University of Bergen

Faculty of Physics and Technology

Search for Dark Matter produced in association with a Higgs boson decaying to tau leptons at $\sqrt{s}$= 13 TeV with the ATLAS detector

Supervisor:

Julia Isabell Djuvsland and Anna Lipniacka

Author:

Erlend Aakvaag

Particle Physics MSc

Bergen, 2020

Abstract

This thesis presents a search for Dark Matter produced in association with the Standard Model Higgs boson decaying into two tau leptons at $\sqrt{s} = 13$ TeV using Run 2 data from the ATLAS detector corresponding to an integrated luminosity of 139 fb^{-1}. Specifically, the thesis documents a search for events predicted by the 2HDM+a, mono-Higgs signal model. The signature exploited here are two hadronically decaying taus with invariant mass consistent with the Standard Model Higgs boson and additional missing transverse energy. This model gives rise to two pseudoscalar spin-0 bosons a, A, and by employing a specific set of benchmark parameters, scans in the $m_a - m_A$ signal grid has been performed in four signal regions. Two signal regions have been defined with a "cut and count" approach, and two with multivariate tools. Background contributions in the four signal regions have been estimated, and associated systematic uncertainties have been evaluated. Expected exclusion limits are set in the $m_a - m_A$ signal grid from exclusion fits performed on simulated Asimov data in the four signal regions. Signal points within the set contours are expected to be excluded at 95% confidence level. This result is complementary to other topology searches for the 2HDM+a model performed by ATLAS in the $m_A - m_a$ grid, [1]. These searches have employed the same set of benchmark parameters as in this analysis.

The analysis has not yet been approved for unblinding by the ATLAS experiment, and the presented material would need be reviewed before looking at real data in the signal regions. Once unblinded however, the aim is to redo the exclusion limits on real data and statistically combine the result with the current ATLAS limits.

Acknowledgements

I would like to thank my supervisors Anna Lipniacka and Julia Isabell Djuvsland for your help, support and patience throughout working on this thesis. I would also like to thank the UiB ATLAS group for all the insights and letting me take part in the ATLAS collaboration this past year. It has been a great and educational ride, so indeed thank you all very much.

Finally, the rest should be trivial, and is left as an exercise for the reader.

,

Contents

Chapter 1

Introduction

The Standard Model is the current theoretical framework for fundamental particle physics. It describes basic constituents of all known matter in terms of elementary particles, and how these elementary particle interact with each other. The theory in its current form was formulated in the 1970s and has since then been subject to stringent experimental testing, showing time and again an unprecedented agreement with data. With the finding of the Higgs boson in 2012, all particles predicted by the Standard Model have been experimental verified. As such, one could be forgiven for wondering if there is anything left to discover in fundamental particle physics. The answer is off course "yes", as there still are plenty of observed phenomena the Standard Model can not explain. With the increased energy and luminosity of the particle beams at the LHC, this has paved the way for a new era of fundamental physics, namely the search for theories beyond the Standard Model.

One such mystery is the particle nature of Dark Matter (DM). Even though several astrophysical experiments have inferred its existence, no such particles have been experimentally observed. However, if DM particles interact with the particles of the Standard Model, they could in principle be produced in collisions at the LHC. This analysis studies the case where Dark Matter particles are produced in association with a Higgs boson in the ATLAS detector. Being a scalar particle, the Higgs boson differs from the other particles in the Standard Model, and by extending the Higgs sector one obtain new particles in the theory which provide interactions that serves as a link between the Standard Model and the Dark sector. Since Dark Matter does not interact electromagnetically it is elusive, and hence will not be directly recorded by the ATLAS detector. However, its existence can be inferred by an imbalance in the transverse momentum, referred to as "missing transverse energy".

The Higgs boson in this analysis is expected to be the Standard Model Higgs with corresponding branching fraction and mass, and the specific Dark Matter model that is studied is known as 2HDM+a. This model contains a set of free parameters which among others are masses of two new pseudoscalar particles, m_A and m_a. By using well motivated choice of benchmark parameters and only varying m_A and m_a, scans over this parameter space have been performed. This parameter space has previously been studied in various channels by ATLAS [1], but not for $h(\tau\tau)$, which is the specific final state studied in this analysis. As such, this analysis is complimentary and aims for a statistically combination of potential exclusion limits with the

current results from ATLAS. This thesis is also supplementary to the ongoing ATLAS analysis in [2], where in addition to 2HDM+a, the $h(\tau\tau)$ final state is searched for in the context of the Z'2HDM model.

The thesis is divided into three parts. The first part gives a theoretical overview of the Standard Model and Dark Matter. The second part explains the experimental setup on which this analysis is based and the utilized software tools. The third part goes through the analysis procedure, where selections are optimized with respect to the Asimov significance in order to obtain signal rich regions with minimal background contamination. This is done both with a cut based approach and by employing Boosted Decsion Trees. Background contributions in these regions are estimated, and corresponding systematic uncertainties are evaluated. Finally, profile likelihood fits are performed on the Asimov simulated data in order to set expected exclusion contours in the 2HDM+a mass plane.

Part I

Theory

Chapter 2

The Standard Model of particle physics

The idea of particles has from a philosophical point of view its origins back in the time of the ancient Greeks around the 5th century BC. Here, Democritus and his contemporaries laid seed to the notion that nature is built up by indivisible parts, atoms as he called it, that makes up the macroscopic world we observe at our scale. It would take over two thousand years before the first developments in atomic theory, followed by the construction of quantum theory which revolutionized physics in the early 20th century. Still, the same questioned remained and is still what drives the field of particle physics today, namely, what are the basic constituents of matter, and how do they group together to form the structures that we see? The current answer to these questions can be found within the Standard Model of particle physics. It is considered to be one the most successful physical theories to date, and describes all know fundamental particles and their interactions[1].

2.1 Overview of the Standard Model

Fundamentally all particles can be grouped into two categories: *fermions* and *bosons*. Simply put, the main difference between the two is their spin quantum number, which leads to widely different properties.

2.1.1 The Standard Model fermions

The Standard Model (SM) fermions are spin-$\frac{1}{2}$ particles, which means that they follow Fermi-Dirac statistics and obey the Pauli exclusion principle. All known matter is built from them, and in total there are 12 different sorts (popularly called flavours) of elementary fermions divided into two groups: 6 quarks and 6 leptons. The lepton sector consists of the familiar electron (e), as well as two heavier replicas called the muon (μ) and the tau (τ). They all have charge $Q = -1$ and are collectively known as the charged leptons. Each charged lepton also has an associated electrically neutral neutrino ν_ℓ, where $\ell{=}e, \mu, \tau$. The charged leptons interact both electromagnetically and weakly, whereas the neutrinos are only subject to the weak interaction as they are electrically neutral. The quantum number related to the weak interaction is the weak isospin, T_3, with value $T_3 = +1/2$ for the charged leptons and $T_3 = -1/2$ for the neutrinos.

[1] Gravity is not yet explained by the Standard Model.

Together, the charged leptons and neutrinos naturally fall into three generations classified according to their to their quantum numbers. In the Standard Model, each generation is represented in terms of a left-handed doublet and a right-handed singlet as a consequence of the weak isopsin symmetry of the $SU(2)_L$ group[2]. This will be further explained in section 2.2.5. In addition, each lepton has a corresponding anti-lepton ($\bar{\ell}$) with identical mass, but opposite charge and lepton quantum number. The antiparticle of the electron, called the positron, carries charge $Q = 1$ and has electron quantum number $L_e = -1$, whereas the anti-electron neutrino remains neutral and also has electron quantum number $L_e = -1$. The same is true for the heavier μ-lepton and τ-lepton respectively. A summary of the leptons and their properties can be found in table 2.1.

Generation	Particle	Mass [MeV/c^2]	Charge [e]	Weak isospin T_3	Anti-particle
I	e	0.511	-1	$+\frac{1}{2}$	$\bar{e}$
	ν_e	$< 2.2 \times 10^{-4}$	0	$-\frac{1}{2}$	$\bar{\nu}_e$
II	μ	105.7	-1	$+\frac{1}{2}$	$\bar{\mu}$
	ν_μ	<0.19	0	$-\frac{1}{2}$	$\bar{\nu}_\mu$
III	τ	1777	-1	$+\frac{1}{2}$	$\bar{\tau}$
	ν_τ	<18.2	0	$-\frac{1}{2}$	$\bar{\nu}_\tau$

Table 2.1: Properties of the leptons sector. The mass values are taken from PDG [3].

The grouping of quarks is similar to that of leptons. They also fall naturally into three generations categorized by their electric charge and weak isospin. Each generation contains a up-type (up, charm, top) quark with charge $Q = +2/3$ and $T_3 = +1/2$, as well as a down-type quark (down, strange, bottom) with charge $Q = -1/3$ and $T_3 = -1/2$. Together, as for the leptons, the up-type and down-type quarks form a left-handed doublet and a right-handed singlet for each generation. Unlike the leptons however, quarks carry an additional color charge and can interact through the strong interaction. There are three such colors; red, green and blue (r, g, b), where adding them all together produce a color neutral state. These color neutral states are collectively called baryons. The proton and neutron are examples of such configurations, with quark content (uud) and (udd) respectively. For each quark q there is also a corresponding anti-quark $\bar{q}$ with opposite electric charge Q and anti-color $\bar{r}, \bar{g}, \bar{b}$. Combining quarks and antiquarks can also lead to a color neutral quark state, allowing color-neutral $q\bar{q}$ configurations. Such configurations are known as mesons. Collectively all possible color neutral states are referred to as hadrons. A summary of the quarks and their properties can be found in table 2.2

[2]Two particles forming a $SU(2)$ doublet mean that they transform into each other under a $SU(2)$ transformation. In other words, as far as the weak interaction is concerned, a left-handed charged lepton and a left-handed neutrino in a doublet are interchangeable. The right handed charged lepton singlets transforms trivially (do not change) under $SU(2)$, and hence do not interact weakly.

Generation	Particle	Mass [MeV/c^2]	Charge [e]	Weak isospin T_3	Anti-particle
I	u	$2.2^{+0.5}_{-0.4}$	$+\frac{2}{3}$	$+\frac{1}{2}$	$\bar{u}$
	d	$4.7^{+0.5}_{-0.3}$	$-\frac{1}{3}$	$-\frac{1}{2}$	$\bar{d}$
II	c	1275^{+25}_{-35}	$+\frac{2}{3}$	$+\frac{1}{2}$	$\bar{c}$
	s	95^{+9}_{-3}	$-\frac{1}{3}$	$-\frac{1}{2}$	$\bar{s}$
III	t	$(173\pm0.4)\times10^3$	$+\frac{2}{3}$	$+\frac{1}{2}$	$\bar{t}$
	b	4180^{+40}_{-30}	$-\frac{1}{3}$	$-\frac{1}{2}$	$\bar{b}$

Table 2.2: Properties of the quark sector. The mass values are taken from PDG [3].

2.1.2 The Standard Model bosons

The universe appear to be governed by four fundamental interactions, three of which are explained by the Standard Model: the strong, weak and electromagnetic interaction. The strong interaction works only at very short distances, binding quarks together to form nucleons, and nucleons together to form nuclei. The electromagnetic interaction is responsible for the attraction between electrons and nuclei needed for the creation of atoms and molecules. The weak interaction underlies processes like β-decay, allowing protons to transmute into neutrons and vice versa. These interactions are in the Standard Model transmitted by spin-1 bosons, collectively called gauge bosons. For example, we describe the electromagnetic interaction through the exchange of a photon which is a quantum state of the electromagnetic field. Electron-electron scattering comes from the emission and reabsorption of a virtual photon as shown in figure 2.1. This pictorial representation of the process is called a Feynman diagram, which is a useful way of visualizing arcane and abstract formulas. Energy conservation dictates the energy of the initial and final state, ee, has to be equal to the energy of the intermediate state, $ee\gamma$. Classically this is not possible, but due to Heisenberg's uncertainty relation the intermediate state can exist with "borrowed energy" if only for a very short amount of time:

$$\Delta E \Delta t \geq \tfrac{\hbar}{2}$$

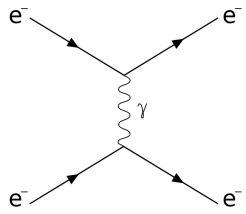


Figure 2.1: Feynman diagram of electron-electron scattering.

The exchanged photon is called virtual, as no such particle can be observed in nature. This way of modelling fundamental interactions has proven to be extremely useful and is an example of a quantum field theory. Here the photon γ is the spin-1 gauge boson mediating the electromagnetic interaction. Photons are themselves neutral and only interacts with electrically charged particles like electrons. In general the potential energy of interactions such as the one figure 2.1 depends of the mass of the exchanged particle, and falls off like $\exp(-mr)/r$. As the photon is massless this reduces to the inverse power law $1/r$, indicating the strength and range of the electromagnetic interaction. The full underlying theory of electromagnetism is a $U(1)$ gauge theory known as Quantum Electrodynamics.

The treatment of the weak and strong interactions are closely analogous to that of the electromagnetic interaction. The strong interaction is transmitted by 8 massless spin-1 gluons. They mediate the strong interaction between quarks and antiquarks in a similar fashion to how photons mediate the electromagnetic interaction between electrically charged particles. What distinguishes them however, is that gluons themselves carry both color and anticolor which leads to gluon self-interactions. As the gluons are massless one could naively expect the strong interaction to fall off as $1/r$ as well, but due to the effects of color the behaviour is very much different at large distances. The full underlying theory describing the strong interaction is the $SU(3)_c$ gauge theory known as Quantum Chromodynamics.

The weak interaction is transmitted by the massive spin-1 $W^{\pm}$ and Z^0 gauge bosons. The W^+ and W^- gauge bosons are electrically charged as indicated by the superscripts, while the Z^0 is electrically neutral. Being massive, their potential is damped by the factor $\exp(-mr)$ resulting in interactions that only manifest itself at very short distances, $\mathcal{O}(10^{-18})$ m. The charged bosons act on the fermion doublets of the Standard Model through a charged current interaction, transforming the components of the doublet into each other through the emission of a $W^{\pm}$. The neutral Z^0 gauge boson can couple to all SM particles in a similar way as in figure 2.1. In fact all Z^0 interactions involving charged particles can also occur via the exchange of a virtual photon. A more detailed explanation of weak interaction, as well as its unification with electromagnetism will be addressed later in this chapter when introducing the underlying $SU(2)_L \times U(1)$ gauge theory of the electroweak interaction.

Finally, the last boson of the Standard Model is the spin-0 Higgs boson. Being a scalar particle it differs from the other particles in the SM. All the above theories are gauge theories, meaning that they are invariant under so-called gauge transformations. Introducing these transformation in electroweak theory requires all the SM particles to massless. This is indeed a problem, as we have clear experimental evidence that some of these particles do have mass. Enter the Brout-Englert-Higgs mechanism, which fixes this problem by spontaneously breaking the $SU(2)_L \times U(1)$ symmetry leading to the acquisition of mass by the particles in the theory. Each particle interacts with the Higgs field at different strengths, and the mass of said particle is proportional to this coupling. Massless particles like the photon do not interact with the Higgs field at all, and remain massless. The Higgs boson (h) properties, as well as properties of the other gauge bosons are summarized in table 2.3. Together with the fermions this lays out the current picture of the physical world.

Interaction	Mediator	Mass [GeV/c^2]	Spin	Charge $Q[e]$	T_3	Range[fm]	Relative strength
Electromagnetic	γ	0	1	0	0	∞	10^{-2}
Weak	Z^0	91.2	1	0	0	10^{-3}	10^{-6}
	$W^{\pm}$	80.4	1	±1	±1	10^{-3}	10^{-6}
Strong	g	0	1	0	0	10^{-15}	1
Scalar sector	h	125.2	0	0	$-\frac{1}{2}$	-	-

Table 2.3: Properties of the boson sector. The mass values are taken from PDG [3].

2.2 Gauge Theories of The Standard Model

2.2.1 General Field theory and symmetries

Before we can discuss the gauge theories of the Standard Model some concepts first needs to be addressed. Gauge theories are quantum field theories that builds on some fundamental mathematical tools. One such quantity of interest is the Lagrangian, which in its simplest terms is defined as the difference in kinetic and potential energy:

$$L = T - V \tag{2.1}$$

In field theory it is more convenient to work with the Lagrangian density, which is a function of the fields ϕ and their derivatives $\partial_\mu \phi$ [3]. Taking the space-time integral of $\mathcal{L}$ over some arbitrary space-time region Ω leads to a new quantity known as the action:

$$S = \int_\Omega d^4x \ \mathcal{L}(\phi, \partial_\mu \phi) \tag{2.2}$$

Hamilton's principle states that the space-time integral over $\mathcal{L}$ is stationary. This means that any infinitesimal variation of the action δS is zero:

$$\delta S = \int_\Omega d^4x \ \left(\frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta(\partial_\mu \phi) \right) = 0 \tag{2.3}$$

From this one can obtain the Euler-Lagrange equations of motion, which describe the dynamics of any system in terms of field ϕ:

$$\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \right) = 0 \tag{2.4}$$

For example fermion fields are described in terms of four-component spinors,ψ, with Lagrangian $\mathcal{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi$. Inserting this into 2.4 yields the well-known Dirac equation:

$$(i\partial\!\!\!/ - m)\psi = 0 \tag{2.5}$$

The predictive power of physics comes from understanding conservation laws, which is closely linked with the concept of symmetry. If a quantity is symmetric in physics it means that it does not change if one applies a transformation associated with that symmetry. Mathematically this can be expressed as shifting the field ϕ by some arbitrary infinitesimal phase ϵ.

$$\phi \to e^{i\epsilon}\phi = \phi + i\delta\phi \tag{2.6}$$

The above transformation is global, and holds for all spacetime points under consideration. The corresponding Lagrangian is symmetric if it is left invariant, i.e does not change by the transformation. If ϕ obeys the equations of motion, the following result can be obtained from equation 2.3:

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \delta\phi \right) = \partial_\mu j^\mu = 0 \tag{2.7}$$

[3] Taking the spatial volume integral of the Lagrangian density leads back to the Lagrangian $L = \int \mathcal{L} \ d^3x$. For this whole thesis the Lagrangian density will be referred to as the Lagrangian.

Here, j^μ is known as the conserved current, which is the central result of what is known as Noether's theorem. This theorem states that for every continuous symmetry of the Lagrangian there is a conserved current, whose form can be derived from the Lagrangian using equation 2.3. The conserved charge associated with the conserved current can be found by integrating the zeroth component:

$$Q = \int d^3x j^0 \tag{2.8}$$

Time translation symmetry of spacetime is an example of the use Noether's theorem leading to the conserved charge of energy[4]. The Standard Model is a theory built on symmetry, making Noether's theorem a central part of the theory. It will later be shown that enforcing the Standard Model Lagrangians to be invariant under certain transformations require the introduction of vector fields in the theory. These vector fields will be associated with the interactions of the Standard Model. The symmetry groups responsible for this are treated in the mathematical framework of group theory.

2.2.2 Quick summary of relevant Group theory

Definition 2.1. *A group is defined as a set G of elements $g \in G$ with an operation $\circ$ that satisfies:*

- *Closure:* $g, h \in G \rightarrow g \circ h \in G$
- *Associativity:* $f \circ (g \circ h) = (f \circ g) \circ h \quad \forall f, g, h \in G$
- *Existence of identity element* $e : g \circ e = g \quad \forall g \in G$
- *Existence of an inverse element* g^{-1} *for each* $g \in G : g^{-1} \circ g = e$

.

A group is said to be Abelian if the operation on group elements is commutative:

- $g \circ h = h \circ g \quad \forall g, h \in G$

Groups that do not meet this requirement are called non-Abelian. In the Standard Model only the group $U(1)$ of electromagnetism is Abelian. A group of dimension n is said to be unitary if it can be represented by $n \times n$ unitary matrices that satisfy the following condition:

$$U^\dagger U = I \tag{2.9}$$

Here, I is the $n \times n$ identity matrix and $U^\dagger$ is the conjugate transpose[5]. A unitary group of order n is denoted $U(n)$. A group is called special unitary, denoted $SU(n)$, if it also satisfies:

$$\det U = 1 \tag{2.10}$$

Especially interesting for particle physics are so-called Lie groups, which are groups whose elements can be written on the form:

$$g = e^{j\alpha_j T_j} \tag{2.11}$$

[4]Charge in this sense meaning a Noether's charge, i.e a conserved quantity, not to be confused with discrete charges.
[5]The conjugate transpose is the combined operation of taking the complex conjugate and the transpose of a matrix.

Here, $j = 1, .., n$ with n being the dimension of the group. α_j are arbitrary real numbers determining g, and T_j are the generators of the group. In general any special unitary group $SU(n)$ has $n^2 - 1$ generators associated with it. The commutation relation between the generators, often called the Lie algebra of the group, is given by

$$[T_i, T_j] = if_{ijk}T_k \quad \forall i, j = 1...n \tag{2.12}$$

,where f_{ijk} are real totally anti-symmetric structure constants.

With the vocabulary of group theory we can now apply this to the Standard Model. There are three gauge groups used to describe transformations in the Standard Model: $U(1)$, $SU(2)$ and $SU(3)$. These are internal symmetries that leads to the conservation of certain quantum charges, and the quantum field theories built from them will be covered in the following sections. In addition to the internal symmetries, these quantum field theories are subject to external space-time symmetries. This give rise to energy-momentum conservation in space-time, and is contained in the 10 dimensional Poincare' group. This group will not be covered in this thesis.

2.2.3 Quantum Electrodynamics (QED)

Quantum Electrodynamics (QED) is the quantum field treatment of the electromagnetic interaction. Classically the electromagnetic field is expressed in terms of its electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$, whose dynamics is described by the Maxwell equations:

$$\mathbf{E}(\mathbf{x}) = -\nabla\phi(\mathbf{x}) - \frac{\partial \mathbf{A}(\mathbf{x})}{\partial t} \qquad \mathbf{B}(\mathbf{x}) = \nabla \times \mathbf{A}(\mathbf{x}) \tag{2.13}$$

Here, $\phi(\mathbf{x})$ is the scalar potential, and $\mathbf{A}(\mathbf{x})$ is the vector potential. Together they make up the four-vector potential $A_\mu(\mathbf{x}) = (\phi(\mathbf{x}), \mathbf{A}(x))$, which is more convenient to work with. The Lagrangian for the free photon field is given by:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.14}$$

where $F_{\mu\nu}$ is the electromagnetic field tensor:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{2.15}$$

It can be shown that the free photon Lagrangian in equation 2.14, as well as the fields in equation 2.13 are invariant under the following transformation of the 4-potential:

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu f(x) \tag{2.16}$$

f is here an arbitrary differentiable scalar function with space-time dependence. The importance of this transformation will soon become evident.

The free fermion fields are on the other hand described by the Dirac Lagrangian for spin-$\frac{1}{2}$ particles:

$$\mathcal{L}_0 = \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x) \tag{2.17}$$

Here, m is the mass of the particle, γ^μ are the Dirac gamma matrices ([4], p.46)[6] and $\bar{\psi}(x) \equiv \psi^\dagger(x)\gamma^0$. The Dirac Lagrangian is invariant under a global $U(1)$ transformation, that is, it does not change if we change

[6]The Dirac gamma matrices are a minimal set of 4×4 matrices satisfying the Clifford algebra; $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$, where $\{...\}$ is the anticommutator and $g^{\mu\nu}$ is the metric tensor.

the field by the infinitesimal phase θ:

$$\psi(x) \to \psi'(x) = e^{-qei\theta}\psi(x) \tag{2.18}$$

Going back to Noether's theorem, inserting the Dirac Lagrangian with the $U(1)$ phase transformation into equation 2.7 gives rise to the follwoing conserved current:

$$j^\mu = -qe\bar{\psi}\gamma^\mu\psi \tag{2.19}$$

The associated conserved electric charge yields:

$$Q = \int d^3x j^0 = qe \int d^3x \psi^\dagger\psi \tag{2.20}$$

The transformations considered so far have been global, meaning they are applied to all space-time points simultaneously and leads certain conservation laws. However, since the fields in the theory are locally defined, it would also seem natural to require invariance of the Lagrangian under local transformations. These are referred to as gauge transformations, and can be obtained by "promoting" θ to $\theta(x)$:

$$\psi(x) \to \psi'(x) = e^{iqe\theta(x)}\psi(x) \tag{2.21}$$

The Dirac Lagrangian itself is not invariant under the above transformation, but it can be if one introduces a new field which transforms as the photon field in equation 2.16. This can be done by substituting the derivative, ∂_μ, with the following covariant derivative D_μ:

$$\partial_\mu \to D_\mu = \partial_\mu + iqeA_\mu(x) \tag{2.22}$$

such that the Dirac Lagrangian yields:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + iqe\bar{\psi}A^\mu\psi = \mathcal{L}_0 + \mathcal{L}_{int} \tag{2.23}$$

This is indeed invariant under a local $U(1)$ gauge transformation, and the introduction of the photon field, A^μ, gives the Lagrangian an additional interacting term, $\mathcal{L}_{int}$. Thus, requiring the Dirac Lagrangian to be invariant under a local $U(1)$ transformation introduces an interaction between the photon field and the fermion fields. To get the complete Lagrangian for QED we just add the free photon Lagrangian from equation 2.14:

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.24}$$

The above Lagrangian for QED fully specifies all dynamics of electromagnetism in a way that is consistent with both quantum theory and special relativity. It has been subject to stringent experimental precision testing, where for one the predicted value for the anomalous magnetic moment of the electron agrees with the observed value to ten parts in billion [5]. As such, QED has proven to be a tremendously successful theory and was the template for the other gauge theories that followed in the Standard Model. We obtained this Lagrangian from enforcing local gauge invariance. This was done under a one dimensional transformation, but in general this can be done in any number of dimensions, which will become evident in the case of QCD and electroweak theory.

2.2.4 Quantum Chromodynamics (QCD)

Quarks are fermions that carry electric charge, so they couple to the electromagnetic field as described in the previous section. However quarks also carry color which introduces an additional coupling with the gluon fields. The theory describing this interaction is known as Quantum Chromodynamics, or QCD. It is a non-Abelian gauge theory built on the $SU(3)_c$ symmetry group. Quarks can carry three colors, denoted r, g, b, i.e red, green and blue respectively. Hence, the fermion Lagrangian in QCD contains a three component spinor Ψ^f in color space, taking the following form for each quark flavour:

$$\mathcal{L}_0 = \bar{\Psi}^f(i\gamma^\mu\partial_\mu - m_f)\Psi^f \tag{2.25}$$

The spinor term is now $\Psi^f = (\psi_r^f, \psi_g^f, \psi_b^f)$, where f denotes the flavour. $SU(3)_c$ is a special unitary group with dimension $n = 3$ and eight generators. These generators must follow the commutation relations given in equation 2.12. A set of matrices satisfying this relation are the eight 3×3 Gell Mann matrices ([6], p. 223), denoted λ_i. Following the same procedure as in QED, we require the Lagrangian to be locally invariant under the transformation of the gauge group. First, we do a local $SU(3)$ transformation of the fields Ψ^f:

$$\Psi^f(x) \to \Psi^{f\prime}(x) = e^{ig_s\lambda_j w_j(x)/2}\Psi^f \quad j = 1, 2..., 8 \tag{2.26}$$

Here, $w_j(x)$ are arbitrary real differentiable functions and g_s is the coupling constant. As was the case in QED, the Lagrangian is not invariant under this transformation. However if we in analogy with equation 2.22 and 2.23 replace the derivative ∂_μ with the following covariant derivative D_μ:

$$\partial_\mu \to D^\mu = \partial_\mu + \frac{1}{2}ig_s\lambda_j A_j^\mu \tag{2.27}$$

$\mathcal{L}_0$ changes to

$$\mathcal{L} = \bar{\Psi}^f(i\gamma_\mu D^\mu - m_f)\Psi^f = \mathcal{L}_0 + \mathcal{L}_{int} \tag{2.28}$$

where $\mathcal{L}_{int}$ is given by:

$$\mathcal{L}_{int} = -\frac{1}{2}g_s\bar{\Psi}^f\gamma_\mu\lambda_j\Psi^f A_j^\mu \tag{2.29}$$

The eight fields, A_j^μ, are identified as gluon fields, the mediators of the strong interaction. Introducing a term for the free gluons gives the full QCD Lagrangian:

$$\mathcal{L}^{QCD} = \overline{\Psi}(x)^f(i\gamma_\mu D^\mu - m_f)\Psi(x)^f - \frac{1}{4}G(x)_{i\mu\nu}G(x)_i^{\mu\nu} \tag{2.30}$$

$G_i^{\mu\nu}$ could in principle have the same form as the electromagnetic field strength tensor, $F^{\mu\nu}$, in equation 2.15, but due to non-abilian features of QCD this would not be gauge invariant. Instead an extra term is added to the expression for the free gluon fields:

$$G_i^{\mu\nu}(x) = F_i^{\mu\nu} + g_s f_{ijk}A_j^\mu(x)A_k^\nu(x) \tag{2.31}$$

Here, f_ijk is the structure constant of the Lie algebra. This extra term describes gluon self-interaction and gives rise to some interesting features in QCD. One such feature is confinement. Since gluons attract each other they reinforce the field as two quarks are torn apart with potential proportional to the distance. This changes the inverse square law at large r, where lattice QCD calculations suggests an infinite color

ionization potential. Hence, probing hadrons at high energy does not have the effect of separation, but rather the creation of quark-antiquark pairs, which again form more hadrons. This makes quarks impossible to isolate and explain why they only exist in color neutral states. The study of quarks and gluons is therefore done indirectly, mostly in scattering experiments where quarks or gluons can recoil energetically from hard collisions. The broken lines of force behind them lead to a cone of hadrons known as a jet, and by studying such jets one can obtain information about the initial particles.

2.2.5 Standard Electroweak theory

Electroweak theory is the theory describing the unification of the electromagnetic and weak interaction. In section 2.2.3 we saw how to treat the electromagnetic interaction as a $U(1)$ gauge theory. However, QED is actually a manifestation of a higher symmetry group which is broken at everyday energy scale. For sufficiently high temperatures, $\mathcal{O}(10^2)$ GeV, it is unified with the $SU(2)$ gauge group, describing the weak interaction. The resulting $SU(2) \times U(1)$ gauge theory was formulated by Glashow, Weinberg and Salam in the late sixties, for which they were awarded the Nobel prize in physics.

The theory of the weak interaction was initially proposed as an attempt to explain the apparent flavour change in β-decays. At quark level the process can be identified by the emission of a charged gauge boson changing an u-quark to a d-quark. Detailed studies of such flavour changing processes has shown that only left-handed (right-handed) fermions (anti-fermions) participate in the weak interaction. This comes from the experimental fact that the charged current weak interaction is parity violating, i.e not invariant under spatial reflection. To account for the left-handed structure of the interaction, the fermion fields need to be decomposed into its chiral states. The chirality of a fermion is defined by the operator $P_{L/R}$, projecting the fermion fields into a left or right-handed components:

$$P_{L/R}\psi(x) = \frac{1}{2}(1 \mp \gamma_5)\psi(x) = \psi(x)^{L/R} \tag{2.32}$$

For the remainder of the section we will consider lepton interactions, but the electroweak treatment of quarks is analogous. Using the projection operators, the Dirac Lagrangian can be decomposed into the following form for the leptons:

$$\mathcal{L} = \bar{\psi}_\ell^L i\partial\!\!\!/ \psi_\ell^L + \bar{\psi}_\ell^R i\partial\!\!\!/ \bar{\psi}_\ell^R + \bar{\psi}_{\nu_\ell}^L i\partial\!\!\!/ \psi_{\nu_\ell}^L = \bar{\Psi}_\ell^L i\partial\!\!\!/ \Psi_\ell^L + \bar{\psi}_\ell^R i\partial\!\!\!/ \psi_\ell^R \tag{2.33}$$

where:

$$\Psi_\ell^L(x) = \begin{pmatrix} \psi_{\nu_\ell}^L(x) \\ \psi_\ell^L(x) \end{pmatrix} \tag{2.34}$$

The mass term $m\bar{\psi}\psi$ is not gauge invariant under $SU(2)_L$ and does not appear in the Lagrangian. In section 2.3 the Higgs field will be introduced to the theory to resolve this. For now though, the fermions are taken to be massless. Furthermore, no right-handed term for the neutrinos appear in equation 2.33. Originally neutrinos were though to be massless which led to the assumption of left-handed neutrino states only. Later experiments has shown that they have a small but non-negligible mass, and although not observed, there

is theoretical motivation for the existence neutrinos with right chirality[7]. However, this will not be covered here, and as such, only massless left-handed neutrinos are considered.

Following the procedure of QED and QCD, the Lagrangian is required to be invariant under a local transformation of the symmetry group. $SU(2)$ has three generators, and a particular set of matrices satisfying the Lie algebra in equation 2.12 are the three Pauli matrices, τ_i:

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{2.35}$$

The free electroweak Dirac Lagrangian is invariant under the following global $SU(2)_L$ transformations:

$$\begin{aligned} \Psi_\ell^L &\to e^{i\frac{\tau_j}{2}\alpha_j}\Psi_\ell^L \quad j = 1,2,3 \\ \psi_\ell^R &\to \psi_\ell^R \end{aligned} \tag{2.36}$$

Inserting the this into the free Lagrangian and applying Noether's theorem (equation 2.7) leads to three conserved currents J_i^μ for the left-handed spinors.

$$J_i^\mu = \bar{\Psi}_\ell^L \gamma^\mu \alpha_i \frac{\tau_i}{2} \Psi_\ell^L \tag{2.37}$$

The three corresponding conserved charges are known as three different components of the weak isospin, denoted T_i. The third component, T_3, can be related to the electric charge Q by the Gell-Mann-Nishjima formula:

$$Q = T_3 + \frac{1}{2}Y \tag{2.38}$$

Here, Y is called the weak hypercharge. Since both Q and T_3 are conserved, so must Y. This introduces an additional $U(1)_Y$ symmetry to the theory which is needed for the unification with QED. $U(1)_Y$ transforms both left and right-handed spinors[8]. Applying the combined $SU(2)_L \times U(1)_Y$ local transformation to the spinors,

$$\begin{aligned} \Psi_\ell^L &\to e^{ig\frac{\tau_j}{2}\alpha_j(x)} e^{g'\frac{Y}{2}\beta(x)} \Psi_\ell^L \\ \psi_\ell^R &\to e^{g'\frac{Y}{2}\beta(x)} \psi_\ell^R \end{aligned} \tag{2.39}$$

and enforcing the electroweak Lagrangian to be locally gauge invariant yields the following covariant derivatives:

$$\begin{aligned} D^\mu \Psi_\ell^L &= \big(\partial^\mu + ig\frac{\tau_j}{2}W_j^\mu + ig'\frac{Y}{2}B^\mu\big)\Psi_\ell^L \quad j = 1,2,3 \\ D^\mu \psi_\ell^R &= \big(\partial^\mu + ig'\frac{Y}{2}B^\mu\big)\psi_\ell^R \end{aligned} \tag{2.40}$$

This introduces four gauge bosons, where the fields W_j^μ originates from $SU(2)_L$, and B^μ from $U(1)_Y$. Hence,the full leptonic electroweak Lagranigan is given as:

$$\mathcal{L}_{EW}^{lep} = \bar{\Psi}_\ell^L \gamma_\mu \big(\partial^\mu + ig\frac{\tau_j}{2}W_j^\mu + ig'\frac{Y}{2}B^\mu\big)\Psi_\ell^L + \bar{\psi}_\ell^R i\gamma_\mu \big(\partial^\mu + ig'\frac{Y}{2}B^\mu\big)\psi_\ell^R + \mathcal{L}_{kin} \tag{2.41}$$

Here, $\mathcal{L}_{kin}$ is the gauge invariant kinetic term describing the free fields:

$$\mathcal{L} = -\frac{1}{4}B_{\mu\nu}(x)B^{\mu\nu}(x) - \frac{1}{4}W_{\mu\nu}^j(x)W_j^{\mu\nu}(x) \tag{2.42}$$

[7] Usually this is referred to as "sterile neutrinos", and the search for them is an active area of particle physics.

[8] $Y_L = -1$ for left-handed leptons and $Y_R = -2$ for right-handed leptons.

with:

$$\begin{aligned} B^{\mu\nu} &= \partial^\mu B^\nu + \partial^\nu B^\mu \\ W_j^{\mu\nu} &= \partial^\mu W_j^\nu - \partial^\nu W_j^\mu - g\epsilon_{jkl} W_k^\mu W_l^\nu \end{aligned} \tag{2.43}$$

The two gauge fields W_1^μ and $W^2\mu$ can be written as a linear combination of the physical fields W^μ and $W^{\mu\dagger}$. These are interpreted as the charged gauge bosons $W^\pm$, with:

$$\begin{aligned} W^\mu &= \frac{1}{\sqrt{2}}(W_1^\mu - iW_2^\mu) \\ W^{\mu\dagger} &= \frac{1}{\sqrt{2}}(W_1^\mu + iW_2^\mu) \end{aligned} \tag{2.44}$$

For the electromagnetic interaction to be unified with the weak interaction in this model, the electromagnetic term $ieQA^\mu$ must be contained in the neutral term $i(g\frac{\tau_3}{2}W_3^\mu + g'\frac{Y}{2}B^\mu)$ in equation 2.41. Therefore the fields W_3^μ and B^μ must be linear combination of A^μ and an additional neutral field Z^μ.

$$\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_w & \sin\theta_w \\ -\sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} B^\mu \\ W_3^\mu \end{pmatrix} \tag{2.45}$$

Here, θ_w is the weak mixing angle. In the unbroken theory, i.e at high temperatures, writing out this linear combination is purely a formal exercise. There is no physical argument for using one set of fields rather then the other. However, at lower energies where the symmetry is broken, A and Z become physically distinct mass eigenstates. θ_w is the angle by which this breaking of symmetry rotates the original W_3 and B gauge boson plane into A and Z.

Experimentally it has been observed that both Z^0 and $W^\pm$ are very massive particles. However, naively adding mass terms to the Lagrangians leads to conflicts with gauge invariance for both the fermions and gauge bosons. Hence, an additional mechanism, namely the Higgs mechanism, needs to be introduced to the theory to account for this.

2.3 The Brout-Englert-Higgs Mechaninsm

The Brout-Englert-Higgs mechanism, or more commonly just the Higgs mechanism, circumvents the zero-mass constraint in the electroweak model, and restores the fermion and gauge boson mass terms through a process known as spontaneous symmetry breaking. This is done by introducing a $SU(2)$ complex scalar doublet Φ to the theory:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{2.46}$$

whose potential takes the following form:

$$V(\Phi) = -\mu^2|\Phi|^2 - \lambda|\Phi|^4 \tag{2.47}$$

Here, λ is a real positive constant determining the strength of ϕ self interactions. For $\mu < 0$, the ground state of the potential does not occur at 0. Rather there are infinite possible ground states at $|\Phi|^2 = -\frac{1}{2}\mu^2/\lambda$. This can be seen in figure 2.2 where one should imagine the potential being wrapped around the vertical axis.

In the context of QFT, a non-zero ground state corresponds to a non-vanishing vacuum expectation value (vev) of $|\Phi|$, which is given by the modulus $v = (-\mu^2/2\lambda)^{\frac{1}{2}}$. This vev selects a preferred direction in the weak isospin-hypercharge space, which in turn "spontaneously breaks" the $SU(2)_L \times U(1)_Y$ gauge symmetry. Due to the ambiguity of the ground state any direction is valid, but to keep the redundancy and degrees of freedom in the theory at a minimum it is convenient to work with the so-called unitary gauge. The fields are then redefined to be:

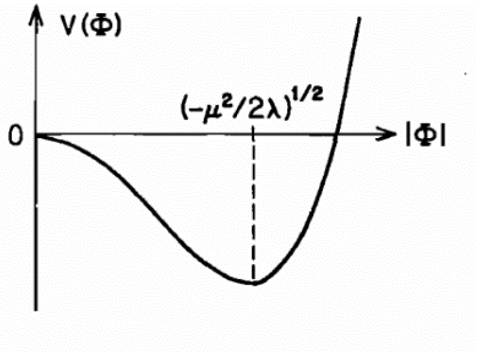


Figure 2.2: Potential of a scalar field with $\mu < 0$. Taken from [7].

$$\Phi = exp\left(\frac{i\xi_j(x)\tau_j}{2v}\right)\begin{pmatrix}0\\ \frac{v+h}{\sqrt{2}}\end{pmatrix} \quad j = 1,2,3 \tag{2.48}$$

Here, ξ_j are three real fields, τ_j the three Pauli matrices and h is the scalar Higgs field. By applying a $SU(2)_L$ gauge transformation with $\alpha_j(x) = \xi_j(x)$, we can remove the phase factor term in the above expression. As such the ξ degrees of freedom seem to vanish, but do essentially reappear as the longitudinal degrees of freedom of the massive gauge bosons when they acquire mass. By applying the electroweak covariant derivative in equation 2.40 to Φ, couplings with the Higgs field generate mass terms for the $W^{\pm}$ and Z^0 gauge bosons. Mass terms for the fermions are generated through a Yukawa coupling. For example, the Yukawa term for the leptons is given by:

$$\mathcal{L}_y = y_\ell \Psi_\ell^L \Phi \Psi_\ell^R + y_\ell \Psi_\ell^R \Phi \Psi_\ell^L \tag{2.49}$$

Plugging in the values for Φ after symmetry breaking yields:

$$\begin{aligned}\mathcal{L}_y &= \frac{y_\ell}{\sqrt{2}}\begin{pmatrix}\bar{\nu}_\ell^L & \bar{\ell}^L\end{pmatrix}\cdot\begin{pmatrix}0\\ v+h\end{pmatrix}\ell^R + \frac{y_\ell}{\sqrt{2}}\ell^R\begin{pmatrix}0 & v+h\end{pmatrix}\cdot\begin{pmatrix}\bar{\nu}_\ell^L\\ \bar{\ell}^L\end{pmatrix} \\ &= \frac{y_\ell v}{\sqrt{2}}(\bar{\ell}\ell) + \frac{y_\ell}{\sqrt{2}}h(\bar{\ell}\ell)\end{aligned} \tag{2.50}$$

The first term in the above equation can be interpreted ass a mass term for the considered lepton and the second as an interaction between the lepton field and the Higgs field. The mass of the lepton is hence given as $m_\ell = y_\ell v/\sqrt{2}$, making it directly proportional to the Higgs coupling. This is true for all fermions in the Standard Model, with the Higgs coupling strongest to top quarks, then to bottom quarks etc. However, the top quark decay is kinematically forbidden as the top quark is heavier than the Higgs mass of 125 GeV, giving bottom quarks the highest decay branching fraction. The decay studied in this thesis is $h \to \tau\tau$, with a corresponding branching fraction of 6.27×10^{-2} [3].

In 2012, CERN reported findings of a scalar spin-0 boson with mass 125 GeV at an excess of 5σ [8]. This turned out to be the aforementioned Higgs boson, and two of the creators of the theory, namely, Peter Higgs and Francois Englert were awarded the Nobel prize in physics the following year[9]. With this finding, all particles predicted by the Standard Model have been experimentally verified.

[9]Robert Brout (1928-2011), a Belgian theoretical physicist, was also heavily involved in the creation of the theory, but he sadly passed away a year before the Higgs finding in 2012 and hence did not receive the price.

2.4 Shortcomings in the Standard Model

The Standard Model has proven to be an extremely successful theory, showing time and time again astonishing agreement with experimental data. However, we know it is not the full explanation of the universe as there still are theoretical "unnaturalness" and observed phenomena it can not account for. The SM as it stand today is hence not complete, and in the following a short summary of some of its problems are listed:

- **Free parameters**: As it stands today, the SM contains 19 parameters that needs to be integrated to the theory from experimental data. These include the masses of the elementary particles, the gauge couplings and the Higgs vacuum expectation value among others.

- **Gravity**: The SM does not explain gravity. In principle this can be solved by adding a spin-2 graviton to the model as the gravitational mediator, but this only works in effective theories that are valid over some specific range of scales [9]. General relativity, the theory of gravitation, seem to have problems with extreme scales, and the two theories become incompatible in the high energy limit.

- **Energy densities in the universe**: Baryonic matter explained by the SM only makes up about 5% of the total energy density of the universe. About 26% of the universe should be Dark Matter, which is the main topic of this thesis and will be explained in the next chapter. The remaining 69% should be dark energy, which is a constant energy density of the vacuum [10]. Attempts to calculate dark energy from the vacuum energy in the SM leads to a mismatch of $\mathcal{O}(10^{120})$, and predicts the radius of the universe to be ~ 32.7 Km [11].

- **Hierarchy problem** The measured mass of the SM Higgs boson is ~ 125 GeV. The mass parameter in the Higgs mass term, $m^2 h^\dagger h$ can be altered by radiative corrections, with the corrective terms being proportional to the scale at which it interacts. In principle this scale can go all the way up to the Planck scale $\Lambda \sim 10^{19}$ GeV, which leads to SM prediction for the Higgs mass well beyond the measured value. Currently this is resolved by numerical cancellations in the corrective terms, but numerical cancellations are deemed highly unnatural and require extreme fine tuning of the SM parameters.

- **Neutrino masses**: According to the Standard Model all neutrinos are massless. However, it has been experimentally verified from neutrino oscillation experiments that neutrinos do have mass. This can in principle be added to the Standard Model through the Higgs mechanism, assuming that right-handed neutrinos exists. However, as of today, there is now experimental evidence for this.

- **Grand Unification**: The coupling constants for the three interactions in the SM appear to be changing with energy. By extrapolating these one can obtain an almost obtain unification at the scale $\mathcal{O}(10^{16})$ GeV. If, however, exact unification of the three interactions is possible it would mean that they are all only distinct low energy aspects of the the same underlying theory. In general unification can be obtained by increasing the symmetry group of the SM. Many such theories exists, like $SU(5)$ and $SO(10)$, which indeed seem to give a scale where all interactions are unified. If this is the case, it could explain the origin of the coupling values, why there are three generations of fermions and why there are so many free parameters in the SM. Collectively, these theories are referred to as Grand Unified

Theories (GUTs), and they seem to be a natural ingredient in the search for the holy grail of physics, namely a theory of everything.

Chapter 3

Dark Matter

Early observations of the motion of stars and galaxies pointed towards gravitational effects coming from non-visible matter. This led to the introduction of the phenomena known as "Dark Matter" (DM), meaning matter that interacts gravitational, but not electromagnetically. In the Standard Model there is in fact a particle with these properties, namely the neutrinos. However, calculations of upper limits on neutrino masses, as well as the neutrino abundance in the universe, suggest that neutrinos alone can not account for the observed gravitational effects. Hence, the particle nature of DM is still very much a mystery and not explained by the Standard Model. Regardless, the DM identity is the subject of many theories beyond the Standard Model. One of the leading DM candidates are so-called Weakly Interacting Massive Particles (WIMPs) [12], which do satisfy the DM constraints set by cosmological observations. They do arise in many theories like the lightest supersymmetric particle (LSP) in supersymmetric (SUSY) theories [13], and the lightest CP odd particle in some two Higgs doublet models (2HDMs) [14]. It is worth noting that none of these models were initially proposed to solve the Dark Matter problem, but seem to give rise particles exhibiting WIMP properties naturally. The outline of this chapter is as follows; first a summary of evidence for DM, followed by WIMP mono-X theories and their signatures in collider searches, and finally an overview of the 2HDM+a signal model searched for in this analysis .

3.1 Astrophysical evidence for Dark Matter

The most straightforward evidence for Dark Matter comes from the studies of stellar rotation curves in galaxies. The tangential stellar velocities (v) from Newtonian mechanics is simply given by:

$$v = \sqrt{\frac{G_N M(r)}{r}} \tag{3.1}$$

with G_N being the gravitational constant, $M(r)$ the mass density, and r the distance from the galactic center. For the visible matter within the galaxy, the matter density $M(r)$ is expected to be constant for large r, i.e $v \propto 1/\sqrt{r}$. However, from studies of galactic rotation curves, like NGC 3198 [15], the observed tangential velocities behaves very much differently, indicating an almost constant velocity profile at large radii as seen in figure 3.1. By equation 3.1 this implies the existence of an increasing non-visible matter density as r

increases. Figure 3.1 shows strong evidence for this by fitting the observations with the combined expected Newtonian velocity distributions from visible and Dark Matter. While the disk curve models the data well at short r, the agreement falls of for as one moves further from the galactic center, where the discrepancy can be explained by a halo of Dark Matter enveloping the galactic disk.

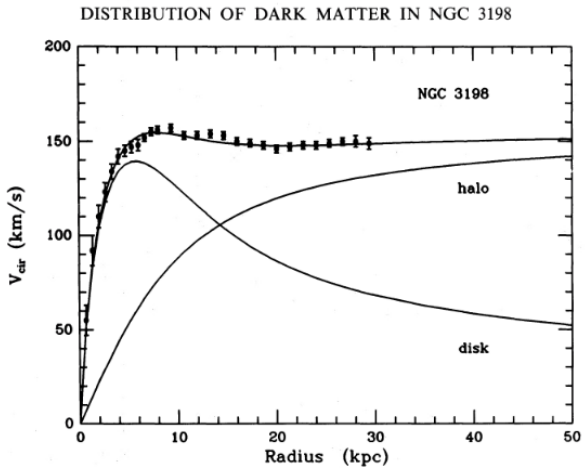


Figure 3.1: Rotation curve of NGC 3198 showing agreement between data and combined fit from velocity distributions of visible matter ("disk") and Dark Matter ("halo") [15] .

While the stellar distributions can be explained by a Dark Matter halo, it is not the only possible explanation. Alternative models to the DM picture have been put forward by explaining the phenomena by using modified theories of gravity [16]. Here no new form of matter needs to be introduced, but results from gravitational lensing experiments such as studies of the bullet cluster [17] seem to disfavor this perspective.

Gravitational lensing arises in general relativity when light rays pass through regions of space subject to the presence of massive objects. These massive objects cause spacetime to curve. As light rays always follow the shortest path between two points on a curve this causes a deflection by an angle $d\theta$. The amount of deflection is proportional to the mass of the lens, and hence provides a tool for probing the masses of galactic objects. Figure 3.2 shows this effect in the Bullet cluster, containing two colliding clusters of galaxies. The major matter components in the cluster pair, i.e stars, gas and Dark Matter behave differently. The stars pass mostly straight through the collision only affected by the gravitational interaction with other galaxy objects. The gaseous matter, observed in X−ray, interacts and slows down (Highlighted in red), and represent the majority of baryonic matter in the cluster pair. In the absence of Dark Matter the gravitational potential would trace this region.

However, the total mass distribution of the cluster pair inferred by gravitational lensing (Highlighted in blue) shows that the major matter component pass through unaffected by the collision. As the stars alone can not account for this mass, it fits well with the assumption that the galaxies mostly are comprised of Dark Matter, which may interact gravitationally but not electromagnetically.

The Standard Model of cosmology (ΛCDM) also relies on the existence of Dark Matter in order to preserve the uniformity of the photons in the Cosmic Microwave Background (CMB). These photons originated from the

Figure 3.2: Observation from the Chandra X-ray Observatory showing the Bullet cluster. The red region corresponds to hot gas observed in X−ray, total matter distribution inferred by gravitational lensing is highlighted in blue[17].

era of the universe when the first neutral atoms were formed, about 300 000 years after the Big Bang. ΛCDM can be parametrized in terms of the energy densities of the universe, i.e $\Omega_b, \Omega_{DM}, \Omega_\Lambda$ for baryonic matter, Dark Matter and Dark Energy respectively. Values for these parameters can be obtained by measuring the temperature fluctuations in the CMB. Measurements from the WMAP[18] and Planck experiment [10] fitted to the ΛCDM model shows that less than 5% of the energy content of the universe comes from baryonic matter as described by the Standard Model in section 2. The rest is "dark", where roughly 26% comes from Dark Matter, and the remaining 69% from Dark Energy[10].

From these results it is clear that Dark Matter plays a significant role in understanding the universe. Large parts of physics is therefore devoted to the understanding of the particle nature of Dark Matter, making searching for it a main priority of the field today.

3.2 Weakly Interacting Particles (WIMPs)

As discussed in the previous section there are strong experimental evidence in favor of the existence of Dark Matter. However, these experiments does not provide any insight into what constitutes the particle nature of DM. One attractive hypothetical category of particles that has received much recognition are so-called Weakly Interacting Massive Particles (WIMPs), denoted χ. There is no precise definition of what would constitute a WIMP, but they are subject to constraints set by observations. In general a WIMP would have to be a new fundamental particle beyond the Standard Model with mass range around the weak scale, i.e 1 to 10^5 GeV, only subject to gravitational interaction and either the weak interaction or possibly a new interaction weaker or as weak as the weak interaction. The lifetime of a WIMP needs be comparable to the age of the universe as to not have decayed at present times. Furthermore, the particle needs to be non-relativistic, often referred to as "cold", in order to explain the structure formations and clustering in the universe. Particles exhibiting these characteristics could explain the observed abundance of DM assuming

[10]Dark Energy will not be covered in this thesis, but in essence it acts as a sort of "anti-gravity", responsible for the present day acceleration of the universal expansion.

they went through a mechanism known as "thermal freeze out". That is, if the temperature in the early universe was sufficiently high to keep WIMPs in thermal equilibrium with the Standard Model particles this would have given rise to the following processes:

$$\chi\chi \leftrightarrow \text{SM}$$

indicating WIMP pair creation and annihilation. The corresponding WIMP annihilation would be given by:

$$\Gamma_{ann} = n_\chi \langle \sigma(\chi\chi \rightarrow SM) v_\chi \rangle \tag{3.2}$$

where n_χ is the WIMP number density, σ the annihilation cross section, v the relative velocity between the WIMPs, and the brackets $\langle ... \rangle$ indicate thermal averaging. As the universe expanded and consequently cooled, the WIMP abundance dropped exponentially, making it increasingly "difficult" for the WIMPs to interact as the space between them got larger. At some point the annihilation rate Γ_{ann} dropped bellow the Hubble expansion rate of the universe, forcing the WIMPs out of thermal equilibrium. As a consequence, the WIMPs would no longer be able to annihilate, keeping the Dark Matter abundance essentially unchanged after this point, as shown in figure 3.3.

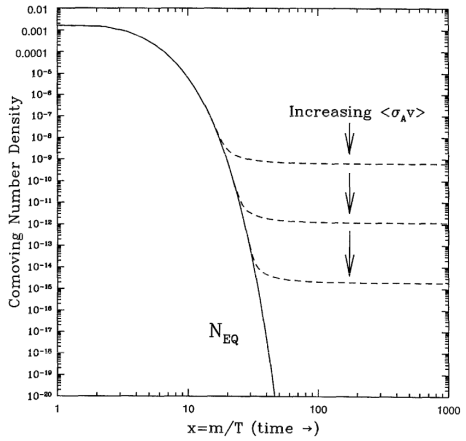


Figure 3.3: DM abundance as a function of WIMP mass and temperature (m_χ/T). Decreasing temperature indicate the time evolution. The solid and dashed line corresponds to the abundance before and after freeze out, respectively [13].

By applying the freeze out condition $\Gamma_{ann} < H$ to equation 3.2 one can calculate the expected abundance of DM today, referred to as the "relic abundance". The measured relic abundance from WMAP [18] yields $\Omega_c h^2 = 0.1131 \pm 0.0034$, whereas the freeze out model predicts [13]:

$$\Omega_c h^2 \approx \frac{3 \times 10^{-27} cm^3 s^{-1}}{\langle \sigma_{ann} v_\chi \rangle} \tag{3.3}$$

The measured relic density can be obtained from the above expression by requiring an annihilation cross section of $\langle \sigma_{ann} v \rangle \approx 3 \times 10^{-26} cm^3 s^{-1}$, which indeed is at the weak scale. This strongly suggests that if a stable particle exists, and it interacts via a force at the weak scale, it most likely would be Dark Matter. This apparent coincidence is known as the "WIMP miracle", and has provided strong motivation for WIMP searches. Throughout the rest of this thesis DM is assumed to be of WIMP nature, and the two terms may be used interchangeably.

3.3 Mono-X collider searches

If DM particles interact with SM particles, it could be possible to both produce DM and study its properties in a laboratory setting. This would imply the production of the mediator connecting the dark and SM sector. The pseudo Feynman diagram in figure 3.4 serves as an illustrative point for three approaches to DM detection. The interaction in the diagram is not specified, and depends on the specific model postulated.

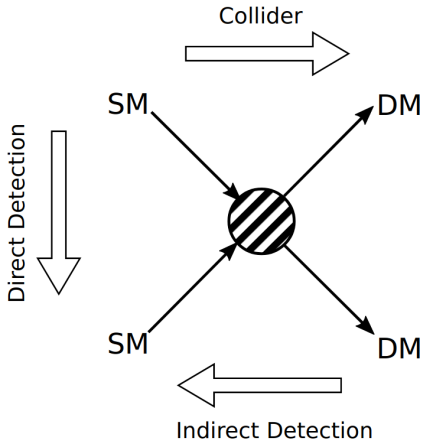


Figure 3.4: Pseudo Feynman diagram showing potential interaction between dark and SM sector. Taken from [19].

Treating the time axis in the diagram from left to right would imply DM annihilation into SM particles. This is a processes one can imagine takes place in parts of the universe where DM is at its densest. Therefore these types of searches are primarily carried out by astrophysical experiments, and referred to as "indirect searches". If one "flips" the diagram, and treats the time axis as vertical, this would involve a scattering process. Here, DM scatters of SM particles similarly to how neutrinos scatters of atomic nuclei. These types of searches are known as "direct searches". The problem with both direct and indirect searches is that they are subject to many astrophysical uncertainties, as they rely on cosmic data. When taking the time axis from left to right, the diagram represent SM annihilation into SM particles. This is a process likely to occur at particle colliders, and hence referred to as "collider searches". In contrast to indirect and direct searches, collider searches provides the possibility to study of DM in a controlled environment. Colliders are also particular sensitive to low DM masses of a few GeV, which is a region inaccessible with current direct detection techniques.

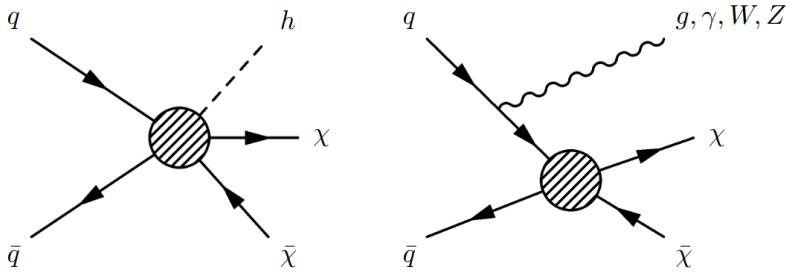


Figure 3.5: Feynman diagrams showing DM production is association with an ISR gluon, photon, Z, W or a Higgs boson.

While figure 3.4 qualitatively illustrates how DM might be produced, it does not give any insight into how DM might be detected in collider experiments. Particle detectors rely on electromagnetic interactions of charged particles, making DM invisible to such an apparatus. However, if DM is produced in association with a SM particle, it is possible to infer the the presence of DM particles by the imbalance in the total transverse momentum, referred to as "missing transverse energy" E_T^{miss}. For example, at the LHC initial state particles can radiate hard gluons, photons or heavier particle like massive gauge and Higgs bosons through a process known as initial state radiation (ISR). Feynman diagrams for such processes can be seen in figure 3.5.

If the DM particle pair is produced with enough momentum, it will recoil with the ISR object. This induces a distinct $E_T^{miss} + X$ signature in the detector, with X being the recoiling SM particle. Such signatures are commonly referred to as "mono-X", and offers a spectrum of processes for probing DM. There is a wide variety of mono-X searches that has been performed at the LHC, and a review from 2018 be found in [20]. For this thesis however, the primary focus will be mono-Higgs. It should be noted that even though this same ISR mechanism could give rise to a mono-Higgs signature, it is heavily suppressed by the Yukawa coupling of Higgs to lighter quarks. We shall see in section 3.6 that an extension of the Higgs sector is necessary to construct a model that predicts a mono-Higgs signature, which will be the theoretical foundation this thesis is based on.

3.4 Effective Field theories

The processes in figure 3.5 can be formulated mathematically as contact interactions in the framework of effective field theories (EFTs). EFTs are simplifications of complete quantum field theories, and is built on the assumption that the energy scale of the mediator is much larger than the momentum transfer involved in the interaction. This makes it possible to integrate out the mediator and paramterize the interaction in terms of effective operators. To illustrate this it can be illuminating to consider a massive scalar mediator S interacting with two quarks with coupling strength g_q, and with DM with coupling strength g_χ. This could be a process occurring in a collider, where DM pair production is induced by quark pair annihilation. The corresponding Lagrangian would contain a mediator mass term as well as S coupling terms:

$$\mathcal{L} \supset \frac{1}{2}M^2S^2 - g_q q\bar{q}S - g_\chi \chi\bar{\chi}S \tag{3.4}$$

The full theory would involve a propagator term for $S \propto \frac{1}{Q^2-M^2}$, where Q is the momentum transfer and M the mass of the mediator. Expanding this expression in terms of $\frac{Q}{M}$ yields:

$$\frac{1}{Q^2 - M^2} = -\frac{1}{M^2} + \frac{Q^2}{M^2} + \mathcal{O}(\frac{Q^4}{M^4}) \tag{3.5}$$

Assuming that the energy scale of the interaction is much lower than the mediator mass, $Q << M$, simplifies the propagator term to be $\sim -1/M^2$. Effectively this reduces the interaction to a contact interaction, which can be seen figure 3.6 where the Feynman diagram on the right omits the mediator and is solely described by the initial and final state particles.

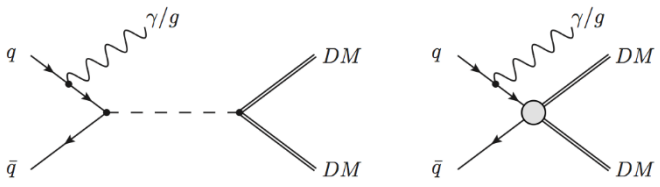


Figure 3.6: Feynman diagrams for DM pair production with a scalar mediator. The left side indicates a more complete theory where the mediator is considered, while the right side correspond to the limit where this is reduced to the EFT contact interaction.

The effective coupling factor can be obtained by multiplying the propagator term $-1/M^2$ with the quark and DM coupling terms:

$$\frac{1}{\Lambda^2} = \frac{g_q g_\chi}{M^2} \tag{3.6}$$

where Λ represent the relevant energy scale. The contact interaction operator $\mathcal{O}_s$ can now be parametrized in terms of the effective coupling and the particle fields involved in the interaction:

$$O_S = \frac{1}{\Lambda^2}(\chi\bar{\chi})(q\bar{q}) \tag{3.7}$$

The above equation parametrizes the theory in terms of a single contact interaction, and is the mathematical formulation represented by the diagram on the right in figure 3.6. However, there are several constraints on the validity of EFTs that needs to be considered. The energy scale Λ sets the maximum limit to where EFTs can be trusted, which will break down and lead to nonphysical results at higher energies. Hence, Λ is constrained by $\Lambda >> \frac{Q}{\sqrt{g_q g_\chi}}$. This is unproblematic in direct search experiments where the momentum transfers from nuclear scattering is of order KeV. In contrast typical momentum transfers in colliders are in the GeV-TeV range. An EFT approach in this energy regime would imply mediator masses beyond the energy reach of current colliders, and their comparison to direct searches lead to nonphysical DM masses. A more complete theory is therefore needed to make quantitative predictions in a collider setting. This is discussed in greater detail in [21], and motivates the use of simplified models as an alternative to the EFT picture.

3.5 Simplified models

An alternative approach to EFTs are so called simplified models. These models expands the contact interaction from EFT, including additional degrees of freedom and at least one mediating particle. By increasing the number of parameters in the theory one gains more theoretical control than in EFT. Simplified models are designed to involve as few parameters as possible while still being complete enough to accurately describe relevant physics at the energy scales probed by colliders. Compared to EFTs which integrates out all but the DM particle, a simplified model can be thought of as the limit of a more complete theory which integrates out all but the lightest degrees of freedom in the dark sector. The resulting model is then often composed of one stable DM candidate and at least one mediator, and can usually be represented by tree-level Feynman diagrams. An overview of DM simplified models used in early LHC Run-2 searches can be found in

[22]. For this thesis however the set of models in focus will be those where the mediator sector is composed of an extended Higgs sector with an additional spin-0 particle.

3.6 The Two-Higgs-Doublet-Model and 2HDM+a

A minimal extension to the Higgs sector in sector 2.3 is the Standard Model with two Higgs doublets. The following will be a breif overview, for a more complete description see [23] section 4.1 . This model introduces two complex $SU(2)_L$ doublets, Φ_1 and Φ_2:

$$\Phi_1 = \begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix} \quad \Phi_2 = \begin{pmatrix} \phi_2^+ \\ \phi_2^0 \end{pmatrix} \tag{3.8}$$

In general both the doublets acquire vacuum expectation values (vevs) v_1 and v_2 after $SU(2)_L \times U(1)_Y$ symmetry breaking. The vevs must lies in the neutral component of the doublets as to retain the $U(1)_{em}$ symmetry of electromagnetism:

$$\Phi_1 = \begin{pmatrix} \phi_1^+ \\ (h_1 + v_1 + i\eta_1)/\sqrt{2} \end{pmatrix} \quad \Phi_2 = \begin{pmatrix} \phi_2^+ \\ (h_2 + v_2 + i\eta_2)/\sqrt{2} \end{pmatrix} \tag{3.9}$$

and are related to the total vev v from the SM Higgs sector by the following relation:

$$v = \sqrt{v_1^2 + v_2^2} \approx 246.22 \text{ GeV} \tag{3.10}$$

In total there are eight fields in the model consisting of two charged scalar fields $\phi_1^\pm$ and $\phi_2^\pm$, two CP-even real scalar fields, h_1 and h_2, and two CP-odd real scalar fields η_1 and η_2. A parametrization of the aforementioned fields reveals five physical Higgs fields, and three unphysical Goldstone bosons which can be gauged away. A particular method for deriving this result can be found in [24]. The five orthogonal states after symmetry breaking correspond to the five physical Higgs fields, and are given as follows:

$$H^\pm = -\sin\beta\phi_1^\pm + \cos\beta\phi_2^\pm, \qquad A^0 = -\sin\beta\eta_1 + \cos\beta\eta_2 \tag{3.11}$$

$$h^0 = -\sin\alpha h_1 + \cos\alpha h_2, \qquad H^0 = \cos\alpha h_1 + \sin\alpha h_2 \tag{3.12}$$

The fields in equation 3.11 are the CP-odd mass eigenstates, where $H^\pm$ are charged scalars and A^0 is a neutral pseuodoscalar. The fields in equation 3.12 are the CP-even mass eigenstates, with h^0 and H^0 both being neutral scalars. The mixing angle β is defined according to the vevs:

$$\tan\beta = \frac{v_1}{v_2} \tag{3.13}$$

and the mixing angle α can be shown to be $\alpha = \beta - \pi/2$ by assuming that h^0 is the SM Higgs boson with mass $m = 125$ GeV. This assumption is known as the alignment limit. There are now several ways to construct a Yukawa structure for the model, depending on the type of 2HDM in question. Considered here, is the Type-II 2HDM, where Φ_1 couples to down-type quarks and fermions, and Φ_2 couples to up-type quarks. The corresponding Lagrangian describing the coupling between the doublets and the fermions can be written as follows:

$$\mathcal{L}_Y = -\big(\bar{Q}Y_u\tilde{\Phi}_2 u_R + \bar{Q}Y_d\Phi_1 d_R + \bar{L}Y_\ell\Phi_1\ell_R + h.c\big) \tag{3.14}$$

Here, Y_f are Yukawa matrices acting on the three fermion generations f. Q and L are left handed quark and lepton doublets, u_R, d_R and ℓ_R are the right handed fermion singlets, and $\tilde{\Phi}_1$ is the conjugate Higgs doublet[11].

To Obtain the 2HDM+a model an additional pseudoscalar singlet P is added to the model which provides a coupling to a DM particle χ. Following [25] the only possible way to construct this $P-\chi$ coupling is given by:

$$\mathcal{L}_\chi = -iy_\chi P\bar{\chi}\gamma_5\chi \tag{3.15}$$

where χ is a dark Dirac fermion, and y_χ is the coupling strength. The potential of the scalars can now be written out in three terms: describing the potential for the Higgs doublets, interactions between the Higgs doublets and the pseudoscalar singlet and a potential term for the pseudoscalar singlet respectively:

$$V = V_\Phi + V_{\Phi P} + V_P \tag{3.16}$$

which explicitly yields [25]:

$$\begin{aligned} V_\Phi =& \mu_1\Phi_1^\dagger\Phi_2 + \mu_2\Phi_2^\dagger\Phi_2 + \big(\mu_3\Phi_1^\dagger\Phi_2 + h.c\big) + \lambda_1(\Phi_1^\dagger\Phi_1)^2 + \lambda_2(\Phi_2^\dagger\Phi_2)^2 \\ &+ \lambda_3(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) + \lambda_4(\Phi_1^\dagger\Phi_2)(\Phi_2^\dagger\Phi_1) + \big[\lambda_5(\Phi_1^\dagger\Phi_2)^2 + h.c\big] \end{aligned} \tag{3.17}$$

$$V_{\Phi P} = P\big(ib_p\Phi_1^\dagger\Phi_2 + h.c\big) + P^2\big(\lambda_{p_1}\Phi_1^\dagger\Phi_1 + \lambda_{P2}\Phi_2^\dagger\Phi_2\big) \tag{3.18}$$

$$V_P = \frac{1}{2}M_p^2P^2 \tag{3.19}$$

Here, the μ's are parameters with mass dimension, the λ's are dimensionless parameters determining quartic coupling strengths, and b_P determines the trilinear portal coupling strength of P and the Higgs doublets. By rotating to the mass basis the parameters $m_P, \mu_1, \mu_2, \mu_3, b_p, \lambda_1, \lambda_2, \lambda_4, \lambda_5$ can be rewritten in terms of masses, vevs, couplings and mixing angles of the physical fields. In addition to the physical mass eigenstates $H^\pm$, h and h^0 from equation 3.11 and 3.12, the rotation reveals two additional pseudoscalar mass eigenstates, namely a and A. These are orthogonal states coming from mixtures of the CP-odd pseudoscalars A^0 and P with mixing angle θ:

$$\begin{aligned} A &= \cos\theta A^0 + \sin\theta P \\ a &= \cos\theta P - \sin\theta A^0 \end{aligned} \tag{3.20}$$

The two fields above provide a link between the dark and extended Higgs sector. The coupling of a and A to the SM fermions occurs via the Yukawa couplings in equation 3.14. The physics of the broken 2HDM+a simplified model is hence fully captured by the mixing angles β, α, θ, the vev v, the quartic couplings $\lambda_3, \lambda_{P1}, \lambda_{P2}$, DM coupling y_χ and the masses M_h, M_H, $M_{H^\pm}, M_A, M_a, m_\chi$. These parameters are subject to constraints that arise from Higgs and flavour physics, mono-H LHC searches and electroweak precision

[11] This comes from a property of $SU(2)$: $\tilde{\Phi} = i\tau_2\Phi_1$ transforms the same way as Φ under a $SU(2)$ transformation where τ_2 is the second Pauli matrix.

measurements. This is discussed in [25] and motivates a particular choice of benchmark parameters:

$$\begin{gathered} M_H = M_A = M_{H^\pm}, \quad m_\chi = 10\ GeV \\ \cos(\beta-\alpha) = 0, \quad \tan\beta = 1, \quad \sin\theta = 0.35, \\ y_{\chi=1}, \quad \lambda_1 = \lambda_{P1} = \lambda_{P2} = 3 \end{gathered} \tag{3.21}$$

The 2HDM+a model give rise to a variety mono-X signatures from the allowed couplings in the above Lagrangians. One such signature is where the mediator a is resonantly produced along with a SM particle X. The a then decays into a DM pair $\chi\bar{\chi}$, inducing a mono-X signal. There are several channels that allow for this, but some relevant examples are $h + E_T^{miss}$, $Z + E_T^{miss}$ and $tW + E_T^{miss}$. Feynman diagrams for these processes can be seen in figure 3.7.

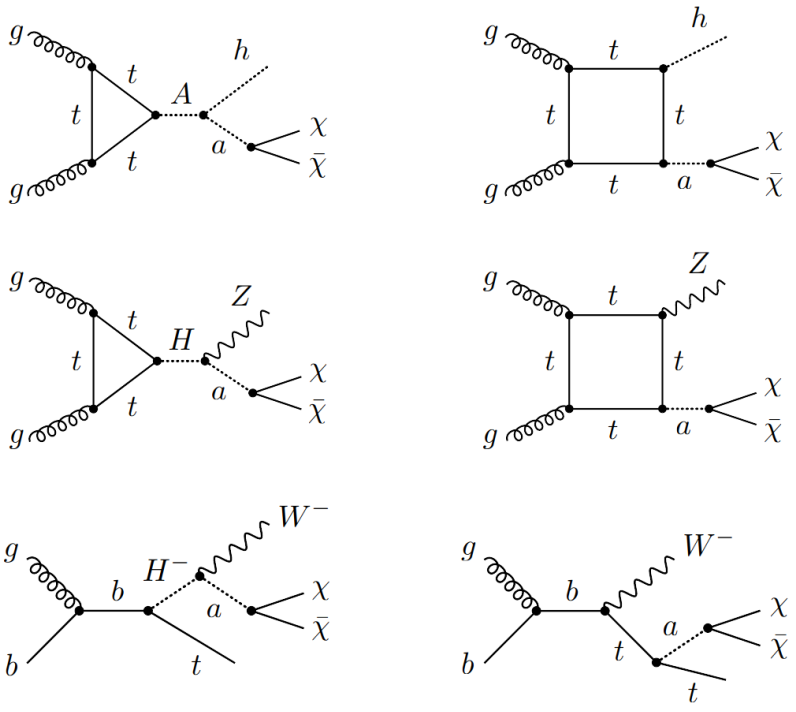


Figure 3.7: Peuodo-Feynman diagrams giving rise to $h + E_T^{miss}$ (top row), $Z + E_T^{miss}$ (second row) and $tW + E_T^{miss}$ (bottom row) signals in the 2HDM+a model.

From the mono-Higgs diagram in the top left one can see that it is possible for the pseudoscalar A to be produced on shell if the corresponding mass satisfies $M_A > m_a + m_h$. This allows for resonant mono-Higgs production, which will dominate over the diagram on the top right. Similar resonant enhancements can emerge from the mono-Z and $tW + E_T^{miss}$ channels as well, if the mass of the corresponding heavy Higgs, H or $H^\pm$, exceeds the mass of the mediator a and the visible particle Z or W. In the case of mono-Higgs the the resonant peak position of E_T^{miss} can be obtained from [26]:

$$E_T^{miss}{}_{max} = \frac{1}{2m_A}\sqrt{(M_A^2 - m_h^2 - m_a^2)^2 - 4m_h^2 m_a^2} \tag{3.22}$$

Using the benchmark parameters in 3.21 figure 3.8 shows expected distributions of E_T^{miss} for different masses of m_A and m_a. As expected from equation 3.22 the peak shifts to higher or lower E_T^{miss} values from variations of m_A or m_a respectively.

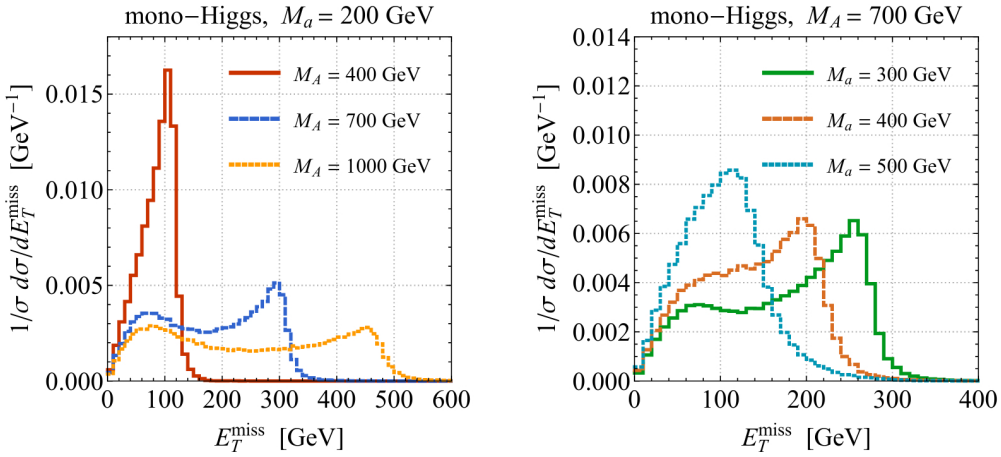


Figure 3.8: Normalized E_T^{miss} distributions for the mono-Higgs signal in the 2HDM+a model. The distribution on the left (right) varies m_A (m_a) while keeping m_a (m_A) fixed. The result shown corresponds to the benchmark parameters in 3.21 and is taken from [27].

For large mass splittings a tail in the lower spectrum of E_T^{miss} can also be observed. This comes from inferences between the two mono-Higgs diagrams in figure 3.7, where the tail events predominantly arises from the box diagram on right.

Scans of the 2HDM+a parameter space have been performed in both ATLAS and CMS for several channels. Exlucions in the $m_A - m_a$ plane from ATLAS searches in mono-Higgs, mono-Z and $h(inv)$ channels can be seen in figure 3.9. The colored regions correspond to points in the mass plane excluded by data at 95% Confidence Level (CLs). For the searches in the figure to be complimentary, all have employed the same benchmark parameters as in 3.21. The exclusion sensitivity comes predominately from the $h(b\bar{b}) + E_T^{miss}$ and $Z(\ell\ell) + E_T^{miss}$ channels, with a maximum reach of m_a up to 340 GeV and m_A up to 1.35 TeV. Areas within this region have also been excluded by $h(\gamma\gamma) + E_T^{miss}$ and $Z(q\bar{q}) + E_T^{miss}$ analyses. Constraints of very low m_a are set by the branching fraction of $h(inv)$, excluding masses $m_a < 100$ GeV for m_A between 200 GeV and 800 GeV.

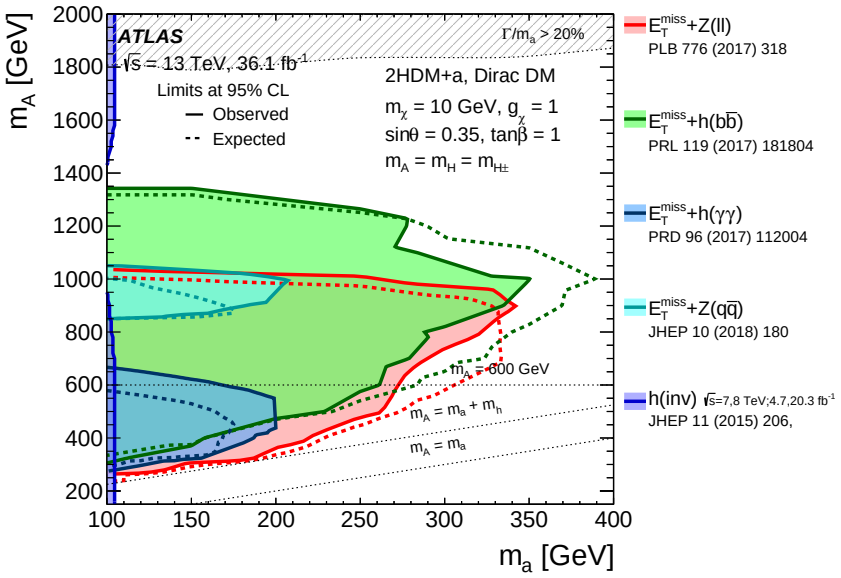


Figure 3.9: Regions in the $m_a - m_A$ plane excluded by data at 95% CL by mono-X analyses. The dashed grey regions at the top of the figure indicate the region where the width of any of the Higgs bosons exceeds 20% of its mass. Contributions from $b\bar{b}$ production is not included here as it is negligible in this region of phase space. This result is taken from [1].

For this analysis a m_a, m_A parameter scan has been performed for $h(\tau\tau) + E_T^{miss}$, a channel that as of yet not have been explored by ATLAS. This analysis therefore aims to be complimentary to the aforementioned searches, and hence employ the same benchmark parameters as in 3.21. For a proper treatment of the signal a good understanding of the final state $h \to \tau\tau + E_T^{miss}$ is required, especially the identification and reconstruction of these objects in the ATLAS detector. This experimental setup will be the focus of the next chapter.

Part II

Experimental framework

Chapter 4

The Experiment

4.1 The Large Hadron Collider (LHC)

The Large Hadron Collider [28] [29] (LHC) is the largest and most advanced particle collider in the world. It was built and is operated by the European Organization for Nuclear Research (CERN) near Geneva, Switzerland. The LHC consists of a 27 Km ring of superconducting 7.7 T dipole magnets designed to collide protons[12] at center-of-mass energy of $\sqrt{s}$=14 TeV every 25 ns with an instantaneous luminosity $\mathcal{L} = 2 \times 10^{34} \text{cm}^{-2} s^{-1}$. The particle collisions occur at four different interaction points in the LHC ring where detectors are placed to record the collision data for physics analyses. The largest detectors, ATLAS [31] and CMS [32], are so-called multipurpose detectors designed to measure and search for a wide range of physical processes. The two other detectors, ALICE [33] and LHCb [34] are on the other hand more specialized, designed to investigate specific problems in physics today [13]. A schematic layout of the LHC and the CERN accelerator complex can be seen in figure 4.1.

The LHC began operations in 2010, before temporarily going offline in 2013 for a two year upgrade. This first period of data taking saw center-of-mass energies between 7 and 8 TeV with an instantaneous luminosity $\mathcal{L} = 0.8 \times 10^{34} \text{cm}^{-2} s^{-1}$, and is referred to as Run 1. Following the first LHC upgrade, Run 2 began in 2015 with almost double the center-of-mass energy, $\sqrt{s} = 13$ TeV, and instantaneous luminosity increased to the designed value of $\mathcal{L} = 2 \times 10^{34} \text{cm}^{-2} s^{-1}$. Run 2 ended in 2018, and the LHC is expected to be offline until 2022 [39] for upgrades. The following running period, Run 3, will further increase center-of-masss energy to the designed value of $\sqrt{s}$=14 TeV. In the long term plans are put forward to increase the luminosity to more than double of the designed value with High Luminosity LHC (HL-LHC) from 2027, and the operation will permanently shut down some time in 2040s.

The protons used for collisions at the LHC start their journey from a gas of hydrogen atoms which is injected into a metal cylinder known as a Duoplasmatron. Here, the hydrogen atoms are ionized and separated into its electron and proton constituents. The protons are fed into a linear particle accelerator,

[12] The LHC also have a dedicated heavy ion program, where collisions of heavy nuclei such as Pb and Xe are studied at various energies [30].

[13] In addition to the mentioned four there are also three smaller detectors in the LHC ring; LHCf [35], MoEDAL[36] and Totem[37].

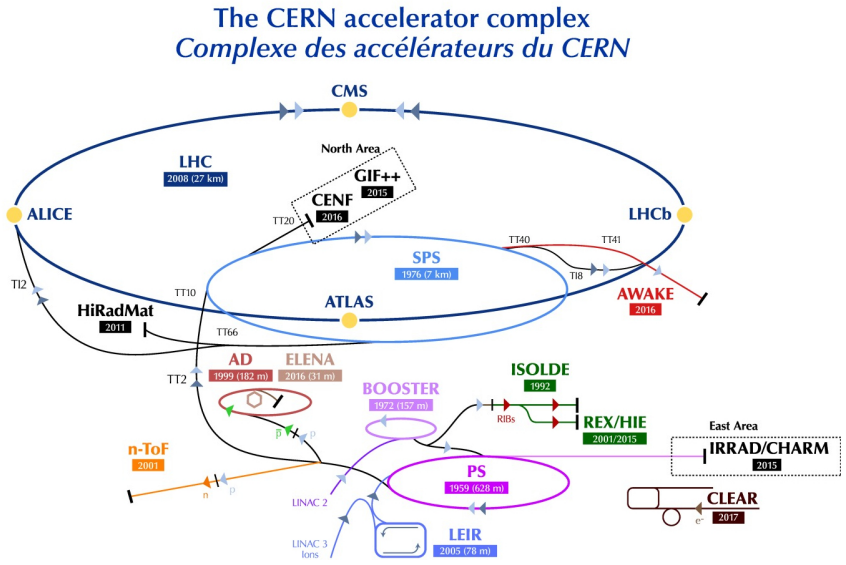


Figure 4.1: The CERN accelerator complex. The LHC (dark blue) is the last in a chain of several smaller smaller particle accelerators; the synchrotron Booster (bright purple), the PS (dark purple) and the SPS (bright blue). Highlighted with yellow dots are also the four main detector experiments at the LHC; ATLAS, ALICE, CMS and LHCb. Taken from [38].

LINAC 2, which accelerates them to 50 MeV using oscillating electric fields. The 50 MeV output beam from LINAC 2 then goes through a chain of synchrotron accelerators before ending up in the LHC. The first of these is the Proton Synchrotron Booster where the beam is accelerated to 1.4 GeV, followed by the Proton Synchrotron (PS) where it is accelerated to 25 GeV, then the Super Proton Synchrotron (SPS) where it is accelerated to 450 GeV, and finally the LHC accelerator where it is accelerated to the collision energy, which was 6.5 TeV during Run 2. This chain can be seen schematically in figure 4.1. To accelerate the particles the synchrontrons utilize radio frequency (RF) cavities, where standing electromagnetic waves are synchronized to the right frequency to boost the beam each time it passes by. To bend the beam in the LHC main ring over 1000 superconducting 7.7 T dipole magnets are installed, and nearly 400 superconducting quadrupole magnets are used to focus the beam into around 2500 bunches, 25 ns apart. Each bunch contains roughly 100 billion protons. The beam is split into two oppositely moving beams when entering the LHC to make collision possible. The two beams are then squeezed and focused by magnets at each collision point to ensure a high collision rate.

The collision rate is described by two quantities, namely the cross section σ and the instantaneous luminosity $\mathcal{L}$. The cross section is a measure related to the probability of a physical process taking place, and the instantaneous luminosity is defined as the the number of events detected (N) in a certain time (t) per unit cross section (σ). The collision intensity can hence be written out as:

$$\frac{dN}{dt} = \mathcal{L}\sigma \tag{4.1}$$

This expression holds independent on whether σ is the total collision cross section or the cross section for some specific process. In the case of a colliding beam experiment such as the LHC the instantaneous luminosity is given as follows:

$$\mathcal{L} = \frac{n_b}{2\pi} \frac{f_{rev} N_1 N_2}{\sigma_x \sigma_y} \tag{4.2}$$

where there are n_b bunches circulating in the collider with N_1 and N_2 particles per bunch for the two opposite directions. f_{rev} is the circulation frequency, and $\sigma_{x,y}$ is the bunch widths transverse to the velocity in the collision region (Typically assumed to be Gaussian). This off course only a simplification, and several technical factors such as beam crossing angles and beam offsets needs to be considered in real life. The amount of data collected over a certain time period can be found by integrating the instantaneous luminosity:

$$\int_0^T \mathcal{L} dt \tag{4.3}$$

With the amount of data collected by ATLAS and CMS this is usually given in inverse femtobarns, fb^{-1}. By using equation 4.1, it is possible to calculate the number of events of a certain process given the integrated luminosity if the cross section is known.

4.2 The ATLAS Detector

The ATLAS (A Toroidal LHC ApparatuS) detector is largest of the four main LHC experiments, and is the apparatus on which this thesis is based on. It is shaped like cylinder around the LHC beam axis and is roughly 44 meters long with a diameter of 25 meters. The weight of the machine is approximately 7000 tons. The scale can be seen in figure 4.2 where four tiny humans also are shown for comparison. The detector is a complicated instrument with several layers designed to take advantage of the high energies available at the LHC and observe physical phenomena occurring in this energy regime. One such observation is probably one of the greatest triumphs in modern physics, namely the discovery of the Higgs boson in 2012 [8]. When protons from the LHC beams collide in the center of the detector, this leads to the formation of many new particles that fly out in all directions. ATLAS is comprised of several sub-detectors, each specializing in measuring certain kinematic properties of these particles. By combining the measurements of the sub-systems it is possible to reconstruct the physics that occurred in the detector. An overview of this setup will be covered in the following.

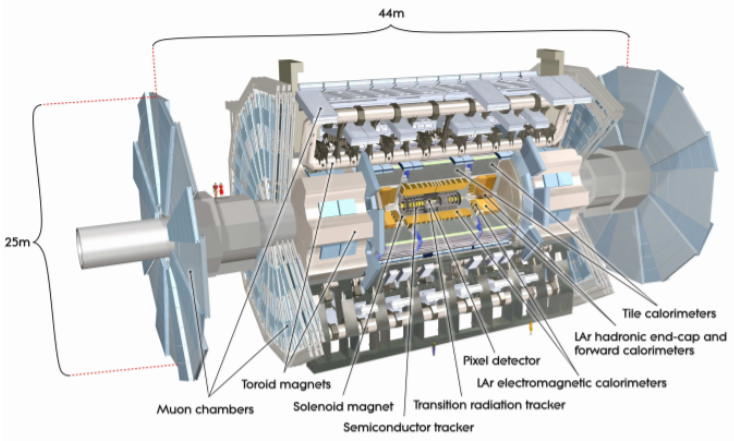


Figure 4.2: An illustration showing the dimensions and some subsystems of the ATLAS detector. Taken from [31].

4.2.1 ATLAS coordinates

The coordinate system in ATLAS is defined such that the nominal interaction point (IP) in the center of the detector is at the origin, and the beam direction is parallel to the z-axis. An illustration is shown in figure 4.3. The $x-y$ plane is transverse to the beam direction, where the positive x-direction points inward to the center of the LHC ring, and positive y upwards. The coordinate system can also be expressed in polar coordinates, where the azimuthal angle ϕ goes around the beam axis, and the polar angle θ is the angle from the IP to the $x-y$ plane.

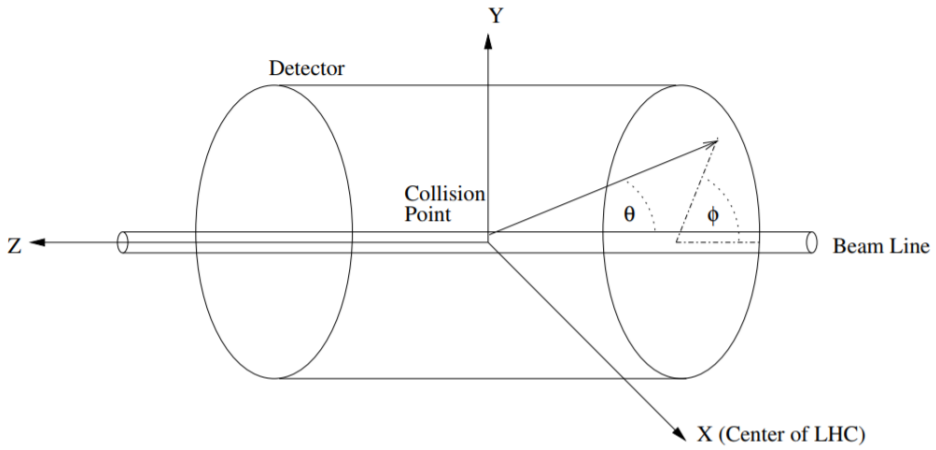


Figure 4.3: Illustration of the ATLAS coordinate system. The point labeled "collision point" in the figure corresponds to the IP. Taken from [40].

Instead of θ however, polar angle measurements are more often made using pseudo-rapidity, which is defined as:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{4.4}$$

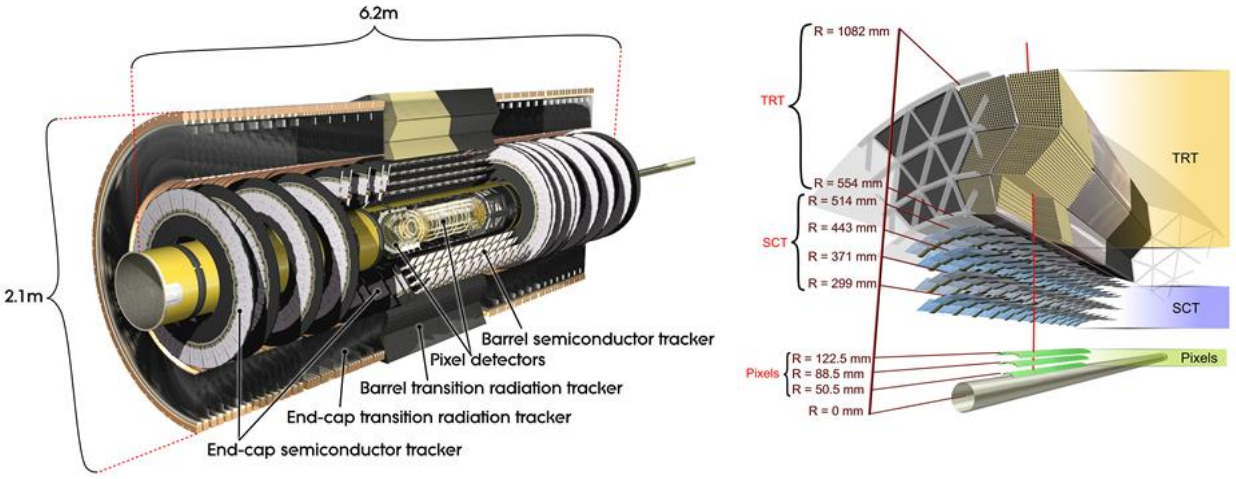
where $\eta = 0$ along the y-axis and increasing in absolute value closer to the beam line. Changes in η is a Lorentz invariant quantity for massless particles. As heavy particles decay quickly into lighter ones, most of the particles traversing the sub-detectors in ATLAS are light leptons, photons and hadrons, which often can be treated as massless compared to their energies. Hence, pseudo-rapidity is approximately Lorentz-invariant for all particles detected in ATLAS. The angular distance ΔR is defined in the $\phi-\eta$ space to be be:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{4.5}$$

From these coordinates it is possible to construct kinematic variables from the measurements of the physics objects in ATLAS. The transverse momentum p_T is an extensively used variable due to momentum conservation in the transverse plane, i.e $\sum_i \boldsymbol{p}_T = 0$ both before and after collision. However, some particles such as neutrinos and some hypothesised particles beyond the Standard Model will not be detected in ATLAS. The missing momentum contribution from such probes are added to the missing transverse momentum, $\boldsymbol{E}_T^{miss}$, which together with the total measured $\boldsymbol{p}_T$ should add up to zero. This is further explained in section 4.5.5.

4.2.2 Inner Detector

The ATLAS Inner Detector (ID) is the sub-detector system closest to the LHC beam-line and consists of three different sensor systems, namely the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). These three sensor systems are designed to to measure trajectories, momentum and charge of charged particles, while also minimizing particle energy loss. All sensors in the ID are immersed in a solenoid magnetic field parallel to the beam line operating at 2T. This magnetic field forces the trajectory of charged particles to curve, whose track is reconstructed in the ID through sophisticated tracking algorithms. The momentum of the particles is related to the amount of curvature of the associated track, and can hence be inferred and measured. The three subsystems in the ID are built on two different technologies, namely silicon sensors and straw drift tubes. When charged particle pass through the silicon sensors, they create electron-hole pairs which can be collected and recorded by applying an electric field. The straw drift tubes work in a similar way, where charged particles ionize the gas within the straws, liberating electrons. An applied electric field makes the electrons drift to a wire at the center of the straw, where they are recorded. The charge is recorded locally in the silicon sensors and straw tube wires, making it possible to identify the position of a traversing charged particle. Silicon sensors are used in both the Pixel detector and SCT, with silicon pixels for the former and silicon strips for the latter. Straw drift tubes are used in the TRT. Illustrations of the ID, and the three sub-systems can be seen in figure 4.4.



(a) (b)

Figure 4.4: An illustration of the ATLAS ID (a), and a cross sectional view of the barrel region(b), showing the Pixel Detector, SCT and TRT. Taken from [31].

The Pixel Detector

The Pixel detector [41] is the innermost part of the ID, and provides high precision position measurements close to the interaction point. It consists of silicon pixel sensors that are arranged in four layers around the beam-line and three discs in each end cap region. Initially, during Run 1, there were only three central layers, but during the long shutdown before Run 2 a new layer, namely the Insertable B-layer (IBL), was added to

the innermost region of the ID [42]. The Pixel detector covers the full range of ϕ and has a pseudo-rapidity coverage of $|\eta| < 2.5$. The three outer layers have a pixel size of $R\phi \times z = 50 \times 400 \mu\text{m}^2$ and an intrinsic spatial resolution of $14 \times 115 \mu\text{m}^2$. The IBL has an even finer granularity, with a pixel size of $R\phi \times z = 50 \times 250 \mu\text{m}^2$ and an intrinsic spatial resolution of $14 \times 72 \mu\text{m}^2$.

The Semi-Conductor Tracker (SCT)

The The Semi-Conductor Tracker [43] surrounds the Pixel detector. It is made out of silicon micro-strip sensors, arranged in four cylindrical layers in the barrel region, and nine discs in each end-cap region. The intrinsic spatial resolution of the micro-strips is 17 μm in the $R - \phi$ plane, and 580 μm in z-direction. It covers the full ϕ range, and has a pseudo-rapidity coverage of $|\eta| < 2.5$.

Transition Radiation Tracker (TRT)

In the outer part of the ID, enclosing the SCT we find the Transition Radiation Tracker [44], which is the largest sensor system in the ID. It consists of drift straw tubes filled with a xenon gas mixture, arranged in 70 barrel layers and 280 end-cap layers (140 on each end). The TRT provides position measurement with an intrinsic resolution of ~ 130 μm in both the $R - \phi$ plane and z-direction, and has pseudo-rapidity coverage of $|\eta| < 2.0$.

In addition to tracking, the TRT also provides particle discrimination through the detection of transition radiation. Charged particles emit transition radiation when passing through the polymer fibers between the straw tubes. The probability of emittance is proportional to the Lorentz factor $\gamma = \frac{E}{m}$. Therefore, more transition radiation photons will be emitted from electrons than from heavier charged pions at fixed momentum, and the number of photons detected provides some discrimination between the two.

4.2.3 Calorimeters

Moving radially outwards from the ID, one finds the calorimeters. These devices are designed to measure energies from both charged and neutral particles. There are two calorimeter systems in ATLAS; the Electromagnetic and Hadronic calorimeter. An illustration of the ATLAS calorimeter systems can be seen in figure 4.5. Both are sampling shower detectors, arranged in alternating layers of passive and active materials. The particles traversing ATLAS can interact with the passive material, either electromagnetically or hadronically, resulting in the production of a cascade of secondary particles. These cascades, more commonly referred to as particle showers, induce signals in the active material which is proportional to the energy deposited. Generally two types of showers occur, depending on the particle in question. Electrons and photons primarily interact through bremsstrahlung and pair production respectively in the passive layers. The sequential repetition of these two processes is called an electromagnetic shower. Hadrons on the other hand, typically loose energy through nuclear inelastic collisions with the passive material, causing what is called a hadronic shower.

By combining the shower induced signals in the active layers it is possible to infer the energy of most initial particles. The fractional energy resolution of the calorimeter systems takes the form:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{4.6}$$

Here, a is a stochastic term corresponding to fluctuations in the inducement of the signal, b is a noise term from the readout electronics, and c is a constant term corresponding to the non-uniformity of the detector response.

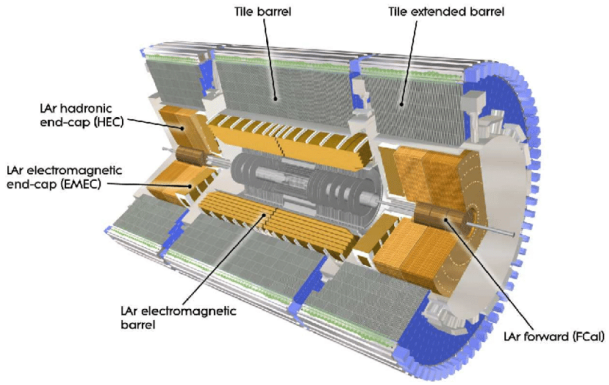


Figure 4.5: An illustration of the ATLAS calorimeters. Both the electromagnetic Liquid Argon calorimeter and the Tile Hadronic Calorimeter are shown. Taken from [45].

The Electromagnetic Calorimeter

The ATLAS LAr Electromagnetic Calorimeter[46] is mainly used for measuring the energy of electrons and photons. It is divided into a barrel part ($|\eta| < 1.475$) and two end-cap parts ($1.375 < |\eta| < 3.2$).

It is based on an accordion shape geometry, and consists of several layers of lead and stainless steal acting as the passive material. In between the layers there is liquid argon (LAr), cooled to -185° and acting as the active material. Immersed in the LAr is a copper grid which acts like an electrode to measure particles that pass through. Whenever an electron or photon pass through the passive material, it interacts and produces a shower of low energy electrons, positrons and photons. This shower enters the LAr chamber and ionizes the atoms, where free electrons will be attracted to the copper electrodes. By measuring the amount of charge deposited on the electrodes it is possible to infer the energy of the initial particle as it entered the EM calorimeter. The fractional energy resolution in the electromagnetic calorimeter after noise subtraction has been measured to be:

$$\frac{\sigma_E}{E} = \frac{10.1\%}{\sqrt{E}} \oplus 0.17\% \tag{4.7}$$

and

$$\frac{\sigma_E}{E} = \frac{12.1\%}{\sqrt{E}} \oplus 0.4\% \tag{4.8}$$

in the barrel [47] and end-cap region [48] respectively.

The Hadronic Calorimeter

The hadronic calorimeter [49] is located just outside the EM calorimeter. This device is constructed to measure the energy of hadronically interacting particles with a pseudorapidity range of $|\eta| < 4.9$. It is composed of three sub-systems; the Tile calorimeter which covers $|\eta|<1.7$, the LAr hadronic end-cap calorimeter which covers $1.5<|\eta|<3.2$ and the LAr forward calorimeter which covers $3.1 < |\eta| < 4.9$.

The Tile Calorimeter (TileCal) consists of a large array of interleaved steel and scintillating tiles acting as passive and active material respectively. When a high energy hadron passes through the steel it interacts with the atomic nuclei of the material, leading to the production of a hadronic shower. The charged particles in the shower radiate ultraviolet light when they enter the scintillating material, with the number of radiated photons being proportional to the deposited energy. The radiated UV light is subsequently converted visible light, whose measured intensity makes it possible to infer the energy of the incident particle.

On each of the barrel ends, the hadronic calorimeter is divided into two main components, the Hadronic End-cap Calorimeter (HEC) and the Forward Calorimeter (FCal). Both have a similar detector design to the LAr EM-calorimeter. The main difference is that the passive material in HEC is copper. The fractional energy resolution after noise subtraction has been measured to be:

$$\frac{\sigma_E}{E} = \frac{52.0\%}{\sqrt{E}} \oplus 3.0\% \tag{4.9}$$

,

$$\frac{\sigma_E}{E} = \frac{70.6\%}{\sqrt{E}} \oplus 5.8\% \tag{4.10}$$

and

$$\frac{\sigma_E}{E} = \frac{94.2\%}{\sqrt{E}} \oplus 7.5\% \tag{4.11}$$

for TileCal, HEC and FCal respectively [8].

4.2.4 The Muon Spectrometer

The ATLAS detector is designed to fully absorb most incident particles. However, there are cases where particles pass through the calorimeter systems without leaving much signal. Neutrinos, for example, do not interact at all with the ATLAS sub-detectors and are primarily recorded as E_T^{miss}. Muons, being minimum ionizing particles in the LHC energy regime, only deposit small energy fractions in the calorimeters. The Muon Spectrometer (MS) [50] is therefore placed in the outermost region of ATLAS to account for this. The MS is designed to capture signal from muons and provide precise measurements of their momenta within a range of $|\eta| < 2.7$. As can be seen in figure 4.6, it is composed of detector systems arranged in three layers in the barrel, at each end-cap and in the transition region in-between the end-caps and barrel respectively. The barrel region has a coverage of $|\eta| < 1.05$, and the end-cap regions cover $1< |\eta|<2.8$.

The spectrometer is immersed in a strong magnetic field provided by superconducting toroidal magnets with strength ranging from 0.15 to 2.5 T. Magnetic bending in the barrel region is preformed by a large barrel toroid that is strongest in the range $|\eta| < 1.4$. The end-cap regions have smaller toroid magnets with highest field strength in the range $1.6 < |\eta| < 2.7$. Magnetic bending in the transition region comes from a combination of the barrel end end-cap fields. As was the case in the ID, this allows for tracking

and high precision momentum measurements. These measurements are obtained from four different detector technologies.

The barrel region utilizes Resistive Plate Chambers (RPC) [51] and Muon Drift Tubes (MDT) [52] and the end-cap regions utilize MDTs, Thin Gap Chambers (TGC) [53] and Cathode Strip Chambers (CSC) [54]. Both the RPC and TGC provide fast responses to muon hits, making them ideal for triggering on muons. The MDT and CSC provide high spatial resolution measurements, and is used for precise reconstruction of muon tracks. The combined muon momentum resolution in the MS is p_T dependent, with $\sigma_{p_T}/p_T \sim 3\%$ at $p_T = 100$ GeV, and $\sigma_{p_T}/p_T \sim 10\%$ at $p_T = 1$ TeV.

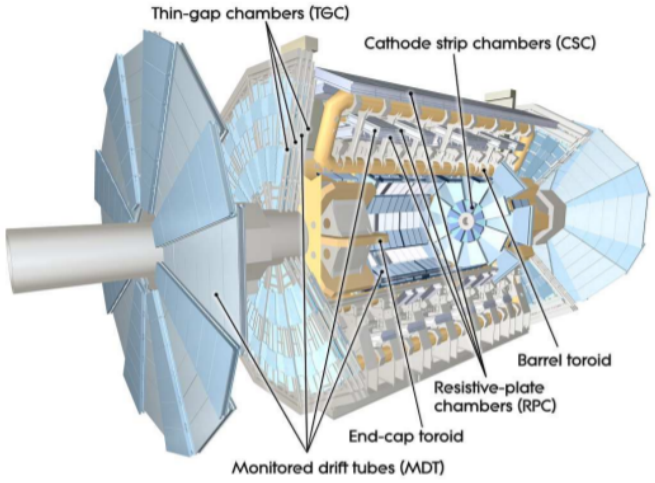


Figure 4.6: Illustration of the ATLAS Muon Spectrometer. Taken from ref [31].

4.3 Pile-up

When bunches of protons cross each other in the collision regions there are interactions between several protons. The most energetic interaction is usually the one of interest, and is referred to the hard-scatter event. The other interactions are also recorded and can contaminate the hard-scatter event, which consequently makes reconstruction a challenge. These events are collectively referred to as pile-up, and the number of pile-up events increases with increasing luminosity. During LHC Run 2, the total mean time averaged number of interactions per crossing $\langle\mu\rangle$ was around 30, as can be seen in figure 4.7.

There are however ways to deal with this. Pile-up contamination can be can be suppressed by using information from the high resolution tracking systems in the ID. If the resolution is sufficient to reconstruct individual interaction vertices, then the primary vertex which is associated with the hard scatter event can be identified. The event is then kept, and tracks that are not consistent with the primary vertex are vetoed. For example, the ATLAS experiment utilizes the Jet-Vertex-Tagger to identify and subsequently veto pile-up jets. This is a 2-D likelihood variable, that is estimated using the k-nearest neighbour algorithm [55]. The remaining pile-up effects that can not be accounted for are dealt with by an average pile-up energy correction.

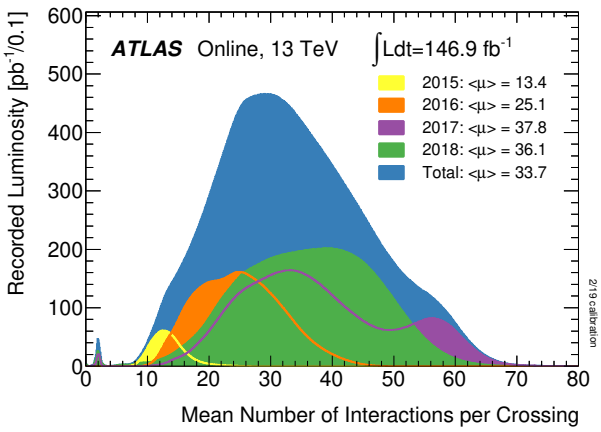


Figure 4.7: Mean number of interactions per crossing for the full Run 2 dataset [56].

4.4 The ATLAS trigger system

The event rate in ATLAS is roughly 40 MHz, with collisions occurring every 25 ns. As each full event corresponds to several MB of data, recording all the data from collisions would require a bandwidth of many TB per second. This is simply not feasible with current technology, and ATLAS therefore needs to prioritize what data should be recorded. Furthermore, processes hinting at new physics are rare with most of the events being Standard Model background processes. As such, a trigger system is employed to record a reduced data set with events deemed interesting, such as events with high E_T^{miss} or Higgs production.

The trigger system used for Run 2 [57] consists of a hardware-based first level trigger (L1), and a software based high level trigger (HLT). The L1 trigger uses custom electronics to quickly analyze events and determine Regions-of-Interest (RoIs), i.e interesting physics, by taking inputs from the coarse granularity calorimeter and MS information. If L1 decides that an event should be kept, it sends a Level 1 Acceptance signal (L1A) to the full ATLAS detector. This initiates a full read out of all sub-detector data in the RoIs determined by L1. The decision time for L1A is roughly 2.5 μs. The nominal rate of L1A is 100KHz, reducing the initial rate of 40MHz by a factor of 400. The readout data is then analyzed by the more sophisticated HLT. By using the full detector information in the RoIs, several selection algorithms determine if the event should be accepted or not. If accepted, the event is fully reconstructed and sent to long term storage. The HLT further reduces the rate from 100KHz to approximately 1kHz within a processing time of 200 ms. Hence, the full trigger system efficiently reduces the ATLAS event rate from 40 MHz to 1KHz, which is a reduction by a factor of 40 000.

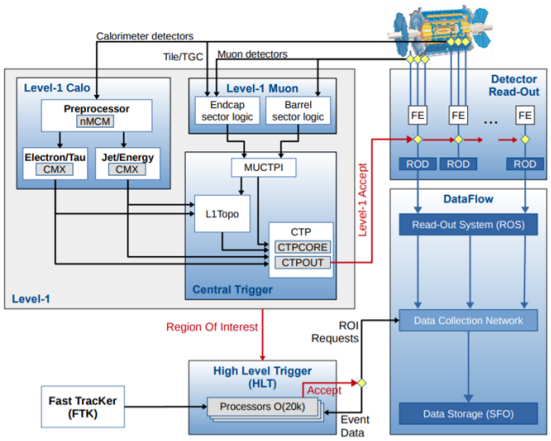


Figure 4.8: Schematic layout of the ATLAS trigger system used for Run 2, showing the flow of an event through the L1 and HLT triggers. Taken from [57].

4.5 Object reconstruction and identification

For the recorded signals from ATLAS to be of any use for physics analyses, it needs to be interpreted in terms of physical objects. The data format used for analysis contains physics objects like electrons, photons, E_T^{miss} and jets, built from the combined information of the sub-detectors. This phase of object building is called reconstruction. For some objects there is another additional phase, called identification, where a selection is applied to the object to decide whether the selected candidate can be identified as given particle. The relevant objects used in this analysis will briefly be discussed in the following.

4.5.1 Electrons

Electrons and photons are reconstructed from tracks in the ID and energy deposits in the EM calorimeter [58]. Electrons are required to have $p_T > 10$ GeV, $|\eta| < 2.47$ and pass the `LooseAndBLayerLLH` identification criteria described in [59].

4.5.2 Muons

The reconstruction of muons [60] use information from the MS, ID and occasionally also from the small energy deposits in the calorimeters if the MS is unavailable. Local track segments in the MS are combined to create a global MS track which is matched with the reconstructed track in the ID. Four muon reconstruction techniques are used to define four identification working points; `Loose`, `Medium`, `Tight` and `High-pT` muons. In this analysis Medium ID muons are used for vetoes in the signal regions, which is the standard muon working point in ATLAS. These signal muons are also required to have $p_T > 25$ GeV.

4.5.3 Jets

Quarks and gluons carry color charge and initiate hadronic showers when accelerated. The quarks and gluons within these showers combines into color neutral hadron states due to quark confinement, where again

these new hadrons can decay further. What one is left with is a cone-like shower of particles which originate from the initial quark or gluon, referred to as a jet. Jets are reconstructed in ATLAS by using information from the calorimeters, where the final state particles are clustered together according to a jet algorithm. Many such algorithms exist, but this analysis employ the anti-kt algorithm [61] with a distance parameter $R = 0.4$ in the $\eta - \phi$ plane. This parameter effectively sets the maximum size of the radius of the jet cone. The jets are also required to have $p_T > 20$ GeV and $|\eta| < 2.8$.

4.5.4 Tau leptons

The tau lepton is the heaviest known lepton, and the only lepton with sufficient mass to decay hadronically. As such, tau decays are grouped into two categories, namely hadronic and leptonic decays. The latter has an associated decay fraction of $\sim 35\%$ [3], but can mot be distinguished from prompt electrons and muons. Therefore, only hadronically decaying taus are considered for reconstruction. Furthermore, lepton number conservation requires the presence of a neutrino in the decay product of a hadronically decaying tau. These are elusive to the detector making an exact reconstruction impossible. Instead, only the visible part of the decay, $\tau_{\text{had-vis}}$, is used.

The reconstruction of the visible part of the tau leptons [62], employ the $\tau_{\text{had-vis}}$ reconstruction algorithm which is seeded from jet-objects that has been reconstructed with the anti-kt algorithm. These jets are required to have $p_T > 10$ GeV and $|\eta| < 2.5$, and are further studied to properly discriminate them from jets originating from incident quarks and gluons. A key characteristic used to discriminate tau decays is that they typically result in one or three charged hadrons as a consequence of charge conservation. This will correspondingly either leave 1 or 3 tracks in the ID, referred to as 1-prong and 3-prong τ's respectively. Only tau candidates associated with such tracks and with p_T >20 GeV are eligible for tau reconstruction. Additionally, candidates in the transition region, $1.37 < |\eta| < 1.52$ are rejected. Variables that exploit the tau decay features are fed into Boosted Decsion Trees (BDTs) and Artificial Neural Networks (ANNs), which define three working points for both 1-prong and 3-prong tau leptons, namely `Loose`, `Medium` and `Tight`. In this analysis the medium working point is used for signal taus, which corresponds to efficiencies of 75% and 60% for 1-prong and 3-prong taus respectively. In addition, the BDT loose working point is used for tau-electron discrimination.

4.5.5 Missing transverse energy

The collisions in ATLAS take place along the z-axis. Hence, energy-momentum conservation requires the transverse momenta to sum to zero after the collision. However, some particles like neutrinos and Dark Matter do not leave any signals and pass through ATLAS undetected. From the imbalance in transverse momentum it is possible to infer the existence of these particles. The vector representing this imbalance is referred to as missing transverse energy, $\mathbf{E}_T^{miss}$, with absolute value E_T^{miss}. It is calculated from the negative vector sum of transverse momenta of reconstructed electrons, muons, taus and jets [63]. In addition, the transverse momenta associated with unused tracks is also considered in the calculation. The former is referred to as the

hard term and the latter is referred to as the soft term. $\mathbf{E}_T^{miss}$ is given as the negative vector sum of the two:

$$\mathbf{E}_T^{miss} = -\Bigg(\underbrace{\sum \mathbf{p}_T^e + \sum \mathbf{p}_T^\mu + \sum \mathbf{p}_T^{\text{jets}} + \sum \mathbf{p}_T^\tau}_{\text{hard term}}\Bigg) - \underbrace{\sum \mathbf{p}_T^{\text{unused tracks}}}_{\text{soft term}} \tag{4.12}$$

4.5.6 Overlap removal

After the reconstruction of the above objects is complete, a procedure known as overlap removal is initiated. The purpose of this is to remove ambiguous objects that simultaneously have been reconstructed by different reconstruction algorithms. This could for example be objects that are both identified as a tau and an electron. Only one of the objects are kept to maintain a consistency between the kinematic variables. The procedure is done sequentially and is shown in table 4.1. Each step will be explained briefly in the following. If two electrons share the same track, the one with the lowest p_T is discarded. If a loose τ is within $\Delta R = 0.2$ of a remaining electron or muon, it is discarded. If calorimeter tagged muons share tracks with any electrons, they are discarded. If remaining electrons share tracks with non calorimeter tagged muons, they are discarded. If a jet is within $\Delta R = 0.2$ of any remaining electrons, it is discarded. If any remaining electrons is within $\Delta R = 0.4$ of remaining jets, they are discarded. If any of the remaining jets have less than 3 tracks and is within $\Delta R = 0.2$ of any remaining muons or the muon ID track is ghost associated [64], the corresponding jet is discarded. If any remaining muons are closer than $\Delta R = 0.4$ to any remaining jets, they are discarded. If any remaining jets is closer than $\Delta R = 0.2$ to any remaining VeryLoose τ leptons, the jet is discarded. The remaining objects after this procedure are the ones used in this analysis.

Reject	Against	Criteria
Electron	Electron	shared track, $p_{\mathrm{T},1} < p_{\mathrm{T},2}$
VeryLoose Tau	Electron	$\Delta R < 0.2$
VeryLoose Tau	Muon	$\Delta R < 0.2$
Muon	Electron	is Calo-Muon and shared ID track
Electron	Muon	shared ID track
Jet	Electron	$\Delta R < 0.2$
Electron	Jet	$\Delta R < 0.4$
Jet	Muon	`NumTrack` < 3 and (ghost-associated or $\Delta R < 0.2$)
Muon	Jet	$\Delta R < 0.4$
Jet	VeryLoose Tau	$\Delta R < 0.2$

Table 4.1: The sequential rejection procedure used for overlap removal.

Chapter 5

Simulation and analysis tools

5.1 Event Generators

To be able to make an experimental statement about a theoretical model with respect to observed data, there needs to a way to link the two. This is achieved with Monte Carlo (MC) generators, which in essence are the software equivalents of particle colliders. These offer a highly sophisticated framework that in simple terms produce the particle physics events with the same probability as expected from a given theoretical model. Particle physics is probabilistic by nature, and as such, one can calculate the number of events expected from a theoretical model to in a certain region of phase space. This can then be be compared to corresponding ATLAS data, and hence makes it possible to test the validity of said model. Typically, MC samples are generated in several stages, where for each stage, a generator that specializes in experimental or theoretical aspects of the model is employed. The fist step corresponds to the collisions of partons[14] at the LHC. This is implemented with so-called parton distributions functions (PDFs), which are defined as the probability density of finding a parton i with momentum fraction x_j with respect the momentum of proton j at some momentum transfer Q^2. These are determined experimentally and make it possible to calculate the cross section for a model. In the case of a pp-collision with two interacting partons, the cross section for the process $pp \to X$ is given by:

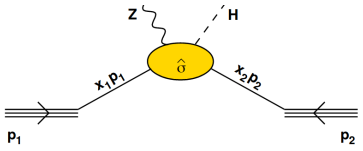


Figure 5.1: Schematic illustration of a hadron-hadron collision modelled using PDFs with $pp \to Z + h$ [65].

$$\sigma_{pp\to X} = \sum_{ij} \int dx_1 dx_2 f_1(x_1, Q^2) f_2(x_2, Q^2) \hat{\sigma}_{ij}(x_1 p_1, x_2 p_2, Q^2) \tag{5.1}$$

Here, $f(x_k, Q^2)$ are the PDFs and p_k the momentum of the initial protons, $k = 1, 2$. $\hat{\sigma}_{ij}$ is the parton level cross section for the process $i + j \to X$. Schematically this can be seen in figure 5.1, where two hadrons p_1, p_2 collide, and initiate parton-parton interactions with momentum fraction $x_1 p_1$ and $x_2 p_2$ respectively. The cross section $\hat{\sigma}$ corresponds to the process $i + j \to Z + h$, and can in principle be calculated from the SM Lagrangians. However, in practice such cross sections are divergent for QCD processes, where large

[14] In the parton model each proton is regarded as a system of particles called partons, i.e quarks anti-quarks and gluons.

cascades of gluon showers are produced. This is handled in another stage of the MC generation by dedicated simulation software specialized for shower modelling. The interactions within the ATLAS detector also need to be accounted for. This is implemented in a further stage where the full detector response for each predicted event is simulated. As this is very computational expensive, it is also the most time consuming part of the MC generation. In this last step, the MC data format goes through a step called "digitization" where the output format is the same as for ATLAS data, and can hence be fed into the same reconstruction software. In this way reconstruction of data and MC is consistent, allowing direct comparison between the two. The data taking from 2015-2018 is simulated in the production campaigns MC16a, MC16d and MC16e, corresponding to data taking in 2015+2016, 2017 and 2018 respectively. Pile-up effects from each year of data taking is also considered in the associated MC campaign. An overview of the full ATLAS MC infrastructure can be found in [66], and the MC generators employed in this analysis will be listed in the following:

- **PYTHIA8** is a multi-purpose generator used for shower modelling in this analysis [67].
- **Powheg** provides a framework for accurately calculating next-to-leading order (NLO) QCD processes. It has been used in combination with the shower modelling from PYTHIA8 in this analysis [68].
- **MadGraph** also provides NLO QCD calculations, and have been used in combination with the shower modelling from PYTHIA8 in this analysis [69].
- **aMCAtNlo** uses the MadGraph platform, and has also been used in combination with the shower modelling from PYTHIA8 [70].
- **SHERPA** is a multi-purpose generator, used for both matrix element and shower modelling in this analysis [71].
- **GEANT4** is used for generating detector responses in the ATLAS detector [72]. It uses a complicated detector geometry, making especially simulations of the calorimeter response very computational expensive. It is employed for almost all background samples in this analysis.
- **ATLFASTII** is a faster alternative to GEANT4 where the calorimeter response is simulated with FastCaloSim [73]. It is employed for all signal points and a few background samples in this analysis.

5.2 The ATLAS Run 2 event data model

The raw data output of ATLAS and MC generation data goes through three stages before being stored in the proper file format used for analysis [74]. In each stage, uninteresting events not matching certain selection criteria are rejected. The three file formats are given sequentially in the following:

- **Analysis Object Data**(**AOD**) is produced directly from the reconstructed objects and contain information about corresponding physics objects like tracks, tau leptons and jets. The contents of these files are referred to as xAOD.

- **Derived Analysis Object Data (DAOD):** AODs are too big to analyse directly. Therefore they are centrally reduced according to the need of specific physics groups. The new file format, referred to as DAODs, still contain xAOD objects but are much smaller.
- **ROOT NTuples** contain the information relevant for a specific analysis. These are stored in ROOT trees, and physics analyses are generally done with this file format using the ROOT software framework.

5.3 Software Tools

5.3.1 ROOT

ROOT is a object oriented framework for data processing, developed by Rene Brun and Fons Rademarker at CERN in 1995 [75]. The Framework is based on the C++ programming language and provides statistical methods, visualization and storage libraries for data analysis of high-energy-physics (HEP) data at the LHC. Although originally designed for HEP, it also has a big user base in other scientific fields outside particle physics. ROOT aims to make programming with sophisticated analysis tools accessible to the average user, while also allowing for analysis-specific flexibility. The framework comes with a built-in C++ interpreter, CINT, which makes it possible to run C++ macros directly from the prompt without compiling. This is indeed very useful in situations where rapid development is more important than execution time. In addition, ROOT provides a variety of libraries and environments which allow the user to mainly focus on the physics, rather than spend too much time on the underlying programming structures.

5.3.2 TMVA: Toolkit for Multivariate Data Analysis

TMVA (Toolkit for Multivariate Analysis) [76] is a ROOT-integrated environment for processing, evaluation and application of multivariate classification and regression methods. As such, TMVA provides a set of multivariate algorithms in a common interface, making it possible to test different algorithms on the same dataset. The family of algorithms integrated in TMVA are those related to supervised learning. That is, algorithms that are trained on data sets where the output is known (usually provided by MC) to produce an inferred function which can be used to map new examples. This is done in three steps:

- Training: Find optimal classification parameters using data with known labels (Signal and background MC events).
- Testing: Evaluate performance of the classifier on new unseen data.
- Application: Apply the trained classifier to real data where labels are not known.

This analysis utilizes multivariate tools, or more specifically Boosted Decsion Trees that are implemented in TMVA to define two signal regions. This is the explained in section 8.2.

5.3.3 HistFitter

HistFitter[77] is a framework for statistical analysis used extensively by the ATLAS collaboration. It is written in python and C++, where the former is used for the configuration and the latter for computational intensive calculations. HistFitter use the software packages HistFactory[78] and RooStats[79] internally, which are based on ROOT and RooFit[79], to construct parametric models and perform statistical tests on data. The underlying statistical models will be discussed in section 6.1. In essence, HistFitter performs complete statistical analyses from single user-defined configuration files. This is done by combining tools from several sources in a common interface. The HistFitter framework has been used for the systematic evaluation in section 9.2.1 and the statistical tests performed in chapter 10.

Part III

Search for the 2HDM+a signal

Chapter 6

Analysis overview and Statistical models

This analysis consider a SM Higgs bosons decaying into two tau leptons produced back-to-back with E_T^{miss} in search for a DM signal. The signal model in question is the 2HDM+a model explained in section 3.6. The model has been recommended by the ATLAS/CMS DM Forum, which refers to it as "“the simplest gauge-invariant and renormalisable extension of the simplified pseudoscalar model” [27]. Feynman diagrams for two possible signal processes that generate a $h + E_T^{miss}$ final state through gluon-gluon fusion (ggF) are shown in figure 6.1. The search for this signal is done in the $m_a - m_A$ parameter space with the benchmark parameters from equation 3.21.

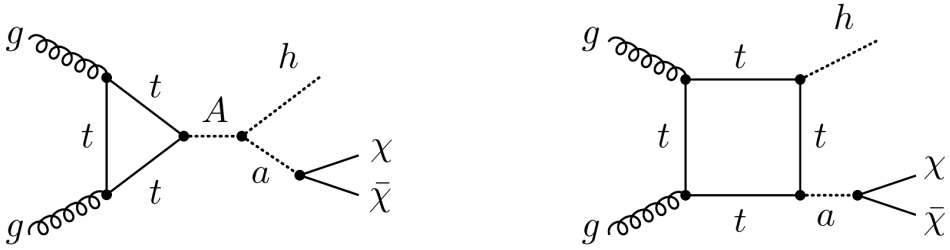


Figure 6.1: Feynman diagrams for ggF induced 2HDM+a signals that give rise to a $h + E_T^{miss}$ signature.

This thesis is supplementary to the ongoing ATLAS analysis in [2], where in addition to 2HDM+a, a search for mono-$h(\tau\tau)$ final states with the Z'2HDM model is performed. The author has contributed to parts of this analysis, some of which will be included in the following chapters. Included results not implemented by the author will be explicitly stated. At the time of writing the analysis is waiting for unblinding approval. As such, the results presented in this thesis will be based on MC samples.

The strategy of the analysis has been to optimize the signal selection to exclude as much of the parameter space as possible. To quantify this, and make useful predictions, statistical methods are used. Firstly, in chapter 8, selections in the phase space are implemented by optimizing cuts on kinematic variables with respect to the Asimov significance. This is done in two ways, namely one "cut and count" approach, and one multivariate approach. The background contributions in the signal regions are evaluated in chapter 9, where systematic uncertainties are included and constrained as nuisance parameters in likelihood fits. The results are discussed in chapter 10 where exclusion fits and hypothesis tests are done on Asimov simulated

data to obtain the expected exclusion contours for the 2HDM+a signal. All fits are done in the framework of HistFitter, whereas the Asimov significance estimation is implemented by the author in ROOT. The theoretical formulation of these statistical methods will be briefly discussed in the following. More information can be found in [77] and [80].

6.1 Statistical Procedure

When searching for new physical processes, it is assumed that the number of measured events follow a Poisson distribution with mean $s+b$, where s is the nominal number of signal events, and b is the number of background events. These measurements are binned and done in dedicated signal regions, designed to have high contributions from the 2HDM+a signal and low background contamination. Two hypotheses are then defined, the null hypothesis H_0 with $s = 0$ (background only) and some hypothesis H_1 with $s + b$ (signal and background). Both hypotheses are modelled by Poisson likelihood functions:

$$L(\mu, b) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s+b)} \tag{6.1}$$

with μ being the signal strength ($\mu = 0$ for H_0 and $\mu = 1$ for H_1). In a binned distribution the likelihood function would be the product of individual likelihood functions L_j for each bin j. In general there are also additional nuisance parameters $\boldsymbol{\theta}$ that needs to be accounted for. To test some hypothesised value of μ one considers the profile likelihood ratio:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} \tag{6.2}$$

Here, $L(\mu, \hat{\hat{\boldsymbol{\theta}}})$ in the numerator specifies the profile likelihood maximized for $\boldsymbol{\theta}$ under the assumption of a hypothesis with signal strength μ. That is, the value of $\boldsymbol{\theta}$ that maximizes L for a specified value of μ. The denominator gives the unconditional maximum likelihood value for L where $\hat{\boldsymbol{\theta}}$ and $\hat{\mu}$ are maximum likelihood estimates.

$\lambda(\mu)$ will always be between 0 and 1, with values closer to 1 indicating a better agreement between the hypothesised value of μ and the data. This can equivalently be written as the more convenient test statistic t_μ, which is just the negative logarithm of the likelihood ratio:

$$t_\mu = -2 \ln \lambda(\mu) \tag{6.3}$$

To quantify the compatibility of a signal model with data one introduces what is known as the p-value. This quantity gives the probability, under the assumption of a hypothesis H with some value for μ, of finding data of equal or greater incompatibility with the prediction of H. It is related to the test statistic by:

$$p_\mu = \int_{t_{\mu,obs}}^{\infty} f(t_\mu|\mu) dt_\mu \tag{6.4}$$

where $f(t_\mu|\mu)$ is the probability distribution function (PDF) of the test statistic given a fixed value μ. In particle physics it is more common to use the equivalent Gaussian significance when optimizing selections, which can be related to the p-value by:

$$Z = \Phi^{-1}(1 - p) \tag{6.5}$$

Here Φ^{-1} is the inverse of the cumulative Gaussian distribution. The HEP standard is that a significance of at least $Z = 5\sigma$ (p-value=$2.87 \cdot 10^{-7}$) is enough to constitute a discovery. In contrast the statistical sensitivity needed to exclude a signal hypothesis is set by a threshold value of $Z = 1.64\sigma$ (p-value=0.05). From Wilks' theorem [81], the discovery significance ($\mu = 0$) can be approximated to be:

$$Z_0 \approx \sqrt{t_0} = \sqrt{2\left(\ln \frac{n}{b} + b - n \right)} \tag{6.6}$$

The expected sensitivity of an experiment is characterized by the median significance which is estimated under the assumption of the Asimov dataset [80]. This assumption sets all quantities equal to their expected value, and the median Z_0 can be calculated accordingly. Assuming the $s + b$ hypothesis ($\mu = 1$), this gives:

$$\text{median } Z_0 = \sqrt{2(s+b)\ln(1+\frac{s}{b}) - s} \tag{6.7}$$

The above expression is a good approximation in many cases of s and b, and is referred to as the Asimov significance. The formula can further be modified to include the variance of of the background, σ_b^2, and following the derivation in [82], this modifies equation 6.7 to be:

$$Z_A = \left[2\left((s+b)\ln\left[\frac{(s+b)(b+\sigma_b^2)}{b^2 + (s+b)\sigma_b^2} \right] - \frac{b^2}{\sigma_b^2}\ln\left[1 + \frac{\sigma_b^2 s}{b(b+\sigma_b^2)} \right] \right) \right]^{1/2} \tag{6.8}$$

which reduces to $s/\sqrt{b+\sigma_b}$ in the large statistics limit. As the 2HDM+a signal points generally have small cross section, this approximation is not employed. Instead, the optimization is based on equation 6.8, where selections are imposed to define signal regions where Z_A is maximized. As these calculations are done in binned histograms, it is convenient to express s and b as functions integrated over each possible cut value x_{cut} in a binned the distribution $f(x)$ such that:

$$s = \int_{x_{cut}}^{\infty} f_s(x)dx$$
$$b = \int_{x_{cut}}^{\infty} f_b(x)dx$$

Following the optimization phase, fits to data are performed in the signal regions. Usually there are two ways of doing this depending on the goal of the analysis, namely discovery or exclusion fits. This analysis employ the latter, where the test statistic is redefined to be:

$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \leq \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{6.9}$$

Here, the value of q_μ is set to zero if $\hat{\mu} > \mu$. It is done this way as to not consider upwards fluctuations of data to represent incompatibility with the hypothesised signal strength μ. The corresponding p-value can be obtained by substituting q_μ for t_μ in equation 6.4.

When the expected number of signal events is much lower than background, the probability functions $f(q_\mu|\mu = 0)$ and $f(q_\mu|\mu = 1)$ will mostly overlap and the signal hypothesis might be rejected if a sizeable downward fluctuation in data is present. In the limit where the two distributions coincide the p-value for the signal tend to 0.05. However, an analysis should not exclude signal models has no sensitivity to. The CL_s method [83] is therefore used instead of the p-value in order to not exclude models where sensitivity is absent. This quantity is given as:

$$CL_s = \frac{p_\mu}{1 - p_0} \tag{6.10}$$

where p_μ and p_0 are the p-values corresponding to a model with signal strength μ and the background-only model respectively. If $CL_s < 0.05$, a specific signal hypothesis is excluded at 95% Confidence Level (CL).

An exclusion fit is run under the assumption of a specific signal hypothesis. If no excess from background is observed in the data, the fit tries to set exclusion limits on the model. If it is observed from the data that $CL_s < 0.05$, the signal model under study is excluded at 95% confidence level. The fit is model dependent, and can be done for a grid of signal points in order to obtain an exclusion contour.

The fits in this analysis are implemented in the HistFitter framework. As the data is blinded, the fits are performed by setting the "data" to the expected background MC as to get the expected exclusion contours for the $m_A - m_a$ grid. Systematic uncertainties are included in the fit as nuisance parameters which are constrained by a Gaussian function. Effectively, this changes equation 6.1 to:

$$L(\mu, \boldsymbol{\theta}) = L_0 \times C_{syst}(\boldsymbol{\theta}^0, \boldsymbol{\theta}) \tag{6.11}$$

where C_{syst} is the Gaussian function that depends on central values $\boldsymbol{\theta}^0$ around which the nuisance parameters $\boldsymbol{\theta}$ can be varied [77].

6.2 Important kinematic variables

Certain kinematic variables provide better separation between the 2HDM+a signal and SM background than others. Relevant variables that have been used in this analysis are listed here:

- E_T^{miss}: The missing transverse energy in an event, defined as the imbalance in the transverse momentum.
- N_{bjet} : Number of bjets in an event.
- $p_T(\tau_1)$: The transverse momentum of the leading tau lepton in an event .
- $p_T(\tau_2)$: The transverse momentum of the sub-leading tau lepton in an event.
- $p_T(\tau\tau)$: The vector sum of $p_T(\tau_1)$ and $p_T(\tau_2)$.
- $m_{inv}(\tau\tau)$: The invariant mass of the tau pair in an event, defined as:

$$m_{\tau\tau} = 2p_T(\tau_1)p_T(\tau_2)[\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)] \tag{6.12}$$

- $\Delta R(\tau\tau)$: The angular separation of the tau pair in an event, defined as:

$$\Delta R(\tau\tau) = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{6.13}$$

- $m_T(\tau_1, E_T^{miss})$: The transverse mass between the leading tau lepton and E_T^{miss} in an event, defined as:

$$m_T(\tau_1, E_T^{miss}) = \sqrt{2p_T(\tau_1)E_T^{miss}(1 - \cos\Delta\phi(p_T(\tau_1), p_T^{miss}))} \tag{6.14}$$

- m_T^{tot}: The total transverse mass in an event, defined as:

$$m_T^{tot} = \sqrt{(p_T(\tau_1) + p_T(\tau_2) + p_T^{miss})^2 - (p_x(\tau_1) + p_x(\tau_2) + p_x^{miss})^2 - (p_y(\tau_1) + p_y(\tau_2) + p_y^{miss})^2} \tag{6.15}$$

Chapter 7

Dataset and simulated samples

7.1 Dataset

The dataset used in this analysis was recorded with the ATLAS detector during run 2, corresponding to data taking in the years 2015-2018 with a center of mass energy $\sqrt{s} = 13$ TeV. The data is required to be part of the so-called Good Run List (GRL), including only events recorded when the full detector was operational. The following GRL is employed;

- `data15_13TeV.periodAllYear_DetStatus-v89-pro21-02_Unknown_PHYS_StandardGRL_All_Good_25ns.xml`
- `data16_13TeV.periodAllYear_DetStatus-v89-pro21-01_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml`
- `data17_13TeV.periodAllYear_DetStatus-v99-pro22-01_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml`
- `data18_13TeV.periodAllYear_DetStatus-v102-pro22-04_Unknown_PHYS_StandardGRL_All_Good_25ns_Triggerno17e33prim.xml`

The resulting data has a luminosity of $3.22\,\text{fb}^{-1}$, $32.99\,\text{fb}^{-1}$, $44.31\,\text{fb}^{-1}$ and $58.45\,\text{fb}^{-1}$ respectively. Figure 7.1 shows the cumulative luminosity versus time in ATLAS, where the blue area corresponds to the total certified good quality data. This amounts to a total luminosity of $138.97\,\text{fb}^{-1}$.

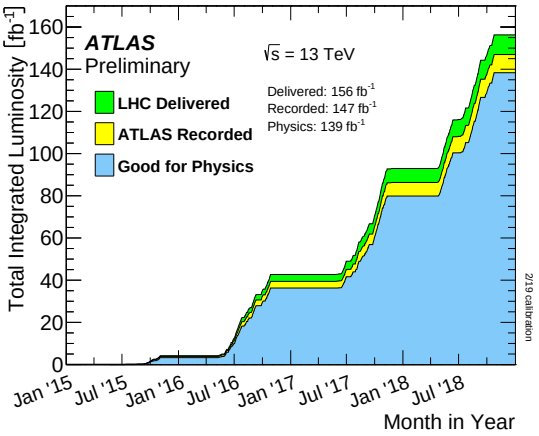


Figure 7.1: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at 13 TeV centre-of-mass energy in 2015-2018. Taken from [56].

7.2 MC simulations

7.2.1 Standard Model background simulations

A summary of the background samples used in the analysis can be found in table 7.1. Diboson and triboson, Z+jets and W+jets processes are simulated with the Sherpa event generator. All W+jets and diboson/triboson samples and most of the Z+jets samples employ the full detector simulation. Six Z+jets samples are simulated with the ATLFASTII detector simulation as they only cover a small fraction of the phase space, and correspond to an extension of the full simulation data set. Higgs boson production is simulated using the Powheg generator with Pythia8 for showering. All Higgs samples employ the full detector simulation and correspond to the processes; Wh, Zh, $ggZH$, ggH. For processes including top quarks, three different simulation scenarios are considered depending on the particular process. Processes including singletops, tW or top pair production are generated with Powheg. Events containing more than two top quarks, as well as Zt and $t\bar{t}WW$ processes are generated with MadGraph. Finally aMCAtNlo is used to generate ttW, ttZ and tWZ events. All top processes considered here use Pythia8 for showering and employ the full detector simulation.

Table 7.1: Summary of Monte Carlo samples used in the analysis.

Process	Generator	Shower Model	Detector sim
Diboson	Sherpa	Sherpa	Full
Z+jets	Sherpa	Sherpa	Full and ATLFASTII
W+jets	Sherpa	Sherpa	Full
Higgs	Powheg	Pythia8	Full
Singletops, tW, top pair production	Powheg	Pythia8	Full
Multitop, Zt, $t\bar{t}WW$	MadGraph	Pythia8	Full
ttW, ttZ, tWZ	aMCAtNlo	Pythia8	Full
2HDM+a	MadGraph	Pythia8	Full and ATLFASTII

7.2.2 2HDM+a signal simulations

The 2HDM+a signal samples used in this analysis are simulated with the MadGraph event generator and Pythia8 for showering. In total 33 signal points with the benchmark parameter values in 3.21 have been simulated in the $m_A - m_a$ grid, and is summarized in table 7.2. Quark ($b\bar{b}$) and gluon (ggF) initiated processes are simulated separately and combined according to their cross section. Technically the $b\bar{b}$ induced processes are simulated with $\tan\beta = 10$, and event weights are employed that reweights the samples to $\tan\beta = 1$. It should however be noted that contributions from $b\bar{b}$ are negligible for this particular value of $\tan\beta$ [27]. For ggF initiated processes the samples are simulated with $\tan\beta = 1$, and no reweighting. Points have also been simulated in the $m_A - \tan\beta$ grid, but scans over this parameter space will not be investigated in this analysis.

Table 7.2: Overview of the simulated points in the m_A-m_a grid of the 2HDM+a signal model. All points are simulated with the benchmark parameters in equation 3.21.

m_a [GeV] / m_A [GeV]	100	150	200	250	300	350	400	500
1200	x	x	-	-	-	-	-	-
1000	x	x	x	x	x	-	-	-
800	x	x	x	x	x	x	x	x
700	-	-	-	x	x	x	x	-
600	x	x	x	x	x	x	x	-
500	-	x	x	x	-	-	-	-
400	x	x	x	-	-	-	-	-
300	x	x	-	-	-	-	-	-

The normalized E_T^{miss} distributions for the simulations at truth level in figure 7.2 can bee compared to the theoretical prediction in figure 3.8. The difference in the shapes can be explained by the fact that E_T^{miss} is not perfectly reconstructed for the simulated signal samples. The reconstruction relies on $p_T(\tau_{\text{had-vis}})$ which only considers the visible part of the hadronically decaying tau leptons. The theoretical prediction is not constrained by this, and assumes perfect reconstruction of all physics objects. Still, the distributions have similar dependencies on the pseudoscalar masses.

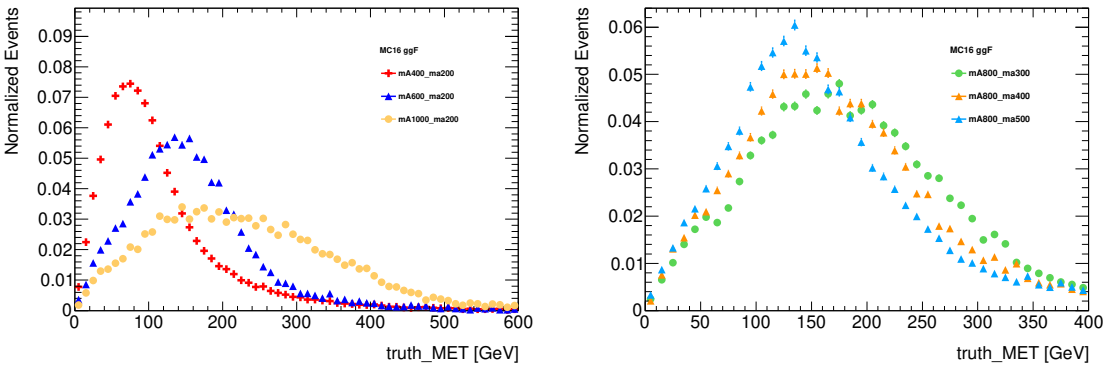


Figure 7.2: Normalized distributions of E_T^{miss} at truth level for ggF induced signal processes. The figure shows the effect on the E_T^{miss} peak from variations in m_A (left) and m_a (right).

All signal points employ the ATLFAST II detector simulation for an efficient use of computing power. To check that this does not affect the kinematic distributions the point $[m_A, m_a] = [700, 250]$ employ the full detector simulation in addition. A comparison between ATLFAST II and the full simulation for E_T^{miss} and $p_T(\tau_1)$ can be seen in figure 7.3. The upper panel shows the kinematic distributions for the two simulations, while the lower panel shows the ratio. The ratio is close to 1, indicating a good agreement. Hence no additional correction or uncertainty is needed within the present statistics. The only selection criteria applied is $N_\tau \geq 2$.

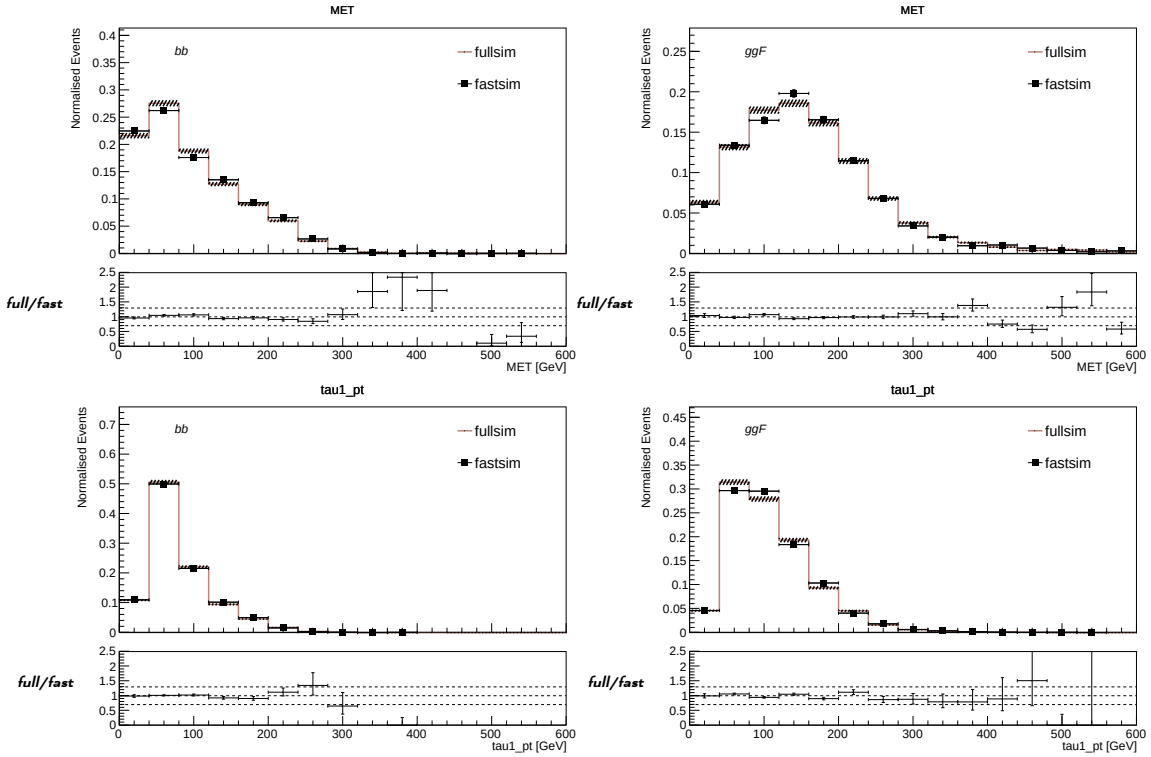


Figure 7.3: Comparison of the ATLFAST II and full detector simulation of the 2HDM+a signal. The upper plots show the E_T^{miss} distributions while the lower plots shows the leading τ p_T distributions. The red lines with shaded uncertainty show the full detector simulation and the black dots the ATLFAST II detector simulation.

7.3 Event Preselection

The preselection is implemented to ensure that the analyzed data only consists of properly reconstructed events. The following requirments are imposed:

- The event should be in a luminosity block that is contained in the Good Run Lists listed in section 7.1 to ensure that the full detector was ready to record data.
- The event should not contain any quality issue as flagged by the LAr, Tile, SCT and Core EventInfo flags.
- The event should have a primary vertex with at least two associated tracks.
- The event should not contain a badly-reconstructed muon, defined as a baseline muon found before overlap removal with large curvature error $\sigma(q/p)/|q/p| > 0.2$.
- The event should not contain any "loose bad jets" after overlap removal, where the bad jet definition is provided by the JetCleaningTool.

7.4 Triggers

The events in this analysis utilize the E_T^{miss}+ditau triggers listed in table 7.3. This trigger chain fires for events with $E_T^{miss} > 50$ GeV and transverse momentum of the two leading taus greater than a threshold

that depends on the running period. For the running period 2015-2017 this threshold is 35 GeV for $p_T(\tau_1)$ and 25 GeV for $p_T(\tau_2)$, while for the running period 2017-2018 the threshold of $p_T(\tau_1)$ is raised to to 60 GeV. The difference between `medium1_tracktwo_xe50` and `medium1_tracktwoEF_xe50` is an improved efficiency, while `mediumRNN_tracktwoMVA_xe50` is again an improved version of the latter which relies on Recurrent Neural Networks (RNNs) for tau lepton identification.

Running period	Trigger name
2015 – 2017	`HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_xe50`
parts of 2017	`HLT_tau60_medium1_tracktwo_tau25_medium1_tracktwo_xe50`
2018 period B-K	`HLT_tau60_medium1_tracktwoEF_tau25_medium1_tracktwoEF_xe50`
2018 from period K on	`HLT_tau60_medium1_tracktwoEF_tau25_medium1_tracktwoEF_xe50` OR
	`HLT_tau60_mediumRNN_tracktwoMVA_tau25_mediumRNN_tracktwoMVA_xe50`

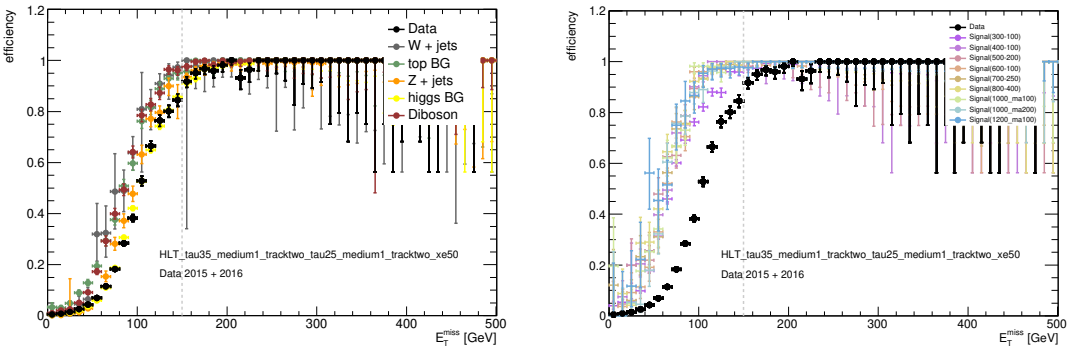
Table 7.3: Triggers used during the different running periods. The running periods B and K correspond data taking periods with different pile-up conditions, thus also different trigger conditions. More info can be found in [84].

The offline reconstructed tau objects are matched to the online objects where the trigger fired. Only events containing at least two tau objects that are matched to a positive trigger decision are considered valid. Furthermore, dedicated scale factors are employed to correct for the trigger efficiency of the τ part of the trigger decisions.

To avoid trigger turn-on effects additional cuts on the trigger variables are applied. These are referred to as "plateau cuts", and require at least two tau leptons with p_T 5 GeV above the trigger threshold, namely $p_T(\tau_1) > 40$ or 65 GeV and $p_T(\tau_2) > 30$ GeV. For E_T^{miss} the turn-on is expected to be signature dependent, and hence needs to be studied to obtain a selection with optimal efficiency. The trigger turn-on effects on E_T^{miss} for the backgrounds, data and some chosen signal points are shown in figure 7.4-7.6 for the three trigger scenarios respectively.

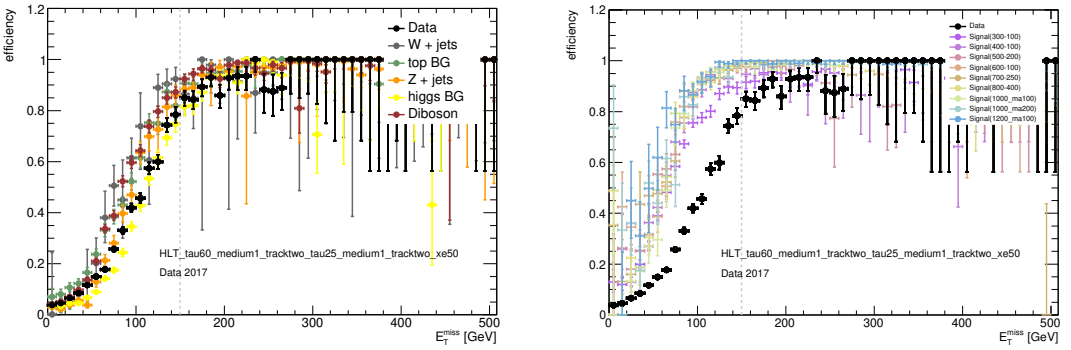
The efficiencies in the plots are calculated by dividing the number of events triggered by the ones quoted in table 7.3 with number of events triggered by corresponding $p_T(\tau_1)/p_T(\tau_2)$ triggers without any E_T^{miss} requirements. It can bee seen in the figures that the data has a somewhat slower turn-on than the backgrounds. This comes from substantial multijet contributions in the data. In addition, the different turn-on curves for signal and background come from the fact that E_T^{miss} is genuine for the signal processes, while for the background this mostly arises due to mismeasurements.

Overall, data and MC roughly agrees, with a full efficiency reached around $E_T^{miss} = 150$ GeV for most of the samples. Hence, an additional plateau requirement with $E_T^{miss} > 150$ GeV is included in the selection.



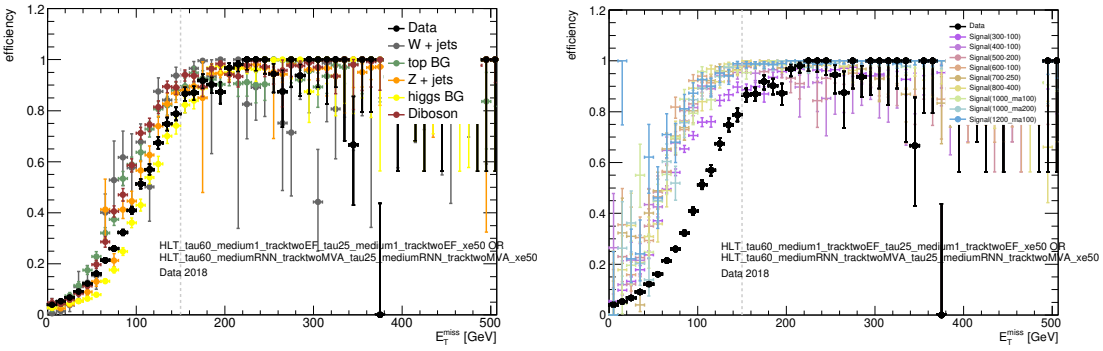
(a) Background (b) Signal

Figure 7.4: Trigger turn-on curves for background (left) and signal (right) with the trigger `HLT_tau35_medium1_tracktwo_tau25_medium1_tracktwo_xe50`.



(a) Background (b) Signal

Figure 7.5: Trigger turn-on curves for background (left) and signal (right) with the trigger `HLT_tau60_medium1_tracktwo_tau25_medium1_tracktwo_xe50`.



(a) Background (b) Signal

Figure 7.6: Trigger turn-on curves for background (left) and signal (right) with the combined triggers `HLT_tau60_medium1_tracktwoEF_tau25_medium1_tracktwoEF_xe50 OR` `HLT_tau60_mediumRNN_tracktwoMVA_tau25_mediumRNN_tracktwoMVA_xe50`, where either one can fire.

Chapter 8

Signal selection and optimization

Further selection criteria are applied in addition to the preselection and trigger requirements to efficiently select signal events while rejecting much of the backgrounds. Four signal regions are defined, two of which are cut-based and two selected from BDT scores. These regions are optimized with respect to the Asimov statistical significance estimation, which one aims to maximize.

8.1 Cut based Selection

Two signal regions are defined such that the selection is optimized with respect to the Asimov significance defined in equation 6.8. The selections are shown in table 8.1, and a background uncertainty of 15% is assumed at this stage. Both SRs follow the same cut flow with exception of the cut on $m_T(\tau_1, E_T^{miss})$. The kinematic distribution for this variable depends on the mass of the heavy pseudoscalar m_A. Hence two orthogonal cuts has been implemented to account for this, targeting signal points with $m_A < 600$ GeV and $m_A \geq 600$ GeV respectively. The efficiencies of the cuts with respect to the nominal prediction are shown in figure 8.1.

Generally a stable efficiency is observed for the targeted signal point. The last selection requirements on $m_T(\tau_1, E_T^{miss})$ is somewhat more drastic, but the corresponding background rejection is sufficiently high as to motivate the selections. The full cut flow will be explained sequentially in the following.

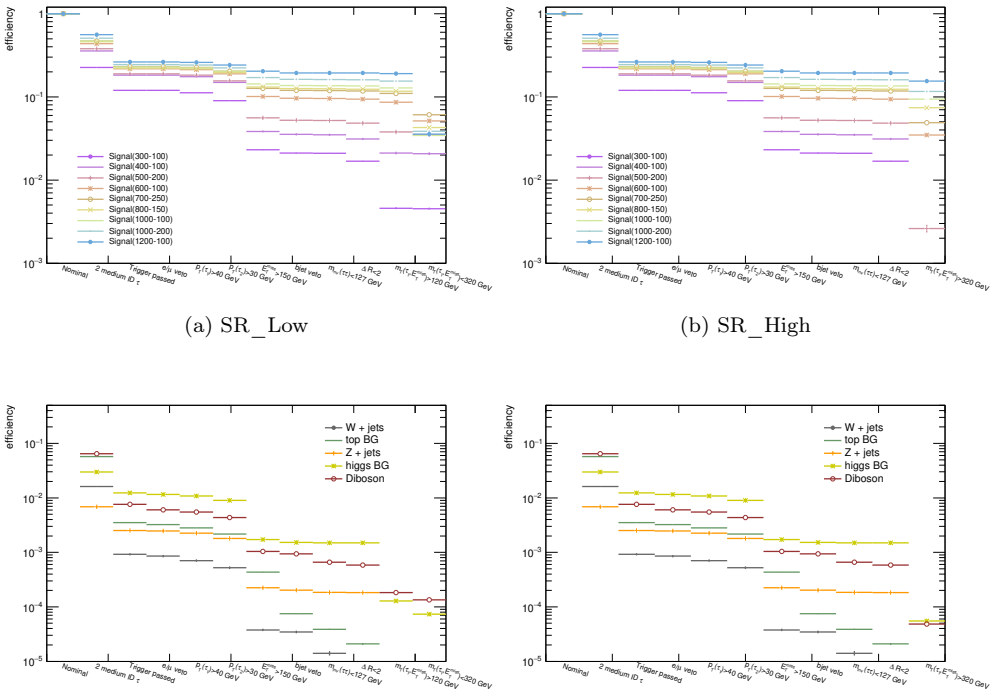
Events considered in this analysis must contain exactly two medium RNN ID tau leptons with opposite charge since they originate from the decay of the same SM Higgs boson. The utilized triggers require $E_T^{miss} >$ 150 GeV and p_T of the tau leptons to be 5 GeV above the trigger threshold for an event to be considered. The effect of these cuts is shown in figure 8.2 (a), (b) and (c).

The majority of top quark decays contain b-quarks, since the top decays via $t \to bW$. Therefore selecting events with a bjet veto provide an efficient background suppression of top processes. The effect of this selection in shown figure 8.2 (d).

The invariant mass of the tau lepton pair is expected to be distributed around the mass of the SM Higgs boson at 125 GeV. To allow for some fluctuations a selection on $m_{inv}(\tau\tau) < 127$ GeV has been implemented. This removes contributions from all background processes without having much effect on the signal efficiency as can be seen in figure 8.2 (e). The signal events contain tau leptons that originate from the same particle.

Table 8.1: Signal regions selection requirements.

Selection requirement		Reason for requirement
Pre-selection		Pre-selection
Passed trigger		Trigger requirement
$N_\tau = 2$, medium ID, opposite charge		Signal signature
$N_\mu = 0$, $N_e = 0$		Signal signature
$p_T(\tau_1) > 40$ GeV or 65 GeV (parts of 2017 and all of 2018 data due to trigger)		Trigger requirement
$p_T(\tau_2) > 30$ GeV		Trigger requirement
$E_T^{miss} > 150$ GeV		Trigger requirement
b-jet veto		Top background suppression
$m_{inv}(\tau\tau) < 127$ GeV		No signal expected outside this region
$\Delta R(\tau\tau) < 2$		General background suppression
SR_High	SR_Low	
$m_T(\tau_1, E_T^{miss}) > 320$ GeV	$m_T(\tau_1, E_T^{miss}) > 120$ GeV	Z+jets suppression
	$m_T(\tau_1, E_T^{miss})$<320	Orthogonality requirement



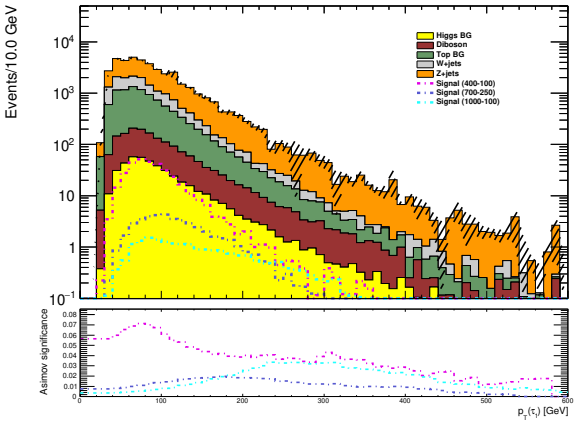
(a) SR_Low

(b) SR_High

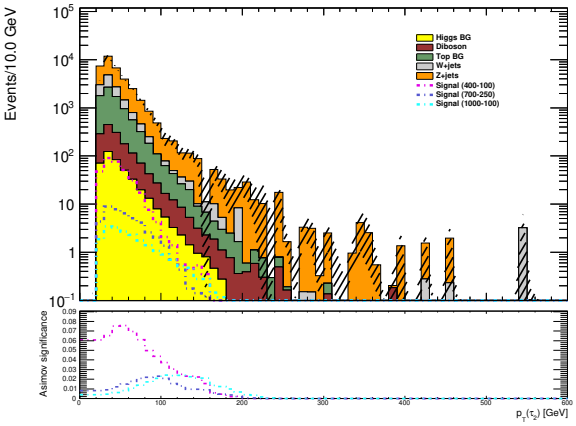
(c) SR_Low

(d) SR_High

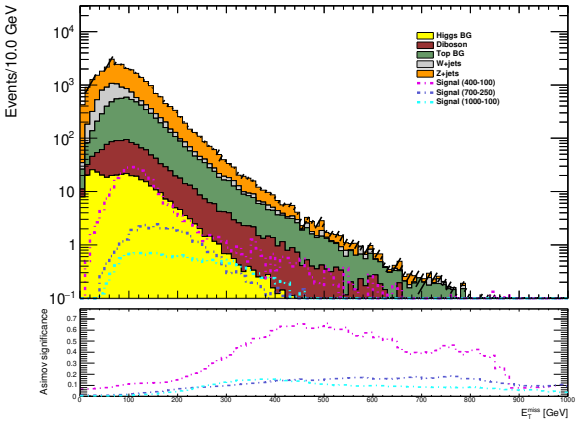
Figure 8.1: Efficiencies of the cuts compared to the nominal prediction. The upper row show the signal efficiencies and the lower row the background efficiencies.



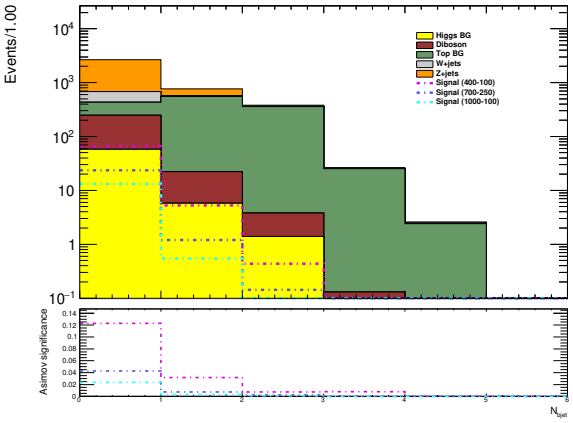
(a) $p_T(\tau_1)$



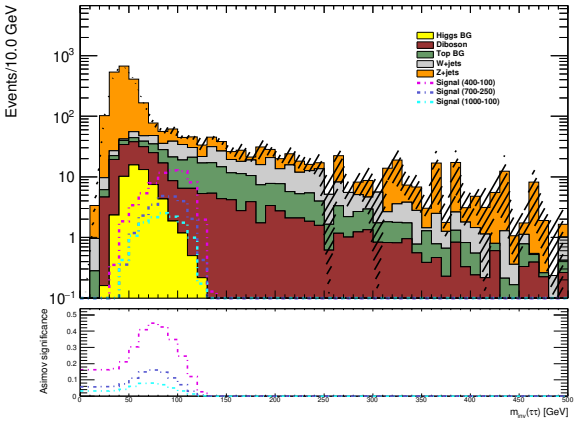
(b) $p_T(\tau_2)$



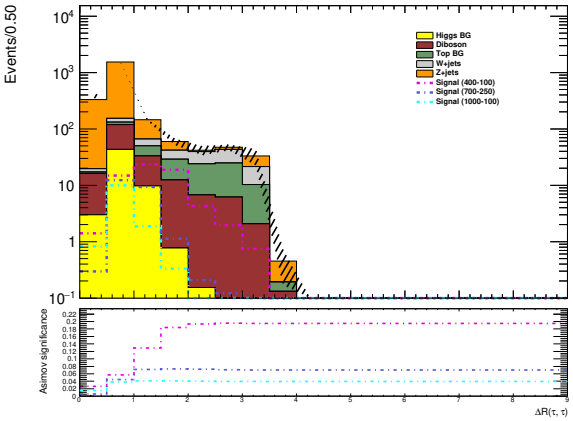
(c) E_T^{miss}



(d) N_{bjet}



(e) $\mathrm{m}_{inv}(\tau\tau)$

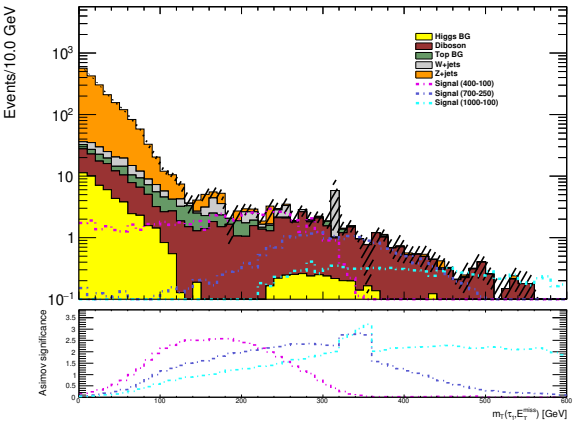


(f) $\Delta R(\tau\tau)$

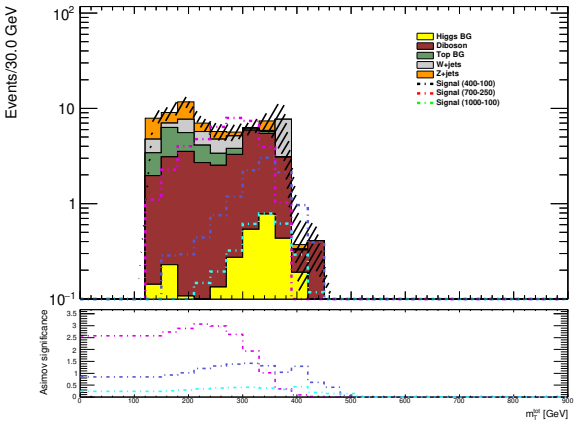
Figure 8.2: Kinematic distribution in consecutive stages of the selection shown in table 8.1. Three exemplary signal points; $[m_A, m_a]$ = [400, 100], [700, 250] and [1000, 100] GeV are also shown. The lower panel shows the Asimov significance defined in equation 6.8, where the signal and background yield is integrated from each bin to the last bin of the histogram (This is reversed for $\Delta R(\tau\tau)$).

They are therefore not expected to have a large angular separation, ΔR. Hence, only events with $\Delta R(\tau\tau) < 2$ are included. This suppresses diboson and top contribution, and to a lesser extent also Z+jets and W+jets. The effect of this selection can be seen in figure 8.2 (f).

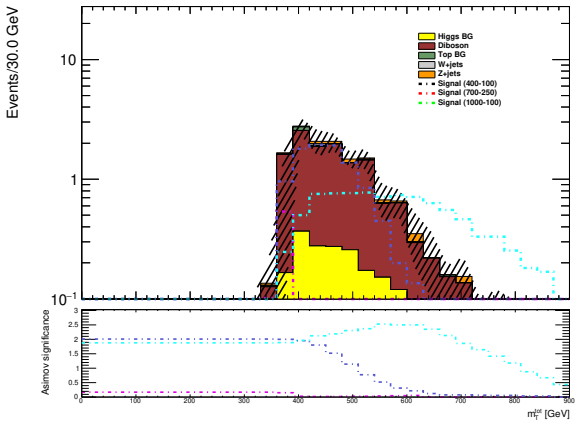
To suppress Z+jets background further selections are applied to $m_T(\tau_1, E_T^{miss})$. As seen in figure 8.3 (a) the kinematic distribution for the lower pseudoscalar mass point [400, 100] GeV peaks and falls off for smaller values of $m_T(\tau_1, E_T^{miss})$ than the two higher mass points. This is a behaviour that is common for most of the signal samples with $m_A < 600$ GeV. As this analysis aims to be sensitive to a wide range of points in the mass grid, two signal regions are defined; one optimized for points with $m_A < 600$ GeV (SR_Low) and one optimized for $m_A \geq 600$ GeV (SR_High). Both regions set a lower bound on $m_T(\tau_1, E_T^{miss})$ to suppress the Z+jets background, only including events with $m_T(\tau_1, E_T^{miss}) > 120$ GeV and $m_T(\tau_1, E_T^{miss}) > 320$ GeV for SR_Low and SR_High respectively. To make the regions orthogonal SR_Low sets an upper bound of $m_T(\tau_1, E_T^{miss}) < 320$ GeV.



(a) $m_T(\tau_1, E_T^{miss})$



(b) m_T^{tot} SR_Low

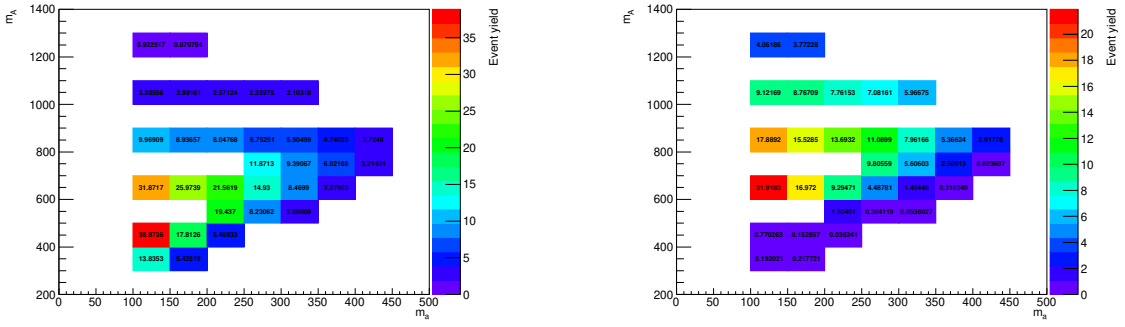


(c) m_T^{tot} SR_High

Figure 8.3: Kinematic distribution in consecutive stages of the selection shown in table 8.1. Three exemplary signal points; $[m_A, m_a]$ = [400, 100], [700, 250] and [1000, 100] GeV are also shown. The lower panel shows the Asimov significance defined in equation 6.8, where the signal and background yield is integrated from each bin to the last bin of the histograms.

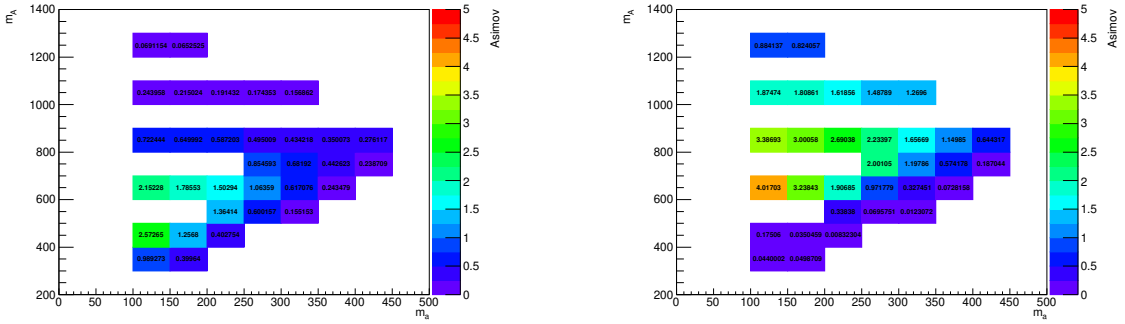
The m_T^{tot} distribution after the full selections for the two signal regions can be seen in figure 8.3 (b) and

(c). From the example points in the figures one can see that SR_Low is indeed optimized for the signal point with low m_A and SR_High for the two signal points with higher m_A. However, some samples, especially those with larger m_a or m_A in the mass grid suffer from low statistics after applying the signal region selections. This can be seen in figure 8.4.



(a) SR_Low (b) SR_High

Figure 8.4: Expected number of events for all simulated signal samples in the SR_Low and SR_High signal regions.



(a) SR_Low (b) SR_High

Figure 8.5: Expected sensitivity estimated with Asimov significance for all simulated signal samples in the SR_Low and SR_High signal regions.

For signal points with high m_a and low m_A, E_T^{miss} tends to peak at lower values as a result of the resonant Higgs production. As such, large portions of these events are removed by the plateau cut, $E_T^{miss} > 150$ GeV. Furthermore, the remaining events also tend to have a background like distributions, making background discrimination difficult. This last point is also the case for the samples with the highest m_A in mass grid, which tend to have flat distributions with no clear peaks. The expected sensitivity for all the samples in the two signal regions is calculated using the Asimov significance from equation 6.8 and is shown in figure 8.5. The background uncertainty in the calculation is assumed to be 15%. The sensitivity calculations here also include contributions from QCD processes, which were omitted in the optimization study. For both signal regions the sensitivity values fall off for increasing m_a due to lack of statistics. As such the two signal regions are mainly sensitive to the the points where $m_a < 300$ GeV, $m_A > 300$ GeV and $m_A < 1000$ GeV, and does not provide sufficient exclusion power to the other points in the mass grid. This last fact motivated

the use of multivariate analysis tools as an alternative to the cut based approach for defining signal regions. As multivariate tools employ more dimensions, using such methods may provide better separation between signal and background, and hence increase the sensitivity power of the analysis.

8.2 BDT based selection

To further increase the sensitivity of the analysis two additional signal regions have been defined by using Boosted Decision Trees (BDTs) with adaptive boosting. The BDTs are implemented in the framework of TMVA explained in section 5.3.2. The BDT with adaptive boosting was chosen because it preformed at least as good or better than other classifiers while still being relatively easy to optimize. In TMVA there are built in hyper-parameter optimization methods for BDTs which are integrated well in the configuration phase. Furthermore, BDTs are often referred to in the literature as the best "out of the box" classifiers as they in practice require little tuning to achieve reasonable results [76]. This is due to the simplicity of the method where each tree in the training step involves only a 1-D cut optimization. The trees are also insensitive to poorly discriminating variables, only splitting on the best discriminator at each step. In for example Artificial Neural Networks (ANNs) it is more difficult to deal with such variables. While ANNs in general preforms better in idealized situations, they require large sets of training data to obtain an optimized performance. BDTs, on the other hand, can preform well on even small training data sets. This last point is the case in this analysis where the amount of signal statistics generally is low.

8.2.1 Boosted Decision Trees

To understand how boosted decision trees work, one first have to be familiar with the concept of a regular decision tree. A schematic representation of a decision tree is shown in figure 8.6, which works as follows. First an event is passed to the root node where a cut on the variable xi is applied. The event is then passed into one of the two next nodes based on the value of xi. This process is repeated at each node, where the applied cuts either can be on a new variable or the same as before. In the end the event will reach the bottom of the tree, referred to as a leaf node, where either -1 (background) or 1 (signal) is outputted. In general, branches in a decision tree are constructed as to maximize information gain between two nodes. In the context of particle physics, this corresponds to applying the cuts at each node that maximizes the discrimination power between background and signal, and is referred to as separation gain. A popular choice to realise the separation gain, which also is applied in this analysis, is the Gini index defined as the following:

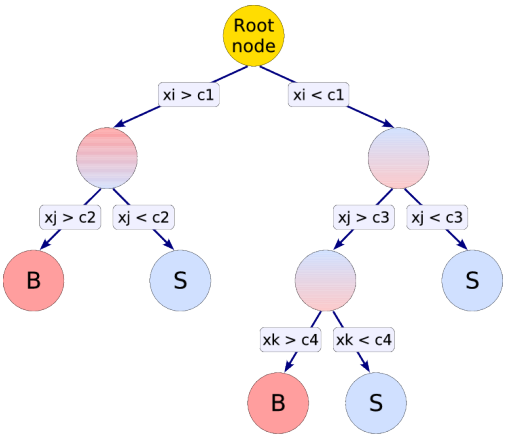


Figure 8.6: Schematic representation of a decision tree with three variables, taken from [76].

$$Gini(x) = \sum_{i,j\in[classes]}^{i\neq j} P_j(x)P_i(x) \tag{8.1}$$

where the sum is over any number of classes, and P_i corresponds to the probability that a random sample in distribution x is assigned to class i. In the case of two classes, signal and background, then $i = s$, $j = b$

and $P_s = 1 - P_b = s/(s+b)$. The gini index can be rewritten accordingly:

$$Gini(x) = 2P_s(x)(1 - P_s(x)) = \frac{2sb}{(s+b)^2} \tag{8.2}$$

This is at its maximum when $s = b$, corresponding to a Gini value of 1/2. The separation gain follows from this and is defined as:

$$SG(x, a) = Gini(x) - Gini(x|a) \tag{8.3}$$

Here, a is some specified selection on distribution x and $Gini(x|a)$ is the conditional Gini index for x given selection a. For all distributions x fed into the DT, one wants to find the selection that corresponds to the purest s or b, i.e maximizes the separation gain. In the case of this analysis x would correspond to the kinematic variables listed in section 6.2, and the procedure in 8.3 can be redone multiple times until a final tree is constructed.

DTs are really prone to overtraining as the number of nodes increases for each iteration. Furthermore, they are rather unstable and only small perturbations in the data may lead to widely different trees. Therefore a max depth of the tree is often predefined, usually with only a handful of branches. This reduces the classification power greatly, making them rather weak by them self with a correct classification rate of little over 50%. However, combining several of them in an ensemble, in a procedure known as boosting, has proven to be extremely successful[15].

The idea of boosting starts with first doing an iteration of the training samples where misclassified events are reweighted to be be more important in the next iteration. This procedure is repeated N times, and the end result of all classifiers are averaged. The final classifier is hence a linear combination of the base classifiers, and can be written out as follows:

$$Y = \sum_{m}^{N} \alpha_m b(\mathbf{x}; a_m) \tag{8.4}$$

Here, $b(\mathbf{x}, a_m)$ corresponds to the N base classifiers where a_m are the parameters used in the training of classifier m, $\mathbf{x}$ are the input variables and α_m is the weight of classifier m. There are two main techniques for boosting algorithms, namely Adaptive (AdaBoost) and Gradient Boost. Since this analysis implements the former it will be the focus here. AdaBoost sets the weight in equation 8.4 to be the following,

$$\alpha_m = \ln\left(\frac{1 - err_{1-m}}{err_{1-m}}\right) \tag{8.5}$$

where err_m is the misclassification rate in iteration m. All the event weights are renormalized accordingly such that the sum of weighted events stays constant. For each iteration the misclassified events by classifier b_m are reweighted by:

$$w = w_0 e^{\alpha_m} \tag{8.6}$$

Here, w_0 is the original weight. The output of each base classifier b_m is either 1 or -1 corresponding to signal or background respectively. The final classifier takes the average of the weighted outputs of the base classifiers and is given by:

[15] Other popular ensemble methods include bagging and random forest, where especially the latter is used widely outside the HEP community [85].

$$Y_{boost}(\mathbf{x}) = \frac{1}{N}\sum_{m}^{N} \alpha_m b_m(\mathbf{x}) \tag{8.7}$$

The output of the final classifier is thus between -1 and 1, where higher values are more signal like. It can be shown that the misclassification rate of Y_{boost} decreases rapidly with number of trees if the misclassification rate of the base learners are slightly better than random guessing [86]. In principle one can reach perfect classification by training on enough trees. However, this is not desirable as too many trees will lead to BDTs that are oversensitive to the training data. In other words, the BDTs "remember" a huge number of examples instead of extracting the relevant features. This can lead to a substantial increase in the misclassification rate when the BDTs are being evaluated on new data samples, and is referred to as overtraining. As such, one usually defines a separate testing dataset for evaluation and optimization such that the misclassification rate is minimal for both the test and train data. The effect of overtraining is visualized in figure 8.7 where the misclassification rate in the testing and training phase are shown for increasing number of trees. In general all classifiers will eventually overtrain if the model capacity is large enough to memorize the training data. However, adaboost have shown to be particular resistant to overtraining if the decision trees are small with a max depth of around 5 [87].

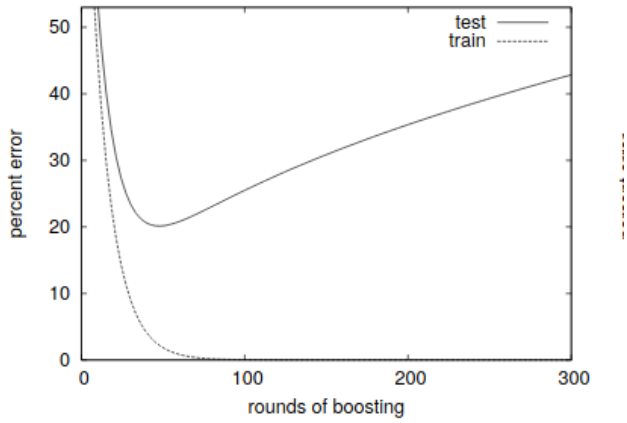
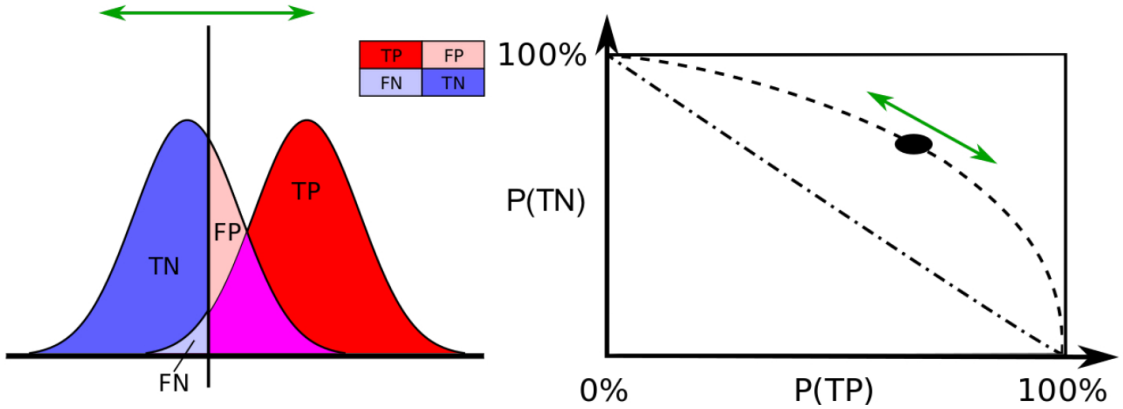


Figure 8.7: A dummy example of the misclassification rate for training and testing with increasing number of boosting rounds (number of trees). Overtraining is present in the region where the two diverge- Taken from [86].

Performance and ROC curves

In this analysis ROC (Receiver Operator Characteristic) curves are used as performance measure. These curves give the signal efficiencies vs background rejection for each value of the classifier response. Perfect separation between signal and background would give a rectangular ROC curve with unit integral.

Figure 8.8 illustrates how such curves work in general. The plot on the left hand side side show two dummy distributions, where the blue and red can be interpreted as background and signal respectively. A working point is defined in the middle of the two distributions (vertical black line), and the part of the distribution to the right of this line will be identified as true positive if it corresponds to the signal and false positive if it corresponds background. The opposite is the case to the left of the line, where signal and background will correspond to false and true negatives respectively. This working point can be



(a) (b)

Figure 8.8: Illustration of how the ROC curve (right) operates based on the two dummy distributions (left). Taken from [88].

mapped to the black do in the ROC curve on the right, where the x-axis represents the percentage of true positives and the y-axis the percentage of true negatives. Moving the working point to the right will move the point on the ROC curve to the left as this would decrease the percentage of true positives. Hence, it can be interpreted as the signal efficiency, while the percentage of false negatives can be interpreted as the background rejection(1−background efficiency). In context of this analysis the working point corresponds to a specific cut on the BDT response, and the ROC curve is obtained by plotting the signal efficiency (P(TP)) and background rejection (P(TN)) at every possible cut value.

8.2.2 Training and evaluation of the BDTs

Before the training step, certain selection criteria have been applied to the MC samples. These can be be found in table 8.2. In addition to the pre-selection, trigger and signal signature requirements, cuts on $m_{inv}(\tau\tau)$ and number of bjets have also been included. The low statistics mainly affect the signal samples, making the dataset heavily unbalanced with far more background events. As these two cuts mainly suppress background without having much effect on the signal samples, they serve as a prepossessing step for making the data more balanced without loosing too many signal events.

After the selection, the signal and background samples are split up into 50% training events and 50% testing events. For the signal samples this is done in two regions in the mass plane, for $m_A < 600$ GeV and $m_A \geq 600$ GeV respectively. Hence two BDTs are constructed and trained on the combined samples in the two mass regions separately. The amount of unweighted training and testing events for the two regions can be seen in table: To tune the BDT configuration parameters the built-in TMVA function "OptimizeConfigParameters()" has been used, with a scan over number of trees, learning rate and maximum three depth. No scans were preformed for the other parameters in the configuration, but the chosen parameter values were motivated by varying them and seeing which set of values gave the best performance. The best preforming parameter values are listed in table 8.4.

Two BDTs have been trained separately on the low and high m_A signal samples with the configuration parameters listed in table 8.4. The two BDTs will be referred to as BDT_low and BDT_high for low and

Table 8.2: Event selection before training

Selection requirement	Reason for requirement
Pre-selection	Pre-selection
Passed trigger	Trigger requirement
$N_\tau = 2$, medium ID, opposite charge	Signal signature
$N_\mu = 0$, $N_e = 0$	Signal signature
$p_T(\tau_1) > 40$ GeV or 65 GeV (parts of 2017 and all of 2018 data due to trigger)	Trigger requirements
$p_T(\tau_2) > 30$ GeV	Trigger requirements
$E_T^{miss} > 150$ GeV	Trigger requirements
b-jet veto	Top background suppression
$m_{inv}(\tau\tau) < 127$ GeV	No signal expected outside this region

Table 8.3: Number of unweighted events used for training and testing.

Region of mass plane	N_{Signal}	$N_{Background}$
$m_A < 600$ GeV	13 740 events	62 410 events
$m_A \geq 600$ GeV	56 409 events	62 410 events

Table 8.4: The optimized configuration parameters for the BDTs. All other parts of the configuration is set to default, more info about the configuration parameters can be found in the TMVA user guide [76].

Name	Description	Value for high (low) m_A samples	Scan
NTrees	Number of trees in Forest	1200 (700)	400-1600 for each 100
MaxDepth	Max depth of the decision tree allowed	3	3-6 for each 1
MinNodeSize	Percentage of training events required in a leaf node	5%	-
nCuts	Number of grid points in variable range used in finding optimal cut in node splitting	20	-
BoostType	Boosting type for the trees in the forest	"AdaBoost"	-
UsedBaggedBoost	Use only a random sub-sample of all events for growing trees in each boost iteration	"True"	-
AdaBoostBeta	Learning rate for AdaBoost algorithm	0.7	0.4-0.8 for each 0.1
BaggedSampleFraction	Relative size of bagged sample to original sample	0.5	-
NegWeightsTreatment	How to treat negative weights in training	"nonegweightsintraining"	-
SeparationType	Separation criterion for node splitting	"gini"	-

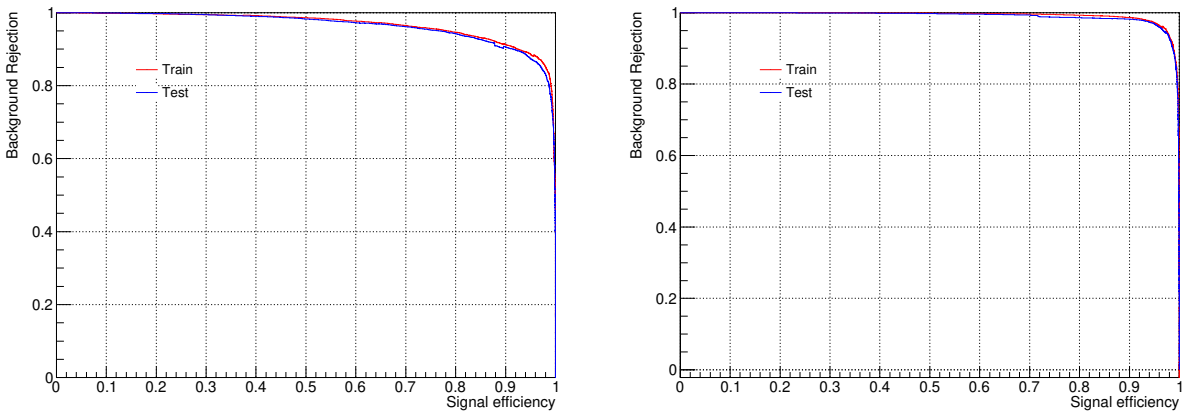
high m_A respectively in the following. In total 8 input variables have been used in the training. These are ranked by counting how often the variables are used to split tree nodes, and by weighting each split node by the squared separation gain and number of events in that node. The used variables and corresponding ranking in BDT_low and BDT_high are listed in table 8.5.

The performance of the BDTs are evaluated by inspecting the ROC curves from BDT responses on the test sets. To also check for overtraining these are plotted with the ROC curves from the BDT response on the training set and can be seen in figure 8.9. The corresponding ROC curve integrals (AUCs) [16] are shown in table 8.6.

[16] AUC stands for "Area Under the Curve"

Table 8.5: Ranking of variables used for training.

Ranking	Variables BDT_low	Variables BDT_high
1	$\Delta R(\tau\tau)$	$m_T(\tau_1, E_T^{miss})$
2	E_T^{miss}	m_T^{tot}
3	$m_T(\tau_1, E_T^{miss})$	$\Delta R(\tau\tau)$
4	m_T^{tot}	E_T^{miss}
5	$m_{inv}(\tau\tau)$	$p_T(\tau_2)$
6	$p_T(\tau_2)$	$m_{inv}(\tau\tau)$
7	$p_T(\tau_1)$	$p_T(\tau_1)$
8	$p_T(\tau\tau)$	$p_T(\tau\tau)$



(a) BDT_low

(b) BDT_high

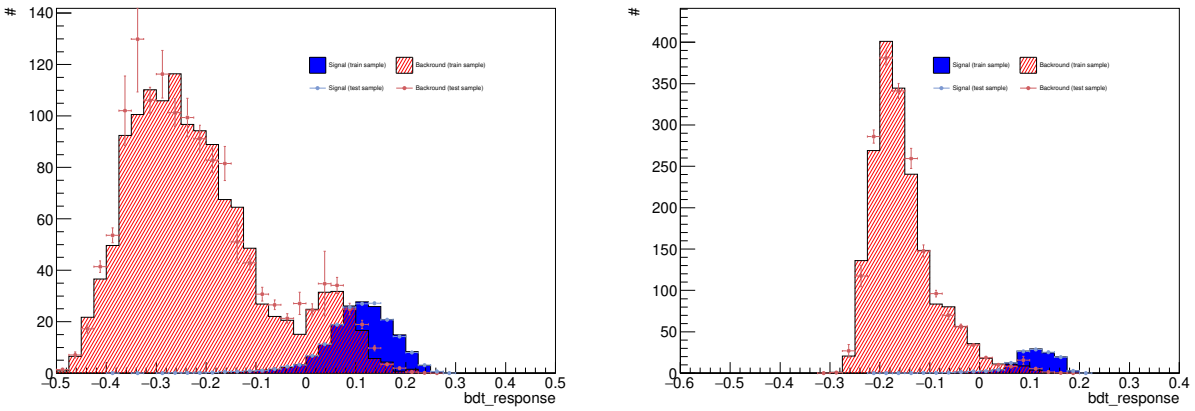
Figure 8.9: ROC curves for the BDT responses evaluated on train and test data

Table 8.6: The ROC curve integrals (AUCs) for the BDT responses evaluated on train and test data.

Classifier	AUC train	AUC test
BDT_low	0.968	0.965
BDT_high	0.992	0.990

An agreement is observed between the two and the plots indicate an almost perfect separation for the BDT_high response. It should however be noted that the BDTs are trained and evaluated on several combined signal samples and the actual separation power might vary for individual signal points. Furthermore, even if agreeing ROC curves is a good sign more checks needs to be done to ensure that no overtraining is present. The most straightforward way to verify this is by visually comparing the BDT response on the on the train and test samples as can be seen in figure 8.10. An agreement is observed between the test and training data for both BDTs. In addition TMVA provides a Kolmogorov-Smirnov (KS) test [89] in the evaluation phase which gives the probability that the train and test distributions are the same. Hypothetically a value

of 1 would be a perfect match between the two, but this kind of test is not always reliable when comparing binned histograms as it was originally designed for continuous distributions. A rule of thumb is therefore to accept probabilities larger than 0.1 and do further checks. As such three points in the BDT responses are chosen such that cuts at these points yields background efficiencies of 1%,10% and 30% respectively. The corresponding signal efficiencies are checked at each point and by convention overtraining is assumed to be negligible if deviations are less than 5% between training and testing. The KS probabilities and efficiencies for the two classifiers are listed in table 8.7. The KS test gives values above 0.1 for both BDTs, and the signal efficiencies in training and testing are within the ±5% interval. This combined with an observed agreement between the test and train samples in the ROC curves and response distributions indicate that no overtraining is present.



(a) BDT_low

(b) BDT_high

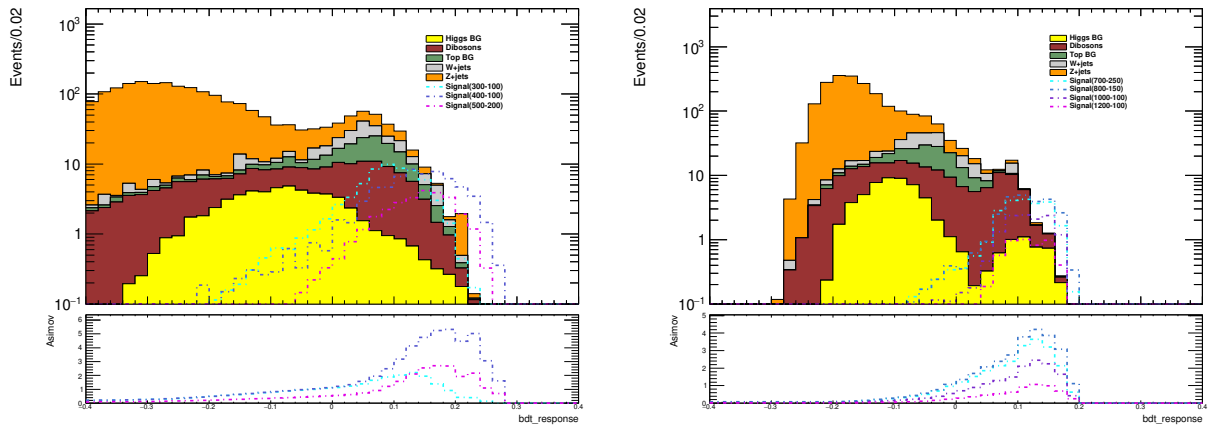
Figure 8.10: BDT response distributions for the train and test samples.

	BDT_low	**BDT_high**
Background efficiency	Signal efficiency in training (testing)	Signal efficiency in training (testing)
1%	39.6% (42.0%)	75.6% (79.9%)
10%	91.5% (91.6%)	98.0% (98.2%)
30%	99.5% (99.4%)	100.0% (100.0%)
	KS test signal (background)	KS test signal (background)
	0.869 (0.150)	0.884 (0.534)

Table 8.7: The upper part of the table shows three background efficiency points with corresponding signal efficiencies in training and testing. The lower part of the table shows the KS values for background and signal. The middle column are the values for BDT_low and the right column the values for BDT_high.

8.2.3 Selection

To optimize the selection on the BDT responses the Asimov significance estimator from equation 6.8 is used. The BDT response is evaluated on the full background dataset and a set of benchmark signal points are used to define a working point for the cut value. The full distributions for the BDTs can be seen in figure 8.11 where three and four signal points are shown for BDT_low and BDT_high respectively. This motivates the selection criteria that can be seen in table 8.8. The corresponding efficiencies, including the bjet veto and the $m_{inv}(\tau\tau)$ selection can be seen in figure 8.12. The selection is relatively stable for the targeted signal points, while simultaneously rejecting a significant amount of background. In principle one could cut even harder on the BDT_high response, but increasing the cut value further would drastically limit the statistics for both background and signal, and hence would not be optimal. With these cuts two additional signal regions are defined, namely SR_BDT_Low and SR_BDT_High .



(a) BDT_low

(b) BDT_high

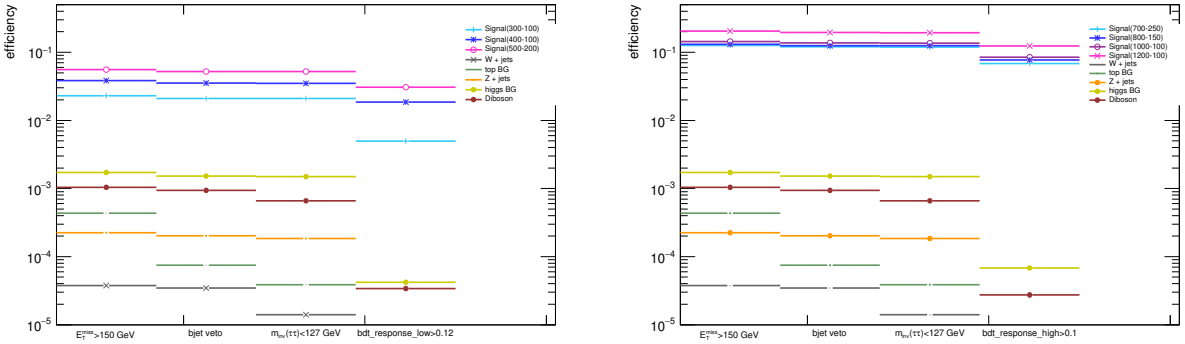
Figure 8.11: BDT response evaluated on the full background dataset, and three and four benchmark signal points for BDT_low and BDT_high respectively.

Classifier	Selection
BDT_low	response>0.12
BDT_high	repsonse > 0.10

Table 8.8: Selection criteria for the BDT responses.

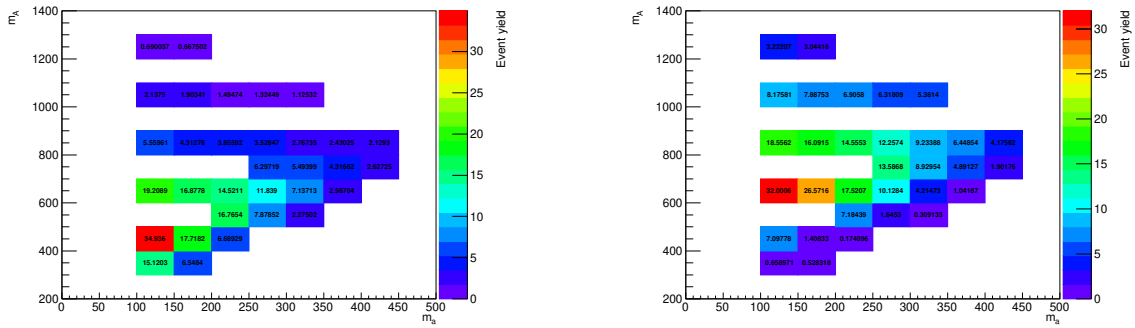
The expected sensitivity and event yields for the two signal regions can be seen figure 8.13 and 8.14 respectively. Comparing these results to the corresponding results for the cut based selection in figure 8.4 and 8.5 one can see that the exclusion power indeed has increased and is inclusive to more points in the mass grid . This is especially the case for low m_A in SR_BDT_Low and high m_a for SR_BDT_High . It should however be noted that SR_BDT_Low and SR_BDT_High are not orthogonal and hence can not be combined. Furthermore, the sensitivity in the highest mass region, i.e $m_A \geq 1000$ is still not optimal, with only the points $[m_A, m_a] = [1000, 100]$ and $[1000, 150]$ GeV having sufficient sensitivity to be excluded. This

is also the case for points with highest m_a along the m_A axis where no sufficient sensitivity is obtained. Still, the general increase in sensitivity makes the utilization of multivariate tools a success and a useful addition to the analysis.



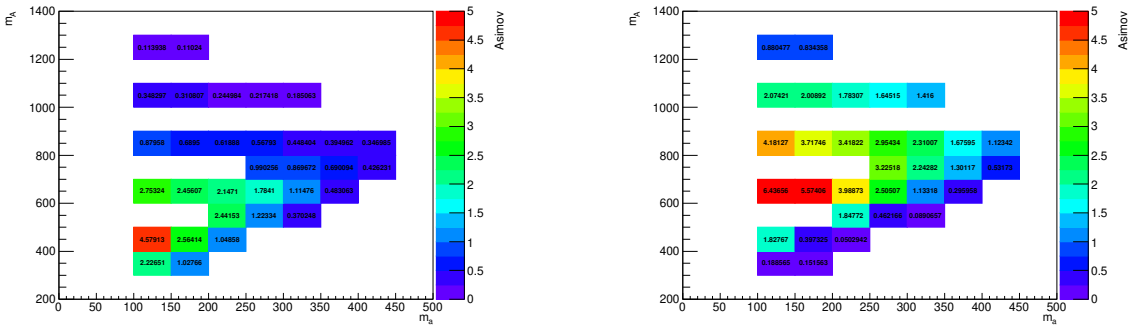
(a) SR_BDT_Low (b) SR_BDT_High

Figure 8.12: Efficiencies for the response selections, including bjet veto and $m_{\tau\tau} > 127$ GeV.



(a) SR_BDT_Low (b) SR_BDT_High

Figure 8.13: Expected number of events for all signal samples in the SR_BDT_Low and SR_BDT_High signal regions.



(a) SR_BDT_Low

(b) SR_BDT_High

Figure 8.14: Expected sensitivity estimated with Asimov significance for all simulated signal samples in the SR_BDT_Low and SR_BDT_High signal regions.

Chapter 9

Backgrounds and systematic uncertainties

9.1 Background and their estimation

When studying physical processes in multipurpose experiments such as ATLAS several background processes needs to be considered. Even if strict selections are implemented, the amount of data recorded is so immense that one can be statistically certain that there exists SM background with signal-like distributions. This can either occur because tails in these distributions may produce events in the considered phase space or because of mismeasurements in the detector where for example a jet can mistakenly be reconstructed as a τ. Furthermore, the variety of allowed kinematics of final state processes in the Standard Model makes it difficult for a signal process to be exclusively defined in any region of the phase space. As such, relevant background contributions entering the signal regions need to be considered. The relevant background processes considered in this analysis will be discussed in the following.

- Diboson production (WW,WZ and ZZ processes): These events contain genuine τ-leptons that originate from the boson decays. It is the dominant background source the signal regions with contributions between 25% and 60%. It is estimated using MC, but cross checks are done in three control regions to ensure the validity of the simulation. This is the topic of section 9.1.3.

- Multijet processes: These events may contain τ-leptons that are misidentified from jets as a result of the jet abundance in pp collisions. The background source is estimated using the data driven fake factor method in [2], and the result is extrapolated to the SRs in this analysis. This is explained in section 9.1.1. These events are mostly negligible in the SRs, and only enters SR_High with a contribution of about 15%.

- Higgs processes: These events mainly arise from Zh decays, where the Z decays to neutrinos and the Higgs to τ-leptons. It is estimated using MC, and contributes between 4% and 25% in the SRs.

- Top quark processes: These events are mostly suppressed by the bjet veto in the SRs since top quarks mainly decay via $t(Wb)$. However some contribution is present in SR_Low ($\sim$ 14%) and SR_BDT_High ($\sim$ 25%). The background is estimated using MC.

- Z+jets production: These events contain τ-leptons from $Z-$decays. No sizeable E_T^{miss} is expected for these processes, but statistical fluctuations or mismeasurements may pass the selection criteria. Z+jets background is mostly relevant in the low mass SRs, with a yield contribution of about 25% in both SR_Low and SR_BDT_Low .

- W+jets production: These events contain τ-leptons where one originate from the W and one from a misidentified jet. Typically the E_T^{miss} is bellow the selection criteria, but some events might pass due statistical fluctuations or mismeasurements. The background is estimated using MC, but the tau promotion method has been employed to increase the statistical precision. This is explained in section 9.1.2.

9.1.1 Multijet background estimation

The contribution from multijet processes in the signal regions are estimated through the fake factor method. The method is explained in the supporting note of the analysis, [2], and the result is evaluated in the signal regions in this thesis. As the author did not implement this first hand, only a short overview of the method and results will be described in the following. The fake factor is the ratio between number of events that contains τ-leptons that satisfy the signal ID requirement of medium ID and the number of events that satisfy the requirement of τ-leptons with loose but not medium ID. This fraction is calculated in a specific measurement region that does not overlap with the signal regions and then applied to the events after the signal region selections. This analysis require 2 medium ID τ-leptons, so either of these can fail or pass the criteria. Hence four regions are defined, which are shown in figure 9.1. These regions correspond to either two τ-leptons with very loose ID (lower left), one with very loose ID and one with medium ID (lower right), one with medium ID and one with very loose ID (upper left) or two with medium ID (upper right). The latter region is part of the signal region requirements, and the number of events in this region can be estimated from the remaining three regions with the fake factors FF and FF_{loose}. FF corresponds to the fraction of events containing two medium ID τ-leptons to events containing one medium ID τ-lepton and one loose but not medium ID τ-lepton. In figure 9.1 this is represented by the white arrows. FF_{loose} corresponds to the fraction of events containing one medium ID τ-lepton and one loose but not medium ID τ-lepton to the events containing two loose but not medium ID τ-leptons. This is represented by the gray arrows in figure 9.1.

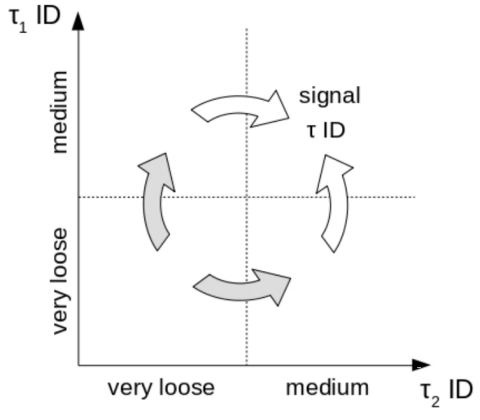


Figure 9.1: The four regions defined by the fake factor method.

Hence, the two fake factors take the following form:

$$\begin{aligned} FF &= \frac{\text{QCD events with 2 medium ID } \tau\text{-leptons}}{\text{QCD events with one medium ID and one loose but not medium ID } \tau\text{-lepton}} \\ FF_{loose} &= \frac{\text{QCD events with 1 medium ID and one loose but not medium ID } \tau\text{-lepton}}{\text{QCD events with two loose and not medium ID } \tau\text{-leptons}} \end{aligned} \tag{9.1}$$

The number of events in the signal regions can then be extrapolated from the event yields in the three other regions with the fake factors according to:

$$\begin{aligned} N_{QCD} =& N(ID, ID) = N(ID, \not{ID}) \cdot FF(\tau_2) + N(\not{ID}, ID) \cdot FF(\tau_1) \\ &- \frac{1}{2}(N(\not{ID}, ID) \cdot FF_{loose}(\tau_1) \cdot FF(\tau_2) + N(\not{ID}, \not{ID}) \cdot FF_{loose} \cdot FF(\tau_1)) \end{aligned} \tag{9.2}$$

where $N(ID, ID)$ corresponds to the number of QCD events satisfying the two medium ID τ-leptons criterion, $\not{ID}$ represents the loose but not medium ID configuration and τ_1,τ_2 corresponds to the leading and subleading τ-lepton respectively. Using the implementation from [2] the expected contribution from QCD events can be estimated in the signal regions defined in this thesis. These events only enter the SR_High region, with an expected yield of zero for the other three signal regions. The expected number of QCD events in SR_High is:

$$N_{QCD} = 2.84 \pm 2.75$$

Here, the uncertainty is evaluated from the relative uncertainty quoted in the reference [2].

9.1.2 W+jets background estimation

W+jets contribution comes predominantly from $W \rightarrow \tau\nu$ and is modelled using MC. These events typically pass the 2 medium RNN ID τ signal region requirement if they contain one real τ from the W-decay and one misidentified fake τ from a jet. However, the low misidentification rate makes the signal region requirement discard many of these events. This might be an advantage in general, but limited statistics result in an increase in statistical uncertainty for the W+jets MC samples and consequently also for the total background. Hence, to increase the statistical precision of the W+jets background the so-called "tau promotion" method has been employed. The following will be a brief summary of the method and a more complete description can be found in [90].

In short this methods recovers some of the discarded events by promoting fake τ-leptons satisfying the loose ID requirement to medium ID signal τ-leptons. To only promote fake τ-leptons that does not originate from the W-decay truth matching requirements are imposed. In the $W(\tau\nu)$ MC samples truth variables are kept which can be matched with the reconstructed objects. A reconstructed τ-lepton is considered real if it is matched with a true $\tau-$lepton within a spatial separation of $\Delta R < 0.2$, and considered fake otherwise.

Only τ-leptons satisfying the latter requirement are considered for promotion. For each event containing only one real medium ID τ-lepton and no medium ID fake τ-leptons, one fake τ-lepton is chosen at random and artificially promoted to a signal τ-lepton. The events are then reweighed to restore the correct normalization and match the nominal prediction. Figure 9.2 depicts this two-step procedure schematically. The events containing two signal τ-leptons are represented by the green boxes. The smaller boxes represents individual events and the size of the boxes the corresponding event weights. Fake $\tau-$leptons are promoted and reweighted such that the last green box contains more events than the initial one.

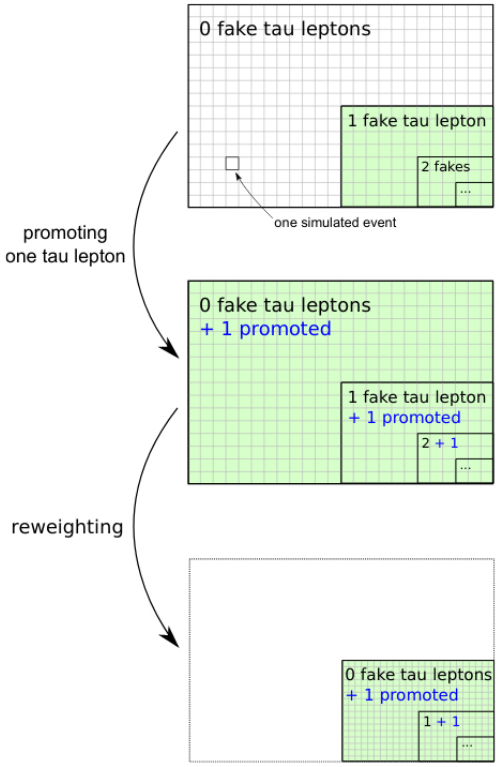


Figure 9.2: Schematic representation of the tau promotion method, taken from [90].

This increases the statistical precision of the W+jets background. The reweighting factor is computed from the fake rate ϵ, given as:

$$\epsilon = \frac{\text{Number of medium ID fake } \tau\text{-leptons}}{\text{Number of reconstructed fake } \tau\text{-leptons}} \tag{9.3}$$

As the fake rate has some dependence on the $p_T(\tau)$ and E_T^{miss}, ϵ is computed in distinct bins of these quantities. The fake rate also depends on the number of tracks associated with the τ-leptons, and possibly on data and pileup conditions in the three MC campaigns. Hence, the computation is done separately for 1-prong and 3-prong τ-leptons in each of the three MC campaigns. Furthermore, the $W(\tau\nu)$ samples are simulated in 6 ranges of $\max[p_T(W), H_T]$, where $p_T(W)$ is the transverse momentum of the W boson at parton level and H_T the scalar p_T sum of all jets at parton level with $p_T > 20$ GeV. This splitting is preserved in the fake rate computation which is done separately for the 6 ranges. The resulting fake rates are stored in 36 2D histograms in the $p_T(\tau) - E_T^{miss}$ plane, whose bin width gets larger for increasing values of the two variables due to decreasing statistics. The 2D histograms for the MC16d campaign can be seen in figure 9.3 and 9.4 for 1-prong and 3-prong τ-leptons respectively. It can be seen from these figures that some bins have fake rate values of either exactly 1 or 0. These values are not physical, but a result of lack of events in the bins when computing the average fake rate. However, this is rare and only the case for less than 3% of the τ-leptons.

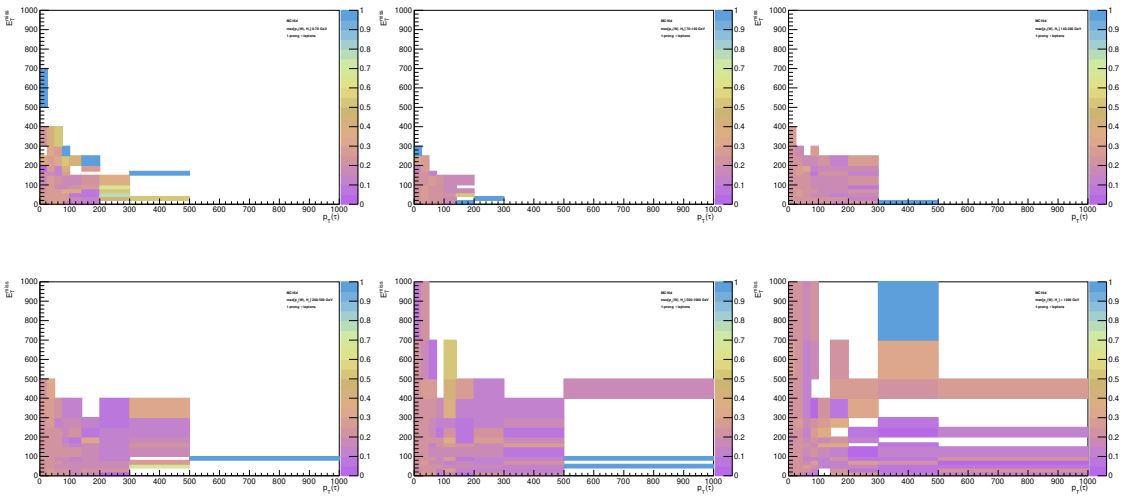


Figure 9.3: The 2D histograms storing the fake rate ϵ for 1-prong τ-leptons in bins of $p_T(\tau)$ and E_T^{miss} for different slices of $max[p_T(W), H_T]$ in the MC16d campaign.

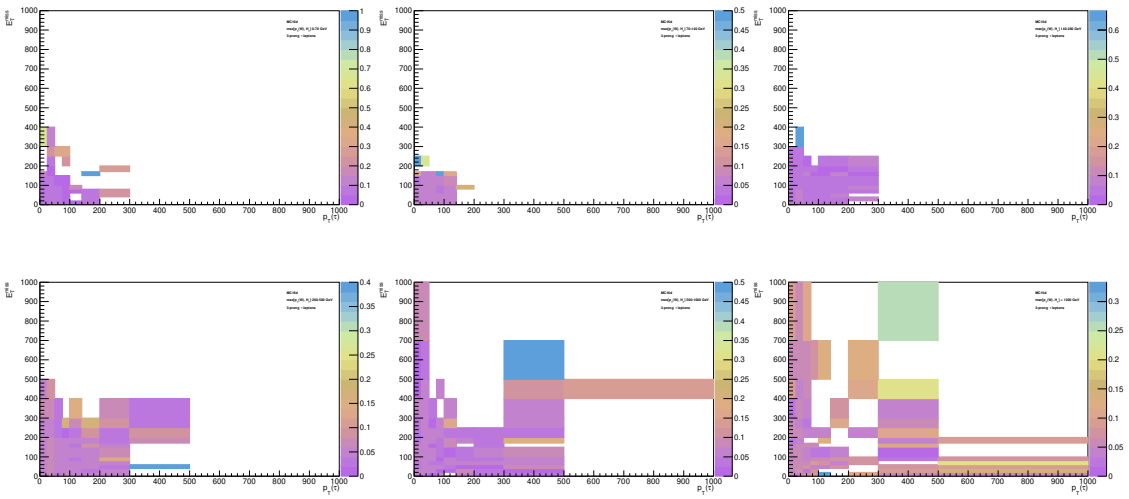


Figure 9.4: The 2D histograms storing the fake rate ϵ for 3-prong τ-leptons in bins of $p_T(\tau)$ and E_T^{miss} for different slices of $max[p_T(W), H_T]$ in the MC16d campaign.

The reweighting factor w is introduced to correctly normalize the promoted events. This is computed as follows. Assume that there is a total of n reconstructed fake τ-leptons in an event, k of which are signal τ-leptons passing the medium ID requirement. After randomly promoting one of the $n-k$ loose ID τ-leptons there are now $k+1$ medium ID signal τ-leptons in the event. The probability of observing k fake signal τ-leptons, and $n-k$ fake non-signal τ-leptons is given by the binomial probability distribution:

$$P(k) = \epsilon^k (1-\epsilon)^{n-k} \binom{n}{k} \tag{9.4}$$

After promoting one of $n-k$ non-signal τ-leptons, this changes to:

$$P(k+1) = \epsilon^{k+1} (1-\epsilon)^{n-k-1} \binom{n}{k+1} \tag{9.5}$$

The reweighting factor w is then the ratio of these two probabilities, since it needs to correct the probability of a promoted event in equation 9.5 to the nominal one in equation 9.4:

$$\begin{aligned} P(k+1) &= P(k) \cdot w \\ \rightarrow w &= \frac{P(k+1)}{P(k)} = \frac{\epsilon}{1-\epsilon} \cdot \frac{n-k}{k+1} \end{aligned} \tag{9.6}$$

Thus, w depends on both the fake rate ϵ and a factor C, which can be interpreted as:

$$C = \frac{n-k}{k+1} = \frac{\text{Number of non-signal } \tau\text{-leptons before promotion}}{\text{Number of signal } \tau\text{-leptons after promotion}} \tag{9.7}$$

As this analysis require exactly 2 medium ID τ-leptons C has a value of 1/2 in all cases considered here, even though it in principle may vary from event to event. For w to finite, fake rates with value $\epsilon = 1$ are set to zero in the above calculation.

The predicted event yields for the promoted sample generally agree with the nominal one. Still, to enforce an exact agreement between the two, the scale factor $SF^{offl} = \frac{N_{nominal}}{N_{promoted}} = 1.01$ is applied to the promoted

samples. The event yields in the ratio are computed by requiring 2 medium ID τ-leptons, where for the promoted sample one is promoted. In addition, the requirement of a positive trigger decision is omitted for the promoted sample but rather incorporated in an additional trigger scale factor SF^{trig} instead. It is done this way because the trigger efficiency may depend on the τ-lepton ID. The SF^{trig} scale factor is computed from the nominal sample as the ratio of the number of events satisfying the trigger plateau cuts, signal definition and a positive trigger decision to the number of events satisfying the same requirements without any requirement on the trigger decision. This gives a scale factor of $SF^{trig} = 0.66$. Both scale factors are applied to all events in all campaigns and are the same in all cases.

The expected event yields and the corresponding statistical uncertainties for the promoted and nominal $W(\tau\nu)$ samples after applying the sequential signal region requirements are listed in table 9.1.

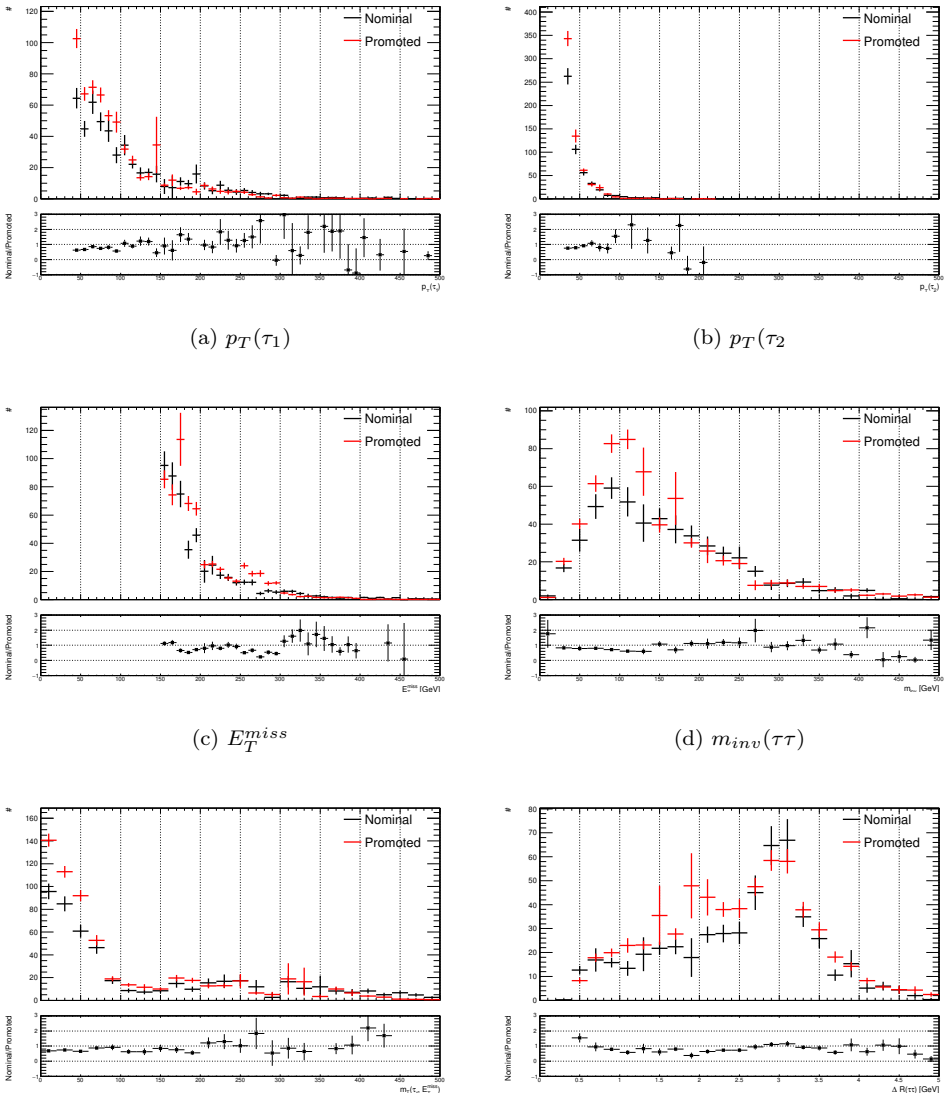
Region	Requirements	Nominal Sample		Promoted Sample		Ratio
	Nominal (Promoted) sample	yield	rel. unc. [%]	yield	rel. unc. [%]	$\frac{\text{Promoted}}{\text{Nominal}}$
All signal regions	2 signal τ-leptons (1 signal and 1 promoted τ)	496700 ± 5200	1.1	491700 ± 3100	0.6	1
	Leading τp_T	111900 ± 2100	1.9	120000 ± 1200	1	1.1
	Subleading τp_T	51600 ± 1400	2.7	56970 ± 790	1.4	1.1
	$E_T^{miss} > 150$ GeV	555.6 ± 22.7	4.1	674.9 ± 23.6	3.5	1.2
	Light lepton veto	555.3 ± 22.7	4.1	674.9 ± 23.6	3.5	1.2
	bjet veto	506.9 ± 22.4	4.4	613.8 ± 23.3	3.8	1.2
	Trigger ($SF^{trig} = 0.66$)	336.5 ± 18.7	5.6	407.4 ± 15.5	3.8	1.2
	Opposite charge	247.6 ± 15.9	6.4	261.0 ± 12.1	4.6	1.1
	$m_{inv}(\tau\tau) <127$ GeV	97.5 ± 11.6	12	134.99 ± 4.93	3.7	1.4
SR_High	$\Delta R(\tau\tau) <2$	53.88 ± 9.67	18	72.92 ± 3.33	4.6	1.4
	$m_T(\tau_1, E_T^{miss}) >320$ GeV	-6.83 ± 7.09	-103.8	0.58 ± 0.33	56.8	-0.1
SR_Low	$\Delta R(\tau\tau) <2$	53.88 ± 9.67	18	72.92 ± 3.33	4.6	1.4
	$m_T(\tau_1, E_T^{miss}) >120$ GeV	4.89 ± 8.64	176.6	11.15 ± 2.15	19.3	2.3
	$m_T(\tau_1, E_T^{miss}) <320$ GeV	11.72 ± 4.93	42.1	10.56 ± 2.13	20.2	0.9
SR_BDT_High	bdt_response_high> 0.1	0.12 ± 0.11	100	1.72 ± 1.07	62.3	14.4
SR_BDT_Low	bdt_response_low> 0.12	6.15 ± 1.48	24.1	4.41 ± 1.42	32.1	0.7

Table 9.1: Comparison of event yields from the nominal and promoted samples after applying the selection for the signal regions. Only statistical uncertainties are shown.

It can be seen from the table that a general decrease in statistical uncertainty is obtained. This is especially the case in SR_High , where the nominal W+jets sample predicts a negative event yield, which is off course unphysical, and results in a substantial statistical uncertainty. The promotion method resolves this problem and decreases the statistical uncertainty considerably. An agreement between the nominal and promoted sample is observed in the beginning of the table, but due to lack of statistics in the later stages of selection the two yields start to deviate further down the cut flow. An additional systematic uncertainty of 25% is therefore added to the promoted sample as suggested in section 6.3 in [90]. According to the reference this conservative uncertainty reflects the level of agreement observed in the regions where the statistical uncertainty of the nominal sample is not exceedingly large.

A comparison of the promoted and nominal samples are shown for some relevant kinematic variables in figure 9.5. The selection in table 9.1 is applied up to the trigger cut, and is the same for all distributions.

The upper panel shows the expected event yields for the two samples, and the lower panel their ratio. A good agreement is observed between the promoted and nominal samples which ensures the validity of the method.



(a) $p_T(\tau_1)$ (b) $p_T(\tau_2$

(c) E_T^{miss} (d) $m_{inv}(\tau\tau)$

(e) $m_T(\tau_1, E_T^{miss})$ (f) $\Delta R(\tau\tau)$

Figure 9.5: Comparison of the nominal and promoted sample for six relevant kinematic variables. The upper panel shows the kinematic distributions for the two samples, and the lower panel their ratio.

9.1.3 Diboson background cross checks

The main background contribution in the signal regions come from diboson production. These processes contain true taus and possibly neutrinos originating from WW, WZ or ZZ decays which give rise to to signal-like signatures. In general diboson MC is modelled well, but to ensure the validity of the MC this will be double checked in the following.

Three control regions have been defined kinematically close to the signal region, one incorporating WZ and ZZ processes and two incorporating WW processes. It should be noted that these regions were defined in an earlier stage of the analysis based on the signal selection used in [2], which was optimized for the Z'2HDM analysis. Since then, the signal regions have been redefined to the selections discussed in chapter 8, and as such, only the WW/ZZ region is orthogonal to the SRs in this analysis by definition. From the SRs, it seems that SR_Low is kinematically closest to WW region 1 and that SR_High is kinematically closest to WW region 2. The 2HDM+a signal contamination in these two regions has been studied, and is generally low with relative contributions mostly around $5-8\%$, but for some specific signal points with bigger cross sections relative contributions reach as high as 20%. This is however mostly within the statistical uncertainties of the total background in WW region 1 and 2, which is roughly 18% and 22% for the two regions respectively. As such, due to time constraints and difficulties related to finding pure diboson regions, the WW regions are defined as is in [2]. For the final analysis however, orthogonal CRs and SRs are required, and the two WW regions would have to be redefined accordingly.

The control region and initial signal region definitions are summarized in table 9.2, and kinematic distributions for E_T^{miss} and m_T^{tot} can be seen in figure 9.6, 9.7 and 9.8.

All regions utilize a ditau selection which constrains the possible diboson decays. For the WZ and ZZ processes this implies possible light lepton final states, like $WZ(\nu\ell\tau\tau)$ and $ZZ(\tau\tau\ell\ell)$. As such a selection with at least one light lepton is used to define a combined control region for WZ and ZZ. The same approach does not work for WW since $WW(\tau\nu\tau\nu)$ is the only allowed tree level process as a result of the ditau selection. Hence a new variable, m_{T2}, is introduced to separate these events. This variable, called stransverse mass, is especially efficient for isolating events when there is E_T^{miss} arising from more than one particle in the decay, as is the case for the two neutrinos in the considered WW-decay. The missing energy splitting between the two neutrinos is unknown, so the missing transverse momentum $\boldsymbol{E}_T^{miss}$ is spitted in all possible combinations $\boldsymbol{q_T}$ and $\boldsymbol{E}_T - \boldsymbol{q_T}$ for the two neutrinos respectively. The m_T variable from equation 6.14 is computed for each possibility and combined to m_{T2} which is given by the following min-max relation:

$$m_{T2} = \min_{\boldsymbol{q_T}} \left\{ \max \left[m_T(\boldsymbol{p}_T(\tau_1), \boldsymbol{q}_T), m_T(\boldsymbol{p}_T(\tau_2), \boldsymbol{E}_T^{miss} - \boldsymbol{q}_T) \right] \right\} \tag{9.8}$$

Selections based on this variable should in principle provide some separation between the considered WZ/ZZ and WW processes. As such, a cut on $m_{T2} > 25$ GeV is included in two selections, which correspondingly defines two WW regions with rather limited statistics. However, the regions have $\sim 50\%$ diboson contribution, and were the optimal selections found that did not directly overlap with the initial signal region. Agreements within uncertainties are observed between data and MC, indicating that the diboson MC is modelled correctly in all three control regions.

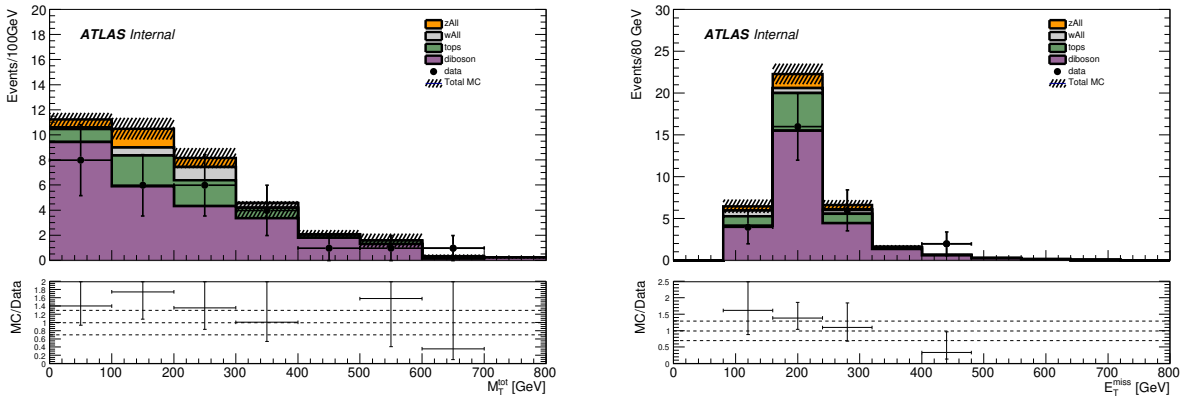
$N_\tau = 2$ (Medium ID, opposite charge) Trigger passed $p_T(\tau_1)$, $p_T(\tau_2) >$ threshold b-jet veto			
WZ/ZZ region	WW region 1	WW region 2	SR_init
$N_e + N_\mu \geq 1$ $E_T^{miss} > 150\,\text{GeV}$	N_e, $N_\mu = 0$ $E_T^{miss} > 150\,\text{GeV}$ $m_{T2} > 25$ GeV 150GeV$<m_T^{tot}<$250GeV $m_{inv}(\tau\tau) < 90\,\text{GeV}$ $p_T(\tau\tau) > 75\,\text{GeV}$	N_e, $N_\mu = 0$ $E_T^{miss} > 250\,\text{GeV}$ $m_{T2} > 25$ GeV $m_T^{tot} < 250\,\text{GeV}$	N_e, $N_\mu = 0$ $E_T^{miss} > 150\,\text{GeV}$ - $m_T^{tot} > 250$ GeV $m_{inv}(\tau\tau) < 127\,\text{GeV}$ $p_T(\tau\tau) > 50\,\text{GeV}$ $m_T(\tau_1, E_T^{miss}) > 150\,\text{GeV}$ $\Delta R(\tau\tau) < 2$

Table 9.2: Control region selection requirements. The original signal region on which the control region definitions are based is shown in the rightmost column (SR_init).

Table 9.3 shows that the dominant diboson processes in the control regions agrees with the ones in the signal regions. The processes corresponding to these Data Sample IDs (DSIDs) are listed in table 9.4 and shows that the dominant processes are those involving 2 charged leptons and 2 neutrinos or 4 charged leptons. This is as to expect from how the control regions were defined and assures the validity of the selections in table 9.2. The agreement of the DSIDs between the control and signal regions indicate that diboson processes also are well modelled in the signal regions. However again, more certainty can only be obtained by redefining the WW regions to be orthogonal to the SRs used in this analysis. Still, no further correction to the diboson background is applied and the diboson MC is used as is for the rest of the analysis.

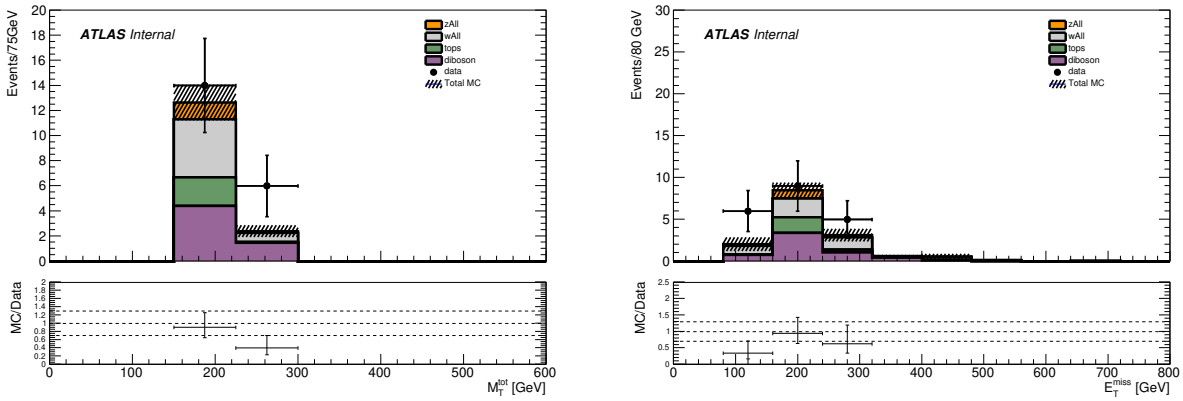
DSID	SR_Low [%]	SR_High [%]	SR_BDT_Low [%]	SR_BDT_High [%]	WZ/ZZ region [%]	WW region 1 [%]	WW region 2[%]	SR_init[%]
345723	1.6	3.0	0.8	3.4	0.0	0.6	0.1	2.3
363358	0.2	0.0	0.0	0.0	0.4	1.1	0.2	0.0
363359	0.2	0.0	0.0	0.0	0.2	0	4.3	0.0
364250	0.8	1.3	0.6	0.8	6.5	0.6	0.4	1.0
364253	35.4	29.8	25.5	31.2	80.1	51.5	33.1	41.6
364254	54.6	62.2	63.2	58.7	4.8	40.0	56.2	51.2
364284	1.4	1.1	1.3	1.4	4.4	2.4	2.0	1.7

Table 9.3: The main diboson DSID contributions in the signal and control regions.



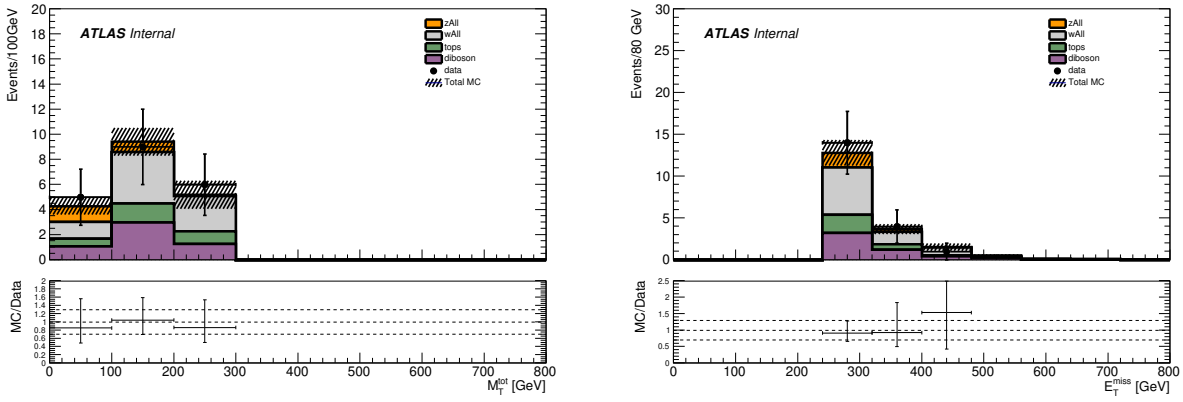
(a) m_T^{tot} (b) E_T^{miss}

Figure 9.6: Kinematic distributions in the WZ/ZZ control region (Upper panel) and ratio of MC and background (Lower Panel). The shaded area corresponds to the uncertainty in the total MC.



(a) m_T^{tot} (b) E_T^{miss}

Figure 9.7: Kinematic distributions in the WW control region 1 (Upper panel) and ratio of MC and background (Lower Panel). The shaded area corresponds to the uncertainty in the total MC.



(a) m_T^{tot} (b) E_T^{miss}

Figure 9.8: Kinematic distributions in the WW control region 2 (Upper panel) and ratio of MC and background (Lower Panel). The shaded area corresponds to the uncertainty in the total MC

DSID	Process
345723	gg→(ZZ→)$\ell\ell\nu\nu$ + 0,1j
363358	W(→qq)Z(→ $\ell\ell$) + 0,1j@NLO + 2,3j@LO
363359	W^+(→qq)W^-(→ $\ell\nu$) + 0,1j@NLO + 2,3j@LO
364250	$\ell\ell\ell\ell$ + 0,1j@NLO + 2,3j@LO
364253	$\ell\ell\ell\nu$ + 0,1j@NLO + 2,3j@LO
364254	$\ell\ell\nu\nu$ + 0,1j@NLO + 2,3j@LO
364284	$\ell\ell\ell\nu$ + jj + 0j@LO

Table 9.4: The corresponding processes to the DSIDs in Table 9.3.

9.2 Systematic uncertainties

Two categories of systematic uncertainties are considered in this analysis, namely experimental uncertainties which arise from uncertainties in reconstruction, identification and calibration of physics objects, and theoretical uncertainties that arise from uncertainties in the theoretical predictions for the signal and background models. These are evaluated in the framework of HistFitter, where systematic uncertainties are implemented as nuisance parameters which are constrained in a profile likelihood fit to "blinded" data, i.e the MC background expectation. The statistical tools used in the fit is briefly explained in section 6.1. Only uncertainty sources that contribute with more than 1% post-fit in the signal regions are considered. The following will briefly explain the sources of systematic uncertainties considered in this analysis and how they are estimated, a more complete description can be found in the supporting note [2] and the HistFitter paper [77].

Experimental uncertainties

Experimental uncertainties are estimated by varying the nominal MC prediction of a source by $\pm 1\sigma$, where the variations either are stored as weights or in separate ROOT trees. This way either two additional weights or trees are included in the samples corresponding to up and down variations of the nominal prediction. The experimental uncertainties considered in this analysis are the following:

- **Luminosity:** The total integrated luminosity for the datasets used in this analysis corresponds to 139 fb^{-1} with an associated uncertainty of 1.7%, [91]. This is added as a user input in HistFitter.
- **Pileup**: The pileup conditions of the simulated events are reweighted to match the data conditions. This reweighting brings with it some uncertainty which is provided by the pileup reweighting tool.
- **Jets**: Uncertainties related to the Jet Energy Scale (JES), Jet Vertex Tagger (JVT) efficiency, eta inter-calibration, response and pileup are considered. These are constrained by 16 different nuisance parameters.
- **Missing transverse energy:** Uncertainties related to the E_T^{miss}-soft term and scale are considered. This is constrained by one nuisance parameter.

- **Tau leptons:** Uncertainties in τ reconstruction and identification efficiencies and τ energy calibration are considered. These are constrained by 10 different nuisance parameters.

- **Flavour tagging:** Uncertainties related to efficiency of flavour tagging for b, c and light jets are considered. These are correspondingly constrained by three nuisance parameters.

Theoretical uncertainties

The theoretical uncertainties are implemented by employing generator weights that are computed during the event generation. The theoretical uncertainties considered in this analysis are the following:

- **PDF and α_s:** The generators operate based on Parton Distribution Functions (PDFs) that has some related uncertainty. As such, MC samples are generated with the `NNPDF3.0_NNLO` PDF set that contains the nominal set and 100 PDF variation weights. The PDF uncertainty in the SRs is evaluated as the RMS of the variations normalized to N_0, where N_0 is the nominal predicted yield and N_i is the yield corresponding to variation weight w_i.

$$\delta^{PDF} = \frac{1}{N_0}\sqrt{\frac{1}{100}\sum_{i=1}^{100}(N_i - N_0)^2} \tag{9.9}$$

  The uncertainties in the strong coupling constant α_s are computed by evaluating the difference in yields from α_s evaluated at two energy scales and taking the average of the deviation:

$$\delta^{\alpha_s} = \frac{1}{N_0}\frac{N(\alpha_s^{\text{down}}) - N(\alpha_s^{\text{up}})}{2} \tag{9.10}$$

  Here, the up-down values are: $\alpha_s^{\text{up}} = 0.119$ and $\alpha_s^{\text{down}} = 0.117$. The nominal value N_0 corresponds to $\alpha_s = 0.118$ which is determined experimentally at the Z mass peak. The two uncertainties are combined as follows:

$$\delta^{\alpha_s+PDF} = \sqrt{(\delta^{PDF})^2 + (\delta^{\alpha_s})^2} \tag{9.11}$$

  The above calculations are done externally and the resulting uncertainties are handed to HistFitter as user inputs.

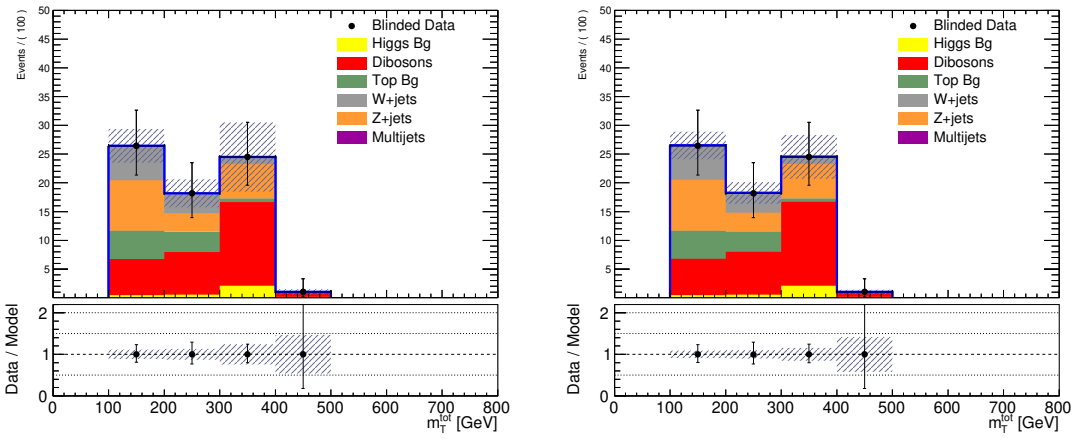
- **Scale uncertainties:** Uncertainties due to lack of higher order corrections are computed by varying the renormalization scale μ_r and factorization scale μ_f. In total seven variations are considered which are stored in 7 corresponding weights.

- **Generator uncertainties:** The uncertainty related to the choice of generator (Sherpa) is investigated by comparing it with the alternative Powheg generator. This is only done for the diboson sample as it is the dominant background in the SRs. The change in event yield is bellow 1% in the four SRs and hence assumed to be negligible .

9.2.1 Evaluation of systematic uncertainties

The evaluation of the different sources of systematic uncertainties is done with a background-only[17] profile likelihood fit on the blinded data in the signal regions, i.e with the data corresponding to the background expectation. As such the user defined systematic sources as well as the variation weights and trees are treated as nuisance parameters. Figure 9.9 to 9.12 show the pre- and post- fit distributions in the 4 signal regions, where 8-binned histograms for m_T^{tot} and the BDT scores are used in the fit for the cut- and BDT-based regions respectively. The resulting top 20 main systematic uncertainty sources are listed in table 9.5.

In all signal regions theoretical uncertainties related to the diboson samples, as well as experimental uncertainties related to tau energy calibration and tau RNN identification rank among the dominant systematic contributions. This is to expect as all SRs impose a medium tau ID criteria and have substantial diboson yields. The uncertainty related to the fake factor method is also sizable, but only relevant for SR_High where QCD background is present. This is also the case for theoretical uncertainties related to the Z+jets background and the promotion method uncertainty related to W+jets background, which with jet related systematics mostly are relevant in the low m_A signal regions where the Z/W+jets background have substantial yields.

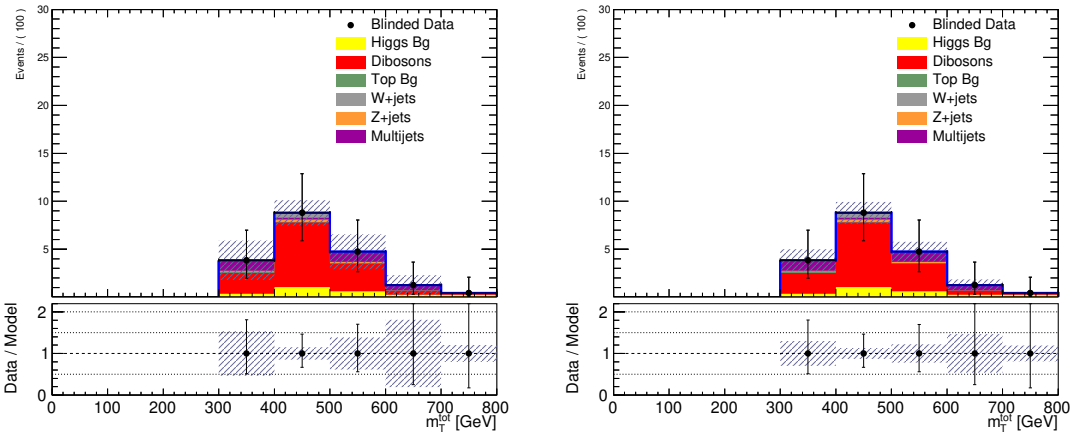


(a) SR_Low before fit

(b) SR_Low after fit

Figure 9.9: Distribution of m_T^{tot} in SR_Low before (left) and after (right) the fit to the "blinded" data. The shaded band represents the total uncertainty.

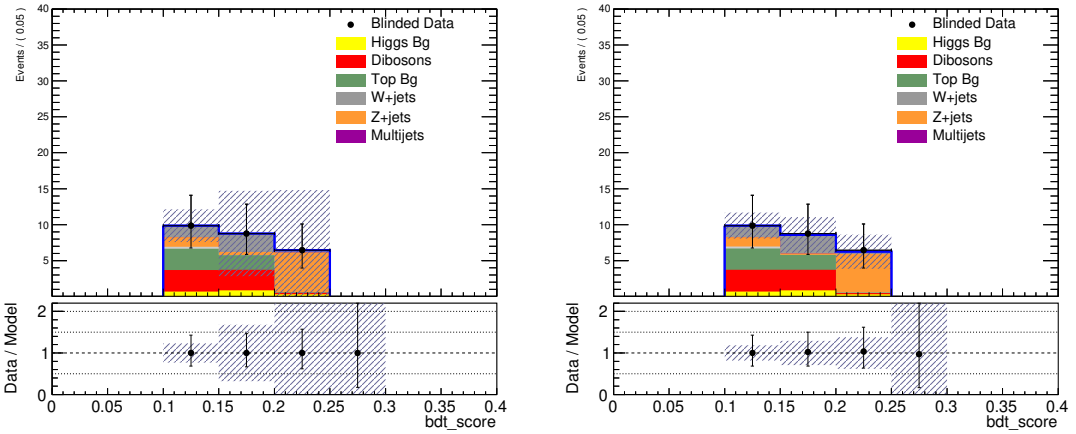
[17] Technically, Histfitter defines a background-only fit as a fit performed on data in CRs. However, the background fit used in this analysis is essentially the same, with exception of being performed to the MC background expectation ("blinded" data) in the SRs. As such the same naming convention will be used in this and the next chapter.



(a) SR_High before fit

(b) SR_High after fit

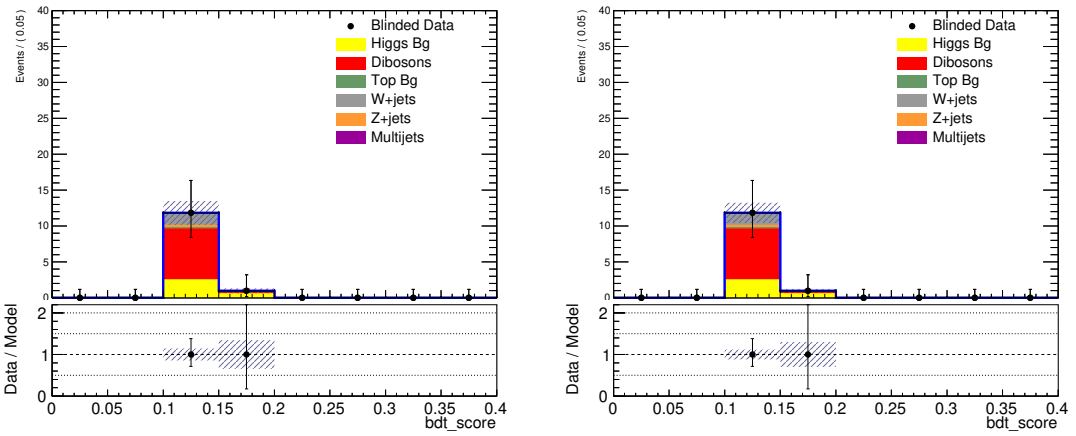
Figure 9.10: Distribution of m_T^{tot} in SR_Low before (left) and after (right) the fit to the "blinded data". The shaded band represents the total uncertainty.



(a) SR_BDT_Low before fit

(b) SR_BDT_Low after fit

Figure 9.11: Distribution of the BDT score in SR_BDT_Low before (left) and after (right) the fit to the "blinded" data. The shaded band represents the total uncertainty.



(a) SR_BDT_High before fit

(b) SR_BDT_High after fit

Figure 9.12: Distribution of the BDT score in SR_BDT_Low before (left) and after (right) the fit to the "blinded" data. The shaded band represents the total uncertainty.

Uncertainty of channel	SR_High	SR_Low	SR_BDT_High	SR_BDT_Low
Total background expectation	19.10	70.29	12.76	24.71
Total statistical ($\sqrt{N_{\text{exp}}}$)	±4.37	±8.38	±3.57	±4.97
Total background systematic	±2.91 [15.23%]	±5.46 [7.77%]	±1.51 [11.84%]	±3.56 [14.39%]
fakeFactorMethod_uncert	±2.76 [14.4%]	±0.00 [0.00%]	±0.00 [0.00%]	±0.01 [0.02%]
diboson_BG_PDF	±1.04 [5.5%]	±2.45 [3.5%]	±0.61 [4.8%]	±0.51 [2.1%]
tau_TES_PHYSICSLIST	±0.76 [4.0%]	±0.51 [0.72%]	±0.93 [7.3%]	±0.74 [3.0%]
diboson_BG_scales_rf	±0.19 [0.99%]	±1.50 [2.1%]	±0.32 [2.5%]	±0.72 [2.9%]
tau_RNNID_SYST	±0.43 [2.3%]	±1.33 [1.9%]	±0.31 [2.4%]	±0.36 [1.5%]
tau_ELEOLR_TOTAL	±0.44 [2.3%]	±1.03 [1.5%]	±0.28 [2.2%]	±0.29 [1.2%]
diboson_BG_renormalisation_scale	±0.30 [1.6%]	±1.49 [2.1%]	±0.20 [1.6%]	±0.56 [2.3%]
lumi	±0.26 [1.4%]	±1.00 [1.4%]	±0.18 [1.4%]	±0.34 [1.4%]
met_SoftTrk_Scale	±0.24 [1.2%]	±1.00 [1.4%]	±0.08 [0.65%]	±0.22 [0.87%]
jet_Pileup_RhoTopology	±0.14 [0.71%]	±1.04 [1.5%]	±0.16 [1.3%]	±0.27 [1.1%]
diboson_BG_factorisation_scale	±0.56 [2.9%]	±0.11 [0.16%]	±0.17 [1.3%]	±0.17 [0.69%]
Systematic_promotionMethod	±0.08 [0.42%]	±1.43 [2.0%]	±0.23 [1.8%]	±0.60 [2.4%]
jet_EffectiveNP_Modelling1	±0.21 [1.1%]	±0.88 [1.2%]	±0.13 [1.0%]	±0.58 [2.4%]
tau_RECO_TOTAL	±0.20 [1.0%]	±0.63 [0.89%]	±0.14 [1.1%]	±0.17 [0.68%]
tau_RNNID_1PRONGSTATSYSTPTGE40	±0.18 [0.97%]	±0.58 [0.82%]	±0.13 [0.99%]	±0.16 [0.65%]
pileup_reweight	±0.18 [0.96%]	±0.08 [0.12%]	±0.19 [1.5%]	±0.43 [1.7%]
jet_Flavor_Response	±0.15 [0.79%]	±1.38 [2.0%]	±0.13 [0.98%]	±0.46 [1.8%]
z_BG_PDF	±0.04 [0.20%]	±1.41 [2.0%]	±0.03 [0.20%]	±0.54 [2.2%]
z_BG_renormalisation	±0.06 [0.29%]	±1.24 [1.8%]	±0.04 [0.33%]	±0.08 [0.34%]
z_BG_scales_rf	±0.05 [0.25%]	±0.98 [1.4%]	±0.04 [0.32%]	±0.08 [0.31%]

Table 9.5: Breakdown of the 20 most dominant systematic uncertainties on background estimates in the four signal regions. Note that the individual uncertainties can be correlated, and do not necessarily add up quadratically to the total background uncertainty. The percentages show the size of the uncertainty relative to the total expected background.

Chapter 10

Results

The expected background event yields in the four signal regions are shown in table 10.1, where the corresponding uncertainties both incorporates statistical and systematic uncertainties. W_{Light}+jets corresponds to the background from W decays into light leptons, which was not considered in the promotion method.

Channel	SR_High	SR_Low	SR_BDT_High	SR_BDT_Low
Expected MC bkg events	19.10 ± 2.91	70.29 ± 5.46	12.76 ± 1.51	24.71 ± 3.56
Expected higgs events	2.50 ± 0.29	3.35 ± 0.45	3.21 ± 0.37	1.88 ± 0.27
Expected diboson events	12.38 ± 1.68	29.14 ± 3.66	7.25 ± 1.24	6.08 ± 1.28
Expected top events	0.29 ± 0.16	9.00 ± 1.17	0.24 ± 0.09	5.10 ± 1.05
Expected W+jets events	0.60 ± 0.08	10.69 ± 1.43	1.75 ± 0.23	4.47 ± 0.60
Expected W_{Light}+jets events	0.00 ± 0.00	0.11 ± 0.01	0.00 ± 0.00	$0.26^{+0.34}_{-0.26}$
Expected Z+jets events	0.49 ± 0.14	18.01 ± 2.82	0.32 ± 0.11	6.92 ± 3.06
Expected qcd events	2.84 ± 2.76	0.00 ± 0.00	0.00 ± 0.00	0.01 ± 0.01

Table 10.1: Expected event yields in the signal regions for the different background sources. Both statistical and systematic uncertainties are included. Uncertainties on the fitted yields are symmetric by construction, where the negative error is truncated when reaching to zero event yield.

10.1 Fit setup

The sensitivity of the analysis is evaluated in HistFitter, where fits are performed on the background MC prediction to eventually set expected limits on the 2HDM+a signal models. The background-only fit from section 9.2.1 is evaluated as a test for the fit setup. The event yields in the SRs before and after the fit are unchanged as expected since the blinded data corresponds to the MC prediction. Figure 10.1 shows the values of systematic and statistical uncertainty parameters used in the fit. All have values close to 1, indicating that they do not have a strong influence on the fit.

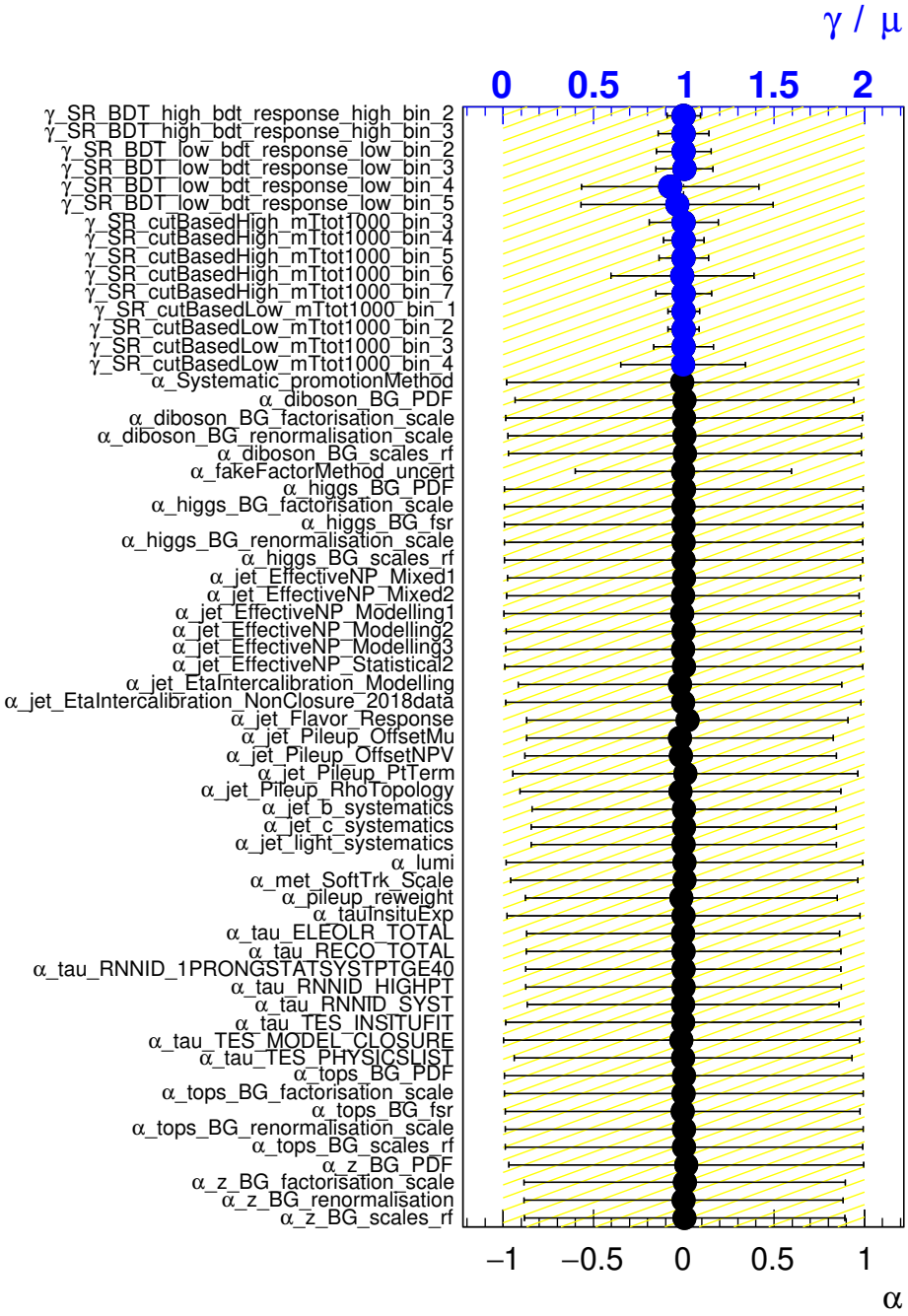


Figure 10.1: Fit parameters for in the background-only fit. Parameters related to statistical and systematic uncertainties are shown in blue and black respectively.

Furthermore, Histfitter also computes correlations between all nuisance parameters and plots this in a correlation matrix post-fit. This is shown in figure 10.2 for the background-only fit. To make the plot readable only correlations larger than 10% are included. In general values are distributed around and close to 0, which indicates that there are no strong correlation between the uncertainty sources used in the fit.

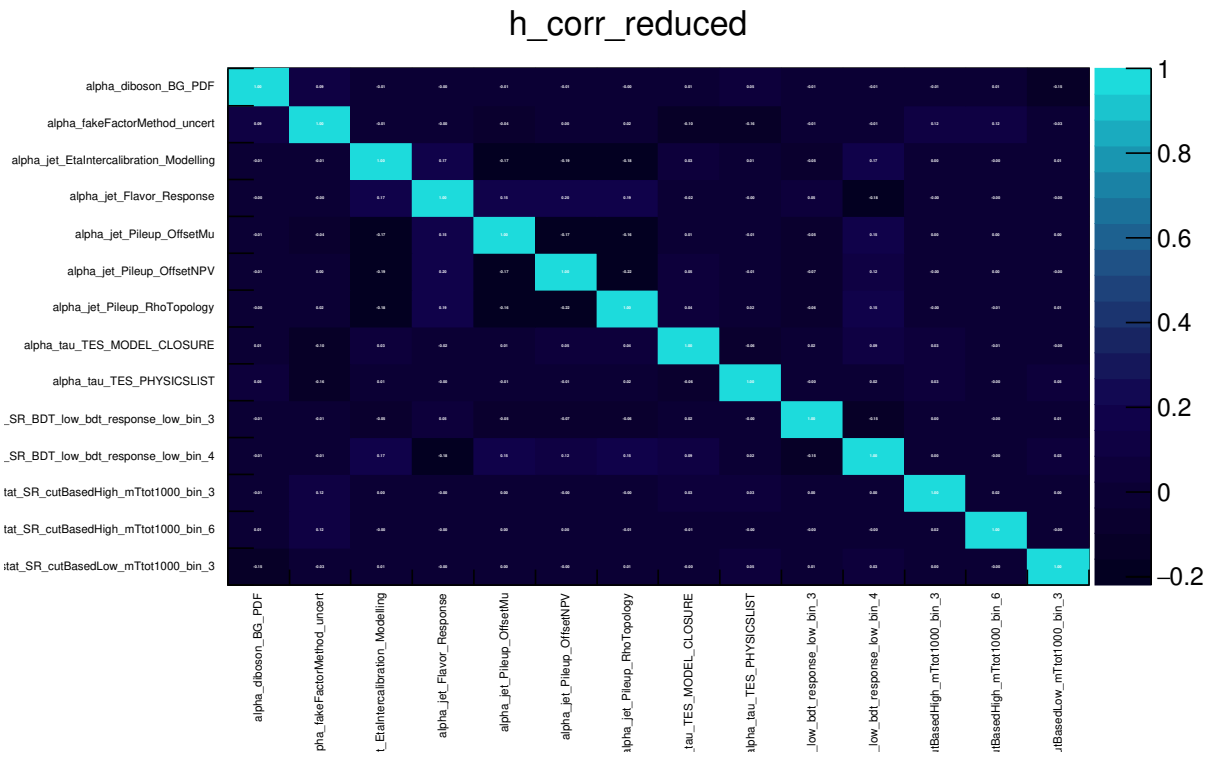
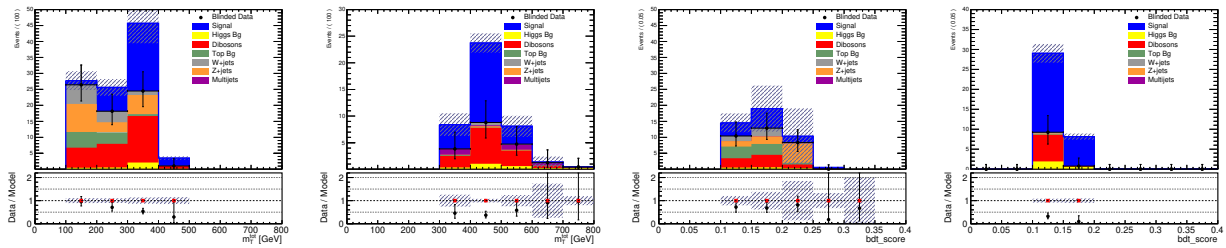


Figure 10.2: Correlation of different the systematic sources which are included as nuisance parameters in the background-only fit. Only correlations larger than 20% are included.

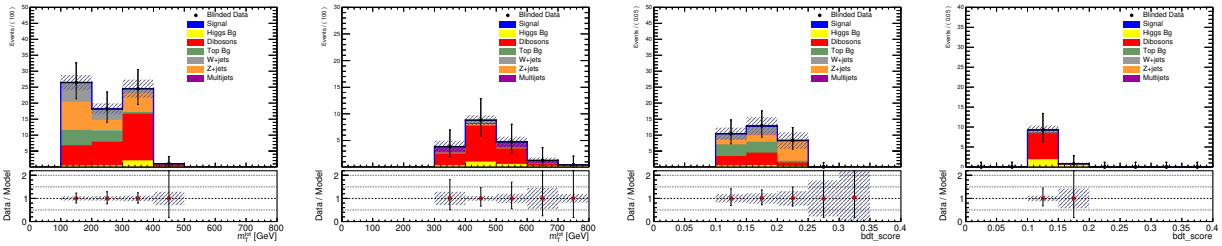
10.1.1 Exclusion fit including the signal

An additional test of the fit setup is done by running an exclusion fit on the 2HDM+a signal point $[m_A, m_a] = [600, 100]$ GeV as described in section 6.1. The contribution from bb production is also included at this stage for completeness, even though this in principle is negligible. Figure 10.3 and 10.4 show the distributions before and after the fit respectively. Again the data corresponds to the background MC prediction. The signal contributions are "fitted away" in all four signal regions as can be seen in figure 10.4 which is expected behaviour. In addition, a scan on the upper limit of the signal strength μ is performed for 20 possible values. The result from running this scan on the $[m_A, m_a] = [600, 100]$ GeV signal point is shown in figure 10.5. Values for μ where the p-value goes bellow 0.05 are indicated by the red line. These μ



(a) SR_Low (b) SR_High (c) SR_BDT_Low (d) SR_BDT_High

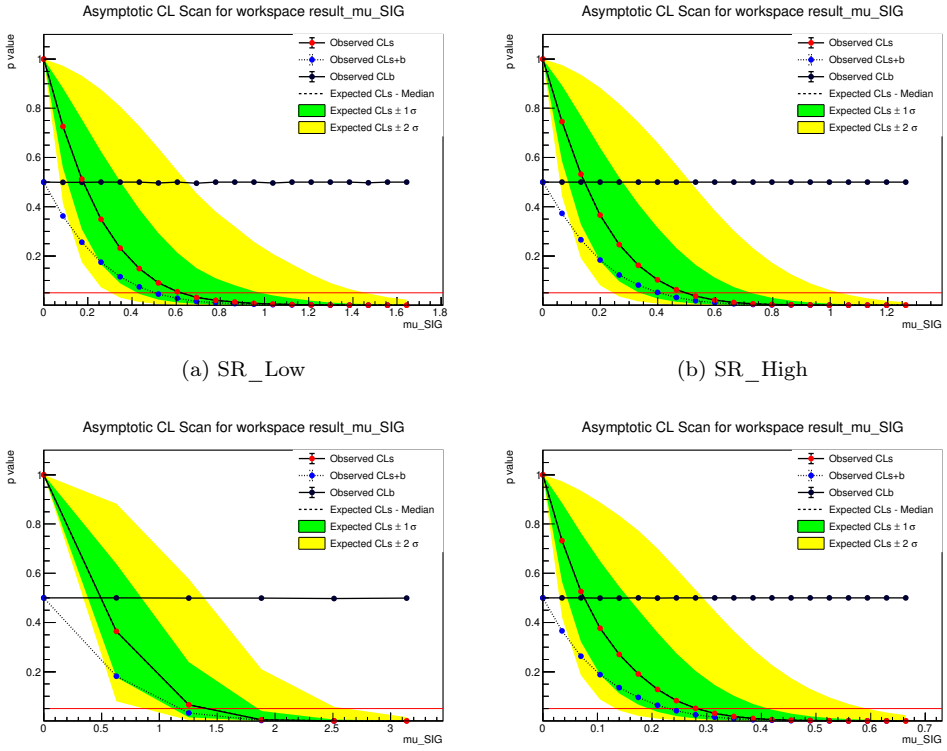
Figure 10.3: Signal region distributions before the fit to the "blinded" data with signal point $[m_A, m_a]$=[600,100] GeV. The axis and legend labels are the same as in figure 9.10.



(a) SR_Low (b) SR_High (c) SR_BDT_Low (d) SR_BDT_High

Figure 10.4: Signal region distributions after the fit to the "blinded" data with signal point $[m_A, m_a]$=[600,100] GeV. The axis and legend labels are the same as in figure 9.10.

values are excluded at 95% confidence level on the blinded data. The signal scan show that all signal regions except SR_BDT_Low have sufficient sensitivity to exclude the $[m_A, m_a] = [600, 100]$ GeV signal point. This is again expected behaviour from the signal yields seen in figure 10.3. All in all the results of the tests look reasonable indicating a valid fit setup.



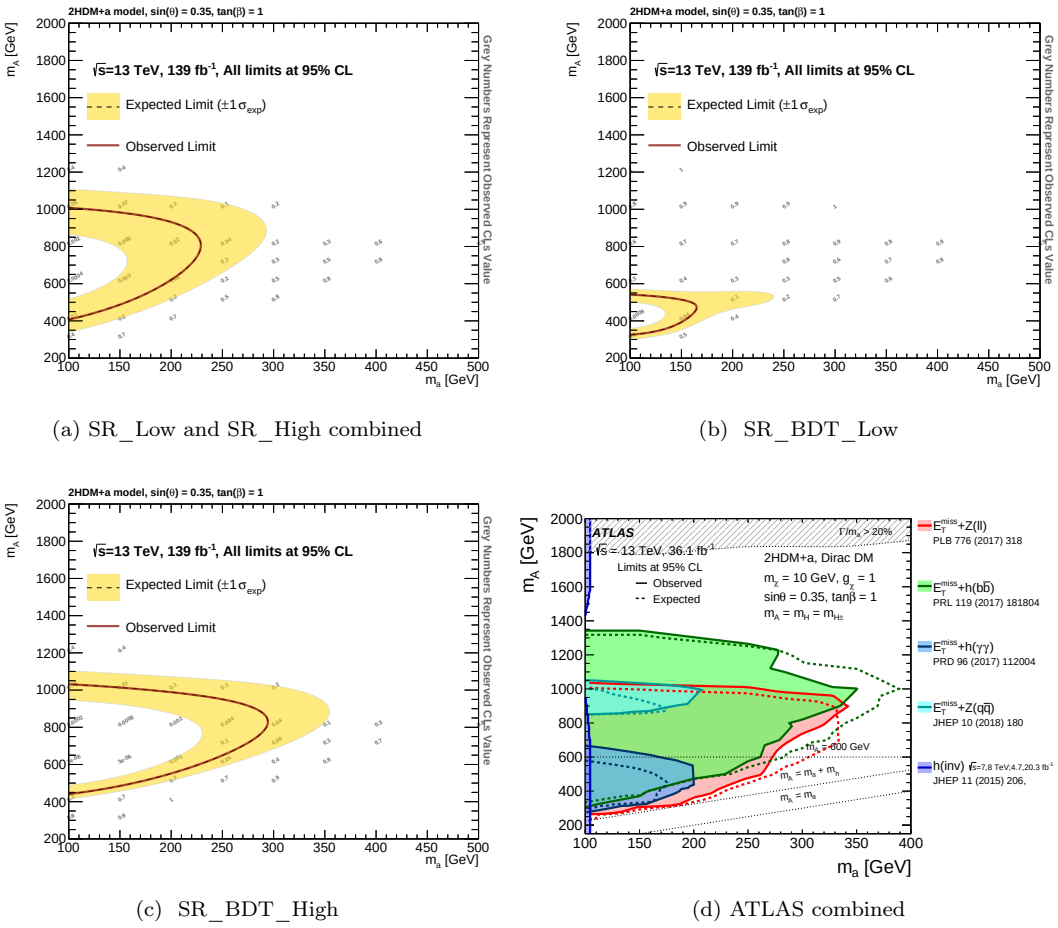
(a) SR_Low (b) SR_High

(c) SR_BDT_Low (d) SR_BDT_High

Figure 10.5: Upper limit scan for signal strength μ on the signal point $[m_A, m_a] = [600, 100]$ GeV in the four signal regions. Values of μ with to p-values bellow the red line are excluded at 95% confidence level.

10.2 Expected exclusion contours for the 2HDM+a signal

The exclusion fit is run together with a hypothesis test for all 33 2HDM+a signal points in the $m_A - m_a$ grid with $\sin(\theta) = 0.35$ and $\tan(\beta) = 1$. The result is combined to form expected exclusion contours, which is interpolated to determine where $CL_s = 0.05$, i.e exclusion at 95% confidence level. The expected contours



(a) SR_Low and SR_High combined

(b) SR_BDT_Low

(c) SR_BDT_High

(d) ATLAS combined

Figure 10.6: Expected exclusion contours for the 2HDM+a signal with the full-but blinded- Run 2 statistics 139 fb^{-1} in the four signal regions (top row and left bottom row). The combined exclusion from ATLAS [1] with a data set luminosity of 36.1 fb^{-1} is shown in the bottom right row.

with a $\pm 1\sigma$ band in the four signal regions are shown in figure 10.6, where also the combined contour from mono-X analyses in ATLAS [1] at 39.1 fb^{-1} is shown for comparison. The contour in the top left row is combined for SR_Low and SR_High . It can be seen from the figure that SR_BDT_High gives the best exclusion power in this analysis, with sensitivity up to $m_a = 300$ GeV. The combined contour for SR_High and SR_Low is also sensitive to many of the same central signal points, but suffers from low statistics for increasing m_a. SR_BDT_Low is sensitive to the lowest values of m_A, but the expected sensitivity estimated from Asimov significance shown in figure 8.14 is heavily reduced due the effect of the systematic uncertainties. Hence, the assumed 15% background uncertainty in the optimization studies in chapter 8 was an underestimate for SR_BDT_Low and consequently gave overestimated values for the expected sensitivity.

The ATLAS result on the bottom right of the figure show exclusion contours from different mono-Higgs, mono-Z and $h(inv)$ searches. The initial hope of this analysis was to be competitive with current limits in the low E_T^{miss} regime, especially with respect to $h(bb)$. These events have a higher higgs branching fraction ($\sim 5.8 \cdot 10^{-1}$), but are subject to substantial top and QCD background which generally require hard cuts on E_T^{miss}. The $h(\tau\tau)$ branching fraction is roughly an order of magnitude lower ($\sim 6.7 \cdot 10^{-2}$), but the $\tau\tau$ final state has less background and should in principle also retain sensitivity for lower E_T^{miss} values. However, as argued in section 8.1, the utilized E_T^{miss} trigger plateau cut discards large portions of the events in this E_T^{miss} regime. As such, no specific region in the signal grid has been targeted in this analysis.

By comparing the obtained contours in this analysis to the combined from ATLAS, a similar sensitivity for the central m_A signal points can be seen. The sensitivity from SR_BDT_High somewhat supplements $h(bb)$ sensitivity at higher m_a, but this analysis is in general not expected to be competitive with the current limits in any parts of the signal grid. It is however a study of a new 2HDM+a topology that as of yet not has been investigated in ATLAS. As such the expected contours in this analysis could be statistically combined with the other ATLAS results once fitted to data, and enhance the sensitivity in the regions where exclusion is obtained. This would off course require an unblinding approval from ATLAS, which at the time of writing is still pending.

Chapter 11

Conclusion and outlook

This thesis presented a search for the 2HDM+a mono-Higgs signal model where Dark Matter is produced in association with two tau leptons in the ATLAS detector. The analysis has employed the full Run 2 dataset collected by the ATLAS detector from 2015-2018, corresponding to an integrated luminosity 139 fb^{-1} and a center-of-mass energy of $\sqrt{s} = 13$ TeV. The analysis supplementary to the analysis in [2], which in addition to the 2HDM+a model, search for the same $h(\tau\tau)$ final state and with the Z'2HDM model.

The selection of signal has been optimized with respect to the Asimov significance with the hope of eventually setting limits on the model in the $m_A - m_a$ mass grid. This has been done in two ways, namely with a cut based and multivariate approach. The first approach involved a series of sequential cuts on a set of kinematic variables which led to the definition of two orthogonal signal regions targeting low and high m_A respectively. Due to a lack of sensitivity in parts of the mass grid, an additional multivariate approach was employed by the utilization of Boosted Decision Trees. These BDTs have been trained separately on samples with low and high m_A, and cuts on the corresponding BDT responses defines two additional non-orthogonal signal regions. Overall, this gives an enhancement of the sensitivity in comparison with the cut based approach.

The background in the four signal regions have been estimated with MC simulations, and corresponding systematic uncertainties arising from the experimental and theoretical framework of the analysis have been taken into consideration. To statistically test the signal models profile likelihood fits have been employed. As of writing this the analysis is still not been approved for unblinding by ATLAS, and as such the fits are only performed on Monte Carlo data, using an Asimov statistical approach. This has been done on all the 33 simulated signal samples to obtain expected contours where points in the mass $m_A - m_a$ are excluded at 95% confidence level. Overall, the expected exclusion power is predominantly in the central m_A, low m_a region, and the analysis is not really competitive with the current limits in any part of the mass grid. However, this is the first look into the $h(\tau\tau)$ final state for the 2HDM+a signal model with ATLAS data and as such the analysis is complimentary to previously performed ATLAS searches. Once the analysis gets approval for unblinding the current fit setup can in principle be run on the real data and the corresponding contours statistically combined with the current ATLAS results. This would require some additional work in the background estimation, where the diboson control regions would need to be redefined

to properly verify the validity of the MC estimation. In addition, the two signal regions optimized for low m_A have considerable top and Z+jets contributions, and as such the corresponding MC estimations should be investigated more thoroughly. Future work could also involve a study of alternative triggers to loosen the E_T^{miss} plateau selection requirement. This discards substantial amounts of signal events, especially from the signal points with high m_a and low m_A, making it hard to obtain sufficient sensitivity in this part of the mass grid. Triggers that allow for a looser E_T^{miss} requirement could increase the statistical sensitivity of this analysis.

Bibliography

[1] M. Aaboud et al. "Constraints on mediator-based dark matter and scalar dark energy models using

$$\sqrt{s}$$

= 13 TeV pp collision data collected by the ATLAS detector". In: *Journal of High Energy Physics* 2019.5 (2019). ISSN: 1029-8479. DOI: 10.1007/jhep05(2019)142. URL: http://dx.doi.org/10.1007/JHEP05(2019)142.

[2] Erlend Aakvaag, Julia Isabell Djuvsland, and Anna Lipniacka. *Search for Dark Matter produced in association with a Higgs boson decaying to tau leptons at s = 13 TeV with the ATLAS detector*. Tech. rep. ATL-COM-PHYS-2020-031. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2707037.

[3] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[4] Tommy Ohlsson. *Relativistic Quantum Physics: From Advanced Quantum Mechanics to Introductory Quantum Field Theory*. Cambridge University Press, 2011. DOI: 10.1017/CBO9781139032681.

[5] D. Hanneke, S. Fogwell Hoogerheide, and G. Gabrielse. "Cavity control of a single-electron quantum cyclotron: Measuring the electron magnetic moment". In: *Physical Review A* 83.5 (2011). ISSN: 1094-1622. DOI: 10.1103/physreva.83.052122. URL: http://dx.doi.org/10.1103/PhysRevA.83.052122.

[6] Franz Mandl and Graham Shaw. *QUANTUM FIELD THEORY*. Jan. 1985.

[7] Vernon D. Barger and R.J.N. Phillips. *COLLIDER PHYSICS*. 1987. ISBN: 978-0-201-14945-6.

[8] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[9] John F. Donoghue. "The effective field theory treatment of quantum gravity". In: (2012). DOI: 10.1063/1.4756964. URL: http://dx.doi.org/10.1063/1.4756964.

[10] Planck Collaboration et al. "Planck 2015 results. XIII. Cosmological parameters". In: 594, A13 (Sept. 2016), A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589 [astro-ph.CO].

[11] Adrià Gómez-Valent. *Vacuum energy in Quantum Field Theory and Cosmology*. 2017. arXiv: 1710.01978 [gr-qc].

[12] Giorgio Arcadi et al. "The waning of the WIMP? A review of models, searches, and constraints". In: *European Physical Journal C* 78.3, 203 (Mar. 2018), p. 203. DOI: 10.1140/epjc/s10052-018-5662-y. arXiv: 1703.07364 [hep-ph].

[13] G. Jungman, M. Kamionkowski, and K. Griest. "Supersymmetric dark matter". In: 267 (Mar. 1996), pp. 195–373. DOI: 10.1016/0370-1573(95)00058-5. arXiv: hep-ph/9506380 [hep-ph].

[14] G. C. Branco et al. "Theory and phenomenology of two-Higgs-doublet models". In: 516.1-2 (July 2012), pp. 1–102. DOI: 10.1016/j.physrep.2012.02.002. arXiv: 1106.0034 [hep-ph].

[15] T. S. van Albada et al. "Distribution of dark matter in the spiral galaxy NGC 3198." In: 295 (Aug. 1985), pp. 305–313. DOI: 10.1086/163375.

[16] T. D. Lee and C. N. Yang. "Question of Parity Conservation in Weak Interactions". In: *Phys. Rev.* 104 (1 1956), pp. 254–258. DOI: 10.1103/PhysRev.104.254. URL: https://link.aps.org/doi/10.1103/PhysRev.104.254.

[17] H Tananbaum et al. "Highlights and Discoveries from the Chandra X-ray Observatory". In: *Reports on Progress in Physics* 77 (May 2014). DOI: 10.1088/0034-4885/77/6/066902.

[18] C. L. Bennett et al. "Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results". In: 208.2, 20 (Oct. 2013), p. 20. DOI: 10.1088/0067-0049/208/2/20. arXiv: 1212.5225 [astro-ph.CO].

[19] Stefano Giagu. "WIMP Dark Matter Searches With the ATLAS Detector at the LHC". In: *Front. in Phys.* 7 (2019), p. 75. DOI: 10.3389/fphy.2019.00075.

[20] Nicolò Trevisani. "Collider Searches for Dark Matter (ATLAS + CMS)". In: *Universe* 4.11 (2018), p. 131. DOI: 10.3390/universe4110131.

[21] Giorgio Busoni et al. "On the validity of the effective field theory for dark matter searches at the LHC, part II: complete analysis for the s-channel". In: *Journal of Cosmology and Astroparticle Physics* 2014.06 (2014), 060–060. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2014/06/060. URL: http://dx.doi.org/10.1088/1475-7516/2014/06/060.

[22] Jalal Abdallah et al. "Simplified models for dark matter searches at the LHC". In: *Physics of the Dark Universe* 9-10 (2015), pp. 8 –23. ISSN: 2212-6864. DOI: https://doi.org/10.1016/j.dark.2015.08.001. URL: http://www.sciencedirect.com/science/article/pii/S2212686415000163.

[23] John F. Gunion et al. *The Higgs Hunter's Guide*. Vol. 80. 2000.

[24] Heather E. Logan. "TASI 2013 lectures on Higgs physics within and beyond the Standard Model". In: (June 2014). arXiv: 1406.1786 [hep-ph].

[25] Martin Bauer, Ulrich Haisch, and Felix Kahlhoefer. "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators". In: *Journal of High Energy Physics* 2017.5 (2017). ISSN: 1029-8479. DOI: 10.1007/jhep05(2017)138. URL: http://dx.doi.org/10.1007/JHEP05(2017)138.

[26] Jose Miguel No. "Looking through the pseudoscalar portal into dark matter: Novel mono-Higgs and mono-Zsignatures at the LHC". In: *Physical Review D* 93.3 (2016). ISSN: 2470-0029. DOI: 10.1103/physrevd.93.031701. URL: http://dx.doi.org/10.1103/PhysRevD.93.031701.

[27] Tomohiro Abe et al. *LHC Dark Matter Working Group: Next-generation spin-0 dark matter models.* 2018. arXiv: 1810.09420 [hep-ex].

[28] Oliver Sim Brüning et al. *LHC Design Report.* CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1. URL: http://cds.cern.ch/record/782076.

[29] J. T. Boyd. *LHC Run-2 and Future Prospects.* 2020. arXiv: 2001.04370 [hep-ex].

[30] John Jowett. "Colliding Heavy Ions in the LHC". In: CERN-ACC-2018-121 (2018), TUXGBD2. 6 p. DOI: 10.18429/JACoW-IPAC2018-TUXGBD2. URL: http://cds.cern.ch/record/2648704.

[31] G. et al. Aad. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* (2008). URL: http://nordberg.web.cern.ch/nordberg/PAPERS/JINST08.pdf.

[32] S. Chatrchyan et al. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[33] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[34] Jr. Alves A.Augusto et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[35] O. Adriani et al. "The LHCf detector at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08006. DOI: 10.1088/1748-0221/3/08/S08006.

[36] James Pinfold. "The MoEDAL Experiment at the LHC". In: *EPJ Web of Conferences* 71 (Apr. 2014), p. 00111. DOI: 10.1051/epjconf/20147100111.

[37] Stefano Lami. "The TOTEM experiment at the LHC". In: (2009). Ed. by M. Campanelli, pp. 38–44.

[38] Esma Mobs. "The CERN accelerator complex. Complexe des accélérateurs du CERN". In: (2016). General Photo. URL: https://cds.cern.ch/record/2197559.

[39] *LHC-project-schedule Project Schedule.* https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. Accessed: 2020-11-01.

[40] Matthias Schott and Monica Dunford. "Review of single vector boson production in pp collisions at $\sqrt{s} = 7$ TeV". In: *The European Physical Journal C* (2014). DOI: 10.1140/epjc/s10052-014-2916-1. URL: https://doi.org/10.1140/epjc/s10052-014-2916-1.

[41] M.S. Alam et al. "ATLAS pixel detector: Technical design report". In: (May 1998).

[42] Alessandro La Rosa. *The ATLAS Insertable B-Layer: from construction to operation.* 2016. arXiv: 1610.01994 [physics.ins-det].

[43] Georges Aad et al. "Operation and performance of the ATLAS semiconductor tracker". In: *JINST* 9 (2014), P08009. DOI: 10.1088/1748-0221/9/08/P08009. arXiv: 1404.7473 [hep-ex].

[44] E. Abat et al. "The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance". In: *JINST* 3 (2008), P02013. DOI: 10.1088/1748-0221/3/02/P02013.

[45] Maria Fiascaris. "Performance of the ATLAS Tile Calorimeter in pp collisions at the LHC". In: *Journal of Physics: Conference Series* 587 (Feb. 2015). DOI: 10.1088/1742-6596/587/1/012010.

[46] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331061.

[47] M. Aharrouche et al. "Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 568.2 (2006), 601–623. ISSN: 0168-9002. DOI: 10.1016/j.nima.2006.07.053. URL: http://dx.doi.org/10.1016/j.nima.2006.07.053.

[48] T. Barillari. "The ATLAS liquid argon hadronic end-cap calorimeter: construction and selected beam test results". In: *Nuclear Physics B - Proceedings Supplements* 150 (2006), 102–105. ISSN: 0920-5632. DOI: 10.1016/j.nuclphysbps.2004.10.087. URL: http://dx.doi.org/10.1016/j.nuclphysbps.2004.10.087.

[49] *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331062.

[50] *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331068.

[51] G. Aielli et al. "The RPC first level muon trigger in the barrel of the ATLAS experiment". In: *Nucl. Phys. B Proc. Suppl.* 158 (2006). Ed. by S.K. Park, S.P. Ratti, and R. Santonico, pp. 11–15. DOI: 10.1016/j.nuclphysbps.2006.07.031.

[52] F. Bauer et al. "Construction and test of MDT chambers for the ATLAS muon spectrometer". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 461.1-3 (2001), 17–20. ISSN: 0168-9002. DOI: 10.1016/s0168-9002(00)01156-6. URL: http://dx.doi.org/10.1016/S0168-9002(00)01156-6.

[53] Satoshi Tanaka et al. "Techniques developed for the ATLAS thin gap chambers". In: *Nuclear Science, IEEE Transactions on* 51 (July 2004), pp. 934 –938. DOI: 10.1109/TNS.2004.829533.

[54] Theodoros Argyropoulos et al. "Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance". In: *IEEE Trans. Nucl. Sci.* 56 (2009), pp. 1568–1574. DOI: 10.1109/TNS.2009.2020861.

[55] Georges Aad et al. "Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *Eur. Phys. J. C* 76.11 (2016), p. 581. DOI: 10.1140/epjc/s10052-016-4395-z. arXiv: 1510.03823 [hep-ex].

[56] *ATLAS Public results Lumonosity public results run 2*. https://project-hl-lhc-industry.web.cern.ch/content/project-schedule. Accessed: 2020-08-31.

[57] Aranzazu Ruiz-Martinez and ATLAS Collaboration. *The Run-2 ATLAS Trigger System*. Tech. rep. ATL-DAQ-PROC-2016-003. Geneva: CERN, 2016. DOI: 10.1088/1742-6596/762/1/012003. URL: http://cds.cern.ch/record/2133909.

[58] Morad Aaboud et al. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79.8 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [physics.ins-det].

[59] M. Aaboud et al. "Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data". In: *The European Physical Journal C* 77.3 (2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-4756-2. URL: http://dx.doi.org/10.1140/epjc/s10052-017-4756-2.

[60] G. Aad et al. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at

$$\sqrt{s}$$

s =13 TeV". In: *The European Physical Journal C* 76.5 (2016). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-016-4120-y. URL: http://dx.doi.org/10.1140/epjc/s10052-016-4120-y.

[61] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-ktjet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (2008), 063–063. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2008/04/063. URL: http://dx.doi.org/10.1088/1126-6708/2008/04/063.

[62] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*. Tech. rep. ATL-PHYS-PUB-2015-045. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2064383.

[63] *E_T^{miss} performance in the ATLAS detector using 2015-2016 LHC p-p collisions*. Tech. rep. ATLAS-CONF-2018-023. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2625233.

[64] G. Aad et al. "Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector". In: *Journal of High Energy Physics* 2013.9 (2013). ISSN: 1029-8479. DOI: 10.1007/jhep09(2013)076. URL: http://dx.doi.org/10.1007/JHEP09(2013)076.

[65] *QCD (for LHC)Lecture 2: Parton Distribution Functions Parton Distribution Functions*. https://gsalam.web.cern.ch/gsalam/repository/talks/2009-Bautzen-lecture2.pdf. Accessed: 2020-11-01.

[66] G. Aad et al. "The ATLAS Simulation Infrastructure". In: *The European Physical Journal C* 70.3 (2010), 823–874. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1429-9. URL: http://dx.doi.org/10.1140/epjc/s10052-010-1429-9.

[67] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. "A brief introduction to PYTHIA 8.1". In: *Computer Physics Communications* 178.11 (2008), 852–867. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2008.01.036. URL: http://dx.doi.org/10.1016/j.cpc.2008.01.036.

[68] Paolo Nason. "A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms". In: *Journal of High Energy Physics* 2004.11 (2004), 040–040. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2004/11/040. URL: http://dx.doi.org/10.1088/1126-6708/2004/11/040.

[69] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *Journal of High Energy Physics* 2014.7 (2014). ISSN: 1029-8479. DOI: 10.1007/jhep07(2014)079. URL: http://dx.doi.org/10.1007/JHEP07(2014)079.

[70] *particle wiki MadGraph5_aMCatNLO*. https://particle.wiki/wiki/MadGraph5_aMC@NLO. Accessed: 2020-08-31.

[71] T Gleisberg et al. "Event generation with SHERPA 1.1". In: *Journal of High Energy Physics* 2009.02 (2009), 007–007. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2009/02/007. URL: http://dx.doi.org/10.1088/1126-6708/2009/02/007.

[72] S. Agostinelli et al. "GEANT4–a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[73] Takashi Yamanaka. "The ATLAS calorimeter simulation FastCaloSim". In: *J. Phys. Conf. Ser.* 331 (2011). Ed. by Simon C. Lin, p. 032053. DOI: 10.1088/1742-6596/331/3/032053.

[74] A Buckley et al. "Implementation of the ATLAS Run 2 event data model". In: *Journal of Physics: Conference Series* 664.7 (2015), p. 072045. DOI: 10.1088/1742-6596/664/7/072045. URL: https://doi.org/10.1088%2F1742-6596%2F664%2F7%2F072045.

[75] *The ROOT framework The ROOT framework*. https://root.cern/. Accessed: 2020-08-31.

[76] Andreas Hoecker et al. "TMVA: Toolkit for Multivariate Data Analysis". In: *PoS* ACAT (2007), p. 040. arXiv: physics/0703039.

[77] M. Baak et al. "HistFitter software framework for statistical data analysis". In: *The European Physical Journal C* 75.4 (2015). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3327-7. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3327-7.

[78] Kyle Cranmer et al. "HistFactory: A tool for creating statistical models for use with RooFit and RooStats". In: (June 2012).

[79] *Roostat documentation*. https://root.cern/doc/master/group__Roostats.html. Accessed: 2020-11-01.

[80] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[81] S.S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360.

[82] *Glen Cowan Discovery sensitivity for a counting experimentwith background uncertainty*. https://www.pp.rhul.ac.uk/~cowan/stat/notes/medsigNote.pdf. Accessed: 2020-08-31.

[83] Alexander L. Read. "Presentation of search results: The CL(s) technique". In: *J. Phys. G* 28 (2002). Ed. by M.R. Whalley and L. Lyons, pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.

[84] *Running periods Data and Monte Carlo Datasets for Analysis*. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/DataMCForAnalysis?sortcol=1;table=1;up=1#sorted_table, note = Accessed: 2020-08-31.

[85] *Decision Tree Ensembles- Bagging and Boosting Decision Tree Ensembles- Bagging and Boosting*. https://towardsdatascience.com/decision-tree-ensembles-bagging-and-boosting-266a8ba60fd9. Accessed: 2020-08-31.

[86] R. Schapire. "Explaining AdaBoost". In: *Empirical Inference*. 2013.

[87] J. Quinlan. "Bagging, Boosting, and C4.5". In: *AAAI/IAAI, Vol. 1*. 1996.

[88] *ROC curve wiki ROC curve*. https://en.wikipedia.org/wiki/Receiver_operating_characteristic. Accessed: 2020-11-01.

[89] "Kolmogorov–Smirnov Test". In: *The Concise Encyclopedia of Statistics*. New York, NY: Springer New York, 2008, pp. 283–287. ISBN: 978-0-387-32833-1. DOI: 10.1007/978-0-387-32833-1_214. URL: https://doi.org/10.1007/978-0-387-32833-1_214.

[90] Yang Liu et al. *Improved description of the di-tau final state in events with associated production of a W boson and jets in the ATLAS detector using the tau-promotion method*. Tech. rep. ATL-COM-PHYS-2019-769. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2680467.

[91] *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2019-021. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2677054.

List of Figures

List of Tables