into training and testing data sets. It requires a careful balancing of distributing available data. Too low training statistics results in weak classifiers while too much decreases statistics available for further analysis, eventually rendering the gained performance of the BDT useless.

In this thesis two different splitting strategies are presented, the classic splitting model and the K-Fold cross validation model [94].

### 6.2.1 The classic splitting model

In the classic splitting model, the available background and signal datasets are split into training and testing subsets. After training, the BDT is validated with the testing data set which is also used for further statistical analysis, while the training set has to be discarded.

As the multi-jet background is estimated using data from the fail-ID region weighted by fake factors, the number of events for this background is very limited, especially in the B-TAG region. A representative subsample corresponding to  $3 \text{ fb}^{-1}$  of the fail-ID region is chosen for training. Events for training are selected such that the fraction of the two high level triggers that can trigger the event are the same as in the full data set and are evenly distributed over the 2015 and 2016 data taking periods. These

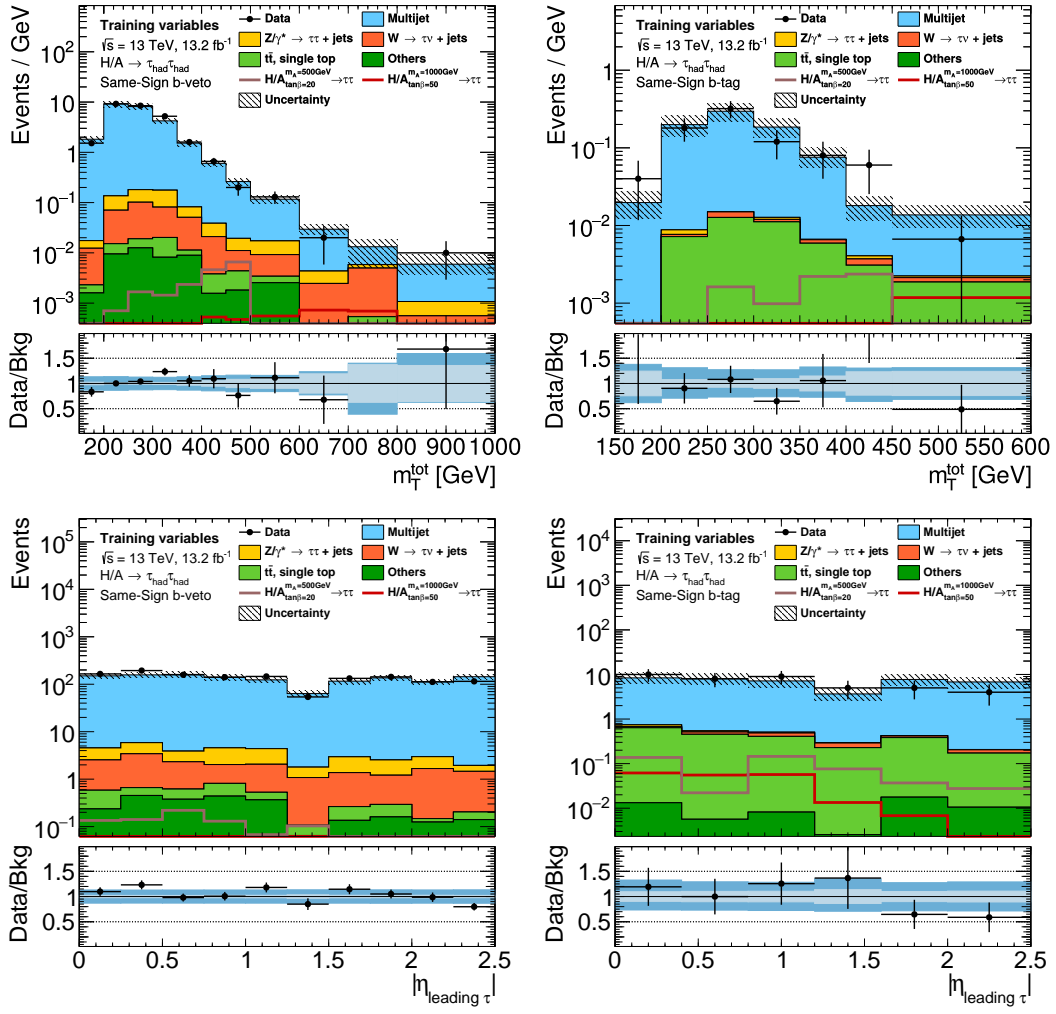


Figure 6.2.3: Same-sign distributions of the main separating training variables  $m_T^{\text{tot}}$  and  $|\eta_{\text{leading } \tau}|$ .

provisions are made to minimize artificial differences in pileup or  $p_T$  distributions between the training and test event sets that might impact the performance of the BDT negatively. The remaining  $10.2 \text{ fb}^{-1}$  of data are used for testing. Because the fail-ID region is orthogonal to the signal region, the signal region data is unaffected by this selection and the full  $13.2 \text{ fb}^{-1}$  can be used in further statistical analysis.

The signal samples, top and  $Z \rightarrow \tau\tau$  samples are produced by Monte Carlo simulation and thus have significantly more statistics than the data driven multi-jet background. This allows to split the Monte Carlo samples in half using events with odd event number for testing and with even event number for training. Event numbers in simulated event samples are given in increasing order by the Monte Carlo generator to individual events. Because the pile-up changes with respect to MC event numbers, selecting odd and even event numbers ensures the same pile-up profile for testing and training sets. Special measures have been taken for the bbH sample. As pointed out in Chapter 4, the bbH samples have a significant amount of events with negative event weights. To study the impact of the event weights, all distributions of the variables used to train the BDTs are compared using correct event weights and the absolute value of

the weights. The distributions for  $m_T^{\text{tot}}$ ,  $E_{T,\text{mis}}$  and  $|\eta_{\text{leading } \tau}|$  of the 1000 GeV bbH sample are shown in Figure 6.2.4, as they are the most important training variables. The remaining distributions for the 1000 GeV bbH sample are listed in appendix Figure A.4.2 for B-TAG and A.4.1 for B-VETO.

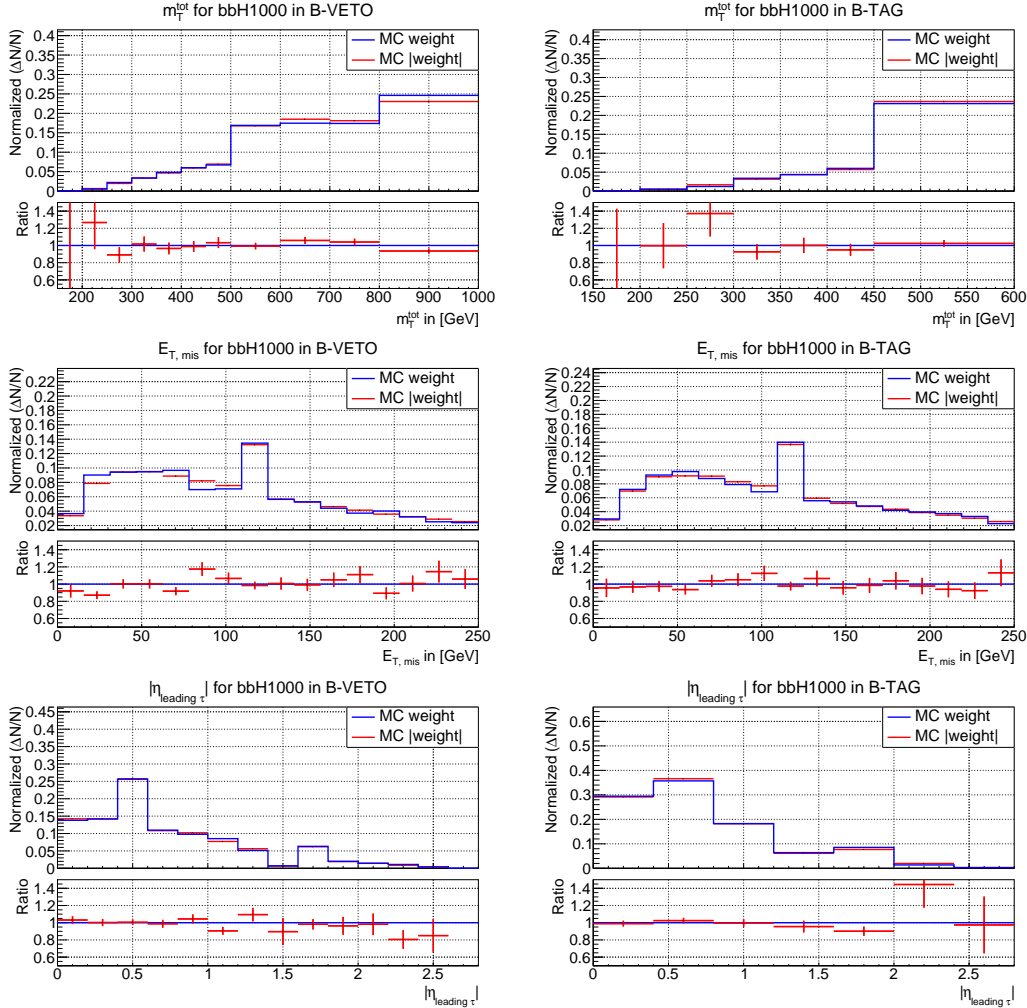


Figure 6.2.4: Shape of the distribution for the 1000 GeV bbH sample in B-VETO (left) and B-TAG (right). The distributions are normalized to unity and the shapes of the distributions are compared using the absolute value of the MC weight (red) and the nominal weight (blue).

Only slight differences in the shape of the distributions weighted by the absolute and nominal value of the event weights are observed. Since these small differences might only influence the performance slightly if the overall distribution shape is conserved, the absolute value of the event weights in the bbH samples are used for training <sup>1</sup>. This increases the training statistics for the signal samples significantly which results in higher stability and performance of the classifier.

In the training set the MC backgrounds are scaled such that they match the  $3 \text{ fb}^{-1}$  of luminosity from the multi-jet background. The signal samples are scaled to the same cross section of 1 pb. Because background and signal distributions are normalized

<sup>1</sup>The absolute value of the Monte Carlo weights are only applied in the BDT training. In further analysis the correct weight treatment is used.

before training, only the relative scale between the subprocesses is relevant. The TMVA framework allows detailed configuration of the BDT training. By varying the number of trees (`nTrees`), their maximum depth (`MaxDepth`) and the minimum fraction of events that should pass a node (`MinNodeSize`), different separation performance and overtraining is achieved. The configuration of each BDT is chosen such that the BDT shows a minimal amount of overtraining. Under this premise, the BDT configuration with the best performance is selected. By prioritizing minimal overtraining, the classifier has a more stable response and is less prone to fluctuations of characteristics in the test samples and data of the signal region. Configurations proven to yield good performance while showing little to no overtraining have usually 100 to 300 shallow trees with a maximum depth between 2 and 4. More trees or deeper trees often result in significantly higher overtraining, while for far less or shallower trees, performance breaks down significantly. The training configurations are listed in Table A.3.1 and A.3.2 of the Appendix.

### 6.2.2 The K-Fold cross validation model

Another approach to split the training and test sets is the so called K-Fold cross validation model. Here the entire data set is split into subsets of equal size, including the multi-jet sample. Individual BDTs are trained on each subsets which are tested on the remaining sets. This method is known as K-Fold cross validation, where  $K = 2$  corresponds to the number of independent sub-samples derived from the original background and signal set. In the case of this analysis, two subsets are created by using odd and even event numbers as splitting criterion to ensure that the subsets are representative. One BDT is trained on the even event number and applied to the odd event number subset, further referred to as  $KF_O$ , while the other one is trained and validated vice versa and called  $KF_E$ . On data from the signal region, the combined BDT response is calculated using the arithmetic mean of the individual BDTs. This procedure doubles the amount of BDTs needed to 32. However, the statistics of multi-jet background available for training is more than twice the amount used in the classic splitting model, since  $6.2 \text{ fb}^{-1}$  of data from the fail-ID region can be used for training. This allows the classifier to better resolve general features of the background. In addition, by combining the two BDTs, the entire background and signal samples can be used in further analysis, increasing the available sample size significantly.

Like in the classic splitting model, the training configuration is optimized by selecting configurations that yield the best performance and minimize overtraining. Although the multi-jet background training sample size is significantly increased, the best configurations are found using similar configurations as in the classic splitting model. The configuration in TMVA is listed in Table A.3.1 and A.3.3. The set of variables used to train the BDTs is identical to the classic splitting model and also for the  $bbH$  signal sample the absolute value of Monte Carlo generator weights are used.

## 6.3 Training results and validation

After training the BDTs, the classifiers are checked for overtraining by comparing the BDT response between the training and test data set. If the response distributions for testing and training deviate from one another and the BDT performs better on the training set, then the BDT is overtrained. To minimize overtraining is crucial, since heavily overtrained BDTs tend to have lower average performance and often have less

stable response which can impact the background modeling negatively. To ensure that the background distribution of the BDT response is modeled correctly, the response distributions are validated in the same-sign validation region.

To compare the performance of the BDTs, ROC-Curves are used. Classifiers with higher inverse background efficiency at a given signal efficiency perform better in separating signal from background.

The validation procedure has to be done for each individual BDT. To illustrate the procedure, validation plots for the 500 GeV and 1000 GeV signal mass BDT are shown in this Section. The validation plots for the other BDTs can be found in the Appendix A.

### 6.3.1 The classic splitting model

For checking whether a classifier is overtrained, the BDT response distributions for the testing and training set are compared for the 500 GeV and 1000 GeV BDT in Figure 6.3.1.

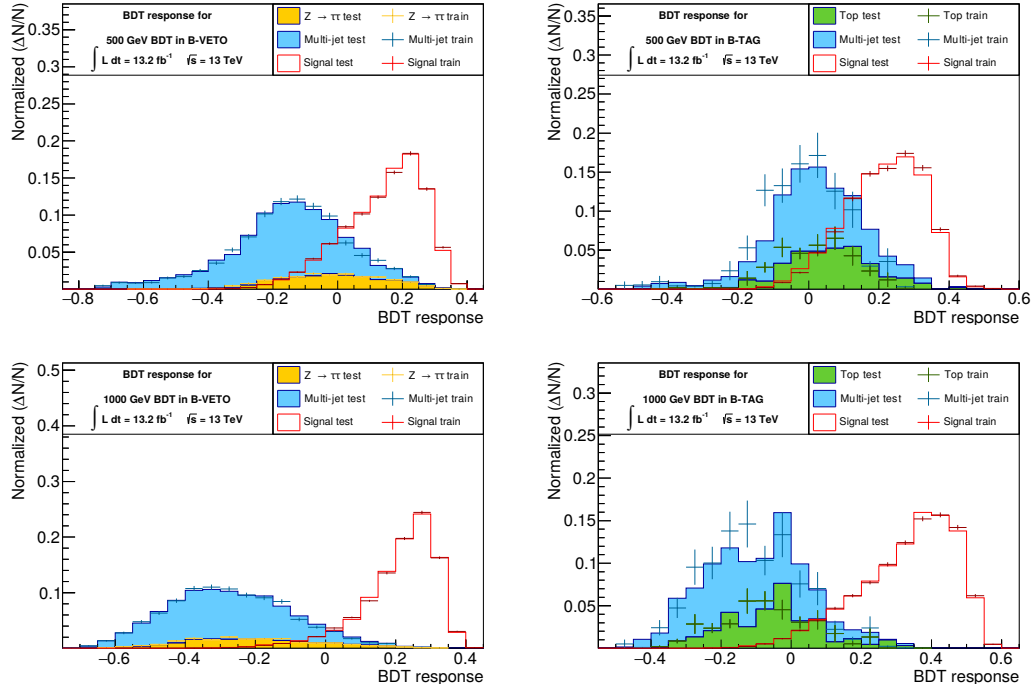


Figure 6.3.1: Normalized response distributions of the 500 GeV (top) and 1000 GeV (bottom) signal mass BDT in B-VETO (left) and B-TAG (right). The classifier show no significant overtraining.

The training configurations yield low overtraining of the classifier, especially considering the low statistics in the B-TAG region. Notable is the separation power of the BDT against multi-jet background, while top decay and  $Z \rightarrow \tau\tau$  background is still shifted towards the signal distributions.

Variable correlations yield important information about the redundancy of variables, as strongly correlated variables hold almost identical separation information. However, since individual BDTs are trained for different signal mass hypotheses, the correlation between variables of different signal samples can vary significantly. The variable cor-

relations for the 500 GeV and 1000 GeV signal sample in Figure 6.3.2 clearly show a strong negative correlation between total transverse mass and the absolute value of the pseudorapidity of  $\tau_{\text{had,vis}}$  for high signal masses, while for lower masses the variables are less correlated. Another correlated set of variables are  $m_{\text{T}}^{\text{tot}}$ ,  $E_{\text{T,mis}}$  and  $\sum m_{\tau,\text{vis}}$ ,

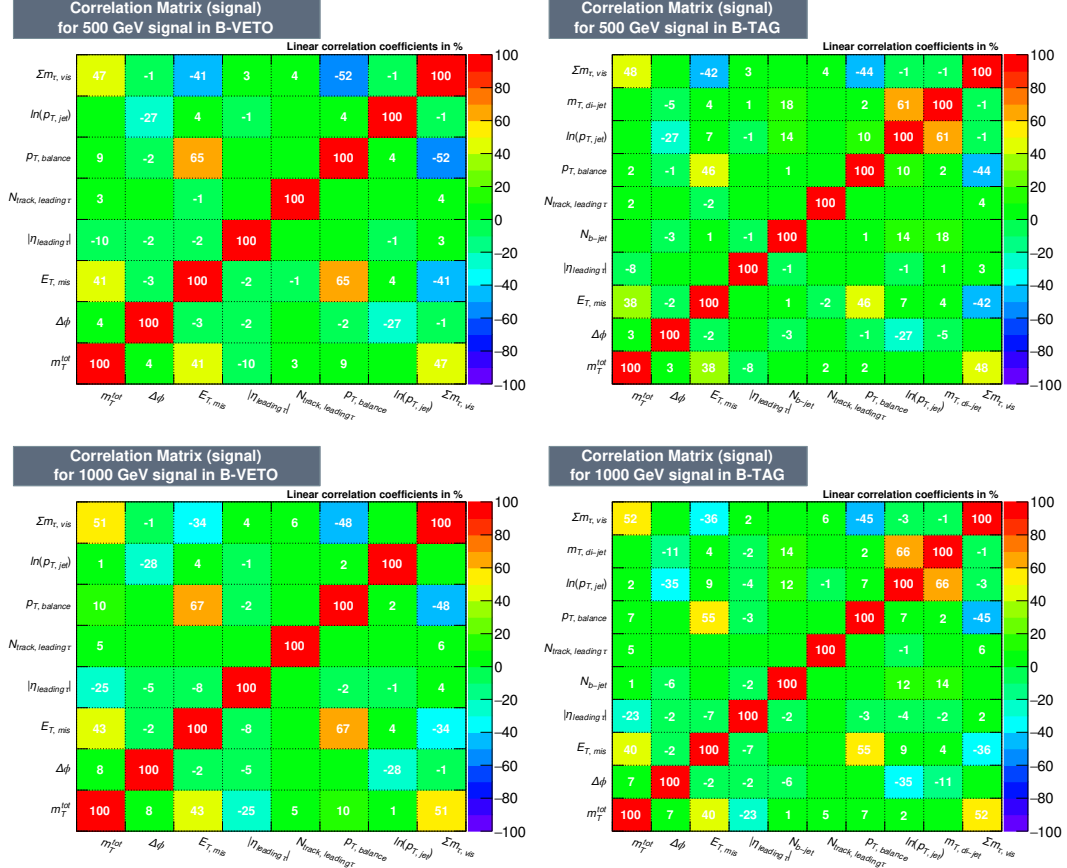


Figure 6.3.2: Signal correlation of variables used for 500 GeV (top) and 1000 GeV (bottom) signal mass BDTs that are trained using the classic splitting model.

as the variables are either very similar or directly used in the equation to calculate another variable. Using these variables independently can increase the performance of the classifiers, because for different signal mass BDTs one of the variables might be more suitable to separate background and signal than the others.

The correlations of variables for the backgrounds shown in Figure 6.3.3 are identical for all BDTs since the same background sets are used for training and testing. Some of the obvious correlations between variables like  $m_{\text{T}}^{\text{tot}}$  and  $E_{\text{T,mis}}$  are also present in the background sets, while others, such as  $p_{\text{T,balance}}$  and  $\sum m_{\tau,\text{vis}}$ , are far less correlated. This is mostly due to the vastly different phase space of the signal and the dominant multi-jet background.

As shown in Figure 6.3.4, the background suppression is strongly enhanced for higher signal masses. This behavior is expected because major separation variables like  $m_{\text{T}}^{\text{tot}}$  or  $\sum m_{\tau,\text{vis}}$  strongly correlate with the signal mass. For higher masses the background and signal distributions of these variables deviate stronger from one another and thus signal and background are easier to separate.

The validation of the background modeling of the BDTs can be seen in the Figures

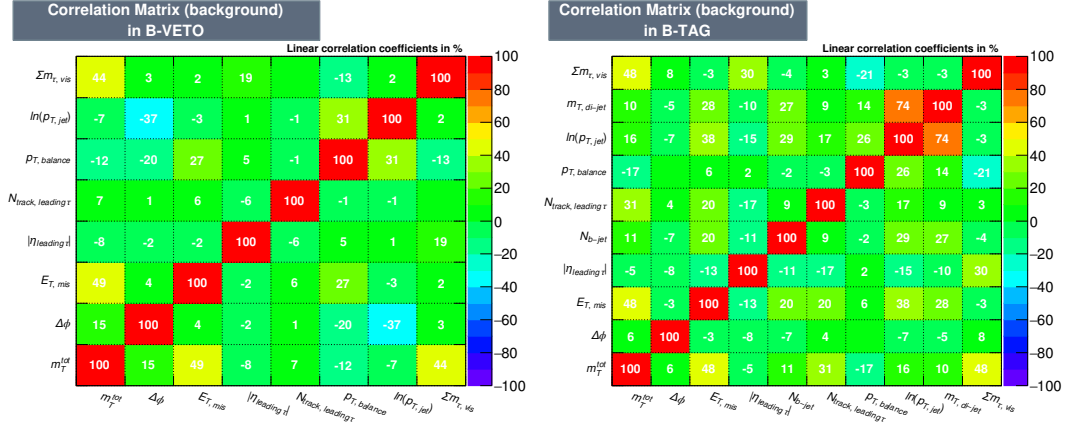


Figure 6.3.3: Background correlation of training variables.

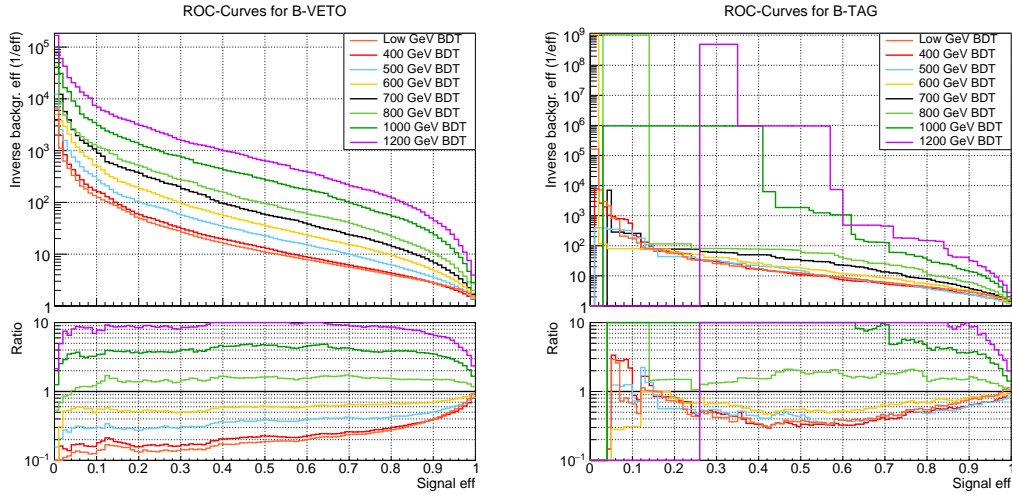


Figure 6.3.4: ROC-Curve for BDTs trained on the classic splitting model in B-VETO (left) and B-TAG (right)

6.3.5, A.5.1, A.5.2 and A.5.3, showing that the data and background distributions in the same-sign region of the BDTs are in very good alignment.

For statistical analysis to determine if a significant signal signature is found, the signal region distributions are used. For the 500 GeV and 1000 GeV signal mass BDT these can be found in Figure 6.3.6, while the others are listed in Figure A.6.1, A.6.2 and A.6.3.

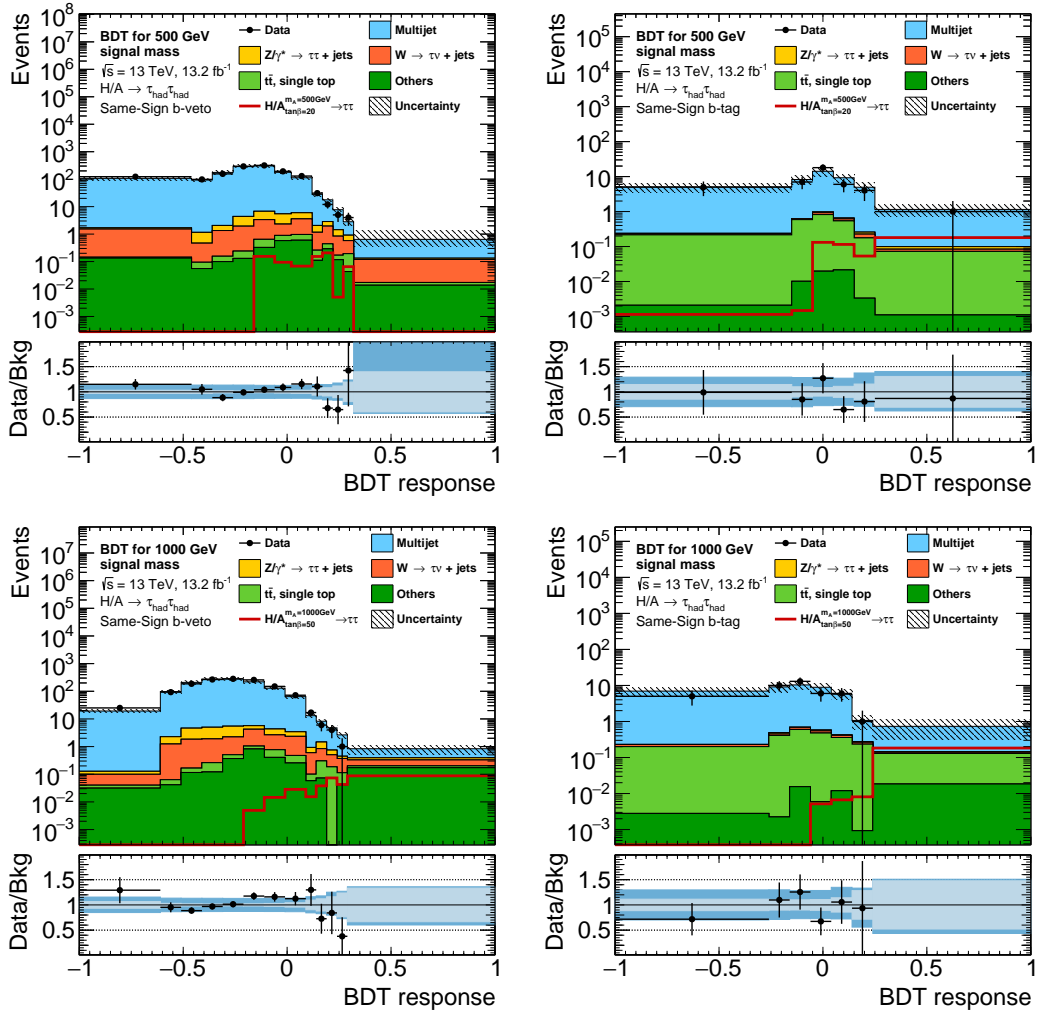


Figure 6.3.5: BDT response for 500 GeV (top) and 1000 GeV (bottom) signal mass BDTs trained with the classic splitting in B-VETO (left) and B-TAG (right) in the same-sign validation region. No significant deviation between background estimation and data distributions are observed.



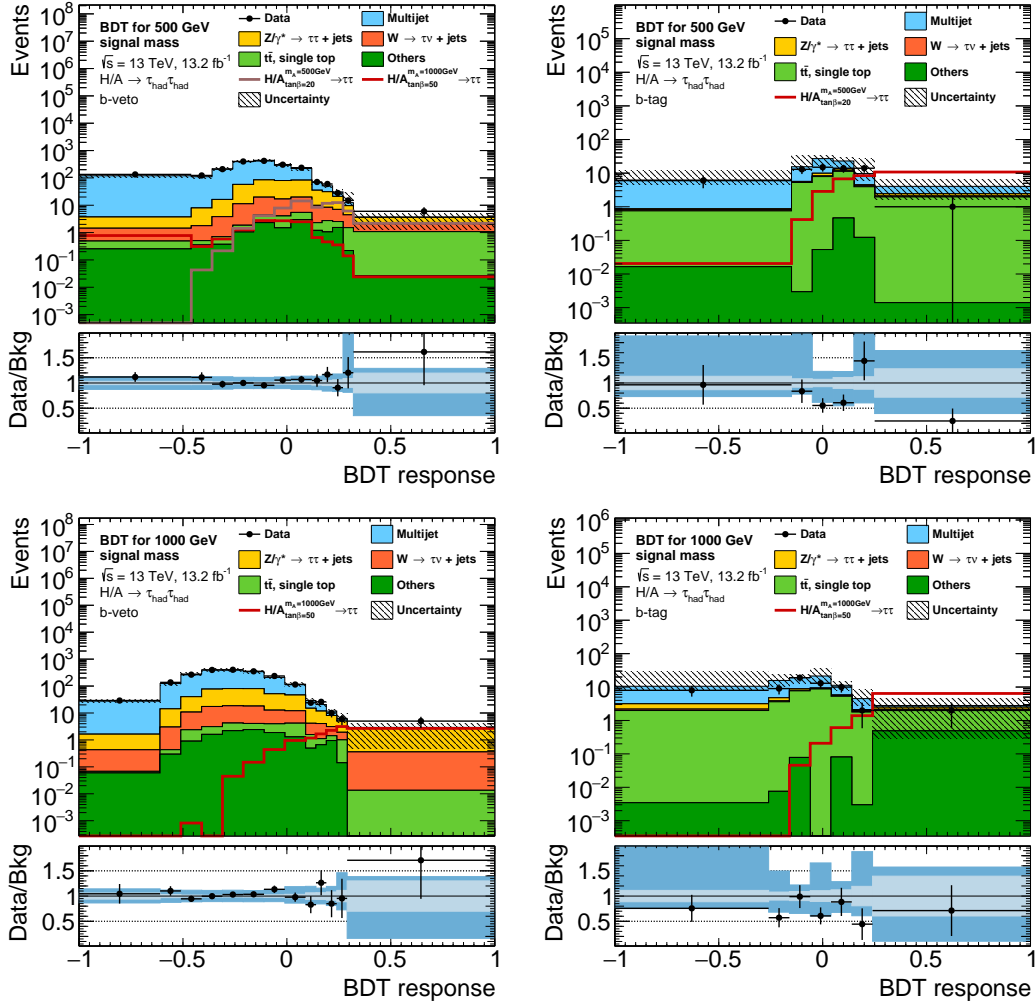


Figure 6.3.6: Signal region BDT response for 500 GeV (top) and 1000 GeV (bottom) signal mass BDTs trained using the classic splitting method.

### 6.3.2 The K-Fold cross validation model

Since the K-Fold cross validation model uses two individual BDTs for each mass point, in addition to minimize overtraining and maximize separation performance, it is also crucial that the BDTs behave similarly. That means, they select similar phase space in the same detector region and have almost the same response distribution shape on the test set. To check shape deviations the response distributions for the 500 GeV and 1000 GeV signal mass BDTs are shown in Figure 6.3.7 and 6.3.8.

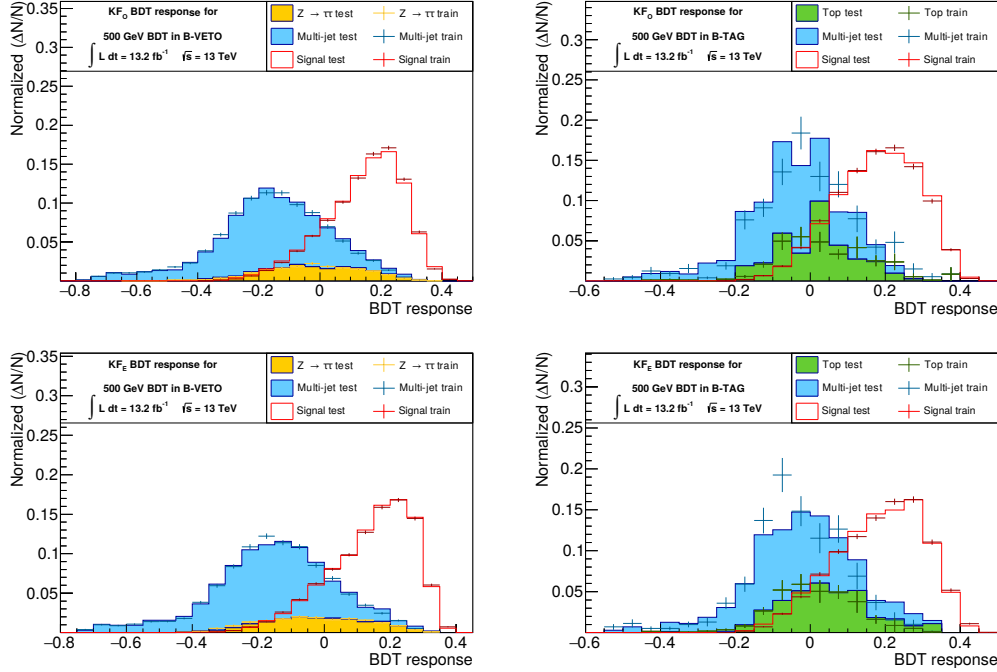


Figure 6.3.7: Normalized response distributions of the 500 GeV signal mass BDTs trained with K-Fold method. The  $KF_O$  BDTs (top) are compared with the  $KF_E$  BDTs (bottom). Both classifiers have similar response distributions.

The shape of the response for lower mass  $KF_O$  and  $KF_E$  BDTs are similar, while larger deviations are observed towards high signal masses. A more peak like structure of one BDT response can overlap with the broader background distribution of the other one, as seen in the 1000 GeV signal mass classifier. Thus the performance of the better performing BDT can be compromised by the higher background contamination of the weaker BDT resulting in overall weaker combined classifier. The shape difference in the response distributions might be due to the lack of background data available for training so that both BDTs can not converge to select the same phase space.

Similar to the classic splitting model BDTs, the classifiers perform well in separating multi-jet background from signal, while significant proportions of the top and  $Z$  decay background is classified as more signal like.

The background modeling of the BDTs are validated in the same-sign validation regions where the  $KF_O$  and  $KF_E$  BDTs are combined on data. The plots in Figure 6.3.9, A.5.4, A.5.2 and A.5.3 show no significant deviation between background modeling and data distributions. The response distributions in signal region that are used in the statistical analysis are shown in Figure 6.3.10, A.6.4, A.6.5 and A.6.6.

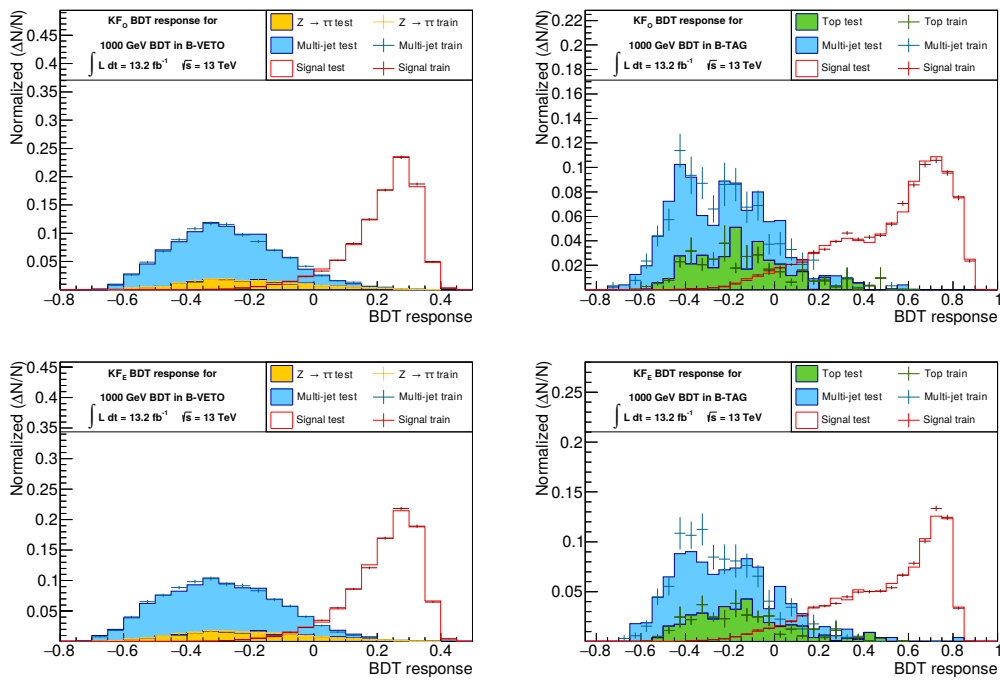


Figure 6.3.8: Normalized response distributions of the 1000 GeV signal mass BDTs train with K-Fold method. The  $KF_O$  BDTs (top) are compared with the  $KF_E$  BDTs (bottom). Slight differences of the signal shape distributions between the  $KF_O$  and  $KF_E$  BDTs in the B-TAG region are observed.

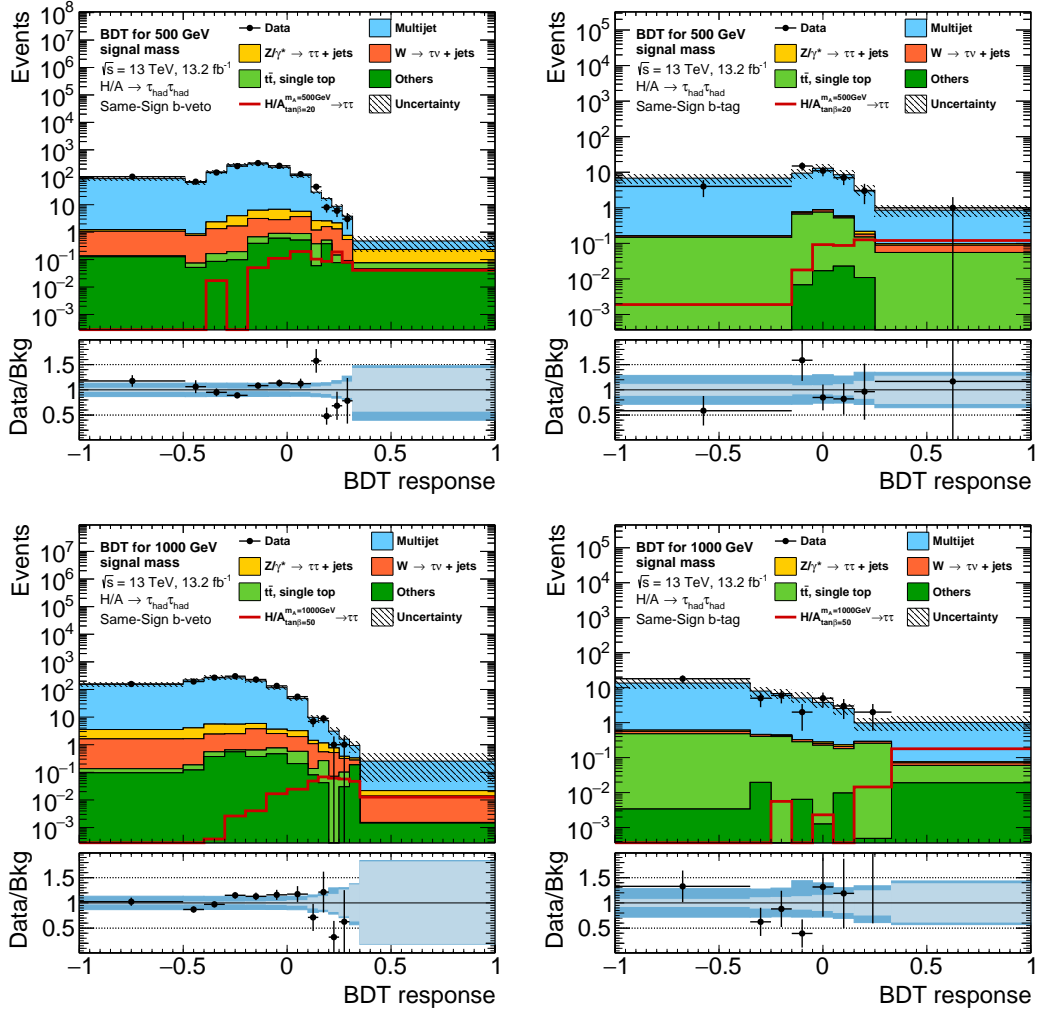


Figure 6.3.9: BDT response for 500 GeV (top) and 1000 GeV (bottom) signal mass BDTs trained using the K-Fold method in B-VETO (left) and B-TAG (right) in the same-sign validation region. No significant deviations between background estimation and data distributions are observed.

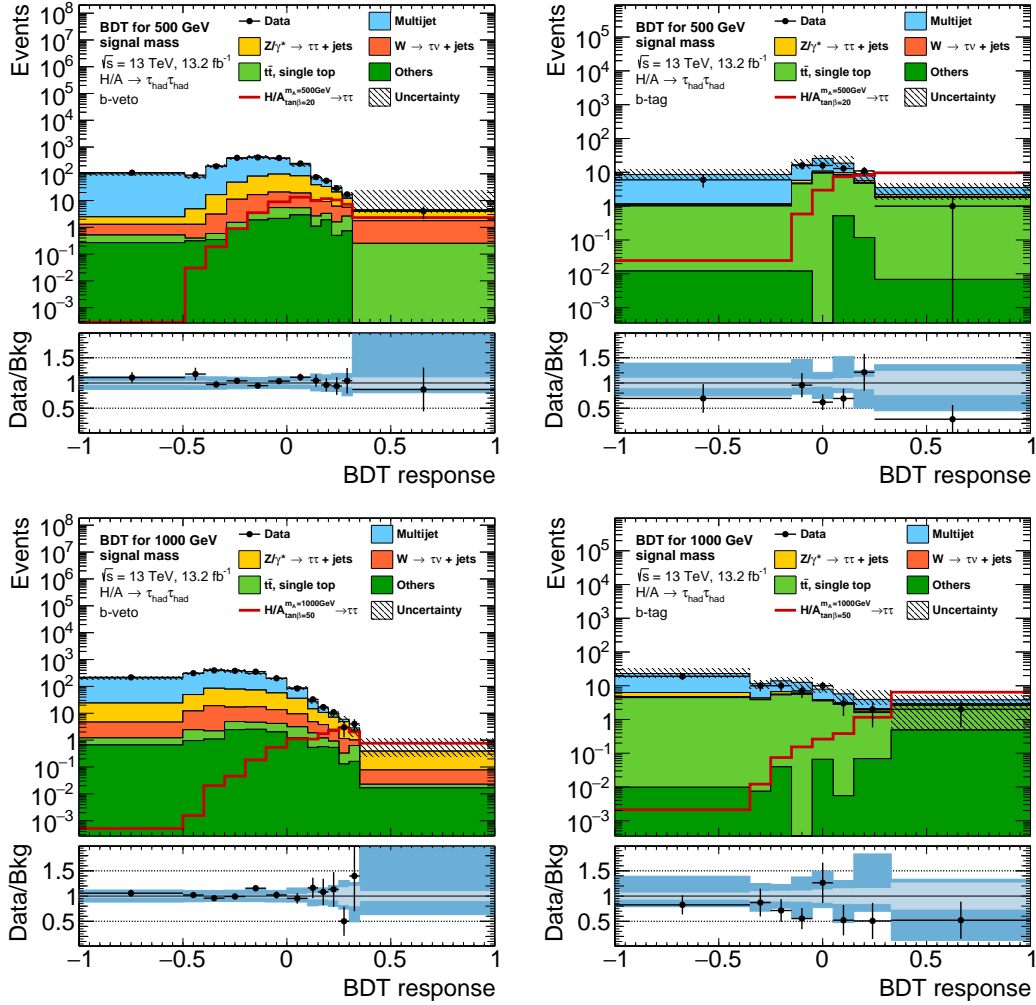


Figure 6.3.10: Signal region BDT response for 500 GeV (top) and 1000 GeV (bottom) signal mass BDT trained using the K-Fold method.



## 7 Limit setting

Since no significant deviation between data and the SM background estimation is observed in the signal region, the classifier distributions presented in Section 6 can be used to set limits in the parameter space of the MSSM, excluding certain areas in the  $m_A$ - $\tan(\beta)$  plane. In this section the theory of limit calculations is briefly discussed and the set up of the limit machinery is described.

### 7.1 Theory of limit setting

The calculation of exclusion limits on a signal hypothesis is done using a profile likelihood ratio test. To define the nominal signal hypothesis  $H_1$  and the background hypothesis  $H_0$  that is tested against, a parameter of interest  $\mu$  is introduced which acts like a signal strength scale factor. Let  $\mathcal{D}$  be a distribution with  $M$  number of bins and  $\{x_1, \dots, x_M\}$  the measured content of the bins containing observed data  $n$ , background  $b$  and signal  $s$  yields. Then the number of expected events in the  $i$ -th bin is:

$$E_\mu[n_i] = \mu s_i + b_i, \quad (7.1)$$

with  $E_\mu[n_i]$  as the expected number of events given a signal strength  $\mu$ . The background hypothesis is defined by setting  $\mu = 0$ , while  $\mu = 1$  corresponds to the nominal signal hypothesis.

The likelihood function is calculated as the product of the probability density functions of the individual bins in  $\mathcal{D}$  where for each bin a Poisson probability distribution is assumed [95]:

$$L_{\mathcal{D}}(\mu) = \mathcal{N}_c \prod_{i=1}^M \text{Pois}(n_{i,\text{obs}} | \mu s_i + b_i) = \prod_{i=1}^M \text{Pois}(n_{i,\text{obs}} | E_\mu[n_i]), \quad (7.2)$$

with

$$\text{Pois}(m|n) = \frac{n^n}{m!} e^{-n}, \quad (7.3)$$

and  $n_{i,\text{obs}}$  denoting the measured observed data in bin  $i$ . The factor  $\mathcal{N}_c$  is a constant that can be ignored. In addition, the nominal event yields are also effected by systematical and statistical uncertainties which have to be considered in the likelihood construction.

To account for the systematic uncertainties, each uncertainty parameter  $p$  is associated with a nuisance parameters  $\alpha_p$ . For  $k$  independent and uncorrelated uncertainties, a set of nuisance parameters  $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_k\}$  can be defined. Systematic uncertainties can impact the distributions in two different ways, either as normalization uncertainty, which shifts the entire distribution by a constant factor and can be modeled by varying an overall factor for all bins, or as shape uncertainty, that changes the shape of the distribution and thus has to be modeled for each bin independently. To implement shape and normalization systematics, the constraints on the nuisance parameters and their influence on the distributions are treated separately.

The nuisance parameters  $\boldsymbol{\alpha}$  are assumed to be normally distributed and thus are constrained by Gaussian distributions around a mean value of  $m = 0$  and a standard deviation of  $\sigma = 1$ . To model the influence of normalization uncertainty for each background or signal sample, transformation functions  $\eta_{sp}(\alpha_p)$ <sup>1</sup> that depend on the nuisance parameters are constructed. Let the upper and lower bounds of relative normalization uncertainty be  $\eta_{sp}^+$  and  $\eta_{sp}^-$ , then  $\eta_{sp}(\alpha_p)$  is fitted such that the following condition is fulfilled:

$$\eta_{sp}(\alpha_p) = \begin{cases} \eta_{sp}^+ & , \alpha_p = +1 \\ 1 & , \alpha_p = 0 \\ \eta_{sp}^- & , \alpha_p = -1 \end{cases} . \quad (7.4)$$

That means a shift of  $\pm 1\sigma$  in the nuisance parameter corresponds to a scale factor shift of  $\eta_{sp}^\pm$  of the sample distribution  $s$ .

A similar approach is done to model shape systematics on every sample, but for each bin  $b$  separately. The transformation function  $\zeta_{spb}(\alpha_p)$  is fitted with the same requirements as in equation (7.4).

Statistical uncertainties are taken into account by multiplying the expected bin content of the signal or background with a factor  $\gamma_b$  that is constrained by a Poisson distribution. Using the so called Beeston-Barlow treatment [96] for statistical uncertainties, the  $\gamma_b$  are defined on the combined statistics of the background and signal samples.

The content of bin  $b$  for signal and background is then defined by including the transformation parameter  $\eta_{sp}$ ,  $\zeta_{spb}$  and the statistical nuisance factor  $\gamma_b$ :

$$b_{sb}(\boldsymbol{\alpha}, \gamma_b) = b_{sb}^{nom} \times \gamma_b \prod_{p=1}^k \eta_{sp}(\alpha_p) \zeta_{spb}(\alpha_p), \quad (7.5)$$

with the nominal bin content for background  $b_{sb}^{nom}$ . The signal yield  $s_{sb}(\boldsymbol{\alpha}, \gamma_b)$  is calculated likewise.

Using equation (7.5) and defining the entire set of nuisance parameters  $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \gamma_b)$ , the full profile likelihood function is constructed:

$$L(\mu, \boldsymbol{\theta}) = \prod_{b=1}^M \text{Pois}(n_{b, \text{obs}} | \mu s_b(\boldsymbol{\theta}) + b_b(\boldsymbol{\theta})) \times \text{Pois}(m_b | \gamma_b \tau_b) \times \prod_{p=1}^k \text{Gauss}(\alpha_p | 1, 0), \quad (7.6)$$

with the definition of a Gaussian distribution:

$$\text{Gauss}(x | \sigma, m) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}. \quad (7.7)$$

The second Poisson term is the constraint term for the statistical uncertainty parameter with the bin width  $m_b = (\frac{\delta_b}{s_b^{nom}/b_b^{nom}})^2$  and  $\tau_b = m_b^{-1}$  where  $\delta_b$  is the statistical uncertainty of the bin  $b$ .

For a statistical test that rejects the nominal signal hypothesis  $H_1$  in favor of the background hypothesis  $H_0$ , the Neyman-Pearson lemma states that the likelihood ratio is the most powerful test [97]. In the case of calculating upper limits the likelihood ratio is defined [98]:

$$\tilde{\lambda}(\mu) = \begin{cases} \frac{L(\mu, \hat{\boldsymbol{\theta}}(\mu))}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} & , \hat{\mu} \geq 0 \\ \frac{L(\mu, \hat{\boldsymbol{\theta}}(\mu))}{L(0, \hat{\boldsymbol{\theta}}(0))} & , \hat{\mu} < 0 \end{cases}, \quad (7.8)$$

<sup>1</sup>For convenience individual background and signal contributions, like  $Z \rightarrow \tau\tau$  or ggH, are called samples and entities dependent on samples are marked with subscript  $s$ .



where  $\hat{\mu}$  and  $\hat{\theta}$  are the set of signal scale factor and nuisance parameters that maximize the likelihood function, so called Maximum Likelihood Estimators (MLE), whereas  $\hat{\theta}(\mu)$  are conditional MLE given a certain  $\mu$ . In this search a hypothetical signal should only contribute with positive  $\mu$ . When observing  $\hat{\mu} < 0$  in equation (7.8), the MLE of the numerator is set to the closest possible value  $\hat{\mu} = 0$ . For the statistical analysis it is convenient to use the logarithmic likelihood ratio  $\tilde{q}_\mu$  as test statistic:

$$\tilde{q}_\mu = \begin{cases} -2 \ln \tilde{\lambda}(\mu) & , \hat{\mu} \leq \mu \\ 0 & , \hat{\mu} > \mu \end{cases}, \quad (7.9)$$

so that an increase in  $\tilde{q}_\mu$  corresponds to an increase of incompatibility of the  $\mu$  hypothesis with data. The case  $\hat{\mu} > \mu$  should not be treated as less compatible with the MLE and thus is not included when constructing the rejection intervals. Calculating the value of  $\tilde{q}_{\mu, \text{obs}}$  is now straight forward by evaluating the test statistic using the observed data distributions. The construction of the likelihood functions from the distribution histograms is done by the HistFactory framework described in Ref. [99] where more detailed information about the likelihood construction and uncertainty treatment can be found.

To quantify the probability that a particular  $\mu$  value is compatible with the observed data, the p-value can be introduced:

$$p_\mu = \int_{\tilde{q}_{\mu, \text{obs}}}^{\infty} f(\tilde{q}_\mu | \mu, \hat{\theta}(\mu, \text{obs})) d\tilde{q}_\mu, \quad (7.10)$$

with  $\tilde{q}_{\mu, \text{obs}}$  the value of test statistic using observed data. The term  $f(\tilde{q}_\mu | \mu, \hat{\theta}(\mu, \text{obs}))$  denotes the probability density of  $\tilde{q}_\mu$  under the conditional MLE of nuisance parameters that are calculated using observed data. To determinate the probability density of the test statistic, pseudo experiments, also called toy Monte Carlos, can be produced by generating random numbers for bin entries of the distributions and nuisance parameters using the underlying constraining probability functions. This however is very computationally expensive since for every statistical test, toy Monte Carlos in the order of a few million have to be produced to populate the tails of the probability function.

Alternatively, by using Wilks theorem [100] and Walds approximation [101] for logarithmic likelihood ratios, the probability distribution of  $\tilde{q}_\mu$  approaches an asymptotic formula which is independent of the nuisance parameters if the sample size is sufficiently large. Wald approximates the test statistic by [98]:

$$\tilde{q}_\mu = \begin{cases} \frac{\mu^2 - 2\mu\hat{\mu}}{\sigma^2} & , \hat{\mu} < 0 \\ \frac{(\mu - \hat{\mu})^2}{\sigma^2} & , 0 \leq \hat{\mu} \leq \mu \\ 0 & , \hat{\mu} > \mu \end{cases}, \quad (7.11)$$

where  $\hat{\mu}$  is assumed to be Gaussian distributed around the mean  $\mu'$  with standard deviation  $\sigma$ . Using this approximation, the asymptotic formula for the probability density function reads:

$$\begin{aligned} f(\tilde{q}_\mu | \mu') &= \Phi\left(\frac{\mu' - \mu}{\sigma}\right) \delta(\tilde{q}_\mu) \\ &+ \begin{cases} \frac{1}{2\sqrt{2\pi}\tilde{q}_\mu} \exp\left[-\frac{1}{2}\left(\tilde{q}_\mu - \frac{\mu - \mu'}{\sigma}\right)^2\right] & , 0 < \tilde{q}_\mu \leq \mu^2/\sigma^2, \\ \frac{1}{\sqrt{2\pi}(2\mu/\sigma)} \exp\left[-\frac{1}{2(2\mu/\sigma)^2}\left(\tilde{q}_\mu - \frac{\mu^2 - 2\mu\mu'}{\sigma^2}\right)^2\right] & , \tilde{q}_\mu > \mu^2/\sigma^2, \end{cases} \end{aligned} \quad (7.12)$$

with  $\Phi(x)$  being the cumulative function of the normal Gaussian distribution. This formula requires knowledge of the standard deviation  $\sigma$  of  $\hat{\mu}$ , which can be derived from the data set. Using the asymptotic formula for estimating the probability density of the test statistics, the computation time needed is reduced significantly. The p-value described in equation (7.10) is calculated by setting  $\mu' = \mu$  and is a measure of the statistical probability that the observed data can be described by the hypothesis that is tested. A cut-off parameter  $\alpha = 0.05$  is set so that if  $p_\mu < \alpha$ , the hypothesis tested is discarded. That means an upper limit can be set by finding the  $\mu$  that fulfills  $p_\mu = \alpha$ . Values of  $\mu$  that do not pass the threshold are excluded at a confidence level ( $CL$ ) [102]:

$$CL = 1 - \alpha = 95 \%. \quad (7.13)$$

However the  $CL$  method can lead to unnaturally strong limits when the expected signal is small, background contaminations are large and measurements suffer from high uncertainties [103]. Thus the modified  $CL_s$  method is used that include background measurements to calculate upper limits [104, 105]. There the p-value is defined as ratio:

$$p'_\mu = \frac{p_\mu}{1 - p_b} \quad (7.14)$$

with

$$p_b = 1 - \int_{\tilde{q}_{\mu, \text{obs}}}^{\infty} f(\tilde{q}_\mu | 0, \hat{\boldsymbol{\theta}}(0, \text{obs})) d\tilde{q}_\mu. \quad (7.15)$$

An illustration of the test statistic probability function with the different p-values can be found in Figure 7.1.1.

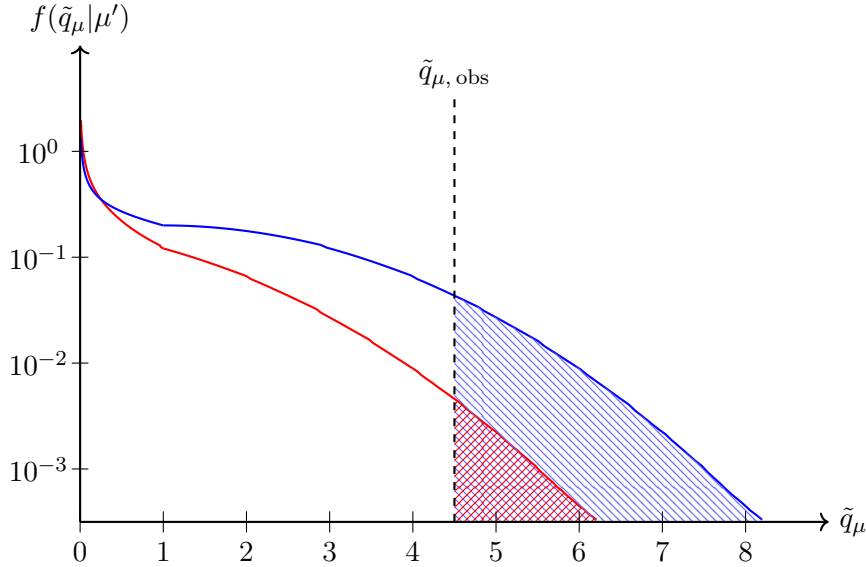


Figure 7.1.1: Schematic plots of the probability density functions approximated with the asymptotic formula. The blue line corresponds to the background only hypothesis with  $\mu' = 0$  and the red line to the signal hypothesis with parameter  $\mu' = \mu$ . From the right side of the observed  $\tilde{q}_\mu$  the blue shaded area resembles  $1 - p_b$  while the red one is  $p_\mu$ .

Using the modified p-value with the constrain  $p'_\mu = \alpha$  to find the upper limit on  $\mu$ ,

the corresponding  $CL_s$  confidence level is defined as:

$$CL_s = 1 - \alpha. \quad (7.16)$$

To estimate the sensitivity of an analysis based on background and signal estimation, an upper limit can be set using an artificial data set instead of the observed data, called Asimov data set. The Asimov data set is pseudo data defined such, that if the likelihood is maximized, all MLE of the nuisance parameters are at their nominal value. This holds true if the Asimov data of bin  $i$  is set to the expected value:

$$n_{i, A} = E_{\mu'}[n_i] = \mu' s_i + b_i. \quad (7.17)$$

As the nominal signal hypothesis is tested and rejected in favor of the background hypothesis, the Asimov data set is generated with  $\mu' = 0$ . The standard deviation of  $\hat{\mu}$  of the Asimov data set is given by:

$$\sigma_A = \frac{\mu - \mu'}{\sqrt{\tilde{q}_{\mu, A}}} = \frac{\mu}{\sqrt{\tilde{q}_{\mu, A}}}. \quad (7.18)$$

Instead of using  $\tilde{q}_{\mu, A, \text{obs}}$  of the artificial data set, the median of the corresponding probability density  $f(\tilde{q}_{\mu, A}|0)$  is used, as it is less prone to statistical fluctuations of the background estimations. The resulting  $CL_s$  limit on  $\mu$  is called expected limit and corresponds to the lowest possible upper rejection limit on the signal hypothesis assuming the data follows the background distribution. Additionally, one can find the  $\pm N\sigma$  bands of the limit by varying  $\hat{\mu}$  in equation (7.11) by  $N\sigma_A$  around the mean  $\mu'$ . Consequentially, by comparing the expected limit of two analysis methods, the method that is most sensitive to a potential signal is determinate by the method that delivers the lowest expected limit with similar or even smaller error bands.

## 7.2 Limit setting results and uncertainties

Since no clearly visible excess of data over the SM background is found in the signal region, the BDT response distributions are used to set an upper limit in the parameter space of the investigated theory, excluding certain parameter configurations. In the  $m_A$ - $\tan(\beta)$  parameter plane of the  $m_h^{\text{mod}+}$  MSSM scenario,  $CL_s = 95\%$  upper limits are set with signal distributions scaled to the cross section times branching ratio of the parameter point predicted by theory.

For each  $m_A$  point from 300 GeV to 1200 GeV, the  $\tan(\beta)$  space is scanned from 5 to 60 in steps of 5  $\tan(\beta)$  points. The resulting limit plots are contour graphs for the nominal signal hypothesis with  $\mu = 1$ . The nominal signal hypothesis is excluded in favor of the background only hypothesis with respect to the corresponding confidence level.

Monte Carlo generated signal and background distributions, as well as the data driven multi-jet background estimation, are effected by the systematic uncertainties listed and explained in Appendix B.1.

The systematic uncertainties on Monte Carlo samples are mainly uncertainties on scale factors. These scale factors parametrize the difference between Monte Carlo simulations and measurements on recorded data. The most important scale factor systematics arise from the tau energy scale calibration scale factors, trigger scale factors, flavor tagging scale factors and pile-up reweighting. Also systematics on multi-jet

fake factors and fake rates are taken into account, which are mainly determined by statistical uncertainties in the control regions.

Also cross section uncertainties of the Monte Carlo generated backgrounds, provided by theoretical calculations, are considered. For top background, a event generator depended uncertainty on factorization and renormalisation modeling is estimated by varying the generator configuration. Additionally, uncertainties on shower generation and hard scatter generation are estimated by comparing the simulated nominal top event sample with samples showered by the Herwig++ [106] shower generator for shower uncertainty and for hard scattering uncertainties, events generated with POWHEG-BOX and MADGRAPH5\_aMC@NLO, both showered with Herwig++, are compared. The normalization shift of the combined generator uncertainty is factored out and summarized in the top background normalization uncertainty, while the impact on the shape of the distribution is modeled in the top background shape uncertainty. Further systematic uncertainties on the signal samples associated with final and initial state radiation, modeling of multi parton interactions, parton distribution functions, as well as normalization and factorization scale are estimated by an acceptance study. This study is done on particle level without detector simulation, using an event selection similar to that in the main analysis. From the upper and lower variations of the model parameters, the combined acceptance uncertainty is calculated from the event yields.

A detailed description of the top background and signal acceptance uncertainty estimations can be found in Ref. [87].

Each systematic uncertainty is modeled with an up and down variation. To propagate them, the classifier responses are calculated by varying the up and down values of the corresponding scale factors, resulting in different BDT distribution. This is done for each systematic uncertainty independently. Only normalization uncertainties with an at least 1% normalization shift and shape uncertainties are considered in the limit setting. Normalization uncertainties with lower impact are neglected, as they are small compared to major ones. Pruning of minor uncertainties reduces the amount of free parameters that have to be fitted in order to find the MLE, thus increases the speed of likelihood minimization and stabilizes the fitting. The limits calculated using the response distributions of the BDTs trained on different training configurations are presented in the following two sections.

### 7.2.1 Limit results of the BDTs trained with the classic splitting model

The resulting upper limits in the  $m_A$ - $\tan(\beta)$  plane for B-TAG and B-VETO using the response distribution of the BDTs trained on the classic splitting model are shown in Figure 7.2.1. The black dotted line corresponds to the expected upper limit using the BDT response distributions with the corresponding  $\pm 1\sigma$  (green) and  $\pm 2\sigma$  (yellow) bands. Since the same data set and uncertainties are used to calculate the limits using the  $m_T^{\text{tot}}$  distributions, both limits are directly comparable.

The expected limit in the B-VETO category is similar to that derived from the  $m_T^{\text{tot}}$  distribution. However, the expected limit in B-TAG is approximately  $4 \tan(\beta)$  points lower for the mass region above 800 GeV signal mass. This lowers the combined limit of both channels drastically in the high mass regions, as shown in Figure 7.2.2. The numerical values for the combined limits are listed in Table B.2.1.

Although the combined expected limit experienced an improvement in the high mass region, the observed limits of the analysis using BDT response distributions are comparable to those set using  $m_T^{\text{tot}}$ . Both observed limits vary within the  $\pm 1\sigma$  bands of the

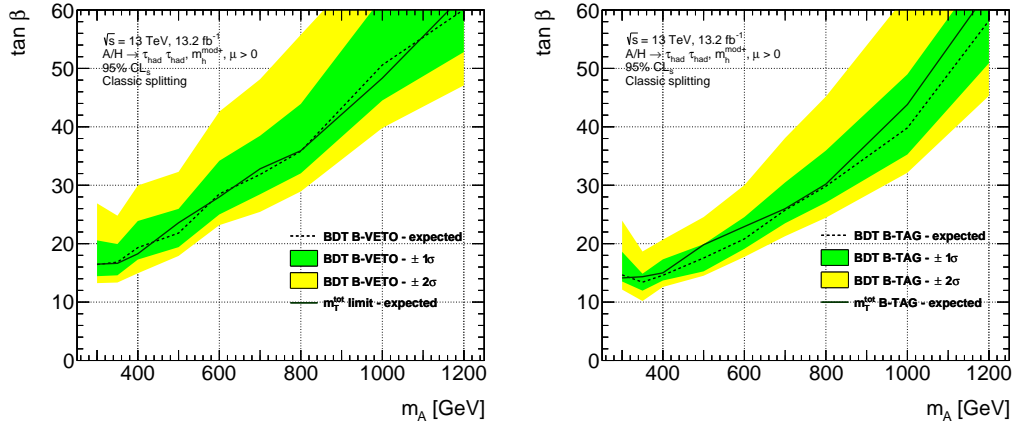


Figure 7.2.1: B-VETO (left) and B-TAG (right) expected limits calculated using the response distributions of the BDTs trained with the classic splitting model (black dotted line) in comparison to the  $m_T^{\text{tot}}$  limit (green solid line).

expected limit, which indicates that no significant deviation from the SM consistent with a heavy neutral Higgs boson signal is observed.

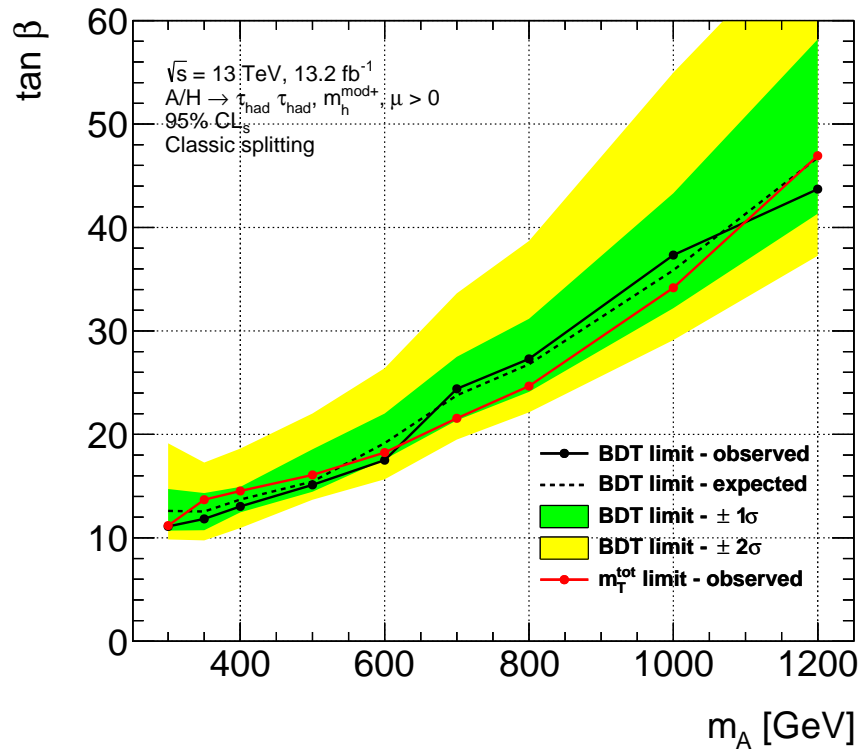
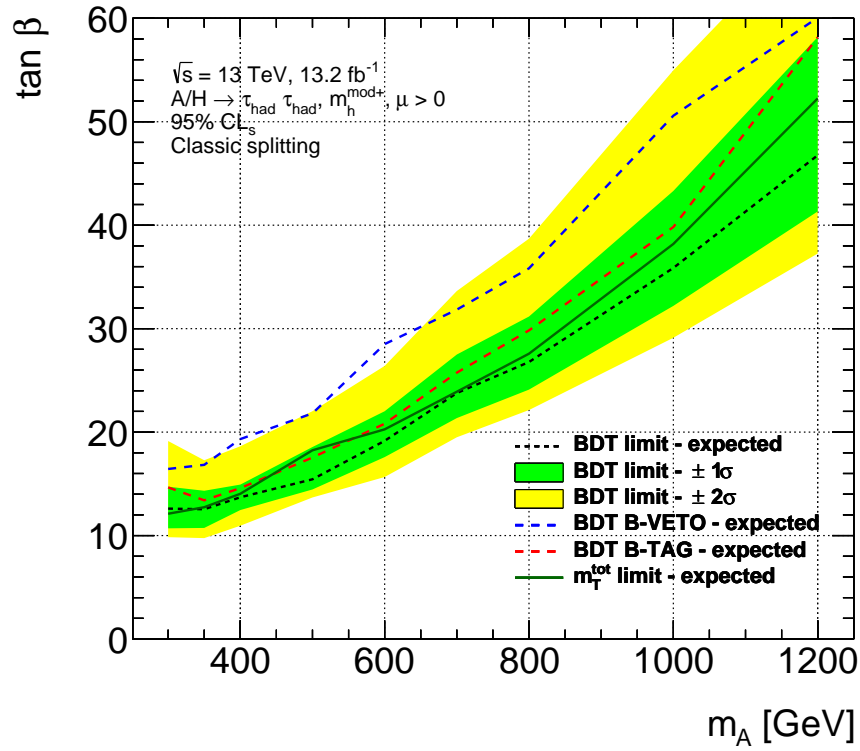


Figure 7.2.2: Combined expected and observed limits calculated using the response distributions of the BDTs trained with the classic splitting model in comparison to the  $m_T^{\text{tot}}$  limit.

## 7.2.2 Limit results of the BDTs trained with K-Fold model

Expected limits set in the  $m_A$ - $\tan(\beta)$  using BDTs trained on the K-Fold model in B-TAG and B-VETO are presented in Figure 7.2.3. Similar to the classic splitting model, the B-VETO limit of the BDTs are comparable to the one using the  $m_T^{\text{tot}}$  distribution, while in the B-TAG region improvements are achieved. Compared to the expected limit obtained from BDTs trained on the classic splitting model, here the improvements are mainly located in the lower mass region  $m_A \leq 600$  GeV, whereas the limit for the high mass, while being comparable or better than the  $m_T^{\text{tot}}$  limit, did not improve as much. Consequentially the sensitivity of the analysis using both

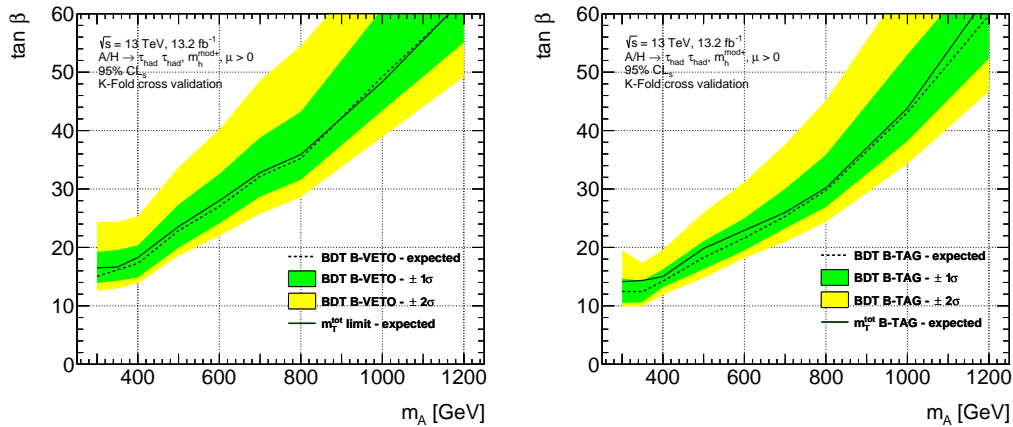


Figure 7.2.3: B-VETO (left) and B-TAG (right) expected limits calculated using the response distributions of the BDTs trained on the K-Fold cross validation method in comparison to the  $m_T^{\text{tot}}$  limit.

channels increases in the lower mass region. Comparing the combined BDT expected limit with the  $m_T^{\text{tot}}$  limit in Figure 7.2.4, former is 1 to 3  $\tan(\beta)$  points lower over the entire mass range, with exception of the 1000 GeV point that shows similar sensitivity as the  $m_T^{\text{tot}}$  limit. This might be due to the shape difference observed between the individual BDTs, so that the combined classifier may experience a performance loss compared to the best single BDT. The numerical values of the combined limits are summarized in Table B.2.2.

The combined observed limit of the BDT fluctuates within the  $\pm 1\sigma$  bands around the expected limit, with exception of the  $m_A = 1200$  GeV point which is stronger constrained, almost reaching the border of the  $-2\sigma$  band. This might be due to the empty last bin in the response distribution of the 1200 GeV BDT in B-TAG shown in Figure A.6.6. Since the BDT shifts the signal distribution strongly to higher response values, the last bin is the most sensitive bin containing most of the signal. If no data is contained in this bin, despite non-negligible background expectation, this results in a stronger observed upper limit, because an empty bin is more in alignment with the background-only hypothesis than with the nominal signal hypothesis.

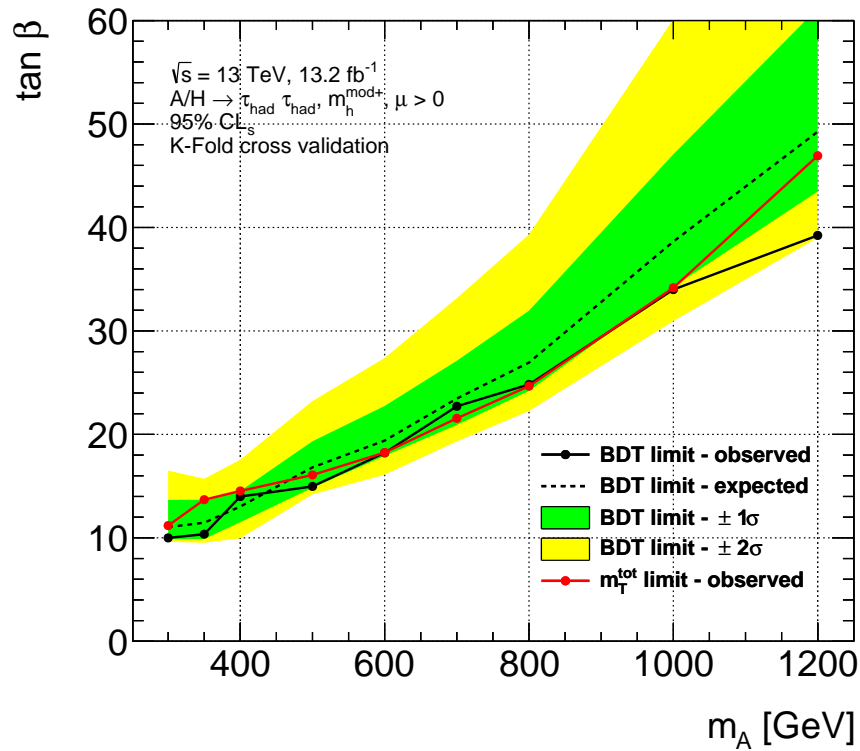
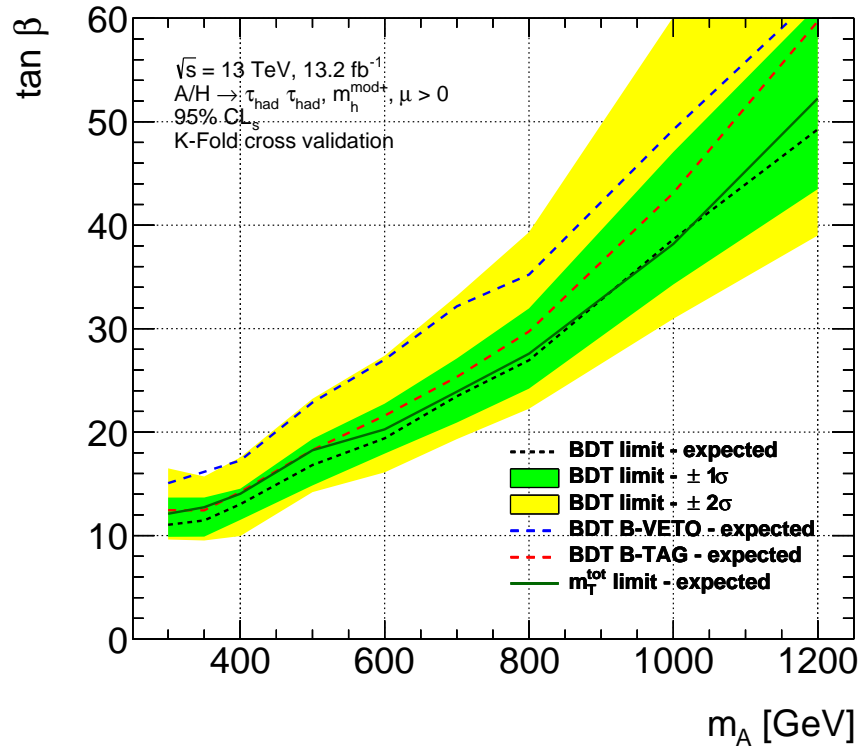


Figure 7.2.4: Combined expected and observed limits calculated using the response distributions of the BDTs trained on the K-Fold cross validation method in comparison to the  $m_T^{\text{tot}}$  limit.



## 8 Summary

In this study the search for heavy neutral MSSM Higgs bosons decaying into a pair of hadronically decaying tau leptons at the ATLAS detector is performed using  $13.2 \text{ fb}^{-1}$  of data taken at a center of mass energy of  $\sqrt{s} = 13 \text{ TeV}$ . In this thesis a multivariate analysis method, the Boosted Decision Tree, is introduced to separate signal from background.

The BDT algorithm requires a set of signal and background events to train on it. To enhance the separation power, individual BDTs are trained on the different Higgs masses that are investigated, ranging from 300 GeV up to 1200 GeV. This allows the BDTs to specialize for specific signal masses. Furthermore, since the B-TAG and B-VETO categories consist of different background compositions with different decay topologies, individual BDTs are trained for both categories separately, thus increasing the specialization and separation power of the classifier even more.

A major problem in the training of the BDTs is the limited number of background events, so two different approaches of splitting the available event sample into training and testing subsets is investigated and compared. In the classic splitting model the training set gets discarded once the BDT has been trained. This procedure is simple to implement but results in lower background statistics for further analysis. To counter this disadvantage of the classic split model, a K-Fold cross-validation method is introduced with  $K = 2$ . There the entire background and signal set is split evenly in half, where two BDTs are trained on each one of the subsets and then applied on the other. Using this approach the entire background and data set can be used since no events are classified and trained by the same BDT. However, BDTs trained on a pair of data sets are required to select similar phase spaces, consequentially making the training and selection of the classifiers even harder. As seen in the case of the 1000 GeV signal mass K-Fold BDT, too large deviations between the individual sub-BDTs can deteriorate the performance of the combined classifier.

All BDTs are trained on a sample of the dominant signal processes and the two leading backgrounds in each category. Considering the low background statistics, the BDTs show no significant overtraining. The classifiers were validated in the same-sign validation region where none of the BDTs show a mismodeling of the background.

Since no significant excess of data over the background estimation in the signal region is observed, upper limits in the  $m_A$ - $\tan(\beta)$  plane of the examined  $m_h^{\text{mod}+}$  scenario of the MSSM are set. In this thesis the BDT response distributions are used to calculate  $CL_s = 95\%$  upper limits. The derived limits are then compared to the previously published result in Ref. [87] that rely on the total transverse mass distribution of the tau decay products as discriminating variable.

The resulting combined expected limit using the BDTs trained on the classic model show drastic improvements at high  $m_A$ , lowering the limit by 5.49  $\tan(\beta)$  points at  $m_A = 1200 \text{ GeV}$  compared to the  $m_T^{\text{tot}}$  limit in the  $m_h^{\text{mod}+}$  scenario.

The expected limit set by using the BDTs trained with the K-Fold method performs better in the lower mass region since the entire background and signal data set can be used to set limits. This reduces the statistical uncertainty of the background and signal estimations which results in better expected limit. However, for high signal

masses the K-Fold BDTs do not improve the limit as much as the BDTs trained on the classic model. The loss in separation power in the higher mass regions might be due to the problems encountered in finding a set of two BDTs that select similar phase spaces.

For both training setups, the main driving factor in improving the limit is a performance increase in the B-TAG region.

Over all, both concepts of training and applying the BDTs show significant improvements in the limit setting when compared to the previously published results. This proves the feasibility of using multivariate methods in the analysis presented and should therefore substitute the total transverse mass as discriminating variable in future analyses in this decay channel.

With further data taken by the ATLAS detector more multi-jet background events will be available for training, resulting in a more stable BDT training, especially for B-TAG type BDTs. This might also improve the capability of resolving more features of the training data, subsequently improving the separation power.

With further background and signal events available for training, even more classifiers could be combined using the K-Fold cross validation method with  $K = 3$  or higher, thus improving the stability of the combined classifier. This however is currently not possible due to statistical limitations.

Also by introducing new variables that separate signal and background well, further improvements could be achieved. Especially variables that hold information about the substructure of the tau and b-jet decay might yield further separation power.

# A BDT configuration and validation

## A.1 Validation of the training variables

In order to train reliable classifiers, it is important to examine the background modeling of the variables used in training. If significant systematic mismodeling is observed, BDTs trained with these variables might suffer from performance loss and bias the analysis. To validate the background modeling, the same-sign validation region is used. Figure A.1.1, A.1.2 shows the same-sign distributions of the training variables in B-TAG and the distributions for B-VETO are listed in Figure A.1.3, A.1.4. No deviation in the background modeling is found.

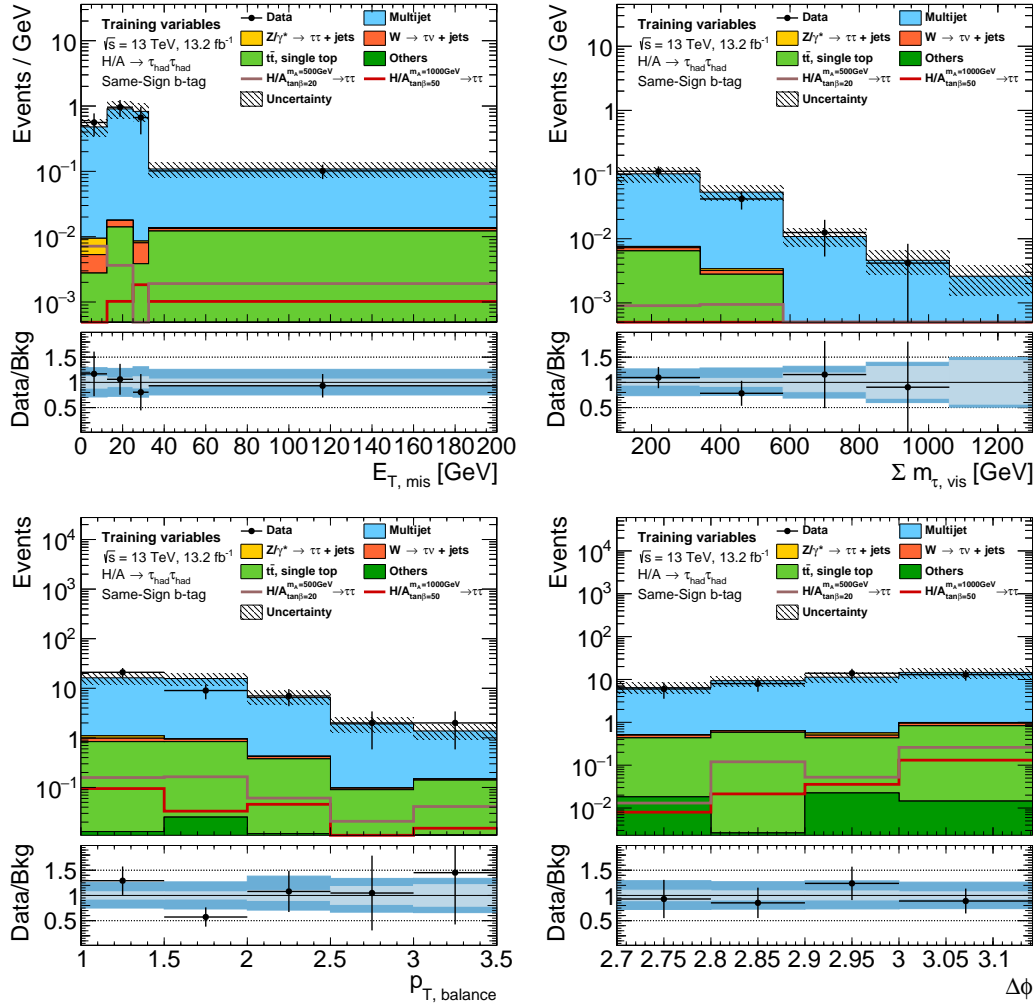


Figure A.1.1: Same-sign distributions of variables used for training in B-TAG.

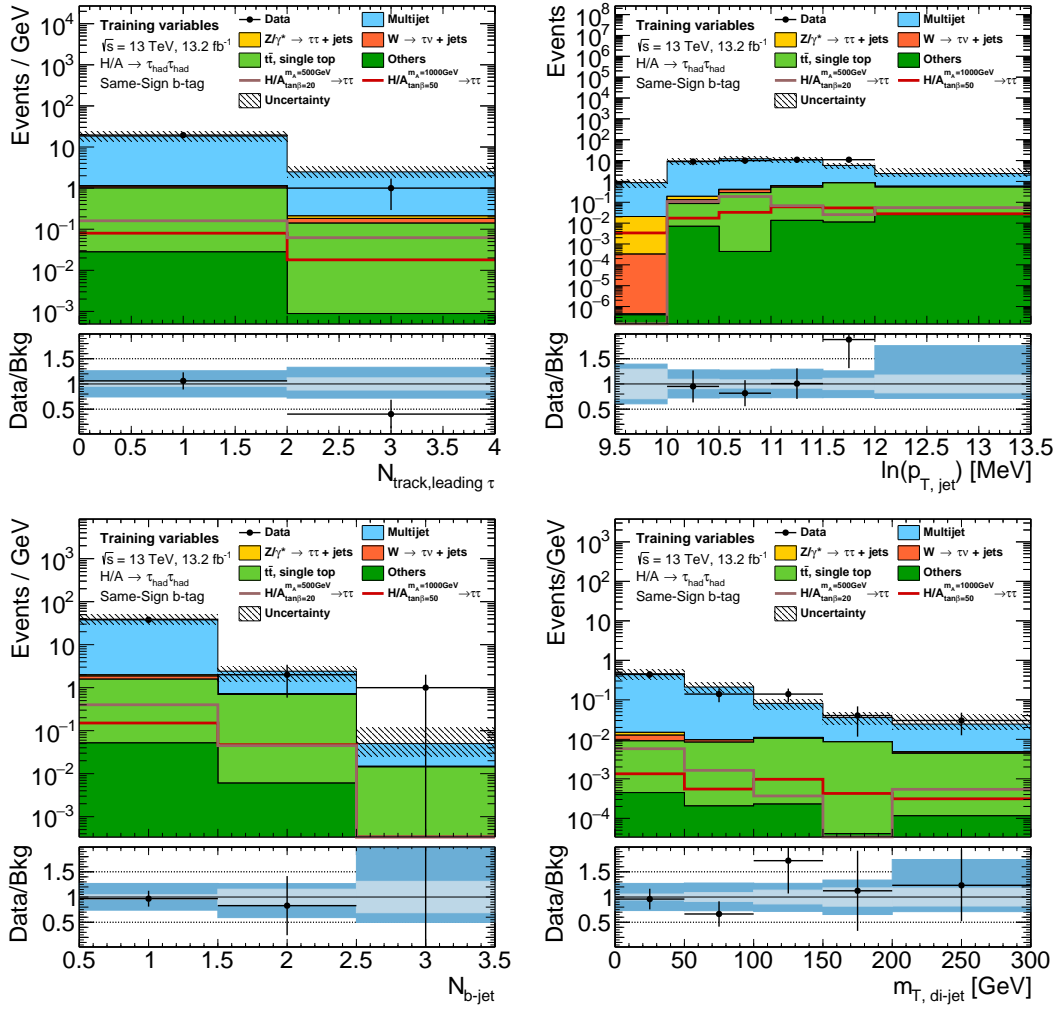


Figure A.1.2: Same-sign distributions of variables used for training in B-TAG.

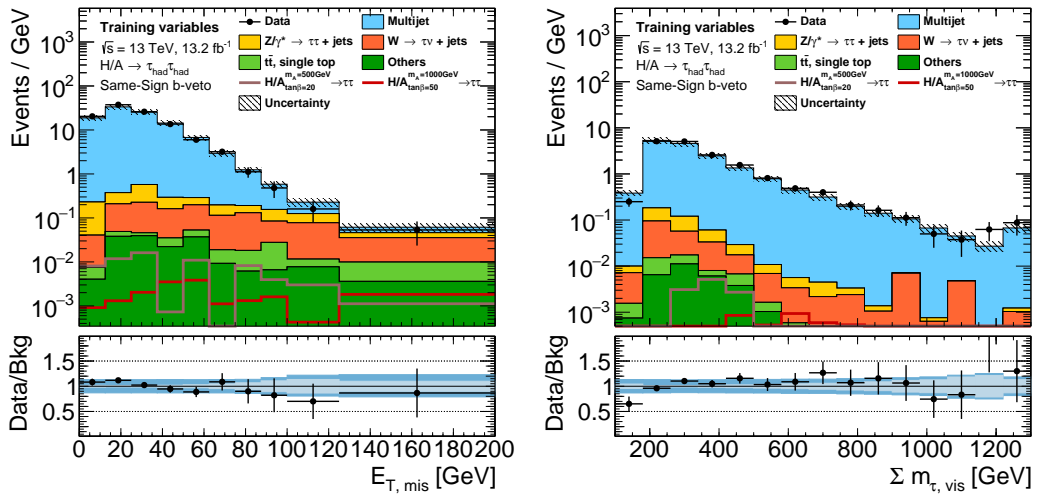


Figure A.1.3: Same-sign distributions of variables used for training in B-VETO.

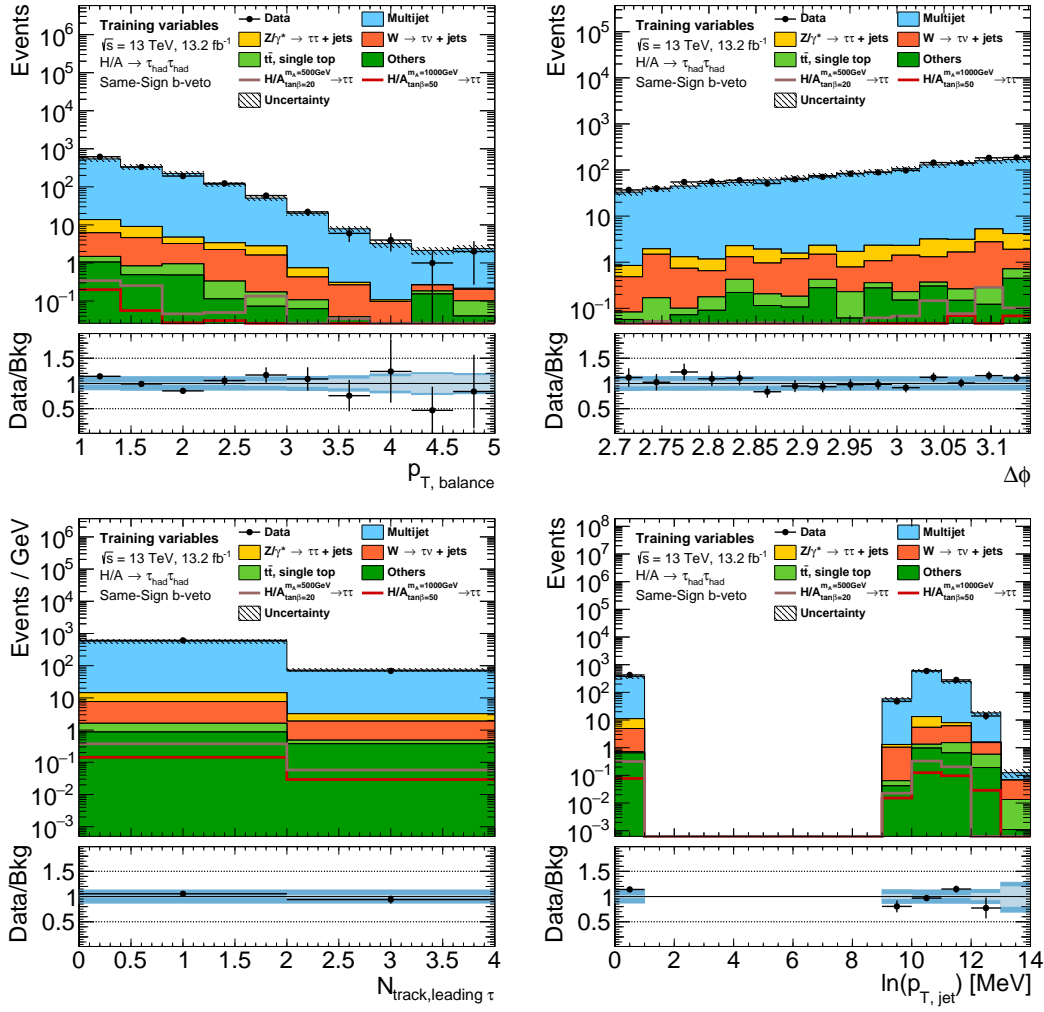


Figure A.1.4: Same-sign distributions of variables used for training in B-VETO.

## A.2 Signal region variable distributions

Plots presented here compare the background and signal distributions of the variables used for training. The signal cross section is scaled to 1 pb in B-VETO and 0.2 pb in B-TAG so that the different shapes of the distributions can be compared better.

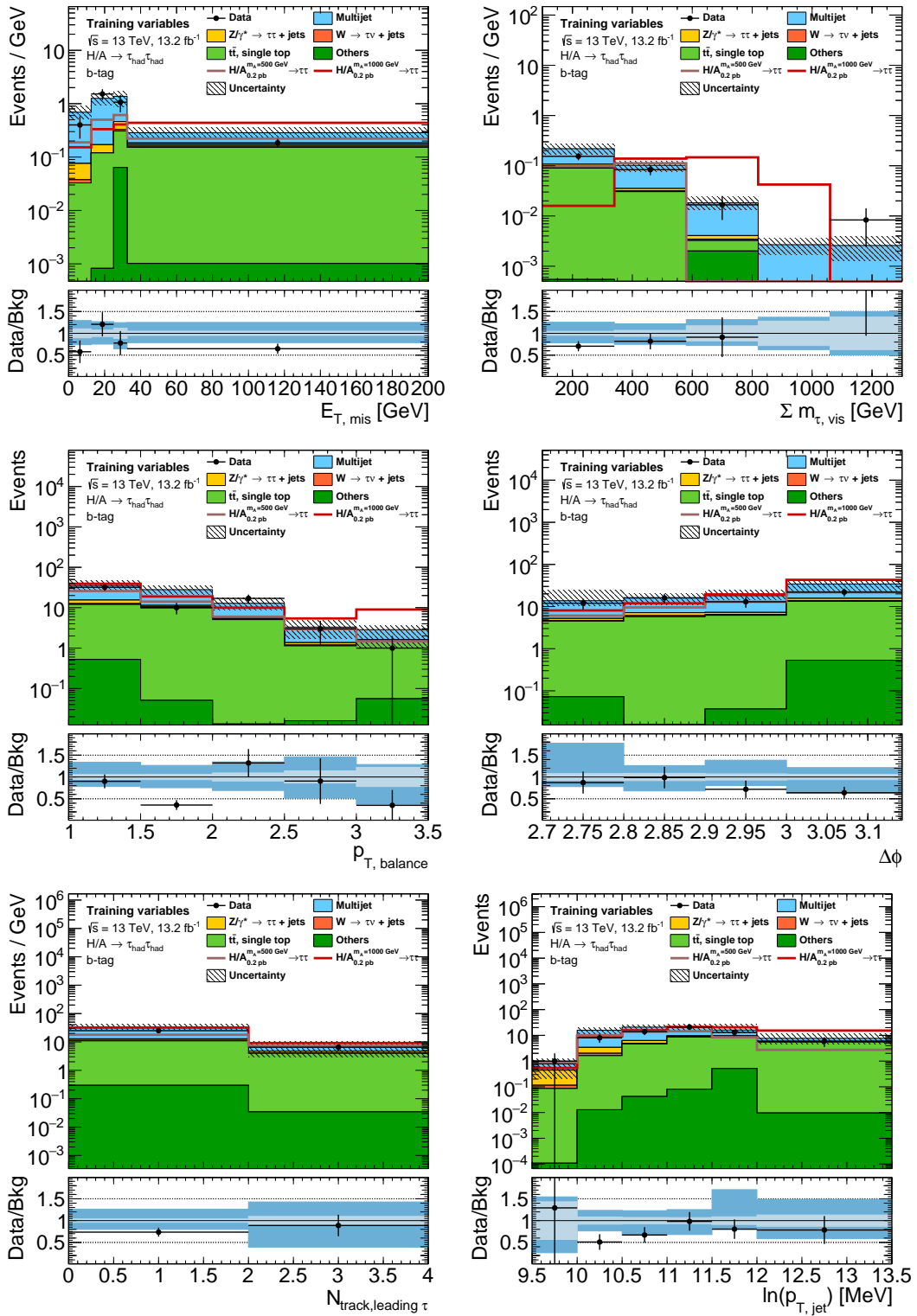


Figure A.2.1: Signal region distribution of the training variables in B-TAG.

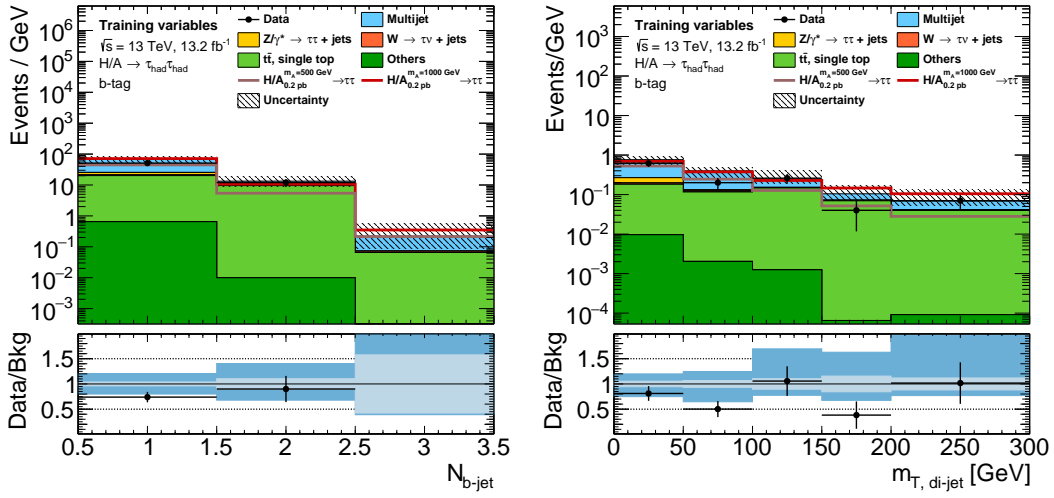


Figure A.2.2: Signal region distribution of the training variables in B-TAG.

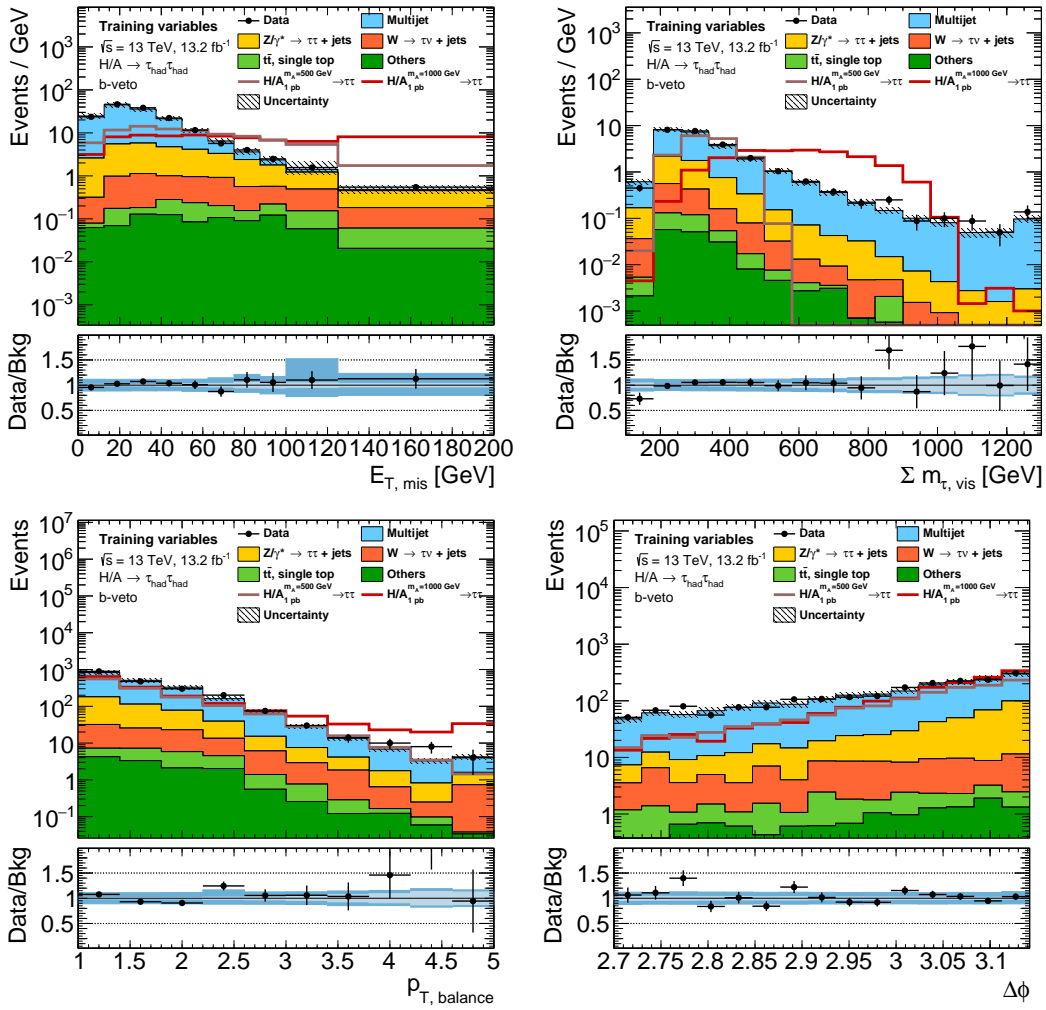


Figure A.2.3: Signal region distribution of the training variables in B-VETO.

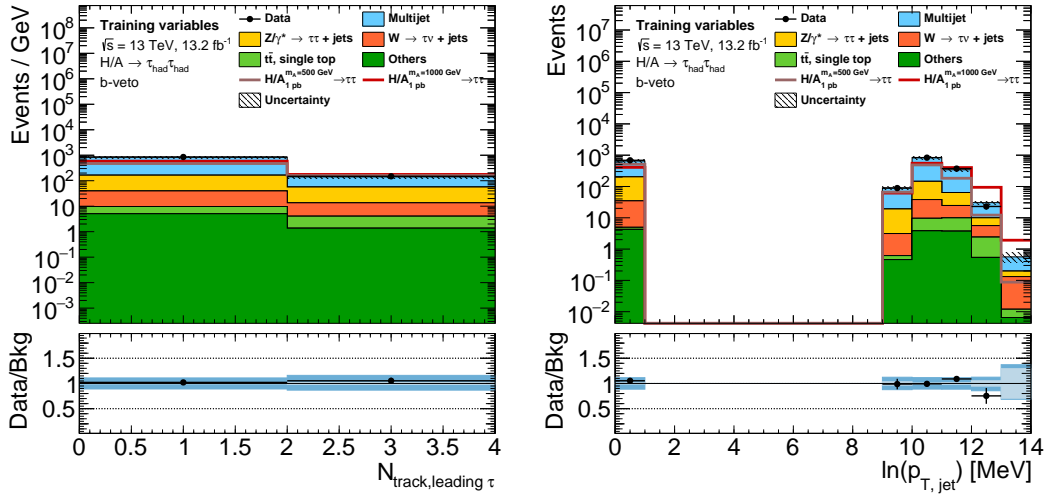


Figure A.2.4: Signal region distribution of the training variables in B-VETO.

### A.3 Configuration of the BDTs

The configuration for each BDT has been selected to minimize overtraining and subsequently improve the performance of the classifier. However, there are common configurations in TMVA that are used for all BDTs trained with the classic splitting and the K-Fold model. These common configurations are listed in Table A.3.1. The configuration for the individual BDTs for each mass point differ in the number of trees trained (NTrees), the minimum fraction of events that have to pass through a node in order to split it further (MinNodeSize) and the maximum depth of the trees (MaxDepth). The individual configurations are listed in Table A.3.2 for the classic sample splitting model and in Table A.3.3 for the K-Fold model.



TMVA configurable	Value	Description
nCuts	20	Number of cuts used to search for optimal separation
BoostType	AdaBoost	Boosting algorithm used for training
UseBaggedBoost	True	Only random subsamples used to train for each iteration
AdaBoostBeta	0.5	Learning rate $\beta$ for adaptive boosting
UseRandomisedTrees	False	At each node choose randomly under the best separating variables
BaggedSampleFraction	0.6	Size of bagged samples. Used whenever bagging is used
UseYesNoLeaf	True	Only use binary output for leaves (Signal/Background)
NegWeightTreatment	InverseBoostNegWeights	Determinate how events with negative weights are treated
NodePurityLimit	0.5	Nodes with purity $>$ are treated as signal nodes
SeparationType	GiniIndex	Statistical metric to determinate node splitting
SigToBkgFraction	1	Ration between signal and background for training
PruneMethod	NoPruning	Pruning method used to remove insignificant nodes

Table A.3.1: Common configuration for BDT training used for B-TAG and B-VETO type BDTs in the classic splitting model and the K-Fold model.

TMVA configurable	Values for BDTs of the different mass regions						Description
	Low		400 GeV		500 GeV		
NTrees	B-TAG 300	B-VETO 300	B-TAG 300	B-VETO 300	B-TAG 300	B-VETO 300	Number of trees grown
MaxDepth	2	2	2	2	2	3	Maximum depth of individual trees
MinNodeSize	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%	Minimum fraction of events that have to pass a node

TMVA configurable	Values for BDTs of the different mass regions											
	600 GeV		700 GeV		800 GeV		1000 GeV		1200 GeV			
NTrees	B-TAG 300	B-VETO 300	B-TAG 300	B-VETO 100	B-TAG 300	B-VETO 300	B-TAG 300	B-VETO 300	B-TAG 300	B-VETO 300	B-TAG 300	B-VETO 300
MaxDepth	2	4	2	4	2	4	2	4	2	4	2	4
MinNodeSize	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%

Table A.3.2: Configuration for BDTs trained with classic splitting model. The Low mass BDT are trained with signal data from 300 GeV and 350 GeV signal samples.

TMVA configurable	Values for BDTs of the different mass regions											
	Low			400 GeV			500 GeV			600 GeV		
	B-TAG	B-VETO		B-TAG	B-VETO		B-TAG	B-VETO		B-TAG	B-VETO	
NTrees	300	300		300	300		300	300		300	300	
MaxDepth	2	2		4	4		2	4		2	4	
MinNodeSize	1.0 %	1.0 %		5.0 %	1.0 %		1.0 %	1.0 %		1.0 %	1.0 %	

TMVA configurable	Values for BDTs of the different mass regions											
	700 GeV			800 GeV			1000 GeV			1200 GeV		
	B-TAG	B-VETO		B-TAG	B-VETO		B-TAG	B-VETO		B-TAG	B-VETO	
NTrees	300	300		300	300		100	300		100	300	
MaxDepth	2	4		2	4		2	4		2	4	
MinNodeSize	1.0 %	1.0 %		1.0 %	1.0 %		1.0 %	1.0 %		1.0 %	1.0 %	

Table A.3.3: Configuration for BDTs trained with K-Fold model. The Low mass BDT are trained with signal data from 300 GeV and 350 GeV signal samples.

## A.4 Event weights of the b-associated signal production Monte Carlo

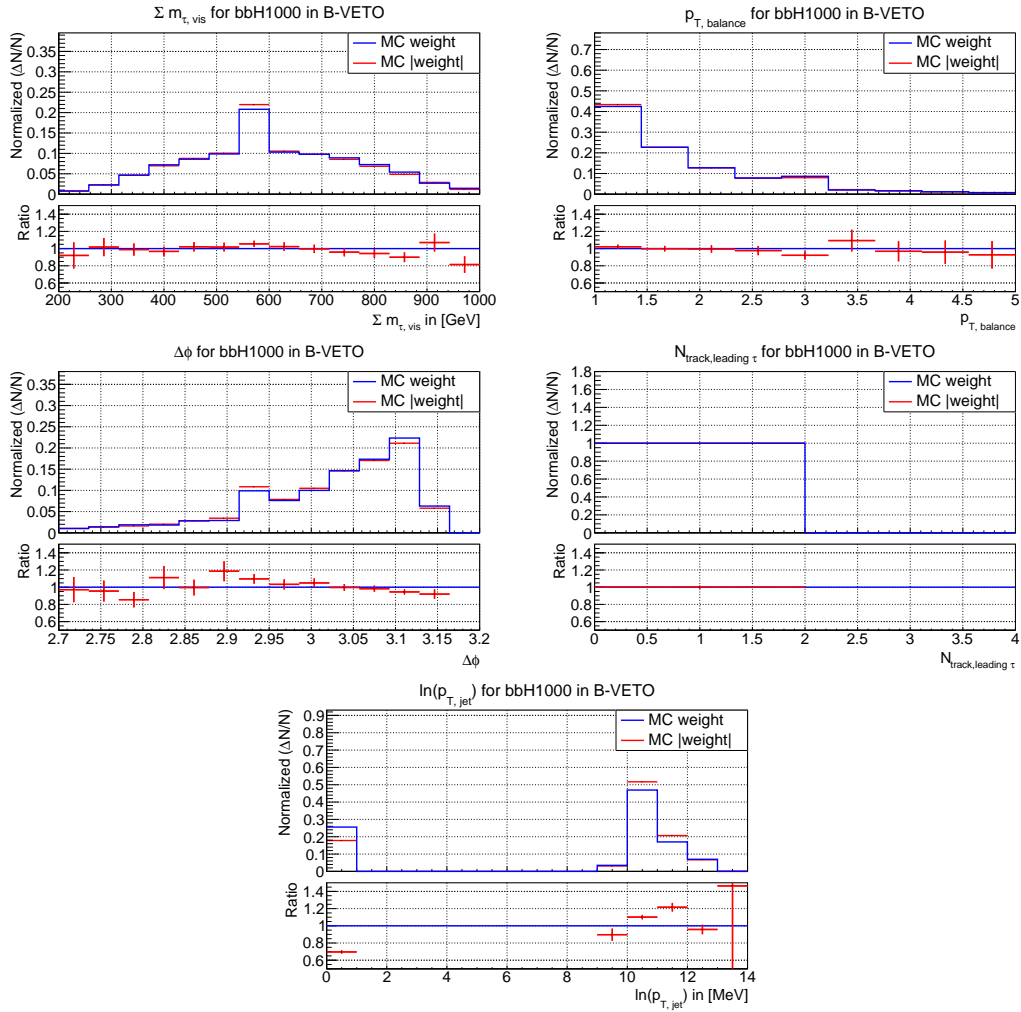


Figure A.4.1: Shape comparison between negative and positive MC weights for 1000 GeV signal mass in B-VETO

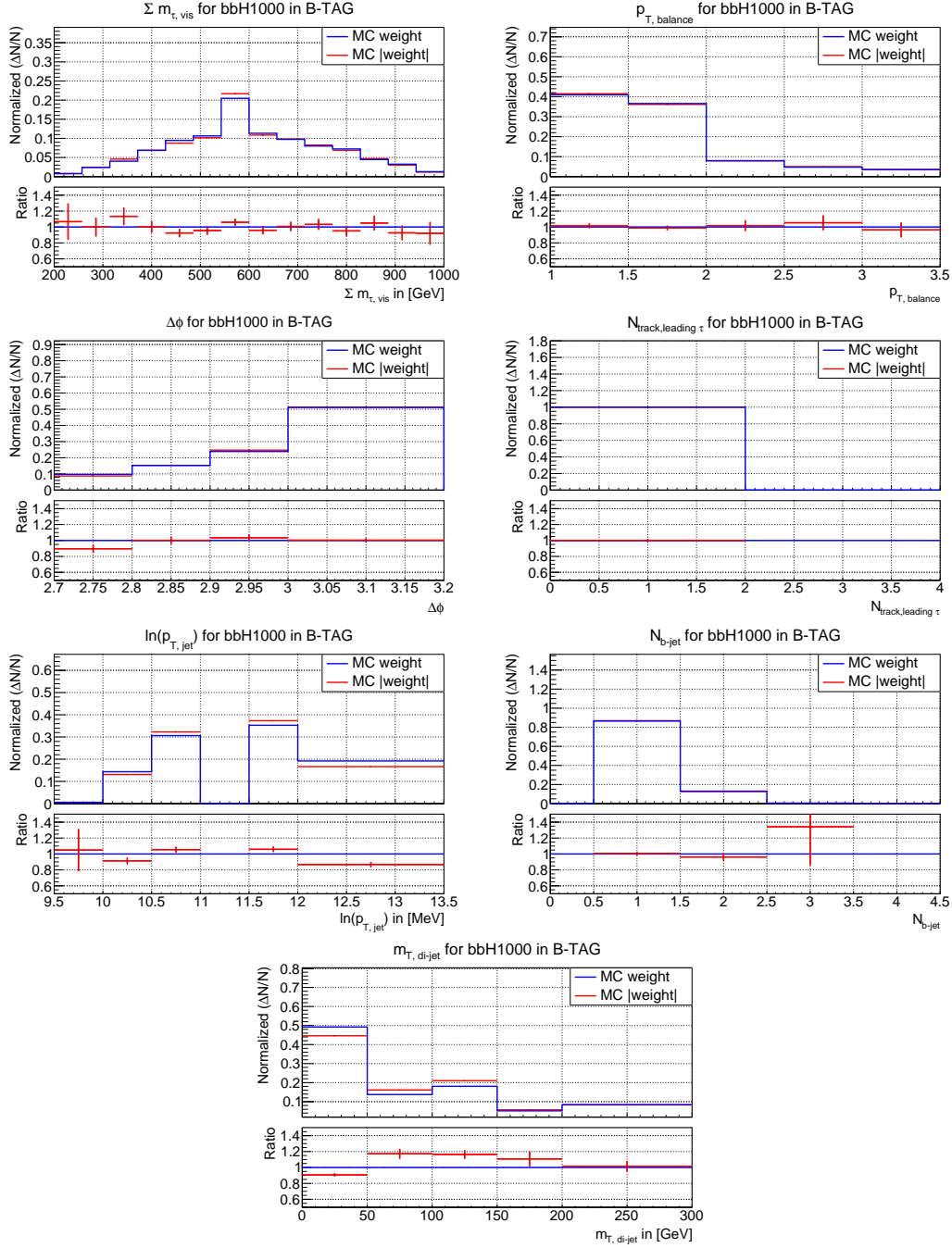


Figure A.4.2: Shape comparison between negative and positive MC weights for 1000 GeV signal mass in B-TAG

## A.5 Validation of the BDTs

In order to check whether the classifiers model the background correctly or cause a systematic mis-modeling, the BDT distributions for data and background are compared in the same-sign validation region. Since in this region almost no signal yields are expected, this region can be used to validate background modeling. The data distribution should match the background estimation. The same-sign distributions are shown in Figure A.5.1 and A.5.2 for the BDTs trained using the classic splitting model and in Figure A.5.4 and A.5.5 for the combined K-Fold BDTs. No significant background mis-modeling is observed.

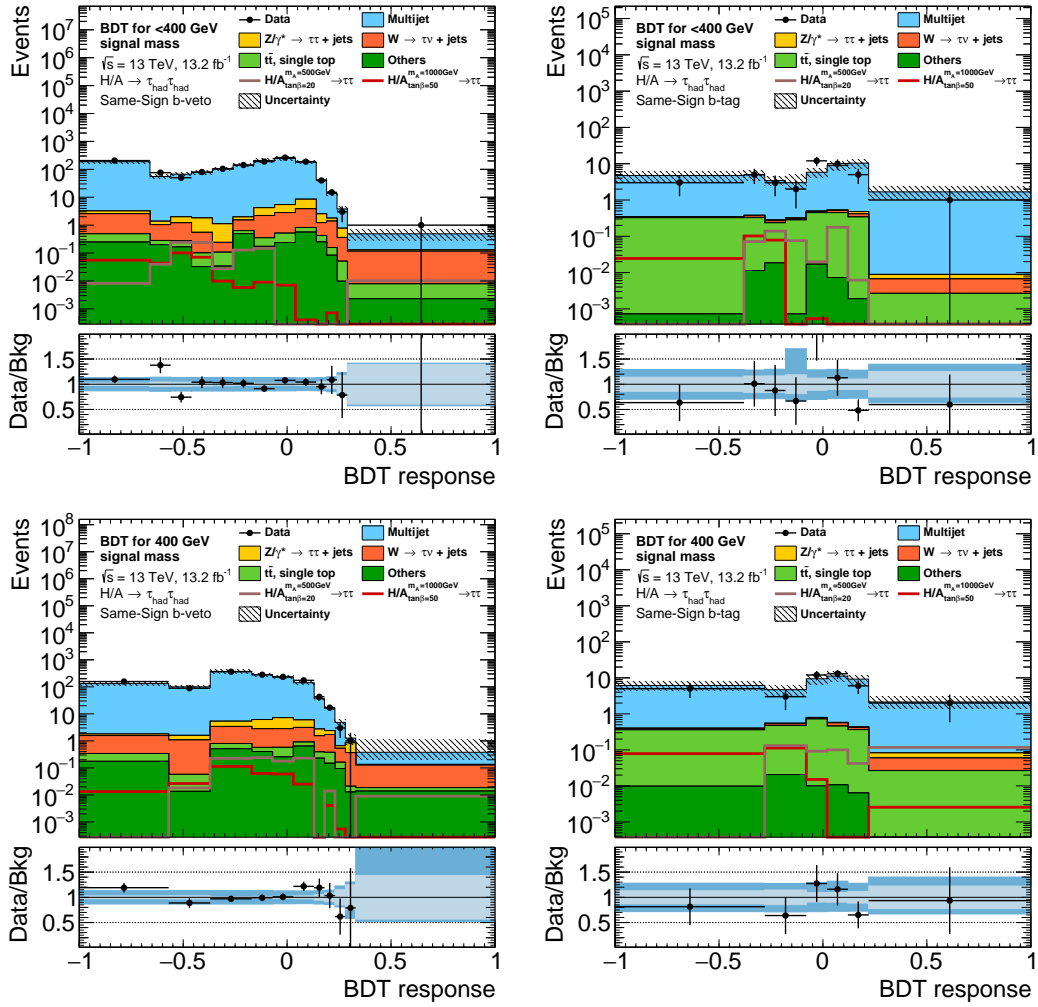


Figure A.5.1: Same-sign validation of BDTs trained with classic splitting for 300, 350, 400 GeV signal mass in B-VETO (left) and B-TAG (right).

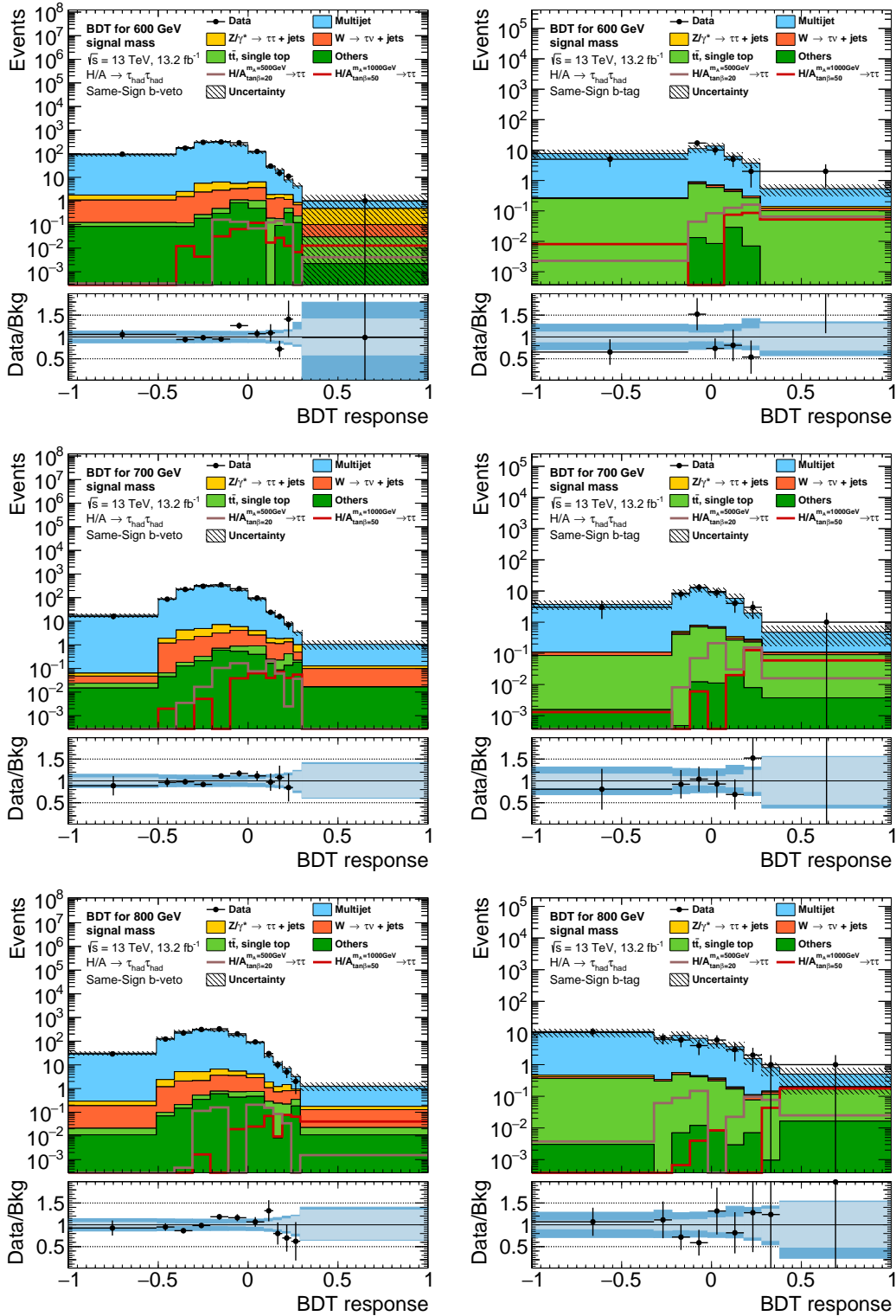


Figure A.5.2: Same-sign validation of BDTs trained with classic splitting for 600, 700, 800 GeV signal mass in B-VETO (left) and B-TAG (right).

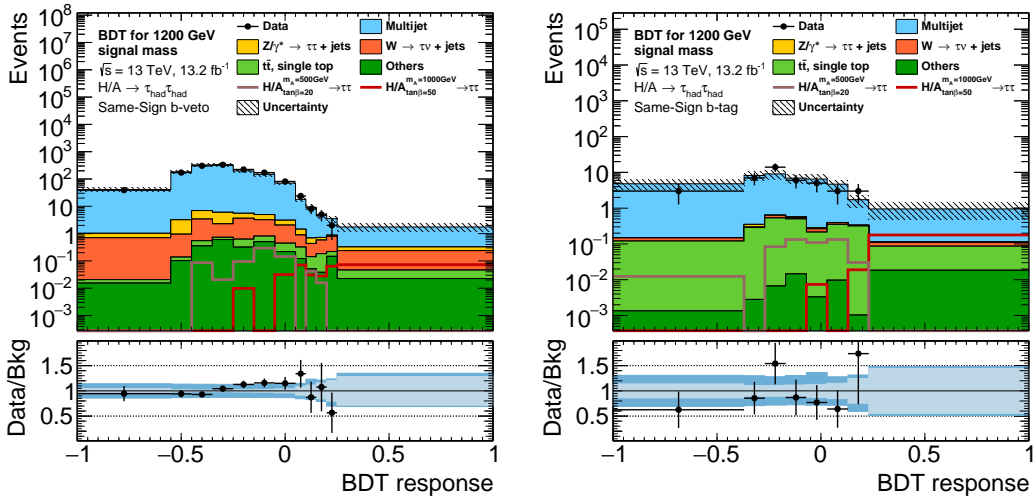


Figure A.5.3: Same-sign validation of BDTs trained with classic splitting for 1200 GeV signal mass in B-VETO (left) and B-TAG (right).

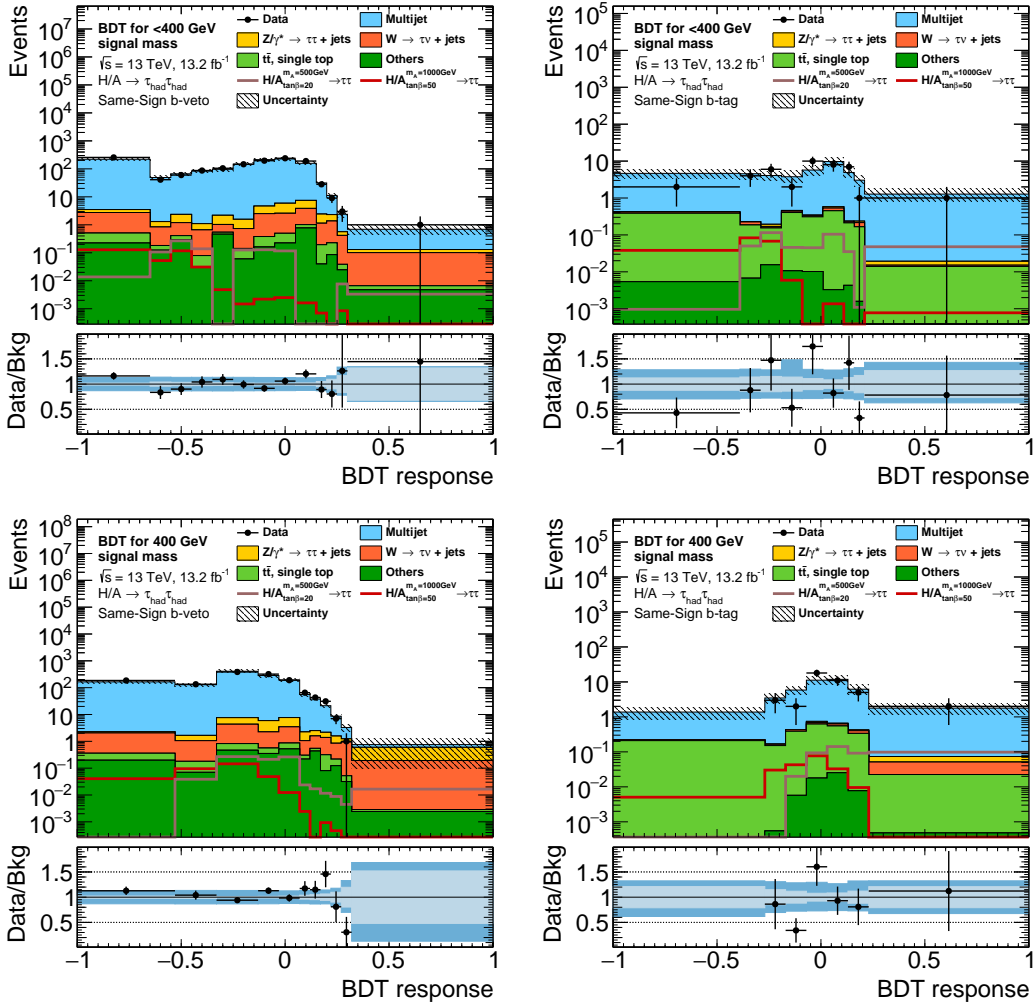


Figure A.5.4: Same-sign validation of BDTs trained using K-Fold method for 300, 350, 400 GeV signal mass in B-VETO (left) and B-TAG (right).



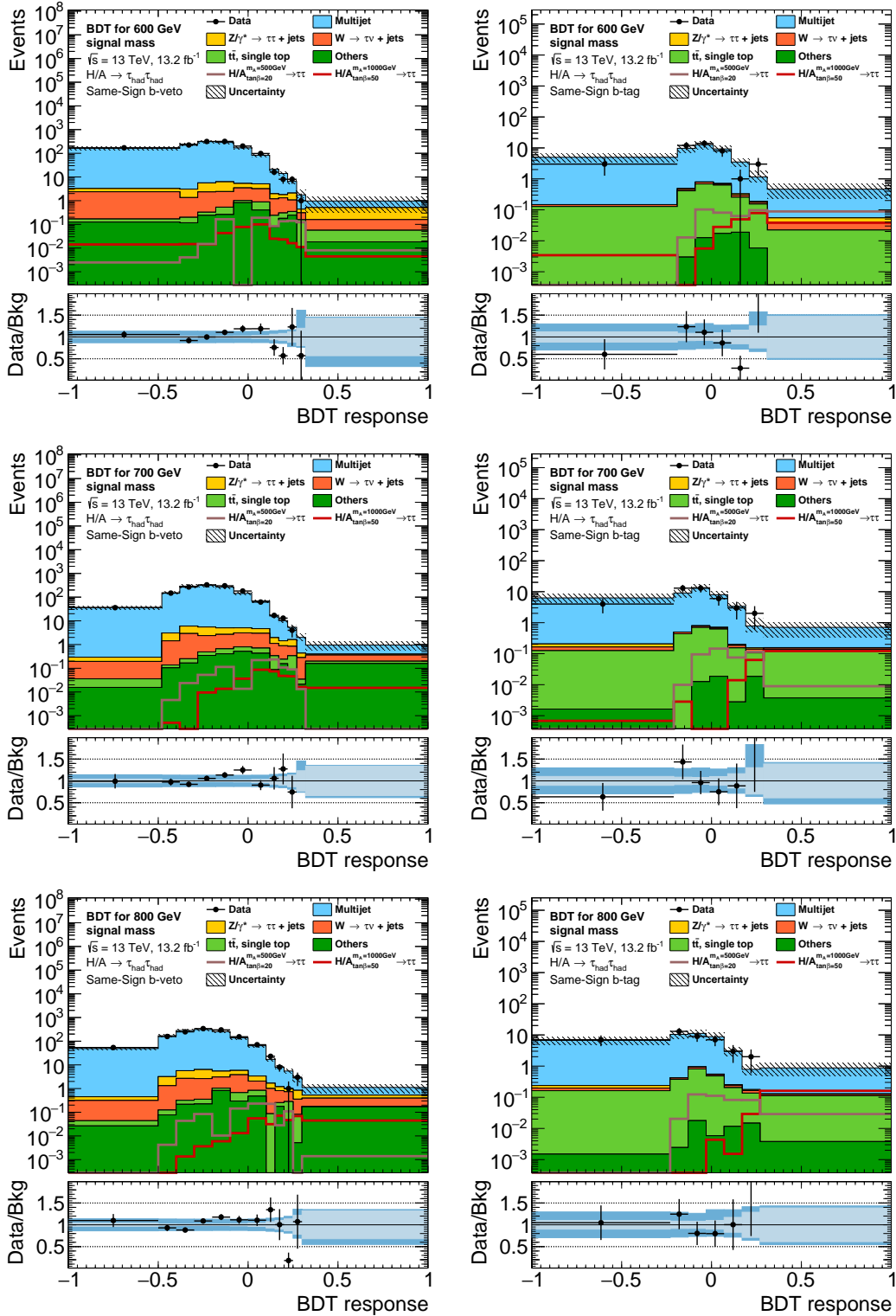


Figure A.5.5: Same-sign validation of BDTs trained using K-Fold method for 600, 700, 800 GeV signal mass in B-VETO (left) and B-TAG (right).

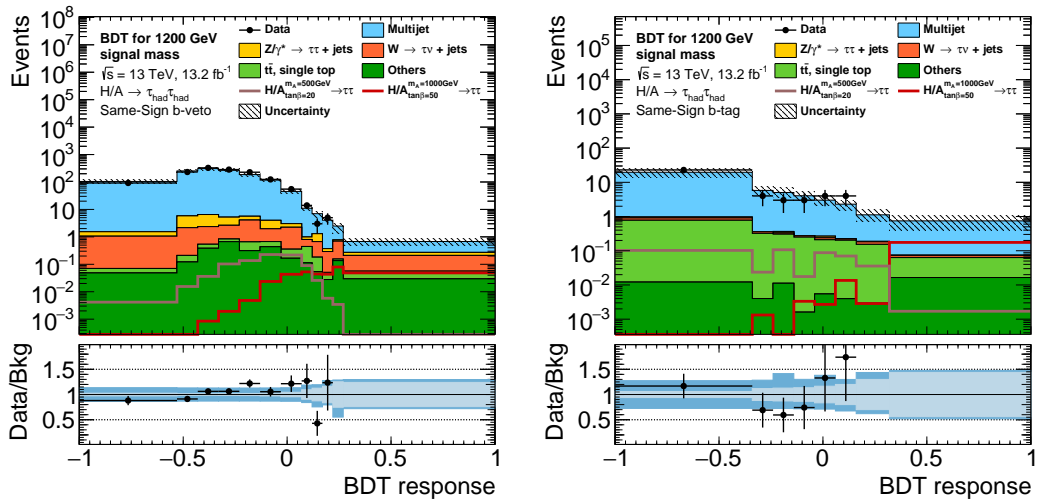


Figure A.5.6: Same-sign validation of BDTs trained using K-Fold method for 1200 GeV signal mass in B-VETO (left) and B-TAG (right).

## A.6 Signal region distributions of the BDTs

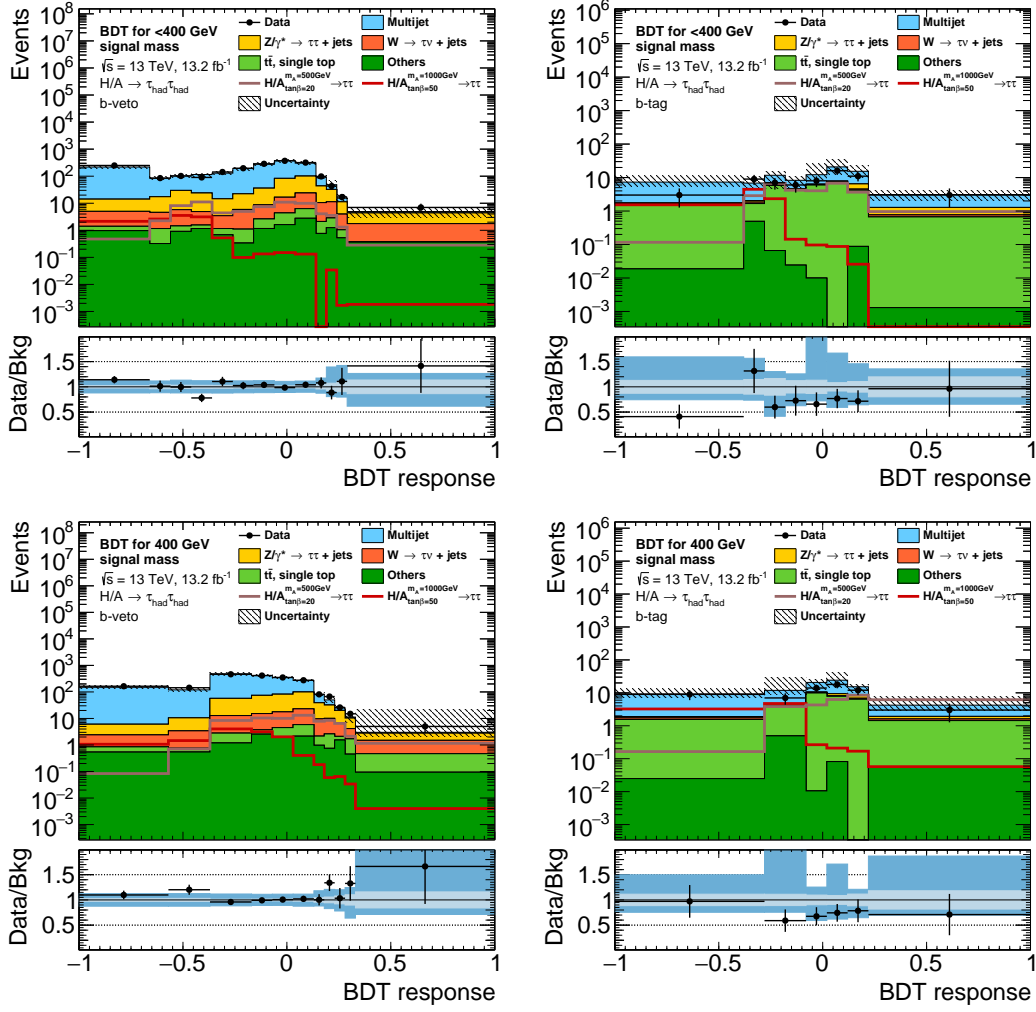


Figure A.6.1: Signal region BDT response of BDTs trained with classic splitting for 300, 350, 400 GeV signal mass in B-VETO (left) and B-TAG (right).

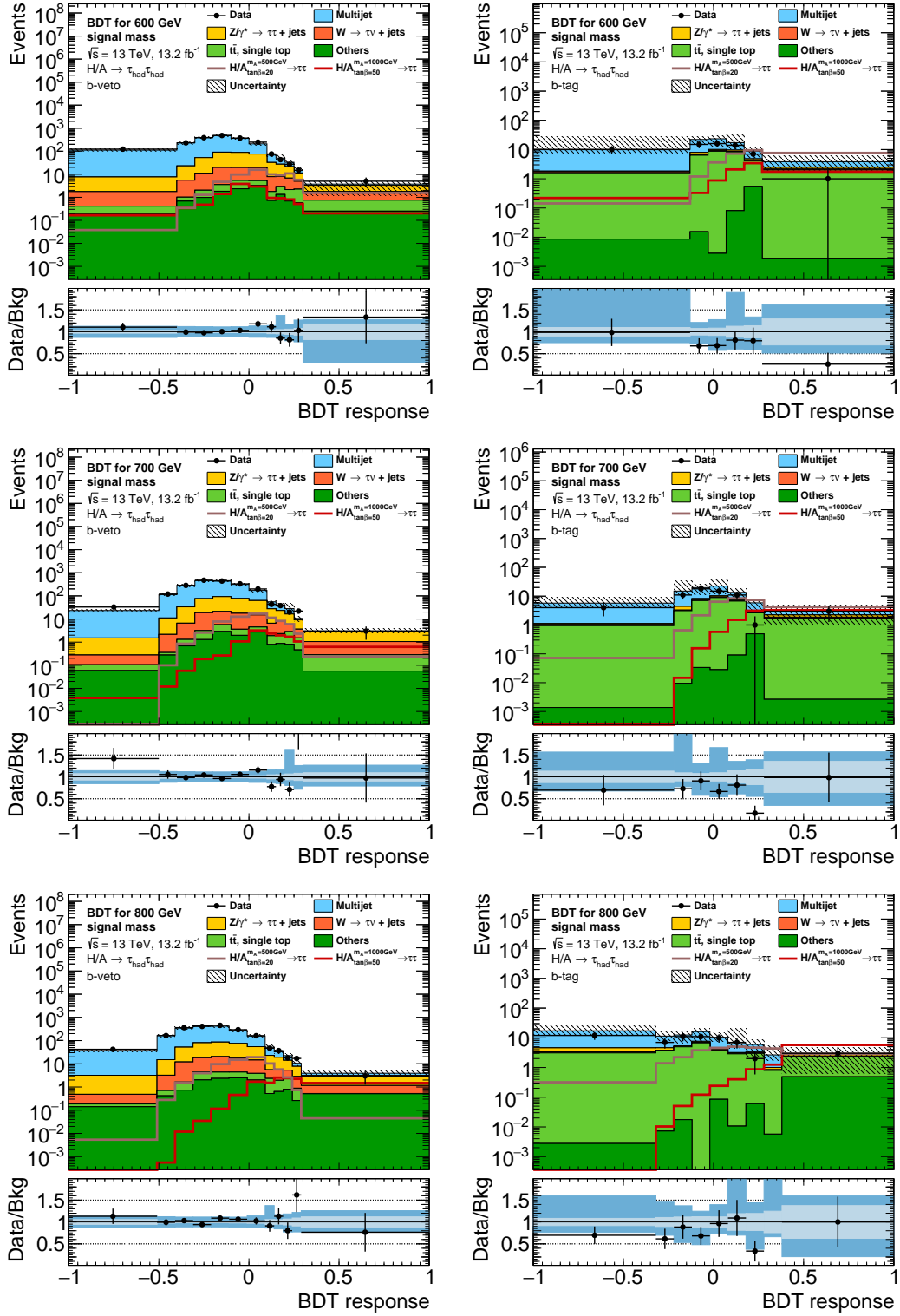


Figure A.6.2: Signal region BDT response of BDTs trained with classic splitting for 600, 700, 800 GeV signal mass in B-VETO (left) and B-TAG (right).

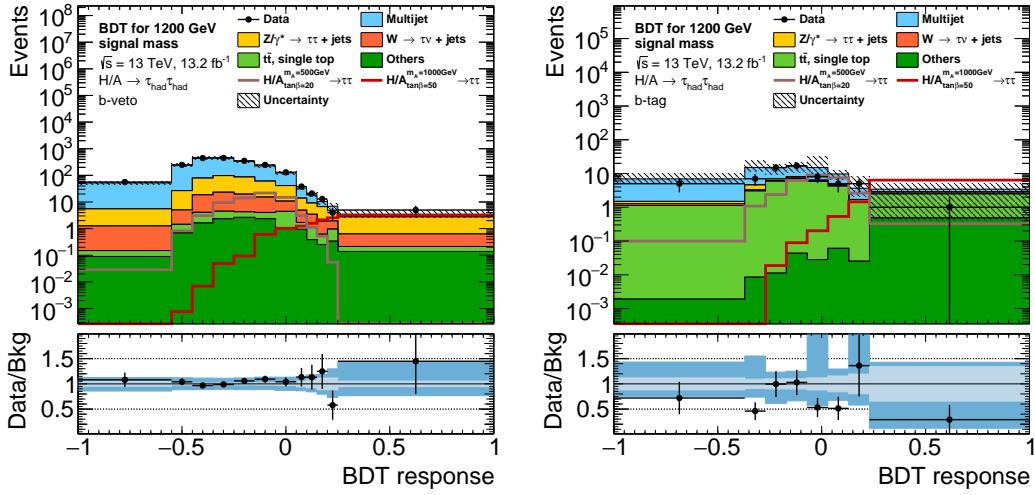


Figure A.6.3: Signal region BDT response of BDTs trained with classic splitting for 1200 GeV signal mass in B-VETO (left) and B-TAG (right).

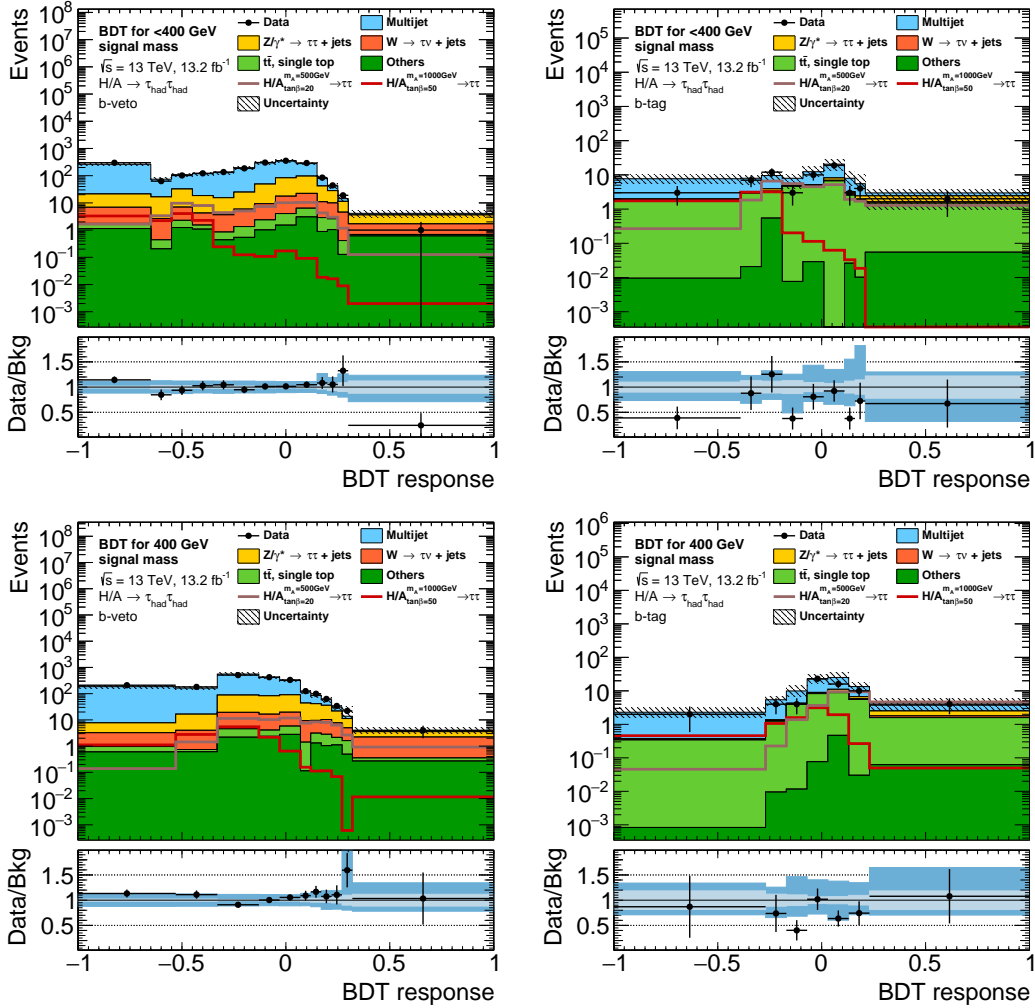


Figure A.6.4: Signal region BDT response of BDTs trained using K-Fold method for 300, 350, 400 GeV signal mass in B-VETO (left) and B-TAG (right).

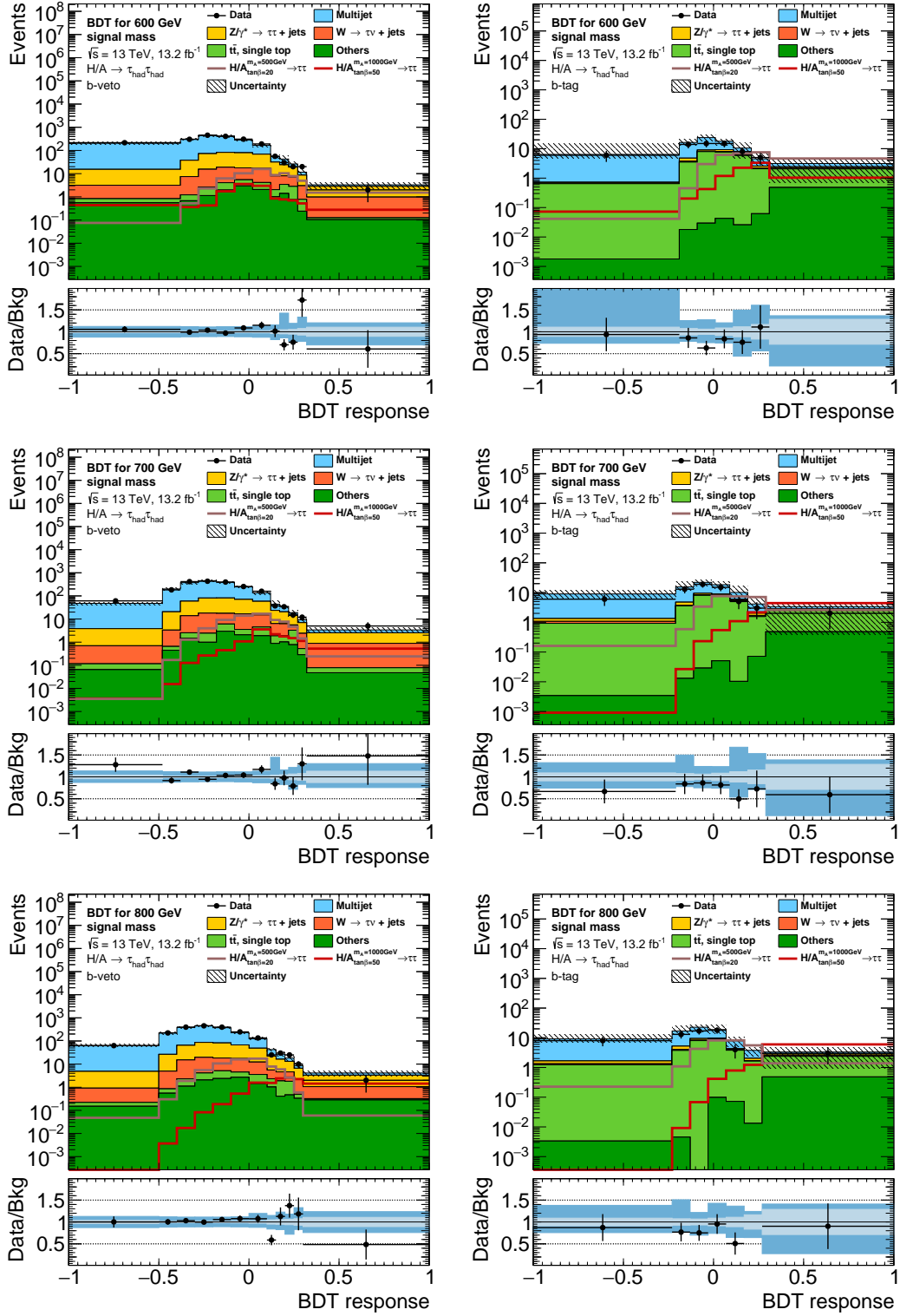


Figure A.6.5: Signal region BDT response of BDTs trained using K-Fold method for 600, 700, 800 GeV signal mass in B-VETO (left) and B-TAG (right).

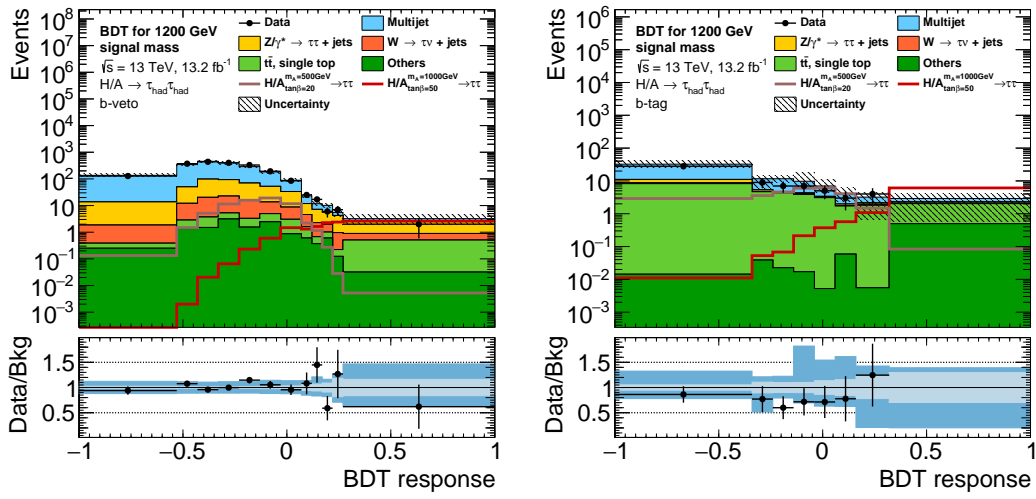


Figure A.6.6: Signal region BDT response of BDTs trained using K-Fold method for 1200 GeV signal mass in B-VETO (left) and B-TAG (right).





# B Limit

## B.1 Uncertainties

To set upper limits on the signal strength parameter  $\mu$  in a given confidence interval, systematic uncertainties have to be considered. Most of the occurring uncertainties can be treated as normalization uncertainty. The transformation function  $\eta_{sp}(\alpha_p)$  models the impact of the normalization shift using the associated nuisance parameter. The upper and lower values  $\eta_{sp}^{\pm}$  of the uncertainties, corresponding to a  $\pm 1\sigma$  shift of  $\alpha_p$ , are listed in Table B.1.1 for the B-TAG region and in Table B.1.2 for the B-VETO region.

The Uncertainties are described below:

### **Fake factor Uncertainty:**

The uncertainties on the fake factor measurements are driven by the statistical uncertainties in the control region where the fake factor is measured. It is the leading uncertainty for B-TAG and B-VETO category.

### **Fake rate uncertainty:**

Similar to the fake factor measurements, an uncertainty, mainly driven by statistical uncertainty in the fake rate region, is applied for fake rates.

### **Flavor tagging efficiency:**

Since b-tags are applied on jets, efficiency scale factors on the flavor tagging have to be applied. The uncertainties on the corresponding scale factors are combined in independent sets of nuisance parameters.

### **Trigger efficiency:**

Trigger decisions are emulated on Monte Carlo samples. To account for differences in the efficiency between recorded data and simulated samples scale factors are applied. Since different trigger setups are used in the data taking periods of 2015 and 2016, independent nuisance parameters are used to model the systematic uncertainties of the scale factors.

### **Tau-ID efficiency:**

Uncertainties on the tau identification based on a BDT are modeled by two different nuisance parameters, one for high- $p_T$  tau candidates and one combined total nuisance parameter.

### **Tau energy scale:**

The Tau energy scale calibrates the energy derived from the calorimeter deposits to the true value of the tau lepton using MC samples. Differences between data and MC are modeled using scale factors with corresponding systematic uncertainties. The different sources of uncertainties, namely the modeling of the detector material, calorimeter

performance and calibration, pile-up and interactions of the decay products with the detector, are modeled using a set of independent nuisance parameters.

**Jet reconstruction and calibration nuisance parameter:**

For the reconstruction and calibration of jets, scale factors on the jet energy resolution and vertex tagging are applied. Uncertainties on these scale factors are covered by the corresponding nuisance parameters.

**Cross section uncertainty:**

Events simulated using Monte Carlo generators have to be scaled to match a certain cross section. Uncertainties on these cross sections mainly arise from theoretical calculations.

**Luminosity uncertainty:**

All backgrounds and signal distributions are scaled to match the expected luminosity recorded with the ATLAS detector. Thus for all processes the uncertainty on the luminosity measurement has to be taken into account.

**Pile-up reweighting uncertainty:**

Since the MC event samples are processed in advance with a expected pile-up profile, the actual profile is modeled using scale factors. This procedure is known as pile-up reweighting (PRW). These scale factors are affected by systematic uncertainties that have to be propagated for all MC samples.

Name	$\eta_{sp}^+$	$\eta_{sp}^-$
<b>Top</b>		
Cross Section Uncertainty Top	1.060	0.940
Fake Rate Systematic	1.044	0.939
Flavor Tagging Eff. NP B 0	0.968	1.031
Jet Vertex Tagger Eff. NP	1.043	0.959
Luminosity Uncertainty	1.028	0.972
PRW Scale Factor Uncertainty	0.977	0.991
Tau Energy Scale NP. Detector	1.164	0.887
Tau Energy Scale NP. Insitu	1.166	0.903
Tau Energy Scale NP. Model	1.053	0.993
Tau-ID Eff. NP High-PT	1.011	0.989
Tau-ID Eff. NP Total	1.093	0.947
Tau-Reconstruction Eff. NP Total	1.039	0.961
Top Normalisation Uncertainty	1.259	0.732
Trigger Eff. NP. Data 2015	1.031	0.957
Trigger Eff. NP. MC 2015	1.030	0.947
Trigger Eff. NP. Total 2016	1.160	0.851
t $\bar{t}$ shower shape Uncertainty	1.334	0.666
<b>Z <math>\rightarrow</math> <math>\tau\tau</math></b>		
Cross Section Uncertainty Z-Boson	1.050	0.950
Flavor Tagging Eff. NP B 0	0.982	1.018
Flavor Tagging Eff. NP C 0	0.957	1.043
Flavor Tagging Eff. NP C 1	1.021	0.979
Flavor Tagging Eff. NP Light 0	0.884	1.117
Flavor Tagging Eff. NP Light 1	0.980	1.020
Flavor Tagging Eff. NP Light 2	0.984	1.016
Flavor Tagging Eff. NP Light 3	0.990	1.010
Flavor Tagging Eff. NP extrapol C	1.016	0.984
Jet Grouped NP	1.019	0.993
Jet Vertex Tagger Eff. NP	1.026	0.974
Luminosity Uncertainty	1.028	0.972
PRW Scale Factor Uncertainty	1.041	1.021
Tau Energy Scale NP. Detector	1.112	0.778
Tau Energy Scale NP. Insitu	0.988	0.818
Tau Energy Scale NP. Model	0.996	1.000
Tau-ID Eff. NP High-PT	1.016	0.985
Tau-ID Eff. NP Total	1.103	0.942
Tau-Reconstruction Eff. NP Total	1.044	0.957
Trigger Eff. NP. Data 2015	1.030	0.956
Trigger Eff. NP. MC 2015	1.033	0.940
Trigger Eff. NP. Total 2016	1.175	0.839
<b>W <math>\rightarrow</math> <math>\tau\nu</math></b>		
Cross Section Uncertainty W-Boson	1.050	0.950
Fake Rate Systematic	1.108	0.892
Flavor Tagging Eff. NP B 0	0.968	1.031
Flavor Tagging Eff. NP C 0	0.980	1.018
Flavor Tagging Eff. NP C 1	1.013	0.987
Flavor Tagging Eff. NP Light 0	0.948	1.051
Flavor Tagging Eff. NP Light 2	0.989	1.011
Jet Energy Resolution NP	1.037	0.963
Jet Vertex Tagger Eff. NP	1.050	0.952
Luminosity Uncertainty	1.028	0.972
PRW Scale Factor Uncertainty	1.049	0.985
Tau Energy Scale NP. Detector	1.052	0.886
Tau Energy Scale NP. Insitu	1.032	0.905
Tau Energy Scale NP. Model	1.031	0.986
Tau-ID Eff. NP Total	1.051	0.970
Tau-Reconstruction Eff. NP Total	1.022	0.978
Trigger Eff. NP. Data 2015	1.019	0.974
Trigger Eff. NP. MC 2015	1.017	0.967
Trigger Eff. NP. Total 2016	1.089	0.918

Name	$\eta_{sp}^+$	$\eta_{sp}^-$
<b>Multi-jet</b>		
Fake Factor Uncertainty	1.259	0.741
Luminosity Uncertainty	1.040	0.960
<b>Others</b>		
Cross Section Uncertainty Di-Boson	1.060	0.940
Fake Rate Systematic	1.027	0.974
Flavor Tagging Eff. NP B 0	0.991	1.008
Flavor Tagging Eff. NP Light 0	0.686	1.314
Flavor Tagging Eff. NP Light 1	0.987	1.013
Flavor Tagging Eff. NP Light 2	1.021	0.979
Jet Grouped NP	1.102	0.920
Jet Vertex Tagger Eff. NP	1.014	0.986
Luminosity Uncertainty	1.028	0.972
PRW Scale Factor Uncertainty	0.950	0.924
Tau Energy Scale NP. Detector	1.045	1.030
Tau-ID Eff. NP High-PT	1.053	0.948
Tau-ID Eff. NP Total	1.093	0.945
Tau-Reconstruction Eff. NP High-PT	1.030	0.970
Tau-Reconstruction Eff. NP Total	1.041	0.960
Trigger Eff. NP. Data 2015	1.037	0.952
Trigger Eff. NP. MC 2015	1.031	0.955
Trigger Eff. NP. Total 2016	1.195	0.811
<b>ggH for <math>m_{A,H} = 500</math> GeV</b>		
Acceptance Uncertainty ggH	1.155	0.845
Flavor Tagging Eff. NP B 0	0.968	1.031
Flavor Tagging Eff. NP C 0	0.972	1.028
Flavor Tagging Eff. NP C 1	1.019	0.981
Flavor Tagging Eff. NP Light 0	0.943	1.057
Jet Energy Resolution NP	0.990	1.010
Jet Grouped NP	1.044	0.963
Jet Vertex Tagger Eff. NP	1.037	0.964
Luminosity Uncertainty	1.028	0.972
PRW Scale Factor Uncertainty	0.987	1.004
Tau Energy Scale NP. Detector	1.025	0.949
Tau Energy Scale NP. Insitu	0.955	0.909
Tau-ID Eff. NP High-PT	1.020	0.980
Tau-ID Eff. NP Total	1.108	0.940
Tau-Reconstruction Eff. NP Total	1.044	0.956
Trigger Eff. NP. Data 2015	1.035	0.953
Trigger Eff. NP. MC 2015	1.028	0.949
Trigger Eff. NP. Total 2016	1.231	0.783
<b>bbH for <math>m_{A,H} = 500</math> GeV</b>		
Acceptance Uncertainty bbH	1.220	0.780
Flavor Tagging Eff. NP B 0	0.951	1.048
Jet Grouped NP	1.022	0.973
Jet Vertex Tagger Eff. NP	1.036	0.966
Luminosity Uncertainty	1.028	0.972
Tau Energy Scale NP. Detector	1.022	0.918
Tau Energy Scale NP. Insitu	1.025	0.914
Tau-ID Eff. NP High-PT	1.022	0.978
Tau-ID Eff. NP Total	1.158	0.891
Tau-Reconstruction Eff. NP Total	1.146	0.860
Trigger Eff. NP. Data 2015	1.037	0.951
Trigger Eff. NP. MC 2015	1.032	0.948
Trigger Eff. NP. Total 2016	1.220	0.791

Table B.1.1: Table of systematic uncertainties for B-TAG region

Name	$\eta_{sp}^+$	$\eta_{sp}^-$
<b>Top</b>		
Cross Section Uncertainty Top	1.060	0.940
Fake Rate Systematic	1.055	0.925
Flavor Tagging Eff. NP B 0	1.084	0.919
Flavor Tagging Eff. NP B 1	0.980	1.020
Jet Energy Resolution NP	0.982	1.018
Jet Vertex Tagger Eff. NP	1.038	0.963
Luminosity Uncertainty	1.028	0.972
PRW Scale Factor Uncertainty	0.979	1.006
Tau Energy Scale NP. Detector	1.004	0.895
Tau Energy Scale NP. Insitu	1.106	0.950
Tau Energy Scale NP. Model	1.010	0.988
Tau-ID Eff. NP Total	1.089	0.949
Tau-Reconstruction Eff. NP Total	1.038	0.963
Top Normalisation Uncertainty	1.259	0.732
Trigger Eff. NP. Data 2015	1.029	0.960
Trigger Eff. NP. MC 2015	1.026	0.951
Trigger Eff. NP. Total 2016	1.174	0.839
t $\bar{t}$ shower shape Uncertainty	1.047	0.953
<b>Z<math>\rightarrow\tau\tau</math></b>		
Cross Section Uncertainty Z-Boson	1.050	0.950
Jet Vertex Tagger Eff. NP	1.013	0.987
Luminosity Uncertainty	1.028	0.972
Tau Energy Scale NP. Detector	1.125	0.885
Tau Energy Scale NP. Insitu	1.112	0.898
Tau Energy Scale NP. Model	1.014	0.982
Tau-ID Eff. NP High-PT	1.016	0.985
Tau-ID Eff. NP Total	1.105	0.942
Tau-Reconstruction Eff. NP Total	1.046	0.955
Trigger Eff. NP. Data 2015	1.033	0.953
Trigger Eff. NP. MC 2015	1.033	0.941
Trigger Eff. NP. Total 2016	1.177	0.835
<b>W<math>\rightarrow\tau\nu</math></b>		
Cross Section Uncertainty W-Boson	1.050	0.950
Fake Rate Systematic	1.110	0.890
Jet Vertex Tagger Eff. NP	1.023	0.978
Luminosity Uncertainty	1.028	0.972
Tau Energy Scale NP. Detector	1.063	0.934
Tau Energy Scale NP. Insitu	1.080	0.921
Tau-ID Eff. NP Total	1.050	0.970
Tau-Reconstruction Eff. NP Total	1.023	0.977
Trigger Eff. NP. Data 2015	1.010	0.985
Trigger Eff. NP. MC 2015	1.011	0.980
Trigger Eff. NP. Total 2016	1.055	0.949

Name	$\eta_{sp}^+$	$\eta_{sp}^-$
<b>Multi-jet</b>		
Fake Factor Uncertainty	1.126	0.874
Luminosity Uncertainty	1.040	0.960
<b>Others</b>		
Cross Section Uncertainty Di-Boson	1.060	0.940
Fake Rate Systematic	1.057	0.945
Jet Vertex Tagger Eff. NP	1.023	0.977
Luminosity Uncertainty	1.028	0.972
PRW Scale Factor Uncertainty	1.074	0.955
Tau Energy Scale NP. Detector	1.122	0.929
Tau Energy Scale NP. Insitu	1.218	0.960
Tau Energy Scale NP. Model	0.999	1.000
Tau-ID Eff. NP High-PT	1.010	0.990
Tau-ID Eff. NP Total	1.073	0.959
Tau-Reconstruction Eff. NP Total	1.032	0.969
Trigger Eff. NP. Data 2015	1.024	0.967
Trigger Eff. NP. MC 2015	1.023	0.960
Trigger Eff. NP. Total 2016	1.127	0.882
<b>ggH for <math>m_{A,H} = 500</math> GeV</b>		
Acceptance Uncertainty ggH	1.155	0.845
Jet Vertex Tagger Eff. NP	1.018	0.983
Luminosity Uncertainty	1.028	0.972
PRW Scale Factor Uncertainty	1.013	0.997
Tau Energy Scale NP. Detector	1.029	0.965
Tau Energy Scale NP. Insitu	1.024	0.963
Tau-ID Eff. NP High-PT	1.020	0.980
Tau-ID Eff. NP Total	1.107	0.940
Tau-Reconstruction Eff. NP Total	1.044	0.957
Trigger Eff. NP. Data 2015	1.037	0.950
Trigger Eff. NP. MC 2015	1.032	0.947
Trigger Eff. NP. Total 2016	1.218	0.793
<b>bbH for <math>m_{A,H} = 500</math> GeV</b>		
Acceptance Uncertainty bbH	1.220	0.780
Flavor Tagging Eff. NP B 0	1.023	0.977
Jet Grouped NP	0.985	1.016
Jet Vertex Tagger Eff. NP	1.019	0.982
Luminosity Uncertainty	1.028	0.972
Tau Energy Scale NP. Detector	1.026	0.959
Tau Energy Scale NP. Insitu	1.039	0.970
Tau-ID Eff. NP High-PT	1.020	0.980
Tau-ID Eff. NP Total	1.158	0.891
Tau-Reconstruction Eff. NP Total	1.147	0.859
Trigger Eff. NP. Data 2015	1.036	0.951
Trigger Eff. NP. MC 2015	1.032	0.947
Trigger Eff. NP. Total 2016	1.216	0.795

Table B.1.2: Table of systematic uncertainties for B-VETO region

## B.2 Limit result tables

$m_A$ [GeV]	300	350	400	500	600	700	800	1000	1200
BDT exp	12.61	12.55	13.7	15.44	19.15	23.78	26.77	35.87	46.76
+1 $\sigma$ /2 $\sigma$	+2.12/6.53	+1.78/4.73	+1.23/4.93	+3.12/6.58	+2.86/7.24	+3.72/9.86	+4.4/11.92	+7.44/19.1	+11.42/22.72
-1 $\sigma$ /2 $\sigma$	-1.91/2.76	-1.81/2.79	-1.24/2.75	-0.98/1.76	-1.57/3.5	-2.42/4.28	-2.67/4.64	-3.67/6.75	-5.43/9.5
$m_{\text{T}}^{\text{tot}}$ exp	12.11	12.74	14.04	18.24	20.27	23.9	27.58	38.19	52.25
+1 $\sigma$ /2 $\sigma$	+1.85/3.99	+1.56/4.0	+1.43/4.79	+1.92/6.18	+3.42/8.2	+3.59/9.45	+4.54/11.53	+7.35/18.78	+10.82/18.93
-1 $\sigma$ /2 $\sigma$	-2.1/2.33	-1.75/2.93	-1.16/2.58	-2.03/3.51	-1.39/2.88	-2.37/4.33	-3.02/4.85	-4.25/7.7	-6.35/11.08
BDT obs	11.1	11.83	13.04	15.12	17.51	24.4	27.3	37.33	43.71
$m_{\text{T}}^{\text{tot}}$ obs	11.19	13.69	14.53	16.08	18.25	21.56	24.67	34.17	46.93

Table B.2.1: Table of the upper limit values. Compared are the expected and observed limits using the response distribution of the BDTs trained on the classic splitting method with the limit derived from the  $m_{\text{T}}^{\text{tot}}$  distribution

$m_A$ [GeV]	300	350	400	500	600	700	800	1000	1200
BDT exp	11.04	11.46	13.03	16.82	19.4	23.46	26.95	38.64	49.26
+1 $\sigma$ /2 $\sigma$	+2.63/5.45	+2.21/4.25	+1.5/4.53	+2.52/6.4	+3.35/7.99	+3.66/9.71	+5.01/12.38	+8.44/21.5	+11.94/21.66
-1 $\sigma$ /2 $\sigma$	-1.17/1.43	-1.57/1.94	-1.51/3.07	-1.97/2.62	-1.48/3.29	-2.54/4.14	-2.74/4.69	-4.37/7.68	-5.78/10.26
$m_{\Gamma}^{\text{tot}}$ exp	12.11	12.74	14.04	18.24	20.27	23.9	27.58	38.19	52.25
+1 $\sigma$ /2 $\sigma$	+1.85/3.99	+1.56/4.0	+1.43/4.79	+1.92/6.18	+3.42/8.2	+3.59/9.45	+4.54/11.53	+7.35/18.78	+10.82/18.93
-1 $\sigma$ /2 $\sigma$	-2.1/2.33	-1.75/2.93	-1.16/2.58	-2.03/3.51	-1.39/2.88	-2.37/4.33	-3.02/4.85	-4.25/7.7	-6.35/11.08
BDT obs	9.99	10.36	13.99	14.97	18.21	22.72	24.83	34.0	39.23
$m_{\Gamma}^{\text{tot}}$ obs	11.19	13.69	14.53	16.08	18.25	21.56	24.67	34.17	46.93

Table B.2.2: Table of the upper limit values. Compared are the expected and observed limits using the response distribution of the BDTs trained on the K-Fold method with the limit derived from the  $m_{\Gamma}^{\text{tot}}$  distribution





# Bibliography

- [1] J. J. Thomson, “Cathode Rays”, *Philosophical Magazine* *44*.293 (1897) .
- [2] S. Tomonaga, “On a relativistically invariant formulation of the quantum theory of wave fields”, *Prog. Theor. Phys.* **1** (1946) 27–42.
- [3] J. Schwinger, “Quantum electrodynamics. i. a covariant formulation”, *Phys. Rev.* **74** (Nov, 1948) 1439–1461.  
<https://link.aps.org/doi/10.1103/PhysRev.74.1439>.
- [4] R. P. Feynman, “Mathematical formulation of the quantum theory of electromagnetic interaction”, *Phys. Rev.* **80** (1950) 440–457.
- [5] A. Salam, “Weak and electromagnetic interactions in elementary particle physics: relativistic groups and analyticity”, 1968. Almqvist and Wiksell, Stockholm proceedings of the eighth Nobel symposium, 1968. p. 367.
- [6] S. Weinberg, “A Model of Leptons”, *Phys. Rev. Lett.* **19** (1967) 1264–1266.
- [7] S. L. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl. Phys.* **22** (1961) 579–588.
- [8] P. W. Higgs, “Broken symmetries and the masses of gauge bosons”, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509.  
<http://link.aps.org/doi/10.1103/PhysRevLett.13.508>.
- [9] P. W. Higgs, “Spontaneous Symmetry Breakdown without Massless Bosons”, *Phys. Rev.* **145** (1966) 1156–1163.
- [10] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons”, *Phys. Rev. Lett.* **13** (Aug, 1964) 321–323.  
<http://link.aps.org/doi/10.1103/PhysRevLett.13.321>.
- [11] **ATLAS** Collaboration, G. Aad *et al.*, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett.* **B716** (2012) 1–29 CERN-PH-EP-2012-218, [arXiv:1207.7214](https://arxiv.org/abs/1207.7214) [hep-ex].
- [12] **CMS** Collaboration, S. Chatrchyan *et al.*, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett.* **B716** (2012) 30–61 CMS-HIG-12-028, CERN-PH-EP-2012-220, [arXiv:1207.7235](https://arxiv.org/abs/1207.7235) [hep-ex].
- [13] D. Griffiths, “Introduction to elementary particles”. New York: Wiley, 2008.
- [14] J. Greensite, “An introduction to the confinement problem”. Springer Berlin Heidelberg, 2011.

- [15] **Particle Data Group** Collaboration, K. A. Olive *et al.*, “Review of Particle Physics”, *Chin. Phys.* **C38** (2014) 090001.
- [16] M. Gell-Mann, “A schematic model of baryons and mesons”, *Physics Letters* **8** no. 3, (1964) 214 – 215. <http://www.sciencedirect.com/science/article/pii/S0031916364920013>.
- [17] G. Zweig, “An SU(3) model for strong interaction symmetry and its breaking. Version 2”, *Developments in the Quark Theory of Hadrons, Volume 1. Edited by D. Lichtenberg and S. Rosen. pp. 22-101* (1964) 22–101. <http://inspirehep.net/record/4674/files/cern-th-412.pdf>.
- [18] M. Drewes, “The Phenomenology of Right Handed Neutrinos”, *Int. J. Mod. Phys.* **E22** (2013) 1330019 TUM-HEP-881-13, [arXiv:1303.6912](https://arxiv.org/abs/1303.6912) [hep-ph].
- [19] M. C. Gonzalez-Garcia and M. Maltoni, “Phenomenology with Massive Neutrinos”, *Phys. Rept.* **460** (2008) 1–129 YITP-SB-07-14, UB-ECM-PF-07-08, [arXiv:0704.1800](https://arxiv.org/abs/0704.1800) [hep-ph].
- [20] B. T. Cleveland, T. Daily, J. Raymond Davis, J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman, “Measurement of the solar electron neutrino flux with the homestake chlorine detector”, *The Astrophysical Journal* **496** no. 1, (1998) 505. <http://stacks.iop.org/0004-637X/496/i=1/a=505>.
- [21] U. Seljak, A. Slosar, and P. McDonald, “Cosmological parameters from combining the Lyman-alpha forest with CMB, galaxy clustering and SN constraints”, *JCAP* **0610** (2006) 014, [arXiv:astro-ph/0604335](https://arxiv.org/abs/astro-ph/0604335) [astro-ph].
- [22] A. Djouadi, “The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model”, *Phys. Rept.* **457** (2008) 1–216 LPT-ORSAY-05-17, [arXiv:hep-ph/0503172](https://arxiv.org/abs/hep-ph/0503172) [hep-ph].
- [23] **ATLAS** Collaboration, G. Aad *et al.*, “Evidence for the spin-0 nature of the Higgs boson using ATLAS data”, *Phys. Lett.* **B726** (2013) 120–144 CERN-PH-EP-2013-102, [arXiv:1307.1432](https://arxiv.org/abs/1307.1432) [hep-ex].
- [24] **CMS** Collaboration, C. Collaboration, “Properties of the observed Higgs-like resonance using the diphoton channel”, CMS-PAS-HIG-13-016, 2013. <https://cds.cern.ch/record/1558930>.
- [25] H. Georgi, “Towards a grand unified theory of flavor”, *Nuclear Physics B* **156** no. 1, (1979) 126 – 134.
- [26] E. Corbelli and P. Salucci, “The Extended Rotation Curve and the Dark Matter Halo of M33”, *Mon. Not. Roy. Astron. Soc.* **311** (2000) 441–447, [arXiv:astro-ph/9909252](https://arxiv.org/abs/astro-ph/9909252) [astro-ph].
- [27] D. C. et al., “A direct empirical proof of the existence of dark matter”, *The Astrophysical Journal Letters* **648** no. 2, (2006) L109. <http://stacks.iop.org/1538-4357/648/i=2/a=L109>.
- [28] S. P. Martin, “A Supersymmetry primer”, FERMILAB-PUB-97-425-T, [arXiv:hep-ph/9709356](https://arxiv.org/abs/hep-ph/9709356) [hep-ph], 1997. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

- [29] A. Djouadi, “The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model”, *Phys. Rept.* **459** (2008) 1–241 LPT-ORSAY-05-18, [arXiv:hep-ph/0503173](#) [hep-ph].
- [30] **ATLAS, CMS** Collaboration, G. Aad *et al.*, “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 ATLAS-HIGG-2014-14, CMS-HIG-14-042, CERN-PH-EP-2015-075, [arXiv:1503.07589](#) [hep-ex].
- [31] M. Carena, S. Heinemeyer, O. Stål, C. E. M. Wagner, and G. Weiglein, “MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios after the Discovery of a Higgs-like Particle”, *Eur. Phys. J.* **C73** no. 9, (2013) 2552 ANL-HEP-PR-13-12, EFI-13-2, DESY-13-024, FERMILAB-PUB-13-041-T, [arXiv:1302.7033](#) [hep-ph].
- [32] H. Bahl and W. Hollik, “Precise prediction for the light MSSM Higgs boson mass combining effective field theory and fixed-order calculations”, *Eur. Phys. J.* **C76** no. 9, (2016) 499, [arXiv:1608.01880](#) [hep-ph].
- [33] T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, and G. Weiglein, “High-Precision Predictions for the Light CP -Even Higgs Boson Mass of the Minimal Supersymmetric Standard Model”, *Phys. Rev. Lett.* **112** no. 14, (2014) 141801 DESY-13-248, FR-PHENO-2013-018, MPP-2013-317, [arXiv:1312.4937](#) [hep-ph].
- [34] M. Frank, T. Hahn, S. Heinemeyer, W. Hollik, H. Rzehak, and G. Weiglein, “The Higgs Boson Masses and Mixings of the Complex MSSM in the Feynman-Diagrammatic Approach”, *JHEP* **02** (2007) 047 DCPT-06-160, IPPP-06-80, MPP-2006-158, PSI-PR-06-14, [arXiv:hep-ph/0611326](#) [hep-ph].
- [35] G. Degrassi, S. Heinemeyer, W. Hollik, P. Slavich, and G. Weiglein, “Towards high precision predictions for the MSSM Higgs sector”, *Eur. Phys. J.* **C28** (2003) 133–143 DCPT-02-126, IPPP-02-63, LMU-11-02, MPI-PHT-2002-73, RM3-TH-02-19, [arXiv:hep-ph/0212020](#) [hep-ph].
- [36] S. Heinemeyer, W. Hollik, and G. Weiglein, “The Masses of the neutral CP -even Higgs bosons in the MSSM: Accurate analysis at the two loop level”, *Eur. Phys. J.* **C9** (1999) 343–366 KA-TP-17-1998, DESY-98-194, CERN-TH-98-405, [arXiv:hep-ph/9812472](#) [hep-ph].
- [37] S. Heinemeyer, W. Hollik, and G. Weiglein, “FeynHiggs: A Program for the calculation of the masses of the neutral CP even Higgs bosons in the MSSM”, *Comput. Phys. Commun.* **124** (2000) 76–89 KA-TP-16-1998, DESY-98-193, CERN-TH-98-389, [arXiv:hep-ph/9812320](#) [hep-ph].
- [38] R. V. Harlander, S. Liebler, and H. Mantler, “SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM”, *Comput. Phys. Commun.* **184** (2013) 1605–1617 WUB-12-28, LPN12-134, [arXiv:1212.3249](#) [hep-ph].
- [39] L. Evans and P. Bryant, “LHC Machine”, *JINST* **3** (2008) S08001.

- [40] “LEP design report”. CERN, Geneva, 1984.  
<http://cds.cern.ch/record/102083>. Copies shelved as reports in LEP, PS and SPS libraries.
- [41] C. De Melis, “The CERN accelerator complex. Complexe des accélérateurs du CERN”, OPEN-PHO-ACCEL-2016-001, Jan, 2016.  
<https://cds.cern.ch/record/2119882>. General Photo.
- [42] **ATLAS** Collaboration, G. Aad *et al.*, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* **3** (2008) S08003.
- [43] **CMS** Collaboration, S. Chatrchyan *et al.*, “The CMS Experiment at the CERN LHC”, *JINST* **3** (2008) S08004.
- [44] **ALICE** Collaboration, K. Aamodt *et al.*, “The ALICE experiment at the CERN LHC”, *JINST* **3** (2008) S08002.
- [45] **LHCb** Collaboration, A. A. Alves, Jr. *et al.*, “The LHCb Detector at the LHC”, *JINST* **3** (2008) S08005.
- [46] **LHCf** Collaboration, O. Adriani *et al.*, “The LHCf detector at the CERN Large Hadron Collider”, *JINST* **3** (2008) S08006.
- [47] **TOTEM** Collaboration, G. Anelli *et al.*, “The TOTEM experiment at the CERN Large Hadron Collider”, *JINST* **3** (2008) S08007.
- [48] J. L. Pinfold, “The MoEDAL Experiment at the LHC a New Light on the Terascale Frontier”, *J. Phys. Conf. Ser.* **631** no. 1, (2015) 012014.
- [49] **ATLAS** Collaboration, “Accelerator performance and statistics”, 2017.  
<https://acc-stats.web.cern.ch/acc-stats/#lhc/super-table>. Accessed: 2017-03-13.
- [50] **ATLAS** Collaboration, “Atlas experiment - public results”, 2017.  
<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>. Accessed: 2017-03-13.
- [51] **ATLAS** Collaboration, G. Aad *et al.*, “Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics”, SLAC-R-980, CERN-OPEN-2008-020, [arXiv:0901.0512](https://arxiv.org/abs/0901.0512) [hep-ex], 2009.
- [52] **ATLAS** Collaboration, “Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at  $\sqrt{s}=13$  TeV”, tech. rep., CERN, Geneva, Jul, 2015.  
<http://cds.cern.ch/record/2037904>.
- [53] M. Capeans *et al.*, “ATLAS Insertable B-Layer Technical Design Report”, tech. rep., Sep, 2010. <https://cds.cern.ch/record/1291633>.
- [54] Y. Nakahama, “The ATLAS Trigger System: Ready for Run-2”, *J. Phys. Conf. Ser.* **664** no. 8, (2015) 082037.
- [55] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026 FERMILAB-PUB-06-052-CD-T, LU-TP-06-13, [arXiv:hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175) [hep-ph].

- [56] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, “An Introduction to PYTHIA 8.2”, *Comput. Phys. Commun.* **191** (2015) 159–177 LU-TP-14-36, MCNET-14-22, CERN-PH-TH-2014-190, FERMILAB-PUB-14-316-CD, DESY-14-178, SLAC-PUB-16122, [arXiv:1410.3012](https://arxiv.org/abs/1410.3012) [hep-ph].
- [57] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert, and J. Winter, “Event generation with SHERPA 1.1”, *JHEP* **02** (2009) 007 FERMILAB-PUB-08-477-T, SLAC-PUB-13420, ZU-TH-17-08, DCPT-08-138, IPPP-08-69, EDINBURGH-2008-30, MCNET-08-14, [arXiv:0811.4622](https://arxiv.org/abs/0811.4622) [hep-ph].
- [58] LHC Higgs Cross Section Working Group, S. Heinemeyer, C. Mariotti, G. Passarino, and R. Tanaka (Eds.), “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”, *CERN-2013-004* (CERN, Geneva, 2013) , [arXiv:1307.1347](https://arxiv.org/abs/1307.1347) [hep-ph].
- [59] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, “FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order”, *Comput. Phys. Commun.* **182** (2011) 2388–2403 ANL-HEP-PR-10-60, [arXiv:1011.3540](https://arxiv.org/abs/1011.3540) [hep-ph].
- [60] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, “High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO”, *Phys. Rev.* **D69** (2004) 094008 SLAC-PUB-10288, UH-511-1042-03, [arXiv:hep-ph/0312266](https://arxiv.org/abs/hep-ph/0312266) [hep-ph].
- [61] K. Melnikov and F. Petriello, “Electroweak gauge boson production at hadron colliders through  $O(\alpha(s)^2)$ ”, *Phys. Rev.* **D74** (2006) 114017 UH-511-1092-06, [arXiv:hep-ph/0609070](https://arxiv.org/abs/hep-ph/0609070) [hep-ph].
- [62] E. Re, “Single-top Wt-channel production matched with parton showers using the POWHEG method”, *Eur. Phys. J.* **C71** (2011) 1547, [arXiv:1009.2450](https://arxiv.org/abs/1009.2450) [hep-ph].
- [63] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, “Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations”, *JHEP* **03** (2013) 015 NIKHEF-2012-021, CERN-PH-TH-2012-329, [arXiv:1212.3460](https://arxiv.org/abs/1212.3460) [hep-ph].
- [64] D. J. Lange, “The evtgen particle decay simulation package”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **462** no. 12, (2001) 152 – 155. <http://www.sciencedirect.com/science/article/pii/S0168900201000894>. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.
- [65] S. Agostinelli *et al.*, “Geant4a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** no. 3, (2003) 250 – 303. <http://www.sciencedirect.com/science/article/pii/S0168900203013688>.

- [66] **ATLAS** Collaboration, G. Aad *et al.*, “The ATLAS Simulation Infrastructure”, *Eur. Phys. J.* **C70** (2010) 823–874, [arXiv:1005.4568](#) [[physics.ins-det](#)].
- [67] **ATLAS** Collaboration, A. Collaboration, “The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim”, ATL-PHYS-PUB-2010-013, ATL-COM-PHYS-2010-838, 2010.
- [68] P. Nason, “A New method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP* **11** (2004) 040 BICOCCA-FT-04-11, [arXiv:hep-ph/0409146](#) [[hep-ph](#)].
- [69] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070 BICOCCA-FT-07-9, GEF-TH-21-2007, [arXiv:0709.2092](#) [[hep-ph](#)].
- [70] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP* **06** (2010) 043 DESY-10-018, SFB-CPP-10-22, IPPP-10-11, DCPT-10-22, [arXiv:1002.2581](#) [[hep-ph](#)].
- [71] E. Bagnaschi, G. Degrossi, P. Slavich, and A. Vicini, “Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM”, *JHEP* **02** (2012) 088, [arXiv:1111.2854](#) [[hep-ph](#)].
- [72] J. Alwall *et al.*, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079 CERN-PH-TH-2014-064, CP3-14-18, LPN14-066, MCNET-14-09, ZU-TH-14-14, [arXiv:1405.0301](#) [[hep-ph](#)].
- [73] M. Wiesemann, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, and P. Torrielli, “Higgs production in association with bottom quarks”, *JHEP* **02** (2015) 132 CERN-PH-TH-2014-182, CP3-14-64, LPN14-114, MCNET-14-20, ZU-TH-33-14, [arXiv:1409.5301](#) [[hep-ph](#)].
- [74] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO vector-boson production matched with shower in POWHEG”, *JHEP* **07** (2008) 060, [arXiv:0805.4802](#) [[hep-ph](#)].
- [75] S. Alioli, P. Nason, C. Oleari, and E. Re, “Vector boson plus one jet production in POWHEG”, *JHEP* **01** (2011) 095, [arXiv:1009.5594](#) [[hep-ph](#)].
- [76] T. Sjöstrand, S. Mrenna, and P. Skands, “A brief introduction to pythia 8.1”, *Computer Physics Communications* **178** no. 11, (2008) 852 – 867.
- [77] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with shower in POWHEG: s- and t-channel contributions”, *JHEP* **09** (2009) 111, [arXiv:0907.4076](#) [[hep-ph](#)]. [Erratum: *JHEP*02,011(2010)].
- [78] A. Buckley, T. Eifert, M. Elsing, D. Gillberg, K. Koeneke, A. Krasznahorkay, E. Moyses, M. Nowak, S. Snyder, and P. van Gemmeren, “Implementation of the ATLAS Run 2 event data model”, *J. Phys. Conf. Ser.* **664** no. 7, (2015) 072045.

- [79] **ATLAS** Collaboration, G. Aad *et al.*, “Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data”, *Eur. Phys. J. C* **74** no. 10, (2014) 3071 CERN-PH-EP-2014-153, [arXiv:1407.5063 \[hep-ex\]](#).
- [80] **ATLAS** Collaboration, “Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data”, ATLAS-CONF-2016-024, 2016.
- [81] **ATLAS** Collaboration, G. Aad *et al.*, “Muon reconstruction performance of the ATLAS detector in protonproton collision data at  $\sqrt{s} = 13$  TeV”, *Eur. Phys. J. C* **76** no. 5, (2016) 292 CERN-EP-2016-033, [arXiv:1603.05598 \[hep-ex\]](#).
- [82] **ATLAS** Collaboration, G. Aad *et al.*, “Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1”, CERN-PH-EP-2015-304, [arXiv:1603.02934 \[hep-ex\]](#), 2016.
- [83] M. Cacciari, G. P. Salam, and G. Soyez, “The Anti-k(t) jet clustering algorithm”, *JHEP* **04** (2008) 063 LPTHE-07-03, [arXiv:0802.1189 \[hep-ph\]](#).
- [84] **ATLAS** Collaboration, “Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC”, tech. rep., CERN, Geneva, Nov, 2015. <http://cds.cern.ch/record/2064383>.
- [85] **ATLAS** Collaboration, “Expected performance of the ATLAS *b*-tagging algorithms in Run-2”, tech. rep., CERN, Geneva, Jul, 2015. <http://cds.cern.ch/record/2037697>.
- [86] **ATLAS** Collaboration, “Commissioning of the ATLAS *b*-tagging algorithms using  $t\bar{t}$  events in early Run-2 data”, tech. rep., CERN, Geneva, Aug, 2015. <http://cds.cern.ch/record/2047871>.
- [87] **ATLAS** Collaboration, “Search for Minimal Supersymmetric Standard Model Higgs Bosons  $H/A$  in the  $\tau\tau$  final state in up to  $13.3 \text{ fb}^{-1}$  of pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS Detector”, tech. rep., CERN, Geneva, Aug, 2016. <https://cds.cern.ch/record/2206278>.
- [88] **ATLAS** Collaboration, “Tagging and suppression of pileup jets”, ATLAS-CONF-2014-018, 2014. <http://cds.cern.ch/record/1700870>.
- [89] T. Golling, H. S. Hayward, P. U. E. Onyisi, H. J. Stelzer, and P. Waller, “The ATLAS Data Quality Defect Database System”, *Eur. Phys. J. C* **72** (2012) 1960 EFI-11-30, [arXiv:1110.6119 \[physics.ins-det\]](#).
- [90] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, “Classification and regression trees”. New York: Chapman and Hall, 1993.
- [91] L. Ceriani and P. Verme, “The origins of the gini index: extracts from *variabilità e mutabilità* (1912) by corrado gini”, *The Journal of Economic Inequality* **10** no. 3, (2012) 421–443. <http://dx.doi.org/10.1007/s10888-011-9188-x>.
- [92] Y. Freund and R. E. Schapire, “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting”, *J. Comput. Syst. Sci.* **55** no. 1, (1997) 119–139.

- [93] A. Hoecker *et al.*, “TMVA - Toolkit for Multivariate Data Analysis”, *ArXiv Physics e-prints* (2007) , [arXiv:physics/0703039](https://arxiv.org/abs/physics/0703039) [physics].
- [94] R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection”, 1995.  
<http://dl.acm.org/citation.cfm?id=1643031.1643047>.
- [95] R. J. Barlow, “Statistics: A guide to the use of statistical methods in the physical sciences”. New York: Wiley, 1989.
- [96] R. Barlow and C. Beeston, “Fitting using finite Monte Carlo samples”, *Comput. Phys. Commun.* **77** no. MAN-HEP-93-1. 2, (Jun, 1993) 219–228. 16 p. <https://cds.cern.ch/record/249779>.
- [97] J. Neyman and E. S. Pearson, “On the problem of the most efficient tests of statistical hypotheses”, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **231** (1933) 289–337. <http://www.jstor.org/stable/91247>.
- [98] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics”, *Eur. Phys. J.* **C71** (2011) 1554, [arXiv:1007.1727](https://arxiv.org/abs/1007.1727) [physics.data-an]. [Erratum: *Eur. Phys. J.* **C73**,2501(2013)].
- [99] **ROOT** Collaboration, K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, “HistFactory: A tool for creating statistical models for use with RooFit and RooStats”, CERN-OPEN-2012-016, 2012.
- [100] S. S. Wilks, “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”, *Annals Math. Statist.* **9** no. 1, (1938) 60–62.
- [101] A. Wald, “Tests of statistical hypotheses concerning several parameters when the number of observations is large”, *Transactions of the American Mathematical Society* **54** no. 3, (1943) 426–482.  
<http://www.jstor.org/stable/1990256>.
- [102] J. Neyman, “Outline of a theory of statistical estimation based on the classical theory of probability”, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **236** no. 767, (1937) 333–380.
- [103] G. Zech, “Upper limits in experiments with background or measurement errors”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **277** no. 2, (1989) 608 – 610.  
<http://www.sciencedirect.com/science/article/pii/016890028990795X>.
- [104] A. L. Read, “Presentation of search results: the CLs technique”, *Journal of Physics G: Nuclear and Particle Physics* **28** no. 10, (2002) 2693.  
<http://stacks.iop.org/0954-3899/28/i=10/a=313>.
- [105] A. L. Read, “Modified frequentist analysis of search results (The CL(s) method)”, in *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*, pp. 81–101. 2000.  
<http://weblib.cern.ch/abstract?CERN-OPEN-2000-205>.



- [106] M. Bahr *et al.*, “Herwig++ Physics and Manual”, *Eur. Phys. J.* **C58** (2008) 639–707 CERN-PH-TH-2008-038, CAVENDISH-HEP-08-03, KA-TP-05-2008, DCPT-08-22, IPPP-08-11, CP3-08-05, [arXiv:0803.0883](https://arxiv.org/abs/0803.0883) [hep-ph].



# Erklärung

Hiermit versichere ich, die vorliegende Arbeit selbständig, ohne fremde Hilfe und ohne Benutzung anderer als der von mir angegebenen Quellen angefertigt zu haben. Alle aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche gekennzeichnet. Die Arbeit wurde noch keiner Prüfungsbehörde in gleicher oder ähnlicher Form vorgelegt.

---

Ort und Datum

---

Max Märker



# Danksagung

Ich möchte all jenen danken, die mich durch das letzte Jahr meiner Master-Studiumszeit begleitet haben. In erster Linie möchte ich Prof. Arno Straessner und Dr. Wolfgang Mader danken für die Betreuung und Ermöglichung meine Arbeit in diesem interessanten Forschungsgebiet anfertigen zu dürfen. Ich danke allen Mitarbeitern des IKTP für das wunderbare Arbeitsklima und die Möglichkeit stets nach Rat zu fragen wenn ich nicht weiter wusste. Ein besonderer Dank gilt meinen Büro-Mitinsassen, vor allem Lorenz Hauswald, Sebastian Wahrmond, Dirk Duschinger und David Kirchmeier für die vielen Hilfestellungen bezüglich der Analyse und ausgedehnten Diskussionen und Gespräche, welche immer interessante und ungeahnte Ausgänge fanden. Besonders möchte ich mich bei Lorenz und Wolfgang für das zügige Korrigieren und die konstruktiven Kommentare zu meiner Arbeit bedanken.

Auch möchte ich mich bei meinem Mitbewohner Eric Mier bedanken, mit dem ich seit dem beginn des Studiums zusammen gewohnt habe und ohne dem ich die Strapazen des Studiums unmöglich hätte überleben können.

Ein besonderer Dank gilt meiner Familie, die mich stets begleitet hat und mir in meiner Schulzeit als auch im Studium mit Rat und Tat zur Seite stand.

Mein größter Dank geht an meine Freundin Chrissy die mir stets viel Verständnis, Geduld und Interesse für meine Arbeit entgegen brachte, auch wenn ich bis spät abends noch Arbeiten musste.