Detector-corrected measurements of missing transverse energy with the ATLAS detector and constraints on new physics models using Contur

Yoran Yeh
Department of Physics and Astronomy
UCL

CERN-THESIS-2024-319
//2024

A thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
of
University College London

November 2024

Declaration

I, Yoran Yeh, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Parts of this work have been published in the following articles.

1. ATLAS Collaboration, *Differential cross-sections for events with missing transverse momentum and jets measured with the ATLAS detector in 13 TeV proton-proton collisions*, JHEP 08 (2024) 223.
   Main contributions from the author include the development of the analysis software, unfolding, preparation of unfolded distributions with experimental data and theoretical predictions and statistical interpretations. Material of this publication is found in Chapter 5.

2. J. Butterworth, H. Debnath, P. Fileviez Pérez, Y. Yeh, *Dark Matter from Anomaly Cancellation at the LHC*, Phys. Rev D 110 (2024) 075001.
   Main contributions from the author include development of the re-interpretation software package CONTUR, the implementation a BSM signal model in CONTUR, and deriving BSM parameter constraints from collider measurements in CONTUR. Material of this publication is found in Chapter 6.

3. A. Buckley, L. Corpe, M. Filipovich, C. Gütschow, N. Rozinsky, S. Thor, Y. Yeh, J. Yellen, *Consistent, multidimensional differential histogramming and summary statistics with YODA 2*, arXiv: 2312.15070 [hep-ph], in preparation for journal submission to Scipost Physics.
   Main contribution from the author is a complete restructuring of the visualisation and plotting mechanism. Material of this publication is found in Chapter 6.

4. C. Bierlich, A. Buckley, J. Butterworth, C. Gütschow, L. Lönnblad, T. Procter, P. Richardson, Y. Yeh, *Robust Independent Validation of Experiment and Theory: RIVET version 4 release note*, arXiv: 2404.15984 [hep-ph], in preparation for journal submission to Scipost Physics.
   Main contribution from the author is a complete restructuring of the visualisation and plotting mechanism. Material of this publication is found in Chapter 6.

Yoran Yeh

Abstract

This thesis first presents differential cross-section measurements of events with large missing transverse energy in association with jets, using $140\,\mathrm{fb}^{-1}$ of data from $\sqrt{s} = 13\,\mathrm{TeV}$ proton-proton collisions delivered by the Large Hadron Collider (LHC) and measured with the ATLAS detector from 2015 to 2018. Auxiliary measurements with leptons or photons are made in the same kinematic regions, where correlated theoretical and experimental uncertainties cancel in ratios between cross-sections. This analysis is designed to be maximally model-independent and re-interpretable in the context of a wide range of theoretical predictions: the phase space is defined only in terms of the final state particles and results are corrected for detector effects with an iterative Bayesian method. Experimental data are compared to state-of-the-art Standard Model predictions in a likelihood fit that shows adequate modelling of most observables except the invariant dijet mass. The measured cross-sections are also compared to theories beyond the Standard Model. Constraints on a benchmark dark matter model show competitive limits to a dedicated beyond the Standard Model search using the same dataset, proving the effectiveness of using precision measurements in re-interpretation studies. Secondly, improvements are made to several open-source software tools that are at the heart of the LHC re-interpretation ecosystem: RIVET, YODA and CONTUR. The latter is used to set the first collider limits from precision measurements on a beyond the Standard Model scenario, in which a particle dark matter candidate arises naturally from gauge anomaly cancellation. Large parts of parameter space are constrained by the presented ATLAS measurement of missing transverse energy. Collider constraints are combined with results from cosmology and direct dark matter searches, while sensitivity studies for the future High-Luminosity LHC project show promising prospects to either discover or rule out this model in the near future.

Impact statement

Modern society is shaped by the desire of physicists to fathom Nature at subatomic scales. Material science, electronics and medicine are only a few examples of fields that are driven by advancements in our understanding of the building blocks of matter. The work presented in this thesis subjects our current best theory of fundamental particles, the Standard Model, to empirical tests using experimental data collected at the Large Hadron Collider (LHC) at CERN, and furthermore looks at what lies beyond the Standard Model.

The presence of dark matter – inferred from astrophysical observations – is one of the big gaps in our understanding of the Universe and occupies many researchers worldwide. This thesis shines light on studies of dark matter at particle accelerators, which is complementary to the domains of astrophysics and cosmology that looks for signs of dark matter in the cosmos with direct and indirect detection experiments. More specifically, the measurement of missing transverse energy with the ATLAS experiment at CERN that is performed in this work is used as a probe for the creation of dark matter at the Large Hadron Collider. The research carried out in this doctoral study argues that analysis preservation and re-interpretation are instrumental moving forward in the LHC program. The longevity and impact of experimental results is significantly enhanced if they can be re-used to squeeze as much information as possible from LHC collisions. The missing transverse energy measurement is designed to be optimally re-usable, and the effectiveness of this approach is demonstrated by re-interpreting the experimental data in terms of three different theoretical models for dark matter candidates. This approach also shows that precision cross-section measurements of Standard Model processes are complementary and competitive in terms of obtained limits to detector-level searches for physics beyond the Standard Model, with precision cross-section measurements having a wider range of applications from effective field theory fits to the tuning of Monte Carlo event generators.

From an experimental point of view, building the most powerful particle accelerator leads to many byproducts that quickly find applications outside the academic world or collider physics. The incredible demands in terms of technology and computation at the LHC have lead to innovations in for example medical physics and grid computing, and most famously the invention of the World Wide Web.

Acknowledgements

This work would not have been possible without the invaluable support, guidance, and encouragement of numerous people who contributed to both the project and the writing of this thesis.

First, I would like to thank Jon Butterworth for being the best supervisor I could have hoped for during my time at UCL. Your door was always open, and your depth of knowledge across so many topics made our discussions on physics and any other subject both inspiring and stimulating. Thanks to your engagement, vision and unwavering support during my PhD I look back on the past few years with a big smile.

I owe deep gratitude to Christian Gütschow, who showed me the ropes of being a particle physicist from the very start of my PhD onward. I am grateful for the physics insights and coding wizardry you shared through your mentorship, for your guidance on many collaborative projects, and, just as importantly, for the pints we shared after long days of work.

I also want to extend a heartfelt "thank you" and virtual bouquet of flowers to Louie Corpe, who taught me all the ins and outs of conducting experimental measurements and navigating the world of ATLAS. Your boundless enthusiasm and optimism were contagious, helping me to communicate the importance of our work to other colleagues and the broader community. I am also grateful for the opportunity to include some illustrations from your book, *Particle Physics for Babies*, in this thesis.

A big thank you to the members of the ATLAS $p_{\mathrm{T}}^{\mathrm{miss}}$+jets team, especially Jon, Chris, Louie, Martin Habedank, Aidan Kelly, Martin Klassen, Matouš Vozák, Sebastian Weber, and everyone else who contributed to this measurement. Tuesday mornings were never dull, thanks to the insightful discussions we had about this challenging analysis.

I am thankful to Pavel Fileviez Pérez and Hridoy Debnath for our collaborative project on dark matter candidates from anomaly cancellation. It was a true privilege to work with you, and your insights broadened my horizon beyond the realm of collider physics.

A huge thanks to the RIVET and YODA team, especially Andy Buckley, Christian Gütschow and Jamie Yellen for our team effort in redesigning the plotting infrastructure. Our in-person hackathons struck a healthy balance between work, adventurous outdoor activities and Scotch.

The CONTUR toolkit is always evolving due the contribution of many talented, motivated people. I want to thank Jon, Tony Yue, Peng Wang and Sihyun Jeon in particular for our shared work on this beautiful project.

A special thanks to the High Energy Physics group at UCL for providing such an engaging,

motivating, and friendly environment to be a part of. The friendship and camaraderie with my fellow doctoral students in the corridors of the physics building kept me sane during this PhD. To the incredible group of people from my time in Geneva, thank you for making it memorable with a good dose of swims in the lake, adventures in the Alps, and questionable drinks and snacks in DS17.

Jack and Jonny, I owe my appreciation for baked beans, rugby, and University Challenge to you; living with both of you has been an absolute delight. I'm grateful for our friendship and hope we can reunite on the football pitch or in the pub soon.

To my dear friends from home, I am thankful for all the laughter, distractions from work, and visits in London. Thanks to you, home was never far away, and I would not have been able to complete this work without all of you.

I want to thank my sister, Liza, and Daniel for helping me navigate the academic world since I started university and for being role models. Thank you to Kolja and Romy for always bringing a smile to everyone's face and for the fun weekend trips to London.

Susan, I had four years time to figure out what to write here, but I'm coming to realise that words fall short in expressing my gratitude for the happiness you bring into my life. I am forever grateful for your support and this work would not have been possible without you. Carlos and Monique, thank you for making me feel part of the family.

Finally, I would like to thank my parents, Eveline and Chow Fang, who always stood by my side and taught me from a young age to let personal interests be my guiding principle in life. Living in different countries was not always easy, but you supported me every step of the way. This PhD thesis is dedicated to you.

Contents

Chapter 1

Introduction

With the word 'ATLAS' in the title of this thesis, it is tempting to start by drawing parallels between modern society and Ancient Greece: we still organise Olympic games every four years, still run marathons and still ponder big questions of the Universe. This thesis is concerned with the latter. While the roots of modern physics, natural philosophy and astronomy indeed date back centuries to earlier civilizations, the most fundamental questions about the Universe remain partly unanswered. What are we made of? Why is there something rather than nothing, and what is the origin of this space and time? Finally, what does the Universe look like at the smallest or largest possible scales?

In fundamental physics we seek answers to such questions. More specifically, the description of Nature at the smallest scale is in the realm of particle physics. Ancient Greek philosophers already suggested that matter was made of small, indivisible objects called 'atomos'. This turned out to be a reasonable proposal. The Standard Model (SM) is the pinnacle of modern science and particle physics, providing the most accurate description to date of Nature's smallest scales. In the SM, point-like elementary particles interact via the exchange of force-carriers. The SM has been able to classify existing particles in a theoretical framework, but also to predict the existence of new particles that were then observed in experiments at later times.

Despite its immense successes, the SM also has shortcomings. For example, cosmological observations show that all SM particles make up a mere 5% of the total content of the Universe. The majority of the matter and energy content are in the form of 'dark matter' (DM) and 'dark energy', which the SM does not accommodate. The hunt for dark matter is one of the great challenges that (astro)particle physicists are currently occupied with. There are a multitude of theories that go beyond the Standard Model (BSM) to explain such unexplained phenomena, but as of yet none have been confirmed in observations.

Experimentally, the Standard Model is tested to unprecedented levels at the Large Hadron Collider (LHC). This is an era of measuring SM interactions at higher energies and with better precision than ever before. The crown jewel of both the LHC and the SM is the discovery of the Higgs boson in 2012, the existence of which was predicted nearly fifty years before. The LHC collides protons or heavy-ions to probe the Standard Model at energies way above the electroweak symmetry breaking scale and to open the way for precision studies of the strong interaction.

The LHC is also used as a laboratory to look for signs of physics beyond the SM. For example, since the beginning of LHC operations much time and effort has been spent to discover supersymmetry (SUSY). While SUSY has not been found experimentally, nor any other BSM theory, the null results have taught us much about which properties new physics can and cannot have. One of the important lessons to draw from a decade of LHC results, in my opinion, is that we need to further develop flexible approaches to look for signs of BSM physics.

Traditional searches for new physics are optimised to have a high signal-over-background ratio for a single BSM configuration. The parameter space of the model in question is constrained if no excess over the Standard Model background predictions is observed in the data. Such analyses are often a one-trick pony, because the obtained limits and experimental data are typically difficult to re-use, or "re-interpret", in the context of a *different* BSM model. Supermarkets like to advertise that they are replacing single-use carrier bags by re-usable bags. As LHC physicists, why should we not do the same? Over the last decade, this idea has become more popular and the work in this thesis build further on existing strategies for analysis re-interpretation.

Re-interpretation is the physics equivalent of re-usable canvas bags. This domain of HEP has significant importance as the LHC may very well be the last hadron collider for this generation of physicists, and we must do our very best to leave a legacy of physics data that future generations can benefit from. In fact, history has already proven the merits of re-usable data: results from deep inelastic scattering experiments are instrumental to LHC physics through their impact on parton distribution functions. The preservation of experimental results is key in the advancement of particle physics, as re-interpretation does not only have strong implications for BSM interpretation, but also in the context of SM theory predictions, effective field theories and indeed parton distribution functions. Major parts of the work for this PhD project involve the development of open-source preservation and re-interpretation tools in the context of collider physics.

In this thesis, we discuss why certain representation levels of experimental data are more suitable for re-interpretation than others, and for the same reason motivate an approach that keeps irreducible SM background processes as part of signal in the data. These ideas are exemplified in the design and results of a precision measurement of proton-proton collisions leading to transverse missing momentum and jets in the final state with the ATLAS detector, the meat and potatoes of this thesis. The analysis results are preserved in public databases for later re-use. In fact, I immediately demonstrate the validity of this approach by re-interpreting the public results from this new ATLAS measurement in the context of a BSM scenario that predicts a DM candidate from gauge anomaly cancellation.

Finally, the field of particle physics is strongly connected with other domains in physics. Advances in collider physics and the understanding of radioactive decays shaped the landscape of medical physics with radiation therapy, medical imaging and nuclear medicine. Perhaps counter-intuitively, particle physics provide the foundations for cosmology, which is the study of our Universe at the *largest* scales. Furthermore, an understanding of physical processes in astronomy shares overlap with particle physics. The focus of this thesis is predominantly on the realm of particle physics itself, but at the very end we see the connection between particle physics and cosmology through a theory of dark matter.

The structure of the thesis is as follows. Chapter 2 summarises the theoretical framework for this thesis, including a review of the Standard Model and motivations to look beyond the Standard

Model. Experimental techniques used in the LHC and the ATLAS experiment are discussed in Chapter 3. The aim of Chapter 4 is to find a suitable common ground to compare theoretical predictions to experimental observations. Chapter 5 contains the motivation, design and results of the missing energy plus jets measurement with the ATLAS detector. In Chapter 6 we discuss analysis preservation and re-interpretation, with details on a phenomenological study of a DM model using the CONTUR framework.

Chapter 2

Theoretical framework

The main goal of particle physics is to understand the fundamental nature of matter, time and space. Centuries of research have culminated in our best scientific description of the Universe at the smallest scale: the Standard Model (SM). This is a quantum field theory that accurately describes all known fundamental forces – except for gravity – and coherently classifies every fundamental particle that so far has been observed in nature. A major milestone in the success story of the SM was the discovery of the Higgs boson in 2012. However, the Higgs discovery is not the final chapter in particle physics. There are various shortcomings that motivate physics *beyond* the SM (BSM), and numerous attempts to unify new theories with currently unexplained phenomena in nature.

Section 2.1 is a review of the most important aspects of the SM relevant to this thesis. We present the particle content of the SM and their interactions via force-carrying mediators, before introducing the mechanism of spontaneous symmetry breaking that leads to the prediction of the Higgs boson. Several tests show that predictions from the Standard Model agree well with many experimental observations. Section 2.2 steps into the realm of BSM physics, with an emphasis on dark matter (DM). We discuss important shortcomings of the SM, after which the focus shifts to different models and experimental tests for DM.

2.1 The Standard Model

The Standard Model is the accepted paradigm in which particle physicists operate. While we call it "standard" now, it took some of the brightest minds more than a century to bring this theory of quantum fields where it is today, leading to the award of multiple Noble prizes along the way. From a theoretical point of view, I hope to convey with this section that the Standard Model is an elegant and powerful description of nature rooted in gauge theory. The real great success of the Standard Model is arguably its predictive power: theoretical predictions from the SM have stood the test of time through myriad experimental measurements made in the last decades.

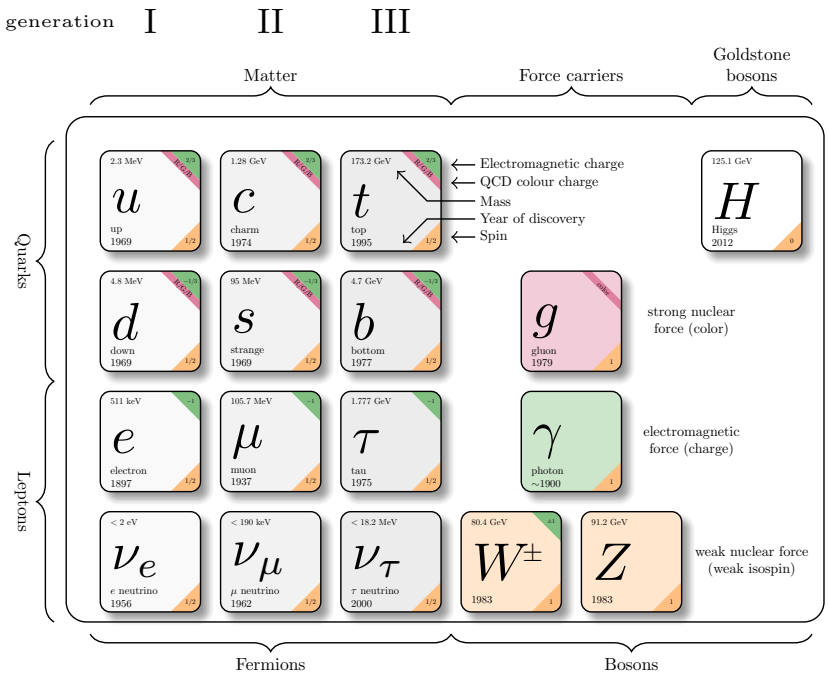


Figure 2.1: Fundamental particles in the Standard Model. Graphic adapted from Ref. [2].

2.1.1 Fundamental particles

Quarks and leptons are elementary spin-$1/2$ particles that describe matter. Quarks come in six flavours, each with a different mass. Three "up-type" quarks carry an electromagnetic charge of $+2/3$, while "down-type" quarks have charge $-1/3$. Except mass and charge they carry the same quantum numbers and can (in principle) undergo the same types of interaction with other particles in the SM. There are also six types of leptons. Three of these are negatively charged (-1), each of which has an electrically neutral counterpart – the neutrino's. Both quarks and leptons have corresponding anti-particles which will have the same mass but opposite quantum numbers. The fermions are categorised into three "generations", where the only difference between the generations is the particle mass. This ordering is illustrated in Figure 2.1.

The first generation of fermions is also the first part of the SM that humans uncovered. The electron is a charged lepton and the first elementary particle to be observed experimentally, by J. J. Thomson in 1897 [1]. The up- and down-quarks bind to form protons and neutrons. The protons and neutrons, together with electrons, form atoms. These make up all everyday objects, but also materials in Earth, the Sun and larger astronomical objects, as well as the molecules that form DNA in animal life and plants. The (anti-)electron neutrinos are emitted in beta decays, the process in which a neutron is converted to a proton in an atomic nucleus, along with a charged electron and the corresponding neutrino: $n \to p + e^- + \bar{\nu}_e$. The heavier fermions from the second and third generations were all discovered over the last century.

2.1.2 Forces

In the SM, elementary particles interact via the exchange of force-carriers from three fundamental forces. The *electromagnetic force* couples to electrically charged particles via the exchanges of massless photons. The photon itself carries no electric charge. The *weak nuclear force*, or electroweak (EW) force, is responsible for beta decays and carried by the electrically charged $W^{\pm}$ and and electrically neutral Z bosons. These bosons have non-zero mass generated by the Higgs field and couple only to fermions with a negative chirality ('left handed') or anti-fermions with positive chirality ('right handed'). These interactions therefore violate parity symmetry. The *strong nuclear force* binds nuclei together via the exchange of massless gluons. Gluons only couple to to particles that carry a colour charge, i.e. quarks and the gluons itself. The strong force is associated to *asymptotic freedom* and confines quarks and gluons to form bound multiquark states known as hadrons. The SM is written as a Lagrangian field theory, where the force carriers arise naturally by imposing a set of local gauge symmetries of the system. In the following we discuss this formalism for the three fundamental forces of the SM.

Quantum electrodynamics

The kinematics of free a fermion ψ with mass m is described with the Dirac Lagrangian

$$\mathcal{L}_{\text{Dirac}} = i\bar{\psi}(x)\gamma^{\mu}\partial_{\mu}\psi(x) - m\bar{\psi}(x)\psi(x). \tag{2.1}$$

This is not invariant under a local $U(1)$ transformation,

$$\psi(x) \xrightarrow{U(1)} \psi'(x) \equiv e^{i\alpha(x)}\psi(x)\,, \tag{2.2}$$

due to the derivative operating on the exponential term in Equation 2.1 that depends on x through the local phase $\alpha(x)$. To retain local $U(1)$ symmetry, one needs to consider the covariant derivative

$$D_{\mu}\psi(x) \equiv \big[\partial_{\mu} - ieA_{\mu}(x)\big]\psi(x)\,. \tag{2.3}$$

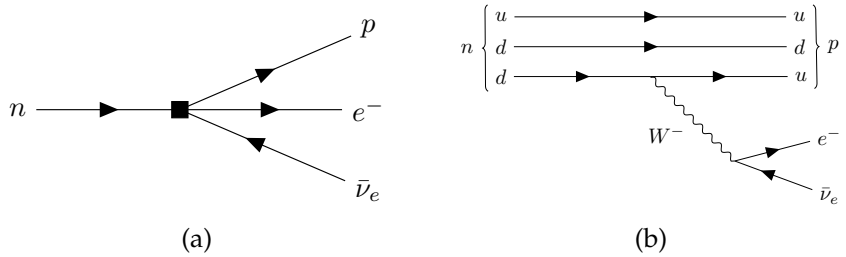
In this expression, $A_{\mu}(x)$ is a new spin-1 field that transforms under $U(1)$ as

$$A_{\mu}(x) \xrightarrow{U(1)} A'_{\mu}(x) \equiv A_{\mu}(x) + \frac{1}{e}\partial_{\mu}\alpha(x). \tag{2.4}$$

Replacing ∂_{μ} with D_{μ} in Equation 2.1 leads to $\mathcal{L}'_{\text{Dirac}} = \mathcal{L}_{\text{Dirac}}$, implying gauge invariance under local $U(1)$ transformations. The price to pay for this invariance is the introduction of a new degree of freedom, i.e. the massless field A_{μ} that interacts with electrically charged, fermionic fields: this is interpreted as the photon field that carries the electromagnetic force. The constant e is interpreted as the electromagnetic coupling. This relativistic quantum field theory of electrodynamics based on the symmetry group $U(1)$ is also called quantum electrodynamics (QED). In the next sections we see that QED is closely intertwined with the weak force in electroweak unification.

Weak force

Enrico Fermi proposed an explanation of the beta decay process of a nucleus via the reaction $n \to p + e^{-} + \bar{\nu}_{e}$, through a four-point interaction as shown in Figure 2.2a. Currently this process



(a) (b)

Figure 2.2: Feynman diagrams for nuclear beta decay (a) in Fermi's four-point theory and (b) as an electroweak interaction mediated by a massive W^- boson.

is understood at a more fundamental level through the exchange of a charged, massive W boson, as indicated in Figure 2.2b. Nevertheless, Fermi's model is very accurate for beta decays, because it occurs in the limit where the energy scale of the W mass is much larger than the energy scale of the decay process. Developments of a theory for weak interactions valid at higher energy scales followed the strategy of the successful model of QED: the existence of extra bosons that carry the weak force are inferred by imposing additional local gauge symmetries of the Standard Model Lagrangian. Before the development of a fundamental theory, weak processes were experimentally understood quite well. For example, the famous Wu experiment [3] indicated that only particles with a left-handed chirality interact via the weak force. This was taken into account when the theory was developed.

Electroweak interactions follow from imposing a local $SU(2)$ symmetry,

$$\psi(x) \xrightarrow{SU(2)} \psi'(x) = \exp\left[\mathrm{ig_w}\boldsymbol{\alpha}(\mathrm{x}) \cdot \mathbf{T}\right]\psi(\mathrm{x}), \tag{2.5}$$

where g_w indicates the charged weak coupling and $\mathbf{T}$ are the three generators of the group. The requirement of invariance of the Lagrangian under $SU(2)$ introduces three gauge fields W^1_μ, W^2_μ and W^3_μ. The generators of $SU(2)$ are 2×2 matrices and fermions that interact weakly are placed in doublets of weak isospin. These carry a weak isospin $I_W = \frac{1}{2}$ and isospin component $I^3_W = \pm\frac{1}{2}$. Physical W vector bosons are given by the linear combinations

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu). \tag{2.6}$$

We know from experiment that the weak force only couples to left-handed particles. For this reason, the local symmetry is often referred to as $SU(2)_\mathrm{L}$ where the L refers to left-handedness. To ensure that right-handed particles are not affected by local $SU(2)$ transformation in Equation 2.5, they are placed in an isospin singlet with $I_W = 0$. As example, the isospin doublet in the first generation of leptons and quarks are

$$\psi_L = \begin{pmatrix} \nu_L \\ e^-_L \end{pmatrix} \quad \text{and} \quad \psi_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}. \tag{2.7}$$

Because the gauge symmetry is only imposed on left-handed states, the weak interaction is said to *violate parity maximally*.

Electroweak unification

In addition to charged $W^{\pm}$ bosons, this predicts the existence of a third force carrier of the weak interaction. The W^3_μ boson can not be interpreted as the neutrally charged Z boson, because from experimental observations we know that the Z boson couples to both left-handed and right-handed fermions. This would contradict our statement that $SU(2)$ symmetry is imposed on left-handed states only. The Glashow-Weinberg-Salam (GWS) [4–6] mechanism introduces a new $U(1)_{\mathrm{Y}}$ local gauge symmetry in addition to $SU(2)_{\mathrm{L}}$ with associated vector gauge field B_μ [5]. This field couples to the hypercharge Y. In electroweak (EW) unification, the physical Z boson and the photon field A_μ are formed through a rotation in the plane of W^3_μ and B_μ,

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix}, \tag{2.8}$$

where θ_W is the Weinberg angle. The Weinberg angle is a combination of the couplings constants g_W and g':

$$\sin\theta_W = \frac{g'}{\sqrt{g_W^2 + g'^2}} \quad \text{and} \quad \cos\theta_W = \frac{g_W}{\sqrt{g_W^2 + g'^2}}. \tag{2.9}$$

Similar to the theory of QED, the partial derivative for EW theory

$$D_\mu = \partial_\mu + i g_W \boldsymbol{W_\mu} \cdot \mathbf{T} + \frac{i}{2} g' Y B_\mu \tag{2.10}$$

ensures that the $SU(2)_{\mathrm{L}} \times U(1)_{\mathrm{Y}}$ gauge symmetry is preserved.

Hypercharge is a quantum number that is shared between two particles in weak isospin doublets, for example $Y_{e_L} = Y_{\nu_L}$. It is related to weak isospin and electric charge via

$$I_3 + \frac{Y}{2} = Q. \tag{2.11}$$

Quantum chromodynamics

Quantum chromodynamics (QCD) is the quantum description of the strong force with massless gluons as the force carriers. QCD poses that quark fields are symmetric under local $SU(3)$ transformations, i.e.

$$\psi(x) \xrightarrow{SU(3)} \psi'(x) = \exp\Big[i g_s \boldsymbol{\alpha}(x) \cdot \hat{\boldsymbol{T}}\Big] \psi(x). \tag{2.12}$$

QCD is an example of a non-Abelian gauge theory, because the generators of the group $\hat{\boldsymbol{T}} = \{T^a\}$ do not commute. There are in total eight generators of the $SU(3)$ symmetry group, that are proportional to the hermitian and traceless Gell-Mann matrices λ^a:

$$T^a = \frac{1}{2}\lambda^a, \tag{2.13}$$

The quantum numbers associated to the $SU(3)$ symmetry are three conserved *colour* charges. The three colour-states red (r), green (g) and blue (b) can be represented by three-component vectors:

$$r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \tag{2.14}$$

The gauge coupling constant of the strong force is normally indicated with g_s.

Gluons are the force-mediators of the strong interaction that couple to the colour charge of a quark-antiquark pair. Eight gluons correspond to the eight linearly independent colour combinations. In contrast to the neutrally charged photon in QED, gluons actually carry a colour charge themselves due to the non-Abelian structure of QCD. The fact that gluons carry colour charge allows them to have a triple or quartic self-interaction.

The strong coupling constant g_s evolves with the energy scale of the momentum transfer Q^2 approximately as [7]

$$g_s(Q^2) \propto \frac{1}{\ln\left(Q^2/\Lambda^2_{\mathrm{QCD}}\right)}. \tag{2.15}$$

At high energy scales the strong coupling is small enough such that perturbative calculations of QCD can be made. However, the coupling increases for lower energy scales to the point where perturbativity breaks down at the perturbative QCD scale Λ_{QCD}. The property that the coupling becomes asymptotically weaker at higher energy scales is called *asymptotic freedom*. This leads to the 'elastic' behaviour of QCD that confines quarks in bound hadron-states and the non-observation of isolated quarks in nature. As we discuss in Section 4.1.5, there are phenomenological models to describe the hadronisation of quarks into bound hadron states. Quarks and gluons are collectively referred to as *partons*.

2.1.3 Spontaneous symmetry breaking

There is one major problem with the EW theory proposed in Section 2.1.2, namely experimental observations show that the W and Z electroweak gauge bosons have a non-zero mass. However, adding a mass term to the EW Lagrangian spoils the $SU(2)_{\mathrm{L}} \times U(1)_{\mathrm{Y}}$ symmetry and massive gauge bosons cannot be accommodated in this version of the theory. The Higgs mechanism resolve this issue by introducing a scalar field that interacts with the gauge bosons, through which non-zero mass terms for W and Z bosons arise naturally as the Higgs field acquires a minimum value.

An isospin doublet with two complex scalar fields,

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{2.16}$$

is added to the Lagrangian in the following combination:

$$\mathcal{L}_{\mathrm{Higgs}} = (D_\mu \phi)^\dagger (D_\mu \phi) - V(\phi). \tag{2.17}$$

The kinetic term contains the covariant derivative as defined in Equation 2.10. The Higgs poten-

tial has the form

$$V(\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2, \text{ with } \mu^2 < 0, \tag{2.18}$$

which due to its shape is also referred to as the "Mexican hat potential". There is an infinite set of minima satisfying

$$\phi^\dagger\phi = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2}. \tag{2.19}$$

We choose $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = v$. This choice retains invariance under $U(1)_{\mathrm{EM}}$, such that the massless photon can be associated to the gauge boson from this local symmetry. Expending the field around the vacuum expectation value v, we choose the unitary gauge

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.20}$$

Writing out the terms with the covariant derivative in Equation 2.17 explicitly gives

$$(D_\mu\phi)^\dagger(D_\mu\phi) = \frac{1}{8}v^2\Big[g_W^2(W^+)^2 + g_W(W^-)^2 + (g_W^2 + g'^2)Z_\mu^2 + 0\cdot A_\mu^2\Big], \tag{2.21}$$

where interaction terms with the Higgs field $h(x)$ have been omitted. Representing the physical gauge bosons in the same decomposition as Equations 2.6 and 2.8, we now identify mass terms for the W and Z gauge bosons, while the photon remains massless:

$$m_{W^+} = m_{W^-} = \frac{1}{2}vg_w, \tag{2.22}$$

$$m_Z = \frac{1}{2}v\sqrt{g_w^2 + g'^2}, \tag{2.23}$$

$$m_\gamma = 0. \tag{2.24}$$

2.1.4 Yukawa interactions

In the Standard Model, also the quarks and leptons also obtain their mass via interactions with the Higgs field. This is motivated by the same reason as the Higgs mechanism for massive gauge bosons: mass terms of the form $m\bar{\psi}\psi$ spoil the electroweak gauge symmetry. Let us adapt the notation that groups the fermions in Figure 2.1 into five different representations according to their quantum numbers:

Left handed quarks	$Q_{Li}(3,2)_{+1/3}$
Right handed up-type quarks	$U_{Ri}(3,1)_{+4/3}$
Right handed down-type quarks	$D_{Ri}(3,1)_{+1/3}$
Left handed leptons	$L_{Li}(1,2)_{-1}$
Right handed leptons	$l_{Ri}(1,1)_{-2}$.

In this notation, $Q_{Li}^I(3,2)_{+1/3}$ indicates an $SU(3)_{\mathrm{C}}$ triplet, $SU(2)_{\mathrm{L}}$ doublet with hypercharge $Y = +1/3$. The subscript $i = 1,2,3$ is the generation index. The Yukawa Lagrangian is an in-

teraction term between the Higgs field and fermions that gives rise to fermion masses [8]:

$$-\mathcal{L}_{\text{Yukawa}} = Y_{ij}\bar{\psi}_{Li}\phi\psi_{Rj} + \text{h.c.} \tag{2.25}$$

$$+ Y^d_{ij}\bar{Q}_{Li}\phi D_{Rj} + Y^u_{ij}\bar{Q}_{Li}\tilde{\phi}U_{Rj} + \text{h.c.}. \tag{2.26}$$

An important observation is that the coupling of fermions to the Higgs field is proportional to the fermion mass.

In above expressions, Y_{ij} are arbitrary, complex matrices. Since these matrices are generally not diagonal, this couples different generations i and j. The physical fermion states evolve as eigenstates of the free-particle Hamiltonian and are labelled as the *mass eigenstates*. This is distinct from the *weak eigenstates* that correspond to the flavour states of fermions that undergo weak interactions. To describe a physical system we wish to describe interactions in the mass basis, and translate weak eigenstates to the mass eigenstates by a (unitary) matrix rotation.

For quarks, the change of basis is made by operating with the 3×3 Cabibbo-Kobayashi-Maskawa (CKM) matrix on down-type quarks:

$$\begin{pmatrix} d^I \\ s^I \\ d^I \end{pmatrix} = \hat{V}_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{\text{ud}} & V_{\text{us}} & V_{\text{ub}} \\ V_{\text{cd}} & V_{\text{cs}} & V_{\text{cb}} \\ V_{\text{td}} & V_{\text{ts}} & V_{\text{tb}} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{2.27}$$

The CKM matrix is not fully diagonal, and as a result charged-current interactions involving the weak force can mix quarks of different generations. The couplings of weak, charged interactions are proportional to the CKM elements. The Wolfenstein parametrisation decomposes the CKM matrix into four parameters that can be determined from fits to experimental data. Decomposition of the CKM matrix in the expansion parameter $\lambda \approx 0.22$ shows that interactions between quark-pairs of the three generations are suppressed up to different orders in λ:

$$\hat{V}_{\text{CKM}} \approx \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(\lambda) & \mathcal{O}(\lambda^3) \\ \mathcal{O}(\lambda) & \mathcal{O}(1) & \mathcal{O}(\lambda^2) \\ \mathcal{O}(\lambda^3) & \mathcal{O}(\lambda^2) & \mathcal{O}(1) \end{pmatrix}. \tag{2.28}$$

An equivalent change of basis for leptons is made with the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix that mixes mass and flavour states of neutrinos.

2.1.5 Successes of the Standard Model

Let us take a step back and appreciate that the Standard Model is outstandingly successful by looking at some examples. First and foremost: *all* particles in the SM have been observed experimentally, with the discovery of the Higgs boson in 2012 as one of the biggest scientific achievements of the twenty-first century. It is important to state that many particles, including the W boson, Z boson, charm quark and top quark, have been predicted before being observed experimentally. The LHC has – so far – been very favourable to the SM. It succeeded in finding the Higgs boson, but no additional particles or signatures from beyond the SM have been discovered. Excitement of discrepancies between theory and experiment, such as the "$750\,\text{GeV}$ diphoton excess" by ATLAS [9, 10] and CMS [11, 12] or the R_K measurement by LHCb [13, 14], took the

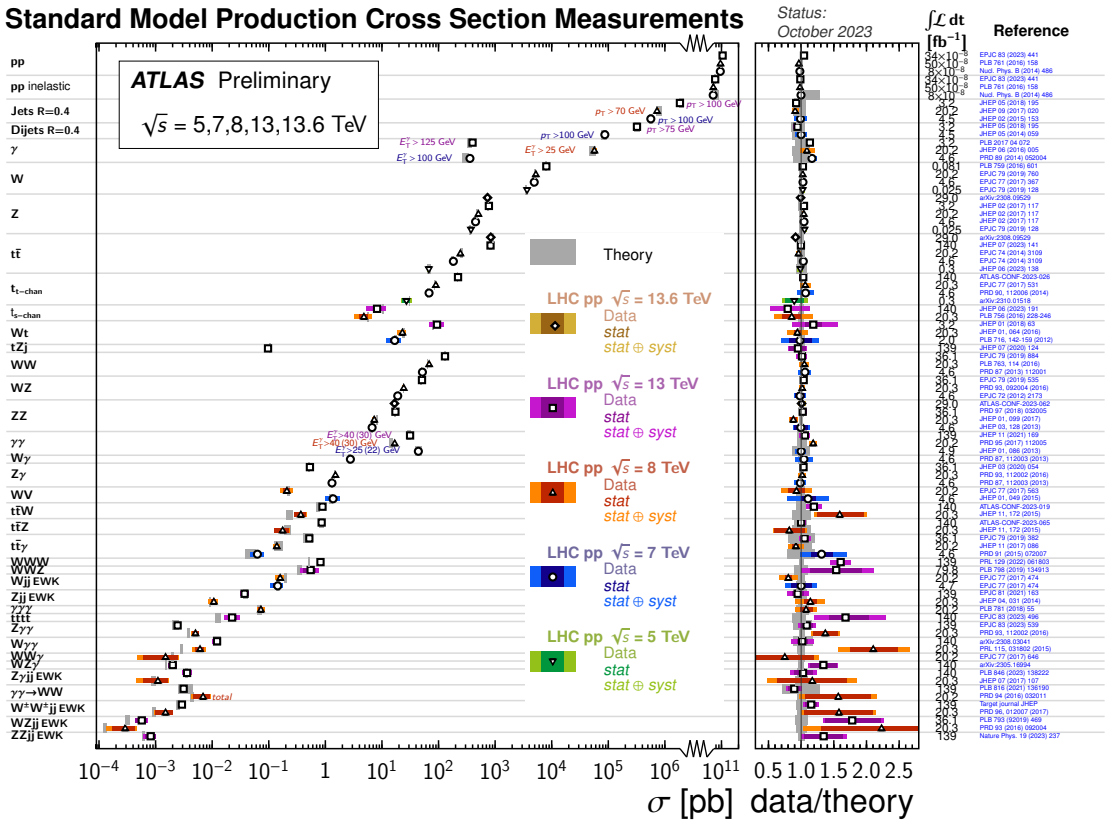


Figure 2.3: Comparison between theoretical predictions (bars) and experimental ATLAS measurements (symbols) of various proton-proton scattering processes. Figure from Ref. [16].

physics community by storm but went away as the SM prevailed in the end. This is exemplified in Figure 2.3, where LHC measurements of SM processes show agreement between theory and experiment across an astonishing 15 orders of magnitude. Another example is the measurement of the electron magnetic moment, where experimental data agree with state-of-the-art SM predictions up to 1 part in 10^{12} [15].

2.2 Beyond the Standard Model

We have discussed the fundamentals of the Standard Model and seen that it can described many processes observed in nature with great accuracy. However, it faces several problems that motivate searches for BSM physics. In this section, we first summarise the biggest open questions in particle physics. We focus our attention on quests for dark matter and describe a theoretical model that predicts a dark matter candidate particle from anomaly cancellation. Finally, we provide an overview of how dark matter (and other dark signatures) could be measured at particle colliders.

2.2.1 Shortcomings of the Standard Model

There are multiple open questions from a theoretical point of view. The first obvious problem with the SM is that it does not describe gravity, and attempts to unite gravitational forces with

the SM are so far unsuccessful. Secondly, naturalness problems such as the hierarchy problem (why is gravity is so much weaker than the weak force?) and the strong CP problem (why does the strong interaction not seem to violate CP invariance experimentally, while it theoretically could?) are not explained by the SM. There are also experimental hints that make us believe that the SM is incomplete. For example, according to the SM matter and antimatter should have been produced in equal amounts in the Big Bang. However, cosmological observations show that there is *much* more matter than antimatter. This large imbalance is referred to as the *matter-antimatter asymmetry*. Lastly, the Standard Model has no explanation for *dark energy* and *dark matter*. The latter has a central role in this thesis; both in the ATLAS measurement of $p_{\mathrm{T}}^{\mathrm{miss}}$+jets in Section 5 and in phenomenological studies of a dark matter model in Section 6.

2.2.2 Dark matter

The search for dark matter is one of the biggest challenges and open-ended questions in physics and astronomy. In this section, we review the motivations for dark matter, what it could be and how it could be discovered.

Evidence from astrophysical observations

There is compelling evidence for the existence of dark matter, and first notions of it date back more than one hundred years.[1] Pioneering work in the field of dark matter include Fritz Zwicky's observations in 1933 of galaxies orbiting the Coma Cluster [18]. He observed that the gravitational effect of visible clusters was too small to explain the high orbital velocities, and concluded from this that dark matter must be present in greater amounts than visible matter. Similar discrepancies between observed and expected orbital velocities is seen at galactic rotation curves. The rotation velocity $v_r(r)$ of matter within galaxies is expected to decrease with the distance to the centre of the galaxy r. Experimental measurements of rotation curves in Figure 2.4 however, show that rotational velocities at large distance to the centre of the galaxy remain constant. This suggests the presence of additional matter in the galaxy.

Further evidence comes from cosmological measurements of the large-scale structure in the universe. The ΛCDM (Lambda cold dark matter) model, also "the standard model of cosmology", accurately describes the evolution of our universe. Cosmological parameters in this model can be fitted by using measurements of the cosmic microwave background (CMB). CMB radiation is an emission of black body radiation across the universe and is a remnant from early stages of the expanding universe. While the CMB is very uniform, there are small fluctuations. These CMB power spectra are consistent with the ΛCDM model when approximately 5% of the energy-matter density of the universe is in the form of visible, baryonic dark matter. The majority is attributed to dark energy (72%), while the final 23% has the form of dark matter [20, 21].

Non-particle dark matter

Non-particle dark matter hypotheses suggest that whatever fraction of the universe we classify as dark matter does *not* come in the form of (yet unidentified) fundamental particles, and that the astronomical observations can be explained in a different way than hypothesising new particles.

[1] A comprehensive guide with historical context on how dark matter came to matter in the twentieth century is given in Ref [17].

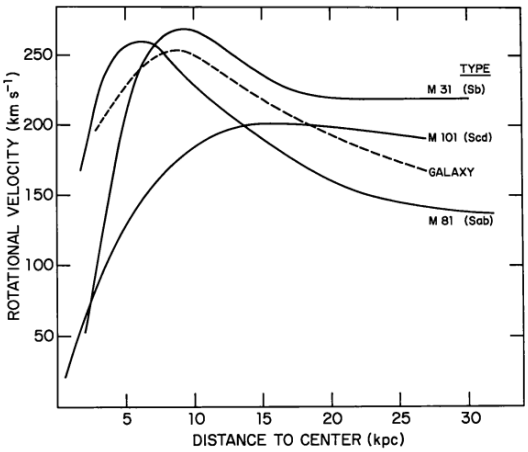


Figure 2.4: Galactic rotation curves indicating the rotational velocity in kilometre per second as a function of distance to the galaxy centre in kiloparsec ($1\,\mathrm{kpc} \approx 3 \times 10^{19}\,\mathrm{m}$). At large radii, galaxies M31, M81 and M101 exhibit flat rotation curves. Figure taken from Ref. [19].

One such proposal is that of Massive Astronomical Compact Halo Objects (MACHOs), stellar bodies that are hard to detect as they do not directly emit light. MACHO candidates include for example brown dwarfs, black holes and neutron stars. Collaborations such as the EROS-2 project [22] look for gravitational lensing effects of these objects to find evidence for (or constrain the parameter space of) MACHOs. Another dark matter proposal is that of primordial black holes, hypothetical stellar bodies that would have been created in the early Universe [23, 24]. The theory of Modified Newtonian Dynamics (MOND) explains the observed motion of large-scale structures by modifying Newton's laws in the limit of small acceleration [25]. While this can successfully describe the observed galaxy rotation curves, it is not a relativistic theory and it does not accommodate gravitational lensing and gravitational waves.

Particle dark matter

Many theories explain the astrophysical observations via the existence of new, fundamental particles beyond the SM. From the experimental observations we can infer some characteristics of particle DM candidates. First and foremost: it does not interact electromagnetically. It does not absorb, reflect or emit light and therefore it can not be detected by conventional telescopes, hence the name "dark". Because it does not interact via the electromagnetic force, DM is electrically neutral. There are a myriad of particle DM candidates proposed in the literature that fit this description. In this thesis, we restrict the discussion to the general class of Weakly Interacting Massive Particles (WIMPs), take a whistle-stop tour through a few other candidates in the "zoo" of dark matter, and finally zoom in further on a small set of DM theories that are most relevant to this work.

Although there is no official definition of WIMPs, they are normally taken to be particles in the range from approximately $1\,\mathrm{GeV}$ to $10\,\mathrm{TeV}$ range that interact feebly. In this context, "feebly" means that they interact weaker than, or as weak as the weak nuclear force with SM particles. In the search for dark matter, WIMPs have historically taken centre stage mainly for two reasons. The first is the "WIMP miracle", alluding to the coincidence that weak-scale WIMPs with couplings to the SM of $\mathcal{O}(1)$ freeze out with the DM relic density observed in the Universe. Depend-

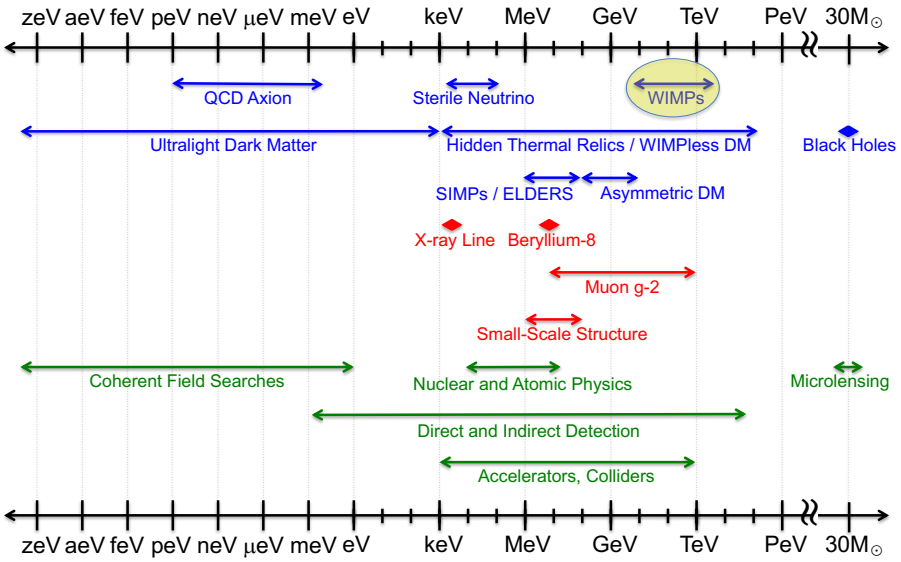


Figure 2.5: Overview of the particle dark matter landscape in terms of mass. Blue arrows indicate mass rangers of proposed candidates, red arrows indicate observed anomalies and green arrows demarcate the reach of different detection methods. Figure taken from Ref. [27].

ing on assumptions made in the calculation of interaction between WIMPs and visible matter, this restricts WIMP masses to the GeV to TeV mass range. This is consistent with the Lee-Weinberg lower bound, restricting hypothetical particles with feeble interactions to a lower mass of at least 2 GeV to be consistent with the observed relic density in the universe [26]. From a model-building perspective it is very attractive that WIMPs provide answers at both the smallest length scales (in particle physics) and the largest length scales (in cosmology). Second, there is a high complementarity between different experimental detection mechanisms of WIMPs, providing a falsifiable theory with enhanced discovery potential. WIMPs are probed via direct detection, indirect detection and collider searches [27], which makes them appealing from an experimental viewpoint.

As illustrated in Figure 2.5, WIMPs are only one proposed candidate in a DM landscape that offers many orders of magnitude in mass. Models of DM include (but are not limited to) supersymmetric (SUSY) particles[2] [28, 29], axions and axion-like particles [30, 31] or sterile neutrinos [32]. In the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets analysis we investigate the two-Higgs-doublet model with a pseudo-scalar mediator (2HDM+a), a model that extends the SM Higgs sector with an additional Higgs doublet [33]. These are examples of UV complete theories (i.e. valid up to high energy scales) that are theoretically motivated and self-consistent. Full models (can) have many free parameters and a complicated phenomenology that is difficult to test and constrain with a finite dataset.

The simplified model approach [34], which only captures the main features of SM-DM interactions, is complementary to full model searches. It is a more model-agnostic approach that is useful to constrain certain regions of DM parameter space in a global way. For experimentalists it is also a useful model to benchmark the sensitivity of different analysis or phase spaces, as exemplified in the reinterpretation part of Chapter 5. Simplified models include s-channel mediators or t-channel mediators that couple SM quarks to Majorana or Dirac DM. DM effective

[2] These are typically WIMPs.

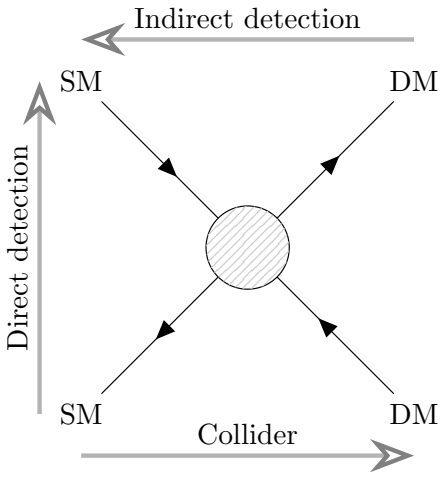


Figure 2.6: Different detection mechanisms of dark matter. If dark matter is produced at colliders, this is via the scattering of SM particles that create DM in the final state. Direct detection experiments attempt to measure the direct scattering of DM particles coming from the cosmos to SM particles on Earth. Indirect searches look for smoking gun evidence of dark matter, particularly the annihilation of DM into SM particles.

field theory (DM EFT) is an even more model-independent approach that models interactions between DM and SM field with four-point operators [35]. Of course, the EFT approach is only valid up to the relevant energy scale of the DM mediator.

In Section 6.3 we focus on a model that predicts a DM candidate from gauge anomaly cancellation, and revise constraints on this model from different experimental measurements.

Detection of particle dark matter

In our discussion of DM so far, we discussed *why* it exists, *what* it could be, but not *how* to discover it. As summarised in Figure 2.6, dark matter detection is broadly classified in three categories: direct detection, indirect detection and collider searches.

Direct search experiments are designed to detect the rare events where SM particles are struck by DM particles travelling through Earth. The main principle is to have a container filled with a homogeneous material which nuclei serve as "SM targets". Interactions with dark matter can lead to ionisation electrons scintillation photons, phonons, Cherenkov radiation and more signatures that can be used to identify such a scattering event. These experiments are located underground to suppress background from cosmic rays. Direct detection groups like the XENON collaboration [36–38] and LZ [39–41] use liquid xenon as the active target. Not only does it have a relatively large atomic number (54), beneficial as the DM-SM cross section is small, it is also attractive because it is stable, suffers from low background contamination and it can be purified very well. Over the past decade these detectors have grown in size because the interaction probability in the detector scales with the interaction volume of the SM medium.

Direct detection experiments are mostly aimed at detecting WIMPs, reporting exclusions of WIMP-models parametrised in the WIMP cross-section and dark matter mass. So far, no evidence of dark matter has been found. The non-observation is converted into exclusions in the plane of WIMP cross-section and dark matter mass. Current bounds are summarised in Figure 2.7.

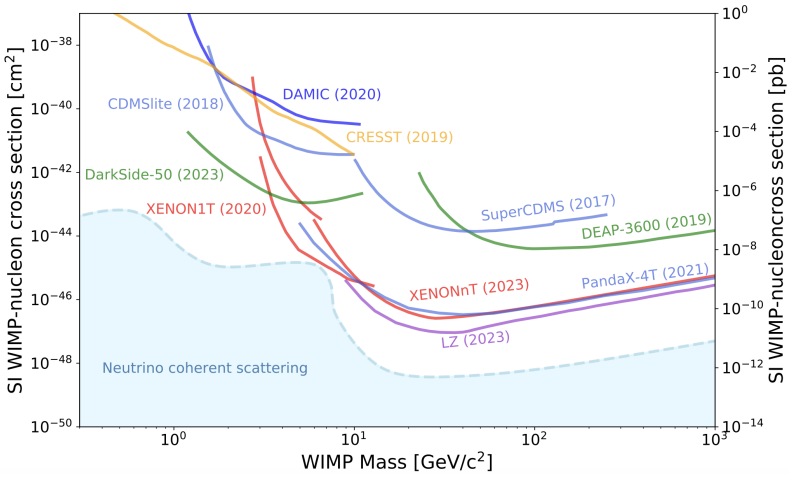


Figure 2.7: Upper limits on the nucleon-DM cross section as function of the DM mass from different direct search programs. The blue shaded area indicates the "neutrino floor": a region where nuclear recoils from neutrinos become an overshadowing background in direct searches. Figure taken from [42].

Indirect searches look for the existence of dark matter via the annihilation of DM-DM into SM particles, which could for example be photons, (anti)neutrinos, (anti)lepton and (anti)protons. This gives a wide scope of measurement tools: charged-particle detectors, neutrino detectors and γ-ray detectors are all suitable for indirect DM searches. The annihilation products would be mainly originate in places with high DM density, e.g. in galactic centres and galaxy clusters. Both space and ground based experiments, including Fermi-LAT [43], H.E.S.S. [44] and IceCube [45], make measurements at these astrophysical scales to hunt for remnants of DM.

Where direct and indirect searches attempt to detect DM signatures coming from outer space, colliders provide a controlled environment where DM can be produced and the initial state of the SM-to-DM process is known very well. The guiding principle of collider searches is that if DM is produced in a proton-proton collisions, it can leave the experiment undetected. The main signatures of DM production in colliders is therefore some sort of "missing energy".

As we discuss more carefully in Chapter 4, the big unknown in the initial state of proton-proton collisions is which fraction of momentum is carried by each individual parton that makes up the proton. This makes it difficult to infer any information of how much energy would have gone into the creation of DM particle in the longitudinal direction. However, incoming SM particles in collider beams carry no transverse momentum component and conservation of momentum dictates that the sum of transverse momenta of the final state particles add up to zero. Therefore, the total sum of transverse momenta is a robust alternative to the total missing transverse momentum.

The fact (or assumption) that DM leaves no detector signature is both a blessing and a curse. Consider the simplest case of the production of a DM candidate χ via $pp \to \chi\chi$. In this scenario, the χ pair will be produced back-to-back and lead to no momentum imbalance in the transverse plane. What is more: the detector will notice absolutely nothing. There is no distinction between the production of pure $pp \to \chi\chi$ or "nothing" happening. It is only possible to detect DM in collider searches if it is produced in association with at least some SM particle in the final state

that are measured by the detector. If DM is produced in association with a highly energetic jet, for example via $pp \to \chi\chi j$, the detector can trigger on the jet, add up the momenta of all particles and notice there is a missing momentum component in the transverse plan, $p_{\mathrm{T}}^{\mathrm{miss}}$. So while the main trigger for DM production at colliders is a large $p_{\mathrm{T}}^{\mathrm{miss}}$, there are SM background contributions from neutrinos that are produced during the collision and like DM leave the detector undetected. The search for DM production in final states with large $p_{\mathrm{T}}^{\mathrm{miss}}$ is the focus of the ATLAS analysis in Chapter 5 and Section 6.3.

While $p_{\mathrm{T}}^{\mathrm{miss}}$ signatures are a well-motivated and popular approach to collider DM searches, there is a multitude of ways at which LHC physicists extend reach of the DM search program. To give an idea we briefly review two of them. First, the community has shown a grown interest in the interplay between DM and long-lived particles (LLPs) [46]. LLP signatures, which could come from a (heavy) DM particle that decays in the detector volume, leads to exotic signatures that are very different from SM processes: there will be displaced decay vertices leading to displaced leptons, jets or combinations of both; disappearing tracks; trackless jets in electromagnetic calorimeters; and more. Secondly, models strongly coupled (QCD-like) dark sectors are currently under study at the LHC. A distinctive signature of these models are semi-visible jets, that arise when parton evolutions include dark sector emissions [47, 48].

Chapter 3

Collider experiments at CERN

CERN ("Conseil Européen pour la Recherche Nucléaire", or European Council for Nuclear Research) is an international organization based in Geneva, Switzerland with research facilities that help to deepen our understanding of particle physics. It was founded in 1954 in the wake of the Second World War, both to re-establish Europe as a powerhouse in fundamental physics research and to strengthen ties between countries on the European continent. Seventy years later countries from across the globe work together on the particle physics experiments housed at CERN.

The (current) flagship project of CERN is the Large Hadron Collider; the largest particle accelerator ever built that accelerates protons to relativistic velocities. Most famously, the LHC enabled the discovery of the Higgs boson in 2012. But the realm of CERN extends beyond the LHC; it is also home to multiple fixed-target experiments, antimatter experiments and non-collider experiments. CERN's history includes many other notable scientific achievements, such as the discovery of W and Z bosons, or the determination of the number of neutrino families at the Large Electron-Positron Collider (LEP).

Alike to human spaceflight programs, collaborations between motivated physicists and engineers – who dream of specific physics goals – give many byproducts. For example, the World Wide Web was born at CERN (in 1989) in the development of an automated information-sharing system. Technologies developed at CERN, both in software and hardware, often lead to innovations in different sectors, including medical imaging, grid computing and cryogenics.

Section 3.1 provides an introduction to the main operating principles of the LHC, while details on the ATLAS experiment are given in Section 3.2.

3.1 Large Hadron Collider

The LHC is the world's highest energy particle collider, situated approximately 100 metres underground in a $26.7\,\mathrm{km}$ tunnel that was constructed for the Large Electron-Positron collider between 1984 and 1989. It is designed to accelerate charged particles in two adjacent beam pipes – with particle travelling opposite direction – and to collide them at a centre-of-mass energy of up to $\sqrt{s} = 14\,\mathrm{TeV}$. The tunnel is not entirely circular but instead is made up out eight straight sections and eight arcs [49].

There are four main collision points, each of which is home to a dedicated particle detector to reconstruct (different properties of) charged-particle collisions. The LHC mostly accelerates protons, but there are also dedicated heavy-ion runs with lead-proton or lead-lead collisions. ATLAS [50] and CMS [51] are "general-purpose" detectors that cover a wide range of physics; from Higgs boson physics to BSM searches to precision SM measurements. LHCb [52] is a forward-backward asymmetric detector with excellent (secondary) vertex reconstruction designed for making precision measurements of heavy-flavour decays. ALICE [53] focuses on heavy-ion collisions to probe the quark-gluon plasma.

The final centre-of-mass energy is limited by two factors. First, charged particles in the beam pipe are kept in their trajectory using bending dipole magnets. Even though the superconducting magnets are incredibly powerful and a generate magnetic field of up to $\sim 8\,\mathrm{T}$, the finite strength of the magnets is a natural limit on the particle velocity in circular colliders. Depending how you look at it, the size of the accelerator can also be a limiting factor, as one could achieve a higher centre-of-mass energy with the same magnets but at a larger radius. The second fundamental limitation is synchrotron radiation: electromagnetic radiation emitted by charged particles travelling in curved trajectories.

The acceleration of protons through electromagnetic fields relies on the principle of radio frequency (RF) system. In the LHC, each beam segment consists of eight cavities in disk-shaped conductors. Each conductor is individually driven by a RF system that operates at $400\,\mathrm{MHz}$, which then provides an alternating electromagnetic field [49]. Every time a charged particle traverses a cavity and its electromagnetic field it accelerates a little bit. To ensure the RF cavities accelerate (and not accidentally slow down) the beam, the RF frequency must always be an integer multiple of the frequency at which bunches revolve the LHC. The RF mechanism also keeps charged particles tightly packed in *bunches*. The proton bunches are kept in orbit by 1232 dipole magnets and 474 quadrupole magnets are used to focus the beams into a small longitudinal area to increase the interaction probability at the collision points.

The CERN accelerator complex consists of multiple components and pre-accelerators to ultimately achieve a proton-proton centre-of-mass energy of $\sqrt{s} = 14\,\mathrm{TeV}$ in the LHC ring. The proton source is hydrogen gas, which is passed through an electromagnetic field to isolate the constituent protons. These are then inserted into Linear accelerator 2 (Linac2), which accelerates the proton beam up to an energy of $50\,\mathrm{MeV}$.[1] The next stage is the Proton Synchrotron Booster (PSB, or sometimes "Booster") a small circular accelerator achieving a beam energy of $1.4\,\mathrm{GeV}$[2] for injection into the Proton Synchrotron (PS) [56]. The PS has a circumference of $628\,\mathrm{m}$ and increases the particle energy up to $26\,\mathrm{GeV}$. The final stage before injection into the LHC ring is the Super Proton Synchrotron (SPS) [57], which delivers proton beams with an energy of $450\,\mathrm{GeV}$. The beam injection from SPS into the LHC takes several minutes and during this time the beam energy is kept at $450\,\mathrm{GeV}$. When the LHC beam reaches the desired beam intensity, the ramp-up phase accelerates protons to their final energy of up to $7\,\mathrm{TeV}$.[3]

[1] Linac2 was the first step in the acceleration of protons for the LHC between 1978 and 2020. If it ain't broke, don't fix it. Its successor Linac4, connected to the pre-accelerator complex at CERN during Long Shutdown 2 in 2020, accelerates a beam of negative hydrogen ions (H^-) up to an energy of $160\,\mathrm{MeV}$ [54]. The two additional electrons of the hydrogen ions are stripped after leaving Linac4. This will ultimately achieve an increase of beam intensity by a factor two [55].

[2] Also the Proton Synchrotron Booster underwent major upgrades in 2020, and its current beam energy is $2\,\mathrm{GeV}$.

[3] The LHC is designed to run at a maximum collision energy of $\sqrt{s} = 14\,\mathrm{TeV}$, but operated at $\sqrt{s} = 13\,\mathrm{TeV}$ during Run-2 and is currently running at $\sqrt{s} = 13.6\,\mathrm{TeV}$ for Run-3.

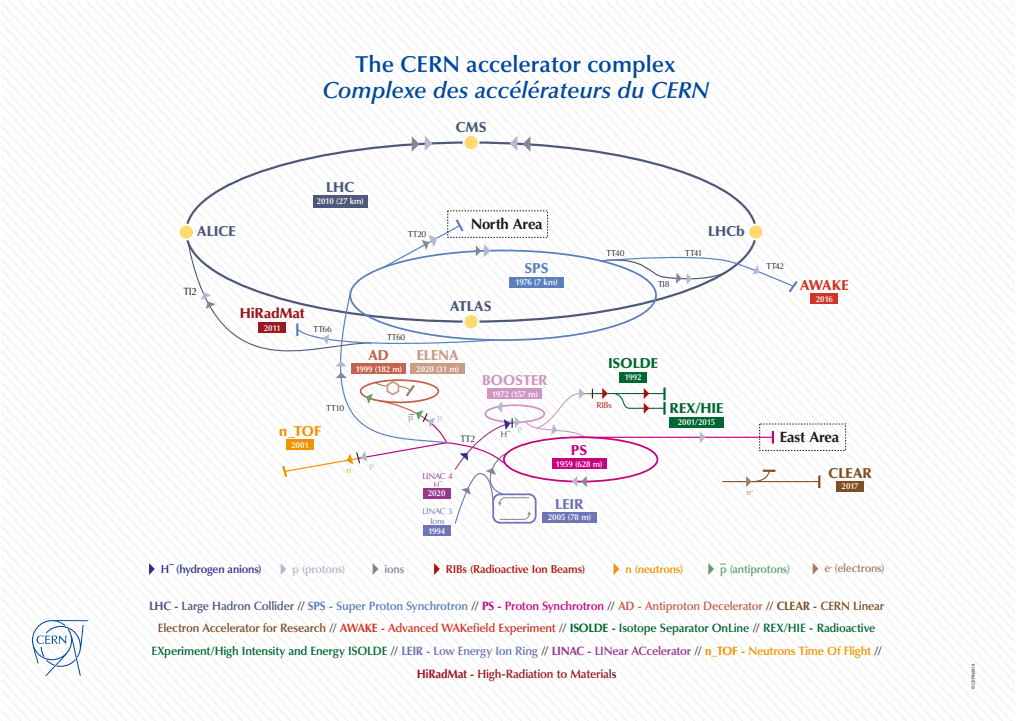


Figure 3.1: The CERN accelerator complex with the four main LHC detectors ATLAS, CMS, LHCb and ALICE, and the pre-accelerators including the SPS, PS, PSB and the Linac [58]. The Low Energy Ion Ring (LEIR) is used in heavy-ion runs to inject the heavy-ions into the PS.

The operation of the LHC is organised in "runs" with temporary "long shutdowns" in-between runs for upgrades and maintenance of both the LHC accelerator complex and the particle detectors. Within each run, there is a technical stop towards to end of the year (year-end technical stop; YETS) until early in the next year, normally for smaller technical work on the detectors and repairs on the LHC and its injectors. During Run-1, the first operational run of the LHC from 2010 to 2013, the LHC delivered beam energies between $3.5\,\text{TeV}$ and $4\,\text{TeV}$ [59]. The main datasets that are analyses in this thesis were collected during Run-2, which was the period of operation between 2015 and 2018. During this time, the total centre-of-mass energy of proton-proton collisions was $13\,\text{TeV}$. At the time of writing, the CERN experiments are in the third year of data-taking in Run-3, with the highest beam energy achieved at the LHC so far of $6.8\,\text{TeV}$. The next scheduled phase, after a Long Shutdown 3 from 2026 to 2028, is the High Luminosity LHC (HL-LHC) project [60, 61]. At the HL-LHC, there is no big jump in beam energy but instead an increase of the collision rate by a factor of five compared to the original design of the LHC and the current instantaneous luminosity during Run-3. Significant upgrades are planned to all major experiments to cope with the high-density environment to come from 2029, with a projected delivered luminosity of $3000\,\text{fb}^{-1}$ (compared to e.g. $156\,\text{fb}^{-1}$ LHC delivered luminosity during Run-2).

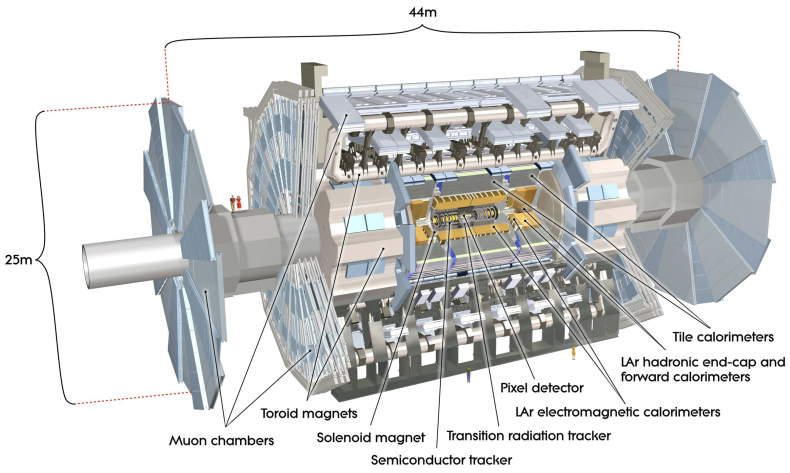


Figure 3.2: Schematic overview of the ATLAS detector [50].

3.2 ATLAS experiment

A Toroidal LHC Apparatus (ATLAS) [50] is a general-purpose detector built to measure proton-proton collisions at the LHC. Situated in a cavern approximately 100 meters underground, it has a forward-backward symmetrical and cylindrical shape around the interaction point (IP), with a diameter of $25\,\mathrm{m}$ and length of $44\,\mathrm{m}$. ATLAS has a layered structure consisting of multiple sub-detectors, each designed to reconstruct specific aspects of the final state. By combining the information from the different detector parts, full events are reconstructed and used for physics analyses.

Closest to the beam pipe is the inner detector (ID), providing accurate measurements of the momentum and charge of charged particles. Moving radially outwards from the IP, the electromagnetic (EM) and hadronic calorimeters measure energy deposits of EM and hadronic showers respectively. The outer layer is the muon spectrometer (MS). ATLAS owes its name to the superconducting toroid magnets, which provide magnetic field up to $4\,\mathrm{T}$ across the MS to measure the charge and momentum of traversing muons. To extend the fiducial region of the detector in a higher angular range, end-caps for the ID, calorimeters and MS are placed on either side of the detector. A schematic overview is shown in Figure 3.2. More technical details are provided in the following sections.

3.2.1 Coordinate system and geometrical parameters

To describe the design of ATLAS and the physics we do with it, we will first need a common set of coordinates and parameters to discuss spatial directions, dimensions and observables. ATLAS uses a three-dimensional coordinate system with the interaction point placed at the origin. The positive x-axis points to the centre of the LHC ring, while the z direction points along the tunnel (counter-clockwise looking at the LHC from above) and the y-axis is perpendicular to the x, z-axes pointing upwards. The azimuthal angle ϕ indicates the rotation around the beam axis and the polar angle θ is the angle from the beam axis. The radial distance to the beam pipe is denoted by R. Figure 3.3 illustrates the spatial coordinates in a three-dimensional simplification of ATLAS.

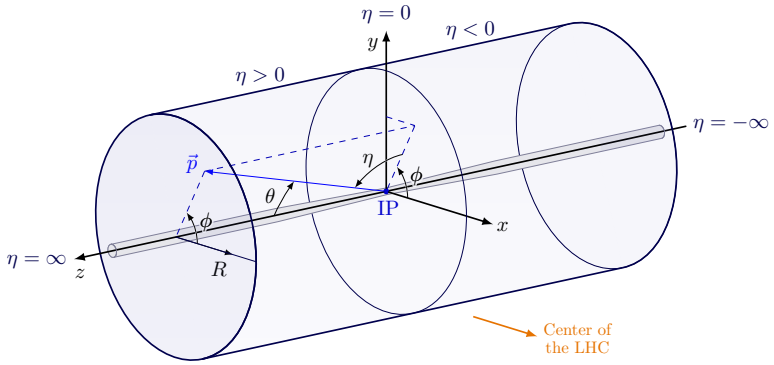


Figure 3.3: Three-dimensional view of the ATLAS coordinate system, indicating the x, y and z directions, polar angles θ and ϕ, pseudorapidity η and radial distance R. Figure adopted from Ref. [62].

A useful quantity is the **rapidity**

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{3.1}$$

where E is the energy of a particle and p_z its momentum in the z-direction. Why is this useful? First, the rapidity *difference* between two objects is a Lorentz invariant quantity under boosts in the $\vec{z}$-direction. Second, it tells us if an object is produced forwards (large y) or centrally (small y). When the momentum along the beam axis is zero, in other words when it is produced *centrally*, the rapidity will evaluate to zero. A positive rapidity corresponds to a particle moving in the positive $\vec{z}$ direction and vice versa.

More specifically in the massless case $E \gg m$, where m is the mass of a particle, this formula can be rewritten to define the **pseudorapidity**

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \tag{3.2}$$

Physicists like this re-parametrization of the polar angle, not only because the polar angle is one of the key quantities we are interested in, but also because it helps to visualise the relation between rapidity and "forwardness". It is also used as an approximation of the rapidity for massive (but relativistic) particles. Objects travelling with $\eta = 0$ move perpendicular to the beam pipe ($\theta = \pi/2$), while $|\eta| \to \infty$ corresponds to objects travelling along the beam pipe ($\theta = 0, \pi$).

The azimuthal angle ϕ and rapidity y allow for an unambiguous direction in which an object is travelling from the interaction point. The **angular difference**

$$\Delta R \equiv \sqrt{\Delta y^2 + \Delta \phi^2} \tag{3.3}$$

is a dimensionless quantity to measure differences in direction of travel.

3.2.2 Inner detector

The ATLAS ID currently consists of four layers that together provide charged track measurements and both primary and secondary vertex reconstruction in the pseudorapidity range $|\eta| < 2.5$. The goal is to minimally perturb the incoming particles while precisely determining their

charge and momenta. Moving radially outwards from the interaction point, the four layers are the Insertable B-layer (IBL), the Pixel detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). The layers are arranged in a barrel for central track measurements and two end caps on either side of the barrel to extend the track reconstruction to the more forward regions. Figure 3.4 illustrates the layout of the ID with the barrel, end caps and various layers.

Reconstruction of charged particle momenta is achieved by surrounding the ID with a central solenoid, which generates an axial $2\,\mathrm{T}$ magnetic field with its field lines aligned with the beam pipe. This magnetic field bends the trajectory of charged particles. Track reconstruction algorithms use the curvature and bending direction to precisely measure charge and momenta of tracks above a nominal p_{T} of $0.5\,\mathrm{GeV}$ [50].

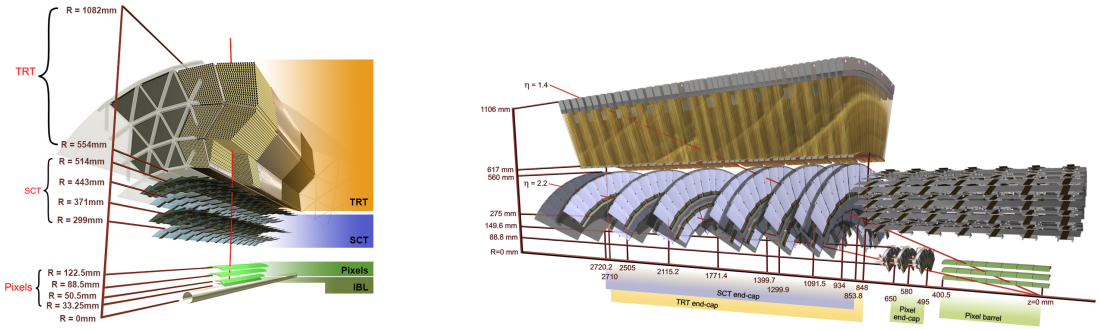
Because the ID is the component closest to the collision point only $3.3\,\mathrm{cm}$ from the LHC beam line, each layer is designed to resolve a large track density and endure large amounts of radiation at design luminosity. To this end, the innermost part covering the vertex region uses pixel silicon technology. High-granularity pixel modules are arranged in three barrel layers and two end caps with three disks each, which is referred to as the **Pixel detector**. The barrel, which covers the $|\eta| < 1.5$ region, has 1456 pixel sensors installed on a total of 112 barrel staves. The end-caps extend the coverage to $|\eta| < 2.5$ and are comprised of 48 end-cap sectors and 288 pixel sensors divided between the six disks. Each pixel sensor consists of roughly, 47232 pixels, such that there is a total of around 80 million pixels in the barrel and end-cap segments of the Pixel detector combined. The nominal pixel size is $50 \times 400\,\mu\mathrm{m}^2$ in the ϕ and z-direction respectively, providing an active area of $1.7\,\mathrm{m}^2$ ($1.45\,\mathrm{m}^2$ in the barrel and $0.28\,\mathrm{m}^2$ in both end-caps combined) [63]. The fruits of our labour are intrinsic accuracies in both the barrel and end-caps of $10\,\mu\mathrm{m}$ and $115\,\mu\mathrm{m}$ in the ϕ and longitudinal directions respectively.[4]

During the long shutdown before Run 2, another pixel layer was installed as the layer that is now closest to the beam pipe [64]. The motivation for this new layer, the **Insertable B-layer** (IBL), was to improve the flavour-tagging efficiency and preserve tracking precision under higher instantaneous and integrated luminosity conditions after the shutdown. The IBL consists of a single layer of pixel modules with a pixel size of $50 \times 250\,\mu\mathrm{m}^2$. The cell size reduction in the z direction improves the resolution of the track parameter z_0 across the entire p_{T} spectrum. With the IBL, b-tagging performance has significantly increased compared to Run 1 by having one additional high-precision hit closer to the interaction point, as this reduces the impact of multiple scattering in track reconstruction [65].

Surrounding the pixel modules is the **Semiconductor Tracker** (SCT). Four layers of microstrip pairs provide eight measurements per track in the SCT barrel, where the SCT end-caps consists of nine double layers per side. The SCT subdetector covers an area of $63\,\mathrm{m}^2$ of silicon. The 4088 modules, divided between the SCT barrel and end-caps, provide a tracking resolution of $17\,\mu\mathrm{m}$ and $580\,\mu\mathrm{m}$ in the ϕ and longitudinal directions respectively [50, 67].

Straw-like drift tubes are used in **Transition Radiation Tracker** (TRT), which is the outermost component of the ID [68]. The tubes with a diameter of $4\,\mathrm{mm}$ are filled with a xenon-based gas mixture and gold-plated tungsten wires of $31\,\mu\mathrm{m}$ are arranged at the centre of the tubes. When

[4] For the barrel components, the longitudinal direction refers to the z direction. For the end-caps, the longitudinal direction is used to indicated distances in R.



(a) (b)

Figure 3.4: (a) Three-dimensional cross section of the barrel of the ATLAS ID. Moving radially outwards from the interaction point, this figure shows the IBL, Pixel detector, SCT and TRT [66]. (b) Cross section of the barrel layers and an end-cap on one side of the detector [50].

a charged particle traverses the TRT, it ionises the gas in the straws. The free electrons drift towards the central wire which is kept at a voltage of $-1.5\,\mathrm{kV}$, where the signal is amplified and read out. In the TRT barrel, 52544 straw tubes of length $1.5\,\mathrm{m}$ are arranged radially around the LHC beam pipe. This provides $R-\phi$ information with a resolution of approximately $120\mu\mathrm{m}$ for charged particle tracks with $|\eta| < 2$ and $p_\mathrm{T} > 0.5\,\mathrm{GeV}$. In each end-cap, 122880 straws are arranged perpendicular to the beam axis to cover the full range in ϕ. Figure 3.4b shows the end-cap configuration for a small region in ϕ. A charged particle track leaves typically more than 30 hits (except in the barrel-end-cap transition region where this number decrease to 22) in the TRT. This large number of hits and the extended measured track length complement the silicon-based layers of the ID, allowing very precise momentum measurements and robust pattern recognition [50].

3.2.3 Calorimeters

In contrast to tracking detectors like the ATLAS ID, where particles by design maintain most of their energy and momentum as they traverse it, calorimeters bring objects to a full stop with the goal to measure their total energy. The ATLAS calorimeters provided a coverage of up to $|\eta| < 4.9$. Figure 3.5 shows a schematic overview of the ATLAS calorimeter system.

To fully absorb these high-energetic particles, a calorimeters need to consist of a high-density material such as lead to maximize the interaction between the incoming radiation and the material. Cascades of secondary particles are produced in this interaction, and in their turn the secondary particles undergo a similar reaction until all the energy in these daughter particles is fully absorbed (and measured) in the calorimeter. The energy of these *showers* can be measured with a scintillating material, where the number of measured photons is proportional to the energy of the interacting particles. A **homogeneous** calorimeter consists a single component that both induces the particle showers and measures the energy. An example a material that has these properties is leaded glass. ATLAS uses **sampling calorimeters**, where layers of dense material are interleaved with a material to detect the energy emitted by particles. This means that the scintillating materials measure only a fraction of the energy. The full energy is reconstructed after final calibrations [69].

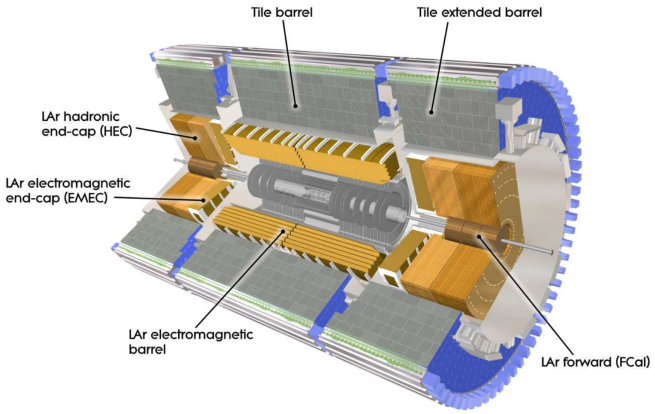


Figure 3.5: Three-dimensional cross section of the ATLAS calorimeter system [66].

The **electromagnetic calorimeters** (ECal) are lead-liquid argon detectors optimised to measure the energy of particle that react via the electromagnetic force: electrons and photons. Moving radially outwards, the absorbers and electrodes are shaped into an accordion-like structure to provide a continuous coverage in the ϕ without cracks. This is immersed into liquid Argon which acts as the active material. High-energy electrons lose most of their energy via bremsstrahlung. Photons create e^+e^- pairs that consequently yield energy deposits in the EM clusters. In the ATLAS ECal, the lead plates are $1.53\,\mathrm{mm}$ thick for $|\eta| < 0.8$ and $1.13\,\mathrm{mm}$ for $|\eta| > 0.8$. In the end-caps, the thickness varies between $1.7\,\mathrm{mm}$ and $2.2\,\mathrm{mm}$, where thinner lead strips lead to a smaller sampling fraction. The barrel part is made of two half-barrels and provides coverage up to $|\eta| < 1.475$. Together, they are $6.4\,\mathrm{m}$ long and have an inner and outer diameter of $2.8\,\mathrm{m}$ and $4\,\mathrm{m}$ respectively. The EM end-caps consist of two wheels that cover the region $1.375 < |\eta| < 3.2$.

Hadrons mainly interact via the strong and nuclear force and deposit less energy via electromagnetic interactions. Most of the hadron's energy can bypass the ECal, so in order to detect them we need to add another layer to our detector: the **hadronic calorimeter** (HCal). Like the ECal, it is a sampling calorimeter but consisting of scintillator tiles and steel as the absorber medium. Tile layers are spatially oriented in the radial direction to reconstruct hadronic showers initiated by the steel layers.[5] These clusters are the main inputs for particle and jet reconstruction. The tile calorimeter consist of a central barrel and two extended barrels and covers the range of $|\eta| < 1.7$, surrounding the ECal barrel. The radial depth of the barrel is approximately $7.4\,\lambda$ interaction lengths [50].

The hadronic end-cap calorimeters (HEC) use liquid-argon sampling technology with copper as the absorbing material and provide coverage of the HCal in $1.5 < |\eta| < 3.2$. Finally, the forward calorimeter (FCal) is a combined electromagnetic and hadronic calorimeter, using copper and tungsten as the absorbing materials respectively.

[5] It is impossible to unsee the resemblance between this layered structure and the Indonesian delicacy *spekkoek* after reading Ivo van Vulpen's popular science book on particle physics [70].

3.2.4 Muon spectrometer

Following particle trajectories from the IP outwards, the only SM particles left outside the calorimeters are neutrinos and muons.[6] Neutrinos are too weakly interacting to be measured in ATLAS, so we forget about them for the moment. Muons, due to their high mass relative to electron, interact much more weakly with the calorimeters and escape with small energy losses. As the name suggests, the **muon spectrometer** (MS) is designed to measure the momentum of muons in the range $|\eta| < 2.7$ and is the outer layer of the ATLAS detector. The MS uses different precision chambers to achieve good momentum resolution in this region and to trigger on muon in the pseudo-rapidity region $|\eta| < 2.4$ [50]. Figure 3.6 shows a schematic overview of the ATLAS calorimeter system.

The monitored drift tubes (MDTs) detect muons with a technique similar to the TRT. Drift tubes are filled with argon gas, where muons ionise the gas and electrons resulting from ionisation leave a signal in the central tungsten wire. The location where a muon passed through can be determined with high precision from the electron drift time. The MS also uses cathode-strip chambers (CSCs). They consist of multiple layers of multiwire proportional chambers filled with an Argon gas mixture, that are segmented in stacks to provide the transverse and radial coordinate. CSCs are characterised by a high response time, a high spatial resolution and high counting rate capabilities. This makes them ideal candidates for the innermost tracking layer of the MS [71].

In the barrel region, the magnetic field lines of the toroid magnets curve the trajectories of the electrically-charged muons. Three concentric cylindrical shells surround the calorimeters with radii of approximately $5\,\mathrm{m}$, $7.5\,\mathrm{m}$ and $10\,\mathrm{m}$ to the beam axis. In the ϕ direction, the barrel region is divided in eight octants to provide full ϕ coverage. In the two end-cap regions, muon segments are organised in wheels that are placed parallel to the end-cap magnets, i.e. like caps for the central barrel. The wheels placed *between* the ID and end-cap magnet are referred to as the small wheels, while the wheels *outside* the end-cap (i.e. at larger $|z|$) are the big wheels. CSCs are placed in the muon small wheels either side of the barrel and cover the pseudo-rapidity region $2 < |\eta| < 2.7$. MDTs are employed in the barrel region and big wheels to cover the region $|\eta| < 2.7$.

In the previous section we have seen how ATLAS measures electrons, photons, muons and (light) hadrons in the form of hadronic jets. What about heavier SM particles, such as third-generation quarks, massive gauge bosons and the Higgs boson? These heavy particles have very short lifetimes so that they decay into lighter particles that *can* be directly detected in ATLAS. Figure 3.7 illustrates the footsteps that the various, light SM particles leave as they traverse the various subdetectors.

3.2.5 Magnet system

Magnets are used to bend trajectories of charged particles for precise momentum measurements in the ID and MS. In ATLAS, the magnet systems consists of one solenoid magnet and three toroid magnets (one barrel and two end-caps). The largest component of the magnet system is

[6] If muons were measured in the calorimeters, there would be less need for the muon spectrometer. And since ATLAS owes its name to the toroidal shape of the magnet which surrounds the muon system, we would need to come up with a new acronym for or continue under the name "ALAS". Alas…

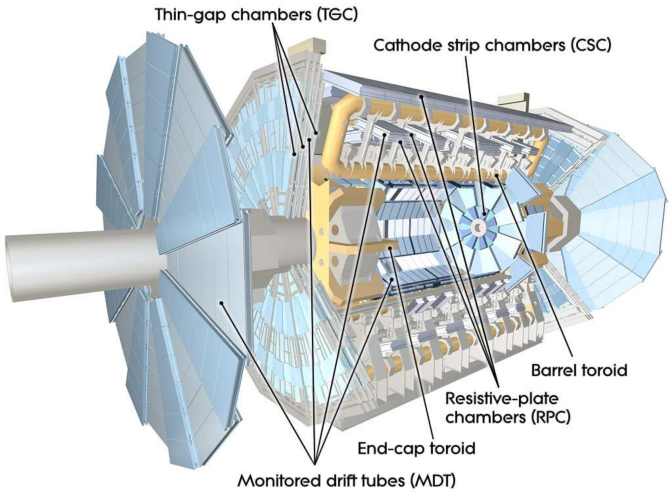


Figure 3.6: Three-dimensional cross section of the ATLAS muon spectrometer [66].

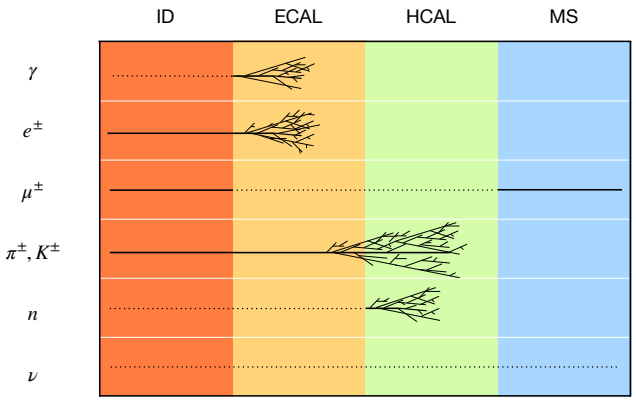
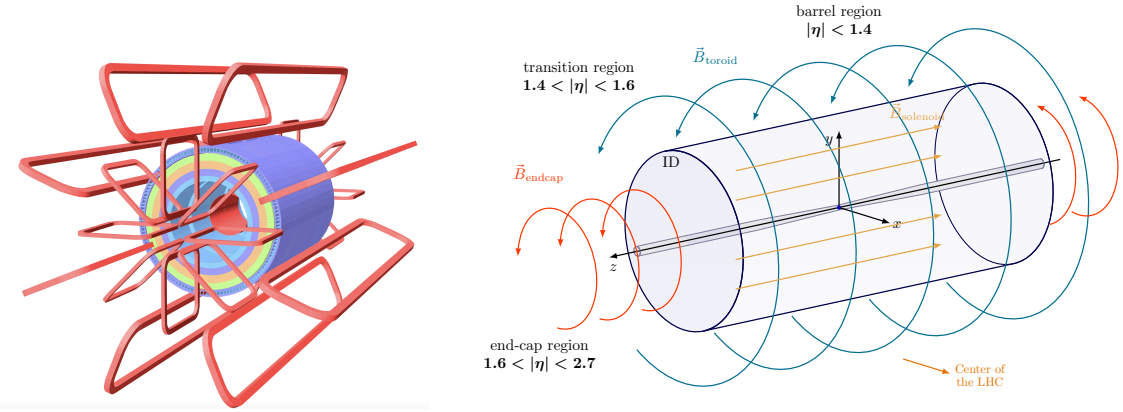


Figure 3.7: Signatures from various SM particles in the ATLAS sub-detectors. Black solid lines indicate a detected signal; either tracking information for the ID and MS or energy depositions in the calorimeter systems. Dashed lines indicate the path of a particle where zero to very little interaction takes place with the detector layer.



(a) (b)

Figure 3.8: (a) Three-dimensional view of the windings of the ATLAS magnet system. The toroids are illustrated as red lines and the central solenoid sits between the ID and calorimeter layers [73]. (b) Direction of the magnetic field lines, corresponding to the solenoid magnet ($\vec{B}_{\text{solenoid}}$), the central toroid ($\vec{B}_{\text{toroid}}$) and end-cap toroids ($\vec{B}_{\text{endcap}}$.)

barrel toroid system. It consists of eight rectangular coils that are placed at different positions in ϕ around the beam axis and just outside the calorimeters. The barrel toroid has inner and outer diameters of $9.4\,\text{m}$ and $20.1\,\text{m}$ respectively, with a length of $25.3\,\text{m}$ in z. The coils surround the MS to provide a magnetic field of up to $4\,\text{T}$ in the ϕ direction. The toroid end-caps are also formed of eight coils and are positioned between the muon small wheel and big wheels on either side of the barrel to bend muons in the forward region. They have a length of $5\,\text{m}$, with inner and outer diameters of $1.65\,\text{m}$ and $10.7\,\text{m}$ respectively. The toroids are complementary to one another; over the range $|\eta| < 1.4$, muon track are bent by the barrel toroid. In the range $1.6 < |\eta| < 2.7$ magnetic deflections are provided by the end-caps. The region over $1.4 < |\eta| < 1.6$ is normally referred to as the *transition region*, since magnetic bending of muons is a result of a combination of the toroid and end-cap magnetic fields.

The solenoid magnet surrounds the ID and provide a $2\,\text{T}$ magnetic field parallel to the beam-pipe to curve charged-particle tracks inside it. A total of $9\,\text{km}$ superconducting wire are wrapped in a cylindrical shape with a length and diameter of $5.3\,\text{m}$ and $2.4\,\text{m}$ respectively [50, 72]. Figure 3.8 shows the structure of the ATLAS magnet system and a rough sketch of the magnetic field lines.

3.2.6 Trigger system and data acquisition

One of the big challenges in ATLAS is processing huge amounts of data during LHC runs. The proton bunch-crossing rate at the interaction point is $40\,\text{MHz}$, so that new events are delivered to ATLAS every $25\,\text{ns}$. Each bunch consists of around 10^{11} protons and as a result multiple protons can interact in a single bunch crossing. The distribution of the mean number of interactions for different years of data-taking in Run-2 is shown in Figure 3.10. With an average $\langle\mu\rangle$ of 33.7, millions of readout channels and limited computer resources we need about think carefully on how to process this information. Fortunately, there is no need to record every single event to do great physics. As shown in Figure 3.11 the total pp collision cross-section is large ($\sim 10^{11}$ pb), but some of these interactions are more interesting than others. For example, the production cross-section of W and Z bosons is already six orders of magnitude smaller than the total pp cross-

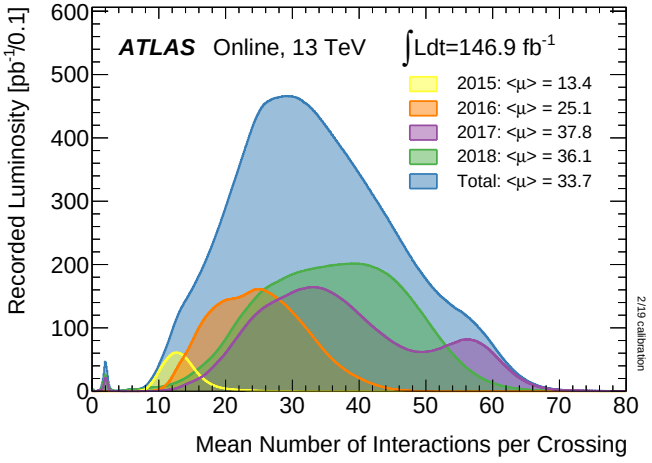


Figure 3.9

Figure 3.10: Mean number of interactions per proton bunch-crossing as recording by ATLAS during Run-2. Figure taken from Ref. [74].

section. The ATLAS **trigger system** receives information from the readout channels during data-taking to decide if an event is recorded or not. This is a two-step process, feeding information to the hardware based Level-1 (L1) trigger first and then to the software based Higher Level Trigger (HLT).

The L1 trigger consists of a Central Trigger Processor (CTP), which combines information from the calorimeter trigger (L1Calo) and the MS trigger (L1Muon). The L1 trigger algorithm isolates one or more regions-of-interest (RoIs) in ϕ, η where trigger thresholds are exceeded. In the RoIs, the system identifies signatures that resemble those coming from jets, muons, electromagnetic clusters or taus. If one or more RoIs meet the L1 trigger criteria, information on these RoIs is sent to the HLT for further processing. This decision must be made within $25\,\mathrm{ns}$. In the meantime, the full event read-outs are stored in pipelines, ready to be sent to the HLT or discarded forever. In this step of the trigger system, the event rate is reduced from $40\,\mathrm{MHz}$ to $100\,\mathrm{kHz}$. Since muons leave very little signal in the calorimeters, the $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers are used to select muon events in addition to the main L1Muon trigger.

In the HLT, events are reconstructed with algorithms much like the ones used in offline processing. Higher-precision reconstruction of full events and reconstruction of the Level-1 regions of interest allow for a refined decision based on additional selection criteria. The HLT reduces the total rate from $100\,\mathrm{kHz}$ to roughly $1\,\mathrm{kHz}$. Events that pass the HLT are not only passed to the output stream for physics analyses, but also to special output streams for detector monitoring during data-taking and detector calibration.

The **trigger menu** is the combined set of trigger chains (at both L1 and the HLT) that are used during data taking. This is a compromise between different physics groups about which events are interesting enough to record! Here, "interesting" is completely determined by the physics goal in mind. The trigger chains in the menu connect selection criteria in the HLT, L1 objects and data streams. As instantaneous luminosity conditions can change during LHC runs, back-up triggers are available to keep the bandwidth maximally saturated and record as much physics as

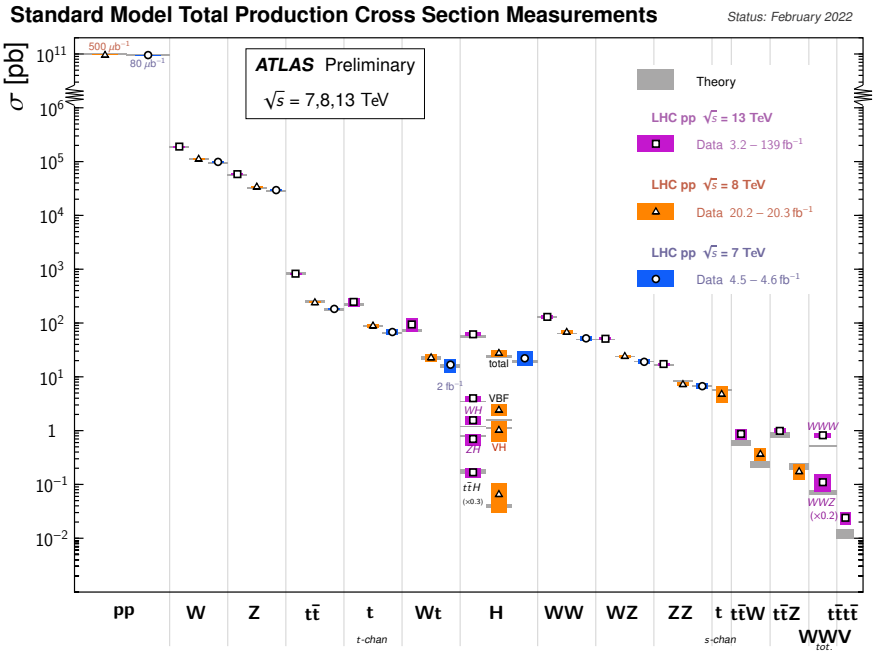


Figure 3.11: Predicted Standard Model cross-sections and as measured by ATLAS of different final states. The production cross-sections of the most rare and most common processes differ by multiple orders of magnitude. Figure taken from Ref. [75].

possible.

Reducing the event rate from $40\,\mathrm{MHz}$ to $1\,\mathrm{kHz}$ comes at a cost: trigger thresholds must be relatively high as not to oversaturate the data acquisition system. An example is the transverse momentum threshold of reconstructed objects. But what if we want to measure low p_T leptons for example? In this scenario, triggers can be *prescaled* by a factor N, meaning that only one out of N events fulfilling the trigger criteria will be recorded. As the instantaneous luminosity decreases during an LHC fill, prescales can be relaxed to maintain the maximum readout capacity of the data acquisition system. Triggers with a prescale factor set to $N = 1$ are referred to as *unprescaled*. Noise, reconstruction efficiencies and pile-up can lead to a signature that is falsely recognised by the trigger as some real object, introducing the notion of "fake" objects. In most physics analyses it is desirable to reduce the relative abundance of fakes, and raising the trigger thresholds may help to achieve this.

3.2.7 Luminosity

Luminosity is one of the most important parameters in collider physics. In this section it is explained why that is the case and how luminosity is measured at ATLAS.

Quantum field theory enables calculations of some cross-section σ, which can be thought of as the probabilities that two particles have an certain interaction mediated by the force carriers in the Lagrangian. Cross-section have the dimension of area[7], in particle physics normally measured in the barn unit (b) where $1\,\mathrm{b} = 10^{-28}\,\mathrm{m}^2$. In reality, ATLAS measures events so we need to

[7] Although cross-section calculations follow the quantum rules, an analogy with the classical world can be made: two basketballs are more likely to collide than two golf balls due to their larger surface area.

relate this number of events to the underlying cross-section. The events are a result of colliding two proton beams, so we can expect the total number of events to depend on characteristics of the beam and on how much they overlap in the interaction point. The event rate is directly proportional to the interaction probability, such that

$$\frac{dN}{dt} = \sigma\mathcal{L}(t) \Rightarrow N = \sigma L \tag{3.4}$$

In this expression $\mathcal{L}(t)$ is the instantaneous luminosity, which captures all the experimental parameters of the colliding beams. This formalism beautifully factors out the event rate into two parts: one where humans have some control over, the luminosity, and one deterministic quantity that relies only on the rules of quantum physics, the cross-section. In general, beam parameters may change over time which introduces the time-dependence. The integral over the instantaneous luminosity is called the integrated luminosity or just "luminosity" for short:

$$L = \int_0^T \mathcal{L}(t)dt. \tag{3.5}$$

At the LHC, luminosity can be thought of the brightness of the proton beams and their overlap. If the beams have a convolved size of $\Sigma_{x,y}$, with N_i protons per bunch in beam i and a revolution frequency f_r, the instantaneous luminosity for a single colliding bunch pair is given by

$$\mathcal{L}_\mathrm{b} = \frac{f_\mathrm{r} N_1 N_2}{2\pi\Sigma_x\Sigma_y}. \tag{3.6}$$

The luminosity quantifies the number of pp interactions in a dataset and is used as a measure for the size of the collected dataset. It has units of inverse area and is for current LHC datasets quoted in units of inverse femtobarns (fb^{-1}). For differential SM cross-section measurements, like the one in this thesis, it is the vital parameter to convert the measured number of events in each bin to a fiducial cross-section and to ultimately make a comparison to theory predictions. It is often also one of the leading sources of the final uncertainty. During data taking, luminosity measurements are used to adjust the trigger and provide feedback to the LHC operators to optimize collisions at the interaction points, which is illustrated in Figure 3.12.

ATLAS measures the delivered luminosity experimentally using dedicated luminosity-sensitive detectors. The guiding principle is that the luminosity is related to the average number of produced particles. For inelastic proton-proton collisions with cross-section $\sigma_\mathrm{inel.}$, the instantaneous luminosity for a single pair of colliding bunches $\mathcal{L}_\mathrm{b}$ can be expressed as

$$\mathcal{L}_\mathrm{b} = \frac{\mu f_\mathrm{r}}{\sigma_\mathrm{inel.}}, \tag{3.7}$$

where the average number of inelastic interactions per bunch crossing is described by the parameter μ. Having multiple interactions per bunch crossing is also referred to as *pile-up*. There are multiple bunches going around the LHC tunnel, so that the total instantaneous luminosity is given by

$$\mathcal{L}_\mathrm{inst.} = n_\mathrm{b}\frac{\langle\mu\rangle f_\mathrm{r}}{\sigma_\mathrm{inel.}}, \tag{3.8}$$

where $\langle\mu\rangle$ is averaged over *all* colliding bunches. The number of colliding bunch pairs n_b reached

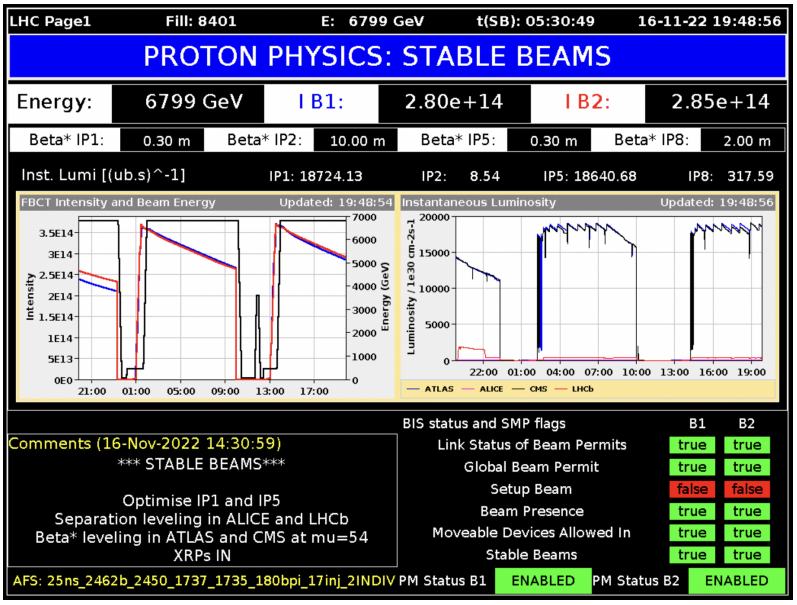


Figure 3.12: Screenshot from LHC Page1 [76] during Run-3 data taking at $\sqrt{s} = 13.6\,\text{TeV}$ on 16 November 2022. The left panel shows how the beam intensities drop over time. The recorded instantaneous luminosity at ATLAS, as shown in the right panel, would decrease following the same trend. However, below a given value the trigger keys are changed in the ATLAS control room so that the luminosity jumps back up to saturate the data-acquisition system. This creates the sawtooth pattern, until the trigger thresholds and prescales reach a level with only a very small proportion of rare events (with a small production cross-section). At that point it is more worthwhile to dump the beam and restart a new fill cycle.

a maximum of 2544 during Run-2 data taking in 2017 and 2018 [77]. It is an important observation that the combination μ_X/σ_X is constant for any events of any type X: if the cross-section changes by some factor, the average number of interactions for process X will scale inversely by the same factor. Defining a "visible" number of interactions per bunch crossing and the corresponding cross-section, $\mu_{\text{vis.}}$ and $\sigma_{\text{vis.}}$ respectively, this can now be directly related to the pile-up parameter μ as follows:

$$\frac{\mu}{\sigma_{\text{inel.}}} = \frac{\mu_{\text{vis.}}}{\sigma_{\text{vis.}}}. \tag{3.9}$$

The visible cross-section is a calibration constant determined in dedicated scans. A real-time measurement of $\mu_{\text{vis.}}$ is then enough to determine the instantaneous luminosity for every bunch at that given moment.[8] A luminosity block (LB) is the time interval over which $\mu_{\text{vis.}}$ is measured. The integrated luminosity is obtained by multiplying the instantaneous luminosity by the duration of the LB.

In 1968, Dutch physicist Simon van der Meer proposed a technique to evaluate the luminosity at an intersection point by displacing the colliding beams with respect to each other [78]. The van der Meer (vdM) method is still the main luminosity calibration method used at the LHC. In this method, the uncalibrated luminosity $R(\Delta x)$ is measured as a function of the beam separation Δx, with $\Delta x^{\text{max}} = -0$ corresponding with head-on collisions. The displacement is varied step-wise in a scan and this scan is repeated in the other transverse direction y. The convolved beam size

[8] This assumes that the calibration constant $\sigma_{\text{vis.}}$ is indeed constant in time. In practice, no luminosity algorithm is perfectly stable in time. Hence different algorithms are used for cross-checks and obtaining the most precise luminosity measurement as possible.

Data sample	2015	2016	2017	2018	Combined
Integrated luminosity (fb^{-1})	3.24	33.40	44.63	58.79	140.07
Total uncertainty (fb^{-1})	0.04	0.30	0.50	0.64	1.17
Total uncertainty (%)	1.13	0.89	1.13	1.10	0.83

Table 3.1: Total integrated luminosity per year of data taking at ATLAS during Run-2 at $\sqrt{s} = 13\,\mathrm{TeV}$.

is then expressed as

$$\Sigma_x = \frac{1}{\sqrt{2\pi}} \frac{\int R(\Delta x)\mathrm{d}\Delta x}{R(\Delta x^{\max})}, \tag{3.10}$$

and likewise for the y-direction. In this ratio the overall normalisation cancels out. This allows different methods to be used to determine the scan curve, as long as the corresponding quantities scale with the luminosity. Combined with Equation. (3.6) this fixes the calibration as

$$\sigma_{\mathrm{vis.}} = \mu_{\mathrm{vis.}}^{\max} \frac{2\pi\Sigma_x\Sigma_y}{N_1 N_2}, \tag{3.11}$$

where $\mu_{\mathrm{vis.}}^{\max}$ is the maximum visible interaction rate on the scan curve.

In Run-2 the primary detector to measure the visible number of interactions was the LUCID-2 Cherenkov detector[9], consisting of 16 photomultiplier tubes (PMTs) in the forward regions ($5.561 < |\eta| < 5.641$) of the ATLAS detector [79]. Charged particles from collision events going through the quartz window of the PMTs produce Cherenkov light that can be detected. A 'hit' is recorded when the signal in the PMTs reaches above a certain threshold. From the average probability to have a hit in any PMT,

$$P^{\mathrm{hit}} = 1 - \exp(-\mu_{\mathrm{vis.}}), \tag{3.12}$$

the visible number of interactions $\mu_{\mathrm{vis.}}$ is probed.

During Run-2 the total recorded luminosity by ATLAS is $140\,\mathrm{fb}^{-1}$. Not all recorded data passes the data quality requirements. Figure 3.13 shows the luminosity for the delivered, recorded and "good for physics" data.[10] A breakdown of the integrated luminosities per year of data taking and the corresponding uncertainties are summarised in Table 3.1.

[9] This is complemented by offline track-counting algorithms, calorimeter-based algorithms and measurements from the ATLAS beam conditions monitor.

[10] This luminosity plot was released by ATLAS before the final Run-2 luminosity calibrations were finalised, which changed the Good for Physics total luminosity from $139\,\mathrm{fb}^{-1}$ to $140\,\mathrm{fb}^{-1}$.

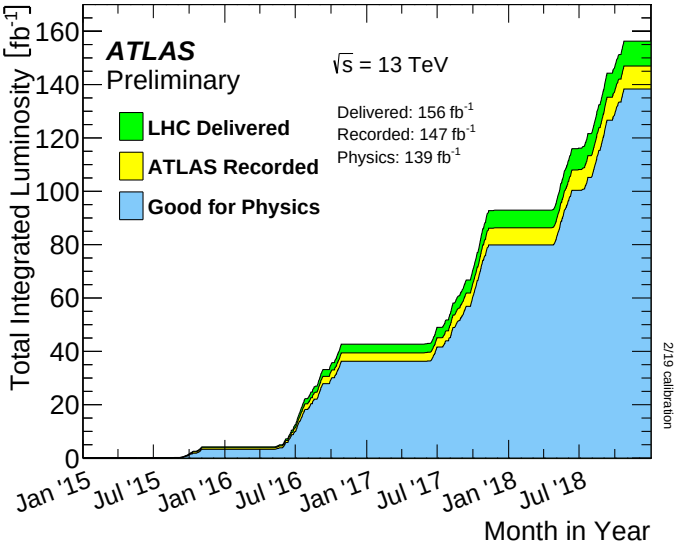


Figure 3.13: Cumulative integrated luminosity for Run-2, including data taking from 2015 to 2018 [80]. The plot shows the total luminosity delivered by the LHC in green, the amount of data that was actually recorded in yellow and the data that passed data quality requirement for physics analyses in blue.

Chapter 4

Theory meets experiment

The main goal in particle physics is to verify that empirical observations match the theoretical predictions, and thus ultimately find an accurate model to describe Nature at the most fundamental scale. We have seen that the Standard Model is the prevailing paradigm in particle physics, while the LHC and its experiments enable experimental measurements of proton-proton collisions. Unfortunately, experimental measurements come in the form of electronic signals in photomultiplier tubes and silicon strips, while our prediction stems from a quantum field theory with a set of rules to describe interactions between fundamental particles. This raises two important questions: *where* and *how* do we tie together these two extreme ends? The goal of this chapter is to address these points.

The answer to the first part of this question is: "somewhere in the middle". More specifically, the representation of theory and experiment in the form of (individual) events and their associated properties is a natural choice. As we will discuss later in this chapter, this definition is still ambiguous, and experimentalists have a choice to make in which representation they compare the data from their experiment to theoretical predictions.

The question of "how" needs answers from both theory and experiment. From the theoretical side, Monte Carlo event generators provide the possibility to translate fundamental interactions from a QFT into simulated collisions like those measured experimentally. Additionally, the interaction between those events and the subcomponents of the detector can also be simulated. Approaching the event-level from the experimental side, model-based assumptions are needed to interpret detector signals as the presence of physical objects, and consequently to reconstruct (SM) particles and their kinematic properties. Ultimately, the combination of physical objects and detector activity can be interpreted as some (SM) process that led to the measurement of this event.

In Section 4.1 we describe event generation and simulation from a theoretical starting point. Object and event reconstruction in ATLAS, starting from detector outputs, is discussed in Section 4.2. Finally we discuss precise definitions of the middle ground between theory and experiment in Section 4.3, and techniques to correct experimental data for detector effects in Section 4.4.

4.1 From Lagrangians to reconstructed events

In Section 2 we have seen that the SM provides a description to understand quantum-mechanical interactions between particles, but more work is required to map theoretical models to a representation that can be compared with real-world data measured at the Large Hadron Collider. In ATLAS, and any other LHC experiment, we study "events". At the most basic level, an event can be defined as a proton-proton collision leading to a number of outgoing particles. In such events, conservation laws dictate conservation of kinematic quantities such as energy and momentum, and of discrete quantum numbers such as lepton number or spin, between the initial and final state. But there are more steps that hinder a direct comparison between a theoretical model and LHC data. For example, collider experiments measure collimated sprays of particles (jets), deal with multiple particle collisions occurring simultaneously (pile-up) and have limited resolutions. Over the last decades, a huge amount of work went into the generation and simulation of collision events. This section takes a whistle-stop tour through all these developments. Taking a (B)SM Lagrangian as the starting point, we summarise the steps to translate this to predictions of fully exclusive final states in terms of hadrons and leptons, and to simulate the interaction of these particles in a detector.

For complicated final states like this, it is typically not possible to derive an analytic function to capture all these effects from first principles. Numerical methods based on pseudo-random numbers provide the best solution to date [81]. Because the computations rely on Markov Chain Monte Carlo techniques, the software algorithms that provide such calculations to simulate events are called Monte Carlo (MC) event generators. There are many MC event generators publicly available. For example, HERWIG [82–84], SHERPA [85, 86] and PYTHIA [87, 88] are *general-purpose* generators that handle the calculations of matrix elements (MEs), parton shower (PS) and subsequent hadronisation. Some generators are tailored to a specific part of this "generator chain". For example POWHEG-BOX [89–91] and MADGRAPH5a_MC@NLO [92] are designed to calculate next-to-leading-order (NLO) matrix elements. These are then interfaced to an external programs for the parton shower and subsequent hadronisation.

MC generators used different underlying methods to simulate particle collisions. For example, there are different models for hadronisation and different implementations of the parton shower. To smoothly link together the various steps in event generation using different generators, common file formats have been introduced. The Les Houches Event (LHE) file format is used to store partonic events [93], which can be used as input to generators that shower and hadronise the events in the LHE file. The HepMC3 file format [94] is the *de facto* standard event record for events after parton showering and hadronisation.

Models depend on theoretical inputs such as the parton distribution function (PDF) or QCD scale choices, i.e. the renormalisation scale μ_R and factorisation scale μ_F.[1] The *nominal* samples corresponds to the default choice of parameters. In the case of PDF or scale choices, the parameters of the calculation in the model are varied and from this a multiplicative reweighting factor is obtained and applied to the nominal configuration. Each event then has multiple weights associated to different uncertainties, that are propagated through the MC chain as event weights.

[1] Renormalisation in QCD introduces an arbitrary, unphysical renormalisation scale μ_R. The dependence of physical observables must be independent of this scale, but truncation of the perturbative series spoils this cancellation. The factorisation scale μ_F is introduced to absorb collinear divergence of parton splittings into bare PDFs.

Ultimately, when events are binned into histograms corresponding to physical observables, this gives different variations of the histograms that can be combined into uncertainty bands. There are algorithmic variations however, such as the choice of parton shower or hadronisation model, that cannot be cast into multiweights because they rely on fundamentally different algorithms to run the event generation. In such cases one relies on the explicit variation of the different algorithms and (for example) taking the difference as a systematic uncertainty.

The different steps in the simulation of proton-proton collisions are summarised in Figure 4.1. The structure of this section loosely follows this diagram top to bottom. First, we introduce parton distribution functions that describe the contents of a proton in Section 4.1.1 and multiple-parton interactions in Section 4.1.2. Then, we see MC generators at work in the computation of matrix elements (Section 4.1.3) and the simulation of parton showers (Section 4.1.4). Hadronisation is covered in Section 4.1.5. Finally, the simulation of detector interactions of final state particles is discussed in Section 4.1.6. This part of the thesis benefits greatly from the detailed works of literature reviews that are available on MC event generators, in particularly Refs. [81, 95, 96].

4.1.1 Parton distribution functions

In the 1960s and 1970s deep inelastic scattering experiments provided empirical evidence for models that describe protons as composite particles composed of smaller constituents, which are dubbed "partons" [97–99]. Partons correspond to (anti-)quarks and gluons, and parton distribution functions (PDFs) describe the fraction of the proton's momentum that each parton carries. When a proton carries a momentum P, the partons will have a momentum p_i of $p_i = x_i P$, where $0 \leq x_i \leq 1$ is the momentum fraction for each parton. PDFs are instrumental in the simulation of proton-proton collisions, because at the most fundamental level the *partons* undergo interactions that we describe with quantum field theories. In order to generate realistic events, accurate models to describe the momenta of partons are crucial in LHC physics.

The collision cross-section factorises in terms of the analytical partonic cross-section that we can actually calculate in perturbation theory, or in explicit form

$$\sigma_{\text{interaction}} = \sum_{i,j \in q,\bar{q},g} \int_0^1 \mathrm{d}x_1 \int_0^1 \mathrm{d}x_2 f_i(x_1, \mu_F) f_j(x_2, \mu_F) \hat{\sigma}_{ij} \,. \tag{4.1}$$

The partonic cross-section ($\hat{\sigma}_{ij}$) is typically distinguished from the proton-proton cross-section with a hat, where i, j label the type of the incoming partons. The parton distribution functions $f_i(x_n, \mu_F)$ give the probability of finding a parton of species i with momentum fraction x_n in the incoming proton n that is probed at a scale μ_F. The integrals sum over all momentum fractions x_n in the PDF, where the discrete sum adds up all different possible combinations of interacting partons. These are probability densities so they must satisfy the normalisation condition

$$\sum_i \int_x f_i(x_i, \mu_F) \mathrm{d}x_i \equiv 1 \,, \tag{4.2}$$

to ensure that the momenta of the partons add up to the total momentum of the proton.

Parton distribution functions depend on non-perturbative strong interactions and cannot be cal-

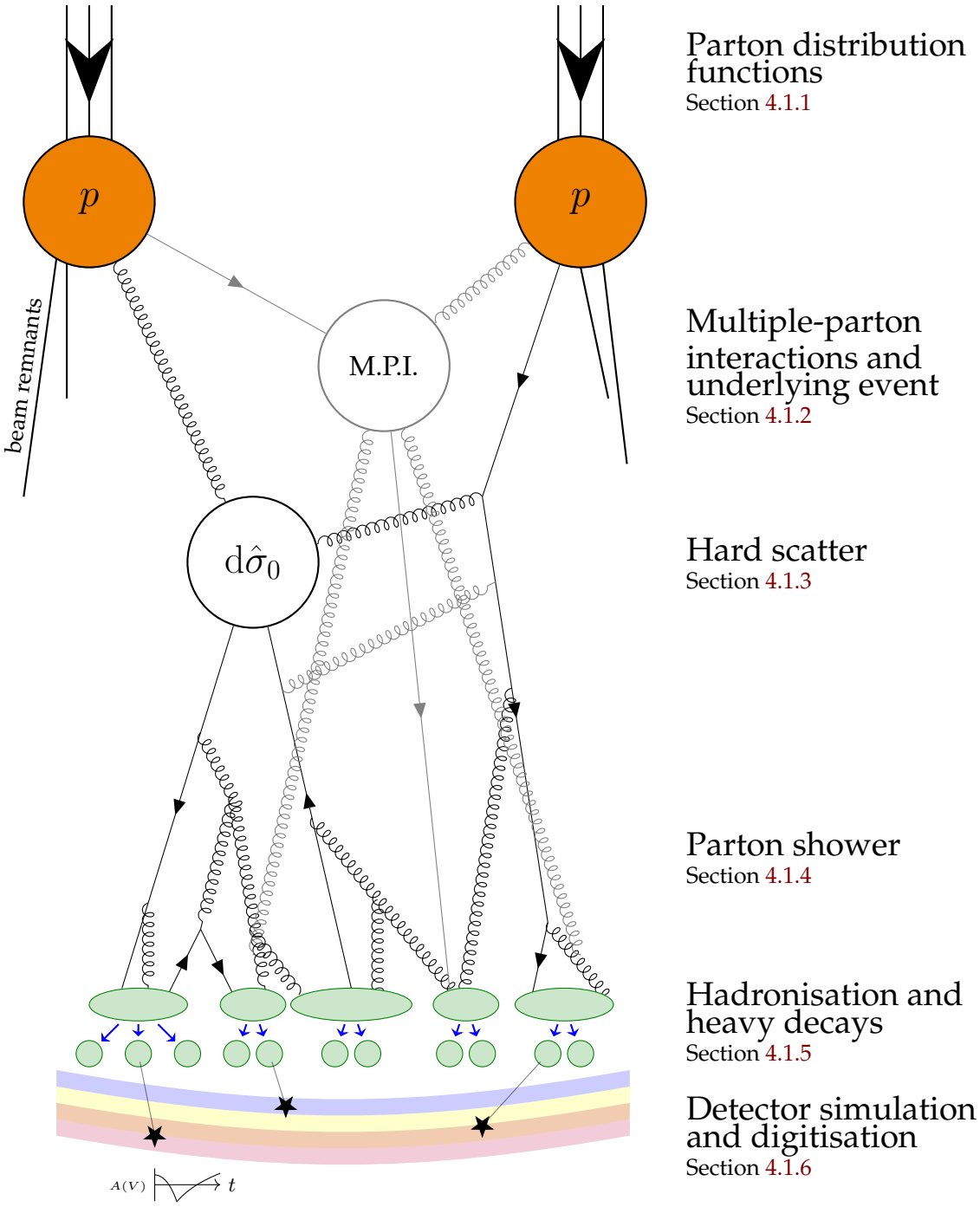


Figure 4.1: Simplified representation of the simulation of proton-proton collisions. Two incoming protons (orange circles) with an internal structure of quarks and gluons cause a hard scatter event as well as multiple-parton interactions. Partons undergo showering (multiple quark and gluon lines), hadronisation (green ellipses) and hadron decays (blue arrows), after which interactions of final-state particles with the detector (black stars) is emulated. After pile-up overlay, the final step is the digitisation of detector signals, indicated at the very bottom by some small negative voltage in one of the readout channels.

culated analytically from first principles for that reason. The extraction of PDFs from experimental data is an active field of research. PDF collaborations like MSHT [100–103], CTEQ [104–108] and NNPDF [109–112] perform global fits to deep inelastic scattering (DIS) and LHC data to provide PDF sets that can be employed in MC event generators.

The resulting PDFs from these groups may vary, due to the use of different methodologies to extract them. Each group provides fits for different orders in α_S. However, the general recipe to extract PDFs from experimental data follows the same methodology.[2] The main idea is that the evolution of PDFs as a function of their factorisation scale μ_F is calculable in perturbation theory. One starts with a scale Q_0^2 where perturbation theory is valid, for which the PDFs are parametrised in a functional form of x_n. With the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations, PDFs are evolved to scale Q^2 corresponding to a measured dataset. These are then used to calculate cross-sections and evaluate the agreement with datasets in a goodness-of-fit test. The overall agreement is calculated by summing over the results for different datasets. The free parameters in the PDFs are then optimised in an iterative procedure [69]. Uncertainties in the PDF sets, for example from the choice of parametrisation or to account for missing higher-orders in the perturbative series, are propagated to the final MC uncertainty estimates.

4.1.2 Multiple-parton interactions and underlying event

When two hadrons collide, nothing prevents more than one parton-parton scattering to occur. Most *multiple-parton interactions* (MPIs) are relatively soft, but high-energy partons from MPI can lead to events with for example heavy SM particles or QCD jets with large transverse momenta. MPIs are observed experimentally (e.g. in ATLAS [113]), and models to describe MPI effects do so by treating interactions independently or by accounting for colour flows between multiple interactions [114, 115].

Underlying event is an umbrella term for all the interactions that contribute to the measured activity in addition to the main hard-scattering event of interest. *Pile-up events* originate from other proton-proton collisions, either in the same bunch crossing (in-time pile-up) or from a different bunch-crossing (out-of-time pile-up), and are overlaid on hard-scatter events in the detector simulation step. Furthermore, remnants of the colliding beams can leave signatures in the detector and must be included in hadronisation models to conserve colour. In Section 4.1.4 we discuss how QCD leads to initial and final state radiation. Additionally QED radiation (in the form of photons) is emitted from outgoing particles.

4.1.3 Calculation of hard-scatter processes

The partonic cross-section for a process $i \to f$,

$$\hat{\sigma}_{i\to f} = \hat{\sigma}_{i\to f}(\mathcal{M}_{i\to f}, \mathrm{d}\Phi_f; \mu_R, \mu_F) \,, \tag{4.3}$$

depends on the matrix-element (M.E.) $\mathcal{M}_{i\to f}$, a kinematic phase-space factor $\mathrm{d}\Phi_f$, renormalisation scale μ_R and factorisation scale μ_F. In this notation, i and f indicate the initial and final state of the scattering process respectively. The matrix-element is written as a sum over all pos-

[2] An exception is the NNPDF collaboration, where "NN" stands for neural network, that uses a different approach based on techniques in machine learning.

sible Feynman diagrams for $i \to f$. The M.E. is squared and averaged over the initial-state spin and colour, and summed over the quantum numbers in the final state. The differential phase-space element $\mathrm{d}\Phi_f$, indicating the number of momentum states accessible to a given final state, is integrated over all possible outgoing momenta.

Matrix-elements are calculated using MC sampling techniques. This becomes computationally more heavy if the number of final-state particles is increased or when higher-orders of QCD or EW couplings in the Feynman diagrams are included (through $\mathcal{M}_{i\to f}$). We cannot calculate matrix-elements up to an arbitrarily large order in QCD and EW scales, which include "real emissions" with additional particles in the final state, as well as virtual corrections via loop diagrams. Adding high-multiplicity partons ("legs") as real corrections to the M.E. calculation is possible with current event generators. At leading order this is fully automated up to an arbitrary number of legs, but actually computing these diagrams is restricted by computing power. The conventional notation for multi-leg calculations is $X + \mathrm{n}j@(\mathrm{N})\mathrm{LO}$, where X is the Born level process, n is the additional number of partons j with respect to the Born level process, and the perturbative order in the strong coupling at which the process is calculated is indicated after the at sign @. Going to higher orders in QCD is mainly restricted by the virtual corrections in loop diagrams, and is not fully automated up to an arbitrarily high order of α_S.

In QCD, renormalisation and factorisation scales (μ_R and μ_F) are introduced to deal with infinities that arise in the calculations of loop diagrams. In an exact calculation, i.e. a calculation that includes all orders of α_S, the dependence on these scales cancel. However, the truncation of the perturbative series leads to a result that depends on the choice of μ_R and μ_F. These scale-dependencies are quantified as *scale uncertainties*, which are in fact often the dominant source of uncertainty in the final MC predictions. The scale uncertainties can be estimated by varying in the renormalisation and factorisation scale by some factor r. Assume we can calculate a cross-section at momentum-scale Q up to some order k in α_s, i.e.

$$\sigma(Q) \sim \sum_{n=0}^{k} c_n(Q,\mu)\alpha_s^n(\mu) + \sum_{n=k+1}^{\infty} c_n(Q,\mu)\alpha_s^n(\mu) \equiv \sigma_k(Q,\mu) + \Delta_k(Q,\mu), \tag{4.4}$$

where $\sigma_k(Q,\mu)$ represents the calculable part and $\Delta_k(Q,\mu)$ is a term that is not calculated analytically. The usual procedure is to calculate $\Delta_k(Q,\mu)$ to obtain a scale uncertainty for the truncation of the series in k. This is estimated by fluctuating the nominal scale by a factor r and assigning the difference of this result with the nominal result as the scale uncertainty, i.e.

$$\Delta_k(Q,\mu) \sim \max\Big[\sigma_k(Q,\frac{\mu}{r}),\sigma_k(Q,r\mu)\Big] - \sigma_k(Q,\mu) \sim \alpha_s(\mu)^{k+1}, \tag{4.5}$$

where typically $r = 2$.

From this equation, we see that the inclusion of higher orders of α_s in the M.E. calculation leads to a reduction of the scale uncertainties of a factor α_s . Calculating matrix-elements to higher fixed-order is an active field of research, and we will see in the next chapter that these need to be matched to a parton shower to produce the exclusive final states for the experiments.

Event generation takes up a significant portion of the LHC's computing budget, where demands of higher perturbative accuracy cause major concern for staying within the available computational budget [116]. Ongoing efforts try to address this issue, for example by making matrix-

element calculation run more efficiently [117] and developing GPU-portable MC event generators for M.E. calculation [118].

4.1.4 Parton shower

After calculation of the hard-scatter in proton-proton collisions, we are one step closer to emulating LHC events as they occur one hundred metres underground near the Franco-Swiss border. The highly energetic outgoing partons from the hard-scatter and/or MPI may radiate soft, collinear quarks and gluons to form collimated sprays of particles. It is computationally too expensive to include these radiated particles in the M.E. calculation analytically. *Parton shower* algorithms are developed to model these QCD-splittings efficiently and perturbatively. In a parton shower, quarks radiate gluons, while gluons split to quark-antiquark pairs or additional gluons. Progressively this evolves the momentum scale of partons from the hard scale Q^2 to a threshold scale at which perturbative QCD no longer holds and partons hadronise, Λ_{QCD}.

The emission distribution of one additional parton can be written in a factorised version, independent of the underlying hard process as

$$\mathrm{d}\sigma \approx \sigma_0 \sum_{\text{partons},i} \frac{\alpha_s}{2\pi} \frac{\mathrm{d}\theta^2}{\theta^2} \mathrm{d}z P_{ji}(z,\phi)\mathrm{d}\phi, \tag{4.6}$$

where θ and z are the emission angle and energy fraction of the emitted the parton relative to its parent particle, and $P_{ji}(z,\sigma)$ a set of flavour- and spin-dependent functions [81]. Because the $n{+}1$-parton process factorises in terms of the n-parton one, this formula can be applied iteratively to model multiple parton emissions.

The Sudakov form factor describes the probability that a parton with virtuality $t \equiv p_i^2$ evolves from t_0 to t without emitting resolvable partons [69]:

$$\Delta_i(t_0,t) = \exp\Big[-\sum_j \int_{t_0}^{t} \frac{\mathrm{d}t'}{t'} \int_{z_{\min}}^{z_{\max}} \mathrm{d}z P_{ji}(z)\Big]. \tag{4.7}$$

Free partons in the final state do not exist, and therefore do not become fully on-shell. This motivates a cut-off value for the virtuality t_{cutoff}. Considering a parton of virtuality t_p, parton-shower algorithms to simulate the additional radiation are defined as follows. First, we generate the virtuality of a potential splitting t_{p+1} by solving $\Delta_i(t_p, t_{p+1}) = r_1$, where r_1 is a random variable uniformly distributed between 0 and 1. If the virtuality t_{p+1} is less than the resolution parameter t_{cutoff} the stopping criterion for the parton shower is reached. Otherwise, the energy fraction z and azimuthal angle ϕ for the emitted parton are generated. This procedure is repeated for all partons, including emitted partons.

Equation 4.6 has two divergences. The collinear limit $\theta \to 0$ is obvious to see, where the soft limit $z \to 0$ is implied through $P_{ji}(z,\phi)$. To solve the collinear divergence we impose a resolution criterion, because in any physical measurement it is not possible to resolve a single parton from an exactly collinear pair of partons with the same quantum numbers. The obvious choice is to say that partons are only resolvable if the relative transverse momentum k_T lies above some cut-off value, and that otherwise two partons can not be distinguished experimentally. After a change of variables this is equivalent to a minimum cut on the virtuality. Wide-angle soft emissions can

be included by using the angle θ as the ordering variable in the parton shower algorithm. This construction of a parton shower is called an angular-ordered shower, implemented for example in HERWIG [119]. An alternative model is the k_T-ordered shower or dipole shower, where emitted partons recoil off a pair of partons and k_T is used as the ordering variable [120–122]. These models are implemented in HERWIG, PYTHIA and SHERPA. The algorithmic choices made inside the parton showers can be associated with a systematic uncertainty on the final MC predictions.

We have now seen how parton showers describe QCD radiation of outgoing particle from the hard scatter, i.e. final-state radiation (FSR). This description works similarly for parton splittings of the *incoming* partons, leading to initial-state radiation (ISR) [123]. The difference for ISR is that the radiation is now simulated in a backwards step, dressing only the partons that take part in the hard scatter with additional radiation.

Matching and merging

Parton showers provide modelling of soft and collinear radiation, while matrix elements are accurate in the regime where partons are hard and well separated. Taking the best of both worlds, these methods can be combined using a suitable matching prescription [89, 90, 124] in order to provide the best MC prediction across those two different regions of phase-space, taking care to avoid double-counting of phase space described by both matrix element and the parton shower. These matching procedures have been extended to NNLO-accurate matrix-element calculations more recently, see e.g. Refs [125–127].

In merged predictions, matrix elements for diagrams with extra partons are calculated and combined with the matrix-element calculation for the Born level process. In this case, there is not only a possible overlap between the parton shower and the matrix elements, but also between different matrix elements. Techniques of *merging* are designed to remove this overlap, which introduces an ordering procedure with associated cut-off scale Q_{cut}. The cut-off scale is chosen arbitrarily, so a smooth transition with no dependence from the chosen cut-off value is desired. There are various merging algorithms that implement this , including CKKW [128, 129], MLM [130] and CKKW-L [131] at LO and MEPS@NLO [132] at NLO.

4.1.5 Hadronisation

Due to the confinement property of QCD, free quarks and gluons are not observed in nature. As the parton shower evolution approaches the non-perturbative regime interactions between partons become strong and they are recombined in colour-neutral hadron states. This process is called hadronisation. Because this occurs at an energy scale of the order the QCD scale (Λ_{QCD}), where the strong coupling constant α_s becomes large, it cannot be calculated in perturbation theory. There are two main phenomenological models that describe hadronisation: the Lund string model and the cluster model. In some MC generators, such as HERWIG 7 and SHERPA, the cluster model is implemented. Others, for example PYTHIA8, employ the Lund string model. Since physics is in this case limited to phenomenological models, it is important to understand any differences on the final results when using different models. If the differences are clearly understood, this can be used to assess the systematic uncertainty associated with the choice of a particular hadronisation model.

The *Lund string model* relies on the fact that the QCD potential between a quark and antiquark becomes linear at large distances. As the quarks are pulled apart, the energy stored in the gluon field increases. Eventually the energy stored in this "QCD string" can become so large that the string breaks into two new strings. The process of string-breaking can continue until the strings do not have enough energy to be separated into substrings. When the energy falls below this threshold, quarks and antiquarks combine into colourless hadrons. [133]

The *cluster model* is based on the observed *pre-confinement* property of parton showers. This means that colourless clusters can be formed from the partons in the parton shower. Any gluons from the parton shower are split into quark-pairs. The mass distributions of these clusters can be calculated from first principles and has no dependence on the momentum transfer Q of the hard process. This means that the cluster formation has the same properties in all events. The clusters are finally matched to a spectrum of hadrons, where heavy clusters may "decay" into lighter clusters first. [134]

As part of this PhD work, the effect of using different parton shower and hadronisation models was studied in $t\bar{t}$ and dijet events. This study, which shows discrepancies in mass-dependent observables for both $t\bar{t}$ and dijet events, is documented in Appendix A.

Regardless of the hadronisation model that is used, heavy hadrons and excited states have short lifetimes and decay within the detector volume. In fact, heavy decays are the main process of interest for flavour-experiments like LHCb. There are dedicated software tools available to decay excited states, one of which is the EVTGEN package [135] that focuses on heavy-quark states.

So far we have obtained a simulated event-record with detailed information of the outgoing particles and their associated four-momenta. This concludes the "theoretical" aspect of event generation, which has so far been independent of the detector that is used to measure events empirically.

4.1.6 Detector simulation and digitisation

The only experimental access we have to measure proton-proton collisions is through a detector. Therefore, the last steps in the production chain are the simulation of interactions between final-state particles and the detector components, and finally the digitisation of detector signals. To do this, there are various fundamental complications that we need to acknowledge. First, due to the detector geometry and used materials there is a finite experimental resolution. Second, we have seen in Section 4.1.2 that due to MPI and the underlying event, different interactions give rise to multiple detector signatures in a single bunch crossing. These can be hard-scatters or soft QCD (minimum-bias) events.

The main software for full simulation of the ATLAS detector is GEANT4 [136–138]. This fully simulates magnetic bending in the ID and MS, hits in the trackers and energy depositions in the calorimeters. With many particles in the event record as a result of the parton shower, hadronisation and hadron decays this requires significant CPU resources: the processing of a single event typically takes several minutes. While GEANT4 is still the default simulation software for centrally produced ATLAS samples, fast simulation (FastSim) workflows like AtlFast3 [139] mitigate the computational overload of detector simulation with efficiency maps, smearing tools and machine learning techniques for the most expensive aspects of the full detector simulation.

Development of FastSim technologies is needed to keep up MC production with growing experimental datasets and is an active research area for all major LHC experiments. These methods are not yet as accurate as the full GEANT4-based detector simulation, but great improvement in both speed and accuracy has been achieved over the last decade [139].

Up until this point the simulation of pile-up was dealt with separately from the simulation of the hard-process and MPI. Digitisation is the final step, where detector hits are converted into electronic signals and timings ("digits") that are read out from the experiment. This is also where pile-up is introduced, because the detector material and consequently the TDAQ system respond to the sum of activity that is measured. Additional pile-up collisions are overlaid over the hard-scattering event, to match conditions conditions between "real" data-taking with ATLAS and detector simulations. Pile-up events are based on minimum bias events, which includes mostly soft QCD processes, simulated with Pythia8. At analysis stage, MC events are weighted to reproduce the distribution of the average number of interactions per bunch crossing $\langle\mu\rangle$ in data, as shown in Figure 3.10, in a process known as "pile-up reweighting" [69].

4.2 From electronic signals to reconstructed events

The goal of this section is to describe how electronic signals are turned into reconstructed events. Ultimately this allows us to apply selection criteria, or "cuts", on events during analysis, to demarcate a region of phase-space where we can map the experimental measurement onto theory. The first step in this process is to determine how different stable, final-state particles interact with the detector materials. In this context, "stable" means that a particle does not decay within the ATLAS detector volume. We then classify digital detector signals, either from experimental data or MC, signals as e.g. jet or lepton candidates and infer kinematic properties of the underlying objects via the calibration procedure.

Many particles created in LHC collisions have lifetime so short that they decay into lighter particles before they can be detected by ATLAS. Therefore, we first discuss the reconstruction of detector-stable objects in Section 4.2.1. Then, Section 4.2.2 summarises the procedure to classify more complex objects from stable objects, such as heavy-particle candidates that decayed within the detector volume but also kinematic features such as the sum of (missing) transverse momentum.[3] Cuts on objects or more general event properties define measurement regions to scrutinise a certain region of phase-space. Section 4.2.3 discusses the typical event selections that are common to most LHC experiments.

4.2.1 Reconstruction of detector-stable objects

Information from different detector subsystems is combined to reconstruct object such as lepton, photon and jet candidates. This section covers the objects that are stable within the detector volume.

[3] We never identify the underlying process in a *single* event, but in a large enough sample in a well-defined phase-space we infer relative fractions of contributing processes.

Electrons

In ATLAS, electrons are defined as objects consisting of a topological cluster with energy deposits in the EM calorimeter that are matched to a track in the ID. This discussion focuses on the reconstruction of *prompt* electrons, referring to electrons that do not originate from b-hadron decays.

EM calorimeter cells have expected noise from pile-up and electronics. We can define a significance N as the ratio between deposited energy in cells with respect to the noise level. Proto-clusters are initiated from cells in the calorimeters where energy depositions reach above the expected noise with significance of 4. In an iterative procedure, neighbouring cells with a significance of 2 are added to the cluster(s). When two proto-clusters contain the same cell above noise threshold, they are merged. In a last step, when no neighbouring cells match the predefined threshold, the surrounding cells are added regardless of the energy deposited in them. This algorithm is known as '4-2-0' topo-cluster reconstruction. [140].

Track reconstruction starts with hits in the ID layers. The pixel and SCT subdetectors assemble clusters from these hits to form three-dimensional space-points. Three space-points are combined to form track seeds. Additional clusters are added to form track candidates. Clusters can be shared between multiple track candidates. These ambiguities are resolved by performing fits to the track candidates to obtain a likelihood metric, where track candidates with a likelihood below a predefined threshold are (iteratively) discarded in the track-finding algorithm [141–144].

Topological clusters that share an overlap within $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ are combined into a single cluster, and otherwise they are treated as two separate clusters. When clusters are matched to ID tracks they are considered as electron candidates. After this cluster-to-track matching, the energies of electrons are recalibrated because the initial cluster calibration is performed before the final matching [145].

The electron identification algorithm is prone to the reconstruction of non-electron objects as (fake) electrons. For example, hadronic objects can be misidentified as electrons. Alternatively, non-prompt electrons from heavy-flavour decays can falsely be classified as prompt. To improve the purity of selected electrons in analysis, different quality criteria on the identification and isolation of candidate electrons can be imposed.

Identification algorithms use a likelihood discriminant, derived in a tag-and-probe method in $Z, J/\psi \to ee$ decays, and additional kinematic criteria based on $|\eta|$ and the four-momentum of electron candidates to distinguish electrons from fake electrons. Different requirements define three working points, optimised such that electrons in the chosen category are reconstructed with a predefined efficiency. The `Loose`, `Medium` and `Tight` working points correspond to an average reconstruction efficiency of respectively 93%, 88% and 80%, that increases with E_{T} [140]. The trade-off between keeping high efficiency but reducing the number of fakes ultimately determines the choice of the working point, and is a choice specific to any given analysis.

Isolation variables take into account activity near leptons from either nearby tracks or nearby deposits in the calorimeters, and can be translated to quality cuts to reduce contributions from non-prompt electron candidates. The general idea is as simple as effective: for both track and cluster variables the total amount of activity (i.e. tracks or energy depositions) around candidate electrons are summed up, from which the electron contribution is excluded. The implementation

of isolation criteria depends highly on the needs of the analysis. The `Loose` and `Tight` working points cut on both the calorimeter an track isolation variable, whereas HighPtCaloOnly cuts only on the calorimeter isolation variable [140].

Photons

The story of photon reconstruction is very similar to the one of electron reconstruction, with the main difference that electrically neutral photons do not leave tracks in the inner-detector. As photons traverse detector material, they can be converted into a pair of electrons before reaching EMCalo. A distinction is made between converted and unconverted photons. A converted photon is a cluster matched to a conversion vertex, where unconverted photons are clusters matched to neither an electron track nor a conversion vertex. Similarly to electrons, prompt and isolated photons are selected by applying identification and isolation criteria. These criteria are defined by a set of cuts on shower shape variables and calorimeter/track isolation variables respectively.

Muons

The reconstruction of muons in ATLAS depends mainly on track information from the ID and MS, but energy deposits in the calorimeter are factored in to account for any energy losses between the ID and MS, as well as tagging of muon candidates without using information from the MS.

Tracking muons in the ID follows the same prescriptions as for electrons. In the MS, hits in individual segments are combined into short, straight-line local track seeds. The information from multiple segments is then combined to form initial track candidates, including loose constraints of the magnetic bending of the muon trajectory and its relative orientation to the interaction point. Muon trajectories are then obtained in a global χ^2 fit, after which outlier hits are removed and hits along the trajectory originally not assigned to the track included [146].

A classification of muon candidates can be made depending on which reconstruction algorithm is used to combine information from the ID, calorimeters and the MS. "Combined muons" are formed from a combined track fit to ID and MS hits, including energy depositions in the calorimeters. Alternatively, ID tracks can be extrapolated to the MS and if this satisfies tight angular matching to at least one MS track we identify this as a "SegmentTagged muon" candidate [146]. Combined or SegmentTagged muon candidates are considered in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets in Chapter 5.

As is the case with electron and photons, `Loose`, `Medium` and `Tight` identification working points are defined to match different requirements on the efficiency of (prompt) muon identification, as well as fake-muon and non-prompt muon background rejections. The main non-prompt background for muons comes from heavy-flavour hadron decays, which can be mitigated using criteria based on isolation requirements in the ID, calorimeter or a combination of both using a particle flow algorithm [147].

Jets

As highly energetic particles produced in a proton-proton collision move away from the interaction point, they radiate energy in the form of soft quarks and gluons, ultimately leading to a collimated spray of particles when these quarks and gluons hadronise and form bound states.

These collimated sprays of particles are known as jets. It is important to note that jets are no fundamental representation of nature and are *only* defined through jet algorithms that humans have designed. This section summarises the main jet clustering algorithm that is used at the LHC and experimental reconstruction techniques in ATLAS during Run-2 data-taking.

The field of jet physics emerged from early electron-positron colliders in the 1970s and 1980s [148], and continued to be a subject of study in electron-proton colliders and hadron colliders in the following decades. Pioneering algorithms to cluster energy depositions in detectors into jets date back more than four decades [149], where current algorithms are broadly classified in two categories. Sequential algorithms identify nearby energy depositions, cluster them, and sequentially add neighbouring energy deposition until this reaches below some predefined threshold. This leads to jets with a fluctuating area in (ϕ, η)-space. Notable examples include the k_T [150], anti-k_T [151] and Cambridge/Aachen (C/A) [152] algorithms. Cone algorithms combine particles within a fixed, conical angular region in (ϕ, η)-space to find stable cones that point in the same direction as the momentum of its contents. An example is SISCone [153].

The selling point of cone algorithms is the regular shape of the resulting jets, advantageous in experiments because of the simplified treatment of jet energy calibrations and subtractions of noise from pile-up or underlying events. However, many cone algorithms are infrared unsafe, meaning that soft or collinear emissions from final state particles lead to divergences in the jet definition. Sequential clustering offers a more natural treatment of soft and collinear radiation and is discussed below.

Clustering algorithms first define two distance metrics,

$$d_{ij} \equiv \min(k_{\mathrm{T},i}^{2p}, k_{\mathrm{T},j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \quad ; \quad \Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i + \phi_j)^2 \tag{4.8}$$

$$d_{iB} \equiv k_{\mathrm{T},i}^{2p}, \tag{4.9}$$

where the first distance variable is between two entities i and j, and the second one between entity i and the beam axis (B). The tweakable parameter R is known as the jet radius parameter, where p determines the k_T ordering of the clustering algorithm. Second, the minimum of the entire set $\{d_{ij}, d_{iB}\}$ is found. If $d_{ij} < d_{iB}$, entities i and j are combined into a single entity (ij), after which removing i and j from the list of entities. Conversely, if $d_{ij} > d_{iB}$, entity i is called a jet. This procedure is repeated until all entities are combined into jets (inclusive clustering) or a fixed number of jets has been found (exclusive clustering) [154]. For simulated events, these "entities" are stable final-state particles. As we will discuss in more detail later, experimental entities constitute energy depositions in the calorimeters, charged-particle tracks or a combination of the two.

The k_T, Cambridge/Aachen and anti-k_T algorithms are defined by $p = 1, 0, -1$ respectively. By construction, the sequential algorithms are infrared safe. The anti-k_T algorithm is the main jet algorithm used at the LHC. The default tool that implements the anti-k_T algorithm, among many others is FastJet [155]. Opposite to the k_T algorithm, it *first* clusters high-k_T objects and gradually adds softer radiation to jet candidates. Generally, this leads to jet areas that are much more regular than when using $p = 1, 0$. The value of R effectively determines the size of the jet and a priori has no "correct" value. However, R must be large enough to include all the perturbative radiation emitted from a parton, and small enough such that it does not include too

much additional hadronic activity from other sources, such as pile-up, the underlying event or neighbouring jets. Typically $R = 0.4$ is used, but large-R jets enable jet substructure analyses to study the Standard Model and BSM signatures in innovative ways. Jet substructure analyses were initially proposed as a new Higgs search channel [156], but are now a common pathway to study a wide range of phenomena in collider physics [157].

Experimentally, detector hits are the starting point for jet reconstruction. The energy of jets is in the form of charged particles (mostly $\pi^{\pm}$ mesons), photons (from $\phi \to \gamma\gamma$ decays) and neutral hadrons (neutrons and kaons). In Run-1 and at the start of Run-2, ATLAS only used energy depositions in the calorimeter to reconstruct jets [158]. This was improved towards the end of Run-2, by leveraging information from ID tracks in addition to calorimeter hits. The methodology of combining information from the calorimeter and inner detector is known as the particle flow (PFlow) algorithm. Jets reconstructed with this algorithm are referred to as PFlow jets. By combining information from both systems, the PFlow algorithm improves not only the precision of the energy and angular resolution of jets, but also significantly reduces the contribution from pile-up interactions in the ID region. Furthermore it extends the acceptance of the detector to softer particles, as p_{T} requirements of tracks are lower than the noise thresholds to trigger topoclusters [147, 159]. Jet vertex tagging (JVT) criteria match jets with the primary vertex to reduce jets emerging from pile-up interactions [160, 161], which is extended to forward jets during Run-2 [162]. Flavour-tagging algorithms exploit properties of b- and c-hadrons, such as their long lifetime, high mass and decay multiplicity to identify the flavour-content of a jet [163].

Due to detector effects, the experimentally measured,"detector-level" energy of a jet can differ from the true "particle-level" energy. Jet energy scale (JES) calibrations restore the jet energy to the particle-level. This includes pile-up corrections, absolute jet energy scale calibrations based on dijet MC events, and a series of multiplicative corrections (known as global sequential calibration) to improve the jet resolution. Finally, *in situ* jet calibrations account for mismodelling effects in both simulated detector interactions and the underlying physics process. These corrections lead to JES uncertainties that are propagated into physics analyses. The jet response is measured in both data and MC, and the ratio is applied to data as an additional correction [159]. The jet energy resolution (JER) can be parametrised as a function of transverse momentum,

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{N}{p_{\mathrm{T}}} \oplus \frac{S}{\sqrt{p_{\mathrm{T}}}} \oplus C, \tag{4.10}$$

with a noise (N) term, stochastic (S) term and a constant (C) term. JER measurements are made in well-understood dijet events, where systematic uncertainties related to JER are propagated through analyses by smearing jets according to a Gaussian function dependent on the 1σ variations in the uncertainty component [163]. From the entire jet calibration chain, this paragraph highlights only the uncertainties related JES and JER because these are important inputs to the calculation of missing transverse momentum. Consequently they have a big impact in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets analysis in Chapter 5.

Jets initiating from the primary vertex with a *b*- or *c*-quarks are known as heavy-flavoured jets. This type of jets leaves a distinctive signature because the primary heavy-flavour hadron typically travels a few millimetres before it decays leptonically, semi-leptonically or hadronically. A large multiplicity of tracks will therefore point to a secondary vertex, located at a different

position from the primary vertex. The *impact parameter* is a distance metric that quantifies this effect for each track, and is defined as the closest point of approach between the primary vertex and the tracks (pointing to the secondary vertex) extrapolated backwards. Flavour-tagging algorithms exploit these characteristics to discriminate *b*- and *c*-jets from light-quark jets. Information from impact parameters, displayed secondary vertices and decay chains of heavy-flavour hadrons is combined and used as input for (multivariate) machine learning algorithms to provide a single *b*- or *c*-tagging discriminant. Flavour-tagging is a rapidly evolving field that goes hand in hand with developments in artificial intelligence. ATLAS used a combination of low-level and high-level taggers during Run-2 [163], while advancements in graph neural network approaches show even better performance in efficiency and light-quark rejection [164]. The performance of such algorithms is quantified by the efficiency of reconstructing heavy-flavoured jets and the rejection factor of jets originating from lighter quarks. Naturally there is a trade-off between signal efficiency and background rejection, and the optimal "working point" is based on the goal of the analysis in question.

4.2.2 Reconstruction of complex objects

The ability to tag candidates for stable particles provides us with the power to zoom in further on the underlying physics processes occurring in proton-proton collisions, in particular providing smoking-gun evidence to work out which heavy particles may have been present in events before promptly decaying. The main recipe here is adding four-momenta of final-state objects and working out kinematic and invariant quantities of the parent particles, but detector resolutions and decays into invisible particles complicate the procedure.

τ leptons

The τ lepton has a lifetime of $(290.3 \pm 0.5) \times 10^{-13}$ s [165] so that it decays before reaching the innermost layers of the ATLAS detector. Approximately two-thirds of τ leptons decay hadronically via $\tau^- \to W^-(\to \text{hadrons})\nu_\tau$. The remaining third decays weakly into electrons or muons with an equal probability via
$\tau^- \to W^-(\to \ell\bar{\nu}_\ell)\nu_\tau$. Leptonically decaying τ leptons are difficult to distinguish from events with electron or muons, because phenomenologically they only differ in the multiplicity of neutrinos. Hadronically decaying τ leptons contains mostly pions such that their signature resembles that of a narrow jet, with either one or three charged tracks depending on the pion multiplicity. These hadronic decays are probed in ATLAS by reconstructing the visible component and separating this from quark- and gluon-initiated jets using machine learning algorithms [166, 167].

Top quark

With a mass of approximately $173\,\text{GeV}$, the top quark is the heaviest fundamental particle in the SM. As a result it has a lifetime of $\mathcal{O}(10^{-25}\,\text{s})$, so short that it decays before getting a chance to form a bound hadron state. This is the only quark in the SM with this property, such that it behaves as if it were a "free" particle. The values of the CKM matrix in Equation 2.27 indicate that the coupling between top quarks and quarks from the first or third generations are highly suppressed. Therefore, the top quark decays mostly via $t \to Wb$. In hadron colliders, top quarks are mainly produced in pairs (also referred to as $t\bar{t}$ production), or in association with a b-quark

or W-boson (single-top production). [4] Events with top quarks are identified by tagging on the presence of its decay products: leptonic W decays (for reasons that we discuss below) and b-jets.

Missing transverse momentum

Missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$) is a measure of the momentum imbalance in the transverse plane. Inner detector tracks, energy deposits in the calorimeter and muons identified in the muon spectrometer are used to reconstruct $p_{\mathrm{T}}^{\mathrm{miss}}$. The reconstruction process involves using tracks to recover momentum from low-transverse-momentum (p_{T}) charged particles that may not be detected by the calorimeters. Additionally, muons reconstructed from the inner detector help recover muons in areas not covered by the muon spectrometer. Calorimeter energy deposits are associated with calibrated high-p_{T} parent objects in a specific order: electrons (e), photons (γ), hadronically decaying τ-leptons (τ), muons (μ), and jets. The $p_{\mathrm{T}}^{\mathrm{miss}}$ is computed by summing the negative momenta of the calibrated reconstructed objects, each projected onto the x and y directions:

$$p_{x(y)}^{\mathrm{miss}} = \underbrace{p_{x(y)}^{\mathrm{miss},e} + p_{x(y)}^{\mathrm{miss},\gamma} + p_{x(y)}^{\mathrm{miss},\tau} + p_{x(y)}^{\mathrm{miss},\mu} + p_{x(y)}^{\mathrm{miss,jets}}}_{\text{hard term}} + p_{x(y)}^{\mathrm{miss,soft}}. \tag{4.11}$$

The soft term, $p_{x(y)}^{\mathrm{miss,soft}}$, is derived from tracks associated with the primary vertex but not with any high-p_{T} objects.

W and Z bosons

W-bosons decay hadronically or leptonically ($W^- \to \ell^- \bar{\nu}_\ell$) with relative branching fractions of 67% and 33% respectively [165]. Due to a higher detector resolution for leptons than for jets, events with W-bosons are predominantly measured via the leptonic decay. This poses a challenge because a significant fraction of the energy leaves the detector in the form of a neutrino. The $p_{\mathrm{T}}^{\mathrm{miss}}$ in an event can be used as a proxy for the neutrino's transverse momentum. Therefore, a combination of lepton requirements and a $p_{\mathrm{T}}^{\mathrm{miss}}$ cut can enrich a sample in leptonically decaying W events. The longitudinal component of $p_{\mathrm{T}}^{\mathrm{miss}}$ (in the z-direction) is unknown so that the invariant mass of the W decay products can not be determined exactly. However, the transverse mass

$$m_{\mathrm{T}} = \sqrt{2 p_{\mathrm{T}}^{\ell} p_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos\Delta\phi)} \tag{4.12}$$

is a quantity that uses all other measured information of potential leptonic W decays, except from longitudinal directions. In this equation, p_{T}^{ℓ} is the transverse momentum of the charged lepton and $\Delta\phi$ this azimuthal angular difference between the direction of p_{T}^{ℓ} and $p_{\mathrm{T}}^{\mathrm{miss}}$. The m_{T} distribution does not peak at m_W but at a lower value instead, because it only uses transverse information of the decay and loses energy transmitted in the longitudinal direction. Nonetheless, cuts on the transverse mass in an event can be implemented to enhance the purity of $W \to \ell\nu$ decays. Distributions of m_{T} and p_{T}^{ℓ} are also used in fits for precise determinations of m_W, for example in the 2023 ATLAS result [169].

Z-bosons couple to both electrically charged and neutral fermions. Decay rates into pairs of quarks (excluding the top-quark), charged leptons and neutrinos are 69.9%, 10.1% and 20.0% respectively. Hadronic decays can be reconstructed by measuring its decay products as jets, but at

[4] Recently, ATLAS also discovered the rare production of four top quarks [168].

colliders this signature is difficult to distinguish from QCD multijet backgrounds. The decay into charged leptons is more distinctive, with a fully visible final-state can be reconstructed precisely due to good lepton resolutions. Constraints on the invariant mass of the dilepton system ($m_{\ell\ell}$) can improve the signal efficiency of events with leptonic Z decays. As we will see in Chapter 5, the associated production of a Z-boson with jets, where the Z decays into a pair of neutrinos, is the major SM contribution to signals of $p_{\mathrm{T}}^{\mathrm{miss}}$+jets.

Higgs boson

The coupling of the Higgs boson (H) to other particles in the SM is proportional to their respective mass. The top quark is too heavy for a Higgs boson to decay into ($m_H < 2m_t$). Therefore the predominant fermionic decay mode is into a pair of b-quarks with a branching fraction of more than 50%. Although $H \to b\bar{b}$ is the predominant decay mode that is now actively studied at the LHC, it is difficult to distinguish it from an overwhelming QCD background. For that reason, the main focus in the early LHC days was on its decay into electroweak bosons that have higher signal-to-background ratios with better experimental precision. It was the mainly the combination of the $H \to ZZ^{(*)} \to 4\ell$ and $H \to \gamma\gamma$ channels – both with sub-percent branching fractions – that ultimately led to the discovery of the Higgs particle in 2012 [170, 171]. Because the Higgs boson does not couple directly to massless particles, the decay into a pair of photons occurs via loops diagrams. The $\gamma\gamma$ (via loop-induced processes) and 4ℓ final-states are reconstructed using techniques illustrated earlier in this chapter.

4.2.3 Event selections

With a better grasp of reconstructing complex objects and particles that decay in the detector, we now enter the realm of physics analysis, where regions of phase-space are defined to zoom in on specific processes. We discuss conventional analysis techniques and address technical hurdles to ensure that any conclusions drawn from this are robust to the "messy" environment in a collider.

Phase-space selection

Phase-space definitions for analysis are a combination of physics goals ("what we want") and technical constraints ("what the detector allows us to measure"). In practice, it is an iterative procedure to optimise the selections, cuts and general analysis strategy.

Selections can be categorised in two types. First, analyses can filter on the particle content in an event. For example, a measurement of the $W \to e\nu$ process is enhanced in signal purity by requiring events to have one prompt electron. Particle content can instead also be used as a *veto*, or in other words not selecting an event if it contains a specific particle. Second, selections can be imposed on the kinematic properties in the event. An example is a cut the minimal electron p_{T}, as the electron trigger may only reach full efficiency above a given value of p_{T}. These cuts can also be modified by physical constraints of the detector. For example, the use of particle flow jets is restricted to the region with ID coverage $|\eta| < 2.5$. In some analyses kinematic properties are used as inputs into machine learning algorithms, such as neural networks or boosted decision trees, trained on simulated MC to optimise discrimination between signal and backgrounds. The analysis strategy and event selections for the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets measurements are discussed at great length in Section 5.1 and 5.2.

Good run lists

Only events in the Good Run List (GRL) of ATLAS data are considered. GRLs keep track of various conditions during data-taking and select good luminosity-blocks, roughly the equivalent of 1-2 minutes of data-taking, corresponding to periods with no detector problems and all subdetector systems performing in good condition. Events are also discarded if there were problems in the LAr system, SCT, Tile calorimeter or partial information is missing in any other way. The latter can for example occur when software corresponding to a particular subsystem is restarted leading to incomplete events.

Overlap removal

Various objects such as jets and leptons are constructed independently from the same detector information. This can lead to ambiguities and a double counting of physics objects. Overlap removal (OR) is a sequential procedure that aims to eliminate this type of ambiguities. Two objects are compared in a pre-defined order and if there is geometrical overlap one of the two objects is removed. The geometrical overlap includes close matching between objects in ΔR or association to the same ID track. The overlap procedure used in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets analysis in Chapter 5 is summarised in Table 4.1.

Reject	Against	Overlap Criteria
τ	Electron	$\Delta R < 0.2$
τ	Muon	$\Delta R < 0.2$
Electron	Muon	shared ID track
Jet	Electron	$\Delta R < 0.2$
Electron	Jet	$\Delta R < 0.4$
Jet	Muon	Number of tracks < 3 and $\Delta R < 0.2$
Muon	Jet	$\Delta R < 0.4$
Jet	τ	$\Delta R < 0.2$

Table 4.1: An ordered sequence of overlap removal criteria and decisions. The first column shows which object is rejected against the object in the second column. Conditions in the third column have to be fulfilled in order for the object to be removed. The sequence goes from the top to the bottom, where the objects removed in the previous stage are not considered in the following decisions. ΔR is calculated in the $\phi - y$ phase space.

Jet cleaning

Jet cleaning comprises a set of jet cuts to suppress fake jet backgrounds, mainly from calorimeter noise, cosmic-ray showers and beam-induced backgrounds due to proton losses upstream of the interaction point. The goal is to reduce these background while maintaining a high selection efficiency of collision-induced jets. Three discriminating features in jet reconstructions are used to achieve this. First, the pulse shapes in LAr calorimeter cells for fake jets look different than those from real jets. The measured pulse shapes are compared to expected pulse shapes from simulation, to ultimately obtain signal shape quality variables. Second, energy ratio variables exploit the expected direction of the shower development in the detector of jets from proton-proton collisions. This mainly allows a suppression from beam-induced backgrounds or calorimeter noise,

that are expected to be more longitudinally localised. For example, the fraction of the jet energy deposited in the EM calorimeter (f_{EM}) is peaked close to zero or one for beam-induced jets that respectively deposit none or most of their energy in the EM calorimeter, while this is a continuous spectrum for real jets. Third, track based variables exploit the fact that most real jets in $|\eta| < 2.5$ contain hadrons that leave tracks in the ID. The fraction of jet energy that was reconstructed in the ID can also be used to separate this from fake jets that tend to leave no tracks in the ID. The `Tight` and `Loose` jet cleaning criteria have a reconstruction efficiency of 99.0% and 99.5% for $p_{\mathrm{T}} > 100\,\mathrm{GeV}$ respectively [172], while reducing fake jets by a factor of 10000 and removing all fake jets with $p_{\mathrm{T}} > 1\,\mathrm{TeV}$ [173].

4.3 Confronting theory with experimental data

So far we have discussed how to generate fully simulated proton-proton collision events from a fundamental theory and how to reconstruct common objects from detector signals. This means that we are really closing in on making an apples-to-apples comparison between theory and experiment. Two representation levels of the data are best suited for this: "particle-level" and "detector-level". This section summarises the pros and cons for both sides, and outlines the complementarity of precision Standard Model measurement and (BSM-driven) searches at the LHC.

4.3.1 Particle and detector level

The representation of collision events with a full set of stable final-state particles is called the particle-level. From here onwards this is synonymously referred to as the "truth-level" representation. In MC simulation, the particle-level is conventionally represented as a tree-like event record after the steps of hadronisation, heavy-flavour decays and QED FSR. However, as we discuss in more details in Section 4.4, experimental data can be corrected back to the particle-level via MC based corrections in a process called unfolding. Detector effects include geometric acceptance, finite resolutions and inefficiencies. The representation where detector effects have had an impact on data is the detector-level. This is also referred to as reconstruction-level. The detector-level is directly accessible in the experimental dataset after object and event reconstructions defined in Section 4.2. Detector simulation is needed to map MC particle-level predictions to the detector-level. The notion of particle-level and detector-level, and how to convert between them, is summarised in Figure 4.2.

There are numerous reasons why one might want to re-use experimental data from a published analysis later on. Consider a BSM search as an example. The measured phase-space may be subject to SM backgrounds. If the SM background modelling is improved, the search results can be re-evaluated using this updated SM prediction. Or, for instance, one might want to evaluate the compatibility of the data with a set of BSM parameters different from those used in the original analysis. Experimental results from detector-level analyses, once published, are difficult to re-use. The main reason is that in addition to event generation an additional step of MC detector simulation is required. Detector simulation is computationally expensive and the necessary GEANT4 configuration to describe ATLAS (or any other detector) is not readily available to people outside the collaboration. Furthermore, detector-level results are at the risk of becoming obsolete

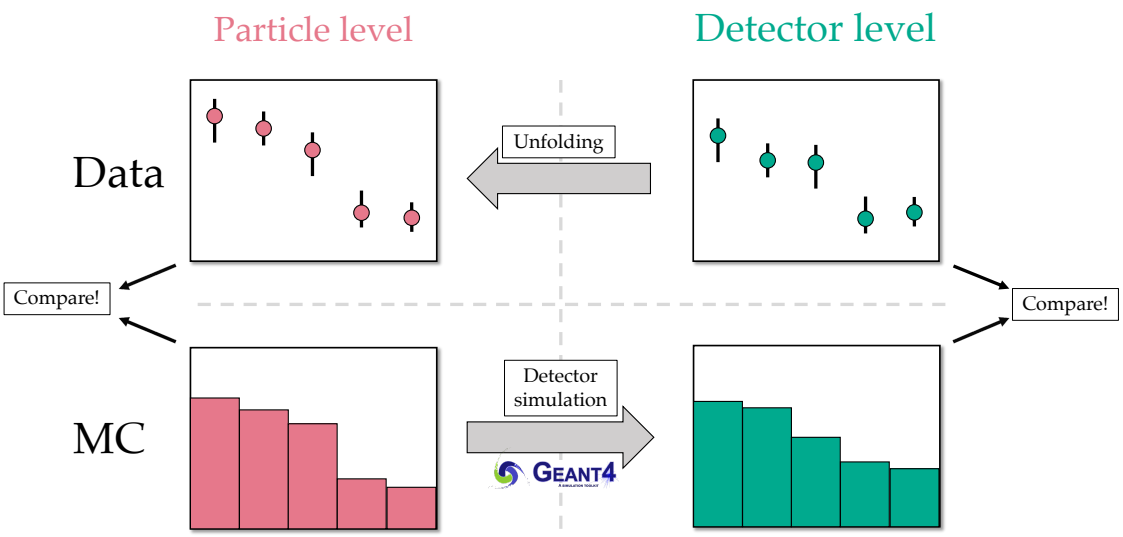


Figure 4.2: Predictions of the underlying theory, i.e. the event record of final-state particles and their associated momenta, are made with MC event generators. These are particle-level predictions indicated in the bottom left histogram. Because data is only accessible through a detector, physicists initially measure detector-level data, as indicated in the top right diagram. The truth MC prediction cannot be validated to detector-level data, because the latter accounts for detector effects while the former does not. There are two approaches to make a fair comparison. First, the MC truth predictions can be put through the computationally expensive detector simulation to make a comparison at detector level. The second option is to correct the data for detector effects (i.e. *unfold*) and validate truth MC predictions to the unfolded data at the particle level.

when the experiment is decommissioned and its detector simulation no longer maintained. In summary, the widespread usability and longevity of results is hindered if they are presented at detector-level.

So this is where unfolding comes in. Particle-level data is easily comparable to new theory models and directly usable by the entire community, and it provides shelf-stable results than can be re-used far into the future. As demonstrated in Chapter 5, it is possible for particle-level analyses to obtain signal sensitivities comparable with those from their detector-level counterparts. The price to pay is that unfolding requires careful robustness checks and physics cross-checks, and additionally encodes an assumption on the response of data events in the measurement. Whether results are best presented at the particle-level or at the detector level depends on the goal of the analysis, as discussed below.

Measurements and searches

Measurements are driven by known or expected SM processes and – while this is not a strict definition – are most often presented at the particle-level. What matters most is getting the best possible estimate of nature without detector dependence, or in other words what would be measured with a perfect detector in infinite resolution and perfect reconstruction efficiency. Results are often presented as fiducial cross-sections, by dividing the recorded number of events by the detector in a given bin by the detector-dependent integrated luminosity, $\sigma_{\text{fid.}} \sim N_{\text{events}}/\mathcal{L}_{\text{int.}}$. This enables a direct comparison between fiducial cross-sections collected under different beam conditions and with variable integrated luminosities. These analyses have a high impact through their reusability and long shelf life. Particle-level measurements have a wide range of applications; from PDF fits to MC event generator tuning and the validation of theory predictions. These measurements can also be sensitive to BSM physics, either directly with BSM final states (e.g. $p_{\text{T}}^{\text{miss}}$), or through loop effects and EFTs.

Searches are detector-level analyses optimised to discover BSM physics or rare SM processes. They provide relatively quick discovery potential, but are not easy to reuse because any new predictions require costly detector simulation or approximations thereof. This also poses limitations to statistical combinations between results that are obtained under different running conditions or between different experiments. The main goal of searches is to maximise the significance of the analysis to a specific class of physics models that is being tested against the SM-only hypothesis. So far all BSM searches have found null results that are then typically converted into limits of the parameter space for the model that was investigated. Results can be presented as (differential) event counts and obtained limits. Despite the myriad of searches carried out by the experiments, only a subset of all BSM scenarios and parameter configurations can be tested this way directly. Furthermore, more often than not the search models are simplified versions that do not capture the full phenomenology of the full model.

Measurements and searches both have pros and cons. Unfolding introduces a level of model-dependence on the experimental data that needs to be accounted for with uncertainties. However, the same is true in detector-level analyses where (SM) backgrounds in the signal region are typically constrained via control regions. This model-dependence of unfolding poses the risk of smearing any effects, such as contributions from new physics, that are present in experimental data but not in simulation. But again, the same is true for corrections applied in detector-level

searches. Unfolding imposes restrictions on the minimum number of events in each bin for the procedure to be well-behaved statistically. This is normally fine if you are doing a measurement: the data is your signal, and the goal is to unfold and measure the data as it is. In a search this potentially poses a problem. If the goal is to discover a specific signal model, it may be necessary to go to extreme regions of phase-space to suppress backgrounds and isolate pure signal event counts.

In summary, particle-level and detector-level analyses serve different purposes and are in many ways complementary. Re-interpretability of measurements and the limited scope of BSM searches are compelling arguments that in order to get as much information from hadronic collisions at the LHC, the focus should shift more towards making re-interpretable measurements. There are community efforts, see e.g. Ref. [174] to argue this point.

4.3.2 Final-state measurements

A conventional approach is to define a phase-space to measure an exclusive hard-scatter process $pp \to X$, with selections to enhance purity of X (or its decay products). Irreducible background processes $pp \to Y$ that lead to the same final-state particles are then subtracted at the detector-level before unfolding. The drawback of this method is that precious LHC data is contaminated with model assumptions of the irreducible backgrounds: if the $pp \to Y$ cross-section is mismodelled, this mismodelling is directly transferred onto the experimental data during subtraction.

If improved $pp \to Y$ predictions come available, the "contaminated" dataset has become completely obsolete. Because the experimental data has a dependence on the mismodelled $pp \to Y$ predictions, it is not possible to investigate the impact of the updated $pp \to Y$ modelling. This leads to situations where entire analyses would have to be redone each time better estimates for the subtracted backgrounds become available in the future. A solution is to define analyses in terms of their final-state particles and remain agnostic about the underlying partonic processes that contribute to this final-state. This removes a degree of model-dependence of experimental data, making measurements more re-interpretable for future use.

An example of such a measurement is the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets analysis in Chapter 5, which in the signal region accepts any process that contributes to final-states with large $p_{\mathrm{T}}^{\mathrm{miss}}$, regardless of the underlying process. The dominant SM process that contributes is $Z \to \nu\nu$+jets, but additional contributions that are kept as signal include $W \to e\nu$+jets or processes involving top quarks. The $p_{\mathrm{T}}^{\mathrm{miss}}$+jets analysis and this thesis, particularly Section 5.7.3, prove that measurements of the final-state i) are compatible with final-process measurements after unfolding ii) can be made without compromising on the precision of the measurement and iii) demonstrate high sensitivity to BSM signatures. The additional level of re-interpretability makes final-state measurements indispensable assets in the precision era of the LHC, getting as much information from current proton-proton collisions and preserving that information for a lifetime. The virtue of final-state definitions is also underpinned by the Reinterpretation Forum for LHC physics in Ref. [174].

4.4 Correcting for detector effects: unfolding

Unfolding is a technique to correct detector-level data for detector effects and determine a probabilistic map to the particle-level. The crux of the problem is that unfolding operates on data that

is not deterministic: interactions with the detector are instead of probabilistic nature and there is an ambiguity where different particle-level distributions could lead to the exact same detector-level results. In this thesis we focus on practical implementations of unfolding at the LHC, but for Refs. [175–178] provide formal discussions on unfolding as an inverse problem.

Different statistical methods exist to address this, e.g. iterative Bayesian unfolding (IBU) [179] or singular value decomposition (SVD) [180]. The common denominator is that all unfolding methods rely on particle-level and detector-level knowledge of MC simulations, and in particular the link between them, to estimate and invert the detector response. Technically these are implemented in computational libraries, for instance RooUnfold [181]. Section 4.4.1 gives an intuitive introduction to the concept of unfolding, whereas Section 4.4.2 provides details on the iterative Bayesian unfolding algorithm that is used as the unfolding method in Chapter 5. Practical tests of the unfolding procedure are summarised in Section 4.4.3

4.4.1 General concept

Unfolding procedures rely on having both MC truth and detector-level predictions. As the simplest possible case of unfolding, consider the measurement of a total fiducial cross-section:

$$\sigma_{\text{fid.}}^{\text{detector-level}} = \frac{N_{\text{S}} - N_{\text{B}}}{\mathcal{L}_{\text{int.}}} \,, \tag{4.13}$$

where N_{S} and N_{B} are the number of signal and background events respectively, and $\mathcal{L}_{\text{int.}}$ is the integrated luminosity. This can be corrected for the detector response as

$$\sigma_{\text{fid.}}^{\text{particle-level}} = \frac{\sigma_{\text{fid.}}^{\text{detector-level}}}{C} \,. \tag{4.14}$$

Here, C is defined as the correction factor

$$C = \frac{N_{\text{MC,Reco}}^{\text{fid.}}}{N_{\text{MC,Truth}}^{\text{fid.}}} \,, \tag{4.15}$$

where the numerator and denominator give the number of selected MC events at reconstruction and truth level respectively. This correction factor can differ from unity when events are reconstructed but do not pass the selection cuts at truth level, or vice versa. This could happen, for example, when the four-momentum of a particle is mismeasured in the MC detector simulation and consequently an event fails to meet an event selection cut. The C-factor in (4.14) corrects for such effects.

Differential cross-section measurements are more complicated than the unfolding of a total cross-section. In principle, a distribution can be unfolded by calculating a C-factor for every bin. However, a layer of complexion is added by the possibility of bin-migrations. Bin-migrations refer to events that are reconstructed in a bin different from the particle-level bin. This can occur, for example, due to finite energy resolutions or detector miscalibrations. When doing a bin-by-bin unfolding using C-factors, migrations are not accounted for.

4.4.2 Iterative Bayesian unfolding

Detector effects and bin migrations are captured by the *response matrix*, that gives the probability of an event generated in particle-level bin j being reconstructed in detector-level bin i. The response matrix is purely derived from MC simulations by tracking particle-level events through detector smearing. This matrix is then constructed for every observable that is unfolded and measured in an analysis. The purity is defined as the diagonal elements of the matrix, which is the fraction of events that are reconstructed in the same detector-level bin. The relation between the number of MC events in particle-level (p) and detector-level (d) bins can be expressed mathematically as

$$\mathrm{d}_i^{\mathrm{MC}} = \mathcal{R}_{ij}\left(\epsilon_{\mathrm{F}}^j \mathrm{p}_j^{\mathrm{MC}}\right) + \mathrm{Fakes}_i \,, \tag{4.16}$$

where $\mathcal{R}_{ij}$ denote the elements of the response matrix and ϵ_{F} the fiducial efficiency, i.e. the events that pass particle-level selections but fail detector-level selections. In principle, relation (4.16) can be inverted to express detector-level data as the unfolded number of events in a particle level bin (u):

$$\mathrm{u}_i^{\mathrm{data}} = \frac{1}{\epsilon_{\mathrm{F}}^i}\mathcal{R}_{ij}^{-1}\left(\mathrm{d}_j^{\mathrm{data}} - \mathrm{Fakes}_j\right) \equiv \frac{1}{\epsilon_{\mathrm{F}}}\mathcal{R}_{ij}^{-1}(\mathrm{d}_j) \,. \tag{4.17}$$

For experimental data a term with "fakes" is included, that refers to instrumental noise at detector-level that pass the event selections as reducible backgrounds. This contributes to the total detector-level yields, but an estimate of this background can be made and subtracted before unfolding. However, simple matrix inversion does not work in practice because large oscillations can occur between neighbouring bins if the off-diagonal elements become non-negligible [176].

The matrix-inversion problem is a largely solved problem through the use of regularisation methods. One popular example, and the method used later on in Chapter 5, is the iterative Bayesian unfolding (IBU) technique [179]. To understand IBU, it helps to interpret the response matrix elements $\mathcal{R}_{ij}$ as conditional probabilities:

$$\mathcal{R}_{ij} = P\left(\mathrm{d}_i \mid \mathrm{p}_j\right) \,. \tag{4.18}$$

However, the aim of unfolding is to answer the *opposite* question: what is the probability that an event originates from particle-level bin j, given that it is reconstructed in detector-level bin i? The answer is found by using Bayes' theorem:

$$P\left(\mathrm{p}_j \mid \mathrm{d}_i\right) = \frac{P\left(\mathrm{d}_i \mid \mathrm{p}_j\right) \cdot P(\mathrm{p}_j)}{\mathrm{d}_i} \tag{4.19}$$

$$= \frac{R_{ij} \cdot P(\mathrm{p}_j)}{\sum_j P\left(\mathrm{d}_i \mid \mathrm{p}_j\right) \cdot P(\mathrm{p}_j)} \equiv \theta_{ji} \,, \tag{4.20}$$

where θ_{ij} can be though of as the unfolding matrix and $P(\mathrm{p}_j)$ is a prior truth distribution.[5] It is

[5] The use of Bayes' theorem in this step explains the nomenclature of the iterative *Bayesian* technique. However, note this is technically not a Bayesian method and the posterior should not be interpreted as a Bayesian statistic.

used to construct determine particle-level yields from the detector-level yields as

$$\mathrm{u}_i = \sum_j \frac{1}{\epsilon_{\mathrm{F}}^j} \theta_{ij} \mathrm{d}_j \ . \tag{4.21}$$

At first sight, the need for a prior truth distribution to obtain the unfolded result may seem worrying. However, the regularisation bias that is introduced with the truth prior is reduced with an iterative procedure: the posterior is used as the prior in a next unfolding iteration. Increasing the number of iterations k reduces the bias. However, statistical fluctuations are amplified for every iteration and hence there is an optimal point for the number of iterations. This bias can be quantified as follows. By fluctuating detector level MC events according to Poisson statistics an alternative *pseudo-dataset* is created. These fluctuated detector-level histograms can be unfolded and compared with the truth MC, which should be in agreement within the statistical uncertainties. Any non-negligible differences between the two can be defined as the regularisation bias systematic. If the MC models the data reasonably well at truth level the unfolding converges typically after two or three iterations. Systematic uncertainties on experimental data are propagated point-wise through the unfolding, by re-running the unfolding process with alternative versions of the response matrix corresponding to the $\pm 1\sigma$ variations of the associated nuisance parameters.

IBU is a popular method in high energy physics because it solves the problems arise with simple matrix inversion. The technique is not only widely used in ATLAS, see for example Refs. [182–184], but also in other experiments such as ALICE [185] or Super-Kamiokande [186].

4.4.3 Unfolding optimisation and tests

The binning of observables can be optimised to achieve higher purity and a minimum number of events in each bin. This reduces the statistical uncertainties, in addition to dependence of bin-migrations particularly in steeply falling distributions. In IBU the statistical uncertainty from the unfolding increases with the number of unfolding iterations. However, a sufficient number of iterations is necessary to reduce the regularisation bias. A good MC modelling of the detector-level data can help the unfolding to converge quicker and hence reduce the final statistical uncertainty on the unfolded measurement.

Various tests can be made to test the robustness of the unfolding to the various inputs and to determine associated uncertainties to the procedure. A closure test is a simple test in which detector-level MC is treated as pseudodata and "unfolded' back to the particle-level with the associated response matrices. By definition, the resulting particle-level distributions should match *exactly* the original particle-level MC prediction. The regularisation bias assesses the dependence of the unfolding on mismodelling effects, by reweighting particle-level and detector-level MC to match the data, unfolding it with the nominal response matrix and comparing the unfolded pseudodata to particle-level MC. Any residual differences are applied as an uncertainty on the unfolded data, which is labelled the *basic unfolding bias* or *basic unfolding uncertainty*. This uncertainty is restricted to reweighting for observables that are part of the main measurement in the analysis. In addition, one can derive similar uncertainty estimates for the modelling "hidden variables". These refer to observables that are not presented in the analysis as the main results, but can have an impact on the unfolded distributions indirectly. For instance, the measurement

of a $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution is correlated to the hidden variable jet p_{T}.

In signal-injection tests a BSM signal is added to the pseudodata to assess if the SM-based response matrix does not hide away any new physics that may be present in the data. Finally, different SM processes interact differently with the detector. To quantify the effect that the composition of multiple SM processes contributing to the same final state may have on the unfolded results, process composition tests modify the relative SM contributions in the MC that is used for unfolding and compare the nominal unfolded distributions to those corresponding to modified compositions.

Chapter 5

Missing transverse energy in association with jets

All the techniques outlined in the previous chapters ultimately serve two main goals: making precision measurements of the SM and to search for new physics beyond the SM. The $p_{\mathrm{T}}^{\mathrm{miss}}$+jets analysis presented in this chapter does both. Measurements of various observables are made in a phase space that is populated by well-known SM processes, but where sensitivity to BSM physics is expected. It is the first *inclusive* measurement of $p_{\mathrm{T}}^{\mathrm{miss}}$ with the full ATLAS Run-2 dataset. Results are corrected for detector effects and by comparing these to both SM and BSM predictions a quantitative interpretation is made. It extends and supersedes a previous ATLAS measurement of $p_{\mathrm{T}}^{\mathrm{miss}}$ made with a $3.2\,\mathrm{fb}^{-1}$ dataset [187] not only with better statistics, but also with more robust object definitions (e.g. using PFlow jets), additional measurements regions and data-driven techniques to estimate reducible and irreducible backgrounds. From the outset, this analysis has been designed with re-interpretability in mind. The selected fiducial phase-space is optimally re-interpretable, and the analysis logic and detailed numerical results are preserved on Rivet and HEPData respectively.

The analysis strategy is outlined in Section 5.1, while specific details of the measured phase space and observables are found in 5.2. Information on the used datasets, theoretical predictions and instrumental backgrounds is summarised in Section 5.3. Thereafter Section 5.4 gives an overview of the systematics included in the analysis, with detector-level comparison for data and MC shown in Section 5.5. After unfolding the detector-level data as detailed in Section 5.6, particle-level results are obtained that are presented in Section 5.7. The chapter concludes with the interpretation of the measured data in terms of both the SM and BSM scenarios in Section 5.8.

5.1 Analysis strategy

The measurement is made with events containing missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$) in association with energetic jets. This is the main **signal region** (SR). As discussed in Section 5.2.1, $p_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the negative sum of the transverse momenta of all particles detected by ATLAS. As the incoming protons have no transverse momentum, conservation of momentum dictates

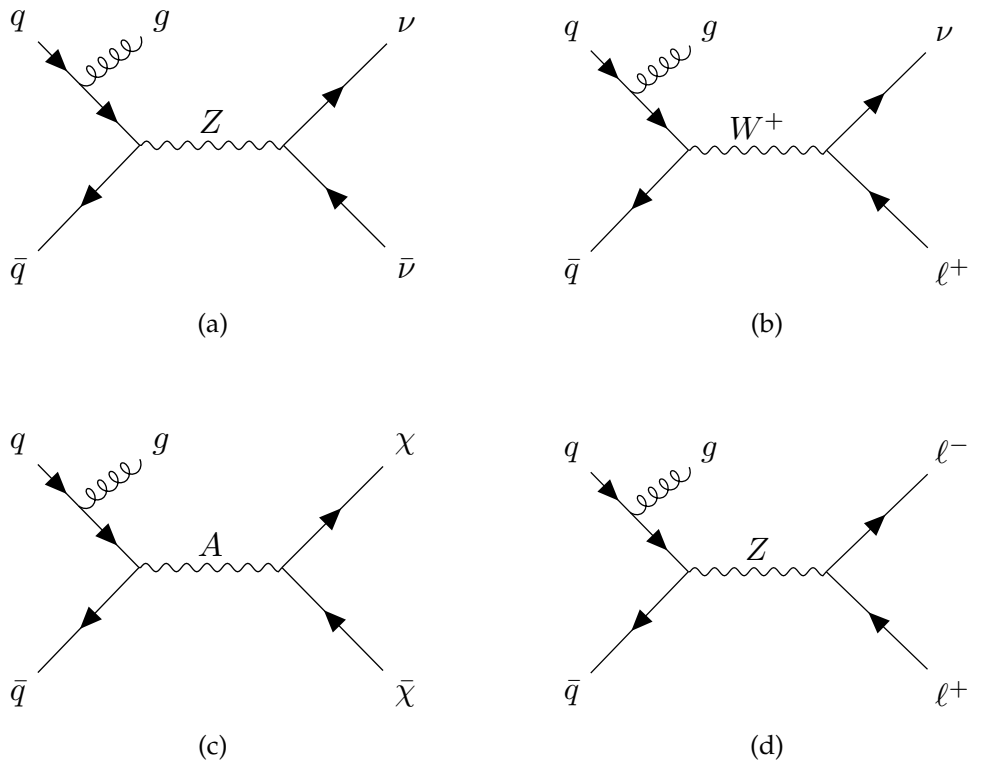


Figure 5.1: Feynman diagrams for the inclusive monojet (a) $Z \to \nu\nu$, (b) $W \to \ell\nu$, (c) $A \to \chi\chi$ and (d) $Z \to \ell\ell$ processes.

that a momentum imbalance in the transverse plan (and therefore a non-zero $p_{\mathrm{T}}^{\mathrm{miss}}$) can only occur when some particles produced in the collision leave the detector unseen. The $Z \to \nu\nu$+jets process, as illustrated in Figure 5.1a, is a SM process leading to a final state with $p_{\mathrm{T}}^{\mathrm{miss}}$ because neutrinos are not detected. The event display of a *monojet* event, with one energetic jet "balanced" by a missing transverse energy component, recorded by the ATLAS detector is illustrated in Figure 5.2.

However, the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets fiducial phase space has more SM contributions than $Z \to \nu\nu$, for example from W boson processes where the lepton is outside the detector acceptance as seen in Figure 5.1b. Furthermore, the finite resolution of the detector and mismeasurement of the "true" energy of the particles can also lead to *fake* $p_{\mathrm{T}}^{\mathrm{miss}}$. In terms of BSM physics, scenarios with dark matter (DM) particles that leave the ATLAS detector unseen can lead to a final state with missing energy. The Feynman diagram in Figure 5.1c shows an example for a mediator A that couples directly to both SM quarks and DM particles χ.

While a precision $p_{\mathrm{T}}^{\mathrm{miss}}$ measurement is the primary goal of this analysis, notice how the main contributions of $p_{\mathrm{T}}^{\mathrm{miss}}$+jets with Z and W bosons have counterparts in final states with *visible* leptons instead. For example, $Z \to \nu\nu$ is kinematically very similar to the $Z \to \ell\ell$ process shown in Figure 5.1d. For $W \to \ell\nu$, the experimental reconstruction of ℓ by the detector dictates whether we see $p_{\mathrm{T}}^{\mathrm{miss}}$+jets or 1ℓ+jets, but otherwise the process is the same. This motivates four **auxiliary measurements** (AMs) of regions containing charged leptons, targeting the main background processes in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets SR. These are 2μ+jets, $2e$+jets, 1μ+jets and $1e$+jets, with major contributions from $Z \to \mu\mu$, $Z \to ee$, $W \to \mu\nu$ and $W \to e\nu$ respectively. In addition to the

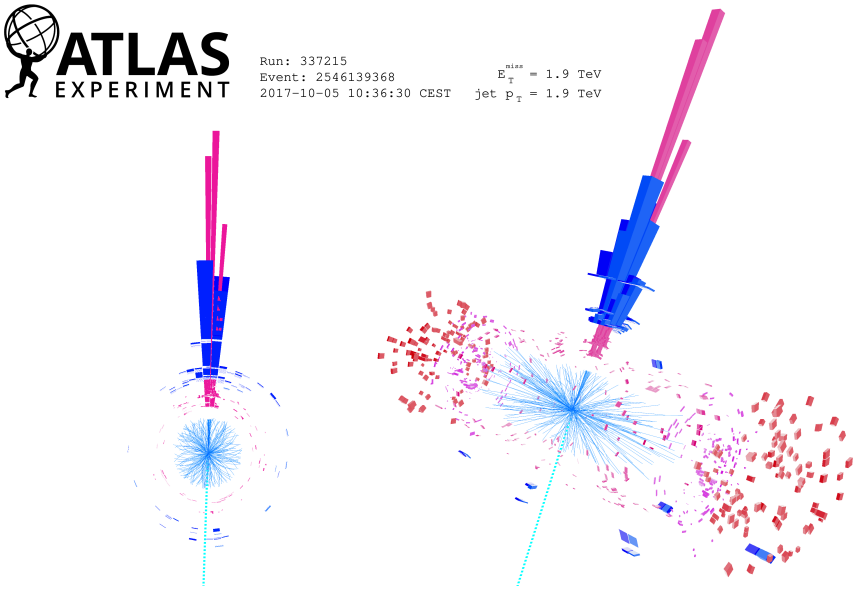


Figure 5.2: Event display of a monojet event recorded by ATLAS during Run-2. Blue and pink energy towers are reconstructed as a highly energetic jet with $p_{\mathrm{T}}^{\mathrm{jet}} = 1.9\,\mathrm{TeV}$ and a missing transverse momentum vector is indicated as a cyan dashed line. Figure taken from Ref. [188].

production of jets in association with massive W or Z bosons, we define a γ+jets AM to measure prompt, inclusive photon production. This latter process is not quite as similar to the processes in the SR, but has better statistics as the production cross-section is larger relative to the other AMs and provides a high purity with very little background contributions.

The AMs are used to constrain both theoretical and experimental systematic uncertainties. This is exploited in two different ways. First, in a similar spirit to the $3.2\,\mathrm{fb}^{-1}$ ATLAS measurement of $p_{\mathrm{T}}^{\mathrm{miss}}$+jets, an explicit ratio R^{miss} is calculated between the fiducial cross sections of each auxiliary region (AR) and the signal region with missing transverse momentum [187]:

$$R^{\mathrm{miss}} = \frac{\sigma_{\mathrm{fid.}}\left(\text{Signal Region}\right)}{\sigma_{\mathrm{fid.}}\left(\text{Auxiliary Region}\right)} \tag{5.1}$$

The cancellation of uncertainties by calculating these ratios is seen explicitly in Section 5.4. The second way, and this is where the current analysis supersedes the previous ATLAS measurement, is by providing differential distributions of the SR and AMs not just as a ratio but also individually. Systematic and statistical uncertainties are correlated between the different regions and this is exploited in a combined fit, for example by profiling over the relevant nuisance parameters, accounting for uncertainties in the covariance matrix or a combination of both.

The individual measurements significantly enhance the re-interpretability of this work. In terms of MC modelling SM processes, there is the additional freedom to compare predictions *only* to a single region. This has immediately proven to be useful. In Section 5.7 the unfolded cross-section of the lepton regions are compared to theoretical predictions from the High Energy Jets (HEJ) framework, which is not yet available in the $Z \to \nu\nu$ channel. In terms of limit setting for BSM scenarios, using R^{miss} works only if BSM contributes in either the numerator or denominator. In models where new physics signals appear in both the SR and the AMs, there may be

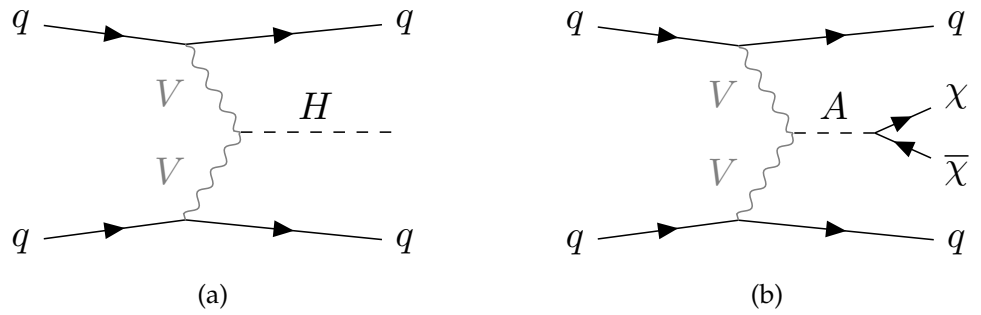


Figure 5.3: Feynman diagrams for VBF topologies in (a) Higgs production and (b) production of a DM mediator A.

unintelligible cancellation effects in the ratio. A separate measurement disentangles this issue.

But there are more ways in which this analysis improves the preceding $3.2\,\mathrm{fb}^{-1}$ measurement. As the first full Run-2 ATLAS measurement of $p_{\mathrm{T}}^{\mathrm{miss}}$+jets, corresponding to a total integrated luminosity of $140\,\mathrm{fb}^{-1}$, the statistical uncertainty will be greatly reduced. Furthermore it introduces the single lepton and photon channels.

Two different jet topologies are measured for every measurement region. The first one is referred to as the ≥ 1 jet (or "monojet") topology. And as the name suggests, the jet selection is intended to be as inclusive as possible by requiring the presence of only one energetic jet in addition to $p_{\mathrm{T}}^{\mathrm{miss}}$, the lepton(s) or a photon. The diagrams in Figure 5.1 are all examples of monojet topologies. The second phase space is enhanced in events undergoing vector boson fusion (VBF) and requires the presence of at least two energetic jets. The measurement of this additional topology is motivated by scenarios where BSM particles do not couple directly to quarks. The VBF topology additionally allows the measurement of dijet observables, in this analysis the dijet mass m_{jj} and the angular separation $\Delta\phi_{jj}$. VBF-like topologies for both SM and BSM are illustrated in Figure 5.3. Precise phase-space definitions and a motivation for the measured observables are presented in Section 5.2.

The measurement is performed in two complementary approaches. In the first approach, only the dominant SM contribution, which is always a vector boson plus jets process, is considered to be signal. Subdominant contributions from other SM processes are subtracted from the data. This is the "traditional" approach of a specific process measurement and is in this thesis referred to as the dominant boson process (DBP) approach. In the second approach however, the phase space is defined in terms of the fiducial final state. In other words, all sources contributing to this final state are considered to be signal and no subdominant SM processes are subtracted from the data. This is referred to as the *target final state* (TFS) approach or the inclusive $p_{\mathrm{T}}^{\mathrm{miss}}$ measurement. In this TFS approach, modelling of the subdominant SM processes is disentangled from the measurement itself. This degree of model-independence gives the analysis results a much bigger impact, as future updates and improvements to SM theoretical predictions of the subdominant backgrounds can be included and compared to the published data.

Each observable is unfolded for all combinations of final states, jet topologies, observables and background subtraction methods. The impact of statistical and experimental uncertainties are determined by re-running the the analysis with the variation applied and repeating the unfolding. The R^{miss} ratios are then constructed after unfolding. The particle-level measurements are

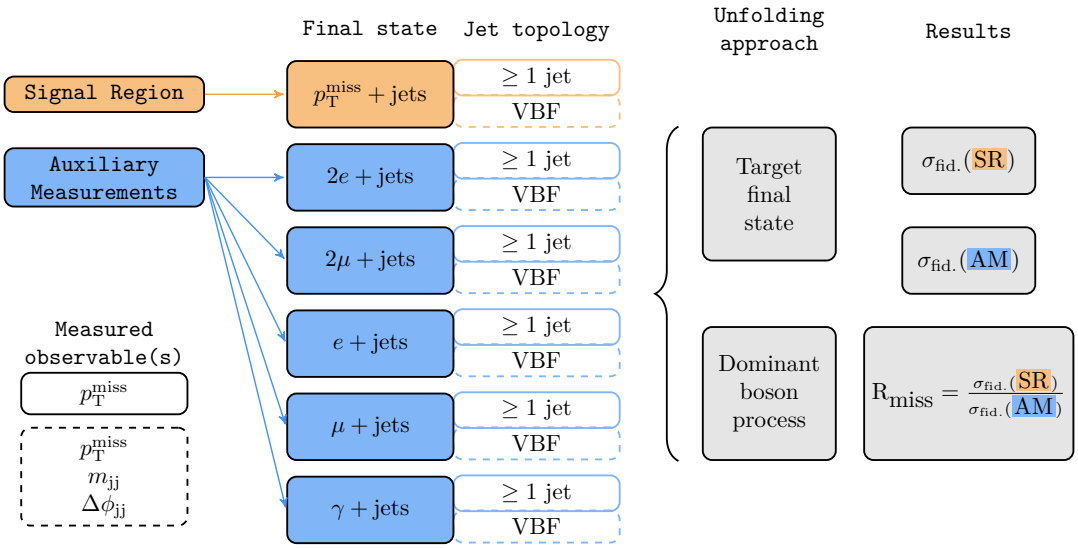


Figure 5.4: Schematic overview of the measured regions, jet topologies and observables.

compared directly with different SM predictions, and are used to constrain BSM parameters by comparing them to SM+BSM predictions. Figure 5.4 gives a global overview of all the measured regions and phase spaces.

5.2 Phase space definitions and measured observables

The measurement is made in the fiducial volume of the detector, closely reflecting the actual acceptance at detector-level, to minimize model-dependence and highly uncertain theory-based extrapolations. Indeed, the motivating principle of this work is to extract measurements based on what was actually observed in the ATLAS experiment. This section gives the object definitions and event selections to define this fiducial phase space. We first discuss the choices and definition of detector-level objects and event selections and follow this by the particle-level equivalent. The final part is an overview of the event variables that are measured as the final, unfolded differential distributions.

5.2.1 Detector-level objects

The following detector-level objects are considered in this analysis, which follow mainly from practical constraints of reconstruction and calibrations at detector-level.

Electrons are reconstructed from tracks in the ID and clusters of energy deposits in the EM calorimeters. The EM calorimeter barrel-to-endcap transition region, $1.37 < |\eta| < 1.52$, is excluded in the analysis. Electron candidates are reconstructed from topo-clusters matched to ID tracks. Three different categories of electrons are selected. First, *veto electrons* are defined by a set of baseline requirements and are used to estimate the fraction of fake leptons. Requirements on electron candidates include kinematic cuts, such as $p_{\mathrm{T}} > 7\,\mathrm{GeV}$ and $|\eta| < 2.47$, and good object quality, such as passing the `Loose` ID likelihood selection criteria. They need to pass the overlap criteria in Table 4.1. Second, *selected electrons in* $1e$*+jets* must satisfy, in addition to above criteria, the `Tight` identification working point to reduce non-prompt backgrounds. The leading elec-

tron is required to have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ corresponding to a trigger efficiency of 88% for the `Tight` electron candidates. Finally, *selected electrons in* $2e$*+jets* are also required to pass additional criteria with respect to the veto selection. The `Medium` identification working point is used, electrons are required to be associated to a primary vertex and the leading electron transverse momentum cut is $p_{\mathrm{T}} > 80\,\mathrm{GeV}$. At this transverse momentum, the efficiency to select electron candidates reaches a plateau of approximately 93% [189].

Topo-clusters in the EM not matched to any track or matched to conversion vertices are reconstructed as unconverted or converted **photon** candidates respectively. Conversion vertices are defined from one or two tracks, that are consistent with a massless particle decaying within the ID. Photon candidates in the the γ+jets region must satisfy `Tight` working points for both identification and isolation, rejecting photons from non-prompt backgrounds efficiently.

Similarly for **muons**, the *veto muon* category requires candidates to have $p_{\mathrm{T}} > 7\,\mathrm{GeV}$ and $|\eta| < 2.5$. In addition, they must pass the `Loose` identification requirements and the overlap removal criteria in Table 4.1. *Selected muons in* 1μ*+jets and* 2μ*+jets* must pass the veto selections, as well as the `Medium` identification working point. The *FixedCutLoose* isolation working point [190] is imposed to reject muons from semi-leptonic hadron decays. The transverse momentum is $p_{\mathrm{T}} > 7\,\mathrm{GeV}$ in the 1μ+jets region and $p_{\mathrm{T}} > 80\,\mathrm{GeV}$ ($p_{\mathrm{T}} > 7\,\mathrm{GeV}$) for the leading (sub-leading) muon in the 2μ+jets region.

Hadronically decaying **taus** must satisfy `Loose` selection criteria and pass the overlap removal in Table 4.1. Candidates must also have one or three associated tracks, corresponding to the one-prong and three-prong decay modes respectively, and carry an absolute electromagnetic charge of one. Taus that decay leptonically into electrons and muons are included in the electron and muon reconstruction respectively.

Jets are reconstructed with the anti-k_T algorithm with a distance parameter $R = 0.4$. Inputs to the jet clustering algorithm include both energy deposits in the calorimeter and charged tracks measured in the inner detector, combined using the particle flow algorithm described in Section 4.2.1. All jets are required to have $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|y| < 4.4$. In the central region a `Tight` set of jet-vertex-tagger (JVT), based on a combination of track-based variables, is applied. Jets in the forward region ($|\eta| > 2.5$) must pass the similar `Tight` track-based fJVT working point. The (f)JVT requirements reduce the rate of fake jets originating from pile-up. Finally, jets must pass the overlap removal procedure outlined in Section 5.2.3.

The ordering of jets in terms of their transverse momentum or (pseudo)rapidity can differ between particle level and detector level for a single event. The calculation of observables involving angles between jets are often defined in terms of this jet ordering. This leads to non-negligible, off-diagonal contributions to the response matrices used in the unfolding procedure. To suppress this effect a truth-matching requirement is applied, matching the leading and sub-leading jets at particle level and detector level within $\Delta R < 0.1$. This increases the stability of the unfolding, and causing the iterative procedure that is used to converge faster leading to reduced statistical uncertainties on the final measurement. The truth-matching is taken into account when deriving the matching efficiencies in the unfolding procedure, to ensure that truth-matching does not lead to unphysical changes of the final results after unfolding.

Missing transverse momentum ($p_{\mathrm{T}}^{\mathrm{miss}}$) is calculated from Equation 4.11. In the auxiliary mea-

surement regions, the x and y components of the momenta from the respective leptons and photons are added to $p_{\mathrm{T}}^{\mathrm{miss}}$, effectively treating them as invisible. This makes the event selection as similar as possible to the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region.

5.2.2 Detector-level event selections

The phase-space regions are defined as a set of cuts on detector-level objects.

There are six regions classified in terms of the number of leptons and photons, that each have two subregions (≥ 1 jet, VBF) giving a total of twelve combinations. The event selections for the phase-space regions in terms of lepton content are specified in Table 5.1. In spite of lepton-flavour universality, the experimental acceptance of electrons and muons is different which is reflected in the different cuts between the two lepton categories. Because high $p_{\mathrm{T}}^{\mathrm{recoil}}$ events tend to have one high p_{T} lepton, the p_{T} cut on the leading lepton is much larger than on the sub-leading lepton in the di-lepton regions. This reduces events from the $t\bar{t}$ background which tends to have lower p_{T} leptons. The cut on di-lepton mass $m_{\ell\ell}$ increases the relative rate of di-lepton production via a Z boson propagator instead of a photon propagator. Additional cuts on real $p_{\mathrm{T}}^{\mathrm{miss}}$ and m_{T} suppress the background contributions from multijet events in the $1e$+jets region.

Attribute	$p_{\mathrm{T}}^{\mathrm{miss}}$+jets	$1e$+jets	$2e$+jets	1μ+jets	2μ+jets
lepton rapidity	-	$\lvert y\rvert \leq 1.37$ or $1.52 \leq \lvert y\rvert \leq 2.47$		$\lvert y\rvert \leq 2.5$	
leading lepton p_{T} (GeV)	-	>30	>80	>7	>80
sub-leading lepton p_{T} (GeV)	-	-	>7	-	>7
$m_{\ell\ell}$ (GeV)	-	-	(66 , 116)	-	(66 , 116)
m_{T} (GeV)	-	(30, 100)	-	-	-
$p_{\mathrm{T}}^{\mathrm{miss}}$ (GeV)	>200	>60	-	-	-
$p_{\mathrm{T}}^{\mathrm{recoil}}$ (GeV)	>200	>200	>200	>200	>200

Table 5.1: Requirements defining the five principal phase space regions of the measurement. For the inclusive $p_{\mathrm{T}}^{\mathrm{miss}}$ measurement, $p_{\mathrm{T}}^{\mathrm{miss}} \equiv p_{\mathrm{T}}^{\mathrm{recoil}}$. In the $Z \to \nu\nu$ measurement, it corresponds to the p_{T} of the Z boson. The dilepton mass and transverse mass are indicated with $m_{\ell\ell}$ and m_{T} respectively.

The definitions of the two jet topologies are summarised in Table 5.2. Additional leptons and hadronically decaying taus are vetoed. Furthermore, the cut on $\Delta\phi\left(\mathrm{jet}, p_{\mathrm{T}}^{\mathrm{miss}}\right)$ vetoes events with $p_{\mathrm{T}}^{\mathrm{miss}}$ closely aligned with jets, greatly reducing the instrumental (fake) $p_{\mathrm{T}}^{\mathrm{miss}}$ background from multijet events. The single jet in the ≥ 1 jet topologies requires a p_{T} of at least $120\,\mathrm{GeV}$ and needs to be within the tracker region. The selections for this topology are very similar to the monojet signature studied in the ATLAS search with Run 2 data [191].

VBF processes typically result in two high-energy forward jets, coming from the the two quarks that participate in the interaction and carry a significant fraction of the proton's energy. Through a colour-neutral interaction an additional vector boson is produced, which decay products are distributed more widely across the detector. These are highly distinctive features of VBF processes and by applying appropriate cuts the purity of VBF events can be enhanced. First, the m_{jj} and $|\Delta y_{jj}|$ cuts exploits the fact that the jets have a large energy and scatter off primarily in opposite directions, respectively. The veto on jets produced in the gap between the forward jets

Attribute	≥ 1 jet	VBF
(Additional) muons	None with p_{T} >7 GeV, $\lvert\eta\rvert < 2.5$	
(Additional) electrons	None with p_{T} >7 GeV, $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.47$	
Hadronic τ	None with p_{T} >20 GeV, $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.47$	
$\Delta\phi\left(\mathrm{jet}, p_{\mathrm{T}}^{\mathrm{miss}}\right)$	>0.4 for four leading p_{T} jets	
leading jet p_{T} (GeV)	>120	>80
sub-leading jet p_{T} (GeV)	-	>50
lead jet $\lvert y\rvert$	<2.4	<4.4
sub-lead jet $\lvert y\rvert$	-	<4.4
m_{jj} (GeV)	-	>200
$\lvert\Delta y_{jj}\rvert$	-	>1
In-gap jets	-	None with $p_{\mathrm{T}} > 30$ GeV

Table 5.2: A summary of the fiducial selections applied to the hadronic recoil system to define the subregions of the measurement. The veto on 'in-gap jets' is applied to jets with a rapidity lying between the rapidities of the leading and the sub-leading jets.

reduces multijet backgrounds, but keeps VBF events where the absence of QCD activity in this region is in fact a distinctive VBF signature [192, 193].

Events are required to have at least one primary vertex with at least two associated tracks. Additional event quality selections outlined in Section 4.2.3 are imposed: only events from the Run-2 GRL are included and the recommended overlap removal procedure is applied at an event-by-event basis. The leading jets must pass tight cleaning, while additional jet criteria suppress non-collision backgrounds mainly caused by the calorimeters.

5.2.3 Particle-level objects

The following particle-level objects are used in this analysis.

Electrons and muons are required to be prompt, meaning that they do not originate from the decay of a hadron. This analysis measures "dressed leptons", objects created by adding the four-momenta of photons within a cone of $\Delta R < 0.1$ to the four-momentum of the lepton. Spatial cuts follow the coverage region of the ATLAS detector. Hence the particle-level acceptance of muons is $|\eta| < 2.5$, where the electron acceptance of $0 < |\eta| < 1.37$ or $1.52 < |\eta| < 2.47$ excludes the crack region at the end of the ECAL barrel.

In the signal region, $p_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the magnitude of a vector, which is the negative sum of two-momenta in the x and y direction of all visible final-state particles with $|\eta| < 5$. Muons with $|\eta| > 2.5$ or $p_{\mathrm{T}} < 7$ GeV are excluded from this sum, which ensures a matching to detector-level criteria on muon selection. The pseudo-rapidity cut corresponds to the ID region to have charged muon tracks while the p_{T} cut matches the baseline detector-level selection criterion in this analysis for muons. In the auxiliary measurements, there is no "true" $p_{\mathrm{T}}^{\mathrm{miss}}$ but rather a boson recoiling of the hadronic system. By treating identified charged dressed leptons (from either a W or Z boson) and isolated photons in these regions as invisible, a quantity corresponding to the momentum of the bosonic recoil system ($p_{\mathrm{T}}^{\mathrm{recoil}}$) is calculated in the same way as $p_{\mathrm{T}}^{\mathrm{miss}}$ in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region. By marking the bosonic (decay) products invisible for the $p_{\mathrm{T}}^{\mathrm{recoil}}$ calculation,

the kinematics of this observable stay as similar as possible to the $p_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Note that in the signal region $p_{\mathrm{T}}^{\mathrm{recoil}} \equiv p_{\mathrm{T}}^{\mathrm{miss}}$.

Jets in the **target final state approach** are reconstructed using the anti-k_T jet algorithm with a radius parameter of $\Delta R = 0.4$. All stable final-state particles are used as input to this algorithm, excluding neutrinos, dressed muons and other invisible objects. Jets in the **dominant boson process definition** are also constructed with the anti-k_T jet algorithm with a radius parameter of $\Delta R = 0.4$, using all stable final-state particles as input except prompt leptons from W, Z, Higgs and tau decays and prompt photons from Higgs decays. In both scenarios, the minimum leading jet p_{T} is $120\,\mathrm{GeV}$ or $80\,\mathrm{GeV}$ for the ≥ 1 jet or VBF region respectively, with a rapidity cut of $|y| < 4.4$. An overlap removal procedure is applied to remove any jets within $\Delta R < 0.2$ range of selected charged leptons. Jets containing a hadron from a tau-decay are classified as hadronically decaying taus.

5.2.4 Particle-level event selections

It is important to emphasise that particle-level event selections are defined as closely to detector-level selections to minimise extrapolations in the unfolding. Therefore, the same event selections from Table 5.1 and Table 5.2 are applied at particle-level.

5.2.5 Triggers

Events for this measurement are selected using the two-state (hardware and software based) trigger system described in Section 3.2.6. To find events that fulfil the definition of the fiducial phase space, a set of missing energy, single-photon or single-lepton triggers is used. Events for the signal region require a transverse momentum imbalance in the calorimeter system and hadronic jets in the final state. This analysis uses the lowest un-prescaled $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers for the different years of Run-2 data-taking. As the amount of pile-up increased throughout these years, the minimum $p_{\mathrm{T}}^{\mathrm{miss}}$ threshold was raised from $70\,\mathrm{GeV}$ to $120\,\mathrm{GeV}$ to suppress pile-up contributions. The $p_{\mathrm{T}}^{\mathrm{miss}}$ cut of $200\,\mathrm{GeV}$ in offline reconstruction ensures that all $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers have reached an efficiency over 99% [194].

Muons leave very little energy in the calorimeter, and therefore calorimeter-based $p_{\mathrm{T}}^{\mathrm{miss}}$ triggers are used to select events with high-p_{T} muons. Because the HLT does not use information from the muon systems in the calculation of the trigger-level $p_{\mathrm{T}}^{\mathrm{miss}}$, the same triggers from the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region are used in the 1μ+jets and 2μ+jets regions.

The selection strategy for the $1e$+jets and $2e$+jets region is implemented using a combination of low- and high-p_{T} single-electron triggers. Events in the low-p_{T} region are selected with two single-electron triggers, employing `Tight` selection criteria and minimum p_{T} thresholds of $24(26)\,\mathrm{GeV}$ in 2015–2016 (2017–2018). Because the relative rate of single-electron triggers is lower in the high-p_{T} region, less restrictive electron identification is required to enhance trigger efficiency in this regime. Events satisfying either the low- or high-p_{T} threshold trigger, with an efficiency of approximately 97% for electrons with $p_{\mathrm{T}} \geq 80\,\mathrm{GeV}$, are retained. To account for small mismodelling effects, within 5% for $p_{\mathrm{T}} \leq 60\,\mathrm{GeV}$ electrons and $\leq 1\%$ for high-p_{T} electrons, simulated events are re-weighted by data-driven scale factors. Both statistical and systematic uncertainties in deriving the trigger scale factors are propagated to the measured observables.

Region	ATLAS trigger	Trigger requirement in HLT
$p_{\mathrm{T}}^{\mathrm{miss}}$+jets, 1$\mu$+jets & 2$\mu$+jets	**2015**:	
	HLT_xe70_L1XE50	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 70\,\mathrm{GeV}$
	2016:	
	HLT_xe90_mht_L1XE50	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 90\,\mathrm{GeV}$
	HLT_xe110_mht_L1XE50	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 110\,\mathrm{GeV}$
	2017:	
	HLT_xe110_pufit_L1XE55	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 110\,\mathrm{GeV}$
	2018:	
	HLT_xe110_pufit_xe65_L1XE50	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 110\,\mathrm{GeV}$
	HLT_xe110_pufit_xe70_L1XE50	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 110\,\mathrm{GeV}$
	HLT_xe120_pufit_L1XE50	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 120\,\mathrm{GeV}$
1e+jets & 2e+jets	**2015**:	
	HLT_e24_lhmedium_L1EM20VH	`Medium` electron $E_{\mathrm{T}} \geq 24\,\mathrm{GeV}$
	HLT_e60_lhmedium	`Medium` electron $E_{\mathrm{T}} \geq 60\,\mathrm{GeV}$
	HLT_e120_lhloose	`Loose` electron $E_{\mathrm{T}} \geq 120\,\mathrm{GeV}$
	2016:	
	HLT_e26_lhtight_nod0_ivarloose	`Tight` electron $E_{\mathrm{T}} \geq 26\,\mathrm{GeV}$
	HLT_e24_lhmedium_nod0_L1EM20VH	`Medium` electron $E_{\mathrm{T}} \geq 24\,\mathrm{GeV}$
	HLT_e60_lhmedium	`Medium` electron $E_{\mathrm{T}} \geq 60\,\mathrm{GeV}$
	HLT_e60_lhmedium_nod0	`Medium` electron $E_{\mathrm{T}} \geq 60\,\mathrm{GeV}$
	HLT_e140_lhloose_nod0	`Loose` electron $E_{\mathrm{T}} \geq 140\,\mathrm{GeV}$
	HLT_e300_etcut	$E_{\mathrm{T}} \geq 300\,\mathrm{GeV}$
	2017:	
	HLT_e26_lhtight_nod0_ivarloose	`Tight` electron $p_{\mathrm{T}} \geq 26\,\mathrm{GeV}$
	HLT_e60_lhmedium_nod0	`Medium` electron $E_{\mathrm{T}} \geq 60\,\mathrm{GeV}$
	HLT_e140_lhloose_nod0	`Loose` electron $E_{\mathrm{T}} \geq 140\,\mathrm{GeV}$
	HLT_e300_etcut	$E_{\mathrm{T}} \geq 300\,\mathrm{GeV}$
	2018:	
	HLT_e26_lhtight_nod0_ivarloose	`Tight` electron $E_{\mathrm{T}} \geq 26\,\mathrm{GeV}$
	HLT_e60_lhmedium_nod0	`Medium` electron $E_{\mathrm{T}} \geq 60\,\mathrm{GeV}$
	HLT_e140_lhloose_nod0	`Loose` electron $E_{\mathrm{T}} \geq 140\,\mathrm{GeV}$
	HLT_e300_etcut	$E_{\mathrm{T}} \geq 24\,\mathrm{GeV}$
γ+jets	**2015**:	
	HLT_g120_loose	`Loose` photon $p_{\mathrm{T}} \geq 120\,\mathrm{GeV}$
	HLT_g200_etcut	$E_{\mathrm{T}} \geq 200\,\mathrm{GeV}$
	2016:	
	HLT_g140_loose	`Loose` photon $p_{\mathrm{T}} \geq 140\,\mathrm{GeV}$
	HLT_g300_etcut	$E_{\mathrm{T}} \geq 300\,\mathrm{GeV}$
	2017:	
	HLT_g140_loose	`Loose` photon $p_{\mathrm{T}} \geq 140\,\mathrm{GeV}$
	HLT_g300_etcut	$E_{\mathrm{T}} \geq 300\,\mathrm{GeV}$
	2018:	
	HLT_g140_loose	`Loose` photon $p_{\mathrm{T}} \geq 140\,\mathrm{GeV}$
	HLT_g300_etcut	$E_{\mathrm{T}} \geq 300\,\mathrm{GeV}$

Table 5.3: A summary of the triggers used in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region, the lepton-based regions and the single photon region [194, 195].

To select events in the γ+jets region, trigger-level requirements include a photon candidate with a minimum p_{T} of $120(140)$ GeV in 2015 (2016-18) and `Loose` photon identification. To improve the trigger efficiency at high-p_{T}, only a p_{T} selection is used in the range above $200(300)$ GeV in 2015 (2016-18). The used triggers are fully efficient in the phase space of the photon auxiliary measurement.

A summary of all the triggers used is presented in Table 5.3.

5.2.6 Definition of observables

The monojet region measures $p_{\mathrm{T}}^{\mathrm{recoil}}$ in both the signal region and auxiliary measurement regions. In the signal region, $p_{\mathrm{T}}^{\mathrm{recoil}}$ is by definition equal to $p_{\mathrm{T}}^{\mathrm{miss}}$. This is the first Run-2, inclusive measurement of this observable with ATLAS data. Because many BSM models involve particles with large masses and energy scales, their decay into (stable) final state particles leads to a spectrum with a relatively large contribution in the tails compared to SM predictions. Any deviations between the measured data and SM predictions can be a sign for new physics. This signature has been studied extensively by various LHC experiments, in the context of dark matter, supersymmetry, axions, invisible Higgs decays, dark energy and more [191, 196–199]. Measuring $p_{\mathrm{T}}^{\mathrm{miss}}$ in an inclusive ≥ 1 jet region is motivated by the fact that different BSM models have different jet topologies, providing sensitivity to a broad range of theories.

For the same reasons, $p_{\mathrm{T}}^{\mathrm{miss}}$ is measured in the VBF region. Additionally, the VBF region is sensitive to BSM models that, in terms of SM particles, only couple to electroweak gauge bosons rather than quarks or gluons. The requirement of at least two jets opens the floor to an additional measurement of dijet observables, in this analysis the invariant mass of the two leading jets m_{jj} and the angle between them $\Delta\phi_{jj}$. The smoothly falling spectrum of the dijet mass is sensitive to a local excess from BSM phenomena [200] which has been an active research area in ATLAS since the start [201–204]. The angular observable $\Delta\phi_{jj}$ is defined as the angle between the leading (ϕ_1) and sub-leading (ϕ_2) jets: $\Delta\phi_{jj} \equiv \phi_1 - \phi_2$. This observables probes the CP properties the particle produced by the VBF [205, 206], leading to more insights about both the (SM) Higgs boson and/or BSM particles produced in this process, see Figure 5.3.

5.3 Datasets, Monte Carlo samples and detector backgrounds

This section describes the datasets used to make an unfolded, inclusive measurement of the observables and regions of phase-space described in earlier sections. The first ingredient is the experimental data, as recorded by the detector, which is to be unfolded. The second ingredient is theoretical predictions from MC event generators. The simulated events are used in two different ways. First, in the unfolding procedure, and secondly to provide a state-of-the-art comparison of the SM and BSM models with the measured experimental data. Datasets are accessible through distributed computing resources at many physical locations that together form the LHC computing grid (LCG)[1], and may contain particle-level information (MC), detector-level information (data and MC) or both (MC).

After describing the experimental datasets used in this analysis (Section 5.3.1), the event gener-

[1] In practice, LHC physicists like to refer to this higher power as "The Grid".

ation software in ATLAS (Section 5.3.2) and data processing steps (Section 5.3.3) are introduced. MC samples for the unfolding procedure are discussed in Section 5.3.4-5.3.6. This is followed by a description of the truth-level samples used in the comparisons between unfolded data and MC in Section 5.3.7, including SM-only samples to quantify the agreement between data and the SM, and BSM samples to set limits on the parameter space of new physics. In addition to contributions from scattering events in proton-proton collisions, instrumental backgrounds contribute fake 'events' to the experimental data in the measurement regions. This section concludes by a description of methods to estimate instrumental backgrounds in Section 5.3.8.

5.3.1 Experimental dataset

The LHC collided protons at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ from 2015 to 2018. During this period, referred to as Run-2, a total integrated luminosity of $140\,\text{fb}^{-1}$ data for physics analyses was recorded. A detailed breakdown of the luminosity (and the luminosity uncertainty) per year of data-taking is presented in Table 3.1. The average pile-up per year was 18, 25, 38 and 36 in the years 2015, 2016, 2017 and 2018 respectively. The data is stored in containers that are not only subdivided into different years, corresponding to different pile-up profiles, but also in periods of data-taking under similar conditions (e.g. between two technical stops).

5.3.2 MC event generation in ATLAS

MC event generation in ATLAS is done through the dedicated ATLAS generator software AthGeneration. This provides direct, out-of-the-box access to numerous state-of-the-art generators, and more importantly ensures consistent usage of these tools and generator tunes in the collaboration. The generator tools and event generation are configured via steering scripts in Python, referred to as jobOptions (JOs), to define commands for AthGeneration to run. MC datasets for analysis are produced and stored centrally by the Physics Modelling Group, where experts on different generators verify configurations from the generator and provided JOs.

Events with a fully generated set of final state particles, in the EVNT format, are then used as input to Geant4 [136, 138, 207], for simulation of the interaction between final-state particles and the ATLAS detector. Pile-up collisions are overlaid and pile-up reweighting is applied, following the description in Section 4.1.6. The simulated interactions in the HITS format are digitized to simulate detector outputs and stored as RDO files [137]. These outputs are fed through same reconstruction and analysis chain as the experimental data, to evaluate efficiencies and migration matrices used to correct for detector effects. The various stages of data-processing for both data and simulation are summarised in Figure 5.5.

5.3.3 From raw data to derivations

After reconstruction and calibration from the trigger and data acquisition system, the primary outputs are Analysis Object Data (AOD) files containing physics outputs like electrons, muons and tracks. Although the primary AOD files are fully calibrated and aligned between data and MC, they are not used directly in analysis. Analyses may only need a subset of the many triggers that the main AODs contain. The ATLAS Derivation Framework takes the primary AOD and creates derived AOD (DAOD) datasets. These select events with a particular physics goals and

drop unneeded information, significantly reducing the size and therefore increasing the speed it takes to run analysis selections over the datasets.

This analysis uses the "EXOT5 derivation" from the Exotics Working Group, tailored to monojet signatures [208]. The EXOT5 skimming includes the requirement of online $p_{\mathrm{T}}^{\mathrm{miss}}$, single electron, single muon and single photon triggers. The offline requirement is a logical OR on either calibrated jets (detector and particle level) or uncalibrated detector-level jets. More specifically it selects a single high p_{T} jet ($p_{\mathrm{T}}(j1) > 100\,\mathrm{GeV}$), or otherwise events with two high p_{T} jets ($p_{\mathrm{T}}(j1), p_{\mathrm{T}}(j2) > 100\,\mathrm{GeV}$) with additional requirements on dijet observables ($m_{jj} > 150\,\mathrm{GeV}$ and $\Delta\eta_{\mathrm{jj}} > 2.5$).

5.3.4 MC samples used in unfolding procedure

Due to the inclusive nature of the measurement, various SM and BSM processes can contribute in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets phase-space. The samples in this section are put through ATLAS detector simulation to get both particle-level and detector-level information, and are used *only* in the unfolding procedure.

Standard Model samples

The main V+jets samples, including a single W, Z/γ^* boson or prompt photon in association with jets, are produced with the SHERPA v2.2.1 [85, 86] parton shower MC generator. This setup provides NLO-accurate matrix-elements for up to two jets and LO-accurate matrix elements for up to four jets, calculated with the COMIX [210] and OPENLOOPS [211] libraries. The default SHERPA parton shower (based on Catani-Seymour dipole [121]) is used with a cluster hadronisation model [134]. The V+jets samples are normalised to a next-to-next-to-leading order (NNLO) prediction [212]. A similar setup, but now using SHERPA v.2.2.11, is used for electroweak production of a W or Z/γ^* boson in association with two jets and up to one additional parton emission at LO accuracy.

Diboson events with fully-leptonic decays are generated with SHERPA v2.2.2. Multiple matrix-elements are matched and merged with the SHERPA parton shower based on the Catani-Seymour ordering using the MEPS@NLO prescription [132]. Virtual QCD corrections for matrix elements at NLO accuracy are provided by the OPENLOOPS library. Triboson samples are simulated the same setup, where semi-leptonically decaying diboson samples have nearly the same setup but use SHERPA v2.1.1 instead.

Events with top quarks are divided in four categories: $t\bar{t}$, s-channel and t-channel single-top production, and tW associated production,. On-shell $t\bar{t}$ events are modelled using POWHEG-BOX [89, 90, 213, 214] v2 at NLO accuracy, interfaced with PYTHIA 8.230. The inclusive NLO $t\bar{t}$ production cross-section is reweighted to NNLO (QCD) predictions including the resummation of NNLL soft-gluon terms. Single top s-channel and t-channel production are modelled using POWHEG-BOX at NLO in the five and four flavour scheme respectively. Associated production of a top-quark and a W boson is also modelled at NLO in QCD using POWHEG-BOX in the five flavour scheme. The diagram removal (DR) scheme [215] is used in nominal predictions to handle overlap with $t\bar{t}$ production. An alternative prediction with the diagram subtraction (DS) method [215] is used as uncertainty estimate for the interference between $t\bar{t}$ and tW production.

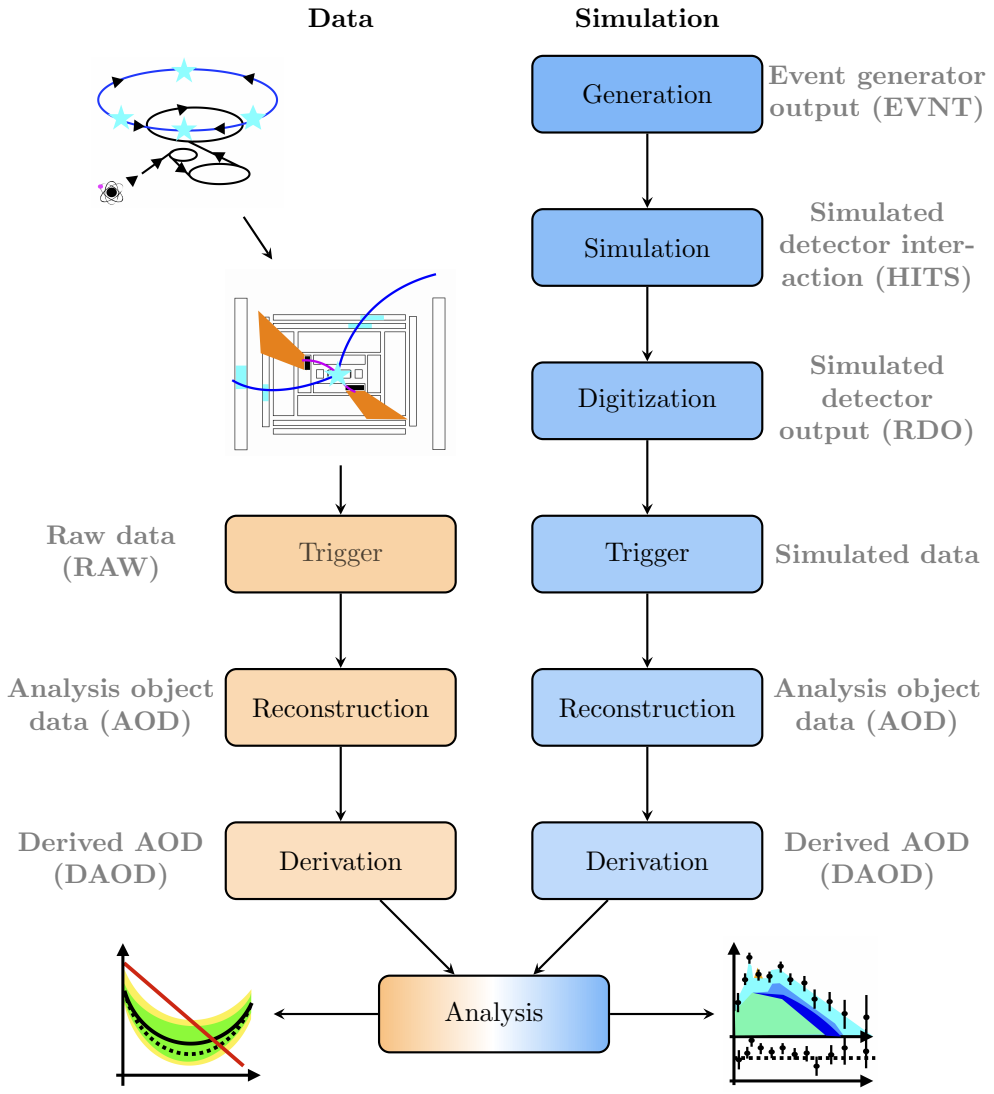


Figure 5.5: Data processing flowchart from experimental data (left column) and event generation (right column) to derivations that are used in analysis. The data formats for each processing step are indicated in parentheses. The illustrations are adapted from Ref. [209], kindly provided by Louie Corpe.

A summary of all fully simulated MC samples used in the unfolding is presented in Table 5.4.

Physics Process	Generator	Parton Shower	Accuracy of cross-section	Tune	PDF Set
QCD V+jets	SHERPA v2.2.1	SHERPA v2.2.1	NNLO	SHERPA default	NNPDF3.0NNLO
γ+jets	SHERPA v2.2.1	SHERPA v2.2.1	NLO	SHERPA default	NNPDF3.0NNLO
EWK V+dijets	SHERPA v2.2.11	SHERPA v2.2.11	LO	SHERPA default	NNPDF3.0NNLO
VV (fully leptonic)	SHERPA v2.2.2	SHERPA v2.2.2	NLO	SHERPA default	NNPDF3.0NNLO
VVV	SHERPA v2.2.2	SHERPA v2.2.2	NLO	SHERPA default	NNPDF3.0NNLO
VV (semi-leptonic)	SHERPA v2.2.1	SHERPA v2.2.2	NLO	SHERPA default	NNPDF3.0NNLO
$t\bar{t}$	POWHEG-BOX v2	PYTHIA v8.230	NNLO+NNLL	A14	NNPDF3.0NLO
Single top (Wt)	POWHEG-BOX v2	PYTHIA v8.230	NLO	A14	NNPDF3.0NLO
Single top (t-channel)	POWHEG-BOX v2	PYTHIA v8.230	NLO	A14	NNPDF3.0NLOnf4
Single top (s-channel)	POWHEG-BOX v2	PYTHIA v8.230	NLO	A14	NNPDF3.0NLO

Table 5.4: Simulated background event samples used in the unfolding procedure of this analysis. The table shows corresponding matrix element and parton shower generators, perturbative accuracy, underlying-event tune and PDF set.

Beyond the SM samples

BSM samples are used in signal injection studies, to test if the unfolding procedure is not biased to the Standard Model and consequently "unfolds away" any signal that might be present in the experimental data. In this analysis, this is done with Weakly Interacting Massive Particle (WIMP) signals of type $\chi\bar{\chi}$ + jet, where χ is a DM candidate. Events are simulated with POWHEG-BOX, interfaced to PYTHIA 8.205 with the A14 tune [216].

Two variations of this simplified model are considered. In the first one, $\chi\bar{\chi}$ production with spin-1 axial-vector mediator exchanges [34] is simulated at NLO precision for different masses of the DM candidates and the mediator. In this model, the couplings of the mediator to SM quarks and DM particles are set to $g_q = 1/4$ and $g_\chi = 1$ respectively. For this choice of parameters the narrow-width approximation remains valid. The choice of $g_\chi = 1$ increases the probability of the mediator decaying into DM particles, leading to missing energy signatures, instead of decaying back to a pair of quarks which would lead to a dijet signature instead. The second model has a similar final state, but is mediated by a spin-0 pseudo-scalar particle with a quark loop at LO accuracy. Couplings to the mediator are both set to $g_q = g_\chi = 1$, and are motivated by the same reasons as the axial-vector mediated scenario.

5.3.5 Normalisation of subdominant SM backgrounds

SM irreducible backgrounds are treated as signal in the final state measurement, as opposed to the dominant boson approach where they are subtracted from the data. In both cases, a thorough validation of the shape and normalization of the SM processes is needed to make a robust measurement in both the signal and auxiliary measurement regions. The relative contribution of the dominant process varies for the different regions and this is summarised in Table 5.5. Misreconstruction of one or more leptons, when the lepton is out of acceptance or due to a detector inefficiency, is typically the reason for background contamination from different processes.

Because it is not possible to measure leptons outside the detector acceptance, the shapes of these backgrounds are normally taken from MC simulation. The MC modelling needs to be validated in the phase-space of this measurement. The regions mainly suffering from large irreducible SM backgrounds are $p_{\mathrm{T}}^{\mathrm{miss}}$+jets, $1e$+jets and 1μ+jets. A set of background control regions (BCRs)

is introduced for the dominant SM background processes in these regions: $W \to \ell\nu$ and top processes. The $2e$+jets, 2μ+jets and γ+jets regions have a high purity of the respective signal process. In the latter regions, predictions are purely MC based taking shape uncertainties on their cross-section (scale, PDF and α_s) as systematic uncertainties.

Process	$p_{\mathrm{T}}^{\mathrm{miss}}$+jets SR	$2e$+jets AM	2μ+jets AM	$1e$+jets AM	1μ+jets AM	γ+jets AM
$Z \to \nu\nu$ + jets	**55%**	0%	0%	0%	0%	0%
$Z \to ee$ + jets	0%	**94%**	0%	0%	0%	0%
$Z \to \mu\mu$ + jets	0%	0%	**95%**	0%	2%	0%
$W \to e\nu$ + jets	6%	0%	0%	**68%**	0%	0%
$W \to \mu\nu$ + jets	9%	0%	0%	0%	**67%**	0%
$W \to \tau\nu$ + jets	20%	0%	0%	5%	7%	0%
γ + jets	0%	0%	0%	0%	0%	$>$ **99%**
top	7%	3%	2%	25%	21%	0%
rest	3%	3%	3%	2%	3%	$<$ 1%

Table 5.5: SM process composition for the signal region, the four lepton AM regions and the gamma AM region, excluding instrumental detector backgrounds. Di- and triboson samples are including in "rest". For each measurement region, the process contribution from the dominant process is highlighted in bold. This includes only theoretical SM predictions from MC generators, but does not include any potential BSM physics that can contribute in this region.

The main background processes in $p_{\mathrm{T}}^{\mathrm{miss}}$+jets are $W \to e\nu$, $W \to \mu\nu$, $W \to \tau\nu$ and top. For each process, a BCR is defined by the "standard" $p_{\mathrm{T}}^{\mathrm{miss}}$+jets event selections and requiring the respective lepton(s) in addition to this as shown in Table 5.6. The $p_{\mathrm{T}}^{\mathrm{miss}}$+jets event selection has no cuts on lepton rapidity or transverse mass, which is how the BCRs for $W \to e\nu$ and $W \to \mu\nu$ differ from the $1e$+jets and 2μ+jets auxiliary measurements. The opposite-charge requirement for the top region suppresses contributions from $Z \to \ell\ell$ decays. Comparison between data and MC in the BCR regions shows a good shape agreement with varying normalization offsets. We then introduce normalization constants of the different MCs and fit those a data-driven approach, where the data, MC yields and normalization constants are related in the following way:

$$\mathrm{Data}^{\mathrm{Y}} - \mathrm{MC}^{\mathrm{Y}}_{\mathrm{bkg.}} = k_{W\to e\nu}\mathrm{MC}^{\mathrm{Y}}_{W\to e\nu} + k_{W\to\mu\nu}\mathrm{MC}^{\mathrm{Y}}_{W\to\mu\nu} + k_{W\to\tau\nu}\mathrm{MC}^{\mathrm{Y}}_{W\to\tau\nu} + k_{\mathrm{T}op}\mathrm{MC}^{\mathrm{Y}}_{\mathrm{T}op}, \quad (5.2)$$

where Y indicates the background region, normalization factors are indicated by k and $\mathrm{MC}^{\mathrm{Y}}_{\mathrm{bkg.}}$ refers to the remaining backgrounds in BCR Y from dibosons, tribosons and $Z \to \ell\ell$. A bootstrapping method is used to obtain a statistical uncertainty on the scale factors and preserve any correlations, by propagating bootstrap replicas of the scale factors to the rest of the analysis chain. Systematic uncertainties on the scale factors are obtained by changing the definition of the BCR, e.g. by increasing or reducing the lepton p_{T} cut to $40\,\mathrm{GeV}$ or $20\,\mathrm{GeV}$ respectively.

The definitions of the $W \to e\nu$+jets and $W \to \mu\nu$+jets BCRs for the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region are not orthogonal to the $1e$+jets and 1μ+jets auxiliary measurement regions. As the normalization factors for these backgrounds are applied the SR in both the TFS and DBP approach, this introduces correlations between the SR and the single-lepton auxiliary measurements. These turn out to be very small and localised only to a few bins, but nonetheless these correlations are fully accounted for using a bootstrap method.

A similar procedure is applied to validate and normalize backgrounds the $1e$+jets and 1μ+jets region, which in this case are mainly events from $W \to \tau\nu$+jets and top processes (see Table 5.5).

Background control region	$W \to e\nu$+jets	$W \to \mu\nu$+jets	$W \to \tau\nu$+jets	Top
Selections	1 `Tight`, isolated electron	1 `Medium`, isolated muon	1 `Medium` τ	1 `Medium`, isolated electron 1 `Medium`, isolated muon opposite charge
	lepton $p_{\mathrm{T}} \geq 30\,\mathrm{GeV}$ pass $p_{\mathrm{T}}^{\mathrm{miss}}$+jets event selections			

Table 5.6: Definition of background control regions (BCRs) for the dominant $p_{\mathrm{T}}^{\mathrm{miss}}$+jets SR backgrounds. For $W \to e\nu$+jets and $W \to \mu\nu$+jets the lepton identification working points are chosen to be consistent with the AM regions where applicable.

Background control region	Top for $1e$+jets	Top for 1μ+jets	$W \to \tau\nu$+jets
Selections	$1e$+jets event selection $n_{jets} \geq 6$	1μ+jets event selection $n_{jets} \geq 6$	1 `Medium` τ $30 < m_T < 100\,\mathrm{GeV}$ true $p_{\mathrm{T}}^{\mathrm{miss}} \geq 160\,\mathrm{GeV}$
	lepton $p_{\mathrm{T}} \geq 30\,\mathrm{GeV}$		

Table 5.7: Definition of background control regions (BCRs) for the dominant $1e$+jets and 1μ+jets backgrounds.

The definitions of the BCRs for 1ℓ+jets are summarised in Table 5.7 and resulting scale factors applied to subdominant SM backgrounds in the signal region and $1e$+jets are presented in Table 5.8. The fit favours a reduction of the top-quark MC samples. Because the top contribution is much larger in the single-lepton channels than in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets SR (as seen in Table 5.5), this is effect is more pronounced in the single-lepton regions leading to lower normalization factors than the SR. A potential reason why this is different between the zero-lepton and single-lepton regions, is that the $t\bar{t}$ contribution in the former is mainly due to the leptonic decay channel with out-of-acceptance leptons, while the single-lepton channel $t\bar{t}$ events have both in-acceptance and out-of-acceptance leptons. Nonetheless each background contribution, i.e. the top backgrounds but also the QCD V+jets, is fluctuated up and down with 30%, leading to a negligible effect on the final unfolded results and hence no additional systematic uncertainty is associated to this relatively low normalization factor in the single-lepton regions.

5.3.6 MC reweighting for dijet mass distributions

The reconstruction-level m_{jj} distributions, comparing detector-level data to SM predictions from MC generators simulated for the ATLAS detector, are presented in the left column of Figure 5.6. This shows a clear shape disagreement between the data and MC. Historically, accurate modelling of the m_{jj}-shape in V+jets processes is proven to be notoriously difficult. Examples of this date back to the previous generation of colliders experiments at Tevatron [217]. This is still a relevant problem at the LHC: examples in ATLAS include Refs. [218, 219] for Zjj and Wjj production at $\sqrt{s} = 7\,\mathrm{TeV}$ and $\sqrt{s} = 8\,\mathrm{TeV}$, but it is also observed at $\sqrt{s} = 13\,\mathrm{TeV}$ for a measurement of Zjj events [220].

Ultimately, this is an issue in Monte Carlo event generators that is also apparent in the SHERPA v2.2.1

Region	Background MC	≥ 1 jet phase-space	VBF phase-space
$p_{\mathrm{T}}^{\mathrm{miss}}$+jets region	$W \to e\nu$+jets	$1.09 \pm 0.01 \pm 0.01$	$1.10 \pm 0.01 \pm 0.02$
	$W \to \mu\nu$+jets	$1.08 \pm 0.01 \pm 0.03$	$1.11 \pm 0.01 \pm 0.04$
	$W \to \tau\nu$+jets	$1.08 \pm 0.01 \pm 0.03$	$1.13 \pm 0.01 \pm 0.06$
	Top	$0.97 \pm 0.01 \pm 0.01$	$0.98 \pm 0.01 \pm 0.03$
$1e$+jets region	$W \to \tau\nu$+jets	$1.10 \pm 0.01 \pm 0.01$	$1.10 \pm 0.01 \pm 0.02$
	Top	$0.78 \pm 0.01 \pm 0.02$	$0.70 \pm 0.04 \pm 0.05$
1μ+jets region	$W \to \tau\nu$+jets	$1.11 \pm 0.01 \pm 0.02$	$1.10 \pm 0.01 \pm 0.03$
	Top	$0.75 \pm 0.01 \pm 0.01$	$0.70 \pm 0.01 \pm 0.03$

Table 5.8: Background scale factor results in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets, $1e$+jets and 1μ+jets regions. The first uncertainty corresponds to the statistical uncertainty from the bootstrap method, where the second uncertainty is the systematic uncertainty from variations of the BCR definition.).

QCD V+jets samples derived for this phase-space. It is important to make the distinction between models used to predict a theory (in this case the SM) and those used to predict the effect of the detector. For the unfolding, we are interested in the latter case. This motivates a data-driven correction on the m_{jj} shape of the SHERPA v2.2.1 QCD V+jets samples.

More specifically, an event-by-event reweighting procedure is applied to the MC sample to account for the m_{jj} mismodelling, because a better agreement between data and MC modelling improves the robustness and fidelity of the unfolding. The SHERPA v2.2.1 QCD V+jets samples are well-understood and have been the main samples since the start of the analysis to validate and optimise the unfolding procedure. Therefore, these are chosen to model the detector response in the phase-space of this measurement. After the experimental data is unfolded, updated SHERPA 2.2.11 QCD V+jets samples are used to assess the agreement between data and MC at particle-level.

In the VBF region, the reweighting functions are a fit to the raw, per-bin data/MC ratio for the QCD V+jets processes, derived as follows:

$$\mathrm{RW}^i = \frac{\mathrm{data}^i - \mathrm{MC}^i_{\text{non V+jets}}}{\mathrm{MC}^i_{\text{V+jets}}}, \tag{5.3}$$

where i indicates the bin index and RW is the per-bin value for the fit. The numerator gives an approximation of the V+jets contribution in the actual data, and dividing this number by the V+jets contribution in pure MC this ratio isolates the V+jets mismodelling.

The m_{jj} RW ratios are derived using the non-reweighted MC vs. data m_{jj} histograms in Figure 5.6, using Equation 5.3. The plots in the right column of Figure 5.6 show the ratio RW^i per bin in blue. A fit is made to derive a more continuous reweighting function that mitigates statistical fluctuation between bins, which is implemented by interpolating an exponential function for the low m_{jj} region and a second order polynomial in the higher m_{jj} region. The functions are interpolated with a hyperbolic tangent. This combination is found to provide the best fit to the non-trivial shape of RW^i, where also higher order polynomials or a combination of solely exponential functions have been studied. Still, in most cases the highest m_{jj} bin, which suffers from low statistics and therefore fluctuates the most, is not fitted easily by the polynomial. Therefore, an additional degree is added to the fit, by fitting the last bin to a constant value and interpo-

lating this with the polynomial using again a hyperbolic tangent. The full functional form is the following:

$$F(m_{jj}) = (1 - \mathrm{s}_1) \cdot p_3 \exp(p_4 m_{jj}) \tag{5.4}$$

$$+ \mathrm{s}_1 \cdot (1 - \mathrm{s}_2) \cdot \left(p_5 + p_6 \cdot (m_{jj} - p_7) + p_4 \cdot (m_{jj} - p_8)^2\right) \tag{5.5}$$

$$+ \mathrm{s}_2 \cdot p_9 \tag{5.6}$$

with the swap functions defined as

$$\mathrm{s}_{1,2}(m_{jj}) \equiv 0.5 \left(1 + \tanh\left(\frac{m_{jj} - p_{1,2}}{\gamma}\right)\right) \tag{5.7}$$

The γ parameter in the denominator of the hyperbolic tangent defines the "width" of the transition region between the different fit functions. The fit to the RW^i histograms are indicated in orange in Figure 5.6. The large fluctuation in the highest m_{jj} bin of the $2e$+jets region is not included in the fit because it is contained within the large statistical uncertainty of that bin.

In the analysis code, which iterates over all the events in the MC samples, the m_{jj} value is calculated on an event-by-event basis. Then, if the given event passes the VBF region cuts and the event is from a QCD V+jets sample, the event is reweighted according to the reweighting function for the corresponding region, using the fit parameters obtained from the non-reweighted spectrum.

The following logic is applied on event-by-event basis to decide on the m_{jj} reweighting:

1. If it passes the VBF cuts **and** is a QCD V+jets sample: reweight event according to the corresponding VBF RW function, fill histograms in VBF region according to the updated weight.

2. Else, if it passes the VBF cuts **and not** a QCD V+jets sample: apply no m_{jj} reweighting factor and fill VBF histogram. Note that in this scenario the event weight can still be different from one, due to lepton scale factors or other weights applied in the analysis.

3. Else, if VBF cuts are not satisfied: event is vetoed and does not increment the VBF histograms.

Validation plots for the m_{jj} reweighting in VBF are presented in Section 5.5. The histograms are scaled such that the integral matches that of the non-reweighted histograms in Figure 5.6. The impact of m_{jj} reweighting on the unfolded data is studied by using m_{jj} reweighted MC to non-reweighted MC in the unfolding procedure and comparing the results, shown in Figure 5.7. The residual difference is negligible compared to the statistical uncertainty on the data in discrepant bins, and hence no systematic uncertainty is associated to this reweighting procedure.

It is worth repeating that the m_{jj} reweighting is applied only to the MC which is used as inputs for the migration matrices, which are used in unfolding. The reweighting does not directly impact the data yields and therefore does not bias the final unfolded result. The reweighted MC is only used to predict the detector response of the data, and in Section 5.6.5 we answer the question if this reweighted MC is robust to BSM possibly being hidden in the measured (unfolded) data.

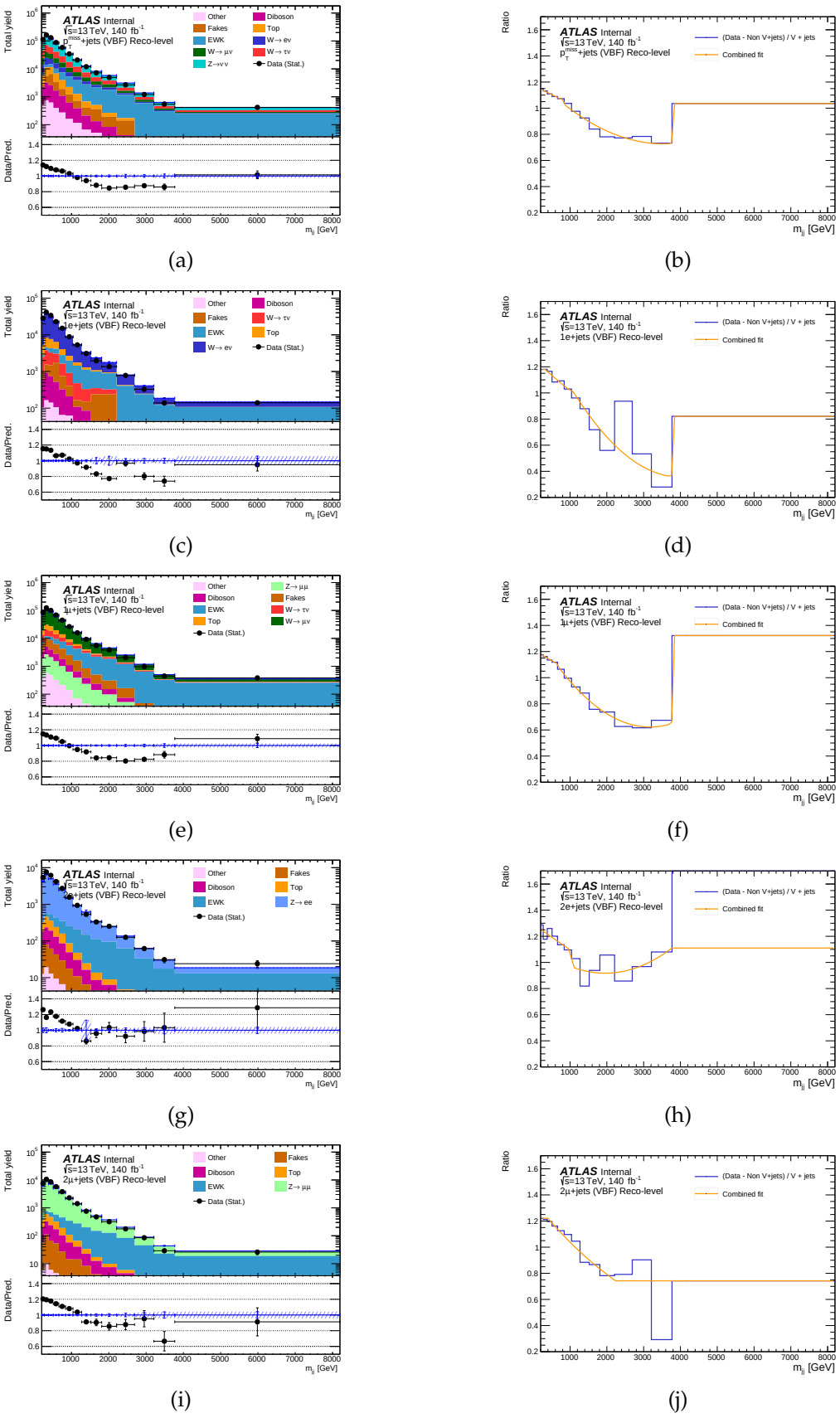
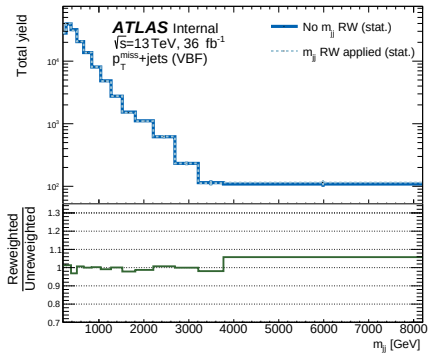
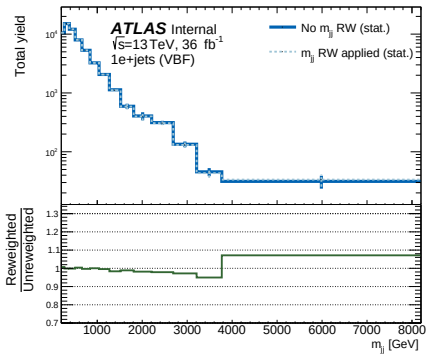


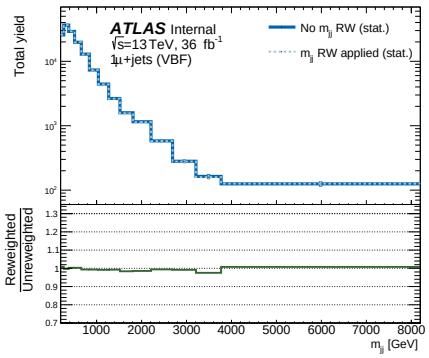
Figure 5.6: Reweighting of m_{jj} observable, where figures on the left show the detector-level data and MC, with the corresponding binned ratio of data with non-V+jets processes subtracted to the V+jets contribution, as in Equation 5.3, indicated as blue histograms in the figures on the right. The orange lines indicate the best fits using Equation 5.4.



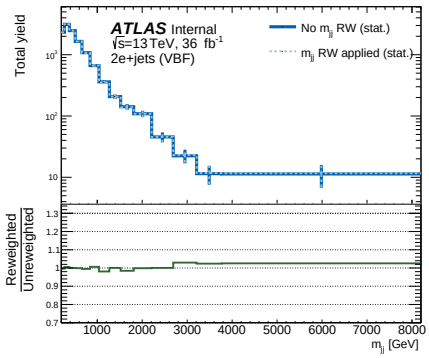
(a)



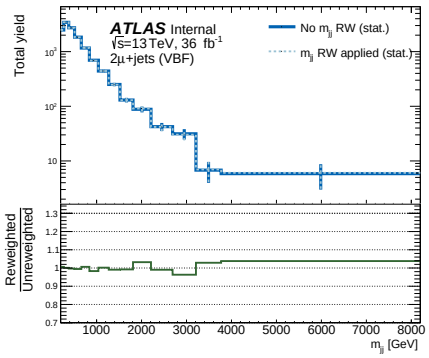
(b)



(c)



(d)



(e)

Figure 5.7: Comparison between unfolded data with and without m_{jj} reweighting applied to the MC used in the unfolding procedure.

5.3.7 MC samples for particle level comparisons

Particle-level predictions are used to quantify the agreement between the experimental data and the SM in a fit, and to put bounds on the parameter space of different BSM theories. For these predictions, no data-driven scale factors are applied, no m_{jj} reweighting is applied and pile-up events are not added.

Standard Model samples

Particle-level predictions are produced using the same generators and configurations as the samples used for unfolding, with the following exceptions.

As mentioned in Section 5.3.6, SHERPA v2.2.1 suffers from known modelling issues for QCD V+jets processes. Therefore, the single weak boson MC samples are replaced with calculations produced with SHERPA v2.2.11 and OPENLOOPS 2 [211, 221, 222], extending the matrix-element-level description of additional emissions up to five jets at LO. The predictions are reweighted to account for NLO EW corrections [223, 224], based on parton-level prediction for $W/Z+$jets from Ref. [225]. Diboson samples also use SHERPA v2.2.11 and OPENLOOPS 2 for matrix-element calculations. The PDF4LHC PDF set [226] was used, with QED effects from LUX$_{\mathrm{QED}}$ [227]. Top-quark pair production samples are produced with SHERPA v2.2.11, using LO-accurate matrix elements up to four additional partons and NLO-accurate matrix elements for up to one additional parton. This setup constitutes the 'nominal' NLO QCD V+jets prediction used in all regions of phase-space and for all measured observables, and will be referred to as the "NLO prediction", referring to the highest order in the QCD $V+$jets calculation. Additional predictions are presented for both the ≥ 1 jet and VBF region.

For ≥ 1 jet, an alternative $V+$jets prediction is obtained by extending the reweighting procedure to include NNLO QCD corrections from Ref. [225]. The stack that uses these NNLO-reweighted $V+$jets predictions is referred to as the "NNLO prediction". The NNLO predictions are used in individual $p_{\mathrm{T}}^{\mathrm{recoil}}$ cross-section measurements and the corresponding R^{miss} ratios.

In the VBF region, alternative $V+$jets samples are obtained using the High Energy Jets (HEJ) framework [228, 229]. This framework provides an all-order (in α_s) description of wide-angle QCD emissions for all relevant Standard Model processes. These corrections are relevant to the VBF phase-space, as they improve the modelling for dijet pairs with a larger invariant mass or rapidity gap. The framework is implemented in a generator [230], which is used for predictions of processes with charged leptons of $p_{\mathrm{T}}^{\mathrm{recoil}}$, $\Delta\phi_{jj}$ and m_{jj}. At the time of writing, the framework is not yet implemented for $Z \to \nu\nu+$jets processes and therefore no alternative prediction for the $p_{\mathrm{T}}^{\mathrm{miss}}+$jets VBF region is available. Consequently, HEJ predictions are not included for the R^{miss} observable.

A summary of all particle-level MC samples used for the interpretation of unfolded data is presented in Table 5.9.

Beyond the SM samples

The motivation for making a particle-level, inclusive measurement of $p_{\mathrm{T}}^{\mathrm{miss}}$ is the sensitivity to BSM physics and the ease of comparing different BSM scenarios to the experimental data. This is demonstrated in the analysis with two models.

Physics process	Variation	Generator	Parton shower	Accuracy of cross-section	Tune	PDF set
QCD V+jets	NLO	SHERPA v2.2.11	SHERPA v2.2.11	NLO	SHERPA default	PDF4LHC
	NNLO	SHERPA v2.2.11	SHERPA v2.2.11	NNLO	SHERPA default	PDF4LHC
	HEJ	HEJ 2.2	-	NLO+NLL	-	NNPDF3.0NNLO
γ+jets	-	SHERPA v2.2.1	SHERPA v2.2.1	NLO	SHERPA default	NNPDF3.0NNLO
EWK V+dijets	-	SHERPA v2.2.11	SHERPA v2.2.11	LO	SHERPA default	NNPDF3.0NNLO
VV (fully leptonic)	-	SHERPA v2.2.11	SHERPA v2.2.11	NLO	SHERPA default	NNPDF3.0NNLO
VVV	-	SHERPA v2.2.2	SHERPA v2.2.2	NLO	SHERPA default	NNPDF3.0NNLO
VV (semi-leptonic)	-	SHERPA v2.2.11	SHERPA v2.2.11	NLO	SHERPA default	NNPDF3.0NNLO
$t\bar{t}$	-	SHERPA v2.2.11	SHERPA v2.2.11	NLO	A14	NNPDF3.0NLO
Single top (Wt)	-	POWHEG-BOX v2	PYTHIA v8.230	NLO	A14	NNPDF3.0NLO
Single top (t-channel)	-	POWHEG-BOX v2	PYTHIA v8.230	NLO	A14	NNPDF3.0NLOnf4
Single top (s-channel)	-	POWHEG-BOX v2	PYTHIA v8.230	NLO	A14	NNPDF3.0NLO

Table 5.9: Particle-level event samples used to interpret unfolded data. The table shows corresponding matrix element and parton shower generators, perturbative accuracy, underlying-event tune and PDF set. The three different QCD V+jets setups (NLO, NNLO and HEJ) are explained in more detail in the text.

First, a simplified model which extends the SM with an additional Dirac fermion DM candidate χ is considered [34]. The DM particles can be produced at the LHC via a mediator Z' gauge boson. Signal events with $\chi\bar{\chi}$ + jets final states are simulated using the same generator setup as the BSM samples used for signal injection tests in the unfolding.

Second, signal samples for the two Higgs doublet model with a pseudo-scalar mediator (2HDM+a) [33] are simulated with MADGRAPH5a_MC@NLO [92] with leading order accuracy and at one-loop level. This takes into account all processes that produce a DM pair in association with up to two SM particles or another DM pair. The matrix-element calculation is interfaced to PYTHIA 8.245 for the parton shower and hadronisation, using the ATLAS A14 tune.

5.3.8 Instrumental (fake) backgrounds

ATLAS triggers on tracks and energy depositions in the subdetectors and uses these signals to reconstruct objects. For different reasons, a combination of signals be falsely identified as a particular signal. For example, due to finite resolution of the hadronic calorimeters a multijet event with a mismeasured p_{T} of one of the jets can be classified as an event with large missing transverse momentum. Or, leptons from heavy flavour decays can be mis-tagged as prompt leptons. And what about cosmic muons accidentally traversing the muon spectrometer? These sources of background are referred to as instrumental or "fake" backgrounds. This section discusses the important sources of instrumental noise in this measurement and methods obtain robust estimates of the different fake contributions.

Fake missing transverse momentum

When jets are mis-reconstructed or mis-calibrated, this can appear in the detector as a momentum imbalance and hence give rise to fake $p_{\mathrm{T}}^{\mathrm{miss}}$. Furthermore, heavy flavour hadrons inside the jets can produce neutrinos that give rise to $p_{\mathrm{T}}^{\mathrm{miss}}$. In those cases, the $p_{\mathrm{T}}^{\mathrm{miss}}$ will be aligned with the direction of the jet. The majority of this background is therefore removed by the $\Delta\phi(\text{jet}, p_{\mathrm{T}}^{\mathrm{miss}}) > 0.4$ requirement, which is applied to $p_{\mathrm{T}}^{\mathrm{miss}}$ and the leading four jets. However, multijet events may still result in fake $p_{\mathrm{T}}^{\mathrm{miss}}$ signatures when it receives contributions from several jets and the resulting fake $p_{\mathrm{T}}^{\mathrm{miss}}$ is not aligned with any of the jets. To estimate this residual multijet background, a data-driven method is used. This is motivated by the fact that jet production cross-section is

large, but the probability to generate fake $p_{\mathrm{T}}^{\mathrm{miss}}$ is low. This would require a very large simulated sample in a simulation-based approach and accurate modelling of the high-p_{T} tails.

To this end the jet smearing method is used, similar to the procedure outlined in Ref. [231]. The main idea is to start with a large sample with well-measured jets collected using a set of inclusive jet triggers with different $p_{\mathrm{T}}^{\mathrm{jet}}$ thresholds. In this context "well-measured" means that seed events require at least two jets, no leptons and low $p_{\mathrm{T}}^{\mathrm{miss}}$. A simple $p_{\mathrm{T}}^{\mathrm{miss}}$ cut would bias the distribution of the leading jet p_{T} as they are correlated. Therefore, a cut on the significance of the missing transverse energy S is used:

$$S = \frac{p_{\mathrm{T}}^{\mathrm{miss}} - 8\,\mathrm{GeV}}{\sqrt{\sum E_{\mathrm{T}}}} < 0.05\sqrt{\mathrm{GeV}}, \tag{5.8}$$

where $\sqrt{\sum E_{\mathrm{T}}}$ is the scalar sum of the event transverse energy.

By fluctuating the four-momenta of these jets multiple times, with jet response-functions that are constrained by data, a pseudo-dataset is obtained with the shape of the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ background. For each seed event, all the jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ are smeared 2000 times. Each fluctuation is then considered as a separate event, for which the altered four-momenta are stored and the $p_{\mathrm{T}}^{\mathrm{miss}}$ vector recalculated. This provides distributions for the fake multijet background with fake $p_{\mathrm{T}}^{\mathrm{miss}}$ up to about $2\,\mathrm{TeV}$ after applying the event selection cuts for the ≥ 1 jet and VBF regions.

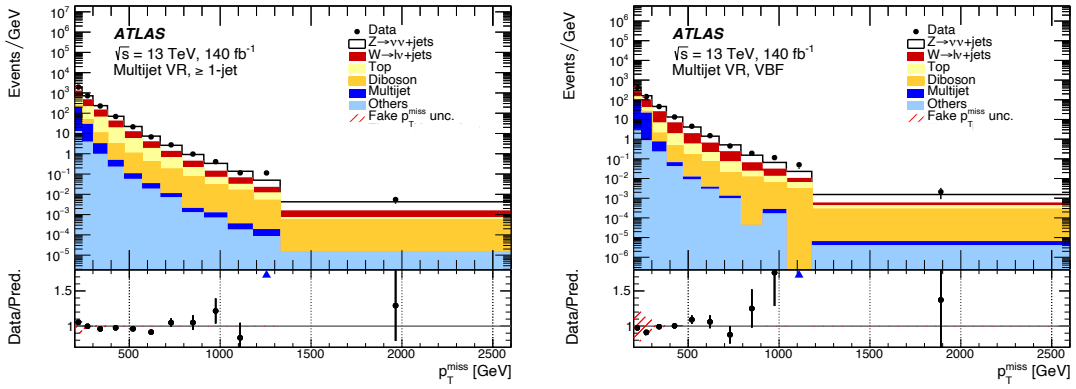
The normalization is extracted from a control region which is defined by *inverting* the cut on the azimuthal angle between $p_{\mathrm{T}}^{\mathrm{miss}}$ and jets to $\Delta\phi(\mathrm{jet}, p_{\mathrm{T}}^{\mathrm{miss}}) < 0.3$. By purposely aligning $p_{\mathrm{T}}^{\mathrm{miss}}$ with jets, this control region is enriched in fake the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ background. The approach is validated in a region defined by the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets requirements and $0.3 < \Delta\phi(\mathrm{jet}, p_{\mathrm{T}}^{\mathrm{miss}}) < 0.4$. In this validation region, the multijet background represents a significant fraction of the first few bins leading to a good agreement between data and MC in this region, as illustrated in Figure 5.8. In the main $p_{\mathrm{T}}^{\mathrm{miss}}$+jets event selection, the multijet background contribution is much smaller than in this validation region. In the ≥ 1 jet region, it reaches up to 1% of the total event yield for $p_{\mathrm{T}}^{\mathrm{recoil}} \leq 300\,\mathrm{GeV}$ but falls steeply as a function of $p_{\mathrm{T}}^{\mathrm{recoil}}$. The multijet background is around 1-2% for $m_{jj} < 2\,\mathrm{TeV}$ and is negligible for larger values.

Fake leptons

The main source of detector backgrounds in the lepton auxiliary measurements comes from "fake leptons". These originate from jets that are misidentified as prompt leptons or leptons from heavy flavour decays. Even though leptons from heavy flavour decays are actual leptons, they are classified as fake because they do not originate from as part of the prompt decay of the hard scatter process of interest. The contribution from fake leptons is strongly suppressed by the reconstruction algorithms, but any remaining contributions are estimated using a data-driven method [233] summarised below.

The main goal is to derive an estimate of the number of prompt leptons ($N_{\mathrm{S}}^{\mathrm{P}}$) and the number of background (fake) leptons ($N_{\mathrm{S}}^{\mathrm{B}}$) passing the signal selections (S), but the ATLAS reconstruction algorithms will only give a number of electrons for different selection criteria without knowing the origin of the lepton. We define three categories of lepton identification:

1. Leptons that pass the loose criteria (L),



(a) ≥ 1 jet selection in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region

(b) VBF selection in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region

Figure 5.8: The event yield in the multijet background validation region for the $p_{\mathrm{T}}^{\mathrm{miss}}$ observable in the (a) ≥ 1 jet validation region and (b) the VBF validation region. Black markers denote the data, and different SM backgrounds are shown as coloured histograms. The red hatched band shows the full uncertainty assigned the multijet estimate, which is 100% on the estimated yield, and vertical lines represent the statistical uncertainty on experimental data. In both plots, there is a sizeable contribution from the multijet background in the lowest three $p_{\mathrm{T}}^{\mathrm{miss}}$ bins, where good agreement between data and MC is observed within systematic and statistical uncertainties of the multijet background. Figure from Ref. [232].

2. Leptons that pass the signal selections (S),

3. Leptons that pass the loose criteria but failed the signal selection (!S),

With these definitions, we can decompose the total number of leptons that pass the signal selections as

$$N_{\mathrm{S}} = N_{\mathrm{S}}^{\mathrm{P}} + N_{\mathrm{S}}^{\mathrm{B}} = N_{\mathrm{L}}^{\mathrm{P}} \frac{N_{\mathrm{S}}^{\mathrm{P}}}{N_{\mathrm{L}}^{\mathrm{P}}} + N_{\mathrm{L}}^{\mathrm{B}} \frac{N_{\mathrm{S}}^{\mathrm{B}}}{N_{\mathrm{L}}^{\mathrm{B}}} \equiv N_{\mathrm{L}}^{\mathrm{P}} \epsilon^{\mathrm{P}} + N_{\mathrm{L}}^{\mathrm{B}} \epsilon^{\mathrm{B}}. \tag{5.9}$$

Here ϵ^{P} and ϵ^{B} are defined as the efficiencies with which with a prompt or background lepton respectively passes the signal selection, given that the `Loose` criteria are satisfied:

$$\epsilon^{\mathrm{P}} \equiv \frac{N_{\mathrm{S}}^{\mathrm{P}}}{N_{\mathrm{L}}^{\mathrm{P}}}, \quad \epsilon^{\mathrm{B}} \equiv \frac{N_{\mathrm{S}}^{\mathrm{B}}}{N_{\mathrm{L}}^{\mathrm{B}}}. \tag{5.10}$$

Similarly, and expression is derived for the number of leptons that pass the loose criteria but fail the signal selection:

$$N_{\mathrm{!S}} = N_{\mathrm{!S}}^{\mathrm{P}} + N_{\mathrm{!S}}^{\mathrm{B}} = \left(N_{\mathrm{L}}^{\mathrm{P}} - N_{\mathrm{S}}^{\mathrm{P}}\right) + \left(N_{\mathrm{L}}^{\mathrm{B}} - N_{\mathrm{S}}^{\mathrm{B}}\right) = N_{\mathrm{L}}^{\mathrm{P}}\left(1 - \epsilon^{\mathrm{P}}\right) + N_{\mathrm{L}}^{\mathrm{B}}\left(1 - \epsilon^{\mathrm{B}}\right). \tag{5.11}$$

Equations 5.9 and 5.11 can be combined in the following matrix equation:

$$\begin{pmatrix} N_{\mathrm{S}} \\ N_{\mathrm{!S}} \end{pmatrix} = \begin{pmatrix} \epsilon^{\mathrm{P}} & \epsilon^{\mathrm{B}} \\ 1 - \epsilon^{\mathrm{P}} & 1 - \epsilon^{\mathrm{B}} \end{pmatrix} \begin{pmatrix} N_{\mathrm{L}}^{\mathrm{P}} \\ N_{\mathrm{L}}^{\mathrm{B}} \end{pmatrix}, \tag{5.12}$$

To estimate fake electrons in the $1e$+jets region, the "matrix method" is used. In this method, the matrix equation is solved directly by inverting the matrix in Equation 5.12, to obtain the number of background electrons in data that satisfy `Tight` requirements:

$$N_{\mathrm{S}}^{\mathrm{B}} = \epsilon^{\mathrm{B}} N_{\mathrm{L}}^{\mathrm{B}} = \frac{\epsilon^{\mathrm{B}}}{\epsilon^{\mathrm{P}} - \epsilon^{\mathrm{B}}} \left(\left(\epsilon^{\mathrm{P}} - 1\right) N_{\mathrm{S}} + \epsilon^{\mathrm{P}} N_{\mathrm{!S}} \right), \tag{5.13}$$

where the right-hand side of this equation only contains components that can be obtained using event yields in measured data. The efficiency ϵ_P is estimated from simulation, and changes from 85% to 95% as the electron p_{T} increases from $100\,\mathrm{GeV}$ to $600\,\mathrm{GeV}$. The efficiency ϵ_B is estimated in a data-driven way, in a dedicated region that is enriched in fake leptons by removing the cuts on m_T, real $p_{\mathrm{T}}^{\mathrm{miss}}$ and $p_{\mathrm{T}}^{\mathrm{recoil}}$. In this region, events with "real" prompt electrons are simulated and subtracted from the data. The efficiency ϵ_B falls steeply from 15% for electron $p_{\mathrm{T}} \sim 50\,\mathrm{GeV}$ to approximately 1% for $p_{\mathrm{T}} \geq 500\,\mathrm{GeV}$.

The resulting efficiencies ϵ_P and ϵ_B are binned in electron p_{T} and η. The final estimates of the background contributions are validated in a dedicated validation region with the nominal $1e$+jets event selection but no requirement on m_T or $p_{\mathrm{T}}^{\mathrm{miss}}$, which shows good agreement between data and prediction after including the fake estimates. The relative contribution ranges from approximately 5% at $p_{\mathrm{T}}^{\mathrm{recoil}} < 500\,\mathrm{GeV}$ to nearly 20% at $p_{\mathrm{T}}^{\mathrm{recoil}} > 1500\,\mathrm{GeV}$ in the validation region, and drops to the 1% level in the actual $1e$+jets AM.

The single-muon, di-muon and di-electron event samples have a high purity in the 1μ+jets, 2μ+jets and $2e$+jets regions respectively. This allows the use of the 'fake-factor' method to estimate the fakes in these regions. This method starts from the same assumption in Equation 5.12, where the number of non-prompt leptons satisfying `Tight` requirements is

$$N_{\mathrm{S}}^{\mathrm{B}} = \epsilon^{\mathrm{B}} N_{\mathrm{L}}^{\mathrm{B}} = \frac{\epsilon^{\mathrm{B}}}{1 - \epsilon^{\mathrm{B}}} \left(N_{\mathrm{!S}} - N_{\mathrm{!}S}^{\mathrm{P}}\right) \equiv F\left(N_{\mathrm{!S}} - N_{\mathrm{!}S}^{\mathrm{P}}\right), \tag{5.14}$$

where the 'fake-factor' is defined as $F \equiv \epsilon^{\mathrm{B}}/(1 - \epsilon^{\mathrm{B}})$. The relation in the second step of Equation 5.14 can be obtained by multiplying $N_{\mathrm{L}}^{\mathrm{B}}$ with $(1 - \epsilon^{\mathrm{B}})$:

$$N_{\mathrm{L}}^{\mathrm{B}}\left(1 - \epsilon^{\mathrm{B}}\right) = N_{\mathrm{L}}^{\mathrm{B}} - \epsilon^{\mathrm{B}} N_{\mathrm{L}}^{\mathrm{B}} = N_{\mathrm{L}}^{\mathrm{B}} - N_{\mathrm{S}}^{\mathrm{B}} = N_{\mathrm{!S}}^{\mathrm{B}} = N_{\mathrm{!S}} - N_{\mathrm{!S}}^{\mathrm{P}} \tag{5.15}$$

This method uses simulation rather than data to estimate the real prompt-lepton contribution ($N_{\mathrm{!}S}^{\mathrm{P}}$) of the loose lepton sample, and relies on the fact that real-lepton kinematics and efficiencies are well modelled in MC generators and simulation.

The efficiencies ϵ^{B} for the 1μ+jets, 2μ+jets and $2e$+jets are determined in dedicated control regions enriched in fake leptons, measuring events with two different flavour same charge leptons. Fake contributions come for example from W+jets events where a fake lepton originates from a quark decay. Event selections on jet p_{T} ($p_{\mathrm{T}} > 100\,\mathrm{GeV}$) and m_{jj} ($m_{jj} > 150\,\mathrm{GeV}$) make the fake regions kinematically more similar to the measured lepton regions. The fake efficiencies are constructed by subtracting the prompt MC component from data, for both `Loose` and signal criteria, and dividing them following Equation 5.10.

For muons, the fake efficiency is found to be a flat 5% up to a lepton p_{T} of $30\,\mathrm{GeV}$, after which it linearly increases to approximately 55% for lepton transverse momenta up to $1\,\mathrm{TeV}$. The to-

tal fake muon background is 5% and $< 1\%$ in the 1μ+jets and 2μ+jets event selection respectively. The ϵ^{B} for electrons in the $2e$+jets selections rises up to approximately 25% for lepton $pt \leq 20\,\text{GeV}$, and goes up to 31% for a p_{T} of $110\,\text{GeV}$. Across the differential distributions at detector level, the fake electron background in the $2e$+jets region does not exceed 1% of the total estimate.

Non-collision background

Muons produced in cosmic-ray showers, beam-pipe interactions from the proton-proton beam or interactions between residual molecules and the proton-proton beam can lead to background signatures in the ATLAS detector. They can leave significant energy depositions in the calorimeters that are constructed as jets. If this happens in-time with a proton-proton collision, this can lead to a large $p_{\text{T}}^{\text{miss}}$ signature. This type of background is referred to as the non-collision background (NCB), because it does not originate from the colliding proton beams as opposed to the fake multijet background. A characteristic signature of this type of fake jets is that it peaks at $\phi = 0, \pi$. The contribution of this background is reduced significantly by the jet selections on the leading jet, such as the (f)JVT requirement. The residuals are estimated in a data-driven approach, exploiting the different timing properties between signal jets from hadronic collisions and jets from the NCB. The NCB estimate is at the order of per mille relative to the total number of events expected from SM processes for the lowest $p_{\text{T}}^{\text{miss}}$ region in the ≥ 1 jet topology, i.e. $200\,\text{GeV} < p_{\text{T}}^{\text{miss}} < 240\,\text{GeV}$, and decreases further as $p_{\text{T}}^{\text{miss}}$ increases. This estimate is equally small in the VBF region.

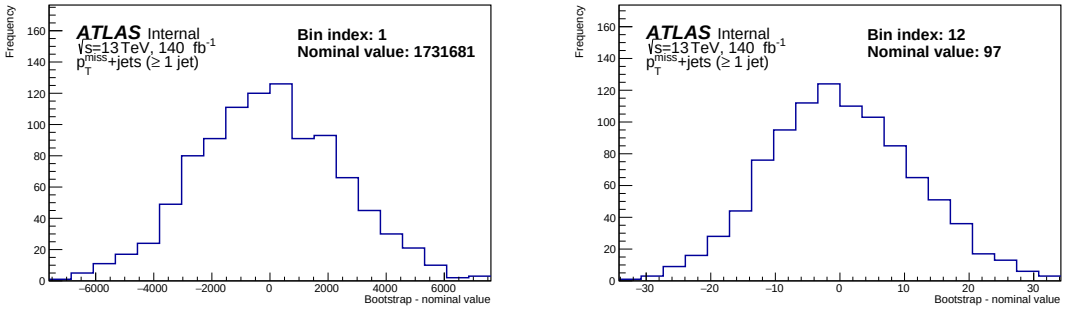
5.4 Uncertainties

5.4.1 Statistical uncertainties

The bootstrap method [234, 235] is used to propagate statistical uncertainties on detector-level data through the unfolding and preserve correlations between bins across different observables, phase-space regions and jet topologies. It is a powerful technique that generates pseudo-datasets from the nominal dataset, based on random numbers generated from a Poisson distribution corresponding to statistical fluctuations. This can be thought of the outcome of the experiment in a parallel universe, or an answer to the question what would happen if one was to repeat the experiment under identical conditions.

In this description, the *nominal* refers to unmodified values of events and histograms from data or MC event generation, before the bootstrap method is applied. In collider experiments the total number of events in the nominal dataset N_{nominal}, or in other words the size of the dataset, is a Poisson distributed variable. In the bootstrap method approach used in this analysis, for each event in the nominal dataset a unique weight is sampled from a Poisson distribution with a mean of unity. When applying the unique weights to each event when filling the histograms in the same way as for the nominal dataset, an alternative histogram referred to as "bootstrap replica" is obtained.

This procedure of resampling from a Poisson distribution can be repeated for any desirable number of bootstrap replicas, to generate a given number of alternative datasets and assess the impact



(a) (b)

Figure 5.9: Distribution of one thousand unfolded bootstrap replicas for both the (a) lowest ($200\,\text{GeV} < p_{\text{T}}^{\text{miss}} < 240\,\text{GeV}$) and (b) the highest ($p_{\text{T}}^{\text{miss}} > 2.6\,\text{TeV}$) bin of the $p_{\text{T}}^{\text{miss}}$ observable in the $p_{\text{T}}^{\text{miss}}$+jets ≥ 1 jet region. The "nominal value" indicates the nominal number of events reconstructed in a given bin. For every bin, including those not shown in this figure, the shape follows a Poisson distribution. This indicates that the chosen number of bootstrap replicas, $N_{\text{replica}} = 1000$, yields a sufficiently large number of data points to reliably estimate the statistical uncertainties and correlations with Equation 5.16.

of statistical fluctuations on the measurement. It allows us to track correlations between different bins and in addition, as we will see in the next paragraph, is used to propagate statistical uncertainties through the unfolding and calculate statistical uncertainties on experimental data as the RMS from the bootstrap ensemble.

The sums of independent Poisson variables with a mean of unity is itself a Poisson variable, which has a mean equal to the summed means: $\text{Pois}(N) = \sum_{i=1}^{N} \text{Pois}(1)$. As a result, the size of the replica datasets will follow a Poisson distribution with mean N_{nominal}. In this analysis, one thousand bootstrap replicas are generated for each nominal histogram. With a Poisson distribution of mean $\lambda = 1$ there is a 36.8% probability of sampling zero, which essentially removes a given event for the bootstrap replica. There is the same 36.8% probability of sampling a weight of 1, assigning the same weight to an event as in the nominal dataset, while an event will be double-counted with a probability of 18.4%. Figure 5.9 shows the number of events in different bins of the $p_{\text{T}}^{\text{miss}}$ observable in $p_{\text{T}}^{\text{miss}}$+jets region for the ≥ 1 jet topology generated for the thousand generated bootstrap replicas. In each bin, this clearly replicates a Poisson distribution centred around the nominal value it was sampled from.

After generating N_{replica} bootstrap replicas, the statistical covariance between any two variables, which can be the bin values in the case of a binned analysis, is calculated as

$$\text{COV}(a,b) = \frac{1}{N_{\text{replica}}} \sum_{i=0}^{N_{\text{replica}}} (a_i - \bar{a})(b_i - \bar{b}), \tag{5.16}$$

where a_i, b_i correspond to the measured variables in bin a, b for the bootstrap replica with index i and $\bar{a}, \bar{b}$ indicate the nominal values. From this formula, it is a straightforward generalisation to derive the statistical error for each variable as $\sigma_a = \sqrt{\text{COV}(a,a)}$ or $\sigma_b = \sqrt{\text{COV}(b,b)}$, and the statistical correlation as $\rho_{a,b} = \text{COV}(a,b)/(\sigma_a \sigma_b)$.

Equation 5.16 captures the key features and beauty of the bootstrap method, allowing the corre-

lation between different bins to be understood qualitatively. Bins a and b in Equation. 5.16 can be within the same histogram, but importantly they can also be bins of in different observables or different measurement regions altogether. Consider an event with label X that is measured in bin y and bin z of two different observables, and the sampled Poisson weight for this event is w_X^i. There can be other events that end up in bin y and z, but also events that end up in either y and not in z and vice versa. An example if an event that passes the cuts for the VBF region and is reconstructed in bins of $p_{\mathrm{T}}^{\mathrm{recoil}}$, $\Delta\phi_{jj}$ and m_{jj}. There is also overlap between the ≥ 1 jet region and the VBF region as they are not completely orthogonal, so a single event can end up in ≥ 1 jet distributions as well as VBF distributions. For the bootstrap replica i, the bootstrapped value of y_i can then be expressed as $y_i = w_X^i + \sum_{i=0}^{N^y_{\mathrm{nominal}}-1} \mathrm{Pois}(1) \sim w_X^i + \mathrm{Pois}(N_{\mathrm{nominal}} - 1)$, and similarly for z_i. The mean of $\mathrm{Pois}(N^y_{\mathrm{nominal}} - 1)$ is by definition $N^y_{\mathrm{nominal}} - 1$, and we can therefore express the difference between some replica value y_i and the average $\bar{y}$ as follows:

$$(y_i - \bar{y}) \xRightarrow[\text{average}]{\text{large } N_{\text{replica}}} w_X^i + N^y_{\mathrm{nominal}} - 1 - N^y_{\mathrm{nominal}} = w_X^i - 1 \,. \tag{5.17}$$

If a large number of replicas is generated, this term will converge to a positive or negative number for $w_X^i > 1$ and $w_X^i = 0$ respectively. The same argument holds for z, and as a result both terms will be positive for $w_X^i > 1$ or both negative for $w_X^i < 1$. In both cases, the contribution to the sum of Equation 5.16 is positive, indicating a correlation between the two bins. If there are more events that end up in both bin y and z, this effect is more pronounced and the covariance increases. If no events from bin y end up in bin z, this effect is washed away by the random weights and the covariance averages to zero. This is true at detector level and by unfolding each bootstrap replica separately these correlations are propagated to the particle-level measurement.

Correlations between bins within a histogram are not present at detector-level and enter during the unfolding procedure as migration of events between different bins can occur. By unfolding each bootstrap replica of the ensemble separately, these correlations are accounted for. The self-correlations within histograms and correlation across different histograms for the unfolded data are presented in Section 5.7. These correlations are a key input for the fitting framework that interprets the data in terms of the SM and BSM scenarios, and are also published on HEPData [236] for re-interpretation purposes in Ref. [237].

5.4.2 Experimental uncertainties

Objects in ATLAS are constructed using tools where the output depends on calibrations, energy scales and and energy resolutions that each have associated systematic uncertainties. To characterise the impact on the final results, a nuisance parameter (NP) is introduced for each experimental uncertainty. The NP has a nominal value and uncertainty that is derived from calibrations for example. Each NP is varied by one standard deviation around its nominal value. The effect on the analysis is evaluated by changing NPs one at a time, while keeping the other NPs at their nominal value, and re-evaluating the measurement for this changed set of parameters. This yields a unique set of results, including measured differential distributions and response matrices, for each systematic uncertainty.

In this measurement, systematic uncertainties are propagated through the unfolding procedure by repeating the unfolding for each NP with the associated response matrix and the nominal

distributions of the data. In other words, experimental uncertainties are transferred onto the data through the response matrices. The relative shift between the nominal distribution and the distribution for a given NP shift is taken as the error on the unfolded measurement associated to the corresponding systematic uncertainty. The resulting total, experimental error for a bin x is calculated from the differences added in quadrature as

$$\left(\sigma_{\text{exp.}}^{\text{tot}}\right)^2 = \sum_{n}^{N_{\text{syst}}} \left(x_{\text{nominal}} - x_{\text{systematic n}}\right)^2, \tag{5.18}$$

where the sum runs over all the systematic uncertainties N_{syst}. The experimental uncertainties evaluated in this analysis are detailed below.

Jet uncertainties related to the jet energy scale (JES) and jet energy resolution (JER) are evaluated using samples dijet events in $\sqrt{s} = 13\,\text{TeV}$ proton-proton collisions as outlined in [159]. Two additional steps are taken to evaluate these uncertainties. First, JES uncertainties include a "Jet Flavour" category related to energy scales whether the jet was initiated from a quark or a gluon. The default assumption of a 50/50 split between quarks and gluons, and an associated uncertainty of 100% each way on this, leads to needlessly large values for this uncertainty. A gluon-fraction map is derived for the regions of phase space used in this measurement, with corresponding custom JES-flavour uncertainties which are less conservative. This reduces the overall JES uncertainty, the dominant source at low $p_{\text{T}}^{\text{miss}}$, in the $p_{\text{T}}^{\text{miss}}$+jets region from values between 5% and 2% to a range between 3% and 0.5%. Second, JER uncertainties make use of random number generators to add additional smearing to jets, which therefore has a stochastic component. A bootstrap method is used to evaluate the statistical component on JER uncertainties, to mitigate spurious effects in sparsely populated bins of the measurement. An additional smoothing is applied to control the shape of these uncertainties. There are over 40 NPs related to jet uncertainties with up and down variations.

Electron uncertainties mainly come from the reconstruction efficiency and calibrations, and are obtained from tag and probe measurement of $J/\psi, Z \to e^+e^-$ events [140]. This includes a total of around 120 NPs with up and down variations, corresponding to reconstruction, identification, isolation and trigger efficiency variations.

Muon uncertainties are derived in a similar way to the electron uncertainties, using $J/\psi/Z \to \mu^+\mu^-$ events as outlined in [146]. Further uncertainties include τ-lepton calibration uncertainties [238] and uncertainties on the soft $p_{\text{T}}^{\text{miss}}$ term [239].

The luminosity uncertainties for the full ATLAS Run-2 data-set summarised in Table 3.1 are included and lead to an additional luminosity uncertainty of 0.83% across all measured bins.

Uncertainties on the multijet background in the $p_{\text{T}}^{\text{miss}}$+jets region are subject to statistical variations due to the smearing method. The uncertainty on this background is taken to be 100% of the estimated yield. This has a small impact on the total uncertainty, because this background represent only a tiny fraction of events in the signal region. Fake lepton background uncertainties are dominated by theoretical uncertainties, transferred onto the fake lepton contributions via the MC estimates in the fake-enriched lepton regions. Further sources of uncertainty are statistical errors, the smoothing procedure and systematic uncertainties from the method used, estimated by varying selections of the fake-enriched regions.

5.4.3 Unfolding uncertainties

The unfolding procedure introduces sources of systematic uncertainty that are described in more detail in Section 5.6. These include the regularisation bias of the iterative Bayesian unfolding procedure, the impact of hidden variables, the SM sample composition of the MC that is used for the unfolding and the subtraction scheme used to account for interference terms in different top processes.

5.4.4 Theoretical uncertainties

The uncertainties on the modelling of SM processes with MC generators are predominantly introduced as scale variations. These are estimated using the prescription from Ref. [225]. In this method, the renormalisation scale μ_R and factorisation scale μ_F are varied by factors of 0.5 and 2.0 to estimate the uncertainty associated to missing higher-order QCD corrections. The envelope of in total seven factor-two rescalings,

$$\frac{(\mu_R, \mu_F)}{\mu_0} = (1,1),\ (2,2),\ (0.5,0.5),\ (2,1),\ (1,2),\ (0.5,1),\ (1,0.5), \tag{5.19}$$

where μ_0 indicates the nominal scale, is taken as the uncertainty. This pure-QCD uncertainty is labelled as $\delta^{(1)}K_{(\mathrm{N})\mathrm{NLO}}$. Constant scale variations affect the overall normalisation, but tend to underestimate shape uncertainties. For this reason, an additional uncertainty to account for possible differences in the shape of p_{T}^{V} and m_{jj} distributions in V+jets channels is included. These uncertainties, $\delta^{(2)}K_{(\mathrm{N})\mathrm{NLO}}$ and $\delta^{(4)}K_{(\mathrm{N})\mathrm{NLO}}$, are derived by supplementing the standard scale uncertainty $\delta^{(1)}K_{(\mathrm{N})\mathrm{NLO}}$ with a shape distortion

$$\omega_{\mathrm{shape}}(x) = \tanh\left[\ln\left(\frac{x}{x_0}\right)\right] = \frac{x^2 - x_0^2}{x^2 + x_0^2}, \tag{5.20}$$

where $x0$ is the midpoint of the variable (p_{T}^{V} or m_{jj}) in logarithmic scale. This modifies the standard scale uncertainty by a factor of $1 \leq \sqrt{1+\omega_{\mathrm{shape}}^2(x)} \leq \sqrt{2}$ [225]. These QCD uncertainties are taken to be correlated between different observable bins, but uncorrelated between different processes. The uncertainty of residual de-correlations between different processes are estimated in a conservative approach detailed in Ref. [225], and are indicated with $\delta^{(3)}K_{(\mathrm{N})\mathrm{NLO}}$. Similarly, truncation of the perturbative series of the EW coupling leads to pure EW uncertainties that contribute to the error bars of the SM prediction. An uncertainty related to non-factorised mixing of QCD and EW effects is included as δK_{mix}, estimated from the ansatz that this term is proportional to the factorised contribution of EW corrections [225].

Following the PDF4LHC prescription in Ref. [226], PDF uncertainties for the Hessian set are calculated as the squared difference corresponding to the predictions of the nominal set and the error sets. Combined PDF + α_s uncertainties are evaluated by shifting the strong coupling constant around its nominal value of $\alpha_s = 0.118$ with ± 0.001, calculating the "PDF only" uncertainty for the modified values of α_s and computing the final uncertainty from the residual differences.

Uncertainties on the inference between $t\bar{t}$ and tW production – as introduced in Section 5.3.4 – are evaluated by using two different SM predictions for tW during unfolding, both corresponding to a different diagram subtraction scheme.

Process	Region							
	$p_{\mathrm{T}}^{\mathrm{miss}}$+jets		1e+jets		2e+jets		1μ+jets		2μ+jets	
$Z \to \nu\nu$	1777796.0	(52.5%)	10.8		0.0		615.1		0.0	
$W \to e\nu$	222353.1	(6.6%)	668757.3	(69.1%)	288.4		79.1		0.0	
$Z \to ee$	0.7		1359.3		139013.8	(91.9%)	0.3		0.0	
$W \to \mu\nu$	322193.0	(9.5%)	46.9		0.0		1887747.7	(67.1%)	80.7	
$Z \to \mu\mu$	9448.2		1.7		0.0		48357.3	(1.7%)	197438.1	(92.8%)
$W \to \tau\nu$	678058.6	(20.0%)	53216.2	(5.5%)	21.1		202825.5	(7.2%)	7.1	
$Z \to \tau\tau$	7104.6		2359.2		135.7		13345.0		129.2	
Top	239772.5	(7.1%)	186560.6	(19.3%)	4021.2	(2.7%)	437914.8	(15.6%)	4623.2	(2.2%)
Fake	27597.3		21135.9	(2.2%)	1747.2	(1.2%)	126263.4	(4.5%)	2097.5	(1.0%)
EWK	40620.3	(1.2%)	14233.3	(1.5%)	2150.8	(1.4%)	38820.3	(1.4%)	2910.6	(1.4%)
Diboson	63318.2	(1.9%)	19468.9	(2.0%)	3908.9	(2.6%)	58663.5	(2.1%)	5371.1	(2.5%)
Triboson	6.1		3.2		1.8		16.1		2.6	
Total	3388268.66		967153.44		151289.03		2814648.22		212660.05	

Table 5.10: Event yields for the ≥ 1 jet region for different SM processes and the instrumental detector background indicated as "Fake". The fractional contribution to the total yield is indicated for the processes with a contribution larger than 1%. These numbers are obtained from a simulated MC event sample that is reweighted according to the m_{jj} reweighting procedure as detailed in Section 5.3.6.

5.5 Detector-level results

In this section, comparison between reconstruction-level data and the SM predictions are shown for the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region and the 1e+jets, 2e+jets, 1μ+jets and 2μ+jets auxiliary measurements. The different contributions to the total SM prediction in the fiducial phase-space is indicated as a stacked histogram. The figures also show the contribution from instrumental (fake) backgrounds described in Section 5.3.8.

In each plot a ratio of data to MC is shown. Black markers show the statistical uncertainty on the data and hatched blue region shows the total systematic uncertainties and statistical uncertainties added in quadrature.

Observables for the ≥ 1 jet and VBF regions are presented in Section 5.5.1 and Section 5.5.2 respectively.

5.5.1 ≥ 1 jet region

Comparisons between experimental data and MC for the $p_{\mathrm{T}}^{\mathrm{miss}}$ observables measured in the ≥ 1 jet topology are shown in Figure 5.10, with the corresponding total event yields summarised in Table 5.10.

Every region has a different, dominant QCD V+jets process that corresponds to the lepton content of the final state. The differential histograms and event yield obtained using the MC samples from Table 5.4, corresponding to the Sherpa 2.2.1 QCD V+jets samples. "Top" combines contributions from single-top DS and $t\bar{t}$ processes. Small components that contribute less than 1% of the total event yield are grouped into the "Other" category. "Fakes" refers to the fake, instrumental background for each measurement region, where the estimation techniques have been described in Section 5.3.8.

The shape is reasonably well-modelled across the spectrum of $p_{\mathrm{T}}^{\mathrm{recoil}}$, but there is a constant offset varying between 10% to 20% where MC simulation under-predicts the event yield. This is understood as a MC mismodelling of the total cross-section, which may be improved by including

higher order diagrams in QCD and EW in the cross-section calculation. The constant offset does not impact the robustness of the unfolding: it is more important that the shape is well-modelled and that the response of the detector is accurately represented. In Section 5.6.3 we assess the impact of the discrepancy between data and unfolding MC on the final results. Any residual bias after two iterations of the iterative unfolding procedure is added to the final unfolded spectra as a systematic uncertainty.

Since the SM predictions underestimate the experimentally measured data, it is tempting to test the compatibility of these observables with BSM signals. However, here it is important to reiterate that these MC predictions are only used to model the detector response and not to assess the compatibility with the SM. In Section 5.8 we quantitatively assess the agreement between data and the state-of-the-art MC predictions at particle-level, including a discussion on the compatibility with two different BSM scenarios.

A breakdown of the systematic uncertainties is shown in Figure 5.11. Only sources that contribute more than 1% in at least one of the measurement bins are displayed. For the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region, JES and statistical uncertainties are the dominant contributions at low and high $p_{\mathrm{T}}^{\mathrm{miss}}$ respectively. Because the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ contribution is very tiny, as can be seen in Table 5.10, even the conservative 100% uncertainty on the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ estimate does not contribute much to the overall uncertainty.

In the single-lepton channels, uncertainties associated to the fake lepton background play a more important role, and it is in fact the dominant source of uncertainty at low $p_{\mathrm{T}}^{\mathrm{recoil}}$ in the 1μ+jets region. The "Top variation" uncertainty corresponds to using a different subtraction method to account for overlap in single-top event generation. This effect is most pronounced in the high $p_{\mathrm{T}}^{\mathrm{recoil}}$ tails of the single-lepton regions. Section 5.6.4 describes how the top-variation uncertainty is propagated into the final, unfolded measurement.

In the di-lepton regions, dominant uncertainties are the JES, statistical uncertainties and uncertainties associated to the lepton reconstruction efficiencies.

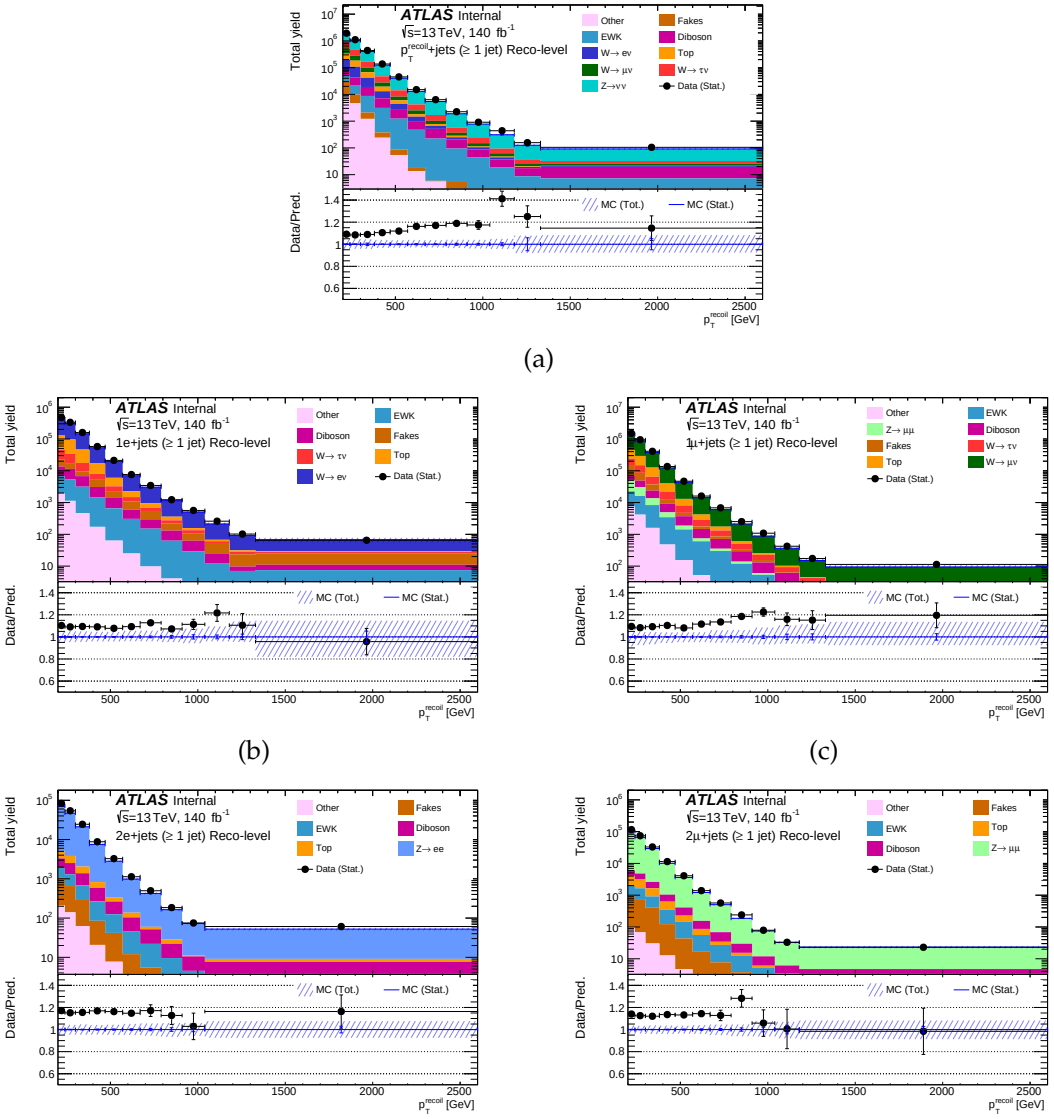


Figure 5.10: Reconstruction-level comparisons between data and MC for $p_{\mathrm{T}}^{\mathrm{miss}}$ in the $\geq$ 1 jet phase space. Different components of the SM predictions are included in the stacked histograms and experimental data is indicated as a black marker including the statistical uncertainty. The bottom panel shows the ratio to the total SM prediction, blue hatched regions mark the combined systematic and statistical uncertainty for the SM prediction. The m_{jj} reweighting is only applied to events in the VBF region, hence no reweighting is applied to the MC sample shown in this figure. The event yield per bin drops multiple orders of magnitude as a function of $p_{\mathrm{T}}^{\mathrm{recoil}}$. The leading-order MC prediction models this shape of the data well across the $p_{\mathrm{T}}^{\mathrm{recoil}}$ spectrum in the five different regions, but there is an overall offset between 10-20% in the total cross-section.

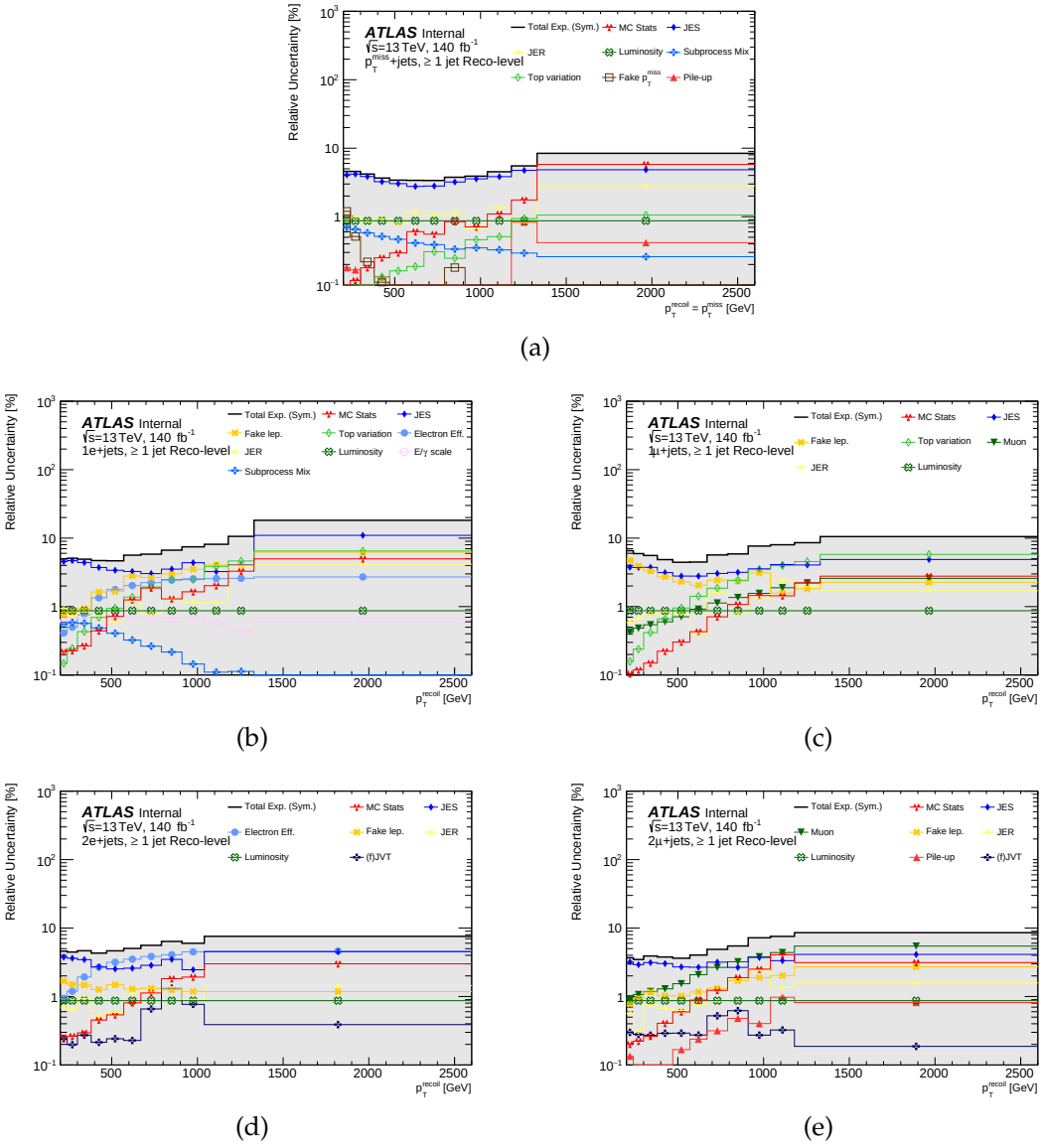


Figure 5.11: Relative contribution from different systematic uncertainty components at reconstruction-level for $p_{\mathrm{T}}^{\mathrm{recoil}}$ in the $\geq$ 1 jet phase space. Only systematics with a contribution of $\geq 1\%$ in at least one of the bins are shown in the figure. For low values of $p_{\mathrm{T}}^{\mathrm{recoil}}$, the jet energy scale (JES) uncertainties dominate, where statistical uncertainties become dominant at large $p_{\mathrm{T}}^{\mathrm{recoil}}$. For illustrative purposes this figure shows the symmetrised uncertainties, calculated as the average of the asymmetric error in each bin. Total Exp. (Sym.) is the combination of statistical and systematic uncertainties and indicates the symmetrised total experimental uncertainty. In subfigure (a), the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty is largest at low $p_{\mathrm{T}}^{\mathrm{miss}}$ bins because that is where the fake multijet background is largest and the uncertainty is taken to be 100% of the fake yield. The single-lepton regions in subfigure (b) and (c) have a large top background, which has an additional modelling uncertainty in light green, which is discussed in Section 5.6.4. Due to the selection of two electron or two muons, the electron and muon uncertainties are large in the di-lepton channels shown in figure (d) and (e) respectively.

Process	Region									
	$p_{\mathrm{T}}^{\mathrm{miss}}$+jets		1e+jets		2e+jets		1μ+jets		2μ+jets	
$Z \to \nu\nu$	292547.1	(49.9%)	0.5		0.0		117.2		0.0	
$W \to e\nu$	40.49.7	(7.1%)	108840.1	(72.9%)	45.0		11.2		0.0	
$Z \to ee$	1.3		305.9		22445.5	(87.5%)	0.2		0.0	
$W \to \mu\nu$	59352.6	(10.1%)	5.7		0.0		312026.6	(69.4%)	12.7	
$Z \to \mu\mu$	1777.2		0.2		0.0		8399.0	(1.9%)	32142.2	(88.4%)
$W \to \tau\nu$	123538.8	(21.0%)	9062.1	(6.1%)	3.5		33496.3	(7.4%)	0.0	
$Z \to \tau\tau$	1236.7		401.8		21.3		2150.5		21.3	
Top	21484.9	(3.7%)	14528.5	(9.7%)	784.8	(3.1%)	37349.8	(8.3%)	1011.3	(2.8%)
Fake	10775.3	(1.8%)	4051.9	(2.7%)	265.9	(1.0%)	21855.1	(4.9%)	327.4	
EWK	26645.5	(4.5%)	9175.0	(6.1%)	1480.2	(5.8%)	25406.7	(5.6%)	2910.6	(1.4%)
Diboson	9319.7	(1.6%)	2881.7	(1.9%)	606.9	(2.4%)	8952.1	(2.0%)	827.1	(2.3%)
Triboson	1.3		0.8		0.4		3.3		0.5	
Total	587627.02		149254.21		25653.44		449767.90		36344.52	

Table 5.11: Event yields for the VBF region for different SM processes and the instrumental detector background indicated as "Fake". The fractional contribution to the total yield is indicated for the processes with a contribution larger than 1%.

5.5.2 VBF region

$p_{\mathrm{T}}^{\mathrm{miss}}$

Comparisons between experimental data and MC for the $p_{\mathrm{T}}^{\mathrm{recoil}}$ observables measured in the VBF topology are shown in Figure 5.12. The total event yields for the VBF region are presented in Table 5.11.

The fiducial cross-section for the VBF region is generally smaller than the ≥ 1 jet region due to stricter event-selections. Therefore, the $p_{\mathrm{T}}^{\mathrm{recoil}}$ distributions in the VBF region have coarser binning towards the tail to populate every bin with a sufficient number of events. Similarly to the ≥ 1 jet region, the overall shape is reasonably well-modelled and the data has an offset of 10–20% with respect to the modelled backgrounds.

Compared to the ≥ 1 jet region, Table 5.11 show a larger relative contribution from electroweak V+dijet processes. As can be seen from the differential distribution in Figure 5.12 this is more pronounced for larger values of $p_{\mathrm{T}}^{\mathrm{recoil}}$.

A breakdown of the systematic uncertainties is shown in Figure 5.13. Features similar to the ≥ 1 jet uncertainty plots are present: JES dominates in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region, where the single-lepton and di-lepton regions have additional, large uncertainties from the fake-lepton estimation and lepton efficiencies respectively.

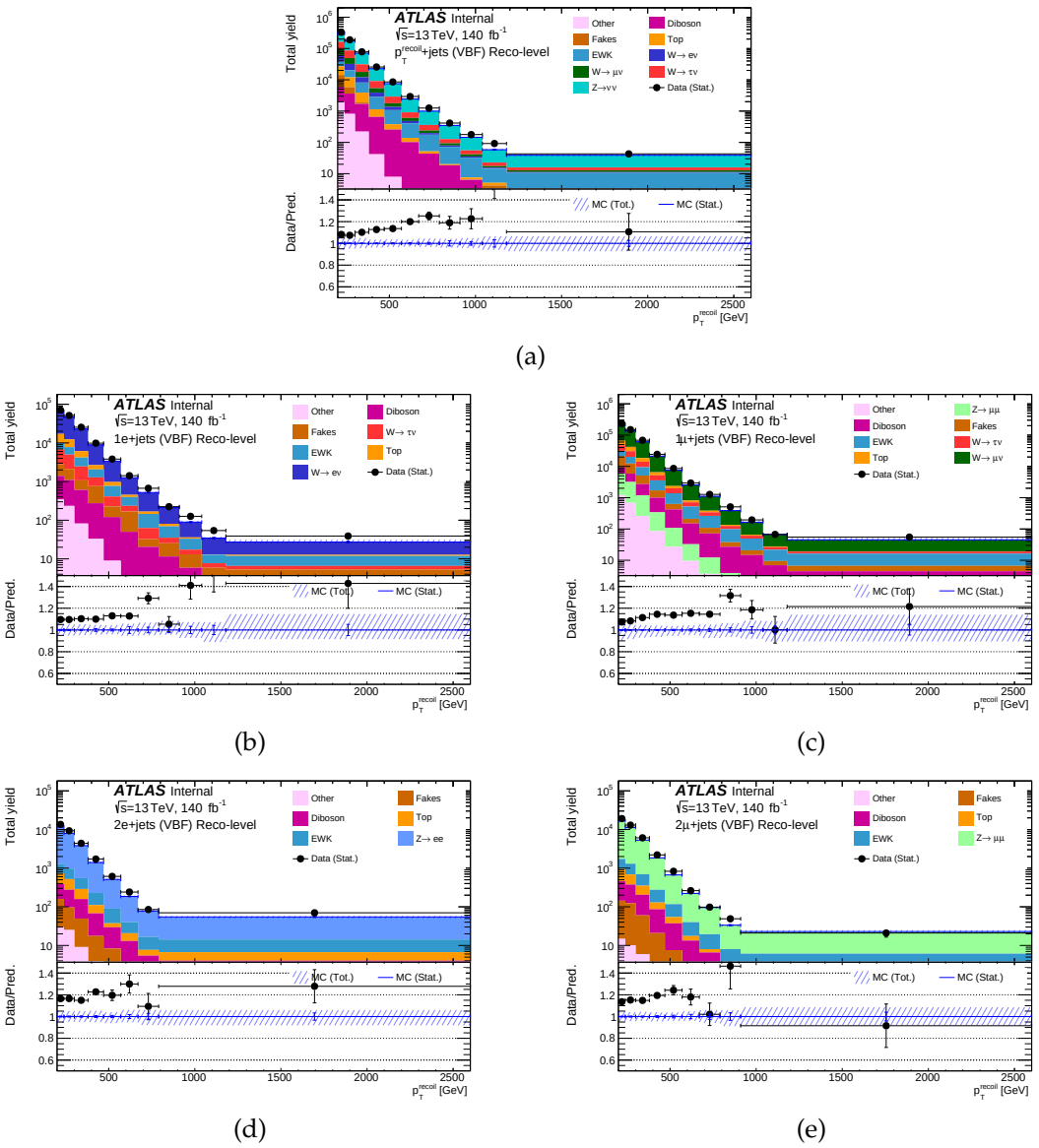


Figure 5.12: Reconstruction-level comparisons between data and MC for $p_{\mathrm{T}}^{\mathrm{miss}}$ in the VBF phase space. Different components of the SM predictions are included in the stacked histograms and experimental data is indicated as a black marker including the statistical uncertainty. The bottom panel shows the ratio to the total SM prediction, blue hatched regions mark the combined systematic and statistical uncertainty for the SM prediction. The m_{jj} reweighting discussed in Section 5.3.6 is applied to the MC sample shown in this figure. Similar to the ≥ 1 jet region, the MC models the shape of the data well across the $p_{\mathrm{T}}^{\mathrm{recoil}}$ spectrum in the different regions, but there is an overall offset between 10-20% in the total cross-section.

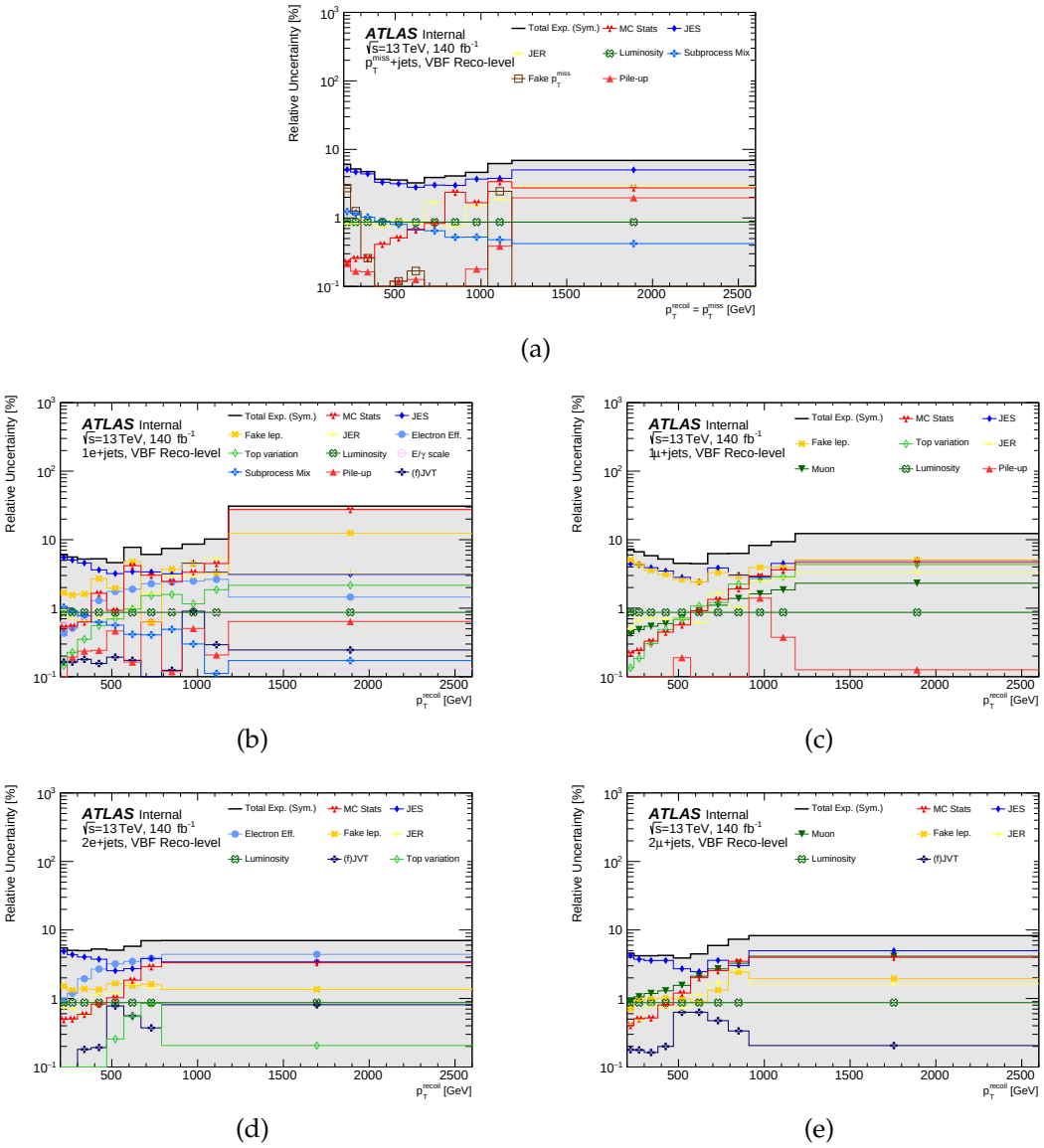


Figure 5.13: Relative contribution from different systematic uncertainty components at reconstruction-level for $p_{\mathrm{T}}^{\mathrm{recoil}}$ in the VBF phase space. Only systematics with a contribution of $\geq 1\%$ in at least one of the bins are shown in the figure. For low values of $p_{\mathrm{T}}^{\mathrm{recoil}}$, the jet energy scale (JES) uncertainties dominate, where statistical uncertainties become dominant at large $p_{\mathrm{T}}^{\mathrm{recoil}}$. For illustrative purposes this figure shows the symmetrised uncertainties, calculated as the average of the asymmetric error in each bin. Total Exp. (Sym.) is the combination of statistical and systematic uncertainties and indicates the symmetrised total experimental uncertainty. Subfigure (a), like the ≥ 1 jet region, shows largest uncertainties on the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ background in low $p_{\mathrm{T}}^{\mathrm{miss}}$ bins, but there are some events at approximately $1000\,\mathrm{GeV}$ where the total MC cross-section is relatively low, and hence the 100% uncertainty on the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ yield leads to a relatively large uncertainty here. The single-lepton channels in figure (b) and (c) are prone to fake lepton contributions across the $p_{\mathrm{T}}^{\mathrm{recoil}}$ spectrum, and also to the modelling of the top-background particularly at high $p_{\mathrm{T}}^{\mathrm{recoil}}$. Similar to the ≥ 1 jet region, the di-lepton channels in subfigures (d) and (e) have relatively large contributions from electron and muon efficiency uncertainties as the phase-space requires the presence of two leptons.

m_{jj}

Comparisons between experimental data and MC for the m_{jj} observables measured in the VBF topology are shown in Figure 5.14.

The QCD V+jets events populating these distributions have been reweighted following the procedure outlined in Section 5.3.6 to improve the modelling of the experimental data. To keep the cross-section constant, the histograms are renormalised to the original cross-section of the unreweighted m_{jj} distributions. This significantly improves the shape of the MC distributions, that now give a reasonably flat ratio with respect to the data. The resulting shape is not 100% flat due to the additional non-QCD V+jets backgrounds that not have been reweighted, and the MC under-predicts slightly towards lower values of m_{jj}.

A breakdown of the systematic uncertainties is shown in Figure 5.15. The high m_{jj} tails have reasonably good MC statistics and therefore the JES is the dominant source of uncertainty across the spectrum, except for the 2μ+jets region where the JER is dominant in the highest m_{jj} bin.

$\Delta\phi_{jj}$

Comparisons between experimental data and MC for the $\Delta\phi_{jj}$ observables measured in the VBF topology are shown in Figure 5.16.

Jets have a cone size of $\Delta R = 0.4$ which limits how closely they can be aligned in the detector and still be resolved as separate jets. This slightly reduces the phase-space for small values ($\Delta\phi_{jj} \approx 0$) which is observed as a slightly lower event yield in this region. The relatively small number of events at $|\Delta\phi_{jj}| \approx 1$ can be understood qualitatively from the $\Delta\phi\left(\text{jet}, p_{\text{T}}^{\text{miss}}\right) > 0.4$ cut. At leading order, when exactly two jets and one vector boson are produced to match most event selection criteria, momentum conservation requires the vector boson to be aligned with one of the two jets, which is then rejected as $\Delta\phi\left(\text{jet}, p_{\text{T}}^{\text{miss}}\right) < 0.4$.

The leading contribution to the multijet events, used to estimate the fake backgrounds, are well-measured dijet events with additional lower p_{T} jets. Due to momentum conservation, the two leading jets are produced roughly back-to-back with $\Delta\phi_{jj} \approx \pm 1$. If a spurious signal appears elsewhere in the detector, this can be recorded as a fake $p_{\text{T}}^{\text{miss}}$ signature that is not aligned with any of the leading jets. These type of events escape the $\Delta\phi\left(\text{jet}, p_{\text{T}}^{\text{miss}}\right) > 0.4$ cut and populate the histogram bins mostly at $\Delta\phi_{jj} \approx \pm 1$. When the fake $p_{\text{T}}^{\text{miss}}$ signature comes from the mismeasurement of one of the leading jets, it will be vetoed due to the $\Delta\phi\left(\text{jet}, p_{\text{T}}^{\text{miss}}\right) > 0.4$ requirement. Therefore the fake contribution does not populate the central bins at values of $\Delta\phi_{jj}$ close to zero.

A breakdown of the systematic uncertainties is shown in Figure 5.17. The jet energy scale uncertainties are the dominant source across most bins. Because the uncertainty of the fake $p_{\text{T}}^{\text{miss}}$ background is taken to be 100% of the yield, its relative contribution is largest in bins with large fake $p_{\text{T}}^{\text{miss}}$. This becomes the dominant source of uncertainty at $\Delta\phi_{jj} \approx \pm 1$ in the signal region.

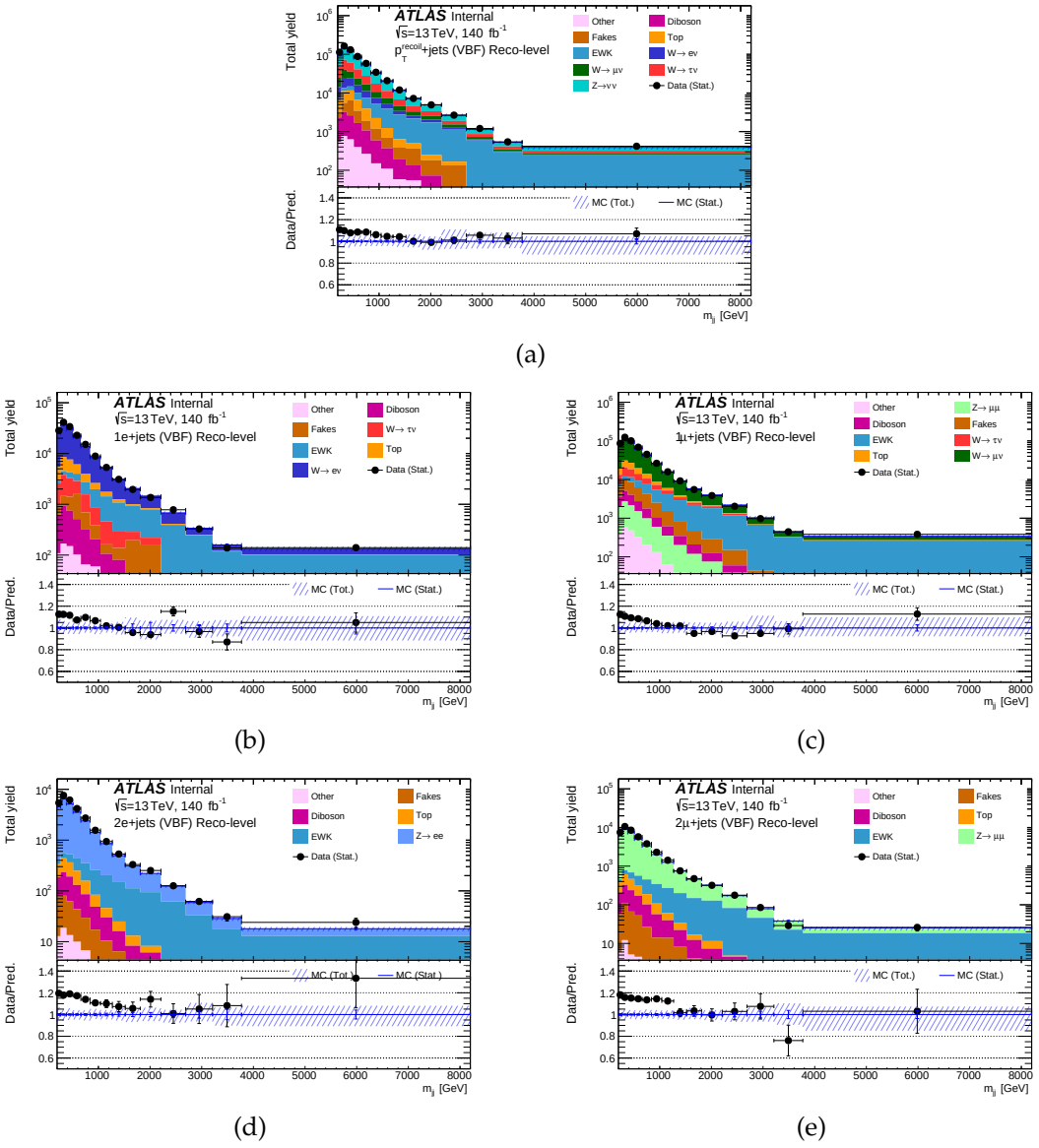


Figure 5.14: Reconstruction-level comparisons between data and MC for m_{jj} in the VBF phase space. Different components of the SM predictions are included in the stacked histograms and experimental data is indicated as a black marker including the statistical uncertainty. The bottom panel shows the ratio to the total SM prediction, blue hatched regions mark the combined systematic and statistical uncertainty for the SM prediction. The m_{jj} reweighting discussed in Section 5.3.6 is applied to the MC sample shown in this figure, leading to a relative flat spectrum that also partly mitigates the overall cross-section offset observed in the $p_{\mathrm{T}}^{\mathrm{recoil}}$ distributions.

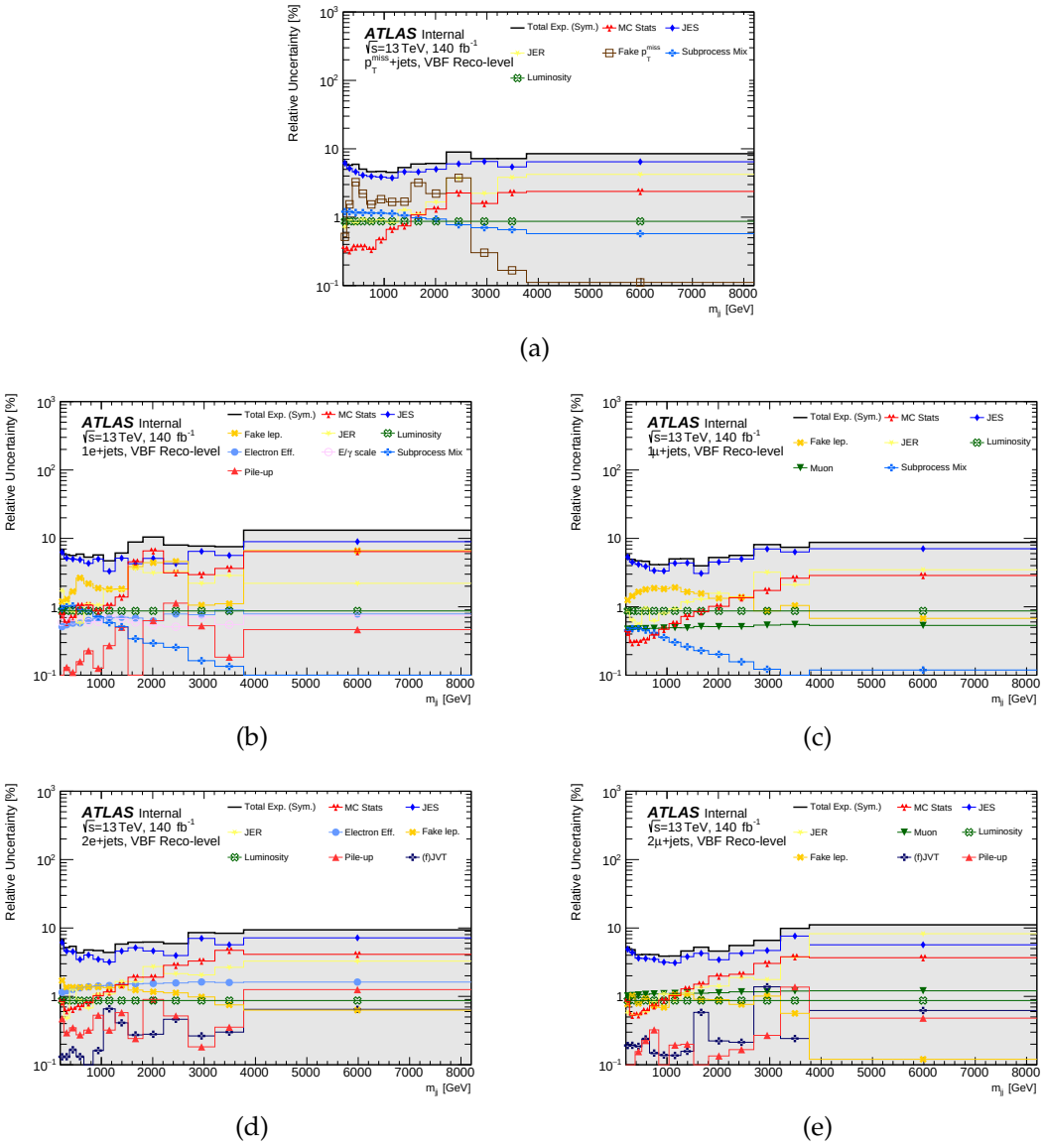


Figure 5.15: Relative contribution from different systematic uncertainty components at reconstruction-level for m_{jj} in the VBF phase space. Only systematics with a contribution of $\geq 1\%$ in at least one of the bins are shown in the figure. The jet energy scale (JES) uncertainties are the main source of uncertainty. For illustrative purposes this figure shows the symmetrised uncertainties, calculated as the average of the asymmetric error in each bin. Total Exp. (Sym.) is the combination of statistical and systematic uncertainties and indicates the symmetrised total experimental uncertainty. The fake $p_{\mathrm{T}}^{\mathrm{miss}}$ background in figure (a) becomes increasingly smaller at large m_{jj} leading to a reduced associated uncertainty.

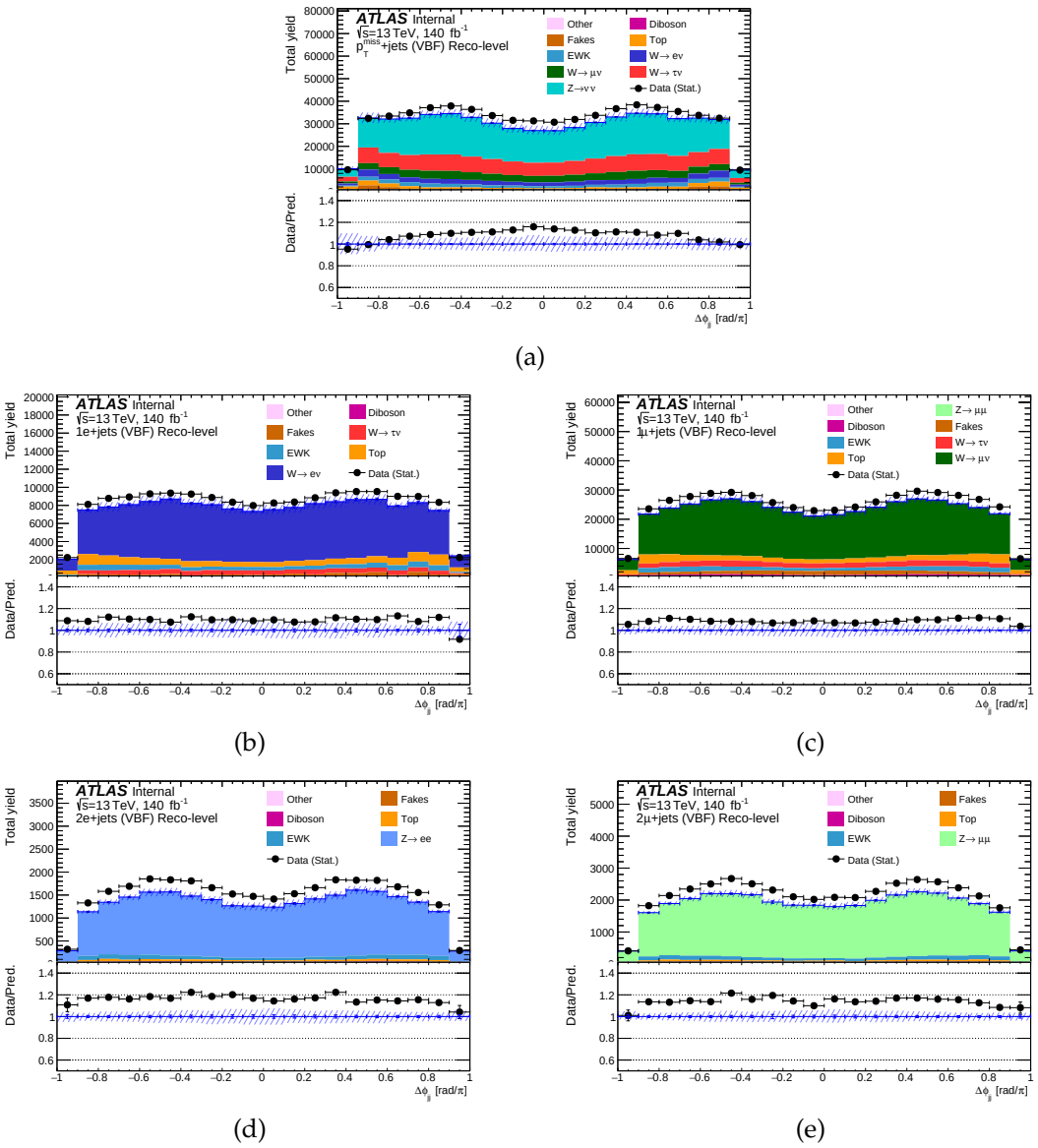


Figure 5.16: Reconstruction-level comparisons between data and MC for $\Delta\phi_{jj}$ in the VBF phase space. Different components of the SM predictions are included in the stacked histograms and experimental data is indicated as a black marker including the statistical uncertainty. The bottom panel shows the ratio to the total SM prediction, blue hatched regions mark the combined systematic and statistical uncertainty for the SM prediction. The m_{jj} reweighting discussed in Section 5.3.6 is applied to the MC sample shown in this figure.

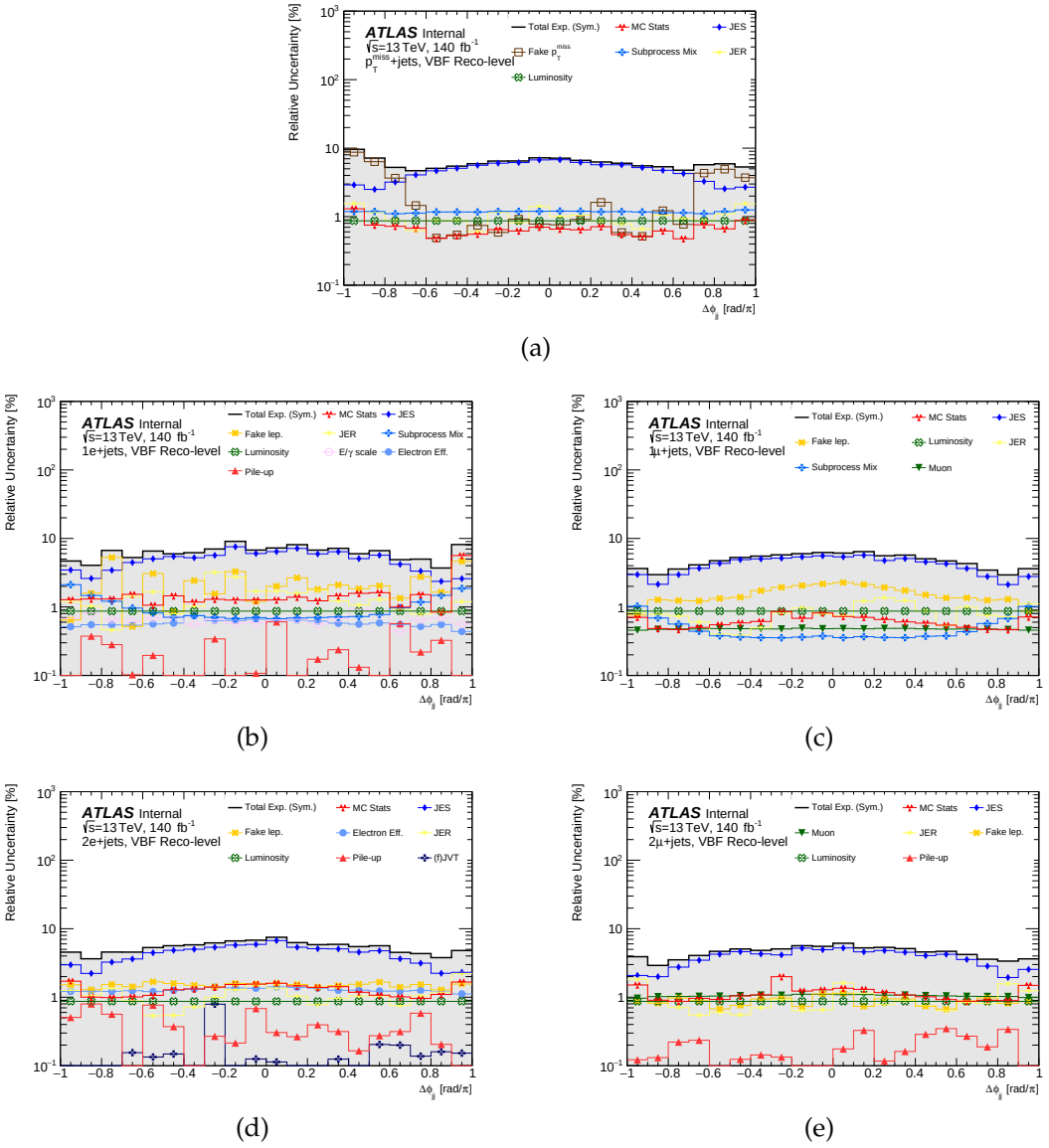


Figure 5.17: Relative contribution from different systematic uncertainty components at reconstruction-level for $\Delta\phi_{jj}$ in the VBF phase space. Only systematics with a contribution of $\geq 1\%$ in at least one of the bins are shown in the figure. For illustrative purposes this figure shows the symmetrised uncertainties, calculated as the average of the asymmetric error in each bin. Total Exp. (Sym.) is the combination of statistical and systematic uncertainties and indicates the symmetrised total experimental uncertainty. The fake $p_{\mathrm{T}}^{\mathrm{miss}}$ background is largest around $|\Delta\phi_{jj}| \approx 1$, as in this region the event topology is such that there is a larger chance that the mismeasurement of a jet leads to a fake imbalance in the transverse plane. As the uncertainty is proportional to the yield and this region has a low cross-section due to the cut on $\Delta\phi_{jj}$ between jets and $p_{\mathrm{T}}^{\mathrm{miss}}$, the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty dominates in this region.

5.6 Unfolding

The detector-level results in Section 5.5 are corrected for detector effects to present them in terms of particle level objects as defined in Section 5.2.3. Two complementary measurements are defined: one in terms of the final state (the "target final state" or "inclusive measurement") and one in terms of the dominant boson process. This thesis considers only the unfolding of the former measurement, where all processes in Section 5.5 are considered to be signal, except for the fake contributions which are subtracted from the experimental data.

The iterative Bayesian unfolding technique is used to correct data for detector inefficiencies and resolutions, and to account for bin migrations between detector-level and particle-level. The prior used in this iterative procedure is taken as the particle-level MC predictions. Further details on the unfolding technique are presented in Section 5.6.1, together with the inputs needed to perform the unfolding. Results of closure tests and optimisation studies for the number of unfolding iterations are given in Section 5.6.2 and 5.6.3 respectively. The robustness of the unfolding procedure is scrutinised further via SM process composition variations, BSM signal injection tests and tests of bin migrations in Section 5.6.4 to 5.6.6. This analysis uses the VIPUnfolding framework [240] as the practical implementation of the unfolding, which is an interface to the RooUnfold package [181].

5.6.1 Method and diagnostic plots

In addition to detector-level data, the iterative Bayesian unfolding method, see Section 4.4.2, uses the following simulation-based quantities to correct data for detector effects.

As the name suggests, the **migration matrix** accounts for events in a specific bin in a particle-level distribution that can migrate to a different, neighbouring bin in the measured distribution. This happens when the measured variable is reconstructed with a higher or lower value than its true value. The migration matrix is filled by selecting events that pass the particle-level selections and tracking their respective bins, and then selecting the corresponding detector-level bin where the event is reconstructed. The entries are normalised by total number of particle-level events in a bin to convert to a probability that an event in particle-level bin i is reconstructed in detector-level bin j. This normalisation is chosen to give intuition how particle-level events 'distribute' when they are reconstructed at detector-level, and ultimately with the unfolding we want to invert this probability to see which bin particle-level events from experimental data – in a given bin – would have originally originated from.

The migration matrices for all measured observables in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region are presented in Figure 5.18. Corresponding plots for the single-lepton and di-lepton channels are shown in Appendix B, and show similar trends to the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region.

When calculating observables involving angles between jets, such as $\Delta\phi_{jj}$, ambiguities can occur. Since it is defined as the difference in ϕ of the leading and sub-leading jet, a sign-flip can occur if the leading jet and sub-leading jet change order at the detector-level. Therefore we require both the leading and sub-leading jets to be uniquely matched between detector and particle level within $\Delta R < 0.1$. The small off-diagonal components in the corners of the $\Delta\phi_{jj}$ migration matrices are the result of a sign flip when $|\Delta\phi_{jj}| \approx 1$.

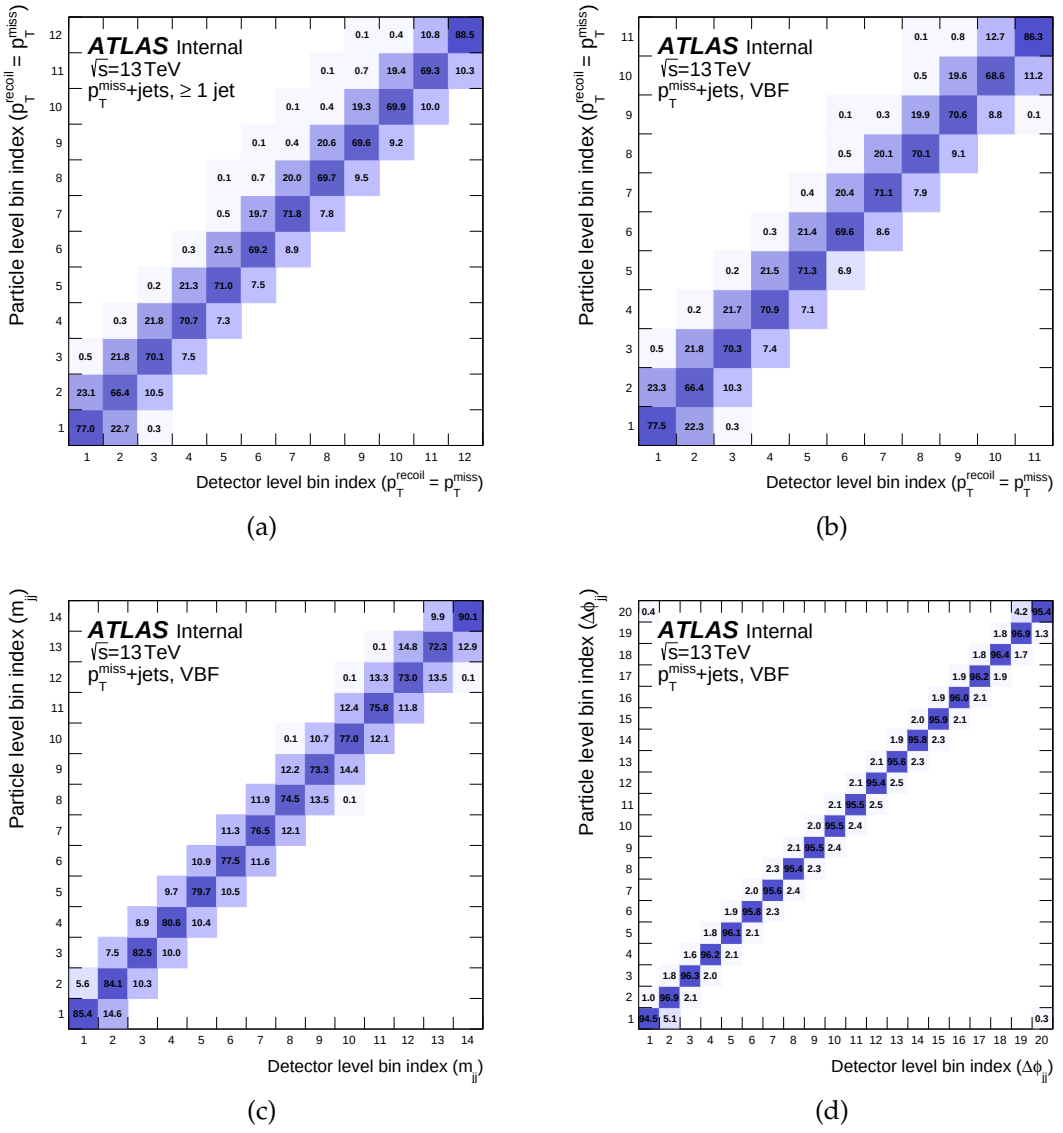


Figure 5.18: Migration matrices for the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region in the measured observables for the ≥ 1 jet and VBF phase space. The corresponding matrices for the single-lepton and di-lepton channels are presented in Appendix B. In all plots, the diagonal entries (i.e. the purity) is observed to be at least 60%, which is found to lead to a reliable convergence of the unfolding within a small number of iterations such that the statistical uncertainties from the unfolding procedure are reduced. For a given bin in the $p_{\mathrm{T}}^{\mathrm{miss}}$ observable in figure (a) and (b), there is more migration to higher $p_{\mathrm{T}}^{\mathrm{miss}}$ bins than to lower $p_{\mathrm{T}}^{\mathrm{miss}}$ bins. This is explained by the fact that these are steeply falling distributions and that reconstruction of the magnitude of a vector, i.e. $p_{\mathrm{T}}^{\mathrm{miss}}$, at a slightly different value leads to relatively more migrations to high $p_{\mathrm{T}}^{\mathrm{miss}}$ bins. The m_{jj} distributions fall slightly less steep and are reconstructed from a combination of two four-momenta, which mitigates this effect and hence the off-diagonal elements in subfigure (c) indicate that events migrate to either a higher or lower m_{jj} bin with approximately equal probability. The $\Delta\phi_{jj}$ migration matrix in subfigure (d) has a very high purity, indicating that angular information on jets is retained very accurately between particle-level and detector-level. Small migrations between bin index 1 and 20, corresponding to $\Delta\phi_{jj} = -1$ and $\Delta\phi_{jj} = 1$ respectively, occur when one of the two jets flips sign in $\Delta\phi_{jj}$ between particle-level and detector-level.

The diagonal elements of the migration matrices indicate probabilities that an event is reconstructed at the same bin at detector-level. **Purity** is defined as the fraction of entries in a detector-level bin that are in the same bin at particle-level. This quantity is calculated by taking the ratio of the number of diagonal elements to the number of events on the detector-level projection of the migration matrix. Conceptually, the purity will only be 100% if all events in a detector-level bin originate from the same particle-level bin. It does not drop below 100% if events from that particle-level bin drifted off to another detector-level bin, as long as no events from other particle-level bins seep in. The binning in this analysis is optimised to have a purity of at least 60% in every bin with at least 20 reconstructed events. This ensures that the binning does not introduce too many migrations between detector-level and particle-level due to resolution effects. Unfolding is inherently a statistical procedure, so that a minimum number of events in each bin leads to more robust unfolding and it reduces the final statistical uncertainty on the measurement.

The purity for the measured observables is given in Figure 5.19. The purity is highest for the $\Delta\phi_{jj}$ observable, indicating that jets are largely constructed at the same radial angle between detector-level and particle-level. The purity fluctuates around 0.7 (70%) for $p_{\mathrm{T}}^{\mathrm{recoil}}$ up to the second-to-last bin. This quantity is more sensitive to bin migrations because it is constructed from the sum of transverse momenta of all measured particles and the boson p_{T}. Similarly, a non-negligible amount of bin migration is observed in the m_{jj} spectrum which inherently depends of the reconstructed momenta of the two leading jets.

The **matching efficiency** is defined as the fraction of events passing the particle-level cuts that also pass the detector-level selections. This accounts for the efficiency and acceptance of the detector, resulting in particle-level events that are not reconstructed and therefore not included in the measured (detector-level) distribution. It can be calculated directly as the number of events which pass both the particle and detector-level selections divided by the number of events passing the particle-level selection, as a function of the particle-level value of the variable being considered. The reconstruction matching efficiency plots are shown in Figure 5.20.

The efficiency is lowest for the $1e$+jets region due to cuts on the real-$p_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse mass of the W. The $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region, with no leptons to reconstruct, has the highest efficiency. This pattern is seen across all the different distributions.

Where the matching efficiency corrects for events that are "missed" by the detector, the opposite can also happen. If an event fails the particle-level selections, it can still pass the detector-level selection due to the finite detector resolution. This is quantified by the **fiducial fraction**, which is calculated as the number of events which pass both the particle and detector-level selections divided by the number of events passing the detector-level selection, as a function of the detector-level value of the variable being considered. Plots for the fiducial fraction are illustrated in Figure 5.21.

The $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region suffers from events where the W decays leptonically. If they fall within the fiducial (particle-level) phase space definitions, the particle-lepton veto rejects this event. However, due to lepton reconstruction efficiencies at the detector-level, these event can escape the detector-level lepton veto and be included in the detector-level measurement. The $1e$+jets region has a real $p_{\mathrm{T}}^{\mathrm{miss}}$ cut of $60\,\mathrm{GeV}$. Since $p_{\mathrm{T}}^{\mathrm{miss}}$ is a sum over all visible particles in the event, this variable is very dependent on detector resolution and can vary between particle and detector level, which offers an explanation for the low fiducial fraction in this region.

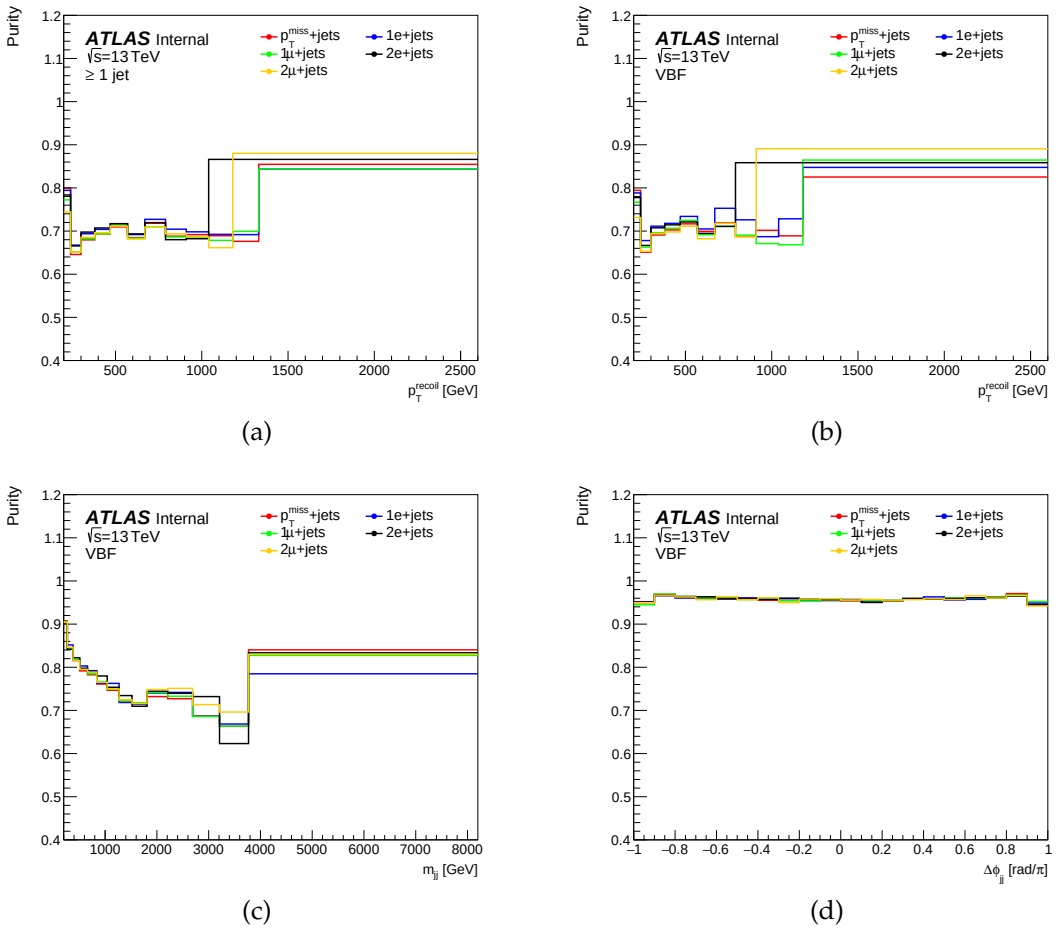


Figure 5.19: These histograms show the purity in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region and auxiliary regions for the unfolded observables, which is at least 60% in every bin. The purity is very similar for the signal region and auxiliary regions because they have largely the same binning, and kinematic selections for the lepton regions are defined to match the signal region as closely as possible. The highest $p_{\mathrm{T}}^{\mathrm{recoil}}$ and m_{jj} bins in figures (a), (b) and (c) have high purity because the binning is very coarse and inclusive, i.e. if some event would be reconstructed in at a higher value at detector-level, it would not migrate to a higher $p_{\mathrm{T}}^{\mathrm{recoil}}$ or m_{jj} bin because there is none. Similarly, the lowest $p_{\mathrm{T}}^{\mathrm{recoil}}$ and m_{jj} bin have relative high purities because migrations values below the lower bin edge lead to migrations out phase space.

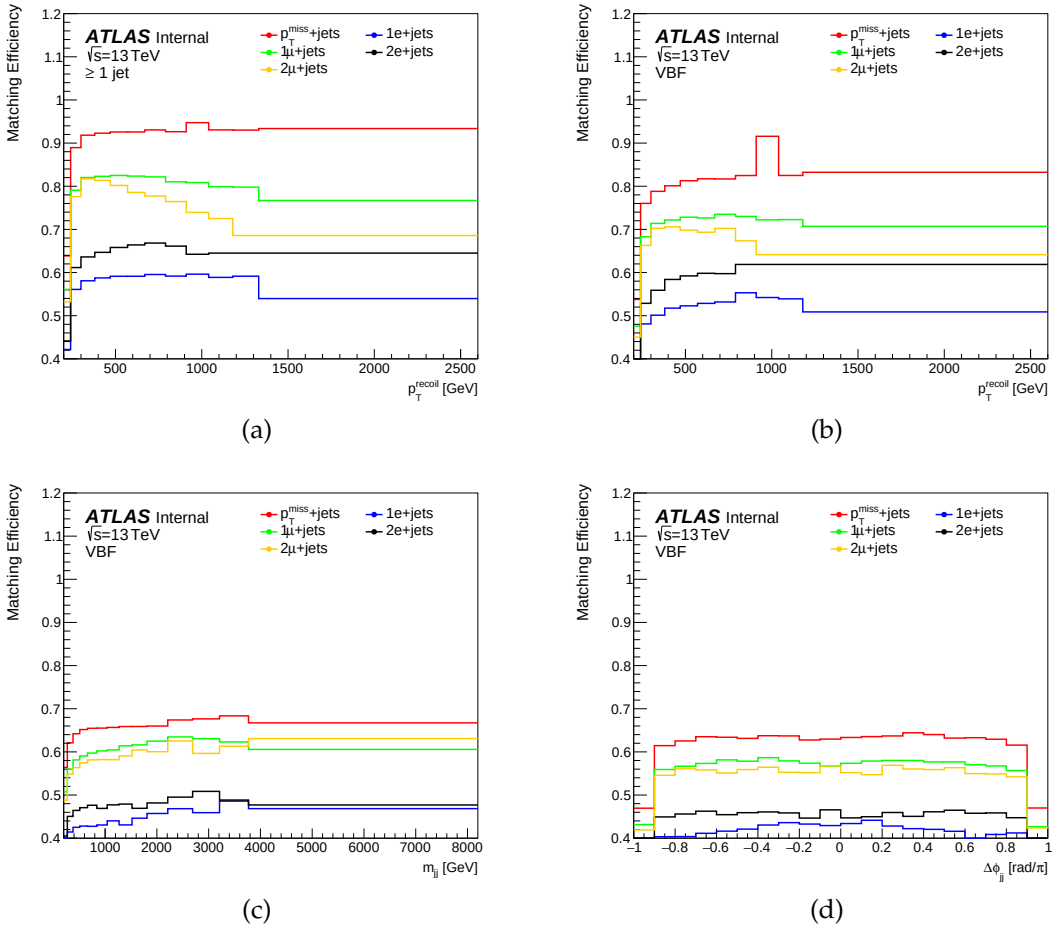


Figure 5.20: Matching efficiency for the observables that are unfolded in this measurement. The $p_{\mathrm{T}}^{\mathrm{recoil}}$ observables in (a) and (b), as well as the m_{jj} observable in (c) have a the lowest efficiency for the lowest $p_{\mathrm{T}}^{\mathrm{recoil}}$ or m_{jj} bin respectively. Particle-level events in this bin that are reconstructed below the minimum value of 200 GeV cannot migrate to a lower bin but instead fail the minimum $p_{\mathrm{T}}^{\mathrm{recoil}}$ or m_{jj} requirement and are hence not matched between particle and detector-level. As explained in the text, the efficiency is lowest for $1e$+jets due to cuts on real $p_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse mass of the W. The reconstruction efficiency is slightly better for $2e$+jets. At low $p_{\mathrm{T}}^{\mathrm{recoil}}$ the efficiency is similar between 1μ+jets and 2μ+jets but due to detector resolution high p_{T} muon events are lost, which is more pronounced in the region that has two muons to reconstruct. The efficiency in all observables is largest for the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region which has no leptons to reconstruct. As $p_{\mathrm{T}}^{\mathrm{recoil}}$ increases there is a non-trivial interplay between having larger bin widths but a reduced number of events, that can lead to fluctuations in the efficiency. This is apparent in subfigure (b) as a fluctuation in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets SR efficiency at approximately $p_{\mathrm{T}}^{\mathrm{recoil}} = 1000\,\mathrm{GeV}$, which is a fluctuation of less than 10% and therefore has no significant impact on the unfolded results.

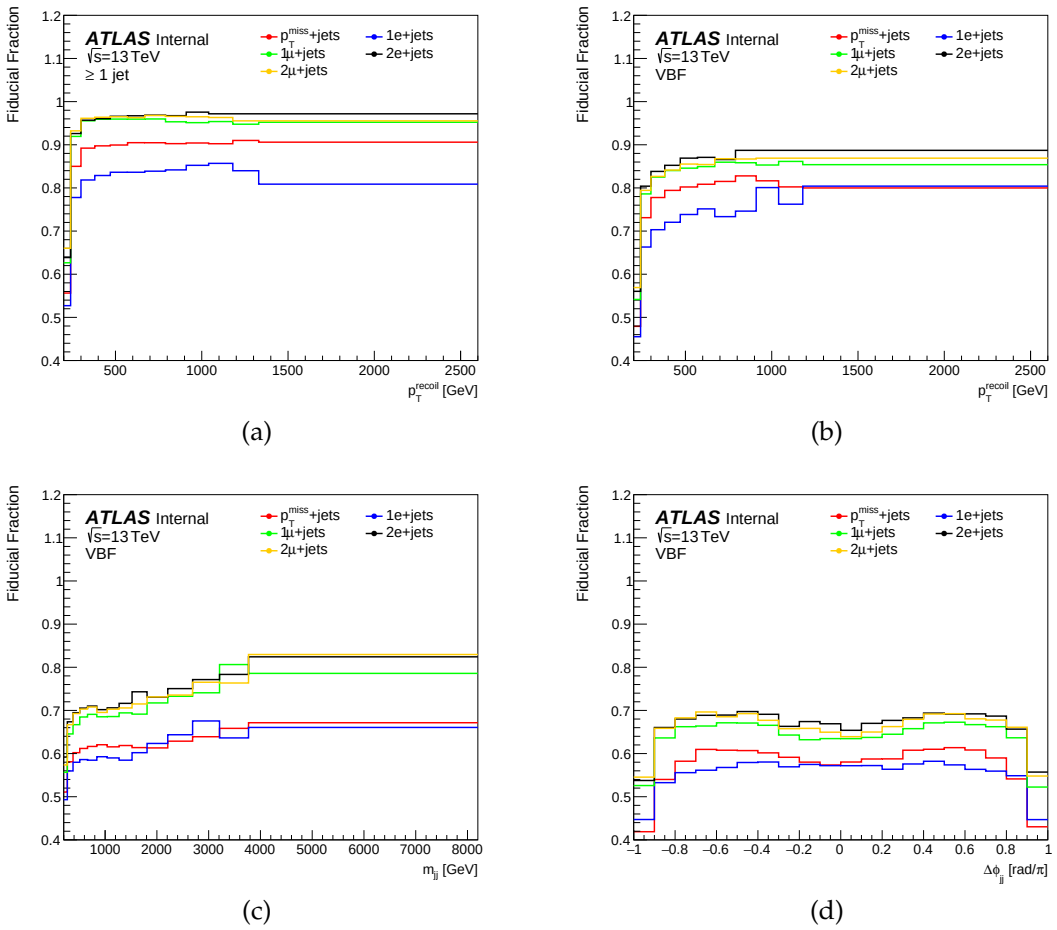


Figure 5.21: Fiducial fraction for the observables that are unfolded in this measurement. Similar to the matching efficiency, low values of the fiducial fraction at the lower $p_{\mathrm{T}}^{\mathrm{recoil}}$ and m_{jj} bounds in subfigures (a), (b) and (c) are due to migrations into the phase-space from events that sit slightly below the cut-off value of 200 GeV at particle-level, but are reconstructed at higher energies at detector-level. The $1e$+jets fiducial fraction is lower than the other regions, which could be explained by migrations in the real $p_{\mathrm{T}}^{\mathrm{miss}}$ cut and the W transverse mass cuts that are only applied in this region. The lower value for the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region is due to events with a leptonically decaying W, which are vetoed at particle-level but can be reconstructed at detector-level if the lepton is out-of-acceptance.

Every distribution, $p_{\mathrm{T}}^{\mathrm{recoil}}$ in the ≥ 1 jet region and $p_{\mathrm{T}}^{\mathrm{recoil}}$, m_{jj} and $\Delta\phi_{jj}$ in the VBF region, is unfolded individually. The simulation-based quantities described above are re-derived for each source of systematic uncertainty. After unfolding the nominal, experimental dataset for each source of uncertainty, the total systematic uncertainty is calculated as the sum of squared differences between the nominal and systematic histograms cf. Equation 5.18. Correlations are taken into account using the bootstrap method described in Section 5.4.1.

5.6.2 Closure tests

The detector-level distributions for MC, particle-level distributions for MC and migration matrices are derived using the same set of simulated MC events. Therefore, if you "unfold" a pseudo-dataset that corresponds to the detector-level MC back to the particle-level using the migration matrices from the same event sample, you should get back precisely the particle-level predictions. This is called a *closure test* and is the first check in making sure the unfolding procedure is set up correctly.

The results for the closure test in this analysis are presented in Figure 5.22 for the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region. As expected, the ratio between particle-level MC and "unfolded" detector-level MC show exact closure. It is also verified that the auxiliary lepton measurements satisfy the closure text precisely.

5.6.3 Regularisation bias and optimisation of the number of unfolding iterations

The iterative Bayesian unfolding technique and the prior used in this method introduce a bias to the unfolded result. The amount of bias depends on the regularisation, which is determined mainly by the number of unfolding iterations and the choice of binning, in addition to the shape agreement between the MC that is used for the unfolding and the experimental data.

Because we do not have information about the particle-level spectra for the experimental data, in fact this is what we are trying to measure, a data-driven approach is needed to estimate the size of the unfolding bias. To achieve this, truth-level events from the MC used for unfolding are reweighted so that the reweighted reconstruction-level distributions match the data. The reweighted distributions are then unfolded using the migration matrix from the un-reweighted MC sample. The residual differences between the unfolded, reweighted MC and the reweighted truth-level MC are defined as the regularisation bias (or the *basic unfolding uncertainties*).

The statistical uncertainties on the unfolded spectra increase with each iteration, so the optimal number of iterations minimises the regularisation bias while keeping the statistical uncertainties as low as possible. The optimal number of iterations is then found by comparing the size of statistical uncertainties and the residual bias for a different number of unfolding iterations. In this analysis, the optimal number of iterations is two, as for larger values the regularisation bias does not decrease further while the statistical uncertainty increases. The residual bias is propagated as a systematic uncertainty to the unfolded data.

The unfolding of the measured $p_{\mathrm{T}}^{\mathrm{miss}}$, m_{jj} and $\Delta\phi_{jj}$ variables can implicitly depend on other kinematic variables, such as the p_{T}, ϕ or η of reconstructed objects. Differences between data and MC of these *hidden variables* can affect the unfolded results. The dependence on the modelling of

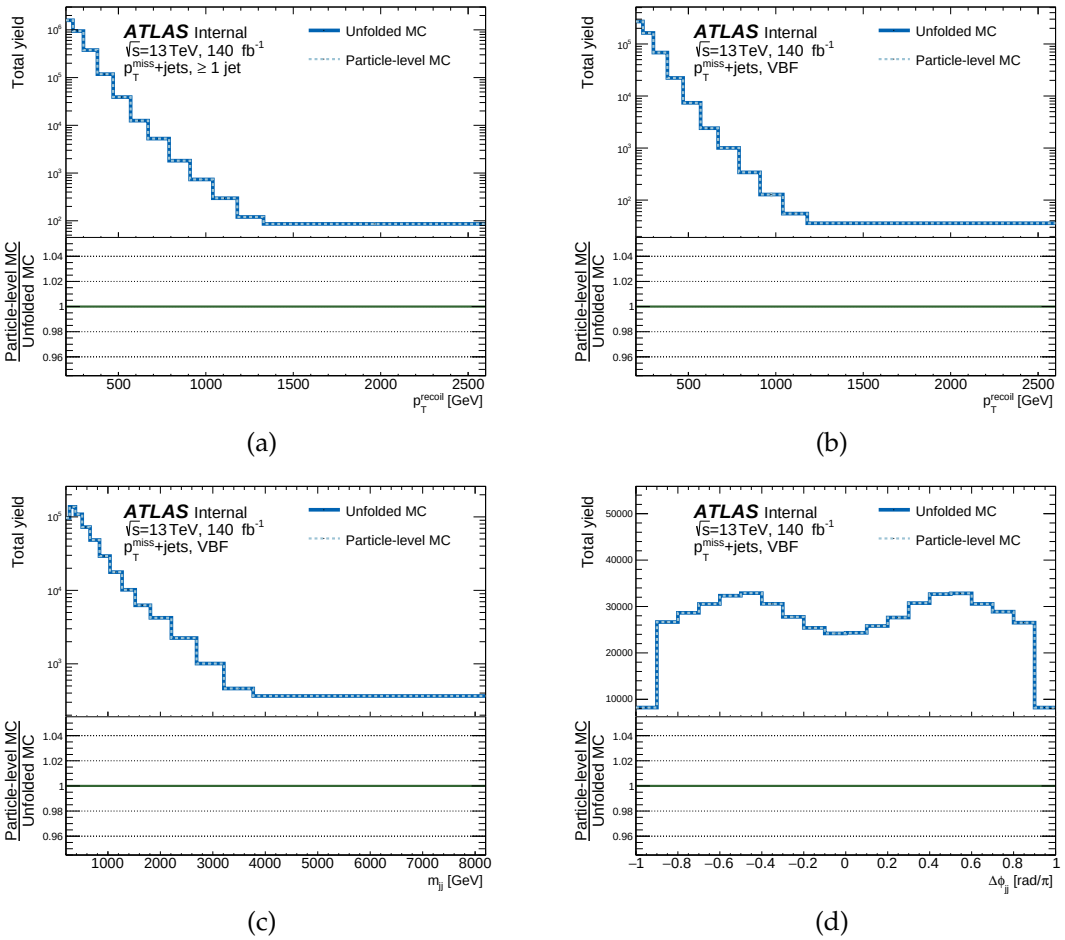


Figure 5.22: Closure tests for all measure observables in the ≥ 1 jet region comparing unfolded detector-level MC predictions to its corresponding truth-level histograms. In the bottom panel, ratio between these two components shows exact closure..

the shape of hidden variables is determined in a similar way as the regularisation bias: a pseudo-dataset is obtained for each variable and the residual bias is taken as the associated uncertainty on the unfolded measurement. Hidden variables include jet η, N_{jets} and ϕ (all regions); lepton p_{T}, η and ϕ (for regions that measure at least one lepton); and $m_{\ell\ell}$ (di-lepton regions). The leading jet p_{T} is not explicitly included. Due to its strong correlations to $p_{\text{T}}^{\text{recoil}}$ any mismodelling effects of the leading jet-p_{T} is implicitly accounted for in the regularisation bias uncertainty.

5.6.4 Process composition tests

Physics processes generally have a different detector response. This requires caution in inclusive measurements like the $p_{\text{T}}^{\text{recoil}}$ + jets measurement presented in this thesis. Due to the stochastic nature of quantum mechanics, we do not know what underlying process our events originate in the experimental dataset. As the MC used to unfold the data consists of various Standard Model processes and we aim to unfold a stack of different processes – which we are agnostic to – simultaneously, the unfolded results depend on the modelling of the process composition. This section describes two uncertainties that ensure the measurement is robust to SM modelling effects.

First, the relative contributions from different SM backgrounds in this regions of phase-space are constrained by dedicated control regions as detailed in Section 5.3.5, with corresponding normalisation factors presented in Table 5.6 and 5.7. The normalisation factor uncertainties indicate the variance of the contribution for each considered SM process. A systematic uncertainty on the process composition is then found by varying each SM contribution within the $\pm 1\sigma$ uncertainties on the normalisation factors, while keeping the other backgrounds fixed at the nominal value. Any residual differences with the unfolded results using the nominal SM composition are taken as the systematic uncertainty for that process. The combined uncertainties for this variation are smaller than 1%.

An additional dependence on the subtraction scheme used to account for interference terms in different top processes – Diagram Subtraction (DS) or Diagram Removal (DR) – is evaluated by using MC samples corresponding to either scheme and comparing the results. The single-top process only has a sizeable contribution in the high $p_{\text{T}}^{\text{recoil}}$ tails of the distributions measured in the $1e$+jets and 1μ+jets regions. The effect of the variations in the single-top contributions for these two regions is shown at detector-level in Figure 5.23. The effect on the experimental data after unfolding is shown in Figure 5.24. The event yield at detector level is up to 14% different in the most extreme bin, but studies of top-enhanced regions show that the agreement between data and MC is well within the theory uncertainty in these regions for both DS and DR samples. The residual effect on the unfolded results is very small because the large discrepancy occurs only bins that suffer from low statistics and already have a large associated statistical uncertainty. The residual difference between the nominal DS sample and the DR variation at truth-level is included as a systematic uncertainty.

5.6.5 BSM signal injection

The Standard Model predictions used in the unfolding procedure are not representative of the experimentally measured data if the data contains BSM physics. This has two main reasons. First, new physics signals may exhibit features that are not present in the SM predictions, such

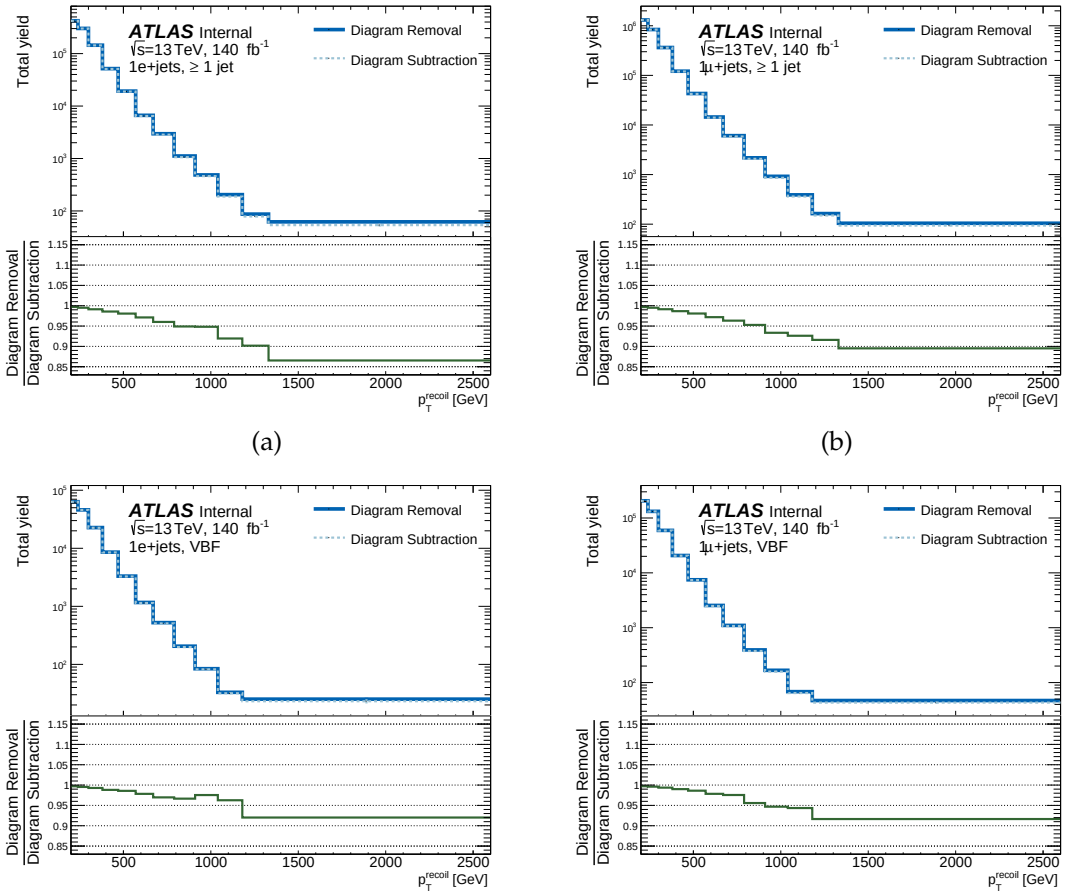


Figure 5.23: Evaluation of single-top DR and single-top DS modelling effects at detector-level. The plots compare the total detector-level MC yields when different single-top samples are included. This is limited to the $p_{\mathrm{T}}^{\mathrm{recoil}}$ observables for the single-lepton regions, because other observables or regions show no significant discrepancy when using different MC samples. For each region the MC predictions derived with the DR method can be up to 14% larger for the high $p_{\mathrm{T}}^{\mathrm{recoil}}$ bins.

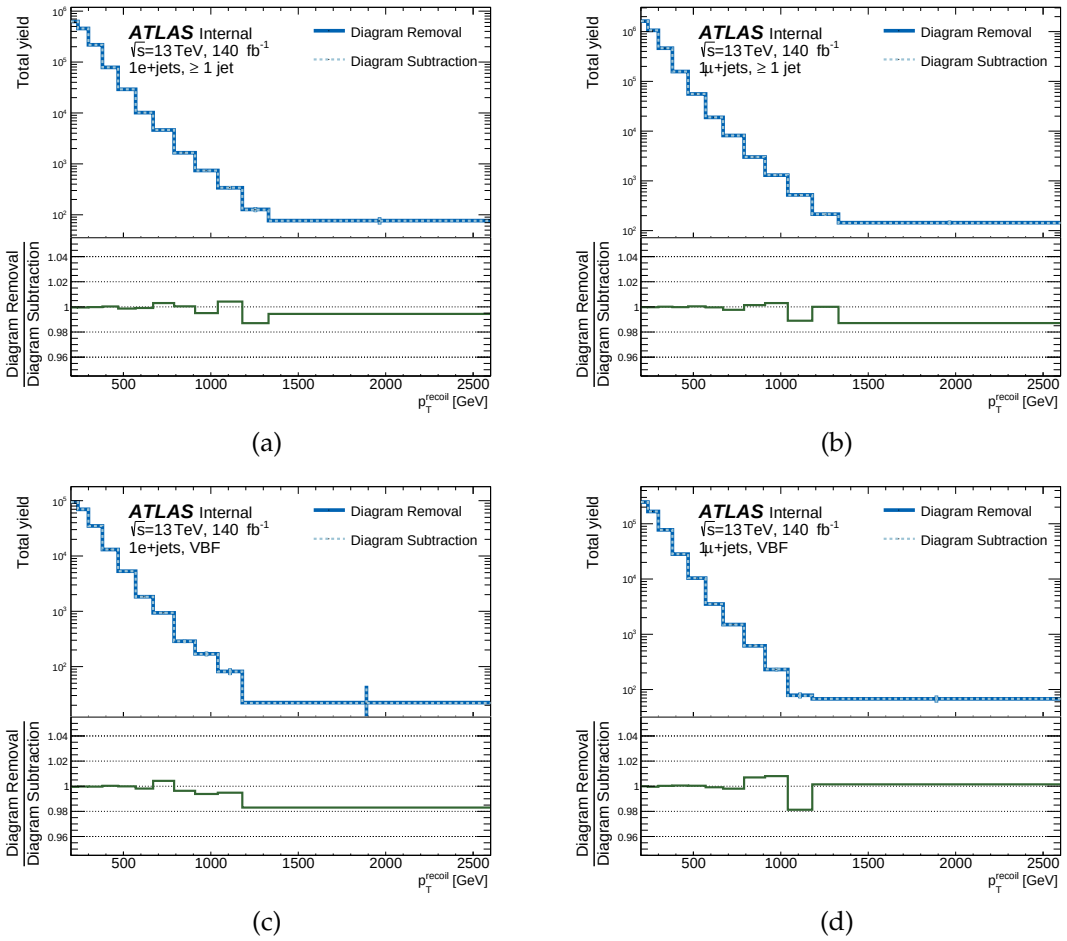


Figure 5.24: Evaluation of single-top DR and single-top DS modelling effects after unfolding. The plots show differential distributions for unfolded, experimental data when different MC samples corresponding to Figure 5.23 are used. The solid and dashed line correspond to scenarios where single-top DR and single-top DS samples are used as inputs for the unfolding procedure, respectively. While the difference in MC yields can be up to 14% in some bins at detector-level, these plots demonstrate that these effects are mostly mitigated by the unfolding procedure. After unfolding, the residual difference in unfolded data when using the two different schemes is smaller than the statistical uncertainties on the experimental data, which is shown as vertical error bars in the top panels.

as bumps in a mass spectrum or longer tails of distributions of energy spectra. The unfolding is biased towards the SM prediction – because this is used as the prior – and could therefore potentially wash away any BSM effects that might be hiding in the data. Furthermore, the detector response for BSM processes is not guaranteed to be modelled well by the response matrices which are derived with SM event samples. In order to test the robustness to potential new physics contributions, and to make sure any new physics signal would not be smeared away during unfolding, *signal injection* tests are performed.

In these tests, various BSM signals are injected into the Standard Model MC predictions. This combination is then treated as pseudodata and unfolded with the (nominal) MC inputs from the SM event samples. The unfolded pseudo-data can then be compared to the corresponding distributions of SM plus signal at particle-level. A small difference gives assurance that the unfolding procedures leaves the BSM signals intact when mapping experimental data from detector-level to particle-level.

Myriad different models can be tested, but in this DM-sensitive phase space three different configurations of a simplified Dark Matter model are considered, where a Dirac DM fermion χ couples to a mediator particle A. The configurations where the DM mediator particle is an axial-vector mediator are referred to as "DMA", while DM models with a pseudo-scalar mediator are referred to as "DMP". The mass scales and quantum numbers are chosen to cover a range of different BSM scenarios:

1. s-channel Dark Matter model mediated by an spin-1 axial-vector mediator (DMA) with mass $M_{\mathrm{A}} = 700\,\mathrm{GeV}$ and a Dark Matter mass of $M_{\chi} = 1\,\mathrm{GeV}$.

2. s-channel Dark Matter model mediated by an spin-1 axial-vector mediator (DMA) with mass $M_{\mathrm{A}} = 700\,\mathrm{GeV}$ and a Dark Matter mass of $M_{\chi} = 355\,\mathrm{GeV}$.

3. s-channel Dark Matter model mediated by an spin-0 pseudo-scalar mediator (DMP) with mass $M_{\mathrm{A}} = 50\,\mathrm{GeV}$ and a Dark Matter mass of $M_{\chi} = 1\,\mathrm{GeV}$.

Due to the large mediator mass in the first two scenarios, a large contribution is expected in the tails of the $p_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region. The third scenario has a lower mediator mass and can also contribute in the lower $p_{\mathrm{T}}^{\mathrm{miss}}$ regions. Figure 5.25 show the detector-level comparison between the SM prediction to the SM prediction with the BSM signal from scenario 1 added on top in the left column. The right column shows the final comparison between the unfolded pseudodata – composed of the SM prediction plus the BSM contributions – and the particle-level SM plus signal predictions. Similar results for scenario 2 and scenario 3 are shown in Figure 5.26 and 5.27 respectively. The lepton+jets measurement regions show no contributions from the BSM signal events and are therefore not included.

As can be expected, the unfolded vs. truth level comparisons depend strongly on the BSM signal strength. Even for the DMA signal with $M_{\mathrm{A}} = 700\,\mathrm{GeV}$ and $M_{\chi} = 1\,\mathrm{GeV}$, where the BSM contribution is almost an order of magnitude larger than the SM, there is an offset after unfolding of at most 15%. The model introducing this bias is so extreme that if it was present in reality, the discrepancy would already be clearly present in the data at detector-level. The other two samples, with more moderate contributions, generally give a good closure between the unfolded and truth-level MC predictions.

A secondary effect that influences the comparisons between unfolded pseudodata and particle-level MC, is that the BSM processes can have a different fiducial fraction and reconstruction efficiency than the SM. In the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution for the ≥ 1 jet region, a larger number of BSM events is reconstructed at detector-level than at particle level. For the DMA sample with $M_{\chi} = 1\,\mathrm{GeV}$, the tails of the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution are completely dominated by the BSM contribution. It turns out that the relative difference in event yield between particle-level and detector-level can be up to 20%. Therefore, the larger yield at detector-level directly results in the unfolded BSM yield being larger than the corresponding truth predictions in Figure 5.25b. Conversely, a lower reconstruction efficiency for BSM events in the VBF region results in a larger BSM yield at particle-level than at detector-level. This outweighs the naive expectation that the SM bias – from using SM predictions as the prior – brings the unfolded pseudodata closer to the SM, and consequently the ratio in the bottom panel of Figure 5.25d is larger than one.

The main goal of the signal injection test is to verify that the unfolding procedure does not unfold away any new physics contributions in the data. These tests show a good closure, even for extreme scenarios, and we therefore conclude that any BSM in the data would be preserved in the unfolding with SM inputs. The main difference between unfolded pseudodata with BSM and truth-level histograms comes from differences in fiducial fraction and efficiency for these extreme BSM events, for which the measurement is not optimised and which goes beyond the actual purpose of the signal injection test.

5.6.6 Phase space migrations via underflow bins

Events can migrate in or out of the fiducial phase-space between detector-level and particle level. Due to resolution effects, events can fall in or out the lowest bin of the measured variables, the underflow bin, and affect the unfolded measurement. This effect is tested by adding two additional bins below the nominal measurement range in the $p_{\mathrm{T}}^{\mathrm{miss}}$ and m_{jj} observables. The nominal range, starting at $200\,\mathrm{GeV}$, is for both variables extended by two additional bins with bin range [140–170] GeV and [170-200] GeV. The unfolding is performed by including these two bins, and removing the two bins after the unfolding. Comparing this to the nominal result, the effect of including the underflow bins in the unfolding is very tiny and therefore neglected.

Additionally, comparisons of data and simulations are made where the various cuts that define the fiducial phase-space are reduced below their nominal value. This tests the modelling of migrations in and out of phase-space due to mismodelling on variables that are used to cut on. Simulated events are reweighted such that the reconstructed distributions match the data. Changes to the unfolded measurement are negligible and no uncertainty for this effect is deemed necessary to include.

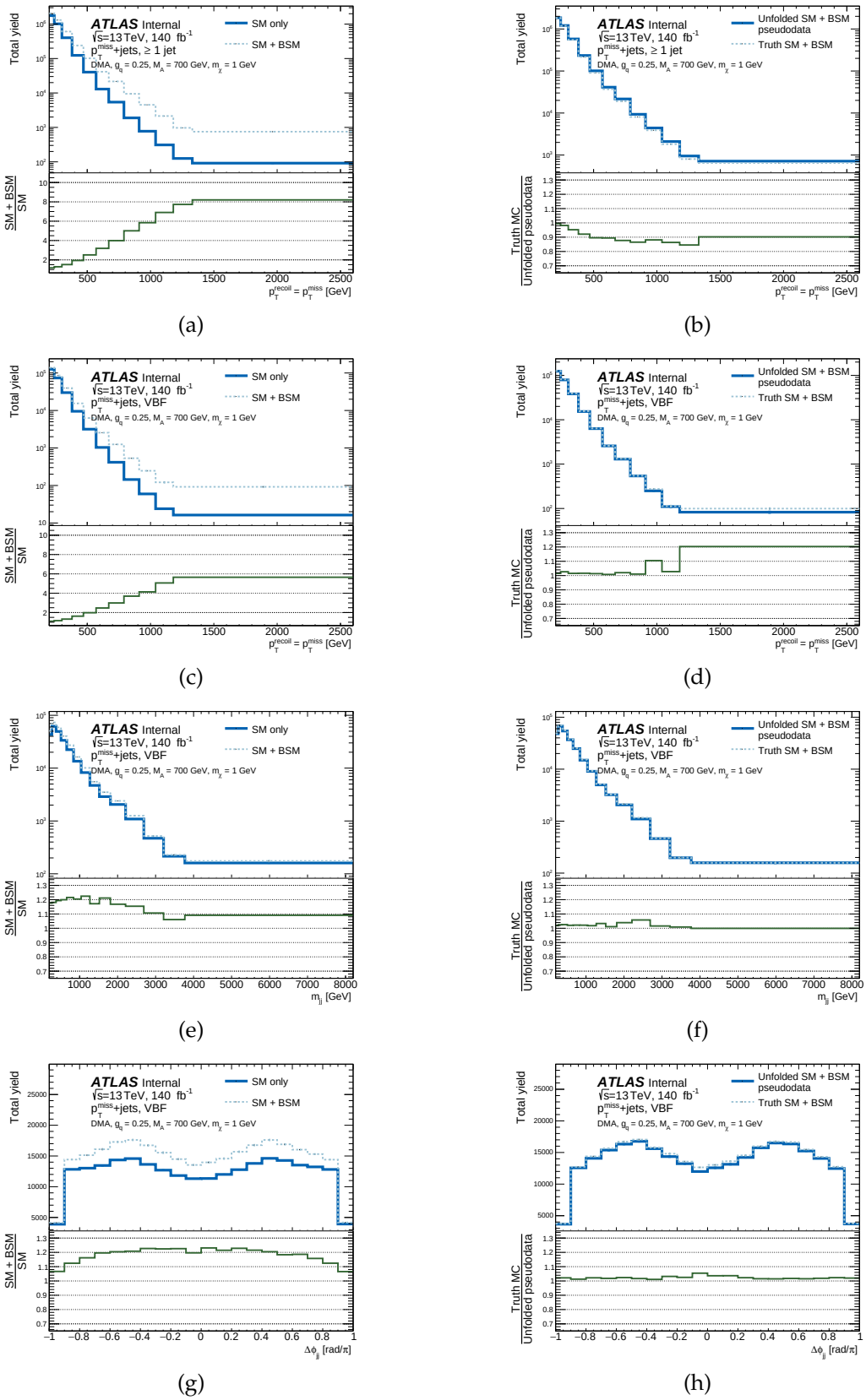


Figure 5.25: Signal injection tests of the axial-vector mediated DM scenario with $M_{\mathrm{A}} = 700\,\mathrm{GeV}$ and $M_{\chi} = 1\,\mathrm{GeV}$, for different observables in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region. The figures in the left column show the detector-level comparisons between SM only and SM+BSM, to indicate the relative size of the BSM contribution. The figures on the right column show the difference between unfolded pseudodata with BSM included and the corresponding particle-level SM+BSM yields. Even BSM contributions that lead to discrepancies of several orders of magnitude, for example for $p_{\mathrm{T}}^{\mathrm{miss}}$ for the ≥ 1 jet region in subfigure (a), lead to agreement within one standard deviation of the statistical uncertainties in the corresponding bins in subfigure (b).

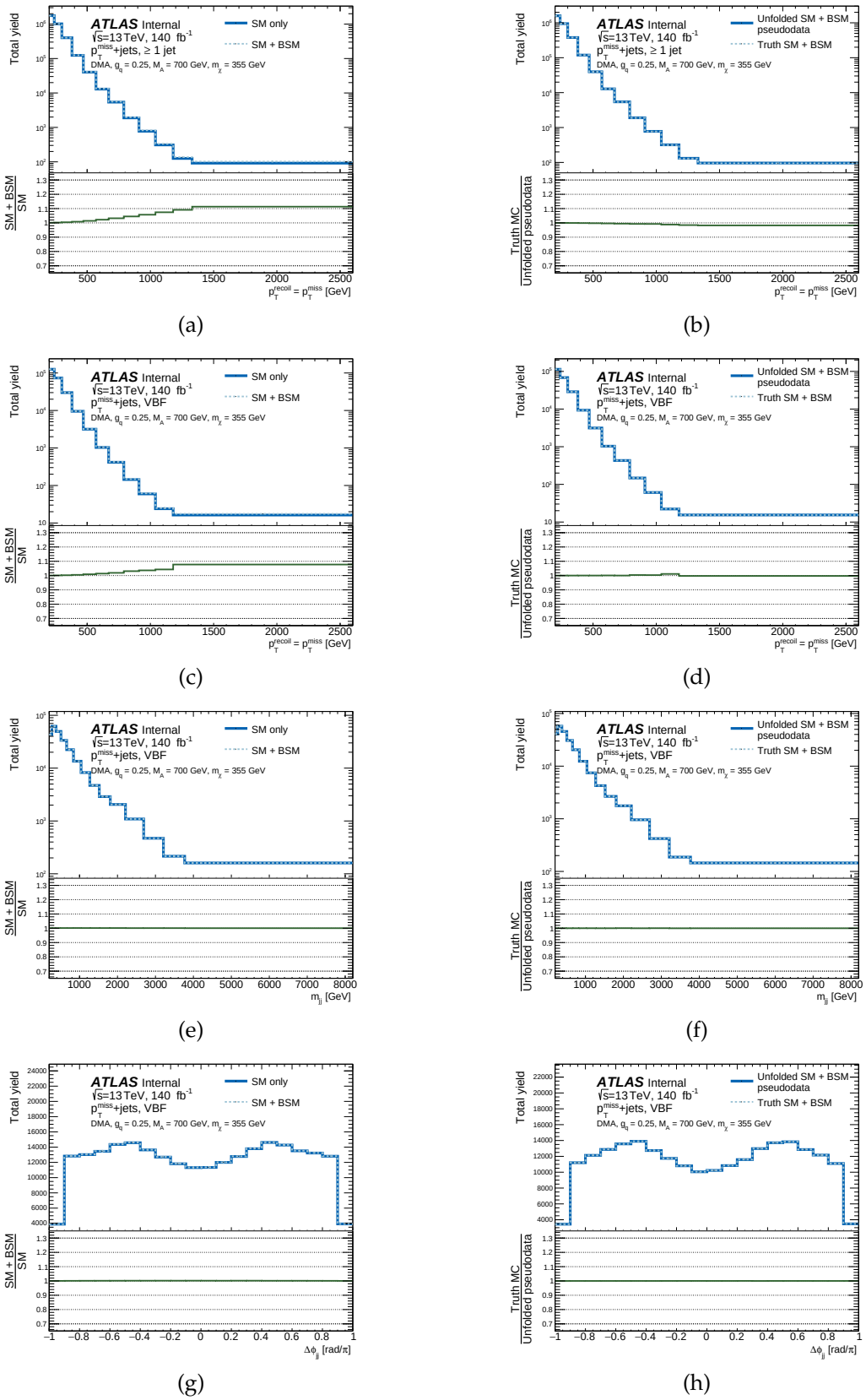


Figure 5.26: Signal injection tests of the axial-vector mediated DM scenario with $M_{\mathrm{A}} = 700\,\mathrm{GeV}$ and $M_{\chi} = 355\,\mathrm{GeV}$, for different observables in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region. The left column shows the detector-level comparisons between SM only and SM+BSM, to indicate the relative size of the BSM contribution. The right column shows the difference between unfolded pseudodata with BSM included and the corresponding particle-level SM+BSM yields. Compared to the scenario with $M_{\chi} = 1\,\mathrm{GeV}$, as shown in Figure 5.25, the contribution of this BSM configuration is smaller. Discrepancies between unfolded pseudodata and particle-level SM+BSM yields are up to 2% in the most discrepant bin, in the $p_{\mathrm{T}}^{\mathrm{recoil}}$ observable for the ≥ 1 jet region in subfigure (b).

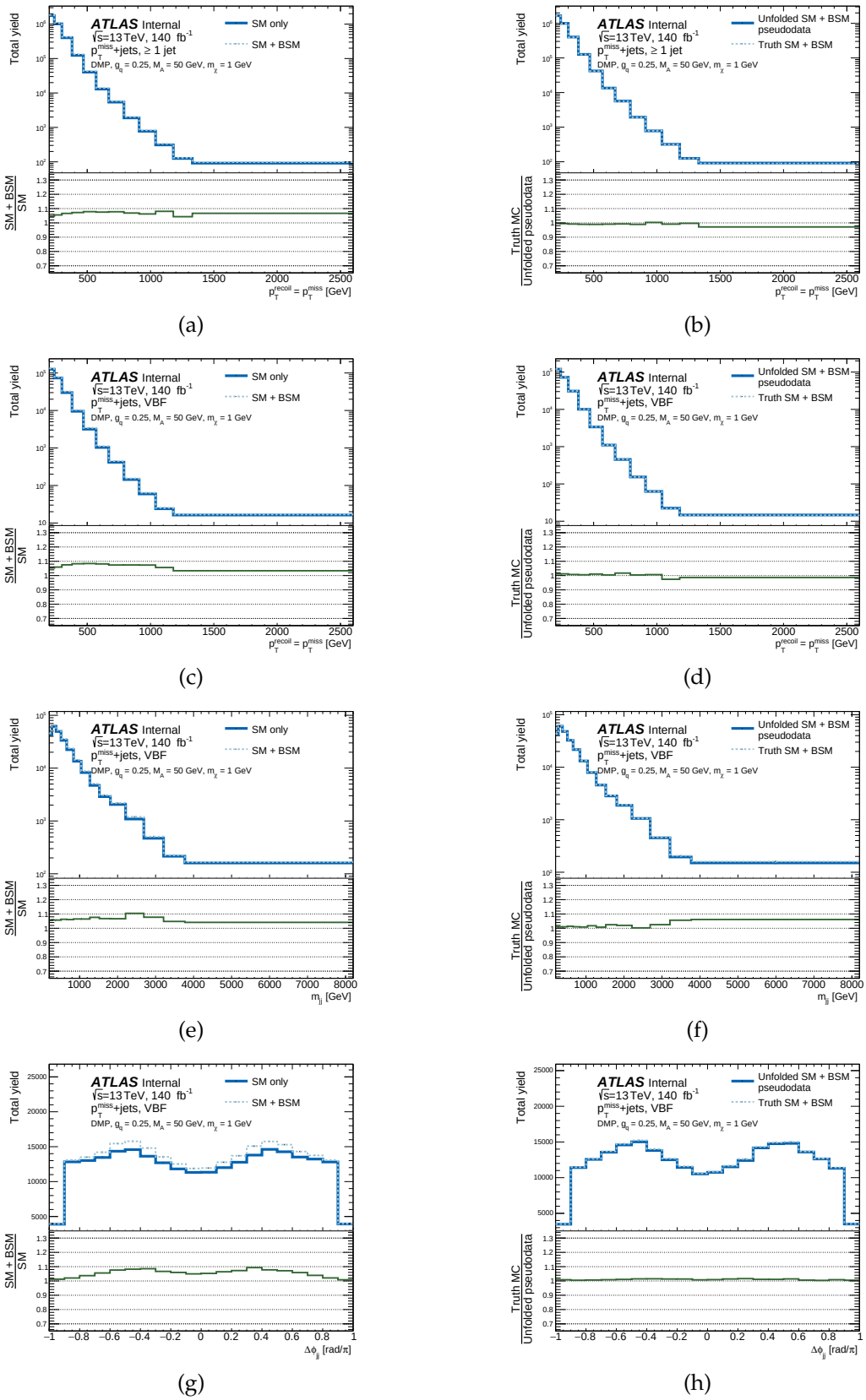


Figure 5.27: Signal injection tests of the pseudo-scalar mediated DM scenario with $M_{\mathrm{A}} = 50\,\mathrm{GeV}$ and $M_{\chi} = 1\,\mathrm{GeV}$, for different observables in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region. The left column shows the detector-level comparisons between SM only and SM+BSM to indicate the relative size of the BSM contribution. The right column shows the difference between unfolded pseudodata with BSM and the corresponding particle-level SM+BSM yields. The mediator mass is lower compared to scenario 1 and 2, which leads to BSM signals at low $p_{\mathrm{T}}^{\mathrm{recoil}}$ and across the m_{jj} and $\Delta\phi_{jj}$ distributions. Discrepancies between unfolded pseudodata and particle-level SM+BSM yields is up to 7% in the most discrepant bin in subfigure (f), but the residual differences at particle-level are within statistical uncertainties of the experimental data.

5.7 Particle-level results

With detector-level results at hand and after thorough testing of the unfolding method, the unfolded results at particle-level are shown in this section. First, the differential cross-section measurements for the individual regions are presented in Section 5.7.1. The R^{miss} ratio are constructed from the signal region and the lepton regions after unfolding and are presented in Section 5.7.2. A consistency check between the particle-level results for the inclusive measurement and the dominant process measurement is presented in Section 5.7.3. Additionally, Section 5.7.4 shows the results of a cross-check comparing the unfolded data from the 2μ+jets region to the $2e$+jets region. Finally, the statistical correlations between bins in a histogram and different observables are shown in Section 5.7.5.

The histograms of unfolded data with a detailed breakdown of the systematic and statistical uncertainties are published on HEPData in Ref. [237]. In addition, the SM predictions with the individual process components are also preserved. This enables a wide range of phenomenological re-interpretation studies. For example, due to the sensitivity to BSM physics, the data can be used to constrain new physics using tools such as CONTUR [241]. In addition, improved SM predictions can be validated directly to the data without the need for detector simulation. If these new predictions only constitute a subset of the processes that contribute to this inclusive measurement, they can be combined with the other SM processes used in this measurement that are preserved on HEPData.

5.7.1 Unfolded, differential distributions

$\geq$ 1 jet

Comparisons between unfolded, experimental data and MC at the particle-level for the $p_{\text{T}}^{\text{miss}}$ observables measured in the ≥ 1 jet topology are shown in Figure 5.28.

Like the corresponding plots at detector-level in Section 5.5.1, the differential cross-section drops five to six orders of magnitude across the $p_{\text{T}}^{\text{miss}}$ spectrum for the signal region and the lepton regions. When a single charged muon is required, the cross-section is similar to the signal region, although it is lower for the $1e$+jets due to tighter selection requirements on electrons. The requirement of two charged leptons reduces the cross-sections roughly by an order of magnitude compared to the signal region and the single lepton regions.

The data is compared with SM predictions from Table 5.9. The MEPS@NLO prediction provides adequate modelling, where the error band with a relative up and down uncertainty of approximately 20% is dominated by scale uncertainties. The prediction reweighted to NNLO in α_S improves modelling of the overall cross-section and reduces the size of the theoretical uncertainties across all bins.

The relative contribution of different SM processes to the total MEPS@NLO cross-section is shown in the bottom panel. For all regions, the QCD V+jets process is the dominant process across the measurement bins. In the signal region, electroweak Vjj and diboson process contribute more toward larger values of $p_{\text{T}}^{\text{miss}}$, while the relative contributions of top and $W+$jets drop to sub-per mille and sub-percent values respectively. The trend of increasing electroweak Vjj and diboson with $p_{\text{T}}^{\text{recoil}}$ is also observed in the lepton auxiliary regions.

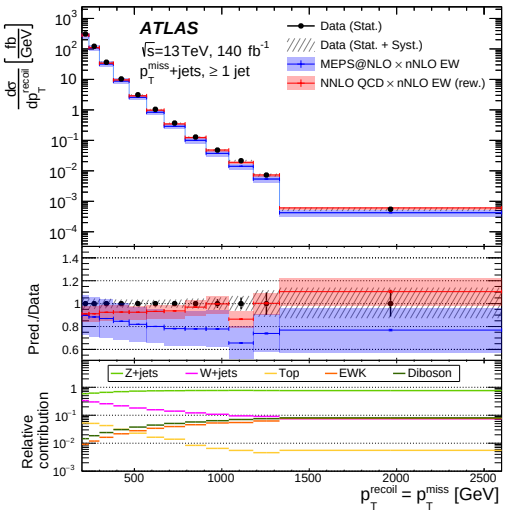
A breakdown of the corresponding systematic and statistical uncertainties on the experimental data is shown in Figure 5.29. Statistical uncertainties are evaluated from the bootstrap replicas using Equation 5.16. Similar to the detector-level uncertainty plots, JES uncertainties dominate at low $p_{\mathrm{T}}^{\mathrm{recoil}}$ while the leading uncertainty in high $p_{\mathrm{T}}^{\mathrm{recoil}}$ bins is attributed to low statistics due to sparsely populated bins. The uncertainties from fake background are only sizeable for low $p_{\mathrm{T}}^{\mathrm{recoil}}$ values in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets and 1μ+jets regions, which have a larger fraction of fake events than the other regions. The unfolding uncertainty is at most 2% and negligible to the size of the total uncertainty in the associated bins.

VBF

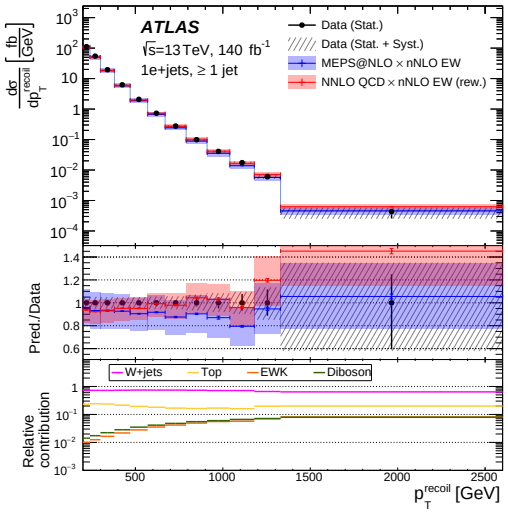
Comparisons between unfolded, experimental data and MC at the particle-level for the $p_{\mathrm{T}}^{\mathrm{miss}}$ observables measured in the VBF topology are shown in Figure 5.30. A breakdown of the corresponding systematic and statistical uncertainties is shown in Figure 5.31

Due to the jet requirements, the cross-sections for the VBF measurements are smaller than ≥ 1 jet measurements. Similar to detector-level, there is an observed increase of EWK Vjj processes, of up to approximately 15% in the highest $p_{\mathrm{T}}^{\mathrm{recoil}}$ bin. The final $p_{\mathrm{T}}^{\mathrm{recoil}}$ bin of the $1e$+jets has a large relative statistical uncertainty possibly due to a very low number of events being reconstructed in that bin in particle-level. This seems unsatisfactory, however there are enough MC statistics to determine the migrations accurately and the bin width is set to have at least 20 events in each bin and a purity of more than 60%: it would be bad practice to "re-optimise" the binning after unfolding and unblinding.

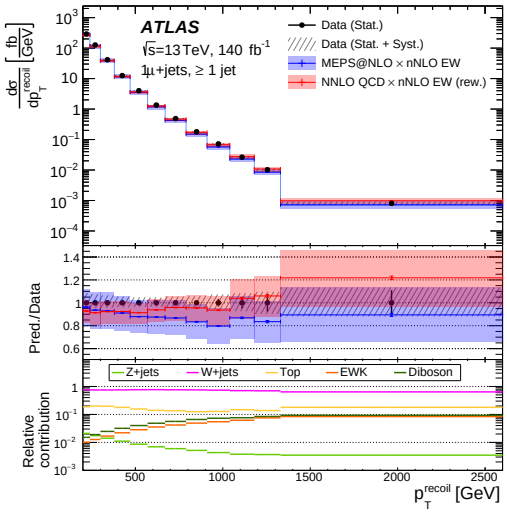
The nominal values of the MEPS@NLO prediction slightly under-predict the data, but the prediction agrees with the data within the errors due to the large theoretical uncertainty of up to 20%. The alternative HEJ prediction, which is available for the lepton regions, improve modelling in 2ℓ + jets and over-predict the data for the ℓ + jets regions. Due to the event sample size the alternative high-energy jet (HEJ) prediction suffers from relatively large statistical uncertainties for high $p_{\mathrm{T}}^{\mathrm{recoil}}$ bins.



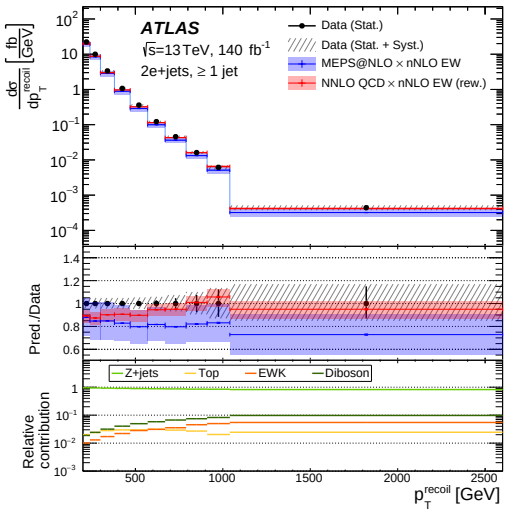
(a) $p_{\mathrm{T}}^{\mathrm{miss}}$+jets



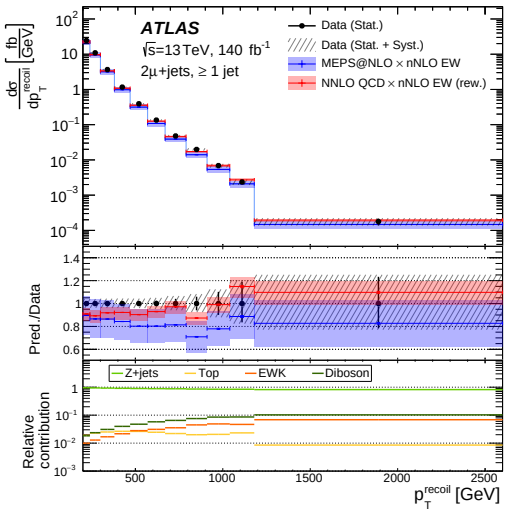
(b) $1e$+jets



(c) 1μ+jets

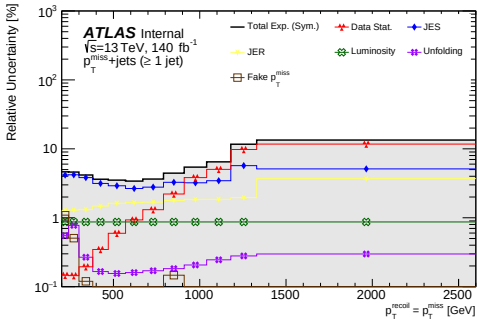


(d) $2e$+jets

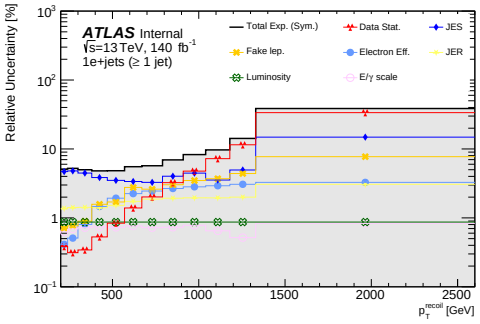


(e) 2μ+jets

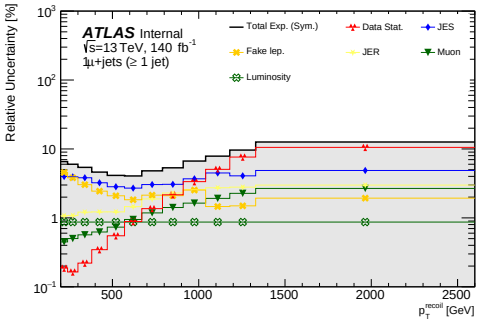
Figure 5.28: Particle-level comparisons between unfolded data and MC for $p_{\mathrm{T}}^{\mathrm{miss}}$ in the ≥ 1 jet phase space. The middle panel shows the ratios of the predictions to the data, with statistical uncertainties as solid markers and shaded bands to indicate the combined statistical and systematic uncertainties. The lower panels show the relative contributions from different SM processes relative to the total MEPS@NLO prediction. Figures from Ref. [232].



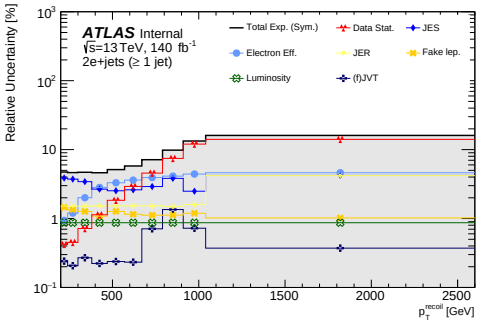
(a) $p_{\mathrm{T}}^{\mathrm{miss}}$+jets



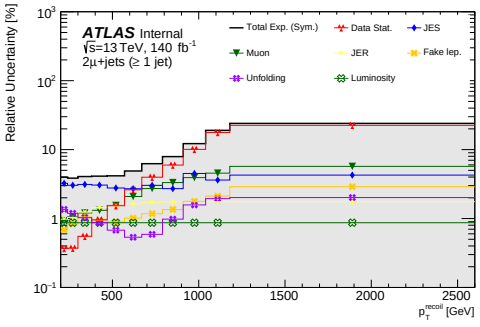
(b) $1e$+jets



(c) 1μ+jets

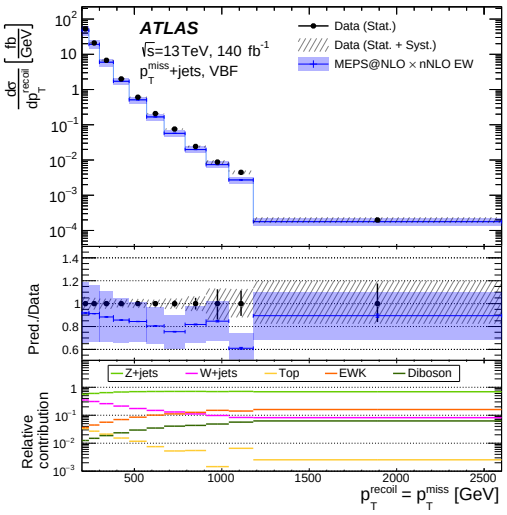


(d) $2e$+jets

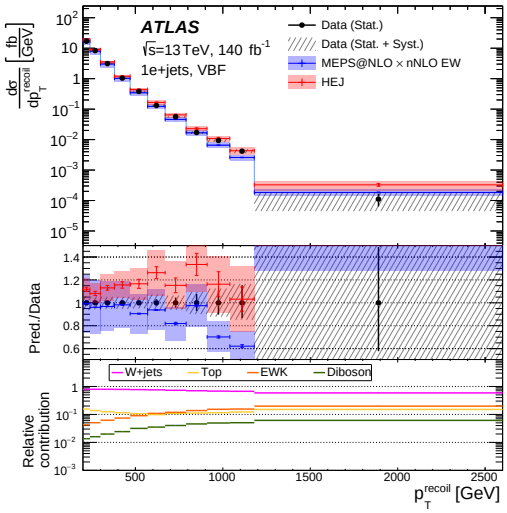


(e) 2μ+jets

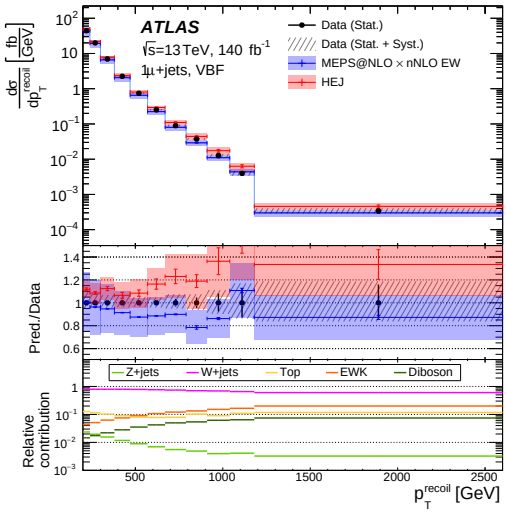
Figure 5.29: Relative contribution from different systematic uncertainty components at particle-level for $p_{\mathrm{T}}^{\mathrm{recoil}}$ in the ≥ 1 jet phase space. Only systematics with a contribution of $\geq 1\%$ in at least one of the bins are shown in the figure. For illustrative purposes this figure shows the symmetrised uncertainties, calculated as the average of the asymmetric error in each bin. Total Exp. (Sym.) is the combination of statistical and systematic uncertainties and indicates the symmetrised total experimental uncertainty. The general trend across all plots is that the jet energy scale (JES) uncertainties are large at low $p_{\mathrm{T}}^{\mathrm{recoil}}$, where statistical uncertainties become increasingly larger with $p_{\mathrm{T}}^{\mathrm{recoil}}$. Subfigure (a) shows a small uncertainty related to fake $p_{\mathrm{T}}^{\mathrm{miss}}$ backgrounds at low $p_{\mathrm{T}}^{\mathrm{miss}}$, while the lepton regions have associated lepton and fake lepton systematic uncertainties across the full $p_{\mathrm{T}}^{\mathrm{recoil}}$ range.



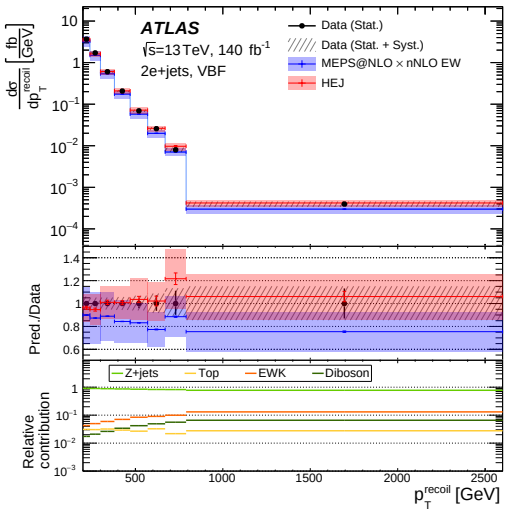
(a) $p_{\mathrm{T}}^{\mathrm{miss}}$+jets



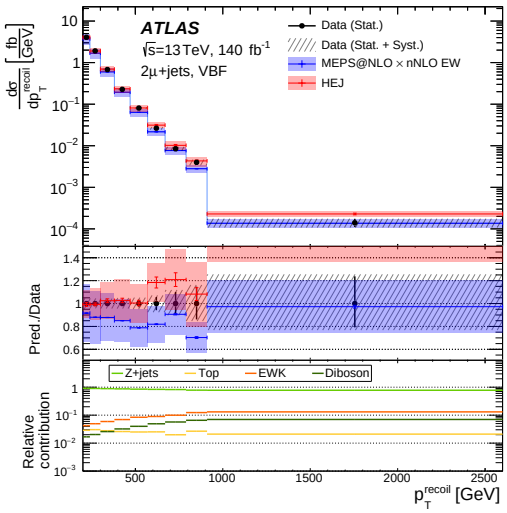
(b) $1e$+jets



(c) 1μ+jets

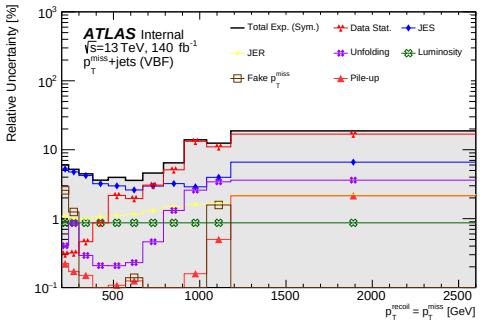


(d) $2e$+jets

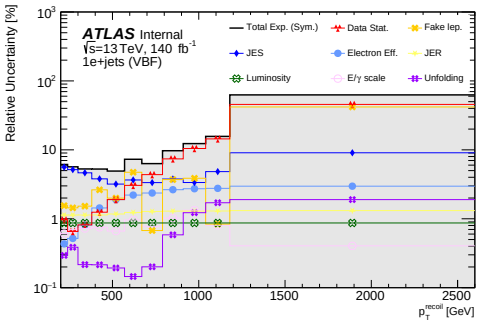


(e) 2μ+jets

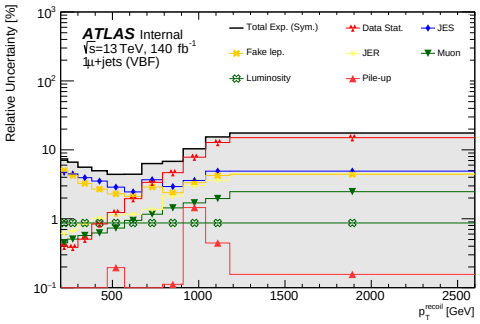
Figure 5.30: Particle-level comparisons between unfolded data and MC for $p_{\mathrm{T}}^{\mathrm{miss}}$ in the VBF phase space. The middle panel shows the ratios of the predictions to the data, with statistical uncertainties as solid markers and shaded bands to indicate the combined statistical and systematic uncertainties. The lower panels show the relative contributions from different SM processes relative to the total MEPS@NLO prediction. The HEJ predictions are available for the lepton regions and over-predict the single-lepton regions while nominal values of the two-lepton region predictions are very accurate at low $p_{\mathrm{T}}^{\mathrm{recoil}}$. Figures from Ref. [232].



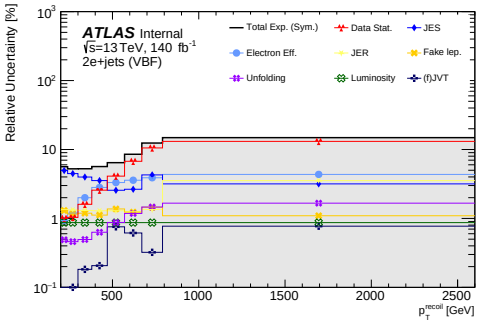
(a) $p_{\mathrm{T}}^{\mathrm{miss}}$+jets



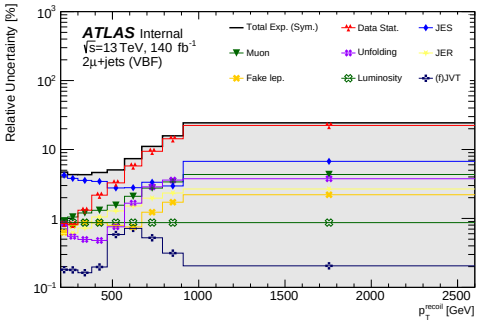
(b) $1e$+jets



(c) 1μ+jets



(d) $2e$+jets

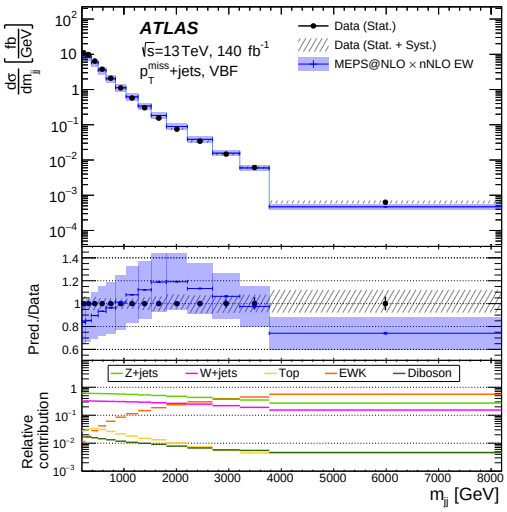


(e) 2μ+jets

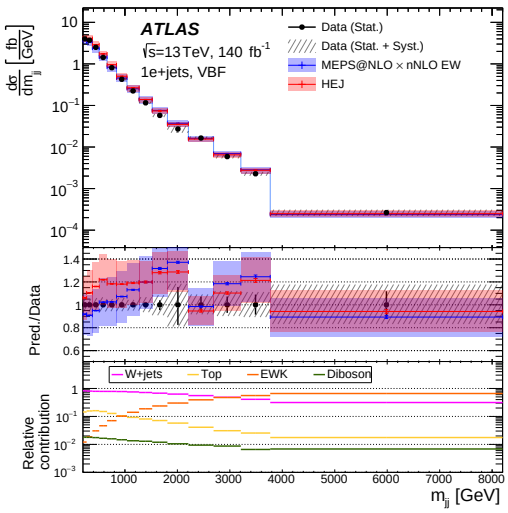
Figure 5.31: Relative contribution from different systematic uncertainty components at particle-level for $p_{\mathrm{T}}^{\mathrm{recoil}}$ in the VBF phase space. Only systematics with a contribution of $\geq 1\%$ in at least one of the bins are shown in the figure. For illustrative purposes this figure shows the symmetrised uncertainties, calculated as the average of the asymmetric error in each bin. Total Exp. (Sym.) is the combination of statistical and systematic uncertainties and indicates the symmetrised total experimental uncertainty. The relatively large uncertainties associated to the fake electron contribution in figure (b) are a propagation of theory uncertainties on the modelling of the fake-enriched region using the matrix method defined in Section 5.3.8.

Comparisons between unfolded, experimental data and MC at the particle-level for the m_{jj} observables measured in the VBF topology are shown in Figure 5.32, with the corresponding uncertainty breakdown plots in Figure 5.33. As m_{jj} increases, EWK Vjj becomes the dominant SM process in the last bin.

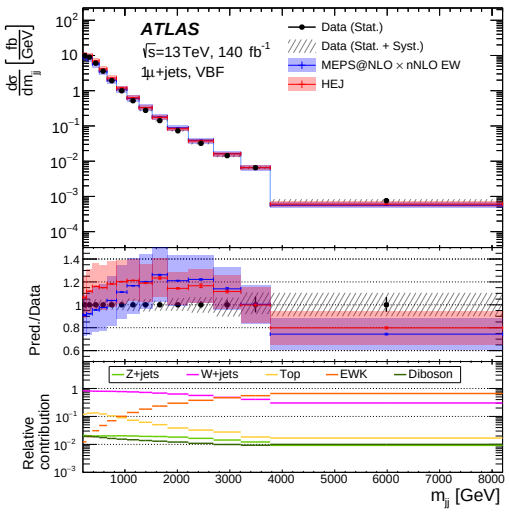
The MEPS@NLO prediction mis-models the shape of the distributions in a similar way as the detector-level MC before the m_{jj} reweighting is applied. Due to the relatively large theory uncertainties, the data and MC agree within the errors bars in most bins. The HEJ prediction, which again is included in the lepton auxiliary measurements, models the shape moderately better than the nominal MEPS@NLO prediction.



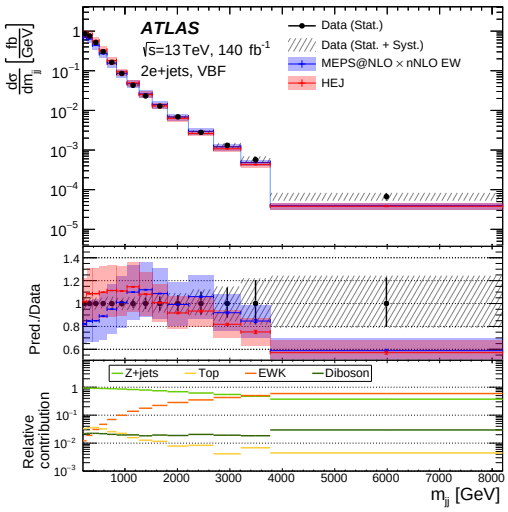
(a) $p_{\mathrm{T}}^{\mathrm{miss}}$+jets



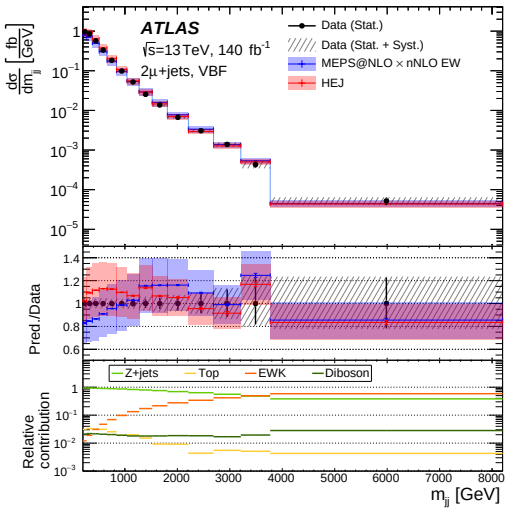
(b) $1e$+jets



(c) 1μ+jets

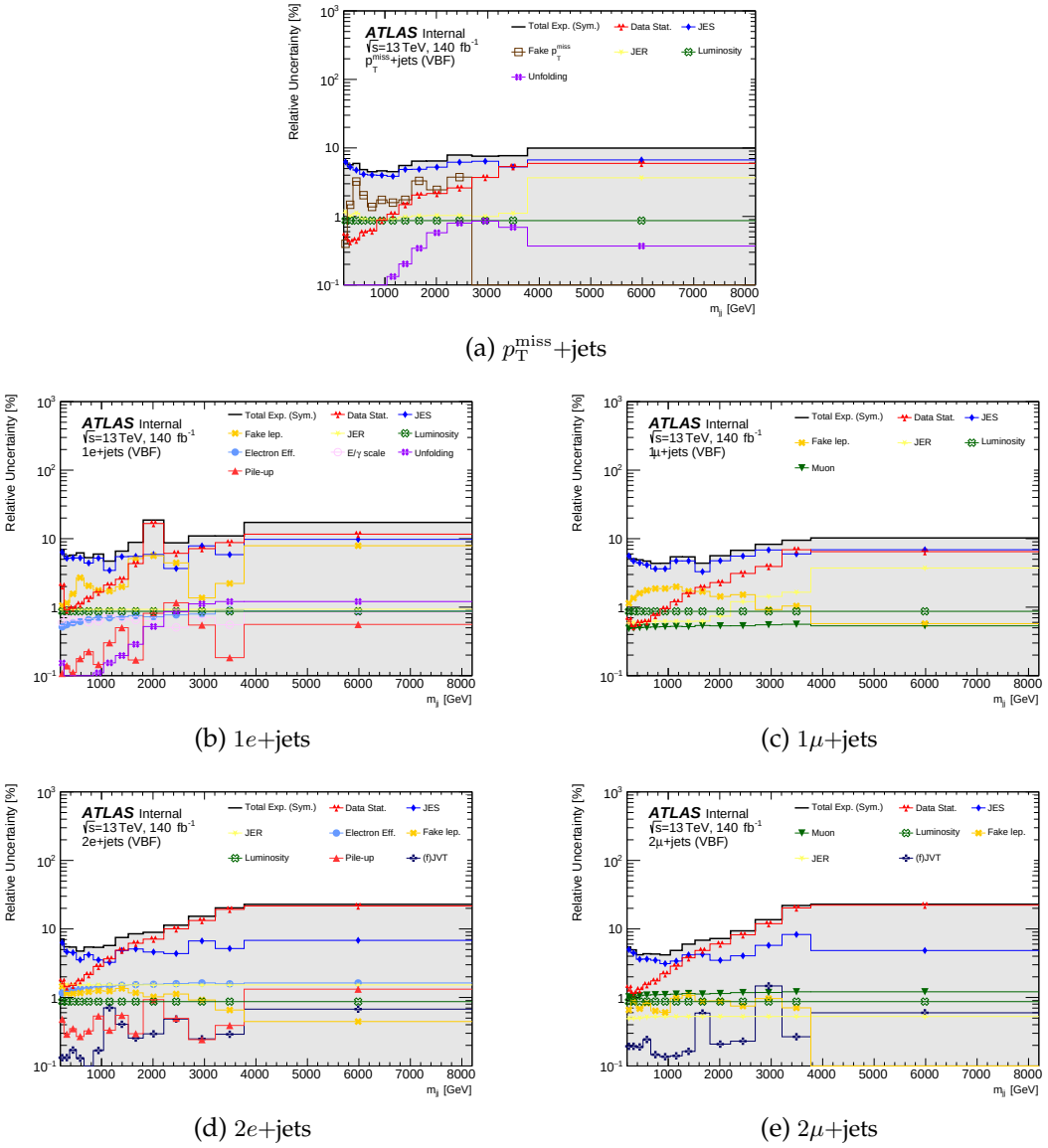


(d) $2e$+jets



(e) 2μ+jets

Figure 5.32: Particle-level comparisons between unfolded data and MC for $p_{\mathrm{T}}^{\mathrm{miss}}$ in the VBF phase space. The middle panel shows the ratios of the predictions to the data, with statistical uncertainties as solid markers and shaded bands to indicate the combined statistical and systematic uncertainties. The lower panels show the relative contributions from different SM processes relative to the total MEPS@NLO prediction. As indicated in the text, there still a shape mismodelling of the MEPS@NLO prediction, although it largely agrees with data within the theory uncertainties. The HEJ predictions, available for the lepton regions, is intended to improve modelling of VBF-like topologies. While the overall fiducial cross-section is off by up to 20%, the ratio to data is more flat across the m_{jj} spectrum. Figures from Ref. [232].



(a) $p_{\mathrm{T}}^{\mathrm{miss}}$+jets

(b) $1e$+jets

(c) 1μ+jets

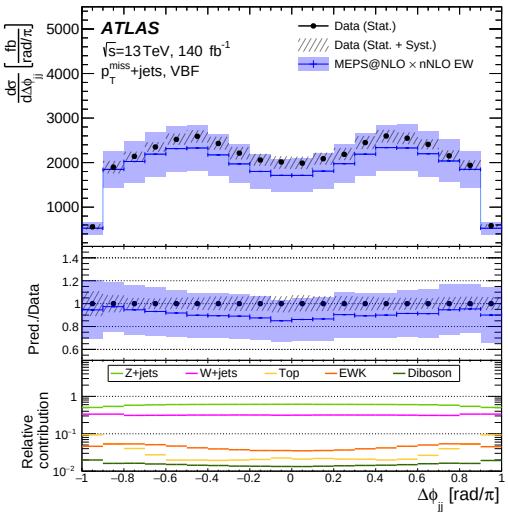
(d) $2e$+jets

(e) 2μ+jets

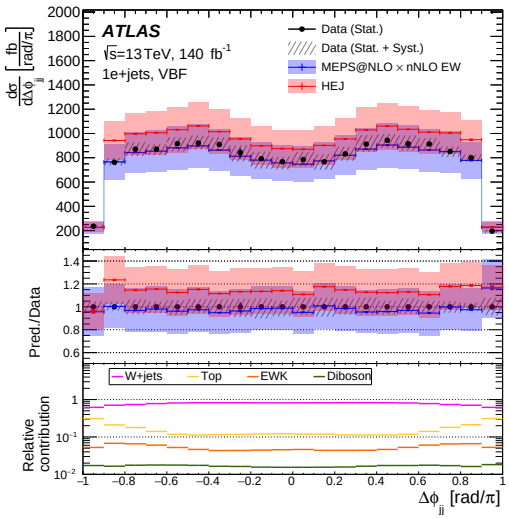
Figure 5.33: Relative contribution from different systematic uncertainty components at particle-level for m_{jj} in the VBF phase space. Only systematics with a contribution of $\geq 1\%$ in at least one of the bins are shown in the figure. For illustrative purposes this figure shows the symmetrised uncertainties, calculated as the average of the asymmetric error in each bin. Total Exp. (Sym.) is the combination of statistical and systematic uncertainties and indicates the symmetrised total experimental uncertainty. In subfigure (a), the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty drops to zero because above a few GeV the $\Delta\phi\left(\mathrm{jet}, p_{\mathrm{T}}^{\mathrm{miss}}\right)$ becomes fully efficient to reject all fake $p_{\mathrm{T}}^{\mathrm{miss}}$ contributions. Subfigures (b) to (d) additionally show the lepton and fake lepton systematic uncertainties, where particularly the fake electron uncertainty for $1e$+jets can lead to relatively large uncertainties on the experimental data.

Comparisons between unfolded, experimental data and MC at the particle-level for the $\Delta\phi_{jj}$ observables measured in the VBF topology are shown in Figure 5.34, with corresponding uncertainty breakdown plots in Figure 5.35.

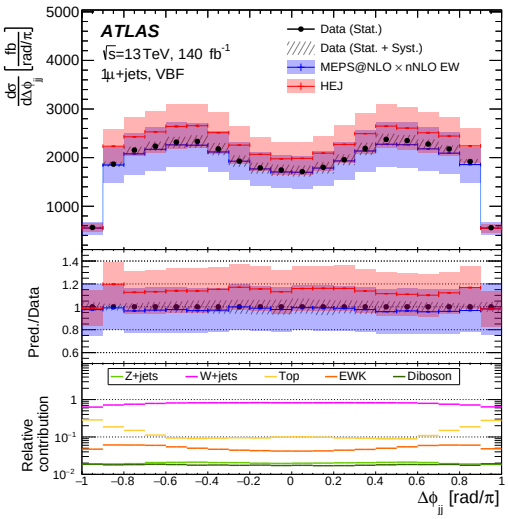
The overall shape is well modelled by both the MEPS@NLO prediction and the HEJ prediction. Both data and MC exhibit similar features to the detector-level distributions for $\Delta\phi_{jj}$ that are shown in Figure 5.16. For every region, each bin has a leading contribution from its corresponding QCD V+jets process, particularly in the two-lepton regions where no subdominant SM process has a contribution larger than 8%. In the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region, the at $|\Delta\phi_{jj}| \approx 1$ has a large relative contribution from fake $p_{\mathrm{T}}^{\mathrm{miss}}$. Since the associated systematic uncertainty is directly proportional to the fake background, this uncertainty is in fact the dominant source of uncertainty here. Several bins in the $1e$+jets region suffer from large systematic uncertainties of the fake lepton estimate, propagated from theory uncertainties of the background modelling in the fake-enriched electron background. Due to the low acceptance of events in the bins at at $|\Delta\phi_{jj}| \approx 1$ for di-lepton regions, there is only a very small number of events and the statistical error is the dominant component of the uncertainty band.



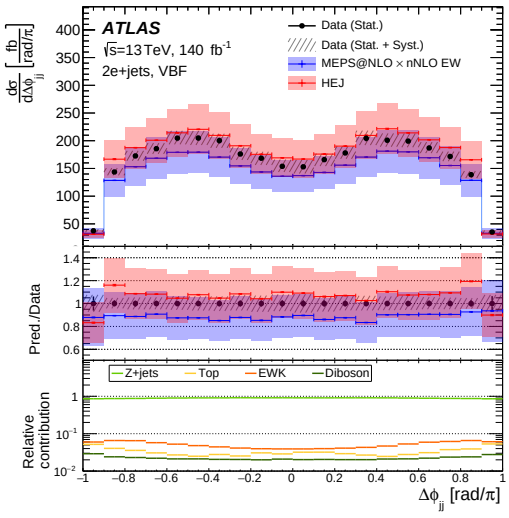
(a) $p_{\mathrm{T}}^{\mathrm{miss}}$+jets



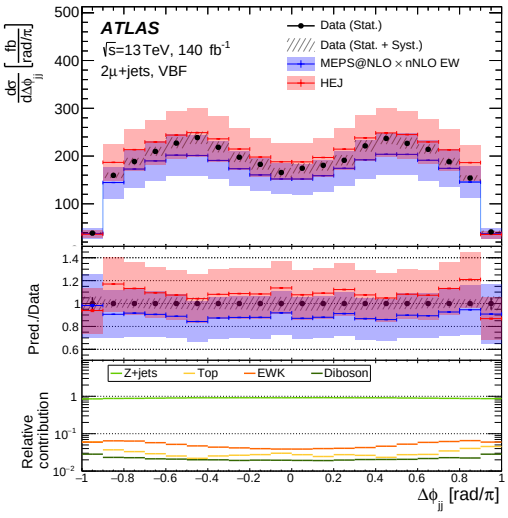
(b) $1e$+jets



(c) 1μ+jets

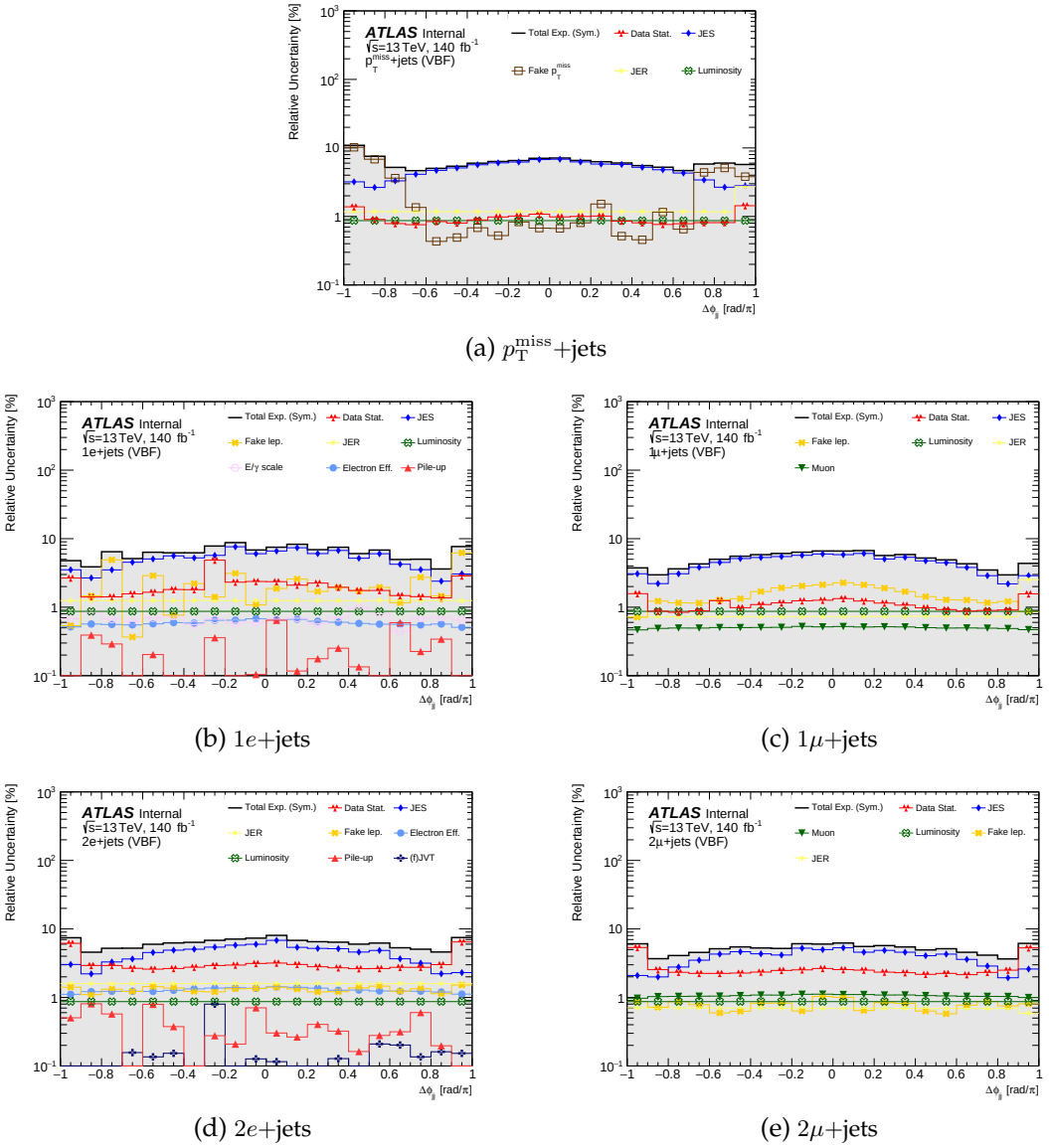


(d) $2e$+jets



(e) 2μ+jets

Figure 5.34: Particle-level comparisons between unfolded data and MC for $p_{\mathrm{T}}^{\mathrm{miss}}$ in the VBF phase space. The middle panel shows the ratios of the predictions to the data, with statistical uncertainties as solid markers and shaded bands to indicate the combined statistical and systematic uncertainties. The lower panels show the relative contributions from different SM processes relative to the total MEPS@NLO prediction. Figures from Ref. [232].



(a) $p_{\mathrm{T}}^{\mathrm{miss}}$+jets

(b) $1e$+jets

(c) 1μ+jets

(d) $2e$+jets

(e) 2μ+jets

Figure 5.35: Relative contribution from different systematic uncertainty components at particle-level for $\Delta\phi_{jj}$ in the VBF phase space. Only systematics with a contribution of $\geq 1\%$ in at least one of the bins are shown in the figure. For illustrative purposes this figure shows the symmetrised uncertainties, calculated as the average of the asymmetric error in each bin. Total Exp. (Sym.) is the combination of statistical and systematic uncertainties and indicates the symmetrised total experimental uncertainty. In subfigure (a) the fake $p_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty becomes dominant source of uncertainty in the region $|\Delta\phi_{jj}| \approx 1$, as the relative contribution from fake $p_{\mathrm{T}}^{\mathrm{miss}}$ is high which is directly correlated to the size of the error bar. Any asymmetries in the fake yield, that are relatively small as can be seen in corresponding Figure 5.16a, are therefore amplified in the uncertainties as can be seen in this plot.

5.7.2 Unfolded, differential R^{miss} ratios

The ratios R^{miss} are constructed after unfolding with the $p_{\text{T}}^{\text{miss}}$+jets signal region in the numerator and the auxiliary lepton regions in the denominator.

The bootstrap replicas from the separate differential distributions are used to calculate the statistical uncertainties on the combined R^{miss} ratio, but are also used to keep track of the correlations between the R^{miss} ratios and the histograms in the numerator of denominator.

Systematic uncertainties are propagated to R^{miss} by calculating the ratio for each nuisance parameter first, and then evaluating the difference with the nominal R^{miss} value similar to Equation 5.18. Here the cancellation of correlated uncertainties between the SR and the lepton regions becomes clear. If a nuisance parameter increases the differential cross-section in a given bin by m% and n% in the signal region and lepton region respectively, the relative difference with respect to the R^{miss} nominal will be n/m, which dissipates completely if this nuisance parameter is fully correlated between these two bins, i.e. $n = m$.

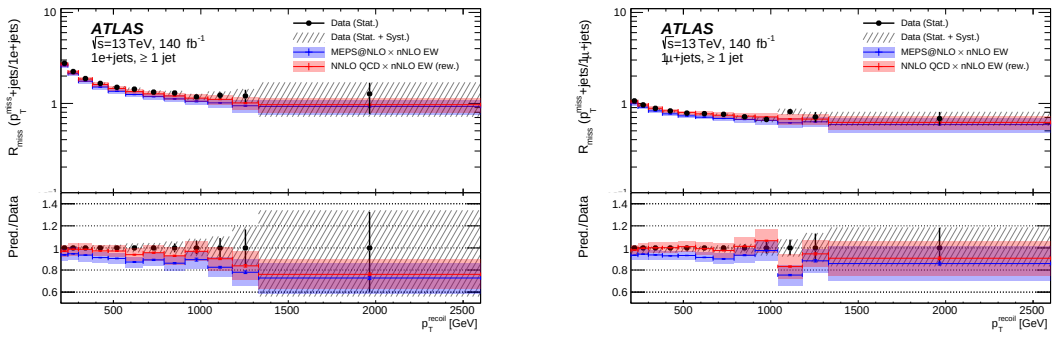
Comparisons between R^{miss} ratios for unfolded data and MC at the particle-level for $p_{\text{T}}^{\text{recoil}}$ in the ≥ 1 jet topology are shown in Figure 5.36. As indicated in Section 5.7.1, differential cross-sections in the single-lepton regions are the same order of magnitude as the signal region and the di-lepton channels are approximately an order of magnitude smaller, which is reflected in the values of R^{miss}. Furthermore, we observe that the size of the ratio decreases with $p_{\text{T}}^{\text{recoil}}$. This can be understood because the efficiency for the lepton reconstruction is higher at large $p_{\text{T}}^{\text{recoil}}$, increasing the acceptance and therefore the fiducial cross-sections relative to the $p_{\text{T}}^{\text{miss}}$+jets region. A better agreement between data and SM predictions, for both the MEPS@NLO and NNLO samples, is observed as the mismodelling of the total cross-section cancels out.

A breakdown of the R^{miss} uncertainties on the measured data is shown in Figure 5.37. The dominant jet uncertainties, shared between the signal region and the lepton regions, cancel in the ratio. For example, the jet energy scale (JES) uncertainty is approximately 1% in the low $p_{\text{T}}^{\text{recoil}}$ region for R^{miss}, while it can be up to 3-6% for the separate fiducial cross-sections. Uncertainties from nuisance parameters that are not correlated between the numerator or denominator in R^{miss}, for example from fake contributions or lepton efficiencies, do not cancel and therefore become the leading uncertainties together with the statistical uncertainty.

Similar R^{miss} plots as a function of $p_{\text{T}}^{\text{recoil}}$ measured in the VBF region is shown in Figure 5.38 and Figure 5.39. In these plots we observe same main features as in the ≥ 1 jet region: a cancellation of jet uncertainties and the ratio decreasing as a function of $p_{\text{T}}^{\text{recoil}}$ due to different efficiencies in $p_{\text{T}}^{\text{miss}}$+jets and lepton regions. Because the HEJ prediction is not available for the $p_{\text{T}}^{\text{miss}}$+jets signal region, no R^{miss} ratio is constructed from HEJ predictions in these plots.

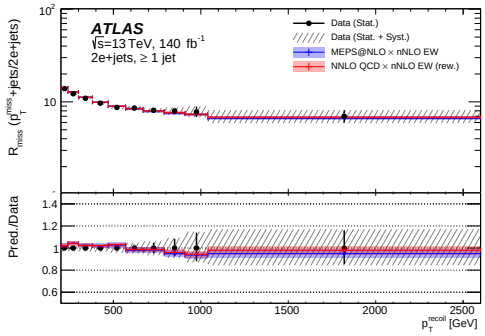
Results for R^{miss} as a function of m_{jj} in the VBF region are shown in Figure 5.40 and 5.41. Unlike the ratios for $p_{\text{T}}^{\text{recoil}}$, the ratio does not decrease as a function of m_{jj} because the lepton reconstruction efficiency does not necessarily improve as a function of m_{jj} like it does for $p_{\text{T}}^{\text{recoil}}$. The mismodelling of the shape of m_{jj} in the individual regions cancels out in R^{miss}, where we observe a flat ratio between MC and data as a function of m_{jj}.

Results for R^{miss} as a function of $\Delta\phi_{jj}$ in the VBF region are shown in Figure 5.42 and 5.43. Like in the individual cross-section measurements, the ratio between data and MC is flat across the

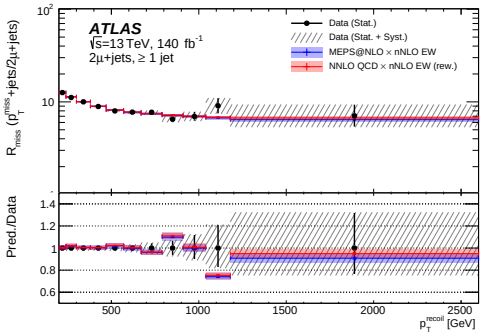


(a) R^{miss} for $1e$+jets

(b) R^{miss} for 1μ+jets

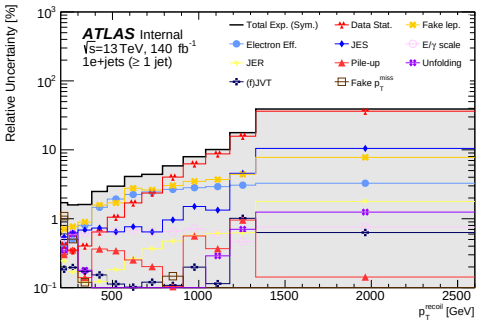


(c) R^{miss} for $2e$+jets

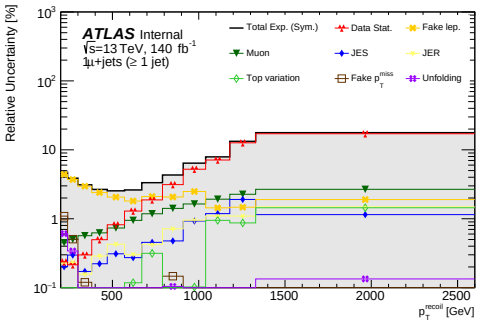


(d) R^{miss} for 2μ+jets

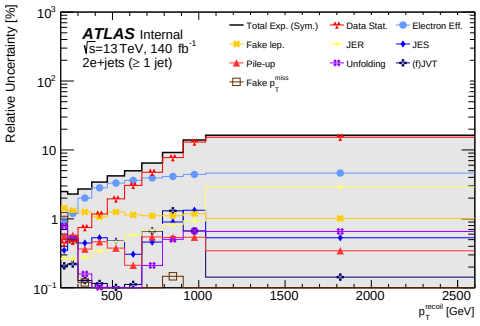
Figure 5.36: Particle-level R^{miss} ratios for unfolded data and MC as a function of $p_{\text{T}}^{\text{recoil}}$ in the ≥ 1 jet region. There is a good agreement between data and MC because mismodelling of the total cross-section of individual regions cancels out in the ratio between the $p_{\text{T}}^{\text{miss}}$+jets signal region and the lepton regions. The ratios decrease as a function of $p_{\text{T}}^{\text{recoil}}$ because the efficiency for the lepton reconstruction is higher at large $p_{\text{T}}^{\text{recoil}}$, increasing the acceptance and therefore the fiducial cross-sections relative to the $p_{\text{T}}^{\text{miss}}$+jets region. Absolute values of the ratio reflect the fact that the fiducial cross-section for the single-lepton regions is higher than those of the di-lepton regions. Figures from Ref. [232].



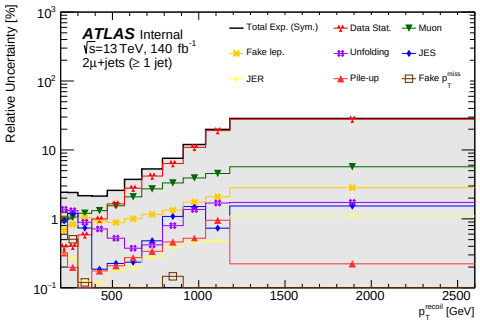
(a) R^{miss} for $1e$+jets



(b) R^{miss} for 1μ+jets



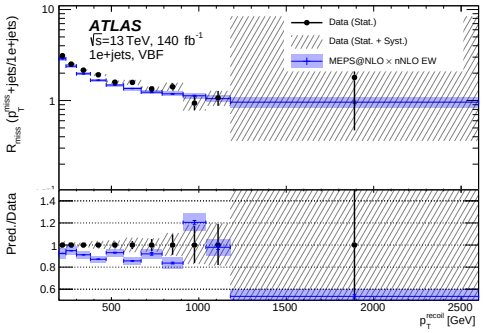
(c) R^{miss} for $2e$+jets



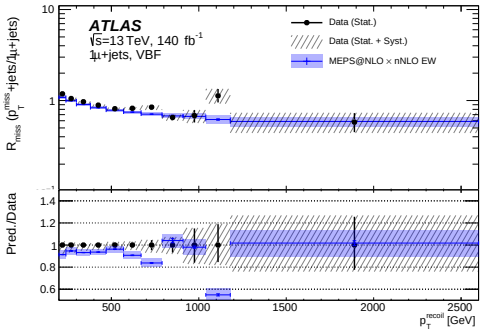
(d) R^{miss} for 2μ+jets

Figure 5.37: Relative contribution from different systematic uncertainty components at particle-level for R^{miss} as a function of $p_{\text{T}}^{\text{recoil}}$ in the ≥ 1 jet phase space. The ratios R^{miss} are constructed to leverage a cancellation of uncertainties, which is observed particularly for JER and JES systematics, when comparing these relative uncertainties to the separate cross-section measurements. Because the R^{miss} ratios "mix" lepton regions with the $p_{\text{T}}^{\text{miss}}$+jets signal region, there are both (fake) lepton uncertainties as well as fake $p_{\text{T}}^{\text{miss}}$ uncertainties, but in the individual measurements these are largely subdominant to the jet uncertainties that cancel out. Like in the individual regions, statistical uncertainties become the dominant uncertainty component as $p_{\text{T}}^{\text{recoil}}$ increases.

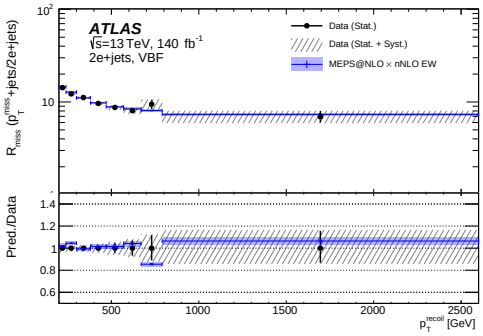
spectrum of $\Delta\phi_{jj}$. In the region $|\Delta\phi_{jj}| \approx 1$ the ratio increases slightly for the di-lepton regions, which indicates that the fiducial cross-section of this bin drops more steeply in di-lepton regions than in the signal region (and can be seen qualitatively in Figure 5.34). This effect is potentially the result of sub-leading jets that contribute to $p_{\mathrm{T}}^{\mathrm{miss}}$ but that escape the $\Delta\phi\left(\mathrm{jet}, p_{\mathrm{T}}^{\mathrm{miss}}\right)$ veto in the signal region. Such events can be vetoed if the Z boson decays into two muons due to the overlap removal procedure.



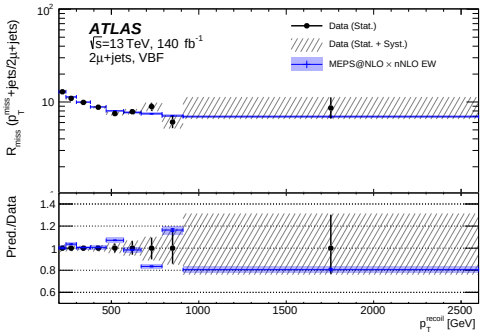
(a) R^{miss} for $1e$+jets



(b) R^{miss} for 1μ+jets

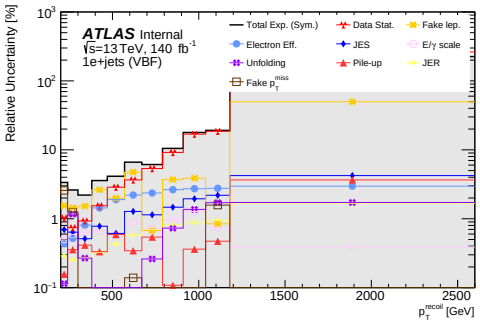


(c) R^{miss} for $2e$+jets

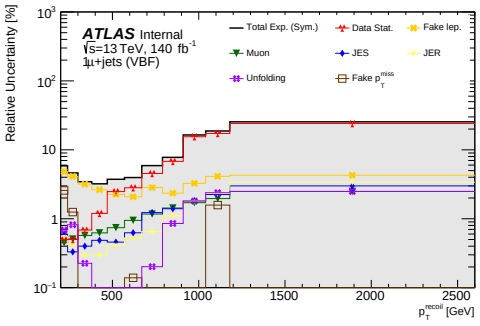


(d) R^{miss} for 2μ+jets

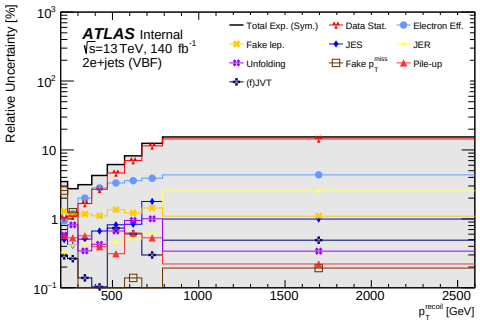
Figure 5.38: Particle-level R^{miss} ratios for unfolded data and MC as a function of $p_{\text{T}}^{\text{recoil}}$ in the VBF region. There is a good agreement between data and MC because mismodelling of the total cross-section of individual regions cancels out in the ratio between the $p_{\text{T}}^{\text{miss}}$+jets signal region and the lepton regions. The ratios decrease as a function of $p_{\text{T}}^{\text{recoil}}$ because the efficiency for the lepton reconstruction is higher at large $p_{\text{T}}^{\text{recoil}}$, increasing the acceptance and therefore the fiducial cross-sections relative to the $p_{\text{T}}^{\text{miss}}$+jets region. Absolute values of the ratio reflect the fact that the fiducial cross-section for the single-lepton regions is higher than those of the di-lepton regions. There is a outlier at approximately two standard deviations in subfigure (b). This particular bin, centred at $1180\,\text{GeV}$ is under-predicted in $p_{\text{T}}^{\text{miss}}$+jets (in Figure 5.30a) while it is over-predicted in 1μ+jets (in Figure 5.30c). Figures from Ref. [232].



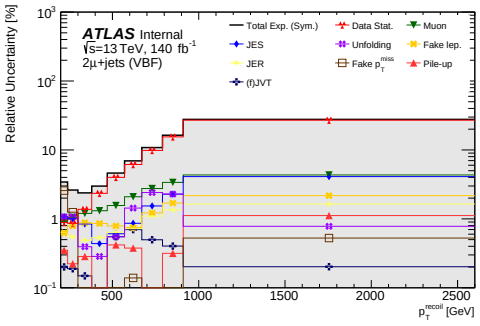
(a) R^{miss} for $1e$+jets



(b) R^{miss} for 1μ+jets

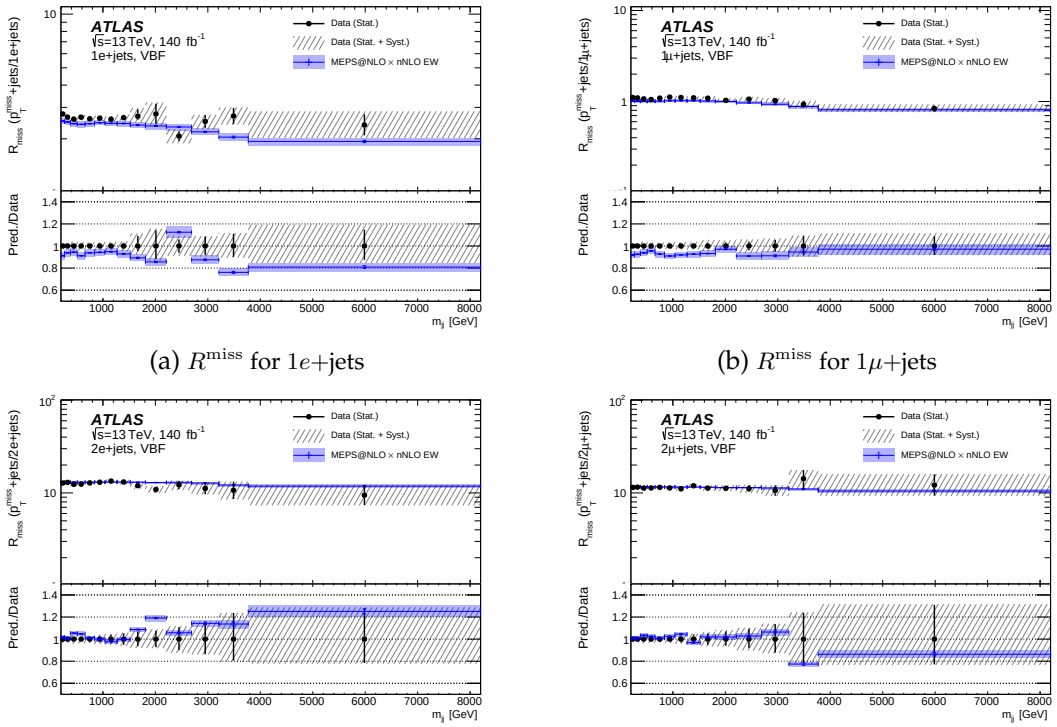


(c) R^{miss} for $2e$+jets



(d) R^{miss} for 2μ+jets

Figure 5.39: Relative contribution from different systematic uncertainty components at particle-level for R^{miss} as a function of $p_{\text{T}}^{\text{recoil}}$ in the VBF phase space. The ratios R^{miss} are constructed to leverage a cancellation of uncertainties, which is observed particularly for JER and JES systematics, when comparing these relative uncertainties to the separate cross-section measurements. Because the R^{miss} ratios "mix" lepton regions with the $p_{\text{T}}^{\text{miss}}$+jets signal region, there are both (fake) lepton uncertainties as well as fake $p_{\text{T}}^{\text{miss}}$ uncertainties, but in the individual measurements these are largely subdominant to the jet uncertainties that cancel out. Like in the individual regions, statistical uncertainties become the dominant uncertainty component as $p_{\text{T}}^{\text{recoil}}$ increases.



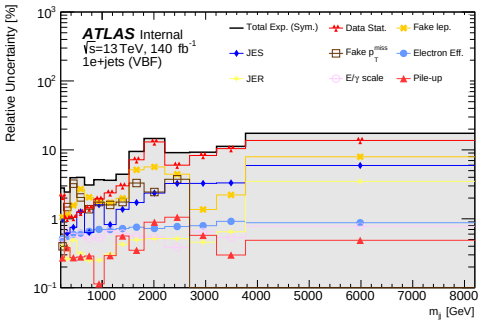
(a) R^{miss} for $1e$+jets

(b) R^{miss} for 1μ+jets

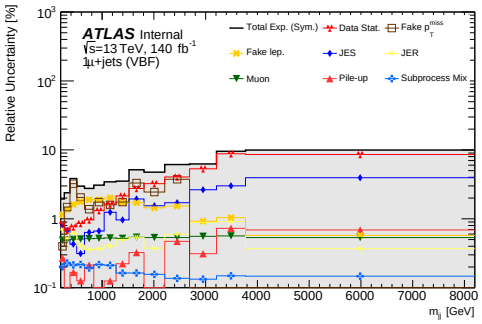
(c) R^{miss} for $2e$+jets

(d) R^{miss} for 2μ+jets

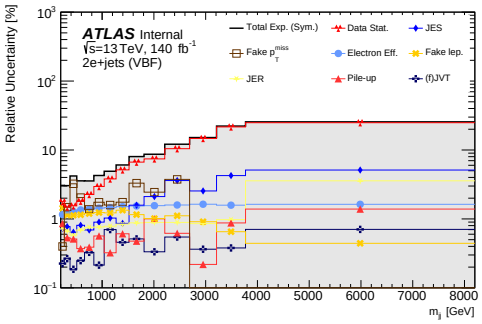
Figure 5.40: Particle-level R^{miss} ratios for unfolded data and MC as a function of m_{jj} in the VBF region. There is a good agreement between data and MC because shape mismodelling effects of the individual cross-sections cancel out in the ratio between the $p_{\text{T}}^{\text{miss}}$+jets signal region and the lepton regions. The ratio remains flat as a function of m_{jj} because there is no dependence of reconstruction efficiency between the signal region or the lepton regions. The absolute values of the ratio reflect the fact that the fiducial cross-section for the single-lepton regions is higher than those of the di-lepton regions. Figures from Ref. [232].



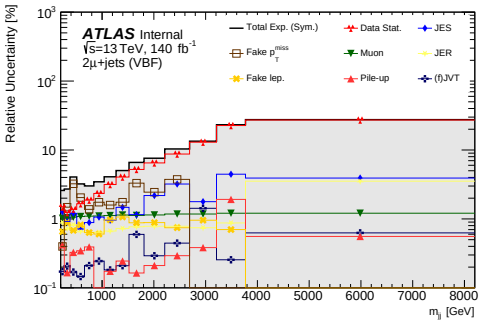
(a) R^{miss} for $1e$+jets



(b) R^{miss} for 1μ+jets

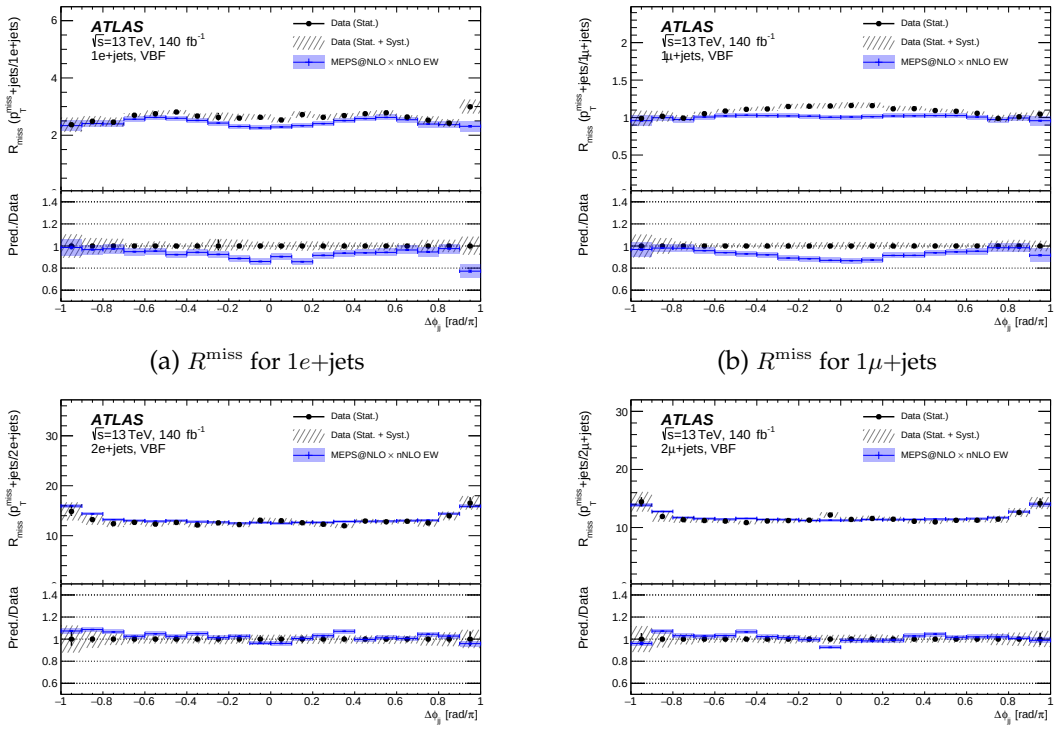


(c) R^{miss} for $2e$+jets



(d) R^{miss} for 2μ+jets

Figure 5.41: Relative contribution from different systematic uncertainty components at particle-level for R^{miss} as a function of m_{jj} in the VBF phase space. The ratios R^{miss} are constructed to leverage a cancellation of uncertainties, which is observed particularly for JER and JES systematics, when comparing these relative uncertainties to the separate cross-section measurements. Because the R^{miss} ratios "mix" lepton regions with the $p_{\text{T}}^{\text{miss}}$+jets signal region, there are both (fake) lepton uncertainties as well as fake $p_{\text{T}}^{\text{miss}}$ uncertainties, but in the individual measurements these are largely subdominant to the jet uncertainties that cancel out. Like in the individual regions, statistical uncertainties become the dominant uncertainty component as m_{jj} increases, and the fake background systematics can become dominant at low m_{jj}.



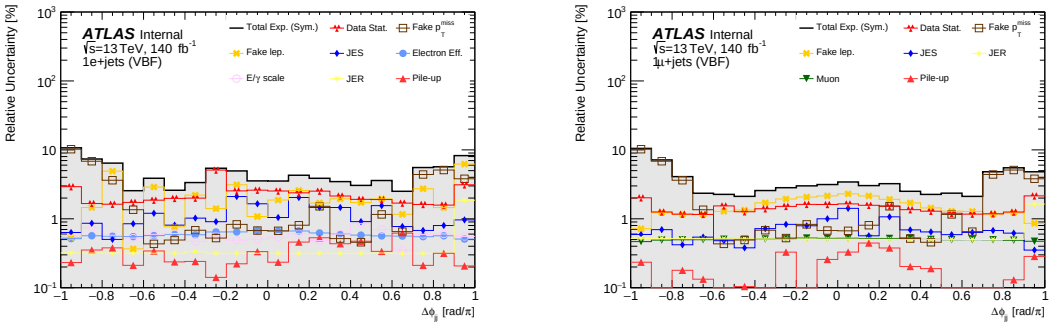
(a) R^{miss} for $1e$+jets

(b) R^{miss} for 1μ+jets

(c) R^{miss} for $2e$+jets

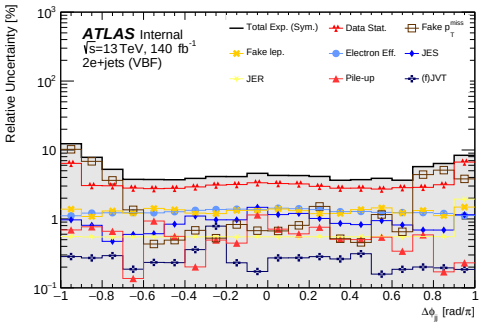
(d) R^{miss} for 2μ+jets

Figure 5.42: Particle-level R^{miss} ratios for unfolded data and MC as a function of $\Delta\phi_{jj}$ in the VBF region. Like in the individual cross-section measurement of the $p_{\text{T}}^{\text{miss}}$+jets signal region and the lepton regions there is a good modelling of data across $\Delta\phi_{jj}$. The absolute values of the ratio reflect the fact that the fiducial cross-section for the single-lepton regions is higher than those of the di-lepton regions. The fiducial cross-section of the single-electron region is lower than the single-muon region due to additional cuts on real $p_{\text{T}}^{\text{miss}}$ and the m_T of the W boson, leading to a higher ratio R^{miss} of figure (a) compared to subfigure (b). Figures from Ref. [232].

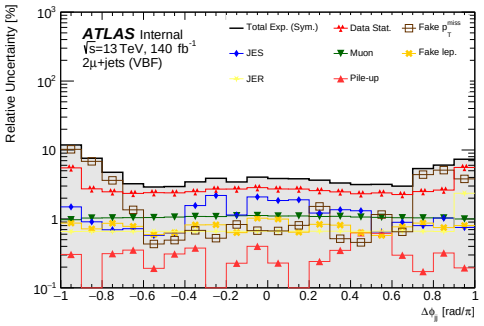


(a) R^{miss} for $1e$+jets

(b) R^{miss} for 1μ+jets



(c) R^{miss} for $2e$+jets



(d) R^{miss} for 2μ+jets

Figure 5.43: Relative contribution from different systematic uncertainty components at particle-level for R^{miss} as a function of $\Delta\phi_{jj}$ in the VBF phase space. The ratios R^{miss} are constructed to leverage a cancellation of uncertainties, which is observed particularly for JER and JES systematics, when comparing these relative uncertainties to the separate cross-section measurements. Because the R^{miss} ratios "mix" lepton regions with the $p_{\text{T}}^{\text{miss}}$+jets signal region, there are both (fake) lepton uncertainties as well as fake $p_{\text{T}}^{\text{miss}}$ uncertainties, but in the individual measurements these are largely subdominant to the jet uncertainties that cancel out. The fake $p_{\text{T}}^{\text{miss}}$ systematic uncertainty is dominant in the region $|\Delta\phi_{jj}| \approx 1$, while particularly the fake lepton uncertainties are important across the spectrum in the single-lepton regions.

5.7.3 Consistency checks between inclusive and dominant boson process unfolding

In this section, a validation study is presented to verify that the unfolded data from the inclusive measurement, the target final state (TFS), and dominant boson process (DBP) mechanisms are self-consistent. To first order, the sum of truth-level subdominant SM processes and unfolded DBP data should add up to the same differential cross-section as the TFS unfolded data – because the subdominant SM processes are precisely what was subtracted before unfolding in the DBP approach. In other words, one expects the following ratio to be consistent with unity:

$$\mathcal{R}_{\text{TFS}}^{\text{DBP}} \equiv \frac{\text{Data}_{\text{DBP}}^{\text{unfolded}} + \text{MC}^{\text{subdom. SM proc.}}}{\text{Data}_{\text{TFS}}^{\text{unfolded}}} \tag{5.21}$$

It should be noted that because these two approaches have a different admixture of SM processes, the detector response of the MC which is used in the unfolding differs between the DBP and TFS methods. This may cause slight a slight offset from one in the ratio. The background composition is constrained up to a set of scale factors with a given uncertainty as described in section 5.3.5. When taking these uncertainties into account, together with the statistical uncertainties on data and MC, any deviations from one should fall within these error bands.

Results for $p_{\text{T}}^{\text{miss}}$, $\Delta\phi_{jj}$ and m_{jj} are shown in Figures 5.44–5.47. The top panel shows the both the DBP and TFS unfolded data individually, the subdominant SM process predictions at particle-level and the subdominant SM process predictions added to the DBP unfolded data. The middle panel shows the ratio $\mathcal{R}_{\text{TFS}}^{\text{DBP}}$. The uncertainty bands consists of three components. First, it includes the systematic uncertainties relating to the composition of the MC admixture, which is applied to the unfolded DBP data, unfolded TFS data and the subdominant theory stack. Second, statistical uncertainties on the MC prediction are included. Lastly, statistical uncertainties on the unfolded data are included. Because the DBP and TFS unfolded data are correlated, a bootstrapping method is used to account for these correlations. The uncertainty band for the $p_{\text{T}}^{\text{miss}}$+jets, $1e$+jets and 1μ+jets regions is dominated by the composition uncertainty in for low values of $p_{\text{T}}^{\text{recoil}}$ and m_{jj}. In the tails, the statistical uncertainties on the MC and the statistical uncertainty from the bootstrap replicas of the data become dominant.

There is good closure in the middle panel within systematic and statistical uncertainties for all observables. This shows consistency between unfolding of an inclusive, final-state particle measurement and a background-subtracted process measurement, where any discrepancies from having a different subprocess mix are included as systematic uncertainties.

Bottom panels show a comparison between the TFS and DBP unfolded data, to give an estimate of the relative size of the dominant SM process at truth-level. The di-lepton measurements have a high dominant process purity. This is reflected as the ratio in the bottom panel, $\text{Data}_{\text{DBP}}^{\text{unfolded}}/\text{Data}_{\text{TFS}}^{\text{unfolded}}$, being closer to one than the other regions. It also means that the detector response for the TFS approach is very similar to the DBP method, such that the nominal values of $\mathcal{R}_{\text{TFS}}^{\text{DBP}}$ are expected to be closer to one than the signal region or single-lepton regions, which is indeed observed in this test. The uncertainties on the ratio are smaller than for the signal region and single-lepton regions because the systematic for the composition uncertainty is not applied here. This check also shows good closure in the $2e$+jets and 2μ+jets regions.

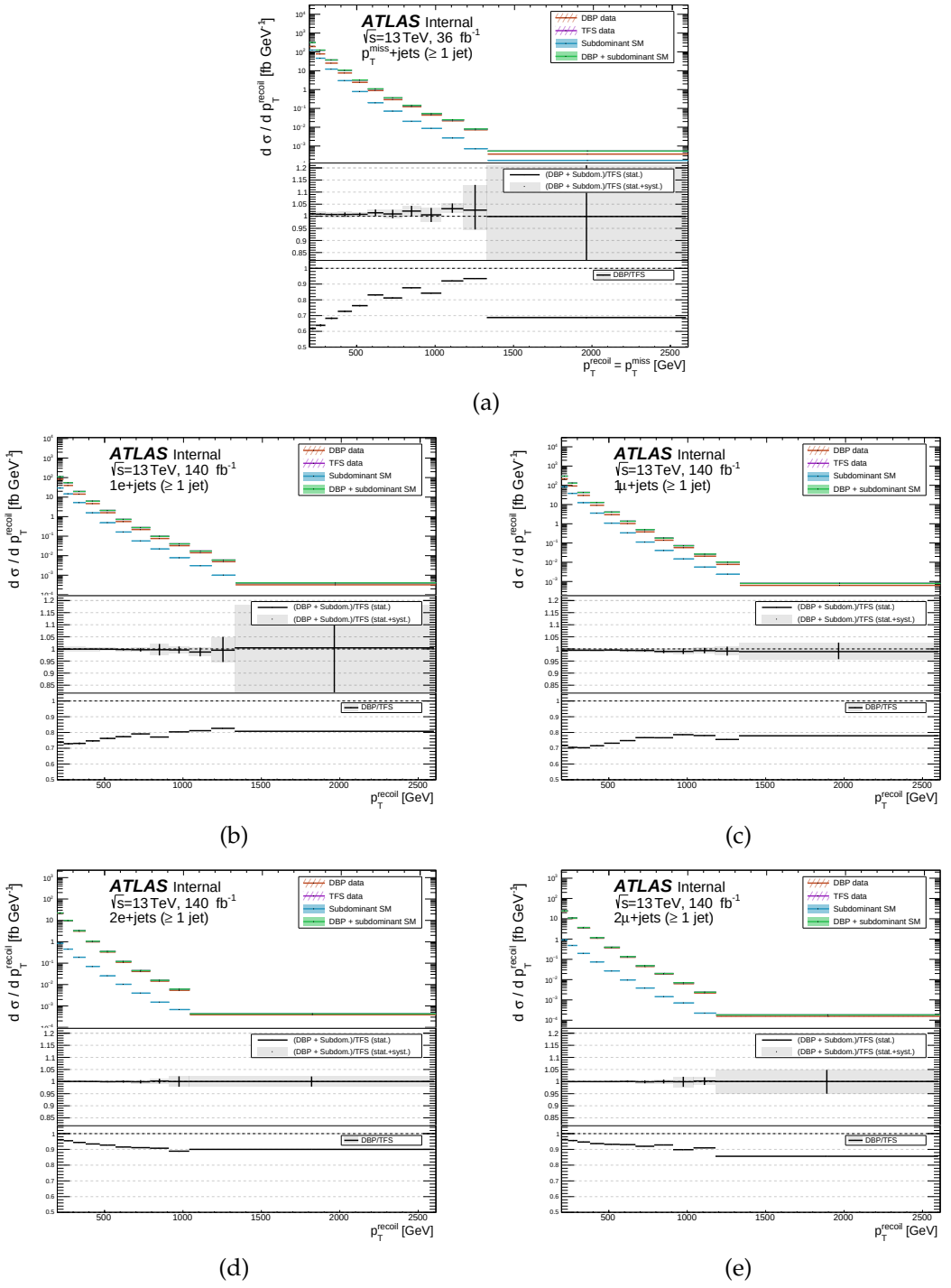


Figure 5.44: Comparison between the DBP and TFS unfolded data the $p_{\mathrm{T}}^{\mathrm{miss}}$ observable for the ≥ 1 jet phase space. The top panel shows the unfolded, background-subtracted data in red and the unfolded, inclusive data in purple. The particle-level MC distributions that correspond to subdominant SM contributions that are subtracted at detector-level in the former, are indicated in blue. The sum of the unfolded, background-subtracted data and the subdominant SM prediction is shown in green. The middle panel shows the ratio between the green and purple curves. This ratio is expected to be one if the inclusive unfolding versus the background-subtracted unfolding are consistent with each other despite having a different subprocess mix that is used in the unfolding procedure, c.f. Equation 5.21. This is the case within the uncertainty band, which is constructed from the statistical uncertainty and systematic uncertainties associated with the subprocesses mix. Statistical correlations are taken into account when deriving this ratio. The bottom panel shows the ratio between the inclusive and background-subtracted data. Deviations from one give a measure of what fraction makes up the "subdominant background" processes. The two-lepton channels in subfigure (d) and (e) have the highest signal purity.

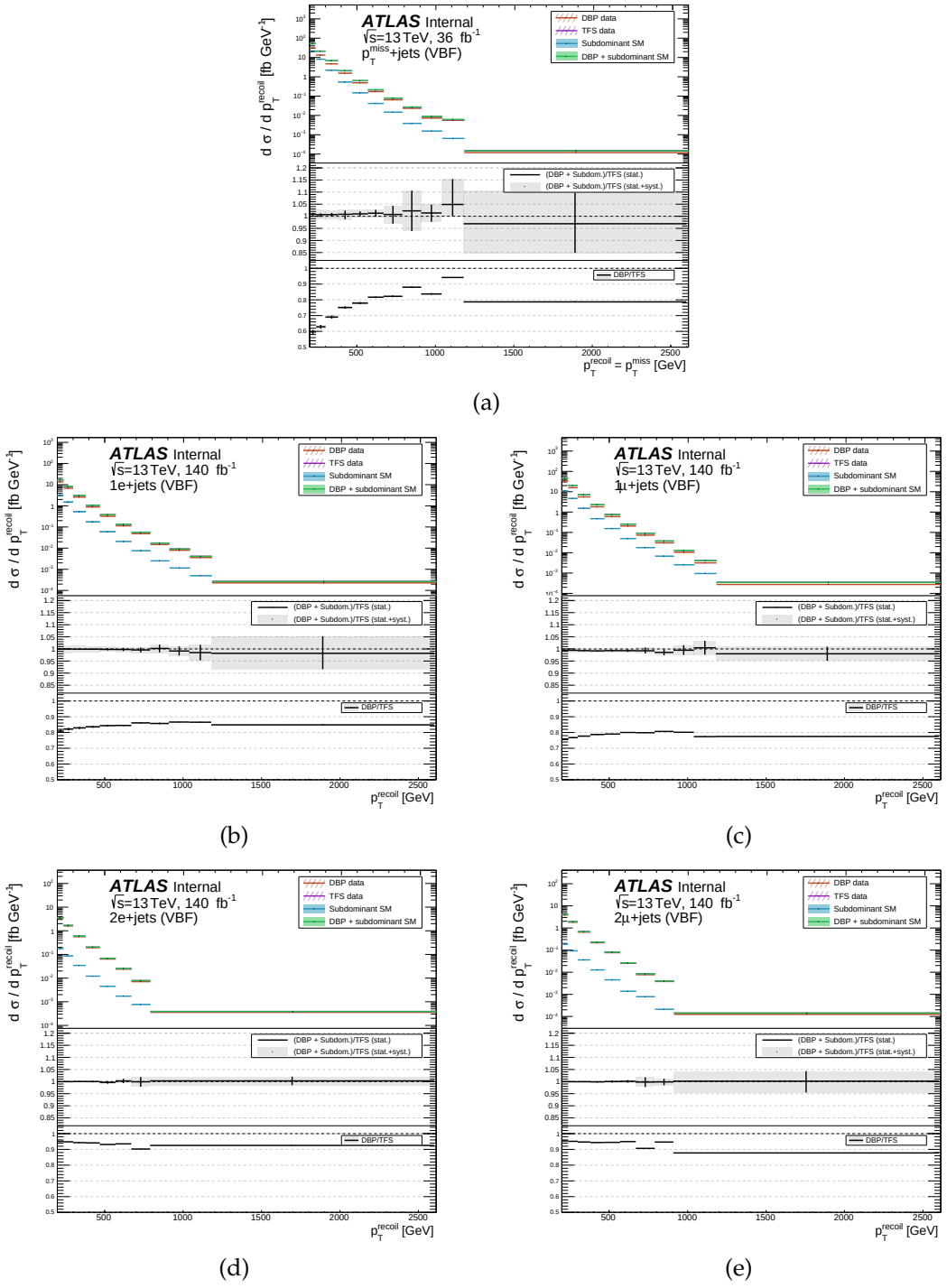


Figure 5.45: Comparison between the DBP and TFS unfolded data the $p_{\mathrm{T}}^{\mathrm{miss}}$ observable for the VBF phase space. The top panel shows the unfolded, background-subtracted data in red and the unfolded, inclusive data in purple. The particle-level MC distributions that correspond to subdominant SM contributions that are subtracted at detector-level in the former, are indicated in blue. The sum of the unfolded, background-subtracted data and the subdominant SM prediction is shown in green. The middle panel shows the ratio between the green and purple curves. This ratio is expected to be one if the inclusive unfolding versus the background-subtracted unfolding are consistent with each other despite having a different subprocess mix that is used in the unfolding procedure, c.f. Equation 5.21. This is the case within the uncertainty band, which is constructed from the statistical uncertainty and systematic uncertainties associated with the subprocesses mix. Statistical correlations are taken into account when deriving this ratio. The bottom panel shows the ratio between the inclusive and background-subtracted data. Deviations from one give a measure of what fraction makes up the "subdominant background" processes. The two-lepton channels in subfigure (d) and (e) have the highest signal purity.

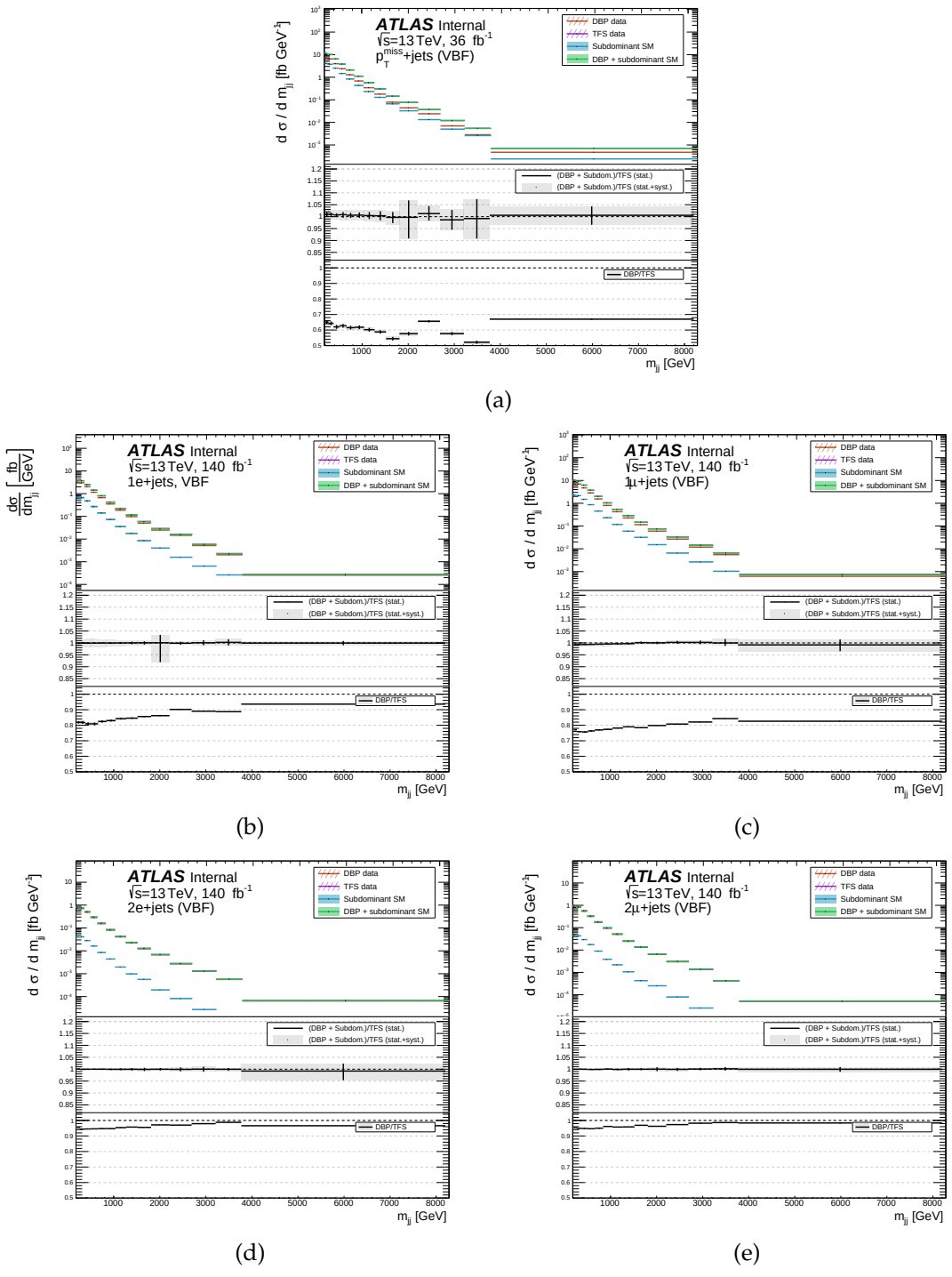


Figure 5.46: Comparison between the DBP and TFS unfolded data the m_{jj} observable for the VBF phase space. The top panel shows the unfolded, background-subtracted data in red and the unfolded, inclusive data in purple. The particle-level MC distributions that correspond to subdominant SM contributions that are subtracted at detector-level in the former, are indicated in blue. The sum of the unfolded, background-subtracted data and the subdominant SM prediction is shown in green. The middle panel shows the ratio between the green and purple curves. This ratio is expected to be one if the inclusive unfolding versus the background-subtracted unfolding are consistent with each other despite having a different subprocess mix that is used in the unfolding procedure, c.f. Equation 5.21. This is the case within the uncertainty band, which is constructed from the statistical uncertainty and systematic uncertainties associated with the subprocesses mix. Statistical correlations are taken into account when deriving this ratio. The bottom panel shows the ratio between the inclusive and background-subtracted data. Deviations from one give a measure of what fraction makes up the "subdominant background" processes. The two-lepton channels in subfigure (d) and (e) have the highest signal purity.

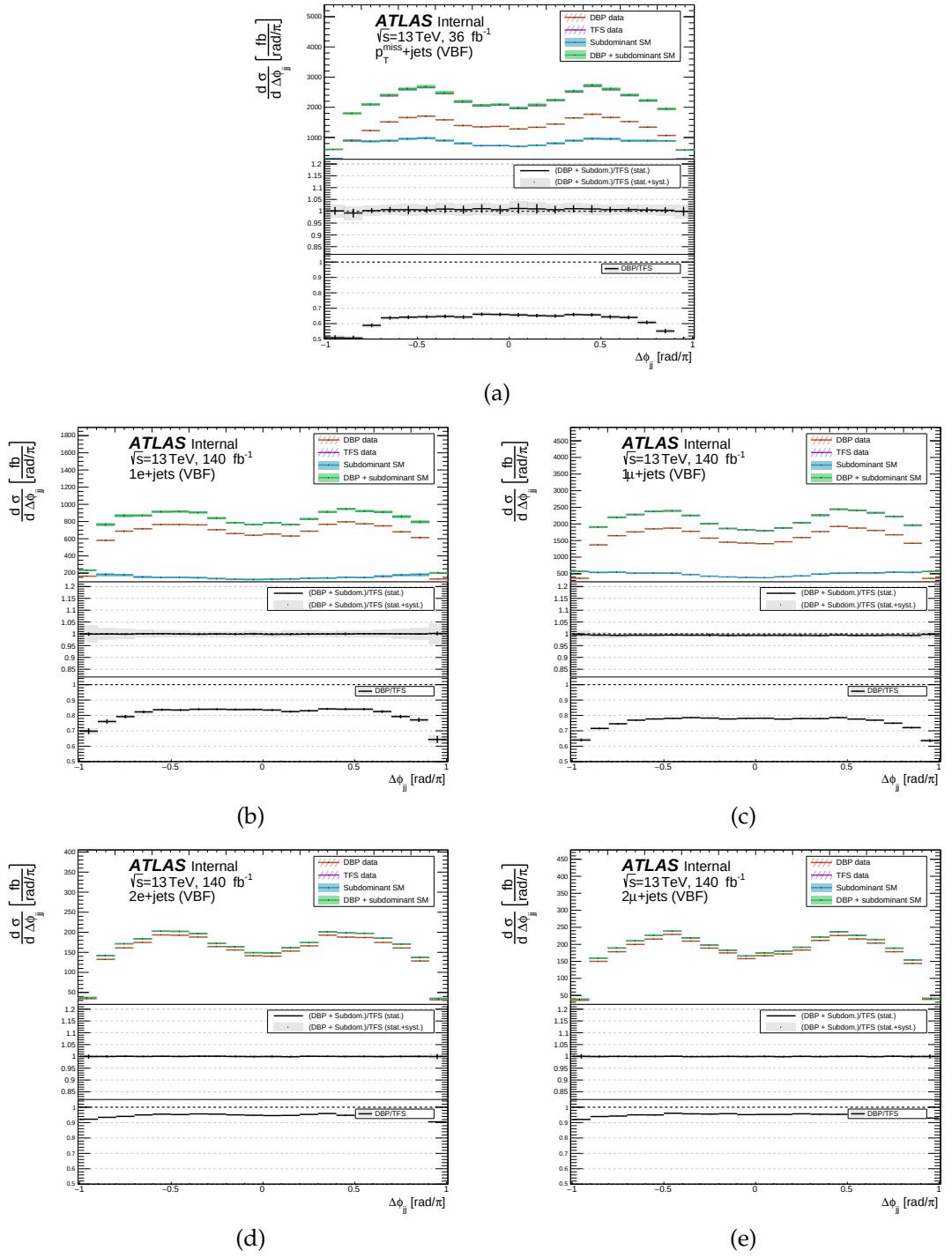


Figure 5.47: Comparison between the DBP and TFS unfolded data the $\Delta\phi_{jj}$ observable for the VBF phase space. The top panel shows the unfolded, background-subtracted data in red and the unfolded, inclusive data in purple. The particle-level MC distributions that correspond to subdominant SM contributions that are subtracted at detector-level in the former, are indicated in blue. The sum of the unfolded, background-subtracted data and the subdominant SM prediction is shown in green. The middle panel shows the ratio between the green and purple curves. This ratio is expected to be one if the inclusive unfolding versus the background-subtracted unfolding are consistent with each other despite having a different subprocess mix that is used in the unfolding procedure, c.f. Equation 5.21. This is the case within the uncertainty band, which is constructed from the statistical uncertainty and systematic uncertainties associated with the subprocesses mix. Statistical correlations are taken into account when deriving this ratio. The bottom panel shows the ratio between the inclusive and background-subtracted data. Deviations from one give a measure of what fraction makes up the "subdominant background" processes. The two-lepton channels in subfigure (d) and (e) have the highest signal purity.

5.7.4 Lepton flavour universality check in dilepton channels

Lepton flavour universality requires that the fiducial cross-sections for $1e$+jets and $2e$+jets should be consistent with the 1μ+jets and 2μ+jets cross-sections respectively. This can be verified directly with the unfolded cross-sections presented in Section 5.7. The catch for the single-lepton regions is that their fiducial phase-space is significantly different due to the m_T and real $p_{\mathrm{T}}^{\mathrm{miss}}$ cut for the electron region. Therefore, this section only shows the results of a cross-check comparing the unfolded data from the 2μ+jets region to the $2e$+jets region. To make an apples-to-apples comparison between the $Z \to ee$ and $Z \to \mu\mu$ production cross-sections and to remove effects from electroweak, top or diboson processes that contribute in these regions, results from the dominant boson process unfolding are used in this test.

The only difference in the fiducial phase-space definitions for 2μ+jets and $2e$+jets is the lepton rapidity cut. To account for this, simulation-based scale factors are derived for each bin and applied to the unfolded data. The corrections are derived by running the analysis code for the 2μ+jets region once with the pseudorapidity cut of $|\eta| < 2.5$ and once with $0 < |\eta| < 1.37$, $1.52 < |\eta| < 2.47$ – corresponding to the cuts for the 2μ+jets and $2e$+jets regions respectively. The ratio between the resulting histograms gives an estimate of the enlarged phase-space of muons compared to electrons due to a larger acceptance in lepton pseudorapidity. This ratio is applied as bin-by-bin scale factors to the muon yields when these are compared to the electron counterparts.

Figure 5.48 shows the unfolded data for the $2e$+jets and 2μ+jets from the DBP approach for the $p_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\phi_{jj}$ and m_{jj} observables. The bottom panel shows the ratio between the electron and muons channels, indicating that the dilepton regions are consistent for the two different jet topologies - monojet and VBF - across the unfolded observables. The error bar includes statistical uncertainties derived from the bootstrap method and the experimental lepton systematics.

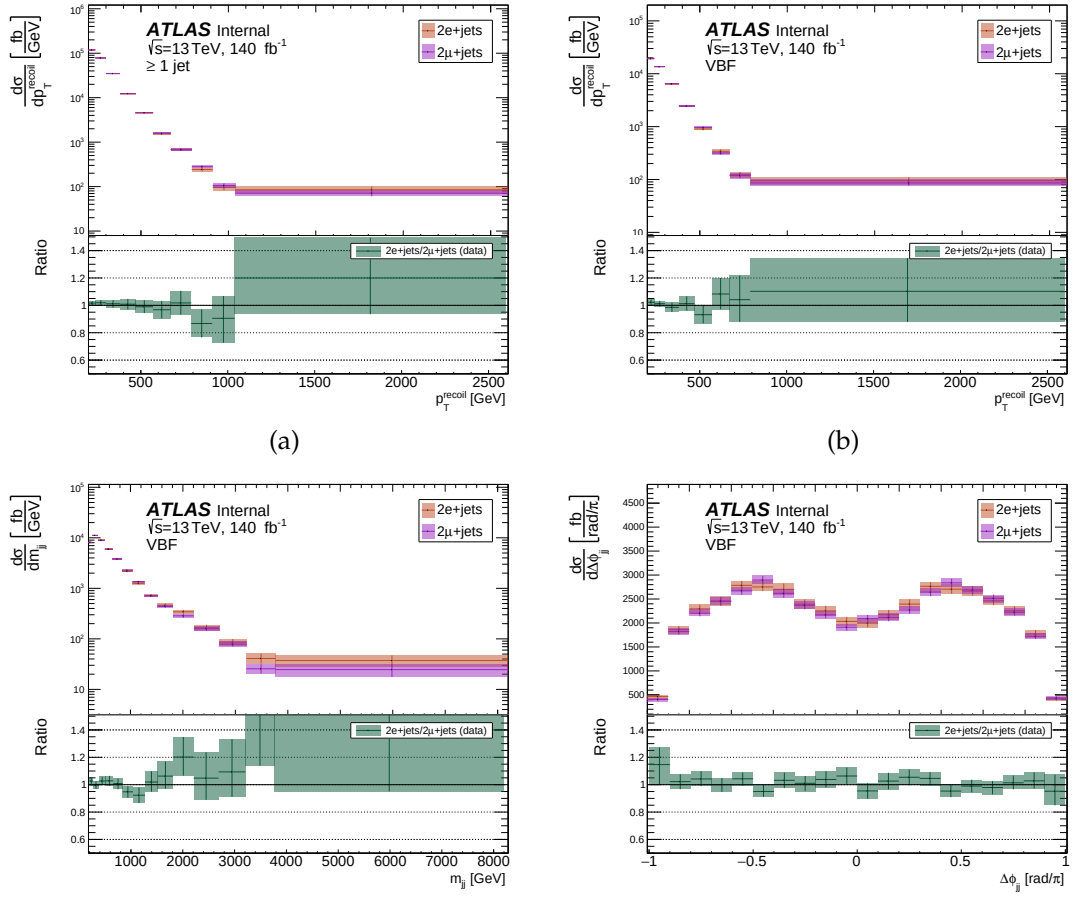
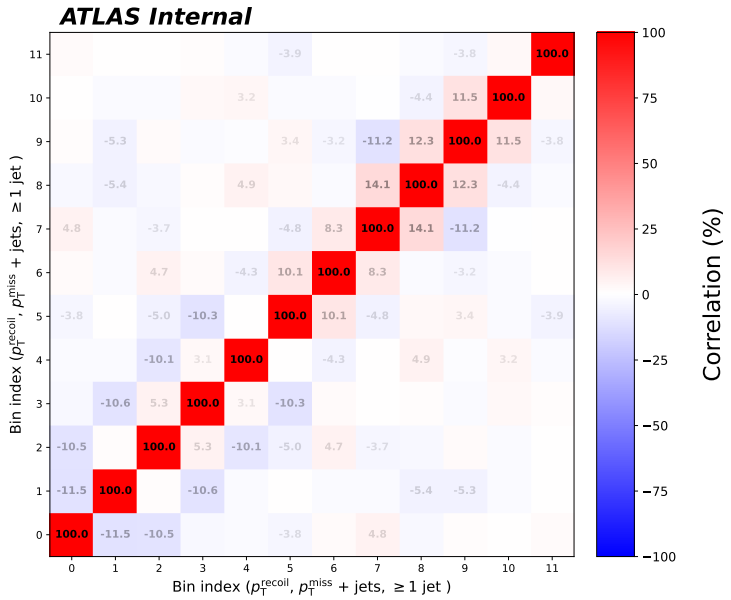


Figure 5.48: Comparison between the $2e$+jets and 2μ+jets region for the (a) $p_{\mathrm{T}}^{\mathrm{recoil}}$ in the ≥ 1 jet, (b) $p_{\mathrm{T}}^{\mathrm{recoil}}$, (c) m_{jj}, (d) $\Delta\phi_{jj}$ in the VBF phase-space. The uncertainty bands include the statistical uncertainties derived using the bootstrap method and the experimental lepton systematics. Lepton flavour universality dictates that the ratio in the bottom panel should be equal to one, but this test is limited by correction factors to account for different experimental reconstruction of electrons and muons. Nonetheless, good closure is observed in almost all bins of the measured distributions.



(a)

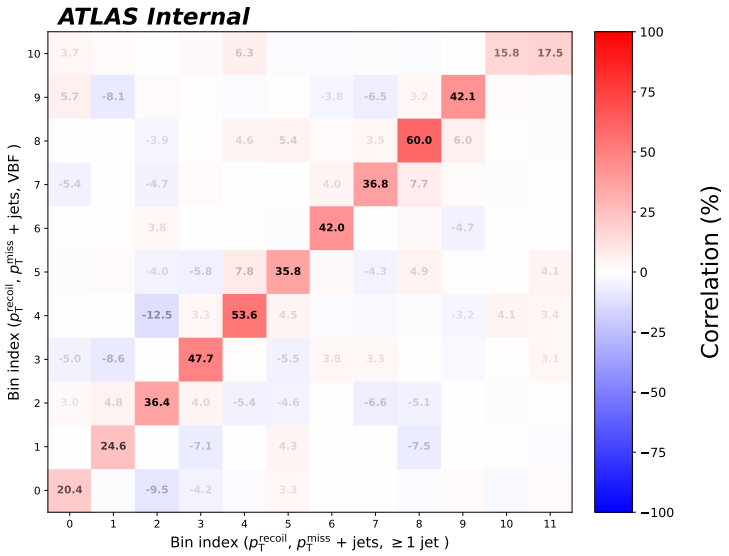
Figure 5.49: Statistical correlations between bins in the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region of the ≥ 1 jet phase-space. The correlation matrix is not completely diagonal, because events can migrate to different bins between detector-level and particle-level during the unfolding procedure. The size of the off-diagonal elements captures the size of the bin migrations from unfolding, while the diagonal elements are by construction fully correlated (100%) as they correspond to the same bin on the horizontal and vertical axes. Correlation matrices for the auxiliary lepton regions and the observables in the VBF region exhibit the same main features and are published with the HEPData entry of the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets analysis in Ref. [237].

5.7.5 Statistical correlations

With the bootstrap method, statistical correlations between any combination of observable, phase-space and jet topology can be constructed. This section shows a selection of the correlation matrices that are published on HEPData in Ref. [237] for the purpose of re-interpretation.

Figure 5.49 shows the statistical correlation *within* the unfolded $p_{\mathrm{T}}^{\mathrm{recoil}}$ distributions for the ≥ 1 jet region. The diagonal of these correlation matrices are 100% by construction, because these indicate the same bin on the horizontal and vertical axes. At detector-level, there are no statistical correlations between different bins of the measured the data: an event is in one bin or another. However, during unfolding events can migrate from one bin in detector-level and to another bin at particle-level. The way that events migrate between particle-level bins is captured in the (non-zero) off-diagonal values observed in these correlation matrices. These "self-correlations" are published on HEPData for every measured histogram, including the R^{miss} ratios.

By construction, the fiducial event-selections in Table 5.2 are not mutually exclusive between ≥ 1 jet and VBF. This can lead to correlations between $p_{\mathrm{T}}^{\mathrm{miss}}$ in the ≥ 1 jet region to $p_{\mathrm{T}}^{\mathrm{miss}}$, m_{jj} and $\Delta\phi_{jj}$ in the VBF region. The statistical correlations are strongest between the $p_{\mathrm{T}}^{\mathrm{miss}}$ observables for both regions, which is shown in Figure 5.50 for the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region. Such correlation matrices contain important information when multiple histograms are used in statistical combinations for either a SM fit or BSM limit setting. If there is some effect, e.g. mismodelling of the cross-section or a strong BSM signal, that impacts strongly correlated bins, there is a risk of accounting for



(a)

Figure 5.50: Statistical correlations between $p_{\mathrm{T}}^{\mathrm{miss}}$ in the ≥ 1 jet phase-space (horizontal axis) and $p_{\mathrm{T}}^{\mathrm{miss}}$ in the VBF phase-space (vertical axis), both measured in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets signal region. Table 5.2 shows that the phase-space definitions of the two different jet topologies are not fully orthogonal, such that a subset of events passes both ≥ 1 jet and VBF event selections. As a result, observables in the two jet topologies can be correlated. As seen in this figure, the correlation between $p_{\mathrm{T}}^{\mathrm{miss}}$ of the two observables is strongest along the diagonal, which corresponds to the same binning in $p_{\mathrm{T}}^{\mathrm{miss}}$ (except for bin 10 and 11). Other combinations of $p_{\mathrm{T}}^{\mathrm{miss}}$, m_{jj} and $\Delta\phi_{jj}$ in the ≥ 1 jet and VBF topologies are less correlated, with the corresponding correlation matrices published in Ref. [237] for re-interpretation studies.

this effect twice. By including the correlations in the fit, the effect from overlap in phase-space between different distributions is mitigated. The $p_{\mathrm{T}}^{\mathrm{miss}}$ observable has largely the same binning in the ≥ 1 jet region and the VBF region , except for the largest $p_{\mathrm{T}}^{\mathrm{miss}}$ bin in VBF which is slightly larger to ensure a good number of events in an otherwise statistically limited bin. For each event the real $p_{\mathrm{T}}^{\mathrm{miss}}$ component is calculated *once* and will be the same in both the ≥ 1 jet and VBF region. Consequently, we expect any correlations to be largest along the diagonal of this correlation matrix, which is indeed observed. In this work, we only highlight this correlation matrix because the correlations between other combinations of $p_{\mathrm{T}}^{\mathrm{miss}}$, m_{jj} and $\Delta\phi_{jj}$ in the ≥ 1 jet and VBF topologies are smaller. However, the full set of correlation matrices published in Ref. [237] for re-interpretation studies.

5.8 Interpretation

In this section we use the unfolded, experimental data to investigate the agreement with SM predictions from state-of-the-art MC event generators and to set limits on two BSM scenarios that predict DM candidates. We immediately reap the benefits of performing an unfolded measurement, as no detector simulation is necessary for either SM or BSM samples that are used in these statistical tests.

First, we set the stage and introduce the statistical framework in Section 5.8.1. Using this framework, a quantitative comparison with SM predictions is presented in Section 5.8.2, while we take a closer look at new physics models in Section 5.8.3.

5.8.1 Statistical framework and methodology

We can regard differential cross-section measurements as a combination of Poisson counting experiments for multiple bins. The probability distribution function (p.d.f.) of a random variable x with expectation value μ follows a *Poisson distribution*,

$$\mathcal{L}(x|\mu) = \frac{\mu^x e^- \mu}{x!} \xrightarrow{\mu\to\infty} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} , \tag{5.22}$$

approaching a Gaussian function in the large-sample limit. It is suggestively written as a conditional probability of observing some experimental value x based on a theoretical model expressed in μ. The combined likelihood of observing multiple (k) data points $\vec{x}$ as a set of theoretical parameters $\vec{\mu}$ and *nuisance parameters* $\vec{\theta}$ is then the product of the individual likelihood functions for each bin:

$$\mathcal{L}(\vec{x}|\vec{\mu},\vec{\theta}) = \prod_i \mathcal{L}(x_i|\mu_i) = \frac{1}{\sqrt{(2\pi)^k \det \mathrm{Cov}}} \cdot e^{-\frac{1}{2}\chi^2(\vec{x},\vec{\mu},\vec{\theta})} \cdot \prod_i \mathcal{G}\big(\theta^{(i)}\big), \tag{5.23}$$

which accounts for the correlations between bins via the covariance matrix (Cov) and where χ^2 is

$$\chi^2(\vec{x},\vec{\mu},\vec{\theta}) \equiv \big(\vec{x} - \vec{p}(\vec{\mu}) + \sum_i \vec{\theta}^{(i)} \cdot \vec{\epsilon}^{\,(i)}\big)^{\mathrm{T}} \mathrm{Cov}^{-1} \big(\vec{x} - \vec{p}(\vec{\mu}) + \sum_i \vec{\theta}^{(i)} \cdot \vec{\epsilon}^{\,(i)}\big) . \tag{5.24}$$

The theory prediction is indicated by $p(\vec{\mu})$, which can be pure SM or SM+BSM. Systematic uncertainties are included as nuisance parameters $\vec{\theta}$ and are expected to be normally distributed, i.e. have a mean of zero and standard deviation of one. The absolute uncertainty amplitudes are $\vec{\epsilon}^{\,(i)}$ and the standard Gaussian constraint term is

$$\mathcal{G}\big(\theta^{(i)}\big) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\theta_i^2} . \tag{5.25}$$

Now let us take a step back and remind ourselves that there are two main goals in this reinterpretation chapter. First, we want to quantify how well the data agrees with the Standard Model in a *goodness-of-fit* test, where in this case the SM corresponds to our null hypothesis $\mathcal{H}_0$. Second, we want to see if there is an alternative hypothesis $\mathcal{H}_1$, in this case SM with an extra contribution from BSM physics, that is more favourable than the null hypothesis. Although these are inherently different objectives, our general strategy is the same: for a given model – SM only or SM+BSM – we want to minimise the χ^2 to maximise the likelihood in a *fit*. In both tests the

systematic uncertainties provide degrees of freedom to vary data and theory within their errors and achieve better *post-fit* agreement, from which we can quantify the agreement between data and the hypotheses.

There are two approaches to include the effect of systematic uncertainties. First, in the "pull approach" the effect of Gaussian systematic uncertainties is embedded through shifts on the difference $\vec{x} - \vec{p}(\vec{\mu})$ conform the set of NPs $\vec{\theta}$. In a fitting procedure, the value the likelihood is maximised (or similarly χ is minimised) by reducing the difference in $\vec{x} - \vec{p}(\vec{\mu})$. This is a balancing act with the external constraint term $\mathcal{G}\left(\theta^{(i)}\right)$ that penalises NP shifts away from the nominal value of the systematic uncertainty. The constraint term also indicates the Gaussian nature of the uncertainties. The "pulls" are the resulting values of the data, theory and NPs that maximise the likelihood function: $\bar{x}, \bar{\mu}$ and $\bar{\theta}_i$ respectively. Profiling systematics in this method has multiple advantages. First, it allows us to implement asymmetric or one-sided uncertainties. Second, the post-fit values and NPs pulls contain information about modelling effects and demonstrate how each systematic is constrained by the data. Lastly, plots showing the agreement between data and theory can be made post-fit. The disadvantage is that it is computationally expensive: this measurement has more than three hundred experimental nuisance parameters and additional theoretical uncertainties, in addition to combining all the different phase space regions and observables in a single fit.

The second, alternative "covariance approach" includes systematic uncertainties in the covariance matrix as $\mathrm{Cov} = \mathrm{Cov}_{\mathrm{stat}} + \sigma_{\mathrm{source}} \mathrm{Cov}_{\mathrm{syst.,source}}$, which takes into account any (anti-)correlations in the measurement. The downside is that this approach symmetrises the systematic uncertainties and we lose information on pulls on individual NPs, but in return we greatly reduce the complexity of the likelihood function. The calculation of χ^2 is proven to be equivalent for both approaches in the case of Gaussian systematic uncertainties [242].

Ideally, i.e. with infinite computing power, all systematic uncertainties are profiled over in the pull approach since this provides the maximum amount of information. However, to perform goodness-of-fit tests and hypothesis tests we need to compute the likelihood function *many* times to model the p.d.f. of χ^2 under $\mathcal{H}_0$ or $\mathcal{H}_1$, as explained in more detail in the next paragraph. Therefore, it becomes infeasible to perform fits to thousands of toy experiments with hundreds of measurements. We make a compromise between retaining full information and speed. The pull approach is used when performing individual goodness-of-fits tests, because the main goal is to assess the modelling, impact and constraints of systematic uncertainties. Experimental and theoretical systematic uncertainties are added to the covariance matrix if their impact is less than 1% for every measurement in the likelihood function. The covariance approach is used when deriving observed distributions of toys. The effect of individual systematic uncertainties on the unfolded data is stored on HEPData [237], from which correlations between different nuisance parameters can be constructed in full.

In addition to the observed value of χ^2 per degrees of freedom, the agreement between data and the SM can be quantified via the p-value. This is the probability of obtaining a result as likely or less likely than the observed figure of merit $\chi^2_{\mathrm{obs.}}$ under the assumption of the SM only hypothesis $\mathcal{H}_0$:

$$p = 1 - \int_{\chi^2_{\mathrm{obs.}}}^{\infty} f\left(\chi^2_k\right) \mathrm{d}\chi^2_k \,, \tag{5.26}$$

where $f\left(\chi^2_k\right)$ indicates the p.d.f. of χ^2 with k measurements.

The Neyman-Pearson lemma states that the likelihood ratio is the most powerful hypothesis test at some significance threshold α. Therefore, to test an alternative hypothesis with BSM $\mathcal{H}_1$ to the SM only hypothesis $\mathcal{H}_0$, we define the likelihood ratio as

$$\lambda(\vec{\mu}) = \frac{\text{argmax}_{\vec{\theta}}\mathcal{L}(\vec{x}|\vec{\mu},\vec{\theta})}{\mathcal{L}(\vec{x}|\vec{\mu}_{\text{SM}},\vec{\theta})} \,, \tag{5.27}$$

where the likelihoods are evaluated by marginalising the NPs $\vec{\theta}$. The notation of $\vec{\mu}_{\text{SM}}$ makes it explicit that it contains only a SM contribution while $\vec{\mu}$ is a combination of SM and BSM signal. This is re-parametrised as the *test statistic*

$$t_{\vec{\mu}} = -2\ln\lambda(\vec{\mu}) = \chi^2_{\mu} - \chi^2_{\mu_{\text{SM}}} \,, \tag{5.28}$$

which will be positive if the SM is more compatible with the experimental data and negative when the alternative BSM hypothesis is more favourable. The test statistic $t_{\vec{\mu}}$ needs context in the form of an underlying p.d.f., based on the hypothesis that is assumed to be the "true" hypothesis, before any conclusion can be drawn from the numerical calculation of the observed $t_{\vec{\mu},\text{obs.}}$. The confidence level (CL) of the hypothesis $\mathcal{H}_1$ being true is defined as the probability $t_{\vec{\mu}}$ being larger or equal than $t_{\vec{\mu},\text{obs.}}$, assuming the alternative, signal hypothesis is true:

$$\text{CL}(\vec{\mu}) = \int_{t_{\vec{\mu},\text{obs.}}}^{\infty} f(t_{\vec{\mu}}|\vec{\mu})\text{d}t_{\vec{\mu}} \,. \tag{5.29}$$

In other words: under the assumption of this BSM scenario $\vec{\mu}$, what is the probability of getting a result *less likely* under these assumptions? In the frequentist approach, a hypothesis is excluded at a *confidence level* of $1-\alpha$ when $CL < \alpha$. The parameter α is the size of the test and is typically taken to be $\alpha = 0.05$ in particle physics experiments, such that the calculated limits are at 0.95 CL (or 95% CL). Lowering α raises the bar in terms of exclusion: to exclude a set of parameters $\vec{\mu}$, a smaller test size requires the value of CL to be smaller, necessitating either an increase in the observed agreement with the SM (i.e. decreased $\chi^2_{\mu_{\text{SM}}}$) or a more incompatibility with the signal model (i.e. increased χ^2_{μ}).

The p.d.f. of the test statistic $f(t_{\vec{\mu}}|\vec{\mu})$ is obtained by simulating pseudo-experiments (or "toy"-experiments) according to the expected statistical and systematic covariances. For the *observed limits*, the experimental data is fluctuated many times and in each case the test statistic is calculated, exposing the underlying distribution of $f(t_{\vec{\mu}}|\vec{\mu})$ for the signal hypothesis. Similarly, we can compute *expected limits* to gauge the sensitivity of a measurement to a particular signal. To calculate expected limits, no experimental data but only theoretical predictions are used. Under a hypothesis $\mathcal{H}_1$, we "expect" to measure a distribution that follows exactly the prediction of that hypothesis. By throwing toys according to this predicted distribution, we obtain the distribution of the expected test statistic. The expected sensitivity is then evaluated by calculating the CL exclusion where $f(t_{\vec{\mu}}|\vec{\mu})$ is evaluated for the null-hypothesis ($\mathcal{H}_0$), which answers the following question: "If we measured precisely the SM (background) only prediction with our experiment, how (un)likely are we to exclude this particular model?".

With Equation 5.29 we have a measure to exclude models of SM plus BSM signal that agree

poorly with experimental data. However, if the null hypothesis is more favourable than the alternative hypothesis it is not guaranteed that the null hypothesis itself has good agreement with the data. Consider an example where 3 events are measured in data, 9 events are expected from the SM and 2 BSM signal events are expected: the observation of 3 events is more favourable for the SM than the BSM, but still unrepresentative for the SM on its own. Additionally, in a pure frequentist CL there is a risk of falsely excluding a correct hypothesis in a fraction α of times. The problems are mitigated by defining a modified frequentist confidence level

$$\mathrm{CL_s} = \frac{\mathrm{CL}(\vec{\mu})}{\mathrm{CL_b}} \,, \tag{5.30}$$

which explicitly takes the confidence level of the background-only hypothesis $\mathrm{CL_b}$ into account. We exclude a hypothesis when $\mathrm{CL_s} < \alpha$. Since $\mathrm{CL_b} \leq 1$, as a result $\mathrm{CL_s} \geq \mathrm{CL}(\vec{\mu})$ for all $\vec{\mu}$. This means that the $\mathrm{CL_s}$ method over-covers and is more conservative than the CL method.

5.8.2 Standard Model fit

The unfolded data is compared to state-of-the-art SM theory predictions in a combined fit with two main goals. The first objective is to quantify if the data is modelled adequately within statistical and systematic uncertainties. The second aim is to gain insight in the modelling and impact of different systematics before we use the same experimental data to infer limits on parameters of different BSM models in a hypothesis test.

Fits are performed using the inclusive measurement using the same MEPS@NLO prediction shown in the data and MC comparison plots in Section 5.7.1. First, a fit it made to include all the $p_{\mathrm{T}}^{\mathrm{recoil}}$ distributions for the signal region and the auxiliary lepton measurements. This provides adequate modelling with a $\chi^2/\mathrm{n.d.o.f.} \approx 101/57$. The differential cross-section measurement for $\Delta\phi_{jj}$ show good modelling, but m_{jj} is known to be poorly modelled. Therefore, a simultaneous fit including all three observables ($p_{\mathrm{T}}^{\mathrm{recoil}}$, $\Delta\phi_{jj}$ and m_{jj}), two jet topologies (≥ 1 jet and VBF) and measurement regions (signal and leptons) yields a poor agreement of $\chi^2/\mathrm{n.d.o.f.} \approx 390/70$.

Fits are also performed using R^{miss} ratios instead. As seen in Section 5.7.2, mismodelling effects and systematic uncertainties cancel to a large degree in these ratios. The fit configuration including the R^{miss} is expected to show improved agreement between data and MC. Indeed, we observe $\chi^2/\mathrm{n.d.o.f.} \approx 62/56$ in a simultaneous fit to all regions and all observables.

For above reasons, R^{miss} ratios are used in the hypothesis tests presented in the next section. Specifically, the hypothesis tests use R^{miss} values of $p_{\mathrm{T}}^{\mathrm{recoil}}$ in the ≥ 1 jet region only (with $\chi^2/\mathrm{n.d.o.f.} \approx 48/45$ in the SM fit), and $p_{\mathrm{T}}^{\mathrm{recoil}}$ from both the ≥ 1 jet and VBF region (with $\chi^2/\mathrm{n.d.o.f.} \approx 110/84$ in the SM fit).

QCD scale uncertainties are the dominant SM theory uncertainty, giving a band of size 15–20% for the MEPS@NLO predictions. The shape of the uncertainty within this band is generally not well known. Following the recommendations from Ref. [243], different decorrelation schemes are tested and, if there are any differences in the final results, the approach leading to the most conservative results is taken. In Section 5.4.4, the default correlation scheme from Ref. [225] was introduced. The shape of a hyperbolic tangent allows tails of QCD scale uncertainty to vary separately from the centre of the distribution, where the mid-point at 50% in logarithmic space

is chosen arbitrarily. We tested alternative schemes where this point was varied to 0%, 25% and 75% of the distribution in logarithmic space. Additionally, for each mid-point an alternative decorrelation function abs$(\tanh(x))$ was tested, which varies the high and low tails in the same direction rather than the opposite direction. Simultaneous fits to all regions in the ≥ 1 jet phase-space show that the resulting values of $\chi^2/\text{n.d.o.f.}$ are unaffected by using different correlation schemes. Therefore, the default mid-point of 50% as suggested in Ref. [225] is used. All other sources of systematic uncertainties, including the experimental systematics, are taken to be fully correlated between different bins, observables and regions.

Figure 5.51a and 5.51b show the nuisance parameter post-fit pulls for the $p_{\text{T}}^{\text{recoil}}$ observable in the ≥ 1 jet region for a SM-only fit to individual cross-section distributions and the R^{miss} ratios respectively. When using individual distributions, there are significant pulls on the QCD scale uncertainties ($\delta^{(1,2)}K_{\text{NLO}}$) to compensate for the overall cross-section offset that is observed between data and MC. These large pulls suggest that it is unsafe to use this fit configuration directly to constrain BSM, as the tightly constrained post-fit parameters can lead to overestimation or underestimation on constraints of the BSM parameter space.

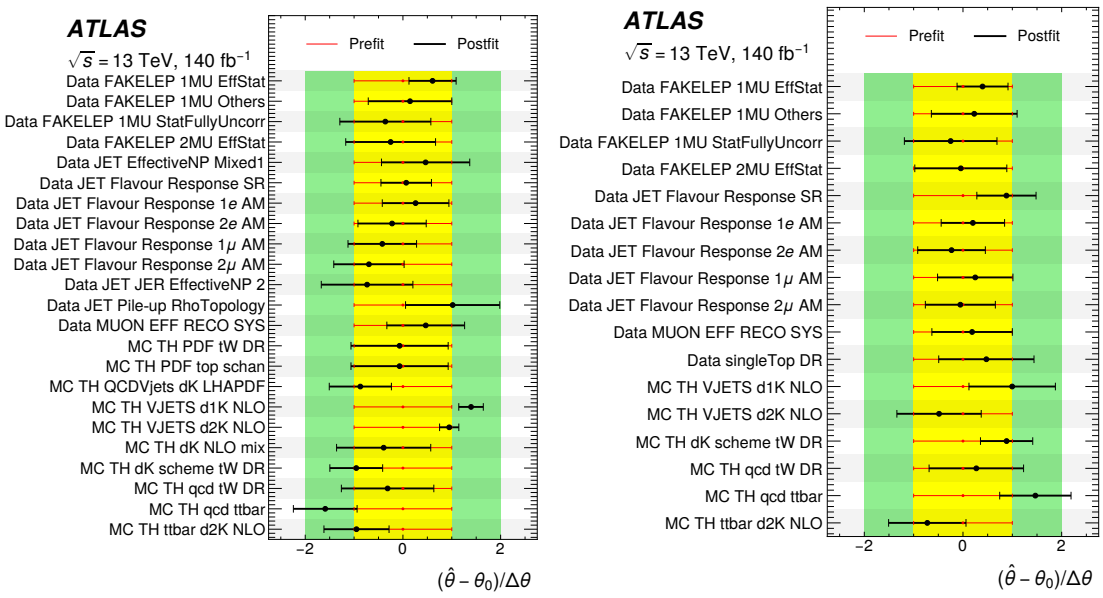
As seen in Figure 5.51b, the constraints on QCD scale uncertainties are relaxed when using R^{miss} instead. This is understood because the offset – present in the signal region as well as the lepton regions – cancels to a large degree in the R^{miss} ratio, as seen in the fiducial cross-section plots in Section 5.7. Some parameters pulled slightly even for R^{miss}, for example the jet flavour response, top modelling and $\delta^{(1,2)}K_{\text{NLO}}$, but all parameters still agree with the pre-fit value within uncertainties and only one parameter (QCD uncertainty related to $t\bar{t}$ modelling) is pulled by more than one standard deviation from the pre-fit value.

The good $\chi^2/\text{n.d.o.f.}$ fit values together with the pull plot motivates the choice for using the R^{miss} fit configuration in deriving constraints on BSM models in the next section. The caveat is, however, that R^{miss} is not a suitable quantity for all BSM interpretations. Depending on the model, BSM signal can appear in both the signal region and the lepton regions and a cancellation in the ratio can lead to reduced sensitivity in such scenarios.

5.8.3 Limits on new physics scenarios

One important motivation for the measurement of missing energy with jets final state is its sensitivity to a broad range of BSM physics, particularly of $p_{\text{T}}^{\text{miss}}$ in the signal region for models that predict DM candidates. In the context of the ATLAS analysis we consider two BSM scenarios, but with the publication of unfolded data on HEPData together with detailed information on systematic uncertainties and statistical correlations this is a powerful measurement in re-interpretation efforts.

First, we set limits on the parameter space of a common benchmark simplified model in which a DM candidate couple directly to an axial-vector mediator Z'. This is one of the two models that was used in the signal injection tests. We show that limits obtained using our particle-level measurement are competitive with a dedicated detector-level monojet search that was highly optimised to BSM sensitivity. Second, we consider a more complicated model that introduces an additional Higgs doublet, along with a pseudo-scalar a which couples to the DM candidate.



(a) (b)

Figure 5.51: Nuisance parameter post-fit pulls, expressed as the difference between the best-fit value $\hat{\theta}$ and the pre-fit value θ_0 normalised by the pre-fit uncertainty ($\Delta\theta$) for each NP. Nuisance parameters with impacts of $< 1\%$ across the measured spectra are not shown. The post-fit values correspond to the SM only fits to the MEPS@NLO predictions, using the signal region and auxiliary measurement regions for (a) the $p_{\mathrm{T}}^{\mathrm{recoil}}$ observable in the inclusive jet phase space, using the cross-section distributions and (b) the $p_{\mathrm{T}}^{\mathrm{recoil}}$ observable in the inclusive jet phase space, using the R_{miss} distributions. In the raw fit, some NPs representing QCD scale uncertainties and their shapes are constrained by construction: these uncertainties are taken to be fully correlated between boson channels and regions of phase space, meaning that the data in the $1e$+jets, 1μ+jets, $2e$+jets and 2μ+jets regions can constrain the theory uncertainties in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region. This follows the approach of Ref. [225], which proposes an aggressive correlation model based on state-of-the-art QCD calculations at NNLO and electroweak at NLO. Figures from Ref. [232].

Fermionic DM with axial-vector mediator

A common, simplified extension of the SM involves an additional U(1) symmetry that predicts the existence of a fermionic Dark Matter candidate χ that couples to an axial-vector mediator Z' and is only charged under this additional gauge group [34], also referred to as the "DMA" model. If SM quarks are charged under this gauge group, Z' mediators can be produced in LHC proton-proton collisions for example via s-channel diagrams. The Z' can decay back into two SM quarks but the distinctive signature is a decay to two χ fermions leading to a $p_{\mathrm{T}}^{\mathrm{miss}}$ contribution larger than the SM prediction. This model has been studied extensively by both ATLAS [187, 191] and CMS [244]. To illustrate the power of unfolded, precision SM measurements we compare our results to those from the full Run-2 dataset ATLAS monojet search [191], that uses the same dataset obtained under the same detector conditions.

The parameter settings to generate DMA events at the particle level include a coupling of Z' to SM quarks $g_q = 0.25$ and the coupling of Z' to the DM candidate $g_\chi = 1.0$. These parameters are often chosen in benchmark models because in this regime the narrow-width approximation remains valid. Also, setting $g_\chi = 1.0$ ensures that Z' decays predominantly to DM particles, which leads to a signal that is more distinctive above multijet background than when it would decay back to a pair of SM quarks. In this analysis, we evaluate 95% $\mathrm{CL_s}$ in the plane of m_χ and $m_{Z'}$. As detailed in the previous section, the fitting procedure for this hypothesis test uses R^{miss} as inputs, where the predicted R^{miss} is recalculated to include the BSM signal.

As this signal leads to a real $p_{\mathrm{T}}^{\mathrm{miss}}$ signature, the DMA model only contributes to the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region and not to the auxiliary lepton regions. Figure 5.52 shows what this model looks like in the $p_{\mathrm{T}}^{\mathrm{miss}}$ observable in the ≥ 1 jet region, for three different masses of M_χ and a fixed mediator mass of $m_{Z'} = 1750\,\mathrm{GeV}$. For each value of M_χ, the relative signal strength gets increasingly stronger at larger $p_{\mathrm{T}}^{\mathrm{miss}}$. The phase-space factor for $Z' \to \chi\chi$ is larger for smaller values of M_χ, leading to increased fiducial cross-sections as M_χ gets smaller.

The obtained limits are shown in Figure 5.53. As the excess data – observed in the individual cross-section measurements – cancels in the R^{miss} ratio, we see that the observed and expected limits agree relatively well when including a BSM signal. Similar to the ATLAS monojet search with $139\,\mathrm{fb}^{-1}$ of data [191], masses for the mediator are excluded up to approximately $2\,\mathrm{TeV}$ in the region $m_\chi < m_{Z'}/2$. The area above $M_\chi > m_{Z'}/2$ is not excluded because in this region the on-shell decay of $Z' \to \chi\chi$ is not accessible due to energy conservation. At mediator masses above $2\,\mathrm{TeV}$ the production cross-section becomes small due to the finite beam energy of $\sqrt{s} = 13\,\mathrm{TeV}$, ultimately leading to reduced sensitivity in the differential distributions of the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets measurement region. There is a discrepancy in the expected limits, that can be attributed to differences in the event selections and the choice of binning. In the unfolded measurement we enforce relatively broad bins in the $p_{\mathrm{T}}^{\mathrm{miss}}$ tails to maintain a sufficient number of events for robust unfolding, while this is not a strict requirement for the search. The key point is that measurement can be used for setting constraints with a sensitivity reaching as far as for dedicated searches, but measurements have the advantage of disabling the need for detector simulation when re-interpreting the data.

In the same spirit as the SM fits, different correlation models of the QCD scale uncertainties are tested to evaluate their impact on the BSM models. This test shows that the obtained limits are insensitive to the correlation scheme. Therefore, the default 50% midpoint in logarithmic space

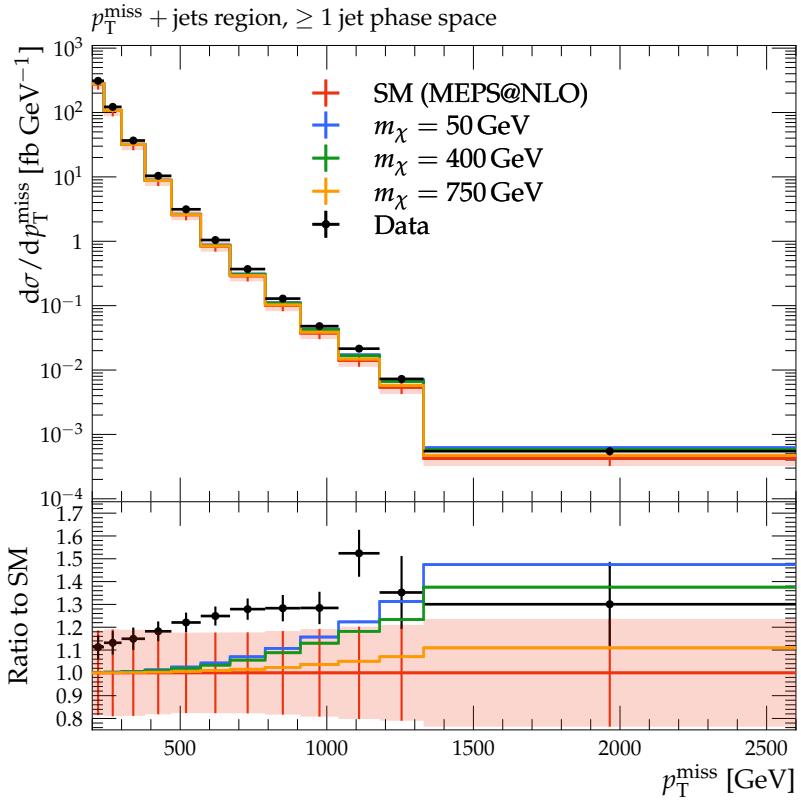


Figure 5.52: Differential distribution of the $p_{\mathrm{T}}^{\mathrm{miss}}$ observable in the ≥ 1 jet phase-space, with the pre-fit SM MEPS@NLO prediction indicated in red and pre-fit unfolded data as black markers. Combined systematic and statistical uncertainties are indicated as vertical bars for both data and MC. Additionally, this figure shows the contribution from three different parameter choices of the DMA model that is explained in more detail in the text. For each BSM signal, the mediator mass is fixed to $M_{Z'} = 1750\,\mathrm{GeV}$. The blue, green, and yellow histograms correspond to the signal for $M_\chi = 50, 400, 750\,\mathrm{GeV}$ respectively, which are added to the SM prediction. The DM masses are chosen to illustrate the shape of BSM contributions for a range of different parameter points in the regime $M_\chi < m_{Z'}/2$. BSM contributions increase with $p_{\mathrm{T}}^{\mathrm{miss}}$ relative to the SM prediction. The $Z' \to \chi\chi$ fiducial cross-section is larger as M_χ becomes smaller.

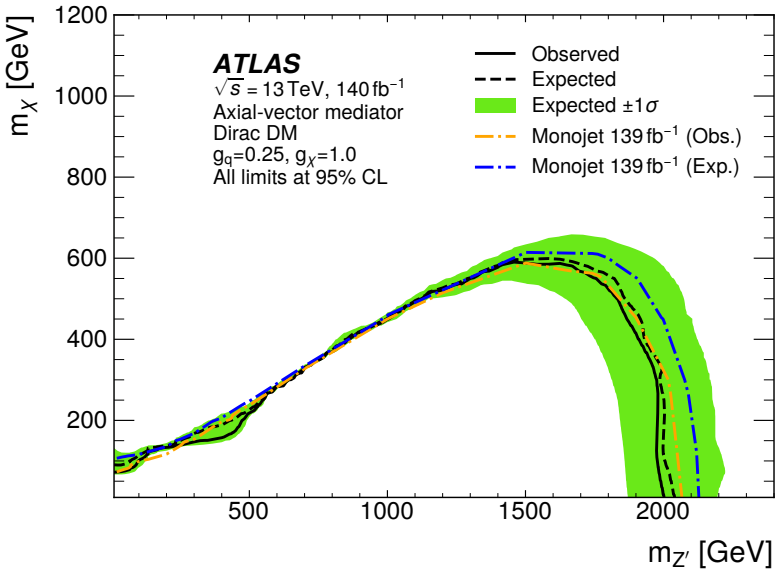


Figure 5.53: Exclusion limits at 95% $\mathrm{CL_s}$ in the plane of Dark Matter mass (m_χ) and mediator mass ($m_{Z'}$) for the simplified DM model with an axial-vector mediator. Expected limits at $\pm 1\sigma$ are indicated as a green shaded region, where the observed and expected limits from this analysis, evaluated using the particle-level R^{miss} measurements in the ≥ 1 jet region, are indicated as a solid and dashed black line respectively. The area under the black solid line is excluded, showing a sensitivity of up to 600 GeV for M_χ and up to 2 TeV for the mediator mass $M_{Z'}$. The results are competitive with the limits from the ATLAS monojet search [191]. Figure from Ref. [232].

for the (de)correlation function is used. All other sources of systematic uncertainties, including the experimental systematics, are taken to be fully correlated between different bins, observables and regions. Additionally, limits from the R^{miss} ratios are consistent with limits obtained by using the individual cross-section measurements. This is expected since both methods use information from the lepton regions to constrain the signal region in an (eventually) equivalent way.

2HDM+*a*

This study of a second DM scenario demonstrates the flexibility of an unfolded measurement and the sensitivity of an inclusive $p_{\mathrm{T}}^{\mathrm{miss}}$ final state. This model predicts an additional Higgs doublet and a pseudo-scalar a that couples to DM [33, 245]. In this analysis, we perform a scan in the plane of the pseudo-scalar mass m_a and $\tan\beta$, the ratio between the vacuum expectation values of the two Higgs doublets. This model is referred to as the 2HDM+a model, which is studied extensively in ATLAS at different model configurations [246]. The sensitivity depends on the choice of multiple free parameters, but the shared feature is 2HDM+a appears predominantly in analyses that have some form of missing energy in the final state, such as $E_{\mathrm{T}}^{\mathrm{miss}}$+Higgs, $E_{\mathrm{T}}^{\mathrm{miss}}+Z$ or indeed $E_{\mathrm{T}}^{\mathrm{miss}}+$ jets.

The 95% $\mathrm{CL_s}$ limits, obtained from the R^{miss} values of both the ≥ 1 jet and VBF regions, are shown in Figure 5.54. There are three main regions of sensitivity. First, for low values of $\tan\beta < 0.7$ masses of $m_a < 520$ GeV are excluded due to loop-induced production of the pseudo-scalar and its consequent decay into DM particle, as $pp \to a(\chi\bar{\chi})$+jets. Second, at large $\tan\beta \geq 10$, b-quark induced production of the pseudo-scalar yields another excluded region. Lastly, there is

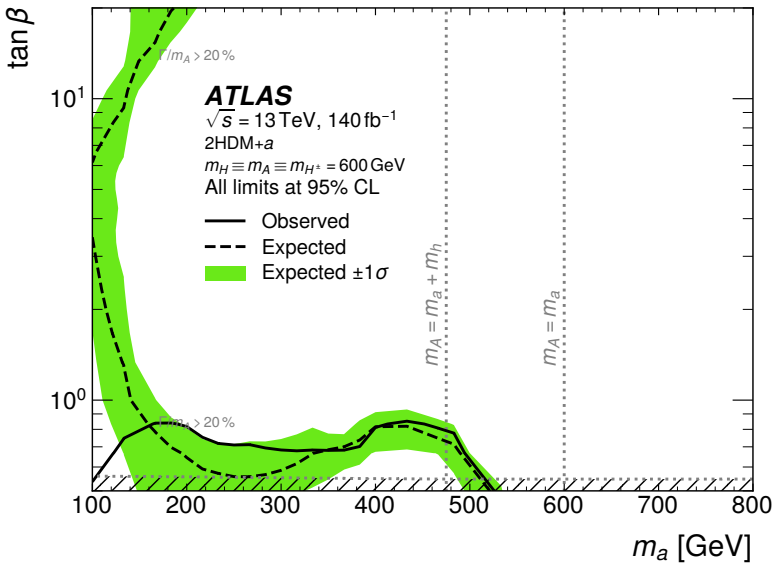


Figure 5.54: Exclusion limits at 95% $\mathrm{CL_s}$ in the plane of $\tan\beta$ and the pseudo-scalar mass m_a for the 2HDM+*a* model. Expected limits at $\pm1\sigma$ are indicated as a green shaded region, where the observed and expected limits from this analysis, evaluated using the particle-level R^{miss} measurements in the ≥ 1 jet and VBF regions, are indicated as a solid and dashed black line respectively. The area below the solid black line is excluded, i.e. of values up to $\tan\beta \approx 0.8$ at $m_a < 500\,\mathrm{GeV}$. The area to the left and below the dashed line is expected to be excluded. At low values for m_a there is a discrepancy between these observed and expected limits, which is understood qualitatively from the bad pre-fit agreement between data and the SM due to the normalization offset of the MC predictions in combination with the shape of the 2HDM+*a* signal in the differential distributions used as inputs to the fit. Figure from Ref. [232].

sensitivity at small m_a for processes that are independent of $\tan\beta$, for example $pp \to H \to aZ$ and $pp \to H^{\pm} \to aW^{\pm}$. Additionally, leptonic decays of W and Z bosons in the aforementioned processes can lead to small contributions to the auxiliary lepton regions.

At low values for m_a there is a discrepancy between observed and expected limits, which is understood qualitatively as follows. There is bad pre-fit agreement between data and SM mainly due to the total cross-section mismodelling of the MC, exemplified by the relatively large pull on the overall normalization of the MC in Figure 5.51. This is reduced for R^{miss}, which is used in the limit-setting, but some residual effects remain. As a consequence, the addition of BSM signal can lead to a better overall fit to the data. This is particularly a problem in the 2HDM+*a* and not so much for the exclusion for the DMA model for the two following reasons. First, the 2HDM+*a* signal in the differential distributions is less peaked and more broadly distributed between bins than the simplified DMA model shown in Figure 5.52. As a result, discrepancies between data and MC can be solved in many bins at once due to the addition of 2HDM+*a* signals. Secondly, the 2HDM+*a* model does not only contribute to the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region, but also the lepton regions. The discrepancies between data and MC in different measurement regions are then mitigated by the BSM simultaneously, which reduces the observed exclusions.

Chapter 6

Analysis preservation and reinterpretation

In the previous chapter we have seen how results from a physics analysis are interpreted in terms of the SM and two BSM scenarios. However, only a limited number of theoretical models can be tested directly by the experiments. This can have practical reasons, for example not having enough person power in the analysis team to carefully examine that one extra BSM model, or the decision to publish a measurement as soon as possible to meet a conference deadline. But there can also be more fundamentally restricting reasons. For example, SM to a fixed order in perturbation theory may not be available at time of the publication or ideas for a new physics scenario were only proposed after the measurement was made.

To pre-empt such situations and maximise scientific value of publications, experiments can define their analyses and store the data in such a way that the experimental data can later be reused. How experiments can do this and how results can be reused is discussed in Section 6.1. Now is a critical time for shifting our approach to analysis preservation, because Run-3 (and Run-2) may be the final word on many topics until they are surpassed well into the HL-LHC era. If we do not to a good job of this now, some data might be lost forever. Careful analysis preservation leaves a potential goldmine for the next generation of physicists, just like how results from deep inelastic scattering experiments are instrumental in the current day through the determination of parton distribution functions. In Section 6.2 we highlight current tools and e-infrastructure that shape the analysis preservation and reinterpretation landscape. Lastly, these tools are taken to the test by reusing existing LHC data to examine the compatibility of a Dark Matter model in Section 6.3.

6.1 General overview

Discussions on analysis preservation and reinterpretation have played a big role in the past decade. Motivations and recommendations for the LHC started to emerge during Run-1, see e.g. Refs. [247–249], with more recent recommendations for the HEP community provided for example in Refs. [174, 250, 251]. The LHC Re-interpretation Forum and Les Houches meetings are instrumental platforms for continued discussion of this field and the development of the

necessary tools to facilitate this, with yearly workshops where experimentalists and phenomenologists alike meet. The necessity of data and analysis preservation is also underpinned by the 2013 update to the European Strategy for Particle Physics [252], as well as a dedicated Data Preservation study group [253].

This section highlights main techniques for analysis and (meta)data preservation in Section 6.1.1, with applications to a variety of physics problems summarised in Section 6.1.2.

6.1.1 Preservation of LHC analyses

The reusability is determined by both the analysis design and the auxiliary material that is published along with the journal publication.

Analysis design

As discussed in Section 4.3.1, particle-level measurements are easier to interpret than particle-level searches. Therefore it is generally to the benefit of the HEP community to publish results in the particle-level representation. A key point here is the validation of the unfolding under different circumstances. Unfolding tests, as illustrated in Section 4.4.3 are instrumental to ensure that model-assumptions during the unfolding still hold when reinterpreting the data under a *new* model.

Unfolding is a statistical procedure that is unstable if the number of binned events becomes too small. Therefore, particle-level measurements are not always possible in low-statistics regions of phase space that are optimised for sensitivity to a rare signal process. This is where detector-level searches still punch their weight in terms of BSM reinterpretation, and the community still benefits of the publication of online materials for searches. But since full detector simulation is CPU expensive, concessions are made to the scope of reinterpretation studies (by only considering a subset of parameter configurations) or accuracy (by reverting to fast simulation or smearing tools).

A second choice in analysis design, alluded to in Section 4.3.2, is the definition in terms of final-state particles instead of physical processes. This removes the model-dependence of measured data on irreducible SM background processes. In this definition, future improvements to SM predictions from MC event generators are simply reinterpreted and validated against the experimental data. Additionally, model-dependence is reduced by defining particle-level objects close to the detector definitions.

The physics programme of experiments should cover as much phase space as possible, with a broad range of final states and observables. Because different analyses use overlapping datasets and are generally not statistically orthogonal, it is important to account for the possibility of events that populate the same bins in different searches or measurements. This can be preempted by the use of bootstrap methods [235] to correlate bins between different analyses, while novel ideas are developed to quantify the amount of overlap between different regions [254].

Online publication of auxiliary (meta)data

CERN recognises the necessity to make LHC data publicly accessible in different levels of abstraction for a variety of use cases, enshrined in the CERN Open Data Policy [255]. The key point

is that experiments provide, in addition to the journal publication, digitised versions of the data and auxiliary metadata to maximise the reinterpretation potential of analyses. The standardised public, online repository for the storage of this information is HEPData [236, 256]. There is a varying degree of how much data is published by different analyses (and groups) within the same experiment, let alone between different experiments. ATLAS provides HEPData recommendations for "bare minimum" to "maximal" reinterpretation potential [257]. This section describes the different types of data that can be published to online repositories. This gradually gets more complex, but in turn offers more flexibility in terms of reinterpretation.

The absolute minimum that analyses can provide are the numerical values of the experimental data, including a simple breakdown of the systematic and statistical uncertainties. This should cover all the measured distributions from a paper that are suited for reinterpretation, and it is recommended to also provide the corresponding SM and BSM predictions used in the analysis.

The analysis logic should be clearly explained in the paper, but this can still lead to ambiguities to precisely reproduce the event selection. To validate the selection, experiments should provide a detailed cut-flow table for each signal region or a runnable code snippet that encapsulates the analysis logic. Section 6.2 highlights RIVET as the industry standard for the preservation of particle-level analyses.

For analyses dominated by systematic uncertainties, a simple breakdown is usually not sufficient. In this case the dominant sources of uncertainty should be provided individually, allowing correlations between different bins to be extracted. This can alternatively be done by providing explicit correlation matrices or even the full statistical workspace containing the likelihood functions. For recombination of results in different histograms, explicit statistical correlations need to be provided, for example obtained by bootstrap methods.

Analyses defined in terms of final-state particles can have contributions from different SM processes. In this case, it is beneficial to include a breakdown of the predictions for different SM processes, ideally with a decomposition of the associated theoretical uncertainties.

In the case of likelihood fits, the publication of the statistical workspace contains valuable information for reproducibility and future (re-)combinations. Since a few years ATLAS started to publish workspaces using a serialised JSON format [258], that represent `pyhf` [259, 260] workspaces[1]. One example of an analysis that published the full likelihood is the ATLAS $\sqrt{s} = 13\,\text{TeV}$ supersymmetry search with τ leptons in the final state [262].

Over the last few years, non-analytic variables from machine learning algorithms have taken a central role in many BSM searches. This outpaced the online preservation infrastructure, so the storage of trained models in a stable way has become an important point of attention in workshops at Les Houches and the Reinterpretation forum [263].

6.1.2 Reusing results from LHC analyses

Experimental data, if preserved in an accessible way, can be used for many different purposes. These are broadly categorised into reinterpretation for SM applications and for BSM searches. This section focuses on BSM reinterpretations of particle-level measurements, which is later on

[1] The `pyhf` package is a Python implementation of HistFactory [261] that works independently of ROOT.

used to constrain a BSM model that predicts a DM candidate from anomaly cancellation in Section 6.3.

In the context of the SM, LHC measurements are important inputs for Monte Carlo event generators. Section 4.1.1 already discussed fits for parton distribution functions, but there are other parts in the MC event generation chain that introduce free parameters that are either fundamental to the model or related to the implementation in MC models. These parameters are "tuned" using experimental data. This tuning includes parameters for MPIs, initial- and final-state radiation, and hadronisation [264, 265]. Of course, experimental data is also used for direct SM-to-data comparisons. State-of-the-art SM predictions can be validated and their (shape) modelling assessed. This can also be in the form of global (SM) effective field theory fits, see e.g. Refs. [266, 267], to interpret experimental measurements with minimal assumptions of the underlying physics at high mass scales.

Reinterpretation for BSM allows models to be validated to regions of phase space that were not included in the original measurement. As a result, in simultaneous fits to multiple measurements, a larger region of BSM parameter space can be excluded. There are various tools that automate this procedure. The general idea is to generate BSM events, compare the results with experimental datasets and ultimately infer some statistical limits from this. The main difference between tools is the choice in terms of the representation level of the data, using either particle-level measurements, detector-level searches, or both. Reinterpretations can conveniently be made at the particle-level, because simulation of generated BSM samples is not needed, but the scope of publicly accessible, re-interpretable particle-level measurements is smaller compared to the plethora of detector-level BSM searches.

6.2 Tools

There is a profusion of open-source software packages that enable analysis preservation and reinterpretation of LHC datasets. The main tools that are currently being used and developed are summarised in Table 6.1. While the subtle distinctions between various tools are interesting to discuss, this section restricts the topic to RIVET, YODA, HEPData and CONTUR, central to the phenomenological studies in Section 6.3.

6.2.1 RIVET, YODA and HEPData

Together, RIVET, YODA and HEPData form the holy trinity of particle-level analysis preservation.

RIVET (Robust Independent Validation of Experiment and Theory) consists of a computational core that can reproduce analysis logic from LHC measurements, mapping particle-level or detector-level events in HepMC3 format [94] to binned histograms. This mechanism includes the full reconstruction of the fiducial phase-space including jet clustering (using FastJet [155]) and dressed leptons, calculation of event properties and application of event selections, and ultimately a projected onto experimentally measured observables. The analysis logic for each publication is stored as a C++ snippet, or a "RIVET routine", that should exactly reproduce the results from the code that was used to produce the results for the final paper. Indeed, careful validation of the RIVET routine is required before it is accepted into the RIVET library, which currently consists of

Objective	Tool	Representation level	Description	Refs.
Data storage	HEPData	Particle & detector	Online repository for experimental datasets	[236]
Analysis preservation	RIVET	Particle & detector	Preservation of analysis code in runnable C++ snippets	[268, 269]
	SimpleAnalysis	Detector	Implementation of ATLAS SUSY searches	[270]
BSM limits	CONTUR	Particle	Multidimensional BSM parameter scanning using LHC measurements	[241, 271]
	MadAnalysis	Particle & detector	Recast of LHC searches and measurements	[272–274]
	Gambit	Particle & detector	Global fits combining LHC data with results outside collider physics	[275]
	Checkmate	Detector	Implementation of LHC searches	[276]
	SModelS	Detector	Simplified-model approach to reinterpreting LHC search results	[277, 278]

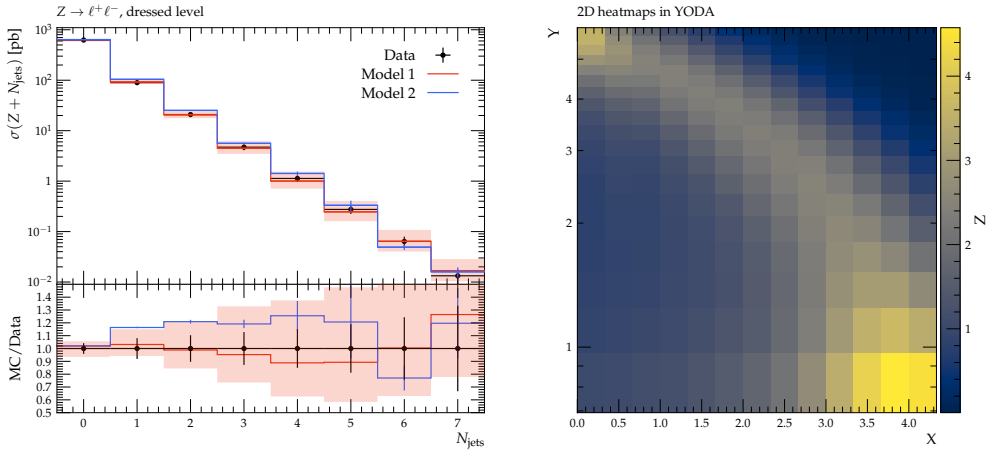
Table 6.1: Overview of several publicly accessible tools to enable reinterpretation of LHC physics results.

more than 1500 preserved analyses. Rivet also includes smearing functions to emulate detector effects, so that detector-level can be included in addition to the particle-level measurements.

RIVET uses the YODA (Yet more Objects for Data Analysis) statistics library [279], which emerged from the development of Monte Carlo event generators as a functional tool developed for flexible, lightweight and multidimensional histogramming. Because datasets in HEP are often very large, it is computationally infeasible to perform histogramming in one go. Therefore, continuous aggregation is a desirable property to enable histograms being filled in multiple steps. This rules out many existing histogramming tools, and HEP software relies in many cases on the ROOT software package [280] for these continuous aggregation features. However, ROOT has many additional dependencies and features beyond histogramming. Therefore it is desirable to have a more lightweight package for a flexible, maintainable implementation of histogram features, provided by YODA.

Although RIVET is also used for purely simulation based comparisons, e.g. in Figure A.1 to study two different MC setups, the power of RIVET lies in its ability to project simulations onto experimental data. In fact, RIVET is the default tool used by developers of MC generators to validate results and tune parameters to existing measurements. The experimental data is obtained from HEPData in the YODA format, and stored to the RIVET repository. This closes the circle of direct data versus MC comparisons at the particle-level.

Both YODA and RIVET offer full functionality of visualising the different YODA data-types. As part of this thesis work, the plotting infrastructure in RIVET and YODA were completely restructured



(a) (b)

Figure 6.1: Example plots generated with the RIVET and YODA plotting functionality. a) 1D-histogram comparing multiple MC predictions to data with uncertainty bands and b) 2D-histogram with values indicated by a colour map. Figures taken from Ref. [279].

to use modern, widely-used computing packages. This Python-based API is interfaced to the computational library NumPy [281] and data-visualisation toolkit Matplotlib [282]. Users can interact with the API directly, but YODA and RIVET both also provide plotting scripts that can be used to generate plots out-of-the-box in a common style.

Most of the plotting functionality is part of the YODA software. This module supports customisation of plot features, such as ratio plots with multiple inputs, switching between logarithmic and linear axes, tweaking of labels and more. An example of 1D and 2D histograms are shown in Figure 6.1. The plotting functionality is designed to be agnostic to any underlying physics, because YODA is a histogramming package not restricted to applications in HEP. Plots can be generated in different graphical formats, but a distinctive feature of this new module is that it also produces a set of executable Python scripts, one corresponding to each graphical output. The Python scripts reproduce the graphical outputs via plotting-calls to Matplotlib and NumPy. The scripts are designed to be easy to understand by the user, and can therefore be refined and tweaked further to offer maximum flexibility. In addition, this text-based intermediate format is future-proof and easily exchangeable between different software systems.

Practical HEP applications of RIVET demand the propagation of MC errors, including uncertainties on the choice of renormalisation and factorisation scales or on the PDF set. To this end, the RIVET plotting module first calculates error bands, for example as the uncertainty envelope or from a PDF prescription. In a second step these combined bands are passed to the YODA plot API for the generation of graphical outputs and associated plotting scripts.

6.2.2 CONTUR

CONTUR (Constraints On New Theories Using RIVET) [241] is a framework that enables comparisons between BSM theories and precision LHC measurements that are available in the RIVET analysis library. The philosophy behind CONTUR is that many models, or subsets of the parameter space thereof, can already ruled out by existing measurements without defining a dedicated

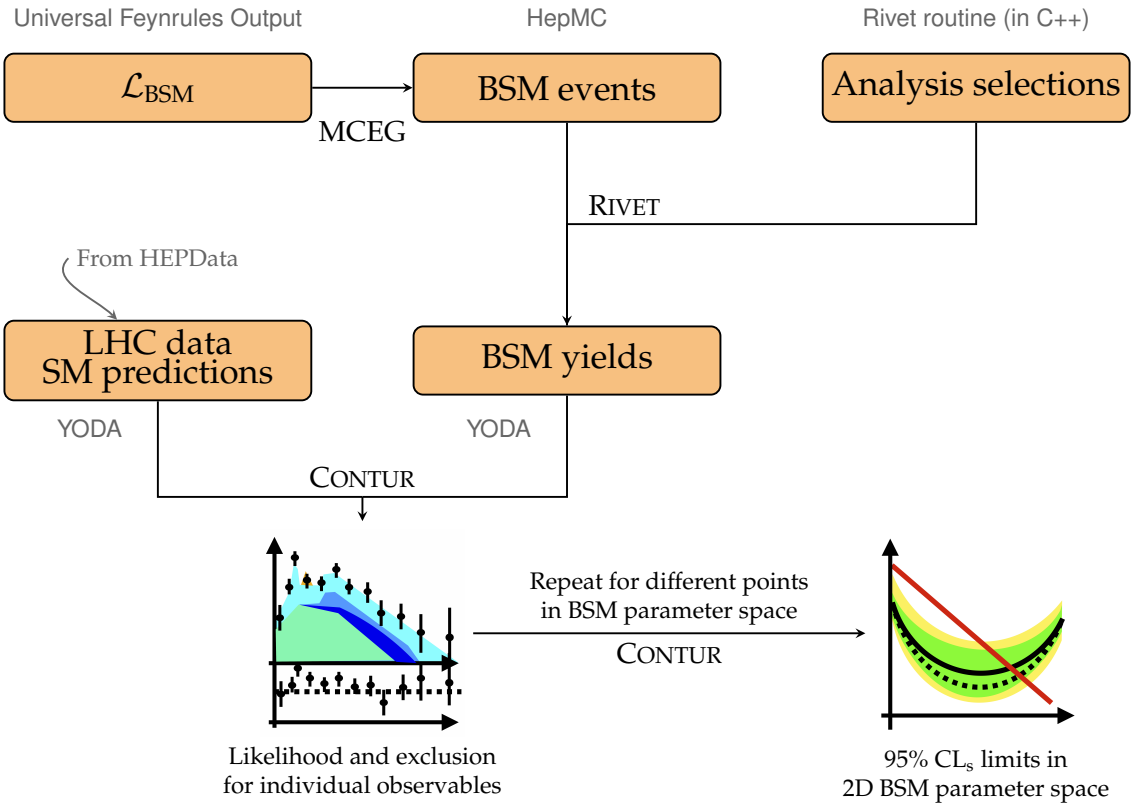


Figure 6.2: Diagrammatic representation of the CONTUR workflow. Starting from a set of BSM Feynman rules, a Monte Carlo event generator (MCEG) can generate BSM collision events at the particle level. Using RIVET this is projected onto observables, that are compared to experimental data and SM predictions from HEPData. CONTUR infers statistical limits from this, which can be repeated for multiple points in BSM parameter space to obtain multidimensional exclusion limits. Additionally, the CONTUR infrastructure enables automatic steering of the event generation and running RIVET over the set of events. Therefore, the arrows with "MCEG" and "RIVET " are effectively hidden from the user. The illustrations of the histogram and exclusion plot are kindly provided by L. Corpe [209].

experimental search for it. This exploits an exhaustive list of measurements and final states, offering different run conditions from different beam energies and experiments. CONTUR links together multiple established tools (e.g. RIVET and different MC event generators) for a streamlined and efficient workflow, while the modular design makes it straightforward to include the latest versions of the used software packages.

The CONTUR workflow is as follows. A set of BSM Feynman rules, indicating the new BSM particles and their interactions, is fed into MC event generators, such as HERWIG or MADGRAPH, to simulate proton-proton collisions under the assumption of the new BSM processes. These events can then be projected onto observables, corresponding to LHC measurements, using RIVET. The BSM contribution can be stacked on top of the SM predictions, and the agreement with experimental data evaluated. If there is a large BSM signal that exceeds the uncertainty bands, this is converted to a confidence level intervals on the parameters of the BSM model. By repeating this procedure for multiple points in parameter space, multidimensional exclusion diagrams can be obtained. This is illustrated schematically in Figure 6.2.

CONTUR uses a subset of particle-level measurements in the RIVET library. These measurements can partially overlap if the same dataset and a similar final state is used. As a result, the injection of BSM can leads to correlated signals in distributions from different analyses. Without a proper treatment of these correlations with bootstraps, double counting can lead to an overestimation of the resulting exclusion. To mitigate this, CONTUR divides the measurements in orthogonal categories or "pools", that are separated by beam energy, experiment and final state. Only the best exclusion is taken and combined with exclusions from the orthogonal pools to form a combined exclusion. This per-pool-best-exclusion can be calculated either from a histogram within that pool, if statistical correlations between different bins are available on HEPData, or a single bin, if correlations are not available.

CONTUR limits are presented as 68% or 95% $\mathrm{CL_s}$ intervals, where $\mathrm{CL_s} = p_{s+b}/1 - p_b$. The p-values for the background only (null) hypothesis b and the alternative hypothesis $s + b$ are calculated from the test statistic based on the likelihood ratio

$$t_\mu = -2 \ln \frac{\mathcal{L}(\vec{x}|\mu\vec{s} + \vec{b})}{\mathcal{L}(\vec{x}|\hat{\mu}\vec{s} + \vec{b})}. \tag{6.1}$$

In this equation, s indicates the number BSM signal events, b the SM background and x the number of events in experimental data. The parameter of the signal strength that maximises the maximises the likelihood is written as $\hat{\mu}$. The signal hypothesis is excluded when the signal strength $\vec{\mu} = 1$ is excluded. The null (SM) hypothesis corresponds to $\mu = 0$. Similar to Section 5.8.1, likelihoods are written as Poisson distributions. The test statistic approaches

$$t_{\mu,\mathrm{obs}} \approx \chi^2_{\mu,\mathrm{obs}} = \left(\vec{x} - \vec{\mu} \cdot \vec{s} - \vec{b}\right)^{\mathrm{T}} \mathrm{Cov.}^{-1} \left(\vec{x} - \vec{\mu} \cdot \vec{s} - \vec{b}\right) \tag{6.2}$$

in the large-sample limit, i.e. where Poisson likelihoods approximate normal distributions, where the covariance matrix Cov. includes the statistical and systematic uncertainties on the data, SM prediction and BSM signal. The probability distribution function for the test statistics follows a central chi-square distribution for n degrees of freedom, such that

$$p_\mu = \int_{t_{\mu,\mathrm{obs}}}^{\infty} f(t_\mu|\mu)\mathrm{d}t_\mu = \int_{\chi^2_{\mu,\mathrm{obs}}}^{\infty} \chi^2_n. \tag{6.3}$$

In the alternative hypothesis μ offers one additional degree of freedom such that $n = 1$. First, the $\mathrm{CL_s}$ value is calculated for each histogram in the CONTUR database following Equation 6.2. Again, this takes into account all bins if statistical correlations for the data are provided on HEPData, or otherwise it evaluates the $\mathrm{CL_s}$ exclusion for each bin individually and continues with the bin that has the highest exclusion. Then, the most sensitive histograms (or bins within those) from each pools are concatenated into a single list from which the global likelihood (combining all the different pools) is obtained.

Additionally, CONTUR calculates expected limits using the same statistical procedure, but replacing the nominal values of the experimentally measured data with the SM predictions, i.e. $\vec{s}_{\mathrm{nom.}} \mapsto \vec{x}_{\mathrm{nom.}}$, while keeping the size of the experimental uncertainties. Furthermore, the sensitivity of the HL-LHC dataset can be estimated by scaling the total experimental uncertainties of the $\sqrt{s} = 13\,\mathrm{TeV}$ LHC measurements by the square root of the ratio to the projected integrated luminosity of $3000\,\mathrm{fb}^{-1}$ during the HL-LHC. The underlying assumption is that the systematic

errors scale down with the statistical uncertainties. This projected sensitivity is conservative in the sense that it uses only currently available SM predictions (that may become more precise in the future) and assumes that the HL-LHC programme does not expand the phase-space regions that are currently measured.

CONTUR is a powerful approach to limit setting, because on a cluster of PCs it is possible to generate 100,000 particle-level events in $\mathcal{O}(\text{hours})$ for $N \times N$ two-dimensional grids in BSM parameter space. With this efficient feedback loop it becomes feasible to study different parameter configurations, or iteratively optimise the grid or targeted phase-space of a BSM model. Many LHC final states are considered at the same time for nearly no extra computing costs, which sometimes leads to unexpected collider signals of the considered BSM theory.

It should be emphasised that CONTUR is a package that only sets exclusions, and that it is not a discovery tool. While large deviations between expected and observed limits *can* be interpreted as a sign of new physics, CONTUR can not make any quantitative inference of this suspicion.

One of the main limitations in CONTUR is its dependence on the preservation and re-interpretability of LHC measurements. Experimental results can only be re-used from model-independent particle-level measurements, where a RIVET routine is provided together with sufficient auxiliary (meta)data on HEPData. Only a subset of analyses that satisfy the requirements of "model-independent" and "particle-level" also provide the SM predictions on HEPData. To make use of those analyses in CONTUR, SM predictions need to be estimated by digitising plots in the original publication or by rerunning the event generation.[2] Furthermore, missing statistical correlations weaken the exclusion limits because in a conservative approach only the most sensitive bin of a histogram is taken into account. Initiatives in ATLAS – following recommendations from Les Houches and Reinterpretation Forum discussions – can promote the publication to HEPData of SM predictions and correlations for every measurement.

6.3 Analysing the gauged baryon number model with CONTUR

This section presents a CONTUR study of a model in which baryon number conservation is implied by a local $\mathrm{U}(1)_\mathrm{B}$ gauge symmetry. The formal definition and new particles associated to this BSM scenario is presented first, observing that the model provides an interesting collider phenomenology with a wide range of final states, including the $p_\mathrm{T}^\mathrm{miss}$+jets final state presented in Chapter 5, making it suitable for reinterpretation using CONTUR.

6.3.1 Motivation and theoretical model

In the Standard Model, global baryon number (B) and total lepton number (L) are accidental symmetries of the full theory. It is possible to construct a model where B, L or both are conserved at the quantum level by extending the SM gauge group with additional Abelian $\mathrm{U}(1)'$ gauge groups as follows [283–286]:

$$\mathcal{G}_\mathrm{SM} \otimes \mathrm{U}(1)'_\mathrm{B} \otimes \mathrm{U}(1)'_\mathrm{L}. \tag{6.4}$$

[2] For measurements with missing SM predictions, CONTUR provides an option to assume that the SM prediction exactly matches the measured data. While this option provides insight into the sensitivity of a measurement to a given model, any constraints from this mode are not included in the global exclusion.

Imposing new gauge symmetries can lead to *gauge anomalies*, which are the non-cancellation of divergences in the full theory. The divergences give rise to many problems, including non-unitarity which in turn leads to probabilistic issues in the quantum field theory (QFT) interpretation. Extra fermionic degrees can be added to cancel the anomaly and to study symmetry breaking, which poses constraints on model at hand [287–290].

In this project, we restrict ourselves to the case where only baryon number is a local gauge symmetry. In addition to new fermionic degrees of freedom, the SM gauge Higgs sector is extended with a $\mathrm{U}(1)_\mathrm{B}$ gauge leptophobic Z_B boson and a Higgs singlet S_B. The local baryonic symmetry $\mathrm{U}(1)_\mathrm{B}$ is broken when S_B acquires a vacuum expectation value and this guarantees the stability of the lightest fermionic field. This fermionic field is the dark matter candidate labelled with χ. A priori, the DM candidate can be represented by either a Majorana or Dirac field. This study is restricted to the case where χ is a Majorana particle, as this is the most generic scenario to cancel gauge anomalies via additional fermionic fields. The model for gauge cancellation using Dirac fields instead is illustrated in Ref. [291].

For Majorana DM, the model has the following interactions:

$$\begin{aligned}\mathcal{L} \supset & -\frac{g_\mathrm{B}}{3}\bar{q}\gamma^\mu q Z_\mu^B \;+\; (D_\mu S_\mathrm{B})^\dagger (D^\mu S_\mathrm{B}) + i\bar{\chi}_L\gamma^\mu D_\mu \chi_L \\ & - V(H, S_\mathrm{B}) - \left(\lambda_\chi \chi_L^T C \chi_L S_\mathrm{B} \;+\; \mathrm{h.c.}\right) \\ & + \mathcal{L}_{\text{anomaly cancellation}},\end{aligned} \tag{6.5}$$

where $\mathcal{L}_{\text{anomaly cancellation}}$ contains the heavy fermionic fields for anomaly cancellation and

$$D^\mu S_\mathrm{B} = \partial^\mu S_\mathrm{B} - 3 i g_\mathrm{B} Z_B^\mu S_\mathrm{B}, \tag{6.6}$$

$$D^\mu \chi_L = \partial^\mu \chi_L - \frac{3}{2} i g_\mathrm{B} Z_B^\mu \chi_L. \tag{6.7}$$

The scalar potential is given by

$$V = -\mu_H^2 H^\dagger H + \lambda_H (H^\dagger H)^2 - \mu_B^2 S_B^* S_\mathrm{B} + \lambda_B (S_B^* S_\mathrm{B})^2 + \lambda_{HB}(H^\dagger H)(S_B^* S_\mathrm{B}), \tag{6.8}$$

with the Higgs bosons

$$H = \begin{pmatrix} h^+ \\ \frac{1}{\sqrt{2}}(h_0 + i a_0) \end{pmatrix} \text{ and } \quad S_\mathrm{B} = \frac{1}{\sqrt{2}}(s_B + i a_B). \tag{6.9}$$

Spontaneous symmetry breaking of baryon number and the electroweak sector occur when the new Higgs boson and the SM boson acquire a vacuum expectation value respectively: $\langle s_B \rangle = v_B$ and $\langle h_0 \rangle = v_0$. The physical Higgs bosons are coupled via the *mixing angle* θ_B:

$$\begin{aligned} h &= h_0 \cos\theta_\mathrm{B} - s_B \sin\theta_\mathrm{B}, \\ h_\mathrm{B} &= s_B \cos\theta_\mathrm{B} + h_0 \sin\theta_\mathrm{B}, \end{aligned} \tag{6.10}$$

6.3.2 Collider signatures

The leptophobic Z_B boson couples to SM quarks and the DM particle χ. The main production mode in a proton-collider is therefore the s-channel $q\bar{q} \to Z_\mathrm{B}$ process. This could be visible as both

resonant and non-resonant production, with Z_B decaying into a pair of quarks or an invisible final state with two χ DM particles. The decay into light quarks leads to a dijet signature, but the dijet production cross-section via Z_B, $\mathcal{O}(10^{-3}\,\mathrm{nb})$ for $g_\mathrm{B} = 1.0$, is orders of magnitude smaller than the SM predicted dijet cross-section of $\mathcal{O}(10^{2}\,\mathrm{nb})$. This makes dijet signatures a difficult final-state to infer limits from using CONTUR. However, for sufficiently high M_{Z_B} the distinctive production of a top-quark pair via $pp \to Z^*_\mathrm{B} \to t\bar{t}$ is allowed. The total $t\bar{t}$ production cross-section is much smaller, with fiducial cuts reducing the phase-space of ATLAS measurements to only $\mathcal{O}(10^{-3}\,\mathrm{nb})$ [292] where BSM signals can have a large impact. Furthermore, if $M_{Z_\mathrm{B}} > 2\chi$, invisible particles are produced in the final state via $Z_\mathrm{B} \to \chi\chi$. If Z_B is produced in association with a highly energetic jet, this leads to a distinctive $p_\mathrm{T}^\mathrm{miss}$+jets signature as $q\bar{q} \to gZ^*_\mathrm{B} \to g\chi\chi$, $gq \to qZ^*_\mathrm{B} \to q\chi\chi$ or $g\bar{q} \to \bar{q}Z^*_\mathrm{B} \to \bar{q}\chi\chi$. At slightly lower cross-sections the Z_B is also produced with a photon instead of a jet, leading to either γ + jets or $\gamma + p_\mathrm{T}^\mathrm{miss}$ final states.

The new Higgs boson h_B is produced mainly via $pp \to Z^*_\mathrm{B} \to Z_\mathrm{B}h_\mathrm{B}$. This production cross-section is relatively small, compared to $Z_\mathrm{B} \to q\bar{q}/\chi\chi$ because the Z^*_B is very off-shell. However, the h_B decay can lead to distinctive signatures that offer good discrimination from SM backgrounds. At low M_{h_B} ($< 150\,\mathrm{GeV}$), the new Higgs decays mainly into a pair of photons. This final-state is precisely measured at the LHC with small fiducial cross-sections of $\mathcal{O}(10^{-1}\,\mathrm{nb})$. At larger masses, h_B can decay as a resonance and non-resonance into pairs of SM EW gauge bosons, i.e. $h_\mathrm{B} \to ZZ$ and $h_\mathrm{B} \to WW$. If kinematically allowed, it can also decay into χ and Z_B, but this depends on the assumptions on the masses of the respective BSM particles.

6.3.3 Generator setup

For the event generation we use HERWIG v7.2.3 [83]. The model-specific features of the gauged baryon number theory are stored in the Universal Feynman Output (UFO) format [293]. The UFO format specifies particles, parameters, vertices and couplings as Python objects, that collectively form the set of Feynman rules used by the event generator. A complete UFO model normally comprises one directory with a Python script for each class of properties, with for example a `particle.py` and `parameters.py`. UFO models are a generic class not limited to a specific generator, and the UFO model is converted into objects that HERWIG can implement using HERWIG's `ufo2herwig` command.

HERWIG has a built-in mechanism to calculate the hard process for $2 \to 2$ scattering, based on the Feynman rules of the BSM model of interest. To this end, we construct a `herwig.in` file corresponding to this BSM model that is then passed to HERWIG for the event generation. In the `herwig.in` file, we use the HPConstructor (from "hard process constructor") object to specify the incoming and outgoing particles for inclusive $2 \to 2$ scattering. SM gluons, quarks and anti-quarks – excluding the (anti-)top quarks – are considered as incoming particles. The h_B and Z_B bosons are added to HPConstructor as outgoing particles. The decays of h_B and Z_B are simulated after the hard scatter process. The constructor for resonant processes ResConstructor defines s-channel processes with an intermediate resonance. To this end, both h_B and Z_B are added as intermediate particles. Incoming particles include again the light quarks, light anti-quarks and gluons. By default CONTUR adds all SM particles as outgoing particles to ResConstructor, although the actual final-state particles that can be constructed from the h_B and Z_B resonances depends on the vertices in the UFO model.

The described LO setup leads to final states with jets, for example via resonant dijet production or the production of heavy bosons in association with jets. Many experimental analyses have a jet p_{T} cut to select only hard and well-measured jets. By default, a large fraction of events generated with HERWIG will have jets below these thresholds. To improve statistics in the BSM event samples and to avoid generating a lot events that are ultimately discarded through jet p_{T} cuts in the RIVET routines anyway, a jet p_{T} cut of $150\,\mathrm{GeV}$ is imposed in the hard scatter process. The impact of the additional jet p_{T} cut at generator-level is studied using the RIVET routine of the $\sqrt{s} = 13\,\mathrm{TeV}$ ATLAS $p_{\mathrm{T}}^{\mathrm{miss}}$+jets analysis, which has a jet p_{T} cut at $120\,\mathrm{GeV}$ cf. Table 5.2 and a cut on $p_{\mathrm{T}}^{\mathrm{miss}}$ at $200\,\mathrm{GeV}$. Before adding a jet p_{T} cut in HERWIG, only 3% of events that pass the cut on $\Delta\phi\left(\mathrm{jet}, p_{\mathrm{T}}^{\mathrm{miss}}\right)$ also pass the signal region ≥ 1 jet selections. The generator-level cut in HERWIG improves this figure to 89%. It may seem counter-intuitive to have a jet p_{T} cut at matrix-element level *larger* than the jet p_{T} cut of the experimental analysis, but in this LO setup the jet p_{T} should balance the $p_{\mathrm{T}}^{\mathrm{miss}}$ from $Z_{\mathrm{B}} \to \nu\nu$ in LO event generation, and as indicated $p_{\mathrm{T}}^{\mathrm{miss}}$ has a higher experimental cut at $200\,\mathrm{GeV}$. Furthermore we are not just analysing the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets final state, but also any other final state with jets, and with lower jet p_{T} cuts it is increasingly more CPU expensive to have reasonable statistics in the phase-space of experimental analyses.

6.3.4 Results

Free parameters in the gauged baryon number model are the coupling g_{B}, masses $M_{Z_{\mathrm{B}}}$, $M_{h_{\mathrm{B}}}$ and M_{χ}, and the mixing angle $\sin\theta_{\mathrm{B}}$. The latter is constrained to be small from direct DM detection constraints and is fixed to $\sin\theta_{\mathrm{B}} = 10^{-2}$ throughout these studies. Different combinations of two-dimensional parameter scans are performed for the BSM coupling and particle masses.

Coarse grained scan for g_{B} and $M_{Z_{\mathrm{B}}}$

First, we examine bounds on Z_{B} as a function of its mass for different values of g_{B}. The values of the coupling are varied between $0.1 < g_{\mathrm{B}} < 1.0$ in increments of 0.1, the mass covers a range of 13 linearly spaced point on the interval $50\,\mathrm{GeV} < M_{Z_{\mathrm{B}}} < 3000\,\mathrm{GeV}$, and the additional free parameters are fixed to $M_{\chi} = 100\,\mathrm{GeV}$ and $M_{h_{\mathrm{B}}} = 200\,\mathrm{GeV}$.

The results are presented in Figure 6.3. The right plot in Figure 6.3a shows the $\mathrm{CL_s}$ exclusion for each individual parameter point on the 10×13 grid, for which events have been generated using HERWIG and compared to existing LHC measurements. The colour coding indicate the level of exclusion. The exclusions for individual grid points are interpolated to obtain exclusion contours at 95% and 68%, shown in the left panel of Figure 6.3a as solid and dashed lines respectively. The left panel also includes the 95% expected exclusion limits as a dotted line. The red dotted line corresponds to the estimated sensitivity from the HL-LHC programme. Figure 6.3a shows the same observed and expected exclusion lines, but the colour coding indicates which measurement pool provided the most exclusion it a given point in parameter space.

For intermediate masses $M_{Z_{\mathrm{B}}} \sim 1500\,\mathrm{GeV}$ the value of g_{B} is excluded down to almost 0.2, while low and high-mass regions allow values of g_{B} up to 0.5 and 0.4 respectively. Couplings of $g_{\mathrm{B}} > 0.5$ are excluded across the entire range of $M_{Z_{\mathrm{B}}}$.

In the region $M_{Z_{\mathrm{B}}} \sim 100\,\mathrm{GeV}$ the main exclusion comes from a precision ATLAS measurement of $Z(\to \ell^-\ell^+) + \mathrm{jet(s)}$ at $\sqrt{s} = 8\,\mathrm{TeV}$ [294]. Figure 6.4 shows one of the differential measurements

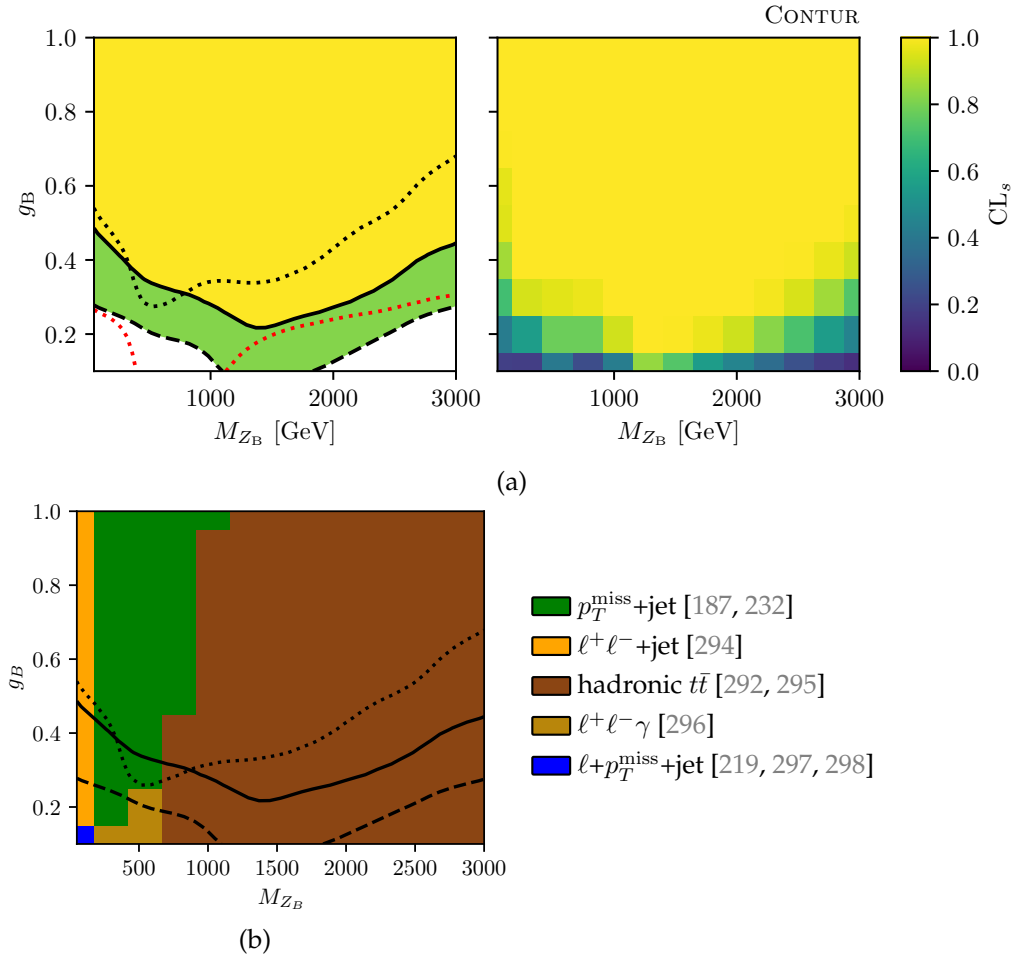


Figure 6.3: CONTUR scan in the g_B-M_{Z_B} plane, indicating the 95% observed exclusion (black solid line), the 68% observed exclusion (black dashed line) and the 95% expected exclusion (dotted line). The red dotted line is the projected HL-LHC sensitivity, and the region above the red dotted lines would be excluded in the HL-LHC scenario at 95% confidence level. Subfigure (a) shows the individual grid points in the right panel, that are interpolated to obtain contour lines in the left panel. Figure (b) shows the same contour lines, but the colour coding indicates which pools give the most sensitivity for a particular point on the plane (see text). The fixed parameters for this scan are $M_{h_B} = 200\,\text{GeV}$, $\sin\theta_B = 0.01$, and $M_\chi = 100\,\text{GeV}$. Figures taken from Ref. [299].

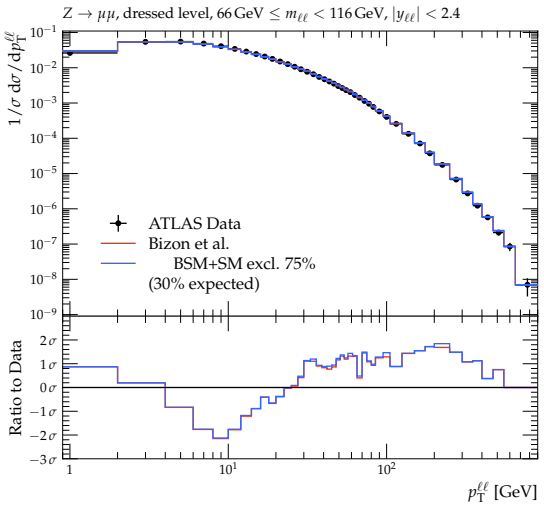


Figure 6.4: Comparison of the gauged baryon model signal to SM predictions and ATLAS data from Ref. [294], for $M_\chi = 100\,\text{GeV}$, $\sin\theta_B = 0.01$, $M_{h_B} = 200\,\text{GeV}$, $g_B = 0.6$ and $M_{Z_B} = 100\,\text{GeV}$. Data is indicated with black markers, the SM prediction as a red histogram and the SM+BSM prediction as a blue histogram. We observe that the BSM contribution leads to an excess in the tail of the p_T distribution of the $\ell\ell$ system, which is primarily due to ZZ_B production where the SM Z decays leptonically.

from this analysis that contributes to the exclusion for $g_B = 0.6, M_{Z_B} = 100\,\text{GeV}$. The plot shows the experimentally observed data as black markers, the SM prediction as a red line and the BSM signal superimposed on the SM background in blue. The ratio panel shows the agreement of the SM and SM+BSM prediction relative to the data, where the vertical axis is scaled to the total uncertainty to show agreement expressed as the number of standard deviations it is away from the data. The BSM leads to an excess in the tail of the p_T distribution of the $\ell\ell$ system, which is primarily due to ZZ_B production where the SM Z decays leptonically.

For larger masses the $Z_B \to \chi\chi$ decay switches on, leading to an exclusion primarily from measurements with a p_T^{miss}+jets final state. The most sensitive LHC measurement of p_T^{miss}+jets is the one presented in Chapter 5, so that our focus on re-interpretability is already paying off in this study. The corresponding particle-level distributions with the BSM signal injected on top of it are shown in Figure 6.5. The ≥ 1 jet p_T^{miss} distribution is the most sensitive observable, because theoretical uncertainties of the NNLO reweighted SM prediction are much smaller than those for the MEPS@NLO prediction in the VBF distributions.

At sufficiently high M_{Z_B}, the new Z_B boson can decay into a top-quark pair. As shown in Figure 6.6, this emerges as a resonance at M_{Z_B} of the invariant mass of the $t\bar{t}$ system and leads to an excess in the high p_T tails of the top-quarks. There is an overestimation of the data from the Standard Model POWHEG +PYTHIA8 prediction. Adding BSM on top of this only brings the SM+BSM yields further away from the experimental data. Although the CL_s method accounts for this in the observed limits, the 95% observed exclusions are stronger than the 95% expected exclusions. This feature also seen in Figure 6.3, where the solid lines excludes smaller values of g_B than the dotted line. The $t\bar{t}$ signal gives strong exclusions, and is fully efficient around $M_{Z_B} = 1500\,\text{GeV}$. For higher masses the production-cross section of M_{Z_B} decreases, ultimately reducing the sensitivity reach in g_B.

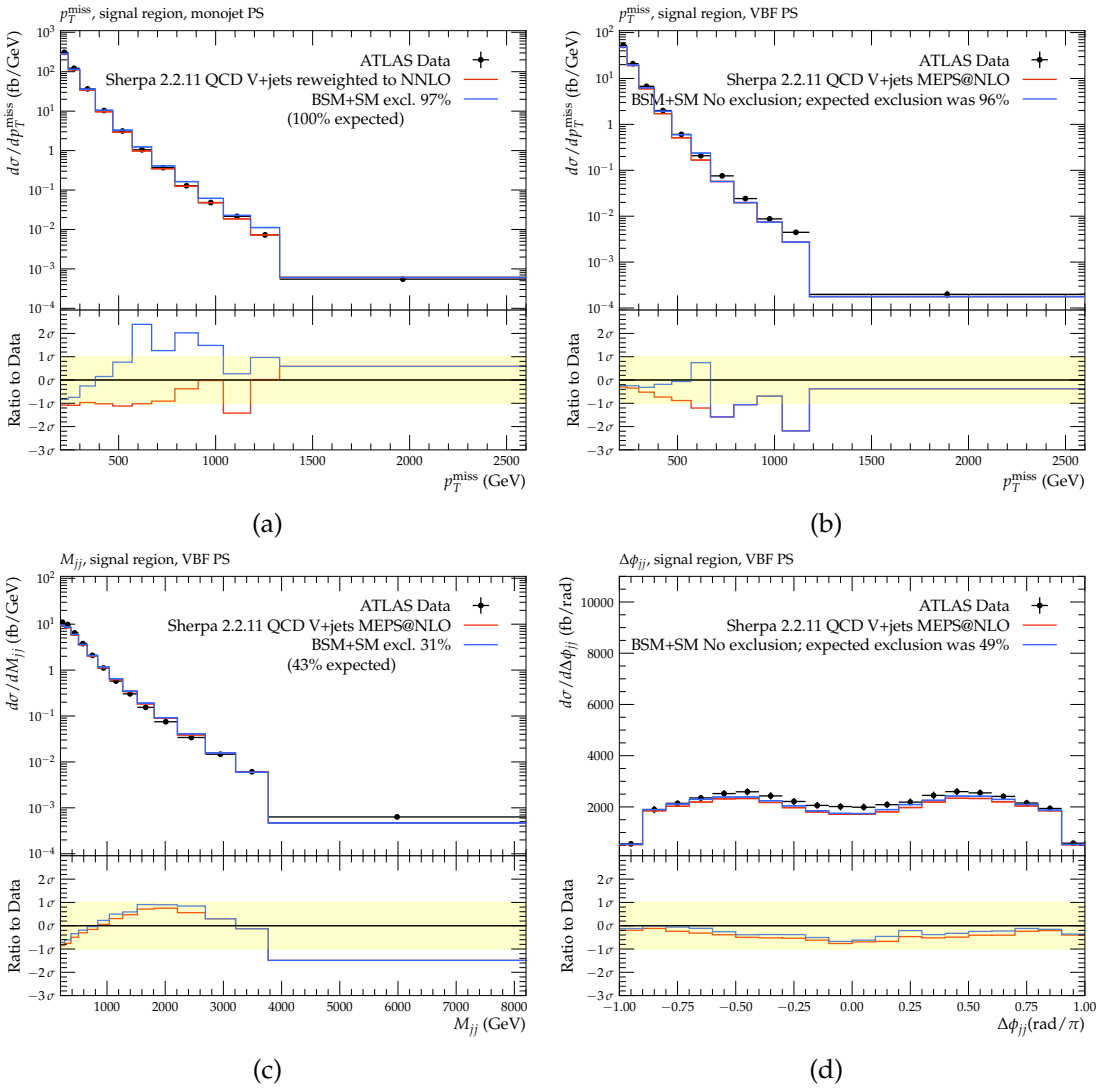
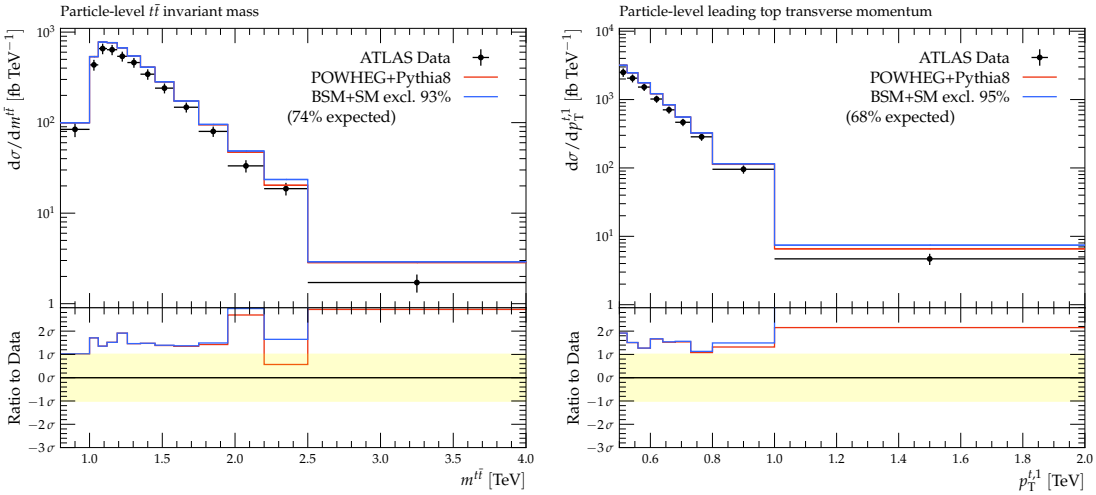


Figure 6.5: Comparison of the gauged baryon model signal to SM predictions and ATLAS data from Ref. [232], for $M_\chi = 100\,\text{GeV}$, $\sin\theta_\text{B} = 0.01$, $M_{h_\text{B}} = 200\,\text{GeV}$, $g_\text{B} = 0.4$ and $M_{Z_\text{B}} = 542\,\text{GeV}$. Data is indicated with black markers, the SM prediction as a red histogram and the SM+BSM prediction as a blue histogram. For this particular parameter point, most of the constraining power comes from the p_T^miss observable in the monojet phase-space. The VBF phase-space, as explained in the chapter on the p_T^miss+jets analysis, only has NLO samples available for the SM predictions, while the monojet phase-space additionally has NNLO reweighted samples with much smaller theoretical uncertainties. CONTUR uses the NNLO samples for the monojet phase-space, which leads to stronger constraints for this region compared to the VBF phase-space.



(a) (b)

Figure 6.6: Comparison of the gauged baryon model signal to SM predictions and ATLAS data from Ref. [292], for $M_\chi = 100\,\text{GeV}$, $\sin\theta_\text{B} = 0.01$, $M_{h_\text{B}} = 200\,\text{GeV}$, $g_\text{B} = 0.4$ and $M_{Z_\text{B}} = 2508\,\text{GeV}$. Data is indicated with black markers, the SM prediction as a red histogram and the SM+BSM prediction as a blue histogram. The SM only prediction already over-predicts the data across the spectrum. By adding additional BSM on top the disagreement with data only becomes larger. Therefore the observed exclusions are normally stronger than the expected conclusion from this measurement, as can be seen explicitly in for example Figure 6.3.

CONTUR considers particle-level measurements, but the gauged baryon model can also be confronted with data from detector-level searches. Search bounds in the same g_B-M_{Z_B} plane are presented in Figure 6.7, with the CONTUR measurement constraints superimposed as a dark green line. The search bounds were obtained under the assumption that M_{Z_B} is decoupled from the Higgs sector and that it decays only into SM quarks. This rules out $g_\text{B} > 0.2$ for intermediate M_{Z_B} masses, while $g_\text{B} < 0.3$ in the mass range $M_{Z_\text{B}} \lessapprox 500\,\text{GeV}$.

Low M_{Z_B} region

Figure 6.8 shows similar two-dimensional scans in the g_B-M_{Z_B} plane, now zooming in on the region of low M_{Z_B} mass below 500 GeV to resolve interpolation effects. As expected from the full-range scan, there are constraints from $Z(\to \ell^-\ell^+)$ + jet(s) at $\sqrt{s} = 8\,\text{TeV}$, while p_T^miss+jets dominates where $M_{Z_\text{B}} > 2M_\chi$. A new signature emerges, as this high granularity scan exposes diphoton measurements as the most sensitive final state in the region $150\,\text{GeV} < M_{Z_\text{B}} < 250\,\text{GeV}$. This photon signal is due to $h_\text{B} \to \gamma\gamma$ decays, which has branching fraction of approximately 25% for the choice $M_{h_\text{B}} = 200\,\text{GeV}$ and $M_\chi = 100\,\text{GeV}$. Although the production of $pp \to Z_\text{B}^* \to Z_\text{B} h_\text{B}$ has a relatively small cross-section, the distinctive $h_\text{B} \to \gamma\gamma$ decay delivers strong exclusions due to the excellent resolution on photon measurements. Figure 6.9 shows exclusions obtained from Ref. [316]. This confirms that the diphoton pair is the result of a M_{h_B} decay, as the invariant diphoton mass peaks sharply at $M_{h_\text{B}} = 200\,\text{GeV}$. Additionally, the new BSM model leads to a sharp excess in the p_T tails of the reconstructed photons.

The decay $Z_\text{B} \to \chi\chi$ can be switched off by setting $M_\chi \gg M_{Z_\text{B}}/2$. The results of a scan with $M_\chi = 10\,\text{TeV}$ is shown in Figure 6.8b, showing that the diphoton signature alone is capable of excluding a large part of parameter space. Furthermore, exclusions arise from h_B decays into two

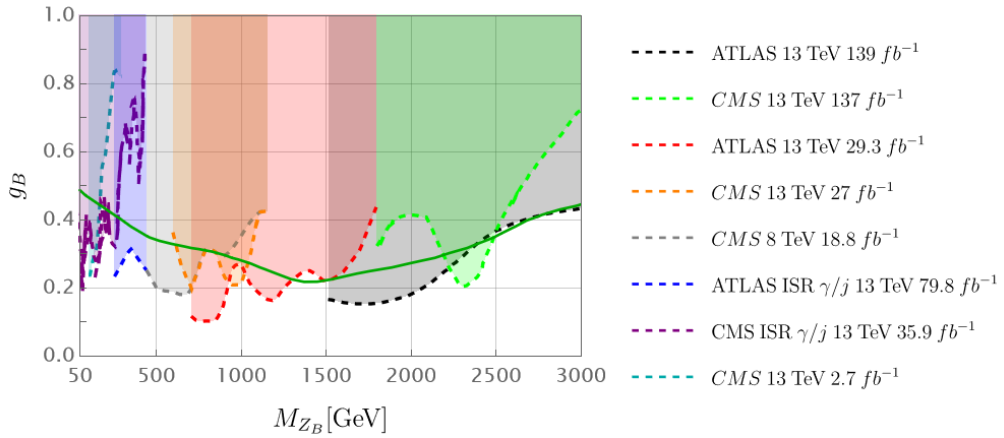


Figure 6.7: Experimental bounds on g_B-M_{Z_B} from LHC searches [300–307] (see legend) and CONTUR constraints corresponding to the 95% confidence level exclusion shown in Figure 6.3 (in dark green). The exclusion contours using searches only are slightly better than the CONTUR constraints, because it can target regions of phase-phase that are not covered by LHC measurements (yet). Figure taken from Ref. [299].

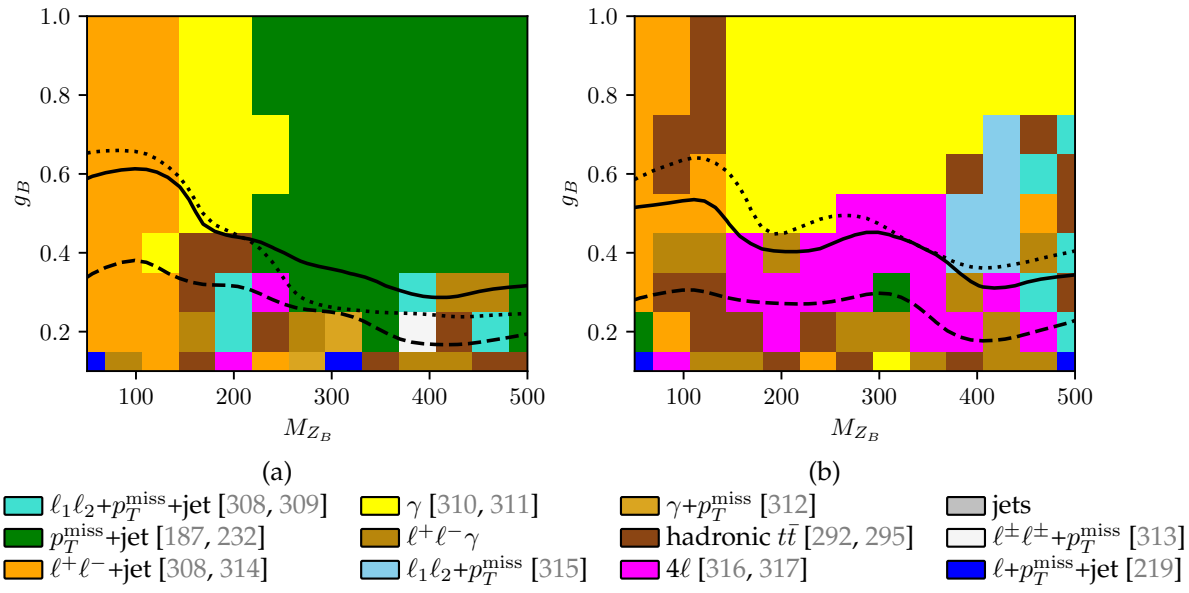
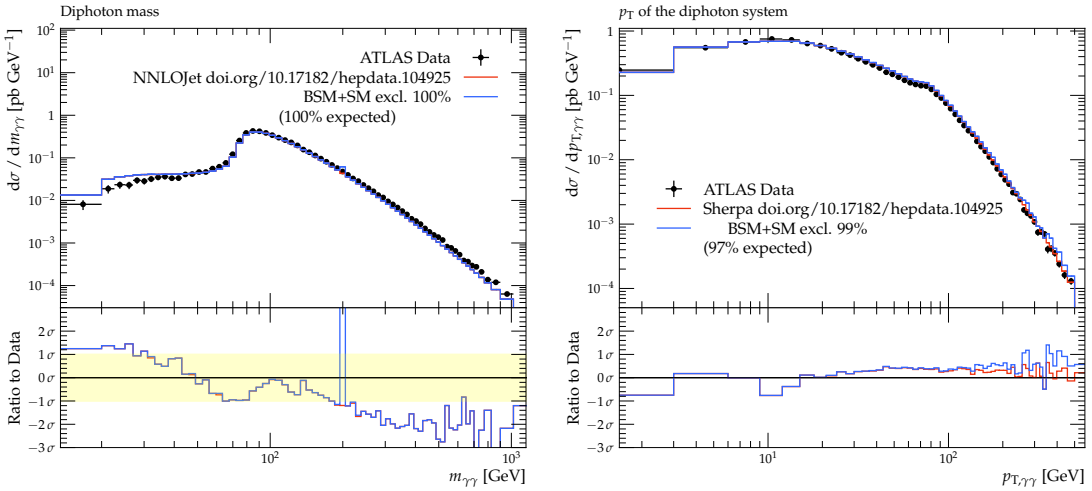


Figure 6.8: (a) Results from CONTUR, as in Figure 6.3b, but now zooming in on the low M_{Z_B} region. (b) as in (a), but now with $M_\chi = 10\,\text{TeV}$. An explanation of different measurements contributing to the overall exclusion is given in the text. Figures taken from Ref. [299].



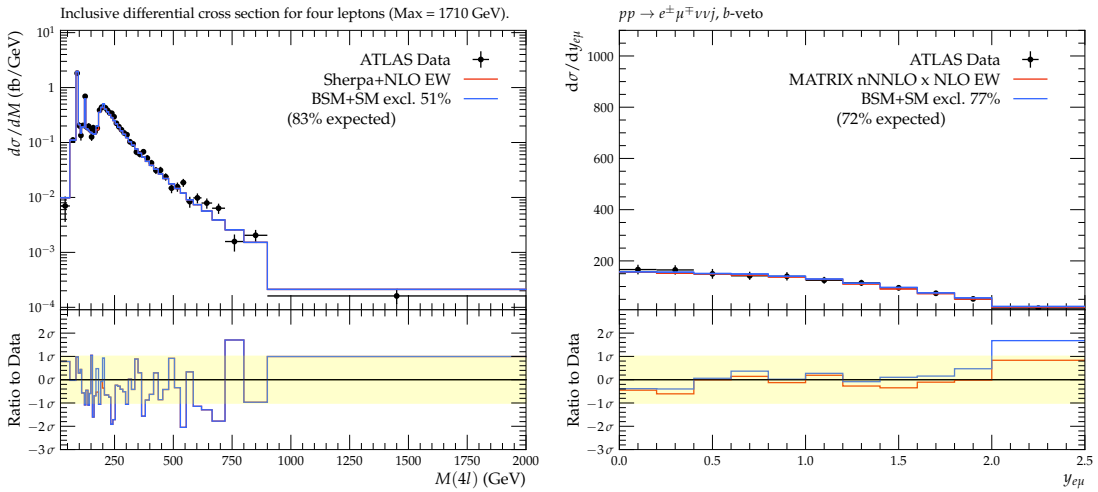
(a) (b)

Figure 6.9: Comparison of the gauged baryon model signal to SM predictions and ATLAS data from Ref. [316], for $M_\chi = 10\,\text{TeV}$, $\sin\theta_\text{B} = 0.01$, $M_{h_\text{B}} = 200\,\text{GeV}$, $g_\text{B} = 0.6$ and $M_{Z_\text{B}} = 225\,\text{GeV}$. Data is indicated with black markers, the SM prediction as a red histogram and the SM+BSM prediction as a blue histogram. In subfigure (a), we see a sharp excess precisely at $m_{\gamma\gamma}$, i.e. peaking around the chosen parameter $M_{h_\text{B}} = 200\,\text{GeV}$. This particular value of 200 GeV is just an illustration: for different parameter values of M_{h_B}, the $m_{\gamma\gamma}$ peak shifts accordingly. In figure (b) we see that the diphoton decay can also lead to an excess in the tails of the transverse momentum of the diphoton system.

heavy gauge bosons, i.e. $ZZ \to 4\ell$ in pink or $WW \to \ell_1\nu_v\ell_2\nu_2$ in light blue. Indeed, Figure 6.10a confirms that the invariant mass plots of the four-lepton mass have a sharp peak in the BSM signal at $M_{h_\text{B}} = 200\,\text{GeV}$. The invariant mass of the WW system is not fully reconstructed experimentally, but kinematic distributions such as the rapidity of the e, μ system give exclusions of the BSM parameter space as shown in Figure 6.10b.

Fixed coupling

Furthermore, we investigate the dependence on other model parameters for fixed values of g_B, that are guided by the allowed parameter space from the searches. We perform multiple scans, for M_{Z_B} against both M_{h_B} and M_χ, all across values between 100 GeV and 5 TeV as shown in Figure 6.11. We consider $g_\text{B} = 0.15$ because this is unconstrained across most of the M_{Z_B} range from the searches, but also $g_\text{B} = 0.25$ which is still viable in the low M_{Z_B} region. Figure 6.11a fixes $M_\chi = 100\,\text{GeV}, g_\text{B} = 0.15$ which results in an exclusion only at 68% confidence intervals. For this small value of g_B, the production cross-section of h_B becomes so small that it effectively switches of. Therefore, Figure 6.11a shows no dependence on M_{h_B} in terms of the exclusion limits. Nonetheless we observe that the dominant exclusion of some parameter points – albeit below 68% CL_s– comes from the diphoton measurement. The dependence on the mass of Z_B is similar to that of Figure 6.3, where $Z_\text{B} \to t\bar{t}$ leads to exclusions for sufficiently large M_{Z_B}. This scan in the M_{Z_B}-M_{h_B} plane also consistent wit Figure 6.3 in that the p_T^miss+jets signal is too weak at $g_\text{B} = 0.15$ to yield exclusions at 68% CL_s. The projected HL-LHC sensitivity in red, however, excluded a large range of $300\,\text{GeV} < M_{Z_\text{B}} < 1500\,\text{GeV}$ at 95% CL_s. A similar picture emerges in Figure 6.11a with $M_\chi = 10\,\text{TeV}, g_\text{B} = 0.15$. There is a slight increase of sensitivity at high M_{h_B}



(a) (b)

Figure 6.10: Comparison of the gauged baryon model signal to SM predictions and ATLAS data, for (a) Ref. [317] with $M_{Z_B} = 312.5\,\text{GeV}$, and (b) Ref. [315] with $M_{Z_B} = 425\,\text{GeV}$. In both figures, $M_\chi = 10\,\text{TeV}$, $\sin\theta_B = 0.01$, $M_{h_B} = 200\,\text{GeV}$, $g_B = 0.5$. Data is indicated with black markers, the SM prediction as a red histogram and the SM+BSM prediction as a blue histogram. Subfigure (a) shows an excess of BSM signal at $M(4\ell) \approx 200\,\text{GeV}$, consistent with the decay $h_B \to ZZ \to 4\ell$. The decay $h_B \to WW$, where both W bosons decay leptonically, leads to excesses in kinematic distributions of the final-state leptons as illustrated in subfigure (b).

mass, because $h_B \to \chi\chi$ is forbidden which leads to an increase branching ratio of its decay into two gauge bosons.

We also consider the sensitivity in a plane of M_χ-M_{Z_B}, for $g_B = 0.15$ and $g_B = 0.25$ for Figure 6.11c and 6.11d respectively. This demonstrates a similar dependence on M_{Z_B} as observed before. For values of $g_B = 0.15$ the BSM signal is too small to lead to exclusions at 95% CL_s, but this is recovered with $g_B = 0.25$. In both scenarios, the limits at $M_{Z_B} = 1000\,\text{GeV}$ from the hadronic $t\bar{t}$ measurements are slightly weaker in the low M_χ region. This is explained because Z_B decays predominantly into DM in this region, reducing the cross-section of $t\bar{t}$-production via the BSM resonance. For $g_B = 0.25$, the p_T^{miss}+jets signature excludes a small region at low M_χ at 68% CL_s, which is consistent with the limits from Figure 6.8a. The projected HL-LHC sensitivity shows that that Z_B with a mass as low as 200 GeV across the full M_χ range can be discovered or rejected in the near future.

6.3.5 Additional dark matter constraints

Dark matter models are constrained by cosmological limitations on the relic density and the scattering of SM nucleons as probed by direct detection experiments. Since the gauged baryon model predicts a DM candidate, both types of DM constraints can be imposed on the parameter space of this model. This demonstrates a direct connection between areas of physics at the largest scale, cosmology, and the smallest scale, particle physics. The Majorana DM candidate can annihilate via the following processes:

$$\chi\chi \to q\bar{q}, Z_B Z_B, Z_B h_B, h_B h_B, hh, hh_B, WW, ZZ. \tag{6.11}$$

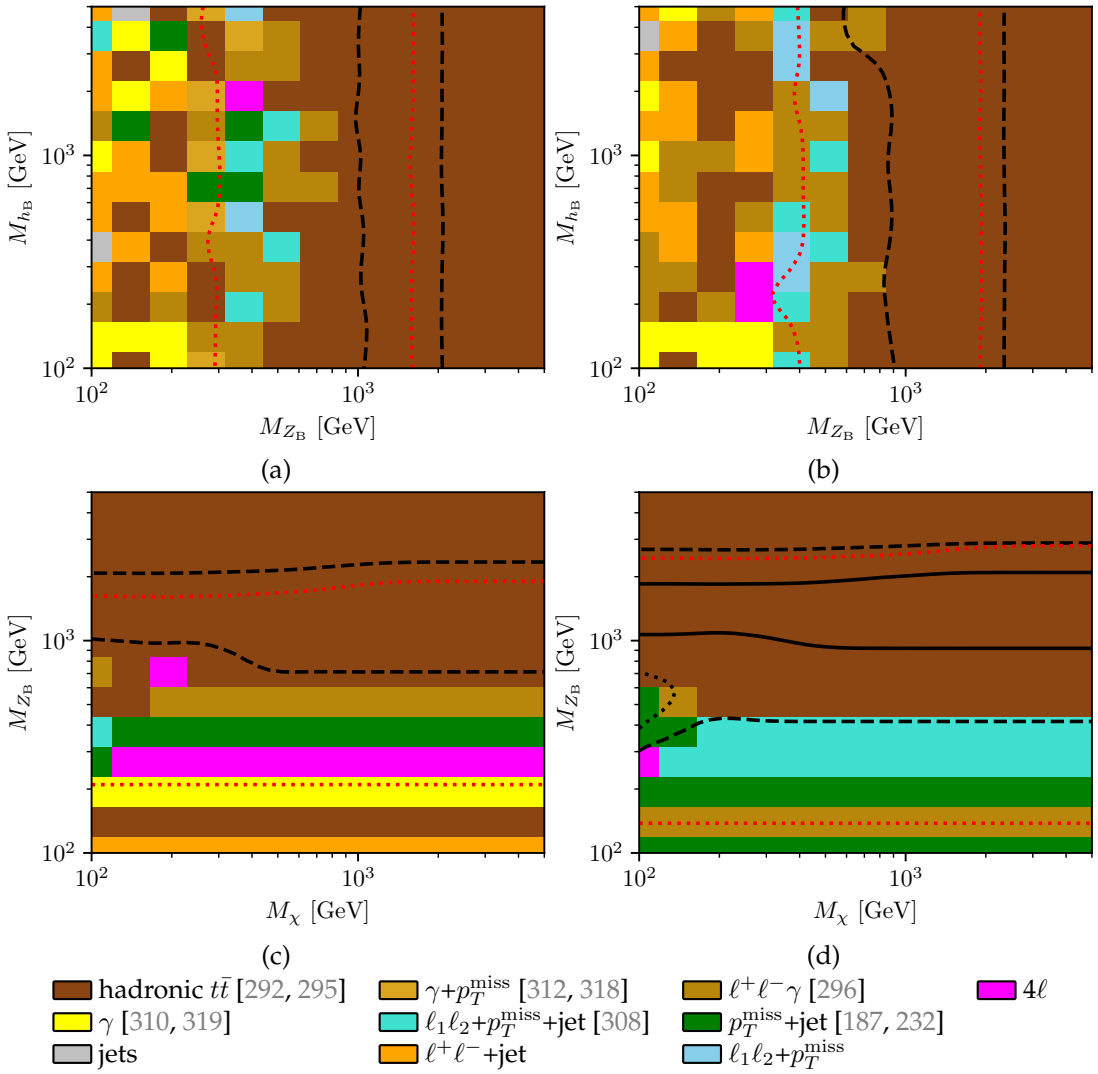
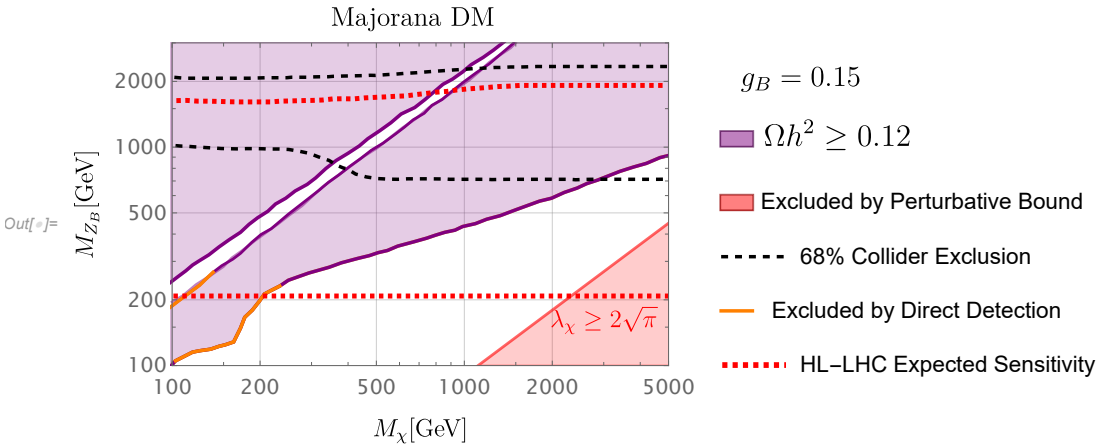


Figure 6.11: CONTUR scans in the M_{Z_B}-M_{h_B} and M_χ-M_{Z_B} planes, indicating the 95% observed exclusion (solid line), the 68% observed exclusion (dashed line) and the 95% expected exclusion (dotted line). The red dotted line is the projected HL-LHC sensitivity. The excluded areas lie *between* the solid or dashed lines respectively. The mass M_χ is fixed to 100 GeV and 10 TeV for subfigures (a) and (b) respectively, both with $\sin\theta_B = 0.01, g_B = 0.15$. Subfigures (c) and (d) are generated with couplings $g_B = 0.15$ and $g_B = 0.25$ respectively, with the other parameters fixed to $\sin\theta_B = 0.01, M_{h_B} = 200$ GeV. In subfigure (b) we see an increased sensitivity at large M_{h_B}, around $M_{Z_B} \approx 600$ GeV compared to subfigure (a). Because the $Z_B \to \chi\chi$ decay is effectively switched off in figure (b), there is a larger decay width into final states (e.g. $\gamma + p_T^{\text{miss}}$) that are more constraining. In figures (c) and (d), the hadronic $t\bar{t}$ measurement yields most part of the exclusions. At lower M_χ, the $Z_B \to \chi\chi$ becomes the dominant decay mode, leading to a reduced sensitivity of the $t\bar{t}$. This is not completely counterbalanced by an increased sensitivity from the p_T^{miss}+jets measurements, although the expected exclusions from the p_T^{miss}+jets analysis pool become visible as a dotted line around $M_\chi \approx 100$ GeV in subfigure (d). At large M_{Z_B} we observe a (counter-intuitive) discrepancy that the current *observed* limits are stronger than the expected HL-LHC sensitivity. This happens because the SM predictions of the hadronic $t\bar{t}$ measurement overestimate the data, as shown in Figure 6.6, leading to stronger observed than expected exclusions. Figures taken from Ref. [299].

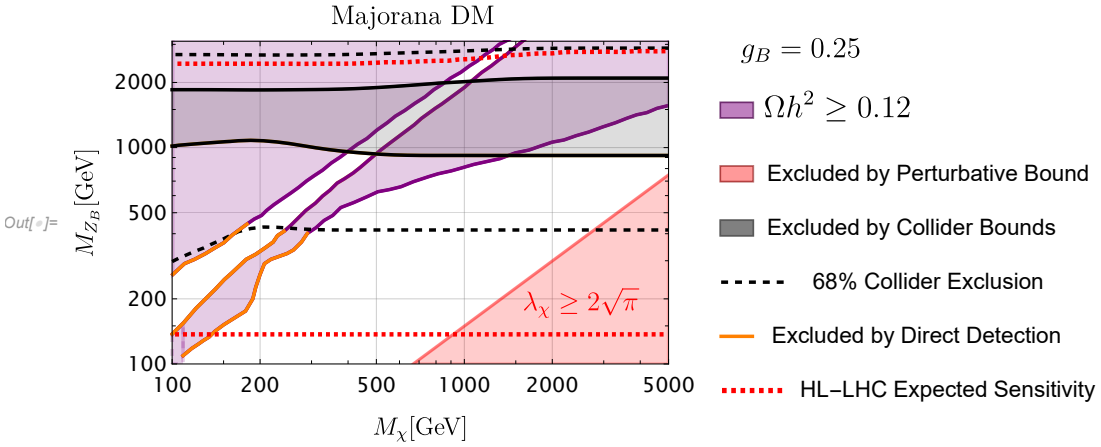
Constraints on the relic density from cosmological observations, $\Omega h^2 \leq 0.12$ [20], can be satisfied for both $g_{\mathrm{B}} = 0.15$ and $g_{\mathrm{B}} = 0.25$ even when M_χ and $M_{Z_{\mathrm{B}}}$ are close to the electroweak scale.

The cross-section of nucleon-DM scattering $\sigma_{\chi,N}$ – via the exchange of H, h_{B} or Z_{B} gauge bosons – can be calculated analytically and compared to direct detection bounds in the plane of M_χ-$\sigma_{\chi,N}$. This excludes values of $100\,\mathrm{GeV} < M_\chi < 200\,\mathrm{GeV}$ and $100\,\mathrm{GeV} < M_\chi < 400\,\mathrm{GeV}$ for $g_{\mathrm{B}} = 0.15$ and $g_{\mathrm{B}} = 0.25$ respectively, but values of $M_{Z_{\mathrm{B}}}$ are still allowed to be close to the EW scale.

The combined set of constraints from collider physics, cosmology and direct-detection experiments are complementary, but still leave a large window of allowed parameter space of the gauged baryon model, as shown in Figure 6.12. All in all we see a well-motivated, anomaly-free model that is compatible with constraints from different kinds experimental measurements. Even more so, particle masses can be well below the TeV-scale and hence well within reach of the Large Hadron Collider at CERN. As demonstrated in Figure 6.12, the projected sensitivities of the HL-LHC program cover a large part of the parameter space that is still allowed. This model a viable candidate for further study in the LHC program, where a discovery is possible in the short to intermediate term.



(a)



(b)

Figure 6.12: Summary of different bounds in the M_χ-M_{Z_B} plane for a coupling of $g_B = 0.15$ and $g_B = 0.25$ for figure (a) and (b) respectively. For the other parameters $\sin\theta_B = 0.01$ and $M_{h_B} = 200\,\text{GeV}$ were assumed. The white regions in the 2D parameter space are still compatible with experimental observations, where the different exclusions, indicated with coloured regions and lines are derived from the following. Observed collider constraints using CONTUR are shown between black solid lines as a grey area for the 95% confidence level. The corresponding constraints from CONTUR at 68% confidence level is the region between the dashed black lines. The red dotted line is the projected HL-LHC sensitivity, and the area between these two lines would be excluded at the HL-LHC under the assumptions given in the text. The purple shaded region is excluded, as in this region the model would lead to a larger relic DM abundance than the experimentally measured value. The solid purple line along the edges of the purple regions indicate the model parameters for which Ωh^2 matches the experimentally measured value of 0.12 exactly. The perturbative bound, which excludes the red shaded region, follows from the requirement of having a new Higgs potential that is bounded from below. The orange line at $\Omega h^2 = 0.12$ is ruled out from limits on the DM-nucleon cross-section in direct-detection experiments. Figures taken from Ref. [299].

Chapter 7

Conclusion

This thesis motivates a shift in approach to new physics searches at the Large Hadron Collider, particularly during Run-3 data-taking and the High-Luminosity LHC era. Where traditional searches normally have a short shelf life, precision Standard Model measurements are a complementary approach that allow swift and simple reinterpretations of experimental data for a multitude of beyond the Standard Model scenarios. In this work, the first inclusive $p_{\mathrm{T}}^{\mathrm{miss}}$ measurement with the full ATLAS Run-2 dataset of $140\,\mathrm{fb}^{-1}$ is presented. This analysis is optimised to be maximally re-interpretable. Using a benchmark DM model it is shown that such measurements are competitive in terms of BSM sensitivity to dedicated collider BSM searches. To back up the claim that re-interpretable measurements are powerful assets, the published ATLAS $p_{\mathrm{T}}^{\mathrm{miss}}$+jets results are immediately used to set stringent constraints on a theoretical model that predicts a DM candidate from anomaly cancellation.

Differential cross-section measurements of missing energy in association with jets are performed using the ATLAS detector. In addition to the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets final state, four auxiliary regions with charged leptons are measured. These lepton regions are used to constrain experimental and theoretical systematics in the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region. Results are also presented as ratios between the $p_{\mathrm{T}}^{\mathrm{miss}}$ region and the lepton regions (R^{miss}). In addition to an inclusive monojet topology, each region is also measured in a VBF-enhanced phase space defined by additional jet selections. The analysis is performed with re-interpretability in mind, exemplified by two important design choices. First, the data is corrected for detector effects in an iterative Bayesian unfolding technique. Second, the phase-space is defined in terms of the final state particles rather than final state processes. In other words, subdominant SM processes that enter the measurement regions are not subtracted from data but rather kept as signal.

Experimental data is compared to theoretical NLO QCD V+jets predictions. These show adequate shape modelling for the $p_{\mathrm{T}}^{\mathrm{miss}}$ observable, but there is an overall offset of 10–20% of the total cross-section. MC predictions from NNLO reweighted QCD V+jets samples show improved modelling of the total cross-section and significantly smaller theoretical uncertainties. The shape of m_{jj} in the VBF region is mismodelled by the NLO predictions, and alternative predictions from HEJ – only available for the lepton regions – show slight improvements in the shape modelling. The agreement between data and MC is quantified with a χ^2 fit, which is best for the R^{miss} results where mismodelling effects largely cancel.

The ATLAS $p_{\mathrm{T}}^{\mathrm{miss}}$+jets measurement is interpreted in terms of two BSM scenarios. First, a pseudo-scalar simplified DM model is tested to benchmark the sensitivity against a detector-level mono-jet search using the same data. Competitive limits are obtained, and hence this work shows that precision SM measurement can be as good of an asset in BSM limit setting than dedicated BSM searches. To demonstrate the sensitivity and flexibility of the measurement, limits are also set on parameters of the 2HDM+a model.

The analysis logic of the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets measurement is preserved in RIVET, while numerical (meta)data will be publicly available on HEPData. Using the CONTUR software, a re-interpretation of these public results is made for a BSM scenario in which a DM candidate is predicted from gauge anomaly cancellation. CONTUR includes a large set of LHC differential cross-section measurements from the RIVET library to compare new BSM models to a wide range of already measured final states. This work presents the first limits on this model using particle-level measurements. The CONTUR framework lends itself optimally for exploratory studies of a new model, because CPU-intensive detector simulation is not required which results in a quick feedback-loop. Large regions of the parameter space are excluded by the new $p_{\mathrm{T}}^{\mathrm{miss}}$+jets analysis, and these collider constraints are combined with results from DM searches and cosmological measurements of the cosmic microwave background. Projected sensitivities show that a large range of the parameter space is accessible during HL-LHC data-taking, so that a discovery is possible in the medium term.

Model-independent cross-section measurements are proven to be extremely powerful in a myriad of different ways, but only if sufficient (digital) material is published outside the experimental collaborations. The "handover" from experimentalists to the outside world can be a bottleneck where useful and labour-intensive physics data is lost into the void. At the dawn of the High-Luminosity LHC program – where we will see the last hadronic collisions for a very long time – it is critical to follow up on collective efforts to publish full statistical models, machine learning algorithms and any other tools used in analysis to benefit re-interpretation. This thesis serve to demonstrate how re-interpretable measurements are performed and re-used, and to convey the benefits of this approach in an urgent message to the rest of the particle physics community.

References

[1] J. J. Thomson. "Cathode rays". In: *Phil. Mag. Ser. 5* 44 (1897), pp. 293–316. DOI: 10.1080/14786449708621070.

[2] C. Burgard and D. Galbraith. *Example: Standard model of physics*. https://texample.net/tikz/examples/model-physics/. Accessed: 1 May 2024.

[3] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 Feb. 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[4] S. L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[5] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264.

[6] Abdus Salam and John Clive Ward. "Electromagnetic and weak interactions". In: *Phys. Lett.* 13 (1964), pp. 168–171. DOI: 10.1016/0031-9163(64)90711-5.

[7] Alexandre Deur, Stanley J. Brodsky, and Guy F. de Teramond. "The QCD Running Coupling". In: *Nucl. Phys.* 90 (2016), p. 1. DOI: 10.1016/j.ppnp.2016.04.003. arXiv: 1604.08082 [hep-ph].

[8] Yuval Grossman. "Introduction to flavor physics". In: *2009 European School of High-Energy Physics*. 2010, pp. 111–144. DOI: 10.5170/CERN-2010-002.111. arXiv: 1006.3534 [hep-ph].

[9] ATLAS Collaboration. "Search for resonances in diphoton events at $\sqrt{s}$=13 TeV with the ATLAS detector". In: *JHEP* 09 (2016), p. 001. DOI: 10.1007/JHEP09(2016)001. arXiv: 1606.03833 [hep-ex].

[10] ATLAS Collaboration. "Search for scalar diphoton resonances with 15.4 fb^{-1} of data collected at $\sqrt{s}$=13 TeV in 2015 and 2016 with the ATLAS detector". In: *ATLAS-CONF-2016-059* (Aug. 2016).

[11] CMS Collaboration. "Search for Resonant Production of High-Mass Photon Pairs in Proton-Proton Collisions at $\sqrt{s}$ =8 and 13 TeV". In: *Phys. Rev. Lett.* 117.5 (2016), p. 051802. DOI: 10.1103/PhysRevLett.117.051802. arXiv: 1606.04093 [hep-ex].

[12] CMS Collaboration. "Search for resonant production of high mass photon pairs using 12.9 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV and combined interpretation of searches at 8 and 13 TeV". In: *CMS-PAS-EXO-16-027* (2016).

[13] LHCb Collaboration. "Test of lepton universality in beauty-quark decays". In: *Nature Phys.* 18.3 (2022). [Addendum: Nature Phys. 19, (2023)], pp. 277–282. DOI: 10.1038/s41567-023-02095-3. arXiv: 2103.11769 [hep-ex].

[14] LHC Collaboration. "Test of lepton universality in $b \to s\ell^+\ell^-$ decays". In: *Phys. Rev. Lett.* 131.5 (2023), p. 051803. DOI: 10.1103/PhysRevLett.131.051803. arXiv: 2212.09152 [hep-ex].

[15] X. Fan et al. "Measurement of the Electron Magnetic Moment". In: *Phys. Rev. Lett.* 130.7 (2023), p. 071801. DOI: 10.1103/PhysRevLett.130.071801. arXiv: 2209.13084 [physics.atom-ph].

[16] ATLAS Collaboration. *Standard Model Summary Plots October 2023*. Tech. rep. Geneva: CERN, 2023. URL: https://cds.cern.ch/record/2882448.

[17] Gianfranco Bertone and Dan Hooper. "History of dark matter". In: *Rev. Mod. Phys.* 90.4 (2018), p. 045002. DOI: 10.1103/RevModPhys.90.045002. arXiv: 1605.04909 [astro-ph.CO].

[18] F. Zwicky. "Die Rotverschiebung von extragalaktischen Nebeln". In: *Helv. Phys. Acta* 6 (1933), pp. 110–127. DOI: 10.1007/s10714-008-0707-4.

[19] M.S. Roberts and A.H. Rots. "Comparison of Rotation Curves Of Different Galaxy Types". English. In: *Astron. Astrophys.* 26.3 (1973), pp. 483–485. ISSN: 0004-6361.

[20] Planck Collaboration. "Planck 2018 results. VI. Cosmological parameters". In: *Astron. Astrophys.* 641 (2020). [Erratum: Astron.Astrophys. 652, C4 (2021)], A6. DOI: 10.1051/0004-6361/201833910. arXiv: 1807.06209 [astro-ph.CO].

[21] Mark Thomson. *Modern particle physics*. New York: Cambridge University Press, 2013. ISBN: 978-1-107-03426-6. DOI: 10.1017/CBO9781139525367.

[22] P. Tisserand et al. "Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds". In: *Astron. Astrophys.* 469 (2007), pp. 387–404. DOI: 10.1051/0004-6361:20066017. arXiv: astro-ph/0607207.

[23] B. J. Carr and S. W. Hawking. "Black holes in the early Universe". In: *Mon. Not. Roy. Astron. Soc.* 168 (1974), pp. 399–415. DOI: 10.1093/mnras/168.2.399.

[24] B. J. Carr et al. "Constraints on primordial black holes". In: *Rept. Prog. Phys.* 84.11 (2021), p. 116902. DOI: 10.1088/1361-6633/ac1e31. arXiv: 2002.12778 [astro-ph.CO].

[25] M. Milgrom. "A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis." In: *Astrophys. J.* 270 (July 1983), pp. 365–370. DOI: 10.1086/161130.

[26] Benjamin W. Lee and Steven Weinberg. "Cosmological Lower Bound on Heavy Neutrino Masses". In: *Phys. Rev. Lett.* 39 (1977). Ed. by M. A. Srednicki, pp. 165–168. DOI: 10.1103/PhysRevLett.39.165.

[27] J. L. Feng. "The WIMP paradigm: Theme and variations". In: *SciPost Phys. Lect. Notes* 71 (2023), p. 1. DOI: 10.21468/SciPostPhysLectNotes.71. arXiv: 2212.02479 [hep-ph].

[28] Hans Peter Nilles. "Supersymmetry, Supergravity and Particle Physics". In: *Phys. Rept.* 110 (1984), pp. 1–162. DOI: 10.1016/0370-1573(84)90008-5.

[29] Howard E. Haber and Gordon L. Kane. "The Search for Supersymmetry: Probing Physics Beyond the Standard Model". In: *Phys. Rept.* 117 (1985), pp. 75–263. DOI: 10.1016/0370-1573(85)90051-1.

[30] R. D. Peccei and Helen R. Quinn. "CP Conservation in the Presence of Instantons". In: *Phys. Rev. Lett.* 38 (1977), pp. 1440–1443. DOI: 10.1103/PhysRevLett.38.1440.

[31] Steven Weinberg. "A New Light Boson?" In: *Phys. Rev. Lett.* 40 (1978), pp. 223–226. DOI: 10.1103/PhysRevLett.40.223.

[32] Scott Dodelson and Lawrence M. Widrow. "Sterile-neutrinos as dark matter". In: *Phys. Rev. Lett.* 72 (1994), pp. 17–20. DOI: 10.1103/PhysRevLett.72.17. arXiv: hep-ph/9303287.

[33] Martin Bauer, Ulrich Haisch, and Felix Kahlhoefer. "Simplified dark matter models with two Higgs doublets: I. Pseudoscalar mediators". In: *JHEP* 05 (2017), p. 138. DOI: 10.1007/JHEP05(2017)138. arXiv: 1701.07427 [hep-ph].

[34] Jalal Abdallah et al. "Simplified Models for Dark Matter Searches at the LHC". In: *Phys. Dark Univ.* 9-10 (2015), pp. 8–23. DOI: 10.1016/j.dark.2015.08.001. arXiv: 1506.03116 [hep-ph].

[35] Andrea De Simone and Thomas Jacques. "Simplified models vs. effective field theory approaches in dark matter searches". In: *Eur. Phys. J. C* 76.7 (2016), p. 367. DOI: 10.1140/epjc/s10052-016-4208-4. arXiv: 1603.08002 [hep-ph].

[36] XENON Collaboration. "The XENON1T Dark Matter Experiment". In: *Eur. Phys. J. C* 77.12 (2017), p. 881. DOI: 10.1140/epjc/s10052-017-5326-3. arXiv: 1708.07051 [astro-ph.IM].

[37] XENON Collaboration. "Dark Matter Search Results from a One Ton-Year Exposure of XENON1T". In: *Phys. Rev. Lett.* 121.11 (2018), p. 111302. DOI: 10.1103/PhysRevLett.121.111302. arXiv: 1805.12562 [astro-ph.CO].

[38] XENON Collaboration. "The XENONnT dark matter experiment". In: *Eur. Phys. J. C* 84.8 (2024), p. 784. DOI: 10.1140/epjc/s10052-024-12982-5. arXiv: 2402.10446 [physics.ins-det].

[39] LZ Experiment. "LUX-ZEPLIN (LZ) Conceptual Design Report". In: *LBNL-190005, FERMILAB-TM-2621-AE-E-PPD* (Sept. 2015). arXiv: 1509.02910 [physics.ins-det].

[40] LZ Experiment. "The LUX-ZEPLIN (LZ) Experiment". In: *Nucl. Instrum. Meth. A* 953 (2020), p. 163047. DOI: 10.1016/j.nima.2019.163047.

[41] LZ Experiment. "First Dark Matter Search Results from the LUX-ZEPLIN (LZ) Experiment". In: *Phys. Rev. Lett.* 131.4 (2023), p. 041002. DOI: 10.1103/PhysRevLett.131.041002. arXiv: 2207.03764 [hep-ex].

[42] Particle Data Group. "Review of Particle Physics". In: *Phys. Rev. D* 110 (2024), p. 030001.

[43] M. Ackermann et al. "Searching for Dark Matter Annihilation from Milky Way Dwarf Spheroidal Galaxies with Six Years of Fermi Large Area Telescope Data". In: *Phys. Rev. Lett.* 115.23 (2015), p. 231301. DOI: 10.1103/PhysRevLett.115.231301. arXiv: 1503.02641 [astro-ph.HE].

[44] F. Aharonian et al. "The energy spectrum of cosmic-ray electrons at TeV energies". In: *Phys. Rev. Lett.* 101 (2008), p. 261104. DOI: 10.1103/PhysRevLett.101.261104. arXiv: 0811.3894 [astro-ph].

[45] M. G. Aartsen et al. "Search for dark matter annihilations in the Sun with the 79-string IceCube detector". In: *Phys. Rev. Lett.* 110.13 (2013), p. 131302. DOI: 10.1103/PhysRevLett.110.131302. arXiv: 1212.4097 [astro-ph.HE].

[46] Juliette Alimena et al. "Searching for long-lived particles beyond the Standard Model at the Large Hadron Collider". In: *J. Phys. G* 47.9 (2020), p. 090501. DOI: 10.1088/1361-6471/ab4574. arXiv: 1903.04497 [hep-ex].

[47] Guillaume Albouy et al. "Theory, phenomenology, and experimental avenues for dark showers: a Snowmass 2021 report". In: *Eur. Phys. J. C* 82.12 (2022), p. 1132. DOI: 10.1140/epjc/s10052-022-11048-8. arXiv: 2203.09503 [hep-ph].

[48] Deepak Kar and Sukanya Sinha. "Exploring jet substructure in semi-visible jets". In: *SciPost Phys.* 10.4 (2021), p. 084. DOI: 10.21468/SciPostPhys.10.4.084. arXiv: 2007.11597 [hep-ph].

[49] Lyndon Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[50] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

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

[52] LHCb Collaboration. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

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

[54] J. Vollaire et al. *Linac4 design report*. Ed. by Maurizio Vretenar. Vol. 6/2020. CERN Yellow Reports: Monographs. Geneva: CERN, Sept. 2020. ISBN: 978-92-9083-579-0, 978-92-9083-580-6. DOI: 10.23731/CYRM-2020-006.

[55] Jean-Paul Burnet et al. *Fifty years of the CERN Proton Synchrotron : Volume 2*. Ed. by Simone Gilardoni and Django Manglunki. CERN-2013-005. Aug. 2013. DOI: 10.5170/CERN-2013-005. arXiv: 1309.6923 [physics.acc-ph].

[56] J. B. ADAMS. "The Cern Proton Synchrotron". In: *Nature* 185.4713 (1960), pp. 568–572. DOI: 10.1038/185568a0. URL: https://doi.org/10.1038/185568a0.

[57] CERN Proton-Antiproton Project. "First Proton Anti-proton Collisions in the CERN SPS Collider". In: *Phys. Lett. B* 107 (1981), pp. 306–309. DOI: 10.1016/0370-2693(81)90836-4.

[58] Renilde Vanden Broeck. "The CERN Accelerator Complex. Complexe des accélérateurs du CERN". 2019. URL: https://cds.cern.ch/record/2693837.

[59] R Alemany-Fernandez et al. *Operation and Configuration of the LHC in Run 1*. 2013. URL: https://cds.cern.ch/record/1631030.

[60] G Apollinari et al. *High-Luminosity Large Hadron Collider (HL-LHC) : Preliminary Design Report*. Dec. 2015. ISBN: 9789290834236. DOI: 10.5170/CERN-2015-005.

[61] I. Zurbano Fernandez et al. *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*. Ed. by I. Béjar Alonso et al. Vol. 10/2020. Dec. 2020. ISBN: 978-92-9083-587-5. DOI: 10.23731/CYRM-2020-0010.

[62] I. Neutelings. *CMS coordinate system*. https://tikz.net/axis3d_cms/. Accessed: 18 August 2023.

[63] ATLAS Collaboration. "ATLAS pixel detector electronics and sensors". In: *JINST* 3 (2008), P07007. DOI: 10.1088/1748-0221/3/07/P07007. URL: https://cds.cern.ch/record/1119279.

[64] Nicoletta Garelli. "Performance of the ATLAS Detector in Run-2". In: *EPJ Web Conf.* 164 (2017), p. 01021. DOI: 10.1051/epjconf/201716401021. URL: https://cds.cern.ch/record/2310929.

[65] ATLAS IBL Collaboration. "Production and Integration of the ATLAS Insertable B-Layer". In: *JINST* 13.05 (2018), T05008. DOI: 10.1088/1748-0221/13/05/T05008. arXiv: 1803.00844 [physics.ins-det].

[66] ATLAS Collaboration. "Alignment of the ATLAS Inner Detector in Run-2". In: *Eur. Phys. J. C* 80.12 (2020), p. 1194. DOI: 10.1140/epjc/s10052-020-08700-6. arXiv: 2007.07624 [hep-ex].

[67] A Abdesselam and T Akimoto. *The Barrel Modules of the ATLAS SemiConductor Tracker*. Tech. rep. Geneva: CERN, 2006. DOI: 10.1016/j.nima.2006.08.036. URL: https://cds.cern.ch/record/974073.

[68] A Vogel. *ATLAS Transition Radiation Tracker (TRT): Straw Tube Gaseous Detectors at High Rates*. Tech. rep. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1537991.

[69] A. Buckley, C. White, and M. White. *Practical Collider Physics*. IOP, Dec. 2021. ISBN: 978-0-7503-2444-1, 978-0-7503-2443-4, 978-0-7503-2442-7. DOI: 10.1088/978-0-7503-2444-1.

[70] I. van Vulpen. *De Melodie van de Natuur*. Atlas Contact, 2018. ISBN: 978 90 450 3601.

[71] Konstantinos Nikolopoulos et al. "Cathode strip chambers in ATLAS : Installation, commissioning and in situ performance". In: *2008 IEEE Nuclear Science Symposium and Medical Imaging Conference and 16th International Workshop on Room-Temperature Semiconductor X-Ray and Gamma-Ray Detectors*. 2008, pp. 2819–2824. DOI: 10.1109/NSSMIC.2008.4774958.

[72] A. Yamamoto et al. "The ATLAS central solenoid". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 584.1 (2008), pp. 53–74. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2007.09.047. URL: https://www.sciencedirect.com/science/article/pii/S0168900207020414.

[73] Steffen Kaiser. "Search for the Higgs Boson in the Process $pp \to Hqq$, $H \to WW$ with the ATLAS Detector". Presented on 22 Feb 2010. 2010. URL: https://cds.cern.ch/record/1254897.

[74] ATLAS Collaboration. "ATLAS data quality operations and performance for 2015–2018 data-taking". In: *JINST* 15.04 (2020), P04003. DOI: 10.1088/1748-0221/15/04/P04003. arXiv: 1911.04632 [physics.ins-det].

[75] ATLAS Collaboration. "Standard Model Summary Plots February 2022". In: *ATL-PHYS-PUB-2022-009* (2022).

[76] CERN. *LHC Page1*. https://op-webtools.web.cern.ch/vistar/. [Online; accessed 16 November 2022].

[77] ATLAS Collaboration. "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 83.10 (2023), p. 982. DOI: 10.1140/epjc/s10052-023-11747-w. arXiv: 2212.09379 [hep-ex].

[78] S van der Meer. *Calibration of the effective beam height in the ISR*. Tech. rep. Geneva: CERN, 1968. URL: https://cds.cern.ch/record/296752.

[79] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017.

[80] ATLAS Collaboration. *Public ATLAS Luminosity Results for Run-2 of the LHC*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. Accessed: 9 April 2024.

[81] Andy Buckley et al. "General-purpose event generators for LHC physics". In: *Phys. Rept.* 504 (2011), pp. 145–233. DOI: 10.1016/j.physrep.2011.03.005. arXiv: 1101.2599 [hep-ph].

[82] Johannes Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J. C* 76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].

[83] Johannes Bellm et al. "Herwig 7.2 release note". In: *Eur. Phys. J. C* 80.5 (2020), p. 452. DOI: 10.1140/epjc/s10052-020-8011-x. arXiv: 1912.06509 [hep-ph].

[84] Gavin Bewick et al. *Herwig 7.3 Release Note*. arXiv: 2312.05175 [hep-ph].

[85] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[86] Enrico Bothmann et al. "Event Generation with Sherpa 2.2". In: *SciPost Phys.* 7.3 (2019), p. 034. DOI: 10.21468/SciPostPhys.7.3.034. arXiv: 1905.09127 [hep-ph].

[87] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[88] Torbjörn Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[89] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[90] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[91] Carlo Oleari. "The POWHEG-BOX". In: *Nucl. Phys. B Proc. Suppl.* 205-206 (2010). Ed. by Johannes Blümlein, Sven-Olaf Moch, and Tord Riemann, pp. 36–41. DOI: 10.1016/j.nuclphysbps.2010.08.016. arXiv: 1007.3893 [hep-ph].

[92] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].

[93] Johan Alwall et al. "A Standard format for Les Houches event files". In: *Comput. Phys. Commun.* 176 (2007), pp. 300–304. DOI: 10.1016/j.cpc.2006.11.010. arXiv: hep-ph/0609017.

[94] Andy Buckley et al. "The HepMC3 event record library for Monte Carlo event generators". In: *Comput. Phys. Commun.* 260 (2021), p. 107310. DOI: 10.1016/j.cpc.2020.107310. arXiv: 1912.08005 [hep-ph].

[95] J. M. Campbell et al. "Event Generators for High-Energy Physics Experiments". In: *Snowmass 2021*. Mar. 2022. arXiv: 2203.11110 [hep-ph].

[96] Christian Bierlich et al. "A comprehensive guide to the physics and usage of PYTHIA 8.3". In: *SciPost Phys. Codeb.* 2022 (2022), p. 8. DOI: 10.21468/SciPostPhysCodeb.8. arXiv: 2203.11601 [hep-ph].

[97] J. D. Bjorken and E. A. Paschos. "Inelastic Electron Proton and gamma Proton Scattering, and the Structure of the Nucleon". In: *Phys. Rev.* 185 (1969), pp. 1975–1982. DOI: 10.1103/PhysRev.185.1975.

[98] E. A. Paschos. "Theoretical Ideas on Inelastic Electron–Nucleon Scattering". In: *Proceedings of Boulder Conference on High Energy Physics. /Mahanthappa, K. T. (ed.). Boulder, Colo.*

Colorado Associated University Press (Jan. 1970), pp. 103–34. URL: https://www.osti.gov/biblio/4177049.

[99] Sidney D. Drell, Donald J. Levy, and Tung-Mow Yan. "A Field-Theoretic Model for Electron-Nucleon Deep Inelastic Scattering". In: *Phys. Rev. Lett.* 22 (14 Apr. 1969), pp. 744–749. DOI: 10.1103/PhysRevLett.22.744. URL: https://link.aps.org/doi/10.1103/PhysRevLett.22.744.

[100] S. Bailey et al. "Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs". In: *Eur. Phys. J. C* 81.4 (2021), p. 341. DOI: 10.1140/epjc/s10052-021-09057-0. arXiv: 2012.04684 [hep-ph].

[101] T. Cridge et al. "An investigation of the α_S and heavy quark mass dependence in the MSHT20 global PDF analysis". In: *Eur. Phys. J. C* 81.8 (2021), p. 744. DOI: 10.1140/epjc/s10052-021-09533-7. arXiv: 2106.10289 [hep-ph].

[102] T. Cridge et al. "QED parton distribution functions in the MSHT20 fit". In: *Eur. Phys. J. C* 82.1 (2022), p. 90. DOI: 10.1140/epjc/s10052-022-10028-2. arXiv: 2111.05357 [hep-ph].

[103] J. McGowan et al. "Approximate N^{3}LO parton distribution functions with theoretical uncertainties: MSHT20aN^{3}LO PDFs". In: *Eur. Phys. J. C* 83.3 (2023). [Erratum: Eur.Phys.J.C 83, 302 (2023)], p. 185. DOI: 10.1140/epjc/s10052-023-11236-0. arXiv: 2207.04739 [hep-ph].

[104] H. L. Lai et al. "Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions". In: *Eur. Phys. J. C* 12 (2000), pp. 375–392. DOI: 10.1007/s100529900196. arXiv: hep-ph/9903282.

[105] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 07 (2002), p. 012. DOI: 10.1088/1126-6708/2002/07/012. arXiv: hep-ph/0201195.

[106] Hung-Liang Lai et al. "New parton distributions for collider physics". In: *Phys. Rev. D* 82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

[107] Jun Gao et al. "CT10 next-to-next-to-leading order global analysis of QCD". In: *Phys. Rev. D* 89.3 (2014), p. 033009. DOI: 10.1103/PhysRevD.89.033009. arXiv: 1302.6246 [hep-ph].

[108] Sayipjamal Dulat et al. "New parton distribution functions from a global analysis of quantum chromodynamics". In: *Phys. Rev. D* 93.3 (2016), p. 033006. DOI: 10.1103/PhysRevD.93.033006. arXiv: 1506.07443 [hep-ph].

[109] Luigi Del Debbio et al. "Unbiased determination of the proton structure function F(2)**p with faithful uncertainty estimation". In: *JHEP* 03 (2005), p. 080. DOI: 10.1088/1126-6708/2005/03/080. arXiv: hep-ph/0501067.

[110] Richard D. Ball et al. "A Determination of parton distributions with faithful uncertainty estimation". In: *Nucl. Phys. B* 809 (2009). [Erratum: Nucl.Phys.B 816, 293 (2009)], pp. 1–63. DOI: 10.1016/j.nuclphysb.2008.09.037. arXiv: 0808.1231 [hep-ph].

[111] Richard D. Ball et al. "Parton distributions for the LHC Run II". In: *JHEP* 04 (2015), p. 040. DOI: 10.1007/JHEP04(2015)040. arXiv: 1410.8849 [hep-ph].

[112] Richard D. Ball et al. "Parton distributions from high-precision collider data". In: *Eur. Phys. J. C* 77.10 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5. arXiv: 1706.00428 [hep-ph].

[113] ATLAS Collaboration. "Measurement of hard double-parton interactions in $W(\rightarrow l\nu)$+ 2 jet events at $\sqrt{s}$=7 TeV with the ATLAS detector". In: *New J. Phys.* 15 (2013), p. 033038. DOI: 10.1088/1367-2630/15/3/033038. arXiv: 1301.6872 [hep-ex].

[114] Torbjörn Sjöstrand and Maria van Zijl. "A multiple-interaction model for the event structure in hadron collisions". In: *Phys. Rev. D* 36 (7 Oct. 1987), pp. 2019–2041. DOI: 10.1103/PhysRevD.36.2019. URL: https://link.aps.org/doi/10.1103/PhysRevD.36.2019.

[115] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175.

[116] ATLAS Collaboration. *ATLAS Software and Computing HL-LHC Roadmap*. Tech. rep. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2802918.

[117] Enrico Bothmann et al. "Accelerating LHC event generation with simplified pilot runs and fast PDFs". In: *Eur. Phys. J. C* 82.12 (2022), p. 1128. DOI: 10.1140/epjc/s10052-022-11087-1. arXiv: 2209.00843 [hep-ph].

[118] Enrico Bothmann et al. "QCD tree amplitudes on modern GPUs: A case study for novel event generators". In: *PoS* ICHEP2022 (Nov. 2022), p. 222. DOI: 10.22323/1.414.0222.

[119] Stefan Gieseke, P. Stephens, and Bryan Webber. "New formalism for QCD parton showers". In: *JHEP* 12 (2003), p. 045. DOI: 10.1088/1126-6708/2003/12/045. arXiv: hep-ph/0310083.

[120] Jan-Christopher Winter and Frank Krauss. "Initial-state showering based on colour dipoles connected to incoming parton lines". In: *JHEP* 07 (2008), p. 040. DOI: 10.1088/1126-6708/2008/07/040. arXiv: 0712.3913 [hep-ph].

[121] Steffen Schumann and Frank Krauss. "A Parton shower algorithm based on Catani-Seymour dipole factorisation". In: *JHEP* 03 (2008), p. 038. DOI: 10.1088/1126-6708/2008/03/038. arXiv: 0709.1027 [hep-ph].

[122] Simon Platzer and Stefan Gieseke. "Coherent Parton Showers with Local Recoils". In: *JHEP* 01 (2011), p. 024. DOI: 10.1007/JHEP01(2011)024. arXiv: 0909.5593 [hep-ph].

[123] Torbjorn Sjostrand. "A Model for Initial State Parton Showers". In: *Phys. Lett. B* 157 (1985), pp. 321–325. DOI: 10.1016/0370-2693(85)90674-4.

[124] Stefano Frixione and Bryan R. Webber. "Matching NLO QCD computations and parton shower simulations". In: *JHEP* 06 (2002), p. 029. DOI: 10.1088/1126-6708/2002/06/029. arXiv: hep-ph/0204244.

[125] Pier Francesco Monni et al. "MiNNLO$_{PS}$: a new method to match NNLO QCD to parton showers". In: *JHEP* 05 (2020). [Erratum: JHEP 02, 031 (2022)], p. 143. DOI: 10.1007/JHEP05(2020)143. arXiv: 1908.06987 [hep-ph].

[126] Keith Hamilton et al. "Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching". In: *JHEP* 05 (2013), p. 082. DOI: 10.1007/JHEP05(2013)082. arXiv: 1212.4504 [hep-ph].

[127] Simone Alioli et al. "Matching NNLO predictions to parton showers using N3LL color-singlet transverse momentum resummation in geneva". In: *Phys. Rev. D* 104.9 (2021), p. 094020. DOI: 10.1103/PhysRevD.104.094020. arXiv: 2102.08390 [hep-ph].

[128] S. Catani et al. "QCD matrix elements + parton showers". In: *JHEP* 11 (2001), p. 063. DOI: 10.1088/1126-6708/2001/11/063. arXiv: hep-ph/0109231.

[129] F. Krauss. "Matrix elements and parton showers in hadronic interactions". In: *JHEP* 08 (2002), p. 015. DOI: 10.1088/1126-6708/2002/08/015. arXiv: hep-ph/0205283.

[130] Johan Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". In: *Eur. Phys. J. C* 53 (2008), pp. 473–500. DOI: 10.1140/epjc/s10052-007-0490-5. arXiv: 0706.2569 [hep-ph].

[131] Leif Lonnblad. "Correcting the color dipole cascade model with fixed order matrix elements". In: *JHEP* 05 (2002), p. 046. DOI: 10.1088/1126-6708/2002/05/046. arXiv: hep-ph/0112284.

[132] Stefan Hoeche et al. "QCD matrix elements + parton showers: The NLO case". In: *JHEP* 04 (2013), p. 027. DOI: 10.1007/JHEP04(2013)027. arXiv: 1207.5030 [hep-ph].

[133] Bo Andersson et al. "Parton Fragmentation and String Dynamics". In: *Phys. Rept.* 97 (1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[134] Jan-Christopher Winter, Frank Krauss, and Gerhard Soff. "A Modified cluster hadronization model". In: *Eur. Phys. J. C* 36 (2004), pp. 381–395. DOI: 10.1140/epjc/s2004-01960-8. arXiv: hep-ph/0311085.

[135] D. J. Lange. "The EvtGen particle decay simulation package". In: *Nucl. Instrum. Meth. A* 462 (2001). Ed. by S. Erhan, P. Schlein, and Y. Rozen, pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.

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

[137] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C* 70 (2010), pp. 823–874. DOI: 10.1140/epjc/s10052-010-1429-9.

[138] J. Allison et al. "Recent developments in Geant4". In: *Nucl. Instrum. Meth. A* 835 (2016), pp. 186–225. DOI: 10.1016/j.nima.2016.06.125.

[139] ATLAS Collaboration. "AtlFast3: The Next Generation of Fast Simulation in ATLAS". In: *Comput. Softw. Big Sci.* 6.1 (2022), p. 7. DOI: 10.1007/s41781-021-00079-7. arXiv: 2109.02551 [hep-ex].

[140] ATLAS Collaboration. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". In: *JINST* 14.12 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[141] ATLAS Collaboration. "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016". In: *Eur. Phys. J. C* 79.3 (2019), p. 205. DOI: 10.1140/epjc/s10052-019-6650-6. arXiv: 1810.05087 [hep-ex].

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

[143] T G Cornelissen et al. "The global χ^2 track fitter in ATLAS". In: *Journal of Physics: Conference Series* 119.3 (July 2008), p. 032013. DOI: 10.1088/1742-6596/119/3/032013. URL: https://dx.doi.org/10.1088/1742-6596/119/3/032013.

[144] ATLAS Collaboration. *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. Geneva: CERN, 2012. URL: https://cds.cern.ch/record/1449796.

[145] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *JINST* 14.03 (2019), P03017. DOI: 10.1088/1748-0221/14/03/P03017. arXiv: 1812.03848 [hep-ex].

[146] ATLAS Collaboration. "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 81.7 (2021), p. 578. DOI: 10.1140/epjc/s10052-021-09233-2. arXiv: 2012.00578 [hep-ex].

[147] ATLAS Collaboration. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *Eur. Phys. J. C* 77.7 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[148] Ahmed Ali and Gustav Kramer. "Jets and QCD: A Historical Review of the Discovery of the Quark and Gluon Jets and its Impact on QCD". In: *Eur. Phys. J. H* 36 (2011), pp. 245–326. DOI: 10.1140/epjh/e2011-10047-1. arXiv: 1012.2288 [hep-ph].

[149] H. J. Daum, H. Meyer, and J. Burger. "A Cluster Algorithm for Jet Studies". In: *Z. Phys. C* 8 (1981), p. 167. DOI: 10.1007/BF01547881.

[150] Stephen D. Ellis and Davison E. Soper. "Successive combination jet algorithm for hadron collisions". In: *Phys. Rev. D* 48 (1993), pp. 3160–3166. DOI: 10.1103/PhysRevD.48.3160. arXiv: hep-ph/9305266.

[151] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[152] Yuri L. Dokshitzer et al. "Better jet clustering algorithms". In: *JHEP* 08 (1997), p. 001. DOI: 10.1088/1126-6708/1997/08/001. arXiv: hep-ph/9707323.

[153] Gavin P. Salam and Gregory Soyez. "A Practical Seedless Infrared-Safe Cone jet algorithm". In: *JHEP* 05 (2007), p. 086. DOI: 10.1088/1126-6708/2007/05/086. arXiv: 0704.0292 [hep-ph].

[154] Ryan Atkin. "Review of jet reconstruction algorithms". In: *J. Phys. Conf. Ser.* 645.1 (2015). Ed. by Alan S. Cornell and Bruce Mellado, p. 012008. DOI: 10.1088/1742-6596/645/1/012008.

[155] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet User Manual". In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[156] Jonathan M. Butterworth et al. "Jet substructure as a new Higgs search channel at the LHC". In: *Phys. Rev. Lett.* 100 (2008), p. 242001. DOI: 10.1103/PhysRevLett.100.242001. arXiv: 0802.2470 [hep-ph].

[157] Roman Kogler et al. "Jet Substructure at the Large Hadron Collider: Experimental Review". In: *Rev. Mod. Phys.* 91.4 (2019), p. 045003. DOI: 10.1103/RevModPhys.91.045003. arXiv: 1803.06991 [hep-ex].

[158] ATLAS Collaboration. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77 (2017), p. 490. DOI: 10.1140/epjc/s10052-017-5004-5. arXiv: 1603.02934 [hep-ex].

[159] ATLAS Collaboration. "Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 81.8 (2021), p. 689. DOI: 10.1140/epjc/s10052-021-09402-3. arXiv: 2007.02645 [hep-ex].

[160] ATLAS Collaboration. *Tagging and suppression of pileup jets with the ATLAS detector*. Tech. rep. Geneva: CERN, 2014. URL: https://cds.cern.ch/record/1700870.

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

[162] ATLAS Collaboration. *Forward jet vertex tagging using the particle flow algorithm*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2683100.

[163] ATLAS Collaboration. "ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset". In: *Eur. Phys. J. C* 83.7 (2023), p. 681. DOI: 10.1140/epjc/s10052-023-11699-1. arXiv: 2211.16345 [physics.data-an].

[164] ATLAS Collaboration. *Graph Neural Network Jet Flavour Tagging with the ATLAS Detector*. ATL-PHYS-PUB-2022-027. 2022.

[165] Particle Data Group. "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[166] ATLAS Collaboration. *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2688062.

[167] ATLAS Collaboration. "Reconstruction, Identification, and Calibration of hadronically decaying tau leptons with the ATLAS detector for the LHC Run 3 and reprocessed Run 2 data". In: *ATL-PHYS-PUB-2022-044* (2022).

[168] ATLAS Collaboration. "Observation of four-top-quark production in the multilepton final state with the ATLAS detector". In: *Eur. Phys. J. C* 83.6 (2023). [Erratum: Eur.Phys.J.C 84, 156 (2024)], p. 496. DOI: 10.1140/epjc/s10052-023-11573-0. arXiv: 2303.15061 [hep-ex].

[169] ATLAS Collaboration. "Improved W boson Mass Measurement using 7 TeV Proton-Proton Collisions with the ATLAS Detector". In: *ATLAS-CONF-2023-004* (2023).

[170] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[171] CMS Collaboration. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[172] ATLAS Collaboration. "Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector". In: *ATLAS-CONF-2015-029* (July 2015).

[173] Tancredi Carli. *Jet cleaning in full run-2 data*. Tech. rep. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2846229.

[174] Waleed Abdallah et al. "Reinterpretation of LHC Results for New Physics: Status and Recommendations after Run 2". In: *SciPost Phys.* 9.2 (2020), p. 022. DOI: 10.21468/SciPostPhys.9.2.022. arXiv: 2003.07868 [hep-ph].

[175] G. Cowan. *Statistical data analysis*. Oxford University Press, USA, 1998. ISBN: 9780198501558.

[176] G. Cowan. "A survey of unfolding methods for particle physics". In: *Conf. Proc. C* 0203181 (2002). Ed. by M. R. Whalley and L. Lyons, pp. 248–257.

[177] Volker Blobel. "Unfolding Methods in Particle Physics". In: *PHYSTAT 2011*. Geneva: CERN, 2011, pp. 240–251. DOI: 10.5170/CERN-2011-006.240.

[178] Particle Data Group. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[179] G. D'Agostini. "Improved iterative Bayesian unfolding". In: Oct. 2010. arXiv: 1010.0632 [physics.data-an].

[180] Andreas Höcker and Vakhtang Kartvelishvili. "SVD approach to data unfolding". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 372.3 (1996), pp. 469–481. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(95)01478-0. URL: https://www.sciencedirect.com/science/article/pii/0168900295014780.

[181] Tim Adye. "Unfolding algorithms and tests using RooUnfold". In: *PHYSTAT 2011*. Geneva: CERN, 2011, pp. 313–318. DOI: 10.5170/CERN-2011-006.313. arXiv: 1105.1160 [physics.data-an].

[182] ATLAS Collaboration. "Measurement of WZ production in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 72 (2012), p. 2173. DOI: 10.1140/epjc/s10052-012-2173-0. arXiv: 1208.1390 [hep-ex].

[183] ATLAS Collaboration. "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector". In: *JHEP* 09 (2014), p. 145. DOI: 10.1007/JHEP09(2014)145. arXiv: 1406.3660 [hep-ex].

[184] ATLAS Collaboration. "Precision measurement and interpretation of inclusive W^+ , W^- and Z/γ^* production cross sections with the ATLAS detector". In: *Eur. Phys. J. C* 77.6 (2017), p. 367. DOI: 10.1140/epjc/s10052-017-4911-9. arXiv: 1612.03016 [hep-ex].

[185] ALICE Collaboration. "First measurement of jet mass in Pb–Pb and p–Pb collisions at the LHC". In: *Phys. Lett. B* 776 (2018), pp. 249–264. DOI: 10.1016/j.physletb.2017.11.044. arXiv: 1702.00804 [nucl-ex].

[186] E. Richard et al. "Measurements of the atmospheric neutrino flux by Super-Kamiokande: energy spectra, geomagnetic effects, and solar modulation". In: *Phys. Rev. D* 94.5 (2016), p. 052001. DOI: 10.1103/PhysRevD.94.052001. arXiv: 1510.08127 [hep-ex].

[187] ATLAS Collaboration. "Measurement of detector-corrected observables sensitive to the anomalous production of events with jets and large missing transverse momentum in pp collisions at $\sqrt{\mathbf{s}} = \mathbf{13}$ TeV using the ATLAS detector". In: *Eur. Phys. J. C* 77.11 (2017), p. 765. DOI: 10.1140/epjc/s10052-017-5315-6. arXiv: 1707.03263 [hep-ex].

[188] ATLAS Collaboration. "ATLAS Event display: Highest-Momentum Monojet (1.9 TeV)". 2020. URL: https://cds.cern.ch/record/2725235.

[189] ATLAS Collaboration. "Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment". In: *JHEP* 05 (2024), p. 162. DOI: 10.1007/JHEP05(2024)162. arXiv: 2308.13362 [hep-ex].

[190] ATLAS Collaboration. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J. C* 76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[191] ATLAS Collaboration. "Search for new phenomena in events with an energetic jet and missing transverse momentum in pp collisions at $\sqrt{s}$ =13 TeV with the ATLAS detector". In: *Phys. Rev. D* 103.11 (2021), p. 112006. DOI: 10.1103/PhysRevD.103.112006. arXiv: 2102.10874 [hep-ex].

[192] J. D. Bjorken. "Rapidity gaps and jets as a new-physics signature in very-high-energy hadron-hadron collisions". In: *Phys. Rev. D* 47 (1 Jan. 1993), pp. 101–113. DOI: 10.1103/

PhysRevD.47.101. URL: https://link.aps.org/doi/10.1103/PhysRevD.47.101.

[193] D. Zeppenfeld et al. "Measuring Higgs boson couplings at the CERN LHC". In: *Phys. Rev. D* 62 (2000), p. 013009. DOI: 10.1103/PhysRevD.62.013009. arXiv: hep-ph/0002036.

[194] ATLAS Collaboration. "Performance of the missing transverse momentum triggers for the ATLAS detector during Run-2 data taking". In: *JHEP* 08 (2020), p. 080. DOI: 10.1007/JHEP08(2020)080. arXiv: 2005.09554 [hep-ex].

[195] ATLAS Collaboration. "Performance of electron and photon triggers in ATLAS during LHC Run 2". In: *Eur. Phys. J. C* 80.1 (2020), p. 47. DOI: 10.1140/epjc/s10052-019-7500-2. arXiv: 1909.00761 [hep-ex].

[196] ATLAS Collaboration. "Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: *Phys. Rev. D* 94.3 (2016), p. 032005. DOI: 10.1103/PhysRevD.94.032005. arXiv: 1604.07773 [hep-ex].

[197] ATLAS Collaboration. "Constraints on mediator-based dark matter and scalar dark energy models using $\sqrt{s} = 13$ TeV pp collision data collected by the ATLAS detector". In: *JHEP* 05 (2019), p. 142. DOI: 10.1007/JHEP05(2019)142. arXiv: 1903.01400 [hep-ex].

[198] Ken Mimasu and Verónica Sanz. "ALPs at Colliders". In: *JHEP* 06 (2015), p. 173. DOI: 10.1007/JHEP06(2015)173. arXiv: 1409.4792 [hep-ph].

[199] ATLAS Collaboration. "Search for new phenomena in final states with an energetic jet and large missing transverse momentum in pp collisions at $\sqrt{s}$ =8 TeV with the ATLAS detector". In: *Eur. Phys. J. C* 75.7 (2015). [Erratum: Eur.Phys.J.C 75, 408 (2015)], p. 299. DOI: 10.1140/epjc/s10052-015-3517-3. arXiv: 1502.01518 [hep-ex].

[200] Robert M. Harris and Konstantinos Kousouris. "Searches for Dijet Resonances at Hadron Colliders". In: *Int. J. Mod. Phys. A* 26 (2011), pp. 5005–5055. DOI: 10.1142/S0217751X11054905. arXiv: 1110.5302 [hep-ex].

[201] ATLAS Collaboration. "Search for new phenomena in the dijet mass distribution using $p - p$ collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Phys. Rev. D* 91.5 (2015), p. 052007. DOI: 10.1103/PhysRevD.91.052007. arXiv: 1407.1376 [hep-ex].

[202] ATLAS Collaboration. "Search for new phenomena in dijet mass and angular distributions from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 754 (2016), pp. 302–322. DOI: 10.1016/j.physletb.2016.01.032. arXiv: 1512.01530 [hep-ex].

[203] 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". In: *Phys. Rev. D* 96.5 (2017), p. 052004. DOI: 10.1103/PhysRevD.96.052004. arXiv: 1703.09127 [hep-ex].

[204] ATLAS Collaboration. "Search for low-mass dijet resonances using trigger-level jets with the ATLAS detector in pp collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. Lett.* 121.8 (2018), p. 081801. DOI: 10.1103/PhysRevLett.121.081801. arXiv: 1804.03496 [hep-ex].

[205] Florian U. Bernlochner et al. "Angles on CP-violation in Higgs boson interactions". In: *Phys. Lett. B* 790 (2019), pp. 372–379. DOI: 10.1016/j.physletb.2019.01.043. arXiv: 1808.06577 [hep-ph].

[206] Tilman Plehn, David L. Rainwater, and Dieter Zeppenfeld. "Determining the Structure of Higgs Couplings at the LHC". In: *Phys. Rev. Lett.* 88 (2002), p. 051801. DOI: 10.1103/PhysRevLett.88.051801. arXiv: hep-ph/0105325.

[207] John Allison et al. "Geant4 developments and applications". In: *IEEE Trans. Nucl. Sci.* 53 (2006), p. 270. DOI: 10.1109/TNS.2006.869826.

[208] ATLAS Exotics Working Group. *EXOT derivations*. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/DerivationframeworkExotics.

[209] Louie Corpe. *Particle Physics for Babies*. ATLAS Experiment and International Particle Physics Outreach Group. ISBN: 9781399921084.

[210] Tanju Gleisberg and Stefan Hoeche. "Comix, a new matrix element generator". In: *JHEP* 12 (2008), p. 039. DOI: 10.1088/1126-6708/2008/12/039. arXiv: 0808.3674 [hep-ph].

[211] Fabio Cascioli, Philipp Maierhofer, and Stefano Pozzorini. "Scattering Amplitudes with Open Loops". In: *Phys. Rev. Lett.* 108 (2012), p. 111601. DOI: 10.1103/PhysRevLett.108.111601. arXiv: 1111.5206 [hep-ph].

[212] Charalampos Anastasiou et al. "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO". In: *Phys. Rev. D* 69 (2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266.

[213] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction". In: *JHEP* 09 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph].

[214] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[215] Stefano Frixione et al. "Single-top hadroproduction in association with a W boson". In: *JHEP* 07 (2008), p. 029. DOI: 10.1088/1126-6708/2008/07/029. arXiv: 0805.3067 [hep-ph].

[216] ATLAS Collaboration. "ATLAS Pythia 8 tunes to 7 TeV data". In: *ATL-PHYS-PUB-2014-021* (Nov. 2014).

[217] D0 Collaboration. "Studies of W Boson Plus Jets Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV". In: *Phys. Rev. D* 88.9 (2013), p. 092001. DOI: 10.1103/PhysRevD.88.092001. arXiv: 1302.6508 [hep-ex].

[218] ATLAS Collaboration. "Measurement of the electroweak production of dijets in association with a Z-boson and distributions sensitive to vector boson fusion in proton-proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector". In: *JHEP* 04 (2014), p. 031. DOI: 10.1007/JHEP04(2014)031. arXiv: 1401.7610 [hep-ex].

[219] ATLAS Collaboration. "Measurements of electroweak Wjj production and constraints on anomalous gauge couplings with the ATLAS detector". In: *Eur. Phys. J. C* 77.7 (2017), p. 474. DOI: 10.1140/epjc/s10052-017-5007-2. arXiv: 1703.04362 [hep-ex].

[220] ATLAS Collaboration. "Measurement of the cross-section for electroweak production of dijets in association with a Z boson in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector". In: *Phys. Lett. B* 775 (2017), pp. 206–228. DOI: 10.1016/j.physletb.2017.10.040. arXiv: 1709.10264 [hep-ex].

[221] Federico Buccioni et al. "OpenLoops 2". In: *Eur. Phys. J. C* 79.10 (2019), p. 866. DOI: 10.1140/epjc/s10052-019-7306-2. arXiv: 1907.13071 [hep-ph].

[222] F. Buccioni, S. Pozzorini, and M. Zoller. "On-the-fly reduction of open loops". In: *Eur. Phys. J. C* 78.1 (2018), p. 70. DOI: 10.1140/epjc/s10052-018-5562-1. arXiv: 1710.11452 [hep-ph].

[223] Ansgar Denner et al. "Electroweak corrections to W + jet hadroproduction including leptonic W-boson decays". In: *JHEP* 08 (2009), p. 075. DOI: 10.1088/1126-6708/2009/08/075. arXiv: 0906.1656 [hep-ph].

[224] Stefan Kallweit et al. "NLO QCD+EW predictions for V + jets including off-shell vector-boson decays and multijet merging". In: *JHEP* 04 (2016), p. 021. DOI: 10.1007/JHEP04(2016)021. arXiv: 1511.08692 [hep-ph].

[225] J. M. Lindert et al. "Precise predictions for $V+$ jets dark matter backgrounds". In: *Eur. Phys. J. C* 77.12 (2017), p. 829. DOI: 10.1140/epjc/s10052-017-5389-1. arXiv: 1705.04664 [hep-ph].

[226] Jon Butterworth et al. "PDF4LHC recommendations for LHC Run II". In: *J. Phys. G* 43 (2016), p. 023001. DOI: 10.1088/0954-3899/43/2/023001. arXiv: 1510.03865 [hep-ph].

[227] Valerio Bertone et al. "Illuminating the photon content of the proton within a global PDF analysis". In: *SciPost Phys.* 5.1 (2018), p. 008. DOI: 10.21468/SciPostPhys.5.1.008. arXiv: 1712.07053 [hep-ph].

[228] Jeppe R. Andersen and Jennifer M. Smillie. "Constructing All-Order Corrections to Multi-Jet Rates". In: *JHEP* 01 (2010), p. 039. DOI: 10.1007/JHEP01(2010)039. arXiv: 0908.2786 [hep-ph].

[229] Jeppe R. Andersen and Jennifer M. Smillie. "High Energy Description of Processes with Multiple Hard Jets". In: *Nucl. Phys. B Proc. Suppl.* 205-206 (2010). Ed. by Johannes Blümlein, Sven-Olaf Moch, and Tord Riemann, pp. 205–210. DOI: 10.1016/j.nuclphysbps.2010.08.044. arXiv: 1007.4449 [hep-ph].

[230] Jeppe R. Andersen et al. *HEJ 2.2: W boson pairs and Higgs boson plus jet production at high energies*. Mar. 2023. DOI: 10.21468/SciPostPhysCodeb.21. arXiv: 2303.15778 [hep-ph].

[231] ATLAS Collaboration. "Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using 4.7 fb^{-1} of $\sqrt{s} = 7$ TeV proton-proton collision data". In: *Phys. Rev. D* 87.1 (2013), p. 012008. DOI: 10.1103/PhysRevD.87.012008. arXiv: 1208.0949 [hep-ex].

[232] ATLAS Collaboration. "Differential cross-sections for events with missing transverse momentum and jets measured with the ATLAS detector in 13 TeV proton-proton collisions". In: *JHEP* 08 (2024), p. 223. DOI: 10.1007/JHEP08(2024)223. arXiv: 2403.02793 [hep-ex].

[233] ATLAS Collaboration. "Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC". In: *JINST* 18.11 (2023), T11004. DOI: 10.1088/1748-0221/18/11/T11004. arXiv: 2211.16178 [hep-ex].

[234] B. Efron. "Bootstrap Methods: Another Look at the Jackknife". In: *The Annals of Statistics* 7.1 (1979), pp. 1–26. DOI: 10.1214/aos/1176344552. URL: https://doi.org/10.1214/aos/1176344552.

[235] ATLAS Collaboration. *Evaluating statistical uncertainties and correlations using the bootstrap method*. ATL-PHYS-PUB-2021-011, URL: https://cds.cern.ch/record/2759945/. 2021. URL: https://cds.cern.ch/record/2759945/.

[236] Eamonn Maguire, Lukas Heinrich, and Graeme Watt. "HEPData: a repository for high energy physics data". In: *J. Phys. Conf. Ser.* 898.10 (2017). Ed. by Richard Mount and Craig Tull, p. 102006. DOI: 10.1088/1742-6596/898/10/102006. arXiv: 1704.05473 [hep-ex].

[237] ATLAS Collaboration. *Differential cross-sections for events with missing transverse momentum and jets measured with the ATLAS detector in 13 TeV proton-proton collisions*. HEPData (collection). https://doi.org/10.17182/hepdata.149990. 2024.

[238] ATLAS Collaboration. "Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13\ TeV$". In: *ATLAS-CONF-2017-029* (May 2017).

[239] ATLAS Collaboration. "Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *Eur. Phys. J. C* 78.11 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[240] V. Lang and S. Richter. *VIPUnfolding*. Software and auxiliary material available at https://gitlab.cern.ch/atlas-physics/sm/StandardModelTools_Unfolding/VIPUnfolding/. 2021.

[241] A. Buckley et al. "Testing new physics models with global comparisons to collider measurements: the Contur toolkit". In: *SciPost Phys. Core* 4 (2021), p. 013. DOI: 10.21468/SciPostPhysCore.4.2.013. arXiv: 2102.04377 [hep-ph].

[242] G. L. Fogli et al. "Getting the most from the statistical analysis of solar neutrino oscillations". In: *Phys. Rev. D* 66 (2002), p. 053010. DOI: 10.1103/PhysRevD.66.053010. arXiv: hep-ph/0206162.

[243] Amoroso, Simone and Gütschow, Christian and Hawkings, Richard and Marshall, Zach and Malaescu, Bogdan and Verkerke, Wouter. *Recommendations on the treatment of theoretical systematic uncertainties in statistical analysis of ATLAS data*. Tech. rep. Geneva: CERN, 2020. URL: https://cds.cern.ch/record/2715689.

[244] CMS Collaboration. "Search for new particles in events with energetic jets and large missing transverse momentum in proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *JHEP* 11 (2021), p. 153. DOI: 10.1007/JHEP11(2021)153. arXiv: 2107.13021 [hep-ex].

[245] Tomohiro Abe et al. "LHC Dark Matter Working Group: Next-generation spin-0 dark matter models". In: *Phys. Dark Univ.* 27 (2020), p. 100351. DOI: 10.1016/j.dark.2019.100351. arXiv: 1810.09420 [hep-ex].

[246] ATLAS Collaboration. "Combination and summary of ATLAS dark matter searches interpreted in a 2HDM with a pseudo-scalar mediator using 139 fb^{-1} of $\sqrt{s} = 13$ TeV pp collision data". In: *Sci. Bull.* 69 (2024), p. 3005. DOI: 10.1016/j.scib.2024.06.003. arXiv: 2306.00641 [hep-ex].

[247] Siegfried Bethke. "Data Preservation in High Energy Physics - why, how and when?" In: *Nucl. Phys. B Proc. Suppl.* 207-208 (2010). Ed. by Stephan Narison, pp. 156–159. DOI: 10.1016/j.nuclphysbps.2010.10.040. arXiv: 1009.3763 [hep-ex].

[248] S. Kraml et al. "Searches for New Physics: Les Houches Recommendations for the Presentation of LHC Results". In: *Eur. Phys. J. C* 72 (2012), p. 1976. DOI: 10.1140/epjc/s10052-012-1976-3. arXiv: 1203.2489 [hep-ph].

[249] Zaven Akopov et al. *Status Report of the DPHEP Study Group: Towards a Global Effort for Sustainable Data Preservation in High Energy Physics*. arXiv: 1205.4667 [hep-ex].

[250] Kyle Cranmer et al. "Publishing statistical models: Getting the most out of particle physics experiments". In: *SciPost Phys.* 12.1 (2022), p. 037. DOI: 10.21468/SciPostPhys.12.1.037. arXiv: 2109.04981 [hep-ph].

[251] Stephen Bailey et al. "Data and Analysis Preservation, Recasting, and Reinterpretation". In: (Mar. 2022). arXiv: 2203.10057 [hep-ph].

[252] *The European Strategy for Particle Physics Update 2013. La stratégie européenne pour la physique des particules Mise à jour 2013. 16th Session of European Strategy Council*. 2013. URL: https://cds.cern.ch/record/1567258.

[253] Jamie Shiers. *Collaborative Long-Term Data Preservation: From Hundreds of PB to Tens of EB*. 2018. DOI: 10.17181/CERN.ZP6P.7Z3Y.

[254] Jack Y. Araz et al. "Strength in numbers: Optimal and scalable combination of LHC new-physics searches". In: *SciPost Phys.* 14.4 (2023), p. 077. DOI: 10.21468/SciPostPhys.14.4.077. arXiv: 2209.00025 [hep-ph].

[255] *CERN Open Data Policy for the LHC Experiments*. http://opendata.cern.ch/docs/cern-open-data-policy-for-lhc-experiments. Accessed: 24 April 2024.

[256] Andy Buckley and Mike Whalley. "HepData reloaded: Reinventing the HEP data archive". In: *PoS* ACAT2010 (2010). Ed. by T. Speer et al., p. 067. DOI: 10.22323/1.093.0067. arXiv: 1006.0517 [hep-ex].

[257] Louie Dartmoor Corpe. *Update on ATLAS recommendations for content to be included in HEPData entries*. Tech. rep. Geneva: CERN, 2019. URL: https://cds.cern.ch/record/2701436.

[258] ATLAS Collaboration. "Reproducing searches for new physics with the ATLAS experiment through publication of full statistical likelihoods". In: *ATL-PHYS-PUB-2019-029* (2019). URL: https://cds.cern.ch/record/2684863.

[259] Lukas Heinrich, Matthew Feickert, and Giordon Stark. *pyhf: v0.7.6*. Version 0.7.6. https://github.com/scikit-hep/pyhf/releases/tag/v0.7.6. DOI: 10.5281/zenodo.1169739. URL: https://doi.org/10.5281/zenodo.1169739.

[260] Lukas Heinrich et al. "pyhf: pure-Python implementation of HistFactory statistical models". In: *Journal of Open Source Software* 6.58 (2021), p. 2823. DOI: 10.21105/joss.02823. URL: https://doi.org/10.21105/joss.02823.

[261] Kyle Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. New York: New York U., 2012. DOI: 10.17181/CERN-OPEN-2012-016. URL: https://cds.cern.ch/record/1456844.

[262] ATLAS Collaboration. *Search for direct stau production in events with two hadronic τ-leptons in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*. HEPData (collection). https://doi.org/10.17182/hepdata.92006. 2020.

[263] Jack Y. Araz et al. "Les Houches guide to reusable ML models in LHC analyses". In: (Dec. 2023). arXiv: 2312.14575 [hep-ph].

[264] Peter Skands, Stefano Carrazza, and Juan Rojo. "Tuning PYTHIA 8.1: the Monash 2013 Tune". In: *Eur. Phys. J. C* 74.8 (2014), p. 3024. DOI: 10.1140/epjc/s10052-014-3024-y. arXiv: 1404.5630 [hep-ph].

[265] Andy Buckley and Holger Schulz. "Tuning of MC generator MPI models". In: *Adv. Ser. Direct. High Energy Phys.* 29 (2018). Ed. by Paolo Bartalini and Jonathan Richard Gaunt, pp. 281–301. DOI: 10.1142/9789813227767_0013. arXiv: 1806.11182 [hep-ph].

[266] John Ellis et al. "Updated Global SMEFT Fit to Higgs, Diboson and Electroweak Data". In: *JHEP* 06 (2018), p. 146. DOI: 10.1007/JHEP06(2018)146. arXiv: 1803.03252 [hep-ph].

[267] Nathan P. Hartland et al. "A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector". In: *JHEP* 04 (2019), p. 100. DOI: 10.1007/JHEP04(2019)100. arXiv: 1901.05965 [hep-ph].

[268] Andy Buckley et al. "Rivet user manual". In: *Comput. Phys. Commun.* 184 (2013), pp. 2803–2819. DOI: 10.1016/j.cpc.2013.05.021. arXiv: 1003.0694 [hep-ph].

[269] Christian Bierlich et al. "Robust Independent Validation of Experiment and Theory: Rivet version 3". In: *SciPost Phys.* 8 (2020), p. 026. DOI: 10.21468/SciPostPhys.8.2.026. arXiv: 1912.05451 [hep-ph].

[270] ATLAS Collaboration. *SimpleAnalysis: Truth-level Analysis Framework*. 2022. DOI: 10.17181/CERN.R6S3.0QKV.

[271] Jonathan M. Butterworth et al. "Constraining new physics with collider measurements of Standard Model signatures". In: *JHEP* 03 (2017), p. 078. DOI: 10.1007/JHEP03(2017)078. arXiv: 1606.05296 [hep-ph].

[272] Eric Conte, Benjamin Fuks, and Guillaume Serret. "MadAnalysis 5, A User-Friendly Framework for Collider Phenomenology". In: *Comput. Phys. Commun.* 184 (2013), pp. 222–256. DOI: 10.1016/j.cpc.2012.09.009. arXiv: 1206.1599 [hep-ph].

[273] B. Dumont et al. "Toward a public analysis database for LHC new physics searches using MADANALYSIS 5". In: *Eur. Phys. J. C* 75.2 (2015), p. 56. DOI: 10.1140/epjc/s10052-014-3242-3. arXiv: 1407.3278 [hep-ph].

[274] Eric Conte and Benjamin Fuks. "Confronting new physics theories to LHC data with MADANALYSIS 5". In: *Int. J. Mod. Phys. A* 33.28 (2018), p. 1830027. DOI: 10.1142/S0217751X18300272. arXiv: 1808.00480 [hep-ph].

[275] Peter Athron et al. "GAMBIT: The Global and Modular Beyond-the-Standard-Model Inference Tool". In: *Eur. Phys. J. C* 77.11 (2017). [Addendum: Eur.Phys.J.C 78, 98 (2018)], p. 784. DOI: 10.1140/epjc/s10052-017-5321-8. arXiv: 1705.07908 [hep-ph].

[276] Manuel Drees et al. "CheckMATE: Confronting your Favourite New Physics Model with LHC Data". In: *Comput. Phys. Commun.* 187 (2015), pp. 227–265. DOI: 10.1016/j.cpc.2014.10.018. arXiv: 1312.2591 [hep-ph].

[277] Sabine Kraml et al. "SModelS: a tool for interpreting simplified-model results from the LHC and its application to supersymmetry". In: *Eur. Phys. J. C* 74 (2014), p. 2868. DOI: 10.1140/epjc/s10052-014-2868-5. arXiv: 1312.4175 [hep-ph].

[278] Mohammad Mahdi Altakach et al. "SModelS v2.3: Enabling global likelihood analyses". In: *SciPost Phys.* 15.5 (2023), p. 185. DOI: 10.21468/SciPostPhys.15.5.185. arXiv: 2306.17676 [hep-ph].

[279] Andy Buckley et al. *Consistent, multidimensional differential histogramming and summary statistics with YODA 2*. Dec. 2023. arXiv: 2312.15070 [hep-ph].

[280] R. Brun and F. Rademakers. "ROOT: An object oriented data analysis framework". In: *Nucl. Instrum. Meth. A* 389 (1997). Ed. by M. Werlen and D. Perret-Gallix, pp. 81–86. DOI: 10.1016/S0168-9002(97)00048-X.

[281] Charles R. Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2.

[282] J. D. Hunter. "Matplotlib: A 2D graphics environment". In: *Computing in Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.

[283] Pavel Fileviez Perez and Mark B. Wise. "Breaking Local Baryon and Lepton Number at the TeV Scale". In: *JHEP* 08 (2011), p. 068. DOI: 10.1007/JHEP08(2011)068. arXiv: 1106.0343 [hep-ph].

[284] Michael Duerr, Pavel Fileviez Perez, and Mark B. Wise. "Gauge Theory for Baryon and Lepton Numbers with Leptoquarks". In: *Phys. Rev. Lett.* 110 (2013), p. 231801. DOI: 10.1103/PhysRevLett.110.231801. arXiv: 1304.0576 [hep-ph].

[285] Pavel Fileviez Perez, Sebastian Ohmer, and Hiren H. Patel. "Minimal Theory for Lepto-Baryons". In: *Phys. Lett. B* 735 (2014), pp. 283–287. DOI: 10.1016/j.physletb.2014.06.057. arXiv: 1403.8029 [hep-ph].

[286] S. Amrith et al. "LHC Constraints on a $B - L$ Gauge Model using Contur". In: *JHEP* 05 (2019), p. 154. DOI: 10.1007/JHEP05(2019)154. arXiv: 1811.11452 [hep-ph].

[287] Pavel Fileviez Perez and Mark B. Wise. "Baryon and lepton number as local gauge symmetries". In: *Phys. Rev. D* 82 (2010). [Erratum: Phys.Rev.D 82, 079901 (2010)], p. 011901. DOI: 10.1103/PhysRevD.82.079901. arXiv: 1002.1754 [hep-ph].

[288] Pavel Fileviez Perez. "New Paradigm for Baryon and Lepton Number Violation". In: *Phys. Rept.* 597 (2015), pp. 1–30. DOI: 10.1016/j.physrep.2015.09.001. arXiv: 1501.01886 [hep-ph].

[289] B. C. Allanach, Joe Davighi, and Scott Melville. "An Anomaly-free Atlas: charting the space of flavour-dependent gauged $U(1)$ extensions of the Standard Model". In: *JHEP* 02 (2019). [Erratum: JHEP 08, 064 (2019)], p. 082. DOI: 10.1007/JHEP02(2019)082. arXiv: 1812.04602 [hep-ph].

[290] John Ellis, Malcolm Fairbairn, and Patrick Tunney. "Anomaly-Free Dark Matter Models are not so Simple". In: *JHEP* 08 (2017), p. 053. DOI: 10.1007/JHEP08(2017)053. arXiv: 1704.03850 [hep-ph].

[291] Pavel Fileviez Perez and Alexis D. Plascencia. "Theory of Dirac dark matter: Higgs boson decays and EDMs". In: *Phys. Rev. D* 105.9 (2022), p. 095021. DOI: 10.1103/PhysRevD.105.095021. arXiv: 2112.02103 [hep-ph].

[292] ATLAS Collaboration. "Differential $t\bar{t}$ cross-section measurements using boosted top quarks in the all-hadronic final state with 139 fb^{-1} of ATLAS data". In: *JHEP* 04 (2023), p. 080. DOI: 10.1007/JHEP04(2023)080. arXiv: 2205.02817 [hep-ex].

[293] Luc Darmé et al. "UFO 2.0: the 'Universal Feynman Output' format". In: *Eur. Phys. J. C* 83.7 (2023), p. 631. DOI: 10.1140/epjc/s10052-023-11780-9. arXiv: 2304.09883 [hep-ph].

[294] ATLAS Collaboration. "Measurement of the transverse momentum and ϕ^*_η distributions of Drell–Yan lepton pairs in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 76.5 (2016), p. 291. DOI: 10.1140/epjc/s10052-016-4070-4. arXiv: 1512.02192 [hep-ex].

[295] ATLAS Collaboration. "Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: *JHEP* 01 (2021), p. 033. DOI: 10.1007/JHEP01(2021)033. arXiv: 2006.09274 [hep-ex].

[296] ATLAS Collaboration. "Measurement of the $Z(\to \ell^+\ell^-)\gamma$ production cross-section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 03 (2020), p. 054. DOI: 10.1007/JHEP03(2020)054. arXiv: 1911.04813 [hep-ex].

[297] ATLAS Collaboration. "Measurement of differential cross sections and W^+/W^- cross-section ratios for W boson production in association with jets at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 05 (2018). [Erratum: JHEP 10, 048 (2020)], p. 077. DOI: 10.1007/JHEP05(2018)077. arXiv: 1711.03296 [hep-ex].

[298] ATLAS Collaboration. "Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in $\sqrt{s}$ = 8 TeV proton-proton collisions using the ATLAS detector". In: *Phys. Rev. D* 93.3 (2016), p. 032009. DOI: 10.1103/PhysRevD.93.032009. arXiv: 1510.03818 [hep-ex].

[299] Jon Butterworth et al. "Dark matter from anomaly cancellation at the LHC". In: *Phys. Rev. D* 110.7 (2024), p. 075001. DOI: 10.1103/PhysRevD.110.075001. arXiv: 2405.03749 [hep-ph].

[300] ATLAS Collaboration. "Search for new resonances in mass distributions of jet pairs using 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 03 (2020), p. 145. DOI: 10.1007/JHEP03(2020)145. arXiv: 1910.08447 [hep-ex].

[301] CMS Collaboration. "Search for high mass dijet resonances with a new background prediction method in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP* 05 (2020), p. 033. DOI: 10.1007/JHEP05(2020)033. arXiv: 1911.03947 [hep-ex].

[302] ATLAS Collaboration. "Search for Low-Mass Dijet Resonances Using Trigger-Level Jets with the ATLAS Detector in pp Collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. Lett.* 121 (8 Aug. 2018), p. 081801. DOI: 10.1103/PhysRevLett.121.081801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.081801.

[303] CMS Collaboration. "Search for narrow and broad dijet resonances in proton-proton collisions at $\sqrt{s} = 13$ TeV and constraints on dark matter mediators and other new particles". In: *JHEP* 08 (2018), p. 130. DOI: 10.1007/JHEP08(2018)130. arXiv: 1806.00843 [hep-ex].

[304] CMS Collaboration. "Search for narrow resonances in dijet final states at $\sqrt{s} = 8$ TeV with the novel CMS technique of data scouting". In: *Phys. Rev. Lett.* 117.3 (2016), p. 031802. DOI: 10.1103/PhysRevLett.117.031802. arXiv: 1604.08907 [hep-ex].

[305] ATLAS Collaboration. "Search for low-mass resonances decaying into two jets and produced in association with a photon using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 795 (2019), pp. 56–75. DOI: 10.1016/j.physletb.2019.03.067. arXiv: 1901.10917 [hep-ex].

[306] CMS Collaboration. "Search for low mass vector resonances decaying into quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D* 100 (11 Dec. 2019), p. 112007. DOI: 10.1103/PhysRevD.100.112007. URL: https://link.aps.org/doi/10.1103/PhysRevD.100.112007.

[307] CMS Collaboration. "Search for Low Mass Vector Resonances Decaying to Quark-Antiquark Pairs in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. Lett.* 119 (11 Sept. 2017), p. 111802. DOI: 10.1103/PhysRevLett.119.111802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.111802.

[308] ATLAS Collaboration. "Searches for scalar leptoquarks and differential cross-section measurements in dilepton-dijet events in proton-proton collisions at a centre-of-mass energy

of $\sqrt{s}$ = 13 TeV with the ATLAS experiment". In: *Eur. Phys. J. C* 79.9 (2019), p. 733. DOI: 10.1140/epjc/s10052-019-7181-x. arXiv: 1902.00377 [hep-ex].

[309] ATLAS Collaboration. "Measurements of $W^+W^-+ \geq 1$ jet production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 06 (2021), p. 003. DOI: 10.1007/JHEP06(2021)003. arXiv: 2103.10319 [hep-ex].

[310] ATLAS Collaboration. "Measurement of the production cross section of pairs of isolated photons in pp collisions at 13 TeV with the ATLAS detector". In: *JHEP* 11 (2021), p. 169. DOI: 10.1007/JHEP11(2021)169. arXiv: 2107.09330 [hep-ex].

[311] ATLAS Collaboration. "Measurement of the cross section for isolated-photon plus jet production in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector". In: *Phys. Lett. B* 780 (2018), pp. 578–602. DOI: 10.1016/j.physletb.2018.03.035. arXiv: 1801.00112 [hep-ex].

[312] ATLAS Collaboration. "Measurement of the $Z\gamma \to \nu\bar{\nu}\gamma$ production cross section in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector and limits on anomalous triple gauge-boson couplings". In: *JHEP* 12 (2018), p. 010. DOI: 10.1007/JHEP12(2018)010. arXiv: 1810.04995 [hep-ex].

[313] ATLAS Collaboration. "Observation of electroweak production of a same-sign W boson pair in association with two jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Rev. Lett.* 123.16 (2019), p. 161801. DOI: 10.1103/PhysRevLett.123.161801. arXiv: 1906.03203 [hep-ex].

[314] ATLAS Collaboration. "Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 07 (2020), p. 044. DOI: 10.1007/JHEP07(2020)044. arXiv: 2003.11960 [hep-ex].

[315] ATLAS Collaboration. "Measurement of fiducial and differential W^+W^- production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 79.10 (2019), p. 884. DOI: 10.1140/epjc/s10052-019-7371-6. arXiv: 1905.04242 [hep-ex].

[316] ATLAS Collaboration. "Measurement of the four-lepton invariant mass spectrum in 13 TeV proton-proton collisions with the ATLAS detector". In: *JHEP* 04 (2019), p. 048. DOI: 10.1007/JHEP04(2019)048. arXiv: 1902.05892 [hep-ex].

[317] ATLAS Collaboration. "Measurements of differential cross-sections in four-lepton events in 13 TeV proton-proton collisions with the ATLAS detector". In: *JHEP* 07 (2021), p. 005. DOI: 10.1007/JHEP07(2021)005. arXiv: 2103.01918 [hep-ex].

[318] ATLAS Collaboration. "Measurements of $Z\gamma$ and $Z\gamma\gamma$ production in pp collisions at $\sqrt{s} =$ 8 TeV with the ATLAS detector". In: *Phys. Rev. D* 93.11 (2016), p. 112002. DOI: 10.1103/PhysRevD.93.112002. arXiv: 1604.05232 [hep-ex].

[319] ATLAS Collaboration. "Measurements of fiducial and differential cross sections for Higgs boson production in the diphoton decay channel at $\sqrt{s} = 8$ TeV with ATLAS". In: *JHEP* 09 (2014), p. 112. DOI: 10.1007/JHEP09(2014)112. arXiv: 1407.4222 [hep-ex].

[320] ATLAS Collaboration. "Measurement of inclusive jet and dijet cross-sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 05 (2018), p. 195. DOI: 10.1007/JHEP05(2018)195. arXiv: 1711.02692 [hep-ex].

[321] F. Siegert. *MC_JETS analysis*. https://rivet.hepforge.org/analyses/MC_JETS.html.

[322] A. Buckley et al. *MC_TTBAR analysis*. https://rivet.hepforge.org/analyses/MC_TTBAR.html.

Appendix A

Comparison of Lund string and cluster hadronisation model

As part of this thesis work, the effects of using the Lund string or cluster hadronisation model is studied using both a set of $t\bar{t}$ and dijet events. The general idea is to generate a set of events with the same matrix-element calculation and parton shower, and in this way isolate the effect of changing the underlying hadronisation models.

Figure A.1 shows a selection of POWHEG-BOX [89–91] $t\bar{t}$ events that are showered and hadronised using HERWIG 7.1.6 [82–84], comparing the cluster and Lund string hadronisation models. This RIVET analysis is designed for purely MC based $t\bar{t}$ production studies [322], characterising the final state with top-quarks via reconstructions of leptons and jets instead of partonic tops, and ultimately plotting variables such as jet p_{T} or reconstructed masses for top quarks and W bosons. It is good to emphasise that both sets of plots constitute the *exact* same set of events up until the parton shower, with the only difference being the hadronisation model that was used.

Most observables show agreement within statistical uncertainties, such as the jet p_{T} and the scalar sum of transverse momenta H_{T}. However, there is a shape difference around the reconstructed W and top masses. By default, and in this study, the analysis is run in the single-lepton decay mode of the $t\bar{t}$ system. The hadronically decaying W boson is reconstructed from the pair of non-b-tagged jets that best matches the W mass. If the two hadronisation models yield hadrons with different kinematic properties and distribute them differently between jets, this can explain the mass discrepancy around the M_{W} peak. Considering both sets of histograms differ *only* in the hadronisation model, this effect can be related to the fact that the cluster model populates the tail of the jet multiplicity distribution more strongly. If some $t\bar{t}$ event leads to *less* reconstructed jets in the Lund string model than in the cluster model, the (less numerous) jets in the Lund string will carry more energy because of energy conservation: the energy has to go somewhere. This would then explain the mass peak of the W mass is shifted towards *higher* values for the Lund string model compared to the cluster model. A jet substructure analysis can shed further light on this hypothesis and effects in jet reconstruction from using different hadronisation models. The top mass is reconstructed from the combined four-momenta of the W boson and a b-jet, propagating the effect of the hadronic W mass reconstruction and potential mismodelling of the b-jet mass to

Sample	Herwig7 version	Parton shower	Hadronisation model
1	7.1.6	Angular-ordered	Cluster
2	7.1.6	Angular-ordered	Lund string
3	7.2.1	Angular-ordered	Cluster
4	7.1.6	Dipole	Cluster

Table A.1: Four different sets of dijet samples.

the discrepancy that is seen around the reconstructed mass of the top quark.

Furthermore, four sets of dijet samples, c.f. Table A.1, are produced for validation and compared to a dijet cross-section measurement in ATLAS [320] using Rivet in Figure A.2. In this context, the most important configurations to compare are again the blue and red histograms, using the exact same setup except for the hadronisation model. Figures A.2a and A.2b show consistent shape modelling of the jet p_{T} distributions between the two different hadronisation models. Both MC predictions underestimate the experimental data, which can be expected because the setups only include LO terms in the matrix-element calculation. The jet mass plots in Figures A.2c and A.2d have a slight shape difference in the low-mass region, where the Lund string model has a larger fiducial cross-section around $200\,\mathrm{GeV}$. The RIVET analysis for the dijet cross-section measurement does not go down to lower dijet mass, but we can investigate the low-mass m_{jj} region with the purely based MC study [321]. Figure A.3 shows the jet-mass distributions of the leading and sub-leading jets in p_{T}. Here, we see again that the Lund string model peaks at a slightly higher mass compared to the cluster model and that the Lund string model has a harder jet p_{T} spectrum.

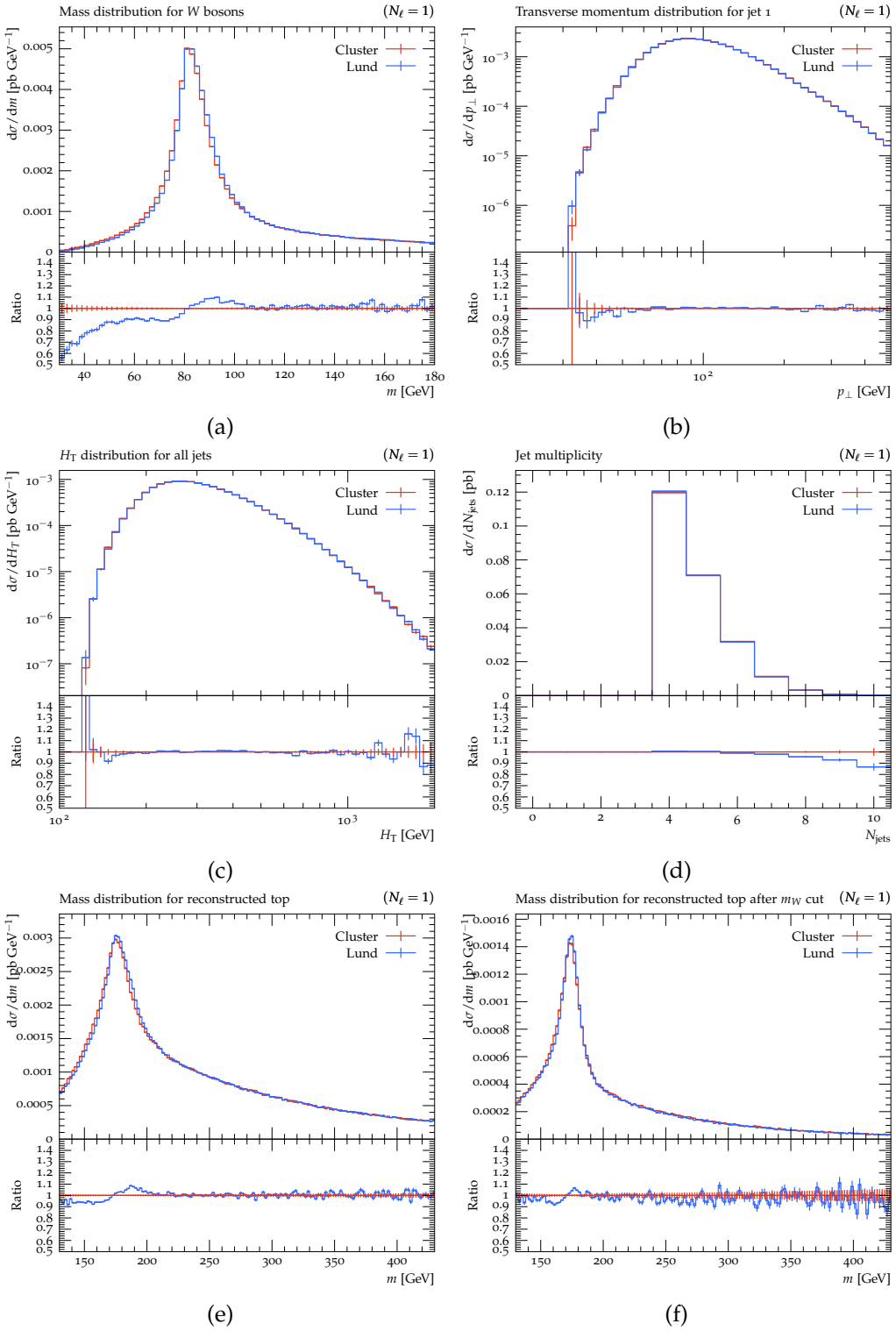


Figure A.1: Observables for $t\bar{t}$ production samples with the Cluster and Lund string hadronisation models, generated with POWHEG+Herwig 7.1 and corresponding to the MC_TTBAR Rivet analysis [322], that generates multiple particle-level quantities from input MC samples. The figures show a discrepancy between mass-dependent observables for the Lund string model (in blue) compared to the cluster model (in red), as well as the tail of the jet multiplicity in subfigure (d).

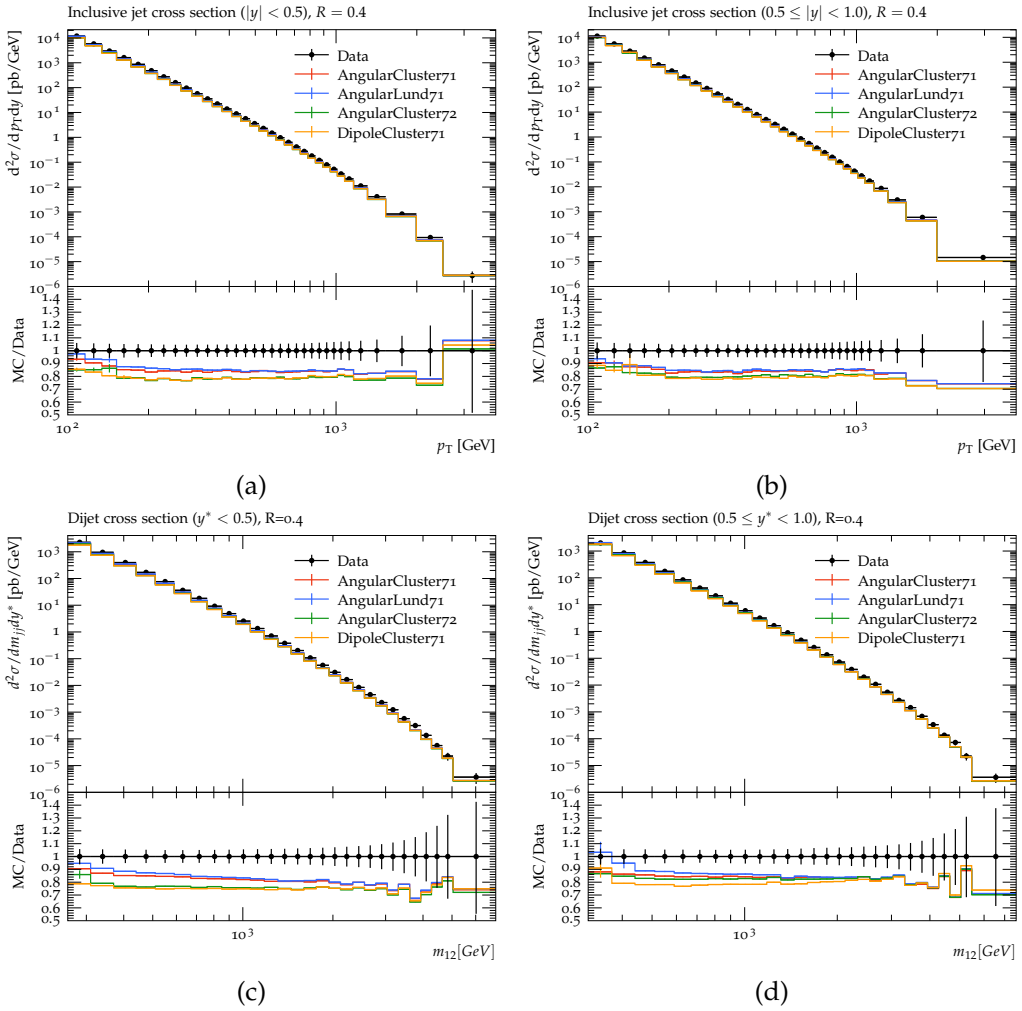


Figure A.2: Observables for four dijet production samples, corresponding to Table A.1, compared to a $3.2\,\mathrm{fb}^{-1}$ inclusive measurement of dijets from ATLAS [320] using RIVET. Subfigure (a) and (b) show the doubly-differential distribution in rapidity and jet p_T as a function of jet p_T, while subfigure (c) and (c) show the corresponding differential distributions (in the same rapidity region) as a function of the jet mass. The shape of the jet p_T are consistent between the Lund string (blue) and cluster hadronisation model (red), but the low end of the mass distribution shows a shape discrepancy as explained in the text.

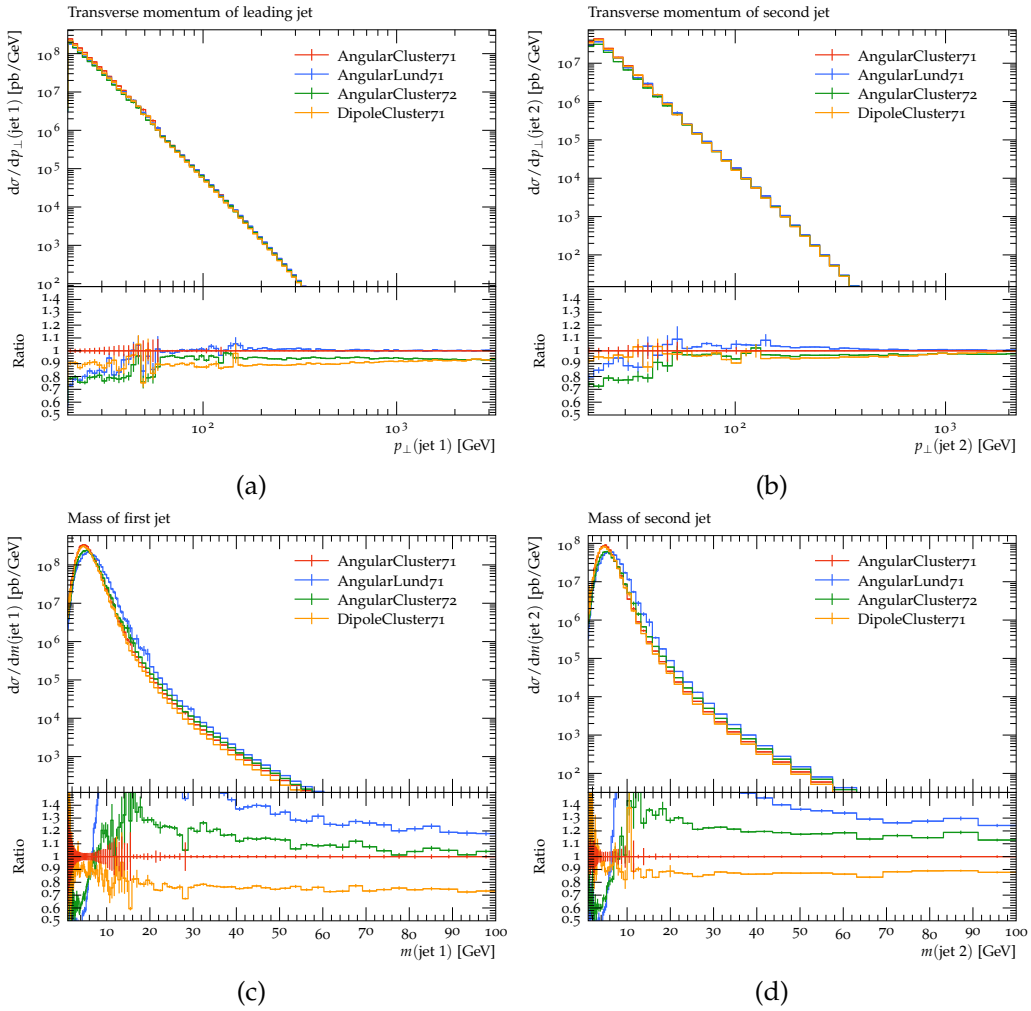


Figure A.3: Observables for dijet production samples indicated in Table A.1. The observables are generated with the MC_DIJET RIVET analysis [321], that generates multiple particle-level quantities from input MC samples. The emphasis is on a comparison between the Cluster and Lund string hadronisation models, generated with POWHEG+Herwig 7.1, in red and blue histograms respectively. The differential distributions as a function of jet p_{T} in subfigure (a) and (b) show that the Lund string model has a harder p_{T} spectrum than the cluster model for the leading and sub-leading jets (in p_{T}) respectively. Consequently, the reconstructed jet mass of the leading and sub-leading jets peaks at a higher value for sample generated with the Lund string model compared to the cluster model.

Appendix B

Migration matrices for auxiliary lepton regions

This section highlights the migration matrices used in the unfolding for the single-lepton and di-lepton channels. The matrices for the $p_{\mathrm{T}}^{\mathrm{recoil}}$ observable for the ≥ 1 jet and VBF region are shown in Figure B.1 and B.2 respectively, while the m_{jj} and $\Delta\phi_{jj}$ observables for the VBF region are shown in Figure B.3 and B.4 respectively. Like the $p_{\mathrm{T}}^{\mathrm{miss}}$+jets region matrices presented in the main body of the text, all matrices are mostly diagonal where the diagonal entries that correspond to the purity are at least 60%.

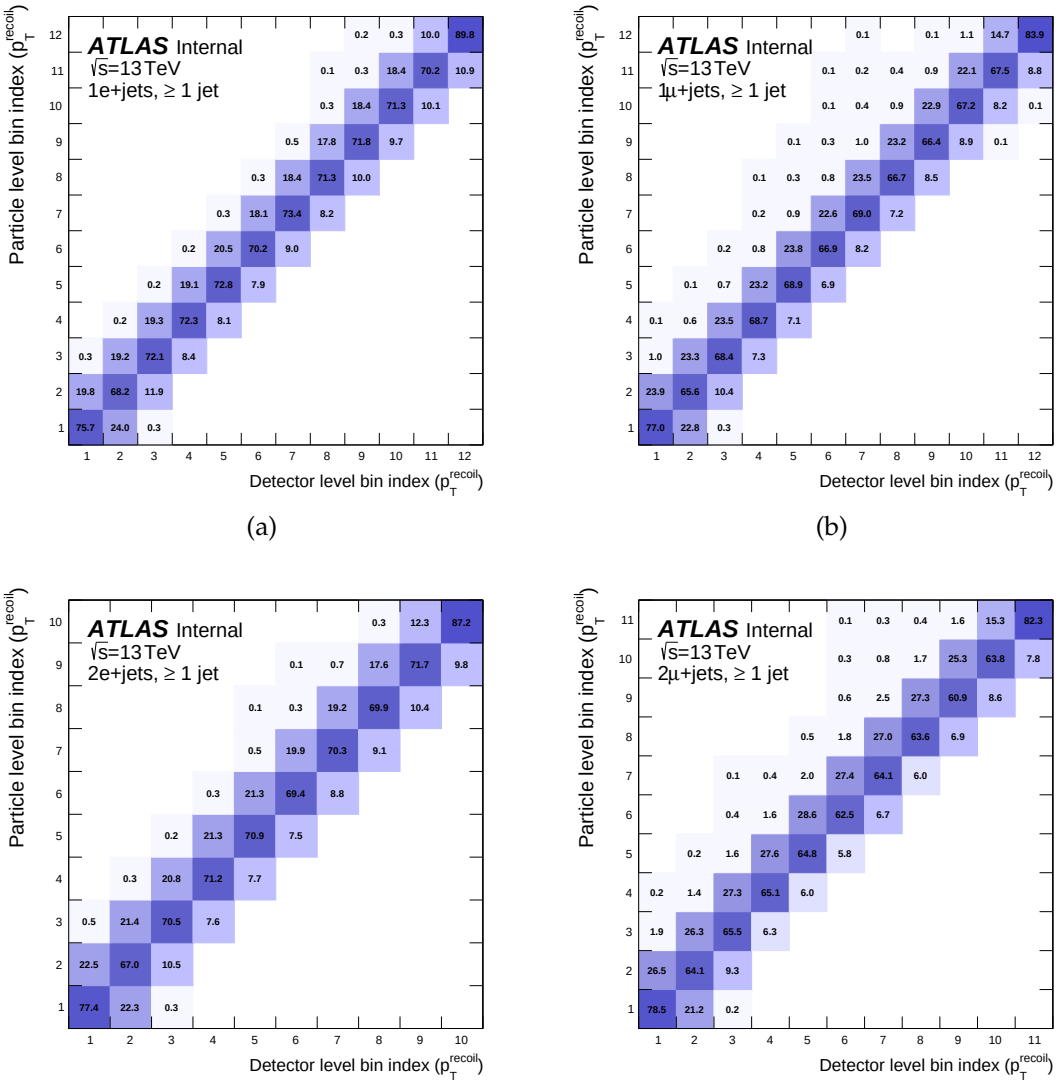


Figure B.1: Migration matrices for the $p_{\mathrm{T}}^{\mathrm{recoil}}$ observable in the ≥ 1 jet region for the single-lepton and di-lepton auxiliary regions. In all plots, the diagonal entries (i.e. the purity) is observed to be at least 60%, which is found to lead to a reliable convergence of the unfolding within a small number of iterations such that the statistical uncertainties from the unfolding procedure are reduced. The off-diagonal elements indicate that there is more migration to higher $p_{\mathrm{T}}^{\mathrm{miss}}$ bins than to lower $p_{\mathrm{T}}^{\mathrm{miss}}$ bins. This is explained by the fact that these are steeply falling distributions and that reconstruction of the magnitude of a vector, i.e. $p_{\mathrm{T}}^{\mathrm{miss}}$, at a slightly different value leads to relatively more migrations to high $p_{\mathrm{T}}^{\mathrm{miss}}$ bins.

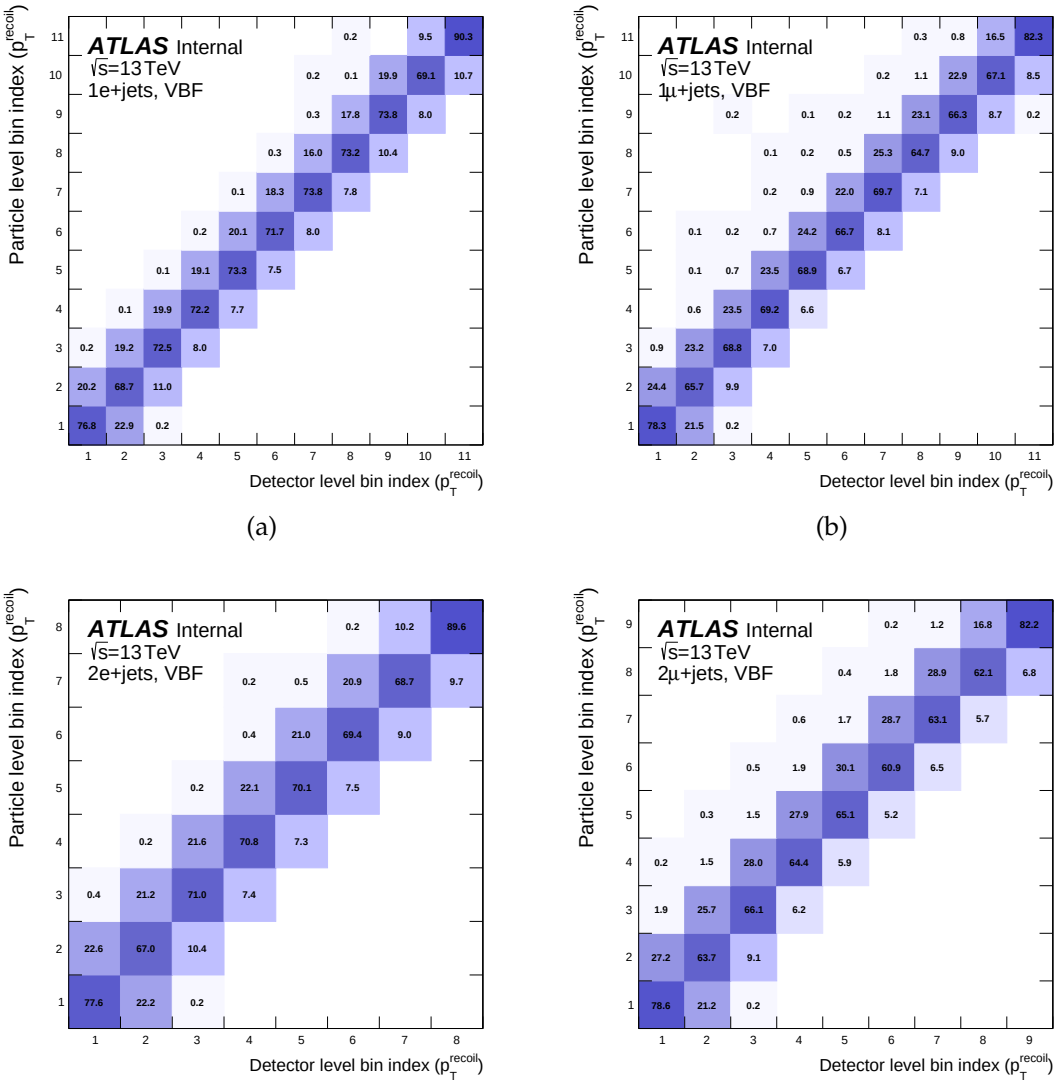


Figure B.2: Migration matrices for the $p_{\mathrm{T}}^{\mathrm{recoil}}$ observable in the VBF region for the single-lepton and di-lepton auxiliary regions. In all plots, the diagonal entries (i.e. the purity) is observed to be at least 60%, which is found to lead to a reliable convergence of the unfolding within a small number of iterations such that the statistical uncertainties from the unfolding procedure are reduced. The off-diagonal elements indicate that there is more migration to higher $p_{\mathrm{T}}^{\mathrm{miss}}$ bins than to lower $p_{\mathrm{T}}^{\mathrm{miss}}$ bins. This is explained by the fact that these are steeply falling distributions and that reconstruction of the magnitude of a vector, i.e. $p_{\mathrm{T}}^{\mathrm{miss}}$, at a slightly different value leads to relatively more migrations to high $p_{\mathrm{T}}^{\mathrm{miss}}$ bins.

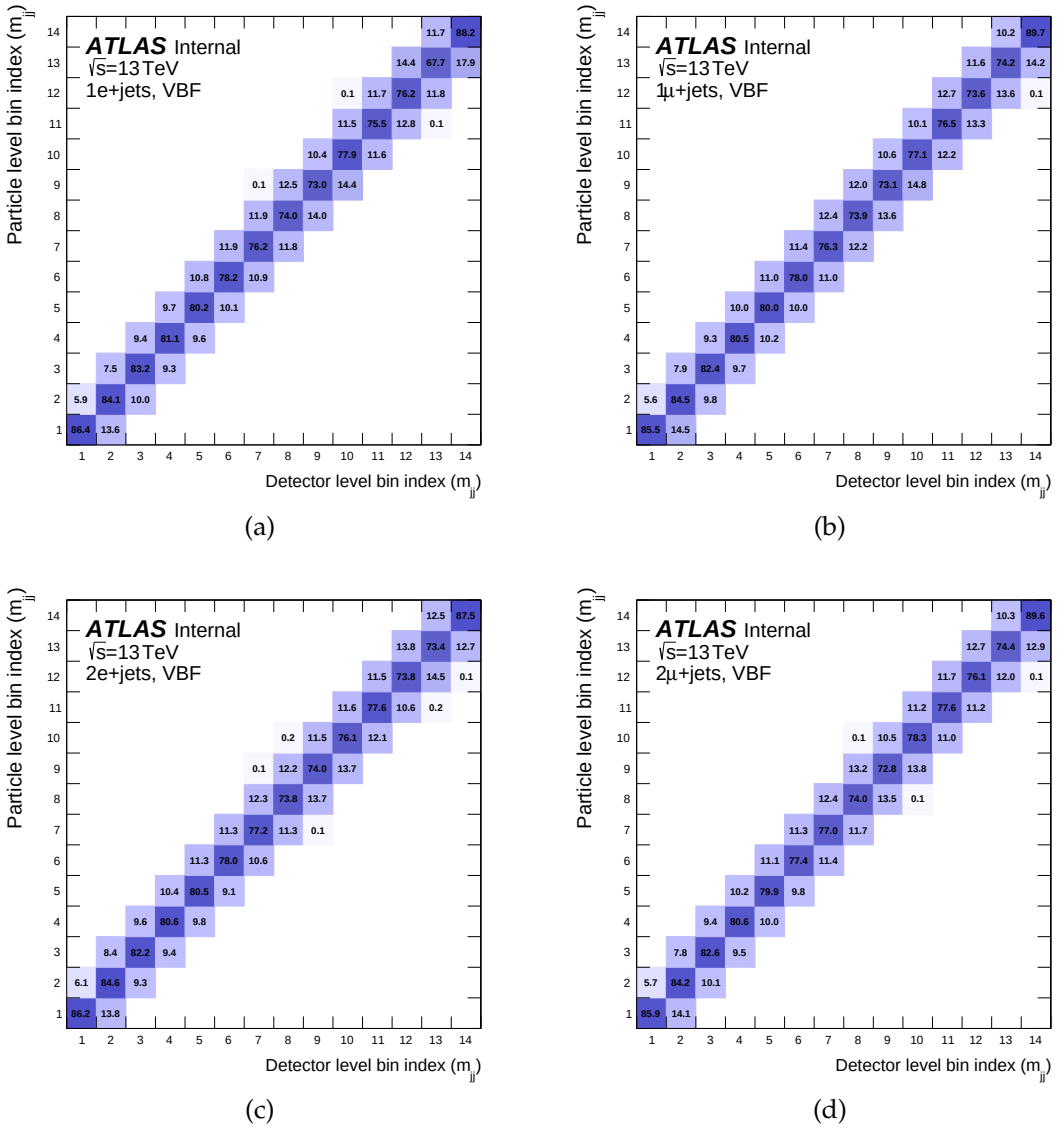


Figure B.3: Migration matrices for the m_{jj} observable in the VBF region for the single-lepton and di-lepton auxiliary regions. In all plots, the diagonal entries (i.e. the purity) is observed to be at least 60%, which is found to lead to a reliable convergence of the unfolding within a small number of iterations such that the statistical uncertainties from the unfolding procedure are reduced.

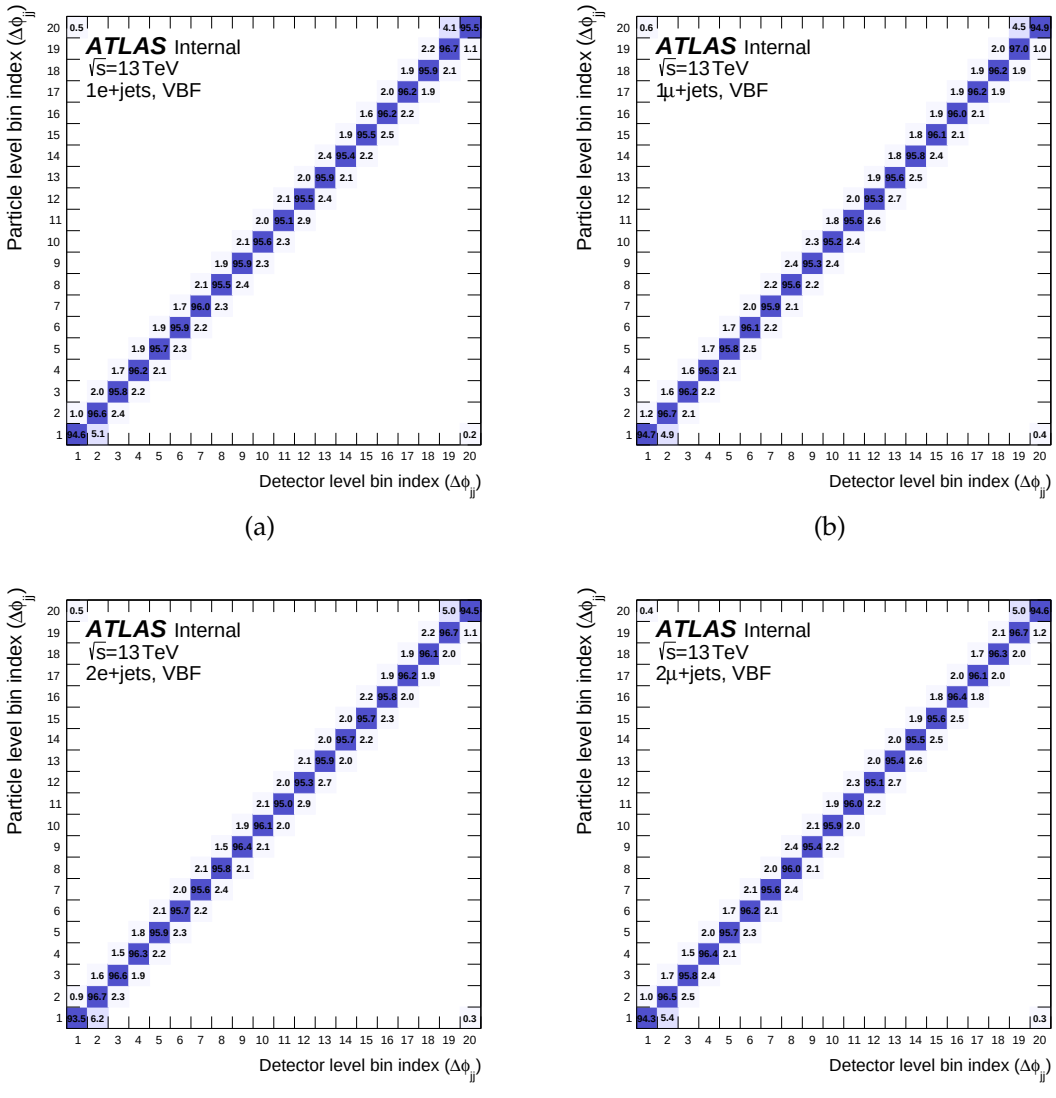


Figure B.4: Migration matrices for the $\Delta\phi_{jj}$ observable in the VBF region for the single-lepton and di-lepton auxiliary regions. In all plots, the diagonal entries (i.e. the purity) is observed to be at least 94%, which is found to lead to a reliable convergence of the unfolding within a small number of iterations such that the statistical uncertainties from the unfolding procedure are reduced.