The Henryk Niewodniczański Institute of Nuclear Physics
Polish Academy of Sciences

# Search for charged Higgs boson in the $\tau$+jets final state with proton-proton collision data recorded with the ATLAS experiment

Thesis submitted for the degree of
Doctor Philosophiae

CERN-THESIS-2018-340
05/11/2018

Author:
Anna Bożena Kowalewska

Supervisor:
dr hab. Anna Kaczmarska
Co-Supervisor:
dr Paweł Malecki

Kraków, June 2018

This work was supported in part by National Science Centre grant
2017/25/B/ST2/01234

*To the memory of all those brave and steadfast thanks to whom I know the meening of: freedom, faith and dignity.*

## Abstract

The thesis presents a search for charged Higgs bosons, predicted in many theories beyond the Standard Model, decaying via $H^{\pm} \to \tau^{\pm}\nu$ in the mass range between 90 and 2000 GeV. It is based on 36 fb$^{-1}$ of data collected with the ATLAS detector at center-of-mass energy $\sqrt{s} = 13$ TeV in the years 2015 and 2016.

As an improvement to the former searches, a multivariate analysis is applied and used to separate the $H^{\pm}$ signal from the Standard Model background processes. The Boosted Decision Trees, BDT, output score is used as the discriminating variable in the final statistical analysis. Furthermore, since at low $H^{\pm}$ masses i.e. between the $W$-boson and top-quark masses, the kinematics of the $t \to bH^{\pm}$ and $t \to bW^{\pm}$ decay products can be very similar, variable sensitive to the the $\tau$ lepton polarisation is additionally used to improve discrimination power. The data are found to be in agreement with the background-only hypothesis. Model-independent limits are set on the production cross-section times branching fraction $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ between 4.2 pb and 0.0025 pb for the charged Higgs boson mass range of 90–2000 GeV, corresponding to upper limits between 0.25% and 0.031% for the branching fraction $\mathrm{BR}(t \to bH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ in the mass range 90–160 GeV. In the context of the hMSSM scenario, all $\tan\beta$ values are excluded for $m_{H^{\pm}}$ below 160 GeV, whereas the $H^{\pm}$ mass range up to 1100 GeV is excluded at $\tan\beta = 60$.

The presented thesis constitutes also the first complete implementation of the embedding method in the ATLAS Run-2 analysis. The technique makes use of the lepton universality of the $W$ boson decay by constructing hybrid samples with simulated hadronically decaying $\tau$ leptons replacing muons from $W^{\pm} \to \mu^{\pm}\nu$ events in data. The main advantage of the embedding method is that one relies on the simulations only for the well understood electroweak decays of $W$ boson and $\tau$ lepton and all other event properties are taken directly from data. It has been proven on 2015 data that the technique presents a significant improvement in the precision of background estimation with true hadronically decaying $\tau$ leptons. Compared to the fully simulated events, the uncertainty on the yield of the dominant $t\bar{t}$ background containing true $\tau$ leptons reduces from 26% to the mere 11%.

## Streszczenie

Prezentowana praca zawiera opis poszukiwań naładowanych bozonów Higgsa w kanale rozpadu $H^{\pm} \to \tau^{\pm}\nu$ w przedziale masowym 90-2000 GeV. Ich istnienie przewidziane jest w wielu rozszerzeniach Modelu Standardowego. Analiza wykonana została z użyciem 36 fb$^{-1}$ danych zebranych przez detektor ATLAS w latach 2015-2016 przy energii w środku masy $\sqrt{s} = 13$ TeV.

W celu wyodrębnienia sygnału hipotetycznego $H^{\pm}$ od procesów opisywanych przez Model Standardowy zastosowana została analiza wielowymiarowa będąca rozwinięciem poprzednich metod poszukiwawczych. W końcowej analizie statystycznej, jako zmiennej dyskryminacyjnej użyto wyniku klasyfikacji algorytmu opartego o wzmacniane drzewa decyzyjne (*Boosted Decision Trees*, BDT). Dodatkowo, ponieważ dla niskich mas naładowanego bozonu Higgsa, tj. pomiędzy masą bozonu W, a masą kwarka szczytowego, kinematyka procesów $t \to bH^{\pm}$ oraz $t \to bW^{\pm}$ może być bardzo podobna, dla rozróżnienia sygnału od tła wprowadzona została zmienna wrażliwa na polaryzację leptonu tau. Analizowane dane okazały się zgodne z przewidywaniami Modelu Standardowego. W związku z tym, wyznaczono niezależne od modelu ograniczenie na produkcję naładowanego bozonu Higgsa $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ wynoszące od 4.2 pb do 0.0025 pb dla mas $H^{\pm}$ w przedziale 90–2000 GeV. Odpowiada to górnemu limitowi dla stosunku rozgałęzień $\mathrm{BR}(t \to bH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ o wartości od 0.25% do 0.031% dla mas $H^{\pm}$ w przedziale 90–160 GeV. W interpretacji modelu hMSSM wykluczono wszystkie wartości $\tan\beta$ dla mas $H^{\pm}$ poniżej 160 GeV. Podczas gdy, cały przedział mas $H^{\pm}$ do 1100 GeV jest wykluczony dla $\tan\beta \geq 60$ GeV.

Praca zawiera również opis pierwszego kompletnego wdrożenia tzw. metody embedding dla danych zebranych przez detektor ATLAS w okresie Run-2. Technika ta oparta jest na uniwersalności leptonowej rozpadu bozonu W. Pozwala to na skonstruowanie próbek hybrydowych ze zdarzeń z danych, w których zastępuje się miony pochodzące z procesu $W^{\pm} \to \mu^{\pm}\nu$, pochodzącymi z symulacji hadronowo rozpadającymi się leptonami $\tau$. Główną zaletą metody embedding jest jej ograniczona zależność od symulacji Monte Carlo jedynie do dobrze zrozumianych rozpadów bozonu W oraz leptonu $\tau$. Wszystkie pozostałe własności zdarzenia pochodzą bezpośrednio z danych. Wykazano, przy użyciu części danych z 2015 roku, że technika ta znacząco poprawia precyzję oszacowania tła zawierającego hadronowo rozpadające się leptony $\tau$. W porównaniu do zdarzeń w pełni pochodzących z symulacji Monte Carlo systematyczna niepewność ilości dominującego tła, pochodzącego z zawierających leptony $\tau$ procesów $t\bar{t}$, zostaje zredukowana z 26% do zaledwie 11%.

# Contents


**1 Introduction** **1**

**2 Theory Overview** **7**

2.1 The Standard Model . . . 7

2.2 Fundamental Interactions . . . 8

2.2.1 Quantum Electrodynamics . . . 8

2.2.2 Quantum Chromodynamics . . . 10

2.2.3 Weak Interactions and Electroweak Unification . . . 11

2.3 Spontaneous Symmetry-breaking and Higgs Mechanism . . . 12

2.4 Beyond the Standard Model . . . 15

2.4.1 Open questions in the SM . . . 15

2.4.2 Charged Higgs Bosons in BSM physics . . . 18

2.4.3 Charged Higgs Boson Production Processes in type-II 2HDMs . . . 21

2.4.4 Charged Higgs Boson Decays . . . 22

2.4.5 Experimental Constraints on Charged Higgs . . . 24

2.5 Tau lepton and Tau polarimetry . . . 24

**3 The Experiment** **29**

3.1 The Large Hadron Collider . . . 29

3.2 The ATLAS experiment at the LHC . . . 30

3.2.1 The ATLAS coordinate system . . . 31

3.2.2 The Inner Detector . . . 32

3.2.3 The Calorimeter System . . . 34

3.2.4 The Muon Spectrometer . . . 36

3.2.5 The Trigger System . . . 37

3.3 Data and Monte Carlo simulations . . . 38

3.3.1 Data . . . 38

3.3.2 Event Simulation - MC event generators . . . 41
3.3.3 Detector Simulation . . . 43
3.3.4 Data formats in ATLAS . . . 43
3.4 Physics Objects Reconstruction and Identification . . . 44
3.4.1 Tracks and Vertices . . . 45
3.4.2 Jets . . . 46
3.4.3 $B$-jet tagging . . . 47
3.4.4 Hadronically decaying $\tau$ leptons . . . 48
3.4.5 Electrons . . . 49
3.4.6 Muons . . . 49
3.4.7 Missing Transverse Energy . . . 50

**4 An Outline of the Analysis** **51**
4.1 Event Selection . . . 52
4.1.1 Simulation Samples . . . 52
4.1.2 Event Preselection . . . 54
4.1.3 Removal of geometric overlaps between objects . . . 55
4.1.4 Final event selection . . . 55
4.2 Trigger efficiency measurement . . . 63
4.3 Estimation of Background Contributions . . . 64
4.4 Systematic uncertainties . . . 69

**5 Multivariate analysis** **73**
5.1 Multivariate analysis in particle physics . . . 73
5.1.1 Multivariate Treatment of Data . . . 74
5.1.2 Boosted Decision Trees . . . 79
5.2 BDT in the current $H^{\pm}$ searches . . . 83
5.3 Studies of BDT tunning for full Run-2 data $H^{\pm}$ searches . . . 95
5.3.1 Comparison with the direct use of $m_T$ . . . 95
5.3.2 Impact of using FFs in the BDT training . . . 96
5.3.3 Comparison with cut-based analysis . . . 97
5.3.4 Optimisation of variable sets and mass ranges . . . 102
5.3.5 Study of the impact of $\Upsilon$ variable and that of its components . . . 108
5.3.6 Additional study with 1-prong and 3-prong events . . . 114

**6 Embedding Method** **117**
6.1 Motivation . . . 117
6.2 The significance of the embedding method . . . 118
6.3 Selection of $\mu$+jets events . . . 119
6.4 Implementation of the embedding method . . . 122
6.5 Validation and usage of the technique . . . 125
6.5.1 Corrections Applied to Embedded Samples . . . 125
6.5.2 Systematic Uncertainties . . . 133
6.5.3 Effect of all Systematic Uncertainties on the Embedded Sample . . . 136
6.5.4 Embedded MC simulation - Closure-Check . . . 136
6.5.5 The final step: embedded data in action . . . 140
6.5.6 Conclusion . . . 142
6.5.7 $\Upsilon$ modeling for embedded data . . . 143

**7 Statistical Interpretation of Results** **145**
7.1 The Likelihood Function and Test Statistic . . . 145
7.2 The $CL_s$ Method . . . 147
7.3 Model-independent Limits . . . 148

**8 Conclusions** **155**

**A Collisions at Hadron Colliders and PDFs** **159**

**B Comparison of the Expected Limits** **161**

**C Additional results for MVA** **163**

# Chapter 1

# Introduction

The Universe is outstanding, isn't it? It is enough to look mindfully around once to become aware of how enormous and rich it is.

Our current understanding of the most fundamental constituents of matter and the properties of the interactions between them, except gravity, is described by the theory named the *Standard Model* (SM) of particle physics. The presence of gravity tells us that the SM is an effective theory with a maximum cut-off scale around the Planck mass ($M_P$), which comprises two parts, both born in the 1960's: quantum chromodynamics (QCD) and the electroweak theory of Glashow-Salam-Weinberg (G-S-W). Throughout the past decades, many predictions of the SM were tested to a very high precision in numerous experiments. Nevertheless, it cannot describe such phenomena as neutrino oscillations which imply massive neutrinos, the existence of non-luminous, dark matter or the matter-antimatter asymmetry observed in the Universe. One could imagine, that it is physics above the Planck scale that is responsible for these phenomena, i.e: experimental inconsistencies are not sufficient to put the Standard Model in trouble at lower energies. There are, however, other reasons to go beyond it, e.g.:

- The *hierarchy problem* - we do not know why the electroweak scale is so much smaller than the cut-off one, i.e. typically $M_P$, and what stabilizes the Higgs mass term with respect to quantum corrections.

- We do not have any explanation in the SM of why the electron charge is equal but opposite in sign to the proton charge, as experiments suggest: $Q_e + Q_p < 10^{-21}$ [1].

- We do not understand why the value of the SM parameter that would lead to CP violation in the strong interactions is so small.

- The fermion mass spectrum ranges from $\sim$ 170 GeV, for the top-quark, to $\sim 10^{-3}$ GeV, for the electron. We do not know why there exists such a large difference in masses. And why the luminous matter is made of three families of fermions.

In July 2012 the ATLAS [2] and the CMS [3] Collaborations at the Large Hadron Collider (LHC) [4] announced the discovery of the last missing piece of the SM of particle physics - the neutral Higgs boson [5, 6]. This much anticipated discovery established an important landmark in our quest for better understanding of the World at its most fundamental level for at least two reasons. First, all what has been part of the SM, arguably the most succesfull theory ever put in place, was finally measured. Secondly, the first ever fundamental spin-0 (i.e. scalar) particle was observed. Up to that point all that had been known were handful of particles with spin-1/2 out of which the matter was formed and four spin-1 mediators of interactions[1]. Therefore, after the Higgs discovery, one can ask the following question: are there any more fundamental spin-0 particles? Indeed, many theories extending the SM do predict additional scalars. These are postulated to solve or explain some still puzzling features of the theory or account for observations that cannot be explained within the SM itself. These scalar particles can come in many different varieties with vastly different properties and consequently requiring different experimental search strategies.

The charged Higgs boson, $H^{\pm}$, is a particle appearing in theories with a Higgs sector extended with respect to the SM e.g., by adding a second doublet [7, 8] or a triplet [9, 10] to its scalar sector. It is often a very important ingredient, sometimes even a necessary one, like in the case of all supersymmetric extensions of the SM, where among other reasons, the non-cancellation of the gauge anomaly would otherwise make these theories inconsistent. However, charged Higgs boson is interesting also in its own right, among others, in broad class of the so-called Two Higgs Doublet Models (2HDMs). A discovery of a charged Higgs boson would be an unambiguous sign of physics beyond the SM.

This thesis describes a search for charged Higgs boson in proton-proton ($pp$) collisions based on 36 fb$^{-1}$ of data taken with the ATLAS detector at center-of-mass energy $\sqrt{s} = 13$ TeV in the years 2015 and 2016. The presented analysis is probing a charged Higgs mass range between 90 and 2000 GeV.

At the LHC one of the two main signals of $H^{\pm}$, in a type-II 2HDM model, is a decay to a tau lepton, $\tau^{\pm}$, and tau (anti-)neutrino $\nu_{\tau}$ ($\bar{\nu}_{\tau}$). In the scenario where the charged Higgs boson mass is smaller than the top quark mass ($m_{top} = 172.5$ GeV) this decay channel is practically the only

[1] There are also strong theoretical reasons to expect existence of spin-2 particles, mediators of gravitational interactions. However, none was directly or indirectly observed so far and there is no satisfactionary theory of quantum gravity yet.

one possible, although $H^+ \to c\bar{s}, c\bar{b}$ may also become sizeable in specific region of parameters space. In the heavy Higgs scenario ($m_{H^\pm} > m_{top}$) the dominant decay is $H^\pm \to t\bar{b}(\bar{t}b)$, however the branching fraction of $H^\pm \to \tau^\pm\nu$ can reach $\mathcal{O}(10\%)$.

Unfortunately, it is hard to distinguish the signatures of the $H^\pm \to \tau^\pm\nu$ process, $t\bar{t} \to (W^\mp\bar{b})(bH^\pm) \to (q\bar{q}'\bar{b})(b\tau^\pm\nu)$, from the ones of the SM $W^\pm \to \tau^\pm\nu$ process, $t\bar{t} \to (W^\mp\bar{b})(bW^\pm) \to (q\bar{q}'\bar{b})(b\tau^\pm\nu)$. Therefore, efficient signal-background discrimination is a crucial part of the search and is addressed in this work both from the side of statistical analysis and accurate background estimation.

The first task of the presented thesis was an application of machine learning methods to discrimination of the $H^\pm \to \tau^\pm\nu$ signal events. The recent developement of such methods lead to increasing interest in data analysis and potentially can supplement or even replace the currently adopted cut-based methods. Thus, as an improvement to the former searches a multivariate analysis, MVA, has been applied and used in the current analysis to separate the $H^\pm$ signal from the SM background processes. Where the output score of the Boosted Decision Trees, BDT, is used as the final discriminant variable in statistical analysis of the results. It relies on usage of an algorithm automatically generating the decision trees based on the Monte Carlo (MC) simulation samples, for which the true result of the classification is known. The trees created with this algorithm are used for performing the classification of the actual data.

Design and optimisation of the algorithm as well as determination of the set of variables used in the BDT were important steps towards implementing machine learning methods and assessing their potential in improving the model-independent limits setting on the production of charged Higgs boson. A valuable step in composing optimal set of discriminating variables used by the BDT was a study of $\tau$ polarisation observable. The information about $\tau$ spin is contained in the kinematics of its decay products. Therefore, measurements of the kinematics give additional information about the characteristics of the process from which the $\tau$ leptons arise. In particular, a decay $W^\pm \to \tau^\pm\nu$ leads to $\tau$ leptons with polarisation -1, while a decay of charged Higgs boson $H^\pm \to \tau^\pm\nu$, in models with only left-handed neutrino, to polarisation +1. Hence, the polarisation variable can play a role of an efficient discriminant. Its important advantage is that its signal distribution is independent of the $H^\pm$ mass.

Additionally, in case of positive signal detection coming from charged Higgs, study of $\tau$ polarisation can bring information about its couplings. In particular, e.g. in the theories with large extra dimensions (LED), the decay channel of potential new Higgs boson would be modified to $H^\pm \to \tau_R^\pm\bar{\nu} + \tau_L^\pm\psi$, where $\psi$ is a neutrino propagating in the extra dimension [11].

The second task of the presented thesis was a state-of-the-art determination of the background for the decay processes of the charged Higgs boson. As was mentioned above, it is

crucial to simulate or determine directly from the data the contribution to the measured signal coming from the decay of the $W$ boson in processes $pp \to t\bar{t}$ (top-antitop quark). Due to significant systematic uncertainties of the MC simulations related to the insufficient theoretical understanding of the details of the proton-proton collision, a new, data-driven technique of background estimation was developed, the so-called *embedding method*. The embedding method uses the lepton universality of the $W$ boson decay by constructing hybrid background samples with simulated $\tau$ leptons replacing reconstructed muons in $W^{\pm} \to \mu^{\pm}\nu$ events in data. The latter are replaced, on the level of reconstructed tracks and calorimeter cells, with $\tau$ leptons simulated by the MC methods, where the $\tau$ kinematics is obtained from kinematics of the measured muons. Therefore, one relies on the simulations only for the well understood electroweak decays of $W$ boson and $\tau$ lepton decays. All other properties of an event are taken directly from data including the contribution from the so-called pile-up[2], underlying event and missing transverse momentum. It should be stressed, that the presented thesis constitutes first complete implementation of the embedding technique in the ATLAS analysis of Run-2 data.

This dissertation has been organised as follows. The particle physics framework is introduced in Chapter 2 (see also Appendix A), where a brief overview of the SM and its open issues are discussed. The Chapter also gives a short description of the Higgs sectors of the SM and its minimal extension along with discussion of the signal processes. At the end of the Chapter 2 the description of $\tau$ lepton and its polarimetry can be found. In Chapter 3 the LHC and ATLAS detector are introduced together with information about the data and MC simulations as well as the procedures for the reconstruction and identification of final state objects in the ATLAS detector. The analysis strategy and final states searched for together with the simulation samples used are detailed in Chapter 4. Chapter 5 presents the MVA technique and studies of BDT application to the $H^{\pm}$ search, performed by the Author. The application of the embedding method, data-driven estimation of the dominant background with true hadronically decaying $\tau$ leptons, performed by the Author, is discussed in Chapter 6. The statistical methods used to assess the agreement of the data with the background estimate together with the final results of the $H^{\pm}$ search are described in Chapter 7. Finally, the conclusions are given in Chapter 8.

In the presented thesis, units with $\hbar = c = 1$ are assumed (unless specifed otherwise), hence all masses and momenta are given in energy units.

[2]The phenomena of so-called *pile-up* is described in Sec. 3.3.

## Statement of the Author's Contribution

The work contained in this thesis is a combination of the Author's personal research in the $H^{\pm} \to \tau^{\pm}\nu$ channel, with the work done in collaboration with other members of the ATLAS Higgs Beyond-the-SM (HBSM) research group. The Author's personal contributions pertaining to the presented analysis are:

- design, optimisation and application of the BDT method to the $H^{\pm} \to \tau^{\pm} \nu$ search. Improvement of separation between signal and dominant $t\bar{t}$ pair production background on the basis of the $\tau$ lepton polarisation. These two steps, done for the first time, are described in details in Chapter 5 and results are published in a Ref. [12].

- the studies of the embedding technique presented in Chapter 6. Too low statistics of the data used for embedding prevented application of the obtained results in the offcial studies leading to a journal publication. It should be however stressed, that presented analysis was a pioneer work and in spite of using a small subset of Run-2 data its results give a clear recommendation for its application to the full Run-2 data when available.

The Author participated in the remaining part of the $H^{\pm} \to \tau^{\pm}\nu$ search by taking part in discussions during weekly meetings of the analysis group.

# Chapter 2

# Theory Overview

## 2.1 The Standard Model

> A theory with "...accuracy akin to measuring the distance between New York and Los Angeles and being off by the width of a human hair" Richard P. Feynman.

The Standard Model is built from three theories: Quantum Electrodynamics (QED), the Glashow-Salam-Weinberg theory of electroweak processes and Quantum Chromodynamics (QCD). They are realisations of the framework of the theory of quantum fields in which both matter and interactions are described in terms of four-dimensional fields in flat Minkowski space-time and where excitations of these fields manifest themselves as particles. Moreover, all of the fundamental interactions derive from one general principle, requirement of the *local gauge invariance*.

All matter described by the SM is made of three types of elementary particles: leptons, quarks and interaction mediators. They are divided into two main classes: fermions and bosons. Fermions, which contain leptons and quarks, all having half-integer spin, follow Fermi-Dirac statistics, while bosons, the spin integer mediators, follow Bose-Einstein statistics.

Ordinary matter is constituted by the so-called *first generation* or *first family*, see Table 2.1 (left). Besides these particles that make up ordinary matter, i.e. electrons, protons and neutrons (both built from quarks), in particle accelerators and high energy astrophysics processes more particles have been found, adding up two more generations (see Table 2.1 middle and right). The three generations are almost identical except for the flavour quantum numbers and masses of the constituents. The masses range from a few MeV to 172.5 GeV for the top quark. For light quarks, exact value of the mass is difficult to determine due to confinement, which describes the fact (observation) that quarks cannot exist as free particles. It is also now clear, with the observation of neutrino oscillations, that at least two of them do have masses. Since in the SM

| | | 1st family | | | 2nd family | | | 3rd family | |
|---|---|---|---|---|---|---|---|---|---|
| | | q(e) | mass (MeV) | | q(e) | mass (MeV) | | q(e) | mass (MeV) |
| Quarks | u | 2/3 | $\sim 2.3$ | c | 2/3 | 1275 | t | 2/3 | 173200 |
| | d | -1/3 | $\sim 4.8$ | s | -1/3 | 95 | b | -1/3 | 4500 |
| Leptons | $\nu_e$ | 0 | $< 2 \times 10^{-6}$ | $\nu_\mu$ | 0 | $< 0.2$ | $\nu_\tau$ | 0 | $< 18$ |
| | e | -1 | 0.511 | $\mu$ | -1 | 106 | $\tau$ | -1 | 1777 |

Table 2.1: List of elementary fermions in the SM presented in the form of three families (generations), together with their charges (in units of the electron charge) and their masses (in MeV) [1].

neutrinos are assumed to be massless their oscillations are considered as one of the hints for physics beyond the Standard Model.

There are six leptons, classified according to their charge (q), *electron number* $(L_e)$, *muon number* $(L_\mu)$ and *tau number* $(L_\tau)$. There are also six antileptons, with identical masses but opposite quantum numbers. Thus, in total twelve leptons exist. Similarly, there are six "flavours" of quarks: *u*- "upness", *d*- "downness"; *c*- "charm", *s*- "strangeness"; *t*- "topness (truth)", *b*- "bottomness (beauty)", which are also classified by charge. Again, there are also six antiquarks with opposite quantum numbers. What is more, (anti-)quarks have the so-called *colour charge*: red (r), green (g), blue (b), i.e. they can feel the strong force, while leptons are neutral under it. Hence, in total there are thirty six quarks and antiquarks [13]. Their binding together in different combinations gives rise to observed multitude of colourless hadronic states.

Finally, every interaction has its mediator (*gauge boson*): the massless *photon* for the electromagnetic force, three masive bosons $W^+$, $W^-$ and $Z^0$ for the weak force and eight massless *gluons* with different colour charges (like the quarks, they do not exist as isolated object) to mediate the strong interaction, summarised in Table 2.2.

This all above adding up to a "large?" number of supposedly elementary particles: 12 leptons, 36 quarks, 12 mediators and (as will be shown later) one *Higgs* particle, the only fundamental scalar, i.e. spin $s = 0$ particle described by the SM. Therefore, one ends up with 61 particles to contend with [13].

## 2.2 Fundamental Interactions

### 2.2.1 Quantum Electrodynamics

The quantum theory of electrodynamics (QED) was perfected by Tomonaga, Feynman and Schwinger in the 1940s. It describes interaction between particles carrying electromagnetic

| Interactions | Particle | $q$(e) | mass (MeV) | range (m) | coupling constant |
|---|---|---|---|---|---|
| weak | $W^+$ | +1 | 80.4 | $10^{-18}$ | $\sim 10^{-5}$ $\nearrow$ |
| weak | $W^-$ | -1 | 80.4 | $10^{-18}$ | $\sim 10^{-5}$ $\nearrow$ |
| weak | $Z^0$ | 0 | 91.2 | $10^{-18}$ | $\sim 10^{-5}$ $\nearrow$ |
| electromagnetic | $\gamma$ | 0 | 0 | $\infty$ | $\alpha \approx \frac{1}{137}$ $\nearrow$ |
| strong | g | 0 | 0 | $\sim 10^{-15}$ | 1 $\searrow$ |

Table 2.2: Gauge bosons described by the SM, where masses are quoted following [1] and their charges $q$ are in units of the electron charge. A coupling constant describes the strength of the force exert in an interaction and here is given at the weak scale. The $\nearrow$ ($\searrow$) means that the strenght of the interaction increases (decreases) with decreasing distance. See end of the Section 2.2.2 for more details.

charge mediated by the exchange of a photon in a quantized, relativistically invariant way [13].

Let's consider the Dirac Lagrangian for a spinor field, $\psi$ (a free particle of spin 1/2 with mass $m$)[1]

$$\mathscr{L} = i(\hbar c)\bar{\psi}\gamma^\mu \partial_\mu \psi - (mc^2)\bar{\psi}\psi, \tag{2.1}$$

where $\gamma^\mu$ denote four unitary traceless $4\times 4$ Dirac matrices, i.e. in Dirac representation $\gamma^0 \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $\gamma^i = \left(\begin{smallmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{smallmatrix}\right)$ with $\sigma^i$ being the Pauli matrices. It is invariant under *global phase transformation* ($\psi \to e^{i\theta}\psi$) but not if the phase factor is a function of space-time point, i.e. under *local phase (gauge) transformation*:

$$\psi \to e^{i\theta(x)}\psi. \tag{2.2}$$

In order to ensure the *local phase (gauge) invariance*, the so-called covariant derivative

$$D_\mu \equiv \partial_\mu + i\frac{q}{\hbar c}A_\mu \tag{2.3}$$

and a new *massless* vector field $A_\mu$

$$A_\mu \to A_\mu + \partial_\mu \lambda, \tag{2.4}$$

where $\lambda \equiv -\frac{\hbar c}{q}\theta(x)$, need to be introduced. These lead to the Lagrangian for quantum electrodynamics:

$$\mathscr{L}_{QED} = [i\hbar c\bar{\psi}\gamma^\mu \partial_\mu \psi - mc^2\bar{\psi}\psi] - (q\bar{\psi}\gamma^\mu\psi)A_\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{2.5}$$

[1] In classical mechanics, $\mathscr{L}$ is the difference between kinetic and potential energy. In relativistic field theory $\mathscr{L}$ is taken as axiomatic. Note also that the form of the Lagrangian for a particular system is not unique. It can always be multiplied by a constant or shifted by a constant, or a divergence of an arbitrary vector function. After applying the Euler-Lagrange equations such terms cancel out, therefore do not affect the field equations. Note that for pedagogical reasons till the end of this section $\hbar$ and $c$ are written explicitly.

with $F^{\mu\nu} \equiv (\partial^\mu A^\nu - \partial^\nu A^\mu)$. Note that any additional mass term of the form $\left(\frac{m_A c}{\hbar}\right)^2$ must be zero, i.e. $(m_A = 0)$, otherwise the invariance will be lost.

Two last terms in $\mathscr{L}_{QED}$ reproduce the Maxwell Lagrangian $\mathscr{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{c}J^\mu A_\mu$, with the current density

$$J^\mu = cq\bar{\psi}\gamma^\mu\psi. \tag{2.6}$$

Hence, indeed the theory of quantum electrodynamics can be generated by the requirement of local gauge invariance applied to the free Dirac Lagrangian Eq. (2.1). It is worth mentioning that the symmetry group of QED is $U(1)$, where $U = e^{i\theta(x)}$ and $U^\dagger U = 1$.

### 2.2.2 Quantum Chromodynamics

Quantum chromodynamics describes the strong interactions between quarks and gluons. The free Lagrangian for a particular flavour where, as it was mentioned above, each flavour of quark comes in three colours, for massive quarks is given by

$$\mathscr{L} = i\hbar c\bar{\psi}\gamma^\mu\partial_\mu\psi - (mc^2)\bar{\psi}\psi. \tag{2.7}$$

The notation was simplified by introducing three-component column vector $\psi \equiv \begin{pmatrix}\psi_r\\ \psi_b\\ \psi_g\end{pmatrix}$, where each element of which is itself a four-component Dirac spinor. The symmetry group of QCD is $SU(3)$, thus the Lagrangian (2.7) needs to be modified to become invariant under local $SU(3)$ gauge transformation:

$$\psi \to S\psi, \qquad S \equiv e^{-i\phi_a(x)\lambda_a/\hbar c}, \tag{2.8}$$

where $\lambda_a$ represents eight linearly independent $3\times 3$ Gell-Mann matrices. As for QED, maneuver is to replace the ordinary derivative, $\partial$, by the "covariant derivative": $D_\mu \equiv \partial_\mu + i\frac{g}{\hbar c}\lambda_a G^a_\mu$, and assign to the eight gauge fields $G^a_\mu$ the following transformation rule

$$D_\mu\psi \to S(D_\mu\psi), \qquad G^\mu_a \to G^\mu_a - \frac{\hbar c}{g}\partial_\mu\phi_a - f_{abc}\phi_b G^c_\mu. \tag{2.9}$$

The third term of Eq. (2.9) with the real structure constant $f_{abc}$ has its source in non-abelian structure of QCD and leads to self-interactions among gluons. The complete Lagrangian, invariant under local $SU(3)$ gauge transformations and describing three equal-mass Dirac fields, the three colours of a given quark flavour, in interaction with eight massless vector fields (gluons) is given by

$$\mathscr{L}_{QCD} = [i\hbar c\bar{\psi}\gamma^\mu\partial_\mu\psi - mc^2\bar{\psi}\psi] - (g\bar{\psi}\gamma^\mu\lambda_a\psi)G^a_\mu - \frac{1}{4}G^{\mu\nu}_a G^a_{\mu\nu}, \tag{2.10}$$

where $G^{\mu\nu}_a = \partial^\mu G^\nu_a - \partial^\nu G^\mu_a - g f_{abc}G^\mu_b G^\nu_c$. It was derived from the requirement that the global $SU(3)$ symmetry of the free QCD Lagrangian (see Eq. 2.7) should hold locally [13]. Note that

in the same manner that the electric current acts as the source for electromagnetic fields, the eight colour currents: $J_a^\mu \equiv cg(\bar{\psi}\gamma^\mu\lambda_a\psi)$, act as sources for the colour fields, $G_a^\mu$.

What distincts mostly, besides self-interactions of gluon fields, the QCD from QED are the phenomenons known as *asymptotic freedom* and *confinment*. The strenght of the strong interaction decreases with increasing energy scale of a process (see also Table 2.2), i.e. at small spatial distances, quarks and gluons become asymptotically free (rattle around without interacting much). Whereas, at large spatial distances (low energies), the coupling becomes very strong causing that quarks and gluons can exist only in the bound states, i.e. in the form of colourless combinations, e.i. hadrons. In high energy physics a typical reference scale at which to present the value of the strong coupling strength is the Z boson mass squared.

### 2.2.3 Weak Interactions and Electroweak Unification

The weak interactions are described by $SU(2)$ local phase transformations: $\psi \to \psi' = e^{i\vec{\phi}(x)} \cdot \frac{\vec{\sigma}}{2}\psi$, where the three Pauli matrices are the generators of the $SU(2)$ symmetry with three gauge bosons, $W_1^\mu$, $W_2^\mu$, $W_3^\mu$ and where the fermions form weak isospin doublets, $\psi = \binom{\nu_e}{e}_L$. These interactions acts only on left-handed particles and right-handed anti-particles. Their form is specified by the gauge symmetry, i.e. there exists one term for each of the three $SU(2)$ generators:

$$j_\mu^1 = g_W \bar{\psi}_L \gamma^\mu \frac{1}{2}\sigma_1 \psi_L, \qquad j_\mu^2 = g_W \bar{\psi}_L \gamma^\mu \frac{1}{2}\sigma_2 \psi_L, \qquad j_\mu^3 = g_W \bar{\psi}_L \gamma^\mu \frac{1}{2}\sigma_3 \psi_L, \tag{2.11}$$

where $g_W$ is the coupling constant for weak interactions. The neutral current interactions are represented by the third current $j_\mu^3$. Whereas, the charged current $W^+/W^-$ interactions enters as a linear combinations of $W_1$ and $W_2$:

$$W^{\pm\mu} = \frac{1}{\sqrt{2}}\left(W_1^\mu \pm W_2^\mu\right) \tag{2.12}$$

and then

$$j_\pm^\mu = \frac{g_W}{\sqrt{2}}\left(j_1^\mu \pm i j_2^\mu\right) = \frac{g_W}{\sqrt{2}}\bar{\psi}_L\gamma^\mu\frac{1}{2}(\sigma_1 \pm i\sigma_2)\psi_L. \tag{2.13}$$

Note that the gauge quantum numbers are the same across different families. In case of leptons it leads to the so-called *lepton universality* - in high energy limit, when the lepton masses can be neglected, the cross sections for processes mediated by the weak interaction are the same for all families.

The charged weak interactions are mediated by $W^\pm$ bosons and neutral ones by $Z$ boson. In the Glashow-Salam-Weinberg electromagnetic and weak interactions are unified into a common electroweak interactions. The combined gauge group is $SU(2)_L \times U(1)_Y$ with the weak

hypercharge $Y$ - the charge of this symmetry. It is connected to electric charge $Q$ and the 3rd component of weak isospin $I_3$ by the following formula: $Y = 2Q - 2I_3$. The subscript $L$ ascribe that only left-handed fermions couple to the weak isospin current in terms of chirality. As G-S-W theory states, an isotriplet of vector fields $W_\mu^i$ couples to the weak isospin with coupling strenght $g_W$ and an isosinglet vector field $B_\mu$ couples to the weak hypercharge with a coupling strenght $\frac{g'}{2}$. Whereas the $W_\mu^i$ respect the $SU(2)_L$ symmetry, $B_\mu$ couples as well as to left- and right-handed fermions. The massive charged $W^\pm$ bosons fields are given by Eq. 2.12, whereas the fields of the photon, $A_\mu$, and Z boson, $Z_\mu$, are created from $W_\mu^3$ and $B_\mu$ as follows:

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W \tag{2.14}$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W \tag{2.15}$$

with the *weak mixing angle* $\theta_W$. Note that coupling constants $g_W$, $g'$ and electric charge $q$ are related:

$$g' \cos\theta_W = g_W \sin\theta_W = q. \tag{2.16}$$

In Table 2.3 fermion doublets and singlets are presented together with assigned $Y$ and $I_3$ value. Quarks are given in the interaction basis, i.e. where $d'$, $s'$ and $b'$ are eigenstates of the weak interactions. These do not coincide with the mass eigenstates $d$, $s$ and $b$, and the mixing between the two is described by the Cabibbo-Kobayashi-Maskawa matrix [14, 15]. The existence of this mixing is an essential ingredient of the SM to be able to correctly describe the CP violation due to the weak interactions.

## 2.3 Spontaneous Symmetry-breaking and Higgs Mechanism

With the term *spontaneous symmetry breaking* one describes the situation, where the vacuum state has lower symmetry than the theory itself.

Consider the Lagrangian, which is invariand under a global $U(1)$:

$$\mathscr{L} = \partial_\mu \varphi^* \partial^\mu \varphi - V(\varphi\varphi^*), \tag{2.17}$$

$$\varphi(x) \longrightarrow \varphi'(x) = e^{i\alpha}\varphi(x), \tag{2.18}$$

where $\varphi(x)$ is a complex scalar field and $\alpha$ a real constant parameter. Assuming that the potential can be expanded in the powers of the field $\varphi$: $V(\varphi\varphi^*) = m^2|\varphi|^2 + \lambda|\varphi|^4 + ...$, where $m$ has the interpretation of the mass of the excitations of field $\varphi$, while $\lambda$ characterizes the strenght of the self-interaction.

In order to find the state that corresponds to the vacuum one needs to minimize the potential $V(\varphi\varphi^*) = m^2|\varphi|^2 + \lambda|\varphi|^4$. We choose $\lambda > 0$, as otherwise the potential would be unbounded

| Leptons | | | Hypercharge Y | $I_3$ |
|---|---|---|---|---|
| $\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$ | $\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$ | $\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$ | $-1$<br>$-1$ | $\frac{1}{2}$<br>$-\frac{1}{2}$ |
| $e_R^-$ | $\mu_R^-$ | $\tau_R^-$ | -2 | 0 |
| Quarks | | | Y | $I_3$ |
| $\begin{pmatrix} u \\ d' \end{pmatrix}_L$ | $\begin{pmatrix} c \\ s' \end{pmatrix}_L$ | $\begin{pmatrix} t \\ b' \end{pmatrix}_L$ | $\frac{1}{3}$<br>$\frac{1}{3}$ | $\frac{1}{2}$<br>$-\frac{1}{2}$ |
| $\begin{pmatrix} u \\ d' \end{pmatrix}_R$ | $\begin{pmatrix} c \\ s' \end{pmatrix}_R$ | $\begin{pmatrix} t \\ b' \end{pmatrix}_R$ | $\frac{4}{3}$<br>$-\frac{2}{3}$ | 0<br>0 |

Table 2.3: Doublets of left(right)-handed and singlets of right-handed fermions where $d'$, $s'$ and $b'$ are linear combinations of the physical quarks, $d$, $s$ and $b$ whoose mixing is described by Cabibbo-Kobayashi-Maskawa matrix. The respective values for hypercharge $Y$ and the 3rd-component of isospin $I_3$ are also shown.

from below and there would be no stable vacuum. In the case of $m^2 > 0$ the minimum of the potential is located at the field value $\varphi = 0$. The shape of the potential is shown in Fig. 2.1(a).

For $m^2 < 0$ the minimum of the potential is located at:

$$|\varphi_{min}|^2 = \frac{-m^2}{2\lambda}. \tag{2.19}$$

It follows that in the plane of the complex field $\varphi$ exists a complete circle of radious $v$:

$$v := \sqrt{\frac{-m^2}{2\lambda}} \geq 0, \tag{2.20}$$

in every point of which the potential takes its minimal value, see Fig. 2.1(b). Hence, there is an infinite number of states with lowest energy, i.e. the vacuum is *degenerate*. After a definitive ground state is chosen, the Hamiltonian ceases to be invariant under the rotation Eq.(2.18). The symmetry is *spontaneously* broken.

Expanding the Lagrangian around the true minimum (corresponding to the true vacuum) and using the two real degrees of freedom $\varphi_1$ and $\varphi_2$

$$\varphi = v + \frac{\varphi_1 + i\varphi_2}{\sqrt{2}} \tag{2.21}$$

one arrives at:

$$\mathscr{L} = \frac{1}{2}\left(\partial_\mu \varphi_1\right)^2 + \frac{1}{2}\left(\partial_\mu \varphi_2\right)^2 - 2\lambda v^2 \varphi_1^2 - \sqrt{2}\lambda \varphi_1 \left(\varphi_1^2 + \varphi_2^2\right) - \frac{\lambda}{4}\left(\varphi_1^2 + \varphi_2^2\right)^2. \tag{2.22}$$

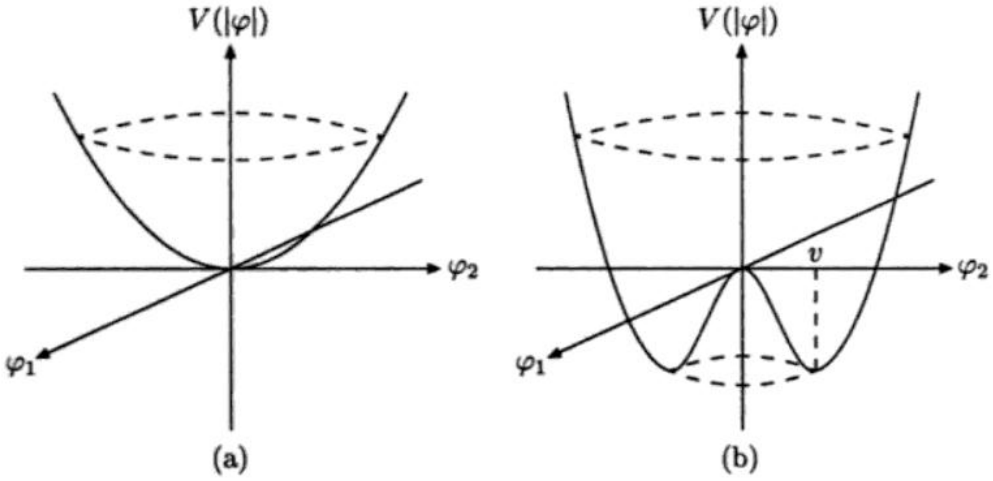



Figure 2.1: The potential for the $\varphi$ field in the basis of two real degrees of freedom $\varphi_1$ and $\varphi_2$. Two cases are shown: without (a) and with (b) spontaneous symmetry breaking. From [16].

Therefore, as a consequence of choosing the vacuum state and using the description of the new fields $\varphi_1$ and $\varphi_2$, describing excitations around the new minimum of the potential, one obtaines [2]: massless field $\varphi_2 \rightarrow m^2_{\varphi_2} = 0$ and massive field $\varphi_1$, whoose mass term is $-\frac{1}{2}m^2_{\varphi_1}\varphi_1^2$, where $m^2_{\varphi_1} = 4\lambda v^2$.
That is, an excitation of the $\varphi_1$ degree of freedom "costs" energy $m^2_{\varphi_1} = 4\lambda v^2$, while an excitation of $\varphi_2$ does not "cost" anything $\rightarrow m^2_{\varphi_2} = 0$.

After making the above system invariant under local gauge transformations, $\phi \rightarrow e^{i\alpha(x)}\phi$, by the introducion of massless gauge field $A^\mu$ and replacing the derivatives in Equation 2.17 with covariant derivatives, $D_\mu = \partial_\mu + i\frac{q}{\hbar c}A_\mu$, as well as after selection of a convinient gauge, $\alpha = -\tan^{-1}(\phi_2/\phi_1)$, and rewriting the fields in terms of fluctuations about a particular ground state , $\eta \equiv \phi_1 - m/\lambda$, ones ends up with the Lagrangian

$$\mathscr{L} = \left[\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - m^2\eta^2\right] + \left[\frac{1}{2}\left(\frac{q}{\hbar c}\frac{m}{\lambda}\right)^2 A_\mu A^\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}\right] + \\ + \left\{\frac{m}{\lambda}(\frac{q}{\hbar c})^2\eta(A_\mu A^\mu) + \frac{1}{2}(\frac{q}{\hbar c})^2\eta^2(A_\mu A^\mu) - \lambda m\eta^3 - \frac{1}{4}\lambda^2\eta^4\right\} + (\frac{m^2}{2\lambda})^2 \quad (2.23)$$

with a single *massive* scalar $\eta-$ the *Higgs* particle and a *massive* gauge field $A^\mu$.

That is, the Higgs mechanism (more precisly: Brout-Englert-Higgs mechanism for which the 2013 Nobel Prize was awarded jointly to Francois Englert and Peter W. Higgs) brings mass

[2]The remaining terms in the Lagrangian are the kinetic ones and representing the interactions.

to the gauge fields. In the SM, this mechanism is responsible for the masses of the weak interactions gauge bosons, $W^{\pm}$ and $Z^0$. Additionaly, due to Yukawa-type couplings of left and right chiral components of the fermion fields to the Higgs field spontaneous symmetry breaking also allows to incorporate non-zero masses of leptons and quarks, without explicit breaking of the gauge symmetry.

#### The Higgs Boson

Almost 50 years after the Brout-Englert-Higgs mechanism had been proposed, in July 2012 the Higgs boson ($H$) was discovered by the ATLAS and CMS Collaborations at the LHC [5, 6]. The combination of the measurements in the $\gamma\gamma$ and $ZZ$ decay channels based on the full dataset taken at $\sqrt{s} = 7$ TeV and 8 TeV by the two Collaborations results in the $H$ mass of $125.09 \pm 0.24$ GeV [17]. The measurements of the spin and the parity quantum numbers of the new boson favour values expected for the SM Higgs, i.e. spin $s = 0$ and $J^P = 0^+$ over the alternative hypotheses [18, 19]. Fig. 2.2 shows the current combination of the measured signal strengths $\mu$, i.e. the ratio of observed yield of the Higgs signal decaying to pair of particles to the yield expected in the SM. In ATLAS a combination over all observed channels results in a value of

$$\mu = 1.18^{+0.15}_{-0.14}, \tag{2.24}$$

which is consistent with the predicted coupling strengths in the SM (in general, any deviation from $\mu = 1$ could be an indication for new physics).

## 2.4 Beyond the Standard Model

### 2.4.1 Open questions in the SM

The Standard Model, despite its great successes, is certainly not the ultimate theory. The most important reason is gravity. It is not incorporated in the SM and in fact cannot be. Our understanding of gravitational interactions is still classical and relies on the fact that the space-time is curved and dynamical. This is very challenging to merge with quantum theory on a static flat Minkowski space.

Indeed, no theory of quantum gravity exists yet, but nevertheless we know that it is inevitable. At all experimentally accessible energy scales, gravity is very weak. However, its importance clearly grows with energy, such that around the scale of the Planck mass $m_{Pl} = 10^{19}$ GeV it cannot be neglected any more. As a consequence, at this point a new, more complete theory is needed to replace our low energy description.

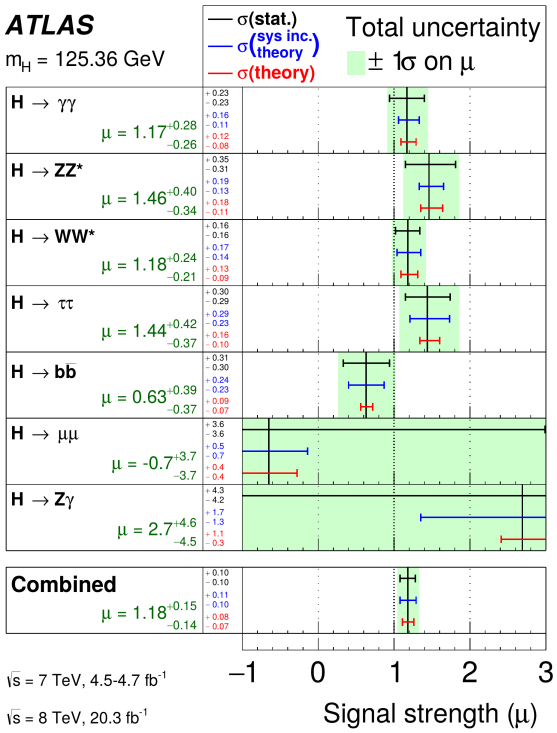



Figure 2.2: The observed signal strengths and uncertainties for five different Higgs boson decay channels and their combination measured in ATLAS experiment [20].

This obliges us to look at the SM as an effective theory, valid only up to some energy scale. At higher energies it should be replaced by a new one. It is conceivable that this new physics is also related to the other open questions of the SM, so that they can provide hints where to look. In the following, some of those issues are briefly summarized.

### Hierarchy Problem

The naturalness issue of the Standard Model, the so-called *hierarchy problem* can be stated as: why does the electroweak scale is so small compared to the cut-off one, typically considered to be the Planck scale? The cut-off is present, as the SM is necessarily an effective theory valid only up to some energy scale - which at most can be the Planck scale where the effects of gravity are of the same size as of all the other interactions. Such a cut-off, however, means that the quadratic divergences appearing in the one loop corrections to the Higgs boson mass squared cannot be simply absorbed in renormalized quantities. Indeed, in order for the Higgs mass to be at the electroweak scale, as observed, the parameters in the SM Lagrangian need to be extremely precisely *fine-tuned*, so that an incredible cancellation occurs. Although this might turn out to be just a coincidence, a more "natural" reason would be that there, in fact, exists some sort of a mechanism that eliminates or strongly suppresses such quadratic corrections in the first place.

### Matter-Antimatter Asymmetry

Presumably, the Big Bang created matter and antimatter in equal amounts, but the Universe is made of matter (mostly electrons, protons and neutrons). Therefore, during cosmic evolution some process(es) must have favored matter over antimatter. That is, there must have been an interaction that violates conservation of baryon and lepton number. A period when the Universe was in fact out of thermal equilibrum, as well as there must have been a charge conjugation parity (CP) symmetry violation in order to have net change in baryon number. It is known from the experiment (first discovered by Cronin and Fitch in the $K^0 - \bar{K}^0$ system [21]) that CP violation happens in the weak interactions of quarks. Unfortunately, this is *nowhere near* enough to account for the matter dominance of the Universe and still this is an uncompleted puzzle, far from clear solution [13].

### Cosmological problems

Astronomical evidences point out that the ordinary matter, described by the SM, represents around 5% of the mass (energy) content of the Universe. The rest is the so-called *Dark Matter* (DM), i.e. matter not interacting with light at any wavelength, $\sim 20\%$ and *Dark Energy* $\sim 75\%$.

In 1933 Fritz Zwicky [22] measured velocity dispersions of some of the galaxies of the Coma cluster and then estimated its mass with the use of the virial theorem. The results showed that the ratio of the mass to the total luminosity obtained this way was more than two orders of magnitude larger than the mass to the luminosity ratio locally. This discrepancy was surprising, as such a large value by any means could not be explained with the standard astrophysical objects. This led Zwicky to suggest that at least in the observed cluster, there exists some non-luminous matter component which holds the cluster together. Additional hints for the existance of DM was given by gravitational lensing effects [23].

One of the major challenges of modern particle physics is to understand the nature of DM and hopefully incorporate it in the more fundamental and complete theory. One kind of dark matter candidate particles are weakly interacting massive particles (WIMPs), appearing in many extensions of the SM. However, so far no dark matter particle has been found.

The Dark Energy was discovered due to the astonishing observation made by the two independent projects, the Supernova Cosmology Project [24] and the High-Z Supernova Search Team [25], that the expansion of the Universe is accelerating. This discovery was awarded the Nobel Prize in 2011. However, the nature of dark energy is still unknown apart from the fact that it acts as the negative pressure on the Universe.

#### Neutrino masses

On contrary to what the Standard Model predicts, it is now known that neutrinos are massive. This is a major issue for the original SM, but it can be easily extended in order to take neutrino masses into account. It is sufficient to add a new right-handed neutrino $\nu_R$ for every generation and introduce a Dirac mass term. The only conceptual problem with this solution is that the resulting $\nu_R$ would be *sterile*, i.e. it would not interact at all with other SM particles (except for gravitationally).

However, one can also treat neutrino masses as a hint when looking for extension of the SM. For example, if neutrinos are Majorana particles, i.e. that they are their own antiparticles, the smallness of their masses can be elegantly explained via the see-saw mechanism [26]. This idea can also lead to thermal leptogenesis as a very promising mechanism for producing the baryon asymmetry [27].

#### Others

There are also other open questions, e.g.:

- Charge quantization: why the hydrogen atom is neutral? The charges of quarks and leptons could be totally unrelated, but somehow the sum of the electromagnetic charges of three valence proton quarks is exactly opposite to the electron charge.
- Gauge coupling unification: the evolution of the SM gauge couplings with the energy shows the tendency to bring them very close to each other at some very high scale. This happens at some point roughly from $10^{12}$ up to $10^{16}$ GeV. However, this unification is not exact within the SM itself.
- The flavour puzzle: why the flavour sector seems so complicated and without any guiding principle? The enourmous hierarchies in leptons and quark masses call for a deeper understanding, again necessarily going beyond the SM.

### 2.4.2 Charged Higgs Bosons in BSM physics

#### SUSY and MSSM- briefly

Supersymmetric (SUSY) extensions of the Standard Model adds a new symmetry to the SM, which relates bosons to fermions and fermions to bosons. Since the supersymmetry generators $Q_\alpha$ are spinors, from the conservation of angular momentum it follows:

$$Q_\alpha |fermion\rangle = |boson\rangle$$

$$Q_\alpha|boson\rangle = |fermion\rangle.$$

It means that every SM particle has an associated supersymmetric partner whose spin is different by $\frac{1}{2}$ and with otherwise identical quantum numbers. All SM particles and their supersymmetric partners are arranged in the so-called supermultiplets, which contain both fermions and bosons with equal masses.

In order to preserve renormalizability and give masses to the different fermions at least two Higgs doublets are required within SUSY. To give mass to both up-type and down-type quarks and charged leptons, a Higgs doublet with weak hypercharge $Y_{H_1} = -1$ and another one with $Y_{H_2} = 1$ is required:

$$H_1 = \begin{pmatrix} H_1^0 \\ H_1^- \end{pmatrix} \qquad and \qquad H_2 = \begin{pmatrix} H_2^+ \\ H_2^0 \end{pmatrix}. \tag{2.25}$$

The *Minimal Supersymmetric Standard Model* is the minimal possible extension of the SM in terms of particles and their interactions, i.e. it introduces the least possible number of new fields while preserving supersymmetry and all the desirable features of the SM.

The first step in constructing the MSSM is to decide in what supermultiplets the SM fields should be embedded. Since fermions belong to the fundamental representation of the gauge group, while gauge bosons to the adjoint one, then they cannot create a common supermultiplet. Therefore, SM quarks and leptons are placed in (separate) chiral multiplets, while gauge bosons in the vector ones. The remaining degrees of freedom are filled by superpartners: additional particles added to the SM. They consist of scalar partners of quarks and leptons, with the corresponding names with added prefix ”s” (short for scalar).

In fact, the left- and right-handed chiral states of quarks and leptons have different gauge transformation properties, so each must have its own complex scalar partner. Note that the index $L$ and $R$ in the name of squarks and sleptons denote the chirality of its corresponding superpartner.

The vector multiplet apart from gauge bosons needs also to contain fermions. They are called gauginos and belong to the adjoint representation of the gauge group. These are the so-called gluinos, charginos and neutralinos.

### The Higgs Sector of the MSSM- a type II 2HDM

Finally, to complete the MSSM one needs to specify the Higgs sector. Because Higgs is a scalar it has to reside in another chiral multiplet and be accompanied by a fermion superpartner, called Higgsino. In fact, as it was already mentioned, it turns out that one Higgs doublet is not enough to provide masses to both up-type and down-type quarks so one needs two Higgs doublets.

Second reason is that second doublet is needed to cancel the gauge anomaly. Thus, finally the MSSM Higgs sector comprises two scalar Higgs doublets with opposite hypercharge (see Eq. 2.25) and four Higgsinos: two neutral and two charged. After *the electroweak symmetry breaking* ones is left with five spin-0 particles: neutral scalars $h$, $H^0$, complex scalar $H^{\pm}$ and a pseudoscalar $A$ (see e.g. [8,28]).

The Higgs sector has two free parameters, one of which is the $\tan\beta \equiv \frac{v_2}{v_1}$ i.e. the ratio of the vacuum expectation values of two Higgs fields (see Eq. 2.28), which satisfy $(v_1+v_2)^2 = v^2$ with $v \approx 246$ GeV and the second one is conveniently chosen to be the pseudoscalar mass $m_A$. Then the remaining masses are expressed as:

$$m^2_{h,H^0} = \frac{1}{2}\left[m^2_A + m^2_Z \mp \sqrt{(m^2_A+m^2_Z)^2 - 4m^2_Z m^2_A \cos^2 2\beta}\right], \tag{2.26}$$

$$m^2_{H^{\pm}} = m^2_A + m^2_W\,, \tag{2.27}$$

with, at the tree level, the following constraints imposed on the Higgs boson masses: $m_{H^0} > max(m_A, m_Z)$, $m_{H^{\pm}} > m_W$ and $m_h \leq min(m_A, m_Z)\cdot|\cos 2\beta| \leq m_Z$.

The vacuum expectation values of the neutral fields are defined as follows

$$\langle H^0_1\rangle = \frac{v_1}{\sqrt{2}} \qquad \text{and} \qquad \langle H^0_2\rangle = \frac{v_2}{\sqrt{2}}. \tag{2.28}$$

Finally, the couplings of the charged Higgs bosons to $\tau$ leptons are presented in Eq. 2.29 - 2.30.

$$g_{H^+\tau^-\bar{\nu}_\tau} = \frac{-i}{\sqrt{2}v}(m_\tau \tan\beta(1+\gamma_5)) \tag{2.29}$$

$$g_{H^-\tau^+\nu_\tau} = \frac{-i}{\sqrt{2}v}(m_\tau \tan\beta(1-\gamma_5)) \tag{2.30}$$

Morover, for the recipe on how the $\tan\beta$ can be determined by measuring the signal rate in the $H^{\pm}\to\tau^{\pm}\nu$ channel, see Eq. 2.31.

$$\Gamma(H^{\pm}\to\tau^{\pm}\nu_\tau) \simeq \frac{m_{H^{\pm}}}{8\pi v^2}\left[m^2_\tau \tan^2\beta \times \left(1-\frac{m^2_\tau}{m^2_{H^{\pm}}}\right)\right] \times \left(1-\frac{m^2_\tau}{m^2_{H^{\pm}}}\right). \tag{2.31}$$

### Benchmark Scenarios for MSSM Higgs Boson Searches

The various benchmark scenarios for MSSM Higgs boson searches at hadron colliders are explained in details in [29], where some parameters, apart from $\tan\beta$ and $m_{H^{\pm}}$, are fixed to certain values (due to the large number of free parameters in the MSSM it is too complex to scan of the whole parameter space). The example benchmarks are the so-called $m_h^{\text{mod}+}$ and $m_h^{\text{mod}-}$ scenarios, in which the lightest CP-even Higgs boson, $h$, has a mass close to the Higgs boson found at the LHC, following from an appropriate choice of the top squark mixing parameter.

### 2.4.3 Charged Higgs Boson Production Processes in type-II 2HDMs

In a type-II 2HDM, which corresponds to the Higgs sector of the MSSM, the production of the charged Higgs boson depends on its mass $m_{H^\pm}$. For low masses ($m_{H^\pm} < m_t$), the primary production mechanism is through the decay of a top quark, $t \to bH^\pm$. The leading source of top quarks at the LHC is $t\bar{t}$ production. For $H^\pm$ masses above the top-quark mass ($m_{H^\pm} > m_t$), the leading $H^\pm$ production mode at the LHC is in association with a top quark, which can be described as either $gb \to tH^\pm$ (referred to as the 5-flavour scheme, or 5FS) or $gg \to tbH^\pm$ (referred to as the 4-flavour scheme, or 4FS). In Fig.2.3 the $t\bar{t}$ production followed by the top quark decay (upper plot) and top quark associated production (bottom left-5FS and bottom right-4FS) are shown.

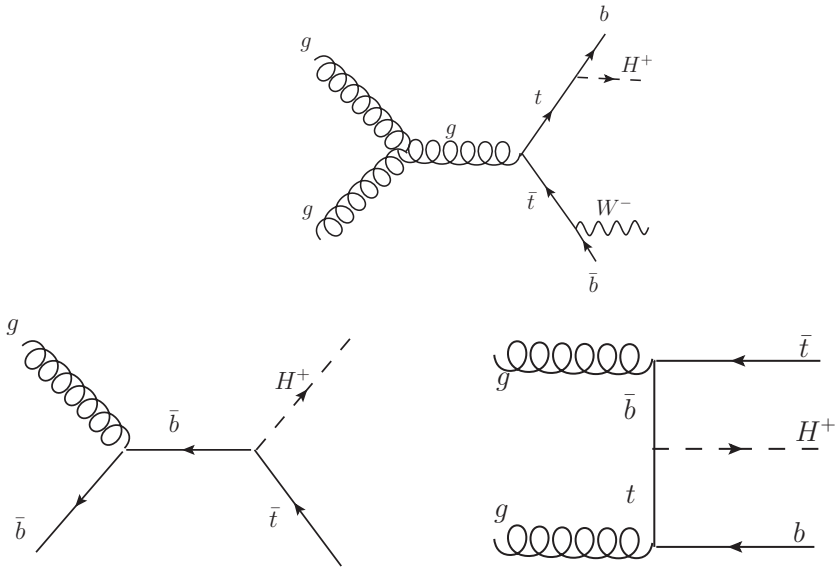



Figure 2.3: Leading Feynman diagrams for the charged Higgs boson production processes in the proton-proton collision: top quark decay (upper) and associated top production in five-flavour (5FS- bottom left) and four-flavour (4FS- bottom right) scheme.

In the five-flavour (5FS) scheme, the b-quark is considered as an active flavour inside the proton, in contrast to the four-flavour (4FS) one. In spite of the fact that the 4FS and 5FS cross sections agree when computed to all orders in perturbation theory, any finite order summation yields different results in the two schemes due to different ordering of the perturbation expansion. In particular, at leading order the 4FS and 5FS schemes predictions can differ significantly. In order to avoid dependence on the chosen approximation cross sections calculated in the 4FS and 5FS approximations are combined and matched [30, 31]. The difference between the two schemes is logarithmic thus the dependence of the relative weight is controlled by a logarithmic

term. In Eq. 2.32 the matched cross section is presented, with the weight $w = \log \frac{m_H^{\pm}}{m_b} - 2$, where $m_b$ is the b-quark mass:

$$\sigma_{matched} = \frac{\sigma_{4FS} + w\sigma_{5FS}}{1+w}. \tag{2.32}$$

Whereas, the theoretical uncertainties are combined according to

$$\Delta\sigma_{matched} = \frac{\Delta\sigma_{4FS} + w\Delta\sigma_{5FS}}{1+w}. \tag{2.33}$$

The production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV for heavy charged Higgs bosons in a type-II 2HDM are shown in Fig. 2.4, whereas an overview of the cross section estimation in the intermediate-mass region can be found in [32]. Also a two-dimentional plot of the production cross section as a function of $\tan\beta$ and $m_{H^{\pm}}$ for 4FS in the 2DHM is presented in Fig. 2.5.

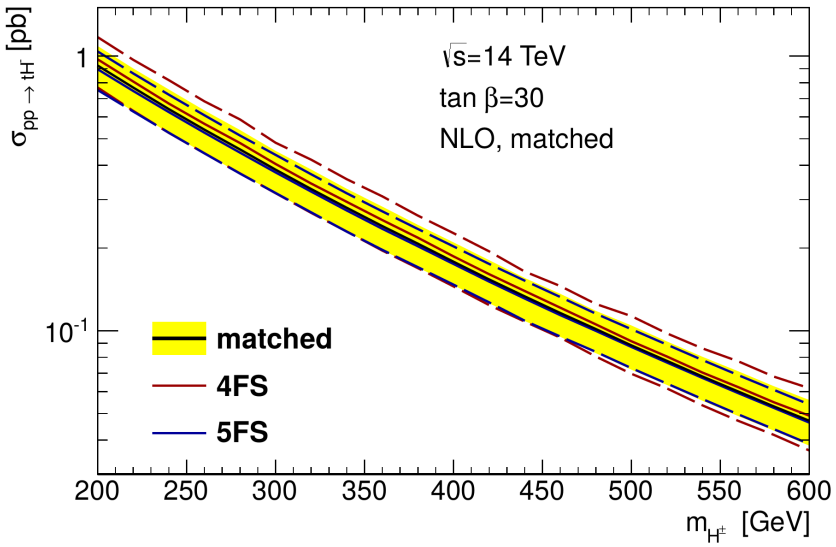



Figure 2.4: Production cross section for charged Higgs bosons as a function of $m_{H^{\pm}}$ for $\tan\beta = 30$ at a center-of-mass energy of $\sqrt{s} = 13$ TeV. Values are shown for the 5FS, 4FS schemes and matched, where the red and blue dashed lines indicate the systematic uncertainties on the 4FS and 5FS, respectively, and the yellow band shows the uncertainty on the matched cross sections [31].

### 2.4.4 Charged Higgs Boson Decays

In the proton-proton collision the charged Higgs boson signals can be most efficiently looked for in the decays $H^{\pm} \to \tau^{\pm}\nu$ and, in the heavy Higgs scenario, also in $H^{+} \to t\bar{b}$ ($H^{-} \to \bar{t}b$) for

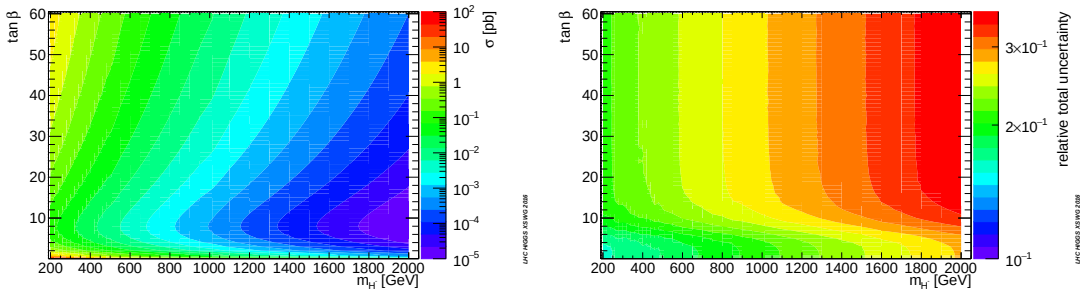



Figure 2.5: Two-dimensional plot of the charged Higgs boson production cross section (left) and average relative uncertainty (right) as a function of $\tan\beta$ and $m_{H^\pm}$ values in the 4FS of the 2DHM [33].

a broad mass range of $m_{H^\pm}$ in the MSSM. The partial decay width for $H^\pm \to l^\pm \nu$ is defined by

$$\Gamma(H^\pm \to l^\pm \nu) = \frac{G_F m_{H^\pm}}{4\sqrt{2}\pi} m_l^2 \tan^2\beta \left(1 - \frac{m_l^2}{m_{H^\pm}^2}\right)^3, \tag{2.34}$$

where $l$ denotes any charged lepton with mass $m_l$ and $G_F$ is the Fermi constant.

The branching fractions for $H^\pm$ decays into SM particles are displayed in Fig. 2.6, as a function of $m_{H^\pm}$, for $\tan\beta = 10$ and $\tan\beta = 50$, in the $m_h^{\text{mod+}}$ scenario. The $H^\pm$ decays to SUSY particles are taken into account, but are not shown in the plots (their effects are visible as kinks in the lines of other channels, in particular for $\tan\beta = 10$). For $\tan\beta > 3$, light charged Higgs bosons decay mainly via $H^\pm \to \tau^\pm \nu$. Above the top-quark mass, the branching fraction BR($H^\pm \to \tau^\pm \nu$) can still be substantial (at least 10%) depending on the value of $\tan\beta$ [33].

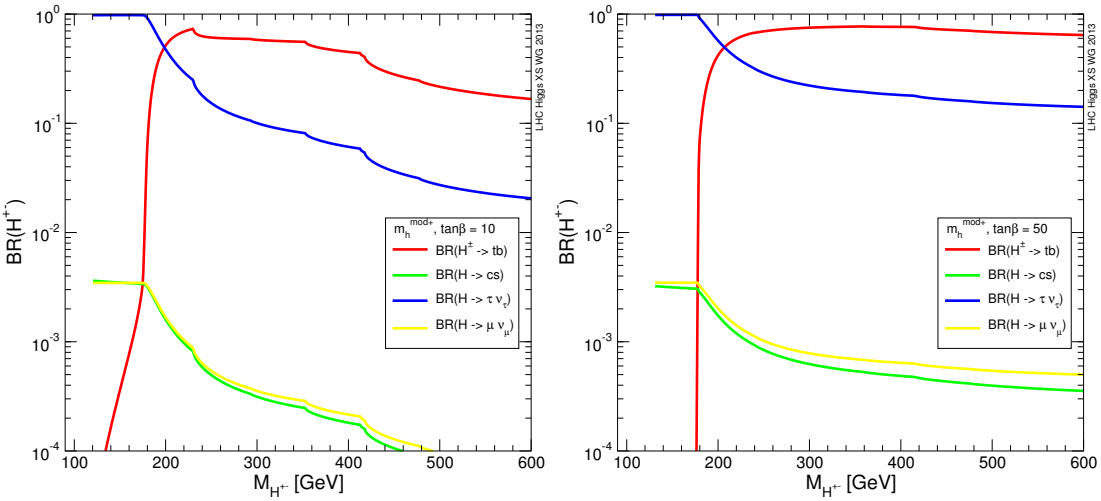



Figure 2.6: Branching fractions of the charged Higgs boson as a function of $m_{H^\pm}$, for $\tan\beta$ values of 10 (left) and 50 (right), in the $m_h^{\text{mod+}}$ scenario of the MSSM [33].

In the presented thesis, only final states with hadronically decaying $\tau$ lepton are considered.

### 2.4.5 Experimental Constraints on Charged Higgs

The ATLAS [34–39] and CMS collaborations presented [40–42] model independent limits for the existence of the charged Higgs bosons. In particular, interpreted in the context of the MSSM, $H^{\pm}$ was excluded for nearly all values of $\tan\beta > 1$ in the mass range $80 < m_{H^{\pm}} < 160$, while in the case of large $\tan\beta$ values also in the region $200 < m_{H^{\pm}} < 250$ GeV [39]. These searches are being continued and updated in the Run-2, see [12, 33, 43] and provide the most sensitive model independent constraints on theories with more than one Higgs doublet.

It is worth mentioning that another way in which such theories can be tested is via flavour physics processes, e.g. by doing fits to the leptonic and semileptonic decays, in a channel where bottom quark decays to a strange quark and a photon ($b \to s\gamma$), $B$ mesons mixings and $Z$ boson decays to bottom and anti-bottom quark pair ($Z \to b\bar{b}$). Typically, such limits are even stronger, but they are also model dependent, e.g. from the analysis in [44] the limit reads $m_{H^{\pm}} > 316$ GeV, for all values of $\tan\beta$, but can only be applied to Type-II 2HDMs. A complete list of current limits for the charged Higgs bosons can be found in [1].

## 2.5 Tau lepton and Tau polarimetry

The $\tau$ lepton was discovered in 1975 by Martin Perl (awarded with a Nobel Prize in 1995) and his collaborators at the Stanford Linear Accelerator [45]. It is the first discovered member of the third quark-lepton family with the measured rest mass of $1776.82 \pm 0.16$ MeV [1] what is about 3480 more than the mass of electron. Being so heavy, $\tau$ has a very short mean lifetime ($(290.6 \pm 1.0) \times 10^{-15}$ s) what corresponds to a decay length of $87.11\mu$m [1], and possibility (as the only lepton) to decay both leptonically and hadronically. To the lowest order all decays of $\tau$ leptons are included in the following four processes (in this Chapter processes with $\tau^-$ are shown as an example), also shown in Fig. 2.7:

$$\tau^- \to \nu_\tau \bar{\nu}_e e^- \tag{2.35}$$

$$\tau^- \to \nu_\tau \bar{\nu}_\mu \mu^- \tag{2.36}$$

$$\tau^- \to \nu_\tau \bar{u} d \to \nu_\tau \,\text{hadrons} \tag{2.37}$$

$$\tau^- \to \nu_\tau \bar{u} s \to \nu_\tau \,\text{hadrons}. \tag{2.38}$$

In Table 2.4 the branching fractions for both hadronic and leptonic decays are listed. In 35.2% of the cases $\tau$ lepton decays leptonically and in 64.8% of the time into one or more hadrons. Considering only hadronically decaying $\tau$ leptons, decays with only one charged particle (the so-called *1-prong*) occur in about 72% of the time and with three charged particles

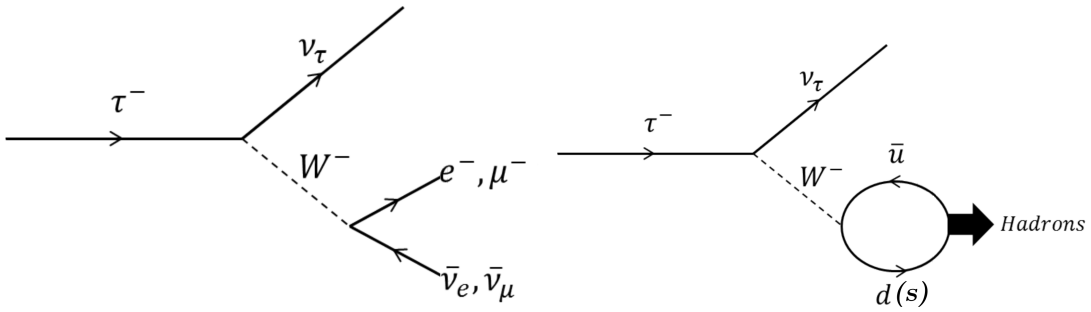



Figure 2.7: Leptonic and hadronic decay of $\tau$ lepton. The virtual $W^-$ created in this reaction then couples to an additional pair of leptons, $e^-\bar{\nu}_e$, $\mu^-\bar{\nu}_\mu$ or quarks, $\bar{u}d$, $\bar{u}s$. All other quark pairings, such as $\bar{c}d, \bar{c}s$, are too massive to be produced on shell.

(the so-called *3-prong*) in about 23% of the time[3]. As it is presented in Table 2.4 (right) the hadronic final states are dominated by $\pi^\pm$ and $\pi^0$ mesons, but there is also a small fraction of decays containing $K^\pm$ and $K^0$ mesons.

It is worth mentioning that measurements of leptonic branching ratios of $\tau$ decays and those of the lifetime enable precise tests of the lepton universality, one of the fundamental building blocks of the Standard Model.

| **Leptonic Decay Modes** | |
|---|---|
| Decay Mode | Branching Fraction |
| $e^-\bar{\nu}_e\nu_\tau$ | $17.85 \pm 0.05$ % |
| $\mu^-\bar{\nu}_\mu\nu_\tau$ | $17.36 \pm 0.05$ % |
| Total Leptonic | 35.21 % |

| **Hadronic Decay Modes** | |
|---|---|
| Decay Mode | Branching Fraction |
| $\pi^-\nu_\tau$ | $10.91 \pm 0.07$ % |
| $\pi^-\pi^0\nu_\tau$ | $25.52 \pm 0.10$ % |
| $\pi^-\pi^+\pi^-\nu_\tau$ | $8.99 \pm 0.06$ % |
| $\pi^-\pi^0\pi^0\nu_\tau$ | $9.27 \pm 0.12$ % |
| $h^-\omega\nu_\tau$ | $1.99 \pm 0.08$ % |
| $\pi^-\pi^+\pi^-\pi^0\nu_\tau$ | $2.70 \pm 0.08$ % |
| $\pi^-3\pi^0\nu_\tau$ | $1.04 \pm 0.07$ % |
| Total Hadronic | 64.79 % |

Table 2.4: The branching ratios for the hadronic and leptonic $\tau$ decay modes [1].

[3] The 5-prong decay has only a fraction of about 0.1% and higher-prong decays are even more suppressed.

### Tau polarimetry

The $\tau$ leptons are the only leptons whose spin information is preserved in kinematics of its decay products recorded by the ATLAS detector due to the short-enough lifetime of $\tau$'s and their parity-violating weak decays.

The $\tau$ polarization, $P_\tau$, the measure of the asymmetry of the cross section for positive ($\sigma_+$) and negative ($\sigma_-$) helicity of $\tau$ lepton, is defined by [46]

$$P_\tau = \frac{\sigma_+ - \sigma_-}{\sigma_+ + \sigma_-}. \tag{2.39}$$

The value of $P_\tau$ provides an insight into the Lorentz structure of the $\tau$ production mechanism, where the positive and negative helicity states and right(left)-handed chiral states are equivalent in the assumed relativistic limit. In particular, $P_\tau$ is a measure of the degree of parity violation in the interaction, i.e. in $W^\pm \to \tau^\pm \nu$ decays, the $W^-$ is expected to couple to a left-handed $\tau^-$ and the $W^+$ to a right-handed $\tau^+$ what corresponds to the $\tau$ polarization equal $P_\tau = -1$. Whereas, the parity-conserving decay results in a value of $P_\tau = 0$ what is the case for the decay of the SM scalar Higgs boson to $\tau$ lepton pairs. On the other hand, an MSSM charged Higgs boson decaying via $H^\pm \to \tau^\pm \nu$, being a spin-0 particle, would lead to a prediction of $P_\tau = 1$.

$\tau$ leptons always couple to a left-handed $\tau$ neutrino, $\nu_\tau$, since parity is maximally violated in their charged-current weak decays. The angular distribution of the $\tau$ decay products, due to the angular momentum conservation, depends strongly on the $\tau$ spin orientation. Since, there is only one neutrino in the final state in the hadronic decay modes they are especially well suited to determine the $\tau$ spin orientation (see Fig 2.8 where the decays of $W^\pm \to \tau_L \nu$ and $H^\pm \to \tau_R \nu$ are illustrated).

The angle $\theta$ between the $\tau$ flight direction and hadronic decay products in its rest frame is the basic observable sensitive to $\tau$ polarization. In the relativistic limit, i.e. $E \gg m_\tau$, the angle $\theta$ is related to the ratio of the energy of the hadronic decay products to the $\tau$ energy in the laboratory frame. Though $\cos\theta$ distribution is difficult to measure experimentally, for distribution sensitive to the polarization state of $\tau$, i.e. one-prong hadronic decays via an intermediate $\rho$ meson, what happend in about 25%: $\tau^\pm \to \rho^\pm(\to \pi^\pm \pi^0)\nu_\tau$, an additional observable $\cos\psi$ is defined in the $\rho$ rest frame, where $\psi$ is the angle between the flight direction of the $\rho$ meson and the charged pion. This observable is related to the kinematics of the final state charged and neutral pions as follows[4] [46]:

$$\cos\psi = \frac{m_\rho}{\sqrt{m_\rho^2 - 4m_\pi^2}} \frac{E_{\pi^-} - E_{\pi^0}}{|\mathbf{p}_{\pi^-} + \mathbf{p}_{\pi^0}|}, \tag{2.40}$$

[4]This procedure is also valid for one-prong hadronic decays of $\tau$ leptons via an intermediate $a_1$ meson what happens in about 9% of cases: $\tau^\pm \to a_1^\pm(\to \pi^\pm \pi^0 \pi^0)\nu_\tau$.

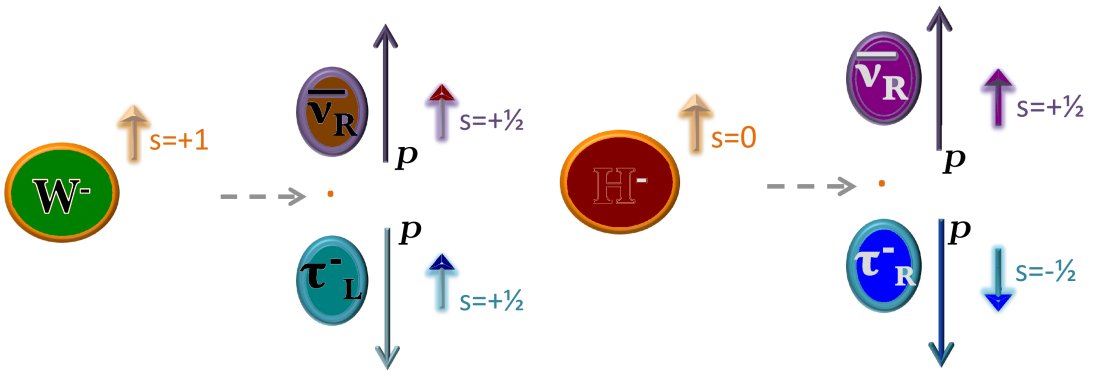



Figure 2.8: The $W^- \to \tau_L^- \bar{\nu}_{\tau,R}$ (left) and $H^- \to \tau_R^- \bar{\nu}_{\tau,R}$ (right) decays, in the rest frame of the bosons. Due to W boson having spin 1, the $\frac{1}{2}$-spin directions of the subsequent leptons must be compensated. Moreover, since the state of the $\nu$ is fixed by nature, the spin of the $\tau^-$ lepton has to be oriented opposite to the direction of flight. Whereas, the $H^\pm$ spin$= 0$ causes that the vectors of the spin of the subsequent leptons must be in opposite directions. Inspired by [47].

where the particle energies and momenta are measured in the laboratory frame and $m_\rho$, $m_\pi$ denote the masses of $\rho$ and $\pi$ mesons, respectively.

This is why a suitable observable for $\tau$ polarization, the "charged asymmetry" ($\Upsilon$), is given by

$$\Upsilon = \frac{E_T^{\pi^\pm} - E_T^{\pi^0}}{p_T} \approx 2\frac{p_T^{trk}}{p_T} - 1 \tag{2.41}$$

and measures the energy sharing between the $\pi^\pm$ and $\pi^0$ in the $\tau$ decay relative to the visible momentum of $\tau$. Experimentally, the energy associated with $\pi^\pm$ is given by the transverse momentum of the single track $(p_T^{trk})$ associated with the $\tau$ hadronic visible decay candidate. The energy ascribed to $\pi^0$ is calculated as the difference between the transverse momentum of visible products of hadronic $\tau$ decay measured in the calorimeter and the transverse momentum of the track of the $\tau$ candidate.

In the context of presented analysis it is worth to emphasize that:

- in $\tau^\pm \to \rho^\pm \nu_\tau$ decays, to the conserve angular momentum, *transversely polarised* $\rho$ is favored in left-handed $\tau$ decays leading to a symmetric energy sharing between $\pi^\pm$ and $\pi^0$,
- whereas right-handed $\tau$ leptons preferentially decay to *longitudinally polarised* $\rho$ what leads to an asymmetric energy sharing.

# Chapter 3

# The Experiment

Below the ATLAS detector at the LHC, that is the experimental apparatus pertaining to the presented analysis is described. The experimental data were taken during running periods in 2015 and 2016 from proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV.

Section 3.1 introduces the LHC machine while Sec. 3.2 gives an overview of the components of the ATLAS detector operating during the Run-2 period of the LHC. Section 3.3 contains discussion of used data and its format, as well as discussion of Monte Carlo events and detector simulation. Finally, Section 3.4 presents the general procedures for reconstructing and identifying the physics objects.

## 3.1 The Large Hadron Collider

The LHC, the world's most powerful particle accelerator and collider for protons and heavy ions and thus the strongest telescope looking inside the structure of matter, currently collides protons with unprecedented centre-of-mass energy of 13 TeV, with the number of collisions exceeds 600 millions per second. It was designed as a discovery machine, therefore primary goals for the LHC are the tests and verification of the SM in the quest of deeper understanding of the fundamental rules governing the Universe. Especially, to study the properties of the Higgs boson and to find any hints of new physics.

The collider is situated at CERN (the European Organization for Nuclear Research) near Geneva in the 27 km long tunnel of the former Large Electron-Positron collider located between 50 and 175 meters below ground level and crossing the France - Switzerland border. Both protons and heavy ions are accelerated in two beam pipes in opposite direction. In proton-proton running mode both beams can contain up to 2808 bunches with $10^{11}$ protons in each bunch, with the 25 ns time distance between them. A magnetic field of up to 8.3 T generated by 1232 superconducting dipole magnets is used to bend the beams, where 392 quadrupole

magnets are used to focus them.

The LHC has been collecting data in two periods so far:

- *Run-1 period*, covering years 2011-2012: data taken at $\sqrt{s} = 7$ and 8 TeV respectively,
- *Run-2 period*, covering years 2015-2018: data taken at $\sqrt{s} = 13$ TeV.

The four major experiments at the LHC, ATLAS, CMS, LHCb [48] and ALICE [49] are located in the places where the beams are collided, in so-called *interaction points* (IP). Whereas, ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) were designed as multipurpose detectors going to perform a wide range of measurements and searches, the LHCb (Large Hadron Collider beauty) focuses on flavour physics and ALICE (A Large Ion Collider Experiment) is specialized in measurements of heavy-ion collisions, which are used to investigate quark-gluon plasma and QCD processes. In Fig. 3.1 a schematic overview of the LHC and its experiments can be found.

## 3.2 The ATLAS experiment at the LHC

The 40 meters long ATLAS detector is the largest particle detector at the LHC. It has the diameter of 25 meters and weights about 7000 tons. It consists of several layers of different sub-detector systems, which identify particles, and measure their momenta and energies. A schematic view of the ATLAS detector is schown in Fig. 3.2.

The physics program ranges from precise measurements of the SM predictions to searches for new, yet unrevealed phenomena. The accuracy needed for these physics measurements requires very high resolution and efficiencies of various sub-detectors. The high interaction rate and large particle flux necessitate radiation-hard elements with very fast readout.

The basic design criteria of the detector took into account the following requirements [2]:

- excellent electromagnetic calorimetry for electron and photon identification and measurements, complemented by full-coverage hadronic calorimetry for accurate jet and missing transverse energy ($E_T^{miss}$) measurements;
- high-precision muon momentum measurements, with the capability of accurate measurements at the highest collision rates using the external muon spectrometer alone;
- efficient charged particle tracking at high luminosity for high transverse momentum ($p_T$) lepton-momentum measurements, electron and photon identification, $\tau$ lepton and heavy-flavour identification and full event reconstruction capability at lower luminosity;

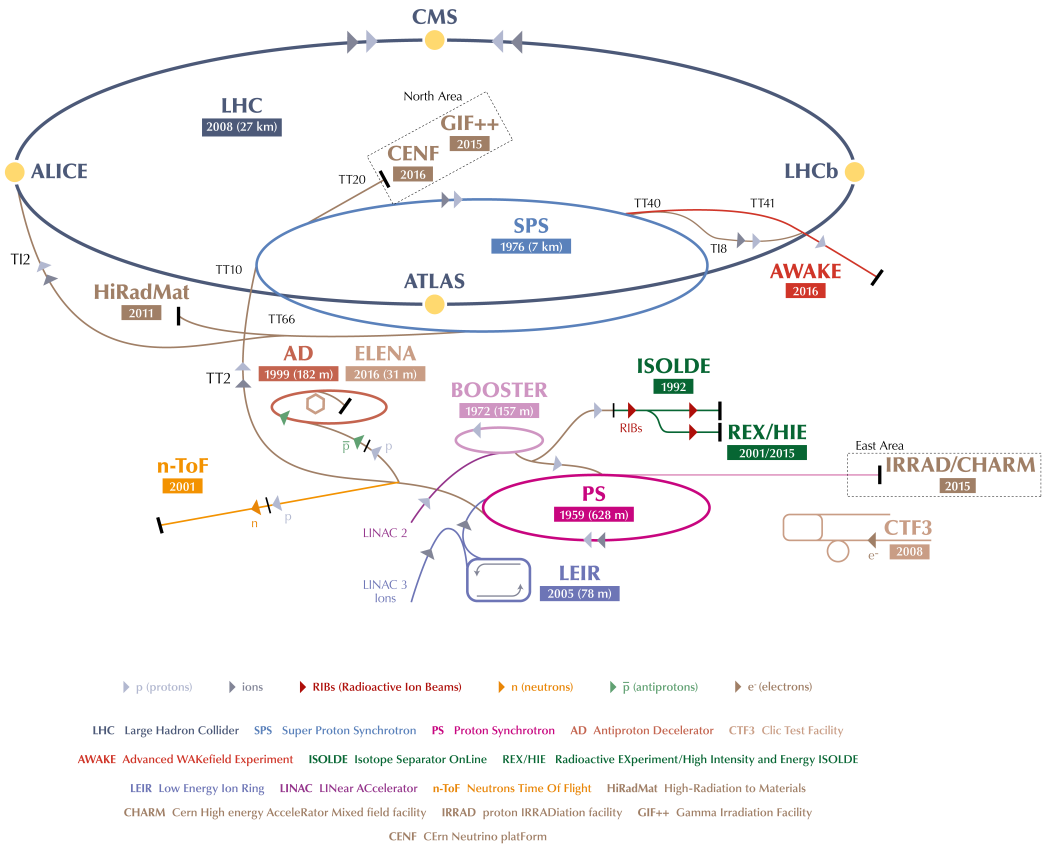



Figure 3.1: The CERN accelerator complex, including the LHC and old accelerators now used as LHC pre-accelerators as well as various other experiments at CERN. The four main experiments (ATLAS, CMS, LHCb and ALICE) are shown as a yellow dots [50].

- maximal coverage of the solid angle (as close to $4\pi$ as possible)
- triggering and measurements of particles at low-$p_T$ thresholds, providing high triggering efficiencies for most physics processes of interest at the LHC.

It is worth mentioning that ATLAS is an international Collaboration which consists of about 3000 scientists from about 182 institutions around the world representing 38 countries, where more than a good third of them are PhD students.

### 3.2.1 The ATLAS coordinate system

ATLAS uses a right-handed coordinate system where the beam line defines as the $z-$axis. The $x-$axis points from the nominal interaction point to the center of the ring and the $y-$axis points

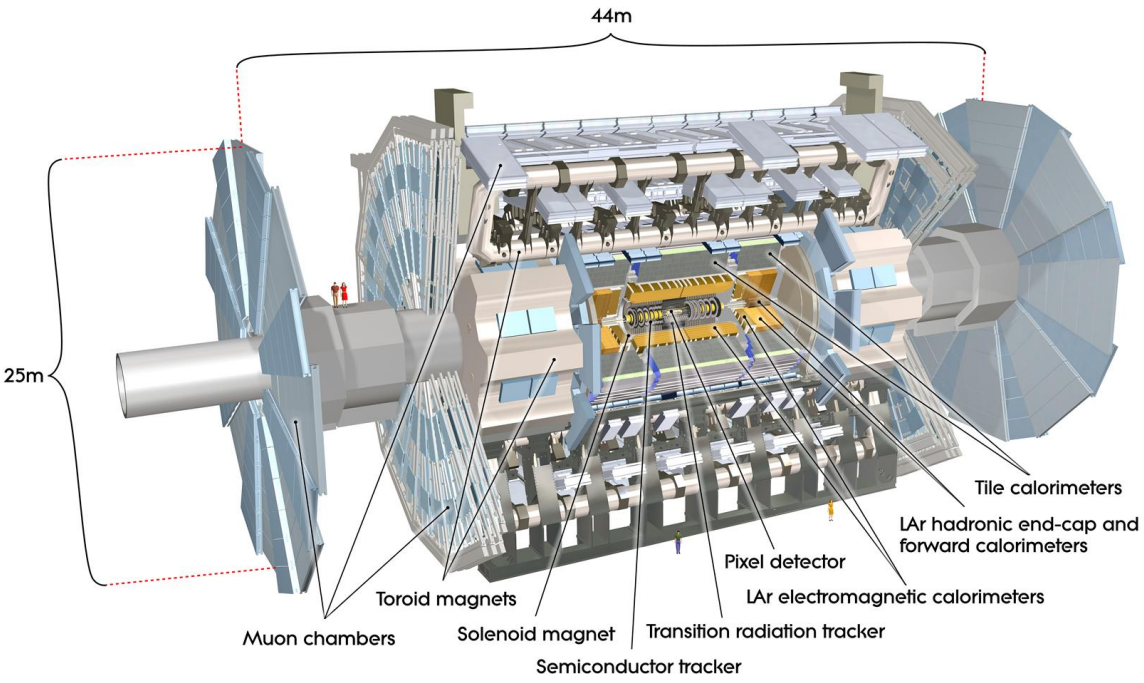



Figure 3.2: Schematic cut-away view of the ATLAS detector showing the different layers of the inner detector, the calorimeters, toroid magnets and muon spectrometer [2].

upwards. The origin of the coordinate system is located at the nominal interaction point.

The azimuthal angle $\phi$ is measured from $x-$axis and the polar angle $\theta$ is defined with respect to the $z-$axis direction. The pseudorapidity is given by

$$\eta = -\log\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{3.1}$$

The angular separation of two objects in the $\eta - \phi$ plane is described as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \tag{3.2}$$

where $\Delta\eta$ is a difference in pseudorapidity and $\Delta\phi$ in azimuthal angle. Therefore, $\Delta R$ is Lorentz invariant under boosts along the $z-$axis, which is convenient when working with objects originating from hadron collisions.

### 3.2.2 The Inner Detector

The Inner Detector (ID) is the innermost part of ATLAS. It is of cylindrical shape 6.2 m long having diameter of 2.1 m and consists of several subsystems: the pixel detector, semiconductor tracker (SCT) and transition radiation tracker (TRT), see Fig. 3.3. The ID is placed in 2T

solenoid magnetic field. It is used to measure the trajectories (tracks) and the momenta of charged particles with transverse momentum above $p_T > 0.5$ GeV. Using the reconstructed tracks the interaction vertex can be found.

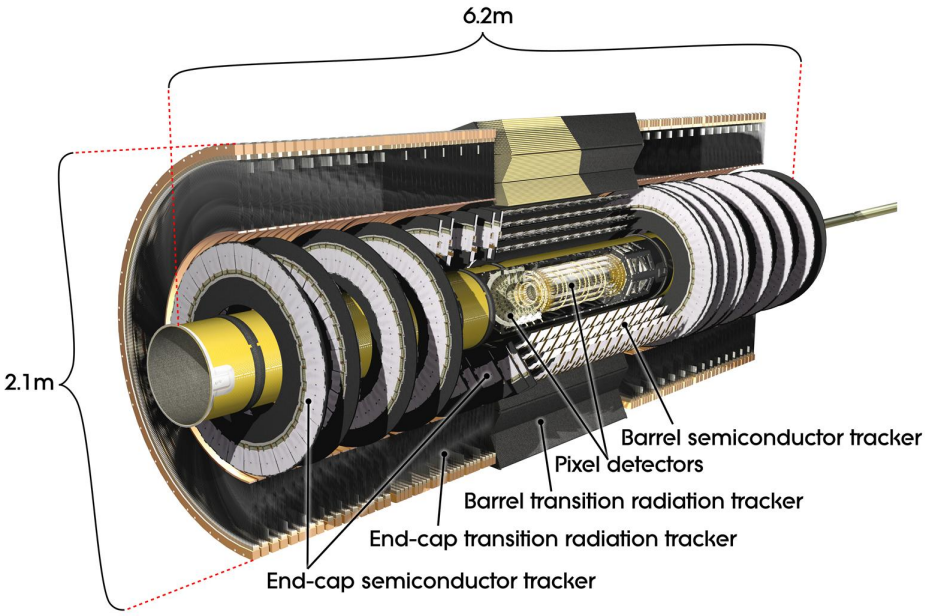



Figure 3.3: Cut-away view of the ATLAS tracking detectors. In this longitudinal view the different tracking layers around the LHC beam pipe are shown. The interaction point is in the centre of tracking detector [2].

Since ID is placed in the vicinity of the beam pipe and the interaction point, the detector material is exposed to huge amounts of radiation and high temperatures what requires fine detector granularity and radiation hardness. Additionally, the pixel detector and SCT are cooled down to around -7° C to mitigate damages.

The *pixel detector* which is composed of four barrel layers and two end-caps with three discs each. It has the finest granularity of the ID system sub-components. The pixel detector is placed in the immediate vicinity of the beam pipe (at the radious of $\sim 3$ cm) and offers precise measurement of the charged particle tracks. This feature allows for precise reconstruction of vertices, which is important for tagging of short-lived objects such as B-hadrons or $\tau$ leptons. The barrel and end-cap layers of the pixel detector are made of small silicon semiconductors called *pixels* and each of the pixel layers is segmented in $R$, $\phi$ and $z$. The innermost barrel layer is the insertable $B$-layer (IBL), which was added during the shutdown period between Run-1 and Run-2 to recover the loss of sensitivity due to radiation damage [51]. The intrinsic accuracy in the pixel detector is 10 $\mu$m in the $R-\phi$ plane in the barrel and endcap, and 115 $\mu$m

along the $z-$direction.

The *semiconductor tracker* uses similar concepts as the pixel detector. However, by using larger semiconductors having the strip-like geometry the SCT has a worse resolution. It covers up to $|\eta| < 2.5$. The SCT is build out of four double layers of silicon strip detectors in the barrel part and nine layers in the end-caps. This ensures that every charged particle traverses at least four layers of detectors. For each double layer in the barrel region one set of the silicon strip modules is aligned to the beam axis and the other set is rotated by 40 mrad, which enables to measure the position along the beam axis. A hit resolution of 17 $\mu$ m in the $R-\phi$ plane and 580 $\mu$ m along the $z-$axis is achieved.

*The transition radiation tracker*, a straw-tube tracker, is the outermost part of the inner detector covering $|\eta| < 2.5$. It is made of gas-filled tubes (mixture of xenon (70%), carbon dioxide (27%) and oxygen (3%)), which are stabilized by carbon fibers and with a gold plated tungsten wire. In the TRT barrel the tubes are aligned along the $z-$axis, whereas in the end-caps they are positioned radially. Hence, the position measurement in $R-\phi$ plane in the barrel region is possible with a nominal hit resolution of 130 $\mu$m. Through the high number of measured points per track, usually 36 points, the TRT contributes considerably to the measurement of tracks. Additionally, it can be used for particle identification since the layers of straws are interleaved with polypropylene fibres (barrel region) and foils (end-caps). The X-ray range transition radiation is emitted as a charged particle crosses a boundary between the media of differing dielectric constants, and can be subsequently absorbed by the gas in the straw tube. Since the transition radiation depends on the radiating particle Lorentz gamma factor, then typically it is the largest for electrons allowing their discrimination from other particles.

### 3.2.3 The Calorimeter System

The ATLAS calorimeters, see Fig. 3.4, consist of an electromagnetic (EM) calorimeter covering the region $|\eta| < 3.2$, a hadronic barrel calorimeter covering $|\eta| < 1.7$, hadronic end-cap calorimeters covering $1.5 < |\eta| < 3.2$, and forward calorimeters covering $3.1 < |\eta| < 4.9$ region [2]. This system surrounds the solenoid which produces the magnetic field inside the ID.

The EM and hadronic calorimeters are the *sampling* calorimeters. They are made of alternating layers of active and absorbing material. Absorbers induce the particle showers while active material is used to detect the shower particles producing a signal proportional to the initial energy. The energy of the passing particle can be calculated using these measurements[1].

[1]Electrons or positrons through bremsstrahlung lose $1/e$ of their total energy in a single radiation length, $X_0$. For hadronic showers hadrons will lose energy through inelastic interactions, which is parameterised as the mean

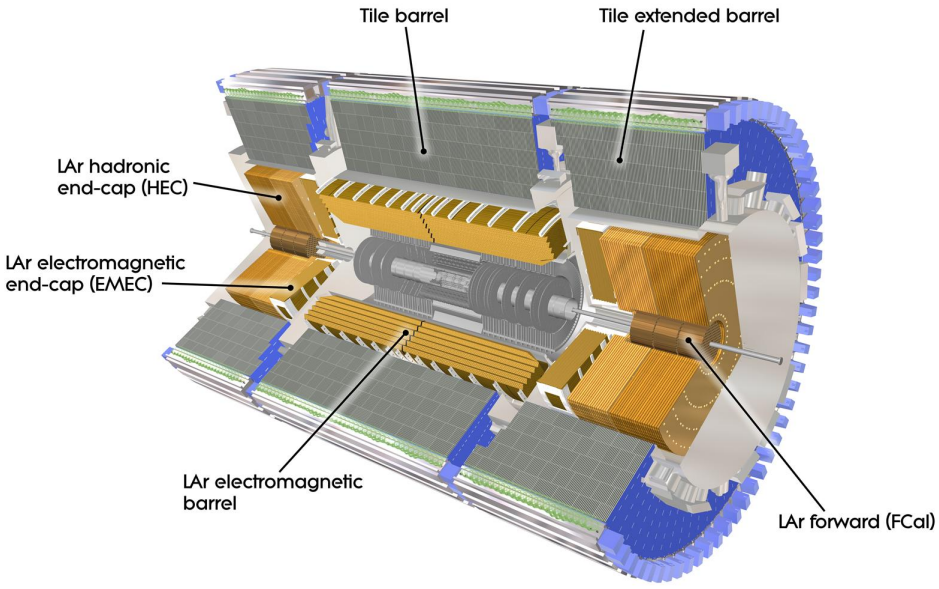



Figure 3.4: Cut-away view of the ATLAS calorimetry. The three distinct cylinders, barrel and end-caps, are visible. The smaller radial regions use the LAr technology requiring cryostats, whereas the outer cylinders use scintillator tiles embedded in an iron absorber structure. The end-caps can be moved longitudinally along the LHC beam line for creating access space to maintain the barrel region [2].

To obtain the best EM energy measurement resolution and to avoid energy leakage the EM calorimeter is 22-24$X_0$ thick. Where the total thickness of the calorimeters is about 10$\lambda$. It is worth also mentioning that a punch-through into the muon system is prevented, what is essential for the calorimeters.

*The electromagnetic calorimeter* consists of the barrel region which covers $|\eta| < 1.475$ and the end-cap region within $1.375 < |\eta| < 3.2$, as can be seen in Fig.3.4. It uses liquid argon (LAr) as the active material and lead as the absorber. The barrel part is composed of three layers of modules, where the first one has a good segmentation in pseudorapidity which allows precise measurement of the impact position of electrons and photons. The second and third layer are used to collect the bulk and tail of the electromagnetic showers and have coarser structure. In order to avoid cracks and enable a full $\phi$ coverage in all layers the calorimeter modules are arranged into the accordion-shaped architecture. Each of the end-caps is divided into an outer and inner wheel. The outer wheels cover $1.375 < |\eta| < 2.5$ and the inner wheels extend further to $2.5 < |\eta| < 3.2$. The inner wheel is made of three layers of modules, while the outer one uses

free path ($\lambda$) and gives the characteristic scale of it, where $\lambda \sim 35A^{1/3}$ gm$^2$ and A is the atomic weight of the absorption material.

only two layers. In order to account for energy losses of electrons and photons before they reach the calorimeter system an additional active LAr layer, the so-called presampler detector, is used in the $|\eta| < 1.8$ region. The EM energy resolution is: $\sigma(E)/E = 10\%/\sqrt{E(\text{GeV})} \oplus 0.7\%$.

*The hadronic calorimeter* used to measure the energy of hadrons is located outside the EM calorimeter. It is composed of three different parts: a tile calorimeter for the barrel part, LAr end-cap calorimeters at both sides and LAr forward calorimeters also at both sides to cover the high$-|\eta|$ region. The tile calorimeter consists of a central barrel ($|\eta| < 1.0$) and two extended barrels ($0.8 < |\eta| < 1.7$). It is composed of scintillating plastic tiles which play the role of the active material and absorber plates made of steel. The scintillating light produced by the shower particles is transmitted to photomultiplier tubes via the wavelength shifting fibres. LAr end-cap calorimeters cover $1.5 < |\eta| < 3.2$ and are equipped with liquid argon, as an active material, and copper absorbers. The forward calorimeter extends the coverage to $3.1 < |\eta| < 4.9$ and is placed in the forward region of the detector. The first module with copper absorber is used for electromagnetic calorimetry, while the other two modules with tungsten absorber are used for hadronic calorimetry. In all three modules, LAr is used as the active material. The forward calorimeter also shields the muon spectrometer against non-muon particles in high $\eta$ region and allows for a precise $E_T^{miss}$ measurement. The hadronic calorimeter energy resolution is: $\sigma(E)/E = 100\%/\sqrt{E(\text{GeV})} \oplus 10\%$.

### 3.2.4 The Muon Spectrometer

The ATLAS muon spectrometer (MS), shown in Fig. 3.5, is the outermost part of the ATLAS detector. It detects mainly muons, since all other charged particles should be absorbed in the calorimeters. MS uses the high precision tracking chambers to measure the muon tracks which are bent in the toroidal magnetic field. The measurement of the track curvature is used to determine the muon transverse momentum, from a few GeV up to the TeV, and charge of the muons. The MS is composed of three regions: the barrel, the end-cap and the transition region between the two aforementioned ones. The barrel toroid provides the magnetic field of about 1.5 to 5.5 T and covers the range of $0 < |\eta| < 1.4$, the end-cap region covers $1.6 < |\eta| < 2.7$ with the magnetic field of 1 to 7.5 T and the transition region covering the $1.4 < |\eta| < 1.6$ region where magnetic fields of both systems are used. In the barrel region three layers of muon chambers are used, whereas four wheels perpendicular to the beam axis are installed in the end-cap region.

In most parts of the MS the trajectories of the muons are measured by monitored drift tubes, which provide a spatial resolution of 35 $\mu$m per chamber. The one exception is the range of $2.0 < |\eta| < 2.7$ in the forward region, where the cathode strip-chambers are used in the

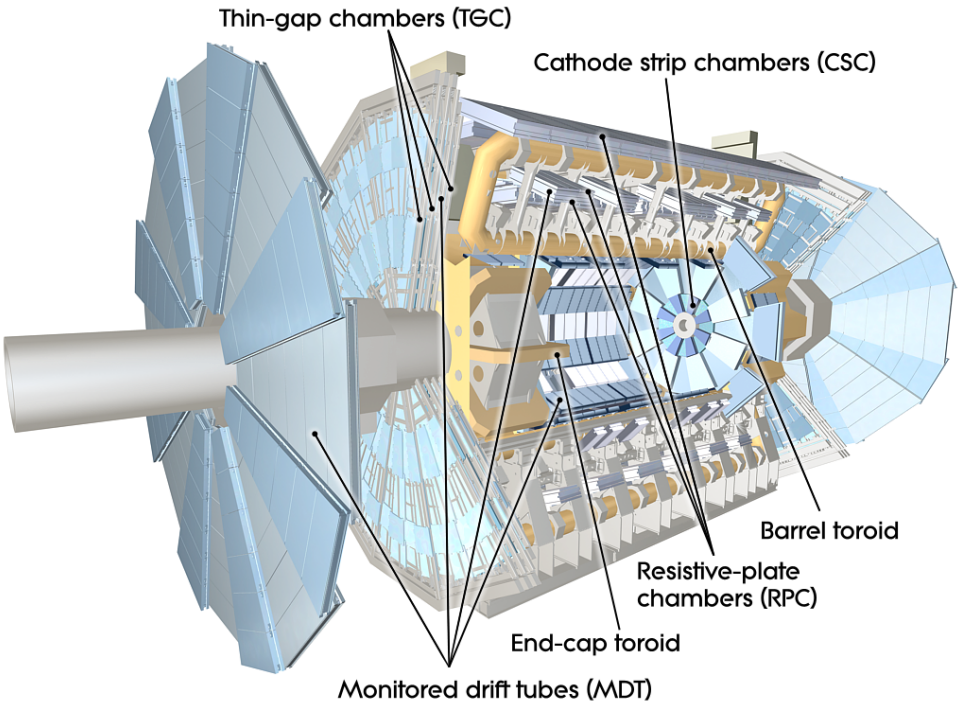



Figure 3.5: Schematic cut-away overview of the ATLAS muon spectrometer with its different regions and the toroid system [2].

innermost layer and provide a higher rate capability and better time resolution.

The MS is also designed to trigger on muons with dedicated trigger chambers for particles in the range $|\eta| < 2.4$. In order to achive a response time of a few nanoseconds, the resistive plate chambers are used in the barrel region and thin gap chambers in the end-cap regions.

### 3.2.5 The Trigger System

The ATLAS trigger and data acquisition system [52] are designed to select interesting events created with very high rate in pp interactions at the LHC and to read out the detector information with minimal dead time due to hardware and software limitations. In order to attain this, the system is composed of the hardware-based level-1 trigger *(L1)* and the software-based high-level trigger *(HLT)*. The L1 trigger consists of the L1 calorimeter trigger system (L1Calo), the L1 muon trigger system (L1Muon) and L1 topological trigger modules (L1Topo). The L1Calo triggers on high-$E_T$ objects coming from electrons, photons, jets and taus, and events with large total $E_T^{miss}$. The L1Muon triggers on muons for each of the predefined $p_T$ thresholds. The L1Topo combines information from L1Calo and/or L1Muon into topological variables, which are based on geometric properties of the event. The maximum L1 accept rate is 100 kHz. The

L1 trigger defines also the Region-of-Interest (RoI), which determines the locations in $\eta$ and $\phi$ where relevant features are identified.

The L1 trigger decision and RoI are passed to the HLT trigger, where the event recording rate of around 1.5 kHz is achieved with the decision time of 200 ms. The HLT runs on computing cluster accessing data from the RoI, applying offline-like algorithms using the full event information. Events accepted by the HLT are transfered to local storage at the experimental site and exported to the Tier-0 facility at CERN's computing center for offline reconstruction.

## 3.3 Data and Monte Carlo simulations

### 3.3.1 Data

The total amount of data from proton-proton collisions delivered by the LHC and recorded by the ATLAS experiment in years 2015 and 2016 at 13 TeV centre-of-mass energy during stable beams is 36 fb$^{-1}$, see Fig. 3.6. The delivered luminosity is the luminosity evaluated before any trigger decision, which accelerator delivered to certain experiment. The recorded luminosity is actual recorded disk data, after corrections for the dead time and operational problems in sub-detectors filtered by a data acquisition system. Precise measurement of the luminosity is an essential task, since the uncertainty on the delivered luminosity influences precision of the physical processes cross section determination. Moreover, it is an important ingredient of the searches for the physics beyond the SM by normalising the background calculations. The integrated luminosity is given by the time integral over the instantaneous luminosity $\mathscr{L}$, which can be defined by beam and machine parameters[2] [53]:

$$\mathscr{L} = \frac{f_r n_b N_1 N_2}{2\pi \sum_x \sum_y}, \tag{3.3}$$

where $f_r$ is the revolution frequency; $n_b$ gives the number of colliding bunches in each beam; $N_1$ and $N_2$ denote the number of protons in each bunch in beam 1 and 2; $\sum_1$ and $\sum_2$ are the horizontal and vertical beam widths at the interaction point assuming gaussian distributed shapes of the beams and head-on collisions. The beam widths are extracted in a van der Meer (VdM) scan during which two beams are first centered on each other and then they are displaced in discrete steps of known distances at the same time recording the relative change of the event counting rate.

[2]Note that luminosity in terms of beam densities $\rho_1$ and $\rho_2$ is given by: $\mathscr{L} = f_r n_b N_1 N_2 \int \rho_1(x,y)\rho_2(x,y)dxdy$. Hence, only if the integral factorises into independent $x$ and $y$ components the equation 3.3 is true.

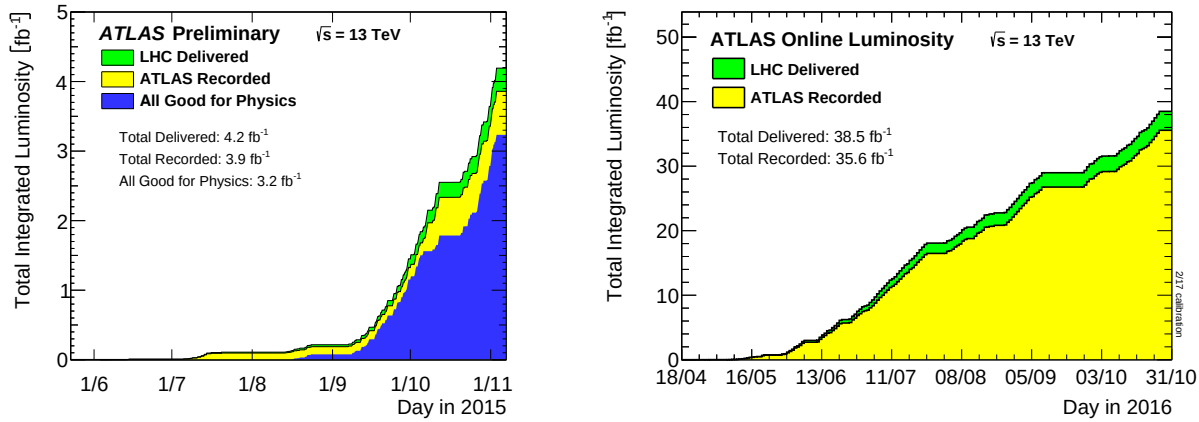



Figure 3.6: Integrated luminosity from pp collisions at $\sqrt{s} = 13$ TeV during stable beams in 2015 (left) and 2016 (right). The total integrated luminosity delivered by the LHC is shown in green, the amount of data recorded by ATLAS is shown in yellow and the subset that is used for physics analyses is shown in blue [54].

The number of events, $N$, expected for a certain process is given by a product of integrated luminosity and cross section, $\sigma$, for this process:

$$N = \sigma \int dt \mathscr{L}. \tag{3.4}$$

ATLAS monitors the delivered luminosity by measuring $\mu_{vis}$, the visible number of interaction per bunch crossing, with a variety of independent detectors and using several different algorithms [55].

$$\mathscr{L} = \frac{\mu_{vis} f_r}{\sigma_{vis}}, \tag{3.5}$$

where $\mu_{vis} = \varepsilon\mu$ and $\varepsilon$ gives the efficiency of a particular detector and algorithm, and $\mu$ is the average number of inelastic interactions per bunch crossing. The visible cross section for the same detector and algorithm is defined by $\sigma_{vis} = \varepsilon\sigma_{inel}$ where $\sigma_{inel}$ is the pp inelastic cross section. Since $\mu_{vis}$ is a directly measurable quantity, the calibration of the luminosity for a particular detector and the algorithm amounts to determine $\sigma_{vis}$, which can by obtained during VdM scans. Using equations 3.3 and 3.5 one gets

$$\sigma_{vis} = \mu_{vis}^{max} \frac{2\pi \sum_x \sum_y}{N_1 N_2}, \tag{3.6}$$

where $\mu_{vis}^{max}$ is the visible number of interaction per bunch crossing reported at the peak of the scan curve by the particular algorithm. The measurement of $\mu_{vis}$ and a value of $\sigma_{vis}$ are extracted for each algorithm and each detector. Systematic uncertainties on the luminosity measurement are estimated in part by comparing the luminosity measurements from all algorithms and detectors.

The data events are classified into runs corresponding to the data taking periods of the ATLAS data acquisition system. These runs are further divided into the luminosity blocks (LBs), the basic time unit for storing the luminosity information for physics use, corresponding to a few minutes of data taking each. The data quality of the LBs is ensured by the so-called *Good Run Lists* (GRL). Only data from these lists are used for physics analyses.

Using the GRL requirements, the dataset used for the presented analysis corresponds to 3.2 $fb^{-1}$ taken in 2015 and 32.9 $fb^{-1}$ taken in 2016. The uncertainty on the luminosity measurement is 2.1% and is derived following the methodology similar to that detailed in Ref. [55].

Due to the presence of $10^{11}$ protons in each bunch it is very probable that more than one interaction occurs per bunch crossing. This phenomena is called *in-time* pile-up. What is more, the interactions which happen directly before or after the interaction of interest can also be recorded due to the low time distance of 25 ns between each bunch crossing. This is called *out-of-time* pile-up. Fig. 3.7 presents a distribution of the mean number of interactions per bunch crossing, i.e the mean value of the pile-up for data taken in 2015 and 2016, which was 23.7 [54].

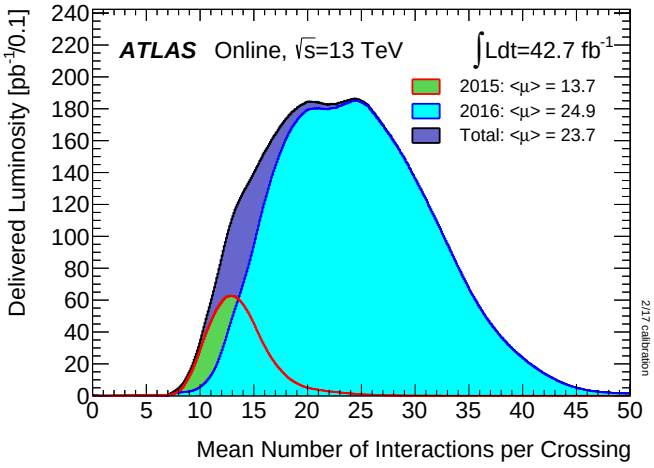



Figure 3.7: Distribution of the mean number of interactions per bunch crossing, i.e. *pile-up* weighted by luminosity for data taken in 2015 and 2016 [54].

Finally, Fig. 3.8 presents the cross sections for production of various particles as a function of the collision energy. Total $\sigma_{inel}$ was measured to be $78.1 \pm 2.9$ mb of $pp$ collisions at a center-of-mass energy $\sqrt{s}$ of 13 TeV with the ATLAS detector [56].

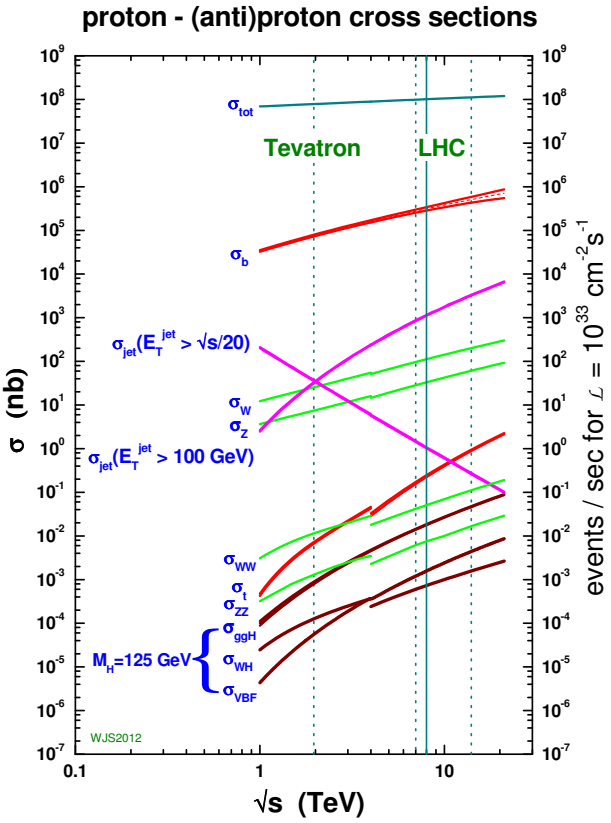



Figure 3.8: Cross sections for $\sqrt{s}$ in range 0.1-20 TeV [57].

### 3.3.2 Event Simulation - MC event generators

A vast majority of recorded events come from known processes described by the SM. Therefore, a main challenge is to find and identify the signals that could arise from processes which have not been yet discovered.

The most commonly employed technique for determining the number and characteristics of the background and signal events are the computer simulations. They are based on the combination of very precise theoretical calculations and advanced numerical methods using the Monte Carlo techniques to generate random events. Hence, event generators are essential tools for particle physics phenomenology at hadron colliders, what is schematically shown in Fig. 3.9. They are used in the planning of new experiments, detector design and performance studies, and in the extraction of theoretical parameters from the measurements themselves.

For the proton-proton collisions several event generators are available and very often for the simulation of the different parts of the collisions different generator is used. Generaly, we can split the way how the generators treat the collision into two groups: parton shower (PS) and matrix element (ME) generators.

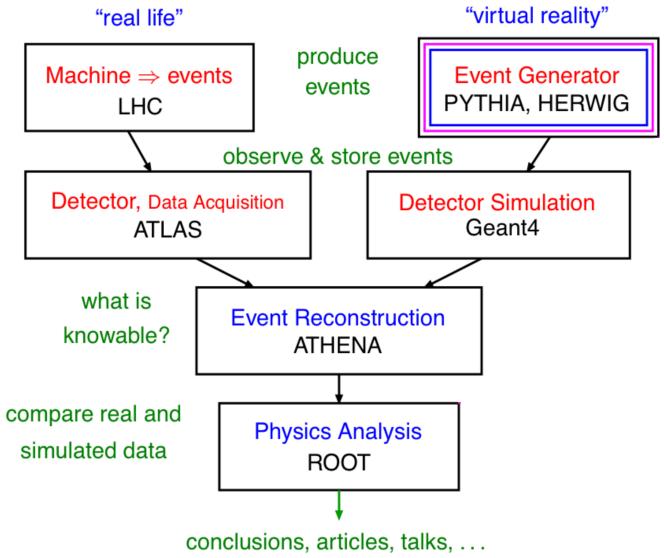



Figure 3.9: Schematic view on the process of simulating particle physics [58].

Parton showers approximate higher-order real-emission corrections to the hard scattering by simulating the branching of a single parton into two partons. It means that in this approach only the lowest order matrix elements are implemented, i.e. basic $2 \to 2$ processes. Initial and final state radiation are added on to the basic process where the showers are assumed to be universal, i.e. the shower evolution does not depend on the details of the hard scattering, but only on the main features: energies and flavours of incoming and outgoing partons, and on the overall $Q^2$ scale for the hard scattering. They locally conserve flavor and four momentum and they respect unitarity, which means that a parton may split into two partons $g \to gg$ or $g \to qq$, or it can radiate gluons $q \to gq$. At the level of sufficiently low ($\sim 1$ GeV) values of the invariant mass of the partons, the pertubative gluon and quark radiation breaks down and the partons are combined into the color neutral states [59, 60].

The second part of treating the hadron-hadron collision are the matrix element generators which calculate the matrix elements to a given order in perturbation theory with different number of final state partons, e.g. $W \to \mu\nu$+jets with additional jets created from initial and final state radiation. Therefore, the emphasis is on the use of exact higher-order matrix elements and on the selecting the kinematic variables in an efficient way. The hadronisation step is not included in the ME generators, thus in the above case the final state is partonic. In order to get an observable state, we have to use a parton shower algorithm to create jets.

The approaches mentioned above have different merits and shortcomings. While fixed-order matrix elements are excellent when simulating well separated, hard partons, they have problems when trying to describe collinear and soft partons, due to the occurrence of large logarithms. Also, obtaining the correct matrix element becomes very cumbersome when we have more

than a handful of partons. With parton showers it is the other way around: hard, wide-angle emissions are poorly approximated, while soft and collinear parton emissions are well described even for very many partons [61].

It is worth remembering that the perfect event generator does not exist. This reflects the limited understanding of physics in many areas. Indeed, a perfect generator can only be constructed once everything is already known, in which case experiments are redundant [60].

### 3.3.3 Detector Simulation

After simulating the physics process using the event generator, the interaction of generated particles with the detector material and the detector response has to be simulated in order to compare our physics predictions to data.

In ATLAS experiment the detector simulation is done with Geant4 [62] toolkit. It was used for simulation of all background Monte Carlo samples used in presented analysis. For signal events a fast detector simulation was used, which is based on parametrisation of the performance of the electromagnetic and hadronic calorimeters [63] and on Geant4 for all other detector components. It is a full detector simulation where all physics processes that can occur in interactions between the particle and the detector material are simulated and every final state particle is propagated through the detector and the response from every active element of the detector is simulated. In order to achive the same structure as in the recorded data, the signal for the final output of the simulation event is digitalised, which allows running of all the reconstruction algorithms to run on simulated events as well.

### 3.3.4 Data formats in ATLAS

The output of the trigger (the Event Filter- final stage of the HLT) are the so-called *RAW data* organised into inclusive streams. Here, the term inclusive means that events can end up in one or more streams, depending on which triggers they pass. Then, after the physics objects reconstruction step Event Summary Data, ESD, are produced. Their reduced content, used in analyses is called Analysis Object Data, AOD. They contain physics objects and other elements of analysis interest. AODs are further brought under selection called derivation resulting in Derived Analysis Object Data, DAOD, to obtain events with the physical observables with appriopriate features for further analysis but with reduced size.

In summary, events simulated by MC generators are processed in several steps:

1. *event generation*: simulation of the interaction between the quarks and gluons in the colliding protons, and also the subsequent parton showering and hadronization and decays

into stable particles,

2. *detector simulation*: calculation of how the particles from the generator interact with the detector material, i.e. how they shower into secondaries and how much energy they deposit in each sensitive element,

3. *digitisation*: turning the simulated energy deposits into the detector response in such a way that they "look" like the raw data from the real detector,

4. *reconstruction of physics objects*: the process is the same as for data.

The above description is presented in Fig. 3.10.

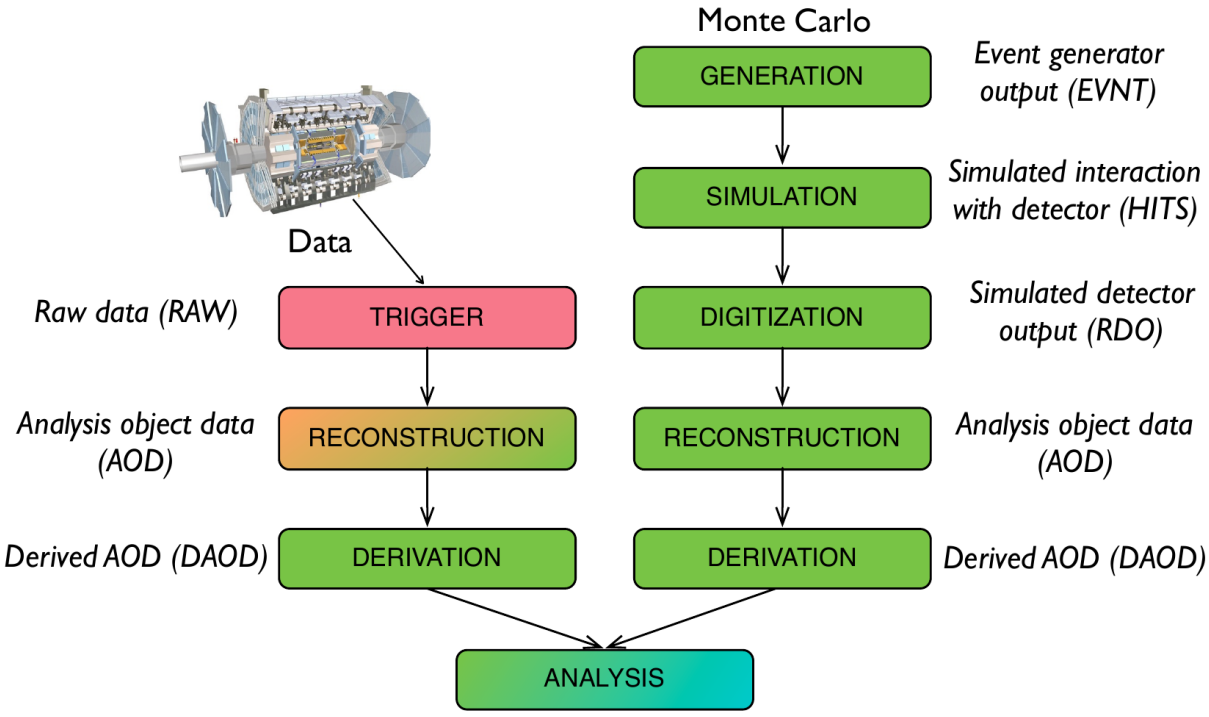



Figure 3.10: Schematic presentation of data and Monte Carlo samples processing flow and data formats used by ATLAS Collaboration [64].

## 3.4 Physics Objects Reconstruction and Identification

In this Section, the general procedures for reconstructing and identyfing the physics objects, i.e. electrons, jets, hadronically decaying $\tau$ leptons, muons, missing transverse energy, etc. used in the analysis presented in this thesis are described.

The process of reconstruction and identification is done for every particle traversing the ATLAS detector using information from the inner detector, calorimeters, and the muon spectrometer. The ATLAS offline reconstruction software processes the RAW data. The building blocks of the object reconstruction are the tracks in the inner detector and muon spectrometer, and clusters of the cells with energy deposits in the calorimeters. By constructing these low-level objects and combining them with dedicated algorithms, the offline software reconstructs electrons, muons, photons, jets, b-jets, hadronic $\tau$ decays and the missing transverse energy. The same reconstruction is done for both data and simulation. Potential differences in reconstruction, identification and trigger efficiencies due to MC modelling are measured and correction factors (the so-called *scale factors*) are then applied to the simulated events at the analysis level.

### 3.4.1 Tracks and Vertices

Charged particle traverses the ID leaving a chain of hits in the pixel, SCT and TRT detectors which is used to form a *track* that represents the trajectory of the particle [65,66]. The curvature of the track in the magnetic field is used to calculate the transverse momentum and to determine the sign of the charge. The application of the quality criteria based on the number of hits in the sub-detectors and depending on the transverse momentum, $p_T$, and pseudorapidity, $\eta$, is also done. Depending on the track $\eta$ and different selection criteria the track reconstruction efficiencies range between 63% and 91% [67].

Tracks are the fundamental (where charged particle trajectories are the fundamental ingredient for the reconstruction and identification of other physics objects) objects. They are also essential also in finding the location of primary vertices of the collisions and secondary or tertiary vertices of the particle decays. Since multiple interactions are expected during one bunch crossing there are also multiple reconstructed vertices. The vertex with the largest scalar sum $\sum p_{T,\text{track}}^2$ and at least three associated tracks is chosen as the primary one, which corresponds to the point where the interaction was the hardest, i.e. to the location of the hard scattering process. Secondary vertices occur at some distance away from the primary one. The efficiency to reconstruct a vertex depends on the number of tracks which are associated with it. The efficiency for two tracks is equal 83%, for three tracks 97%, and for more than 4 tracks close to 100% [68].

In [67–69] detailed information about tracking and vertexing in ATLAS for Run-2 of the LHC can be found.

### 3.4.2 Jets

Due to confinement the partons can only exist in a colourless state, which means that any quark or gluon created in the fragmentation of a parton during high energy collision must hadronise. The hadronisation process produces a shower in the particle detector. The energy deposits and tracks from these showers can be clustered together into narrow cones, called *the jets*, to measure the momentum and energy of the original quark or gluon.

In the analysis presented below the jets are reconstructed using the anti$-k_T$ jet sequential clustering algorithm [70, 71] based on the energy clusters in calorimeters, with a distance parameter of $R = 0.4$. The anti$-k_T$ algorithm defines the following distance measure for two objects $i$ and $j$, such as clusters, that potentially belong to a jet:

$$d_{ij} = \min\left(\frac{1}{\mathrm{p}_{\mathrm{Ti}}^2}, \frac{1}{\mathrm{p}_{\mathrm{Tj}}^2}\right) \frac{(\eta_\mathrm{i} - \eta_\mathrm{j})^2 + (\phi_\mathrm{i} - \phi_\mathrm{j})^2}{\mathrm{R}^2}, \tag{3.7}$$

where $p_{Ti}$ is the transverse momentum of the $i$-th object, $\eta_i$ and $\phi_i$ its pseudorapidity and azimuthal angle, respectively. The minimum distance $d_{min}$ of all $d_{ij}$ is identified by the algorithm. Then if $d_{min}$ is below a certain threshold named $d_{cut}$, particles $i$ and $j$ are combined into a new particle called pseudojet. This step is repeated until there are no cases left where $d_{cut}$ is above $d_{min}$. Then the algorithm considers all created pseudojets to be jets. Since the number of hard anti$-k_T$ jets is unaffected by soft gluon emissions and collinear splitting the algorithm is safe against infrared and ultraviolet divergences. Fig. 3.11 shows an example of jet clustering using the anti$-k_T$ algorithm.

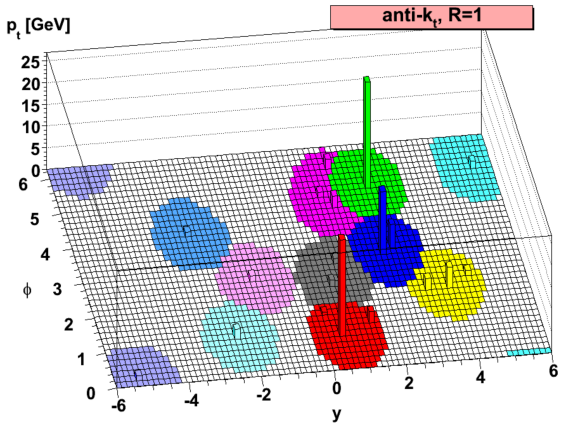



Figure 3.11: The clustering done with the anti$-k_T$ algorithm, where the coloured areas show the clustered jets from that method [71].

For further pile-up suppression a multi-variate technique, the *jet vertex tagger* (JVT) is used. This tagger uses the jet track and vertex information to differentiate between the jets

from the hard scattering process and pile-up and is applied in the presented analysis to jets with $p_T < 60$ GeV and $|\eta| < 2.4$ [72]. A cut on the JVT output is chosen in the way that 92% efficiency for jets from the hard scattering event is achieved.

In terms of the properties of the energy depositions in the calorimeter system, the energy of a reconstructed jet is calibrated with what is called the *jet energy scale* (JES) calibration. This calibration restores the jet energy scale to that of truth jets reconstructed at the particle-level energy scale. Details about jet calibration procedure can be found in Ref. [73, 74].

### 3.4.3 *B*-jet tagging

Since the top quark usually decays into a W boson and a b-quark, it is very important to efficiently identify the jets resulting from hadronization and decays of the b-quark. In order to identify the *b*-jets, the multivariate-based algorithm *MV2c10* is used in the presented analysis [75, 76]. This algorithm exploits the fact that *b*-flavoured hadrons have quite a long mean life time of $\sim 1.5$ ps and combines the impact parameter information with the explicit identification of secondary (several millimeters away from the primary vertex due to time dilation) and tertiary vertices, where the secondary vertex is reconstructed with the tracks of the charged particles within a jet, see Fig. 3.12. A cut on the MV2c10 output to yield a 70% efficiency of

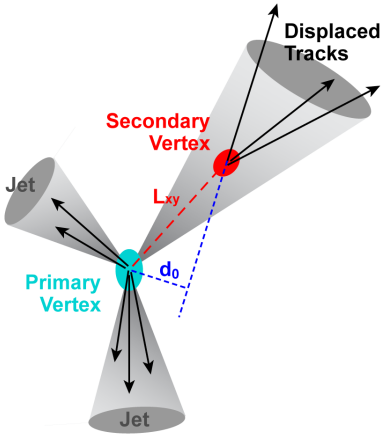



Figure 3.12: Illustration of an event with a *b*-jet, which shows the *b* hadron formed from the *b* quark decays at a secondary vertex at a distance $L_{xy}$ from the primary vertex [77].

tagging *b*-jets from $t\bar{t}$ process is found to be the best choice, i.e. fitting working point for the analysis. It has rejection factors of 13, 56 and 380 against *c*-jets, hadronic $\tau$ decays and jets from light quarks or gluons, respectively. In order to compensate for differences between data and simulation in the *b*-tagging efficiency for *b*-, *c*- and light-quark jets correction factors are applied to the simulated events.

### 3.4.4 Hadronically decaying $\tau$ leptons

$\tau$ leptons decay into hadrons ($\tau \to$ hadrons, denoted as $\tau_{\text{had}}$) in about 65% of the cases, or into leptons ($\tau^{\pm} \to \ell^{\pm}\nu_{\tau}\bar{\nu}_{\ell}$, where $\ell = e, \mu$) which final state is not considered in this thesis. Having a proper decay length of 87 $\mu$m, $\tau$ leptons usually decay before reaching any active material in the ATLAS detector and therefore they can only be identified via their decay products. Most of the time the hadronic decay products are either one or three charged pions or charged kaons with one or more additional neutral pion. Hence, decay depending on the number of charged pions tracks is called 1- or 3-prong, respectively, where all visible decay products are denoted as $\tau_{\text{had-vis}}$. Decays with more than three charged particles happen in less then in 1% of cases and are not relevant for the presented analysis.

The reconstruction of $\tau_{\text{had-vis}}$ objects starts from the anti-$k_T$ jets with a transverse energy $E_T > 10$ GeV. Then the associated tracks reconstructed in the inner detector, with $p_T > 1$ GeV and within a cone of $\Delta R = 0.2$ around the axis of the $\tau_{\text{had-vis}}$ candidate are found. The $\tau_{\text{had}}$ energy is obtained by the tau-specific calibration scheme [78, 79].

Hadronically decaying $\tau$ leptons leave traces in the detector similar to jets. Fortunately, the shower profile of those decays is narrower than that of the quark- or gluon-initiated jets. Moreover, to distinguish these objects also information about the number of associated tracks and the vertex of the $\tau$ lepton decay is used. These properties are used in a boosted decision tree algorithm in order to distinguish $\tau_{\text{had-vis}}$ candidates from the quark- or gluon-initiated jets, separately for one and three charged-particle tracks [79, 80].

The reconstruction efficiency is defined as the fraction of 1-prong (3-prong) hadronic $\tau$ decays which are reconstructed as 1-track (3-track) $\tau_{\text{had-vis}}$ candidates. The identification efficiency is defined the same way, with additional condition of fulfilling the BDT selection criteria. The total efficiency is the product of the reconstruction and of the identification efficiencies [79]. In the analysis presented here, the "*medium*" identification efficiency value is used, i.e. a working point with reconstruction and identification efficiency of 55% (40%) for 1-prong (3-prong) hadronic $\tau$ decays in $Z \to \tau\tau$ events is chosen, corresponding to the rejection factors of about 50 and 100 in multi-jet events, for 1- and 3-track $\tau_{\text{had-vis}}$ candidates, respectively (see Fig. 3.13). An additional likelihood-based requirement is used to reduce the amount of electrons misidentified as $\tau_{\text{had-vis}}$ candidates, providing a flat 95% efficiency in $\tau_{\text{had-vis}}$ $p_T$ and $\eta$, as well as the background rejection factor within the range $20 - 200$ depending on the $\eta$. To account for differences of efficiencies in data and simulation correction factors are derived and applied to simulated events.

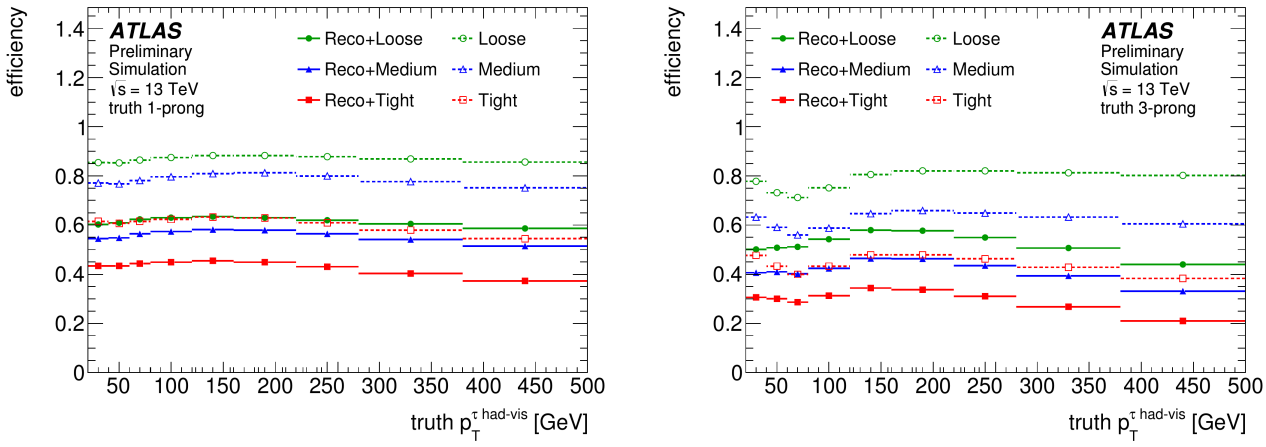



Figure 3.13: Efficiency for $\tau_{\text{had-vis}}$ identification (open symbols) and combined reconstruction and identification efficiency (full symbols) as a function of the $\tau_{\text{had-vis}}$ $p_T$ , for 1-track (*left*) and 3-track (*right*) $\tau_{\text{had-vis}}$ candidates [79].

### 3.4.5 Electrons

Electrons are reconstructed by matching the clustered energy deposits in the electromagnetic calorimeter to tracks reconstructed in the inner detector [81]. In order to improve efficiency of identyfing electrons while rejecting background electrons three identyfication criteria: *loose, medium* and *tight* are defined via likelihoods. The backgrounds arise from hadronic jets mistaken for electrons, electrons from photons conversions, $\pi^0$ Dalitz decays and from semileptonic heavy-flavour decays. They are based on the calorimetric cluster shapes, tracks and track-to-cluster matching variables. For *loose* identification criteria the efficiency is $\sim 80\%$ [82]. In addition to the likelihood identification some cuts applied on the transverse energy, $E_T$, $|\eta|$, and track and calorimeter based isolation must be satisfied. The track isolation criteria from the lepton selection requires that the scalar sum of the track momenta within the cone of radius $\Delta R = 0.2$ that are not associated with the lepton is restricted to a chosen fraction. This is done similary in the calorimeter considering instead the energy deposits within a cone of radius $\Delta R = 0.3$ that are not associated with the lepton being below a chosen threshold. The isolation cuts are used to reduce the number of jets that are misidentified as leptons. Corrections are applied to the energy resolution, reconstruction and identification efficiencies and calorimeter isolation to account for mismodelling in MC simulations.

### 3.4.6 Muons

It is worth mentioning that muons are quite easily distinguishable from other particles, leaving minimal deposits in the calorimeters (causing a clear signature in the detectors) and traverse the

muon spectrometer, thus a very high and pure reconstruction efficiency can be achieved.

They are identified in the muon spectrometer and for $|\eta| < 2.5$ matched to the ID tracks. Hits in the MS are used to reconstruct track segments, which are then combined between multiple layers of the MS to identify the tracks. In next step they are extrapolated to the primary vertex, matched to tracks in the ID and then combined into a single track. In order to suppress mainly muons from hadronic decays isolation criteria are imposed.

Note that *loose*, *medium* and *tight* categories, *working points*, are also used to identify muons [83].

### 3.4.7 Missing Transverse Energy

In proton-proton collisions the exact momentum of the initial partons is unknown, however an approximation that the partons carry no transverse momentum is made. Hence, the transverse momentum in the final state should also be zero due to the energy and momentum conservation. Using this rule in the plane transverse to the beam axis, the missing transverse energy denoted as $\mathbf{E}_T^{miss}$ in the $x-y$ plane is reconstructed from the vector sum of the transverse momenta of reconstructed and fully-calibrated objects, with an additional term that is calculated using the inner-detector tracks not associated to any of the selected objects and matched to the primary vertex in order to make it more resilient to pile-up [84]. In ideal situation $\mathbf{E}_T^{miss}$ arises from weakly-interacting, stable particles produced in the collision, which in the SM are the neutrinos. However, in experimental reality it comes also from mis-measurements in the calorimeters. If New Physcis scenarios are realised by the Nature, large $E_T^{miss}$ can be an indication of weakly interacting exotic particles.

The $\mathbf{E}_T^{miss}$ calculation is based on the energy deposited in the calorimeters and muons from MS. As was mentioned above for the reconstruction of the missing transverse energy first the vectorial quantity $\mathbf{E}_T^{miss}$ is calculated using the reconstructed and calibrated physics objects

$$\mathbf{E}_T^{miss} = \mathbf{E}_T^{miss,e} + \mathbf{E}_T^{miss,\gamma} + \mathbf{E}_T^{miss,\tau} + \mathbf{E}_T^{miss,jet} + \mathbf{E}_T^{miss,soft} + \mathbf{E}_T^{miss,\mu}, \tag{3.8}$$

with the missing transverse energy $\mathbf{E}_T^{\text{miss,species}} = -\sum \mathbf{p}_T^{\text{species}}$ for each species of object, i.e: *e:* electrons; $\gamma$: photons; $\tau$: $\tau$ leptons; *jet*: jets; *soft*: soft objects; $\mu$: muons. Note that in presented analysis the scalar missing transverse energy $E_T^{miss}$ is used,

$$E_T^{miss} = |\mathbf{E}_T^{miss}| = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2}. \tag{3.9}$$

The contributions of $E_T^{miss,soft}$ originate from ID tracks associated with the primary vertex of the hard interaction, which are not used in the reconstruction of the other, high $p_T$ objects as well as the reconstruction of photons, which are needed to calculate $E_T^{miss,\gamma}$ are described in [84, 85].

# Chapter 4

# An Outline of the Analysis

> "Hallo, Rabbit, isn't that you? "No," said Rabbit, (...). But isn't that Rabbit's voice?
> "I don't think so," said Rabbit. "It isn't meant to be. "Oh!" said Pooh." A. A. Milne.

In the analysis presented in this thesis the results of a search for charged Higgs boson, $H^{\pm}$, using 36.1 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector are described. The charged Higgs boson is searched for in topologies in which it is produced in association with a top-quark. The fully hadronic final state is considered, i.e. $\tau$ lepton and $W$ boson (from $t \to bW$) decay hadronically ($\tau_{\text{had-vis}}$+jets channel).

The most significant background contributions to the search come from multi-jets, $t\bar{t}$ or single top quark and $W$+jets events. Less important contributions arise from $Z/\gamma^*$+jets and diboson production. In the analysis backgrounds are categorised based on the type of the reconstructed object that is identified as the $\tau_{\text{had-vis}}$ candidate, not according to their production mode. Hence, there are two types of events: those with true $\tau_{\text{had}}$ and those with another object identified as the $\tau_{\text{had-vis}}$ candidate. Background processes with true $\tau$ lepton decaying hadronically selected and identified as a $\tau_{\text{had-vis}}$ candidate are estimated from simulation. Backgrounds arising from events in which an isolated electron or muon is misidentified as a $\tau_{\text{had-vis}}$ candidate ($l \to \tau$), are also estimated with simulation, with the exception that a correction factor is applied for events where an electron is misidentified as a $\tau_{\text{had-vis}}$ candidate in order to account for the actual misidentification rate [80]. Events where a jet is misidentified as a $\tau_{\text{had-vis}}$ candidate, the so-called jet $\to \tau$ backgrounds, are estimated using data-driven fake-factor method (see Section 4.3).

The presented analysis uses events passing the $E_T^{miss}$ trigger with a threshold at 70 GeV (`HLT_xe70_tc_lcw`), 90 GeV (`HLT_xe90_mht_L1XE50`) or 110 GeV (`HLT_xe110_mht_L1XE50`), depending on the data-taking period (see Section 4.2 and Chapter 3). The efficiency of such triggers is measured in data and used to reweight the simulated events, with the same method as in Ref. [86]. At least one vertex with two or more associated tracks with $p_T > 40$ MeV is

required and events for which any jet with $p_T > 25$ GeV fails the appriopriate quality cuts are discarded. This ensures that there are no jet-like signals due to instrumental effects, like noise in EM calorimeter, or non-collision backgrounds.

In previous search [38,39,43,86–88] for charged Higgs bosons produced in association with a top-quark and decaying via $H^\pm \to \tau^\pm \nu$ channel the transverse mass, $m_T$, of the highest-$p_T$ $\tau_{\text{had-vis}}$ candidate and $E_T^{miss}$ was used as a final discriminating variable between the signal and background in the cut based analysis:

$$m_T = \sqrt{2 p_T^\tau E_T^{miss} (1 - \cos \Delta\phi_{\tau_{\text{had-vis}}, E_T^{miss}})}, \tag{4.1}$$

where $\Delta\phi_{\tau_{\text{had-vis}}, E_T^{miss}}$ is the azimuthal angle between $\tau_{\text{had-vis}}$ candidate and the direction of the missing transverse energy, and $p_T^\tau$ is the transverse momentum of $\tau_{\text{had-vis}}$. This formula corresponds to $m_T$ of the $W$ boson in SM top quark decays ($t \to bW^\pm \to b\tau^\pm\nu$) as well as to the $m_T$ of the $H^\pm$ in the signal events ($t \to bH^\pm \to b\tau^\pm\nu$).

As an improvement to the former searches a multivariate analysis has been applied and used in the current analysis to separate the $H^\pm$ signal from the SM background processes, where the output score of BDT is used as the final discrimination variable in statistical analysis of the results (see Chapter 5.2).

Note that in parallel to the search of the $H^\pm$ decays to $\tau^\pm\nu$ in $\tau_{\text{had-vis}}$+jets channel being a subject of the presented thesis, the $\tau_{\text{had-vis}}$+lepton channel where top quark decaying semi-leptonically was also carried out. This channel improves the analysis sensitivity at low $H^\pm$ masses. The final conclusions presented in the Chapter 7 combine the results of analysis of both channels.

## 4.1 Event Selection

### 4.1.1 Simulation Samples

Simulated events of $H^\pm$ signal are generated in three different, separate mass regions. In the mass range between $90 - 160$ GeV, i.e. below the top-quark mass, $t\bar{t}$ events with one top-quark decaying into a charged Higgs boson and a $b$-quark are generated at the leading order (LO) with MADGRAPH5 [89]. Both $t\bar{t}$ events with two $\bar{t} \to \bar{b}H^\pm$ decays and single-top-quark events with a subsequent decay $\bar{t} \to \bar{b}H^\pm$ have a negligible contribution and are not simulated. In the so-called intermediate-mass region, $160 - 180$ GeV, LO non resonant, single-top-quark resonant and double-top-quark resonant processes with a $W$ boson, a charged Higgs boson and

two $b$-quarks in the final state are generated in the 4FS with MADGRAPH5. In the high mass range, i.e. above the top-quark mass, $200-2000$ GeV, simulated events of $H^{\pm}$ production in association with top-quark are generated in the 4FS at next-to-leading order (NLO) with MADGRAPH5+AMC@NLO [90, 91]. For all the cases above, the NNPDF2.3 LO [92] parton distribution function sets are used and for the underlying event [93] the parton-level generator is interfaced to PYTHIA V8.186 with A14 tune [94].

The background processes of the SM include production of $t\bar{t}$ pairs, single top-quarks, $W$+jets, $Z/\gamma^{*}$+jets and electroweak gauge boson pairs ($WW/WZ/ZZ$), as well as multi-jet events. The $t\bar{t}$ events constitute the main background in the low- and intermediate-mass $H^{\pm}$ search, while multi-jet events dominate for large charged Higgs boson masses. The backgrounds are categorised based on the type of generator-level objects reconstructed as a $\tau_{\text{had-vis}}$ candidate. Only simulated events with true $\tau$ lepton decaying hadronically or with an electron or muon misidentified as $\tau_{\text{had-vis}}$ are kept. Backgrounds arising from a quark- or gluon-initiated jets misidentified as $\tau_{\text{had-vis}}$ are estimated using the data-driven method.

For the generation of $t\bar{t}$ and single top-quarks in the $Wt$- and $s$-channels, the POWHEG-BOX V2 [95–97] generator, with the CT10 [98] PDF set in the matrix-element calculations, is used. The single-top-quark events in electroweak $t$-channel are generated using the POWHEG-BOX V1 generator. This generator uses the 4FS for the NLO matrix-element calculations together with the fixed 4-flavor PDF set CT10f4 [99]. For this process the top-quarks are decayed with preservation of all spin correlations by using MADSPIN [100]. For all processes, the parton shower, fragmentation, and the underlying event are simulated using PYTHIA V6.428 [101] with the CTEQ6L1 [102] PDF set and the corresponding Perugia 2012 tune [103]. The top-quark mass is set to $m_{top} = 172.5$ GeV for all the relevant signal and background simulation samples. The sample of $t\bar{t}$ events is normalised to the next-to-next-to-leading order (NNLO) cross-section, including soft-gluon resummation to next-to-next-to-leading-log (NNLL) order (for more information see Ref. [104] and references therein). The normalisation of single top-quark events sample uses an approximate calculation at NLO in QCD for the $s$- and $t$-channels [105, 106] and NLO+NNLL calculation for the $Wt$-channel [107].

Events containing $W$ or $Z$ boson with associated jets are simulated with SHERPA V2.2.1 [108] together with the NNPDF3.0 NNLO PDF set [109]. The $W/Z$+jets events are normalised to the NNLO cross-sections calculated using FEWZ [110–112]. Diboson processes ($WW$, $WZ$ and $ZZ$) are simulated using the POWHEG-BOX V2 generator, interfaced to the PYTHIA V8.186 parton shower model. For the hard-scatter process the CT10nlo PDF set is used, while for the parton shower the CTEQL1 PDF set is used. The non-perturbative effects are modelled using the AZNLO [113] tune. The NLO generator cross-sections are used in this case.

In order to take into account the photon radiation from charged leptons PHOTOS++ V3.52 [114] together with PYTHIA6 is used. For proper simulation of the bottom- and charm-hadron decays EVTGEN V1.2.0 [115] is employed. Finally, all simulated events are overlaid with additional minimum-bias events generated with PYTHIA V8.186 using the A2 [116] tune and the MSTW2008LO PDF set [117] to simulate the effect of pile-up. Simulated events are then weighted to the same number of collisions per bunch crossing as the data.

The full list of simulated SM backgrounds with their cross-sections and names of generators used is presented in Table 4.1.

| Background process | Generator & parton shower | Cross-section (in pb) |
|---|---|---|
| $t\bar{t}$ with at least one lepton $\ell$ | POWHEG & PYTHIA6 | 451.66 |
| Single top-quark $t$-channel | POWHEG & PYTHIA6 | 70.43* |
| Single top-quark $s$-channel | | 3.35* |
| Single top-quark $Wt$-channel | | 71.67 |
| $W(\ell\nu)+$jets | SHERPA | $2.0\times10^4$ |
| $Z/\gamma^*(\ell\ell)+$jets | SHERPA | $2.1\times10^3$ |
| $WW$ | POWHEG & PYTHIA8 | 54.81 |
| $WZ$ | | 16.30 |
| $ZZ$ | | 8.95 |

Table 4.1: MC generators and cross sections for the main SM background samples at $\sqrt{s} = 13$ TeV. Here, $\ell$ refers to the three lepton families $e$, $\mu$ and $\tau$. All background cross sections are normalised to NNLO predictions, except for diboson events, where the NLO prediction is used. A '*' indicates that the quoted cross section for the sample is without leptonic/hadronic branching ratios. From [33].

### 4.1.2 Event Preselection

The following preselection is done on reconstructed physics objects:

- the $\tau_{\text{had-vis}}$ candidates are required to have the transverse momentum, $p_T^\tau > 40$ GeV and to

be within $|\eta| < 2.3$ (with $1.37 < |\eta| < 1.52$, i.e. the transition region between the barrel and end-cap calorimeters excluded), and have one or three tracks. Moreover, the *medium* identification efficiency working point is used.

- *Loose* likelihood-based identification selection requirement has to be met by electron candidates. The transverse energy $E_T = E_{clus}/cosh(\eta_{track})$, computed using the calorimeter cluster energy $E_{clus}$ and the direction of the electron track $\eta_{track}$, is required to be greater than 20 GeV. The pseudorapidity range for the electromagnetic cluster covers the fiducial volume of the detector, i.e. $|\eta| < 2.47$ with ($1.37 < |\eta| < 1.52$ is excluded). Additionally, $E_T$ and $\eta-$dependent calorimeter isolation requirements are imposed with a *Loose* selection criteria, where both the identification and isolation working points are determined by standard Egamma Working Group tools [81];
- only the jets with $p_T > 25$ GeV and within $|\eta| < 2.5$ are used;
- muons with $p_T > 20$ GeV, $|\eta| < 2.5$ and *Loose* identification criteria are used.

### 4.1.3 Removal of geometric overlaps between objects

When several objects overlap geometrically, the following procedure is applied. First, a $\tau_{\text{had-vis}}$ object is removed if found within $\Delta R < 0.2$ of either an electron or a muon with identification criteria looser than the nominal ones and with transverse momentum above 20 GeV or 7 GeV, respectively. Then, any electron sharing an inner-detector track with a muon is discarded. Next, electrons and muons are removed if found within $\Delta R < 0.4$ of a $b$-tagged jet. Finally, the jets are discarded if they are within $\Delta R < 0.2$ of the highest-$p_T$ $\tau_{\text{had-vis}}$ candidate or the remaining electrons and muons.

### 4.1.4 Final event selection

The event selection, following *pre-selection* and *removal of geometric overlaps between objects* described above, is optimised for the following production processes and final states:

- for the search of $H^{\pm}$ in the $t\bar{t}$ production channel followed by the top quark decay, which is only open for low charged Higgs masses:

$$t\bar{t} \to (W^{\mp}\bar{b})(bH^{\pm}) \to (q\bar{q}'\bar{b})(b\tau^{\pm}_{had}\nu),$$

- for the search of $H^{\pm}$ in the top quark associated production channel, which is open for whole charged Higgs mass range:

$$g\bar{b} \to \bar{t}H^{+} \to (W^{-}\bar{b})H^{+} \to (q\bar{q}'\bar{b})(\tau^{\pm}_{had}\nu)$$

$$gb \to tH^- \to (W^+b)H^- \to (q\bar{q}'b)(\tau^{\mp}_{had}\nu)$$

in the 5FS and

$$gg \to \bar{t}bH^{\pm} \to (W^-\bar{b})bH^{\pm} \to (q\bar{q}'\bar{b})b(\tau^{\pm}_{had}\nu)$$

in the 4FS case where the final state is the same as for low-mass search.

The signal region (SR), i.e. the final event selection, is defined by the following requirements:

- at least one $\tau_{\text{had-vis}}$ candidate with $p_T^{\tau} > 40$ GeV;
- no electron or muon with $E_T$ or $p_T$ above 20 GeV, respectively;
- at least three jets with $p_T > 25$ GeV, of which at least one is $b$-tagged using MV2c10 algorithm with working point corresponding to b-tagging efficiency $\sim 70\%$;
- $E_T^{miss} > 150$ GeV: this cut ensures compatibility with the used trigger selection;
- $m_T > 50$ GeV: this requirement is used in order to reject events with wrongly determined $E_T^{miss}$, i.e. where $\tau_{\text{had-vis}}$ is nearly aligned with the direction of the missing transverse energy.

Table 4.2 shows the expected numbers of events in the SR originating from different backgrounds, together with an expectation for the signal with $m_{H^{\pm}} = 200$ GeV and $m_{H^{\pm}} = 1000$ GeV and the number of events in data. The predicted distributions of kinematic variables in the SR are shown in Figures 4.1 and 4.2. In all plots, the $j \to \tau$ background includes all processes in which the selected $\tau_{\text{had-vis}}$ candidate is from a quark- or gluon-initiated jet, while the $l \to \tau$ background includes all processes in which a lepton (electron or muon) is reconstructed and identified as the $\tau_{\text{had-vis}}$ object. All other backgrounds correspond to events where the $\tau_{\text{had-vis}}$ object matches a hadronic $\tau$ decay at the generator level. The latter two backgrounds are derived from simulation, while the former is estimated from data using the fake-factor (FF) method (see Section 4.3).

In order to probe the modeling of the $t\bar{t}$ background in a multi-jet environment an additional signal-depleted control region (CR) is defined. The $t\bar{t}$ enriched CR has the same event selection as the SR, except for $m_T < 100$ GeV and requirement that at least two jets must be $b$-tagged. It contains $\sim$90% of top backgrounds, i.e. both $t\bar{t}$ and single-top-quark events. Note that this CR is not fully disjoint with the SR, however the signal contamination is expected to be very small (about 70 events for the $H^{\pm}$ signal at 200 GeV, based on the cross section predicted at $\tan\beta = 40$ in the hMSSM benchmark scenario).

Figures 4.3-4.4 show predicted and measured distributions of kinematic variables for the electroweak, top and multi-jet backgrounds in the $t\bar{t}$ enriched CR. In Fig. 4.3 are presented: the number of reconstructed jets, the number of $b$-tagged jets, the $\tau_{\text{had-vis}}$ candidate $p_T$, the $\tau_{\text{had-vis}}$ candidate $\eta$, $E_T^{miss}$ and the transverse mass of the $\tau_{\text{had-vis}}$ candidate, and $E_T^{miss}$. Whereas, in Fig 4.4 the $p_T$ of the hardest $b$-jet, the $\Delta R$ between the $\tau_{\text{had-vis}}$ candidate and the hardest $b$-jet, the difference in azimuthal angle between the $\tau_{\text{had-vis}}$ candidate and $E_T^{miss}$, the difference in azimuthal angle between the hardest $b$-jet and $E_T^{miss}$, $\tau$ polarisation variable $\Upsilon$ are shown. For both Figures the $j \to \tau$ background is estimated using the FF method. As can be seen in all presented distributions a good agreement of predicted backgrounds with data is observed.

| Selection | top | $W \to \tau\nu$ | $Z \to \tau\tau$ | Diboson | $l \to \tau$ |
|---|---|---|---|---|---|
| Trigger | 134368 | 140456 | 22593 | 4710 | 10205 |
| Medium $\tau_{\text{had-vis}}$, $p_{\text{T}}^{\tau} > 40$ GeV | 67797 | 73912 | 12371 | 2703 | 3109 |
| Lepton veto | 60680 | 73895 | 11318 | 2492 | 2879 |
| $E_{\text{T}}^{\text{miss}} > 150$ GeV | $26739 \pm 6315$ | $33127 \pm 4963$ | $3395 \pm 495$ | $1374 \pm 164$ | $841 \pm 117$ |
| $\geq 1$ b-jet | 21511 | 3753 | 485 | 195 | 405 |
| $m_{\text{T}} > 50$ GeV | $7660 \pm 1747$ | $1047 \pm 188$ | $84 \pm 50$ | $63 \pm 7$ | $265 \pm 35$ |

| Selection | $j \to \tau$ | Total bkg | $H^{\pm}(200)$ | $H^{\pm}(1000)$ | data ($36.1\,\text{fb}^{-1}$) |
|---|---|---|---|---|---|
| Trigger | 2016098 | 2328431 | 2268 | 9676 | 2404104 |
| Medium $\tau_{\text{had-vis}}$, $p_{\text{T}}^{\tau} > 40$ GeV | 145716 | 305608 | 1455 | 6756 | 256668 |
| Lepton veto | 139352 | 290616 | 1353 | 6085 | 240067 |
| $E_{\text{T}}^{\text{miss}} > 150$ GeV | $7033 \pm 1013$ | $72509 \pm 8113$ | $680 \pm 84$ | $5759 \pm 410$ | 67673 |
| $\geq 1$ b-jet | 2884 | 29234 | 499 | 4257 | 27461 |
| $m_{\text{T}} > 50$ GeV | $2372 \pm 257$ | $11491 \pm 1777$ | $427 \pm 55$ | $4220 \pm 319$ | 11021 |

Table 4.2: Expected event yields after cumulative selection cuts and comparison with $36.1\,\text{fb}^{-1}$ of data for $\tau_{\text{had-vis}}$+jets channel. The values shown for the signal correspond to $\sigma(pp \to [b]tH^{\pm}) \times \text{Br}(H^{\pm} \to \tau\nu) = 1$ pb. Both the statistical and systematic uncertainties are shown for the final selection cut in the last row. From [33].

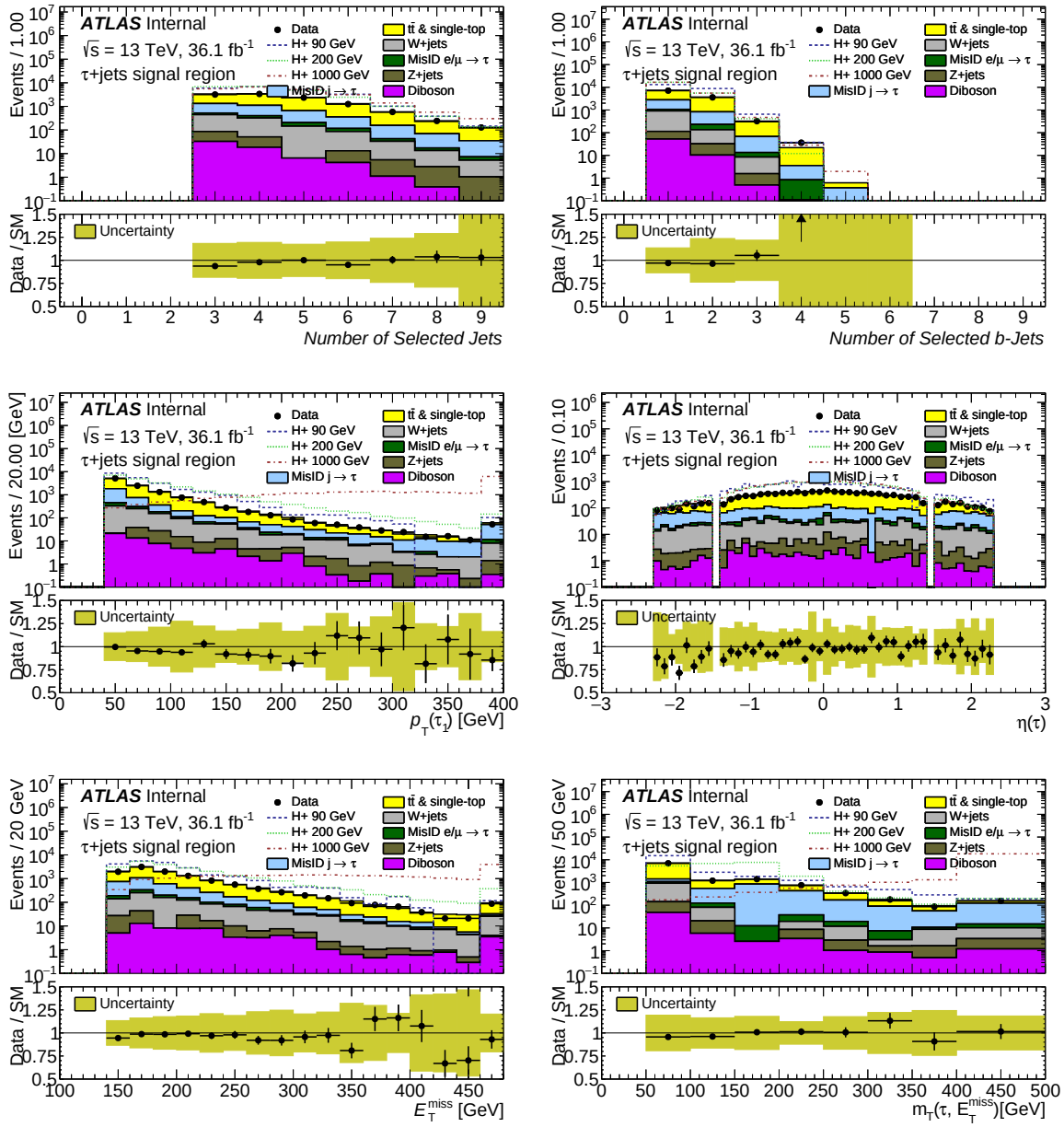



Figure 4.1: Predicted total background distributions along with a few simulated signal samples (with $m_{H^\pm} = 90, 200, 1000$ GeV), after full event selection. Shown are (top left) the number of reconstructed jets, (top right) the number of $b$-tagged jets, (middle left) the $\tau_{\text{had-vis}}$ candidate $p_T$, (middle right) the $\tau_{\text{had-vis}}$ candidate $\eta$, (bottom left) $E_T^{miss}$ and (bottom right) the transverse mass of the $\tau_{\text{had-vis}}$ candidate and $E_T^{miss}$. The background is stacked, while the signal is overlaid. When plotting the transverse momenta and mass, the last bin contains overflow. The $j \to \tau$ background is estimated using the FF method. The uncertainty band in the ratio plots includes both statistical and systematic uncertainties on the background prediction. The signal distributions are all scaled to the integral of the total background. From [33].

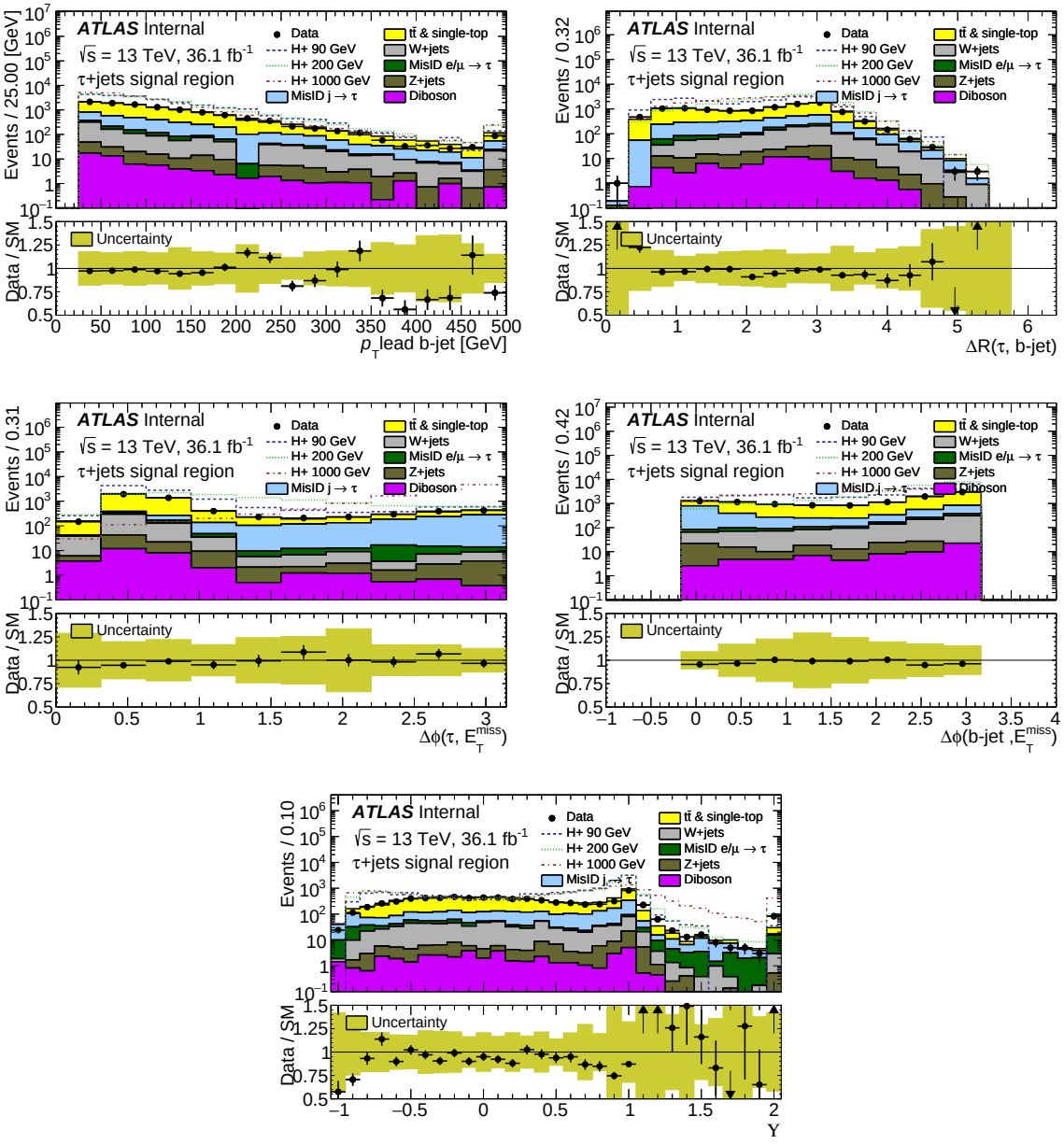



Figure 4.2: Predicted total background distributions along with a few simulated signal samples (with $m_{H^\pm} = 90, 200, 1000$ GeV), after full event selection. Shown are (top left) the $p_T$ of the hardest $b$-jet, (top right) the $\Delta R$ between the $\tau_{\text{had-vis}}$ candidate and the hardest $b$-jet, (middle left) the difference in azimuthal angle between the $\tau_{\text{had-vis}}$ candidate and $E_T^{miss}$, (middle right) the difference in azimuthal angle between the hardest $b$-jet and $E_T^{miss}$, (bottom) $\tau$ polarisation variable $\Upsilon$. The background is stacked, while the signal is overlaid. When plotting the $b$-jet transverse momentum, the last bin contains overflow. The $j \to \tau$ background is estimated using the FF method. The uncertainty band in the ratio plots includes both statistical and systematic uncertainties on the background prediction. The signal distributions are all scaled to the integral of the total background. From [33].

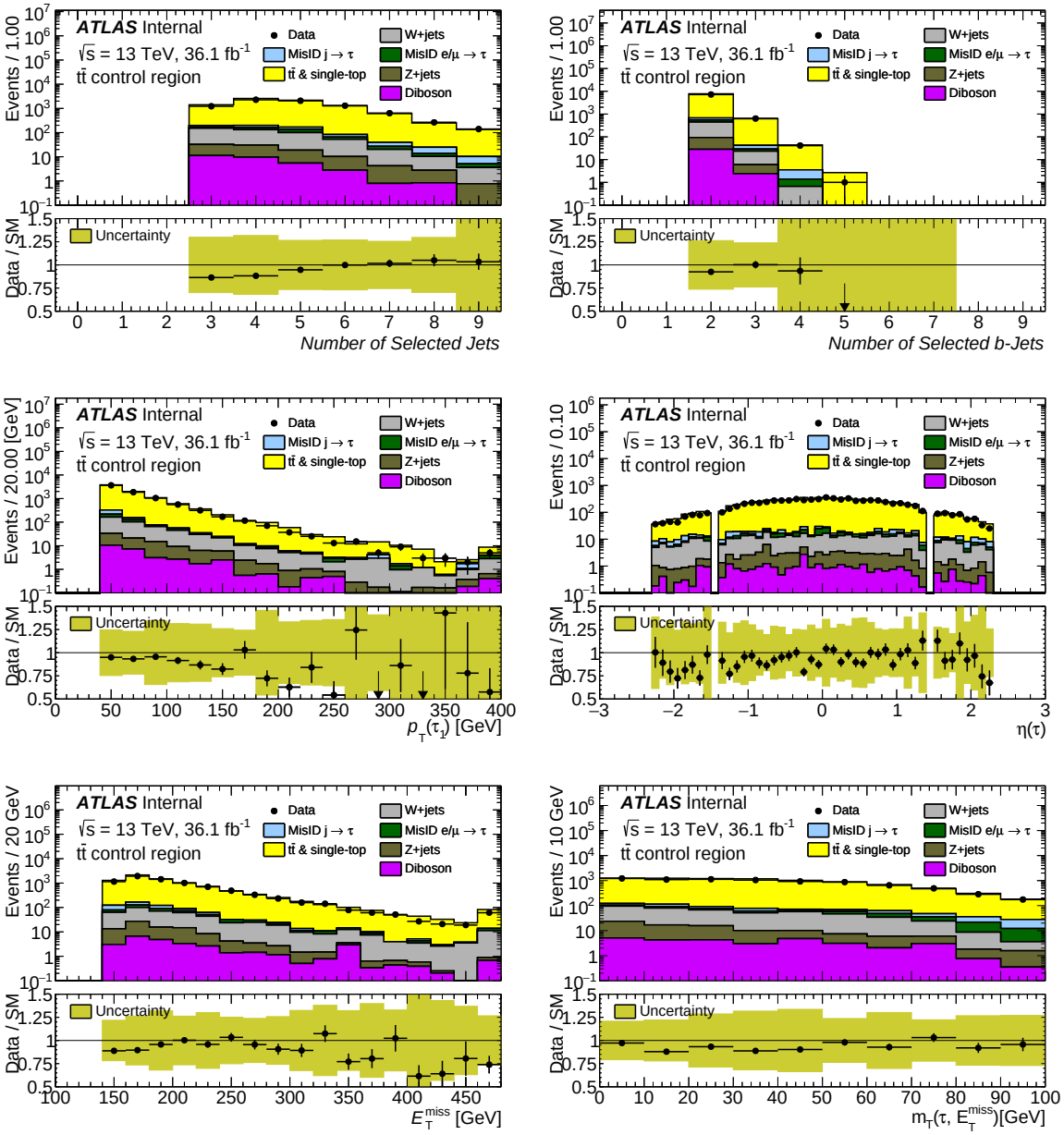



Figure 4.3: Predicted and measured distribution for the electroweak, top and multi-jet backgrounds in the $t\bar{t}$ enriched control region. Shown are (top left) the number of reconstructed jets, (top right) the number of $b$-tagged jets, (middle left) the $\tau_{\text{had-vis}}$ candidate $p_T$, (middle right) the $\tau_{\text{had-vis}}$ candidate $\eta$, (bottom left) $E_T^{miss}$ and (bottom right) the transverse mass of the $\tau_{\text{had-vis}}$ candidate and $E_T^{miss}$. When plotting the transverse momenta and mass, the last bin contains overflow. The $j \to \tau$ background is estimated using the FF method. The uncertainty band in the ratio plots includes both statistical and systematic uncertainties on the background prediction. From [33].

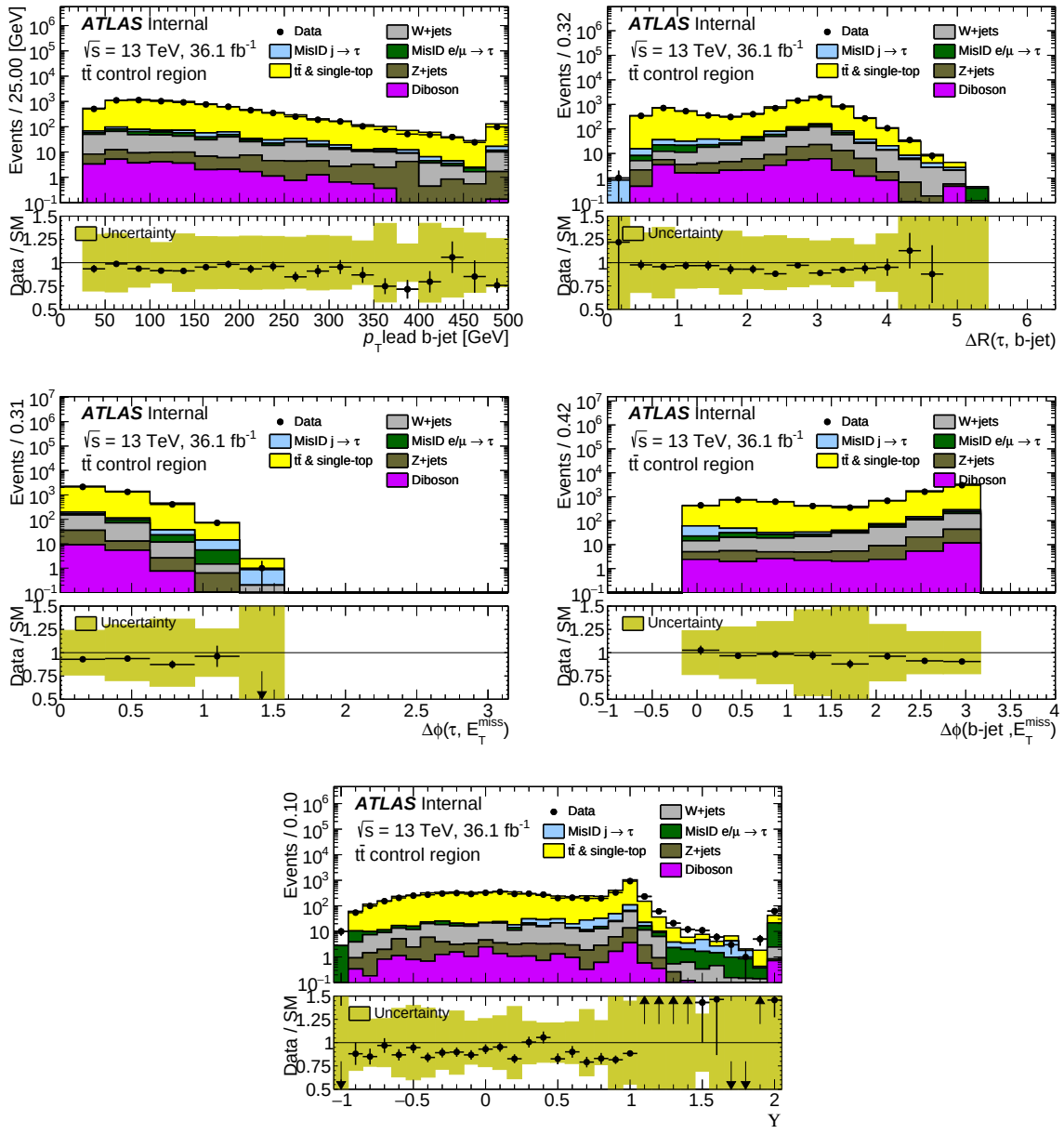



Figure 4.4: Predicted and measured distribution for the electroweak, top and multi-jet backgrounds in the $t\bar{t}$ enriched control region. Shown are (top left) the $p_T$ of the hardest $b$-jet, (top right) the $\Delta R$ between the $\tau_{\text{had-vis}}$ candidate and the hardest $b$-jet, (middle left) the difference in azimuthal angle between the $\tau_{\text{had-vis}}$ candidate and $E_T^{miss}$, (middle right) the difference in azimuthal angle between the hardest $b$-jet and $E_T^{miss}$, (bottom) $\tau$ polarisation variable $\Upsilon$. When plotting the $b$-jet transverse momentum, the last bin contains overflow. The $j \to \tau$ background is estimated using the FF method. The uncertainty band in the ratio plots includes both statistical and systematic uncertainties on the background prediction. From [33].

## 4.2 Trigger efficiency measurement

The $E_T^{miss}$ trigger is not well described in simulation. The strategy for the treatment of the $E_T^{miss}$ trigger in simulation is to derive the trigger efficiency from data in bins of the $E_T^{miss}$ values. The binned $E_T^{miss}$-dependent efficiency is transformed into a continuous efficiency by fitting it with the error function. This is done to remove the bias caused by the binning. Simulated events are weighted using the efficiency curve, based on the $E_T^{miss}$ in the event.

For the efficiency measurement of triggers with $E_T^{miss} > 70$ GeV threshold, the full 2015 data sample is used, with an integrated luminosity of 3.2 fb$^{-1}$ (more detailes are given in Chapter 6). For the triggers with $E_T^{miss} > 90$ GeV threshold efficiency measurement, 6.11 fb$^{-1}$ of 2016 data is used and, for the triggers with $E_T^{miss} > 110$ GeV threshold efficiency measurement, 26.75 fb$^{-1}$ of 2016 data is used. The event selection applied in the CR to measure the trigger efficiency ($e$+$\tau_{\text{had-vis}}$ selection) is disjoint with that in the signal region, while retaining as many similarities as possible:

- exactly one selected electron with *loose* identification and $p_T > 26$ GeV;
- at least one selected $\tau_{\text{had-vis}}$ object with *loose* identification and $p_T > 30$ GeV;
- at least two selected jets, with $p_T > 25$ GeV;
- at least one of them $b$-tagged corresponding to b-tagging efficiency $\sim 70\%$.

The given trigger efficiency is defined as a ratio of events fulfilling event selection described above and trigger requirement to events passing event selection. It is derived from data and fitted with the error function applying the following parameterisation:

$$F(x) = p_0 \cdot \left[1 + \text{erf}\left(\frac{\text{x} - \text{p}_1}{\text{p}_2}\right)\right] + p_3, \tag{4.2}$$

where $p_0, p_1, p_2, p_3$ are the fit parameters. Different choices of the parameters and binnings are examined to find the optimal fit. The fit result obtained after this optimisation is shown in Fig. 4.5 for `HLT_xe90_mht_L1XE50` and `HLT_xe110_mht_L1XE50` triggers.

The measured efficiency of the $E_T^{miss}$ trigger shows a small dependence on the identification criteria (*loose* or nominal) of the electron and $\tau_{\text{had-vis}}$ candidates, as well as on the minimum number of jets used in the definition of the CR. The corresponding small variations of the fitted function of the $E_T^{miss}$ trigger efficiency, together with the statistical uncertainty, are accounted for as systematic uncertainties, which amount to an uncertainty of 1.4% on the event yield in the SR. The stability of the fit is probed by increasing the errors on the fitted points by a factor of 4

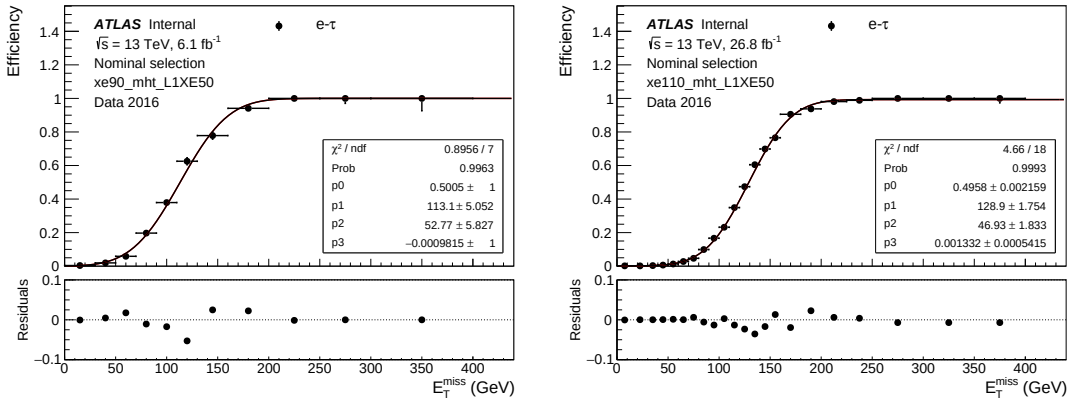



Figure 4.5: Trigger efficiency and corresponding fit with the $\tau_{\text{had-vis}}$ selection for `HLT_xe90_mht_L1XE50` (top left) and `HLT_xe110_mht_L1XE50` (top right) triggers. From [33].

and then redoing the fit. The difference with the nominal fit, present mainly for $E_T^{miss} < 100$ GeV ($\sim 10-15\%$), is used as a systematic uncertainty.

For $E_T^{miss} = 150$ GeV, the total systematic uncertainty on measured trigger efficiency is about 4.5%, whereas the effect on the yields in the whole SR ($E_T^{miss} > 150$ GeV) is 1.4%.

## 4.3 Estimation of Background Contributions

In this search, the dominant background processes are $t\bar{t}$, single top-quark, $W$+jets, $Z/\gamma^*$+jets, diboson and multi-jet events. They are categorised based on the object that gives rise to the identified $\tau_{\text{had-vis}}$ candidate. The contribution from background processes with true $\tau$ lepton decaying hadronically selected and identified as a $\tau_{\text{had-vis}}$ candidate are estimated from simulation. Backgrounds arising from events in which an isolated electron or muon is misidentified as a $\tau_{\text{had-vis}}$ candidate contribute at the level of 3% to the total background, with misidentified muons contributing about one order of magnitude less than the misidentified electrons. The contribution of these backgrounds are also estimated using simulation. However, a correction factor is applied to events where an electron is misidentified as a $\tau_{\text{had-vis}}$ candidate in order to account for the actual misidentification rate known from $Z \to e^+e^-$ events in data measured in function of $p_T$ and number of tracks [80].

### Data-driven fake-factor method

Contribution of background processes where a jet initiated by a quark or a gluon is reconstructed and selected as a $\tau_{\text{had-vis}}$ candidate are estimated with a data-driven technique called the fake-

factor method. For this purpose, an anti-$\tau_{\text{had-vis}}$ selection is defined by inverting the $\tau_{\text{had-vis}}$ identification criteria while maintaining a *loose* requirement on the $\tau_{\text{had-vis}}$ BDT output score to ensure that the fractions of gluon- and quark-initiated jets mimicking $\tau_{\text{had-vis}}$ candidates are similar in both the signal and the corresponding anti-$\tau_{\text{had-vis}}$ control regions. Afterwards, a fake factor (FF) is defined as the ratio between the number of misidentified $\tau_{\text{had-vis}}$ candidates (i.e. jets) fulfilling the nominal $\tau_{\text{had-vis}}$ selection or the anti-$\tau_{\text{had-vis}}$ selection:

$$\text{FF} = \frac{N^{\text{CR}}_{\tau_{\text{had-vis}}}}{N^{\text{CR}}_{\text{anti-}\tau_{\text{had-vis}}}}. \tag{4.3}$$

In the signal region, after subtracting the $\tau_{\text{had-vis}}$ candidates matching a true $\tau_{\text{had}}$ at the generator level but fulfilling the anti-$\tau_{\text{had-vis}}$ selection, the number of events with a misidentified $\tau_{\text{had-vis}}$ candidate ($N^{\tau_{\text{had-vis}}}_{\text{fakes}}$) is derived from the sub-set of anti-$\tau_{\text{had-vis}}$ candidates in the following way:

$$N^{\tau_{\text{had-vis}}}_{\text{fakes}} = \sum_{\text{i}} N_{\text{anti-}\tau_{\text{had-vis}}}(\text{i})\,\text{FF}(\text{i}), \tag{4.4}$$

where the index $i$ refers to a given bin in the parameterisation of the FF, in terms of $p^{\tau}_{T}$ and number of associated tracks (where 1- or 3-prong $\tau_{\text{had-vis}}$ candidates are considered).

In order to account for potentially different sources of misidentified $\tau_{\text{had-vis}}$ candidates in the SR and the corresponding anti-$\tau_{\text{had-vis}}$ CR, FFs are first computed in two regions of the data with different fractions of quark- and gluon-initiated jets, and then combined. A first CR, enriched in gluon-initiated jets (referred to as multi-jet CR), is defined by applying the same event selection as for the SR, but with

- a $b$-jet veto and
- $E^{miss}_{T} < 80$ GeV.

Events of this type are collected using a combination of multi-jet triggers instead of the $E^{miss}_{T}$ trigger. A second CR, enriched in quark-initiated jets (referred to as the $W$+jets CR), is defined by applying the following requirements:

- exactly one lepton (electron or muon) matched to the single-lepton trigger object, with $E_T$ or $p_T$ above 30 GeV, respectively;
- no $b$-tagged jets;
- no cut on $E^{miss}_{T}$;
- exactly one $\tau_{\text{had-vis}}$ candidate with $p^{\tau}_{T} > 30$ GeV and an opposite electric charge to that of selected lepton;
- 60 $GeV < m_T(\ell, E^{miss}_{T}) < 160$ GeV,

where the transverse mass of the lepton and the missing transverse momentum is obtained by replacing the $\tau_{\text{had-vis}}$ candidate by a lepton in Eq. (4.1). The FFs measured in these two control regions are shown in Fig. 4.6 (left), as a function of $p_T^\tau$.

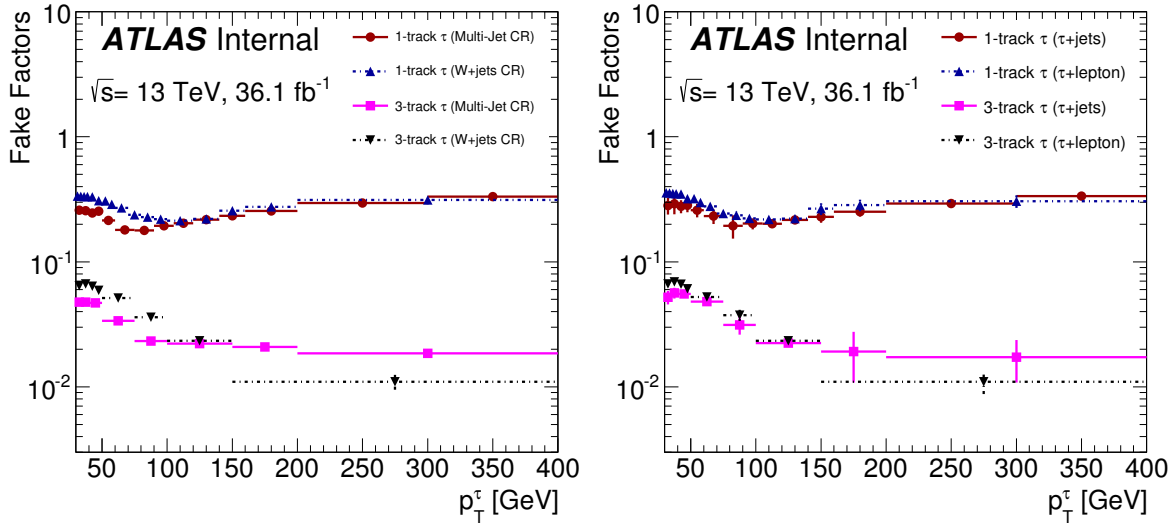



Figure 4.6: Fake factors parameterised as a function of $p_T^\tau$ and the number of charged $\tau$ decay products (two categories: 1-prong and 3-prong), as obtained in the multi-jet and $W$+jets CRs (left), as well as after reweighting by $\alpha_{\text{MJ}}$ (right). The errors shown come from the statistical uncertainty in a given $p_T^\tau$ bin (left) and with additional systematical uncertainties obtained from the combination in a given $p_T^\tau$ bin (right). Note that this thesis describes searching $H^\pm$ in $\tau_{\text{had-vis}}$+jets channel, whereas on the right plot also results for $\tau_{\text{had-vis}}$+lepton channel are shown. From [33].

In the anti-$\tau_{\text{had-vis}}$ regions corresponding to the nominal event selections, the fractions of quark- and gluon-initiated jets misidentified as $\tau_{\text{had-vis}}$ candidates are then measured using a template-fit approach. It is based on variables that are sensitive to the difference between these two types of jets. For 3-prong $\tau_{\text{had-vis}}$ candidates as a template the $\tau_{\text{had-vis}}$ BDT score is used. For 1-prong $\tau_{\text{had-vis}}$ candidates, the so-called $\tau_{\text{had-vis}}$ width $w_\tau$ is used instead. It is defined as:

$$w_\tau = \frac{\sum \left[ p_T^{\text{track}} \Delta R\left(\tau_{\text{had-vis}}, \text{track}\right) \right]}{\sum p_T^{\text{track}}}, \tag{4.5}$$

for tracks satisfying requirement $\Delta R\left(\tau_{\text{had-vis}}, \text{track}\right) < 0.4$. Two templates, denoted by $f_{\text{multi-jet}}$ and $f_{W\text{+jets}}$, are obtained in the multi-jet CR and $W$+jets CR, respectively. Each corresponds to a linear combination of templates of gluon- and quark-initiated jets. Note that the fraction of gluon-initiated jets is by construction larger in the multi-jet control region. Then, a linear combination of the two templates is defined as:

$$f(x|\alpha_{\text{MJ}}) = \alpha_{\text{MJ}}\, f_{\text{multi-jet}}(x) + (1 - \alpha_{\text{MJ}})\, f_{W\text{+jets}}(x),$$

with one free parameter $\alpha_{\text{MJ}}$ ($x$ are bins of the $\tau_{\text{had-vis}}$ width or BDT score). This linear combination is then fitted to the normalised distribution of the $\tau_{\text{had-vis}}$ width or BDT score measured in the SR, by varying the parameter $\alpha_{\text{MJ}}$ followed by minimising a $\chi^2$-function. From the best-fit value of $\alpha_{\text{MJ}}$ combined fake factors are obtained as follows:

$$\text{FF}^{\text{comb}}(\text{i}) = \alpha_{\text{MJ}}(\text{i})\,\text{FF}^{\text{multi-jet}}(\text{i}) + [1-\alpha_{\text{MJ}}(\text{i})]\,\text{FF}^{\text{W+jets}}(\text{i}), \tag{4.6}$$

where the index $i$ again refers to each bin in the parameterisation of the FF, in terms of $p_T^\tau$ and number of associated tracks. Usually, the best-fit value of $\alpha_{\text{MJ}}$ is constrained between 0 and 1, except when it is extracted from a region where the initial fraction of gluon-initiated jets is larger than the corresponding fraction in the multi-jet CR, or lower than the corresponding fraction in the $W$+jets CR. The combined FFs are shown in Fig. 4.6 (right).

Table 4.3 shows the different sources of fake $\tau_{\text{had-vis}}$ objects in the $\tau$- and anti-$\tau$ regions, as well as the total fraction of fake $\tau_{\text{had-vis}}$ candidates in the SR. It also provides the same information estimated from the samples of simulated events corresponding to the multi-jet and $W$+jets control regions.

| | $\tau$+jets SR | | multi-jet CR | | $W(\ell\nu)$+jets CR | |
|---|---|---|---|---|---|---|
| | $\tau$ | anti-$\tau$ | $\tau$ | anti-$\tau$ | $\tau$ | anti-$\tau$ |
| electron | 2.9 | 2.1 | — | — | 2.7 | — |
| light-quark | 6.7 | 43.6 | 56.3 | 50.7 | 54.5 | 61.0 |
| $c$-quark | 2.0 | 13.4 | 7.3 | 9.7 | 11.1 | 13.1 |
| $b$-quark | 1.5 | 17.3 | 1.4 | 1.2 | 2.2 | 4.1 |
| gluon | 0.4 | 3.6 | 22.9 | 32.6 | 7.9 | 14.0 |
| other | 0.2 | 0.7 | 6.2 | 5.6 | 8.6 | 7.4 |
| fraction of fakes | 13.7 | 80.7 | 94.4 | 99.8 | 87.0 | 99.6 |

Table 4.3: Fraction (ordered by sources and total in the last row) of fake $\tau_{\text{had-vis}}$ objects (in %) for various event selections. The row "other" represents jets that are not matched to any of the sources outlined above. For the signal regions, a sample of $t\bar{t}$ events with at least one leptonically decaying top quark is used. For the multi-jet and $W$+jets control regions, the corresponding simulated samples are used to identify the sources of fake $\tau_{\text{had-vis}}$ objects. From [33].

It is worth noting that the distribution of the $\tau$ polarisation variable $\Upsilon$ is found to be different for $\tau_{\text{had-vis}}$ and anti-$\tau_{\text{had-vis}}$ candidates. This variable is strongly correlated to the leading-track momentum fraction, which is one of the input variables to the BDT used for the identification of $\tau_{\text{had-vis}}$ candidates [78]. On the other hand, $\Upsilon$ shows weak correlation with other variables

used as input to the final BDT discriminant. Hence, in order to properly model the shape of $\Upsilon$ in the signal regions, a Smirnov transformation [118] is used in the control regions where FFs are computed, based on the measured distributions of $\Upsilon$ for $\tau_{\text{had-vis}}$ and anti-$\tau_{\text{had-vis}}$ candidates. It is then applied to the anti-$\tau_{\text{had-vis}}$ candidates in the SR.

The Smirnov transformation works as follows: shapes of the $\Upsilon$ variable are obtained for $\tau_{\text{had-vis}}$ and anti-$\tau_{\text{had-vis}}$ in the CR, then a cumulative distribution function is calculated from these shapes ($F(\Upsilon)$). Using an inverse transformation of $F$, the corrected value is obtained:

$$\Upsilon_{\text{corr}} = F_{\tau}^{-1}((F_{\text{anti-}\tau}(\Upsilon)), \tag{4.7}$$

where $F_{\tau}^{-1}(x)$ is the inverse of the cumulative distribution function of $\tau_{\text{had-vis}}$ candidates and $F_{\text{anti-}\tau}(x)$ is the cumulative distribution function of anti-$\tau_{\text{had-vis}}$ candidates.

Figure 4.7 shows the distribution of $\Upsilon$ and the ($F(\Upsilon)$) in the $W$+jets CR for 1-prong objects (the procedure is only applied to 1-prong objects, since $\Upsilon$ is not used in the training of the final BDT discriminant for 3-prong $\tau_{\text{had-vis}}$ candidates).

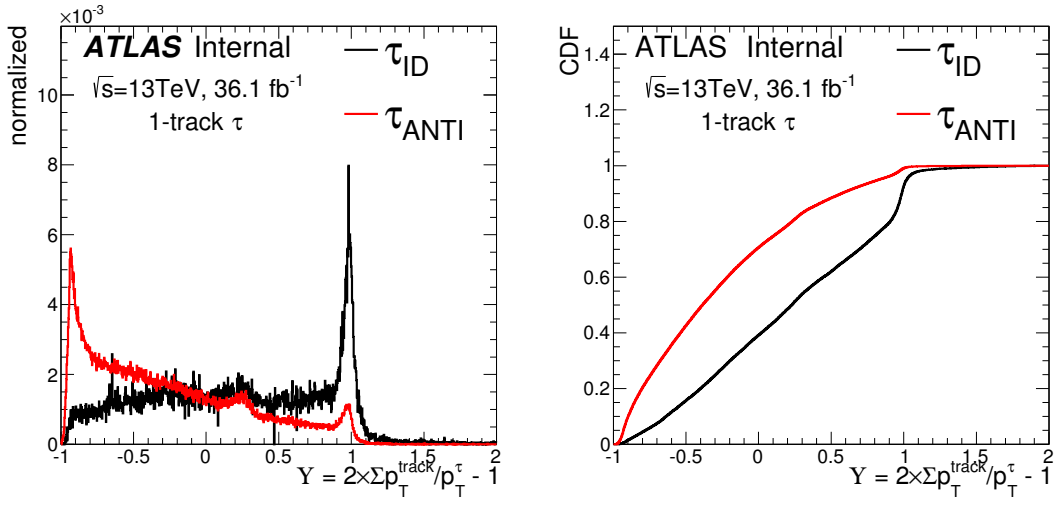



Figure 4.7: Left: Distribution of the $\Upsilon$ variable for $\tau_{\text{had-vis}}$ (black) and anti-$\tau_{\text{had-vis}}$ (red) candidates in the $W$+jets CR. Right: $F(\Upsilon)$ of $\Upsilon$ for $\tau_{\text{had-vis}}$ (black) and anti-$\tau_{\text{had-vis}}$ (red) candidates in the $W$+jets CR. From [33].

The distributions of $\Upsilon$ before and after the Smirnov transformation in the $\tau_{\text{had-vis}}$+jets $b$-veto are shown in Fig. 4.8. The clear improvement in modelling of $\Upsilon$ variable after application of the transform is visible.

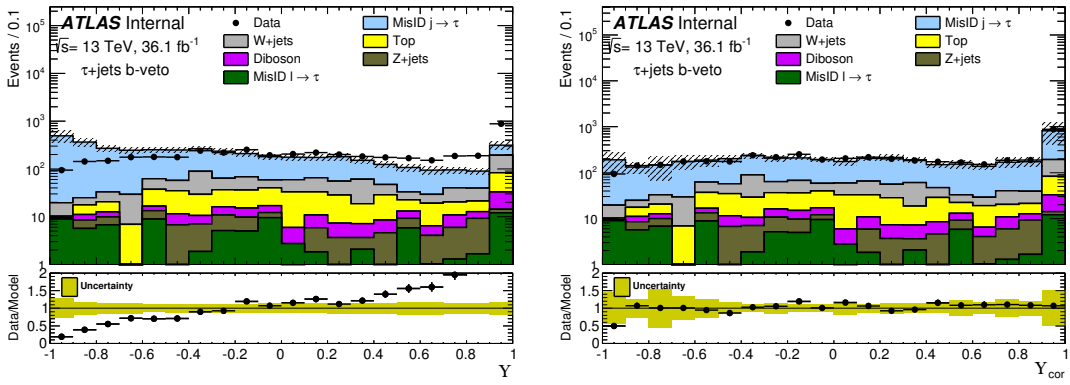



Figure 4.8: Distribution of Υ variable before (left) and after (right) Smirnov transformation in the signal-like region with a *b*-jet veto applied. From [33].

## 4.4 Systematic uncertainties

The normalisation of the background and signal processes and the shape of the BDT score distribution used as the final discriminant are affected by several sources of systematic uncertainty. Individual sources of systematic uncertainty are assumed to be uncorrelated. However, when applied to different samples of simulated events, correlations of a given systematic uncertainty are taken into account across processes.

All instrumental sources of systematic uncertainty are considered, i.e. from the reconstruction and object identification, from the energy scales and resolutions of electrons, muons, (*b*-tagged) jets and $\tau_{\text{had-vis}}$ candidates. Their impact on the reconstructed $E_T^{miss}$ is also included. The dominant systematic uncertainties come from the jet energy scale (between 1% and 4.5% depending on the jet $E_{\text{T}}$), the *b*-tagging efficiency (ranging from 2% to 10% depending on the jet $E_{\text{T}}$), the reconstruction and identification efficiencies of $\tau_{\text{had-vis}}$ candidates (5-6%), as well as their energy scale (2-3%). The uncertainty of 2.1% on the integrated luminosity is applied directly to the event yields of all simulated events. The corresponding small variations of the fit function of the $E_T^{miss}$ trigger efficiency, together with the statistical uncertainty used for the fit function, are accounted for as systematic uncertainties, which amount to an uncertainty of 1.4% on the event yield in the SR.

In the estimation of backgrounds with jets misidentified as $\tau_{\text{had}}$ candidates, the dominant sources of systematic uncertainty are:

- the requirement on the $\tau_{\text{had}}$ BDT output score used in the definition of the anti-$\tau_{\text{had}}$ control sample, which modifies the corresponding fractions of quark- and gluon-initiated jets, as well as the event topology (assessed by considering the shape of the final discriminant

obtained for two alternative cuts on the BDT output score that are symmetric around the nominal cut value);

- the level of contamination of $\tau_{\text{had}}$ candidates matching a true $\tau_{\text{had}}$ decay at the generator level and fulfilling the anti-$\tau_{\text{had}}$ selection (varied by 50%);

- the statistical limitation due to the size of the control sample;

- the statistical error on the best-fit value of $\alpha_{\text{MJ}}$,

- the error on the best-fit value of $\alpha_{\text{MJ}}$, obtained by considering the values of $\alpha_{\text{MJ}}$ at the edges of the band within $\sqrt{2/\text{ndf}}$ of the smallest $\chi^2$ value, where ndf is the number of degrees of freedom in the template fit;

- the modelling of heavy-flavor jets mimicking $\tau_{\text{had}}$ candidates, obtained by computing the fake factors separately for light- and heavy-quark-initiated jets, as in Ref. [86], and comparing those with the nominal predictions, then using the difference as a systematic uncertainty;

- for the $\Upsilon$ distribution only, the systematic uncertainty on the Smirnov transformation is taken as the difference between the resulting $\Upsilon_{\text{corr}}$ obtained in the multi-jet and $W$+jets control regions.

The dominant background with a $\tau_{\text{had}}$ candidate matched to a true $\tau_{\text{had}}$ object at the generator level is the production of $t\bar{t}$ pairs and single-top-quark events. A normalisation factor is computed for this background by including the control region of the $\tau_{\text{had-vis}}$+lepton channel with an $e\mu$ pair and at least one $b$-jet as a single-bin distribution in the statistical analysis. However, other $t\bar{t}$ modelling uncertainties are considered. These are systematic uncertainties due to the choice of the parton shower and hadronisation models; the systematic uncertainties arising from initial- and final-state parton radiation; the uncertainty due to the choice of matrix-element generator. The impacts of these three systematic uncertainties on the event yield of the $t\bar{t}$ background are 14%, 4%, 13%, respectively. The corresponding uncertainties for the $W/Z$+jets and diboson backgrounds, are 35%, 40% and 50%, respectively.

In the end, systematic uncertainties in the $H^{\pm}$ signal generation are estimated in the following steps:

- the one due to the missing higher-order corrections is assessed by varying the factorisation and renormalisation scale up and down by a factor of two;

- the largest variation of the signal acceptance is symmetrised and taken as the scale uncertainty, $4-8\%$ depending on the $H^{\pm}$ mass hypothesis;
- the variation of the signal acceptance with various Parton Distribution Function sets is estimated using LHAPDF [119], and it is found to be negligible for all signal samples;
- underlying-event, jet-structure and different aspects of extra jet production effects are covered by adding in quadrature the departues from a sub-set of tune variations. This uncertainty amounts to $8-10\%$.

In the low- and intermediate-mass $H^{\pm}$ search, the main systematic uncertainties arise from the estimation of the background with misidentified $\tau_{\text{had-vis}}$ candidates, as well as the reconstruction and identification of $\tau_{\text{had-vis}}$ candidates. For large $H^{\pm}$ masses, systematic uncertainties from the signal modelling and the estimation of the background with misidentified $\tau_{\text{had-vis}}$ candidates dominate. It should be pointed out that the search in this region is also limited by the number of selected events.

# Chapter 5

# Multivariate analysis

> "There is no magic in MVA-Methods: ... no "artificial intelligence" ... just "fitting decision boundaries" in a given model".

In this Chapter the techniques of MVA and studies of Boosted Decision Trees application to $H^{\pm}$ search are discussed. The general introduction to the subject is contained in Section 5.1. Section 5.2 presents the application of BDT to the current $H^{\pm}$ search. This method was designed and optimised by the Author and for the first time applied to the analysis which was published by the ATLAS Collaboration [12]. Finally, Section 5.3 describes studies performed by the Author on possible improvements of the BDT performance in $H^{\pm} \to \tau^{\pm}\nu$ analysis when full Run-2 dataset would be available.

## 5.1 Multivariate analysis in particle physics

Taking into account intellectual and financial investments in the accelerator facilities and experiments it is of the great importance to make the best use of the output of this investment, i.e the data collected. Hence, it is crucial to use the most efficient techniques for the analysis of these data at all stages of the experiment.

One of the most challenging tasks in data analysis is to identify events that are rare and simultanously obscured by the wide variety of processes, known as backgrounds, that can mimic the searched process, known as a signal. This is like "finding needles in a haystack" for which the conventional approach of selecting events by using cuts on individual kinematic variables can be insufficient. Having in mind also that, together with improvements in detector design and increasing the number of variables which have to be taken into account, the usage of *multivariate methods* starts to be essential.

The multivariate analysis (MVA) is a set of statistical methods that simultaneously analyse multiple measurements or variables describing a given object that can be dependent or corre-

lated in various ways. In conventional statistical techniques, parameters of a given mathematical model are found either analytically or numerically in the goal of providing predictions for future data. However, dealing with vast amounts of data puts pressure on the development of automated algorithms for learning from data named *machine learning* (ML), where an approximating function $f$ is inferred from the given data without requiring a priori information about it. One of the most powerful approach to obtain the approximation of the unknown function is *supervised learning* in which a training data set, inputs (feature vectors) and the corresponding desired outputs (targets), is used. The training data set $\{y,x\}$, where $y$ is the target and $x$ is a real vector, encodes information about the input-output relationship to be learned. Hence, under the usual main assumption that data are generated from a probabilistic distribution ($p(x,y)$), MVA together with ML allows getting information about statistical dependencies beetween variables using just a training sample, without any explicit knowledge about the observed processes. Therefore Machine Learning allows to construct the highly optimised classifiers.

The multivariate analysis in particle physics is mainly used for: *classification* - the process of assigning objects or events to one of the possible discrete classes and *parameter estimation (regression)* - extraction of one or more parameters by fitting a model to data such as measurements of track parameters, vertices or physical parameters like production cross sections, branching ratios and masses [120]. For example, in the identification of particles (e.g. electrons, $\tau$ lepton, photons, $b$-jets) and in signal and background discrimination as can be seen in several completed analysis [121–123].

It is worth mentioning that classification of objects such as identification of particles and events, or discrimination of signal events from those arising from background processes is one of the most important analysis tasks in high energy physics. Where optimal discrimination between classes is crucial to obtain the signal-enhanced samples for precision physics measurements. Therefore, it is important to remember that good understanding of the inputs before starting playing with multivariate techniques is so desirable, e.g. finding observables with a good separation power between signal and background with little correlations amongst each other. Extraction of features is a bit of an art that requires intuition concerning both the goal of the specific analysis and also what the machine learning algorithms are capable of.

In this analysis binary classification is used, which is also the most common use of ML in high energy physics.

### 5.1.1 Multivariate Treatment of Data

In particle physics data characterising an object, like a particle (e.g. $\tau$ lepton) or an event, generally use multiple quantities such as e.g. the four-vectors (four-momentum), energy deposited

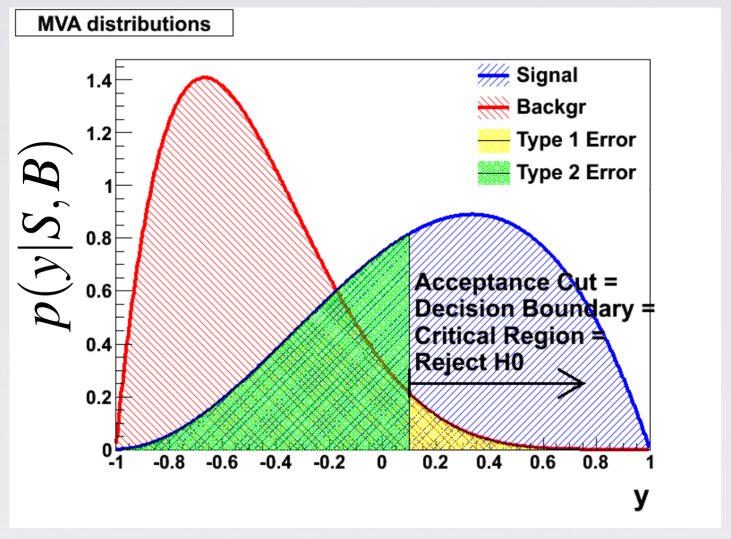



Figure 5.1: The probability distributions functions: $p(y|S)$ and $p(y|B)$ of MVA variable $y$ for signal and background events. The classification is based on a cut on the MVA variable $y$, i.e.: if the measured $y$ is above the indicated threshold value (*vertical line*) then the event is selected as signal, otherwise it is rejected as background. *Type 1 Error* represents misclassified background events as signal one (rejection of the null hypothesis (H0) when it is true); *Type 2 Error* represents misclassified signal events as background (acceptation of the null hypothesis when it is not true). In the context of event classification and selection of the *signal* events, the null hypotesis (H0) is that an event is a *background* and is either rejected or not, depending on the value of the *test statistic* $y(\mathbf{x})$ which is the MVA $y$ variable here [124].

in the calorimeter cells or width of the electromagnetic cascade, allude to *feature variables*. These variables can be represented by vectors $\mathbf{x} = (x_1, x_2, ..., x_d)$ in a d - dimensional feature space. Having selected a set of variables, a proper transformation can be applied to them to yield a representation of the data which exhibits certain desirable properties. Thus, the goal is to construct a function $y = f(\mathbf{x})$ with properties that are useful for subsequent decision-making, what means the extraction of a map: $f : \Re^d \to \Re^n$ (with $n < d$), i.e.

$$x \longrightarrow \boxed{f} \longrightarrow y,$$

where predictor is a function $f$ that maps an input $\mathbf{x}$ to an output $y$. For selection the output space is 1-dimensional and the output is just a real number. Function $f$ is called *a classifier*, and $y$ is called a *class* or a *target*. For binary classification $y \in \{-1 \text{ or } 0, +1\}$, e.g. signal $= +1$ and background $= -1$ or 0. In general MVA methods combine information of all observables $\mathbf{x}$ of an event into one or more output variables $y$ and this variables can then be used to decide if the event is selected as signal or rejected as background, as indicated in Fig. 5.1.

The starting point of machine learning are the data, which are the main resource that one can use to address the information complexity of the prediction task. In high energy physics training

sample set usually comes from Monte Carlo simulations. The process of learning is about taking the training sample set in the goal of producing a predictor function $f(\mathbf{x},\mathbf{w})$, where $\mathbf{w}$ are the weights vectors, i.e. for each feature $i$ there exist a real number $w_i$ representing contribution of the feature to prediction. Therefore, MVA methods attribute a probability for an event observed with features $\mathbf{x}$, resulting in $y(\mathbf{x}) = y$ to belong to a certain class, rather then assigning a definite class membership to it:

$$P(S|y) = \frac{P(y|S)\cdot f_s}{P(y|S)\cdot f_s + P(y|B)\cdot f_b} \tag{5.1}$$

where $f_s$ and $f_b = (1-f_s)$ are fraction of signal and background events in the sample, respectively. $P(y|x) = \frac{P(x|y)P(y)}{P(x)}$, where $P(y|x)$ is the probability of an hypotesis $y$ in the light of the data $x$ corresponds to probability of data $x$ under the assumption of hypotesis $y$ multiplied by the previous probability of hypotesis $y$ and divided by probability of data $x$ (where $P(x) \neq 0$).

It is important to mention that in all approaches to functional approximation the information loss is taking place and the question is how to minimize the influence of this process. In other words the learning action can be treated as an optimisation problem, where the information loss is quantified by a loss function: $L(y, f(\mathbf{x},\mathbf{w}))$. It quantifies how good the usage of the weight parameter $w_i$ is to make a prediction on $x_i$ when the correct output is $y$ and $L(y, f(\mathbf{x},\mathbf{w}))$ is the object we want to minimize. Therefore, a typical loss function represents the total number of misclassified events or, equivalently, the sum of type 1 and type 2 errors (see Fig. 5.1)[1].

In practice, the minimization of the loss function is alwayes averaged over the training data set. The learning algorithm minimizes the average loss, called the risk, quantified by the risk function $R(\mathbf{w})$ that measures the cost of mistakes made in the predictions, and finds the best parameters $\mathbf{w}$. The empirical risk, an approximation to true risk, is defined as the average loss over all (N) predictions:

$$R(\mathbf{w}) = \frac{1}{N}\sum_{i}^{N} L\left\{y_i, f(\mathbf{x_i},\mathbf{w})\right\}. \tag{5.2}$$

Where, a commonly, but not always, used risk function is the mean square error given by:

$$R(\mathbf{w}) = E(\mathbf{w}) = \frac{1}{N}\sum_{i}^{N}(y_i - f(\mathbf{x_i},\mathbf{w}))^2. \tag{5.3}$$

The risk minimization can be performed using various algorithms, which attempt to find the global minimum (usually only a local minimum is attainable) of the risk function in the parameter space. It is worth to mention that the generic method is that of the *gradient descent* and the performance of the classifier (or estimator) is usually evaluated using the test data sets independent of the training sets [120].

[1] Loss function might be for example $\chi^2 = \sum(x_i - y_i)^2$, so it is not always based on the number of misclassified events (see Eq. 5.2 and 5.3).

### Stochastic gradient descent

As was mentioned above, having defined a group of different functions that correspond to training loss, one would like to obtain an algorithm that outputs the weight $\mathbf{w}$ where the risk function achieves the minimum value. Hence, it is optimisation problem ilustrated on Fig. 5.2 where the

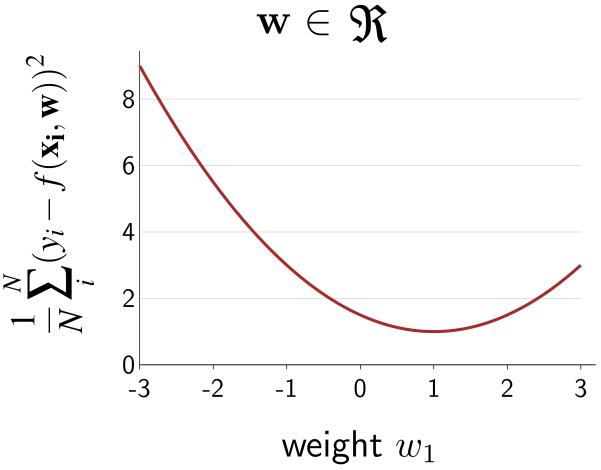



Figure 5.2: Example gradient plot for weight $w_1$.

gradient is the direction that increases the loss the most:

$$\frac{1}{N}\sum_i^N 2(y_i - f(\mathbf{x_i}, \mathbf{w}))\frac{\partial f(\mathbf{x_i})}{\partial \mathbf{x_i}}. \tag{5.4}$$

Now, the iterative optimisation called the *gradient descent* tries to tweak $\mathbf{w}$ so that the risk function value decreases. This procedure has two parameters, the step size $\eta$ (which specifies how aggressively one wants to pursue a direction) and the number of iterations $T$. Alas each iteration requires going over all training examples, which is expensive when one has lots of data (MC samples) in large-scale machine learning applications where the training loss is a sum over the training data. Therefore, in many cases the *stochastic gradient descent* is applied. Instead of looping through all the training examples to compute a single gradient and making one step, the algorithm loops through the random subset of examples $(x_i, y_i)$ and updates the weights $\mathbf{w}$ based on each of them [125].

In the presented analysis the stochastic gradient-boosted decision trees named *FastBDT* was used (see Section 5.1.2 below).

### ROC curve and Neymann-Pearson lemma

An important role, in making the decision which algorithm from binary classifiers is the best, is fulfilled by a receiver operating characteristic curve, *ROC curve*, which illustrates the diagnostic

ability of each of binary classifiers. It shows the relation between the signal efficiency and the background rejection, which is visualised on Fig. 5.3, where the algorithm with the largest area underneath the curve has on average the best performance.

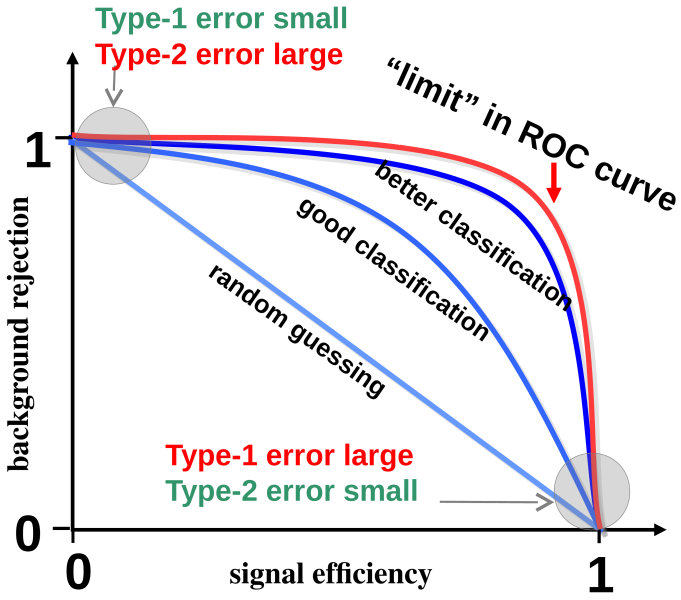



Figure 5.3: The ROC curve, showing the background rejection as a function of the signal efficiency achieved by varying the cut on the MVA output variable $y$.

The Neymann-Pearson lemma states that the best classification quality is provided by Bayes optimal classifier

$$\frac{P(y=1(S)|x)}{P(y=0(B)|x)}, \tag{5.5}$$

which maximises the area under the ROC curve ("limit" in ROC curve: see Fig. 5.3). Thus, according to the Neymann-Pearson lemma, selection algorithm based on the ratio of the particle distributions functions, i.e. the likelihood ratio Eq. (5.5), is optimal for retaining the highest signal efficiency for a given background efficiency. However, typically one does not know the exact probability distribution functions and therefore it is very desirable to construct a suitable variable which can be used as a multivariate classifier.

### Final remarks

Within the particle physics community several methods are particularly relevant and popular, i.e.: Naive Bayes Classifier, k-Nearest Neighbour, Fisher Linear Discriminant, Neural Networks and Decision Trees (DT) which was used in presented analysis (for review see [124] and [126]).

All of listed classifiers are trained using *"simulated data"* (MC samples), where the class membership of an event is known and from these training data the classifiers *"learn"* how the parameters of the decision boundaries are chosen to get optimal separation between the signal and background events.

### 5.1.2 Boosted Decision Trees

Of all the above listed learning methods, decision trees are closest to meet the requirements for serving as an off-the-shelf procedure[2]. They are relatively fast to construct and produce interpretable (human readable) models, i.e. the DT algorithm can be interpreted as selection rules applied to certain objects. This is due to the fact that each training step involves only a one-dimensional cut optimisation. What is more, decision trees are invariant under strictly monotonous transformations of the individual predictors what causes that the scaling is not a problem. They are also resistant to the effects of predictor outliers what protects from the inclusion of many irrelevant predictor variables, i.e. during the training step decision trees algorithm ignores non-discriminating variables as for each node splitting only the best discriminating variable is used. On the other hand, the DT method is vulnerable to overtrain coming from "learning" statistical fluctuations "by heart".

As it is presented on Fig. 5.4 decision trees are tree-structured classifiers that consist of a series of binary splits. The bulding or *training* of a DT is a process that defines the splitting criteria for each node. The training starts with the *root node* and is built up of repeating splits and nodes down to the final or *leaf* nodes, i.e. the split results in two subsets of training events that each goes through the same algorithm of determining the next splitting iteration, where at each node, the split is determined by finding the variable and corresponding cut value that provides the best separation between signal and background. Wherein the separation is often measured in terms of the *Gini index* [126] (with $p$ denoting purity, $S$ signal and $B$ background):

$$Gini_{index} = \sum_{i=S,B} p_i(1-p_i) = 2p(1-p) = \frac{2SB}{(S+B)^2}. \tag{5.6}$$

For most decision trees the split criteria are simple cuts on individual features (observables) as is shown on Fig. 5.4. At the end of the process, the leaf nodes are classified as the signal or background according to the class the majority of events belongs to, i.e according to their purity, $\frac{S}{S+B}$: close to 1 for signal and $0(-1)$ for background. The best split variables are determined

[2] An off-the-shelf method is one that can be directly applied to the data without requiring a great deal of time - consuming data preprocessing or careful tuning of the learning procedure.

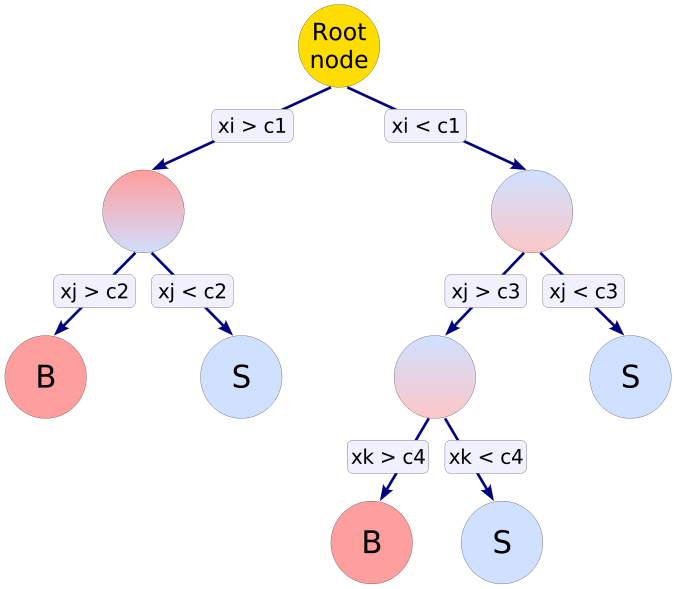



Figure 5.4: Schematic view of a decision tree. Starting from the *root node*, a sequence of binary splits using the discriminating variables $x_i$ is applied to the data. Each split uses the variable that at this node gives the best separation between signal and background when being cut on. The same variable may thus be used at several nodes, while others might not be used at all. The leaf nodes at the bottom end of the tree are labeled S for signal and B for background depending on the majority of events that end up in the respective nodes [127].

by comparing the Gini index before and after the split. The latter is defined by the sum of the indices of the two daughter nodes, weighted by the respective fraction of events in the nodes:

$$p = \frac{\sum_{i \in signal_{events}} w_S^i}{\sum_{i \in signal_{events}} w_S^i + \sum_{i \in background_{events}} w_B^i}. \tag{5.7}$$

The best split is then performed to separate the training sample into two daughter nodes for which the whole procedure is re-iterated [126]. Therefore, one ends with the separation, denoted by $\langle S^2 \rangle$, and defined as [127]

$$\langle S^2 \rangle = \int_{-1}^{1} \frac{(\hat{y_S}(y) - \hat{y_B}(y))^2}{\hat{y_S}(y) + \hat{y_B}(y)} dy, \tag{5.8}$$

where the probability density functions of the output of the classifier are denoted as $\hat{y_S}$ (for signal) and $\hat{y_B}$ (for background). Note that for a complete overlap between the signal and background distribution the separation is equal zero. On the other hand, distributions without any overlap gives a separation of one.

As it was mentioned above a shortcoming of decision trees is their instability with respect to statistical fluctuations in the training sample from which the tree structure is derived. This problem is overcame by constructing a forest of decision trees and classifying an event on a majority vote of the classifications done by each tree in the forest. All trees are trained with data samples that are derived from the training events by reweighting the events according to the *boost* procedure (see Fig. 5.5), i.e. trees are trained in sequence, and misclassified events are reweighted (boosted) in the training of subsequent trees.

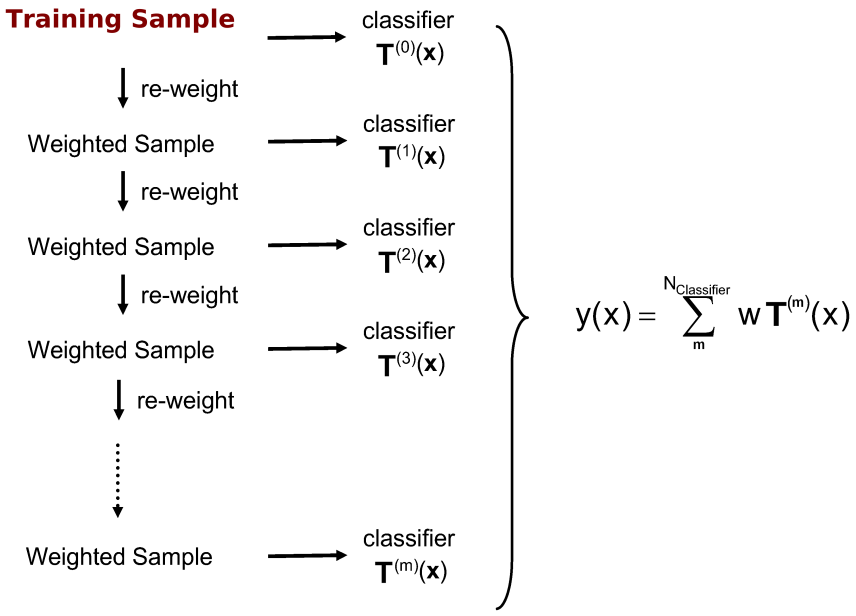



Figure 5.5: Scheme of the boosting procedure. Note that this method can be applied also to other classifiers.

The statistical stability of the classifier is increased by the boosting recipe, which is also able to improve the separation performance compared to a single decision tree. What is worth to mention, the boosting performs in the best way if it is applied to trees (classifiers) that, taken individually, have not much classification power (so called "weak classifiers"), i.e. the individual tree growing is stopped at quite an early stage, resulting in trees which have only a few split levels. Therefore, by limiting the tree depth during the tree building process (training), the tendency of overtraining for simple decision trees is almost completely eliminated.

In the presented analysis the output score of *Boosted Decision Trees* was used in order to separate the $H^{\pm}$ signal from the Standard Model background processes. The training of the BDT was performed using the FastBDT [128] library via the TMVA toolkit [127], where

FastBDT algorithm uses stochastic gradient boosting for training [129]. What is more, in order to be completely safe from overtraining issue *K-Fold Cross-Validation* method was used (see Sec. 5.1.2).

### Hyperparameters

This section gives an overview of the hyperparameters, i.e. those parameters which are not determined automatically by the BDT algorithm and needed to be set by the user. FastBDT offers five different hyperparameters:

- `NTrees`: number of trees which are used for building the classifier and during the boosting procedure;
- `Shrinkage`: the learning rate for gradient boost procedure;
- `RandRatio`: in FastBDT every tree is trained on a subset of the events and the relative size of this subset is determined by this parameter;
- `NTreeLayers`: depth of the trees which gives the information about how many consecutive cuts are performed in a single tree;
- `NCutLevel`: Number of Cut Levels used to control the number of bins in the training step.

Note that generally, a large number of trees and a low learning rate are preferred, since they decrease the overtraining chance and lead to a lower misclassification rate.

### K-Fold Cross-Validation

Generally, there exists a necessity to validate[3] the stability of chosen machine learning technique, i.e a kind of assurance that the algorithm has got most of the patterns from the data correctly (it is low on bias and variance). In an ideal situation, i.e having enough data, one would set aside a validation set and use it to assess the performance of the chosen prediction model. But data are often scarce and this is simply not possible. To solve this issue, K-Fold Cross-Validation method uses a part of the available data to fit the model and a different part to test it, i.e the data are divided into $k$ subsets and each time from $k$ ones, one of the $k$ subsets of data is used as the test set (validation set) and the other $k-1$ subsets are put together to form a training set (see Fig. 5.6). Hence, every data point gets to be in a validation set exactly once and

[3]The process of deciding whether the numerical results quantifying hypothesized relationships between variables are acceptable as descriptions of the data is known as *validation*.

gets to be in a training set $k-1$ times. Then the error estimation is averaged over all $k$ trials to get the total effectiveness of the prediction model. This together significantly reduces the bias since most of the data are used for training, and also significantly reduces variance as most of the data are also being used in the validation set [130].

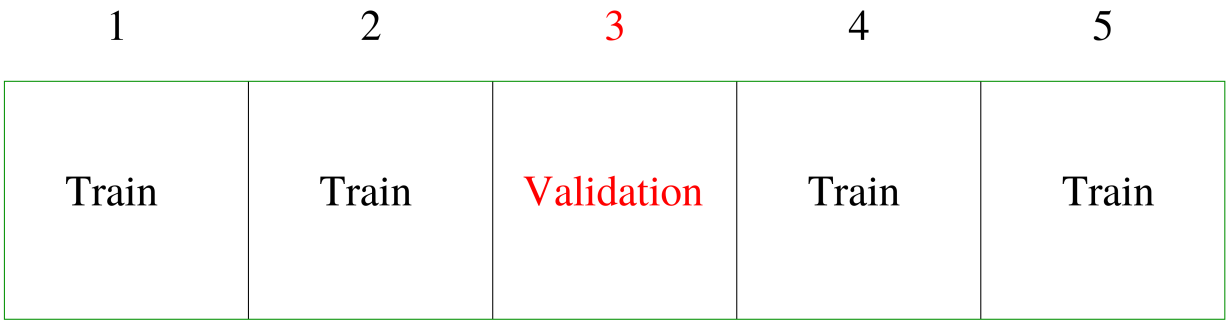



Figure 5.6: For the k-th part (third above), the model is fitting to the other $K-1$ parts of the data, and calculation of the prediction error of the fitted model during predicting the k-th part of the data is done. The procedure is repeated for $k = 1, 2, ..., K$ and combination of the $K$ estimates of prediction error is prepared [126].

The presented above $k$-fold training method is used in the analysis described below. The input samples are divided into 5 equally populated subsets and each 4 of them is used to independently train one of the 5 BDT discriminants. Each of the BDT discriminants is then applied to the remaining subset that is not used for training. Therefore, each time $4/5$ of the available statistics is used for training and the analysis remains immune to overtraining [126]. The final step is to take an average of such trained BDT to perform the analysis.

## 5.2 BDT in the current $H^{\pm}$ searches

**Multi-variate discriminant**

Following the event selections described in Sec. 4.1, the output score of *Boosted Decision Trees* method is used to separate the $H^{\pm}$ signal from the SM background processes. The training of the BDT is performed using the FastBDT [131] library via the TMVA toolkit [132]. The $k$-fold method is used for training and classification: events are divided into $k$ sets, and they are classified using a BDT trained on the signal and backgrounds from the other $k-1$ sets, thereby allowing every event to be classified with a BDT that does not include this event in its training set. While $k = 2$ is sufficient to ensure independence of training and classification of the sets,

$k = 5$ is used in this analysis to increase the size of the training set for each BDT (see previous Secion for more details). Optimisation of the hyperparameters described in Sec. 5.1.2 was done using "grid-search", i.e. by constructing all of the hyperparameters possible combinations and chosing the setup for which the FastBDT algorithm turned out to have the largest area underneath the ROC curve. As a result of optimisation the chosen hyperparameters are:

- for 1-prong: NTrees= 1000; Shrinkage= 0.05; RandRatio= 0.6; NTreeLayers= 3; NCutLevel= 7,
- for 3-prong: NTrees= 500; Shrinkage= 0.02; RandRatio= 0.4; NTreeLayers= 5; NCutLevel= 9.

The signal samples are divided into five $H^{\pm}$ mass bins, in which the kinematic distributions of the input variables and the event topology are found to be similar enough to ensure that the higher statistics from an inclusive training improves the performance:

- 90–120 GeV,
- 130–160 GeV (using the low-mass 160 GeV sample),
- 160–180 GeV (using the intermediate-mass 160 GeV sample),
- 200–400 GeV,
- 500–2000 GeV.

All available $H^{\pm}$ signal samples corresponding to a given mass bin are combined into one inclusive signal sample.

The BDT is trained on $H^{\pm} \to \tau^{\pm}\nu$ signal and top ($t\bar{t}$ and single-top quark) background MC samples using the final (SR) selection defined in Sec. 4.1.4. For the first four $H^{\pm}$ mass ranges, events arising from j $\to \tau$ fakes are taken from the simulated background samples. In the $H^{\pm}$ mass range 500–2000 GeV, the misidentified $\tau_{\text{had-vis}}$ candidates estimated with a data-driven method (see Section 4.3) are included in the training, as the multi-jet background with a misidentified $\tau_{\text{had-vis}}$ candidate dominates in this mass range. At the same time, the j $\to \tau$ fakes in the simulated background samples are excluded from the training to avoid double-counting.

The transverse mass of the $\tau_{\text{had-vis}}$ candidate and $E_T^{miss}$ is known to strongly discriminate the signal from background, particularly for high $H^{\pm}$ masses. In the MVA approach, this quantity is replaced by its three components - $p_T^{\tau}$, $E_T^{miss}$ and $\Delta\phi_{\tau,E_T^{miss}}$, carrying equivalent information. This way the MVA can also benefit from using potential correlations between those three constituents.

At low $H^{\pm}$ masses the kinematics of the $t \to bH^{\pm}$ and $t \to bW^{\pm}$ decay products can be very similar. In that case, the polarisation of the $\tau$ lepton can serve as a discriminating variable: in all SM background processes, the $\tau_{\text{had-vis}}$ object originates from a vector-boson decay, whereas for signal process it stems from the decay of a scalar $H^{\pm}$ boson [133]. The $\tau$ lepton decay mode with the highest branching ratio is via an intermediate $\rho$ resonance (about half of all hadronic 1-prong $\tau$ lepton decays). The polarisation of the $\tau_{\text{had-vis}}$ candidates in this decay mode can be measured by the asymmetry of energies carried by the charged and neutral pions from the $\tau$ lepton decay measured in the laboratory frame. For this purpose, the variable $\Upsilon$, as discussed in detail in Section 2.5, is used:

$$\Upsilon = \frac{E_T^{\pi^{\pm}} - E_T^{\pi^0}}{E_T^{\tau}} \approx 2\frac{p_T^{\tau\text{-track}}}{p_T^{\tau}} - 1. \tag{5.9}$$

It is defined for $\tau_{\text{had-vis}}$ candidates with only one associated track with transverse momentum $p_T^{\tau\text{-track}}$. For $H^{\pm}$ masses in the range 90–400 GeV, the BDT training is performed separately for events with selected 1- or 3-prong $\tau_{\text{had-vis}}$ candidate, and $\Upsilon$ is included in the final BDT discriminant for events where $\tau_{\text{had-vis}}$ has only one associated track. The importance of the other kinematic variables in the BDT training becomes dominant at large $H^{\pm}$ masses, in which case the BDT discriminant is inclusive in the number of tracks associated to the $\tau_{\text{had-vis}}$ candidate and does not contain the $\Upsilon$ variable.

In total, seven variables are used as an input to the BDT technique in the presented analysis:

- $p_T^{\tau}$, $p_T^{b\text{-jet}}$, $E_T^{miss}$,
- $\Delta\phi_{\tau,E_T^{miss}}$ - azimuthal angle between $\tau_{\text{had-vis}}$ candidate and the direction of the missing transverse energy,
- $\Delta\phi_{b\text{-jet},E_T^{miss}}$ - azimuthal angle between $b$-jet candidate and the direction of the missing transverse energy,
- $\Delta R_{b\text{-jet},\tau}$ - distance between reconstructed $b$-jet candidate and $\tau_{\text{had-vis}}$,
- $\Upsilon$ - only for 1-prong $\tau_{\text{had-vis}}$ candidates in mass ranges from 90 to 400 GeV as is explained above,

where as a $b$-jet candidate in above definitions, $b$-jet with the largest $p_T$ in the event is considered.

The distributions of all variables used in the SR as an input to the BDT discriminant are presented in Figures 4.1-4.2. Whereas, Figures 4.3 and 4.4 show distributions of those variables in the CR enriched in $t\bar{t}$ events. All variables have a good modelling both in the CR and SR.

Distributions of the chosen set of variables as used in BDT-training are shown in Fig. 5.7 for 90-120 GeV and 200-400 GeV, $H^{\pm}$ mass bins and in the remaining three mass bins can be found in Appendix C.

The correlation matrices of input variables are presented in Fig. 5.9 for the top backgrounds and in Fig. 5.10 for the signal. It can be seen that the $\Upsilon$ variable is uncorrelated with the other variables used in the training. Therefore, the Smirnov transformation discussed in Section 4.3 can be applied to the $\Upsilon$ distribution for $\mathrm{j} \to \tau$ background estimation with FFs. Note that procedure of the Smirnov transformation for $\Upsilon$ cannot introduce any bias in BDT trainings since it is only applied to $\mathrm{j} \to \tau$ background which is used for the trainings only in high $H^{\pm}$ mass range $(500 - 2000\ \mathrm{GeV})$ where $\Upsilon$ is not used.

Tables 5.1 - 5.5 show the ranking of the variables used in the BDT training for each $H^{\pm}$ mass range. Note that this is a measure of how often a given variable is used to split decision tree nodes. Thus, a variable can be ranked lower or higher in a specific set and a value of "Variable Importance" gives some information on whether an impact of a variable on the BDT output is significant. It can be seen that for low mass $H^{\pm}$ search $\Upsilon$ is the most important variable, while for high mass the components of $m_T$.

The BDT score was first validated in the $t\bar{t}$ enriched CR, see Fig. 5.11, where predictions and measurements are found to be in good agreement. The same good modelling of BDT score distributions can be observed in the SR for five $H^{\pm}$ mass ranges, which is presented in Fig. 5.12.

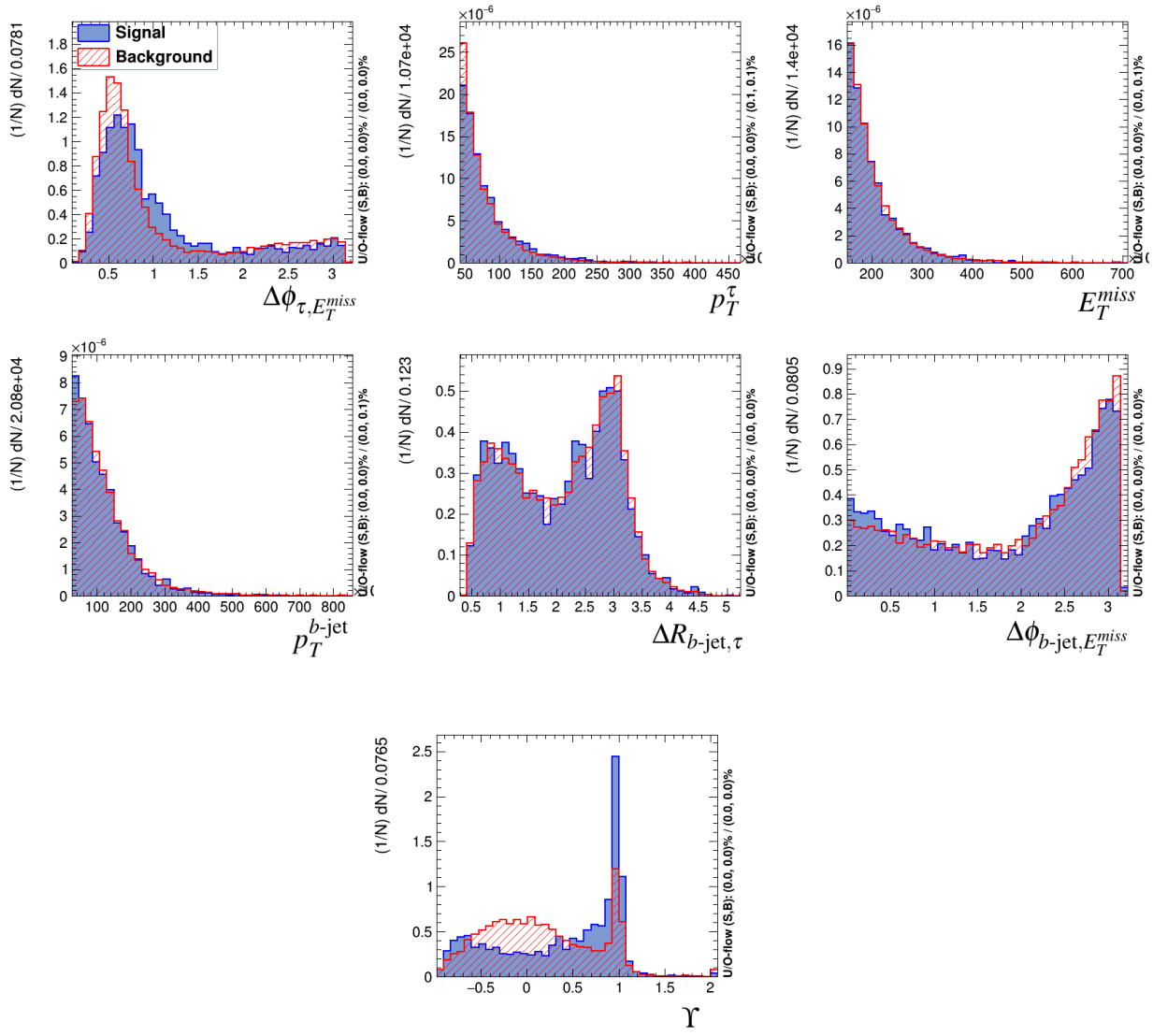



Figure 5.7: Distributions of input variables for BDT-training within 90-120 GeV mass bin. Signal is presented as a blue, solid histogram; top background (as estimated from MC) is presented as a red, dashed histogram. Histograms are arbitrarily normalised.

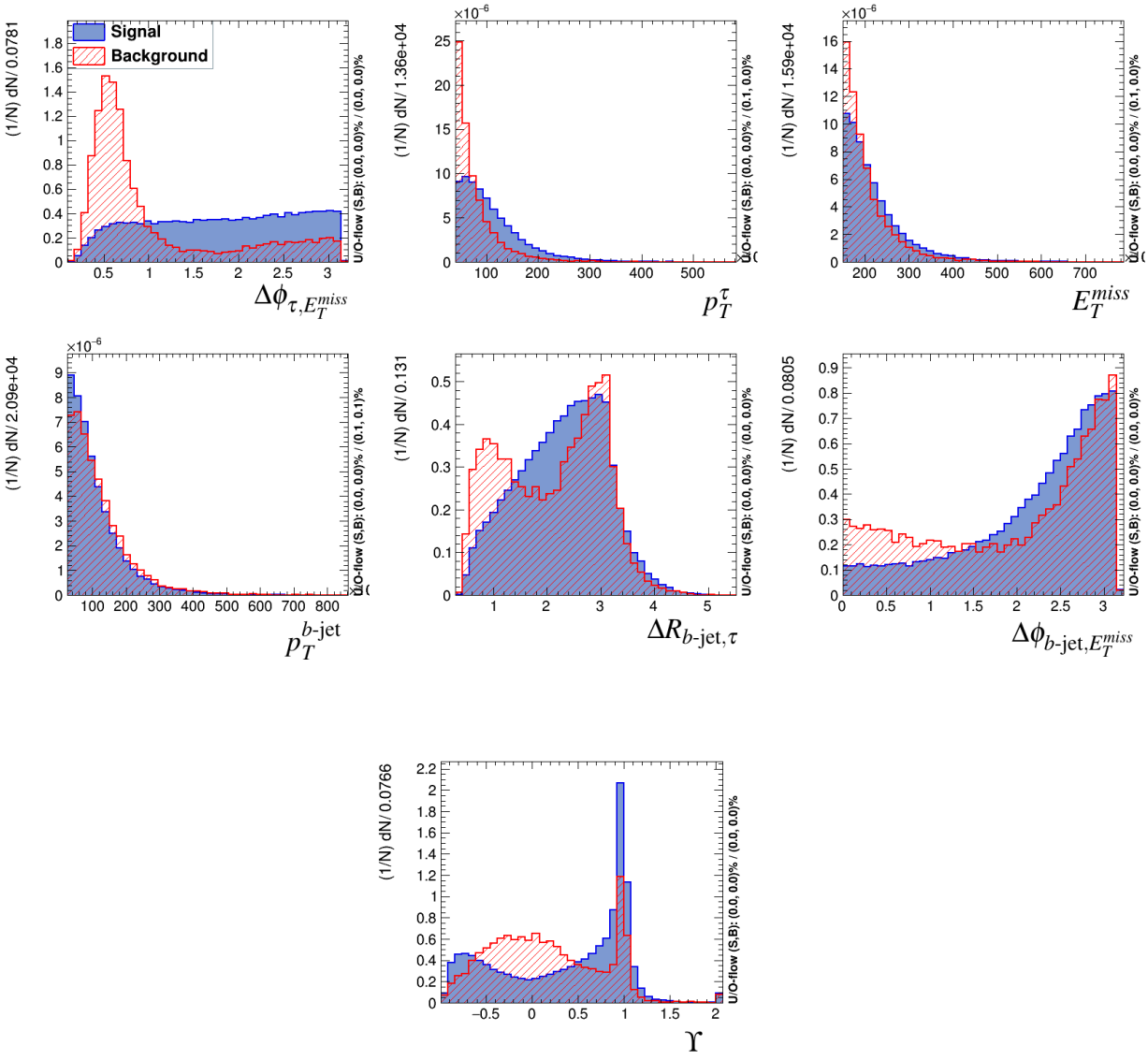



Figure 5.8: Distributions of input variables for BDT-training within 200-400 GeV mass bin. Signal is presented as a blue, solid histogram; top background (as estimated from MC) is presented as a red, dashed histogram. Histograms are arbitrarily normalised.

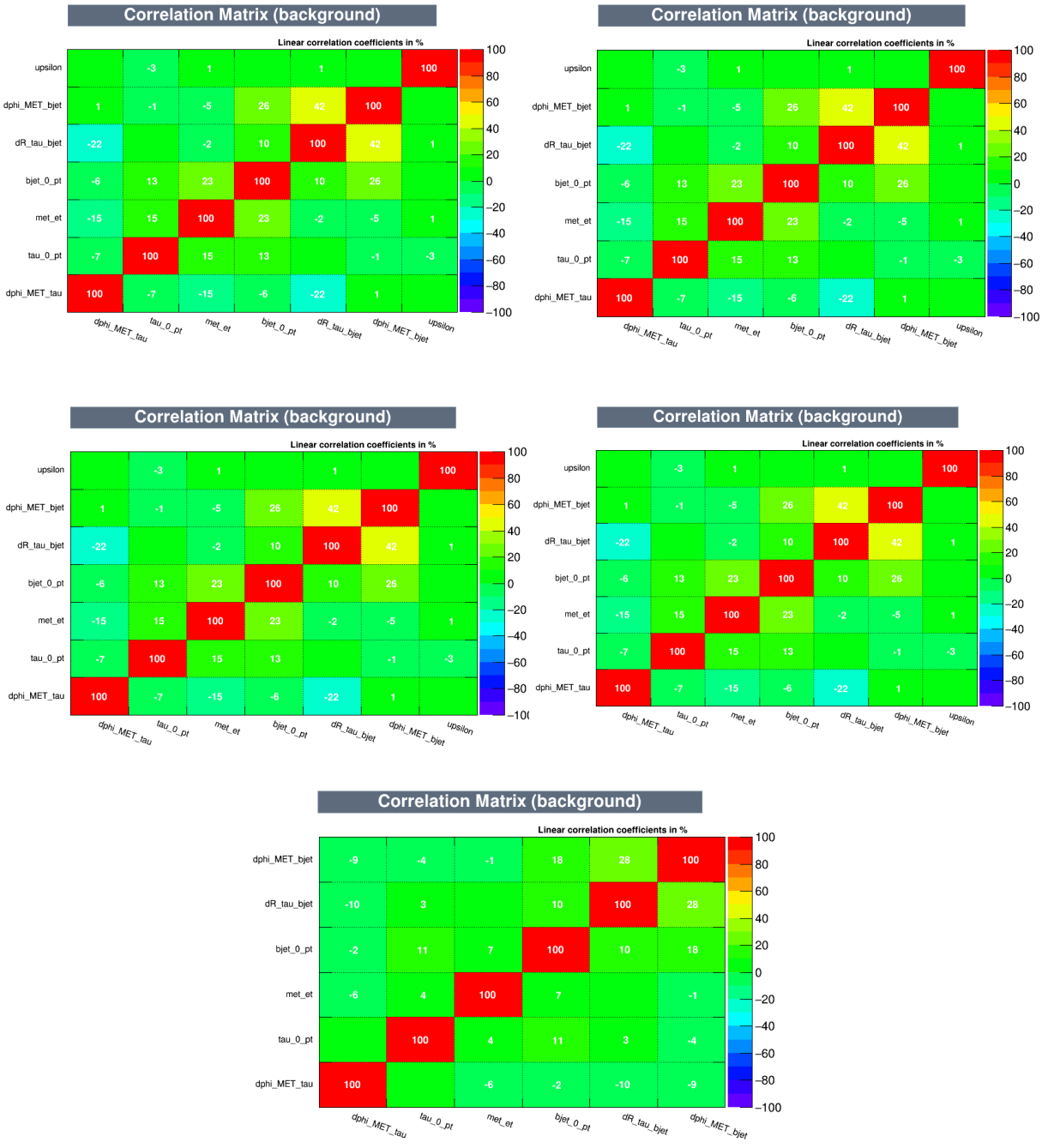



Figure 5.9: Correlation matrix of BDT input variables for the top backgrounds. The five $H^{\pm}$ mass ranges used in the BDT training are presented, 90–120 GeV (top left), 130–160 GeV (top right), 160–180 GeV (middle left), 200–400 GeV (middle right) and 500–2000 GeV (bottom). Where: upsilon= $\Upsilon$; dphi_MET_bjet= $\Delta\phi_{b\text{-jet},E_T^{miss}}$; dR_tau_bjet= $\Delta R_{b\text{-jet},\tau}$; bjet_0_pt= $p_T^{b\text{-jet}}$; met_et= $E_T^{miss}$; tau_0_pt= $p_T^{\tau}$ and dphi_MET_tau= $\Delta\phi_{\tau,E_T^{miss}}$.

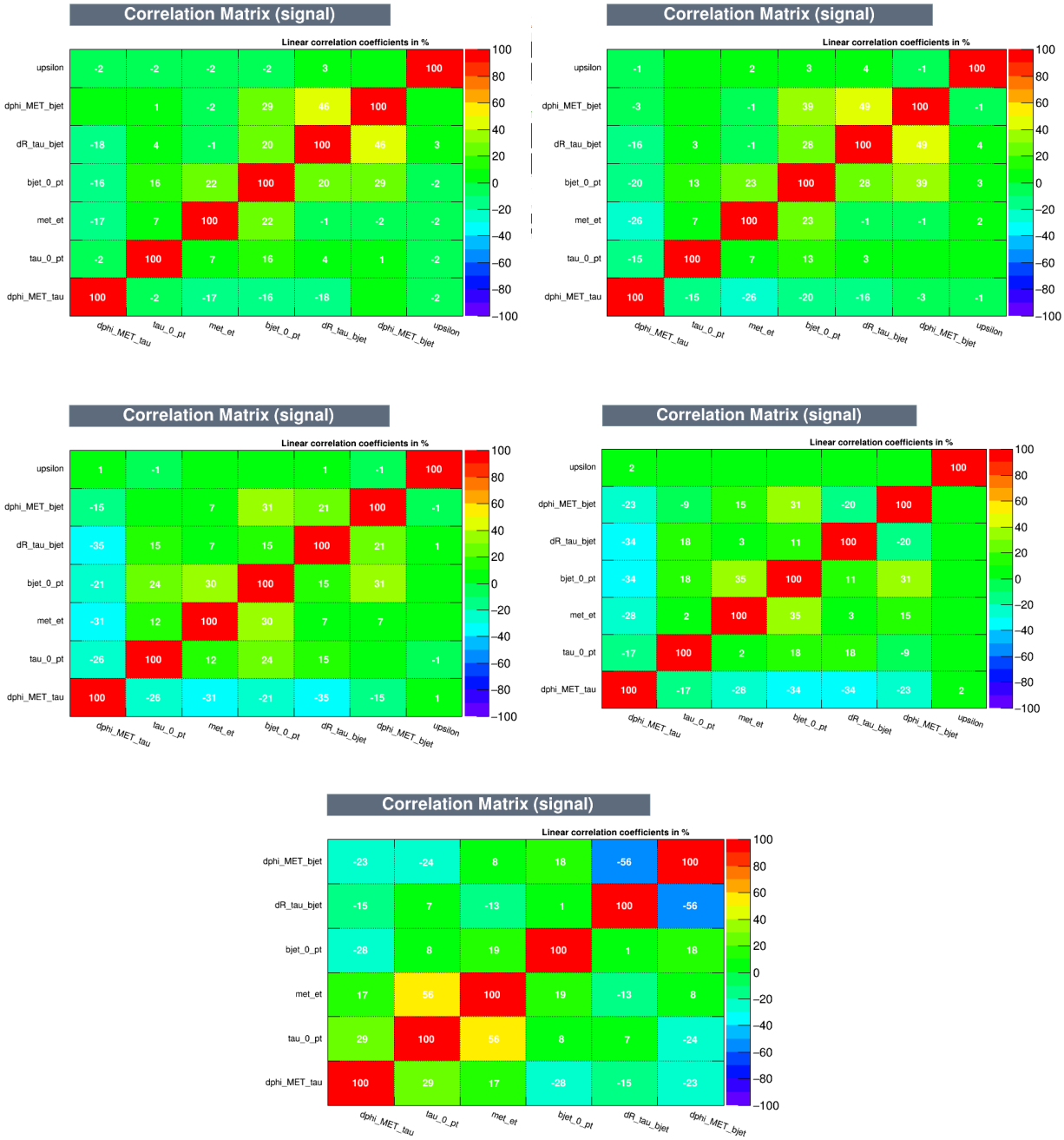



Figure 5.10: Correlation matrix of BDT input variables for the signal. The five $H^{\pm}$ mass ranges used in the BDT training are presented, 90–120 GeV (top left), 130–160 GeV (top right), 160–180 GeV (middle left), 200–400 GeV (middle right) and 500–2000 GeV (bottom). Where: upsilon= $\Upsilon$; dphi_MET_bjet= $\Delta\phi_{b\text{-jet},E_T^{miss}}$; dR_tau_bjet= $\Delta R_{b\text{-jet},\tau}$; bjet_0_pt= $p_T^{b\text{-jet}}$; met_et= $E_T^{miss}$; tau_0_pt= $p_T^{\tau}$ and dphi_MET_tau= $\Delta\phi_{\tau,E_T^{miss}}$.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Upsilon$ | 0.210 |
| 2 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.190 |
| 3 | $p_T^{\tau}$ | 0.140 |
| 4 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.122 |
| 5 | $p_T^{b\text{-jet}}$ | 0.115 |
| 6 | $E_T^{miss}$ | 0.113 |
| 7 | $\Delta R_{\tau,b\text{-jet}}$ | 0.110 |

Table 5.1: Ranking of variables used in the BDT training for the $90 \leq m_{H^{\pm}} \leq 120$ GeV mass range. The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.276 |
| 2 | $\Upsilon$ | 0.180 |
| 3 | $p_T^{\tau}$ | 0.125 |
| 4 | $p_T^{b\text{-jet}}$ | 0.118 |
| 5 | $\Delta R_{\tau,b\text{-jet}}$ | 0.103 |
| 6 | $E_T^{miss}$ | 0.100 |
| 7 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.099 |

Table 5.2: Ranking of variables used in the BDT training for the $130 \leq m_{H^{\pm}} \leq 160$ GeV mass range. The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.319 |
| 2 | $p_T^{\tau}$ | 0.167 |
| 3 | $\Upsilon$ | 0.165 |
| 4 | $\Delta R_{\tau,b\text{-jet}}$ | 0.098 |
| 5 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.098 |
| 6 | $E_T^{miss}$ | 0.091 |
| 7 | $p_T^{b\text{-jet}}$ | 0.063 |

Table 5.3: Ranking of variables used in the BDT training for the $160 \leq m_{H^{\pm}} \leq 180$ GeV mass range. The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.385 |
| 2 | $p_T^{\tau}$ | 0.234 |
| 3 | $\Delta R_{\tau,b\text{-jet}}$ | 0.098 |
| 4 | $\Upsilon$ | 0.093 |
| 5 | $E_T^{miss}$ | 0.090 |
| 6 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.052 |
| 7 | $p_T^{b\text{-jet}}$ | 0.049 |

Table 5.4: Ranking of variables used in the BDT training for the $200 \leq m_{H^{\pm}} \leq 400$ GeV mass range. The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $p_T^{\tau}$ | 0.522 |
| 2 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.276 |
| 3 | $E_T^{miss}$ | 0.122 |
| 4 | $\Delta R_{\tau,b\text{-jet}}$ | 0.030 |
| 5 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.029 |
| 6 | $p_T^{b\text{-jet}}$ | 0.022 |

Table 5.5: Ranking of variables used in the BDT training for the $500 \leq m_{H^{\pm}} \leq 2000$ GeV mass range. The top variable is best ranked.

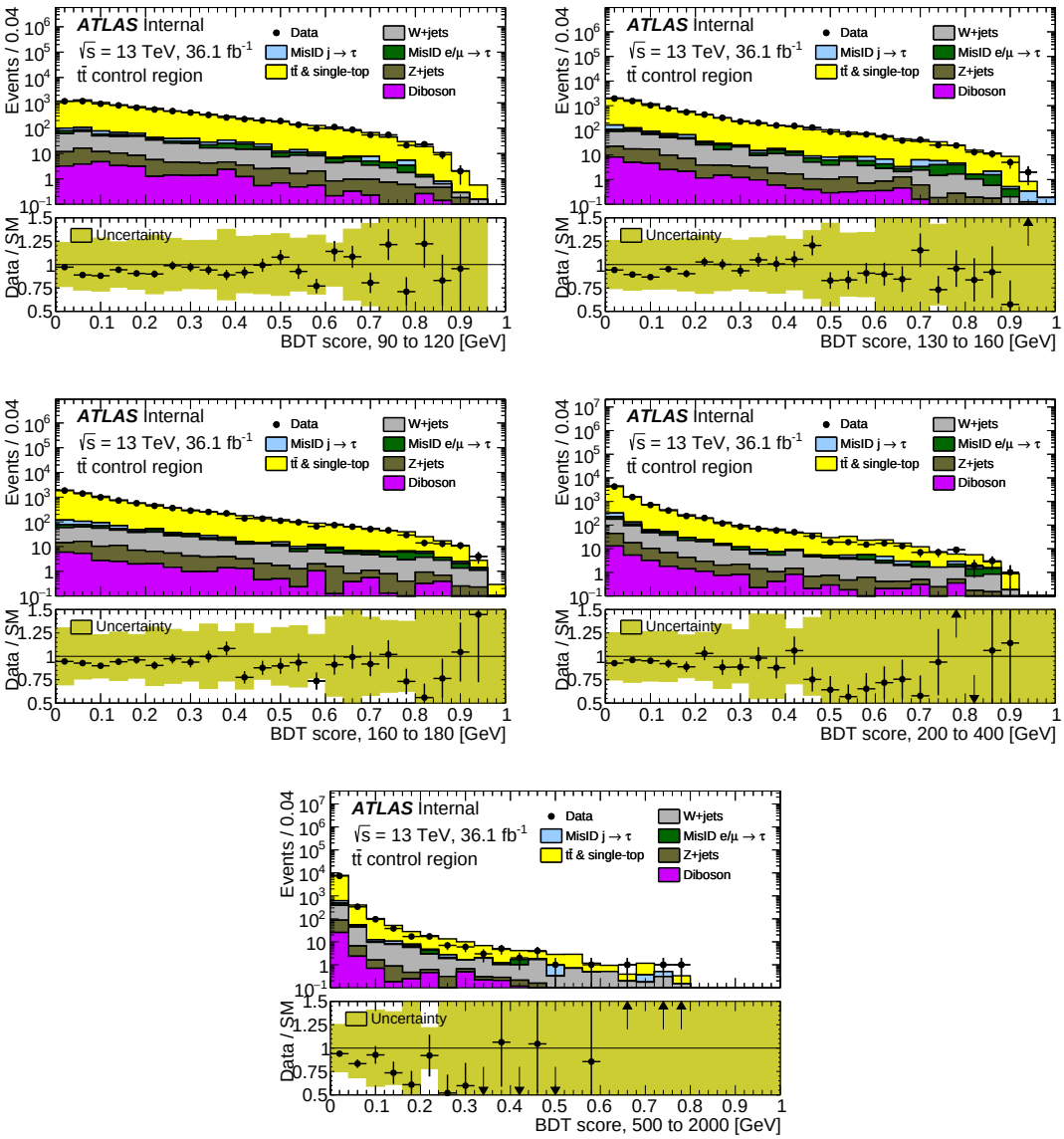



Figure 5.11: Predicted and measured BDT score distributions in the $t\bar{t}$ enriched control region. Shown are five $H^{\pm}$ mass range trainings. The $j \to \tau$ background is estimated using the FF method. The uncertainty bands in the ratio plots include both the statistical and systematic uncertainties of simulated events, added in quadrature. From [33].

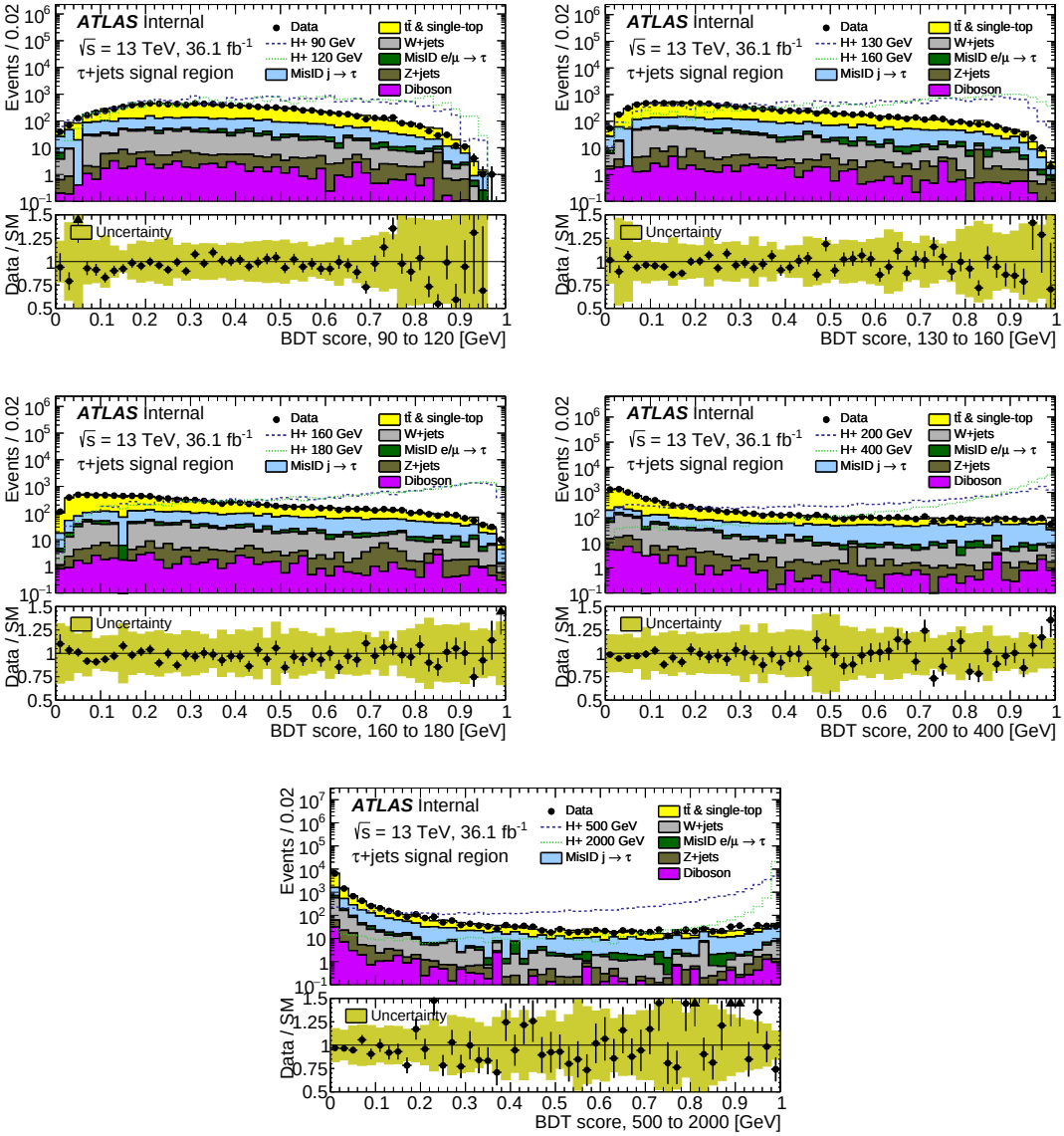



Figure 5.12: BDT score distributions in signal region. Shown are five $H^{\pm}$ mass range trainings. The $j \to \tau$ background is estimated using the FF method. The uncertainty bands in the ratio plots include both the statistical and systematic uncertainties of simulated events, added in quadrature. The normalisation of the signal samples corresponds to the integral of the background. From [33].

## 5.3 Studies of BDT tunning for full Run-2 data $H^{\pm}$ searches

The BDT used for the published result [12] is optimised for the limited number of $H^{\pm}$ mass points and based on comparison of the signal and background separations after trainings. The goal of all tests presented in this Section is to study possible improvements in performance of the BDT presented in previous Section in application to the full Run-2 dataset. The studies are done with the following BDT setup:

- for 1-prong and 3-prong: NTrees= 300; Shrinkage= 0.05; RandRatio= 0.5; NTreeLayers= 3; NCutLevel= 8,

which has been used before the final optimisation of the hyperparameters is performed for the BDT applied in the analysis. In order to quantify the outcome of the tests two comparisons are made. First, the separation between BDT outputs for the signal and the full background model is investigated based on Kolmogorov-Smirnov test (KS) [134]. Second, final comparison is based on setting the expected limits on $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$. The limits stem from the profile likelihood ratio fit for the background-only and the background+signal hypoteses. In order to get limits presented in this Section, the fits are performed only in the SR. No systematic uncertainties are considered.

### 5.3.1 Comparison with the direct use of $m_T$

In the *default* BDT ("default", i.e. as in published result [12], see Sec. 5.2) instead of $m_T$ variable, its components: $\Delta\phi_{\tau,E_T^{miss}}$, $p_T$ and $E_T^{miss}$ are used. Such an approach, using more basic variables in place of complex ones, can be problematic if statistics of the samples used for MVA training are not sufficiently large. To study this, an alternative BDT is trained using directly $m_T$ variable. The KS tests obtained for default, $KS_{\mathrm{default}}$, and alternative, $KS_{\mathrm{altBDT}}$, BDT output (Kolmogorov distance) give respectively:

- for $m_{H^{\pm}}$ equal 500 GeV - $KS_{\mathrm{default}} = 0.8174$ and $KS_{\mathrm{altBDT}} = 0.8229$,
- for $m_{H^{\pm}}$ equal 1000 GeV - $KS_{\mathrm{default}} = 0.9167$ and $KS_{\mathrm{altBDT}} = 0.9199$,
- for $m_{H^{\pm}}$ equal 1800 GeV - $KS_{\mathrm{default}} = 0.9505$ and $KS_{\mathrm{altBDT}} = 0.9541$.

For all mass points only small increase of the signal-background separation is observed.

Similarly, a moderate improvement on the expected limits is found, as presented in Fig. 5.13(a). Larger gain by using directly $m_T$ as an input variable is observed for high $H^{\pm}$ masses, where $m_T$ is the highest ranked variable (see Tables 5.1 - 5.5 and 5.6 - 5.10 ), and statistics of the background sample used for training is low. However, with higher statistics the

*default* settings are expected to give stronger limit, especially that $m_T$, $p_T^\tau$ and $E_T^{miss}$ are highly correlated, see Fig. 5.15.

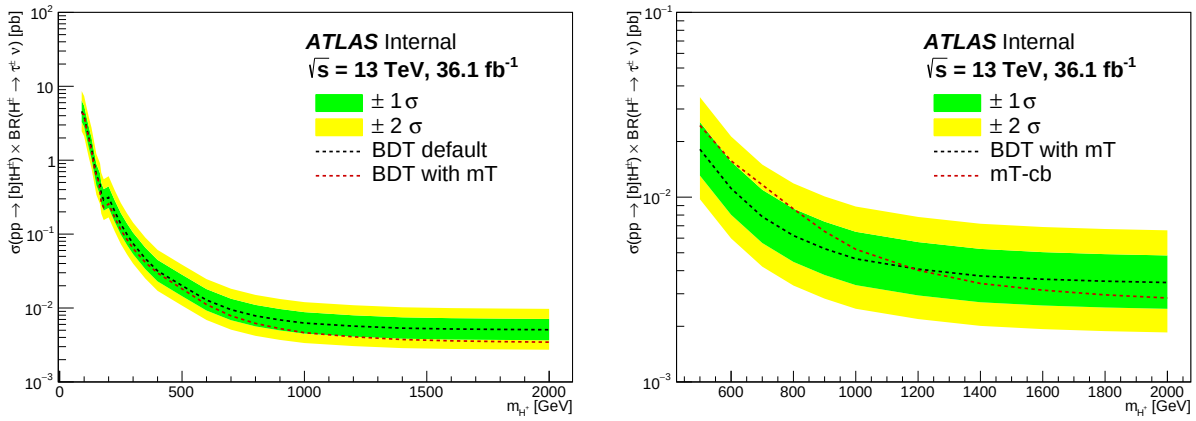
Signal and background events distributions of $\Delta\phi_{\tau,E_T^{miss}}$ used for *default* BDT-training and $m_T$ used in place of $\Delta\phi_{\tau,E_T^{miss}}$ in *alternative* BDT training, for the $500 \leq m_{H^\pm} \leq 2000$ GeV mass bin, are presented in Fig. 5.14. Good separation between signal and background can be observed for both cases.

### 5.3.2 Impact of using FFs in the BDT training

In the presented analysis for the first four $H^\pm$ mass ranges, events arising from $\mathrm{j} \to \tau$ fakes, taken from the simulated background samples, are included in the training. As it is presented above, the most important discriminanting variable in high $H^\pm$ masses is $m_T$. Its distribution for misidentified $\tau_{\text{had-vis}}$ candidates, as taken from top MC samples and as estimated from data using FFs after SR selection, is shown in Fig. 5.16. High $m_T$ region is populated almost only by fake $\tau_{\text{had-vis}}$ candidates coming mainly from QCD processes and estimated with the FF method. Hence, in the $H^\pm$ mass range $500-2000$ GeV, the misidentified $\tau_{\text{had-vis}}$ candidates determined with the data-driven method have to be included in the BDT training. In Fig. 5.17, distributions of BDT score for training in the high mass range, for events with the misidentified $\tau_{\text{had-vis}}$ candidates as estimated with FF method and as taken from the simulated top background samples are presented. The only MC top backgrounds are used for training, an enhancement around BDT score $\sim 1$ with signal-like background events, is present, see 5.17 (b). This effect is due to the absence of multi-jet MC events for high $m_T$ in the training. This causes the BDT algorithm to "learn" that high $m_T$ events are signal-like. This enhancement of signal-like background events is reduced by introducing fake $\tau$ contribution to the background estimated from data with the FF method in the training, see Fig. 5.17 (a), since it provides realistic background modelling in the high $m_T$ region. This should lead to stronger final limits, as it prepares the algorithm for encountering such signal-like events in the real background. However, it is important to note that a definitive conclusion on the impact of including in the BDT training background events estimated with FF method can be achieved only with a large statistics and therefore requires full Run-2 data. With limited statistics, addition of rare signal-like background events at the training stage makes it harder for the BDT to construct well performing decision tree and can negatively impact the classification output. Nevertheless, in the high mass region the physically correct approach is indeed to include the dominant background from multi-jets and therefore use the FF method and this solution is chosen for the analysis.

### 5.3.3 Comparison with cut-based analysis

Limited statistics of the background events at high $H^{\pm}$ masses above 1 TeV makes it problematic for the BDT algorithm to perform better then the cut-based analysis with $m_T$ as the final discriminating variable [133]. This effect can be seen in Fig. 5.13 (b), where expected limits for BDT training with $m_T$ variable are compared to the limits obtained with the cut-based analysis. Given sufficiently large statistics, the BDT should be at least as sensitive as the cut-based approach.




$(a)$ $(b)$

Figure 5.13: Expected 95% CL exclusion limits on $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ for charged Higgs boson production as a function of $m_{H^{\pm}}$ in 36.1 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 13$ TeV for (a) *default* BDT output and *alternative* BDT output where $m_T$ in place of $\Delta\phi_{\tau,E_T^{miss}}$ is used as an input variable; (b) BDT output, where $m_T$ in place of $\Delta\phi_{\tau,E_T^{miss}}$ is used as an input variable and cut based analysis limit based on $m_T$ variable, $m_T$-cb.

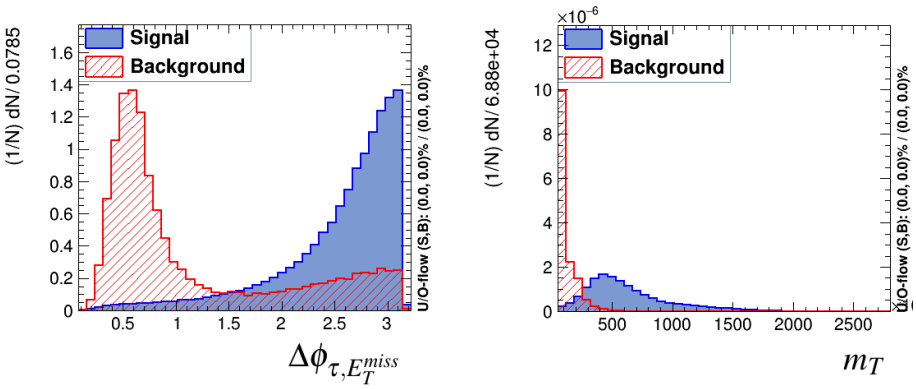



Figure 5.14: Distributions of $\Delta\phi_{\tau,E_T^{miss}}$ used for *default* BDT-training and $m_T$ used in place of $\Delta\phi_{\tau,E_T^{miss}}$ in *alternative* BDT training for the $500 \leq m_{H^{\pm}} \leq 2000$ GeV mass range. Signal is presented as a blue, solid histogram; top background (as estimated from MC) together with j $\to \tau$ background (as estimated with FFs) are presented as a red, dashed histogram. Histograms are arbitrarily normalised.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Upsilon$ | 0.281 |
| 2 | $m_T$ | 0.268 |
| 3 | $p_T^{\tau}$ | 0.112 |
| 4 | $p_T^{b\text{-jet}}$ | 0.101 |
| 5 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.083 |
| 6 | $E_T^{miss}$ | 0.08 |
| 7 | $\Delta R_{\tau,b\text{-jet}}$ | 0.076 |

Table 5.6: Ranking of variables used in an *alternative* BDT training for the $90 \leq m_{H^{\pm}} \leq 120$ GeV mass range, where $m_T$ in place of $\Delta\phi_{\tau,E_T^{miss}}$ is used (see Table 5.1 for comparison). The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $m_T$ | 0.379 |
| 2 | $\Upsilon$ | 0.222 |
| 3 | $p_T^{b\text{-jet}}$ | 0.116 |
| 4 | $\Delta R_{\tau,b\text{-jet}}$ | 0.085 |
| 5 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.08 |
| 6 | $p_T^{\tau}$ | 0.071 |
| 7 | $E_T^{miss}$ | 0.048 |

Table 5.7: Ranking of variables used in an *alternative* BDT training for the $130 \leq m_{H^{\pm}} \leq 160$ GeV mass range, where $m_T$ in place of $\Delta\phi_{\tau,E_T^{miss}}$ is used (see Table 5.2 for comparison). The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $m_T$ | 0.395 |
| 2 | $\Upsilon$ | 0.173 |
| 3 | $\Delta R_{\tau,b\text{-jet}}$ | 0.121 |
| 4 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.120 |
| 5 | $p_T^{\tau}$ | 0.095 |
| 6 | $E_T^{miss}$ | 0.053 |
| 7 | $p_T^{b\text{-jet}}$ | 0.045 |

Table 5.8: Ranking of variables used in an *alternative* BDT training for the $160 \leq m_{H^{\pm}} \leq 180$ GeV mass range, where $m_T$ in place of $\Delta\phi_{\tau,E_T^{miss}}$ is used (see Table 5.3 for comparison). The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $m_T$ | 0.524 |
| 2 | $p_T^{\tau}$ | 0.123 |
| 3 | $\Delta R_{\tau,b\text{-jet}}$ | 0.104 |
| 4 | $\Upsilon$ | 0.093 |
| 5 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.064 |
| 6 | $E_T^{miss}$ | 0.049 |
| 7 | $p_T^{b\text{-jet}}$ | 0.043 |

Table 5.9: Ranking of variables used in an *alternative* BDT training for the $200 \leq m_{H^{\pm}} \leq 400$ GeV mass range, where $m_T$ in place of $\Delta\phi_{\tau,E_T^{miss}}$ is used (see Table 5.4 for comparison). The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $m_T$ | 0.766 |
| 2 | $p_T^{\tau}$ | 0.102 |
| 3 | $E_T^{miss}$ | 0.068 |
| 4 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.026 |
| 5 | $\Delta R_{\tau,b\text{-jet}}$ | 0.023 |
| 6 | $p_T^{b\text{-jet}}$ | 0.015 |

Table 5.10: Ranking of variables used in an *alternative* BDT training for the $500 \leq m_{H^{\pm}} \leq 2000$ GeV mass range, where $m_T$ in place of $\Delta\phi_{\tau,E_T^{miss}}$ is used (see Table 5.5 for comparison). The top variable is best ranked.

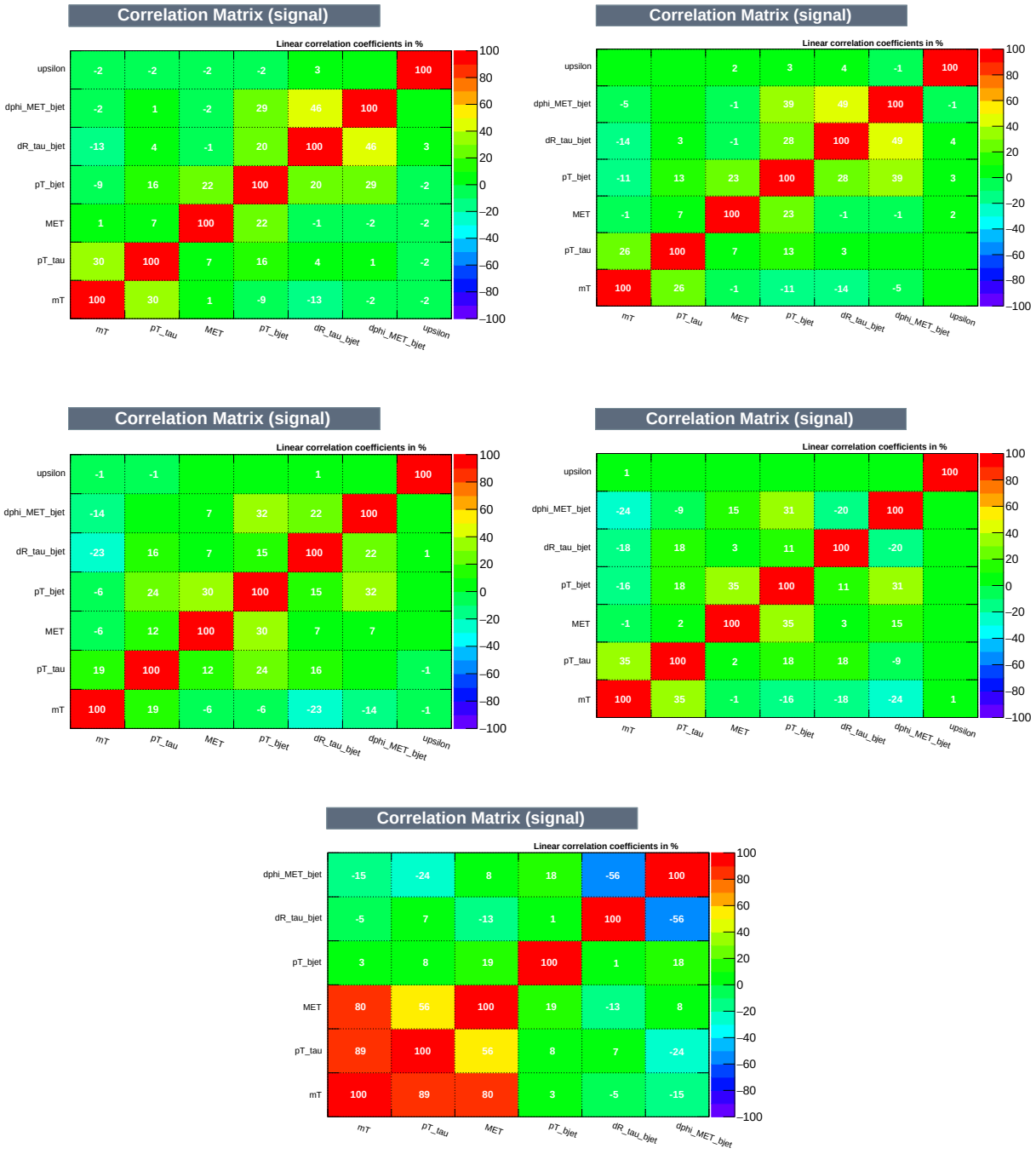



Figure 5.15: Correlation matrix of *alternative* BDT input variables, where $m_T$ in place of $\Delta\phi_{\tau,E_T^{miss}}$ is used (see Fig. 5.10 for comparison), for the signal. The five $H^{\pm}$ mass ranges used in the BDT training are presented, 90–120 GeV (top left), 130–160 GeV (top right), 160–180 GeV (middle left), 200–400 GeV (middle right) and 500–2000 GeV (bottom). Where: upsilon= $\Upsilon$; dphi_MET_bjet= $\Delta\phi_{b\text{-jet},E_T^{miss}}$; dR_tau_bjet= $\Delta R_{b\text{-jet},\tau}$; pT_bjet= $p_T^{b\text{-jet}}$; MET= $E_T^{miss}$; pT_tau= $p_T^{\tau}$ and mT= $m_T$.

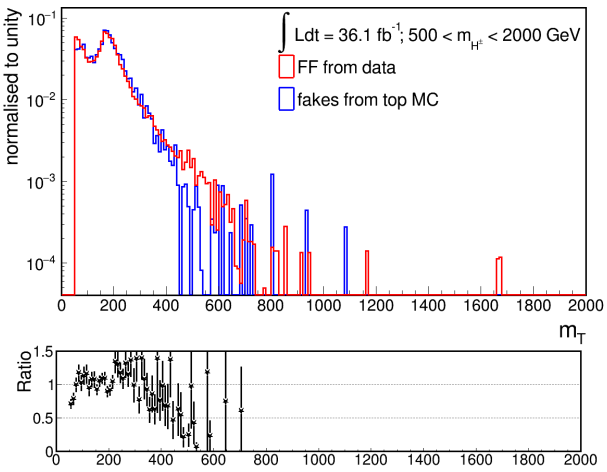
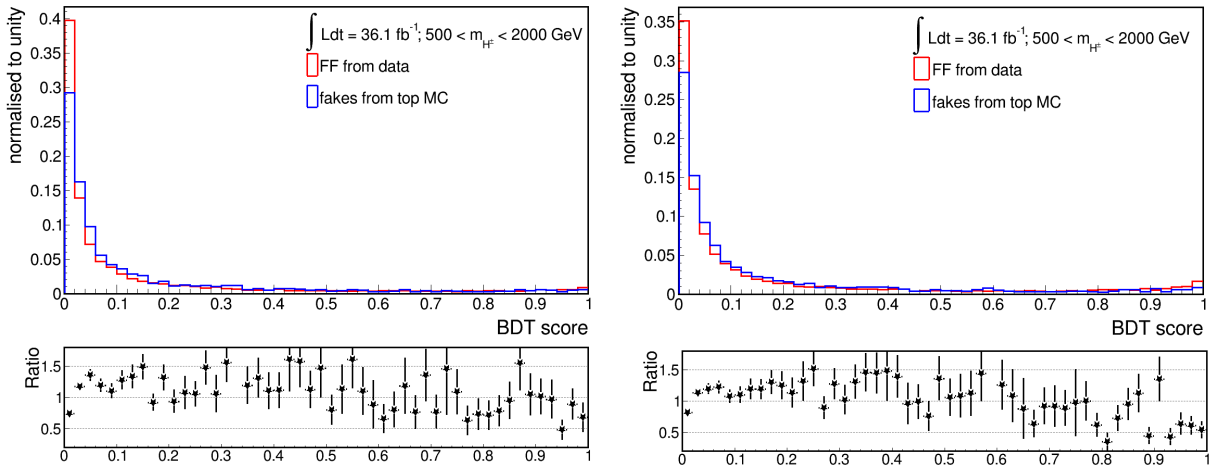



Figure 5.16: Distributions of $m_T$ for misidentified $\tau_{\text{had-vis}}$ candidates as taken from top MC samples and as estimated with the FF method, after SR selection. Histograms are normalised to unit area.




$(a)$ $(b)$

Figure 5.17: Juxtaposition of BDT score, for training in the mass range 500–2000 GeV, for events with the misidentified $\tau_{\text{had-vis}}$ candidates as estimated with the FF method and as taken from the simulated top background samples for two training cases: (a) in the training the misidentified $\tau_{\text{had-vis}}$ candidates estimated with the FF method are included, (b) in the training events the misidentified $\tau_{\text{had-vis}}$ candidates are taken from the simulated top background samples are included.

### 5.3.4 Optimisation of variable sets and mass ranges

As presented in Sec. 5.2, in the current analysis seven variables in total are used as an input to the BDT. Fig. 5.18(a) shows the difference in the expected limits when using a set of $6/7$ (*default set*) or $8/9$ variables for events with 1-prong/3-prong $\tau_{\text{had-vis}}$, respectively. Where two variables were added in the trainings with respect to the *default* set:

- $\Delta p_T^{b\text{-jet,jet}}$ - difference in the transverse momentum between the reconstructed leading (in terms of $p_T$) $b$-jet and leading light-jet (non $b$-tagged),
- $p_T^{\text{sublight-jet}}$ - the transverse momentum of the next-to-leading (in terms of $p_T$) non $b$-tagged jet.

The difference between these two BDTs is found to be almost negligible. Thus, taking into consideration also the fact that two more variables cause more systematic uncertainties, the set of $6/7$ variables is chosen in the analysis. Note that for those studies kfold$=2$, thus lower statistics of signal samples for the training is used.

After further optimisation of the BDT algorithm and application of kfold$=5$, which resulted in high signal statistics delivered to the training, it is found that the quality of the separation coming from the set of $8/9$ variables improved limit above the $H^{\pm}$ mass of 180 GeV, see Fig. 5.18(b). As can be seen from Tables 5.11 - 5.15, two added variables seem to play less relevant role in separation. Nevertheless, they turned out to be important for the total performance of the BDT discrimination power. However, note that still the difference in results between in total 7 or 9 variables used as an input to BDT training is within one sigma of its statistical uncertainty. This is consistent with the results obtained from KS tests for *default* and *alternative* BDT output:

- for $m_{H^{\pm}}$ equal 90 GeV - $KS_{\text{default}} = 0.2575$ and $KS_{\text{altBDT}} = 0.2547$,
- for $m_{H^{\pm}}$ equal 200 GeV - $KS_{\text{default}} = 0.5275$ and $KS_{\text{altBDT}} = 0.5500$,
- for $m_{H^{\pm}}$ equal 1200 GeV - $KS_{\text{default}} = 0.9280$ and $KS_{\text{altBDT}} = 0.9338$.

Next test concerned $H^{\pm}$ mass bins ranges used for training. In Fig. 5.19 the expected limits for two cases are presented:

- (a) comparison of outputs of the *default* and *alternative* BDT with $\Upsilon$ variable used in the training for charged Higgs masses in the range $90-225$ GeV ( with respect to default $90-400$ GeV). Mass ranges for trainings events with 3-prong $\tau_{\text{had-vis}}$ candidates without $\Upsilon$ are also respectively changed. The comparison shows that the expected limits

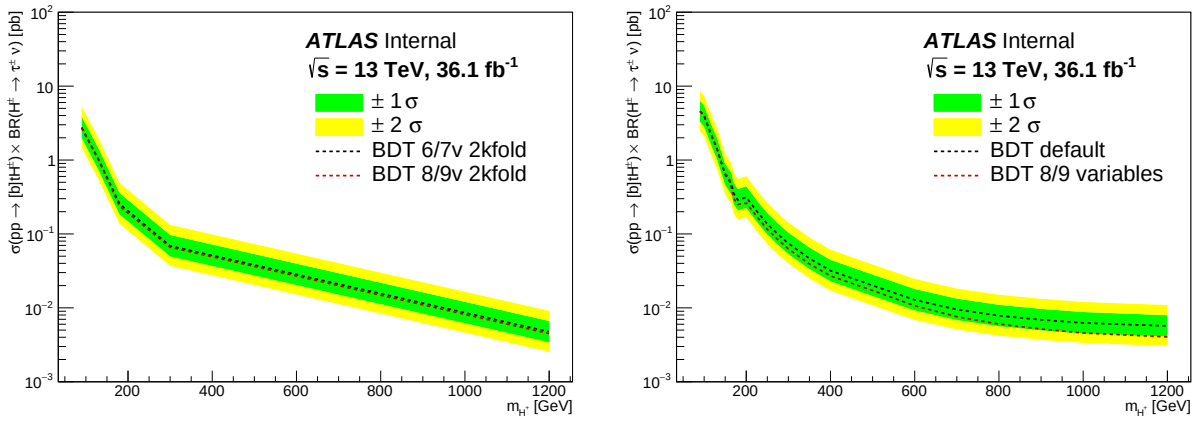
$\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ obtained using both BDTs are within one sigma of its statistical uncertainty. This is consistent with the results obtained from KS tests for *default* and the *alternative* BDT output:

  - for $m_{H^{\pm}}$ equal 180 GeV - $KS_{\mathrm{default}} = 0.5195$ and $KS_{\mathrm{altBDT}} = 0.5242$,
  - for $m_{H^{\pm}}$ equal 200 GeV - $KS_{\mathrm{default}} = 0.5275$ and $KS_{\mathrm{altBDT}} = 0.5628$,
  - for $m_{H^{\pm}}$ equal 250 GeV - $KS_{\mathrm{default}} = 0.6298$ and $KS_{\mathrm{altBDT}} = 0.6042$,
  - for $m_{H^{\pm}}$ equal 350 GeV - $KS_{\mathrm{default}} = 0.7348$ and $KS_{\mathrm{altBDT}} = 0.7263$,

- (b) comparison of *default* BDT output with an *alternative* BDT with $\Upsilon$ variable used in the training for $H^{\pm}$ masses in the range $90-400$ GeV divided into two mass bins: $90-120$, $130-400$. Mass ranges for trainings events with 3-prong $\tau_{\text{had-vis}}$ candidates without $\Upsilon$ are also respectively changed. The comparison shows that the expected limits $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ obtained with using both BDTs are within one sigma of its statistical uncertainty. This is consistent with the results obtained from KS tests for *default* and the *alternative* BDT output:

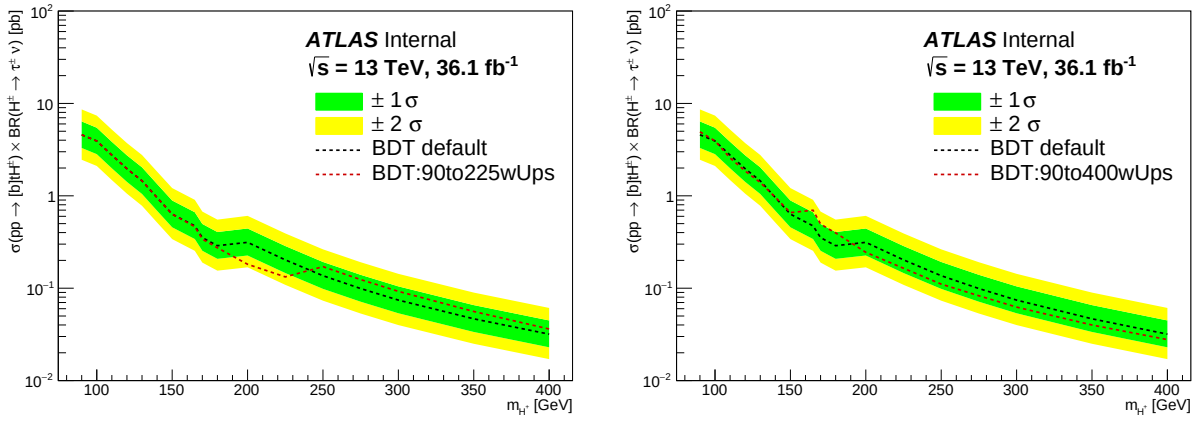
  - for $m_{H^{\pm}}$ equal 165 GeV - $KS_{\mathrm{default}} = 0.4939$ and $KS_{\mathrm{altBDT}} = 0.4482$,
  - for $m_{H^{\pm}}$ equal 180 GeV - $KS_{\mathrm{default}} = 0.5195$ and $KS_{\mathrm{altBDT}} = 0.4908$,
  - for $m_{H^{\pm}}$ equal 200 GeV - $KS_{\mathrm{default}} = 0.5275$ and $KS_{\mathrm{altBDT}} = 0.5430$,

Thus, when full Run-2 dataset will be available it is worthwhile to revisit optimisation of the mass bins division since there is a window here to slightly improve the performance of the BDT.




$(a)$ $(b)$

Figure 5.18: Juxtaposition of expected 95% CL exclusion limits on $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ for charged Higgs boson production as a function of $m_{H^{\pm}}$ in 36.1 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 13$ TeV for two cases: (a) BDT output built by set of $6/7$ input variables and BDT built by set of $8/9$ input variables, with k-fold$=2$ training method. Here, only 5 mass points: 90, 130, 180, 300 and 1200 GeV are used for the limits study; (b) *default* BDT output and an *alternative* BDT built by set of $8/9$ input variables with kfold$=5$ used, and all available mass points used for the limits study.




$(b)$ $(c)$

Figure 5.19: Juxtaposition of expected 95% CL exclusion limits on $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ for two cases: (a) *default* BDT output and an *alternative* BDT where in place of using $\Upsilon$ in training for $H^{\pm}$ masses in the range 90-400 GeV it is used for $H^{\pm}$ masses in the range 90-225 GeV for events where $\tau_{\text{had-vis}}$ has only one associated track; (b) *default* BDT output and an *alternative* BDT where $\Upsilon$ is used in the training for events where $\tau_{\text{had-vis}}$ has only one associated track, for $H^{\pm}$ masses in the range 90-400 GeV divided into two mass bins: 90-120 and 130-400 GeV.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Upsilon$ | 0.252 |
| 2 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.193 |
| 3 | $p_T^{\tau}$ | 0.123 |
| 4 | $p_T^{b\text{-jet}}$ | 0.088 |
| 5 | $E_T^{miss}$ | 0.08 |
| 6 | $\Delta R_{\tau,b\text{-jet}}$ | 0.074 |
| 7 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.068 |
| 8 | $\Delta p_T^{b\text{-jet},\,jet}$ | 0.064 |
| 9 | $p_T^{sublight\text{-jet}}$ | 0.059 |

Table 5.11: Ranking of 9 variables used in an *alternative* BDT training for the $90 \leq m_{H^{\pm}} \leq 120$ GeV mass range. The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.304 |
| 2 | $\Upsilon$ | 0.200 |
| 3 | $p_T^{\tau}$ | 0.122 |
| 4 | $p_T^{b\text{-jet}}$ | 0.100 |
| 5 | $E_T^{miss}$ | 0.079 |
| 6 | $\Delta R_{\tau,b\text{-jet}}$ | 0.067 |
| 7 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.065 |
| 8 | $\Delta p_T^{b\text{-jet},\,jet}$ | 0.034 |
| 9 | $p_T^{sublight\text{-jet}}$ | 0.028 |

Table 5.12: Ranking of 9 variables used in an *alternative* BDT training for the $130 \leq m_{H^{\pm}} \leq 160$ GeV mass range. The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.326 |
| 2 | $p_T^{\tau}$ | 0.16 |
| 3 | $\Upsilon$ | 0.155 |
| 4 | $E_T^{miss}$ | 0.093 |
| 5 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.081 |
| 6 | $\Delta R_{\tau,b\text{-jet}}$ | 0.068 |
| 7 | $p_T^{sublight\text{-jet}}$ | 0.054 |
| 8 | $\Delta p_T^{b\text{-jet},\,jet}$ | 0.034 |
| 9 | $p_T^{b\text{-jet}}$ | 0.03 |

Table 5.13: Ranking of 9 variables used in an *alternative* BDT training for the $160 \leq m_{H^{\pm}} \leq 180$ GeV mass range. The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.387 |
| 2 | $p_T^{\tau}$ | 0.244 |
| 3 | $E_T^{miss}$ | 0.092 |
| 4 | $\Upsilon$ | 0.087 |
| 5 | $\Delta R_{\tau,b\text{-jet}}$ | 0.083 |
| 6 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.039 |
| 7 | $p_T^{b\text{-jet}}$ | 0.034 |
| 8 | $\Delta p_T^{b\text{-jet},\,jet}$ | 0.03 |
| 9 | $p_T^{sublight\text{-jet}}$ | 0.004 |

Table 5.14: Ranking of 9 variables used in an *alternative* BDT training for the $200 \leq m_{H^{\pm}} \leq 400$ GeV mass range. The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $p_T^{\tau}$ | 0.562 |
| 2 | $\Delta\phi_{\tau, E_T^{miss}}$ | 0.26 |
| 3 | $E_T^{miss}$ | 0.122 |
| 4 | $\Delta\phi_{b\text{-jet}, E_T^{miss}}$ | 0.018 |
| 5 | $\Delta R_{\tau, b\text{-jet}}$ | 0.016 |
| 6 | $p_T^{b\text{-jet}}$ | 0.013 |
| 7 | $\Delta p_T^{b\text{-jet}, jet}$ | 0.007 |
| 8 | $p_T^{sublight\text{-jet}}$ | 0.002 |

Table 5.15: Ranking of 8 variables used in an *alternative* BDT training for the $500 \leq m_{H^{\pm}} \leq 2000$ GeV mass range. The top variable is best ranked.

### 5.3.5 Study of the impact of $\Upsilon$ variable and that of its components

As it is discussed in Chapter 4, at low $H^{\pm}$ masses the kinematics of the products of $t \to bH^{\pm}$ and $t \to bW^{\pm}$ decays can be very similar. Therefore, the $\tau$ lepton polarisation sensitive variable, $\Upsilon$, can serve as a very powerful discriminating variable (see also Tables 5.1 - 5.3).

It is also investigated whether using $\Upsilon$ as an input variable to the BDT discriminant only for events with $\tau_{\text{had-vis}}$ decays to one charged and one neutral meson gives comparable results to that obtained taking into account all 1-prong decays, see Fig. 5.20(a). Obtained expected limits on $\sigma(pp \to [b]tH^{\pm}) \times \text{BR}(H^{\pm} \to \tau^{\pm}\nu)$ show that the results are within one sigma of its statistical uncertainty, which is consistent with the outcome obtained from KS tests for *default* and the *alternative* BDT output:

- for $m_{H^{\pm}}$ equal 90 GeV - $KS_{\text{default}} = 0.2575$ and $KS_{\text{altBDT}} = 0.2503$,
- for $m_{H^{\pm}}$ equal 130 GeV - $KS_{\text{default}} = 0.3850$ and $KS_{\text{altBDT}} = 0.3983$,
- for $m_{H^{\pm}}$ equal 200 GeV - $KS_{\text{default}} = 0.5275$ and $KS_{\text{altBDT}} = 0.5394$.

Whereas, in Fig. 5.20(b)-(d) and Tables 5.16 - 5.19 the results of using $p_T^{\tau\text{-track}}$ in place of the $\Upsilon$ for $H^{\pm}$ masses in the range $90-400$ GeV in the final BDT discriminant for events where $\tau_{\text{had-vis}}$ has only one associated track, can be found. The results indicate slightly better performance of the BDT discriminant achieved by using directly $\Upsilon$ variable not just its components (see Eq. 5.9) as an input variable in trainings in $H^{\pm}$ mass range $90-160$ GeV. For the current analysis, see Sec. 5.2, the direct usage of $\Upsilon$ observable is chosen. This is especially relevant for low $H^{\pm}$ mass region. However, since the difference between the power of discrimination for $\Upsilon$ and its components is indeed small, it seems useful to study it again on full Run-2 data, taking also into account the modelling of the variables and systematic uncertainties connected with them[4]. The comparison shows that the results are within one sigma of its statistical uncertainty, which is consistent with the outcome obtained from KS tests for *default* and the *alternative* BDT output:

- for $m_{H^{\pm}}$ equal 90 GeV - $KS_{\text{default}} = 0.2575$ and $KS_{\text{altBDT}} = 0.2220$,
- for $m_{H^{\pm}}$ equal 130 GeV - $KS_{\text{default}} = 0.3850$ and $KS_{\text{altBDT}} = 0.3711$,
- for $m_{H^{\pm}}$ equal 165 GeV - $KS_{\text{default}} = 0.4939$ and $KS_{\text{altBDT}} = 0.5013$,
- for $m_{H^{\pm}}$ equal 200 GeV - $KS_{\text{default}} = 0.5275$ and $KS_{\text{altBDT}} = 0.5335$,
- for $m_{H^{\pm}}$ equal 350 GeV - $KS_{\text{default}} = 0.7348$ and $KS_{\text{altBDT}} = 0.7356$.

[4]Preliminary studies the Author performed on Run-1 data suggest that also the cut-based analysis can potentially be made more sensitive by using $\Upsilon$ as a final discriminating variable.

The correlation of $p_T^{\tau\text{-track}}$ is rather mild with other variables as presented in Fig. 5.21 - 5.22, which corresponds to correlation of $\Upsilon$ variable presented in Fig. 5.9- 5.10.

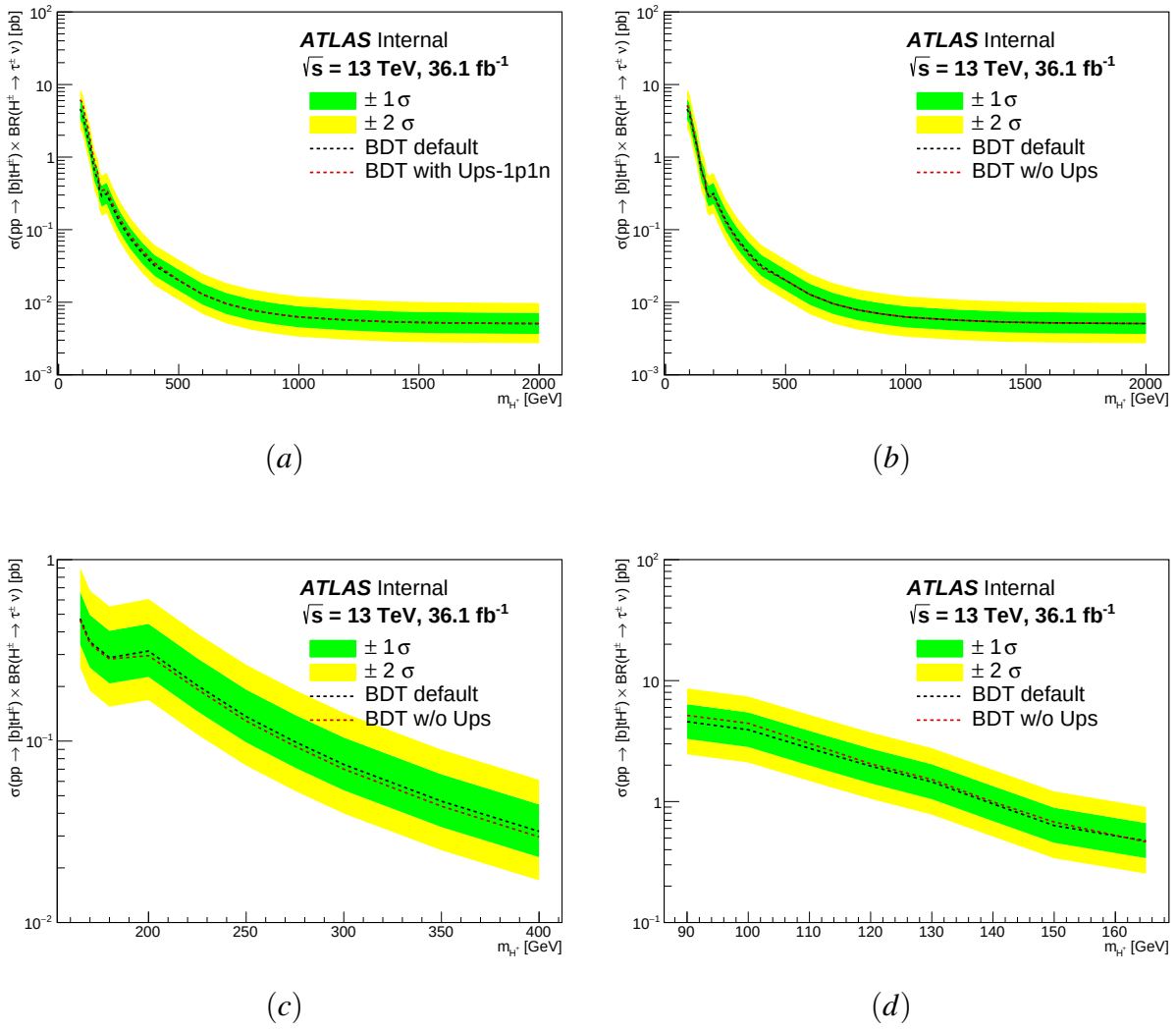



Figure 5.20: Juxtaposition of expected 95% CL exclusion limits on $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ for charged Higgs boson production as a function of $m_{H^{\pm}}$ in 36.1 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 13$ TeV for two cases: (a) *default* BDT output and an *alternative* BDT where $\Upsilon$ for $H^{\pm}$ masses in the range $90-400$ GeV is used as an input variable in the final BDT discriminant for events where $\tau_{\text{had-vis}}$ decays only on one charged and one neutral meson (BDT with Ups-1p1n); (b)-(d) *default* BDT output and an *alternative* BDT where $p_T^{\tau\text{-track}}$ in place of using $\Upsilon$ for $H^{\pm}$ masses in the range $90-400$ GeV is used as an input variable in the final BDT discriminant for events where $\tau_{\text{had-vis}}$ has only one associated track (BDT $w/o$ Ups).

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $p_T^{\tau\text{-track}}$ | 0.27 |
| 2 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.232 |
| 3 | $p_T^{\tau}$ | 0.143 |
| 4 | $p_T^{b\text{-jet}}$ | 0.103 |
| 5 | $E_T^{miss}$ | 0.094 |
| 6 | $\Delta R_{\tau,b\text{-jet}}$ | 0.083 |
| 7 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.076 |

Table 5.16: Ranking of variables used in an *alternative* BDT training for the $90 \leq m_{H^{\pm}} \leq 120$ GeV mass range, where $p_T^{\tau\text{-track}}$ in place of $\Upsilon$ is used (see Table 5.1). The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.334 |
| 2 | $p_T^{\tau\text{-track}}$ | 0.212 |
| 3 | $p_T^{b\text{-jet}}$ | 0.120 |
| 4 | $p_T^{\tau}$ | 0.09 |
| 5 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.085 |
| 6 | $E_T^{miss}$ | 0.082 |
| 7 | $\Delta R_{\tau,b\text{-jet}}$ | 0.077 |

Table 5.17: Ranking of variables used in an *alternative* BDT training for the $130 \leq m_{H^{\pm}} \leq 160$ GeV mass range, where $p_T^{\tau\text{-track}}$ in place of $\Upsilon$ is used (see Table 5.2). The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.383 |
| 2 | $p_T^{\tau\text{-track}}$ | 0.175 |
| 3 | $p_T^{\tau}$ | 0.131 |
| 4 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.095 |
| 5 | $\Delta R_{\tau,b\text{-jet}}$ | 0.085 |
| 6 | $E_T^{miss}$ | 0.082 |
| 7 | $p_T^{b\text{-jet}}$ | 0.048 |

Table 5.18: Ranking of variables used in an *alternative* BDT training for the $160 \leq m_{H^{\pm}} \leq 180$ GeV mass range, where $p_T^{\tau\text{-track}}$ in place of $\Upsilon$ is used (see Table 5.3). The top variable is best ranked. The ranking is shown for 1-prong $\tau_{\text{had-vis}}$ candidates.

| Rank | Variable | Variable Importance |
|---|---|---|
| 1 | $\Delta\phi_{\tau,E_T^{miss}}$ | 0.404 |
| 2 | $p_T^{\tau}$ | 0.206 |
| 3 | $p_T^{\tau\text{-track}}$ | 0.118 |
| 4 | $\Delta R_{\tau,b\text{-jet}}$ | 0.093 |
| 5 | $E_T^{miss}$ | 0.090 |
| 6 | $\Delta\phi_{b\text{-jet},E_T^{miss}}$ | 0.046 |
| 7 | $p_T^{b\text{-jet}}$ | 0.042 |

Table 5.19: Ranking of variables used in an *alternative* BDT training for the $200 \leq m_{H^{\pm}} \leq 400$ GeV mass range, where $p_T^{\tau\text{-track}}$ in place of $\Upsilon$ is used (see Table 5.4 for comparison). The top variable is best ranked. The ranking is shown for events with 1-prong $\tau_{\text{had-vis}}$ candidates.

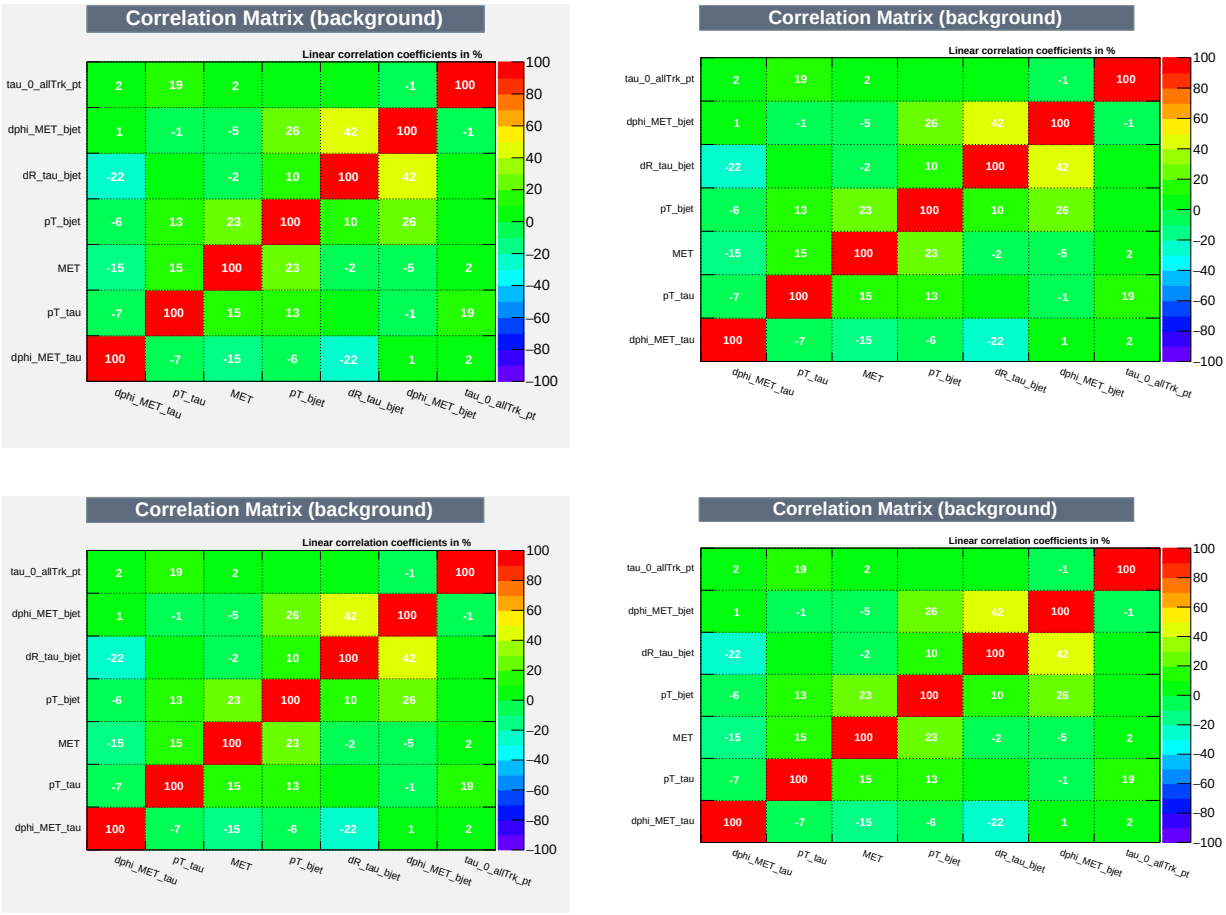



Figure 5.21: Correlation matrix of BDT input variables, where $p_T^{\tau\text{-track}}$ in place of $\Upsilon$ is used (see Fig. 5.9 for comparison), for the top backgrounds. The five $H^{\pm}$ mass ranges used in the BDT training are presented, 90–120 GeV (top left), 130–160 GeV (top right), 160–180 GeV (middle left) and 200–400 GeV (middle right). Where: tau_0_allTrk_pt= $p_T^{\tau\text{-track}}$; dphi_MET_bjet= $\Delta\phi_{b\text{-jet},E_T^{miss}}$; dR_tau_bjet= $\Delta R_{b\text{-jet},\tau}$; pT_bjet= $p_T^{b\text{-jet}}$; MET= $E_T^{miss}$; pT_tau= $p_T^{\tau}$ and dphi_MET_tau= $\Delta\phi_{\tau,E_T^{miss}}$.

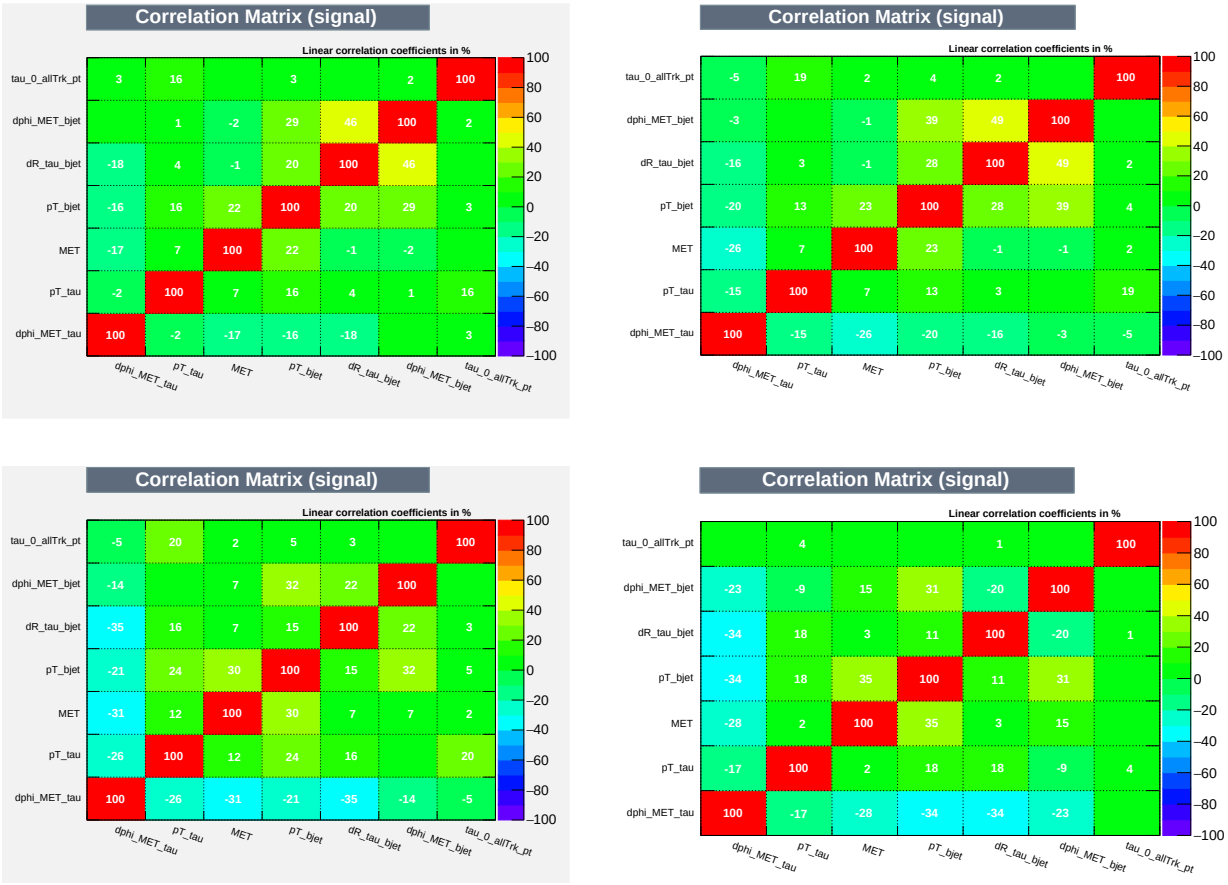



Figure 5.22: Correlation matrix of BDT input variables, where $p_T^{\tau\text{-track}}$ in place of $\Upsilon$ is used (see Fig. 5.10 for comparison), for the signal. The five $H^{\pm}$ mass ranges used in the BDT training are presented, 90–120 GeV (top left), 130–160 GeV (top right), 160–180 GeV (middle left) and 200–400 GeV (middle right). Where: tau_0_allTrk_pt= $p_T^{\tau\text{-track}}$; dphi_MET_bjet= $\Delta\phi_{b\text{-jet},E_T^{miss}}$; dR_tau_bjet= $\Delta R_{b\text{-jet},\tau}$; pT_bjet= $p_T^{b\text{-jet}}$; MET= $E_T^{miss}$; pT_tau= $p_T^{\tau}$ and dphi_MET_tau= $\Delta\phi_{\tau,E_T^{miss}}$.

### 5.3.6 Additional study with 1-prong and 3-prong events

Often in data analysis one has to deal with events that fall into different categories. For example $\tau$ lepton decays into final states with either *one* or *three* charged particles. This will inevitably result in different optimal decision boundaries in these different event categories. Thus, it can be better to help the classifiers by dividing the sample manually in order to train individualy each event category, as it is the case for $\tau$ leptons decaying into final state with one (1-prong) or three (3-prong) charged paricles [124]. Therefore, the following check, for $H^{\pm}$ masses in the range $200-400$ GeV, is also performed. It compares the BDT output when the BDT training is done for all events together with the BDT output when the BDT training is performed separately for events with a selected 1- or 3-prong $\tau_{\text{had-vis}}$ object with the $p_T^{\tau\text{-track}}$ in place of $\Upsilon$ included in the final BDT discriminant for events where $\tau_{\text{had-vis}}$ has only one associated track. The result presented in Fig. 5.23 suggests that having large enough statistics in the high $H^{\pm}$ mass region it would be interesting to study there the performance of the BDT after the training done separately for events with a selected 1- or 3-prong $\tau_{\text{had-vis}}$ object.

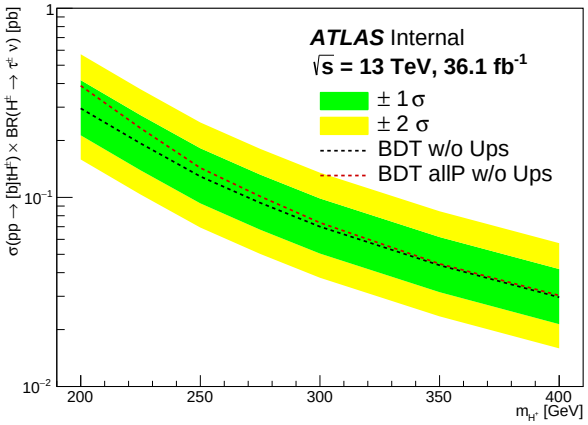



Figure 5.23: Expected 95% CL exclusion limits on $\sigma(pp \to [b]tH^{\pm}) \times \text{BR}(H^{\pm} \to \tau^{\pm}\nu)$ for charged Higgs boson production as a function of $m_{H^{\pm}}$ in 36.1 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 13$ TeV for comparison of BDT output when the BDT training is performed separately for events with a selected 1- or 3-prong $\tau_{\text{had-vis}}$ object. This study is done for $H^{\pm}$ masses in the range $200-400$ GeV with the $p_T^{\tau\text{-track}}$ in place of $\Upsilon$ included in the final BDT discriminant (BDT w/o Ups), with the BDT output when the BDT training is performed for all events together (BDT allP w/o Ups).

**Final remarks**

In the presented analysis for the first time MVA method (FastBDT) is implemented into the search for charged Higgs boson in the $H^{\pm} \to \tau^{\pm}\nu$ channel in the ATLAS experiment. The performance in terms of the signal - background separation is compared to that of the nominal cut-based analysis. Also, other studies discussed above are carried out with the goal of finding the best configuration of the BDT discriminant. It is shown that there is some space for improvement of the optimisation of the variable sets and the mass ranges compared to the choice taken for analysis presented in Sec. 5.2. Performed studies show a possible path for further ehnancement of the limits when the full Run-2 data becomes available and give more confidence that the obtained exclusion limits are robust.

# Chapter 6

# Embedding Method

In this Chapter the embedding technique and the first results of its application to the ATLAS data from LHC Run-2 is described in the following order: Section 6.1 - motivation; Section 6.2 - the significance of the method; Section 6.3 - selection of samples to be used for embedding; Section 6.4 - implementation of the embedding method and Section 6.5 - validation and usage of the technique.

The analysis described in this Chapter constitutes first complete implementation of the embedding technique in the Run-2 data analysis which has been performed by the Author.

## 6.1 Motivation

A precise understanding of the background processes is crucial for making any observations of new, yet undiscovered phenomena. In the ATLAS experiment one of the two main signals of the $H^{\pm}$ particle that could be produced in the proton-proton collisions e.g. in the process $pp \to t\bar{t} \to bH^{\pm}\bar{b}W^{\mp}$ (right diagram in Fig. 6.1), is its decay to a tau lepton, $\tau^{\pm}$, and tau (anti-)neutrino $\nu_{\tau}$ ($\bar{\nu}_{\tau}$). Unfortunately, it is very hard to distinguish it from the SM top quark pair production process, $pp \to t\bar{t} \to bW^{\pm}\bar{b}W^{\mp}$, where $W^{\pm}$ boson decays to $\tau^{\pm}$ and tau (anti-)neutrino (left diagram in Fig. 6.1). Therefore, it is crucial to determine as precisely as possible the expected number and characteristics of the background events. This can be achieved on the basis of the MC simulations or using the data-driven techniques.

It is important to note that even though there are other processes in the proton-proton collisions that can produce the $H^{\pm}$ boson, e.g. top quark associated production relevant when $H^{\pm}$ boson would be heavier than the top quark, all of them have corresponding process with $W^{\pm}$ in place of $H^{\pm}$. Therefore, in all cases the contribution from $W^{\pm} \to \tau^{\pm}\nu$ process would significantly contribute to the irreducible background. However, the background with true hadronically decaying $\tau$ leptons, i.e. the one studied in detail in this Chapter, is the most important for

the relatively light charged Higgs search, where the main process contributing to the irreducible background is the top quark production.

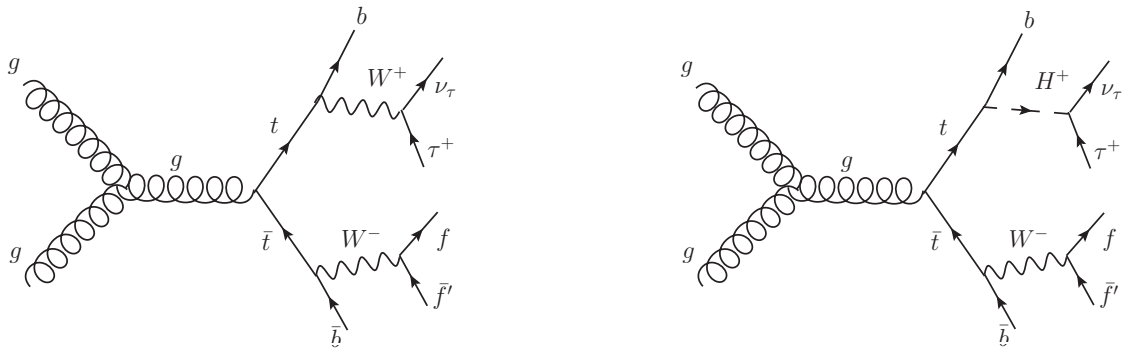



Figure 6.1: Feynman diagrams for the dominant process giving the SM background, top quark pair production, (left) and the charged Higgs boson signal production for $m_{H^\pm} < m_{top}$ (right).

The most commonly employed technique for the background estimation in collider physics is the use of the MC simulations. Relying on a detailed understanding of the SM physics available, one generates large amount of simulated events and uses them to obtain distributions that are expected to arise from the SM processes alone. The main disadvantage of such an approach is that it comes with a sizeable theoretical uncertainty. In particular, MC simulations are still not capable of accurate modelling of many effects in the proton-proton collision at the same time such as pile-up, underlying event or light and b-jet production, as well as they are not able to describe the missing transverse energy with sufficient precision. Therefore, there is an increasing need for development of the data-driven techniques for the background estimation, which are as free as possible from relying on simulations. One of the examples of such a technique is the so-called *embedding method*.

## 6.2 The significance of the embedding method

The main physical principle on which the embedding method is based is the lepton universality of the $W$ boson decay, as explained in the Chapter 2. It guarantees that the $W^\pm \to \tau^\pm \nu$ and $W^\pm \to \mu^\pm \nu$ decays have the same amplitude and therefore the same cross sections to a very good approximation. The difference in the cross sections comes only from the phase space factors, related to the mass differences between $\mu$ and $\tau$ leptons. In the ultra-relativistic limit, like the one at the LHC, this effect is negligible. This allows the use of information from the selected measured events with $\mu$ leptons from $W^\pm \to \mu^\pm \nu$ processes by constructing hybrid events, where the detector signature of a muon is replaced by that of a $\tau$ lepton simulated with

the use of MC methods. This replacement is done directly at the level of the reconstructed tracks and calorimeter cells. Therefore, the $\tau$ lepton kinematics is obtained from the measured kinematics of the muons and one takes from simulation only well understood electroweak $W$ boson decays and decays of the $\tau$ lepton together with the detector response to them. Hence, all the other aspects of a given event are taken directly from the data including the contribution from pile-up, underlying event as well as the missing transverse energy determination exept for contribution from the neutrinos resulting from $\tau$ decays.

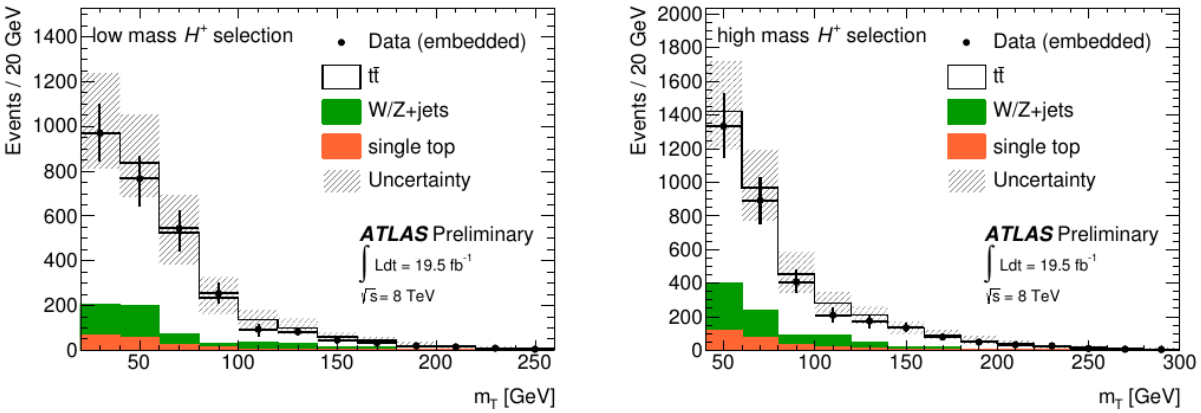
In high energy physics the embedding method is still a relatively new technique. Nevertheless, it has already been successfully used in the LHC Run-1 for the studies of the background in the neutral and charged Higgs boson searches, in the channels $H^0 \to \tau^+\tau^-$ (both SM and BSM) and $H^\pm \to \tau^\pm\nu$ respectively, both by the ATLAS [34–39] and CMS collaborations [40–42]. It has been also employed for the measurements of the $W^\pm \to \tau^\pm\nu$ production cross section in ATLAS experiment [135].

For all the processes where the embedding method has been used so far, it was found to provide a result with a significantly higher level of precision then the MC simulations due to their large systematic uncertainties related to the lack of a very detailed understanding of all the effects taking place during proton-proton collision. As an example, Fig. 6.2 presents the comparison of the backgrounds with true $\tau$ leptons obtained through embedding (black points) with simulation (histogram) for the transverse mass distribution for the low and high mass charged Higgs boson search for the analysis of Run-1 data[1]. The statistical and systematical uncertainties of the embedding method are marked by the black error bars, while those of the simulation by gray hashed area. The improvement in the background determination was significant [39], which resulted in considerably better final limits on $\mathrm{B}(t \to bH^\pm) \times \mathrm{B}(H^\pm \to \tau^\pm\nu)$ for the low mass $H^\pm$ selection case (see Appendix B, Fig.B.1).

## 6.3 Selection of $\mu$+jets events

In search for charged Higgs boson in $\tau_{\text{had-vis}}$+jets topology, the first step in the implementation of the embedding technique is to select the $\mu$+jets data events which will be used for embedding of $\tau$ leptons simulated by MC methods. This selection needs to be consistent with the SM processes giving a muon, jets and $E_T^{miss}$ in the final state. The dominant process of this kind is top quark pair production, $pp \to t\bar{t} \to bW^\pm(\mu\nu)\bar{b}W^\mp(q\bar{q})$, (see Fig. 6.3), while the other are single top quark production, W+jets and Z+jets processes. Moreover, the $\mu$+jets selection has to

[1] Low-mass $H^\pm$ selection was for 90 GeV $\leq m_{H^\pm} \leq$ 160 GeV, so for $m_{H^\pm} < m_{top}$; high-mass $H^\pm$ selection was for 180 GeV $\leq m_{H^\pm} \leq$ 1000 GeV for the analysis of Run-1 data.




(a) Low mass $H^+$ selection (b) High mass $H^+$ selection

Figure 6.2: The results of the validation of the embedding method performed for the analysis of Run-1 data. Transverse mass, $m_T$, distributions for low mass $H^{\pm}$ selection (left plot) and high mass $H^{\pm}$ selection (right plot) for embedded data (points) and MC simulations (histograms). For MC samples only events with true $\tau$ leptons are considered. From [39].

be similar and simultanously looser (in order not to introduce any bias) to the signal $\tau_{\text{had-vis}}$+jets selection.

The selection of an appropriate $\mu$+jets sample from the collected data was carried out basing upon the kinematic cuts and muon selection trigger to achieve optimal signal to background ratio and sample purity.

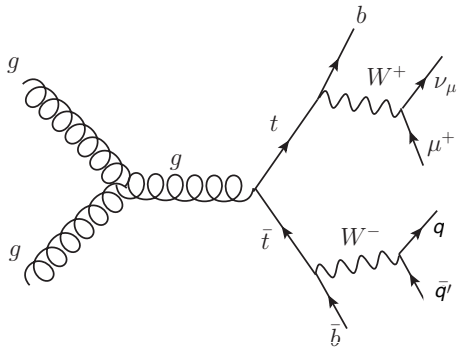



Figure 6.3: Feynman diagram of the Standard Model process of top quark pair production with $W^{\pm} \to \mu^{\pm}\nu$ decay used for the embedding.

An event to be selected for the embedding procedure has to fulfill the following conditions:

- an event is triggered by the $\mu$ trigger with the transverse momentum threshold on the $\mu$ trigger object of 20 GeV,

- an event contains exactly one $\mu$ passing "*tight*" identification criteria within $|\eta| < 2.5$ and with $p_T > 30$ GeV,
- an event does not contain within $|\eta| < 2.47$ an isolated electron with $p_T > 20$ GeV and passing "*loose*" identification criteria,
- an event contains at least three reconstructed jets with $p_T > 25$ GeV and within $|\eta| < 2.5$,
- and at least one of these jets is the $b$-tagged one, where working point giving 70% of $b$-tagging efficiency is chosen,
- an event has the missing transverse energy $E_T^{miss} > 45$ GeV.

Note that the neutrino from the hadronic $\tau$ lepton decay in an embedded event contributes to the final $E_T^{miss}$ of the event. If this contribution is significant, it can occur that the $E_T^{miss}$ in the original, $\mu$+jets event is too small for the event to pass the $\mu$+jets event selection, but the final $E_T^{miss}$ in the embedded event would be large enough for the event to pass the final $\tau_{\text{had-vis}}$+jets SR selection [136, 137]. However, it was shown during Run-1 data studies that $E_T^{miss}$ cut in the $\mu$+jets sample selection removes only events with the very low $m_T$. Hence, in Run-2 data analysis where we take in the final $\tau_{\text{had-vis}}$+jets SR selection only events with $m_T > 50$ GeV and with $E_T^{miss} > 150$ GeV this effect is assumed to be negligible and no additional systematics uncertainty is considered.

The above selection criteria are chosen to be more loose than the SR selection. In such a way they do not introduce any bias in how the potential signal events reconstructed from the $\tau$ lepton tracks and $E_T^{miss}$ in $H^{\pm} \to \tau^{\pm}\nu$ case are selected. At the same time the cuts optimisation in the $\mu$+jets sample is also done with the goal of obtaining DRAW format data writting rate at the acceptably low level of few Hz[2].

It is worth mentioning that one can have contamination from muons coming from the charged Higgs decay signal process[3]: $H^{\pm} \to \tau^{\pm}\nu \to \mu^{\pm}\nu\bar{\nu}\nu$. Such muons, after being replaced by the simulated $\tau$ leptons, could conceivably satisfy all the signal criteria and therefore erroneously enhance the signal. Notwithstanding, this contamination was found to be insignificant in the analysis carried out by the ATLAS collaboration on the data collected at LHC Run-1, mostly due to its much softer distribution in transverse mass and therefore did not have any appreciable effect on the signal strength nor final limits [39, 136].

[2]Technical issue imposed by the Collaboration.

[3]Note that the cross section for the direct process $H^{\pm} \to \mu^{\pm}\nu$ is very small and therefore, its potential effect on the background estimation is negligible.

On the other hand the contribution from muons being products of the $\tau$ lepton decay in the process $W^{\pm} \to \tau^{\pm}\nu \to \mu^{\pm}\nu\bar{\nu}\nu$ is taken into account in the normalization factor (see Section 6.5.1 below).

## 6.4 Implementation of the embedding method

In order to implement the embedding method the following steps are performed (they are also presented below in the form of a flowchart, in Fig. 6.4, intuitively showing the whole procedure [138]):

1. From the selected $\mu$+jets sample the kinematics of a given $W^{\pm} \to \mu^{\pm}\nu$ process is determined and momentum of the original $\mu$, $\vec{p}_{\mu}$, is extracted and rescaled in order to take into account the mass difference between the $\tau$ and $\mu$ leptons:

$$\vec{p}_{\tau} = \frac{\sqrt{E_{\mu}^2 - m_{\tau}^2}}{\sqrt{\vec{p}_{\mu} \cdot \vec{p}_{\mu}}} \vec{p}_{\mu}, \tag{6.1}$$

   where $\vec{p}_{\tau}$ is the $\tau$ momentum, $m_{\tau}$ is the $\tau$ mass and $E_{\mu}$ is the energy of $\mu$. This rescaling is important as the lepton universality of the $W$ boson decays holds in its rest frame and then the transformation to the laboratory frame has a non-negligible dependence on the lepton mass, even though all the events are ultra-relativistic. The momentum obtained after these transformations is processed by the TAUOLA [139] - Monte Carlo generator for the simulation of the $\tau$ lepton production and decay. Whereas, photon final state radiation is generated by PHOTOS [114].

2. For the $\tau$ decay products a realistic ATLAS detector simulation has to be performed including the signal digitisation and followed by the full reconstruction of an event.

3. All the tracks together with the energy depositions in the calorimeter in close vicinity of the muons are to be replaced by the results of the simulated $\tau$ decay obtained in the previous step (i.e. the simulated hadronic $\tau$ decay is embedded in the real proton-proton collision event). The stored information about the energy deposited in the calorimeter and the muon spectrometer is removed with the use of a simulated $W^{\pm} \to \mu^{\pm}\nu$ decay with exactly the same kinematics as the considered real event. An illustration of the replacement procedure for the tracks is given in Fig. 6.5.

4. The resulting hybrid event undergoes full reconstruction procedure with the aim of reconstructing $\tau$-jets, electrons and muons, as well as the missing transverse energy and other high-level objects.

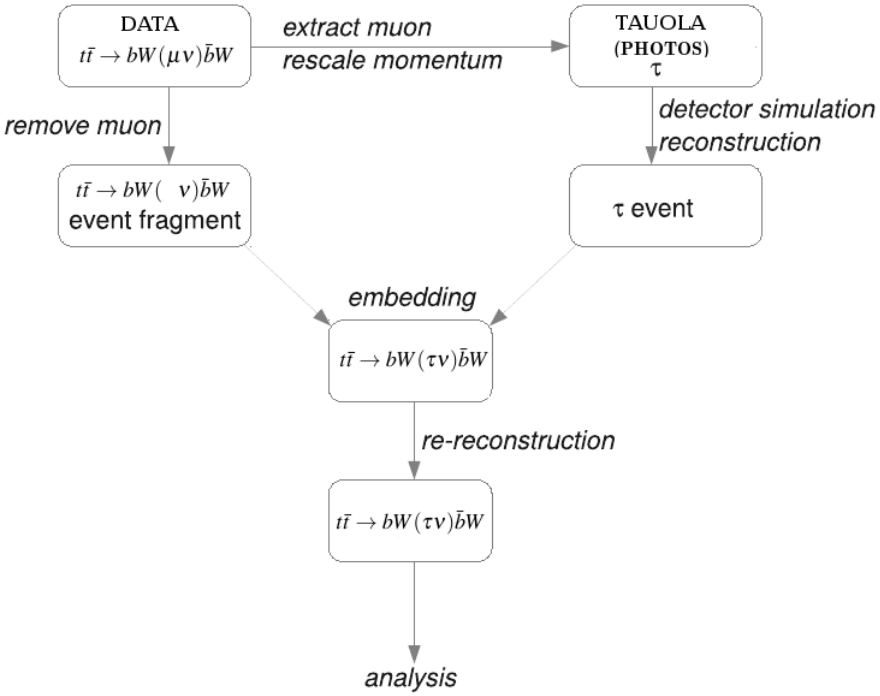
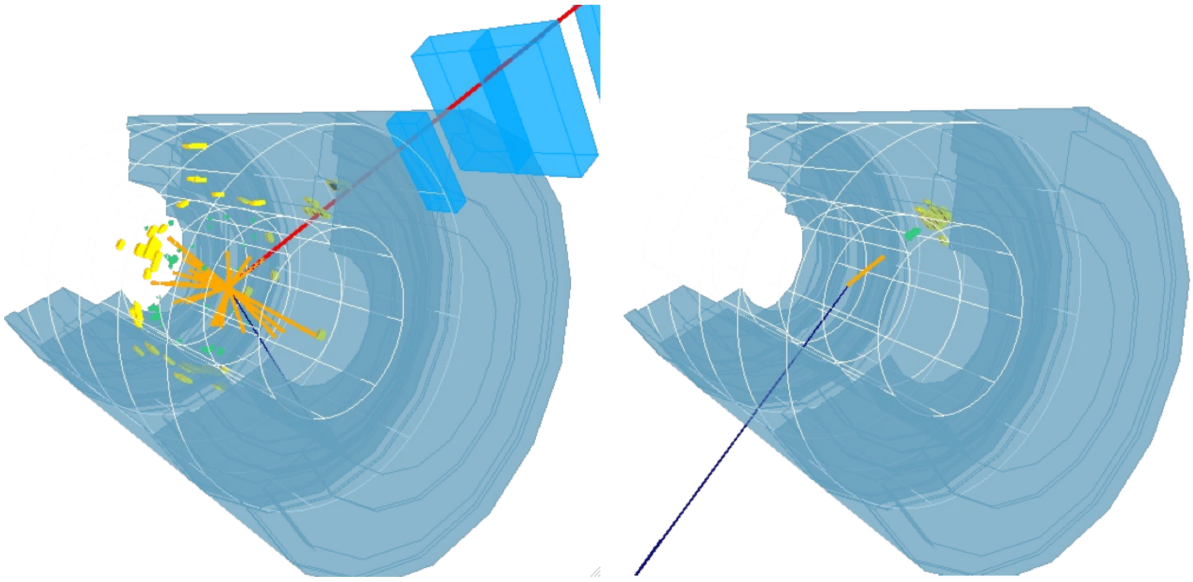


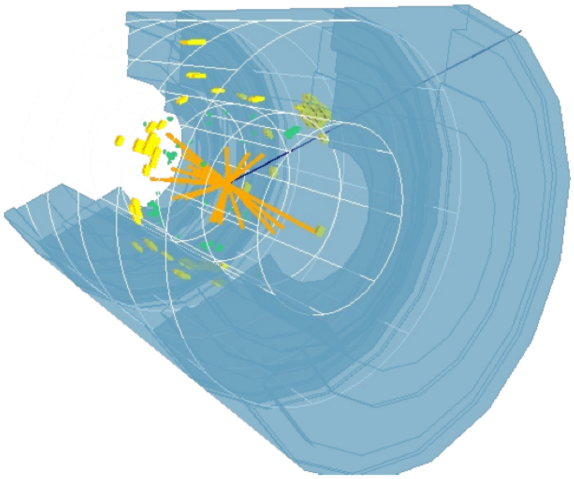

Figure 6.4: The flowchart of the embedding procedure.

Due to the absence of data in format appropriate as an input for embedding e.i. DRAW data - see Chapter 3 for details, produced by the Collaboration a private production was carried out by the Author. As such production is very demanding on computing resources, only 0.9 fb$^{-1}$ from 3.2 fb$^{-1}$ of the 2015 data were processed. This provides 14470 events after $\mu$+jets selection, i.e. events used for the embedding step, and *198 events* in the signal region after the SR selection.

The MC embedding is performed on raw simulation samples (RDO) files of $t\bar{t}$ MC samples used by standard analysis and described in Chapter 4. In order to be conceptually as close as possible to DRAW data set used for embedding the filtration was done on RDO $t\bar{t}$ samples with the same $\mu$+jets selection as on DRAW (see Section 6.3). This selection leaves 275180 events with *3419 events* after the SR selection.



(a) $W \to \mu\nu$ event (b) simulated $\tau_{\text{had-vis}}$ decay



(c) embedded hybrid event

Figure 6.5: ATLAS Event Display showing (a) $W^{\pm} \to \mu^{\pm}\nu$ event, (b) simulated hadronic $\tau$ lepton decay and (c) the embedded hybrid event in a cut-away view of the ATLAS detector. The muon is shown in red and traverses the muon chambers, $E_T^{miss}$ is shown in dark blue. The $\tau_{\text{had-vis}}$ candidate can be identified by its track (orange) and deposits in the calorimeters (green and yellow). From [137].

## 6.5 Validation and usage of the technique

The embedding method, as implemented and used for the Run-1 data, has been shown to significantly improve the precision of the background estimation with true hadronically decaying $\tau$ leptons. The aim of the work described in this Chapter has been to investigate whether for the analysis for Run-2 data a similar result holds or not. In other words, if a proper implementation of the embedding technique for Run-2 data will give rise to the background estimation with significantly smaller systematic uncertainties than the corresponding background estimation from MC simulations.

First test of the embedding procedure relies on a comparison of various kinematic and angular distributions obtained from *simulated hybrid events* (embedding method applied to MC samples) and $t\bar{t}$ samples simulated with the MC generators.

After the above closure-check has been done the shape of the transverse mass distribution, see Eq. 4.1, after applying the SR selection and appropriate corrections (described in the Section below), has been obtained and is discussed at the end of the Chapter.

### 6.5.1 Corrections Applied to Embedded Samples

The normalisation of the embedded sample giving the final number of background events with true hadronically decaying $\tau$ leptons is obtained from the number of embedded events in the following way:

$$N_\tau = N_{\text{embedded}} \cdot \left(1 - c_{\tau\to\mu}\right) \cdot \frac{\varepsilon^{\text{E}_\text{T}^\text{miss}\text{-trigger}}}{\varepsilon^{\mu\text{-reco,trigger}}} \cdot \text{BR}(\tau \to \text{hadrons}\, \nu), \tag{6.2}$$

where $N_{\text{embedded}}$ is the number of embedded events in the signal region; $c_{\tau\to\mu}$ represents the fraction of events in which the selected $\mu$ is a decay product of a $\tau$ lepton; $\varepsilon^{\text{E}_\text{T}^\text{miss}\text{-trigger}}$ and $\varepsilon^{\mu\text{-reco,trigger}}$ are efficiencies of the $E_T^{miss}$ trigger, the $\mu$ trigger and $\mu$ identification, respectively. Finally, $\text{BR}(\tau \to \text{hadrons}\, \nu)$ is the branching fraction of $\tau$ lepton decay into hadrons and neutrino.

- **Correction for the $\tau \to \mu$ decays**

  The contribution from muons being products of the $\tau$ leptons decay in the process

  $$W^\pm \to \tau^\pm \nu \to \mu^\pm \nu \bar{\nu} \nu$$

  are taken into account in the normalization factor $c_{\tau\to\mu}$, which is found to be $7.7 \pm 0.5\%$ for the 2015 data. The estimation of this factor is based on the simulated *hybrid events*,

i.e. embedded $t\bar{t}$ Monte Carlo simulations. The amount of events, in which the original muon came from a $\tau$ or a direct $W$ decay and pass the SR selection after the embedding are compared. The muons from $\tau$ lepton decays typically have lower transverse momenta than the muons from $W$ decays and therefore they are less frequently passing the SR selection after embedding. This effect is strengthened by the requirement on the transverse momentum of $\tau$'s being greater than 40 GeV in the SR selection compared to $p_T^{\mu} > 30$ GeV in the $\mu$+jets selection. It is assumed, as in the Run-1 data analysis, that the $c_{\tau\to\mu}$ factor is independent of the kinematic variables, i.e. it is only a normalisation factor. In principle this is not the case, since $p_T$ spectrum of muons coming from decay of $\tau$'s is in fact expected to be softer that from the $W$ boson decay. However, taking this into account leads to a small modification of already subdominant correction factor.

- **Correction for the muon reconstruction and trigger efficiencies**

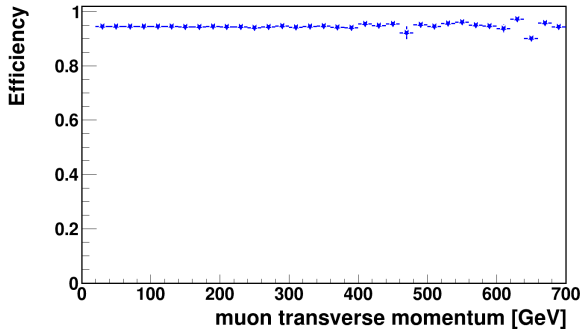
  The $\mu$ trigger and reconstruction efficiencies, $\varepsilon^{\mu\text{-reco,trigger}}$, are corrected for to remove efficiencies related with the original $\mu$+jets selection in the sample used as an input for embedding.

  In order to remove these effects, the inverse of the $\mu$ reconstruction and $\mu$ trigger efficiencies are applied as the correction factors to the embedded events. The efficiencies are shown in Fig. 6.6 as functions of pseudorapidity and transverse momentum of the original muon for data taken in 2015. More details about used functions can be found in [83, 140]. Their impact on the events with embedded $\tau$ lepton is shown in Fig. 6.7. These plots present comparison of the distributions of $p_T^{\tau}$, $\eta^{\tau}$, $E_T^{miss}$, $m_T$, $\Delta\phi(\tau, E_T^{miss})$ and $\Upsilon$ before and after applying the correction factors for events passing the final selection on $H^{\pm}$ search. It is found that the impact of the correction factors on the $\mu$ reconstruction and trigger efficiencies is about 25% as a function of $p_T^{\tau}$ and between 20% and 50% as a function of $\eta^{\tau}$ on the normalisation of the embedded data sample.

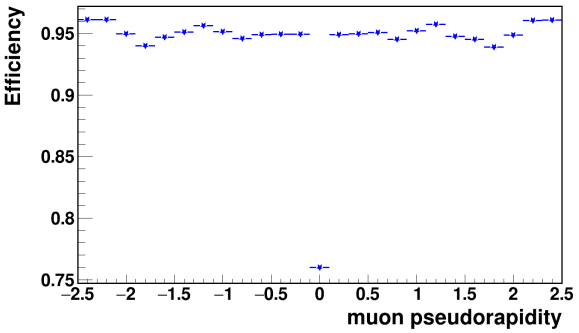
- **Application of the $E_T^{miss}$ trigger efficiency**

  The $E_T^{miss}$ trigger efficiency, $\varepsilon^{\mathrm{E_T^{miss}\text{-trigger}}}$, is applied to the embedded simulation events (see Fig.6.9) and embedded data ( see Fig.6.10) since no trigger information is accessible in embedded samples. The strategy for the treatment of the $E_T^{miss}$ trigger [33, 87, 141] is to derive the trigger efficiency from data in bins of $E_T^{miss}$. Then the binned $E_T^{miss}$-dependent efficiency is transformed into a continuous efficiency by fitting to it the error function

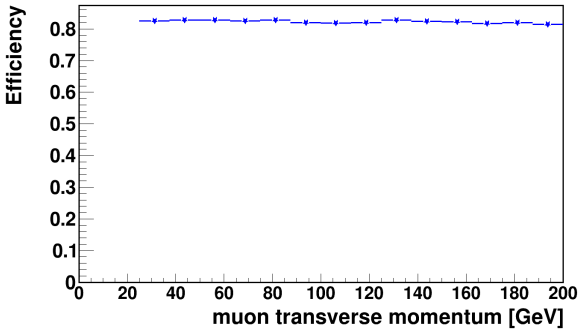
$$F(x) = p_0 \cdot \left[1 + \mathrm{erf}\left(\frac{\mathrm{x} - \mathrm{p_1}}{\mathrm{p_2}}\right)\right] + p_3, \tag{6.3}$$



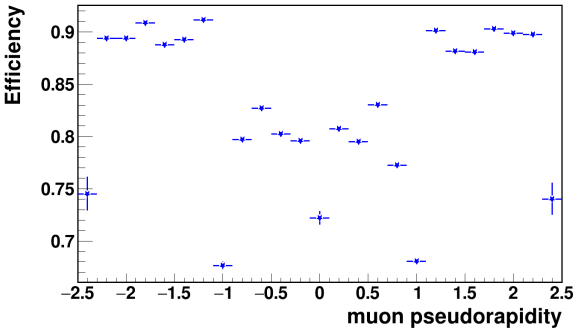

$(a)$ $\mu$ reconstruction efficiency vs. $p_T^\mu$




$(b)$ $\mu$ reconstruction efficiency vs. $\eta^\mu$




$(c)$ $\mu$ trigger efficiency vs. $p_T^\mu$




$(d)$ $\mu$ trigger efficiency vs. $\eta^\mu$

Figure 6.6: Respectively, the muon reconstruction and trigger efficiencies used in $\mu$+jets selection as a functions of muon transverse momentum $(p_T^\mu)$ : (a), (c) and muon pseudorapidity $(\eta^\mu)$ : (b), (d).

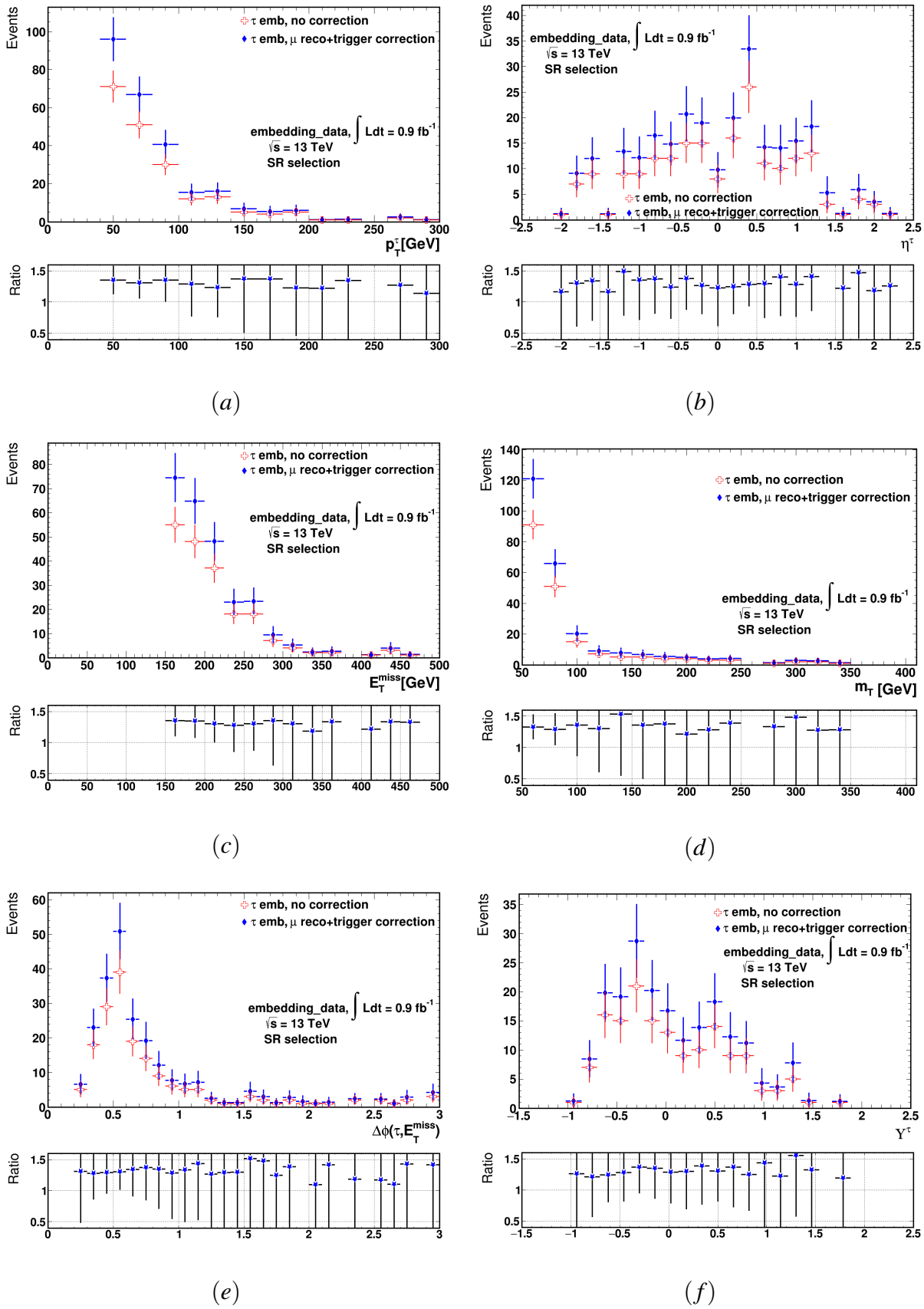



Figure 6.7: Juxtaposition of (a) the embedded $\tau$ transverse momentum $p_T^\tau$ and (b) pseudorapidity $\eta^\tau$, (c) $E_T^{miss}$, (d) $m_T$, (e) $\Delta\phi(\tau, E_T^{miss})$ and (f) $\Upsilon^\tau$ observables before and after applying the correction factors on the $\mu$ reconstruction and trigger efficiencies for the data embedding.

where $p_0, p_1, p_2, p_3$ are the fit parameters. This is done to remove any potential bias caused by the binning. Embedded events are weighted using the efficiency curve, based on the $E_T^{miss}$ in the event. The efficiency to apply to embedded simulation and data events is

$$\varepsilon = \frac{\texttt{event selection} + \texttt{HLT_xe70_tc_lcw}}{\texttt{event selection}}, \tag{6.4}$$

where HLT_xe70_tc_lcw is the $E_T^{miss}$ trigger used to collect events in the $\tau_{\text{had-vis}}$+jets channel for the 2015 data. Corresponding fit is performed in the region with the $e$+$\tau_{\text{had-vis}}$ selection (see Fig. 6.8), i.e. selection differing from that in SR by requiring exactly one selected electron with a *loose* identification and $p_T > 26$ GeV; at least one selected $\tau_{\text{had-vis}}$ object with a *loose* identification and $p_T > 30$ GeV; at least two selected jets; and $E_T^{miss} > 100$ GeV [33].

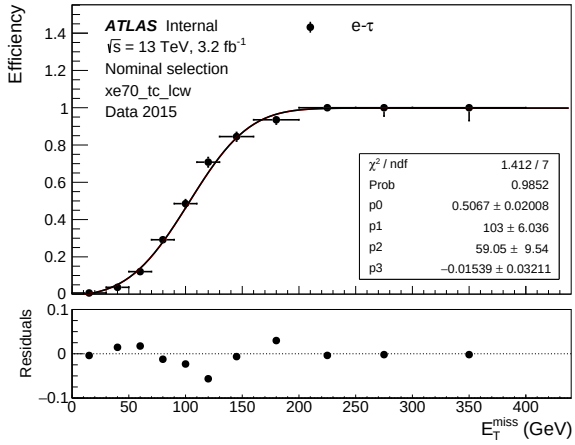



Figure 6.8: Trigger efficiency and the corresponding fit in case of the $e$+$\tau_{\text{had-vis}}$ selection. From [33].

The plots in Fig. 6.9 - 6.10 present comparison of the distributions of $p_T^\tau$, $\eta^\tau$, $E_T^{miss}$, $m_T$, $\Delta\phi(\tau, E_T^{miss})$ and $\Upsilon$ (for embedded data) before and after applying the correction factors on the $E_T^{miss}$ trigger efficiency for embedded simulation and data events passing the SR selection on $H^\pm$ search.

It is found that the impact of the correction on $E_T^{miss}$ trigger efficiency is about $-5\%$ as a function of $p_T^\tau$ and up to $-10\%$ as a function of $E_T^{miss}$ on the normalisation of the data embedding sample, as presented in Fig. 6.10.

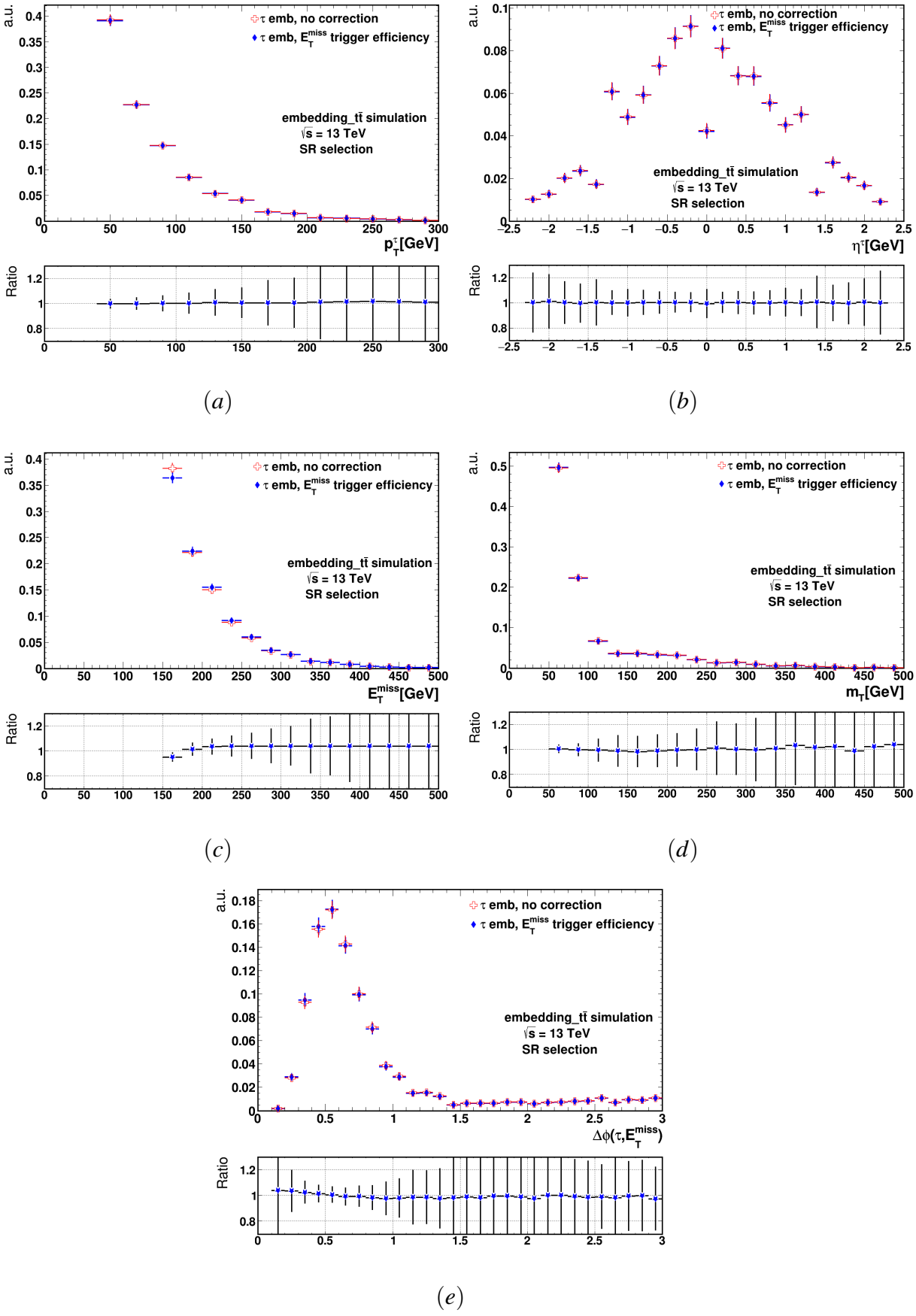



Figure 6.9: Juxtaposition of (a) $p_T^\tau$, (b) $\eta^\tau$, (c) $E_T^{miss}$, (d) $m_T$ and (e) $\Delta\phi(\tau, E_T^{miss})$ distributions before and after applying the correction on $E_T^{miss}$ trigger efficiency for embedded $t\bar{t}$ MC simulation.

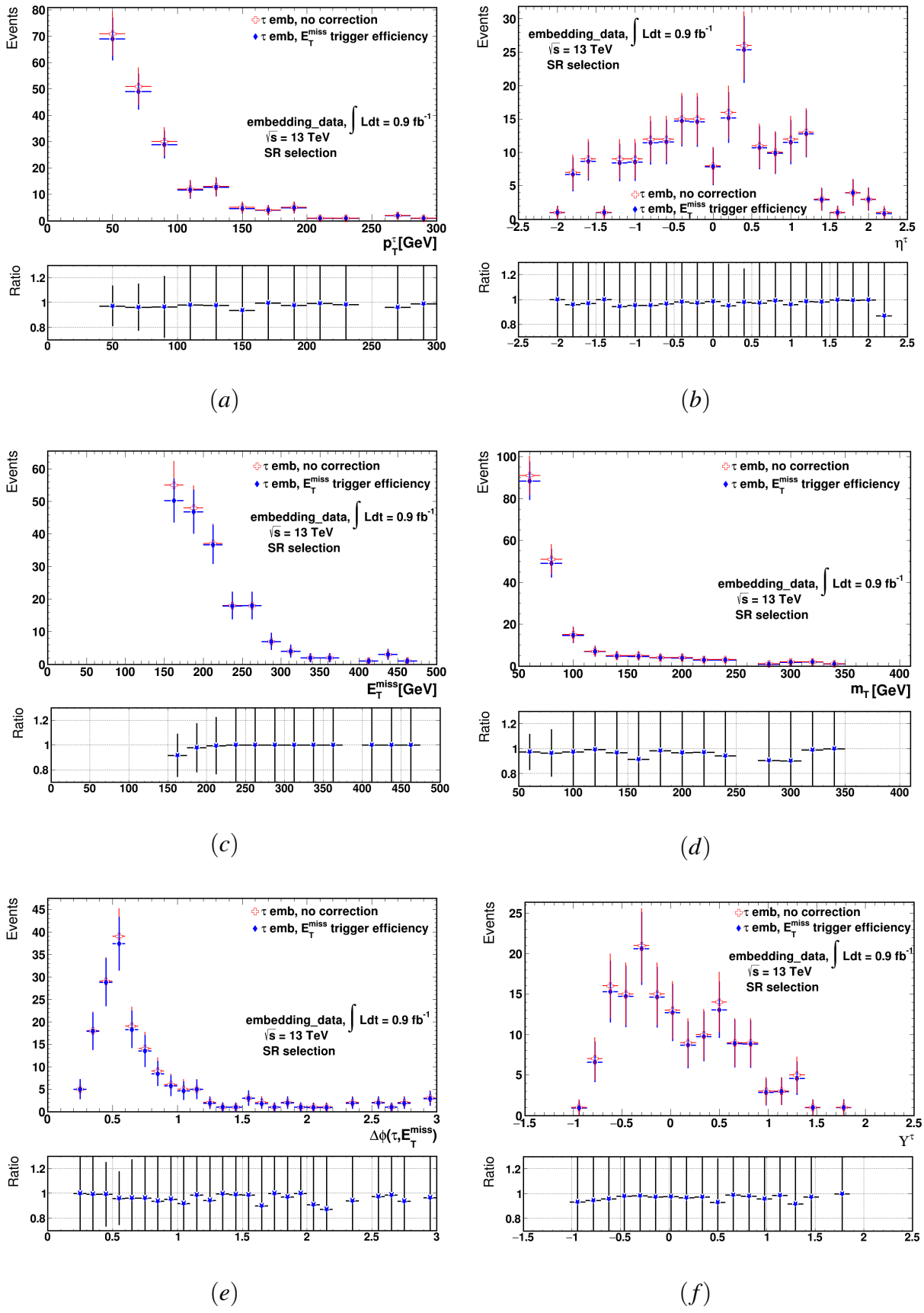



Figure 6.10: Juxtaposition of (a) $p_T^\tau$, (b) $\eta^\tau$, (c) $E_T^{miss}$, (d) $m_T$ , (e) $\Delta\phi(\tau, E_T^{miss})$ and (f) $\Upsilon^\tau$ observables before and after applying the correction on $E_T^{miss}$ trigger efficiency for data embedding.

- **$\tau_{\text{had-vis}}$ Identification Scale Factors**

  As the reconstruction and identification efficiencies for $\tau_{\text{had}}$ are different in data and simulation, and the $\tau_{\text{had-vis}}$ in the embedded data are simulated, dedicated correction factors (also referred to as the *scale factors*) need to be applied. They are defined as the ratio of the efficiency in data ($\varepsilon_{Data}$) to the efficiency in simulation ($\varepsilon_{MC}$) for $\tau_{\text{had-vis}}$ candidates to pass a certain level of identification. In the analysis presented here, the "*tight*" identification is used. Corresponding scale factors are shown in Fig. 6.11, separately for one and three tracks (prongs) $\tau_{\text{had-vis}}$ candidates. The performance of used algorithms is measured with Z boson or top quark decays to $\tau$ leptons and uses the full 2015 dataset of $pp$ collisions collected at the LHC, at $\sqrt{s} = 13$ TeV corresponding to the integrated luminosity of 3.2 fb$^{-1}$ [142].

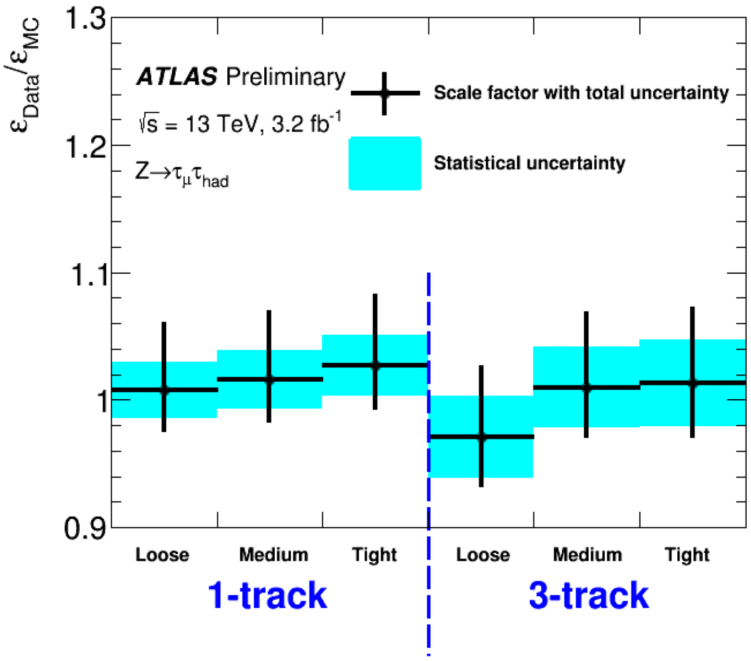



Figure 6.11: The scale factors ($\varepsilon_{Data}/\varepsilon_{MC}$) needed to bring the $\tau$ identification efficiency in simulation ($\varepsilon_{MC}$) to the level observed in data ($\varepsilon_{Data}$) for one and three tracks (prongs) $\tau_{\text{had-vis}}$ candidates with $p_T > 20$ GeV. The combined systematic and statistical uncertainties are shown. From [142].

### 6.5.2 Systematic Uncertainties

As it has been already mentioned, the most commonly employed technique for the background estimation in collider physics is the use of the Monte Carlo simulations. However, in this approach one has to deal with sizeable theoretical uncertainties. Estimating the background with true hadronically decaying $\tau$ leptons using the embedding method one is independent of theoretical cross sections and their uncertainties or of the choice of generator model. What is more, since everything except the $\tau$ in the embedded data is taken directly from collision data, no additional uncertainties related e.g. to the jet energy scale, b-tagging efficiency or pile-up need to be considered. Nevertheless, the embedding procedure comes with its own systematic uncertainties which are discussed below.

- **Uncertainties related to embedding method**
  The only systematic uncertainty directly related to the embedding procedure itself is the one concerning the amount of energy that is subtracted in calorimeter cells attributed to a muon. Its effect is estimated by changing the subtracted energy by $\pm 20\%$ [143]. Another uncertainty taken into account is connected to the normalisation uncertainty which comes from the correction factor for muons being products of $\tau$ lepton decay collected in the $\mu$+jets sample, i.e. the uncertainty of the $c_{\tau\to\mu}$ factor, which is estimated to be around 0.5% in the 2015 data.

- **Uncertainties related to $E_T^{miss}$ trigger**
  The efficiency of the $E_T^{miss}$ trigger is measured in the $e$+$\tau_{\text{had-vis}}$ region of the data and its parameterisation shows a small dependence on the identification criteria (*loose* or nominal) of the electron and $\tau_{\text{had-vis}}$ candidates, as well as on the minimum number of jets used in the definition of this region. The corresponding small variations of the values of the fit function to the $E_T^{miss}$ trigger efficiency, together with the statistical uncertainty used for the fit function, are accounted for as systematic uncertainties, which amount to the uncertainty of 1.4% [33] on the event yield in the signal region, see Fig. 6.12 and 6.13.

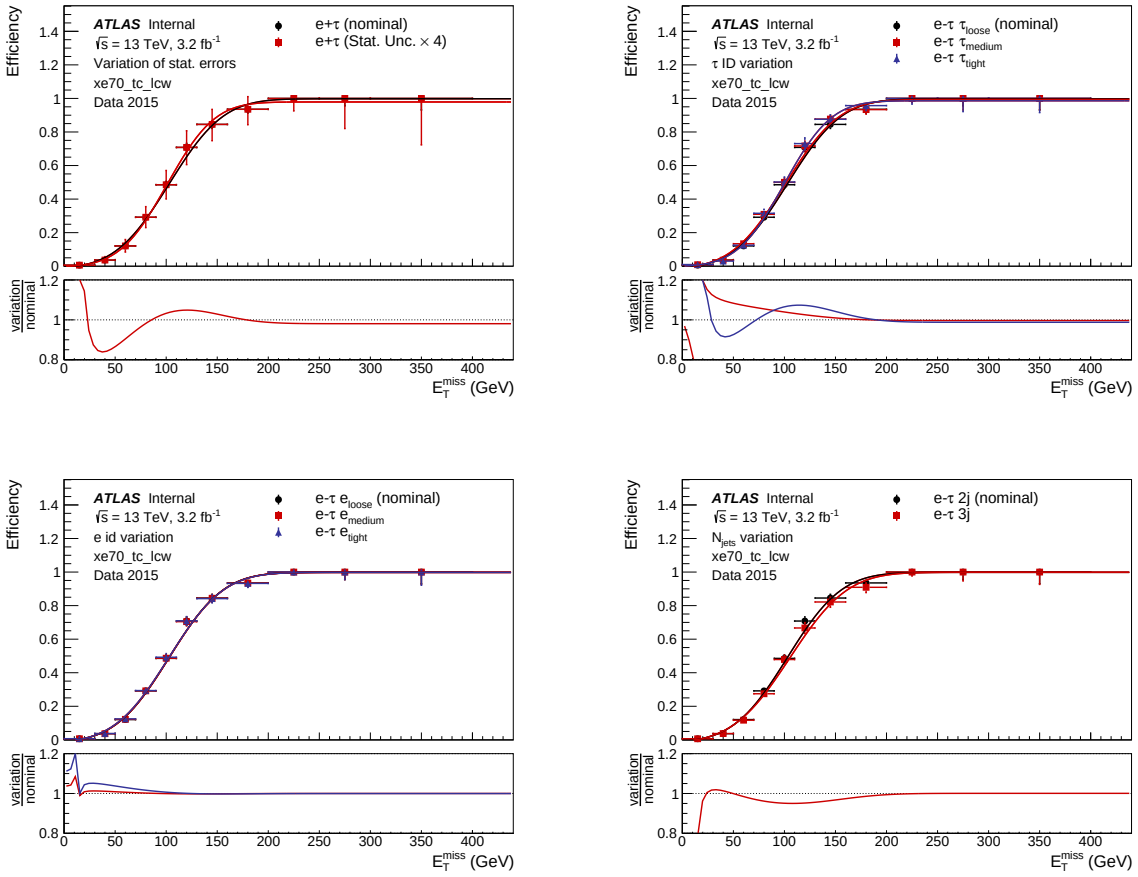



Figure 6.12: Trigger efficiencies in data and corresponding fits for the $e$+$\tau_{\text{had-vis}}$ selection and `HLT_xe70_tc_lcw` trigger (a) where the errors on the fitted points have been increased to 4× Stat. error. (b) when varying the $\tau_{\text{had-vis}}$ identification between *loose* (nominal), *medium* and *tight*; (c) when varying the *e* identification between *loose* (nominal), *medium* and *tight*; (d) when increasing the requirement on the number of jets to $\geq 3$. The ratio plots show the fit of the variation divided by the fit of the nominal case. From [33].

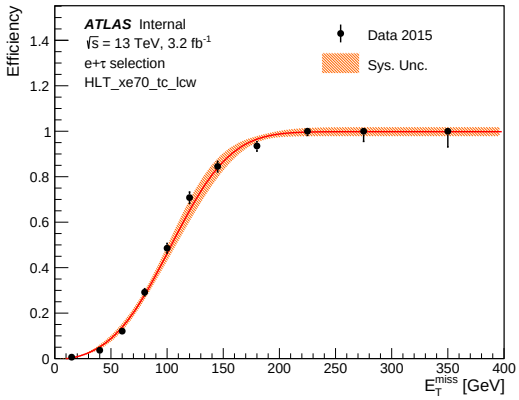



Figure 6.13: Trigger efficiencies in data and corresponding fits for the $e$+$\tau_{\text{had-vis}}$ selection and `HLT_xe70_tc_lcw` trigger with the combined systematic uncertainties arising from the statistical uncertainty that used in the fit, the identification of the electron and the $\tau_{\text{had-vis}}$ candidate, as well as the number of jets required in that selection. From [33].

- **Uncertainties related to $\tau_{\text{had-vis}}$**

  As it has been discussed above, since the products of a hadronic $\tau$ decay are simulated uncertainties on the identification correction factors as well as those on the $\tau$ energy scale need to be taken into account. They are evaluated by shifting $\tau$ identification efficiency and $\tau$ energy scale one standard deviation and checking the effect on the final distributions and yields. Systematic uncertainties depend on the pseudorapidity, transverse momentum and charged particle multiplicity of the $\tau_{\text{had-vis}}$ candidates.

  The uncertainties on the measurement of the correction factor for $\tau$ identification efficiency are approximately 5% (6%) for 1-(3-) prong $\tau_{\text{had-vis}}$ candidates, inclusive in $\eta$ and for a visible transverse momentum greater than 20 GeV. The reconstructed $\tau$ energy scale correction factor is measured with a precision of approximately (2-3)% [142].

- **Uncertainties related with $\mu$ reconstruction and trigger**

  Uncertainties related to the $\mu$ trigger and reconstruction efficiencies, used for normalising the embedded samples, are estimated by shifting these efficiencies by one standard deviation and evaluating the effects on the total yields and final distributions. The $\mu$ trigger efficiency uncertainties depend on $p_T^{\mu}$, $\eta^{\mu}$, $\phi^{\mu}$ and the data taking period. The uncertainties on the $\mu$ reconstruction efficiency depend on the same variables [137, 143].

### 6.5.3 Effect of all Systematic Uncertainties on the Embedded Sample

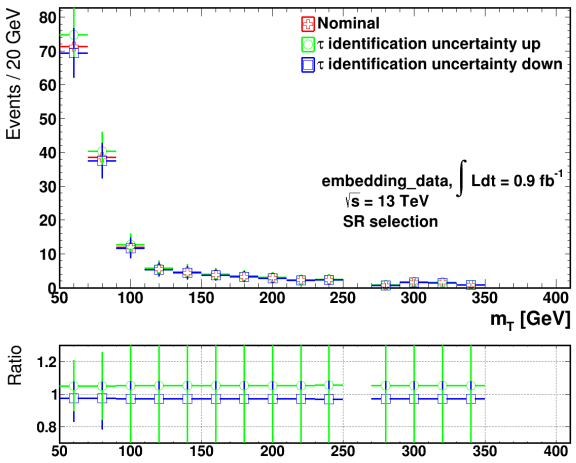
In Table 6.1 the effects of systematic uncertainties on the yield at the true $\tau$ background for the embedded samples in the SR are summarised. The dominant detector-related systematic uncertainties for this search arise from the energy scale of the simulated $\tau_{\text{had}}$ decays, (TES), from the reconstruction and identification of the simulated $\tau_{\text{had-vis}}$ candidates, ($\tau$ ID, $\tau$ RECO), and the embedding procedure itself. The variations in the transverse mass, $m_T$, shape and yields after applying the SR selection with regard to the individual systematic uncertainties are shown for the 2015 data in Fig. 6.14. Whereas, the effect of the embedding uncertainties on the shape of $p_T^{\tau}$, $E_T^{miss}$, $\eta^{\tau}$, $m_T$ and $\Delta\phi(\tau, E_T^{miss})$ distributions after applying the SR selection for embedded $t\bar{t}$ MC simulations is presented in Fig. 6.15. In both cases good agreement is observed.

| Uncertainty | $H^{\pm}$ search 2015 embedded data [%] |
|---|---|
| $\tau$ ID, $\tau$ RECO | $+5.2, -5.1$ |
| TES | $+2.4, -7.7$ |
| $E_T^{miss}$ trigger | $\pm 3.2$ |
| $\mu$ ID, trigger | $+2.7, -3.0$ |
| embedding | $\pm 3.9$ |

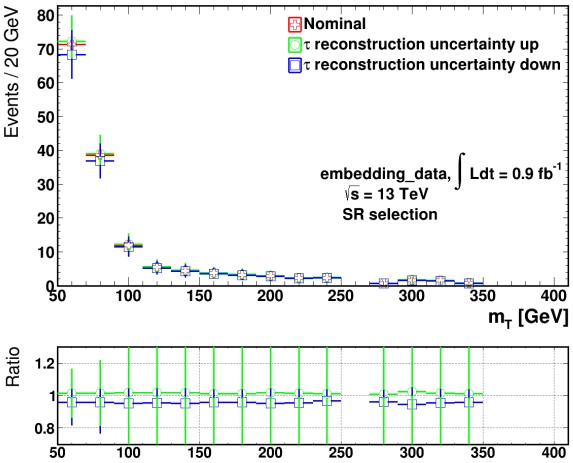
Table 6.1: Systematic uncertainties on the yield at the true $\tau$ background for the embedded data sample in the SR: $\tau$ lepton identification and reconstruction ($\tau$ ID, $\tau$ RECO), (TES) $\tau$ energy scale, ($E_T^{miss}$) missing transverse energy trigger, ($\mu$ ID, trigger ) muon identification and trigger and uncertainty related with embedding method. All variations given are percentages of the yield of the true $\tau$ background.

### 6.5.4 Embedded MC simulation - Closure-Check

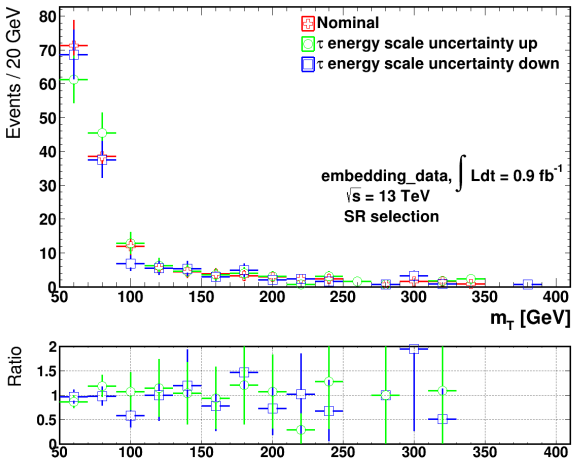
In order to validate the embedding method in the SR, a comparison of the embedded MC simulation with $t\bar{t}$ MC simulation used in analysis for events with reconstructed $\tau$ matched to true $\tau$ lepton decaying hadronically has been performed. The result of this closure-check is presented in Fig. 6.16. The resulting distributions are in a reasonable agreement within combined in quadrature statistical and systematic uncertainties related to the embedding procedure. Visible fluctuations reflect limited number of events available for the analysis.



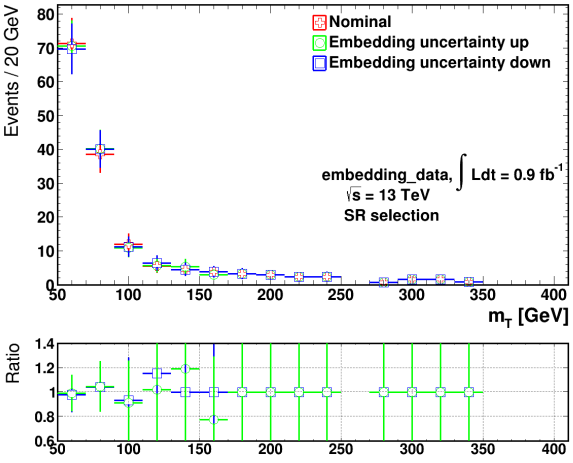

(a) $\tau$ *ID* uncertainty



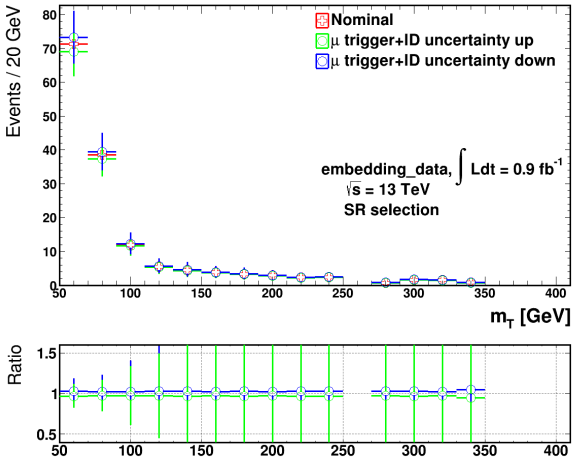

(b) $\tau$ *reco* uncertainty



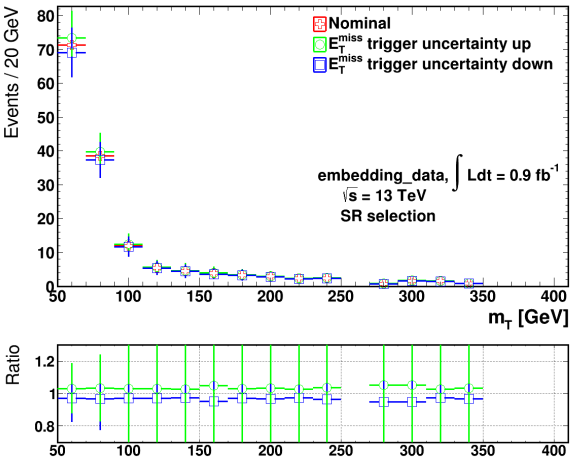

(c) *TES* uncertainty




(d) $\tau$ *emb* uncertainty

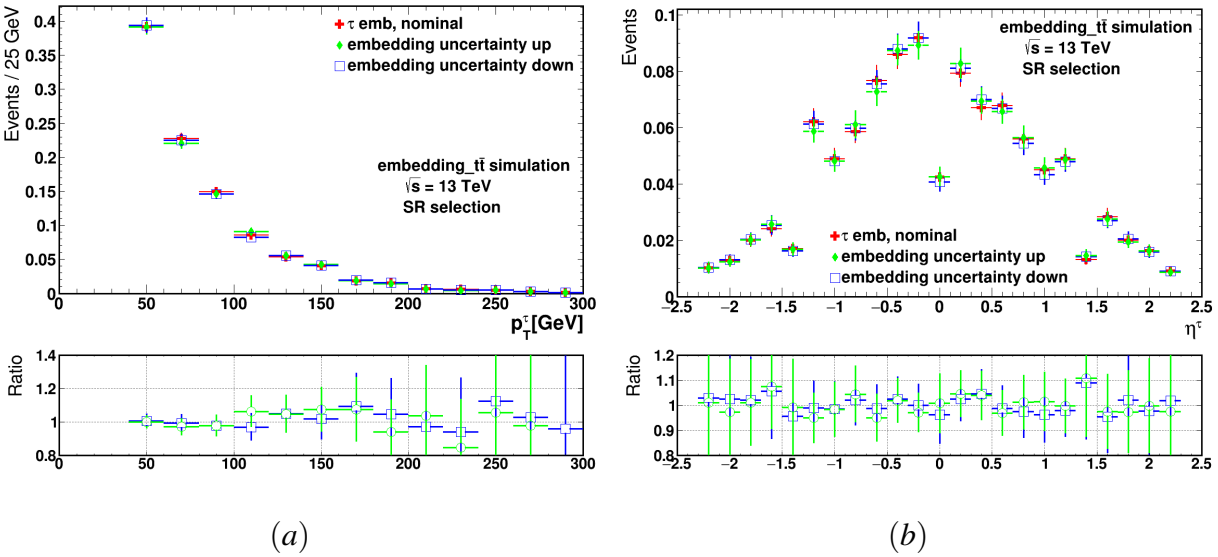



(e) $\mu$ *reco+trigger* uncertainty




(f) $E_T^{miss}$ *trigger* uncertainty

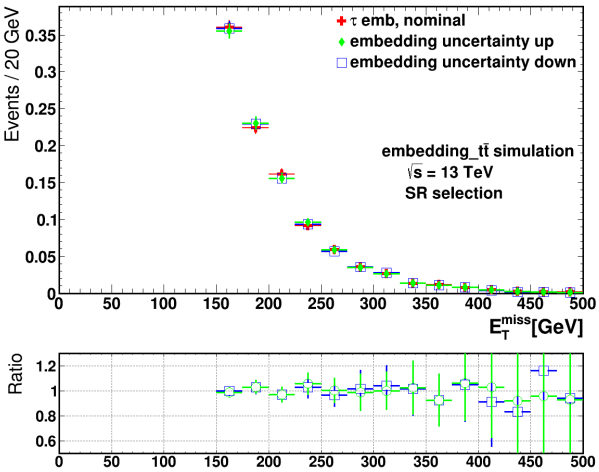
Figure 6.14: Effects of the (a) $\tau$ identification: *ID*, (b) $\tau$ reconstruction: *reco*, (c) $\tau$ energy scale: *TES*, (d) embedding: *emb*, (e) $\mu$ reconstruction and trigger uncertainties and (f) $E_T^{miss}$ trigger uncertainties on the shape of $m_T$ distribution and yield after applying the SR selection.



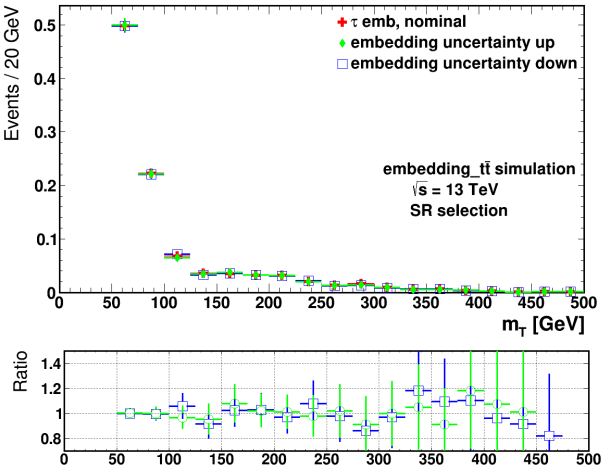

(a)



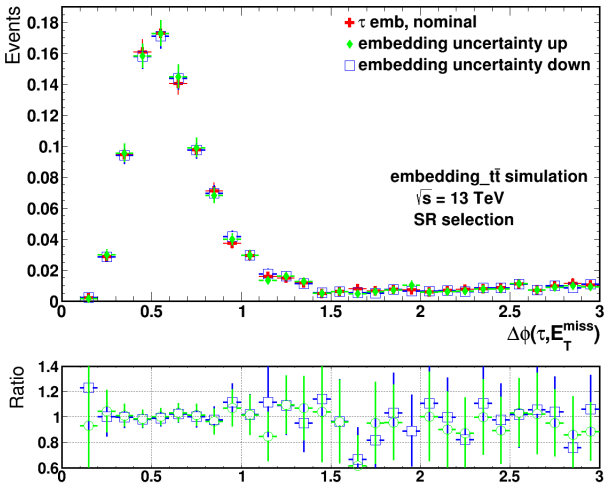

(b)




(c)




(d)




(e)

Figure 6.15: Effect of the embedding uncertainties on the shape of (a) $p_T$, (c) $E_T^{miss}$, (b) $\eta$, (d) $m_T$ and (e) $\Delta\phi$ distributions after applying the SR selection for embedded $t\bar{t}$ MC simulation.

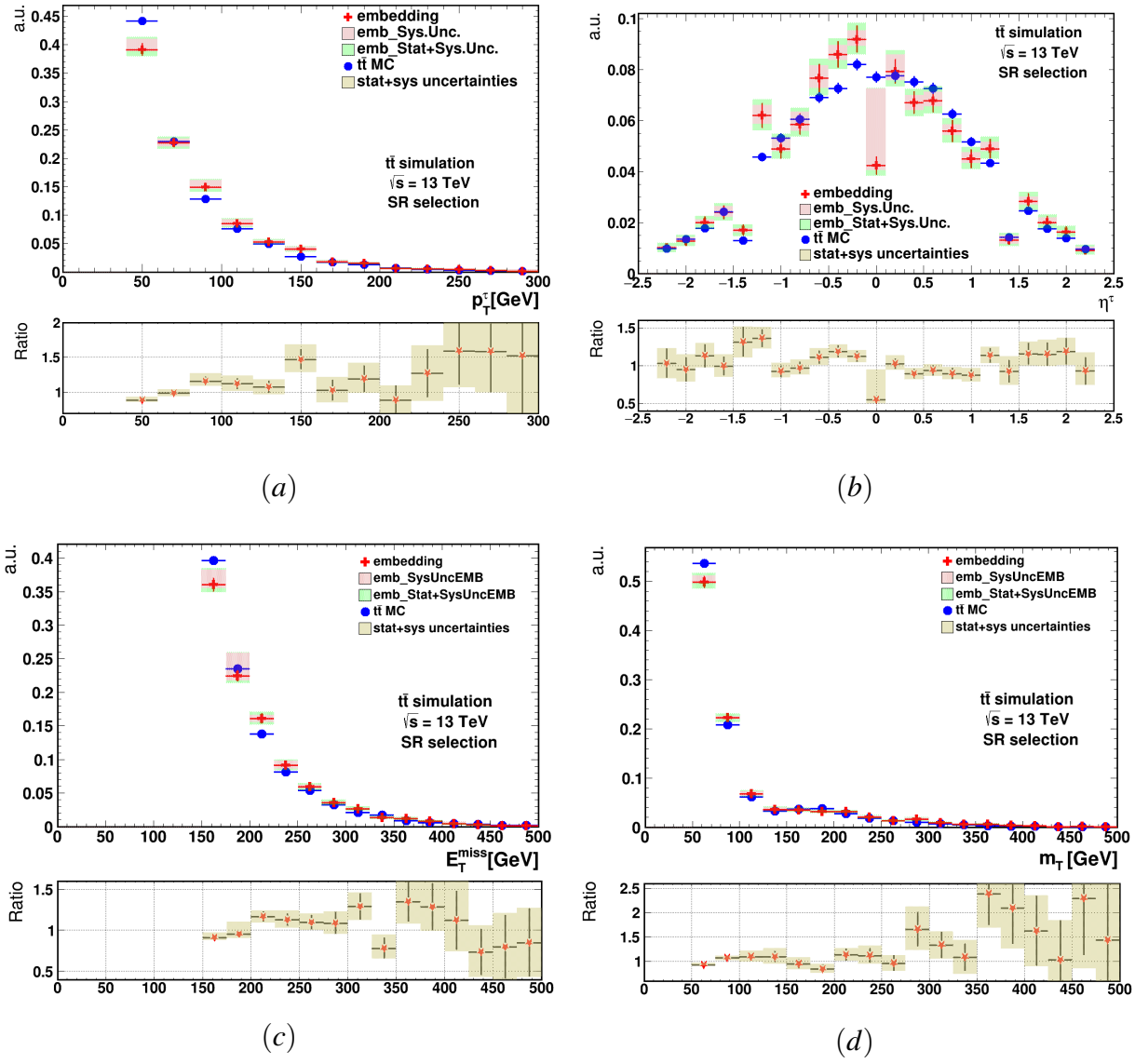
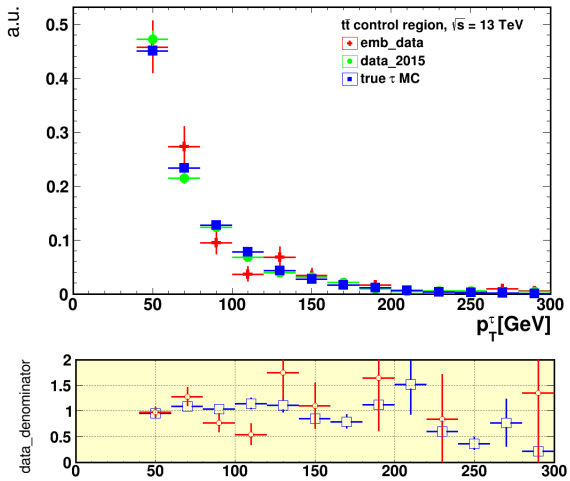


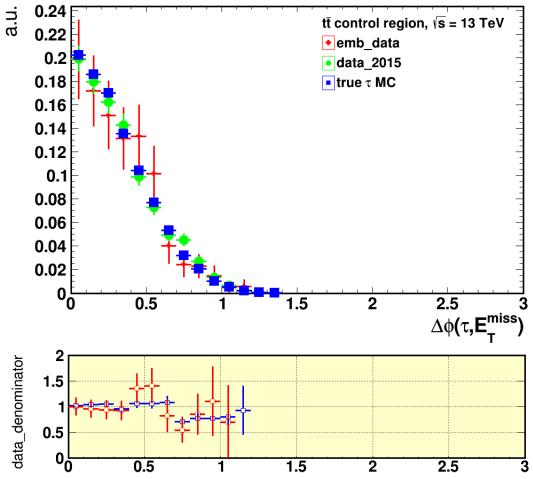

Figure 6.16: Juxtaposition of (a) $p_T^\tau$, (b) $\eta^\tau$, (c) $E_T^{miss}$ and $m_T$ of default and MC embedded $t\bar{t}$ simulation after applying the SR selection with all corrections and systematic uncertainties related to the embedding procedure. The hatched red areas indicate the embedding-related systematic uncertainties, whereas the hatched green areas indicate the embedding-related statistical and systematic uncertainties.

### 6.5.5 The final step: embedded data in action

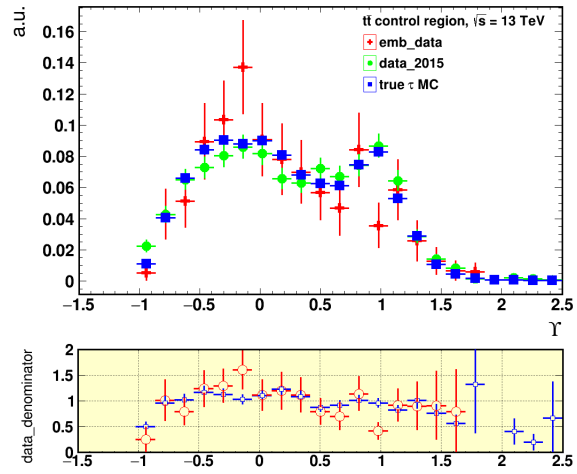
Fig. 6.17 shows the comparison of $p_T$, $\Delta\phi(\tau, E_T^{miss})$, $\Upsilon$ and $m_T$ distributions obtained for real and embedded data and true $\tau$ leptons from MC simulations. These distributions are obtained in enriched with $t\bar{t}$ events CR where the event selection is an analogue to the SR selection, except requiring $m_T < 100$ GeV.



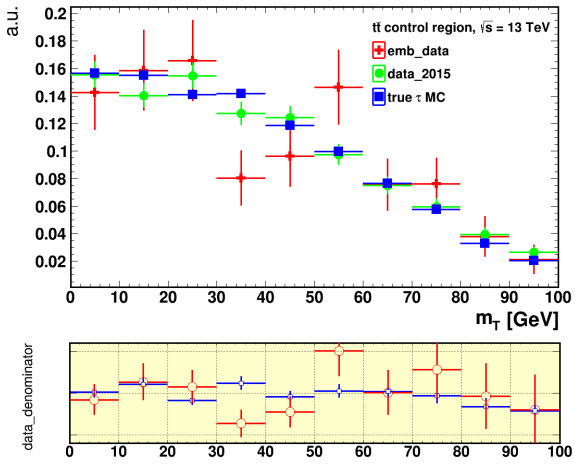

$(a)$ $p_T^\tau$ in $t\bar{t}$ enriched region




$(b)$ $\Delta\phi(\tau, E_T^{miss})$ in $t\bar{t}$ enriched region




$(c)$ $\Upsilon$ in $t\bar{t}$ enriched region




$(d)$ $m_T$ in $t\bar{t}$ enriched region

Figure 6.17: Juxtaposition between 2015 data, embedding data and true $\tau$ leptons from MC simulations of following distributions: (a) $p_T^\tau$, (b) $\Delta\phi(\tau, E_T^{miss})$, (c) $\Upsilon$ and (d) $m_T$, in $t\bar{t}$ enriched CR where 2015 data are always on denominator in ratio plots. Only statistical uncertainties are shown.

The resulting distributions demonstrate that the modeling of the embedded data is in an agreement within the statistical uncertainty with 2015 data and MC simulations.

As stated at the beginning of this Chapter, the main goal of this analysis has been to find an answer if a proper implementation of the embedding technique for Run-2 data will give a rise to the background estimation for selected events with true hadronically decaying $\tau$ leptons with significantly smaller systematic uncertainties than the corresponding background estimation from MC simulations. The answer is positive, i.e as it had been proven for the Run-1 2011 and 2012 data [39], also for Run-2 2015 data the embedding method presents significant improvement in the precision of the background estimation with true hadronically decaying $\tau$ leptons, as is presented in Fig. 6.18. The resulting distributions are in a good agreement within the combined statistical and systematic uncertainties, where visible fluctuations are a result of a very limited number of events available for the analysis.

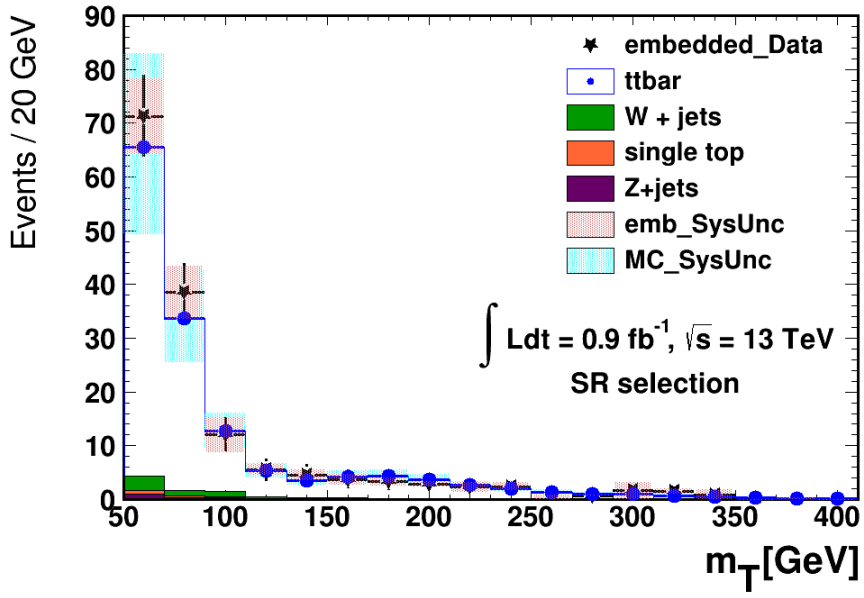



Figure 6.18: Transverse mass, $m_T$, distribution of the $\mu \to \tau$ embedded 2015 data sample for events with a true $\tau_{\text{had-vis}}$ compared to the default MC simulations with true $\tau$ lepton decaying hadronically after applying the SR selection. Systematic uncertainties are shown for the embedded data in hatched red bands and in hatched light blue bands for MC simulations.

### 6.5.6 Conclusion

For the background estimation, with true hadronically decaying $\tau$ leptons, using MC simulations the dominant detector-related systematic uncertainties arise from the jet energy scale, the reconstruction and identification of $\tau_{\text{had-vis}}$ candidates, the $\tau_{\text{had-vis}}$ energy scale and from the b-tagging efficiency. Their impacts on the predicted event yield for the dominant background $t\bar{t}$ process are, 11.0%, +5.7/-3.3%, 3.6% and 1.4%, respectively [88]. Which together with the uncertainty[4] on the final $t\bar{t}$ shape and yield due to the final state radiation (FSR) of 7%, the choice due to the matrix-element (ME) generator 15%, and due to the parton shower and underlying event (PSUE) 16% [87] gives 26% variation on the yield of the true $\tau$ background.

Whereas, *the embedding method* gives 11% variation on the yield of the true $\tau$ background.

Impact of systematic uncertainties on the embedded data and MC reference samples (for the dominant background $t\bar{t}$ process) in the SR is presented in Table 6.2.

| Uncertainty | embedded samples | reference (MC) samples |
|---|---|---|
| $\tau$ ID, $\tau$ RECO | $+5.2\%, -5.1\%$ | $+5.7/-3.3\%$ |
| TES | $+2.4\%, -7.7\%$ | $\pm 3.6\%$ |
| $E_T^{miss}$ trigger | $\pm 3.2\%$ | $\pm 3.0\%$ |
| $\mu$ ID, trigger | $+2.7\%, -3.0\%$ | - |
| embedding | $\pm 3.9$ | - |
| jet energy scale | - | $\pm 11\%$ |
| b-jet tag efficiency | - | $\pm 1.5\%$ |
| $t\bar{t}$ cross section | - | 6% |
| FSR | - | 7% |
| ME | - | 15% |
| PSUE | - | 16% |

Table 6.2: Impact of systematic uncertainties on the embedded data and MC reference samples [87] in the SR: $\tau$ lepton identification and reconstruction ($\tau$ ID, $\tau$ RECO), (TES) tau energy scale, ($E_T^{miss}$) missing transverse energy trigger, ($\mu$ ID, trigger ) muon identification and trigger, (embedding) uncertainty related with embedding method, final state radiation (FSR), matrix-element (ME) generator and parton shower, and underlying event (PSUE). All variations given are percentages on the yield of the true $\tau$ background.

[4] Systematic uncertainty due to the $t\bar{t}$ production cross section causing a difference of around 6% on the event yield is not included.

In the presented analysis the available number of events on which the embedding procedure is implemented and studied is very limited and insufficient to infer any meaningful limits on the observations of the charged Higgs boson. However, studies how the improvement in systematic uncertainties from using embedding method translates to the improvement of the limits for the observations of the charged Higgs has been performed for Run-1 data analysis and are briefly discussed in Appendix B.

### 6.5.7 $\Upsilon$ modeling for embedded data

As it is discussed in Chapter 4, at low $H^{\pm}$ masses the kinematics of the $t \to bH^{\pm}$ and $t \to bW^{\pm}$ decay products can be very similar. In such a case, the $\tau$ lepton polarisation, represented by the $\Upsilon$ variable, turned out to be crucial as a discriminating observable between signal and background processes. Therefore, it has been worthwhile to study modeling of the $\Upsilon$ variable also for the embedding method implementation. The result is shown in Fig. 6.19 for the $\mu \to \tau$ embedded 2015 data sample of events with a true 1-prong $\tau_{\text{had-vis}}$ and compared to default simulation after applying the SR selection of $H^{\pm}$ search. The resulting distributions of the $\Upsilon$ observable, sensitive to the $\tau$ polarisation, are in a reasonable agreement within combined statistical and systematic uncertainties. Where visible fluctuations are a result of a very limited number of events available for the analysis. Hence, $\Upsilon$ variable can be used with current embedding implementation.

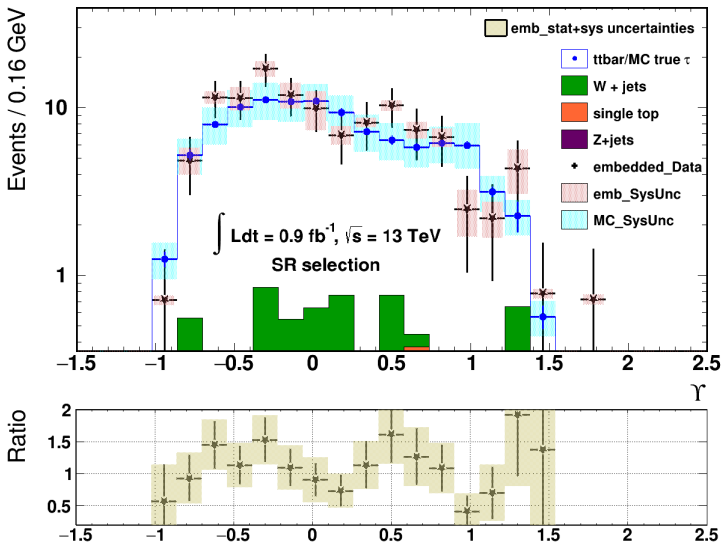



Figure 6.19: $\Upsilon$ distributions in logarithmic scale of the $\mu \to \tau$ embedded 2015 data sample of events with a true 1-prong $\tau_{\text{had-vis}}$ compared to default simulation after applying the SR selection.

# Chapter 7

# Statistical Interpretation of Results

## 7.1 The Likelihood Function and Test Statistic

The compatibility of the data with predictions for the signal and background rates is estimated using hypothesis tests with the data tested against the signal+background hypothesis. The statistical interpretation presented in this thesis is based on a simultaneous fit of the parameter of interest (or signal strength) $\mu \equiv \sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ by means of the negative log-likelihood minimisation. The binned likelihood function $\mathscr{L}(\mu, \theta)$ is constructed as a product of the Poissonian probability terms over all bins and regions considered in the analysis [144]

$$L(\mu, \theta) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} e^{-(\mu s_i + b_i)}. \tag{7.1}$$

Assuming a certain value of the signal cross section $\sigma_s$ the expected number of signal events given a signal-strength parameter $\mu$ is denoted by $\mu s_i$ (where $\mu = 0$ results in the background-only hypothesis) and $b_i$ denotes the number of the expected background events in i-th bin. Statistical and systematic uncertainties are incorporated into the likelihood via a vector of the so-called *nuisance parameters* $\theta$. Thus, the expected number of events in a given bin depends on the signal strength and on $\theta$. The latter allows some variations of the expected signal and background event yields according to the corresponding uncertainties, and the fitted values of the nuisance parameters correspond to the deviation from the nominal expectations that give the best fit to data.

The ratio of profiled log-likelihoods to test a hypothetical signal strength $\mu$ is given by

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})}, \tag{7.2}$$

where $\hat{\hat{\theta}}$ represents the value of $\theta$ that maximizes $L$ for a specific signal-strength parameter

$\mu$ (called: *the conditional maximum-likelihood estimator of* $\theta$), and $\hat{\mu}$, $\hat{\theta}$ are the likelihood estimators for the maximized unconditional likelihood function.

The test statistic is defined as the profile likelihood ratio:

$$\tilde{q}_\mu = \begin{cases} -2\log\frac{\mathscr{L}(\mu,\hat{\hat{\theta}}(\mu))}{\mathscr{L}(0,\hat{\hat{\theta}}(0))}, & \hat{\mu} < 0 \\ -2\log\frac{\mathscr{L}(\mu,\hat{\hat{\theta}}(\mu))}{\mathscr{L}(\hat{\mu},\hat{\theta})}, & 0 \leq \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{7.3}$$

Here, $\hat{\mu}$ and $\hat{\theta}$ are the values of the parameters that maximise the likelihood function, while $\hat{\hat{\theta}}(\mu)$ are the values of the nuisance parameters that maximise the likelihood function for a given value of the signal strength $\mu$ (as stated above). Data where $\hat{\mu} > \mu$ are not considered as less compatible with $\mu$ than the measured data. Thus there are not included in the rejection region of the test. The greater the value $\tilde{q}_\mu$ is the greater is the incompatibility of data and hypothetical value of $\mu$. This test statistic is used to measure the compatibility of the data with the signal+background hypothesis for the signal strength equal to $\mu$. The agreement of the data and the presumed signal-strength parameter $\mu$ is represented by the $p$-value which is estimated by integrating the distribution of $\tilde{q}_\mu$ above the observed value $\tilde{q}_{\mu,\text{obs}}$

$$p_\mu = \int_{\tilde{q}_{\mu,\text{obs}}}^{\infty} f(\tilde{q}_\mu|\mu)dq_\mu, \tag{7.4}$$

where $f(\tilde{q}_\mu|\mu)dq_\mu$ is the probability density function of $\tilde{q}_{\mu,\text{obs}}$ assuming $\mu$.

Throughout the statistical analysis, also in order to derive the *expected limits*, the asymptotic approximation is used [144]. An artificial data set, the so-called *Asimov data set*, is defined such that using it to evaluate estimators for all parameters yields the best estimates of these parameters. The Asimov likelihood is defined by

$$\lambda_A(\mu) = \frac{L_A(\mu,\hat{\hat{\theta}})}{L_A(\hat{\mu},\hat{\theta})} = \frac{L_A(\mu,\hat{\hat{\theta}})}{L_A(\mu',\hat{\theta})}, \tag{7.5}$$

where $\mu'$ is the mean of $\hat{\mu}$, which follows Gaussian distribution. The Asimov data set satysfies

$$-2\log\lambda_A(\mu) \approx \frac{(\mu-\mu')^2}{\sigma^2}, \tag{7.6}$$

where $\mu' = \hat{\mu}$, which then may be used to derive the standard deviation $\sigma$ of the expected upper limit:

$$\sigma^2 = \frac{(\mu-\mu')^2}{-2\log\lambda_A(\mu)}. \tag{7.7}$$

Note that the *expected limits* are used to estimate the sensitivity of an experiment. They are given by the median limits assuming the background-only hypothesis

$$\sigma^2 = \frac{\mu^2}{-2\log\lambda_A(\mu)}. \tag{7.8}$$

The asymptotic approximation is also used to derive the $1\sigma$ and $2\sigma$ uncertainty bands.

## 7.2 The $CL_s$ Method

In situations where the measurements are conducted at the sensitivity threshold of the detector, it has been shown that the most robust method for determining the exclusion limits is the so-called $CL_s$ method [145]. It is defined by

$$CL_s = \frac{p_{s+b}}{1-p_b}, \tag{7.9}$$

where $p_{s+b}$ is the probability of finding a value of $\tilde{q}$ that is equally or less compatible with the signal+background hypothesis than the observed value $\tilde{q}_{\text{obs}}$ (see Eq. 7.4). The p-value of the background-only hypothesis, $p_b$, is defined as

$$p_b = \int_0^{\tilde{q}_{\mu,\text{obs}}} f(\tilde{q}_\mu|b)d\tilde{q}_\mu, \tag{7.10}$$

where $f(\tilde{q}_\mu|b)$ is the probability density function of the test-statistic for the background-only hypothesis. The signal+background hypothesis is excluded at a given confidence level by using $CL_s$, typicaly taken to be 95%, if

$$CL_s \leq 0.05. \tag{7.11}$$

Thanks to the $CL_s$ method one can be sure that the signal is not incorrectly excluded based on fluctuations in the observed data - otherwise one could run into problems for searches with low sensitivity. In such cases, the value of $p_{s+b}$ can be very low and the usual CL method [124] could give an exclusion, even though this is purely an effect of low sensitivity and not the signal strength. In the $CL_s$ method the value of $1-p_b$ decreases together with $p_{s+b}$ preventing the exclusion condition $CL_s < 0.05$ from being met. Therefore, the method is conservative. It is worth noting that exclusion intervals from the $CL_s$ method had been shown to be in agreement with the ones from Bayesian methods for Poisson or Gauss distributions with a constant prior for its mean [145].

In the presented analysis upper (exclusion) limits on the signal production cross section are derived with the usage of $\text{CL}_s$ method. The statistical interpretation is based on a simultaneous fit of the parameter of interest (or signal strength), e.g. $\mu \equiv \sigma(pp \to tbH^+) \times BR(H^+ \to \tau^+\nu)$, and the nuisance parameters $\theta$ that encode statistical and systematic uncertainties. The three signal regions and one validation region are included in the simultaneous fit:

- in the three signal regions of the $\tau_{\text{had-vis}}$+jets, $\tau_{\text{had-vis}}$+$e$ and $\tau_{\text{had-vis}}$+$\mu$ channels, the BDT score distributions are used;

- in the control region of the $\tau_{\text{had-vis}}$+lepton channel enriched in $t\bar{t}$ events (the same event selection as the SR, but with the requirement of an $e\mu$ pair instead of the $e/\mu$+$\tau_{\text{had-vis}}$ pair), a one-bin BDT score distribution is used.

Where, for a given signal hypothesis, values of the production cross section times branching fraction $\mu \equiv \sigma(pp \to tbH^+) \times \text{BR}(H^+ \to \tau^+\nu)$ for the full mass range investigated, as well as on $\text{BR}(t \to bH^+) \times \text{BR}(H^+ \to \tau^+\nu)$ in the low $H^\pm$ mass range, yielding $\text{CL}_s < 0.05$ are excluded at the 95% confidence level.

## 7.3 Model-independent Limits

In Table 7.1 the expected number of events for all SM processes and the measured event yields in the SR are shown, prior to using the multi-variate discriminant and applying the statistical fitting procedure. The contribution from hypothetical charged Higgs bosons are also shown, assuming a mass of 170 or 1000 GeV, and with $\sigma(pp \to [b]tH^\pm) \times \text{BR}(H^\pm \to \tau^\pm\nu)$ set to the prediction from the hMSSM scenario [146, 147] for $\tan\beta = 40$, as computed using Refs. [31, 91, 148, 149] for the production cross-section and HDECAY [150] for the branching fractions. The signal acceptances for a charged Higgs boson mass hypothesis of 170 GeV are 0.9%, 0.6% and 0.5% in the signal regions of the $\tau_{\text{had-vis}}$+jets, $\tau_{\text{had-vis}}$+electron and $\tau_{\text{had-vis}}$+muon channels, respectively. They become 11.6%, 0.9% and 1.2% for a charged Higgs boson mass of 1 TeV. The event yields observed in 36 fb$^{-1}$ of data collected at 13 TeV are consistent with the expected SM backgrounds.

The BDT score distributions in the five charged Higgs boson mass ranges used in the presented analysis to distinguish between the hypothetical signals and the SM backgrounds, are shown in Fig. 7.1. All plots are obtained after the statistical fitting procedure with the background-only hypothesis, where the binning shown in the Figures is also used in the statistical analysis.

The data are found to be consistent with the background-only hypothesis. In light of this fact exclusion limits are set at the 95% confidence level, by using the $\text{CL}_s$ procedure, on $\sigma(pp \to [b]tH^\pm) \times \text{BR}(H^\pm \to \tau^\pm\nu)$ for the full investigated mass range, and also on $\text{BR}(t \to bH^\pm) \times \text{BR}(H^\pm \to \tau^\pm\nu)$ in the low $H^\pm$ mass range. The expected and observed exclusion limits as a function of the $H^\pm$ mass hypothesis are presented in Fig. 7.2. The observed limits range from 4.2 pb and 0.0025 pb over the mass range considered in this search. Between the explicitely considered $H^\pm$ mass regions the limits are interpolated linearly. The potential bias from such interpolation is found to be smaller than the statistical uncertainty. For the $H^\pm$ mass range between 90 and 160 GeV, if one assumes that the production cross-section is equal to that of $t\bar{t}$

| Sample | Event yields $\tau_{\text{had-vis}}$+jets |
|---|---|
| True $\tau_{\text{had}}$ | |
| $t\bar{t}$ & single-top-quark | $7700 \pm 60 \pm 1800$ |
| $W \to \tau\nu$ | $1050 \pm 30 \pm 180$ |
| $Z \to \tau\tau$ | $84 \pm 42 \pm 28$ |
| Diboson ($WW, WZ, ZZ$) | $63.2 \pm 4.6 \pm 7.2$ |
| Misidentified $e, \mu \to \tau_{\text{had-vis}}$ | $265 \pm 12 \pm 35$ |
| Misidentified jet $\to \tau_{\text{had-vis}}$ | $2370 \pm 20 \pm 260$ |
| All backgrounds | $11500 \pm 80 \pm 1800$ |
| $H^{\pm}$ (170 GeV), hMSSM $\tan\beta = 40$ | $1400 \pm 14 \pm 170$ |
| $H^{\pm}$ (1000 GeV), hMSSM $\tan\beta = 40$ | $10.0 \pm 0.1 \pm 0.6$ |
| Data | 11021 |

Table 7.1: Expected event yields for the backgrounds and a hypothetical $H^{\pm}$ signal after all $\tau_{\text{had-vis}}$+jets selection criteria, and comparison with 36 fb$^{-1}$ of data. All yields are evaluated prior to using the multi-variate discriminant and applying the statistical fitting procedure. The values shown for the signal assume a charged Higgs boson mass of 170 or 1000 GeV, with $\sigma(pp \to [b]tH^{\pm}) \times \text{BR}(H^{\pm} \to \tau^{\pm}\nu)$ corresponding to $\tan\beta = 40$ in the hMSSM benchmark scenario. Statistical and systematic uncertainties are quoted, respectively. From [12].

pairs, this translates into observed limits for the branching fraction $\text{BR}(t \to bH^{\pm}) \times \text{BR}(H^{\pm} \to \tau^{\pm}\nu)$ ranging from 0.25% to 0.031%.

In Fig. 7.3 the final upper limits obtained for the $\tau_{\text{had-vis}}$+jets channel are also shown. As can be seen the sensitivity of the combined analysis at high $H^{\pm}$ mass is driven by the $\tau_{\text{had-vis}}$+jets channel.

The importance of the various sources of systematic uncertainty is estimated by comparing the expected 95% CL limits on $\sigma(pp \to [b]tH^{\pm}) \times \text{BR}(H^{\pm} \to \tau^{\pm}\nu)$ when taking only statistical uncertainties into account to those obtained when a certain set of systematic uncertainties is added in the limit-setting procedure. This is summarised in Table 7.2 for the $H^{\pm}$ masses of 170 GeV and 1000 GeV.

Fig. 7.4 shows the 95% CL exclusion limits on $\tan\beta$ as a function of $H^{\pm}$ mass assuming the hMSSM scenario. For $m_{H^{\pm}} \leq 160$ GeV all $\tan\beta$ values are excluded. At $\tan\beta = 60$, above which no reliable theoretical predictions exist, the charged Higgs boson mass range up to 1100 GeV is excluded, hence significantly improving previous results based on the dataset collected in 2015 with an integrated luminosity of 3.2 fb$^{-1}$.

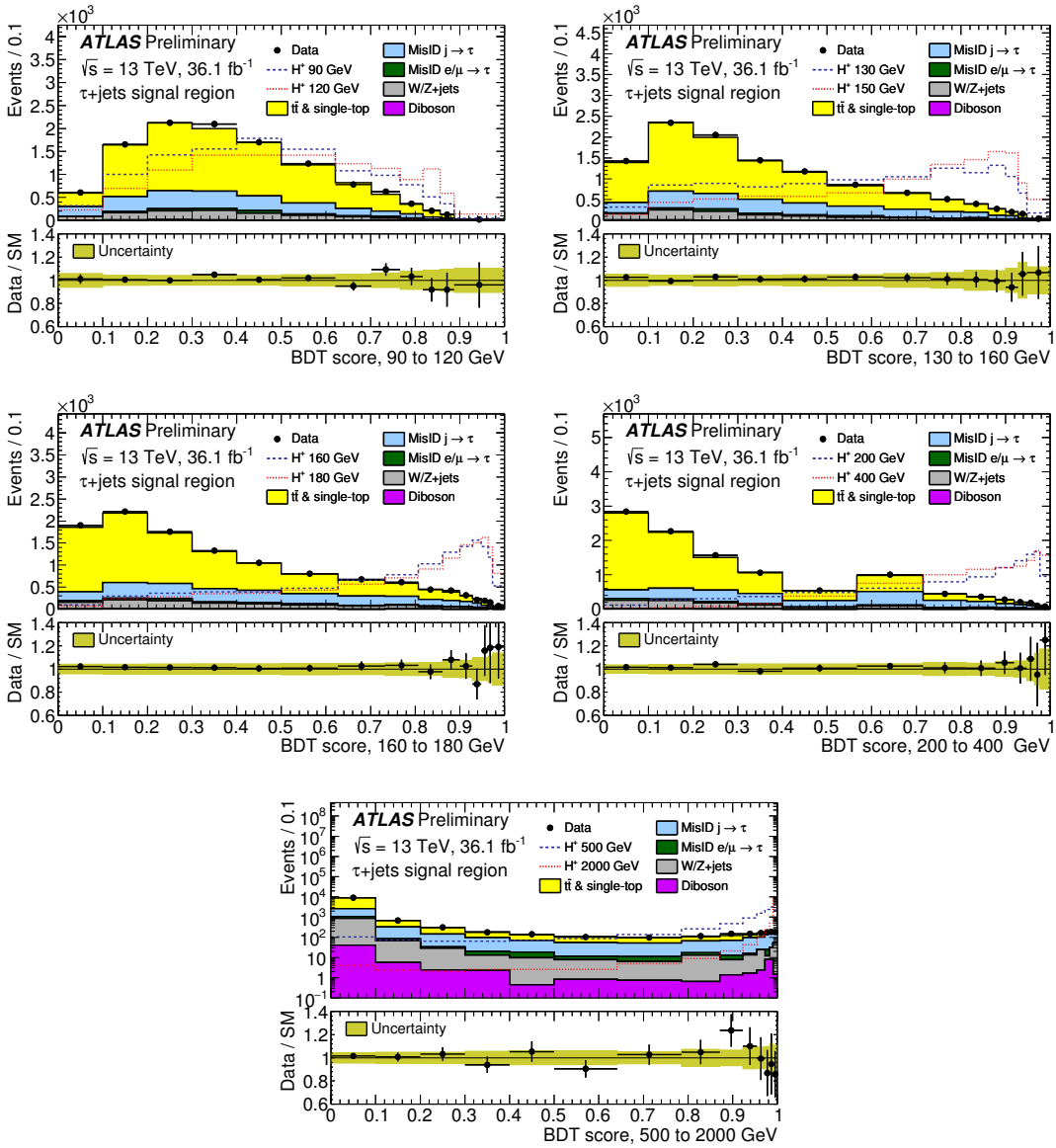



Figure 7.1: BDT score distributions in the five mass ranges used for the BDT trainings, after a fit to the data with the background-only hypothesis. The uncertainty bands in the ratio plots include both the statistical and systematic uncertainties. The normalisation of the signal (shown for illustration) corresponds to the integral of the background. From [12].

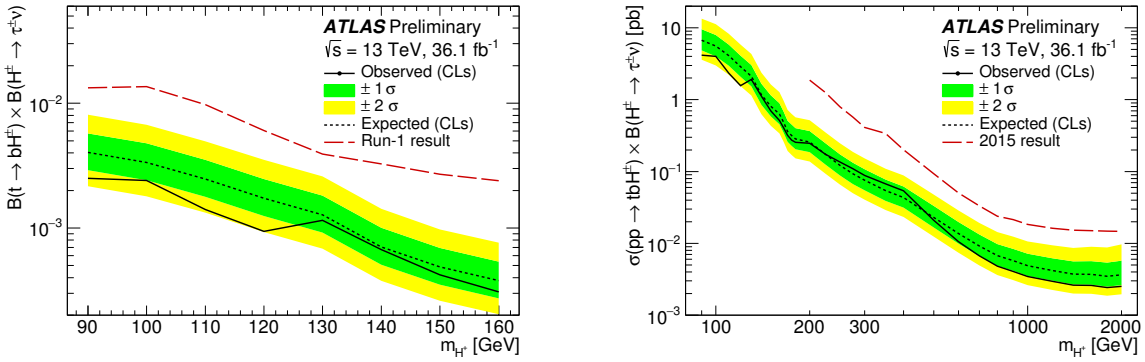



Figure 7.2: Observed and expected 95% CL exclusion limits on (right) $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{B}(H^{\pm} \to \tau^{\pm}\nu)$ and (left) $\mathrm{B}(t \to bH^{\pm}) \times \mathrm{B}(H^{\pm} \to \tau^{\pm}\nu)$ as a function of the charged Higgs boson mass in 36.1 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 13$ TeV, after combination of the $\tau_{\text{had-vis}}$+jets and $\tau_{\text{had-vis}}$+lepton final states. In the case of the expected limits, one- and two-standard-deviation uncertainty bands are also shown. In the $H^{\pm}$ mass range of 90–160 GeV, the limits obtained with the Run-1 data [39] are also shown. From [12].

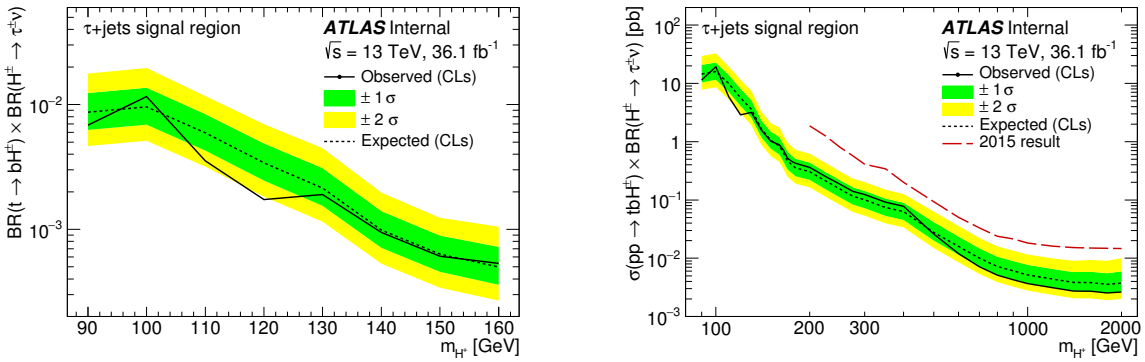



Figure 7.3: Observed and expected 95% CL exclusion limits on $\mathrm{BR}(t \to bH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ (left) and $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ (right) for charged Higgs boson production as function of $m_{H^{\pm}}$ in 36.1 fb$^{-1}$ of $pp$ collision data at $\sqrt{s} = 13$ TeV for the $\tau_{\text{had-vis}}$+jets channel. From [33].

| Source of systematic uncertainty | Impact on the expected limit (stat. only) in % | |
|---|---|---|
| | $m_{H^+} = 170\ GeV$ | $m_{H^+} = 1000\ GeV$ |
| Experimental | | |
| luminosity | 2.9 | 0.2 |
| trigger | 1.3 | <0.1 |
| $\tau_{\text{had-vis}}$ | 14.6 | 0.3 |
| jet | 16.9 | 0.2 |
| electron | 10.1 | 0.1 |
| muon | 1.1 | <0.1 |
| $E_T^{miss}$ | 9.9 | <0.1 |
| Fake-factor method | 20.3 | 2.7 |
| $\Upsilon$ modelling | 0.8 | – |
| Signal and background models | | |
| $t\bar{t}$ modelling | 6.3 | 0.1 |
| $W/Z$+jets modelling | 1.1 | <0.1 |
| cross-sections $(W/Z/VV/t)$ | 9.6 | 0.4 |
| $H^+$ signal modelling | 2.5 | 6.4 |
| All | 52.1 | 13.8 |

Table 7.2: Impact of systematic uncertainties on the expected 95% CL limit on $\sigma(pp \to [b]tH^{\pm}) \times \text{BR}(H^{\pm} \to \tau^{\pm}\nu)$, for two $H^{\pm}$ mass hypotheses: 170 and 1000 GeV. The impact is obtained by comparing the expected limit considering only statistical uncertainties (stat. only) with the expected limit when a certain set of systematic uncertainties is added in the limit-setting procedure. From [12].

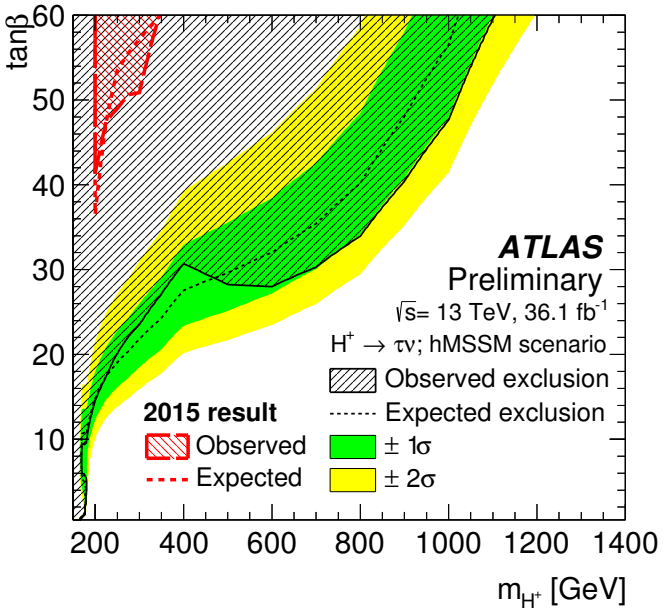



Figure 7.4: 95% CL exclusion limits on $\tan\beta$ as a function of $m_{H^\pm}$, shown in the context of the hMSSM, for the regions in which theoretical predictions are available ($0.5 \leq \tan\beta \leq 60$). In the case of the expected limits, one- and two-standard-deviation uncertainty bands are also shown. As a comparison, the red curves show the observed and expected exclusion limits based on the dataset of 3.2 fb$^{-1}$ collected in 2015 at $\sqrt{s} = 13$ TeV. From [12].

# Chapter 8

# Conclusions

In this thesis search whole mass range of the charged Higgs boson decaying via $H^{\pm} \to \tau^{\pm}\nu$ is considered. Performed investigations are based on 36 fb$^{-1}$ of data taken with the ATLAS detector at center-of-mass energy of $\sqrt{s} = 13$ TeV in the years 2015 and 2016. The studied final states containing a hadronically decaying $\tau$ lepton, the missing transverse energy $E_T^{miss}$ due to neutrinos and jets including at least one b-tagged jet arise from the processes:

$$t\bar{t} \to (W^{\mp}\bar{b})(bH^{\pm}) \to (q\bar{q}'\bar{b})(b\tau^{\pm}_{had}\nu)$$

for the search of $H^{\pm}$ in the $t\bar{t}$ production channel followed by the top quark decay, which is only open for $m_{H^{\pm}} \leq m_{top}$. In the case of the search for $H^{\pm}$ in the top quark associated production channel the considered processes are:

$$g\bar{b} \to \bar{t}H^{+} \to (W^{-}\bar{b})H^{+} \to (q\bar{q}'\bar{b})(\tau^{\pm}_{had}\nu)$$
$$gb \to tH^{-} \to (W^{+}b)H^{-} \to (q\bar{q}'b)(\tau^{\mp}_{had}\nu)$$

in the five flavour scheme and

$$gg \to \bar{t}bH^{\pm} \to (W^{-}\bar{b})bH^{\pm} \to (q\bar{q}'\bar{b})b(\tau^{\pm}_{had}\nu)$$

in the four flavour scheme case where the final state is the same as for low-mass search.

In previous analysis searching for charged Higgs bosons produced in association with a top-quark and decaying via $H^{\pm} \to \tau^{\pm}\nu$ channel the transverse mass of the highest-$p_T$ $\tau_{\text{had-vis}}$ candidate and $E_T^{miss}$ had been used as a final discriminating variable between signal and background in the cut based analysis. As an improvement to these former searches a multivariate analysis is applied and used in the analysis described in presented thesis to separate the $H^{\pm}$ signal from the Standard Model background processes. The output score of Boosted Decision Trees is used as the final discrimination variable in statistical analysis of the results. Moreover,

since at low $H^{\pm}$ masses i.e. between the $W$-boson and top-quark masses, the kinematics of the $t \to bH^{\pm}$ and $t \to bW^{\pm}$ decay products can be very similar, the polarisation of the $\tau$ lepton is used as a discriminating variable. In particular, the decay $W^{\pm} \to \tau^{\pm}\nu$ leads to $\tau$ leptons with polarisation -1, while decay of the charged Higgs boson, which couples to the only left-handed neutrino, $H^{\pm} \to \tau^{\pm}\nu$ to polarisation +1. In addition, an important advantage of using the $\tau$ polarimetry arises from the fact that its events signal distribution is independent of the $H^{\pm}$ mass. This is especially important at low and intermediate (160 - 200 GeV) $H^{\pm}$ masses searches, where it is difficult to distinguish transverse mass coming from $W^{\pm}$ from the one from $H^{\pm}$.

Comparison of the background estimates and data shows that the obtained results are in agreement with the background-only hypothesis. The model-independent limits set on the production cross-section times branching fraction, $\sigma(pp \to [b]tH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$, are between 4.2 pb and 0.0025 pb for whole charged Higgs boson mass range (90–2000 GeV). This corresponds to upper limits between 0.25% and 0.031% for the branching fraction $\mathrm{BR}(t \to bH^{\pm}) \times \mathrm{BR}(H^{\pm} \to \tau^{\pm}\nu)$ in the mass range 90–160 GeV. In the context of the hMSSM scenario it is found that all $\tan\beta$ values are excluded for $m_{H^{\pm}}$ below 160 GeV, whereas the $H^{\pm}$ mass range up to 1100 GeV is excluded at $\tan\beta = 60$.

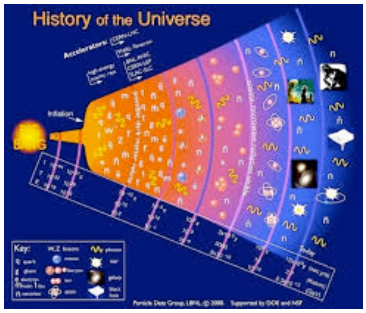
Additional studies have been carried out in attempt to find the best set of BDT discriminants. It is demonstrated that there is some space for optimisation of variable sets and mass ranges with respect to the choice taken for presented analysis. Results of those studies can be used for full Run-2 data search.

Another important part of the presented thesis has been to find an answer whether a proper implementation of the *embedding technique* for Run-2 data will give a rise to the background estimation for selected events with true hadronically decaying $\tau$ leptons; whether this will lead to significantly smaller systematic uncertainties than the corresponding background estimation coming from Monte Carlo simulations. The embedding method heavily uses the lepton universality of the $W$ boson decay by constructing hybrid background samples with simulated hadronically decaying $\tau$ leptons based on measured and selected $W^{\pm} \to \mu^{\pm}\nu$ data events. The decay products of the $\tau$ lepton are then merged with the original event except for the muon. The main advantage of the technique is that one relies on the simulations only for the well understood electroweak decays of $W$ boson and $\tau$ lepton decays and all the other aspects of a given event are taken directly from the data. Therefore, many systematic uncertainties associated with the simulations do not have to be considered, especially the theoretical $t\bar{t}$ cross section and its uncertainties. To study the effects of removing the original muon and replacing it by the simulated $\tau$ the $\mu \to \tau$ embedded simulation is compared to default simulation with hadronically decaying $\tau$ leptons. Many tests and cross-checks have been performed all being successful.

In the end the $\mu \to \tau$ embedded data are compared to simulations resulting a positive answer, i.e as it had been proven in the 2011 and 2012 Run-1 data, for 2015 data in the Run-2 the embedding method also presents significant improvement in the precision of background estimation with true hadronically decaying $\tau$ leptons, resulting in $\pm 11\%$ variation on the yield. This number has to be compared to $\pm 26\%$ given by the background estimation using the MC simulations and including the uncertainties on the final $t\bar{t}$ shape and yield.

### Coda

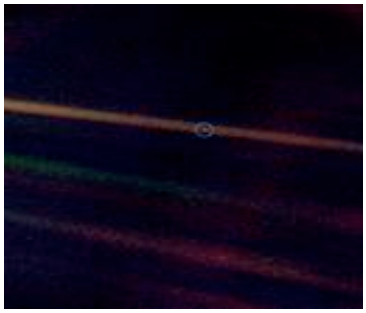
A discovery of a charged Higgs boson would be an undeniable proof of the existence of the physics beyond the Standard Model. It would also show us a way we should follow in the quest of more deep understanding of the fundamental rules governing our Universe.




The evolution of the Universe depends on its

matter composition and fundamental interactions

*Pale Blue Dot: The Earth from a distance of about 6. 4 bln km*



"*It has been said that astronomy is a humbling and character-building experience. There is perhaps no better demonstration of the folly of human conceits than this distant image of our tiny world. To me, it underscores our responsibility to deal more kindly with one another, and to preserve and cherish the pale blue dot, the only home we have ever known.*" Carl Sagan, "Pale Blue Dot: A Vision of the Human Future in Space".

# Appendix A

# Collisions at Hadron Colliders and PDFs

At the elementary level not the protons are colliding but their constituents i.e. the gluons $(g)$, the quarks $(q)$ and anti-quarks $(\bar{q})$, often also referred to as *partons*. The protons consists of three valence quarks: two up quarks $(u)$ and one down quark $(d)$, and a sea of quark-anti-quark pairs held together by gluons. Typically, the quarks and gluons will collide with relatively small energies, the so-called *soft collisions*, whereas we are interested in the *hard collisions*, i.e. scattering between partons at high $Q^2$. Where $Q$ is the momentum transfer between scattering partons.

In the same $pp$ interaction in addition to hard scattering process, multiple soft interactions named *underlying event* may occur, having a non-negligible impact on the event. They are related to the final state parton interactions calculated by using phenomenological multi parton interactions model. The Fig. A.1 below presents a schematic view of a proton-proton collision.

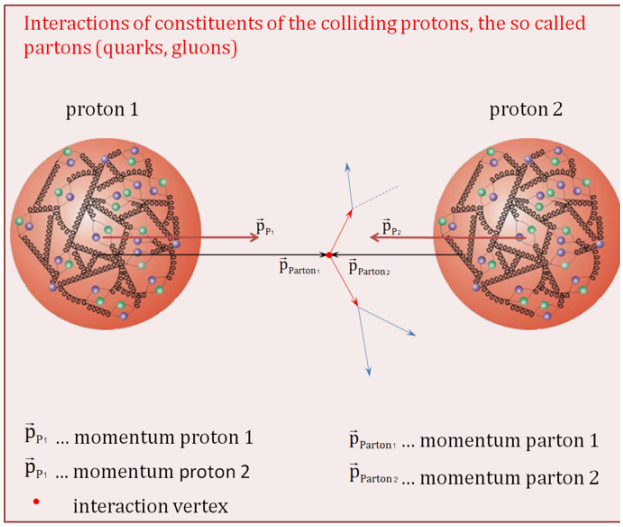
The distribution of momentum that the quarks and gluons carry is quantified by the parton distribution functions (PDFs), which can to the lowest order be described as the probability of finding a given parton with a specific momentum fraction inside the proton.

These parton distribution functions, $f_i(x,Q)$, are typically parameterised by $x$ which is the momentum fraction that the parton carries, $x = \frac{|\vec{p}_{parton}|}{|\vec{p}_{proton}|}$ e.g:

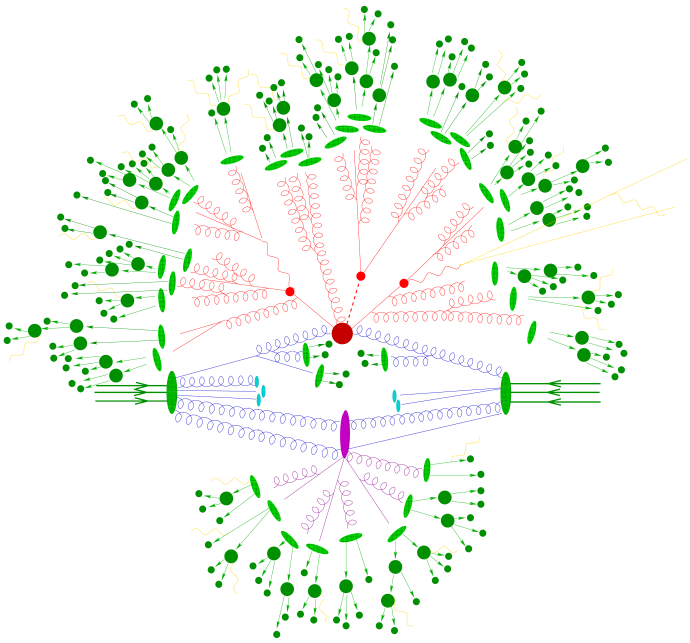
$$\sigma(pp \to WX) = \sum_{ij} \int f_i(x,Q) f_i(\bar{x},Q) \sigma(q_i \bar{q}_j \to W) dx d\bar{x} dQ, \tag{A.1}$$

where $X$ denotes an arbitrary particle.

The parton distribution functions for two different momentum transfers $Q = 2$ GeV and $Q = 100$ GeV are shown on Fig. A.2. It is visible that typically the dominant parton is the gluon which at the medium values of $x$ has an order of magnitude larger values of the PDF than the quarks. Since up and down quarks are the valence ones, for high values of $x$ they become to play the dominant role in parton distribution functions.




(a)



(b)

Figure A.1: (a) schematic overview of a $pp$ collision. (b) sketch of a hadron-hadron collision as simulated by a Monte-Carlo event generator. The red blob in the center represents the hard collision, surrounded by a tree-like structure representing Bremsstrahlung as simulated by parton showers. The purple blob indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation. From [59, 151].

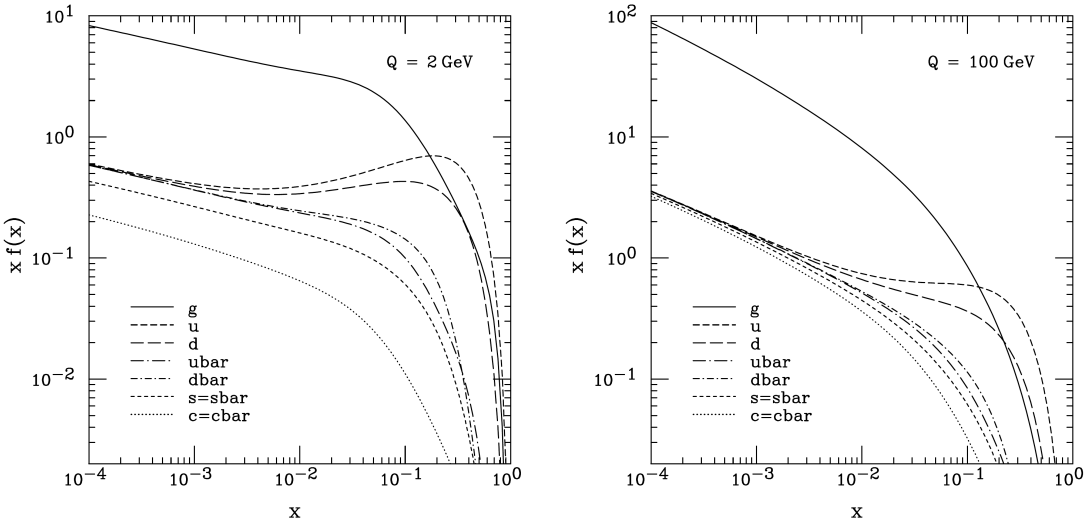
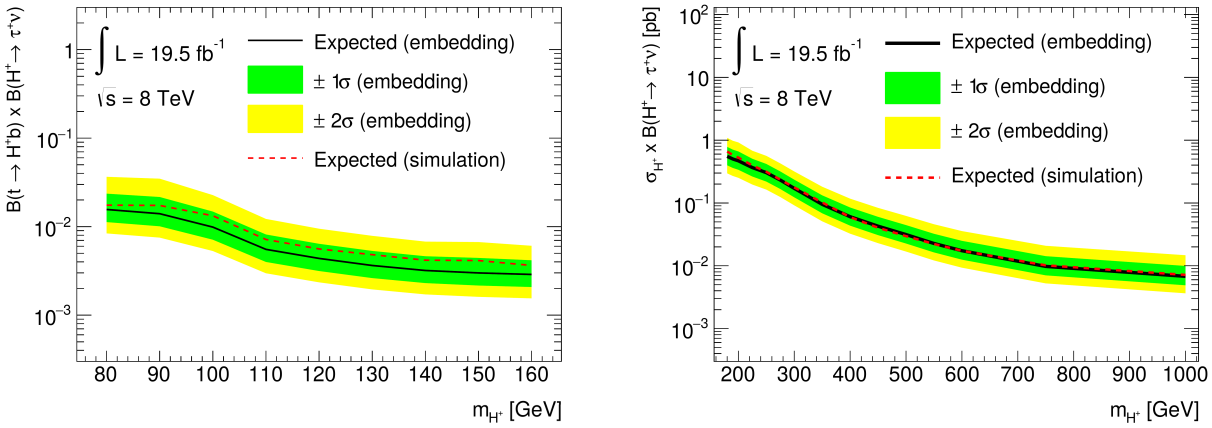



Figure A.2: Overview of the *CTEQ6M* parton distribution functions at $Q = 2$ and 100 GeV. From [102].

# Appendix B

# Comparison of the Expected Limits

In this thesis the available number of embedded events has been very limited and insufficient to use them for setting meaningful limits on observation (or exclusion of the existence) of charged Higgs boson. However, due to the fact that variations on the yield of the true $\tau$ background are of a similar significance as in the corresponding Run-1 analysis, one can expect also similar improvement of the limits. Therefore, in order to give a rough estimate of the significance of the embedding method for the final limits, the results from the Run-1 data [136, 137] are quoted here.




(a) low-mass $H^+$ search (b) high-mass $H^+$ search

Figure B.1: Expected limits on $\mathrm{B}(t \to bH^\pm) \times \mathrm{B}(H^\pm \to \tau^\pm \nu)$ for the low mass $H^\pm$ selection and $\sigma(pp \to \bar{t}(t)H^\pm + X) \times \mathrm{B}(H^\pm \to \tau^\pm \nu)$ for the high mass $H^\pm$ selection, for the 2012 data using embedding (solid black line) or simulation (dashed red line) for the background contribution with true $\tau_{had}$ for the (a) low-mass and (b) high-mass $H^\pm$ search. From [137].

Figure B.1 shows the comparison of the expected limits on the observation of the charged Higgs particle between two cases: when one uses the embedded data or MC simulation for the estimate of the background contribution with true hadronically decaying $\tau$ leptons.

For the low-mass $H^{\pm}$ search the background contribution with true $\tau_{had}$ is the most important one. Since the systematic uncertainties on the embedded data are smaller there than those on simulation, therefore the expected limits using the embedded data are stronger than those using MC simulation.

In the high $H^{\pm}$ masses case, where the contribution of background events with jets misidentified as $\tau_{\text{had-vis}}$ candidates is more important one does not observe improvements in the limits using the embedded data compared to MC simulation.

# Appendix C

# Additional results for MVA

Distributions of the chosen set of variables, i.e. $p_T^{\tau}$, $p_T^{b\text{-jet}}$, $E_T^{miss}$, $\Delta\phi_{\tau,E_T^{miss}}$, $\Delta\phi_{b\text{-jet},E_T^{miss}}$, $\Delta R_{b\text{-jet},\tau}$ and $\Upsilon$, as used for *default* BDT-training in the presented analysis for the three mass bins, i.e. 130-160, 160-180 and 500-2000 GeV (see Section 5.2) are shown below.

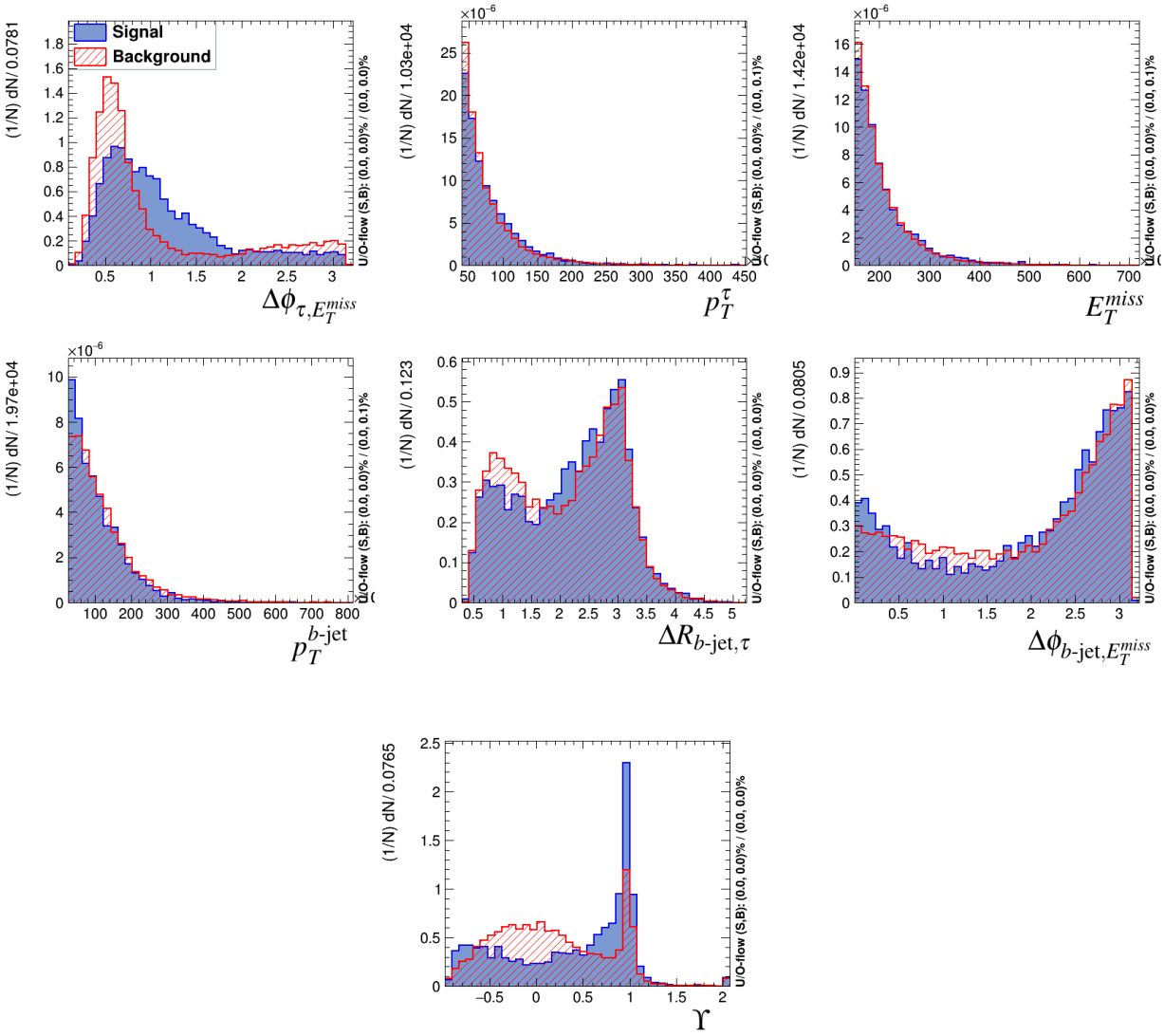



Figure C.1: Distributions of input variables for BDT-training within 130-160 GeV mass bin. Signal is presented as a blue, solid histogram; top background (as estimated from MC) is presented as a red, dashed histogram. Histograms are arbitrarily normalised.

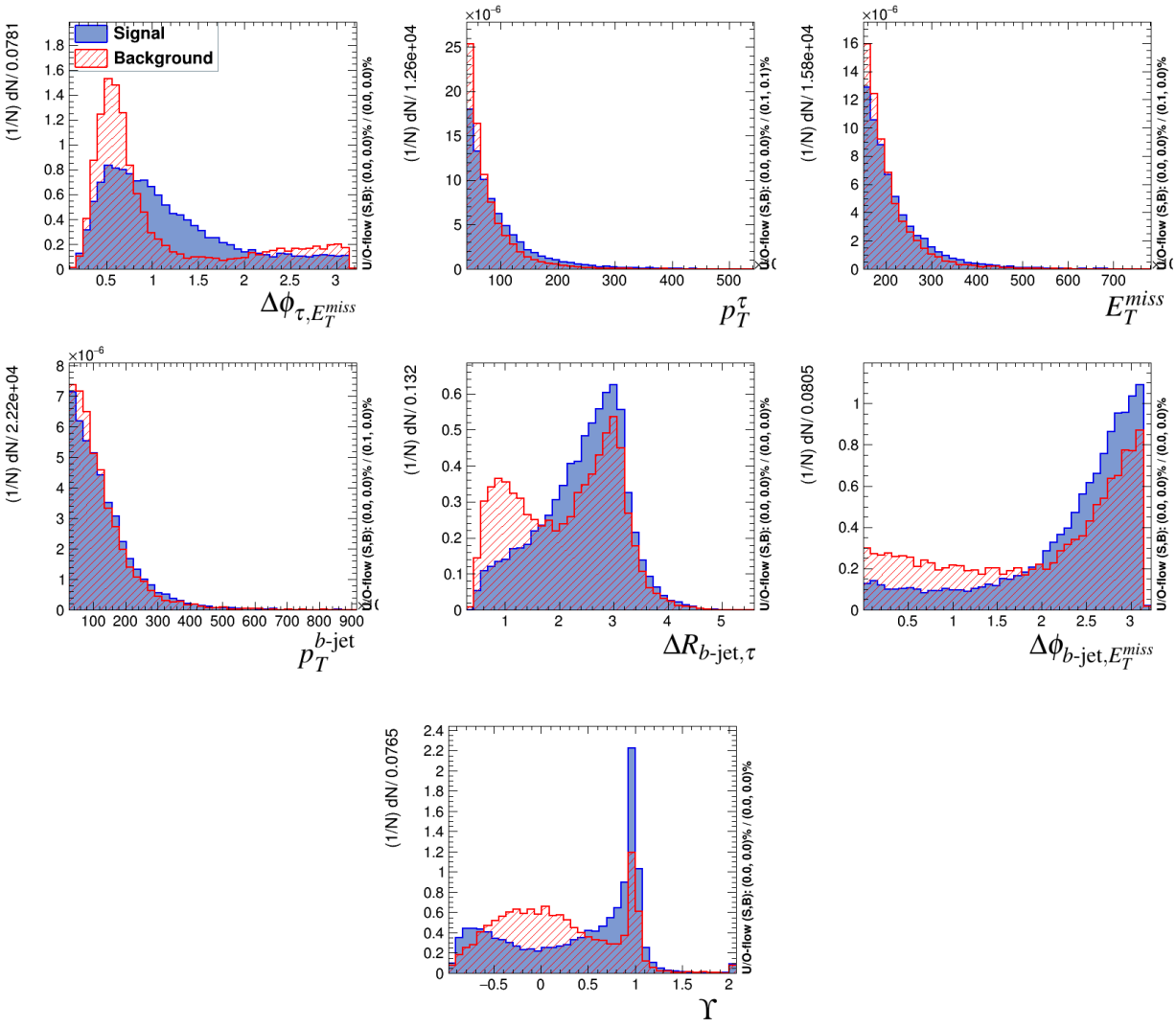



Figure C.2: Distributions of input variables for BDT-training within 160-180 GeV mass bin. Signal is presented as a blue, solid histogram; top background (as estimated from MC) is presented as a red, dashed histogram. Histograms are arbitrarily normalised.

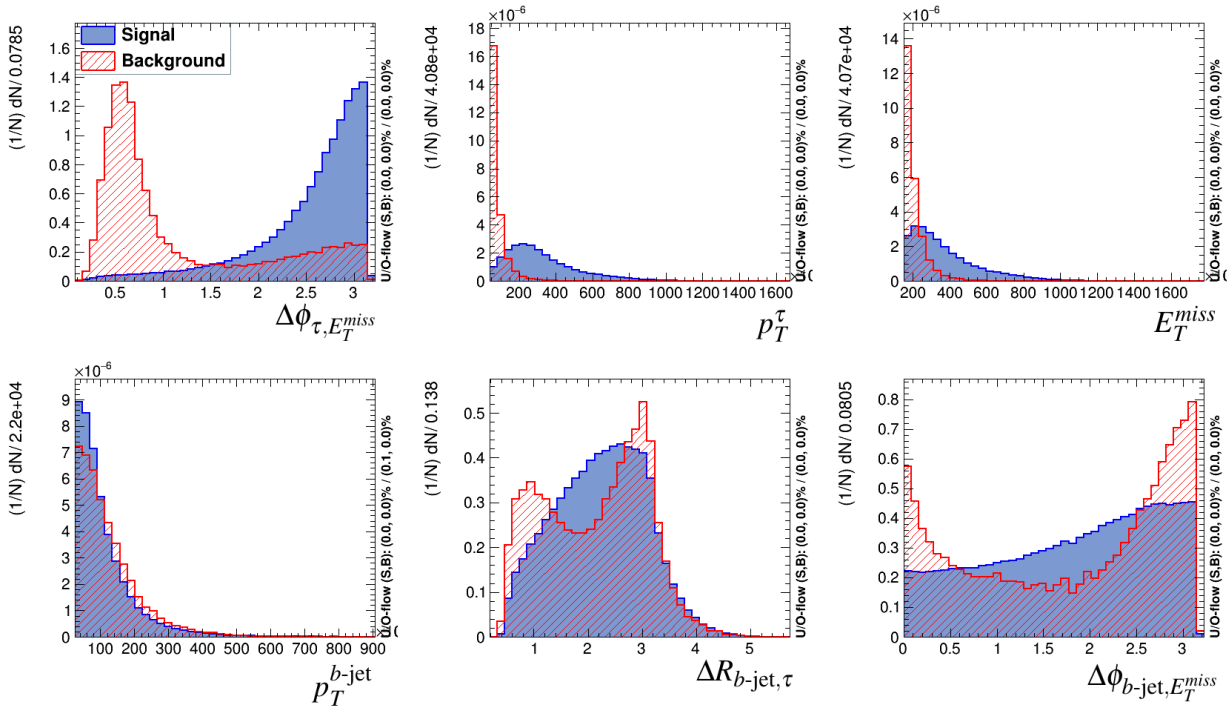



Figure C.3: Distributions of input variables for BDT-training within 500-2000 GeV mass bin. Signal is presented as a blue, solid histogram; top background (as estimated from MC) is presented as a red, dashed histogram. Histograms are arbitrarily normalised.

# Bibliography


[1] K.A. Olive et al. (Particle Data Group) *Chin. Phys. C* **38** (2014) 090001.

[2] **ATLAS** Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[3] **CMS** Collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[4] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.

[5] **ATLAS** Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1 [arXiv:1207.7214 [hep-ex]].

[6] **CMS** Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30 [arXiv:1207.7235 [hep-ex]].

[7] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. II. The Higgs bosons in the minimal supersymmetric model*, *Phys. Rept.* **459** (2008) 1 [arXiv:hep-ph/0503173 [hep-ph]].

[8] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher and J. P. Silva, *Theory and phenomenology of two-Higgs-doublet models*, *Phys. Rept.* **516** (2012) 1 [arXiv:1106.0034 [hep-ph]].

[9] G. Senjanovic and R. N. Mohapatra, *Exact Left-Right Symmetry and Spontaneous Violation of Parity*, *Phys. Rev.* **D12** (1975) 1502.

[10] J. F. Gunion, R. Vega and J. Wudka, *Higgs triplets in the standard model*, *Phys. Rev.* **D42** (1990) 1673.

[11] K. Agashe, N. G. Deshpande and G. H. Wu, *Charged Higgs decays in models with singlet neutrino in large extra dimensions*, *Phys. Lett.* **B489** (2000) 367 [arXiv:hep-ph/0006122 [hep-ph]].

[12] **ATLAS** Collaboration, *Search for charged Higgs bosons decaying via $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ in the $\tau$+jets and $\tau$+lepton final states with 36 $fb^{-1}$ of pp collision data recorded at $\sqrt{s}$=13 TeV with the ATLAS experiment*, submitted to JHEP (2018) [arXiv:1807.07915 [hep-ex]].

[13] D. J. Griffiths, *Introduction to elementary particles; 2nd rev. version*. Wiley, New York, NY, 2008.

[14] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[15] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[16] T. Morii, L. C. S and S. N. Mukherjee, *The physics of the Standard Model and beyond*. World Scientific, Singapore, 2004.

[17] **ATLAS, CMS** Collaboration, *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015) 191803 [arXiv:1503.07589 [hep-ex]].

[18] **ATLAS** Collaboration, *Study of the spin and parity of the Higgs boson in diboson decays with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 476 [arXiv:1506.05669 [hep-ex]].

[19] **CMS** Collaboration, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*, Eur. Phys. J. **C75** (2015) 212 [arXiv:1412.8662 [hep-ex]].

[20] **ATLAS** Collaboration, *Measurements of the Higgs boson production and decay rates and coupling strengths using pp collision data at $\sqrt{s} = 7$ and 8 TeV in the ATLAS experiment*, Eur. Phys. J. **C76** (2016) 6 [arXiv:1507.04548 [hep-ex]].

[21] J. H. Christenson, J. W. Cronin, V. L. Fitch and R. Turlay, *Evidence for the $2\pi$ Decay of the $K_2^0$ Meson*, Phys. Rev. Lett. **13** (1964) 138.

[22] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, Helvetica Physica Acta **6** (1933) 110.

[23] D. M. Wittman, J. A. Tyson, D. Kirkman, I. Dell'Antonio and G. Bernstein, *Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales*, *Nature* **405** (2000) 143 [arXiv:astro-ph/0003014 [astro-ph]].

[24] S. Perlmutter et al., *Measurements of Omega and Lambda from 42 high redshift supernovae*, *Astrophys. J.* **517** (1999) 565 [arXiv:astro-ph/9812133 [astro-ph]].

[25] A. G. Riess et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron. J.* **116** (1998) 1009 [arXiv:astro-ph/9805201 [astro-ph]].

[26] R. N. Mohapatra and G. Senjanovic, *Neutrino Masses and Mixings in Gauge Models with Spontaneous Parity Violation*, *Phys. Rev.* **D23** (1981) 165.

[27] M. Fukugita and T. Yanagida, *Baryogenesis Without Grand Unification*, *Phys. Lett.* **B174** (1986) 45.

[28] P. Fayet, *Supersymmetry and Weak, Electromagnetic and Strong Interactions*, *Phys. Lett.* **64B** (1976) 159.

[29] M. Carena, S. Heinemeyer, O. Stl, 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** (2013) 2552 [arXiv:1302.7033 [hep-ph]].

[30] R. Harlander, M. Kramer and M. Schumacher, *Bottom-quark associated Higgs-boson production: reconciling the four- and five-flavour scheme approach*, arXiv:[1112.3478] [hep-ph].

[31] M. Flechl, R. Klees, M. Kramer, M. Spira and M. Ubiali, *Improved cross-section predictions for heavy charged Higgs boson production at the LHC*, *Phys. Rev.* **D91** (2015) 075015 [arXiv:1409.5615 [hep-ph]].

[32] M. Flechl, *Charged Higgs boson production around the top quark mass threshold*, *PoS* **CHARGED2016** (2016) 013.

[33] M. Pitt et al., *Search for charged Higgs bosons in the $\tau$+jets and $\tau$+lepton final states with 36.1 $fb^{-1}$ of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, *ATL-COM-PHYS-2016-1452* (2018) .

[34] **ATLAS** Collaboration, *Search for neutral MSSM Higgs bosons decaying to $\tau^+\tau^-$ pairs in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *Phys. Lett.* **B705** (2011) 174 [arXiv:1107.5003 [hep-ex]].

[35] **ATLAS** Collaboration, *Search for the Standard Model Higgs boson in the H to $\tau^+\tau^-$ decay mode in $\sqrt{s} = 7$ TeV pp collisions with ATLAS*, *JHEP* **09** (2012) 070 [arXiv:1206.5971 [hep-ex]].

[36] **ATLAS** Collaboration, *Search for the neutral Higgs bosons of the Minimal Supersymmetric Standard Model in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **02** (2013) 095 [arXiv:1211.6956 [hep-ex]].

[37] **ATLAS** Collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, *JHEP* **04** (2015) 117 [arXiv:1501.04943 [hep-ex]].

[38] **ATLAS** Collaboration, *Search for charged Higgs bosons decaying via $H^+ \to \tau\nu$ in top quark pair events using pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *JHEP* **06** (2012) 039 [arXiv:1204.2760 [hep-ex]].

[39] **ATLAS** Collaboration, *Search for charged Higgs bosons decaying via $H^\pm \to \tau^\pm\nu$ in fully hadronic final states using pp collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **03** (2015) 088 [arXiv:1412.6663 [hep-ex]].

[40] **CMS** Collaboration, *Search for neutral Higgs bosons decaying to tau pairs in pp collisions at $\sqrt{s} = 7$ TeV*, *Phys. Lett.* **B713** (2012) 68 [arXiv:1202.4083 [hep-ex]].

[41] **CMS** Collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of $\tau$ leptons*, *JHEP* **05** (2014) 104 [arXiv:1401.5041 [hep-ex]].

[42] **CMS** Collaboration, *Search for charged Higgs bosons with the H+ to tau nu decay channel in the fully hadronic final state at $\sqrt{s} = 8$ TeV*, *CMS-PAS-HIG-14-020* (2014) .

[43] **ATLAS** Collaboration, *Search for charged Higgs bosons produced in association with a top quark and decaying via $H^\pm \to \tau\nu$ using pp collision data recorded at $\sqrt{s} = 13$ TeV by the ATLAS detector*, *Phys. Lett.* **B759** (2016) 555 [arXiv:1603.09203 [hep-ex]].

[44] O. Deschamps, S. Descotes-Genon, S. Monteil, V. Niess, S. T'Jampens and V. Tisserand, *The Two Higgs Doublet of Type II facing flavour physics data*, *Phys. Rev.* **D82** (2010) 073012 [arXiv:0907.5135 [hep-ph]].

[45] M. L. Perl, *The Discovery of the tau lepton*, in *The Rise of the standard model: Particle physics in the 1960s and 1970s. Proceedings, Conference, Stanford, USA, June 24-27, 1992*, pp. 79–100, 1992, http://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=SLAC-PUB-5937.

[46] **ATLAS** Collaboration, *Measurement of $\tau$ polarisation in $Z/\gamma^* \to \tau\tau$ decays in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Eur. Phys. J.* **C78** (2018) 163 [arXiv:1709.03490 [hep-ex]].

[47] S. Bedikian, *A Search for the Charged Higgs: Using Tau Polarimetry with Proton-Proton Collisions at the ATLAS Detector*, Ph.D. thesis, Yale U., 2013-03-14.

[48] **LHCb** Collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.

[49] **ALICE** Collaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002.

[50] E. Mobs, *The CERN accelerator complex. Complexe des accelerateurs du CERN*, *OPEN-PHO-ACCEL-2016-009* (2016) .

[51] **ATLAS** Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, *CERN-LHCC-2010-013* (2010) .

[52] **ATLAS** Collaboration, *Performance of the ATLAS Trigger System in 2015*, *Eur. Phys. J.* **C77** (2017) 317 [arXiv:1611.09661 [hep-ex]].

[53] P. Grafström and W. Kozanecki, *Luminosity determination at proton colliders*, *Prog. Part. Nucl. Phys.* **81** (2015) 97.

[54] "https://twiki.cern.ch/twiki/bin/view/atlaspublic/luminositypublicresults."

[55] **ATLAS** Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, *Eur. Phys. J.* **C76** (2016) 653 [arXiv:1608.03953 [hep-ex]].

[56] **ATLAS** Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$TeV with the ATLAS Detector at the LHC*, *Phys. Rev. Lett.* **117** (2016) 182002 [arXiv:1606.02625 [hep-ex]].

[57] J. M. Campbell, J. W. Huston and W. J. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, *Rept. Prog. Phys.* **70** (2007) 89 [arXiv:hep-ph/0611148 [hep-ph]].

[58] C. Leonidopoulos, *Introduction to Collider Physics*, *Scottish Universities Physics Alliance Lectures, Edinburgh* (2014) .

[59] S. Höche, *Introduction to parton-shower event generators*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, pp. 235–295, 2015, arXiv:1411.4085 [hep-ph], DOI.

[60] T. Sjöstrand, *Monte Carlo event generation for LHC*, in *1991 CERN School of Computing Ystad, Sweden, August 23-September 2, 1991*, pp. 0227–249, 1991.

[61] A. Buckley et al., *General-purpose event generators for LHC physics*, *Phys. Rept.* **504** (2011) 145 [arXiv:1101.2599 [hep-ph]].

[62] **GEANT4** Collaboration, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250.

[63] **ATLAS** Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, *ATL-PHYS-PUB-2010-013* (2010) .

[64] J. Catmore, *Lectures at Joint Oslo/Bergen/NBI ATLAS Software Tutorial, https://indico.cern.ch/event/472469/* .

[65] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse and A. Salzburger, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, Mar, 2007.

[66] **ATLAS** Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment*, *ATLAS-CONF-2012-042* (2012) .

[67] **ATLAS** Collaboration, *Early Inner Detector Tracking Performance in the 2015 data at* $\sqrt{s}$ *= 13 TeV*, *ATL-PHYS-PUB-2015-051* (2015) .

[68] **ATLAS** Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at* $\sqrt{s} = 13$ *TeV* , *ATL-PHYS-PUB-2015-026* (2015) .

[69] **ATLAS** Collaboration, *The Optimization of ATLAS Track Reconstruction in Dense Environments*, *ATL-PHYS-PUB-2015-006* (2015) .

[70] **ATLAS** Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J.* **C77** (2017) 490 [arXiv:1603.02934 [hep-ex]].

[71] M. Cacciari, G. P. Salam and G. Soyez, *The Anti-k(t) jet clustering algorithm*, *JHEP* **04** (2008) 063 [arXiv:0802.1189 [hep-ph]].

[72] **ATLAS** Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J.* **C76** (2016) 581 [arXiv:1510.03823 [hep-ex]].

[73] **ATLAS** Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev.* **D96** (2017) 072002 [arXiv:1703.09665 [hep-ex]].

[74] **ATLAS** Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, *ATL-PHYS-PUB-2015-015* (2015) .

[75] **ATLAS** Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, *JINST* **11** (2016) P04008 [arXiv:1512.01094 [hep-ex]].

[76] **ATLAS** Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, *ATL-PHYS-PUB-2016-012* (2016) .

[77] **D0** Collaboration, *Observation of Single Top Quark Production*, *Phys. Rev. Lett.* **103** (2009) 092001 [arXiv:0903.0850 [hep-ex]].

[78] **ATLAS** Collaboration, G. Aad et al., *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV*, *Eur. Phys. J.* **C75** (2015) 303 [arXiv:1412.7086 [hep-ex]].

[79] **ATLAS** Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, *ATL-PHYS-PUB-2015-045* (2015) .

[80] **ATLAS** Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, *ATLAS-CONF-2017-029* (2017) .

[81] **ATLAS** Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, *ATLAS-CONF-2014-032* (2014) .

[82] Di Ciaccio et al., *Support Note for Electron ID: efficiency measurements*, *ATL-COM-PHYS-2017-1352* (2017) .

[83] **ATLAS** Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV*, *Eur. Phys. J.* **C76** (2016) 292 [arXiv:1603.05598 [hep-ex]].

[84] **ATLAS** Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV*, arXiv:[1802.08168] [hep-ex].

[85] **ATLAS** Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, *Eur. Phys. J.* **C76** (2016) 666 [arXiv:1606.01813 [hep-ex]].

[86] **ATLAS** Collaboration, *Search for charged Higgs bosons produced in association with a top quark and decaying via $H^{\pm} \to \tau\nu$ using pp collision data recorded at $\sqrt{s} = 13$ TeV by the ATLAS detector*, *Phys. Lett.* **B759** (2016) 555 [arXiv:1603.09203 [hep-ex]].

[87] J. Griffiths et al., *Search for charged Higgs bosons in the $\tau$+jets final state using 14.7 $fb^{-1}$ of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, *ATL-COM-PHYS-2016-501* (2016) .

[88] **ATLAS** Collaboration, *Search for charged Higgs bosons in the $\tau$+jets final state using 14.7 $fb^{-1}$ of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, *ATLAS-CONF-2016-088* (2016) .

[89] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *MadGraph 5 : Going Beyond*, *JHEP* **06** (2011) 128 [arXiv:1106.0522 [hep-ph]].

[90] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer 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 [arXiv:1405.0301 [hep-ph]].

[91] C. Degrande, M. Ubiali, M. Wiesemann and M. Zaro, *Heavy charged Higgs boson production at the LHC*, *JHEP* **10** (2015) 145 [arXiv:1507.02549 [hep-ph]].

[92] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys.* **B867** (2013) 244 [arXiv:1207.1303 [hep-ph]].

[93] **ATLAS** Collaboration, *ATLAS Run 1 Pythia8 tunes*, *ATL-PHYS-PUB-2014-021* (2014) .

[94] T. Sjöstrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [arXiv:0710.3820 [hep-ph]].

[95] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [arXiv:hep-ph/0409146 [hep-ph]].

[96] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [arXiv:0709.2092 [hep-ph]].

[97] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [arXiv:1002.2581 [hep-ph]].

[98] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin et al., *New parton distributions for collider physics*, *Phys. Rev.* **D82** (2010) 074024 [arXiv:1007.2241 [hep-ph]].

[99] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev.* **D93** (2016) 033006 [arXiv:1506.07443 [hep-ph]].

[100] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, *JHEP* **03** (2013) 015 [arXiv:1212.3460 [hep-ph]].

[101] T. Sjöstrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026 [arXiv:hep-ph/0603175 [hep-ph]].

[102] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky and W. K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, *JHEP* **07** (2002) 012 [arXiv:hep-ph/0201195 [hep-ph]].

[103] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, *Phys. Rev.* **D82** (2010) 074018 [arXiv:1005.3457 [hep-ph]].

[104] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930 [arXiv:1112.5675 [hep-ph]].

[105] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, *Comput. Phys. Commun.* **191** (2015) 74 [arXiv:1406.4403 [hep-ph]].

[106] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer and M. Wiedermann, *HATHOR: HAdronic Top and Heavy quarks cross section calculatoR*, *Comput. Phys. Commun.* **182** (2011) 1034 [arXiv:1007.1327 [hep-ph]].

[107] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, *Phys. Rev.* **D82** (2010) 054018 [arXiv:1005.4451 [hep-ph]].

[108] T. Gleisberg, S. Hoeche, F. Krauss, M. Schonherr, S. Schumann, F. Siegert et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [arXiv:0811.4622 [hep-ph]].

[109] **NNPDF** Collaboration, R. D. Ball et al., *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [arXiv:1410.8849 [hep-ph]].

[110] 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 [arXiv:hep-ph/0312266 [hep-ph]].

[111] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through O(alpha(s)**2)*, *Phys. Rev.* **D74** (2006) 114017 [arXiv:hep-ph/0609070 [hep-ph]].

[112] 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 [arXiv:1011.3540 [hep-ph]].

[113] **ATLAS** Collaboration, *Measurement of the $Z/\gamma^*$ boson transverse momentum distribution in $pp$ collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, *JHEP* **09** (2014) 145 [arXiv:1406.3660 [hep-ex]].

[114] N. Davidson, T. Przedzinski and Z. Was, *PHOTOS interface in C++: Technical and Physics Documentation*, *Comput. Phys. Commun.* **199** (2016) 86 [arXiv:1011.0937 [hep-ph]].

[115] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152.

[116] **ATLAS** Collaboration, *Summary of ATLAS Pythia 8 tunes*, *ATL-PHYS-PUB-2012-003* (2012) .

[117] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J.* **C63** (2009) 189 [arXiv:0901.0002 [hep-ph]].

[118] L. Devroye, ed., *Non-Uniform Random Variate Generation*. Springer-Verlag New York, 1986.

[119] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J.* **C75** (2015) 132 [arXiv:1412.7420 [hep-ph]].

[120] P. C. Bhat, *Advanced analysis methods in high-energy physics*, *AIP Conf. Proc.* **583** (2001) 22 [arXiv:hep-ex/0106099 [hep-ex]].

[121] **D0** Collaboration, *Search for single top quarks in the tau+jets channel using 4.8 $fb^{-1}$ of p p-bar collision data*, *Phys. Lett.* **B690** (2010) 5 [arXiv:0912.1066 [hep-ex]].

[122] **ATLAS** Collaboration, *Measurement of the top quark pair cross section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV using final states with an electron or a muon and a hadronically decaying $\tau$ lepton*, *Phys. Lett.* **B717** (2012) 89 [arXiv:1205.2067 [hep-ex]].

[123] **ATLAS** Collaboration, *Evidence for the associated production of a W boson and a top quark in ATLAS at $\sqrt{s} = 7$ TeV*, *Phys. Lett.* **B716** (2012) 142 [arXiv:1205.5764 [hep-ex]].

[124] O. Behnke, K. Krninger, T. Schrner-Sadenius and G. Schott, eds., *Data analysis in high energy physics*. Wiley-VCH, Weinheim, Germany, 2013.

[125] P. Liang and S. Ermo, *Machine learning*, *Lectures, Stanford University, California* (2017) .

[126] T. Hastie, R. Tibshirani and J. Friedman, *The Elements of Statistical Learning*, Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.

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

[128] T. Keck, *Fastbdt: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification*, *CoRR* **abs/1609.06119** (2016) [arXiv:1609.06119].

[129] J. H. Friedman, *Stochastic gradient boosting*, *Computational Statistics and Data Analysis* **38** (2002) 367 .

[130] G. Prashant, "https://towardsdatascience.com/cross-validation-in-machine-learning-72924a69872f."

[131] T. Keck, *Fastbdt: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification*, *CoRR* (2016) [arXiv:1609.06119].

[132] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, [physics/0703039].

[133] D. P. Roy, *The hadronic tau decay signature of a heavy charged Higgs boson at LHC*, *Physics Letters B* **459** (1999) 607 [hep-ph/9905542].

[134] N. Smirnov, *Table for estimating the goodness of fit of empirical distributions*, *Annals of Mathematical Statistics **19**: 279-281* (1948) .

[135] **ATLAS** Collaboration, *Measurement of the $W \to \tau\nu_\tau$ cross section in $pp$ collisions at $\sqrt{s}$ =7 TeV with the ATLAS experiment*, *Phys. Lett.* **B706** (2012) 276 [arXiv:1108.4101 [hep-ex]].

[136] Bosman, M et al., *Search for charged Higgs bosons in the $\tau$+jets final state with pp collision data recorded at $\sqrt{s} = 8$ TeV with the ATLAS experiment* , *ATL-COM-PHYS-2014-211* (2014) .

[137] A. Kopp, *Search for charged Higgs bosons decaying via $H^\pm \to \tau^\pm\nu$ in fully hadronic final states with the ATLAS detector at the LHC*, Ph.D. thesis, Freiburg U., 2015-09-11.

[138] C. Isaksson, M. Flechl, N. Mser and M. Schmitz, *Embedding Technique for the ttbar Background Estimation in Charged Higgs Boson Searches*, Oct, 2009.

[139] S. Jadach, Z. Was, R. Decker and J. H. Kuhn, *The tau decay library TAUOLA: Version 2.4*, *Comput. Phys. Commun.* **76** (1993) 361.

[140] F. Sforza, "https://twiki.cern.ch/twiki/bin/view/atlasprotected/mcpanalysisguidelinesmc15."

[141] **ATLAS** Collaboration, *Study of the material of the ATLAS inner detector for Run 2 of the LHC*, *JINST* **12** (2017) P12009 [arXiv:1707.02826 [hep-ex]].

[142] **ATLAS** Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at* $\sqrt{s} = 13$ TeV, *ATLAS-CONF-2017-029* (2017) .

[143] **ATLAS** Collaboration, *Modelling* $Z \to \tau\tau$ *processes in ATLAS with* $\tau$*-embedded* $Z \to \mu\mu$ *data*, *JINST* **10** (2015) P09018 [arXiv:1506.05623 [hep-ex]].

[144] 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 [physics.data-an]].

[145] A. L. Read, *Presentation of search results: The CL(s) technique*, *J. Phys.* **G28** (2002) 2693.

[146] A. Djouadi, L. Maiani, G. Moreau, A. Polosa, J. Quevillon and V. Riquer, *The post-Higgs MSSM scenario: Habemus MSSM?*, *Eur. Phys. J.* **C73** (2013) 2650 [arXiv:1307.5205 [hep-ph]].

[147] A. Djouadi, L. Maiani, A. Polosa, J. Quevillon and V. Riquer, *Fully covering the MSSM Higgs sector at the LHC*, *JHEP* **06** (2015) 168 [arXiv:1502.05653 [hep-ph]].

[148] J. R. Andersen et al., *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, arXiv:[1307.1347] [hep-ph].

[149] S. Dittmaier, M. Kramer, M. Spira and M. Walser, *Charged-Higgs-boson production at the LHC: NLO supersymmetric QCD corrections*, *Phys. Rev.* **D83** (2011) 055005 [arXiv:0906.2648 [hep-ph]].

[150] A. Djouadi, J. Kalinowski and M. Spira, *HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension*, *Comput. Phys. Commun.* **108** (1998) 56 [arXiv:hep-ph/9704448 [hep-ph]].

[151] "http://sherpa.hepforge.org/dokuwiki/lib/exe/detail.php?media=sketch.gif.."

## Podziękowania/Acknowledgements

Throughout the last years people closest to me have supported me in many different ways. Most of them are not mentioned here not because I forgot about my gratitude for them but rather because I want to keep this to the more directly work-related people.

Dziękuję prof. Tadeuszowi Lesiakowi za wyciągniętą do mnie pomocną dłoń kiedy pojawiłam się na studiach doktoranckich w Instytucie Fizyki Jądrowej w Krakowie.

Za otwarte serce, czas i słowa wsparcia podarowane mi wtedy przez Anię Kaczmarską.

Dziękuję Ani i Pani Barbarze Wosiek, jak również Panu Lesiakowi za możliwość pracy poza Instytutem.

Many thanks to Anna Kopp, Dirk Sammel, Duc Bao Ta, Elias Coniavitis, Justin Griffiths, Heather Russell, Mikael Martensson and Michael Pitt. Without them the results presented here would not have been possible. Dziękuję też Pawłowi Bruckmanowi za pokazanie mi jak używa się maszynki do robienia limitów.

Dziękuję Paniom: Marii i Beacie z Sekretariatu Oddziału Fizyki i Astrofizyki Cząstek; Paniom: Iwonie, Renacie i Jowicie z Międzynarodowego Studium Doktoranckiego oraz Pani Ewie Bartel z Działu Planowania Ekonomicznego za niezawodną pomoc.

Last but not least, Dziękuję Andrzejowi za jego niesłabnące wsparcie, cierpliwość, ogromną znajomość fizyki cząstek i niezastąpiony ciepły uśmiech bliskości.