Search for Charged Higgs through a Reinterpretation of a Dark Meson Search with the ATLAS Experiment Based on 140 fb^{-1} of LHC $\sqrt{s} = 13$ TeV Proton-Proton Collisions

by

Anni Xiong

A dissertation accepted and approved in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in Physics

Dissertation Committee:

Eric Torrence, Chair

Stephanie Majewski, Advisor

Tim Cohen, Core Member

Josef Dufek (Earth Science), Institutional Representative

University of Oregon

Summer 2025

DISSERTATION ABSTRACT

Anni Xiong

Doctor of Philosophy in Physics

Title: Search for Charged Higgs through a Reinterpretation of a Dark Meson search with the ATLAS experiment based on 140 fb^{-1} of LHC $\sqrt{s} = 13$ TeV Proton-Proton Collisions

This thesis presents a search for single charged Higgs production through a reinterpretation of an existing Dark Meson search. The data used in the analysis is recorded by the ATLAS experiment at the Large Hadron Collider (LHC) during Run 2 with an integrated luminosity of 140 fb^{-1}. The LHC is a particle accelerator located in Geneva, Switzerland that collides two beams of protons. ATLAS is one of the four detectors that measures collisions delivered by the LHC.

The discovery of the neutral Higgs Boson of mass 125 GeV by the ATLAS and CMS experiments in 2012 prompted discussions on whether extensions of the Standard Model (SM) scalar sector exist. The two-Higgs-doublet model (2HDM) is one of the simplest extensions to the SM that predicts the existence of charged Higgs.

A reinterpretation is an analysis method commonly used in High energy Physics to speed up the search process, when certain conditions apply. Different theory models often share similar experimental signatures for signals as well as background processes. It is therefore possible to recycle a previously established analysis for a new signal. In this thesis, the search for dark meson is used to search for a different hypothetical particle, the charged Higgs predicted in 2HDM.

The search for dark meson, which has previously been published through ATLAS collaboration, studied decays of dark mesons to top and bottom quarks.

These top and bottom quarks subsequently decay to all jets (fully-hadronic) and (exactly) 1-lepton final states. The similarities of decay products between the dark mesons and the charged Higgs makes the dark meson search potentially sensitive to the charged Higgs signal.

Additionally, this thesis presents a project concerning the hardware electronics of the ATLAS experiment, regarding the upgrade to the Global Trigger system for High-Luminosity Large Hadron Collider (HL-LHC). HL-LHC is the next generation LHC that will deliver much more intensive proton collisions. A software simulation program for the tau trigger was upgraded to accept full-granularity data from the calorimeter and muon systems.

This dissertation includes previously published coauthored material.

CURRICULUM VITAE

NAME OF AUTHOR: Anni Xiong

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR, USA
University of Washington, Seattle, WA, USA

DEGREES AWARDED:

Bachelor of Science, Physics, 2018, University of Washington, Seattle
Doctor of Philosophy, Physics, 2025, University of Oregon, Eugene

AREAS OF SPECIAL INTEREST:

Experimental High Energy Particle Physics

PROFESSIONAL EXPERIENCE:

Graduate Research Assistant, University of Oregon, ATLAS Collaboration,
2021 – Present
Graduate Teaching Assistant, University of Oregon, Department of Physics,
2019 – 2021

PUBLICATIONS:

Search for dark mesons decaying to top and bottom quarks in proton-proton
collisions at $\sqrt{s} = 13$ Tev with the atlas detector. Journal of High Energy Physics,
2024(9),968 September 2024. ISSN 1029-8479. doi: 10.1007/jhep09(2024)005.
http://dx.doi.org/10.1007/JHEP09(2024)005

ACKNOWLEDGEMENTS

To every reader of this dissertation and to my childhood dream of becoming a physicist.

TABLE OF CONTENTS

LIST OF FIGURES

13

14

15

16

17

LIST OF TABLES

26

CHAPTER I

INTRODUCTION

The idea that matter is made of small building blocks can be dated back to 6th century BCE. Ancient Indian philosophers in Janisim advocated the idea that the ajiva (non living part of universe) consists of matter or pudgala, of definite or indefinite shape which is made up of tiny uncountable and invisible particles called permanu. Ancient Greek Philosophers such as Democritus envisioned atomism by proposing that the physical universe is made of fundamental indivisible components known as atoms.

Thousands of years later, the development of Quantum Field Theory (QFT) provided mathematical structure for the subatomic particles and allowed predictions of new particles and their properties. Details of modern theories of particle Physics is included in Chapter II. Modern technologies in particle detection allowed elementary particles to be observed and measured to high precision. Modern radiation/ particle detection is a very broad topic, Chapter III of this thesis will focus on technologies and experimental setup of the ATLAS experiment and the Large Hadron collider at The European Organization for Nuclear Research (CERN). Chapter IV presents a qualification task regarding the trigger upgrade of the ATLAS Data Acquisition system.

Chapter V presents a dark meson search. It is a previously published co-authored paper as supplementary material for the reinterpretation. The paper can be found at this link. The credit of this chapter belongs to the ATLAS Collaboration. Finally Chapter VI presents the a search for charged Higgs through a reinterpretation of the dark meson search.

CHAPTER II

THEORETICAL MOTIVATION

This chapter serves as the theoretical motivation for the relevant physics analyses in the thesis. The chapter includes an introduction to the Standard Model (SM) of particle physics in Section 2.1 as well as a Beyond Standard Model theories: two Higgs-doublet model (2HDM) in Section 2.4 Branco et al. [2012]. These theories are closely related to the Physics data analysis that will be discussed in further detail in Chapter VI.

2.1 The Standard Model

The Standard Model (SM) of particle Physics is one manifestation of Quantum Field Theory; it attempts to explain and make predictions of all the phenomena of elementary particles in terms of their properties and interaction. As shown in Figure 2.1, there are 3 types of particles categorized by their spin: spin $\frac{1}{2}$ fermions, called leptons and quarks, spin-1 bosons, called the gauge bosons which mediate the forces and at least one spin-0 particle called Higgs Boson that gives mass to other elementary particles.

Single quarks cannot be observed in nature, but the quarks come together by the strong force and form bound states; these bound states are called hadrons and can be observed. Hadrons can be further classified into baryons and mesons. Baryons are made of three quarks and mesons are made of two quarks. Common examples for baryons are protons p and neutrons n; that of mesons are the triplet of pions (π^-, π^0, π^+).

For every charged particle, whether it is elementary or composite, there exists a corresponding particle with the same mass but opposite charge called the anti-particle.
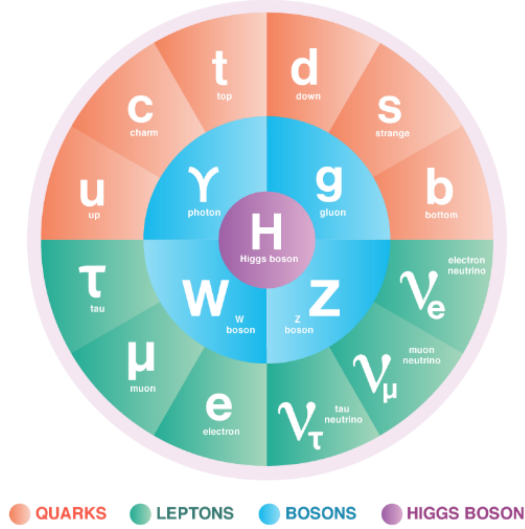
Figure 2.1. Standard model particles organized in different categories: the outermost are quarks and leptons, the middle layer includes the force carriers and the center one is Higgs Boson.

2.1.1 leptons and the weak interactions

There are 6 leptons that appear in 3 pairs, each known as a generation. They are often written in the following form:

$$\begin{pmatrix} e^- \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu^- \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau^- \\ \nu_\tau \end{pmatrix}$$

with their corresponding anti-particles

$$\begin{pmatrix} e^+ \\ \overline{\nu}_e \end{pmatrix}, \begin{pmatrix} \mu^+ \\ \overline{\nu}_\mu \end{pmatrix}, \begin{pmatrix} \tau^+ \\ \overline{\nu}_\tau \end{pmatrix}$$

In the order of increasing mass, there are 3 charged leptons (e^-, μ^-, τ^-) and associated with them, 3 neutron leptons called the neutrinos with close to zero

(a) $e^- \rightarrow e^- + \gamma$ (b) $e^- + \gamma \rightarrow e^-$
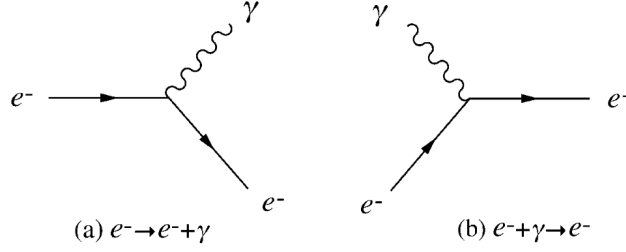
Figure 2.2. Feynman vertex of the most basic Electromagnetic interaction through the emission and absorption of a photon B.R.Martin.

mass [1]. The charged leptons interact through both electromagnetic force and weak force while the neutral leptons are observed to only interact through the weak force. Electromagnetic interactions can be understood as an emission or absorption of a photon demonstrated by the process $e^- \rightarrow e^- + \gamma$ and $\gamma + e^- \rightarrow e^-$ shown by the Feynman vertex 2.2. Electron-electron scattering $e^- + e^- \rightarrow e^- + e^-$ through the exchange of a photon is can be understood by piecing together the two basic processes.

Each lepton has a quantum number associated with it called "lepton number" (usually written as L). In general, each lepton gets a +1 and each anti-particle gets a -1. More specifically, Lepton number in each generation is represented as in Equation 2.1 through 2.3

$$L_e = N(e^-) - N(e^+) + N(\nu_e) - N(\bar{\nu}_e) \tag{2.1}$$

$$L_\mu = N(\mu^-) - N(\mu^+) + N(\nu_\mu) - N(\bar{\nu}_\mu) \tag{2.2}$$

$$L_\tau = N(\tau^-) - N(\tau^+) + N(\nu_\tau) - N(\bar{\nu}_\tau) \tag{2.3}$$

---

[1]SM technically predicts zero neutrino mass but the measurement suggests otherwise; this remains an open question in particle physics
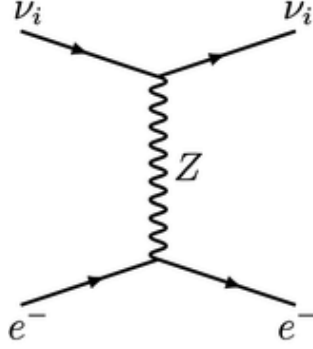
Figure 2.3. Feynman diagram of elastic $\nu_\mu e^-$ scattering B.R.Martin.

In SM, lepton number in each generation is individually conserved in all the interactions.

Similar to electro magnetic processes, the weak interaction involving only leptons is mediated by W and Z bosons $W^\pm$ and Z^0 being emitted by one lepton and absorbed by another. One simple example is the $\nu_\mu e^-$ scattering process shown in Fig 2.3.

So far, the experimental data has been consistent with the assumption that the possible interactions of the lepton and its neutrino are identical among all the leptons when the mass differences are taken into account. This assumption is called the **lepton universality**. One simple example is the similar decay mode for muon and the tau lepton as shown in Equation 2.4 and 2.5.

$$\mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu \tag{2.4}$$

$$\tau^- \rightarrow e^- + \overline{\nu}_e + \nu_\tau \tag{2.5}$$

2.1.2 Quarks and the strong force

Quarks and their bound states (hadrons) mostly interact by strong force. The QCD (quantum Chromodynamics) is the theory of strong interactions. Analogous to electric charge, quarks also have another property called "color charge". 3 quarks were originally proposed to exist, but 6 quarks are now known to exist. Like the leptons, they can be written in 3 pairs, denoted:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}$$

Each generation has a quark with charge $+\frac{2}{3}$: the up (u), charm (c) and top (t) quark; and a quark with charge $-\frac{1}{3}$: the down (d), strange (s) and bottom (b) quark. The corresponding anti-particles with opposite charges are as shown in 2.1.2

$$\begin{pmatrix} \bar{u} \\ \bar{d} \end{pmatrix}, \begin{pmatrix} \bar{c} \\ \bar{s} \end{pmatrix}, \begin{pmatrix} \bar{t} \\ \bar{b} \end{pmatrix}$$

Because isolated free quarks have never been observed, the mass of the most quarks are inferred indirectly from the observed masses of their bound state. The masses and other properties are summarized in Figure 2.4. The heaviest quark, the top quark, has mass close to that of a gold atom !

In QCD, each quark carries also a color charge and the fundamental process $q \rightarrow q + g$ is mediated by spin-1 gluons.

Although no free quarks are observed, more than 200 hadrons (including the anti-particles) have been discovered through decades of study. Figure 2.5 gives some examples; baryons have half integer spin and are assumed to be bound states of 3 quarks (3q); antibaryons are bound states of 3 anti-quarks and the mesons are bound

Name	Symbol	Mass	Q	B	S	C	$\tilde{B}$	T
Down	d	$m_d \approx 0.3$	$-1/3$	$1/3$	0	0	0	0
Up	u	$m_u \approx m_d$	$2/3$	$1/3$	0	0	0	0
Strange	s	$m_s \approx 0.5$	$-1/3$	$1/3$	-1	0	0	0
Charmed	c	$m_c \approx 1.5$	$2/3$	$1/3$	0	1	0	0
Bottom	b	$m_b \approx 4.5$	$-1/3$	$1/3$	0	0	-1	0
Top	t	$m_t \approx 174$	$2/3$	$1/3$	0	0	0	1

Figure 2.4. The approximate masses of quarks in GeV and their electric charge in e; baryon number B, strangeness S, charm C, bottom $\tilde{B}$ and top T are also shown in Ref.B.R.Martin.

Particle		Mass (MeV/c^2)	Q	S	C	$\tilde{B}$
p	uud	938	1	0	0	0
n	udd	940	0	0	0	0
Λ	uds	1116	0	-1	0	0
Λ_c	udc	2285	1	0	1	0
Λ_b	udb	5624	0	0	0	-1

Figure 2.5. some examples of baryons with their quark compositions, electric charge Q, strangeness S, charm C and bottom $\overline{B}$ number B.R.Martin.

states of a quark and an anti-quark($q\bar{q}$). In strong and electromagnetic interactions, the quark number of each flavor is separately conserved.

The lightest known mesons are the triplet of pions (π^+, π^0, π^-). They are copiously produced in many hadronic interaction. And they are some of the most commonly detected particles from proton proton collisions produced by the LHC. Some examples of mesons are shown in Table 2.6. Analogous triplet of pions can exist in the dark sector of a beyond standard model theory. Later in the chapter, a triplet of "dark" pions in the theory behind dark meson will be discussed.

Hadrons have typical effective radii r of order 1 fm, with an associated time scale r/c of order 10^{-23} s. The vast majority are highly unstable and decay to lighter hadrons by the strong interaction with lifetimes of this order.

34

Particle		Mass (MeV/c^2)	Q	S	C	$\tilde{B}$
π^+	$u\bar{d}$	140	1	0	0	0
K^-	$s\bar{u}$	494	−1	−1	0	0
D^-	$d\bar{c}$	1869	−1	0	−1	0
D_s^+	$c\bar{s}$	1969	1	1	1	0
B^-	$b\bar{u}$	5279	−1	0	0	−1
Υ	$b\bar{b}$	9460	0	0	0	0

Figure 2.6. some examples of mesons with their quark compositions, electric charge Q, strangeness S, charm C and bottom $\overline{B}$ numberB.R.Martin.

2.1.3   Higgs boson

The Higgs Boson is an essential part of the SM. W, Z bosons as well as the fundamental fermions acquire mass through the Higgs mechanism. Without it, the SM is not a consistent theory M.Thomson [2013]. The coupling between the fermion fields and the Higgs field is referred to as the **Yukawa coupling**.

The standard model has the minimal Higgs model which consists of two complex scalar fields:

$$H = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.6}$$

The Lagrangian for this doublet of complex scalar field is

$$L = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi) \tag{2.7}$$

The Higgs potential is

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \tag{2.8}$$
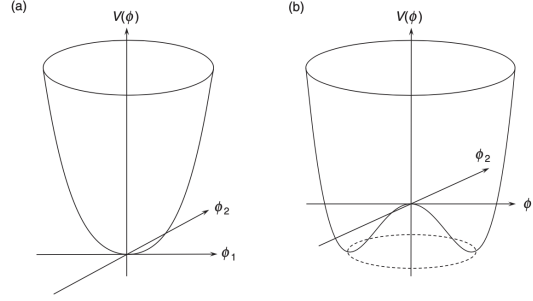
Figure 2.7. The $V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$ potential for a complex scalar field for (left) $\mu^2 > 0$ and (right) $\mu^2 < 0$ M.Thomson [2013].

The Higgs potential is visually plotted in 2.7. When $\mu^2 > 0$, the potential has only one minimum at 0 VEV (vacuum expectation value), while when $\mu^2 < 0$, the potential has infinite set of minima at non zero VEV value, therefore "breaking the symmetry". In SM, the Higgs mechanism is developed by breaking the $SU(2)_L \times U(1)_Y$ local gauge symmetry of the electroweak sector of the Standard Model. And using the covariant derivative, masses of the photon, Z boson and two W bosons can be derived by calculating eigenvalues of the mass matrix.

The Higgs boson was discovered in 2012 by ATLAS Aad and CMS experiment Cha [2012]. Results from multiple channels were combined to make the discovery: H$\rightarrow \gamma\gamma$, H$\rightarrow ZZ \rightarrow 4l$, H$\rightarrow WW \rightarrow e\mu\nu\mu$ in 8 TeV data as well as previous results. Clear evidence of a neutral boson with mass 126.0 $\pm$ 0.4 (stat) $\pm$ 0.4 (sys) GeV Aad was observed. Since then the discovery prompted further discussions on whether extensions of the SM scalar sector exist, other than the observed SM doublet. Additional Higgs doublet(s) can be added as simple extensions to the SM, two Higgs doublet model is one such theory and it will be discussed later in the chapter.

2.2 Dark matter and the theory beyond standard model

So far, SM is the most successful theory tested with extremely high precision, but it is not the final theory of particle Physics. Several phenomena remain unexplained, such as the nature of dark matter, dark energy, gravity, matter anti-matter asymmetry and the neutrino mass. A large number of Beyond Standard Model (BSM) theories have been proposed to cover these issues in the SM.

The nature of dark matter is one of the most important research topics in particle physics, the evidence for dark matter is included in this section to motivate the theory behind the dark mesons and many other BSM theories.

2.3 Evidence of dark matter

The nature of dark matter is one of the hardest problems in modern particle physics. Here includes some main astronomical observations that provide strong evidence for its existence.

Galaxy rotation curve. : A spiral galaxy contains about 2/3 of its mass (M) in the galactic core Persic et al. [1996] Salucci et al. [2007], assuming the galaxy having a spherically symmetric density. Classical mechanics predicts that the velocity of objects moving outside of the core should fall like $\propto \sqrt{M/r}$ where r is the distance away from the core. This is the so-called "Keplerian fall". However, the experimental measurement tells a different story. In 1980, Vera Rubin and her colleague measured rotational speed of spiral galaxies versus distance from the galactic core (rotation curve) in Ref. Rubin et al. [1980]. The measurement was sampled over a large amount of galaxy radii, masses and luminosities. Observation showed the galaxies move consistently faster than what is expected from classical mechanics. Instead
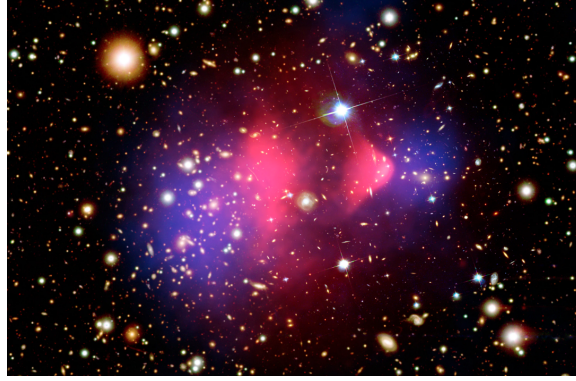
Figure 2.8. The "bullet cluster" 1E0657-56 ;
Credit: X-ray:NASA/CXC/CfA/M.Markevitch et al.; Optical:NASA/STScI; Magellan/U.Arizona/D.Clowe et al.; Lensing Map:NASA/STScI; ESO WFI; Magellan/U.Arizona/D.Clowe et al.

of the Keplerian fall, most rotation curves rise slowly even at the farthest measure point. These findings suggested (assuming Newtonian Mechanics holds) large amount of invisible matter is exist in these galaxies.

Observation of non-luminous matter through gravitational lensing effect. Another type of evidence for dark matter comes was discovered through the gravitational lensing effect such as the bullet clusters Clowe et al. [2006]. Two large clusters undergo violent collisions when they move toward each other as shown in Fig 2.8. This is an image of 1E 0657-56 taken by Chandra X-ray observatory, where the two pink clumps in the image contain most of the visible baryonic matter. The concentration of mass was measured and the blue area marked where most mass is in the cluster. Large amounts of invisible matter can be identified using gravitational lensing effect (light from distant objects is distorted when passing by massive objects). The clear separation of baryonic matter from the the invisible matter gave a direct evidence of dark matter. There are also other similar evidences which come from the gravitational lensing effect Massey et al. [2010].

Cosmic Microwave Background (CMB) and Large Scale structure of the universe.
Other compelling evidence for dark matter is astrophysical measurements related
to the large- scale structure of the Universe. On scales much larger than the size of
galaxy, galaxies are not uniformly distributed in space, but displays coherent structure
with galaxies residing in groups which lie at the interactions of long filamentsCoil
[2013]. Vast regions are empty space exist in between. CMB is the faint radiation
exists in the empty space in between stars and galaxies and is strongest in the
microwave region of the EM spectrum. It is widely believed to come from the early
universe and a strong evidence for the Big Bang theory. The CMB provides the most
precise measurement for the abundance of dark matter. Ground and space based
experiment such as Cosmic Microwave Background Explorer (COBE) Efstathiou
et al. [1992] have been trying to understand the formation of the early universe
by precise measurement of the anisotropy in space. These measurements provided
firm experimental basis for the ΛCDM model which has became the standard model
of cosmology. Within the ΛCDM model, only 5% of the energy-matter density of
the normal universe is baryonic matter. A further 23% is in the form of cold (non-
relativistic) dark matter (CDM) M.Thomson [2013]

From the above observations, a few properties can be deduced for dark matter.
Dark Matter should have mass due to its gravitational effect; it is stable and does not
decay to other normal matter; and it does not have electromagnetic interactions and
maybe only interact weakly, so it does not emit or absorb light like known matter.
These properties make it very challenging to be detected.

2.4 Charged Higgs in two-Higgs-doublet-model (2HDM)

Charged Higgs bosons (often denoted as H+) could appear when a 2nd Higgs doublet or triplet is added as extensions to the scalar sector of the SM Branco et al. [2012]. The 2HDM model studied in this thesis is one of the simplest extension to the SM that adds exactly one Higgs doublet. A total of five Higgs bosons are predicted by 2HDM: three neutral Higgs bosons and a pair of charged Higgs bosons.

Each of the two Higgs doublets in 2HDM has a (VEV) ν_1 and ν_2 and one most important parameter is $\tan\beta$. It is the ratio between the two VEVs.

$$tan\beta = \frac{\nu_2}{\nu_1} \tag{2.9}$$

There are various theoretical motivations for 2HDM that make charged Higgs worth searching for. 2HDM is the simplest extension to the SM and provides extra sources of CP-violation that is important to answer the matter-antimatter asymmetry problem. Another important motivation is SUSY. In SUSY, additional Higgs doublet is needed to give mass simultaneously to the charge $2/3$ and charge $-1/3$ quarks. Thus two Higgs doublet models can be considered as part of Minimal Supersymmetric Standard Model (although in general, 2HDM models can exist outside of SUSY). 2HDM also provides extra source of CP-violation

$$\mathcal{L}_{H^\pm} = -H^+ \left(\frac{\sqrt{2}\,V_{ud}}{v} \bar{u} \left(m_u X P_L + m_d Y P_R \right) d + \frac{\sqrt{2}\,m_\ell}{v} Z \bar{\nu}_L \ell_R \right) + \text{H.c.}$$

Figure 2.9. General form of Yukawa couplings Branco et al. [2012].

There are various types of 2HDM models depending on the Yukawa couplings to fermions and Fig 2.9 give the most general form of Yukawa couplings. V_{ud} is the

	Type I	Type II	Lepton-specific	Flipped
X	$\cot\beta$	$\cot\beta$	$\cot\beta$	$\cot\beta$
Y	$\cot\beta$	$-\tan\beta$	$\cot\beta$	$-\tan\beta$
Z	$\cot\beta$	$-\tan\beta$	$-\tan\beta$	$\cot\beta$

Figure 2.10. The parameters X, Y and Z in 2.9 for the four models without flavour-changing neutral currents Branco et al. [2012].

element of the CKM matrix, corresponding to the up and down quarks. The values of X, Y, Z are defined in Fig 2.10 which shows how each type of the 2HDM is different from one another. Ref.Branco et al. [2012] gave a comprehensive introduction to a whole suite of 2HDM models. The one analyzed in the reinterpretation is the type II model which is MSSM(Mimium super-symmetric model) like. In this case, the coupling of the charged Higgs to the top and bottom quarks is governed by either the bottom quark mass times $\tan\beta$, which may be large, or by the top-quark mass times $\cot\beta$. As a result, one expects the presence of charged Higgs to influence the b-quark decays. This makes the B physics an important stage to search for charged Higgs indirectly. Charged Higgs can be searched indirectly through B hadrons where it appears virtually in the loop. A very strong bound on the mass of the charged Higgs come from studies of B decays.

The re-interpretation in this thesis focused on the direct production of $H^\pm$. To explore the potential of searches at the LHC, the branching ratio is an important guide. In the branching ratio plot Fig 2.11 for the case of $\tan\beta = 1$, one can see that for lower mass of $H^\pm$, the decay into $\tau^+\nu_\tau$ dominates; while at the mass range larger than the top quark mass, the decays into tb are completely dominant. This is why the $H \to tb$ is the main channel to search for heavy charged Higgs.

The charged Higgs production cross section increases as $\tan\beta$ decreases, so it is easier to extract exclusion limit with samples of smaller $\tan\beta$ to achieve larger signal
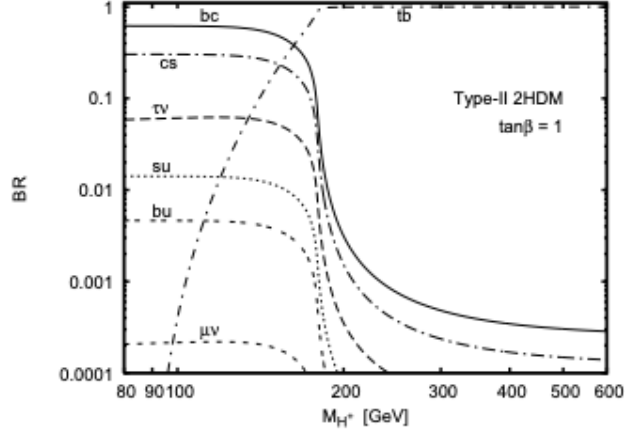
Figure 2.11. The branching ratios of the charged Higgs in the type II 2HDM Logan and MacLennan [2010].

yield. However, when $\tan \beta < 1$, the top Yukawa coupling becomes so large that the theory is no longer perturbative MARTIN [1998]. Therefore, the data analysis in chapter VI focused on models with $\tan \beta \geq 1$.

2.5 The theory behind the dark meson search

The theory behind dark mesons is a Beyond Standard Model theory that could provide explanation to the origin of dark matter. There is a range of similar models described in Ref.Kribs and Neil [2016]. This introduction will focus only on the phenomenology that is relevant to the dark meson search in Chapter V, published in 2024 in Ref. dar [2024]. The published paper is also included in this thesis, see the introduction in Chapter V for more details about the theory.

Dark Mesons could emerge by extending the SM with a new strongly interacting dark sector. The full phenomenology of dark mesons that could be sought at the LHC is described in Ref. Kribs et al. [2019b]. This type of extensions can generate a stable dark matter candidate as a heavy baryon. The middle part of Figure 2.12
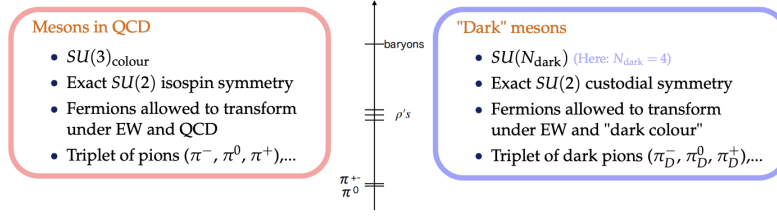
Figure 2.12. Comparison of dark meson and standard model mesons; the middle part includes a spectrum of dark sector particle masses. Credit: Jochen Jens Heinrich.

ranks the masses of dark sector particles showing the ρ particle as the dark matter candidate.

The new sector is implemented with a SU(2) group structure that is analogous to the (Quantum Chromo-dynamics) QCD meson and baryon sector in the SM (the so called custodial symmetry). Figure 2.12 shows how the theory of dark sector compares and contrasts with the standard model QCD. Interactions between the Higgs boson and the dark sector breaks the dark sector global symmetry and therefore allows the dark mesons to decay back to pure SM particles.

In Chapter V, the two relevant sets of dark mesons are a lighter pseudo-scalar triplet of dark pions, π_D and an additional triplet of dark rho vector mesons, ρ_D. The triplets are assumed to have the same mass and the dark sector can be described by 3 parameters: the mass of the dark pions m_{π_D}, the mass of the dark rhos m_{ρ_D} and the number of dark colors N_D. N_D was fixed to $N_D = 4$ in the search because the N_D does not have big influence on the phenomenology when N_D is not excessively large. Once $N_D = 4$ is fixed, the dark pion production cross section depends trivially on the ratio between the dark pion mass and the dark rho mass. This is denoted as η and will appear frequently when introducing the dark meson search in chapter V.

$$\eta = \frac{m_{\pi D}}{m_{\rho D}} \tag{2.10}$$

The dark meson search only considers models with $\eta < 0.5$ where $\rho_D^{\pm,0} \to \pi_D^\pm \pi_D^{0,\mp}$ has a branching fraction of nearly 1.0, while for models with $\eta > 0.5$, the production mode shown in 2.13 is kinetically forbidden. Another reason that make $\eta < 0.5$ model more valuable to study is the much weaker exclusion given by previous studies.

The dark sector allows two distinct models of kinetic mixing with the SM, $SU(2)_L$ and $SU(2)_R$ which corresponds to mixing with either W or B(meson) field. ρ_D could be produced through either type of mixing. The $SU(2)_L$ models give considerably larger cross sections than the $SU(2)_R$ models. Therefore the search focused on $SU(2)_L$ models. This is why all analysis figures of the dark meson search in Chapter V carry $SU(2)_L$ labels.

As shown on Fig 2.13, dark pions are either pair-produced via Drell-Yan type processes in 2.13(b) or resonantly produced by a kinetic mixing of SM electroweak gauge bosons and a dark rho particle ρ_D shown in 2.13(a). The resonant production is the main production mode in most parameter space of the search and 2.13(c) illustrates a pair of resonantly produced dark pion decay into top and bottom quarks. This process is the most typical event analyzed in the search of dark mesons at Chapter V.
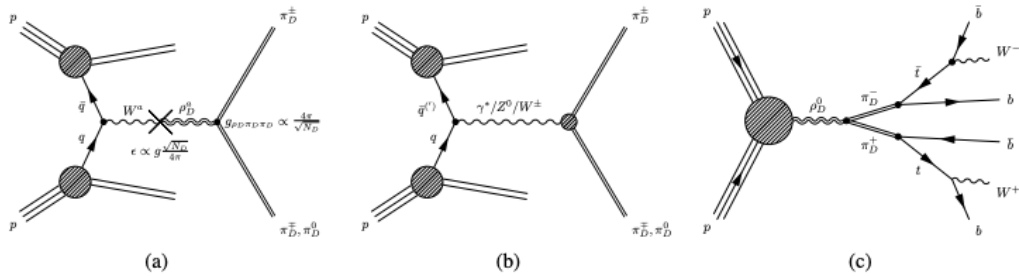
Figure 2.13. Leading order dark pion production dar [2024].

CHAPTER III

THE EXPERIMENTAL SET UP

The data analysis in this thesis was collected by the ATLAS experiment from proton proton collisions produced by the Large Hadron Collider(LHC) during the Run 2 data taking period. The following section gives a brief introduction to the LHC, the ATLAS experiment and its sub-detectors. The tau trigger development in Chapter IV is developed for the environment of the High Luminosity LHC (HL-LHC), so a short introduction to the Trigger and Data Acquisition (TDAQ) system in HL-LHC is provided on section 3.2.62.

3.1 THE LARGE HADRON COLLIDER

The LHC, located on the border between France and Switzerland, is the world's biggest and most powerful particle accelerator Evans and Bryant [2008]. It is housed in the 26.7 km Large Electron Positron (LEP) collider tunnel. The LHC ring consists of super conducting magnets that extends 27km. It delivers bunches of protons every 25ns (at the frequency of 40MHz) and collides them at near the speed of light. During Run 2, every bunch is packed with 10^{11} protons and is accelerated up to the energy $\sqrt{s} = 6.5$ TeV with the final collision energy reaching $\sqrt{s} = 13$ TeV at collision points.

The proton beams go through multiple stages of acceleration before finally entering the LHC tunnel De Melis [2016]. As shown on Figure 3.1, the journey of protons starts from a bottle of hydrogen gas at LINAC2 where protons are extracted by stripping electrons from the hydrogen molecules. LINAC2 accelerates protons to an energy of 50 MeV. Then the proton beams are injected to Proton Synchrotron Booster (PSB), which accelerates the protons to 1.4 GeV. The Proton Synchrotron (PS) boosts the

beam to 25 GeV and right before entering the LHC, the Super Proton Synchrotron (SPS) accelerates the beams to 450 GeV.

After the beam is injected into the LHC ring, the (Radio Frequency) RF cavities housed inside the LHC ring bring the beam energy from 450 GeV to 6.5 TeVEvans and Bryant [2008]. The process can happen in only 20 min, after the beams pass through the RF cavities 10 million times. RF cavities are metallic chambers working in a superconducting state. They contain a longitudinal oscillating voltage applied across an isolated gap in the vacuum chamber. Each proton receives a forward push when going through a cavity and gains 2 MeVof energy. There are 8 RF cavities for each of the two beams. Therefore, each beam's energy is boosted by 16 MeVafter going through all the cavities once in the LHC ring. Other than accelerating the protons, the RF cavities keep the protons tightly packed in **bunches**, instead of having protons evenly distributed along the LHC ring. The compact bunches increase the chance of head-on collisions and ensures high luminosity when the bunches of protons collide.

Various types of magnets are used at the LHC to direct and focus the beams to obtain the best quality of proton collisions. Dipole magnets are used to bend the beams and quadrupole magnets are used to focus the beams. Just prior to the collision, quadrupole magnets "squeeze" the particles closer to increase the chances of interactions.

Inside the tunnel, two beams of protons are made to collide at 4 points along the LHC ring. These 4 points are where the main particle detectors are located: ATLAS, CMS, LHCb and ALICE. Their locations relative to the LHC ring are shown on Figure 3.1. ATLAS and CMS are general-purpose detectors that explore a wide range of physics. ALICE is dedicated to study heavy ion physics while the LHCb

investigates imparity between matter and antimatter through measuring bottom quarks.
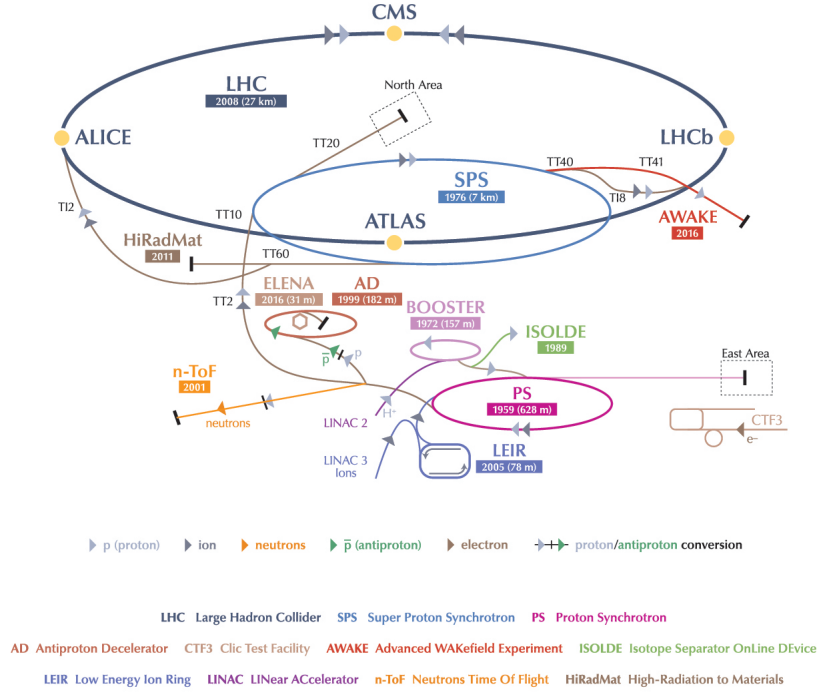
Figure 3.1. The CERN accelerator complex De Melis [2016].

Luminosity $\mathcal{L}$ is the most important parameter for a particle collider. It measures the rate of collisions. Instantaneous luminosity is the number of collisions per unit area per second. The integrated luminosity is the time integral of the instantaneous luminosity; it is used to express the amount of data collected by LHC detectors. With cross section σ, the number of collisions per unit time can be expressed as

$$\frac{dN_{collision}}{dt} = \sigma \mathcal{L} \tag{3.1}$$
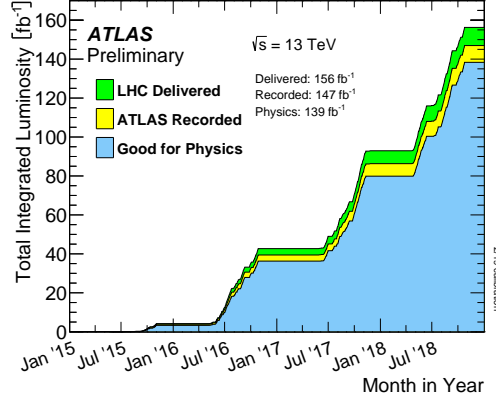
Figure 3.2. Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at 13 TeV centre-of-mass energy in 2015-2018 (Image: ATLAS Collaboration/CERN).

The design luminosity of the LHC is $\mathcal{L} = 10^{34}$ cm^{-1}s^{-1}. It means 10^{34} collisions are produced by the LHC collisions per unit area per second. During Run 2 data taking, peak luminosity reached $\mathcal{L} = 2.0 \times 10^{34}$ cm^{-1}s^{-1}. According to 3.2, by the end of Run 2, the LHC has delivered 156 fb^{-1}of data to each experiment and ATLAS recorded 139 fb^{-1}of data certified good for Physics analysis 3.2

After the current data taking period Run 3 and Long Shutdown 3 (LS3), the High Luminosity LHC (HL-LHC) plans to start operation by late 2020's as shown in 3.3. To fully exploit the physics potential, it plans to collide particles at center of mass energy $\sqrt{s} = 14$ TeV; at the same time, the accelerator is expected to reach an instantaneous luminosity of $\mathcal{L} = 7.5 \times 10^{34} cm^{-1}s^{-1}$. During 12 years of operation, up to 3000 fb^{-1} of data maybe be collected. As shown in Fig 3.3, this is almost 10 times the total amount of data collected from the previous 3 runs. Figure 3.4 shows mean number of interactions per bunch crossing (often referred to as μ or pile up interactions) during Run 2. The average number of interactions is 33.7 during the entire period of Run 2 data taking.
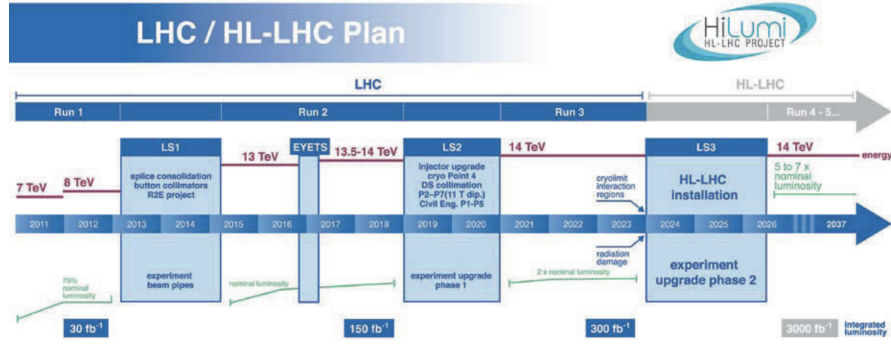
Figure 3.3. LHC baseline plan for the next decade and beyond showing the energy of the collisions (upper red line) and luminosity (lower green lines)G. et al. [2017].
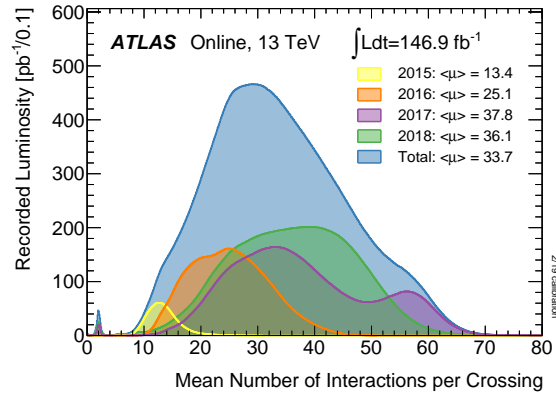
Figure 3.4. Luminosity-weighted distribution of the mean number of interactions per crossing for the 2015 – 2018 pp collision data at 13 TeV centre-of-mass energy. All data recorded by ATLAS during stable beams is shown (Image: ATLAS Collaboration/CERN).

The HL-LHC will achieve much higher collision rate, which largely increases the chance of particle interactions. For example, it expects to produce at least 15 million Higgs bosons per year, compared to around three million collected during LHC operation in 2017. Through the properties of Higgs bosons, the mechanism of electroweak symmetry breaking will be explored in more detail. The higher collision rate will also enhance the probability of discovering rare processes Beyond the Standard Model (BSM). A much higher luminosity results in 200 inelastic proton-proton collisions per beam crossing ($\mu = 200$). This is about 4 times higher than previous runs. Chapter IV of this thesis will present a study on tau trigger simulation as part of this upgrade.

3.2 The ATLAS detector

ATLAS (A Toroidal LHC Apparatus) experiment is a general purpose particle detector designed to study a wide range of particle interactions Bianchi and Collaboration [2024]. Shown in Fig 3.5, it is nearly a 4π detector shaped like a cylinder positioned along the beam line that completely surrounds the collision point to capture as many debris from the collision as possible.

Various sub-detectors are placed layer by layer like an onion, surrounding the beam line. The silicon tracker is placed at the innermost layer, closest to the collision point. It is surrounded by a thin superconducting solenoid providing a 2 Telsa axial magnetic field, followed by electromagnetic and hadronic calorimeters. Finally, a muon spectrometer is placed at the outer most layer of the detector. The following sections will include a more in-depth discussion about each sub-detector.
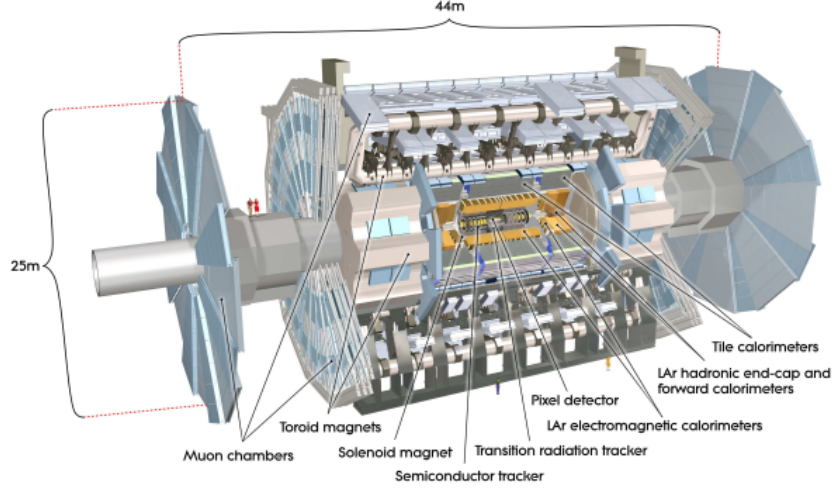
Figure 3.5. The cut-away view of the configuration of the ATLAS detector indicating the locations of the larger detector sub-systems Bianchi and Collaboration [2024].
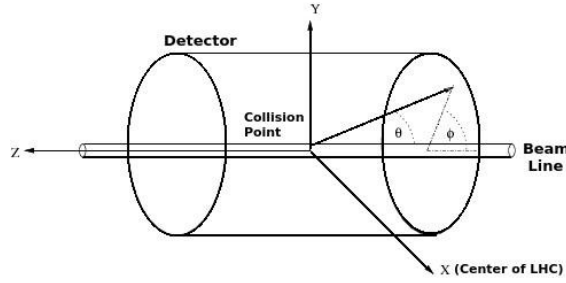
Figure 3.6. Coordinate system used by the ATLAS and CMS experiments at the LHC.

3.2.1   Coordinate system used in ATLAS analyses

ATLAS uses a right-handed coordinate system with its origin at the interaction point (IP) at the center of the detector and the $z-$axis along the beam pipe. The $x-$axis points from the IP to the center of the LHC ring, and the $y-$axis points upwards. In the plane transverse to the beam pipe, polar coordinate is used with (r, ϕ) where the ϕ is the azimuthal angle around the beam pipe. The coordinate system is illustrated in Fig 3.6.

In special relativity, rapidity is a quantity converted from classical velocity. It is expressed with energy and momentum as $y = \frac{1}{2}\ln(\frac{E+p_z}{E-p_z})$ and is used more in the theoretical context. Pseudorapidity, represented by η, is used to measure the angle of something with respect to the y axis. It is defined by $\eta = -\ln\tan(\frac{\theta}{2})$. It is more conveniently used in experimental setting because it only depends on the polar angle θ. In relativistic limit however, pseudorapidity is approximately equal to rapidity.

If a particle has $\eta = 0$, it is located in the central position of the detector; if a particle has a large η, it is at a more forward region of the detector, closer to the beam pipe. The angular distance between two points in $\eta - \phi$ space is calculated by adding the two coordinates in quadrature. This is expressed as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

The collision partons have initially zero momentum in the transverse plane but the initial momentum in the longitudinal direction is unknown. Therefore, Physics analyses with LHC experiments often use quantities in the transverse plane (i.e the $x - y$ plane). The common observables are transverse momentum (p_{T}) and transverse energy (E_{T}). These quantities will be mentioned often in the later chapters.

3.2.2 Magnets used in the ATLAS experiment

To measure the momentum of charged particles passing through various parts of the ATLAS detector, a specific magnetic field is set up through a combination of solenoid and toroid superconducting magnets Aad [2008a]ten Kate [1999]. Fig 3.7 shows the ATLAS magnet system consisting of 3 main parts listed below:

- a central solenoid (CS)

- a Barrel Toroid (BT)

- and two End-Cap Toroids (ECT)
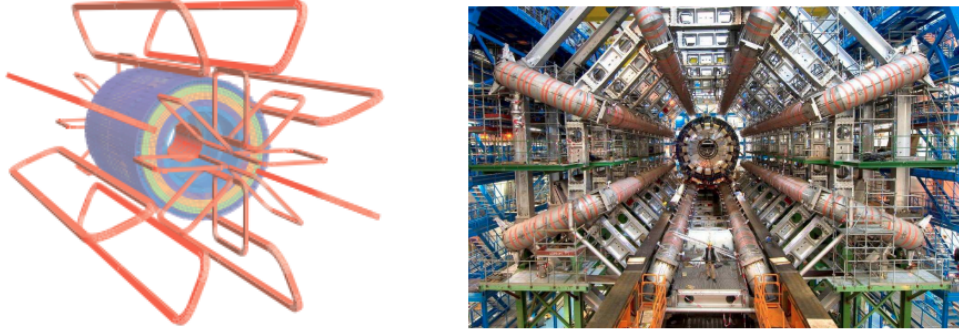
Figure 3.7. ATLAS magnet system, overview (left), installed barrel toroids in the ATLAS underground cavern (right) Aad [2008a].

The central solenoid provides the magnetic field for the inner detector and the other 3 toroids provide the magnetic field for the muon spectrometer both around the outer layer and the forward direction of ATLAS. As shown in Fig 3.7, each toroid consists of 8 coils installed radially around the beam line in a symmetric fashion. The end-cap toroid is rotated by 22.5° with respect to the barrel toroid system to achieve the best bending power at the interface between the two systems

3.2.3 Inner detector

The inner detector (ID) is located at the innermost layer, closest to the beam pipe Aad [2008b] Pernegger [2015]. It was designed to provide excellent momentum resolution and able to measure tracks to high precision for charged particles with p_T down to 0.5 GeV. Fig 3.8 shows a cross sectional plan view of the ID. The ID is enclosed in a cylindrical shell of length ±3512 mm and of radius 1150 mm, which covers the region with pseudorapidity $|\eta| < 2.5$. The high-radiation environment imposes stringent conditions on the inner-detector sensors. At design luminosity, the pixel inner vertexing layer must be replaced after approximately three years of operation.
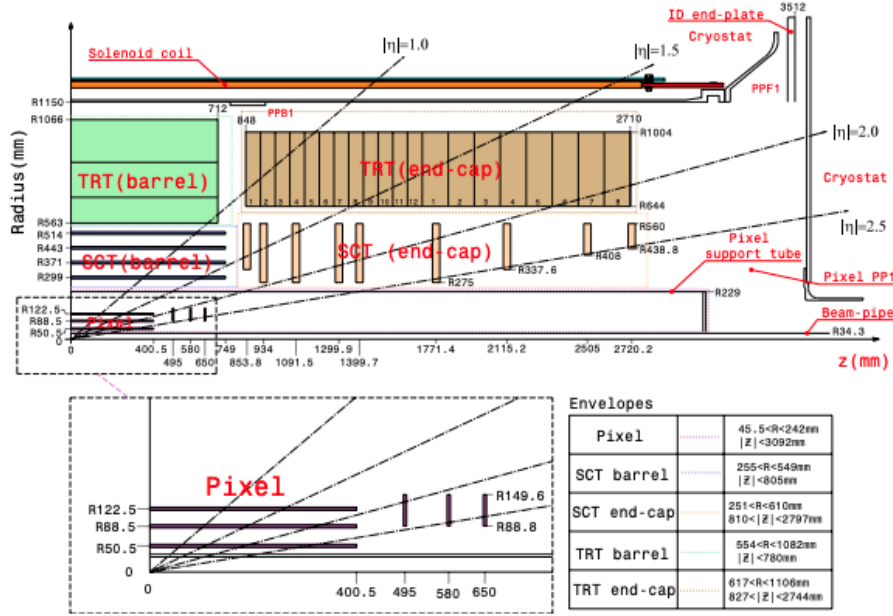
Figure 3.8. Plan view of a quarter-section of the ATLAS inner detector showing each of the major detector elements with its active dimensions and envelopes. The labels PP1, PPB1 and PPF1 indicate the patch-panels for the ID services Aad [2008a].

The ID has 3 separate sub-detectors as shown from Fig 3.10: silicon pixel detector, Silicon micro-strip (SCT) and the transition radiation tracker (TRT). At regions closer to the beam pipe, high-resolution pattern recognition capabilities are available using discrete space-points from pixel layers and SCT layers.

Pixel detector. The pixel detector is only 3.3 cm away from the beam pipe and it is what the particles first interact with. The Pixel Detector is incredibly compact, with over 92 million pixels and almost 2000 detector elements. Fig 3.9 shows the active part of the pixel system which consists of three barrel layers: Layer 0 (so-called b-layer), Layer 1 and Layer 2. Two identical endcap regions, each with three disk layers are also shown.
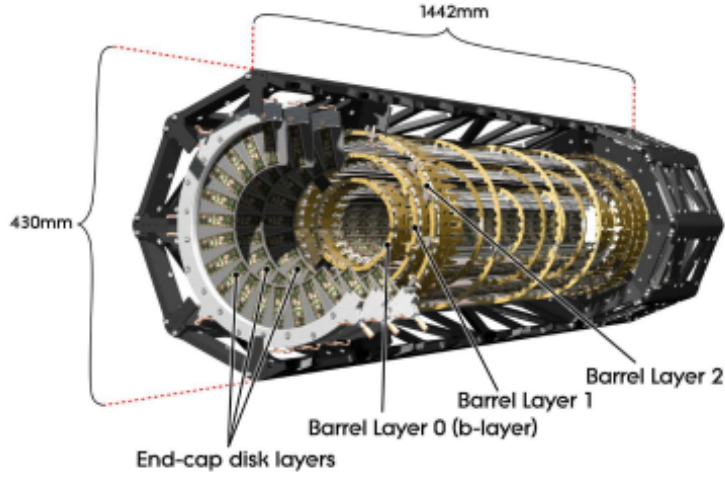
Figure 3.9. A schematic view of the active region of the pixel detector consisting of barrel and end-cap layers Aad [2008b].

Insertable B Layer(IBL). To prepare for the Run 2 data taking, the pixel detector was upgraded from a 3-layer system to a 4-layer structure by adding the IBL during the LHC Long Shutdown 1 ("LS1") Pernegger [2015]. The IBL was added at the very inner most layer of the Pixel detector and is designed to maintain and even improve the performance on tracking and pattern recognition with pile-up $(\mu) > 50.$.

SCT. There are in total 4088 SCT modules covering 4 layers in the barrel region as well as two end-caps region with 9 disk layers Aad [2008a] as shown in Fig 3.8. The SCT modules use silicon strip detector technology that has been commonly used in high energy physics experiments in the past few decades.

TRT. A transition radiation detector uses the radiation emitted by particles traversing through different materials to perform particle identification or tracking. In ATLAS, the TRT is made of polyimide drift tubes ("straws") of 4 mm diameter. Each tube was a 30 μm gold-plated tungsten wire in its center. It consists of many
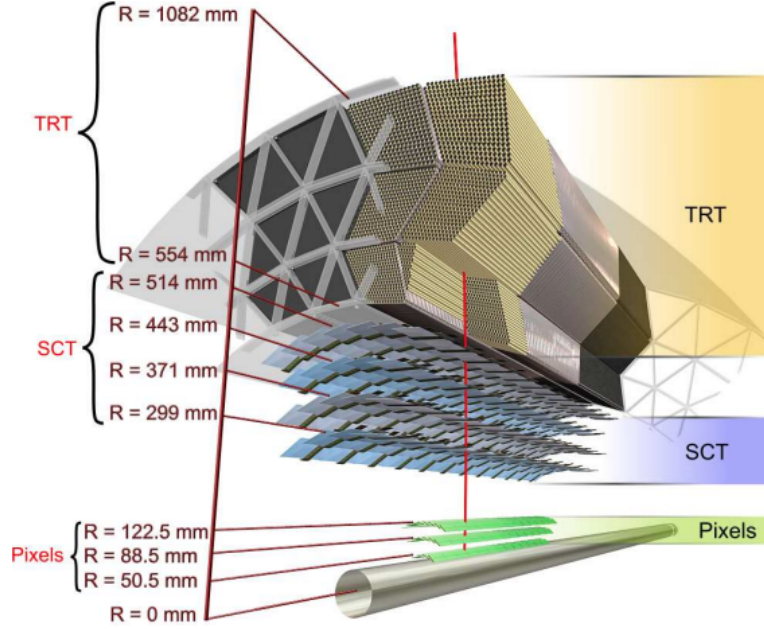
Figure 3.10. Drawing showing the sensors and structural elements traversed by a charged track with 10 GeV p_{T} in the barrel inner detector (η= 0.3) Aad [2008a].

layers of straw tube elements interleaved with transition radiation material. With an average of 36 hits per track, it enhances the tracking performance as well as the momentum resolution in the region $|\eta| < 2.0$.

3.2.4   Calorimetry

Calorimeters in particle physics Fabjan and Gianotti [2003] are used to measure the energy of particles when they interact with a block of material. When a particle enters a detector, secondary particles are produced progressively and form a "shower". Figure 3.14 is a schematic diagram showing how a particle initiates a "shower" in a detection material while losing its initial energy in the cascading process. When the cascaded particles do not have enough energy left to emit new particles, the shower stops. The particle showers are absorbed by the detector and their energy can be
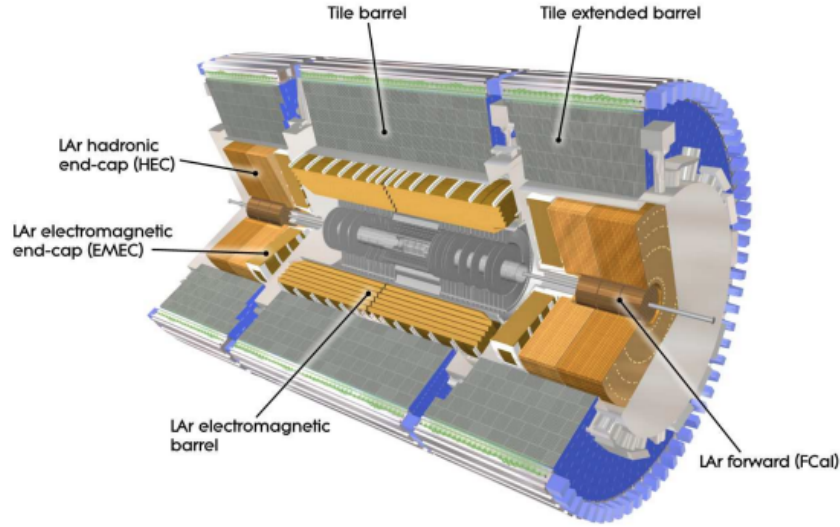
Figure 3.11. Main sections of the ATLAS calorimeter system Aad [2008a].

transformed into a measurable quantity. The energy of the particles can either be absorbed in its entirety or sampled. ATLAS uses sampling calorimeters across its calorimeter system.

The main segments of the ATLAS calorimeter system Aad [2008a] is shown on Fig 3.11, all sections collectively cover the region $|\eta| < 4.9$. Each section uses a different technique that is best suited to detect the physics processes there. The main parameters of the calorimeter system are shown on Table 3.13.

Electromagnetic calorimeters (ECAL) and hadronic calorimeter (HCAL) are two main types of calorimeters used in ATLAS. ECAL is used to measure energies of the electrons, photons through electromagnetic interactions. Measurements from ECAL is also used to reconstruct jets. HCAL is used to measure energies of the hadrons. The two are constructed differently due to different shapes of the shower profile. Figure3.15 presents simulated profiles of electromagnetic and hadronic showers from a 10 GeVelectron and a 10 GeVpion, both traversing through copper. By overlaying
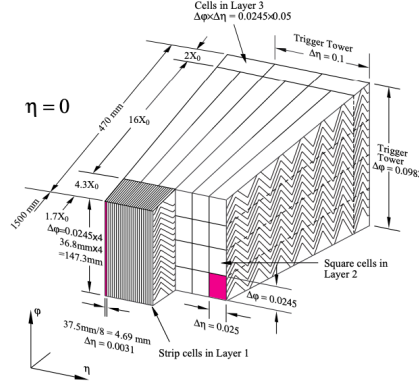
Figure 3.12. Sketch of the LAr electromagnetic barrel region of the calorimeter where different layers with their granularity are clearly shown Aad [2008a].

the two types of showers, one can see that the electromagnetic showers are shorter and narrower than Hadronic showers. Therefore, ECAL tends to be built in a more compact way and much smaller in volume compared to HCAL.

3.2.41 ATLAS Liquid Argon calorimetrs. The Liquid Argon (LAr) calorimeters in ATLAS is a sampling calorimeter which has lead (absorber) and liquid argon(detector) arranged in an accordion shape with copper-tungsten sensors.

EM calorimeter. The EM calorimeter is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components (EMEC) ($1.375 < |\eta| < 3.2$), each in their own cryostat. The EM calorimeter is located closest to the beam pipe and has the finest granularity. It is a lead-LAr detector with accordion-shaped Kapton electrodes and lead absorber plates over its full coverage. The accordion geometry provides complete ϕ symmetry without azimuthal cracks. The barrel region of the EM calorimeter can be further divided into a pre-sampler and 3 parts as shown on Fig 3.12. Each layer has a different granularity. The Tau trigger algorithm developed for the phase II upgrade in chapter IV uses simulated data mainly from this part of the calorimeters.

Table 1.3: Main parameters of the calorimeter system.

	Barrel		End-cap	
EM calorimeter				
Number of layers and $\|\eta\|$ coverage				
Presampler	1	$\|\eta\| < 1.52$	1	$1.5 < \|\eta\| < 1.8$
Calorimeter	3	$\|\eta\| < 1.35$	2	$1.375 < \|\eta\| < 1.5$
	2	$1.35 < \|\eta\| < 1.475$	3	$1.5 < \|\eta\| < 2.5$
			2	$2.5 < \|\eta\| < 3.2$
Granularity $\Delta\eta \times \Delta\phi$ versus $\|\eta\|$				
Presampler	0.025×0.1	$\|\eta\| < 1.52$	0.025×0.1	$1.5 < \|\eta\| < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$\|\eta\| < 1.40$	0.050×0.1	$1.375 < \|\eta\| < 1.425$
	0.025×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.1	$1.425 < \|\eta\| < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \|\eta\| < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \|\eta\| < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \|\eta\| < 2.4$
			0.025×0.1	$2.4 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 2nd layer	0.025×0.025	$\|\eta\| < 1.40$	0.050×0.025	$1.375 < \|\eta\| < 1.425$
	0.075×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.025	$1.425 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 3rd layer	0.050×0.025	$\|\eta\| < 1.35$	0.050×0.025	$1.5 < \|\eta\| < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	
LAr hadronic end-cap				
$\|\eta\|$ coverage			$1.5 < \|\eta\| < 3.2$	
Number of layers			4	
Granularity $\Delta\eta \times \Delta\phi$			0.1×0.1	$1.5 < \|\eta\| < 2.5$
			0.2×0.2	$2.5 < \|\eta\| < 3.2$
Readout channels			5632 (both sides)	
LAr forward calorimeter				
$\|\eta\|$ coverage			$3.1 < \|\eta\| < 4.9$	
Number of layers			3	
Granularity $\Delta x \times \Delta y$ (cm)			FCal1: 3.0×2.6	$3.15 < \|\eta\| < 4.30$
			FCal1: $\sim$ four times finer	$3.10 < \|\eta\| < 3.15,$
				$4.30 < \|\eta\| < 4.83$
			FCal2: 3.3×4.2	$3.24 < \|\eta\| < 4.50$
			FCal2: $\sim$ four times finer	$3.20 < \|\eta\| < 3.24,$
				$4.50 < \|\eta\| < 4.81$
			FCal3: 5.4×4.7	$3.32 < \|\eta\| < 4.60$
			FCal3: $\sim$ four times finer	$3.29 < \|\eta\| < 3.32,$
				$4.60 < \|\eta\| < 4.75$
Readout channels			3524 (both sides)	
Scintillator tile calorimeter				
	Barrel		Extended barrel	
$\|\eta\|$ coverage	$\|\eta\| < 1.0$		$0.8 < \|\eta\| < 1.7$	
Number of layers	3		3	
Granularity $\Delta\eta \times \Delta\phi$	0.1×0.1		0.1×0.1	
Last layer	0.2×0.1		0.2×0.1	
Readout channels	5760		4092 (both sides)	

Figure 3.13. Main parameters of the ATLAS calorimeter system Aad [2008a].

Tile calorimeter. is located immediately outside of the barrel EM calorimeter. It has a barrel section covering $\|\eta\| < 1.0$ as well as two extended barrel regions spanning over $0.8 < \|\eta\| < 1.7$. It is a sampling calorimeter that uses steel as the absorber and scintillating tiles as the active material.

The hadronic end-cap (HEC). The HEC consists of two wheels per end-cap, located directly behind the electromagnetic calorimeter at the end-cap (EMEC). The HEC covers the region over $1.5 < \|\eta\| < 3.2$, overlapping partially with both the forward calorimeter (which can be seen in Figure 3.16 and the tile calorimeter. Each wheel is divided into two segments in depth, for a total of four layers per end-cap. One wheel is constructed with 32 identical wedge-shaped modules. Each of them is an
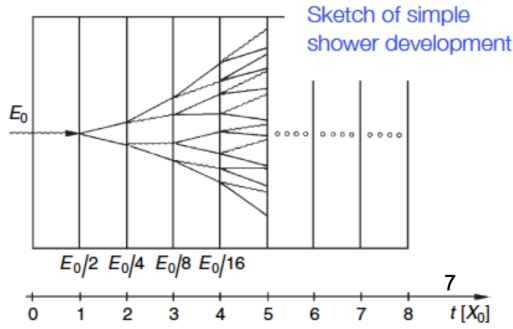
Figure 3.14.   A sketch that shows how a shower initiated by a particle is developed while traversing through materialsho.
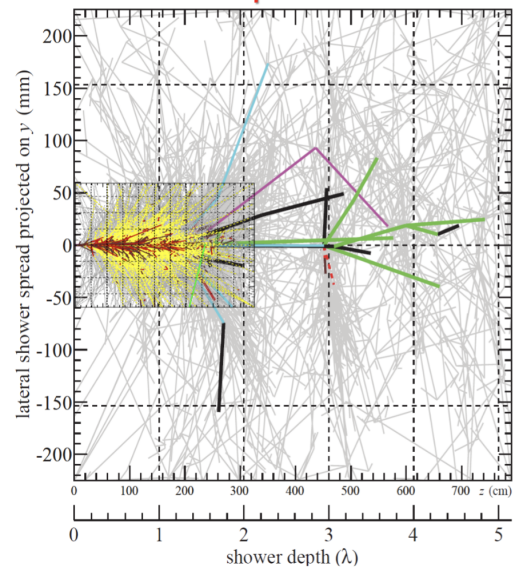
Figure 3.15.   Simulation of a 10 GeVelectron(left) and pion (middle) traversing through copper generated by GEANT4 software; the right most figure has the two overlaying on top of each other for easier comparison, credit: P. Loch, U.Arizona.

alternation of 25 mm thick copper absorbers (50 mm in the second wheel), and signal collection gap.

FCal.   Aad [2008a] FCal is at the most forward region of the detector, covering $3.1 < |\eta| < 4.9$. Each end cap consists of 3 modules arranged as shown in Figure 3.16: the one closer to the beam pipe is made of copper and is optimized for EM measurement; the other two are made of tungsten and are optimized for hadronic interactions. Each module is a medal matrix with regularly spaced electrode in the shape of a rod parallel to the beam pipe as shown in Fig 3.17. The LAr in the gap between the rod and the tube is the sensitive medium. FCal is exposed to high
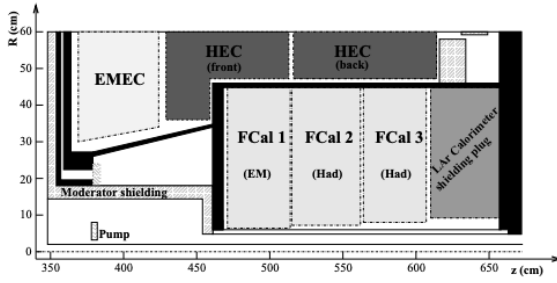
Figure 3.16. Schematic diagram showing the three FCAL models located in the end-cap cryostat. The material in front of the FCAL and the shielding plug behind it are also shown. Aad [2008a].
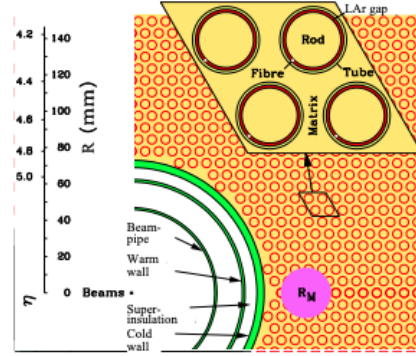
Figure 3.17. Electrode structure of FCal1 with matrix of copper plates and tubes and rods with LAr gap Aad [2008a].

particle fluxes due to its really forwarded location. This results in a detector design with LAr gap width (FCal1)0.269 mm, (FCal2)0.376 mm and (FCal3)0.508 mm.

3.2.5 Muon system

Muons are extremely penetrating particles. Cosmic ray muons created in the atmosphere can penetrate through earth for more than a mile before being stopped. so the ATLAS muon system Aad [2008a] is located at the outer most layer of the detector. Fig 3.18 shows the layout of the muon spectrometer. While flying outward, muons created from the interaction points get deflected by magnetic field from two different sources: the large toroid magnets in the barrel region are used for $|\eta| < 1.4$; while the two smaller end-cap magnets are used for $1.6 < |\eta| < 2.7$. In the transition region $1.4 < |\eta| < 1.6$, muons can be influenced by magnetic fields from both sources. The muon system is equipped with its own tracking system so the tracks are measured with high precision.

There are 4 types of muon chambers, Monitored Drift Tubes (MDT), Thin-Gap Chambers (TGC), Cathode Strip Chambers (CSC) and Resistive-Plate Chambers
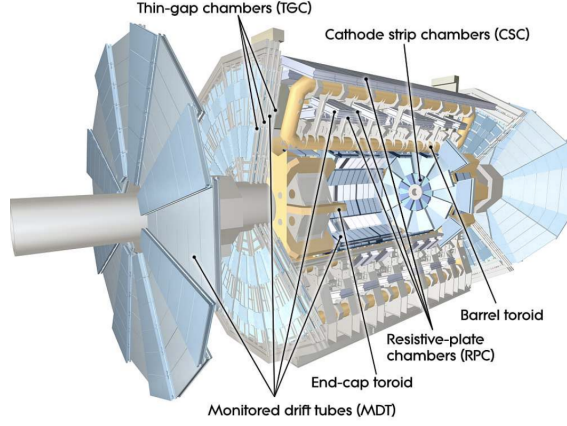
Figure 3.18. Layout of the ATLAS muon systemAad [2008a].

(RPC). The MDT takes the main role of measuring tracks in the main bending direction of the magnetic field, over a wide range of η. The CSC are multi-wire proportional chambers with cathodes segmented into strips. They are used at large η over $2 < |\eta| < 2.7$ in the inner-most plane with much higher interaction rates. RPCs are used in the barrel and TGCs are used in the end-cap regions. A trigger system is dedicated for the muon system; it serves the purpose of providing bunch-crossing information, p_{T} threshold as well as keeping good measurements of muon coordinates.

The alignment between different muon chamber is critical for the performance of the muon system, especially for muons with high momenta. The internal deformations and relative positions of the MDT chambers are monitored by approximately 12000 precision-mounted alignment sensors, all based on the optical monitoring of deviations from straight lines. The optical system is really sensitive and able to detect relative changes in chamber position at the level of 20 μm. Although it cannot detect some specific geometrical distortions, additional optical lines are used to compensate for it. Once fully functional, the optical alignment system can determine the absolute chamber positions in the barrel region to an accuracy of a few hundred microns.
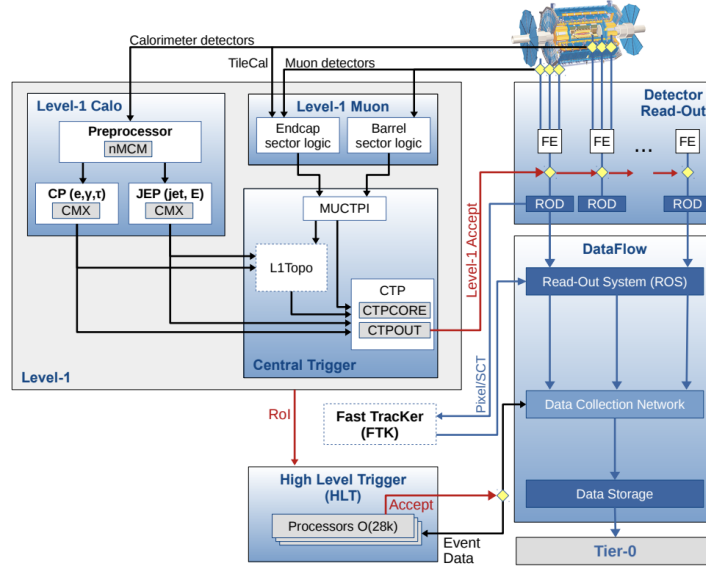
Figure 3.19. The ATLAS TDAQ system in Run2 with emphasis on the components relevant for triggeringAab [2017].

3.2.6   Trigger and Data Acquisition system (TDAQ)

The huge bunch crossing rate of the LHC leads to more than 1 billion of particle collisions per second in the middle of ATLAS, leading to massive amount of data generated by the detector. ATLAS has a combined data volume of more than 60 million megabytes per second which is equivalent to 5400 simultaneous streams of 4K video. However, only a small fraction of the events may contain information interesting for physics analysis. The ATLAS trigger Abolins et al. [2016], Martínez [2016] filters the events down to 1 kHz (from 40 MHz initially) by rapidly identifying signatures of physics objects including electrons, muons, photons, tau leptons, jets and B meson candidates. Events that passed the trigger system are stored later through the data acquisition system.

Fig 3.19 shows the schematic diagram of the ATLAS TDAQ system during Run 2. It illustrates the data pathway after being generated by ATLAS. In general, the triggering is done in two main steps:

1st level hardware trigger. The hardware trigger executes preliminary selecting tasks by programming hardware electronics. In Run 2, Level-1 Calo and Level-1 Muon belong to this level of trigger. The 1st level trigger elements accept data directly from calorimeters and muon detectors and are designed to run in sync with the LHC bunch crossing. Finally, the Central Trigger Processor(CTP) makes the Level-1 Accept decision. After the Level-1 trigger decision, the event rate is reduced to 100 kHz.

Software-based High Level Trigger (HLT). After the L1 triggers, the events are processed by the HLT which includes information from more sub-detectors. The HLT algorithm can access information from finer granularity calorimeters, muon system and tracking information from the inner detector which was not available to Level-1 triggers. Most HLT triggers use a two-stage approach with a fast first-pass reconstruction to reject the majority of events and a slower precision reconstruction for the remaining events.

3.2.61 ATLAS Run 3 trigger system. A schematic diagram of the ATLAS Run 3 trigger system run [2024] is shown in Figure3.20. L1CALO has been upgraded to perform on-detector digitization of transverse energies from the LAr calorimeters. This change allows the trigger to receive finer granularity data from the LAr calorimeter. Prior to Run 3, the ATLAS L1Calo trigger takes trigger towers (0.1 × 0.1) as input. For Run 3, the calorimeter info is provided in the form of SuperCells
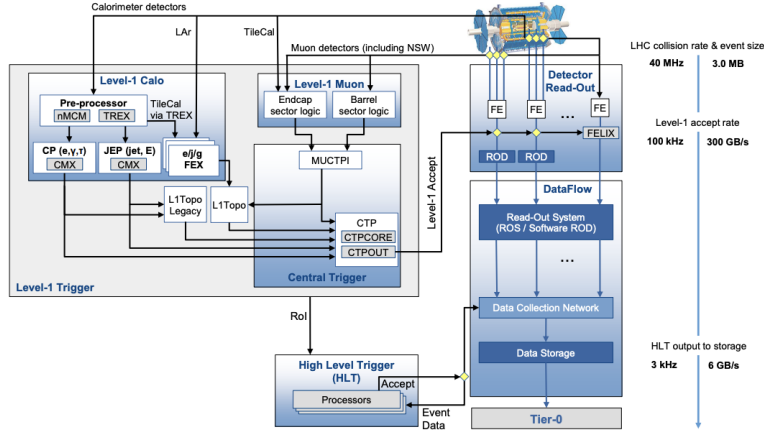
Figure 3.20. ATLAS Run 3 trigger system run [2024].

(containing sums of four or eight calorimeter cells). Electromagnetic (eFEX) and jet (jFEX) feature extractors as well as a global feature extractor (gFEX) are newly added to the L1Calo system.

A new read-out system was also introduced to the DAQ system in Run 3, called Front-End Link exchange (FELIX). It takes input from Level-1 Calo, Level-1 Muon and the Central Trigger. And it is responsible for validation before data is sent for storage. See run [2024] for complete list of changes made in the Run 3 trigger system.

3.2.62   Changes of the trigger system in preparation for HL-LHC.   A short section is dedicated to the future trigger system in preparation for Run-4 pha due to its strong relevance to the qualification task.

The HL-LHC expects to produce up to 200 proton-proton collisions per bunch crossing ($\mu = 200$), where μ is the number of inelastic collisions produced by a single collision. The high pile-up level poses special challenges to the Trigger and Data Acquisition (TDAQ) systems of the ATLAS detector CER [2017]. During LS3, the TDAQ system will go through a major upgrade. The trigger system therefore introduced the concept of a Global Trigger system to handle this new challenge. The
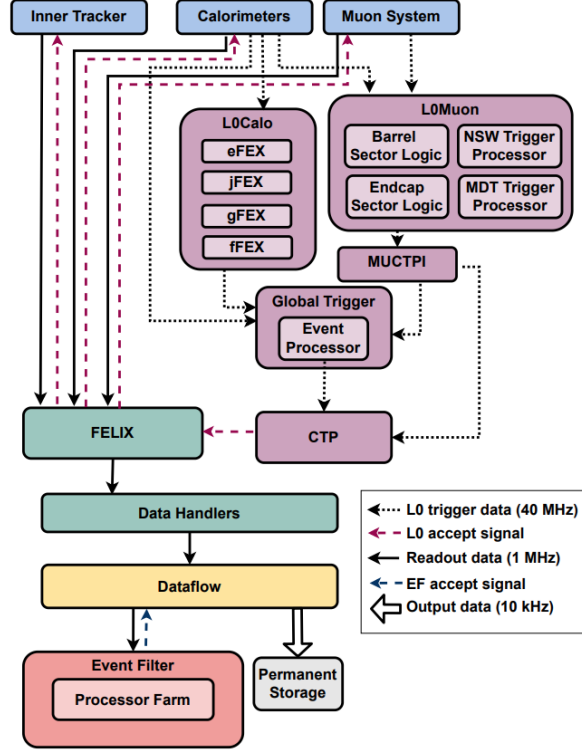
Figure 3.21. The TDAQ architecture for the HL-LHC illustrating the trigger and data acquisition readout chain with the hardware level-0 trigger and high-level trigger, known as the Event Filter. Events selected by the Event Filter are then sent to permanent storage CER [2017].

goal is to bring off-line event filtering capability to the Level-0 trigger system (Level-1 system in Run-2). The overall TDAQ structure is illustrated in Figure 3.21.

The schematic diagram Fig 3.22 shows how data would flow to the Global Trigger System. Different from the Run 2 scheme, full granularity data from the LAr calorimeter are streamed to the Global Trigger at 40MHz. The calorimeter data are fed to the Global Event Processor (GEP) where all the trigger algorithms will be implemented, after an array of Multiplexers, as shown in Figure 3.22. There will be many GEP modules running in parallel. Every GEP is hosted on a Xilinx Versal Premium 1802. Each FPGA is divided into 4 Super Logic Regions (SLRs) made up

66
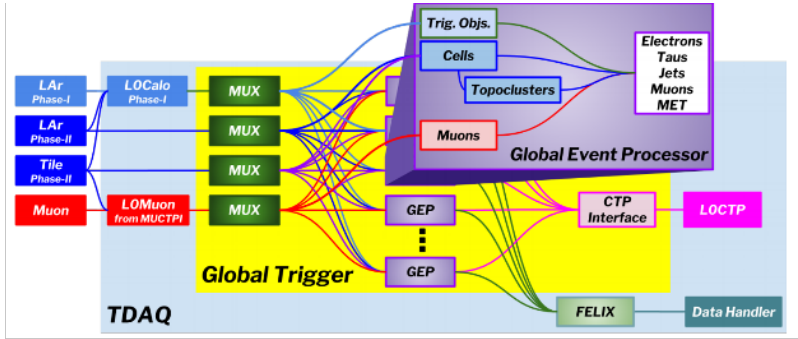
Figure 3.22. Illustration of the Global Trigger system processing and trigger decision flow.

of silicon die slices. All 4 SLRs are interconnected to allow results to flow from one to another. The GEP has $4\mu s$ to run all the trigger algorithms on an event.

After the Global Trigger, the Central Trigger Processor (CTP) makes the final trigger decision on an event and throws out unwanted events. After the Global Trigger, the event rate will be reduced to 1MHz. Eventually, after the Event Filter decision, 10kHz of the data will be stored on tape.

Similar to the Run-2 and Run-3 scheme, after the Global Trigger, the Central Trigger Processor (CTP) makes the final trigger decision on whether or not to keep an event. After the Global Trigger, the event rate will be reduced to 1MHz. Eventually, after the Event Filter decision, 10kHz of the data will be stored on tape.

3.2.7 Particle reconstruction

Each elementary particle interacts with detection material differently so strategies have been developed to identify each particle. Fig 3.23 illustrates how different particles interact with different layers of the detector. This section will describe the identification strategies Aad [2008a] for the most commonly mentioned particles/physics objects.
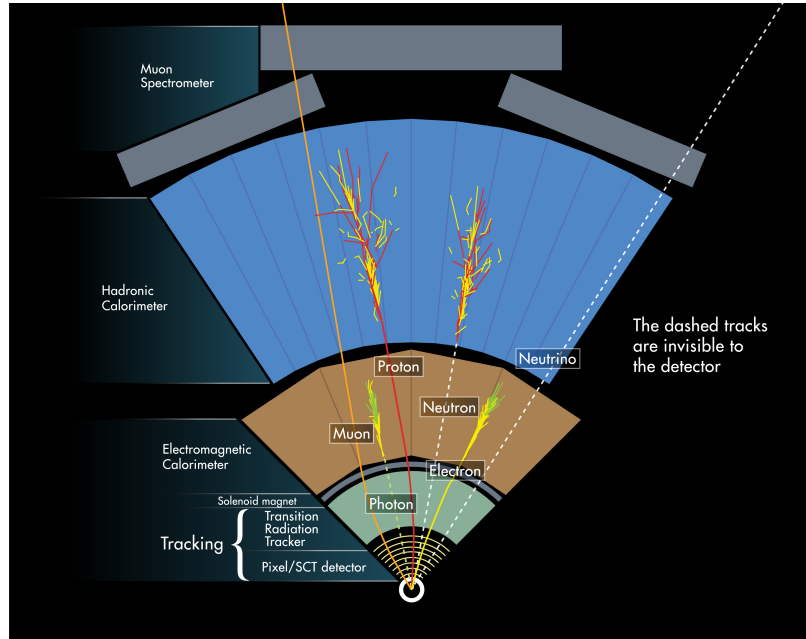
Figure 3.23. How ATLAS detects different particles: diagram of particle paths in the detectorPequenao and Schaffner [2013].

Electrons and photons. Electrons leave tracks in the inner detector and produce showers in the electromagnetic calorimeters. For standard reconstruction of electrons and photons, a seed cluster is taken from the energy deposits on electromagnetic calorimeter cells and a loosely matching track is searched for among all reconstructed tracks in the ID. Electrons and photons can be distinguished by requiring electrons to have associated tracks and no associated conversion while photons should meet the opposite of the these conditions. Isolated high-p_T electrons are identified based on cuts and the shower shapes, on information from the reconstructed track and on the combined reconstruction. Different criteria are used to define electrons in data analysis depending on the signal efficiency and jet rejection requirements of the specific physics sample.
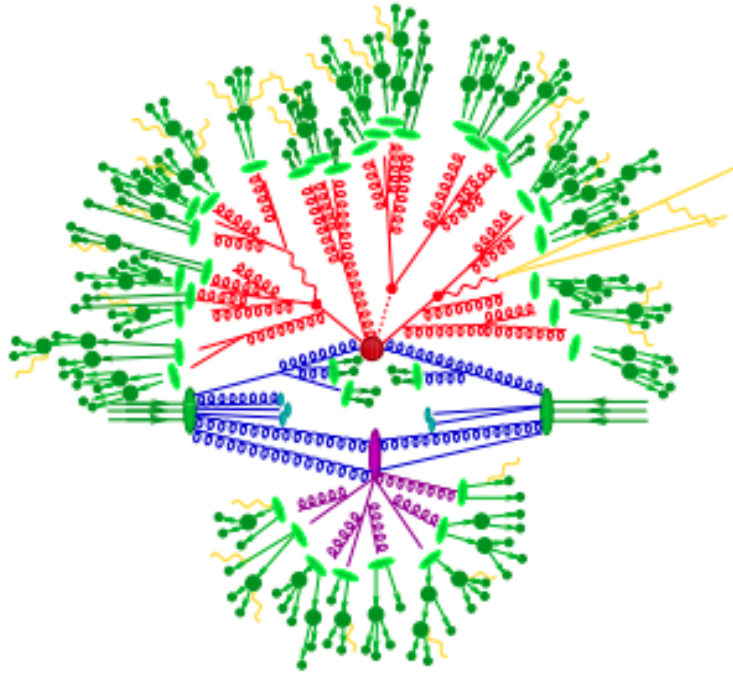
Figure 3.24. Pictorial representation of a $t\bar{t}$H event. The hard interaction (big red blob) is followed by the decay of both top quarks and the Higgs boson (small read blobs); (red lines) represent additional QCD radiations and the (purple blob) represents secondary interaction before the final state partons hadronise (light green blob) and hadrons decay (dark green blob); photon radiations are marked by (yellow lines) Gleisberg et al. [2009].

Jets. Jet formation is illustrated on Fig 3.24 through a $t\bar{t}$ H event. LHC collisions produce single partons with color charges; colored objects are not stable and cannot be observed, but when the colored objects interact with each other through strong interactions and hadronise, the combined colorless objects (such as pions $\pi^{\pm}$) are stable and can be detected through energy deposits on the calorimeters. Jets are collimated sprays of partons from hadronization and can be identified through clustering algorithms using calorimeter deposits. There are many different jet algorithms but ideally, a jet algorithm should have the following general properties: Ellis [2010]

– Fully specified: include definition in detail of how to map particle list onto jet list

– Theoretically well behaved: the algorithm should be infrared and collinear safe and insensitive to the addition of low energy partons.

– Detector independence: there should be no dependence on cell type, numbers or size relating to calorimeters

– Order independence: the algorithms should behave equally at the parton, particle, and detector levels.

The majority of ATLAS analyses use an anti-k_t jet clustering algorithm with R=0.4 as small R jet definition. The details of the algorithm are described in Cacciari et al. [2008a]. A distance is introduced between two entities as in Equation 3.2 where $\Delta_{i,j}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$. The algorithm proceeds by starting with the initial list of entities/protojets; the algorithm recursively groups pairs of closest protojets together using d_{ij} to form new protojets. The distances are recalculated at each iteration and the procedure repeated until no entities are left.

70
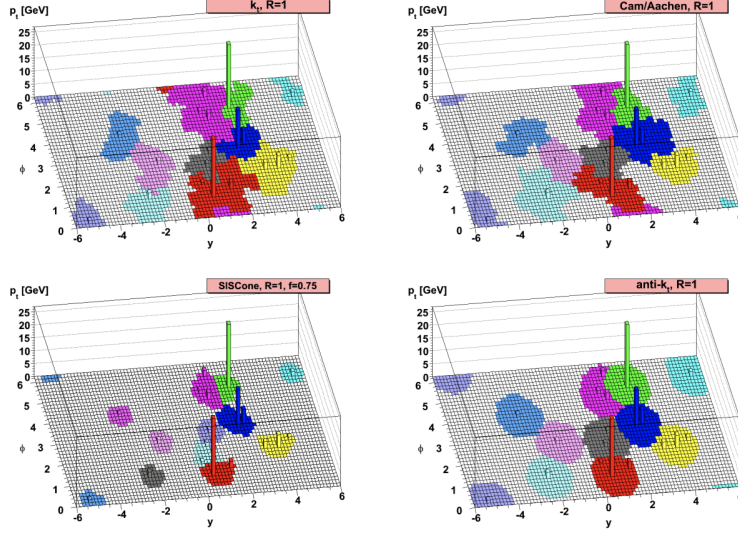
Figure 3.25. Comparison of the $anti-K_T$ with other jet algorithm Cacciari et al. [2008a].

$$d_{ij} = min(k_{t,i}^{-2}, k_{t,i}^{-2})\frac{\Delta_{ij}^2}{R^2} \tag{3.2}$$

Using a parton-level event plus around 10^4 random soft particles, the behavior of 4 different jet algorithms are compared in Fig 3.25. The anti-k_t algorithm gives the characteristic circular shapes for most of the hard jets and only the soft jets end up with irregular shapes. These characteristics meet the expectations of a jet algorithm and has been the most popular jet reconstruction method in ATLAS analyses.

Missing transverse momentum. A good measurement of the missing transverse energy E_T^{miss} is crucial for studying many new physics phenomenon beyond standard model such as Supersymmetry and the signals studied in this thesis. Standard model neutrinos or other beyond standard model particles (if they exist) escape the ATLAS detector without being detected. Using momentum conservation in the transverse plane, the missing momentum can be measured through momentum

71

imbalance of all the observed particles. Therefore, $E_{\mathrm{T}}^{\mathrm{miss}}$ can be mathematically represented as the vector sum of jets, electrons, muons and the soft energy terms $E_{\mathrm{T}}^{\mathrm{miss}} = \left| \sum_{jets} p_{\mathrm{T}} + \sum_{leptons} p_{\mathrm{T}} + \sum_{soft} p_{\mathrm{T}} \right|$ where the soft term accounts for tracks with p_{T} greater than 500 MeV and calorimeter clusters that are not selected as objects.

Muons. Aad [2008a] LHC collisions produce muons with a wide range of energies. While the muon spectrometer is the main sub-detector that performs precision measurements, inner detector and calorimeters are also used to improve muon identification efficiency and momentum resolution. The inner detector provides the best measurements at low to intermediate momenta while the muon spectrometer takes over for energies above 30 GeV. To achieve the desired performance combining multiple sub-detectors, the inner detector and the muon spectrometer need to be calibrated and aligned with respect to each other.

Tau leptons. Aad [2008a] Hadronic decays of tau leptons play an important role in probing new physics beyond standard model. The large number of hadronic jets is the main source of background for identifying taus and makes the reconstruction very challenging. To efficiently reconstruct tau leptons and reject jet background, they are measured from two complementary sources: calorimeters and the tracks of from inner detector. In general, taus can be identified by requiring narrow calorimeter deposits with one or three tracks with total charge matching that of a tau lepton. See chapter IV for more details of calorimeter based tau triggering.

CHAPTER IV

TAU TRIGGER FOR THE ATLAS PHASE II UPGRADE

In the era of the HL-LHC, ATLAS plans to use a Hardware-based single level trigger system (level 0), which includes both the legacy system L0Calo and L0Muon and the Global trigger, as shown in Figure 4.1. Full granularity data from the LAr calorimeter are streamed to the global trigger at 40MHz. The high pile-up level and full granularity data will create challenging environment for the data acquisition(TDAQ) system, so upgrading the trigger system is a crucial part of the upgrade that leads to the operation of HL_LHC (Phase II upgrade). See Ref. tda for the Technical Design Report of TDAQ system for the Phase II upgrade.
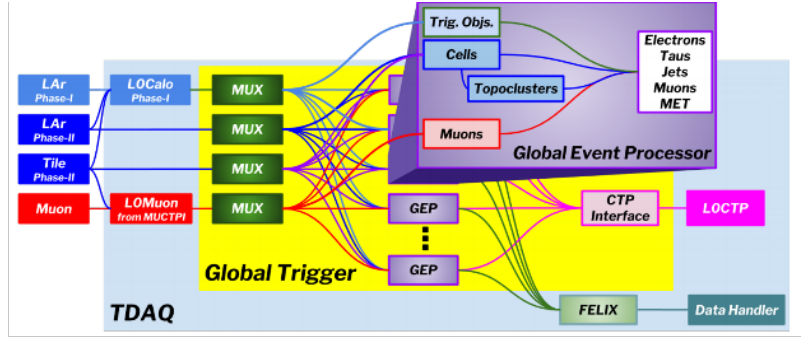
Figure 4.1. Dataflow and trigger algorithms (in the purple box) in the Global Trigger system from Ref. tda.

The goal of the task is to develop a tau trigger algorithm like the eFEX (Electron Feature Extractor) tau algorithm being commissioned in Run 3 of ATLAS data taking. That version of the tau trigger was developed by a graduate student at University of Oregon, Nicholas Luongo. Script eFE is his work of the eFEX tau trigger code. Note there is another version of eFEX tau trigger code by Denis in Ref. L1c. The difference between the two is that Ref. L1c was developed for the purpose of performance study and Ref. eFE is more intended for firmware emulation. Before introducing the Tau

algorithm for the Global Trigger, it is important to understand the eFEX tau trigger in the phase I upgrade.

4.1 The eFEX Tau algorithm

There are two criteria that determine whether to trigger an event on taus: the reconstructed tau transverse energy (E_T), $E_\mathrm{T}^\mathrm{reco}$ in short and the rejection of jets based on the so-called f_Core variable. The following is the content of the algorithm in which all calculations are based on Super cells (SC). Refer to Figure 4.3 for the sizes of SC and how they compare to the sizes of the elementary cells.

1. **Seeding** This 1st step if the algorithm is called seeding. Elementary calorimeter cells are combined into courser granularity cells (super cells, as described above). Trigger Towers are built by summing these super cells from each layer. Seeds are then defined as the towers with local maximum energy. By the end of this step, the algorithm should return a list of seeds, together with their ϕ and η info.

2. **Clustering**, the 2nd step is to determine the tau candidate energy. The reconstructed tau E_T is obtained by taking the seeds from the 1st step and developing clusters around them: cell energies are summed up around each seed according to a rectangular window.

   Each layer of the detector has a different granularity so the window choice for each layer is different. In Phase-I upgrade, data from EM0, EM1, EM2, EM3 and HAD (Hadronic Calorimeter) layers are used in the tau trigger. Because, EM1 and EM2 have finer granularity than the other 3 layers, 5×2 SC window was used for the EM1 and EM2 layer and 3×2 SC window was used for

74
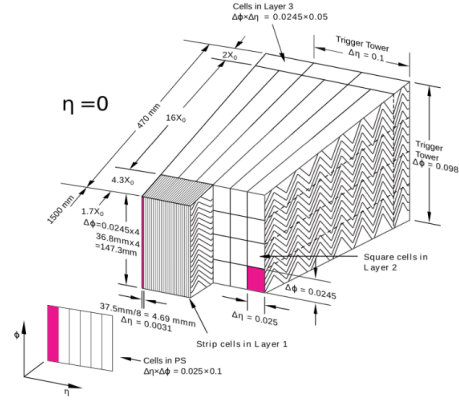
Figure 4.2. The segmentation of readout cells in LAr EM Barrel calorimeter lAr.
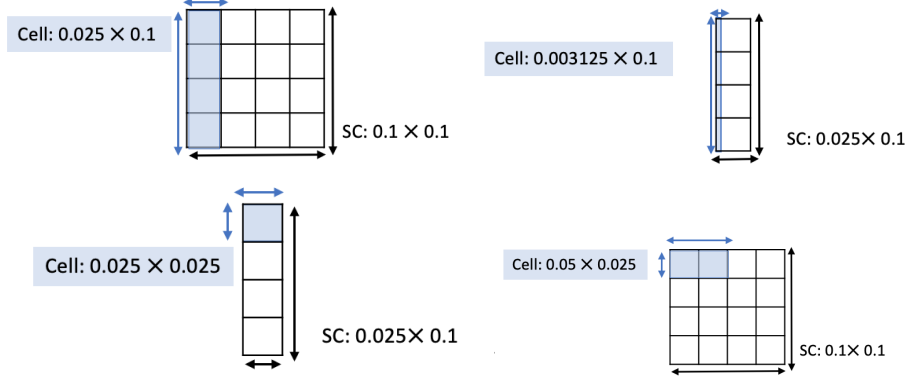
Figure 4.3. Super cell(SC) and cell size comparison in EM 0 layer(Pre-sampler), EM 1 layer(Front), EM 2 layer(Middle), EM 3 layer(Back), in the order of left to right and upper to lower.

the EM0, EM3 and HAD window. The reconstructed E_{T} of tau candidate is then calculated by summing the energy reconstructed from each layer of the calorimeter

$$E_{\mathrm{T}}(\tau) = E_{\mathrm{T}}(L0) + E_{\mathrm{T}}(L1) + E_{\mathrm{T}}(L2) + E_{\mathrm{T}}(L3) + E_{\mathrm{T}}(HAD)$$

3. **Isolation variable** f_{Core} The hadronically-decaying taus result in jets that are very similar to jets originated from quarks, so it is very challenging to distinguish

75

taus from jets. However, jets originated from quarks are more spread out on the calorimeter due to the subsequent parton shower that results in many more pions. The variable f_{Core} measures how concentrated the shower profiles is and can be used to distinguish taus from jets. It is calculated solely based on EM2 layer and is defined by the ratio between the inner part of the window(3×2 SC) and the total energy of the window.

$$f_{\text{Core}} = \frac{E_T(3 \times 2)}{E_T(9 \times 2)}$$

The f_{Core} for taus is more likely to be larger than that of the quark jets. So a selection based on the value of this variable will reduce the amount of background events from quark jets while keeping events from the potential tau candidates. According to the eFEX tau trigger study in Ref. nic done by a senior graduate student Nicholas Luongo, adding the isolation variable reduces the trigger rate by about 10%.

4. **Assign trigger bit**

   Finally, based on the isolation variable and reconstructed E_{T}, the Boolean trigger bit for taus is assigned to each event.

The Run 4 version of the algorithm will largely remain the same as the Run 3 version except the clusters and trigger towers will be built with elementary cells. All of the major changes with respect to the eFEX version of the algorithm are explained in the following section.

4.2 Athena tau trigger for HL-LHC

For the HL-LHC, the tau algorithm was developed under the Run4 GEP simulation framework using Athena release 21.9.26 (see the wiki page Ref. twi about how to set up the Athena release before compilation). The framework Ref. run was written by Maria Mazza and others from Michigan State University (MSU). Maria's concluding presentation for developing the framework can be found in Ref. tal.

The Athena tau trigger code itself and the necessary changes on the framework are currently stored and updated under a fork tau, please refer to it for the latest version. The reference code used to develop the Run 4 tau trigger in this work is located Ref. L1c in git.

The framework is responsible for preparing the input data to various trigger algorithms, coordinating each algorithm to be ran in the correct order (see L0GepSimulationSequence.py) and finally, store the results of each algorithm in Athena format (see GepOutputReader). The framework determines when to run the tau algorithm and it provides the input cell data to the tau algorithm. However, for simplicity, the output n-tuples from the tau algorithm are saved independently from the results of the other algorithms. And the output tree is defined inside of the algorithm code, see the trigger code itself Ref. tau for more details.

Prior to the addition of the tau trigger algorithm, there were already several algorithms added under the framework, such as the topo clustering algorithm and the anti-kt algorithm.etc. The tau trigger was then written with a similar structure to those existing algorithms. The detailed instructions of running the framework is given in the Twiki page twi. The data sample that was used to test the algorithm is a $Z \rightarrow \tau\tau$ sample with $\mu = 200$. The information of the signal sample listed in

77

the Rucio website can be found in Ref. zta. The background sample is a dijet JZ0
sample with $\mu = 200$; the dijet sample information in Rucio can be found in JZ0

4.2.1   Detailed description of the algorithm

4.2.11   Build towers using elementary cells.   To seed the algorithm in a coarser
granularity, the 0th step of the algorithm is to build trigger towers. The trigger
towers were built by feeding elementary cells from all layers (sampling=0 through 20)
available (including Tile cells) into a 2D histogram with bin width 0.1×0.1 in the η
and ϕ direction.

The elementary cell data of the algorithm were accessed using `CustomCaloCell`
provided by the Run 4 simulation framework mentioned above. The list of elementary
cell information that can be accessed with this class can be found in Ref. cus.

if the tower E_T is less than a threshold, the tower will not be considered in the
local maximum finding, this threshold is called `m_towerEtCut` and implemented as
a property. Figure 4.4a shows the energy distribution of the towers with 0.1×0.1
granularity.

(a) E_T distribution of the trigger towers   (b) E_T distribution of the seeds
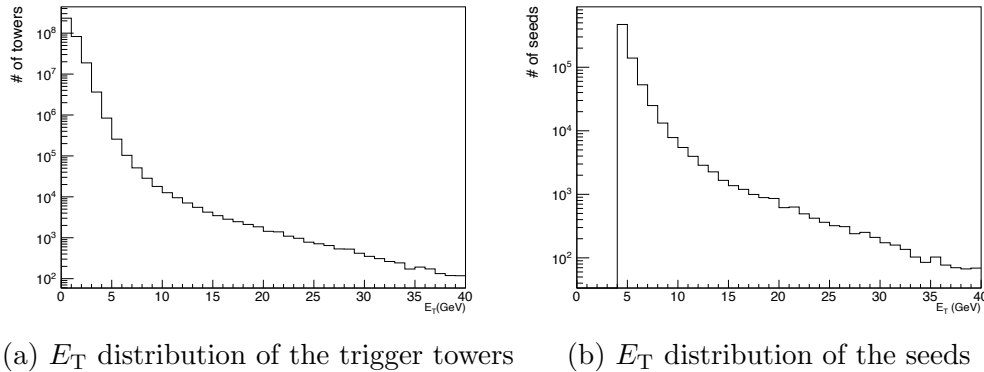
Figure 4.4. $Z\tau\tau$ Transverse energy distribution of the trigger towers and that of the
seeds.

78

4.2.12 **Seeding.** The seed finding code is nearly identical to that of the eFEX tau algorithm, see L246 in Ref. L1c for the reference used. In this step, every tower is compared to its surrounding 8 towers in a 3×3 window in the 0.1×0.1 granularity. The tower that has a local maximum E_{T} will be used as a seed. These 8 surrounding towers are divided into two groups (binsAbove and binsBelow) shown in Figure 4.5. The binsAbove includes the 4 cells that are above and to the right of the seed tower(in red) and the binsBelow includes the 4 cells below and to the left of the seed tower.

There is a reason for dividing the 8 cells into two groups. If we checked all cells less than we would end up with no seeds from that region and if we checked all cells less than or equal to we would end up with two right next to each other. To make sure that only one seed is found when two towers side by side have exactly equal energies, a strictly less than comparison is used for binsAbove whereas less than or equal to is used for binsBelow.

if the seed E_{T} is less than a threshold, the seed will not be used to form clusters in the next step. This threshold is declared as a property called `m_SeedEtCut`.

Figure 4.5. eFEX: central tower seed cell condition.

The same comparison process is done on every single tower from the 2d histogram. By the end of this step, the algorithm outputs TLorentz vectors of all the local maxima in tower granularity.

To later form clusters in each layer, the algorithm needs to provide the location of the seed in different elementary cell granularity, from each layer included. Cell maps were built for all layers by filling 2D histograms with cellη, cellϕ, cellID. Seed cells were then found by referring to the 2D histograms according to the the local maxima in tower granularity.

After getting a rough location of the seed cell, a re-seeding step is performed to get a more accurate location of the seed. In this re-seeding step, the surrounding cells are checked in a window of 3×3 of elementary cells in layer 2. The cell with the largest E_T among the surrounding cells is used as the final seed cell. This re seeding process is done only on layer 2 and the same seed location (η and ϕ) will be adopted for other layers. This is to make sure that the position of the seed is aligned between different layers.

4.2.13 Clustering. The goal of the clustering is to calculate the tau candidate E_T through reconstruction (E_T^reco). To focus on the barrel region of LAr, clusters are only built within $-1.3 < \eta < 1.3$ of the calorimeter. The clustering window size in each layer is modeled after the eFEX version, as shown in Figure 4.6. Table 4.1 lists the width and height of the windows in the number of elementary cells.

Table 4.1. Clustering window sizes used in each layer and for isolation calculations. Window sizes are repesented as number of cells in $\eta \times$ number of cells in ϕ

	Size in $\delta\eta \times \delta\phi$	Window sizes
EM0	0.3×0.2	12×2
EM1	0.125×0.2	40×2
EM2	0.125×0.2	5×8
EM3	0.3×0.2	6×8
f_Core inner sum (Layer 2)	0.075×0.125	3×5
f_Core outer sum (Layer 2)	0.275×0.425	11×17

EM3 layer EM2 layer EM1 layer EM0 layer

Cell: 0.05 × 0.025 Cell: 0.025 × 0.025 Cell: 0.003125 × 0.1 Cell: 0.025 × 0.1

0.3× 0.2 0.125× 0.2 0.125× 0.2 0.3× 0.2

Summing area, each grid has size of (0.025 * 0.025) Elementary cell in each layer (shown above) Seed tower (0.1 *0.1)
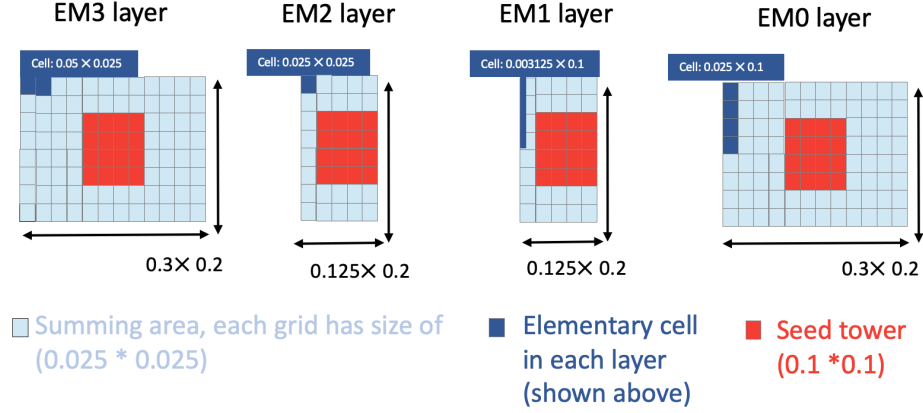
Figure 4.6. Cluster sizes in each layer within the LAr barrel (red); the sizes of the elementary cells are shown in dark blue. See Table 4.1 for the sizes in terms of the number of cells.

However, the clustering method used in this work is very different from that of the eFEX version. Instead of using 2D histogram bins to refer to the cells, cell ID's were used. The `getneighbour` function from the Run4 simulation framework provides a list of IDs for 8 surrounding cells given any elementary cell. So this function was used in the algorithm to construct the clusters.

While the cluster in the eFEX version is asymmetric around the seed, the cluster in this work is developed by adding rings of cells going further and further away from the seed. Therefore the clusters are symmetrical around the seed , as shown in the left part of Figure 4.7. The clustering method takes in a window width and height as arguments. The position of the cell inside each ring relative to the seed location ($\delta\eta$ and $\delta\phi$) will be checked before including the ring to the cluster. The clustering is stopped as soon as one of the cells in the ring goes beyond the size of the window as set in the arguments of the function.

This process often results in a window that is larger than the window size desired. So afterwards, the window size is reduced down to the desired size by checking the

distance between the seed and the clustered cells, as shown in the right part of Figure 4.7. Only cells within the region of the window are added to the final cluster.

After all cells within the clustering window are identified, the cell E_T needs to meet two criteria before a cell is added to the tau reconstructed E_T . The following criteria are used whenever a calculation of the reconstructed tau E_T is performed.

1. The cell $E_T > 0$, because only positive E_T cells are processed by the FPGA.

2. The cell $E_T > 2\sigma$ noise level, because the Liquid Argon Calorimeter does not send cells with E_T below 2σ to the Global Trigger.
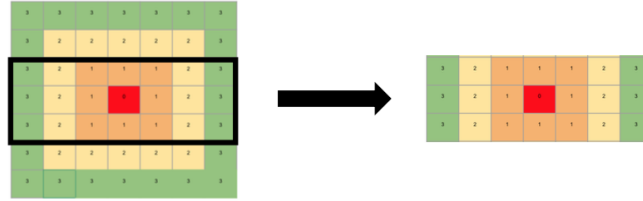
Figure 4.7.   Clustering process of choosing a window larger than desired, then reducing down to the window with desired size (black box).

From the implementation of the clustering algorithm, the cluster from each layer looks like Figure 4.8; the sizes of the resulting clusters are similar to what was intended as specified in Table 4.1 and visually shown in Figure 4.6

After the clustering is done on each layer, the final reconstructed tau candidate E_T is calculated simply by adding the cluster sum from each layer. Figure 4.9 shows the fraction of energy contributed by each layer, as expected, layer 2 tends to contribute the most energy to the E_T^{reco} . Layer 3 is furthest away from the interaction point among all the layers included here so tau candidates leave the least amount of energy in layer 3. As shown in the figure, layer 3 most frequently contributes zero to the reconstructed E_T.
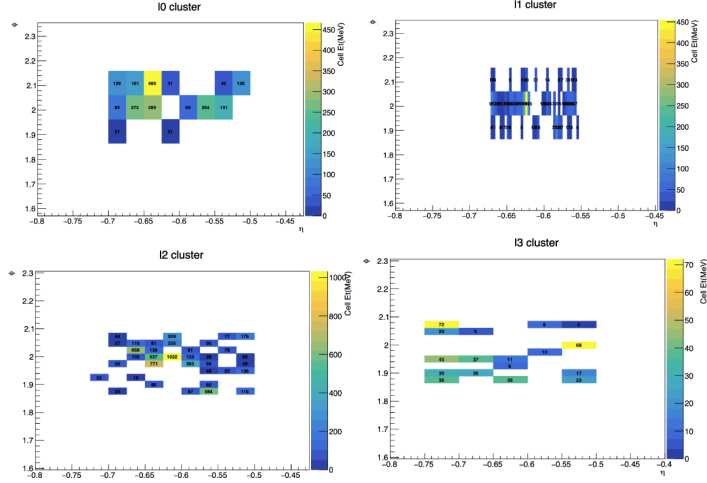
Figure 4.8. Cluster size constructed in each layer; all positive E_T cells are shown within the clustering window, the negative E_T cells are shown as white space within the window. No 2σ selection applied here for better showing the shape of the cluster.
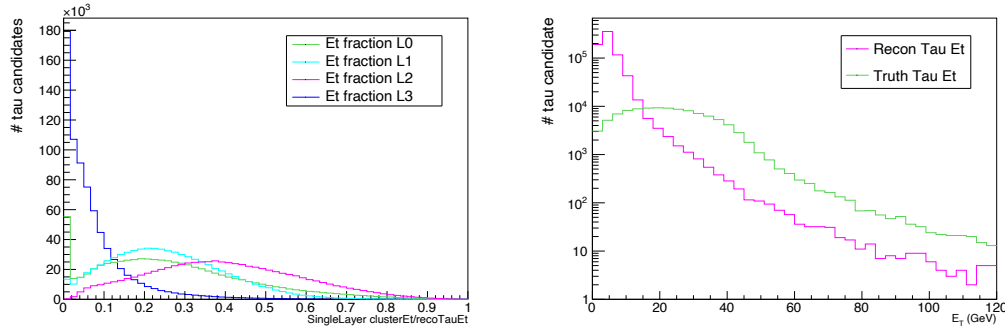
Figure 4.9. Fraction of reconstructed E_T contributed by each layer is on the (left); Reconstructed tau candidate E_T and visible truth tau E_T are on the (right).

Finally, Figure 4.9 shows the distribution of the reconstructed tau E_T and the visible truth tau E_T (with neutrino T 4-vector subtracted).

4.3 ISOLATION

In the beginning of the development, the window sizes were applied to the calculation of the f_{Core} and it was found that it does not provide separation between signal and the background when using elementary cells. Thus, the outer window was

enlarged and the inner window was reduced so that the windowed sum between the two can have a larger difference. Similar to the eFEX version of the algorithm, the isolation is calculated only on the EM2 layer:

$$f_{\text{Core}} = \frac{\text{Inner window (yellow)}}{\text{outer window (white)}} = \frac{\text{cell}(3 \times 5)}{\text{cell}(11 \times 17)} = \frac{\delta\eta \times \delta\phi = 0.075 \times 0.125}{\delta\eta \times \delta\phi = 0.275 \times 0.425}$$

The new window choice is shown in Figure 4.7.
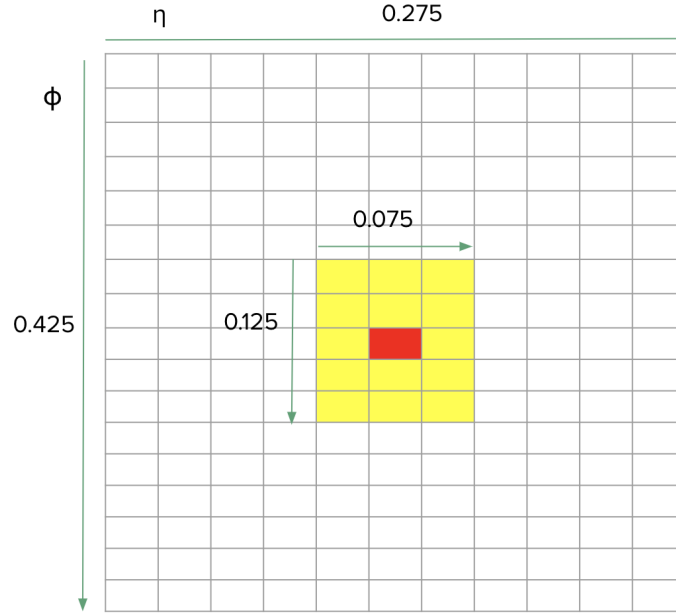
Figure 4.10. The window choice for isolation in layer 2 (ϕ vs. η); each grid is an elementary cell in layer EM2.

Figure 4.11 shows the distribution of inner and outer window cluster E_{T} from the above mentioned method.
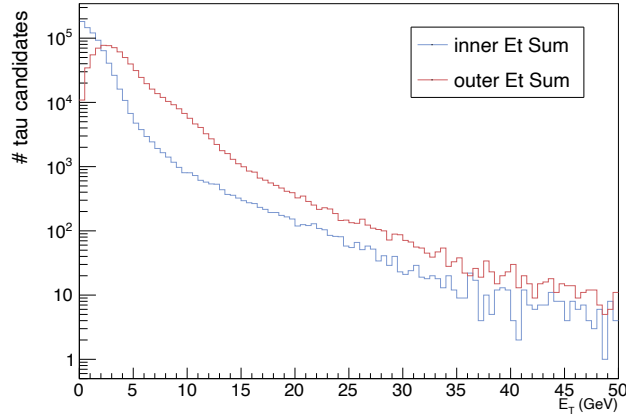
Figure 4.11. Inner and outer window sums for calculating the isolation between tau candidates and quark jets.

4.4 Performance of the algorithm

4.4.1 Truth matching results

To know how closely the reconstruction match with the truth information, reconstructed tau η, ϕ and E_t values were compared with that of the truth taus. Truth taus and reconstructed tau are compared in pairs in Figure 4.12a and 4.12b. In each event, there are many reconstructed taus while there are only 2 truth taus. Here, only one truth tau and one reconstructed tau are selected from each event. the truth tau is chosen to be the one that has larger pt among the 2 truth taus per event. And the truth matching was done so that the reconstructed tau plotted here is the one that is closest to the lead Ttruth tau (when the ΔR is the smallest).

As shown by Figure 4.12a and Figure 4.12b, the reconstructed tau positions (or the seed locations) have reasonably good agreement with the truth tau positions.

Then Figure 4.12c shows how the reconstructed tau E_{T}'s are compared to the visible part (subtract neutrino pt) truth tau E_{T}. The reconstructed E_{T} are always smaller than that from the truth information. So there could be underestimation
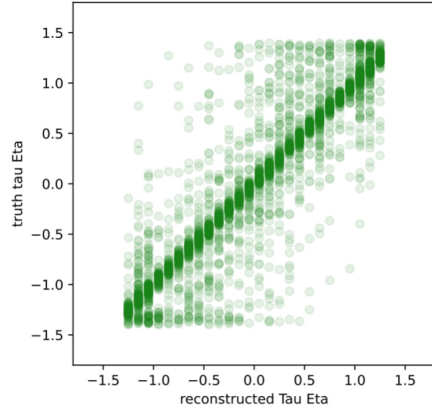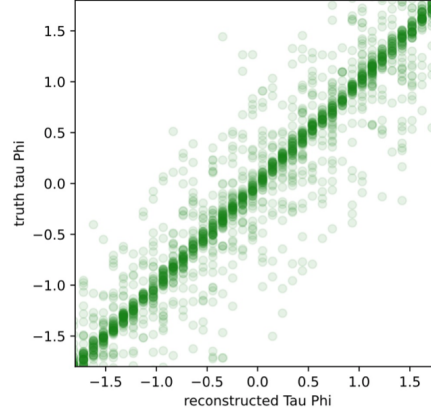
(a) Compare   truth   tau   and (b) Compare   truth   tau   and
reconstructed tau ϕ
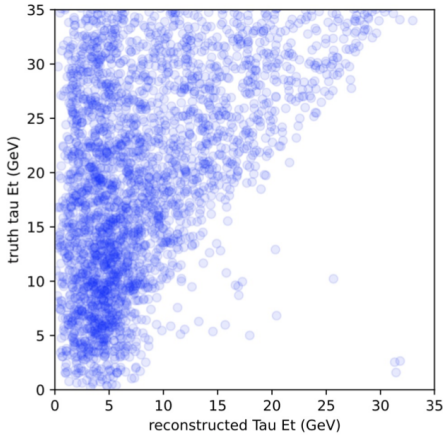
(c)   compare   truth   tau   and
reconstructed tau ϕ

Figure 4.12. Transverse energy distribution of the trigger towers and that of the seeds for the $Z\tau\tau$ sample.

with the reconstruction process. One cause could come from not including all the calorimeter layers available. The other cause could be the loss of energy of taus before its decay products hit the detector material. More investigations can be done to understand the cause better.

4.4.11   FCore distributions of the Ztautau and the dijet sample. f_{Core} is intended to distinguish taus from quark jets, so it is important to study the f_{Core}

distributions of the $Z\tau\tau$ sample and the dijet background sample and how does the selection on E_T^{reco} influence the f_{Core} distributions.

(a) isolation of $Z\tau\tau$ (b) isolation of Dijet JZ0W
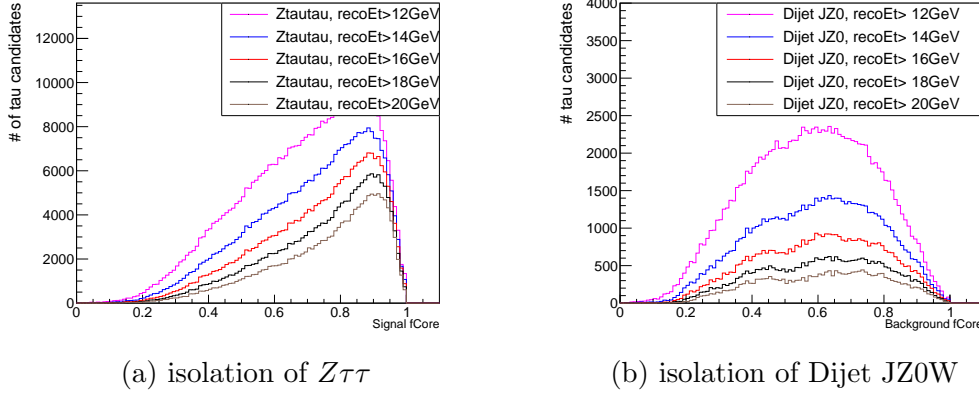
Figure 4.13. Distribution of the isolation variable with different E_T^{reco} requirements.

Figure 4.13a and Figure 4.13b are the f_{Core} variable dsitributions for the signal and background sample. f_{Core} for signal clearly peak toward larger values and the background sample have smaller f_{Core} values. This trend is consistent with the theory mentioned before.

Figure 4.14a shows the raw f_{Core} distributions for signal and background samples Figure 4.14b and the following figures show the distributions with various E_T^{reco} requirements. Overall, the $Z\tau\tau$ events have larger f_{Core} values compared to that of the dijet JZ0W events. This is expected because taus result in more concentrated shower profile on the calorimeters. Without any E_T^{reco} requirements, Figure 4.14a shows visible f_{Core} difference between the signal and the background events. The Figure 4.14b and the following figures show that the distinction between the signal and background events is bigger after adding a selection criteria on E_T^{reco} . Overall, the figures show that using the f_{Core} as a trigger threshold gives discrimination power between signal and background events and the f_{Core} variable is useful to reduces the amount of background events.

4.4.2   Trigger efficiency and background rate

Calculating trigger efficiencies and background rate reduction are some of the most important ways of judging the performance of the algorithm. This section shows the relevant plots using full statistics available, which includes around 1 million signal events and 0.6 million of background events.

The trigger efficiency in this work is calculated on event level. It is defined as the ratio of the number of triggered events to the number of total events.To be more specific, if an event has at least one tau candidate that passes the E_T^reco and f_Core requirement, this event is counted as a triggered event. The trigger efficiency with various different combinations of E_T^reco and f_Core cut are shown on Figure 4.16b and Figure 4.15b.

Depending on the capacity of computing farm as well as the physics priority in HL-LHC era, it might be determined to trigger on tau candidates on a specific rate. Calculating the background rate at different reconEt and f_Core requirements helps determine the proper trigger thresholds to be used when running the trigger.

Because real taus will only take a very small fraction of the total number of events, majority of the events are considered "background" as seen by the tau trigger. The overall event rate can be approximated using the background rate.

The following equation approximates the rate of background events that will be triggered by the algorithm:

$$\text{Background Rate} = \frac{\int_\text{Trigger Threshold}^{\max E_T} recoE_\text{T}^\text{bkg} dx}{\int_0^{maxEt} recoE_\text{T}^\text{bkg} dx} \times 30700\,[\text{kHz}]$$

where the first factor gives the fraction of events that are triggered and the second factor turns it into a rate depending on the number of bunches in the LHC ring.

The relation between the number of bunches and the frequency of the bunches can be estimated by

$$N(\text{protonBunches}) = \frac{\text{Circumference(LHC)}}{\text{Period} \times c \times 0.99} = \text{frequency} \frac{\text{Circumference(LHC)}}{c \times 0.99}$$

Where c is the speed of light $2.9979 * 10^8 m/s$ and the circumference of the LHC is 26659 m. Frequency is what we want to use in the rate calculation, the proton bunch colliding frequency. According to a presentation in 2020 from the HiLumi WP2 meeting in Ref. hil, the baseline filling scheme for the HL-LHC is $4 \times 72b$, which leads to 2747 collisions per LHC orbit for the ATLAS experiment. But the scheme could be optimized so that there are 2808 collisions in the ring and higher in the future. 30700kHz, as used in the background rate calculation in this work, corresponds to roughly 2758 collisions per LHC orbit. Given the uncertainty of the actual filling scheme that will be used, this is a reasonable estimation compared to the numbers in Ref. hil. But this calculation can be made more accurate in the future.

Figure 4.15a and Figure 4.16a are the triggered background rates with different f_{Core} and $E_{\text{T}}^{\text{reco}}$ requirements. It is apparent that more stringent f_{Core} or $E_{\text{T}}^{\text{reco}}$ requirements result in reduction of the background rate.

Figure 4.15b and Figure 4.16b are the trigger efficiencies with either varying f_{Core} and fixed $E_{\text{T}}^{\text{reco}}$ or varying $E_{\text{T}}^{\text{reco}}$ and fixed f_{Core}. The efficiencies are plotted against the leading truth tau pt from each event. The selection conditions correspond to the ones used in the background rate plots. They show that trend that efficiency is lower with larger trigger threshold (either f_{Core} or $E_{\text{T}}^{\text{reco}}$).
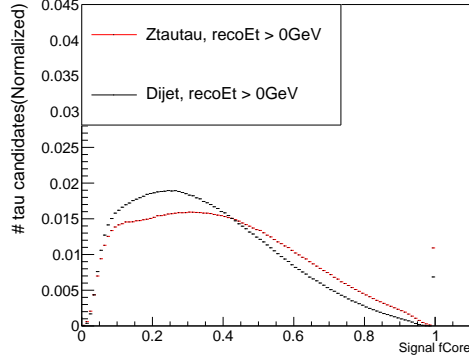
(a) No $E_{\mathrm{T}}^{\mathrm{reco}}$ requirement
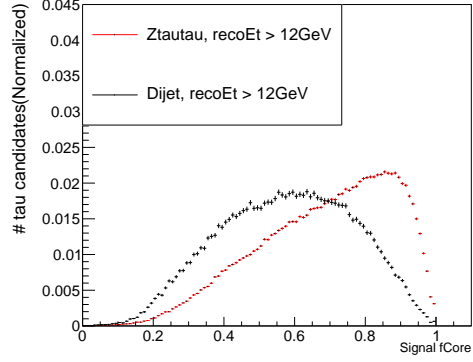
(b) $E_{\mathrm{T}}^{\mathrm{reco}} > 12\mathrm{GeV}$
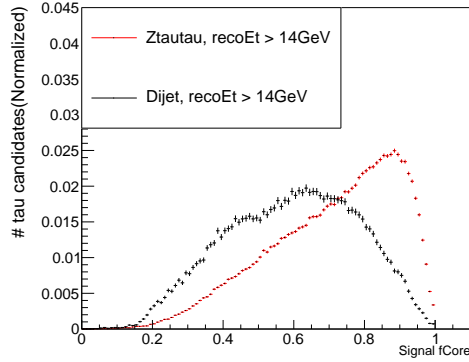
(c) $E_{\mathrm{T}}^{\mathrm{reco}} > 14\mathrm{GeV}$
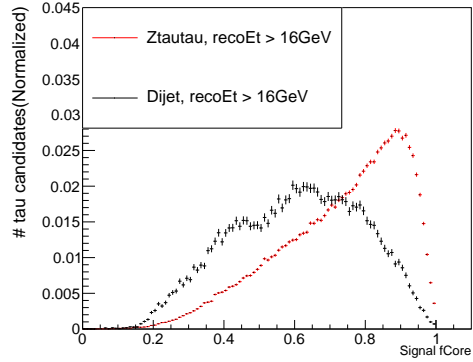
(d) $E_{\mathrm{T}}^{\mathrm{reco}} > 16\mathrm{GeV}$
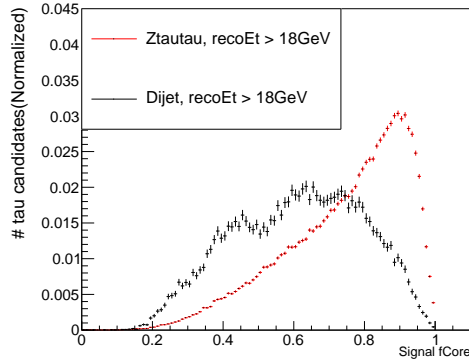
(e) $E_{\mathrm{T}}^{\mathrm{reco}} > 18\mathrm{GeV}$

Figure 4.14. Distribution of the isolation variable when adding different $E_{\mathrm{T}}^{\mathrm{reco}}$ requirements.

90
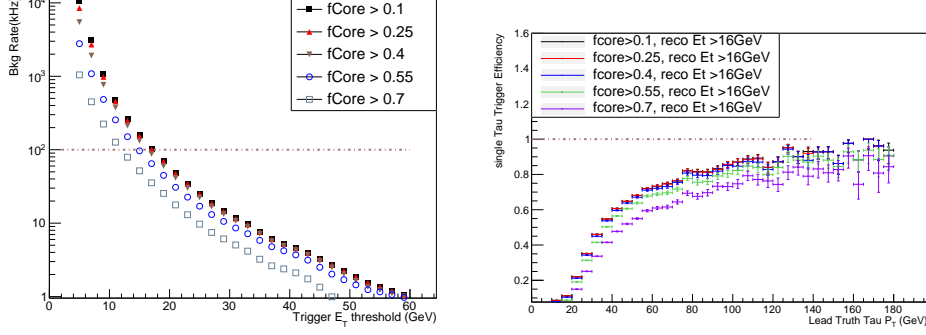
(a) Background rate with varying f_{Core} (b) Trigger efficiency with varying f_{Core}
; There is a 10GeV selection on the
truth tau T

Figure 4.15. Background rate versus threshold of trigger E_{T} with various f_{Core} requirements, no $E_{\mathrm{T}}^{\mathrm{reco}}$ selection applied beforehand (left), trigger efficiency with fixed $E_{\mathrm{T}}^{\mathrm{reco}}$ selection and different f_{Core} (right) requirement.
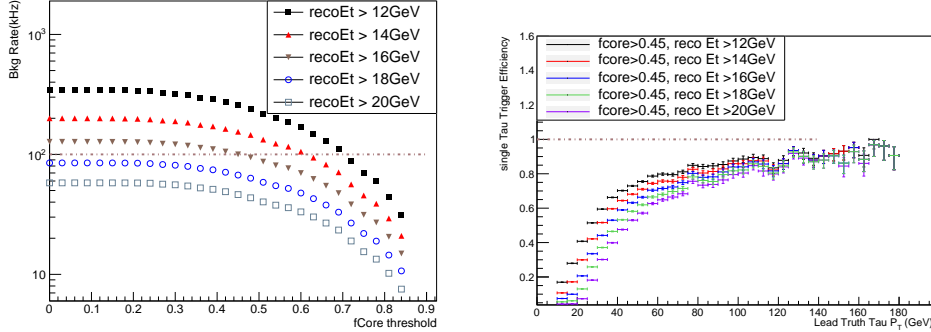
(a) Background rate with varying $E_{\mathrm{T}}^{\mathrm{reco}}$ (b) Trigger efficiency with varying $E_{\mathrm{T}}^{\mathrm{reco}}$
; There is a 10GeV selection on the
truth tau T

Figure 4.16. Background rate versus threshold of f_{Core} , no f_{Core} selection applied beforehand (left), trigger efficiency with fixed f_{Core} cut and different $E_{\mathrm{T}}^{\mathrm{reco}}$ requirement (right).

91

4.4.3 Performance study conclusion

So far, the content of the tau trigger algorithm has been articulated and calculations of the trigger rate and trigger efficiencies for various combinations of f_{Core} and E_T^{reco} have been shown. To choose the best f_{Core} and E_T^{reco} threshold for the tau trigger, let us look for a trigger rate that could be allocated to the tau trigger alone. Figure 4.17 shows the rate for a di-tau trigger in Run 3.
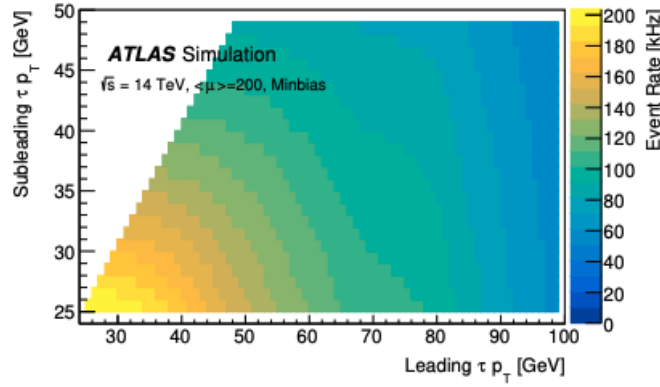
Figure 4.17.   Rate as a function of leading and sub-leading online (uncalibrated) τT for a di τ trigger implemented in the Global Trigger using a recursive neural network tda.

This plot does not provide an explicit trigger rate that will be allocated to the tau trigger, but it shows a range of rate that might be used. Before the specific rate is determined, let us use 100kHz as an estimation. As marked by red lines in Figure 4.15a and Figure 4.16a, 3 combinations of threshold will lead to a total of 100kHz rate for the tau trigger.

Table 4.2. f_{Core} and E_T^{reco} threshold that result in the tau trigger rate of 100kHz.

Options	E_T^{reco} Threshold (GeV)	f_{Core} Threshold
Option 1	16	0.45
Option 2	14	0.6
Option 3	12	0.7

To choose the best combination, efficiency curves for these 3 combinations are shown in Figure 4.18.
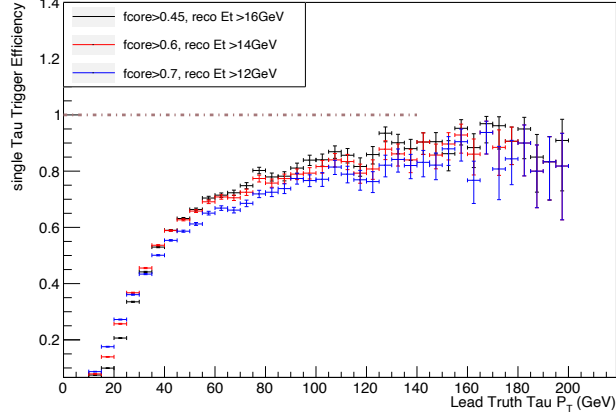
Figure 4.18. Efficiency curves with 3 options of E_T^{reco} and f_{Core} threshold (with 998600 $Z\tau\tau$ events).

The black curve (with $f_{Core} = 0.45$ and $E_T^{reco} = 16$ GeV) gives the best efficiency. However, the three options do not produce drastically different results. This makes sense because according to Figure 4.15b and Figure 4.16b, smaller thresholds (E_T^{reco} or f_{Core}) gives better efficiency. So the efficiency curve remains about the same when the f_{Core} threshold decreases and the E_T^{reco} threshold increases.

Therefore, the recommended trigger threshold from this work of tau trigger are 16GeV for E_T^{reco} and 0.45 for f_{Core} .

4.5 Future work for the performance study of the tau trigger

The 1st version of the Athena based tau trigger algorithm using only elementary cells of the calorimeter has been implemented under the run4 simulation framework. Future work needs to be done to fully study and optimize the performance of the tau trigger for HL LHC.

1. The current algorithm only includes cells from the barrel region of the LAr. So future work should expand the coverage in η and include the hadronic layers in the reconstruction of the tau candidates.

2. The di-tau final state is crucial to many ATLAS searches, so the future work should study di-tau trigger rate and efficiencies.

3. The current algorithm sets a minimum seed E_{T} requirement to be 4 GeV. A future study should explore the effect of relaxing this requirement so more seeds are allowed to proceed to the clustering stage.

4. The current algorithm is built under Athena release 21.9 which is currently out of date. Future work should migrate the framework and the algorithm so they can run under Athena release 22.

5. A different direction worth exploring is to study the performance when using topoclusters as the input to the algorithm instead of elementary cells.

4.6   Tau Trigger Firmware

The framework I used to develop the tau firmware comes from Kirk Weedman, the former engineer in our group. He developed a method to sum up cell energies in an RTL (Register Transfer Level) design based on star-shaped windows (shown in Figure 4.20) using System Verilog. "RTL" is a term used among engineers to refer to the code used by the FPGA to perform certain calculations, as opposed to a software code. Kirk's code and documentation are stored in the centralized repository for GEP here Ref. gep [a]. The main repository for all GEP related firmware is located in Ref. gep [b].
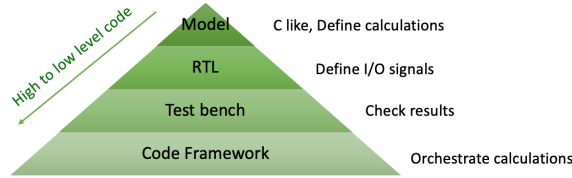
Figure 4.19. Structure of the System Verilog code for the tau firmware, arranged from high to low level code.

To simplify the development process, Kirk used the macro language feature of the System Verilog to make text substitution of the lower level code with high level code. This way, the developer only need to interact with the higher level section of the code. The code of the project is arranged as illustrated in Figure 4.19. The following are what the 4 sections do:

1. **Model file** contains high-level code that resembles C code; defines the calculations with cell data, give immediate results

2. **RTL file**: defines Input/Output signals, sequential logic on the FPGA board;

3. **Test bench file**: Performs Assertion based validation by Comparing results from the Model and the hardware simulation;

4. **Code framework**: handles timing issues and determine when to do operations on the signals
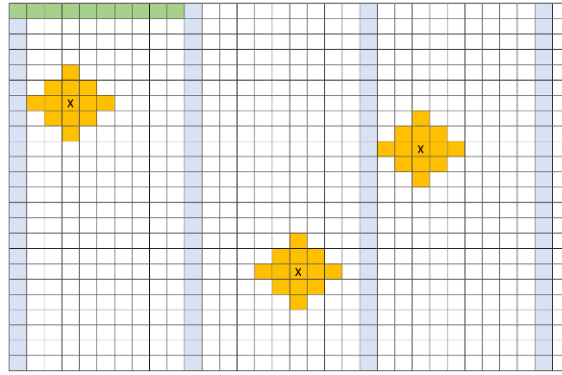
Figure 4.20. Algorithm processing unit and the data streaming method, the horizontal and the vertical axes are η and ϕ direction correspondingly.

Figure 4.20 shows how the input cell data are streamed in to the firmware. Evenly spaced cell data become available every clock cycle, as shown by the blue region. Every blue bar is followed by 10 cells , so it takes a total of 10 clock cycles to take in all the cell data. The intention is to use elementary cell data here for the tau trigger, so the reader can think of each grid as one elementary cell. But as seen by the firmware, it can run with cell data regardless of the physical granularity's of the grid.

As data come in, intermediate steps of the sum are performed as soon as the relevant cell data become available. Compared to the strategy of doing calculations only after all the cell data is available, this method significantly shortens the time it takes to complete all the calculations.

The algorithm is based on APUs (Algorithm Processing Units). An algorithm for a small slice of the cell data (in the angular space) is written, then the complete algorithm is done by repeating the APU and connecting the results returned from each slice. An APU here is represented by the region marked by green in figure 4.20. The System Verilog code is first written to produce calculation results from the 10 eta cells. To calculate the energy sums for the window, results from each APU are connected by their η and ϕ positions. Subsections of the window are then brought together to produce the energy sum for a complete pattern. In the end, the algorithm returns a 2D array of energy sums; each value in the 2D array represents a windowed sum centered on each cell.

For the tau trigger, the previous scheme was adopted but used rectangular windows instead, as shown in Figure 4.22a. Because rectangular windows are used by the Run 3 tau firmware. So far, the firmware that preforms window based sums has been implemented for both 1D ($E_T[\eta]$) and 2D cell data ($E_T[\eta][\phi]$).

4.6.1   Timing diagram of the firmware

The visual representation of the firmware is best illustrated by a timing diagram.
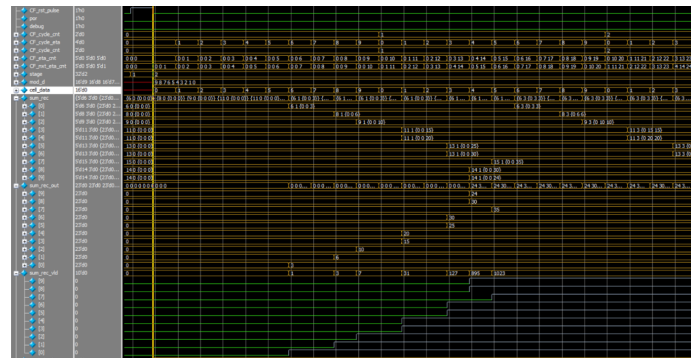
Figure 4.21. Timing diagram for the firmware implementation with a single APU.

97

Figure 4.21 is taken from the firmware implementation of a single APU (when the calculations are done on 1D cell data with only 10 cell energies). It shows timing information of the firmware that happened within 2 and half bunch crossing's. Every BC creates the input cell E_T for one event.

The horizontal axis of Figure 4.21 is time. Each bar represents a clock cycle and each event takes in 10 cell energies as input to the firmware. The vertical section in grey shows what information is plotted with respect to clock cycles. Each small bar with a number next to it shows either an input is taken in or an output is produced. The `cell data` section highlighted in white shows the cell energy taken in each clock cycle. The numbers $0-9$ indicate the cell E_T. For simplicity, the same data are fed in for each event. The following two sections (`sum rec` and `sum rec out`) show how the cell energies are being summed as time progresses. The `sum rec` gives results from the C-like model file (the "model file" in Section-4.6). The `sum rec out` gives RTL results (from the "RTL" in Section-4.6) from the hardware simulation. You can see intermediate steps of the calculation are produced while the input cell data are streamed in. The section on the bottom with many green lines performs property assertions to make sure the intermediate calculations are done correctly. Here, the results from the model file and the RTL are being compared to make sure they are matched. The firmware simulation will be stopped promptly if any one of the results does not match.

4.6.2   C simulation of the firmware

The Athena based simulation described previously produce the final results of a tau trigger algorithm but it doesn't model the firmware exactly. To quickly validate

the results produced by the firmware, a standalone C++ program was developed separately to preform exactly the same calculation as the firmware.

4.6.3   C simulation for windowed sums

To quickly confirm that the firmware gives the right results, I wrote a standalone C++ program to reproduce the firmware results (5×5 windowed sums). I used the same input data file format as the topo-clustering algorithm. The input cell energy is encoded by 10-bit binary numbers with no sign bit. All the cell data are arranged into rows and columns according to their positions in the η and ϕ space. The sums are performed as in Figure 4.22a. Because the detector is cylindrical, the sums wrap around in the ϕ direction but not in the η direction.
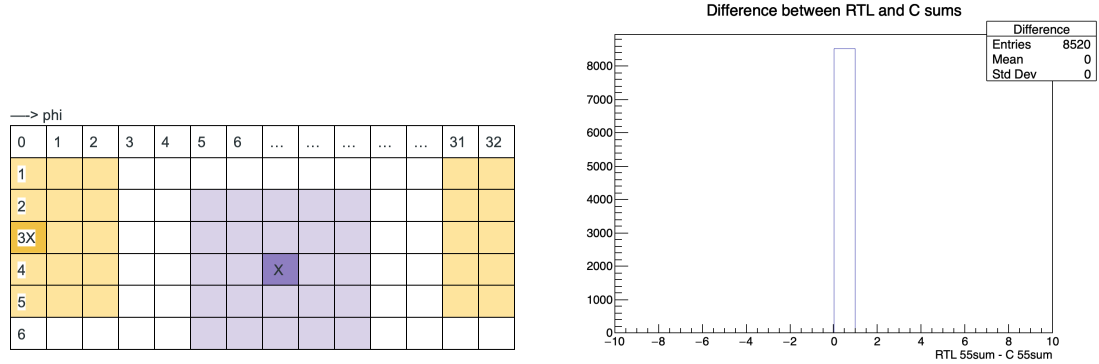
(a) Summing scheme for reconstructing tau E_T using the calorimeter cells

(b) Difference of summing results between RTL and C++ calculations

Figure 4.22.

The C++ code produced results using both 1D and 2D cell data. The output from both the C++ code and the firmware are written into text files of the same format and compared. The difference between the RTL and the C++ code sums is shown in Figure 4.22b.

The test data set is taken from simulated data of layer 2 of the LAr calorimeter. The dimension of the data set is 80 cells in the η direction and 213 cells in the ϕ

direction. There are then $80 \times 213 = 8520$ sums being calculated. The difference between these sums results in a delta function centered at 0, meaning the results on two sides match exactly. In conclusion, we have finished validating the firmware which performs a summing algorithm of the cell data. The initial design is working as expected.

CHAPTER V

THE SEARCH FOR DARK MESON

5.1 TITLE

Search for dark mesons decaying to top and bottom quarks in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

5.2 DISCLAIMER

This chapter is a previously published co-authored paper, the credit belongs to the ATLAS collaboration. See dar [2024] for a link to the publication with DOI 10.1007/jhep09(2024)005.

5.3 ABSTRACT

A search for dark mesons originating from strongly-coupled, $SU(2)$ dark flavor symmetry conserving models and decaying gaugephobically to pure Standard Model final states containing top and bottom quarks is presented. The search targets fully hadronic final states and final states with exactly one electron or muon and multiple jets. The analyzed data sample corresponds to an integrated luminosity of 140 fb^{-1} of proton–proton collisions collected at $\sqrt{s} = 13$ TeV with the ATLAS detector at the Large Hadron Collider. No significant excess over the Standard Model background expectation is observed and the results are used to set the first direct constraints on this type of model. The two-dimensional signal space of dark pion masses m_{π_D} and dark rho-meson masses m_{ρ_D} is scanned. For $m_{\pi_D}/m_{\rho_D} = 0.45$, dark pions with masses $m_{\pi_D} < 940$ GeV are excluded at the 95% CL, while for $m_{\pi_D}/m_{\rho_D} = 0.25$ masses $m_{\pi_D} < 740$ GeV are excluded.

5.4 Introduction

The Standard Model (SM) can be extended by a new strongly-coupled, confining gauge theory with fermion representation which transforms under the electroweak group. The appeal of such an extension is that dark matter can arise in the form of composite mesons or baryons of the new strongly-coupled theory. In addition, these models often exhibit an automatic accidental symmetry protecting against dark matter decay. Consequently, candidates for strongly-coupled dark matter include dark mesons, dark quarkonia-like states, glueballs and dark baryons Kribs and Neil [2016], Butterworth et al. [2022]. The search presented here targets one set of models incorporating this concept: Stealth Dark Matter Lattice Strong Dynamics (LSD) Collaboration et al. [2015]. Here, the new strongly-coupled dark sector consists of vector-like fermions that can transform under the new dark group but also interact with both the electroweak sector of the SM and the Higgs boson. The result is the emergence of a quantum chromodynamics (QCD)-like dark sector as the direct analog to the QCD meson and baryon sector. This leads to several intriguing phenomenological consequences: as long as the vector-like mass is dominant over the chiral mass, the new dark sector is only weakly constrained by precision electroweak or Higgs coupling measurements, while the Higgs interactions break the dark sector global symmetry and thus allow dark mesons to decay into pure SM states Kribs et al. [2019a]. This search focuses on the low-energy effective theories developed in Ref. Kribs et al. [2019c], which incorporate the leading interactions between dark mesons of a strongly-coupled $SU(2)$ dark flavor symmetry preserving dark sector and the SM. These models contain a stable dark scalar baryon which could account for the stable dark matter observed in cosmological measurements Lattice Strong Dynamics (LSD) Collaboration et al. [2015].
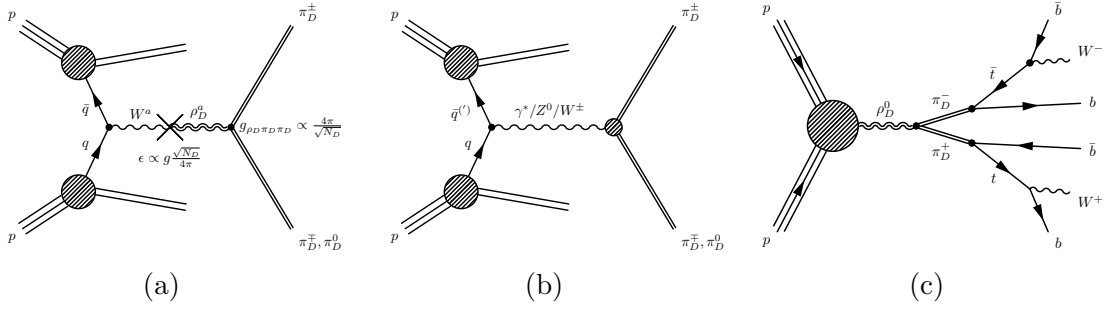
Figure 5.1. Examples of leading Feynman diagrams of dark pion pair production. The diagram in a shows resonant production via kinetic mixing with the W-field resulting in either a neutral or charged dark rho meson, a mixing with the B-field that can only result in a neutral dark rho meson is also possible, b shows Drell–Yan-type pair production of dark pions, and c shows an illustrative diagram of the dark pion decay into a top and a bottom quark for dark pion production.

The simplified model targeted in this search contains only the two phenomenologically relevant sets of dark mesons: a lighter pseudoscalar triplet of dark pions, π_D, and an additional triplet of dark rho vector mesons, ρ_D, which are both expected at a scale around or slightly above the electroweak scale. Following standard theoretical assumptions, the triplets are completely mass-degenerate and the dark sector can be fully described by three parameters: the mass of the dark pions m_{π_D}, the mass of the dark rho mesons m_{ρ_D}, and the number of dark colors N_D. Since the phenomenological consequences remain unchanged for values of N_D that are not excessively large, N_D is fixed to $N_D = 4$ throughout this search following the typical choice made for Stealth Dark Matter Lattice Strong Dynamics (LSD) Collaboration et al. [2015].

Contrary to QCD, the vector-like nature of the dark sector allows to either gauge the full $SU(2)_L$ weak interaction symmetry group or just the underlying $U(1)$ group, which leads to two distinct models of kinetic mixing of dark mesons with the SM, $SU(2)_L$ and $SU(2)_R$. The phenomenological consequences manifest themselves in the allowed decay channels and production cross-sections of dark pions, where the $SU(2)_L$ models result in considerably larger cross sections than the $SU(2)_R$ models. Dark

pions are always pair-produced either via Drell–Yan-type processes or resonantly via kinetic mixing of SM electroweak gauge bosons with the ρ_D that then subsequently decays into a pair of dark pions, as shown in Figure 5.1. The kinetic mixing parameter ϵ depends on the number of dark colors as shown in Figure 5.1a (see also Ref. Kribs et al. [2019a]). Throughout nearly all of the parameter space investigated in this search, the resonant production dominates the production of dark pions. Once the choice of $N_D = 4$ is made, the production cross-section depends trivially on the ratio of the dark pion and dark rho-meson masses, for which the symbol $\eta_D = m_{\pi_D}/m_{\rho_D}$ is used, equivalent to the η defined in Ref. Kribs et al. [2019a]. For gaugephobic $SU(2)_L$ models, a given dark pion mass and η_D-parameter fully specify the model, including the dark pion decay branching fractions.

This search considers only models with $\eta_D < 0.5$ where the decay $\rho_D^{\pm,0} \to \pi_D^{\pm}\pi_D^{0,\mp}$ has a branching fraction of nearly 1.0, while for models with $\eta_D > 0.5$ this decay is kinematically forbidden and the dark rho meson can only decay to pairs of leptons or quarks. Previous searches for resonances in the dilepton spectrum both in ATLAS ATLAS Collaboration [2017a] and CMS CMS Collaboration [2018] have placed strong constraints on such models Kribs et al. [2019a]. The bounds for models with $\eta_D < 0.5$ are considerably weaker Butterworth et al. [2022]. This is the first search in any collider experiment optimized for this specific type of model.

Figure 5.2 shows the pair-production cross-sections for dark pions in $SU(2)_L$ and $SU(2)_R$ models. The contribution of resonant production to the total production cross-section is indicated by the dashed lines. A variety of different decay channels are open to dark pions in the available parameter space. The most relevant channels and their branching fractions are shown in Figure 5.3. For gaugephobic models the
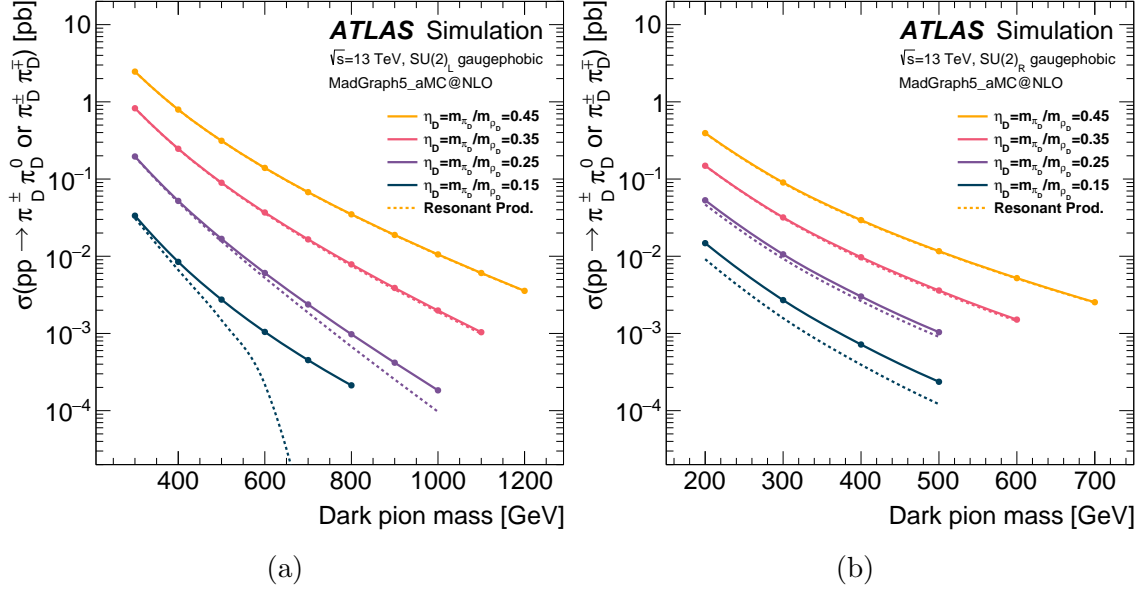
Figure 5.2. Pair-production cross-sections for dark pions as a function of dark pion mass for four different values of η_D in a an $SU(2)_L$ model and b an $SU(2)_R$ model. The dashed colored lines indicate the contribution of the resonant production mode to the total dark pion production cross-section.

decay to top and bottom quarks dominates at high masses, while decays to bottom and charm quarks, τ-leptons and gauge bosons are relevant at lower dark pion masses.

This search is the result of the analysis of 140 fb^{-1} of proton–proton (pp) collisions collected by the ATLAS detector during Run 2 of the Large Hadron Collider (LHC). Since the dark pions are pair-produced in the model considered, the experimental signatures are either three top quarks and one bottom quark ($tttb$) or two top quarks and two bottom quarks ($ttbb$)[1]. About one third of dark pions in the $SU(2)_L$ models are neutral, resulting in the $tttb$ event signature being twice as likely as the $ttbb$ signature. These processes can give rise to several different final states depending on the hadronic or semileptonic decay mode of each of the top quarks. The search is performed in the *all-hadronic* channel, targeting fully hadronic top quark decays

---

[1]The label "$tttb$" is used to indicate both $t\bar{t}t\bar{b}$ as well as its charge conjugate, $t\bar{t}\bar{t}b$; whereas "$ttbb$" refers to $t\bar{t}b\bar{b}$.
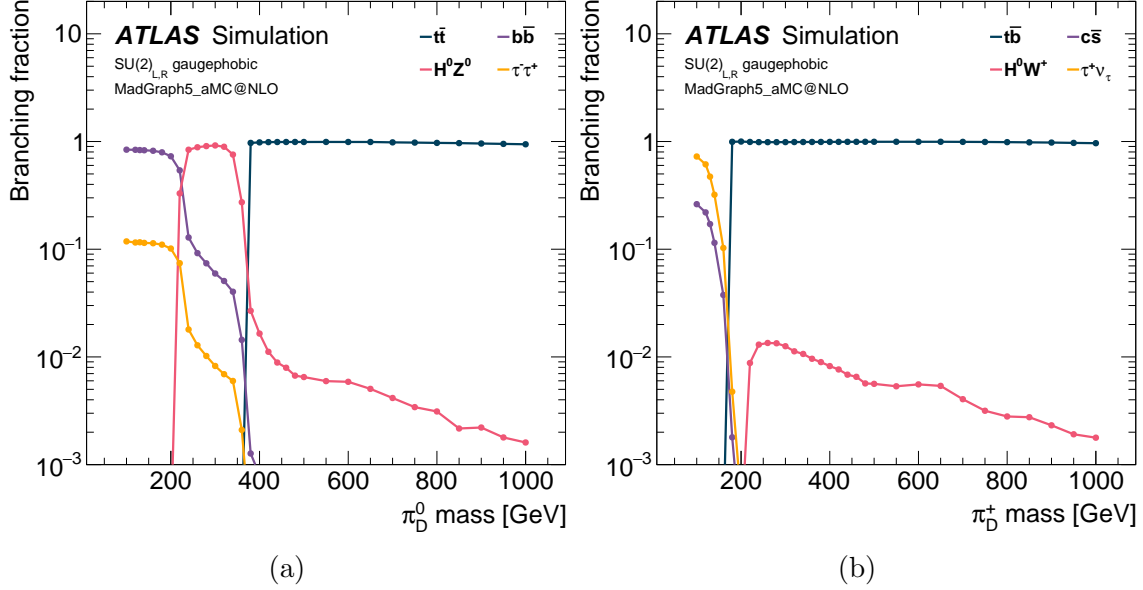
Figure 5.3. Branching fractions of the available decays of dark pions from gaugephobic $SU(2)_L$ and $SU(2)_R$ models are shown for a neutral dark pions and b positively charged dark pions. Channels with small branching fractions are suppressed for clarity.

where the signal results in eight to ten jets of which at least four originate from bottom quarks, and in the *one-lepton* channel, corresponding to final states with exactly one electron or muon in addition to up to four jets from b-quarks.

The results are interpreted as limits on the production cross-section of dark pion pairs as a function of m_{π_D} and η_D.

5.5   ATLAS DETECTOR

The ATLAS detector ATLAS Collaboration [2008] at the LHC covers nearly the entire solid angle around the collision point.[2] It consists of an inner tracking detector

---

[2]ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the center of the LHC ring, and the y-axis points upwards. Polar coordinates (r, ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln \tan(\theta/2)$ and is equal to the rapidity $y = \frac{1}{2} \ln \left(\frac{E + p_z c}{E - p_z c} \right)$ in the relativistic limit. Angular distance is measured in units of $\Delta R \equiv \sqrt{(\Delta y)^2 + (\Delta \phi)^2}$.

surrounded by a thin superconducting solenoid, electromagnetic and hadronic calorimeters, and a muon spectrometer incorporating three large superconducting air-core toroidal magnets.

The inner-detector system (ID) is immersed in a $2\,\text{T}$ axial magnetic field and provides charged-particle tracking in the range $|\eta| < 2.5$. The high-granularity silicon pixel detector covers the vertex region and typically provides four measurements per track, the first hit generally being in the insertable B-layer (IBL) installed before Run 2 ATLAS Collaboration [2010a], Abbott et al. [2018]. It is followed by the SemiConductor Tracker (SCT), which usually provides eight measurements per track. These silicon detectors are complemented by the transition radiation tracker (TRT), which enables radially extended track reconstruction up to $|\eta| = 2.0$. The TRT also provides electron identification information based on the fraction of hits (typically 30 in total) above a higher energy-deposit threshold corresponding to transition radiation.

The calorimeter system covers the pseudorapidity range $|\eta| < 4.9$. Within the region $|\eta| < 3.2$, electromagnetic calorimetry is provided by barrel and endcap high-granularity lead/liquid-argon (LAr) calorimeters, with an additional thin LAr presampler covering $|\eta| < 1.8$ to correct for energy loss in material upstream of the calorimeters. Hadronic calorimetry is provided by the steel/scintillator-tile calorimeter, segmented into three barrel structures within $|\eta| < 1.7$, and two copper/LAr hadronic endcap calorimeters. The solid angle coverage is completed with forward copper/LAr and tungsten/LAr calorimeter modules optimized for electromagnetic and hadronic energy measurements respectively.

The muon spectrometer (MS) comprises separate trigger and high-precision tracking chambers measuring the deflection of muons in a magnetic field generated

107

by the superconducting air-core toroidal magnets. The field integral of the toroids ranges between 2.0 and 6.0 T m across most of the detector. Three layers of precision chambers, each consisting of layers of monitored drift tubes, cover the region $|\eta| < 2.7$, complemented by cathode-strip chambers in the forward region, where the background is highest. The muon trigger system covers the range $|\eta| < 2.4$ with resistive-plate chambers in the barrel, and thin-gap chambers in the endcap regions.

The luminosity is measured mainly by the LUCID–2 Avoni et al. [2018] detector that records Cherenkov light produced in the quartz windows of photomultipliers located close to the beampipe.

Events are selected by the first-level trigger system implemented in custom hardware, followed by selections made by algorithms implemented in software in the high-level trigger ATLAS Collaboration [2017b]. The first-level trigger accepts events from the 40 MHz bunch crossings at a rate below 100 kHz, which the high-level trigger further reduces in order to record complete events to disk at about 1 kHz.

A software suite ATLAS Collaboration [2024a] is used in data simulation, in the reconstruction and analysis of real and simulated data, in detector operations, and in the trigger and data acquisition systems of the experiment.

5.6 Data and simulated event samples

This analysis is performed using data from pp collisions with $\sqrt{s} = 13$ TeV recorded by the ATLAS detector in 2015–2018. Only events for which all relevant subsystems were operational are considered. The data correspond to an integrated luminosity of $140.1 \pm 1.2\,\text{fb}^{-1}$ ATLAS Collaboration [2023a].

Monte Carlo (MC) simulated event samples are used for the estimate of background from SM processes and to model the targeted signal models. The

details of the event generation are provided in Sections 5.6.1 and 5.6.2 for signal and background samples, respectively. The generation of all simulated event samples includes the effect of multiple pp interactions per bunch crossing, as well as changes in detector response due to interactions in bunch crossings before or after the one containing the hard interaction, modeled by overlaying simulated inelastic events on the physics event. These two effects are referred to as pileup. The simulated event samples are processed with the GEANT4-based ATLAS detector simulation Agostinelli et al. [2003], ATLAS Collaboration [2010b]. All samples are weighted to match the pileup distribution observed in data and are processed with the same reconstruction algorithms as data.

5.6.1 Signal samples

Signal samples are generated in a grid over a two-dimensional space, varying the dark pion mass m_{π_D} between $300 - 1200$ GeV and the η_D-parameter between $0.15 - 0.45$. The matrix element calculation for the pair production of dark pions is performed at next-to-leading order (NLO) in QCD based on the model described in Ref. Kribs et al. [2019a] using MadGraph5_aMC@NLO v2.4.3 Alwall et al. [2014a] interfaced with PYTHIA.212 Sjöstrand et al. [2015] for the modeling of parton showering, hadronization and underlying event using the A14 set of tuned parameters ("tune") ATLAS Collaboration [2014] and the NNPDF2.3LO Collaboration et al. [2013] set of parton distribution functions (PDF).

The decays of bottom and charm hadrons are simulated using the EvtGen v1.6.0 program Lange [2001]. An additional set of signal samples, with parameter values near the expected exclusion contour of the all-hadronic channel, is generated at NLO in QCD using MadGraph5_aMC@NLO v2.9.9 Alwall et al. [2014a] interfaced with

PYTHIA.306 Bierlich et al. [2022] using the A14 tune and NNPDF2.3lo PDF set. Kinematic distributions match in both setups. All signal cross-sections are extracted from MadGraph5_aMC@NLO v2.9.9. The number of dark colors N_D is set to 4 for all signal points. The dark pion decays are simulated using the narrow width approximation and contain all possible decay channels from Figure 5.3.

5.6.2 Background samples

The dominant SM background process in the all-hadronic channel is multijet production. This background is estimated with data-driven methods while MC simulation is used to estimate the remaining SM processes. The background in the one-lepton channel is estimated from MC simulations and is dominated by top quark pair-production ($t\bar{t}$), often in association with heavy-flavor quarks ($t\bar{t}$+HF). Other important backgrounds are the production of a vector boson in association with jets (W/Z + jets) and single top-quark production (single top) which is dominated by the associated production of a top quark with a W boson but also contains single top-quark production in the $s-$ and $t-$channels. Smaller background contributions stem from $t\bar{t}$ produced in association with additional bosons or quarks ($t\bar{t}t\bar{t}$, $t\bar{t}$ +W/Z, $t\bar{t}H$, and other $t\bar{t}$+X) and multiboson production. The configurations used to produce the background samples are described below and are summarized in Table 5.1. For all background samples, except those generated with SHERPA, the EvtGen v.1.6.0 or v1.7.0 program is used to simulate the decays of bottom and charm hadrons.

5.6.21 $t\bar{t}$ background. The production of $t\bar{t}$ events is modeled using the PowhegBox v2 Frixione et al. [2007a], Nason [2004], Frixione et al. [2007b], Alioli et al. [2010] generator that provides matrix elements at NLO in QCD with the

Table 5.1. Overview of the configuration of all nominal background samples used in the analysis; details and definitions are provided in the text.

Process	Generator	PDF	Showering	Tune	Cross section
$t\bar{t}$	PowhegBox v2	NNPDF3.0NLO	PYTHIA	A14	NNLO+NNLL
$t\bar{t}b\bar{b}$	Powheg Box Res	NNPDF3.0NLO	PYTHIA	A14	NLO
W/Z + jets	SHERPA v2.2.11	NNPDF3.0NNLO	SHERPA	Def.	NLO
Single top	PowhegBox v2	NNPDF3.0NLO	PYTHIA	A14	NLO+NNLL
$t\bar{t}t\bar{t}$	MadGraph5_aMC@NLO v2.4.3	NNPDF3.1NLO	PYTHIA	A14	NLO
$t\bar{t}$ +W/Z	MadGraph5_aMC@NLO v2.3.3	NNPDF3.0NLO	PYTHIA	A14	NLO
$t\bar{t}H$	PowhegBox v2	NNPDF3.0NLO	PYTHIA	A14	NLO
Other $t\bar{t}$+X	MadGraph5_aMC@NLO	NNPDF2.3LO	PYTHIA	A14	NLO
Multiboson	SHERPA v2.2.1/v2.2.2	NNPDF3.0NNLO	SHERPA	Def.	NLO

NNPDF3.0NLO Collaboration et al. [2015], Carrazza et al. [2015] set PDFs and the h_{damp} parameter, which controls the matching in PowHegand effectively regulates the high-Tradiation against which the $t\bar{t}$ system recoils, set to 1.5 m_{top} ATLAS Collaboration [2016a]. The functional form of the renormalization and factorization scales are set to the default scale $\sqrt{m_{\mathrm{top}}^2 + p_{\mathrm{T}}^2}$. The events are interfaced with PYTHIA.230 for the parton shower and hadronization, using the A14 set of tuned parameters and the NNPDF2.3LO PDF set. The $t\bar{t}$ sample is normalized to the cross section prediction at next-to-next-to-leading order (NNLO) in QCD including the resummation of next-to-next-to-leading-logarithmic (NNLL) soft-gluon terms calculated using Top++2.0 Beneke et al. [2012], Cacciari et al. [2012a], Bärnreuther et al. [2012], Czakon and Mitov [2012, 2013], Czakon et al. [2013], Czakon and Mitov [2014]. For pp collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV, this cross section corresponds to $\sigma(t\bar{t})_{\mathrm{NNLO+NNLL}} = 832 \pm 51$ pb using a top-quark mass of $m_{\mathrm{top}} = 172.5$ GeV.

The inclusive $t\bar{t}$ sample described above is complemented by a dedicated sample in which a pair of top quarks is produced in association with two b-quarks. Events are simulated with the Powheg Box Res Ježo et al. [2018] generator and OpenLoops 1 Buccioni et al. [2019], Cascioli et al. [2012], Denner et al. [2017], using

a pre-release of the implementation of this process in POWHEG BOX RES provided by the authors Ježo [2019], with the NNPDF3.0NLO PDF set. It is interfaced with PYTHIA.240, using the A14 set of tuned parameters and the NNPDF2.3LO PDF set. The four-flavor scheme is used with the b-quark mass set to 4.95 GeV. The factorization scale and h_{damp} parameter are both set to $0.5 \times \Sigma_{i=t,\bar{t},b,\bar{b},j} m_{\mathrm{T},i}$, and the renormalization scale is set to $\sqrt[4]{m_{\mathrm{T}}(t) \cdot m_{\mathrm{T}}(\bar{t}) \cdot m_{\mathrm{T}}(b) \cdot m_{\mathrm{T}}(\bar{b})}$. This $t\bar{t}b\bar{b}$ sample is used in the one-lepton channel where the dominant background comes from $t\bar{t}$ production, whereas the multijet-dominated all-hadronic channel relies on the five-flavor scheme inclusive sample alone.

Previous studies have seen improved agreement between data and prediction in $t\bar{t}$ events, particularly for the top-quark Tdistribution, when comparing with NNLO calculations ATLAS Collaboration [2016b]. Top-quark pair differential calculations at NNLO QCD accuracy including electroweak (EW) corrections have become available Czakon et al. [2017]. Hence, a small improvement to the modeling is incorporated by correcting the $t\bar{t}$ and the $t\bar{t}b\bar{b}$ samples to match their top/antitop-quark Tand the top-quark mass distribution to the accuracy predicted at NNLO in QCD and NLO in EW.

Events in the $t\bar{t}$ and $t\bar{t}b\bar{b}$ samples are classified according to the flavor of the particle jets not originating from the top quark. The particle jets are reconstructed from the simulated stable particles using the anti-k_{t}algorithm Cacciari et al. [2008b, 2012b] with a radius parameter $R = 0.4$, and are required to have T > 15 GeV and $|\eta| < 2.5$. Events are labeled as $t\bar{t} + \geq 1b$ if at least one particle jet is matched within $\Delta R < 0.4$ to b-hadrons with T > 5 GeV that do not arise from the decay of top quarks. In the remaining events, if at least one particle jet is matched within $\Delta R < 0.4$ to additional c-hadrons with T > 5 GeV, the events are classified as

$t\bar{t} + {\geq}1c$. All other events are labeled as $t\bar{t} + $ light. The $t\bar{t}b\bar{b}$ sample is used for the $t\bar{t} + {\geq}1b$ category meaning that all $t\bar{t} + $ light and $t\bar{t} + {\geq}1c$ events are rejected from this sample. Likewise, only the $t\bar{t} + $ light and $t\bar{t} + {\geq}1c$ events are retained from the inclusive $t\bar{t}$ sample.

5.6.22   Other backgrounds.   The production of $W/Z + $ jetsis simulated with the SHERPA v2.2.11 Bothmann et al. [2019] generator using NLO matrix elements for up to two partons, and leading-order (LO) matrix elements for up to five partons calculated with the Comix Gleisberg and Höche [2008] and OPENLOOPS 1 libraries. They are matched with the SHERPAparton shower Schumann and Krauss [2008] using the MEPS@NLO prescription Höche et al. [2012, 2013], Catani et al. [2001], Höche et al. [2009] and the set of tuned parameters developed by the SHERPAauthors. The HESSIAN NNPDF3.0NNLO PDF set is used and the samples are normalized to a prediction that is NNLO in QCD Anastasiou et al. [2004].

The associated production of a top quark with a W bosons (tW) is modeled using the POWHEGBOX v2 Re [2011], Nason [2004], Frixione et al. [2007b], Alioli et al. [2010] generator at NLO in QCD using the five-flavor scheme and the NNPDF3.0NLO PDF set. The diagram removal scheme Frixione et al. [2008] is used to remove interference and overlap with $t\bar{t}$ production. The related uncertainty is estimated by comparing with an alternative sample generated using the diagram subtraction scheme Frixione et al. [2008], ATLAS Collaboration [2016a]. Single top-quark t-channel production is modeled using the POWHEGBOX v2 Frederix et al. [2012], Nason [2004], Frixione et al. [2007b], Alioli et al. [2010] generator at NLO in QCD using the four-flavor scheme and the corresponding NNPDF3.0NLO PDF set. Single top-quark s-channel production is modeled using the POWHEGBOX v2 Alioli et al. [2009], Nason [2004], Frixione et al. [2007b], Alioli et al. [2010] generator at NLO in QCD in the five-flavor

scheme with the NNPDF3.0NLO PDF set. All single top-quark events are processed through PYTHIA.230 using the A14 tune and the NNPDF2.3LO PDF set.

The production of $t\bar{t}t\bar{t}$ events is modeled with the MADGRAPH5_aMC@NLO v2.4.3 generator which provides matrix elements at NLO in QCD with the NNPDF3.1NLO Collaboration et al. [2015] PDF set. The functional form of the renormalization and factorization scales is set to $0.25 \times \sum_i \sqrt{m_i^2 + p_{\mathrm{T},i}^2}$, where the sum runs over all the particles generated from the matrix element calculation, following Ref. Frederix et al. [2018]. Top quarks are decayed at LO using MADSPIN Frixione et al. [2007c], Artoisenet et al. [2013] to preserve all spin correlations. The events are interfaced with PYTHIA.230 for the parton shower and hadronization, using the A14 set of tuned parameters and the NNPDF2.3LO PDF set.

The production of $t\bar{t}+W/Z$ events is modeled with the MADGRAPH5_aMC@NLO v2.3.3 generator at NLO in QCD with the NNPDF3.0NLO PDF set. The events are interfaced to PYTHIA.210 using the A14 tune and the NNPDF2.3LO PDF set.

The production of $t\bar{t}H$ events is modeled using the POWHEGBOX v2 Frixione et al. [2007a], Nason [2004], Frixione et al. [2007b], Alioli et al. [2010], Hartanto et al. [2015] generator at NLO in QCD with the NNPDF3.0NLO PDF set. The events are interfaced to PYTHIA.230 using the A14 tune and the NNPDF2.3LO PDF set.

Further rare top-quark-pair backgrounds ttt, $t\bar{t}ZZ$, $t\bar{t}WW$, $t\bar{t}WZ$, $t\bar{t}WH$ and $t\bar{t}HH$ are all produced using the NLO in QCD MADGRAPH5_aMC@NLO generator interfaced with PYTHIAusing the A14 set of tuned parameters and scaled to NLO cross sections de Florian et al. [2017].

Samples of diboson final states (VV) are simulated with the SHERPA v2.2.1 or v2.2.2 Bothmann et al. [2019] generator depending on the process, including off-shell effects and Higgs-boson contributions, where appropriate. Semileptonic final

states, where one boson decays leptonically and the other hadronically, are generated using matrix elements at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. Samples for the loop-induced processes $gg \to VV$ are generated using LO-accurate matrix elements for up to one additional parton emission. The matrix element calculations are matched and merged with the SHERPAparton shower based on Catani–Seymour dipole factorization Gleisberg and Höche [2008], Schumann and Krauss [2008] using the MEPS@NLO prescription. The virtual QCD corrections are provided by the OPENLOOPS 1 library. The NNPDF3.0NNLO PDF set is used, along with the dedicated set of tuned parton-shower parameters developed by the SHERPAauthors.

5.7   OBJECT AND EVENT SELECTIONS

For each event, interaction vertices are reconstructed from ID charged particle tracks, where the tracks are required to have transverse momenta (T) greater than 500 MeV ATLAS Collaboration [2015a]. Candidates for the primary vertex are required to have at least two associated tracks. If multiple vertices are reconstructed, the vertex with the largest sum of the squares of the transverse momenta of associated tracks is taken as the primary vertex. Events that fail the primary vertex reconstruction are rejected.

Electrons are reconstructed from energy deposits in the electromagnetic calorimeter that are matched to charged-particle tracks in the ID ATLAS Collaboration [2019a]. They are identified using a likelihood-based (LH) identification which employs calorimeter and tracking information to discriminate between electrons and jets and that combines this likelihood and the likelihood of it originating from background processes into a single discriminant. Only electron candidates with

115

T > 10 GeV within $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ are considered. Electrons are required to be well isolated using criteria based on the properties of the topological clusters in the calorimeter and of ID tracks around the reconstructed electron. Further requirements of $|z_0 \sin\theta| < 0.5$ mm and $|d_0|/\sigma(d_0) < 5$ are placed on the longitudinal and transverse impact parameters to select electrons originating from the primary vertex. Electrons are further categorized as "baseline" or "signal". For the all-hadronic channel, baseline electrons are identified by the *LooseAndBLayer* likelihood-based identification working point and are required to fulfill the *Loose* isolation criteria ATLAS Collaboration [2019b,a]. Events containing a baseline electron candidate satisfying these baseline criteria are rejected. For the one-lepton channel, baseline electrons are identified with the *Medium* working point ATLAS Collaboration [2019a] and are not subject to any isolation requirement. Signal electrons are identified by the *Tight* working point and are subject to the *Tight* track-based isolation criteria ATLAS Collaboration [2019a]. The T-requirement of the signal electrons is increased to T > 28 GeV. Signal electrons constitute a subset of the baseline electrons.

Muon candidates are reconstructed by combining tracks in the MS with tracks in the ID and are subject to cut-based identification criteria which are based on the numbers of hits in the different ID and MS subsystems, and on the significance of the charge-to-momentum ratio q/p ATLAS Collaboration [2021a]. All muon candidates are required to be within the acceptance region of the ID at $|\eta| < 2.5$ and to have T > 10 GeV. Muons are required to satisfy isolation requirements based on the properties of ID tracks around the reconstructed muon ATLAS Collaboration [2021a]. Similarly to electrons, requirements on the longitudinal and transverse impact parameters, $|z_0 \sin\theta| < 0.5$ mm and $|d_0|/\sigma(d_0) < 3$, are also applied. Baseline muons are identified

in the all-hadronic channel by the *Loose* quality working point and are required to fulfill the *Loose* isolation criteria ATLAS Collaboration [2021a]. Events containing a muon candidate satisfying these baseline criteria are rejected. In the one-lepton channel, baseline muons are identified by the *Medium* quality working point ATLAS Collaboration [2021a] and are not subject to any isolation criteria. For the selection of signal muons the *Medium* quality working point is applied and the muon candidates are required to fulfill the *Tight* isolation criteria based on the $p_T^{\mathrm{varcone30}}$ variable defined in Ref. ATLAS Collaboration [2021a] and have $p_{\mathrm{T}} > 28$ GeV. Signal muons constitute a subset of the baseline muons. Events are selected for the one-lepton channel if they contain exactly one signal and no additional baseline leptons (electrons or muons).

Jet candidates are reconstructed using a particle-flow reconstruction algorithm ATLAS Collaboration [2017c] combining charged particle tracks from the ID and three-dimensional topological energy clusters ATLAS Collaboration [2017d] in the calorimeter. Jets are reconstructed using the anti-k_{t} algorithm Cacciari et al. [2008b] implemented in the FastJet package Cacciari et al. [2012b] with a fixed radius parameter $R = 0.4$ using charged constituents associated with the primary vertex and neutral particle flow objects as inputs. In order to minimize the contribution from pileup jets, a requirement on the jet-vertex tagger ATLAS Collaboration [2016c] is made for jets with Tbelow 60 GeV.

Slight differences in the efficiency of the association of jets to vertices in data and simulation are addressed by applying scale factors to simulation. Jet energy scale corrections, derived from MC simulation, are used to calibrate the average energies of jet candidates to the scale of their constituent particles ATLAS Collaboration [2021b]. Remaining differences between data and simulated events are evaluated and corrected for using in situ techniques, which exploit the transverse momentum

balance between a jet and a reference object such as a photon, Z boson, or multijet system in data. After these calibrations, all jets in the event with $p_T > 20$ GeV must satisfy a set of loose jet-quality requirements ATLAS Collaboration [2013] designed to reject jets originating from sporadic bursts of detector noise, large coherent noise or isolated pathological cells in the calorimeter system, hardware issues, beam-induced background or cosmic-ray muons ATLAS Collaboration [2015b]. In the one-lepton channel, the jets are required to satisfy $|\eta| < 2.5$, while in the all-hadronic channel, they are required to satisfy $|\eta| < 2.8$ to match the η range of the H_T trigger. If these jet requirements are not met, the jet is discarded.

Jets are tagged as containing a b-hadron (b-tagged) by a deep neural network algorithm trained on a simulated hybrid sample composed of $t\bar{t}$ and $Z' \rightarrow t\bar{t}$ events ATLAS Collaboration [2023b, 2017e] at a working point corresponding to a 77% b-jet efficiency, as measured on an inclusive $t\bar{t}$ sample. This working point has a rejection factor of 5 and 170 on charm and light-flavored jets, respectively. Flavor-tagging efficiency differences between data and simulation are corrected by a reweighting procedure detailed in Refs. ATLAS Collaboration [2019c, 2023c, 2022a].

To resolve any reconstruction ambiguities between electrons, muons and jets, an overlap removal procedure is applied in a prioritized sequence, based on baseline leptons and jets, as follows. First, if a electron shares the same ID track with another electron, the electron with lower p_T is discarded, and then any electron sharing the same ID track with a muon is rejected. Next, jets are rejected if they lie within $\Delta R = 0.2$ of an electron. Similarly, jets within $\Delta R = 0.2$ of a muon are rejected if the jet has fewer than three associated tracks or if the muon is matched to the jet through ghost association Cacciari et al. [2008b]. Finally, electrons that are close to

a remaining jet are discarded if their distance from the jet is $\Delta R < 0.4$, while for muons the distance is $\Delta R < \min(0.4, 0.04 + 10 \ \text{GeV}/\text{T})$.

Large-R jets are reclustered from the calibrated $R = 0.4$ jets described above using the anti-k_t algorithm with a fixed radius parameter of $R = 1.2$ Nachman et al. [2015]. The large-R jets aim to fully contain the dark pion decay products and the R parameter is optimized for the dark pion mass range of the targeted signal points. For the one-lepton channel the signal lepton is added to the $R = 0.4$ jet collection before the reclustering, which then proceeds in the same way as in the all-hadronic channel. After reclustering, the large-R jet containing the lepton is referred to as J^{lep} and the leading fully hadronic large-R jet J^{had}. Both large-R jets originate from the same reclustering procedure, ensuring there is no overlap between the two.

Events are selected for the all-hadronic channel using triggers on H_{T}, defined as the scalar sum of the transverse momenta of all the reconstructed jets in the event with $|\eta_{\text{jet}}| < 2.8$ ATLAS Collaboration [2016d]. The H_{T}-trigger threshold was 850 GeV in 2015 and the first half of 2016, and was increased to 1000 GeV in the latter half of 2016 for the remainder of Run 2. Since the trigger decision is based on H_{T} computed from trigger-level jet momenta (which lack a detailed calibration), the triggers show a slow onset behavior with respect to H_{T} computed from jet momenta of fully calibrated jets. A requirement of $H_{\text{T}} > 1150$ GeV ensures that the trigger is fully efficient to minimize systematic uncertainties resulting from imprecise modeling of the onset behavior in simulation. The H_{T} variable computed from fully calibrated jets is used for the remainder of this search.

Events are selected for the one-lepton channel using a combination of single-lepton triggers ATLAS Collaboration [2020a,b,c, 2022b]. The single-lepton triggers require the presence of a muon or an electron with Thigher than a certain threshold and,

Table 5.2. Summary of the preselection criteria for the all-hadronic and one-lepton channels, in terms of the number of baseline and signal leptons, $R = 0.4$ and $R = 1.2$ jets, number of b-jets, and H_T. The definitions of the physics objects for the two channels are given in the text. Signal leptons constitute a subset of the baseline leptons.

Variable	All-hadronic channel	One-lepton channel
N_lep(baseline)	0	1
N_lep(signal)	-	1
$N_\mathrm{jets}(R = 0.4)$	≥ 6	≥ 5
$N_\mathrm{jets}(R = 1.2)$	≥ 2	-
$N_\mathrm{b-jets}$	≥ 3	≥ 3
H_T	≥ 1150 GeV	≥ 300 GeV

in some cases, impose identification and lepton-isolation requirements. The lowest Tthreshold was 24 (20) GeVfor electrons (muons) during the 2015 data-taking period and 26 GeVfor both electrons and muons in the data-taking periods from 2016 to 2018. The efficiencies of the single-lepton triggers range between 20% and 50% in the simulated signal samples. To account for small differences in the single-lepton trigger efficiency between data and simulation, all triggered simulated events receive an event weight to match data.

The analysis strategy relies on the reconstruction of each dark pion using a large-R jet. In the all-hadronic channel, events are required to have six or more $R = 0.4$ jets with T > 25 GeV, in addition to the $H_\mathrm{T} > 1150$ GeV requirement and the lepton veto. At least three jets within the ID acceptance ($|\eta| < 2.5$) must be b-tagged. At least two large-R jets with jet mass $m_\mathrm{jet,R=1.2} > 190$ GeV are required in all events. Events for the one-lepton channel are required to have at least five jets, out of which at least three have to be b-tagged, and to have $H_\mathrm{T} > 300$ GeV. Here the H_T is defined similar to the all-hadronic channel, however only jets with $|\eta_\mathrm{jet}| < 2.5$ are considered. Events passing all selection requirements listed here are considered preselected for the analysis. The preselection requirements are summarized for both the analysis channels in Table 5.2.

5.8 Analysis strategy

Preselected events are separated into signal, control, and validation regions based on the properties of the large-R and small-R jets in the event, as well as the signal lepton for the one-lepton channel. Signal regions (SRs) are signal-enriched regions while the control regions (CRs) are used to estimate the SM background contributions. The validation regions (VRs) are used to validate the background estimation methods. The analysis strategies for the all-hadronic and one-lepton channels are detailed below.

5.8.1 All-hadronic channel

In the all-hadronic channel, the leading two large-R jets define the overall SR, where the leading large-R jet satisfies $m_{\mathrm{jet,R=1.2}} > 300$ GeV and the sub-leading large-R jet satisfies $m_{\mathrm{jet,R=1.2}} > 250$ GeV as shown in Figure 5.4 for distributions after preselection. To suppress multijet events containing, for example, gluon to $b\bar{b}$ splitting, a selection on $m_{bb}/p_{T,bb}$, defined as the ratio of the mass to the transverse momentum of the pair of b-tagged jets closest to the center of the large-R jet, is applied to both large-R jets. In signal events $m_{bb}/p_{T,bb}$ is expected to take on larger values than in background events, thus a cut of $m_{bb}/p_{T,bb} > 0.25$ is required. Further, both large-R jets must satisfy a bb_i tag, where the ΔR between the leading ($i = 1$) or sub-leading ($i = 2$) large-R jet and the second closest b-tagged jet, defined as $\Delta R\,(j, b_2)$, is less than 1.0 and thus both b-tagged jets are well contained within the volume of the large-R jet. This variable is designed to suppress $t\bar{t}$ events where the second closest b-jet arises from the other top quark in the event and can thus have a large ΔR with the large-R jet. In signal, on the other hand, the decay products of the dark pion always include two b-quarks no matter whether the $tttb$ or $ttbb$ final

121
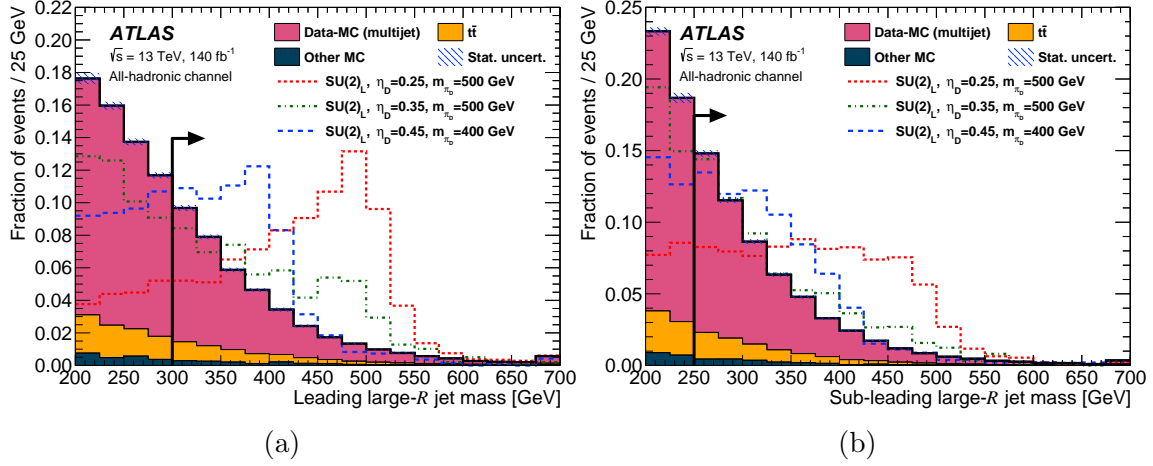
Figure 5.4. Mass of the a leading and b sub-leading large-R jet for all simulated backgrounds overlaid with three example distributions for various signal points after preselection in the all-hadronic channel. Also shown is a simplified data-driven estimate of the multijet background which was created by taking the event yields for data and subtracting all simulated backgrounds from it. Statistical uncertainties stemming from MC are indicated by the shaded region. The SR is to the right of the vertical line in both the subfigures. Individual SR bins select sub-regions of leading and sub-leading large-R jet mass for improved background discrimination. The last bin contains the overflow.

state is considered and the ΔR therefore tends to be small. The overall SR is then subdivided into nine separate bins in the leading versus sub-leading large-R jet mass plane. A large-R jet is considered $\pi_{D,i}$ tagged if its mass falls into one of the nine mass bins. The SR requires both large-R jets to satisfy both tagging selections (i.e. both jets must be bb_i and $\pi_{D,i}$ tagged). The events where the two leading large-R jets satisfy only one or two out of the four possible tags form the CRs used for the data-driven multijet extrapolation; events that satisfy three tags allow for a validation of the method and thus form the VRs. The SR selection criteria are summarized in Table 5.3.

Table 5.3. Summary of selection criteria for the SR ("Tag selection"). Nine bins are defined in the leading large-R jet vs. sub-leading large-R jet mass plane. The inverted selection ("Anti-tag selection") is also defined for use in the data-driven multijet extrapolation described in Section 5.9.1.

	Tag	Variable	Tag selection	Anti-tag selection
Both large-R jets		$m_{bb}/p_{T,bb}$	> 0.25	> 0.25
Leading large-R jet	bb_1	$\Delta R\,(j, b_2)$	< 1.0	≥ 1.0
Sub-leading large-R jet	bb_2	$\Delta R\,(j, b_2)$	< 1.0	≥ 1.0
Leading large-R jet	$\pi_{D,1}$	$m_{\mathrm{jet,R=1.2}}$	[$300 - 325$ GeV, $325 - 400$ GeV, > 400 GeV]	≤ 300 GeV
Sub-leading large-R jet	$\pi_{D,2}$	$m_{\mathrm{jet,R=1.2}}$	[$250 - 300$ GeV, $300 - 350$ GeV, > 350 GeV]	≤ 250 GeV

5.8.2 One-lepton channel

Events satisfying the preselection requirements for the one-lepton channel are categorized into SRs, CRs and VRs based on two kinematic variables defined in terms of the properties of the small-R and large-R jets. The first variable, $\Delta R(l, b_2)$, is defined as the angle between the lepton in the event and the second closest b-jet to this lepton and aims to suppress $t\bar{t}$ background similar to $\Delta R\,(j, b_2)$ in the all-hadronic channel. The kinematics of the second b-jet as distinguishing characteristic of signal events is also utilized for the second variable, $m_{bb_{\Delta R_{min}}}$, defined as the invariant mass of the two b-jets in the event that are closest to each other. This is effective for discriminating high dark pion mass signal points against background in which the two b-quarks closest to each other come e.g. from gluon splitting. The distributions of these variables in signal and background MC simulations are shown in Figure 5.5 for preselected events.
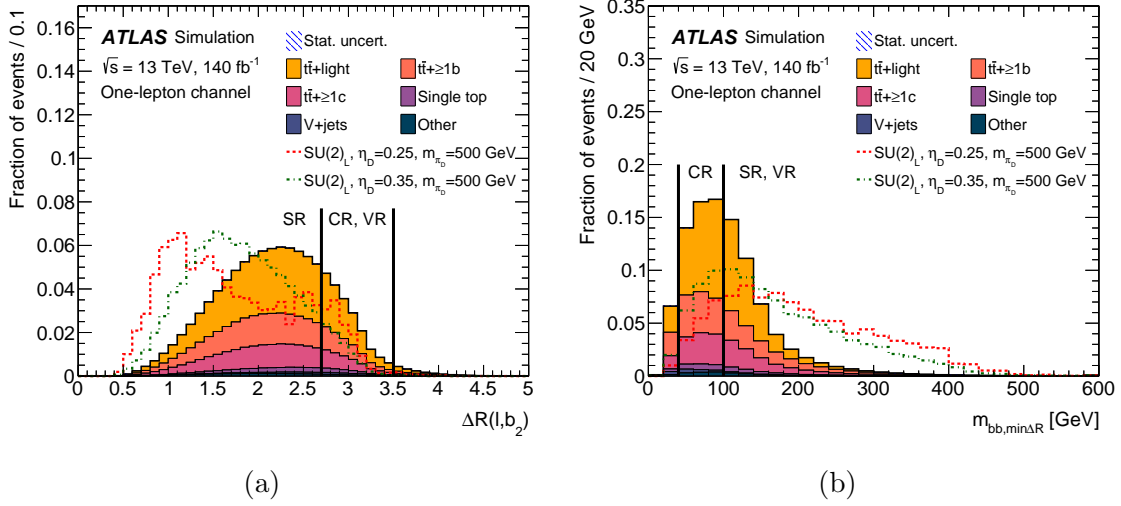
Figure 5.5. Normalized distributions of a $\Delta R(l, b_2)$ and b $m_{bb_{\Delta R_{min}}}$ for all simulated backgrounds with two example signal distributions overlaid after the one-lepton channel preselection. Statistical uncertainties stemming from MC are indicated by a shaded region, but are not visible on the scale of the y-axis. The vertical dashed lines indicate the selection requirements applied to events in the SR, CR and VR, as indicated by the labels.

The SR is defined by requiring $\Delta R(l, b_2) < 2.7$ and $m_{bb_{\Delta R_{min}}} > 100$ GeV. A CR for the $t\bar{t}$ +HF background is defined by the requirements $2.7 < \Delta R(l, b_2) < 3.5$, and 40 GeV $< m_{bb_{\Delta R_{min}}} < 100$ GeV, thus ensuring orthogonality to the SR. This region is used to correct for mismodeling in $t\bar{t}$ +HF events and has a background composition similar to that in the SR. Typical signal contamination in the CR, from signal points that are not already excluded through re-interpretation of other collider searches Butterworth et al. [2022], is below 1%. The $t\bar{t}$ +HF background estimate is validated in a VR defined by the requirements $2.7 < \Delta R(l, b_2) < 3.5$, and $m_{bb_{\Delta R_{min}}} > 100$ GeV, making it orthogonal to both the SR and the CR while also exhibiting a background composition similar to that of the SR and CR.

The statistical analysis in the one-lepton channel relies on a profile-likelihood fit to the distribution of the sum of the masses of the reclustered jets, described in

124
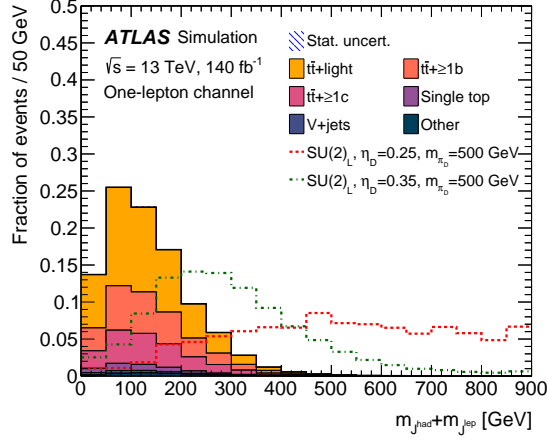
Figure 5.6. Normalized distributions of $m_{\mathbb{J}had} + m_{\mathbb{J}lep}$ for all simulated backgrounds overlaid with two example distributions for various signal points after the one-lepton channel preselection. Statistical uncertainties stemming from MC are indicated by a shaded region, but are not visible on the scale of the y-axis.

Section 5.7, $m_{\mathbb{J}had} + m_{\mathbb{J}lep}$. This discriminating variable is shown in Figure 5.6 for preselected events. For the fit, all events are further classified into regions split into bins based on the number of jets and b-jets in the event. Six bins are defined labeled XR_5j3b, XR_5j4b, XR_6j3b, XR_6j4b, XR_7j3b and XR_7j4b, where the number before the 'j' indicates the number of jets, the number before the 'b' the number of b-jets and 'X' can take on the values 'S' for an SR, 'C' for a CR and 'V' for a VR bin. In all cases, the highest jet or b-jet multiplicity is inclusive, e.g. the region CR_7j4b is a CR that contains events with ≥ 7 jets and ≥ 4 b-jets.

5.9 Background estimate

5.9.1 All-hadronic channel

The dominant background for the analysis in the all-hadronic channel originates from multijet events and constitutes about $75\% - 85\%$ of the total background in each SR bin. A data-driven method is used to estimate this background, while

125

Leading large-R jet

Sub-leading large-R jet	$\bar{\pi}_{D,1}\bar{bb}_1$	$\bar{\pi}_{D,1}bb_1$	$\pi_{D,1}\bar{bb}_1$	$\pi_{D,1}bb_1$
$\pi_{D,2}bb_2$	J	K	L	S
$\pi_{D,2}\bar{bb}_2$	B	D	H	N
$\bar{\pi}_{D,2}bb_2$	E	F	G	M
$\bar{\pi}_{D,2}\bar{bb}_2$	A	C	I	O

Figure 5.7. Region labels for the 16 regions used in the data-driven multijet estimate in the all-hadronic channel. Region S labels the SR, regions B, C, E and I are used for the ABCD extrapolation, regions D, F, G, H, J and O are used to compute correlation correction factors, and regions K, L, M and N are validation regions. The background estimate is performed independently for all nine SR bins.

MC simulation is used to account for the remaining sub-dominant SM processes as described in Section 5.6. Typical ABCD multijet extrapolations are based on two discriminating variables. Here, however, an extended ABCD method is employed that relies on four instead of two discriminating variables, which allows the correction of correlations between pairs of discriminating variables and provide validation regions close to the SR selection.

The multijet background is estimated by extrapolating from regions with small leading and sub-leading $m_{\text{jet,R}=1.2}$ and large $\Delta R\left(j, b_2\right)$ to SR bins with large leading and sub-leading $m_{\text{jet,R}=1.2}$ and small $\Delta R\left(j, b_2\right)$. The method is similar to the one detailed in Ref. ATLAS Collaboration [2018]. To this end, two additional anti-tags denoted by a slashed tag label (with orthogonal selections to the already described bb_i and $\pi_{D,i}$ tags) are defined. The tag $\bar{bb}_i$ inverts the bb_i selection, while $\bar{\pi}_{D,i}$ places upper requirements on the large-R jet mass, as summarized in Table 5.3. The combinations of possible tags and anti-tags in an event result in 16 separate regions shown in Figure 5.7. The extended ABCD method extrapolates from regions with one tag to each of the nine SR bins. Two-tag regions are used to determine correlation correction factors and three-tag regions are used for validation of the multijet estimate.

126

The concept of the extended ABCD method is detailed here by explicitly stating the computations for one VR, however, analogous derivations have to be carried out for all SRs and VRs. Considering only region K in Figure 5.7, a 2-variable ABCD estimate for this region would be computed in the standard way through $\hat{K} = \frac{J \cdot D}{B}$, where J, D and B are the number of data events minus the number of simulated events in the respective regions. If $\pi_{D,1}$ and $\pi_{D,2}$ are uncorrelated, $\hat{K}$ would be a valid estimate. However, if $\pi_{D,1}$ and $\pi_{D,2}$ are correlated, then $\hat{K}$ needs to be corrected by a correlation factor (k-factor). As long as there is no significant additional three-tag correlation with the bb_2 tag, this correlation factor, $k_{\pi_{D,1},\pi_{D,2}}$, can be measured from $\frac{F \cdot A}{C \cdot E}$ since $\frac{K \cdot B}{J \cdot D} = \frac{F \cdot A}{C \cdot E}$. Thus,

$$\hat{K} = \frac{J \cdot D}{B} \cdot \frac{F \cdot A}{C \cdot E} = \frac{J \cdot D}{B} \cdot k_{\pi_{D,1},\pi_{D,2}} \ . \tag{5.1}$$

One can also consider the estimate of region K using a three-variable ABCD estimate computed with two correlation correction factors:

$$\hat{K} = \frac{J \cdot C}{A} \cdot k_{\pi_{D,1},\pi_{D,2}} \cdot k_{\pi_{D,1},bb_2} \ , \tag{5.2}$$

where $k_{\pi_{D,1},bb_2} = \frac{D \cdot A}{B \cdot C}$. Substituting this correlation factor into Eq. (5.2) yields once again Eq. (5.1). All other k-factors can be defined according to the same principle as $k_{\pi_{D,1},\pi_{D,2}}$ and $k_{\pi_{D,1},bb_2}$.

The final multijet background estimate requires a four-variable ABCD estimate $\hat{S}'$ that is computed from data minus event counts from MC simulated backgrounds in the regions B, C, E, and I where exactly one tag is applied and region A with no

applied tags according to

$$\hat{S}' = \frac{B \cdot C \cdot E \cdot I}{A^3} \ . \tag{5.3}$$

This estimate is then multiplied by six k-factors to correct for correlations between tags,

$$\hat{S} = \hat{S}' \cdot k_{\pi_{D,1}, bb_1} \cdot k_{\pi_{D,2}, bb_2} \cdot k_{\pi_{D,1}, bb_2} \cdot k_{\pi_{D,2}, bb_1} \cdot k_{\pi_{D,1}, \pi_{D,2}} \cdot k_{bb_1, bb_2} \ . \tag{5.4}$$

If the selection criteria defined by these tags are independent from each other then the expectation value of the corresponding k-factor will be 1. Correlation factors around 1.5 are observed between the $\pi_{D,1}$ and bb_1 tags as well as the $\pi_{D,2}$ and bb_2 tags. The bb_1 and bb_2 tags are highly correlated with a k-factor of 0.1 due to the preselection requirement of three b-tagged jets. This could be mitigated by requiring four b-tagged jets, however, this would result in low statistics and high signal contamination for the four-variable ABCD extrapolation regions.

Typical signal contaminations in the extrapolation regions are less than 5% and small compared to the uncertainty applied to the multijet background discussed in Section 5.10. The method is validated through the closure of the estimate in the four 3-tag validation regions K, L, M and N. Figure 5.8 compares data to estimated background yields for each validation region in each bin of the SR. If significant 3-tag correlations occurred, a discrepancy between multijet estimate and data yields should be visible in the validation regions. However, all data yields are compatible with the background estimate within the uncertainty bands for all validation regions. The uncertainty includes a non-closure systematic uncertainty, which by design covers non-linear correlations and the impact of multijet estimation regions with low statistics.
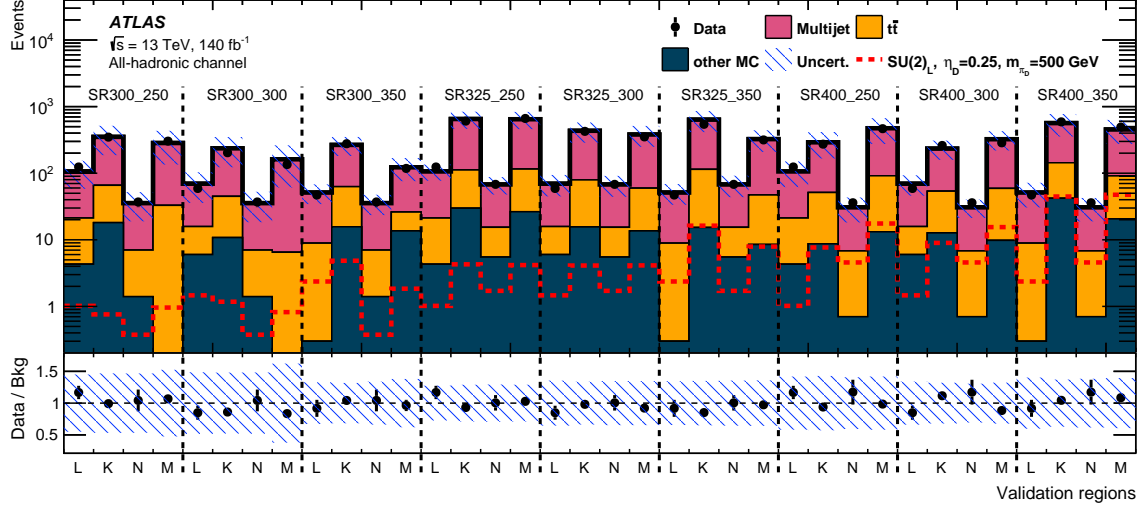
Figure 5.8. The four all-hadronic channel validation regions K, L, M and N from Figure 5.7 for each of the nine SR mass bins with the following naming convention: the leading large-R jet lower mass boundary in GeVis followed by the sub-leading large-R jet lower mass boundary. The shaded region indicates the uncertainty on the background estimate in each bin that includes the statistical uncertainties from the limited sample sizes in data and simulation as well as a multijet non-closure systematic uncertainty, as detailed in Section 5.10.

It ranges between 33% and 57% and is derived and discussed in Section 5.10. The procedure was further validated by performing signal injection tests as well as stability tests over time, under variation of the selection criteria and with scaled simulated background contributions. In all cases the resulting background estimates where stable and thus consistent with the nominal estimate.

5.9.2 One-lepton channel

The background in the one-lepton channel is estimated from MC simulations using the samples described in Section 5.6 and is dominated by $t\bar{t}$ production in association with light-flavor or heavy-flavor quarks. This background is estimated by using the inclusive $t\bar{t}$ sample complemented by the dedicated $t\bar{t}b\bar{b}$ sample, with events categorized into $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$ and $t\bar{t} + $ light as described in Section 5.6.

The $t\bar{t}$ +HF background that populates the SR is known from previous studies to be underestimated by the current MC predictions ATLAS Collaboration [2019d]. This mismodeling is corrected by keeping the normalization of these backgrounds unconstrained in the profile-likelihood fit to the $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ distribution. To this end, the $t\bar{t} + \geq 1c$ and $t\bar{t} + $ light contributions are combined and two normalization factors are fit depending on whether in addition to $t\bar{t}$ any b-quarks are present in the event. In this categorization, the combination of the $t\bar{t} + \geq 1c$ and $t\bar{t} + $ light backgrounds is referred to as $t\bar{t} + 0b$. The resulting background estimates for the SR, CR and VR are presented in Section 5.11.2.

5.10 Systematic uncertainties

The predicted signal and background event yields in the SR bins are affected by various sources of systematic uncertainties stemming from instrumentation, the data-driven multijet estimation and theoretical considerations. For the all-hadronic channel, the total uncertainty is dominated by the uncertainty on the multijet background estimate. In the one-lepton channel, the theoretical sources of uncertainty on the background modeling dominate the total uncertainty.

5.10.1 Experimental uncertainties

The uncertainty in the combined 2015–2018 integrated luminosity is 0.83% ATLAS Collaboration [2023a], obtained using the LUCID-2 detector Avoni et al. [2018] for the primary luminosity measurements, complemented by measurements using the ID and calorimeters. A systematic variation that might be introduced by the reweighting of simulated samples to match the pileup profile observed in data is estimated by varying the scale factor applied to the pileup distributions. The onset of the H_{T} trigger used

in the all-hadronic channel was studied and potential effects of mismodeling were evaluated. All three of these uncertainties were found to have a negligible impact on the analysis ($< 1\%$ in the all-hadronic channel, and ranking from $< 1\%$ to a few percent depending on the signal in the one-lepton channel).

Slight performance differences between data and simulation in lepton reconstruction, identification and isolation are corrected by the application of scale factors that are estimated from tag-and-probe experiments in data and simulation ATLAS Collaboration [2019a, 2021a]. The impact of lepton momentum scale corrections is evaluated by $\pm 1\sigma$ scale variations. For resolution uncertainties the lepton energy or momentum is smeared. In total seven (twelve) separate variations for electrons (muons) are considered. The impact of lepton uncertainties is less than 1% in both channels.

The determination of the jet energy scale and resolution is done by combining information from collision data, test beam data and simulation as described in Ref. ATLAS Collaboration [2021b]. Effects from jet flavor composition, single-particle response and pileup are considered. In the one-lepton channel, 29 parameters are evaluated for scale variations, while 13 parameters are evaluated for jet p_T resolution systematic uncertainties. In the all-hadronic channel the variations are simplified since the data-driven background estimation method largely compensates yield changes by different systematic variations and causes most of the systematic uncertainties to be negligibly small. As such, 23 parameters arise from scale variations, while for jet p_T resolution systematic uncertainties 8 parameters are evaluated. The impact of these uncertainties on the final result is small, ranging between $< 1\%$ to about 4% in the all-hadronic channel, and remaining $< 10\%$ in the one-lepton channel.

Uncertainties on the corrections of b-tagging efficiency differences between data and simulation are derived from dedicated flavor-enriched subsets of the data ATLAS Collaboration [2019c]. Also considered are uncertainties due to the mis-tagging of c-jets ATLAS Collaboration [2022a] and light-flavor jets ATLAS Collaboration [2023c]. Additionally, variations to extrapolate the measured uncertainties to the high-T region are considered for both channels ATLAS Collaboration [2021c]. The impact of flavor-tagging systematic uncertainties is less than 1% on the final result in the all-hadronic channel, while a larger contribution is observed in the one-lepton channel, reaching values of the order of 10%.

5.10.2 Modeling uncertainties in background simulations

For the all-hadronic channel, uncertainties in modeling of the $t\bar{t}$ background are included. For the one-lepton channel, uncertainties in modeling the $t\bar{t}$, $t\bar{t}b\bar{b}$, and single top-quark backgrounds are included. All other simulated backgrounds are negligible in both channels and no systematic uncertainties are assigned to them. For the single top-quark background, a 30% normalization uncertainty is applied ATLAS Collaboration [2024b] with an impact ranging between $< 1\%$ and a few percent. Details on the $t\bar{t}$ and $t\bar{t}b\bar{b}$ uncertainties are described below.

5.10.21 $t\bar{t}$ **uncertainties.** Several uncertainties in the theoretical modeling of the $t\bar{t}$ background samples are considered. In the one-lepton channel the $t\bar{t}$ theory systematic uncertainties apply only on the $t\bar{t}$ + light and $t\bar{t}$ + $\geq 1c$ background components as they are estimated from the five-flavor scheme $t\bar{t}$ sample, while the $t\bar{t} + \geq 1b$ background has dedicated systematic uncertainties.

Missing higher order contributions in perturbative expansion of the $t\bar{t}$ production cross-section are estimated by adding in quadrature contributions from renormalization and factorization scale variations, which are obtained by independently varying the parameters μ_R and μ_F by a factor 0.5 and 2.0 and taking the envelope. Uncertainties on the choice of PDF set used for event simulation are estimated by using the PDF4LHC and NNPDF error sets following the PDF4LHC prescription Botje et al. [2011] and taking the envelope. The initial-state radiation (ISR) modeling is estimated by variations of the strong coupling constant α_S through the VAR3C tune variation. The amount of final-state radiation (FSR) in an event is estimated by varying the factorization scale by factors 0.5 and 2.0 inside PYTHIA. To assess the uncertainty in the matching of NLO matrix elements to the parton shower, the nominal sample is compared to an alternative sample obtained setting the `pthard` PYTHIAparameter to 1 (the default is 0). This parameter regulates the definition of the vetoed region of the showering, important to avoid holes and overlaps in the phase space filled by PowHegand PYTHIA. This recommendation follows the description included in Ref. Höche et al. [2022]. The alternative sample was produced using PowHeginterfaced with PYTHIA.306 using the NNPDF2.3LO PDF set and the A14 set of tuned parameters. The impacts of using a different parton shower and hadronization model were evaluated by comparing the nominal $t\bar{t}$ sample with another event sample produced with the POWHEGBOX v2 generator. For the parton shower variation the NNPDF3.0NLO PDF set was used, while events in the sample used to estimate the impact of the hadronization model were interfaced with HERWIG 7.04 Bähr et al. [2008], Bellm et al. [2016], using the H7UE set of tuned parameters Bellm et al. [2016] and the MMHT2014LO PDF set Harland-Lang et al.

[2015]. The impact of a variation of the h_{damp} parameter is assessed by comparing the nominal samples to an alternative set of samples for which h_{damp} is set to $3m_{\mathrm{top}}$.

All alternative samples used to derive the systematic uncertainties are corrected to match the NNLO in QCD and NLO in EW predictions of the top/antitop-quark Tand the top-quark mass distribution using the procedure outlined in Section 5.6.21. A systematic uncertainty associated with this reweighting itself is derived from the maximum and minimum 7-point scale variations, independently for the top/antitop-quark Tand the top-quark mass. The variations are taken into account in the final statistical fit by including them as scale variations on the $t\bar{t}$ background.

To avoid over-constraining the $t\bar{t}$ modeling nuisance parameters in the one-lepton channel, the theoretical systematic uncertainties are treated as uncorrelated among $t\bar{t} + \mathrm{light}$ and $t\bar{t} + \geq 1c$, jet and b-jet multiplicity bins, and between their shape and acceptance components.

The impact of $t\bar{t}$ modeling uncertainties on the final result is found to range between $< 1\%$ and 10%.

5.10.22 $t\bar{t}b\bar{b}$ **uncertainties.** Theory uncertainties on the $t\bar{t}b\bar{b}$ sample are only applied to the $t\bar{t}+\geq 1b$ component in the one-lepton channel, as all other $t\bar{t}$ components are estimated from the bulk $t\bar{t}$ sample.

The scale, PDF, ISR, and FSR uncertainties for the $t\bar{t}b\bar{b}$ sample are derived in the same way as the bulk $t\bar{t}$ sample. The impacts of using a different parton shower and hadronization model are evaluated by comparing the nominal $t\bar{t}$ sample to another sample produced with the PowhegBox v2 generator. For the parton shower variation the NNPDF3.0NLO PDF set was used, while events in the sample estimating the impact of the hadronization model were interfaced with HERWIG 7.1,

using the H7.1-DEFAULT set of tuned parameters and the MMHT2014LO PDF set Harland-Lang et al. [2015].

The matching uncertainty is evaluated by comparing the nominal sample with an alternative sample obtained by setting the `pthard` PYTHIA parameter to 1. The alternative sample was produced using PowHeginterfaced with PYTHIA.307 using the NNPDF2.3LO PDF set and the A14 set of tuned parameters.

The $t\bar{t}b\bar{b}$ uncertainty nuisance parameters are treated as uncorrelated between the different regions of jet and b-jet multiplicities and are further split up into their shape and acceptance components to avoid over-constraining them in the fit. They constitute the dominant systematic contribution in the one-lepton channel, with an impact on the final results ranging from 1% to about 30%.

5.10.3 *Systematics study for one-lepton channel*

An additional section is added here to the published paper to present a systematics study in support of the statistical analysis.

The 1-lepton channel is dominated by systematic uncertainties over statistical uncertainties. The work focused on estimating the theory uncertainties of the $t\bar{t}$ and $t\bar{t}b\bar{b}$ background. Three kinds of theory systematics are calculated in the study:

1. **Scale uncertainties** due to missing higher order in the perturbative expansion of the production cross section; it can be estimated by an envelope of renormalization (μ_R) and factorization scale (μ_F) variations

2. **PDF and alternative PDF uncertainties** the standard deviation of the nominal PDF set and weights of the alternative PDF sets are used to calculate the combined PDF uncertainties

135

Uncertainty type	$t\bar{t}$	$t\bar{t}b\bar{b}$
Scale	20-30%	50-70%
PDF	2-4%	2-4%
Alternative PDF	2-4%	2-30%
Strong coupling	1-2%	3%

Table 5.4. Theory uncertainties summary table

3. **strong coupling(α_s) uncertainties** up and down variations are taken directly from the weights

Figures 5.9 presents the up and down varied distribution of $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ due to scale and PDF uncertainties. The percent variations are shown on the ratio plots at the lower panels. Table 5.4 summarized the impact that each type of uncertainty makes.

Renomalization and factorization scale variations give the leading contribution to the theoretical uncertainties.

5.10.4   Data-driven background estimation uncertainties

In the all-hadronic channel, the dominant systematic uncertainty originates from the multijet estimation method. Deviations of the ratio between data and SM estimate from unity in the VRs, denoted by $k_{VR} = (\text{data} - \text{MC})/\text{multijet}$, are used to estimate a non-closure systematic uncertainty. These k-factors are a measure of the remaining correlations between 3-tags in the multijet estimation method and are expected to be close to unity. The statistical uncertainties on the k-factors $\sigma_{k_{VR}}$ are added in quadrature so that the final multijet uncertainty is calculated according to:

$$\sigma_{ABCD} = \sqrt{\left(1 - \prod_{VR} k_{VR}\right)^2 + \sum_{VR} \sigma_{k_{VR}}^2} \quad, \tag{5.5}$$

136
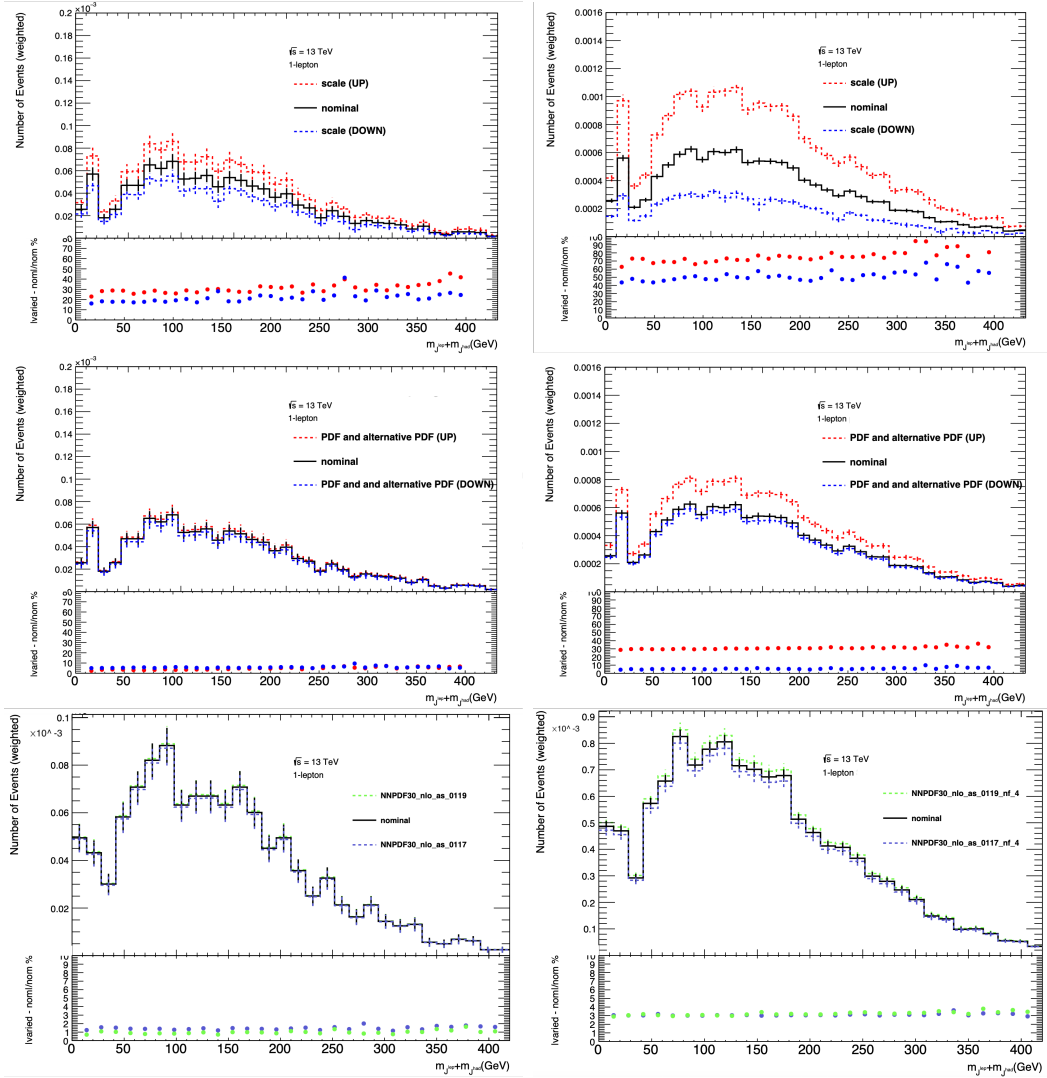
Figure 5.9. Up and down variation due to scale (top row), PDF (middle row) and strong coupling (bottom row) uncertainty for the $t\bar{t}$ (left column) and $t\bar{t}b\bar{b}$ (right column) background processes, in the signal region of the one-lepton channel

137

where the first term under the square root describes the non-closure component of the systematic uncertainty calculated from the k-factors and the second term the statistical uncertainty summed up in quadrature over all four VRs. The statistical component dominates in most regions and is driven by region H in Figure 5.7 which has the lowest number of events. The computation is done separately for each bin of the SR producing values ranging from 33% to 57%. The uncertainty in each SR mass bin is treated as fully uncorrelated with the other SR mass bins thus yielding a conservative estimate.

5.11 Statistical analysis and results

Signal hypotheses are evaluated against the data by performing profile-likelihood fits in each analysis channel. The set of nuisance parameters θ that corresponds to the systematic uncertainties is scaled in such a fashion that, before the fit, all individual uncertainties j have $\theta_j = 0$. In addition, the uncertainties of these nuisance parameters σ are scaled to fulfill $\sigma_j = 1$ before fitting. Defining γ as the set of statistical uncertainties with γ_i the uncertainty on the number of predicted events of a specific bin i, the set is scaled to be a factor around 1. The profile-likelihood fit is then performed by minimizing the quantity $q(\mu) = -2\log\mathcal{L}$, where the likelihood $\mathcal{L}$ is defined as

$$\mathcal{L}(\mu, \theta | S, B, N) = \prod_{i \in \text{bins}} \mathcal{P}\left(N_i | \mu S_i + B_i\right) \times \mathcal{P}\left(S_i + B_i | \gamma_i\right) \prod_{j \in \text{syst.}} \mathcal{G}\left(\theta_j, \sigma_j\right) \qquad (5.6)$$

with N being the number of data events, S the predicted number of signal events and B the expected number of background events. Statistical and systematic uncertainties are taken to be Poissonian ($\mathcal{P}$) and Gaussian ($\mathcal{G}$) distributed parameters, respectively.
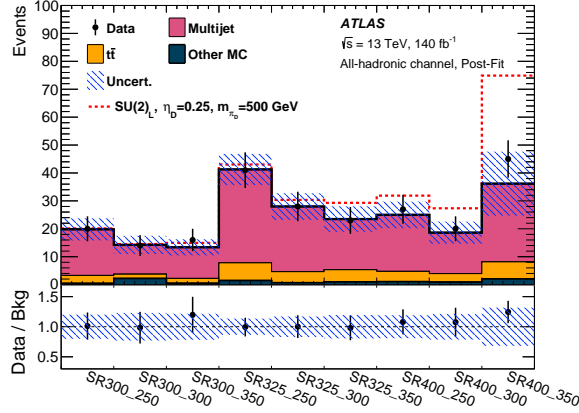
Figure 5.10. Comparison of the events yields in data and the background prediction after the fit under the background-only hypothesis in all bins of the all-hadronic SR. The dashed line is stacked onto the SM background and shows an example $SU(2)_L$ signal with $\eta_D = 0.25$ and $m_{\pi_D} = 500$ GeV. The uncertainty bands contain all statistical and systematic uncertainties. The horizontal axis labels reflect the different SR bins, with the first number indicating the lower boundary of the leading large-R jet mass bin and the second number the lower boundary of the sub-leading large-R jet mass bin.

The fit is evaluated using the ROOFIT Verkerke and Kirkby [2003a] package with the minimization conducted with the MINUIT2 package Hatlo et al. [2005a], which yields the optimal values for μ and the set of parameters θ_i, γ_i and σ_i in a signal plus background fit.

5.11.1   All-hadronic channel

The fit in the all-hadronic channel is performed over all nine SR bins simultaneously. Since the multijet estimate constrains background to the data, most systematic uncertainties (other than from the multijet estimate) have limited impact on the final result and all systematic uncertainties below 1% are ignored. Surviving uncertainties are kept as nuisance parameters in the fit. Because simulated backgrounds constitute less than 20% of the total SM background in the SR, no CRs are used to scale the simulated estimates.

Figure 5.10 shows the distributions in all SR bins after fitting. No significant pulls or constraints are observed for the fitted nuisance parameters. The results are also summarized in Table 5.5. As can be seen the data are very well described by the background-only hypothesis with no significant excess of events above the SM prediction.

5.11.2   One-lepton channel

In the one-lepton channel, a profile-likelihood fit is performed to the binned $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ spectrum in the six SRs and the six CRs defined based on jet and b-jet multiplicity requirements, simultaneously. As already indicated in Section 5.9.2, two independent normalization factors are used for the $t\bar{t}$ background depending on whether in addition to $t\bar{t}$ any b-quarks are present in the event. These two normalization factors are denoted by $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t} + 0b)$. To avoid over-constraining the background modeling nuisance parameters, the theoretical systematic uncertainties are treated as uncorrelated among different $t\bar{t}$ flavors, jet and b-jet multiplicity bins, and between their shape and acceptance components.

The event yields for all SRs, CRs and VRs are shown after a fit to the data under the background-only hypothesis in Figure 5.11. The corresponding distributions of $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ are shown in Figure 5.12 for all jet and b-jet multiplicity bins of the SR. The SM background estimate is in good agreement with data after the background-only fit. The post-fit normalization factors for the $t\bar{t}$ background components are $k(t\bar{t} + 0b) = 0.92^{+0.09}_{-0.08}$ and $k(t\bar{t}+ \geq 1b) = 1.60^{+0.19}_{-0.17}$. No pulls greater than 1σ are observed for the fitted nuisance parameters, with only some constraints in the nuisance parameters describing the theoretical modeling. Data and post-fit background event
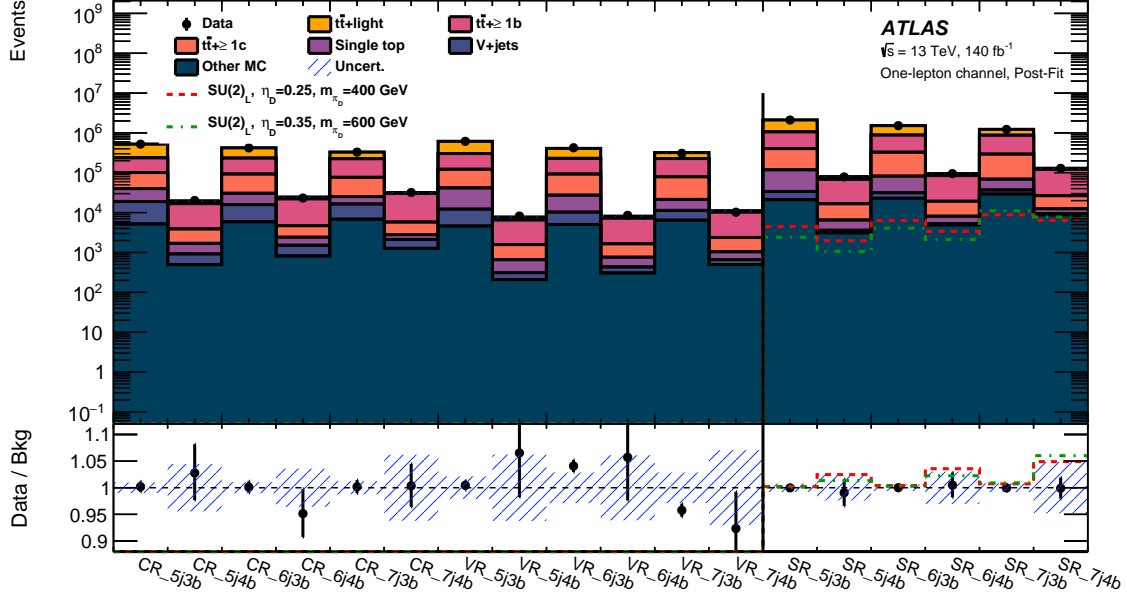
Figure 5.11. Comparison of the events yields in data and the background prediction in the SRs, CRs, and VRs in the one-lepton channel, after the fit to data in all CRs and SRs under the background-only hypothesis. The uncertainty bands contain all statistical and systematic uncertainties. Two example $SU(2)_L$ signal points are overlaid in the SRs.

yields are shown in Table 5.6. Good agreement between the background prediction and data is observed in all regions.

5.12 INTERPRETATION

The results are used to set 95% confidence level (CL) upper limits on the dark pion pair production cross-sections following the CL_S technique Read [2002]. Upper limits on the dark pion production cross-sections are shown in Figure 5.13 for four slices of η_D. The impacts of the different systematic uncertainties, described in Section 5.10, on the signal strength, defined as the signal cross-section normalized to the theoretical prediction, of two benchmark signals are summarized in Table 5.7: the analysis is limited by the systematic uncertainty, with dominant contribution from the theoretical sources of uncertainty on the background modeling.

141
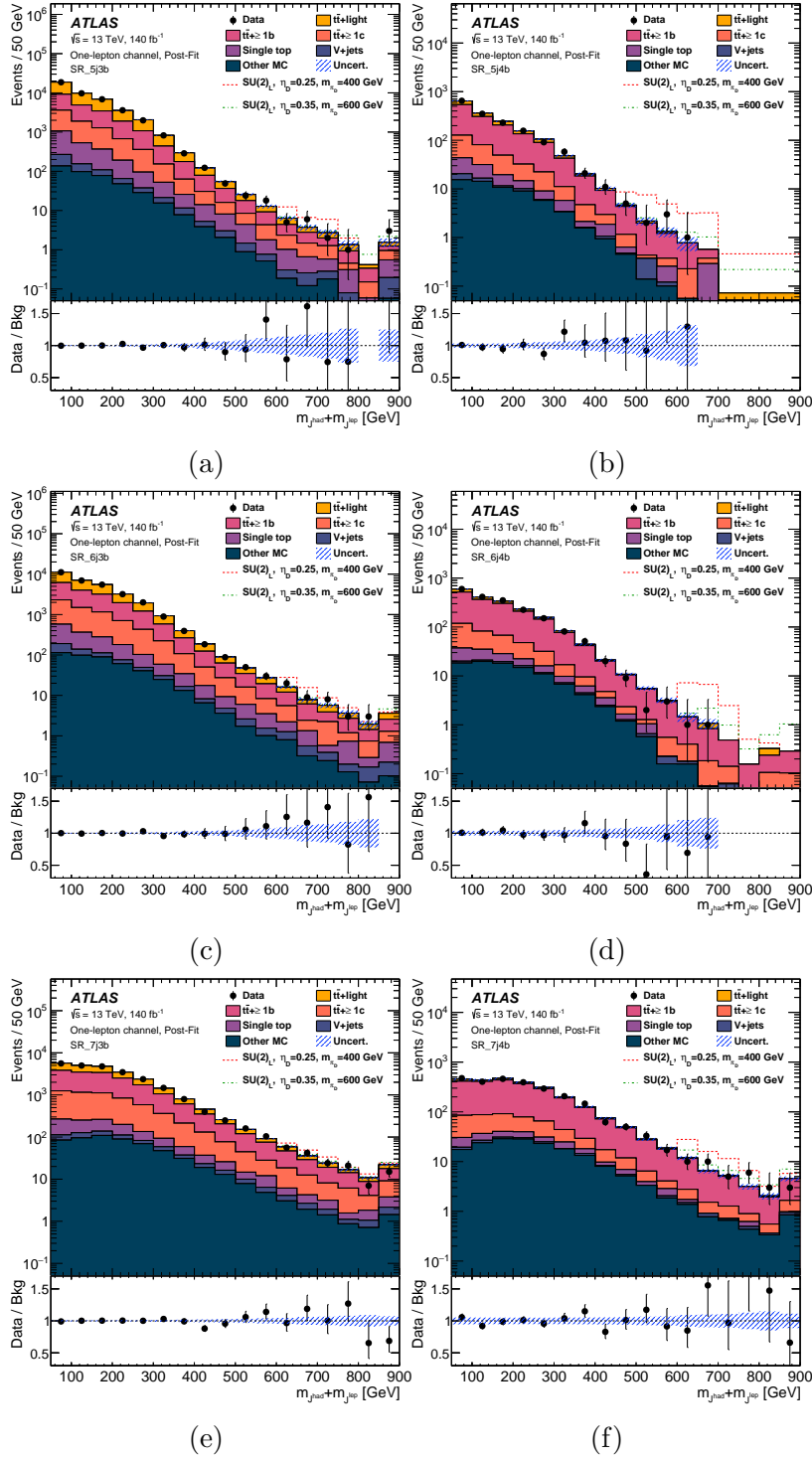
Figure 5.12. Distributions of $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ in a SR_5j3b, b SR_5j4b, c SR_6j3b, d SR_6j4b, e SR_7j3b, and f SR_7j4b, after the fit to data in all CRs and SRs under the background-only hypothesis. Two example $SU(2)_L$ signal point with $\eta_D = 0.25$ and $m_{\pi_D} = 400$ GeV is stacked on top of the background. The uncertainty bands contain all statistical and systematic uncertainties. The dashed line in the ratio panel shows the ratio of signal plus background event yields over just background event yields.

Using the predicted dark pion pair production cross-sections, the limits can be translated into limits on dark pion masses in the two-dimensional η_D–m_{π_D} plane. The exclusion contour obtained from the all-hadronic channel for the $SU(2)_L$ model is shown in Figure 5.14a. No dark pion masses can be excluded for $\eta_D = 0.15$. For $\eta_D = 0.25$ the exclusion covers the mass range 280 GeV $< m_{\pi_D} < 520$ GeV (expected 280 GeV $< m_{\pi_D} < 540$ GeV), while for $\eta_D = 0.35$, dark pions with masses $m_{\pi_D} < 430$ GeV are excluded (expected $m_{\pi_D} < 450$ GeV). The exclusion contour obtained from the one-lepton channel for $SU(2)_L$ is shown in Figure 5.14b and is observed to fully cover the all-hadronic limit and significantly extend the probed phase space for this dark meson model. Since the all-hadronic limit is completely contained within the one-lepton limit, there is no expected gain from a combination of the two channels. Dark pion masses with $m_{\pi_D} < 940$ GeV can be excluded for $\eta_D = 0.45$, $m_{\pi_D} < 720$ GeV are excluded for $\eta_D = 0.35$ and for $\eta_D = 0.25$ the mass region $m_{\pi_D} < 740$ GeV is excluded. The results significantly extend the phase space previously excluded through re-interpretation of other collider searches Butterworth et al. [2022]. In the $SU(2)_R$ model, cross sections are significantly smaller than for $SU(2)_L$ and therefore none of the channels have sensitivity to this model with the current data sample.
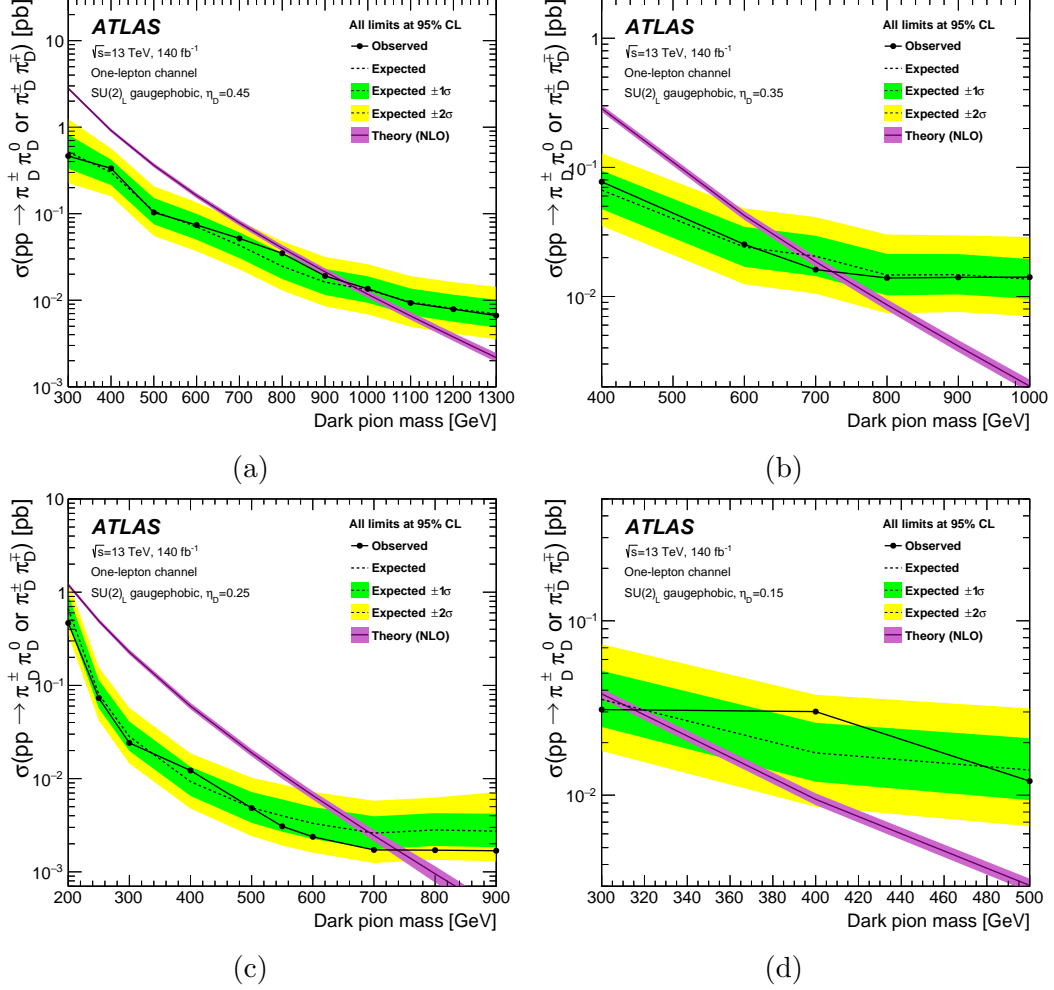
Figure 5.13. Observed (solid line) and expected (dashed line) limits on the dark pion production cross-section as a function of dark pion mass using the CL_S method for all $SU(2)_L$ models in four slices of η_D: a $\eta_D{=}0.45$, b $\eta_D{=}0.35$, c $\eta_D{=}0.25$, and d $\eta_D{=}0.15$. The surrounding shaded bands correspond to one and two standard deviations around the expected limit. The overlaid theory line shows the theoretical dark pion cross-section prediction Kribs et al. [2019a].
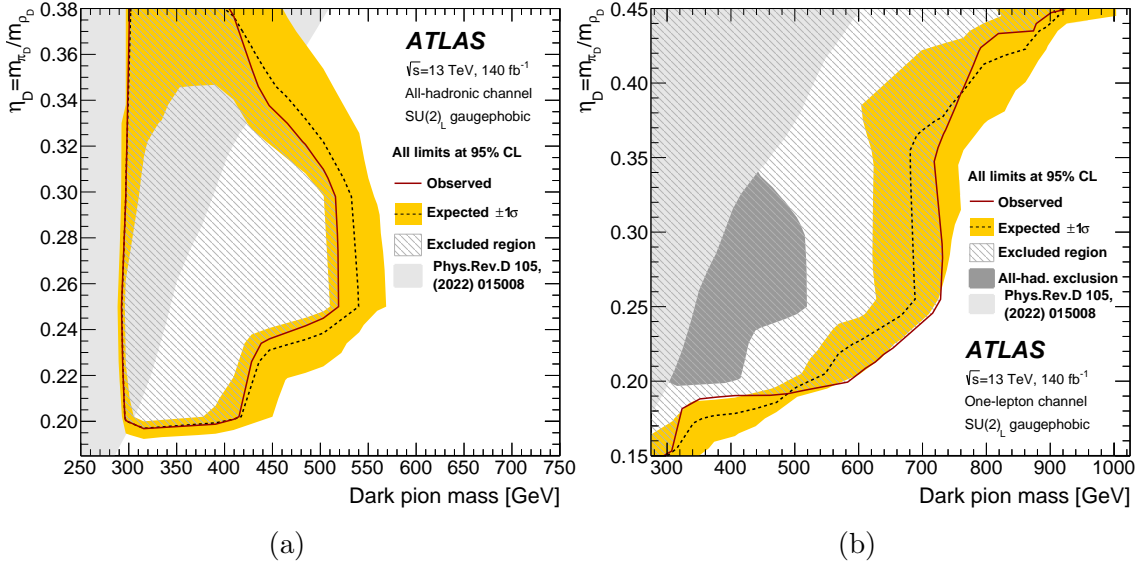
Figure 5.14. Observed (solid line) and expected (dashed line) exclusion contours at 95% CL in the η_D–m_{π_D} plane for $SU(2)_L$ signal models in the a all-hadronic and b one-lepton channel. Masses that are within the contours are excluded, as indicated by the hatched area. An uncertainty band corresponding to the $\pm 1\sigma$ variation on the expected limit is also indicated. The shaded area in a and the innermost shaded area in b indicates the phase space previously excluded through re-interpretation of other collider searches presented in Ref. Butterworth et al. [2022]. The outermost shaded area in b indicates the phase space excluded by the analysis in the all-hadronic channel and is identical to the observed limit shown in a.

Table 5.5. Observed and predicted event yields after the fit under the background-only hypothesis in all nine all-hadronic SR mass bins. The name of each column corresponds to the SR bins: the leading large-R jet lower mass boundary in GeVis followed by the sub-leading large-R jet lower mass boundary; the same convention is used in Figure 5.10. The quoted uncertainties contain statistical and systematic components. The total post-fit uncertainty can be smaller than the sum in quadrature of the different components due to correlations resulting from the fit to data.

	SR300_250	SR300_300	SR300_350
V+jets	< 0.01	1.97 ± 0.89	0.28 ± 0.06
Single top	0.12 ± 0.07	0.00 ± 0.03	< 0.01
$t\bar{t} + X$	0.30 ± 0.04	0.21 ± 0.09	0.17 ± 0.04
$t\bar{t}$	2.9 ± 1.9	1.6 ± 1.1	1.78 ± 0.76
Multijet	16.5 ± 4.3	10.5 ± 3.6	11.1 ± 3.1
Total SM	19.8 ± 4.0	14.2 ± 3.3	13.4 ± 2.9
Data	20	14	16

	SR325_250	SR325_300	SR325_350
V+jets	0.74 ± 0.64	0.12 ± 0.18	0.19 ± 0.16
Single top	0.36 ± 0.06	0.12 ± 0.13	0.27 ± 0.15
$t\bar{t} + X$	0.35 ± 0.06	0.44 ± 0.11	0.50 ± 0.07
$t\bar{t}$	6.4 ± 3.6	4.0 ± 2.0	4.4 ± 2.1
Multijet	33.4 ± 6.7	23.3 ± 5.2	18.1 ± 5.0
Total SM	41.2 ± 5.6	28.0 ± 4.8	23.5 ± 4.6
Data	41	28	23

	SR400_250	SR400_300	SR400_350
V+jets	0.71 ± 0.55	< 0.01	1.19 ± 0.30
Single top	< 0.01	0.47 ± 0.08	0.11 ± 0.02
$t\bar{t} + X$	0.34 ± 0.07	0.40 ± 0.07	0.73 ± 0.10
$t\bar{t}$	3.7 ± 1.8	3.1 ± 2.5	6.1 ± 3.9
Multijet	20.2 ± 5.3	14.7 ± 4.6	28 ± 12
Total SM	25.0 ± 4.8	18.6 ± 4.0	36 ± 11
Data	27	20	45

Table 5.6. Observed and predicted event yields after the fit under the background-only hypothesis in the six SR bins in the one-lepton channel. The name of each column corresponds to the bins described in Section 5.8.2, identified with the number of jets followed by the number of b-jets. The quoted uncertainties contain statistical and systematic components. The total post-fit uncertainty can be smaller than the sum in quadrature of the different components due to correlations resulting from the fit to data. (note: this table is changed from the original paper to correct for typo's)

	SR_5j3b	SR_5j4b	SR_6j3b
$t\bar{t} + \geq 1b$	13119 ± 2600	1020 ± 180	11011 ± 1800
$t\bar{t} + \geq 1c$	5737 ± 500	204 ± 40	4914 ± 600
$t\bar{t} + \text{light}$	20983 ± 1900	235 ± 70	12889 ± 1200
Single top	1667 ± 500	57 ± 18	993 ± 330
V+jets	258 ± 30	10 ± 1.5	179 ± 22
Other MC	420 ± 15	62 ± 4	456 ± 17
Total SM	42183 ± 100	1588 ± 20	30441 ± 110
Data	42191	1577	30469

	SR_6j4b	SR_7j3b	SR_7j4b
$t\bar{t} + \geq 1b$	1310 ± 120	11536 ± 2500	1876 ± 400
$t\bar{t} + \geq 1c$	223 ± 45	4493 ± 800	281 ± 90
$t\bar{t} + \text{light}$	197 ± 50	7095 ± 1500	156 ± 50
Single top	56 ± 18	632 ± 220	60 ± 21
V+jets	7.6 ± 1.2	155 ± 20	14 ± 1.9
Other MC	98 ± 6	581 ± 40	176 ± 14
Total SM	1891 ± 22	24492 ± 120	2563 ± 40
Data	1906	24521	2582

Table 5.7. Impact of different categories of systematic uncertainty in the one-lepton channel, for two signal benchmarks, relative to the total uncertainty on the fitted signal strength. For each category, the fit is repeated with the corresponding group of nuisance parameters fixed to their best-fit values and the impact for each category is evaluated as the quadrature difference between the signal strength uncertainty in the new fit and in the nominal one, divided by the uncertainty in the nominal fit. The contribution from the statistical uncertainty and the systematic one, further separated into the global instrumental and theoretical uncertainties, are shown. The total systematic uncertainty is different from the sum in quadrature of the different groups due to the correlations among the nuisance parameters in the fit.

Category	$SU(2)_L,\ \eta_D = 0.25,$ $m_{\pi_D} = 400$ GeV	$SU(2)_L,\ \eta_D = 0.35,$ $m_{\pi_D} = 700$ GeV
Luminosity	0.03	0.05
Pileup	0.05	0.09
Flavor tagging	0.28	0.26
Leptons	0.01	0.04
Jets	0.08	0.14
$t\bar{t} + {\geq}1b$	0.26	0.53
$t\bar{t} + {\geq}1c$	0.12	0.18
$t\bar{t} + \text{light}$	0.13	0.17
Top TNNLO reweighting	0.08	0.09
Single top	0.06	0.06
Statistical	0.28	0.24
Instrumental	0.30	0.30
Theory	0.38	0.63

5.13 Conclusion

Results from a search for dark mesons originating from strongly-coupled, $SU(2)$ dark flavor symmetry conserving models and decaying to top and bottom quarks are reported. The analysis is based on the full Run 2 data sample of proton-proton collisions at $\sqrt{s} = 13$ TeV recorded by the ATLAS detector at the LHC, corresponding to an integrated luminosity of 140 fb^{-1} . The search is performed in the all-hadronic channel, where the event signature results in eight to ten jets of which at least four originate from bottom quarks, and in the one-lepton channel, corresponding to final states with one electron or muon in addition to jets. In the all-hadronic channel, a data-driven technique is used to estimate the predominant multijet background. In the one-lepton channel, the dominant background comes from $t\bar{t}$ production and is estimated from MC simulations.

No excess above the SM background expectation is observed. The strongest exclusion limits are obtained from the one-lepton channel. For $SU(2)_L$ signals with $m_{\pi_D}/m_{\rho_D} = 0.45$, dark pions with masses $m_{\pi_D} < 940$ GeV are excluded, while for $m_{\pi_D}/m_{\rho_D} = 0.25$ the exclusion covers the mass range $m_{\pi_D} < 740$ GeV. These results constitute the first direct collider constraints on this type of model, and significantly extend the phase space previously excluded through re-interpretation of other collider searches.

CHAPTER VI

SEARCH FOR CHARGED HIGGS THROUGH A REINTERPRETATION OF

THE DARK MESON ANALYSIS

This chapter will present the main physics analysis of this thesis: a search for charged Higgs through a reinterpretation of the dark meson search. The reinterpretation shares the same pre-selection, analysis region definition and background estimation of the dark meson search as described in the previous chapter V. The number of events after applying each pre-selection and signal region (SR) selection are shown in cutflow tables for all-hadronic in Section 6.7.1 and one-lepton channel in Section 6.8.1. A final statistical analysis in one lepton channel was performed using profile-likelihood fits in Section 6.8.3. Finally, a 1D exclusion region was extracted from the fitting results as shown in Section 6.8.34.

6.1   MOTIVATION OF THE REINTERPRETATION

The hypothetical physics process analyzed in this thesis is shown on left of Figure 6.1. It is the leading order Feynman diagram of the single production of heavy charged Higgs followed by the charged Higgs decay to a top and bottom quark. This process shares the same final state particles as the main process analyzed in the dark meson search. This is shown more clearly by a side-by-side comparison in Figure 6.1

The similarity the background processes used in the dark meson search can be recycled for this charged Higgs process. Both signal particles decay to top and bottom quarks. If the decay products of charged Higgs is reasonably collimated, they could possibly be reconstructed using the same large R jets in the all-hadronic channel. These initial observations lead to the assessment that the dark meson search could be a good analysis framework to search for charged Higgs signals.

150

(a) Pair production of dark mesons and their decay

(b) Single charged Higgs production and its decay
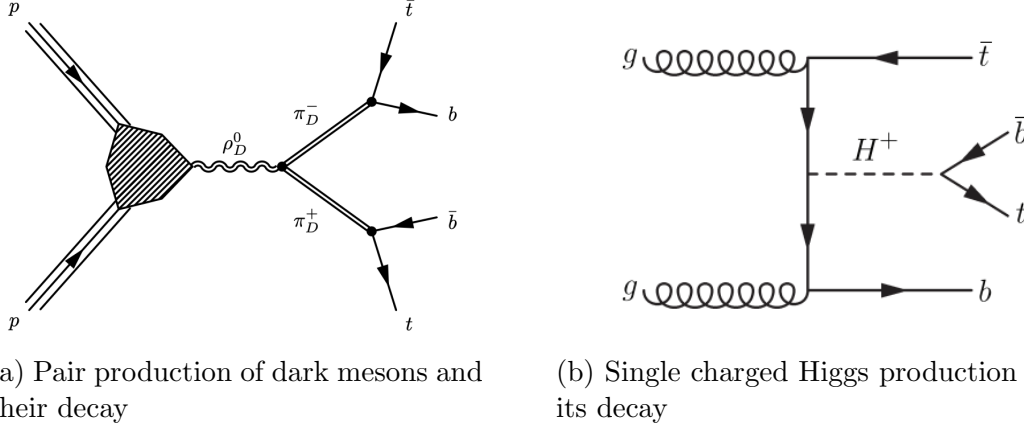
Figure 6.1. Comparison of the process analyzed in the thesis and that analyzed in the dark meson search dar [2024] bar [2021].

6.2 Existing searches of H^+ at the LHC experiments

ATLAS has performed searches of charged Higgs in proton-proton collisions at $\sqrt{s} = 7, 8$ and 13 TeV with data samples ranging from 2.9 to 36 fb^{-1}, probing mass range both below (for example in the $\tau\nu$ decay mode hpl [2018b]) and above the top quark mass (for example in the tb decay mode hpl [2018c]). The CMS experiment also conducted the search of charged Higgs decaying to top and bottom quarks. The all-hadronic and one-lepton channel results are presented by Ref.Collboration [2020a] and Ref.Collboration [2020b].

The ATLAS search in Ref.bar [2021] is used as the main reference when developing the reinterpretation in this thesis. It studied the leading order heavy charged Higgs production decaying to a top and a bottom quark. The main physics process studied is shown in the Feynman diagram 6.1(b). Events with one charged lepton (e, μ) and jets in the final state are considered in the analysis. A Neutral network algorithm was used to enhance the separation between signal and the background. After pre-selections,
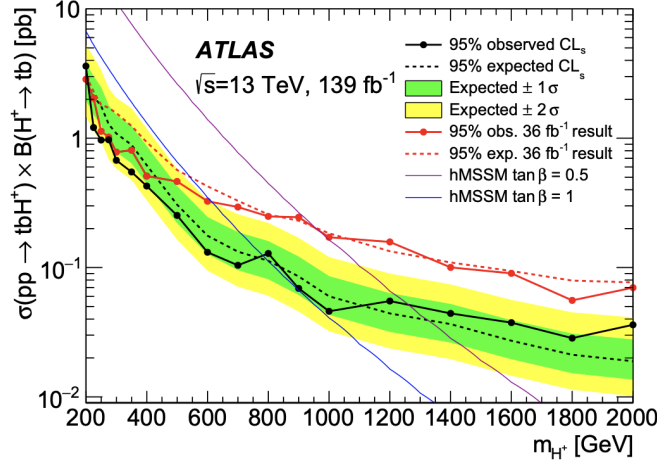
Figure 6.2. Observed and expected upper limits to produce $H^+ \rightarrow tb$ in association with a top quark and a bottom quark. The bands surrounding the expected limit show the 68% and 95% confidence intervals. The red lines show the observed and expected 95% CL exclusion limits obtained with the 36 fb^{-1}data sample hpl [2018a]. Theory predictions are shown for two representative values of $\tan \beta$ in the hMSSM benchmark scenario. Uncertainties in the predicted H^+ cross-sections or branching ratios are not considered bar [2021].

the events are categorized into 4 regions according to the number of reconstructed jets and b-jets.

A binned maximum-likelihood fit to the data is performed simultaneously on the Machine learning output distributions in the four analysis regions, and each mass hypothesis is tested separately. The final statistical analysis result was interpreted with the theory of 2HDM and hMSSM Branco et al. [2012]. And the final exclusion plot is shown in Fig 6.2

For $\tan \beta$ values below 1, charged Higgs masses between 200 and 790 GeV are observed to be excluded at 95% Confidence Level (CL). The caveat with hMSSM models though is that the theory is not perturbative when $\tan \beta$ is smaller than 1 MARTIN [1998].
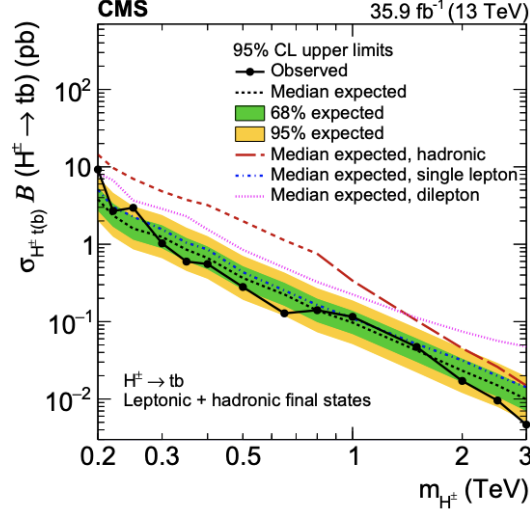
Figure 6.3. Upper limits at 95% CL on the product of the H^+ production cross section and branching fraction as a function of m_H^+ for the process $\sigma_H^+ t(b) B(H^+ \to \mathrm{tb})$. The median expected limit (dashed line), 68% (inner green band), and 95% (outer yellow band) confidence interval expected limits are also shown (left). The relative expected contributions of each channel to the overall combination are shown (right). The black solid line corresponds to the combined expected limits while the dashed, dotted and dash-dotted lines represent the contributing channels Collboration [2020a].

The CMS experiment conducted the search with the same process in these channels: Ref.Collboration [2020a] and Ref.Collboration [2020b]. The result from all three channels is presented in Figure 6.3.

The single-lepton final state has the best sensitivity in the whole m_H^+ range from 0.2 to 3 TeV, while the dilepton channel contributes in the low m_H^+ regime, i.e., $\leq$ 1.5 TeV, and the all-jet channel improves the overall sensitivity by 20–25% at larger values of m_H^+.

To include other studies about additional Higgs boson beyond the SM, a summary plot 6.4 shows the overall status of charged Higgs searches when interpreted with hMSSM Branco et al. [2012]. The teal colors region shows the result for the same process with $H^+ \to \mathrm{tb}$. The region with $\tan \beta < 20$ and $m_H^+ < 800$ GeV is excluded.
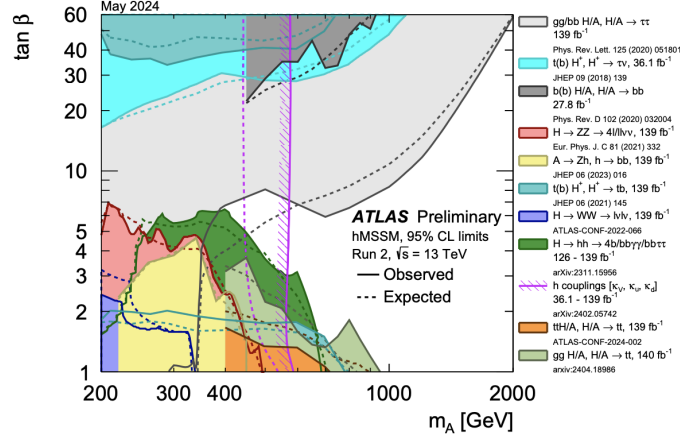
Figure 6.4. The expected and/or observed limits at 95% confidence-level (CL) in the hMSSM, for several analyses; $m_A = m_H^+$ based on proton-proton collision data recorded by the ATLAS experiment at $\sqrt{s}$=13 TeV ATL [2024].

6.3 SIGNAL SIMULATION

The simulated signals used in the reinterpretation share with those used in the analysis mentioned above bar [2021]. The production and decay of H^+ were simulated through the Monte Carlo method. The $pp \rightarrow tb\, H^+$ process followed by $H^+ \rightarrow tb$ decay was modelled with MadGraph5_aMC@NLO Degrande et al. [2015a] at next-to-leading order (NLO) in QCD Alwall et al. [2014b] using a four-flavour scheme (4FS) implementation with the NNPDF2.3NLO Ball et al. [2013] parton distribution function (PDF). The parton shower of the events is generated using Pythia 8.1.2Sjöstrand et al. [2008]. A set of underlying-event (UE) parameters were tuned to ATLAS data called A14tune ATLAS Collaboration [2014]. 11 signal samples of different H^+ mass from 200 GeV to 2000 GeV are analyzed in the reinterpretation, as listed in table 6.1. The list of signal samples with full MC sample names can be found in appendix A

154

Table 6.1. Simulated H^+ samples; cross sections are in pb and take into account the production of both H^+ and H^-.

Signal	DSID	Simulation type	$\tan \beta = 2$	$\tan \beta = 1$
H200	451290	full	1.69	6.73
H300	451294	full	0.71	2.84
H400	450595	full	0.33	1.33
H500	450596	full	0.17	0.66
H600	450597	fast	0.088	0.350
H800	450002	fast	0.028	0.112
H1000	450004	fast	0.0103	0.0409
H1200	450598	fast	0.00413	0.0164
H1400	450599	fast	0.00178	0.00710
H1600	450600	full	8.15458×10^{-4}	0.00325
H1800	450601	full	3.88126×10^{-4}	0.00155
H2000	450602	full	1.916558×10^{-4}	0.0007637

6.4 Data and background simulations

The analysis in this thesis uses the same ATLAS data and background processes. Please see appendices in Gledhill et al. [2020] for the list of background samples.

6.5 H^+ cross section

The heavy charged Higgs production cross section in association with heavy quarks (for the process of $pp \rightarrow tH^+ + X$) at the LHC are presented in Table 6.1 and Figure 6.5. They are Santander-matched cross sections for 2HDM type-II but without SUSYQCD corrections Degrande et al. [2015b],CERN [2017],Flechl et al. [2015],Dittmaier et al. [2011],Berger et al. [2005]. There are two methods to calculate the associated production cross section of the charged Higgs: 4FS and 5FS, depending on whether bottom quarks are involved in the initial state. The Santander-matched cross section is a method of calculating the cross section by combining the two different methods. Figure 6.5 presents how cross sections vary with respect to
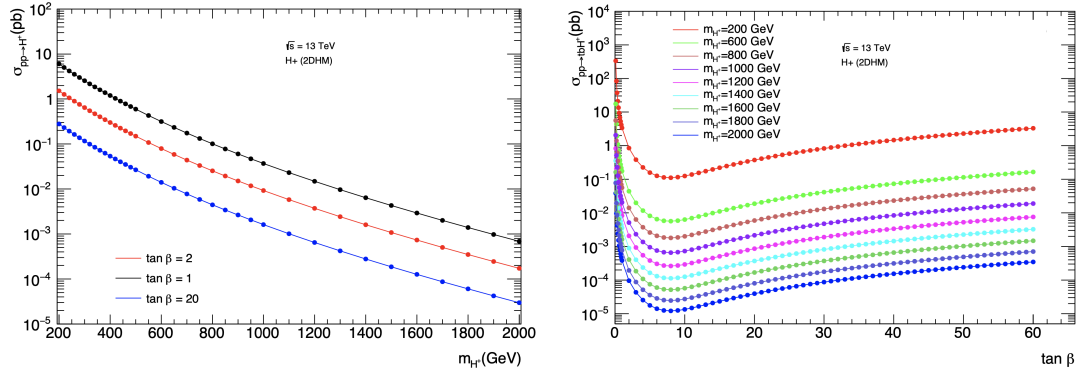
Figure 6.5. Santander-matched production cross section of single charged Higgs with respect to H^+ mass and $\tan \beta$. The caveat with hMSSM models is that the theory is not perturbative when $\tan \beta$ is smaller than 1 MARTIN [1998].

H^+ mass and $\tan \beta$. Please see this <u>cern twiki link</u> for a table of the cross sections and the uncertainties, for m_H^+ from 200 GeV to 2000 GeV and for $\tan \beta$ from 0.1 to 60. These values are used in the entire analysis flow of the reinterpretation. Several factors contribute to the total uncertainty of the cross sections: estimation of (Parton Distribution Function) PDF, strong coupling (α strong), renormalization and factorization scale uncertainty from missing higher order contributions in perturbative expansion of the cross-section calculation, and bottom quark mass measurement.

6.6   Flow of analysis

The re-interpretation in all-hadronic and one-lepton channel follow the same procedure with the following steps:

1. N-tuple production with AnalysisTop 21.2.200 with documentation reconstruction group

2. Flatten the nominal tree of the n-tuples

3. 1D histogram production using the flat n-tuples

156

4. Flatten the systematic trees of the n-tuples for jet and lepton systematics

5. Statistical analysis with TRExFitter-00-04-26 with documentation tre

6. Limit extraction

6.7 ALL-HADRONIC CHANNEL

This section presents reinterpretation results in the all-hadronic channel of the dark meson search. Six signal samples are analyzed at this stage with the framework from the dark meson search.

6.7.1 Pre-selections and signal region selection Cutflow

Raw and weighted number of events are counted after applying each pre-selection criteria as well as the SR selection criteria. Table B.1 in appendix B presents the raw event count and Table 6.2 presents the weighted event counts. The selection criteria are defined by the original dark meson search. The pre-selection criteria are shown in Table 5.2 and the all-hadronic signal region criteria are shown in Table 5.3. Previous dark meson version of the cutflow table was reproduced before that of the H^+ signals to ensure consistent settings.

The simulated background and signal events need to be normalized so they can be directly compared to data. To scale them to data, the weight includes MC weight from the initial AOD, integrated luminosity, signal cross section, k-factor and weight associated to each campaign of the ATLAS data collection (in short $MCweight \times \int Ldt \times campaign\ weight \times kfactor \times \sigma_{signal}$).

Two selection criteria cut off significant amount of events: Trigger and the requirement on large R jet mass ($m_{\mathrm{jet},R=1.2}^{1,2} > 190$ GeV). The H_{T} trigger rejects

Table 6.2. All-hadronic pre-selection and signal region cutflow (separated by the horizontal bar) for charged Higgs signals with $\tan\beta = 1$, at 140 fb^{-1} integrated luminosity; a factor of two is included to account for both H^+ and H^-

Cut	H600	H1200	H1400	H1600	H1800	H2000
Derivation	15575	734.0	318.1	191.1	90.9	44.9
GRL	15575	734.0	318.1	191.1	90.9	44.9
Good Calo	15575	734.0	318.1	191.1	90.9	44.9
Primary vertex	15575	734.0	318.1	191.1	90.9	44.9
Trigger	806	429.3	230.2	155.3	79.0	40.4
Bad muon veto	806	429.3	230.1	155.2	79.0	40.4
Jet cleaning	802	427.3	229.1	154.5	78.6	40.3
$H_T > 1150$ GeV	365	292.5	183.4	135.0	72.0	38.0
$N_{\mathrm{jets}} \geq 6$	340	251.6	152.9	110.7	58.4	30.8
$N_{\mathrm{jets,\ R=1.2}} \geq 2$	325	244.5	148.1	105.9	55.7	29.2
$N_{\mathrm{b\text{-}jets}} \geq 3$	205	132.8	77.9	52.3	26.1	13.3
electron veto	153	96.9	58.7	39.4	19.8	10.1
muon veto	88	64.2	39.1	26.9	13.5	7.1
$m_{\mathrm{jet,R=1.2}}^{1,2} > 190$ GeV	11.29	7.12	4.92	3.87	2.37	1.36
$p_{T,\mathrm{jet}}^6 > 25$ GeV	11.5	6.85	4.81	3.77	2.3	1.33
preselected	11.5	6.85	4.81	3.77	2.3	1.33
$lead.m_{bb}/p_{T,bb} > 0.25$	10.91	6.43	4.33	3.45	2.05	1.18
$sub.m_{bb}/p_{T,bb} > 0.25$	10.68	6.07	4.06	3.2	1.82	1.07
$lead.m_{\mathrm{jet,R=1.2}} > 300 GeV$	4.85	2.63	1.79	1.5	0.87	0.57
$sub.m_{\mathrm{jet,R=1.2}} > 250 GeV$	2.81	1.62	0.98	0.83	0.56	0.36
$lead.\Delta R\,(j, b_2) < 1.0$	0.73	0.49	0.29	0.29	0.16	0.09
$sub.\Delta R\,(j, b_2) < 1.0$	-0.0	0.1	0.04	0.06	0.01	0.02

158

large number of events particularly when H^+ mass is small, where H_{T} is defined as the scalar sum of all jet p_{T}'s in an event. And the large R jet mass requirement indiscriminately eliminates 80-90% of the events with respect to the remaining events from the last stage of selection. However, as can be seen from the weighted cutflow table of the dark meson signal C.1 in the Appendix C, the same requirements do not cut off significant number of events of dark meson signals. These results can be explained by comparing the 1D distributions of two signals as shown in the following section.

6.7.2   1D distributions of the discriminating variables

This section presents 1D distributions with pre-selections. Each plot in Figure 6.6 includes a discriminating variable comparing the dark meson benchmark signal with multiple H^+ signals of similar and higher mass. While the distributions of $m_{bb}/p_{T,bb}$ and $\Delta R\,(j, b_2)$ do not have drastic difference between H^+ and dark meson, the large R jet mass distributions of H^+ in Figure 6.6 reside on much smaller values compared to the dark meson signal. The large R jet mass is meant to reconstruct the mass of the target signal, so the reconstructed large R jet mass should ideally be close or equal to the mass of the target signal. While the H^+ mass in the plots have mass larger than 600 GeV , the main bulk of all the distributions is smaller than 300 GeV. One can see that with the positions of the vertical lines; majority of the events will be rejected once the SR selections are applied. In Figure 6.7, the same behavior can be observed through 2D histograms of pre-selected H^+ events in leading and sub-leading large R jet mass. These plots suggest a significant underestimation of the reconstructed H^+ mass when the dark meson selection criteria are applied. Figure 6.8 plots ttbar events in leading and subleading large R jet mass. It shows how SR is defined to exclude

the background ttbar events by requiring large R jet masses. Comparing the Figure 6.7 with the dark meson SR in Figure 6.8, the charged Higgs signal events mostly fall in the background region of the all-hadronic channel.

6.7.3   All-hadronic channel results

Yields of H^+ in all SRs of the all-hadronic channel are put side to side with dark meson results in Table 6.3. For $\mathrm{m}_H^+ = 600$ GeV through 2000 GeV, the number of events passing the SR selection will not be enough to extract limits. The main reason is the underestimation of the large R jet mass and the stringent requirement on H_T trigger.

Figure 6.9 presents data yields of the dark meson benchmark model in all signal regions defined by the leading and subleading large R jet mass. The yields from signal events are stacked on top of total background events. The results from 6.9 are copied to Figure 6.10 to show how yields of the dark meson (indicated by the red dashed line) compare to yields of various charged Higgs models. The charged Higgs yields are consistently lower than that dark meson yields.
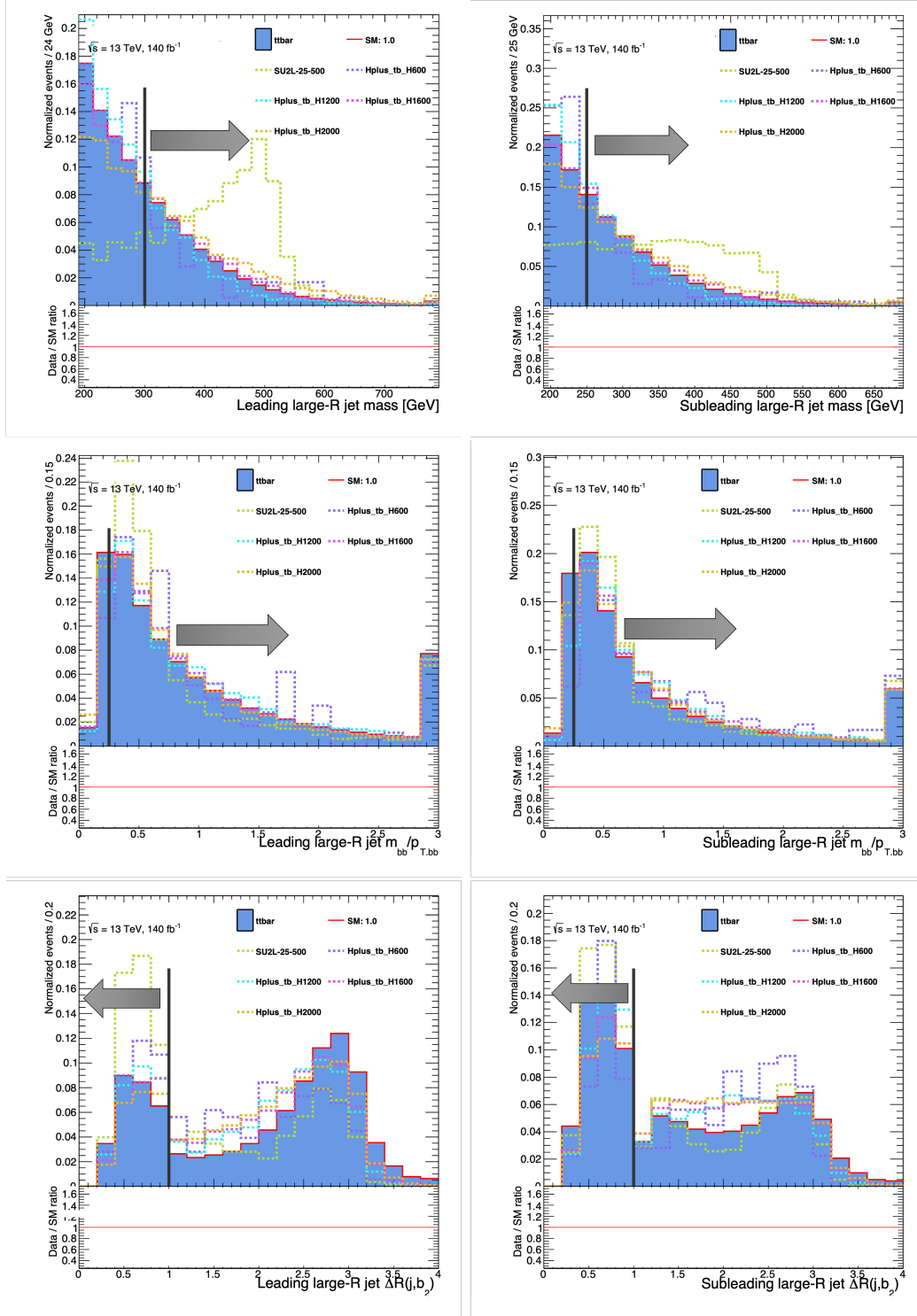
Figure 6.6. Plots of the all-hadronic preselection variables with the last bin including overflow events. H^+ samples with $m_H^+ = 600, 1200, 1600, 2000$ are added for comparison with the dark meson benchmark sample SU2L-25-500; Multijet background is not calculated at the preselection stage. SR selections are indicated with a vertical line.

161
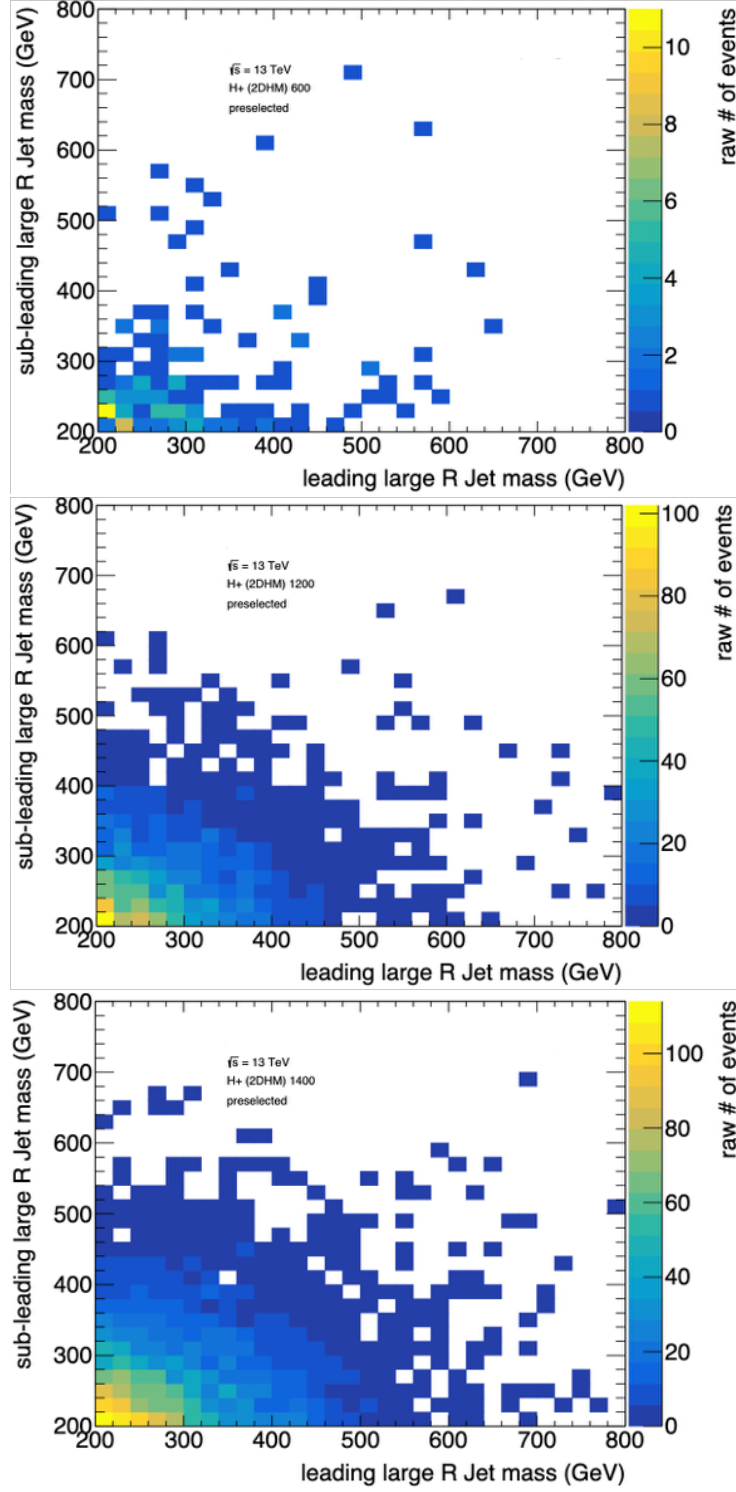
Figure 6.7. 2D histograms of the leading and sub-leading large R jet mass in the all-hadronic pre-selection; H^+ samples with $m_H^+ = 600, 1200, 1400$ are included.

162
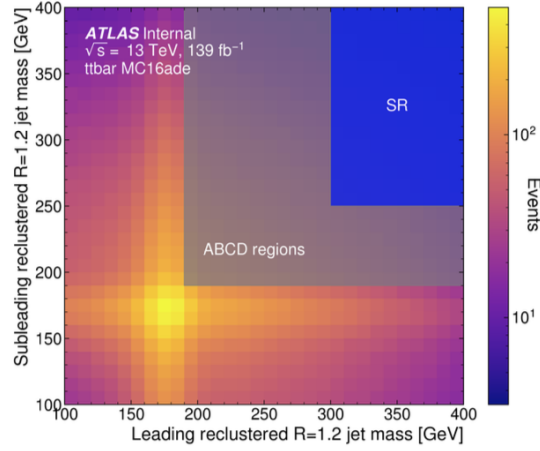
Figure 6.8. SR selection based on leading and sub-leading large R jet mass of $t\bar{t}$ events. The ABCD region was chosen to exclude $t\bar{t}$ from all regions. The SR bins are open ended for the largest leading and subleading jet masses Gledhill et al. [2020].
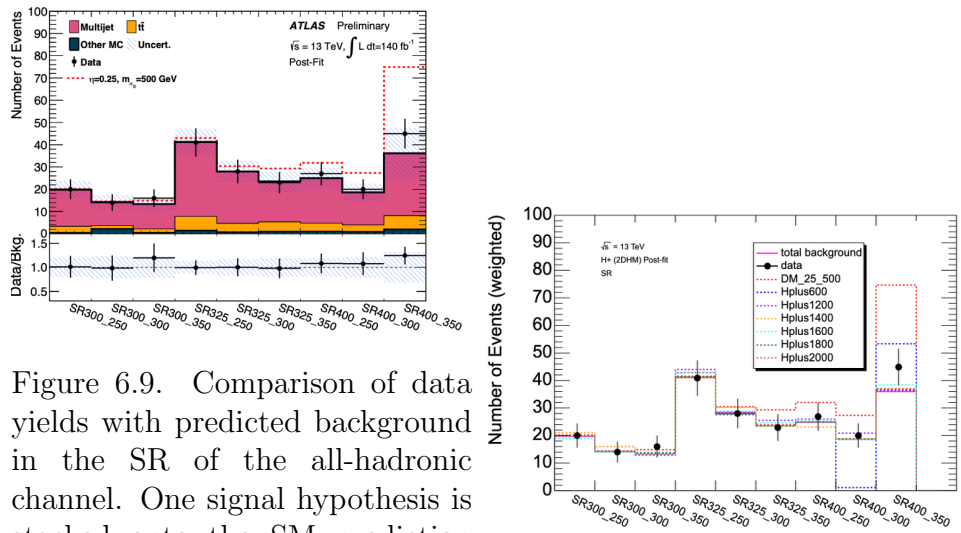
Figure 6.9. Comparison of data yields with predicted background in the SR of the all-hadronic channel. One signal hypothesis is stacked onto the SM prediction to illustrate how signal would manifest itself. The shown uncertainty is a combination of statistical and systematic uncertainties. The x-axis labels reflect the different SR bins dar [2024].

Figure 6.10. Comparison of data yields with predicted background in the SR of the all-hadronic channel. Charged Higgs hypotheses are stacked onto the SM prediction to illustrate how they would manifest itself.

Table 6.3. Event yields in all nine all-hadronic signal regions with three dark meson and six charged higgs signal points at the full 140 fb^{-1} integrated luminosity. The quoted uncertainties contain statistical and systematical components. The name of the signal regions indicate the lower mass boundary on the leading large-R jet in GeV followed by the sub-leading large-R jet lower mass boundary. $tan\beta = 0.1$ for the H^+ samples

	SR300_250	SR300_300	SR300_350	SR325_250	SR325_300
$t\bar{t}$	3.09 ± 2.13	1.71 ± 1.01	1.78 ± 0.73	6.83 ± 3.98	4.27 ± 2.22
V+jets	0.00 ± 0.00	2.00 ± 1.83	0.28 ± 0.29	0.72 ± 0.94	0.13 ± 0.24
Single top	0.12 ± 0.14	0.00 ± 0.03	0.00 ± 0.00	0.36 ± 0.22	0.12 ± 0.17
$t\bar{t} + X$	0.30 ± 0.07	0.22 ± 0.06	0.17 ± 0.05	0.35 ± 0.10	0.44 ± 0.14
Multiboson	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
QCD	15.86 ± 8.67	11.62 ± 6.95	9.04 ± 3.76	35.24 ± 12.05	25.27 ± 11.10
Total background	19.38 ± 8.82	15.55 ± 7.20	11.26 ± 3.66	43.49 ± 11.52	30.24 ± 10.75
$\eta = 0.25$, $m_{\pi_D} = 500$ GeV	0.36 ± 0.25	0.30 ± 0.23	1.59 ± 0.54	1.80 ± 0.45	2.38 ± 0.55
$\eta = 0.35$, $m_{\pi_D} = 500$ GeV	2.77 ± 1.46	0.35 ± 0.36	1.47 ± 1.06	4.44 ± 1.81	4.36 ± 1.77
$\eta = 0.45$, $m_{\pi_D} = 400$ GeV	2.57 ± 1.39	3.09 ± 1.63	0.40 ± 0.90	5.87 ± 1.90	8.97 ± 2.56
Data	20	14	16	41	28
$m_{H^+} = 600$ GeV	0	0	0	0	0
$m_{H^+} = 1200$ GeV	-0.13	1.75	-0.41	2.75	0.47
$m_{H^+} = 1400$ GeV	1.23	1.69	-0.13	-0.26	2.25
$m_{H^+} = 1600$ GeV	-0.89	-0.27	-0.15	1.71	0.84
$m_{H^+} = 1800$ GeV	0.13	0.12	0.3	0.32	-0.54
$m_{H^+} = 2000$ GeV	0.07	0.06	-0.01	0.05	0.37

	SR325_350	SR400_250	SR400_300	SR400_350
$t\bar{t}$	4.64 ± 2.24	3.87 ± 1.89	3.41 ± 2.91	6.61 ± 4.28
V+jets	0.19 ± 0.34	0.72 ± 0.90	0.00 ± 0.28	1.18 ± 0.76
Single top	0.28 ± 0.25	0.00 ± 0.00	0.47 ± 0.26	0.11 ± 0.11
$t\bar{t} + X$	0.50 ± 0.09	0.34 ± 0.08	0.40 ± 0.10	0.73 ± 0.13
Multiboson	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
QCD	27.34 ± 12.38	18.75 ± 8.69	15.40 ± 6.32	26.40 ± 13.10
Total background	32.95 ± 11.94	23.68 ± 8.39	19.68 ± 6.10	35.03 ± 12.32
$\eta = 0.25$, $m_{\pi_D} = 500$ GeV	5.82 ± 1.09	6.90 ± 1.01	8.76 ± 1.40	38.74 ± 3.59
$\eta = 0.35$, $m_{\pi_D} = 500$ GeV	5.48 ± 2.03	4.91 ± 1.82	6.99 ± 2.19	12.78 ± 2.82
$\eta = 0.45$, $m_{\pi_D} = 400$ GeV	10.16 ± 2.54	1.47 ± 0.82	5.50 ± 2.59	3.59 ± 1.50
$m_{H^+} = 600 GeV$	0	0	-17.6	17.3
$m_{H^+} = 1200 GeV$	2.12	1.0	2.4	0.53
$m_{H^+} = 1400 GeV$	0.04	-1.89	-0.0	1.2
$m_{H^+} = 1600 GeV$	1.21	0.81	-0.05	2.37
$m_{H^+} = 1800 GeV$	0.01	-0.01	0.21	0.8
$m_{H^+} = 2000 GeV$	0.6	0.22	0.15	0.81
Data	23	27	20	45

6.7.4   All-hadronic channel conclusion and outlook

The target Feynman diagram of charged Higgs has only one large-radius (large-R) jet, whereas the all-hadronic channel selection criteria require at least two such jets. Consequently, H^+ events from single production often fail to meet the selection requirements of the all-hadronic channel. It is concluded that the all-hadronic channel of the analysis is not a good fit to analyze the charged Higgs signal events.

Although an optimization towards H^+ does not fit the time frame of this thesis, different reconstructions methods are explored in collaboration with an undergraduate student Kalel Chester. shows the results of multiple alternative methods to better reconstruct H^+ signals. The new method attempts to rebuild the H^+ mass by summing the two leading p_T small jets of R=0.4. The solid lines represent the distributions of the H^+ candidate mass using new reconstructing methods and the dashed lines are those from the dark meson analysis. The new methods show much better results compared to the old method. More study in this direction should be explored if a future iteration of this H^+ search was to be performed. In addition to new reconstruction methods, adjusting some of the selection criteria (such as smaller H_T trigger threshold or large R jet mass requirement) could save more signal events and gain better significance for charged Higgs.

6.8   One-lepton channel

In contract to all-hadronic channel, one-lepton channel has no requirement on larger R jet mass and it has a much smaller requirement on H_T (one-lepton channel requires the H_T to be larger than 300GeV, instead of the 1150 GeV in all-hadronic channel). These two conditions make it much easier for H^+ events to pass the selection
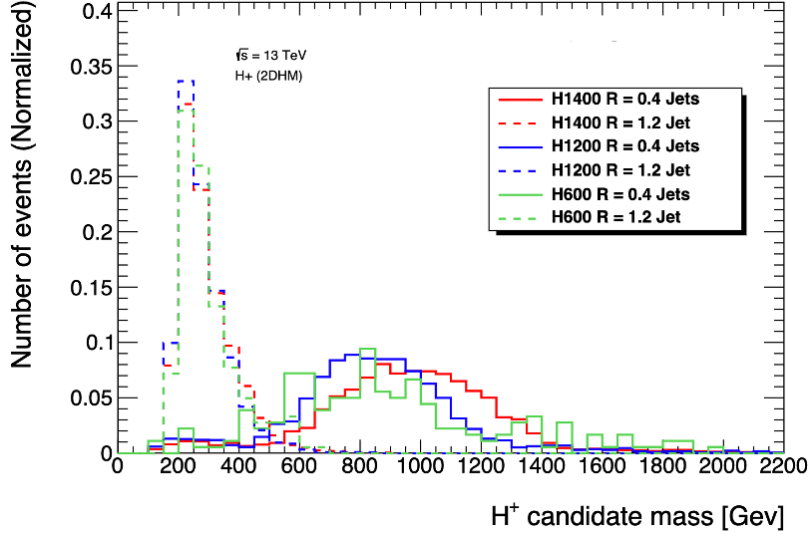
Figure 6.11. Comparison of different methods of reconstructing H^+ using the candidate mass. Credit: Kalel Chester.

criteria. Three more samples of smaller H^+ mass: 200GeV, 300GeV and 400GeV are added to the analysis to increase the number of events after selections. H^+ signals are then analyzed with the one lepton channel of the dark meson search.

6.8.1 Pre-selections and signal region selections

The analysis region definitions have been descried in Section 5.8.2 of the dark meson search paper in Chapter V. They are now presented in Table 6.4 for easier reference.

Table 6.4. Analysis region definitions for the 1-lepton channel

Region	Preselection	N_{jets}	$\Delta R(l, b_2)$	$m_{bb_{\Delta R_{min}}}$
SR	See Table 5.2	≥ 5	< 2.7	> 100 GeV
CR	See Table 5.2	≥ 5	$> 2.7, < 3.5$	> 40 GeV, $< 100 GeV$
VR	See Table 5.2	≥ 5	$> 2.7, < 3.5$	> 100 GeV

Table 6.6 presents number of events after each pre-selection and SR criteria according to Table 6.4. The one-lepton channel uses single lepton trigger. Slight

166

performance differences between data and simulation in lepton reconstruction, identification, isolation and triggering are corrected by the application of scale factors. The Scale Factor (SF) uncertainties of leptons are propagated through a set of event weights. So compared to the all-hadronic version, a few more weights are added. To scale the simulation events to be comparable to data, additional weights are included compared to all-hadronic channel: lepton scale factor weight (weight_leptonSF), b-tagging weight (weight_bTagSF_DL1r_77) and jet vertex tagging weight (weight_jvt).

Number of events remain after SR selections is comparable to that of dark meson signals. This can be seen by comparing to the equivalent dark meson results in Table 6.5. The significances are then calculated to evaluate possibilities of extracting limits. "B" in the significance calculation represents the post-fit background yield in each signal region, obtained from post-fit histograms. ΔB is the Poisson error of the background yield. Significances are evaluated for each SR and H^+ masses. Three types of significance definition are used and as shown in table 6.7, 6.8 and 6.9. A significance bigger than 2 usually indicates that the signal could be excluded at 95% confidence level. So, the one lepton channel analysis expects to have sensitivity for samples with H^+ mass 300 GeV and 400 GeV .

Table 6.5. Expected event yields of several dark meson signals for a luminosity of 140 fb^{-1} , for background and a selection of signal samples, in the 1-lepton channel after each selection step requirement. Ref.Gledhill et al. [2020]

Process	Preselection	Jets_ N≥5	deltaRLep2ndClosestBJet<2.7	bb_ m_ for_ minDeltaR>100	s/sqrt(s+b)
ttlight	175081 ± 151 (100.00%)	129749 ± 130 (74.11%)	100835 ± 114 (77.72%)	46245 ± 77 (45.86%)	
ttge1b	74674 ± 33 (100.00%)	63852 ± 30 (85.51%)	51085 ± 27 (80.01%)	22765 ± 18 (44.56%)	
tt1c	55305 ± 84 (100.00%)	47831 ± 78 (86.49%)	38015 ± 69 (79.48%)	17809 ± 47 (46.85%)	
Single top	12954 ± 37 (100.00%)	9405 ± 32 (72.60%)	6971 ± 27 (74.12%)	3582 ± 20 (51.38%)	
ttPlusX	5234 ± 9 (100.00%)	4872 ± 9 (93.08%)	4096 ± 8 (84.06%)	1756 ± 5 (42.87%)	
V + jets	5233 ± 46 (100.00%)	3911 ± 38 (74.74%)	2223 ± 28 (56.83%)	617 ± 13 (27.77%)	
Multiboson	403 ± 7 (100.00%)	311 ± 7 (77.13%)	220 ± 6 (70.84%)	57 ± 3 (25.71%)	
Total predicted bkg	328884 ± 186 (100.00%)	259931 ± 162 (79.03%)	203445 ± 142 (78.27%)	92831 ± 96 (45.63%)	
$SU(2)_L$ 0.45 600	2899 ± 48 (100.00%)	2801 ± 47 (96.62%)	2569 ± 45 (91.72%)	1775 ± 37 (69.07%)	5.825e+00
$SU(2)_R$ 0.45 300	932 ± 7 (100.00%)	821 ± 6 (88.13%)	758 ± 6 (92.30%)	267 ± 4 (35.21%)	8.757e-01
$SU(2)_L$ 0.25 500	334 ± 6 (100.00%)	317 ± 5 (95.09%)	277 ± 5 (87.28%)	229 ± 5 (82.55%)	7.505e-01
$SU(2)_L$ 0.45 1000	225 ± 4 (100.00%)	221 ± 4 (97.88%)	202 ± 3 (91.40%)	167 ± 3 (82.86%)	5.482e-01
$SU(2)_L$ 0.25 700	41 ± 1 (100.00%)	39 ± 1 (95.71%)	34 ± 1 (88.56%)	30 ± 1 (85.89%)	9.687e-02
$SU(2)_R$ 0.15 400	12 ± 0 (100.00%)	11 ± 0 (89.22%)	10 ± 0 (88.90%)	8 ± 0 (77.58%)	2.465e-02

Table 6.6. 1-lepton channel preselection cutflow for all charged higgs signals, data and all considered simulated backgrounds when $tan\beta = 1$ at 140 fb^{-1} integrated luminosity. The weighted counts are multiplied by 2 to account for negatively charged Higgs

Cut/m_{H+}	200	300	400	600	800	1000	1200	1400	1600	1800	2000
Derivation	276354	154520	54650	14616	4707	1720	688	298	178	85	42
GRL	276354	154520	54650	14616	4707	1720	688	298	178	85	42
Good Calo	276354	154520	54650	14616	4707	1720	688	298	178	85	42
Primary vertex	276354	154520	54650	14616	4707	1720	688	298	178	85	42
Trigger	117630	68106	24538	6950	2342	893	369	162	100	48	24
Bad muon veto	117627	68102	24542	6950	2342	893	369	162	100	48	24
Jet cleaning	117145	67788	24442	6916	2332	889	367	161	99	48	24
$t\bar{t}b\bar{b}$ overlapRem	117145	67788	24442	6916	2332	889	367	161	99	48	24
$N_l^{loose} = 1$	90897	52527	18931	5227	1712	643	257	111	67	31	15
$N_l^{tight} = 1$	82971	47873	17326	4761	1557	585	233	101	60	28	14
$N_{jets} \geq 4$	52418	38796	14761	4257	1432	545	219	95	57	27	13
$N_{b\text{-jets}} \geq 3$	13291	15596	6065	1821	622	239	97	40	24	11	5
Epicycle cuts	8114	11799	5450	1784	619	238	97	40	24	11	5
preselected	8114	11799	5450	1784	619	238	97	40	24	11	5
$Jets_N \geq 5$	7013	10020	4553	1558	551	215	89	37	22	10	5
$\Delta R(l, b_2) < 2.7$	5648	8402	3792	1268	447	171	71	29	17	8	4
$m_{bb_{\Delta R_{min}}} > 100$	1722	3934	2193	842	327	129	56	24	14	7	3
SR 5j, 3b	586	1328	711	231	68	26	11	4	2	1	0
SR 5j, $\geq$ 4b	18	186	119	28	14	5	2	1	0	0	0
SR 6j, 3b	493	1013	573	216	78	29	12	5	3	1	1
SR 6j, $\geq$ 4b	92	277	145	49	22	7	3	1	1	0	0
SR 7j, 3b	458	833	466	220	103	44	19	8	5	3	1
SR $\geq$ 7j, $\geq$ 4b	75	297	179	98	43	18	10	4	2	1	1

Table 6.7. $S/\sqrt{B}$ significance of the charged higgs signals with $tan\beta = 1$ in the six Signal regions of the one-lepton channel.

Significance	H300	H400	H600	H800	H1000	H1200	H1400	H1600	H1800	H2000
SR 5j, 3b	6.47	3.46	1.13	0.33	0.13	0.05	0.02	0.01	0.0	0.0
SR 5j, $\geq$ 4b	4.66	2.99	0.7	0.34	0.13	0.04	0.02	0.01	0.01	0.0
SR 6j, 3b	5.81	3.29	1.24	0.44	0.16	0.07	0.03	0.02	0.01	0.0
SR 6j, $\geq$ 4b	6.37	3.33	1.13	0.51	0.17	0.08	0.03	0.02	0.01	0.0
SR 7j, 3b	5.32	2.97	1.4	0.66	0.28	0.12	0.05	0.04	0.02	0.01
SR $\geq$ 7j, $\geq$ 4b	5.86	3.53	1.94	0.84	0.35	0.19	0.08	0.05	0.02	0.01

Table 6.8. $S/\sqrt{S+B}$ significance of the charged higgs signals with $tan\beta = 1$ in the six Signal regions of the one-lepton channel.

Significance	H300	H400	H600	H800	H1000	H1200	H1400	H1600	H1800	H2000
SR 5j, 3b	6.37	3.43	1.12	0.33	0.13	0.05	0.02	0.01	0.0	0.0
SR 5j, $\geq$ 4b	4.41	2.88	0.7	0.34	0.13	0.04	0.02	0.01	0.01	0.0
SR 6j, 3b	5.71	3.25	1.23	0.44	0.16	0.07	0.03	0.02	0.01	0.0
SR 6j, $\geq$ 4b	5.95	3.21	1.11	0.5	0.17	0.08	0.03	0.02	0.01	0.0
SR 7j, 3b	5.23	2.95	1.4	0.66	0.28	0.12	0.05	0.04	0.02	0.01
SR $\geq$ 7j, $\geq$ 4b	5.55	3.42	1.91	0.83	0.35	0.19	0.08	0.05	0.02	0.01

Table 6.9. $S/\sqrt{S+B+\Delta B}$ significance of the charged higgs signals with $tan\beta = 1$ in the six Signal regions of the one-lepton channel.

Significance	H300	H400	H600	H800	H1000	H1200	H1400	H1600	H1800	H2000
SR 5j, 3b	6.35	3.43	1.12	0.33	0.13	0.05	0.02	0.01	0.0	0.0
SR 5j, $\geq$ 4b	4.36	2.85	0.69	0.34	0.13	0.04	0.02	0.01	0.0	0.0
SR 6j, 3b	5.7	3.25	1.23	0.44	0.16	0.07	0.03	0.02	0.01	0.0
SR 6j, $\geq$ 4b	5.89	3.18	1.1	0.5	0.16	0.07	0.03	0.02	0.01	0.0
SR 7j, 3b	5.22	2.94	1.39	0.65	0.28	0.12	0.05	0.03	0.02	0.01
SR $\geq$ 7j, $\geq$ 4b	5.5	3.39	1.89	0.83	0.35	0.19	0.08	0.05	0.02	0.01

6.8.2 1D distributions of the discriminating variables

Figure 6.12 and Figure 6.13 presents the 1D distributions of the discriminating variables comparing dark meson signals with H^+ signals. The multijet background is not included in these plots. Comparing to dark meson of similar masses, the charged Higgs signals have smaller number of jets and smaller H_T. $m_{bb_{\Delta R_{min}}}$ is the invariant mass of two b quarks that are closest to each other. $\Delta R(l, b_2)$ is the angle between the lepton in the event and the second closest b-jet to this lepton. Charged Higgs signals have slightly bigger values on $\Delta R(l, b_2)$. $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ is the mass of a R=1.2 re-clustered jet with a lepton and that of the leading all-hadronic jet. This variable is best at differentiating signal and background events, so it is used as the main variable to extract fitting results later in statistical analysis. Charged Higgs signals showed

169

smaller $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ values compared to dark meson ones. This is probably because the charged Higgs diagram has only one large Radius jet.

It is important to note that the H^+ lines do not have significant difference with that of the dark meson signals. So, there will not be significant reduction of events from particular selection requirement.

6.8.3 Statistical analysis

The statistical analysis for the H^+ search is done by profile likelihood fits over six signal regions categorized according to the jet and b-jet multiplicity : (5j, 3b), (5j,$\geq$ 4b), (6j, 3b), (6j,$\geq$ 4b), ($\geq$ 7j, 3b), ($\geq$ 7j, $\geq$ 4b). The regions are defined with a mixture of exclusive and inclusive number of jets and b-jets, for example (5j, 3b) is a region with events that exactly equal to 5 jets and equal to 3 b-jets. TRExFitter is the software package used for the fitting process, and it builds based on the packages RootFit Verkerke and Kirkby [2003b] and Minuit Hatlo et al. [2005b]. The fit is performed to the binned $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ distributions of signal and control regions in all jet and b-jet multiplicity regions simultaneously. $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ is the primary variable to perform the fit on because it provides the best separation between signal and background.

6.8.31 Statistical background. This section is a statistics introduction to the theory behind the profile likelihood fit based on an ATLAS tutorial Mironova. The goal of a fit is to solve for a set of parameters (such as the scaling factor μ) that provide the best fit to the recorded data. This process is done by maximizing the likelihood function defined as
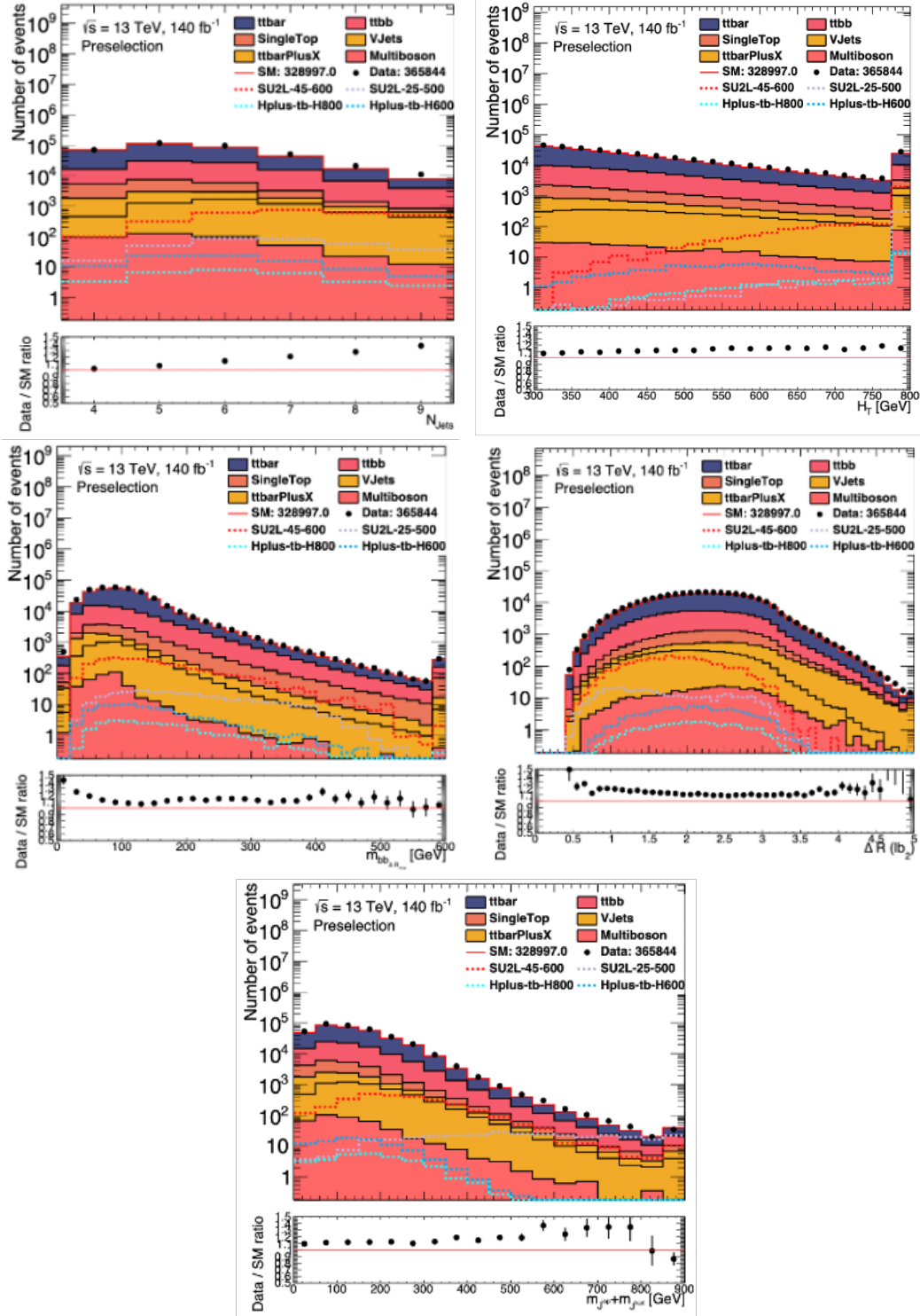
Figure 6.12. Plots of the one-lepton pre-selection variables with the last bin including overflow events. H^+ samples with $m_H^+ = 600,\ 800$ are added for comparison with the dark meson sample SU2L-25-500 and SU2L-45-600; Multijet background is not calculated at the pre-selection stage.
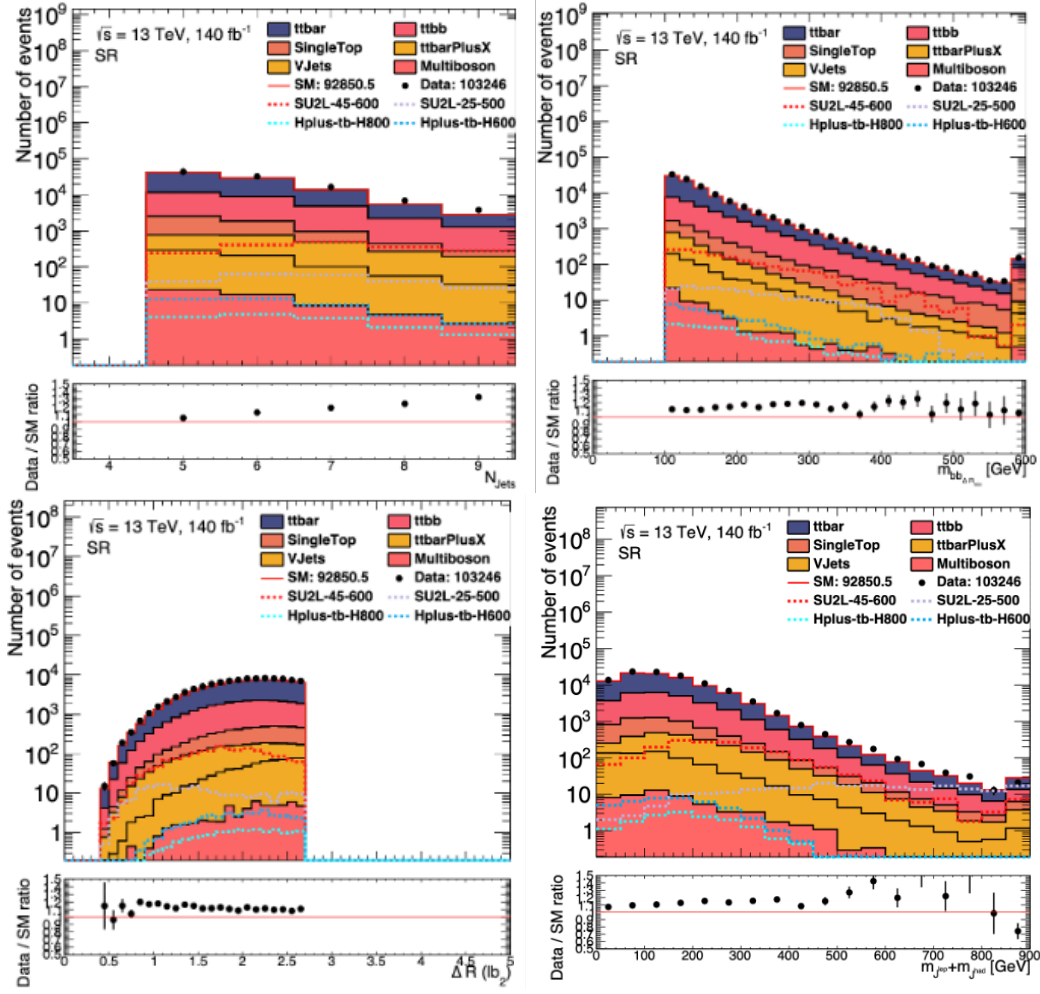
Figure 6.13. Plots of the one-lepton SR variables with the last bin including overflow events. H^+ samples with $m_H^+ = 600, 800$ are added for comparison with the dark meson sample SU2L-25-500 and SU2L-45-600; Multijet background is not calculated at the pre-selection stage.

$$L(\vec{\mu}, \vec{\theta}, \vec{\gamma}, \vec{\tau}) = \prod_{i \in bins} Pois(N_i | \vec{\mu} S_i(\vec{\theta}) + B_i(\vec{\theta}, \vec{\gamma}, \vec{\tau}) \times \prod_{\theta \in \theta} \frac{1}{\sqrt{2\pi}} e^{-\theta^2/2}$$

$$\times \prod_{i \in bins} Gauss(\beta_i | \gamma_i \beta_i, \sqrt{\gamma_i \beta_i}) \tag{6.1}$$

Where N_i is the number of data events. s_i is the number of expected signal events. b_i is the number of total expected background events. μ is the signal normalization factor that is used to scale the nominal signal yield. τ is the floating normalization factor used to scale the background yields. θ are the nuisance parameters to include systematic uncertainties in the fit. The γ nuisance parameters are used to account for uncertainties on the MC distributions, there is one parameter for the total statistical uncertainty for each bin.

6.8.32 Validation with the dark meson search. The reinterpretation aims to utilize the same fitting configuration as that used for the dark meson search. Before performing a signal plus background fit using the charged Higgs signal, fit was repeated for dark meson signals to make sure that the fitting configuration consistent with the dark meson settings.

Figure 6.14 and Figure 6.15 are examples of unblinded signal plus background pre-fit and post-fit plots with dark meson signals. The post-fit distributions show good agreement with data and the normalization factors are matching with those obtained previously. This confirms that the data and the systematics samples of the background processes are the same. And the fitting configuration is consistent with the previous settings.

6.8.33 Signal plus background fit on unblinded data. Signal plus background fit is performed on ten H^+ signal samples. Here presents results from
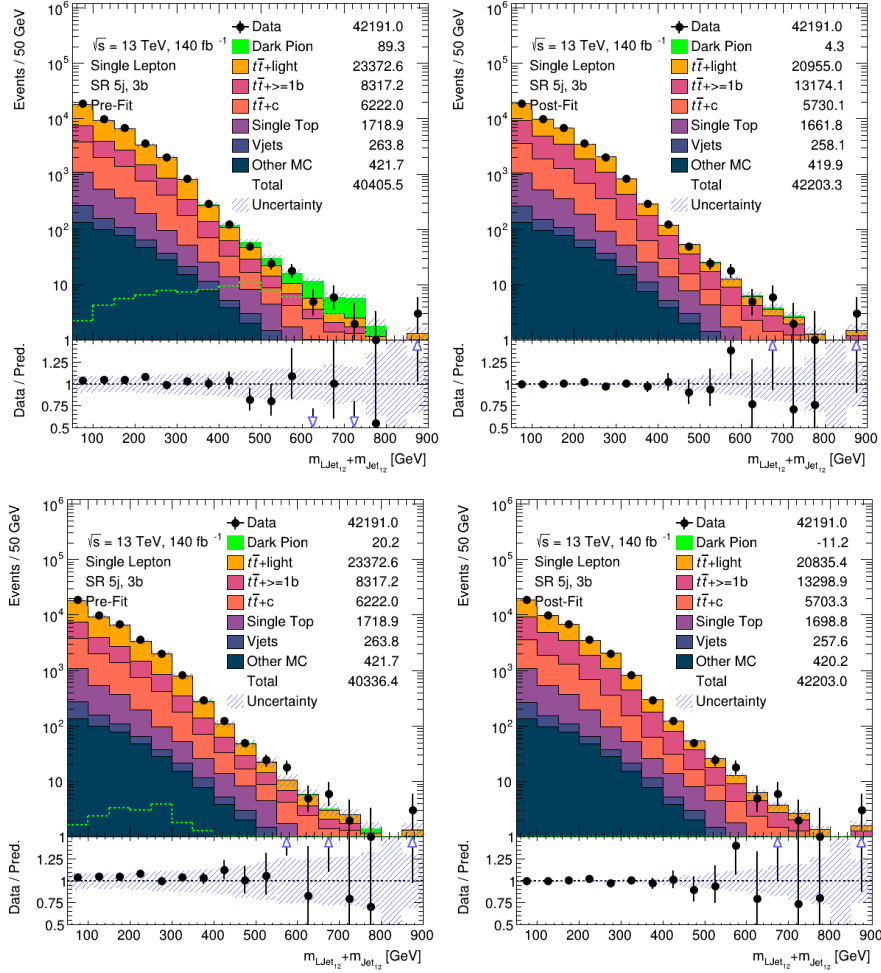
Figure 6.14. Comparison of pre-fit (left) and post-fit (right) $m_{\mathbb{J}^{had}}+m_{\mathbb{J}^{lep}}$ distributions in the SRs. The upper row shows the SR with five jets and exactly three b-jets for dark meson signal 25 400 and the bottom row shows that of the dark meson signal 35 700. The uncertainty bands contain all statistical and systematic uncertainties.

174
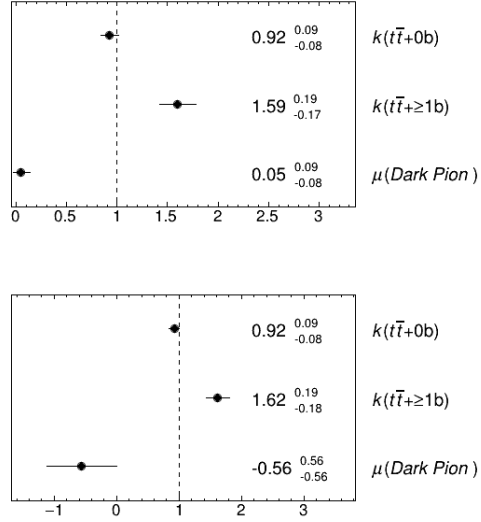
Figure 6.15. Normalization factors for the two fitted $t\bar{t}$ background categories and the dark meson signals of mass 400 GeV, $\eta = 0.25$ (left) and 700 GeV, $\eta = 0.35$ (right), as determined in the signal plus background fit in the signal and control regions of the 1-lepton channel.

the sample with H^+ mass 300 GeV. Figure 6.16 and 6.17 are the pre-fit and post-fit distributions of $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ variable with signal line overlaid on top of the background distributions. The lower panels are ratio plots showing how the predicted compared to recorded data. The signal sample is stacked on top of the background samples to clearly show the total amount of simulated events in 6.18 and 6.19. Comparing the lower panels of the left and right column, it is easy to see that the total simulated distribution is well matched with data after the fitting process. That suggests that the fitting behaves as expected. However, the charged Higgs signal $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ distributions are very similar to that of the total background distributions. This is an undesired feature for the fitting process because this could make it very difficult for the fitter to differentiate signal from the background distribution.

Figure 6.20 presents summary plots showing total number of data, background and signal events in each bin of the fit.

175
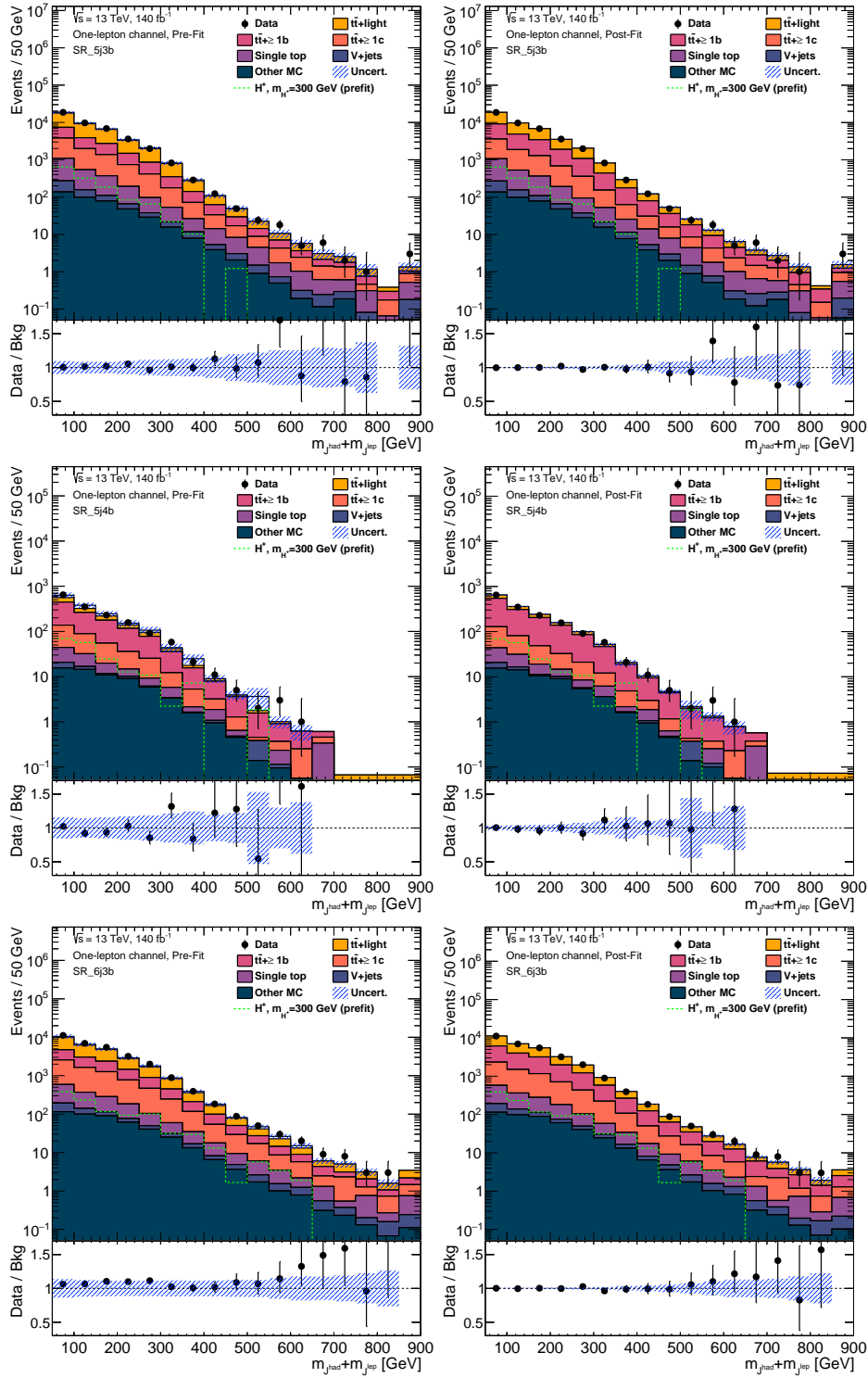
Figure 6.16. Comparison of pre-fit (left) and post-fit (right) $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ distributions in the **signal regions** with H^+ sample of mass 300 GeV (overlaid). The upper row shows the SR with five jets and exactly three bjets, the middle row the SR containing events with five jets and at least four bjets. The bottom row is the SR with exactly six jets and three bjets. The uncertainty bands contain all statistical and systematic uncertainties.
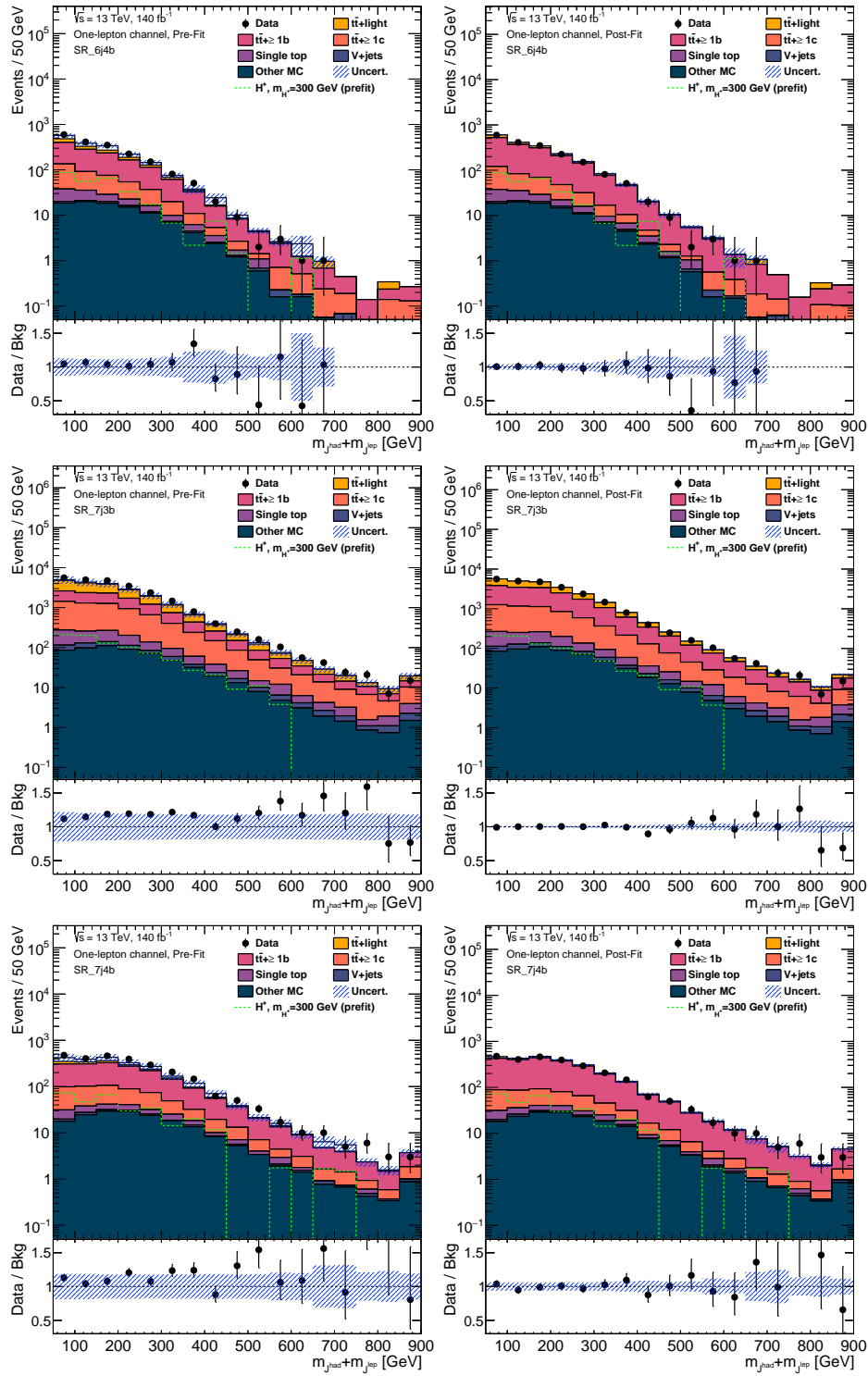
Figure 6.17. Comparison of pre-fit (left) and post-fit (right) $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$ distributions in the **signal regions** with H^+ sample of mass 300 GeV (overlaid). The upper row shows the SR with six jets and at least four bjets, the middle row the SR containing events with at least seven jets and exactly three bjets. The bottom row is the SR with at least seven jets and at least four bjets. The uncertainty bands contain all statistical and systematic uncertainties.

177
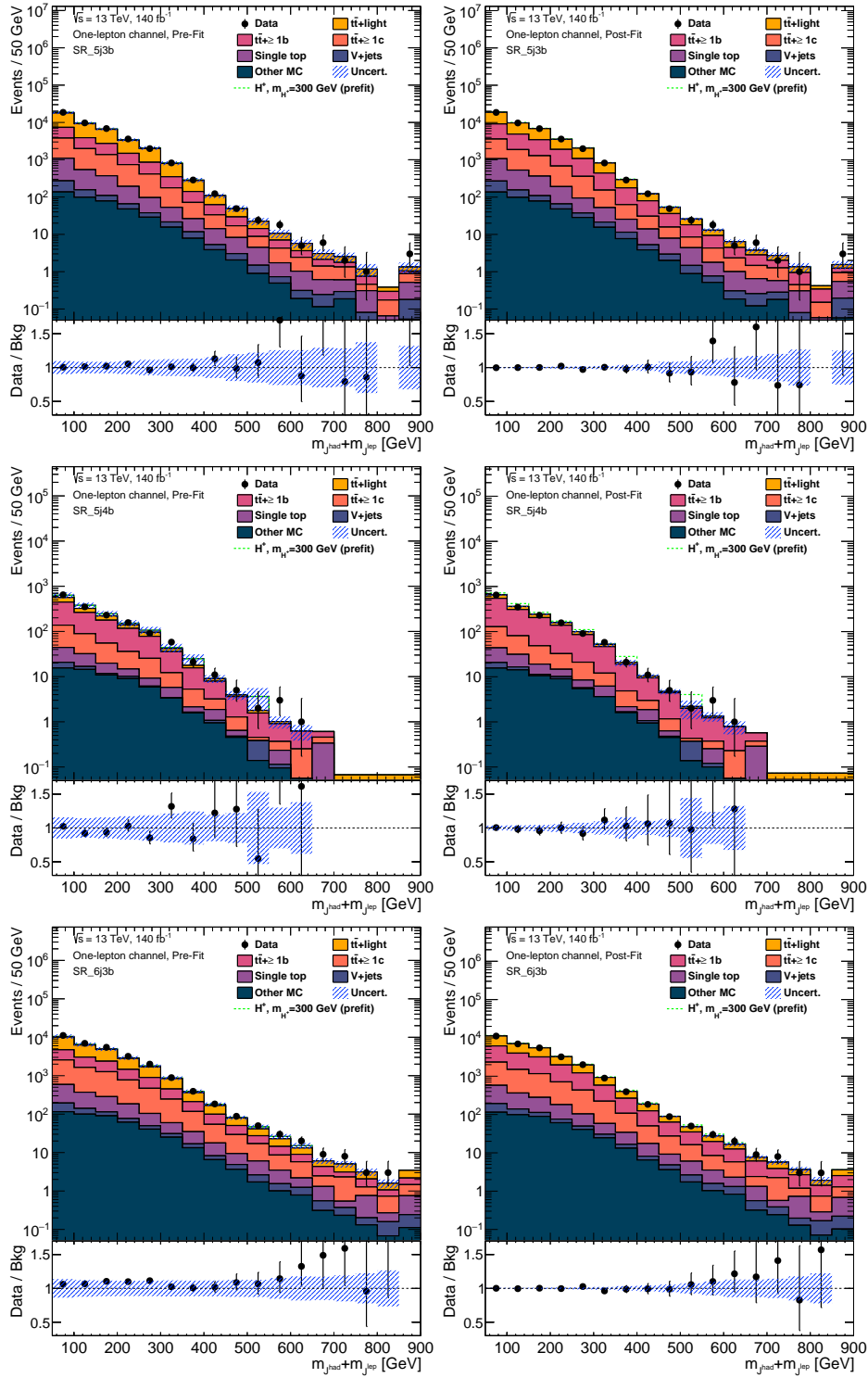
Figure 6.18. Comparison of pre-fit (left) and post-fit (right) $m_{\mathbb{J}^{had}}+m_{\mathbb{J}^{lep}}$ distributions in the **signal regions** with H^+ sample of mass 300 GeV stacked. The upper row shows the SR with five jets and exactly three bjets, the middle row the SR containing events with five jets and at least four bjets. The bottom row is the SR with exactly six jets and three bjets. The uncertainty bands contain all statistical and systematic uncertainties.

178
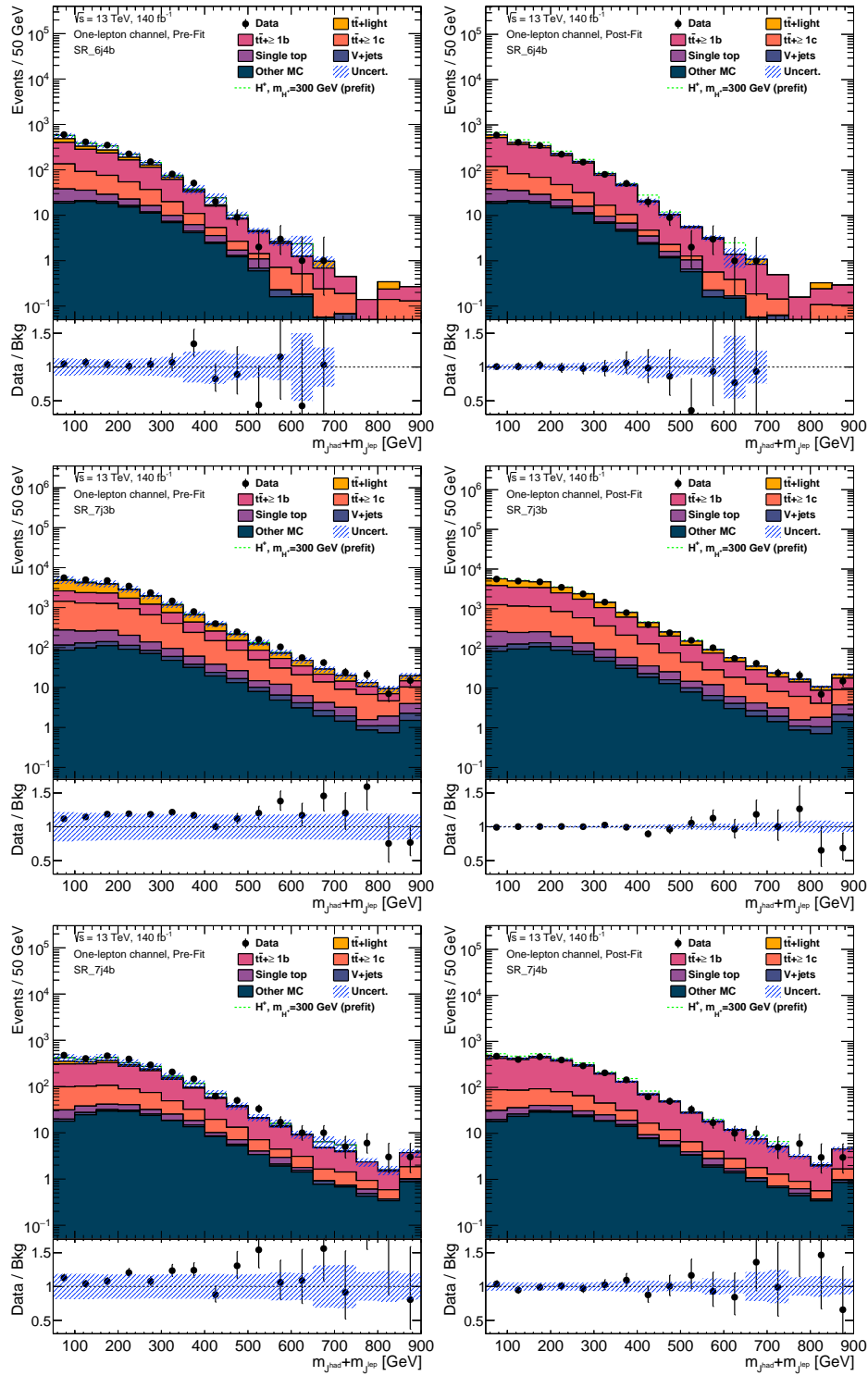
Figure 6.19. Comparison of pre-fit (left) and post-fit (right) $m_{\mathbb{J}had}+m_{\mathbb{J}lep}$ distributions in the **signal regions** with H^+ sample of mass 300 GeV stacked. The upper row shows the SR with six jets and at least four bjets, the middle row the SR containing events with at least seven jets and exactly three bjets. The bottom row is the SR with at least seven jets and at least four bjets. The uncertainty bands contain all statistical and systematic uncertainties.
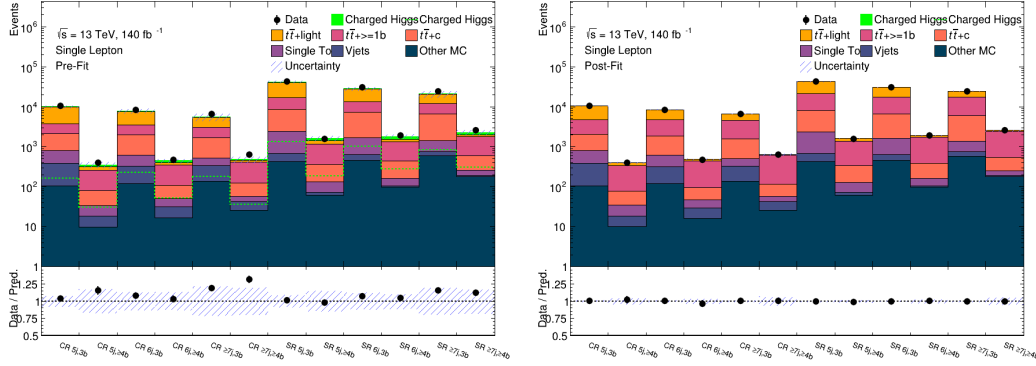
Figure 6.20. Comparison of event yields in all signal, control and validation regions before fitting (left) and after fitting (right). The yields shown here are the sum over all bins in the corresponding regions. Signal scale factors are applied to the post-fit signal line.

Figure 6.21 presents the pull-plots from both instrumental and theoretical nuisance parameters (NPs). They are varied during the likelihood minimization and provide additional degree of freedom for modifying the fitting distribution. The pull plots show the post fit values of the nuisance parameters. None of the NPs is highly pulled or constrained.

A correlation matrix shows how much the movement of a NP influence the other. A strong correlation between two NP means they are not moved independently. Figure 6.22 is the correlation matrix when fitting with the 300 GeV signal sample. Comparing to the dark meson results, the correlation diagram does not differ significantly.

Finally, Figure 6.23 shows the set of normalization factors that provide the best fit to the data. They include two types of background and the 300 GeV charged Higgs sample. The very small magnitude of the signal strength is consistent with background only hypothesis, suggesting that a charged Higgs signal of 300 GeV is not likely to exist in experimental data.
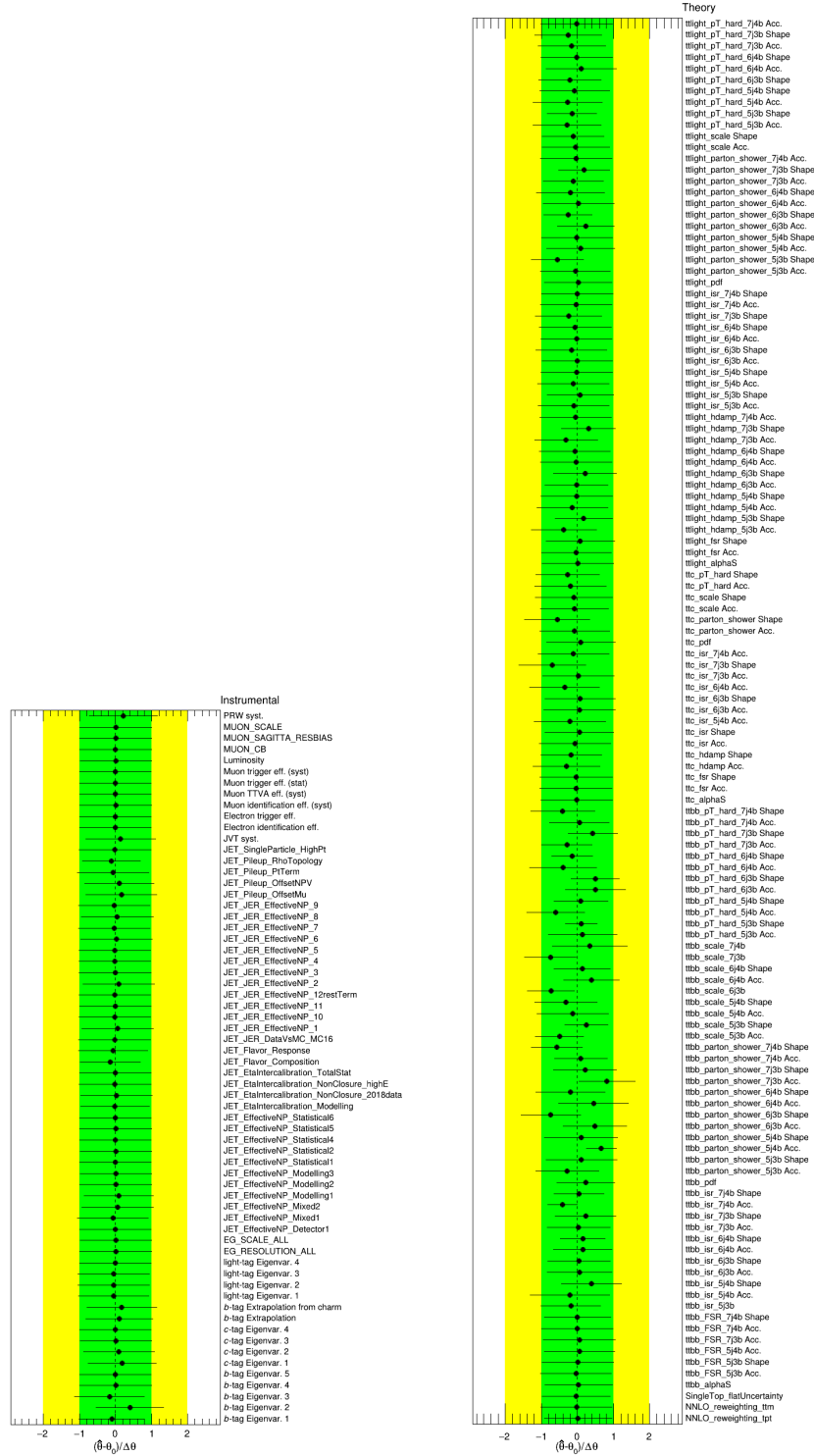
Figure 6.21. Pulls and constraints on the nuisance parameters of the un-blinded 300 GeV charged Higgs signal plus background fit in the control regions and SRs of the 1-lepton channel.

Figure 6.22. Correlation matrix of the signal-plus-background unblinded fit in the control regions and SRs of the 1-lepton channel, for a signal with charged Higgs sample of mass 300 GeV; Shown are only parameters that have at least a 20% correlation with any of the other nuisance parameters. All shown numbers are given in percentages.
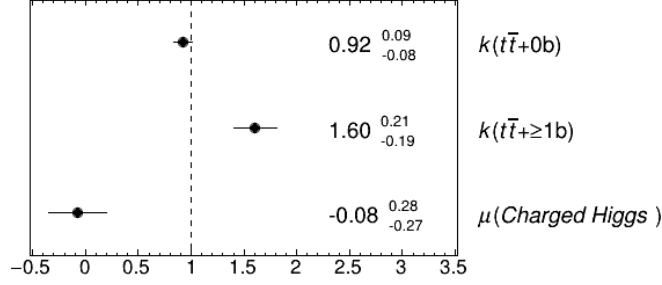
Figure 6.23. Normalization factors for the two fitted $t\bar{t}$ background categories and the 300 GeV H^+signal as determined in the signal plus background fit in the signal and control regions of the 1-lepton channel.

6.8.34 Limit setting. Finally, the exclusion limit in Figure 6.24 can be extracted from fitting results of all eleven signal simulations of H^+. And the limit plot is shown below. All the observed points are within 2 σ deviation from the expected points. Traditionally, only a deviation larger than 3 σ is considered significant in a wide range of particle physics experiments. Therefore, no significant excess above the expected Standard Model background is found and observed (expected) upper limits at 95% confidence level are set on the pp $\rightarrow tbH^+$ production cross-section. Signals with H+ masses from 200 to 475 GeV are excluded...

6.8.4 One-lepton channel conclusion and outlook

In summary, charged Higgs signals have comparable number of events after applying one-lepton selection criteria from the dark meson analysis. m_H^+ 200 to 475 GeV are excluded at 95% confidence level. For one-lepton channel, I think the most challenging part comes from the fitting process. In future iteration of the analysis, a different variable, other than $m_{\mathbb{J}^{had}} + m_{\mathbb{J}^{lep}}$, should be explored in the fitting process. If it gives a better distinction between the signal and the background. This would
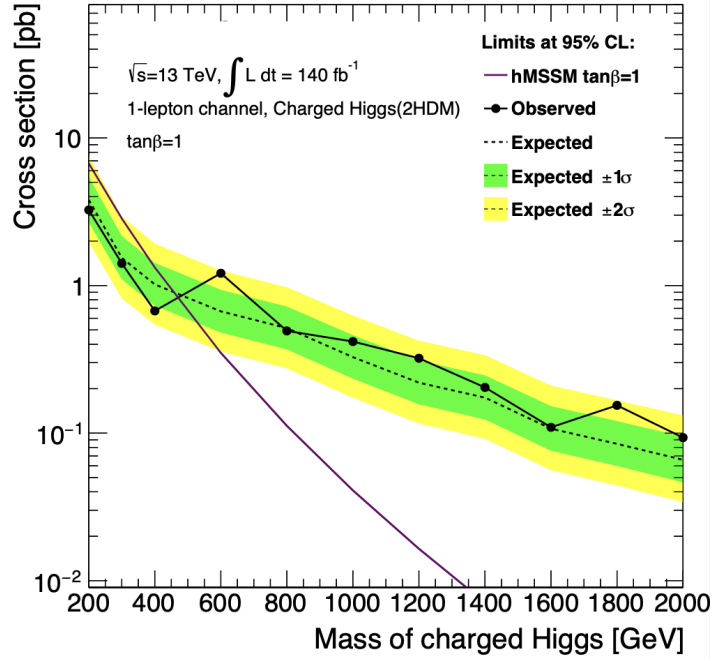
Figure 6.24. Observed and expected upper limits to produce $H^+ \rightarrow tb$ in association with a top quark and a bottom quark. The bands surrounding the expected limit show the 68% and 95% confidence intervals. Theory predictions are shown for one representative values of $\tan \beta$ in the hMSSM benchmark scenario.

help the fit to distinguish signals of heavier charged Higgs with smaller statistics and

potentially expand the final exclusion region.

CHAPTER VII

In summary, this thesis presents a project about trigger update for the future ATLAS trigger and a reinterpretation to search for the hypothetical charged Higgs. A previously published analysis for dark meson was used as the framework to perform the reinterpretation. The dark meson analysis is found to be not sensitive to charged Higgs signals in the all-hadronic channel, while it is sensitive in the one-lepton channel. A statistical analysis is presented and exclusion region was extracted. No excess of events are observed beyond expected standard model events. Although the current results do not exceed the performance of the previous charged Higgs search, further adjustments on the selection criteria and the fitting process can be made to enlarge the exclusion region.

Appendix A

LIST OF CHARGED HIGGS SIGNAL SAMPLES

This appendix includes the uniquely identifiable names of charged Higgs signal samples.

- mc16_13TeV.451290.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO _Hplus4FS_H200_tb.deriv.DAOD_TOPQ1.
  e7834_e5984_s3126_r10201_r10210_p6628_tid43254581_00

- mc16_13TeV.451294.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H300_tb.deriv.DAOD_TOPQ1.
  e7834_e5984_s3126_r10724_r10726_p6628_tid43254583_00

- mc16_13TeV.450595.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H400_tb.deriv.DAOD_TOPQ1.
  e7429_e5984_s3126_r10201_r10210_p6628_tid43254579_00

- mc16_13TeV.450597.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H600_tb.deriv.DAOD_TOPQ1.
  e7429_a875_r10201_p4514

- mc16_13TeV.450002.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H800_tb.deriv.DAOD_TOPQ1.
  e7137_e5984_a875_r10201_r10210_p4514

- mc16_13TeV.450004.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H1000_tb.deriv.DAOD_TOPQ1.
  e7137_a875_r10201_p4514

- mc16_13TeV.450598.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H1200_tb.deriv.DAOD_TOPQ1.
  e7429_a875_r10201_p4514

- mc16_13TeV.450599.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H1400_tb.deriv.DAOD_TOPQ1.
  e7429_a875_r10201_p4514

- mc16_13TeV.450600.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H1600_tb.DAOD_TOPQ1.
  e7429_s3126_r10724_p4514

- mc16_13TeV.450601.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H1800_tb.deriv.DAOD_TOPQ1.
  e7429_e5984_s3126_r10724_r10726_p6109_tid37809143_00

- mc16_13TeV.450602.aMcAtNloPythia8EvtGen
  _A14NNPDF23LO_Hplus4FS_H2000_tb.deriv.DAOD_TOPQ1.
  e7429_e5984_s3126_r10724_r10726_p6109_tid37809140_00

Appendix B

RAW CUTFLOW TABLES

This appendix includes raw event counts of the H^+ signals when analyzed with both all hadronic and one lepton channel

Table B.1. Raw event counts of the all-hadronic preselection cutflow for charged Higgs signals of different masses

Cut	H600	H1200	H1400	H1600	H1800	H2000
Derivation	350000	314000	320000	419000	420000	420000
GRL	350000	314000	320000	419000	420000	420000
Good Calo	350000	314000	320000	419000	420000	420000
Primary vertex	350000	314000	320000	419000	420000	420000
Trigger	14595	183870	233725	342029	365797	380245
Bad muon veto	14594	183825	233662	341937	365689	380133
Jet cleaning	14520	182966	232485	340182	363926	378289
$H_T > 1150$ GeV	6437	125355	187324	299842	335670	358066
$N_{\text{jets}} \geq 6$	5760	108986	159624	253553	282915	301772
$N_{\text{jets, R=1.2}} \geq 2$	5537	105810	153364	241827	268958	286929
$N_{\text{b-jets}} \geq 3$	3405	56616	78073	116459	123906	127365
electron veto	2481	41894	58402	87481	94020	97327
muon veto	1589	27189	38942	59502	65268	68432
$m^{1,2}_{\text{jet},R=1.2} > 190$ GeV	181	3073	4927	9198	11767	14046
$p^6_{T,\text{jet}} > 25$ GeV	178	3013	4845	9007	11526	13740
preselected	178	3013	4845	9007	11526	13740
$lead.m_{bb}/p_{T,bb} > 0.25$	166	2746	4425	8138	10170	12078
$sub.m_{bb}/p_{T,bb} > 0.25$	160	2595	4098	7481	9210	10828
$lead.m_{\text{jet,R=1.2}} > 300 GeV$	55	848	1548	3278	4440	5702
$sub.m_{\text{jet,R=1.2}} > 250 GeV$	30	447	868	1861	2761	3635
$lead.\Delta R\,(j, b_2) < 1.0$	10	147	292	571	731	894
$sub.\Delta R\,(j, b_2) < 1.0$	2	31	57	107	118	148

Appendix C

WEIGHTED DARK MESON CUTFLOW TABLES

This appendix includes number of background and dark meson signal events after each selection The tables are taken from Ref. Gledhill et al. [2020].

C.0.1 all-hadronic channel

C.0.2 one-lepton channel

Table C.1. All-hadronic preselection cutflow for data and all considered simulated backgrounds at 140 fb^{-1} integrated luminosity.Gledhill et al. [2020]

Cut	data	All MC	$t\bar{t}$ allhad	$t\bar{t}$ nonallhad	V+jets	$t\bar{t} + X$	single top	multiboson	Signal-25-500
Derivation	559740475	74686277.3	700250.6	485751.5	52752296.7	234104.3	20387641.6	126232.6	2650.2
GRL	417309494	74686277.3	700250.6	485751.5	52752296.7	234104.3	20387641.6	126232.6	2650.2
Good Calo	412694354	74686277.3	700250.6	485751.5	52752296.7	234104.3	20387641.6	126232.6	2650.2
Primary vertex	412692263	74686277.3	700250.6	485751.5	52752296.7	234104.3	20387641.4	126232.6	2650.2
Trigger	205205611	3413076.8	650091.2	419676.5	2188774.2	21031.8	132246.8	1256.3	2464.4
Bad muon veto	205203444	3412881.3	650081.8	419566.8	2188717.7	21028.4	132231.3	1255.3	2464.2
Jet cleaning	202840649	3401193.6	647806.6	418105.5	2181341.8	20951.7	131736.3	1251.7	2457.2
$H_T > 1150$ GeV	87703211	1701555	399278.5	258393.7	970739.3	10700.4	61920.1	523.0	2270.1
$N_{\text{jets}} \geq 6$	45304007	1207437.7	363798.8	207361.1	581846.0	9943.0	44267.4	221.4	2039.9
$N_{\text{jets, R=1.2}} \geq 2$	45005443	1196438.1	363209.0	201599.1	578272.9	9792.5	43347.0	217.6	1996.0
$N_{\text{b-jets}} \geq 3$	576239	89564.5	48886.3	24779.6	8697.4	2696.3	4484.4	20.5	1340.9
electron veto	568304	80167.3	82489.7	18600.3	8682.5	2359.6	4094.0	20.5	1133.9
muon veto	561409	74916.9	48686.5	12526.6	8667.2	2048.2	3707.3	20.5	941.6
$m^{1,2}_{\text{jet}, R=1.2} > 190$ GeV	69879	82489.7	7205.5	1660.0	1734.0	409.0	384.2	0.8	560.0
$p^6_{T,\text{jet}} > 25$ GeV	67339	11220.6	7127.8	1629.1	1687.4	405.6	375.0	0.8	549.2
preselected	67339	11220.6	7127.8	1629.1	1687.4	405.6	375.0	0.8	549.2
lead. $m_{bb}/p_{T,bb} > 0.25$	59299	10048	6364.7	1472.5	1507.8	365.1	337.1	0.8	492.9
sub. $m_{bb}/p_{T,bb} > 0.25$	52917	9075.8	5782.9	1309.9	1347.9	330.3	304.3	0.4	449.9
lead. $m_{\text{jet,R=1.2}} > 300$ GeV	20018	3447.5	2156.6	514.2	510.0	138.7	128.0	0.0	363.6
sub. $m_{\text{jet,R=1.2}} > 250$ GeV	11444	2040	1234.5	302.5	343.6	83.3	76.0	0.0	302.7
lead. $\Delta R(j, b_2) < 1.0$	3365	575.7	411.1	63.9	52.5	26.5	21.7	0.0	167.0
sub. $\Delta R(j, b_2) < 1.0$	234	46.2	28.3	7.9	5.0	3.4	1.5	0.0	66.6

191

Table C.2. one-lepton channel preselection cutflow for data and all considered simulated backgrounds at 140 fb^{-1} integrated luminosity. **Note:** There is an inconsistency of this table with the values given in table 6.5. The values in table 6.5 are correct, the reason for the mismatch is that some files have been crashing when reading the leptonSF weight in the cutflow building. We are investigating, the mismatch however is small ($< 1\%$, note that for $t\bar{t}$ the values in table 6.5 are after the NNLO reweighting, the values given here are before). Gledhill et al. [2020]

Cut	data	All MC	$t\bar{t}$	$t\bar{t}b\bar{b}$	V+jets	$t\bar{t}+X$	single top	multiboson	Signal-25-500
Derivation	4347158651	4449107658.8	104993139.7	4014381.1	4316572778.5	185274.0	18613595.3	4728490.3	2450.4
GRL	4249705009	4449107658.8	104993139.7	4014381.1	4316572778.5	185274.0	18613595.3	4728490.3	2450.4
Good Calo	4247341451	4449107658.8	104993139.7	4014381.1	4316572778.5	185274.0	18613595.3	4728490.3	2450.4
Primary vertex	4247339918	4449106386.9	104993139.2	4014381.1	4316571507.1	185274.0	18613595.2	4728490.3	2450.4
Trigger	3166828982	2780217810.8	25367729.3	967237.7	2745045665.1	60180.9	5502868.2	3274129.7	1092.3
Bad muon veto	3166691696	2780150073.9	25367271.1	967220.4	2744978584.6	60179.7	5502768.9	3274049.2	1092.3
Jet cleaning	3132324418	2771444941.7	25262502.8	963585.8	2736419885.1	59922.1	5479419.9	3259625.9	1088.5
$t\bar{t}b\bar{b}$ overlap removal	3132324418	2770592223.1	24665456.0	707914.0	2736419885.1	59922.1	5479419.9	3259625.9	1088.5
$N_l^{\text{loose}} = 1$	2306534763	2514142564.2	19195491.9	510214.2	2486613701.5	43144.9	4730259.9	3049751.9	603.2
$N_l^{\text{tight}} = 1$	1798122804	2218596399.9	17496673.3	459568.5	2193489534.8	39260.7	4311915.3	2799447.2	549.2
$N_{\text{jets}} \geq 4$	22092576	23276233.1	9461911.1	322578.8	12460786.6	34203.4	698734.1	298019.1	539.5
$N_{\text{b-jets}} \geq 3$	615481	572020.0	427609.3	106588.1	9733.9	5883.2	21599.1	606.4	333.7
Epicycle cuts	365844	341994.5	240123.8	78120.5	5144.4	5238.1	12965.6	402.2	333.7
preselected	365844	341994.5	240123.8	78120.5	5144.4	5238.1	12965.6	402.2	333.7

Bibliography

G.Shaw B.R.Martin. Particle physics. 2008.

M.Thomson. *Modern Particle Physics.* Cambridge University Press, 2013.

G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, Marc Sher, and João P. Silva. Theory and phenomenology of two-higgs-doublet models. *Physics Reports*, 516 (1–2):1–102, July 2012. ISSN 0370-1573. doi: 10.1016/j.physrep.2012.02.002. URL `http://dx.doi.org/10.1016/j.physrep.2012.02.002`.

Heather E. Logan and Deanna MacLennan. Charged higgs phenomenology in the flipped two-higgs-doublet model. *Physical Review D*, 81(7), April 2010. ISSN 1550-2368. doi: 10.1103/physrevd.81.075016. URL `http://dx.doi.org/10.1103/PhysRevD.81.075016`.

Search for dark mesons decaying to top and bottom quarks in proton-proton collisions at $\sqrt{\mathrm{s}} = 13$ TeVwith the atlas detector. *Journal of High Energy Physics*, 2024(9), September 2024. ISSN 1029-8479. doi: 10.1007/jhep09(2024)005. URL `http://dx.doi.org/10.1007/JHEP09(2024)005`.

Cinzia De Melis. The CERN accelerator complex. Complexe des accélérateurs du CERN. 2016. URL `https://cds.cern.ch/record/2119882`. General Photo.

Apollinari G., Béjar Alonso I., Brüning O., Fessia P., Lamont M., Rossi L., and Tavian L. *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1.* CERN Yellow Reports: Monographs. CERN, Geneva, 2017. doi: 10.23731/CYRM-2017-004. URL `https://cds.cern.ch/record/2284929`.

Riccardo Maria Bianchi and ATLAS Collaboration. Images from the paper: "The ATLAS Experiment at the CERN Large Hadron Collider: A Description of the Detector Configuration for Run 3". 2024. URL `https://cds.cern.ch/record/2900788`. General Photo.

The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008a. doi: 10.1088/1748-0221/3/08/S08003. URL `https://cds.cern.ch/record/1129811`. Also published by CERN Geneva in 2010.

ATLAS pixel detector electronics and sensors. *JINST*, 3:P07007, 2008b. doi: 10.1088/1748-0221/3/07/P07007. URL `https://cds.cern.ch/record/1119279`.

Calorimaters (energy measurement) desy. URL `https://www.desy.de/~garutti/LECTURES/ParticleDetectorSS12/L10_Calorimetry.pdf`.

Performance of the ATLAS Trigger System in 2015. Performance of the ATLAS Trigger System in 2015. *Eur. Phys. J. C*, 77(5):317, 2017. doi: 10.1140/epjc/s10052-017-4852-3. URL `https://cds.cern.ch/record/2235584`. 77 pages in total, author list starting page 61, 50 figures, 1 table. Published in Eur. Phys. J. C. All figures including auxiliary figures are available at http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/TRIG-2016-01/.

The atlas trigger system for lhc run 3 and trigger performance in 2022. *Journal of Instrumentation*, 19(06):P06029, June 2024. ISSN 1748-0221. doi: 10.1088/1748-0221/19/06/p06029. URL `http://dx.doi.org/10.1088/1748-0221/19/06/P06029`.

Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System. Technical report, CERN, Geneva, 2017. URL `https://cds.cern.ch/record/2285584`.

Joao Pequenao and Paul Schaffner. How ATLAS detects particles: diagram of particle paths in the detector. 2013. URL `https://cds.cern.ch/record/1505342`.

T Gleisberg, S Höche, F Krauss, M Schönherr, S Schumann, F Siegert, and J Winter. Event generation with sherpa 1.1. *Journal of High Energy Physics*, 2009(02): 007–007, February 2009. ISSN 1029-8479. doi: 10.1088/1126-6708/2009/02/007. URL `http://dx.doi.org/10.1088/1126-6708/2009/02/007`.

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

Phase II upgrade TDAQ TDR . URL `https://cds.cern.ch/record/2285584/files/ATLAS-TDR-029.pdf`.

Segmentation of the barrel region of the EM LAr . URL `https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08003`.

Graham D. Kribs, Adam Martin, Bryan Ostdiek, and Tom Tong. Dark Mesons at the LHC. *JHEP*, 07:133, 2019a. doi: 10.1007/JHEP07(2019)133.

J. M. Butterworth, L. Corpe, X. Kong, S. Kulkarni, and M. Thomas. New sensitivity of LHC measurements to composite dark matter models. *Phys. Rev. D*, 105(1): 015008, 2022. doi: 10.1103/PhysRevD.105.015008.

Search for charged higgs bosons decaying into a top quark and a bottom quark at $\sqrt{s}$ = 13 TeVwith the atlas detector. *Journal of High Energy Physics*, 2021(6), June 2021. ISSN 1029-8479. doi: 10.1007/jhep06(2021)145. URL `http://dx.doi.org/10.1007/JHEP06(2021)145`.

Search for charged higgs bosons decaying into top and bottom quarks at $\sqrt{s} = 13$ TeVwith the atlas detector. *Journal of High Energy Physics*, 2018(11), November 2018a. ISSN 1029-8479. doi: 10.1007/jhep11(2018)085. URL `http://dx.doi.org/10.1007/JHEP11(2018)085`.

CMS Collboration. Search for charged higgs bosons decaying into a top and a bottom quark in the all-jet final state of pp collisions at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2020(7), July 2020a. ISSN 1029-8479. doi: 10.1007/jhep07(2020)126. URL `http://dx.doi.org/10.1007/JHEP07(2020)126`.

Summary plots for beyond Standard Model Higgs boson benchmarks for direct and indirect searches. Technical report, CERN, Geneva, 2024. URL `https://cds.cern.ch/record/2898861`. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2024-008.

STEPHEN P. MARTIN. *A SUPERSYMMETRY PRIMER*, page 1–98. WORLD SCIENTIFIC, July 1998. doi: 10.1142/9789812839657_0001. URL `http://dx.doi.org/10.1142/9789812839657_0001`.

Galen Rhodes Gledhill, Olga Sunneborn Gudnadottir, Jochen Jens Heinrich, Stephanie Majewski, Rebeca Gonzalez Suarez, Giulia Ripellino, Anni Xiong, Timothy Thankachen Mathew, and Federica Piazza. Search for dark mesons

decaying to top and bottom quarks with the ATLAS detector in 139 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2020. URL `https://cds.cern.ch/record/2711578`.

Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1). ISSN 0370-2693. doi: 10.1016/j.physletb.2012.08.020. URL `http://dx.doi.org/10.1016/j.physletb.2012.08.020`.

Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, September 2012. ISSN 0370-2693. doi: 10.1016/j.physletb.2012.08.021. URL `http://dx.doi.org/10.1016/j.physletb.2012.08.021`.

Massimo Persic, Paolo Salucci, and Fulvio Stel. The Universal rotation curve of spiral galaxies: 1. The Dark matter connection. *Mon. Not. Roy. Astron. Soc.*, 281:27, 1996. doi: 10.1093/mnras/278.1.27.

P. Salucci, A. Lapi, C. Tonini, G. Gentile, I. Yegorova, and U. Klein. The universal rotation curve of spiral galaxies - ii. the dark matter distribution out to the virial radius. *Monthly Notices of the Royal Astronomical Society*, 378(1):41–47, June 2007. ISSN 1365-2966. doi: 10.1111/j.1365-2966.2007.11696.x. URL `http://dx.doi.org/10.1111/j.1365-2966.2007.11696.x`.

V. C. Rubin, W. K. Ford, Jr., and N. Thonnard. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc). *The Astrophysical Journal*, 238:471–487, June 1980. doi: 10.1086/158003.

Douglas Clowe, Maruša Bradač, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *The Astrophysical Journal*, 648(2):L109–L113, August 2006. ISSN 1538-4357. doi: 10.1086/508162. URL `http://dx.doi.org/10.1086/508162`.

Richard Massey, Thomas Kitching, and Johan Richard. The dark matter of gravitational lensing. *Reports on Progress in Physics*, 73(8):086901, July 2010. ISSN 1361-6633. doi: 10.1088/0034-4885/73/8/086901. URL `http://dx.doi.org/10.1088/0034-4885/73/8/086901`.

Alison L. Coil. *The Large-Scale Structure of the Universe*, page 387–421. Springer Netherlands, 2013. ISBN 9789400756090. doi: 10.1007/978-94-007-5609-0_8. URL `http://dx.doi.org/10.1007/978-94-007-5609-0_8`.

G. Efstathiou, J. R. Bond, and S. D. M. White. COBE background radiation anisotropies and large-scale structure in the universe. *Monthly Notices of the Royal Astronomical Society*, 258(1):1P–6P, September 1992. doi: 10.1093/mnras/258.1.1P.

Graham D. Kribs and Ethan T. Neil. Review of strongly-coupled composite dark matter models and lattice simulations. *Int. J. Mod. Phys. A*, 31(22):1643004, 2016. doi: 10.1142/S0217751X16430041.

Graham D. Kribs, Adam Martin, Bryan Ostdiek, and Tom Tong. Dark mesons at the lhc. *Journal of High Energy Physics*, 2019(7), July 2019b. ISSN 1029-8479. doi: 10.1007/jhep07(2019)133. URL `http://dx.doi.org/10.1007/JHEP07(2019)133`.

Lyndon R Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001, 2008. doi: 10.1088/1748-0221/3/08/S08001. URL https://cds.cern.ch/record/1129806. This report is an abridged version of the LHC Design Report (CERN-2004-003).

H H J ten Kate. Superconducting magnet system for the ATLAS detector at CERN. *IEEE Trans. Appl. Supercond.*, 9(2, pt.1):841–846, 1999. doi: 10.1109/77.783428. URL https://cds.cern.ch/record/409763.

Heinz Pernegger. The Pixel Detector of the ATLAS Experiment for LHC Run-2. Technical Report 06, CERN, Geneva, 2015. URL https://cds.cern.ch/record/1985432.

C. W. Fabjan and F. Gianotti. Calorimetry for particle physics. *Rev. Mod. Phys.*, 75:1243–1286, 2003. doi: 10.1103/RevModPhys.75.1243.

M. Abolins et al. The ATLAS Data Acquisition and High Level Trigger system. *JINST*, 11(06):P06008, 2016. doi: 10.1088/1748-0221/11/06/P06008.

A. Ruiz Martínez. The Run-2 ATLAS Trigger System. *J. Phys. Conf. Ser.*, 762(1): 012003, 2016. doi: 10.1088/1742-6596/762/1/012003.

Phase II upgrade TDAQ TDR . URL https://cds.cern.ch/record/2285584/files/ATLAS-TDR-029.pdf.

S. Ellis. "What is a Jet?". "What is a Jet?". 2010. URL https://cds.cern.ch/record/1276292.

L1Calo eFEX tau trigger algorithm by Nicholas. URL https://gitlab.cern.ch/atlas/athena/-/blob/master/Trigger/TrigT1/L1CaloFEX/L1CaloFEXSim/src/eFEXtauAlgo.cxx.

L1Calo eFEX tau trigger algorithm by Denis. URL `https://gitlab.cern.ch/atlas/athena/-/blob/master/Trigger/TrigT1/L1CaloFEX/TrigT1CaloFexPerf/src/EFexTauAlgorithm.cxx`.

L1CALO conference talk by Nicholas. URL `https://indico.cern.ch/event/867020/contributions/3726146/attachments/2003208/3344698/L1CALOJoint03132020.pdf`.

run4 simulation framework twiki page. URL `https://twiki.cern.ch/twiki/bin/viewauth/Atlas/GlobalTriggerPerformance#Default_algorithms`.

Run4 GEP simulation framework git repository. URL `https://gitlab.cern.ch/atlas/athena/-/tree/21.9/Trigger/TrigL0GepPerf`.

Presentation by Maria Mazza about the Run 4 simulation framework. URL `https://indico.cern.ch/event/1008889/contributions/4279982/attachments/2215109/3749858/TDAQ_20210324.pdf`.

Athena Tau Trigger for HL LHC . URL `https://gitlab.cern.ch/axiong/athena/-/blob/21.9.26/Trigger/TrigL0GepPerf/src/GepTauAlg.cxx`.

ztautau signal sample. URL `https://rucio-ui.cern.ch/did?scope=mc15_14TeV&name=mc15_14TeV.800281.Py8EG_AZ_Ztautau.recon.AOD.e8379_s3735_s3738_r13046`.

JZ0 dijet background sample. URL `https://rucio-ui.cern.ch/did?scope=mc15_14TeV&name=mc15_14TeV.800281.Py8EG_AZ_Ztautau.recon.AOD.e8379_s3735_s3738_r13046`.

CustomCaloCell header file. URL `https://gitlab.cern.ch/atlas/athena/-/blob/21.9/Trigger/TrigL0GepPerf/TrigL0GepPerf/CustomCaloCell.h`.

The LHC filling scheme optimization for HL LHC . URL `https://indico.cern.ch/event/788818/contributions/3277254/attachments/1786584/2909004/002_filling_schemes.pdf`.

Archival Repository of Kirk's APU method of implementing an algorithm , a. URL `https://gitlab.cern.ch/atlas-tdaq-p2-firmware/global-trigger/gep-fw/-/tree/master/Sources/AlgoComputeCells/GepAlgTau1`.

Main repository for the Global trigger firmware, b. URL `https://gitlab.cern.ch/atlas-tdaq-p2-firmware/global-trigger/gep-fw`.

Lattice Strong Dynamics (LSD) Collaboration, Thomas Appelquist, et al. Stealth Dark Matter: Dark scalar baryons through the Higgs portal. *Phys. Rev. D*, 92(7): 075030, 2015. doi: 10.1103/PhysRevD.92.075030.

Graham D. Kribs, Adam Martin, and Tom Tong. Effective Theories of Dark Mesons with Custodial Symmetry. *JHEP*, 08:020, 2019c. doi: 10.1007/JHEP08(2019)020.

ATLAS Collaboration. Search for new high-mass phenomena in the dilepton final state using $36\,\mathrm{fb}^{-1}$ of proton–proton collision data at $\sqrt{s} = 13\,\mathrm{TeV}$ with the ATLAS detector. *JHEP*, 10:182, 2017a. doi: 10.1007/JHEP10(2017)182.

CMS Collaboration. Search for high-mass resonances in dilepton final states in proton–proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$. *JHEP*, 06:120, 2018. doi: 10.1007/JHEP06(2018)120.

ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi: 10.1088/1748-0221/3/08/S08003.

ATLAS Collaboration. ATLAS Insertable B-Layer: Technical Design Report, 2010a. URL `https://cds.cern.ch/record/1291633`.

B. Abbott et al. Production and integration of the ATLAS Insertable B-Layer. *JINST*, 13:T05008, 2018. doi: 10.1088/1748-0221/13/05/T05008.

G. Avoni et al. The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS. *JINST*, 13:P07017, 2018. doi: 10.1088/1748-0221/13/07/P07017.

ATLAS Collaboration. Performance of the ATLAS trigger system in 2015. *Eur. Phys. J. C*, 77:317, 2017b. doi: 10.1140/epjc/s10052-017-4852-3.

ATLAS Collaboration. Software and computing for Run 3 of the ATLAS experiment at the LHC. 2024a.

ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 83:982, 2023a. doi: 10.1140/epjc/s10052-023-11747-w.

S. Agostinelli et al. GEANT4 – a simulation toolkit. *Nucl. Instrum. Meth. A*, 506: 250, 2003. doi: 10.1016/S0168-9002(03)01368-8.

ATLAS Collaboration. The ATLAS Simulation Infrastructure. *Eur. Phys. J. C*, 70: 823, 2010b. doi: 10.1140/epjc/s10052-010-1429-9.

J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014a. doi: 10.1007/JHEP07(2014)079.

Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z.

Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159, 2015. doi: 10.1016/j.cpc.2015.01.024.

ATLAS Collaboration. ATLAS Pythia 8 tunes to 7 TeV data. ATL-PHYS-PUB-2014-021, 2014. URL `https://cds.cern.ch/record/1966419`.

NNPDF Collaboration, Richard D. Ball, et al. Parton distributions with LHC data. *Nucl. Phys. B*, 867:244, 2013. doi: 10.1016/j.nuclphysb.2012.10.003.

D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth. A*, 462:152, 2001. doi: 10.1016/S0168-9002(01)00089-4.

Christian Bierlich, Smita Chakraborty, Nishita Desai, Leif Gellersen, Ilkka Helenius, Philip Ilten, Leif Lönnblad, Stephen Mrenna, Stefan Prestel, Christian T. Preuss, Torbjörn Sjöstrand, Peter Skands, Marius Utheim, and Rob Verheyen. A comprehensive guide to the physics and usage of pythia 8.3, 2022. URL `https://arxiv.org/abs/2203.11601`.

Stefano Frixione, Giovanni Ridolfi, and Paolo Nason. A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *JHEP*, 09:126, 2007a. doi: 10.1088/1126-6708/2007/09/126.

Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004. doi: 10.1088/1126-6708/2004/11/040.

Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007b. doi: 10.1088/1126-6708/2007/11/070.

Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010. doi: 10.1007/JHEP06(2010)043.

NNPDF Collaboration, Richard D. Ball, et al. Parton distributions for the LHC run II. *JHEP*, 04:040, 2015. doi: 10.1007/JHEP04(2015)040.

Stefano Carrazza, Stefano Forte, Zahari Kassabov, Jose Ignacio Latorre, and Juan Rojo. An unbiased Hessian representation for Monte Carlo PDFs. *Eur. Phys. J. C*, 75(8):369, 2015. doi: 10.1140/epjc/s10052-015-3590-7.

ATLAS Collaboration. Studies on top-quark Monte Carlo modelling for Top2016. ATL-PHYS-PUB-2016-020, 2016a. URL https://cds.cern.ch/record/2216168.

M. Beneke, P. Falgari, S. Klein, and C. Schwinn. Hadronic top-quark pair production with NNLL threshold resummation. *Nucl. Phys. B*, 855:695–741, 2012. doi: 10.1016/j.nuclphysb.2011.10.021.

Matteo Cacciari, Michal Czakon, Michelangelo Mangano, Alexander Mitov, and Paolo Nason. Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation. *Phys. Lett. B*, 710:612–622, 2012a. doi: 10.1016/j.physletb.2012.03.013.

Peter Bärnreuther, Michal Czakon, and Alexander Mitov. Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$. *Phys. Rev. Lett.*, 109:132001, 2012. doi: 10.1103/PhysRevLett.109.132001.

Michal Czakon and Alexander Mitov. NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels. *JHEP*, 12:054, 2012. doi: 10.1007/JHEP12(2012)054.

Michal Czakon and Alexander Mitov. NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction. *JHEP*, 01:080, 2013. doi: 10.1007/JHEP01(2013)080.

Michal Czakon, Paul Fiedler, and Alexander Mitov. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$. *Phys. Rev. Lett.*, 110:252004, 2013. doi: 10.1103/PhysRevLett.110.252004.

Michal Czakon and Alexander Mitov. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Comput. Phys. Commun.*, 185:2930, 2014. doi: 10.1016/j.cpc.2014.06.021.

Tomáš Ježo, Jonas M. Lindert, Niccolo Moretti, and Stefano Pozzorini. New NLOPS predictions for $t\bar{t} + b$-jet production at the LHC. *Eur. Phys. J. C*, 78(6):502, 2018. doi: 10.1140/epjc/s10052-018-5956-0.

Federico Buccioni, Jean-Nicolas Lang, Jonas M. Lindert, Philipp Maierhöfer, Stefano Pozzorini, Hantian Zhang, and Max F. Zoller. OpenLoops 2. *Eur. Phys. J. C*, 79 (10):866, 2019. doi: 10.1140/epjc/s10052-019-7306-2.

Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini. Scattering Amplitudes with Open Loops. *Phys. Rev. Lett.*, 108:111601, 2012. doi: 10.1103/PhysRevLett. 108.111601.

Ansgar Denner, Stefan Dittmaier, and Lars Hofer. COLLIER: A fortran-based complex one-loop library in extended regularizations. *Comput. Phys. Commun.*, 212:220–238, 2017. doi: 10.1016/j.cpc.2016.10.013.

Tomas Ježo. Powheg-box-res ttbb source code, 2019. URL `https://gitlab.cern.ch/tjezo/powheg-box-res_ttbb/`.

ATLAS Collaboration. Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ using the ATLAS detector. *Eur. Phys. J. C*, 76:538, 2016b. doi: 10.1140/epjc/s10052-016-4366-4.

Michal Czakon, David Heymes, Alexander Mitov, Davide Pagani, Ioannis Tsinikos, and Marco Zaro. Top-pair production at the LHC through NNLO QCD and NLO EW. *JHEP*, 10:186, 2017. doi: 10.1007/JHEP10(2017)186.

Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008b. doi: 10.1088/1126-6708/2008/04/063.

Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur. Phys. J. C*, 72:1896, 2012b. doi: 10.1140/epjc/s10052-012-1896-2.

Enrico Bothmann et al. Event generation with Sherpa 2.2. *SciPost Phys.*, 7(3):034, 2019. doi: 10.21468/SciPostPhys.7.3.034.

Tanju Gleisberg and Stefan Höche. Comix, a new matrix element generator. *JHEP*, 12:039, 2008. doi: 10.1088/1126-6708/2008/12/039.

Steffen Schumann and Frank Krauss. A parton shower algorithm based on Catani–Seymour dipole factorisation. *JHEP*, 03:038, 2008. doi: 10.1088/1126-6708/2008/03/038.

Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. A critical appraisal of NLO+PS matching methods. *JHEP*, 09:049, 2012. doi: 10.1007/JHEP09(2012)049.

Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. QCD matrix elements + parton showers. The NLO case. *JHEP*, 04:027, 2013. doi: 10.1007/JHEP04(2013)027.

S. Catani, F. Krauss, B. R. Webber, and R. Kuhn. QCD Matrix Elements + Parton Showers. *JHEP*, 11:063, 2001. doi: 10.1088/1126-6708/2001/11/063.

Stefan Höche, Frank Krauss, Steffen Schumann, and Frank Siegert. QCD matrix elements and truncated showers. *JHEP*, 05:053, 2009. doi: 10.1088/1126-6708/2009/05/053.

Charalampos Anastasiou, Lance Dixon, Kirill Melnikov, and Frank Petriello. High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order. *Phys. Rev. D*, 69:094008, 2004. doi: 10.1103/PhysRevD.69.094008.

Emanuele Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J. C*, 71:1547, 2011. doi: 10.1140/epjc/s10052-011-1547-z.

Stefano Frixione, Eric Laenen, Patrick Motylinski, Chris White, and Bryan R. Webber. Single-top hadroproduction in association with a W boson. *JHEP*, 07:029, 2008. doi: 10.1088/1126-6708/2008/07/029.

Rikkert Frederix, Emanuele Re, and Paolo Torrielli. Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO. *JHEP*, 09:130, 2012. doi: 10.1007/JHEP09(2012)130.

Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO single-top production matched with shower in POWHEG: s- and t-channel contributions. *JHEP*, 09:111, 2009. doi: 10.1088/1126-6708/2009/09/111.

Rikkert Frederix, Davide Pagani, and Marco Zaro. Large NLO corrections in $t\bar{t}W^{\pm}$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions. *JHEP*, 02:031, 2018. doi: 10.1007/JHEP02(2018)031.

Stefano Frixione, Eric Laenen, Patrick Motylinski, and Bryan R. Webber. Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations. *JHEP*, 04:081, 2007c. doi: 10.1088/1126-6708/2007/04/081.

Pierre Artoisenet, Rikkert Frederix, Olivier Mattelaer, and Robbert Rietkerk. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *JHEP*, 03:015, 2013. doi: 10.1007/JHEP03(2013)015.

Heribertus B. Hartanto, Barbara Jäger, Laura Reina, and Doreen Wackeroth. Higgs boson production in association with top quarks in the POWHEG BOX. *Phys. Rev. D*, 91(9):094003, 2015. doi: 10.1103/PhysRevD.91.094003.

D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2017. doi: 10.23731/CYRM-2017-002.

ATLAS Collaboration. Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV. ATL-PHYS-PUB-2015-026, 2015a. URL `https://cds.cern.ch/record/2037717`.

ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data. *JINST*, 14:P12006, 2019a. doi: 10.1088/1748-0221/14/12/P12006.

ATLAS Collaboration. Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 79:639, 2019b. doi: 10.1140/epjc/s10052-019-7140-6.

ATLAS Collaboration. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13\,\text{TeV}$. *Eur. Phys. J. C*, 81:578, 2021a. doi: 10.1140/epjc/s10052-021-09233-2.

ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J. C*, 77:466, 2017c. doi: 10.1140/epjc/s10052-017-5031-2.

ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C*, 77:490, 2017d. doi: 10.1140/epjc/s10052-017-5004-5.

ATLAS Collaboration. Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ using the ATLAS detector. *Eur. Phys. J. C*, 76:581, 2016c. doi: 10.1140/epjc/s10052-016-4395-z.

ATLAS Collaboration. Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector. *Eur. Phys. J. C*, 81:689, 2021b. doi: 10.1140/epjc/s10052-021-09402-3.

ATLAS Collaboration. Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton–proton run. *JINST*, 8: P07004, 2013. doi: 10.1088/1748-0221/8/07/P07004.

ATLAS Collaboration. Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector. ATLAS-CONF-2015-029, 2015b. URL `https://cds.cern.ch/record/2037702`.

ATLAS Collaboration. ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset. *Eur. Phys. J. C*, 83:681, 2023b. doi: 10.1140/epjc/s10052-023-11699-1.

ATLAS Collaboration. Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run. ATL-PHYS-PUB-2017-013, 2017e. URL `https://cds.cern.ch/record/2273281`.

ATLAS Collaboration. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13\,\text{TeV}$. *Eur. Phys. J. C*, 79:970, 2019c. doi: 10.1140/epjc/s10052-019-7450-8.

ATLAS Collaboration. Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using $139\,\text{fb}^{-1}$ of ATLAS proton–proton collision data at $\sqrt{s} = 13\,\text{TeV}$. *Eur. Phys. J. C*, 83:728, 2023c. doi: 10.1140/epjc/s10052-023-11736-z.

ATLAS Collaboration. Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13\,\text{TeV}$ collected with the ATLAS detector. *Eur. Phys. J. C*, 82:95, 2022a. doi: 10.1140/epjc/s10052-021-09843-w.

Benjamin Nachman, Pascal Nef, Ariel Schwartzman, Maximilian Swiatlowski, and Chaowaroj Wanotayaroj. Jets from jets: re-clustering as a tool for large radius jet reconstruction and grooming at the LHC. *JHEP*, 02:075, feb 2015. doi: 10.1007/jhep02(2015)075.

ATLAS Collaboration. The performance of the jet trigger for the ATLAS detector during 2011 data taking. *Eur. Phys. J. C*, 76:526, 2016d. doi: 10.1140/epjc/s10052-016-4325-0.

ATLAS Collaboration. Performance of the ATLAS muon triggers in Run 2. *JINST*, 15:P09015, 2020a. doi: 10.1088/1748-0221/15/09/p09015.

ATLAS Collaboration. Performance of electron and photon triggers in ATLAS during LHC Run 2. *Eur. Phys. J. C*, 80:47, 2020b. doi: 10.1140/epjc/s10052-019-7500-2.

ATLAS Collaboration. Operation of the ATLAS trigger system in Run 2. *JINST*, 15:P10004, 2020c. doi: 10.1088/1748-0221/15/10/P10004.

ATLAS Collaboration. The ATLAS inner detector trigger performance in pp collisions at 13 TeV during LHC Run 2. *Eur. Phys. J. C*, 82:206, 2022b. doi: 10.1140/epjc/s10052-021-09920-0.

ATLAS Collaboration. Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector. *Phys. Rev. D*, 98:012003, 2018. doi: 10.1103/PhysRevD.98.012003.

ATLAS Collaboration. Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 04:046, 2019d. doi: 10.1007/JHEP04(2019)046.

ATLAS Collaboration. Simulation-based extrapolation of b-tagging calibrations towards high transverse momenta in the ATLAS experiment. ATL-PHYS-PUB-2021-003, 2021c. URL `https://cds.cern.ch/record/2753444`.

ATLAS Collaboration. Search for top-philic heavy resonances in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 84:157, 2024b. doi: 10.1140/epjc/s10052-023-12318-9.

Michiel Botje, Jon Butterworth, Amanda Cooper-Sarkar, Albert de Roeck, Joel Feltesse, Stefano Forte, Alexander Glazov, Joey Huston, Ronan McNulty, Torbjorn Sjostrand, and Robert Thorne. The pdf4lhc working group interim recommendations, 2011.

Stefan Höche, Stephen Mrenna, Shay Payne, Christian Tobias Preuss, and Peter Skands. A Study of QCD Radiation in VBF Higgs Production with Vincia and Pythia. *SciPost Phys.*, 12(1):010, 2022. doi: 10.21468/SciPostPhys.12.1.010.

M. Bähr et al. Herwig++ physics and manual. *Eur. Phys. J. C*, 58:639, 2008. doi: 10.1140/epjc/s10052-008-0798-9.

Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J. C*, 76 (4):196, 2016. doi: 10.1140/epjc/s10052-016-4018-8.

L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne. Parton distributions in the LHC era: MMHT 2014 PDFs. *Eur. Phys. J. C*, 75(5):204, 2015. doi: 10.1140/epjc/s10052-015-3397-6.

Wouter Verkerke and David Kirkby. The RooFit toolkit for data modeling, 2003a.

M. Hatlo, F. James, P. Mato, L. Moneta, M. Winkler, and A. Zsenei. Developments of mathematical software libraries for the lhc experiments. *IEEE Trans. Nucl. Sci.*, 52(6):2818–2822, 2005a. doi: 10.1109/TNS.2005.860152.

Alexander L. Read. Presentation of search results: the CL_S technique. *J. Phys. G*, 28:2693, 2002. doi: 10.1088/0954-3899/28/10/313.

Search for charged higgs bosons decaying via $H^\pm \to \tau^\pm \nu_\tau$ in the τ+jets and τ+lepton final states with 36 fb^{-1}of pp collision data recorded at $\sqrt{s} = 13$ TeVwith the atlas

experiment. *Journal of High Energy Physics*, 2018(9), September 2018b. ISSN 1029-8479. doi: 10.1007/jhep09(2018)139. URL `http://dx.doi.org/10.1007/JHEP09(2018)139`.

Search for charged higgs bosons decaying into top and bottom quarks at $\sqrt{s} = 13$ TeVwith the atlas detector. *Journal of High Energy Physics*, 2018(11), November 2018c. ISSN 1029-8479. doi: 10.1007/jhep11(2018)085. URL `http://dx.doi.org/10.1007/JHEP11(2018)085`.

CMS Collboration. Search for a charged higgs boson decaying into top and bottom quarks in events with electrons or muons in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2020(1), January 2020b. ISSN 1029-8479. doi: 10.1007/jhep01(2020)096. URL `http://dx.doi.org/10.1007/JHEP01(2020)096`.

Celine Degrande, Maria Ubiali, Marius Wiesemann, and Marco Zaro. Heavy charged higgs boson production at the lhc, 2015a. URL `https://arxiv.org/abs/1507.02549`.

J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), July 2014b. ISSN 1029-8479. doi: 10.1007/jhep07(2014)079. URL `http://dx.doi.org/10.1007/JHEP07(2014)079`.

Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, José I. Latorre, Juan Rojo, and Maria Ubiali. Parton distributions with lhc data. *Nuclear Physics*

B, 867(2):244–289, February 2013. ISSN 0550-3213. doi: 10.1016/j.nuclphysb.2012. 10.003. URL `http://dx.doi.org/10.1016/j.nuclphysb.2012.10.003`.

Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to pythia 8.1. *Computer Physics Communications*, 178(11):852–867, June 2008. ISSN 0010-4655. doi: 10.1016/j.cpc.2008.01.036. URL `http://dx.doi.org/10.1016/j.cpc.2008.01.036`.

Celine Degrande, Maria Ubiali, Marius Wiesemann, and Marco Zaro. Heavy charged higgs boson production at the lhc, 2015b. URL `https://arxiv.org/abs/1507.02549`.

CERN. Cern yellow reports: Monographs, vol 2 (2017): Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector, 2017. URL `https://e-publishing.cern.ch/index.php/CYRM/issue/view/32`.

Martin Flechl, Richard Klees, Michael Krämer, Michael Spira, and Maria Ubiali. Improved cross-section predictions for heavy charged higgs boson production at the lhc. *Physical Review D*, 91(7), April 2015. ISSN 1550-2368. doi: 10.1103/physrevd.91.075015. URL `http://dx.doi.org/10.1103/PhysRevD.91.075015`.

Stefan Dittmaier, Michael Krämer, Michael Spira, and Manuel Walser. Charged-higgs-boson production at the lhc: Next-to-leading-order supersymmetric qcd corrections. *Physical Review D*, 83(5), March 2011. ISSN 1550-2368. doi: 10.1103/physrevd.83.055005. URL `http://dx.doi.org/10.1103/PhysRevD.83.055005`.

Edmond L Berger, Tao Han, Jing Jiang, and Tilman Plehn. Associated production of a top quark and a charged higgs boson. *Physical Review D*, 71(11), June 2005.

ISSN 1550-2368. doi: 10.1103/physrevd.71.115012. URL `http://dx.doi.org/10.1103/PhysRevD.71.115012`.

Top reconstruction group. Analysistop documentation. URL `https://atlas-topq.docs.cern.ch/Reco/`.

Documentation for trexfitter. URL `https://trexfitter-docs.web.cern.ch/trexfitter-docs/latest/`.

Wouter Verkerke and David Kirkby. The roofit toolkit for data modeling, 2003b. URL `https://arxiv.org/abs/physics/0306116`.

M. Hatlo, F. James, P. Mato, L. Moneta, M. Winkler, and A. Zsenei. Developments of mathematical software libraries for the lhc experiments. *IEEE Transactions on Nuclear Science*, 52(6):2818–2822, 2005b. doi: 10.1109/TNS.2005.860152.

Maria Mironova. Introduction to statistical analysis.