# Search for a light charged Higgs boson in $t \to H^{\pm} b$ decays, with $H^{\pm} \to c s$, in $pp$ collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

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

vorgelegt von
Christian Nass
aus
Koblenz

Bonn, 2025

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

| | |
|---|---|
| Gutachter/Betreuer: | Prof. Dr. Jochen Dingfelder |
| Gutachter: | Prof. Dr. Florian Bernlochner |
| Tag der Promotion: | 12.12.2025 |
| Erscheinungsjahr: | 2026 |

# Contents

CHAPTER 1

# Introduction

The goal of physicists is to grasp the Universe and express their knowledge in mathematical models. Particle physics does this at the most fundamental level by studying the fundamental particles of the Universe and their interactions. Today, our best understanding is encoded in the Standard Model (SM) of particle physics. The SM was developed in the 1970s and ever since has been very successful in describing experimental results, though not all observations. Among its shortcomings are the inability to provide a dark-matter candidate and to describe the baryon asymmetry in the Universe. As a consequence, there has to be new physics beyond the Standard Model (BSM).

New physics can be searched for either indirectly by performing precision measurements of observables that can be precisely predicted by the SM or by direct searches for new particles predicted by BSM theories. In this thesis a direct search for a new electrically charged scalar particle is performed, a charged Higgs boson $H^{\pm}$. The SM contains only a single scalar particle, the SM Higgs boson $H$. The Higgs boson is uncharged and was the last missing piece of the SM which was discovered in 2012 by the ATLAS and CMS collaborations [1, 2] at CERN. The Higgs boson was predicted by the Higgs mechanism, developed by Brout, Englert and Higgs [3, 4]. The Higgs mechanism in the SM is the simplest realisation of a scalar sector which respects the electroweak symmetry and can explain the existence of the massive electroweak gauge bosons $W^{\pm}$ and $Z$. However, a more complex scalar sector, leading to more than one Higgs boson, would be able to explain these observations, too, and in addition could provide solutions to shortcomings of the SM.

For instance, an extended scalar sector can add new charge conjugation parity (CP) symmetry violation to the SM [5–8]. The CP violation present in the SM is not sufficient to explain the astronomically observed baryon asymmetry of the Universe [9, 10]. New symmetries can also promote particles from an extended scalar sector to be dark-matter candidates [11, 12]. An extended scalar sector can also yield an answer to why CP-symmetry is preserved in the strong interaction (strong CP problem) [13, 14]. In any case, many sophisticated BSM models require an extended scalar sector. For example, supersymmetric models (SUSY) [15, 16] and models with a Type-II seesaw mechanism [17–21]. A charged Higgs boson shows up in all models that add at least one additional doublet of complex scalar fields to the Higgs sector.

When performing a search for a new particle, the production and decay topology has to be specified in order to develop a dedicated analysis strategy. This thesis focuses on the search for a charged Higgs boson with a mass smaller than the mass of the top quark ($m_{H^{\pm}} < m_t$) decaying into a charm- and a

strange-quark $cs$[1]. In this mass regime charged Higgs bosons are expected to be mainly produced in the decay of top quarks, $t \to H^{\pm}b$, and many phenomenology studies advocate searching for a charged Higgs boson in the $cs$ decay channel [11, 22].

Data from the Large Hadron Collider (LHC) at CERN is analysed in this search. The dataset corresponds to 140 fb$^{-1}$ of $pp$ collisions collected at a centre-of-mass energy of $\sqrt{s} = 13$ TeV[2]. In $pp$ collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV top quarks are predominantly produced in pairs. Therefore, this thesis focuses on the $t\bar{t} \to WbH^{\pm}b$ process, with $W \to \ell\nu_{\ell}$ ($\ell = e, \mu$) and $H^{\pm} \to cs$. Only events with a lepton – $e$ or $\mu$, including those from leptonically decaying $\tau$-leptons – in the final state are considered to simplify the event selection and reconstruction. Figure 1.1 shows the dominant Feynman diagram of the signal process.

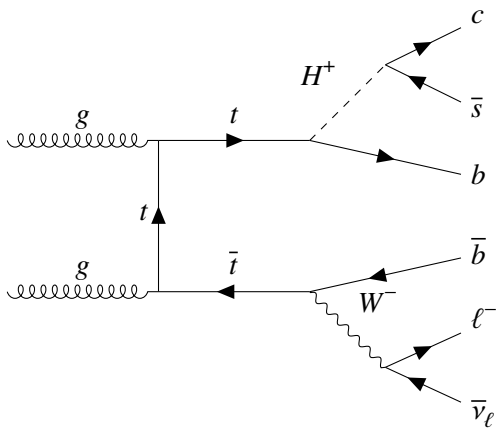



Figure 1.1: Depiction of the signal process via the dominant Feynman diagram.

A potential signal in the data is identified by first reconstructing the $t\bar{t}$ event topology and then training a Boosted Decision Tree (BDT) to separate the signal from the dominant SM $t\bar{t}$ background. The BDT exploits differences in kinematic properties of the $t\bar{t}$ system and the flavour composition of the events. A dedicated flavour-tagging scheme is adopted which facilitates simultaneous tagging of $b$- and $c$-quark-initiated jets.

The thesis is structured as follows: Chapter 2 describes the theoretical framework foundations for the analysis presented in this thesis, i.e. the SM and popular BSM models predicting a charged Higgs boson. Chapter 3 introduces the experimental setup and techniques used to reconstruct particles in ATLAS. Chapters 4, 5 and 6 present the methods used to extract a potential signal from data. The results of the search for $H^{\pm} \to cs$ decays are presented in Chapter 7, with concluding remarks in Chapter 8.

[1] Unless explicitly stated otherwise, charge conjugation is implied in this thesis; the notation $cs$ is used in place of $c\bar{s}/\bar{c}s$.

[2] Throughout the thesis, energies, momenta and masses are expressed in natural units ($\hbar = c = \epsilon_0 = 1$).

CHAPTER 2

# The Standard Model and theories beyond

This chapter gives an introduction to the foundations of particle physics. The current knowledge of the elementary particles and their interactions is gathered in the Standard Model of particle physics. Section 2.1 introduces the basic theoretical concept of the SM and focuses on its experimental implications to understand the physics processes occurring in particle collisions. This section is based on References [23–26], which also serve as references for further reading. Despite its success in describing experimental observations, certain observations in the Universe are incompatible with the SM predictions. Theories beyond the SM that predict the existence of a charged Higgs boson are described in Section 2.2. Lastly, Section 2.3 summarises the results of preceding searches for a charged Higgs boson.

## 2.1 The Standard Model of particle physics

The SM is a theoretical model describing the known fundamental particles and three of the four known fundamental interactions: strong, weak and electromagnetic. The gravitational force is not part of the SM, but is known to be negligible at microscopic scales. The formulation of the SM was finalised in the 1970s. Since then, it has been very successful in describing the experimental data with high precision, even predicting new particles, like the Higgs boson.

The main underlying theory of the SM is relativistic quantum field theory (QFT), in which a Lagrangian defines the dynamics of the system. QFT combines the theories of quantum mechanics and special relativity, which enables the description of physics processes at subatomic length scales and energies much larger than the energy of the system at rest. In QFT the fundamental objects are quantum fields that pervade space and time. Particles are excitations of their underlying quantum field. The particles of the SM are introduced in Section 2.1.1.

The SM is constructed by postulating symmetries of the Universe, which, according to Noether's theorem, lead to conservation laws, and then working out the most general renormalisable Lagrangian that adheres to these symmetries. The postulated symmetries are, on the one hand, the Poincaré (space-time) symmetry, which ensures that the physics does not depend on the reference frame of an observer and implies for instance energy and momentum conservation. On the other hand, the local

$$\mathrm{SU}(3) \times \mathrm{SU}(2)_L \times \mathrm{U}(1)$$

gauge symmetry is postulated, which is the essential part of the SM. SU(3) is the symmetry of the strong

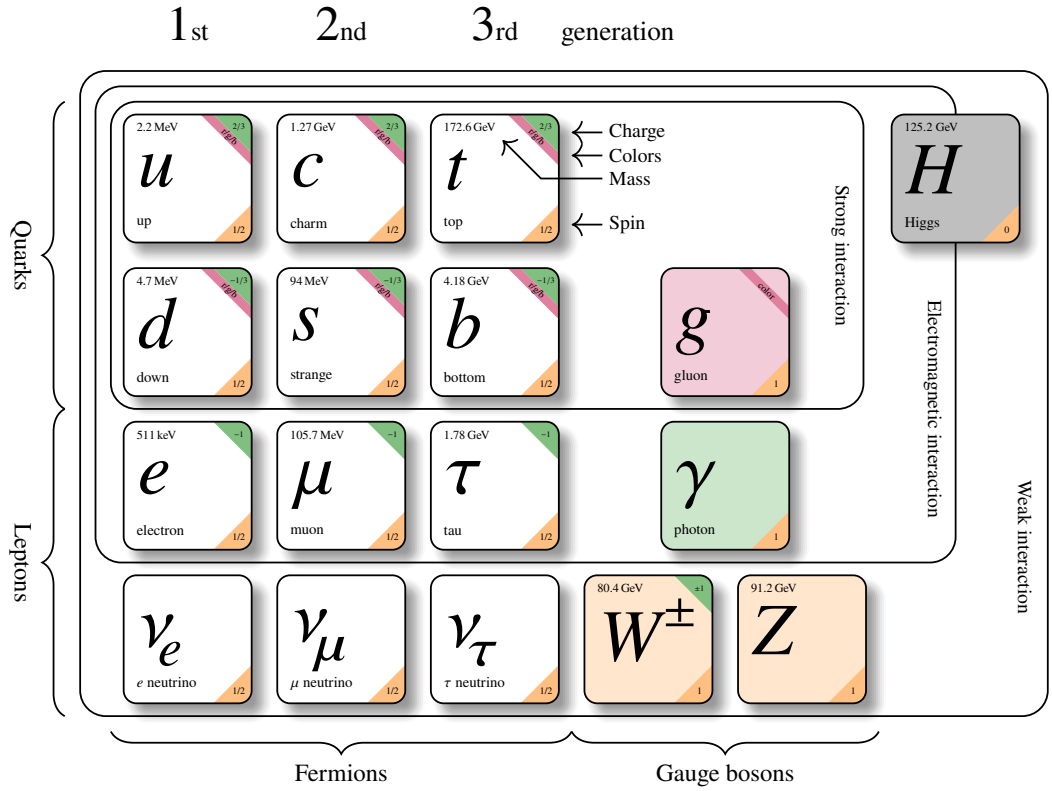



Figure 2.1: Fundamental particles and interactions of the SM. The diagram is adapted from Reference [27]. The particle masses are taken from Reference [28].

interaction, and $SU(2)_L \times U(1)$ is the symmetry of the unified weak and electromagnetic interaction. The local gauge invariance requirement of the Lagrangian with respect to these symmetries defines the exact nature of the interactions. The interactions of the particles and the Higgs mechanism, which breaks the electroweak symmetry and explains how elementary particles acquire mass, are described in Section 2.1.2.

### 2.1.1 Particles of the SM

The particles of the SM are shown in Figure 2.1. Also indicated are essential properties of the particles: electric charge, colour charge, mass and spin, as well as the interactions they participate in. Based on these properties, the particles are grouped into the following categories:

**Gauge bosons** are bosons with spin 1. Bosons are integer spin particles that follow the Bose-Einstein statistics, from which they take their name. Gauge bosons mediate force between particles. Gluons $g$ are the mediators of the strong interaction. Gluons are massless and carry colour and anticolour. The photon $\gamma$ is the mediator of the electromagnetic interaction. The photon is massless and carries no electric charge itself. The $W^{\pm}$ and $Z$ bosons are massive and are the mediators of the charged- and neutral-current weak interactions, respectively. The $W^+$ and $W^-$ carry electric charge and are each other's antiparticle. Gluons, photons and the $Z$ boson are their

own antiparticles.

**Fermions** are particles with half-integer spin. Fermions obey the Pauli exclusion principle and follow the Fermi–Dirac statistics, from which they take their name. Fermions are the building blocks of the Universe (matter particles). The fermions in the SM all have spin 1/2. The SM contains three generations of fermions that share the same properties apart from their mass, which increases from one generation to the next. Ordinary matter primarily consists of particles from the first generation. For every fermion there exists an antiparticle with identical properties except for the additive quantum numbers, which are of opposite sign[1].

**Quarks** $q$ are fermions that carry a colour charge. Quarks participate in all three interactions. Quarks are divided into up-type quarks with electric charge 2/3 (up $u$, charm $c$, top $t$) and down-type quarks with electric charge -1/3 (down $d$, strange $s$, bottom $b$).

**Leptons** are fermions that don't carry a colour charge. Hence, they don't participate in the strong interaction. Charged leptons $\ell$ (electron $e$, muon $\mu$, tau $\tau$) carry an electric charge of -1 and therewith participate in the weak and electromagnetic interactions. Neutrinos $\nu$ (electron neutrino $\nu_e$, muon neutrino $\nu_\mu$, tau neutrino $\nu_\tau$) carry no electric charge and therewith participate only in the weak interaction. In the SM, neutrinos are massless[2].

The **Higgs boson** arises as part of the Higgs mechanism, described in more detail in Section 2.1.2. It is the only scalar particle of the SM (spin 0). The Higgs boson couples to every massive particle. The coupling strength is proportional to the mass of the particle.

### 2.1.2 Interactions of the SM

Figure 2.2 serves as an overview of all interactions in the SM. It shows the SM Lagrangian ($\mathcal{L}_{\text{SM}}$) in its most concise form and Feynman diagrams of the interaction vertices related to the terms in the Lagrangian. The boxes indicate the coupling strengths of the vertices. The given parameters will be introduced below.

**(1)** describes the kinetic energy of gauge bosons and interactions among each other.

**(2)** encodes the kinetic energy of the fermions and interaction of the fermions with the gauge bosons.

**(3)** represents the Yukawa interactions of fermions with the Higgs field.

**(4)** describes interactions of the Higgs field with the electroweak gauge bosons and the Higgs boson self-interactions.

[1] Antifermions are marked by a bar above their acronym.

[2] From the observation of neutrino oscillations it is known that neutrinos have a non-vanishing mass [29, 30]. The weakest upper limit on the neutrino mass is set for $\nu_\tau$ as $m_{\nu_\tau} < 18.2\,\text{MeV}$ [28]. Observations from cosmology even constrain the summed neutrino mass of all flavours to be $\sum m_\nu < 0.12\,\text{eV}$ [31]. Given that $pp$ collision data with $\sqrt{s} = 13\,\text{TeV}$ is analysed in this thesis, the neutrino masses can be neglected to a very good approximation in this thesis.

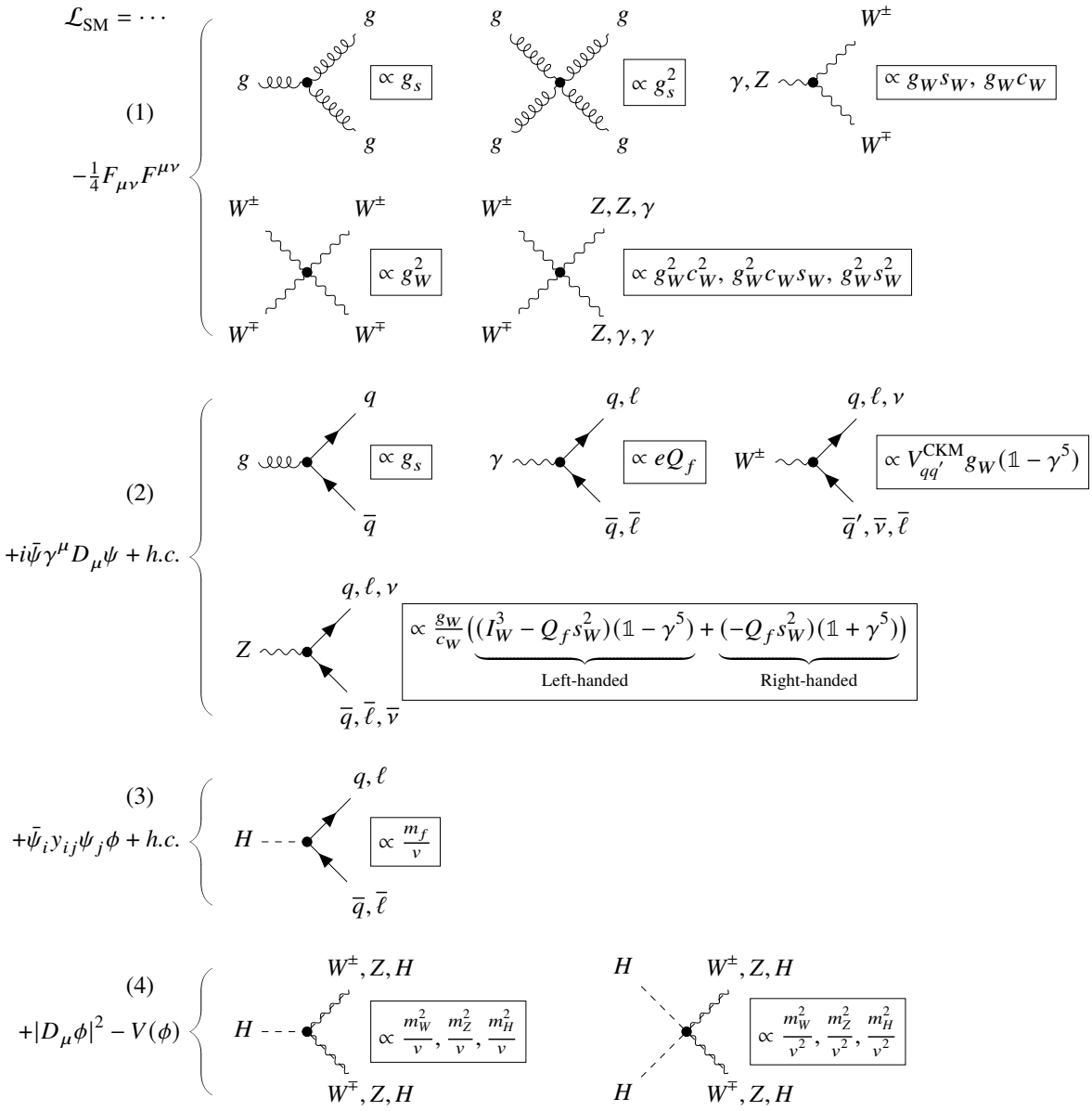



Figure 2.2: Interaction vertices of the SM Lagrangian. $s_W \coloneqq \sin\theta_W$, $c_W \coloneqq \cos\theta_W$.

### Yang-Mills Lagrangian

The Yang-Mills theory [32] is a QFT-based gauge theory based on special unitary groups SU(N). The most general renormalisable Lagrangian invariant under gauge transformations is

$$\mathcal{L}_{\mathrm{YM}} = -\frac{1}{4}(F^a_{\mu\nu})^2 + \bar{\psi} i\gamma^\mu D_\mu \psi - m\bar{\psi}\psi, \tag{2.1}$$

where the index $a$ is summed over the generators of the gauge group and $m$ is the mass of the fermion field $\psi$. $F^a_{\mu\nu}$ is the field tensor defined as

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a{}_{bc} A^b_\mu A^c_\nu \tag{2.2}$$

where $A^a_\mu$ are the gauge boson fields, $g$ is the coupling constant, and $f^{abc}$ are the structure constants of the gauge group. The first term in Equation 2.1 describes the gauge boson fields. If the structure constant of the symmetry group does not vanish, terms cubic and quartic in the boson fields occur, describing triple and quartic self-interaction vertices of the gauge bosons.

The covariant derivative $D_\mu$ is defined in terms of the representation matrices $t_a$ (generators of the group) as

$$D_\mu = \partial_\mu - igA^a_\mu t_a.$$

The second and third terms in Equation 2.1 describe the free fermion field ($\bar{\psi}[i\gamma^\mu\partial_\mu - m]\psi$) and interactions of the fermions with the gauge bosons ($g\bar{\psi}\gamma^\mu A^a_\mu t_a \psi$).

In contradiction to its common name, coupling constant, $g$ does depend on the energy scale being probed. Infinities in the calculation of observable quantities are absorbed in the definition of $g$ (renormalisation) – $g$ is a running coupling constant.

### Strong interaction

The theory describing the strong interaction is quantum chromodynamics (QCD). QCD is a non-Abelian Yang-Mills [32] gauge theory with symmetry group SU(3). The fundamental charge of QCD is colour charge, with three distinct colours: red, green, and blue (r, g, b) and anticolours ($\bar{\mathrm{r}}$, $\bar{\mathrm{g}}$, $\bar{\mathrm{b}}$). Colour charge is the conserved quantity of QCD.

The Lagrangian of QCD is derived by enforcing invariance under SU(3) local phase transformations,

$$\psi(x) \rightarrow \psi'(x) = \exp\left[ig_s\alpha_a(x)T^a\right]\psi(x) \qquad\qquad a = 1, 2, \cdots, 8.$$

$g_s$ is the coupling constant of the strong interaction, $\alpha_a(x)$ are transformation functions dependent on the spacetime coordinate $x$, and $T^a = \lambda^a/2$ are the generators of the group, with $\lambda^a$ being the Gell-Mann matrices. The local phase transformation can be thought of as a rotation in colour space.
The boson fields are identified as the eight gluon fields ($A^a_\mu \equiv G^a_\mu$) and, in order for the Lagrangian to be

gauge invariant, have to transform as

$$G^k_\mu \to G^{k\prime}_\mu = G^k_\mu - \partial_\mu \alpha^k(x) - g_s f_{ij}{}^k \alpha^i(x) G^j_\mu, \qquad \text{with} \quad D_\mu = \partial_\mu - \frac{i}{2} g_s G^a_\mu \lambda_a. \tag{2.3}$$

The non-zero components of the SU(3) structure constants are

$$f_{123} = 1, \quad f_{458} = f_{678} = \sqrt{3}/2, \quad f_{147} = f_{165} = f_{246} = f_{257} = f_{345} = f_{376} = 1/2.$$

Inserting Equation 2.3 into 2.1 yields the full Lagrangian of QCD,

$$\begin{aligned} \mathcal{L}_{\text{QCD}} = & \underbrace{-\frac{1}{4}(\partial_\mu G^i_\nu - \partial_\nu G^i_\mu)(\partial^\mu G^{i\nu} - \partial^\nu G^{i\mu})}_{\text{Gluon kinetic term}} \\ & + \underbrace{\frac{1}{2} g_s f^{ijk}(\partial_\mu G^i_\nu - \partial_\nu G^i_\mu) G^j_\mu G^k_\nu - \frac{1}{4} g_s^2 f_{ijk} f^{imn} G^{j\mu} G^{k\nu} G_{m\mu} G_{n\nu}}_{\text{Gluon self-interaction term}} \\ & + \underbrace{\bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi}_{\text{Free fermion term}} - \underbrace{\bar{\psi} \frac{i}{2} g_s G^i_\mu \lambda_i \psi}_{\text{Gluon fermion interaction term}} \end{aligned} \tag{2.4}$$

The Feynman diagrams of the interaction terms are visualised in Figure 2.2.

The self-interactions of the massless gauge bosons have particular implications on the strong interaction. The strong coupling strength is larger at small energies (large distances) and smaller at high energies (small distances). The dimensionless coupling constant of the strong interaction $\alpha_s$ varies depending on the considered energy scale, e.g. $\alpha_s(200\,\text{MeV}) \approx 1.0$ and $\alpha_s(m_Z) \approx 0.118$.

At large distances the coupling becomes so large that coloured objects (hadrons) are always confined to colour-neutral states (colour confinement). This behaviour can be explained in a quantitative picture by two free quarks that are being separated. The quarks interact via the exchange of virtual gluons, which themselves interact via gluons. The interaction squeezes the gluons into a "flux tube" between the quarks (cf. Figure 2.3). This is in contrast to the electromagnetic field lines where the field lines spread out. The energy stored in the colour field increases almost linearly at larger distances, which is experimentally measured to be $\sim 1\,\text{GeV/fm}$ [23]. That being so, if quarks are separated at macroscopic distances, it is energetically favourable to break the flux tube by producing a quark-antiquark pair from the vacuum.

Thus, quarks are bound in either $q\bar{q}$ states (mesons) with colour and anticolour, or $qqq$ / $\bar{q}\bar{q}\bar{q}$ states (baryons / antibaryons) with colour rgb / $\bar{\text{r}}\bar{\text{g}}\bar{\text{b}}$.[3] In contrast, at high energies the quarks inside hadrons can be seen as quasi-free particles (asymptotic freedom).

### Electroweak interaction

The theory of electroweak interaction unifies the weak and electromagnetic interactions. The underlying mathematical concept is similar to QCD. It is a non-Abelian Yang-Mills gauge theory, but the electroweak symmetry group is $\text{SU}(2)_L \times \text{U}(1)$. The subscript "$L$" indicates that the interaction respecting the SU(2)

[3] More complex states like tetraquarks ($q\bar{q}q\bar{q}$), pentaquarks ($qqqq\bar{q}$) or glueballs are also possible but insignificant for this thesis.

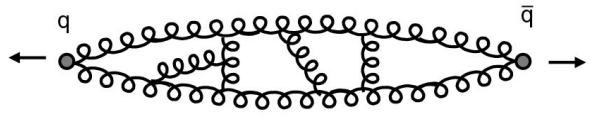



Figure 2.3: Gluon interactions between two separating quarks. [23]

symmetry couples exclusively to left-handed chiral particles and right-handed chiral antiparticles[4]. The chiral projection operators are defined in terms of the Dirac matrices

$$\psi_{L/R} = P_{L/R}\psi, \qquad \text{with} \quad P_L = \frac{\mathbb{1} - \gamma_5}{2} \quad \text{and} \quad P_R = \frac{\mathbb{1} + \gamma_5}{2}.$$

A major difference with respect to QCD is that the electroweak symmetry is broken by the masses of the gauge bosons. The mechanism of symmetry breaking is described in Section 2.1.2. The consequence is that the gauge fields do not correspond to the physically observed fields.

The fundamental charges of the unbroken symmetry are the weak isospin $I_W$ of the $\mathrm{SU}(2)_L$ symmetry and the weak hypercharge $Y$ of the U(1) symmetry. Left-handed fermions of the same generation are placed in weak isospin doublets,

$$Q_L^i \coloneqq \begin{pmatrix} U^i \\ D^i \end{pmatrix}_L = \left( \begin{pmatrix} \psi_u \\ \psi_d \end{pmatrix}_L \quad \begin{pmatrix} \psi_c \\ \psi_s \end{pmatrix}_L \quad \begin{pmatrix} \psi_t \\ \psi_b \end{pmatrix}_L \right), L_L^i \coloneqq \begin{pmatrix} N^i \\ E^i \end{pmatrix}_L = \left( \begin{pmatrix} \psi_e \\ \psi_{\nu_e} \end{pmatrix}_L \quad \begin{pmatrix} \psi_\mu \\ \psi_{\nu_\mu} \end{pmatrix}_L \quad \begin{pmatrix} \psi_\tau \\ \psi_{\nu_\tau} \end{pmatrix}_L \right),$$
$$\text{with} \quad I_W = \frac{1}{2}, \text{ and third component } I_W^3 = \begin{pmatrix} 1/2 \\ -1/2 \end{pmatrix}. \quad (2.5)$$

Right-handed fermions are placed in weak isospin singlets,

$$U_R^i = \begin{pmatrix} \psi_{u,R} & \psi_{c,R} & \psi_{t,R} \end{pmatrix}, D_R^i = \begin{pmatrix} \psi_{d,R} & \psi_{s,R} & \psi_{b,R} \end{pmatrix}, E_R^i = \begin{pmatrix} \psi_{e,R} & \psi_{\mu,R} & \psi_{\tau,R} \end{pmatrix},$$
$$\text{with} \quad I_W = I_W^3 = 0. \quad (2.6)$$

Neutrinos, being massless, cannot possess right-handed states. Owing to the fact that the weak isospin singlets are unaffected by the $\mathrm{SU}(2)_L$ local gauge transformation, they do not couple to the gauge bosons of the symmetry. The weak hypercharge of a particle is defined as

$$Y = 2(Q - I_W^3)$$

with $Q$ being the electric charge of the particle.

The Lagrangian of the electroweak interaction is derived by enforcing invariance under $\mathrm{SU}(2)_L \times \mathrm{U}(1)$ local phase transformations,

$$\psi(x) \to \psi'(x) = \exp\left[ig_W \alpha_a(x) T^a\right] \exp\left[ig_Y \beta(x)/2\right] \psi_L(x) + \exp\left[ig_Y \beta(x)/2\right] \psi_R(x) \quad \text{with} \quad a = 1, 2, 3.$$

[4]The chirality of a particle is defined by whether it transforms in a right- or left-handed representation of the Poincaré group. For massless particles, chirality and helicity, the projection of the spin onto the direction of motion, are the same.

$g_W$ and $g_Y$ are the coupling constants of the $SU(2)_L$ and U(1) symmetries, respectively, and $\alpha_a(x)$ and $\beta(x)$ are transformation functions. The generators of the SU(2) group are the Pauli matrices $\sigma^a$ ($T^a = \sigma^a/2$). There are three boson fields related to the $SU(2)_{I_W}$ symmetry ($A^a_\mu \equiv W^a_\mu$) and one boson field related to the $U(1)_Y$ symmetry ($A_\mu \equiv B_\mu$). The field transformation properties and the covariant derivative are given by

$$W^k_\mu \to W^{k\prime}_\mu = W^k_\mu - \partial_\mu \alpha^k(x) - g_W \epsilon_{ij}{}^k \alpha^i(x) W^j_\mu, \qquad B_\mu \to B'_\mu = B_\mu - \partial_\mu \beta(x),$$
$$\text{with} \quad D_\mu = \partial_\mu - \frac{i}{2}\left(g_W W^a_\mu \sigma_a - g_Y B_\mu Y\right). \tag{2.7}$$

The full electroweak Lagrangian is retrieved by inserting Equation 2.7 into 2.1. The resulting terms are very similar to the ones of the QCD Lagrangian (Equation 2.4), but with two main differences: (i) the $W_\mu$ boson fields couple only to left-handed fermion fields, and (ii) there are no self-interactions of $B_\mu$ bosons.

As mentioned beforehand, the electroweak gauge fields do not correspond to the physical fields, i.e. the weak eigenstates differ from the mass eigenstates that are fundamental for the free-particle propagation. The four fields mix and build the physical $W^\pm$ bosons, the $Z$ boson, and the $\gamma$. The $W^\pm$ bosons, as they couple fermions from the different weak isospin doublets, are related to the weak isospin raising and lowering operators $\sigma_\pm = (\sigma_1 \pm \sigma_2)/2$, and therefore

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^{(1)}_\mu \mp W^{(2)}_\mu\right).$$

The $W^\pm$ bosons mediate the weak charged current, and the coupling strength is determined by $g_W I^3_W$. $W^\pm$ bosons do not couple to right-handed particles, and parity is thus maximally violated in weak charged-current interactions.

The $W^{(3)}_\mu$ and $B_\mu$ fields mix to build the $Z$ boson ($Z_\mu$) and the photon ($A_\mu$) fields. The mixing is described by a rotation from the weak to the mass eigenbasis,

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = R^T_{\theta_W} \begin{pmatrix} B_\mu \\ W^{(3)}_\mu \end{pmatrix} \qquad \text{with} \quad R_{\theta_W} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \tag{2.8}$$

The corresponding rotation angle is the weak mixing angle $\theta_W$ and is experimentally measured to be $\sin^2\theta_W = 0.23129 \pm 0.00004$ [28]. The coupling strength of the $\gamma$ to other particles is determined by the well-known electromagnetic coupling constant $e$ with

$$e = g_Y \cos\theta_W = g_W \sin\theta_W$$

The photon couples only to particles with electric charge and couples equally to left- and right-handed particles. The $Z$ boson couples to all particles with weak isospin or weak hypercharge. The coupling strength of the $Z$ boson differs for left- and right-handed particles. All the interaction vertices of the electroweak interactions and their coupling strengths are summarised in Figure 2.2.

As the eigenstates of the weak interaction do not correspond to the physical mass eigenstates, the weak charged-current interaction can couple fermions from different generations, i.e. it is capable of changing

the flavour of fermions. The Cabibbo-Kobayashi-Maskawa (CKM) matrix links the two eigenstates in the quark sector[5] by representing the weak eigenstates as a mixture of the mass eigenstates. By convention, the up-type quarks' weak and mass eigenstates are identical, and the mixing occurs for the down-type quarks,

$$\underbrace{\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}}_{\text{Weak eigenstates}} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{\text{CKM}}} \underbrace{\begin{pmatrix} d \\ s \\ b \end{pmatrix}}_{\text{Mass eigenstates}} \qquad |V_{\text{CKM}}| \approx \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.042 \\ 0.009 & 0.041 & 0.999 \end{pmatrix} \tag{2.9}$$

The CKM matrix is a unitary matrix. The magnitudes of the CKM matrix elements, omitting uncertainties, are taken from [28], estimated from global fits of all SM parameters. The matrix elements describe the coupling strengths of the $W$ boson to different quark pairs. The transition from up-type quark ($i$) to down-type quark ($j$) is proportional to $|V_{ij}|^2$. Transitions within the same generation are likely and transitions between different generations are suppressed.

### The Higgs mechanism

The last term in the Yang-Mills Lagrangian (Equation 2.1), which corresponds to the mass of the fermion field, is actually not gauge invariant under the electroweak symmetry group $\mathrm{SU}(2)_L \times \mathrm{U}(1)$. Therefore, the fermion fields have to be massless ($m = 0$). The same is true for Dirac mass terms for the bosons of the electroweak symmetry. But either are known to be massive.

In the SM particles acquire their mass by interacting with the Higgs field. In the SM the Higgs mechanism is embedded in the electroweak symmetry and the Higgs field needs to have the necessary degrees of freedom to explain the masses of the electroweak gauge boson. Therefore, the Higgs field in the SM consists of a charged and a neutral complex scalar field placed in a weak isospin doublet ($Y = 1$)

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

The Lagrangian for a scalar field is given by the Klein-Gordon Lagrangian,

$$\mathcal{L}_H = (D_\mu \phi)^2 - V(\phi). \tag{2.10}$$

The form of the covariant derivative is unchanged with respect to the one derived in the electroweak Lagrangian (Equation 2.7), as in the SM the Higgs mechanism is embedded in the electroweak symmetry. The Higgs potential is given by

$$V(\phi) = \mu^2 |\phi|^2 + \lambda |\phi|^4.$$

The Higgs potential can contain only terms quadratic in $\phi$ to be local gauge invariant. $\mu$ and $\lambda$ are free scalar parameters of the theory. In order for the potential to have a well-defined minimum, $\lambda$ must be

[5] A similar mixture in the lepton sector only occurs if neutrinos are massive. The corresponding mixing matrix is called Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix.

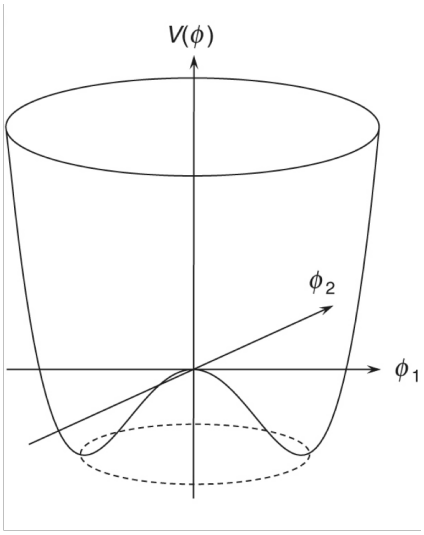



Figure 2.4: Sketch of the Higgs potential with $\lambda > 0$ and $\mu^2 < 0$ for a complex Higgs field $\phi = \phi_1 + i\phi_2$. A Higgs field with fewer dimensions than in the SM is chosen, as the concept is identical and it facilitates sketching the potential in three dimensions. The dashed line represents the set of minima. The figure is taken from Reference [23].

positive. If $\mu^2 < 0$, the potential has a non-zero vacuum expectation value $v$ given by

$$\phi\phi^\dagger = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}.$$

This condition is satisfied by an infinite set of points in the Higgs potential, indicated by the dashed line in Figure 2.4. The choice of the vacuum state breaks the symmetry of the Lagrangian (EWSB).

One of the physical fields after symmetry breaking is the photon field, which is massless and couples to electric charge. Therefore, the vacuum state chosen by nature must have a vanishing vacuum expectation value for the charged scalar field $\phi^+ = 0$. In perturbation theory, the fields are expressed as small perturbations about this vacuum state,

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ v + \eta(x) + i\phi_4(x) \end{pmatrix}.$$

An excitation of the field $\eta(x)$ corresponds to an excitation in the direction where the potential changes to first order quadratic. Excitations of the fields $\phi_1(x), \phi_3(x), \phi_4(x)$ are in the direction where the potential does not change (Goldstone fields). When inserting this field into $\mathcal{L}_H$ (Equation 2.10) and requiring gauge invariance, one can identify $\eta(x)$ as being a massive field and the Goldstone fields as being massless. In addition, mass terms of the gauge bosons arise, i.e. terms quadratic in the gauge fields. The Goldstone fields are eliminated from the Lagrangian by choosing an appropriate gauge transformation. The transformation does not change the physical predication, but after transformation the fields are identical to the physical fields (unitarity gauge). The Higgs field in the unitarity gauge is given by

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{2.11}$$

where $\eta(x)$ was identified to be the physical Higgs field $h(x)$. Substituting the Higgs field in unitarity gauge (Equation 2.11) into the scalar Lagrangian (Equation 2.10) gives the Higgs sector Lagrangian, $\mathcal{L}_H$.

Eventually, the additional degrees of freedom of the Goldstone fields give mass to the gauge bosons (longitudinal polarisation states). From the Lagrangian, the mass of the $W^{\pm}$ bosons, the Higgs boson and, after disentangling the mixing of the fields, the masses of the $Z$ and $\gamma$ are

$$m_W = \frac{1}{2} v g_W, \qquad m_H = \sqrt{2\lambda} v, \qquad m_Z = \frac{1}{2} v \sqrt{g_W^2 + g_Y^2}, \qquad m_\gamma = 0.$$

The vacuum expectation value is measured to be approximately 246 GeV [28]. The Lagrangian also contains triple and quartic gauge couplings of the Higgs boson with the $W^{\pm}$, $Z$ boson and itself (cf. Figure 2.2).

#### Yukawa interaction

In the SM the fermions acquire mass by interacting with the Higgs field (Yukawa interaction). The term in the SM Lagrangian describing the Yukawa interactions has the following form:

$$\mathcal{L}_y = -\bar{Q}_L^i y_{ij}^U \phi U_R^j - \bar{Q}_L^i y_{ij}^D \phi D_R^j + \bar{L}_L^i y_{ij}^E \phi E_R^j + h.c.^{6}. \tag{2.12}$$

The Yukawa Lagrangian is gauge invariant under the electroweak $\mathrm{SU}(2)_L \times \mathrm{U}(1)$ symmetry, because $Q_L^i$, $U_R^j$, etc. are the weak isospin doublets and singlets defined in Equations 2.5 and 2.6, respectively. The *Yukawa couplings* $y_{ij}^U$, $y_{ij}^D$, and $y_{ij}^E$ are complex $3 \times 3$ matrices in the weak interaction basis and are diagonal in the mass basis. The fact that $y_{ij}^U$ and $y_{ij}^D$ cannot be diagonalised simultaneously gives rise to the CKM matrix introduced previously.

Writing the Higgs field in unitarity gauge (Equation 2.11) the Yukawa Lagrangian reads as

$$\mathcal{L}_y = \underbrace{-\frac{y_{ij}^U}{\sqrt{2}} v (\bar{Q}_L^i U_R^j + \bar{U}_R^i Q_L^j)}_{\text{Fermion mass term}} - \underbrace{\frac{y_{ij}^U}{\sqrt{2}} h (\bar{Q}_L^i U_R^j + \bar{U}_R^i Q_L^j)}_{\text{Fermion interaction with the Higgs}} + \cdots (\text{equivalents terms for } y_{ij}^D \text{ and } y_{ij}^E). \tag{2.13}$$

With $\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L = \bar{\psi}\psi$ the first term represents a classical Dirac mass term, but it is introduced to the SM Lagrangian in a gauge-invariant way through the interaction with the non-zero expectation value of the Higgs field. The Yukawa couplings are free parameters of the SM and are chosen to match the observed fermion masses, i.e. in mass basis

$$y_{ij} = \sqrt{2} \frac{m_i}{v} \delta_{ij}. \tag{2.14}$$

The second term in Equation 2.13 describes the coupling of the Higgs boson to fermions, which consequently is proportional to the mass of the fermions.

[6] *h.c.* stands for "Hermitian conjugate" and means that there is an additional Hermitian conjugate term for each preceding term.

### 2.1.3 Limitations of the SM

The SM has been very successful in describing experimental data from particle physics experiments. But for several reasons it is known that the SM is not the desired theory of everything. There are, on the one hand, physical observations that cannot be explained by the SM and require the introduction of new physics. On the other hand, there are structural or aesthetic weaknesses of the SM that do not require the existence of new physics, but new physics might give profound answers to these questions. The most crucial limitations are outlined below.

#### Gravitation

The gravitational force is not part of the SM. Gravity is well described by the theory of general relativity. So far it has not been possible to unite general relativity and QFT due to renormalisability issues. Above the Planck scale ($\Lambda_P \sim 10^{19}$ GeV) quantum effects of gravity are expected to dominate. Hence, the SM is not capable of describing the physics of the very early Universe.

#### Dark matter and dark energy

Many cosmological and astrophysical observations suggest the existence of non-baryonic, electrically neutral matter (dark matter). So far dark matter has been observed only indirectly by its gravitational impact on ordinary matter, among which are the galaxy rotational curves [33], gravitational lensing [34], and the large-scale structure formation of the Universe [35]. The SM does not provide a candidate for dark matter that could explain all these observations.

In the Lambda-Cold-Dark-Matter model, which is the most renowned model in cosmology and is based on the Big Bang theory, the Universe consists of ordinary matter, dark matter and dark energy. Dark energy describes a form of energy that is uniformly distributed in the Universe and that has a constant energy density with respect to the cosmological scale factor. Dark energy is thought to be responsible for the accelerating expansion of the Universe. Once again, the SM cannot explain its origin. Measurements of the cosmological microwave background conclude that ordinary matter makes up only $\sim 5\,\%$ of the mass-energy content of the Universe, whereas dark matter makes up $\sim 27\,\%$ and dark energy even $\sim 68\,\%$ [31].

#### Baryon asymmetry

If the theory of the Big Bang is correct, the evolution of the Universe started with quantum fluctuations in a singularity. From there on, matter could be produced as matter-antimatter pairs. Though today's Universe consists almost entirely of matter/baryons. Sakharov formulated three conditions that must be satisfied to result in a baryon asymmetry in the Universe [36]: (i) baryon number violation; (ii) departure from thermal equilibrium; (iii) C and CP violation. CP violation occurs in the SM in weak interactions as part of the CKM (PMNS) matrix. But its magnitude is not sufficient to explain the observed size of baryon asymmetry in the Universe.

#### Gauge hierarchy problem

The mass of the Higgs boson depends on its bare mass $m_0$ and higher-order corrections due to quantum loops in the propagator of the Higgs boson,

$$m_H^2 = m_0^2 - \frac{g^2}{8\pi^2}\Lambda^2 + \cdots ,$$

where $g$ refers to the coupling of the considered loop particle and the Higgs boson – in the case of fermions this is the Yukawa coupling – and $\Lambda$ is the considered ultraviolet cutoff scale. If the SM is a valid theory up to an energy scale where the three forces of the SM might unify ($\Lambda_{\text{GUT}} \sim 10^{16}\,\text{GeV}$) or even the Planck scale where gravity becomes dominant, all the large terms in the Higgs mass have to cancel each other almost perfectly to end up at the observed (comparatively small) mass of 125 GeV (fine-tuning). In certain models beyond the SM this cancellation appears naturally, e.g. Supersymmetry. [37, 38]

**Strong CP problem**

One of the free parameters of the SM is the strong CP phase, which allows for CP violation in the QCD Lagrangian. This phase is experimentally known to be very small, $\theta_{\text{CP}} \simeq 0$. A popular explanation for this result is given by Peccei and Quinn [13]. By postulating a new U(1) symmetry, one could eliminate the CP-violating phase from the Lagrangian. Imposing this symmetry, while requiring gauge invariance of the Lagrangian, requires an extended Higgs sector, at least a second Higgs doublet.

## 2.2 Physics beyond the Standard Model

The limitations of the SM portrayed in the previous section demonstrate the desperate need for physics beyond the SM. Given the success of the SM, it is natural to think of simple extensions of the SM, which leave predictions of already validated precision measurements unchanged and still solve one or even multiple of the open questions. This is indeed true for a wide range of extensions of the scalar sector.

A restriction on extended scalar sectors is given by the $\rho$-parameter, which is at tree-level given by [22]

$$\rho = \frac{\sum_{i=1}^{n} \left[ I_{w,i}(I_{w,i}+1) - \frac{1}{4}Y_i^2 \right] v_i}{\sum_{i=1}^{n} \frac{1}{2}Y_i^2 v_i} \tag{2.15}$$

where the index $i$ runs over the number of scalar multiplets $n$. The $\rho$-parameter is experimentally measured to be close to 1 [28]. Electroweak SU(2) singlets (with $Y = 0$) and doublets (with $Y = \pm 1$) are compatible with $\rho = 1$, as well as certain larger multiplets or multiplets with vanishing vacuum expectation value.

The simplest extension in which a charged Higgs boson appears is the addition of a second weak isospin doublet of complex scalar fields to the scalar sector [39]. The phenomenologies of such two-Higgs-doublet models (2HDMs) are discussed thoroughly in Section 2.2.1. Also discussed in more detail are supersymmetric models in Section 2.2.2, as those are very popular and all of them predict the existence of a charged scalar. The theoretical concept and motivations of other popular models predicting a charged Higgs boson are covered in Section 2.2.3.

### 2.2.1 Two-Higgs-Doublet models

#### Introduction to 2HDMs

Two-Higgs-doublet models predict the Higgs field to be made up of two (in the SM it is only one) weak isospin doublets of complex scalar fields,

$$\Phi = \Phi_1 + \Phi_2 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix} + \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_5 + i\phi_6 \\ \phi_7 + i\phi_8 \end{pmatrix}.$$

The 2HDM Lagrangian reads as

$$\mathcal{L}_{\text{2HDM}} = (D_\mu \Phi_1)^2 + (D_\mu \Phi_2)^2 - V_{\text{2HDM}}(\Phi) + \mathcal{L}_{y,\text{2HDM}}. \tag{2.16}$$

As the Lagrangian still has to respect the electroweak symmetry, the form of the covariant derivative $D_\mu$ is unchanged with respect to the SM case (Equation 2.7). On the contrary, the most general potential has become more complex [22],

$$\begin{aligned} V_{\text{2HDM}}(\Phi) =& m_{11}^2 \Phi_1^2 + m_{22}^2 \Phi_2^2 - m_{12}^2 (\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1) \\ &+ \frac{\lambda_1}{2} \Phi_1^4 + \frac{\lambda_2}{2} \Phi_2^4 + \lambda_3 \Phi_1^2 \Phi_2^2 + \lambda_4 (\Phi_1^\dagger \Phi_2)(\Phi_2^\dagger \Phi_1) \\ &+ \left[ \frac{\lambda_5}{2} (\Phi_1^\dagger \Phi_2)^2 + \lambda_6 \Phi_1^2 (\Phi_1^\dagger \Phi_2) + \lambda_7 \Phi_2^2 (\Phi_1^\dagger \Phi_2) + h.c. \right]. \end{aligned} \tag{2.17}$$

The parameters $m_{12}$, $\lambda_5$, $\lambda_6$, and $\lambda_7$ are complex, which means there are in total 14 real parameters. In the most general case, without additional symmetries, three degrees of freedom can be eliminated by choosing the appropriate gauge transformation. Thus, the potential has 11 physical degrees of freedom. In comparison, the SM Higgs potential has only two degrees of freedom ($\mu$ and $\lambda$).

The existence of a second Higgs doublet also modifies the Yukawa interactions. The Yukawa Lagrangian in 2HDMs is given by

$$\begin{aligned} \mathcal{L}_{y,\text{2HDM}} =& - \bar{Q}_L^i [y_{ij}^{U,1} \tilde{\Phi}_1 + y_{ij}^{U,2} \tilde{\Phi}_2] U_R^j - \bar{Q}_L^i [y_{ij}^{D,1} \Phi_1 + y_{ij}^{D,2} \Phi_2] D_R^j \\ &+ \bar{L}_L^i [y_{ij}^{E,1} \Phi_1 + y_{ij}^{E,2} \Phi_2] E_R^j + h.c., \end{aligned} \tag{2.18}$$

where $\tilde{\Phi}_k = \epsilon_{ij} \Phi_k^*$ and $\epsilon_{ij}$ denotes the totally antisymmetric tensor in two dimensions.

The large number of free parameters in the 2HDM potential and the increased complexity of the Yukawa couplings make phenomenological studies difficult. The aim of the following section is to first give a general introduction to 2HDMs and then describe specific models in more detail.

#### Motivation for 2HDMs

2HDMs are simple and very popular extensions of the SM for several reasons. Firstly, the addition of a second Higgs doublet to the SM introduces new sources of CP violation, which, as mentioned before, are needed to describe the baryon asymmetry in the Universe. CP can either be explicitly violated by two complex angles in the new scalar potential $V_{\text{2HDM}}(\Phi)$, which do not vanish, or spontaneously by the minima of the potential which do have a complex phase.

Another reason why it makes sense to discuss 2HDMs thoroughly is that new physics models where new symmetries are imposed often require the addition of a second Higgs doublet. Among them is SUSY, which needs in its minimal extension (MSSM) at least two Higgs doublets, and the U(1) symmetry in the Peccei and Quinn model.

Generally, an extended Higgs sector in connection with a new symmetry can address various open questions. For example, the new symmetry can promote a particle from the extended scalar sector to be a dark-matter candidate [11, 12, 40]. Or a 2HDM in which all quarks and charged leptons couple to one Higgs doublet (with $v_1 \sim$ GeV) and the neutrinos to the other Higgs doublet (with $v_2 \sim$ eV) – neutrino-specific 2HDM – does give a natural explanation for the large mass difference between neutrinos and the other fermions [41, 42].

In addition, 2HDMs offer explanations for discrepancies between certain experimental measurements and SM predictions. An example is the $4.2\sigma$ [43] discrepancy in the measurement of the muon anomalous magnetic moment (muon g-2) [44, 45] [7]. A 2HDM model with an additional $\mathrm{U}(1)_{L_\mu - L_\tau}$ symmetry [48] – $L_\mu$ and $L_\tau$ are the muon and tau lepton numbers, respectively – can justify the discrepancy. Other anomalies are the $3.1\sigma$ [49] discrepancy on the value of $R(D^{(*)}) = \mathrm{BR}(B^- \to D^{(*)}\tau^-\overline{\nu})/\mathrm{BR}(B^- \to D^{(*)}\ell\overline{\nu})$ averaged over multiple experiments, and the (controversial) $W$-boson mass measurement from CDF-II where a $7\sigma$ discrepancy [50][8] is seen. Multiple phenomenological studies on 2HDMs exist to address these anomalies [53–56].

### Minima of the general 2HDM potential

With the introduction of the second Higgs doublet, the shape of the 2HDM potential (Equation 2.17) can become more complex. It can have three types of minima [22],

$$\begin{aligned}
&\text{CP-conserving} && \langle\Phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v_1\end{pmatrix}, && \langle\Phi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v_2\end{pmatrix},\\
&\text{CP-violating} && \langle\Phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ \frac{v_1}{\sqrt{2}}e^{i\theta}\end{pmatrix}, && \langle\Phi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v_2\end{pmatrix},\\
&\text{charge-violating} && \langle\Phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}\frac{\alpha}{\sqrt{2}}\\ \frac{v_1}{\sqrt{2}}e^{i\theta}\end{pmatrix}, && \langle\Phi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v_2\end{pmatrix}.
\end{aligned}$$

$v_1$, $v_2$, $\theta$, and $\alpha$ are real parameters, with $v_1$ and $v_2$ being the vacuum expectation values of the two doublets. If the minimum is CP-/charge-violating, CP/charge is spontaneously broken.

In general, not all parametrisations of the potential ensure that there is a stable minimum at all. But there can be multiple sets of degenerate minima, i.e. multiple sets of $v_1, v_2$, and tunnelling between the minima is possible. Phenomenological studies have proven that different types of minima cannot coexist [22]. Up to two CP-conserving minima can coexist [57]. In contrast, if there is a minimum violating CP or breaking charge conservation, it is the only minimum of the potential and therefore global and stable. The latter one results in massive photons and can therefore be completely excluded anyhow.

[7] New lattice QCD calculations on the hadronic vacuum polarisation and hadronic light-by-light scattering suggest that the discrepancy is less [46, 47].

[8] Measurements of the $W$-boson mass from ATLAS [51] and CMS [52] agree with the SM prediction.

### CP-conserving 2HDMs

For a further introduction to the mathematical concept, the focus is set on models with CP-conservation – $V_{\text{2HDM}}(\Phi)$ is entirely real and the minimum is CP-conserving – to ease the mathematical formulations.

Similar to the SM Higgs mechanism, the doublets are expanded about the minima,

$$\Phi_a = \begin{pmatrix} \phi_a^+ \\ (v_a + \rho_a + i\eta_a)/\sqrt{2} \end{pmatrix}, \quad a = 1, 2.$$

When inserting the expanded fields into the 2HDM Lagrangian from Equation 2.16, terms quadratic in the fields arise from which mass matrices $\mathcal{M}$ can be extracted. Similar to the weak mixing of the $W_\mu$ and $B_\mu$ fields, a diagonalisation of these matrices can be achieved by multiplication with a rotation matrix with an appropriate rotation angle. The rotation angle diagonalising the $\mathcal{M}_{\phi^\pm\text{mass}}$ and $\mathcal{M}_{\eta\text{mass}}$ matrices is $\beta$ and is given by

$$\tan\beta = \frac{v_2}{v_1}. \tag{2.19}$$

The corresponding rotation angle for $\mathcal{M}_{\rho\text{mass}}$ is denoted as $\alpha$. The physical fields are

$$\begin{pmatrix} G^\pm \\ H^\pm \end{pmatrix} = R_\beta^T \begin{pmatrix} \phi_1^\pm \\ \phi_2^\pm \end{pmatrix}, \qquad \begin{pmatrix} G^0 \\ A \end{pmatrix} = R_\beta^T \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, \qquad \begin{pmatrix} H \\ h \end{pmatrix} = R_\alpha^T \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}. \tag{2.20}$$

$G^\pm$ and $G^0$ are the Goldstone fields, with eigenvalues of zero which eventually give mass to the $W^\pm$ and $Z$ bosons. The remaining five fields give rise to scalar bosons: two charged scalars $H^\pm$, two CP-even neutral scalars $h$ and $H$, whereby $H$ is defined to be the heavier one of the two, and one CP-odd pseudoscalar $A$. The mass term of the charged scalar field reads as

$$m_{H^\pm}^2 = \left[ \frac{m_{12}^2}{v_1 v_2} - (\lambda_4 + \lambda_5) \right] (v_1^2 + v_2^2). \tag{2.21}$$

In addition, terms trilinear and quartic in the fields are present in the 2HDM Lagrangian, giving rise to three- and four-point interaction vertices between the bosons. The Feynman diagrams can be seen in Figure 2.2 bracket (4).

The SM Higgs boson can be expressed as

$$H^{\text{SM}} = h \sin(\alpha - \beta) - H \cos(\alpha - \beta).$$

The scenario $\sin(\alpha - \beta) \to 1$ is supported by experimental measurements [58], which means $h$ has the same couplings as the SM Higgs boson (alignment limit).

In summary, the CP-conserving 2HDMs have 7 free parameters, which can be parametrised by

- the masses of the new particles: $m_h, m_H, m_A, m_{H^\pm}$,
- the mixing angles: $\alpha, \beta$,
- and the diagonal term of the mass matrix of the Higgs doublets $m_{12}$.

Table 2.1: Models with natural flavour conservation and the respective couplings of the right-handed singlets to the Higgs doublets. The right column shows a possible discrete $\mathbb{Z}_2$ symmetry which can be imposed to obtain the desired couplings.

| Models | $U_R^i$ | $D_R^i$ | $E_R^i$ | Imposed $\mathbb{Z}_2$ symmetry |
|---|---|---|---|---|
| Type-I | $\Phi_2$ | $\Phi_2$ | $\Phi_2$ | $\Phi_1 \to -\Phi_1$ |
| Type-II | $\Phi_2$ | $\Phi_1$ | $\Phi_1$ | $\Phi_1 \to -\Phi_1, D_R^i \to -D_R^i$ |
| Type-X (lepton-specific) | $\Phi_2$ | $\Phi_2$ | $\Phi_1$ | $\Phi_1 \to -\Phi_1, E_R^i \to -E_R^i$ |
| Type-Y (flipped) | $\Phi_2$ | $\Phi_1$ | $\Phi_2$ | $\Phi_1 \to -\Phi_1, D_R^i \to -D_R^i, E_R^i \to -E_R^i$ |

Considering the observed Higgs boson to be the light CP-even boson and the alignment limit, the parameters $m_h$ and $\sin(\alpha - \beta)$ are known, and 5 free parameters are left. The connections of $m_{H^\pm}$ to the other parameters are non-trivial. The couplings of the charged Higgs boson are determined by the rotation angle $\beta$.

It should be noted again that these calculations are applicable only for the CP-conserving minima. For CP-violating minima the calculations would be more involved as more parameters of the potential are unequal to zero, e.g. the mixing of the neutral Higgs states has to be described by a three-dimensional rotation matrix, resulting in three mixing angles $\alpha_i$. Nonetheless, the prediction of the five additional bosons ($H^\pm$, $H$, $h$, $A$) is common for all models. The clear distinction of the CP-even and -odd states of the neutral Higgs bosons can be made only for the CP-conserving case.

#### 2HDMs with natural flavour conservation

The Yukawa coupling matrices $y_{ij}^{X,1}$ and $y_{ij}^{X,2}$ ($X \in [U, D, E]$) are in general not simultaneously diagonalisable in the mass basis. Therefore, couplings of the Higgs field with fermions of different generations appear, e.g. $\bar{\psi}_d \phi \psi_s$. The result is flavour-changing neutral currents (FCNC) mediated by the neutral Higgs scalars.

Models with large FCNCs are disfavoured as experiments have set strict limits on FCNCs [59]. In popular 2HDMs, natural – in the sense that the Yukawa matrices are not aligned by coincidence – flavour conservation (NFC) is achieved by imposing new discrete or continuous symmetries. In case all fermions with the same quantum numbers couple to the same Higgs doublet – one of the matrices $y_{ij}^{X,1}$ or $y_{ij}^{X,2}$ vanishes for each flavour $X \in [U, D, E]$ – FCNCs will be absent. This can be achieved through the introduction of a new symmetry, forcing the right-handed states with the same charge to couple to the same Higgs doublet. The possible combinations of couplings are presented in Table 2.1, whereas, without loss of generality, up-type quarks are assumed to couple to $\Phi_2$.

With the help of the relations in Equation 2.20, the Yukawa interactions from Equation 2.18 can be expressed in terms of the physical fields. The Yukawa couplings can be replaced by the masses of the fermions, similar to Equation 2.14, but with $v = \sqrt{v_1^2 + v_2^2}$. The coupling strength of the charged Higgs boson to fermions is proportional to the masses of the fermions, $V_{\mathrm{CKM}}$, and $\tan\beta$ or $\cot\beta$, depending on the model being considered.

Other models with FCNCs at tree-level, like Type-III models and models with minimal flavour violation, are discussed in References [22, 60].

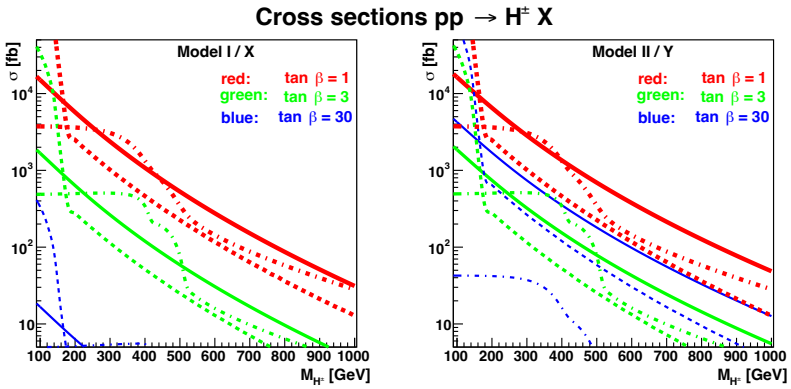



Figure 2.5: Production cross-sections of charged Higgs bosons in $pp$ collisions at 14 TeV for different 2HDM models. Solid: $gb \to tH^{\pm}$; Dashed: $gg \to tH^{\pm}b$; Dash-dotted: $gg \to H/h/A \to H^{\pm}W^{\mp}$. In the calculation $m_H = 500\,\text{GeV}$, $m_A = 600\,\text{GeV}$, and $\sin(\beta - \alpha) = 1$ is assumed. The figure is taken from Reference [11].

### Charged Higgs boson production and decay

The production and decay channels of the charged Higgs boson are determined by its coupling to other particles. As the charged Higgs boson couples to mass, it will predominantly be produced in connection with heavy fermions and bosons. The by far heaviest fermion is the top quark. For this reason, the production and decay modes are usually discussed in two different regimes, depending on the mass of the charged Higgs boson being smaller ($m_{H^{\pm}} < m_t$ – light charged Higgs boson) or larger ($m_{H^{\pm}} > m_t$ – heavy charged Higgs boson) than the top-quark mass.

For the discussion of the production of a charged Higgs boson, the focus is being set on $pp$ collisions at the Large Hadron Collider (LHC). Figure 2.5 shows the cross-sections for different charged Higgs boson production modes. If the charged Higgs boson is light, it will be mainly produced in $gg \to t\bar{t}$ events, where one top quark decays into a charged Higgs boson and a bottom quark,

$$gg \to t\bar{t} \to tH^{\pm}b.$$

Feynman diagrams of the main $t\bar{t}$ production modes at the LHC are drawn in Figure 2.6.

In the heavy charged Higgs boson scenario the process $gb \to tH^{\pm}$ usually dominates. But if $m_{H^{\pm}} + m_{W^{\pm}}$ is below the mass of a neutral Higgs boson, the production in connection with a $W$ boson can dominate, $gg \to H^{\pm}W^{\mp}$. Exemplary Feynman diagrams of these processes are drawn in Figure 2.7. Also shown there is a possible Feynman diagram with two charged Higgs bosons in the final state.

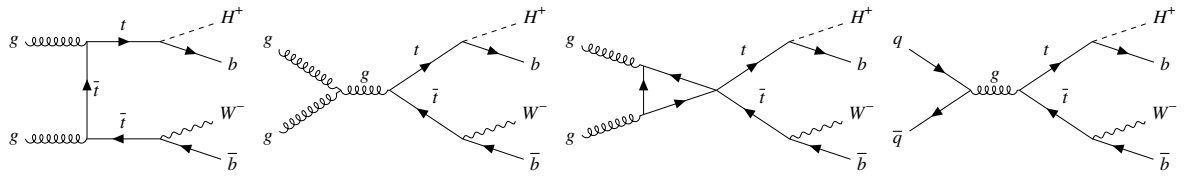



Figure 2.6: Feynman diagrams of $t\bar{t}$ production modes in $pp$ collisions with one top quark decaying into a charged Higgs boson and a bottom quark.

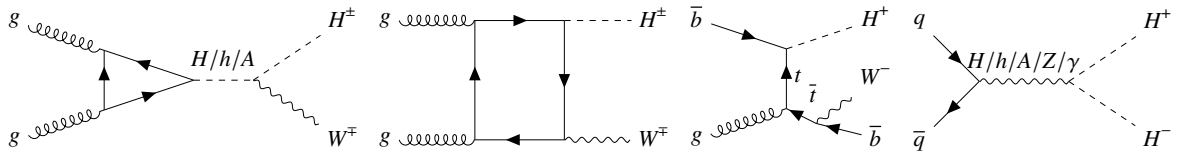



Figure 2.7: Exemplary Feynman diagrams of charged Higgs boson production modes in $pp$ collisions.

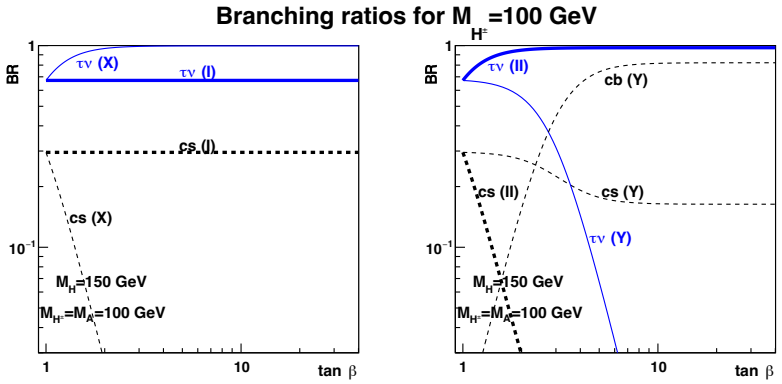



Figure 2.8: Branching ratios of a charged Higgs boson with $m_{H^\pm} = m_A = 100\,\text{GeV}$ and $m_H = 150\,\text{GeV}$ for models with NFC as defined in Table 2.1. The figure is taken from Reference [11]. Calculations were carried out with the 2HDMC [61] and HDECAY [62] software.

Regarding the decay of the charged Higgs boson, in the light-mass regime, the dominant fermionic decay modes are

$$H^\pm \to cb, \qquad H^\pm \to cs, \qquad H^\pm \to \tau\nu_\tau.$$

Those are the heaviest accessible fermion pairs, whereas the decay to $cb$ quarks is suppressed by the CKM matrix element. As the masses of these fermion pairs are roughly of the same order, the dominant decay mode is primarily determined by $\tan\beta$ and the Yukawa couplings. Figure 2.8 shows the branching ratios of a light charged Higgs boson as a function of $\tan\beta$ for models with NFC. The $cs$-channel is relevant in the 2HDM Type-I model for all $\tan\beta$ values and dominates in Type-II and Type-X models for small $\tan\beta$ values ($\tan\beta \lessapprox 1$). In the heavy charged Higgs boson scenario the dominant decay mode to fermions is always $H^\pm \to tb$.

The couplings of the charged Higgs boson to bosons are proportional to $\sin(\beta-\alpha)$ or $\cos(\beta-\alpha)$. The possible decay modes are into a $W^\pm$ boson plus a neutral gauge or Higgs boson,

$$\begin{aligned} &H^\pm \to W^\pm Z, && H^\pm \to W^\pm \gamma, \\ &H^\pm \to W^\pm H, && H^\pm \to W^\pm h, && H^\pm \to W^\pm A. \end{aligned}$$

Not all of these decay channels may be accessible, depending on the masses of the neutral and charged Higgs bosons. The decay $H^\pm \to W^\pm A$ is disfavoured in some models because the mass terms of $H^\pm$ and $A$ are very similar. $H^\pm \to W^\pm\gamma$ is small in most models. $H^\pm \to W^\pm Z$ is only relevant for a heavy charged Higgs boson, and the same applies for $H^\pm \to W^\pm H$ and $H^\pm \to W^\pm h$, if the found Higgs boson is $h$. The

latter three decay modes compete with the $tb$-channel in the heavy charged Higgs boson scenario. For the outlined reasons, all bosonic decay modes are disfavoured in most light charged Higgs boson scenarios.

As a summary, it is to be noted that the production and decay channels heavily depend on the free parameters of the models. When searching for a charged Higgs boson, it is best to cover all of them. In the light charged Higgs boson scenario, the main production mode is in top quark decays, and the relevant decay modes are $\tau\nu_\tau$, $cb$, and $cs$.

### 2.2.2 Supersymmetry

Supersymmetry (SUSY) models have been developed based on the idea of combining the Poincaré symmetry and the internal local gauge symmetry of the SM. This is possible by the introduction of fermionic symmetry generators, which relate particles of different spins, to the Poincaré group (super-Poincaré group) [63]. The fermionic generators enable transformations of bosons to fermions and vice versa (supersymmetry transformations).[9] This makes the equations for forces and matter identical.

In SUSY models, every particle needs to have a supersymmetric partner particle (sparticle) with differing spin. Since the particles of the SM cannot be their own superpartners, for every SM fermion and gauge boson a new sparticle with half a unit less spin is predicted. For the scalar sector, the partner of the Higgs field is a new weak isospin doublet with spin-1/2 and opposite hypercharge. A second Higgs doublet is required to cancel chiral anomalies, and because conjugate chiral multiplets cannot couple together in the Lagrangian, i.e. a single doublet is unable to give mass simultaneously to up-type and down-type quarks [22, 64, 65].

There are plenty of reasons why SUSY models are of interest. Among the most notable ones is the natural solution it provides to the gauge hierarchy problem, introduced in Section 2.1.3. For every loop of particles, which might give large corrections to the Higgs boson mass, there is a counter term from the corresponding sparticles. If SUSY is an unbroken symmetry, the masses of particles and sparticles are identical, and the cancellations of the higher-order terms ($\Lambda^2, \cdots$) are exact.

Similar loop corrections show up in the propagators of the forces, which make the coupling strengths of the three interactions depend on the energy scale being considered (cf. Section 2.1.2). At an energy scale of about $10^{16}$ GeV, the coupling strengths of the three interactions are almost of the same size. This led to the development of Grand Unified Theories (GUT), in which all the forces are unified into a single symmetry group, e.g. SU(5), which is broken at low energies. [66] The new sparticle loops in SUSY models will affect the running of the coupling strength and can cause all three couplings to have the exact same strength at a certain energy scale, which would facilitate GUT models.

Another important motivation for SUSY models is that among the predicted sparticles, there can be a stable, heavy and only weakly interacting particle, which is a perfect candidate for dark matter – given $R$-parity is conserved (definition see below). Lastly, in models with local supersymmetry gravity arises naturally, making it possible to unite all four fundamental interactions in one model (Supergravity) [67].

Below, a few SUSY models will be introduced to get an overview of the active research areas. In the discussed models, SUSY is always (softly) broken, resulting in the sparticle masses being orders

[9] The discussion of supersymmetric models in this section is limited to models with one supersymmetric transformation ($N = 1$) in four dimensions.

Table 2.2: The SM particles and their supersymmetric partners in the MSSM. The Table is adapted from Reference [23].

| Particle | | Spin | | Sparticle | | Spin |
|---|---|---|---|---|---|---|
| Quark | $q$ | 1/2 | | Squark | $\widetilde{q}_L, \widetilde{q}_R$ | 0 |
| Lepton | $\ell$ | 1/2 | | Slepton | $\widetilde{\ell}_L, \widetilde{\ell}_R$ | 0 |
| Neutrino | $\nu$ | 1/2 | | Sneutrino | $\widetilde{\nu}_L, \widetilde{\nu}_R(?)$ | 0 |
| Gluon | $g$ | 1 | | Gluino | $\widetilde{g}$ | 1/2 |
| Photon | $\gamma$ | 1 | $\widetilde{\gamma}$ | Neutralino | $\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$ | 1/2 |
| $Z$ boson | $Z$ | 1 | $\widetilde{Z}$ | | | |
| Higgs | $H$ | 0 | $\widetilde{H}_1, \widetilde{H}_2$ | | | |
| | | | $\widetilde{H}^{\pm}$ | Chargino | $\tilde{\chi}_1^{\pm}, \tilde{\chi}_2^{\pm}$ | 1/2 |
| $W$ boson | $W^{\pm}$ | 1 | $\widetilde{W}^{\pm}$ | | | |

of magnitudes larger than their SM counterparts. The reason for this is that so far no sparticles have been found, e.g. a particle like the electron with a mass of 511 keV but with zero spin. But in order for the SUSY models to solve the fine-tuning problem, a mass scale for sparticles of only about 1 TeV is favoured.

#### Minimal Supersymmetric Standard Model

The minimal supersymmetric Standard Model (MSSM) is the minimal – in terms of additional particles and forces – extension of the SM realising SUSY. The particle content of the MSSM is presented in Table 2.2. The MSSM already contains 178 independent parameters [26]. The electroweak scalar partners mix to the physical neutralino and chargino states. A U(1) $R$-symmetry, or $R$-parity of the form

$$R = (-)^{3(B-L)+2s},$$

is imposed to ensure baryon ($B$) and lepton number ($L$) conservation – otherwise the proton wouldn't be stable. In this equation $s$ is the spin of the particles; hence, SM particles and their superpartners have opposite $R$-parities. The $R$-parity makes the lightest supersymmetric particle (LSP), typically $\tilde{\chi}_1^0$, stable. The LSP is a candidate for dark matter.

In the MSSM, the Higgs sector is identical to the type-II 2HDM. Because of relations among the parameters of the MSSM, the MSSM Higgs sector has at tree-level only two free parameters, $\tan\beta$ and $m_A$. The mass of the charged Higgs boson is at tree level given by

$$m_{H^{\pm}}^2 = m_A^2 + m_W^2.$$

Due to the new particles, additional production and decay channels of the charged Higgs boson open up. [65, 68]

#### Other SUSY models

Another well-studied SUSY model is the next-to-minimal supersymmetric Standard Model (NMSSSM) [69]. The NMSSM is of interest as it addresses the $\mu$-problem[10] of the MSSM. The Higgs sector of the NMSSM is a type-II 2HDM with an additional singlet.

The U(1) extended MSSM (USSM) tackles the $\mu$-problem by an additional U(1) gauge symmetry [71]. USSM models predict the existence of additional (compared to the MSSM) neutral gauge bosons.

Also to mention are supersymmetries in string theory models (Superstring theory) [72, 73]. Generally, SUSY models with much more complicated Higgs sectors are possible, e.g. Higgs triplets.

### 2.2.3 Other models with an extended Higgs sector

#### Inert 2HDM

The inert 2HDM is a special 2HDM model with an unbroken $\mathbb{Z}_2$ symmetry, e.g. $\Phi_2 \rightarrow -\Phi_2$. Unbroken means that $\Phi_2$ has a zero vacuum expectation value. In this model $\Phi_2$ is called the inert doublet, as it doesn't couple to fermions. $\Phi_1$ behaves as in the SM, yielding a Higgs boson and giving masses to the electroweak gauge bosons. From $\Phi_2$ arise two charged and two neutral scalars. They can only interact with the electroweak gauge bosons and the SM Higgs boson. The lightest neutral scalar from $\Phi_2$ serves as a candidate for dark matter [74, 75]. Initially, the inert model was proposed to explain neutrino masses [76].

#### 2HDM plus singlets (N2HDMs)

The simplest model predicting a charged Higgs boson beyond the 2HDM is the N2HDM. The N2HDM has, compared to the 2HDM, an additional real or complex singlet. The additional singlet leads to one or two additional neutral scalars, respectively. The phenomenology of the charged Higgs boson is unchanged with respect to 2HDMs, except that new production and decay channels involving the new neutral scalars may be possible. N2HDMs are popular because with the imposition of a simple $\mathbb{Z}_2$-symmetry the particles from the singlet can be good candidates for dark matter. Models with more than just one additional singlet are also compatible with $\rho = 1$.

#### 2HDM plus doublets (MHDMs)

Another class of models are multi-Higgs-doublet-models (MHDMs). In these models the Higgs sector contains three (3HDM) [77] or even more Higgs doublets. Each additional doublet adds another pair of charged Higgs bosons to the particle content. In the 3HDM the phenomenology of the charged Higgs bosons is described by five parameters: The masses of the charged Higgs bosons, the two ratios of the vev's, and the mixing angle between $H_1^\pm$ and $H_2^\pm$. Due to the larger parameter space, only very weak exclusions have been set on 3HDMs and on MHDMs.

#### Models with Higgs triplets – Georgi-Machacek model

If an increased coupling of the SM Higgs boson with the electroweak gauge bosons is measured, it can only be explained by Higgs triplets or higher-order multiplets [65]. But scalar sectors with Higgs triplets or higher-order multiplets are generally constrained by the $\rho = 1$ measurements (cf. Equation 2.15). The $\rho$ issue can be circumvented by either giving the triplet field a vacuum expectation value that is much smaller compared to another multiplet, or the fields and vacuum expectation values can be arranged so that a custodial SU(2) symmetry is maintained [20, 21].

[10] The $\mu$-problem is discussed in Reference [70].

The most prominent example of the latter case is the Georgi-Machacek model [78].
In the Georgi-Machacek model the SM Higgs doublet is extended by a real and complex triplet. The resulting particles can be grouped in two singlets ($H_1^0, H_1^{0\prime}$), a triplet ($H_3^0, H_3^\pm$), and a fiveplet ($H_5^0, H_5^\pm, H_5^{\pm\pm}$). The coupling of $H_3^\pm$ to fermions is given by [21]

$$g_{H_3^\pm ff} = \frac{g_W}{m_W}\frac{v_3}{v_2} V_{\mathrm{CKM}} \left(m_{U^i} P_L - m_{D^i} P_R\right),$$

where $v_2$ and $v_3$ are the vacuum expectation values of the doublets and triplets, respectively, $m$ are the masses of the fermions and $P_L/P_R$ are the chiral projection operators. The charged Higgs bosons from the fiveplet do not couple to fermions.
Higgs triplets also show up in left–right symmetric models[11] and are able to generate small neutrino masses naturally, e.g. in the type-II seesaw mechanism [79, 80]. Common for all models with triplets are doubly charged Higgs bosons.

## 2.3 Searches for a charged Higgs boson

The following section summarises the most recent searches for a charged Higgs boson. This includes direct searches and constraints from indirect searches, e.g. precision Higgs measurements. The most stringent exclusions on models with an extended scalar sector are set by flavour physics precision measurements. These constraints are presented in Section 2.3.1.

Deviations in precision measurements, e.g. muon g-2 and $R(D^{(*)})$ measurements, can also hint towards certain BSM theories. In order to confirm an extended Higgs sector, a direct search for new particles predicted by the theory is inevitable. Direct searches for a charged Higgs boson, their excesses and constraints on BSM models are presented in Section 2.3.2.

### 2.3.1 Constraints on 2HDMs from flavour-physics precision measurements

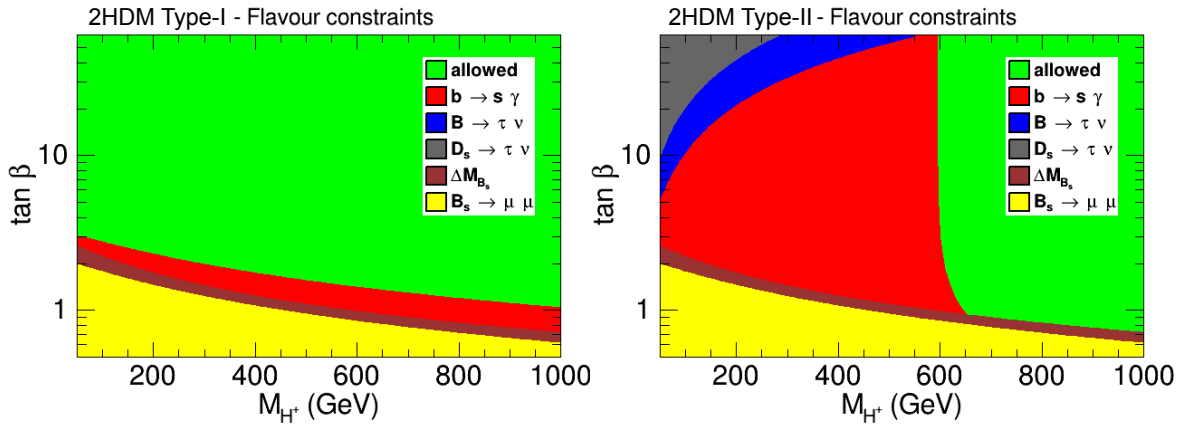
Flavour physics measurements are affected by the presence of a charged Higgs boson because in any vertex with a $W$ boson and two fermions, the $W$ boson may be replaced by a charged Higgs boson. The charged Higgs boson contribution is especially large if the fermions are heavy. Figure 2.9 summarises the most important constraints from flavour-physics measurements on $m_{H^\pm}$ and $\tan\beta$ in CP-conserving 2HDM models (cf. Table 2.1). The different contributions are:

**$b \to s\gamma$**

A $b$ quark can, at the one-loop level, decay to an $s$- or $d$-quark and a photon. A Feynman diagram of this process is depicted in Figure 2.10. In case a charged Higgs boson exists, it can, besides the SM $W$ boson, act as a mediator particle in the loop and modify the branching ratio of this process. Experimentally this is measured in the decays of $B$-mesons. In order to cancel theoretical uncertainties, the ratio

$$R_\gamma \equiv \frac{\mathcal{B}_{s\gamma} + \mathcal{B}_{d\gamma}}{\mathcal{B}_{c\ell\nu}}$$

[11] In left-right symmetric models parity is an exact symmetry and is only broken spontaneously due to the form of the scalar potential.




(a) (b)

Figure 2.9: Constraints from flavour-physics measurements on $m_{H^\pm}$ and $\tan\beta$, assuming $m_h = 125\,\text{GeV}$, in the CP-conserving (a) type-I and (b) type-II 2HDM models. The exclusions for the type-X and type-Y models are similar to type-I and type-II, respectively. The figure is adapted from Reference [81].

is measured. $R_\gamma$ compares the branching ratios ($\mathcal{B}$) of $b$-quark decays into another down-type quark and a photon to the decay into a $c$ quark, lepton and neutrino. In the 2HDM type-II and type-Y models, $R_\gamma$ is only allowed to be larger than in the SM, setting very tight limits for a large parameter space. The most precise measurements have been conducted by the Belle II collaboration at the SuperKEKB collider in Japan. Charged Higgs boson masses below 580 GeV are excluded for $\tan\beta$ values above $\sim 1$ [82–84]. In 2HDM type-I and type-X models, as the $H^\pm tb$ coupling is proportional to $\cot\beta$, the exclusion is limited to small values of $\tan\beta \lesssim 2$ [82].

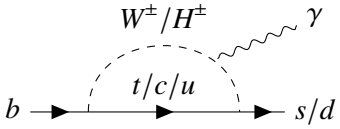



Figure 2.10: Feynman diagram of the process $b \to s\gamma$ with a $W^\pm$ or $H^\pm$ in the loop.

**$B/D_s \to \tau\nu_\tau$**

The decay of a charged $B^\pm$ or $D_s$ meson into a $\tau$ lepton and the respective neutrino proceeds in the SM through the exchange of a $W$ boson. The exchange of a charged Higgs boson, which, compared to the $W$ boson, is not helicity suppressed, modifies the branching ratio as [11]

$$R_{B_\tau} = \frac{\mathcal{B}(B \to \tau\nu_\tau)_{\text{2HDM}}}{\mathcal{B}(B \to \tau\nu_\tau)_{\text{SM}}} = \left(1 - \frac{m_B^2}{m_{H^\pm}^2}\tan\beta^2\right)^2 .$$

The measurement of $R_{B_\tau}$, which is in agreement with the SM prediction [28], adds constraints on large $\tan\beta$ values (cf. Figure 2.9).

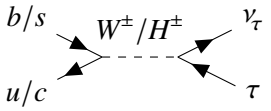



Figure 2.11: Feynman diagram of the process $B \to \tau\nu_\tau$ through a $W^\pm$ or $H^\pm$ exchange.

**$\Delta m_{B_s^0}$**

Neutral meson states can oscillate at second order, e.g. via a diagram as drawn in Figure 2.12. The two flavour eigenstates mix to two physical mass eigenstates, a lighter and a heavier one. In BSM theories the oscillations can also occur via the exchange of charged Higgs bosons. This has implications on the mass difference between the two mass eigenstates. Measurements of the mass difference exclude small values of $\tan\beta$ for all CP-conserving 2HDMs (cf. Figure 2.9). [85–87]

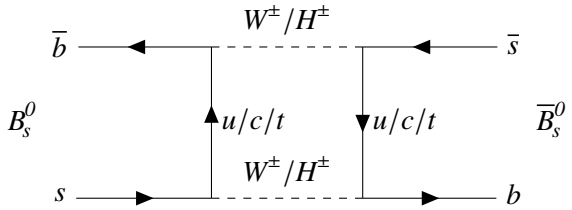



Figure 2.12: A possible Feynman diagram of the oscillation from $B_s^0$ to $\overline{B}_s^0$.

**$B_s^0 \to \mu^-\mu^+$**

$B_s^0$-meson decays to a $\mu^-\mu^+$ pair are highly suppressed because the tree-level process would involve flavour-changing neutral currents, which are forbidden in the SM. Higher-order diagrams involving $W$ bosons, e.g. box diagrams (see Figure 2.12), are allowed. For 2HDM models without NFC, the branching ratio of this process can be drastically increased. But also in models with NFC, a charged Higgs boson can be exchanged instead of the $W$ boson in the higher-order diagrams and enhance the branching ratio. Measurements of the branching ratio mainly exclude small values of $\tan\beta$ (cf. Figure 2.9) [11, 88].

These exclusions are always calculated for certain models and cannot reject all models with a second Higgs doublet. In models with a more complex Higgs sector, the discussed exclusions are much weaker due to the increased complexity of the models, e.g. in MHDMs [11].

### 2.3.2 Direct searches for a charged Higgs boson

Direct searches for a charged Higgs boson have first been performed at the Fermi National Accelerator Laboratory (FermiLab) by the CDF and D0 collaborations starting in the 1990s. At the Tevatron collider at FermiLab, protons and anti-protons were collided at centre-of-mass energies of up to 1.96 TeV. Charged Higgs bosons have been searched for in decays of top quarks, but no excess has been found [104–107].

Direct searches for a charged Higgs boson were also performed at the Large Electron-Positron Collider (LEP) at CERN. At LEP, electrons and positrons were collided with centre-of-mass energies of up to $\sqrt{s}_{\text{LEP}} = 209\,\text{GeV}$. Charged Higgs bosons with masses below $\sqrt{s}_{\text{LEP}}/2$ could have been produced in pairs. The measurements from all experiments at LEP combined could exclude charged Higgs bosons with masses below 72.5 GeV and 80 GeV in the 2HDM type-I/X and type-II/Y models, respectively [108].

Table 2.3: Summary of recent searches for a charged Higgs boson at the LHC. $\sqrt{s}$ is the centre-of-mass energy of the $pp$ collisions, $\mathcal{L}$ is the integrated luminosity. The table is divided into searches for a light (top) and heavy (bottom) charged Higgs boson. An exception is the $H^\pm \to \tau\nu_\tau$ channel, which is being searched for in the light and heavy charged Higgs boson mass regime. 95 % CL upper limits on the branching ratio $t \to H^\pm b$ are shown for decays of a light charged Higgs boson to SM particles. The upper limits from the $H^\pm \to \tau\nu_\tau$ channel are for charged Higgs boson masses up to 160 GeV.

| Process | Experiment Reference | $\sqrt{s}$ [TeV] | $\mathcal{L}$ [fb$^{-1}$] | Search range $m_{H^\pm}$ [GeV] | Limits on $\mathcal{B}(t \to H^\pm b)$ [%] Expected | Observed |
|---|---|---|---|---|---|---|
| $H^\pm \to cs$ | ATLAS [89] | 7 | 4.7 | 90 – 150 | 1.5 – 8 | 1.2 – 5.1 |
| | CMS [90] | 13 | 35.9 | 80 – 160 | 0.29 – 2.39 | 0.25 – 1.68 |
| $H^\pm \to cb$ | ATLAS [91] | 13 | 139 | 60 – 160 | 0.09 – 0.25 | 0.15 – 0.42 |
| | CMS [92] | 8 | 19.7 | 90 – 150 | 0.53 – 0.83 | 0.33 – 1.3 |
| $H^\pm \to \tau\nu_\tau$ | ATLAS [93] | 13 | 140 | 80 – 3 000 | 0.02 – 0.28 | 0.02 – 0.27 |
| | CMS [94] | 13 | 36.1 | 80 – 3 000 | 0.12 – 0.35 | 0.08 – 0.36 |
| $H^\pm \to WA$ | ATLAS [95] | 13 | 139 | 100 – 160 | – | – |
| | CMS [96] | 13 | 35.9 | 100 – 160 | – | – |
| $H^\pm \to tb$ | ATLAS [97] | 13 | 139 | 200 – 2 000 | – | – |
| | CMS [98, 99] | 13 | 35.9 | 200 – 3 000 | – | – |
| $H^\pm \to Wh$ | ATLAS [100] | 13 | 140 | 250 – 3 000 | – | – |
| $H^\pm \to WH$ | CMS [101] | 13 | 138 | 300 – 700 | – | – |
| $H^\pm \to WZ$ | ATLAS [102] | 13 | 140 | 200 – 3 000 | – | – |
| | CMS [103] | 13 | 137 | 200 – 3 000 | – | – |

The most recent direct searches for a charged Higgs boson have been performed in $pp$ collisions at the LHC at CERN. The measurements are summarised in Table 2.3. In the search for a light charged Higgs boson, the channels $H^\pm \to cs, cb, \tau\nu_\tau$, which have been proposed in Section 2.2.1, are covered by the ATLAS and CMS collaborations:

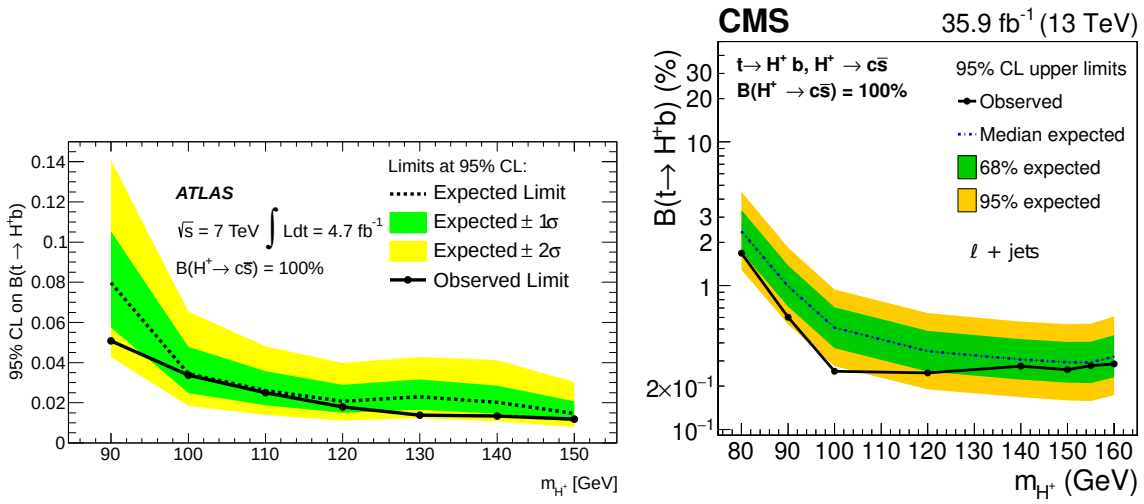
**$H^\pm \to cs$**

The previous search for $H^\pm \to cs$ decays performed by the ATLAS collaboration used a partial LHC Run-1 dataset of $pp$ collisions at $\sqrt{s} = 7$ TeV [89]. A kinematic fitter [107, 109] was used to reconstruct the full $t\bar{t}$ system, from which $W/H^\pm$ dijet candidates are extracted. No significant excess compatible with a signal from a charged Higgs boson with masses between 90 to 150 GeV was found in the dijet mass distribution. Upper limits on the branching ratio $\mathcal{B}(t \to H^\pm b)$, assuming $\mathcal{B}(H^\pm \to cs) = 1.0$, were set at 95 % confidence level (CL) between 1.2 to 5.1 %.

The most recent search by the CMS collaboration for $H^\pm \to cs$ decays was performed on a partial LHC Run-2 dataset ($\sqrt{s} = 13$ TeV) and extended the mass range from 80 to 160 GeV. The analysis strategy is similar to the one from the ATLAS search presented above. The main difference is the usage of charm tagging to identify signal events. No significant excess has been found by CMS either, and limits on $\mathcal{B}(t \to H^\pm b)$ were lowered between 0.25 to 1.68 %. The observed and expected limits with uncertainty bands for both analyses are shown in Figure 2.13.

**$H^\pm \to cb$**

The search for $H^\pm \to cb$ by the ATLAS collaboration was performed using the full LHC Run-2




(a) (b)

Figure 2.13: Expected and observed 95 % CL upper limits on $\mathcal{B}(t \to H^{\pm}b)$, assuming $\mathcal{B}(H^{\pm} \to cs) = 1.0$ for the searches performed by the (a) ATLAS and (b) CMS collaborations. The $\pm 1\sigma$ and $\pm 2\sigma$ variations around the expected upper limit are indicated by the green and yellow bands, respectively.

dataset for a charged Higgs boson mass range from 60 to 160 GeV [91]. The 95 % CL upper limits on the branching ratio $\mathcal{B}(t \to H^{\pm}b)$, assuming $\mathcal{B}(H^{\pm} \to cb) = 1.0$, are derived to be between 0.15 and 0.42 %. These limits are tighter than the ones obtained by the $H^{\pm} \to cs$ searches for two main reasons. On the one hand, $b$-quarks can be identified more reliably compared to $s$-quarks, due to their unique signature in the detector. On the other hand, the main background, SM $t\bar{t}$ events, can be substantially reduced with the help of the $b$-quark identification because $\mathcal{B}(W \to cb) \approx 0.17$ %, whereas $\mathcal{B}(W \to cs) \approx 50$ %.

It is important to note that a local excess with $3\sigma$ significance at a charged Higgs boson mass of 130 GeV was seen, as shown in Figure 2.14. The excess has a global significance of $2.5\sigma$ and the width of the excess is compatible with what is expected from signal injection studies. Since then, it was of great interest whether the excess would also be observed by the ATLAS Run-2 $H^{\pm} \to cs$ (this!) analysis. This would be a strong evidence for the existence of a charged Higgs boson with mass around 130 GeV.

**$H^{\pm} \to \tau\nu_{\tau}$**

The searches for $H^{\pm} \to \tau\nu_{\tau}$ decays have set stringent upper limits on $\mathcal{B}(t \to H^{\pm}b) \times \mathcal{B}(H^{\pm} \to \tau\nu_{\tau})$. This is the case because the momentum and direction of leptons are measured with higher precision than those of jets. The observed limits vary between 0.02 and 0.27 % for charged Higgs boson masses of 80 to 160 GeV [94, 110]. The search for $H^{\pm} \to \tau\nu_{\tau}$ is extended to masses above the top-quark mass. In quarkphobic 2HDM models, the $\tau\nu_{\tau}$ decay channel can dominate also for heavy charged Higgs bosons.

**Searches for heavy charged Higgs bosons**

In the heavy charged Higgs boson scenario, the $H^{\pm} \to tb$ decay always dominates over the $cs$ and $cb$ channels. The search for $H^{\pm} \to tb$ has been performed by the ATLAS and CMS collaborations

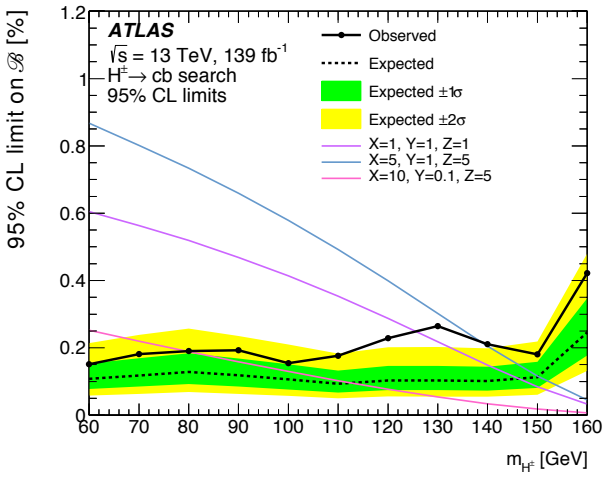
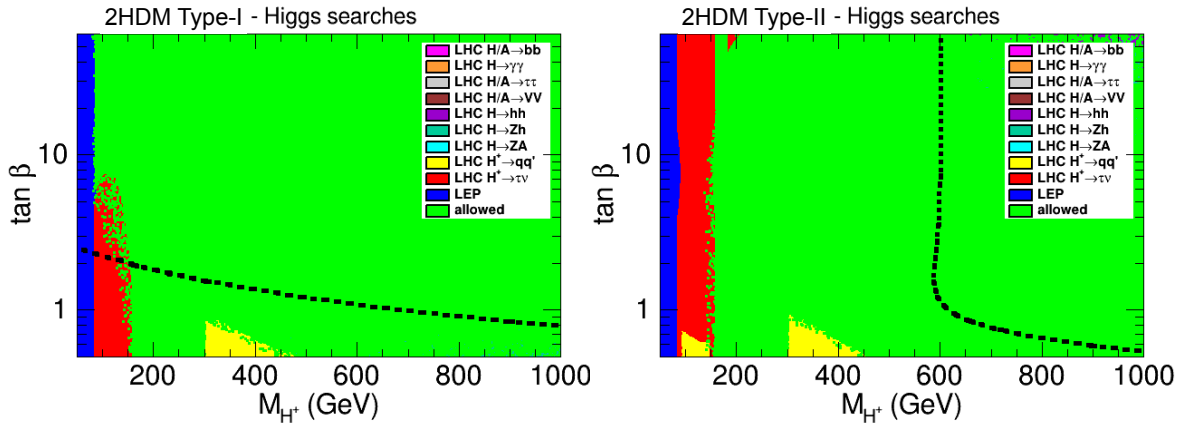



Figure 2.14: Expected and observed 95 % CL upper limits on $\mathcal{B}(t \to H^{\pm}b)$, assuming $\mathcal{B}(H^{\pm} \to cb) = 1.0$ for the searches performed by the ATLAS collaboration. The $\pm 1\sigma$ and $\pm 2\sigma$ variations around the expected upper limit are indicated by the green and yellow bands, respectively.

in a mass range from 200 to 2 000 GeV [97] and 200 to 3 000 GeV [98, 99], respectively. Other searches for a heavy charged Higgs boson are in decays to a $W$ boson and a neutral Higgs boson, $h$ or $H$ [100, 101]. The decay channel $H^{\pm} \to WZ$ opens up in models with Higgs triplets and has been searched for by the ATLAS [102] and CMS [103] collaborations. No significant excess has been found so far.

The constraints on $m_{H^{\pm}}$ and $\tan\beta$ from direct searches in CP-conserving 2HDM models are displayed in Figure 2.15. The direct measurements can exclude new parameter spaces in the 2HDM type-I and type-X models, especially for a light charged Higgs boson. Unfortunately, the plots are not up to date with the description given above. The measurements for $H^{\pm} \to \tau\nu_{\tau}$ and $H^{\pm} \to qq'$ included in the plots are early LHC Run-2 (or even Run-1) measurements performed on a smaller integrated luminosity. The full Run-2 measurements can extend the excluded parameter space further.




(a) (b)

Figure 2.15: Constraints from direct measurements on $m_{H^{\pm}}$ and $\tan\beta$, assuming $m_h = 125\,\text{GeV}$, in the CP-conserving (a) type-I and (b) type-II 2HDM models. The exclusions for the type-X and type-Y models are similar to type-I and type-II, respectively. The region below/left of the dashed line is excluded by flavour physics measurements (cf. Figure 2.15). The figure is adapted from Reference [81].

CHAPTER 3

# The Large Hadron Collider and the ATLAS detector

In this thesis data from proton-proton collisions recorded by the ATLAS detector is being analysed. The ATLAS detector is situated around one of the four interaction points of the Large Hadron Collider (LHC) and records the collisions, acting like a giant, fast and radiation-hard camera. The purpose is reconstructing the emerging particles and thereby reconstructing the physics at very high energies and small distances. The following chapter describes the physics of proton-proton collisions in Section 3.1, how protons are being accelerated and brought to collision at the LHC in Section 3.2, and reports on the structure of the ATLAS detector in Section 3.3. Lastly, Section 3.4 gives details on the reconstruction of particles with the ATLAS detector.

## 3.1 Proton-proton collisions

The proton is the lightest baryon, consisting of three valence quarks *uud*. At smaller length scales, a sea of gluons and quark-antiquark pairs inside the proton becomes visible, as shown schematically in Figure 3.1. As a consequence of Heisenberg's uncertainty principle, the actual observed structure depends on the energy scale $Q$ being probed. At higher energy scales, finer structures become visible.

In high-energy proton-proton collisions, the partons inside the protons interact with each other. The

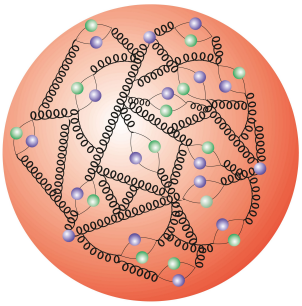


Figure 3.1: Schematic structure of the proton taken from Reference [111]. The blue and green dots represent quarks and antiquarks, respectively. The curly lines represent gluons.

interaction with the highest momentum transfer is referred to as the hard scattering process. To predict the inclusive cross-section of the hard scattering process $\sigma_{ij\to X}$, all possible Feynman diagrams for this process have to be considered, which is a difficult and resource-intensive task. Meaningful predictions are retrieved by using perturbation theory in which only the Feynman diagrams up to a certain order in $\alpha_s$ are considered. Perturbation theory yields good approximations as the coupling constant of the strong interaction is $\ll 1$ for energy scales above $m_Z$.

In order to calculate the inclusive cross-section $pp \to X$, the structure of the proton has to be considered. This is encoded in the parton distribution functions (PDFs). A PDF yields the probability density for finding a certain parton inside the proton with a momentum fraction $x$. This probability depends heavily on the considered energy scale, $\mathrm{PDF}(x, Q^2)$. Predicting the PDFs is challenging as it involves the low-energy, non-perturbative regime. Therefore, the PDFs have been measured for certain energy scales in deep inelastic scattering experiments. The DGLAP evolution equations [112–114] are used to estimate the PDFs at different energy scales.

The inclusive $pp \to X$ cross-section can then be calculated as

$$\sigma_{pp\to X} = \sum_i \sum_j \int \mathrm{d}x_i \, \mathrm{d}x_j \mathrm{PDF}(x_i, \mu_F^2) \times \mathrm{PDF}(x_j, \mu_F^2) \times \hat{\sigma}^{(n)}_{ij\to X}(x_i, x_j, \mu_R),$$
$$\text{with} \quad \hat{\sigma}^{(n)} = \hat{\sigma}^{(0)} + \alpha_s(\mu_R^2)\hat{\sigma}^{(1)} + \cdots + \alpha_s^n \hat{\sigma}^{(n)} + O(\alpha_s^{n+1}). \quad (3.1)$$

The indices $i$ and $j$ run over the various parton types. $\mu_R$ is the renormalisation scale of the strong coupling, already introduced in Section 2.1.2, and defines the upper scale up to which the calculations are valid. $\mu_F$ is the factorisation scale and can be thought of as the scale that separates the high- and low-energy physics. [115]

Figure 3.2 shows a schematic of a $pp \to t\bar{t}$ collision. Block 1, the hard process, has been discussed in the previous paragraph. But, as can be seen in the picture, a real $pp$ collision involves much more than just the hard scattering process. In addition to the hard scattering process, other partons inside the protons can interact with each other. These interactions are referred to as multiparton interactions (MPI). The interacting particles can also radiate off other particles, mainly gluons, before the hard scattering process. This type of radiation is termed Initial-State Radiation (ISR). The same applies to particles produced in the hard scattering process. Accordingly, this radiation is called Final-State Radiation (FSR). MPI, ISR, FSR and the proton remnants that didn't interact with each other built the underlying event (UE).

Moving on in time, the number of particles constantly increases and the average energy per particle decreases until only stable particles are left. In this context, "stable" means that the average flight path $\tau c$, with $\tau$ being the lifetime of the particle, is much larger than the detector radius. Block 2 is about the evolution of the particles produced in the hard-scattering interactions. On the one hand, the produced particles, in this example the top quarks, may be short-lived and decay into other particles, which may decay themselves. On the other hand, the particles radiate partons, leading to a shower of particles (PS).

Block 3 is concerned with the fragmentation and hadronisation of final-state partons. Fragmentation is the process by which an energetic final-state gluon or quark radiates a cascade of softer partons. Due to confinement the final-state partons, except the very short-lived top quark, will be bound in hadrons. The hadrons may be short-lived and decay into other hadrons before reaching the detector. The bundle of arising hadrons is boosted in the direction of the initial parton and forms a so-called *jet*. Hence, final-state partons can not be detected individually but only through hadrons in a jet.

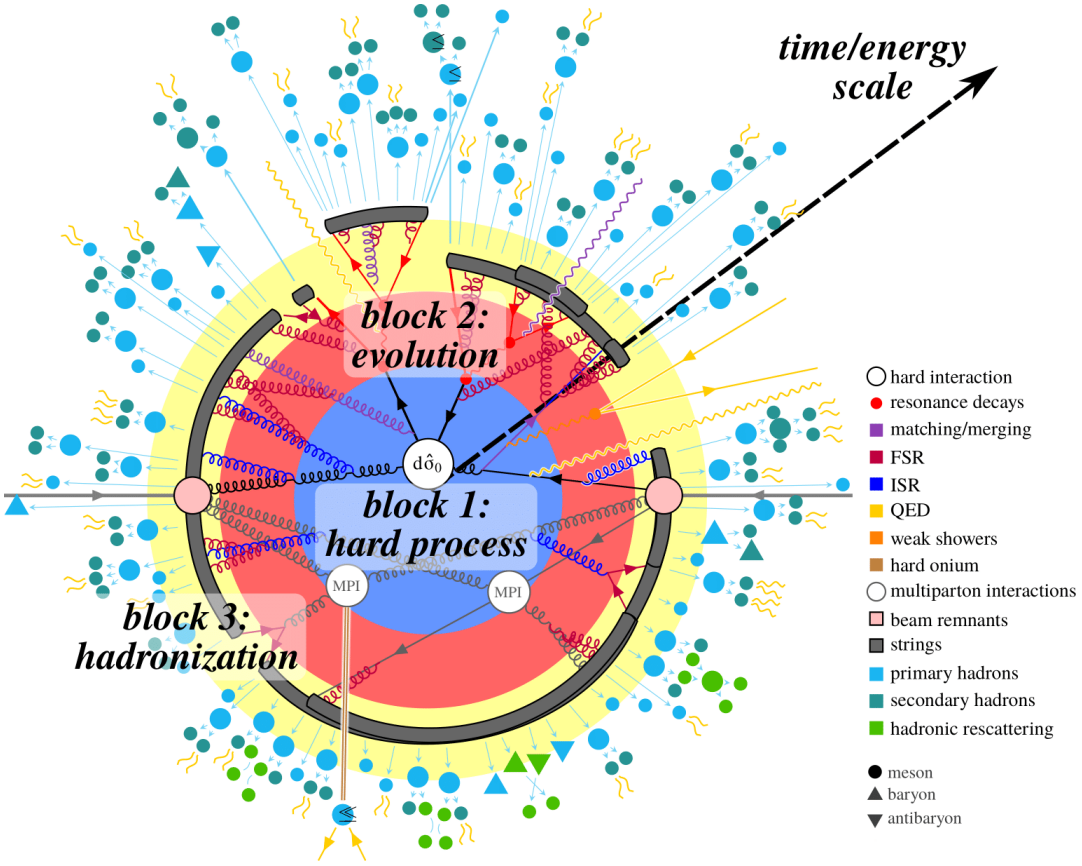



Figure 3.2: Schematic of a $pp$ collision in which a top-quark pair is produced. The figure is taken from Reference [116, 117].

### 3.1.1 Simulation of proton-proton collisions

Accurate simulation of the physics in particle collisions is a key ingredient for nearly any collider experiment. The simulations are important to understand what a physics process looks like in the detector and to develop reconstruction algorithms. Additionally, simulations are important to calibrate the detector response to retrieve the "true" particle four-momentum, calibrate selection efficiencies, etc.

The simulations use repeated random sampling algorithms (Monte Carlo – MC) to simulate single event processes. The simulation consists of multiple steps [118]:

1. First, the simulation of the hard process by **calculating the matrix element** (ME), including ISR and FSR. The calculations are performed up to a certain order in perturbation theory: leading-order (LO), next-to-leading-order (NLO), next-to-next-to-leading-order (NNLO), etc. The outcome is a list of particles produced in the hard scattering process and their four-momenta.

2. The **evolution**, mainly taking care of the decay of unstable particles, comes second.

3. The third step is the **fragmentation and hadronisation**, which are a not well-understood processes. Perturbative calculations are not possible, as these are long-range, low-energy ($\alpha_s \approx 1$) processes.

Event generators use different models like the string [119, 120] or cluster [121] fragmentation models for this task.

4. Next is the **detector simulation**. A model of the ATLAS – or generally any – detector and its magnetic field configuration is developed in GEANT4 [122]. GEANT4 simulates the passage of particles through the detector, simulating processes like bremsstrahlung, Compton scattering, etc. The result is energy deposits in the detector components.
   The detector simulation is usually the computationally most expensive part. Therefore, a fast simulation has been developed in which the calorimeter response is parameterised instead of fully simulating the electromagnetic and hadronic showers [123].

5. In a last step, these energy depositions have to be converted into **digitized** signals. After this step, the information is identical to a real measurement. Any further reconstruction on real data can be applied one-to-one to the simulation. But in simulation the initial physics process and emerging particles are known.

The first three steps combined are known as the event generation. A variety of event generators exist that use different physics models. These models have parameters that are tunable for the given task. Generators may even be specialised in a certain task. Therefore, in a full event generation, it is quite common to interface multiple event generators, each taking care of a different step.

## 3.2 Accelerator physics at the Large Hadron Collider

The LHC is a circular collider in which protons or lead atoms are brought to collision at four interaction points, around which four detectors are situated: ALICE, LHCb, CMS and ATLAS. ALICE and LHCb are designed specifically for heavy-ion collisions and $b$-quark physics, respectively. CMS and ATLAS are general-purpose detectors.

The goal of the LHC campaign was to reach the highest centre-of-mass energies and luminosities to tackle the open questions in particle physics. It was constructed in the same tunnel as the Large Electron-Positron collider (LEP) with a circumference of 26.7 km. The tunnel is between 45 to 170 m underground to reduce the background from cosmic radiation. In LEP, electrons and positrons were collided at centre-of-mass energies up to 209 GeV in order to perform electroweak precision measurements. At the LHC, protons can be accelerated to much higher energies as they lose less energy by synchrotron radiation. The LHC was designed to accelerate the beam protons to 7 TeV, resulting in a centre-of-mass energy of $\sqrt{s} = 14\,\text{TeV}$.[1]

In order to observe interesting physics processes with small cross-sections, it is desirable to have as many well recordable $pp$ collisions as possible in short time scales. The collision rate at an interaction point of an accelerator is quantified by the instantaneous luminosity and can be parametrised for a circular collider with beams that have no spatial offset with respect to each other as

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}}}{4\pi\sigma_x\sigma_y} S.$$

[1] The highest centre-of-mass energy reached so far is 13.6 TeV.

$N_b$ are the number of protons per colliding bunch, $n_b$ are the number of bunches rotating in each direction, $f_{\text{rev}}$ is the revolution frequency of the bunches, and $\sigma_{x,y}$ are the bunch size in the vertical and horizontal plane. $S$ is the geometric reduction factor which accounts for the fact that the bunches are collided under a certain crossing angle and depends on the crossing angle and the bunch length $\sigma_z$. The expected event rate of a certain process with cross-section $\sigma$ can be calculated with the instantaneous luminosity as

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \mathcal{L}\sigma.$$

The luminosity contains all information related to the experiment, and the cross-section contains the information related to the physics process. The total number of expected events can be calculated with the integrated luminosity, which is defined as the instantaneous luminosity integrated over time.

### 3.2.1 Accelerator physics

This section briefly introduces the physics concepts utilised to build a circular particle collider. More details about accelerator physics and the instruments used in modern particle accelerators can be found in References [124, 125].

In particle accelerators, electric fields are used to control the longitudinal motion, and magnetic fields control the transverse motion. The physics governing this is contained in the Lorentz force law.

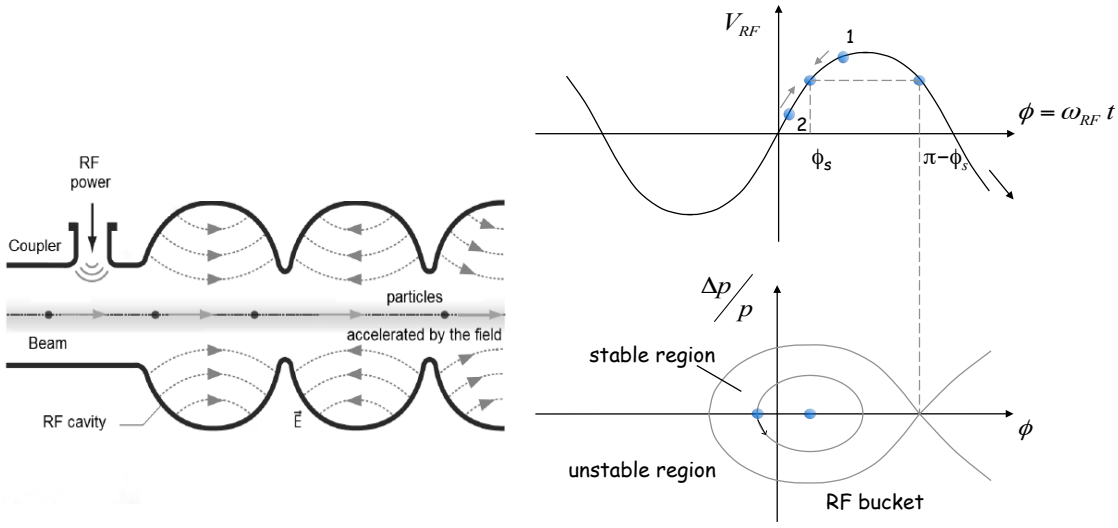
Charged particles are accelerated using radio-frequency (RF) cavities. Cylindrical conductors are charged alternately positive and negative. When choosing the correct frequency, the charged particles passing the conductors will be consistently pushed by the conductors behind them and pulled by the conductors ahead of them. A sketch of this process is given in Figure 3.3(a).

The RF cavities define buckets, confined by the arrival time of the charged particles at the RF cavities, or alternatively their phase, and their momentum. Figure 3.3(b) shows the envelope of an RF bucket in the phase-momentum space. Only particles inside the bucket are consistently accelerated. A particle exactly synchronised with the RF frequency is in the centre of the RF bucket and is called a synchronous particle. Particles with a phase ahead (behind) of the synchronous particle observe a different electric field and undergo a lower (higher) acceleration. In circular colliders particles repeatedly pass through RF cavities. The higher momentum particles will have a longer orbit and a lower revolution frequency, which will delay their phase at the accelerating cavity, and vice versa for lower momentum particles. These two phenomena cause all particles in the bucket to perform a stable longitudinal oscillation, as indicated by the lines with arrows in Figure 3.3(b).

Dipole magnets are used to bend the trajectory of charged particles along a circular path. For the energies considered at the LHC, magnetic fields allow for smaller bending radii compared to electric fields. Relativistic charged particles being accelerated perpendicular to their direction of motion emit synchrotron radiation. The energy loss of the charged particle is proportional to their mass $\sim m^{-4}$.

Quadrupole magnets are used to focus the beams. Like a lens, particles passing the quadrupole magnets are focused in one direction but defocused in the other direction. Quadrupole magnets focusing the beams in the horizontal and vertical planes are placed alternately, resulting in a net focus in both directions.

Other higher-order magnet configurations are also used for focusing, e.g. sextupole magnets correct for chromatic aberrations. In order to ensure a stable operation, other instruments (which will not be




(a) (b)

Figure 3.3: (a) Sketch of particle acceleration in RF cavities. The figure is adapted from Reference [126]. (b) RF voltage as a function of the particle phase $\phi$ and the related phase-space picture, where $p$ is the momentum and $\Delta p$ is the momentum difference with respect to the synchronous particle. The envelope is the RF bucket. The figure is adapted from Reference [127].

introduced in detail here), like beam cleaning collimators, beam position and profile monitors, etc. are needed, too.

### 3.2.2 The Large Hadron Collider

The LHC is part of the CERN accelerator complex, situated near Geneva. Protons are produced from a hydrogen-ion source and pass through multiple accelerators before being injected into the LHC with an energy of 450 GeV. A sketch of the accelerator complex as of 2022[2] is given in Figure 3.4. The path of the protons to the LHC is marked by grey arrows.

Both LHC beam pipes get filled with 2808 proton bunches in the standard LHC fill mode, each containing $1.15 \times 10^{11}$ protons. High vacua are inside the beam pipes to minimise beam losses and background collisions at the interaction points. It takes about 20 min to accelerate the protons to their final collision energy of 6.5 TeV.

The dipole magnets are of importance as the field strength is among the main factors limiting the beam energies. The dipole magnets are superconducting and are made from niobium–titanium (NbTi). An improvement in the heat load of the NbTi fibres allows for cooling to temperatures below 2 K using superfluid helium, which facilitates magnetic field strengths of 8 T. Twin-bore dipole magnets (two-in-one design) in which the windings for the two beam pipes are in a common cryostat are used due to the limited space in the tunnel.

[2] In this thesis data from the years 2015 to 2018 is being analysed. The main difference is that back then LINAC2 was still in operation.

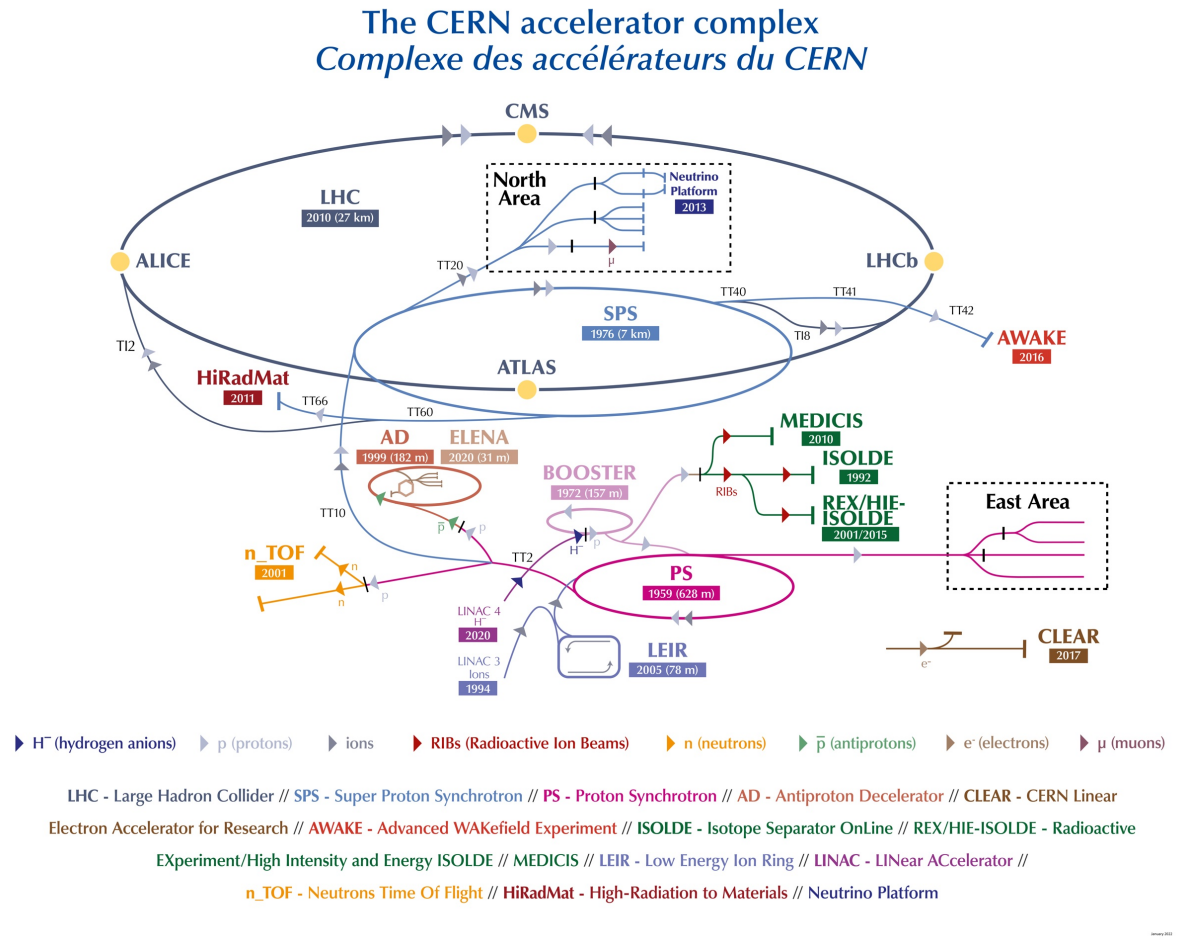



Figure 3.4: The CERN accelerator complex as of January 2022. [128]

The beams are collimated at the interaction points to a transverse width of about 16 µm in both directions to reach high luminosities. A so-called $\beta^*$ levelling procedure is adopted at the ATLAS interaction point in which the transverse beam width is adjusted to keep the luminosity constant for a long period of one LHC fill cycle.

The large instantaneous luminosity delivered by the LHC leads to the effect of *pile-up* in the ATLAS detector. Multiple interactions of the same bunch-crossing overlap in the detector (in-time pile-up). Additionally, the signal readout, processing and dead time of many detector components are larger than the time between consecutive bunch crossings of 25 ns (out-of-time pile-up). The effect of pile-up on the measurements can be diminished by reconstructing the hard scattering interaction vertex.

## 3.3 The ATLAS detector

The ATLAS detector [129] is a general-purpose detector situated around one of the interaction points of the LHC. A schematic of its structure is visualised in Figure 3.5. The detector has the shape of

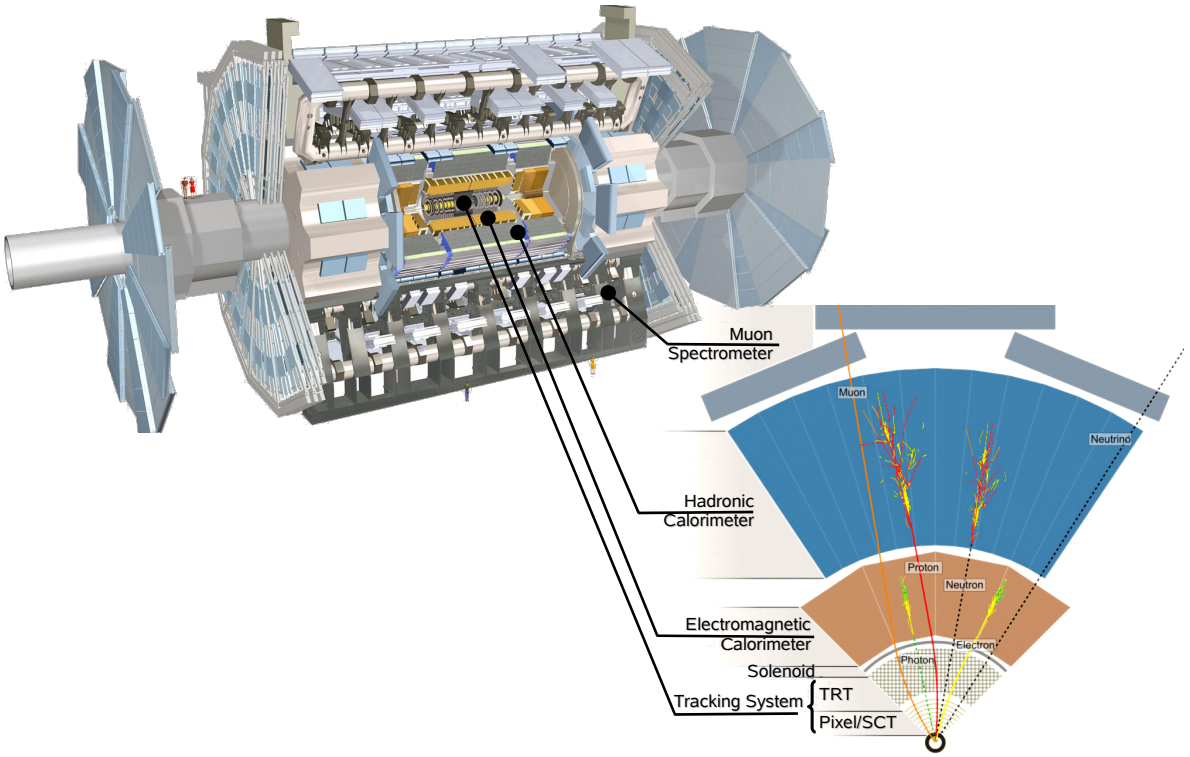



Figure 3.5: Schematic of the ATLAS detector. Displayed on the right-hand side is the cross-section of the central part in the transverse plane and the signatures of different types of particles are shown. [130]

a cylinder and is forward-backward symmetric with respect to the interaction point. It consists of a barrel part and endcaps on each side and covers almost the full solid angle. The beam pipe coincides with the symmetry axis of the cylinder. The detector components are built in layers around the beam pipe. Innermost is the inner tracking detector, which is surrounded by a superconducting solenoid, followed by electromagnetic and hadronic calorimeters. Outermost is the muon spectrometer in which three superconducting air-core toroidal magnets are incorporated. The detector systems as of LHC Run-2 (2015 – 2018) are described in more detail in the following sections. If not cited otherwise, the descriptions are based on Reference [129].

ATLAS uses a cylindrical, right-handed coordinate system. The origin of the coordinate system is defined by the nominal interaction point. The $x$- and $y$-axes are defined to point towards the centre of the LHC and upwards, respectively. The beam pipe defines the $z$-axis. The azimuthal angle $\phi$ is the angle around the $z$-axis and is measured with respect to the $x$-axis. The polar angle $\theta$ is the angle with respect to the $z$-axis. Because the centre-of-mass frame of the hard scattering process is standardly boosted along the $z$-axis (cf. Section 3.1), the Lorentz invariant quantities in $z$-direction, rapidity $y$ and an approximation of it, the pseudorapidity $\eta$, are commonly used to describe the angle of objects with respect to the $z$-axis. The flux of emerging particles from the $pp$ collisions is roughly uniform in $y$ or $\eta$.

They are defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \qquad \eta = -\ln \tan \frac{\theta}{2}.$$

The pseudorapidity is a special case of the rapidity where the mass of the particle is neglected with respect to its momentum. The angular distance between objects is measured in units of

$$\Delta R = \sqrt{(\Delta y)^2 + (\Delta \phi)^2}.$$

### 3.3.1 Inner tracking detector

The main objective of the inner detector is to provide a precise reconstruction of the trajectory of charged particles as well as the reconstruction of the primary interaction vertex and possible secondary vertices from heavy-flavour[3] hadron decays. The inner detector covers the pseudorapidity range $|\eta| < 2.5$. It is immersed in a 2 T magnetic field that is aligned with the beam pipe. The magnetic field is produced by a thin solenoid magnet which surrounds the inner detector. The trajectory of charged particles is bent in the transverse plane by the magnetic field. From the bending radius the particle's momentum and charge can be inferred.

The configuration of the inner detector is shown in Figure 3.6. It is composed of three parts (from the inside out): (i) pixel, (ii) silicon microstrip (SCT), and (iii) transition radiation tracker (TRT). [129, 131]

#### Pixel detector

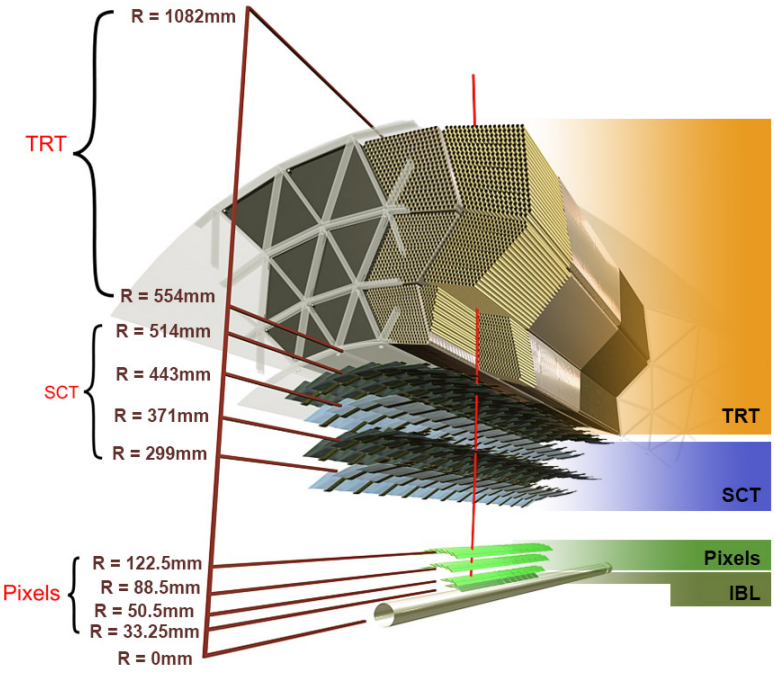
The pixel detector is located closest to the beam pipe. It consists of pixel modules, each containing arrays of semiconductor diode pixels with size $50 \times 400\,\mu\text{m}$. The semiconductor diodes are depleted by applying a reverse bias voltage. A high-energy charged particle traversing the depleted diode region ionises the material and creates electron-hole pairs. The electric field inside the diode makes the electrons (holes) drift towards the anode (kathode), resulting in a current that is readout, amplified and further processed. A signal in a pixel is commonly referred to as a *hit*.

The pixel modules are arranged in four concentric layers around the beam pipe in the barrel region, and as three disc layers perpendicular to the beam pipe in the endcaps. The layer closest to the beam pipe is the insertable B-Layer (IBL) with a distance of only 3.35 mm to its centre. It has been installed during the Long Shutdown 1 (between Run-1 and Run-2) and has pixels with size $50 \times 250\,\mu\text{m}$. A particle with $|\eta| < 2.5$ traverses the IBL and three other layers of the pixel detector, as visualised in Figure 3.6. The intrinsic accuracies in the barrel (endcaps) are $10\,\mu\text{m}$ in $R$–$\phi$ and $115\,\mu\text{m}$ in $z$ ($R$). The pixel detector is located closest to the interaction point to deal with the large particle densities and for precise vertex reconstruction.

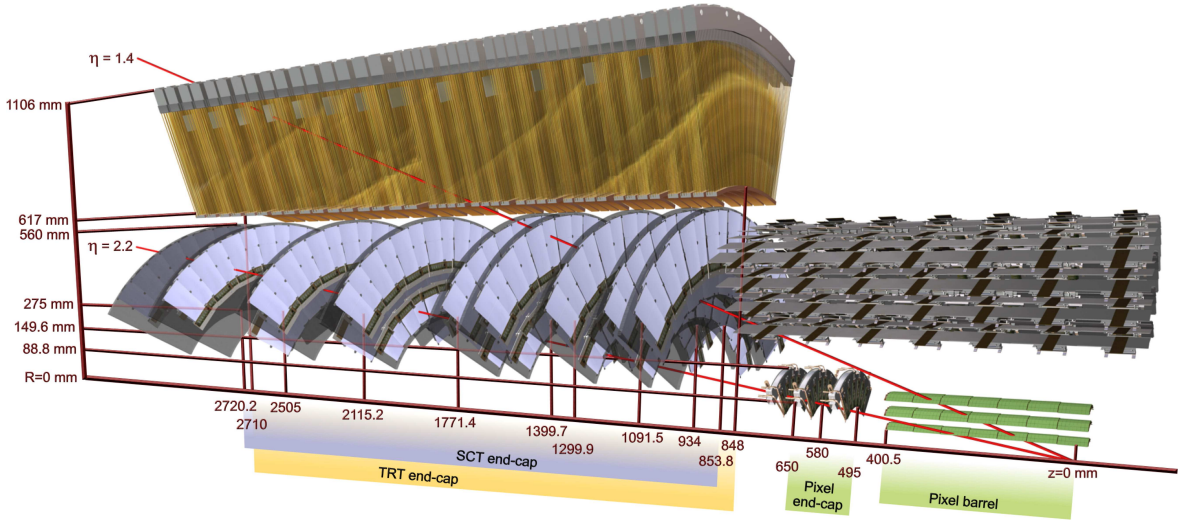
#### SCT

The semiconductor tracker consists of silicon strips with a strip pitch of $80\,\mu\text{m}$. The SCT modules in the barrel are arranged in four detector layers concentric around the beam pipe at radii between 299 to 514 mm. Each endcap contains nine discs perpendicular to the beam pipe. A layer/disc

[3] $b$- and $c$-quarks are commonly categorised as heavy-flavour quarks. $u$, $d$, and $s$ are categorised as light-flavour quarks. The categorisation is motivated by their masses and lifetimes. Most hadrons containing heavy-flavour quarks have a relatively long lifetime. The decays of these hadrons are visible as additional *secondary* vertices close to the beam pipe.




(a)




(b)

Figure 3.6: Schematic of the ATLAS inner detector. The red lines represent 10 GeV trajectories of charged particles with (a) $\eta = 0.3$ [barrel region], and (b) $\eta = 1.4$ and $\eta = 2.2$ [endcap region]. [129, 131]

has two layers of strips orientated at an angle of 40 mrad to each other in order to achieve a two-dimensional measurement in each detector layer. The SCT is designed such that each particle with $|\eta| < 2.5$ traverses four detector layers. The intrinsic accuracies in the barrel (discs) are 17 µm in $R$–$\phi$ and 580 µm in $z$ ($R$).

**TRT**

The TRT is a gaseous detector that spans from $R = 554$ to 1 082 mm in the barrel region. It consists of approximately 300 000 straw tubes with a diameter of 4 mm and a thin gold-plated tungsten wire in its centre. A voltage is applied between the tube wall and the wire. High-energy particles traversing the straws ionise the gas. The freed electrons get accelerated towards the anode wires and ionise atoms on their way to the wire, creating an avalanche of electrons. The electric signal is eventually read out at the end of the wires.
The straws are aligned parallel (perpendicular) to the beam pipe in the barrel (endcaps). Particles with $|\eta| < 2.0$ typically leave 36 hits in the TRT. The straw tubes provide position measurement only in $R$–$\phi$ with an intrinsic accuracy of 130 µm. The TRT is a cost-efficient solution to provide many hits in a large detector volume for a precise track momentum reconstruction.

Fibres and foils are interleaved between the straw tubes. High-energy particles traversing them emit transition radiation. The transition radiation photons are absorbed in the straw tubes, yielding a signal that is larger than the signal of the ionising particle itself. The transition radiation intensity depends on the energy and mass of the traversing particle. Therefore, the signal strength in the straw tubes is used to identify particles with differing masses, mainly separating electrons/positrons from charged hadrons.

Pattern recognition algorithms are utilised to reconstruct the trajectory of charged particle *tracks* from hits in the inner detector (see Section 3.4.1). The inner detector is designed to reconstruct tracks with absolute values of transverse momenta ${p_\mathrm{T}}^4$ down to 0.5 GeV. The track $p_\mathrm{T}$ resolution $\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 0.05\,\% p_\mathrm{T} \oplus 1\,\%$ decreases with $p_\mathrm{T}$. Tracks with high momenta move almost on a straight line through the detector. This makes a precise momentum determination from the track curvature challenging.

### 3.3.2 Calorimeter system

The main objective of the calorimeter system is to measure the energy and direction of electrons/positrons, photons and hadrons. The calorimeter system surrounds the inner detector, extends to $R = 4.25$ m, and covers the pseudorapidity range $|\eta| < 4.9$. It is located radially behind the inner detector tracking system, as the calorimeters perform a destructive measurement of the aforementioned particles.

An overview of the calorimeter system is given in Figure 3.7. The calorimeter system can be divided into two parts (from the inside out): the (i) electromagnetic and (ii) hadronic calorimeters.

**Electromagnetic Calorimeter**

The electromagnetic calorimeter (EMC) can be divided into a barrel part and two endcap components, covering in total the pseudorapidity range $|\eta| < 3.2$. The EMC is a sampling calorimeter with an accordion shape in the radial direction and consists of liquid argon as the

[4]The absolute value of the transverse momentum is defined as $p_\mathrm{T} = |\vec{p}| \sin\theta = \sqrt{p_x^2 + p_y^2}$.

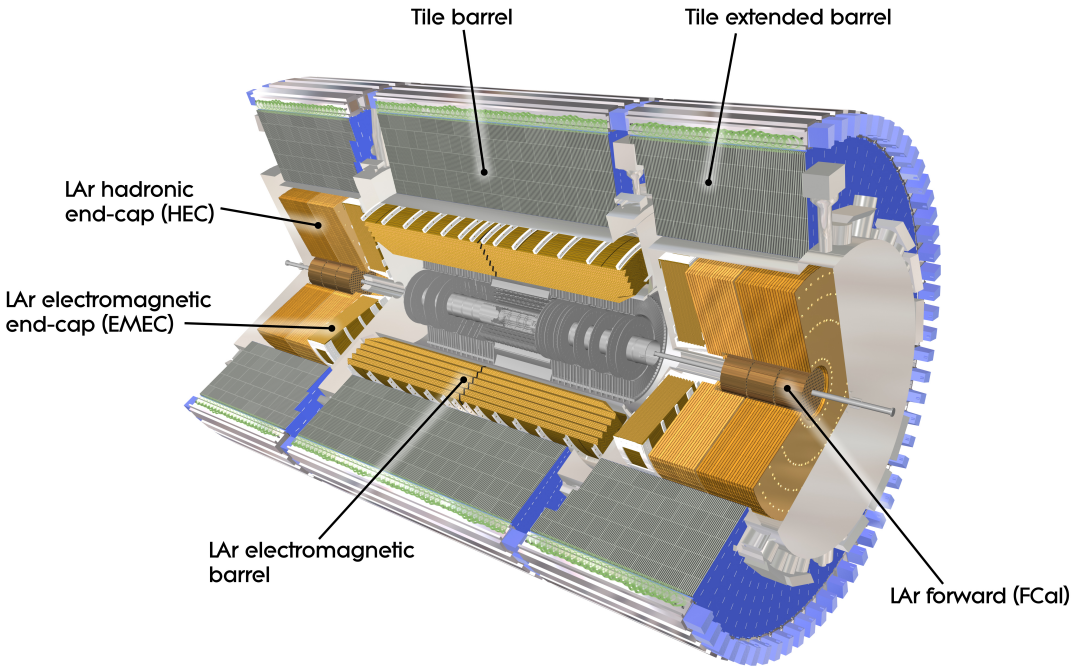



Figure 3.7: Schematic of the ATLAS calorimeter system. [129]

active material and lead plates as absorbers in between.
High-energy electrons and photons penetrating the EMC start an electromagnetic shower, most likely inside the lead plates, because it is a material with a large atomic number and high density. In an electromagnetic shower, electrons/positrons radiate bremsstrahlung, and photons undergo pair production of an electron-positron pair. The particles repeatedly undergo these processes, resulting in many particles with much lower energy than the incoming mother particle. At some point the energy of the particles is so small that photons undergo Compton scattering and the photoelectric effect, and electrons/positrons scatter elastically with nuclei or electrons of the traversing material, ionising the material. The energy of the incoming particle is eventually transformed into a charge signal that gets collected and is proportional to the energy of the incoming particle. Muons and charged hadrons radiate only little bremsstrahlung due to their higher mass and usually don't start an electromagnetic shower.
The EMC is segmented in the $\eta$–$\phi$ plane to determine the direction of the incoming particles. It is also segmented radially into three (two) segments for $|\eta| < 2.5$ ($|\eta| > 2.5$) to get information on the longitudinal shower shape, which is used to differentiate the types of incoming particles. The granularity in $\eta$–$\phi$ varies for the different segments. The first layer has a fine granularity in $\eta$ to separate $\pi^0 \rightarrow \gamma\gamma$ decays from prompt $\gamma$ events. The thickness of the EMC depends on $\eta$ but is never lower than 22 radiation lengths, ensuring that the whole electromagnetic shower will be contained within the EM. The *presampler* is located between the inner detector and the EMC in the range $|\eta| < 1.8$. It consists of an active liquid argon layer and is used to detect interactions of particles with the inner detector to correct for the energy lost upstream of the calorimeter.

#### Hadronic Calorimeter

The hadronic calorimeter is grouped in the tile calorimeter ($|\eta| < 1.7$), the hadronic endcap calorimeter ($1.5 < |\eta| < 3.2$) and the forward calorimeter ($3.1 < |\eta| < 4.9$). The tile calorimeter is a sampling calorimeter with steel plates as the absorber and scintillating tiles as the active material.

Hadrons interact via the strong force with the nuclei of the detector, produce secondary particles and may excite the nuclei. The secondary particles may decay rapidly and produce electromagnetic showers ($\pi^0, \eta$), ionise the detector material or interact strongly again, creating even more new particles. These processes repeat until all energy of the incoming particle is stored in the detector. Some energy, which is absorbed in the target recoil and by breaking nuclear bindings, is not detectable. The energy is stored in the scintillators either directly as photons or as excited electrons that will emit photons when de-exciting. The photons are collected at the edges of the scintillating tiles using wavelength-shifting fibres and photomultiplier tubes. The light intensity is proportional to the energy of the incoming particle.

These hadronic showers have a much larger longitudinal spread than electromagnetic showers and typically already start in the EMC. The EMC and hadronic calorimeters have a combined thickness for all $\eta$ values of at least 9.7 hadronic interaction lengths. This ensures that the hadronic showers are fully contained within the calorimeters and that particles don't punch through to the muon spectrometer. The tile calorimeter has three radial layers, and the granularity in $\eta$–$\phi$ is coarser than in the EMC.

The hadronic endcap calorimeter and forward calorimeter use, as does the EMC, liquid argon as the active material and copper or tungsten as the absorber material. The forward calorimeter has no EMC in front. The forward calorimeter has three radial layers, where the first layer is optimised for measuring electromagnetic showers and the other two layers mainly measure hadronic showers.

The energy resolution of the calorimeters at ATLAS can be parametrised as $\sigma_E/E = a/\sqrt{E} \oplus b$. The resolution improves with increasing energy as statistical fluctuations in the shower developments reduce with increasing shower sizes. The design values for the coefficient $(a, b)$ are $(10\,\%, 0.7\,\%)$ for the EMC, $(50\,\%, 3\,\%)$ for the hadronic tile and endcap calorimeters, and $(100\,\%, 10\,\%)$ for the forward calorimeter.

### 3.3.3 Muon spectrometer

The outermost detector system is the muon spectrometer (MS). Only muons and solely weakly interacting particles (neutrinos) are supposed to reach the MS. An overview of the ATLAS MS is given in Figure 3.8. The MS is immersed in a magnetic field that is produced by large air-core toroids, one in the barrel part and one in each endcap. The objectives of the MS are to measure the trajectory of muons, determine precisely their transverse momentum, and provide fast signals for trigger purposes. To detect the trajectory of muons, different types of drift chambers are utilised:

#### Monitored Drift Tubes

Monitored Drift Tubes (MDTs) are used for a precise measurement in the bending (longitudinal) direction. The working principle of the MDTs is similar to the straw tubes in the inner detector. In the barrel region the layers are concentric around the beam pipe at distances between $R = 5$ to 10 m; in the endcaps the MDTs are arranged perpendicular to the beam pipe in wheels at distances between $|z| = 7.4$ to 21.5 m. The layer dimensions and chamber sizes increase proportionally to

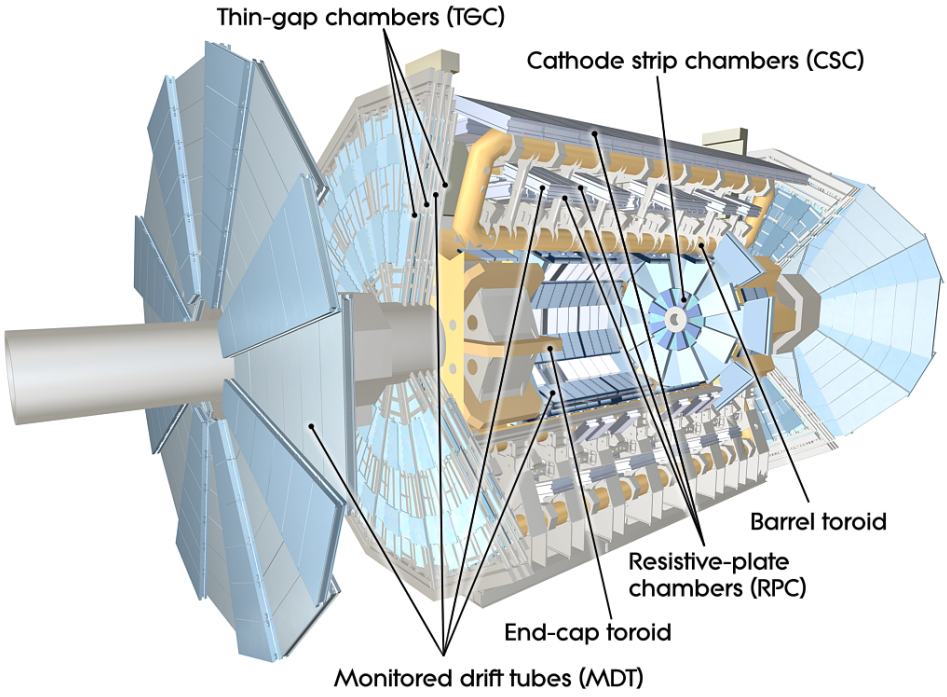



Figure 3.8: Schematic of the ATLAS muon spectrometer. [129]

their distance from the nominal interaction point. Muons with $|\eta| < 2.7$ nominally pass three MDT layers, yielding 20 measurements per track in the bending direction with a precision of 35 µm per layer in $R/z$–$\phi$.

### Cathode Strip Chambers

At the innermost endcap layer between $2.0 < |\eta| < 2.7$ the MDTs are replaced by Cathode Strip Chambers (CSCs) to deal with the increased particle rate. Compared to the MDTs, CSCs have a higher rate capability and timing resolution. CSCs are multiwire proportional chambers where arrays of anodes are crossed orthogonally with cathode stripes. Incident ionising radiation produces free electrons that drift towards the anodes and positively charged ions that drift towards the cathodes. This setup enables the measurement of both coordinates of the track with a precision of 40 µm in the bending plane and 5 mm in the transverse plane.

### Resistive Plate and Thin Gap Chambers

Resistive Plate Chambers (RPCs) in the barrel region ($|\eta| < 1.05$) and Thin Gap Chambers (TGCs) in the endcaps ($1.05 < |\eta| < 2.4$) are installed to provide fast trigger signals. The signal and processing time of RPCs and TGCs (15 to 25 ns) is smaller than the average bunch crossing time at the LHC, allowing for bunch-crossing identification. The transverse coordinate from the RPC and TGC measurements complements the coordinate from the precision MDT measurement to form a space-point.

RPCs consist of two parallel plates with an electric field applied and gas filled in between. RPCs produce fast signals compared to wire chambers, as the primary ionisation electrons don't have to

drift to a region of amplification first. A particle in the given $\eta$-range nominally passes through three RPC chambers, resulting in six measurements per track. Both coordinates can be determined with a precision of 10 mm.

TGCs are multiwire proportional chambers where the distance between the cathode and the wire is smaller than the distance between wires. Small distances between the wires and a large electric field facilitate fast signals. Nine TGC layers, two at the innermost MDT endcap layer and seven at the middle MDT layers, are passed by each particle in the given $\eta$-range. The resolution of the coordinates varies between 2 and 7 mm as the wire granularity changes with $\eta$.

The muon system was designed to achieve a standalone relative track $p_{\mathrm{T}}$ resolution of 10 % for a 1 TeV muon.

### 3.3.4 Luminosity measurement

It is crucial for any analysis to know how many particles collided, i.e. the integrated luminosity of the recorded dataset. The LUCID-2 detector [132] is specifically designed to monitor the instantaneous luminosity and measure the integrated luminosity. It is a Cherenkov detector situated in each forward region of the ATLAS detector ($z = \pm 17$ m). LUCID-2 consists of 16 photomultiplier tubes in each direction installed concentric around the beam pipe and pointing towards the interaction point. The Cherenkov light is produced in quartz windows in front of the photomultipliers. The number of detected particles is proportional to the number of interactions in a bunch crossing. [133]

The total number of protons in each colliding bunch is inferred from measurements of the beam currents. The absolute luminosity scale is calibrated by a van der Meer scan [134, 135]. In a van der Meer scan, beams are displaced horizontally and vertically to measure the combined size of the colliding bunches. The BCM detector consists of diamond sensors at $z = \pm 1.84$ m and is used to check the consistency during some van der Meer scan periods. The LUCID-2 measurements are complemented by bunch-by-bunch measurements from offline measurements of the track multiplicity in the inner detector and bunch-integrated measurements from the calorimeter system [133].

### 3.3.5 Trigger and data acquisition system

It is neither feasible nor desired to record all $pp$ collisions taking place at ATLAS. Physics analyses are generally interested in a certain final state. And these final states usually have a small cross-section compared to the total inelastic $pp$ cross-section. This search analyses final states with a top-quark pair. About one top-quark pair event is expected to arise in $1 \times 10^8$ $pp$ collisions at 13 TeV. To specifically select events interesting for physics analysis, ATLAS uses a trigger system that decides in real-time whether to record or dismiss an event. The following description is based on References [136–138].

An overview of the trigger and data acquisition system in ATLAS is given in Figure 3.9. A two-stage trigger system is used in ATLAS. The Level-1 (L1) trigger system is hardware-based. Only information from the calorimeter system (L1Calo) and the dedicated trigger chambers from the MS (L1Muon) are processed at this stage.

The L1Calo trigger system digitises and calibrates the data. Electrons, photons, $\tau$-lepton and jet candidates above dedicated energy thresholds are identified using calorimeter information but with a coarse granularity. In addition, the total transverse energy and missing transverse energy are calculated.

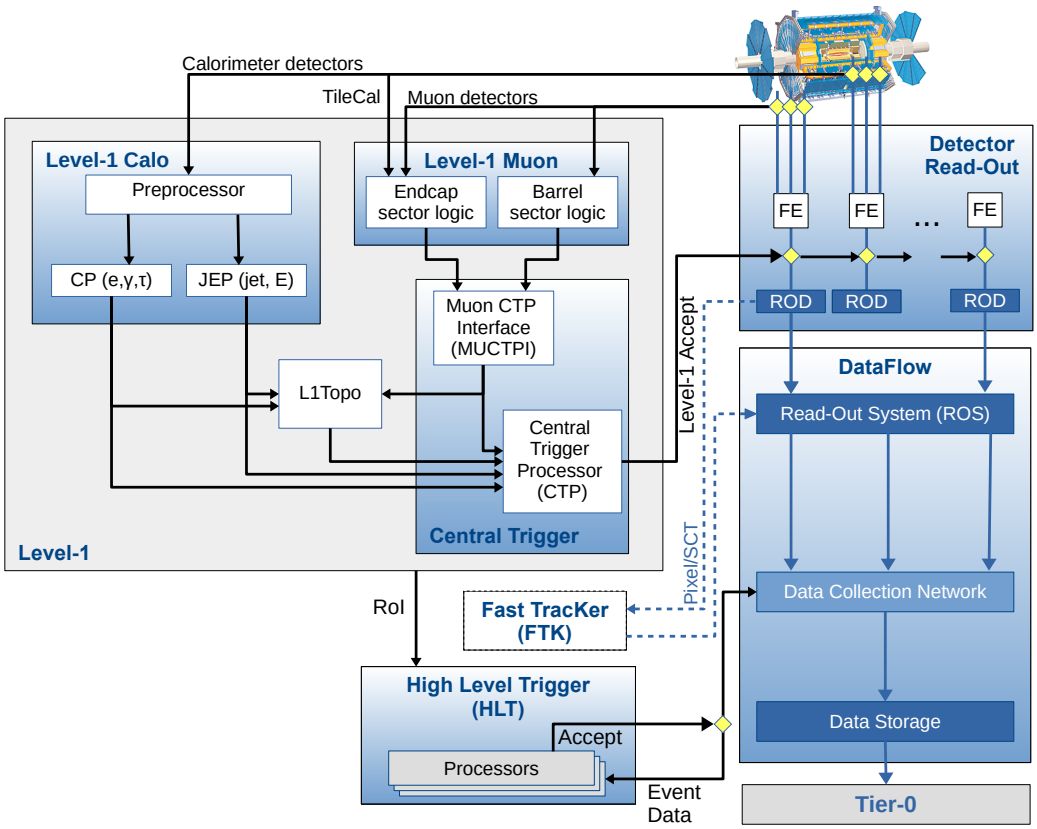



Figure 3.9: Flowchart of the ATLAS Run-2 trigger and data acquisition system. [136]

The L1Muon trigger system identifies muon candidates and estimates their $p_{\mathrm{T}}$ by calculating the deviation in the hit pattern from an infinite momentum muon. Lastly, combinations of candidates are built to calculate their angular distance and invariant mass. Based on all this information an event is either accepted or rejected at the L1 trigger. To avoid overlapping readout windows due to detector latency and to prevent overflowing buffers, so-called *dead time* may be applied, in which events are vetoed. The L1 trigger passing rate at Run-2 was about 90 kHz and is capped at 100 kHz with a latency of 2.5 µs. In comparison, the LHC bunch crossing rate is 40 MHz.

For all events accepted at L1, the whole detector data is read out. The data is analysed by the second trigger stage, the High-Level Trigger (HLT). The HLT system is software-based and works with data from Regions-of-Interest (RoIs). RoIs are regions in $\eta$ and $\phi$ around objects identified by the L1 trigger. In Run-2 the HLT consisted of around 1 500 reconstruction sequences aiming at identifying different signatures. A reconstruction sequence typically has multiple steps to provide early rejection before executing more complex and time-consuming algorithms. The reconstruction algorithms at the HLT are as close as possible to the offline reconstruction algorithms to maximise the overlap between trigger and offline event selection. If an event passes a sequence successfully, it will be stored permanently. The rate of events passing the HLT at Run-2 was about 1.2 kHz.

## 3.4 Object reconstruction

This section introduces the methods used in this analysis to reconstruct physical objects from the ATLAS detector data.

### 3.4.1 Inner detector track and vertex reconstruction

Reconstructing the trajectory of charged particles is a pattern recognition task in which tracks are reconstructed from hits in the ID. The following description, if not stated otherwise, is based on Reference [139].

The track reconstruction can be divided into five steps (inside-out tracks):

1. Formation of three-dimensional *space-points* from hits in the ID. In the SCT, hits from two angled sensors have to be combined to retrieve a space-point. Given the dense track environment at ATLAS, hits from multiple particles can occasionally be reconstructed in one cluster, leading to a *merged cluster* and only one scape-point. Merged clusters are identified using an artificial neural network utilising the measured charges and incident angles from track candidates. The space-points in the different silicon detector layers are visualised by the red points in Figure 3.10(a).

2. Formation of track seeds from sets of three space-points.

3. If at least one additional space-point is compatible with the estimated particle trajectory, a combinatorial Kalman filter [140] is used to build track candidates by including the space-points from the remaining layers. Requirements on the track candidate $p_{\mathrm{T}}$, $\eta$, impact parameters with respect to the beam spot, and number of hits in the pixel and SCT detectors, number of holes and shared space-points are enforced.[5]

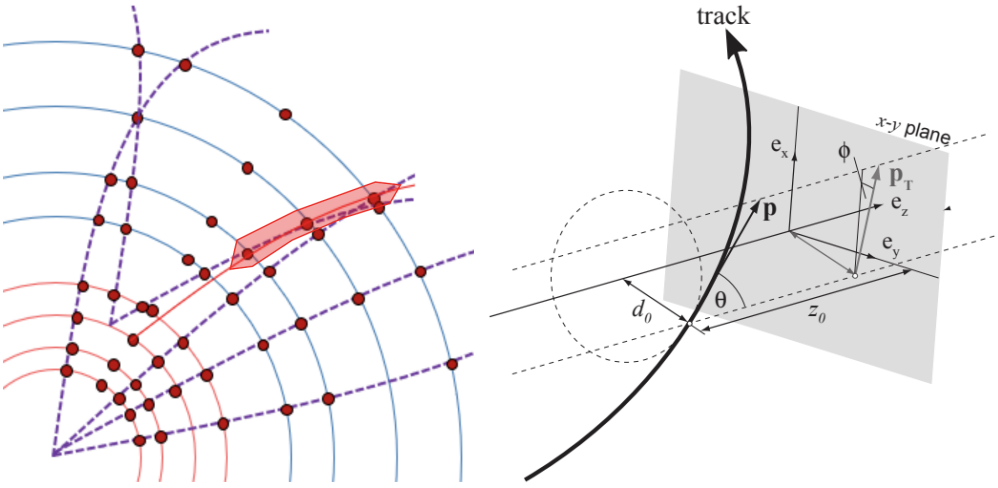
4. Resolving ambiguities in which track candidates share space-points. An example of this is given by the red area in Figure 3.10(a), where the red line represents a second track candidate. This step prioritises track candidates of higher quality, assessed using, among other factors, the track-fit $\chi^2$ and the number of holes.

5. The final step is a high-resolution track fit of all remaining track candidates. In these fits the position and uncertainty of clusters are re-estimated by additional neural networks [141] and information from the TRT is included. The increased track length in the TRT significantly improves the momentum resolution.

A track can be parametrised by five parameters which are defined with respect to the closest approach of the track to the beam-spot position in the transverse plane (perigee): The ratio of charge over momentum $q/p$, the angles $\theta$ and $\phi$, and the transverse ($d_0$) and longitudinal ($z_0$) impact parameters. The parameters are visualised in Figure 3.10(b).

In this analysis only final tracks with $p_{\mathrm{T}} > 500\,\mathrm{MeV}$ and $|\eta| < 2.5$ are considered. The tracking efficiency for charged pions with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ in the central region is on average about 90 % [142], but depends on many factors: the particle's $p_{\mathrm{T}}$ and $\eta$, the surrounding particle density, and impact parameters.

[5] A hole is an absent hit/cluster in a sensor where, considering the fitted trajectory, a cluster is expected. A shared space-point is a space-point associated with multiple tracks but which has not been identified as a merged cluster.




(a) (b)

Figure 3.10: (a) Track finding from space-points (red dots) in the inner silicon detectors. The solid lines represent the layers of the pixel and SCT detector, respectively (transverse plane). The dashed lines represent track candidates. The red area/track represents an ambiguity of track candidates.
(b) Definition of track parameters. [143]

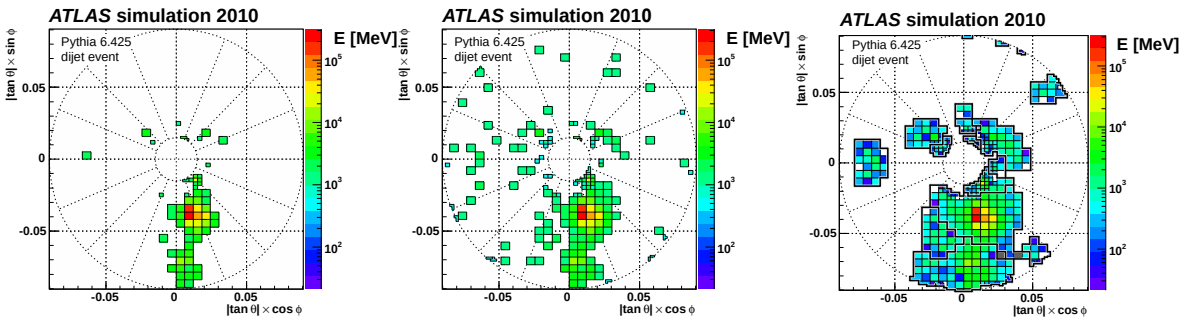
The reconstruction of the hard scattering vertex is important for rejecting pile-up from additional *pp* collisions and for the calculation of track impact parameters with respect to the hard scattering position. The vertex reconstruction description is based on References [144, 145].

Primary vertices are reconstructed with the help of tracks. First, a seed position is defined based on the perigee of tracks. The vertex position is then determined in an iterative procedure by a $\chi^2$ fit using the seed position and the tracks as inputs. After each iteration, tracks with low compatibility are being down-weighted and the vertex position is recomputed. The output of the final fit is a three-dimensional vertex position. Tracks with low compatibility with the vertex are removed from the vertex. A new vertex is determined with the remaining tracks in the event. The procedure is repeated until no tracks are left or no additional vertex can be found. Each vertex is required to have at least two associated tracks.

The hard scatter primary vertex (PV) of the event is defined as the vertex with the highest scalar sum of the $p_{\mathrm{T}}^2$ of associated tracks. A PV position resolution of 20 µm and 30 µm in the transverse and longitudinal directions is achieved, respectively [144].

### 3.4.2 Reconstruction of energy deposits in the calorimeters

Electrons, photons and hadrons penetrating the calorimeters will produce particle showers. The showers typically extend laterally and longitudinally over multiple adjacent cells. In order to reconstruct the total energy deposited by the initial particle, the cells are combined into *topological clusters* [146]. Topological clusters are built around calorimeter cells with significance $\geq 4$, excluding cells in the presampler and first EMC layer. The significance of a cell is defined as the ratio of the cell energy at the EM scale over the expected cell noise from electronic noise and pile-up. The cluster is extended by iteratively adding all neighbouring cells with significance $\geq 2$. In a last step, all neighbouring cells are added regardless of their energy. Clusters containing multiple cells with significance $\geq 4$ may be split up.




(a) (b) (c)

Figure 3.11: Stages of the formation of topological clusters in the first layer of the forward calorimeter for a simulated dijet event. In (a) and (b) only cells with significances greater than 4 and 2, respectively, are shown. In (c) the full topological clusters are shown. [146]

This procedure is shown schematically in Figure 3.11

The energy of topological clusters is calibrated to match the energy of the incoming particles. The calibration accounts for signal losses in inactive detector material, losses due to clustering and the non-compensating calorimeter response. Non-compensating means that the signal in the calorimeters differs for electromagnetic and hadronic showers. The shower type is being identified by the shape and location of the shower in the calorimeters.

### 3.4.3 Muon spectrometer track reconstruction

Standalone MS tracks are reconstructed from hits in the muon spectrometer. The task is similar to the track reconstruction in the inner detector, though the detector occupancy is much reduced. In a first step straight line segments in the bending plane from hits in each MDT layer are searched for by means of a Hough transform [147, 148]. Track candidates are built by combining segments from different layers, where the trigger detectors provide the orthogonal coordinates [147]. The final tracks are determined by repeated global $\chi^2$ fits of all track candidates taking into account the exact magnetic field geometry [147]. Outlier hits are removed, and matching hits are added to the tracks [147]. Ambiguities are resolved by removing tracks in favour of higher-quality tracks. The final fit constrains tracks to agree loosely with the nominal interaction point and accounts for energy loss in the calorimeters [147, 149].

### 3.4.4 Reconstruction of electrons and muons

The goal is to reconstruct and identify *prompt* electrons and muons[6]. In this context prompt means that the leptons are produced directly as a product of the hard-scattering collision. Prompt leptons may emerge from the decay of weak gauge bosons. Given the short lifetime of weak gauge bosons, they decay before travelling macroscopic distances, and the leptons arise immediately at the primary vertex.

*Non-prompt* leptons can appear in the decay of heavy-flavour hadrons, which typically live long enough to travel a macroscopic distance before decaying. Non-prompt electrons can additionally arise

[6] In this section the names electron and muon are used as representatives for the lepton type. When referring to electrons, the same applies to positrons, and similarly for muons and anti-muons.

in the conversion of photons when passing through detector material. A large fraction of non-prompt muons stem from the decay of light hadrons, primarily charged pions.

Besides non-prompt leptons, there are other signatures that can mimic a lepton in the detector. Another particle being identified as a lepton is a so-called *fake lepton*. In the case of electrons, those are mainly light charged hadrons. Fake muons can arise if hadronic showers punch through the calorimeters to the MS.

### Electron identification

Electrons are reconstructed from energy deposits in the EMC that are matched to tracks in the ID. Energy deposits from electrons in the calorimeters are reconstructed by *EM-topo clusters*. EM-topo clusters are topological clusters that only consider the part in the EMC (and at least 50 % of the total cluster energy has to be stored in the EMC) [150]. Tracks that are geometrically matched to – also considering track seeds in a region-of-interest around – an EM-topo cluster are refitted, allowing for a larger energy loss by bremsstrahlung [151].

All EM-topo clusters in a small area around an energetic seed cluster are combined in a *supercluster* to reunite clusters that have been split up and clusters from secondary EM showers [150]. More distant clusters that still have the same best-matched track as the seed cluster are expected to originate from bremsstrahlung and are added to the supercluster, too [150]. The supercluster, together with its geometrically best-matched track, forms the electron candidate.

Prompt electrons are identified using a likelihood discriminant, binned in the transverse energy $E_{\mathrm{T}}$[7] and $\eta$. The likelihood discriminant is constructed from quantities related to the electron track and the shower shape in the EMC [150, 152]. Track quantities like the hits in the silicon detectors and impact parameters are useful in rejecting photon conversions. In addition, a likelihood based on the transition radiation in the TRT is included to reject tracks from charged hadrons. Quantities describing the lateral and longitudinal shower shape are utilised to reject clusters from multiple particle incidents and hadronic showers (jets).

Based on the likelihood, three different working points (WPs) are defined with average electron identification efficiencies of 93 % (Loose), 88 % (Medium), and 80 % (Tight) [150]. The efficiencies increase from low to high electron $E_{\mathrm{T}}$. The background rejection with respect to QCD processes with two jets in the final state is improved for the Medium and Tight WP by a factor of 2 and 3.5, respectively, with respect to the Loose WP [150].

### Muon identification

Muons are reconstructed by combining the information from ID, calorimeters and MS. Different techniques exist to reconstruct muons in ATLAS [147]. Most muons are reconstructed by matching a track in the MS to an ID track and performing a combined track fit based on the hits in both detectors, taking into account the energy loss in the calorimeters. Based on the fitted trajectory, hits may be added or removed and the fit repeated. Muons reconstructed in this way are referred to as combined muons – CB muons.

[7] The transverse energy is defined as $E_{\mathrm{T}} = E \sin\theta$. The $E_{\mathrm{T}}$ value of a cluster is calculated by summing the $E_{\mathrm{T}}$ values of all cells associated to the cluster.

Muons are also reconstructed by a combined track fit from tracks in the ID that are matched to at least three loosely aligned hits in the MS (inside-out muons – IO muons) [147]. This procedure recovers muons where the standalone muon track fit did not yield a muon candidate, e.g. in regions with limited MS coverage and for low $p_{\mathrm{T}}$ muons, which may not reach all MS segments [147].

Other muon reconstruction methods target muons with low-$p_{\mathrm{T}}$, small ($|\eta| < 0.1$) or large $\eta$ ($2.5 < |\eta| < 2.7$) [147]. These methods are not detailed here, as they are not used in this analysis.

High-quality muon candidates are selected based on the number of hits in the ID and MS, the compatibility of ID and MS tracks, and the combined track fit properties [147]. Non-prompt muons from light charged hadrons can be identified by a kink in the trajectory (where the hadron decays into a muon and another particle). Muon tracks from heavy-flavour hadron decays don't agree well with the PV position and are less isolated.

Three standard WPs, *Loose*, *Medium*, and *Tight*, as well as two specialised WPs for low and high $p_{\mathrm{T}}$ muons, are defined [147]. The WPs set different requirements for the number of hits in the ID, hit segments in the MS layers, $\chi^2$ value of the combined track fit, and agreement of ID and MS tracks [147]. The Medium and Tight WP consider in the range $\eta < 2.5$ only CB and IO muons [147]. Efficiencies to pass the described WPs have been determined in $t\bar{t}$ events. Muons in a $p_{\mathrm{T}}$ range from 20 to 100 GeV pass the Loose, Medium, and Tight WP in 99 %, 97 % and 93 % of the cases, with efficiencies to misidentify light charged hadrons as muons of 0.25 %, 0.17 % and 0.12 %, respectively [147].

#### Isolation

The isolation of a lepton is a powerful metric to reject non-prompt leptons from heavy-flavour decays. In this analysis, isolation requirements based on the output of a boosted decision tree (BDT), referred to as *non-prompt-lepton BDT* [153, 154], are applied. The inputs are optimised separately for the electron and muon cases. The BDTs include information describing the track isolation in the ID, the cluster isolation in the calorimeter, and properties of the lepton compared to that of a matched track-jet (for the definition see Section 3.4.5). Common variables are, among others, two dedicated lifetime variables: the output of a prompt lepton recurrent neural network (RNN) and the maximal longitudinal significance of the electron to a secondary vertex. The neural network and secondary vertex algorithm are fed with tracks in a cone around the lepton candidate. Two calibrated working points exist for the non-prompt-lepton BDT referred to as *PLIV-Tight* and *PLIV-VeryTight* [153, 154].

#### Calibration and definitions

The energy (momentum) scale of electrons (muons) is calibrated in simulation and data using $Z \to \ell\bar{\ell}$ and $J/\Psi \to \ell\bar{\ell}$ events [155, 156]. The events are also used to calibrate reconstruction, identification and isolation efficiencies in simulation. The energy resolution of electrons is also improved based on shower development properties using a multivariate regression algorithm. [157]

Leptons in this analysis are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$. Electrons in the barrel-endcap transition region ($1.37 < |\eta| < 1.52$) and with $|\eta| > 2.47$ are excluded. The transverse impact parameters measured with respect to the beam-spot position and the longitudinal impact parameter measured

Table 3.1: Definition of electron and muon collections used in this analysis. Given are the identification and isolation WPs they have to fulfil.

| Collection name | Identification WP | Isolation WP |
|---|---|---|
| Very Loose | Medium | - |
| Loose | Medium | PLIV-Tight |
| Loose-ID | Medium | PLIV-VeryTight |
| Loose-Isol | Tight | PLIV-Tight |
| Tight | Tight | PLIV-VeryTight |

with respect to the PV of electrons (muons) have to satisfy $|d_0|/\sigma(d_0) < 5$ ($|d_0|/\sigma(d_0) < 3$) and $|z_0 \sin\theta| < 0.5\,\text{mm}$.

Multiple lepton definitions that differ by the selected identification and isolation WPs are used in this analysis and are summarised in Table 3.1. Even though the electrons and muons are completely different objects, the selected WP names used to define the electron and muon collections agree with each other.

### 3.4.5 Jets

Quarks (except the top quark) or gluons produced in the hard scattering will fragment and hadronise into colour-neutral states, resulting in a spray of collimated particles, a jet (cf. Section 3.1). This section introduces the techniques used to reconstruct and calibrate jets.

#### Particle-flow algorithm

Jets are reconstructed using information from the inner detector and calorimeters that are combined in a particle-flow algorithm [158][8]. ID tracks matched to the PV and topological clusters are used to measure the charged and neutral constituents of the jet, respectively.

The algorithm matches tracks to, possibly multiple, topological clusters and subtracts the expected deposited energy cell by cell. If the energy of the matched clusters matches the track's $p_{\text{T}}$, given expected shower fluctuations, the clusters are removed completely. The jet reconstruction is then performed on the tracks and the, possibly modified, topological clusters (particle-flow objects).

The advantage of the particle-flow algorithm compared to a pure calorimeter-based reconstruction is the improved momentum (and angular) resolution of the ID tracker for low-$p_{\text{T}}$ particles. Additionally, low-momentum particles are recovered as the ID reconstructs tracks with $p_{\text{T}}$ down to 0.5 GeV, which usually don't form a seed in the topological cluster formation. Lastly, as the tracks can be assigned to a vertex, charged energy depositions from pile-up can be subtracted from the clusters.

#### Jet clustering

The jets are reconstructed by clustering nearby particle-flow objects. The anti-$k_t$ jet clustering algorithm [159, 160] sequentially combines nearby objects based on their energy/momentum and angular

[8] Jets can also be reconstructed using solely calorimeter or ID information. Reconstructed jets based only on ID tracks are commonly referred to as *track-jets*.

distance to each other. Central is the distance measure

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

Here, $i$ and $j$ are indices of particles and pseudo-jets, and $R$ is the radius parameter of the algorithm. The algorithm starts with the hardest object and recombines it with nearby soft objects iteratively into a pseudo-jet. This recombination stops if no more particle with $d_{ij} < p_{\mathrm{T,pseudo\text{-}jet}}^{-2}$ exists. In this case, the pseudo-jet is a final jet. The clustering procedure repeats, excluding the final jets, until no objects are left. Jets clustered using the anti-$k_t$ algorithm typically have a conical/circular shape with radius $R$. The shape only differs in case a more energetic jet is nearby. The radius parameter is chosen based on the use case and optimised to contain all particles of a jet while minimising contributions from other particles and pile-up. The common radius for (small-radius) jets at ATLAS is $R = 0.4$. An important feature of the anti-$k_t$ algorithm is that it is infrared and collinear safe. This means that neither an additional soft radiation nor a collinear splitting in two energetic particles would change the jet properties remarkably.

#### Jet energy calibration and definitions

Jets require a careful calibration, as they are complex objects consisting of multiple particles with different characteristics. Corrections are applied to data and MC in multiple sequential steps, calibrating the jet energy scale (JES) and resolution (JER) [161]:

**A Pile-up correction** is applied as jets span over a relatively large volume in the detector and are therefore particularly sensitive to it. Corrections are applied based on the jet area, the median $p_{\mathrm{T}}$ density of the event, the mean number of interactions in a bunch crossing and the number of reconstructed vertices in the event.

**The overall JES correction** calibrates the absolute JES at reconstruction level to particle level. In simulated dijet events, the energy of reconstructed jets is compared to the energy of jets built from truth-particles (truth-jets). This calibration accounts for energy losses in passive detector material, the non-compensating calorimeter response, and out-of-cone effects resulting from the jet clustering with limited radius. The corrections are energy- and $\eta$-dependent because the detector response depends on the energy of the incoming particle and the detector structure is non-uniform in $\eta$.

**A global sequential calibration** is applied to improve the JER by accounting for the energy-dependent detector response. Corrections are applied sequentially based on observables describing the longitudinal shower shape, hits in the MS, track multiplicity, and the $p_{\mathrm{T}}$ and position of associated ID tracks.

**In-situ jet calibrations** correct for differences between data and simulation which are caused by imperfect simulation of the physics processes and the detector materials. In the in-situ calibrations the jet response is measured with respect to a well-calibrated reference object. This jet response in simulations and data are compared to each other and eventually corrections are applied to data. Three sequential calibrations are applied. The first calibration corrects the JES of forward jets ($0.8 < |\eta| < 4.5$) to match the JES of central jets ($|\eta| < 0.8$). The second calibration uses

$Z/\gamma$+jet events to calibrate the jet response with respect to a calibrated $Z(\rightarrow ee/\mu\mu)/\gamma$. The third calibration uses multijet events to calibrate a high-$p_{\mathrm{T}}$ jet against multiple calibrated lower-$p_{\mathrm{T}}$ jets.

The JER decreases steadily with increasing energy, is smaller for central jets, and varies between 4 and 25 % [161]. The uncertainty on JER varies between 1 and 5 % and is smallest for intermediate jet energies [161]. The energy resolution of jets is improved substantially by the calibration, but compared to the lepton energy and momentum resolutions, it is still about an order of magnitude larger [155, 156, 161].

Jets in this analysis are required to have $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$. Jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 2.4$ additionally have to pass a jet-vertex-tagger [162] requirement in order to suppress jets originating from pile-up. The jet-vertex tagger determines the probability for a jet to originate from a PV based on the tracks associated with the respective jet.

### 3.4.6 Flavour tagging of jets

For this and also many other analyses it is important to know the flavour of the quark that initiated a jet. Because the couplings of quarks to other particles are usually specific for each quark flavour. For example, due to the nature of the weak interaction, the top quark almost exclusively decays into a $W$ boson and a $b$ quark. Identifying $b$-quark-initiated jets reliably is crucial for selecting and reconstructing top-quark events.

Jets containing $b$- or $c$-hadrons (commonly referred to as $b$- or $c$-jets) can be distinguished from jets containing only hadrons of other flavours (light-jets) based on multiple properties. Among them are the larger invariant mass and decay multiplicity of the heavy-flavour hadrons as well as differences in the fragmentation properties. Most important is the lifetime. Heavy-flavour hadrons travel on average a macroscopic distance before decaying ($\langle c\tau \rangle \approx 450\,\mu\mathrm{m}$ for $b$-hadrons with $p/m = 1$) [163]. This leads to distinct signatures in the detector, like displayed secondary vertices from heavy-flavour hadron decays, tracks not pointing towards the PV, etc. (see Figure 3.12).

The mass and lifetime of $c$-hadrons are in between $b$- and light-hadrons. This makes it generally more difficult to differentiate $c$-jets from $b$- and light-jets than differentiating $b$- from light-jets.

In Run-2, ATLAS identified $b$-, $c$-, and light-jets using a feed-forward neural network (high-level tagger – *DL1r* [163]) that combines the outputs from multiple lower-level taggers into one powerful discriminant. The lower-level taggers use information from jets and ID tracks associated with the jet based on the momentum at the perigee and the $\Delta R$ with respect to the jet axis:

**IP2D and IP3D** [165] are log-likelihood discriminants based on the transverse (2D) and transverse+longitudinal (3D) signed impact parameter significance sum of all associated tracks. The sign is positive if the track intersects the jet axis in the transverse plane in front of the PV. Tracks originating from the decay of long-lived heavy-flavour hadrons lead to large values of IP2D and IP3D.

**RNNIP** [166] is a RNN that uses the signed impact parameter significances, number of hits, $\Delta R$ with respect to the jet axis, and the $p_{\mathrm{T}}$ fraction of tracks associated to a jet as inputs. The network yields probabilities of a jet being a $b$-, $c$-, or light-jet. The main advantage with respect to the IPxD taggers is that the RNNIP can consider correlations among the tracks of a jet.

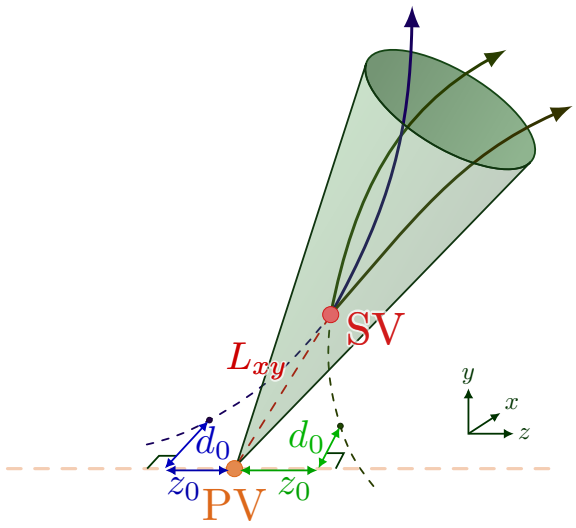



Figure 3.12: Sketch of a heavy-flavour jet and related flavour-tagging quantities. [164] PV and SV are the primary and secondary vertices, respectively. $L_{xy}$ is the distance between the two in the plane transverse to the beam pipe.

**SV1** [167] is a secondary vertex tagging algorithm that reconstructs a single secondary vertex in a jet. The procedure is similar to the primary vertex finding. Vertex candidates compatible with strange hadrons or detector interactions are removed, and the invariant mass of the vertex has to be less than 6 GeV. Properties of the secondary vertex, like the invariant mass, number of associated tracks, energy fraction and decay length significance, are used as inputs for the DL1r tagger.

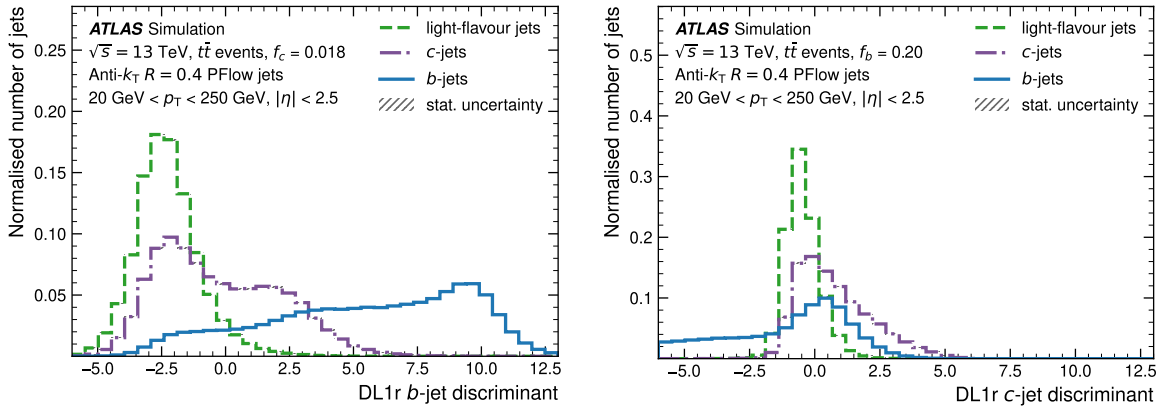
**JetFitter** [168] is a multi-vertex finding algorithm that aims at reconstructing the $b$- and $c$-hadron decay vertices on a common line with the PV. The JetFitter algorithm therewith reconstructs the full $b$-hadron decay chain and can even reconstruct decay vertices with a single associated track. Vertex quantities, similar to the SV1 ones but for both vertices, are used as inputs for the high-level tagger.

The DL1r algorithm also includes the $p_{\mathrm{T}}$ and $|\eta|$ of the jets and takes advantage of all correlations among the input variables. Output of the DL1r tagger are three values which describe the probability of a jet being a $b$- ($p_b$), $c$- ($p_c$) or light-jet ($p_l$). The final $b$- and $c$-tagging discriminants are log-likelihood ratios of the probabilities,

$$\mathcal{D}_{\mathrm{DL1r}} = \ln\left(\frac{p_b}{f_c p_c + (1-f_c)p_l}\right), \qquad \mathcal{D}^c_{\mathrm{DL1r}} = \ln\left(\frac{p_c}{f_b p_b + (1-f_b)p_l}\right).$$

$f_c = 0.018$ and $f_b = 0.3$ are the effective $c$- and $b$-jet fractions in the background hypothesis and can be optimised for the given use case. The values in this analysis are chosen to yield an overall good separation versus the other jet flavours. The distributions of both discriminants are shown in Figure 3.13.

Generator-specific corrections are applied to simulated events to compensate for differences in the tagging efficiencies and inefficiencies between data and simulation [169–171]. The correction factors




(a) (b)

Figure 3.13: Distributions of the (a) $\mathcal{D}_{\text{DL1r}}$ and (b) $\mathcal{D}^{c}_{\text{DL1r}}$ discriminants for the individual jet flavours in $t\bar{t}$ simulated events. [163]

are derived using $t\bar{t}$ and $Z$+jets events in bins of jet $p_{\text{T}}$. Calibrations are available for single (fixed-cut) and multiple (pseudo-continuous) cuts on the flavour-tagging discriminants. The cuts are also referred to as *working points* (WPs). The calibration used in this analysis utilises cuts on $\mathcal{D}_{\text{DL1r}}$ and $\mathcal{D}^{c}_{\text{DL1r}}$ to identify $b$- and $c$-jets, as both types are present in the signal process final state – *pseudo-continuous flavour-tagging* (PCFT) calibration. The calibration cuts and corresponding tagging efficiencies are depicted in Figure 3.14.

Jets passing the loose or tight $b$-tagging WP cut receive PCFT scores of 3 or 4, respectively, and are labelled as $b$-tagged jets. Other jets ($b$-veto) passing the loose or tight $c$-tagging cut receive PCFT scores of 1 or 2, respectively, and are labelled as $c$-tagged jets. The remaining jets are untagged and receive a PCFT score of 0. Roughly 70 % and 60 % of the $b$-jets pass the loose and tight $b$-tagging WPs, respectively. And roughly 45 % and 25 % of the $c$-jets pass the loose and tight $c$-tagging WPs, respectively.

The available PCFT working points have been optimised for the measurement of $VH$ ($V = W, Z$) production with $H \to b\bar{b}$ and $H \to c\bar{c}$ [172]. Studies at an early stage of the analysis showed that a dedicated calibration for this analysis would yield limited improvements. The studies are not further outlined in this thesis, however, they yielded at most a 1 to 2 % improvement in the expected limits on the charged Higgs boson production, considering only statistical uncertainties. For this reason, the flavour-tagging calibration from the $VH$ analysis has been adopted by this analysis.

Another set of pseudo-continuous calibrated WPs has been studied in this analysis, too. The set contains four calibrated $b$-tagging WPs with $b$-jet efficiencies of 85 %, 77 %, 70 %, and 60 %. The latter two WPs are identical to the two $b$-tagging WPs defined above. The calibration has no calibrated $c$-tagging WPs and is referred to as *pseudo-continuous b-tagging* (PCBT) calibration.

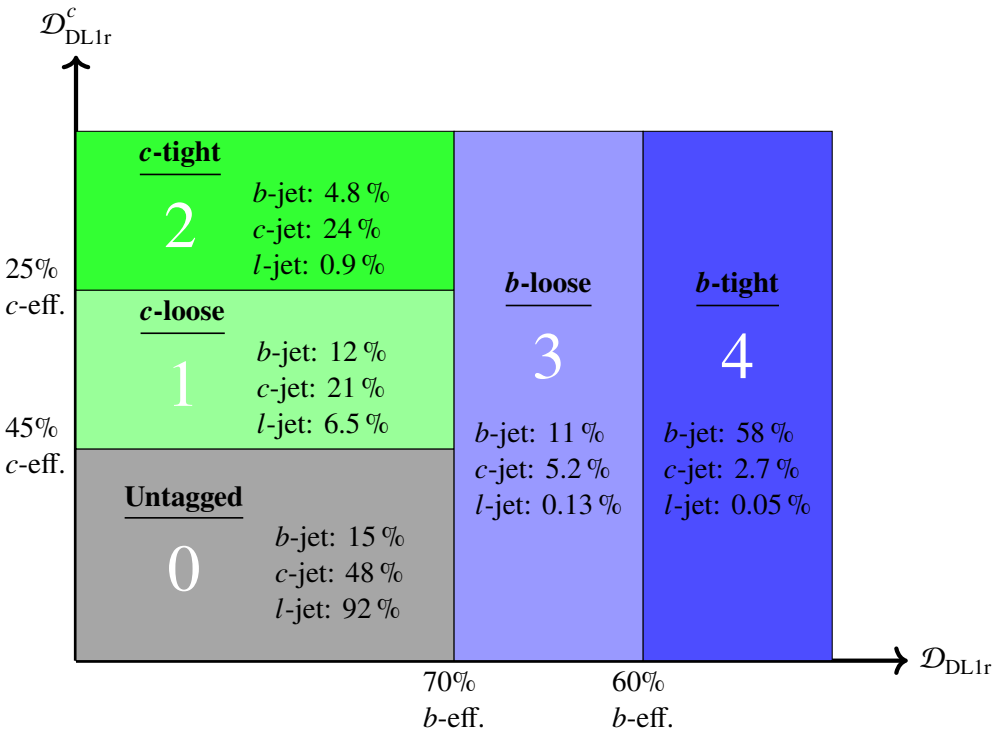



Figure 3.14: Flavour-tagging calibration bins. The white numbers correspond to the PCFT score of jets in the respective bin. The efficiencies for $b$-, $c$-, and light-jets (denoted as "$l$-jet" in the Figure) are derived in $t\bar{t}$ events and are taken from Reference [172].

### 3.4.7 Missing transverse momentum

Solely weakly interacting particles have a tiny interaction cross-section, which makes it impossible to detect them directly on a reliably basis nowadays. Neutrinos fall into this category, but also new, undiscovered particles could leave the detector undetected. Though, these particles can be detected indirectly, by considering the momentum balance of all other particles in the transverse plane. The momentum of the colliding particles along the transverse plane can be neglected with respect to their momentum along the $z$-axis. Due to energy-momentum conservation, the transverse momentum sum of all particles emerging from the collision has to vanish. If this is not the case, a particle must have left the detector undetected.

The missing transverse momentum ($\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$) is defined as the negative vector sum of the transverse momenta of all reconstructed and calibrated objects and all ID tracks which are matched to the PV but not associated with one of the aforementioned objects [173]. Considering the objects used in this analysis, this is

$$\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}} = -\left(\sum p_{\mathrm{T}}^{e} + \sum p_{\mathrm{T}}^{\mu} + \sum p_{\mathrm{T}}^{\mathrm{jet}} + \sum p_{\mathrm{T}}^{\mathrm{unused\ tracks}}\right).$$

The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is often expressed in terms of its absolute value, commonly denoted by $E_{\mathrm{T}}^{\mathrm{miss}}$, and its azimuthal angle, $\phi^{\mathrm{miss}}$.

The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ observable is most difficult to determine precisely, as it depends on all other objects, and the uncertainties add up. Mismeasurements, pile-up, and detector acceptance and resolution effects can fake $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ in the detector.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is a log-likelihood ratio based on the resolution of all objects and quantifies the belief of a $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement to be real – events with a neutrino receive on average larger $E_{\mathrm{T}}^{\mathrm{miss}}$ significances.

### 3.4.8 Signal ambiguity resolution

As a last step, in order to resolve ambiguities in which one detector signature is used in the reconstruction of multiple physical objects, a dedicated, sequential overlap removal (OR) procedure is applied, outlined below [174]:

1. Remove any electron sharing a track with a muon.
   - Electron-Muon OR: This criterion removes fake electron candidates reconstructed from a muon that radiated a high-energy bremsstrahlung photon.
2. Remove any jet within $\Delta R = 0.2$ of an electron.
3. Remove any electron within $\Delta R = 0.4$ of a jet.
   - Electron-Jet OR: Criteria 2 and 3 remove jets overlapping with prompt electrons while aiming to preserve jets with electrons from heavy-flavour decays and light-flavour jets faking electrons.
4. Remove any jet with less than 3 tracks that is within $\Delta R = 0.2$ of a muon.
5. Remove any jet with less than 3 tracks that has a muon ID track ghost-associated [159, 175] with it.
6. Remove any muon within $\Delta R = 0.4$ of a jet.
   - Muon-Jet OR: The motivations behind criteria 4 to 6 are similar to the electron-jet case. The looser requirement of only removing the jet if it has fewer than three associated tracks, compared to the electron-jet case, is owed to the reduced probability of a muon to fake a jet. Only muons with a high energetic FSR or bremsstrahlung photon can potentially fake a jet.

CHAPTER 4

# Event selection and reconstruction

This thesis presents the search for a light charged Higgs boson in the $cs$ final state using $pp$ collision data collected by the ATLAS experiment. As presented in Section 2.2.1, for $m_{H^\pm} < m_t$ the main production mode of a charged Higgs boson is in the decay of a top quark. At the LHC, top quarks are produced predominantly in pairs. The focus in this analysis is set on semileptonic $t\bar{t}$ events; this means one top quark decays into a $W$ boson and a $b$ quark, with the $W$ boson decaying into a lepton and a neutrino, $t\bar{t} \to H^\pm(\to cs)bW(\to \ell\nu)b$. The $H^\pm \to cs$ decay channel is being searched for as it is one of the main decay channels in 2HDM models and even dominates for certain models if $m_{H^\pm} < m_t$ (for a detailed discussion see Section 2.2.1).

The roadmap for physics analyses in ATLAS is generally quite similar. The first step is to define the data sample to be analysed and to produce simulated event samples of the signal process and all other processes, referred to as background processes or simply backgrounds, that might mimic a signal process in the detector. Data and simulated samples are described in Section 4.1.

Based on the final state under investigation, usually a pre-selection on data is defined to obtain a sample enriched in interesting data events (Section 4.2). Thereupon the modelling of simulated event samples is thoroughly validated in signal and background enriched regions (Section 4.3).

The reconstruction of physics objects from data has been described in the previous chapter (Section 3.4). For analyses with final states containing short-lived resonances, it is often beneficial to reconstruct the whole event topology. In this thesis the full $t\bar{t}$-system topology is reconstructed, as described in Section 4.4.

In general, a sophisticated strategy is developed to separate the signal process from background processes with similar characteristics, which is very analysis dependent. The final step is then the statistical interpretation of the results. In this search this means quantifying how certain we are about the existence or non-existence of a $H^\pm \to cs$ signal in data. These two steps are presented in Chapters 5 and 6.

## 4.1 Data and simulated event samples

### Data sample

The data analysed in this thesis corresponds to $pp$ collision data recorded with the ATLAS detector at a centre-of-mass energy of 13 TeV during LHC Run-2 in the years 2015 to 2018. The data events are scrutinised for good operation conditions of all detector components and related hardware and software [176]. The total integrated luminosity of used data samples is $(140.0 \pm 1.2)\,\text{fb}^{-1}$ [133]. The integrated luminosity for the years 2015, 2016, 2017 and 2018 are 4, 39, 50.6 and $63.8\,\text{fb}^{-1}$, respectively [133]. The $\beta^*$ at the interaction point decreased over the years, yielding an increased peak in the mean number of interactions per bunch crossing ($\langle\mu\rangle \approx 16$ to 60) and instantaneous luminosity ($\mathcal{L}_{\text{peak}} \approx 0.5$ to $1.9 \times 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$) [133].

### Simulated samples

MC simulated event samples are used to model the signal process (Section 4.1.1) and SM backgrounds. All SM processes are considered, which can abundantly yield, or mimic, a final state with a prompt, high-energy lepton and jets. The focus of the event selection is set on leptons, as the signatures of prompt electrons and muons are comparatively clean in the detector. The main backgrounds are processes involving a top quark (Section 4.1.2) or a weak boson $V = W, Z$ (Section 4.1.3). All simulated signal and background processes are listed in Table 4.1 with their nominal generator setups. The background from multijet processes is considered in this analysis but not listed in Table 4.1 as it is difficult to model properly and consequently not estimated from simulation. A data-driven estimation is performed as described in Section 4.3.

In the following, generators and settings used in the event generation of the nominal samples are specified. The detector simulation and digitisation steps are the same for all samples. All nominal background samples were passed through the full ATLAS detector simulation. Alternative background samples and signal samples were passed through a fast detector simulation. The effects of pile-up were modelled by overlaying each hard-scatter event with minimum-bias events, simulated using soft QCD processes of Pythia 8.186 [177] with a set of tuned parameters called the A3 tune [178] and the NNPDF2.3lo [179] set of PDFs.

### 4.1.1 Signal samples

#### $H^{\pm} \to cs$

Signal events are generated by first generating top-quark pairs at NLO using the Powheg Box v2 [180–183] generator with the NNPDF3.0nlo [184] PDF set and the $h_{\text{damp}}$[1] parameter set to $1.5\,m_t$ [185]. The top-quark mass is set to 172.5 GeV. The decays of $t \to H^{\pm}b$ and $t \to W^{\pm}b$ were modelled by MadSpin [186], using the Type-II 2HDM configuration for BSM decays, to retain spin correlation effects. Subsequent decays of the $H^{\pm}$ and $W^{\pm}$, PS, hadronisation, and UE were modelled by Pythia 8.307 [187]. Therein, the $W$ boson was forced to decay leptonically ($W^{\pm} \to \ell^{\pm}\nu_{\ell}$), allowing for all three lepton flavours ($\ell = e, \mu, \tau$). The charged Higgs boson was forced to decay into a $cs$-quark pair ($H^{\pm} \to cs$). The Pythia parameters were set to the A14 tune [188], and the NNPDF2.3lo PDF set was used. The

[1] The $h_{\text{damp}}$ parameter is a resummation damping factor and is one of the parameters that controls the matching of Powheg ME to the PS. It regulates the first high-$p_{\text{T}}$ emission against which the $t\bar{t}$ system recoils.

Table 4.1: List of signal and background processes and the generators used to simulate them. The symbol "$q$" is used for $u, d, c, s$ quarks. The subscript "$x$" at the signal sample names is a placeholder for the mass of the charged Higgs boson in GeV. For the Other top and $VV$ processes, only the range of used generator versions is quoted. The exact generator version are given in Section 4.1.

| Name | Process | ME Generator | PS and hadronisation |
|---|---|---|---|
| **Signal** | | | |
| $H_x^{pm}$ | $t\bar{t} \to b\bar{b}W^{\pm}(\to \ell\nu_\ell)H^{\mp}(\to cs)$ | Powheg Box v2 | MadSpin + Pythia 8.307 + EvtGen 1.7.0 |
| $H_{cb,x}^{\pm}$ | $t\bar{t} \to b\bar{b}W^{\pm}(\to \ell\nu_\ell)H^{\mp}(\to cb)$ | | |
| **Top quark** | | | |
| $t\bar{t}(ud)$ | $t\bar{t} \to b\bar{b}W^{\pm}(\to \ell\nu_\ell)W^{\mp}(\to ud/\ell\nu_\ell)$ | Powheg Box v2 | Pythia 8.230 + EvtGen 1.6.0 |
| $t\bar{t}(cs)$ | $t\bar{t} \to b\bar{b}W^{\pm}(\to \ell\nu_\ell)W^{\mp}(\to cs)$ | | |
| $t\bar{t}$ + HF | $t\bar{t} \to b\bar{b}W^{\pm}(\to \ell\nu_\ell)W^{\mp}(\to q\bar{q}/\ell\nu_\ell) + \geq 1c/b$ | | |
| $t\bar{t}$(allHad) | $t\bar{t} \to b\bar{b}W^{\pm}(\to q\bar{q})W^{\mp}(\to q\bar{q})$ | | |
| $tW$ | $tW$ | | |
| Single top | single $t$-quark $s$- & $t$-channel | | |
| Other top | $t\bar{t}H$ | Powheg Box v2 | Pythia 8.230 + EvtGen 1.6.0 |
| | $t\bar{t}W$ | MadGraph5_aMC@NLO 2.3.3 | Pythia 8.210 + EvtGen 1.2.0 |
| | $t\bar{t}Z$ | | |
| | $tWZ$ | MadGraph5_aMC@NLO 2.3.3 | Pythia 8.210 + EvtGen 1.2.0 |
| | $t\bar{t}t\bar{t}$ | MadGraph5_aMC@NLO 2.3.3 | Pythia 8.230 + EvtGen 1.6.0 |
| | $tHbj$ | MadGraph5_aMC@NLO 2.6.2 | Pythia 8.230 + EvtGen 1.6.0 |
| | $tWH$ | MadGraph5_aMC@NLO 2.6.2 | Pythia 8.235 + EvtGen 1.6.0 |
| | $t\bar{t}t$ | MadGraph 2.2.2 | Pythia 8.212 + EvtGen 1.2.0 |
| | $tZq$ | | |
| **Weak boson** | | | |
| $W$+jets | $W$+jets | Sherpa 2.2.11 | Sherpa 2.2.11 |
| $Z$+jets | $Z$+jets | | |
| $VV$ | $WW, WZ, ZZ$ | Sherpa 2.2.1–2.2.2 | Sherpa 2.2.1– 2.2.2 |

EvtGen 1.7.0 program [189] was used to model the decays of bottom and charm hadrons.

Samples were produced for several charged Higgs boson mass hypotheses. The kinematics of the $t \to H^{\pm}(\to cs)b$ decays change with the value of $m_{H^{\pm}}$. Depending on the $m_{H^{\pm}}$ search range and mass resolution, multiple samples are needed to (dis)prove the existence of a charged Higgs boson. In Section 2.2.1 it was shown that $m_{H^{\pm}}$ is only one of many free parameters in 2HDMs, the simplest extension of the SM containing a charged Higgs boson. But other parameters, e.g. $\tan\beta$, mainly change the coupling of the charged Higgs boson to other particles. This results in a modified charged Higgs production cross-section, but the kinematics are unchanged.

Twelve signal samples were generated for charged Higgs boson masses from 60 GeV to 168 GeV with 1.8 million simulated signal events each. Eleven samples in steps of 10 GeV from 60 GeV to 160 GeV and one sample at 168 GeV. The latter mass point was chosen such that $m_{H^{\pm}} + m_b$ is right below the top-quark mass threshold to maintain on-shell $t \to H^{\pm}b$ decays. The lowest mass point selection is motivated by exclusions from previous searches (cf. Section 2.3) and the acceptance of signal events in the event selection (cf. Section 4.2). For smaller charged Higgs boson masses a different analysis strategy with a looser jet $p_{\mathrm{T}}$ requirement is advisable. The mass point step size was chosen to be below the dijet mass resolution [161]. This way the search will be sensitive to any charged Higgs boson in the given mass range. The signal samples are denoted by $H_x^{\pm}$, where "$x$" is the mass of the charged Higgs boson in GeV.

Figure 4.1 shows kinematic distributions of the signal process for different charged Higgs boson

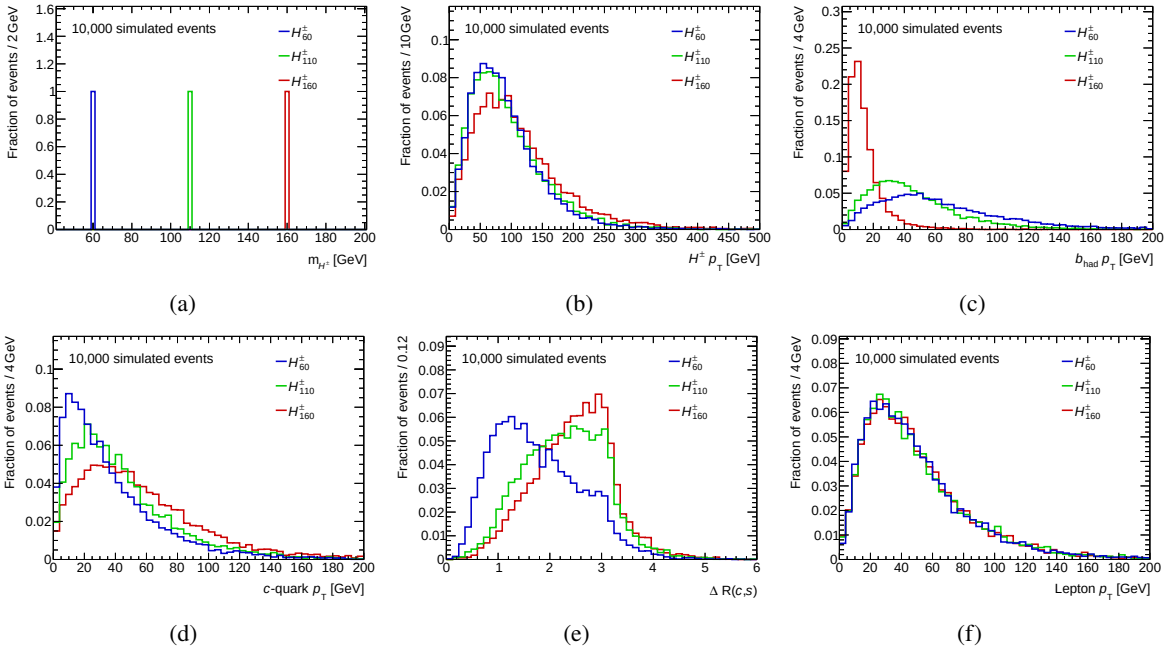



Figure 4.1: Comparison of kinematic quantities of simulated $t\bar{t} \to b\bar{b}H^{\pm}(\to cs)W^{\mp}(\to \ell\nu_{\ell})$ events for different charged Higgs boson masses. $b_{\text{had}}$is the $b$ quark that has been produced in the decay $t \to bH^{\pm}$.

masses using the truth four-momenta from simulation. These plots have been produced to validate the modelling of the signal process. The distributions change as expected with increasing charged Higgs boson mass. For example, the $c$-quark $p_{\text{T}}$ is on average larger for heavier charged Higgs bosons. The lepton $p_{\text{T}}$ is included as a cross-check and is, as expected, unaffected by the charged Higgs boson mass. The decay width of the charged Higgs boson, as can be seen in Figure 4.1(a), is assumed to be narrow (set to zero) in simulation. This assumption is being made because the decay width of the charged Higgs boson is way below 1 GeV for most of the parameter space [190], calculated using FeynHiggs [191] and HDECAY libraries, and thereby at least one order of magnitude smaller than the dijet mass resolution. If the mass difference between the $H^{\pm}$ and the $W$ boson is smaller than either of their total widths, the interference between the decays $t \to H^{\pm}b$ and $t \to W^{\pm}b$ can become relevant. The size and sign of the interference term depends on the model but is at maximum of the order of a few percent of the $H^{\pm}$ contribution [192]. For larger mass differences, the interference term can be omitted with high accuracy. In this analysis, the interference term is neglected for all mass points.

The production of charged Higgs bosons via single-top-quark processes is neglected in this analysis. Generally, the single top-quark production cross-section is about a factor of four smaller than $t\bar{t}$ production in LHC Run-2 [28]. But the main reason is that these events do not contain a prompt lepton and are highly suppressed by the event selection (cf. Section 4.2).

### $\mathbf{H^{\pm} \to cb}$

In addition to the $H^{\pm} \to cs$ signal samples, samples with $H^{\pm} \to cb$ have been produced. The used generators and settings are identical to the $H^{\pm} \to cs$ signal sample generation presented above, except

that the charged Higgs boson is forced to exclusively decay into a $cb$ quark pair. These samples are marked by an additional $cb$ subscript, $H^{\pm}_{cb,x}$. The $H^{\pm}_{cb}$ samples were produced for the 110 GeV, 130 GeV and 150 GeV charged Higgs boson mass points. The mass point selection is motivated by the $3\sigma$ ($2.5\sigma$) local (global) excess found at 130 GeV by the dedicated $H^{\pm} \to cb$ analysis [91]. The $H^{\pm}_{cb}$ samples are solely used to cross-check the sensitivity of the analysis to a potential $H^{\pm} \to cb$ signal. The presumption was established because the final states differ only by having a $b$-jet instead of a light-jet emerging from the $H^{\pm}$ decay.

### 4.1.2 Top-quark samples

The main background in this search are SM $t\bar{t}$ events. The $t\bar{t}$ production processes are identical for SM $t\bar{t}$ and signal events, and the final-state particles often exactly agree.

For the production of $t\bar{t}$ and single-top-quark events in the $tW$-, $s$-, and $t$-channels, a similar generator setup as for the signal process has been used. The main difference is that the events were not interfaced with MadSpin. Powheg Box v2 calculates the ME at NLO using the five-flavour scheme (four-flavour scheme for single-top-quark $t$-channel events) with the NNPDF3.0nlo PDF set and $h_{\text{damp}} = 1.5\, m_t$. PS, hadronisation and UE are modelled using Pythia 8.230 with the A14 tune and NNPDF2.3lo PDF set. Heavy-flavour hadron decays are modelled using EvtGen 1.6.0. This is the nominal generator setup which has been found to show the best agreement with data [185]. The $t\bar{t}$ samples were normalised using cross-sections calculated at NNLO plus next-to-next-to-leading-logarithm (NNLL) accuracy using the Top++ program [193]. The production cross-section of the three single-top-quark channels are calculated at NLO +NNLL and NLO for the $tW$- and $s/t$-channels, respectively, using the Hathor program [194–198].

The simulated $t\bar{t}$ samples are categorised based on the decay of the top-quarks and the flavour of additional jets in the event. In SM $t\bar{t}$ simulation, a branching ratio ($\mathscr{B}$) of $\mathscr{B}(t \to Wb) = 100\,\%$ is assumed. Events in which both $W$ bosons decay hadronically rarely pass the event selection and are therefore of minor interest. These events are labelled as $t\bar{t}(\text{allHad})$.

In other $t\bar{t}$ events at least one $W$ boson decays leptonically. These events are categorised in $t\bar{t}+\text{LF}$ (LF for light-flavour) and $t\bar{t}+\text{HF}$ (HF for heavy-flavour) categories based on the existence of additional – not stemming from decay products of top quarks – heavy-flavour jets in the event. The categorisation is adopted because a precise prediction of heavy-flavour jets in association with a top-quark pair is challenging in current theoretical models due to the scale hierarchy between $t\bar{t}$ and $b\bar{b}$ production from gluon emission and the non-negligible mass of the $b$-quark [199, 200]. The heavy-flavour jets are identified using truth information from simulation as described in Reference [201]. For this purpose particle jets are built from truth stable particles not originating from the $t\bar{t}$ system using the anti-$k_t$ jet clustering algorithm with $R = 0.4$. The particle jets are required to have $p_{\text{T}} > 15$ GeV and $|\eta| < 2.5$. Particle jets with a matched $b$- or $c$-hadron are identified as a heavy-flavour jet, and events with at least one such jet are labelled as $t\bar{t}+\text{HF}$.

The $t\bar{t}+\text{LF}$ category is split further into $t\bar{t}(ud)$ and $t\bar{t}(cs)$ categories based on the decay of the $W$ boson. If the $W$-boson decay products contain a $c$-quark, the event is labelled as $t\bar{t}(cs)$, otherwise as $t\bar{t}(ud)$. The categorisation is motivated by the signal process. If the $W$ boson decays as $W \to cs$, the final states for signal and SM $t\bar{t}$ background are identical, which makes a separation challenging. The $t\bar{t}(ud)$ category also contains dileptonic $t\bar{t}$ events, i.e. events in which both $W$ bosons decay leptonically.

To mention is that in the decay of the $W$ boson a diagonal CKM matrix is assumed. Hence, the $W$

boson can decay hadronically only as $W \to cs$ or $W \to ud$, therefore the nomenclatures $t\bar{t}(ud)$ and $t\bar{t}(cs)$. The SM $t\bar{t}$ samples are produced centrally by ATLAS and used by many analyses. With the current reconstruction techniques, it is barely possible to reliably separate $d$- from $s$-quark-initiated jets, which justifies the use of these samples. The decay $W \to cb$ would show a different signature in the detector, but the corresponding CKM-matrix element is very small $V_{cb} = (40.8 \pm 1.4) \times 10^{-3}$ [28], leading to $\mathscr{B}(W \to cb) \approx 0.17\,\%$. The $H^{\pm} \to cb$ analysis [91] considered the $W \to cb$ background in their analysis, but it had a negligible impact on the final results. For this reason it is not considered in this analysis. Nonetheless, in the future many analyses could benefit by assuming a non-diagonal CKM matrix in $t\bar{t}$ sample productions.

Other background processes considered in this analysis and involving a top quark include, compared to the processes described above, additional weak bosons or top quarks in the final state: $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}t\bar{t}$, $t\bar{t}t$, $tHjb$, $tWH$, $tWZ$ and $tZq$. These processes have a much lower cross-section and are therefore grouped into a single category referred to as *Other top*. The $t\bar{t}H$ process was modelled using the same generators and versions as the exclusive $t\bar{t}$ process. The remaining processes were modelled using MadGraph and MadGraph5_aMC@NLO generators for ME calculation interfaced with Pythia 8 and EvtGen. The exact generator versions are listed in Table 4.1. The $t\bar{t}H$, $t\bar{t}W$, and $t\bar{t}Z$ cross-sections are calculated at NLO QCD and NLO EW accuracy using MadGraph5_aMC@NLO [202].

### 4.1.3 Weak-boson samples

Processes involving a leptonically decaying weak boson contain prompt leptons and are therefore an important background in this analysis. The production of a vector boson $V = W, Z/\gamma^*$ in conjunction with jets ($V$+jets) was modelled using Sherpa 2.2.11 [203] for ME and PS with the NNPDF3.0nnlo PDF set. In the ME calculation contributions from up to five (two) additional parton emissions were considered at LO (NLO), using the Comix [204] and OpenLoops [205–207] libraries.

The production of the diboson processes $VV = WW, WZ, ZZ$ with at least one prompt lepton in the final state are simulated using Sherpa 2.2.1 for ME and PS with the NNPDF3.0nnlo PDF set. The process $WZ \to \ell\nu\nu\bar{\nu}$ was simulated with Sherpa 2.2.2. Contributions with up to three (one) additional parton emissions are calculated at LO (NLO) precision [208]. Loop-induced diboson processes that are initiated via the $gg$ production mode are simulated with LO accuracy with up to one additional parton emission.

Merging of ME and PS was performed for all Sherpa samples using the MEPS@NLO formalism [209–211] with a merging scale of $Q_{\text{cut}} = 20\,\text{GeV}$ and the five-flavour numbering scheme. The default set of tuned parameters developed by the Sherpa authors was used.

## 4.2 Event selection

A basic event selection is applied to the data to reject background events and obtain a sample enriched in signal events. In the signal process a $t\bar{t}$-pair decays into two $b$-quarks, a charged Higgs boson, which decays further into quarks $H^{\pm} \to cs$, and a $W$ boson, which decays into leptons $W \to \ell\nu$. This results in a final state with one lepton and four jets, out of which two are $b$-jets, as visualised in Figure 4.2. The event selection aims at selecting events with exactly this signature.

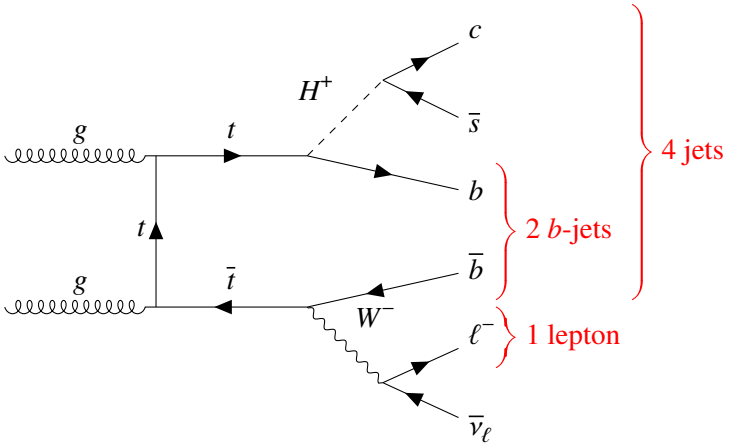



Figure 4.2: A Feynman diagram of the signal process highlighting the classes of final-state particles.

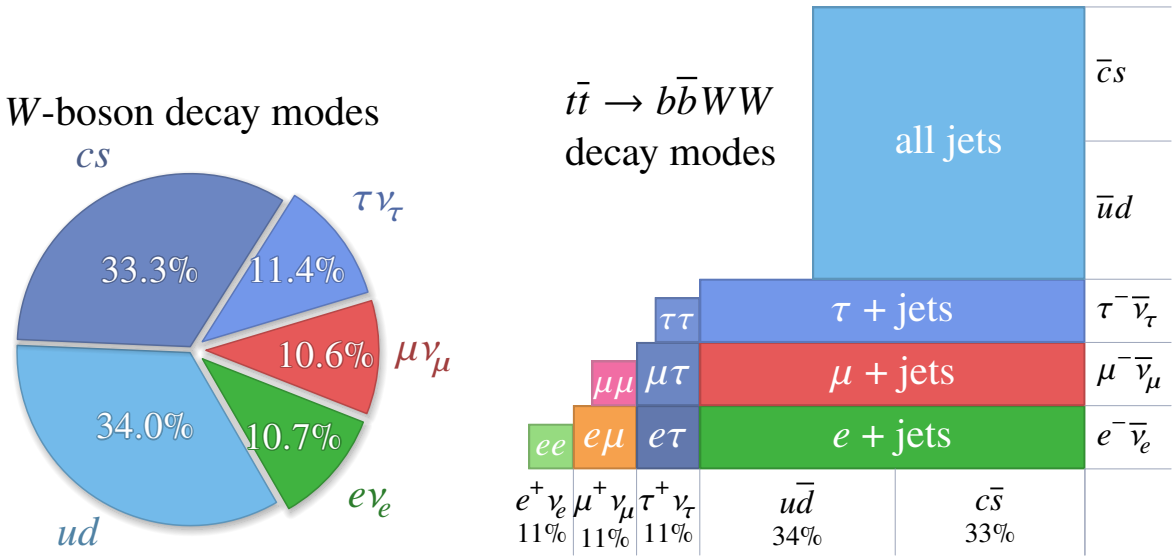
The focus is set on leptonic $W$-boson decays with an electron or muon in the final state, also including decays of a tauon which decays further into lighter leptons, $W \to \tau(\to \ell\nu_\ell)\nu_\tau$. The focus on semileptonic $t\bar{t}$ events with one electron or muon reduces the size of the selected signal sample to about one-fourth, as follows from the $W$-boson branching ratio into leptons of about 11 % per lepton flavour (cf. Figure 4.3) and the branching ratio of $\tau$-leptons into lighter leptons of about 34 % [28]. The advantage of this restriction is that the leptonic final state facilitates an efficient suppression of abundant QCD processes and eases the event reconstruction due to a reduced number of jet combinatorics.

Regarding the main background, Figure 4.3 shows the possible decay modes of SM $t\bar{t}$ events. In the SM the top quark almost exclusively decays into a $W$ boson and a $b$ quark and the further signature is determined by the decay mode of the $W$ boson. In about 45 % of $t\bar{t}$ decays, the final state is fully hadronic; in about 44 %, the final state has one lepton (semileptonic); and in about 11 %, there are two leptons (dileptonic). The semileptonic decay modes have exactly the same final state as the signal process. The other decay modes are attempted to be rejected by the event selection.

The first step in selecting data events is the trigger system. At trigger level the focus is set on electrons and muons due to their clean signatures in the detector and fast reconstruction. The focus on prompt, high-$p_\mathrm{T}$ leptons facilitates an efficient rejection of abundant backgrounds related to QCD processes. Efficient background rejections are crucial for selecting as many signal events as possible. To be specific, it allows for having low lepton $p_\mathrm{T}$ thresholds.

The reconstruction of electrons and muons at trigger level is similar to the offline reconstruction described in Section 3.4, but computationally expensive steps are either replaced by less complex algorithms or additional selections are applied beforehand. Details are given in References [214–216]. This analysis uses the lowest unprescaled electron and muon triggers, which are listed in Table 4.2[2]. Multiple trigger chains are considered for each data-taking period, which impose different ID and isolation requirements. The stricter the requirements, the lower the lepton $p_\mathrm{T}$ requirement can be set. If an event passes any trigger chain in Table 4.2 it is recorded and considered in this analysis. The further

[2] "Unprescaled" means that all events passing the chain were recorded. If the rate of a trigger chain is too large, a prescale is commonly applied, which effectively reduces the number of recorded events. "Lowest" refers to the lepton $p_\mathrm{T}$ threshold. Lowering the $p_\mathrm{T}$ threshold increases the rate of a trigger chain, especially caused by soft QCD interactions.




(a) (b)

Figure 4.3: Main decay modes of (a) a $W$ boson and (b) a top-quark pair. In (b), the $x$- and $y$-axes represent the possible decay products of each $W$ boson. The boxes indicate the classes of final-state particles (neglecting neutrinos). The sizes of the boxes are proportional to the probabilities of the final states appearing in $t\bar{t}$ decays. The $W$-boson branching ratios are taken from Reference [28]. The figures have been adapted from Reference [212, 213].

Table 4.2: List of trigger chains used in the analysis. The isolation is based on surrounding tracks with $p_{\mathrm{T}} > 1$ GeV within a fixed radius $\Delta R$, $p_{\mathrm{T}}$-dependent radius $\Delta R^{\mathrm{var}}$ and the longitudinal distance from the primary vertex $\Delta z$. The electron ID likelihood in the years 2016–2018 used no information related to the transverse impact parameter. The L1 $p_{\mathrm{T}}$ requirements are not explicitly listed but the thresholds are generally a bit looser than the HLT thresholds listed here. One electron chain (marked above with "+ L1 requirements") imposed $\eta$-dependent $p_{\mathrm{T}}$ thresholds at L1 and a veto related to the hadronic activity.

| Year | 2015 | 2016–2018 |
|---|---|---|
| Electron | $\geq 1$ $e$, $p_{\mathrm{T}} > 24$ GeV, medium ID + L1 requirements | $\geq 1$ $e$, $p_{\mathrm{T}} > 26$ GeV, tight ID, $\sum_{\Delta R^{\mathrm{var}}<0.2} p_{\mathrm{T}}^{\mathrm{trk}}/p_{\mathrm{T}} < 0.10$ |
| | $\geq 1$ $e$, $p_{\mathrm{T}} > 60$ GeV, medium ID | $\geq 1$ $e$, $p_{\mathrm{T}} > 60$ GeV, medium ID |
| | $\geq 1$ $e$, $p_{\mathrm{T}} > 120$ GeV, loose ID | $\geq 1$ $e$, $p_{\mathrm{T}} > 140$ GeV, loose ID |
| Muon | $\geq 1$ CB muon, $p_{\mathrm{T}} > 20$ GeV, $\sum_{\Delta R<0.3} p_{\mathrm{T}}^{\mathrm{trk}}/p_{\mathrm{T}} < 0.06$ | $\geq 1$ CB muon with $p_{\mathrm{T}} > 26$ GeV and $\sum_{\Delta z<6(2)\mathrm{mm}} p_{\mathrm{T}}^{\mathrm{trk}}/p_{\mathrm{T}} < 0.07$ |
| | $\geq 1$ CB muon, $p_{\mathrm{T}} > 50$ GeV | $\geq 1$ CB muon with $p_{\mathrm{T}} > 50$ GeV |

selection is based on offline reconstructed objects.

Each event is required to contain exactly one electron or muon with $p_{\mathrm{T}} > 27$ GeV which passes the medium identification and isolation working point (loose lepton definition in Table 3.1). Events with an additional very loose lepton (medium ID and no isolation requirement) with $p_{\mathrm{T}} > 10$ GeV are vetoed to reduce backgrounds with multiple prompt leptons in the final state: dileptonic $t\bar{t}$, $Z$+jets, and $VV$. The offline reconstructed lepton has to be geometrically matched ($\Delta R < 0.1$) to the online reconstructed lepton, which fired the trigger. Furthermore, at least four jets have to be present in each event, out of which at least one jet has to be $b$-tagged, i.e. pass the 70 % $b$-tagging WP. The presence of at least four

jets in an event is essential for a proper $t\bar{t}$-system reconstruction. The $b$-tag requirement substantially reduces backgrounds from multijet and $V$+jets processes and is beneficial for the $t\bar{t}$-system reconstruction. All selection requirements listed above define the preselection.

The signal region is defined by the lepton satisfying the tight lepton definition (tight ID and isolation) and the number of $b$-tagged jets being two or larger. The data in the signal region is a subset of the preselection data. Events passing the preselection but not the signal region requirements are used to estimate the multijet background in the signal region (cf. Section 4.3). In total, a bit more than 9.8 million data events pass the preselection requirements, and 3.6 million events are in the signal region. All selection steps are detailed in Table 4.3.

Table 4.3: Cutflow table for a few selected signal samples. The first lepton in $e/\mu$ and $\mu/e$ is the lepton used in the electron-channel selection, and the second lepton is the one used in the muon-channel selection. In the signal samples the $W$ boson is forced to decay leptonically. The cuts above the dashed line define the preselection region; all cuts together define the signal region.

| Selection steps | Fraction of events passing selection [%] | | |
|---|---|---|---|
| | $H^{\pm}_{80}$ | $H^{\pm}_{130}$ | $H^{\pm}_{160}$ |
| Initial | 100.0 | 100.0 | 100.0 |
| Trigger | 37.8 | 37.5 | 37.5 |
| At least 1 medium ID $e/\mu$ with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ | 35.3 | 35.2 | 35.0 |
| Exactly 1 medium ID $e/\mu$ with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ | 34.8 | 34.7 | 34.7 |
| No medium ID $\mu/e$ with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ | 33.8 | 33.7 | 33.9 |
| Exactly 1 medium ID $e/\mu$ with $p_{\mathrm{T}} > 27\,\mathrm{GeV}$ | 32.9 | 32.7 | 33.0 |
| Lepton matches trigger lepton | 32.6 | 32.5 | 32.8 |
| Jet cleaning | 32.5 | 32.4 | 32.7 |
| At least 1 jet with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ | 32.4 | 32.3 | 32.6 |
| At least 2 jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ | 31.7 | 31.5 | 31.5 |
| At least 3 jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ | 28.0 | 27.8 | 26.5 |
| At least 4 jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ | 19.1 | 19.2 | 15.8 |
| At least 1 $b$-tagged jet | 16.6 | 15.8 | 11.1 |
| Lepton passes medium isolation | 15.5 | 14.7 | 10.3 |
| Lepton passes tight ID | 14.9 | 14.2 | 9.9 |
| Lepton passes tight isolation | 13.6 | 12.9 | 9.1 |
| At least 2 $b$-tagged jets | 6.6 | 5.4 | 1.7 |

The selection efficiencies of three different signal samples after each subsequent selection step are listed in Table 4.3. To keep in mind is that in the signal sample production the $W$ boson is forced to decay leptonically. The selection efficiency of semileptonic $t\bar{t}$ events is similar to the efficiency of the $H^{\pm}_{80}$ signal sample.

A large fraction of signal events is rejected at the trigger step. The main reasons are the lepton $p_{\mathrm{T}}$ thresholds and the $|\eta| < 2.5$ requirement. This loss is inevitable for this analysis, as the lepton $p_{\mathrm{T}}$ thresholds are driven by the large backgrounds at low $p_{\mathrm{T}}$ – feasibility of trigger rates – and the $\eta$ requirement is determined by the tracking volume. Only a few events are rejected by requiring tighter offline lepton identification and isolation requirements.

Another requirement which rejects a significant fraction of signal events is the requirement of a fourth

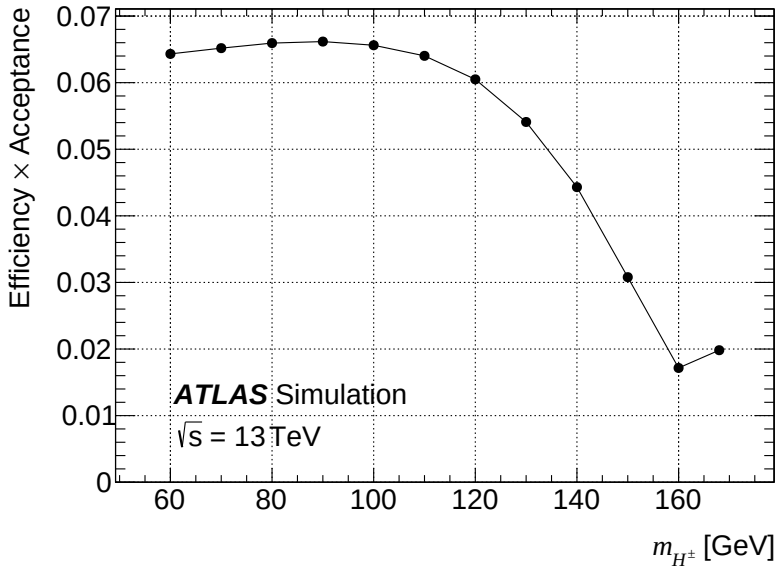



Figure 4.4: Signal selection efficiency times acceptance of all $H^\pm \to cs$ signal samples.

jet. If no additional radiation is present in the event, only a single jet needs to fail the $p_\mathrm{T}$ or $\eta$ requirement for the whole event to be rejected. In a future iteration of the analysis, a lowering of the jet $p_\mathrm{T}$ threshold to 20 GeV, which is the lower bound of jet flavour-tagging calibration, might be worthwhile to investigate. The increase in signal efficiency would be at the cost of increased backgrounds and therefore has to be studied carefully. My proposal for future analyses is to still require three jets with $p_\mathrm{T} > 25$ GeV but the fourth jet needs to have $p_\mathrm{T} > 20$ GeV. This can recover especially signal events if the charged Higgs boson mass is close to the top-quark mass. Because the two jets emerging from the charged Higgs boson have on average a large $p_\mathrm{T}$, but the $b$-jet produced together with the charged Higgs boson has on average a very small $p_\mathrm{T}$.

The last cut that reduces the size of the selected signal sample significantly is the requirement of a second $b$-tagged jet. This is the case because both $b$-quark-initiated jets have to pass the 70 % WP, which has a probability of 49 % – in contrast, the probability of just one of the two $b$-quark-initiated jets being $b$-tagged is 91 %. And both $b$-quark-initiated jets have to pass the initial $p_\mathrm{T}$ and $\eta$ requirements. The signal sample size reduction is especially severe for samples with charged Higgs boson masses close to the top-quark mass due to the previously discussed small momentum of the $b$-jet produced together with the charged Higgs boson.

Figure 4.4 shows the final selection efficiencies of all signal samples. The selection efficiency is between 1.7 and 6.6 %. The drop at larger charged Higgs boson masses is caused by the previously discussed effects, which can also be inferred from Table 4.3. Unexpectedly, the efficiency doesn't drop for the last mass point. This has been traced down to be caused by a sizeable number of off-shell top quarks (with $m_t > 172.5 + 2\Gamma_t$) in this sample, which is not observed for other signal samples. This is caused by the small mass difference $m_t - (m_{H^\pm} + m_b) \approx 0$ and thereby small available phase space.

No selection is enforced regarding the neutrino because 99 % of the backgrounds include a $W$ boson

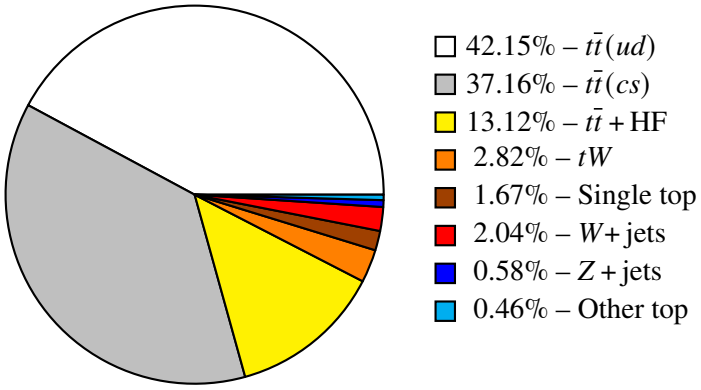



Figure 4.5: Composition of the simulated backgrounds in the signal region. The samples are defined in Table 4.1. The $t\bar{t}$(allHad) background is merged into the "Other top" category; the $VV$ background is merged into the "$Z$ + jets" category.

that produces a high-energy neutrino in the same way as in the signal process. The distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ is almost identical for signal and background processes.

In general, the event selection is rather loose in order to maintain a high signal efficiency. Nonetheless, the signal region is relatively clean in $t\bar{t}$ events. A discrimination of signal and $t\bar{t}$ events is done using a more sophisticated method than just using simple cuts (see Section 5).

Figure 4.5 shows the composition of simulated backgrounds in the signal region. The dominant background with 92.5 % is $t\bar{t}$ production. Single-top-quark and $W$+ jets processes contribute 4.5 % and 2 %, respectively. The contributions from other backgrounds are about 1 %.

Fully hadronic $t\bar{t}$ events are rejected successfully by the lepton requirement, and its contribution is insignificant in the signal region. The fraction of dileptonic events in the $t\bar{t}(ud)$ ($t\bar{t}$ + LF) and $t\bar{t}$ + HF categories is about 17 % (9 %) and 13 %, respectively. When comparing the dileptonic fractions with the $W$-boson branching ratios in Figure 4.3, it becomes clear that more dileptonic than semileptonic $t\bar{t}$ events got rejected by event selection, as was intended.

## 4.3 Modelling of background events

The simulation of the signal process and background processes with a prompt lepton has been presented in Section 4.1. In order to get a complete background expectation, the multijet background is estimated using a data-driven method. The procedure is described and validated in Section 4.3.1. The modelling of simulated background events is then evaluated by comparing the background expectation to data in regions that are not sensitive to a potential signal. Discrepancies between expectation and data are corrected for as described in Section 4.3.2. Validation studies on the modelling of simulated signal events were already presented in Section 4.1.1.

### 4.3.1 Estimation of the multijet background in the signal region

Multijet (MJ) processes don't involve a prompt lepton. Hence, MJ events can only pass the event selection when either a jet fakes a prompt lepton or when a real non-prompt lepton is produced in the

decay of heavy-flavour hadrons. The lepton identification and isolation requirements are designed to reject these types of lepton candidates, and misidentification rates are accordingly small. But due to the overall large MJ cross-section, it is a non-negligible background in this analysis.

A previous $t\bar{t}$ analysis with a similar signal region definition estimated a MJ contribution of around 3 % [217]. The MJ contribution in this analysis is expected to be less due to improvements in $b$-tagging and the non-prompt lepton suppression.

It is difficult to obtain a good MJ estimation from simulation because of the small misidentification rates. It necessitates the simulation of a huge number of events and only a tiny phase space is of relevance, which then has to be modelled well. Available dijet samples simulated by Pythia have been tested for usage in the analysis. Out of many millions of events, only a handful of events passed the signal region selection in the lowest jet $p_{\mathrm{T}}$-slices. This makes a proper MJ estimation from simulation impossible, and a data-driven method has been developed instead.

In LHC Run-2 about $1.1 \times 10^{16}$ [133] proton-proton collisions took place at ATLAS. Most of them are soft QCD interactions. And the lepton triggers selects the MJ processes which actually fake a lepton in the detector. Considering this, the MJ background doesn't have to be simulated but can be extracted from data. The challenging part is the identification of MJ events in data. This is not possible on an event-by-event basis but is done statistically by exploiting the differences between MJ processes and processes with prompt leptons.

#### Methodology

The method used in this analysis to estimate the MJ background is commonly known as the *fake-factor* or *ABCD* method [218, 219]. The concept of the ABCD method is used because it is well established within high-energy particle physics and the concept is simple and easy to apply. In the ABCD method the data is divided into four regions, labelled $A$, $B$, $C$, and $D$. Region A is the signal region as defined in Section 4.2. The other regions are enriched in MJ events. The aim is to estimate the MJ contribution in region A with the help of the other regions. Region B encompasses events where the lepton passes the PLIV-Tight but not the PLIV-VeryTight isolation requirement (Loose-Isol lepton definition in Table 3.1). Events in region C are required to have exactly one $b$-tagged jet. In region D the looser lepton isolation requirement and the jet requirement of one $b$-tagged are applied.

The number of $b$-tags is well suited to define MJ enriched regions, as MJ processes mostly involve light jets. The lepton isolation is well suited, too, as non-prompt leptons stemming from heavy-flavour hadron decays will be close to the jet the hadron is part of. Similarly, tracks from MJ processes faking leptons will be close to the jet remnants.

Figure 4.6 summarises the region definitions and illustrates the functioning principle of the ABCD method. The MJ background in the signal region is estimated by extracting the shape of the MJ background in region B and then rescaling it to the correct normalisation in the signal region with the help of regions C and D. The MJ background in regions B, C and D is obtained by subtracting the MC predictions from the data in the respective region,

$$\mathrm{MJ}_{x,\ell} = \mathrm{Data}_{x,\ell} - \mathrm{MC}_{x,\ell} \qquad x \in B, C, D \quad \ell \in e, \mu.$$

With all other backgrounds being estimated from simulation, the difference between data and prediction

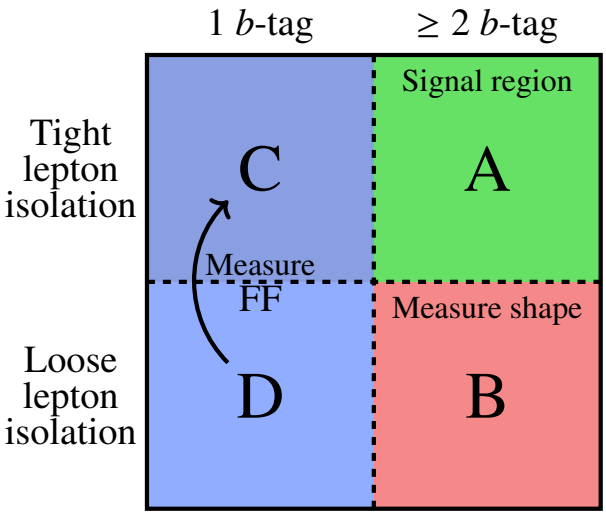



Figure 4.6: Functioning of the ABCD method and definition of the regions.

is expected to be caused by the missing MJ background. The MJ estimation is performed separately for the two lepton channels, as the signatures of electrons and muons are different in the detector. The separate estimation accounts for different probabilities for jets to fake an electron or muon.

Figure 4.7 shows the difference between data and MC – the MJ contribution – in the different regions. Plotted is the transverse mass variable $m_{\mathrm{T,W}}$, as it is sensitive to a possible MJ background. The variable $m_{\mathrm{T,W}}$ is defined as

$$m_{\mathrm{T,W}} = \sqrt{2 p_{\mathrm{T}}^{\ell} E_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos\phi)},$$

with $p_{\mathrm{T}}^{\ell}$ being the transverse momentum of the lepton and $\phi$ the azimuthal opening angle between the lepton and the missing transverse momentum. MJ events receive on average smaller $m_{\mathrm{T,W}}$ values because they do not contain an on-shell $W$ boson. In regions B and D the difference between data and MC prediction is larger than the total uncertainty on the prediction. This clearly shows that the MJ enrichment is statistically significant and underpins a solid region definition. In region B a larger MJ enrichment than in region C is favourable because the shape will be derived from region B. In regions C and D, only the overall normalisation is of interest.

The two core assumptions of the ABCD method are on the one hand that the shapes of the MJ background are the same in regions B and A, and on the other hand that the MJ ratio in regions A to B equals the ratio in regions C to D,

$$\frac{\mathrm{MJ}_{A,\ell}}{\mathrm{MJ}_{B,\ell}} = \frac{\mathrm{MJ}_{C,\ell}}{\mathrm{MJ}_{D,\ell}} \quad \Longrightarrow \quad \mathrm{MJ}_{A,\ell} = \underbrace{\frac{\mathrm{MJ}_{C,\ell}}{\mathrm{MJ}_{D,\ell}}}_{\mathrm{FF}} \mathrm{MJ}_{B,\ell}. \tag{4.1}$$

A premise for the assumptions to hold is that the regions – the two quantities that are used to split the data into the four regions – are uncorrelated. Also shown in Equation 4.1 is that by rearranging the terms, the MJ contribution in region A can be estimated with the help of the other regions. The MJ estimation

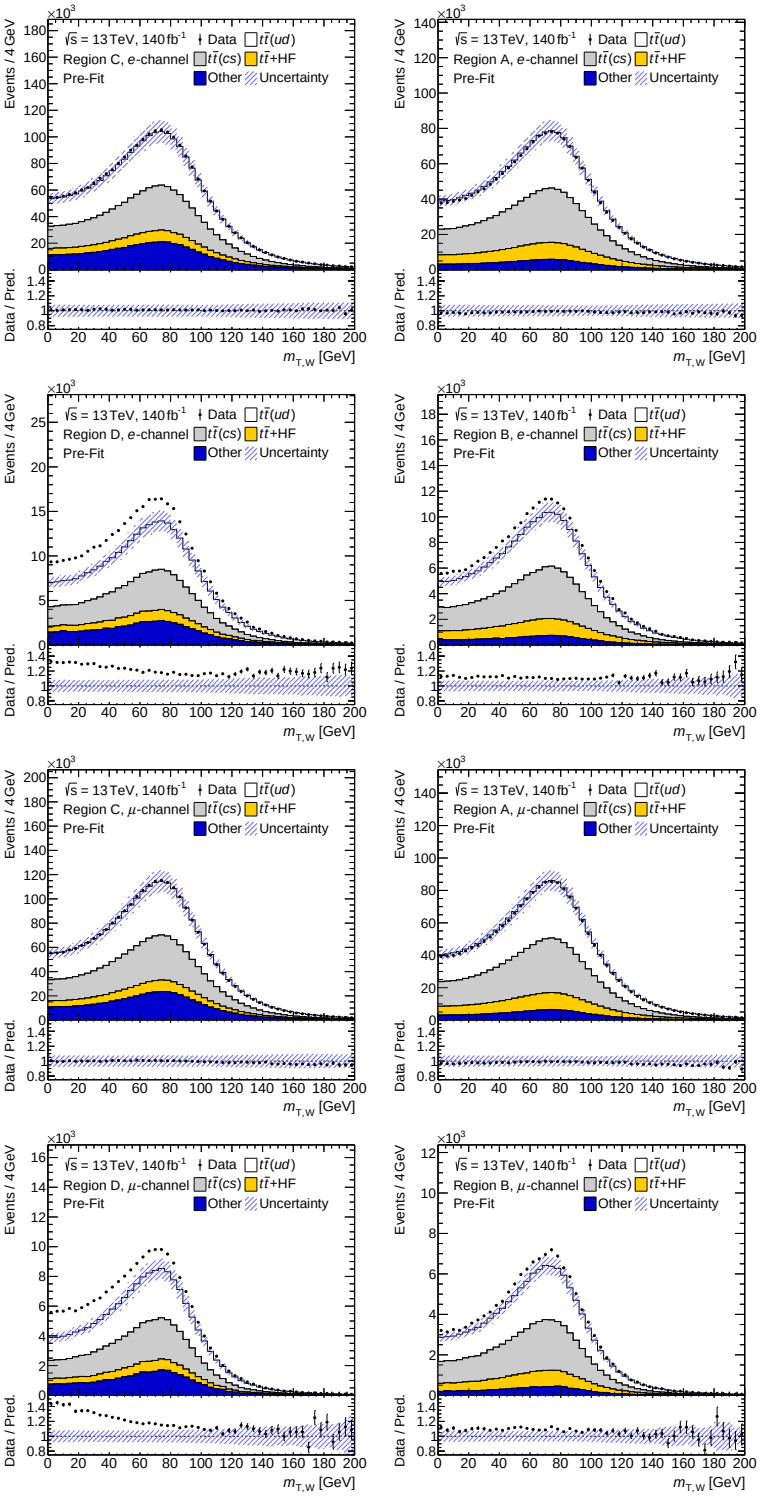



Figure 4.7: Distributions of $m_{\mathrm{T,W}}$ in the four regions used in the ABCD method. The plots are arranged in the same way as the regions are arranged in Figure 4.6. The top (bottom) four plots show the distributions in the $e(\mu)$-channel. The uncertainty band represents the total statistical plus systematic uncertainty.

in region A is the MJ template from region B multiplied by the ratio $\mathrm{MJ}_{C,\ell}/\mathrm{MJ}_{D,\ell}$, which is commonly referred to as the *fake factor* (FF). The FF corrects for the different normalisations in regions A and B. A single FF is used in the analysis because no FF dependence on the lepton kinematics like $p_{\mathrm{T}}$ and $\eta$ has been observed. The derived FFs and their statistical uncertainties are

$$\mathrm{FF}_e = 0.417 \pm 0.027,$$
$$\mathrm{FF}_\mu = -0.04 \pm 0.05.$$

$\mathrm{FF}_\mu$ is negative because in the $\mu$-channel the MC prediction exceeds the number of data events in region C, within uncertainties. The used lepton isolation requirements have been specifically developed to reject non-prompt leptons, and the probability of jet particles reaching the MS and faking a muon is tiny. Therefore, a small to negligible MJ contribution in the $\mu$-channel is expected. In contrast, the pattern of an electron in the detector is much more likely to be faked by jets. This expectation is also reflected in the fake-factor which differs from zero by multiple standard deviations. As a negative MJ contribution is not physical and the fake-factor in the $\mu$-channel agrees within uncertainties with zero, it is set to zero, $\mathrm{FF}_\mu = 0$. This results in a systematic upward shift of the MJ contribution, but this shift is overall small in comparison to the $e$-channel contribution ($\mathrm{MJ}_{B,e} > \mathrm{MJ}_{B,\mu}$ and $\mathrm{FF}_e \gg |\mathrm{FF}_\mu|$) and will be well covered by a dedicated systematic uncertainty in this analysis (cf. Section 6.1).

Similarly, in region B the MC prediction can exceed data for single bins, which results in bins with negative contents for the MJ templates. Since this is unphysical, the content of such bins is set to zero. It has been checked that in such cases the bin content generally agrees with zero within statistical uncertainties.

#### Results

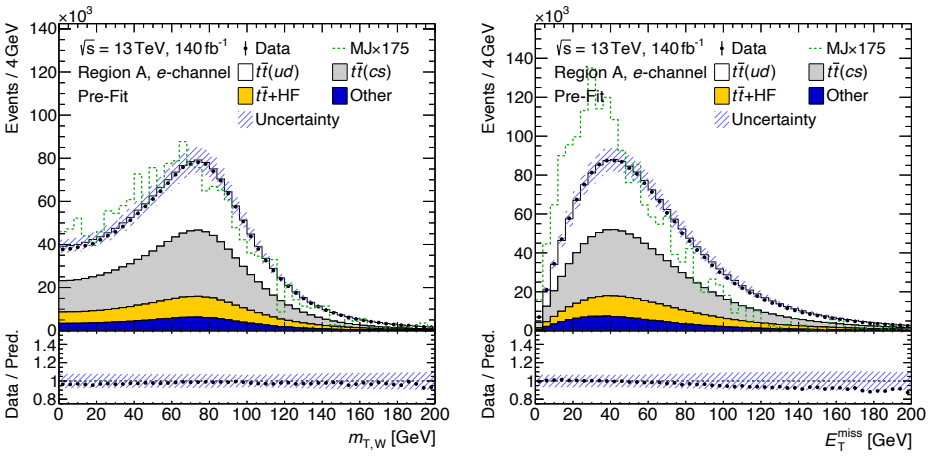
Figure 4.8 shows the $m_{\mathrm{T,W}}$ distribution in the signal region for the $e$-channel, including the estimated MJ background. The distributions of the MJ background are, as expected, more populated at smaller $m_{\mathrm{T,W}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ values compared to the $t\bar{t}$ background because MJ events do not contain an on-shell $W$ boson.

The size of the MJ background in the signal region is about 0.3 % of the total simulated background. This is a rather small contribution and was expected, given the good agreement of data and prediction in the signal region without the MJ background in Figure 4.7. This MJ contribution is even smaller than the overall background normalisation uncertainty from multiple $t\bar{t}$-modelling-related uncertainties. Therefore, the MJ background plays a subordinate role in this analysis. The subsequent analysis parts are performed solely in region A, the signal region.

In a previous stage of the analysis, a looser $b$-tagging requirement and a different lepton isolation requirement were used to define the signal region. At that stage the ABCD method yielded a MJ contribution of about 1.3 % of the total predicted background. Hence, the tighter $b$-tagging requirement and particularly the new lepton isolation requirement could reduce the MJ background by around a factor of 4.

A more recent $t\bar{t}$ analysis [220] applies an almost identical event selection as has been applied at the previous analysis stage. In the referenced analysis, the matrix method[3] is used to estimate the

[3] In the matrix method the fake/non-prompt lepton background in the signal (tight) region is estimated by solving a system of linear equations that use measured efficiencies for real and fake leptons passing loose and tight selection criteria. [154]




(a) (b)

Figure 4.8: Distribution of (a) $m_{\mathrm{T,W}}$ and (b) $E_{\mathrm{T}}^{\mathrm{miss}}$ in the $e$-channel in the signal region with the MJ background included. Due to its small contribution, the MJ background is barely visible in the stack plot. Therefore, it is merged in the "Other" category, and its shape, scaled to the total background prediction, is overlaid.

MJ background and yields an overall MJ contribution of around 1.3 %, too. This proves that the MJ estimation in this analysis yields reliable results.

### Validation studies

The MJ estimation has been validated in multiple ways. Firstly, it has been tested whether the general assumption of the ABCD method from Equation 4.1 is valid with the help of the $t\bar{t}$(allHad) sample. The $t\bar{t}$(allHad) sample is used because $t\bar{t}$(allHad) and MJ events passing the preselection only contain fake and non-prompt leptons. It has been verified that the ratios $t\bar{t}(\mathrm{allHad})_{A,\ell}/t\bar{t}(\mathrm{allHad})_{B,\ell}$ and $t\bar{t}(\mathrm{allHad})_{C,\ell}/t\bar{t}(\mathrm{allHad})_{D,\ell}$ agree with each other. The ratios agree within $2\sigma$ in both lepton channels, considering only the statistical uncertainty. This confirms that the regions are uncorrelated and the FFs yield the correct normalisation.

Another way to prove that the regions are uncorrelated is by performing an alternative MJ estimation in which the requirements defining regions B and C are swapped. The 1$b$-tag requirement defines region B, and the lepton isolation requirement defines region C. The resulting MJ contribution indeed agrees with the nominal approach within 10 %.

An alternative MJ estimation has been performed using a looser lepton ID criterion instead of a looser lepton isolation requirement. The lepton ID is designed to reject fake leptons and is therefore also suited to define MJ-enriched regions. Region B is defined by the lepton passing the Medium ID WP while not passing the tight one (loose-ID lepton definition in Table 3.1). Region C is defined as in the nominal approach using the one $b$-tag requirement, and in region D both requirements are enforced. The estimated MJ contribution using this approach is about 30 % less than in the nominal approach. This alternative estimate is considered a systematic uncertainty of the MJ background in this analysis.

The lepton ID and isolation requirements are not used together to define the four ABCD regions, as those variables seem to be correlated. When swapping the two requirements to define regions B and C – the same validation approach as presented above – the estimated MJ contribution differed by a factor of

2. The variable $E_{\mathrm{T}}^{\mathrm{miss}}$has not been used either to define the regions because of insufficient modelling, which will be discussed in the following section.

### 4.3.2 $t\bar{t}$ modelling correction

In Figure 4.9 a few selected kinematic distributions are presented that show a significant disagreement between data and prediction in the signal region. The deviations are observed in the tails of the distributions, are partially larger than about 10 %, and exceed the total uncertainty on the prediction. The slopes in the ratio graphs indicate a mismodelling of the predictions rather than the disagreement just being caused by statistical fluctuations. With the signal region being dominated by $t\bar{t}$ events (contribution of 92.5 %), the deviation is apparently caused by an inaccurate simulation of $t\bar{t}$ processes. The simulated top-quark $p_{\mathrm{T}}$ seems to be harder in simulation than in data. Indeed, similar mismodellings have been seen in other analyses with a $t\bar{t}$-pair in the final state [91, 217, 220]. The disagreement can be attributed to missing higher-order QCD and EW corrections in $t\bar{t}$ simulation [221, 222]. This affects SM $t\bar{t}$ and signal processes in the same way because the same event generator setup has been used to simulate the events. A data-driven correction is derived to improve the modelling of $t\bar{t}$ simulation. The correction aims at improving the modelling of $p_{\mathrm{T}}$-dependent variables.

#### Methodology

The scalar momentum sum ($S_{\mathrm{T}}$) is defined as the sum of the scalar transverse momenta of all calibrated objects in the event,

$$S_{\mathrm{T}} = p_{\mathrm{T},\ell} + E_{\mathrm{T}}^{\mathrm{miss}} + \sum_{i=1}^{N_{\mathrm{jets}}} p_{\mathrm{T,i}}, \tag{4.2}$$

and is used to correct the simulation. The $S_{\mathrm{T}}$ distribution is shown in Figure 4.9(e). The choice of $S_{\mathrm{T}}$ is motivated by the fact that $S_{\mathrm{T}}$ describes the hardness of an event and is closely related to the transverse momenta of individual top quarks. The corrections are derived in bins of the number of jets in the event ($N_{\mathrm{jets}}$). Events with 9 or more jets are grouped together due to its small selected sample size, resulting in 6 bins: $N_{\mathrm{jets}} = 4, 5, 6, 7, 8, \geq 9$.

A two-dimensional correction is deployed because the two variables are closely correlated, but their mismodelling is in contradiction to their correlation. In general, an event with more jets will have a larger $S_{\mathrm{T}}$. But the simulation prefers more often a smaller value of $N_{\mathrm{jets}}$ and a larger value of $S_{\mathrm{T}}$. Therefore, a simple one-dimensional correction in $S_{\mathrm{T}}$ does not improve the modelling of $N_{\mathrm{jets}}$.

The correction is derived directly in the signal region. The two main reasons behind this choice are that the signal region definition is designed to select semileptonic $t\bar{t}$ events and that the signal events are affected by the mismodelling, too. It has been carefully validated that this choice does not impact the extraction of a potential signal – this is presented in detail later on in this section.

The correction weights are defined by the ratio of data templates, with non-$t\bar{t}$ backgrounds subtracted, to simulated $t\bar{t}$ templates,

$$w_{\mathrm{corr}} = \frac{\mathrm{Data} - \mathrm{MC}_{\text{non-}t\bar{t}}}{\mathrm{MC}_{t\bar{t}}}. \tag{4.3}$$

The data-driven MJ background is part of the $\mathrm{MC}_{\text{non-}t\bar{t}}$ category. The numerator represents an estimate of the $t\bar{t}$ events in data (modified data template). It is assumed that the distributions of $N_{\mathrm{jets}}$ and $S_{\mathrm{T}}$ are

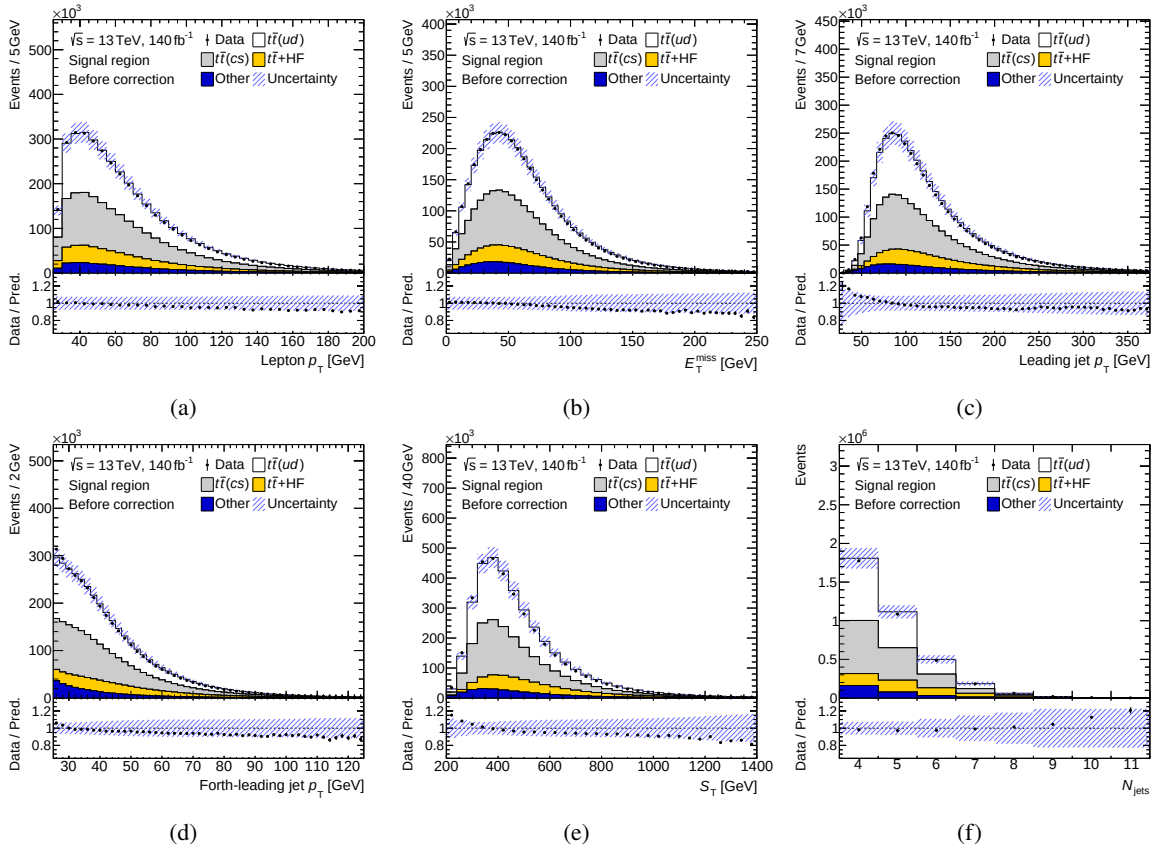



Figure 4.9: Distribution of the (a) lepton $p_{\mathrm{T}}$, (b) $E_{\mathrm{T}}^{\mathrm{miss}}$, (c) leading jet $p_{\mathrm{T}}$, (d) forth-leading jet $p_{\mathrm{T}}$, (e) $S_{\mathrm{T}}$ (definition in Equation 4.2), and (f) the number of jets in an event. The uncertainty band represents the combined statistical plus systematic uncertainty of the prediction.

modelled properly for the non-$t\bar{t}$ backgrounds. The binning of the templates is chosen to be equidistant from 200 to 2 000 GeV with a bin width of 20 GeV. Below 200 GeV and above 2 000 GeV, the sample size is small. The bin width is chosen sufficiently small to resolve the shape. Bins in the tails with low sample size are merged, such that each modified data template bin contains at least ten expected events. The $w_{\mathrm{corr}}$ templates for a few selected $N_{\mathrm{jets}}$ bins are shown in Figure 4.10.

The final correction weights are extracted by fitting functions to the templates to mitigate the effects of statistical fluctuations in data and simulation in the tails. No underlying model exists that describes the disagreement between data and $t\bar{t}$ simulation. Hence, a function that fits the distribution well is chosen. Among the tested functions, a linear-plus-exponential function,

$$f_{\mathrm{corr}}(x, S_{\mathrm{T}}) = x_0 + x_1 S_{\mathrm{T}} + x_2 e^{x_3 S_{\mathrm{T}}}, \tag{4.4}$$

is found to yield the smallest $\chi^2$ values in fits to the $w_{\mathrm{corr}}$ templates across all $N_{\mathrm{jets}}$ bins. The exponential part is dominant in the low $S_{\mathrm{T}}$ region and flattens out towards larger $S_{\mathrm{T}}$ values. The medium and high $S_{\mathrm{T}}$ regions are described solely by the linear part.

To avoid overfitting, the fit is performed separately for even and odd event numbers[4]. Splitting the events by their event number achieves a random split. The correction weights derived from fits to even-numbered events are applied to odd-numbered events and vice versa.

#### Multijet background adjustment

One important remark has to be made on the MJ background. The MJ estimation depends on the $t\bar{t}$ background modelling in regions B, C, and D. The phase space in these regions is similar to the signal region. For this reason it is expected that the $t\bar{t}$ background is similarly mismodelled in these regions. This assumption has been confirmed by producing $w_{\text{corr}}$ in the respective regions. The aim is to consider this effect in the $t\bar{t}$ correction procedure.

The $t\bar{t}$ correction leaves the normalisation in the signal region almost unchanged – the normalisation changes by about $\approx 2\,\%$ – whereas the correction weights are in the range of 0.7 to 1.3. As only the normalisation is of interest in regions C and D, the FFs are assumed to be unaffected by the mismodelling. This assumption simplifies the weight calculation and fitting procedure immensely, as it removes bin-by-bin correlations. The mismodelling in region B and the signal region is assumed to be identical given the similar phase space, which differs only by a looser lepton isolation requirement. With these assumptions the $t\bar{t}$ mismodelling in region B can be directly considered in the weight calculation by replacing the MJ background in region A in Equation 4.3 by its templates in region B,

$$w_{\text{corr}} = \frac{\text{Data}_A - \text{MC}_{\text{non-}t\bar{t},A} - \text{MJ}_A}{\text{MC}_{t\bar{t},A}} = \frac{\left(\text{Data}_A - \text{FF} \times \text{Data}_B\right) - \left(\text{MC}_{\text{non-}t\bar{t},A} - \text{FF} \times \text{MC}_{\text{non-}t\bar{t},B}\right)}{\text{MC}_{t\bar{t},A} - \text{FF} \times \text{MC}_{t\bar{t},B}}, \tag{4.5}$$

$$\text{with} \quad \text{MJ}_A = \text{FF}\frac{\text{Data}_B - \text{MC}_{\text{non-}t\bar{t},B}}{w_{\text{corr}} \times \text{MC}_{t\bar{t},B}}.$$

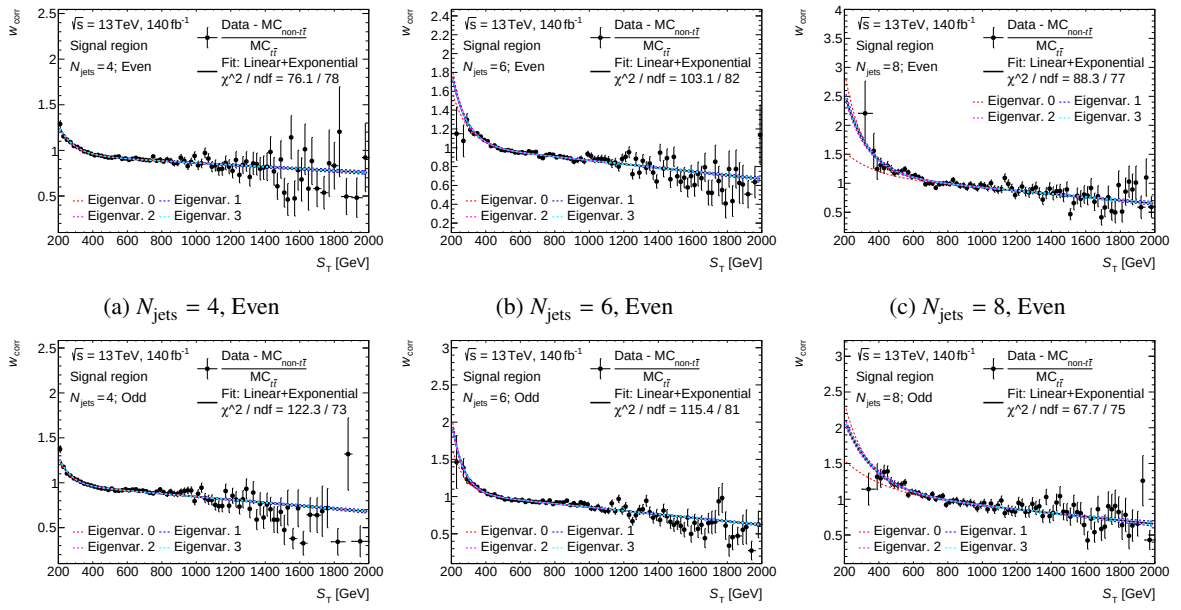
Effectively, the $t\bar{t}$ mismodelling in region B, which biased the $w_{\text{corr}}$ calculation through the MJ estimation, is now correctly accounted for. The $t\bar{t}$ correction weights and MJ background are derived simultaneously. The $w_{\text{corr}}$ templates shown in Figure 4.10 are calculated as given in Equation 4.5.

The application of the $t\bar{t}$ correction in region B changes the estimated MJ background contribution and shape with respect to the results presented in Section 4.3.1. Plots comparing the MJ background before and after the $t\bar{t}$-correction has been applied are available in Appendix A.3. Changes in the MJ background estimation up to 40 % are seen for some high statistics bins.

#### Results

The derived $w_{\text{corr}}$ templates plus fitted functions are shown in Figure 4.10 for the bins with $N_{\text{jets}} = 4, 6, 8$. The fitted $w_{\text{corr}}$ templates for other $N_{\text{jets}}$ bins and a list of all fit parameters are given in Appendix A.1. Shown as dotted lines are variations of the eigenvectors from the fit uncertainty matrix by their eigenvalue $\pm 1\sigma$. Data and MC statistical uncertainties are included in the fit. The fit uncertainties are relatively small due to the overall large sample size. An exception is the region of high $N_{\text{jets}}$ and low $S_{\text{T}}$, as only a few events are available to determine the exponential parameters. The reduced $\chi^2$, defined by the $\chi^2$ value of the fit divided by the number of degrees of freedom in the fit, $\chi^2/\text{ndf}$, of all fits are between 0.5

[4]Every recorded data event has a unique event number. Similarly, event numbers are allocated to simulated events.




(a) $N_{\text{jets}} = 4$, Even (b) $N_{\text{jets}} = 6$, Even (c) $N_{\text{jets}} = 8$, Even

(d) $N_{\text{jets}} = 4$, Odd (e) $N_{\text{jets}} = 6$, Odd (f) $N_{\text{jets}} = 8$, Odd

Figure 4.10: Distribution of $w_{\text{corr}}$ in different $N_{\text{jets}}$bins for even and odd event numbers. The solid black lines are fits of a linear-plus-exponential function to the distributions. The dotted lines show the $\pm 1\sigma$ variations of the eigenvectors of the uncertainty matrix. The uncertainties include data and MC statistical uncertainties.

and 2. This indicates a solid choice of the fit function and generally a good agreement between fit and data. The fits agree well for even and odd parities.

The same quantities as in Figure 4.9 are shown again in Figure 4.11, but this time with the corrections applied. The modelling of all shown variables and other affected kinematic variables is significantly improved, as can be seen when comparing to the red dashed line, which shows the total background prediction before applying the correction. In the subsequent analysis steps the $t\bar{t}$ correction weights have been applied to SM $t\bar{t}$ and signal samples.

The modelling of the top-quark $p_{\text{T}}$-spectrum depends significantly on the assumed event generator model. Therefore, separate corrections are derived and applied to the alternative $t\bar{t}$ samples. The procedure is identical to the nominal approach. Exemplary fits for the alternative sample are shown in Figure A.2 in the Appendix.

### Validation studies

Multiple functions have been tested for fitting the $w_{\text{corr}}$ templates. Thereby the linear-plus-exponential function (cf. Equation 4.4) showed the best overall agreement with the fitted data and reduced $\chi^2$ values. A complete review of all tested functions will not be presented here due to the sheer number of tested functions, but a few functions and the idea behind their choice will be introduced.

Polynomials of different orders have been tested because every function can be dismantled in polynomials, and lower-order polynomials can already provide good approximations. Fits with

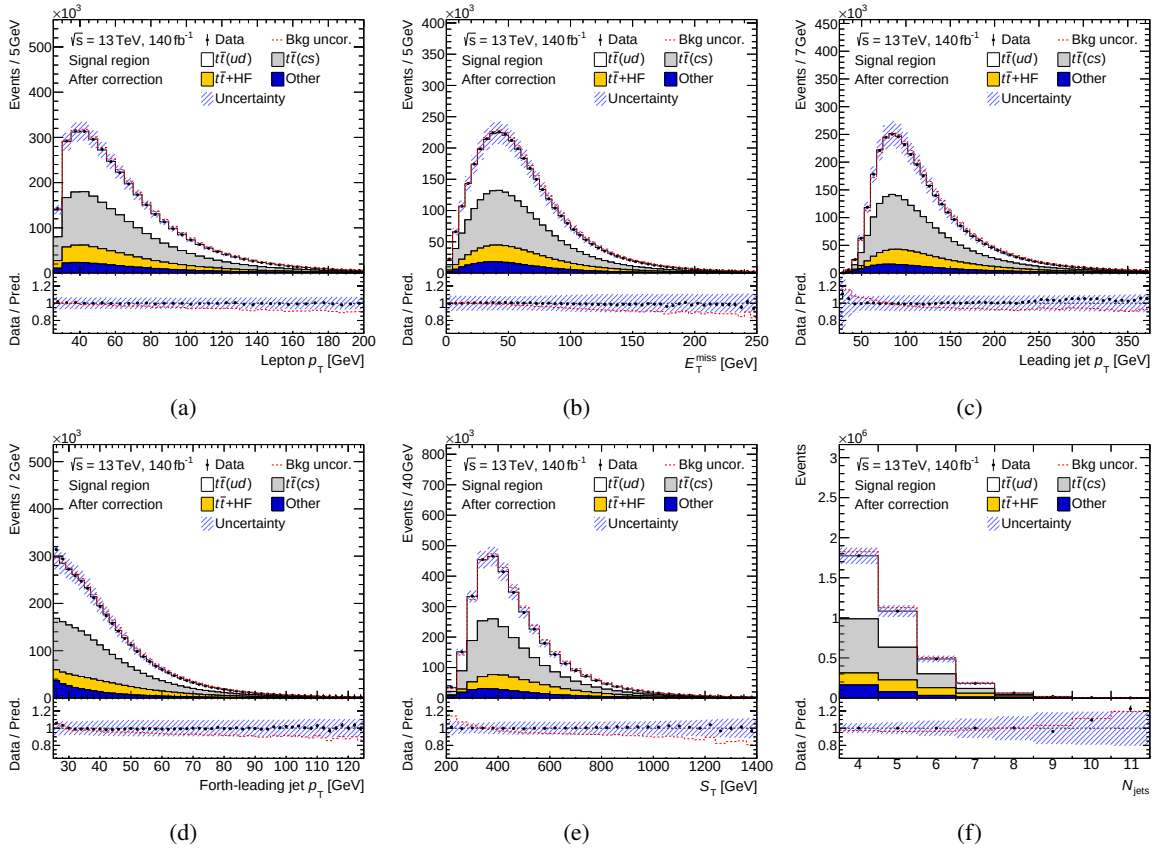
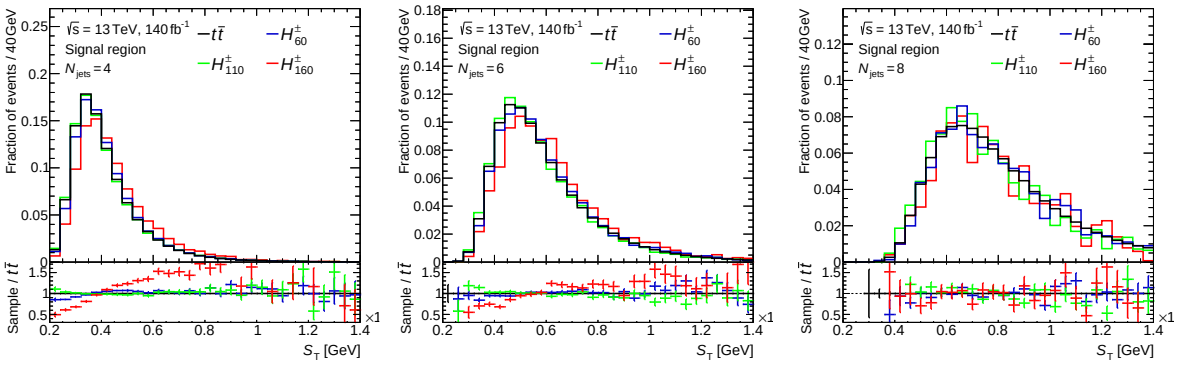



Figure 4.11: Distribution of the (a) lepton $p_{\mathrm{T}}$, (b) $E_{\mathrm{T}}^{\mathrm{miss}}$, (c) leading jet $p_{\mathrm{T}}$, (d) forth-leading jet $p_{\mathrm{T}}$, (e) $S_{\mathrm{T}}$ (definition in Equation 4.2), and (f) the number of jets in an event, after the $t\bar{t}$ correction has been applied. The uncertainty band represents the combined statistical plus systematic uncertainty of the prediction. The red dashed line represents the total background prediction before applying the correction.

polynomials up to the order of eight showed worse agreement than in the nominal fits. Especially the lowest-order polynomials showed bad agreement between fit and data points. The higher-order polynomials often fit statistical fluctuations due to the large number of degrees of freedom and still yielded larger $\chi^2$ values than in the nominal fit. The initial fit parameters have to be tweaked very well in these fits to achieve satisfactory results. In fits with even higher-order polynomials than of order eight, convergence issues showed up.

Additionally, extensions of the nominal function have been tested that might lead to an improved description due to the increased fit function complexity. It has been tested whether using a higher-order polynomial (quadratic or cubic instead of linear) could improve the agreement in the high $S_{\mathrm{T}}$ part. The low $S_{\mathrm{T}}$ part was tried to be improved by the addition of a second exponential function. In case the steep rise at small $S_{\mathrm{T}}$ values is not exponential but translates more quickly into something else, it might be better described by two exponentials. The $\chi^2$ in these fits didn't improve significantly, yielding slightly worse reduced $\chi^2$ values with respect to the nominal fits.

As a final conclusion, the linear-plus-exponential function provided the best reduced $\chi^2$ values without a clear indication that a simpler function can describe the templates as well or that a more complex




(a) (b) (c)

Figure 4.12: Comparison of the $S_T$ distribution of three signal mass points with semileptonic SM $t\bar{t}$ events ($t\bar{t}(ud)$ + $t\bar{t}(cs)$ + $t\bar{t}$ + HF) for events with (a) 4, (b) 6, and (c) 8 jets.

function can improve the agreement.

Another part of the $t\bar{t}$ correction that has been validated thoroughly was the usage of the signal region to derive the correction weights. When deriving the correction in the signal region, care has to be taken that no artefacts are introduced by the correction procedure. This means the estimated signal contribution has to be unchanged; no signal was artificially added or removed.

It has been verified that the distributions of $S_T$ in the outlined $N_{\text{jets}}$ bins are similar for SM $t\bar{t}$ and signal events. The comparison is plotted in Figure 4.12. For the highest mass points (150 to 168 GeV), it can be seen that the $S_T$ distribution is slightly harder than for SM $t\bar{t}$ events. This difference can be attributed to a bias introduced by the event selection. With increasing charged Higgs boson mass the momentum transferred to the $b$-quark that is produced in conjunction with the charged Higgs boson decreases. As a result, mostly events in which the respective top quark has a large momentum pass the event selection.

The small differences in $S_T$ together with the fact that signal and background events are affected in the same way by the mismodelling ensure that the correction is unaffected by a possible signal in the data and the correction will not bias the signal extraction. This has been verified explicitly using a custom dataset with known signal contribution. The $H^{\pm}_{160}$ signal sample has been scaled to the upper limit derived by the CMS analysis [90]. After applying the derived correction weights, the expected signal contribution was extracted in the final fit (the fitting procedure is presented in Section 6.2).

The common strategy is to derive the correction in a dedicated control region. In this analysis the signal region definition has been optimised to select semileptonic $t\bar{t}$ events – the signal region is a $t\bar{t}$ control region. Separating semileptonic $t\bar{t}$ and signal events is complicated due to the similarity of the processes. Only dileptonic $t\bar{t}$ events do not contain a signal. The possibility of deriving the correction in dileptonic $t\bar{t}$ events and applying it to semileptonic $t\bar{t}$ events has been tested. Unfortunately, the idea had to be abandoned because the $S_T$ distribution is systematically shifted towards smaller values compared to the semileptonic case. The two neutrinos in dileptonic events are reconstructed as a single object, $E_T^{\text{miss}}$. If the neutrinos fly in different $\phi$ directions, the summed $E_T^{\text{miss}}$ value partially cancels, resulting in a smaller $S_T$ value. It was tested to retain the same $S_T$ distribution as in semileptonic events with the help of truth information from simulation, which yielded improved but unsatisfactory results.

## 4.4 $t\bar{t}$-system reconstruction

In order to quantify the presence of a potential signal in data, the signal process has to be efficiently and reliably identified. For this purpose the full $t\bar{t}$ system is reconstructed. The $t\bar{t}$ system consists of a semileptonically decaying top quark ($t_{\text{lep}}$) and a hadronically decaying top quark ($t_{\text{had}}$). The $t_{\text{lep}}$ decays into a $b$ quark ($b_{\text{lep}}$) and a $W$ boson ($W_{\text{lep}}$), which in turn decays into a lepton and a neutrino. The $t_{\text{had}}$ decays into a $b$ quark ($b_{\text{had}}$) and a $H^{\pm}$ boson, which in turn decays into a $c$- and an $s$-quark ($j_1$, $j_2$). In SM $t\bar{t}$ events the $t_{\text{had}}$ decays into a $W$ boson ($W_{\text{had}}$), which decays into lighter quarks. Jet $j_1$ ($j_2$) is defined to be the jet with the higher (lower) $p_{\text{T}}$ out of the two jets which originate from the hadronically decaying boson. The $t\bar{t}$ system and nomenclature of objects are visualised in Figure 4.13. The reconstruction uses calibrated physics objects – leptons, $\vec{p}_{\text{T}}^{\text{miss}}$ and jets – as proxies for the lepton, neutrino and quarks arising from the top-quark decays.

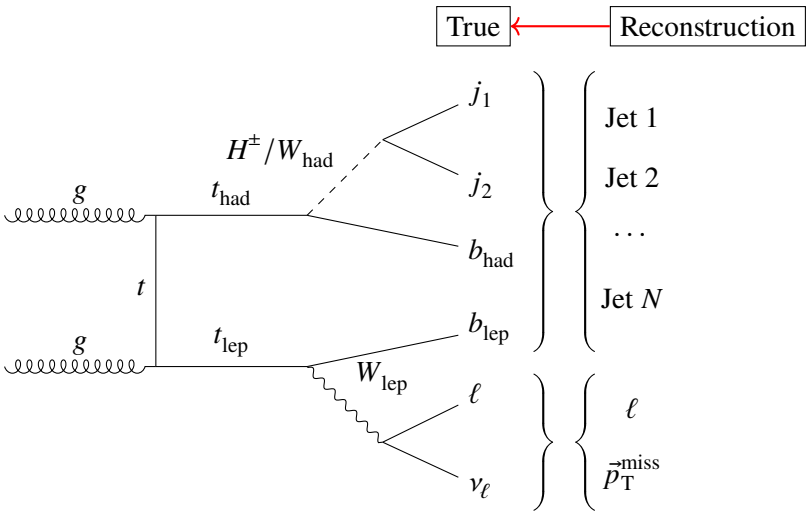



Figure 4.13: Nomenclature of the objects in the $t\bar{t}$ system. Visualised on the right-hand side is the attempted mapping of reconstructed objects onto the respective true objects.

The $t\bar{t}$-system reconstruction is crucial for separating signal and $t\bar{t}$ backgrounds. The identification of the two jets stemming from the hadronically decaying boson ($H^{\pm}/W_{\text{had}}$) facilitates the study of its properties, like the invariant mass. And the $t\bar{t}$-system reconstruction is useful to identify and reject non-$t\bar{t}$ backgrounds.

The $t\bar{t}$-system reconstruction proceeds in two steps: First the $W_{\text{lep}}$ is reconstructed with the help of the reconstructed lepton and $\vec{p}_{\text{T}}^{\text{miss}}$ (Section 4.4.1). In the second step jets are labelled corresponding to their expected origin (Section 4.4.2).

### 4.4.1 $W_{\text{lep}}$ reconstruction

The $W_{\text{lep}}$ decays into a lepton ($\ell$) and a neutrino ($\nu$). The event selection requires exactly one lepton in the event. Thus, the lepton can be unambiguously identified by the single reconstructed lepton. The neutrino cannot be reconstructed directly. It is attempted to be reconstructed using the $\vec{p}_{\text{T}}^{\text{miss}}$measurement and utilising the fact that the lepton and neutrino stem from a common particle with known mass. The

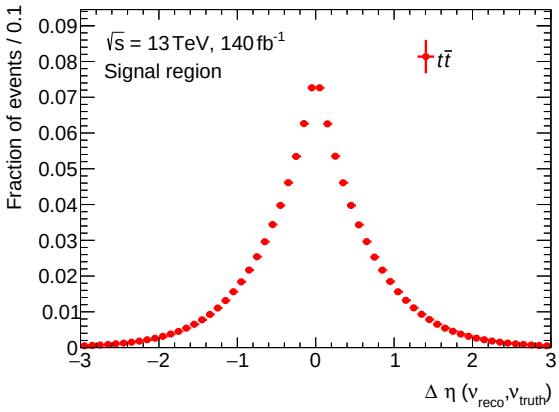



Figure 4.14: Difference between the truth and reconstructed neutrino $\eta$ in SM $t\bar{t}$ simulated events. In case the reconstruction yields two possible solutions, the solution closest to the truth neutrino $\eta$ is chosen.

neutrino is expected to dominate the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$measurement given that it is the only prompt lepton in the event and arises from the decay of an on-shell heavy gauge boson. This assumption is less exact in events where the reconstructed electron or muon stems from the leptonic decay of a $\tau$-lepton, as in this case the lepton is accompanied by two additional neutrinos.

The transverse momentum and $\phi$ values of the neutrino are derived from the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ measurement: $p_{\mathrm{T},\nu} = E_{\mathrm{T}}^{\mathrm{miss}}$; $\phi_\nu = \phi^{\mathrm{miss}}$. The mass of the neutrino is neglected ($m_\nu = 0$). The last missing piece in order to reconstruct the four-momentum of the neutrino is its pseudorapidity, $\eta_\nu$. It is calculated by requiring the invariant mass of the lepton and neutrino to be equal to the true $W$-boson mass, measured to be $m_W = 80.379\,\mathrm{GeV}$ [28],

$$\eta_\nu = \eta_\ell \pm \operatorname{arccosh}\left(\frac{m_W^2}{2p_{\mathrm{T}}^{\nu}p_{\mathrm{T}}^{\ell}} + \cos\left(\phi_\nu - \phi_\ell\right)\right).$$

This equation generally has two solutions. In the jet labelling procedure (Section 4.4.2), one solution will be picked based on the invariant top-quark mass. In case the argument of arccosh is exactly one, there is one solution. In this case the pseudorapidity of neutrino and lepton agree with each other ($\eta_\nu = \eta_\ell$). Due to reconstruction inefficiencies or additional sources of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, it may happen that the argument of arccosh is smaller than one, for which arccos is not defined. In this case the argument of arccosh is set to one, as it is mathematically the closest possible solution. This yields again $\eta_\nu = \eta_\ell$, but this time the invariant mass of the lepton and the neutrino exceeds $m_W$. The latter case occurs for about 35 % of the simulated SM $t\bar{t}$ events.

Figure 4.14 shows the reconstructed resolution of $\eta_\nu$ in simulated SM $t\bar{t}$ events, which is used to validate the neutrino reconstruction. In case the reconstruction yields two possible solutions, the solution closest to the truth neutrino $\eta$ is chosen. The distribution peaks as expected around zero. Deviations from zero are caused by reconstruction inefficiencies and additional sources of $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. In 55 % of the events, the difference between reconstructed and truth $\eta$ is below 0.4.

### 4.4.2 Jet labelling

The quarks from the top-quark decays are expected to be reconstructed as individual jets. The goal is to label in each event four jets as $b_{\text{lep}}$, $b_{\text{had}}$, $j_1$ and $j_2$. But it is a priori not clear which jet corresponds to which initial quark. In addition, there may be more than four jets in an event due to initial- and final-state radiation. This is a combinatorial problem and is in this analysis resolved by comparing the top-quark candidate's mass with the predicted top-quark mass and by utilising flavour-tagging information of jets. The invariant mass of $j_1$ and $j_2$, which form together the hadronically decaying boson $H^{\pm}/W_{\text{had}}$, cannot be utilised as the mass of the charged Higgs boson is a free parameter in theoretical models.

#### Reconstructed top-quark mass probability density functions

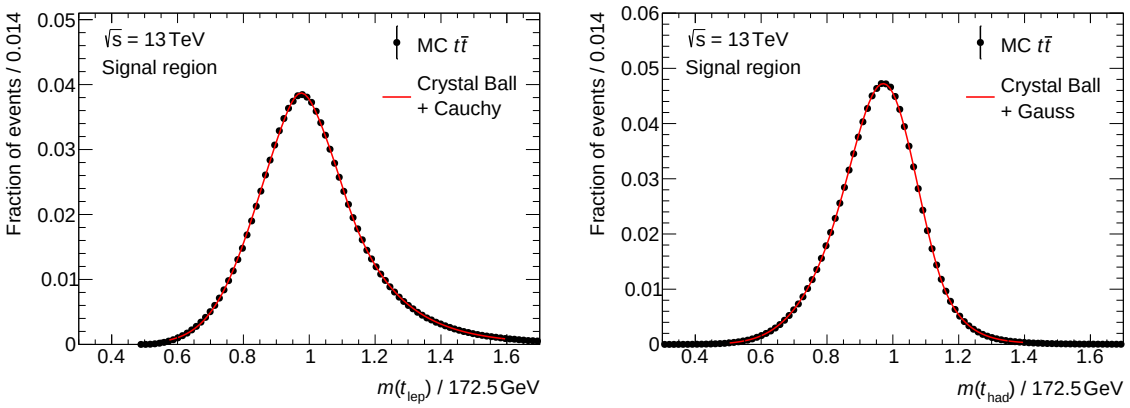
The mass resolutions of semileptonically and hadronically decaying top quarks differ significantly. This is caused by the different four-momenta resolutions of jets, leptons and neutrinos. For this reason, the plain mass difference ($m_{t_{\text{cand}}} - m_{t_{\text{true}}}$) is not well suited for the selection of the best top-quark candidates. Instead, probability density functions of the reconstructed top-quark masses ($\text{PDF}_t$) are used. The $\text{PDF}_t$ are built from SM $t\bar{t}$ simulation by matching reconstructed objects to generator-level (truth) objects.

Truth quarks are geometrically matched ($\Delta R$) to the closest reconstructed jet within $\Delta R = 0.4$. Heavy-flavour quarks are only matched to reconstructed jets that have a respective truth hadron ($p_{\text{T}} > 5\,\text{GeV}$) associated with it. Similarly, light-flavour quarks are only matched to reconstructed jets without an associated heavy-flavour hadron. In dense environments where jets are close to each other, multiple truth quarks can be matched to the same reconstructed jet. If another jet is nearby, ambiguities are resolved by minimising the sum of the four $\Delta R$ values between truth quarks and nearby ($\Delta R \leq 0.4$) reconstructed jets while requiring unique assignments. In about 53 % of the events, at least one truth quark cannot be matched to a reconstructed jet. The main reasons for this are low momenta or large $|\eta|$ values of the truth quarks, resulting in the respective reconstructed jet being outside the detector or reconstruction/calibration acceptance ($|\eta| > 2.5$ or $p_{\text{T}} < 25\,\text{GeV}$). The event still passes the event selection due to additional jets inside the detector acceptance. These events are not considered in the $\text{PDF}_t$.

No matching has to be applied for the lepton, as there is only one reconstructed lepton candidate. The reconstructed neutrino has up to two possible $\eta$ values, and the value closest to the truth-neutrino $\eta$ is selected to build the $\text{PDF}_t$. The $t_{\text{lep}}$-mass and $t_{\text{had}}$-mass $\text{PDF}_t$ are shown in Figure 4.15.

The effects of statistical fluctuations in the tails and the choice of the binning are mitigated by fitting functions to the $\text{PDF}_t$. The top-quark mass distribution is expected to follow a Breit-Wigner distribution due to the small lifetime of the top-quark and the energy-time uncertainty relation. But the measured distribution is smeared by reconstruction effects. Multiple objects with given measurement uncertainties are combined to calculate the top-quark mass. From the central limit theorem follows that they add up to a Gaussian uncertainty, that eventually smears the Breit-Wigner distribution.

The tail in the $\text{PDF}_{t_{\text{lep}}}$ distribution towards higher values is caused by additional sources of $\vec{p}_{\text{T}}^{\,\text{miss}}$ in the events, resulting in inaccurate neutrino reconstructions and thus top-quark mass measurements. The Crystal Ball function [223–225] is commonly used to describe Gaussian Peaks with tails due to losses and consists of a symmetric Gaussian and a one-sided power-law tail and therefore suits the $\text{PDF}_{t_{\text{lep}}}$ distribution well. A combination of the Crystal Ball and Breit-Wigner functions was found to fit the $\text{PDF}_{t_{\text{lep}}}$ distribution best. For $\text{PDF}_{t_{\text{had}}}$ a Crystal Ball + Gaussian function fits best. These fit functions




(a) (b)

Figure 4.15: Reconstructed mass for (a) semileptonically top-quark decays $t_{\text{lep}}$ and (b) hadronically top-quark decays $t_{\text{had}}$ and the respective fits to the distributions.

showed the best reduced $\chi^2$ values equal to 37 for the fit to $t_{\text{lep}}$-mass and 11.6 for the fit to $t_{\text{had}}$-mass $\text{PDF}_t$, respectively. The reduced $\chi^2$ values are relatively large due to the large sample size – small uncertainties – in the central part of the distributions. Though, the agreement between data points and fits visible in Figure 4.15 is sufficient given the purpose of labelling the jets.

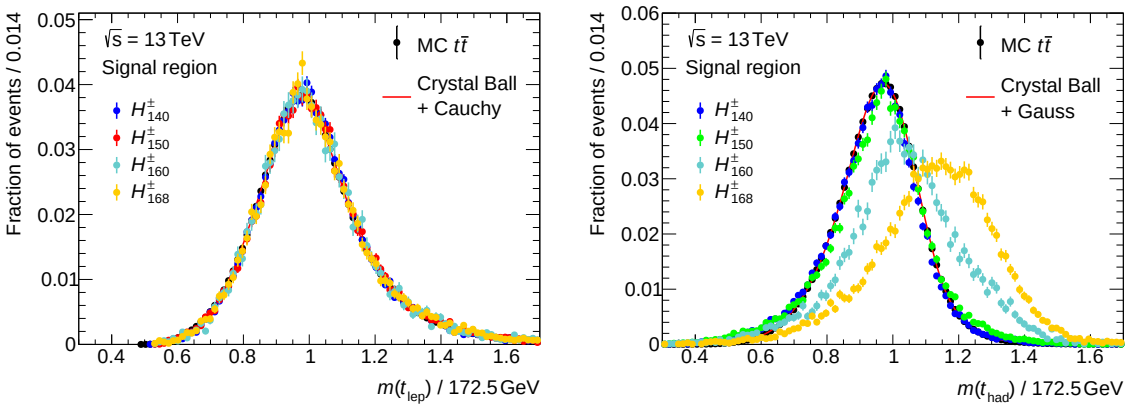
The modelling of both $\text{PDF}_t$ has been validated by evaluating them for random and $p_{\text{T}}$-ordered jet permutations. Good agreement between simulation and data has been seen.

Figure 4.16 shows the SM $t\bar{t}$ $\text{PDF}_t$ overlaid with the $\text{PDF}_t$ derived from signal samples. This is a cross-check to verify that the $\text{PDF}_t$ are equal for $t\bar{t}$ and signal samples, and eventually the method performs equally well in labelling the jets correctly for SM $t\bar{t}$ and signal events. As expected, no difference between SM $t\bar{t}$ and signal samples is seen in the $\text{PDF}_{t_{\text{lep}}}$ distribution. Regarding the $\text{PDF}_{t_{\text{had}}}$, the distributions are equal for all samples with charged Higgs boson masses up to 140 GeV. For larger charged Higgs boson masses the distribution for the signal samples are shifted towards larger values. This comes from an enhanced fraction of off-shell top quarks in these samples. The probability of off-shell top-quark production is increased in these samples because of the small phase-space in the decay of on-shell top-quarks ($m_t \approx m_{H^\pm} + m_b$). As a result of the shifted top-quarks mass distribution, the performance of the jet labelling procedure is slightly reduced for these samples.

The option to use dedicated $\text{PDF}_t$ for each signal mass hypothesis was disregarded because of the extra effort. The background hypothesis would depend on the considered signal mass point. That would result in significantly increased computing resources needed for this analysis for a small improvement in the jet labelling efficiency, as will be shown later in this section.

### Methodology

The jet labelling is performed by building all possible permutations of $b_{\text{lep}}$, $b_{\text{had}}$, $j_1$, $j_2$ jet labellings for all $\eta_\nu$ solutions. The jets are labelled according to the permutation with the largest value of the product




(a) (b)

Figure 4.16: Probability density functions of the reconstructed mass for (a) semileptonically $\mathrm{PDF}_{t_{\mathrm{lep}}}$ and (b) hadronically $\mathrm{PDF}_{t_{\mathrm{had}}}$ decaying top quarks and the respective fits to the distributions. Overlaid are the respective $\mathrm{PDF}_t$ distributions for signal samples with a charged Higgs boson mass close to the top-quark mass.

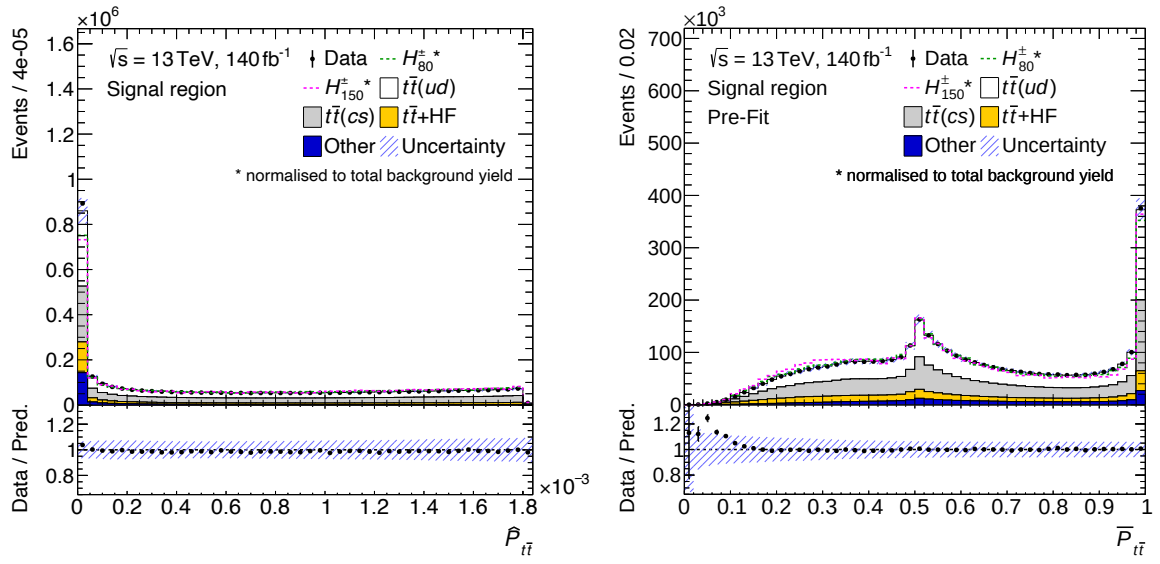
of the $t_{\mathrm{lep}}$-mass and $t_{\mathrm{had}}$-mass $\mathrm{PDF}_t$,

$$P_{t\bar{t}} = \mathrm{PDF}_{t_{\mathrm{lep}}}(m^{\mathrm{cand}}_{t_{\mathrm{lep}}}) \times \mathrm{PDF}_{t_{\mathrm{had}}}(m^{\mathrm{cand}}_{t_{\mathrm{had}}}).$$

The highest $\mathrm{PDF}_t$ product value is denoted by $\hat{P}_{t\bar{t}}$. The larger the value of $\hat{P}_{t\bar{t}}$, the more likely it is that the correct jet permutation has been found. The value of $\hat{P}_{t\bar{t}}$ divided by the sum of $P_{t\bar{t}}$ values of all considered permutations is denoted by $\overline{P}_{t\bar{t}}$. A $\overline{P}_{t\bar{t}}$ value close to 1 hints towards the fact that the correct jet permutation has been found, because in this case there is no other permutation with a high $P_{t\bar{t}}$ value. The distributions of $\hat{P}_{t\bar{t}}$ and $\overline{P}_{t\bar{t}}$ are plotted in Figure 4.17.

Certain permutations are vetoed based on the flavour-tagging score (PCFT score) of jets in order to decrease the probability of having a random permutation in an event with a large $P_{t\bar{t}}$ value. In addition, the top-quark mass requirement is only capable of uniquely assigning the $b_{\mathrm{lep}}$-jet and three other jets to build the $t_{\mathrm{had}}$. All permutations yielding the same $P_{t\bar{t}}$ value except one are vetoed. The applied requirements to achieve this and the physics motivation behind the requirements are listed in the following. The requirements have been optimised to select the correct permutation in SM $t\bar{t}$ events:

1. Consider only up to six jets when building the permutations. Considered first are all $b$-jets followed by the highest-$p_{\mathrm{T}}$ non-$b$-tagged jets.
   - This requirement vetoes unlikely permutations, as additional jets in the event arising from ISR, FSR or similar effects have on average a smaller $p_{\mathrm{T}}$ and less often contain a heavy-flavour hadron than the jets from the top-quark decay products.
2. $b_{\mathrm{lep}}$- and $b_{\mathrm{had}}$-jets have to be $b$-tagged (PCFT score $\geq 3$).
   - The $b$-jet tagging efficiency is relatively large at about 70 % with small misidentification rates of 0.18 % for light-jets and 7.9 % for $c$-jets. Due to the event selection that requires two $b$-jets, most of the events in which the corresponding $b_{\mathrm{lep}}$- or $b_{\mathrm{had}}$-jet is not $b$-tagged




(a) (b)

Figure 4.17: Distributions of (a) $\hat{P}_{t\bar{t}}$ and (b) $\overline{P}_{t\bar{t}}$.

are already filtered out. Therefore, this requirement vetoes many wrong permutations while often keeping correct permutations.

3. PCFT score($b_{\text{lep}}$) ≥ PCFT score($j_1/j_2$) and PCFT score($b_{\text{had}}$) ≥ PCFT score($j_1/j_2$).
   - These requirements again veto unlikely permutations based on the flavour-tagging identification and misidentification rates. In case multiple jets assigned to $t_{\text{had}}$ are $b$-tagged, the latter requirement removes permutations with identical $P_{t\bar{t}}$ values by only retaining the permutation where the jet with the highest PCFT score is labelled as $b_{\text{had}}$. In case the highest PCFT score is shared by multiple jets assigned to $t_{\text{had}}$, the one with the larger $p_{\text{T}}$ will be labelled as $b_{\text{had}}$. This is motivated by the fact that in SM $t\bar{t}$ events the $b_{\text{had}}$-jet has on average a larger $p_{\text{T}}$ than the jets stemming from the decay of the $W$ boson.

4. $p_{\text{T},j_1} > p_{\text{T},j_2}$;
   - This requirement vetoes permutations with the same $P_{t\bar{t}}$ value. Interchanging $j_1$ and $j_2$ does not lead to a different $H^{\pm}/W_{\text{had}}$ nor $t_{\text{had}}$ candidate. For this reason no attempt is being made in classifying these jets further.

An example of the permutations in an event with five jets, out of which two are $b$-tagged, is shown in Table 4.4.

### Results

Figure 4.18 summarises the performance of the jet labelling procedure. The plots contain only events in which all truth quarks could be matched to a reconstructed jet. Figure 4.18(a) shows how often the jets that are matched to a respective truth quark have been labelled correctly. The $u/c$- and $d/s$-jets count as

Table 4.4: Example of the permutations and $P_{t\bar{t}}$ values for a signal event of the $H^{\pm}_{80}$ sample. The event has five jets. The jets are sorted by their $p_{\mathrm{T}}$. The jets with indices 1 and 2 are $b$-jets. Permutation 1 has the highest $P_{t\bar{t}}$ value with $\hat{P}_{t\bar{t}} = 3.2 \times 10^{-4}$ and $\overline{P}_{t\bar{t}}$ is 0.516 for this event.

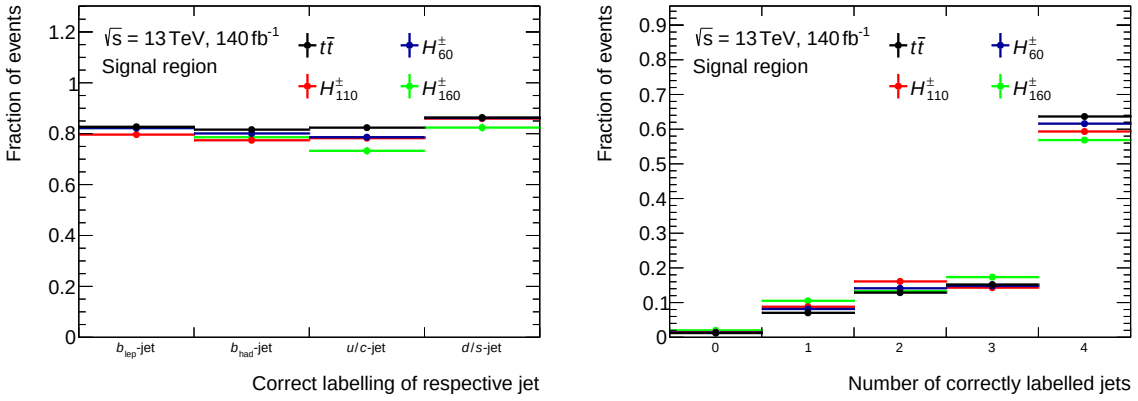
| Permutation | Jet indices | | | | | Top-quark mass [GeV] | | $P_{t\bar{t}}$ |
|---|---|---|---|---|---|---|---|---|
| | $\eta_\nu$ | $b_{\mathrm{lep}}$ | $b_{\mathrm{had}}$ | $j_1$ | $j_2$ | $m_{t_{\mathrm{lep}}}$ | $m_{t_{\mathrm{had}}}$ | |
| 1 | −1.66 | 1 | 2 | 0 | 3 | 211 | 150 | $3.2 \times 10^{-4}$ |
| 2 | 0.54 | | | | | 262 | | $4.6 \times 10^{-5}$ |
| 3 | −1.66 | 1 | 2 | 0 | 4 | 211 | 361 | $4.6 \times 10^{-8}$ |
| 4 | 0.54 | | | | | 262 | | $6.5 \times 10^{-9}$ |
| 5 | −1.66 | 1 | 2 | 3 | 4 | 211 | 294 | $3.3 \times 10^{-7}$ |
| 6 | 0.54 | | | | | 262 | | $4.6 \times 10^{-8}$ |
| 7 | −1.66 | 2 | 1 | 0 | 3 | 95 | 125 | $4.1 \times 10^{-6}$ |
| 8 | 0.54 | | | | | 153 | | $2.5 \times 10^{-4}$ |
| 9 | −1.66 | 2 | 1 | 0 | 4 | 95 | 396 | $1.0 \times 10^{-9}$ |
| 10 | 0.54 | | | | | 153 | | $6.1 \times 10^{-8}$ |
| 11 | −1.66 | 2 | 1 | 3 | 4 | 95 | 359 | $2.3 \times 10^{-9}$ |
| 12 | 0.54 | | | | | 153 | | $1.4 \times 10^{-7}$ |

correctly labelled if they are labelled as either $j_1$ or $j_2$. The probabilities of labelling the true $b_{\mathrm{had}}$-, $b_{\mathrm{lep}}$-, $u/c$- or $d/s$-jet correctly are very similar, at around 80 %.

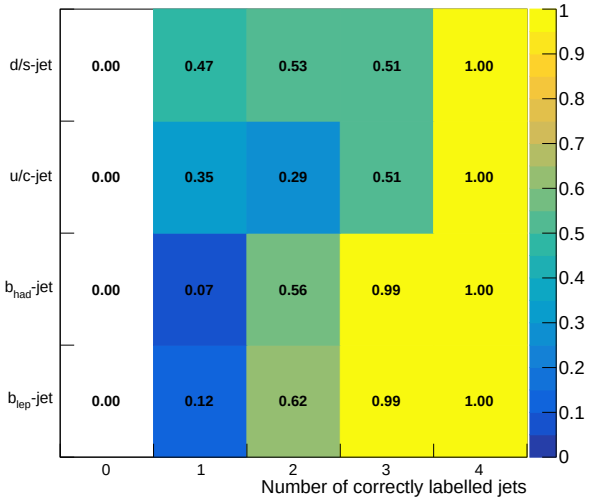
Figure 4.18(b) shows how many jets are labelled correctly in each event. In about 64 % of the SM $t\bar{t}$ events, all jets are labelled correctly. In comparison with the results from Figure 4.18(a), a clear correlation can be seen. Either all jets are labelled correctly ($64\,\% > 0.8^4 \approx 40\,\%$), or if one jet is labelled wrongly, other jets are likely to be labelled wrong, too.

One can also see in Figure 4.18(b) that the efficiency of labelling all jets correctly is overall a few percent smaller for signal samples and is smallest for samples with a large charged Higgs boson mass. The overall smaller efficiency is due to a greater $c$-jet abundance in signal events compared to SM $t\bar{t}$ events ($\mathscr{B}(H^{\pm} \to cs)/\mathscr{B}(W_{\mathrm{had}} \to cs) \approx 2$). The probability of the $c$-quark-initiated jet getting $b$-tagged is larger than for light jets. The additional $b$-tag increases the number of permutations and complicates the jet labelling. The decrease for large charged Higgs boson masses can be attributed to an on-average larger number of jets in the events. This is a bias introduced by the event selection. The probability of the $b_{\mathrm{had}}$-jet to meet the jet-$p_{\mathrm{T}}$ requirement is higher if the top quark is more highly boosted, which is the case when it recoils against additional radiated gluons. The increased number of jets leads to an increased number of wrong permutations that have to be rejected.

Figure 4.18(c) shows in a two-dimensional plot the efficiency of a respective jet being labelled correctly as a function of the total number of correctly matched jets. The efficiencies in one column add up to the number of correctly matched jets. The columns with 0 and 4 correctly labelled jets all have efficiencies of 0 % and 100 %, respectively, because in these cases either none or all jets are labelled correctly. If only one jet is not labelled correctly, it is mostly a jet from the charged Higgs boson decay. Those jets are generally more difficult to identify because additional jets in the event are mostly light-flavour jets.

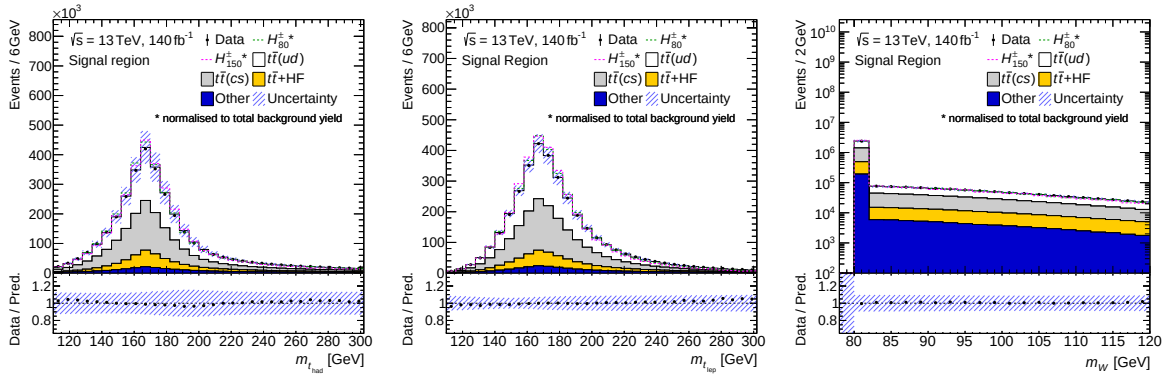



(a) (b)




(c)

Figure 4.18: Performance plots of the jet labelling procedure. In Figure (a) the fraction of events in which the correct label has been assigned to the truth quark matched jets is shown. Figure (b) shows how many jets are labelled correctly in each event. Figure (c) shows the efficiency of a jet being labelled correctly as a function of the total number of correctly matched jets For a meaningful evaluation of the performance only events in which all four truth quarks could be matched to a reconstructed jet are considered in the plots.




(a) (b) (c)

Figure 4.19: Mass distributions of the reconstructed (a) hadronically and (b) semileptonically decaying top quark, and (c) the $W$ boson.

For the column with two correctly assigned jets, the Pearson correlation coefficient showed no clear indication that mostly the same pair of jets is labelled wrong, i.e. usually random combinations have the highest $P_{t\bar{t}}$ value. This is also true for events in which only one jet is labelled correctly. Since no further discrimination is made between $j_1$ and $j_2$ as being the up- or down-type-quark jet, the probability of randomly having one jet from the charged Higgs boson decay being labelled correctly is highest.

Regarding the events with two solutions for the neutrino $\eta$, the correct one is picked in about two thirds of the cases. Figure 4.19 shows the reconstructed top-quark masses for the semileptonically and hadronicically decaying top quark as well as the reconstructed $W$-boson mass.

### 4.4.3 Outlook

The method described above has been developed to reconstruct the $t\bar{t}$ system for this analysis. But as might be expected, there are many analyses with $t\bar{t}$ final states that face the same issue, and different tools have been developed to reconstruct the $t\bar{t}$ system. The most popular one is the Kinematic Likelihood Fitter framework [107, 109, 226] (KLFitter). This outlook will shortly introduce the KLFitter library and why its usage has been refrained from in this analysis. Lastly, other ideas exploiting multivariate analysis methods to reconstruct the $t\bar{t}$ system will be presented.

#### Kinematic likelihood fitter

The KLFitter framework performs a likelihood fit of kinematic quantities to reconstruct top-quark events. The kinematic quantities are allowed to be varied within the detector resolution. The four-momenta of jets, leptons and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ are fitted to match with the $W$-boson and top-quark mass. Based on the likelihood, the best permutation of jets and a set of improved four-momenta are retrieved.

The usage of the KLFitter framework has been tested at an early stage of this analysis using signal and SM $t\bar{t}$ samples. The fraction of events in which all four jets have been labelled correctly is identical to the results from this analysis within percent level. But the resolution of the charged Higgs boson mass improves by a few percent using the newly fitted four-momenta. In order to test how much this analysis

can improve thanks to the improved four-momenta, a BDT has been trained and limits on the branching ratio of a top quark decaying to a charged Higgs boson and a $b$ quark have been calculated, as done in the nominal analysis and described in Sections 5 and 6.2, respectively, but using a reduced background sample size and not considering systematic uncertainties. An improvement between 0 and 4 % in the limits has been seen for the various mass points when using the KLFitter method, which is negligible compared to the uncertainties on the limits. Since the application of the KLFitter increased the runtime per event by a factor of 10 or higher in this analysis, it was decided to use the method presented in the previous section.

#### Multivariate reconstruction methods

It has been tried to tackle the combinatorial jet labelling task by using multivariate analysis techniques. BDTs have been tested explicitly first, due to their good out-of-the-box performance and to get a realistic notion about the room for improvement. Many BDT setups have been tested using different multi-class labels, training data, etc., but only a few selected will be introduced below.

A binary BDT has been trained for each mass point to decide on a jet permutation to be correct or incorrect. In the training, the most likely jet labelling permutations were included. This means each event was included multiple times, with only one permutation being correct. In the end, the BDT was evaluated on all permutations and the jets were labelled based on the permutation that achieved the highest BDT score. Inputs to the BDT were variables describing the kinematics of the top-quark candidates of the respective permutation, like the momentum of the jets, the invariant mass of the top-quark, etc. The BDT managed to select the correct permutation in 75 % of the events. Compared to the results shown in Figure 4.18, this is a clear improvement, but it is not a fair comparison because the invariant mass of the hadronic decaying boson is included in the BDT. A clear statement on the performance can only be made based on the expected limits on $\mathcal{B}(t \to H^{\pm}b)$. As for the KLFitter case, a (in this case second) BDT was trained to separate between signal and background events using the jet permutations from the first BDT and limits were calculated with reduced background statistics and not considering systematic uncertainties. The results were comparable to the previous results despite the larger fraction of correct jet labellings. The reason for this is that the variables discriminating right from wrong permutations and signal from background events are all related to the top-quark kinematics and are very similar. All kinematic properties of the wrong permutations picked by the first BDT look much more like they do in signal events, degrading the separation power of the second BDT. The conclusion is that the rather simple reconstruction in this analysis still collects many correct permutations and one BDT is sufficient to do all the hard work. The final BDT not only separates signal from background events but also indirectly identifies events with wrong jet assignments, as these events receive lower BDT scores.

Another idea was to use one BDT with all the permutations and derive from its output limits on $\mathcal{B}(t \to H^{\pm}b)$ directly, since the quantities to identify correct permutations and signal events are very similar. The BDT had three output classes: (I) correct signal permutation, (II) correct SM $t\bar{t}$ permutation, and (III) wrong permutations. For each event the permutation with the highest BDT score for class I or II was selected and a simultaneous fit of the two classes has been performed. The results and conclusions were similar to the binary BDT training.

An idea that has never been realised due to time constraints was to use a RNN to reconstruct the $t\bar{t}$ system. In an RNN the output from the previous step is fed back as an input for the next step. RNNs are well suited to work on variable-length sequences. In this particular use case, this is the number of jets

that varies for each event. Inputs are $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$, the lepton four-momenta, and the four-momenta plus PCFT score of jets. To improve the performance, it might be beneficial to include preprocessing layers that can calculate the invariant mass of the input four-momenta. Outputs are the jet indices of $b_{\mathrm{lep}}$, $b_{\mathrm{had}}$, $j_1$ and $j_2$ (many-to-many architecture).

A new library to reconstruct event topologies that has quickly risen in popularity is *SPANet* [227]. SPANet uses symmetry-preserving attention neural networks to determine the jet assignments without evaluating all possible permutations. It is advisable to test the performance of SPANet in a future iteration of this search with the LHC Run-3 dataset.

CHAPTER 5

# Identification of $H^{\pm} \rightarrow cs$ events

This chapter describes the analysis strategy to identify signal events in the pre-selected dataset. Firstly, differences between signal and background events and the quantities in which these differences show up are discussed in Section 5.1. The separating powers of these quantities are combined into one powerful discriminant with the help of a BDT. The training and optimisation of the BDT is described in Section 5.2.

## 5.1 Discriminating quantities

The background in this analysis mainly comprises SM $t\bar{t}$ events with a contribution of 92.5 %. Most work has been spent on discriminating signal and SM $t\bar{t}$ processes, not only because of its large contribution but also because $t\bar{t}$ events are most challenging to discriminate from signal events. The only difference between the two processes are the different characteristics of the charged Higgs and $W$ bosons. These are the bosons' masses, couplings, and spins. The implications of these differences are discussed in the following (Sections 5.1.1 to 5.1.3). Useful quantities to identify non-$t\bar{t}$ background are discussed in Section (5.1.4).

### 5.1.1 Boson mass

The mass of the charged Higgs bosons and therewith the mass difference with respect to the $W$ boson is unknown. But a potential mass difference would be visible in many kinematic quantities related to the $t_{\text{had}}$-quark and its decay products.

#### Kinematic differences

Figure 5.1 shows an exemplary decay topology of the $t_{\text{had}}$-quark for a lighter and a heavier charged Higgs boson to visualise the kinematic differences caused by a different charged Higgs boson mass hypothesis. The implications of the charged Higgs boson mass on the decay kinematics discussed below are based on energy and momentum conservation. Only scenarios with $m_{H^{\pm}} < m_t$ are being discussed.

In the $t_{\text{had}}$-quark rest frame the momenta of the $H^{\pm}$ boson and the $b_{\text{had}}$-quark are back-to-back and the absolute values of the momenta are fully determined by the masses of the participating particles. The flight directions of $c$- and $s$-quarks are not fixed with respect to the $H^{\pm}/b_{\text{had}}$ flight direction. In Figure 5.1

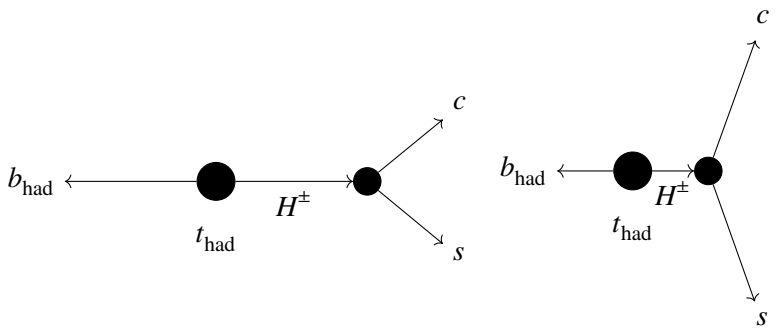



Figure 5.1: Exemplary decay topology of the $t_{\text{had}}$-quark in signal events for a (a) lighter and (b) heavier charged Higgs boson in the $t_{\text{had}}$ rest frame. Lighter means much lighter than the top-quark mass, and heavier is close to the top-quark mass. The topology is drawn in the $t_{\text{had}}$-quark rest frame. The arrows represent the momenta of the particles.

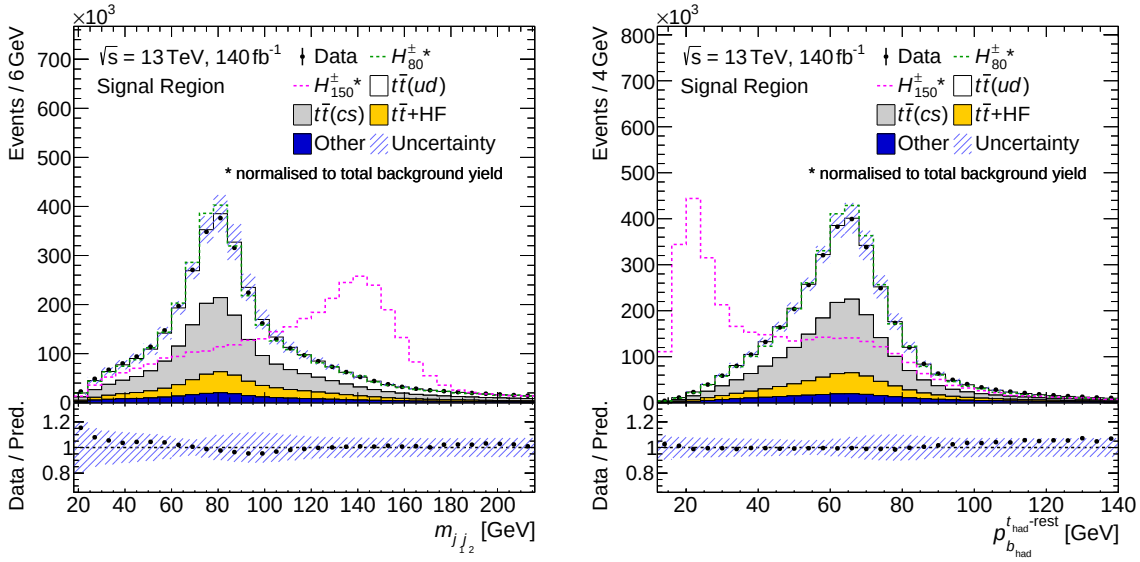
only a specific configuration is drawn. For a heavier $H^{\pm}$, the momenta of $H^{\pm}$ and $b_{\text{had}}$-quark decrease, whereas the momenta of $c$- and $s$-quarks increase on average. The angle between $c$- and $s$-quarks is on average larger, whereas the angle of one of them with respect to the $b_{\text{had}}$-quark is on average smaller.

In the laboratory frame the $t_{\text{had}}$-quark is usually boosted. If $m_{H^{\pm}}$ is large, most of the $t_{\text{had}}$-quark momentum will be transferred to the $H^{\pm}$, and their momenta are aligned. In fact, in the laboratory frame the $p_{\text{T}}$ of $H^{\pm}$ even increases slightly with increasing $m_{H^{\pm}}$, whereas the $p_{\text{T}}$ of the $b_{\text{had}}$-quark decreases rapidly.

#### Affected quantities

The kinematic differences discussed above are visible in many quantities. First and foremost, the dijet mass of $j_1$ and $j_2$ ($m_{j_1 j_2}$), which is a direct measurement of $m_{H^{\pm}}$. A measurement of the $H^{\pm}$ or $b_{\text{had}}$-quark momentum in the $t_{\text{had}}$-quark rest frame ($p_{b_{\text{had}}}^{t_{\text{had}}-\text{rest}}$) is an indirect measurement of $m_{H^{\pm}}$, because the absolute value of the momentum is solely determined by the masses of the participating particles ($t \to H^{\pm} b$), which are all known except for $m_{H^{\pm}}$. The latter two quantities are expected to be measured less precisely than $m_{j_1 j_2}$, as they are calculated from three jets (instead of two). But on the contrary, it directly relates the three quarks and quantifies whether the $b_{\text{had}}$-jet four-momentum matches the expectation, given $j_1$ and $j_2$.

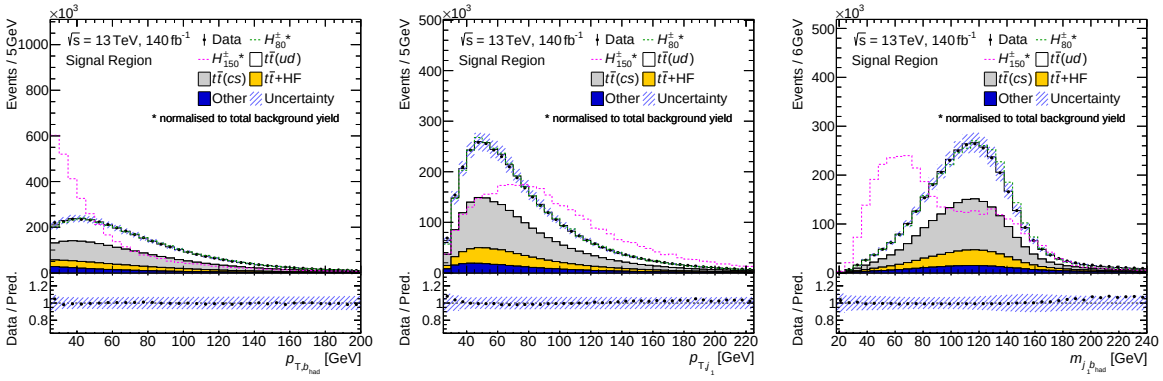
The distributions of $m_{j_1 j_2}$ and $p_{b_{\text{had}}}^{t_{\text{had}}-\text{rest}}$ are shown in Figure 5.2 for the total background prediction. Overlaid are the predictions from the $H_{80}^{\pm}$ and $H_{150}^{\pm}$ signal samples. A strong separation between the background prediction and the $H_{150}^{\pm}$ signal sample is recognisable. In the $m_{j_1 j_2}$ distribution the SM $t\bar{t}$ backgrounds peak as expected around the $W$-boson mass and the signal samples around their respective charged Higgs boson mass values. This confirms the effectiveness of the $t\bar{t}$-system reconstruction. The quantities discussed in this paragraph are very powerful in separating signal and background events, as all events from the same process are expected to have the exact same value, which is only smeared by the natural decay width and reconstruction effects. All other kinematic quantities show differences only on a statistical basis.




(a) (b)

Figure 5.2: Distributions of the (a) dijet mass of $j_1$ and $j_2$, and (b) the absolute value of the $b_{\text{had}}$-jet momentum in the $t_{\text{had}}$ rest frame.

Other quantities that are correlated to the charged Higgs boson mass are the momenta of the three jets stemming from the $t_{\text{had}}$-quark ($b_{\text{had}}$ $j_1$, $j_2$) and the angles between them. This has direct implications on the dijet masses built from the $b_{\text{had}}$-jet and $j_1$ or $j_2$ ($m_{j_1 b_{\text{had}}}$ and $m_{j_2 b_{\text{had}}}$). A few selected quantities are shown in Figure 5.3. They are sensitive to the momentum of the $t_{\text{had}}$-quark in the laboratory frame as well as the orientations of the $t_{\text{had}}$-quark and $H^{\pm}$ decay planes with respect to the laboratory frame. Because of these correlations, quantities like $p_{\text{T},t_{\text{had}}}$ also help in separating signal from background events, even though their distributions don't show a separation between signal and SM $t\bar{t}$ events.




(a) (b) (c)

Figure 5.3: Distributions of the (a) $b_{\text{had}}$-jet $p_{\text{T}}$, (b) $j_1$ $p_{\text{T}}$, and (c) dijet mass of the $b_{\text{had}}$-jet and $j_1$.

Correlations with flavour-tagging quantities are also important when exploiting kinematic differences.

For example, if the $b_{\text{had}}$-jet passes a tighter $b$-tagging WP, it is even more likely that it is labelled correctly, and more emphasis can be put on the $b_{\text{had}}$-jet $p_{\text{T}}$.

If the mass difference $m_{H^{\pm}} - m_W$ is close to or larger than the dijet mass resolution, the $t_{\text{had}}$-quark kinematic quantities yield the largest separating power between signal and SM $t\bar{t}$ events.

### 5.1.2 Boson couplings

Differences in the couplings of the $H^{\pm}$ and $W$ bosons to other particles result in different decay widths and decay properties. The difference in the decay width is challenging to measure in data. The decay width of the charged Higgs boson is expected to be narrow (cf. Section 4.1.1). This means that the difference in the decay width is an order of magnitude smaller than the dijet mass resolution ($\Gamma_W - \Gamma_{H^{\pm}} \ll \Delta m_{jj}$). But differences in the flavour of the decay products can be seen in data with the help of jet flavour-tagging algorithms.

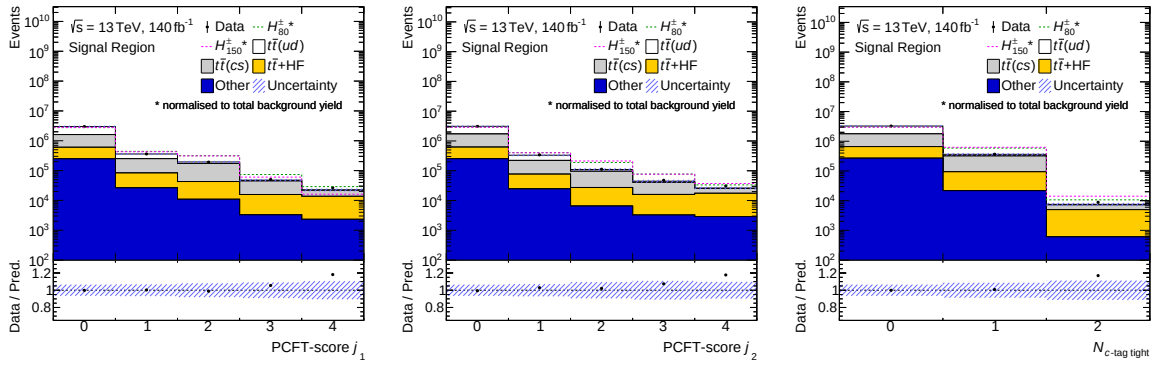
In this analysis the charged Higgs boson is expected to decay solely into a $cs$-quark pair. The probability of a hadronically decaying $W$ boson to decay into a $cs$-quark pair is about 50 %. Otherwise, it decays mainly into a $ud$-quark pair. In data, the charged Higgs boson would be visible by an increased $cs$-quark pair contribution in the hadronic decaying boson's decay products ($j_1$ and $j_2$).

Out of the four mentioned quark types – $u$, $d$, $s$, and $c$ – current flavour-tagging algorithms are only capable of efficiently separating $c$-jets from the other jet types. Studies on identifying $s$-quark jets have been performed showing some identification power with respect to $d$-quark jets [228, 229]. The $s$-jets are identified in two ways. On the one hand, by a larger energy deposition in the hadronic calorimeter from neutral Kaons compared to neutral pions that promptly decay into photons. And on the other hand, by decays of neutral Kaons into charged pion pairs inside the tracking volume. But the identification and rejection rates are very small compared to $b/c$-taggers. And no calibrations are available for these taggers yet. Therefore, only the $c$ quark is utilised to identify the $cs$-quark pair.

This analysis specifically uses a flavour-tagging calibration with $c$-tagging WPs to identify signal events (cf. Section 3.4.6). Figure 5.4 shows the PCFT scores of $j_1$ and $j_2$, and the number of jets in each event passing the tight $c$-tagging WP. In the tight $c$-tag bin (bin 2 of the PCFT score) of $j_1$ and $j_2$, a larger contribution of signal than background events is expected ($\sim 50$ % enhancement; note the logarithmic $y$-axis in Figure 5.4). An enhancement is also seen for the loose $c$- and loose $b$-tag bins – bin 1 and 3, respectively. The $t\bar{t}(ud)$ background is heavily suppressed in these bins. The enhancement is seen for both plotted signal mass hypotheses in Figure 5.4. The flavour-tagging quantities can help to identify a potential signal in data, even if the charged Higgs boson mass is close to the $W$-boson mass.

### 5.1.3 Boson spin

The hadronic decaying boson's spin impacts the decay angles between the $t_{\text{had}}$-quark decay products. Differences are observable because the weak interaction couples only to left-handed particles and because the bosons are produced in the decay of top quarks, entangling the $W$-boson polarisation with the $b$-quark. (If the bosons were produced in unpolarised states, the difference in phase space caused by the different spins would average out).




(a) (b) (c)

Figure 5.4: PCFT scores of (a) $j_1$ and (b) $j_2$, and (c) the number of jets in an event passing the tight $c$-tagging WP.

### Definition of cos ($\theta^*$) – Spin sensitive variable

The top quark decays before it hadronises. Figure 5.5 shows the possible top-quark decay configurations. In Figure 5.5, without loss of generality, it is assumed that the top quark is at rest and its spin points to the right. The possible polarisation states of the emerging $W$ boson are (i) right-handed, momentum and spin are aligned; (ii) left-handed, momentum and spin are opposite; (iii) longitudinal, momentum and spin are perpendicular. The fraction of right-handed $W$ bosons is negligible ($F_{\mathrm{R}} \approx 0$). The reason for this is that in this case the $b$ quark is right-handed, too. However, the weak interaction couples only to left-handed particles. Right-handed $W$ bosons are not completely disallowed, as the $b$ quark is a massive particle, and massive particles comprise left- and right-handed chiral states. Left-handed and longitudinally polarised $W$ bosons have configurations with a left-handed $b$-quark and have fractions of about $F_{\mathrm{L}} \approx 0.3$ and $F_0 \approx 0.7$ (calculated at NNLO [230] and measured at ATLAS [231]), respectively.

The $W$ boson has two decay configurations, whereby one of them is suppressed. The anti-quark is preferably emitted in the direction of the $W$-boson spin. The dominant $W$-boson decay configuration is also shown in Figure 5.5.

It can be seen in Figure 5.5 that the angle between the $b_{\mathrm{had}}$-quark and the emerging up/down-type quarks depends on the $W$-boson helicity state. The angle $\theta^*$ is defined as the angle between the reversed $b_{\mathrm{had}}$-quark momentum and the down-type-quark momentum in the $W$-boson's rest frame. The definition of $\theta^*$ is sketched in Figure 5.6(a). The helicity fraction amplitudes as a function of cos ($\theta^*$) are calculated to be [232]

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\cos(\theta^*)} = \frac{3}{8}(1+\cos(\theta^*))^2 F_{\mathrm{R}} + \frac{3}{8}(1-\cos(\theta^*))^2 F_{\mathrm{L}} + \frac{3}{4}(1-\cos^2(\theta^*))F_0.$$

The individual contributions of the $W$ boson polarisation states and the combined SM prediction are plotted in Figure 5.6(b). In the SM, very few events with a cos ($\theta^*$) value close to 1 are predicted.

For signal events the discussion simplifies immensely because the charged Higgs boson has spin zero. No distinct helicity states exist, and in the decay of the charged Higgs boson, no preferred emittance direction exists, as quark and anti-quark have the same handedness. Anyhow, the coupling of the charged Higgs boson to other particles does not have a spin-specific characteristic, as does the weak interaction

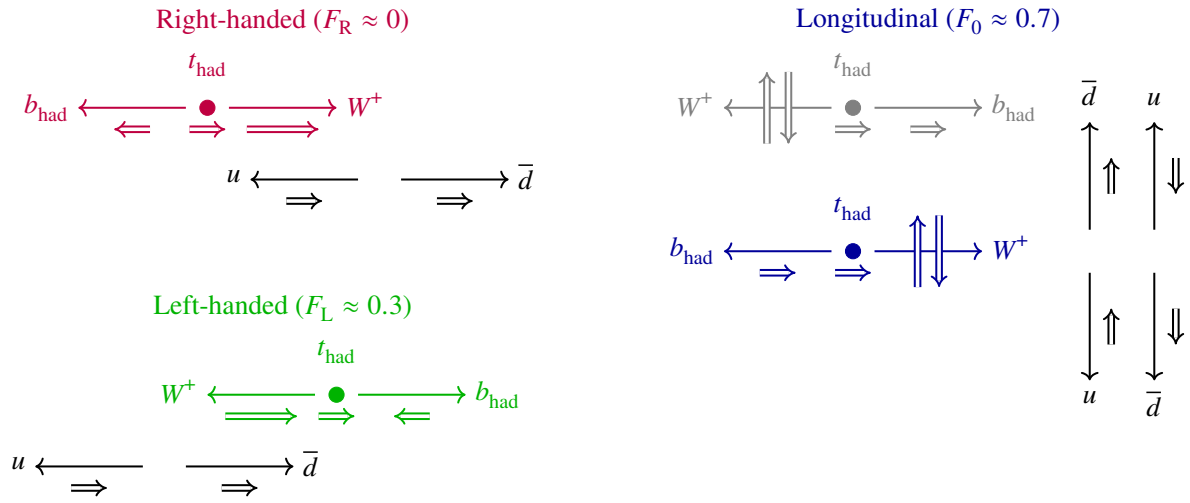
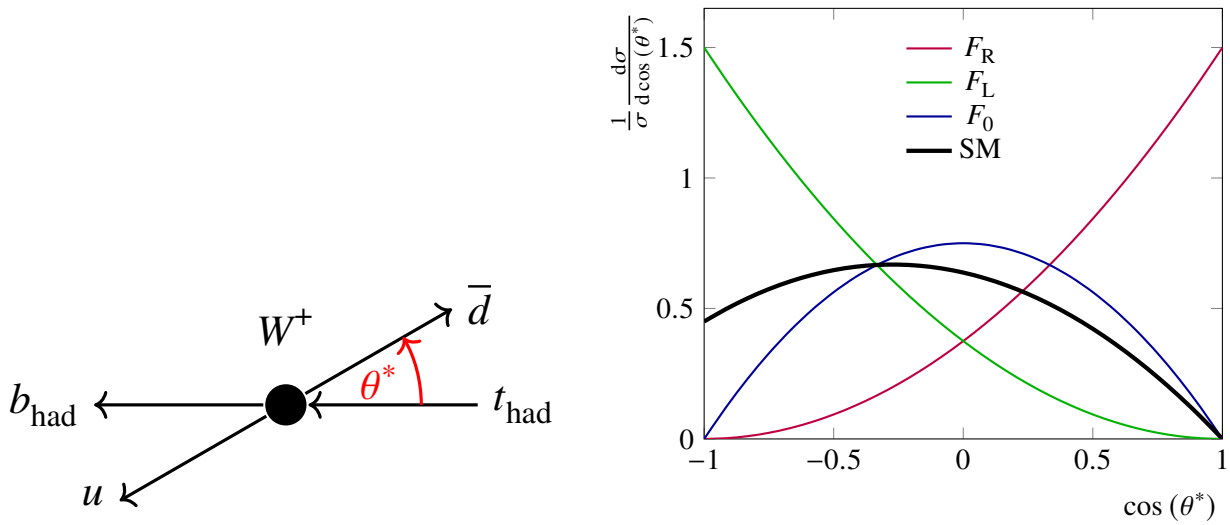



Figure 5.5: Top-quark decay configurations in the top-quark rest frame. Momenta and spins of particles are denoted by single and double arrows, respectively. The configurations are grouped based on the $W$-boson polarisation. Shown in black is the preferred decay configuration of the $W$ boson.




(a) (b)

Figure 5.6: (a) Definition of the angle $\theta^*$. (b) Differential cross-section normalised to the total cross-section as a function of $\cos(\theta^*)$. Shown in black is the distribution as expected in the SM. For comparison the separate distributions are plotted where the respective fraction is set to 1.

($V$–$A$ coupling). For these reasons the distribution of signal events in $\cos(\theta^*)$ is expected to be flat. The anticipated behaviour has been validated to be simulated correctly by plotting the $\cos(\theta^*)$ distribution for signal and SM $t\bar{t}$ simulated events using the truth-quark momenta (see Figure 5.7(a)).

#### Reconstruction of $\cos(\theta^*)$

At the reconstruction level, multiple effects reduce the separation power significantly. First and foremost, a limited jet reconstruction resolution will smear out the distributions. This plays an important role, as three jets are needed to calculate $\cos(\theta^*)$.

In events with $\cos(\theta^*)$ values close to 1 ($-1$) the up-type (down-type) quark flight direction agrees with the $b_{\text{had}}$-quark flight direction. This is also approximately true in the laboratory frame, which makes it extremely difficult to resolve them as two individual jets. If the two jets are not separated, the event either doesn't pass the event selection or a correct jet labelling is not possible. This removes events at the edges of the distribution. As a consequence, a lot of separation power in the highest discriminating region – around $\cos(\theta^*) = 1$ – is lost.

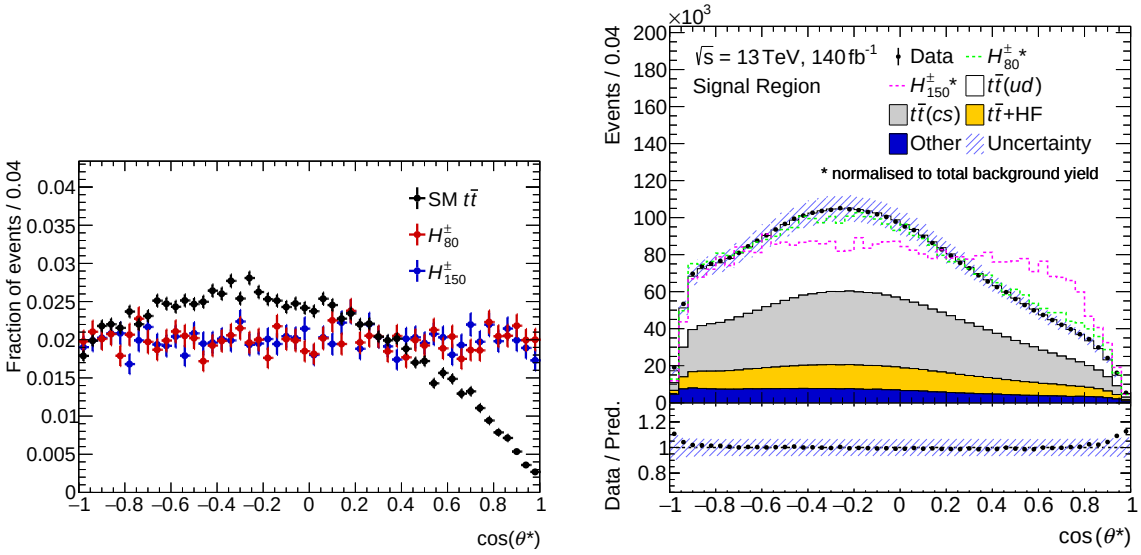
Another reason why the separation power is reduced is that the $t_{\text{had}}$ might be wrongly reconstructed. As discussed in Section 4.4, it may often happen that a respective reconstructed jet is missing or that the jets are wrongly labelled. As a result, many random jets are picked up, which makes the $\cos(\theta^*)$ distribution similar for all processes.

The last and most important reason is that it is difficult to identify the down-type quark jet in the boson's decay products. The flavour-tagging algorithms are exploited for this task. But as discussed previously, the $W$ boson mainly decays in $u$, $d$, $c$, and $s$ quarks, and only $c$-jets can be separated efficiently from the other quark types.

The jet out of $j_1$ and $j_2$ passing the tightest $c$-tagging WP is expected to be the up-type-quark jet. In case none of the two jets pass a $c$-tagging WP, the jet passing the tightest $b$-tagging WP is expected to be the up-type-quark jet. This definition is motivated by the small CKM-matrix element $V_{cb}$, which is much smaller than the $c$-jet misidentification rates of the $b$-tagging WPs. In case both jets are untagged, the one with the larger $p_{\text{T}}$ is expected to be the up-type quark jet. This is an almost random selection, but motivated by the fact that the $p_{\text{T}}$ spectrum of the $c$-jet is slightly harder than for the $s$-jet in signal events. The $p_{\text{T}}$ requirement is the deciding requirement in most events with a $u$-quark and about half of the events with a $c$-quark, as about 45 % of true $c$-jets pass the loosest calibrated $c$-tagging WP. In summary, the up-type-quark jet is defined to be $j_1$ or $j_2$ depending on the PCFT score, prioritising score $2 > 1 > 4 > 3 > 0$, and if both PCFT scores are identical, $j_1$ is expected to be the up-type-quark jet. As we are interested in the down-type-quark jet, the other jet ($j_1$ or $j_2$), which is not expected to be the up-type-quark jet, is of interest.

As a side remark, even if it wouldn't be possible to separate up- and down-type-quark jets at all and one would need to always pick a random jet, some separation power in $\cos(\theta^*)$ would remain due to the specific distribution for SM $t\bar{t}$ events. Selecting the wrong jet just flips the sign of the $\cos(\theta^*)$-value. Therefore, a pure random selection would result in the original $\cos(\theta^*)$ distribution overlaid with the same distribution mirrored around 0, with both distributions weighted by 50 %. In Figure 5.7(a) one can recognise that this decreases the separation power immensely, but nonetheless for SM $t\bar{t}$ processes fewer events are expected close to 1 and $-1$ than for signal processes.

Figure 5.7(b) shows the $\cos(\theta^*)$ distribution of all backgrounds stacked and the signal distributions




(a) (b)

Figure 5.7: Distribution of cos $(\theta^*)$ comparing (a) signal and SM $t\bar{t}$ events at truth level and (b) signal with the total background prediction at reconstruction level.

overlaid for two signal mass point hypotheses. Data and predictions agree well with each other. The previously discussed effects are clearly visible. The differences between signal and background almost vanished. The $H^{\pm}_{80}$ signal sample shows more events above cos $(\theta^*) \approx 0.5$ with an average excess of about 10 %. The distribution of all processes, including the $H^{\pm}_{80}$ signal process, is skewed towards smaller values. This is mainly caused by the $p_{\mathrm{T}}$ requirement when identifying the quark types. This can be illustrated by imagining the decay of the top quark at rest. The $H^{\pm}$ and $b_{\mathrm{had}}$-quark are back-to-back. The $H^{\pm}$ decays further into two quarks. The quark that is emitted opposite to the $H^{\pm}$-boson momentum will have a lower $p_{\mathrm{T}}$. Opposite to the $H^{\pm}$ momentum is in the direction of the $b_{\mathrm{had}}$-quark and therefore cos $(\theta^*)$ is closer to -1. This effect is not as pronounced for the $H^{\pm}_{150}$ signal sample because with increasing charged Higgs boson mass the momentum of the charged Higgs boson decreases (in the $t_{\mathrm{had}}$ rest frame). The momentum difference of a boson's decay particle emitted in or opposite to the boson's flight direction is relatively small. Considering mismeasurements, it happens more often that the quark that is emitted in the charged Higgs boson flight direction has a lower $p_{\mathrm{T}}$. As can be seen when comparing to Figure 5.7(a), this is a pure reconstruction effect. Because of the discussed reconstruction effects, the cos $(\theta^*)$-variable is especially powerful in separating signal and background events when $j_1$ or $j_2$ is $c$-tagged and $\hat{P}_{t\bar{t}}$ is high.

The impact on the decay angles caused by the different spins is also visible in other top-quark kinematic quantities, like $m_{j_1 b_{\mathrm{had}}}$, but the difference is usually not as pronounced. In addition, those variables are much more affected by the charged Higgs boson mass, whereas cos $(\theta^*)$ is independent of the charged Higgs boson mass (not considering reconstruction effects).

### 5.1.4 Non-$t\bar{t}$ backgrounds

Non-$t\bar{t}$ backgrounds mainly involve the production of a single top quark or vector bosons. These backgrounds are comparatively simple to identify. Differences between signal and non-$t\bar{t}$ backgrounds are expected to show up in any quantity describing the $t\bar{t}$ system and in correlations among these quantities. These are the masses of the reconstructed top quarks, $m_{t_{\text{lep}}}$ and $m_{t_{\text{had}}}$, the momenta of objects and angles between them, etc. Not to forget are the jet labelling probabilities $\hat{P}_{t\bar{t}}$ and $\overline{P}_{t\bar{t}}$. Even though in the previous section the jets have been labelled to yield a top-quark mass as close to $m_t$ as possible, the values still differ on average more than for true $t\bar{t}$ events.

## 5.2 Multivariate classifier

A multivariate analysis tool is exploited to combine the previously presented quantities (cf. Section 5.1) into one powerful discriminant. In a supervised training procedure a classifier is developed that identifies signal- and background-like phase space regions. In this analysis, BDTs are utilised. BDTs are a high-level and powerful machine learning model and are popular in high-energy particle physics analyses. The operating principles of BDTs are introduced in Section 5.2.1.

Besides BDTs, the usage of neural networks is also popular. Neural network models like convolutional neural networks are not well applicable to this analysis, as the discussed quantities are mostly high-level quantities, constructed from many reconstructed quantities, and usually without an apparent correlation. A feedforward deep neural network architecture would be best suited for this analysis. If trained well, BDTs and feedforwards neural networks yield similar results, but BDTs provide better "out-of-the-box" performance, making it generally easier to work with and optimise them.

After a quick introduction to BDTs, Sections 5.2.2 and 5.2.3 describe the optimisation of the classification BDT in this analysis, and Section 5.2.4 presents their performance.

### 5.2.1 Operating principle of Boosted Decision Trees

BDTs are an ensemble of decision trees that are turned into a powerful predictor with the help of boosting. The objective of BDTs is to predict a target variable $y$, in our case the class labels 1 for signal and 0 for background events, of unknown events given multiple input features $x$. The following explanations are based on References [233–236] and focus on the use case and implementation in this analysis, i.e. a binary classification task using gradient boosting as implemented in the *XGBoost* framework [236]. The XGBoost framework has been used for training and application of the BDTs in this analyis.

#### Decision tree

A decision tree is a sequence of yes/no decisions based on cuts ($c$) on observables. An exemplary decision tree is visualised in Figure 5.8. The decision tree partitions the phase space in disjoint regions $R_j$ – represented by the leaf nodes. From a mathematical point of view a decision tree can be expressed as

$$T(x;\Theta) = \sum_{j}^{J} w_j I(x \in R_j),$$

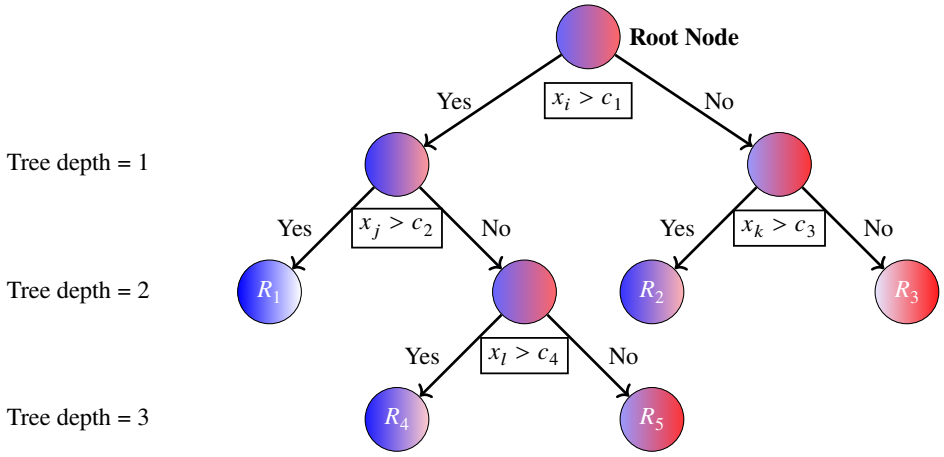



Figure 5.8: Schematic of a decision tree. $x$ refers to observables and $c$ are cut values on them that define the splittings. The colours blue and red should visualise the signal (S) and background (B) contributions at each node.

with parameters $\Theta = \{R_j, w_j\}_j^J$. $J$ is equal to the number of regions, and $w_j$ is a constant score assigned to each region. $I$ is defined to be 1 if the event falls in region $R_j$ and 0 otherwise.

The challenge lies in finding the regions that maximally separate signal and background events ($\hat{\Theta}$). In general, it is intractable to evaluate all possible region definitions. Hence, decision trees are built in a greedy top-to-bottom manner. The nodes are split recursively until a stop criterion is reached, e.g. a maximum tree depth. The stop criteria limit the complexity to avoid overfitting of the tree to retrieve a general predictive model. The optimal node splittings are evaluated by using an impurity measurement of the daughter nodes that has to be minimised. In the gradient boosting use case, suitable objective functions are defined that measure the difference between the prediction and the target.

#### Gradient boosting of decision trees

Boosting is about repeatedly applying a weak predictive algorithm, in our case a decision tree, to a modified version of the data to retrieve a powerful predictor. This is done in a greedy forward stagewise manner by adding new trees to the model without changing the previous trees. The model prediction is given by

$$\hat{y}(x) = \sum_{m=1}^{M} \eta T(x; \Theta_m),$$

where $M$ is the number of iterations, or boosting steps, and $\eta$ is the step size shrinkage parameter ($\eta \in [0., 1.]$), defines the learning rate of the boosting process and prevents the model from overfitting. At each boosting step,

$$\hat{\Theta}_m = \underset{\Theta_m}{\operatorname{argmin}} \sum_{i=1}^{N_{\text{events}}} L(y_i, \hat{y}_i^{(m-1)} + T_m(x_i; \Theta_m)), \tag{5.1}$$

has to be solved, given the current model prediction $\hat{y}_i^{(m-1)}$, and $L$ being the objective (loss) function[1]. Each tree at iteration $m$ corrects for differences – fits the residuals – between the current model prediction and the true prediction $y_i$. The residuals are determined by taking the gradient of the objective function, given the current model.

In XGBoost this is implemented via an objective function that contains a differential loss function ($l$) and a regularisation term ($\Omega$) that controls the complexity of the model to avoid overfitting,

$$L = \sum_i l(y_i, \hat{y}_i) + \sum_j \Omega(T_j) \qquad \text{with} \quad \Omega_j(T_j) = \gamma J + \frac{1}{2}\lambda w_j^2. \tag{5.2}$$

The loss function in this analysis is a binomial log-likelihood – also referred to as logistic loss or cross-entropy – (XGBoost setting "binary:logistic"),

$$l(y_i, \hat{y}_i) = y_i \ln(1 + e^{-\hat{y}_i}) + (1 - y_i) \ln(1 + e^{\hat{y}_i}).$$

The $\gamma$ parameter in Equation 5.2 defines a minimum loss reduction that is required for a node splitting, and $\lambda$ is an L2 regularisation parameter. Another parameter that is used in this analysis to limit the model complexity is the "min_child_weight" parameter in XGBoost which sets a lower bound on the number of (weighted) events in each node.

The tree structure is obtained by inserting the objective from Equation 5.2 into Equation 5.1. The minimisation is performed using a second-order approximation from which a simplified (ignoring constant terms) objective follows,

$$\tilde{L}^m = -\frac{1}{2}\sum_{j=1}^{J} \underbrace{\frac{\left(\sum_{x_i \in R_j} g_i\right)^2}{\sum_{x_i \in R_j} h_i + \lambda}}_{S} + \gamma J, \quad \text{with} \quad g_i = \partial_{\hat{y}_i^{(m-1)}} l(y_i, \hat{y}_i^{(m-1)}), \quad h_i = \partial^2_{\hat{y}_i^{(m-1)}} l(y_i, \hat{y}_i^{(m-1)}).$$

$g_i$ and $h_i$ are the first and second order derivatives of the loss function evaluated at the previous model. $\tilde{L}^m$ is like an impurity score, and the loss reduction (gain) of a split is given by

$$L_{\text{gain}} = \frac{1}{2}(S_{\text{daughter,1}} + S_{\text{daughter,2}} - S_{\text{mother}}) - \gamma.$$

The split that maximises $L_{\text{gain}}$ is used to partition the node. If the maximum gain is negative $L_{\text{gain}} < 0$, the node is not partitioned, and it becomes a final node. The scores for the finale nodes are given by

$$w_j = -\frac{\sum_{x_i \in R_j} g_i}{\sum_{x_i \in R_j} h_i + \lambda}.$$

In the tree-building process, all possible splits have to be evaluated in order to find the optimal split ($L_{\text{gain}}$). But this is unfeasible for large datasets. XGBoost has an approximate algorithm implemented that considers only certain splits. The split points are based on quantiles of the input features. This

[1] For simplicity, event weights are not taken into account in this section. An integration of event weights in the formalism is straightforward.

analysis uses the XGBoost setting "hist", which defines the split points once globally – another option reevaluates them for every node – with "max_bin" option set to 1024. This leads to approximately 1024 possible split points. The max_bin value has been optimised to not show degradation in performance with respect to larger max_bin values while being as low as possible for fast training.

The final probability of a BDT to classify an event as a signal event is referred to as the *BDT score* and is given by

$$\text{BDT score} = \frac{1}{1 + e^{-\hat{y}_M(x)}}$$

where $\hat{y}_M$ is the prediction at the $M$-th (last) tree.

Positive characteristics of BDTs are that their performance is unaffected by weak input features. In addition, input features don't have to be transformed because the loss is unrelated to the absolute feature value.

#### BDT training and application

In this analysis separate BDT trainings are performed per signal mass point hypothesis due to the vastly changing kinematics of the $t_{\text{had}}$-quark decay products. The BDTs are trained using simulated signal[2] and background events. Only events passing the signal region selection (cf. Section 4.2) are considered in the training. All simulated backgrounds – i.e. all backgrounds except the MJ background – are considered in the training, each weighted to their expected abundance in data. Around 20 to 70 thousand signal events enter the training, depending on the mass point, and around 100 million background events. Signal and background events are overall normalised to a common value to obtain class balance in the training.

The data is split into training and testing samples. The training sample is used to build the BDTs as described in the previous section. The evaluation of the BDT on the testing sample yields an unbiased prediction. This is implemented by splitting the data into multiple folds. Each fold is used once for testing, and the other folds are used for training of the BDTs. This procedure is known as cross-validation [233]. The optimal number of folds depends on the specific application.

The performance of a BDT is evaluated using the area under the receiver-operating-characteristics curve (AUC) $\{\text{AUC score} \in \mathbb{R} | 0 \leq \text{AUC score} \leq 1\}$. The receiver-operating-characteristics (ROC) curve plots the true positive rate versus the false positive rate. For any cut on the BDT score, the true positive rate is given by the fraction of signal events with a BDT score greater than the given cut value (signal efficiency), and the false positive rate is given by the fraction of background events with a greater BDT score (1 − background efficiency). The higher the AUC score, the better the classifier. A random classifier achieves an AUC score of 0.5.

Tricky is the determination of the performance of a BDT input feature. Commonly used metrics are (i) the number of splits a feature has been used in, (ii) the total or average gain per split of a feature, and (iii) the total or average sum of second-order derivatives of the loss function with respect to a feature

[2]The BDT training has been carried out with reduced signal sample size using 1 million of the 1.8 million generated signal events (55.6 %). The reason for this is that these studies have been performed for an ATLAS publication, and a signal extension was granted only after the analysis unblinding. A retraining of the BDTs would have changed the analysis. No retraining has been performed for the presentation in this thesis because the change is expected to be small.

(cover). The most precise but time-consuming method is to perform separate trainings where one feature is left out at a time and then compare the performance with the nominal approach.

### 5.2.2 Input feature optimisation

Due to the good out-of-the-box performance of BDTs, the choice of the input features has a larger influence on the BDTs' performance than the set of BDT hyperparameters (tree depth, learning rate, regularisation parameters, etc.). Therefore, the optimal input feature set is determined first using a default hyperparameter set. A dedicated optimisation using the final feature set is performed afterwards as described in Section 5.2.3.

#### Examined features

The first step in determining the optimal input feature set was collecting all quantities that could help in distinguishing signal and background events. A discussion on which quantities show differences and why has been presented in Section 5.1. A comprehensive list of all considered quantities is given in Table 5.1. The list contains 69 quantities and contains any quantity describing the $t\bar{t}$-system kinematics, flavour-tagging information of the $t\bar{t}$-system labelled jets, overall number of heavy-flavour tags, and spin-sensitive quantities.

Any quantity describing the $t\bar{t}$-system kinematics is considered, even if it doesn't show any discriminating power, because it might be helpful through, maybe non-obvious, correlations with other quantities. Combinations of quantities have been tested in an earlier, lightweight feature optimisation study. An example of this is the ratio $p_{\mathrm{T},j_1}/p_{\mathrm{T},b_{\mathrm{had}}}$. The two quantities are anti-correlated, and their ratio shows a larger discriminative power than the separate quantities. Combined quantities built from quantities with linear correlations, like the one presented above, are not included in the final optimisation study, because including them in the training did not translate into an improved BDT performance. The BDT efficiently learns these rather simple correlations itself. This is in contrast to quantities with more complex correlations, like the invariant mass of objects and angular distance measures $\Delta R$ between objects. These quantities are still included in the final optimisation study.

A remark has to be made on the flavour-tagging information. The input feature optimisation has been performed at an earlier stage of the analysis where only the PCBT calibration (cf. Section 3.4.6) had been available. But there were already plans to test new flavour-tagging calibrations in the future which include $c$-tagging WPs. In order to have a fair comparison between the two calibrations later on, it was decided to include all flavour-tagging information in the final input feature set, regardless of their performance.

Using a different flavour-tagging calibration might also change the importance of non-flavour-tagging features in the training due to different correlations. In principle the input feature optimisation has to be performed for each flavour-tagging calibration. In the case of the PCFT calibration the optimisation has not been performed again because the looser $b$-tagging WPs partially also tag $c$-jets. The correlations will be similar, and significant changes in the variable importance are not expected.

#### Optimisation procedure

The final set of input features is obtained by training a BDT with all quantities listed in Table 5.1 and recursively removing the least performing features until a statistically significant loss in performance

Table 5.1: List of variables tested as input features to the BDT. $\cos\theta^*(j_1)$ ($\cos\theta^*(j_2)$) is calculated similar to $\cos(\theta^*)$, except that the angle $\theta^*$ is defined to be always measured with respect to $j_1$ ($j_2$), independent of their flavour-tagging score or $p_{\mathrm{T}}$.

| Category | Features |
|---|---|
| Four-momentum of individual objects | $(E, p_{\mathrm{T}}, \eta, \phi)$ of $\ell$, $j_1$, $j_2$, $b_{\mathrm{had}}$, $b_{\mathrm{lep}}$ |
| $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ and $W$ related | $E_{\mathrm{T}}^{\mathrm{miss}}$, $\phi^{\mathrm{miss}}$, $E_{\mathrm{T}}^{\mathrm{miss}}$ significance, $\Delta\eta(\ell,\nu)$, $m_{\mathrm{T,W}}$, $m$ and $p_{\mathrm{T}}$ of $W$ |
| $\Delta R$ between two objects | $\Delta R(\ell,\nu)$, $\Delta R(j_1,j_2)$, $\Delta R(j_1,b_{\mathrm{had}})$, $\Delta R(j_2,b_{\mathrm{had}})$, $\Delta R(b_{\mathrm{lep}},b_{\mathrm{had}})$, $\Delta R((j_1,j_2),b_{\mathrm{had}})$, $\Delta R((j_1,b_{\mathrm{had}}),j_2)$, $\Delta R((j_2,b_{\mathrm{had}}),j_1)$ |
| Invariant mass of two objects | $m_{j_1 j_2}$, $m_{j_1 b_{\mathrm{had}}}$, $m_{j_2 b_{\mathrm{had}}}$ |
| Four-momentum of $H^{\pm}/W_{\mathrm{had}}$ | $(E, p_{\mathrm{T}}, \eta, \phi)$ of $(j_1,j_2)$ |
| Four-momentum of top quarks | $(m, p_{\mathrm{T}}, \eta, \phi)$ of $t_{\mathrm{had}}$, $t_{\mathrm{lep}}$ |
| Other top-quark kinematics | $p_{b_{\mathrm{had}}}^{t_{\mathrm{had}}-\mathrm{rest}}$, $\Delta R(t_{\mathrm{had}},t_{\mathrm{lep}})$, $m_{t\bar{t}}$ |
| Spin | $\cos(\theta^*)$, $\cos\theta^*(j_1)$, $\cos\theta^*(j_2)$ |
| $t\bar{t}$-system reconstruction | $\hat{P}_{t\bar{t}}$, $\overline{P}_{t\bar{t}}$ |
| Event level | Lepton-channel, $N_{\mathrm{jets}}$, $S_{\mathrm{T}}$ |
| Flavour-tagging scores | PCBT-scores of $j_1$, $j_2$, $b_{\mathrm{had}}$, $b_{\mathrm{lep}}$ |
| Number of heavy-flavour jets | $N_{b\text{-tag }85}$, $N_{b\text{-tag }77}$, $N_{b\text{-tag }70}$, $N_{b\text{-tag }60}$ |

is observed. The aim is to remove non-discriminating features as all features part of the final BDT training have to be well modelled in order to retrieve reliable results. It is assumed that for all signal samples the same set of features is useful in discriminating signal and background events. The top-quark kinematics change with the mass of the charged Higgs boson, but the affected quantities are always the same, independent of the absolute mass value. The only exception is the 80 GeV mass point, where $m_{H^{\pm}} \approx m_W$. For this mass point, top-quark kinematic quantities show little difference between signal and background. But since additional weakly discriminating features in the BDT don't degrade its performance, no separate optimisation is performed for the 80 GeV mass point. The input feature optimisation is performed using the signal sample for the 130 GeV mass point as it lies in between the $W$-boson and $t$-quark masses. The intermediate mass point covers relatively well both cases of almost identical and vastly different top-quark kinematics.

The BDT trainings are carried out using five-fold cross-validation on a randomly shuffled dataset. The performance of an input feature set is measured by calculating the mean and standard deviation of the AUCs from the five folds. The hyperparameters were set to a default configuration in all trainings. This configuration has been determined by performing a lightweight grid search of the hyperparameters in BDT trainings with all features included.

In order to save time and resources, only one fourth of the background events are used in these trainings. The reduced background sample size still includes about 13 million $t\bar{t}$ events and 12 million events of other backgrounds. This suffices to model the backgrounds well and is anyhow much more compared to the signal sample size of a few ten-thousands.

First, a feature ranking is produced based on the performance of features in trainings where only the given feature is excluded from the training. Afterwards, the lowest-ranked features are tested for removal first. If the AUC of the BDT without the given feature did not degrade by more than 0.03 % – this

corresponds to about $0.6\,\sigma$ – the feature was permanently removed from the input feature set. Features that already showed a clear drop in performance during the ranking procedure have not been tested for removal again. The flavour-tagging features have not been tested for removal for the reasons outlined above.

From this primal optimised set, all removed features are tested for re-adding, and kept features are tested for removal. This was done to reevaluate the performance of features on the changed dataset – with respect to where they have been tested for removal first – and to reduce the effects from statistical fluctuations in the training performance. The results turned out to be compatible with the initial optimisation, and the optimised set was left unchanged. This justifies that the used methodology is very robust.

#### Final BDT input feature set

The final feature set is summarised in Table 5.2 (including the new PCFT calibration quantities). The distributions of all features, comparing the stacked background prediction with the signal prediction and data, are shown in Appendix A.4. The optimised set is very much up to expectations. It contains the $p_{\mathrm{T}}$ of the $t_{\mathrm{had}}$-quark associated jets, invariant masses of two of these jets, $p_{b_{\mathrm{had}}}^{t_{\mathrm{had}}-\mathrm{rest}}$, and $\cos(\theta^*)$ to separate signal from SM $t\bar{t}$ events. The quantities related to $t_{\mathrm{lep}}$ and $m_{t\bar{t}}$ can suppress non-$t\bar{t}$ backgrounds. Here it is interesting to note that the lepton $p_{\mathrm{T}}$ and $W$-boson mass are sufficient to reject backgrounds without a $W \to \ell\nu$ decay, and $E_{\mathrm{T}}^{\mathrm{miss}}$ as well as the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance do not significantly improve the rejection of these backgrounds. It is surprising that the mass and momentum of $t_{\mathrm{lep}}$ are included in the final set while this is not the case for $t_{\mathrm{had}}$. This is presumably related to the fact that many other quantities describing the $t_{\mathrm{had}}$-quark kinematics are part of the final set, like the invariant masses of any combination of two jets and all momenta of final jets. The $\Delta R(b_{\mathrm{lep}}, b_{\mathrm{had}})$ quantity helps in rejecting any type of background. Other $\Delta R$ quantities did not provide extra information on the invariant mass quantities and are not part of the final set.

The $N_{\mathrm{jets}}$ and $S_{\mathrm{T}}$ quantities provide information via correlations with other variables. As an example, the $b_{\mathrm{had}}$-jet $p_{\mathrm{T}}$ is expected to be small for signal events with a charged Higgs boson mass close to the top-quark mass. If the event has in addition a large $S_{\mathrm{T}}$ value, it is even less likely to observe such a low $b_{\mathrm{had}}$-jet $p_{\mathrm{T}}$ in $t\bar{t}$ events. The $\overline{P}_{t\bar{t}}$ quantity helps the BDT to identify wrong jet labellings.

The flavour-tagging quantities are helpful in separating $t\bar{t}$ backgrounds, as discussed in Section 5.1.2, and non-$t\bar{t}$ backgrounds. They also add information via correlations with $t\bar{t}$ kinematic quantities. For instance, if the $b_{\mathrm{had}}$-jet passes the tight $b$-tagging WP, the BDT can put more emphasis on the $p_{\mathrm{T},b_{\mathrm{had}}}$ quantity.

Figure A.5 in the Appendix exemplarily shows the linear correlations of the input features for the $H_{130}^{\pm}$ signal sample. Most correlations are small (below 10 %). The largest correlations are observed among some top-quark kinematic quantities and among flavour-tagging quantities, with linear correlations up to about 70 %. The results are similar for other signal and background samples.

Figure 5.9 shows the loss in AUC when removing one of the features from the final feature set. The higher the reduction in AUC, the more unique information it adds to the BDT training. The ordering from top to bottom can be seen as a ranking of the features in the final set. The $m_{j_1 j_2}$ quantity is by far the most important. Surprising are the high rankings of the $m_{t_{\mathrm{lep}}}$ and $p_{\mathrm{T},b_{\mathrm{lep}}}$ features. $m_{t_{\mathrm{lep}}}$ yields unique information in rejecting non-$t\bar{t}$ backgrounds and information about $m_{t_{\mathrm{had}}}$ in correlations with $\overline{P}_{t\bar{t}}$. The

Table 5.2: List of input features used in the final BDT training.

| Variable type | Variable name | Definition |
|---|---|---|
| **Top-quark kinematic variables** | | |
| $t_{\text{had}}$ | $p_{\text{T},j_1}$ | $p_{\text{T}}$ of $j_1$ |
| | $p_{\text{T},j_2}$ | $p_{\text{T}}$ of $j_2$ |
| | $p_{\text{T},b_{\text{had}}}$ | $p_{\text{T}}$ of $b_{\text{had}}$-jet |
| | $p_{b_{\text{had}}}^{t_{\text{had}}-\text{rest}}$ | Momentum of $b_{\text{had}}$-jet in $t_{\text{had}}$ rest frame |
| | $m_{j_1 j_2}$ | Invariant mass of $j_1$+$j_2$ jets |
| | $m_{j_1 b_{\text{had}}}$ | Invariant mass of $j_1$+$b_{\text{had}}$ jets |
| | $m_{j_2 b_{\text{had}}}$ | Invariant mass of $j_2$+$b_{\text{had}}$ jets |
| | $\cos(\theta^*)$ | Boson spin sensitive variable |
| $t_{\text{lep}}$ | $p_{\text{T},b_{\text{lep}}}$ | $p_{\text{T}}$ of $b_{\text{lep}}$-jet |
| | $p_{\text{T},\ell}$ | $p_{\text{T}}$ of lepton |
| | $m_W$ | Invariant mass of reconstructed $W$ boson |
| | $m_{t_{\text{lep}}}$ | Invariant mass of reconstructed $t_{\text{lep}}$ |
| | $p_{\text{T},t_{\text{lep}}}$ | $p_{\text{T}}$ of reconstructed $t_{\text{lep}}$ |
| $t\bar{t}$ system | $\Delta R(b_{\text{lep}}, b_{\text{had}})$ | $\Delta R$ between the $b_{\text{lep}}$- and $b_{\text{had}}$-jet |
| | $t\bar{t}$ mass | Invariant mass of $t_{\text{had}}$+$t_{\text{lep}}$ |
| **Event variables** | | |
| Event level | $N_{\text{jets}}$ | Number of jets in the event |
| | $S_{\text{T}}$ | Scalar $p_{\text{T}}$ sum of all calibrated objects |
| | $\overline{P}_{t\bar{t}}$ | Probability of jets being labelled correctly normalised by all combinations in the event |
| **Flavour-tagging variables** | | |
| Flavour-tagging score | PCFT score $j_1$ | PCFT score of $j_1$ |
| | PCFT score $j_2$ | PCFT score of $j_2$ |
| | PCFT score $b_{\text{had}}$ | PCFT score of $b_{\text{had}}$-jet |
| | PCFT score $b_{\text{lep}}$ | PCFT score of $b_{\text{lep}}$-jet |
| Number of tags | $N_{c\text{-tag loose}}$ | Number of jets passing loose $c$-tag WP ($b$-veto) |
| | $N_{c\text{-tag tight}}$ | Number of jets passing tight $c$-tag WP ($b$-veto) |
| | $N_{b\text{-tag loose}}$ | Number of jets passing 70 % $b$-tag WP |
| | $N_{b\text{-tag tight}}$ | Number of jets passing 60 % $b$-tag WP |

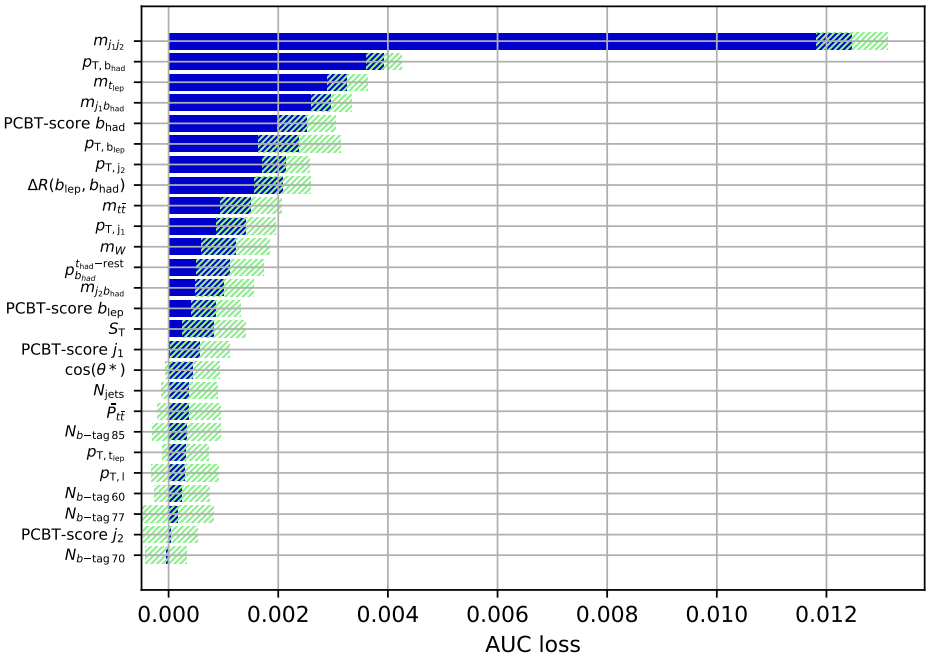


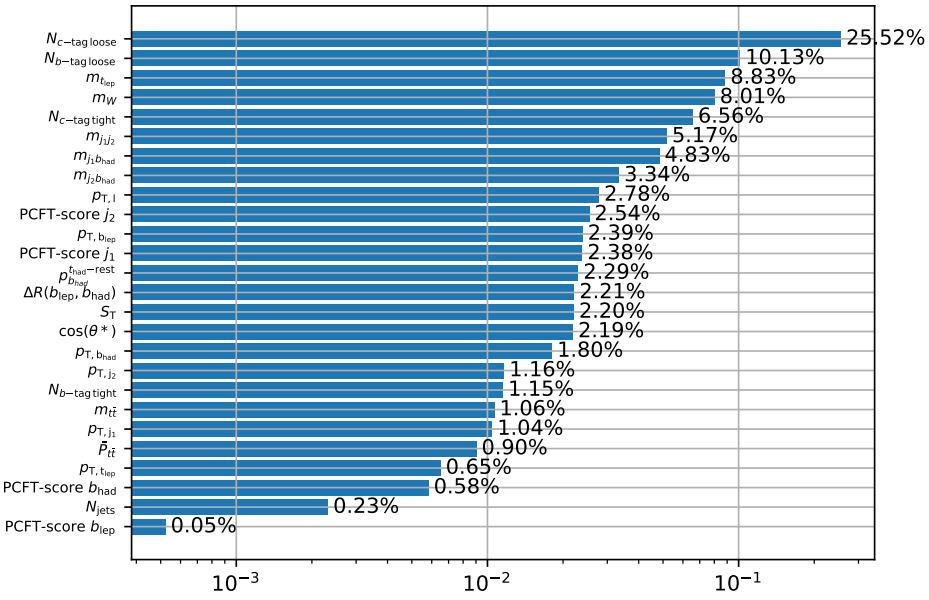

Figure 5.9: Loss in AUC, compared to the training with the optimised feature set, when leaving out one feature in the BDT training. The BDT training setup is identical to the setup used for the feature optimisation studies. The green hashed band represents the standard deviation on the AUC mean from the five-fold cross-validation.

$p_{\mathrm{T},b_{\mathrm{lep}}}$ quantity helps in recovering discriminative power in case of wrong jet labellings. No difference in the $p_{\mathrm{T},b_{\mathrm{lep}}}$ distribution is expected between signal and $t\bar{t}$ backgrounds. But the $p_{\mathrm{T},b_{\mathrm{lep}}}$ distribution is softer for higher charged Higgs boson masses due to an on-average smaller momentum of the true $b_{\mathrm{had}}$-jet, which may get wrongly labelled as $b_{\mathrm{lep}}$. Regarding the flavour-tagging information, it has to be noted that the rankings in Figure 5.9 were produced with the PCBT calibration.

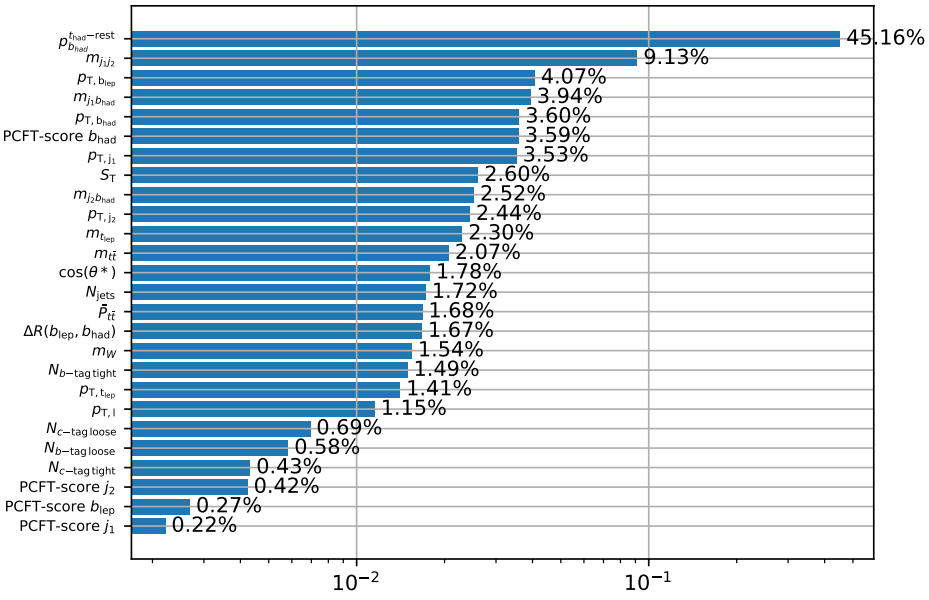
Figure A.6 in the Appendix compares the feature importances for the metrics introduced in Section 5.2.1 – calculated from the final trainings with the PCFT calibration. The ranking of the features depends highly on the considered metric. Figure 5.10 compares the total gain related to a feature – in my opinion the most meaningful metric – for the 80 GeV and 150 GeV signal mass point trainings. In this figure the expected behaviour can be observed that for the 80 GeV mass point the flavour-tagging quantities are most important for separating signal and background events but are not as important for larger charged Higgs boson mass hypotheses.

### 5.2.3 Final BDT training configuration and hyperparameter optimisation

The final BDT trainings are carried out using five-fold cross-training. Using five folds is a compromise between having a lot of training data (80 %) and not having to perform too many trainings. The events are split into the five folds based on their event numbers. This resembles a random splitting. The five-fold cross-training leads to five BDTs that are trained per signal mass point – in total 60 BDTs. The final BDT scores of simulated and real data events are evaluated using the BDT belonging to their event number.




(a)




(b)

Figure 5.10: Feature importance based on the total gain in loss by node splittings based on the given feature for the final BDT training with the (a) $H^{\pm}_{80}$ and (b) $H^{\pm}_{150}$ signal sample. The overall feature importance is normalised to 1.

The last step before performing the final trainings is the optimisation of the BDT hyperparameters. The optimal hyperparameters are expected to depend on the signal mass point hypothesis because the discriminative power of features varies with $m_{H^{\pm}}$. On the other hand, the optimal hyperparameters for each fold of the same signal mass point hypothesis are expected to be similar and only differ by statistical variations.

The optimal set of hyperparameters is determined by performing nested cross-validation. For each fold, the other four folds are used to determine the optimal hyperparameters by performing four-fold cross-validation. This avoids any bias between the hyperparameter optimisation dataset and the final evaluation dataset.

The hyperparameter set optimisation is carried out with the help of the *Hyperopt* [237] tool. The Hyperopt tool performs a Bayesian optimisation on an objective function using Tree-structured Parzen Estimators [238]. The objective that is maximised in this analysis is chosen to be the mean AUC of the four-fold cross-validation with the BDT hyperparameters being the adjustable parameters of the objective function. In an iterative procedure the algorithm learns from previous trainings in which part of the parameter space the optimum lies and can propose profound parameter sets for the next iteration. This makes it not only more efficient than random or grid searches but also leads in general to better results.

In this analysis the number of iterations – tested hyperparameter sets – is limited to 30. This is a rather small number, but the number of iterations is limited by the large dataset size and the long training time of the BDTs – running 30 iterations took around three days. A parallelisation is not possible since the subsequent tested hyperparameter set depends on the previous trainings.

In the Hyperopt tool a search space has to be defined for the hyperparameters. The examined search spaces are summarised in Table 5.3. For the number of boosting rounds, early stopping was used. If the mean AUC on the validation samples didn't improve for five consecutive boosting rounds, the training was stopped, and the number of boosting rounds was set to the boosting round with the highest mean AUC. The utilised search spaces are either *uniform* or *loguniform* (in the *quniform* search space, the drawn values are rounded to integer values). In the loguniform search space, a value according to exp(uniform(low; high)) is drawn so that the logarithm of the returned values are uniformly distributed. Also given in Table 5.3 are the actual search ranges that follow from the given search spaces. With the choice of the appropriate search space, prior knowledge on the expected value of a hyperparameter can be handed over to the Hyperopt algorithm. For example, the learning rate is, in most cases, rather small, around a few percent. By choosing the loguniform search space, also a few larger values can be tested without skewing the search range.

The choice of the search spaces and ranges is based on discussions from textbooks [233] and the previously lightweight hyperparameter search that has been performed for the input feature optimisation studies. The parameters used in the input feature optimisation studies are listed under "Baseline" in Table 5.3.

As anticipated, the optimal set of hyperparameters for the trainings of the different folds with the same signal sample training were found to be very similar and the performances were statistically indistinguishable. For the sake of simplicity, a common set of hyperparameters was adopted for all folds in each signal mass point training. The final hyperparameters for the $H^{\pm}_{80}$ and $H^{\pm}_{130}$ signal samples are listed in Table 5.3, too. The hyperparameters of other signal samples are relatively close to the $H^{\pm}_{130}$ signal sample. For the $H^{\pm}_{80}$ signal sample, the number of boosting rounds is smaller and the learning rate is larger because many input features show almost no separation for this mass point.

Table 5.3: The BDT hyperparameters used in the optimisation and their scan range. Also shown are the optimal hyperparameters for the trainings with the $H^{\pm}_{80}$ and $H^{\pm}_{130}$ signal samples. The baseline values have been used in the input feature optimisation studies.

| Parameter | Search space and range | | Baseline | $H^{\pm}_{80}$ | $H^{\pm}_{130}$ |
|---|---|---|---|---|---|
| Boosting rounds | Early-Stopping 5 rounds | | | 175 | 725 |
| Maximum tree depth | quniform( 3; 15) | 3 to 15 | 8 | 4 | 9 |
| Learning rate ($\eta$) | loguniform( -7; -0.9) | 0.0009 to 0.4066 | 0.05 | 0.0405 | 0.0129 |
| Minimum loss reduction ($\gamma$) | uniform( 0; 200) | 0 to 200 | 5.0 | 9 | 7 |
| Minimum sum of weights in a region (min_child_weight) | loguniform( -5; 11) | 0.0067 to 59 874 | 1.0 | 260 | 2000 |

In general, learning rates between 0.01 to 0.05 were found to show the best results. The number of boosting rounds is anti-correlated to the learning rate. Tree depths between 9 to 14 show the best performance for most trainings. These are rather large tree depths, which is a consequence of almost all input features having a direct correlation with each other. Regarding regularisation, the impact of the $\gamma$ parameter was found to be negligible. But the min_child_weight parameter turned out to be effective with values around a few thousand. This corresponds to a minimum number of (weighted) events in each final region of a few per mille (signal and background events have been normalised to 1 million each).

As already outlined in the beginning, for the final BDT trainings, all four folds, which have been used to determine the hyperparameters, are used as training data.

### 5.2.4 Final BDT scores

The final BDT score distributions for the trainings with the $H^{\pm}_{80}$, $H^{\pm}_{100}$, $H^{\pm}_{130}$, and $H^{\pm}_{150}$ signal samples are shown in Figure 5.11. The BDT scores of the other mass point trainings are shown in Appendix A.3. The agreement between data and prediction is generally good.

The discrimination between signal and background is as expected lowest for the 80 GeV mass point. A three-peak structure is visible for signal and background events that can be attributed to the charm-tagging WPs (from left to right): untagged, loose, and tight $c$-tagged events. The $t\bar{t}(ud)$ contribution is heavily suppressed at large BDT score values compared to the $t\bar{t}(cs)$ background.

With increasing charged Higgs boson mass the discrimination increases, too. But the signal distributions always maintain a relatively large tail towards smaller BDT scores. This is caused by signal events in which not all jets are labelled correctly. These events then exhibit signal- and non-signal-like characteristics. The BDT cannot clearly assign these events to one class, and they receive an intermediate BDT score.

The discrimination is highest for the 150 GeV mass point. One would expect the discrimination to be even larger for the 160 GeV and 168 GeV mass points. But the fraction of events that contain the true reconstructed $b_{\text{had}}$-jet is considerably reduced for these mass points, which decreases their overall discriminative power.

Figure 5.12 shows the ROC curves for selected mass points. All ROC curves are smooth. The AUC increases rapidly for charged Higgs boson mass point hypotheses more distant from the $W$-boson mass. Once the mass difference $m_{H^{\pm}} - m_W$ is larger than the dijet mass resolution, the increase in AUC flattens

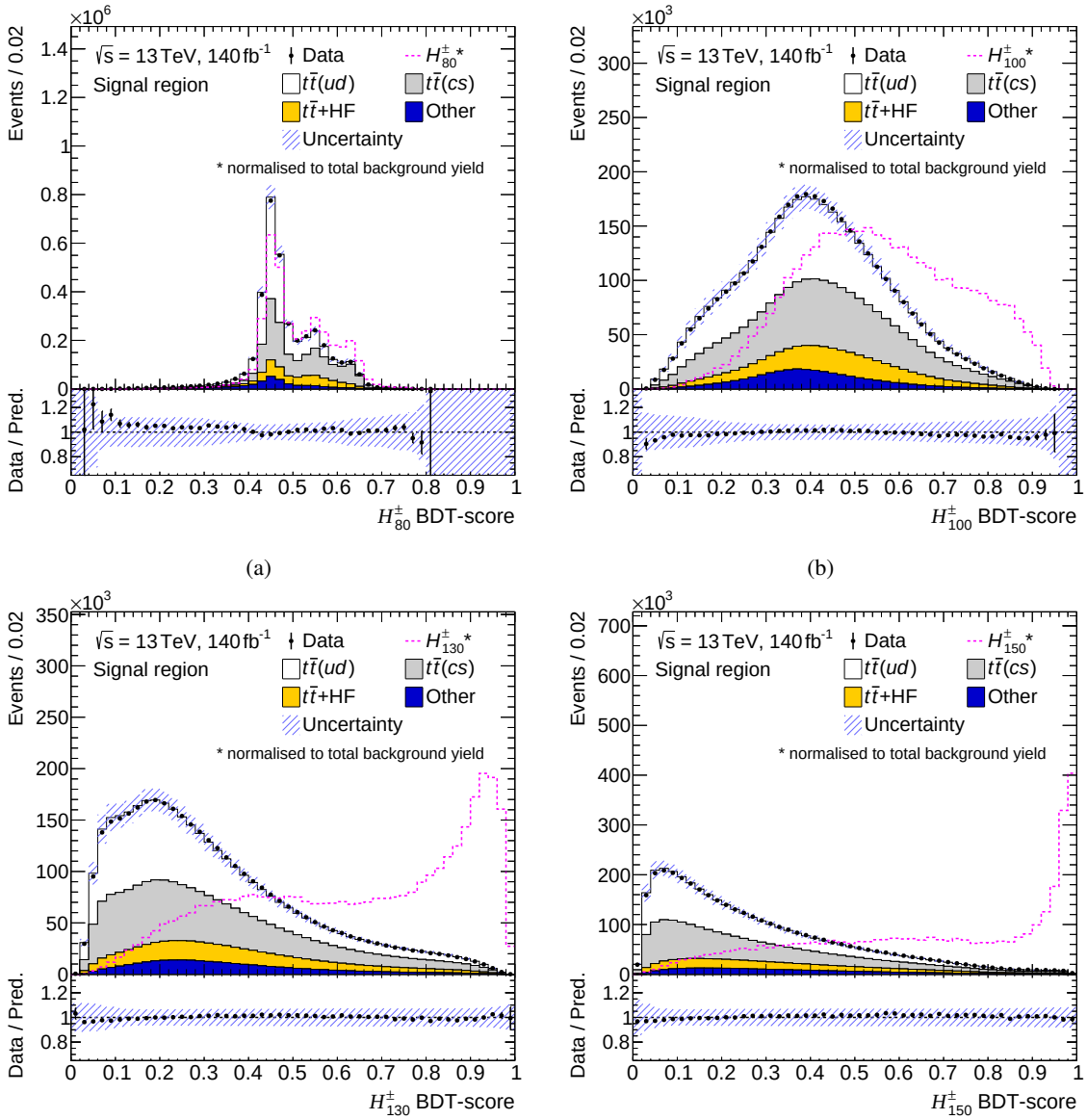



Figure 5.11: BDT score distributions for the trainings with the (a) $H^{\pm}_{80}$, (b) $H^{\pm}_{100}$, (c) $H^{\pm}_{130}$, and (d) $H^{\pm}_{150}$ signal samples.

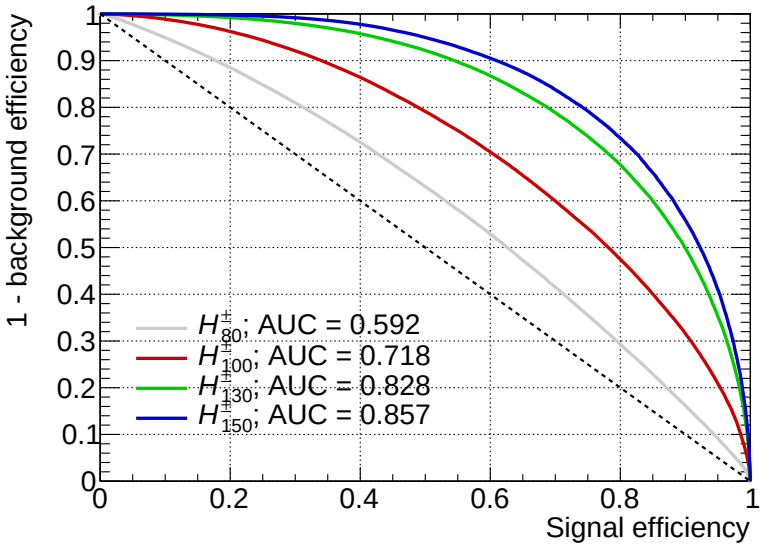



Figure 5.12: ROC curves of selected BDT score distributions. The black dashed line indicates the ROC curve for a distribution without discriminative power.

out. For mass points close to the top-quark mas, the discussed $b_{\text{had}}$-jet effect reduces the performance.

## 5.3 Outlook

In the following, two concepts will be presented that might improve the presented BDT training in order to increase the signal sensitivity of the analysis. On the one hand, the utilised loss function (cross-entropy) favours an overall separation between signal and background. This is beneficial in order for the BDT to learn the different characteristics signal and background events exhibit. A different loss function which promotes gains with very pure signal nodes – and in accordance puts less focus on having pure background nodes – will make the BDT put more focus on the high BDT score region in the training. This will decrease the overall separation of signal and background but increase the signal sensitivity in the high BDT score region. This analysis might profit from this as having a very pure signal region is important in the statistical interpretation of the results as presented in the following chapter.

The other concept is to include systematic uncertainties in the BDT training [239]. The predictions that are being used to train the BDT are based on models that aim at simulating nature. We know these models are not perfect and systematic uncertainties are related to them. The systematic uncertainties are considered in the statistical interpretation of the results and weaken the derived results. By making the BDT aware of these uncertainties in the training, it can put less focus on quantities that are not well modelled/understood. There are multiple ways to achieve this. As an example, the loss function can be modified to penalise the gain in node splits if differences are seen between the nominal and systematically varied quantities. Or the weight of a tree can be scaled down if its response differs significantly for systematic variations. Another way is to augment the training data to include events multiple times with varied systematic uncertainties. Neural networks can also use an adversarial training setup to not be prone to systematic uncertainties.

Another way to improve the analysis in the future might be by using other machine learning tools, which in this thesis have not been explored due to time constraints.

Due to the unknown mass of the charged Higgs boson, parametric neural networks (PNNs) are promising. The general idea of PNNs is to condition the learning of the network on some representation, in this case $m_{H^{\pm}}$. This is realised by including $m_{H^{\pm}}$, similar to the input features, as a learnable parameter in the PNN training. The final prediction of the PNN depends on the specified value for $m_{H^{\pm}}$. There are several advantages of PNNs compared to single mass point classifiers, such as interpolation between the mass points. Another benefit is an increase in signal sample size in the training as the signal events of all mass points are included. BDT trainings have been carried out with the full signal sample size after the signal extension. The extension increases the signal sample size by a factor of 1.8. The trainings showed improvements in the AUC by around 0.2 % for all mass points. This transitions to an improved signal sensitivity (measured using expected upper limits; for its definition see the following chapter) by around 0 to 10 %. Including all signal mass points in the PNN training corresponds to an increase in signal sample size by a factor of 12.

PNNs are already popular and used in many analyses. In the future many analyses will probably move forward and use graph neural networks [240] and transformers [241]. They are becoming more popular in object reconstruction techniques, e.g. flavour-tagging [242]. Complex structures are possible using modern machine learning techniques, and no detailed discussion on the application in this analysis will be given. But as a general idea, at the analysis level the objects – jets, leptons and $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ – can be used as inputs to the machine learning tools, without the need for a $t\bar{t}$-system reconstruction.

In the future, if the development of neural networks continues, the networks will probably be trained using only basic detector information, i.e. hits in the tracking systems and calorimeter deposits.

CHAPTER 6

# Extraction of a potential $H^{\pm} \to cs$ contribution in data

This chapter describes the extraction of a potential signal in data with the help of the BDT scores presented in the previous chapter. A statistical model is developed that quantifies the presence of signal events in ATLAS data (Section 6.2). Systematic uncertainties impacting the analysis are discussed beforehand in Section 6.1.

## 6.1 Systematic uncertainties

Systematic uncertainties impacting this analysis are related to the detector response, reconstruction of objects, and modelling of signal and background processes. They can affect the normalisation and shape of quantities – including the BDT score. This section describes how the systematic uncertainties have been estimated. Table 6.1 summarises all considered uncertainties. The uncertainties are grouped based on their origin. As an indicator of their impact on the analysis, the uncertainties on the normalisation of the total background prediction are given in Table 6.1 in the "Impact" column.

### 6.1.1 General uncertainties

The luminosity of the combined Run-2 ATLAS *pp* collision dataset is determined with a precision of 0.83 % [133]. This uncertainty affects the overall normalisation of all simulated samples, as they are normalised to the measured luminosity.

A correction is applied to all simulated samples that accounts for different pile-up conditions in simulation and data. The correction is based on measurements of the inelastic *pp* cross-section in fiducial detector regions [243]. The uncertainty on this correction is considered a systematic uncertainty in this analysis.

### 6.1.2 Object reconstruction

Calibrations of physics objects correct for differences between simulation and data. The procedures to calibrate the objects are described in Section 3.4. The uncertainties related to the calibrations are

Table 6.1: List of systematic uncertainties considered in this analysis. "General" and "Object reconstruction" components affect all samples; "Process modelling" components affect only the given sample. The "Components" column quotes the number of systematic uncertainties. "N" and "S" stand for normalisation-only and normalisation-plus-shape uncertainty, respectively. "Smo." and "Sym." are shortcuts for smoothing and symmetrisation, respectively, and are related to the treatment in the statistical model (see Section 6.2). The "Impact" column states the normalisation uncertainty on the total background prediction of a given group of uncertainties. The uncertainty groups are indicated by the dashed lines.

| Category | Type | Components | | Smo. | Sym. | Impact |
|---|---|---|---|---|---|---|
| **General** | | | | | | |
| Luminosity | Scale | N | 1 | | | 0.9 % |
| Pile-up | Correction | S | 1 | | | |
| **Object reconstruction** | | | | | | |
| Lepton | Energy/Momentum scale & resolution | S | 7 | ✓ | | |
| | Trigger, reconstruction, identification and isolation efficiencies | S | 14 | | | 0.9 % |
| $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ | Soft-track-term scale & resolution | S | 3 | ✓ | ✓ | |
| Jets | Jet energy scale & resolution | S | 43 | ✓ | | 4 % |
| | Jet-vertex-tagger | S | 1 | | | |
| Flavour-tagging | Efficiencies & inefficiencies | S | 85 | | | 2.9 % |
| **Process modelling** | | | | | | |
| Common | Scale ($\mu_R$, $\mu_F$), $\alpha_s$, PDF | S | 4 | | | |
| Signal & top-quark | ME & PS | S | 1 | ✓ | ✓ | 5 % |
| | ISR & FSR | S | 1 | ✓ | ✓ | |
| | $t\bar{t}$ $h_{\mathrm{damp}}$ | S | 1 | ✓ | ✓ | |
| | $tW$-DS | S | 1 | ✓ | | |
| | Single-top cross-section | N | 1 | | | |
| | $t\bar{t}$ correction | S | 4 | | ✓ | |
| Weak-boson | NLO EW correction | S | 1 | | | 0.9 % |
| | CKKW & QSF | N | 1 | | | |
| | $V$+jets cross-section | N | 1 | | | |
| Multijet | MJ alternative estimation | S | 1 | | | |
| | MJ norm | N | 1 | | | |

considered as systematic uncertainties in this analysis. The uncertainties affect all simulated samples.

Energy (momentum) scale and resolution calibration uncertainties are considered for electrons (muons) [150, 156]. In addition, uncertainties on the scale factors that account for differences in the electron and muon selection efficiencies at trigger level, in the offline reconstruction, and in passing identification and isolation criteria are considered [147, 150].

Jet energy scale calibration uncertainties related to pile-up, flavour, punch-through and in-situ corrections, and jet energy resolution calibration uncertainties are considered [244, 245]. Furthermore, the uncertainty on the jet identification efficiency of the jet-vertex-tagger [162] is considered. Regarding the flavour-tagging of jets, the eigendecomposition of the uncertainties on the calibration scale factors that account for different identification and misidentification rates for $b$, $c$ and light-flavour jets in simulation and data is considered [169, 246, 247].

The energy scale and resolution uncertainties for leptons and jets are propagated to $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$. In addition, uncertainties on the $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ soft-track-term scale and resolution calibration are considered [248, 249].

### 6.1.3 Process modelling

Some modelling uncertainties are common for all samples due to the common usage of perturbation theory and PDFs in the event generation. In the calculation of the partonic cross-section (cf. Equation 3.1), the uncertainty due to missing higher orders in the perturbative expansion – i.e. the uncertainty related to the $O(\alpha_s^{n+1})$ term – is estimated by varying the renormalisation $\mu_R$ and factorisation $\mu_F$ scale parameters by a factor of 2 up and $1/2$ down. Uncertainties related to the used PDFs are considered by dedicated PDF error eigensets. The error eigensets contain multiple sources of uncertainties. Among them are uncertainties related to the measurement of the PDFs and uncertainties related to the assumed functional form in the fits. The PDF uncertainties from all eigensets are combined into one uncertainty by calculating their standard deviation with respect to the nominal set (for Hessian PDF sets, the differences are summed up in quadrature). PDF uncertainties due to missing higher orders in the DGLAP evolution equation are considered by the scale variations discussed above. Uncertainties related to measurements of the strong coupling constant $\alpha_s$ and the extrapolation of $\alpha_s$ to different energy scales based on the renormalisation group equation are assessed by two alternative PDF sets fitted with varied $\alpha_s$ values. For all processes whose cross-sections are taken from simulation, the related cross-section uncertainty is included as a normalisation uncertainty [197, 198, 250, 251].

All other MC modelling uncertainties are assessed by either comparing nominal and alternative event generators or comparing different sets of generator parameters. The modelling uncertainties of minor background processes are neglected. This is the $t\bar{t}$(allHad) process and the processes listed under "Other top" in Table 4.1. Their combined background yield is below 1 %.

#### Signal and top-quark processes

The uncertainty related to the matching of NLO ME to PS in SM $t\bar{t}$ and single-top-quark processes are evaluated with the help of samples generated with MadGraph5_aMC@NLO 2.6.0 and 2.6.2 [252] (instead of Powheg Box) using the NNPDF3.0nlo set of PDFs, respectively. The Pythia and EvtGen versions are unchanged with respect to the nominal samples.

The uncertainty related to the modelling of PS and hadronisation is assessed with samples created by Herwig 7.13 and 7.16 [253, 254] (instead of Pythia) for SM $t\bar{t}$ and single-top-quark processes, respectively. The samples were generated with the default Herwig 7.1 set of tuned parameters and the MMHT2014lo PDF set [255]. As for the nominal samples, the Powheg Box v2 generator is used for ME calculation.

The uncertainty arising from the $h_{\text{damp}}$ parameter (defined in Section 4.1.1) in the SM $t\bar{t}$ simulation is estimated by comparison with an alternative sample where the $h_{\text{damp}}$ parameter is increased by a factor of 2 ($h_{\text{damp}} = 3\,m_t$).

Signal and top-quark processes modelling uncertainties on ISR and FSR are determined by varying the couplings at ISR/FSR vertices – ISR: *Var3c* variation of the A14 PS tune, FSR: vary $\mu_R^{\text{FSR}}$ by a factor of two.

Systematic uncertainties introduced by the $t\bar{t}$ correction (cf. Section 4.3.2) are considered by performing an eigenvalue decomposition of the fit parameter uncertainty matrix and varying the eigenvectors separately by their eigenvalues.

Uncertainties on the single-top quark production cross-sections are calculated using Hathor. An uncertainty on the $t\bar{t}$ production cross-section is not taken into account, as its value is determined directly from fits to the BDT scores.

The interference between $t\bar{t}$ and $tW$ processes is handled using the diagram removal scheme [256]. In the diagram removal scheme, all Feynman diagrams in the NLO $tW$ amplitudes with two on-shell top quarks are simply removed in the nominal $tW$ sample. The uncertainty associated with this choice is assessed by an alternative method in which the NLO $tW$ cross-section is modified by a subtraction term that locally cancels the $t\bar{t}$ contribution (diagram subtraction – *DS*) [185, 256].

#### Weak-boson processes

For processes including a weak boson, uncertainties related to electroweak corrections at NLO are assessed with the help of an alternative electroweak virtual approximation in which the electroweak and QCD components are combined using an exponentiated prescription [257, 258].

The weak-boson samples in this analysis are simulated using the Sherpa event generator for ME and PS. CKKW and QSF are parameters of Sherpa that define the merging/matching scale for jets between ME and PS, and the resummation scale of soft gluon emissions, respectively. The impact of these parameters on the BDT input observables has been estimated at the generator level using Rivet [259] and alternative samples in which their nominal values are varied up and down. Little shape differences have been visible when varying any of the two parameters. But for the QSF parameter, a normalisation difference due to the event selection of about 17 % is visible, which is added as a normalisation uncertainty in this analysis.

The cross-section of single-boson processes is determined at NNLO with an accuracy of 5 %.

#### Multijet process

The multijet background is estimated using a data-driven ABCD method as described in Section 4.3.1. A systematic uncertainty on the applied method is estimated by using an alternative region definition in which a loose lepton identification (instead of isolation) requirement is used to define the regions. In the alternative approach the MJ background shape is changed and the estimated MJ contribution is reduced by about 30 %. In addition, a 50 % normalisation uncertainty is applied. This value is based

on validation studies of the fake factor estimation (Section 4.3.1) and MJ estimation studies from other analyses that have a similar signal region definition.

All systematic uncertainties impacting simulation indirectly impact the MJ background, as its estimation method relies on simulated data. But the uncertainties are not propagated separately to the MJ background due to its small overall contribution and the already very large considered uncertainties.

## 6.2 Statistical interpretation

The goal of this analysis is to search for an excess in data that matches the expected signal pattern. In case no excess is found, upper exclusion limits on the branching ratio of a top quark decaying into a charged Higgs boson and $b$ quark are derived. This is done by performing binned maximum-likelihood fits to the BDT score distributions. A statistical fit model is built in *Histfactory* [260] format with the help of the *TRExFitter* framework. Fitting of the model and the derivation of upper limits are performed using the *pyhf* framework [261]. Uncertainties on the fitted parameters are calculated by utilising the auto-diff functionality of *jax*.

### 6.2.1 Statistical model

The compatibility of predictions and the measured data is evaluated with the help of a statistical model. The model can be expressed as

$$L(n, a|\eta, \chi) = \prod_{c \in \text{channels}} \prod_{b \in \text{bins}_c} \text{Poisson}(n_{cb}|\nu_{cb}(\eta, \chi)) \times \prod_{x \in \chi} q_x(a_x|\chi). \tag{6.1}$$

$L$ can be interpreted as a binned likelihood. For each bin ($b$) in each channel ($c$) – this analysis has only one channel, the signal region – the data count $n_{cb}$ is compared with the total prediction $\nu_{cb}$ using the Poisson distribution, and the probabilities are multiplied. The prediction $\nu_{cb}$ depends on constrained ($\chi$) and unconstrained ($\eta$) modifiers that are the adjustable parameters of the model. The constrained parameters are constrained by auxiliary data measurements ($a_x$), which are incorporated in the likelihood via the constraint terms $q_x$. In the fitting procedure the predictions, to be precise $\eta$ and $\chi$, are fit to data by minimising the likelihood $L$.

#### Unconstrained modifiers

This analysis has three unconstrained modifiers:

**$\mathscr{B}_{\mathrm{H}^{\pm}}$:** $\mathscr{B}(t \to H^{\pm} b)$
$\mathscr{B}_{H^{\pm}}$ is the parameter of interest in the fit. In the SM $\mathscr{B}_{H^{\pm}} = 0$, and it is bounded to $\mathscr{B}_{H^{\pm}} \in [0, 0.1]$. In general, the branching ratio can reach up to 1, but a smaller upper bound increases the fit stability. The upper bound of 0.1 is justified by measurements setting lower bounds on $\mathscr{B}(t \to Wb)$ [262].

**$\mu_{\mathrm{t\bar{t}}}$:** $t\bar{t}$ cross-section scaling factor
The nominal prediction from the simulation is 1 with bounds $\mu_{t\bar{t}} \in [0, \text{inf})$.

**$f_{\mathrm{HF}}$:** Fraction of $t\bar{t}$ + HF events among $t\bar{t}$ background events
The nominal prediction from simulation is 0.1364 with bounds $f_{\text{HF}} \in [0, 1]$.

Besides the parameter of interest, $\mu_{t\bar{t}}$ and $f_{\mathrm{HF}}$ are chosen as free parameters because the analysis is expected to be sensitive to both. In addition, in multiple ATLAS analyses it has been observed that the $f_{\mathrm{HF}}$ value measured in data deviates significantly from its predicted value in simulation [91, 97, 263–265].

The unconstrained parameters modify the predicted yields as follows:

$$\begin{aligned} N_{H^{\pm}}(\text{post-fit}) &= \mu_{t\bar{t}} \times 2(1-\mathscr{B}_{H^{\pm}})\ \mathscr{B}_{H^{\pm}} \times N_{H^{\pm}}(\text{pre-fit}), \\ N_{t\bar{t}+\mathrm{LF}}(\text{post-fit}) &= \mu_{t\bar{t}} \times (1-f_{\mathrm{HF}}) \times (1-\mathscr{B}_{H^{\pm}})^2 \times N_{t\bar{t}+\mathrm{LF}}(\text{pre-fit}), \\ N_{t\bar{t}+\mathrm{HF}}(\text{post-fit}) &= \mu_{t\bar{t}} \times f_{\mathrm{HF}} \times (1-\mathscr{B}_{H^{\pm}})^2 \times N_{t\bar{t}+\mathrm{HF}}(\text{pre-fit}). \end{aligned} \tag{6.2}$$

For the nominal values of $\mu_{t\bar{t}}$ and $f_{\mathrm{HF}}$, and $\mathscr{B}_{H^{\pm}} = 0$, the SM expectation from simulation is recovered. For the above equations to hold, the event yields before the fit are set to the total expected $t\bar{t}$ yield ($N_{H^+}$(pre-fit, unprescaled) is normalised to $\mathscr{B}_{H^{\pm}} = 1\,\%$):

$$\begin{aligned} N_{t\bar{t}+\mathrm{LF}}(\text{pre-fit}) &= \frac{N_{t\bar{t}+\mathrm{LF}}(\text{pre-fit, unprescaled})}{1-f_{\mathrm{HF}}(\text{pre-fit})}; \\ N_{t\bar{t}+\mathrm{HF}}(\text{pre-fit}) &= \frac{N_{t\bar{t}+\mathrm{HF}}(\text{pre-fit, unprescaled})}{f_{\mathrm{HF}}(\text{pre-fit})}; \\ N_{H^+}(\text{pre-fit}) &= \frac{N_{H^+}(\text{pre-fit, unprescaled})}{2 * 0.01 * (1. - 0.01)}. \end{aligned}$$

In Equation 6.2 and throughout the statistical interpretation of the results, it is assumed that the top quark decays solely into a $W$ or $H^{\pm}$ boson and a $b$ quark and that the charged Higgs boson decays solely into $cs$ quarks – i.e. $\mathscr{B}(t \to Wb) + \mathscr{B}(t \to H^{\pm}(\to cs)b) = 1.0$. The branching ratios of other SM top-quark decays are predicted to be tiny and can be neglected with a high precision. Other relevant light charged Higgs boson decays are $H^{\pm} \to \tau\nu_{\tau}$ and $H^{\pm} \to cb$. Events with the former decay mode are suppressed by the event selection and reconstruction. If $\mathscr{B}(H^{\pm} \to cb)$ is sizeable, its contribution is not correctly accounted for in the analysis. Upper limits on $\mathscr{B}(t \to H^{\pm}b)$ assuming $\mathscr{B}(H^{\pm} \to cs) = 1.0$ would presumably be too tight, because $H^{\pm} \to cb$ events receive on average smaller BDT scores than $H^{\pm} \to cs$ signal events. In order to correctly account for a $H^{\pm} \to cb$ contribution, dedicated $H^{\pm} \to cb$ signal samples for each mass point hypothesis are needed, which are currently not available.

### Constrained modifiers

The systematic uncertainties listed in Table 6.1 are implemented as Gaussian-constrained nuisance parameters[1] (NPs) in the statistical model ($q_x(a_x|\chi)$ in Equation 6.1). Their expectation value is zero and $\pm 1$ corresponds to a change by their $\pm 1\sigma$ uncertainty. Normalisation-only components (labelled "N" in Table 6.1) impact the total expected event rate of samples. Normalisation-plus-shape components (labelled "S" in Table 6.1) are included in the form of alternative templates and impact the normalisations and shapes of samples. Those are additive modifiers correlated across all bins. The components are considered to be uncorrelated with each other, but correlations of individual components across samples are maintained.

[1] Commonly all parameters of a fit, constrained and unconstrained, except the parameter of interest, are referred to as nuisance parameters.

Some shape templates undergo modifications to improve the fit stability. For the templates with only one (up) variation – e.g. when comparing to an alternative sample – the respective down variation is extracted by mirroring about the nominal template (one-sided symmetrisation). The systematic uncertainty on FSR in signal and top-quark processes is symmetrised by taking the mean of up and down variations ((up − down)/2 – two-sided symmetrisation). This uncertainty showed large fluctuations between neighbouring bins for up and down variations caused by large event weights (the systematically varied FSR template is retrieved by recalculating the event weights, given the varied $\mu_R^{\mathrm{FSR}}$ parameter). The impact of the large event weights is aimed to be reduced by the symmetrisation, as it averages the variation over more events.

Additionally, shape components that are estimated from two different configurations (two-point systematics – those are all non-reweighting systematic uncertainties; an example is alternative samples, which are produced with a different generator setup) and the FSR systematic uncertainties are smoothed. The difference of two-point systematic uncertainties with respect to nominal usually shows larger fluctuations in tiny phase space regions, as is the high BDT score region. The smoothing is performed using the *TTBARRESONANCE*[2] algorithm implemented in TRExFitter. It performs a rebinning of the template to reduce statistical fluctuations of the systematic variations by merging neighbouring bins with similar differences with respect to nominal. The *353QH twice* smoothing algorithm [267] is applied to the difference between nominal and rebinned systematic templates. The algorithm performs the smoothing by calculating medians from multiple neighbouring bins. Eventually the smoothed differences are transferred to the original systematic template. A few plots showing the original and modified systematic templates and a comparison with the nominal template are shown in the Appendix in Figure A.20. Which systematic uncertainties are smoothed and symmetrised are marked in Table 6.1.

One systematic uncertainty on the prediction not discussed so far is the uncertainty due to the finite size of simulated datasets. This is accounted for by one Gaussian-constrained NP per bin, whose expectation value is one and the variance is defined as the square root of the yield uncertainties of all samples in a bin summed up relative to the total yield in that bin: $\sigma_{cb} = \sqrt{\sum_s \Delta_{scb}} / \sum_s \nu_{scb}$; where $s$ runs over the samples and $\Delta$ is the MC statistical uncertainty [268].

Eventually small NPs that have a negligible impact on the final fit results are pruned to increase the stability of fits. Normalisation components that impact a given sample by less than 0.5 % are pruned for that sample. Systematic components where the template of a given sample has no bin that shows a relative deviation by 0.1 % with respect to the nominal template are pruned for that sample. The NPs related to the MC dataset sizes are not pruned. The quoted cut values (0.5 % and 0.1 %) were determined by measuring their impact on the parameter of interest uncertainty ($\Delta\mathscr{B}_{H^{\pm}}$) and the expected upper limits on $\mathscr{B}_{H^{\pm}}$. These are the largest cut values – remove as many components as possible – that change $\Delta\mathscr{B}_{H^{\pm}}$ and the expected upper limits by much less than 1 %.

Constrained NPs can be *pulled* or *constrained* by the fit. A NP is pulled if its post-fit value deviates from its initial value before the fit. A NP is constrained if its post-fit uncertainty is smaller than its

[2]The name "TTBARRESONANCE" comes from a search for heavy particles decaying into top-quark pairs [266]. The smoothing algorithm got developed by the referred analysis in order to not smooth away single bins with a statistically significant difference with respect to nominal – i.e. the smoothing procedure is more conservative than many other smoothing algorithms.

pre-fit uncertainty. In this case the fit is able to constrain it better than in the previous auxiliary data measurement. To be precise this is an overconstraint, but for the sake of simplicity it will be referred to as a constraint. Underconstraints – post-fit uncertainty is larger than its pre-fit uncertainty – are generally not expected and are also not observed in this analysis.

### 6.2.2 Hypothesis testing

Discovery and exclusion of a potential signal in data are established by performing a likelihood ratio test of two hypotheses. A null hypothesis ($H_0$) has to be rejected in favour of an alternative hypothesis ($H_1$), typically corresponding to the unconstrained model, at a defined significance. The agreement of the hypotheses with the observed data is quantified by the statistical model/likelihood ($L$) introduced in the previous section. The two hypotheses are compared by taking the ratio of their likelihoods. The test statistic $t_\mu$ of the likelihood ratio test is defined as [269]

$$t_\mu = -2\ln(\lambda(\mu)) \quad \text{with} \quad \lambda(\mu) = \frac{L(H_0)}{L(H_1)} = \frac{L(\mu, \hat{\hat{\theta}}(\mu))}{L(\hat{\mu}, \hat{\theta})}.$$

$\mu$ is the parameter of interest and $\theta$ are the nuisance parameters. A hat above a parameter ($\hat{\theta}$) indicates that this is its maximum likelihood estimate. $\hat{\hat{\theta}}(\mu)$ indicates that it is the maximum likelihood estimate of the nuisance parameters for a given $\mu$. A greater value of $t_\mu$ corresponds to a larger incompatibility of $H_0$ with $H_1$.

The outcome of the likelihood ratio test is summarised by the $p$-value,

$$p_\mu = \int_{t_{\mu,\text{obs}}}^{\infty} f(t_\mu|\mu')\,\mathrm{d}t_\mu.$$

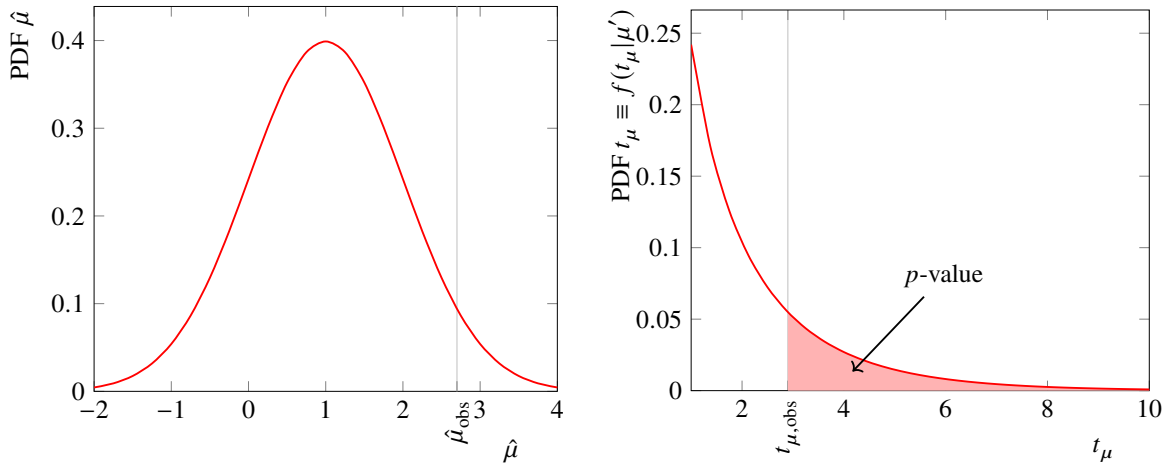
$t_{\mu,\text{obs}}$ is the value of the test statistic observed in data. $f(t_\mu|\mu')$ is the probability density function of $t_\mu$ for an assumed value of $\mu'$ in data. In general $\mu' = \mu$, but in order to estimate expected uncertainties on $\mu$, one needs to know $f(t_\mu|\mu')$ with $\mu' \neq \mu$. The $p$-value corresponds to the probability of having observed data of equal or greater incompatibility than predicted by $H_0$.

The difficulty is that $f(t_\mu|\mu')$ is not known a priori, as it depends on the specific likelihoods of the given problem. In principle, it is all about knowing the distribution of $\hat{\mu}$ for an assumed $\mu'$ in data. $f(t_\mu|\mu')$ is the same but transformed to the test statistic $t_\mu$. This is illustrated in Figure 6.1.

Two common methods exist for estimating $f(t_\mu|\mu')$. One method is to run many pseudo-experiments (toys) e.g. by running MC simulations. In pyhf this is implemented by random sampling from the simulated event statistics. The downside of toys is that producing them is, in general, computationally expensive.

In this analysis an approximation of the sampling distribution (asymptotic approximation) is used because of its fast computation and the overall large data and MC sample sizes in the analysis. In the large statistics limit $f(t_\mu|\mu')$ can be estimated using Wald's approximation [270],

$$t_\mu = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + O(1/\sqrt{N}).$$




(a) (b)

Figure 6.1: (a) Exemplary probability density distribution of $\hat{\mu}$. In this plot $\mu' = 1$ and a standard normal distribution is assumed. (b) The respective probability density distribution in $t_\mu$-space for $\mu' = \mu$. Also indicated is the definition of the $p$-value for an observed value $t_{\mu,\text{obs}}$ of the test statistic in data.

When neglecting the $O$-term, $\hat{\mu}$ follows a Gaussian distribution with a mean $\mu'$ and standard deviation $\sigma$, and $t_\mu$ follows a noncentral chi-square distribution,

$$f(t_\mu;\Lambda) = \frac{1}{2\sqrt{t_\mu}}\frac{1}{\sqrt{2\pi}}\left[\exp\left(-\frac{1}{2}(\sqrt{t_\mu}+\sqrt{\Lambda})^2\right)+\exp\left(-\frac{1}{2}(\sqrt{t_\mu}-\sqrt{\Lambda})^2\right)\right] \tag{6.3}$$

with noncentral parameter $\Lambda = (\mu-\mu')^2/\sigma^2$ [269]. The standard deviation is commonly extracted from the Asimov dataset[3] as

$$\sigma = \frac{(\mu-\mu')^2}{t_{\mu,\text{A}}} \tag{6.4}$$

where $t_{\mu,\text{A}}$ is the test statistic of the Asimov dataset.

The parameter of interest in this analysis is $\mu = \mathscr{B}_{H^\pm}$, which is defined to be $\geq 0$. This modifies the definition of the test statistic to

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

and modifies the distribution of $f(t_\mu|\mu')$ for $\hat{\mu} < 0$ compared to Equation 6.3.

[3] In the Asimov dataset all parameters of the model agree with their expected value (a hypothetical scenario where the data shows no statistical fluctuations). In this analysis this means the Asimov dataset is identical to the initial prediction from the simulation.

### 6.2.3 Discovery test

The discovery of a signal in data is established by rejecting the background-only hypothesis ($H_0$ with $\mu = 0$) in favour of the unconstrained signal-plus-background hypothesis ($H_1$ with $\mu = \hat{\mu}$). The test statistic in this case is given by [269]

$$q_0 \coloneqq t_0 = -2 \ln \begin{cases} \lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}$$

$q_0$ considers only positive values of $\hat{\mu}$ as a discrepancy between the background-only hypothesis and data. A negative $\hat{\mu}$ does not hint towards signal events in data but rather some systematic uncertainties. Therefore, $q_0$ is set to zero in this case.

In the asymptotic approximation the $p$-value can be calculated as [269]

$$p_0 = \int_{q_{0,\mathrm{obs}}}^{\infty} f(q_0|0)\, \mathrm{d}q_0 = 1 - \Phi(\sqrt{q_{0,\mathrm{obs}}}), \tag{6.5}$$

where $\Phi$ is the cumulative distribution function of the standard normal distribution. The significance $Z = \Phi^{-1}(1 - p)$ with $p = p_0$ is commonly referred to as the *discovery significance*. It is convention to claim discovery if $Z_0 \geq 5$. In this case $p_0 \leq 2.87 \times 10^{-7}$, which corresponds to a one-sided integral of a standard normal Gaussian above $5\sigma$. If in nature there were no signal, in only 1 out of $1/p_0$ times one would observe the given data or something more signal-like, i.e. falsely claim discovery (type-I error).

The expected discovery significance depends on the model being considered. In the SM, $Z_0 = 0$ is expected. In the case of BSM models, it also depends on the values of free parameters, in 2HDM models specifically on $\tan\beta$.

### 6.2.4 Upper limit calculation

The upper limit defines the $\mu_{\mathrm{up}}$ value for which the signal-plus-background hypothesis ($H_0$ with $\mu = \mu_{\mathrm{up}}$) gets rejected in favour of the unconstrained signal-plus-background hypothesis ($H_1$ with $\mu = \hat{\mu}$) at a given confidence level (CL). $\mu_{\mathrm{up}}$ can be thought of as the largest $\mu$ value (amount of signal) that is within the CL still compatible with the observed data. Conventionally a CL of 95 % is chosen, which means $H_0$ is rejected in 95 % of the cases ($p \leq 1 - \mathrm{CL} = 0.05$, $Z \geq 1.64$). If in nature there were a signal with $\mu = \mu_{\mathrm{up}}$, in $1/p$ times one would observe the given data or something more background like, i.e. falsely reject the signal hypothesis (type-II error).

The test statistic is given by [269]

$$q_\mu = -2 \ln \begin{cases} \lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{6.6}$$

A smaller value of $\mu$ than what is found in data ($\hat{\mu}$) does not speak against a larger value of $\mu_{\mathrm{up}}$. Consequently, $q_\mu$ is set to zero in these cases.

In the asymptotic approximation the $p$-value can be calculated as [269]

$$p_{s+b} = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q_\mu|\mu)\, \mathrm{d}q_\mu. \tag{6.7}$$

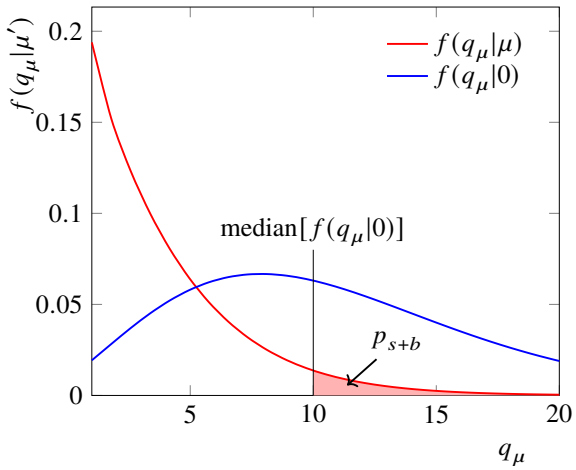



Figure 6.2: Methodology to extract the expected upper limits. The distributions are estimated from the Asimov dataset and not the observed data. The shapes of the distributions depend on $\mu$. $\mu_{\text{up,exp}}$ is defined to be the value of $\mu$ for which $p_{s+b} = \alpha$.

$\mu_{\text{up}}$ is the value of $\mu$ for which $p_{s+b} = 1 - \text{CL} \equiv \alpha$. In the asymptotic approximation this leads to [269]

$$\mu_{\text{up}} = \hat{\mu} + \sigma\Phi^{-1}(1-\alpha). \tag{6.8}$$

Extracting the value of $\mu_{\text{up}}$ that solves this equation is not easy because $\sigma$ depends on $\mu$ itself (cf. Equation 6.4). In practice $\mu_{\text{up}}$ is determined numerically.

Expectations on the upper limit ($\mu_{\text{up,exp}}$) and their uncertainties are derived using the Asimov dataset. This leads to $H_1$ being the background-only hypothesis with $\hat{\mu} = 0$ and $\mu' = 0$.

$\mu_{\text{up,exp}}$ is determined by evaluating the $p$-value at the median significance of the Asimov dataset. The $p$-value is calculated as in Equation 6.7, but the lower bound of the integral is set to median$[f(t_\mu|0)]$. $\mu_{\text{up,exp}}$ is the value of $\mu$ for which $p_{s+b} = \alpha$. This methodology is visualised in Figure 6.2. In the asymptotic approximation one finds [269]

$$\mu_{\text{up,exp}} = \mu' + \sigma(\Phi^{-1}(1-\alpha) \pm M), \tag{6.9}$$

with $\mu' = 0$ in this analysis/Asimov dataset. $M = 0$ corresponds to the expected upper limit and $M = 1, 2$ correspond to the one and two sigma error bands. This result is similar to Equation 6.8 but with $\mu'$ replacing $\hat{\mu}$. The variations $\mu_{\text{up,exp}}$ are simply $\pm\sigma M$ because $\hat{\mu}$ is Gaussian distributed with standard deviation $\sigma$.

A weakness of the upper limit calculation based on $p_{s+b}$ is that, if the signal sensitivity of the experiment is rather weak, fluctuations in the observed data can strongly impact the excluded $\mu$-range. An example of this is a one-bin experiment in which the signal is much smaller than the background contribution ($N_s \ll N_b$) and the observed data shows a significant downward fluctuation that is consistent with the background-only yield. In this case, already relatively small signal contributions will be excluded. In order to mitigate such *spurious exclusions*, it is taken into account how likely it is to observe

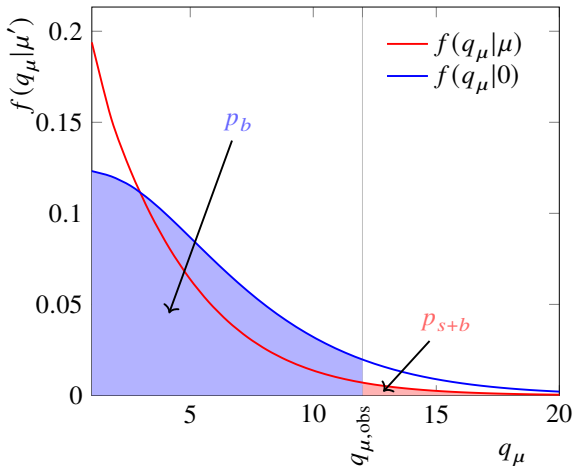



Figure 6.3: Methodology to extract $\mathrm{CL}_s$. $p_{s+b}$ and $p_b$ are represented by the red and blue areas, respectively. In this example a smaller signal sensitivity is assumed compared to the example in Figure 6.2, which makes the two distributions more similar. The large observed value of $q_\mu$ would result in tight $\mathrm{CL}_{s+b}$ exclusion limits that are weakened in the $\mathrm{CL}_s$ method due to the small (large) value of $\mathrm{CL}_b$ ($p_b$).

the given data, assuming there is only background. In this analysis this is done by using the modified frequentist technique [271],

$$\mathrm{CL}_s = \frac{\mathrm{CL}_{s+b}}{\mathrm{CL}_b} = \frac{p_{s+b}}{1 - p_b} \quad \text{with} \quad p_b = \int_{-\infty}^{q_{\mu,\text{obs}}} f(q_\mu|0)\,\mathrm{d}q_\mu .$$

The definition of $p_b$ is visualised in Figure 6.3. $\mu_{\text{up}}$ is defined as the $\mu$ for which $\mathrm{CL}_s = 0.05$. $\mathrm{CL}_s$ does not correspond to a classical frequentist confidence interval and is more conservative than the standard $\mathrm{CL}_{s+b}$ method as $0 \leq p_b \leq 1$. In experiments with good signal and background separation $\mathrm{CL}_s$ approaches $\mathrm{CL}_{s+b}$.

As outlined previously, $\mu_{\text{up}}$ has to be determined numerically. In this analysis this is implemented by scanning over 100 equidistant values of $\mu$ and calculating the respective $\mathrm{CL}_s$ values. The $\mu$ scan bounds are from zero up to a meaningful upper value, which are extracted from previous scans and differ for each mass point. The upper limit at $\mathrm{CL}_s = 0.05$ is finally extracted by performing a linear extrapolation between the scanned $\mathrm{CL}_s$ values.

### 6.2.5 BDT score binning

In the previous section statistical models/hypotheses are compared to quantify the presence of a signal in data. The models are fit to the data in a binned maximum-likelihood approach. The discriminative variable is the BDT score. The outcome of the statistical test is reliant on the binning of the BDT score distribution. Multiple objectives have to be considered when determining the binning.

The main objective is a maximised signal sensitivity. When considering only the signal sensitivity, having many bins is beneficial. Especially in the most discriminating region, which is the high BDT score region. But it has to be taken care that the fit setup is stable, converges and yields reliable and trustworthy results, which can be problematic when the number of bins is too large. Most important in that regard is that the signal and background predictions are well modelled in each bin. In order to use

the asymptotic approximation, it also has to be ensured that a sufficient number of events is in each bin. It is convention to require at least $\sim 10$ data events.

The final BDT score binnings are determined by first producing templates with a very fine binning; 10 000 equal-width bins from 0 to 1. The bins of the templates are then merged based on conditions each bin has to fulfil. The bins are merged iteratively from right to left, i.e. first merging the most discriminating bins. Once a (merged) bin fulfils all conditions listed below, the bin is final and a new merging iteration begins starting from the next to last final bin until all bins are merged. Since the number of bins is limited, it may happen that in the last iteration not all conditions can be fulfilled. In this case the bins from that iteration are merged with the bins from the second-to-last iteration into one final bin.

The rebinning conditions are:

1. At least 20 expected background events;
2. Statistical uncertainty of simulated backgrounds is less than 10 %;
3. Statistical uncertainty of simulated signal is less than 20 %;
4. $S/\sqrt{B} > \gamma$ ($\gamma := \mathit{Separation} \times \mathit{Threshold}$);
5. At maximum 1 000 neighbouring bins may be merged into one final bin.

Conditions 1 to 3 ensure a proper modelling of the prediction. Condition 4 is related to the signal sensitivity. The figure of merit measuring the sensitivity is the number of expected signal events in the single/merged bin under consideration ($S$) divided by the square root of the total background prediction in that bin ($B$). By choosing $\gamma = 1$, one can ensure that each bin has a significant signal contribution, as in this case the signal contribution is larger than the statistical uncertainty on the background. This search is performed model-independently and there is no prescribed value for the signal strength. The signal process in simulation is arbitrarily scaled to $\mathcal{B}_{H^{\pm}} = 1\,\%$, and there is no clear choice for $\gamma$.

It was decided to define $\gamma$ based on the bin with the highest signal sensitivity. This is the first final bin that fulfils conditions 1 to 3. This bin usually dominates the overall signal sensitivity of the analysis. The *Separation* parameter gets defined as $\mathit{Separation} \equiv S_0/\sqrt{B_0}$, where $S_0$ and $B_0$ are the numbers of signal and background events in the first final bin. Further on in the merging procedure, only bins that have a sensitivity comparable to the first bin will be declared as final.

This merging strategy results in relatively few bins for highly discriminating BDT scores and many bins for low discriminating BDT scores. This is the intuitively desired result because when the separation is large in the high BDT score region, the low BDT score bins won't contribute a lot to the overall separation and can be merged. Nonetheless, this effect is mitigated by the *Threshold* parameter. The *Threshold* values are chosen such that each rebinned BDT score has at least 19 bins and at maximum 49 bins. In this way it is assured that some shape information is always kept in the low BDT score region, which is important to constrain NPs. For the same reason condition 5 is applied. At the same time, there are not too many bins such that the rebinned BDT score is less likely to be sensitive to unphysical effects, like inaccurate simulations in tiny phase space regions. Most BDT scores have the *Threshold* parameter set equal to 1. The BDT scores trained with a signal mass point close to the $W$-boson mass have a value larger than 1, and mass points close to the $t$-quark mass have a value smaller than 1.

The developed rebinning strategy is rather conservative. This means it has been put more focus on solid and reliable results in the statistical interpretation, rather than improvements in the expected upper limits. Binning strategies focusing more on the latter have also been tested. In these tests, the parameters for the minimum number of events and statistical uncertainties on predictions have been lowered. In addition, other requirements in place of condition 4 have been tested. Improvements in the expected upper limits were only visible for final binnings in which large NPs could be constrained in the fit. Those were NPs related to the $t\bar{t}$ modelling of ME and PS. Strong constraints on them mean that the final results are only applicable for generators that are very similar to the nominal generator setup. If nature is not that close to the nominal generator setup – or in other words, the nominal generator setup cannot model nature that well – the results will be overconfident. Therefore, strong constraints on the discussed NPs make the results generally less trustworthy and are disfavoured.

The rebinning algorithms implemented in TRExFitter have also been tested. The results on the expected upper limits were always worse compared to the method developed in this analysis.

Figure 6.4 shows the same BDT score distributions as shown in Figure 5.11 but after the rebinning procedure. In Figure 5.11 an equidistant binning was chosen with 100 bins. The rebinned BDT scores for the other mass point trainings are shown in Appendix A.3. For a better visualisation the BDT score bin numbers in Figure 6.4 are plotted with equal size of each bin. When plotting the true BDT score, the highest BDT score bins are tiny compared to the other bins and it is not possible to read off important information.

Interesting to see is that for most BDT scores, the highest bin has exactly 20 expected events. This means condition 1 is the restricting requirement defining the first merged bin. The lower edges of the following merged bins are driven by their signal sensitivity (Condition 4). Only very few signal events are available in the low BDT score region and condition 3 or even condition 5 define the lower edges of the merged bins.

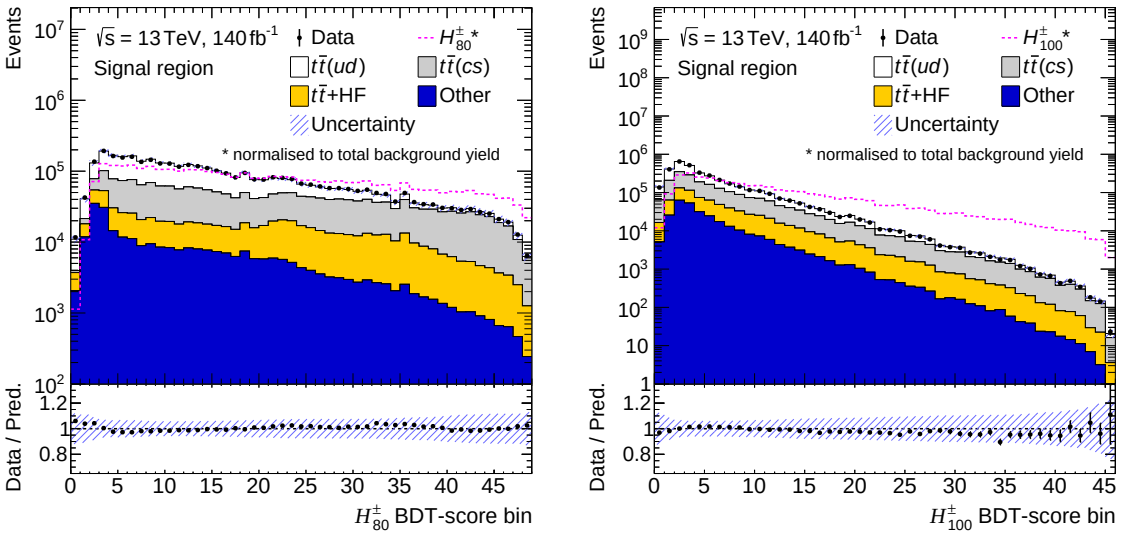
### 6.2.6 Validation of the statistical model and fit setup

The analysis has been developed in a blinded way. This means the data in signal-enriched regions was not looked at – was "blinded" – before finalising the analysis strategy. This includes all steps such as event reconstruction, BDT input feature selection, BDT score rebinning, etc. This is done to prevent any bias in the final results. As an example, the binning could be adapted to smear out statistical fluctuations, or one could specifically decide on a binning in which a bin has a significant upward fluctuation in data.

This is realised by blinding any bin in histograms that might have a significant signal contribution. If a bin has a signal-to-background ratio of $S/B > 5\,\%$, it is blinded. The blinding requirement is tested for all signal samples. The signal templates are scaled to the current best upper limits on $\mathscr{B}_{H^{\pm}}$ in $H^{\pm} \to cs$ decays. These are mainly set by CMS [90] and the mass points that are not covered by the CMS search are scaled to $\mathscr{B}_{H^{\pm}} = 0.5\,\%$, as motivated by a mass-independent search [272].

About 20 % and 50 % of the rebinned BDT score bins were blinded for the $H^{\pm}_{100}$ and $H^{\pm}_{130}$ BDT scores, respectively.

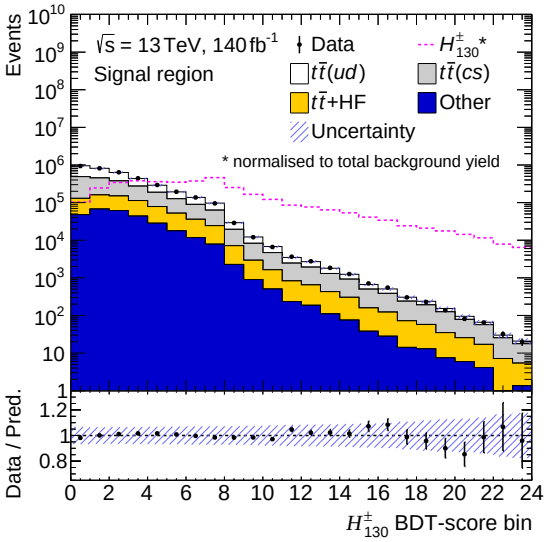
The fit setup has been validated using the Asimov dataset and the unblinded data bins. In fits to the Asimov dataset, correlations of fit parameters have been monitored. (Anti-)correlations close to 100 % can lead to unstable fits. A plot showing the linear correlation coefficients for the $H^{\pm}_{130}$ BDT score is presented in Figure 6.5. Only NPs that are correlated with another NP by at least 20 % are listed.



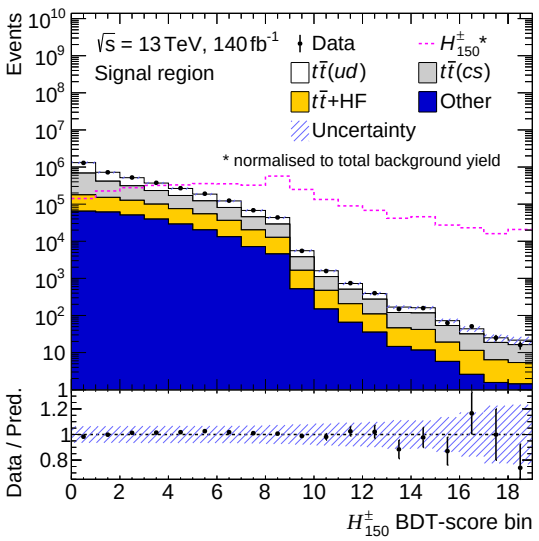

(a)




(b)




(c)




(d)

Figure 6.4: BDT score bins after rebinning for the trainings with the (a) $H^{\pm}_{80}$, (b) $H^{\pm}_{100}$, (c) $H^{\pm}_{130}$ and (d) $H^{\pm}_{150}$ signal samples.

The largest correlation in this plot is 50 %. The correlations are similar for other mass points. The $t\bar{t}$-modelling systematic uncertainties show overall the largest correlation with the parameter of interest, $\mathcal{B}_{H^{\pm}}$.

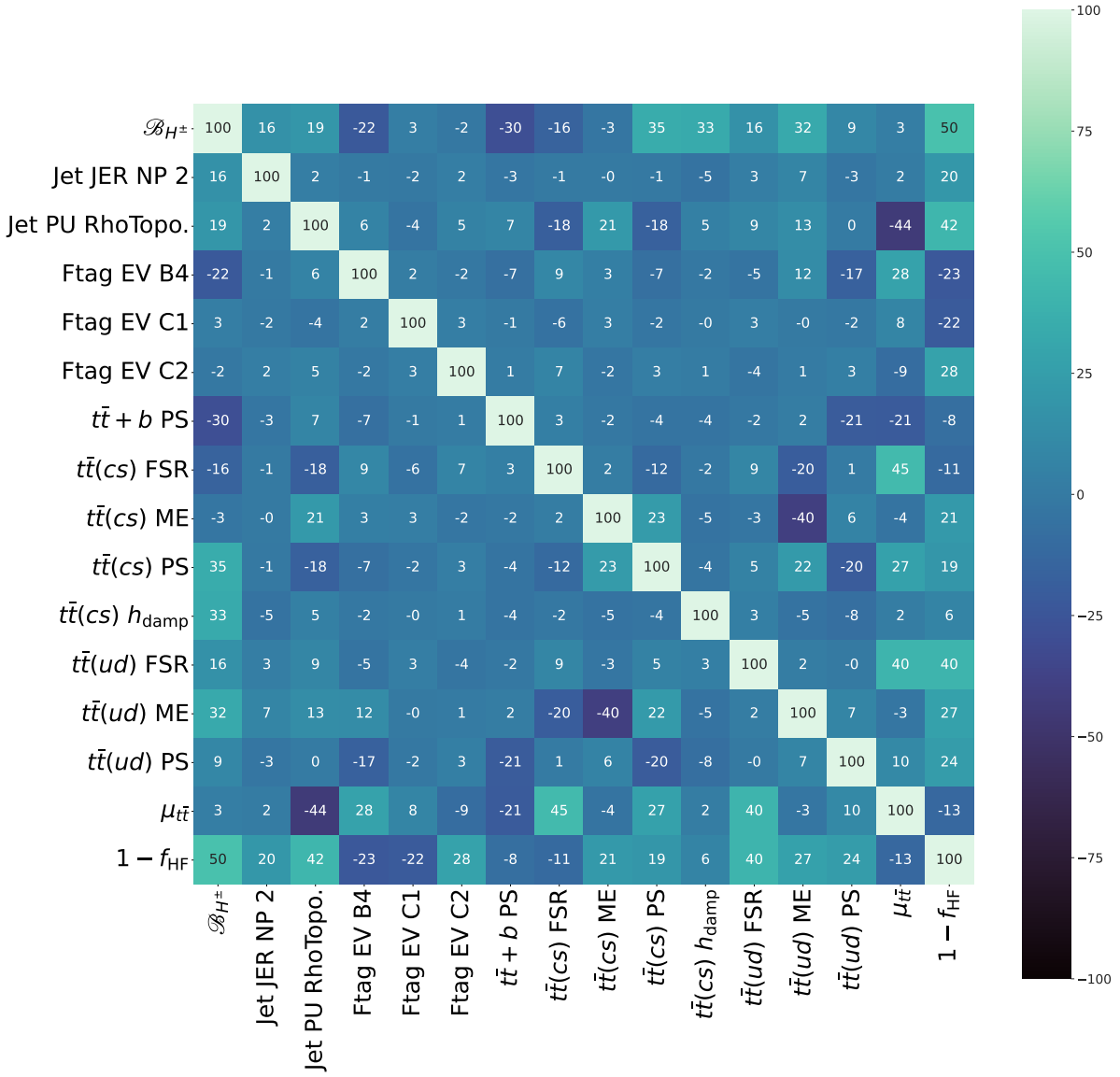



Figure 6.5: Linear correlation coefficients of fit parameters in fits to the Asimov dataset for the $H^{\pm}_{130}$ BDT score. For conciseness and better interpretability, only parameters that are correlated with another NP by at least 20 % are listed. "Ftag EV" are NPs related to the eigenvariations from the flavour-tagging calibration uncertainty matrix. The names of the other NPs are straightforward to identify based on Section 6.1.

In the Asimov fits, it was also assured that NPs are not unexpectedly strongly constrained. Most NPs showed no constraints at all. The largest constraints were seen for a few large $t\bar{t}$-modelling NPs, with constraints up to only 80 % of their initial uncertainty.

Pulls and constraints of NPs have been monitored in the fits to unblinded data bins, too. The constraints

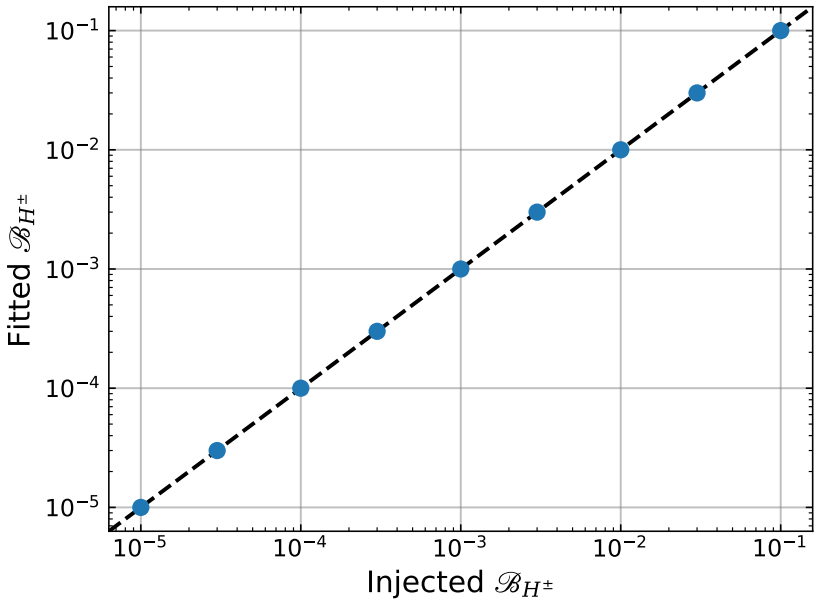



Figure 6.6: Signal injection test for the 130 GeV mass point. The x-axis shows the injected $\mathcal{B}_{H^{\pm}}$-value in the Asimov dataset and the y-axis shows the values from the fit. The dashed line marks the locations for which both values agree with each other.

were weaker than for the Asimov fits due to the reduced fit range. The size of the pulls depended on the number of unblinded bins. For the highly separating BDT scores, the largest pulls were about 20 %. For the lower separating BDT scores, the largest observed pull was 100 %. Given the large number of NPs, having no pull in the central value by more than $1\sigma$ is an acceptable result.

Signal injection tests have been performed to ensure that the fit can reliably identify a possible signal contribution in data. Special Asimov datasets were produced with a fixed signal contribution. Values of $\mathcal{B}_{H^{\pm}}$ up to 10 % – much larger than the current best upper limits on $\mathcal{B}_{H^{\pm}}$ – were tested. The results of the signal injection tests are shown in Figure 6.6 for the 130 GeV mass point. The fit was able to identify the injected value of $\mathcal{B}_{H^{\pm}}$ for each mass point with high precision for $\mathcal{B}_{H^{\pm}}$ values at least one order of magnitude smaller than the current best upper limit for the given mass point.

In addition to the presented validation studies, it was checked that the expected upper limits can be calculated successfully and that the results are in a meaningful range. An example of the $\mathrm{CL}_s$ scan is given in Figure 6.7. The upper limit and uncertainties are given by the $\mu$-values for which the respective dashed line crosses the red line. The dip in the $1\sigma$ and $2\sigma$ error bands at small $\mu$ values is due to the clipping of the test statistic – cf. Equation 6.6; $q_\mu = 0$ for $\hat{\mu} > \mu$.

Given the overall solid validation results of the fit setup, it was decided to unblind the analysis and perform fits to the full dataset. The unblinded results are presented in the following chapter.

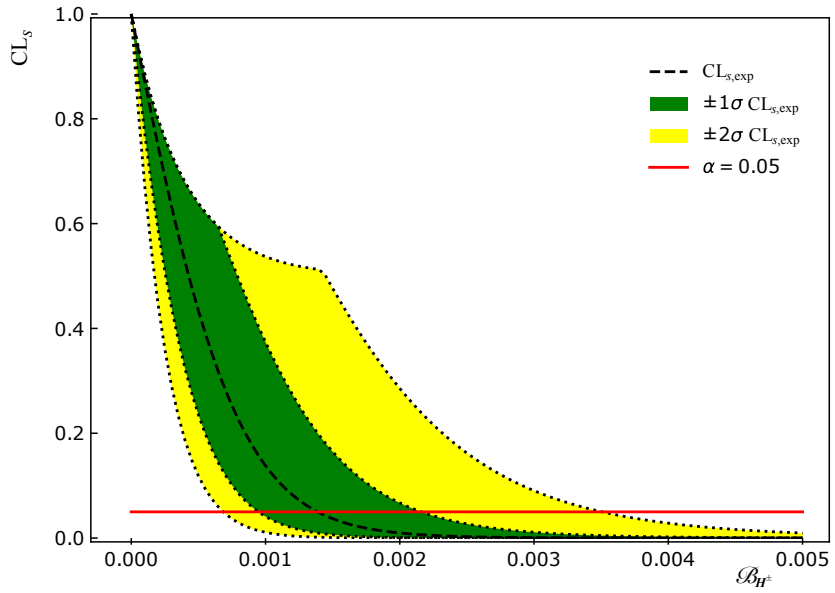



Figure 6.7: $\mathscr{B}_{H^{\pm}}$ scan of the CL$_s$ for the Asimov dataset to determine the expected upper limit and the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties. The red line indicates the 95 % confidence level.

CHAPTER 7

# Results of the search for $H^{\pm} \to cs$

This chapter presents the results of the search for a potential $H^{\pm} \to cs$ signal in data. The final discriminants of the search are the BDT score distributions. The fits of the statistical model to the BDT score distributions are presented in Section 7.1. The impact of systematic uncertainties on the fit results is discussed in Section 7.2. The statistical interpretation of the fit results is presented in Section 7.3 and compared to previous searches in Section 7.4. Lastly, the sensitivity of this analysis to a potential $H^{\pm} \to cb$ signal is evaluated and discussed in Section 7.5, and an interpretation of the results in 2HDMs with NFC is presented in Section 7.6.

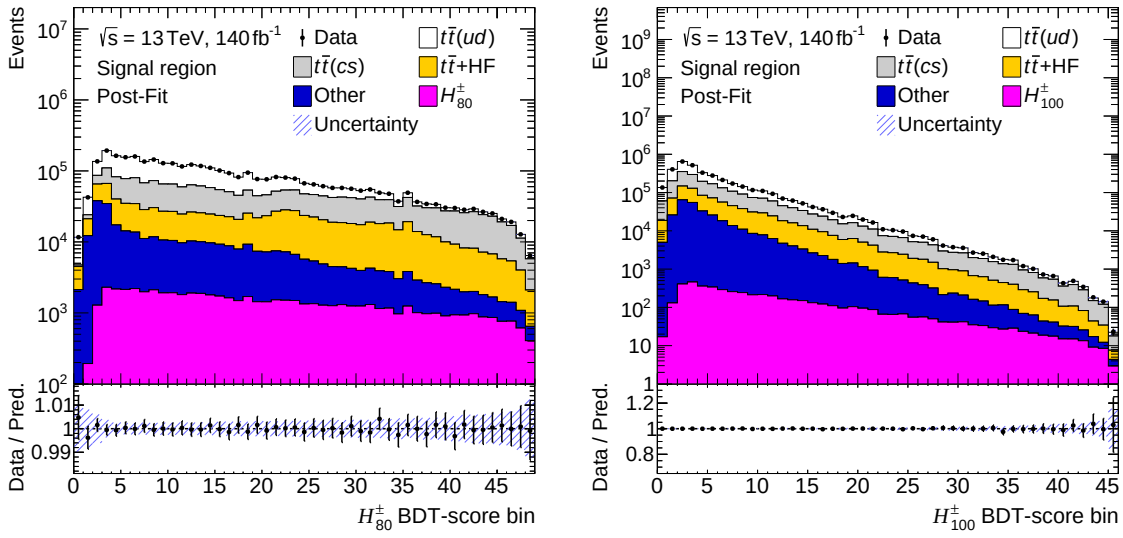
## 7.1 BDT score fit results

#### Post-fit BDT score distributions and yields

Figure 7.1 shows the same BDT scores as in Figure 6.4 but after performing the fit of the statistical model to the data (post-fit), as described in Chapter 6. A good agreement between data and prediction is visible in the ratio panel for all BDT score fits. Striking are the largely reduced uncertainties on the prediction compared to Figure 6.4. This is due to the fit taking into account correlations of NPs and caused by constrains on NPs.

In some BDT scores a non-zero signal contribution has been fitted. It is visible at the bottom of the stack plots. A post-fit yield of 300 signal events is retrieved for the highest bin of the $H^{\pm}_{80}$ BDT score distribution. This amounts to 4.6 % of the total data size in that bin and is larger than the associated total uncertainty in that bin, which is 1.2 %. In the fits to the $H^{\pm}_{100}$ and $H^{\pm}_{130}$ BDT score distributions, a post-fit signal yield of only 2 and 1.5 events, respectively, is retrieved. But for both bins, this amounts to 10 % of the total data size in the respective bins and roughly equals the size of the total bin uncertainty. The post-fit signal yield in the fit to the $H^{\pm}_{150}$ BDT score distribution is zero.

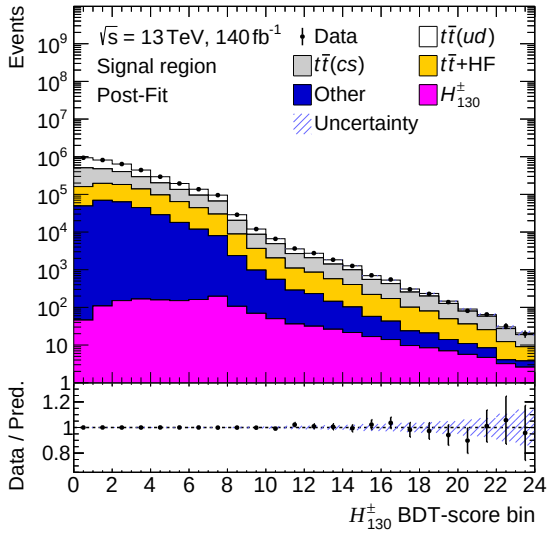
To evaluate the goodness of the fits, it has been tested that the retrieved values for the NPs, determined from the fits to the BDT score distribution, also lead to an improved data-prediction agreement in other quantities. Distributions of a few selected BDT input features, including $m_{j_1 j_2}$, with NPs set to the BDT score fit values, are shown in Figure A.19 in the Appendix. In these plots it can be seen that the overall agreement of predictions with data after applying post-fit NPs is very similar to pre-fit. This is due to



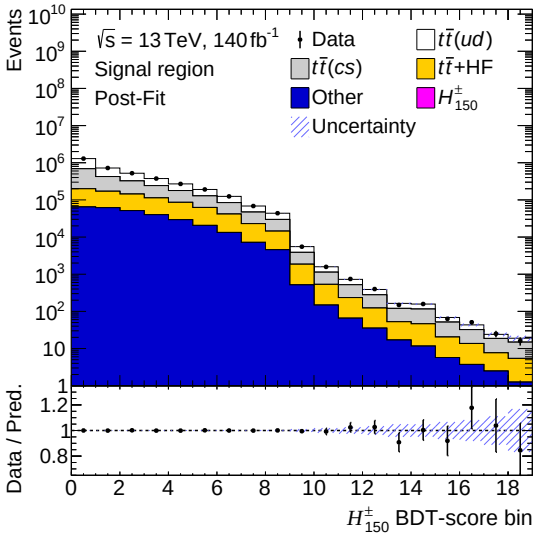

(a)




(b)




(c)




(d)

Figure 7.1: BDT score distributions after fitting to data (post-fit) for the trainings with the (a) $H^{\pm}_{80}$, (b) $H^{\pm}_{100}$, (c) $H^{\pm}_{130}$ and (d) $H^{\pm}_{150}$ signal samples.

Table 7.1: Pre-fit and post-fit yields of the single processes in the fits to the $H^{\pm}_{130}$ BDT score.

| Process | Pre-fit yields | Post-fit yields $H^{\pm}_{130}$ |
|---|---|---|
| $t\bar{t}(ud)$ | 1 500 000 ± 130 000 | 1 400 000 ± 91 000 |
| $t\bar{t}(cs)$ | 1 300 000 ± 110 000 | 1 300 000 ± 110 000 |
| $t\bar{t}$ + HF | 470 000 ± 70 000 | 630 000 ± 190 000 |
| $tW$ | 100 000 ± 23 000 | 110 000 ± 23 000 |
| Single top | 62 000 ± 29 000 | 68 000 ± 29 000 |
| $W$+ jets | 75 000 ± 32 000 | 72 000 ± 29 000 |
| $Z$ + jets & $VV$ | 21 000 ± 9 800 | 21 000 ± 9 700 |
| Other top & $t\bar{t}$(allHad) & $t\bar{t}H$ | 17 000 ± 520 | 17 000 ± 470 |
| Multijet | 12 000 ± 7 500 | 12 000 ± 6 900 |
| $H^{\pm}_{130}$ | - | 1 600 ± 2 500 |
| Total prediction | 3 617 809 ± 242 000 | 3 618 339 ± 11 000 |
| Data | 3 618 340 | |

generally small pulls on NPs, discussed in more detail below.

The pre- and post-fit yields of all processes for the fit with the $H^{\pm}_{130}$ signal sample are listed in Table 7.1. In the fit, a small signal contribution consistent with zero is extracted. The post-fit yields generally agree with the pre-fit yields within the pre-fit uncertainties. The largest shift is seen for the $t\bar{t}$ + HF process and will be discussed in more detail later in this section.

The total pre-fit prediction already agrees well with the total data yield. This is due to the $t\bar{t}$ correction and is even improved after the fit. The uncertainty on a single process neglects the correlations among samples and can therefore be larger than the uncertainty on the total prediction. An example is the $f_{\text{HF}}$ parameter that totally anti-correlates the $t\bar{t}$ + LF and $t\bar{t}$ + HF sample yields (cf. Equation 6.2). The $f_{\text{HF}}$ parameter also causes the $t\bar{t}$ + HF post-fit uncertainty to be larger than its pre-fit value because it projects the uncertainties of the much larger $t\bar{t}$ + LF background onto $t\bar{t}$ + HF.

Figure 7.2 compares the post-fit yields of a few charged Higgs boson mass hypotheses. The predicted $t\bar{t}(ud)$ and $t\bar{t}(cs)$ yields vary by about 10 %, which is consistent with their uncertainties. Their variations are not strictly correlated with each other. The $t\bar{t}$ + HF yield variations are a bit larger, by about 15 %. The remaining backgrounds show only small variations (total variation of $\sim$ 2 %).

#### Unconstrained parameters

The fitted values of the unconstrained fit parameters are presented in Figure 7.3. Of most interest are the values for $\mathscr{B}_{H^{\pm}}$. They are in the range from 0 to 1.43 %. For five mass points, 60, 70, 90, 150 and 168 GeV, a value of zero has been fitted. The largest value is observed for the 80 GeV mass point, but it is associated with a large uncertainty. The 110 GeV mass point shows the most significant deviation from zero with $\mathscr{B}_{H^{\pm}} = (0.147 \pm 0.098)$ %. Interestingly, the neighbouring mass points also have fitted $\mathscr{B}_{H^{\pm}}$ values larger than zero. If a signal were present at a given mass point, it is expected to be observed as smaller enhancements in neighbouring mass point hypotheses, too – spread over multiple neighbouring mass points. This is the case because the mass point step size is smaller than the $m_{j_1 j_2}$ resolution and the

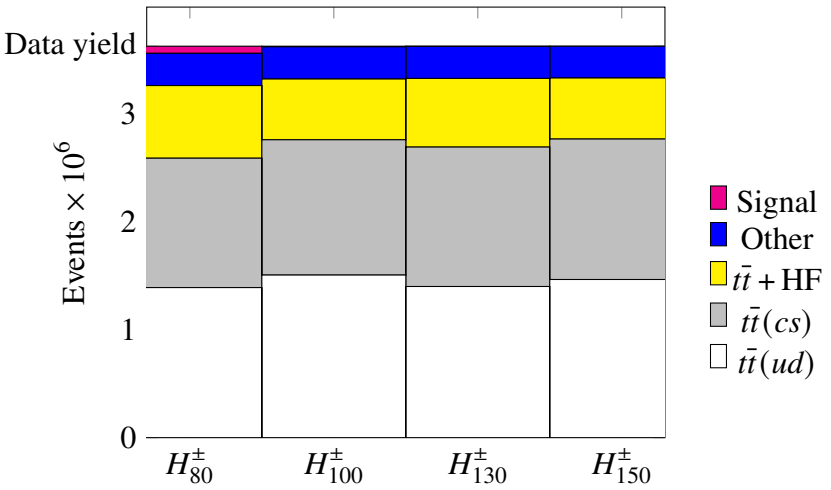
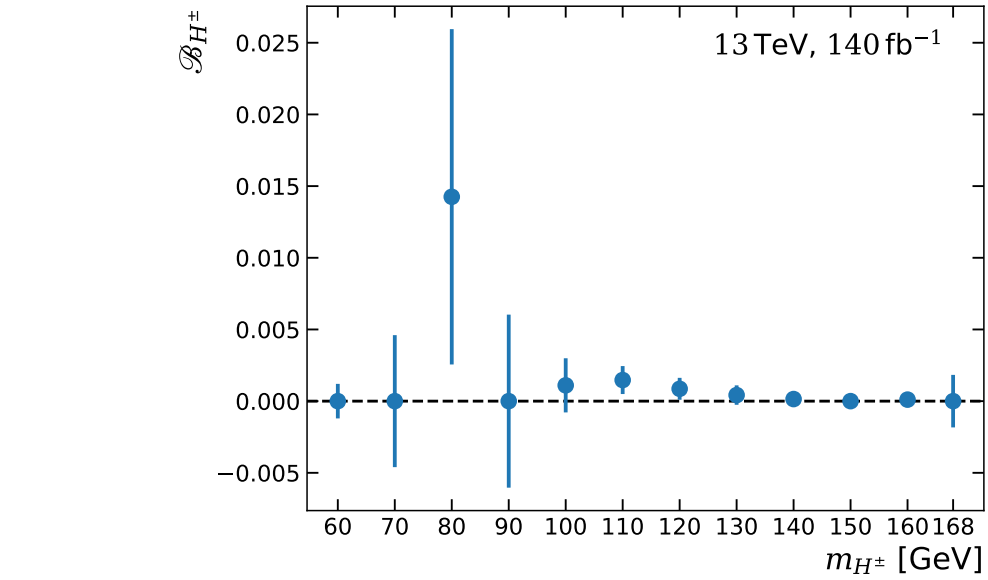



Figure 7.2: Vertical bar chart comparing the post-fit yields of processes for fits to the different BDT score distributions.

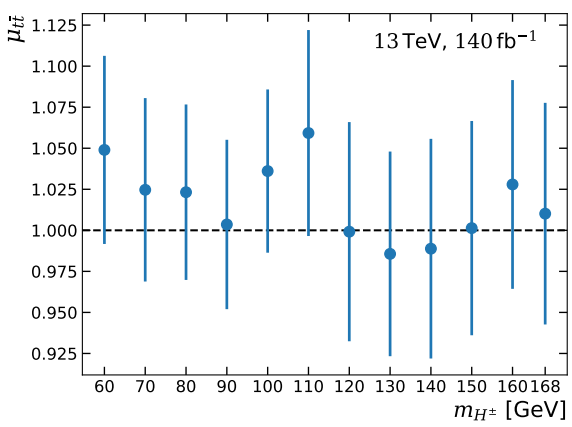
kinematic differences learnt by the BDT are relatively similar. This has been proven explicitly in signal injection tests using Asimov datasets with a fixed signal contribution. An exception here is the 80 GeV mass point The kinematic differences between the charged Higgs boson and $W$ boson are minor and have low importance only in this single mass point BDT training.

The errors on $\mathcal{B}_{H^{\pm}}$ have to be taken with a grain of salt because symmetric errors are assumed in the error calculation, but the extracted values are close to the border of 0, which is the physical lower bound of $\mathcal{B}_{H^{\pm}}$. Though, in fits where $\mathcal{B}_{H^{\pm}}$ has been unbound to allow for negative values, no change in the size of the error bars has been visible. In the unbounded fits, the mass points that previously received values of $\mathcal{B}_{H^{\pm}} = 0$ got negative values for $\mathcal{B}_{H^{\pm}}$, but all agreed well with zero within their uncertainty. Anyhow, the discovery significance presented in Section 7.3 is the meaningful quantity in expressing if there is a signal in the data.

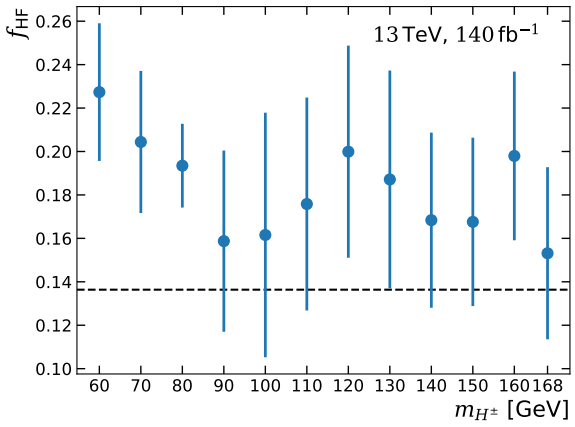
The cause for the large $\mathcal{B}_{H^{\pm}}$ value in the fit to the $H^{\pm}_{80}$ BDT score, in comparison to the other mass points, has been investigated in more detail since the pre-fit agreement in the high BDT score is already very good (cf. Figure 6.4). The $t\bar{t}(ud)$ PS and $t\bar{t}(cs)$ FSR uncertainties show the largest pulls in Figure A.21 (excluding the MC statistical uncertainty) and have large shape differences with respect to the nominal prediction. The $t\bar{t}(ud)$ PS NP is especially conspicuous because it is also heavily constrained. The systematic templates of both uncertainties are shown in Figure A.23. The shape difference of the $t\bar{t}(ud)$ PS uncertainty in the low BDT score region matches exactly the disagreement before the fit (first 10 bins in Figure 6.4), which explains the strong pull. In the high BDT score region, both uncertainties, given their positive pulls, reduce the predicted yields. The higher the BDT score, the larger the reduction is. This is an almost perfect anti-correlation to the signal distribution in the high BDT score region. The signal can compensate for the reduction in the high BDT score region without changing the agreement in the low BDT score region. This explains the large value for $\mathcal{B}_{H^{\pm}}$ and its large uncertainty, given that a pull of the two systematic uncertainties by $1\sigma$ removes the signal contribution almost completely. In future searches for a charged Higgs boson, pulls and constraints on the $t\bar{t}(ud)$ PS uncertainty should be monitored carefully. It is interesting to see whether Herwig (the alternative PS generator) is actually better at modelling the discussed phase space or this observation is just a coincidence.




(a)




(b)




(c)

Figure 7.3: Values of (a) $\mathscr{B}_{H^{\pm}}$, (b) $\mu_{t\bar{t}}$ and (c) $f_{\mathrm{HF}}$ for the tested charged Higgs boson mass hypotheses, extracted from fits to the respective BDT scores.

As mentioned before, the $t\bar{t}$ correction leads to a good pre-fit agreement of the overall $t\bar{t}$ normalisation. This is reflected in the fitted values of $\mu_{t\bar{t}}$ (Figure 7.3(b)) which all agree with their pre-fit prediction of 1 within uncertainties. The size of the uncertainties on $\mu_{t\bar{t}}$ are comparable to the relative uncertainty on the $t\bar{t}$ cross-section from dedicated measurements [220].

The fitted values of $f_{\mathrm{HF}}$ (Figure 7.3(c)) are all larger than predicted by simulation. The value of $f_{\mathrm{HF}}$ is determined most precisely for the 80 GeV mass point because the respective BDTs' predictive power is based predominantly on flavour-tagging information. For the 80 GeV mass point $f_{\mathrm{HF}}$ is measured to be

$f_{\mathrm{HF}} = 0.193 \pm 0.019$ and deviates by about $3\sigma$ from its prediction. The other BDT scores are not as sensitive to $f_{\mathrm{HF}}$ – the shapes of $t\bar{t}$ + LF and $t\bar{t}$ + HF are more similar – and their uncertainties are larger by a factor of 2 to 3. Most mass point fits prefer a smaller value for $f_{\mathrm{HF}}$ compared to the 80 GeV mass point, but all values agree with the 80 GeV mass point value within uncertainties.

The measurement for the 80 GeV mass point corresponds to an increase by a factor of about 1.4 in the $t\bar{t}$ heavy-flavour fraction. Similar findings have been observed in many other analyses with a $t\bar{t}$ final-state in ATLAS [91, 97, 263–265, 273].

### Constrained parameters

Figure 7.4 shows the pulls of constrained parameters in the fit with the $H^{\pm}_{130}$ signal sample. The difference of the fitted and initial value of NPs relative to their initial uncertainties is plotted in order to make the pulls of all NPs comparable. Most NPs are neither pulled nor constrained. Figure 7.4 shows only NPs that are pulled by at least 10 %. Most of the pulls are below 50 %. Only a handful of NPs show pulls that do not agree with the initial value within uncertainty. Just one NP is pulled by more than $1\sigma$ and none is being pulled by more than $2\sigma$. The affected NPs are mainly related to the jet reconstruction and the modelling of simulated $t\bar{t}$ events. The $t\bar{t}$ generator and parton shower uncertainties show the largest constraints, ranging between 75 and 95 % of their initial uncertainty. NPs related to flavour-tagging, lepton reconstruction, $E_{\mathrm{T}}^{\mathrm{miss}}$ and the modelling of non-$t\bar{t}$ processes are almost unaffected. Pulls on NPs for the MC statistical bin uncertainty are commonly observed, as they can improve the agreement in a bin without affecting other bins. In Figure 7.4 it can be seen that mainly the NPs of the high BDT score bins are pulled. This is related to the small MC sample sizes in those bins.

The outlined conclusions are valid for all mass point fits. BDT scores with fewer bins show, in general, less pulls and constraints. Plots similar to the one in Figure 7.4 for the 80, 100 and 150 GeV mass points are shown in Figure A.21 in the Appendix. Interesting to see is that even for the 80 GeV mass point the

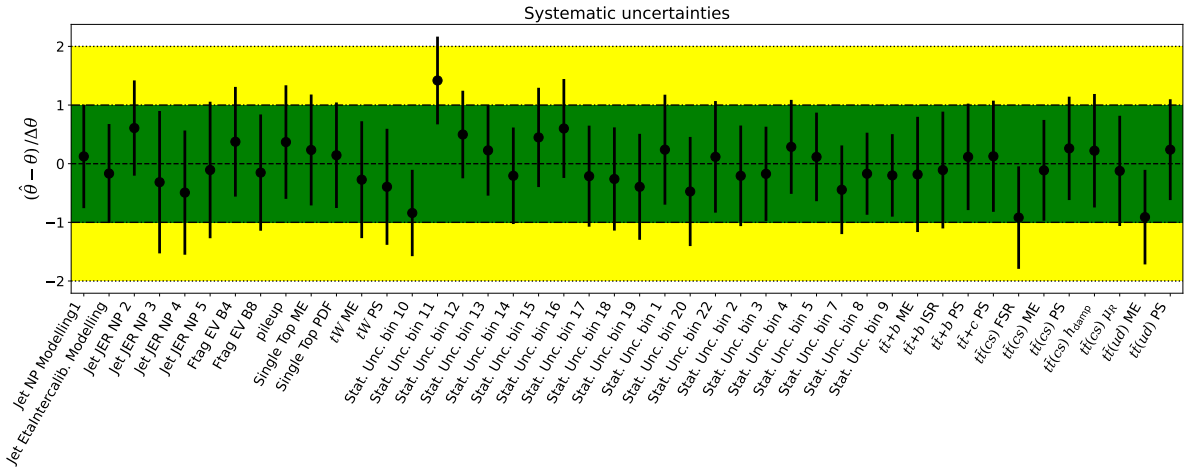



Figure 7.4: Pull plot of constrained parameters in the fit with the $H^{\pm}_{130}$ signal sample. $\theta$ and $\Delta\theta$ are the initial value and its uncertainty, respectively, and $\hat{\theta}$ is the value estimated from the fit. Only parameters that are pulled by more than 10 % of their uncertainty are listed. The NPs corresponding to the MC statistical uncertainties, "Stat. Unc. bin #", are indexed starting from zero. The $H^{\pm}_{130}$ has 24 bins in total, i.e. 23 corresponds to the highest BDT score bin.

flavour-tagging NPs show no significant pulls or constraints. This hints towards a solid flavour-tagging calibration.

## 7.2 Impact of statistical and systematic uncertainties

The impact of systematic uncertainties on the fit results are measured in two different ways. In the first method, the impact of a whole group of systematic uncertainties is measured. This helps to identify the sources that most severly limit the analysis sensitivity, like the reconstruction of an object, the modelling of a process, or the overall statistics. And it gives an indication of how much the analysis can be improved, for example, by increasing the MC sample sizes. In the second method, the individual systematic components that impact the fit results most are investigated.

#### Impact of groups of systematic uncertainties

The impact of a group of systematic uncertainties is determined by fixing the relevant NPs to their post-fit values from the nominal fit and then reperforming the fit. The value of $\mathcal{B}_{H^{\pm}}$ is unchanged in the new fit ($\mathcal{B}_{H^{\pm}} = \hat{\mathcal{B}}_{H^{\pm}}$), but the uncertainty on $\mathcal{B}_{H^{\pm}}$ is reduced ($\Delta\mathcal{B}_{H^{\pm}} < \Delta\hat{\mathcal{B}}_{H^{\pm}}$) because the fit has less flexibility. The impact of NPs on $\Delta\mathcal{B}_{H^{\pm}}$ cannot be retrieved by simply fixing all other NPs that are not part of the group of systematic uncertainties because the data statistical uncertainty is essential in the likelihood and will always contribute to the fitted $\Delta\mathcal{B}_{H^{\pm}}$.

The relative contribution of a group of systematic uncertainties to the total uncertainty (*Relative $\Delta\mathcal{B}_{H^{\pm}}$ contribution*) is measured as

$$\frac{(\Delta\hat{\mathcal{B}}_{H^{\pm}})^2 - (\Delta\mathcal{B}_{H^{\pm}})^2}{(\Delta\hat{\mathcal{B}}_{H^{\pm}})^2}.$$

The numerator is an approximation for the squared uncertainty on $\mathcal{B}_{H^{\pm}}$ caused by the group of systematic uncertainties under consideration and assumes that the NPs are uncorrelated with NPs from the other groups. The relative contribution to the total uncertainty is derived by dividing it by the nominal uncertainty squared. The results on the relative contributions to $\Delta\mathcal{B}_{H^{\pm}}$ are listed in Table 7.2 for fits to the $H^{\pm}_{80}$ and $H^{\pm}_{150}$ BDT score distributions. The $H^{\pm}_{80}$ and $H^{\pm}_{150}$ BDT scores are chosen because one is insensitive and the other is very sensitive to kinematic quantities.

The core conclusion from Table 7.2 is that the analysis sensitivity is limited by systematic uncertainties. The data statistical uncertainty is estimated by fixing all NPs in the fit. It is higher for the $H^{\pm}_{150}$ BDT score because the highest BDT score bins have fewer events ($\sim 20$) than in the $H^{\pm}_{80}$ BDT score ($\sim 6\,500$). This is the result of the binning algorithm and is mainly driven by the differences in the separation power of the BDT scores. The binning also directly relates to the MC statistical uncertainty. Hence, a similar picture as for the data statistical uncertainty is expected. But the MC statistical uncertainty in this analysis is primarily driven by the signal sample size, which is higher in the high BDT score region. In fact, the signal sample size is relatively similar in all bins in the rebinned BDT score distribution. Therefore, the MC statistical uncertainty is largest for BDT scores with many bins.

The group of $t\bar{t}$ modelling uncertainties is ranked the highest for all mass points except 80 GeV. In the fit with the $H^{\pm}_{80}$ signal sample, the flavour-tagging uncertainties are more important, which is expected.

Table 7.2: Impact of data statistical and systematic uncertainties on this analysis. The impact is estimated using fits in which a group of systematic NPs are fixed to their nominal post-fit values. The uncertainty on $\mathscr{B}_{H^{\pm}}$ from the fit ($\Delta\mathscr{B}_{H^{\pm}}$) is compared with the uncertainty from the nominal fit $\Delta\hat{\mathscr{B}}_{H^{\pm}}$. The relative contribution to the uncertainty on $\mathscr{B}_{H^{\pm}}$ is calculated as $\left((\Delta\hat{\mathscr{B}}_{H^{\pm}})^2 - (\Delta\mathscr{B}_{H^{\pm}})^2\right)/(\Delta\hat{\mathscr{B}}_{H^{\pm}})^2$. The groups are presented in descending order based on their relative contribution. The sum of the individual systematic components differs from the total systematic uncertainty due to correlations between uncertainties of different groups. The uncertainty from data statistical uncertainties is determined from fits with all NPs fixed to their post-fit values. The relative reduction on $\Delta\mathscr{B}_{H^{\pm}}$ is calculated as $(\Delta\hat{\mathscr{B}}_{H^{\pm}} - \Delta\mathscr{B}_{H^{\pm}})/\Delta\hat{\mathscr{B}}_{H^{\pm}}$. For reference, $\Delta\hat{\mathscr{B}}_{H^{\pm}}$ is 1.2 % and 0.04 % for the fits with the $H^{\pm}_{80}$ and $H^{\pm}_{150}$ BDT score, respectively.

| $H^{\pm}_{80}$ | | | $H^{\pm}_{150}$ | | |
|---|---|---|---|---|---|
| Category | Relative $\Delta\mathscr{B}_{H^{\pm}}$ contribution | Relative $\Delta\mathscr{B}_{H^{\pm}}$ reduction | Category | Relative $\Delta\mathscr{B}_{H^{\pm}}$ contribution | Relative $\Delta\mathscr{B}_{H^{\pm}}$ reduction |
| **Data statistical** | 0.4 % | | **Data statistical** | 14 % | |
| **Systematic** | 99.6 % | 94 % | **Systematic** | 86 % | 62 % |
| Flavour-tagging | 41 % | 23 % | $t\bar{t}$ modelling | 52 % | 31 % |
| MC statistical | 41 % | 23 % | MC statistical | 12 % | 6.4 % |
| $t\bar{t}$ modelling | 25 % | 13 % | Weak-boson & MJ modelling | 7.1 % | 3.6 % |
| $\mu_{t\bar{t}}$ & $f_{\mathrm{HF}}$ | 4.5 % | 2.3 % | Single-top-quark modelling | 6.3 % | 3.2 % |
| Jet | 3.7 % | 1.9 % | $\mu_{t\bar{t}}$ & $f_{\mathrm{HF}}$ | 5.7 % | 2.9 % |
| Single-top-quark modelling | 2.5 % | 1.2 % | Jet | 5.2 % | 2.6 % |
| Luminosity & Pile-up | 2.3 % | 1.2 % | Flavour-tagging | 3.9 % | 2.0 % |
| Weak-boson & MJ modelling | 1.5 % | 0.8 % | Lepton & $E_{\mathrm{T}}^{\mathrm{miss}}$ | 0.6 % | 0.3 % |
| Signal modelling | 0.7 % | 0.3 % | Luminosity & Pile-up | 0.5 % | 0.3 % |
| Lepton & $E_{\mathrm{T}}^{\mathrm{miss}}$ | 0.5 % | 0.2 % | Signal modelling | 0.2 % | 0.1 % |

The jet uncertainties are especially highly ranked for mass points around the $W$-boson mass, given the dijet mass resolution. For these mass points it plays a vital role whether one can reliably separate the mass peaks from signal and SM $t\bar{t}$ events.

The groups of systematic uncertainties not mentioned so far are subdominant. The modelling uncertainties of single top-quark and $W$+jets processes contributes a few percent to the total uncertainty. Luminosity, pile-up, lepton, $E_{\mathrm{T}}^{\mathrm{miss}}$and signal modelling systematic uncertainties are almost negligible.

The fixation of NPs in the reperformed fits implies perfect knowledge of their values. By calculating

$$\frac{\Delta\hat{\mathscr{B}}_{H^{\pm}} - \Delta\mathscr{B}_{H^{\pm}}}{\Delta\hat{\mathscr{B}}_{H^{\pm}}}$$

one retrieves the expected reduction on $\Delta\mathscr{B}_{H^{\pm}}$ if this would be the case. This is also shown in Table 7.2 in the column "Relative $\Delta\mathscr{B}_{H^{\pm}}$ reduction". The results are closely related to what has already been discussed. From the extra column one can easily infer by how much the analysis could improve if a given group of uncertainties were better understood and constrained. As an example, if the $t\bar{t}$ background could be perfectly modelled, it would reduce the uncertainty on $\mathscr{B}_{H^{\pm}}$ in this analysis by 31 % for the $H^{\pm}_{150}$ BDT score fit. Interesting in this column is that without any systematic uncertainties, the uncertainty on $\mathscr{B}_{H^{\pm}}$ would be only 6 % and 38 % of its current value for the $H^{\pm}_{80}$ and $H^{\pm}_{150}$ BDT score fit, respectively. A similar statement cannot be made about the data statistical uncertainty. This would correspond to the case of infinite data. In this case the systematic uncertainties can be determined perfectly, too, and

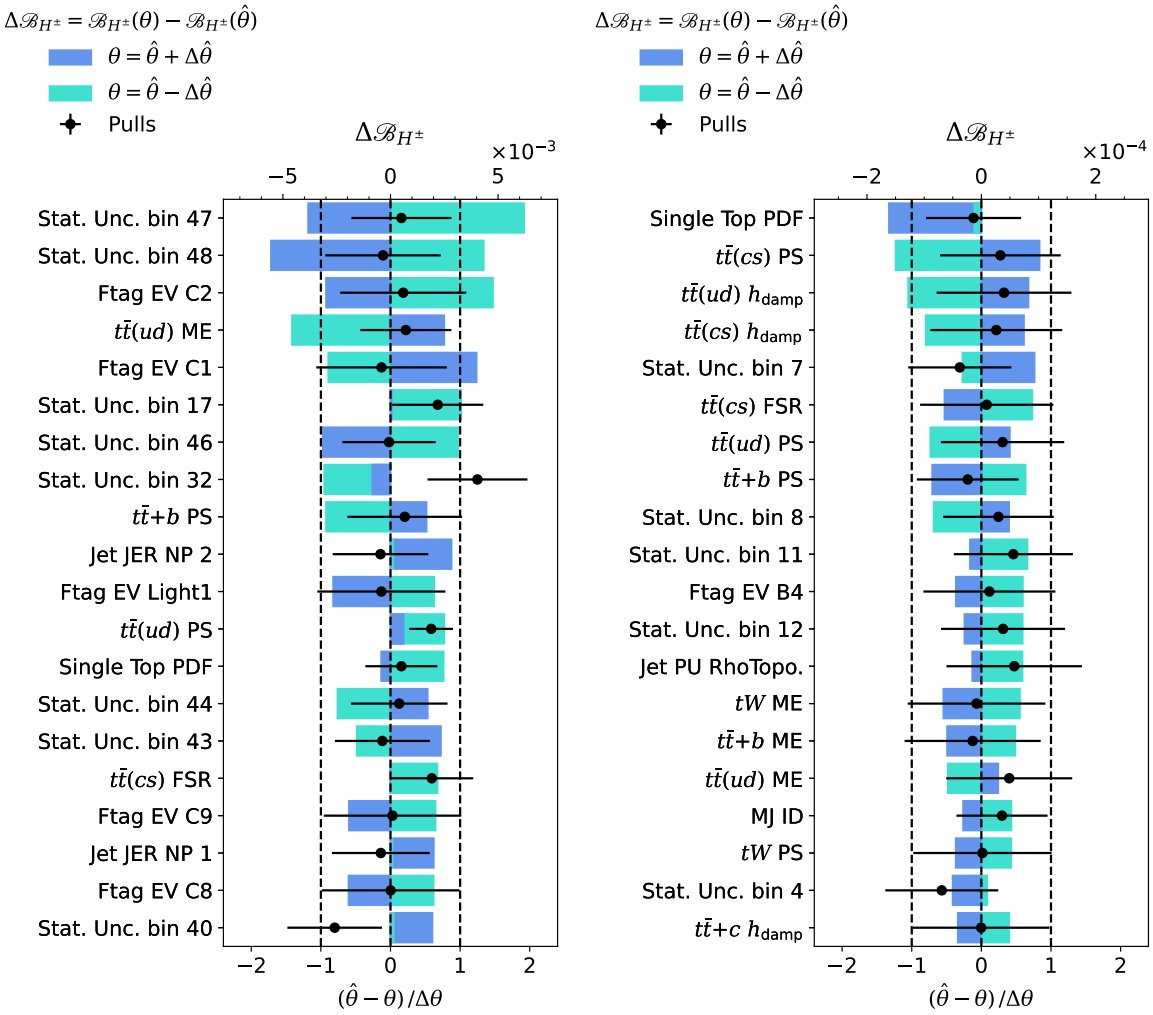



Figure 7.5: Ranking of NPs according to their impact on the measured $\mathcal{B}_{H^\pm}$ values for fits to the (a) $H^{\pm}{}_{80}$ and (b) $H^{\pm}{}_{150}$ BDT score distributions. The impact is determined by fixing the NP under consideration to its nominal post-fit $\pm 1\sigma$ uncertainty value ($\theta = \hat{\theta} \pm \Delta\hat{\theta}$). $\Delta\mathcal{B}_{H^\pm}$ is the difference of the fitted $\mathcal{B}_{H^\pm}$ with respect to the nominal $\mathcal{B}_{H^\pm}$ value ($\hat{\mathcal{B}}_{H^\pm}$). The black points with the uncertainty bars represent the NP pulls relative to their nominal values ($\theta_0$) and uncertainties ($\Delta\theta_0$).

$\Delta\mathcal{B}_{H^\pm}$ would simply vanish.

### Impact of individual systematic uncertainties

In the second method the change in the fitted $\mathcal{B}_{H^\pm}$ value is monitored when fixing one NP to its $\pm 1\sigma$ variation of the nominal fit: $\theta = \hat{\theta} \pm \Delta\hat{\theta}$. The fit is reperformed and the difference with respect to $\hat{\mathcal{B}}_{H^\pm}$ is calculated. The 20 NPs with the largest impact on $\mathcal{B}_{H^\pm}$ are listed in descending order in Figure 7.5, regardless whether they cause an up or down variation on $\mathcal{B}_{H^\pm}$. The pulls of the NPs from the nominal fit are superimposed to read off the values the NPs are fixed to in the new fits. In these fits the $\mathcal{B}_{H^\pm}$ are

allowed to become negative. Otherwise, for mass points in which $\mathscr{B}_{H^{\pm}}$ is fitted to zero, but a $\mathscr{B}_{H^{\pm}}$ even below zero is preferred by data, the plots have little explanatory power. Only NPs for which their impact would be so large that the best-fit $\mathscr{B}_{H^{\pm}}$ becomes positive showed variations. Only a handful of such NPs for the 90 and 168 GeV mass points were seen.

The NPs in Figure 7.5 are mostly part of the groups of systematic uncertainties that are highly ranked in Table 7.2. Most noticeable is that for all mass points, except for 80 GeV, the $t\bar{t}$ ME, PS, $h_{\text{damp}}$ and FSR uncertainties are among the highest-ranked NPs. These uncertainties generally show the largest shape differences in the high BDT score region. But the form of the shape differences varies largely for the different mass points, without a clear trend. Which NP has the highest impact depends on the (anti-)correlation with the signal shape in the high-BDT-score region.

The MC statistical uncertainties of the highest BDT score bins are usually ranked high, as the respective bins are most sensitive to a signal.

Jet uncertainties are, as already pointed out, more important for mass points around the $W$-boson mass. Ranking plots for the 100 GeV and 130 GeV mass points can be found in the Appendix in Figure A.22. Between 4 to 5 jet-related NPs are among the top twenty. These are predominantly related to systematic uncertainties on the jet energy resolution calibration. In comparison, only 1 to 2 jet-related NPs are in Figure 7.5 (80 GeV and 150 GeV mass points).

Flavour-tagging NPs related to $c$-tagging efficiencies are only relevant for the 80 GeV mass point. For other mass points, the $b_{\text{had}}$-jet identification is crucial. NPs related to $b$-tagging efficiencies increase in importance but still have an overall low influence on $\mathscr{B}_{H^{\pm}}$.

Systematic components not discussed so far have a small impact on the analysis and usually don't show up among the highest-ranked NPs.

## 7.3 Discovery significance and upper limits

### Discovery significance

The results of the discovery tests for the different charged Higgs boson mass hypotheses are plotted in Figure 7.6. The local $p_0$ values, calculated as given in Equation 6.5, are shown on the $y$-axis in logarithmic scale. The dashed lines indicate the discovery significance for $Z$ equal to zero, one and two.

First of all, no discovery can be claimed, as the significances for all mass points are below 5. The largest local significance is observed for the 110 GeV mass point with a value of about 1.5. This corresponds to a probability of 5 % to observe the given data under the background-only hypothesis. Given that twelve different mass hypotheses are tested, this is most likely a statistical fluctuation and no sign of a charged Higgs boson with a mass of 110 GeV. A global significance of this search – the observed peak at 110 GeV – would be even smaller than 1.5.

The next highest significances are observed for the 80 and 120 GeV mass points. Even though a significant signal contribution has been predicted by the fit to the 80 GeV mass point, a large $p_0$ value of 0.1 is retrieved. This matches the expectation given the large uncertainty on $\mathscr{B}_{H^{\pm}}$. The differences between signal and background are small for that mass point and are covered by systematic uncertainties.

All other mass points have a discovery significance below 1. The mass points with a measured $\mathscr{B}_{H^{\pm}}$ of zero receive a $p_0$ value of 0.5. This is because $\mathscr{B}_{H^{\pm}}$ is capped at zero and does not allow for downward fluctuations; $f(q_0|0)$ has a delta function at zero with size 0.5.

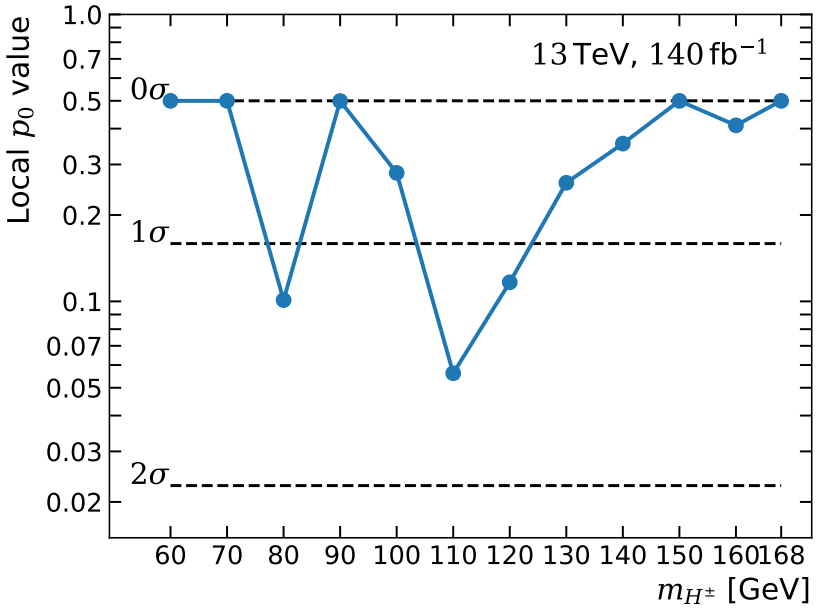



Figure 7.6: Local $p_0$ values for a charged Higgs boson signal at different mass hypotheses. The dashed lines indicate the discovery significance – $2\sigma$ corresponds to $Z = 2$, etc. The calculation of the $p_0$ value and discovery significance is outlined in Section 6.2.2.

The observed distribution of $p_0$ values agrees well with randomly drawing from a standard normal distribution. When drawing from a normal distribution, 50 % of values are below zero, about 34 % are in between 0 and $1\sigma$, and about 16 % are above $1\sigma$. Multiplying the quoted probabilities by twelve, the number of mass points the discovery significance has been calculated for, one obtains the expectations of roughly 6, 4.1 and 1.9 for the three intervals. The observed numbers of mass points exhibiting discovery significances in these intervals are 5, 4, and 3, respectively, and are therefore quite similar. This agreement is another strong indication that no signal is present in the data – or it is so small that the analysis is not sensitive enough to detect it – since a signal would lead to enhanced $p_0$ values in multiple neighbouring mass points.

#### Upper limits on $\mathscr{B}(t \to H^{\pm}b)$

The results of the upper limit calculation on the branhcing ratio of a top-quark decaying into a charged Higgs boson and a $b$-quark, assuming $\mathscr{B}(H^{\pm} \to cs) = 100\,\%$, are presented in Figure 7.7. The limits are calculated using the $CL_s$ technique as described in Section 6.2.2 and quoted at 95 % confidence level. The expectations and their 1 and $2\sigma$ uncertainties are derived from the background-only Asimov dataset. The black solid line represents the observed limits in data. The limits derived at different mass hypotheses are connected by straight lines (linear interpolation). This is motivated by tests in which the signal sample of a neighbouring mass point has been used in the upper limit calculation, e.g. the $H^{\pm}_{120}$ signal sample was fitted to the $H^{\pm}_{130}$ BDT score. The results were worse by just a few percent, or at least much smaller than the $1\sigma$ uncertainty. Hence, the limits in the middle between two mass points hypotheses are overconfident, but only marginally.

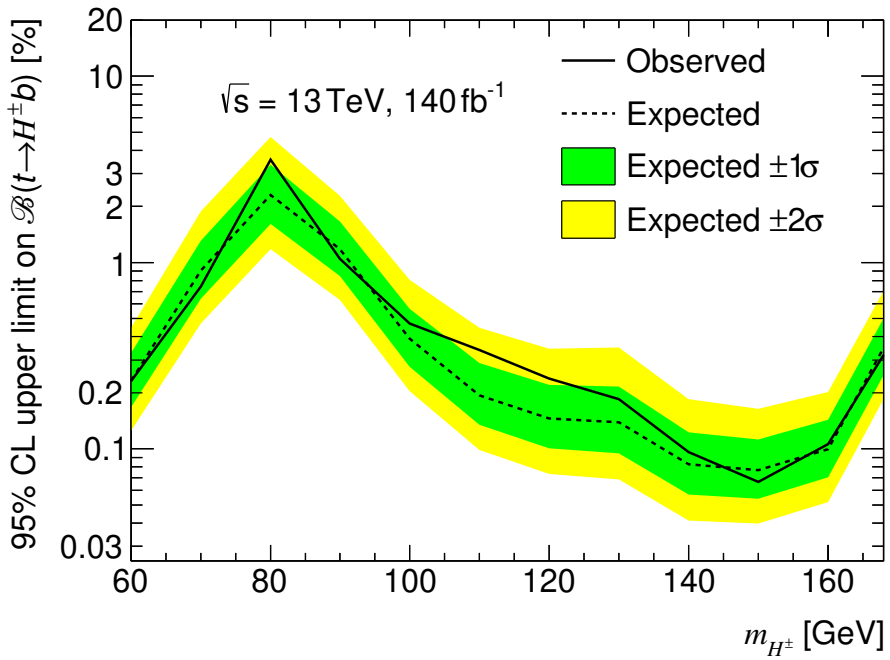



Figure 7.7: Upper limits on $\mathcal{B}(t \to H^{\pm}b)$, assuming $\mathcal{B}(H^{\pm} \to cs) = 100\,\%$, quoted at 95 % confidence level calculated using the $CL_s$ technique (cf. Section 6.2.2). The dashed line represents the expected limits from Asimov data and the green and yellow bands are its 1 and $2\sigma$ error bands, respectively. The solid line represents the limits extracted from the observed data.

The overall shape of the limits is characterised by the $W$-boson mass peak of the $t\bar{t}$ background. Around the $W$-boson mass the limits are worst and rapidly improve for mass points further away. The strongest limits are set for the 150 GeV mass point. The mass points closer to the top-quark mass are further away from the $W$-boson mass peak but suffer from a reduced signal selection efficiency (cf. Figure 4.4) due to the low momentum of the $b_{\text{had}}$-jet. The expected limits are in the range from 0.077 to 2.3 %. The range for the observed limits is a bit larger with 0.066 to 3.6 %.

The small kink in the expected limits at 130 GeV is presumably related to a large impact of the $t\bar{t}$ ME systematic uncertainty of the $t\bar{t}(ud)$ process (cf. Figure A.22). The shape variation of the aforementioned NP is similar to the signal shape in the high BDT score region. This is expected to be a coincidence because the shape variation of the $t\bar{t}$ ME systematic uncertainty of the $t\bar{t}(cs)$ process is much different.

Expected and observed limits agree well within uncertainties. This comes as no surprise, given the discovery significances showed no anomalies. The fact that the observed data even preferred negative $\mathcal{B}_{H^{\pm}}$ values in a few unbounded fits reflect in improved observed limits with respect to the expected one for the affected mass points. To mention in that regard are the 70 and 150 GeV mass points. In both cases the discrepancies are still well covered by the $1\sigma$ uncertainties on the expectation. Similarly, worse observed limits are seen for mass points with positive fitted $\mathcal{B}_{H^{\pm}}$ values. The discrepancy between observed and expected agrees well with the findings from the discovery significance.

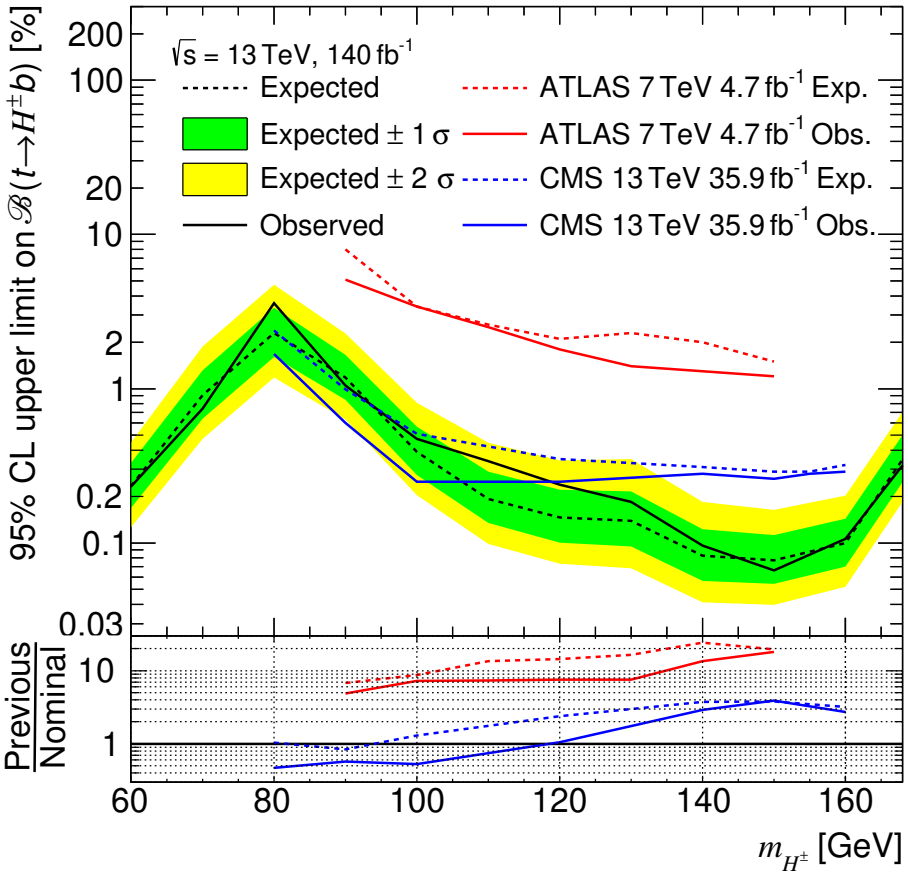



Figure 7.8: Comparison of the upper limits on $\mathcal{B}(t \to H^{\pm}b)$, assuming $\mathcal{B}(H^{\pm} \to cs) = 100\,\%$, of this analysis with previous analyses searching for $H^{\pm} \to cs$ from ATLAS [89] and CMS [90]. The limits are quoted at 95 % confidence level and are calculated using the $CL_s$ technique. The ratio panel compares the limits of the two previous analyses to this analysis ("Nominal").

## 7.4 Comparison with previous results

The upper limits on $\mathcal{B}_{H^{\pm}}$ from this analysis are compared with previous results in Figure 7.8. A comparison is made with the predecessor search performed by ATLAS with a dataset collected at $\sqrt{s} = 7\,\text{TeV}$ and an integrated luminosity of $4.7\,\text{fb}^{-1}$ [89], and with the most recent search performed by CMS with a dataset collected at $\sqrt{s} = 13\,\text{TeV}$ and an integrated luminosity of $35.9\,\text{fb}^{-1}$ [90]. Both analyses and their strategies have been outlined in Section 2.3.2. The CMS search previously set the most stringent limits on $\mathcal{B}(t \to H^{\pm}(\to cs)b)$.

The previous ATLAS analysis set limits in the mass range from 90 to 150 GeV. The analysis presented in this thesis could improve the expected limits in that mass range with respect to the previous ATLAS analysis by factors between 8 to 25, depending on the mass point. The improvement in the observed limits is a bit less with factors between 5 to 20. This substantial improvement is caused by multiple factors. Among them are the increase in luminosity, improvements in flavour tagging, especially the

introduction of $c$-tagging, and the usage of multivariate analysis methods. The BDT can improve the limits especially well for mass points further away from the $W$-boson mass. For these mass points the kinematic quantities are most important and the BDT is able to utilise correlations among them (the previous search utilises only $m_{j_1 j_2}$).

The CMS analysis set limits in the mass range from 80 to 160 GeV. The analysis presented in this thesis could improve the expected limits with respect to the CMS analysis by factors up to 4. The improvements at higher charged Higgs boson masses are, as before, mainly due to the BDT. For the 150 GeV mass point, the limits from the CMS analysis are weaker by around $3.5\sigma$.

Close to the $W$-boson mass, the limits are roughly identical, even though an improvement is expected due to the increased data sample size. Flavour-tagging information is most important for these mass points and the CMS analysis has (on top of a $b$-tagging WP) three $c$-tagging WPs. The additional $c$-tagging WP helps in rejecting $t\bar{t}(ud)$ events. Another reason why the limits are not better, which affects all mass points, is a difference in the set of the systematic uncertainties. The considered uncertainties are in general very similar, but the CMS analysis has no uncertainty due to the choice of the generator type/model for modelling of ME and PS in $t\bar{t}$ events. In Section 7.2 it was shown that the analysis is dominated by systematic uncertainties and that the $t\bar{t}$ ME and PS uncertainties are always among the highest-ranked uncertainties. Henceforth, the more conservative approach in this analysis in comparison to the CMS analysis has a significant impact on the results and has to be taken into consideration when comparing both analyses.

For the observed limits, the CMS limits turn out to be more stringent in the mass range from 80 to 115 GeV. Around 80 GeV the limits are about 50 % smaller. Above that mass range the observed limits from this analysis are more stringent, up to a factor of 3 at 150 GeV.

Throughout this analysis it has been mentioned a few times that at the beginning of the analysis development a different flavour-tagging calibration was used. This calibration is also described in Section 3.4.6 and is referred to as "PCBT". The main difference is that it has four calibrated $b$-tagging WPs and no $c$-tagging WPs. This flavour-tagging calibration had been used because at that time no calibrations with calibrated $b$- and $c$-tagging WPs were available. The same analysis strategy as described in this thesis was applied, except for the usage of a looser $b$-tagging requirement in the event selection.

Figure 7.9 compares the limits from the analysis using the PCBT calibration with the nominal approach. The limits of the PCBT analysis are weaker in the mass range from 60 to 140 GeV by around 10 to 40 %. The differences in the limits fluctuate across the quoted charged Higgs boson mass range and no clear trend is visible. On average the improvement is around 20 % in that mass range. Above that mass range, the limits roughly agree with each other. This result is expected, as the flavour-tagging information is less important for these mass points. The plot shows the improvement in the analysis thanks to the usage of $c$-tagging.

## 7.5 Sensitivity to a $H^{\pm} \to cb$ signal

At the beginning of the thesis it was already mentioned that this analysis was of large interest because of the $3\sigma$ local excess seen at 130 GeV by the $H^{\pm} \to cb$ search [91]. In the previous section it was shown that a similar excess is not observed in the $H^{\pm} \to cs$ search.

Figure 7.10 compares the limits from the two searches with each other. At 130 GeV the expected

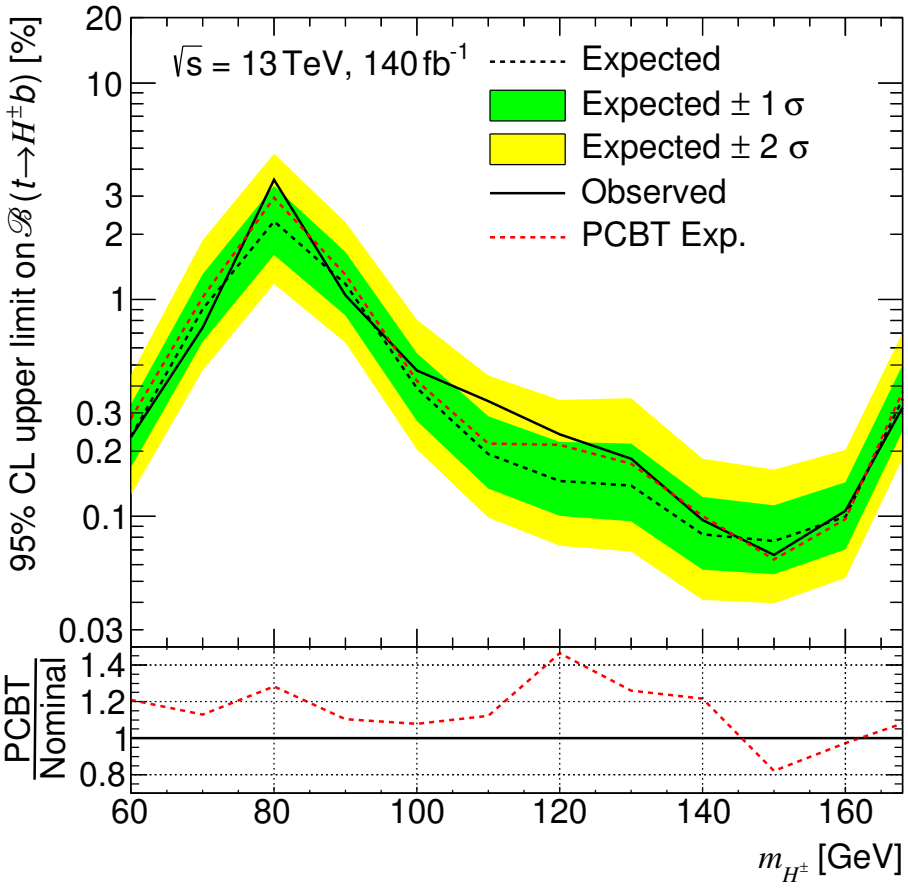



Figure 7.9: Comparison of the expected upper limits on $\mathcal{B}(t \to H^{\pm}b)$, assuming $\mathcal{B}(H^{\pm} \to cs) = 100\,\%$, set with the PCFT (nominal) and the PCBT calibration.

limits from the $H^{\pm} \to cs$ analysis are about 20 % lower, but the observed limits are more stringent by 40 %. This means the sensitivity is at a similar level, but no excess is found. This means either the branching ratio into $cb$ is significantly larger than into $cs$ or the found excess is just a statistical fluctuation (if systematic effects are not underestimated).

The limits in the $H^{\pm} \to cb$ search are much better around the $W$-boson mass because the decay $W \to cb$ is rare compared to $W \to cs$. In the $H^{\pm} \to cb$ search, $b$-tagging is efficiently utilised to separate signal and background events. Interestingly, the limits are almost uniform for the whole mass range. This means that the flavour-tagging information is also superior in the high mass range in the $H^{\pm} \to cb$ search. Above 135 GeV the expected limits from the $H^{\pm} \to cs$ are more stringent than the ones from the $H^{\pm} \to cb$ search. This is unexpected given the disadvantage on the flavour-tagging side and is due to a better utilisation of kinematic differences between signal and background events in this analysis.

The dedicated $H^{\pm} \to cb$ analysis uses the PCBT calibration for flavour-tagging and cannot, compared to this analysis, tag $c$-jets. Considering the better-suited flavour-tagging calibration and the improved utilisation of kinematic differences in this analysis, it is expected that the limits can be considerably improved over the whole mass range when applying this analysis to a $H^{\pm} \to cb$ signal. This conclusion

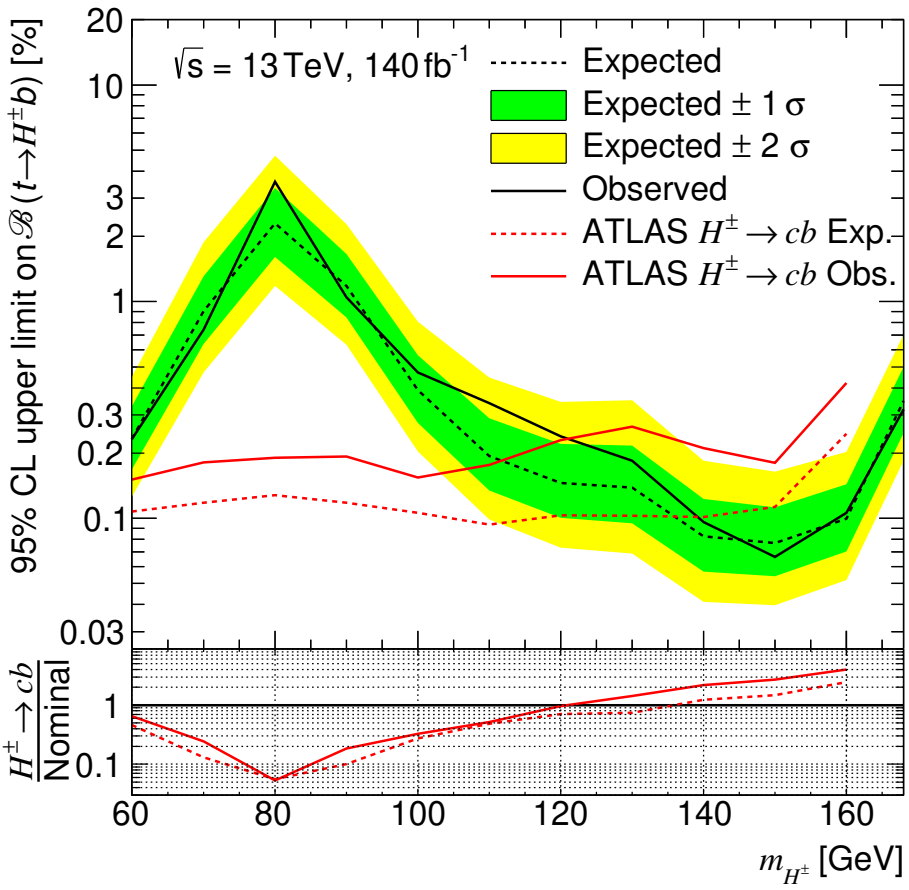



Figure 7.10: Comparison of the upper limits on $\mathcal{B}(t \to H^{\pm}b)$ set by this analysis and by the search for $H^{\pm} \to cb$ [91], assuming $\mathcal{B}(H^{\pm} \to cs) = 100\,\%$ and $\mathcal{B}(H^{\pm} \to cb) = 100\,\%$, respectively.

implies a retraining of the BDTs using $H^{\pm} \to cb$ events as signal.

Given that the upper limits on $\mathcal{B}_{H^{\pm}}$ of this $H^{\pm} \to cs$ analysis are compatible with the most stringent limits on $H^{\pm} \to cb$ at the 130 GeV mass point (Figure 7.10), where the excess was found, and the proximity of the event topologies and final states of these two charged Higgs boson decay channels, it was tested whether this analysis may be as sensitive to a $H^{\pm} \to cb$ signal as it is to a $H^{\pm} \to cs$ signal. If this is the case, it is possible to enhance or decrease the significance of the excess.

It is expected that the limits, without any retraining of the BDTs, would be weaker, as the flavour-tagging information learnt by the BDT will be different. In addition, the jet labelling procedure is optimised for the two jets from the charged Higgs boson being non-$b$-jets. The flavour-tagging vetoes might veto the correct combination much more often, and the assignment of the $b_{\text{had}}$-jet is less often correct. Changes in kinematic quantities due to the larger mass of the $b$-quark compared to the $s$-quark are minor, as $m_{H^{\pm}} \gg m_b - m_s$.

Since the signal samples from the dedicated $H^{\pm} \to cb$ analysis were not available anymore, new $H^{\pm} \to cb$ signal samples were produced, similar to the signal samples from this analysis (cf. Section 4.1.1).

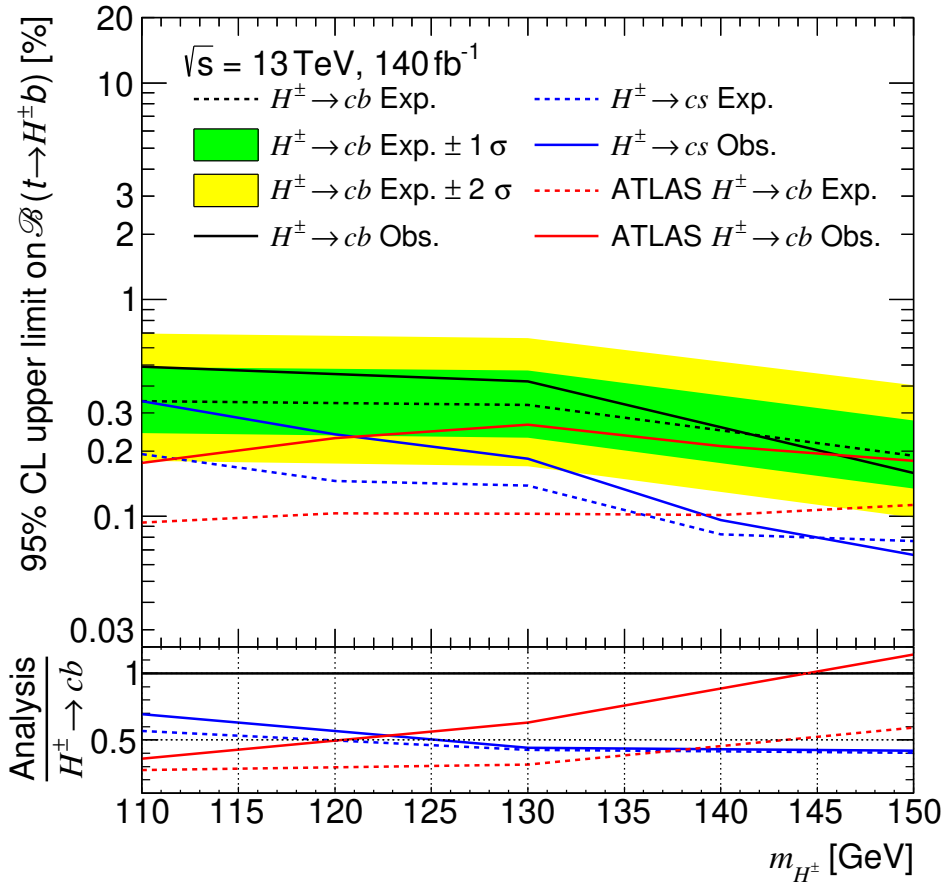



Figure 7.11: Upper limits on $\mathcal{B}(t \to H^{\pm}b)$ for a charged Higgs boson decaying to $cb$ are set by this analysis (denoted "$H^{\pm} \to cb$"). In the limit calculation $\mathcal{B}(H^{\pm} \to cb) = 100\,\%$ is assumed. The limits are compared to the limits for a $H^{\pm} \to cs$ signal from this analysis (denoted "$H^{\pm} \to cs$") and to the $H^{\pm} \to cb$ limits from the dedicated ATLAS analysis [91] (denoted "ATLAS $H^{\pm} \to cb$").

Samples for three mass points around 130 GeV were created: 110 GeV, 130 GeV and 150 GeV.

The resulting limits on $\mathcal{B}_{H^{\pm}}$ for a $H^{\pm} \to cb$ signal are shown in Figure 7.11. Exactly the same analysis strategy is applied to the $H^{\pm} \to cb$ signal sample (without retraining of the BDTs). In the limit calculation, $\mathcal{B}(t \to W^{\pm}b) + \mathcal{B}(t \to H^{\pm}(\to cb)b) = 1.0$ is assumed. The observed (expected) limit is in the range from 0.49 to 0.16 % (0.34 to 0.19 %). The observed limits agree with the expectations within $1\sigma$.

Figure 7.11 also compares the limits from this analysis on a $H^{\pm} \to cs$ and a $H^{\pm} \to cb$ signal. Due to the aforementioned reasons, the limits on $H^{\pm} \to cb$ are weaker by around a factor of 2.

In addition, Figure 7.11 compares the expected and observed limits on a $H^{\pm} \to cb$ signal from this analysis with the dedicated $H^{\pm} \to cb$ analysis. At the 130 GeV mass point, the expected limits from this analysis are a factor of 2.3 weaker. Both the expected and observed upper limits from this analysis are weaker than the observed upper limit from the dedicated analysis by 20 % and 40 %, respectively. This means that the analysis is unfortunately not sensitive enough to confirm the excess at 130 GeV. The

observed limit at the 150 GeV mass point from this analysis was improved by 14 % with respect to the dedicated analysis.

## 7.6 Interpretations in CP-conserving 2HDM models

This section presents the interpretations of the results on CP-conserving 2HDMs with natural flavour conservation, defined in Table 2.1. Exclusion limits on parameters can be set based on the retrieved upper limits on $\mathscr{B}(t \to H^{\pm}b) \times \mathscr{B}(H^{\pm} \to cs)$. Figure 7.12 shows exclusion limits in the $m_{H^{\pm}}$ - $\tan(\beta)$ plane calculated using the 2HDMC program [61][1].

The following assumptions are made in the calculations:

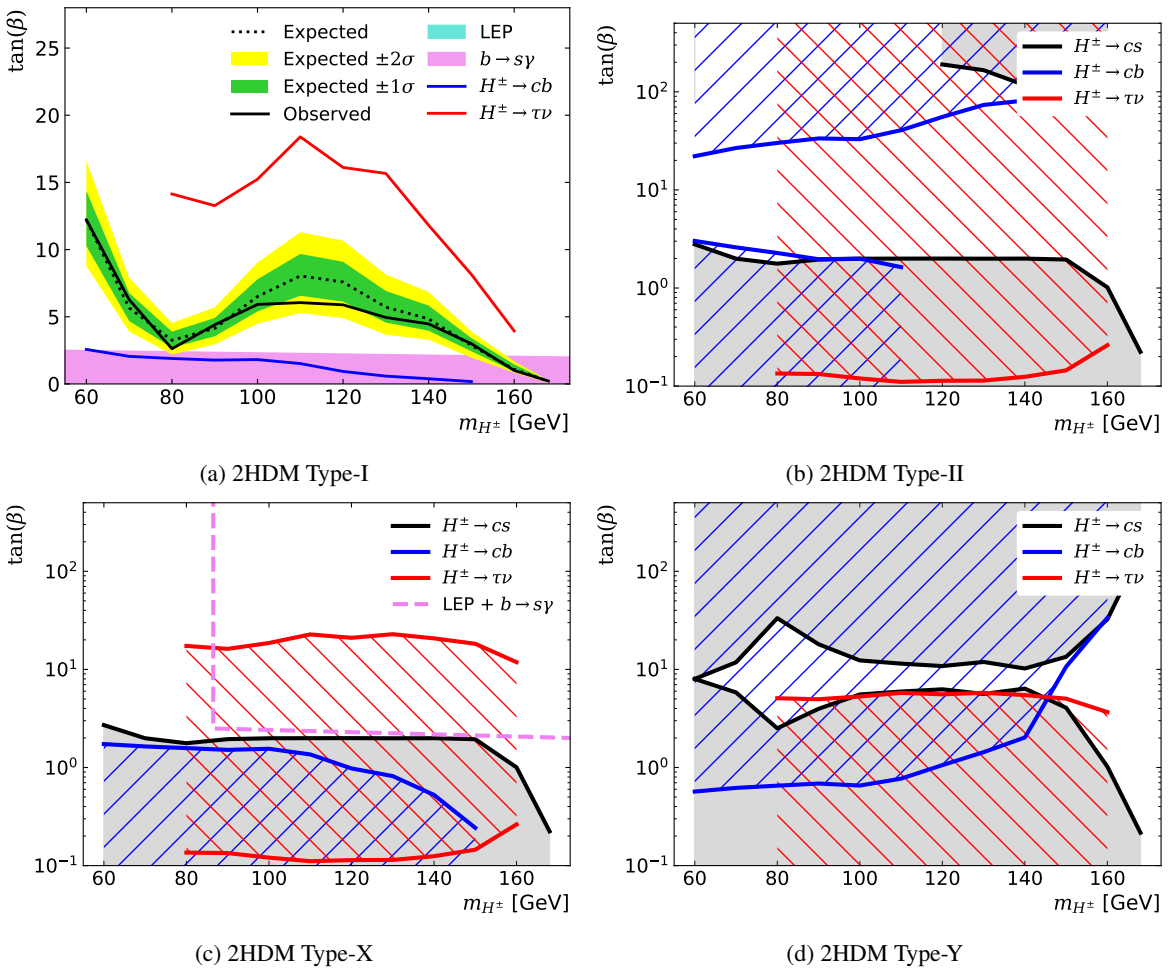
- Alignment limit: $m_h = 125\,\text{GeV}$ and $\sin(\beta - \alpha) = 1$;
- The mass of the pseudoscalar and charged Higgs bosons are identical: $m_A = m_{H^{\pm}}$;
- The mass of the heavy Higgs boson is identical to the mass of the pseudoscalar Higgs boson (but not smaller than $m_h$): $m_H = m_A$ ($m_H = 130\,\text{GeV}$ if $m_A < 130\,\text{GeV}$);
- $m_{12}^2 = m_A^2 \frac{\tan(\beta)}{1+\tan^2(\beta)}$.

Considering these assumptions, the Type-II model resembles the Higgs sector of the MSSM in the decoupling limit ($m_A \gg m_Z$). And the constraints on the MSSM are similar to the derived constraints on the Type-II model, given the SUSY particles are heavy [81].

Anyhow, the discussed assumptions have a minor influence on exclusion limits set by this search. The exclusion values depend only on the derived upper limits and the couplings of fermions to the charged Higgs boson. The couplings determine the branching ratios $\mathscr{B}(t \to H^{\pm}b)$ and $\mathscr{B}(H^{\pm} \to cs)$. The tree-level coupling of the charged Higgs boson to fermions depends only on $\tan(\beta)$. The important assumption that is made here is that no neutral Higgs boson is lighter than $m_{H^{\pm}} - m_W$. If this were not the case, bosonic decay modes of the charged Higgs boson would become relevant and the exclusion limits weaken. This assumption is supported by measurements that indicate that the found Higgs boson is the light scalar Higgs boson, which means that $H^{\pm} \to W^{\pm}H/h$ decay modes are not accessible for a light charged Higgs boson. In addition, the mass terms of charged Higgs bosons and pseudoscalar Higgs boson are very similar, making the decay $H^{\pm} \to W^{\pm}A$ unfavourable in most models.

The exclusion limits in the 2HDM Type-I model are shown in Figure 7.12(a). In Type-I 2HDMs – in the low mass region and under the discussed assumptions – the charged Higgs boson decay branching ratios are independent of $\tan(\beta)$. But the branching ratio $\mathscr{B}(t \to H^{\pm}b)$ steadily decreases with increasing $m_{H^{\pm}}$, because up- and down-type quarks both couple to the same doublet – the charged Higgs boson decouples from fermions for $\tan(\beta) \gg 1$. For these reasons only small values of $\tan(\beta)$ can be excluded. The parameter space below the solid black line can be discarded by this search. The observed exclusion limits vary between 12.2 at the 60 GeV mass point and 0.2 at the 168 GeV mass point.

[1] The default values for the top- and bottom-quark masses are 172.5 GeV and 4.75 GeV, respectively. In order to calculate exclusion limits for the 168 GeV mass point, the bottom-quark mass has been adjusted to its PDG value of 4.18 GeV [28]. Only after this change, $m_{H^{\pm}} + m_b$ is smaller than $m_t$ for the 168 GeV mass point.




(a) 2HDM Type-I

(b) 2HDM Type-II

(c) 2HDM Type-X

(d) 2HDM Type-Y

Figure 7.12: Exclusion limits at 95 % CL of $m_{H^{\pm}}$ - tan($\beta$) values for CP-conserving 2HDMs with natural flavour conservation as defined in Table 2.1. The exclusion limits are calculated using the 2HDMC program [61].

The exclusion limits are in general anti-correlated to the upper limits, i.e. the smaller the limits, the more parameter space can be excluded. This behaviour is nicely visible around the 80 GeV mass point. In addition, the branching ratio $\mathscr{B}(t \to H^{\pm}b)$ decreases with increasing $m_{H^{\pm}}$, as the available phase space reduces. This makes the exclusion limits weaker at high $m_{H^{\pm}}$.

The cyan and violet regions are excluded by previous direct searches at LEP and indirect flavour physics measurements ($b \to s\gamma$) from Belle and BaBar, respectively. Constraints from neutral Higgs searches for $H$ or $A$ are not included in the plots. Those constraints are heavily dependent on the discussed assumptions for $\sin(\beta - \alpha)$, $m_A$, etc., and do not lead to exclusions in general scenarios. The red and blue lines are the exclusion limits from the $H^{\pm} \to \tau\nu_{\tau}$ [93] and $H^{\pm} \to bc$ [91] searches, respectively. The exclusion limits from this search are better than the $H^{\pm} \to bc$ exclusion limits because of the much larger branching ratio $\mathscr{B}(H^{\pm} \to cs) \gg \mathscr{B}(H^{\pm} \to cb)$. However, more values of tan($\beta$) can be excluded by the $H^{\pm} \to \tau\nu_{\tau}$ search compared to this search. This is due to a larger branching ratio, $\mathscr{B}(H^{\pm} \to \tau\nu_{\tau})/\mathscr{B}(H^{\pm} \to cs) \approx 2.3$, and tighter upper limits. The upper limits set on $\mathscr{B}_{H^{\pm}}$ by the

$H^{\pm} \to \tau\nu_{\tau}$ search are tighter for multiple reasons. Tau leptons are less frequently produced than jets at the LHC, which reduces the overall background of e.g. single-top and $W$+jets backgrounds. Furthermore, $\tau$-jet identification techniques can be used to reduce backgrounds. The $H^{\pm} \to \tau\nu_{\tau}$ search focuses on hadronically decaying tau leptons [93]. The tau-jets are easier to separate from other jets than $c$-jets from light- and $b$-quark jets [163, 274]. This is especially important for the rejection of $t\bar{t}$ background events with a hadronically decaying $W$-boson. The last and probably most important reason is that the $W$-boson branching ratio $\mathscr{B}(W \to \tau\nu_{\tau})$ is almost a factor of 7 smaller than $\mathscr{B}(W \to q\bar{q})$, reducing the overall size of the $t\bar{t}$ background.

The exclusion limits for the 2HDM Type-II, Type-X and Type-Y models are shown in Figures 7.12(b) to 7.12(d). For these models the branching ratios of the charged Higgs boson are dependent on $\tan(\beta)$. In models II and Y, $\mathscr{B}(t \to H^{\pm}b)$ does not decrease monotonically. Hence, $\mathscr{B}(t \to H^{\pm}b) \times \mathscr{B}(H^{\pm} \to cs)$ vs. $\tan(\beta)$ is not a monotone function anymore, and it can occur that a combination of low and high $\tan(\beta)$ values or only some intermediate $\tan(\beta)$ values can be excluded. To indicate clearly which regions are excluded, they are filled or hatched. For clarity reasons the expected exclusion limits from this analysis are not drawn.

The exclusion limits from LEP and flavour physics measurements in the Type-X model are identical to the Type-I model and indicated by the violet dashed line. Everything below and left of the line is excluded. For the other two models, these exclusion limits are not drawn because the plotted parameter space is fully excluded by the $b \to s\gamma$ measurements [82].

In the 2HDM Type-II model, the $H^{\pm} \to cs$ search excludes values of $\tan(\beta) < 2$ for almost the whole analysed mass range. The $H^{\pm} \to \tau\nu_{\tau}$ search excludes high values of $\tan(\beta)$. The exclusion at low and high $\tan(\beta)$ of the two searches are in agreement with the expectations based on the $H^{\pm}$ branching ratios, discussed in Section 2.2.1. Both searches combined can fully exclude a 2HDM Type-II model with a low-mass charged Higgs boson between 80 and 160 GeV.

In the 2HDM Type-X model the evolution of the charged Higgs boson branching ratios with $\tan(\beta)$ is similar to the 2HDM Type-II model. For this reason the exclusion limits are alike. The main difference is that, as in 2HDM Type-I, both quark types couple to the same doublet, and the branching ratio $\mathscr{B}(t \to H^{\pm}b)$ reduces significantly with increasing $\tan(\beta)$. As a result, large $\tan(\beta)$ values cannot be excluded.

In the 2HDM Type-Y model, the $H^{\pm} \to cs$ channel is relevant for all $\tan(\beta)$ values. The branching ratio $\mathscr{B}(t \to H^{\pm}b)$ is similar to the Type-II model and is smallest for intermediate $\tan(\beta) \approx 10$ values. This search does exclude the full $\tan(\beta)$ space in the analysed search range, except for a few intermediate $\tan(\beta)$ values. The branching ratio $\mathscr{B}(H^{\pm} \to cb)$ is dominant at $\tan(\beta) > 2$ in the Type-Y model. This enables the $H^{\pm} \to cb$ search to exclude the intermediate $\tan(\beta)$ values not excluded by this search. The two searches combined fully exclude a 2HDM Type-Y model with a low-mass charged Higgs boson between 80 to 150 GeV.

In summary, in the discussed models the $H^{\pm} \to cs$ search only improves the exclusion limits in the 2HDM Type-I model upon the previous exclusion limits set by LEP and from flavour physics measurements. In the Type-I model, the $H^{\pm} \to \tau\nu_{\tau}$ channel is more sensitive. Considering only direct searches, the $H^{\pm} \to cs$ search can mainly exclude new regions thanks to its extended mass range.

It should be noted again that there are other 2HDM models with natural or minimal flavour conservation, models with new symmetries, etc., in which the $H^{\pm} \to cs$ channel can be most sensitive.

CHAPTER 8

# Summary and conclusion

This thesis presents the search for a light charged Higgs boson that decays to a $cs$ quark pair. A charged Higgs boson is predicted by most theories with an extended Higgs sector and the $cs$ channel is one of the dominant decay modes. The focus is set on semileptonic $t\bar{t}$ events where one top quark decays into a charged Higgs boson and a bottom quark, $t\bar{t} \to b\bar{b}H^{\pm}(\to cs)W^{\mp}(\to \ell\nu)$. The search analyses $pp$ collisions data at $\sqrt{s} = 13\,\text{TeV}$ collected with the ATLAS detector between 2015 and 2018. An event selection is applied to filter semileptonic $t\bar{t}$ events and the $t\bar{t}$ system is reconstructed. Signal events in data are identified with the help of a boosted decision tree that utilises top-quark kinematic, flavour-tagging, and spin-related quantities.

No significant signal excess is observed for any of the twelve signal mass hypotheses between 60 to 168 GeV. In that mass range, observed (expected) upper limits at 95 % CL on the branching ratio $\mathcal{B}(t \to H^{\pm}b)$, assuming $\mathcal{B}(H^{\pm} \to cs) = 100\,\%$, are set between 0.066 % (0.077 %) and 3.6 % (2.3 %). These are the first direct limits on $\mathcal{B}(t \to H^{\pm}b)$ in the $cs$ channel in the charged Higgs boson mass range from 60 to 80 GeV and 160 to 168 GeV. The limits from this analysis supersede limits from previous searches in the mass range from 120 to 160 GeV.

The results are interpreted in terms of popular CP-conserving two-Higgs-doublet models with natural flavour conservation. New regions in the $m_{H^{\pm}}$ - $\tan(\beta)$ plane, which are not already excluded by previous direct or indirect searches, cannot be excluded.

The sensitivity of this search to a $H^{\pm} \to cb$ signal is estimated. The analysis sensitivity is not sufficient to assess the validity of the excess at 130 GeV found by the dedicated $H^{\pm} \to cb$ search. The results indicate that when retraining the BDT with the $H^{\pm} \to cb$ signal samples, the sensitivity of this analysis supersedes the dedicated analysis.

In this thesis, several options are given for how the search for a $H^{\pm} \to cs$ signal could be further improved in the future. Among them is the usage of attention based machine learning tools to label the jets in an event. It has been shown that the search is limited by systematic uncertainties, and the ones related to the modelling of the $t\bar{t}$ background are dominant for all charged Higgs boson mass points except the one at 80 GeV. An improved understanding of the $t\bar{t}$ production process will significantly enhance the sensitivity to a charged Higgs boson.

Furthermore, a combination of the $cs$ and $cb$ channels in the search for a charged Higgs boson is sensible. A caveat in the presented $H^{\pm} \to cs$ search is that it does not account for a $H^{\pm} \to cb$ signal. For certain models and parameter spaces the decays of the charged Higgs boson into a $cs$ and $cb$ quark

pair are both relevant. An example is the 2HDM Type-Y model with intermediate and large $\tan(\beta)$ values. For this reason the limits are only quoted assuming $\mathscr{B}(H^{\pm} \to cs) = 100\,\%$. In order to calculate limits on $\mathscr{B}(t \to H^{\pm}b) \times \mathscr{B}(H^{\pm} \to cs)$, the $H^{\pm} \to cb$ signal has to be separated from $H^{\pm} \to cs$. The only difference between the two decay modes is the existence of a $b$-jet instead of a light-jet in the final state. A simple selection based on the number of $b$-jets is not sufficient, as about two thirds of the $H^{\pm} \to cb$ events in the signal region have only two $b$-tagged jets. A better separation can be achieved by reconstructing the $t\bar{t}$ system and testing the flavour-tagging scores of the two jets stemming from the charged Higgs boson. This already comes close to a search for $H^{\pm} \to cb$ with many similarities to the $H^{\pm} \to cs$ signal.

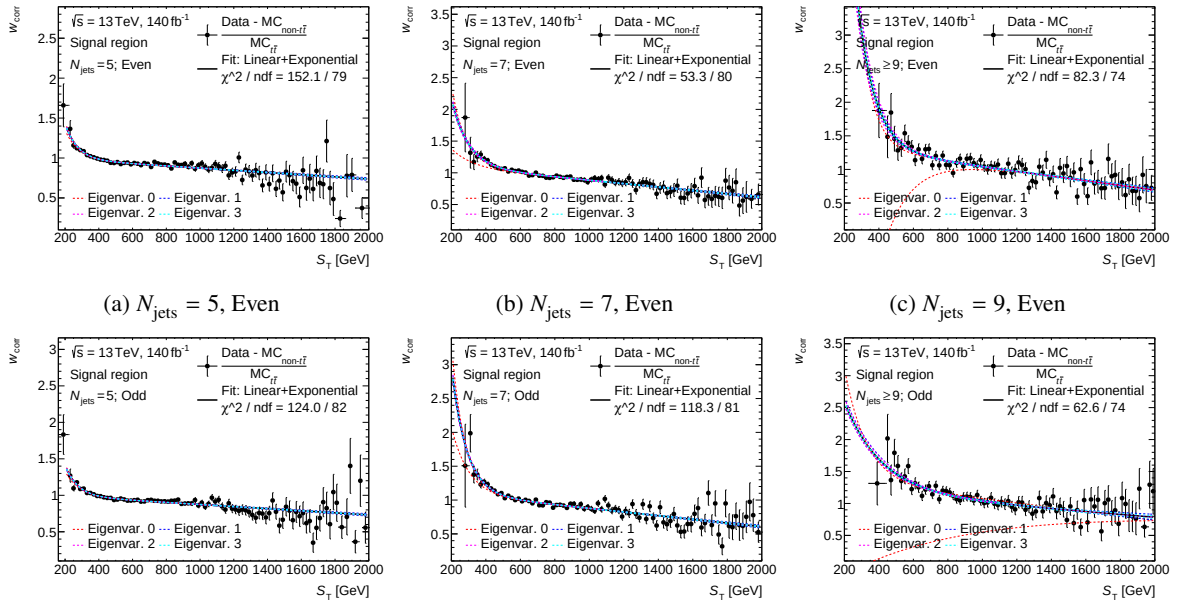
In a combined search, a similar analysis strategy to what has been presented in this thesis can be applied. Just instead of a binary classifier, a multi-class classifier has to be trained with a background and two signal classes. Even the same input features as in this analysis can be utilised because the kinematics of both signals are very similar, and differences in the flavour-tagging composition are considered by the flavour-tagging quantities. In a simultaneous fit of the MVA output score distributions, both signal contributions (branching ratios) can be extracted coherently. Knowing both signal strengths facilitates the calculation of exclusion limits with varied contributions of the two decay channels. This limit calculation with varied contributions is commonly done in searches for new particles with unknown decay properties, e.g. vector-like quarks [275].

This thesis highlights why the search for a charged Higgs boson is of interest and presents a new analysis that yields improved exclusion limits in the $cs$ channel. With an increasing integrated luminosity of the LHC dataset or with the help of new experiments with enhanced centre-of-mass energies, we may obtain certainty about the existence of a charged Higgs boson. For this reason, the search for a charged Higgs boson will remain an important research topic in the future. The search presented in this thesis can guide future analysis to refine their analysis strategy. And thanks to the improved upper limits on the production cross-section of a charged Higgs boson decaying into $cs$ quarks, new parameter spaces of models predicting a charged Higgs boson can be constrained. This can guide theoreticians to develop and refine models with a charged Higgs boson.

APPENDIX A

# Auxiliary material

## A.1 $t\bar{t}$ correction plots



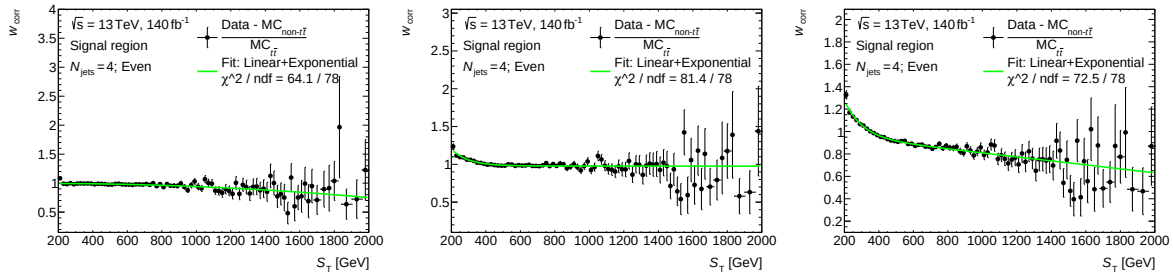

(a) $N_{\text{jets}} = 5$, Even (b) $N_{\text{jets}} = 7$, Even (c) $N_{\text{jets}} = 9$, Even

(d) $N_{\text{jets}} = 5$, Odd (e) $N_{\text{jets}} = 7$, Odd (f) $N_{\text{jets}} = 9$, Odd

Figure A.1: Distribution of $w_{\text{corr}}$ in different $N_{\text{jets}}$bins for both parities. Also shown, as a solid black line, are fits of a linear-plus-exponential function to the distribution. The dotted lines show the $\pm 1\sigma$ variations of the eigenvectors of the uncertainty matrix. The uncertainties include only data and MC statistical uncertainties.

Table A.1: Fit parameters and their uncertainties of the linear-plus-exponential function ($f_{\text{corr}}(x, S_{\text{T}}) = x_0 + x_1 S_{\text{T}} + x_2 e^{x_3 S_{\text{T}}}$) fit to the $t\bar{t}$ correction weight templates. The uncertainties on the parameters may be large compared to the fitted value and total uncertainty, as the fit parameters are highly correlated.

| $N_{\text{jets}}$ | Parity | $x_0$ | $x_1$ | $x_2$ | $x_3$ |
|---|---|---|---|---|---|
| 4 | Even | $0.97 \pm 1.67 \times 10^{-4}$ | $-1.07 \times 10^{-4} \pm 3.56 \times 10^{-10}$ | $3.03 \pm 0.372$ | $-1.12 \times 10^{-2} \pm 7.89 \times 10^{-7}$ |
| | Odd | $1.01 \pm 1.04 \times 10^{-4}$ | $-1.66 \times 10^{-4} \pm 2.60 \times 10^{-10}$ | $5.45 \pm 2.23$ | $-1.42 \times 10^{-2} \pm 1.34 \times 10^{-6}$ |
| 5 | Even | $1.01 \pm 9.92 \times 10^{-5}$ | $-1.37 \times 10^{-4} \pm 2.00 \times 10^{-10}$ | $6.66 \pm 5.56$ | $-1.35 \times 10^{-2} \pm 1.61 \times 10^{-6}$ |
| | Odd | $1.02 \pm 1.03 \times 10^{-4}$ | $-1.43 \times 10^{-4} \pm 2.06 \times 10^{-10}$ | $5.89 \pm 4.48$ | $-1.32 \times 10^{-2} \pm 1.66 \times 10^{-6}$ |
| 6 | Even | $1.06 \pm 1.93 \times 10^{-4}$ | $-1.96 \times 10^{-4} \pm 2.80 \times 10^{-10}$ | $7.78 \pm 9.42$ | $-1.15 \times 10^{-2} \pm 1.47 \times 10^{-6}$ |
| | Odd | $1.09 \pm 1.26 \times 10^{-4}$ | $-2.36 \times 10^{-4} \pm 2.11 \times 10^{-10}$ | $19.2 \pm 103$ | $-1.48 \times 10^{-2} \pm 2.65 \times 10^{-6}$ |
| 7 | Even | $1.15 \pm 5.74 \times 10^{-4}$ | $-2.68 \times 10^{-4} \pm 5.97 \times 10^{-10}$ | $8.39 \pm 47.5$ | $-1.03 \times 10^{-2} \pm 4.80 \times 10^{-6}$ |
| | Odd | $1.15 \pm 4.52 \times 10^{-4}$ | $-2.71 \times 10^{-4} \pm 5.07 \times 10^{-10}$ | $19.3 \pm 141$ | $-1.16 \times 10^{-2} \pm 2.69 \times 10^{-6}$ |
| 8 | Even | $1.20 \pm 2.28 \times 10^{-3}$ | $-2.74 \times 10^{-4} \pm 1.72 \times 10^{-9}$ | $6.41 \pm 25.9$ | $-7.53 \times 10^{-3} \pm 3.53 \times 10^{-6}$ |
| | Odd | $1.16 \pm 4.17 \times 10^{-3}$ | $-2.51 \times 10^{-4} \pm 2.63 \times 10^{-9}$ | $3.18 \pm 4.35$ | $-5.79 \times 10^{-3} \pm 2.71 \times 10^{-6}$ |
| $\geq 9$ | Even | $1.39 \pm 3.55 \times 10^{-3}$ | $-3.47 \times 10^{-4} \pm 2.40 \times 10^{-9}$ | $32.9 \pm 2\,940$ | $-9.76 \times 10^{-3} \pm 1.15 \times 10^{-5}$ |
| | Odd | $1.22 \pm 4.57 \times 10^{-2}$ | $-2.19 \times 10^{-4} \pm 1.92 \times 10^{-8}$ | $3.55 \pm 25.3$ | $-4.62 \times 10^{-3} \pm 1.10 \times 10^{-5}$ |




(a) MadGraph5_aMC@NLO (b) Herwig (c) $h_{\text{damp}} = 3.0$

Figure A.2: Distribution of $w_{\text{corr}}$ for the $N_{\text{jets}}$equal 4 bin for the even parity and the corresponding fit of a linear-plus-exponential function is shown as a solid green line. The uncertainties include data and MC statistical uncertainties.

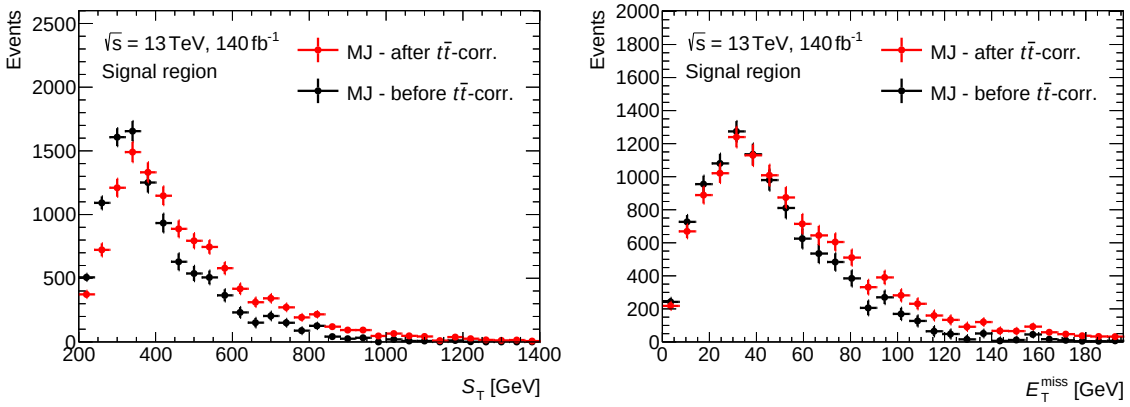



Figure A.3: Comparison of the MJ distribution before and after the $t\bar{t}$-correction has been applied. The uncertainties include data and MC statistical uncertainties.

## A.2 BDT input feature distributions

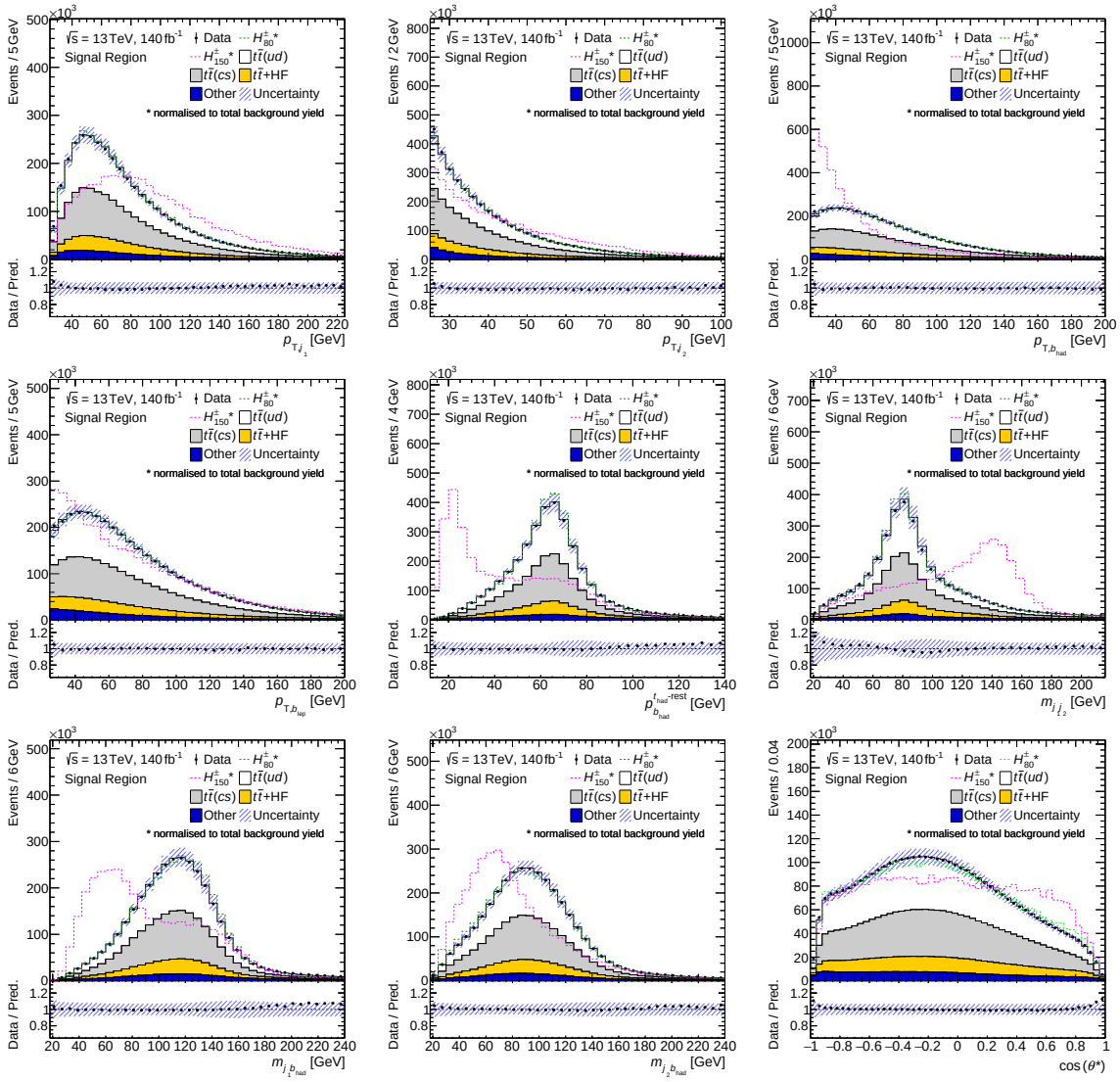



Figure A.4: BDT input features: Hadronic-top kinematics.

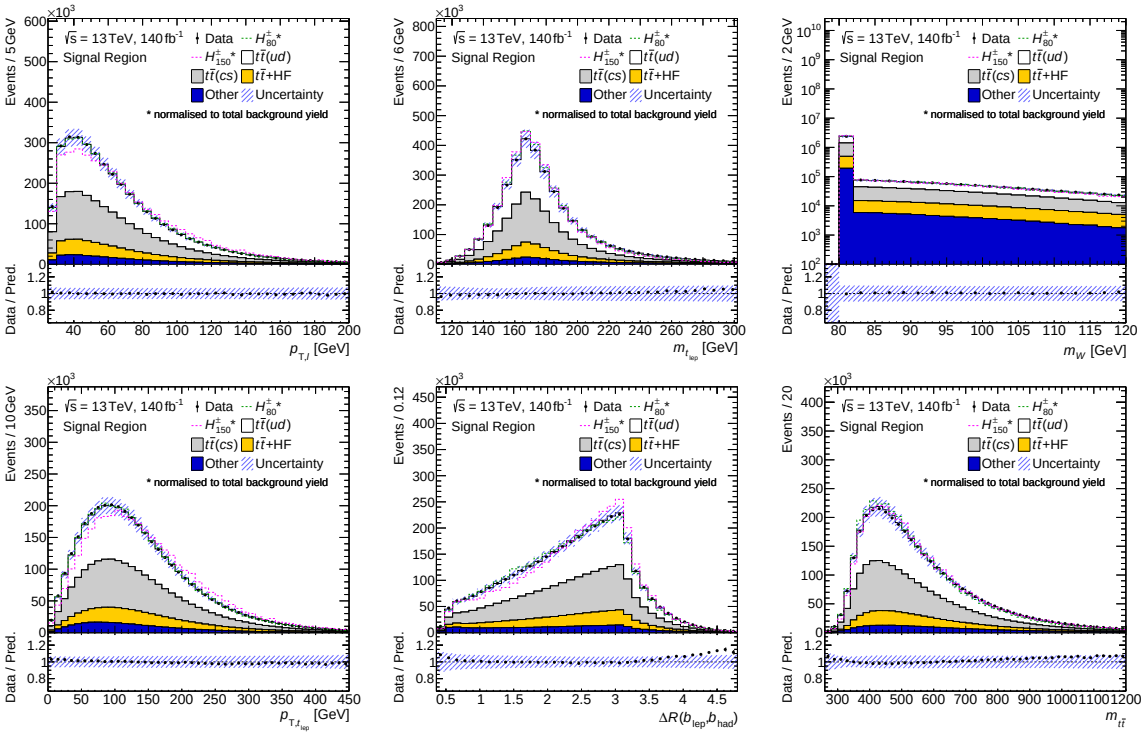



Figure A.4: BDT input features: Leptonic-top and $t\bar{t}$-system kinematics.

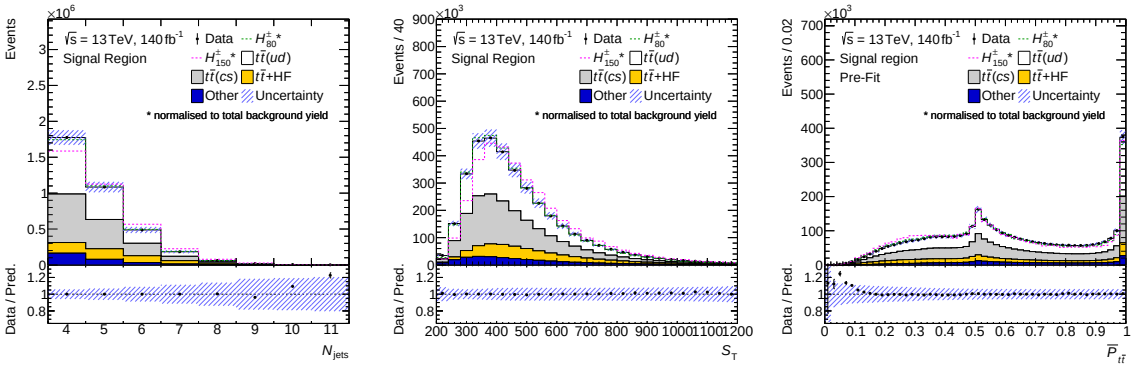



Figure A.4: BDT input features: Event-level properties.

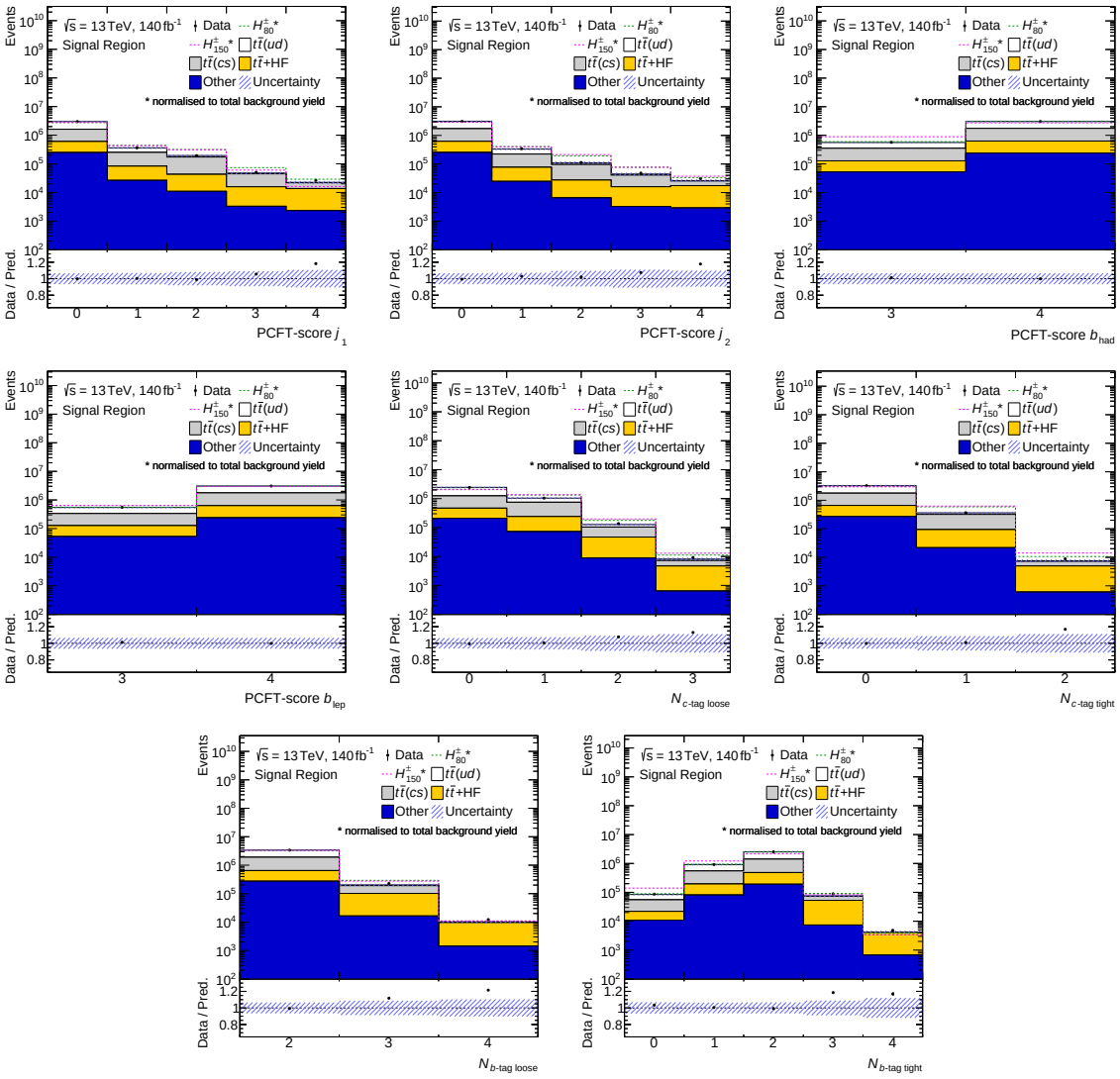



Figure A.4: BDT input features: Flavour tagging properties.

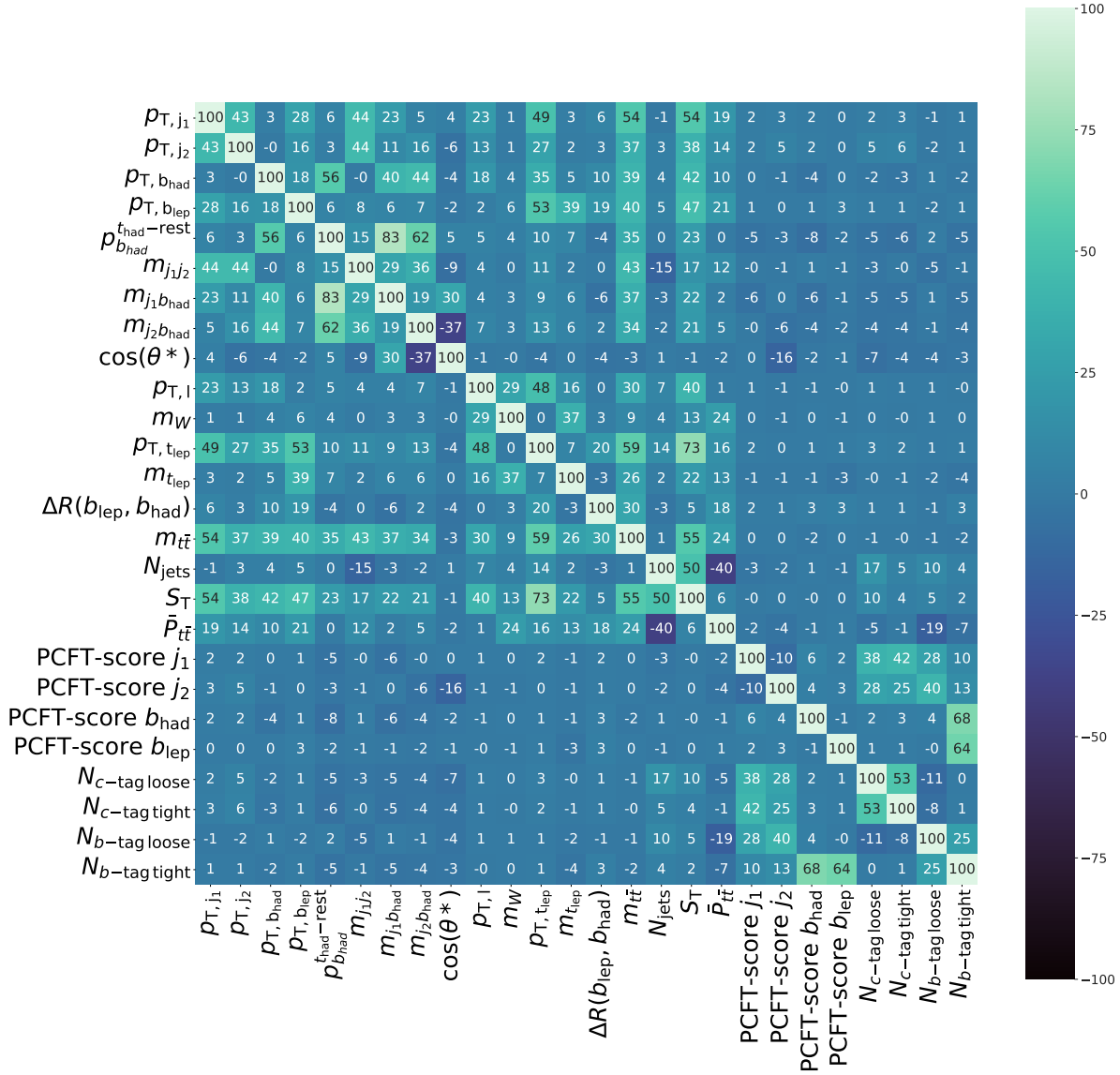



Figure A.5: Linear correlation coefficients of the BDT input features in percent for the $H^{\pm}_{130}$ signal sample.

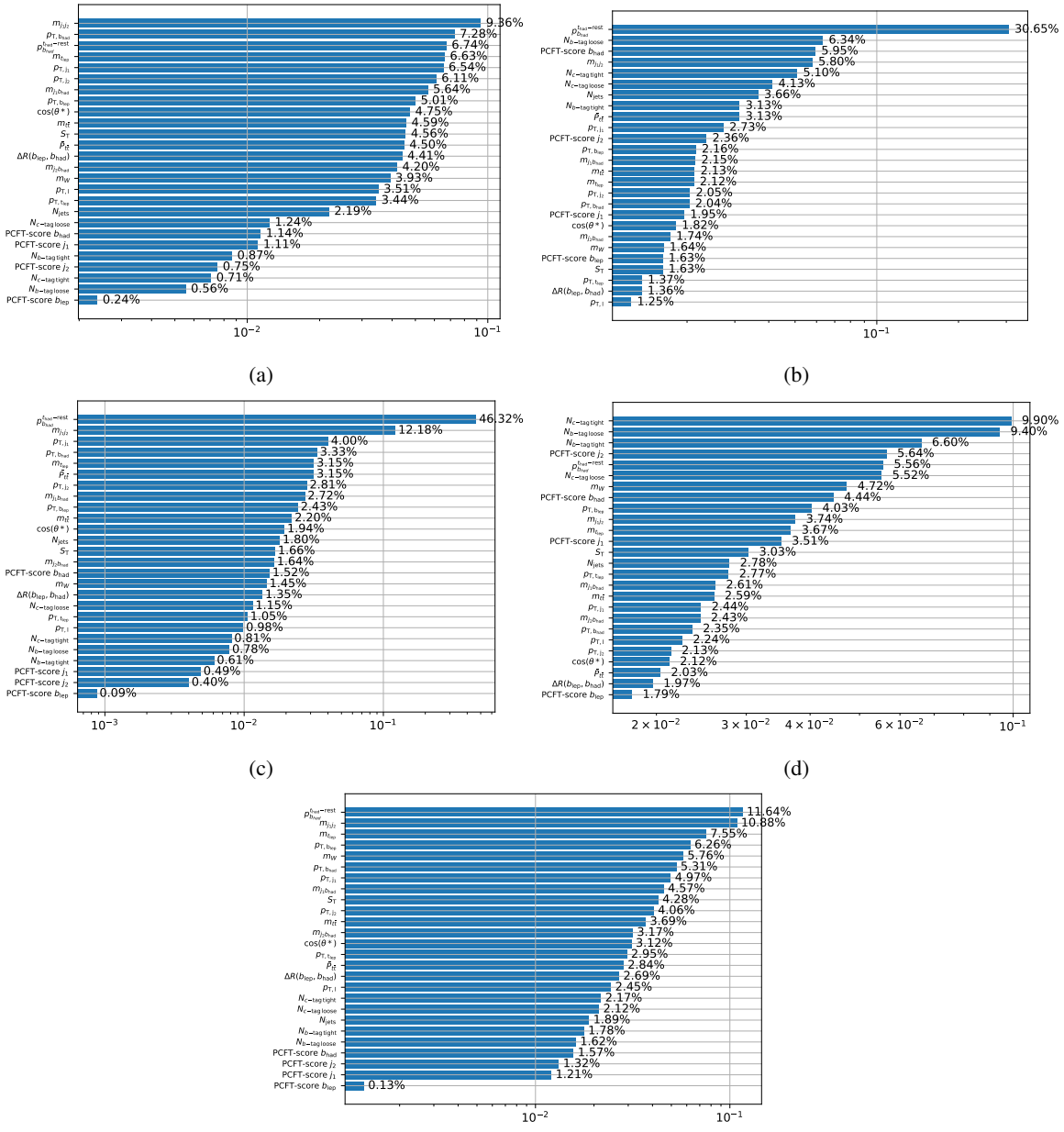



Figure A.6: Feature importance in the final BDT training with the $H^{\pm}_{130}$ signal sample. The overall feature importance is normalised to 1. (a) Importance based on the number of times a feature has been used to split a node. Importance based on the (b) average and (c) total gain in loss by node splittings based on the given feature. Importance based on the (d) average and (e) total coverage (definition see Section 5.2.1).

## A.3 BDT score distributions

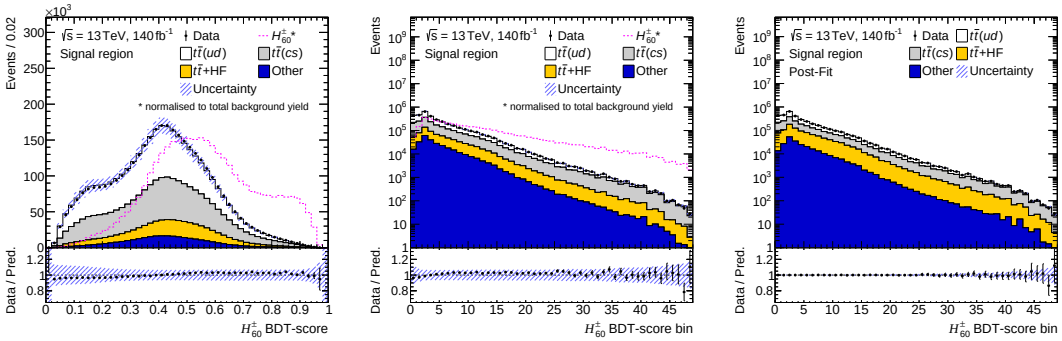



Figure A.7: BDT score from the training with the 60 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

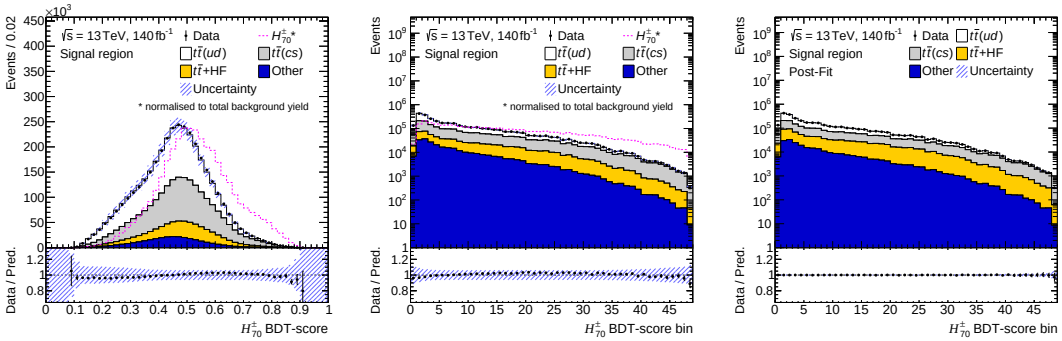



Figure A.8: BDT score from the training with the 70 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

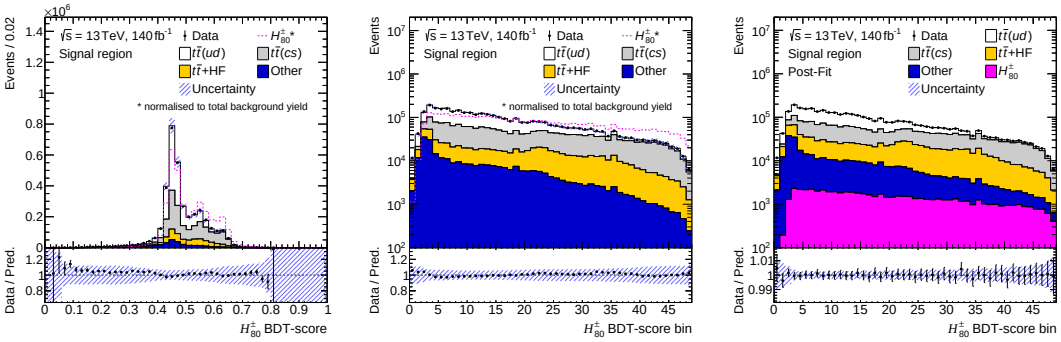



Figure A.9: BDT score from the training with the 80 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

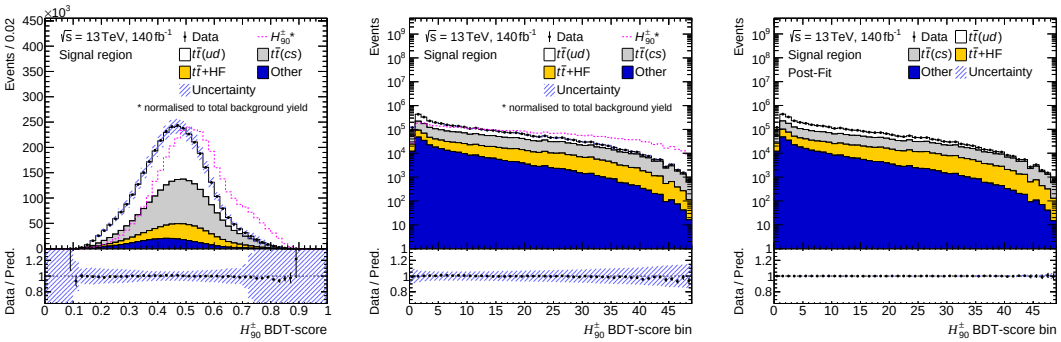



Figure A.10: BDT score from the training with the 90 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

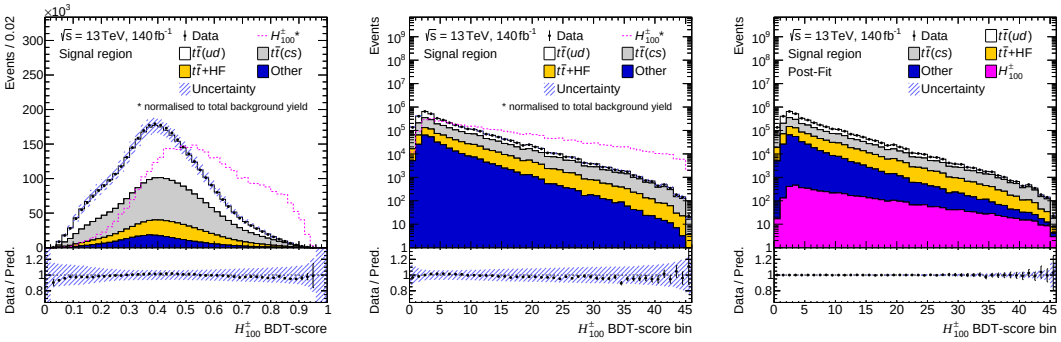



Figure A.11: BDT score from the training with the 100 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

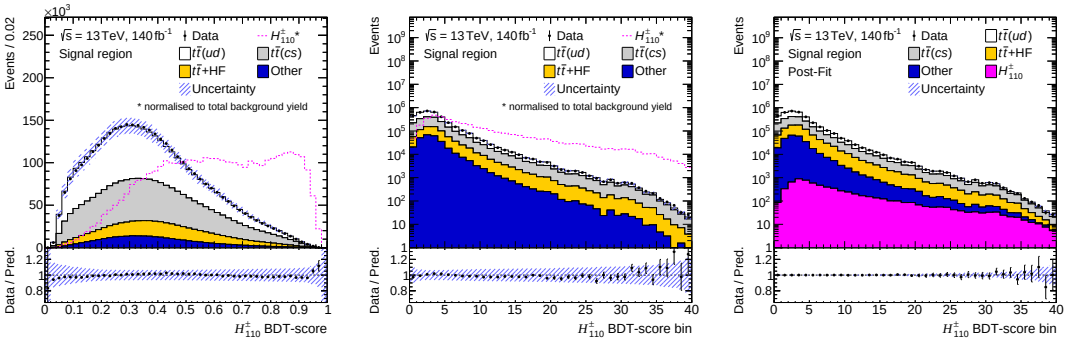



Figure A.12: BDT score from the training with the 110 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

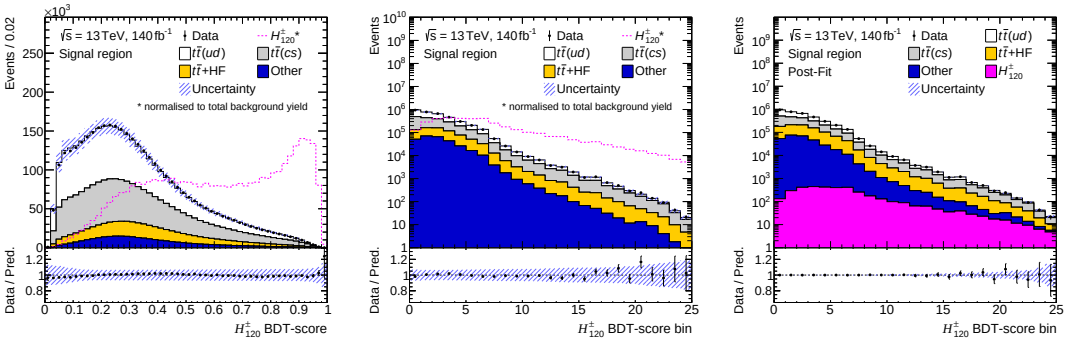



Figure A.13: BDT score from the training with the 120 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

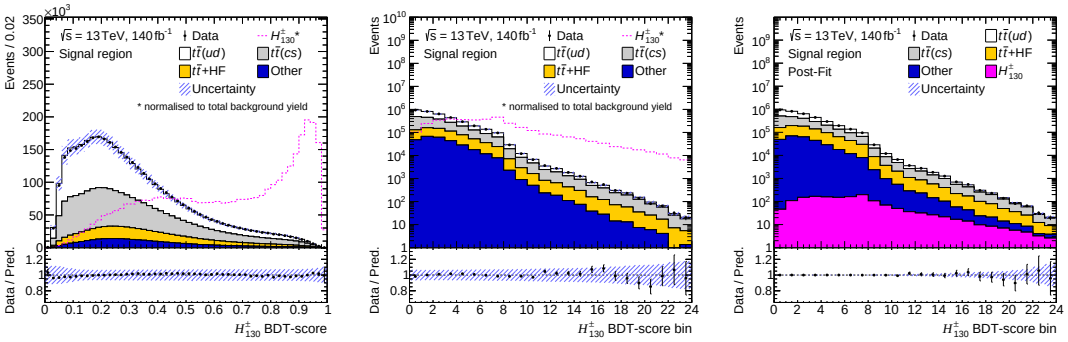



Figure A.14: BDT score from the training with the 130 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

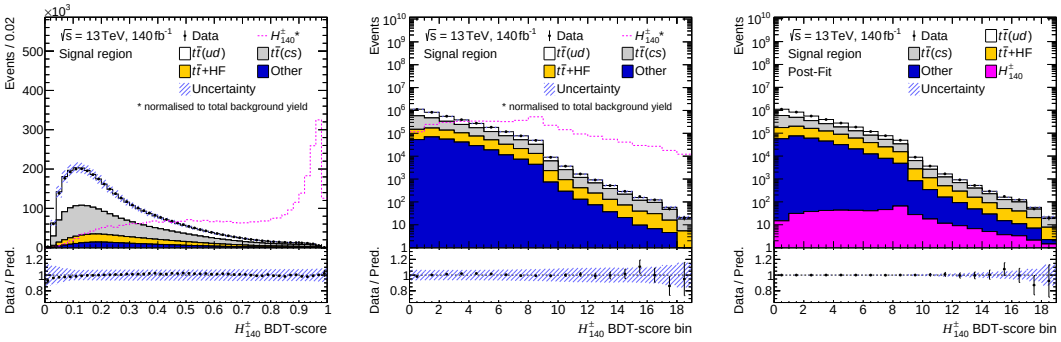



Figure A.15: BDT score from the training with the 140 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

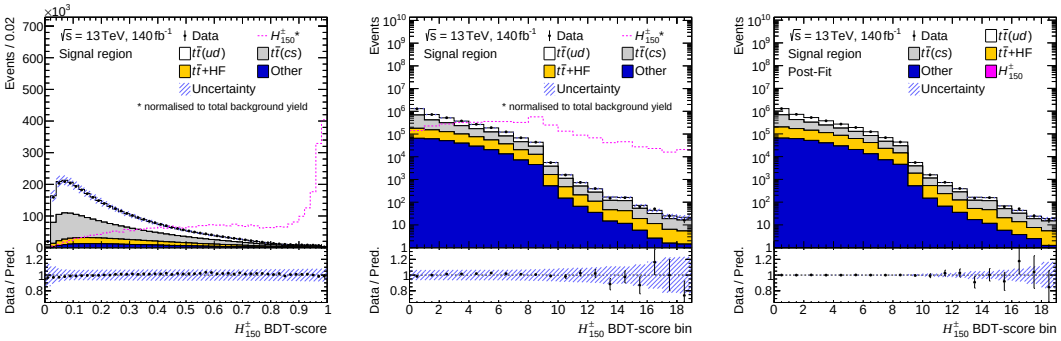



Figure A.16: BDT score from the training with the 150 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

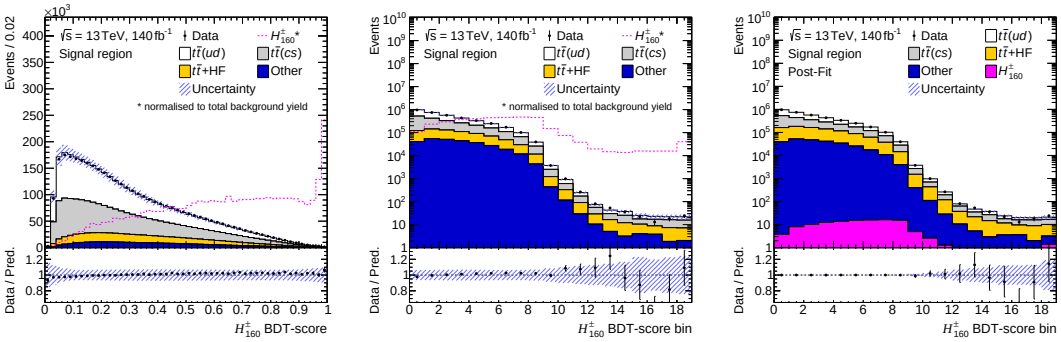



Figure A.17: BDT score from the training with the 160 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

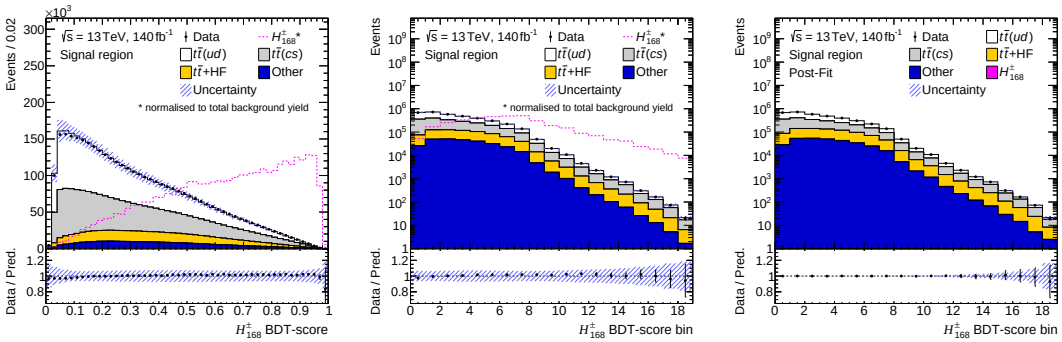



Figure A.18: BDT score from the training with the 168 GeV mass point: (a) Initial, (b) rebinned and (c) post-fit to data.

## A.4 Post-fit BDT input feature distributions

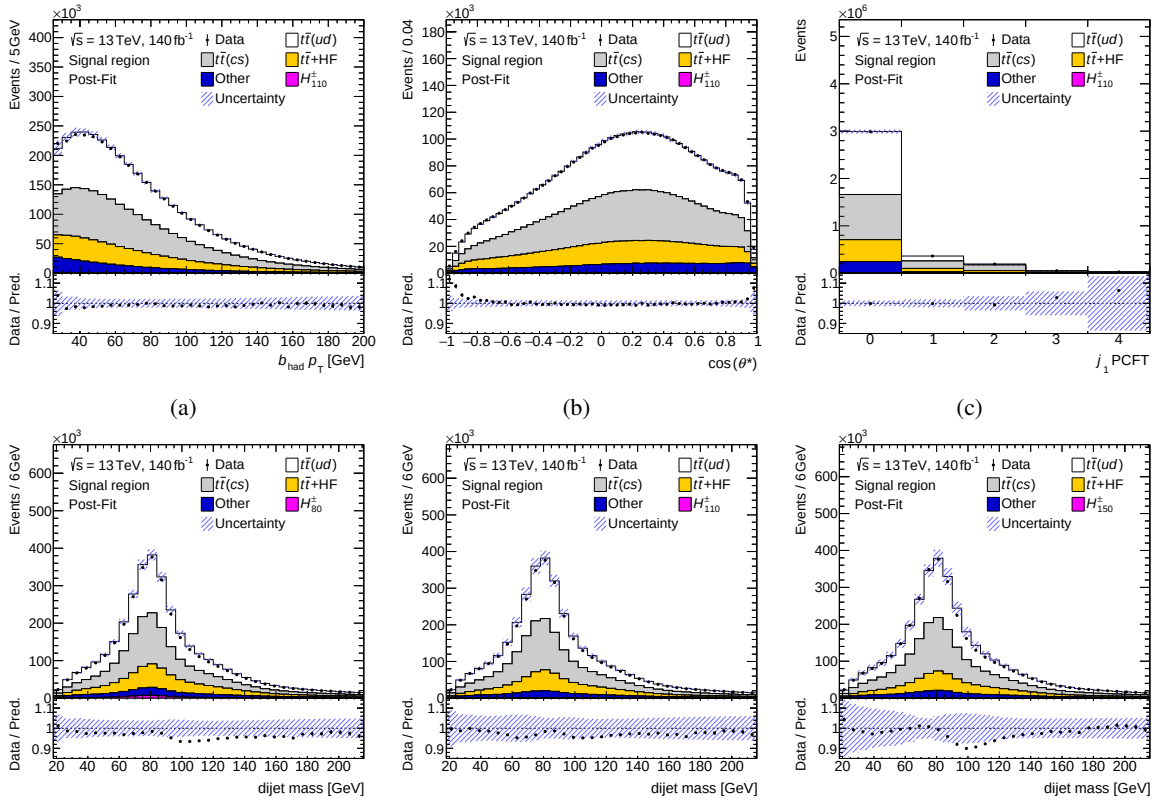



Figure A.19: Post-fit BDT input feature distributions. (a) $b_{\text{had}}$-jet $p_{\text{T}}$, (b) cos ($\theta^*$), (c) PCFT score $j_1$, and (d-f) $m_{j_1 j_2}$. The NPs are set to values extracted from fits to BDT scores. Which BDT score has been fit is indicated by the signal sample in the legend.

## A.5 Systematic uncertainties

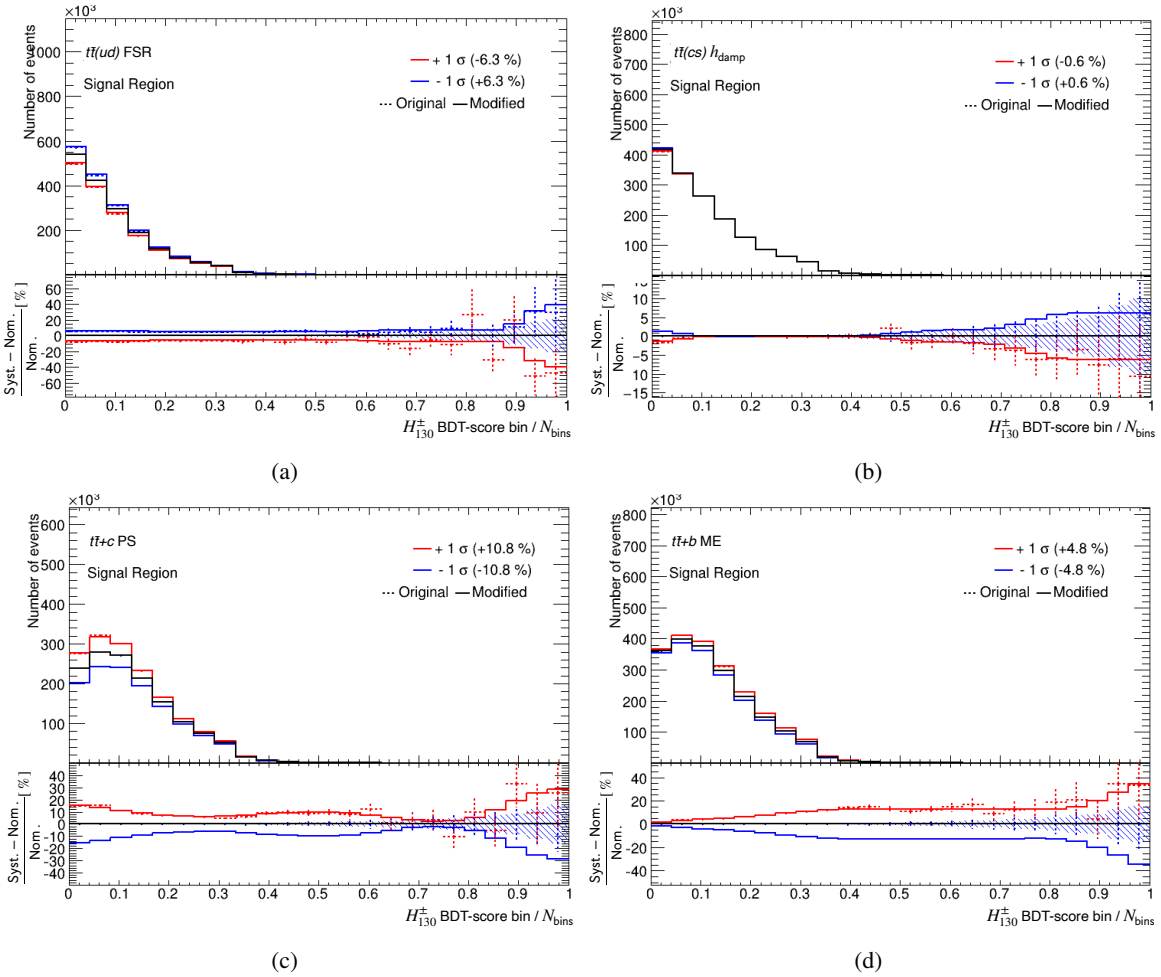
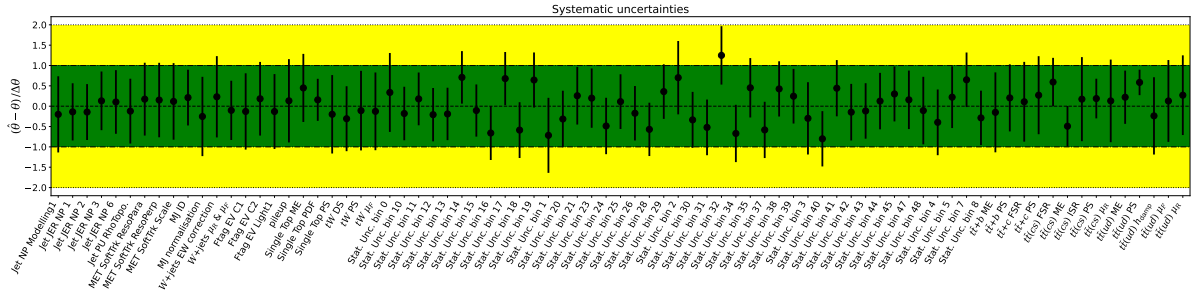
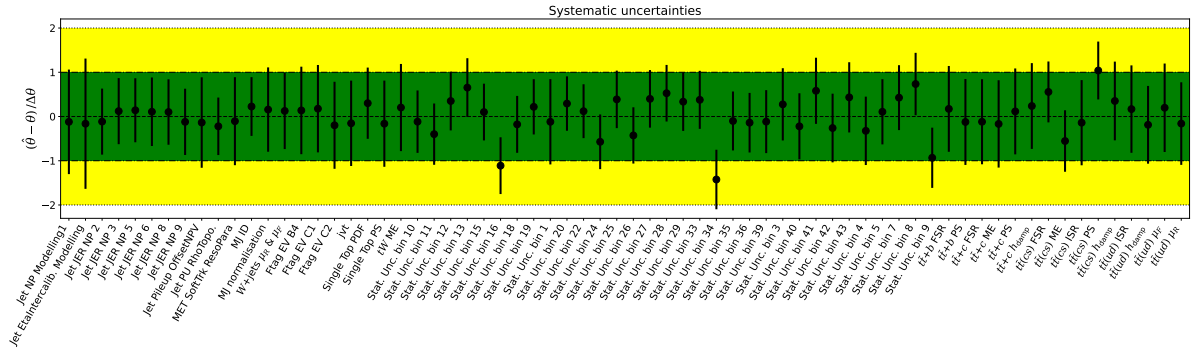


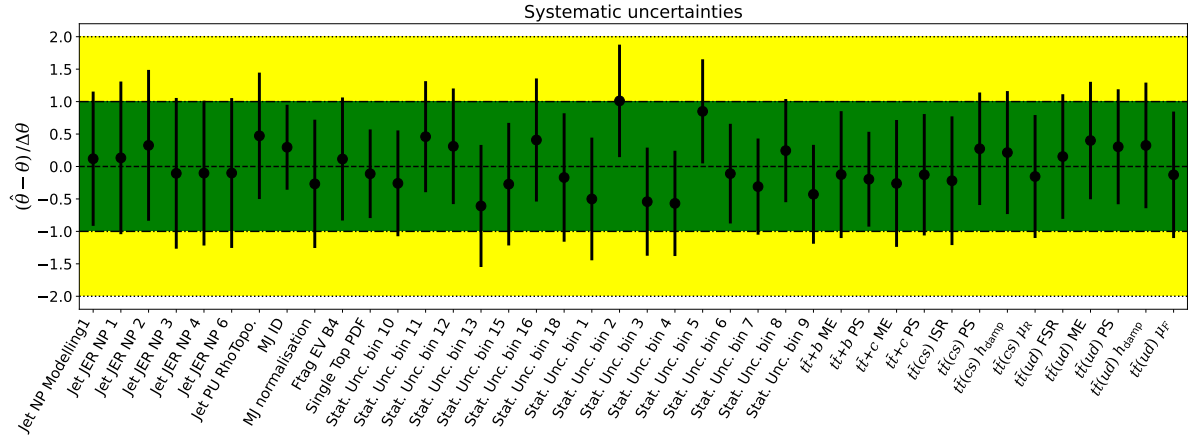

Figure A.20: Systematic uncertainty templates before (original) and after (modified) symmetrisation and smoothing for the $H^{\pm}_{130}$ BDT score. The ratio plot visualises the relative difference with respect to the nominal template. (a) FSR uncertainty of the $t\bar{t}(ud)$ sample; (b) $h_{\text{damp}}$ uncertainty of the $t\bar{t}(cs)$ sample; (c) Alternative parton shower generator uncertainty of the $t\bar{t}+c$ sample; (d) Alternative matrix element generator uncertainty of the $t\bar{t}+b$ sample.




(a)




(b)




(c)

Figure A.21: Pull plot of constrained parameters in the fit with the (a) $H^{\pm}_{80}$, (b) $H^{\pm}_{100}$ and (c) $H^{\pm}_{150}$ signal samples. $\theta$ and $\Delta\theta$ are the initial value and its uncertainty, $\hat{\theta}$ is the value estimated from the fit. Only parameters that are pulled by more than 10 % of their uncertainty are listed. The bin simulated statistical uncertainty, "Stat. Unc. bin #", starts counting at zero for bin 1.

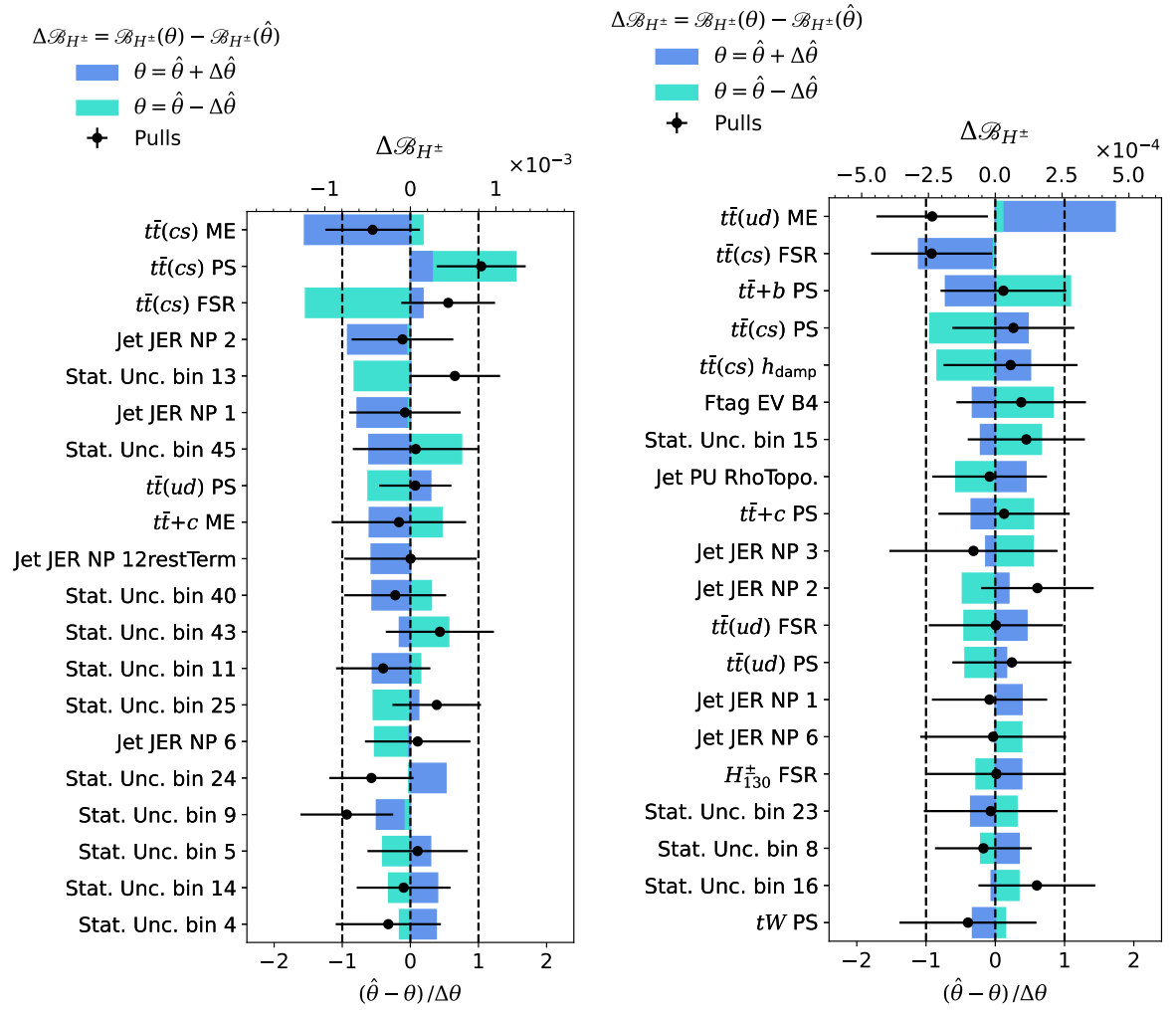



Figure A.22: Ranking of the NPs according to their impact on the measured $\mathscr{B}_{H^\pm}$ values for fits to the (a) $H^\pm_{100}$ and (b) $H^\pm_{130}$ BDT score distributions. The impact is determined by fixing the NP under consideration to its nominal post-fit $\pm 1\sigma$ uncertainty value. $\theta = \hat{\theta} \pm \Delta\hat{\theta}$ is the value to which the NP under consideration is fixed in the fit, where $\hat{\theta}$ is the best-fit value of the NP from the nominal fit and $\Delta\hat{\theta}$ is its $1\sigma$ uncertainty. $\Delta\mathscr{B}_{H^\pm}$ is the difference of the fitted $\mathscr{B}_{H^\pm}$ with respect to the nominal $\mathscr{B}_{H^\pm}$ value ($\hat{\mathscr{B}}_{H^\pm}$). The black points with a horizontal line represent the NP pulls relative to their nominal values ($\theta_0$) and uncertainties ($\Delta\theta_0$).

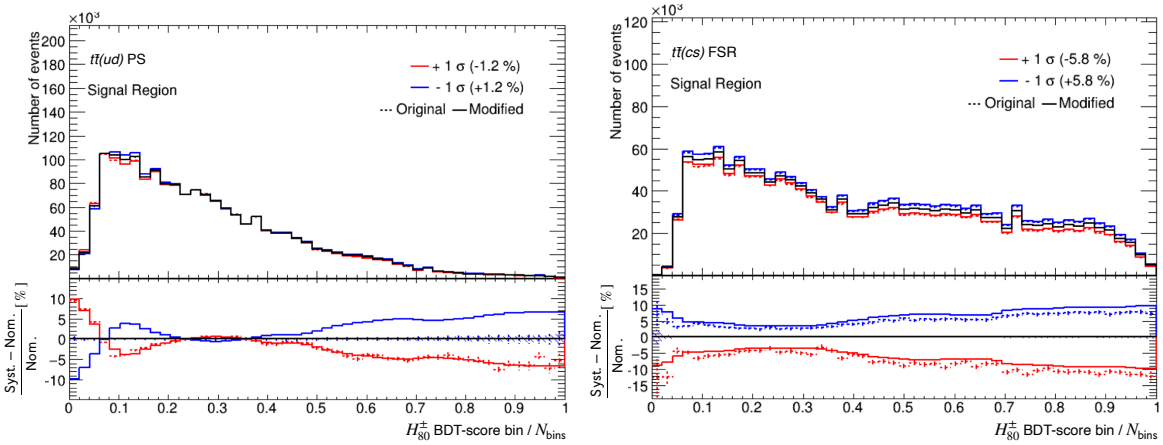



Figure A.23: Systematic uncertainty templates of the (a) $t\bar{t}(ud)$ PS and (b) $t\bar{t}(cs)$ FSR uncertainty for the $H_{80}^{\pm}$ BDT score.

# Bibliography


[1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on p. 1).

[2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex] (cit. on p. 1).

[3] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16 1964) 508, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (cit. on p. 1).

[4] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9 1964) 321, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321 (cit. on p. 1).

[5] J. F. Gunion and H. E. Haber, *Conditions for CP-Violation in the General Two-Higgs-Doublet Model*, Phys. Rev. D **72** (2005) 095002, arXiv: hep-ph/0506227 (cit. on p. 1).

[6] M. Joyce, T. Prokopec and N. Turok, *Non-local Electroweak Baryogenesis. Part II: The Classical Regime*, Physical Review D **53** (1996) 2958, ISSN: 1089-4918, URL: http://dx.doi.org/10.1103/PhysRevD.53.2958 (cit. on p. 1).

[7] J. M. Cline, K. Kainulainen and A. P. Vischer, *Dynamics of Two Higgs Doublet CP Violation and Baryogenesis at the Electroweak Phase Transition*, Physical Review D **54** (1996) 2451, ISSN: 1089-4918, URL: http://dx.doi.org/10.1103/PhysRevD.54.2451 (cit. on p. 1).

[8] L. Fromme, S. J. Huber and M. Seniuch, *Baryogenesis in the Two-Higgs Doublet Model*, Journal of High Energy Physics **2006** (2006) 038, ISSN: 1029-8479, URL: http://dx.doi.org/10.1088/1126-6708/2006/11/038 (cit. on p. 1).

[9] M. Trodden, "Electroweak Baryogenesis: A Brief Review", *33rd Rencontres de Moriond: Electroweak Interactions and Unified Theories*, 1998 471, arXiv: hep-ph/9805252 (cit. on p. 1).

[10] A. Riotto and M. Trodden, *Recent progress in baryogenesis*, Ann. Rev. Nucl. Part. Sci. **49** (1999) 35, arXiv: hep-ph/9901362 (cit. on p. 1).

[11] A. G. Akeroyd et al., *Prospects for charged Higgs searches at the LHC*, Eur. Phys. J. C **77** (2017) 276, arXiv: 1607.01320 [hep-ph] (cit. on pp. 1, 2, 17, 20, 21, 26, 27).

[12] G. Arcadi, A. Djouadi and M. Raidal, *Dark Matter through the Higgs portal*, Phys. Rept. **842** (2020) 1, arXiv: 1903.03616 [hep-ph] (cit. on pp. 1, 17).

[13] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Pseudoparticles*, Phys. Rev. Lett. **38** (1977) 1440 (cit. on pp. 1, 15).

[14] J. E. Kim, *Light Pseudoscalars, Particle Physics and Cosmology*, Phys. Rept. **150** (1987) 1 (cit. on p. 1).

[15] H. Haber and G. Kane, *The search for supersymmetry: Probing physics beyond the standard model*, Physics Reports **117** (1985) 75, ISSN: 0370-1573, URL: https://www.sciencedirect.com/science/article/pii/0370157385900511 (cit. on p. 1).

[16] S. P. Martin, *A Supersymmetry Primer*, Adv. Ser. Direct. High Energy Phys. **18** (1998) 1, arXiv: hep-ph/9709356 (cit. on p. 1).

[17] T. P. Cheng and L.-F. Li, *Neutrino masses, mixings, and oscillations in SU(2)×U(1) models of electroweak interactions*, Phys. Rev. D **22** (1980) 2860 (cit. on p. 1).

[18] J. Schechter and J. W. F. Valle, *Neutrino masses in SU(2) $\bigotimes$ U(1) theories*, Phys. Rev. D **22** (1980) 2227 (cit. on p. 1).

[19] G. Lazarides, Q. Shafi and C. Wetterich, *Proton lifetime and fermion masses in an SO(10) model*, Nucl. Phys. B **181** (1981) 287 (cit. on p. 1).

[20] M. S. Chanowitz and M. Golden, *Higgs boson triplets with $M_W = M_Z cos\,\theta_W$*, Phys. Lett. B **165** (1985) 105 (cit. on pp. 1, 24).

[21] J. F. Gunion, R. Vega and J. Wudka, *Higgs triplets in the standard model*, Phys. Rev. D **42** (1990) 1673 (cit. on pp. 1, 24, 25).

[22] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516** (2012) 1, arXiv: 1106.0034 [hep-ph] (cit. on pp. 2, 15–17, 19, 22).

[23] M. Thomson, *Modern particle physics*, New York: Cambridge University Press, 2013, ISBN: 978-1-107-03426-6 (cit. on pp. 3, 8, 9, 12, 23).

[24] F. Halzen and A. D. Martin, *QUARKS AND LEPTONS: An Introductory Course in Modern Particle Physics*, 1984, ISBN: 978-0-471-88741-6 (cit. on p. 3).

[25] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Reading, USA: Addison-Wesley, 1995, ISBN: 978-0-201-50397-5, 978-0-429-50355-9, 978-0-429-49417-8 (cit. on p. 3).

[26] H. Schopper, ed., *Particle Physics Reference Library. Volume 1: Theory and Experiments*, Springer, 2020, ISBN: 978-3-030-38206-3, 978-3-030-38209-4, 978-3-030-38207-0 (cit. on pp. 3, 23).

[27] D. Galbraith and C. Burgard, *Example: Standard model of physics*, URL: https://texample.net/tikz/examples/model-physics/ (visited on 17/09/2024) (cit. on p. 4).

[28] Particle Data Group, P. Zyla et al., *Review of Particle Physics*, PTEP **2020** (2020) 083C01 (cit. on pp. 4, 5, 10, 11, 13, 15, 26, 63, 65–67, 83, 151).

[29] Y. Fukuda et al., *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv: hep-ex/9807003 (cit. on p. 5).

[30] Q. R. Ahmad et al., *Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301, arXiv: nucl-ex/0204008 (cit. on p. 5).

[31] N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, Astron. Astrophys. **641** (2020) A6, [Erratum: Astron.Astrophys. 652, C4 (2021)], arXiv: 1807.06209 [astro-ph.CO] (cit. on pp. 5, 14).

[32] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, Phys. Rev. **96** (1954) 191, ed. by J.-P. Hsu and D. Fine (cit. on p. 7).

[33] V. C. Rubin and W. K. Ford Jr., *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, Astrophys. J. **159** (1970) 379 (cit. on p. 14).

[34] D. Clowe et al., *A direct empirical proof of the existence of dark matter*, Astrophys. J. Lett. **648** (2006) L109, arXiv: astro-ph/0608407 (cit. on p. 14).

[35] M. Davis, G. Efstathiou, C. S. Frenk and S. D. M. White, *The Evolution of Large Scale Structure in a Universe Dominated by Cold Dark Matter*, Astrophys. J. **292** (1985) 371, ed. by M. A. Srednicki (cit. on p. 14).

[36] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32 (cit. on p. 14).

[37] K. G. Wilson, *The Renormalization Group and Strong Interactions*, Phys. Rev. D **3** (1971) 1818 (cit. on p. 15).

[38] G. F. Giudice, *Naturalness after LHC8*, PoS **EPS-HEP2013** (2013) 163, arXiv: 1307.7879 [hep-ph] (cit. on p. 15).

[39] T. D. Lee, *A Theory of Spontaneous T Violation*, Phys. Rev. D **8** (1973) 1226, ed. by G. Feinberg (cit. on p. 15).

[40] M. Aoki, S. Kanemura and O. Seto, *Neutrino mass, Dark Matter and Baryon Asymmetry via TeV-Scale Physics without Fine-Tuning*, Phys. Rev. Lett. **102** (2009) 051805, arXiv: 0807.0361 [hep-ph] (cit. on p. 17).

[41] S. Gabriel and S. Nandi, *A New two Higgs doublet model*, Phys. Lett. B **655** (2007) 141, arXiv: hep-ph/0610253 (cit. on p. 17).

[42] S. M. Davidson and H. E. Logan, *Dirac neutrinos from a second Higgs doublet*, Phys. Rev. D **80** (2009) 095008, arXiv: 0906.3335 [hep-ph] (cit. on p. 17).

[43] B. Abi et al., *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, Phys. Rev. Lett. **126** (2021) 141801, arXiv: 2104.03281 [hep-ex] (cit. on p. 17).

[44] G. Arcadi, Á. S. de Jesus, T. B. de Melo, F. S. Queiroz and Y. S. Villamizar, *A 2HDM for the g-2 and dark matter*, Nucl. Phys. B **982** (2022) 115882, arXiv: 2104.04456 [hep-ph] (cit. on p. 17).

[45] F. J. Botella, F. Cornet-Gomez, C. Miró and M. Nebot, "Leptonic $g-2$ in 2HDM", *56th Rencontres de Moriond on Electroweak Interactions and Unified Theories*, 2022, arXiv: 2205.07350 [hep-ph] (cit. on p. 17).

[46] H. Wittig, "Progress on $(g-2)_\mu$ from Lattice QCD", *57th Rencontres de Moriond on Electroweak Interactions and Unified Theories*, 2023, arXiv: 2306.04165 [hep-ph] (cit. on p. 17).

[47] S. Kuberski et al., *Hadronic vacuum polarization in the muon g − 2: the short-distance contribution from lattice QCD*, JHEP **03** (2024) 172, arXiv: 2401.11895 [hep-lat] (cit. on p. 17).

[48] W. Altmannshofer, S. Gori, M. Pospelov and I. Yavin, *Quark flavor transitions in $L_\mu - L_\tau$ models*, Phys. Rev. D **89** (2014) 095033, arXiv: 1403.1269 [hep-ph] (cit. on p. 17).

[49] Y. S. Amhis et al., *Averages of b-hadron, c-hadron, and $\tau$-lepton properties as of 2018*, Eur. Phys. J. C **81** (2021) 226, arXiv: 1909.12524 [hep-ex] (cit. on p. 17).

[50] T. Aaltonen et al., *High-precision measurement of the W-boson mass with the CDF II detector*, Science **376** (2022) 170 (cit. on p. 17).

[51] G. Aad et al., *Measurement of the W-boson mass and width with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **84** (2024) 1309, arXiv: 2403.15085 [hep-ex] (cit. on p. 17).

[52] V. Chekhovsky et al., *High-precision measurement of the W boson mass with the CMS experiment at the LHC*, (2024), arXiv: 2412.13872 [hep-ex] (cit. on p. 17).

[53] K. Ghorbani and P. Ghorbani, *W-boson mass anomaly from scale invariant 2HDM*, Nucl. Phys. B **984** (2022) 115980, arXiv: 2204.09001 [hep-ph] (cit. on p. 17).

[54] H. Bahl, J. Braathen and G. Weiglein, *New physics effects on the W-boson mass from a doublet extension of the SM Higgs sector*, Phys. Lett. B **833** (2022) 137295, arXiv: 2204.05269 [hep-ph] (cit. on p. 17).

[55] J. P. Lees et al., *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, Phys. Rev. D **88** (2013) 072012, arXiv: 1303.0571 [hep-ex] (cit. on p. 17).

[56] S. Iguro and K. Tobe, *$R(D^{(*)})$ in a general two Higgs doublet model*, Nucl. Phys. B **925** (2017) 560, arXiv: 1708.06176 [hep-ph] (cit. on p. 17).

[57] I. P. Ivanov, *Minkowski space structure of the Higgs potential in 2HDM. II. Minima, symmetries, and topology*, Phys. Rev. D **77** (2008) 015017, arXiv: 0710.3490 [hep-ph] (cit. on p. 17).

[58] ATLAS and CMS Collaborations,
*Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC $pp$ collision data at $\sqrt{s} = 7$ and 8 TeV*,
JHEP **08** (2016) 045, arXiv: 1606.02266 [hep-ex] (cit. on p. 18).

[59] M. Raidal et al., *Flavour physics of leptons and dipole moments*,
Eur. Phys. J. C **57** (2008) 13, ed. by R. Fleischer, T. Hurth and M. L. Mangano,
arXiv: 0801.1826 [hep-ph] (cit. on p. 19).

[60] G. D'Ambrosio, G. F. Giudice, G. Isidori and A. Strumia,
*Minimal flavor violation: An Effective field theory approach*, Nucl. Phys. B **645** (2002) 155,
arXiv: hep-ph/0207036 (cit. on p. 19).

[61] D. Eriksson, J. Rathsman and O. Stal,
*2HDMC: Two-Higgs-Doublet Model Calculator Physics and Manual*,
Comput. Phys. Commun. **181** (2010) 189, arXiv: 0902.0851 [hep-ph]
(cit. on pp. 21, 151, 152).

[62] 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 (cit. on p. 21).

[63] R. Haag, J. T. Łopuszański and M. Sohnius,
*All possible generators of supersymmetries of the S-matrix*, Nuclear Physics B **88** (1975) 257,
ISSN: 0550-3213,
URL: https://www.sciencedirect.com/science/article/pii/0550321375902795
(cit. on p. 22).

[64] J. F. Gunion and H. E. Haber, *Higgs bosons in supersymmetric models (I)*,
Nuclear Physics B **272** (1986) 1, ISSN: 0550-3213,
URL: https://www.sciencedirect.com/science/article/pii/0550321386903408
(cit. on p. 22).

[65] J. Steggemann, *Extended Scalar Sectors*,
Annual Review of Nuclear and Particle Science **70** (2020) 197, ISSN: 1545-4134,
URL: https://www.annualreviews.org/content/journals/10.1146/annurev-nucl-032620-043846 (cit. on pp. 22–24).

[66] K. Inoue, A. Kakuto, H. Komatsu and S. Takeshita,
*Aspects of Grand Unified Models with Softly Broken Supersymmetry*,
Prog. Theor. Phys. **68** (1982) 927, [Erratum: Prog.Theor.Phys. 70, 330 (1983)] (cit. on p. 22).

[67] H. P. Nilles, *Supersymmetry, Supergravity and Particle Physics*, Phys. Rept. **110** (1984) 1
(cit. on p. 22).

[68] H. E. Haber and L. Stephenson Haskins, "Supersymmetric Theory and Models",
*Theoretical Advanced Study Institute in Elementary Particle Physics: Anticipating the Next Discoveries in Particle Physics*, WSP, 2018 355, arXiv: 1712.05926 [hep-ph] (cit. on p. 23).

[69] U. Ellwanger, C. Hugonie and A. M. Teixeira,
*The Next-to-Minimal Supersymmetric Standard Model*, Phys. Rept. **496** (2010) 1,
arXiv: 0910.1785 [hep-ph] (cit. on p. 24).

[70] J. E. Kim and H. P. Nilles, *The mu Problem and the Strong CP Problem*, Phys. Lett. B **138** (1984) 150 (cit. on p. 24).

[71] M. Endo, K. Hamaguchi, S. Iwamoto, K. Nakayama and N. Yokozaki, *Higgs mass and muon anomalous magnetic moment in the U(1) extended MSSM*, Phys. Rev. D **85** (2012) 095006, arXiv: 1112.6412 [hep-ph] (cit. on p. 24).

[72] P. Athron, S. F. King, D. J. Miller, S. Moretti and R. Nevzorov, *The Constrained Exceptional Supersymmetric Standard Model*, Phys. Rev. D **80** (2009) 035009, arXiv: 0904.2169 [hep-ph] (cit. on p. 24).

[73] S. F. King, S. Moretti and R. Nevzorov, *E(6)SSM*, AIP Conf. Proc. **881** (2007) 138, ed. by S. Khalil, arXiv: hep-ph/0610002 (cit. on p. 24).

[74] E. Ma, *Common origin of neutrino mass, dark matter, and baryogenesis*, Mod. Phys. Lett. A **21** (2006) 1777, arXiv: hep-ph/0605180 (cit. on p. 24).

[75] L. Lopez Honorez, E. Nezri, J. F. Oliver and M. H. G. Tytgat, *The Inert Doublet Model: An Archetype for Dark Matter*, JCAP **02** (2007) 028, arXiv: hep-ph/0612275 (cit. on p. 24).

[76] E. Ma, *Verifiable radiative seesaw mechanism of neutrino mass and dark matter*, Phys. Rev. D **73** (2006) 077301, arXiv: hep-ph/0601225 (cit. on p. 24).

[77] A. G. Akeroyd, S. Moretti, K. Yagyu and E. Yildirim, *Light charged Higgs boson scenario in 3-Higgs doublet models*, Int. J. Mod. Phys. A **32** (2017) 1750145, arXiv: 1605.05881 [hep-ph] (cit. on p. 24).

[78] H. Georgi and M. Machacek, *DOUBLY CHARGED HIGGS BOSONS*, Nucl. Phys. B **262** (1985) 463 (cit. on p. 25).

[79] J. F. Gunion, J. Grifols, A. Mendez, B. Kayser and F. I. Olness, *Higgs Bosons in Left-Right Symmetric Models*, Phys. Rev. D **40** (1989) 1546 (cit. on p. 25).

[80] E. Ma and U. Sarkar, *Neutrino masses and leptogenesis with heavy Higgs triplets*, Phys. Rev. Lett. **80** (1998) 5716, arXiv: hep-ph/9802445 (cit. on p. 25).

[81] A. Arbey, F. Mahmoudi, O. Stal and T. Stefaniak, *Status of the Charged Higgs Boson in Two Higgs Doublet Models*, Eur. Phys. J. C **78** (2018) 182, arXiv: 1706.07414 [hep-ph] (cit. on pp. 26, 31, 151).

[82] M. Misiak and M. Steinhauser, *Weak radiative decays of the B meson and bounds on $M_{H^\pm}$ in the Two-Higgs-Doublet Model*, Eur. Phys. J. C **77** (2017) 201, arXiv: 1702.04571 [hep-ph] (cit. on pp. 26, 153).

[83] T. Saito et al., *Measurement of the $\bar{B} \to X_s\gamma$ Branching Fraction with a Sum of Exclusive Decays*, Phys. Rev. D **91** (2015) 052004, arXiv: 1411.7198 [hep-ex] (cit. on p. 26).

[84] M. Misiak et al., *Updated NNLO QCD predictions for the weak radiative B-meson decays*, Phys. Rev. Lett. **114** (2015) 221801, arXiv: 1503.01789 [hep-ph] (cit. on p. 26).

[85] L. F. Abbott, P. Sikivie and M. B. Wise, *Constraints on Charged Higgs Couplings*, Phys. Rev. D **21** (1980) 1393 (cit. on p. 27).

[86] Z. Ligeti, M. Papucci and G. Perez,
*Implications of the measurement of the $B_s^0 - \bar{B}_s^0$ mass difference*,
Phys. Rev. Lett. **97** (2006) 101801, arXiv: hep-ph/0604112 (cit. on p. 27).

[87] S.-P. Li, X.-Q. Li and Y.-D. Yang, *Muon $g-2$ in a $U(1)$-symmetric Two-Higgs-Doublet Model*, Phys. Rev. D **99** (2019) 035010, arXiv: 1808.02424 [hep-ph] (cit. on p. 27).

[88] R. Aaij et al., *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays*, Phys. Rev. Lett. **118** (2017) 191801,
arXiv: 1703.05747 [hep-ex] (cit. on p. 27).

[89] ATLAS Collaboration, *Search for a light charged Higgs boson in the decay channel $H^+ \to c\bar{s}$ in $t\bar{t}$ events using $pp$ collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*,
Eur. Phys. J. C **73** (2013) 2465, arXiv: 1302.3694 [hep-ex] (cit. on pp. 28, 146).

[90] CMS Collaboration, *Search for a light charged Higgs boson in the $H^\pm \to cs$ channel in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. D **102** (2020) 072001,
arXiv: 2005.08900 [hep-ex] (cit. on pp. 28, 81, 129, 146).

[91] ATLAS Collaboration,
*Search for a light charged Higgs boson in $t \to H^\pm b$ decays, with $H^\pm \to cb$, in the lepton+jets final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
JHEP **09** (2023) 004, arXiv: 2302.11739 [hep-ex]
(cit. on pp. 28, 29, 64, 65, 76, 121, 139, 147, 149, 150, 152).

[92] CMS Collaboration, *Search for a charged Higgs boson decaying to charm and bottom quarks in proton–proton collisions at $\sqrt{s} = 8$ TeV*, JHEP **11** (2018) 115, arXiv: 1808.06575 [hep-ex]
(cit. on p. 28).

[93] G. Aad et al., *Search for charged Higgs bosons produced in top-quark decays or in association with top quarks and decaying via $H^\pm \to \tau^\pm \nu_\tau$ in 13 TeV $pp$ collisions with the ATLAS detector*, (2024), arXiv: 2412.17584 [hep-ex] (cit. on pp. 28, 152, 153).

[94] CMS Collaboration, *Search for charged Higgs bosons in the $H^\pm \to \tau^\pm \nu_\tau$ decay channel in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **07** (2019) 142, arXiv: 1903.04560 [hep-ex]
(cit. on pp. 28, 29).

[95] ATLAS Collaboration, *Search for $H^\pm \to W^\pm A \to W^\pm \mu\mu$ in $pp \to t\bar{t}$ events using an $e\mu\mu$ signature with the ATLAS detector at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2021-047, 2021,
URL: https://cds.cern.ch/record/2780092 (cit. on p. 28).

[96] CMS Collaboration, *Search for a light charged Higgs boson decaying to a W boson and a CP-odd Higgs boson in final states with $e\mu\mu$ or $\mu\mu\mu$ in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Phys. Rev. Lett. **123** (2019) 131802, arXiv: 1905.07453 [hep-ex] (cit. on p. 28).

[97] ATLAS Collaboration, *Search for charged Higgs bosons decaying into a top quark and a bottom quark at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **06** (2021) 145,
arXiv: 2102.10076 [hep-ex] (cit. on pp. 28, 30, 121, 139).

[98] CMS Collaboration, *Search for a charged Higgs boson decaying into top and bottom quarks in events with electrons or muons in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **01** (2020) 096,
arXiv: 1908.09206 [hep-ex] (cit. on pp. 28, 30).

[99] CMS Collaboration, *Search for charged Higgs bosons decaying into a top and a bottom quark in the all-jet final state of $pp$ collisions at $\sqrt{s} = 13$ TeV*, JHEP **07** (2020) 126, arXiv: 2001.07763 [hep-ex] (cit. on pp. 28, 30).

[100] G. Aad et al.,
*Search for a heavy charged Higgs boson decaying into a W boson and a Higgs boson in final states with leptons and b-jets in $\sqrt{s} = 13$ TeV $pp$ collisions with the ATLAS detector*, (2024), arXiv: 2411.03969 [hep-ex] (cit. on pp. 28, 30).

[101] CMS Collaboration, *Search for a charged Higgs boson decaying into a heavy neutral Higgs boson and a W boson in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **09** (2023) 032, arXiv: 2207.01046 [hep-ex] (cit. on pp. 28, 30).

[102] G. Aad et al., *Combination of searches for singly and doubly charged Higgs bosons produced via vector-boson fusion in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2024), arXiv: 2407.10798 [hep-ex] (cit. on pp. 28, 30).

[103] CMS Collaboration, *Search for charged Higgs bosons produced in vector boson fusion processes and decaying into vector boson pairs in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **81** (2021) 723, arXiv: 2104.04762 [hep-ex] (cit. on pp. 28, 30).

[104] F. Abe et al.,
*Search for charged Higgs decays of the top quark using hadronic decays of the tau lepton*, Phys. Rev. Lett. **79** (1997) 357, arXiv: hep-ex/9704003 (cit. on p. 27).

[105] A. Abulencia et al.,
*Search for charged Higgs bosons from top quark decays in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$-TeV.*, Phys. Rev. Lett. **96** (2006) 042003, arXiv: hep-ex/0510065 (cit. on p. 27).

[106] V. M. Abazov et al., *Search for Charged Higgs Bosons in Top Quark Decays*, Phys. Lett. B **682** (2009) 278, arXiv: 0908.1811 [hep-ex] (cit. on p. 27).

[107] T. Aaltonen et al., *Search for charged Higgs bosons in decays of top quarks in p anti-p collisions at s**(1/2) = 1.96 TeV*, Phys. Rev. Lett. **103** (2009) 101803, arXiv: 0907.1269 [hep-ex] (cit. on pp. 27, 28, 90).

[108] G. Abbiendi et al., *Search for Charged Higgs bosons: Combined Results Using LEP Data*, Eur. Phys. J. C **73** (2013) 2463, arXiv: 1301.6065 [hep-ex] (cit. on p. 27).

[109] J. Erdmann et al.,
*A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework*, Nucl. Instrum. Meth. A **748** (2014) 18, arXiv: 1312.5595 [hep-ex] (cit. on pp. 28, 90).

[110] 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*, JHEP **09** (2018) 139, arXiv: 1807.07915 [hep-ex] (cit. on p. 29).

[111] *Das Innere des Protons*, DESY Mediendatenbank,
URL: https://media.desy.de/DESYmediabank/catalog/Presse/r/20640/viewmode=infoview/qsr=proton (visited on 28/10/2024) (cit. on p. 32).

[112] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972) 438 (cit. on p. 33).

[113] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, Sov. Phys. JETP **46** (1977) 641 (cit. on p. 33).

[114] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. B **126** (1977) 298 (cit. on p. 33).

[115] 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 (cit. on p. 33).

[116] Machine Learning for Hadronization collaboration documentation, URL: https://uchep.gitlab.io/mlhad-docs/ (visited on 28/10/2024) (cit. on p. 34).

[117] C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, SciPost Phys. Codebases (2022) 8, URL: https://scipost.org/10.21468/SciPostPhysCodeb.8 (cit. on p. 34).

[118] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det] (cit. on p. 34).

[119] B. Andersson, G. Gustafson and B. Soderberg, *A General Model for Jet Fragmentation*, Z. Phys. C **20** (1983) 317 (cit. on p. 35).

[120] T. Sjöstrand, *Jet fragmentation of multiparton configurations in a string framework*, Nucl. Phys. B **248** (1984) 469 (cit. on p. 35).

[121] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl. Phys. B **238** (1984) 492 (cit. on p. 35).

[122] S. Agostinelli et al., *GEANT4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 35).

[123] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, 2010, URL: https://cds.cern.ch/record/1300517 (cit. on p. 35).

[124] A. Wu Chao, M. Tigner, H. Weise and F. Zimmermann, eds., *Handbook of Accelerator Physics and Engineering*, World Scientific, 2023, ISBN: 978-981-12-6917-2, 978-981-12-7015-4, 978-981-12-6919-6 (cit. on p. 36).

[125] S. Peggs and T. Satogata, *Introduction to Accelerator Dynamics*, Cambridge University Press, 2017 (cit. on p. 36).

[126] E. Cenni, *SRF accelerator cavities*, EASISchool 2 on Cryogenics, 2019, URL: https://indico.cern.ch/event/792215/contributions/3408423 (visited on 01/10/2024) (cit. on p. 37).

[127] F. Tecker, "Longitudinal Beam Dynamics", *CERN Accelerator School: Advanced Accelerator Physics Course*, 2014 1, arXiv: 1601.04901 [physics.acc-ph] (cit. on p. 37).

[128] E. Lopienska, *The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022*, (2022), General Photo, URL: https://cds.cern.ch/record/2800984 (cit. on p. 38).

[129] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 38–41, 43, 45).

[130] Sebastian Fleischmann, *Tau lepton reconstruction with energy flow and the search for R-parity violating supersymmetry at the ATLAS experiment*, PhD thesis: Rheinische Friedrich-Wilhelms-Universität Bonn, 2012, URL: https://hdl.handle.net/20.500.11811/5258 (cit. on p. 39).

[131] G. Aad et al., *Alignment of the ATLAS Inner Detector in Run-2*, Eur. Phys. J. C **80** (2020) 1194, arXiv: 2007.07624 [hep-ex] (cit. on pp. 40, 41).

[132] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST **13** (2018) P07017 (cit. on p. 46).

[133] ATLAS Collaboration, *Luminosity determination in $pp$ collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **83** (2023) 982, arXiv: 2212.09379 [hep-ex] (cit. on pp. 46, 61, 71, 116).

[134] S. van der Meer, *Calibration of the Effective Beam Height in the ISR*, (1968) (cit. on p. 46).

[135] P. Grafström and W. Kozanecki, *Luminosity determination at proton colliders*, Prog. Part. Nucl. Phys. **81** (2015) 97 (cit. on p. 46).

[136] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004, arXiv: 2007.12539 [hep-ex] (cit. on pp. 46, 47).

[137] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex] (cit. on p. 46).

[138] ATLAS Collaboration, *ATLAS TDAQ System Phase-I Upgrade: Technical Design Report*, ATLAS-TDR-023; CERN-LHCC-2013-018, 2013, URL: https://cds.cern.ch/record/1602235 (cit. on p. 46).

[139] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex] (cit. on p. 48).

[140] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. A **262** (1987) 444 (cit. on p. 48).

[141] ATLAS Collaboration, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, JINST **9** (2014) P09009, arXiv: 1406.7690 [hep-ex] (cit. on p. 48).

[142] ATLAS Collaboration, *Early Inner Detector Tracking Performance in the 2015 Data at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-051, 2015, URL: https://cds.cern.ch/record/2110140 (cit. on p. 48).

[143] *Track Seeding and Finding*, ATLAS Software Documentation: tracking-tutorial, URL: https://atlassoftwaredocs.web.cern.ch/internal-links/tracking-tutorial/tracksf/ (visited on 21/11/2024) (cit. on p. 49).

[144] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, Eur. Phys. J. C **77** (2017) 332, arXiv: 1611.10235 [hep-ex] (cit. on p. 49).

[145] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s}$ = 13 TeV*, ATL-PHYS-PUB-2015-026, 2015, URL: https://cds.cern.ch/record/2037717 (cit. on p. 49).

[146] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex] (cit. on pp. 49, 50).

[147] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **81** (2021) 578, arXiv: 2012.00578 [hep-ex] (cit. on pp. 50–52, 118).

[148] J. Illingworth and J. Kittler, *A survey of the hough transform*, Computer Vision, Graphics, and Image Processing **44** (1988) 87, ISSN: 0734-189X, URL: https://www.sciencedirect.com/science/article/pii/S0734189X88800331 (cit. on p. 50).

[149] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv: 1603.05598 [hep-ex] (cit. on p. 50).

[150] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex] (cit. on pp. 51, 118).

[151] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, URL: https://cds.cern.ch/record/1449796 (cit. on p. 51).

[152] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: 1902.04655 [physics.ins-det] (cit. on p. 51).

[153] ATLAS Collaboration, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. D **97** (2018) 072003, arXiv: 1712.08891 [hep-ex] (cit. on p. 52).

[154] ATLAS Collaboration, *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, JINST **18** (2023) T11004, arXiv: 2211.16178 [hep-ex] (cit. on pp. 52, 74).

[155] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data*, (2023), arXiv: 2309.05471 [hep-ex] (cit. on pp. 52, 55).

[156] ATLAS Collaboration, *Studies of the muon momentum calibration and performance of the ATLAS detector with pp collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **83** (2023) 686, arXiv: 2212.07338 [hep-ex] (cit. on pp. 52, 55, 118).

[157] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data*, JINST **14** (2019) P03017, arXiv: 1812.03848 [hep-ex] (cit. on p. 52).

[158] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on p. 53).

[159] M. Cacciari, G. P. Salam and G. Soyez, *The anti-$k_t$ jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on pp. 53, 59).

[160] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph] (cit. on p. 53).

[161] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81** (2021) 689, arXiv: 2007.02645 [hep-ex] (cit. on pp. 54, 55, 62).

[162] 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. C **76** (2016) 581, arXiv: 1510.03823 [hep-ex] (cit. on pp. 55, 118).

[163] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 $pp$ collision dataset*, Eur. Phys. J. C **83** (2023) 681, arXiv: 2211.16345 [physics.data-an] (cit. on pp. 55, 57, 153).

[164] I. Neutelings, *B tagging jets*, URL: https://tikz.net/jet_btag/ (visited on 16/01/2025) (cit. on p. 56).

[165] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017, URL: https://cds.cern.ch/record/2273281 (cit. on p. 55).

[166] ATLAS Collaboration, *Identification of Jets Containing b-Hadrons with Recurrent Neural Networks at the ATLAS Experiment*, ATL-PHYS-PUB-2017-003, 2017, URL: https://cds.cern.ch/record/2255226 (cit. on p. 55).

[167] ATLAS Collaboration, *Secondary vertex finding for jet flavour identification with the ATLAS detector*, ATL-PHYS-PUB-2017-011, 2017, URL: https://cds.cern.ch/record/2270366 (cit. on p. 56).

[168] ATLAS Collaboration, *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*, ATL-PHYS-PUB-2018-025, 2018, URL: https://cds.cern.ch/record/2645405 (cit. on p. 56).

[169] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in $pp$ collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 970, arXiv: 1907.05120 [hep-ex] (cit. on pp. 56, 118).

[170] ATLAS Collaboration, *Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using $pp$ collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector*, Eur. Phys. J. C **82** (2022) 95, arXiv: 2109.10627 [hep-ex] (cit. on p. 56).

[171] ATLAS Collaboration,
*Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using* $139\,fb^{-1}$ *of ATLAS proton–proton collision data at* $\sqrt{s} = 13\,TeV$,
Eur. Phys. J. C **83** (2023) 728, arXiv: 2301.06319 [hep-ex] (cit. on p. 56).

[172] ATLAS Collaboration, *Measurements of WH and ZH production with Higgs boson decays into bottom quarks and direct constraints on the charm Yukawa coupling in* 13 TeV $pp$ *collisions with the ATLAS detector*, (2024), arXiv: 2410.19611 [hep-ex] (cit. on pp. 57, 58).

[173] ATLAS Collaboration, *The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using* $140\,fb^{-1}$ *of* $\sqrt{s} = 13\,TeV$ $pp$ *collisions*, (2024), arXiv: 2402.05858 [hep-ex] (cit. on p. 58).

[174] ATLAS Collaboration, *Search for a light charged Higgs boson in* $t \to H^{\pm}b$ *decays, with* $H^{\pm} \to cs$, *in* $pp$ *collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, (2024), arXiv: 2407.10096 [hep-ex] (cit. on p. 59).

[175] M. Cacciari, G. P. Salam and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv: 0802.1188 [hep-ph] (cit. on p. 59).

[176] ATLAS Collaboration,
*ATLAS data quality operations and performance for 2015–2018 data-taking*,
JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det] (cit. on p. 61).

[177] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*,
Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph] (cit. on p. 61).

[178] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*,
ATL-PHYS-PUB-2016-017, 2016, URL: https://cds.cern.ch/record/2206965
(cit. on p. 61).

[179] NNPDF Collaboration, R. D. Ball et al., *Parton distributions with LHC data*,
Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph] (cit. on p. 61).

[180] S. Frixione, G. Ridolfi and P. Nason,
*A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*,
JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph] (cit. on p. 61).

[181] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*,
JHEP **11** (2004) 040, arXiv: hep-ph/0409146 (cit. on p. 61).

[182] S. Frixione, P. Nason and C. Oleari,
*Matching NLO QCD computations with parton shower simulations: the POWHEG method*,
JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph] (cit. on p. 61).

[183] 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] (cit. on p. 61).

[184] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC run II*,
JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph] (cit. on p. 61).

[185] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, URL: https://cds.cern.ch/record/2216168 (cit. on pp. 61, 64, 119).

[186] 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] (cit. on p. 61).

[187] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph] (cit. on p. 61).

[188] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419 (cit. on p. 61).

[189] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152 (cit. on p. 62).

[190] S. Dittmaier et al., *Handbook of LHC Higgs Cross Sections: 2. Differential Distributions*, (2012), arXiv: 1201.3084 [hep-ph] (cit. on p. 63).

[191] S. Heinemeyer, W. Hollik and G. Weiglein, *FeynHiggs: A Program for the calculation of the masses of the neutral CP even Higgs bosons in the MSSM*, Comput. Phys. Commun. **124** (2000) 76, arXiv: hep-ph/9812320 (cit. on p. 63).

[192] S. M. Moosavi Nejad, S. Abbaspour and R. Farashahian, *Interference effects for the top quark decays $t \to b + W^+/H^+(\to \tau^+\nu_\tau)$*, Phys. Rev. D **99** (2019) 095012, arXiv: 1904.09680 [hep-ph] (cit. on p. 63).

[193] 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] (cit. on p. 64).

[194] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. D **83** (2011) 091503, arXiv: 1103.2792 [hep-ph] (cit. on p. 64).

[195] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, Phys. Rev. D **81** (2010) 054028, arXiv: 1001.5034 [hep-ph] (cit. on p. 64).

[196] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a $W^-$ or $H^-$*, Phys. Rev. D **82** (2010) 054018, arXiv: 1005.4451 [hep-ph] (cit. on p. 64).

[197] M. Aliev et al., *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*, Comput. Phys. Commun. **182** (2011) 1034, arXiv: 1007.1327 [hep-ph] (cit. on pp. 64, 118).

[198] P. Kant 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] (cit. on pp. 64, 118).

[199] G. Aad et al., *Measurement of $t\bar{t}$ production in association with additional b-jets in the $e\mu$ final state in proton-proton collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, (2024), arXiv: 2407.13473 [hep-ex] (cit. on p. 64).

[200] G. Aad et al., *Measurement of top-quark pair production in association with charm quarks in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, (2024), arXiv: 2409.11305 [hep-ex] (cit. on p. 64).

[201] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in $pp$ collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, Eur. Phys. J. C **75** (2015) 349, arXiv: 1503.05066 [hep-ex] (cit. on p. 64).

[202] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph] (cit. on p. 65).

[203] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph] (cit. on p. 65).

[204] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008) 039, arXiv: 0808.3674 [hep-ph] (cit. on p. 65).

[205] F. Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79** (2019) 866, arXiv: 1907.13071 [hep-ph] (cit. on p. 65).

[206] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601, arXiv: 1111.5206 [hep-ph] (cit. on p. 65).

[207] A. Denner, S. Dittmaier and L. Hofer, *Collier: A fortran-based complex one-loop library in extended regularizations*, Comput. Phys. Commun. **212** (2017) 220, arXiv: 1604.06792 [hep-ph] (cit. on p. 65).

[208] ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, ATL-PHYS-PUB-2017-005, 2017, URL: https://cds.cern.ch/record/2261933 (cit. on p. 65).

[209] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP **09** (2012) 049, arXiv: 1111.1220 [hep-ph] (cit. on p. 65).

[210] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, JHEP **04** (2013) 027, arXiv: 1207.5030 [hep-ph] (cit. on p. 65).

[211] S. Catani, F. Krauss, B. R. Webber and R. Kuhn, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001) 063, arXiv: hep-ph/0109231 (cit. on p. 65).

[212] I. Neutelings, *Piechart of SM decays*, URL: https://tikz.net/sm_decay_piechart/ (visited on 12/12/2024) (cit. on p. 67).

[213] I. Neutelings, *Branching fraction matrix of pair decays*, URL: https://tikz.net/sm_decay_matrix/ (visited on 12/12/2024) (cit. on p. 67).

[214] ATLAS Collaboration, *2015 start-up trigger menu and initial performance assessment of the ATLAS trigger using Run-2 data*, ATL-DAQ-PUB-2016-001, 2016, URL: https://cds.cern.ch/record/2136007 (cit. on p. 66).

[215] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015, arXiv: 2004.13447 [physics.ins-det] (cit. on p. 66).

[216] ATLAS Collaboration,
*Performance of electron and photon triggers in ATLAS during LHC Run 2*,
Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex] (cit. on p. 66).

[217] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s}$ = 13 TeV with the ATLAS experiment*, Phys. Lett. B **810** (2020) 135797,
arXiv: 2006.13076 [hep-ex] (cit. on pp. 71, 76).

[218] W. Buttinger, *Background Estimation with the ABCD Method – Featuring the TRooFit Toolkit*, (2018) (cit. on p. 71).

[219] CDF Collaboration,
*A Measurement of $\sigma B(W \to e\nu)$ and $\sigma B(Z^0 \to e^+e^-)$ in $\overline{p}p$ collisions at $\sqrt{s}$ = 1800 GeV*,
Phys. Rev. D **44** (1991) 29 (cit. on p. 71).

[220] ATLAS Collaboration, *Measurement of differential cross-sections in $t\bar{t}$ and $t\bar{t}$+jets production in the lepton+jets final state in $pp$ collisions at $\sqrt{s}$ = 13 TeV using 140 $fb^{-1}$ of ATLAS data*,
JHEP **08** (2024) 182, arXiv: 2406.19701 [hep-ex] (cit. on pp. 74, 76, 138).

[221] M. Czakon et al., *Top-pair production at the LHC through NNLO QCD and NLO EW*,
JHEP **10** (2017) 186, arXiv: 1705.04105 [hep-ph] (cit. on p. 76).

[222] L. Serkin, *Treatment of top-quark backgrounds in extreme phase spaces: the "top $p_T$ reweighting" and novel data-driven estimations in ATLAS and CMS*, tech. rep.,
Proceeding for 13th International Workshop on Top Quark Physics: CERN, 2021,
arXiv: 2105.03977 (cit. on p. 76).

[223] J. E. Gaiser, *Charmonium Spectroscopy From Radiative Decays of the $J/\psi$ and $\psi'$*, Appendix F, PhD thesis, 1982 (cit. on p. 84).

[224] M. Oreglia, *A Study of the Reactions $\psi' \to \gamma\gamma\psi$*, Appendix D, PhD thesis, 1980 (cit. on p. 84).

[225] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*, Appendix E, PhD thesis: Cracow, INP, 1986 (cit. on p. 84).

[226] Placeholder, *KLFitter*,
URL: https://github.com/KLFitter/KLFitter (visited on 17/09/2024) (cit. on p. 90).

[227] A. Shmakov et al., *SPANet: Generalized permutationless set assignment for particle physics using symmetry preserving attention*, SciPost Phys. **12** (2022) 178,
arXiv: 2106.03898 [hep-ex] (cit. on p. 92).

[228] Y. Kats and E. Ofir,
*From strange-quark tagging to fragmentation tagging with machine learning*, (2024),
arXiv: 2408.12377 [hep-ph] (cit. on p. 96).

[229] Y. Nakai, D. Shih and S. Thomas, *Strange Jet Tagging*, (2020), arXiv: 2003.09517 [hep-ph] (cit. on p. 96).

[230] A. Czarnecki, J. G. Korner and J. H. Piclum,
*Helicity fractions of W bosons from top quark decays at NNLO in QCD*,
Phys. Rev. D **81** (2010) 111503, arXiv: 1005.2625 [hep-ph] (cit. on p. 97).

[231] G. Aad et al., *Measurement of the polarisation of W bosons produced in top-quark decays using dilepton events at s=13 TeV with the ATLAS experiment*, Phys. Lett. B **843** (2023) 137829,
arXiv: 2209.14903 [hep-ex] (cit. on p. 97).

[232] J. A. Aguilar-Saavedra and J. Bernabeu,
*Breaking down the entire W boson spin observables from its decay*,
Phys. Rev. D **93** (2016) 011301, arXiv: 1508.04592 [hep-ph] (cit. on p. 97).

[233] T. Hastie, R. Tibshirani and J. Friedman,
*The elements of statistical learning: data mining, inference and prediction*, 2nd ed.,
Springer, 2009 (cit. on pp. 101, 104, 111).

[234] J. Friedman, T. Hastie and R. Tibshirani,
*Additive Logistic Regression: A Statistical View of Boosting*,
The Annals of Statistics **28** (2000) 337 (cit. on p. 101).

[235] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*,
The Annals of Statistics **29** (2001) 1189,
URL: https://doi.org/10.1214/aos/1013203451 (cit. on p. 101).

[236] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, (2016),
arXiv: 1603.02754 [cs.LG] (cit. on p. 101).

[237] J. Bergstra, B. Komer, C. Eliasmith, D. Yamins and D. D. Cox,
*Hyperopt: a Python library for model selection and hyperparameter optimization*,
Comput. Sci. Discov. **8** (2015) 014008 (cit. on p. 111).

[238] J. Bergstra, R. Bardenet, Y. Bengio and B. Kégl,
"Algorithms for Hyper-Parameter Optimization",
*Advances in Neural Information Processing Systems*,
ed. by J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira and K. Weinberger, vol. 24,
Curran Associates, Inc., 2011, URL: https://proceedings.neurips.cc/paper_files/paper/2011/file/86e8f7ab32cfd12577bc2619bc635690-Paper.pdf (cit. on p. 111).

[239] A. Ghosh, B. Nachman and D. Whiteson,
*Uncertainty-aware machine learning for high energy physics*, Phys. Rev. D **104** (2021) 056026,
arXiv: 2105.08742 [physics.data-an] (cit. on p. 114).

[240] J. Zhou et al., *Graph neural networks: A review of methods and applications*,
AI Open **1** (2020) 57, ISSN: 2666-6510,
URL: https://www.sciencedirect.com/science/article/pii/S2666651021000012 (cit. on p. 115).

[241] A. Vaswani et al., *Attention Is All You Need*, 2023, arXiv: 1706.03762 [cs.CL],
URL: https://arxiv.org/abs/1706.03762 (cit. on p. 115).

[242] ATLAS Collaboration, *Graph Neural Network Jet Flavour Tagging with the ATLAS Detector*,
ATL-PHYS-PUB-2022-027, 2022, URL: https://cds.cern.ch/record/2811135 (cit. on p. 115).

[243] 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] (cit. on p. 116).

[244] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex] (cit. on p. 118).

[245] ATLAS Collaboration, *Jet energy resolution in proton–proton collisions at $\sqrt{s}$ = 7 TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2306, arXiv: 1210.6210 [hep-ex] (cit. on p. 118).

[246] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001, 2018, URL: https://cds.cern.ch/record/2306649 (cit. on p. 118).

[247] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s}$ = 13 TeV*, ATLAS-CONF-2018-006, 2018, URL: https://cds.cern.ch/record/2314418 (cit. on p. 118).

[248] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s}$ = 13 TeV*, Eur. Phys. J. C **78** (2018) 903, arXiv: 1802.08168 [hep-ex] (cit. on p. 118).

[249] ATLAS Collaboration, *$E_T^{miss}$ performance in the ATLAS detector using 2015–2016 LHC $pp$ collisions*, ATLAS-CONF-2018-023, 2018, URL: https://cds.cern.ch/record/2625233 (cit. on p. 118).

[250] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189, arXiv: 0901.0002 [hep-ph] (cit. on p. 118).

[251] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Uncertainties on $\alpha_S$ in global PDF analyses and implications for predicted hadronic cross sections*, Eur. Phys. J. C **64** (2009) 653, arXiv: 0905.3531 [hep-ph] (cit. on p. 118).

[252] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph] (cit. on p. 118).

[253] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph] (cit. on p. 119).

[254] J. Bellm et al., *Herwig 7.1 Release Note*, (2017), arXiv: 1705.06919 [hep-ph] (cit. on p. 119).

[255] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 204, arXiv: 1412.3989 [hep-ph] (cit. on p. 119).

[256] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: 0805.3067 [hep-ph] (cit. on p. 119).

[257] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini and M. Schönherr, *NLO electroweak automation and precise predictions for W+multijet production at the LHC*, JHEP **04** (2015) 012, arXiv: 1412.5157 [hep-ph] (cit. on p. 119).

[258] ATLAS Collaboration, *Modelling and computational improvements to the simulation of single vector-boson plus jet processes for the ATLAS experiment*, JHEP **08** (2022) 089, arXiv: 2112.09588 [hep-ex] (cit. on p. 119).

[259] C. Bierlich et al.,
*Robust independent validation of experiment and theory: Rivet version 4 release note*,
SciPost Phys. Codeb. **36** (2024) 1, arXiv: 2404.15984 [hep-ph] (cit. on p. 119).

[260] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke,
*HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep.,
New York U., 2012, URL: https://cds.cern.ch/record/1456844 (cit. on p. 120).

[261] L. Heinrich, M. Feickert, G. Stark and K. Cranmer,
*pyhf: pure-Python implementation of HistFactory statistical models*,
J. Open Source Softw. **6** (2021) 2823 (cit. on p. 120).

[262] D0 Collaboration, *Determination of the Width of the Top Quark*,
Phys. Rev. Lett. **106** (2011) 022001, arXiv: 1009.5686 [hep-ex] (cit. on p. 120).

[263] ATLAS Collaboration, *Measurements of fiducial cross-sections for $t\bar{t}$ production with one or two additional b-jets in $pp$ collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector*,
Eur. Phys. J. C **76** (2016) 11, arXiv: 1508.06868 [hep-ex] (cit. on pp. 121, 139).

[264] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **04** (2019) 046, arXiv: 1811.12113 [hep-ex] (cit. on pp. 121, 139).

[265] ATLAS Collaboration, *Measurement of the $t\bar{t}t\bar{t}$ production cross section in $pp$ collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **11** (2021) 118, arXiv: 2106.11683 [hep-ex] (cit. on pp. 121, 139).

[266] ATLAS Collaboration, *Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in proton–proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*,
Eur. Phys. J. C **78** (2018) 565, arXiv: 1804.10823 [hep-ex] (cit. on p. 122).

[267] *1974 CERN School of Computing, Godoysund, Norway, 11-24 Aug 1974: Proceedings*,
CERN Yellow Reports: School Proceedings, 1974 (cit. on p. 122).

[268] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*,
Comput. Phys. Commun. **77** (1993) 219 (cit. on p. 122).

[269] G. Cowan, K. Cranmer, E. Gross and O. Vitells,
*Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554,
arXiv: 1007.1727 [physics.data-an] (cit. on pp. 123–126),
Erratum: Eur. Phys. J. C **73** (2013) 2501.

[270] A. Wald, *Tests of statistical hypotheses concerning several parameters when the number of observations is large*, Transactions of the American Mathematical Society **54** (1943) 426,
URL: https://api.semanticscholar.org/CorpusID:54174575 (cit. on p. 123).

[271] A. L. Read, *Presentation of search results: the $CL_S$ technique*, J. Phys. G **28** (2002) 2693
(cit. on p. 127).

[272] ATLAS Collaboration, *Search for top quark decays $t \to qH$, with $H \to \gamma\gamma$, in $\sqrt{s}$ = 13 TeV $pp$ collisions using the ATLAS detector*, JHEP **10** (2017) 129, arXiv: 1707.01404 [hep-ex]
(cit. on p. 129).

[273] CMS Collaboration, *Measurement of the cross section for $t\bar{t}$ production with additional jets and b jets in pp collisions at $\sqrt{s}$ = 13 TeV*, JHEP **07** (2020) 125, arXiv: 2003.06467 [hep-ex] (cit. on p. 139).

[274] ATLAS Collaboration,
*Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*,
ATL-PHYS-PUB-2019-033, 2019, URL: https://cds.cern.ch/record/2688062
(cit. on p. 153).

[275] ATLAS Collaboration, *Search for pair-produced vector-like top and bottom partners in events with large missing transverse momentum in pp collisions with the ATLAS detector*,
Eur. Phys. J. C **83** (2023) 719, arXiv: 2212.05263 [hep-ex] (cit. on p. 155).

# List of Figures

# List of Tables

# Acknowledgements

Lastly, I would like to thank all the people that helped me during my postgraduate education.

First of all, Prof. Dr. Jochen Dingfelder, who gave me the opportunity to work on a scientific dissertation in his research group. Jochen gave me a lot of freedom in my studies without any restrictions or time pressure. It was a really nice experience and a pleasure to work on such an exciting topic, such as searching for a new undiscovered particle.

Probably the most important person during my studies was my supervisor, Tatjana Lenz. She provided helpful advice in analysis-related discussion talks and overall guidance in ATLAS related procedures. To mention here is, for example, her help in publishing the analysis results. And one would always get quick feedback from her; it doesn't matter whether it was late at night or early in the morning.

I would also like to thank all the other group members that I have been working together with over the years. For their feedback on the analysis – in this regard I would like to highlight Christopher Deutsch – and enjoyable discussion during lunch breaks. I would also like to thank ATLAS members for providing feedback on the analysis.

Not to forget is the support of my family – Rainer, Maria and Martin. Without them, I would not have been able to perform my physics graduation in the first place.