2018

A Search For Nothing: Dark Matter And Invisible Decays Of The Higgs Boson At The Atlas Detector

William Balunas
University of Pennsylvania, balunas@sas.upenn.edu

Recommended Citation

Balunas, William, "A Search For Nothing: Dark Matter And Invisible Decays Of The Higgs Boson At The Atlas Detector" (2018). *Publicly Accessible Penn Dissertations*. 2723.
https://repository.upenn.edu/edissertations/2723

A Search For Nothing: Dark Matter And Invisible Decays Of The Higgs Boson At The Atlas Detector

Abstract

This thesis presents a search for invisible decays of the Higgs boson using the vector boson fusion channel. This uses 36 fb^−1 of proton-proton collision data at √s = 13 TeV using the ATLAS detector at the Large Hadron Collider. The experimental methods for understanding the signal and background processes as well as detector effects are described in detail. The search is carried out in several regions defined by kinematic requirements on the final-state objects, and the observed event yields are used in a profile-likelihood fit in order to constrain the backgrounds. The results are interpreted using a modified frequentist method and are found to be consistent with the Standard Model expectations. An upper limit of 34% (28% expected) at 95% C.L. is placed on the invisible branching ratio of the Higgs boson. Re-interpretation of these results in terms of dark matter is also discussed, in the context of the Higgs portal and other simplified models.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Physics & Astronomy

First Advisor

Elliot Lipeles

Keywords

ATLAS, Dark Matter, Higgs, LHC, VBF

Subject Categories

Elementary Particles and Fields and String Theory | Physics

A Search for Nothing: Dark Matter and Invisible Decays of the Higgs Boson at the ATLAS Detector

William Balunas

a dissertation
in
Physics and Astronomy

Presented to the Faculties of The University of Pennsylvania
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy
2018

Elliot Lipeles, Associate Professor, Physics
Supervisor of Dissertation

Joshua Klein, Professor, Physics
Graduate Group Chairperson

Dissertation Committee
I. Joseph Kroll, Professor, Physics
Justin Khoury, Associate Professor, Physics
Joshua Klein, Professor, Physics
Mirjam Cvetič, Professor, Physics
Elliot Lipeles, Associate Professor, Physics

A Search for Nothing: Dark Matter and Invisible Decays of the Higgs Boson at the ATLAS Detector

Acknowledgements

There's a long list of people who deserve thanks for the role they've played in making this thesis a reality, and for supporting me in my journey through graduate school more generally. I'll do my best to credit as many as I can here.

First off, thanks to my advisor Elliot Lipeles for his unyielding support and consistently good advice over the past several years. I've learned a huge amount from working with him and I couldn't ask for a better advisor. Special thanks as well for his help in reviewing and editing this thesis.

I'd also like to thank the rest of the faculty in the Penn ATLAS group: Brig Williams, Joe Kroll, and Evelyn Thomson. They've done an excellent job maintaining a great environment to work in, and they always ask the right questions.

I owe a lot of credit to my collaborators on the VBF Invisible analysis; the scope and complexity of this project would make it impossible for any one person alone (at least on the timescale of a Ph.D. thesis). So thanks to Ben Carlson, Jamie Saxon, Alex Madsen, Marta Perego, and Rui Zou, who all devoted large fractions of their time to this project at various points over the last few years. Jamie and Alex are no longer part of ATLAS, but their important contributions are still much appreciated. Thanks to Loan Truong for her work on the technical side of implementing and using the statistical model for the analysis. Thanks to Andy White for organizing and managing the project, as well as contributing some work on the signal modeling. Thanks to Christian Sander and Vincent Kitali for their excellent work on understanding the notorious multijet background. Thanks to Tae Min Hong and Ketevi Assamagan for preparing the publication draft and dealing with the incredible quantity of comments from the collaboration during internal review.

The physics group and subgroup conveners also deserve a mention for their assistance and good suggestions throughout the lifetime of the project. In particular, Klaus Mönig, Gabriel Facini, Arely Cortes-Gonzalez, Marco Vanadia, and Silvia Resconi have been very helpful. Thanks also to our editorial board, especially Hugo Beauchemin, for tirelessly working to ensure that every last detail is correct and well-presented.

There are several others on ATLAS who have given me valuable help on other projects as well. Thanks to Alan Watson, Denis Damazio, Walter Hopkins, and Will Buttinger for their assistance with learning to simulate the ATLAS L1 trigger system. Thanks also to David Strom and Michael Begel for their suggestions, feedback, and ideas for my upgrade performance work. Thanks to Jeff Dandoy and Kate Pachal for their help with all things related to jet calibration.

I also need to thank my colleagues at Penn I haven't already mentioned; I wouldn't have learned nearly as much in graduate school (or enjoyed it nearly as much) without them. That includes Bijan Haney, Khilesh Mistry, Christian Herwig, Elodie Resseguie, Leigh Schaefer, Joey Reichert, Will DiClemente, Ben Rosser, Joana Machado Miguéns, Rob Fletcher, Chris Meyer, and Keisuke Yoshihara. The same goes for former Penn students and postdocs: John Alison, Sarah Heim, Kurt Brendlinger, Alex Tuna, and Josh Kunkle.

On a more personal level, I owe a huge amount of gratitude to my friends for giving me the support and motivation to make it this far. Jason Rocks, Katie Siewert, and Johanna-Laina Fischer, that means you. Elodie Resseguie deserves another shout-out here as well. Thanks also to my friends at CERN (past and present) for making moving to a foreign country such a great experience: Doug Davis, Asmund Folkestad, Nefeli & Timoklia Kousi, Gerhard Rzehorz and many others. Thanks to Maddie McKay for always being able to talk about anything. Special thanks to Ana Trišović for helping me preserve my sanity. And thanks to Milena Bajić, for the things she taught me about myself and so much else.

Perhaps most importantly, I'd like to thank my family, especially my parents Edward and Stephanie, for their unconditional support and love throughout my entire life. None of this would have been possible without their many years of hard work, through good times and bad. This Ph.D. seems easy in comparison to everything they've done for me.

ABSTRACT

A Search for Nothing: Dark Matter and Invisible Decays of the Higgs Boson at the ATLAS Detector

William Balunas

E. Lipeles

This thesis presents a search for invisible decays of the Higgs boson using the vector boson fusion channel. This uses 36 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV using the ATLAS detector at the Large Hadron Collider. The experimental methods for understanding the signal and background processes as well as detector effects are described in detail. The search is carried out in several regions defined by kinematic requirements on the final-state objects, and the observed event yields are used in a profile-likelihood fit in order to constrain the backgrounds. The results are interpreted using a modified frequentist method and are found to be consistent with the Standard Model expectations. An upper limit of 34% (28% expected) at 95% C.L. is placed on the invisible branching ratio of the Higgs boson. Re-interpretation of these results in terms of dark matter is also discussed, in the context of the Higgs portal and other simplified models.

Contents

List of Tables

List of Figures

Preface

This thesis presents the majority of my work on ATLAS as a graduate student at the University of Pennsylvania between 2013 and 2018.

When I first came to Penn in June 2013, the LHC had recently completed its Run 1, and the ATLAS collaboration was starting to focus on looking ahead to the future. I joined the trigger upgade effort, working on simulation software for the new systems that would be put into place for the Phase-I and Phase-II upgrades (currently planned to start in 2019 and 2024 respectively). I wrote the software to simulate the algorithms that would later be implemented in hardware for the L1 trigger, and worked on developing improvements to these algorithms themselves. I focused particularly on pileup mitigation and forward jets, since this will become particularly challenging problems in the future as the LHC instantaneous luminosity continues to increase. I worked on these topics for well over two years, but they are given less emphasis in this thesis (only Chapter 5 is devoted to them), as they don't pertain directly to the nominal subject.

Around early 2015 (around the time the Run 1 paper was undergoing ATLAS internal approval), I began working on the VBF Higgs to Invisible analysis effort for Run 2. I started off with planning out trigger strategies but soon branched out into more of the problems that come with starting up a new analysis (getting the necessary Monte Carlo simulation, writing a data analysis software framework, etc.). Over the next months, a few more people joined the effort and we had formed a small, but functioning, analysis group right around the time ATLAS began taking Run 2 data.

In September 2015, I relocated to CERN and stayed there until December 2017. This was a particularly valuable learning experience; being able to interact so easily with so many physicists on ATLAS (and other projects) taught me an enormous amount about the experiment and the field or particle physics in general. While I was there I contributed to the day-to-day operation of the experiment by taking Trigger shifts in the control room, in addition to continuing my work on trigger upgrade simulation and the VBF Invisible analysis. Along the way, also I worked on some small “side projects” (such as uncertainties for jet energy calibration, among a few others). These are not discussed in this document.

Now, as I finish writing this thesis, the analysis is undergoing ATLAS approval and the paper is being prepared for submission to *Physics Letters B*. This is the culmination of about 3 years of work by a very small group of people (I’d estimate about 3 really active members at any given time on average). I personally have had a hand in nearly every aspect of the project from start to finish; the only notable exception is the actual evaluation of the multijet background in the signal region, which Christian Sander at DESY deserves the credit for. The material presented in this thesis was produced by the analysis team as a group; as a result I of course didn’t personally compute every single number and make every single plot. The text is entirely my own, though a few sections are based heavily on parts of the ATLAS internal note[1] for the analysis (only the parts I myself originally wrote). Since the result has not yet been made public by ATLAS, many of the plots are still labeled “ATLAS Internal”. In general these aren’t meant to be shown outside the collaboration until the approval process is finished, but exceptions are made for theses such as this one. The same holds true for the numbers (such as uncertainties or the final limits); it is possible that there might be some small changes during the final internal approval process, so the publication could have some minor differences relative to what’s presented here.

My intention was to make this thesis focused and concise. As such, I’ve spent most of the body discussing the primary topic (i.e. the search for invisibly decaying Higgs). If, as the reader, you’ve also looked through other ATLAS students’ theses, you’ll likely find that I spend less space on my technical work and more on the analysis. That said, everything here is meant to be self-contained and complete. A new graduate student, for example, should be able to understand everything by reading it in order. Readers who are already familiar with the technical aspects of the ATLAS experiment and particle reconstruction might want to skip Chapters 3 and 4. If you’re already well-versed in the theoretical aspects, you could skip Chapter 2. If you don’t care about trigger algorithms for ATLAS upgrades, then you’ll probably want to skip Chapter 5.

WKB
Philadelphia, April 2018

---

[1] For ATLAS readers, the reference number is `ATL-COM-PHYS-2016-1802`.

Chapter 1

Introduction

The Standard Model (SM) has been remarkably successful at describing nearly all observed physical phenomena, other than gravitation. However, it is still widely viewed as incomplete, even aside from the lack of a quantum description of gravity. This is due to theoretical considerations such as naturalness and unification, as well as more practical problems including the lack of a viable dark matter candidate. For this reason, increasingly high-energy particle colliders have been built over the past several decades in an attempt to explore the physics lying beyond the energy frontier.

The present time represents a turning point in the field of particle physics; until the discovery of the Higgs boson in 2012, there has been fairly definitive theoretical motivation behind the searches being done. For example, the W, Z, and Higgs bosons were all expected in exactly the form they were found. Now, it is no longer clear at all what kinds of particles remain to be found. For this reason, experiments at the energy frontier are now carrying out search programs of unprecedented breadth, ranging from specific supersymmetry models to general model-independent searches for a variety of final states.

One of the few concrete "leads" that particle physicists currently have is that of dark matter. This is something which we know exists from astrophysical measurements, and could easily be explained by some extension to the Standard Model. Measuring the production of dark matter at a collider would provide new insights about the overarching composition of the univserse, as well as the connection between physics at the smallest and largest observable scales.

The primary subject of this thesis is a search for dark matter production via weak vector boson fusion (VBF)[2] using the ATLAS detector. This is primarily interpreted in terms of invisible decays of the Higgs boson. As such, it is referred to as the "VBF Higgs → Invisible" or "VBF+$E_{\mathrm{T}}^{\mathrm{miss}}$"

---

[2] See Section 2.3 for futher information on VBF and other modes of Higgs boson production.

analysis[3]. Invisible decays of the Higgs boson could have important implications with respect to dark matter, as the Higgs can act as a mediator between the known Standard Model and the "dark sector" (see, for example, [1, 2, 3, 4] for general background information).

The search is interpreted in a way which is largely independent of the underlying theory of the dark matter; an upper limit is placed on the branching ratio of the Higgs to invisible particles ($\mathcal{B}(h \to \text{inv.})$), which is a general constraint that can be applied to any theory. In practice, this is only useful for theories which have a dark matter candidate lighter than $m_h/2$. The exact same analysis can also be reinterpreted in terms of other dark matter models, involving beyond-Standard-Model mediators rather than the Higgs. In addition to the model-independent limit on $\mathcal{B}(h \to \text{inv.})$, the result is re-cast in terms of a few specific benchmark models, including those with non-SM mediators.

This thesis begins with several chapters dedicated to explaining the background and general concepts that underlie this search. These are intended to be quite general, but focus on the aspects that are particularly relevant later on. Chapter 2 gives a brief overview of our current theoretical understanding of the relevant physics. Chapter 3 describes the experimental setup used to gather the data, including the LHC and the ATLAS detector itself. Chapter 4 details the methods used to reconstruct the various objects produced in LHC collisions. Chapter 5 presents a set of studies and simulation work for ATLAS hadronic trigger upgrades (this is a self-contained chapter and is not needed to understand the rest).

The rest of this thesis is devoted to the VBF+$E_{\mathrm{T}}^{\mathrm{miss}}$ analysis at 13 TeV. Chapter 6 gives an overview of the analysis, including the overall strategy and methods. Chapter 7 focuses in detail on the techniques used for modeling the signal and precisely understanding the backgrounds. Chapter 8 presents the results of the analysis, along with a detailed description of the statistical methods used and further interpretation.

This is the first result of its kind from ATLAS Run 2, and expands on older results from Run 1 [5]. It uses 36 fb^{-1} of 13 TeV data from 2015 and 2016. The CMS experiment has included a similar Run 2 analysis on a much smaller dataset as part of a combination [6], but much of the methodology is different. The results presented here will also be used as part of a larger ATLAS invisible Higgs combination, much like what was done for Run 1 [7]. Ultimately, the goal is for this work to lay the foundation for a longer-term precision analysis on the full Run-2 dataset (projected to be roughly 120–150 fb^{-1}).

---

[3] $E_{\mathrm{T}}^{\mathrm{miss}}$ refers to "missing transverse momentum", which is a variable used as a proxy for the momentum of undetectable objects. See Section 4.4 for a detailed description of its definition.

Chapter 2

Theoretical Framework

2.1 Introduction to the Standard Model

Modern particle physics is generally interpreted in terms of the Standard Model (SM). This is a quantum field theory which encapsulates our understanding of the electromagnetic, weak, and strong interactions in a single framework. It was developed gradually over several decades, beginning with the first use of quantum field theory to provide a complete description of electrodynamics in the 1920s [8, 9].

The SM is based on the assumption that the matter in the universe (quarks and leptons) consists of fermionic fields. These fields are quantized[4], which implies that their excitations can be thought of as particles. Furthermore, these fields are assumed to be symmetric under local transformations of the unitary groups U(1), SU(2), and SU(3). These symmetries imply the existence of vector gauge fields which mediate the interactions between the other fields (and in some cases, themselves).

The electromagnetic interaction was the first part of the SM to be understood in terms of a quantum field theory, known as "quantum electrodynamics" (QED) [13]. This consists of a U(1) gauge field (corresponding to photons) interacting with the fermions. Since the coupling of this interaction is small ($\alpha \approx 1/137$), it can be described accurately in terms of a perturbative expansion in α. This means cross section calculations at leading or next-to-leading order are sufficient for most practical purposes at all accessible energy scales.

Several decades later, the strong interaction was also successfully described as a quantum field theory: "quantum chromodynamics" (QCD). This is conceptually similar to QED, but the phenomenology is very different, for several reasons. First, the symmetry group for QCD (first proposed

---

[4] The underlying principles of quantum field theory will not be discussed here; a basic understanding is assumed. A detailed treatment can be found in [10, 11, 12] among many others.

in [14]) is SU(3). This has 8 generators, implying 8 mediating gauge bosons (gluons). Additionally, SU(3) is non-Abelian, which implies that the gluons self-interact. In the case of gluons, this leads to an effect called confinement, where no object which is charged under the strong interaction ("colored") can exist independently. Instead, they always form bound states ("hadrons"). The existence of the fundamental quarks had to be inferred from the properties of the hadrons themselves [15, 16]. If a colored object is isolated, it quickly becomes energetically favorable to produce new quark-antiquark pairs, creating new hadrons. When a quark or gluon is produced in a high-energy collision, it typically results in a shower of hadrons (a "jet") because of this. The coupling of this interaction is also large at low energies ($\alpha_s \sim 1$), meaning it cannot be described perturbatively. Due to renormalization effects, the coupling constant "runs", giving it a smaller effective value at higher energy scales. So, for hard interactions at high energies (above tens of GeV), QCD cross sections can be calculated using an expansion in α_s [17, 18]. This property is known as "asymptotic freedom".

The final component of the SM gauge sector is the SU(2) symmetry. This has 3 generators, corresponding to 3 gauge fields. In practice, these SU(2) gauge fields mix with the U(1) component of the SM. This is discussed in more detail in Section 2.2, but the end result is three vector bosons corresponding to the weak interaction ($W^{\pm}$ and Z) and one corresponding to the electromagnetic interaction (γ, the photon). Because of this unification, the two interactions are together referred to as the "electroweak interaction". SU(2) is also non-Abelian, but since the W and Z are massive, this self-interaction doesn't result in confinement like QCD does. The weak gauge fields only interact with the left-handed chiral component of the fermion fields in the SM (the right-handed components are not charged under the weak interaction). The left-handed fermions are SU(2) doublets.

In addition to the fermions and vector (gauge) bosons in the SM, there is a scalar field (the Higgs field). First proposed in 1964 [19, 20, 21, 22], this was discovered only recently, in 2012 [23, 24]. Interactions with this field are the source of the masses of the fermions and weak gauge bosons. This is explained further in Section 2.2. Since the Higgs is relatively newly discovered and its properties are not yet measured precisely, it is often considered as a possible candidate for a mediator to new physics beyond the SM. Currently, the branching fraction of the Higgs to non-SM particles is constrained only to within 34% (at 95% CL), from a combination of ATLAS and CMS results [25]. For the case where the decay products are invisible[5], the current upper limit is 23% [6, 7].

---

[5] "Invisible" means the particles do not interact with the detector and thus cannot be measured directly. Neutrinos are the only invisible SM particles.

2.2 Electroweak Mixing and the Higgs Field

When the theory of the electroweak interaction was first developed [26, 27], the W and Z bosons were predicted to be massless (a typical mass term in the Lagrangian would violate the SU(2) symmetry). However, these were experimentally observed to have masses of 80 GeV and 91 GeV respectively [28, 29, 30, 31]. The Higgs field was added to the SM as a solution to this problem. It "spontaneously breaks" the electroweak symmetry, which results in the W and Z masses as well as the mixing between the U(1) and SU(2) components.

The Higgs field is a complex SU(2) doublet (denoted ϕ), so it has 4 degrees of freedom. The most general Lagrangian for a scalar of this type (excluding its interactions with the fermions) is

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D^\mu \phi) + \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \tag{2.1}$$

where μ and λ are free parameters and D_μ denotes the covariant derivative:

$$D_\mu \phi = (\partial_\mu - \frac{i}{2} g A^a_\mu \sigma^a - \frac{i}{2} g' B_\mu)\phi. \tag{2.2}$$

Here, A^a_μ and B_μ are respectively the SU(2) and U(1) gauge fields, g and g' are their coupling constants, and σ^a are the generators of SU(2). As Equation 2.1 indicates, the Higgs potential has a quartic form. λ must be positive (or the potential would not be bounded from below), but this is not the case for μ^2. In fact, if $\mu^2 < 0$ (which we observe in nature), then the minimum of the Higgs potential is not at $\phi = 0$ but rather at

$$|\langle \phi \rangle| = \sqrt{\frac{-\mu^2}{\lambda}} \tag{2.3}$$

This means the Higgs field has a nonzero vacuum expectation value. Due to the SU(2) symmetry, we are free to choose a "direction" for this value; the most common convention is

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{2.4}$$

Perturbative expansions can then be done around this point. v is experimentally measured to be roughly 246 GeV. Here, this can be substituted back into the kinetic term of Equation 2.1 to find explicitly:

$$\mathcal{L}_{Higgs} \subset \frac{v^2}{8}[g^2 A^1_\mu A^{1\mu} + g^2 A^2_\mu A^{2\mu} + (-gA^3_\mu + g' B_\mu)(-gA^{3\mu} + g' B^\mu)] \tag{2.5}$$

This has exactly the form of a mass term for three gauge bosons. Note that one of the mass eigenstates is a mixture of A^3 and B; this corresponds to the physical Z boson. The fourth vector field (which is the other mixture of A^3 and B orthogonal to the Z) has no mass term; this is the photon.

The Higgs also has direct couplings to the SM fermions. These are known as Yukawa couplings, and have the form:

$$\mathcal{L} \subset -\lambda_\psi (\bar{\psi}_L \phi)\psi_R + h.c. \tag{2.6}$$

where ψ_L is the left-handed fermion doublet and ψ_R is the right-handed singlet. This results in mass terms for the fermions which would otherwise be forbidden by the SU(2) symmetry, with:

$$m_\psi = \frac{1}{\sqrt{2}}\lambda_\psi v \tag{2.7}$$

This implies that large fermion masses correlate directly to large couplings to the Higgs field. Consequently, the top quark has by far the largest Yukawa coupling of the fermions.

2.3 Higgs Boson Production and Decay

At the LHC, we're interested in producing the Higgs boson in proton-proton collisions and studying its decays. So, we must first understand production mechanisms. At a proton-proton collider, the initial state consist of quarks and gluons (the parton density of the other SM particles is negligible here). The largest coupling for Higgs production is via the top quark, but the top content of the proton is effectively zero. However, the top Yukawa coupling can be accessed through a gluon-initiated loop or via $t\bar{t}$ pair production (Figure 2.1). By far the largest cross section is for the gluon-initiated top loop (also called gluon-gluon fusion, or ggF).

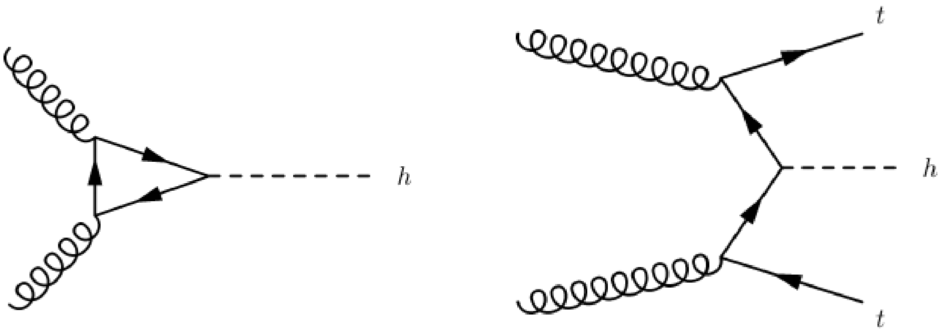

Figure 2.1: Feynman diagrams for Higgs production via gluon-gluon fusion (**left**) and $t\bar{t}$ associated production (**right**).

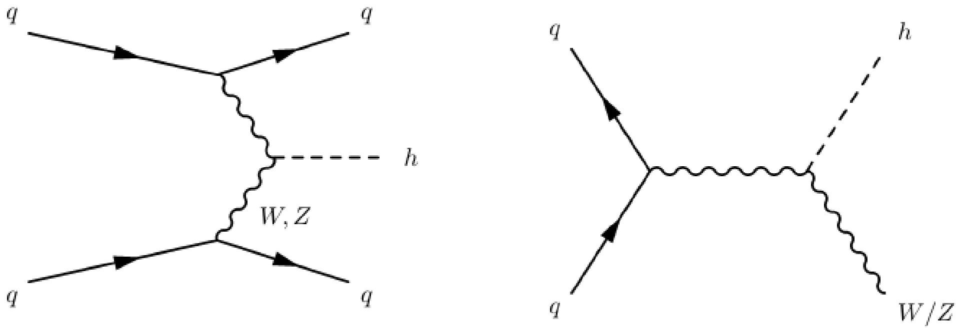

Figure 2.2: Feynman diagrams for Higgs production via weak vector boson fusion (**left**) and W/Z associated production (**right**).

Alternately, the Higgs can be produced via its couplings with the W and Z bosons. The dominant process here is weak vector boson fusion (VBF), where the Higgs propagator is connected to a t-channel W or Z exchanged between two quarks. In this case, the Higgs is always produced in association with two final-state quarks, which tend to have certain kinematic properties. This makes it convenient for experimental purposes, as these two quarks result in hadronic jets which can be used for tagging the events. Apart from VBF, the Higgs can simply be produced via radiation from an s-channel W or Z boson. This is referred to as “associated production” or “Higgs-strahlung”. Both of these processes are shown in Figure 2.2. The inclusive cross sections for these four production modes at the LHC are shown in Table 2.1.

Production Process	Inclusive Cross Section
Gluon-gluon fusion	48.6 pb ± 7%
Vector boson fusion	3.78 pb ± 3%
W/Z associated production	2.25 pb ± 4%
$t\bar{t}$ associated production	507 fb ± 8%

Table 2.1: Inclusive SM cross sections for production of a 125 GeV Higgs boson in 13 TeV pp collisions [32].

Decay mode	Branching ratio
$b\bar{b}$	58.2%
W^+W^-	21.4%
gg	8.2%
$\tau\bar{\tau}$	6.3%
$c\bar{c}$	2.9%
ZZ	2.6%
$\gamma\gamma$	0.23%
$Z\gamma$	0.15%

Table 2.2: SM decays of a 125 GeV Higgs boson [32] (list truncated at 0.1%). All have a 2% relative uncertainty or less.

The decays of the Higgs are governed by its coupling to the other SM particles. The largest coupling is to the top quark, but decays to $t\bar{t}$ are kinematically forbidden due to the large mass of the top. In principle both top quarks could be off-shell, but this is so heavily suppressed that it isn't measurable. So, the most common decays of the Higgs are to b quarks (with a branching fraction of 57.7%), with decays to the lighter fermions being much rarer. Decays to weak vector bosons (W^+W^- and ZZ) are also possible and have a large coupling, although at least one of the vectors needs to be off-shell. These together make up 24.1% of Higgs decays. Finally, although the Higgs has no direct coupling to photons or gluons, decays to these can occur through fermion loops. The SM branching ratios for the dominant decay modes are shown in Table 2.2.

2.4 Limitations of the Standard Model

Although the SM has been remarkably successful in accurately describing experimental observations, it is never regarded as a complete theory of physics. There are several reasons for this. First and foremost, there are some observations which the SM does not (and makes no attempt to) account for. The most obvious example is gravity: our best existing theory for describing this (General Relativity) is completely separate from the SM and is based on a different fundamental framework.

Formulating a quantized description of gravity which is compatible with the SM remains one of the major open problems in physics.

The other major observation which the SM does not explain is that of dark matter (DM). The existence of DM is inferred from several types of astrophysical measurements, including galaxy rotation speeds, gravitational lensing around galaxy clusters, and cosmic microwave background effects. In all of these cases, the magnitude of the gravitational effects are much larger than can be accounted for by the observed luminous matter. This implies that either our description of gravity is incorrect at these scales or there is additional matter which observations in the electromagnetic spectrum do not reveal. The possibility of a modified theory of gravity has been studied extensively, but these solutions face substantial difficulties. For example, "Bullet Cluster" observations are extremely difficult to reconcile with most modified gravity theories. This is a weak-lensing measurement of two galaxy clusters after a collision, which clearly shows a distribution of the lensing potential inconsistent with the visible matter, even under most modified gravity models [33]. The dark matter hypothesis states that these anomalous observations are in fact due to the presence of some as-yet undiscovered form of matter. This idea, and how it relates to what can be observed at energy-frontier colliders, is discussed in more detail in Section 2.5.

Aside from these observations which suggest incompleteness of the theory, there are theoretical issues with the SM as well. Perhaps the most well-known of these come in the form of "naturalness problems", such as that of the Higgs mass. This problem arises because loop corrections to m_H^2 are proportional to Λ^2, where Λ is the cutoff scale[6] used to renormalize the theory. Taking Λ to be on the order of the GUT scale, these corrections are many orders of magnitude larger than the observed mass of 125 GeV. This implies that the bare mass (which is a fundamental constant of nature) cancels out these corrections almost perfectly. Since there is no reason *a priori* why this should be true, this is often considered suggestive of some deeper mechanism at play. Another closely related issue is that gravity's coupling to matter is many orders of magnitude weaker than that of the other three interactions. Still another example is the fact that the strong interaction does not violate CP symmetry (the SM Lagrangian can contain a term for this, but the observed angle associated with it is not measurably different from zero). It should be noted that these naturalness problems have no implications for the consistency of the SM or its ability to accurately describe observations.

---

[6] The "cutoff scale" is the energy at which the SM is no longer accurate as an effective field theory. That is, it corresponds to the scale at which new physics becomes relevant.

2.5 Dark Matter Models

There are many models that try to provide an explanation for dark matter. Arguably the simplest is that it consists of non-luminous objects comprised of SM particles (such as dim dwarf stars or black holes). However, this hypothesis is strongly disfavored by experimental observations [34, 35]. Another is that DM is made up of axions, which are very light particles associated with a field proposed to solve an entirely different problem (in [36, 37], to address the question of why the strong interaction conserves CP symmetry). Collider experiments are generally not sensitive to axions, so they will not be considered further here.

Arguably the most popular DM candidate is the "Weakly Interacting Massive Particle" (WIMP). This is the general term for a new stable (at least on the timescale of the universe) particle which somehow interacts with the SM. This particle is considered to have some new field associated with it, so that it can be described in terms of the same quantum field theory as the SM. This field could be of any type (scalar, spinor, vector, etc.). Assuming all particles were in a high-temperature thermal equilibrium in the early Universe, there must be some mechanism for WIMPs to annihilate into SM particles in order to explain their current abundance.

Based on the observed DM relic abundance, it is possible to estimate its annihilation cross section. This is known from the relationship between the cross section and the DM abundance at the time of freeze-out[7]. Carrying out this calculation results in cross sections which are on the weak scale. Equivalently, assuming an effective coupling around $g \sim 1$, this predicts a WIMP mass on the order of 100 GeV. This apparent coincidence is known as the "WIMP Miracle", and is usually considered suggestive that this particle could be produced at the LHC.

There are three types of methods that can be used to search for WIMPs:

- **Direct detection:** This involves building a detector which is capable of measuring the extremely weak interactions between WIMPs and SM particles, usually scattering from a nucleus in the active material. These detectors are generally located deep underground (to minimize backgrounds from cosmic SM backgrounds) and are kept extremely cold (to minimize thermal noise which would otherwise wash out the signal). Many of the tightest constraints on dark matter come from the LUX [38, 39] and XENON1T [40] experiments.

---

[7] "Freeze-out" refers to the time after which the DM abundance is roughly constant, since the universe's expansion has rendered it too diffuse to continue annihilating at a substantial rate.

- **Indirect detection:** These experiments search for the products of DM annihilation, such as photons or positrons, particularly in the earlier universe. Many (but not all) of these are located on satellites, in order to remove the effects of Earth's atmosphere. Modern examples include Fermi-LAT [41] and AMS [42, 43], among numerous others (many of which have functions beyond just searching for DM). See [44] for a fairly recent review of relevant experiments.

- **Collider searches:** If WIMPs can annihilate into SM particles, then the reverse process can also occur if there is sufficient energy. If dark matter exists on the electroweak scale or below, then it can in principle be produced at the LHC. The precise mode of production depends on the nature of the DM and how it couples to SM particles, so searches are carried out in many channels. Typically these look for an excess of events with missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) in the final state[8]. However, if DM couples to the SM via some undiscovered mediator, then it is also possible (and in some cases, more effective) to instead search for the mediator itself[9].

These three methods all utilize the same underlying coupling between the SM and DM particles. The only difference in the processes they search for is the orientation of the Feynman diagram (Figure 2.3).

In order for the early-universe DM annihilation to occur, there needs to be some interaction (other than gravity) between the WIMP and SM particles. This interaction can be described in

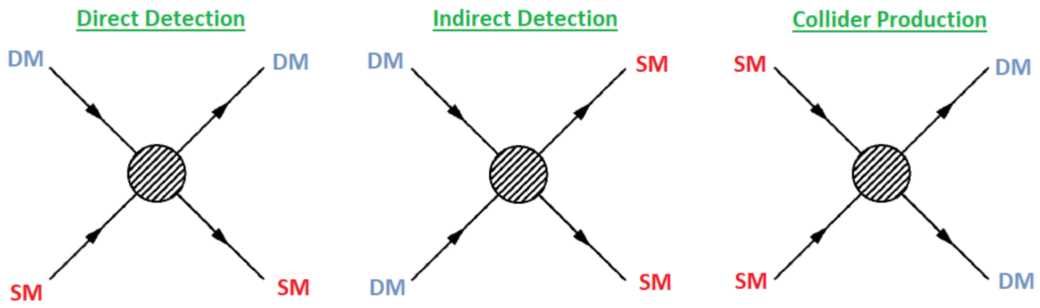

Figure 2.3: Feynman diagrams showing the processes used in direct detection, indirect detection, and collider searches for DM. The SM and DM particles are not necessarily fermions; the arrows only indicate the direction of time.

---

[8] A typical example is the search for DM in association with jets at ATLAS [45].

[9] A typical example is the search for dijet resonances at ATLAS [46].

several ways. The simplest is to postulate the existence of a heavy mediator particle which couples to both WIMPs and the SM, such that it can be described using an effective field theory (EFT). However, this mediator does not necessarily need to be very heavy, and so an EFT is not necessarily valid. To be more general, we can explicitly include this mediator in the model, but this introduces more free parameters (such as the mediator mass and its couplings). This class of theories are known as “simplified models”. They typically include a WIMP, a mediator, and little or nothing else, without attempting to provide a “reason” for adding them, other than to describe observations.

At the other end of the spectrum, it is possible to write down full UV-complete theories that include DM and its interactions with the SM as only one aspect of the whole. The most well-known example is supersymmetry (SUSY), which naturally provides a dark matter candidate in the form of neutral gauginos (in addition to resolving several other problems with the SM). Although this type of theory is the most complete, it lacks generality and can easily introduce dozens of additional free parameters. For this reason, this work interprets experimental results only in terms of simplified models. Note that simplified models are often a good description of the relevant dynamics from a more complete theory, and in those cases they can be directly reinterpreted as such.

One of the most-studied simplified models is the “Higgs portal” model. Here, the mediator is between the SM and the WIMPs is the SM Higgs. This is theoretically well-motivated, as the Higgs couples to *all* particles unless there is a specific reason otherwise (photon and gluon couplings are forbidden by the SM gauge symmetries). It also removes the need for a new mediator; only one particle is added to the SM. For example, the absolute simplest case would have the WIMP as a real scalar field which is not charged under any SM gauge groups. This would be written as:

$$\mathcal{L}_\chi = \frac{1}{2}\partial_\mu \chi \partial^\mu \chi - \frac{m_\chi^2}{2}\chi^2 + \lambda_\chi \phi^\dagger \phi \chi^2 \tag{2.8}$$

where χ denotes our new scalar field and ϕ is the SM Higgs field. This model could also allow for a χ^4 self-interaction (omitted in this example for simplicity).

Other scenarios for what χ is (complex fields, fermions, etc.) require more complicated constructions, but are qualitatively the same in terms of LHC phenomenology. However, this is not the case for direct and indirect DM detection experiments, which are generally quite sensitive to the spin structure of the interaction. So, while collider results typically make no assumption about the the spin of χ, comparisons to other experiments are dependent on this choice.

In the Higgs portal scenario, the coupling between the the Higgs and the DM field χ implies that if $2m_\chi < m_h$, then h can decay into a pair of χ. This would result in an invisible branching ratio $\mathcal{B}(h \to \text{inv.})$ which is larger than the SM prediction (roughly 10^{-3}). Experimental constraints on $\mathcal{B}(h \to \text{inv.})$ are then equivalent to constraints on these models. In practice, the best way to place limits on $\mathcal{B}(h \to \text{inv.})$ is currently through the VBF channel. Although gluon-gluon fusion has a much larger cross section, the very large SM backgrounds produced at hadron colliders make it less sensitive overall.

Outside of the Higgs portal, there are many versions of simplified models which introduce a new mediator particle. This can be anything in general, and is typically considered to be fairly heavy (at least on the order of 100 GeV). These models can be constrained by SM measurements, especially in the case of VBF production. Any (sufficienctly heavy) particle which is produced in VBF and decays invisibly can also decay to W^+W^- or ZZ, so existing measurements in those visible channels can be interpreted as limits on the production cross section for this new mediator.

In recent years it has become common for DM searches at the LHC to interpret their results in terms of a benchmark model of this type. This provides information to the theory community which is easily reinterpreted in terms of other, potentially more complete models. For an in-depth discussion of these models and the channels used to probe them, see [47] and the references therein.

The search for dark matter in the VBF+$E_\mathrm{T}^\mathrm{miss}$ channel can also be reinterpreted in terms of these other mediator models. The Feynman diagram is exactly the same: VBF production of the mediator, which decays to $\chi\bar{\chi}$. In order for this channel to be sensitive, the mediator must couple to W/Z and be heavier than $2m_\chi$. Most of the older simplified models do not consider this case; they typically include only couplings to quarks for simplicity. In fact, the only collider searches sensitive to these weak couplings are VBF+$E_\mathrm{T}^\mathrm{miss}$ and $W/Z + E_\mathrm{T}^\mathrm{miss}$.

Results from other experiments using very different techniques can also constrain the space of DM models. Direct detection experiments place limits on the WIMP-nucleon scattering cross section, shown in Figure 2.4 as a function of the WIMP mass. This is the total cross section for an inelastic collision between a WIMP and an individual nucleon in the detector material, and is agnostic to the actual coupling mechanism. Converting between this quantity and the production cross section at the LHC depends on the spin of χ, since the coupling structure does matter in this case. In some cases (such as light pseudoscalar DM), collider searches such as the one presented here are much more sensitive. In other cases (such as heavier scalar DM), direct detection experiments set the strongest limits.

Astrophysical measurements also provide us with constraints on dark matter. In particular, its relic density in the universe is closely related to the interactions between DM and the SM. This is because, assuming a standard thermal history of the universe, this number is set by the DM self-annihilation cross section. That cross section is in turn determined by the coupling between DM and the lighter SM particles, typically through some mediator. The most precise measurement of the relic density is currently from the Planck satellite [48]:

$$\Omega_{DM}h^2 = 0.1196 \pm 0.0031 \tag{2.9}$$

This constrains the parameters of any dark matter model, assuming that the thermal history of the universe is well-modeled, there is only one DM particle, and there are no additional mechanisms that impact the relic density (such as co-annihilation or effects from new particles at higher scales). The simplest Higgs portal models are already excluded by invisible Higgs constraints under these assumptions, provided m_χ is sufficiently smaller than $m_h/2$ [49].

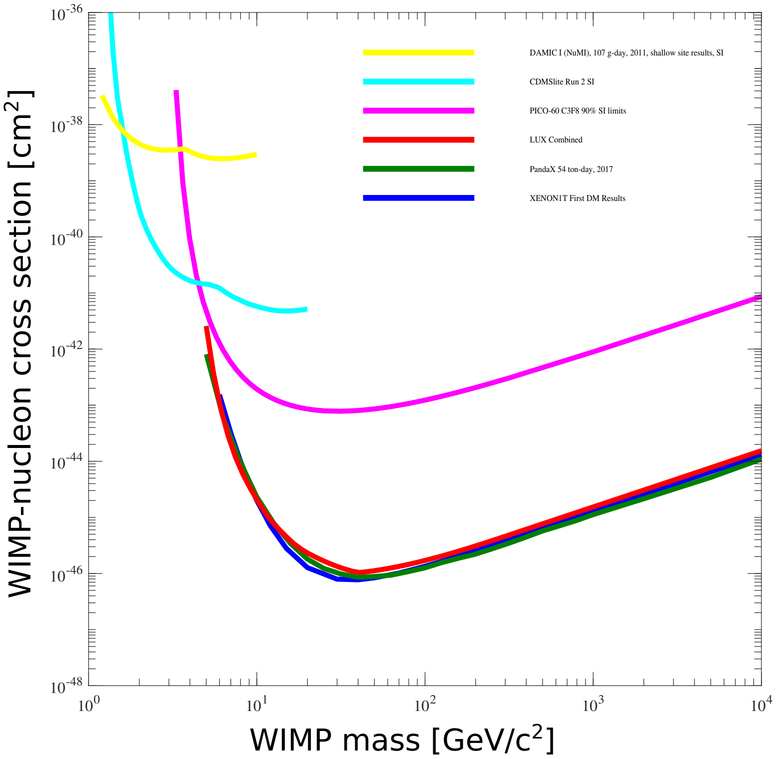

Figure 2.4: Summary of the latest spin-independent WIMP-nucleon scattering cross section limits from direct detection experiments at 90% C.L. [39, 40, 50, 51, 52, 53].

Chapter 3

LHC and the ATLAS Detector

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [54] is an accelerator designed to produce proton-proton (pp) and heavy ion (usually Pb-Pb) collisions at the highest energies currently possible. It consists of a 27 km ring of superconducting magnets, as well as several radiofrequency cavities which accelerate the particles. This work focuses only on physics with pp collisions, so the details of using heavy ions will not be discussed further here. Proton beams are accelerated up to a maximum energy of 6.5 TeV, for a center-of-mass energy of 13 TeV. Since protons are composite objects, the actual momentum transfer between two partons in any hard process is usually much less than this (this is governed by the proton parton distribution function (PDF). Since the exact momentum of the individual partons along the beamline is not known in hadron collisions, particle kinematics are usually described in terms of their momenta tranverse to the beamline (denoted p_{T}).

The protons used in the LHC come from an ionized hydrogen source and are first sent through a series of smaller accelerators. These consist of an inital linear accelerator (Linac), Proton Sychrotron Booster, Proton Synchrotron, and Super Proton Synchrotron. After going through all of these steps, the protons each have an energy of 450 GeV, at which point they are injected into the LHC. They are injected in “bunches” of approximately 10^{11} protons, separated by 25 ns. The LHC has 3564 “slots” for these bunches, however they can’t all be filled (roughly two thirds at most) due to technical limitations. In practice, there are usually many consecutive filled bunches (called a “bunch train”) followed by a smaller number of consecutive empty slots. The exact size of these is often changed, depending on the state of the LHC and what the goals for a particular run are. The two proton beams circulate around the ring in opposite directions and can be focused together (using a set

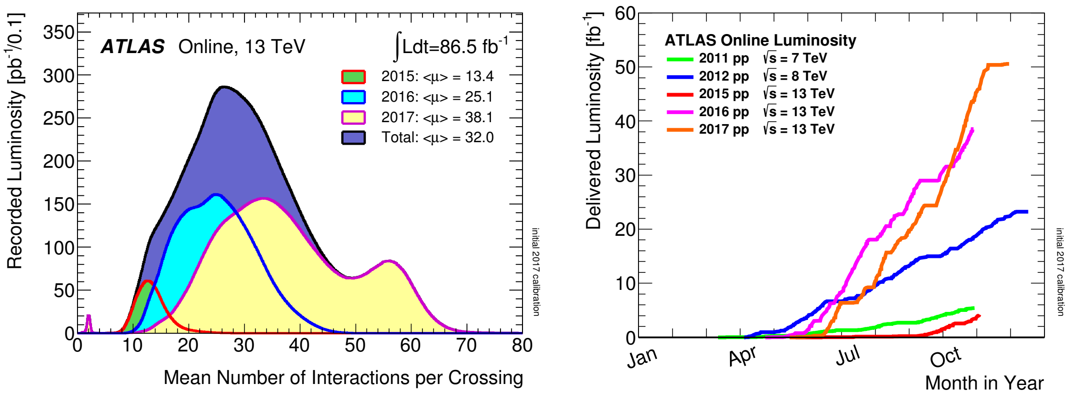

Figure 3.1: **Left:** Luminosity-weighted distribution of the mean number of interactions per bunch crossing for 2015–2017. **Right:** Integrated luminosity of pp collisions provided to ATLAS during stable beams.

of magnets) to produce collisions at several interaction points. There are four of these interaction points with major detector experiments located at each: ATLAS, CMS, LHCb, and ALICE.

Since each colliding bunch contains many protons, there are usually many hard pp interactions each time they cross. The actual number (denoted μ) depends on the beam parameters (such as the crossing angle and beam emittance), but mean values of μ ranged roughly from 10 to 45 in 2016 and as high as 65 in 2017 (Figure 3.1). Having many simultaneous interactions (also called "pileup") has implications for reconstructing what actually happened in a given event. For example, a large amount of pileup can degrade the energy resolution of measured particles, and sometimes ambiguities can arise over which interaction an object originated from (see Chapter 4 for further details). Furthermore, the detector response time for the calorimeters is longer than the bunch spacing of 25 ns, so objects from one collision can influence the reconstruction of objects from others. This is known as "out-of-time pileup".

It is possible to reduce pileup by adjusting beam parameters; however this comes at the cost of decreasing the instantaneous luminosity of the collisions. Instantaneous luminosity is defined in terms of the rate at which events occur ($\frac{dN}{dt}$) and the cross section for the process (σ) as:

$$\frac{dN}{dt} = \sigma \mathcal{L} \tag{3.1}$$

Integrated luminosity is then the integral of this quantity over time, and therefore the total number of events observed is directly propotional to it. Since we often want to search for very rare processes at the LHC, it is crucial to collect data for as many collisions as possible. This means it is generally preferable to use the highest luminosity the accelerator can support. The LHC was designed to achieve an instantaneous luminosity of 1.0×10^{34} cm^{-2}s^{-1}, but values as high as 2.0×10^{34} cm^{-2}s^{-1} have been achieved in 2017. The total integrated luminosity of the datasets collected by ATLAS is shown in Figure 3.1.

3.2 The ATLAS Detector

ATLAS is a general-purpose particle detector located at one of the LHC's interaction points. It consists of several specialized subdetectors which work together to reconstruct leptons, photons, and hadronic jets resulting from the collisions. The innermost layers form the Inner Detector, which provides tracking for all electrically charge particles. Outside of this are the calorimeters, which measure energy deposits from individual particles or showers. The calorimeters consist of electromagnetic layers (which have fine granularity and are specialized for measuring electrons and photons) and hadronic layers (which capture hadronic showers from jets that penetrate the electromagnetic layers). The outermost layers of ATLAS form the muon spectrometer, which provides additional tracking specifically for muons, since they generally escape through the entire detector. This section will give a brief overview of each of these main components; for a more in-depth overall technical description of the ATLAS detector, see [55].

The coordinate system used by ATLAS is a right-handed one with the origin at the nominal interaction point. The z-axis is along the beam pipe, the x-axis points toward the center of the LHC ring, and the y-axis is vertically up. For practical purposes, cylindrical coordinates (r, ϕ) are generally used in the plane transverse to the z-axis. Here, ϕ is the azimuthal angle around the z-axis. Another useful coordinate is the pseudorapidity η, which is defined in terms of the polar angle θ as:

$$\eta = -\ln\tan(\theta/2) \tag{3.2}$$

This variable is used rather than θ because the difference in η between two objects is invariant under Lorentz boosts along the z-axis. This can all be put together to define a distance measure ΔR between any two objects in (η, ϕ) which is invariant under those same boosts:

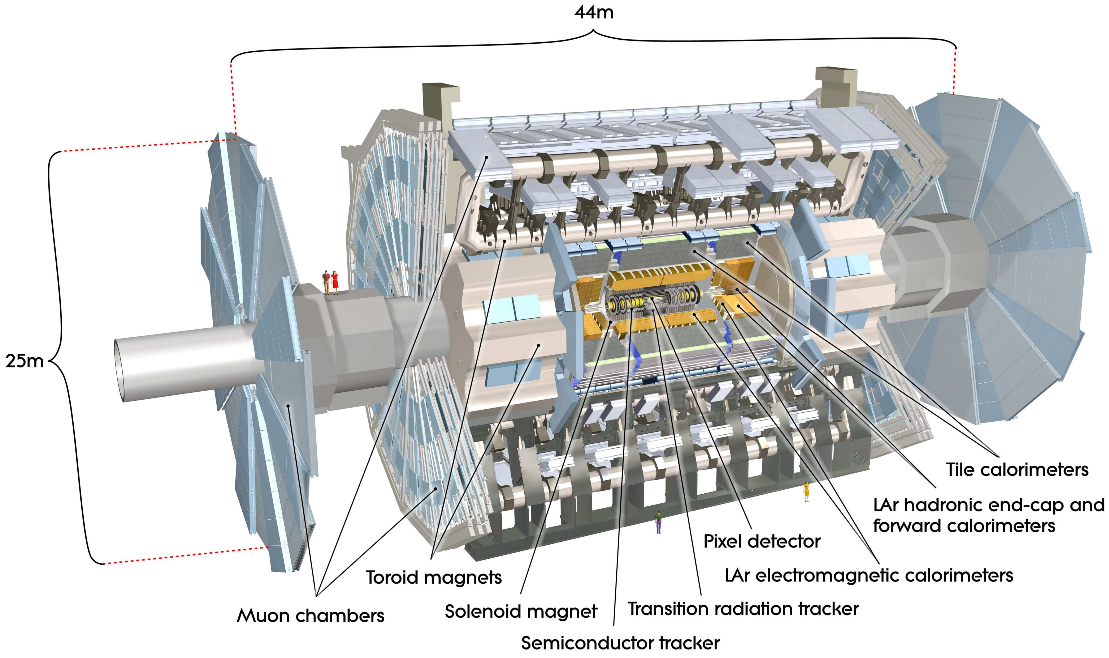

Figure 3.2: General cut-away view of the ATLAS detector [55].

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.3}$$

where $\Delta\eta$ and $\Delta\phi$ are the distances between the two objects in η and ϕ. The cyclical nature of ϕ is always accounted for here. For example, two objects with $\phi_1 = \frac{7\pi}{4}$ and $\phi_2 = \frac{-7\pi}{4}$ would have $\Delta\phi = \frac{\pi}{2}$.

3.2.1 The Inner Detector

The Inner Detector serves the primary purpose of measuring the trajectories of charged particles produced in the collisions at ATLAS. It provides this tracking capability in a pseudorapidity range of $|\eta| < 2.5$. It consists of three subdetectors: the Pixel Detector, the Semiconductor Tracker (SCT), and the Transition Radiation Tracker (TRT). Together, these read out the spatial positions of "hits", where charged particles leave localized energy deposits. These hits are later reconstructed into tracks, which represent that actual path of the particle. The Inner Detector is located inside a solenoid magnet, which produces a uniform 2 T magnetic field in the z-direction. This creates a curvature in the path of any charged particle, which is measured. This allows us to deduce that particle's momentum in the transverse plane p_T, since it is given directly by:

$$r = \frac{p_\mathrm{T}}{qB} \tag{3.4}$$

where r is the radius of the track's curvature, q is the particle's charge (which is always ± 1), and B is the magnetic field. For the complete technical specifications of the Inner Detector, see [56, 57].

3.2.1.1 Pixel Detector

The Pixel detector consists of four cylindrical barrel layers and three disk-shaped endcap layers. Together, these provide tracking information at the smallest radii for $|\eta| < 2.5$. Each layer consists of many sensors, each of which is an instrumented silicon wafer measuring 16.4×60.8 mm [58]. Ionization in the silicon due to charged particles passing through it provides the signal. Each sensor is divided into 46080 individual pixels ($50 \times 400 \mu$m each). In total, the Pixel detector contains 1744

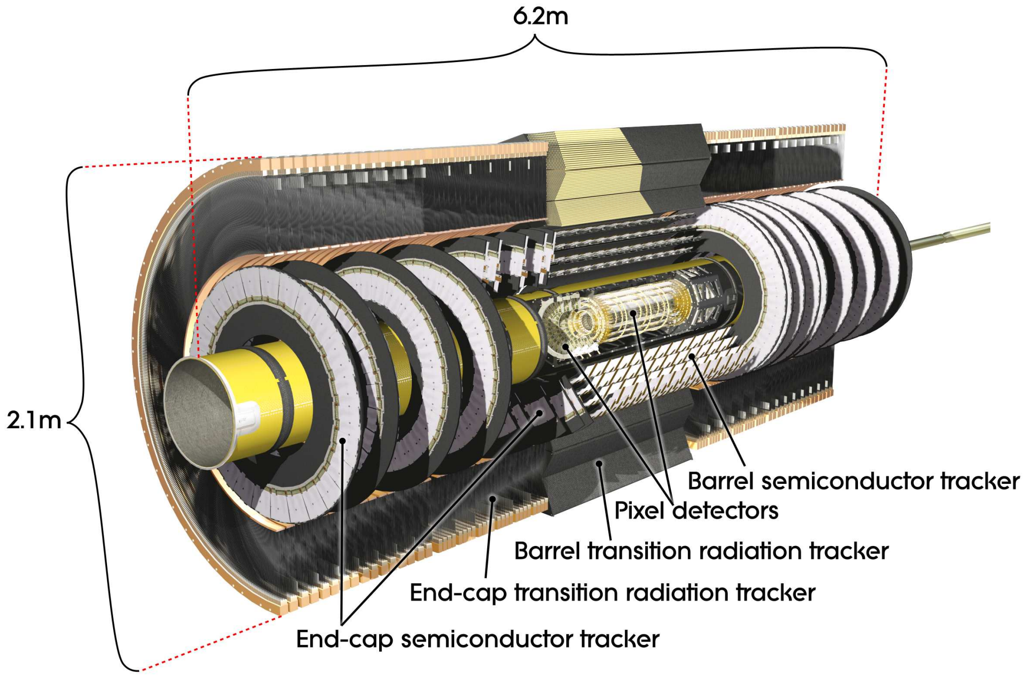

Figure 3.3: General cut-away view of the Inner Detector [55].

of these sensors, giving roughly 80 million individual readout channels.

The four barrel layers are located at radial positions of 33.4 mm, 50.5 mm, 88.5 mm, and 122.5 mm. The innermost of these (the Insertable B-Layer, or IBL) was newly installed in 2014. This provides substantially improved d_0 resolution[10] for ID tracks, which is praticularly beneficial for identifying B-hadron decays. (since their relatively long lifetime leads to seconday vertices which are measurably displaced from the primary interaction point).

3.2.1.2 Semiconductor Tracker

The Semiconductor Tracker uses the same basic technology as the Pixels, but the fundamental unit of silicon is a larger "strip". Each sensor is a rectangle with dimensions between 55 mm and 75 mm (depending on the layer and region), divided up into 768 individual strips [59]. The SCT has four cylindrical barrel layers (at radii of 299 mm, 371 mm, 443 mm, and 514 mm) and nine disk-shaped end-cap layers. This also provides tracking coverage out to $|\eta| = 2.5$. The main reason for using the larger strips here instead of more Pixel layers is cost, as well as the substantially larger quantity of material required.

3.2.1.3 Transition Radiation Tracker

The Transition Radiation Tracker is the outermost component of the ID, and covers only the region $|\eta| < 2.0$. It uses proportional drift tubes (known as "straws") filled with a gas mixture (usually xenon-based) as the basic detector element [60]. Each straw consists of a cylindrical cathode, with an anode wire running along its axis. The straw layers are interleaved with polypropylene fibers (in the barrel) or foils (in the end-cap) [61, 62]. Since this has a different dielectric constant from the other surrounding materials, charges particles crossing those boundaries emit low-energy (keV-scale) photons, known as transition radiation. These photons ionize the gas in the tubes, which creates an electron avalanche due to the high electric field. The resulting electrons are collected on the anode wire to give a readout signal.

The amount of transition radiation (i.e. number of photons) a particle gives off is proportional to its Lorentz factor γ. Since lighter particles have a higher γ for fixed momentum, this can be used to discriminate between charged particles of different mass. To this end, the readout electronics for each straw have three possible outputs for each bunch crossing: zero, "low threshold", and "high threshold". The high threshold is set such that electrons are far more likely than anything else

---

[10] d_0 denotes the distance of closest approach (in the xy plane) of the track to the beam axis.

to produce hits that pass. These high-threshold hits are then used as an input to the electron identification algorithm (see Section 4.2).

3.2.2 The Calorimeters

ATLAS includes two types of calorimeter system for measuring electromagnetic and hadronic showers. These are the Liquid Argon (LAr) calorimeters and the Tile calorimeters. Together, these cover the region with $|\eta| < 4.9$.

The thickness of the calorimeters is determined by the depth of the showers of interest. Electromagnetic showers, originating from photons and electrons, tend to be short (with a characteristic length scale on the order of centimeters, depending on the medium) and very narrow. They consist primarily of photons from bremsstrahlung and e^+e^- pairs, which deposit nearly all of their energy if the first few layers of the LAr calorimeter. The LAr calorimeter, which is primarily designed with these showers in mind, was therefore built with a minimum thickness of 24 radiation lengths.

Hadronic showers (also known as jets), have a much longer and wider shape. These originate from a quark or gluon, which hadronizes and showers via the strong introduction. Hadrons travel much farther through material as they lose energy, resulting in a shower that (for any significant initial-particle energy) typically penetrates through the entire depth of the LAr calorimeter. For this reason, energy deposits in the Tile calorimeter are usually associated with hadronic showers. It is also worth noting that a substantial number of hadrons produced in these showers are pions, which can decay electromagnetically.

The reconstruction and interpretation of these showers is discussed in Sections 4.1 and 4.2 for jets and electrons respectively.

3.2.2.1 Liquid Argon Calorimeters

The Liquid Argon system consists of four individual calorimeters: the Electromagnetic Barrel (EMB), Electromagnetic End-cap (EMEC), Hadronic End-cap (HEC), and Forward Calorimeter (FCAL). These are all sampling calorimeters, which use liquid argon as the active material interleaved with metal absorber layers. The active regions are equipped with electrodes, which measure the ionization of the argon due to passing particles. The EMB and EMEC are primarily designed around reconstructing electrons and photons, while the HEC and FCal are geared toward measuring hadronic showers (jets). A brief description of each is given here; for the complete technical documentation see [63]. The total length of the ouput signal from an energy deposit in the LAr calorimeter is several hundred ns (Figure 3.5). As a result, activity in one event can influence how the following bunch crossings are reconstructed.

The EMB covers the region $|\eta| < 1.4$. It uses lead absorber between the active LAr layers in an "accordion" geometry (Figure 3.6). The exact layout and thickness of the absorber depends on the layer and location within the detector, but it is typically between 1 mm and 1.5 mm thick. The EMB consists of four layers of cells. The first (front) consists of strips which are very finely segmented in η: they are read out at a granularity of 0.003125×0.1. This level of η-resolution is particularly helpful for e/γ reconstruction. The second (middle) layer is thicker, and the cell size is 0.025×0.025 For the third (back) layer, the cells size is 0.05×0.025. In addition to these three primary layers, there is a "presampler", which primarily serves to help correct for energy lost before the shower reaches the calorimeter. This is the innermost layer, and it has a readout granularity of 0.025×0.1.

The EMEC has a very similar design, and covers the region $1.375 < |\eta| < 3.2$. Like the EMB, this has an accordion geometry with four layers, including a presampler. The η granularity of the front layer is slightly coarser in some parts of the EMEC, but the middle and back layers are the same. The lead absorber is slightly thicker here (1.7-2.2 mm).

The HEC is designed somewhat differently, as it is intended only for measuring hadronic objects. It covers the region $1.5 < |\eta| < 3.2$ and sits directly behind the EMEC. This uses copper absorbers which are much thicker than the lead ones used in the EM calorimeters (25-50 mm). The geometry consists of cells which are truly rectangular in (η, ϕ) rather than accordion-like. The ϕ-granularity is 0.1 at small values of η, and 0.2 closer to the beamline. The η-granularity varies similarly.

The FCAL covers the most forward region of ATLAS, $3.2 < |\eta| < 4.9$. It consists of three layers (numbered FCAL1, FCAL2, and FCAL3, in order of distance from the interaction point). The first (FCAL1) is nominally an electromagnetic calorimeter, and uses copper as its absorber. The other two layers are hadronic, and use tungsten absorbers (this is the only significant difference). Rather than having alternating layers of absorber and active material, the FCAL uses an array of tube electrodes embedded into an absorber matrix. The tubes are arranged such that their axes are all parallel to the beamline.

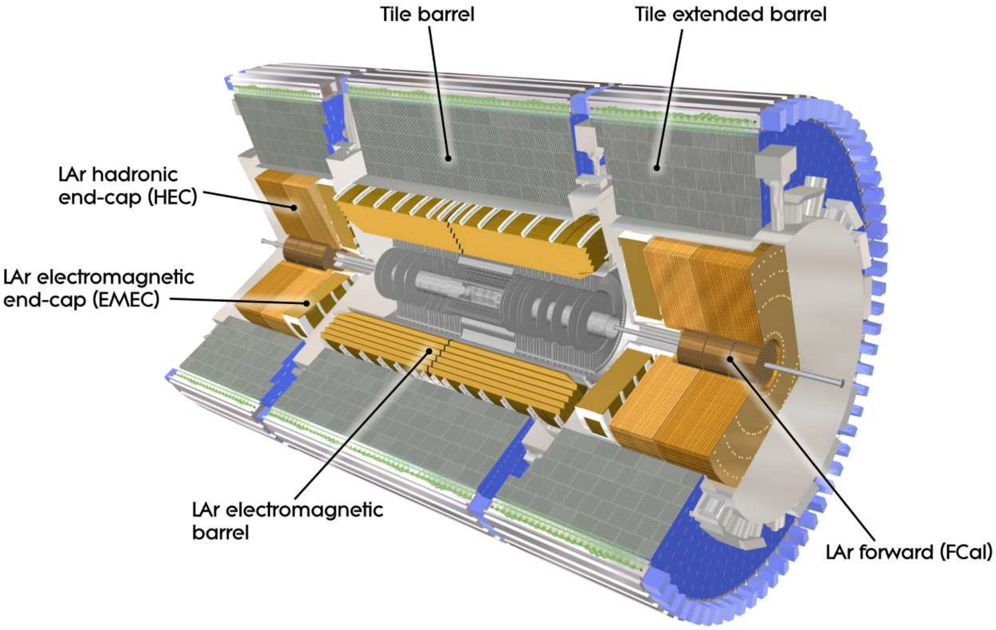

Figure 3.4: General cut-away view of the calorimeters [55].

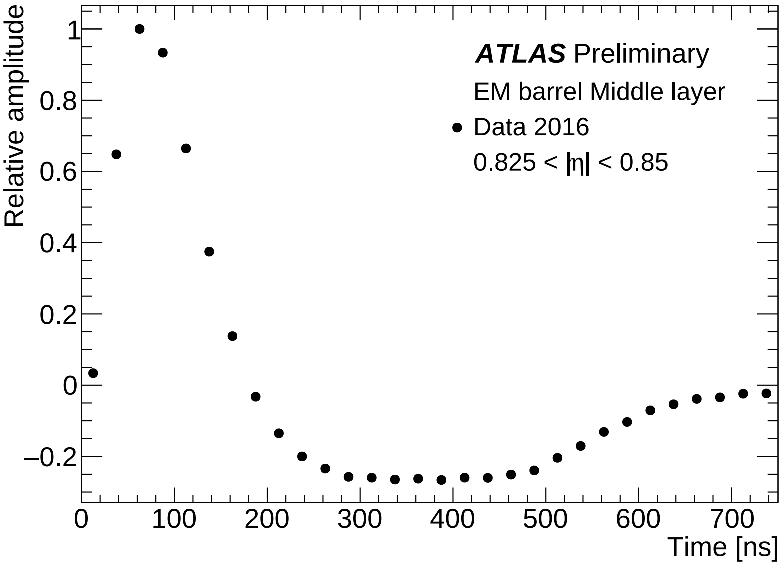

Figure 3.5: A typical ionization pulse shape in the EMB middle layer. This was taken from a special run with isolated bunch crossings in 2016.

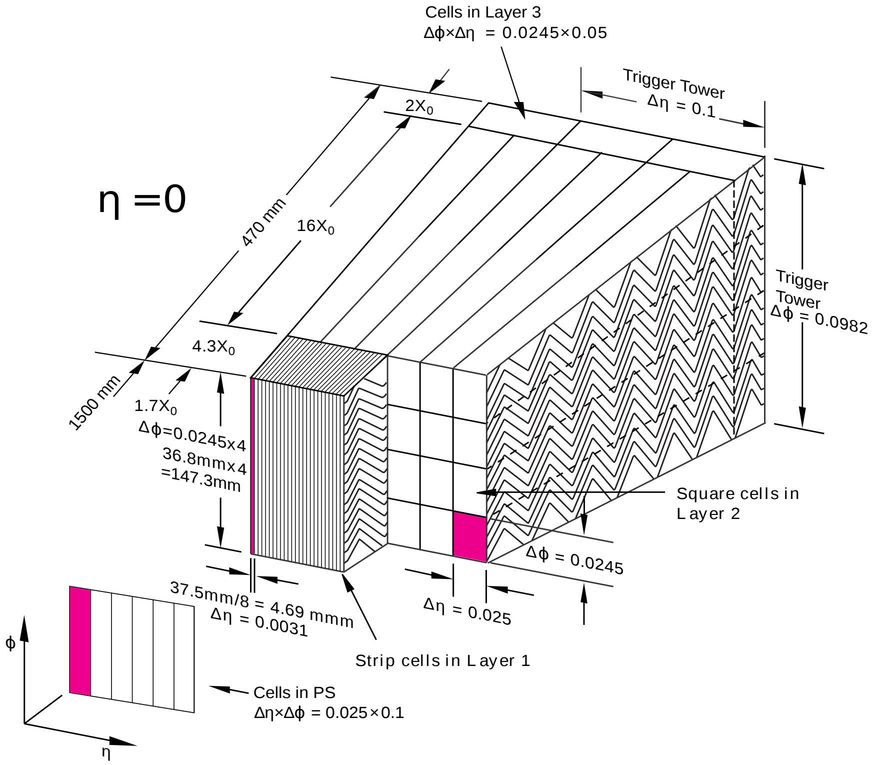

Figure 3.6: Schematic diagram of the accordion structure and layout of the EMB cells [63].

3.2.2.2 Tile Calorimeters

The Tile calorimeter provides coverage for hadronic showers in the region $|\eta| < 1.7$. Is consists of a “barrel” in the center, and an “extended barrel” on each side; both have the same design. It uses a solid plastic scintillator as its active material, arranged in many layers such that the total thickness is at least 9 interaction lengths for hadronic showers. Between these layers are sheets of steel absorber. Groups of the scintillator tiles are read out via photomultiplier tubes (PMTs) with a uniform granularity of 0.1×0.1 in (η, ϕ). Since the Tile calorimeter is only in the central region of the detector and is designed only for hadronic showers, it is much simpler than the LAr systems. The readout pulse for a signal is also much faster for the Tile calorimeter than for LAr, because it is coming directly from PMTs rather than electrodes. For the complete technical documentation on the Tile calorimeters, see [64].

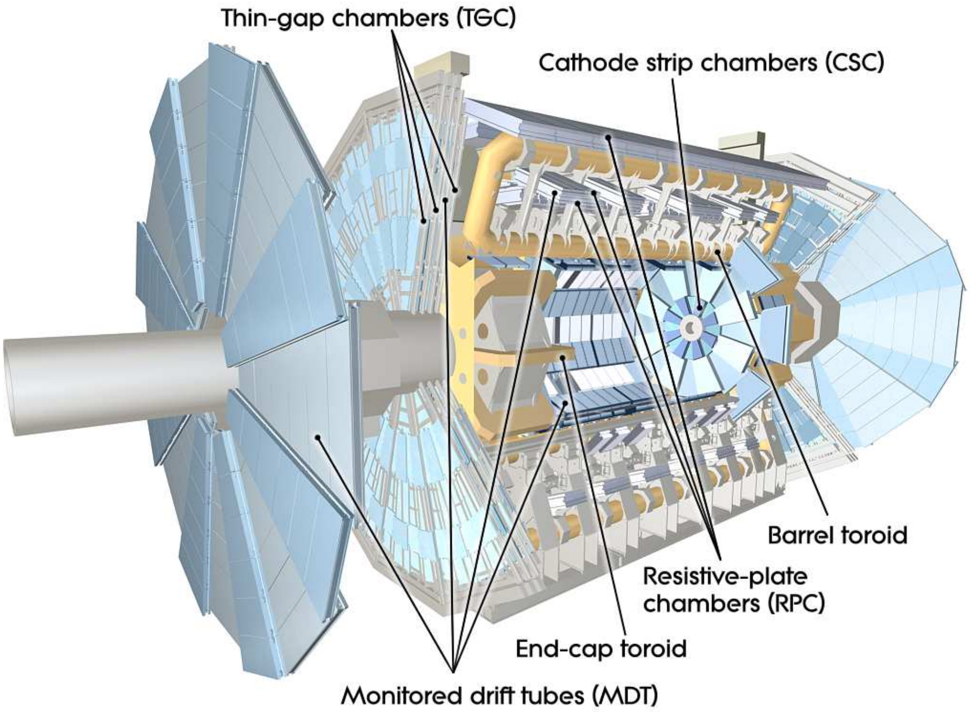

Figure 3.7: General cut-away view of the Muon Spectrometer [55].

3.2.3 The Muon Spectrometer

The Muon Spectrometer is the outermost component of the ATLAS detector, and provides tracking for muon reconstruction in the region $|\eta| < 2.7$ [55]. Most of the system lies in a large magnetic field provided by a set of toroid magnets. This serves to bend the trajectory of the muons, allowing the reconstruction of their momentum. The MS itself consists of four types of chambers, arranged in a barrel ($|\eta| < 1.05$) and two end-cap sections ($1.0 < |\eta| < 2.7$).

The entire η range is covered by Monitored Drift Tube (MDT) chambers. These provide the best resolution for muon tracking (60-80 μm for a single hit). Each MDT consists of an aluminum tube filled with an argon-based gas mixture and an anode wire running along its axis. When a muon passes through, some of the gas is ionized, and the electrons are collected on the wire to generate a readout signal. In the innermost region of the end-caps, the MDTs are replaced by Cathode Strip Chambers (CSCs). These have a similar design, but the cathodes are long strips rather than tubes, and there are many anode wires arranged in a plane in each chamber. Here, the cathode strips are used for readout (by measuring the charge induced on them by the anode wires). These provide finer granularity than MDTs can, which is important in this innermost region due to large rates and background conditions.

Additionally, there are layers of Resistive Plate Chambers (RPCs) in the barrel and Thin Gap Chambers (TGCs) in the endcap. These are primarily used for triggering purposes, but also provide additional tracking information for offline reconstruction with a resolution of 5-10 mm per hit. The RPCs consist of pairs of resistive plates separated by a narrow (2 mm) gap. This gap is filled with a gas mixture (based on tetrafluoroethane). A high voltage (8.9 kV) turns primary ionization electrons into avalanches, which are collected and read out. The TGCs are multiwire proportional chambers (very similar to the CSCs), with a few small differences in the dimensions and gas mixture. Like the RPCs, they are designed to read out very quickly so they can be used for triggering.

See [65] for the full MS technical specifications.

3.2.4 The Trigger System

Since collisions occur every 25 ns and each event requires approximately 1 MB of permanent storage to record, it is far beyond the capabilities of any existing technology to record every event. LHC physics analyses (particularly those studying electroweak/Higgs physics or looking for physics beyond the SM) are typically interested in rare processes. So, a system is required which can identify which events are likely to contain uncommon or interesting phenomena in real time, so that those can be recorded and the rest discarded. This is the function of the ATLAS trigger.

The trigger system consists of two levels: one implemented in hardware (Level-1, or L1) and one implemented in software (the High-Level Trigger, or HLT). In past years, the HLT had two sublevels (known as Level-2 and the Event Filter); however this is no longer necessary and only a single software level is now used. Their purpose is to reconstruct the same objects that are used in analysis (jets, electrons, etc.) with very limited time and resources, and then decide based on those objects whether to record the event. The event selection criteria are typically based on the multiplicity of objects in an event and their p_{T}. For example, any event containing an isolated electron over 26 GeV (according to some trigger-level electron algorithm) is recorded. ATLAS maintains an extensive "menu" of trigger items, each of which is a set of these selection criteria. If an event passes any one of these, it is recorded. Some trigger items are "prescaled", meaning that only a small, randomly-selected sample of events passing the selection criteria are actually recorded. These are typically used to study events which are too common to record all of, such as events containing only low-p_{T} jets which are useful for calibration purposes.

3.2.4.1 The Level-1 Trigger

The Level-1 trigger is implemented entirely in hardware. It uses information read out via dedicated L1 paths for the calorimeters ("L1Calo") and the muon spectrometer ("L1Muon"). Tracking information from the inner detector is not used at L1 (although a potential upgrade to add a hardware-level track trigger is being considered for Run 4). Muons are reconstructed from RPC and TGC information, while jets and electrons/photons are reconstructed exclusively from calorimeter information. There is no distinction between electron and photon objects at L1, due to the lack of ID tracking.

The L1Calo system reads out calorimeter information at a much coarser granularity than is available for the full "offline" reconstruction. Cells are grouped into "trigger towers", which are 0.1×0.1 squares in (η, ϕ) throughout most of the detector [66]. The tower size is larger (and the spatial resolution therefore worse) in the forward regions of the detector. These towers objects are sent to a set of boards containing two processing systems, which execute the actual L1 algorithms to identify the objects. The "Cluster Processor" searches for the localized energy deposits characteristic of electrons, photons, and hadronic τ decays. The "Jet and Energy-Sum Processor" searches for jets and calculates global energy sums, using 2×2 sums of towers called "jet elements". Both were orginally based on ASICS, but were upgraded following run 1 and now use FPGAs instead. The details of these algorithms are explained in Section 4.6.

The Level-1 muon system is based on track hits read out from the RPCs (for $|\eta| < 1.05$) and TGCs (for $1.05 < |\eta| < 2.4$) [66]. For any given η value within these regions, there are at least 3 planes of the muon detectors, each consisting of 2 to 4 layers. The readout electronics use programmable logic to directly construct muon candidates from coincidences across layers. This provides a p_{T}, η, and ϕ value for the muon candidate.

The reconstructed objects are then sent to the Level-1 Central Trigger Processor, which makes the final decision to accept or reject the event. If the event is accepted, a signal is sent to the individual subdetectors to read out their information to the HLT. This entire process needs to be repeated at a rate of 40 MHz in order to keep pace with the collisions. The latency between the collision occurring and the final L1 decision is required to be less than 2.5 μs.

3.2.4.2 The High-Level Trigger

The High-Level Trigger makes the final decision on whether an event is to be recorded to permanent storage. It works exactly the same as the L1 trigger in concept, but is implemented in software running on a large (tens of thousands of cores) processor farm rather than FPGAs and has more information to work with. This allows it to run much more complex algorithms, but it still has finite time and resources. It needs to keep up with the maximum L1 output rate of 100 kHz, and each event typically takes one HLT core 200-300 ms to process, depending on the pileup conditions. Ultimately, about 1 kHz of events are written to permanent storage via the “main physics” stream (i.e. that which is used for analysis). The HLT also sends a small number of events to other output streams for technical purposes such as calibration measurements and debugging abnormal events. To save resources, not all of the information for each event is always saved for these secondary streams. This results in much smaller event sizes, allowing for much higher trigger rates in these cases. In some cases, physics analyses are able to make use of even this limited information, such as ATLAS’s trigger-level dijet resonance search [67].

Compared to L1, the HLT uses more detailed readout from the ID, calorimeters, and MS to reconstruct objects as close as possible to those used offline in analysis. The full event data is read out from the detector and sent to the HLT farm, which runs its own dedicated algorithms to construct HLT objects. For example, the HLT recieves full-granularity calorimeter information, which allows it to reconstruct jets with near-offline precision. The HLT is also capable of carrying out full track ID reconstruction within small regions seeded by the L1 trigger. The details of these algorithms are described in Section 4.6.

CHAPTER 4

Reconstruction Algorithms

This chapter gives an overview of the algorithms used to reconstruct the objects produced in the *pp* collisions at ATLAS. "Reconstruction" refers to the process of converting raw detector data into objects with physical meaning in terms of particles. Of course, no reconstruction algorithm is perfect and the detector has limitations, so reconstructed objects don't always correspond exactly to the "true" objects in the event. For example, true electrons are not always successfully identified, and sometimes an electron is reconstructed where no true electron existed. The performance of these algorithms is often quantified in terms of **signal efficiency** (the probability to reconstruct an object given that a real one was actually present) and **signal purity** (the fraction of reconstructed objects that correspond to a true object).

Here, the focus is on those used for the invisible Higgs search presented in this dissertation (jets, electrons, muons, and missing transverse momentum). ATLAS is also capable of reconstructing photons and hadronic τ decays, but these are outside the scope of this work. Additionally, there are other varieties of jets not discussed here (larger-radius jets with substructure from boosted decays, jets from displaced secondary verticies due to B-hadron decays, etc.).

Several of the quantities used in the reconstruction algorithms depend on the identification of the **primary vertex**. This is the vertex which is chosen as the mostly likely candidate for where the hard interaction occurred. In general, there are many candidates due to pileup, so the one which is chosen is the vertex with the largest scalar sum of associated track p_{T}. The efficiency of this method varies depending on the event topology of interest[11], but in general it is very close to 100%.

---

[11] This is discussed further with respect to the VBF Invisible analysis in Section 4.4.

4.1 Jets

Jets (showers of hadrons originating from a quark or gluon) are reconstructed from energy deposits in the calorimeters. For each event, the energy of each calorimeter cell is read out and the cells are then grouped into topological clusters. The purpose of clustering is to suppress calorimeter noise by excluding cells which are unlikely to have energy from an actual jet in them. The topo-clustering algorithm [68] groups cells based on the ratio (denoted ζ_{cell}) of their energy in that event over an estimated noise value σ_{cell}:

$$\zeta_{cell} = \frac{E_{cell}}{\sigma_{cell}} \tag{4.1}$$

The noise value σ_{cell} is defined as the RMS of that cell's energy distribution, when no hard-scatter object is present. This consists of two components: electronic noise (measured when no *pp* collisions are present at all) and a contribution from pileup (which is estimated based on the expected pileup distribution for each run). Cells with $\zeta > 4$ are designated as "seed" cells and start the clusters. Next, any cells with $\zeta > 2$ neighboring a seed are added to the cluster. This process is repeated iteratively, so that any cell with $\zeta > 2$ neighboring *any part* of the cluster is added. Finally, all cells neighboring the cluster at this stage are added (Figure 4.1).

Since topo-clusters can grow large and merge together, an additional splitting step is performed after the initial clustering. The purpose of this is to separate nearby showers. All cells which are local energy maxima above some threshold are designated as seeds, and the clusters are split spatially between them, potentially resulting in more clusters than the first iteration. It is possible for ambiguity to arise in which cluster a cell belongs to (for example, if it neighbors multiple local maxima). In this case, the cell is shared between the two clusters with the highest-energy local maxima. The energy sharing between the two clusters in this case is defined by a geometrical weighting based on the energies of each cluster and the distance to their centroids[12].

The final clusters are used as inputs to the anti-k_t algorithm[13] [69] to give reconstructed jets. For the purposes of this algorithm, each topo-cluster is considered to have zero mass and energy equal to the sum of its constituent cells. The direction of the topo-cluster is defined as the average (weighted by absolute energy) direction of its consituent cells in (η, ϕ). In the context of this work,

---

[12] See [68] for the precise definitions of the splitting and weighting.

[13] The anti-k_t algorithm is described in Section 4.1.1.

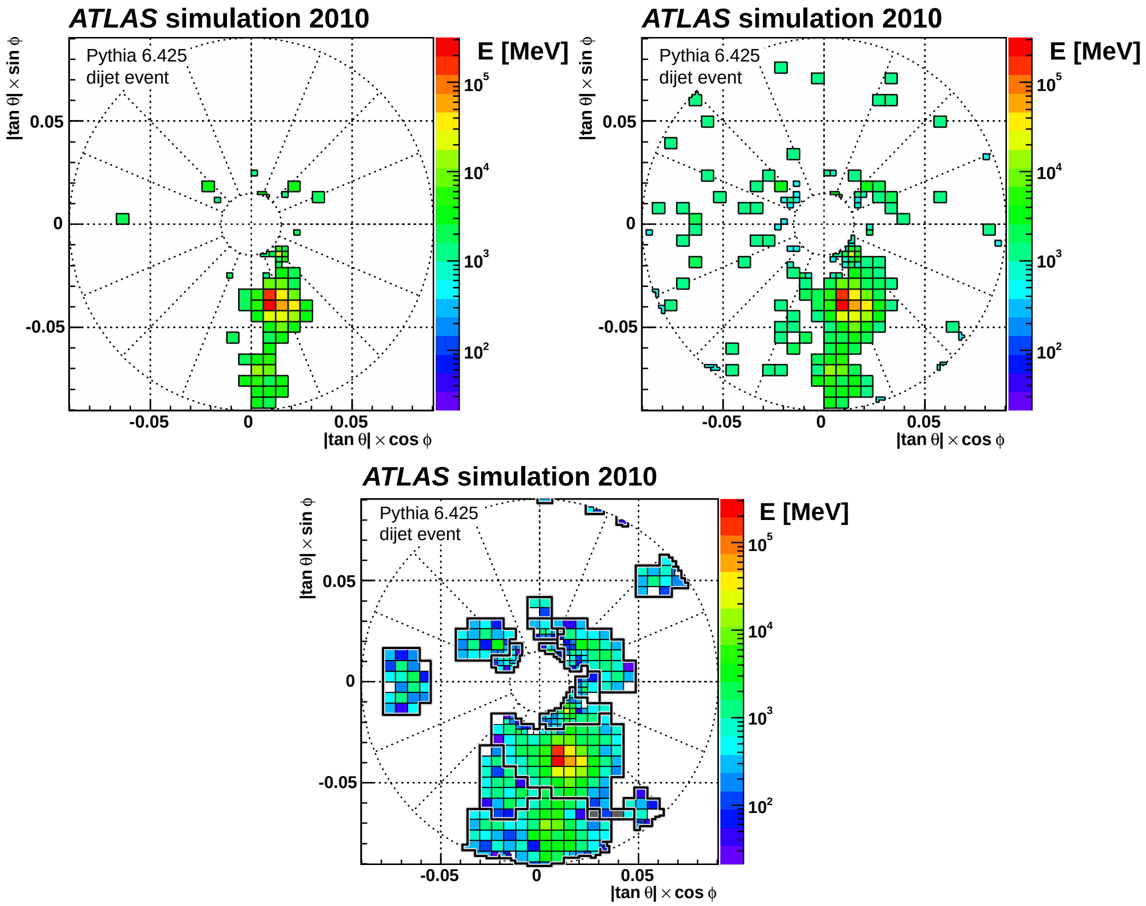

Figure 4.1: Stages of topo-cluster formation for an example simulated event in the first FCAL module [68]. **Upper left:** "Seed" cells passing $\zeta_{cell} > 4$. **Upper right:** "Neighbor" cells passing $\zeta_{cell} > 2$. **Bottom:** All cells in the final topo-clusters.

only jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 4.5$ are considered. This is driven by the difficulty of resolving jets at low p_{T}, and by the spatial extent of the detector at high $|\eta|$.

4.1.1 The Anti-k_t Algorithm

Jets are reconstructed by running the anti-k_t algorithm [69] on topo-clusters. This is a sequential clustering algorithm, which can be run on any input jet constituents (topo-clusters, truth-level particles from MC, etc.). The algorithm is based on defining a distance measure d_{ij} between each pair of input objects, and grouping them together according to that quantity. For anti-k_t, the distance measure is defined as:

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \frac{\Delta R_{ij}^2}{R^2} \tag{4.2}$$

Here, R is a free parameter which defines the characteristic size of the jets. All jets referred to throughout this dissertation are reconstructed with $R = 0.4$. Larger values are sometimes used in other contexts for jets resulting from hadronic decays of boosted objects (for example, top quarks). The pair of constituents with the smallest value of d_{ij} is clustered together to form a single object, and the distance measures are then recalculated with this new combined constituent. This process is repeated iteratively, until the smallest distance measure exceeds $d_{iB} = p_{Ti}^{-2}$ (the distance measure between object i and the beam). At this point, the algorithm is finished, and all of the combined clusters at that point are considered to be jets.

The anti-k_T algorithm has several useful properties. It is infrared- and collinear-safe, meaning the structure of the final jets is insensitive to soft radiation and collinear splitting. The output jets also tend to be circular in (η, ϕ), which is not the case with other sequential clustering algorithms. The exception is when two jets are nearby; in this case the overlapping area tends to be included in only the harder of the two. In some cases, nearby or overlapping jets can be part of a single larger "fat jet", originating from an object that decayed into multiple colored objects (such as top quarks). For these situations, a larger value of R can be chosen, and additional substructure and grooming techniques exist. However, these cases are outside the scope of this dissertation, as the focus here is on light-quark jets associated with vector boson fusion.

ATLAS uses the FastJet implementation [70] of the anti-k_t algorithm. This allows the algorithmic complexity to scale as $N \ln N$, where N is the number of input constituents. Before this method, the runtime of simpler implementations always scaled as N^3, which would have presented serious problems under the high-pileup conditions present at the LHC.

4.1.2 Jet Calibration

After the jets have been reconstructed, their energies (and directions) are calibrated to bring them as close to the "true" values as possible. Jet calibration is done in several steps [71]:

1. **Origin correction:** The angular direction of the topo-clusters is adjusted to originate from the event's primary vertex. This does not affect the jet's energy.

2. **Pileup correction:** An estimate of how much of the jet's p_{T} came from pileup is subtracted off. Initially, this is taken as the product of the per-event pileup p_{T} density ρ and the jet's area. The jet area is defined using "ghost particles", which are infinitesimal additional inputs (uniformly distributed in (η, ϕ) with a very fine granularity) to the anti-k_t algorithm. The number of these ghosts which ends up in a jet gives that jet's area. In addition to this, there is a residual pileup correction to remove any remaining dependence on $\langle\mu\rangle$ and N_{PV}. This residual correction is derived from Monte Carlo.
3. **MC-based calibration:** A re-scaling is applied to adjust the detector-level energy scale to that of the actual particles. In practice, "the actual particles" means visible truth-level particles from MC clustered with anti-k_t. This calibration is a simple scaling factor applied to the jet energy, defined as $c_{JES} = E_{truth}/E_{reco}$. It is calculated directly from MC as a function of the jet's η and p_{T}.
4. **Global sequential calibration:** Additional information about the topology of calorimeter deposits and tracks associated with the jet are used to apply another correction to the jet energy. This is intended to reduce the dependence of the energy response on other jet properties, particularly flavor. This does not affect the mean energy response, but improves the resolution substantially.
5. **Residual in-situ calibration:** A final data-driven correction is applied to cover residual differences between data and MC. At lower energies, these calibrations come from γ+jet, and Z+jet events. At higher energies, multijet events are used. These are calculated by taking a sample of clean, well-measured events and requiring that the p_{T} of all objects cancels out precisely. This is achieved by varying the energy of the jet, and the amount it needs to be varied by gives the final calibration.

After a jet has been reconstructed and calibrated, we are still faced with the problem of knowing whether it actually originated from the hard interaction we're interested in. There are two sources of jets we want to remove. The first is pileup: calorimeter deposits originating from collisions we're not actually interested in can create jets. There are techniques that can be used to suppress these jets; they are discussed further in Section 4.1.3. The second source of "bad" jets is non-collision background. These are reconstructed jets which did not originate from a *pp* collision at all. There are four main ways which these can come about:

- Noise fluctuations in the calorimeter.
- Cosmic muons traveling through the detector during an event.
- Scattering of beam protons with residual gas in the beam pipe.
- Scattering of beam-halo protons with collimators far from the interaction point.

The vast majority of non-collision jets can be removed by imposing requirements on jet properties. This is known as "jet cleaning", and is discussed further in Section 4.1.4.

4.1.3 Pileup Suppression

Under typical conditions for 13 TeV *pp* collisions at ATLAS, each bunch crossing contained several tens of hard interactions. Each of these can produce its own jets, which are reconstructed the same as those from the primary vertex. Additionally, soft QCD emissions from multiple vertices can coincidentally land in the same region of the calorimeters, resulting in "stochastic jets". Both of these cases can interfere with our understanding of the hard interaction of interest, so ideally they should be removed when possible. Particularly for VBF events (which tend to have a specific jet topology), failure to identify which jets are from pileup can seriously degrade the sensitivity of an analysis.

In order to remove pileup jets from consideration, a variable called the Jet Vertex Tagger (JVT) is used [72]. This is a multivariate combination of two track-based variables: corrJVF and R_{pT}. Both of these variables are ratios where the numerator is the scalar sum of the p_{T} of all tracks in the jet originating from the primary vertex ($\sum_k p_{\mathrm{T}}^{trk,k}(\mathrm{PV}_0)$). R_{pT} is the simpler of the two, its denominator is simply the total calibrated p_{T} of the jet:

$$R_{pT} = \frac{\sum_k p_{\mathrm{T}}^{trk,k}(\mathrm{PV}_0)}{p_{\mathrm{T}}^{jet}} \tag{4.3}$$

For corrJVF, the denominator is the p_{T} scalar sum of *all* tracks associated with the jet, regardless of the vertex they originated from. However, tracks from pileup vertices are weighted according to the total numbers of pileup tracks in the event n_{trk}^{PU}. This is intended to remove dependence on the total number of pileup vertices in the event.

$$\mathrm{corrJVF} = \frac{\sum_k p_{\mathrm{T}}^{trk,k}(\mathrm{PV}_0)}{\sum_k p_{\mathrm{T}}^{trk,k}(\mathrm{PV}_0) + \frac{\sum_k p_{\mathrm{T}}^{trk,k}(\mathrm{PV}_0)}{k n_{trk}^{PU}}} \tag{4.4}$$

Here, k is a free parameter. It is chosen based the scaling of the pileup track contribution to jets with N_{PV}, but the performance of the JVT variable is insensitive to the exact value. Note that both of these variables rely on the correct identification of the primary vertex.

These two variables are combined into a two-dimensional likelihood based on a k-nearest-neighbor discriminant [73]. In the work presented here, jets are required to have JVT > 0.59, which corresponds to roughly a 95% efficiency for primary-vertex jets. Since JVT requires tracks associated the jet, it can only be used in cases where these are present. So, jets with $|\eta| > 2.4$ and central jets with no associated tracks do not have any JVT requirement applied. Additionally, jets with $p_{\mathrm{T}} > 60$ GeV do not have the JVT requirement applied, since pileup effects are considered to be small above this scale.

Identifying and removing jets outside of the tracker acceptance is much more difficult. For these, it is not possible to accurately identify the vertex from which a given jet originated. Nevertheless, some methods do exist to attempt to discriminate against these. One, known as "forward JVT" is based on identifying central pileup jets which balance well against a forward jet in p_{T} and ϕ. This method works because the overwhelming majority of pileup events are simple QCD dijets which balance in p_{T}. In this case, it is considered likely that both jets came from the same vertex, so the forward one can be removed. This algorithm was studied in the context of the VBF+$E_{\mathrm{T}}^{\mathrm{miss}}$ analysis, but its performance was found to be insufficient to make it worth using. Removal of any substantial number of pileup jets comes at too high a price in signal acceptance to improve the overall sensitivity.

4.1.4 Jet Cleaning

Jet cleaning is a method for removing objects which are reconstructed as jets but do not actually originate from a *pp* collision in the beamspot. This involves placing requirements on certain properties of the jet which reject these non-collision contributions while maintaining a high efficiency for real jets. Several cleaning variables are defined to discriminate these.

First, there are calorimeter signal quality variables. These are used to remove jets which were due to fluctuations or noise bursts in the calorimeters, which result in topo-clusters where there was no real energy deposit. This occurs almost entirely in the LAr calorimeter. Most calorimeter noise fluctuations/bursts are automatically masked by the data acquisition system, but some still remain and enter jet reconstruction. In order to determine whether a calorimeter signal came from a real energy deposit, a quality variable $Q_{\text{cell}}^{\text{LAr}}$ is defined for each cell. This is a χ^2-like quadratic deviation of the measured signal pulse shape from that which is expected for a real energy deposit To convert this into a jet-level variable, an energy-weighted average of $Q_{\text{cell}}^{\text{LAr}}$ is taken over all of the cells in the jet, denoted $\langle Q \rangle$. This is normalized such that $0 \leq \langle Q \rangle \leq 1$ (note that a small value of $\langle Q \rangle$ implies good pulse shape quality). Additionally, a "good LAr cell fraction" f_Q^{LAr} is defined. This is the fraction of LAr cells in the jet with a poor signal quality (i.e. with a value of $Q_{\text{cell}}^{\text{LAr}}$ which is too large). A very similar variable f_Q^{HEC} is defined as well; this is identical but calculated with respect to the HEC calorimeter instead. See [74] for the precise definitions of these variables and further details.

Next, there are energy ratio variables. Real jets tend to deposit their energy along a significant longitudinal distance (i.e. in several layers in both the EM and hadronic calorimeter layers). Non-collision jets tend to be much more localized, so we define shower shape variables to discriminate between the two. These have very simple definitions: the fraction of the jet's energy which was in the EM layers (f_{EM}), the fraction of the jet's energy which was in the HEC calorimeter (f_{HEC}), and the maximum fraction of the jet's energy which was in any single layer ($f_{\max}$). Jets with very small values of f_{EM} or very large values (i.e. close to 1) of f_{HEC} or $f_{\max}$ are much more likely to be from non-collision processes.

Finally, track variables can also be used to clean jets that are within the tracker acceptance. Here only one is used: the charged fraction f_{ch}. This is defined as the ratio of the p_{T} scalar sum of tracks associated with the jet to the total calibrated p_{T} of the jet. Note that this is actually the same as the variable R_{pT} which was used in the context of pileup suppression. Here it is used to discriminate against objects which have no tracks at all (or one, in the case of cosmic muons) rather than objects with tracks from the wrong vertex.

All 7 of these variables are then used to define a set of jet cleaning criteria[14]. If a jet fails any of these requirements, it is designated as a "bad jet". Here, events which contain *any* bad jets are completely removed from consideration. ATLAS defines two sets of criteria ("Loose" and "Tight"), where Tight has better background rejection but lower signal efficiency. The analysis presented in this dissertation uses the Tight working point in all cases. This is because the final state is fully hadronic and therefore particularly sensitive to these non-collision jets.

---

[14] See [74] for the full list of these criteria.

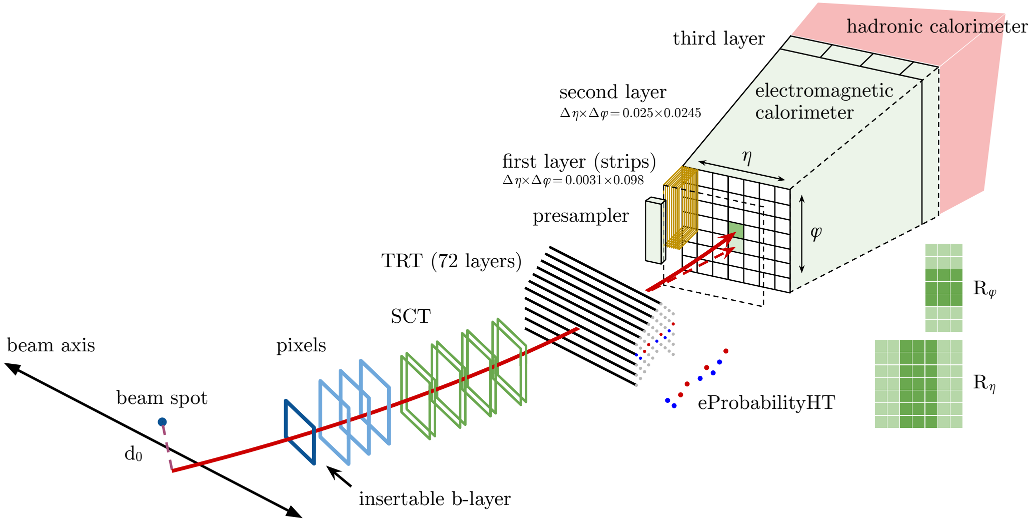

Figure 4.2: Schematic view of electron reconstruction and identification [75].

4.2 Electrons

Electrons, along with muons, are of particular interest at hadron colliders since they can only be produced promptly via W and Z bosons (off-shell photons can also decay into e^+e^-, but this case is not relevant here). The vast majority of the background in this search is from W and Z production, so events with identified leptons are used to constrain it[15]. Furthermore, leptonic W events enter the signal region as a background when the lepton is not successfully reconstructed. So, efficient identification of electrons and muons identification is essential to the sensitivity of the analysis.

Electrons are reconstructed from a combination of calorimeter and Inner Detector information. A typical electron traverses the entire Inner Detector, leaving a track in the form of hits in the individual layers. In the TRT, many of these are "high-threshold" hits[16], which distinguishes electrons from heavier charged particles. Then, the electron deposits its remaining energy entirely in the EM layers of the calorimeter. This tends to be a very narrow shower which does not penetrate into the hadronic layers. This shape distinguishes electrons from jet or individual charged hadrons, which have a much wider and deeper shower.

To construct electron candidates, a sliding window algorithm scans over the EM layers of the calorimeters to find localized energy deposits [76]. These are then called "calorimeter clusters". If a cluster is located very close to an Inner Detector track originating from the primary vertex, that

---

[15] Chapter 6 gives an overview of how these "control samples" are defined and used.

[16] See Section 3.2.1.3 for a description of TRT hits.

object is then considered for electron identification. The tracks themselves are reconstructed by starting with a seed of 3 Inner Detector hits, and then running a Kalman filter algorithm [77] to get an estimate of the full trajectory[17]. After being associated with a cluster, the track is re-fitted to account for energy losses in the tracker and additional corrections are applied to calibrate the electron's energy. The exact procedure is described fully in [79]. Due to the necessity of tracking information, electrons are only reconstructed for $|\eta| < 2.47$. The region with $1.37 < |\eta| < 1.52$ is also excluded, as this is the transition region between detector components and is not fully instrumented.

Once the electron candidate has been established, it is passed through an identification algorithm in order to remove backgrounds. These backgrounds are primarily from charged hadrons (typically associated with jets) and photon conversions ($\gamma \to e^+e^-$ via interactions with detector material). Non-prompt electrons can also arise from semileptonic decays of hadrons containing b and c quarks. All of these backgrounds are reducible, in the sense that the tracks and clusters tend to have properties different from those of real prompt electrons. For example, tracks originating from heavy quark decays tend to have a larger impact parameter with respect to the primary vertex (their comparatively long lifetimes means they often travel measureable distances before decaying). The electron identification algorithm uses a set of variables describing the properties (shape, quality, etc.) of the track the and calorimeter cluster.

The algorithm itself uses a likelihood-based discriminant constructed from the full set of input variables. This is defined by constructing probability distribution functions (PDFs) for each variable, for each hypothesis (i.e. signal or background). These PDFs are determined from either simulation or a very pure sample in data[18] of real electrons (or fake ones, for the background hypothesis). A likelihood ratio between the hypotheses is constructed from the input variables, and this is used as the discriminant. Several working points for the identification ("Loose", "Medium", and "Tight") are defined by choosing different values of the discriminant to cut on. These respectively correspond to increasing background rejection, but lower signal efficiency. For the analysis presented here, the Tight working point is used, and only electrons with $p_{\mathrm{T}} > 7$ GeV are considered. After identification, additional isolation requirement are placed on the electron. This essentially requires that there are no other significant energy deposits or tracks too close to the electron. This is intended to remove real electrons that were part of a shower or decay (that is, non-prompt electrons), as well as electron candidates arising from random energy deposits in jets.

---

[17] See [78] for further details on how this is carried out.

[18] Whether data or simulation is used depends on the particular version of the identification algorithms being used. For the analysis presented in this thesis, simulation is used.

4.3 Muons

Due to lepton universality, muons are always produced exactly the same as electrons at the LHC (up to the mass difference, which is generally negligible in the case of prompt lepton production[19]).

As such, they are used for the same purposes, particularly as a handle on processes involving weak bosons. However, due to their larger mass, muon leave a very different signature in the detector. Where electrons deposit all of their energy in the calorimeters, muons traverse the entire detector and escape. Because of this, muons are reconstructed exclusively from tracking information. The calorimeter deposits are generally small, but can be used to assist with identification in some algorithms.

Tracks from both the Inner Detector and the Muon Spectrometer are used in muon reconstruction. The ID tracks are exactly the same as those used in electron reconstruction (Section 4.2). MS tracks are built by first constructing “segments” from hits in the individual muon subdetectors (MDT, CSC, etc; see Section 3.2.3). A combinatorial fit is then run on these segments to get full track candidates. Finally, the full set of individual hits is used in a χ^2 fit against the path of the track candidate. If the resulting χ^2 value is small enough to satisfy the selection criteria, the candidate is accepted as a MS track.

After the tracking is performed, there are several algorithms for the final muon reconstruction. The definition used here is the “combined muon” algorithm. This requires an ID track to coincide with a MS track, and then performs a global re-fit on the full set of hits from both subdetectors. During this procedure, MS hits may be added or removed in order to improve the fit quality. A more detailed description of this and other muon reconstruction algorithms at ATLAS is given in [80].

After the muon objects have been reconstructed, identification requirements are imposed to reduce backgrounds (analogous to those used for electrons). Most of the background comes from non-prompt muons originating from hadron decays (particularly pions and kaons). These background muons have a discontinuity in their track direction at the point of decay, resulting in a lower-quality fit. The actual variables used in muon identification consist of p_{T} and q/p differences between the ID and MS tracks, as well as the χ^2 of the combined fit. As with electrons, three working points (“Loose”, “Medium”, and “Tight”) are defined with varying signal efficiency and background rejection[20]. This work uses the Medium working point in all contexts. As with electrons, isolation

---

[19] The mass difference does play an important role in some cases, such as photon conversions and meson decays. However, none of these cases are relevant in the context of this dissertation.

[20] See [80] for the full definition of each working point.

requirements are imposed to reject muons that were produced in jets very close to the interaction point (largely due to heavy flavor decays). This effectively requires that there not be a significant amount of track activity in a cone around the muon.

4.4 Missing Transverse Momentum

Not all particles can be directly reconstructed with ATLAS. Neutrinos, as well as any dark matter candidate, do not interact with the detector via electromagnetic or strong interactions. So, their presence can only be inferred from an imbalance of transverse momentum in the event (denoted $E_{\mathrm{T}}^{\mathrm{miss}}$, sometimes also abbreviated as "MET"). As this is a search for the production of invisible particles, the reconstruction of this object is crucially important.

At the simplest level, the $E_{\mathrm{T}}^{\mathrm{miss}}$ in a given event is defined as the negative vector sum of all reconstructed objects in the event. Here, this includes jets above 25 GeV, electrons above 7 GeV, and muons above 7 GeV. However, there is often additional soft activity in the event (such as objects below the reconstruction threshold or soft radiation). For this reason, a "soft term" is also included in the $E_{\mathrm{T}}^{\mathrm{miss}}$ definition. The soft term can be calculated either from calorimeter information (the Cluster Soft Term, CST) or from ID tracks (the Track Soft Term, TST). Under LHC Run 2 conditions, the large quantity of pileup means the track-based calculation generally performs better. For the analysis presented here, the performance is nearly identical, so the TST is used simply because it is better-supported by existing software. The full definition is then:

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{jets} \vec{p}_T^{\,jet} - \sum_{electrons} \vec{p}_T^{\,e} - \sum_{muons} \vec{p}_T^{\,\mu} - \sum_{tracks} \vec{p}_T^{\,tk} \tag{4.5}$$

where the sum over tracks includes only tracks originating from the primary vertex which are not part of another reconstructed object. The uncertainty on this number is computed by propagating the individual uncertainties on each object through the calculation. Note that this definition does not include a terms for photons; this is because photons are not explicitly identified. They are either reconstructed as jets or simply ignored if their p_{T} is too soft. The impact of the ignored photons is negligible for this analysis.

Note that it is possible to reconstruct a significant value of $E_{\mathrm{T}}^{\mathrm{miss}}$ in events that did not contain any real invisible particles. Most of the time, this happens because the p_{T} of one of the objects going into the sum was mismeasured due to detector effects. If a real jet or lepton fails reconstruction/identification criteria, it is still accounted for by the soft terms (except for the neutral component of jets). “Fake” $E_{\mathrm{T}}^{\mathrm{miss}}$ can also come about as a result of pileup. The JVT algorithm[21] attempts to remove pileup jets from consideration, and those which are identified as such are excluded from the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. However, if a pileup jet has $|\eta| > 2.5$, it is outside the tracker volume and will not be rejected. So, if a pileup vertex produces two jets (which balance each other out) and only one is removed by JVT, a large amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ is artificially introduced. These effects need to be carefully handled when evaluating systematic uncertainties.

For the VBF+$E_{\mathrm{T}}^{\mathrm{miss}}$ analysis in particular, correct identification of the primary vertex can be a relevant issue. Events in the signal regions[22] contain two jets, at least one of which is always outside tracking acceptance, and very little else. In about one third of signal-like events, both tagged jets are outside the tracker. As a result, they leave no tracks associated with the “true” primary vertex, and it is possible for a different one to be chosen incorrectly. This interferes with the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ soft term when using the TST definition. The CST definition is not affected by vertexing inefficiency, but suffers from worse resolution under higher pileup conditions. So, a series of checks were performed to ensure that the TST definition can safely be used in this case.

The validation checks consisted of several components. First, the TST and CST $E_{\mathrm{T}}^{\mathrm{miss}}$ were compared in signal MC (Figure 4.3). This was done in the inclusive SR selection, except for the third jet veto. The two definitions give equally good agreement with truth-level $E_{\mathrm{T}}^{\mathrm{miss}}$, indicating that the calculation of the soft term is not important for the signal. This also implies the same for the VBF-produced W and Z background, since the jet kinematics are very similar. The same checks were also carried out for the strong-produced W and Z background MC. Example plots are shown in Figure 4.4. Again, the agreement between the two methods is good, indicating that the soft term makes little difference here.

The missing transverse momentum can also be defined without including the soft term. In this case it is denoted $H_{\mathrm{T}}^{\mathrm{miss}}$ (or MHT), since it is the vector sum of “hard” objects. Since the soft term is similar between signal and W/Z background, the option of using $H_{\mathrm{T}}^{\mathrm{miss}}$ instead of $E_{\mathrm{T}}^{\mathrm{miss}}$ in the analysis selection was considered. This would substantially simplify the calculation and associated systematic uncertainties, as well as allowing a perfectly equal definition across all regions. However, it was found that the soft term is actually quite different for the multijet background, and omitting it would sacrifice a large amount of discrimination power against it.

The $H_{\mathrm{T}}^{\mathrm{miss}}$ variable is kept in the analysis selection in addition to $E_{\mathrm{T}}^{\mathrm{miss}}$. In this case, it includes

---

[21] See Section 4.1.3 for a description of JVT.

[22] The full definition of the analysis selection is detailed in Section 6.2.

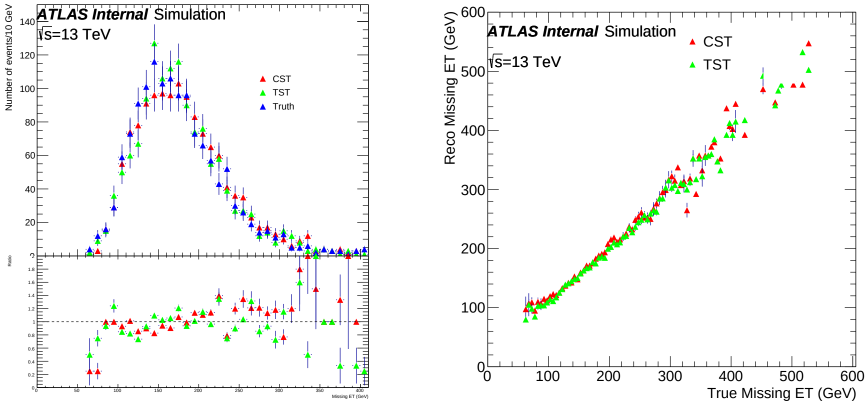

Figure 4.3: **Left:** Comparison of CST and TST $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in signal MC. **Right:** Event-by-event comparison of CST and TST $E_{\mathrm{T}}^{\mathrm{miss}}$ in signal MC.

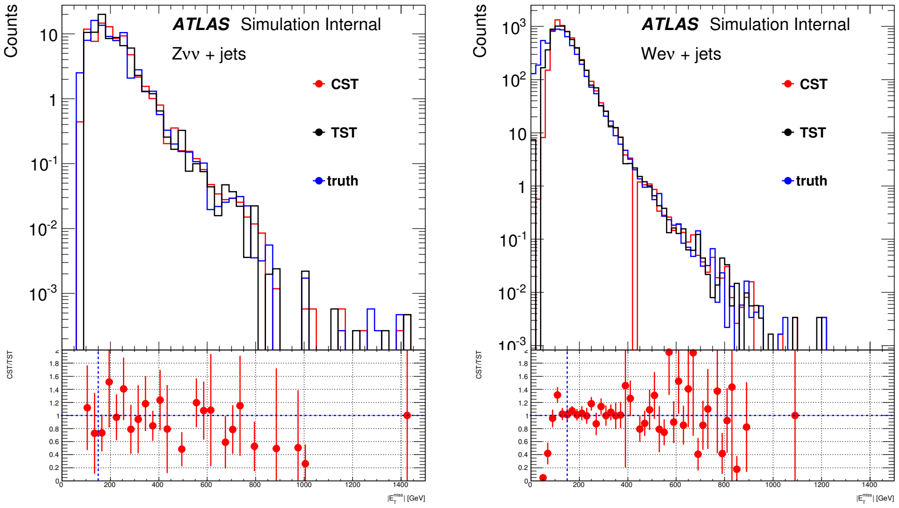

Figure 4.4: Example comparisons of CST and TST $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in background MC. **Left:** $Z \to \nu\nu$ in the signal region (except the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut). **Right:** $W \to e\nu$ after the SR jet p_{T}, $\Delta\eta_{jj}$, and $\Delta\phi_{jj}$ cuts.

the vector sum of *all* jets, even those which failed the JVT requirement[23]. This is used to reject multijet background that enters the signal region due to primary vertex misidentification. In these cases, the JVT requirement would otherwise remove jets from the true primary vertex with $|\eta| < 2.5$ and $p_{\mathrm{T}} < 60$ GeV. This can introduce reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ in events that contained no invisible objects.

4.5 Overlap Removal

In some cases, ambiguities can arise when multiple objects are identified at the same location in (η, ϕ). For example, high-p_{T} electrons are nearly always identified as jets, and jets from b-hadron decays often contain muons which pass the muon identification. In order to avoid double-counting, an overlap removal procedure is applied. This proceeds in the following order:

- If a jet and an electron are $\Delta R < 0.2$ apart, the jet is discarded. This is because electrons are almost always also identified as jets, and the electron ID criteria are much stricter, so this case is very likely to be a real electron rather than a hadronic object.
- If a remaining jet (i.e. one not removed by the previous step) and an electron are $\Delta R < 0.4$ apart, the electron is discarded. If a real prompt electron is identified as a jet, the two reconstructed objects are almost always very close in ΔR. If this is not the case, it is most likely that a hadronic object passed the electron ID, or a real electron was produced as part of the shower.
- If a jet has fewer than three tracks and is $\Delta R < 0.4$ from a muon, the jet is discarded. In this case, the most likely scenario is that the muon is the only actual particle involved and it left energy deposits in the calorimeter, from which the jet object was reconstructed. This should be treated like a muon, not a hadronic object.
- If a jet has more than three tracks and is $\Delta R < 0.4$ from a muon, the muon is discarded. In this case, it is very likely that the muon was produced as part of the showering from a real hadronic objects, and should not be treated as prompt.

---

[23] See Section 4.1.3 for a description of the JVT discriminant, which is intended to remove pileup jets.

4.6 Trigger-Level Objects

The ATLAS trigger system[24] needs to have its own dedicated algorithms for reconstructing the same objects which are used offline in the analysis. This is due to the very limited time, resources, and information that are available to the trigger. As a result, the reconstruction algorithms are substantially different from those used offline, particularly at Level-1.

4.6.1 Level-1 Objects

Jets are reconstructed at Level-1 from jet elements, which are typically 0.2×0.2 in (η, ϕ)[25]. These are arranged in a grid covering the entire area of the detector, and a sliding-window algorithm is used to locate the jet candidates [81]. This scans over all possible 0.8×0.8 "windows", and sums together the energy of all towers in each. Windows with a total energy over some threshold (typically 10 GeV) are considered to be jet candidates. Since this method will frequently identify a single jet many times in neighboring windows, a local-maximum criterion is also imposed. This requires that the "core" of the window (usually a 0.4×0.4 square at its center) contain more energy than any other nearby core. Note that this introduces inefficiency in situations where multiple jets are close to each other. This algorithm is also quite sensitive to pileup, so noise thresholds are placed on each tower before the sliding window is executed. Each tower's energy is only considered if it is above some threshold value (which depends on the pileup conditions and which part of the detector the tower is located in). No energy calibration is applied to L1 jets; as a result they typically have a substantially smaller energy scale than the offline jets they correspond to.

Electrons (and photons, as they are indistinguishable at L1) use a very similar algorithm to jets. However, this one is provided with finer-granularity (0.1×0.1) trigger towers. The "window" here is a 4×4 square of towers [81]. However, the energy is only searched for in 2×1 or 1×2 pairs of trigger towers, since the shower from EM objects is so narrow. For each window, only the EM layers of the central 4 towers are considered. Whichever 1×2 or 2×1 pair of these central 4 has the highest energy is the electron candidate if it passes the p_{T} threshold. Since electrons and photons leave very localized energy deposits in only the EM layers of the calorimeter (and there is a very large jet background), there is also an isolation requirement imposed. This demands that the sum of the energy in the 12 outer towers of the window be below some programmable threshold. The

---

[24] See Section 3.2.4 for an overview of the ATLAS trigger system.

[25] An overview of the Level-1 system is given in Section 3.2.4.1.

energy in the hadronic layers is also required to be below a threshold, to further reject hadronic backgrounds.

Muon candidates are constructed at L1 from hits in the RPCs and TGCs [66]. Coincidences between layers are generated separately in η and ϕ, then combined using programmable logic. In order to form a coincidence, a set of hits must lie within a parametrized envelope for the muon's trajectory (this is a function of the muon's charge and p_{T}). The p_{T} of the muon candidate is determined by the degree of deviation of the hit pattern from a straight trajectory [82].

$E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated using a very simple algorithm at L1. It is defined as the vector sum of all the individual tower E_T values. The same noise thresholds used in the L1 jet trigger are applied here, to reduce pileup sensitivity.

4.6.2 High-Level Trigger Objects

Jets are reconstructed in the HLT using full-granularity calorimeter information. A similar method is used to the offline reconstruction: cells are grouped together using topological clustering, and the anti-k_t algorithm is run over these. The main difference between HLT jet and offline jets is the simpler calibration procedure. In the HLT, a pileup correction is applied by subtracting off the product of the jet area and an estimated event-level pileup density (calculated from the region $|\eta| < 2$). Then, a response correction is applied as a function of η and p_{T}. At this point the jet is considered to be calibrated, and no tracking information is used.

HLT electrons are reconstructed using both calorimeter and ID tracking information. This is done in several steps, to optimize the use of computing resources [83]. First, calorimeter clusters are constructed using a sliding window algorithm within RoIs determined by the L1 trigger. This window uses the second EM layer of the calorimeter (where electrons deposit most of their energy), and is a 3×7 rectangle in the barrel, and a 5×5 square in the end-cap. If a "fast-reconstructed" ID track with $p_{\mathrm{T}} > 1$ GeV matches this cluster, it becomes an electron candidate. Next, a more precise step is carried out. The cluster energy is calibrated and fully-reconstructed tracks are now used. A likelihood-based identification criterion is imposed, very similar to the one used offline (the same variables are used, except the energy loss due to bremsstrahlung). An isolation requirement can then be made, which is based on the total track p_{T} within a cone around the electron (this is optional; only some trigger items use it).

HLT muons are reconstructed using the full set of hits from the entire muon spectrometer. As with offline reconstruction, the algorithm is based on associating an ID track with a MS track. There are two methods for doing this [82]. The first is the “RoI-based” method, which begins by searching for MS tracks only in regions of interest (RoIs) where the L1 trigger found a muon candidate. These are matched with ID tracks to form a combined HLT muon. The second method is the “full-scan” method, which is very similar but scans over the entire detector rather than being restricted to RoIs. This can reconstruct muons which the L1 trigger may have missed, but it is substantially more resource-intensive than using only the RoIs. The HLT also has the ability to apply cone-based isolation requirements to the muons.

$E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction at HLT presents a challenge under current LHC conditions, as the trigger rates are extremely sensitive to pileup. There are several algorithms, all of which are based purely on calorimeter information. The simplest is simply a negative vector sum of all calorimeter cells (analogous to what L1 uses), however this is rarely used in practice due its poor pileup scaling. Another similar algorithm uses the negative vector sum of all topo-clusters. Several variations in this attempt to suppress pileup further by applying various subtraction and thresholding techniques[26].

Alternately, a $H_{\mathrm{T}}^{\mathrm{miss}}$ algorithm can be used, which adds together only the reconstructed (and calibrated) jet objects. For different parts of 2015 and 2016 data-taking, the primary $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger was the topo-cluster-based algorithm or a combination of that with $H_{\mathrm{T}}^{\mathrm{miss}}$.

---

[26] See [82] for further details.

Chapter 5

Hadronic Trigger Upgrades: Algorithms and Simulation

This chapter discusses a set of studies and simulation work on the topic of hadronic trigger algorithms, specifically intended for future upgrades of the ATLAS detector. The material presented here is self-contained and is not directly related to the remainder of this dissertation. However, the techniques studied here could potentially become useful for triggering on a variety of VBF event topologies in the future, when extreme pileup conditions will make current methods less effective.

First, a general overview on ATLAS Level-1 trigger upgrades is given, and then a variety of related studies is presented. They are grouped into three main categories: general performance studies for jet/$E_{\mathrm{T}}^{\mathrm{miss}}$ triggers, algorithms for forward jets, and techniques for mitigating the effects of the pileup which will be present in very large quantities.

5.1 Introduction to Level-1 Trigger Upgrades

The ATLAS detector is planned to undergo two major sets of upgrades in the future: Phase-I (prior to Run 3, which is scheduled to start in 2021), and Phase-II (prior to Run 4, which is scheduled to start in 2026 and coincide with a a major luminosity upgrade to the LHC). A large part of this upgrade program consists of improvements to the Trigger and Data Acquisition (TDAQ) system, especially in the Level-1 trigger[27].

The first aspect of the Phase-I upgrade is a replacement of the readout systems for the LAr calorimeters[28]. Rather than reading out calorimeter information at the level of trigger towers, the upgraded system will provide new objects called “supercells”. These are groups of individual

---

[27] For an overview of the Level-1 trigger system and the algorithms it uses, see Sections 3.2.4.1 and 4.6.1.

[28] The LAr calorimeters and their layout are described in Section 3.2.2.1.

calorimeter cells which are analog-summed and then digitized, such that each supercell has an energy value for every bunch crossing. In the EMB, there are 10 supercells per 0.1×0.1 square in (η, ϕ): one for the presampler, four for the front layer, four for the middle layer, and one for the back layer. The front and middle layers are segmented in η (that is, each supercell is 0.025×0.1 in (η, ϕ)). The primary design reason behind this is to improve the performance of electromagnetic object identification in the L1 trigger, by providing finer-granularity calorimeter readout including depth information.

These supercells will then be sent to three "feature extractors" (or FEXes), boards containing FPGAs which run the actual trigger algorithms. So, these FEXes will replace the current Cluster Processor and Jet Energy Processor (though these will also be kept in parallel, but not used for primary triggers). The Tile Calorimeter will not be undergoing readout upgrades for Phase-I, so it will continue to read out trigger towers through the existing system, and those towers will be sent to the FEXes along with the supercells. The FEXes identify trigger-level objects (or TOBs) such as jets or electrons and send these to the L1 Central Trigger Processor for a final L1 decision to be made. Along the way, the objects are also passed through "L1Topo", which can make topological requirements on objects in the event. For example, invariant mass or $\Delta\phi$ cuts can be applied. This overall architecture is shown in Figure 5.1

There are three FEXes, each of which serves a different specialized purpose. The Electromagnetic Feature Extractor (eFEX) identifies electrons and photons, and also provides capability for hadronic

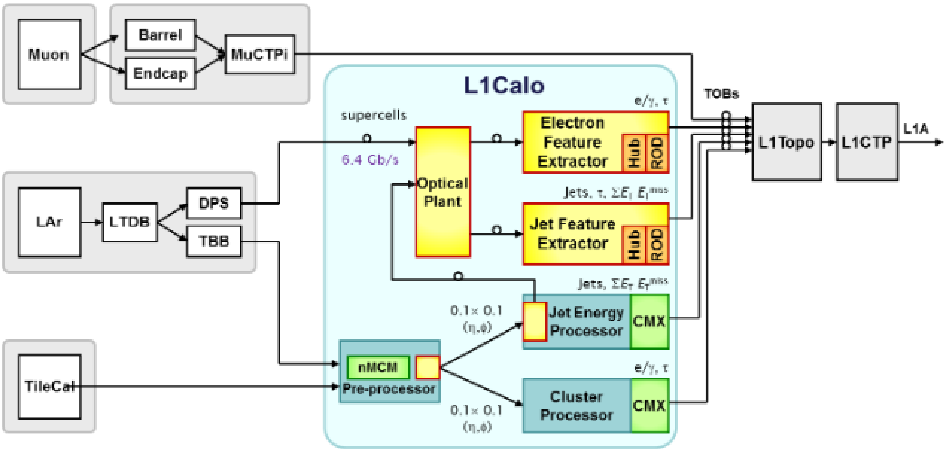

Figure 5.1: Schematic diagram of the planned Phase-I upgrade to the L1 trigger system. Items shown in yellow are new; those in blue/green already exist in the current system. The gFEX is not shown here but is also present alongside the jFEX.

taus. The Jet Feature Extractor (jFEX) identifies jets and calculates energy sums. The Global Feature Extractor (gFEX) is primarily intended for large-R jets characteristic of boosted objects and global event-level quantities. The gFEX is different in that it covers the whole detector with a single module. By contrast, the jFEX and eFEX each break the detector down into regions covered by several different modules. As a result, the algorithms they use must be local (i.e. they cannot depend on what's happening in the rest of the detector). Each jFEX module, for example, covers a ϕ ring of the detector using 4 FPGAs. So, each FPGA only has access to the information from some local region of the detector (see Figure 5.2 for an example). The rest of this chapter will focus primarily on the context of jFEX.

The baseline algorithm for jet-finding in jFEX is a sliding window which is qualitatively similar to that of the current L1 system, but with some notable differences. Before being input to the algorithm, the supercells are summed into "jTowers", which are 0.1×0.1 in the central region of the detector ($|\eta| < 2.5$)[29]. The energies for the electromagnetic and hadronic layers of each jTower are provided separately. A sliding window (nominally 0.9×0.9) runs over these jTowers and searches

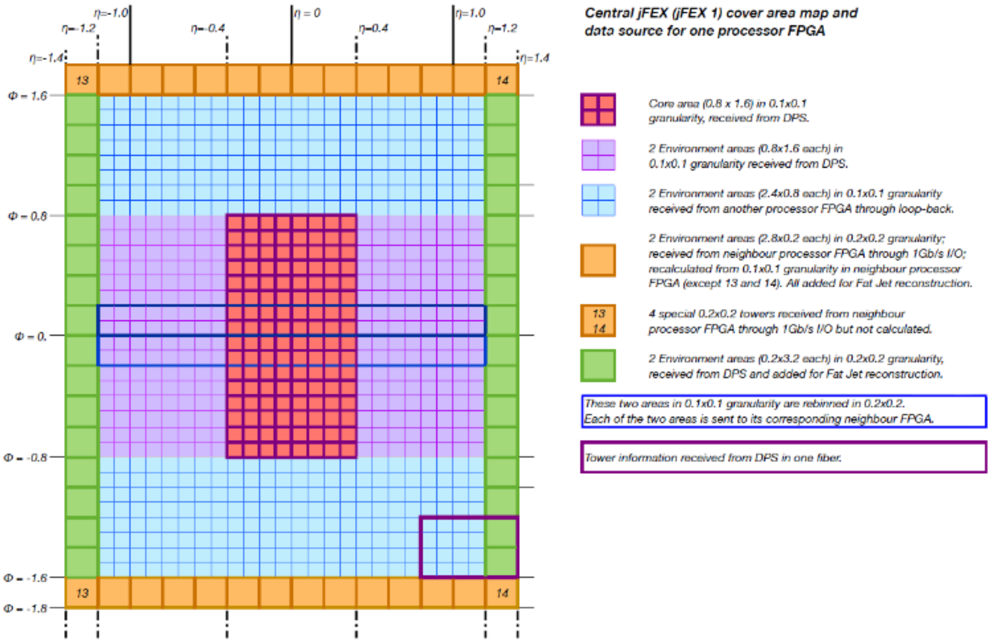

Figure 5.2: η-ϕ coverage map for the central jFEX module. The area covered by one individual FPGA is also shown. "DPS" refers to the upgraded LAr calorimeter readout system.

---

[29] For $2.5 < |\eta| < 3.1$, the jTowers are 0.2×0.2. They are irregular for $|\eta| > 3.1$, due to the geometry of the FCAL (see Section 5.3).

for localized clusters of energy. The exact definition is programmable and the version which will be used is still being decided at the time of writing; possible modifications include a more circular window shape and Gaussian weighting based on the distance of a jTower from the window center. In all cases; a local maximum requirement is imposed on the central towers in the window (again, the exact definition is programmable) in order to avoid double-counting. As a result of this some inefficiency is introduced for jets which are very close to each other in the case of multi-jet triggers. However, if there is sufficient energy, the event can still be triggered by a lower-multiplicity jFEX jet trigger or by gFEX. gFEX is specialized for merged "fat jets" with a larger radius parameter, so if two jets are close enough to introduce inefficiency in jFEX, then gFEX would reconstruct them as one large jet.

Before running the actual sliding window algorithm, noise thresholds are applied to the jTowers (so any jTower with energy below the threshold is discarded). These thresholds serve to suppress electronic noise and stochastic pileup effects, where random soft activity from several different interactions coincides in a window. The thresholds are chosen[30] based on the expected amount of pileup, and they can in general (and should) be different for different towers. These can be applied on the EM energy, hadronic energy, or the sum of the two.

For the ATLAS Phase-II upgrades, the same system using the various FEXes will remain in place. The main difference from this perspective is that the Tile calorimeter readout system will be replaced with something more akin to the supercells used by the LAr calorimeters. The exact parameters of this are not yet determined, so when modeling this system for Phase-II performance studies, offline calorimeter cells are used to provide the Tile information as an approximation.

The remainder of this chapter presents a series of performance studies for jFEX. All of these are based on a simulation I created specifically for the purpose; this simulation has now become the ATLAS standard one and is also used more widely for upgrade performance work. In all of the results presented here, a 0.8×0.8 sliding window was used, with a requirement that the central 0.4×0.4 block have more energy than any of its nearest or next-to-nearest neighbors. Noise thresholds were determined from a significance-based scheme, using the energy distributions for every individual jTower derived from an inclusive minimum-bias[31] MC sample.

---

[30] The exact definition can be freely chosen, but they are typically either significance-based or occupancy-based and determined from simulated jTower energy distributions.

[31] "Minimum-bias" means the MC sample is intended to be representative of an *inclusive* random sample of *pp* collisions in ATLAS. That is, there are no requirements on any particular process or particle production.

5.2 Performance for Phase-I and Phase-II

The performance of trigger algorithms can be quantified in terms of several different variables. One of the most important is the **rate**, which is simply the frequency in time at which the trigger accepts events. Loosening trigger requirements will result in a more inclusive sample being collected, but will also increase the rate. Due to technical restrictions[32], there are limitations on the overall rate of the trigger system. For example, only 100 kHz of events are allowed to pass L1 (this will be increased to 1 MHz with the Phase-II upgrades). As a result, some amount of rate is typically budgeted for each particular trigger item, and this acts as a constraint on the object thresholds used. A typical trigger rate, for a single jet trigger at $\mu = 200$, is shown in Figure 5.3.

The **efficiency** of a trigger is equally important. This is the fraction of signal events[33] which are successfully recorded (so a value of 100% is ideal). Efficiency is generally described as a function of the offline p_{T} of the object of interest, or of a variable such as $E_{\mathrm{T}}^{\mathrm{miss}}$. A typical trigger efficiency has the shape shown in Figure 5.4. The width of the "turn-on" corresponds to how closely the trigger is able to replicate the algorithms used for offline reconstruction. For a hypothetical perfect trigger which implements exactly the offline algorithm, this would be a step function. In cases where no

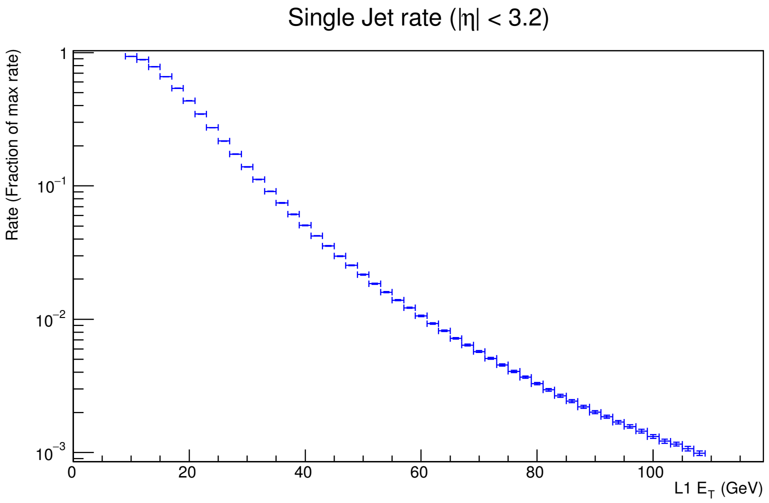

Figure 5.3: Simulated single-jet ($|\eta| < 3.2$) trigger rates for jFEX at $\mu = 200$, as a function of the trigger threshold. Here, the y-axis is shown as a fraction of the total filled beam crossing rate (which depends on the filling scheme, but is typically ~ 30 MHz at nominal luminosities).

---

[32] Specifically, at L1 the limitation is the read-out bandwidth of the detector front-end electronics.

[33] "Signal events" can be anything; the efficiency of a trigger depends on the signal being considered.

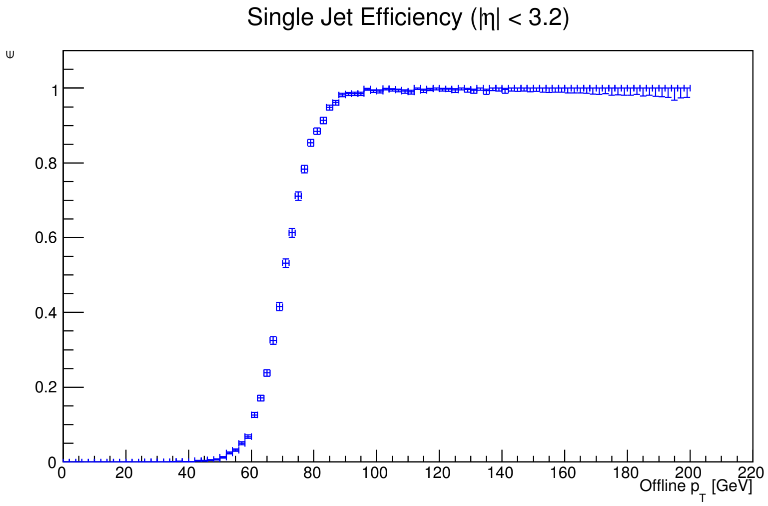

Figure 5.4: Simulated single-jet ($|\eta| < 3.2$) trigger efficiency for jFEX at $\mu = 200$, for a trigger threshold of 70 GeV. This used $hh \to 4b$ MC simulation as the signal, and only considers isolated jets.

reliable description of offline jet reconstruction and calibration is available (such as early studies for Phase-II), the trigger efficiency can be measured with respect to truth-level jets[34] from MC instead.

When evaluating trigger performance, it is necessary to consider both the trigger rate and efficiency. Varying the trigger threshold will change both of these numbers, so a given algorithm can be thought of as a curve in the space of these two variables. For a fixed rate budget, the algorithm which provides the highest efficiency is considered to be the best. Here, the "95% efficiency point" at a given rate is used as a metric to summarize trigger performance and compare between algorithms. This is defined as the offline jet p_{T} for which the trigger is 95% efficient[35]. The results of this method for single-jet triggers in a typical Phase-II scenario are shown in Figure 5.5. Here, the efficiency measurements are derived from MC simulation of $hh \to 4b$ as the signal, and the jets are required to be isolated from other jets to be considered. This isolation requirement removes a large part of the dependence on the choice of signal. The rate measurements are taken from minimum-bias MC simulation, in order to get a representative description of all pp collisions.

---

[34] Truth-level jets in this context are defined by running the anti-k_t algorithm on all visible truth-level final-state particles from the generator.

[35] The choice of 95% is somewhat arbitrary, but is motivated by the guideline that this is roughly the regime above which most analyses prefer to work. For comparisons between similar algorithms, this choice does not matter.

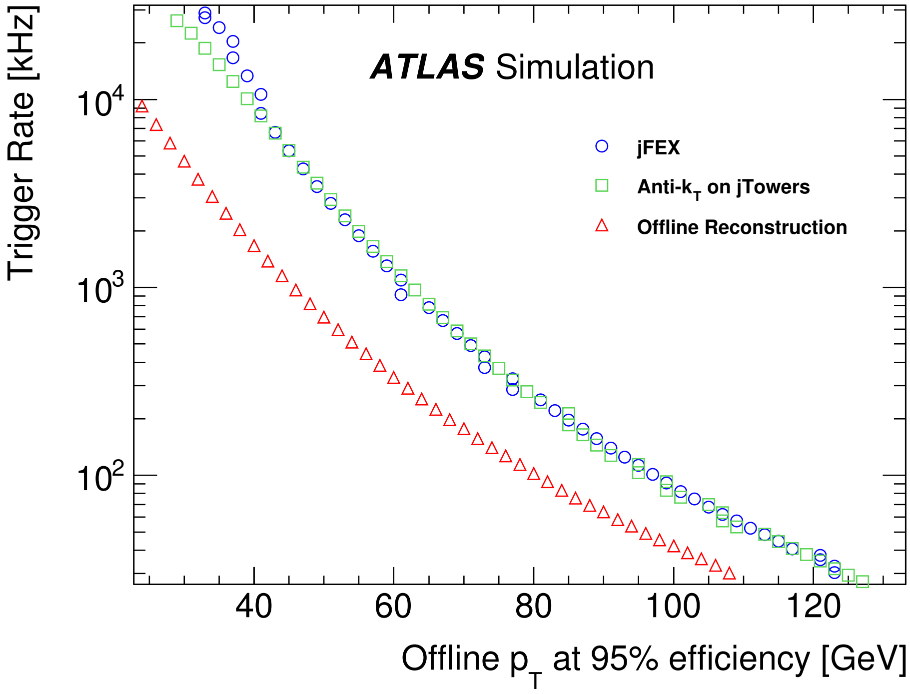

Figure 5.5: Simulated performance summary for single-jet ($|\eta| < 3.2$) triggers with jFEX at $\mu = 200$. Anti-k_t executed on the individual supercells (not possible in practice due to hardware limitations) is also shown. The rates for offline jets are provided as a reference (the "efficiency" for these jets is simply a step function).

Multi-jet triggers can be somewhat more complicated, since the local maximum requirement imposed by the jFEX algorithm introduces inefficiencies when two signal jets are close together. An example of this effect is shown in Figure 5.6. This shows an event with 4 real jets, where two are near each other. The anti-k_t algorithm is successfully able to resolve these into two separate objects, but the trigger algorithm only recognizes it as one. As a result, this event would not pass a 4-jet trigger at Level 1. Additionally, in this example, there is an isolated jet identified by the trigger which does not correspond to a truth-level jet above a 15 GeV threshold. This is most likely due to a noise or pileup fluctuation (and is too soft to pass a realistic trigger threshold anyway). For multijet triggers in general, the efficiency can be measured at jet-level or at event-level, with respect to the lowest offline jet p_T for example.

In order to take a deeper look at the performance of the jet trigger algorithm, the energy scale and resolution can also be examined. This can be thought of in terms of the distribution of the ratio of trigger-level jet p_T to offline (or truth-level) p_T. The average value of this distribution represents

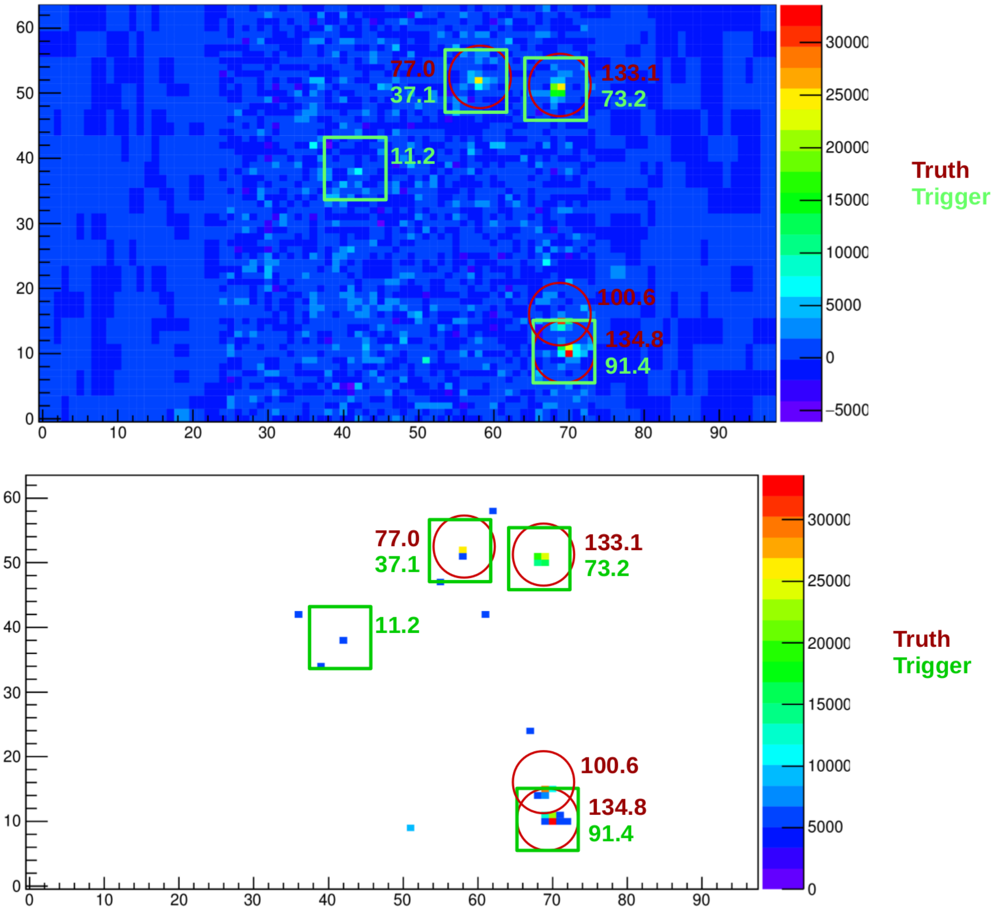

Figure 5.6: Example “event display” showing all jTowers for an event where a jet is lost by jFEX due to proximity to another jet. The x- and y-axes are the tower indices in η and ϕ respectively. This is shown without **(top)** and with **(bottom)** noise thresholds applied. Red circles indicate the truth-level jets, and green squares indicate the jets found by jFEX, along with the corresponding p_{T} in GeV. This event was taken from MC simulation of $hh \to 4b$ at $\mu = 200$.

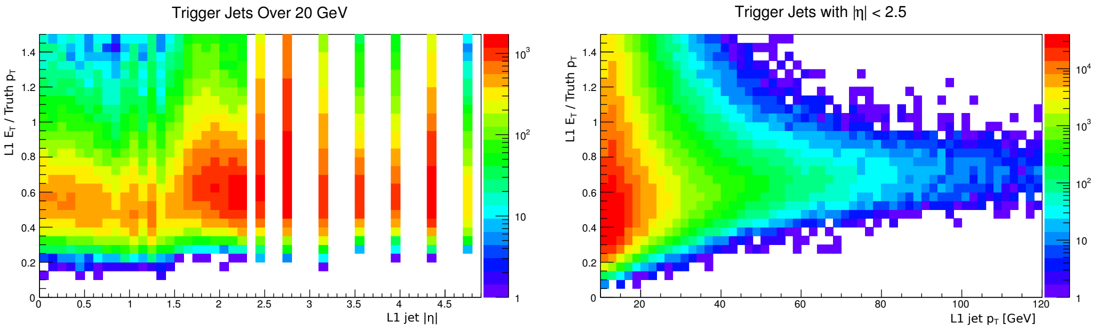

Figure 5.7: jFEX jet response for minimum-bias MC simulation with $\mu = 200$. This is shown as a function of η **(left)** and the L1 E_{T} **(right)**, with the indicated cuts. Points are sparse at high η due to the coarse supercell granularity.

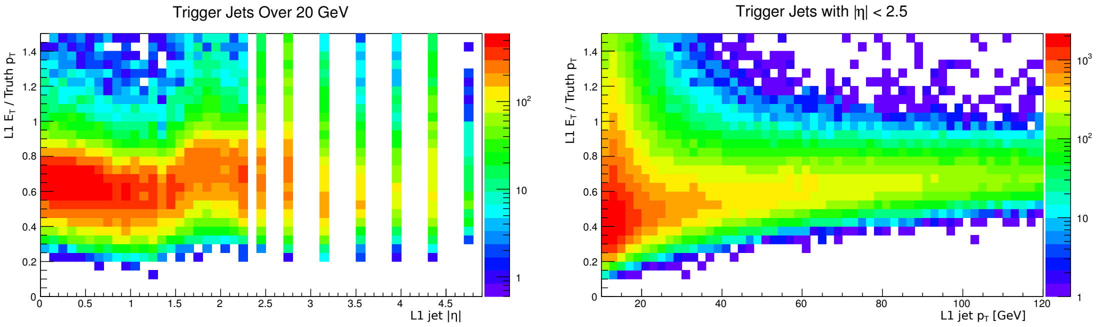

Figure 5.8: jFEX jet response for $hh \to 4b$ signal MC simulation with $\mu = 200$. This is shown as a function of η **(left)** and the L1 E_{T} **(right)**, with the indicated cuts. Points are sparse at high η due to the coarse supercell granularity.

the L1 jet energy scale, and its width represents the resolution. Both of these quantities can in general depend on the p_{T} of the jet and where exactly it is located within the detector, due to the varying response properties of the different calorimeters. These distributions are shown as a function of the jet p_{T} and η in Figures 5.7 and 5.8, for minimum-bias MC and $hh \to 4b$ signal simulation respectively. Note that there are discontinuities in the energy scale at η values corresponding to transition regions between calorimeters.

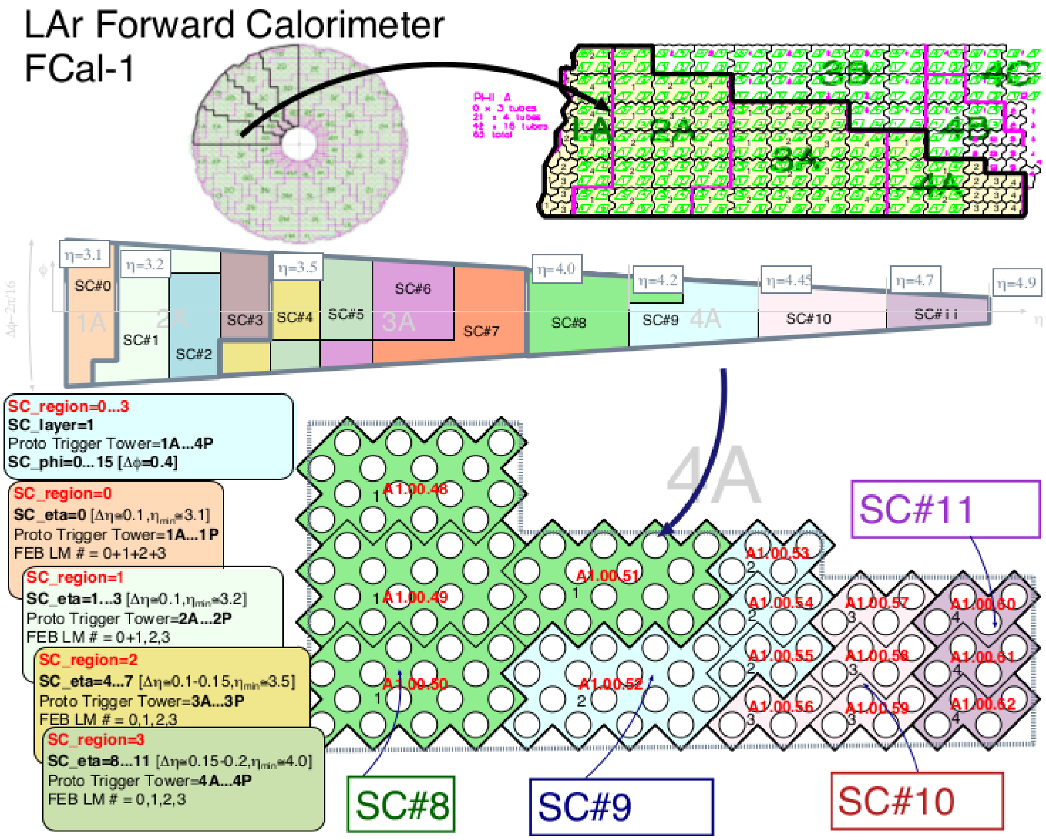

Figure 5.9: Planned supercell layout for FCAL1. Layers 2 and 3 of the FCAL also have 16 wedges in ϕ, but are divided into only 8 and 4 supercells in η respectively.

5.3 Forward Jet Algorithms

Forward jets present a particularly challenging case for the L1 trigger. This is largely due to calorimeter granularity which is much coarser than that in the central regions of the detector. Additional complications are introduced by the fact that the calorimeter cells (and as a result, the supercells) in the FCAL have regular shapes in the x-y plane rather than the η-ϕ plane. This implies that the actual geometry of the FCAL supercells is irregular in both size and shape, so there is no obvious way to map them into square sliding windows as is done in the central region. This geometry is illustrated in Figure 5.9 for a typical piece of the first FCAL layer. The other complication is that pileup effects tend to be worse in the forward region, making it even more difficult to distinguish signal jets from the stochastic background.

This situation means that the L1 trigger typically has much worse performance for forward jets than for central jets (and this is even worse in the current system, where each forward trigger towers spans the entire η range from 3.1 to 4.9). This presents a problem for VBF analyses which want to trigger on forward jets[36]. So, several methods for improving forward jet trigger performance were studied, in the form of modifications to the jFEX sliding window algorithm. There are some restrictions: any such algorithm needs to be implementable in the FPGAs which actually run the algorithm, and the only available information is the jTowers. For the FCAL, it is potentially possible to pass each individual supercell to the jFEX as its own "jTower" (which can overlap with other ones), so these studies treat the supercells as the fundamental objects on which the algorithm runs in the FCAL.

The simplest strategy is to simply bin the supercells into rectangular jTowers (typically 0.2×0.2) like the ones in the rest of the detector and assign them into windows based on that. However, this assumes several symmetries of the FCAL supercells which aren't actually realized in the detector, so the locations of the supercells can be artificially shifted into and out of the "correct" windows. Additionally, a single supercells almost always overlaps two or more of these rectangular jTowers, so in these cases the energy is divided up based on the size of the overlap. This is the closest method to what is used in the current trigger system, so it is considered as the baseline algorithm to compare against.

The approach taken here for alternatives is to consider generalizations of the standard sliding window algorithm. The generalization comes into the definition of which supercells are included in which windows, such that the windows need not be square or even regular. One method to assign towers into windows is based on the anti-k_t algorithm itself. Here, anti-k_t is run over the individual FCAL supercells in a "training sample" to form jets which have shapes as similar as possible to offline jets. These jets are binned into a η-ϕ grid, where each bin is associated with a single window. For each of these bins, all of the jets at that location are used to construct and "average jet shape". In practice this is quantified by the average amount of energy each supercell contributes to jets at that location, denoted ξ. This can be defined either in terms of absolute E_{T} or as a fraction of the average jet energy in the sample at that location. Finally, all supercells with ξ above some threshold are included in the window. This procedure is illustrated in Figure 5.10. The mapping from supercells to windows only needs to be computed once and can be stored in a look-up table for later use.

---

[36] The VBF analysis presented in the later chapters of this document uses a $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger rather than forward jets. However, $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers tend to be extremely pileup-sensitive, so the jet triggers will become increasingly important as the LHC luminosity increases.

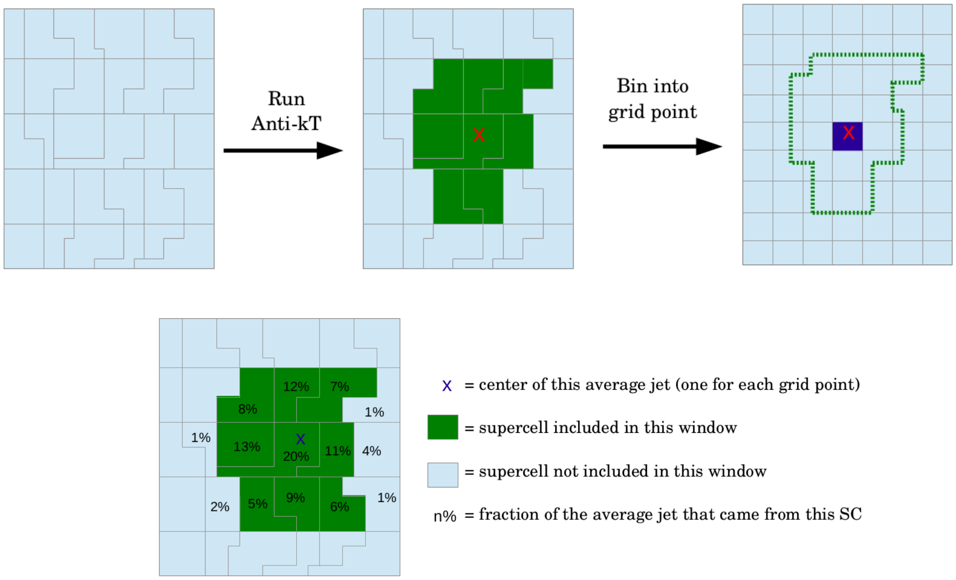

Figure 5.10: Sketch of the anti-k_t-based method for defining the forward jFEX windows. In this example, a threshold of $\xi > 0.05E_{\mathrm{T}}^{\mathrm{jet}}$ is used.

The local maximum requirement in the standard sliding window also needs some analog here, in order to avoid double counting. In this case, the "core" of the jet is defined as the N supercells with the largest ξ, where N is a free parameter. It was found that the performance of the algorithm is not sensitive to the choice of N, and for these studies $N = 4$ was used. This algorithm was tested using a minimum-bias MC simulation as the training sample. It was found that a ξ threshold of zero (i.e. including any supercells which anti-k_t ever clusters into jets at the given location) gave the best performance. This algorithm turned out to have very similar performance to the simple method of binning the FCAL supercells into rectangular jTowers. The results are summarized in Figure 5.11, and a detailed comparison of efficiencies between this and the baseline algorithm is shown in Figure 5.12.

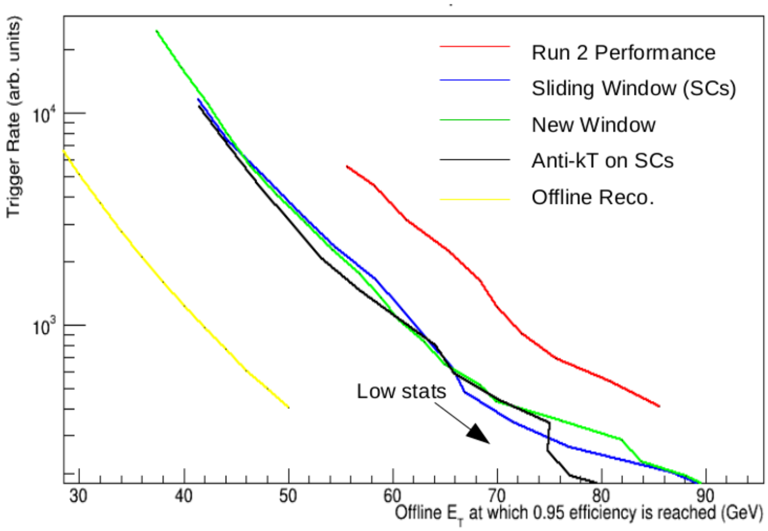

Figure 5.11: Single jet ($|\eta| > 3.2$) performance summary for the anti-k_t-based method for mapping FCAL supercells into windows (denoted "New Window" in the legend). This used a threshold of $\xi > 0$ and $N = 4$ for the local maximum finding. Several other performance curves are also shown for reference; the standard "Sliding Window" (blue curve) uses supercells binned into jTowers.

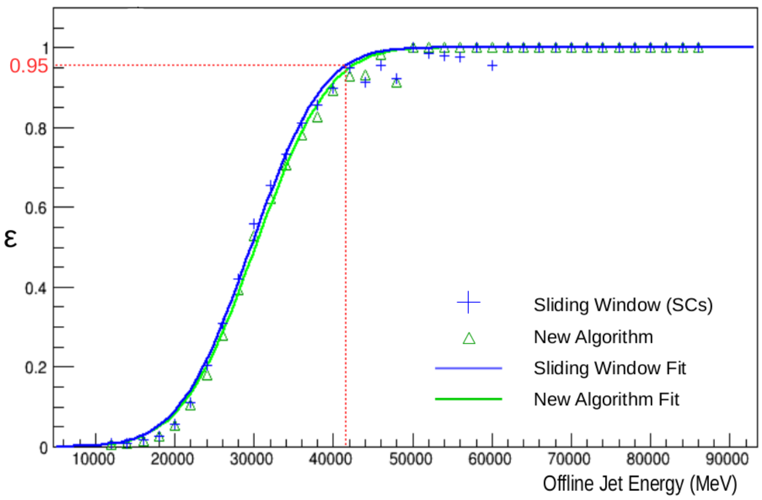

Figure 5.12: Comparison of efficiencies at fixed rate between the anti-k_t-based supercell mapping ("New Algorithm") and the basic rectangular binning. Fits to error functions are also shown, and the 95% efficiency point is indicated as a reference.

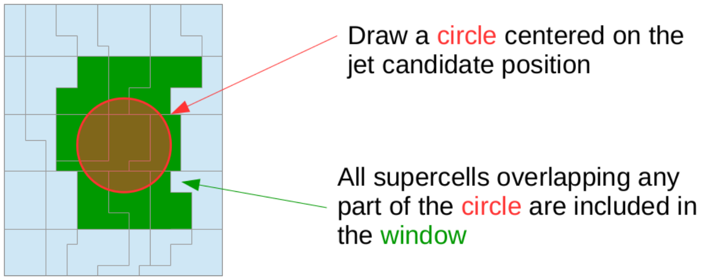

Figure 5.13: Sketch of the geometric method for defining the forward jFEX windows.

Another method for defining the mapping from supercells to windows is to use only the geometry of the supercells themselves. This removes any dependence on a training sample (and the associated statistical effects). The approach implemented here was to draw a circle in η-ϕ space around the center of each window with some variable radius R, representing roughly the shape of a jet in that location. Any supercell which has any overlap with this circle is included in the window. This is illustrated in Figure 5.13.

Here, Gaussian weighting is applied to suppress distant supercell with only a small overlap, which reduces sensitivity to pileup and other noise in the other supercells. This involves multiplying each supercell energy by some factor before adding it to the window's total energy. The factor is computed as a Gaussian function of the distance between the window center and the supercell centroid. It was found that once the Gaussian weighting is applied, the performance of the algorithm is insensitive to the choice of R, provided it is at least ~ 0.3. The characteristic width of the Gaussian weighting function is variable, but typically taken to be roughly 0.2 or 0.3. The performance of this algorithm compared to the baseline one (and several other references) is shown in Figure 5.14.

Both the anti-k_t-based and the geometrical method for mapping supercells into windows provided little improvement over a naive binning of the supercells into rectangular jTowers. Futhermore, the performance of all three of these methods is comparable to actually running the anti-k_t algorithm itself on the supercells. This appears to indicate that the limiting factor for L1 forward jet trigger performance is limited not by how sophisticated the algorithm is, but rather by the amount of information actually available in the supercells. This can be interpreted to mean that the performance can only be subtantially improved in terms of resolving jets by improving the granularity of the FCAL readout.

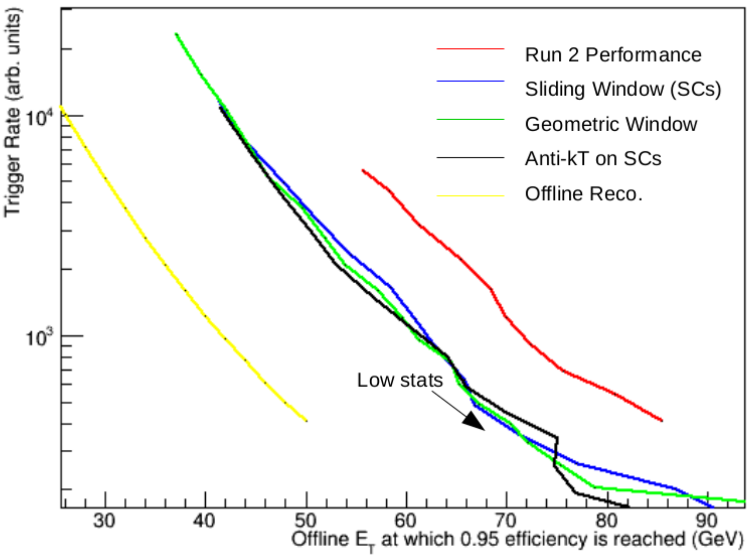

Figure 5.14: Single jet ($|\eta| > 3.2$) performance summary for the geometric method for mapping FCAL supercells into windows. This used a radius parameter of $R = 0.4$, and Gaussian weighting was also applied. Several other performance curves are also shown for reference; the standard "Sliding Window" (blue curve) uses supercells binned into jTowers.

5.4 Pileup Suppression Methods

Hadronic triggers are particularly sensitive to pileup (compared to lepton triggers for example), so as the LHC luminosity increases, it is expected that the performance of the existing algorithms will degrade substantially. This is due to jets from pileup vertices as well as "stochastic jets" formed by coincidences of random soft activity from multiple interactions. Because of this, it is necessary to investigate ways of mitigating these effects for future LHC runs.

One of the simplest methods for dealing with this is area-based pileup subtraction, or "ρ subtraction". This involves calculating an average pileup density per unit area in η-ϕ space (denoted ρ), and subtracting ρA from the energy of any object, where A is the area of the object in question. ρ can be calculated at the event level, but in practice it is typically done in several ϕ-bands per event. This is equivalent to the assumption that the pileup is diffusely distributed throughout the detector is is uniform in ϕ but not η. Objects identified as jets are not included in the calculation of ρ, so that the hard interaction does not influence it. Applying this technique can allow for the reduction of trigger rates while maintaining the same efficiencies.

This technique was implemented in the jFEX simulation and tested in the context of Phase-II conditions, as this is when pileup mitigation will be most relevant. Results from this study for a typical multi-jet trigger are shown in Figure 5.15. This indicates that the ρ subtraction method substantially improves the trigger performance in this high-pileup scenario. Pileup subtraction tends to have less impact on single-jet triggers, as the relevant p_{T} thresholds are generally well above the scale of pileup contributions.

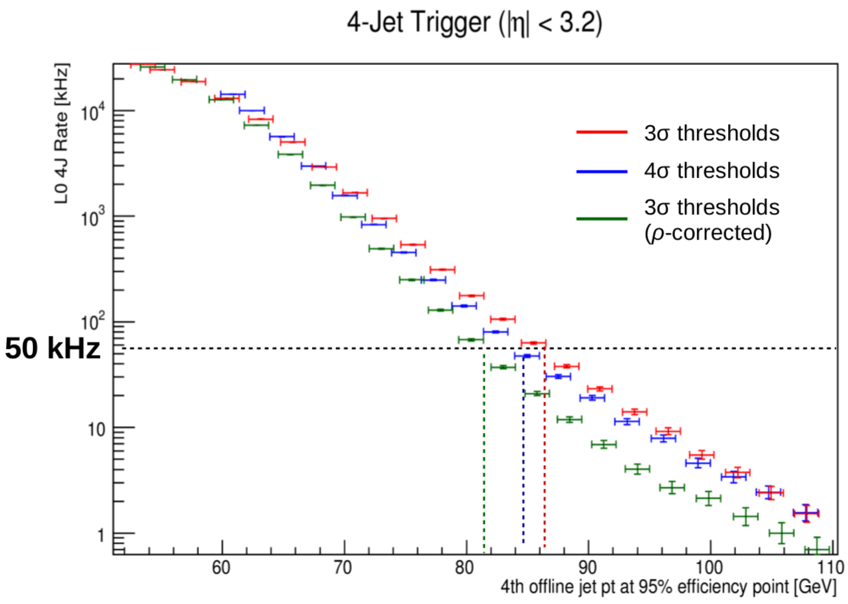

Figure 5.15: Simulated performance summary for 4-jet ($|\eta| < 3.2$) triggers with jFEX at $\mu = 200$. Two different sets of noise thresholds (3σ and 4σ) are shown, as well as 3σ with the additional pileup suppression using ρ subtraction. The nominal rate budget of 50 kHz is also indicated.

Chapter 6

VBF+$E_{\mathrm{T}}^{\mathrm{miss}}$ Analysis Overview

6.1 Introduction

The purpose of this analysis is to search for dark matter production via weak vector boson fusion (VBF). The primary signal model is to have the SM Higgs boson as the mediator between SM and DM particles (Figure 6.1). In this context, the goal is to place an upper limit on the branching ratio of the Higgs boson to invisible decays (or observe such decays, in the case of a positive signal[37]). Results from this search can also be interpreted in terms of other models, with beyond-SM mediators. This will be discussed further in Section 8.3.

Several collider limits already exist on this quantity, from a variety of channels. VBF is in general the most sensitive, as it has the largest production cross section (aside from ggF, which has much larger backgrounds). The ATLAS and CMS experiments have both set limits using the VBF channel at $\sqrt{s} = 8$ TeV [5, 84]. The sensitivity can be improved by combining with limits from other channels (in particular, associated production with Z and ggF). ATLAS has additionally combined these results with constraints from visible channels at 8 TeV, to set the current tightest limit of 23% on $\mathcal{B}(h \to \text{inv.})$ [7]. CMS has also reported results from a very small set of early 13 TeV data, combined with their 8 TeV results [6].

The main challenge in searching for dark matter at the LHC (especially in the VBF channel) is the large amount of background. Roughly half of this background comes from $Z \to \nu\nu$ production with two additional jets. This can be produced via either "strong" or "electroweak" processes (that is, order 2 or 4 in the electroweak coupling), as shown in Figure 6.2. Due to the difference between

---

[37] Observation of a positive signal would not necessarily imply BSM Higgs decays, since other other mediators could produce the same signature.

the strong and electroweak coupling, the strong-produced component has a much larger cross section. However, is can be effectively discriminated against by taking advantage of the different kinematics between the production modes. VBF processes always have two jets in the final state, and they have a strong tendency to be well-separated in η. Due to the lack of color exchange, there also tends to be less gluon radiation relative to the "strong" diagrams. So, the vast majority of the strong background can be removed by making strict kinematic requirements when selecting events. $Z \to \nu\nu$ can also be produced via VBF (which falls under the "electroweak" category). This has nearly identical kinematics to the signal process[38], so it is considered an "irreducible" background.

The rest of the background is predominantly composed of $W \to \ell\nu$, which includes electron, muon, and tau final states. In the electron and muon cases, this occurs because the lepton is not in the detector acceptance (outside of the η range or its p_{T} is too low), or because it fails the identification and isolation criteria (which can happen for a variety of reasons). In the tau case, it either decays leptonically (which reduces to the previous case) or hadronically. If the decay is hadronic, the tau should be reconstructed as a jet, but this can sometimes fail due to acceptance effects (typically in p_{T}). In all of these cases, the lepton is "lost", and the event looks identical to signal. Although most leptons are successfully identified, the total $W \to \ell\nu$ background is roughly the same size as the $Z \to \nu\nu$ background. This is due to the larger W production cross section and leptonic branching ratios. Exactly like the Z case, this can be produced via "strong" or "electroweak" processes, which are both relevant.

The W and Z backgrounds can be estimated directly from Monte Carlo simulation, but the uncertainties associated with doing so are far too large to make this useful. Instead, data-driven methods are employed. In addition to the signal regions (i.e. the kinematic selection within which the search is carried out), control regions are defined which are used to correct for mismodeling in the MC. The control regions are expected to contain no signal contribution. They are designed to capture $Z \to \ell\ell$ and $W \to \ell\nu$ events, which ℓ indicates an electron or muon (which is successfully identified in this case). The principle behind this is that, for example, $Z \to ee$ events have extremely similar properties to $Z \to \nu\nu$ events, except for the cross section. So, the measurement of $Z \to ee$ in the control regions can be used to constrain the $Z \to \nu\nu$ in the signal regions, using assumptions about correlations between the two. In practice, this is implemented in a single simultaneous fit to all signal and control regions. Accurate modeling of the W and Z background is the central problem in this analysis; the methods used to handle this are discussed in full detail in Section 7.2.

---

[38] This is less true for significantly heavier mediators, which tend to have more forward jets.

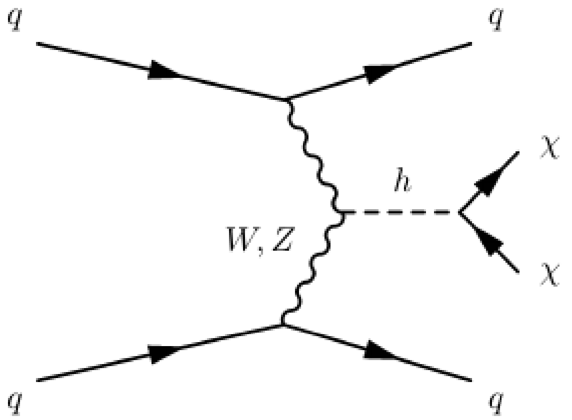

Figure 6.1: Feynman diagram for the Higgs signal model. χ need not be a fermion. Alternative mediator models are exactly the same, but with the Higgs propagator exchanged for the other mediator.

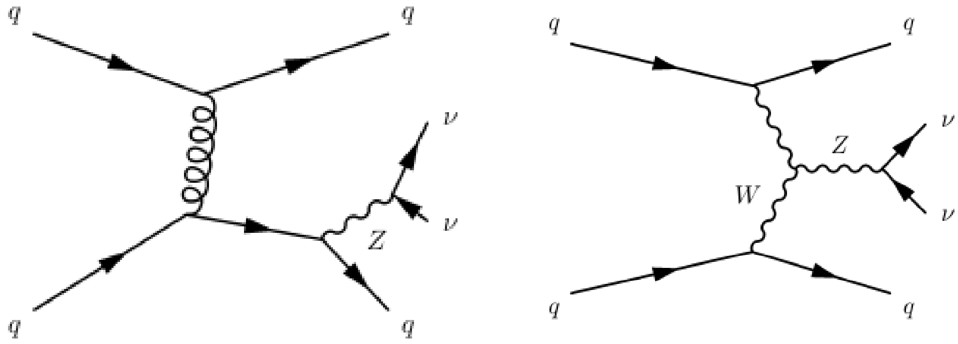

Figure 6.2: Typical leading-order Feynman diagrams for strong **(left)** and electroweak **(right)** production of the $Z \to \nu\nu$ background. The $W \to \ell\nu$ background has analogous diagrams.

The fit which is used to constrain the background estimates is based on the construction of a likelihood function defined over all regions. This includes background normalizations (denoted k_W and k_Z) which are allowed to float freely, meaning no assumption is made about the MC modeling of them. Nuisance parameters which account for systematic uncertainties are also included in the likelihood function. The full details on the implementation of the fit and associated statistical methods are discussed in Section 8.1.

6.2 Event Selection

Events used in the analysis are classified into many orthogonal signal regions (SRs), control regions (CRs), and validation regions (VRs). The signal regions were kept blinded until all definitions were finalized and the analysis strategies validated. Note that the names of the signal and control regions are only indicative of the purpose they were designed for; the only distinction in the statistical model[39] is that the SRs are expected to contain signal while the CRs are not. Validation regions are not used in the fit; their only purpose is to validate the methods used in the analysis. All selected events are required to contain a primary vertex with at least two associated tracks and pass the Tight jet cleaning criteria[40]. The primary vertex must be associated with at least two tracks which each have $p_{\mathrm{T}} > 400$ MeV. In the case of multiple vertices passing this requirement (which is nearly always the case), the one with the largest scalar sum of associated track p_{T} is chosen. The definitions of the signal regions were chosen based on the optimization procedure described in Section 6.2.4.

For an event to be considered part of a signal region, it must also satisfy the following requirements:

- The event passed the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger.
- The event contains no electrons or muons.
- The event has $E_{\mathrm{T}}^{\mathrm{miss}} > 180$ GeV.
- The event contains *exactly* two jets, with no additional jets over 25 GeV.
- The leading jet has $p_{\mathrm{T}} > 80$ GeV.
- The sub-leading jet has $p_{\mathrm{T}} > 50$ GeV.

---

[39] A complete description of the statistical model used in this analysis and the associated fit is given in Section 8.1.

[40] Jet cleaning is described briefly in Section 4.1.4.

- The two jets lie in opposite z-hemispheres: $\eta_{j1}\eta_{j2} < 0$.
- The two jets are well-separated in η: $\Delta\eta_{jj} > 4.8$.
- The two jets are not back-to-back: $\Delta\phi_{jj} < 1.8$.
- The dijet system has a large invariant mass: $m_{jj} > 1$ TeV.
- Neither jet is aligned with the $E_{\mathrm{T}}^{\mathrm{miss}}$ in ϕ: $\Delta\phi_{j,MET} > 1.0$ for both jets.
- The event has $H_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV, where *all* jets are included in the calculation (even those tagged as pileup).

These events are then divided up into three signal regions, depending on the dijet mass. These are denoted SR1 (1.0 TeV $\leq m_{jj} < 1.5$ TeV), SR2 (1.5 TeV $\leq m_{jj} < 2.0$ TeV), and SR3 ($m_{jj} \geq 2.0$ TeV). Together, they are collectively referred to as the “inclusive signal region”.

The jet-related cuts are designed to suppress non-VBF backgrounds as much as possible. The ideal signal event consists of the two jets (corresponding to the final-state quarks from VBF) recoiling against the invisibly decaying Higgs. This topology is illustrated in Figure 6.3. An event display for an actual event passing the SR selection is shown in Figure 6.4. The large amount of $E_{\mathrm{T}}^{\mathrm{miss}}$ is required for two reasons: to distinguish the signal from backgrounds with no invisible particles, and to make it possible to trigger on the events. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is meant to represent the p_{T} of the Higgs boson. Due to the finite η coverage of the calorimeters, the opposite-hemisphere requirement is already implied by the $\Delta\eta_{jj}$ cut. However, this is not the case when the $\Delta\eta_{jj}$ cut is modified for the purposes of defining a validation region, so it is included here for consistency.

6.2.1 Control Regions

In order to constrain the $W \to \ell\nu$ background in the signal regions, W control regions (also called 1-lepton control regions) are defined. These are intended to select a sample of W events as similar as possible to those that enter the SRs, except that the lepton from the W decay is successfully reconstructed. Only electrons and muons are considered here (taus always decay into one of these, or into hadrons which result in additional jets). The selection for these regions is exactly the same as the SRs, except:

- The event passed a single lepton trigger.
- The event contains *exactly* one electron or muon (corresponding to the trigger it passed).

- The lepton has $p_{\mathrm{T}} > 30$ GeV.
- The lepton's p_{T} is explicitly added to the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector-wise before imposing the $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\Delta\phi_{j,MET}$ requirements.

The purpose of adding the lepton to the $E_{\mathrm{T}}^{\mathrm{miss}}$ is so that the variable being cut on is still equal to the W boson p_{T} (up to detector effects), as it is in the SRs. The lepton p_{T} cut is driven by the trigger efficiency; events with softer ones cannot reliably be recorded. The W control regions are also split up by lepton charge, in order to help constrain backgrounds from multijet events where a jet is misidentified as a lepton (see Section 7.4). This works because W^+ and W^- events are produced at different rates at the LHC, but the multijet background is very nearly charge-symmetric.

In order to suppress this background from multijet events, an additional requirement is imposed on the "$E_{\mathrm{T}}^{\mathrm{miss}}$ significance" for the electron CRs only. This is not necessary for the muon CRs, since jets misidentified as muons are already rare enough to be negligible. The $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is defined in terms of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and object transverse momenta as:

$$E_{\mathrm{T}}^{\mathrm{miss}}\mathrm{sig.} = \frac{E_{\mathrm{T}}^{\mathrm{miss}}}{\sqrt{p_{\mathrm{T}}^{j1} + p_{\mathrm{T}}^{j2} + p_{\mathrm{T}}^{\ell}}} \tag{6.1}$$

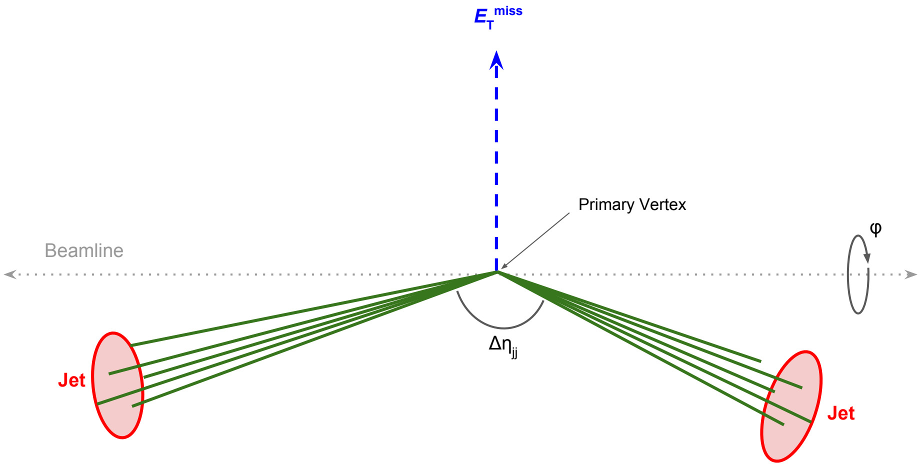

Figure 6.3: Schematic diagram of a typical signal event.

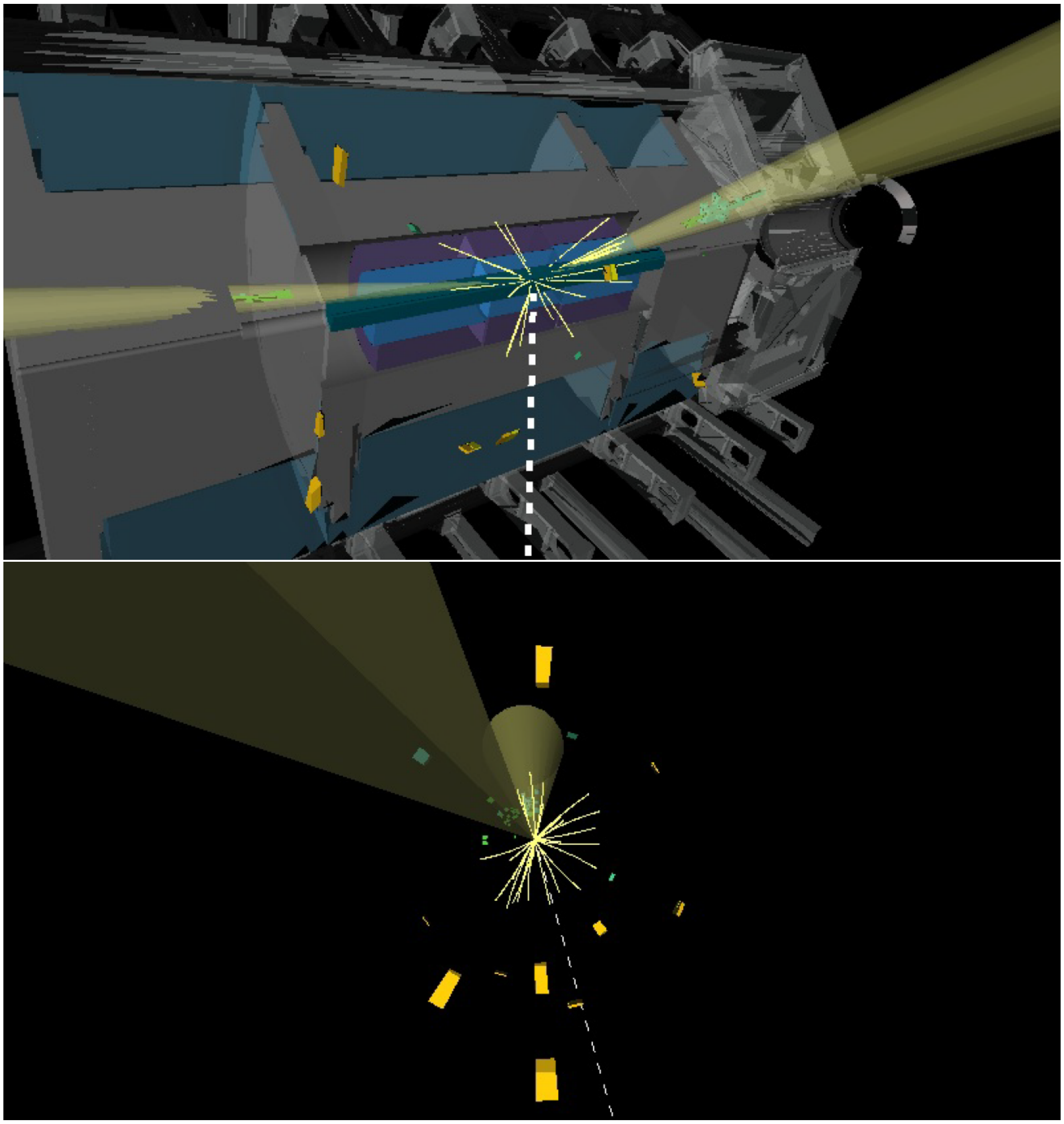

Figure 6.4: Event display for a typical event in the signal region, showing a side view **(top)** and a projection in the xy plane **(bottom)**. The yellow cones indicate the two jets, and the dashed line shows the direction of the $E_{\mathrm{T}}^{\mathrm{miss}}$. Calorimeter deposits with $E_{\mathrm{T}} > 1$ GeV and tracks with $p_{\mathrm{T}} > 2$ GeV are also shown. This is event 894673740 from Run 305723 (Aug. 6, 2016).

The requirement for an event to enter the 1-electron CRs is $E_{\mathrm{T}}^{\mathrm{miss}}$sig. > 4.0 GeV$^{1/2}$. This value was motivated by the comparing data in these regions to MC simulation (which does not include multijet events) as a function of the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance (Figure 6.5). This value is where the disagreement, corresponding to the fake electron contribution, becomes small. Note that the 1-muon CR has no significant disagreement. Applying this cut in the 1-electron region sucessfully removes the data/MC discrepancy at low $E_{\mathrm{T}}^{\mathrm{miss}}$, where multijet events are concentrated (Figure 6.6).

In order to constrain the $Z \to \nu\nu$ background in the signal regions, $Z \to \ell\ell$ control regions (also called 2-lepton control regions) are defined. Again, only electrons and muons are considered here. Since both of these particles have masses which are negligible at the relevant scales, the kinematics of $Z \to \ell\ell$ events are nearly identical to those of $Z \to \nu\nu$ events (small differences can arise due to photon radiation from charged leptons and the detector acceptance for charged leptons). The selection for these regions is exactly the same as the SRs, except:

- The event passed a single lepton trigger.
- The event contains *exactly* two leptons, which have the same flavor (corresponding to the trigger) and opposite charge.
- The leading lepton has $p_{\mathrm{T}} > 30$ GeV.
- The sub-leading lepton has $p_{\mathrm{T}} > 7$ GeV.
- The mass of the dilepton system is near the Z mass: $|m_{ll} - m_Z| < 25$ GeV.
- The p_{T} of the dilepton system is explicitly added to the $E_{\mathrm{T}}^{\mathrm{miss}}$ vector-wise before imposing the $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\Delta\phi_{j,MET}$ requirements.

6.2.2 Validation Regions

A “two-jet” (2J) validation region is defined by loosening the VBF-targeted cuts relative to the SRs. This region allows validation of fitting and background estimation methods in a signal-free environment. The definition is the same as the SRs, except:

- The dijet mass requirement is inverted: $m_{jj} < 1.0$ TeV.
- The dijet $\Delta\eta$ requirement is relaxed: $\Delta\eta_{jj} > 2.5$.

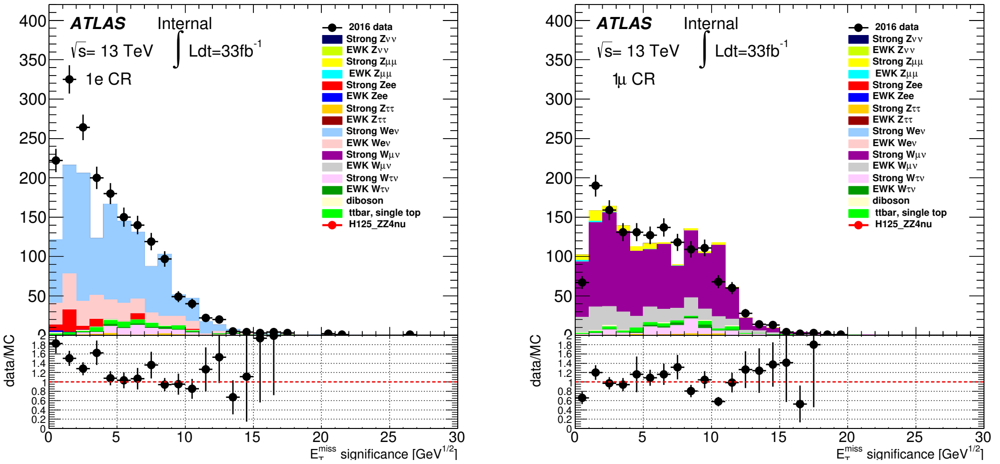

Figure 6.5: **Left:** $E_{\mathrm{T}}^{\mathrm{miss}}$sig. distribution in the inclusive $W \to e\nu$ CR. **Right:** $E_{\mathrm{T}}^{\mathrm{miss}}$sig. distribution in the inclusive $W \to \mu\nu$ CR.

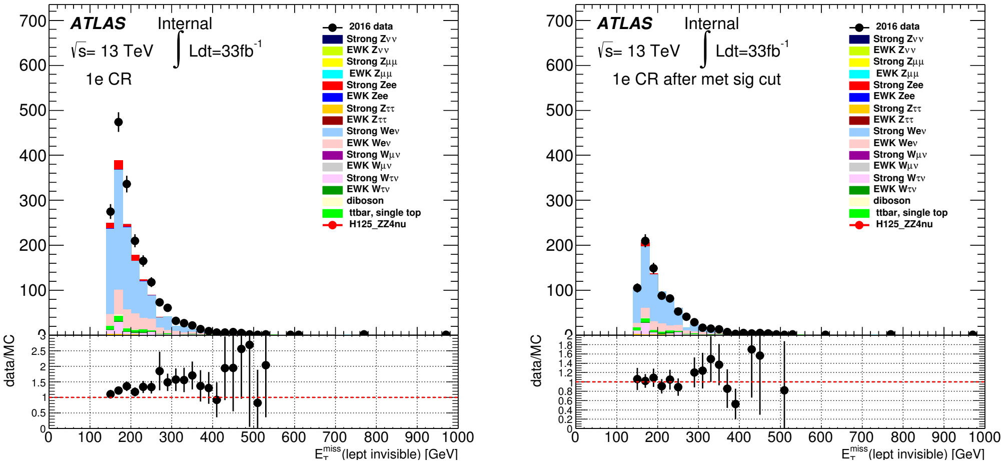

Figure 6.6: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution (including the electron) in the inclusive $W \to e\nu$ CR before (**left**) and after (**right**) applying the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance cut.

These two quantities are very strongly correlated, so inverting the m_{jj} cut without also loosening the $\Delta\eta_{jj}$ cut would result in very low statistics in the VRs (and bias it towards having softer jets). With this definition, the 2J VR has roughly 8% signal contamination, assuming $\mathcal{B}(h \to \text{inv.}) = 100\%$. In practice, this is already constrained to less than 25%, so we expect no more than 2% signal contamination. These signal contamination numbers are taken from MC simulation of signal and background yields in the VR.

Analogously to the W and Z control regions, additional VRs with 1 and 2 leptons are also defined. These have exactly the same lepton requirements as the the CRs; the only differences are the m_{jj} and $\Delta\eta_{jj}$ cuts. They are also separated by charge. None of the VRs are split into multiple m_{jj} bins, so there is only one $E_{\mathrm{T}}^{\mathrm{miss}}$ VR, one $Z \to ee$ VR, etc.

6.2.3 Triggers

Events in the signal region are recorded using an $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger at both L1 and HLT. The trigger threshold was raised several times through the 2015-2016 data-taking period to keep the rates within technical limits as the pileup increased. The exact triggers used are listed in Table 6.1 ("XE" simply means $E_{\mathrm{T}}^{\mathrm{miss}}$, and the numbers are the thresholds in GeV). In the HLT, they all use the $H_{\mathrm{T}}^{\mathrm{miss}}$ version of the algorithm, in order to minimize pileup sensitivity (indicated by "mht" in the trigger name). So, for example, "HLT_xe70_mht_L1XE50" means there is a $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement of 50 GeV at L1, followed by a $E_{\mathrm{T}}^{\mathrm{miss}}$ cut of 70 GeV at HLT using the $H_{\mathrm{T}}^{\mathrm{miss}}$ definition. In addition to these, events from the trigger HLT_noalg_J400 are also included, to catch events with saturated L1 trigger towers that incorrectly failed the L1 $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger due to a firmware bug.

The trigger efficiency, defined with respect to the offline energy scale, is measured in data by using a single muon trigger in data (targeted at $W \to \mu\nu$ events, which also have $E_{\mathrm{T}}^{\mathrm{miss}}$). In order to reduce resolution differences with respect to the signal regions, an additional requirement that these events have exactly two jets is applied. The results of this measurement for one example trigger are shown in Figure 6.7. This indicates that the MC simulation overestimates the efficiency by an amount which depends on the $E_{\mathrm{T}}^{\mathrm{miss}}$. This mismodeling is corrected for in MC using a scale factor, defined as the ratio of efficiency measured in data to effciency predicted by MC. In order to smooth out statistical fluctuations, the scale factor is fit to an error function:

$$\mathrm{SF} = \frac{\epsilon_{\mathrm{data}}}{\epsilon_{\mathrm{MC}}} = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{E_{\mathrm{T}}^{\mathrm{miss}} - p_0}{\sqrt{2}p_1}\right)\right) \tag{6.2}$$

where p_0 and p_1 are free parameters in the fit.

For the W and Z control regions, single lepton triggers are used. These are listed in Table 6.2. The naming scheme is the same as that of the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers, but with a few additional elements. For example, "lhmedium" indicated that the likelihood-based ID was used at the Medium working point. Items such as "ivarloose" correspond to isolation requirements, and "nod0" means that the lepton candidate's d_0 (i.e. transverse displacement from the beam axis) is not taken into account in the likelihood function. For these triggers, efficiency and scale factor measurements are provided by the ATLAS Combined Performance groups, using methods very similar to those used for the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers.

6.2.4 Selection Optimization

In order to choose values for the jet-related cuts (jet p_{T}, $\Delta\eta_{jj}$, $\Delta\phi_{jj}$, m_{jj}, and $E_{\mathrm{T}}^{\mathrm{miss}}$), a multivariate optimization was performed. The optimal cut values are defined as those which result in the smallest expected upper limit on $\mathcal{B}(h \to \text{inv.})$. All of these cuts are simultaneously varied over many values, to cover a parameter space with dimension equal to the number of cuts being optimized. At each point (on a discretized grid) in this parameter space, the expected upper limit on $\mathcal{B}(h \to \text{inv.})$ is estimated, and the point that gives the smallest result is chosen.

The actual estimation of the branching ratio limit for optimization purposes is done as a signal-region-only calculation. This uses a simplified analytical approximation of the expected limit in place of the full fit in order to efficiently scan over optimization points. Here, the quantity of interest is the signal strength parameter μ, defined as:

$$\mu = \frac{N - \sum_b B_b}{S} \tag{6.3}$$

where N is the observed number of events, B_b is the expected yield from background b, and S is the expected signal yield in the case of $\mathcal{B}(h \to \text{inv.}) = 1$. The Standard Model predicts the μ should have a distribution centered at zero, with some uncertainty σ_μ. Note that σ_μ is itself a function of μ, as it depends on N, S, and B, which will be shown explicitly later in this section. Assuming this distribution is Gaussian (i.e. $N \gg 1$ and the systematic contributions are Gaussian), the 95% CL upper limit on $\mathcal{B}(h \to \text{inv.})$ is equal to μ where:

Period	Trigger
2015	HLT_xe70_mht_L1XE50
2016, Run $\leq$ 304008	HLT_xe90_mht_L1XE50
2016, Run $>$ 304008	HLT_xe110_mht_L1XE50

Table 6.1: $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers used to record events for the signal regions.

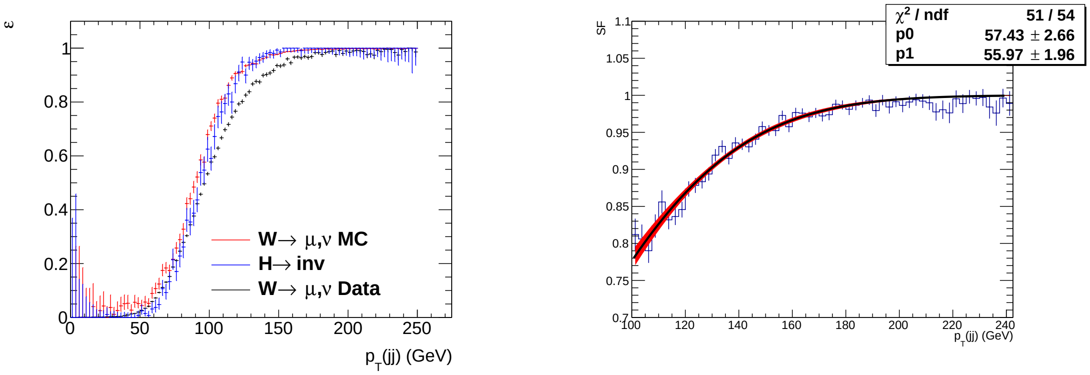

Figure 6.7: **Left:** Trigger efficiency for HLT_xe90_mht_L1XE50, shown for $W \rightarrow \mu\nu$ MC and data as well as signal MC. **Right:** The corresponding scale factor (see text), shown with a fit. Here, these are shown as a function of dijet p_{T}, which is almost perfectly correlated with $E_{\mathrm{T}}^{\mathrm{miss}}$ in the signal regions.

Period	Electron Trigger	Muon Trigger
2015	HLT_e24_lhmedium_L1_EM20VH HLT_e60_lhmedium HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 HLT_mu50
2016, Run $\leq$ 304008	HLT_e24_lhtight_nod0_ivarloose HLT_e60_lhmedium HLT_e140_lhloose	HLT_mu26_ivarmedium HLT_mu50
2016, Run $\leq$ 304008	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium HLT_e140_lhloose	HLT_mu26_ivarmedium HLT_mu50

Table 6.2: Lepton triggers used to record events for the control regions.

$$\frac{\mu}{\sigma_\mu} = 1.64 \tag{6.4}$$

The number 1.64 here corresponds approximately to a 95% one-sided confidence bound on a Gaussian distribution. This equation is solved numerically by varying μ until the condition is satisfied. This can be converted into an expected CL_{s} limit[41] by taking the expected p-value for the background-only hypothesis as 0.5. That is, the average result under the background-only hypothesis should give a p-value of 0.5 with respect to that hypothesis.

6.2.4.1 Unbinned Optimization

The simplest case is the one with a single, unbinned signal region. In practice, there are also systematic uncertainties on the signal and background predictions which can be correlated with each other. These are treated by adding nuisance parameters α to the predicted yields:

$$\mu = \frac{N - \sum_b (B_b + \sum_s \alpha_s \Delta B_{bs})}{S + \sum_s \alpha_a \Delta S_s} \tag{6.5}$$

where s indexes the different systematics and ΔB and ΔS refer to the systematic uncertainties on the backgrounds and signal respectively. The nuisance parameters α are defined to have a Gaussian distribution with mean $\bar{\alpha} = 0$ and width $\sigma_\alpha = 1$. Note that s is the only index on α, meaning that each source of systematic uncertainty is treated as fully correlated between the signal and all the backgrounds (if they're uncorrelated, they are treated as separate nuisance parameters). From this, the full uncertainty on μ is given by:

$$\sigma_\mu = \frac{1}{S}\sqrt{N + \underbrace{\mu^2\sigma_S^2 + \sum_b \sigma_{B_b}^2}_{\text{MC stat. uncertainty}} + \underbrace{\sum_s \left(\mu\Delta S_s + \sum_b \Delta B_{bs}\right)^2}_{\text{Systematic uncertainty}}} \tag{6.6}$$

where σ_S and σ_B refer to the statistical uncertainties on the signal and background yield predictions respectively. Here, these correspond to MC statistical uncertainties.

To check the validity of the approximations used here, the expected limits obtained from this method were compared to calculations done using a SR-only fit with the full statistical formalism

---

[41] CL_{s} is a modified frequentist formalism designed for searches such as this one [85]. It is discussed further in Section 8.1.

(see Chapter 8). The results were found to be consistent to within a factor of 10% or less, depending on the particular cuts being used. Near perfect agreement was found in the regime with a very tight VBF selection (which is expected, because this is where systematic uncertainties matter least).

This optimization procedure was carried out for an integrated luminosity of 40 fb^{-1}. Systematic uncertainties were assumed to be the same as in Run 1 [5], as this was done very early in the analysis process and no newer estimates were yet available. The MC samples used for this optimization also had lower statistics than the final ones. Only the W and Z backgrounds were considered here; the rest were assumed to be negligible. Results from these studies indicated that the optimal definition for an unbinned signal region has substantially stricter VBF requirements compared to those used in the 8 TeV analysis. For the jet-related cuts, the optimization procedure gave the following selection criteria:

- Leading jet $p_{\mathrm{T}} > 80$ GeV
- Subeading jet $p_{\mathrm{T}} > 50$ GeV
- $\Delta\eta_{jj} > 4.8$
- $\Delta\phi_{jj} < 1.8$
- $m_{jj} > 2.0$ TeV
- $E_{\mathrm{T}}^{\mathrm{miss}} > 150$ GeV

These cuts gave an expected branching ratio limit of 29%. In particular, the m_{jj} and $\Delta\phi_{jj}$ cuts are substantially tighter than what was used at 8 TeV. There are two reasons for this. The first is that all the relevant cross sections are increased due to the higher center-of-mass energy, effectively reducing the statistical uncertainties. The second is that the 13 TeV Monte Carlo predicts a much larger amount of strong-produced W/Z+jets background, which can be discriminated against with these cuts. Since systematic uncertainties on the background would be by far the dominant contribution with the cuts used in the Run 1 analysis, it has become especially important to have a high signal-to-background ratio.

The values of the jet p_{T} and $\Delta\eta_{jj}$ cuts were found to have little impact on the expected limit. This is due to the fact that the strict m_{jj}, $\Delta\phi_{jj}$, and $E_{\mathrm{T}}^{\mathrm{miss}}$ cuts already implicitly constrain the values that those variables can (or are likely to) have.

6.2.4.2 Binned Optimization

We can construct a more sophisticated set of signal regions by binning in one or more variables. The same optimization procedure is generalized to accommodate this. Here, μ is calculated as a weighted average over all the bins:

$$\mu = \sum_i w_i \mu_i = \sum_i w_i \frac{N_i - \sum_b (B_{bi} + \sum_s \alpha_s \Delta B_{bsi})}{S_i + \sum_s \alpha_s \Delta S_{si}} \tag{6.7}$$

where i indexes the different bins and the weights w_i are given by:

$$w_i = \frac{\sigma_{\mu_i}^{-2}}{\sum_j \sigma_{\mu_j}^{-2}} \tag{6.8}$$

Thus, bins with the best sensitivity (i.e. smallest σ_{μ_i}) are assigned the largest weights. When calculating the bin weights, the individual σ_{μ_i} are calculated independently of each other. This implies that the weights do not take into account any uncertainty correlations between the bins. Now, the total ucnertainty on μ is:

$$\sigma_\mu = \frac{1}{S} \sqrt{\sum_i w_i^2 \left(N_i + \mu_i^2 \sigma_{S_i}^2 + \sum_b \sigma_{B_{bi}}^2 \right) + \sum_s \left(\sum_i w_i \left(\mu_i \Delta S_{si} + \sum_b \Delta B_{bsi} \right) \right)^2} \tag{6.9}$$

Note that the systematic terms are summed over the different bins before squaring. This is because each source of systematic uncertainty is treated as fully correlated across all bins. This results in additional cross-terms, which has some implications for the effect of binning. In the uncorrelated case, adding a new bin with very poor sensitivity can never make the total uncertainty (and therefor the expected limit) worse, due to the construction of the bin weights. However, in our case correlations can actually cause this to happen.

Due to the fact that the w_i are calculated ignoring correlations across bins, the results from this method can differ substantially from a full fit in cases where systematic uncertainties are important. Since we are in such a regime, we cannot trust this approximation to be fully valid. Therefore, we use this method to select a few candidate scenarios for the signal region definition. The full limit-setting formalism (described in Chapter 8) is then used to evaluate and compare these more fully.

The results of this optimization were very similar to the single-bin case. Binning is only done in m_{jj}, as this has by far the best discriminating power between signal and background. Three bins were chosen, as adding more beyond this did not appreciably improve the sensitivity. This was mainly due to the finite statistics of the data and MC; for a much larger data volume it would be advantageous to extend the binning higher in m_{jj}. The results of this procedure were used as the final values for the jet-related cuts. The one exception is $E_{\mathrm{T}}^{\mathrm{miss}}$, which was chosen at a higher value of 180 GeV for reasons related to the multijet background.

6.2.4.3 Third Jet Veto

One of the most powerful discriminants between the VBF-produced signal and QCD-produced background is a veto on additional jets. An optimization similar to the one described above was done to choose a definition for this veto. There are two components involved: the p_{T} threshold over which a jet is counted, and any geometrical requirements on its position relative to the two selected "VBF jets". Two definitions for the latter component were considered: veto only on jets between the two selected jets in η ("Central Jet Veto"), or veto on jets anywhere in the detector ("Third Jet Veto").

The same procedure used to optimize the other jets cuts was applied to the veto definition. It was found that vetoing on any third jet always performed better than allowing jets outside a certain η range. Additionally, the lowest possible p_{T} threshold on these jets (25 GeV) was found to be best. For any looser veto definition, the loss in background rejection always outweighed the gain in signal acceptance.

6.3 Monte Carlo Simulation

In order to relate theoretical calculations to the observed data, Monte Carlo simulation (MC) is used extensively. This consists of several steps. First, a generator is run which assumes some physical theory (often simply the SM) and simulates pp collisions. This is repeated for many events, and the output is a set of final-state particles and their four-momenta for each one. These particles are then passed through a simulation of ATLAS itself (using `GEANT4` [86]), which models the interactions of the particles with the detector. The detector response is modeled to provide output which mimics that of real data (called "digitization"). At this stage, the resulting detector "hits" are overlaid with those from simulated minimum-bias events, in order to model pileup. Finally, the event reconstruction algorithms are run on this, exactly as they are on data. This procedure is illustrated in Figure 6.8.

There are many different generators, but they all work in the same general way. A parton distribution function (PDF) is provided, which describes the distribution of quarks and gluons in the initial-state protons. Then, a matrix-element calculation is performed for all the possible initial states given some specific final state or set of final states. The order of this calculation and the exact way it is carried out depends on the generator. The phase space of possible final states is then sampled, which gives the output particles and their momenta for each event. Finally, a parton shower algorithm is applied to model hadronization and fragmentation of strongly-interacting objects. This cannot be done from first principles, because it is in a non-perturbative regime and would be far too computationally expensive. Instead, parton showers are developed and tuned based on empirical data. The overall process is illustrated in Figure 6.9.

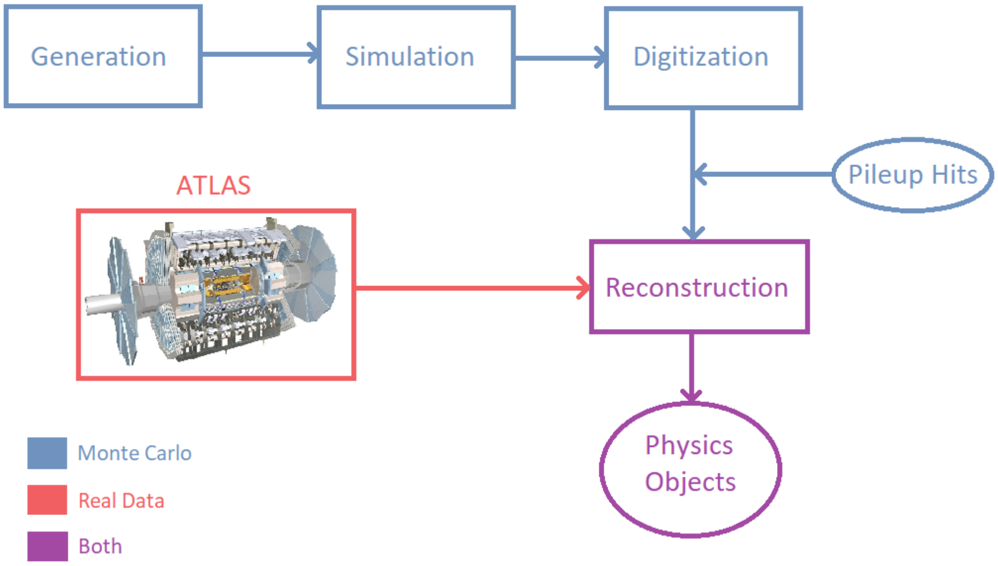

Figure 6.8: Schematic diagram showing the stages of MC generation and reconstruction. After digitization, the simulated events are reconstructed as if they were real data.

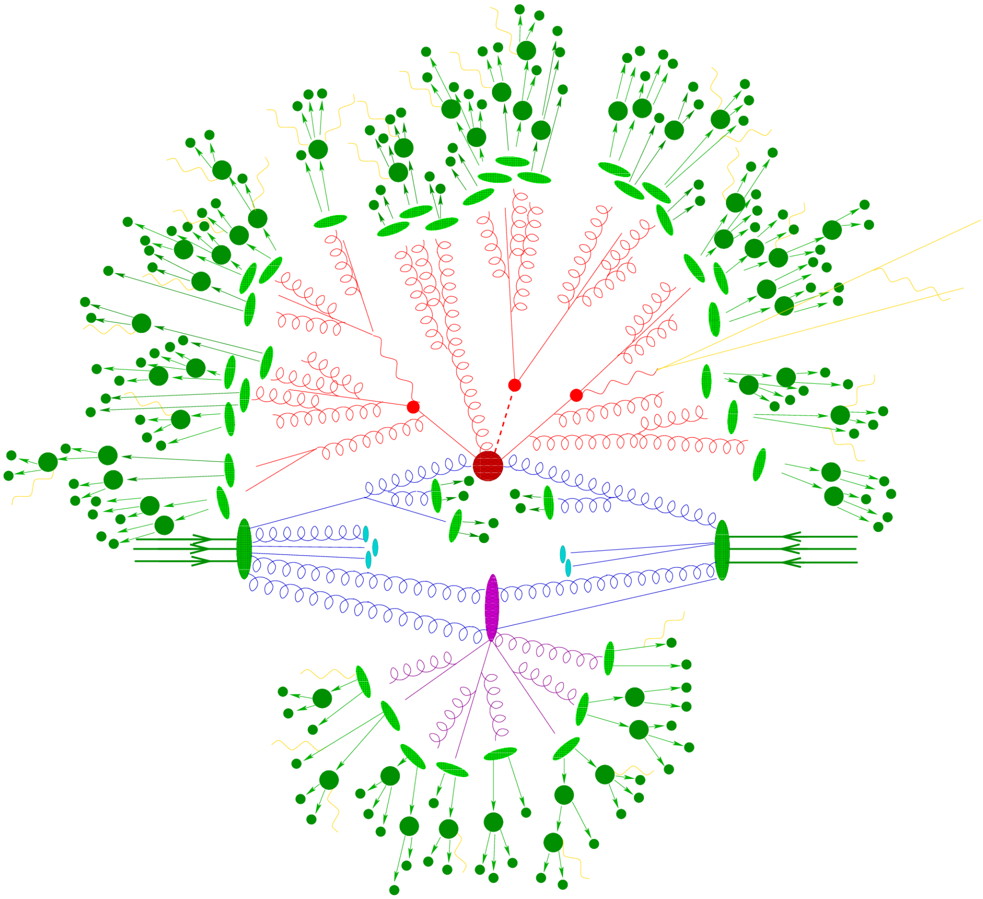

Figure 6.9: Sketch of a *pp* collision as simulated by a MC generator [87]. The red blob in the center represents the hard interaction, and the surrounding red structure represents radiation as simulated by parton showers. The green structures represent hadronization, and the yellow lines show soft photon radiation. The purple structure represents a secondary hard interaction.

6.3.1 Signal Samples

The primary signal model is an invisibly decaying Higgs boson. In the signal regions defined for this analysis, VBF by far the dominant contribution. However, there is also a non-negligible ggF component. Both of these processes are simulated using `POWHEG` [88, 89], interfaced with `PYTHIA` [90, 91] for hadronization and showering. The `NNPDF3.0 NNLO` PDF set [92] is used[42]. The matrix element calculation is done up to NLO in α_S. The factorization and renormalization scales used in this calculation are determined dynamically during the generation of radiation (as is always the case with the `POWHEG` method). They are both set equal to the transverse momentum of the radiated particle.

The invisible decay of the Higgs is simulated by forcing it to decay into four neutrinos (via Z bosons). The kinematic difference between this scenario and decays to a new invisible particle are negligible, since we only consider situations in which the Higgs is produced on-shell (this analysis is not sensitive to off-shell production). The SM branching ratio for the invisible Higgs decay is approximately 1.2×10^{-3}, however this value floats in the final fit to allow for enhancement by non-SM particles. NLO electroweak corrections are also applied to these samples; these are computed separately using `HAWK` [93]. The Higgs production cross sections are taken from [32].

Additional VBF samples with heavier Higgs masses were also generated, for heavy scalar mediator interpretations of the results. The parameters used in the production of these samples are exactly the same as the primary signal sample; the only difference is m_H.

6.3.2 Background Samples

The primary background processes ($Z \to \nu\nu$ and $W \to \ell\nu$), as well as $Z \to \ell\ell$, are modeled using `SHERPA 2.2.1` [87]. These calculations use the `Comix` [94] and `OpenLoops` [95] matrix element generators, and merging is done with the `SHERPA` parton shower [96] using the ME+PS@NLO prescription [97]. The `NNPDF3.0 NNLO` PDF set is used, along with dedicated parton shower tuning developed by the `SHERPA` authors. In all cases, the METS (Matrix Elements and Truncated Showers) method [98, 99] is used to set the factorization, renormalization, and resummation scales dynamically; this is the `SHERPA` default. This same scheme is used to merge final-state jets; the associated scale (also called the CKKW matching scale, from [100]) is fixed to 20 GeV.

The W and Z simulation is split into two components based on the order in the electroweak coupling constant α_{EW}. The "strong-produced" component is of order α_{EW}^2 and the "electroweak"

---

[42] `NNPDF3.0` includes data from ATLAS, CMS, LHCb, and HERA and uses a neural network to extract the PDFs.

component is of order α_{EW}^4. Matrix elements for the strong-produced contribution are calculated at NLO in α_S for up to 2 final-state partons and LO for up to 4 partons. These events are normalized to the NNLO inclusive cross sections, which are computed with FEWZ [101]. A complete description of W and Z event generation for ATLAS Run 2 is given in [102].

The electroweak contribution is further divided into two parts: the diboson diagrams (whch contain two s-channel bosons) and the diagrams which contain t-channel boson propagators (including VBF). The interference between these two diagram types is assumed to be negligible. Matrix elements for the diboson contribution are calculated up to NLO in α_S for up to 3 final-state partons and LO for up to 5 partons. Matrix elements for the t-channel contribution are calculated up to LO in α_S for up to 3 final-state partons. The diboson samples include only the contribution from on-shell bosons; the off-shell component is assumed to be negligible. For both the VBF and diboson samples, the generator cross section is used.

For all $Z \to \ell\ell$ samples, a generator-level filter on the dilepton mass (requiring $m_{ll} > 40$ GeV) is imposed. This serves to remove the contribution from off-shell photons, which is negligible in all of the signal and control regions. The effective integrated luminosity of the primary samples is given in Tables 6.3 and 6.4 for reference. In practice each lepton flavor is generated separately, but they are grouped here for simplicity. The effective luminosity is defined as the number of events generated divided by the cross section for the process being generated (after accounting for any generator-level event weighting). The strong-produced samples are sliced in $\max(p_{\mathrm{T}}^{V}, H_{\mathrm{T}})^{43}$ (at generator level) in order to sufficiently populate the higher-p_{T} region with simulated events. Due to the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut at 180 GeV, the most important slice is 140–280 GeV. Because of detector resolution effects, events from the 70–140 GeV slice can also enter the signal and control regions. The electroweak samples are generated inclusively, since their cross sections are small enough to make this computationally feasible.

Additionally, there is a term of order α_{EW}^3 from interference between the strong and electroweak production modes. This is modeled independently using `MadGraph5_aMC@NLO` [103]. This is calculated at leading order in the QCD coupling (i.e. order α_S^2). As with the primary samples, this simulation uses the `NNPDF3.0 NNLO` PDF set, and showering is done using `PYTHIA`. The contribution from this interference term is found to be small (on the order of 1% of the electroweak term), so it is neglected in the final fit.

There are additional backgrounds from top processes and QCD multijets, but they are very small

---

[43] H_{T} is defined as the scalar sum of the p_{T} of all jets in the event above some threshold (typically on the order of 10 GeV).

$\max(p_{\mathrm{T}}^{V}, H_{\mathrm{T}})$	$Z \to \nu\nu$			$Z \to \ell\ell$			$W \to \ell\nu$		
	LF	CF	BF	LF	CF	BF	LF	CF	BF
0–70 GeV	1.2	6.8	11.3	5.0	22.7	62.9	1.6	4.1	20.9
70–140 GeV	33.6	128	274	79.5	96.0	463	24.0	44.2	104
140–280 GeV	108	245	586	204	325	2045	50.3	77.9	688
280-500 GeV	193	271	1064	420	445	1340	130	133	313
500-1000 GeV	1037			1699			408		
>1000 GeV	6283			6897			3337		

Table 6.3: Effective integrated luminosity for the strong-produced W and Z background MC samples, in fb^{-1}. Below 500 GeV, they are split based on the final-state jets into light flavor (LF), c-filtered (CF), and b-filtered (BF) samples. These flavor filters are orthogonal and complete. The lepton flavors are grouped together here and the smallest number is shown.

$Z \to \nu\nu$	$Z \to \ell\ell$	$W \to \ell\nu$
341	786	147

Table 6.4: Effective integrated luminosity for the VBF-produced W and Z background MC samples, in fb^{-1}. The lepton flavors are grouped together here and the smallest number is shown. These samples are inclusive in jet flavor and boson p_{T}.

relative to the V+jets. The top backgrounds (both pair production and single-top) are simulated using `POWHEG` interfaced with `PYTHIA`, and `EvtGen` to describe the relevant b decays. The QCD multijet background is simulated using `PYTHIA`. However, due to the extremely large cross section for this process, the effective luminosity for these MC samples is substantially smaller than the amount in data. This means a prediction for this background cannot be taken directly from these samples. Instead, data-driven techniques are used to estimate the multijet background in the signal and control regions (see Chapter 7).

6.3.3 Pileup Reweighting

All of the MC samples were generated with a fixed pileup distribution (typically described in terms of the number of interactions per bunch crossing, μ), which is in general different from the distribution the LHC actually provided. Since many event-level quantities are sensitive to this, a pileup reweighting procedure is used to correct the distribution in MC. A set of weights is derived the the data and MC distributions as a function of μ, and then applied event-by-event to MC. In general, different MC samples can be generated with different μ distributions, so the weights can vary from sample to sample. These weights are constructed such the the total normalization of each sample is preserved.

A variation in this reweighting is included to cover the uncertainty on the ratio between the predicted and measured total inelastic cross section in the detector's fiducial volume [104]. This is treated as a systematic uncertainty on the MC predictions. The nominal value of this ratio is 1.09, and it is varied to 1.00 and 1.18 for the purposes of uncertainty calculation.

6.3.4 Scale Factors

Scale factors are applied to reconstructed objects to correct for MC mismodeling of the efficiency to reconstruct/identify them. These are multiplicative weighting factors which are used in addition to pileup reweighting and any generator-level event weights. In general, four types of these are applied to any given MC event:

- Trigger efficiency scale factors (for all triggers used, including $E_{\mathrm{T}}^{\mathrm{miss}}$).
- Reconstruction efficiency scale factors (for electrons and muons).
- Identification efficiency scale factors (for electrons only[44]).
- Isolation efficiency scale factors (for electrons and muons).

In this case, a *veto* is applied to leptons. The efficiency of this veto is just as easily mismodeled, but there are no reconstructed leptons to which to apply scale factors. So, "Anti-ID scale factors" are also defined and applied to MC events. Specifically, they are applied to events which contain leptons within acceptance at truth-level, but no fully identified leptons. These are derived from the efficiencies and scale factors for normally identified leptons:

$$\mathrm{SF}_{\text{anti-ID}} = \prod_i \frac{1 - \mathrm{SF}^i \cdot \epsilon^i}{1 - \epsilon^i} \tag{6.10}$$

where i indexes the truth-level leptons in the event within acceptance. The efficiencies and ID scale factors are taken as functions of the truth lepton's $|\eta|$ and p_{T}.

The efficiencies ϵ are derived from MC by running over the full set of samples and considering all the truth leptons within acceptance (i.e. $|\eta| < 2.4$ for electrons and $|\eta| < 2.7$ for muons. Each truth lepton is considered successfully reconstructed if there is a matching reconstructed lepton within $\Delta R < 0.2$ of it.

---

[44] "Reconstruction" and "identification" are treated as two separate steps for electrons, but only one for muons. So, the muons have one scale factor to cover both.

6.4 Experimental Uncertainties

There are three broad classes of uncertainty involved in this search (and in collider analyses in general): statistical, experimental, and theoretical. This section discusses experimental uncertainties which apply to all aspects of the analysis. Theoretical uncertainties are discussed in detail throughout Chapter 7. Like the theoretical uncertainties, the experimental ones are implemented in the overall fit as one nuisance parameter per independent uncertainty source[45]. Each of these is associated with some variation of the MC predictions for the yields in all regions, which is computed by changing the quantity of interest (e.g. the electron ID efficiency) by 1σ. Here, the nuisance parameters are listed, and each one is given a name (which will identify it later in plots and tables).

In addition to the object-specific uncertainties listed below, there are two additional experimental systematics. The first is the uncertainty on the total integrated luminosity of the dataset, which amounts to an overall normalization. This scales everything by 2.1% (the overall uncertainty on the combined 2015 and 2016 integrated luminosity), and its nuisance parameter is simply denoted **Lumi**. The second is the uncertainty on the "pileup reweighting data scale factor", denoted **PRW_DATASF**. This is effectively the uncertainty on the number of interactions per bunch crossing in data, which comes into play during the pileup reweighting procedure[46]. It is equivalent to the uncertainty on the total inelastic cross-section in the detector's fiducial volume.

6.4.1 Jets

The jet calibration procedure consists of several steps, as described in Section 4.1. Each one of these has several individual sources of uncertainty, originating from measurements used from data or MC simulation. In total, there are 88 nuisance parameters associated with these[47], 75 of which come from the residual *in-situ* calibration. In many cases, this fully-detailed description is unnecessary, so a "category-reduced" set of 29 can be used. This is created by combining subsets of the 75 originating from the *in-situ* step, based on their source (statistical, detector, etc.) to reduce them to 16. Additionally a "strongly-reduced" set of only 4 NPs can be used, by performing a decomposition of the full set. In this case, the loss of information from reducing to 4 NPs is non-negligible, so the category-reduced set of 29 is used. In addition to these energy scale uncertainties, there is an additional uncertainty on the jet energy resolution (JER).

The full set of jet-related nuisance parameters is then:

---

[45] See Section 8.1 for a full description of this procedure.

[46] The pileup reweighting procedure is described in Section 6.3.3.

[47] See [71] for the complete list with details.

- **JET_EffectiveNP_[...]:** The resulting 16 components of the JES uncertainty decomposition.
- **JET_Pileup_[...]:** The four uncertainties associated with the jet pileup correction.
- **JET_EtaIntercalibration_[...]:** The three uncertainties associated with the η-intercalibration component of the JES[48].
- **JET_Flavor_[...]:** The three uncertainties associated with flavor composition and response.
- **JET_PunchThrough_MC15:** The uncertainty in the GSC punch-through correction.
- **JET_SingleParticle_HighPt:** The high-p_{T} uncertainty from single-particle and test-beam measurements.
- **JET_JvtEfficiency:** The uncertainty on the efficiency of the JVT requirement.
- **JET_JER_SINGLE_NP:** A single uncertainty parametrizing knowledge of the jet energy resolution.

In all cases, these are p_{T}- and η-dependent factors by which the jet energy is scaled. Some (such as the flavor-specific uncertainties) can depend on other variables as well. Since this analysis deals mainly with light-flavor jets that are far below the TeV scale, the flavor, punch-through, and single-particle uncertainties are very small.

6.4.2 Electrons

The electron reconstruction and identification procedure (described in Section 4.2) comes with substantially smaller set of systematic uncertainties than the jets do, since they're much simpler objects. Therefore, the complete set of nuisance parameters is used. This includes:

- **EG_SCALE_ALL:** The uncertainty on the electron energy scale.
- **EG_RESOLUTION_ALL:** The uncertainty on the electron energy resolution.
- **EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR:** The uncertainty on the electron reconstruction efficiency scale factors.
- **EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR:** The uncertainty on the electron identification efficiency scale factors.

---

[48] This is described in full detail in [71].

- **EL_EFF_ISO:** The uncertainty on the electron isolation efficiency scale factors.
- **EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR:** The uncertainty on the electron trigger efficiency scale factors.

6.4.3 Muons

As with the electrons, all nuisance parameters associated with muon reconstruction and identification are included in the fit. These are all corrections to the muon momentum, except for the ones which are specified as being efficiency uncertainties. The full set consists of:

- **MUON_SCALE:** The uncertainty on the muon momentum scale.
- **MUON_ID:** The uncertainty on the muon identification efficiency.
- **MUON_EFF_SYS:** The systematic component of the uncertainty on the muon identification efficiency scale factors.
- **MUON_EFF_STAT:** The statistical component of the uncertainty on the muon identification efficiency scale factors.
- **MUON_MS:** The uncertainty associated with the original muon spectrometer track for the muon.
- **MUON_SAGITTA_RHO:** The uncertainty associated with the track correction applied during reconstruction.
- **MUON_SAGITTA_RESBIAS:** The uncertainty originating from residual charge-dependent bias after track correction.
- **MUON_TTVA_SYS:** The systematic component of the track-to-vertex association uncertainties.
- **MUON_TTVA_STAT:** The statistical component of the track-to-vertex association uncertainties.

6.4.4 Missing Transverse Momentum

Since $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction uses the jet, electron, and muon objects, all of the uncertainties on those must be propagated through to the final $E_{\mathrm{T}}^{\mathrm{miss}}$ object. The soft (track-based) term is not included in any of this, so its systematic uncertainties are derived separately in addition to the rest. These are:

- **MET_SoftTrk_Scale:** The uncertainty on the energy scale of the track soft term.
- **MET_SoftTrk_ResoPara:** The uncertainty on the energy resolution of the track soft term (component parallel to the p_{T} of the recoil system).
- **MET_SoftTrk_ResoPerp:** The uncertainty on the energy resolution of the track soft term (component perpendicular to the p_{T} of the recoil system).

Chapter 7

Signal and Background Modeling Techniques

This chapter discusses the complete details of methods used to model the signal processes and estimate backgrounds in the signal and control regions. Background modeling in particular is one of the most challenging aspects of the analysis, and constitutes the bulk of the actual work involved with this search. In general, MC simulation can be used to model all the relevant processes, but this alone is generally not sufficient to provide estimates with the necessary precision. The exact reasons for this depend on the process in question, but in general data-driven methods are employed alongside MC to more tightly constrain the backgrounds. In the case of the signal, there is no option but to rely purely on MC (as there is no other data sample which can be used to constrain it). This is not a problem; this simulation is NLO in α_s and has small associated uncertainties and the kinematics are known to be well-modeled from other Higgs analyses. The details of signal modeling are described in Section 7.1.

For the W and Z backgrounds, data in the control regions is used to constrain these estimates and reduce the impact of these uncertainties. Although the MC is NLO in α_s, the theoretical uncertainties associated with the strong production in particlar would very large and ruin the sensitivity of this search. This is less problematic for the electroweak production; that MC for that is only LO in α_s but there are no QCD interactions involved to first order. In addition to large uncertainties, the existing generators are not entirely trusted to accurately reproduce all kinematic distributions. In particular, some mismodeling of the m_{jj} shape has been observed in past analyses; normalizing each bin to data removes this effect. The techniques used to understand these backgrounds are discussed in Section 7.2.

In the case of the multijet background (both in the signal and control regions), the primary problem with using pure MC is a lack of statistics. Due to the enormous cross-section for the multijet process, it is impossible to simulate enough events to accurately model the data. Fully data-driven methods are therefore used: a “rebalance-and-smear” technique for the signal region and a template fit method for the control regions, discussed in Sections 7.3 and 7.4 respectively.

The remaining backgrounds (mainly from top processes) are so small that a MC-only estimate is easily sufficient. Any uncertainties the modeling would be negligible relative to the size of the total background.

7.1 Higgs → Invisible Signal

The primary signal model (Higgs → invisible) is modeled using `POWHEG`, as described in Section 6.3. The uncertainties on the signal modeling from the generator are small enough that they do not dominate the sensitivity of the analysis. However, they are not negligible, so they are calculated explicitly and described below. Signal yields for both the VBF and ggF modes are taken directly from the Monte Carlo. The only modification is the NLO electroweak correction to the VBF process, which is calculated separately and applied as a correction.

7.1.1 Electroweak Corrections to VBF Higgs Production

The signal MC was generated at leading order in the electroweak coupling, but the NLO corrections can also be accounted for. So, these corrections are calculated explicitly using `HAWK` [93]. They are computed differentially as a function of the Higgs p_{T}, specifically in the phase space used in this analysis. The results are shown in Figure 7.1.

This correction is applied event-by-event to the VBF signal MC, depending on the p_{T} of the Higgs in that event. The final impact on the signal yield is roughly 0.2%.

7.1.2 Theoretical Uncertainties

PDF uncertainties on the signal are computed using the `NNPDF 3.0 NNLO` PDF set, which contains 100 different PDF replicas. Each of these is produced from a different pseudo-dataset generated from the original `NNDPF 3.0` input dataset. These are generated in a way which preserves the statistical properties and correlations in the dataset [105]. The signal yield is computed for each of these replicas, and the standard deviation of these 100 yields is then taken as the uncertainty. The results

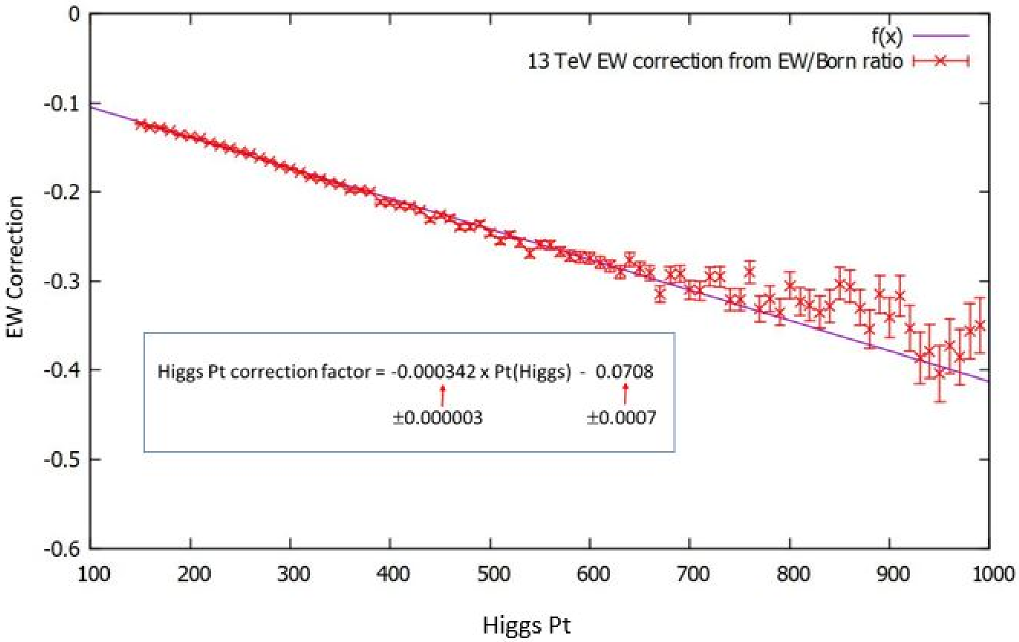

Figure 7.1: NLO electroweak corrections to the VBF signal production, in % of the nominal yield. A linear fit is also shown.

Bin	PDF Uncertainty
1.0 TeV $< m_{jj} <$ 1.5 TeV	0.64 %
1.5 TeV $< m_{jj} <$ 2.0 TeV	1.02 %
$m_{jj} >$ 2.0 TeV	1.53 %

Table 7.1: PDF uncertainties on the VBF signal yield. These are given relative to the total expected yield.

for this procedure are shown in Table 7.1 for the VBF signal. For the ggF component of the signal, the PDF uncertainty is neglected since its impact on the analysis is extremely small.

There are also scale uncertainties on the signal. These arise from the fact that the MC calculation is fixed-order, necessitating choices for the factorization and renormalization scales (μ_F and μ_R). μ_F is the scale at which the matrix element calculation is factorized from the PDF, and μ_R is the scale at which the strong interaction is renormalized. If the MC calculation could be done to infinite precision, the result would be independent of these choices. So, varying these scales and re-computing these results is used as a rough estimate of the scale of the missing higher-order terms.

For the VBF production mode, VBF Higgs events are generated using MCFM [106, 107] (the Higgs p_{T} is taken as the $E_{\mathrm{T}}^{\mathrm{miss}}$). μ_F and μ_R are each varied up and down by a factor of 2, and they are also coherently varied by a factor of 2. The envelope of these 6 variations is taken as the overall scale uncertainty on the VBF signal. This procedure results in a 7.6% uncertainty for the inclusive signal region.

For the ggF component of the signal, the scale uncertainties are estimated using the Stewart-Tackmann procedure [108]. This is necessary here because simple scale variations in exclusive fixed-order calculations such as this one generally underestimate the actual size of the missing higher-order terms, due to cancellation which arise to due the requirement of *exactly* two jets. As a result, this method typically gives much larger uncertainties than simply varying the scales for the ggF+2 jets process. As with the VBF component of the signal, events are generated using MCFM using nominal scales and a 6-point variation set. Then, the procedure consists of evaluating the uncertainty associated with the third jet veto, by subtracting the 3-jet cross section from the inclusive 2-jet cross section. Then, the uncertainties associated with the other cuts are evaluated and added in quadrature. This calculation results in a 68% overall uncertainty on the ggF signal in the inclusive SR. This is a conservative estimate, but since ggF makes up only a small fraction of the signal, it does not substantially harm the sensitivity of the analysis.

Finally, there are uncertainties associated with the modeling of the underlying event and parton shower. To estimate the size of these effects, the same signal samples were alternately showered with `Herwig 7` [109, 110] instead of `PYTHIA`. This was compared at truth level to the nominal sample, and the signal efficiency was found to decrease by 7.7%. Additionally, the parameters used for the `PYTHIA` showering were varied. The effects of any such variations were found to be negligible compared to the difference between `Herwig` and `PYTHIA`. So, the 7.7% difference is taken as the overall parton shower uncertainty on the signal.

7.2 W and Z Backgrounds

The W and Z backgrounds are modeled using the `SHERPA` generator[49]. However, taking the raw MC prediction as the background estimate would result in very large uncertainties. Additionally, the most recent versions of `SHERPA` are known to mismodel the jet kinematics in the VBF phase space. The $W \to \ell\nu$ and $Z \to \ell\ell$ control regions are used to correct for this effect, for the W and Z backgrounds respectively. Since the m_{jj} distribution in MC is known to be inaccurate, these process are considered to be uncorrelated across the three m_{jj} bins. That is, no assumption is made about the shape of the m_{jj} distribution of the background.

---

[49] See Section 6.3 for further details on the MC simulation.

It is, in general, possible to take advantage of the correlation between W and Z processes as well. Since they are associated with the same gauge group and have the same Feynman diagrams at leading order, they should be strongly correlated. The Run 1 ATLAS result used this argument to justify constraining them to the same normalization [5]. This is equivalent to the assumption that the W and Z processes are 100% correlated with each other, and that the MC mismodels both of them by the same amount.

In reality, the W and Z production are very similar, but not quite exactly the same. Differences in their couplings to the various quark flavors, combined with the proton PDF, can introduce small discrepancies. Although higher-order QCD corrections are identical for both, this is not the case for electroweak corrections (for example, the SM has a WWZ vertex but not a ZZW vertex). These effects have been studied in detail (for monojet-like searches) in the inclusive phase space [111], and the non-correlation effects were found to have roughly a 1-2% impact on the event yields (except at very high p_{T}, where the effect is larger). However, the effect could be larger or smaller in the VBF-like region used in this search. This result was obtained by comparing higher order k-factors[50] for the W and Z processes.

In order to correctly account for this and take advantage of the correlation, it is necessary to compute the differences in the kinematic region used in the analysis. This was attempted, but the available MC samples did not have adequate statistics to compute the k-factors with sufficient precision. This can still be done, but it would require a concerted effort on a timescale beyond that of this analysis (probably with the assistance of theorists). For the present, a more conservative approach is used: no assumption is made on the correlation between W and Z production. Each one is assigned a separate normalization in the overall fit, and they are allowed to float independently. Denoting our Z background estimate in the SR are Z_{SR} (and likewise for W_{SR}), this can be thought of conceptually as:

$$Z_{\mathrm{SR}} \sim \overbrace{\frac{Z_{\mathrm{SR}}^{\mathrm{MC}}}{Z_{\mathrm{CR}}^{\mathrm{MC}}}}^{\alpha_Z} \times Z_{\mathrm{CR}}^{\mathrm{data}} = \overbrace{\frac{Z_{\mathrm{CR}}^{\mathrm{data}}}{Z_{\mathrm{CR}}^{\mathrm{MC}}}}^{k_Z} \times Z_{\mathrm{SR}}^{\mathrm{MC}} \tag{7.1}$$

$$W_{\mathrm{SR}} \sim \underbrace{\frac{W_{\mathrm{SR}}^{\mathrm{MC}}}{W_{\mathrm{CR}}^{\mathrm{MC}}}}_{\alpha_W} \times W_{\mathrm{CR}}^{\mathrm{data}} = \underbrace{\frac{W_{\mathrm{CR}}^{\mathrm{data}}}{W_{\mathrm{CR}}^{\mathrm{MC}}}}_{k_W} \times W_{\mathrm{SR}}^{\mathrm{MC}} \tag{7.2}$$

where W and Z here denote the respective yields from data or MC in the indicated region. In

---

[50] A "k-factor" is defined as the ratio of a fixed-order cross section to that obtained from a higher-order calculation.

practice the background estimation procedure is performed as a simulatneous fit of all k parameters in all regions. There are six of these (one for W and Z in each of the 3 m_{jj} bins). This is known as the "Transfer Factor" method. The transfer factors themselves (denoted α_Z and α_W) are defined simply as $\alpha_Z = Z_{\mathrm{SR}}^{\mathrm{MC}}/Z_{\mathrm{CR}}^{\mathrm{MC}}$. The background estimates can be equivalently thought of either as a MC-derived transfer from CR to SR or a data/MC correction to the MC prediction in the SR:

$$Z_{\mathrm{SR}} \sim \alpha_Z \times Z_{\mathrm{CR}}^{\mathrm{data}} = k_Z \times Z_{\mathrm{SR}}^{\mathrm{MC}} \tag{7.3}$$

$$W_{\mathrm{SR}} \sim \alpha_W \times W_{\mathrm{CR}}^{\mathrm{data}} = k_W \times W_{\mathrm{SR}}^{\mathrm{MC}} \tag{7.4}$$

The main benefit of this method is that it allows many of the systematic uncertainties on the background to "cancel out". That is, any systematic bias in the estimation of W and Z processes should affect the signal and control regions in almost exactly the same way. So, for example, the total uncertainty on $Z_{\mathrm{SR}}^{\mathrm{MC}}$ and $Z_{\mathrm{CR}}^{\mathrm{MC}}$ could be large, but the uncertainty on their ratio (and therefore the uncertainty on the actual background estimate) can be much smaller. However, this cancellation is not perfect. Acceptance effects can introduce differences between SR and CR (which can be significant for theoretical uncertainties), and some experimental systematics, such as the lepton identification efficiencies, affect the signal and control regions differently. Additionally, photon radiation from the leptons can introduce small differences between the $Z \to \ell\ell$ and $Z \to \nu\nu$ regions.

7.2.1 Data and MC in the Control Samples

Comparisons of data and MC simulation in the various control regions can be used to check the modeling of the W and Z processes. This is shown in Figures 7.2 and 7.3 for $Z \to \ell\ell$, and in Figures 7.4 and 7.5 for $W \to \ell\nu$. These plots do not include the systematic uncertainties on the MC predictions (which are on the order of 25%). These are all pre-fit distributions, and so the multijet (i.e. fake electron) background in the $W \to e\nu$ region is not included, as no pre-fit estimate is defined[51]. The fact that there is no significant excess in the absence of a fake electron absence indicates that this contribution is quite small.

Even before the fit, the agreement between data and MC is generally good. This indicates that the normalization factors for these backgrounds in the fit should be close to unity. It also validates the background simulation; if it models W and Z well in the control regions then it should also model them well in the signal regions, since they are nearly identical kinematically.

---

[51] See Section 7.4 for the methodology and results of the fake electron estimate.

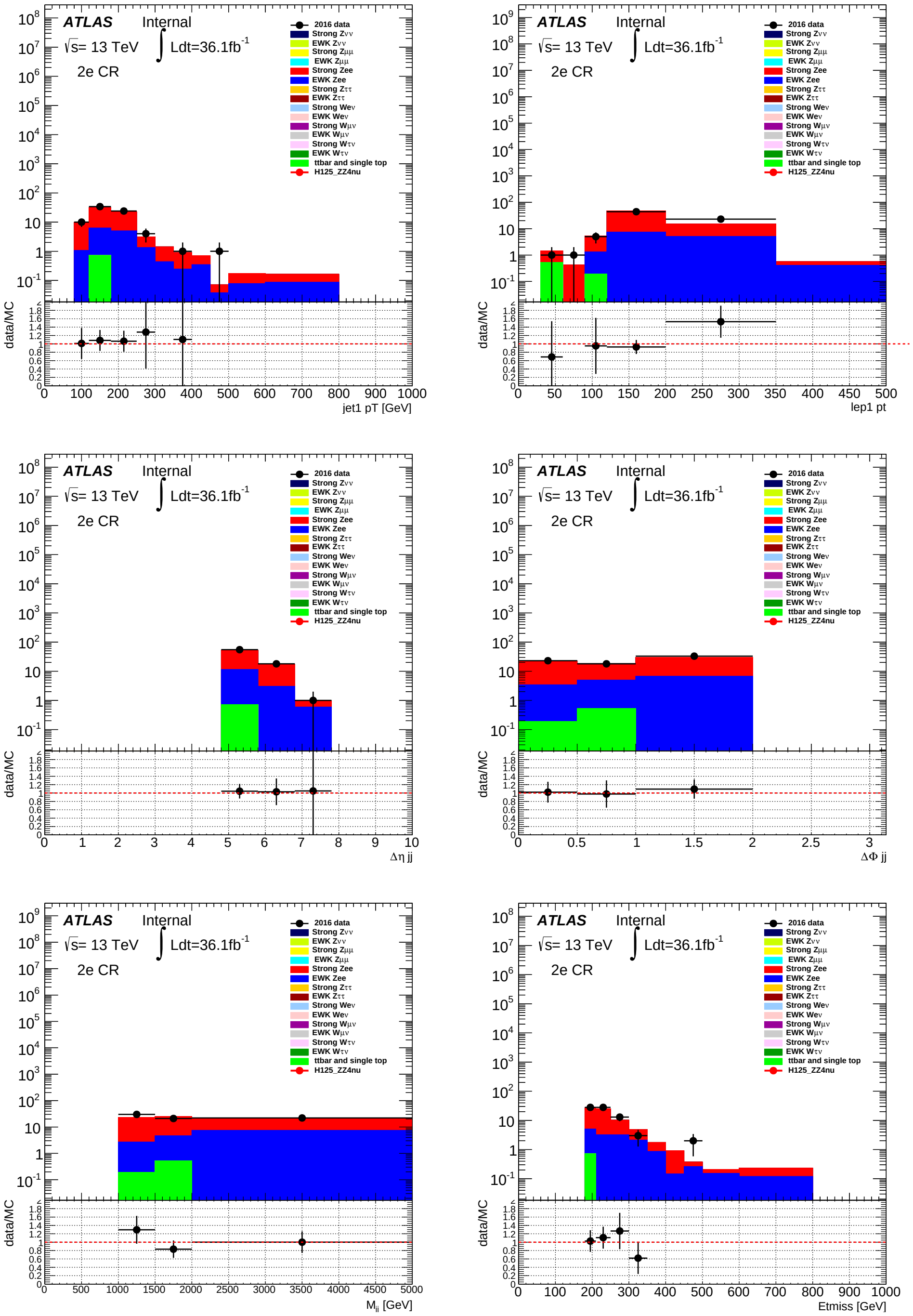

Figure 7.2: Pre-fit kinematic distributions for data and Monte Carlo in the inclusive $Z \to ee$ control region. Error bars include only the statistical uncertainty from the data sample.

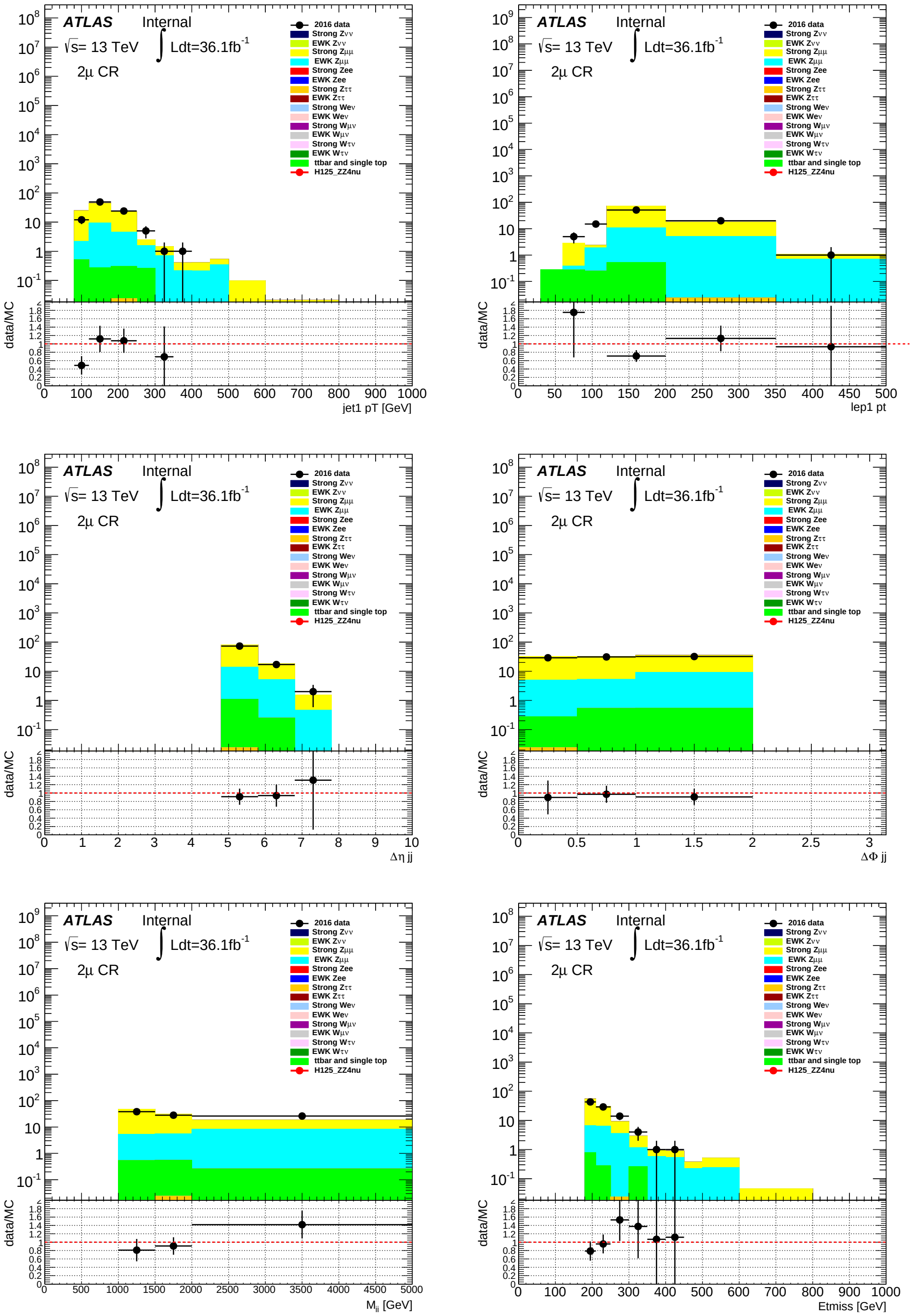

Figure 7.3: Pre-fit kinematic distributions for data and Monte Carlo in the inclusive $Z \rightarrow \mu\mu$ control region. Error bars include only the statistical uncertainty from the data sample.

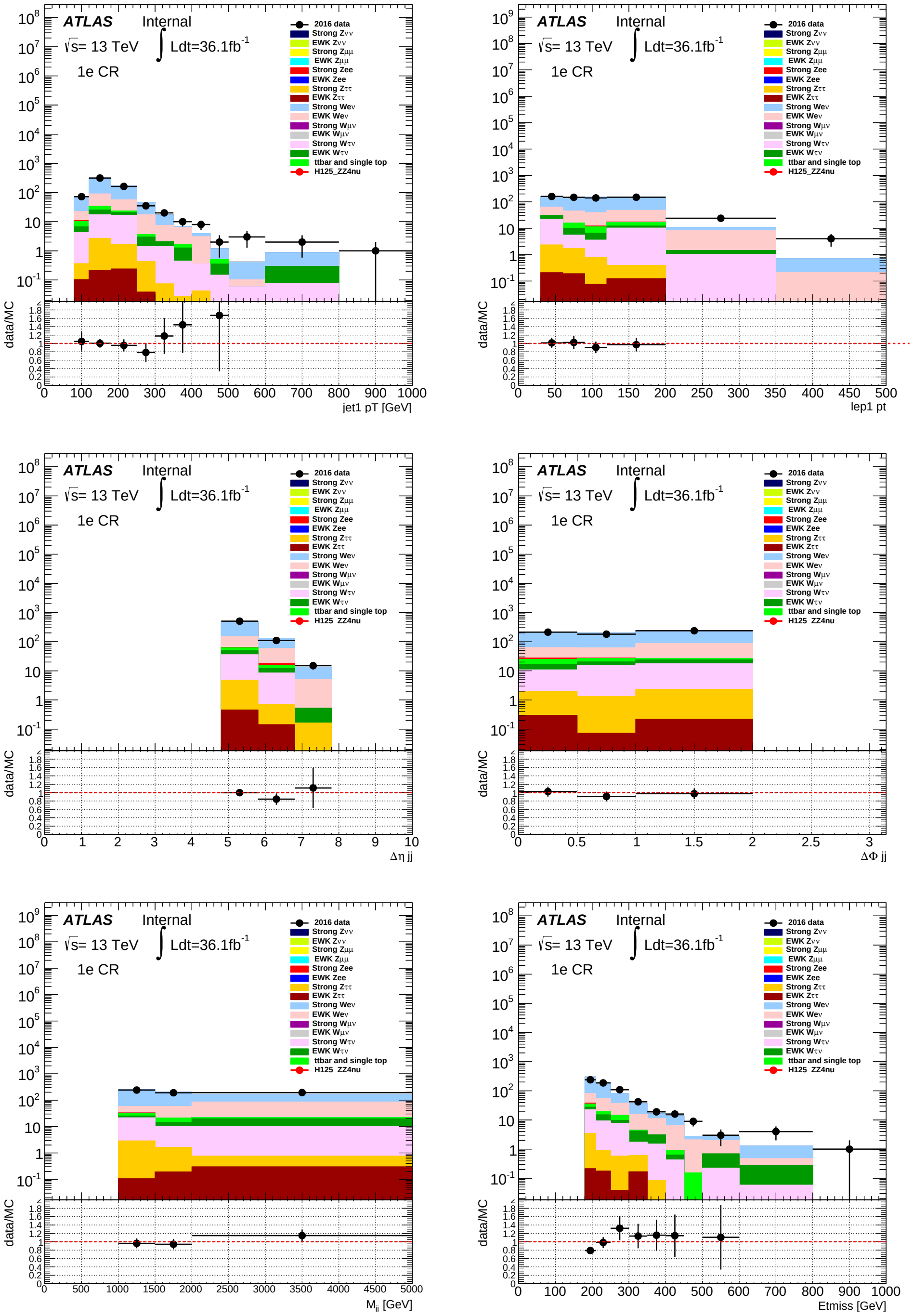

Figure 7.4: Pre-fit kinematic distributions for data and Monte Carlo in the inclusive $W \to e\nu$ control region. Error bars include only the statistical uncertainty from the data sample. The multijet background estimate is not shown.

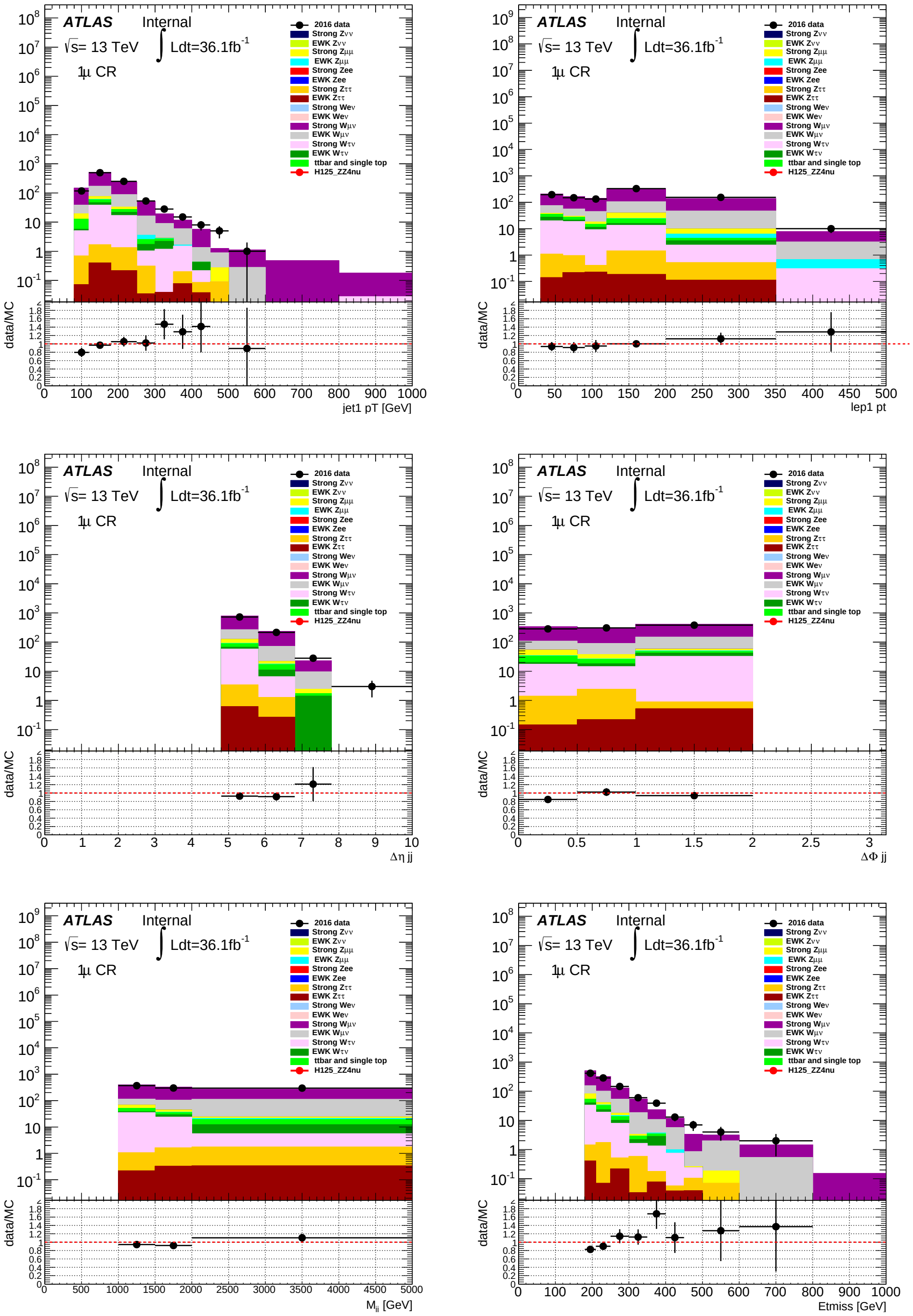

Figure 7.5: Pre-fit kinematic distributions for data and Monte Carlo in the inclusive $W \to \mu\nu$ control region. Error bars include only the statistical uncertainty from the data sample.

7.2.2 Theoretical Uncertainties

Due to the finite precision of our theoretical calculations in the Monte Carlo, there are several sources of uncertainty on the predicted background (and signal) yields. These sources include the finite order of the calculation, the precision to which the initial-state PDF is known, and the limited description of parton showering and hadronization.

For the W and Z background MC, there are four scales which are fixed during event generation: the renormalization, factorization, resummation (sometimes denoted “qsf”), and CKKW matching scales[52]. To estimate the uncertainties associated with these choices, events are generated with each of these scales varied up or down by a factor of two. The CKKW matching scale is an exception; its nominal values is 20 GeV and is varied to 15 and 30 GeV.

For the factorization and renormalization scales (μ_F and μ_R), these uncertainties can be calculated using on-the-fly varied event weights in the reconstructed `SHERPA` MC. This results in full statistical correlation between the nominal and varied samples. In practice, this procedure was carried out at truth level for the strong-produced samples since a much larger quantity of truth-only MC was available. For these, an envelope of a 7-point variation set is used (nominal, μ_F varied up/down, μ_R varied up/down, and both coherently varied up/down). The coherent variations of both tend to be the largest and dominate the envelope in most cases. The results of this procedure are shown in Tables 7.2 and 7.3.

For the resummation and CKKW matching scales, separate varied MC samples need to be generated. Due to computing resource limitations, these cannot be reconstructed using the ATLAS detector simulation. So, the effect of the variations is determined at truth-level and used to define a reweighting which is then applied to the fully-reconstructed nominal sample. These variations are symmetrized, meaning the effect of the variation is taken to be the difference between the up/down variation and the mean of the two variations, rather than comparing directly to the nominal yield. Since the varied samples are statistically uncorrelated with each other and limited in size, there are substantial statistical uncertainties on the size of the variations. To mitigate this effect, the variations were evaluated as a function of m_{jj} and extrapolated from low values into the SRs and CRs using a linear fit. This linear fit was found to describe the variations well, and the results of this method were found to be consistent with simply evaluating the uncertainties in the regions themselves. The statistical uncertainties are then derived from the uncertainties on the fit parameters. This extrapolation method was only used for the strong-production samples, since they

---

[52] See Section 6.3 and the references therein for details on how these scales are set.

were the most affected by statistical fluctuations. The results, and their statistical uncertainties, are shown in Tables 7.4 and 7.5. In many cases, the uncertainty can be as large as the central value for the variation (or larger); the larger of the two is taken as the final uncertainty in order to be conservative.

For the PDF uncertainties, the background yield predictions are evaluated with a full ensemble of 100 PDFs within the `NNPDF` set, as is done with the signal MC. The standard deviation of these yields is taken to be the corresponding PDF uncertainty. These uncertainties are evaluated separately in each signal and control region. Since the full ensemble of PDF weights is stored for each event in the main MC samples, this procedure can be carried out directly on fully reconstruced simulation. The results are shown in Table 7.6.

Selection	Process	1 TeV $< m_{jj} <$ 1.5 TeV	1.5 TeV $< m_{jj} <$ 2 TeV	$m_{jj} >$ 2 TeV
0-lepton (MET)	Strong Z	-19.7%	-20.8%	-21.4%
	Strong W	-20.3%	-23.5%	-21.7%
	EWK Z	-16.4%	-15.5%	-20.6%
	EWK W	- 7.9%	-13.1%	-15.7%
1-lepton ($W \to \ell\nu$)	Strong W	-18.9%	-20.0%	-20.4%
	EWK W	-14.7%	- 8.9%	-14.0%
2-lepton ($Z \to \ell\ell$)	Strong Z	-19.0%	-18.9%	-20.4%
	EWK Z	-17.2%	-15.5%	-18.1%

Table 7.2: Relative changes of the event yields upon applying the “up” variation for the factorization/renormalization scale envelope.

Selection	Process	1 TeV $< m_{jj} <$ 1.5 TeV	1.5 TeV $< m_{jj} <$ 2 TeV	$m_{jj} >$ 2 TeV
0-lepton (MET)	Strong Z	33.2%	35.3%	37.4%
	Strong W	33.7%	41.2%	38.1%
	EWK Z	7.0%	11.2%	12.7%
	EWK W	25.8%	9.8%	16.1%
1-lepton ($W \to \ell\nu$)	Strong W	30.1%	33.7%	34.6%
	EWK W	8.4%	15.7%	25.1%
2-lepton ($Z \to \ell\ell$)	Strong Z	31.5%	31.3%	34.4%
	EWK Z	6.5%	10.6%	15.3%

Table 7.3: Relative changes of the event yields upon applying the “down” variation for the factorization/renormalization scale envelope.

Selection	Process	1 TeV < m_{jj} < 1.5 TeV	1.5 TeV < m_{jj} < 2 TeV	m_{jj} > 2 TeV
0-lepton (MET)	Strong Z	2.6 ± 2.4 %	4.2 ± 3.1 %	6.6 ± 4.0 %
	Strong W	3.0 ± 3.7 %	5.7 ± 4.6 %	9.9 ± 6.0 %
	EWK Z	9.8 ± 3.0 %	7.7 ± 3.0 %	8.7 ± 2.7 %
	EWK W	14.8 ± 3.1 %	12.8 ± 3.0 %	12.3 ± 2.6 %
1-lepton ($W \to \ell\nu$)	Strong W	0.1 ± 1.9 %	1.9 ± 2.3 %	3.5 ± 3.1 %
	EWK W	17.8 ± 3.5 %	14.6 ± 3.3 %	10.3 ± 2.8 %
2-lepton ($Z \to \ell\ell$)	Strong Z	-0.6 ± 2.9 %	0.0 ± 3.6 %	1.0 ± 4.8 %
	EWK Z	4.1 ± 8.3 %	20.1 ± 9.9 %	1.8 ± 7.0 %

Table 7.4: Relative changes of the event yields upon varying the resummation scale by a factor of 2. These are symmetrized; only the "down" variation is shown.

Selection	Process	1 TeV < m_{jj} < 1.5 TeV	1.5 TeV < m_{jj} < 2 TeV	m_{jj} > 2 TeV
0-lepton (MET)	Strong Z	0.0 ± 2.0 %	1.2 ± 2.5 %	3.0 ± 3.3 %
	Strong W	6.8 ± 3.5 %	11.1 ± 4.4 %	17.5 ± 5.7 %
	EWK Z	5.7 ± 3.2 %	1.2 ± 3.1 %	3.0 ± 2.7 %
	EWK W	8.5 ± 3.1 %	-0.4 ± 3.2 %	-0.3 ± 2.7 %
1-lepton ($W \to \ell\nu$)	Strong W	2.1 ± 1.9 %	3.9 ± 2.4 %	6.6 ± 3.1 %
	EWK W	11.4 ± 3.5 %	-1.6 ± 3.6 %	-1.1 ± 3.0 %
2-lepton ($Z \to \ell\ell$)	Strong Z	-0.2 ± 2.6 %	0.0 ± 3.3 %	0.6 ± 4.3 %
	EWK Z	11.9 ± 9.1 %	-8.5 ± 8.5 %	6.1 ± 7.9 %

Table 7.5: Relative changes of the event yields upon varying the CKKW matching scale between 15 GeV and 30 GeV. These are symmetrized; only the "up" variation is shown.

Selection	Process	1 TeV < m_{jj} < 1.5 TeV	1.5 TeV < m_{jj} < 2 TeV	m_{jj} > 2 TeV
0-lepton (MET)	Strong Z	1.4%	2.5%	3.0%
	Strong W	2.3%	1.5%	3.1%
	EWK Z	1.7%	1.9%	3.8%
	EWK W	2.5%	2.0%	2.4%
1-lepton ($W \to \ell\nu$)	Strong W	1.0%	2.1%	2.1%
	EWK W	1.7%	1.5%	4.6%
2-lepton ($Z \to \ell\ell$)	Strong Z	2.0%	1.8%	1.7%
	EWK Z	2.9%	2.2%	3.8%

Table 7.6: Relative changes of the event yields upon applying the PDF variations described in the text.

7.2.3 The Double Ratio Method

In addition to the Transfer Factor strategy, another method known as the "Double Ratio" was considered. This involves introducing additional, less VBF-like control regions. In this case, the 2-jet VRs[53] would simply be used for this purpose. Then, rather than using a simple ratio of SR to CR yields in MC, a more complex double ratio is used. In its simplest form, it can be written as:

$$Z_{\text{SR}} \sim \frac{Z^{\text{data}}_{\text{2J,MET}}/Z^{\text{MC}}_{\text{2J,MET}}}{Z^{\text{data}}_{\text{2J},2\ell}/Z^{\text{MC}}_{\text{2J},2\ell}} \times \frac{Z^{\text{MC}}_{\text{SR}}}{Z^{\text{MC}}_{\text{CR}}} \times Z^{\text{data}}_{\text{CR}} \tag{7.5}$$

$$W_{\text{SR}} \sim \frac{W^{\text{data}}_{\text{2J,MET}}/W^{\text{MC}}_{\text{2J,MET}}}{W^{\text{data}}_{\text{2J},1\ell}/W^{\text{MC}}_{\text{2J},1\ell}} \times \frac{W^{\text{MC}}_{\text{SR}}}{W^{\text{MC}}_{\text{CR}}} \times W^{\text{data}}_{\text{CR}} \tag{7.6}$$

where "2J" indicates these new control regions. As with the Transfer Factor method, this would in practice be implemented in a simultaneous fit to all regions.

In reality, the 2-jet CR with 0 leptons actually contains a mixture of $Z \to \nu\nu$ and $W \to \ell\nu$ where the lepton is lost. So, it isn't possible to get a direct measurement of $Z^{\text{data}}_{\text{2J,MET}}$ or $W^{\text{data}}_{\text{2J,MET}}$, but only their sum. The lost lepton (i.e. W) contribution here can be estimated itself via a Transfer Factor method. This complicates the mathematics slightly, but conceptually the fit works the same way. The precise form of the final fit depends on the assumptions made (for example, whether W and Z are treated as correlated, etc.).

There are two advantages of the Double Ratio method over the Transfer Factor. The first is that it makes a weaker assumption about the MC modeling. The MC need not even mismodel the W and Z yields by the same relative amount; only the ratio between the 2-jet and VBF-like regions of this ratio. This can be thought of as a ratio between W and Z of the efficiency of the selection going from the 2-jet to the VBF-like selection[54]. The second advantage is that better cancellation of the systematic uncertainties can take place in the Double Ratio method. This is because uncertainties that are different between the VBF-like signal and control regions are also different (by a similar amount, in principle) between the 2-jet regions with and without leptons.

In this case, it was found that using the Double Ratio method did not significantly improve the sensitivity of the analysis over the Transfer Factor method. This was due to the fact that the dominant portion of the uncertainties of the analysis are rooted in the limited statistics in the MC samples. This uncertainty cannot be mitigated by any cancellation in these ratios; since

---

[53] See Section 6.2.2 for the definition of these regions.

[54] These aren't "efficiencies" in the strictest sense of the word, since the 2-jet and VBF-like regions are actually orthogonal.

it is statistical in nature, it is uncorrelated between different regions. So, this analysis uses the Transfer Factor method, as it is substantially simpler and achieves roughly the same sensitivity. For future projects (for example, using the full Run 2 dataset), the Double Ratio method may become advantageous to use if MC statistics are no longer a limiting factor.

7.3 Multijets in the Signal Regions

Multijet events (i.e. processes which are mediated purely through gluons and do not involve the electroweak interaction) can enter the signal regions. Although they do not contain any real $E_{\mathrm{T}}^{\mathrm{miss}}$, a substantial amount can sometimes be introduced by detector effects. In particular, this occurs if one or more of the final-state jets is mismeasured or not reconstructed, or if some (but not all) jets from a pileup interaction are included in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Although, detector effects of this magnitude (introducing >180 GeV of $E_{\mathrm{T}}^{\mathrm{miss}}$ where none exists) is very rare, the cross section for multijet production is extremely large. Since it is not technically feasible to generate sufficient Monte Carlo to model this, data-driven methods are necessary to estimate this background. Additionally, uncertainties on any theory-based multijet estimate tend to be large, due to the difficulty of modeling the strong interaction accurately.

The simplest method for estimating this type of background is a so-called “ABCD”-type technique. This involves identifying two uncorrelated cut variables which both discriminate well against multijets, and inverting the requirements on them to define control regions which are enriched in multijets. If the two variables are indeed uncorrelated, an simple extrapolation can be performed from these into the signal regions. The problem with this method here is that there is no suitable choice of two variables in this case. One example would be $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\Delta\phi_{j,MET}$, but these variables are actually quite correlated. In addition, CRs defined by inverting these cuts would not necessarily be free of signal contamination.

Instead, the “Rebalance and Smear” (R+S) technique is used. This is based on the principle of using our knowledge of the detector response to predict the multijet background. First, the jet energy response of the detector is measured independently. Then, a seed sample which is pure in multijets is “smeared” according to this response function many times per event to emulate a large sample of multijet events. From this emulated large sample, the final estimate for the multijet yields in the signal regions are taken.

The seed sample is an inclusive set of events gathered from a set of single-jet triggers at a wide range of p_{T} thresholds (with varying prescales). The samples from the different triggers are “stitched together" by applying a weight to each event equal to its effective trigger prescale. Then, the events are "rebalanced" such that jet resolution effects are not double counted; this also serves to suppress events with real $E_{\mathrm{T}}^{\mathrm{miss}}$. This means the momenta of all the jets in the event (with $p_{\mathrm{T}} > 20$ GeV in this case) are adjusted using a kinematic fit to fulfill the condition $E_{\mathrm{T}}^{\mathrm{miss}} = 0$. Note that this means the vector sum of the jets after the fit is not zero, but rather equal to the soft term[55] of the $E_{\mathrm{T}}^{\mathrm{miss}}$. The soft term is not rebalanced since its response is much more difficult to measure and it is generally only a small component of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in our regions of interest. During this procedure, the JVT discriminant[56] is not used to remove jets associated with pileup vertices. This is done in order to avoid cases where a central pileup jet is rejected but a forward jet from the same vertex is not (since no forward pileup removal is used in this analysis), thereby introducing spurious $E_{\mathrm{T}}^{\mathrm{miss}}$.

To reject non-multijet contributions to the seed sample (from processes with real $E_{\mathrm{T}}^{\mathrm{miss}}$), a cut on the $H_{\mathrm{T}}^{\mathrm{miss}}$ significance is applied:

$$H_{\mathrm{T}}^{\mathrm{miss}}\mathrm{sig.} = \frac{H_{\mathrm{T}}^{\mathrm{miss}}}{\sqrt{H_{\mathrm{T}}}} < 5\ \mathrm{GeV}^{1/2} \tag{7.7}$$

Processes with true $E_{\mathrm{T}}^{\mathrm{miss}}$ are already suppressed by the rebalancing procedure as well. Here, $H_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from *all* jets, regardless of whether they pass the usual JVT requirement. The reason is to avoid the introduction of spurious $E_{\mathrm{T}}^{\mathrm{miss}}$ in cases where the primary vertex is misidentified and jets from the actual hard scatter are incorrectly identified as pileup and removed.

After the rebalanced seed sample has been constructed, each event is smeared many times according to a random sampling of a jet energy response template. These templates are derived from Monte Carlo, but are also constrained by measurements of dijet asymmetry in data They are binned in the jet p_{T} and η, as well as the jet flavor (b or light-flavor). To calculate these, reconstructed jets are matched to truth-level jets within $\Delta R < 0.1$. This tight requirement is intended to minimize the effects of multiple jets being matched to one. Additionally, an isolation requirement is imposed on the matched jets to complement this. The shape of the response is modeled as a Gaussian core with a width corresponding to the usual jet energy resolution (JER), along with non-Gaussian tails on either side. The core is determined by a Gaussian fit constrained to a range equal to twice the distribution's RMS. The result of this fit is subtracted for the original distribution (and negative entries are suppressed) to obtain the tails. A typical example is shown is Figure 7.6. The non-

---

[55] The standard track-based soft term is used. See Section 4.4 for further details on how this is defined.

[56] See Section 4.1.3.

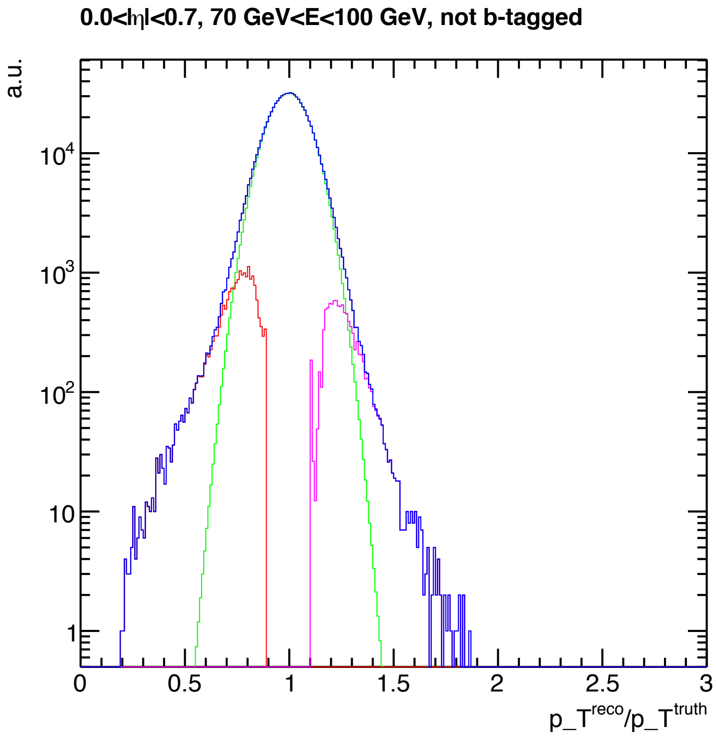

Figure 7.6: Example jet response template, showing the decomposition into the Gaussian core (**green**), upper tail (**pink**), and lower tail (**red**).

Gaussian component is separated out mainly for the purpose of modeling systematic uncertainties on the response function.

After each event is smeared, the jet removal from JVT is re-evaluated. Jets which were originally removed due to JVT but were then smeared above the 60 GeV threshold are added back into the event. Additionally, the primary vertex is re-calculated, since it can change after the smearing in some cases. This is done by checking the $\sum p_{\mathrm{T}}$ of all smeared jets associated with each vertex. If the primary vertex changes, JVT is re-calculated for each jet with respect to this new vertex, and removed/reinstated if necessary. In practice this has little impact after the full SR selection (particularly the cut on $H_{\mathrm{T}}^{\mathrm{miss}}$ computed with no JVT requirement).

The entire seed sample is smeared 20 times, to receive 20 different predictions for the total yield. The variance of this distribution is interpreted as the statistical uncertainty due to the finite size of the seed sample. Each smearing uses a different sampling of the jet response distributions, so each of these smeared versions is different. Events which were triggered only by low-p_{T} triggers have large prescales, which translate into large weights when the seed sample is assembled. To mitigate the impact of statistical fluctuations in those (and to increase the effective sample size), each event is smeared min(1000, prescale) times, and its weight is reduced accordingly.

Before the final prediction of the SRs can be made, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency must also be accounted for. Since the seed sample was collected used single jet triggers, it does not suffer from the same inefficiencies as the SRs do. So, the $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency is calculated from data and parametrized as a function of H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$ (Figure 7.7). Each event from the R+S method is then weighted according to this expected efficiency.

The final results for the R+S prediction of the multijet background in the SRs is given in Table 7.7. The statistical uncertainties are large due to the very small probability that any given event gets smeared into the SR. Systematic uncertainties are associated with our knowledge of the jet response templates. These are modeled as a Gaussian core, with a width equal to the usual jet energy resolution (JER), along with non-Gaussian tails on either side. To calculate the systematic uncertainties on the R+S results, the JER is varied by 10% and the Gaussian tails are varied by 50%, and the final predictions are re-evaluated for each variation.

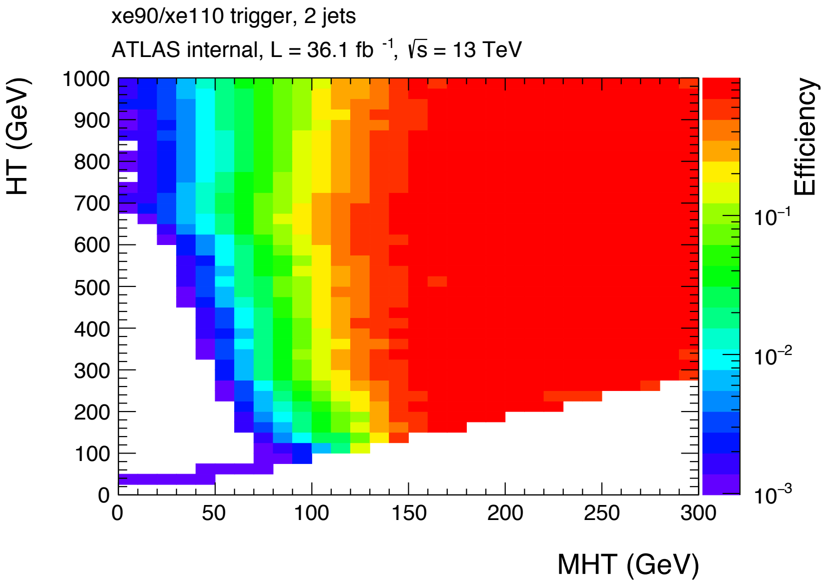

Figure 7.7: $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency for multijet events as a function of H_{T} and $H_{\mathrm{T}}^{\mathrm{miss}}$.

Bin	Nominal	JER up	Tail up	Combined uncertainty
Inclusive	9.82 ± 5.52	8.18 ± 4.38	7.34 ± 4.00	6.35
1.0 TeV $< m_{jj} <$ 1.5 TeV	7.13 ± 4.71	4.52 ± 2.19	5.39 ± 3.60	5.08
1.5 TeV $< m_{jj} <$ 2.0 TeV	2.24 ± 2.80	2.75 ± 3.69	1.40 ± 3.60	5.10
$m_{jj} >$ 2.0 TeV	0.45 ± 0.59	0.90 ± 0.88	0.55 ± 0.71	1.67

Table 7.7: The multijet background prediction in the signal regions, from the R+S method. "Combined uncertainty" indicates the quadrature sum of the statistical uncertainty and systematic variations.

7.3.1 Validation

The simplest way to validate the R+S method is to test it on MC multijet events in the SRs, comparing its prediction to the actual event yields after the selection. This is acts as a closure test; good agreement indicates that the method does not introduce any bias. However, no multijet MC events actually pass the selection for any of the three SRs. This does not imply that 0 events are expected in data, but rather that the size of the multijet samples is insufficient for this test. This is expected, since the multijet cross section is several orders of magnitude larger than any other process involved in this analysis.

Instead, a QCD-enriched validation region is defined which is orthogonal to the SR. This is done by loosening the jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ requirements, and changing the $\Delta\phi$ requirement to a sideband. This significantly increases the expected number of multijet events with respect to the SRs. The selection is the same as the signal region, except:

- The event has $1.8 < \Delta\phi < 2.7$.
- A third jet with $p_{\mathrm{T}} < 50$ GeV is allowed in the event (but no fourth jet).
- The dijet $\Delta\eta$ requirement is relaxed: $\Delta\eta_{jj} > 3.0$.
- This dijet mass requirement is relaxed: $m_{jj} > 600$ GeV.
- The $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is relaxed: $E_{\mathrm{T}}^{\mathrm{miss}} > 100$ GeV.

The comparison between the R+S prediction and the direct event yield from the MC in this region is shown in Figure 7.9. The statistical uncertainties are still fairly large due to limited MC sample size, but the results are consistent.

The same comparison can be carried out with actual data for better statistics, since this VR is orthogonal to the SRs. Good agreement is only observed at low values of $E_{\mathrm{T}}^{\mathrm{miss}}$, while at higher values an excess of data is present. This is attributed to non-QCD processes with real $E_{\mathrm{T}}^{\mathrm{miss}}$, such as $Z \to \nu\nu$ (Figure 7.8). So, kinematic plots are shown only for 100 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 120 GeV (Figure 7.10) and good agreement is found overall. This indicates that the R+S method can reliably model the multijet background without introducing a bias.

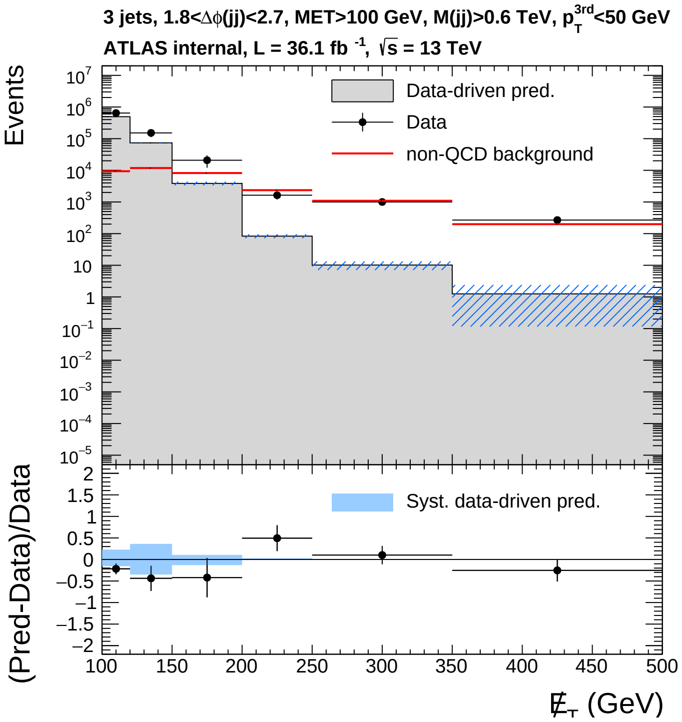

Figure 7.8: $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution comparing data events in the multijet VR with the R+S prediction from data. The MC prediction for non-multijet background is also shown. The ratio plot includes the non-multijet contribution in the numerator but not in the uncertainty bands.

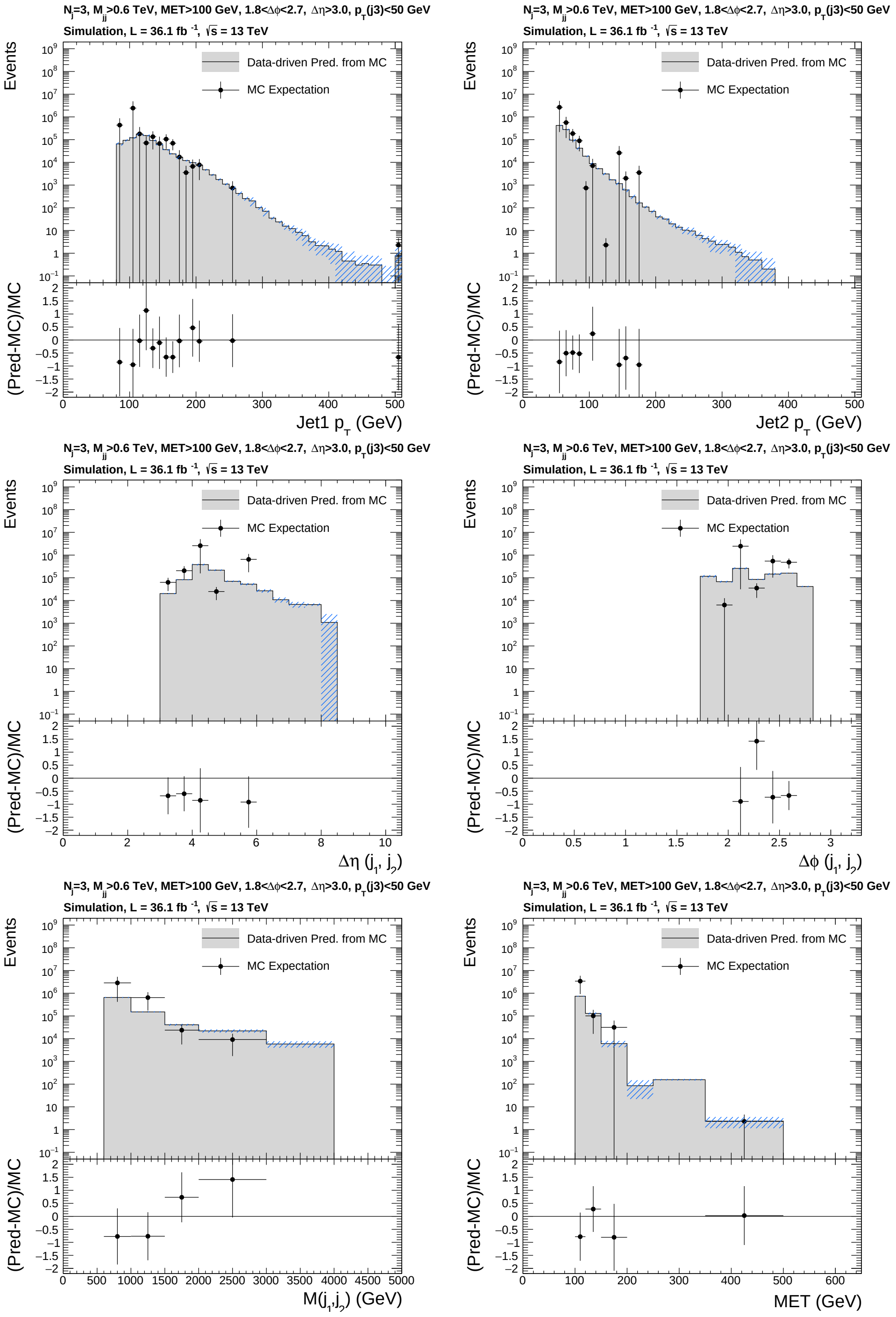

Figure 7.9: Kinematic distributions comparing multijet MC events to the R+S prediction from MC. The selection is the multijet validation region defined in the text.

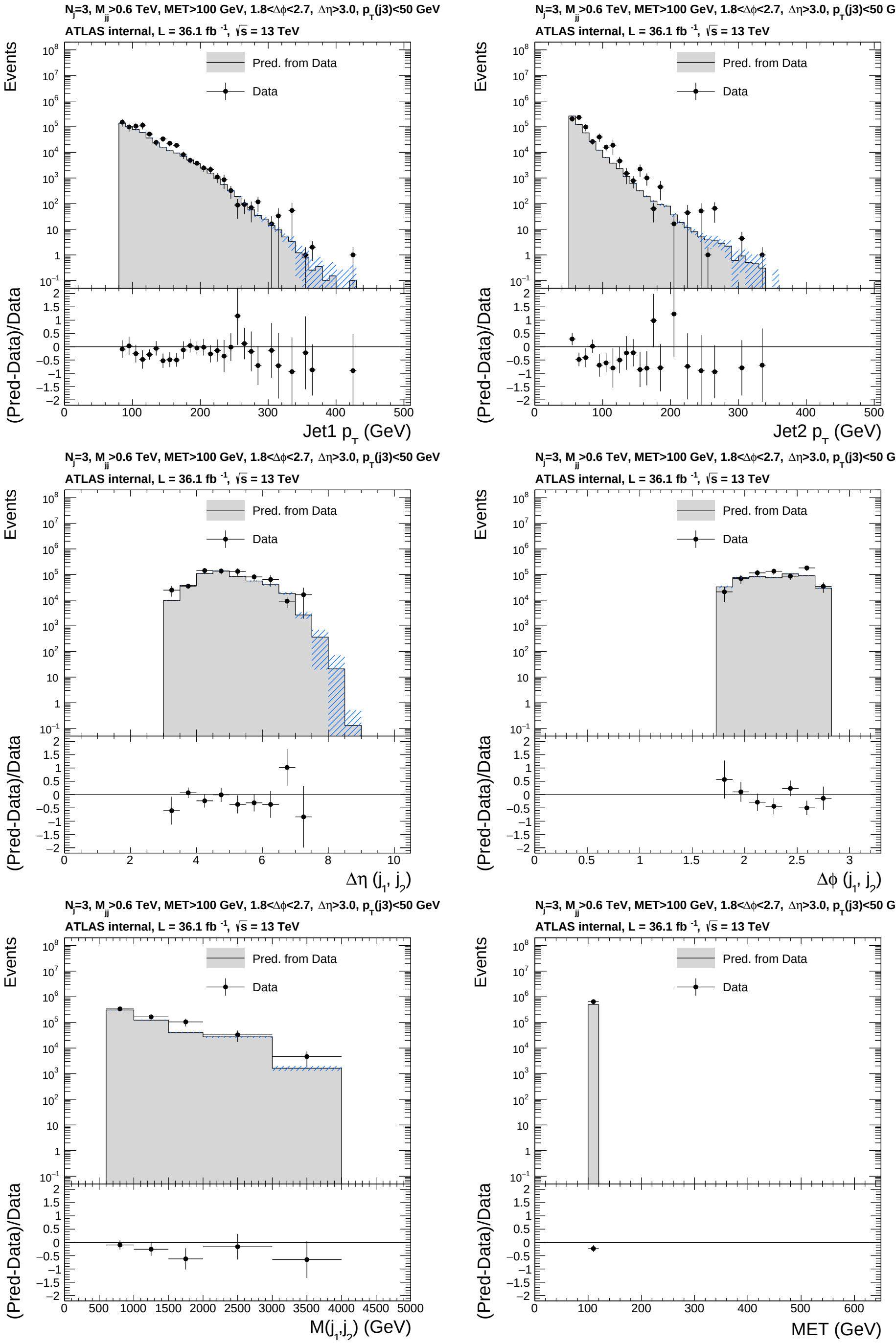

Figure 7.10: Kinematic distributions comparing data events to the R+S prediction from those same events. The selection is the multijet validation region defined in the text. All distributions except $E_{\mathrm{T}}^{\mathrm{miss}}$ have an additional requirement of $E_{\mathrm{T}}^{\mathrm{miss}} < 120$ GeV to suppress non-multijet backgrounds. Non-multijet backgrounds are not shown.

7.4 Fake Electrons

In the 1-electron control regions, there is a non-negligible background from misidentified (or "fake") electrons. These are objects which are reconstructed and identified as electrons, but were in reality something else. These are essentially always jets, which are usually (but not always) removed by isolation and shower shape requirements[57]. As a result, most events in this category are from QCD multijets, where one jet is misidentified as an electron. Since this process has an extremely large cross section and is very difficult to model theoretically, it cannot be effectively estimated from Monte Carlo. Instead, a data-driven technique must be used. Note that this is used only for electrons; the contribution from fake muons in the $W \to \mu\nu$ regions is negligible.

Because the multijet background does not have a prompt neutrino, the $E_{\mathrm{T}}^{\mathrm{miss}}$ tends to be smaller and to point in the direction of the jet that was misidentified as an electron. As a result, it tends to have significantly lower $E_{\mathrm{T}}^{\mathrm{miss}}$ significance[58] than the real W+jets contribution. This makes $E_{\mathrm{T}}^{\mathrm{miss}}$sig. an ideal variable for discriminating between the two. A "template fit" using this distribution is used to estimate this background.

Control samples modeling the jets misidentified as electrons in multijet events are constructed by selecting events that pass the $W \to e\nu$ CR selection, except that the electron identification requirement is invert and fully identified electrons are removed (i.e. they contain objects which are reconstructed as electrons but fail the Tight ID criteria). These are called the "anti-ID" control samples. To obtain the normalization of the multijet background, templates of the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. distribution for processes with prompt leptons are first taken from MC simulation. Shape templates for the multijet background are constructed by by subtracting the prompt-electron contribution (taken from MC) from the observed data in the anti-ID control samples. The anti-ID samples are very pure in multijet events, so there is no need to fit the MC normalization here. Since the multijet contribution is expected to be (very nearly) charge-symmetric, the same shape template and normalization factor is used to model both charge categories. The shape of this $E_{\mathrm{T}}^{\mathrm{miss}}$sig. template is shown in Figure 7.11.

To determine the misidentified lepton background normalization, a fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. distribution in each of the main $W \to e\nu$ control regions is used. The actual W+jets contribution is not charge-symmetric, so the different charge samples are kept separate in the fit. An additional set of control regions is introduced which are the same as the main $W \to e\nu$ regions, but with the $E_{\mathrm{T}}^{\mathrm{miss}}$sig.

---

[57] See Section 4.2 for further details on electron reconstruction and identification.

[58] $E_{\mathrm{T}}^{\mathrm{miss}}$ significance is defined in Equation 6.1.

requirement inverted. This serves to constrain the fake electron normalization (β_e^i). There are two free normalizations for each m_{jj} bin in the fit: one for events with a prompt lepton (k_W^i) and one for events where a jet is misidentified as an electron (β_e^i). These two parameters are constrained by data in two control regions: $W \rightarrow e\nu$ with $E_{\mathrm{T}}^{\mathrm{miss}}$sig. above and below the threshold. In practice, this is all included in the final fit for the whole analysis[59]. With this method, there is no pre-fit estimate of the fake electron background.

The expected background from misidentified electrons as given by the final fit is shown in Table 7.8. In practice, only the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. shape in two bins is used (above and below the cut value of 4 GeV$^{1/2}$). So, the template shape can be specified by a single number: the ratio of the two bins. This is denoted R^i and is also shown in Table 7.8.

Since everything is carried out in a single simultaneous fit, it is in general possible that the k_W^i normalizations could be biased by the fake electron estimate. In order to check this, this fit was also carried out using only the $W \rightarrow \mu\nu$ regions, which are unaffected by fakes. The resulting fit parameters were found to be consistent with those from the full fit within uncertainties (Table 7.9). In addition, another check was performed by fixing k_W^i to 1 and fitting the fake normalization once more. Again, the results were found to be consistent within uncertainties. This indicates that the W normalization is not biased by the fake electron fit.

---

[59] For the full description of how this is implemented in the framework of the fit, see Section 8.1.

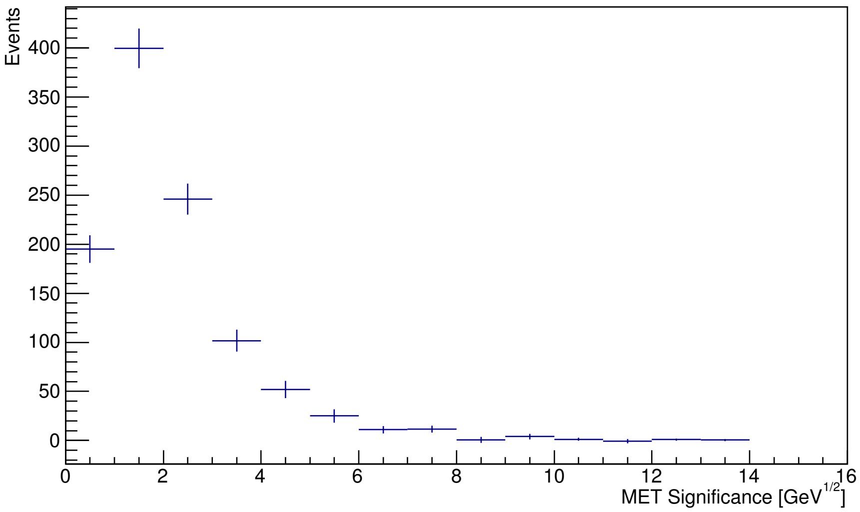

Figure 7.11: $E_{\mathrm{T}}^{\mathrm{miss}}$sig. template shape for fake electrons in the inclusive $W \to e\nu$ control region. In practice, three of these are used (one for each m_{jj} bin).

Bin	R	N_{fake}
1.0 TeV $< m_{jj} <$ 1.5 TeV	9.0 ± 0.4	7.8 ± 3.9
1.5 TeV $< m_{jj} <$ 2.0 TeV	9.7 ± 0.6	3.8 ± 2.3
$m_{jj} >$ 2.0 TeV	5.0 ± 0.4	4.0 ± 2.9

Table 7.8: Expected background yields due to misidentified electrons in the $W \to e\nu$ CRs (N_{fake}). The template ratio of fake electron events below and above the $E_{\mathrm{T}}^{\mathrm{miss}}$sig. cut ($R^i$) is also shown.

Bin	$W \to \mu\nu$ Only		Full Fit	
	k_W	β_e	k_W	β_e
1.0 TeV $< m_{jj} <$ 1.5 TeV	1.15 ± 0.17	10.4 ± 2.9	1.07 ± 0.25	8.4 ± 2.9
1.5 TeV $< m_{jj} <$ 2.0 TeV	1.06 ± 0.15	4.8 ± 1.9	1.01 ± 0.18	4.8 ± 1.7
$m_{jj} >$ 2.0 TeV	1.06 ± 0.15	6.7 ± 2.6	1.19 ± 0.23	5.3 ± 2.7

Table 7.9: Comparison of the fit parameters using the full and muon-only fits. The full fit results are not the same as those reported in Chapter 8, because some uncertainties were updated after this check was performed.

Chapter 8

Results and Interpretation

8.1 Statistical Methods

In order to interpret the raw experimental data, a simultaneous likelihood fit is carried out on the signal and control regions. This is based on frequentist statistics. A model of the background and possible signal is constructed using Monte Carlo and the data-driven techniques described in Chapter 7. This is used to construct probability distributions as a function of the observed data in each region. The model has several parameters, which can be allowed to vary in the fit:

- The **signal strength parameter** (μ): This is the normalization for the signal process. Assuming the SM production cross section, $\mu = 1$ corresponds to $\mathcal{B}(h \to \text{inv.}) = 100\%$. The background-only hypothesis is $\mu = 0$ (in reality the SM cross section is nonzero, but too small to detect). μ does not have a probability distribution associated with it; it does not constrain the fit in any way.

- The **nuisance parameters** (α_i, γ_i): These parametrize the systematic uncertainties. There is one α for each independent uncertainty source. They are usually defined to have a Gaussian probability distribution with a mean $\bar{\alpha} = 0$ and width $\sigma_\alpha = 1$. A value of $\alpha_i = 1$ means systematic i is varied up by 1σ, and the yields predicted by the model are modified accordingly. The γ_i are very similar, but are used to parametrize the statistical uncertainty due to the finite MC sample size. They are assigned Poisson probability distributions rather than Gaussian ones.

- The **normalization parameters** (k^i_W, k^i_Z, β^i_e): These are additional normalizations for the various yields in the signal and control regions. They are used to propagate constraints and corrections between the different regions, since the MC is not assumed to be completely correct. For example, k^i_W is the normalization for $W \to \ell\nu$ MC in the SR and 1-lepton CR ith m_{jj} bin. It serves to apply the constraint from the 1-lepton CR to the W background in the SR. A similar procedure is used for the fake lepton contribution to the 1-electron CRs; β^i_e is the normalization for this. There are 9 total normalization parameters (k_W, k_Z, and β_e for each of the three bins). Note that the R^i (defined in Section 7.4) are not in this category, as they are held fixed in the fit.

The function of the control regions is implemented in the scheme used for the normalization parameters. The exact scheme is shown in Table 8.1. By using the same k^i_W and k^i_Z across the signal and control regions, they constrain each other in the fit. This allows the W and Z backgrounds to be constrained simultaneously. At the same time, the fake lepton fit happens for the 1-electron control regions. The R^i are determined in advance from the template fit[60], and the β^i_e are the normalization which is constrained here.

Using this model, a global likelihood function is constructed for every possible experimental outcome (in practice a finite sampling is used), given a certain hypothesis. This accounts for the Poisson uncertainty due to the finite size of the dataset. In the following equations, α_i or $\vec{\alpha}$ will be used to denote **all** nuisance parameters (including the γ_i) and k^i or $\vec{k}$ will be used to denote **all**

Region	Lepton charge	$E_{\mathrm{T}}^{\mathrm{miss}}$sig.	Signal	Z+jets	W+jets	Top	Multijet
SR			$\mu\times$ MC	$k^i_Z\times$ MC	$k^i_W\times$ MC	MC	(R+S)
$W \to e\nu$	+	> 4 GeV$^{1/2}$	0	$k^i_Z\times$ MC	$k^i_W\times$ MC	MC	β^i_e
$W \to e\nu$	+	< 4 GeV$^{1/2}$	0	$k^i_Z\times$ MC	$k^i_W\times$ MC	MC	$R^i\times\beta^i_e$
$W \to e\nu$	-	> 4 GeV$^{1/2}$	0	$k^i_Z\times$ MC	$k^i_W\times$ MC	MC	β^i_e
$W \to e\nu$	-	< 4 GeV$^{1/2}$	0	$k^i_Z\times$ MC	$k^i_W\times$ MC	MC	$R^i\times\beta^i_e$
$W \to \mu\nu$	+		0	$k^i_Z\times$ MC	$k^i_W\times$ MC	MC	0
$W \to \mu\nu$	-		0	$k^i_Z\times$ MC	$k^i_W\times$ MC	MC	0
$Z \to ee$			0	$k^i_Z\times$ MC	$k^i_W\times$ MC	MC	0
$Z \to \mu\mu$			0	$k^i_Z\times$ MC	$k^i_W\times$ MC	MC	0

Table 8.1: Summary of the fit model for one m_{jj} bin. The entries in the table are the event yields in each region predicted by the model. "MC" refers to the Monte Carlo prediction and is different for each region and process. The "MC" values are implicitly functions of the nuisance parameters α and γ. This is all replicated for the other two values of i. Note that μ does not carry any index. Entries in gray are small and offer little or no constraining power.

---

[60] The details of this method are discussed in Section 7.4.

normalization parameters (including the β_e^i). This is simply to aid readability; they have exactly the same function in the fit. The exact form of the global likelihood function is:

$$L(\text{data}|\mu, \vec{k}, \vec{\alpha}) = \prod_i P(d_i|B_i(\vec{k}, \vec{\alpha}) + \mu S_i(\vec{\alpha})) \prod_s \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\alpha_s^2} \tag{8.1}$$

Here, i indexes all signal and control region bins, and s indexes the sources of systematic uncertainty. d_i denotes the observed yield in region i in data. P represents the Poisson distribution, which gives the probability to observe a number of events x when y are expected under a certain hypothesis for the fit parameters:

$$P(x|y) = \frac{y^x}{x!} e^{-y} \tag{8.2}$$

B_i denotes the total background estimate in bin i; in practice this is a sum over the W, Z, and other background contributions, which each have their own dependence on $\vec{k}$ as specified in Table 8.1. The nuisance parameters are taken as Gaussian uncertainties on the nominal background prediction, so the exact form is:

$$B_i(\vec{k}, \vec{\alpha}) = \sum_p k_p^i B_i^{p,\text{ nominal}} \left(1 + \sum_s \alpha_s \sigma_{sp}^i\right) \tag{8.3}$$

where p indexes the different background processes (W, Z, multijet, etc.) and σ_{sp}^i denotes the relative impact of each uncertainty on the given yield[61]. From the global likelihood function, a test statistic is defined as a function of μ. In the absence of any additional fit parameters beyond μ, it has the form:

$$q(\mu) = -2 \ln \frac{L(\text{data}|\mu)}{L(\text{data}|\hat{\mu})} \tag{8.4}$$

In the denominator, $\hat{\mu}$ is the best-fit value of μ when it is also allowed to float freely in the fit. This likelihood ratio is the most powerful statistical discriminant between two hypotheses [112].

Of course, there are in reality many systematic uncertainties and normalization factors. Uncertainties are included as nuisance parameters in the MC predictions themselves (these are the α_i and γ_i). To handle these, the test statistic is generalized to a profile likelihood ratio:

---

[61] So, σ_{sp}^i is the raw "size of the uncertainty", which is reported in places such as Table 7.2.

$$\tilde{q}(\mu) = -2\ln\frac{L(\text{data}|\mu, \hat{\hat{\vec{\alpha}}}_\mu, \hat{\hat{\vec{k}}}_\mu)}{L(\text{data}|\hat{\mu}, \hat{\vec{\alpha}}, \hat{\vec{k}})} \tag{8.5}$$

Here, $\hat{\hat{\vec{\alpha}}}_\mu$ and $\hat{\hat{\vec{k}}}_\mu$ denote the best-fit values for all the α_i and k_i when μ is held fixed to the value at which the test statistic is being evaluated. Again, all parameters in the denominator are best-fit values from the fit where μ is allowed to float freely. This test statistic summarizes the compatibility of entire measurement with the hypothesis. For example, $\tilde{q}(0)$ is a single number which describes how well the background-only hypothesis describes the data.

In order to set limits on a theoretical model, the quantity of interest is the p-value. This is defined as the probability, given a hypothesis, to obtain the observed data (or more extreme)[62]. To calculate this, it is necessary to know the probability distribution of $\tilde{q}$ for the given hypothesis. This cannot be calculated analytically, but the distribution can be sampled using "toys". This method involves using MC to generate many toy datasets under the hypothesis of interest and evaluting $\tilde{q}$ for each one. However, this can be very computationally intensive. For cases with sufficiently large sample sizes (including this analysis), asymptotic approximations can be used to evaluate the $\tilde{q}$ distributions instead of toys [113].

The typical convention for exclusion limits on a signal hypothesis is $p < 0.05$. To reach this, various hypotheses for μ are scanned over in order to reach the point where $p = 0.05$. In practice, we use not only the p-value for the signal hypothesis, but a quantity called CL_s [85] defined as:

$$CL_\text{s} = \frac{p_{s+b}}{1 - p_b} \tag{8.6}$$

where p_{s+b} and p_b are the p-values for the signal+background hypothesis and the background-only hypothesis respectively. The purpose of using this is that the signal here is small compared to the background, and as a result, the $\tilde{q}$ distributions for the hypotheses with and without signal have a lot of overlap. As a result, a fluctuation in the background can easily be misinterpreted as being related to the signal. Accounting for the compatibility with the background-only hypothesis suppresses this effect. The upper limit on $\mathcal{B}(h \to \text{inv.})$ is then defined as the smallest μ hypothesis for which $CL_\text{s} < 0.05$. Note that this does not provide 95% coverage[63] like the p-value does; it is actually

---

[62] "More extreme" means a larger value of the test statistic.

more conservative since $p_b < 1$. For example, for an expected limit, $p_b = 0.5$, so this would actually correspond to 97.5% coverage.

The **expected limit** is defined the same way, but the observed data yields in the signal regions are not used. Instead, those values are replaced by the yields predicted by the background-only hypothesis (using the best-fit parameters given $\mu = 0$). This gives a description of how sensitive the analysis is, independent of any potential statistical fluctuations in the SRs. Note that the signal region data can still implicitly enter this calculation, as it influences the best-fit values of the nuisance parameters. Those best-fit values can be determined from a CR-only fit to remove this effect, but this is not the standard convention. Additionally, an "Asimov dataset" can be used to compute expected limits. In this case, all signal *and* control region data is replaced by the background-only prediction with all nuisance parameters set to zero. This provides an estimate of the sensitivity of the analysis without being affected by any statistical fluctuations in data at all. However, it also ignores any mismodeling in MC.

8.2 Results

Two types of fit are performed here, as described in the previous section. First, a background-only fit is carried out to check the consistency of the method and ensure that all of the relevant processes are well-modeled. This involves using only data in the control regions (the signal regions are excluded entirely), and fixing μ to zero. Then, the full simultaneous fit to all regions is done to obtain the final results.

8.2.1 Background-Only Fit

A summary of the results of the background-only fit are shown in Figure 8.1. The full $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions are given in Figures 8.2 and 8.3. The fake electron contribution is not shown in Figure 8.2, since the fit gives only an inclusive yield estimate rather than kinematic distributions. The dijet mass distributions are not plotted; since the analysis is binned in this variable, the fit forces overall agreement at the level of each bin (provided there are no major differences between the CRs). Agreement for the m_{jj} distribution is not necessarily expected within each bin, as the MC is known to mismodel this.

After the background-only fit, there is good agreement between data and MC across all CRs. This indicates that the MC is modeling the shape of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions well, and that the

---

[63] "Coverage" refers to the probability of the excluded region containing the true value under the given hypothesis.

fit is stable. Note that the fit does not affect any kinematic shapes within any region; only the normalization. The results of the background-only fit are shown in Table 8.2. All of the W and Z normalization parameters are consistent with unity, which implies that the MC mismodeling of the m_{jj} distribution (at the level of the three bins) is smaller than the uncertainties.

If these normalization factors are then applied to the MC predictions in the SR, they can be compared to data as a cross-check. In the absence of a signal, the agreement should be good despite the SRs not being included in the fit. As shown in Figure 8.4, this is indeed the case.

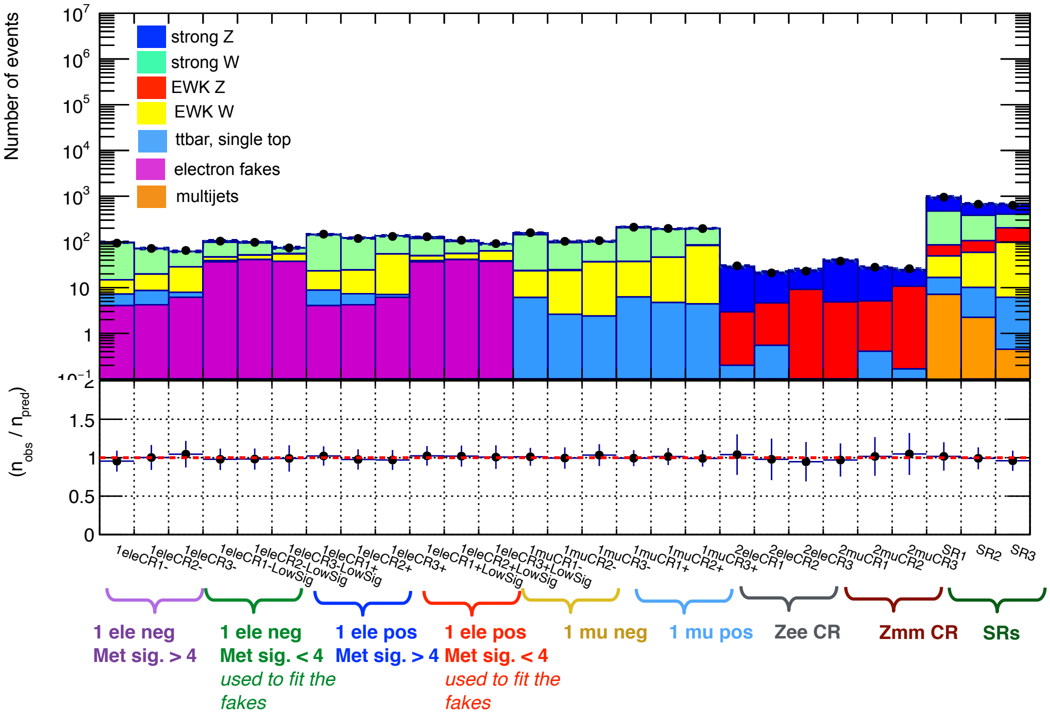

Figure 8.1: Summary plot showing data and post-fit expected yields after the background-only fit. The error bars include both statistical and systematic uncertainties. The SR yields are shown but are not included in the fit.

m_{jj} Bin	k_W	k_Z	β_e
1.0 – 1.5 TeV	1.00 ± 0.24	1.04 ± 0.28	2.2 ± 1.9
1.5 – 2.0 TeV	1.03 ± 0.28	0.89 ± 0.25	1.9 ± 1.8
> 2.0 TeV	1.23 ± 0.25	0.99 ± 0.26	2.0 ± 2.2

Table 8.2: Fit results for the background-only fit to CR data. The numbers shown are the best-fit values for the W and Z normalization factors, as well as the fake electron normalization β_e.

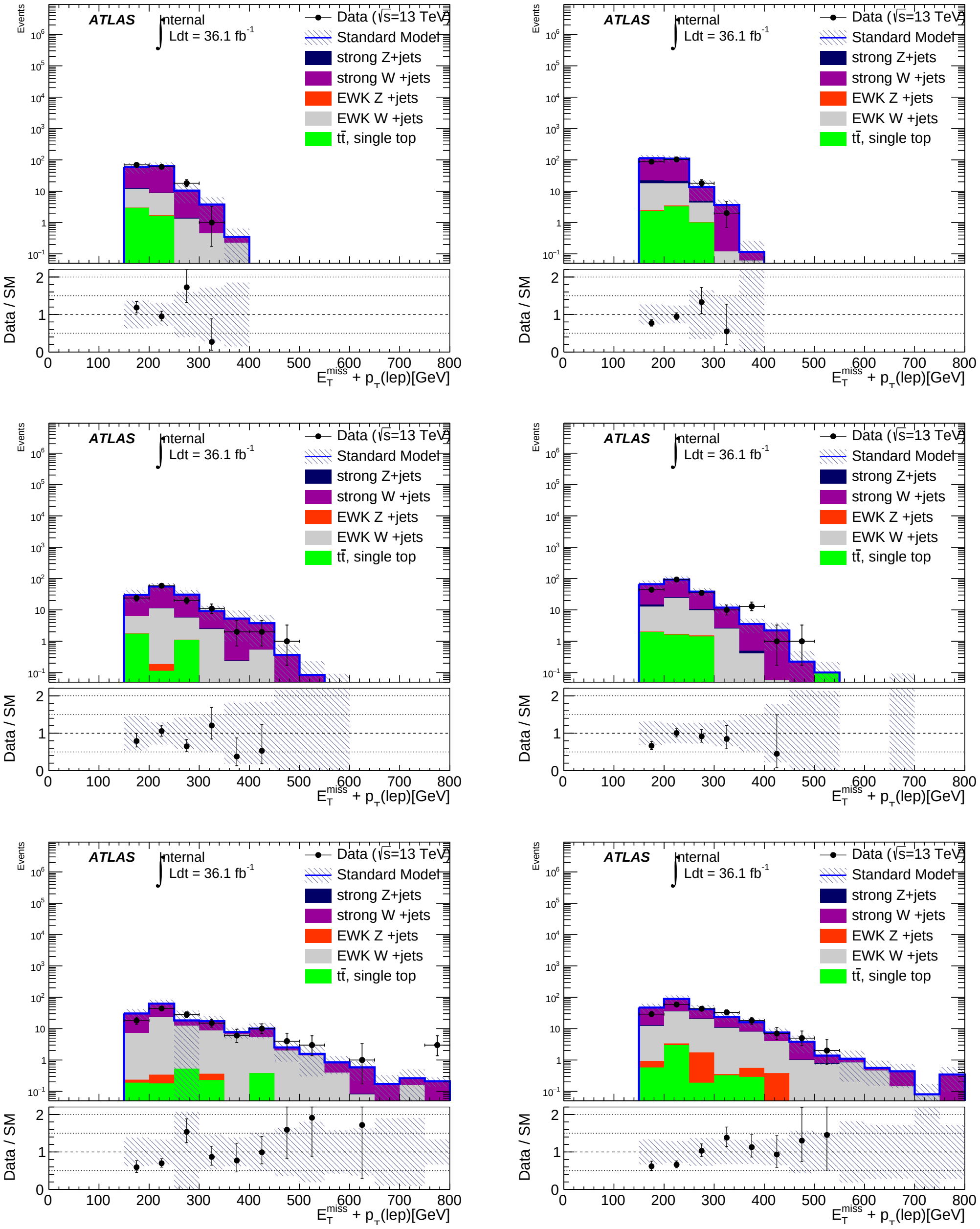

Figure 8.2: Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions (including the lepton p_{T}) for the e^+ **(left)** and μ^+ **(right)** control regions after the background-only fit. The three rows are the three m_{jj} bins: CR1 **(top)**, CR2 **(middle)**, and CR3 **(bottom)**.

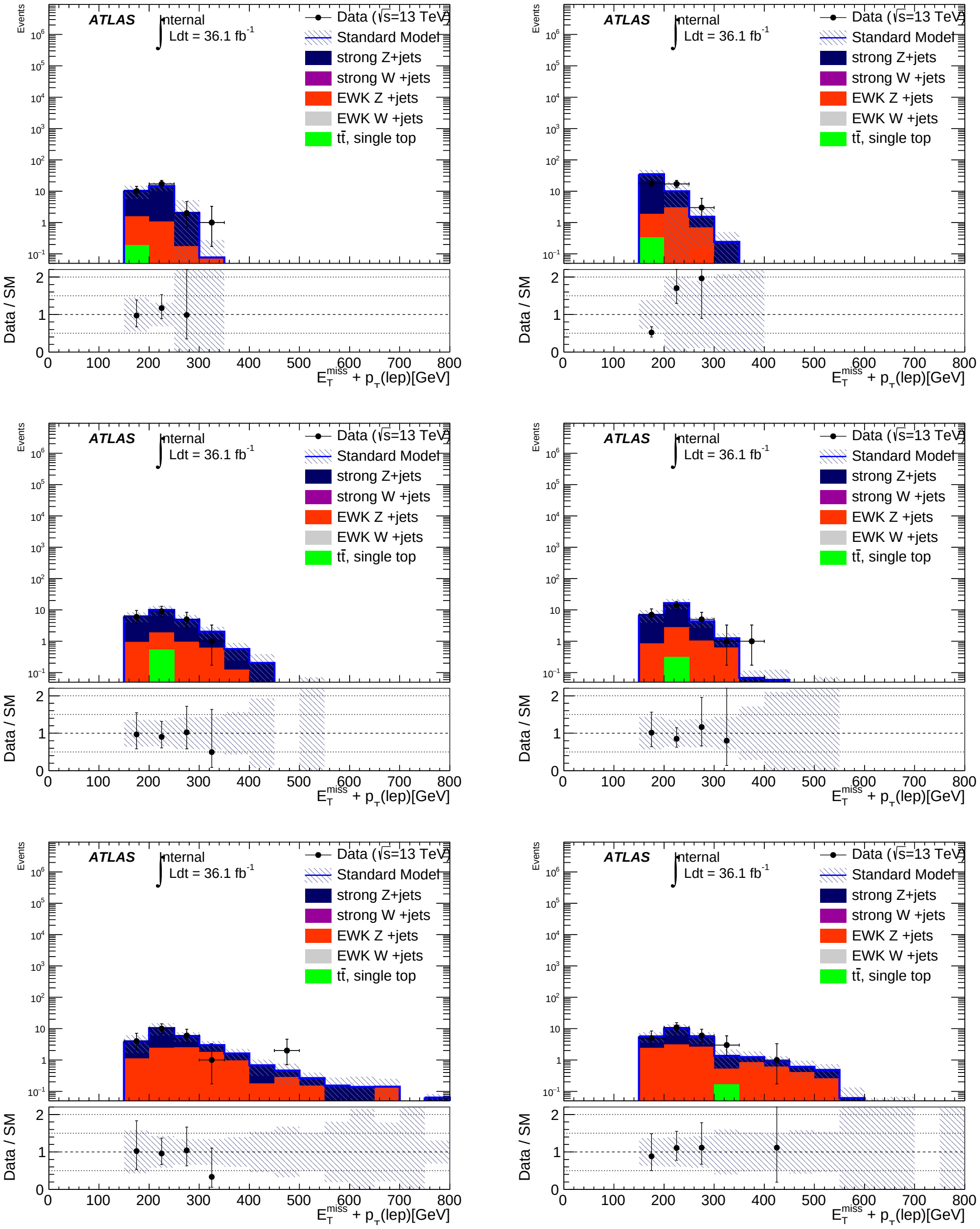

Figure 8.3: Post-fit $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions (including the dilepton p_{T}) for the $2e$ **(left)** and 2μ **(right)** control regions after the background-only fit. The three rows are the three m_{jj} bins: CR1 **(top)**, CR2 **(middle)**, and CR3 **(bottom)**.

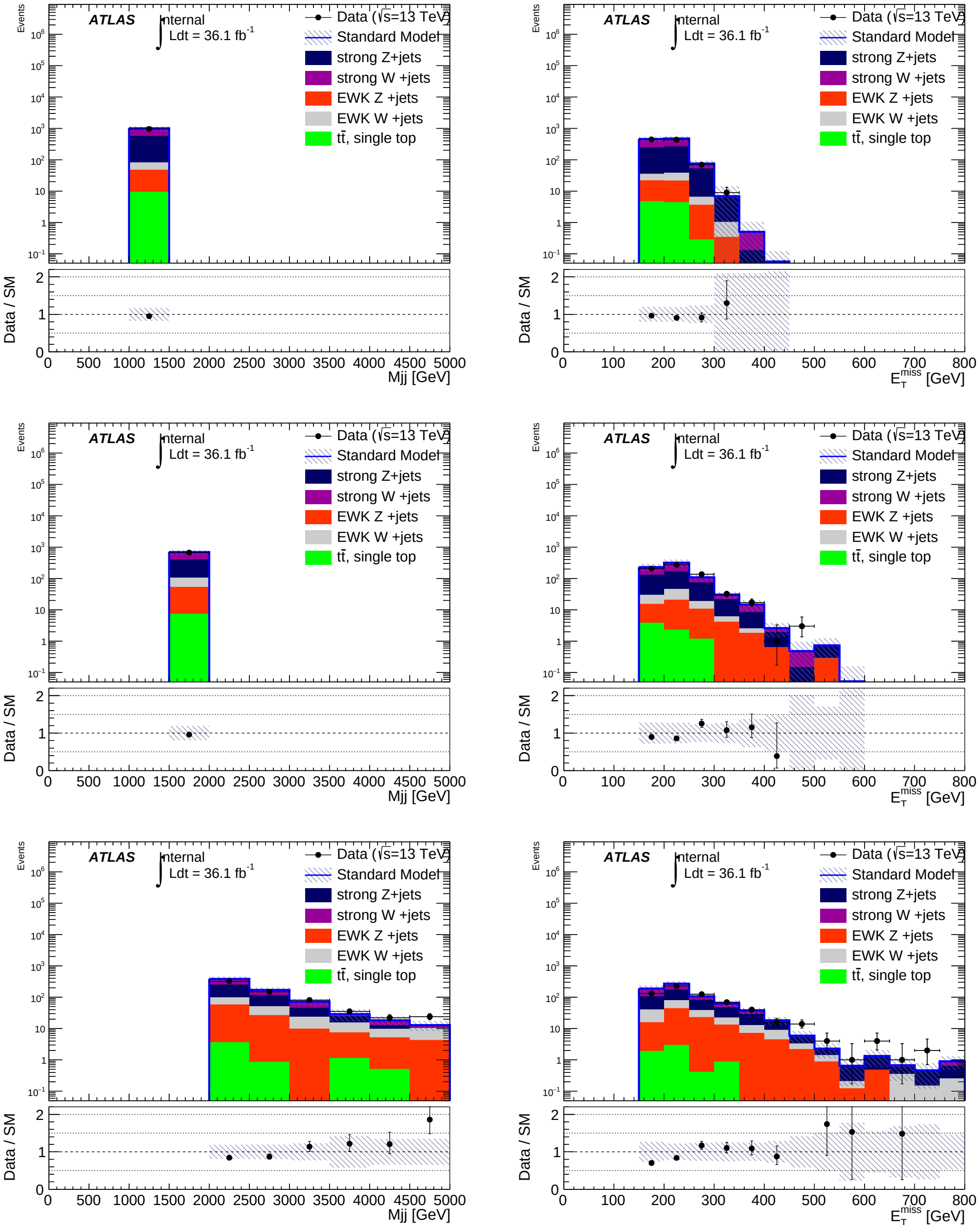

Figure 8.4: Post-fit m_{jj} **(left)** and $E_{\mathrm{T}}^{\mathrm{miss}}$ **(right)** distributions in the signal regions after the background-only fit. The three rows are the three m_{jj} bins: SR1 **(top)**, SR2 **(middle)**, and SR3 **(bottom)**. The data points are shown only for reference, they were not used in the fit.

Region	Monte Carlo					QCD	Signal	Total Background	Data
	Strong Z	EWK Z	Strong W	EWK W	Top				
SR Bin 1	455.4	37.0	442.6	36.8	9.5	7.1	244.2	988.4	952
SR Bin 2	308.3	53.5	287.5	51.8	6.8	2.2	281.2	710.2	667
SR Bin 3	201.3	80.2	176.7	82.1	6.7	0.5	401.6	547.4	633
$W \to \ell\nu$ Bin 1	23.5	1.2	509.8	81.2	19.9		0.0	635.5	611
$W \to \ell\nu$ Bin 2	12.9	1.5	356.1	106.4	13.4		0.0	490.4	491
$W \to \ell\nu$ Bin 3	7.2	2.0	231.5	171.4	8.9		0.0	421.1	500
$Z \to \ell\ell$ Bin 1	66.8	6.9	0.0	0.0	0.7		0.0	74.4	68
$Z \to \ell\ell$ Bin 2	49.9	10.0	0.0	0.0	1.1		0.0	61.0	49
$Z \to \ell\ell$ Bin 3	30.5	17.7	0.0	0.0	0.3		0.0	48.4	49

Table 8.3: Summary of yields used in the final fit in all regions. The MC values are pre-fit (i.e. normalization factors are not applied). The control regions are summed over lepton flavor and charge for simplicity. For the $W \to e\nu$ CRs, only the region above the $E_{\mathrm{T}}^{\mathrm{miss}}$ significance cut is included. Fake electrons are not included, as there is no pre-fit estimate defined. The "Signal" column is normalized to $\mathcal{B}(h \to \text{inv.}) = 100\%$.

m_{jj} Bin	k_W	k_Z	β_e	μ
1.0 – 1.5 TeV	0.99 ± 0.16	0.94 ± 0.24	2.7 ± 1.8	
1.5 – 2.0 TeV	1.01 ± 0.19	0.80 ± 0.20	2.0 ± 1.7	0.07 ± 0.13
> 2.0 TeV	1.18 ± 0.20	1.03 ± 0.26	2.0 ± 2.1	

Table 8.4: Fit results for the full fit to SR and CR data. The numbers shown are the best-fit values for the W and Z normalization factors, as well as the fake electron normalization β_e and the signal strength μ. Note that the same value of μ applies to all three m_{jj} bins.

Observed	Expected				
	Nominal	$+1\sigma$	-1σ	$+2\sigma$	-2σ
0.34	0.28	0.40	0.20	0.56	0.15

Table 8.5: Observed and expected upper limits on $\mathcal{B}(h \to \text{inv.})$, at 95% confidence level (using the CL_{s} method). The $\pm 1\sigma$ and $\pm 2\sigma$ values indicated what the expected limit would be under the corresponding fluctuation of data in the signal regions.

8.2.2 Full Fit

With everything now validated, the signal regions are unblinded and the full fit to data in all SRs and CRs is carried out. This is used to calculate the upper limit on $\mathcal{B}(h \to \text{inv.})$, as described in Section 8.1. A summary of the event yields in all regions is given in Table 8.3. The best-fit values for the main parameters are given in Table 8.4, and the resulting limits are shown in Table 8.5. The observed 95% upper limit on $\mathcal{B}(h \to \text{inv.})$ is **0.34** (using the CL_{s} method), and the expected limit is **0.28**.

8.2.3 Uncertainties and Limiting Factors

The most important uncertainties are shown in Table 8.6. This gives the impact on the W and Z normalization factors associated with a 1σ variation of the nuisance parameter, for one representative m_{jj} bin. The impact on the upper limit on $\mathcal{B}(h \to \text{inv.})$ is also shown, which is defined as the relative change in the expected limit when that uncertainty (or set of uncertainties) is "turned off". This means that nuisance parameter (or set of them) is fixed to the overall best-fit value, and the fit is re-run and the limit re-calculated. This provides a first-order estimate of how much more sensitive the analysis would be if these uncertainties could be reduced to zero.

One of the largest systematic uncertainties is the statistical error on the Monte Carlo predictions due to the finite sample size. In principle this effect can be mitigated by generating more events, but in practice this was not technically feasible. The main limitations come from the strong-produced W and Z samples. As shown in Table 6.3, the effective integrated luminosity for many of the important samples is not substantially larger than in data. This means all of the associated background predictions are subject to statistical fluctuations of roughly the same size as those in data. Furthermore, the same MC is used to evaluate systematic uncertainties. This means the estimates of those uncertainties are themselves also subject to statistical fluctuations. This can potentially result in artifically large values for some systematic uncertainties.

A substantial increase (roughly a factor of 4) to the amount of MC simulation for W and Z+jets was requested, but was not available within the timeline of this project, due to the extremely large quantity of computing resources required. This uses a method based on filtering the events at generator level based on the dijet mass, and simulating detector interactions and reconstruction for only a small fraction of events at low m_{jj}. Due to the finite detector resolution and pileup effects, events with low m_{jj} (or even only 1 jet) at generator level can migrate into the signal and control regions. For future iterations of this search, it will be crucial to utilize any available techniques to mitigate the problem of limited MC statistics, as this will likely become the dominant uncertainty by a wide margin otherwise.

Out of the experimental uncertainties, the jet-related ones are the most important here. In other similar searches which use the Transfer Factor method (such as the most recent ATLAS monojet search [45]), this is not always the case because jet modeling and reconstruction effects "cancel out" between the signal and control regions. However, since this search relies primarily on forward jets (which have larger uncertainties associated with them relative to central jets), the impact is substantially larger. The use of a third jet veto also introduces much more sensitivity to

Source	Transfer Factors (%)		Limit Improvement (%)
	α_Z	α_W	
Experimental			
Jet Energy Scale	15	16	9
Jet Energy Resolution	3	5	3
$E_{\mathrm{T}}^{\mathrm{miss}}$ Soft Term	< 1	< 1	1
Pileup Distribution	< 1	1	1
Luminosity	< 1	< 1	< 1
ℓ ID and veto	< 1	4	3
Theoretical			
PDF	1	1	< 1
Fact./Renorm. Scale	1	2	3
Resummation Scale	3	3	2
CKKW Matching Scale	< 1	5	5
3^{rd} Jet Veto	-	-	1
Statistical			
MC Sample Size	-	-	15
Combined Impact			
Experimental			16
Theoretical			12
Experimental and Theoretical			28
Experimental, Theoretical, and MC Stats			42

Table 8.6: Impact of the dominant uncertainties on the W/Z normalization factors and upper limit on $\mathcal{B}(h \to \text{inv.})$. The normalization uncertainties are shown only for the first bin (1.0 TeV $< m_{jj} <$ 1.5 TeV). Empty entries (indicated by '-') mean that uncertainty or group does not affect the quantity in question in a well-defined way. The exact definition of these numbers is described in the text. The transfer factors are defined in Section 7.2.

jet reconstruction effects relative to analyses which are inclusive in the number of jets.

Theoretical uncertainties, particularly on the W and Z background, are also significant. The resummation and CKKW matching scale uncertainties had to be evaluated using dedicated varied MC samples, which were quite limited in statistics. As a result, these uncertainties are quite possibly dominated by statistical fluctuations[64]. The method of extrapolating these variations from lower m_{jj} helps to mitigate this effect, but it is still present to some extent.

For a more detailed look at the effect of each source of systematic uncertainty, the pull and impact on $\hat{\mu}$ of each NP is shown in Figure 8.5. This includes NPs associated with experimental/theoretical systematics (the α's[65]) and those associated with MC statistical uncertainties (the γ's). The fit does not significantly constrain any of the systematic uncertainties, which would potentially lead to

---

[64] See Tables 7.4 and 7.5 to compare the nominal values of these uncertainties with the statistical errors on them.

[65] The list of experimental α parameters is given in Section 6.4. Theoretical α parameters are described throughout Chapter 7.

an underestimate of their impact. As Table 8.6 suggests, the most important contributions come from MC statistical uncertainties, the jet energy scale, and theoretical uncertainties in the MC.

8.3 Interpretation

An upper limit has been set on the invisible branching ratio of the SM Higgs, which is completely model-independent. However, this can also be interpreted in terms of specific dark matter models. Here, a few example models are discussed in detail. This is not an exhaustive list; the space of DM models is far too large to include a fully comprehensive study here. In principle, if theorists (or experimentalists) are interested in some new model, it can then be checked against the limits presented here and in other similar publications.

8.3.1 The Higgs Portal Model

The "Higgs Portal" model is one of the simplest DM models; it involves adding only one new particle (χ) to the SM. This particle, which is intended as a dark matter candidate, is assumed to couple only to the Higgs (excluding gravity). Since the model is so simple, it only has two free parameters (once the spin of χ is chosen): the new particle's mass m_χ and its coupling to the Higgs λ_χ.

With a few assumptions, the results from this search can be compared to the limits set by DM direct detection experiments (Figure 8.6). For a given value of m_χ, the limit on $\mathcal{B}(h \to \text{inv.})$ can be translated into a limit on λ_χ. This can be used to determine the scattering cross section for DM-nucleon interactions [2]. This relies on knowing the nucleon form factor; the most recent result for this [114] is used here (notably, this provides a very large reduction in the uncertainty on the form factor with respect to previous calculations). This relationship between $\mathcal{B}(h \to \text{inv.})$ and the DM-nucleon scattering cross section is interpreted in an effective field theory framework, assuming that the theory is renormalized around ~ 1 TeV, well above the scale probed here.

In general, χ could be a scalar, Majorana fermion[66], or vector. The fermion and vector cases are non-renormalizable and have non-unitarity effects at some scales. For vector DM, this results in non-physical behavior at low m_χ, so this interpretation is not included here. It is possible to evade these problems with a UV completion of the theory which includes a renormalizable mass generation mechanism [115], but this can modify the relationship between $\mathcal{B}(h \to \text{inv.})$ and the DM-nucleon scattering cross section. The exact Lagrangians for the models used here are:

---

[66] Dirac fermion DM cannot be coupled to the Higgs doublet in the usual way if it is assumed to be a SU(2) singlet.

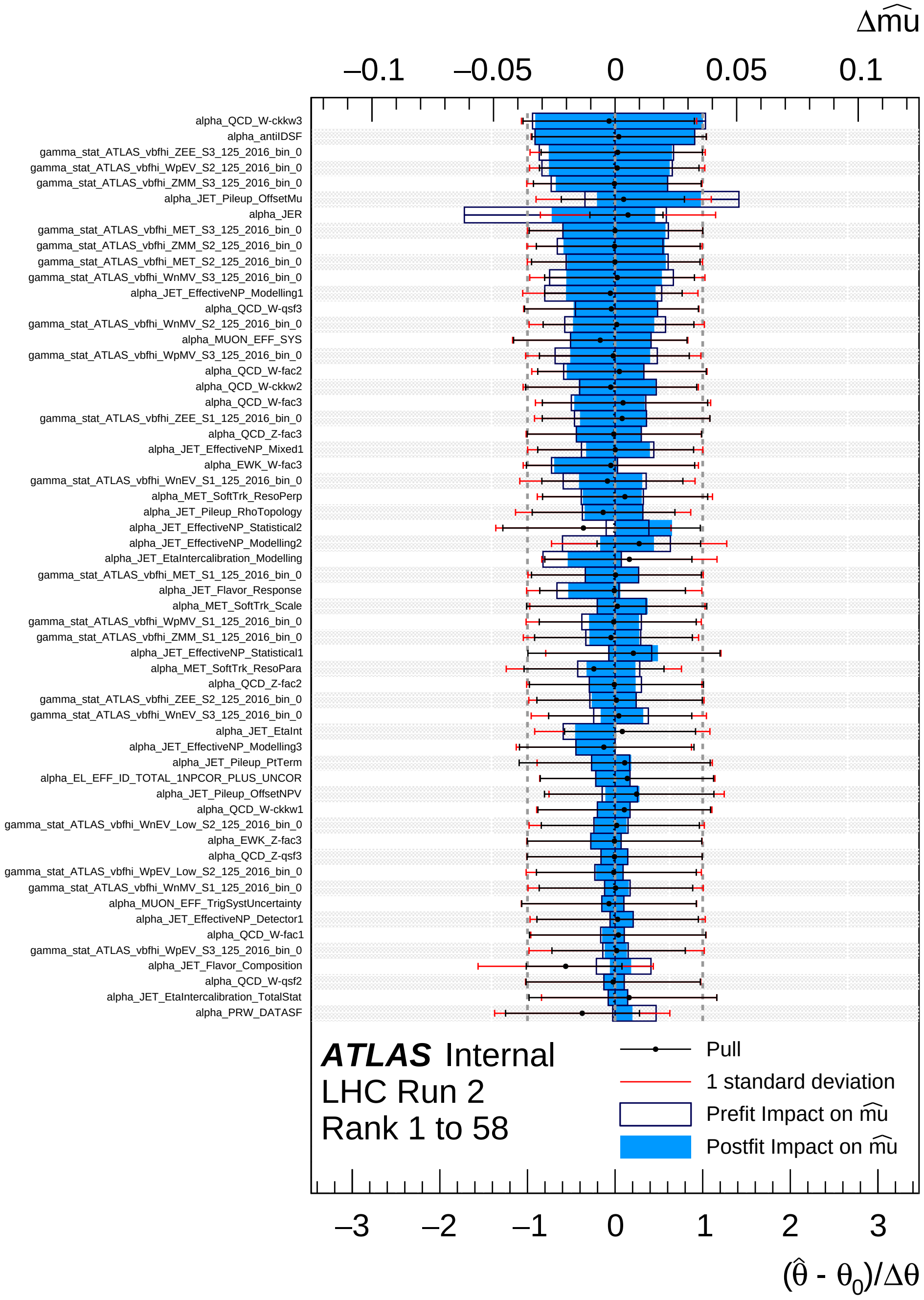

Figure 8.5: Impact of the dominant nuisance parameters on the best-fit value of μ, for the full fit to all regions. This is defined as the change in $\hat{\mu}$ under a 1σ variation of the NP. The pull of each NP (i.e. the deviation of its best-fit value from its nominal value) is also shown.

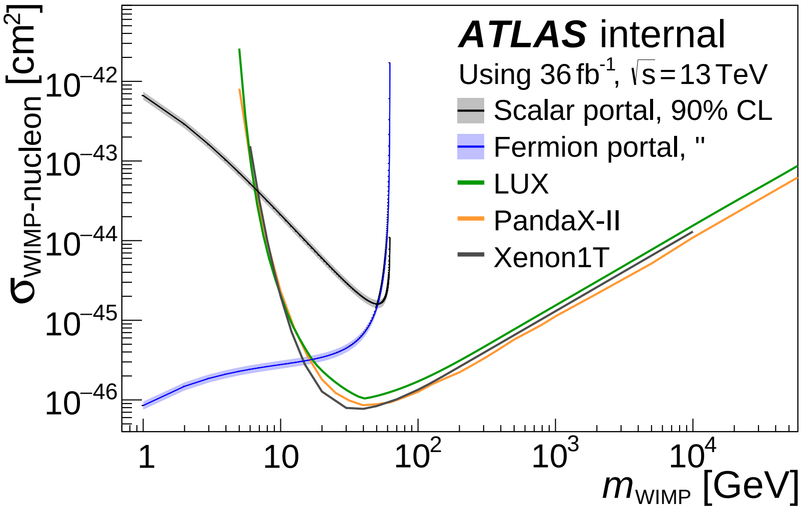

Figure 8.6: Upper limits on the DM-nucleon scattering cross section for scalar and Majorana fermion dark matter. The tightest current limits from direct detection experiments are also shown. The direct detection curves assume a certain local density of dark matter, based on astrophysical measurements.

$$\Delta\mathcal{L}_{\text{scalar}} = -\frac{1}{2}m_\chi^2\chi^2 - \frac{1}{4}\lambda_\chi H^\dagger H\chi^2 \tag{8.7}$$

$$\Delta\mathcal{L}_{\text{fermion}} = -\frac{1}{2}m_\chi\bar{\chi}\chi - \frac{1}{4}\frac{\lambda_\chi}{\Lambda}H^\dagger H\bar{\chi}\chi \tag{8.8}$$

As Figure 8.6 shows, this analysis provides very complementary coverage with respect to direct detection experiments. At low m_χ, direct detection becomes less sensitive due to the difficulty of measuring nuclear recoil from such a light particle. However, colliders are *more* sensitive at low m_χ since it allows more phase space for the decays, and there is no attempt to reconstruct the DM particle itself.

8.3.2 Heavy Scalar Mediators

In addition to the SM Higgs, this result can be interpeted in terms of a completely general scalar mediator. This could be a heavy Higgs, or some entirely different scalar boson. Since there is no spin structure, the kinematics of the signal generally depend only on the mass of the mediator (assuming it is produced on-shell). In order to perform this interpretation, alternate signal MC samples were generated which were the same as the nominal VBF Higgs production, but with the mass of the Higgs modified (all couplings are kept the same). The production cross section of such a scalar is not known (since no assumption about a model is being made), so limits cannot be placed directly on the invisible branching ratio of this particle. Instead, limits are placed on the product of the production cross section (σ) and invisible branching ratio of a general massive scalar ($\mathcal{B}_{\text{inv.}}$).

This is implemented in the fit exactly the same way as the SM Higgs case, but the signal yield prediction is replaced with the modified-mass MC prediction (this is done for each mass point individually). In the fit, μ is now interpreted to also include the production cross section for the new scalar. At higher mediator masses, the kinematics are very similar to the SM higgs case, but the jets tend to be more forward, as shown in Figure 8.7. This results in longer tails for the $\Delta\eta_{jj}$ and m_{jj} distributions relative to the 125 GeV scenario. In principle a dedicated search could be re-optimized for these slightly different kinematics, but this was considered unnecessary here. The upper limits on $\sigma \times \mathcal{B}_{\text{inv.}}$ from this procedure are shown in Figure 8.8. These are shown at a 95% confidence level using the CL_{s} method.

This interpretation makes only one assumption about the underlying theory of the scalar mediator: that it be produced only in VBF. It only needs to couple to W/Z and have mass greater than $2m_\chi$, in order to be produced on-shell via VBF. Heavy Higgs MC samples were used to model it, but there is no difference between this case and a general scalar, aside from the production cross

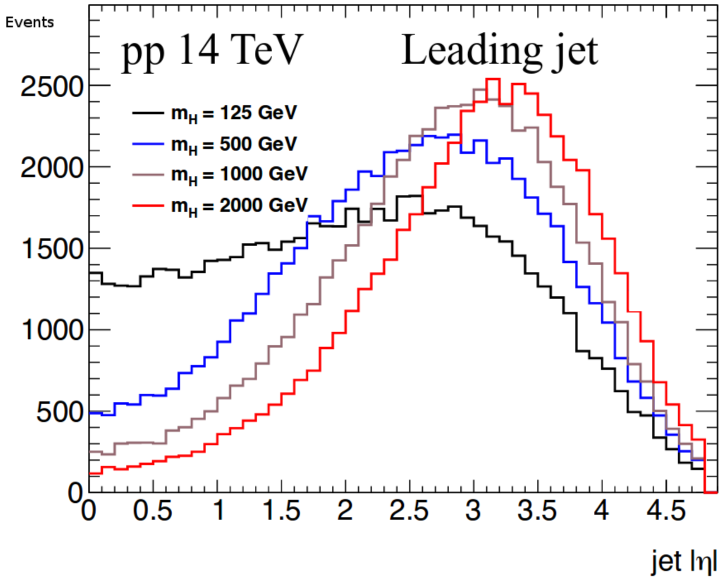

Figure 8.7: Leading jet $|\eta|$ distributions for heavy scalar production via VBF. These are derived from truth-level simulation of VBF Higgs production with modified Higgs mass. The only cut applied in this plot is leading jet $p_{\text{T}} > 20$ GeV.

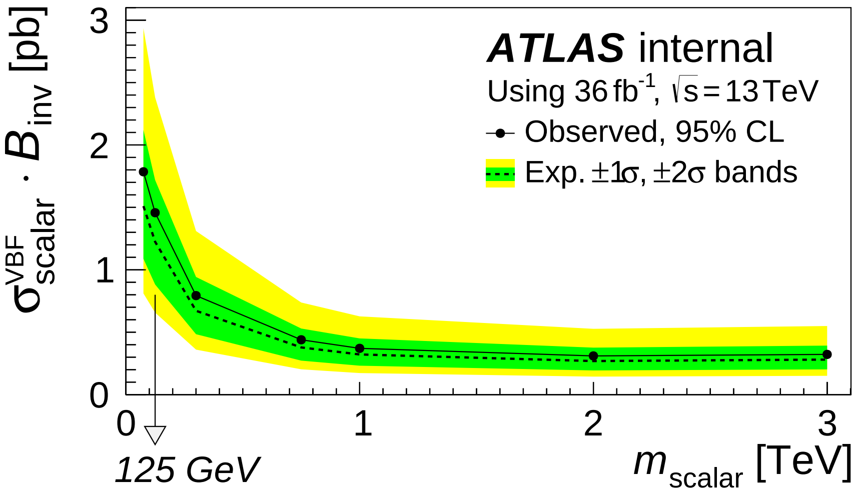

Figure 8.8: Limits on cross section times invisible branching ratio ($\sigma\mathcal{B}_{\text{inv.}}$) for an arbitrary heavy scalar produced via VBF.

section. So, any theory which predicts such a scalar is constrained by these limits. A scalar which can be produced in both VBF and via other mechanisms that would appear in this channel would have stricter limits; a case-by-case reinterpretation would be needed for these.

8.3.3 Two Higgs Doublet Model with a Pseudoscalar

Since the space of dark matter models (even just "simplified models") is so large, it is often useful to interpret results in terms of a few specific benchmark scenarios. One such model which is currently relevant is a two-Higgs-doublet model (2HDM) with an additional pseudoscalar field P [116]. This is considered because it is a simple model with a wide range of phenomenology which many collider searches can be sensitive to. This is a Type II 2HDM[67], where the P mixes with the pseudoscalar degree of freedom from the second Higgs doublet. After electroweak symmetry breaking, this results in two neutral Higgs bosons (h and H), two charged Higgs bosons (H^+ and H^-), and two pseudoscalars (a and A). The observed Higgs boson at 125 GeV is assumed to correspond to h, and the alignment limit is assumed. Finally, a DM candidate χ is included, which is taken to be fermionic (though its spin doesn't really matter) and couple only to P.

If a and χ are sufficiently light, then this model predicts a significant branching ratio of h to a final state consisting entirely of χs. Assuming m_χ is small (less than m_a and $m_h/2$), there are four possible cases:

---

[67] For an overview of two-Higgs-doublet models, see for example [117] and the references therein.

- $\boldsymbol{m_a < 2m_\chi}$: a is too light to decay into a pair of χ, so it will always decay into lighter SM particles instead. This scenario is not excluded by invisible Higgs constraints, but can be by searches in visible channels.

- $\boldsymbol{2m_\chi < m_a < m_h/2}$: h can undergo a 2-body decay to aa which then decays to $\chi\chi\chi\chi$. This results in a significant value of $\mathcal{B}(h \to \text{inv.})$, so it can be excluded depending on the parameter choices.

- $\boldsymbol{m_h/2 < m_a < m_h - 2m_\chi}$: h can undergo a 3-body decay to $a\chi\chi$, which then results in $\chi\chi\chi\chi$. Again, this can result in a significant value of $\mathcal{B}(h \to \text{inv.})$, depending on parameter choices.

- $\boldsymbol{m_a > m_h - 2m_\chi}$: Decays of h to $\chi\chi\chi\chi$ are in principle still possible via aa where both are off-shell, but this is quite suppressed. The existing phonomenology literature does not provide a description for this process, so it is not considered here.

Any set of parameters which results in a value of $\mathcal{B}(h \to \text{inv.})$ which is too large is excluded. Since the current tightest limit on this quantity (23%) comes from the ATLAS run 1 combination of invisible and visible channels [7], that result is used for this interpretation rather than the Run 2 VBF-only result. Since the total decay width of the SM Higgs is roughly 4.1 MeV, this means any invisible decay width over 1.2 MeV is ruled out.

Exclusion contours are calculated numerically from the decay width formulae given in [116]. This is done for a benchmark parameter set[68]: $m_\chi = 10$ GeV, $y_\chi = 1$, $m_A = m_H = m_{H^\pm}$, and $\lambda_{P1} = \lambda_{P2} = \lambda_3 = 3$. For this parameter set (and any set with $y_\chi \sim 1$), a decays overwhelmingly into $\chi\chi$ if kinematically allowed. So, any h decays to aa or $a\chi\chi$ are effectively always invisible decays. The results are shown in Figure 8.9, along with several other dark matter searches.

This VBF analysis provides quite complementary coverage in the parameter space of this model with respect to other collider searches. Most other searches are sensitive to a, A, or H production via their coupling to quarks, which results in a dependence on $\tan\beta$ (the ratio of the vacuum expectation values of the two Higgs doublets). At high values of $\tan\beta$, other analyses lose sensitivity while the VBF channel is unaffected. The "trade-off" is that VBF is only sensitive to low values of m_a (since it relies on decays of the SM Higgs), while other searches have no such limitations, aside from the center-of-mass energy of the collisions. The same principle holds for m_A: the VBF analysis is largely insensitive to this parameter while other searches cannot reach higher masses. There is some loss of

---

[68] See [116] for the exact definition of these parameters.

sensitivity in the region of $m_A \approx 1.2$ TeV; this is due to a cancellation in the effective *haa* coupling, which is implicitly dependent on m_A. The exact value of m_A at which this cancellation occurs depends on the choice of the other parameters, and for some choices it does not occur at all.

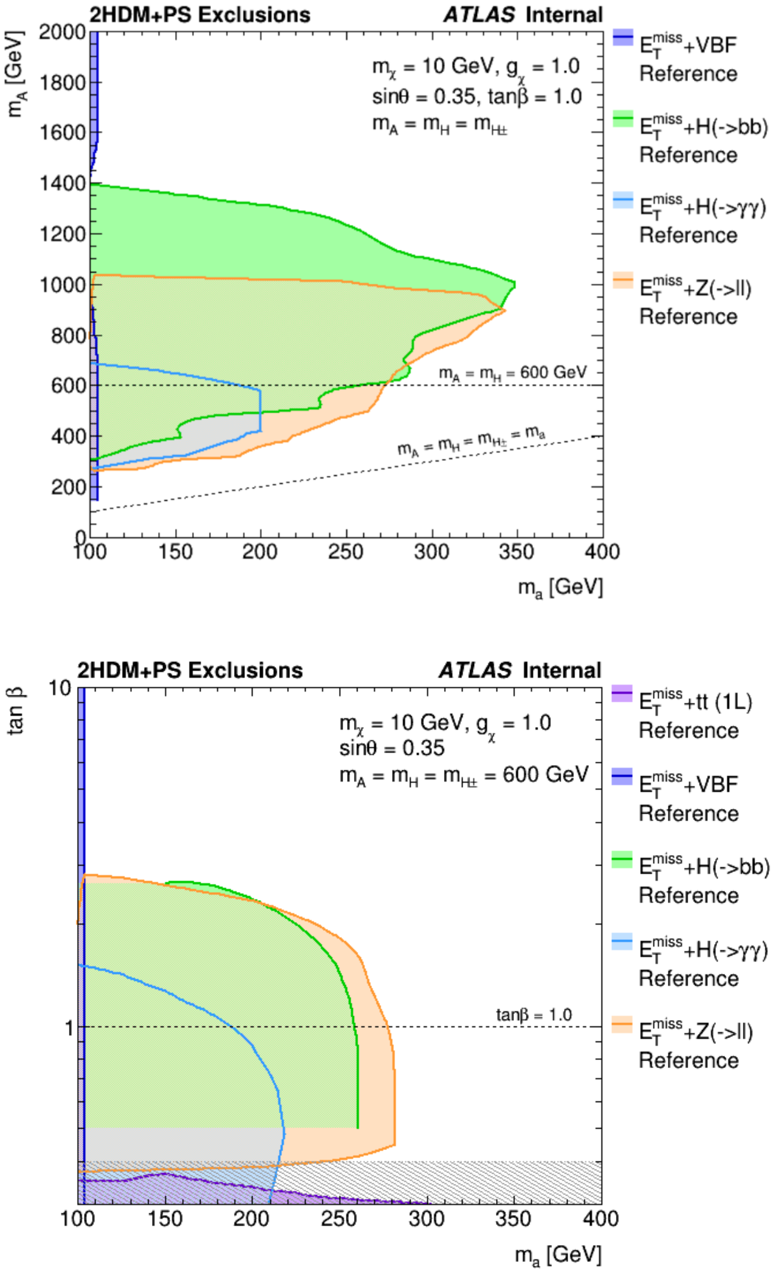

Figure 8.9: Exclusions for the benchmark 2HDM from this analysis and other dark matter searches at ATLAS. **Top:** The m_A-m_a plane, with $\tan\beta$ fixed to 1. **Bottom:** The $\tan\beta$-m_a plane, with m_A fixed to 600 GeV. The other parameters of the model are held fixed to the benchmark values described in the text.

Chapter 9

Conclusion

This thesis has presented a search for dark matter production via VBF in pp collisions at $\sqrt{s} =$ 13 TeV. No evidence of such production was found; the Standard Model predictions describe the data accurately to well within the uncertainties. This was interpreted in terms of a SM Higgs mediator, and an upper limit of 31% was placed on its invisible branching ratio at a 95% confidence interval. This limit was used to place constraints on the DM-nucleon scattering cross section within the context of an effective field theory, in order to draw comparisons with direct detection experiments. The results were also interpreted in terms of non-SM mediator models, including a generic heavy scalar which couples weakly and a benchmark 2HDM with an additional pseudoscalar.

Looking Ahead

The work presented in this thesis is intended to lay the groundwork for an even more sophisticated analysis of the full ATLAS Run 2 dataset. In the process of carrying out this analysis, several points have been identified which will be crucial to maximizing the sensitivity of a similar future search. The most important of these is the estimation of systematic uncertainties, as discussed in Section 8.2.3. In many cases (particularly some of the theoretical uncertainties), the available MC samples had insufficient statistics to estimate these effects accurately. There are two ways to address this for the future: produce more MC simulation (which is not always feasible due to computing limitations), or come up with a different method for estimating the uncertainties rather than computing them directly. This could involve extrapolating from a higher-statistics region as was done for some of the strong-produced W/Z uncertainties, or using physical arguments to impose constraints.

The issue of Monte Carlo statistics will continue to become more problematic as ATLAS gathers ever-increasing quantities of data. Even with 36 fb^{-1}, this is already a significant limiting factor in the sensitivity of the search. Again, this can be solved in principle by brute force, but the resources available may not be able to support this approach. Instead, it will likely become necessary to use more sophisticated techniques for using the existing resources more efficiently. This means using methods like generator-level filtering to simulate more events in the phase space of interest and less elsewhere. Using a lower-order generator to speed up the calculations is possible in concept, however this comes at the cost of substantially increased theoretical uncertainties. Furthermore, it would not reduce the computing cost of modeling detector interactions and reconstruction, which actually takes the majority of the computing resources.

As the size of the ATLAS dataset increases, it will become possible to further leverage the shape of kinematic variables such as m_{jj} to discriminate signal from background. This could include increasing the number of bins and going higher in the distribution to improve the signal-to-background ratio. If reliable calculations of the m_{jj} distribution become available, then it could also become possible to constrain this shape from theory. This would allow the use of low-m_{jj} bins with large statistics to constrain the higher ones.

Finally, if theoretical uncertainties start to dominate the sensitivity of the search in the future (and the reason isn't MC statistics), then it could become necessary to take a closer look at the way they are estimated. Factor-of-two scale variations are ad-hoc and generally do not capture the physics of higher-order effects. Empirically they are known to give effects of roughly the appropriate size, which is not a problem when the uncertainties are subdominant. However, if these arbitrarily-computed uncertainties are driving the analysis, it would call into question the validity of the result. The only remedy here is to look explicitly at higher-order effects, along the lines of the approach taken in [111].

In all, there is still a lot of space left to be explored, in terms of invisible Higgs decays and dark matter in general. In addition to the VBF channel, ATLAS and CMS are both continuing the search for dark matter in a broad variety of final states. Together with the ever-improving results from direct and indirect detection experiments, this makes constraints on dark matter tighter every year. As we continue to explore the high-energy frontier, there could be answers to this and other questions in fundamental physics hidden in the data. Exactly what those might be remains to be discovered.

Bibliography

[1] S. Kanemura, S. Matsumoto, T. Nabeshima, and N. Okada, *Can WIMP Dark Matter overcome the Nightmare Scenario?*, Phys. Rev. D **8**2 (2010) 055026, arXiv:1005.5651 [hep-ph].

[2] A. Djouadi, O. Lebedev, Y. Mambrini, and J. Quevillon, *Implications of LHC searches for Higgs–portal dark matter*, Phys. Lett. B **7**09 (2012) 65–69, arXiv:1112.3299 [hep-ph].

[3] G. Belanger et al., *Status of Invisible Higgs decays*, Phys. Lett. B **7**23 (2013) 340–347, arXiv:1302.5694 [hep-ph].

[4] D. Curtin et al., *Exotic decays of the 125 GeV Higgs boson*, Phys. Rev. D **9**0 (2014) 075004, arXiv:1312.4992 [hep-ph].

[5] ATLAS Collaboration, *Search for invisible decays of a Higgs boson using vector-boson fusion in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, JHEP **0**1 (2016) 172, arXiv:1508.07869 [hep-ex].

[6] CMS Collaboration, *Searches for invisible decays of the Higgs boson in pp collisions at* $\sqrt{s} = 7, 8,$ *and* 13 *TeV*, JHEP **0**2 (2017) 135, arXiv:1610.09218 [hep-ex].

[7] ATLAS Collaboration, *Constraints on new phenomena via Higgs boson couplings and invisible decays with the ATLAS detector*, JHEP **1**1 (2015) 206, arXiv:1509.00672 [hep-ex].

[8] M. Born, W. Heisenberg, and P. Jordan, *Zur Quantenmechanik. II.*, Zeitschrift für Physik **3**5 (1926) no. 8, 557–615.

[9] P. A. M. Dirac, *The quantum theory of the emission and absorption of radiation*, Proceedings of the Royal Society of London A **1**14 (1927) no. 767, 243–265.

[10] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory.* Addison-Wesley, 1995.

[11] S. Weinberg, *The Quantum Theory of Fields, Volume 1: Foundations.* Cambridge University Press, 1995.

[12] M. Srednicki, *Quantum Field Theory.* Cambridge University Press, 2007.

[13] R. P. Feynman, *Mathematical formulation of the quantum theory of electromagnetic interaction*, Phys. Rev. **80** (1950) 440–457.

[14] M. Y. Han and Y. Nambu, *Three-triplet model with double* SU(3) *symmetry*, Phys. Rev. **139** (1965) B1006–B1010.

[15] M. Gell-Mann, *A schematic model of baryons and mesons*, Physics Letters **8** (1964) 214–215.

[16] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, in *Developments in the Quark Theory of Hadrons. Vol. 1. 1964 - 1978*, D. Lichtenberg and S. P. Rosen, eds., pp. 22–101. 1964. `http://inspirehep.net/record/4674/files/cern-th-412.pdf`.

[17] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343–1346.

[18] H. D. Politzer, *Reliable perturbative results for strong interactions?*, Phys. Rev. Lett. **30** (1973) 1346–1349.

[19] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Physics Letters **12** (1964) 132–133.

[20] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[21] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321–323.

[22] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585–587.

[23] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **7**16 (2012) 1, arXiv:1207.7214 [hep-ex].

[24] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **7**16 (2012) 30, arXiv:1207.7235 [hep-ex].

[25] ATLAS and CMS Collaborations, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at* $\sqrt{s} = 7$ *and* 8 *TeV*, JHEP **0**8 (2016) 045, arXiv:1606.02266 [hep-ex].

[26] S. L. Glashow, *The Renormalizability of Vector Meson Interactions*, Nucl. Phys. **10** (1959) 107–117.

[27] A. Salam and J. C. Ward, *Weak and Electromagnetic Interactions*, Nuovo Cimento **11** (1959) 568–577.

[28] UA1 Collaboration, *Experimental observation of isolated large transverse energy electrons with associated missing energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. B **122** (1983) 103–116.

[29] UA2 Collaboration, *Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider*, Phys. Lett. B **122** (1983) 476–485.

[30] UA1 Collaboration, *Experimental observation of lepton pairs of invariant mass around 95 GeV/*c^2 *at the CERN SPS collider*, Phys. Lett. B **126** (1983) 398–410.

[31] UA2 Collaboration, *Evidence for* $Z^0 \rightarrow e^+e^-$ *at the CERN pp collider*, Phys. Lett. B **129** (1983) 130–140.

[32] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, arXiv:1610.07922 [hep-ph].

[33] D. Clowe et al., *A direct empirical proof of the existence of dark matter*, Astrophys. J. **6**48 (2006) L109–L113, arXiv:astro-ph/0608407 [astro-ph].

[34] MACHO Collaboration, *The MACHO project: Microlensing results from 5.7 years of LMC observations*, Astrophys. J. **5**42 (2000) 281–307, arXiv:astro-ph/0001272 [astro-ph].

[35] EROS-2 Collaboration, *Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds*, Astron. Astrophys. **4**69 (2007) 387–404, arXiv:astro-ph/0607207 [astro-ph].

[36] R. Peccei and H. R. Quinn, *CP Conservation in the Presence of Pseudoparticles*, Phys. Rev. Lett. **3**8 (1977) 1440–1443.

[37] R. Peccei and H. R. Quinn, *Constraints imposed by CP conservation in the presence of pseudoparticles*, Phys. Rev. D **1**6 (1977) 1791–1797.

[38] LUX Collaboration, *Results on the Spin-Dependent Scattering of Weakly Interacting Massive Particles on Nucleons from the Run 3 Data of the LUX Experiment*, Phys. Rev. Lett. **1**16 (2016) 161302, arXiv:1602.03489 [hep-ex].

[39] LUX Collaboration, *Results from a Search for Dark Matter in the Complete LUX Exposure*, Phys. Rev. Lett. **1**18 (2017) 021303, arXiv:1608.07648 [astro-ph.CO].

[40] XENON Collaboration, *First Dark Matter Search Results with XENON1T*, Phys. Rev. Lett. **1**19 (2017) 181301, arXiv:1705.06655 [astro-ph.CO].

[41] Fermi-LAT Collaboration, *The Large Area Telescope on the Fermi Gamma-ray Space Telescop Mission*, Astrophys. J. **6**97 (2009) 1071–1102, arXiv:0902.1089 [astro-ph.IM].

[42] AMS Collaboration, *High Statistics Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5–500 GeV with the Alpha Magnetic Spectrometer on the International Space Station*, Phys. Rev. Lett. **1**13 (2014) 121101.

[43] AMS Collaboration, *Antiproton Flux, Antiproton-to-Proton flux Ratio, and Properties of Elementary Particle Fluxes in Primary Cosmic Rays Measured with the Alpha Magnetic Spectrometer on the International Space Station*, Phys. Rev. Lett. **1**17 (2016) 091103.

[44] J. Conrad, *Indirect Detection of WIMP Dark Matter: a compact review*, in *Interplay between Particle and Astroparticle physics (IPA2014) London, United Kingdom, August 18-22, 2014*. 2014. arXiv:1411.1925 [hep-ph].

[45] ATLAS Collaboration, *Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector*, JHEP **0**1 (2018) 126, arXiv:1711.03301 [hep-ex].

[46] ATLAS Collaboration, *Search for new phenomena in dijet events using* $37\ fb^{-1}$ *of pp collision data collected at* $\sqrt{s} = 13\ TeV$ *with the ATLAS detector*, Phys. Rev. D **9**6 (2017) 052004, arXiv:1703.09127 [hep-ex].

[47] J. Abdallah et al., *Simplified Models for Dark Matter Searches at the LHC*, Phys. Dark Univ. **9**-10 (2015) 8–23, arXiv:1506.03116 [hep-ph].

[48] Planck Collaboration, *Planck 2013 Results. XVI. Cosmological parameters*, Astron. Astrophys. **5**71 (2014) A16, arXiv:1303.5076 [astro-ph.CO].

[49] A. Beniwal et al., *Combined analysis of effective Higgs portal dark matter models*, Phys. Rev. D **9**3 (2016) no. 11, 115016, arXiv:1512.06458 [hep-ph].

[50] J. Barreto et al., *Direct Search for Low Mass Dark Matter Particles with CCDs*, Phys. Lett. B **7**11 (2012) 264–269, arXiv:1105.5191 [astro-ph.IM].

[51] SuperCDMS Collaboration, *New Results from the Search for Low-Mass Weakly Interacting Massive Particles with the CDMS Low Ionization Threshold Experiment*, Phys. Rev. Lett. **1**16 (2016) 071301, arXiv:1509.02448 [astro-ph.CO].

[52] PICO Collaboration, *Dark Matter Search Results from the PICO-60* C_3F_8 *Bubble Chamber*, Phys. Rev. Lett. **1**18 (2017) 251301, arXiv:1702.07666 [astro-ph.CO].

[53] PandaX-II Collaboration, *Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment*, Phys. Rev. Lett. **1**19 (2017) 181302, arXiv:1708.06917 [astro-ph.CO].

[54] L. R. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001. `https://cds.cern.ch/record/1129806`. This report is an abridged version of the LHC Design Report (CERN-2004-003).

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

[56] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, Vol. 1.* CERN, Geneva, 1997. `https://cds.cern.ch/record/331063`.

[57] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report, Vol. 2.* CERN, Geneva, 1997. `https://cds.cern.ch/record/331064`.

[58] G. Aad et al., *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007.

[59] A. Ahmad et al., *The Silicon microstrip sensors of the ATLAS semiconductor tracker*, Nucl. Instrum. Meth. **A**578 (2007) 98–118.

[60] E. Abat et al., *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: Design and performance*, JINST **3** (2008) P02013.

[61] E. Abat et al., *The ATLAS TRT Barrel Detector*, JINST **3** (2008) P02014.

[62] E. Abat et al., *The ATLAS TRT end-cap detectors*, JINST **3** (2008) P10003.

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

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

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

[66] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. C **72** (2012) 1849, arXiv:1110.1530 [hep-ex].

[67] ATLAS Collaboration, *Search for light dijet resonances with the ATLAS detector using a Trigger-object Level Analysis in LHC pp collisions at* $\sqrt{s} = 13$ *TeV*, ATLAS-CONF-2016-030, 2016. `https://cds.cern.ch/record/2161135`.

[68] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv:1603.02934 [hep-ex].

[69] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-*k_t *jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[70] M. Cacciari and G. P. Salam, *Dispelling the* N^3 *myth for the* k_t *jet-finder*, Phys. Lett. **B641** (2006) 57–61, arXiv:0512210 [hep-ph].

[71] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv:1703.09665 [hep-ex].

[72] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581, arXiv:1510.03823 [hep-ex].

[73] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, arXiv:physics/0703039.

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

[75] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton–proton collision data*, ATLAS-CONF-2016-024, 2016. `https://cds.cern.ch/record/2157687`.

[76] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, 2008. `https://cds.cern.ch/record/1099735`.

[77] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444–450.

[78] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv:1704.07983 [hep-ex].

[79] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv:1407.5063 [hep-ex].

[80] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **76** (2016) 292, arXiv:1603.05598 [hep-ex].

[81] R. Achenbach et al., *The ATLAS level-1 calorimeter trigger*, JINST **3** (2008) P03001.

[82] ATLAS Collaboration, *Performance of the ATLAS muon trigger in pp collisions at* $\sqrt{s} = 8$ *TeV*, Eur. Phys. J. C **7**5 (2015) 120, arXiv:1408.3179 [hep-ex].

[83] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **7**7 (2017) 317, arXiv:1611.09661 [hep-ex].

[84] CMS Collaboration, *Search for invisible decays of Higgs bosons in the vector boson fusion and associated ZH production modes*, Eur. Phys. J. C **7**4 (2014) 2980, arXiv:1404.1344 [hep-ex].

[85] A. L. Read, *Presentation of search results: the* CL_s *technique*, J. Phys. G **2**8 (2002) no. 10, 2693–2704.

[86] S. Agostinelli et al., *GEANT4 - a simulation toolkit*, Nucl. Instrum. Meth. **A**506 (2003) 250–303.

[87] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **0**2 (2009) 007, arXiv:0811.4622 [hep-ph].

[88] S. Alioli, P. Nason, C. Oleari, and E. Re, *NLO Higgs boson production via gluon fusion matched with shower in POWHEG*, JHEP **0**4 (2009) 002, arXiv:0812.0578 [hep-ph].

[89] P. Nason and C. Oleari, *NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG*, JHEP **0**2 (2010) 037, arXiv:0911.5299 [hep-ph].

[90] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, arXiv:0603175 [hep-ph].

[91] T. Sjostrand, S. Mrenna, and P. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **1**78 (2008) 852–867, arXiv:0710.3820 [hep-ph].

[92] R. D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **0**4 (2015) 040, arXiv:1410.8849 [hep-ph].

[93] A. Denner, S. Dittmaier, S. Kallweit, and A. Mück, *HAWK 2.0: A Monte Carlo program for Higgs production in vector-boson fusion and Higgs strahlung at hadron colliders*, Comput. Phys. Commun. **1**95 (2015) 161–171, arXiv:1412.5390 [hep-ph].

[94] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **1**2 (2008) 039, arXiv:0808.3674 [hep-ph].

[95] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **1**08 (2012) 111601, arXiv:1111.5206 [hep-ph].

[96] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorization*, JHEP **03** (2008) 038, arXiv:0709.1027 [hep-ph].

[97] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, JHEP **0**4 (2013) 027, arXiv:1207.5030 [hep-ph].

[98] S. Höche, F. Krauss, S. Schumann, and F. Siegert, *QCD matrix elements and truncated showers*, JHEP **0**5 (2009) 053, arXiv:0903.1219 [hep-ph].

[99] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, *NLO matrix elements and truncated showers*, JHEP **0**8 (2011) 123, arXiv:1009.1127 [hep-ph].

[100] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **1**1 (2001) 063, arXiv:hep-ph/0109231 [hep-ph].

[101] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. D **6**9 (2004) 094008, arXiv:hep-ph/0312266 [hep-ph].

[102] ATLAS Collaboration, *Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2*, ATL-PHYS-PUB-2016-003, 2016. `https://cds.cern.ch/record/2120133`.

[103] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **0**7 (2014) 079, arXiv:1405.0301 [hep-ph].

[104] ATLAS Collaboration, *Measurement of the Inelastic Proton–Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **1**17 (2016) 182002, arXiv:1606.02625 [hep-ex].

[105] R. D. Ball et al., *A determination of parton distributions with faithful uncertainty estimation*, Nucl. Phys. B **8**09 (2009) 1–63, arXiv:0808.1231 [hep-ph].

[106] J. M. Campbell and R. K. Ellis, *An update on vector boson pair production at hadron colliders*, Phys. Rev. D **6**0 (1999) 113006, arXiv:hep-ph/9905386 [hep-ph].

[107] J. M. Campbell, R. K. Ellis, and C. Williams, *Vector boson pair production at the LHC*, JHEP **0**7 (2011) 018, arXiv:1105.0020 [hep-ph].

[108] I. W. Stewart and F. J. Tackmann, *Theory Uncertainties for Higgs and Other Searches Using Jet Bins*, Phys. Rev. D **8**5 (2012) 034011, arXiv:1107.2117 [hep-ph].

[109] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C **5**8 (2008) 639–707, arXiv:0803.0883 [hep-ph].

[110] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **7**6 (2016) 196, arXiv:1512.01178 [hep-ph].

[111] J. M. Lindert et al., *Precise predictions for V+jets dark matter backgrounds*, Eur. Phys. J. C **7**7 (2017) no. 12, 829, arXiv:1705.04664 [hep-ph].

[112] J. Neyman and E. S. Pearson, *On the Problem of the Most Efficient tests of Statistical Hypothesis*, Philosophical Transactions of the Royal Society A **2**31 (1933) 289–337.

[113] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **7**1 (2011) 1554, arXiv:1007.1727 [physics.data-an]. [Erratum: Eur. Phys. J. C 73 (2013) 2501].

[114] M. Hoferichter, P. Klos, J. Menéndez, and A. Schwenk, *Improved limits for Higgs-portal dark matter from LHC searches*, Phys. Rev. Lett. **1**19 (2017) 181803, arXiv:1708.02245 [hep-ph].

[115] S. Baek, P. Ko, and W.-I. Park, *Invisible Higgs Decay Width vs. Dark Matter Direct Detection Cross Section in Higgs Portal Dark Matter Models*, Phys. Rev. D **9**0 (2014) 055014, `arXiv:1405.3530` [hep-ph].

[116] M. Bauer, U. Haisch, and F. Kahlhoefer, *Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators*, JHEP **0**5 (2017) 138, `arXiv:1701.07427` [hep-ph].

[117] G. C. Branco et al., *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **5**16 (2012) 1–102, `arXiv:1106.0034` [hep-ph].