

# GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Uploaddatum: 27.09.2024

Uploadzeit: 15:41

**Dies ist ein von FlexNow automatisch beim Upload generiertes Deckblatt. Es dient dazu, die Arbeit automatisiert der Prüfungsakte zuordnen zu können.**

**This is a machine generated frontpage added by FlexNow.  
Its purpose is to link your upload to your examination file.**

Matrikelnummer: 21975023



## Master's Thesis

# Untersuchung von Top-Quark Paarproduktionen mit einem Higgs Zerfall in zwei W-Bosonen in einfach-leptonischen Endzuständen unter Nutzung von ATLAS

## Studies of top quark pair production associated with an Higgs boson decaying into two W-bosons in single-leptonic final states using ATLAS

prepared by

**Ireas Tom Raschke**

from Finsterwalde

at the II. Physikalisches Institut

**Thesis Number:** II.Physik-UniGö-MSc-2024/07

**Time Period:** 1st April 2024 until 30th September 2024

**First Referee:** Prof. Dr. Arnulf Quadt

**Second Referee:** Prof. Dr. Ariane Frey



# Contents

<b>1. Introduction</b>	<b>1</b>
<b>2. The Standard Model of Particle Physics</b>	<b>3</b>
2.1. Gauge Symmetry and Spontaneous Symmetry Breaking . . . . .	3
2.2. Particles . . . . .	4
2.3. Beyond the Standard Model . . . . .	9
2.4. Higgs-Associated Top Quark Pair Production . . . . .	10
<b>3. Experimental Setup</b>	<b>13</b>
3.1. Large Hadron Collider . . . . .	13
3.2. The ATLAS Experiment . . . . .	14
3.3. Other Experiments . . . . .	20
<b>4. Truth Matching</b>	<b>23</b>
4.1. Algorithm . . . . .	23
4.2. Optimisation . . . . .	24
<b>5. SPA-Net: Symmetry Preserving Attention Networks</b>	<b>25</b>
5.1. Classical $\chi^2$ Method . . . . .	25
5.2. Modern DNN Method . . . . .	26
5.3. Advantages . . . . .	28
<b>6. Neutrino Weighting</b>	<b>29</b>
6.1. Parameter Estimation . . . . .	29
6.2. Optimisation . . . . .	30
<b>7. Event Selection</b>	<b>33</b>
7.1. Objects of the Target Event Topology . . . . .	33
7.2. Separation Power . . . . .	34
7.3. Region Definition . . . . .	39

*Contents*

<b>8. Results</b>	<b>43</b>
8.1. Truth Matching . . . . .	43
8.2. SPA-Net . . . . .	44
8.3. Neutrino Weighting . . . . .	48
8.4. Data . . . . .	51
<b>9. Discussion</b>	<b>55</b>
9.1. Summary . . . . .	55
9.2. Future Aspects . . . . .	56
<b>Bibliography</b>	<b>59</b>
<b>A. Additional SPA-Net Output Probabilities</b>	<b>65</b>
<b>B. Additional Comparisons of ATLAS Dataset and MC Samples</b>	<b>67</b>

# 1. Introduction

As far back as 400 BC, humans pondered about the fundamental makeup of the natural world. The earliest records of this go back to the Greek philosopher Democritus who proposed that matter cannot be infinitely divisible. Thus, there must exist the smallest, indivisible particle, which he termed the ‘atom’ (from Greek *atomos* meaning ‘indivisible’). Although this idea was not empirically tested but derived from philosophical thoughts, it marked the inception of the quest for a fundamental understanding of the universe.

Over the following millennia, this idea was refined by several researchers using different modern techniques. Today, research into the fundamental structure of nature continues with particle physics superseding nuclear physics. The contemporary aim of particle physics is to develop a comprehensive understanding of fundamental phenomena, centred around formulating and expanding upon the Standard Model of particle physics, which stands as the most successful description of the fundamental nature of the universe.

At the forefront of this quest for knowledge stands the Large Hadron Collider (LHC) situated at CERN (European Organization for Nuclear Research). The LHC, the world’s most powerful particle accelerator, collides protons which reach nearly the speed of light, mimicking conditions that prevailed in the universe mere moments after the Big Bang. Through detailed analysis of these collisions, scientists aim to validate and refine the Standard Model of particle physics.

One of the main parts at the LHC is the ATLAS experiment which takes data using a toroidally shaped particle detector around the collision beam pipe utilising cutting-edge technologies. This data is then used to study the laws of physics and to formulate models based on these observations.

The task of this thesis is to identify, separate and reconstruct the top quark pair production with an associated Higgs boson ( $H$ -boson) in which the  $H$ -boson decays into two  $W$ -bosons. Specifically, the single-leptonic decay channel is targeted. Due to its challenging topology and low cross-section, this particular channel has not been analysed independently before. Hence, a modern approach utilising techniques such as neural networks and *Neutrino Weighting* is deployed. The usage of these techniques for event reconstructions is tested and evaluated in this thesis. The goal is apply the reconstruction to the

## 1. Introduction

ATLAS dataset collected in Run I and Run II and to compare the results to the predicted simulated event samples under the usage of the aforementioned techniques. A successful reconstruction with background separation would allow for several measurements such as the top Yukawa coupling or Higgs branching ratios using this channel.

This study begins by introducing the theoretical foundation of the Standard Model in Ch. 2. The following Ch. 3 introduces the general experimental setup of the LHC and the specifics of the ATLAS detector. Ch. 4 covers the developed truth matching algorithm used for preparing the needed samples for training the neural network. The techniques applied in this project are explained in Ch. 5 and Ch. 6, which cover *SPA-Net* and *Neutrino Weighting*, respectively. The selection of events for this analysis is described in Ch. 7. The main results of this study are listed and explained in Ch. 8. Lastly, Ch. 9 summarises the results and discusses whether the targeted topology can be reconstructed and separated from background. Additionally, potential future tasks for further research are proposed.

# 2. The Standard Model of Particle Physics

The Standard Model of particle physics (SM) [1] stands as the most successful framework in understanding the fundamental particles and their interactions to date. It encompasses all known elementary particles along with their anti-particles, and three out of the four known fundamental forces: the strong, weak, and electromagnetic forces. Notably, gravity remains beyond the scope of the SM.

## 2.1. Gauge Symmetry and Spontaneous Symmetry Breaking

The SM is a renormalisable quantum field theory [2] characterised by an internal gauge symmetry that is denoted as a  $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ . A gauge symmetry is a mathematical symmetry that describes how the Lagrangian is invariant under local transformations. This means that the laws of physics do not change under these transformations. Gauge transformations form a Lie group [3] which necessitates the generation of an underlying gauge field [4].

The  $SU(3)_C$  group symmetry describes the strong interaction, stemming from quantum chromodynamics (QCD) [5–9]. It introduces colour charges and their corresponding mediators, the gluons, which are explained in more detail in Sec. 2.2.

The combined  $SU(2)_L \otimes U(1)_Y$  group symmetry corresponds to the electromagnetic and to the weak force [10–14]. They introduce the electrical charge  $Q$  and the weak isospin  $T$ , respectively. Those two forces further combine at high energies on the order of 100 GeV via the electroweak unification. The latter entails the unification of the electromagnetic force [10–17], originating from quantum electrodynamics (QED), and the weak force, stemming from quantum flavour dynamics (QFD). The charge of the combined interaction is the weak hypercharge  $Y$ . However, the bosons  $W^{0,1,2}$  and  $B^0$  predicted by this symmetry are massless which is contradicting the massive bosons observed. This requires a process that introduces massive bosons.

## 2. The Standard Model of Particle Physics

The Brout-Englert-Higgs Mechanism [18–20] explains how bosons, and fermions as well, acquire their mass. The theory introduces a quantum Higgs-field. Every particle interacting with the field becomes massive. At very high energies, this Higgs-field is symmetric. However, at energies which we observe under normal conditions, this symmetry is spontaneously broken [21]. One of the results of this is a non-zero vacuum expectation value. This vacuum energy introduces additional terms to the Lagrangian, spoiling the electroweak symmetry. These additional terms mix the four massless bosons of the theoretical electroweak symmetry ( $W^1, W^2, W^3, B^0$ ) to generate the massive bosons  $W^\pm, Z^0$  and the massless boson  $\gamma$ . The mass originates from the introduced interactions between the  $W^{1,2}$ -bosons and the Higgs-field. The  $W$ -bosons are associated with the  $SU(2)_L$  symmetry group. The remaining two bosons are associated with the  $U(1)_Y$  symmetry group.

Furthermore, one of the degrees of freedom introduced by the Higgs-field is not defined via the Brout-Englert-Higgs mechanism. It manifests as the scalar Higgs boson  $H$  [18].

## 2.2. Particles

The SM describes particles as elemental units of matter or forces. All particles have various properties such as mass, electrical charge and spin. Furthermore, each particle has an anti-particle that has the same properties as the complementary particle, but its electrical charge is inverted. These anti-particles are typically denoted with a bar on top of the particle identifier. Uncharged particles do not have a complementary anti-particle. At its core, the SM categorises particles into two main groups by their spin. These groups are fermions and bosons. In the following the focus will be on particles, since anti-particles have identical properties but with an inverted electrical charge.

### Fermions

In the SM, fermions are one of the two fundamental classes of elementary particles, the other being bosons. Fermions are particles with half-integer spin. They are the basic constituents of matter and obey the Fermi-Dirac statistics as well as the Pauli exclusion principle. Moreover, Fermions can be subdivided into quarks and leptons as described in the following sections.

### Quarks

Quarks are elementary particles that interact via the strong nuclear force, as well as the weak and electromagnetic forces which were introduced in Sec. 2.1. Quarks are classified

into three generations. In each generation, there is an up-type and a down-type quark defined by their electrical charge and weak isospin. The up-type quarks have a positive electrical charge of  $+2/3$  times the elementary charge  $e$  and a weak isospin of  $+1/2$ . The down-type quarks have an electrical charge of  $-1/3 e$  and a weak isospin of  $-1/2$ .

For each generation, a weak isospin doublet can be defined containing two out of six quark flavours. Each quark flavour has unique properties. The first generation's doublet contains the up ( $u$ ) and down ( $d'$ ) quark. The second generation consists of the charm ( $c$ ) and strange ( $s'$ ) and the final and third generation includes the top ( $t$ ) and bottom ( $b'$ ) quark.

Note that down-type quarks denoted as  $q'$  in the isospin doublets describe the weak eigenstate while  $q$  describes the mass eigenstate. These two eigenstates are related via the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix [22, 23]. The matrix describes the probability amplitude  $V_{ij}$  for the transition from quark flavour  $i$  to quark flavour  $j$  under the weak interaction. Furthermore, it describes the relation of the weak and mass eigenstates of down-type quarks

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.1)$$

The entries in the CKM matrix are not predicted by theory. Hence, the amplitudes must be measured experimentally. The current best fitted values [24] are the following:

$$\begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} = \begin{bmatrix} 0.97373 \pm 0.00031 & 0.2243 \pm 0.02243 & 0.00382 \pm 0.00020 \\ 0.221 \pm 0.221 & 0.975 \pm 0.006 & 0.0408 \pm 0.0014 \\ 0.0086 \pm 0.0002 & 0.0415 \pm 0.0009 & 1.014 \pm 0.029 \end{bmatrix} \quad (2.2)$$

With increasing quark generation, the masses of the quarks also increases. The current best measurements of the masses for the up-type quarks and the down-type mass eigenstates [24] yield

$$\begin{aligned} m_u &= 2.16_{-0.26}^{+0.49} \text{ MeV}, & m_d &= 24.68_{-0.17}^{+0.48} \text{ MeV}, \\ m_c &= 1.27_{-0.02}^{+0.02} \text{ GeV}, & m_s &= 93.4_{-3.4}^{+8.6} \text{ MeV}, \\ m_t &= 172.69_{-0.30}^{+0.30} \text{ GeV}, & m_b &= 4.18_{-0.02}^{+0.03} \text{ GeV}. \end{aligned} \quad (2.3)$$

## 2. The Standard Model of Particle Physics

Quarks also carry a property known as colour charge [5, 6]. It is analogous to the electric charge but associated with the strong nuclear force. However, unlike electric charge, which comes in positive and negative values, colour has six possible states: red, green, blue and their complements anti-red, anti-green, and anti-blue. Quarks always combine to a state in which all three colours are combined, or any colour combined with its complement [6]. These states are called colourless. This is known as colour confinement and is responsible for the fact that isolated quarks are never observed in nature.

### The Top Quark

The top quark will be explained in more detail since its of particular interest for this study. The discovery of the top quark took place in 1995 at FERMILAB through the efforts of the DØ [25] and CDF [26] experiments that both discovered the top quark independently. The top quark is acknowledged as the most massive among all quarks. Producing top quarks needs significant energies owing to high mass of the top quark. Such energies are attainable in hadron colliders. The top quark decays before it hadronises due to its brief lifetime is approximately  $\tau_t \approx 5 \cdot 10^{-25}$  s [24]. It decays into a  $W$ -boson and a down-type quark which then undergoes hadronisation.

The likelihood of decaying into a specific down-type quark is determined by the previously mentioned CKM matrix in Eq. 2.2. Comparing the values of  $V_{td}$  and  $V_{ts}$  to  $V_{tb}$  from Eq. 2.2 shows the predominant top decay mode to  $b$  quarks.

The  $W$ -boson, the second particle from the top decay, also decays further. As explained in the following sections, it can decay either into a charged lepton & neutrino pair or into a quark and anti-quark pair. Thus, the resulting final state from a  $t$  quark decay include one quark and a lepton or three quark, depending on the  $W$ -boson decay mode.

### $b$ tagging

To precisely identify  $t$  quarks, an accurate algorithm to distinguish  $b$  is necessary. The reason for this is the top quarks predominant decay mode into  $b$  quarks which was explained previously. To achieve this,  $b$  tagging is introduced [27, 28].

The accurate identification of works by exploiting the properties of the  $b$  quark. It has a relatively large mass and a long lifetime. The latter is explained by its suppressed decay modes. Eq. 2.2 shows that the  $V_{tb}$  element is the highest values for  $b$  quark row. Thus, the  $b$  flavour couples predominantly to the  $t$  flavour and decay modes into other flavours are suppressed. However, the  $t$  is significantly heavier than the  $b$  quark as seen in Eq. 2.3. Hence, the decay into a top quark is also strongly suppressed.

Due to the long lifetime of the  $b$  quark, it can travel a measurable distance before decaying. Thus, its decay products do not originate from the primary vertex of the collision. Instead, reconstructing these decay products yields a displaced vertex position which is typically referred as ‘secondary vertex’. This reconstruction combined with the hard momentum spectra of the  $b$  quarks decay products yields a sufficiently accurate  $b$  tagging which reaches an efficiency of 70% [29] at the ATLAS experiment. This approach can also be applied, to a lesser extent, to charm tagging.

## Leptons

Leptons are the second group of fermions. They do not experience the strong nuclear force and interact via the weak and electromagnetic forces. Analogue to the quarks, there are three generations of leptons: electron, muon and tau [1]. These are typically denoted as  $e$ ,  $\mu$  and  $\tau$ , respectively.

Each generation consists of a weak isospin doublet containing an uncharged neutrino  $\nu_l$  and its respective charged lepton  $l$  [1]. Charged leptons have an electrical charge of  $-1e$  and a weak isospin of  $-1/2$ . Furthermore, the charged leptons are massive. Their masses increase with higher generations as current measurements [24] show

$$\begin{aligned} m_e &= 0.511 \pm 0.001 \text{ MeV}, \\ m_\mu &= 105.66 \pm 0.01 \text{ MeV}, \\ m_\tau &= 1776.86 \pm 0.12 \text{ MeV}. \end{aligned} \tag{2.4}$$

On the contrary, the uncharged neutrinos have no electrical charge [24], a weak isospin of  $+1/2$  and are predicted to be massless by the SM [1]. However, the later prediction is conflicting with some observed phenomena that are explained in Sec. 2.3.

## Bosons

Bosons are the second group of particles in the SM. Unlike the Fermions, bosons have integer values of spin and obey Bose-Einstein statistics. Bosons can be subdivided into gauge bosons that mediate three of the four fundamental forces and the Higgs boson.

### Gauge Bosons

In the SM, there are 4 different gauge boson types [1] that determine the interactions between particles. While many of the gauge boson properties differ, they all have spin 1.

## 2. The Standard Model of Particle Physics

The photon ( $\gamma$ ) is the massless mediator of the electromagnetic force [11, 13, 14]. It itself is not electrical charged and thus, couples to positively and negatively charged particles equally. The behaviour of photons is determined by QED as introduced in Sec. 2.1.

A gluon ( $g$ ) is a type of massless gauge bosons that is responsible for the mediation of the strong nuclear force [5]. They couple to the colour charge of quarks. And unlike the photon, gluons themselves also carry colour charges which are exchange during a strong interaction. There are eight different possible colour combination for gluons due to the rules of QCD. Since they carry the charge they couple to, gluons are able to couple to other gluons. These processes are called self-couplings.

Lastly, there are  $W^\pm$ - and  $Z^0$ -bosons. These are the mediators of the weak nuclear force. In contrast to the other gauge bosons, these are massive due to symmetry breaking in the Brout-Englert-Higgs Mechanism which is explained in more detail in Sec. 2.1. Their masses are measured to be [24]

$$\begin{aligned} m_W &= 80.38 \pm 0.01 \text{ GeV}, \\ m_Z &= 91.19 \pm 0.01 \text{ GeV}. \end{aligned} \tag{2.5}$$

$W^\pm$ -bosons have an electrical charge of  $\pm 1e$ . Moreover, they have a weak isospin. Its third component of the weak isospin is  $\pm 1$ , respectively. Fermions interacting with the  $W$ -bosons undergo a flavour transition. Charged leptons convert to lepton neutrinos and vice versa. Quarks also change their flavour. Their transition rate from one quark to the other is determined by the previously discussed CKM matrix. Couplings with the  $W$ -bosons are CP violating. The  $Z^0$ -boson, as indicated by its superscript, has no electrical charge and its third component of the weak isospin is 0.

### Higgs Boson

The Higgs boson ( $H$ -boson) was the last particle of the SM discovered in 2012 by the combined efforts of the ATLAS [30] and CMS [31] experiments. The Higgs boson's detection gave strong support to the Brout-Englert-Higgs mechanism discussed in Sec. 2.1. It is electrically neutral and has spin 0, making it the only scalar boson. Moreover, it is the second-heaviest particle ever detected with the combined measurement of ATLAS and CMS [32] resulting in a mass of

$$m_H = 125.09 \pm 0.24 \text{ GeV}. \tag{2.6}$$

Furthermore, it represents the quantum manifestation of the Higgs field and thus, couples to mass of particles. Its production and decay branching ratios are strongly dependent on the mass of the interacting particle. The higher the mass, the stronger is the Higgs coupling resulting in higher branching ratios. The reason for this is the linearity between the Yukawa coupling strength and the (fermion) masses. Hence, the most common production modes are in combination with either  $t$  quarks or  $W$ -/ $Z$ -bosons.

For the decay, the most likely modes are decays into pairs of  $b$  quarks or  $W$ - and  $Z$ -bosons. While the Higgs coupling of the  $t$  is stronger than the  $b$ , the decay into a  $t$  quark pair is strongly suppressed, due to its significantly higher mass  $2m_t > m_H$ . Couplings with the Higgs field (thus also with the  $H$ -boson) are described as Yukawa couplings that also generate the mass of fermions and bosons [18, 20]. Since the  $H$ -boson is massive, it can couple to itself, similar to the gluon. This gives rise to Higgs self-coupling.

## 2.3. Beyond the Standard Model

Despite the incredible range of phenomena that the SM can predict and explain for example the existence and properties of the  $W$ -,  $Z$ - and  $H$ -bosons as well as its precision, certain phenomena contradict its fundamental explanations.

### Neutrino Oscillation

One of these problematic observations is neutrino oscillation [33, 34]. It describes how neutrinos change their flavour while propagating through space. The first indication was found by the Homestake Experiment in 1968 [35]. It detected solar neutrinos using a large tank of fluid and observed a neutrino deficit. In 1998, the Super-Kamiokande experiment studied atmospheric neutrinos and discovered a changing electron/muon neutrinos flux during the runtime [36].

The oscillation originates from the separation of the flavour eigenstates of the neutrinos ( $\nu_e, \nu_\mu, \nu_\tau$ ) and its mass eigenstates ( $\nu_1, \nu_2, \nu_3$ ). If these eigenstates are not equal, then neutrinos could change their flavour. The probability of a neutrino of a given flavour transitioning into another flavour depends on the mixing angles and mass differences between the mass eigenstates. The observed oscillation necessitates at least two massive neutrinos. Contradictory, the SM predicts all neutrinos as massless which would result in their flavour and mass eigenstates to be identical. Consequently, the oscillation would be prohibited and thus, the SM fails to account for neutrino oscillations.

## 2. The Standard Model of Particle Physics

### Gravity

Another obvious problem with the SM is the absence of gravity. While the force of gravity is observed, the SM has no boson responsible for mediating gravity. Possible solutions include the introduction of a hypothetical massless and uncharged graviton [37] or string theory [38]. Without an explanation of gravity, the SM can not achieve to be a unified theory of everything.

### Further Contradictions

There are more contradicting observations that need explanations beyond the SM. These examples contain, but are not limited to, dark matter [39–41], a type of invisible mass needed for stellar observations, prediction of vacuum energy [42], a deviation between QFT and cosmological predictions and the hierarchy problem [43], a relatively low  $H$ -boson mass that necessitates fine-tuning. For each of the problems, several explanations beyond the Standard Model such as supersymmetry [44] or modified gravity [45] are possible.

## 2.4. Higgs-Associated Top Quark Pair Production

The targeted event topology of this thesis is the  $H$ -boson associated top quark pair production in which the  $H$ -boson decays into two  $W$ -bosons. It consists of a fully hadronic decaying  $t\bar{t}$  pair with an additional  $H$ -boson. The latter decays into two  $W$ -bosons which then further decay semi-leptonically, which means that one  $W$ -boson decays hadronically into two quarks while the other decays leptonically into a lepton neutrino pair. In the following, these events will be referred to as semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events. The corresponding Feynman diagram of the targeted topology can be seen in Fig 2.1.

The rarity of the semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events makes its measurement difficult. Fig. 2.2 depicts a summary of several event cross-section measured by ATLAS. It shows the dominating  $t\bar{t}$  background with significantly higher a cross-section of  $(834 \pm 47)$  pb [46] compared to  $t\bar{t}H$  which reaches  $(507 \pm 61)$  fb [47]. These theoretical cross-section refers to all  $t$  quark and  $H$ -boson decay modes. Hence, the  $t\bar{t}H$  events are three orders of magnitude rarer than  $t\bar{t}$  events. Other background processes such as  $tW$ ,  $t\bar{t}Z$  and  $t\bar{t}W$  have cross-sections higher or close to  $t\bar{t}H$ .

However, since this study focuses on a particular  $H$ -boson decay mode, the branching ratio of the Higgs decay must be taken into account. For the previously stated  $H$ -boson mass  $M_H$ , the theoretical branching ratio of  $H \rightarrow WW$  is the second highest at  $(21.5 \pm$

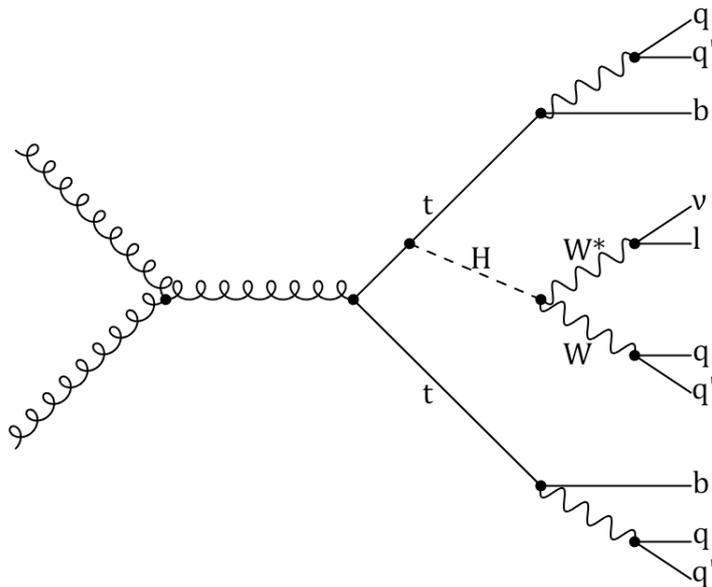


Figure 2.1.: Feynman diagram of the targeted  $t\bar{t}(H \rightarrow WW^*)$  event topology.

1.0) % [48]. Furthermore, the  $W$ -boson decay ratios are  $(67.41 \pm 0.27)$  % hadronically and  $(32.58 \pm 0.47)$  % leptonically [24]. Multiplying the  $t\bar{t}H$  cross-section with these decay ratios and taking the  $W$ -boson permutations into account yields a final cross-section of  $(56.8 \pm 12.3)$  fb for the targeted semileptonic  $t\bar{t}(H \rightarrow WW^*)$  event topology. This results in more than 10000 expected  $t\bar{t}$  events per semileptonic  $t\bar{t}(H \rightarrow WW^*)$  event. Taking other background processes into account, semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events become even more difficult to measure.

To reduce some of the backgrounds, the semileptonic decay mode is chosen. The implemented single lepton limitation helps suppressing  $W$ +jets and QCD multijet backgrounds.

In the targeted topology, the  $H$ -boson production is expected to be a resonance process. Hence, only on-shell  $H$ -boson masses are considered. Due to the insufficient mass of the  $H$ -boson to create two on-shell  $W$ -bosons, one of the two  $W$ -bosons must be off-shell. This is denoted by the asterisk superscript  $W^*$ . The analysis restricts to events where the on-shell  $W$ -boson is decaying hadronically. This simplifies the event reconstruction since the invariant mass of the two selected decay jets should match the SM  $W$ -boson mass  $M_W$  as close as possible. Thus, it can be used to suppress some of the main backgrounds due to the constraint on the decay jet properties. However, it also necessitates that the off-shell  $W^*$ -boson decays leptonically which complicates the leptonic reconstruction.

## 2. The Standard Model of Particle Physics

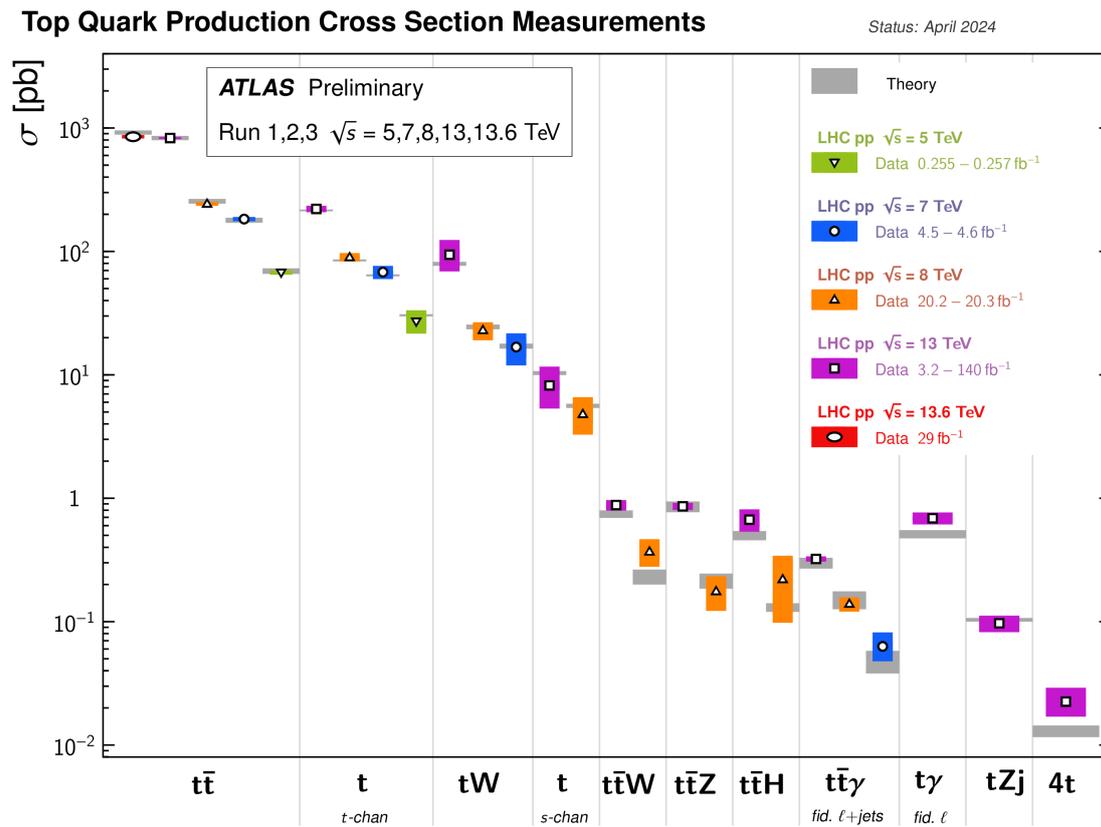


Figure 2.2.: Overview of several  $t$  quark related production cross-section measurements with theoretical predictions calculated at next to leading order or higher [49].

## 3. Experimental Setup

To collect data necessary for this study, a high-energy particle collider is needed. Additionally, a setup for signal detection and suitable reconstruction algorithms are required. This chapter introduces the Large Hadron Collider [50, 51] in Sec. 3.1 which is used for collecting the data. It is one of the most important particle accelerators worldwide and includes many experiments such as the ATLAS experiment [52]. The ATLAS experiment provides the data for the project and is discussed in great detail in Sec. 3.2. Other experiments are briefly described in Sec. 3.3.

### 3.1. Large Hadron Collider

The Large Hadron Collider (LHC) [50, 51] is the world's largest and most powerful particle accelerator. It's located underground near Geneva, Switzerland, crossing the border between Switzerland and France. Specifically, it's situated at the European Organization for Nuclear Research (CERN). The collider spans 27 km in circumference and reaches depths of up to 175 m. Along the collider, there are several detectors build which will be explored in Sec. 3.2 and Sec. 3.3. Fig 3.1 shows an aerial view of the regions above the LHC demonstrating its size and showing the position of the different experiments.

It is used to create and accelerate bunches of protons close to the speed of light. Then a collision of these bunches is induced. During such a collision, partons from the protons interact and give rise to a large multitude of processes. By studying the debris produced by these collisions, physicists can gain insights into fundamental questions about the nature of the universe. Moreover, predictions of the SM introduced in Ch. 2 can be tested with such collisions.

The LHC is operating for certain time periods with long shutdowns in between. The first operation period, Run I, was active from November 2009 until February 2013. It started with a beam energy of 1.2 TeV which was then later increased to 3.5 TeV. In the last year it reached a beam energy of 4 TeV. During the operation,  $5.6 \text{ fb}^{-1}$  had been accumulated by ATLAS and CMS. Then Long Shutdown 1 followed Run I.

In April 2015, Run II started. The second run reached beam energies of 6.5 TeV and

### 3. Experimental Setup



Figure 3.1.: Aerial view above the Large Hadron Collider showing its size. The different experiments are marked along the LHC course. © CERN, Maximilien Brice

lasted until December 2018. It acquired an integrated production of  $160 \text{ fb}^{-1}$ . The Long Shutdown 2 then held on until 2022.

The currently active Run III launched in April 2022 with a centre-of-mass energy of 13.6 TeV, 6.8 TeV per beam. It is planned to operate until the coming Long Shutdown 3 planned for 2026.

## 3.2. The ATLAS Experiment

ATLAS (A Toroidal LHC ApparatuS) [52] is one of the four main experiments at the LHC. Its detector is designed as a general-purpose detector to explore a wide range of physics phenomena [53, 54]. The detector began its operation along the start-up of the LHC.

It spans a height of 25 m and a length of 44 m. It is build almost hermetically around the LHC beam pipe as a toroidal shape which can be seen in Fig 3.2. Constructing the detector so close to the interaction point allows measuring the maximum possible phase space.

The coordinate system used in ATLAS is important for understanding the detector and is introduced in the following section. After that, several components and layers of the detector are explained in more detail.

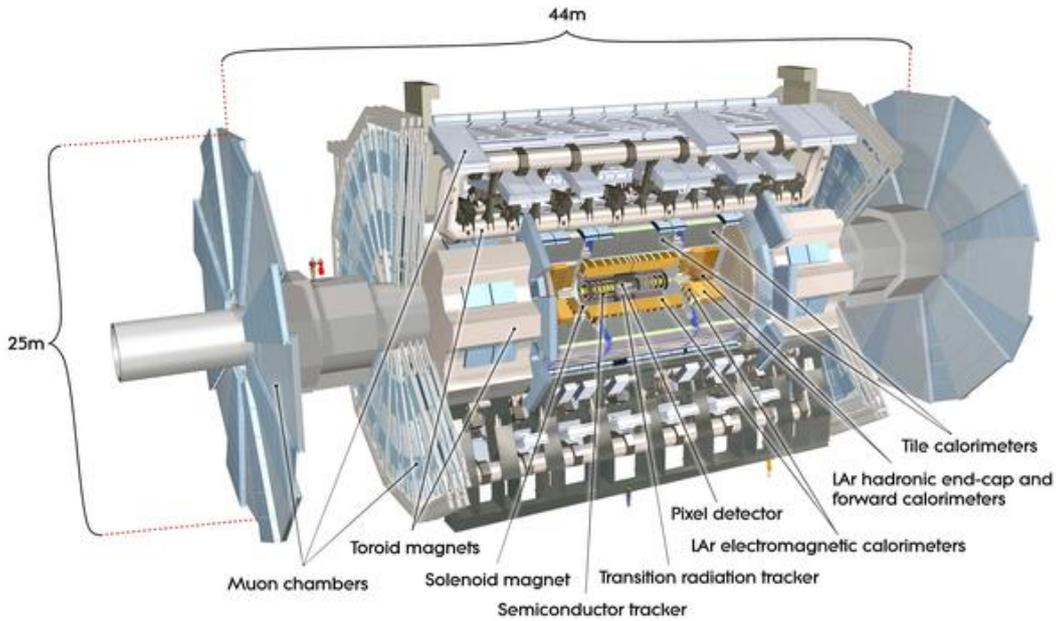


Figure 3.2.: Overview of the ATLAS detector showing its size and structure. Moreover, every important component is labelled. ©CERN

## Coordinate System for ATLAS

The ATLAS detector is built as a cylinder. Thus, a fitting coordinate system must be defined. Every position in the detector is defined by the  $z$  position parallel to the beam axis, the azimuth angle  $\phi$  along the detector and the pseudo-rapidity  $\eta$  with respect to the collision point of the detector. The latter quantity  $\eta$  is defined by

$$\eta = -\ln \left( \tan \frac{\theta}{2} \right) \quad (3.1)$$

with  $\theta$  the polar angle. Fig 3.3 illustrates a 2 dimensional projection of the cylinder at a fixed distance from the centre.

To properly work with spatial separations in this coordinate system, the variable

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

is introduced. It represents a cone around each point in the coordinate system as shown in Fig 3.3. This should not be confused with  $R$  which is typically the radius from the beam pipe. Hence, expressions such as  $R\phi$  correspond to the radial scaled azimuth angle and is used in classifying detector resolutions.

### 3. Experimental Setup

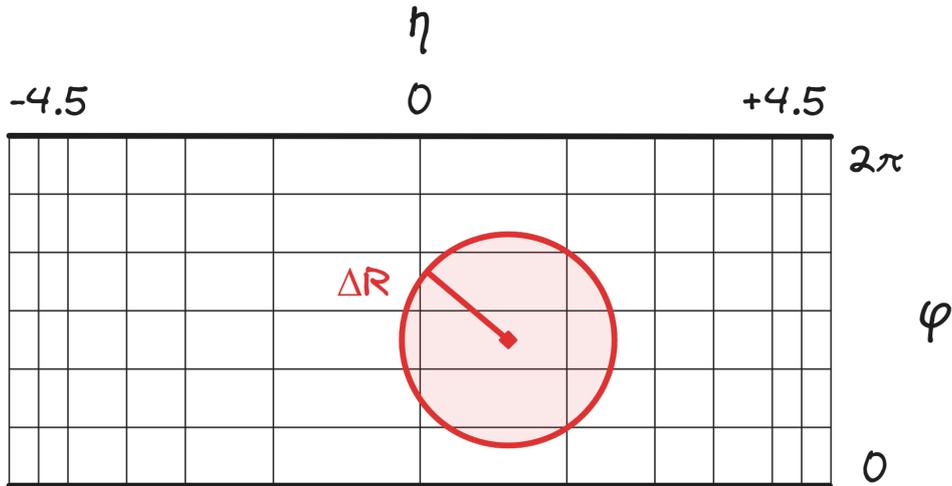


Figure 3.3.: Sketch visualising the ATLAS coordinate system using the azimuth angle  $\phi$  and pseudo-rapidity  $\eta$ . It represents the rolled out detector cylinder. Moreover, it shows how a spatial cone defined by  $\Delta R$  would look like.

## Components of the ATLAS Detector

The detector of the ATLAS experiment is of particular interest for this project. The following sections will cover its elements including the detector modules and magnetic systems in great detail. An overview of the detector is shown in Fig 3.2. The individual layers of the detector are further visualised in Fig 3.4.

### Magnetic System

For the measurement of particle trajectories and their curvature, significant magnetic fields are required. For that reason, ATLAS implemented superconducting magnet system build from Aluminium, Copper and Niobium-Titanium alloy.

It includes a central solenoid magnet surrounding the Inner Detector. It provides a central field of 2 T [53]. Since the central solenoid is placed right before the calorimeter, this magnet is build as thin as possible.

Furthermore, three large air-core toroids generate an additional magnetic field for the outer muon spectrometer. Each consisting of eight symmetrical coils placed radially around the beam axis [53]. These magnet systems are cooled using liquid Helium at 4.5 K [53].

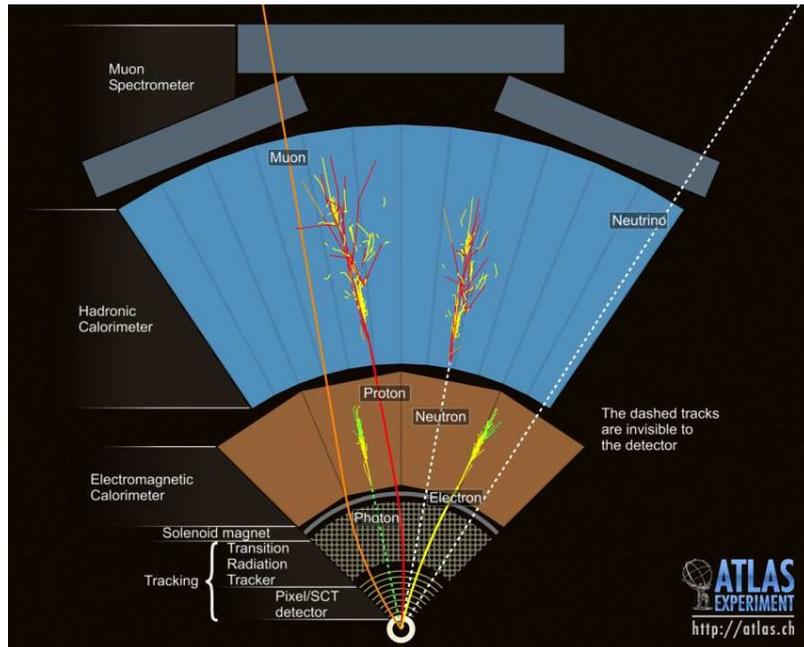


Figure 3.4.: Cross-section of the concentric layers within the ATLAS detector. ©CERN

### The Inner Detector

The Inner Detector [55, 56] is a combined system of high-resolution detectors close to the beam pipe and continuous tracking elements at outer radii. Its task is to measure the particles' momentum. As mentioned before, the complete Inner Detector is surrounded by the 2 T central solenoid magnet. Charged particles are deflected inside the detector, due to the Lorentz force. By precisely measuring the curvature, momentum measurements and particle track reconstruction are achieved.

Its detector modules are placed either in the barrel or the end-cap region. The first region is arranged on concentric cylinders surrounding the beam axis and typically covers the low  $\eta$  region. The later regions detectors are installed on disks perpendicular to the beam axis and cover the remaining space for a full coverage up to  $\eta \pm 2.5$ .

The innermost detector tiles are semiconductor pixel detectors [53, 57]. The first barrel layer is only 4 cm from the beam pipe, covers an area of  $0.2 \text{ m}^2$  and measures the complete space up to  $\eta = \pm 2.5$ . It reaches a resolution of  $\sigma_{R\phi} = 12 \text{ }\mu\text{m}$  and  $\sigma_z = 66 \text{ }\mu\text{m}$ . This layer is followed by two more barrel layers that cover  $\eta = \pm 1.7$  and reach the same resolution. To compensate this lower  $\eta$  acceptance, 5 end-cap disks are placed in each side. These cover the remaining space up to  $\eta = \pm 2.5$ . Each module is only 21.4 mm wide and 62.4 mm long containing 61440 pixels controlled by 16 chips. Every pixel detector has its individual circuit including buffering to store data until the trigger decision. While the precision and radiation resistance of these pixel detectors are remarkable, they also

### 3. Experimental Setup

introduce some material before the calorimeters and are expensive. Hence, the amount of pixel detectors that can be included in the detector is limited.

Directly surrounding the pixel modules, the silicon strips are introduced [53, 58]. It is subdivided into four barrel layers for the  $\eta$ -range up to  $\eta = \pm 1.4$  and 9 end-cap wheels on each side for the remaining space up to  $\eta = \pm 2.5$ . Compared to the pixel modules, they achieve a slightly lower precision of  $\sigma_{R\phi} = 16 \mu\text{m}$  and  $\sigma_{z/R} = 580 \mu\text{m}$ . Each detector is  $6.28 \times 6.40 \text{ cm}^2$  in size and includes 768 strips. Moreover, each module consists of four of these silicon detectors.

Straw tube trackers around the previous modules allow for a lot of measured tracking points [53, 59]. In the barrel regions, these are placed parallel to the beam axis. In contrast to the end-caps where these are aligned perpendicular to the beam axis. Previously, these were filled with a Xenon gas mixture. However, due some leakage this Xenon was changed to Argon to reduce costs. While these achieve significantly worse resolutions of  $\sigma = 170 \mu\text{m}$ , they can cover a greater area at lower cost and less material per tracking point. Hence, they strongly contribute the precise momentum measurements.

#### The Calorimeters

The calorimeters [53, 60] are used to detect charged particles by absorbing them, allowing for a measurement of their deposited energy. The physical process of showering which allow such measurements was previously described in Ch. 2. ATLAS' calorimeter system is subdivided into four parts: electromagnetic calorimeter, hadronic barrel calorimeter, hadronic end-cap calorimeter and hadronic forward calorimeter.

The first of these calorimeters, the electromagnetic calorimeter [53], is built from lead, liquid Argon and kapton electrodes. It covers the pseudorapidity range up to  $\eta = \pm 3.2$  while the barrel region can detect up to  $\eta = \pm 1.475$ . It is the innermost calorimeter and is primarily responsible for detecting photons and electrons.

The hadronic calorimeter [53] is build around the electromagnetic calorimeter and is build from plastic scintillator plates and iron absorber in the barrel region and from copper, lead and liquid Argon in the end-cap regions. This separation is done to optimise the detector for the varying requirements of the great  $\eta$  range. The barrel region reaches up to  $\eta = \pm 1.7$  and the end-caps up to  $\eta = \pm 3.2$ . Its thickness is one of the most important parameters for the calorimeter, since it has to contain the hadronic showers and thus, minimise punch-through into the adjacent Muon Spectrometer.

The forward hadronic calorimeter [53] is implemented to cover the strongly boosted region of  $3.1 < |\eta| = 4.9$ . It as well is build from a lead liquid Argon mixture with additional electrodes from copper or a tungsten alloy.

## The Muon Spectrometer

The outermost measurement part of the ATLAS detector is the Muon Spectrometer [53, 61]. It bends the muon tracks by utilising the previously explained air-core magnets. The magnetic field is mostly orthogonal to the muon trajectories. In the region up to  $\eta = \pm 1.0$  the large barrel toroid magnet provides the needed magnetic field. Within the acceptance range of  $1.4 \leq |\eta| \leq 2.7$  two smaller end-cap magnets bend the tracks. Between of these regions is the transition region where the deflection of the muon trajectories is done by the magnet field of both magnetic systems. The muon spectrometer consists of four parts.

The first part is the monitored drift tube chamber which make up  $800 \text{ m}^3$  in volume. It consists of 30 mm diameter aluminium tubes operating with a mixture of 93% argon and 7%  $\text{CO}_2$ . The single-wire resolution is  $\sim 80 \mu\text{m}$ . To increase the overall resolution, multi-layer pairs of multiple tubes are utilised.

The New Small Wheel project replaced the original Cathode Strip Chambers [61–63]. It was installed after Run II and uses Micro-Mesh Gaseous Structure as well as small-strip thin gap chambers. The former being a detector consisting of a planar electrode and a thin steel mesh forming a drift chamber. It is filled with the same argon- $\text{CO}_2$  mixture as the first drift tube chamber. Its spatial resolution reaches  $73 \mu\text{m}$ . The later thin gap chambers consist of a gold-plated tungsten wire grid placed between two electrodes. The electrodes are made from a carbon-epoxy mixture. The detector module reaches a spatial resolution of  $100 \mu\text{m}$  at the rate of up to  $20 \text{ kHz/cm}^2$ .

The resistive plate chambers [53, 61, 62] are another gaseous detector. It uses a non-flammable  $\text{C}_2\text{H}_2\text{F}_4$   $\text{SF}_6$  gas mixture. Again, the electrodes of the detection chamber are made from carbon. The panels themselves are made polystyrene placed between aluminium sheets. They achieve a spatial resolution of 1 cm and a temporal resolution of 1 ns.

The thin gap chambers [53, 61] are the last chamber of the muon spectrometer. It is a proportional chamber using a flammable gas mixture of 55%  $\text{CO}_2$  and 45% n-pentane. Even though it needs additional safety precautions, it is less sensitive to mechanical deformations, yields nearly Gaussian pulse height distributions and has a small dependence on the incident angle. The distance of both electrodes are 2.8 mm.

## Trigger System

The aforementioned detector systems are able to detect and record significant amounts of data. The initial bunch-crossing rate of up to 40 MHz result in an interaction rate of approximately 1 GHz [53] which is too much to save. Approximately only 0.1 kHz of

### 3. Experimental Setup

data is suited for permanent storage. Thus, a very fast selection must be defined and implemented that discards most of the events to a manageable level. For that purpose, an online multi-level trigger system is introduced. Each trigger level refines the selection done from the previous level.

The first level (L1) trigger [53, 64] is a hardware-based system and makes the initial selection. It is subdivided into the L1 calorimeter trigger and the L1 muon trigger. The L1 calorimeter trigger utilises signals from the calorimeter in the Cluster and Jet/Energy-sum Processor. The Cluster Processor identifies potential  $e$ ,  $\gamma$ ,  $\tau$  candidates. This is supplemented by the Jet/Energy-sum Processor which identifies jet candidates and calculates the total and missing energy. The L1 muon trigger, uses signals from the muon spectrometer to identify  $\mu$  candidates. The LVL1 decision is formed by the Central Trigger Processor and is based on the signal of the previous two trigger systems as well as some additional detector sub-systems. The maximum rate for the L1 trigger is up to 100 kHz, within a decision latency of up to 2.5  $\mu$ s. The selected events of the L1 trigger are temporary saved into readout drivers and then readout buffers. Furthermore, the L1 trigger identifies regions of interest, which are further investigated by the second trigger stage.

The event data stays on the readout buffers until the high level (HLT) trigger [53, 64] either rejects or accepts them. The second trigger level is a software-based system and utilises the aforementioned regions of interest. The HLT trigger uses only the necessary information but has access to the full resolution information from the detectors. Its average rate reaches 1.2 kHz which corresponds to approximately 1.2 GB/s data. Its latency is variable but in the order of magnitude 100  $\mu$ s. Events that pass both triggers are then transferred to an offline event filter system that performs the final selection.

### 3.3. Other Experiments

Besides the discussed ATLAS experiment, there are three more main experiments at the LHC: CMS, ALICE and LHCb. This section will briefly mention each of these experiments and compare them to the ATLAS experiment.

The CMS (Compact Muon Solenoid) [65] detector is a general-purpose similar to the ATLAS detector. As its name suggest, a particular focus is set on the precise measurement of high energy muons over a wide range of angles and momenta. Its main goals include muon mass resolution of around 1% at 100 GeV and the ability to determine the charge of muons with momentum  $p < 1$  TeV. Other main goals include the reconstruction of charged particles, efficient tagging of  $\tau$  leptons and  $b$  quarks, precise missing transverse energies

and isolation of photons and leptons at high luminosities. CMS started its operation with the commissioning of the LHC.

The ALICE (A Large Ion Collider Experiment) [66] focuses on heavy ion collisions. It is designed to measure the physics of strongly interacting matter and the quark-gluon plasma at extreme energies and temperatures. It consists of a central barrel which covers polar angles from  $45^\circ$  to  $135^\circ$  and a forward muon spectrometer covering  $2^\circ$  to  $9^\circ$ . While the ALICE physics program includes the study of light atoms, the main study utilises heavy nuclei Lead-Lead collisions. Its first collisions was recorded in November 2010.

LHCb (Large Hadron Collider beauty) [67] is dedicated to study particles that contain  $b$  quarks, which sometimes are referred as beauty quarks. Unlike the other experiments, its detector is build as a single-arm spectrometer. It covers a forward angle of approximately  $0.5^\circ$  to  $17^\circ$ . The reason lies in LHCb's focus on detecting  $b$  hadrons which are predominantly produced in forward and backward cones. The LHCb experiment started the operation at the same time as ATLAS and CMS.



# 4. Truth Matching

Training a neural network as explained in Ch. 5 requires sufficient training samples. To generate such training samples, a truth matching algorithm is needed. The algorithm takes the truth and reconstructed information of Monte Carlo (MC) simulations as inputs, and outputs a jet-parton assignment that can be used for training. The following section gives a high-level explanation of the algorithm.

## 4.1. Algorithm

To produce the training data, the algorithm begins by scanning the input MC samples. Each sample includes truth information of the partons before interacting with the detector. The truth information also includes the true identification of particles as well as their decay products. Furthermore, the samples include the reconstructed information at detector level. Here, the particles and their properties are simulated as they would appear in a real dataset. Pairing the reconstructed and truth information of each event, allows generating training samples that uses the reconstructed information with true labels.

The algorithm analyses each event by iterating through all truth-level final state partons. Around each of these partons, the algorithm searches possible jet matches. All jets within a spatial cone around a parton are considered as its candidate matches. Here, the cone is defined by  $\Delta R \leq 0.4$  as introduced in Sec. 3.2. For the targeted  $t\bar{t}(H \rightarrow WW^*)$  semileptonic final state, 8 quarks (3 per  $t$  quark + 2 for the hadronic  $W$ -boson from the  $H$ -boson) have to be matched.

If there is no jet within the cone of a parton, then that parton is not matched. If exactly one candidate is found inside the cone, the jet is saved as the final match for the jet-parton assignment. For cases with more than one candidate match, the jet closest to the particle in  $\Delta R$  space is chosen as the final match.

Each matched jet is made unavailable for other partons. Hence, each jet-parton assignment is unique which is required for training the neural network. Afterwards, the algorithm proceeds to combine the matched particles. The four-vectors of the jets are combined up to get the four-vector of the expected resonance particles.

## 4.2. **Optimisation**

Some technical optimisations are implemented to increase the processing speed. The algorithm utilises bit-shifting operations. Each jet gets an integer value assigned. Each expected parton corresponds to one of the bits in the integers bit-representation. Bit values of 1 mean the jet is potentially matched to the corresponding parton, while 0 means that the parton is not included within the jets  $\Delta R$  cone. Each parton has a fixed position in the bit-representation. Hence, checking a potential parton-jet match is done efficiently by bit-operations. Additional minor optimisations such as hashmap implementations and lazy evaluations further increase the efficiency. All optimisations combined cut the processing time by over 60%.

# 5. SPA-Net: Symmetry Preserving Attention Networks

This chapter introduces the transformer neural network *SPA-Net* [68, 69], which is used for the hadronic jet-parton assignment during event reconstruction. Sec. 5.1 first explains the classical approach for solving such assignment problems by using the  $\chi^2$  minimisation technique. Then, Sec. 5.2 presents the implemented modern neural network approach with the main advantages listed in Sec. 5.3

## 5.1. Classical $\chi^2$ Method

Traditionally, to determine the jet-parton assignment of an event, a simple  $\chi^2$  minimisation approach [70] is used. The method tests every possible combination of jets in the event, calculating the masses  $m$  of the resonance particles ( $t, W, H$ ) and then compares it to the SM prediction  $M$ . The method tries to find the best matching mass for all resonance particles. Mathematically, for every assignment the following value is calculated

$$\chi^2 = \sum_{i,j} \frac{(m_{qi} - M_j)^2}{\sigma_j^2}$$

with  $\sigma$  being the SM mass uncertainty. The lowest  $\chi^2$  value corresponds to the best matching jets  $i$  of each resonance particle  $j$ . Thus, it is expected to be the most accurate jet-parton assignment.

However, this approach is not suited for the targeted event topology of this study. For the targeted  $t\bar{t}(H \rightarrow WW^*)$  channel, 8 jets are expected in the semileptonic final state. This leads to high numbers of permutations which slows down the calculation significantly. Even when introducing  $b$ -tagging and requiring the method to only swap  $b$ -tagged jets with other  $b$ -tagged jets, the algorithm would have  $6! = 720$  possible light jet permutations per event. Hence, this method is not implemented and instead, a neural network is trained.

## 5.2. Modern DNN Method

In this study, the transformer-based deep neural network architecture *SPA-Net* (Symmetry Preserving Attention Networks) [68, 69] is used for predicting the jet-parton assignment. *SPA-Net* was originally developed for fully hadronic  $t\bar{t}$  analysis [71]. However, the implemented network is extended to general event topologies [68] and thus, can be used for the targeted  $t\bar{t}(H \rightarrow WW^*)$  topology.

Its task is to predict a jet-parton assignment for all jets in the targeted semileptonic final state. This hadronic assignment can then be combined with the leptonic estimation using *Neutrino Weighting* as explained in Ch. 6 resulting in a full event reconstruction.

### Symmetries

One of *SPA-Net*'s features is the utilisation of symmetries in the event topology. These symmetries are implemented in the loss function during training. The loss does not change under predefined symmetrical permutations. Thus, the neural network treats changes within these permutations as equal. Symmetries allow changes in the assignment that do not change the underlying physics. These symmetries are classified into two groups: jet and particle symmetries.

The jet symmetry describes that the labelling of certain particles is arbitrary and thus, changing their labels does not change the resulting physical behaviour. In the  $t\bar{t}(H \rightarrow WW^*)$  topology, this symmetry is present in the hadronic  $W$ -boson decays which produce two quarks  $q\bar{q}'$ . Under the assumption that the  $W$ -boson couples evenly to both quarks, their labeling  $q$  and  $q'$  are arbitrary and can be swapped without any change in the physical reconstruction.

The particle symmetry depicts the symmetry with respect to the particle origin. Particles that originate from a quark or its respective anti-quark are symmetric, when disregarding their charge. For the targeted  $t\bar{t}(H \rightarrow WW^*)$  topology, the symmetry is applicable in the decays of either the  $t$  or  $\bar{t}$  quark. While there might be subtle differences in the hadronic charge distributions of the  $t$  and  $\bar{t}$  quarks, we do not expect to be able to resolve them. Hence, the jet assignment treats both particles equally.

By considering these symmetries, the number of possible permutations is reduced and the processing speed increases. The run-time is reduced to  $\mathcal{O}(N_{\text{jets}}^3)$  compared to the baseline  $\chi^2$  method with a expected run-time of  $\mathcal{O}(N_{\text{jets}}^C)$  [68]. Here,  $N_{\text{jets}}$  are the number of jets within a given event and  $C = 8$  is the number of final state partons or  $C' = 6$  when separating  $b$ -tagged jets.

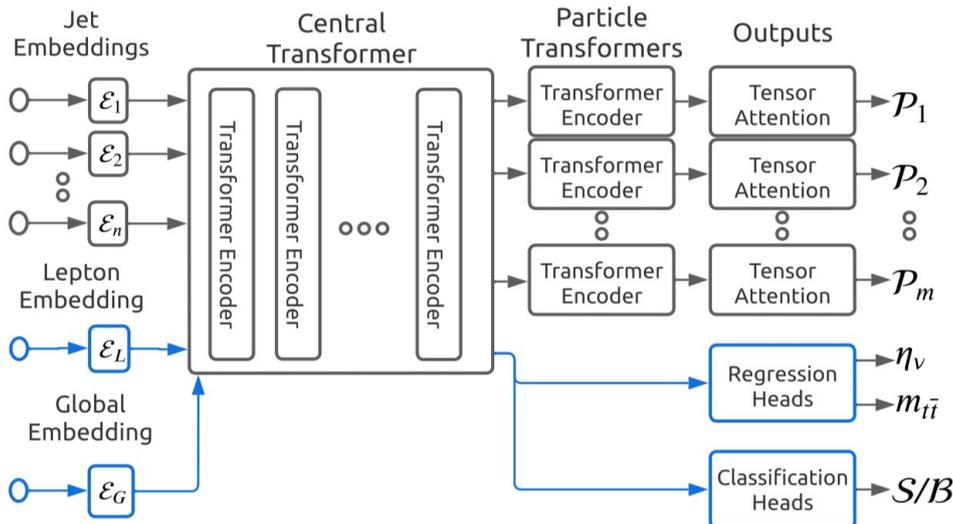


Figure 5.1.: Overview of *SPA-Net*'s architecture including its four regions and the additional embeddings for additional lepton and global information such as the missing transverse energy [69].

## Architecture

*SPA-Net*'s network architecture is visualised in Fig 5.1. It can be subdivided into four components [68]: input embeddings, central transformer, particle transformers and self-attention output layer.

During training, each prediction is split into three steps. First, the amount of valid jet assignment permutations is reduced by splitting up events into sub-structures which are predefined by the event topology. For the targeted event topology, these sub-structures are both  $t$  quark decays and the hadronic  $W$ -boson decay from the  $H$ -boson. Then each jet-parton assignment sub-problem is solved by applying a 'Symmetric Tensor Attention' [68] layer, producing a single tensor. Each tensor is a multi-dimensional array containing an entry for each sub-assignment including the probability that any particular combination is the correct sub-assignment. Lastly, the tensors are combined into a final jet-parton prediction by calculating the combined symmetric loss [68].

Particle symmetries which correspond to permutations in a particles origin are implemented during the combination of these sub-assignments. The symmetric cross entropy loss during training is not sensitive to changes within these particle symmetries. This behaviour is achieved by allowing the network to train any equivalent assignment as long as the overall loss is minimised.

Jet symmetries are included in permutation groups which are used by the Symmetric Tensor Self-Attention layer [68]. The attention refers to *SPA-Net*'s ability to dynamically

## 5. SPA-Net: Symmetry Preserving Attention Networks

weight different parts of the input to improve the overall performance by focusing on relevant features.

During inference, the combined symmetric loss is not calculated. Instead, *SPA-Net* iteratively assigns jets to the expected particles  $p$  by selecting the most likely assignment from each output tensor  $\mathcal{P}_p$ . The used permutations are bijective to ensure a unique jet-parton assignment [68]. Hence, in cases where one jet is assigned more than once, the highest probability score is selected and the others are evaluated again.

### 5.3. Advantages

In comparison to previous techniques, *SPA-Net* offers several advantages. Firstly, *SPA-Net* is agnostic to the number of input objects. Hence, training and prediction can be done using samples containing any jet multiplicity. The architecture also allows predicting and training on partial events by splitting up the individual particle assignments. This advantage is vital for this study as the training samples contain a lot of partial events. Secondly, all predicted jet matches are unique.

Furthermore, each prediction is also rated using an assignment probability that represents *SPA-Net's* confidence that a particular prediction is correct. Additionally, a detection probability is calculated that estimates whether a particular particle is detected for a given input event.

Lastly, the expected symmetries due to the event topology are predefined which leads to significantly higher processing speed by simplifying the training and prediction task. While training is an additional step compared to the classical approach, the prediction is significantly faster and thus, needs less computing resources.

For training *SPA-Net* to the targeted topology, a suitable training sample is created using  $t\bar{t}H$ ,  $t\bar{t}$  and  $t\bar{t} + X$  events. The needed truth-matching algorithm for creating such training samples is explained in Ch. 4.

## 6. Neutrino Weighting

*Neutrino Weighting* is a technique to reconstruct events including neutrinos in their final state. It was first implemented by the DØ collaboration in dileptonic  $t\bar{t}$  events [72]. The kinematic reconstruction with neutrinos in the final state are underconstrained, due to the unmeasured neutrino particle. Hence, the reconstruction relies on additional constraints due to set assumptions. With these constraints, *Neutrino Weighting* yields an estimation of the unknown parameters by calculating the most likely combination of them.

While the original *Neutrino Weighting* was developed for events with two neutrinos in the final state, this study adapts and modifies the original approach to reconstruct the leptonically decaying  $W^*$ -boson in the  $H \rightarrow WW^*$  subprocess [73]. For the targeted event topology, the free parameters are the mass of the off-shell  $W^*$ -boson  $m_{W^*}$  and the pseudo-rapidity of the neutrino  $\nu$   $\eta_\nu$ .

The algorithm outputs an estimation for these free parameters and an output weight  $w$  which is in the range of  $[0, 1]$ . Higher values describe a better agreement between the estimated values and the observed missing event kinematics. This weight can be used in the event selection to identify the targeted semileptonic  $t\bar{t}(H \rightarrow WW^*)$  topology.

### 6.1. Parameter Estimation

To apply the *Neutrino Weighting* algorithm, some assumptions need to be made to constrain the otherwise underconstrained  $H \rightarrow WW^*$  sub-system.

Firstly, the  $H$ -boson is assumed to be on-shell. The used MC samples sets  $m_H^{\text{MC}}$  to 125.0 GeV. When using real data, the measured mass as given in Eq. 2.6 is used. Secondly, the mass of the on-shell  $W$ -boson is fixed at the current best measurements as stated in Eq. 2.5. Lastly, the on-shell  $W$ -boson is further assumed to decay hadronically. Then, the  $H \rightarrow WW^*$  sub-system can be reconstructed by estimating the leptonic  $W^*$ -boson and combining it with the hadronic  $W$ -boson from *SPA-Net*'s prediction as explained before in Ch. 5.

To find the best estimation for the  $W^*$ -boson, several combinations of the off-shell mass  $m_{W^*}$  and the pseudo-rapidity of the neutrino  $\eta_\nu$  are sampled using a grid-based search.

## 6. Neutrino Weighting

The grid samples 100  $\eta_\nu$  points within  $[-3, 3]$  and 100  $m_{W^*}$  points within  $[0, 50]$  GeV. All points are distributed equidistantly. The grid size and granularity was varied optimised, the results are given in Sec. 8.3.

The mass of the  $W^*$ -boson is limited to approximately 50 GeV due to the on-shell constraints. While the pseudorapidity  $\eta$  is not physically limited, the range  $[-3, 3]$  is selected to ensure optimal detector performance.

Due to the assumptions, the  $H \rightarrow WW^*$  sub-system is fully constrained [73] and the neutrino transverse momentum  $p_T^\nu$  can be calculated. This reconstructed neutrino momentum  $p_T^\nu$  can be compared to the observed missing transverse momentum  $p_T^{\text{miss}}$ . For that, the transverse momenta are split into their two components  $p_x$  and  $p_y$  with  $p_T^2 = p_x^2 + p_y^2$ . The weight of one solution is calculated as

$$w = \exp \frac{(p_x^\nu - p_x^{\text{miss}})^2}{\sigma_x^2} \cdot \exp \frac{(p_y^\nu - p_y^{\text{miss}})^2}{\sigma_y^2}. \quad (6.1)$$

The variable  $\sigma_{x/y}$  is the experimental resolution of the missing momentum  $p_{x,y}^{\text{miss}}$  and only serves to scale the weight. Thus, it has no physical impact on the result if taken equal in  $x$  and  $y$  [73]. The implementation sets  $\sigma_{x/y} = 10$  GeV. The value is chosen since it is the worst resolution at the ATLAS detector for events with transverse energy up to approximately 100 GeV [74]. This is expected to be appropriate for the targeted event topology.

For each event the sampled  $m_{W^*}$  and  $\eta_\nu$  can be visualised together with the weight of each grid point to form a distribution as seen in Fig 6.1. The maximum of this distribution corresponds to the most likely combination of the unknown parameters. White regions are kinematically forbidden and thus, no solutions can be found for these regions.

## 6.2. Optimisation

Since the targeted  $t\bar{t}(H \rightarrow WW^*)$  process is rare, it is crucial to maximise the number of reconstructable events. To increase these statistics as well as the accuracy of the estimated values for  $m_{W^*}$  and  $\eta_\nu$ , two optimisations are implemented. These are explained briefly in the following sections. The improvement due to these optimisations is given in Sec. 8.3.

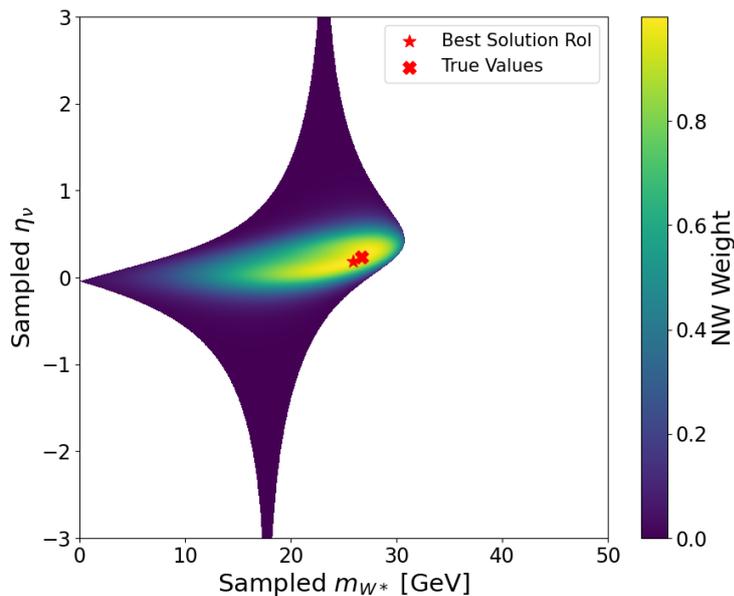


Figure 6.1.: Example of a sampled weight distribution for a cherry-picked event. The highest weighted solution in the region of interest (star) and the true values (cross) are marked.

## Higgs Mass Smearing

Some events do not yield any solution during the initial grid-based search. Hence, there is no estimation for  $m_{W^*}$  or  $\eta_\nu$ . To decrease the amount of events with no solution, the assumed Higgs boson mass  $m_H$  can be altered slightly. This approach is valid since reconstruction effects can change the measured energies of jets, leptons or the missing transverse energy. Then, the precise  $H$ -boson mass might not yield a solution. Therefore, it is viable to also test masses close to the expected  $m_H$  to counteract these reconstruction effects.

If an event has no solution, the assumed  $H$ -boson mass is varied in a range of  $\pm 1$  GeV using 0.1 GeV steps around  $m_H$ . The grid-based search is repeated for each alternative assumption  $m'_H$ . If any solution is found using the alternative Higgs mass  $m'_H$ , it is used to estimate  $\eta_\nu$  and  $m_{W^*}$ .

This additional step increases the processing time for the *Neutrino Weighting* algorithm by 10-50%. The specific time increase is dependent on the sample composition. Due to this additional step, around 1% of events that would be rejected become reconstructable.

## Regions of Interest

As an additional improvement, the *Neutrino Weighting* algorithm defines regions of interest. These regions are used to refine the estimated solution.

If the algorithm finds at least one solution during the grid-based search, an additional finer grid-based search is started around the highest weighted solution. This second grid is the local region of interest. Within the region of interest, the steps for  $m_{W^*}$  and  $\nu_\eta$  are more granular to find an even higher weighted, local solution. The final solution is the highest weighted estimation within the region of interest.

If an event has two solution initially only the highest weighted solution is searched again using the region of interest. Hence, if the second highest weighted solution was the actual solution but was not estimated at precise enough values, the true solution is not found.

The region of interest increase the processing time by  $\sim 15\%$  independent of the sample composition.

# 7. Event Selection

This chapter covers the event selection strategy used to separate the targeted semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events from other background processes. Sec. 7.1 introduces the event topology in greater detail and explains the assumptions used, while Sec. 7.2 defines the separation power and showcases some of the investigated variables. The behaviour of several event topologies is compared with respect to these variables. Lastly, sensitive variables are used in Sec. 7.3 for the final region definition.

## 7.1. Objects of the Target Event Topology

The targeted semileptonic  $t\bar{t}(H \rightarrow WW^*)$  event topology is introduced in Sec. 2.4. Its final state consists of eight jets and a single lepton-neutrino pair. Six of these jets originate from the  $t\bar{t}$  decays, while the remaining two jets stem from the  $H$ -boson decay sub-system. Among the eight jets, exactly two are tagged as  $b$ -jets which originate from the  $t$  quark decay ( $t \rightarrow Wb$ ). For the purposes of event reconstruction, three restrictions are made to improve the background rejection.

Firstly, the  $t\bar{t}$  system is expected to decay fully hadronically. This implies that the observed lepton does not originate from the  $t$  quark decay but instead is produced in the  $H$ -boson decay chain. Only then can the implemented *Neutrino Weighting* technique be applied properly.

Secondly,  $H$ -boson is produced on-shell due to the resonance production. Hence, the mass is expected to equal the value given in Eq. 2.6. The simulated MC samples were tested and the on-shell  $H$ -boson production dominates. Given the energy constraints, the  $H$ -boson cannot decay into two on-shell  $W$ -bosons. Therefore, one of the  $W$ -bosons must be off-shell.

Lastly, the off-shell  $W^*$ -boson is expected to decay leptonically. This is crucial because, without the on-shell constraint on the hadronic  $W$ -boson, it would be almost impossible to assign the jets to the  $W$ -boson, as any combination of jets could potentially reconstruct an off-shell  $W^*$ -boson with an indeterminate mass. This would make background reduction even more difficult. Therefore, the two jets assigned to the on-shell  $W$ -boson are expected

## 7. Event Selection

to have an combined invariant mass close to the SM  $W$ -boson mass close to the value stated in Eq. 2.5.

As explained in Sec. 2.4, the primary background arises from  $t\bar{t}$  events, which have a significantly larger production cross-section compared to the signal  $t\bar{t}(H \rightarrow WW^*)$  events.  $t\bar{t}(H \rightarrow bb)$  and  $t\bar{t}(H \rightarrow \tau\tau)$  events also contribute significantly, due to their comparatively high  $H$ -boson decay ratio of  $(58 \pm 12)\%$  and  $(6.3 \pm 1.7)\%$ , respectively [48]. Additional backgrounds can arise from  $t\bar{t}(H \rightarrow cc)$  events and hadronic or dileptonic  $t\bar{t}(H \rightarrow WW^*)$  decays. These background processes are considered during the variable investigation and in the final region definition.

All events require at least 5 jets. Moreover, the used samples implement a single lepton trigger that requires at least one lepton for selecting an event. This is combined with a lepton restriction which rejects more than one lepton. Therefore, every event contains exactly one lepton and thus, the choice of lepton is unambiguous. This preselection is made to reduce the main  $t\bar{t}$  background and to ensure a viable lepton for the *Neutrino Weighting* algorithm.

## 7.2. Separation Power

To compare the performance of certain variables to separate signal and background, the corresponding distributions are shown together. Here, the signal is defined to only include semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events while the background includes everything else. Naturally, the background contains the majority of all events. For a fair comparison, both signal and background are normalised to unity and binned using the same edges. The choice of variables is physically motivated by identifying measurable differences in the signal and background processes.

Furthermore, the separation power  $\mathcal{S}$  can be calculated using

$$\mathcal{S} = \frac{1}{2} \sum_i^{\text{bins}} \frac{(s_i - b_i)^2}{s_i + b_i}. \quad (7.1)$$

For each bin  $i$ , the variables  $s_i$  and  $b_i$  equal the normalised signal and background fractions, respectively. Since the separation power  $\mathcal{S}$  is dependent on the binning, the absolute value has no physical meaning. However, when choosing the same binning, the separation power of several variables can be compared to identify the best performing ones.

Table 7.1.: Summary of the investigated variables for the event selection sorted by their calculated separation powers  $\mathcal{S}$ .

Investigated Variable	$\mathcal{S}$
Total Jet Multiplicity	21.9 %
<i>SPA-Net</i> $t$ Assignment Probability	21.2 %
<i>SPA-Net</i> $W_{\text{had}}$ Assignment Probability	9.0 %
<i>Neutrino Weighting</i> Reconstructable	1.7 %
Missing Transverse Energy $E_T^{\text{miss}}$	3.0 %
Lepton Energy $E_{\text{lep}}$	1.2 %
<i>Neutrino Weighting</i> Output Weight	0.4 %
$b$ -Jet Multiplicity	0.2 %

## Jet Multiplicities

Fig 7.1 shows two distributions of events per total jet and  $b$ -jet multiplicity, respectively. This analysis uses the 85 % working point for the  $b$ -tagging.

As expected, the distribution of  $t\bar{t}(H \rightarrow bb)$  events is peaking at higher  $b$ -jet multiplicity than any other distribution. While all distributions peak at 2  $b$ -jets, the  $t\bar{t}(H \rightarrow bb)$  event distribution peaks at 3  $b$ -jets due to the two additional  $b$ -quarks in the final state. The  $b$ -jet separation power  $\mathcal{S} = 0.2\%$  is relatively low. The reason for this is that most background processes expect the same number of  $b$ -jets. Only the  $t\bar{t}(H \rightarrow bb)$  background can be separated efficiently from the signal using this variable.

The total number of jets shows similar distributions for all events.  $t\bar{t}$  events slightly favours low jet multiplicities which can be explained by this final state containing only 6 jets. The targeted semileptonic  $t\bar{t}(H \rightarrow WW^*)$  final state contains 8 jets. Moreover, hadronic  $t\bar{t}(H \rightarrow WW^*)$  events contain 10 jets in their final state. Hence, the number of jets can be used to separate some of the backgrounds from the signal and each other. However, due to additional jets or jets lost during reconstruction, the distributions do not clearly peak at their expected jet multiplicities. The separation power  $\mathcal{S} = 21.9\%$  is the highest value of all investigated variables. This supports the importance of this variable for separating signal from background.

## Lepton Energies

The distributions in Fig 7.2 show the behaviour of different event topologies for the lepton energies  $E_{\text{lep}}$  and missing transverse energies  $E_T^{\text{miss}}$ .

For the missing transverse energy  $E_T^{\text{miss}}$ , the targeted semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events peak approximately 10 GeV lower than most other processes. In these events, the missing energy is expected to originate from the neutrino decay within the  $H \rightarrow WW^*$  sub-system.

## 7. Event Selection

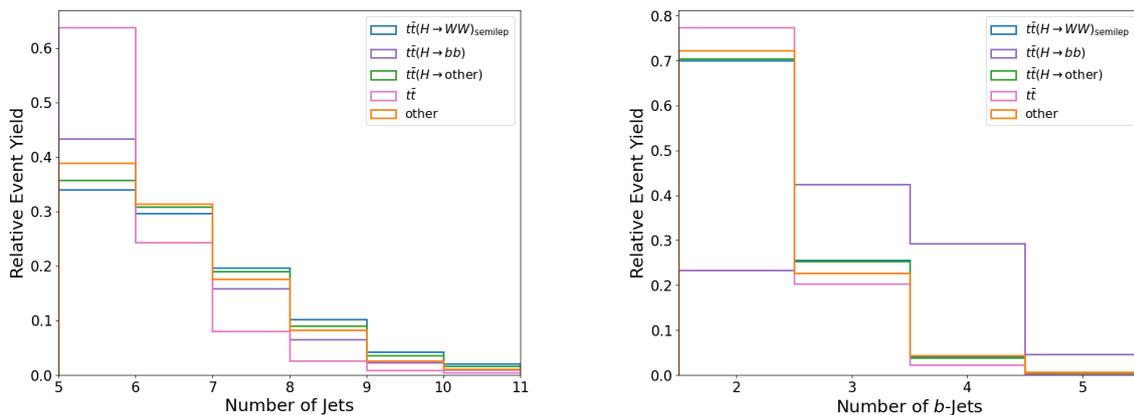


Figure 7.1.: Distributions of different event topologies for (left) the total jet multiplicity and (right) the  $b$ -jet multiplicity per event.

Hence,  $E_T^{\text{miss}}$  is limited by the energy of the leptonic  $W^*$ -boson decay. The separation power reaches  $\mathcal{S} = 3.0\%$  which corresponds to a similar sensitivity to signal and background.

The energies of the lepton  $E_{\text{lep}}$  show a behaviour similar to  $E_T^{\text{miss}}$ . The targeted semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events generally have lower lepton energies. However, here the dileptonic  $t\bar{t}(H \rightarrow WW^*)$  events behave very similar to the semileptonic decays. This similarity is expected since the production of the leptons is identical for both the semileptonic and dileptonic decay mode. Here, the separation power is  $\mathcal{S} = 1.2\%$  which is caused by the similar behaviour of the dileptonic  $t\bar{t}(H \rightarrow WW^*)$  and  $t\bar{t}(H \rightarrow \tau\tau)$  backgrounds. The dileptonic  $t\bar{t}(H \rightarrow WW^*)$  can be separated by  $E_T^{\text{miss}}$  since these events contain two neutrinos and thus, generally higher  $E_T^{\text{miss}}$ .

## SPA-Net Output Probabilities

Fig 7.3 includes two distributions of *SPA-Net*'s output probabilities. These probabilities are calculated during the jet assignment prediction. Events with high probabilities are more likely to be correctly matched. Both the  $t$  and the  $W_{\text{had}}$  assignment probabilities show different behaviours for the event topologies.

High values of the  $t$  assignment probability generally has slightly higher semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events contributions. The signal peaks as most backgrounds just below 0.4. The  $t\bar{t}$  nbackground scores lower with its peak at 0.3. The average of both  $t$  quark assignment probability is  $\mathcal{S} = 21.2\%$  which is the second highest separation power of all considered variables.

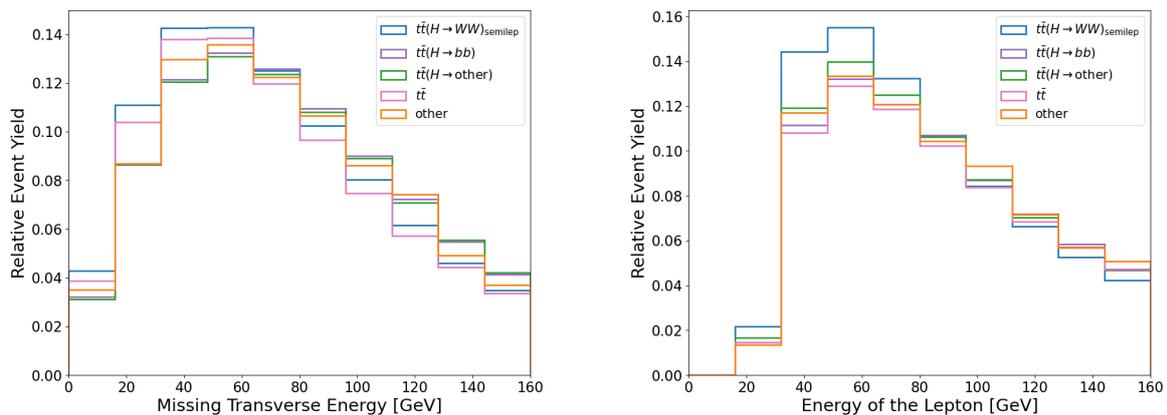


Figure 7.2.: Distributions of different event topologies for (left) the missing energy per event and (right) the energy of the identified lepton.

Similarly, the  $W_{\text{had.}}$  assignment probability also separates the different events topologies. Here,  $t\bar{t}$  peak at around 0.15. Again, higher assignment probability contain higher contributions of  $t\bar{t}(H \rightarrow WW^*)$  and other  $t\bar{t}H$  processes with their peak at 0.25. This variable reaches a separation power of  $\mathcal{S} = 9.0\%$ . Compared to  $\mathcal{S}$  of the  $t$  assignment probability, this variable is less sensitive to signal.

Due to their relatively large separation power of both probabilities, these variables will be used in the region definition.

## Neutrino Weighting Output

After estimating the unknown parameters of the event structure, the *Neutrino Weighting* algorithm outputs a weight. Visualising the distribution of the weights in Fig. 7.4 shows its separating power. Here the weight can be used to identify reconstructable and non-reconstructable events as well as quantify the reconstruction quality. Events with at least one solution are labeled as reconstructable and higher weights correspond to better estimations.

Separating events with and without *Neutrino Weighting* solution yields a separation power  $\mathcal{S} = 1.7\%$ . Looking at the weight distribution of all reconstructable yields a separation of  $\mathcal{S} = 0.4\%$ . In comparison to the *SPA-Net* output probabilities, the separation power is lower and thus, does not separate the signal as efficiently. However, the *Neutrino Weighting* weight can still be used to separate signal events, when combining it with *SPA-Net*'s probabilities.

## 7. Event Selection

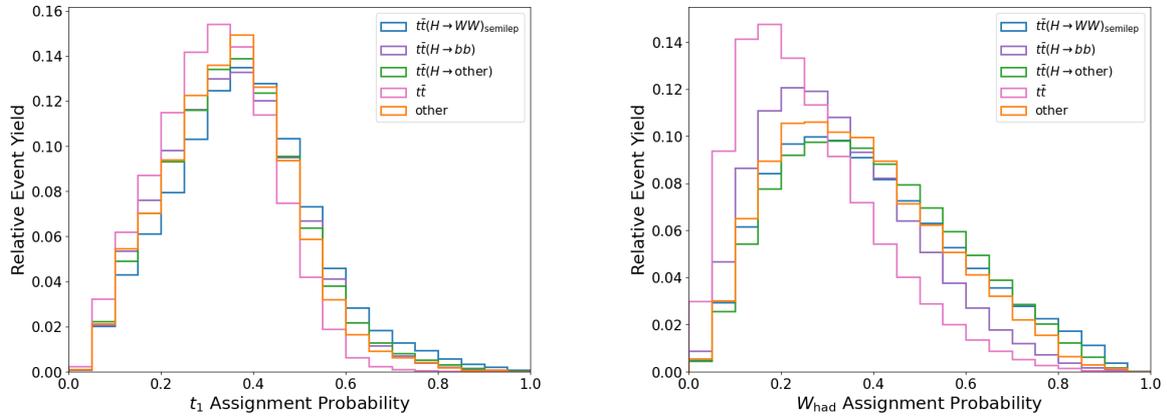


Figure 7.3.: Distributions of different event topologies for *SPA-Net*'s assignment probabilities of (left) one of the  $t$  quarks and (right) the hadronic  $W$ -boson from the Higgs decay.

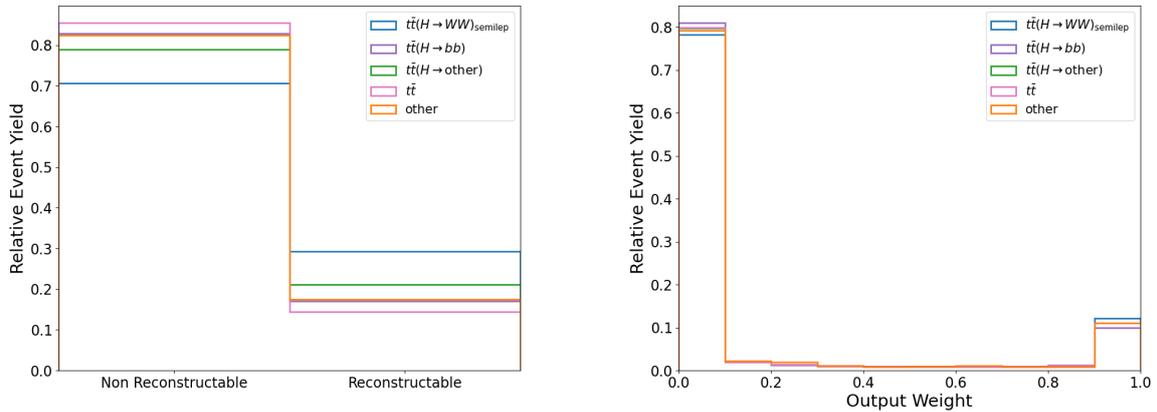


Figure 7.4.: Distributions of different event topologies for *Neutrino Weighting* output weight while (left) including events without any solution and (right) excluding events without any solution. Events with no solution are visualised as negative weight.

Table 7.2.: Overview of the defined signal and control regions and their corresponding cuts. The conditions of the  $CR_{t\bar{t}}$  marked with \* are disjunctive, meaning only one them must be fulfilled.

	$SR$	$CR_{t\bar{t}}$		$CR_{H\rightarrow bb}$
Total Jet Multiplicity	$\geq 7$	$< 8$	$\geq 7$	—
$b$ -Jet Multiplicity	$\leq 3$	$= 2$	—	$\geq 3$
<i>Neutrino Weighting</i> Reconstructable	yes	any	any	any
<i>Neutrino Weighting</i> Output Weight	$\geq 0.2$	—	$< 0.2^*$	—
<i>SPA-Net</i> $t$ Assignment Probability	$\geq 0.4$	—	$< 0.4^*$	—
<i>SPA-Net</i> $W_{\text{had}}$ Assignment Probability	$\geq 0.4$	—	$< 0.4^*$	—
<i>SPA-Net</i> $W_{\text{had}}$ Detection Probability	$\geq 0.7$	—	$< 0.7^*$	—

### 7.3. Region Definition

To separate signal and background, the overall phase space is subdivided in signal regions  $SR$  and control regions  $CR$ . If defined properly, the signal region is expected to include a relatively high purity of the targeted semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events while excluding as much of the background contamination as possible. The control regions contain the remaining events with high background yields. Each control region targets one of multiple background processes. Only events where the *Neutrino Weighting* algorithm yields a prediction for the unknown neutrino and leptonic  $W^*$ -boson are reconstructable. Hence, only the events with a *Neutrino Weighting* solution are selected in the following. Given the rarity of the  $t\bar{t}(H \rightarrow WW^*)$  process, it is important that the union of all regions covers the entire remaining phase space to avoid discarding events.

The  $SR$  and  $CR$  are constructed to be orthogonal. The choice of variables is driven by their respective separation power. The values at which the regions selects events are determined by using a grid search. An overview of all region definitions is given in Tab. 7.2 with the resulting region compositions given in Tab. 7.3. Here, the  $t$  assignment probability refers to the average of both assigned  $t$  quarks. Furthermore, the ratios  $S/B$  and  $S/\sqrt{B}$  are calculated to show the signal purity and its statistical significance. Here,  $S$  and  $B$  correspond to the signal and background events, respectively.

#### Variable Combination

To increase the performance of the outputs from *SPA-Net* and the *Neutrino Weighting* algorithm, the distributions are combined in several 2-dimensional histograms. In Fig. 7.5 the histogram for the average  $t$  assignment probability and the *Neutrino Weighting* weight is shown for signal and background events. These figures show that sig-

## 7. Event Selection

Table 7.3.: Summary of all signal and control region compositions.

	$SR$	$CR_{t\bar{t}}$	$CR_{H\rightarrow bb}$
Event Count	$140 \pm 12$	$4764069 \pm 2183$	$1346247 \pm 1161$
Relative Yield in %	$0.0023 \pm 0.0001$	$77.97 \pm 0.03$	$22.03 \pm 0.02$
$S/B$	$0.008 \pm 0.001$	$0.000 \pm 0.001$	$0.000 \pm 0.001$
$S/\sqrt{B}$	$0.10 \pm 0.01$	$0.22 \pm 0.01$	$0.23 \pm 0.01$
Contribution Signal	0.80 %	0.01 %	0.02 %
Contribution $t\bar{t}(H \rightarrow bb)$	0.12 %	0.03 %	0.27 %
Contribution $t\bar{t}(H \rightarrow \text{other})$	0.24 %	0.03 %	0.06 %
Contribution $t\bar{t}$	98.84 %	99.93 %	99.65 %

nal events have slightly higher contributions in region with high weight and assignment probabilities. Hence, these regions are of particular interest. Due to the low statistics, the region selection must contain more conservative selections to ensure enough events per region.

### Signal Region

One  $SR$  is defined to separate as much signal events from background contamination as possible. As listed in Tab. 7.2, the signal region mainly combines the outputs from *SPA-Net* and *Neutrino Weighting*. In addition to that, a conservative selection on the jet multiplicity is applied to retain enough signal events which expect at least seven jets in the final state of which three or less are  $b$ -tagged.

The selection results in a background rejection of 99.998 % while retaining a signal acceptance of 0.12 % for  $SR$ . After applying the proper event weights derived from SM prediction, the signal region contains 140 events with a signal purity  $S/B$  of 0.80 %. However, due to the low event count in  $SR$ , the significance  $S/\sqrt{B}$  is low compared to the  $CRs$ . Further increases in purity reduced the statistical significance of  $SR$  drastically. The  $SR$  is dominated by the  $t\bar{t}$  background with over 98 % contribution due to its high cross-section.

### Control Regions

Two additional control regions,  $CR_{t\bar{t}}$  and  $CR_{H\rightarrow bb}$ , are defined to separate  $t\bar{t}$  and  $t\bar{t}(H \rightarrow bb)$  events, respectively.

For the  $CR_{H\rightarrow bb}$  control region, the  $b$ -jet multiplicity is of particular interest. Since it is the only background process at leading order  $t\bar{t}$  and  $t\bar{t}H$  with a higher expected number of  $b$ -quarks in the final state, it can be separated by cutting events with at least three  $b$ -jets

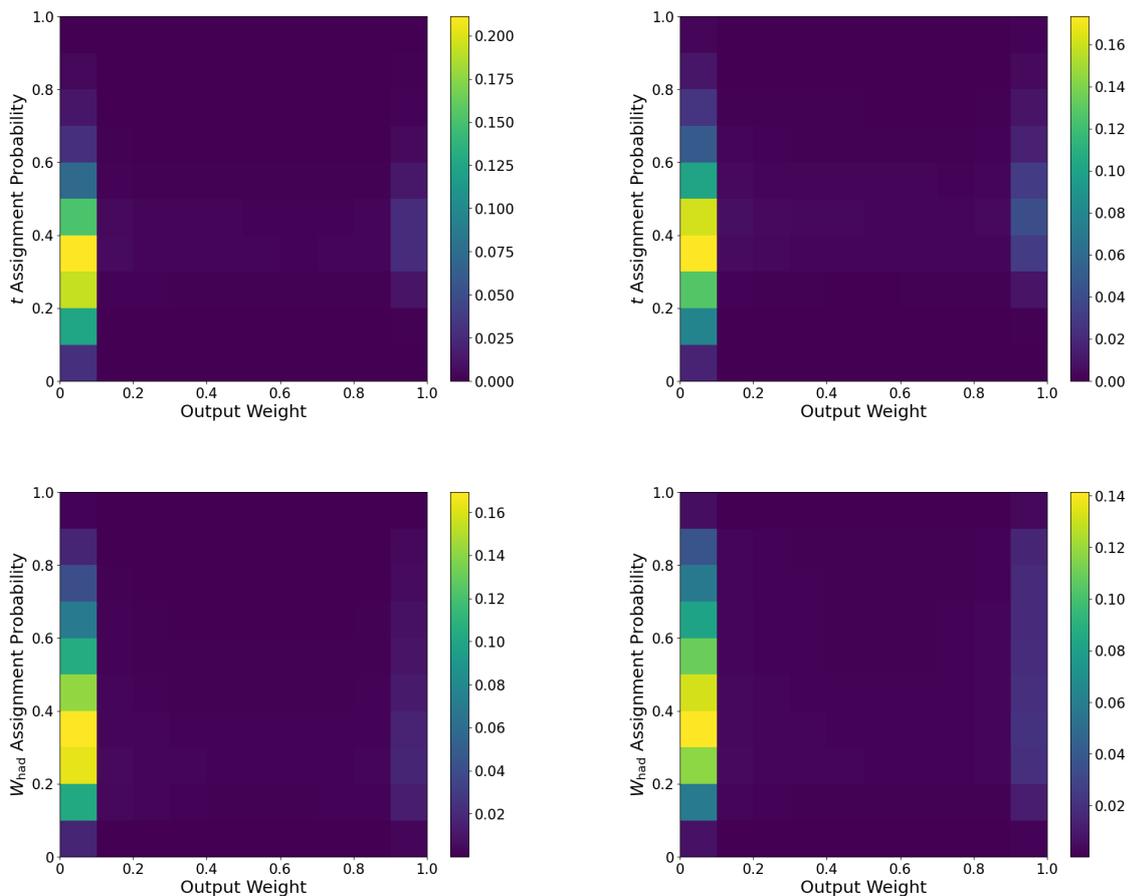


Figure 7.5.: Histogram combining *SPA-Net*'s assignment probability and *Neutrino Weighting* outputs for (left) background events and (right) signal events. Here, the partons are split into (upper)  $t$  quark and (lower)  $W_{\text{had}}$  from  $H$ -boson decay.

detected. The singular selection yields a  $t\bar{t}(H \rightarrow bb)$  acceptance of 75.0% while including 22.03% of all events. Moreover, 20.8% of all signal events are selected by  $CR_{H \rightarrow bb}$  which corresponds to a signal contribution of 0.02%. While the significance  $S/\sqrt{B}$  is relatively high, the signal purity  $S/B$  is low.

The final control region  $CR_{t\bar{t}}$  contains the majority of events with almost five million events. This corresponds to 77.97% of all events. This region accepts 70.0% of signal events which corresponds to a signal contribution of 0.01% due to the high number of other events. As expected, the signal purity is drastically lower compared to  $SR$  while containing higher statistics.



# 8. Results

In this chapter, the results of this study are summarised. First, the validation and performance of the truth matching script are stated in Sec. 8.1. In Sec. 8.2 and 8.3 the results from implementing *SPA-Net* and *Neutrino Weighting* are listed. The final Sec. 8.4 includes the results combining both techniques and using it on real data to compare it to the MC samples. All included uncertainties refer to statistical uncertainties. Additional systematic uncertainties are not included in this study.

## 8.1. Truth Matching

This section briefly summarises the validated performance for the truth matching algorithm. The focus for this algorithm is set on producing unbiased jet-parton matches. While a higher truth matching performance is beneficial for creating sufficient training sets, it is not the highest priority to optimise for the overall analysis since *SPA-Net* is also able to learn using partial events.

### Final Jet Selection

During the matching process, a parton might have more than one potential jet to match. In this case, a metric needs to be defined to decide which jet is chosen as the matched reconstructed object. For that reason, final assignment in the case of multiple jet candidates is tested by using three different assignment schemes.

The first matching scheme compares only the spatial separation between the potential jets and the targeted particle. The closest jet when calculating  $\Delta R$  in respect to the parton is selected as the final match. The second approach utilises only  $|\Delta p_T|$  of the potential jets and the targeted particle. Here, the closest pair in momentum space is selected as the final match. This scheme is motivated by the CP violating properties of the  $W$ -boson which should result in non-symmetric transverse momenta for the decay particles. The last scheme combines the information of both  $\Delta R$  and  $|\Delta p_T|$ . For combined scheme, the  $\Delta R$  was scaled to a range of  $[0,4]$  and  $|\Delta p_t|$  was multiplied by  $10^{-5}$ . This

## 8. Results

Table 8.1.: Overview of the three different matching schemes tested for the final jet matching. It shows the success fraction for each resonance parton in  $t\bar{t}H$  events. The highlighted approach is chosen for the analysis.

	successful matches in %				
	$t\bar{t}H$	$H$	$t\bar{t}$	$t$	$\bar{t}$
$\Delta R$ only	15.5	75.2	18.7	46.5	44.0
$\Delta R$ & $\Delta p_T$	15.4	75.5	18.7	46.2	43.8
$\Delta p_T$ only	15.1	75.4	18.7	46.1	43.6

results in a metric range of approximately 0-10. Then, the pair that minimises the sum of both differences is chosen as the final match.

The success ratios for matching individual particles or the  $t\bar{t}H$  system are summarised in Table 8.1. The  $t$  quark matches are determined by matching all decay products. The combined  $t\bar{t}$  and  $t\bar{t}H$  systems are counted as successful, when all partons of their decay chain are matched. Events where some partons do not exist are considered partially, counting only partons that exist on truth-level. For the test, approximately 1.5 million events were used and the table shows that the  $\Delta R$  scheme performs the best. Hence, it is chosen for this analysis.

### Performance by Process

To further investigate the performance of the truth matching algorithm, its overall performance is analysed. Therefore, the number of successful events are counted when considering different sample compositions.

The summary of the overall performance can be found in Table 8.2. The truth matching algorithm performs best on the semileptonic  $t\bar{t}(H \rightarrow WW^*)$  subset, yet the majority of events is not fully matched. However, this is not problematic since *SPA-Net* is able to learn with partial events. Hence, all events that include at least matched jet are kept to use in training.

## 8.2. SPA-Net

This section covers the jet-parton assignment using *SPA-Net*. The final neural network is optimised to correctly assign jets for the event reconstruction. Additional studies focusing on events with wrong jet assignments were conducted.

Table 8.2.: Overview of the truth matching performance split by different input samples. A successful match means all decay particles of that parton are matched to a jet. If any particle does not exist on truth-level, the event is not counted towards that column. The  $t\bar{t}H$  column only includes events where every particle exists on truth-level. The highlighted row corresponds to the composition used for training *SPA-Net*.

	successful matches in %				
	$t\bar{t}H$	$H$	$t\bar{t}$	$t$	$\bar{t}$
$t\bar{t}(H \rightarrow WW^* \rightarrow l\nu l\nu)$	15.6	75.8	18.8	46.5	44.0
$t\bar{t}(H \rightarrow WW^*)$	15.5	75.2	18.7	46.5	44.0
$t\bar{t}H, t\bar{t}$ & $t\bar{t} + X$ composition	10.3	74.7	9.0	46.5	40.4

Table 8.3.: List of the event composition used in training the final *SPA-Net* model.

Event Topology	Event Count	Contribution
$t\bar{t}$	581683	32.1 %
$t\bar{t} + X$	3925	0.3 %
$t\bar{t}(H \rightarrow bb)$	85054	4.7 %
$t\bar{t}(H \rightarrow \text{other})$	19560	1.0 %
$t\bar{t}(H \rightarrow WW^* \rightarrow qq\bar{q}\bar{q})$	454473	25.1 %
$t\bar{t}(H \rightarrow WW^* \rightarrow l\nu qq)$	545439	30.1 %
$t\bar{t}(H \rightarrow WW^* \rightarrow l\nu l\nu)$	123212	6.8 %
total	1813347	100 %

## Final Model

Several MC sample sets were used to train *SPA-Net*. The final model is trained on a MC set containing 1.8 million MC events. The training sample is composed of  $t\bar{t}$ ,  $t\bar{t} + X$  and several  $t\bar{t}H$  events with additional  $t\bar{t}(H \rightarrow WW^*)$  events injected to improve the network performance for the targeted topology. The detailed event composition is listed in Tab. 8.3. The events are not balanced between the event types.

Moreover, the architecture and other hyperparameters were optimised in a grid-search to find the best performing network. The final parameter values are listed in Tab. 8.4. The following results are based on this final model.

## Assignment Matrix

To better understand *SPA-Net*'s output, the predicted assignment for each jet is listed against its true label. The resulting normalised confusion matrices are shown in Fig. 8.1.

The first matrix includes no partial events which means that every parton has a true jet assignment. The figure shows dominating diagonal terms, with over 70-90 % of events

## 8. Results

Table 8.4.: Summary of the *SPA-Net*'s final architecture and hyperparameters. Parameters marked with \* are not optimised.

Hyperparameter	Value
Epochs	128 (early stopped at 69)
Batch Size	2048
Training / Evaluation Split*	80 % / 20 %
Learning Rate	0.0015
Hidden Layers	16
Transformer Layers	64
Initial Embedding Layers*	16
Position Embedding Layers*	16
Include Skip Connection*	True
Normalise Input Variables*	True
Use Partial Events	True

included. A diagonal matrix would correspond to a perfect prediction where every *SPA-Net* assignment equals the true assignment. The  $(q_1, q_2)$  2x2 submatrices for  $t_1$ ,  $t_2$  and  $W_{\text{had}}$  show the predefined allowed permutations during learning. Hence, an off-diagonal term within these submatrices is considered as a correct, but swapped, assignment since the permutation does not change the underlying physics.

The second matrix includes partial and complete events. Hence, where some particles are not matched on truth-level. The result shows that *SPA-Net* is quite sensitive to partial events. The distribution generally becomes broader with only 30-60% diagonal entries. Furthermore, the  $t$  quark 2x2 submatrices are less visible, since some light quarks get matched as  $b$ -quarks and vice versa. The light jet assigned to the  $t_1 q_2$  and  $t_2 q_2$  particles get assigned to the wrong  $t$  quark in 24-47% of all events. The jets from the  $W_{\text{had}}$ -boson also gets mislabelled more often. In approximately 30% of events, the light jets are wrongly identified as a  $t$  quark jet. Interestingly, *SPA-Net* only rarely swap the  $b$ -tagged jets from the different  $t$  quarks.

The results show that *SPA-Net* is able to predict events correctly, if every expected particle is measured and thus, can be matched. If any particle is not detected, the neural network accuracy reduces significantly.

In Tab. 8.5 shows how many events get assigned correctly, swapped or incorrectly. If a parton is not matched on truth-level, it is not considered for the corresponding entry. Swapped assignments allow permutations within the 2x2 submatrices. The overview shows that *SPA-Net* assigns the  $t$  quarks correctly or swapped in 17.6% of events. For the remaining 84.4%, at least one decay particle is assigned wrong. The hadronic  $W$ -boson

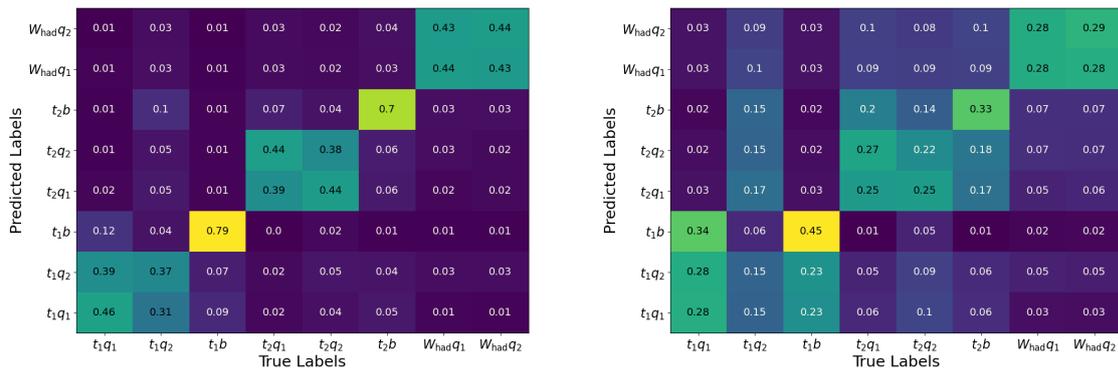


Figure 8.1.: The *SPA-Net* prediction confusion matrices showing the predicted and true labels using (left) full events and (right) full and partial events. Each column is normalised to unity.

reaches 27.9% correct or swapped assignments. The majority of wrong assignments for each parton is caused by the high amount of partial events in the used sample.

## Probability Output

To better identify events with accurate predictions, *SPA-Net* outputs a detection and assignment probability for each resonance particle. These probabilities describe *SPA-Net*'s confidence whether the targeted resonance particle exists and whether it is assigned correctly. These probabilities are visualised in Fig. 8.2 for the hadronic  $W$ -boson from the  $H$ -boson sub-system. In Appx. A, an overview of all output probabilities is given.

The detection probability has similar distribution for correct and wrong assignments. However, the distribution has a different shape for events where the targeted resonance particle  $W_{had}$  does not exist. Hence, the variable can be used to better identify events with a reconstructable hadronic  $W$ -boson from the  $H$ -boson decay.

To differentiate between correct and incorrect predictions, the assignment probability can be used. A particle with high probabilities is more likely to be assigned correctly. Here, the correct and swapped assignments behave very similar which is expected since *SPA-Net* treats swapped assignments as correct.

Both metrics further indicate a well-trained network. These metrics are also used in the event selection, to improve the region definition.

## 8. Results

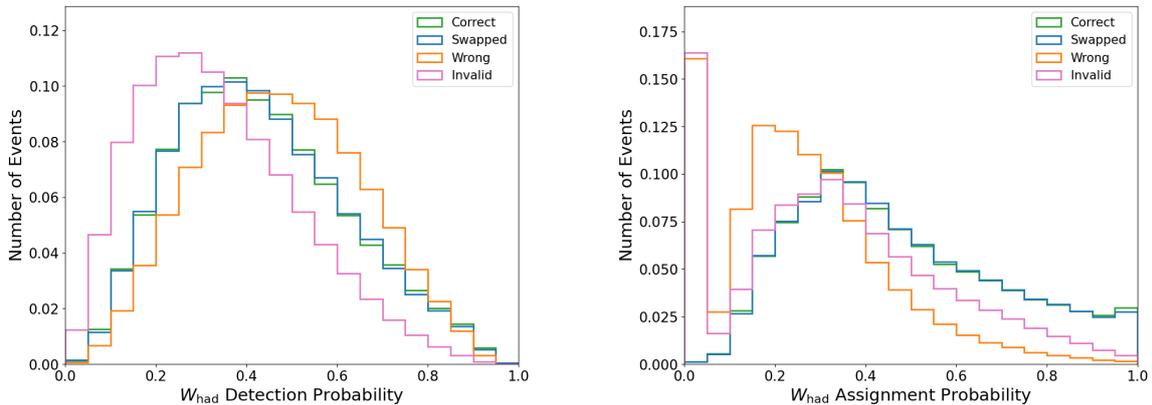


Figure 8.2.: Normalised distributions of *SPA-Net*'s (left) detection and (right) assignment probabilities for the hadronic  $W$ -boson from the  $H$ -boson decay. The events are split into correct, swapped, incorrect and invalid. Invalid events correspond to events where the targeted resonance does not exist. Swapped assignments correspond to swapped assignments within the predefined permutations.

Table 8.5.: Summary of *SPA-Net*'s parton prediction distributions. Both rows contain the relative contribution and absolute number of events for correct, swapped and wrong events. The  $t$  quark row contains the average of  $t_1$  and  $t_2$ .

	Correct	Swapped	Wrong
$t$	8.0 % 333322	7.6 % 31529	84.4 % 351634
$W_{\text{had}}$	13.8 % 42729	14.1 % 43623	72.1 % 222845

### 8.3. Neutrino Weighting

As explained in Ch. 6, the deployed *Neutrino Weighting* algorithm predicts the unknown parameters  $\eta_\nu$  and  $m_{W^*}$ . This is done by calculating the most likely combination of these parameters while using the hadronic  $W$ -boson from the  $H$ -boson decay and missing transverse energy  $E_T^{\text{miss}}$  as input. Assuming the decay mode and on-shell mass of the  $H$ -boson as outlined in Ch. 7, the event can be reconstructed using *Neutrino Weighting* estimations.

#### Estimation Heatmaps

To validate the accuracy of the *Neutrino Weighting* algorithm, the estimated solutions are compared to the true values of the MC samples. Here, the *Neutrino Weighting* algorithm

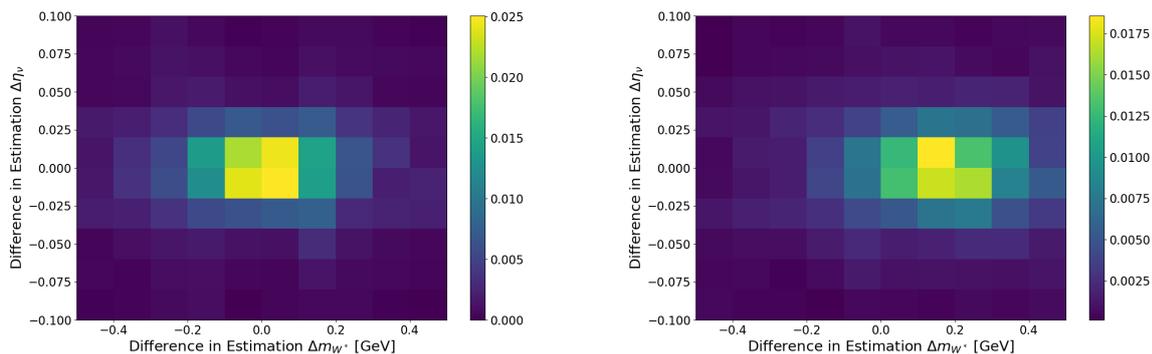


Figure 8.3.: Comparison of two heatmaps showing the difference between the estimated and true values. The two heatmaps use (left) the correct  $m_H = m_H^{\text{MC}}$  and (right) the off-set  $m_H = 125.2 \text{ GeV}$ .

uses the truth information as input. Hence, effects due to simulated detector inefficiencies and reconstruction algorithms are ignored in the comparison.

As an additional validation step, the assumed Higgs mass  $m_H$  is varied. The MC samples for validation use a true mass of  $m_H^{\text{MC}} = 125 \text{ GeV}$ . The *Neutrino Weighting* algorithm is tested by assuming the correct mass of  $m_H = m_H^{\text{MC}}$  and a slightly off-set mass  $m_H = 125.2 \text{ GeV}$ . Fig. 8.3 shows the difference of the estimated values of  $\eta_\nu$  and  $m_{W^*}$  for both assumptions. The heatmaps clearly show the sensibility of the algorithm in respect to the assumed  $H$ -boson mass. Furthermore, for a correct assumption  $m_H = m_H^{\text{MC}}$ , the distribution peaks around  $(\nu_\eta, m_{W^*}) = (0, 0 \text{ GeV})$  as expected. Using a different  $H$ -boson mass significantly changes the estimated mass its decay particle  $W_{\text{lep}}$ . Hence, when using  $m_H = 125.2 \text{ GeV}$  the estimated mass  $m_{W^*}$  is increased by around 200 MeV in respect to the true value due to kinematic conservation laws. This sensitivity is expected to introduce a systematic uncertainty in regards to the assumed  $m_H$  which should be studied. However, this thesis does not include the analysis of systematic uncertainties. When using data, the  $m_H$  assumption needs to be changed as stated in Ch. 6.

Furthermore, in Fig. 8.4 the same distribution is shown. Here, the  $H$ -boson mass is not varied, but the input uses the reconstructed information, including simulated detector effects. The predictions are not centred around  $(\nu_\eta, m_{W^*}) = (0, 0 \text{ GeV})$  which is expected to originate from events where the on-shell  $W_{\text{had}}$  assumption is not fulfilled. Hence, the estimated mass differs by several GeV from the true value. Also, when using reconstructed information as input, the *Neutrino Weighting* estimations deviate stronger from the true values. These increased inaccuracies originate from less accurate inputs. Hence, the dependence on the jet and kinematic accuracies is expected to be an important systematic that should be studied in the future.

## 8. Results

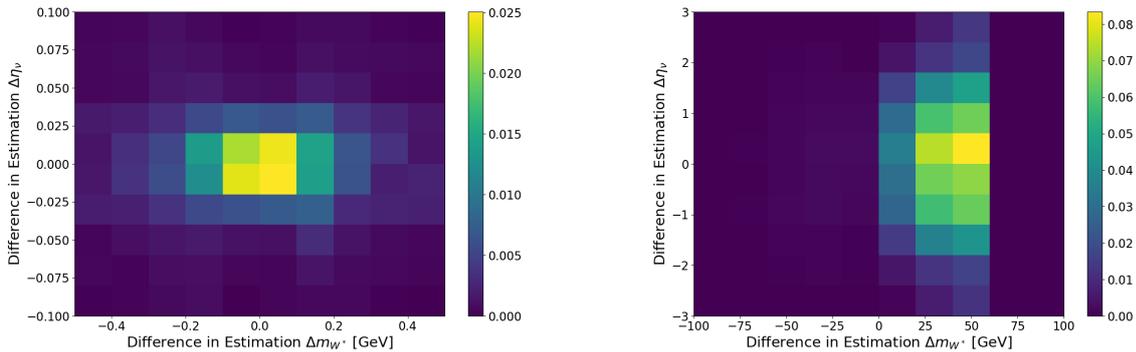


Figure 8.4.: Comparison showing the *Neutrino Weighting* output difference when using (left) true information and (right) reconstructed information as input. Note the different scales.

## Performance

To evaluate the *Neutrino Weighting* algorithm, different setups have been tested. Each setup differs in its input variables, grid size and use of the optimisations which are explained in Sec. 6.2. The results are listed in Tab. 8.6. To quantify the performance, three metrics are calculated.

The average signal weight  $\bar{w}_{\text{signal}}$  from the *Neutrino Weighting* algorithm is an indicator, how good the reconstruction is. It is the arithmetic mean of all calculated weights of all events with at least one solution. Values close to 1 correlate to estimations matching the observed values closely. Lower values correspond to estimations of lower quality. The signal reconstruction efficiency  $\epsilon_{\text{signal}}$  equals the number of signal events containing at least one solution divided by the total number of signal events. Higher efficiencies mean that the *Neutrino Weighting* algorithm finds solutions for signal events more often. Hence, more  $t\bar{t}(H \rightarrow WW^*)$  events become reconstructable, regardless of the solution quality. The reconstruction purity  $p_{\text{signal}}$  is the ratio of reconstructed semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events to all reconstructed events. A high purity means less background events have a *Neutrino Weighting* solution, regardless of their quality.

All metrics should be optimised simultaneously to ensure sufficient numbers of signal events with high quality solutions. The table shows that the algorithm is able to achieve a 100% purity when using the truth values as input. While this result is unrealistic, it further validates the algorithm's estimations. As expected, the average weight and efficiency is the highest when using the exact truth values as input.

Three more test were conducted when using the reconstructed input while ignoring the *SPA-Net* prediction. Instead, the correct assignment is used. These three test give a more

Table 8.6.: Summary of the *Neutrino Weighting* performance on all events using different setups. Each setup is evaluated for the  $tt(H \rightarrow WW^*)$  signal by comparing the average weight  $\bar{w}$ , the reconstruction efficiency  $\epsilon$  and the reconstruction purity  $p$ . The highlighted setup is used.

Setup	$\bar{w}_{\text{signal}}$	$\epsilon_{\text{signal}}$	$p_{\text{signal}}$
truth input (200x200)	0.610	92.2 %	100 %
reconstructed input (100x100)	0.226	9.3 %	77.0 %
reconstructed input (200x200)	0.254	9.3 %	77.2 %
reconstructed input (100x100) + optimisations	0.252	9.4 %	76.8 %
<i>SPA-Net</i> input (100x100) + optimisations	0.231	25.6 %	45.3 %

realistic insight into how well the algorithm works independently of the performance of *SPA-Net*. All three metrics decrease overall which is expected due to the less accurate input. When increasing the grid granularity, the average weight and purity increase slightly. However, this effect can also be achieved by using the aforementioned optimisations. The increase in  $\bar{w}_{\text{signal}}$  originates from the introduced regions of interest while the changes in the efficiency and purity originate from the  $H$ -boson mass smearing. The comparison shows that a 100x100 grid size including the optimisations is sufficient. Additionally, the optimised setup uses fewer computing resources than using finer grid granularities.

Lastly, the final setup is evaluated which uses the 100x100 grid with optimisation but also uses *SPA-Net*'s predictions. Due to the *SPA-Net* performance, some jets are incorrectly matched. The resulting incorrect parton reconstruction leads to more background events being reconstructable and thus, the purity decreasing to 45.3%. Moreover, more signal events are now reconstructable which can be seen in the increase of  $\epsilon_{\text{signal}}$ . This is expected to originate from signal events that do not fulfil the event assumptions. With correct assignments, the *Neutrino Weighting* algorithm does not find a solution. When some partons are assigned incorrectly, the reconstructed properties vary. Thus, the incorrect event properties can yield a solution event when the on-shell or decay assumption is not fulfilled. Overall, the reconstruction quality is slightly seen by the decreased average weight.

## 8.4. Data

This section contains the analysis of the combined Run I and Run II dataset collected by ATLAS using the final *SPA-Net* model and the optimised *Neutrino Weighting* algorithm.

## 8. Results

Table 8.7.: Summary of the event yields using the combined ATLAS dataset of Run I and Run II. The region definition is outlined in Sec. 7.3.

	$SR$	$CR_{t\bar{t}}$	$CR_{H\rightarrow bb}$
Event Count	$67 \pm 8$	$2501793 \pm 1582$	$752538 \pm 868$
Relative Yield in %	$0.0021 \pm 0.0003$	$76.87 \pm 0.05$	$23.12 \pm 0.03$

### Signal and Control Regions

The region yield is compared when using data instead of the MC samples. The results for data can be seen in Tab. 8.7. Compared to the results in Tab. 7.3, the relative yield of each region differs by less than 2%. When taking the statistical uncertainties into account, the  $SR$  contributions are within a  $1\sigma$ -interval. For both  $CR$ s, the statistical uncertainties become relatively small due to their high amount of events per region. Hence, an investigation into systematic uncertainties would be beneficial to evaluate these high-statistic regions. In particular, systematic uncertainties regarding the variables used in the region definition and the theoretical modelling uncertainties for  $t\bar{t}$  and  $t\bar{t}(H \rightarrow bb)$  are expected to be important contributions to the overall uncertainty.

### Comparison

Several variables have been investigated to further compare the ATLAS dataset to the MC simulation. The event distributions of several selected variables are shown in Fig. 8.5. All other investigated variables show similar behaviour as seen in Appx. B.

The MC samples and dataset show similar behaviour for measured quantities and derived variables alike. The normalised event distributions have almost identical shapes and their respective values are mostly within a  $3\sigma$ -interval of each other although only statistical uncertainties are included. Due to the high statistics available, most bins contain very small uncertainties.

The most notable difference is seen in the missing transverse energy  $E_T^{\text{miss}}$  and the  $b$ -jet multiplicity. The MC sample contains higher contributions of events with  $E_T^{\text{miss}} > 70$  GeV and thus, the shape is slightly shifted towards higher  $E_T^{\text{miss}}$ . Moreover, the  $b$ -jet multiplicity shows a deficit of 5-20% for events with  $\geq 3$   $b$ -jets.

However, the mentioned deviations as well as others exceeding a  $3\sigma$ -interval are expected to be resolved when including their respective systematic uncertainties such as  $b$ -tagging efficiencies or  $E_T^{\text{miss}}$  resolution. For the derived quantities such as *SPA-Net*'s probabilities and *Neutrino Weighting* output, systematic uncertainties related to their input variables should resolve significant deviations.

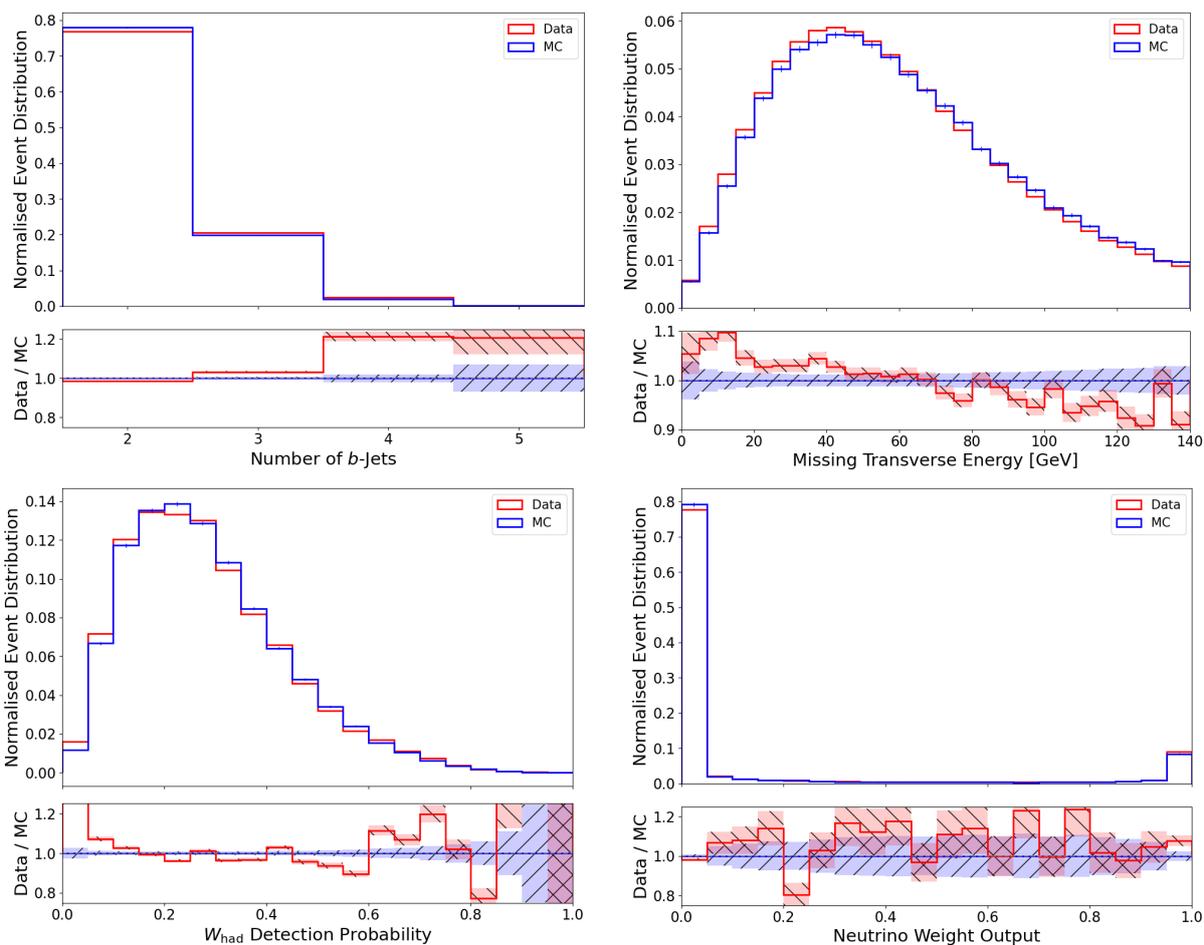


Figure 8.5.: Comparison of the normalised MC and data event distributions for selected variables. These include the (top left) total jet multiplicity, (top right) missing transverse energy, (bottom left)  $SPA\text{-}Net$ 's  $W_{\text{had}}$  detection probability and (bottom right)  $Neutrino\ Weighting$  output.



# 9. Discussion

To conclude this thesis, a summary of the main results is given in Sec. 9.1. This section will reiterate the main measurements with respect to the main goal of this analysis. Lastly, inspiration for future steps of this analysis are explained in Sec. 9.2.

## 9.1. Summary

To summarise this thesis, the main goals of this study are reiterated. The project focused on separating semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events while using and testing modern computational techniques. Additionally, the comparison of the Run I and Run II dataset to the available MC samples was conducted.

### Signal Separation

As explained in Sec. 2.4, the targeted semileptonic  $t\bar{t}(H \rightarrow WW^*)$  event topology is difficult to measure due to its rarity. The theoretical calculation expects over 10000 background events per signal event. This corresponds to a signal purity of less than 0.01 %.

Using *SPA-Net* and *Neutrino Weighting* to analyse and separate the simulated events led to the region definition in Sec. 7.3. For the signal region *SR* the purity of  $t\bar{t}(H \rightarrow WW^*)$  events reaches 0.44 %. However, the main  $t\bar{t}$  background still dominates the *SR* region with over 99 % contribution which is caused by the significant higher cross-section of this process. The  $t\bar{t}(H \rightarrow bb)$  background was efficiently separated by simply selecting an excess in *b*-jets. This singular selection separated 75.0 % of the  $t\bar{t}(H \rightarrow bb)$  background. Still,  $t\bar{t}$  is dominating this region.

This sensitivity study shows that the outputs of the deployed techniques can be used to increase the signal sensitivity significantly, but the event channel remains difficult to separate from background. Thus, the algorithm outputs should be combined with a more intricate kinematic event selection for a more sensitive selection.

## Modern Techniques

The adaptation of *SPA-Net* and *Neutrino Weighting* for event studies were tested and evaluated. It is shown that a properly trained neural network can efficiently assign jets to expected parton as demonstrated in Sec. 8.2. Simultaneously, *SPA-Net* produces output probabilities which can be used for the event definition. Furthermore, as shown in Sec. 8.3, the *Neutrino Weighting* algorithm can be successfully adapted to events with only one neutrino in the final state. When using a truth-based input, the algorithm is able to reconstruct the unknown parameters with  $<1\%$  deviation. While this is unrealistic, this demonstrates its ability to precisely estimate unknown kinematics when using sufficiently accurate input. Moreover, the deployed optimisations as explained in Sec. 6.2 are efficient additions to the *Neutrino Weighting*. They omit the need of additional grid points which reduces the needed computing resources. Additionally, the event acceptance is increased due to the smearing of assumed  $H$ -boson mass. However, the results also show that the algorithm is highly dependent on the input accuracy. Therefore, the *Neutrino Weighting* is only viable when the analysis uses precise measurements and confident jet assignments.

## Data Comparison

For the comparison of the ATLAS dataset of Run I and Run II combined to the MC samples, the regions and several variables were investigated. The results in Sec. 8.4 show similar behaviour with little differences in the overall shape. The normalised values of each distribution mostly match within a  $3\sigma$ -interval when considering only statistical uncertainties. The remaining deviations are expected to resolve when including the respective systematic uncertainties as explained. No unexpected major deviations were observed. Especially the control region comparison will benefit greatly from systematic uncertainties due to the very high statistics in both  $CR$ . In conclusion, this study substantiates that the MC samples for ATLAS dataset is well modelled and hence, can be used reliable when training neural networks and evaluating algorithms.

## 9.2. Future Aspects

This section outlines potential tasks for the future. While this study provides insights into the  $t\bar{t}(H \rightarrow WW^*)$  event separation, there remain opportunities for refinement. These potential future tasks are explained briefly. However, due to time constraints these optimisations exceeded the scope of this work.

## Include Systematic Uncertainties

This study only includes statistical uncertainties for the final results. Hence, systematic uncertainties due to detector calibrations, theoretical models, luminosity and other effects are not included. In the future, this analysis should be complemented by introducing these systematic uncertainties.

Especially the *Neutrino Weighting* performance evaluation would benefit studying the systematic uncertainties introduced by the used input as explained in Sec. 8.3. Varying these would result in a more detailed understanding of the algorithm and its sensitivity in respect to the input.

Combining systematic and statistical uncertainties will yield more realistic uncertainties for the measurements. Moreover, altering the systematic uncertainties would show the sensitivity of the measurements with respect to the investigated uncertainties. Therefore, the reduction of the most sensitive uncertainties should then be prioritised to improve the results of this study.

## Further Optimisation

While this study includes several technical optimisations, additional improvements could be tested in the future.

As seen in this study, *SPA-Net*'s prediction accuracy is limited by the correct  $t\bar{t}$  assignment. Hence, training the neural network to assign these jets correctly should improve the overall performance of this analysis. This could be done, by initially training *SPA-Net* only on  $t\bar{t}$  events and then fine-tuning the parameters for semileptonic  $t\bar{t}(H \rightarrow WW^*)$  events. Since  $t\bar{t}$  events are the main background in this analysis, this could also help to increase the signal purity in the event selection.

The *Neutrino Weighting* algorithm also includes some optimisations such as the regions of interest and  $H$ -boson mass smearing. As an additional improvement, changing the  $\eta_\nu$  sampling to an adaptive approach could improve algorithm by reducing the number of needed samples. The expected  $\eta_\nu$  is depend on the overall event energy, making certain regions more likely to contain a solution.



# Bibliography

- [1] M. K. Gaillard, P. D. Grannis and F. J. Sciulli, *The standard model of particle physics*, Rev. Mod. Phys. **71(2)** (1999)
- [2] F. Wilczek, *Quantum field theory*, Rev. Mod. Phys. **71(2)** (1999)
- [3] A. Iserles et al., *Lie-group methods*, Acta Numer. **9(1)** (2001)
- [4] R. Mills, *Gauge fields*, Am. J. Phys. **57(6)** (1989)
- [5] H. D. Politzer, *Asymptotic freedom: An approach to strong interactions*, Phys. Rept. **14(4)** (1974)
- [6] H. D. Politzer, *Reliable perturbative results for strong interactions?*, Phys. Rept. Lett. **30(26)** (1973)
- [7] D. J. Gross and F. Wilczek, *Asymptotically free gauge theories. I*, Phys. Rev. D **8(10)** (1973)
- [8] W. Marciano and H. Pagels, *Quantum chromodynamics*, Phys. Rept. **36(3)** (1978)
- [9] M. Gell-Mann, *The Eightfold Way: A Theory of strong interaction symmetry*, Am. J. Phys. **33(8)** (1961)
- [10] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19(21)** (1967)
- [11] J. Schwinger, *Quantum electrodynamics. I. A covariant formulation*, Phys. Rev. **74(10)** (1948)
- [12] J. Schwinger, *On quantum-electrodynamics and the magnetic moment of the electron*, Phys. Rev. **73(4)** (1948)
- [13] P. A. M. Dirac, *The quantum theory of the emission and absorption of radiation*, Proc. Roy. Soc. Lond **114(767)** (1927)
- [14] E. Fermi, *Quantum theory of radiation*, Rev. Mod. Phys. **4(1)** (1932)

## BIBLIOGRAPHY

- [15] S. L. Glashow, *Partial-symmetries of weak interactions*, J. Phys. G. **22(4)** (1961)
- [16] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13(2)** (1964)
- [17] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2(7)** (1970)
- [18] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13(16)** (1964)
- [19] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13(9)** (1964)
- [20] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13(20)** (1964)
- [21] T. Brauner, *Spontaneous symmetry breaking and Nambu–Goldstone bosons in quantum many-body systems*, Symmetry **2(2)** (2010)
- [22] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10(12)** (1963)
- [23] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49(2)** (1973)
- [24] Particle Data Group Collaboration, M. Tanabashi et al., *Review of Particle Physics*, Prog. Theor. Phys. **2022(8)** (2022)
- [25] CDF Collaboration, *Observation of top quark production in  $\bar{p}p$  collisions*, Phys. Rev. Lett. **74(14)** (1995)
- [26] DØ Collaboration, *Observation of the top quark*, Phys. Rev. Lett. **74(14)** (1995)
- [27] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **8(4)** (2016)
- [28] CMS Collaboration, *Identification of b-Quark Jets with the CMS Experiment*, JINST **8(4)** (2013)
- [29] 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(11)** (2019)

- [30] 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(1)** (2012)
- [31] CMS Collaboration, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716(1)** (2012)
- [32] ATLAS and CMS Collaborations, *Combined Measurement of the Higgs Boson Mass in  $pp$  Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015)
- [33] SNO Collaboration, *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89(1)** (2002)
- [34] T. Kuo and J. Pantaleone, *Neutrino oscillations in matter*, Rev. Mod. Phys. **61(4)** (1989)
- [35] R. Davis, *A review of the Homestake solar neutrino experiment*, Prog. Part. Nucl. Phys. **32(4)** (1994)
- [36] Super-Kamiokande Collaboration, *Evidence for oscillation of atmospheric neutrinos*, Rev. Mod. Phys. **81(8)** (1998)
- [37] B.R. Holstein, *Graviton physics*, Am. J. Phys. **74(11)** (2006)
- [38] T. Damour and A.M. Polyakov, *String theory and gravity*, Gen. Relat. Gravit. **26** (1994)
- [39] V. Trimble, *Existence and nature of dark matter in the universe*, Annu. Rev. Astron. Astrophys. **25(1)** (1987)
- [40] M. E. Peskin, *Dark matter and particle physics*, J. Phys. Soc. Japan **76(11)** (2007)
- [41] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, Helv. Phys. Acta **6** (1933)
- [42] S. Weinberg, *The cosmological constant problem*, Rev. Mod. Phys. **61(1)** (1989)
- [43] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *The hierarchy problem and new dimensions at a millimeter*, Phys. Lett. B **429(3-4)** (1998)
- [44] P. Fayet and S. Ferrara, *Supersymmetry*, Phys. Rep. **32(5)** (1977)

## BIBLIOGRAPHY

- [45] T. Clifton et. al, *Modified gravity and cosmology*, Phys. Rep. **513(1-3)** (2012)
- [46] ATLAS Collaboration, *Inclusive and differential cross-sections for dilepton  $t\bar{t}$  production measured in  $\sqrt{s} = 13$  TeV pp collisions with the ATLAS detector*, JHEP **07** (2023)
- [47] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018)
- [48] LHC Collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, CERN (2017)
- [49] ATLAS Collaboration, *Top cross section summary plots - April 2024* (2024)
- [50] L. Evans and P. Bryant, *LHC Machine*, JINST **3(8)** (2008)
- [51] L. R. Evans, *The Large Hadron Collider (LHC)*, New J. Phys. **9(9)** (2007)
- [52] ATLAS Collaboration, *The ATLAS experiment at the CERN large hadron collider*, JINST **3(8)** (2008)
- [53] ATLAS Collaboration, *ATLAS: Detector and physics performance technical design report. Volume 1*, CERN-LHCC-99-14, ATLAS-TDR-14 (1999)
- [54] ATLAS Collaboration, *ATLAS: Detector and physics performance technical design report. Volume 2*, CERN-LHCC-99-15, ATLAS-TDR-15 (1999)
- [55] ATLAS Collaboration, *ATLAS inner detector: Technical design report. Vol. 1*, CERN-LHCC-97-16, ATLAS-TDR-4 (1997)
- [56] ATLAS Collaboration, *ATLAS inner detector: Technical design report. Vol. 2*, CERN-LHCC-97-16, ATLAS-TDR-4 (1997)
- [57] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, CERN-LHCC-2017-021, ATLAS-TDR-030 (2017)
- [58] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, CERN-LHCC-2017-005, ATLAS-TDR-025 (2017)
- [59] ATLAS Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, JINST **3(2)** (2008)
- [60] ATLAS Collaboration, *ATLAS Calorimeter: Run 2 performance and Phase-II upgrades*, PoS Proc. Sci. **314(11)** (2017)

- [61] ATLAS Collaboration, *ATLAS muon spectrometer: Technical design report*, CERN-LHCC-97-22, ATLAS-TDR-10 (1997)
- [62] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*, CERN-LHCC-2017-017, ATLAS-TDR-026 (2022)
- [63] ATLAS Collaboration, *New Small Wheel Technical Design Report*, CERN-LHCC-2013-006, ATLAS-TDR-020 (2013)
- [64] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15(10)** (2020)
- [65] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3(8)** (2008)
- [66] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3(8)** (2008)
- [67] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3(8)** (2008)
- [68] A. Shmakov et al., *SPANet: Generalized permutationless set assignment for particle physics using symmetry preserving attention*, SciPost Phys. **12(5)** (2022)
- [69] M. J. Fenton et al., *Reconstruction of Unstable Heavy Particles Using Deep Symmetry-Preserving Attention Networks*, Commun. Phys. **7(1)** (2024)
- [70] S. Baker and R. D. Cousins, *Clarification of the use of chi-square and likelihood functions in fits to histograms*, Nucl. Instr. Meth. Phys. Res. **221(2)** (1984)
- [71] M. J. Fenton et al., *Permutationless many-jet event reconstruction with symmetry preserving attention networks*, Phys. Rev. D **105(11)** (2022)
- [72] DØ Collaboration, *Measurement of the Top Quark Mass Using Dilepton Events*, Phys. Rev. Lett. **80(10)** (1998)
- [73] F. Fabbri, J. Howarth and T. Maurin, *Isolating semi-leptonic  $H \rightarrow WW^*$  decays for Bell inequality tests*, Eur. Phys. J. C **84(1)** (2024)
- [74] 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(11)** (2018)



# A. Additional SPA-Net Output Probabilities

The used implementation of *SPA-Net* outputs a separate assignment and detection probability for each particle. The following figures A.1, A.2 and A.3 show these probabilities for both  $t$  quarks and the hadronic  $W$ -boson from the  $H$ -boson decay chain. This study utilise the  $t$  assignment probability which equals the average of both  $t$  assignment probabilities.

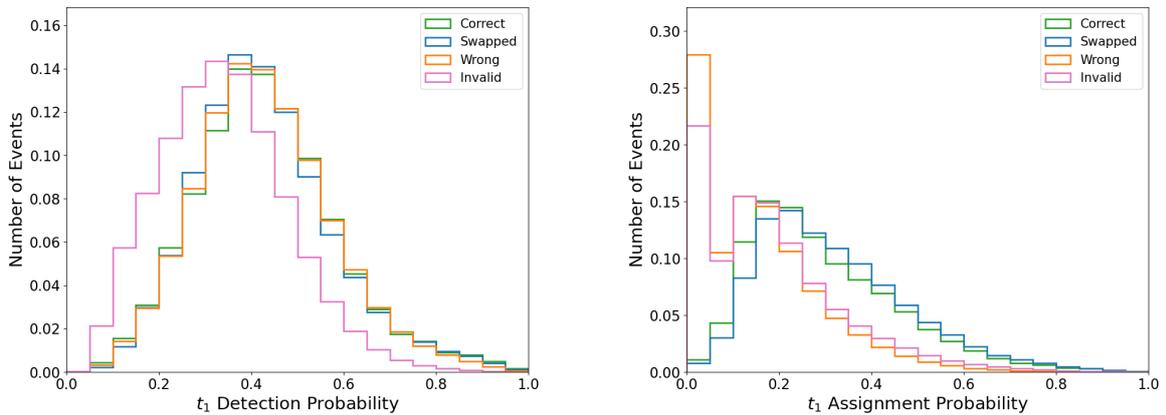


Figure A.1.: Overview of the (left) detection and (right) assignment probability for the first assigned  $t$  quark.

### A. Additional SPA-Net Output Probabilities

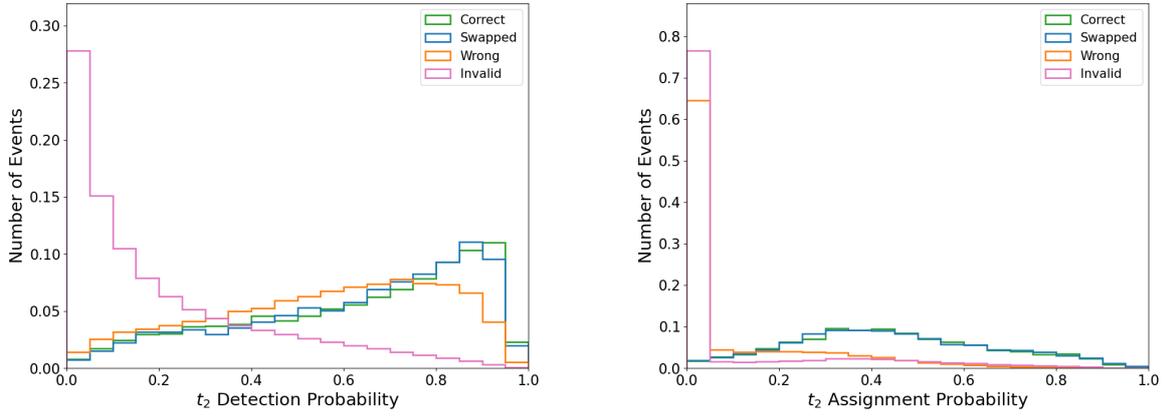


Figure A.2.: Overview of the (left) detection and (right) assignment probability for the second assigned  $t$  quark.

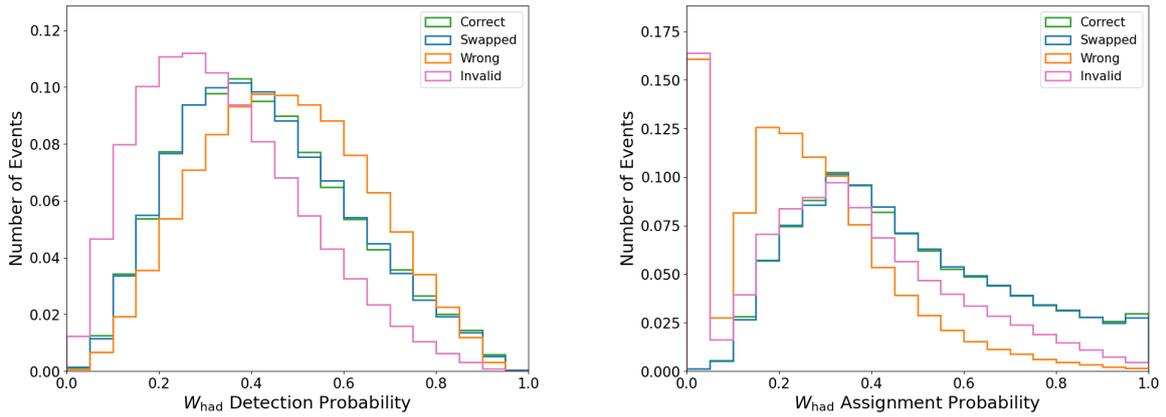


Figure A.3.: Overview of the (left) detection and (right) assignment probability for the hadronic  $W$ -boson from the  $H$ -boson decay.

## B. Additional Comparisons of ATLAS Dataset and MC Samples

The following figures B.1, B.2 and B.3 show comparisons of data and MC simulations for additional investigated variables. These distributions show no significant change in behaviour to what is described in Sec. 8.4 and are included to complement the comparison.

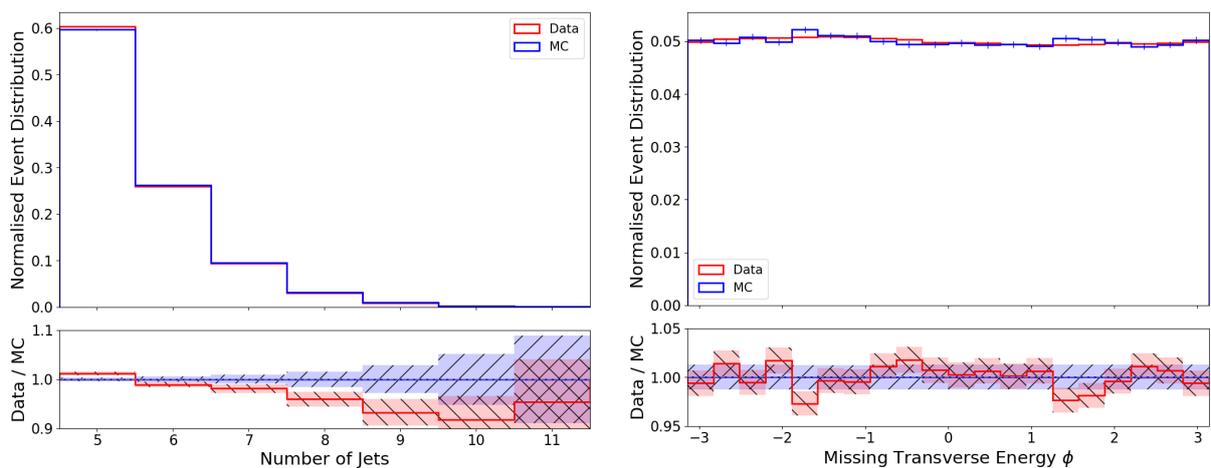


Figure B.1.: Comparison of data vs MC distributions for the (left) total number of jets and (right)  $\phi$  of the missing transverse energy probability.

B. Additional Comparisons of ATLAS Dataset and MC Samples

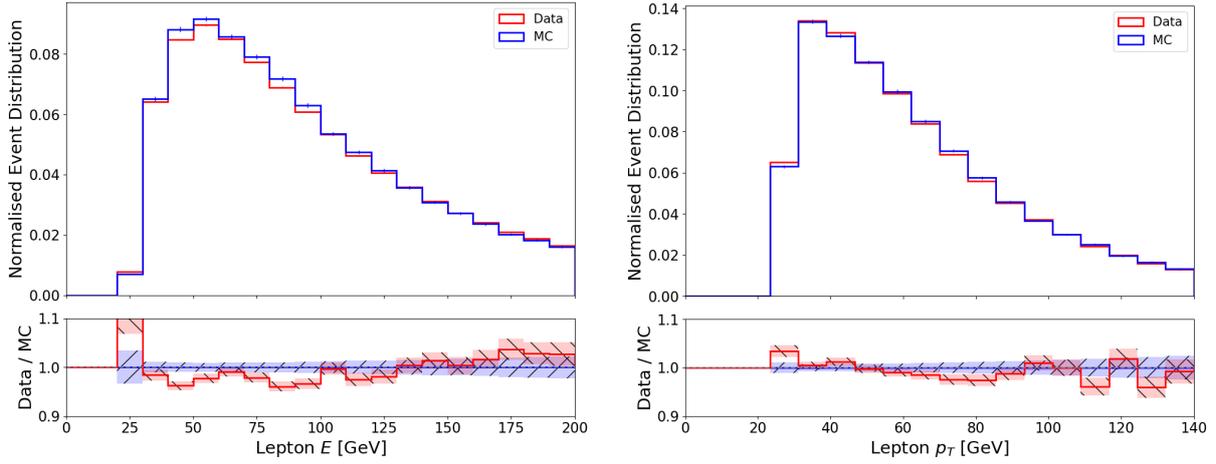


Figure B.2.: Comparison of data vs MC distributions for the (left) energy and (right) transverse momentum of the measured lepton.

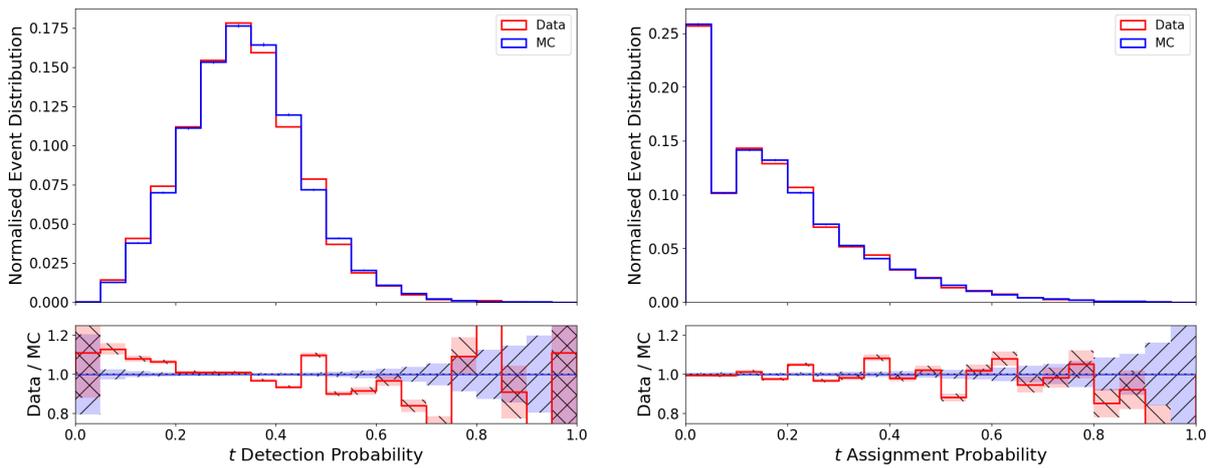


Figure B.3.: Comparison of data vs MC distributions for the average (left) detection and (right) assignment probability of both  $t$  quarks.

# Declaration

**Erklärung** nach §13(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen: Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.

Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestanden Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 27. September 2024



(Ireas Tom Raschke)