

A Study of the Substructure of Jets Initiated by Top Quarks and W Bosons

Alexander Sopio

A dissertation submitted in fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Department of Physics and Astronomy
University College London

2022



Declaration

I, Alexander Soppio, 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 work.

Abstract

Large hadronic jets initiated by the decays of heavy objects such as top quarks and W bosons have long been a very active field of study in particle physics. The Lund Jet Plane is a new and versatile tool for investigating the sub-structure of these complex objects. This thesis presents a collection of studies involving this new observable including contributions to its first-ever measurement in $p - p$ collision data. A brand new measurement of the Lund Jet Plane is also presented, as well as the development of new tagging techniques involving Deep Learning technologies. All measurements use the ATLAS data collected during Run 2, from 2015 to 2018, at $\sqrt{s} = 13\text{TeV}$.

A brand new measurement of the Lund Jet Plane in semi-leptonic $t\bar{t}$ decays is presented. It is the first measurement of this observable for boosted-object topologies using $R = 1.0$ jets. The measurement complements other measurements of the LJP which have only probed small-radius jets predominantly initiated by light quarks and highly energetic gluons. The measurement is unfolded to the particle level using Iterative Bayesian Unfolding. The unfolded measurement is compared to a wide range of Monte Carlo event generators. Measurements of new jet substructure observables, like the Lund Jet Plane, in new jet topologies are valuable for understanding the composition of these jets and can help improve modelling by providing new tunes for Monte Carlo generators.

The performance of various W and top jet taggers is evaluated and calibrated using the data to provide new tagging recommendations for ATLAS analyses. A completely new class of W and top jet taggers is also introduced. These taggers make use the many possible representations of the Lund Jet Plane as input to state-of-the art Deep Neural Network architectures, including Deep Convolutional Neural Networks, Recurrent Neural Networks and Graph Neural Networks.

Impact Statement

The research presented in this thesis was conducted within the ATLAS collaboration. The most important outcome of this research has been a completely new measurement of a little studied observable, the Lund Jet Plane, in hadronic jets Initiated by top quarks and W bosons. This is the first ever measurement of this observable for these objects. The measurement provides new insights into the structure of top quark and W boson initiated jets and can help improve our ability to identify top quark and W boson decays to hadrons. The results can also help improve the tuning of Monte Carlo event generators, which are crucial tools for testing theoretical predictions of proton-proton collisions against measurements from the experiments at the Large Hadron Collider.

Another strong focus of this work is the development of top quark and W boson taggers. These taggers are used in many searches and measurements to help identify interesting physical processes. The studies presented here concern both the evaluation of the tagger performance in data and the development of completely novel tagging tools. Results of the performance studies have become part of the ATLAS analysis recommendations which are followed by a large number of analysis teams across the collaboration. State-of-the-art Machine learning algorithms, including Deep Convolutional Neural Networks, Recurrent Neural Networks and Graph Neural Networks, which are used for many applications outside high energy physics including computer vision and natural language processing, have been utilised in the development of these tools.

Acknowledgements

I would like to thank Mario Campanelli for his thoughtful supervision and for always being available to chat. I would like to thank Federica Fabbri for providing invaluable advice and support with hands-on analysis work; Matt LeBlanc, Ben Nachman and Jennifer Roloff for allowing me to contribute to the first measurement of the Lund Jet Plane; the members of the ATLAS jet- E_T^{miss} group and the W and top jet tagging subgroup for their important feedback during the course of countless meetings. I am also very grateful to the many undergraduate students whose thesis projects Mario has supervised over the last five years who have made many valuable contributions to the development of machine-learning jet taggers at UCL. Lastly, I would like to thank Nikos Konstantinidis and the UCL CDT in Data Intensive Science for organising an excellent PhD programme with interesting partner-projects, fascinating seminars and a terrific community of colleagues from the UCL HEP and Astronomy groups.

Contents

1	Introduction	10
1.1	The Standard Model	10
1.1.1	The Standard Model Lagrangian	13
1.1.2	The electroweak sector	14
1.1.3	The Higgs sector and spontaneous symmetry breaking	18
1.1.4	The QCD sector	21
1.1.5	Beyond the Standard Model	23
1.1.6	Physics of proton-proton collisions	24
1.1.7	Parton showers and Monte Carlo event generators	26
1.2	The ATLAS experiment at the LHC	30
1.2.1	The Large Hadron Collider	30
1.2.2	Overview of the ATLAS detector	32
1.2.3	The Inner Detector	32
1.2.4	The calorimeters	35
1.2.5	Magnet systems	41
1.2.6	Muon systems	41
1.2.7	Triggers and data processing	42
1.3	Jet inputs	43
1.3.1	Topoclusters	44
1.3.2	Tracks	48
1.3.3	Particle Flow Objects (PFO)	48
1.3.4	Track-Calo Clusters (TCC)	50
1.3.5	Unified Flow Objects (UFO)	52
1.4	Jets	54
1.4.1	The purpose of jets	54
1.4.2	IRC safety	54
1.4.3	Jet algorithms	55
1.5	Jet substructure	58
1.5.1	Jet grooming	60
1.5.2	Jet shapes	61
1.5.3	Soft Drop	65
1.6	Lund diagrams	65

1.7	The Lund Jet Plane	66
2	Run 2 signal efficiency measurements for W and top jet taggers	71
2.1	Introduction	71
2.2	Measuring the signal efficiency scale factor in data	71
2.2.1	Data and Monte Carlo samples	72
2.2.2	Background estimation: template fits	73
2.2.3	Selections	74
2.2.4	Cut-based taggers	75
2.2.5	DNN taggers	76
2.3	Systematic uncertainties	79
2.3.1	Sources of systematic uncertainty	79
2.4	Effect of in situ jet mass calibrations	81
2.4.1	Pre-2018 calibration of large- R jet energy and mass	81
2.4.2	In situ calibration of large- R jet energy and mass	81
2.4.3	Signal efficiency and scale factors for different mass calibrations	82
2.5	Alternative Monte Carlo truth jet definitions	89
2.5.1	Early Run 2 truth jet definition	89
2.5.2	Full Run 2 truth jet definition	90
2.5.3	Effect of new truth jet definitions on the signal efficiency	91
2.6	Conclusion	97
2.7	Outlook	98
3	Measurement of the Lund Jet Plane in dijet events	99
3.1	Introduction	99
3.1.1	Data and Monte Carlo samples	99
3.1.2	Object reconstruction and selections	100
3.1.3	Reconstruction of the Lund Jet Plane	101
3.1.4	Iterative Bayesian Unfolding	103
3.2	Effect of unmatched tracks and their composition	110
3.3	Effect of dead TileCal modules in 2018 data	114
3.4	Unfolded results	118
3.5	Conclusion and subsequent developments	122
4	Measurement of the Lund Jet Plane in $t\bar{t}$ events	124
4.1	Introduction	124
4.1.1	Construction of the Lund Jet Plane	125
4.2	Data, objects and selections	128
4.2.1	Dataset	128
4.2.2	Triggers	128
4.2.3	Event selection	128
4.2.4	Object definitions	137

4.2.5	Particle level definitions	139
4.2.6	Monte Carlo simulation	141
4.3	Background estimation	145
4.3.1	Estimation of "fake" lepton background events	145
4.3.2	Difference in the total $t\bar{t}$ cross-section between data and Monte Carlo	146
4.3.3	Detector level signal and background	148
4.3.4	Detector-level background cross-check	151
4.4	Unfolding	153
4.4.1	The response matrix	153
4.4.2	Acceptance and inefficiency factors	157
4.4.3	Comparison of emission matching procedures	159
4.4.4	Technical closure test	162
4.4.5	Data-driven non-closure uncertainty	163
4.4.6	Choice of the number of unfolding iterations	166
4.5	Unfolded results	168
4.6	Uncertainties	170
4.6.1	Statistical uncertainty	170
4.6.2	Systematic uncertainties	173
4.6.3	$t\bar{t}$ modelling systematics	173
4.6.4	Object-reconstruction systematics	174
4.6.5	Background modelling systematics	176
4.6.6	Summary of systematic uncertainties	177
4.7	Comparison of the unfolded results to alternative $t\bar{t}$ MC samples	186
4.7.1	Comparison of $t\bar{t}$ MC/MC ratios	197
4.8	Conclusion	202
5	Deep Learning W and top jet taggers	204
5.1	Introduction	204
5.2	Machine learning in jet tagging	204
5.2.1	Performance metrics	205
5.2.2	Monte Carlo samples	206
5.2.3	Jet definitions and truth-jet labelling	207
5.3	Jet Image taggers	207
5.3.1	Convolutional Neural Networks	210
5.4	Using the Lund Jet Plane for W and top tagging	212
5.4.1	Comparison of CNN performance for Jet Images and LJPs	213
5.4.2	Recurrent Neural Networks	216
5.4.3	Higher-order LJPs	218
5.4.4	RNN tagger comparisons	218
5.4.5	Using UFOs as inputs to the LJP	222

5.4.6	Tagger performance vs. p_T at fixed working points	224
5.4.7	Graph Neural Networks	226
5.5	Conclusion	230
6	Thesis Conclusion	231
	Appendices	235
A	Performance of Run 2 W taggers	235
B	Measurement of the Lund Jet Plane in $t\bar{t}$ events	236
B.1	Complete systematic uncertainties on the LJP	236
B.2	Covariance and correlation matrices	253
B.3	Normalisation of the LJP with uncertainties	258
B.4	Background cross-checks	259
B.5	Smallest & greatest Data-MC LJP for alternative MC samples	260
C	Deep learning W and top taggers	267
C.1	UFO types within W and top jets	267
C.2	Model architectures	268
	Bibliography	273

Chapter 1

Introduction

1.1 The Standard Model

Throughout history, humans have invented countless theories to explain what makes up the world around them. The notion that all matter is composed of discrete indivisible particles – *atoms* – is an ancient one, perhaps going all the way back to the Greek natural philosopher Democritus. In a modern context, when we speak of atoms we are not referring to such indivisible units, but to composite particles made up of even smaller units called elementary particles. Interactions between these particles are governed by four fundamental forces, through which one species of particle may be transformed into another or entirely new particles may be created. The fundamental laws that govern the elementary particles and forces of the Universe are the subject of the field of particle physics [1].

The Standard Model of particle physics is the most complete theory of the Universe's most fundamental constituents and forces that is known to date. It is the product of nearly a century of scientific advances. The mathematical language used to construct the Standard Model is known as Quantum Field Theory (QFT). Within the framework of Quantum Field Theory, both particles and force fields are described by mathematical entities known as quantum fields. For this reason, the particles that constitute matter and the fundamental forces can both be referred to as either fields or particles depending on the context. An introduction to the formalism of QFT and its relationship to the Standard Model can be found in refs. [2–4].

A summary of the particles that make up the Standard Model is shown in Table 1.1. Four characteristic properties of the elementary particles and force carriers: mass, electric charge, and Weak isospin are also shown. The electric charge quantifies the strength with which a particle couples to the electromagnetic force carrier, the photon, and is conserved under all Standard Model interactions. The Weak isospin is associated with the coupling strength of the Weak interaction where a $W^\pm$ or Z boson is exchanged, and the particle mass determines the strength with which the Higgs field H couples to the particle. The strong force, which is mediated by gluons that couple to the colour charge, has been left out of the table for brevity as all types of quarks and gluons can carry many (3 in the case of quarks, 8 in the case of gluons) different colour states.

Table 1.1: Fundamental particles and forces of the Standard Model. Data is taken from [5].

Category	Symbol	Spin	Electric Charge	3 rd Component of Weak Isospin (T_3)	Mass
Quarks	d	1/2	-1/3	-1/2	4.7 MeV
	u	1/2	+2/3	+1/2	2.2 MeV
	s	1/2	-1/3	-1/2	95 MeV
	c	1/2	+2/3	+1/2	1.3 GeV
	b	1/2	-1/3	-1/2	4.2 GeV
	t	1/2	+2/3	+1/2	173 GeV ¹
Leptons	e	1/2	-1	-1/2	0.51 MeV
	ν_e	1/2	0	+1/2	< 2 eV
	μ	1/2	-1	-1/2	105.6 MeV
	ν_μ	1/2	0	+1/2	< 2 eV
	τ	1/2	-1	-1/2	1776 MeV
	ν_τ	1/2	0	+1/2	< 2 eV
Gauge bosons	γ	1	0	0	0
	$W^\pm$	1	± 1	± 1	80.3 GeV
	Z^0	1	0	0	90.2 GeV
	g	1	0	-	0
Higgs boson	H	0	0	-1/2	125.2 GeV

¹There exists no consistent definition for the masses of the six different quarks. While the five light quarks only exist within bound states, the top quark has a lifetime which is too short for it to form hadrons. Hence, the mass of the top quark can be precisely determined by measuring its decay width. For the light quarks, this is not possible. The masses quoted are instead effective quark masses which are defined based on properties of hadrons. More details on quark masses can be found in ref. [5].

1.1.1 The Standard Model Lagrangian

On a basic level, the relationships between the particle and force fields of the Standard Model can be illustrated using the mathematical expression for the Lagrangian density $\mathcal{L}$. In QFT, this equation is used to derive the equations of the fields. However, it is possible to deduce some important facts about the interactions of the Standard Model without performing these derivations explicitly. In a highly abbreviated form, the Standard Model Lagrangian may be written as

$$\mathcal{L}_{SM} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + i\bar{\psi}\not{D}\psi + (h.c.) + \bar{\psi}_i y_{ij} \psi_j \phi + (h.c.) + |D_\mu \phi|^2 - V(\phi), \quad (1.1)$$

where the fermion fields are denoted by ψ , ϕ denotes the Higgs field, y_{ij} is a matrix (mass matrix) describing the couplings of the Higgs field to the fermions and $V(\phi)$ denotes the Higgs potential². The abbreviation $(h.c.)$ denotes the Hermitian conjugate of the preceding term. Standard Model theory requires the Lagrangian to be invariant under a $SU(3)_C \times SU(2)_L \times U(1)_Y$ local gauge symmetry. By defining the gauge covariant derivative D_μ , one obtains the fields of the Standard Model interactions, or gauge fields, that are also contained within the combined field tensor $F^{\mu\nu}$. An introduction to gauge symmetries, local gauges, and how they relate to conservation laws in physics can be found in ref. [6]. The $SU(2)_L \times U(1)_Y$ part describes both the Electroweak and the Higgs sector of the theory and is transformed, through spontaneous symmetry breaking, into the $SU(2)_L \times U(1)_Q$ symmetry group. Each gauge symmetry has its own associated gauge field representing the four gauge bosons, or force carriers, that mediate the electroweak and strong interactions. The strong interaction is described by the $SU(3)$ invariant part of the Lagrangian. The theory of the strong force is also called Quantum Chromodynamics (QCD). The terms of the Standard Model Lagrangian that involve more than one type of field represent interactions and can be grouped by the types of force fields that they include. In this way, one obtains terms containing either the QCD or Electroweak gauge

²As is convention in particle physics, all mathematical expressions in this text are stated in natural units where $\hbar/2\pi = c = \epsilon_0 = \mu_0 = 1$ unless otherwise specified.

fields or the Higgs field:

$$\mathcal{L}_{SM}^{int} = \mathcal{L}_{QCD} + \mathcal{L}_{EW} + \mathcal{L}_{Higgs}. \quad (1.2)$$

The following sections will give a brief description of each of these three terms and introduce the theories of the interactions that they describe.

1.1.2 The electroweak sector

The Electroweak part of the Standard Model is invariant under the $SU(2)_L \times U(1)_Y$ symmetry. The conserved quantities associated with the $SU(2)_L$ and $U(1)_Y$ symmetry groups are *weak isospin* T and *hypercharge* Y . The hypercharge relates to the electric charge Q and the third component of weak isospin T_3 through the formula $Q = T_3 + \frac{1}{2}Y$. These three quantum numbers are conserved for all electroweak processes. The $U(1)_Y$ and $SU(2)_L$ symmetry groups have associated gauge fields B_μ and W_μ^α , where $\alpha = 1, 2, 3$. The kinetic terms in the Lagrangian for the electroweak fields can be written as

$$\mathcal{L}_{EW,gauge} = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W_{\mu\nu}^\alpha W^{\alpha\mu\nu}, \quad (1.3)$$

where we have defined the field-strength tensors

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (1.4)$$

$$W_{\mu\nu}^\alpha = \partial_\mu W_\nu^\alpha - \partial_\nu W_\mu^\alpha - g_W \varepsilon_{\alpha\beta\gamma} W_\mu^\beta W_\nu^\gamma, \quad (1.5)$$

where $\varepsilon_{\alpha\beta\gamma}$ is the three-dimensional Levi-Civita tensor. The coupling constants for the W_μ^α and B_μ fields we denote as g_W and g' .

The fields W_μ^α and B_μ do not correspond to the physical vector bosons, or mass eigenstates, which are

$$A_\mu = \cos \theta_W B_\mu + \sin \theta_W W_\mu^{(3)} \quad (1.6)$$

$$Z_\mu = -\sin \theta_W B_\mu + \cos \theta_W W_\mu^{(3)} \quad (1.7)$$

$$W^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^{(1)} \pm i W_\mu^{(2)} \right), \quad (1.8)$$

where A_μ , Z_μ and $W^\pm$ are the photon, Z boson and $W^\pm$ boson fields and the weak mixing angle, or Weinberg angle, has been defined as

$$\tan \theta_W = \frac{g'}{g_W}. \quad (1.9)$$

In order to write the electroweak Lagrangian for fermions such that it respects the $SU(2)_L \times U(1)_Y$ local gauge symmetry, one defines the covariant derivative

$$D_\mu = \partial_\mu + ig_W T^\alpha W_\mu^\alpha + ig' \frac{1}{2} B_\mu, \quad (1.10)$$

where W_μ^α and B_μ are gauge boson fields and $T^\alpha = \frac{1}{2} \sigma^\alpha$ are the generators of the $SU(2)$ symmetry.

The fermion fields can be represented as vectors containing the weak isospin doublets. These fields have different forms for left- and right-handed particles. For leptons,

we write:

$$l_{L1} = \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \quad l_{L2} = \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \quad l_{L3} = \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \quad (1.11)$$

$$l_{R1} = e_R \quad l_{R2} = \mu_R \quad l_{R3} = \tau_R. \quad (1.12)$$

Applying the covariant derivative defined in equation 1.10 yields

$$D_\mu l_{LN} = \left(\partial_\mu + ig_W \frac{1}{2} W_\mu^\alpha \sigma^\alpha - ig' \frac{1}{2} B_\mu \right) l_{LN} \quad (1.13)$$

$$D_\mu l_{RN} = \left(\partial_\mu - ig' \frac{1}{2} B_\mu \right) l_{RN} \quad (1.14)$$

where $N = 1, 2, 3$. The chiral nature of the weak interaction is evident from these equations, as the $W_{\mu\alpha}^\alpha$ field only couples to the left-handed leptons. The Lagrangian term representing the electroweak interactions of leptons is thus

$$\mathcal{L}_{EW,leptons} = \bar{l}_{LN} i \gamma^\mu D_\mu l_{LN} + \bar{l}_{RN} i \gamma^\mu D_\mu l_{RN}, \quad (1.15)$$

where the barred fields denote the adjoint fields $\bar{\psi} = \psi^\dagger \gamma^0$, and γ^μ are the Dirac gamma matrices.

Quark fields can be represented analogously, with two notable differences. Firstly, unlike for leptons, there are right-handed representations for all particles. Secondly, the quark fields of the $SU(2)$ symmetry are not the same as the mass eigenstates. One can write the left- and right-handed quark fields as

$$q'_{L,R1} = \begin{pmatrix} u' \\ d' \end{pmatrix}_{L,R} \quad q'_{L,R2} = \begin{pmatrix} c' \\ s' \end{pmatrix}_{L,R} \quad q'_{L,R3} = \begin{pmatrix} t' \\ b' \end{pmatrix}_{L,R}, \quad (1.16)$$

where the primes denote $SU(2)$ eigenstates, rather than the mass eigenstates. The quark electroweak Lagrangian term then becomes

$$\mathcal{L}_{EW,quarks} = \bar{q}'_{LN} i \gamma^\mu D_\mu q'_{LN} + \bar{q}'_{RN} i \gamma^\mu D_\mu q'_{RN}. \quad (1.17)$$

From equation 1.17, one can derive Lagrangian terms for the interactions of quarks with the gauge boson fields. For the weak charged current interaction this becomes

$$\mathcal{L}_{EW,quark}^{CC} = -\frac{g_W}{2\sqrt{2}} \left(J^\mu W_\mu + J^{\mu\dagger} W_\mu^\dagger \right), \quad (1.18)$$

where J^μ is the quark current density

$$J^\mu = 2\bar{q}'_{LN} \gamma^\mu q'_{LN} \quad (1.19)$$

$$= \bar{q}_N \gamma^\mu (1 - \gamma^5) V q_N, \quad (1.20)$$

where V is the *Cabbibo-Kobayashi-Maskawa* (CKM) matrix. The factor $\frac{1}{2}(1 - \gamma^5)$ is called the *left-handed chiral projection operator* and is used to project out the left-handed quark spinors q_{LN} , which the charged-current interaction couples to, from a general quark spinor q_N . The effect of the CKM matrix is a mixing of the quark flavours. Its elements encode the mixing probabilities with which any of the three up-type quarks can transform into any of the three down-type quarks, and vice-versa, through interaction with a W boson.

An equivalent mixing of flavours exists also for the leptons. A lepton interacting with a W boson is transformed into a neutrino of the corresponding flavour. However, the flavour eigenstates of the neutrinos do not correspond to the mass eigenstates. This leads to a mixing between the neutrino flavours with probabilities according to the *Pontecorvo-Maki-Nakagawa-Sakata* (PMNS) matrix, in analogy with the CKM matrix for quarks. A consequence of the mixing of neutrino flavours is the possibility of *neutrino oscillations*. Neutrinos that are produced with a given mass are in a superposition of flavour eigenstates. Therefore they will oscillate over time between the different flavours with amplitudes given by the entries of the PMNS matrix and periods given by the differences in mass between the mass eigenstates Δm . Fortunately, the oscillations occur on a scale such that they can be detected in experiment. The theory of neutrino oscillations within

the Standard Model is able to explain why the distribution of flavours observed in solar neutrinos on Earth is different from the distribution predicted by the nuclear interactions that take place inside the Sun – a well known phenomenon called the *solar neutrino problem*. The existence of neutrino oscillations was conclusively confirmed by the Super-Kamiokande [7] and SNO [8] experiments around the year 2000 and measurements of the oscillation lengths and mixing probabilities are the subject of many contemporary neutrino experiments.

1.1.3 The Higgs sector and spontaneous symmetry breaking

The theory of electroweak unification leaves a few questions unanswered. It cannot explain, on its own, the existence of the two massive vector bosons of the Weak interaction. The Higgs mechanism [9] achieves this through the addition of a scalar field and a process called *spontaneous symmetry breaking*, which gives rise to a new particle – the Higgs boson.

The Higgs sector of the Standard Model Lagrangian can be written as

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^2 + \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2, \quad (1.21)$$

where ϕ represents two complex scalar fields placed in a weak isospin doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (1.22)$$

The terms in the Lagrangian $\mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2$, which represent the potential energy of ϕ , are minimised when

$$\phi^\dagger \phi = -\frac{\mu^2}{2\lambda}. \quad (1.23)$$

The shape of the potential, which has famously been dubbed the "Mexican hat", has infinitely many possible ground states along the complex 4-dimensional hypersphere described by equation 1.23. By choosing a ground state on this hypersphere, the invariance of the Lagrangian under the $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken. By defining $v^2 = -\mu^2/\lambda$ and choosing a gauge, the field can be rewritten as

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+h \end{pmatrix}, \quad (1.24)$$

where h is the Higgs field. After application of the covariant derivative (equation 1.10) to the new scalar field and gathering all terms that are quadratic in the gauge boson fields, the Lagrangian contains the terms

$$\frac{1}{8}v^2 g_W^2 \left(W_\mu^{(1)} W^{(1)\mu} + W_\mu^{(2)} W^{(2)\mu} \right) + \frac{1}{8}v^2 \left(g_W W_\mu^{(3)} + g' B_\mu \right) \left(g_W W^{(3)\mu} + g' B^\mu \right). \quad (1.25)$$

Remembering the definitions of the vector boson fields in equations 1.6-1.8, one can deduce the masses of the gauge bosons from the terms in equation 1.25. Specifically, one can compare each term to the form of a mass term $\frac{1}{2}m_\Phi^2 \Phi^\mu \Phi_\mu$ that one would expect to obtain for a generic field. The gauge boson masses are related to the couplings g_W , g' and Higgs vacuum expectation value v through:

$$m_A = 0 \quad (1.26)$$

$$m_Z = \frac{1}{2} \sqrt{g_W^2 + g'^2} \quad (1.27)$$

$$m_W = \frac{1}{2} g_W v. \quad (1.28)$$

The Lagrangian also contains a mass term $\lambda v^2 h^2$ for the Higgs field, which gives the Higgs mass

$$m_H = \sqrt{2\lambda} v. \quad (1.29)$$

As the equations above show, the addition of a massive Higgs boson to the Standard Model provides an explanation for the existence of massive vector bosons, and relates their masses to their coupling to the Higgs field. Similarly, by introducing fermion fields into the Higgs Lagrangian it is possible to show that the Higgs field also interacts with fermions.

Since the right- and left-handed fermion fields transform differently under $SU(2)_L \times U(1)_Y$ symmetry, there is no way to write a mass term of the form $\bar{f}_L f_R$ consistent with electroweak theory. The mass terms instead take the form of the gauge invariant *Yukawa coupling*. For leptons, the Yukawa term in the Lagrangian is

$$\mathcal{L}_{leptons,\phi} = -\sqrt{2} \left(\bar{l}_{Li} y_{ij} \phi l_{Rj} + \bar{l}_{Rj} y_{ij}^* \phi^\dagger l_{Li} \right), \quad (1.30)$$

where y_{ij} are 3×3 matrices in lepton family space and l_i^L, l_i^R are the left- and right-handed lepton isospin doublets (equations 1.11 and 1.12).

For quarks, the Yukawa term is

$$\mathcal{L}_{quarks,\phi} = -\sqrt{2} \left(\bar{q}_{Li} y_{ij}^- \phi q_{Rj} + \bar{q}_{Rj} y_{ij}^+ \phi^c q_{Li} \right), \quad (1.31)$$

where $y_{ij}^\pm$ are 3×3 complex matrices in quark family space and $\phi^c = i\sigma_2 \phi$ is the conjugate scalar field.

After spontaneous symmetry breaking, the scalar doublet ϕ in equations 1.30 and 1.31 is replaced by the fields in equation 1.24. By gathering all the terms that contain the fermion fields and the Higgs field, one obtains the strength of the Yukawa coupling of the Higgs field to fermions. The coupling strength for each fermion f , which can be obtained by diagonalising the matrices y_{ij} , is

$$g_{hf} = \sqrt{2} \frac{m_f}{v}, \quad (1.32)$$

where m_f is the fermion mass.

Just as for the vector bosons, the Higgs couples to fermions with a strength proportional to the particle masses. As a consequence of this, the Higgs boson is more likely to decay to heavy particles. Most commonly, the Higgs is observed decaying to a $b\bar{b}$ pair. This is the most massive fermion anti-fermion pair that can be produced from a decaying Higgs in a tree-level Feynman diagram. The $H \rightarrow b\bar{b}$ decay was first measured at ATLAS in 2018 [10].

1.1.4 The QCD sector

The QCD sector of the Standard Model Lagrangian describes the strong interaction. The corresponding symmetry group is $SU(3)$ and its associated conserved quantum number is called *colour charge*. The gauge bosons of the strong interaction are called gluons. The gluons themselves carry colour charge, meaning that (self-)interactions of gluon fields are possible in addition to interactions between gluon and quark fields. The QCD part of the Lagrangian can be written as

$$\mathcal{L}_{QCD} = -\frac{1}{4}F_a^{\mu\nu}F_{\mu\nu a} - \sum_{f=1}^{n_f} \bar{q}_f (i\gamma^\mu D_\mu - m_q) q_f, \quad (1.33)$$

where q_f is the quark field, m_q the quark mass and f denotes the quark flavour. The QCD coupling strength is identical for all quark flavours. The field strength tensor of the strong interaction is

$$F_{\mu\nu a} = \partial_\mu G_{\nu a} - \partial_\nu G_{\mu a} - g_s f_{abc} G_{\mu b} G_{\nu c} \quad (1.34)$$

and the covariant derivative

$$D_\mu = \partial_\mu + ig_s G_{\mu a} T_a \quad (1.35)$$

where G_μ is the gluon field, f_{abc} are the structure constants of $SU(3)$ and g_s is the strong coupling constant. The tensors $T_a = \frac{1}{2}\lambda_a$ are the generators of the $SU(3)$ symmetry group, where λ_a are the *Gell-Mann matrices*. Since there are 8 generators of $SU(3)$, the index a runs from 1 to 8 and so there are 8 different colour states that can be carried by a gluon.

A particular feature of QCD concerns the running coupling constant. Running coupling constants, which change with the momentum scale μ , are a feature of all renormalisable field theories. However, unlike in QED for example, the strong coupling constant increases as μ decreases. This has important consequences for the validity of perturbation theory, as higher order corrections can no longer be assumed to vanish.

The coupling strength's dependence on momentum transfer μ is quantified by the β function

$$\beta(g_s) = \mu \frac{dg_s}{d\mu}. \quad (1.36)$$

In analogy with the fine-structure constant in QED it is common to define $\alpha_s = g_s^2/4\pi$. The dependence of α_s on μ^2 can be determined by writing the power series

$$\frac{d\alpha_s}{d\ln(\mu^2)} = -\beta_0 \left(\frac{\alpha_s}{4\pi} \right) - \left(\frac{\alpha_s}{4\pi} \right) \sum_n \left(\frac{\alpha_s}{4\pi} \right)^n \beta_n, \quad (1.37)$$

where β_0 are the power series terms of the beta function. The first term can be evaluated as

$$\beta_0 = \frac{11}{3}n_c - \frac{2}{3}n_f \quad (1.38)$$

where $n_c = 3$ is the number of colours and n_f is the number of quark flavours. By discarding the $n > 0$ terms in equation 1.37 and integrating from μ^2 to Λ_{QCD} , one obtains the expression

$$\alpha_s(\mu^2) = \frac{4\pi}{\beta_0 \ln(\mu^2/\Lambda_{QCD}^2)}, \quad (1.39)$$

where Λ_{QCD} is the *QCD energy scale*. One notices that $\alpha_s > 1$ for values of $\ln(\mu^2/\Lambda_{QCD}^2) \lesssim 2$. Below this value, the sum in equation 1.37 is no longer finite and perturbation theory is no longer useful. Thus, the value of Λ_{QCD} can be said to divide strong processes into a *soft* (small μ^2) and a *hard* (large μ^2) regime. On the other hand, for $\mu \gg \Lambda_{QCD}$ the coupling is weak and strong processes may be calculated in perturbation theory with great accuracy. This phenomenon, called asymptotic freedom, enables physicists to test the theory of QCD in great detail and justifies the postulation of the existence of quarks and gluons, which can never be observed in isolation. The value of the strong coupling constant at a given mass scale has been measured experimentally in a multitude of different processes. It has been found that $\alpha_s(M_Z^2) = 0.118$ and hence $\Lambda_{QCD} \approx 90 \text{ MeV}$ for 6 quark flavours [5].

One can conclude from equation 1.39 that, at very short distances where $\mu^2 \gg 1 \text{ GeV}$ the coupling is weak and perturbative calculations are possible, whereas, at longer distances the coupling grows stronger and stronger. This provides an intuitive explanation

for *colour confinement*, the phenomenon that quarks and gluons can only exist as bound states within colour-neutral hadrons. Colour confinement is a consequence of the self-interaction of gluons. As the bonds between quarks are stretched, ever growing networks of gluon fields, called flux tubes, are created, giving rise to an increasing $q - q$ potential.

An additional complication in predicting QCD processes are *soft and collinear divergences*. Perturbation theory predictions of QCD differential cross-sections usually involve factors of the form

$$\frac{d\theta_p}{\theta_p} \frac{dE_p}{E_p} \quad (1.40)$$

where θ_p , E_p are the energy and angular separation of a quark or gluon emission. Simply integrating over these factors to get the total cross-section will yield a non-physical infinite result. Fortunately, these singularities can be removed by temporarily introducing a *regularisation* term to make the integral finite. However, the problem of soft and collinear divergences reappears when QCD theorists attempt to make predictions of final states as they are measured in detector experiments. Experimental cuts have the effect of removing parts of the integral over $d\theta, dE$, which can lead to some divergences not being removed by the regularisation. The cross-section gains a series of *logarithmic enhancements* of the form $(\alpha_s L^2)^n$, where L is a large logarithm of the form $L = \ln(Q^2/M^2)$ or similar. Some logarithmic enhancements can be accounted for through *resummation*, which describes the procedure of analytically summing terms with the largest logarithmic coefficients to all powers of α_s . Resummation problems are highly nontrivial and have to be solved separately for every observable. A resummation of all terms with $\alpha_s^n L^{2n}$ is referred to as *leading-logarithmic* (LL), while one that also includes $\alpha_s^n L^{2n-1}$ is called *next-to-leading-logarithmic* (NLL). The precision of a QCD prediction is frequently expressed as the order of perturbation theory used in the calculation of the matrix element plus a resummation order, e.g. NLO+NLL [11].

1.1.5 Beyond the Standard Model

The Standard Model, despite its immensely powerful ability to accurately predict the physics of the known elementary particles, is not a theory of everything. There are phenomena, such as the existence of dark matter, whose theoretical explanations on the level

of elementary particles are beyond the reach of the Standard Model. The physics of these phenomena is called *Beyond Standard Model*, or BSM, physics. One extension to the Standard Model that provides possible explanations for BSM physics is *Supersymmetry* (SUSY) [12–14], whose predicted heavy super-partner particles are may be candidates for dark matter WIMPs (weakly interacting massive particles). SUSY is a potential solution to many long-standing problems in Standard Model physics, especially the *hierarchy problem*: the issue that the SM cannot be extended to very high energies without giving rise to large corrections to the Higgs boson mass that only cancel if this free parameters are precisely tuned. The possibility that supersymmetric particles may exist at the TeV scale has made supersymmetry the most searched for BSM theory at the LHC. So far, however, LHC experiments have found no evidence of SUSY particles and continue to exclude them at ever increasing energies [15, 16].

The ATLAS and CMS detector experiments at the LHC have proved very effective at excluding BSM physics. For example, the constraints produced by searches for Dark Matter WIMPs are strong enough to complement the results from direct detection experiments [17, 18]. Other BSM theories that have recently been investigated ATLAS and CMS searches include: vector-like quarks [19–22], leptoquarks [23–26], as well as quantum black holes and additional spatial dimensions [27–29].

1.1.6 Physics of proton-proton collisions

Proton-proton collisions are extremely complex processes. In collision experiments with very large centre-of-mass energies, protons cannot merely be considered bound states of one down-type and two up-type quarks. The kinematics of virtual gluon emissions and quark-antiquark pairs, known as sea quarks, must also be considered. In a typical $p - p$ collision, hundreds of particles are produced, and strong processes occur across a wide range of energy scales, spanning both the perturbative and non-perturbative regimes of QCD. Fortunately, the probability of perturbative processes can be calculated to great accuracy without needing to consider non-perturbative effects. The total cross-section of a $p - p$ collision can be written as

$$\sigma_{pp} = \sum_{a,b} \int_0^1 \int d\Phi f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \frac{1}{2x_a x_b s} |M_{a,b \rightarrow n}|^2(\Phi_n, \mu_F, \mu_R), \quad (1.41)$$

where a, b denote partons within the initial-state protons h_1, h_2 , n denotes the final-state particles and the differential $d\Phi_n$ denotes the phase space element over the outgoing particles. The parameter μ_F is an energy called the factorisation scale which has been fixed, in analogy with the renormalisation scale μ_r , to remove divergences due to collinear gluon emissions. The functions $f(x, \mu_F)$ are called the *parton distribution functions* (PDFs). They represent the distributions of the probability of partons a and b , within hadrons h_1 and h_2 , participating in a hard scatter process as a function of momentum fraction x . Knowledge of the PDFs is crucial to modelling the physics of $p - p$ collision events that are studied using LHC detector experiments such as ATLAS. The shapes of the PDFs are independent of the processes in which the partons participate, and so they can be fitted using data from many different experiments. This postulate is a well-known theorem in QCD called the collinear factorisation theorem [30].

During the course of a $p - p$ collision event, there are additional partons being emitted at energies ranging from the scale of the proton mass up to the scale of the hard parton-parton interaction. In classical electromagnetism, charged particle that scatter in the electromagnetic field of another charge emits *bremssstrahlung* as it decelerates. Analogously, the quarks which interact in a hard-scatter process emit gluons before and after the scattering takes places. This radiation is called *initial state radiation* (ISR) when it is emitted by incoming partons. Similarly, there is *final state radiation* (FSR) in processes that produce outgoing partons. Like the outgoing partons of the hard-scatter, gluons due to ISR and FSR produce cascades further emissions as parton showers (see section 1.1.7).

Spectator partons, which do not participate in the hard scatter process, also interact during the collision. This long-distance, i.e. low momentum transfer, interaction produces soft radiation which is referred to as the *underlying event* (UE). The underlying event typically has much lower transverse momentum than the decay products of the hard-scatter process and does not constitute a dominant background effect when such processes are measured. However, in jet substructure, where the distribution of radiation

inside hadronic jets is studied, UE can have a significant impact. The model which has been outlined above, in which only one parton from each of the colliding protons takes part in a hard scatter process, is a simplification. Highly energetic processes involving multiple partons are called *multi-parton interactions* (MPI). At the centre-of-mass energy of the LHC, accurate modelling of the underlying event requires that contributions from MPI are also considered [31].

An important contamination effect, which is not a consequence of interactions of the constituents of the proton but of the fact that protons in real-world accelerator experiments always collide in bunches, is *pileup*. Pileup refers to the contributions of, mostly low-energy, radiation from simultaneous $p - p$ collisions that are picked up by the detector along with the high momentum-transfer process that is of interest to particle physicists. At the end of the second run of the LHC, where the instantaneous luminosity was $L \approx 2 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ and the bunch spacing 25 ns, the number of protons in each bunch was around 10^{11} . The average number of simultaneous collisions, denoted $\langle \mu \rangle$, peaked at around 40.

1.1.7 Parton showers and Monte Carlo event generators

To investigate what happens to final state quarks and gluons that are produced in high energy $p - p$ collisions, it is instructive to consider an emission of a parton from another parton at an angle θ . The emission process could be either a gluon radiating from another parton or a quark-antiquark pair being produced by a gluon. From perturbation theory calculations, it can be shown that the splitting probability is given by

$$dP_{a \rightarrow bc}(z, \theta) = \frac{d\theta^2}{\theta^2} dz P_{ba}(z) \frac{\alpha_s(q^2)}{2\pi}, \quad (1.42)$$

where a denotes the initial parton, b, c denote the final partons. The fraction of energy carried by parton b is denoted $z = E_b/E_a$. The function $P_{ba}(z)$ is called the *Altarelli-Parisi splitting function* and it has a different form depending on the types of partons, gluons or quarks, in the final state. The splitting probability is greater for small angles and goes to infinity for $\theta \rightarrow 0$. This is the collinear divergence property of parton splittings³, already mentioned in section 1.1.4. Collinear divergence is closely linked with

the collinear safety requirement in the reconstruction of hadronic final states as jets (see section 1.4). Starting from the highly-energetic parton, the preference for emissions at small angles causes a cascade of collinear parton emissions as shown in figure 1.1. With each subsequent splitting, the energy of the initial parton is shared between the growing number of showering partons. It can be shown (see e.g. ref. [11] for a derivation), from the calculation of the splitting functions for soft gluon emissions, that the emission angle becomes smaller with each splitting. This important principle is known as *angular ordering* and is central to models of parton showers utilizing the Altarelli-Parisi splitting functions. The shower terminates when the partons' energies and separations in space are small enough that they begin to form hadrons. The kinematics of the spray of final hadrons, which can be reconstructed in a detector as a *jet*, reflect the flow of energy through the parton shower that generated it [11].

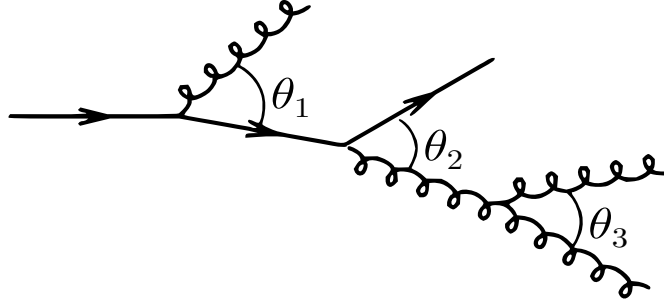


Figure 1.1: Diagram of showering partons. Each emission carries off a fraction z of the initial parton's energy. Emissions take place until the parton energies reach the hadronisation scale. Subsequent emissions at small angles θ obey the relationship $\theta_1 > \theta_2 > \theta_3$. This property is known as angular ordering.

The parton shower and the subsequent hadronisation of the partons involve exchanges of many gluons with low momentum transfers below Λ_{QCD} . It is therefore not possible to accurately predict these processes in fixed-order perturbation theory, like the hard-scatter process. However, they can be simulated using probabilistic models or *Monte Carlo* (MC) event generators. Monte Carlo parton shower algorithms simulate the evolution of a shower from a single hard parton through a series of random splittings on a hardness scale, e.g. the momentum transfer q^2 , until a cut-off Q_0^2 is reached. It is easiest

³This description of the collinear divergence is only strictly true in the massless parton limit. For massive quarks there is actually a suppression of radiation in a cone $\theta_d \lesssim m_q/E_q$ in the direction of the quark momentum. A phenomenon known as the *Dead Cone* effect [32, 33].

to express a parton's evolution through different q in terms of the exponentially decaying probability that it does not generate a splitting between the values Q_1 and Q_2

$$\Delta_a(Q_1^2, Q_2^2) = \exp \left[- \int_{Q_1^2}^{Q_2^2} \frac{dq^2}{q^2} \frac{\alpha(q^2)}{2\pi} \int_{1-Q_1^2/q^2}^{1-Q_2^2/q^2} dz P_{ab}(z) \right], \quad (1.43)$$

where $P_{ij}(z)$ is the splitting function for parton types a and b , like we have already seen in equation (1.42). The quantity $\Delta_a(Q_1^2, Q_2^2)$, called the Sudakov form factor, is the key to generating a numerical parton shower. By generating a random number R , uniformly distributed between 0 and 1, and solving the equation

$$\Delta_a(Q_1^2, Q_2^2) = R \quad (1.44)$$

for Q_2 one obtains the q^2 scale of parton after the splitting. To generate a shower, the process is repeated for every parton until the termination criterion $Q_i^2 < Q_0^2$ is reached. Parton shower simulations are astonishingly powerful predictive tools. Provided that the splitting scale is proportional to the splitting angle θ (in the equations above $q^2 \simeq z(1-z)\theta^2 E^2$ in the limit of small angles), the intrinsic angular ordering property of parton showers ensures that soft gluon emissions are captured by the model. It can be shown that, for global jet observables, an angular ordered MC parton shower, as described above, performs the equivalent of an all-order resummation of the leading logarithmic enhancements to perturbative QCD predictions.

Parton shower models of the angular-ordered kind as introduced above are not the only possible description of parton showers. It is possible to construct alternative models that exhibit different ordering properties. One such example are dipole shower models in which $q\bar{q}$ pairs are emitted at each step. When using this picture, it is usually more convenient to order emissions by transverse momentum p_T or a related . The angular-ordered and dipole shower approaches each have their own advantages and drawbacks. Angular ordered parton showers automatically take soft gluon emissions into consideration, which ensures that collinear divergences, for global observables like the number of jets in an event, are correctly modelled up to leading logarithmic order. This is not necessarily true of dipole showers. However, dipole showers have an advantage over

angular ordered showers in that they are more straightforward to match to the hard subprocess, as their evolution starts from the hardest emissions by construction. In practice, both angular-ordered and dipole shower models can incorporate modifications that help overcome some of their respective disadvantages and bring them to an equivalent level of accuracy. Today's state-of-the-art event generators incorporate both angular-ordered and dipole shower approaches. Herwig [34, 35] uses an angular-ordered parton shower algorithm by default, but can optionally be configured to use a dipole shower. Sherpa [36] and Pythia [37, 38] both use dipole-based approaches to shower modelling [39].

Parton shower simulations on their own are, however, not sufficient to make testable predictions. They need to be connected to a *matrix element generator* that calculates the probability of the hard scatter process according to equation (1.41) and a *hadronisation model* that describes partons at Q_0^2 combining to form hadrons. *General purpose Monte Carlo generators* (GPMC) which are used to make predictions for collider experiments need to perform all of these three simulation steps. In order to accurately simulate an entire $p - p$ collision event, as described in section 1.1.6, the effects of ISR, FSR, UE, pileup etc. must also be modelled using MC generators. This is done using the same components of the GPMC that are used to model the evolution of the final states of the hard scatter process [39].

Thanks to the factorisation theorem, the matrix element (ME) of the hard subprocess can be calculated at the parton level to obtain the production cross-section in $p - p$ collisions according to equation (1.41). In a Monte Carlo event generator, this task is performed by a matrix element generator. Matrix element generators usually include a library of pre-computed MEs from Leading-order perturbation theory for the most common $2 \rightarrow 1$, $2 \rightarrow 2$ and $2 \rightarrow 3$ Standard Model processes. For processes with greater particle multiplicities, where the phase space to be integrated is much larger, the ME is evaluated using MC integration. Next-to-leading order matrix elements, or K -factors which are used to scale LO cross-sections to match NLO corrections, are also available for some processes. In addition to the MEs themselves, ME generators also need to specify a choice of the renormalisation and factorisation scales μ_R and μ_F (which appear in the $p - p$ cross-section equation (1.41)), as well as a set of fitted PDFs. Examples of com-

monly used ME generators include MadGraph [40], Comix [41] and Amegic [42]. ME generator predictions also need to be matched with the subsequent parton showers, which is particularly important for NLO matrix elements where logarithmic enhancements play an important role. This is done using programs like Powheg [43] and MC@NLO [44, 45]. The MC@NLO procedure has also been integrated into the Madgraph event generator package as MadGraph5_aMC@NLO [46].

As hadronisation is a highly non-perturbative process, its theoretical description cannot be derived from first principles. Therefore, there exists a number of different approaches to stochastically model hadronisation. One example is the Lund string model, used in e.g. Pythia [37, 38], which assumes that partons at the hadronisation scale are enveloped in a gluon flux tube called the string. The high concentration of energy on the string causes quark-antiquark pairs to be produced which combine with the final state quarks to form colour-neutral hadrons. Cluster hadronisation models are an alternative class of hadronisation models in which colour neutral parton clusters are formed before splitting into hadrons. One event generator that uses such pre-confined clusters to model hadron formation is Herwig [34, 35]. Sherpa [47] also uses a cluster-based hadronisation model, Ahadic, by default but can also be configured to use Lund strings.

1.2 The ATLAS experiment at the LHC

1.2.1 The Large Hadron Collider

The *Large Hadron Collider* (LHC) is a 27 km long synchrotron located at CERN in Geneva, Switzerland. It is the World's largest particle accelerator of its kind and is able to accelerate protons and heavy ions to energies that have never before been achieved. The two beams of the LHC are made to collide at four points on the accelerator ring. Detector experiments that probe the physics of the collisions are located at each of these four points. Two of the detectors, *ATLAS* and *CMS*, are general-purpose detectors, while *ALICE* specialises on heavy-ion collisions and *LHCb* is built to study the properties of *b*-hadrons. Figure 1.2 shows an overview of the CERN accelerator complex, the LHC and the locations of the four detector experiments.

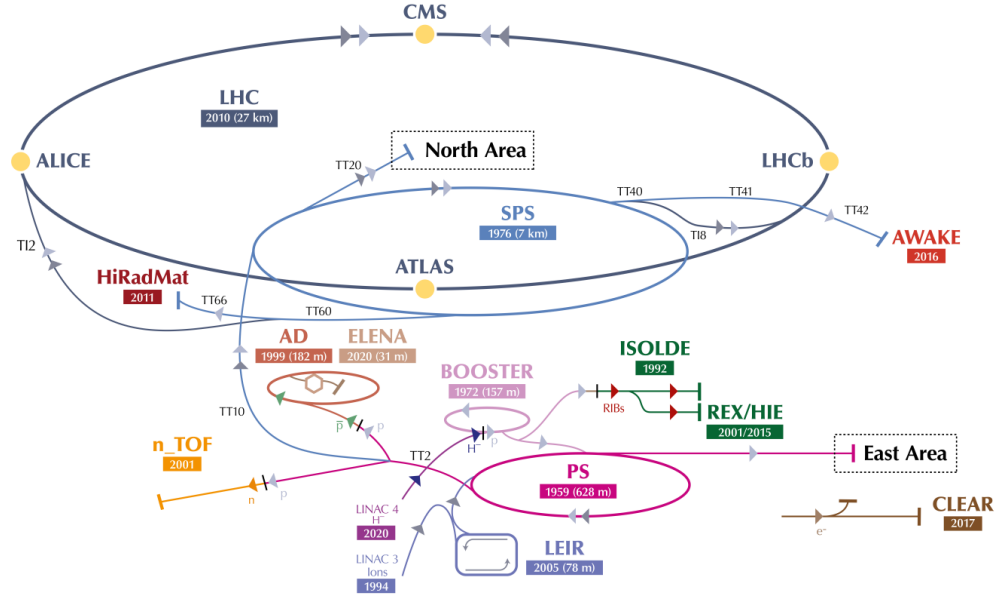


Figure 1.2: An overview of the CERN accelerator complex including the Large Hadron Collider and the accelerators that feed into it. From ref. [52].

The LHC first began taking proton-proton collision data in 2010. During the initial period of operations, known as *Run 1*, the LHC was producing collisions at $\sqrt{s} = 7 \text{ GeV}$ and, from 2012 at 8 GeV . From 2013 until 2015, the first *Long Shutdown* took place to prepare the LHC for higher collision energies and luminosities.

During *Run 2*, from summer 2015 until winter 2018, the centre-of-mass energy was increased to 13 TeV and an unprecedented peak instantaneous luminosity of $1.9 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ was reached. As a consequence, the LHC experiments were able to collect their largest datasets to-date. In the case of the ATLAS detector, the total integrated luminosity delivered was 139 fb^{-1} compared to only 22.7 fb^{-1} in Run 1 [48, 49].

From 2019 to spring 2022, the LHC entered its second long shutdown in preparation for its third period of data-taking. During early Run 3 in summer 2022, the LHC began to produce its first collisions at $\sqrt{s} = 13.6 \text{ TeV}$. Throughout Run 3, which is expected to last until 2026, the LHC is expected to deliver more collisions than in all previous runs. After the end of Run 3, the LHC will undergo further upgrades, called the *High-Luminosity-LHC* (HL-LHC). The HL-LHC is expected to operate at a five times greater luminosity than in Run 2 [50, 51].

1.2.2 Overview of the ATLAS detector

A Toroidal LHC ApparatuS – commonly referred to as *the ATLAS detector* – is a general purpose particle detector located approximately 100m below ground at Point 1 on the Large Hadron Collider. It covers close to a 4π solid angle around the proton-proton collisions and is designed to be hermetic, that is to say no particles (with the exception of neutrinos) are able to escape it. The detector is about 44 metres long and 25 metres in diameter and weighs about 7000 tonnes. The ATLAS detector can be roughly divided into three main sub-detector regions. From the interaction point and outwards, these regions are: the inner detector (ID), the calorimetry and the muon spectrometer. Surrounding the inner detector is a solenoid providing a 2T magnetic field. A toroid magnet, consisting of 8 radially symmetrically spaced elements in the detector barrel and two end caps, surrounds the calorimeters. A sophisticated trigger system filters out events that are relevant for physical studies and reduces the rate of data collection from around 1GHz – the rate of proton-proton collisions – to around 1 kHz [53].

The coordinate system used at ATLAS is a right-handed coordinate system with the origin located at the expected *interaction point* (IP) in the centre of the detector. The x -axis points from the IP to the centre of the LHC ring, the y -axis points upward and the z -axis points along the direction of the beam pipe. Cylindrical coordinates (r, ϕ) are used in the $x - y$ plane, where ϕ is the azimuthal angle around the beam pipe. The *pseudorapidity* is defined in terms of the polar angle θ with respect to the z -axis as

$$\eta = -\ln(\tan \theta/2). \quad (1.45)$$

The pseudorapidity is equivalent to the rapidity $y = \frac{1}{2} \ln \left(\frac{E+p_T}{E-p_T} \right)$ in the limit of massless particles. Differences in rapidity are invariant under a Lorentz boost. Since the sum of particle momenta in the $x - y$ plane is known to be 0, it is useful to define *transverse momentum* as $p_T = p \sin \theta$ and *transverse energy* $E_T = E \sin \theta$.

1.2.3 The Inner Detector

The Inner Detector (ID) is designed to capture the tracks of charged particles by providing high-precision position and momentum measurement. Accurate tracking is crucial to

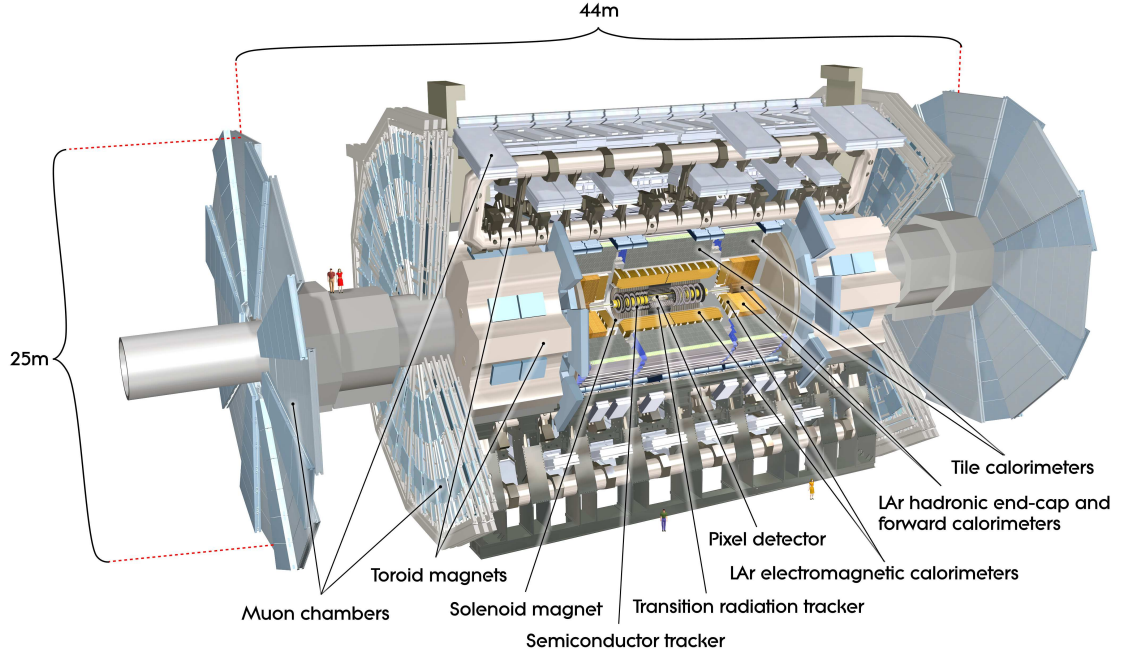


Figure 1.3: An overview of the ATLAS detector. From ref. [53].

particle identification and jet flavour tagging. The ID extends from the LHC beam pipe to about $R = 1.1$ m and is 6.2m in length. It consists of three main sub-detectors: the *Pixel* tracker, the silicon-microstrip *Semiconductor Tracker* (SCT), and the straw-tube *Transition Radiation Tracker* (TRT). The pixel and SCT trackers together cover a pseudorapidity region of about $|\eta| < 2.5$. To enable measurements of particle momentum and charge, the inner detector is surrounded by a 2T solenoid magnet. An overview of the inner detector is shown in figure 1.4 [54].

The innermost sub-detector of the inner detector is the silicon *Pixel* tracker. It surrounds the beam pipe at about $R = 45.5$ mm from the interaction point and extends to $R = 242$ mm. It is 3092 mm in length. Despite being the smallest subsystem of the ATLAS detector, the Pixel tracker has over 80 million readout channels, which is more than every other ATLAS subsystem. It is built up of over 1700 identical modules consisting of $250\mu\text{m}$ thick n-type silicon wafers. In the barrel, the Pixel modules are connected to form staves which are arranged symmetrically about the beam axis. The staves in each layer are rotated 20deg with respect to the R and ϕ axes and partially overlap to improve spatial resolution (see figure 1.4b). Three circular discs on either side of the barrel provide coverage for the $2.0 < |\eta| < 2.5$ region. Approximately 47000 n^+ doped pixels are

implanted on each module. A reverse-bias voltage is applied across the pixels and the wafer, allowing ionisation currents caused by passing particle to be picked up. The size of each pixel is about $50\mu\text{m}$ by $400\mu\text{m}$ along the $(R - \phi)$ and z directions respectively [53].

Since the beginning of Run 2 the Pixel tracker has been augmented by the *insertable b-layer* (IBL), an additional layer of pixels at 31 – 40 mm from the beam axis. The installation of the IBL required replacing the old ATLAS beam pipe ($R_{\text{inner}} = 29, \text{mm}$) by a smaller one ($R_{\text{inner}} = 24.3, \text{mm}$). The beam-pipe is made of beryllium, which has a high transparency to particles as well as the mechanical stiffness required to maintain an ultra-high vacuum. The IBL is the first ATLAS sub-detector to make use of a novel type of 3D CMOS technology, in which electronic connections are run through the bulk of the sensor chip rather than on the surface. The purpose of the IBL has been to provide improved tracking performance until the end of Run 3, and to compensate for radiation damage sustained by the innermost layer (B-layer) of the Pixel tracker [55, 56].

The Semiconductor Tracker begins at $R = 299 \text{mm}$ from the interaction point and extends to around $R = 514 \text{mm}$. The 4000 submodules of the SCT are arranged concentrically around the Pixel tracker and are divided into a barrel and endcap section. The modules in the four layers of the barrel section are tilted at an angle of approximately 11 deg with respect to the ϕ -axis. The modules in the endcaps are arranged in 9 disk layers perpendicularly to the beam axis. On the modules are approximately 16000 identical sensors containing 6.5 million microscopic silicon sensing units called microstrips. The operating principle of the SCT microstrips is similar to that of the Pixels. As it is located further from the interaction point, the SCT has a lower spatial resolution than the Pixel tracker: approximately 2 times lower in the $R - \phi$ direction and 5 times lower in the z -direction [53].

Both the Pixel and SCT detectors are strongly affected by the high radiation dosage they receive over time, which effectively turns the n-type semi-conductor into a p-type. To counteract this, the detectors are kept at a temperature of -5 to -10degC and the operating voltage is slowly increased to compensate for the loss of signal.

The Transition Radiation Tracker begins at $R = 554$ mm and extends to the to around 1082 mm, just within the inner wall of the inner detector cryostat. The TRT consists of 4 mm radius proportional drift-tube sensors called straws which are interleaved with polypropylene fibres. The purpose of the drift tubes is to provide additional tracking information, while the polypropylene fibres enable the identification of electrons through transition radiation. The TRT is only able to measure the $R - \phi$ coordinates of the charged particle tracks. The approximately 290,000 TRT straws are tightly bundled around the SCT barrel and endcap sections and provide a spatial resolution of $130 \mu\text{m}$ [53, 57].

For the High-Luminosity LHC, the ATLAS inner detector is expected to be replaced with an all-silicon *Inner Tracker* (ITk) [58].

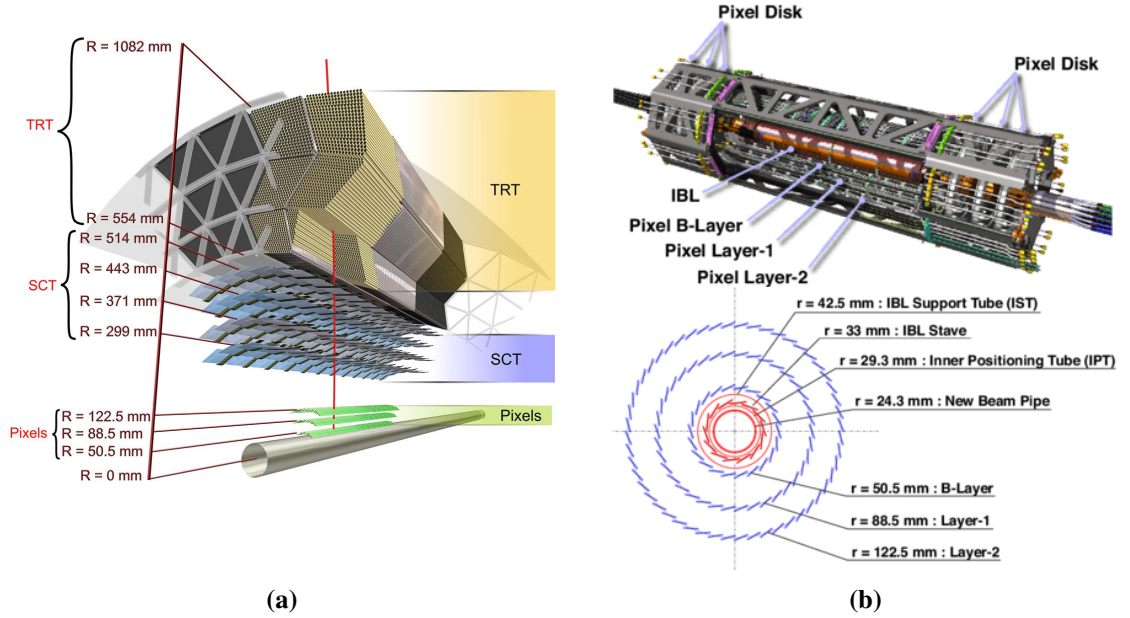


Figure 1.4: An overview of the ATLAS inner detector is shown in (a) [54]. A more detailed schematic and edge-on view of the pixel detector is shown in (b) [59].

1.2.4 The calorimeters

1.2.4.1 Liquid Argon Calorimeter

The liquid argon calorimeter consists of parallel plates of lead and steel absorbers which are submerged in liquid argon. The liquid argon acts as the sensitive medium of the calorimeter. The absorber plates are folded into an accordion shape, as can be seen in figure 1.6. The accordion absorber modules are separated by honeycomb-shaped resin-

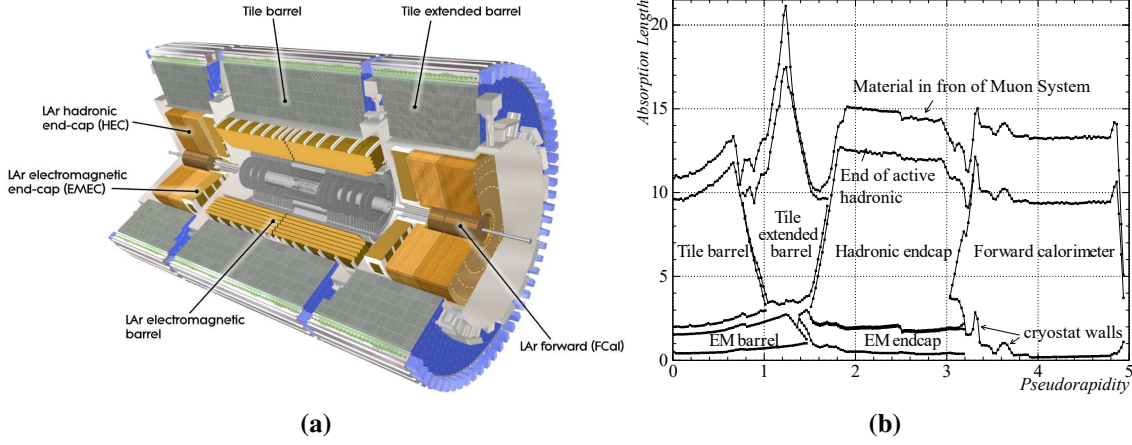


Figure 1.5: An overview of the ATLAS calorimetry is shown in (a) [60]. (b) shows a graph of the amount of material, measured in units of hadronic interaction lengths, provided by the ATLAS calorimetry across the full pseudorapidity range of the detector [61]. Also shown is the dead material in front of the calorimeters, including the walls of the liquid Argon cryostat, and in front of the Muon spectrometer.

impregnated paper spacers to provide room for the liquid argon and ensure a uniform material density throughout the calorimeter. The accordion modules are stacked around the beam axis, so that full coverage in ϕ is achieved.

Readout is performed using a grid of copper electrodes which is interwoven with the honeycomb spacers between the absorber plates [62]. Electrons and photons deposit their energy in the absorbers and the liquid argon by giving rise to electromagnetic showers. A cascade of bremsstrahlung photons and e^+e^- pairs is produced. Charges that are liberated by the shower in the liquid argon are collected on the electrode grid. This enables a signal to be picked up that is proportional to the energy deposited by the particle. The electrodes are separated by a mean distance of 2×2.1 mm, which corresponds to an average drift time of 450 ns when a 2 kV voltage is applied.

The liquid argon electromagnetic calorimeters are divided into a barrel and an end-cap section which cover pseudorapidity ranges of $|\eta| < 1.475$ and $1.375 < |\eta| < 3.2$ respectively. Inside the innermost layers of the barrel and end-cap calorimeters are pre-sampler layers covering a range of $|\eta| < 1.52$ in the barrel and $1.5 < |\eta| < 1.8$ in the endcaps [61]. The spatial resolution of the liquid argon calorimeter varies between each layer. The pre-sampler, which consists of a single 11 mm liquid-argon layer and ac-

counts for the energy lost in the dead material in front of the calorimeters, has an angular granularity of $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$. As can be seen in figure 1.6a, the granularity is approximately 0.025×0.025 in the second and third layer of the liquid argon barrel, while in the first layer this figure is reduced by a factor of 8 along η and increased by a factor of 4 along ϕ . The reason for this very fine segmentation along the η direction is to be able to resolve the two electromagnetic showers that are produced by a neutral pion which promptly decays into a pair of photons. The ability to accurately distinguish photons from neutral pions has been crucial for the discovery of the Higgs boson, for example in the $H \rightarrow \gamma\gamma$ channel.

The readout from the electrode grid is amplified and reshaped using analogue CR – RC² circuits on the Front End Boards (FEBs). The resulting output pulse shape is shown in figure 1.6b. The output of the FEBs is then sampled digitally in 25 ns intervals. This corresponds to the time between bunch crossings and is significantly shorter than the total drift time of 450 ns. Due to the long drift time, cell energies are affected by contributions of soft hadronic radiation from simultaneous bunch crossings, known as in-time pileup, as well as from preceding collisions, known as out-of-time pileup. The energy of a calorimeter cell at a given time is then reconstructed using the Optimal Filtering technique, which reduces the level of noise by assuming perfect knowledge of the signal pulse shape. A notable feature of the shape of the signal pulse, shown in figure 1.6b, is a long negative pedestal trailing a sharp peak. The negative pedestal causes cell energy contributions due to out-of-time pileup to be negative. Because the total time integral of the pulse is 0 and the contributions from pileup are constant over time, the time-averaged out-of-time pileup contributions cancel out the effects of the in-time pileup thus greatly reducing the effect of pileup on the measured signal [63, 60].

During the second long shutdown of the LHC, the liquid argon calorimeter was upgraded to provide higher resolution data to the level 1 triggers. The 0.1×0.1 trigger towers were replaced by smaller *SuperCells* in each layer of the calorimeter with granularities varying between 0.025×0.1 and 0.1×0.1 . This added considerable complexity to the information on which trigger decisions would be based. Therefore, the analogue output from the front-end boards is digitised and processed on the LAr Trigger Digitiser

Board (LTDB) and the LAr Digital Processing System (LDPS) before being fed to the level 1 trigger. The LDPS is constructed using FPGAs and performs a basic feature extraction algorithm to reconstruct shower energies and computes the missing transverse energy of each event [64, 65].

Substantial maintenance and upgrade work will be carried out on the ATLAS liquid argon calorimeters in preparation for the HL-LHC. Both the front-end boards and off-detector electronics will be upgraded. Although some degradation in detector performance is expected due to radiation damage to the organic materials in the calorimeter, the effect is sufficiently small that the inactive components of the LAr calorimeter will not have to be replaced during the third long shutdown [65].

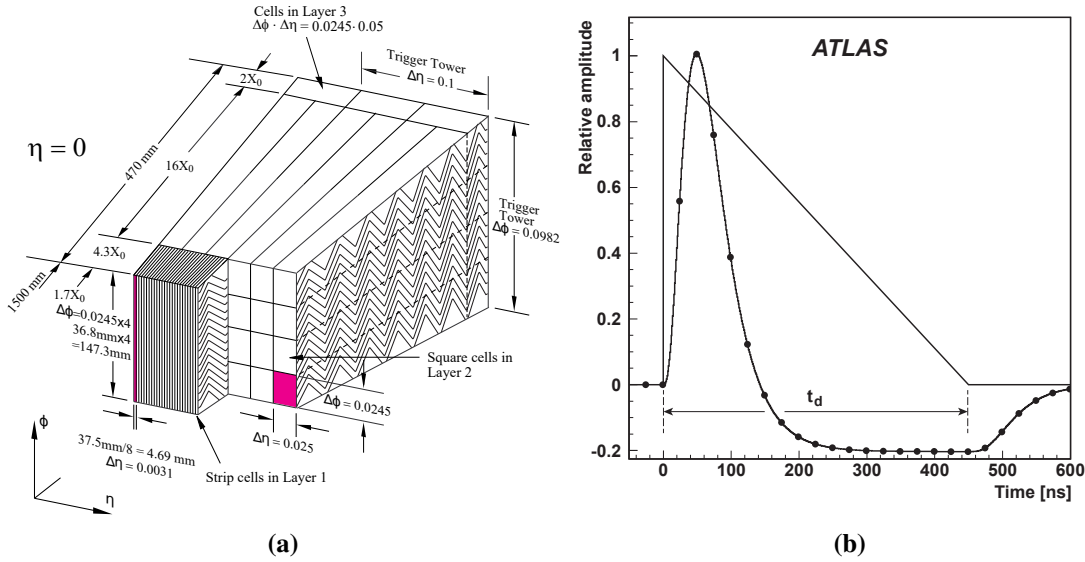


Figure 1.6: (a) Schematic showing the structure of the liquid argon calorimeter in the central barrel region. (b) Pulse shape of the signal in the ATLAS liquid argon calorimeter. Note the bipolar shape of the pulse and the height of the pedestal P , which in this case is around 0.2 amplitude units. From ref. [60].

1.2.4.2 Tile Calorimeter

The hadronic calorimetry comprises the Tile calorimeter as well as liquid argon hadronic endcaps. The Tile calorimeter (TileCal) is located at 2.28 m to 4.23 m from the beam line. It consists of approximately 460,000 solid iron absorber and plastic scintillator tiles, which makes it the heaviest subdetector of ATLAS at around 2900 tons [66]. The iron and scintillator tiles are arranged in a staggered configuration on top of iron master

plates which, in turn, are mounted on a structural girder on the outside of the calorimeter. Hadrons, which are able to traverse the liquid argon calorimeter, deposit their energy by giving rise to hadronic showers in the iron absorber plates. Ionising radiation produced by the hadronic showers causes scintillation to occur in the plastic tiles. Readout is achieved by feeding light produced in the scintillator tiles through wavelength-shifting optical fibres to photomultiplier tubes located inside the structural girders. A small space between the scintillator and steel tiles allows the scintillators to be calibrated individually through the insertion of a radioactive caesium source [66].

The Tile calorimeter is divided into three sections: one barrel and two extended barrel sections. The barrel covers a pseudorapidity range of approximately $|\eta| < 1$, while the extended barrels cover $0.8 < |\eta| < 1.7$. At around $|\eta| = 0.8$ there are gaps in the hadronic calorimeter to provide room for electronics, cables and other services to the LAr calorimeter and ID. Plastic scintillator detector tiles have been placed in the gaps to provide some coverage in these regions. Within the barrel and extended barrel region, the readout of the TileCal is divided along the radial axis into three sampling layers, which are further divided into cells along the $\eta - \phi$ plane. The resolution of cells in the first two layers is about $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, while in the third layer it is around $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$. A cross-section of the TileCal, showing its submodules and division into layers and cells is shown in 1.7a [66].

Substantial upgrade work will be carried out on the Tile calorimeter in preparation for the HL-LHC. All on- and off-detector electronics will be replaced to cope with the higher data rate. Additionally, about 10% of PMTs are going to be replaced as they have sustained a significant amount of radiation damage [67].

1.2.4.3 Hadronic Endcap Calorimeters

The hadronic endcap calorimeters (HEC) cover a pseudorapidity range $1.2 < |\eta| < 3.2$. It is located just outside the electromagnetic endcap and consists of two wheels of outer radius 2.03 m. The thickness of the inner and outer wheels are 0.82 m and 0.96 m respectively. Like the electromagnetic calorimeter, the HEC uses Liquid Argon as its active medium. The HEC consist of parallel plates of copper absorbers, electrodes and cardboard spacers. The thickness of the copper absorber plates is 25 mm in the inner

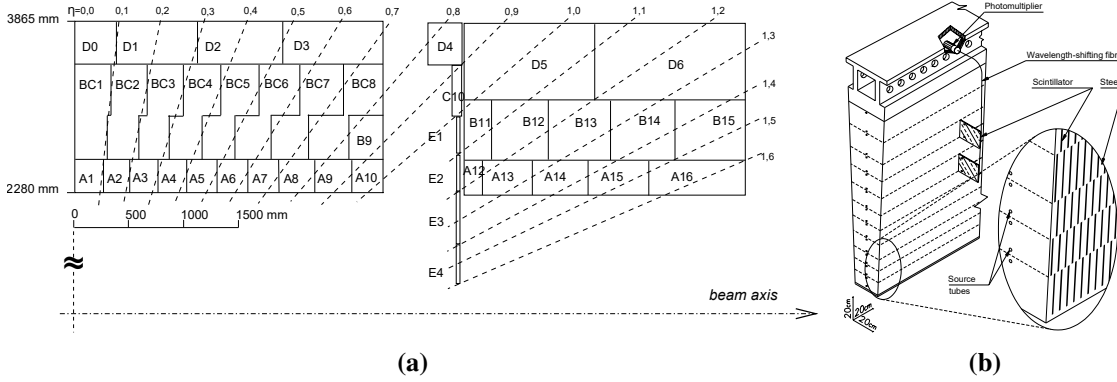


Figure 1.7: (a) Regions of the TileCal, including the gap and crack scintillators (C10 and E1-4). (b) Construction of a single TileCal wedge. From ref. [66].

wheels and 50mm in the outer wheels. The HEC is contained within the same cryostat as the electromagnetic endcap (EMEC) and the forward calorimeter (FCAL). The $\eta - \phi$ granularity of the HECs is approximately 0.1×0.1 for $|\eta| < 2.5$ and 0.2×0.2 for $2.5 < |\eta| < 3.2$. To improve resolution in the transition region between the HECs and the FCALs, the pseudorapidity range $3.1 < |\eta| < 3.2$ is covered by both the HEC and the FCAL [61, 68].

1.2.4.4 Forward Calorimeters

The forward calorimeters (FCAL) cover the highest range of pseudorapidity at $3.1 < |\eta| < 4.9$. Each FCAL consists of three submodules. The innermost layer consists mostly of copper absorbers, while the two outer layers are made from tungsten. Like the electromagnetic and hadronic endcap calorimeters, the forward calorimeters use liquid argon as the sensitive medium. The FCALs are located between the hadronic endcaps and the beam line, at a distance of about 5m from the interaction point.

The FCAL modules consist of blocks of absorber material into which holes have been drilled parallel to the LHC beam axis, each containing a copper electrode rod. A thin gap (0.27mm in FCAL1, 0.37mm in FCAL2, 0.51mm in FCAL3) is maintained between the electrode and the absorber by means of a polymer fibre that has been helically wound around the electrode rods. The exceptionally thin liquid-argon gap in the FCAL reduces the ion drift time by a factor of about 7.5 compared to the remaining liquid-argon calorimetry. A very fast detector response is necessary due to the high particle flux in the forward region [61, 69].

1.2.5 Magnet systems

The ATLAS detector's system of superconducting magnets consists of three subsystems: the Central Solenoid and the Barrel- and Endcap toroids.

The Central Solenoid surrounds the inner detector tracker at 2.44 m from the beam line, and extends to 2.63 m just in front of the liquid argon calorimeter. It consists of niobium-titanium superconductor which is encased in aluminium and wound around a support cylinder. The coil and support structure is contained inside a liquid Argon cryostat. The Central Solenoid produces a uniform magnetic field of 2 T, pointing along the beam axis, inside the inner detector to enable precise measurement of the track momentum. The heavy iron absorber plates of the Tile Calorimeter act as a "return yoke" that cancels out the solenoid B-field outside the Inner Detector [70].

The B-field generated by the toroid magnets is directed tangentially around the beam axis, along the ϕ -direction in the ATLAS coordinate system. The toroid coils are located outside the calorimetry, and are interwoven with the muon systems. The toroid field is necessary for measuring the momenta of muons. The barrel toroid consists of 8 coils with separate cryostats that span the full 25 m length of the detector. The endcap toroids consist of 8 coils inside a common cryostat that are fitted into the spaces within the barrel toroid that extend beyond the length of the Tile calorimeter. The B-field strength inside the toroids is around 3.5 T [70].

1.2.6 Muon systems

The Muon Spectrometer is the outermost subdetector of ATLAS. It consists of barrel and endcap sections covering $|\eta| < 1.05$ and $1.05 < |\eta| < 2.7$ respectively. In order to accurately reconstruct the muon trajectories, the muon spectrometer is emerged in a strong magnetic field originating from the superconducting toroid magnets. Field strengths vary from around 2.5 Tm in the barrel region to around 6 Tm in the endcaps [71].

Four types of detector modules make up the muon spectrometer: Monitored Drift Tubes (MDT), Resistive Plate Chambers (RPC), Thin Gap Chambers (TGC) and Cathode Strip Chambers (CSC). The most common type of module, MDT, consists of a collection of drift tubes containing anode wires in a Ar/CO₂ gas mixture. Muons passing through the drift tubes liberate electrons in the working gas that drift towards the anode, thus

enabling the measurement of the muon position along the tube and distance from the anode wire [71].

The three other detector types operate according to similar principles. The RPCs consist of parallel plate electrodes, while TGCs consist of closely spaced anode wires and cathode plates. CSCs are multi-wire proportional chambers containing planes of alternating cathode and anode wires. RPCs can be found in front of and behind the MDT modules. Each RPC module provides two measurements of the muon (η, ϕ). The RPCs have a very fast response time compared to the MDTs and are used for triggering. The TGCs are located in the endcap section of the muon spectrometer and, like the RPCs, provide very fast position measurements. The CSCs are designed to handle a high particle flux and are used in place of MDTs in the forward region of the muon spectrometer [71].

1.2.7 Triggers and data processing

The ATLAS trigger and data acquisition (TDAQ) system consists of a hardware-based *Level 1 Trigger* (L1) and software-based *High-Level Trigger* (HLT). The L1 trigger is able to process events at the bunch crossing frequency (40MHz) and consists of the L1-Muon and L1-Calorimeter trigger subsystems which accept inputs from the calorimeters and muon systems respectively. A topological trigger [72] (L1Topo) was added in Run 2. This trigger runs on field programmable gate arrays (FPGAs) and is able to make selections based on the kinematic relationships of trigger objects that are produced by the L1Muon and L1Calo. The L1Topo is able to calculate, among other things, the missing transverse momentum which is an important quantity in many physics analyses. The final trigger decision in the L1 trigger is made by the Central Trigger Processor (CTP) which takes as its inputs the outputs of the L1Muon, L1Calo and L1Topo triggers. The CTP also ensures that the amount of time between L1 trigger accepts is sufficiently long so that features of the same event are not captured twice, and limits the number of L1 accepts within a given number of bunch crossings so that the front-end buffers, which store the detector readout until an HLT decision has been made, do not overflow [73, 74].

The frequency of events entering the High-Level Trigger is approximately 100kHz. The software-based HLT takes as its input the full detector information saved in the buffers in addition to Region-of-Interest (ROI) information from the L1 trigger. The

ROIs are used by the HLT to identify the sections of the detector where relevant physics takes place. The HLT then uses this information to construct higher-level objects, such as jets, muons or missing E_T . Stringent cuts are applied to the reconstructed objects. The HLT defines around 1500 different *trigger chains* that combine decisions from the L1-trigger with corresponding cuts on HLT objects to produce a final decision. A list of trigger chains, called the *trigger menu*, is compiled ahead of each run according to current research interests while limiting the final rate of saved events to around 1 kHz. The HLT in Run 2 benefited greatly from the introduction of the L1 Topo as it helped reduce the L1 rate by as much as a factor of 10 for some trigger chains [74, 72].

Events that have been accepted by the HLT are saved on tape at CERN’s own computing facilities (Tier-0) before being sent to offline processing. Offline reconstruction takes place across the LHC computing grid [75] which consists of tens of thousands of computing nodes located at research institutes around the world. Nodes are located at either Tier-1 data centres, which have sufficient storage to process the raw events from Tier-0, or Tier-2 facilities which handle end-user analysis tasks. Computing on the grid for ATLAS is handled by the Production And Distributed Analysis (PAndA) framework [76, 77] which coordinates the production of large Monte Carlo samples as well as offline analysis jobs submitted by ATLAS members. At the end of Run 2 in 2019 ATLAS computing resources were handling over 370PB of data and simulated events. Datasets containing reconstructed objects are stored in a file format call Analysis Object Data (xAOD), which are further processed into derivations, or DAOD, containing only the relevant objects for the analyses of the various ATLAS subgroups [78]. Both xAOD and DAOD files are accessed and processed using the ROOT software framework [79].

1.3 Jet inputs

Jets are objects constructed from the collections of particles that are produced when a quark or gluon is emitted in a hard scatter process. As a consequence of the collinear divergences observed in QCD, the parton shower generated by a highly energetic quark or gluon is made up of partons emitted at small angles with respect to the hard-scatter parton. As the shower evolves down to lower energies, it eventually reaches the hadronisation

scale at around 1 GeV and the partons begin to form hadrons. The result is a collimated spray of hadrons that leaves a, roughly cone-shaped, signature in the detector trackers and calorimeters. The purpose of the concept of a jet is to construct an object from these bundles of hadronic energy that acts as a proxy for the initial hard parton so that its 4-momentum can be reconstructed. A detailed discussion of jets and how they are constructed is given in section [1.4](#).

There are many different ways of combining signatures in the detector to reconstruct the kinematics of hadrons. Hits in the ID are combined to construct charged particle tracks and energy deposits in calorimeter cells energies are merged into topological clusters. These composite objects encode the kinematics of individual particles and can be combined using jet algorithms to reconstruct jets. However, it is also possible to first combine them into composite objects, which may improve jet reconstruction performance. The following section introduces the various types of jet inputs that are used in ATLAS analyses of Run 2 data.

1.3.1 Topoclusters

The term *topocluster* refers to a three-dimensional collection of topologically connected signal cells inside the calorimeters. The purpose of constructing topoclusters is to reconstruct the energy and shape of particle showers inside the calorimeters, while reducing background effects chiefly caused by electronic noise and pileup. This can be achieved by taking advantage of the high spatial granularity of the calorimeters. Thanks to the lateral and longitudinal segmentation of the calorimeters into cells, the energy flow structure of particles showering in the detector can be resolved. The observable which determines the

construction of topoclusters is the cell signal significance

$$\varsigma_{cell}^{EM} = \frac{E_{cell}^{EM}}{\sigma_{cell}^{EM}} \quad (1.46)$$

where E_{cell}^{EM} is the EM-scale energy of the cell and σ_{cell}^{EM} is the average expected cell noise. The cell noise is calculated from electronic noise and pileup contributions:

$$\sigma_{cell}^{EM} = \sqrt{\sigma_{noise}^{electronic} + \sigma_{noise}^{pileup}}. \quad (1.47)$$

In ATLAS Run 2, the dominant source of cell noise is in-time pileup. The value of σ_{noise}^{pileup} depends on the average number of $p-p$ interactions per bunch crossing $\langle\mu\rangle$. For the duration of Run 2 the expected values of $\langle\mu\rangle$ was set to 40⁴.

Topocluster construction is defined by a set of three different cuts $\{S, N, P\}$, satisfying $S > N \geq P$, on the signal significance ς_{cell}^{EM} .

$$|E_{cell}^{EM}| > S \sigma_{noise, cell}^{EM} \quad \Rightarrow \quad |\varsigma_{cell}^{EM}| > S \quad (1.48)$$

$$|E_{cell}^{EM}| > N \sigma_{noise, cell}^{EM} \quad \Rightarrow \quad |\varsigma_{cell}^{EM}| > S \quad (1.49)$$

$$|E_{cell}^{EM}| > P \sigma_{noise, cell}^{EM} \quad \Rightarrow \quad |\varsigma_{cell}^{EM}| > S \quad (1.50)$$

Cells passing the S cut are considered seeds. The seed is then "grown" into a *protocluster* by adding neighbouring cells that pass the threshold N . Finally, all cells neighbouring the protocluster that pass the filter threshold P are also added to the protocluster. The optimal threshold values for hadronic energy reconstruction at ATLAS have been determined to be $S = 4$, $N = 2$ and $P = 0$ [80]. The value of $P = 0$ implies that all cells neighbouring the protocluster are included to form the topocluster. In cases where two protoclusters are sufficiently close together such that neighbouring cells both pass either the S or N cut, the protoclusters are merged into one.

⁴The actual value of $\langle\mu\rangle$ during Run 2 was measured at 33.7 with yearly averages ranging from 13.4 in 2015 to 37.8 in 2017 [49].

On its own, the algorithm above tends to generate very large clusters that are likely due to multiple particles. To mitigate this effect, a final splitting step is introduced in which a search for local signal maxima is performed inside the protoclusters. Local maxima are defined as cells with an energy of $E_{cell}^{EM} > 500 \text{ MeV}$ surrounded by at least 4 neighbouring cells with a smaller signal. If a protocluster is found to contain more than one local maximum, its constituent cells are divided between each local maximum to form new protoclusters. For cells that neighbour two or more local maxima, the cell energies are split between the two largest neighbouring maxima $E_{clus,1}^{EM}$ and $E_{clus,2}^{EM}$ by applying the geometric weights $w_{cell,1}^{geo}$ and $w_{cell,2}^{geo}$. These weights are determined by the distances d_1, d_2 from the cell to each local maximum and are defined as follows:

$$w_{cell,1}^{geo} = \frac{E_{clus,1}^{EM}}{E_{clus,1}^{EM} + rE_{clus,2}^{EM}}, \quad (1.51)$$

$$w_{cell,2}^{geo} = 1 - w_{cell,1}^{geo}, \quad (1.52)$$

where

$$r = \exp(d_1 - d_2). \quad (1.53)$$

1.3.1.1 Local Cluster Weighting

The scale at which energies of Electrons and photons are reconstructed from shower signatures in the Liquid Argon calorimeter is called the *EM-scale*. Hadrons only deposit a fraction of their total energy as electromagnetic showers. The *Local Cluster Weighting*, or LCW scale, brings the energy of a topocluster to the scale of hadronic energy loss, while also compensating for the signal lost in inactive detector material and on cluster boundaries. Based on position (η_{clus}), longitudinal depth (λ_{clus}) the total signal (E_{clus}^{EM}) and signal density (ρ_{clus}), the LCW calibrated cluster can be classified as being due to either electromagnetic or hadronic showers. This is quantified as the EM-probability $\mathcal{P}_{clus}^{EM}$ which represents the probability of the cell signal being purely due to electromagnetic showers. The dependence of $\mathcal{P}_{clus}^{EM}$ on λ_{clus} and ρ_{clus} is shown in figure 1.8 The local

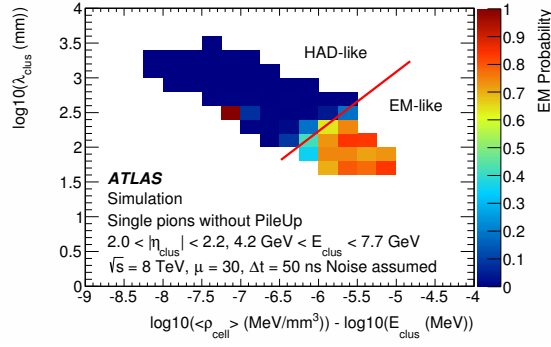


Figure 1.8: The distribution of EM-probability as function of cluster depth λ_{clus} and signal density $\rho_{clus}/E_{clus}^{EM}$. From Ref. [60].

hadronic calibration is then performed by reweighting the cell energy by w_{cell}^{cal} which is calculated from the EM-probability according to the formula

$$w_{cell}^{cal} = P_{clus}^{EM} w_{cell}^{em-cal} (1 - P_{clus}^{EM}) w_{cell}^{had-cal}, \quad (1.54)$$

where w_{cell}^{em-cal} and $w_{cell}^{had-cal}$ are the calorimeter signal weights. With no weighting applied a cell signal is considered to be at EM-scale, therefore we set $w_{cell}^{em-cal} = 1$. The hadronic signal weight is calculated from

$$w_{cell}^{had-cal} = \frac{E_{cell}^{dep}}{E_{cell}^{EM}}, \quad (1.55)$$

where E_{cell}^{EM} is the EM-scale cell and E_{cell}^{dep} is the true deposited energy. Like the EM-probability $\mathcal{P}_{clus}^{EM}$ the value of E_{cell}^{dep} is retrieved from a lookup-table from a set of cluster observables. These observables are the cluster rapidity η_{clus} , the signal strength E_{clus}^{EM} and the cluster depth λ_{clus} . The lookup-tables are constructed by filling histograms from MC simulations in which single pions with a wide range of energies enter the detector from all directions.

Additional corrections to the cluster energies are performed to account for signal losses due to unclustered energy deposits (out-of-cluster corrections) and energy deposited in inactive components of the detector (dead material corrections). A more detailed discussion of how this corrections are performed can be found in ref. [60].

1.3.2 Tracks

Tracks are rarely used as jet inputs on their own, except for the purpose of flavour tagging. However, as they form a part of composite jet inputs and are used in many jet substructure studies, a short description of track reconstruction at ATLAS is provided here. More detailed descriptions can be found in refs. [81–83].

Tracks are constructed from hits in the inner detector. The 2-dimensional images on the Pixel and SCT sensors are transformed into *clusters* by identifying adjacent signal pixels. Clusters are identified by their size as being either due to a single particle or multiple particles, so called *merged clusters*. Clusters that are adjacent on multiple layers of the same sensor are combined into *space points* which are then used as inputs to the 3-dimensional track finding algorithm. The track finding algorithm is a combinatorial scheme which starts by considering a set of three space points as a track seed. Momentum and trajectory estimates are calculated from the three space points of the seeds so that seeds which are incompatible with helical trajectories originating at the beam line can be discarded. A Kalman filter [84, 85] extrapolates the trajectories from the seeds by building track candidates from the remaining space points. Multiple track candidates are built from each seed. An ambiguity solving algorithm is applied to the track candidates to identify the charged particle tracks. The ambiguity solver assigns a score to each candidate track based on the observed and expected density of single-particle clusters and merged clusters in the track. Missing hits along the candidate trajectory, or holes, are also considered in the score calculation. A neural network is also used to help score track candidates. Track candidates are accepted or rejected based on their scores and the number of clusters shared with other tracks. All tracks must have $p_T > 400\text{MeV}$ and $|\eta| < 2.5$ and be compatible with the position of the beam line. A high-resolution track fit is performed on the accepted tracks to determine the final track p_T measurement. Neural networks are used in aiding the ambiguity solver and performing the track fits.

1.3.3 Particle Flow Objects (PFO)

Particle-flow algorithms are a well-established method for combining tracking and calorimeter information to improve the momentum and angular resolution of jets. By including tracking information in addition to calorimeter topoclusters one can improve the

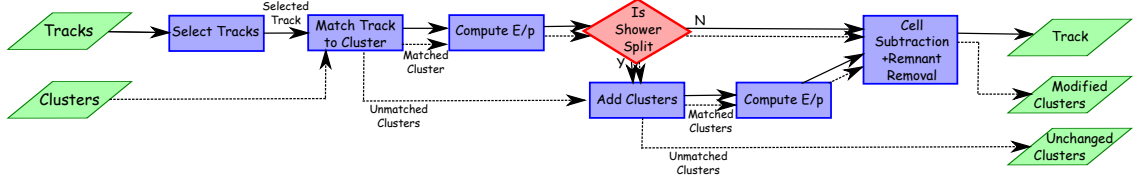


Figure 1.9: A flowchart showing the steps of the particle-flow algorithm. From ref. [86].

p_T resolution at lower p_T and capture additional particles that are not energetic enough to leave a signature in the calorimeters while also benefiting from the high angular resolution of the tracker compared to the calorimeter. As mentioned in section 1.3.1, topoclusters are not always due to a single particle. The goal of particle-flow is to combine tracks and topoclusters into a single object representing an individual particle without double-counting energy/momentum between the tracker and calorimeter measurements. A flowchart of the algorithm is shown in figure 1.9.

The particle flow algorithm first selects a collection candidate clusters satisfying $E^{cluster}/p^{track} > 0.1$. Clusters are matched to tracks based on the angular distance metric

$$\Delta R' = \sqrt{\left(\frac{\Delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\Delta\eta}{\sigma_\eta}\right)^2}, \quad (1.56)$$

where σ_ϕ and σ_η are the angular widths of the topocluster. The cluster with the smallest $\Delta R'$, provided $\Delta R' < 1.64$, is considered matched to the track. If no match is found, the particle that produced the track is considered not to have produced a cluster and a particle-flow object is created from the track only.

After the clusters have been matched to the track, the next step of the algorithm is to determine whether the whole particle energy has been deposited in a single cluster or if multiple clusters have to be associated to the track. This is determined by the following significance metric

$$S(E^{clus}) = \frac{E^{clus} - \langle E_{dep} \rangle}{\sigma(E_{dep})}, \quad (1.57)$$

where $\langle E_{dep} \rangle$ is the mean expected deposited energy of the particle and $\sigma(E_{dep})$ is the associated standard deviation. The value of $\langle E_{dep} \rangle$ and $\sigma(E_{dep})$ is retrieved from a lookup-table against the p_T of the track. The distributions of E_{dep} for different track p_T have

been determined from single pion simulations. If $S(E^{clus}) < -1$ for the matched cluster, clusters within $\Delta R < 0.2$ of the track are also considered matched to the track.

Finally, if $E^{matched\ clusters} > \langle E_{dep} \rangle$, an energy subtraction procedure is performed on the matched clusters by considering the energy deposited in each calorimeter cell. The track is extrapolated to the calorimeter layer where the energy density of the matched clusters is highest. Then, the energy distribution across calorimeter cells is compared to the expected energy profile (also determined from single-pion simulation) in each calorimeter layer. Cells are removed from the matched clusters until the expected energy $\langle E_{dep} \rangle$ is reached. The topoclusters that remain after cell-energies have been subtracted together with the tracks are said to represent the total "energy-flow" of the particles in the event and make up the set of *particle-flow objects* (PFO).

1.3.4 Track-Calo Clusters (TCC)

In order to study the substructure of high p_T jets, for example to distinguish boosted objects from a multijet background, one needs to be able to accurately resolve the angular structure of the jet constituents. While the ATLAS calorimeter has an excellent energy resolution, the angular resolution is limited by the coarse granularity of the calorimeter cells. The ATLAS tracker, on the other hand, has a very high angular resolution for charged particles, and a degraded p_T resolution for highly energetic charged particles. The goal of *Track-Calo-Clusters* (TCC) is to utilise the strengths of the tracker and the calorimetry by combining tracks and calorimeter topocluster measurements into a single object that is optimised for precision jet substructure measurements.

In order to construct TCCs, the first step is to perform a matching between extrapolated tracks and topoclusters. The matching is only performed if the angular uncertainty on the track extrapolation is smaller than the topocluster width. The matching criterion for combining a track and a cluster into a TCC is then

$$\Delta R < \sqrt{\sigma_{cluster}^2 + \sigma_{track}^2}, \quad (1.58)$$

where σ_{track} is the angular resolution for the tracks extrapolated to the calorimeter and $\sigma_{cluster}$ is the topocluster width. Highly boosted objects tend to be resolved as a single

topocluster whose width is typically greater than the uncertainty of its position. Hence, the matching criterion uses the cluster width rather than the cluster angular resolution.

For a direct single match between a track and a topocluster the TCC 4-vector is constructed as follows:

$$TCC_{1track,1cluster} = (p_T^c, \eta^t, \phi^t, m_c = 0). \quad (1.59)$$

Clusters and primary vertex associated tracks for which there is no match are also used to construct TCCs. As for topoclusters, every TCC is considered massless, therefore the 4-momenta of cluster-only and track-only TCCs become

$$TCC_{cluster-only} = (p_T^c, \eta^c, \phi^c, m_c = 0) \quad (1.60)$$

$$TCC_{track-only} = (p_T^t, \eta^t, \phi^t, m_t = 0). \quad (1.61)$$

In cases where there are multiple tracks to one or more clusters, the cluster energies are shared between the tracks so that one TCC object is created for each track. The general formula for the 4-momentum of such a TCC for a given track τ is

$$TCC_\tau = (p_T[\mathbf{M}_\tau], \eta^\tau, \phi^\tau, m[\mathbf{M}_\tau]) \quad (1.62)$$

where

$$\mathbf{M}_\tau = \sum_{c \in C_\tau} \mathbf{p}^c f_c^\tau \mathcal{F}_{c,t}^\tau. \quad (1.63)$$

In this equation, the factors f_c^τ represents the fraction of 4-momentum $\mathbf{p}^c$ shared by cluster c with track τ compared to the other clusters in the set of matched clusters C_τ . The factor $\mathcal{F}_{c,t}^\tau$ represents the fraction of cluster 4-momentum shared with a track t out of the set of tracks T_c that have been matched to cluster c . Explicitly these fractions can be expressed as

$$f_c^\tau = \frac{p_T^c}{p_T [\sum_{k \in C_\tau} \mathbf{p}^k]}, \quad \mathcal{F}_{c,t}^\tau = \frac{p_T^\tau}{p_T [\sum_{t \in T_c} \mathbf{p}^t f_c^t]}. \quad (1.64)$$

At very high p_T , where jets are closely collimated, TCC-jets have shown excellent performance compared to topocluster-jets. Performance improvements were shown in the reconstruction of the combined jet mass and the jet moment D_2 (see section 1.5.2 for a definition) for $p_T > 2\text{TeV}$. By using only combined TCCs, i.e. excluding the track-only and cluster-only objects, it was also shown that one can construct jets that are resilient to high levels of pileup without requiring grooming (see section 1.5.1 for a definition). TCCs are therefore most suitable as jet inputs in the very high p_T regime [87].

1.3.5 Unified Flow Objects (UFO)

The latest development in combined track-cluster jet constituents are unified flow objects (UFO). These objects were designed to combine the advantages of both the TCC and Particle-Flow combination schemes and ensure high precision in substructure measurements across a wide range of p_T . For TCCs, the boosted object tagging performance was found to be superior to LC-Topocluster jets only at very high- p_T . Jets using particle-flow object show improved performance compared to LC-topocluster jets across the whole p_T range, but perform significantly worse than TCCs at high p_T . The degraded performance for particle-flow at high p_T can be explained by the loss of precision in track p_T measurements as the amount of track curvature decreases.

An outline of the UFO construction algorithm is shown in figure 1.10. First, the particle-flow algorithm is run on the tracks and topoclusters in the event. The algorithm then distinguishes between four types of objects:

1. neutral PFOs,
2. charged PFOs for which energy has been subtracted from topoclusters,
3. charged PFOs with no cluster energy subtraction,
4. tracks for which no PFOs have been created.

Primary vertex association is applied to the tracks that are not part of the collection of PFOs (category 4). The algorithm then proceeds by producing TCCs from the objects

in categories 1, 3 and 4. The tracks for which energy has been subtracted from the clusters in particle-flow (category 2), are not used as inputs to TCCs since these are considered well-matched to clusters. These tracks are already isolated enough that they are not part of the kind of dense environment that one might find in very boosted high- p_T jets. The collection of resulting TCCs and category 2 PFOs are the final *unified flow objects* (UFO).

In the latest ATLAS recommendations for UFO based jets, two pileup mitigation procedures are applied to the PFOs before they are input into the UFO algorithm: *constituent subtraction* [88] and *soft-killer* [89]. The former technique utilises ghost-particles of negative p_T to subtract the expected momentum density due to pileup, while the latter uses geometric areas to perform a similar subtraction. Soft-Drop grooming (see section 1.5.1 for a definition) is also applied to the UFO jets.

UFOs were found to increase the background rejection of standard W taggers by up to 120% at 50% signal efficiency compared to LC-topocluster jets. For top taggers this figure was up to 135% [90]. These very encouraging results make UFOs an ideal choice for new types of substructure-based jet taggers.

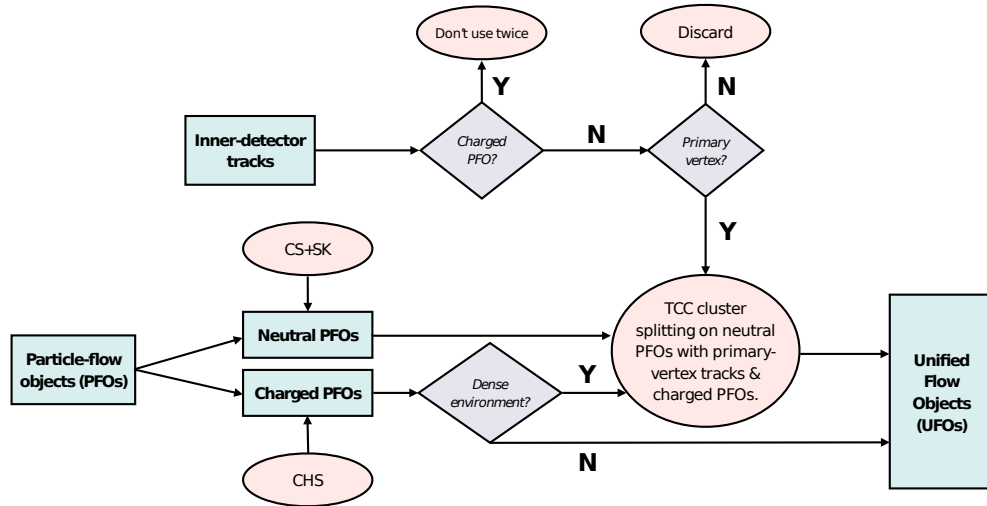


Figure 1.10: A flowchart showing the UFO reconstruction algorithm. From [90].

1.4 Jets

1.4.1 The purpose of jets

Although the concept of a jet seems rather intuitive, there is no common prescription for how a jet should be constructed. A jet must therefore be defined through the recipe, or algorithm, that is used to construct it. There are many different jet algorithms, each with a number of free parameters, and hence a multitude of *jet definitions*, that can be employed to meet the needs of any particular analysis. In 1990, the *Snowmass accords* [91] were formulated. They defined a list of requirements that a good jet definition should satisfy, so that observables reconstructed using jets can be directly compared between theory and experiment. According to the Snowmass accords, a jet definition should:

1. be simple to implement in an experimental analysis;
2. be simple to implement in the theoretical calculation;
3. be defined at any order of perturbation theory;
4. yield a finite cross-section at any order of perturbation theory;
5. yield a cross-section that is relatively insensitive to hadronisation.

Points 1 and 2 are practical requirements. At the time of the Snowmass accords, computing power was much more limited than it is today, and the range of jet algorithms considered practical was strongly constrained by complexity. Points 3 and 4 relate to the concept known as *infrared and collinear safety*, which will be discussed in section 1.4.2. Point number 5 can be related to the broader requirement that a jet definition should capture the physics of the hard-scatter process while being insensitive to non-perturbative processes, including hadronisation but also underlying event and multi-parton interactions. This requirement is not trivial to define and jet physicist have developed many different schemes with the aim of subtracting soft radiation from jets. One category of such schemes, called *grooming algorithms*, is described in greater detail in section 1.5.1.

1.4.2 IRC safety

A very desirable property of jet algorithms and jet observables is Infra-Red and Collinear safety (IRC). A jet algorithm is said to be infrared safe if the outcome is unaffected by soft

gluons that are emitted by the partons throughout the showering process. Collinear safety refers to the requirement that the algorithm should remain unaffected by processes where one particle in the shower splits into two decay products with roughly collinear momenta (see section 1.1.7 for a description of collinear emissions in parton showers). In both the infrared and collinear case, the sequence of events in the parton shower process should not affect how jets are reconstructed from the final state particles. Theoretically, the problem of infrared and collinear emissions can be understood in terms of singularities that they introduce into the analytical calculations of jet observables. If the jet algorithm is unaffected by infrared and collinear emissions, these singularities cancel out against divergences introduced by virtual emissions, or "loop" Feynman diagrams. However, for an IRC unsafe jet algorithm, the singularities no longer cancel and the measured observable is no longer finite. This effect is precisely the logarithmic enhancements introduced in section 1.1.4. Diagrams illustrating the concept of collinear and infrared safety are shown in Figures 1.11 and 1.12 respectively.

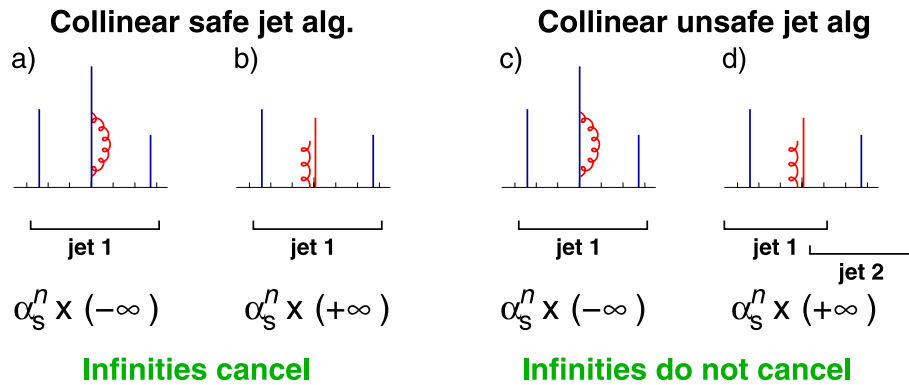


Figure 1.11: An illustration of collinear safety in jet algorithms. On the left, the collinear-safe jet algorithm clusters all emissions in a single jet. The collinear QCD emission shown in b) cancels out the virtual divergences introduced in a). On the right, the collinear emission in c) causes the collinear-unsafe jet algorithm to cluster the emissions as two separate jets and the divergences introduced by c) no longer cancel out. From Ref. [92].

1.4.3 Jet algorithms

The three most commonly used jet algorithms in the LHC era of particle physics are all part of the k_T family [93] of sequential recombination algorithms. In these algorithms,

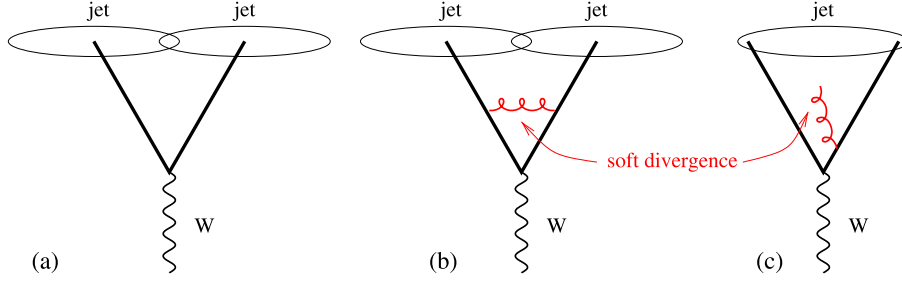


Figure 1.12: An illustration of infrared safety in jet algorithms. A W boson decaying to two jets is shown in (a). A soft emission causes the IR-unsafe algorithm in (c) to cluster the two jets as one, whereas in (b) the IR-safe algorithm remains unaffected. From Ref. [92].

the following variables are defined:

$$d_{ij} = \min \left(p_{Ti}^{2p}, p_{Tj}^{2p} \right) \frac{\Delta R_{ij}^2}{R^2}, \quad \Delta R_{ij} = \sqrt{\Delta \phi_{ij}^2 + \Delta y_{ij}^2}, \quad d_{iB} = p_{Ti}^{2p}, \quad (1.65)$$

where ΔR is the angular distance in rapidity y and azimuth ϕ and $p_{Ti,j}$ are the transverse momenta of the particles, tracks or calorimeter clusters to be recombined into jets. The algorithms then perform the following steps:

1. For each pair of jet constituents i, j , find the minimum of all d_{ij} and d_{iB} .
2. If d_{ij} is a minimum, add the 4-vectors of the jet constituents and return to 1.
3. If d_{iB} is a minimum or there are no more constituents left, the jet is considered finished.

The three most commonly used jet algorithms at the LHC today are distinguished by different values for the parameter p . All three algorithms in the k_T family satisfy IRC safety. The factor $\Delta R_{ij}^2/R^2$ ensures that collinear splittings are clustered right at the beginning of each algorithm while the p_T -dependent factor takes care of soft emissions. Each of the three algorithms derived from the k_T -algorithm has its own advantages and disadvantages depending on its application [92].

1.4.3.1 The k_T algorithm

For the k_T algorithm [93] the momentum weighting is set to $p = 1$. The name of the algorithm comes from the fact that, in the limit of highly collinear jet inputs, the dis-

tance metric reduces to the relative transverse momentum k_T of one jet input to the other. The k_T algorithm begins by combining the jet inputs with the smallest relative transverse momentum into proto-jets and proceeds by combining these until the full jet has been reconstructed. One can think of this clustering sequence as imitating the process of jet formation in reverse, where soft "emissions" with low relative transverse momentum are successively clustered with harder "jet cores" until the full particle shower has been recombined. While this property makes the k_T algorithm intuitive to understand, it also produces jets of irregular shapes and sizes. This makes it difficult to determine if a given particle is contained within a given jet without running it through the jet algorithm, which can complicate experimental calibration procedures and corrections for background effects such as pileup removal.

1.4.3.2 The Cambridge-Aachen algorithm

For the Cambridge-Aachen (CA) algorithm [94, 95] $p = 0$. Hence, this algorithm clusters the jet inputs only based on the angular distance ΔR_{ij} between them. This algorithm was introduced following the k_T algorithm with the motivation that it would reduce the number of "junk jets" or falsely reconstructed jets that are entirely due to low- p_T hadrons. The CA algorithm produces irregularly shaped jets with a smaller average area than k_T jets, which makes it less sensitive to contaminations from the underlying event and pileup. Like the k_T algorithm, it is also sensitive to soft emissions as these tend to be closely separated in ΔR . The CA algorithm is preferred for many jet substructure applications, because it provides an angular-ordered hierarchy of clusterings which closely resembles the angular ordering that is observed in models of collinear gluon emissions that are used in many Monte Carlo event generators (see section 1.1.7). The CA algorithm is also used for the *filtering* jet grooming (introduced in section 1.5.1) technique for removing soft contamination from jets.

1.4.3.3 The Anti- k_T algorithm

The anti- k_T algorithm [96] has $p = -1$ and therefore clusters the hardest jet inputs first. It produces jets that are circle-shaped in $\eta - \phi$ space with radius R , which is practical for jet calibration purposes. The kinematic reconstruction performance of the anti- k_T algorithm is usually better than that of the k_T algorithm and comparable to that of the CA algorithm

for most applications. For these two reasons, the anti- k_T algorithm is usually preferred for reconstructing jets in ATLAS and CMS analyses. Because the merging criterion is inverted compared to the k_T -algorithm, the clustering sequence of the anti- k_T algorithm cannot be said to mimic the process of jet formation. The soft proto-jets that one expects immediately before the final state are never reconstructed, rather the algorithm starts from the hard core of the jet and proceeds by adding progressively softer particles until all the particles satisfying the distance requirement have been added to the jet.

1.4.3.4 Jet area and ghost-association

The area [97] of a jet roughly refers to the region of y, ϕ space in which a particle would be clustered with the jet. It is defined using *ghost particles* with negligible momentum which are clustered with the real jet using the same algorithm used to construct the jet. Because the sequential recombination algorithms are infrared safe and the ghost particles do not contribute to the total 4-momentum of the jet, the definition of the jet is unchanged by the addition of ghost-particles. By locating the boundary inside which ghost particles are joined to the jet, one can define a "catchment area" for the jet. Two definitions of jet area exist: the *passive* area which is defined by clustering a single ghost particle with the final jet at various locations in space and the *active* area which is defined by clustering a collection of ghost particles along with the jet inputs during reconstruction. Figure 1.13 shows the active jet areas for jets reconstructed with the k_T , anti- k_T and Cambridge-Aachen algorithms in a single dijet event. The active area clearly shows how each algorithm produces differently shaped jets. Only the anti- k_T algorithm collects particles within a regular cone.

The jet area definition can also be used as a method for associating inner-detector tracks with jets that have been reconstructed from calorimeter topoclusters. By considering tracks as ghost-particles and using the passive area definition one can decide whether a track should be associated to a given jet. This process is called *ghost association*.

1.5 Jet substructure

The LHC produces many hadronically decaying heavy particles, such as top quarks, W , Z and Higgs bosons. At the LHC, when such heavy particles are produced, their energy

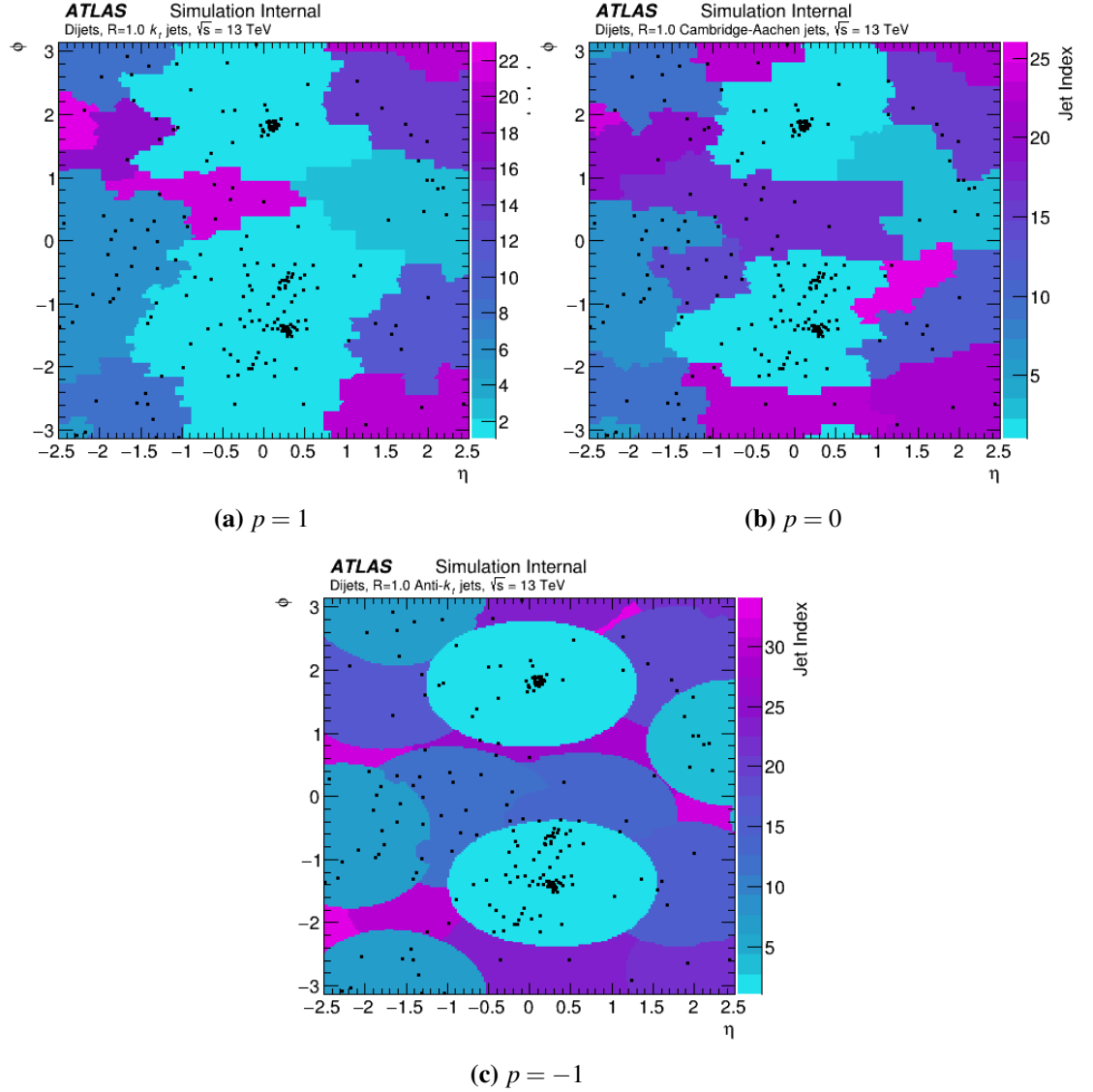


Figure 1.13: A dijet event clustered with the k_T (a), Cambridge-Aachen (b) and Anti- k_T (c) jet algorithms. The value of the distance parameter is $R = 1$, while the value of the momentum weighting parameter p is varied for each of the algorithms. The jet areas, calculated using ghost-association, are shown for each jet. Jets are coloured based on their rank when ordered by p_T from the highest to the lowest value. For each of the three algorithms, the two leading jets seem to capture the two regions where the density of particles is visibly higher

is typically so large that they carry substantial transverse momentum. This produces a spray of decay products which are highly boosted in the direction of the heavy particle that can be reconstructed inside a single jet. Typically, this will be a "fat", or "large- R " jet with size parameter $R = 1.0$ or larger. By looking at the constituents of this large jet and their energy distributions, one may attempt to identify the heavy particle that gave

rise to the jet. One may construct observables that capture the likelihood that a jet originated from a W boson or a top quark. The methods used to construct such observables often make use of the clustering history of the sequential recombination algorithm that is used to reconstruct the jet. Since these jet algorithms are themselves inspired by the process of jet formation, one may hope to gain some insights about the origin of the jet by studying its clustering history. The study of the constituents of jets, their spatial and kinetic distribution, and the process by which they are clustered into a jet make up the field called Jet Substructure (JSS) [30].

Jet Substructure is a relatively new field of study. It was only with the advent of the LHC era of particle physics that one began to consider reconstructing the highly boosted decay products of heavy particles inside a single jet. Some of the earliest JSS methods were devised to reconstruct the (then undiscovered) Higgs boson in its hadronic decay channels [98, 99]. In 2008, just before the LHC began its operations, a method was put forward by which the Higgs boson could be discovered in its vector-boson associated production channel [100]. This proposal made heavy use of jet substructure and laid the groundwork for many of the methods used today.

1.5.1 Jet grooming

Highly boosted objects, such as W and Z bosons or top quarks, may have decay products which are sufficiently collimated that they may not be reconstructed as well separated $R = 0.4$ jets. As a rule of thumb, they are typically contained within a cone of radius $2M/p_T$ where M is the mass and p_T the transverse momentum of the boosted object. The decay products are therefore instead reconstructed as a single large R, typically $R = 1.0$, jet. By studying the substructure of the large R jet one can identify the species of boosted object from which it originates. Before the jet can be *tagged*, however, *grooming* techniques are typically applied to remove contamination due to additional radiation from the beam (the underlying event), and simultaneous proton-proton collisions (pileup). Grooming algorithms typically start from clustered jets and remove the parts which are considered unnecessary. The most common grooming techniques are *trimming*, *filtering* and *pruning* [101].

Filtering [100] is performed by re-clustering the constituents of a jet using the Cambridge-Aachen algorithm with $R_{filter} < R_{jet}$. All but the n_{filter} largest p_T subjects are discarded. Since n_{filter} is a fixed parameter of this algorithm, filtering is most usefully applied to boosted object jets (W, Z, H) where a multi-pronged energy distribution inside the jet is expected. The parameter n_{filter} is set to the number of expected prongs, so that soft radiation thought not to have originated from the boosted objects decay product is removed.

Like filtering, trimming [102] works by clustering the jet constituents into smaller subjects of radius $R_{trim} < R_{jet}$ using either the k_T or Cambridge-Aachen algorithm. Then all subjects which do not pass a defined cut on the p_T fraction f_{trim} of the whole jet are discarded. Since the number of subjects is not fixed, the trimming algorithm may be used to groom jets of any origin, not merely boosted-object decays. Trimming is one of the most widely used grooming algorithms in ATLAS. Trimmed anti- k_T $R = 1.0$ topocluster jets have been the recommended large- R jets in Run 1 and most of Run 2. Jet mass and energy calibrations have been derived for these trimmed jets and they are the most common calibrated jets used in studies of jet substructure.

Pruning [103] is another multi-purpose grooming algorithm which exploits the features of sequential recombination jet algorithms. The constituents of the jet are re-clustered using either the Cambridge-Aachen or k_T algorithm using a distance parameter larger than R_{jet} . Soft and wide angle radiation is removed by requiring that the two proto-jets in each clustering step satisfy $\Delta R_{ij} < R_{prune}$, where $R_{prune} = 2f_{prune}m_{jet}/p_{T,jet}$, and $\min(p_T^1, p_T^2) < z_{prune}p_T^{1+2}$. If the two requirements are not met, the lower p_T proto-jet is discarded and the jet algorithm proceeds using only the higher p_T proto-jet as input to the subsequent clustering step.

1.5.2 Jet shapes

Jet substructure observables, or *jet shapes*, can be used to construct jet taggers, which are successfully used in many analyses to select boosted objects. This section presents some of the jet substructure observables most commonly used for identifying and studying W and top jets. Distributions of these observables have also been measured directly at ATLAS for W , top and QCD initiated jets [104]. To illustrate the utility of jet shapes in

discriminating between different types of jets, a collection of jet shape distributions from this measurement is shown in figure 1.14.

Energy Correlation Functions [105] are constructed from the energies and pair-wise angles of jet constituents. Three orders of energy correlation functions can be defined as follows

$$ECF1 = \sum_{i \in J} p_T \quad (1.66)$$

$$ECF2(\beta^{ECF}) = \sum_{i < j \in J} p_{T_i} p_{T_j} (\Delta R_{ij})^{\beta^{ECF}} \quad (1.67)$$

$$ECF3(\beta^{ECF}) = \sum_{i < j < k \in J} p_{T_i} p_{T_j} p_{T_k} (\Delta R_{ij} \Delta R_{ik} \Delta R_{jk})^{\beta^{ECF}} \quad (1.68)$$

The free parameter β^{ECF} can be chosen to weight the importance of the angular separation of the jet constituents. Most commonly, a value of $\beta^{ECF} = 1$ is chosen, but other integer values are also frequently used. It is also common to normalise the energy correlation functions $ECF2$ and $ECF3$ by $ECF1$, such that one obtains the ratios

$$e_2 = \frac{ECF2}{(ECF1)^2}; \quad e_3 = \frac{ECF3}{(ECF1)^3}. \quad (1.69)$$

These ratios, in turn, are used to construct the variables C_2 and D_2 [106]. These variables have been found to be very useful in identifying two-pronged structures within jets and are commonly used to perform W and Z boson tagging. They are calculated as follows:

$$C_2 = \frac{e_3}{(e_2)^2}; \quad D_2 = \frac{e_3}{(e_2)^3}. \quad (1.70)$$

N-subjettiness [107] is a quantity that describes the degree with which the constituents of the jet are compatible with being composed of N or fewer subjets. For $N = 0, 1, 2, 3$ the subjettiness is defined as

$$\tau_0(\beta^{NS}) = \sum_{i \in J} p_{T_i} \Delta R_i^{\beta^{NS}}, \quad (1.71)$$

$$\tau_1(\beta^{NS}) = \frac{1}{\tau_0(\beta^{NS})} \sum_{i \in J} p_{T_i} \Delta R_{a_1, i}^{\beta^{NS}}, \quad (1.72)$$

$$\tau_2(\beta^{NS}) = \frac{1}{\tau_0(\beta^{NS})} \sum_{i \in J} p_{T_i} \min \left(\Delta R_{a_1, i}^{\beta^{NS}}, \Delta R_{a_2, i}^{\beta^{NS}} \right), \quad (1.73)$$

$$\tau_3(\beta^{NS}) = \frac{1}{\tau_0(\beta^{NS})} \sum_{i \in J} p_{T_i} \min \left(\Delta R_{a_1, i}^{\beta^{NS}}, \Delta R_{a_2, i}^{\beta^{NS}}, \Delta R_{a_3, i}^{\beta^{NS}} \right), \quad (1.74)$$

where ΔR_i is the angular distance between the constituent i and the jet axis. Similarly, $R_{a_n, i}$ is the distance between i and the axis of the n^{th} subjet that has been re-clustered from all jet constituents using the k_T algorithm. The parameter β^{NS} is chosen to weight the importance of the angular distance and is typically set to 1. Ratios of N-subjettinesses have been found to be very good discriminants for identifying two- and three-pronged structures inside jets. These ratios are commonly defined as

$$\tau_{21} = \frac{\tau_2}{\tau_1}; \quad \tau_{32} = \frac{\tau_3}{\tau_2}. \quad (1.75)$$

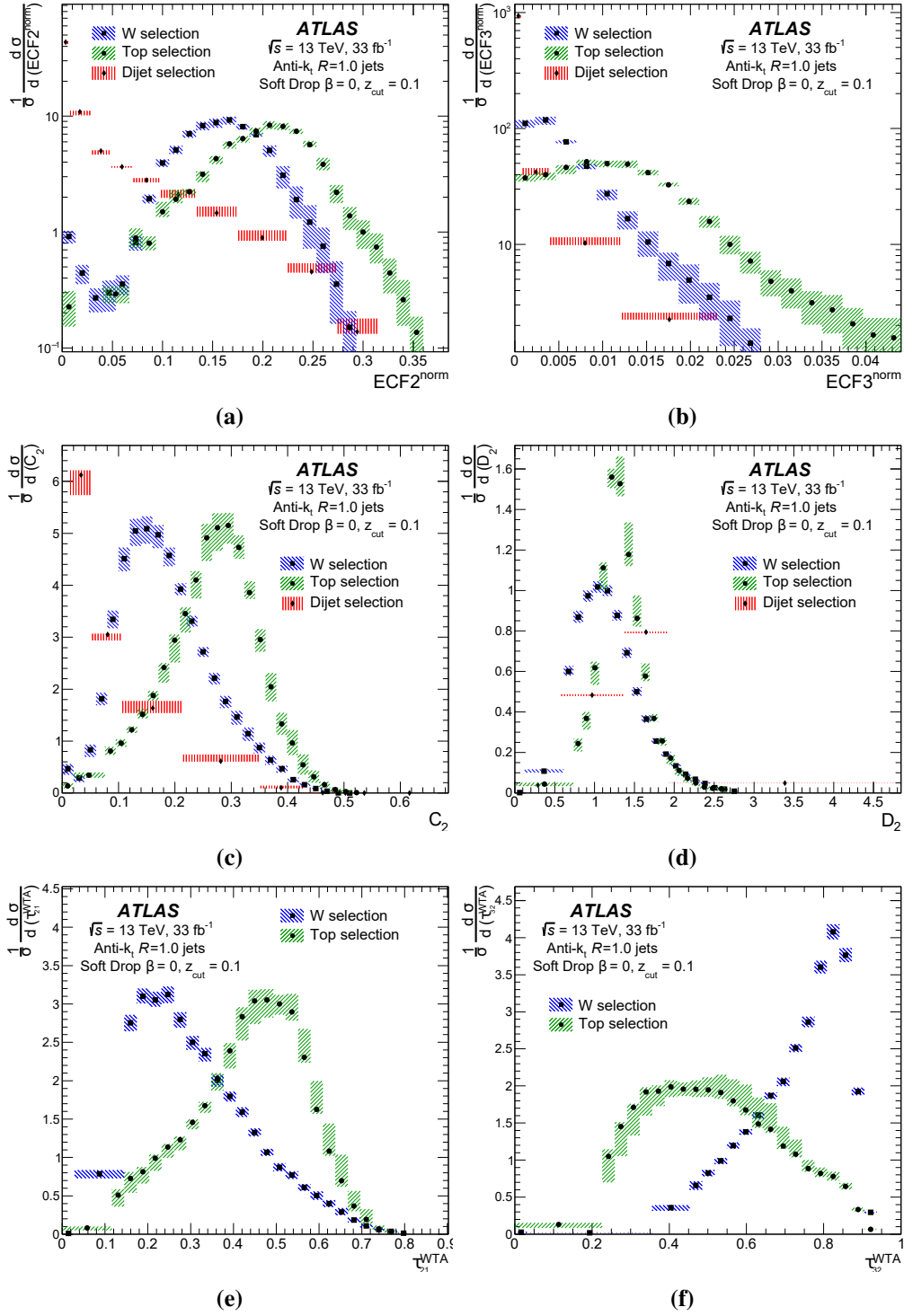


Figure 1.14: A collection of distributions of jet shape variables, as measured by ATLAS in 2016. Trimmed $R = 1.0$ topocluster jets with $p_T > 450 \text{ GeV}$ are selected from semi-leptonic $t\bar{t}$ and dijet events and are subsequently classified through a series of kinematic selections as originating from either top quarks, W bosons, or light quarks. The jet shapes are calculated from the untrimmed jet constituents after application of the Soft Drop algorithm. The variables shown are (a) ECF2 , (b) ECF3 , (c) C_2 , (d) D_2 , (e) τ_{21} , (f) τ_{32} . From ref. [104].

1.5.3 Soft Drop

Soft Drop is an IRC safe jet grooming procedure, which may also act as a tagger [108]. It is based on the Mass Drop tagging procedure [100], one of the earliest boosted-object jet taggers, as well as its later successors [109]. The Soft Drop procedure starts by clustering a jet using the Cambridge-Aachen algorithm. It then goes through the clustering steps backwards, starting with the finished jet j and the two proto-jets j_1, j_2 that it was formed from. It then checks if j_1 and j_2 pass the condition

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{cut} \left(\frac{\Delta R_{12}}{R} \right)^\beta \quad (1.76)$$

where p_{T1}, p_{T2} are the transverse momenta of the two proto-jets, ΔR_{12} is the angular distance between them, R is the jet radius and z_{cut} and β are parameters of the algorithm. If the condition is satisfied, the algorithm stops. Otherwise, the procedure continues by considering the proto-jet with the largest p_T as the new j . When the algorithm is used as a tagger, it simply continues until either the condition is met (tagged) or until there are no more proto-jets left (rejected). For jet grooming, the algorithm stops in case there is only a single jet constituent left to de-cluster.

The Soft Drop algorithm removes the effect of non-global logarithms. These are divergences introduced by gluons outside the jet that emit soft radiation into the jet cone. It can be shown that for $\beta > 0$, these effects are suppressed to the power of $1/\beta$, while for $\beta = 0$ they are completely removed. This means that, when matrix elements are calculated at up to next-to-leading-logarithmic order, the effect of non-global logarithms on Soft Dropped jets can be neglected.

1.6 Lund diagrams

The diagrams now known as Lund Planes were first introduced by Anderson et. al. of the Lund University theory group in a study of Coherence Effects in Deep Inelastic Scattering [110]. In this paper, graphs of rapidity y and transverse momentum k_T are presented in order to show the kinematically allowed regions for a QCD bremsstrahlung gluon emitted from a quark dipole system. The authors note that the allowed region is bounded by the condition $|y| \lesssim \ln(\sqrt{s}/k_T)$, which corresponds to a triangle in the $y -$

$\ln(k_T)$ plane. This result follows from an approximation used to calculate, perturbatively, the differential cross section of semi-hard gluon emissions from hadrons produced in e^+e^- annihilations [111]. The authors then proceeded to extend this approximation to deep inelastic electron-proton scattering and plotted their newly obtained limits on the $y - \ln(k_T)$ plane. This was the birth of the Lund Plane as a tool for visualising the kinematic phase space of QCD emissions. Lund diagrams have since become a common feature in discussions concerning parton-shower generation in the development of Monte Carlo event generators, like the Ariadne and Pythia programs also developed by the Lund group [112].

With the advent of analytical predictions of jet substructure variables with resummations of logarithmic corrections in the LHC era [109, 108], Lund diagrams received renewed attention as theoretical tools for visualising radiation patterns inside jets. For example, they have been used to visualise the impact of novel grooming algorithms on the distribution of QCD radiation inside top quark initiated jets in predictions of top quark masses [113]. This novel application of Lund Diagrams as jet physics tools would eventually become the inspiration for the Lund Jet Planes discussed in the next section.

1.7 The Lund Jet Plane

The *Lund Jet Plane* is an application of the concept Lund Diagrams to represent the kinematic structure within a jet. To understand the construction of the Lund Jet Plane, it is instructive to first consider a particle showering and producing new particles through decays and emissions. Each emission appears as a point on the triangular Lund Diagram associated with its parent particle, and one can plot a new smaller phase space to represent the kinematics of possible second-order emissions. By plotting the emissions from the initial particles on its Lund Diagram one constructs the *Primary Lund Plane*. The Primary Lund Plane does therefore not contain information about second order splittings. However, in analogy with the primary Lund Plane, one can construct a secondary Lund Plane and so on. An illustration of the construction of the Lund Plane from a particle shower is shown in figure 1.15.

As we have seen in Section 1.4.3, the sequential recombination algorithms that are used to reconstruct jets at the LHC often seek to mimic the showering process in reverse. This insight is the key to constructing the observable known as the *Lund Jet Plane* [114]. The constituents of the jet are recombined using a jet algorithm (typically Cambridge-Aachen) and the Lund Plane coordinates are calculated at each step. In the analyses performed at ATLAS the Lund Jet Plane is constructed from the coordinates $\ln(R/\Delta R)$, $\ln(1/z)$ where

$$\Delta R^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2; \quad z = \frac{p_T^j}{p_T^i + p_T^j}, \quad (1.77)$$

where R is the jet radius, or the distance parameter of the Cambridge-Aachen jet. The variable ΔR represents the angular separation between the two proto-jets i, j and z is the momentum fraction of the proto-jet j for which $p_T^j < p_T^i$. The difference in rapidity y , which is used to calculate ΔR , is invariant under Lorentz-boosts along the z -axis. To calculate y from the measured η it is necessary to know the mass of the particle. When the mass of the individual jet constituents is not known it is typically assumed that they have an average mass equivalent to that of a pion. To be consistent with the particle shower analogy above, we often denote the proto-jet with the larger transverse momentum as *core* and the smaller one as *emission*. It possible to define LJPs using other sets of kinematic variables, for example

$$k_T = p_{Tj} \Delta R; \quad \kappa = z \Delta R; \quad (1.78)$$

$$\psi = \tan^{-1} \frac{y_i - y_j}{\phi_i - \phi_j}; \quad m^2 = (p_i + p_j)^2, \quad (1.79)$$

as has been done in ref. [114]. We will, however, use the ΔR and z variables, as the former is a measure of the emission angle and the latter depends only on transverse momentum. This choice is, however, purely a matter of taste. Due to the wide range of ΔR and z that are observed for emissions within jets, it is also convenient to use the logarithms $\ln(\Delta R/R)$ and $\ln(1/z)$ to plot the Lund Jet Plane. In this parametrisation, the per-jet density of

emissions across the Lund Jet Plane is expressed as

$$\rho_{LJP}(\Delta R, z) = \frac{1}{N_{jets}} \times \frac{d^2 N_{emissions}}{d \ln(1/z) d \ln(\Delta R/R)}. \quad (1.80)$$

Figure 1.16 illustrates the dominant effects across different regions of the Lund Jet Plane. The plane has a triangular structure as it is crossed by the line $k_T (= z\Delta R) \lesssim \Lambda_{QCD}$, beyond which non-perturbative effects begin to dominate and the density of emissions decreases rapidly. Emissions of various physical origins dominate in different parts of the plane. The lower left corner of the Lund Jet Plane contains hard and wide-angle emissions. For jets originating from W or top quark decays, large contributions are expected in this region due to the hard splitting that is produced when a W decays to two quarks or when a top quark decays to a b and a W . The upper left corner of the plane contains contributions from soft emissions at large emission angles, and is therefore likely to be dominated by background effects, such as the Underlying Event (UE) and Multi-Parton-Interaction (MPI). Contributions from initial state radiation (ISR) are also expected along the left edge of the plane where the emission angle is large. Perturbative effects dominate where $k_T > \Lambda_{QCD}$, in the lower left corner of the plane. To leading order, the emission density across the Lund Jet Plane has been analytically shown to be

$$\rho_{LJP} \approx \frac{2\alpha_s(k_T)C_i}{\pi}, \quad (1.81)$$

where C_i is one of the Casimir constants from QCD, either $C_F = 4/3$ for a quark-initiated jet or $C_A = 3$ for a gluon-initiated jet. Because the value of α_s is greatest for $k_T \sim 1\text{GeV}$, one can expect the largest contributions to lie along this diagonal of the LJP [114].

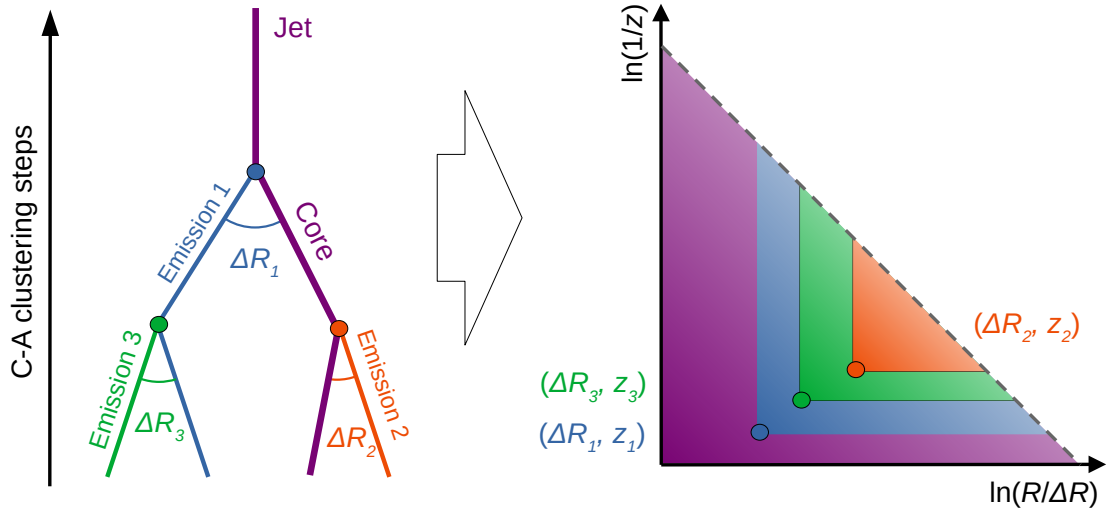


Figure 1.15: Illustration showing the process by which the primary and higher-order Lund Jet Planes are constructed. The Cambridge-Aachen algorithm clusters jet constituents based on angular distance ΔR . At each step, we label the smaller pT proto-jet "emission" and the harder branch "core". The variables $\ln(R/\Delta R)$ and $\ln(1/z)$ are plotted on the primary Lund Jet Plane (purple) for each emission originating from the core branch at each step. Emissions 1 and 2 (blue and orange) have both been emitted from the core branch and are therefore part of the primary LJP. Emission 3 (green), on the other hand, lies in the secondary LJP defined by the phase space of emission 1. The secondary and tertiary LJPs that are generated by each emission are represented by shaded areas in their corresponding colours.

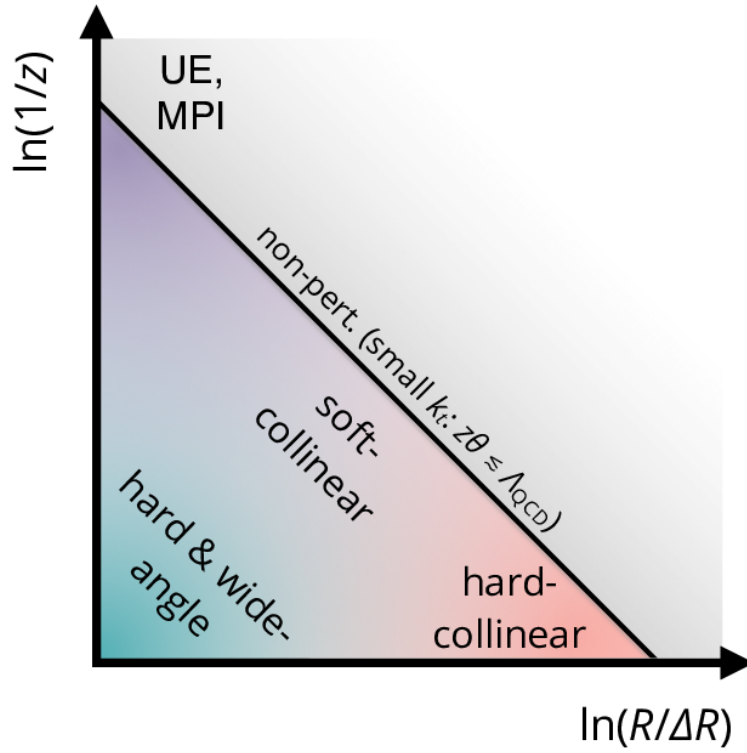


Figure 1.16: Schematic showing the structure of the Lund Jet Plane. The Lund Jet Plane provides an overview of the p_T -fraction and angular distributions of radiation inside a jet. Emissions of different origins, such as hard-scatter processes, underlying event or pileup each have different characteristic values of ΔR and z . Each regions of the Lund Jet Plane is dominated by different types of radiation. From Ref [115]

Chapter 2

Run 2 signal efficiency measurements for W and top jet taggers

2.1 Introduction

The subject of this chapter is the determination of the signal efficiency of W and top taggers in ATLAS Run 2 Monte Carlo and Data. The resulting Data/MC scale factor is used to calibrate associated uncertainties in measurements that make use of these taggers. The studies presented here were carried out in 2018, shortly following the publication of ref. [116]. The studies provide a deeper analysis of the uncertainties associated with these measurement and attempt to improve the results using updated recommendations for jet reconstruction and tagger evaluation. These subsequent developments include testing the effects of new jet calibrations, alternative definitions of the tagger "truth" labels. A new measurement of the tagger efficiency from the larger 2015-2017 dataset was obtained and is presented in the ATLAS conference note ref. [117]. The tagger efficiency and scale factor measurements are part of the tagger recommendations that are followed by ATLAS analyses using W and top taggers. Since the recommendations had to be ready in good time before the first analyses using the full Run 2 dataset were published, the 2018 data was not used in the updated measurement of the tagger efficiencies and scale factors.

2.2 Measuring the signal efficiency scale factor in data

A description of the measurement of the W and top tagger signal efficiencies in early Run 2 data (collected in 2015 and 2016) can be found in ref. [116]. By selecting large- R

topocluster jets in semi-leptonic $t\bar{t}$ decays, a sufficiently high purity of jets containing tops and W -bosons can be achieved, such that the performance of taggers can be evaluated in data. By evaluating the tagging efficiency in both Data and Monte Carlo event, one can extract a Data/MC *scale factor*. These scale factors can then be used to calculate a systematic uncertainty associated with each tagger or to correct, in situ, tagging efficiencies that have been estimated from MC.

The signal efficiency for MC events is defined as

$$\epsilon_{Data}^{sig} = \frac{N_{signal}^{tagged}}{N_{signal}^{tagged} + N_{signal}^{not\,tagged}} \quad (2.1)$$

The signal efficiency in Data is defined as

$$\epsilon_{Data}^{sig} = \frac{N_{fitted\,signal}^{tagged}}{N_{fitted\,signal}^{tagged} + N_{fitted\,signal}^{not\,tagged}} \quad (2.2)$$

The signal efficiency scale factor is defined as

$$SF_{Data/MC} = \frac{\epsilon_{Data}^{sig}}{\epsilon_{MC}^{sig}} \quad (2.3)$$

2.2.1 Data and Monte Carlo samples

The tagger efficiency studies were carried out using a dataset from early Run 2 of ATLAS, collected during 2015, 2016 and 2017 in $p - p$ collisions at $\sqrt{s} = 13$ TeV. The total integrated luminosity is about 36fb^{-1} for the 2015-2016 dataset and about 44fb^{-1} for the 2017 data. The dataset comprises only events from data periods during which all subsystems of the detector were fully operational. Events in the data must have at least one primary vertex, reconstructed from two or more ID tracks, that is consistent with the LHC beam spot.

The nominal $t\bar{t}$ Monte Carlo sample is simulated using the PowhegBox [118, 43, 119, 120] matrix element generator interfaced with Pythia 8.230 [38]. This is an updated version of the configuration used in ref. [116] which made use of Pythia 6 [37] to simulate hadronisation and parton showers. Background events from W/Z + jets and diboson

processes are simulated using Sherpa v. 2.2.1 [47]. Modelling uncertainties are estimated using alternative $t\bar{t}$ and $W/Z + jets$ events. Alternative modelling of parton showers and hadronisation is achieved using Herwig 7.04 [34], while alternative matrix elements are generated using MadGraph5_aMC@NLO v. 2.6.0 [46]. The effect of multiple pp interactions, or pileup, is modelled by overlaying minimum-bias events onto the simulations. These are generated using Pythia v. 8.186 [121] with the A3 tune [122]. Lastly, interactions of the MC particles with the material of the detector are simulated using Geant 4 [123] with an up-to-date model of the ATLAS detector geometry in Run 2.

2.2.2 Background estimation: template fits

The contributions to the number of events from non- $t\bar{t}$ backgrounds is estimated by performing a template fit on the Monte Carlo signal and background distributions of the jet mass. The jet mass observable used in this analysis is the *combined jet mass* m_{comb} [124], which is a linear combination of the jet mass reconstructed from calorimeter topoclusters m_{calo} and the track-assisted mass measurement m_{TA} , which is calculated using the momenta of tracks that have been matched to the large- R jet through ghost-association (see section 1.4.3.4 for a definition). Since only charged particles leave tracks in the inner detector, the momentum of neutral particles is compensated for by scaling the 4-momenta of the tracks by the jet p_T as measured in the calorimeter. By using m_{TA} in the calculation of the jet mass, the effect of the limited spatial resolution of the topoclusters is reduced and a sharper mass peak is observed than for the cluster-based measurement alone. This improvement in the mass resolution is greatest in high p_T events where topoclusters can grow very large.

Two separate sets of combined mass histograms are produced: one for events that pass the tagger and one for events that fail. For Monte Carlo events, separate histograms of the tagged and untagged distributions are filled for the W or top signal jets (as defined in section 2.2.3.5), the background $t\bar{t}$ jets and non- $t\bar{t}$ sources of background. Small background contributions with similar shapes are combined into larger background histograms, or "templates". A linear combination of the background templates and the signal histogram are then fitted to the data using a least-squares fit. The fit is performed simultaneously on the tagged and untagged distributions while conserving the total number of

events $N^{tot} = N^{tagged} + N^{not\,tagged}$. The integral of the fitted signal template N_{signal} is then used to calculate the signal efficiency in Data, as per equation 2.2.

2.2.3 Selections

To select $t\bar{t}$ events where one of the tops decays hadronically and the other to leptons, events are required to pass an inclusive muon trigger. The electron channel, used in ref. [116], was not utilized in these studies due to large associated background modelling uncertainties. The names of the triggers for the 2015 and 2016 data periods can be found in ref. [116].

2.2.3.1 Leptons

Events are required to contain exactly one muon candidate with $p_T > 30\text{GeV}$ that has been matched to its corresponding trigger-level object. The muons are required to have $|\eta| < 2.5$ and must satisfy the "medium" quality requirement as defined in ref. [125].

2.2.3.2 Small- R jets

Small- R jets are reconstructed using anti- k_T with $R = 0.4$ from topoclusters calibrated to EM scale. The jet energies are calibrated to remove the effects of pileup and the jet energy scale is calibrated in situ using both Monte Carlo simulations and data [126, 127]. Pileup-suppression is achieved by requiring that tracks associated to low- p_T jets are associated to the primary vertex.

In order to be considered, small- R jets are required to have $p_T > 20\text{GeV}$ and $|\eta| < 2.5$. At least one small- R jet is required to have $p_T > 25\text{GeV}$ and $\Delta R(lepton, jet) < 1.5$.

2.2.3.3 Missing E_T

Missing transverse energy is reconstructed from small- R jets, the single lepton and inner detector tracks which have not been associated with a small- R jet or the lepton. The missing transverse momentum must satisfy the cuts $E_T^{miss} > 20\text{GeV}$ and $E_T^{miss} + m_T^W > 60\text{GeV}$, where $m_T^W = \sqrt{2p_T^\ell E_T^{miss} (1 - \cos \Delta\phi(\vec{E}_T^{miss}, \ell))}$ is the transverse mass of the leptonically decaying W candidate.

2.2.3.4 *b*-jets

B-tagging is performed on jets reconstructed from tracks using the anti- k_T algorithm with $R = 0.2$. The algorithm used to tag the jets was the multivariate discriminant MV2 [128] with a 70% signal efficiency working point. The b-tagging efficiency was corrected in situ using scale factors that have been obtained from dileptonic $t\bar{t}$ decays [129].

2.2.3.5 Large- R jets

Large- R jets were reconstructed using the anti- $k - T$ algorithm with $R = 1.0$. Trimming with $R_{sub} = 0.2$ and $f_{cut} = 0.05$ is applied in order to remove soft contaminations. The large- R jets were calibrated by applying corrections to the jet energy scale, mass scale, energy resolution and mass resolution. A detailed description of the large- R jet calibrations used in these studies is presented in sections 2.4.1 and 2.4.2.

Only the highest p_T large- R jet in the event is considered for the tagger studies. To ensure that the jet axis matches the direction of the hadronically decaying top quark, the large- R jet is required to be separated by the lepton by an angular distance $\Delta R(lepton, ljet) > 1.5$ and azimuthal angle $\Delta\phi(lepton, ljet) > 2.3$.

In order to obtain large- R jets which are predominantly the result of either a top or W decay, the semi-leptonic $t\bar{t}$ events are further divided into sub-selections based on the p_T of the large- R jet and the presence of a b -jet in the vicinity of the large- R jet axis. The jet selections with corresponding cuts are defined as follows:

- Top jet candidate: $\Delta R(b - jet, ljet) < 1.0$ and $p_T^{ljet} > 350 \text{ GeV}$,
- W jet candidate: $\Delta R(b - jet, ljet) > 1.0$ and $p_T^{ljet} > 200 \text{ GeV}$.

2.2.4 Cut-based taggers

The basic ATLAS W and top taggers use cuts on a set of jet substructure variables to tag large- R jets. These cuts are optimised to provide a constant signal efficiency across a wide range of p_T . The choice of JSS observables is also optimised to give the highest possible background rejection within the optimised p_T range of the tagger. To simplify the optimisation of the cuts, the number of jet variables chosen is kept small. The W tagger uses the combined jet mass m_{comb} and one other JSS variable, while the top tagger

uses two JSS variables out of the list shown in Table 2.1. The values of the cuts are determined in Monte Carlo studies. To ensure a constant signal efficiency across the p_T range, continuous functions of p_T , typically polynomials of order $n > 4$, are fitted to the cut values required to obtain the desired signal efficiency. Due to the smooth dependence of the cut values on p_T , these taggers are also commonly referred to as *smoothed taggers*. For the W -tagger optimal performance is achieved using m_{comb} and the ECF double ratio D_2 , while for the top tagger the optimal choice of JSS variables is m_{comb} and the 3-2 subjettiness ratio τ_{32} .

Working points of $\epsilon_{sig} = 50\%$ and $\epsilon_{sig} = 80\%$ are chosen for the W and top tagger respectively. This choice of working points is largely based on the needs of BSM searches which typically require higher background rejections for processes involving W -jets than for those involving top jets.

One highly performant cut-based tagger is the 3-variable W -tagger. This tagger uses the variables m_{comb} , D_2 (defined in section 1.5.2) and N_{trk} to identify W jets. Three p_T -dependent cuts are made: two on m_{comb} surrounding the mass peak of the W , and one on D_2 . The optimal m_{comb} and D_2 cuts for each p_T and the fitted smooth functions are shown in figure 2.1 for the $\epsilon_{sig} = 50\%$ working point. An additional cut is made on the number of ghost-associated [97] tracks to the ungroomed large- R jet, n_{trk} . This cut was introduced after the publication of ref. [116] as it has been shown to improve tagging performance in diboson resonance searches [130]. As QCD jets tend to be composed of more hadrons, one obtains a higher purity of W jets by putting an upper limit on the number of tracks within the jet. The cut on n_{trk} is kept fixed for all p_T and is optimised for each working point: at 50% the cut value is $n_{trk} < 26$.

2.2.5 DNN taggers

An innovation in boosted object tagging techniques in Run 2 of ATLAS is the use of machine learning algorithms, such as Deep Neural Networks (DNNs) and Boosted Decision Trees (BDTs), to define the JSS variable selections. These multivariate classifiers take a collection of JSS variables and combine them in a highly non-linear way to produce a classifier score between 0 and 1. The classifiers are trained on labelled Monte Carlo samples. A complete description of the training process of the DNN JSS taggers can be

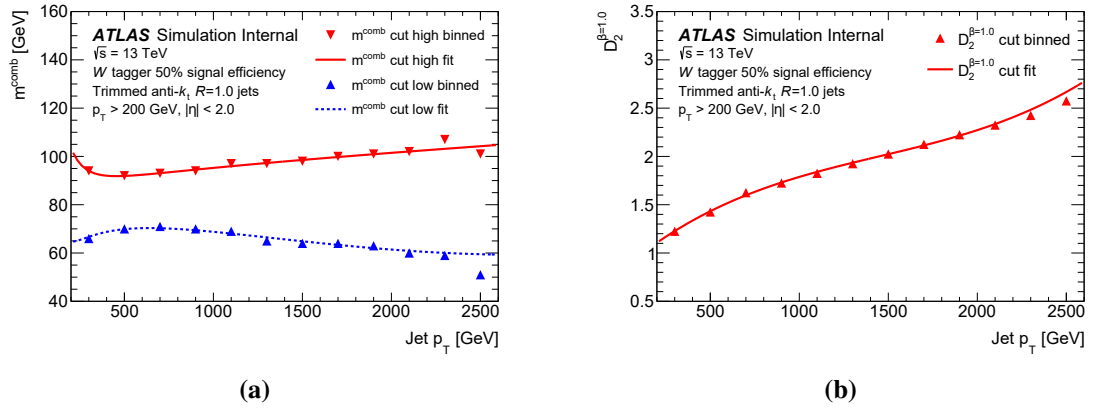


Figure 2.1: Optimal m_{comb} (a) and D_2 (b) cuts as a function of p_T . From ref. [117].

found in ref. [131]. The training set comprises $W' \rightarrow WZ$ and $Z' \rightarrow t\bar{t}$ samples as a source of W and top signal jets and QCD dijet events are used as background.

The choice of input variables to the DNN has been optimised by evaluating the performance of a number of DNNs, each trained with a different collection of input variables. A set of 10 input variables are chosen for the W and top DNN tagger. The choice of input variables is different for each tagger and is optimised by retraining the DNN with different collections of variables and choosing the combination that gives the greatest discrimination power. The optimal choice of input variables for the W and top DNN are shown in Table 2.1.

Table 2.1: Summary of observables used as inputs to the W and top smoothed and DNN taggers.

Observable description	Variable	cut-based W	DNN W	DNN top
transverse momentum	p_T	Y	Y	Y
combined jet mass	m_{comb}	Y	Y	Y
ECF ratio	e_3	N	N	Y
ECF ratio	C_2	N	Y	Y
ECF ratio	D_2	Y	Y	Y
1-Subjettiness	τ_1	N	N	Y
2-Subjettiness	τ_2	N	N	Y
3-Subjettiness	τ_3	N	N	Y
2-1-Subjettiness ratio	τ_{21}	N	Y	Y
3-2-Subjettiness ratio	τ_{32}	N	N	Y
Fox-Wolfram moment	R_2^{FW}	N	Y	N
Planar flow	$\mathcal{P}$	N	Y	N
Angularity	a_3	N	Y	N
Aplanarity	A	N	Y	N
Energy sharing	z_{cut}	N	Y	N
1-2 k_T splitting scale	$\sqrt{d_{12}}$	N	Y	Y
2-3 k_T splitting scale	$\sqrt{d_{23}}$	N	N	Y
k_T splitting ΔR	$KtDR$	N	Y	N
k_T splitting inv. mass	Q_W	N	N	Y
Number of GA tracks	N_{trk}	Y	N	N

2.3 Systematic uncertainties

2.3.1 Sources of systematic uncertainty

2.3.1.1 Modelling systematics

The dominant sources of systematic uncertainty in the measurement of W and top tagger scale factors are related to modelling in Monte Carlo simulation. The modelling systematics are determined using a collection of alternative $t\bar{t}$ Monte Carlo samples. New template fits of the m_{comb} distributions are performed for each alternative sample. A new set of signal efficiencies and scale factors is computed for the alternative samples and the difference to the nominal Powheg+Pythia 8 sample is taken as the size of the uncertainty.

The uncertainty in the modelling of the hard scatter process (labelled *MatrixElement*) is evaluated using the aMC@NLO+Pythia 8 sample, where the matrix element has been calculated using MadGraph5_aMC@NLO [132].

Uncertainty associated with the parton showering and hadronisation model, as well as the matching between the matrix element and the parton shower, (labelled *hadronisation*) is estimated using the Powheg+Herwig 7 sample.

Uncertainty in the amount of initial and final state QCD radiation (labelled *radiation*) is estimated from two alternative Powheg+Pythia 8 samples, radHi and radLo, where the parameters controlling the amount of radiation are scaled by either a factor of 0.5 or 2.0. These parameters are h_{damp} , a parameter of Pythia 8 controlling the amount of high- p_T radiation that the $t\bar{t}$ system recoils against, the factorisation and renormalisation scales, as well as the Var3c parameter of the A14 tune.

2.3.1.2 Large- R jet systematics

The uncertainties associated with the large- R jets are related to the jet substructure observables (JSS), jet mass-scale (JMS), jet mass resolution (JMR), jet energy scale (JES) and jet energy resolution (JER). The JMS, JES and JSS uncertainties are determined using the R_{trk} method [133]. In this method, the quantity $\langle X^{calo\ jet} / X^{track\ jet} \rangle_{Data} / \langle X^{calo\ jet} / X^{track\ jet} \rangle_{MC}$ is computed, where X is either the jet mass, jet p_T or another jet observable. The quantities denoted *track jet* are determined from $R = 1.0$ track jets that have been matched to the calorimeter topocluster jets by a

cut on the $\eta - \phi$ distance. By computing the deviations in the jet mass jet p_T double ratios from 1, one obtains the uncertainties associated with the Monte Carlo derived JES and JMS calibrations.

The uncertainty in the resolution of the detector response for the jet mass, jet energy and other jet observables is estimated from the in situ measurements [134] of the JER and JMR. The conservative estimates used in the ATLAS recommendations are 2% in jet p_T , 20% in jet mass and 15% for other jet observables.

2.3.1.3 Small- R jet systematics

The uncertainty in the small- R jet momentum scale is determined from the in situ calibrations [127]. Contributions to this uncertainty come from sources such as MC mis-modelling, event topology and pileup. There is also a number of uncertainties associated with the η inter-calibration of the JES measurement. As for the b -tagging uncertainties, a global reduction of the number of systematics is performed using eigenvector decompositions of the jet p_T correlation matrix.

2.3.1.4 b -tagging systematics

The flavour tagging uncertainty associated with the MV2 [128] b -tagging algorithm is derived in situ from fully leptonic $t\bar{t}$ decays [129]. The b -tagging efficiency in data is estimated using two different methods: *tag and probe* and the combinatorial *likelihood* method.

Systematic uncertainties are evaluated for a large number of different sources, including a wide range of modelling uncertainties from alternative Monte Carlo samples, jet uncertainties and uncertainties associated with the normalisation factors of background processes that are estimated in MC. The number of systematic uncertainties is reduced using a statistical method in which a covariance matrix is constructed for the bins of the efficiency scale factor vs p_T histogram with each of the systematic variations applied. By calculating the eigenvectors of the sum of the covariance matrices, one obtains a set of histograms corresponding to a new set of statistical systematic variations. The number of new systematics is equivalent to the number of bins used in the efficiency scale factor histograms. The eigenvector systematics with the smallest contribution to the

total uncertainty can be neglected in analyses where the systematic uncertainty contributions from b -tagging are expected to be small.

2.4 Effect of in situ jet mass calibrations

2.4.1 Pre-2018 calibration of large- R jet energy and mass

Until the autumn of 2018, the ATLAS prescription for the calibration of large- R jets was based on the same procedure used to derive the jet energy and mass calibrations for small- R jets [135, 127]. These calibrations did not include in situ corrections to account for mismodelling in the Monte Carlo simulations that are used to derive the calibrations. These large- R jet calibrations are the same as were used in the published analysis [116].

In this calibration scheme, a jet energy scale is derived from Monte Carlo simulation in order to bring the reconstructed jet 4-momenta to the energy scale of particle-level jets. Corrections are applied to compensate for the detector response as a function of p_T and $|\eta|$. The energy scale is determined from Pythia 8 Monte Carlo simulated dijet events by matching reconstructed jets to particle level jets in $\eta - \phi$ space and then calculating the response $R_{JES} = E_{reco}/E_{truth}$. An isolation requirement is also applied to minimise contamination from outside the jet cone. The energy response for a given p_T or η is determined from a polynomial fit of R_{JES} as a function of $\ln(E_{reco})$. Finally, the calibrated detector level jet energy is obtained through a numerical inversion of the functional fit.

The large- R jet mass scale [124] is determined after the jets have been calibrated using the jet energy scale calibration. From there on, the calibration procedure is analogous to the JES determination. The mass response $R_{JMS} = m_{reco}/m_{truth}$ is extracted from Monte Carlo dijet events using isolated detector level jets that have been matched to particle-level jets. A Gaussian is fitted to the jet mass response, which is then inverted to obtain the final calibrated detector level jet mass.

2.4.2 In situ calibration of large- R jet energy and mass

Unlike the calibration procedure outlined in section 2.4.1 which is dependent on the accuracy of the modelling of the Pythia 8 MC generator, the *in situ* jet calibrations are able to correct for the effect of mismodelling. These calibrations are determined from real data by taking advantage of recoil-type event topologies, in which a large- R jet is produced

back-to-back with another well-reconstructed object or, in the case of the jet mass scale, the momentum measurement of inner detector tracks with the masses of known particles. A detailed description of the in situ large- R jet mass and energy calibration can be found in ref. [134].

The large- R jets, which have been constructed from topoclusters using anti- k_T with $R = 1.0$ and trimmed, have their energy response calibrated using response measurements obtained from γ +jets, Z +jets and multijet events in data. In each type of event, the p_T -balance between the large- R jet and the lepton, or a collection of small- R jets in the case of multijet events, is used to determine the energy response of the large- R jet. To compensate for the variations in energy response across different η , a relative inter-calibration is performed by computing the difference between the average momentum response in the central barrel region of the detector ($|\eta| < 0.8$) and the response of individual jets in the more forward region ($0.8 < |\eta| < 2.5$).

The jet mass response, or jet mass scale, is measured using two different methods. The first method is the R^{trk} method, which uses track p_T measurements to calculate the p_T response of the calorimeter topocluster based jets. The second method, which is also used to extract the jet mass resolution (JMR), relies on measuring the mass distributions of W s and tops in semi-leptonic $t\bar{t}$ events. A functional fit is then performed to match the mass spectrum in Monte Carlo particle-level jets to the reconstructed mass spectrum in data. This method is commonly known as forward folding. The ratios of the mass response R_{data}^m/R_{MC}^m resolution $\sigma_{data}^m/\sigma_{MC}^m$ are determined from the free parameters of the fit.

2.4.3 Signal efficiency and scale factors for different mass calibrations

2.4.3.1 Template fit results

Since the tagging efficiencies are derived from calibrated jet mass distributions, it is expected that the new mass calibration will have an effect on the signal efficiencies in both MC and data. The jet mass distributions in a single p_T bin for data and MC are

shown in figures 2.2 and 2.3 for the top and W DNN taggers. The MC distributions have been scaled by a factor determined by the template fit.

The quality of the template fits for each histogram is quantified by the goodness-of-fit metric χ^2/NDF . The χ^2 statistic for the histograms defined as

$$\sum_{i=1}^{NDF} \frac{(x_i - \tilde{x}_i)^2}{\sigma_i^2}, \quad (2.4)$$

where x_i and $\tilde{x}_i$ are the measured and fitted value in each bin i and σ_i^2 is the bin variance. The number of degrees of freedom, NDF , is the number of non-zero bins of the histogram. For a good fit, the value of χ^2/NDF is expected to be less than or close to 1, provided that the value of NDF is sufficiently large.

The values of χ^2/NDF for the m_{comb} histogram in each p_T bin are shown in tables 2.2 and 2.3 for the W and top DNN taggers with the calorimeter-based and in situ mass calibrations. As the number of bins in each histogram, NDF , is identical in each p_T bin, a lower χ^2/NDF generally denotes a better fit. For the 3-variable W tagger, a better fit is achieved using the in situ calibrations in 6 out of 8 p_T bins. For the W DNN tagger, this is also the case for 5 out of 8 p_T bins. For the top DNN tagger, the fit is improved using the in situ calibrations in all p_T bins but one. Hence, it can be concluded that the new large- R jet mass calibrations improve the reconstruction of the mass sufficiently that the measurement of the signal efficiency is also improved.

To comment more precisely on the goodness of the fit, one can calculate the value of the χ^2 -distribution with a given NDF at a desired p -value. In this context, the p -value is approximately the probability with which the fit is expected to agree with the data under repeated observations. The jet mass distributions for the tagged jets is generally narrower than for the untagged jets and is captured by histograms with a smaller number of non-zero bins. Therefore, there is a smaller number of degrees of freedom for the fits to these histograms. The histograms of the tagged mass distributions (labelled pass in tables 2.2 and 2.3) have a NDF between 5 and 18, while the untagged histograms (fail) have between 26 and 60. A very good fit with a p -value of 85% would have a value of χ^2/NDF between 0.40 and 0.66 for the tagged distributions and between 0.72 and 0.81 for the untagged distributions.

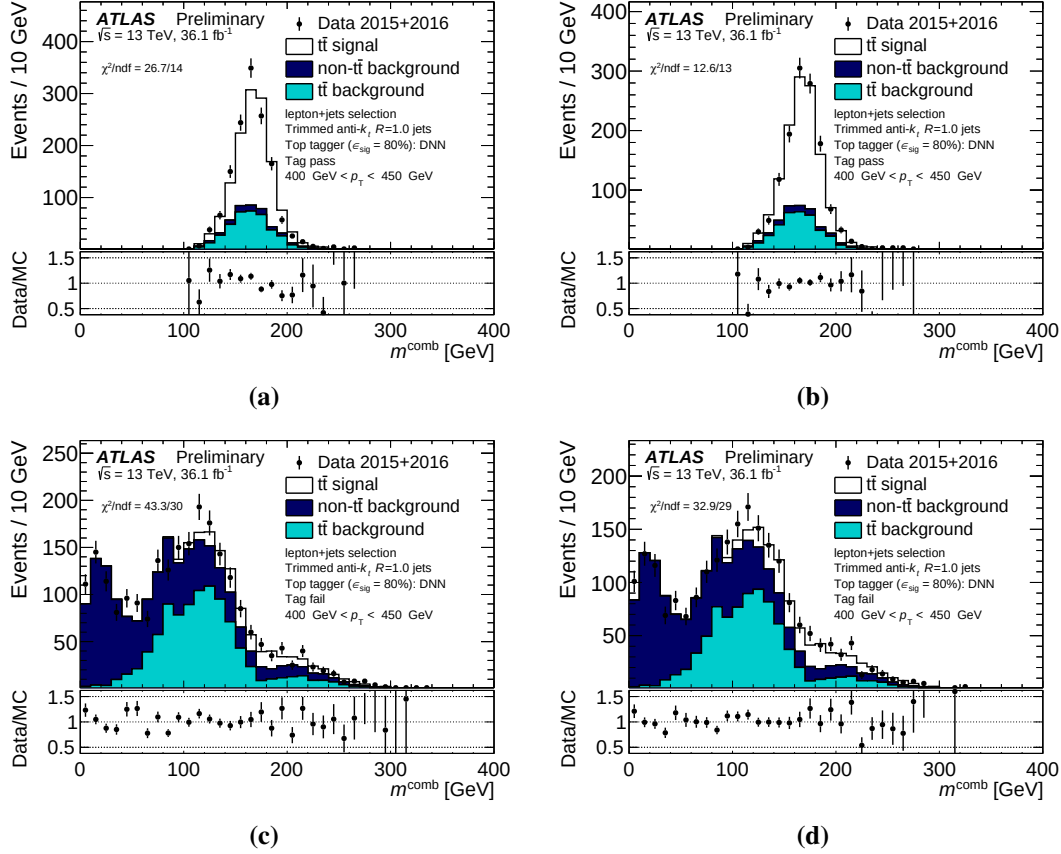


Figure 2.2: Measured and fitted MC m_{comb} distributions in the 400 – 450, GeV p_T -bin for signal jets tagged using the DNN top tagger. Tagged (a, b) and untagged (c, d) distributions are shown for the calorimeter-based (a, c) and in situ (b, d) large- R jet mass calibrations.

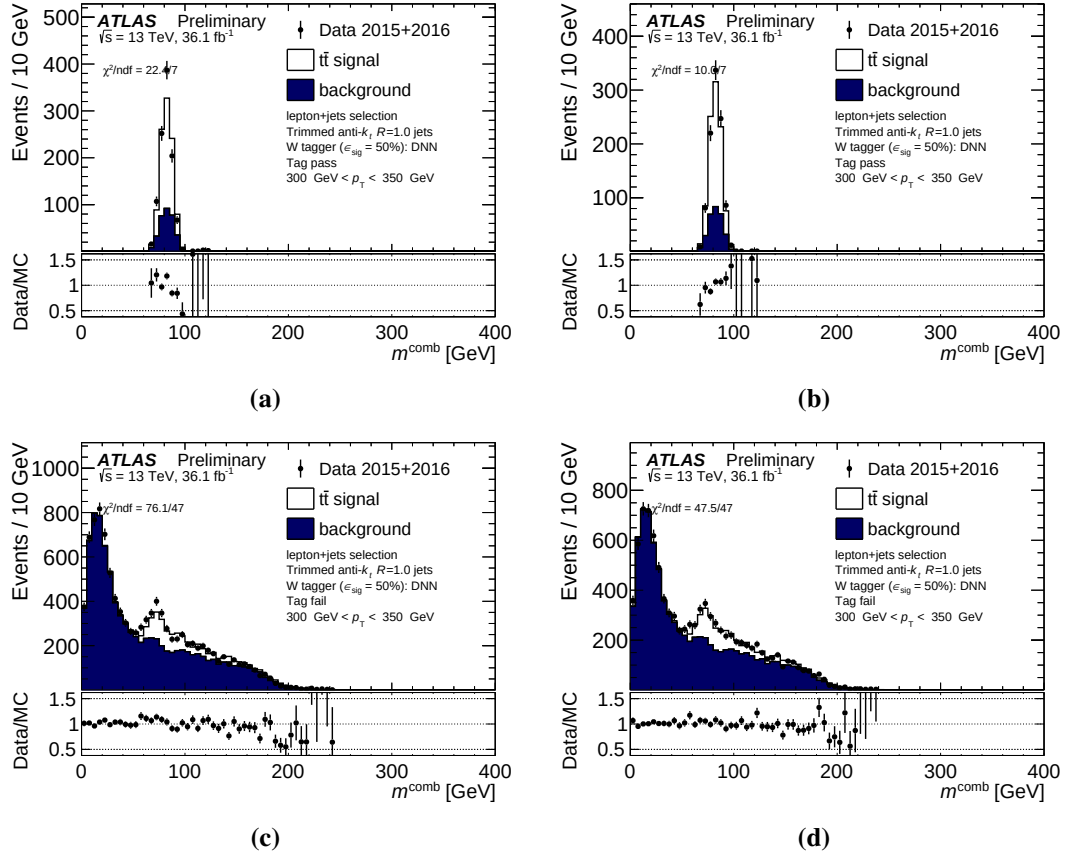


Figure 2.3: Measured and fitted MC m_{comb} distributions in the 300 – 350, GeV p_T -bin for signal jets tagged using the DNN W tagger. Distributions are shown for the old (a, c) and new (b, d) large- R jet calibrations.

Table 2.2: Quality of fit χ^2/NDF for the template fits of the m_{comb} distributions for the W taggers with calo mass and in-situ large- R jet calibrations. Two mass distributions, one for the tagged and one for the untagged jets, are fitted in each p_T bin. Lower χ^2/NDF values indicate a better fit.

p_T [GeV]	Tag	3-var W		W DNN	
		Calo Mass	In-Situ	Calo Mass	In-Situ
200-250	pass	4.80	2.40	2.53	3.63
	fail	1.91	1.38	2.06	1.55
250-300	pass	4.60	0.83	2.34	2.00
	fail	1.13	1.15	1.07	1.33
300-350	pass	4.47	0.28	3.20	1.43
	fail	1.51	1.04	1.62	1.01
350-600	pass	0.90	0.62	1.06	1.06
	fail	0.86	0.91	0.84	0.86

Table 2.3: Quality of fit χ^2/NDF for the template fits of the m_{comb} distributions for the W taggers with calo mass and in-situ large- R jet calibrations. Two mass distributions, one for the tagged and one for the untagged jets, are fitted in each p_T bin. Lower χ^2/NDF values indicate a better fit.

p_T [GeV]	Tag	top DNN	
		Calo Mass	In-Situ
350-400	pass	2.49	0.83
	fail	1.22	1.10
400-450	pass	1.91	0.97
	fail	1.44	1.13
450-500	pass	1.51	0.70
	fail	1.25	0.73
500-600	pass	0.95	0.65
	fail	1.16	0.92
600-800	pass	1.17	0.73
	fail	0.95	1.31
800-1000	pass	0.68	0.58
	fail	0.96	0.75

2.4.3.2 Signal efficiencies and scale factor

The signal efficiencies and corresponding scale factors for the calorimeter-based and in situ large- R jet mass calibrations are shown in figure 2.4. The propagated statistical uncertainties on the signal efficiencies in data and Monte Carlo are also shown. Thanks to the improvement in the quality of the fits with the new large- R jet calibrations, the estimated uncertainty on the signal efficiency in data is reduced in most bins. For the few bins where the quality of the template fit is worse with the in situ calibrations, such as the $600 < p_T < 800\text{GeV}$ bin for the top tagger, there is also a greater post-fit statistical uncertainty.

The signal efficiency in $t\bar{t}$ data and MC for the W DNN tagger is around 30 – 40% across the p_T range. This is significantly lower than the target efficiency of 50%. This discrepancy is due to the different event topologies that are used in training and testing the tagger. One notable difference between the W' training sample and the semileptonic $t\bar{t}$ decays is the polarisation state of the W boson. The W bosons from the BSM $W' \rightarrow WZ$ decays are exclusively longitudinally polarised, while the top decays are expected to produce W s with both longitudinal and transverse polarisations in proportion $\Gamma_L : \Gamma_T = 0.69 : 0.31$ [136]. The polarisation state of the W boson affects the kinematics of the decay partons. Due to conservation of momentum, transversely polarised W bosons are more likely to decay to a pair of quarks where one quark carries a larger fraction of the W s momentum. For longitudinally polarised W bosons, momentum is shared more evenly between the two quarks. One can therefore expect the W jets from longitudinally polarised W bosons to have a more pronounced two-pronged structure, which can make them easier to distinguish for substructure-based taggers. This effect has been observed in previous W tagging studies [137, 138]. As the observed signal efficiency in $t\bar{t}$ is approximately 70% of the target signal efficiency, it is likely that the W DNN tagger correctly classifies the longitudinally polarised W boson jets at a much higher rate than the transversely polarised ones.

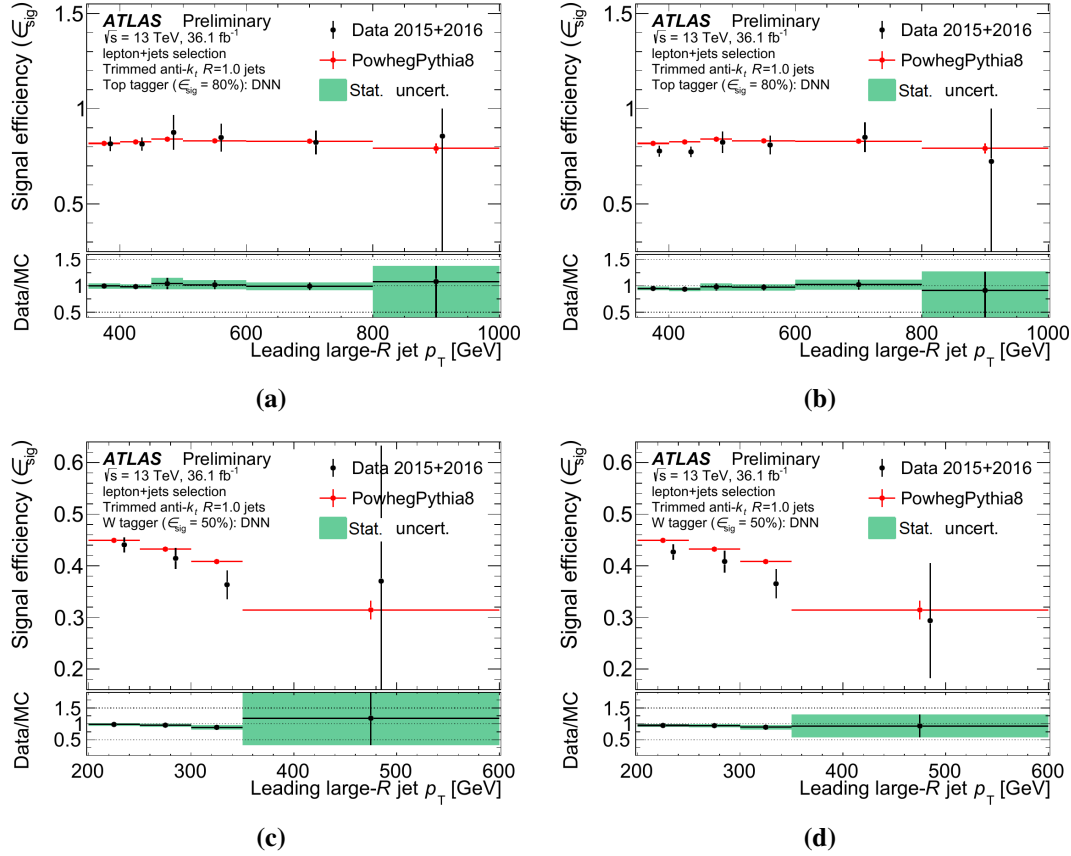


Figure 2.4: Signal efficiencies and Data/MC scale factors for the top (a,b) and W (c,d) DNN taggers with the calorimeter mass based (a,c) and in situ (b,d) large- R jet calibrations. Statistical uncertainties on the Data and Monte Carlo are shown as error bars on the Data and MC efficiencies. The statistical uncertainty on the Data/MC scale factor is shown as a green band in the bottom ratio plots.

2.5 Alternative Monte Carlo truth jet definitions

The truth-definition used to evaluate tagger performance in Monte Carlo had, up until the publication of ref. [116], been based on a matching between the truth W boson or top and the truth-level large- R jets. From the beginning of Run 2, a "containment" requirement, under which the decay partons of the heavy particle must be contained within the large- R detector level jet, was added to the truth definition of top and W jets. The matching between the truth-level partons and jets introduces some theoretical problems. The parton showering process, which links the decay of the heavy particle to the final-state jets, is a non-perturbative process and is modelled differently across different generators. Since the kinematics of individual partons are not observable, there is no consistent definition of parton 4-momenta between different Monte Carlo generators. As modelling uncertainties on the measured tagger efficiencies are derived using alternative Monte Carlo generators, it was proposed that these could be reduced by introducing hadron-level quantities into the containment requirement. It was therefore decided that a new containment requirement, using the kinematics of truth hadrons rather than partons, should be introduced. The new truth-labelling scheme re-defines the containment requirement based on the presence (or absence) of B hadrons within the jets. Although the new truth labelling scheme introduces hadron-level quantities, it is not a purely hadron-level definition as it still requires the truth jet to be matched to the generator-level W boson or top quark. The matching between the truth heavy particles and jets was kept under the new truth jet definition to retain some consistency with the old definition and the Run 1 truth jet definitions which did not impose any containment requirement. Hence, some model dependence is still expected under the new labelling scheme.

2.5.1 Early Run 2 truth jet definition

The truth definition that was originally used to evaluate the tagger performance is defined by first matching the ungroomed truth jets to the W or top quark in the generator record. To determine whether top or W is fully, or only partially, contained within the jet, the parton-level decay products of the respective particle are matched to the reconstructed jet. The truth jets, which are built from stable truth particles using the anti- k_T algorithm with $R = 1.0$, are first matched to the reconstruction level jets by requiring

$\Delta R(j_{truth}, j_{reco}) < 0.75$. The truth jets, in turn, are matched to the truth top or W particle using the criterion $\Delta R(j_{truth}, particle) < 0.75$. The decay partons of the W boson or top quark are obtained from the truth record and are matched to the reconstruction-level jet by requiring $\Delta R(j_{reco}, decay\ parton) < 0.75$ for all decay partons.

In the analysis of tagger efficiencies, two different signal topologies are defined based on the above definitions:

- W : Truth jet is matched to the W boson. Both decay partons of the W boson are contained within the detector level jet.
- Top (fully contained): Truth jet is matched to the top quark. All decay products of the top quark, i.e. the decay products of the W boson and the b quark, are contained within the detector level jet.

For the top jets, top quarks whose decay products are not fully contained within $\Delta R < 0.75$ of the jet are considered part of the background. These background jets can be further divided into the following categories:

- Top (W): Truth jet is matched to the top quark. Only the decay partons from the W are contained inside the detector level jet.
- Top (b): Truth jet is matched to the top quark. Only the b quark is contained within the detector level jet.
- Top (qb): Truth jet is matched to the top quark. Only one quark from the W decay plus the b quark are contained within the detector level jet.

In figures 2.2 and 2.7, where the mass distributions for true signal and background jets are shown, the above three categories of partially contained top jets are labelled $t\bar{t}$ background.

2.5.2 Full Run 2 truth jet definition

The new truth definition is instead based on the presence of b-hadrons within the detector level jet. To be considered a contained top or W jet, the presence or absence of a b-hadron within the detector level is required in addition to a cut on the large- R jet mass and

distance to the top-quark or W . Jets within $\Delta R < 0.75$ of a truth W with no associated B -hadron in the range $50\text{ GeV} < m < 110\text{ GeV}$ are considered W -jets. Jets within $\Delta R < 0.75$ of a truth top with $m > 140\text{ GeV}$ and at least one associated B -hadron are considered top-jets. Ghost-association (see section 1.4.3.4 for a definition) is used to determine whether the B -hadron is associated with the jet. The definitions of the fully contained, top (W), top (b) and top (qb) jet topologies are the same for the new truth definitions, with the exception that every requirement for a contained b -quark is replaced by one for a ghost-associated B -hadron.

Comparisons between the old and new truth jet definitions are shown in figure 2.5. Figures 2.5a and 2.5b show mass distributions for truth jets with a B -hadron match or a B -hadron veto are compared to truth jets containing either all top decay products (qqb), two light quarks (qq) or a b -quark and a light quark (qb). We can see that the new top jet definition based on the matched B hadron differs significantly from the old fully contained top definition and includes a significant number of events that were classified as partially contained top quarks under the old definition. However, at high p_T where all top decay products are likely to be contained within the jet, the old fully-contained and the new top definition are much more similar. In figure 2.5c the mass distribution for jets matching the old fully contained top definition is shown for the Powheg+Pythia, Powheg+Herwig and aMC@NLO+Pythia MC generators. The corresponding distributions for the new top jet definition are shown in figure 2.5c. The differences in the shape of the top jet mass distributions between the various MC generators are no greater for the new truth definition than they were for the old one.

2.5.3 Effect of new truth jet definitions on the signal efficiency

Using the new truth jet definitions, new sets of m_{comb} histograms are filled for each p_T -bin. The new truth definition significantly changes the proportion of signal and background MC events in each bin. Because the estimation of the signal efficiency in data involves performing a new template fit using the new signal and background distributions, the data estimate of the signal efficiency also changes with the new truth definition. The following section presents a study of the impact of the new truth jet definition on the

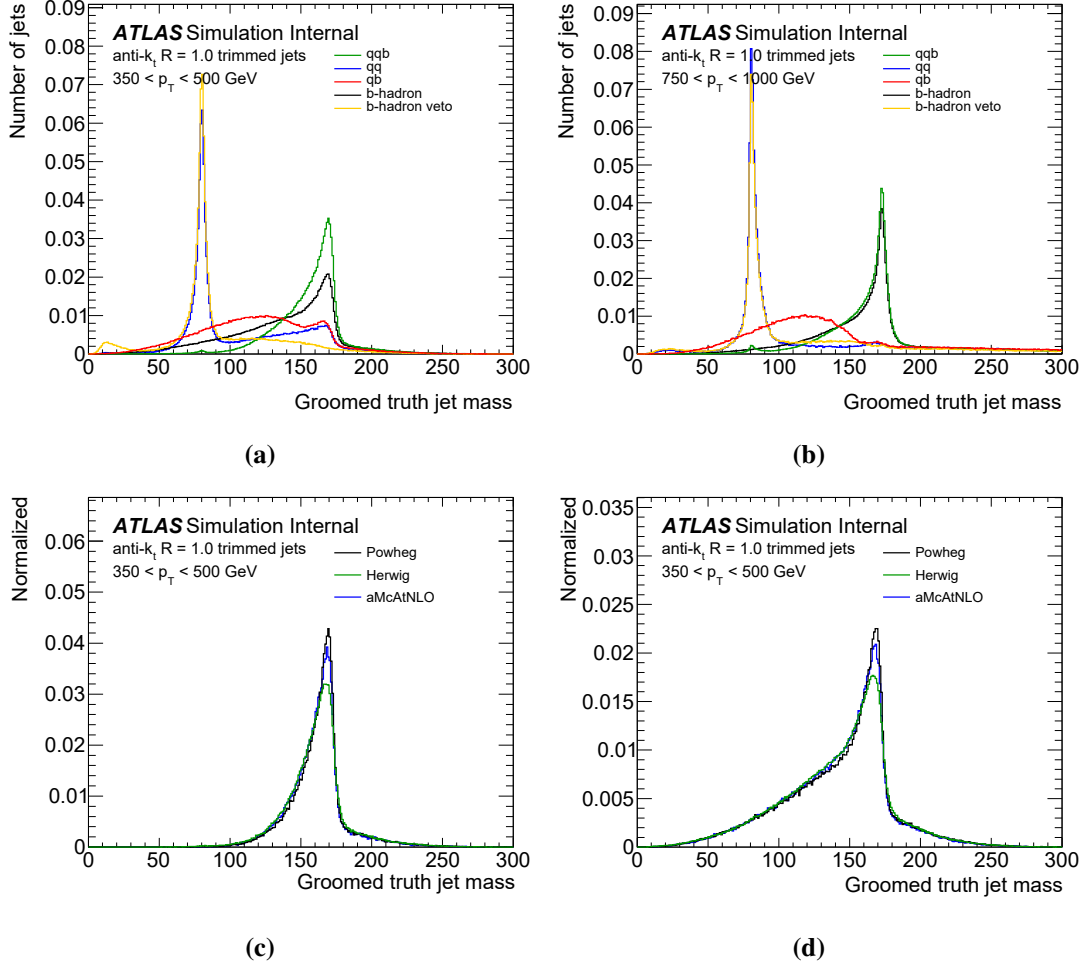


Figure 2.5: Large- R jet mass distributions with old and new truth definitions are shown in (a) for lower jet p_T and in (b) for higher jet p_T . The old fully contained top jet definition is labelled qqb , while the W jet definition is labelled qq . The new truth top and W jets correspond to the B -hadron match and veto distributions respectively. The bottom two plots shown a comparison of the jet mass distributions in the top jet selection with the old (c) and new (d) truth definitions for a variety of Monte Carlo generators. These figures were produced by C. Delitzsch in a preparatory study for ref. [117].

signal efficiency and scale factor using a partial sample containing only 2017 data and MC.

The m_{comb} distributions for tagged and untagged jets in a sample p_T bin, with the old and new truth definitions are shown for the W DNN tagger in figure 2.6 and for the top DNN tagger in figure 2.7. The measured distributions in 2017 data as well as the fitted signal and background MC distributions are shown. With the new truth definition, a somewhat larger portion of MC events are labelled as signal than under the old truth definitions. This is true for both the W and the top tagger and can be clearly seen in the tagged distributions in figures 2.6 and 2.7.

Figure 2.8 shows the signal efficiency and scale factors for 2017 data and MC with the old and new MC truth definitions. Statistical uncertainties on the signal efficiencies and scale factor are also shown. For the W DNN tagger, the signal efficiency with the new truth definition is around 10% lower at $p_T < 600 \text{ GeV}$ and unchanged at higher p_T . For the top-DNN tagger, the change in the efficiency is about a 10% downward in the lowest bin at $200 \text{ GeV} < p_T < 200 < \text{GeV}$, while for higher p_T there is a higher signal efficiency, up to 25% more at $p_T > 400 \text{ GeV}$. With the new truth jet definition the goodness of fit of the template fits is improved for most p_T bins, which is reflected in the smaller uncertainties in figure 2.8. One possible explanation for the improvement to the goodness of fit is the increase in the number of events classified as signal under the new truth labelling scheme. With the new signal distributions, the template fit is able to fit the mass peaks of the W boson and the top quark better than with the old ones.

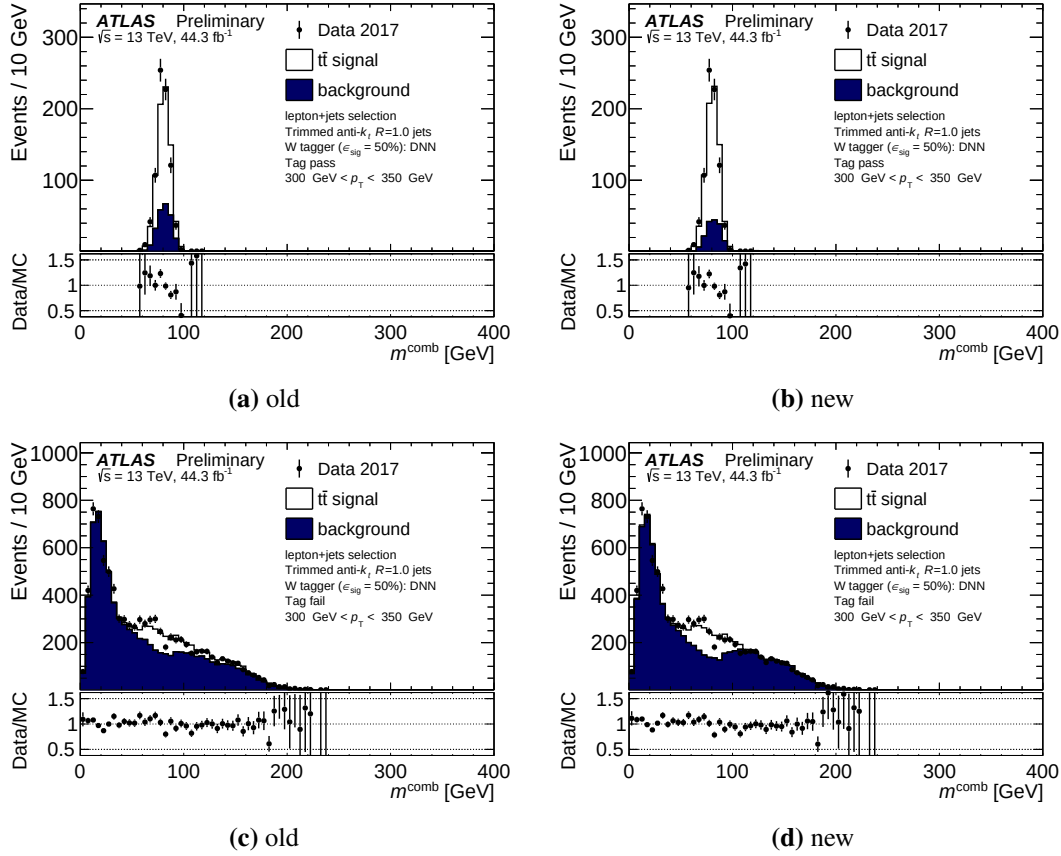


Figure 2.6: Mass distributions for tagged (a,b) and untagged (c,d) jets with the W -DNN tagger using the old (a,c) and new (b,d) truth definitions. The signal and background Monte Carlo distributions shown have been re-scaled in accordance with the template fit.

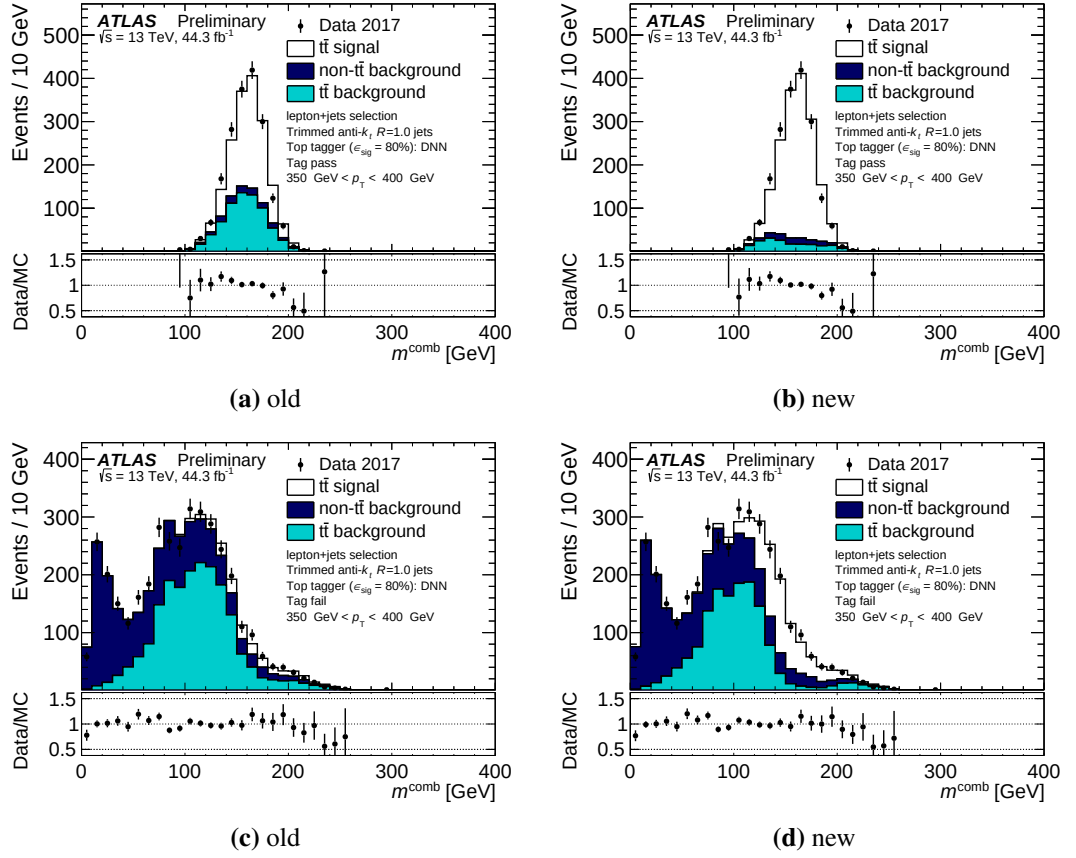


Figure 2.7: Mass distributions for tagged (a,b) and untagged (c,d) jets with the top-DNN tagger using the old (a,c) and new (b,d) truth definitions. The signal and background Monte Carlo distributions shown have been re-scaled in accordance with the template fit.

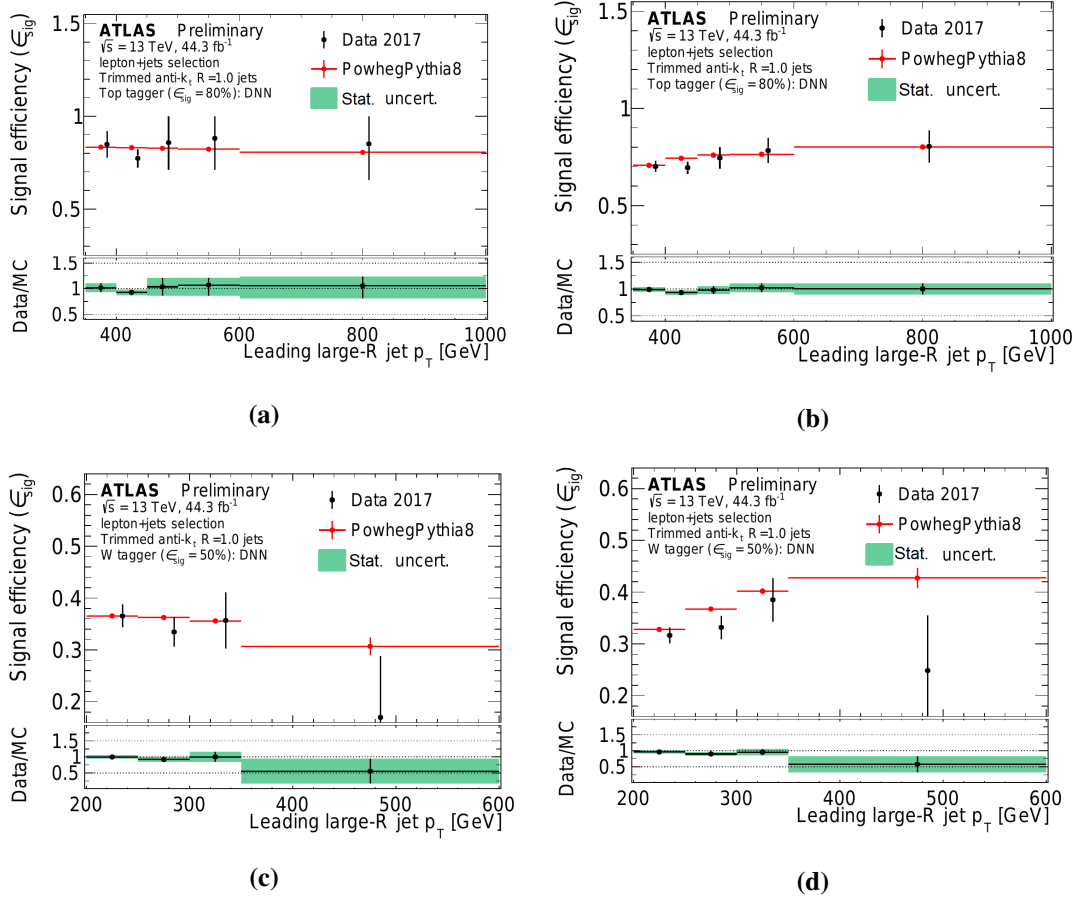


Figure 2.8: Signal efficiency scale factors for the top (a,b) and W (c,d) DNN taggers for the old (a,c) and new (b,d) truth definitions. The scale-factors are evaluated in 2017 data and Monte Carlo. Statistical uncertainties on the Data and Monte Carlo are shown as error bars on the Data and MC efficiencies. The statistical uncertainty on the Data/MC scale factor is shown as a green band in the bottom ratio plots.

2.6 Conclusion

The signal efficiencies in Data and Monte Carlo for the 2015-2017 Run 2 data is presented in the ATLAS PUB-note ref. [117]. The 2015-2017 dataset, at 80.5 fb^{-1} is more than twice the size of the 2015-2016 dataset used in ref. [116]. The Monte Carlo generators have also been updated to the latest versions, notably Pythia 6 has been replaced by Pythia 8. The latest large- R jet mass calibrations and the new truth labelling convention, whose effects were studied in this chapter, are also used.

The in situ jet mass calibrations for $R = 1.0$ jets improves the modelling of the jet mass. It was observed that the template fits that are used to calculate the tagger are also improved as a result. This leads to smaller uncertainties on the measured and predicted signal efficiencies.

The impact of a new truth-labelling scheme, which includes final state hadrons in the definition of top and W jets, on the signal efficiencies was investigated. It was found that the new truth labels change the proportions of signal and background events, leading to shifts in the measured and predicted efficiencies. The goodness of the template fit which is used to derive the measured signal efficiency is also changed as the shape of the signal and background templates changes with the new truth labelling. It is found that, in most p_T bins, the goodness-of-fit is improved leading to smaller uncertainties on the measured signal efficiency.

2.7 Outlook

In addition to the developments presented in this chapter, a number of other changes have been made to the W and top tagger efficiencies in Run 2. The W DNN tagger, which was studied extensively in ref. [116] and subsequent studies in the 2015-2016 data, was superseded by the 3-variable cut-based tagger, introduced in section 2.2.4. With the addition of the cut on n_{trk} , the 3-variable tagger outperforms the W DNN tagger at the $\epsilon_{sig} = 50\%$ working point by achieving a higher peak background rejection in the p_T range studied. This can be seen in the figures presented in appendix A.

The signal efficiencies and scale factors for the 2015-2017 data period are shown in figure 2.9 for the best performing top and W tagger: the top DNN and the 3-variable W tagger. Modelling uncertainties are reduced compared to the published results in ref. [116] thanks to the updated MC samples and jet calibrations. For the cut-based 3-variable tagger, the total uncertainty is greater in the three lowest p_T bins than for the W DNN tagger. This effect was already observed in in ref. [116] for the simpler 2-variable $m + D_2$ tagger, and is largely due to the sensitivity of the m and D_2 variables to the choice of the parton shower model.

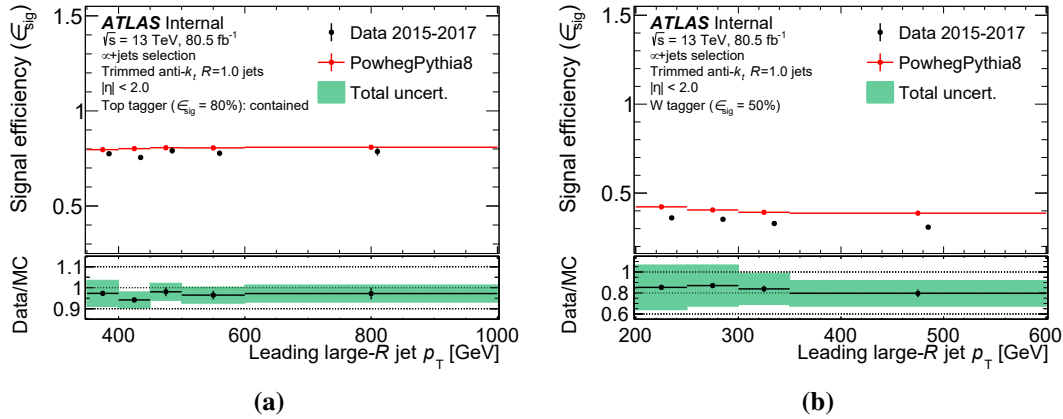


Figure 2.9: Signal efficiencies and scale factors derived from the full 2015-2017 data for the top DNN tagger (a) and 3-variable W tagger (c). Statistical uncertainties on the Data and Monte Carlo signal efficiencies are shown as error bars. The green band in the ratio plot shows the total uncertainty on the scale factor, including both statistical and systematic uncertainties. From ref. [117].

Chapter 3

Measurement of the Lund Jet Plane in dijet events

3.1 Introduction

The Lund Jet Plane was measured by ATLAS in dijet events in 2020 [115]. This was the first ever measurement of the LJP observable and it generated much interest in both the experimental and theoretical jet physics community. The LJP Measurement can help improve modelling by providing tunes for MC generators and ultimately provides new insights about QCD radiation within jets. An introduction to the concept of the Lund Jet Plane and its theoretical background can be found in Section 1.7. This chapter contains a summary of the method and results published in ref. [115], as well as more detailed descriptions of specific contributions to this measurement. A detailed description of all the studies carried out in the analysis are documented in the associated ATLAS internal note. The final section of this chapter, section 3.5, provides a short description of publications that have made use of the LJP measurement, either by using its results directly to improve theoretical understanding, or by applying its methods in new measurements.

3.1.1 Data and Monte Carlo samples

The measurement uses the whole of the Run 2 ATLAS dataset which was collected during 2015-2018 at $\sqrt{s} = 13\text{ TeV}$ and amounts to a total integrated luminosity of 139 fb^{-1} . Events are required to pass the lowest unprescaled jet trigger for each of the different data periods. The names of these triggers are `HLT_j360` and `HLT_j380` for the 2015

and 2016 data periods respectively and `HLT_j420` for 2017-2018 [139]. The suffix `jXXX` in the trigger name denotes the trigger jet p_T . A detector-level cut on jet p_T is applied so that only events on the trigger plateau are used. The dataset contains a very large number of jets, approximately 120 million, after selections have been applied.

A selection of different Monte Carlo generators are used in order to assess the modelling performance for the various physical processes that dominate in different sections of the Lund Jet Plane. MC generators with varying Matrix Element, Parton Showering and Hadronisation models are compared against the data and used to assess modelling uncertainties. The following MC generator configurations are used:

Pythia 8.230: Pythia 8.230 [38] with the NNPDF2.3 LO set of PDFs [140] and the A14 tune [141] using Pythia’s own LO matrix elements.

Powheg+Pythia: Same configuration as Pythia 8.230, but matrix elements are generated using `PowhegBox v2` [118, 43, 119, 120].

Sherpa 2.2.5 (AHADIC): Sherpa 2.2.5 [47] with the CT14NNLO PDF set [142] including $2 \rightarrow 2$ Matrix Elements and using Sherpa’s own AHADIC hadronisation model.

Sherpa 2.2.5 (String): Same configuration as for Sherpa 2.2.5 (AHADIC) but using the Lund string hadronisation model introduced in Pythia 6.4 [37].

Herwig 7 (Ang. ord.): Herwig 7.1.3 [34, 143] with the MMHT2014NLO PDF set [144] and using an angular ordered parton shower model.

Herwig 7 (Dipole): Same configuration as Herwig 7 (Ang. ord.) but with an alternative parton shower and hadronisation model, a dipole parton shower model and cluster hadronisation.

All Monte Carlo events are passed through the Geant 4 [123] simulation of the ATLAS Run 2 detector geometry. Pileup has also been simulated in each sample using minimum bias events.

3.1.2 Object reconstruction and selections

Jets are reconstructed using the anti- k_T algorithm with $R = 0.4$. In each event, the leading and sub-leading jet are considered for constructing the Lund Jet Planes. The leading jet in the event is required to have $p_T^{\text{leading}} > 675\text{GeV}$ and must satisfy the p_T balance relation-

ship $p_T^{leading} < 1.5 p_T^{sub-leading}$. The latter requirement ensures that transverse momentum is conserved between the leading and sub-leading jets, as one would expect for a scattering process with two hadronic final states. As a result of the dijet balance requirement, sub-leading jets will have $p_T > 450 \text{ GeV}$. The reason for choosing such high- p_T jets is to allow for a wider range of jet constituent p_T to be reconstructed. This extends the range of possible z coordinates in the LJP, which allows features in the plane to be resolved in greater detail. Since the Lund Jet Planes will be constructed from tracks associated to jets, the jets are required to be located within $|\eta| < 2.1$ where the acceptance of the ID is greatest.

3.1.3 Reconstruction of the Lund Jet Plane

In this measurement, the Lund Jet Plane is constructed entirely from inner detector tracks. Since tracks are only left behind by charged particles, this neglects any contributions to the LJP that are entirely due to neutral particles. The choice of tracks for constructing the LJP is justified by the high $\eta - \phi$ granularity of the tracker, which ensures a very close correspondence between the reconstructed tracks and the true trajectories of the particles.

The soft-drop splitting momentum fraction z_g , splitting angle r_g and dimensionless relative jet mass $\rho \equiv \log(m^2/p_T^2)$ have been measured for tracks and calorimeter topoclusters in Ref. [145]. Preliminary studies in Monte Carlo simulations showed that the difference in these quantities between charged particles and all particles is very small. It was also found that quantities reconstructed from tracks followed the truth level distributions more closely than those constructed from calorimeter topoclusters. As the LJP variables z and ΔR are calculated in a similar way to z_g and r_g , one can expect the effect of using clusters to reconstruct the LJP to be on a similar scale.

It was decided to use tracks as inputs to the Lund Jet Plane at reconstruction level, which was then unfolded to the truth level using the true distributions for only charged particles in Monte Carlo simulation. Although it had been considered that the measurement using charged particles could be unfolded to obtain an LJP for all particles including neutrals, the analysis team decided against this as such corrections would introduce a high degree of model dependence. A study in Monte Carlo simulated events, which is documented in the ATLAS internal note associated with [115], showed that the difference

between the LJP double ratio of charged particles to all particles in Sherpa and Pythia 8 is limited to around 10%.

An additional advantages of using tracks as inputs are the availability of systematic uncertainties for individual tracks which can easily be propagated to derived quantities such as the LJP.

The reconstruction quality of the tracks must pass the "loose" criterion. The tracks are also required to have $p_T > 500\text{MeV}$ and must be associated with the primary vertex, i.e. the reconstructed vertex with the highest sum of p_T^2 in the event.

At particle level, the charged truth particles are required to be stable. This requirement is defined by the condition $ct > 10\text{mm}$ and excluding neutrinos, which can't be detected, and muons, which are not produced in the hadronic processes being studied. The truth particles have been matched to the tracks at the derivation stage. In order to perform this matching, a `MCTruthProbability` score is calculated for each track-truth pair by comparing which layers of the ID have had energy deposited in them by the track/truth particle. The truth matching probability can be used to identify inner-detector tracks as unmatched or originating from pileup as has been done in ref. [146]¹.

For each jet, all tracks within $\eta - \phi$ distance $\Delta R = 0.4$ are collected and re-clustered using the Cambridge-Aachen algorithm. A very large distance parameter, $R = 10.0$, is chosen for the Cambridge-Aachen algorithm at this stage to make sure that all tracks within the $\Delta R = 0.4$ cone are included in the CA track jet. Since the Cambridge-Aachen algorithm only clusters jet constituents based on angular distance, changing the distance parameter of the algorithm only changes the angular separations at which the clustering sequence terminates and does not affect the order of the sequence. The declustering is performed by iterating through the Cambridge-Aachen clustering sequence, starting with the complete jet and finishing with the input tracks. The Lund Jet Plane is constructed in each step of the declustering, as outlined in section 1.7.

¹Further details on the relevant studies can be found in the accompanying ATLAS internal documentation.

3.1.4 Iterative Bayesian Unfolding

The measurement is unfolded using the iterative Bayesian unfolding procure [147]. The purpose of unfolding a measured probability distribution, e.g. a differential cross-section measurement, is to remove detector effects and recover the "true" distribution at the level of the outgoing particles. The unfolding problem can be simply formulated as a matrix equation

$$\mathbf{A}\mathbf{x} = \mathbf{y} \quad (3.1)$$

where $\mathbf{y}$ is a vector containing the bins of the detector-level measured histogram and $\mathbf{x}$ represents the truth level histogram which we want to obtain. The matrix $\mathbf{A}$, which transforms the truth level histograms into detector-level ones and thereby encodes the effects of the detector on the measurement, is called the *response matrix*.

To perform the unfolding, one first constructs the response matrix using Monte Carlo events in which the truth level quantities are known and detector effects have been simulated. The matrix $\mathbf{A}$ must then be inverted to obtain the truth level histograms. Performing the inversion is not as straightforward as it might seem since, in practice, the response matrix that one obtains from MC simulations is often non-invertible. The solution is instead to reformulate the unfolding transformation in terms of conditional probabilities:

$$x_i = \sum_j P(T_i|M_j)y_j \equiv \theta_{ij}y_j \quad (3.2)$$

where x_i is the truth level probability distribution in bin i , y_j is the measured probability distribution in bin j and $\theta_{ij} = P(T_i|M_j)$ is the conditional probability of an event falling into bin i of the truth level distribution, given that bin j has been measured at the detector-level. By applying Bayes theorem, one can relate θ back to the response matrix $\mathbf{A}$:

$$\theta_{ij} = P(T_i|M_j) = \frac{P(M_j|T_i)P(T_i)}{\sum_i P(M_j|T_i)P(T_i)} = \frac{a_{ji}P(T_i)}{\sum_{i=1}^{n_{bins}} a_{ji}P(T_i)}. \quad (3.3)$$

Here, a_{ji} are the elements of $\mathbf{A}$ and $P(T_i)$ is the probability associated to bin i of the truth level distribution. Thus, the unfolding problem is reduced to estimating the truth level distribution $P(T)$, which is also known as the prior. Like the response matrix, the

distribution $P(T)$ can easily be obtained from Monte Carlo events. However, the choice of prior introduces a bias into the unfolded measurement. To counteract this, one can obtain a new prior by applying the unfolding matrix θ to $P(T)$. This is done multiple times, which is reflected in the name *iterative* Bayesian unfolding. The choice of the number of iterations will depend on the uncertainties associated with the measurement, as each successive application of θ introduces more statistical uncertainty.

The final step in the unfolding procedure is to apply *acceptance* and *inefficiency* factors. The acceptance f is a bin-by-bin correction factor that accounts for reconstructed events that have no associated truth level quantities. The inefficiency factor c corrects for truth level events that have not been reconstructed. The correction factors are calculated as follows

$$f_i = \frac{N_i^{reco \wedge pl}}{N_i^{reco}}, \quad c_i = \frac{N_i^{pl \wedge reco}}{N_i^{pl}}. \quad (3.4)$$

The acceptance and inefficiency correction factors must then be applied before and after unfolding respectively. Given a distribution of measured events $\mathbf{d}$ and expected background events $\mathbf{b}$, the final unfolded signal distribution is calculated as follows

$$y_i = (d_i - b_i) \cdot f_i \quad (3.5)$$

$$x_i = (\theta^n y)_i / c_i \quad (3.6)$$

where θ^n is the unfolding matrix, as defined in 3.3, that is applied repeatedly to the acceptance-corrected measured signal distribution y_i .

The unfolding is performed using the `RooUnfold` [148] software package. This package contains utilities for performing iterative Bayesian unfolding on 2-dimensional histograms, such as the LJP, and implements corrections for detector inefficiency and acceptance.

The 2-dimensional LJP histograms can be simultaneously unfolded by first flattening them into a 1-dimensional histogram. The indices of the histogram bins are

$$i_{resp} = i_{\ln R/\Delta R} + i_{\ln 1/z} \times n_{\ln R/\Delta R} \quad (3.7)$$

where $n_{\ln R/\Delta R}$ is the number of bins along the $\ln R/\Delta R$ of the LJP. It is then possible to define a response matrix for the flattened LJP as defined in Equation 3.3.

To visualise the response, it is sometimes more informative to show response matrices corresponding to slices of the LJP, rather than the whole flattened 2-dimensional plane. From 3.7 it follows that the bin coordinates of the response matrix along the detector level and particle level axes i, j can be converted back to LJP coordinates using the following formulas

$$i_{\ln(R/\Delta R)} = i \mod n_{\ln(R/\Delta R)}, \quad i_{\ln(1/z)} = \lfloor \frac{i}{n_{\ln(R/\Delta R)}} \rfloor, \quad (3.8)$$

$$j_{\ln(R/\Delta R)} = j \mod n_{\ln(R/\Delta R)}, \quad j_{\ln(1/z)} = \lfloor \frac{j}{n_{\ln(R/\Delta R)}} \rfloor. \quad (3.9)$$

The response matrix for a slice through the LJP is then defined by the coordinates i_x, j_x and $i_y, j_y = i_{slice}$, where x is the LJP coordinate along the direction of the slice and y is the LJP coordinate orthogonal to the slice direction where the bin index is fixed to i_{slice} .

3.1.4.1 Constructing the response matrix

The response matrix is constructed using events that pass both the reconstruction and particle level event selections. Additionally, since the Lund Jet Plane is constructed from the kinematics of individual emissions in the CA clustering hierarchy, it is necessary to perform a matching between emissions at the reconstruction and the particle level. The simplest way to perform this matching would be to match truth and detector level emission by the order in which they appear in the CA-declustering. However, because the shape of the reconstructed Lund Jet Plane is changed by the presence of unmatched or unreconstructed tracks, there may be different number of emissions at the particle and re-

construction levels. Therefore, an additional matching condition based on the kinematics of the emissions is applied.

Figure 3.1 shows a comparison of 5 different matching strategies. The simplest of these is matching emissions by order in the CA-declustering. Two more sophisticated methods use the distance between the emissions in either detector coordinates $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ or Lund Jet Plane coordinates $\Delta R_{LJP} = \sqrt{\Delta[\ln(R/\Delta R)]^2 + \Delta[\ln(1/z)]^2}$. Another method uses the truth-particle links to select the reconstruction level emission with the highest fraction of tracks linked to the truth particles of a given truth level emission. The final and most accurate scheme computes the fraction of linked tracks weighted by the track p_T .

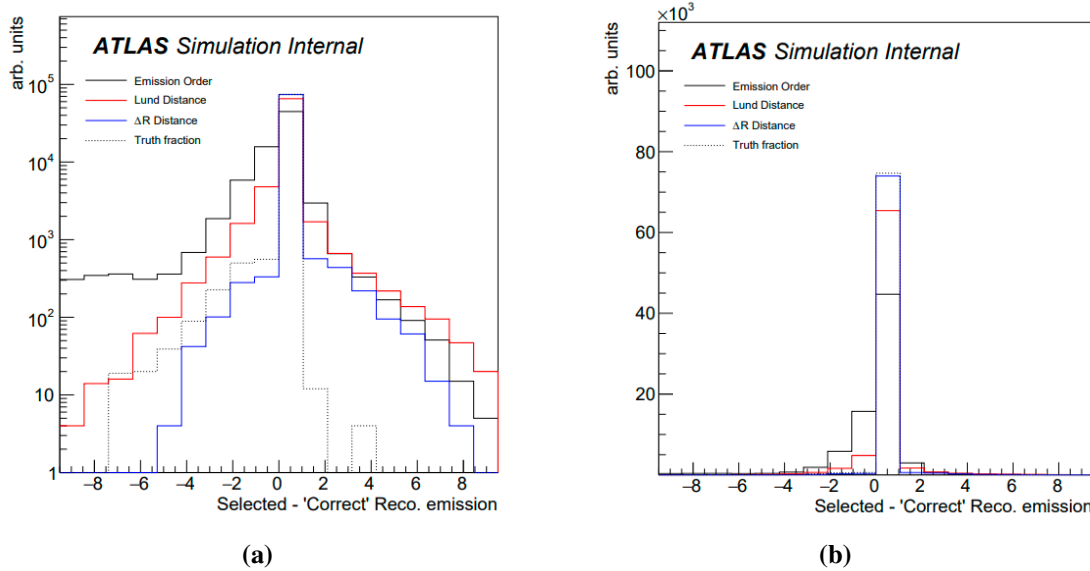


Figure 3.1: Comparison of matching procedures between reconstruction and particle level emissions. The same plot is shown with linear and (a) and log-scale (b) y-axis. The x-axis shows the difference in emission order between each of the different matching schemes and the "correct" matching obtained from the p_T weighted fraction of linked truth particles. This figure is taken from the ATLAS internal note accompanying ref. [115].

3.1.4.2 Shape of the response: migration matrices

Migration matrices can be obtained from the response matrix by normalizing the rows or columns of the matrix so that they add up to 1. One then obtains a conditional probability, Like $P(DL|PL)$ which represents the probability of reconstructing an emission in one bin at detector level given that it appears in another bin at particle level (i.e. truth level). Similarly one can also compute the probability $P(PL|DL)$ of finding an emission in one bin at particle level given that it appears in another bin at the detector level. Migration matrices showing the conditional probability $P(DL|PL)$ in slices through the Lund Jet Plane are shown in figures 3.2 and 3.3. The migration matrices are mostly diagonal and symmetric, except at small splitting angles, or large $\ln(R/\Delta R)$, where splittings with small z at the truth level are slightly more likely to be reconstructed with a larger z at the detector level. However, only around 10% of emissions are mismatched between their particle and detector level $\ln(1/z)$ bins in this region. It is possible that this effect might be the result of the limited spacial resolution of the detector tracks, as there is a greater uncertainty due to track reconstruction in this region – up to 1.5% compared to $< 0.5\%$ in the rest of the plane. A larger collection of migration matrices, including for $P(PL|DL)$, can be found in the ATLAS internal note accompanying ref. [115].

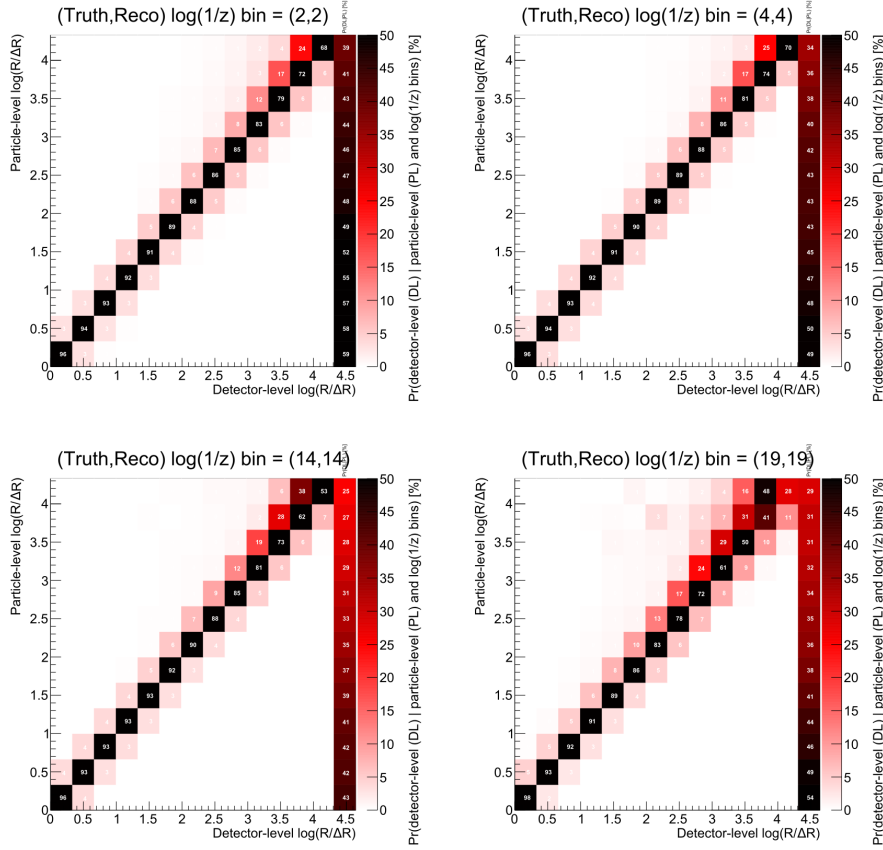


Figure 3.2: Bin-to-bin migrations for slices through the Lund Jet Plane in dijets along the $\ln(R/\Delta R)$ axis. The migrations between bins are quantified here by the conditional probability $P(DL|PL)$, meaning the probability of reconstructing an emission at the detector level given that it appears at the particle (or truth) level. The extra row of bins on the right denotes the reconstruction efficiency in each bin of the slice. This figure is taken from the ATLAS internal note accompanying ref. [115].

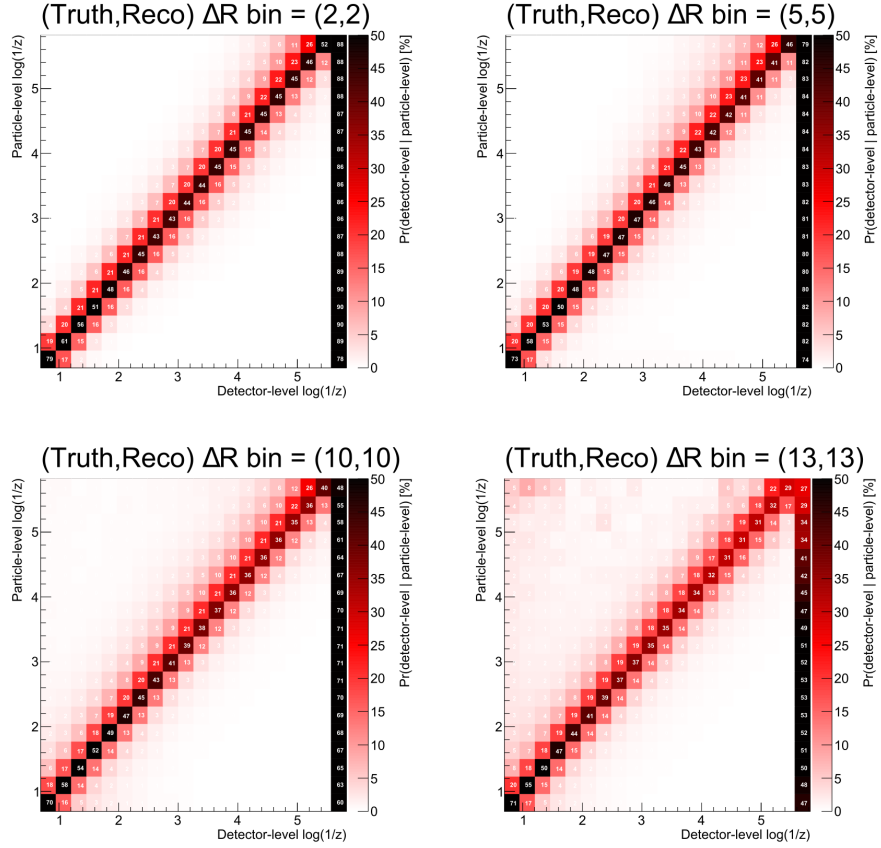


Figure 3.3: Bin-to-bin migrations for slices through the Lund Jet Plane in dijets along the $\ln(1/z)$ axis. The migrations between bins are quantified here by the conditional probability $P(DL|PL)$, meaning the probability of reconstructing an emission at the detector level given that it appears at the particle (or truth) level. The extra row of bins on the right denotes the reconstruction efficiency in each bin of the slice. This figure is taken from the ATLAS internal note accompanying ref. [115].

3.2 Effect of unmatched tracks and their composition

The distribution of tracks without a truth match across the Lund Jet Plane was studied in Monte Carlo events. The absence of track-particle links, as described in Section 3.1.3, is used to identify these tracks. The fraction of unmatched tracks and matching efficiency, which are used to quantify the impact of unmatched and unreconstructed tracks, are defined as follows

$$f_{unmatched} = \frac{\#unlinked\ tracks}{\#reconstructed\ tracks} \quad (3.10)$$

$$\epsilon_{matched} = \frac{\#linked\ tracks}{\#charged\ truth\ particles}. \quad (3.11)$$

Figure 3.4a shows the fraction of unmatched tracks across the LJP. A wide range and fine binning is chosen for the $\ln(1/z)$ and $\ln(R/\Delta R)$ values, to show the large variation in the rate of unmatched tracks that exists across the plane. In the lower left corner of the plane, which represents the fiducial region of the dijets measurement, the rate of unmatched tracks is limited to $< 10\%$. Most unmatched tracks are found along the diagonal of the plane, where soft and wide-angle radiation is dominant.

Some emissions in the Lund Jet Plane are entirely due to secondary decays in the detector. These are identified in the MC truth record using the PDGID of the linked truth particles. Secondary particles are produced in the Geant 4 detector simulation and have a PDGID of 200000 or greater. An emission is labelled as secondary if this condition is met for all of its linked truth particles. Using the MC truth record, it is also possible to further subdivide secondary emissions into hadronic and electromagnetic decays. Electromagnetic secondary decays encompass all processes, such as photon conversion, in which one or more e^+ or e^- are produced. Processes in which a hadron is produced are labelled as hadronic. For a small number of secondary decays it is not possible to categorise them as either electronic or hadronic from their MC truth record, these are labelled as "unknown" secondary decays. The distribution of secondary decays across the LJP is shown in figure 3.4b. The density of secondary decays is highest in the region above the diagonal of the LJP where perturbative QCD emissions are highly suppressed.

In the regions of the LJP that are dominated by interesting physics, on the other hand, there is only a very small secondary decays.

Tracks which are due to pileup are identified using the z coordinate of the linked truth particle at its point of production. This value is compared to the location of the primary vertex and if $|z_{prod}^{particle} - z^{PV}| < 0.1$ for all truth particles linked to the emission it is labelled as pileup. The distribution of pileup tracks across the LJP is shown in figure 3.4c. The fraction of pileup tracks across the whole of LJP is very low, typically less than 5%, due to the primary vertex association requirement on the tracks. There is a somewhat higher contribution from pileup tracks at small R and large z , suggesting that hard, collinear pileup tracks are more easily misidentified by the PV-association algorithm.

The truth matching efficiency for LJP emissions is shown in figure 3.4d. The matching efficiency is highest along the diagonal of the plane where there are few tracks without a truth match or tracks due to secondary decays. An unusual feature of figure 3.4d is the drop in efficiency in the lower left hand corner of the plane, around $\ln(R/\Delta R) < 1.5$, $\ln(1/z) < 2.0$. Emissions in this region are at wide angles and large p_T and should therefore be well-reconstructed. However, emissions in this region are also more likely to be proto-jets composed of many particles rather than a single particle. One possible hypothesis for the observed drop in the matching efficiency could be composite nature of these emissions, since the matching efficiency for an emission in the LJP is calculated as the average for its constituent tracks. This seems, however, not quite enough to explain the observed drops to efficiencies of < 0.5 . The cause of this effect is therefore not well understood.

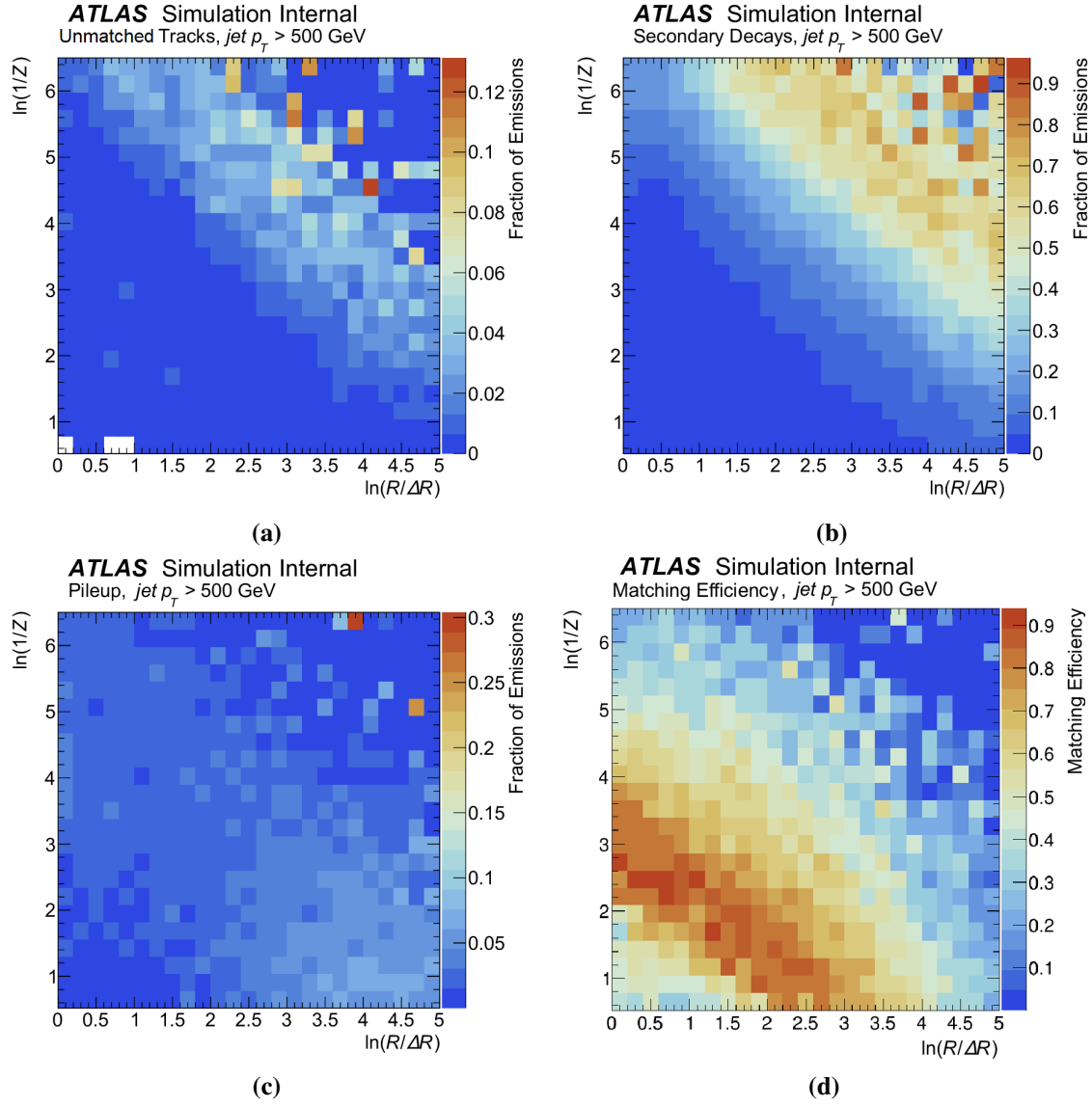


Figure 3.4: Distribution across the Lund Jet plane of unmatched (a), secondary (b) and pileup (c) rates, as well as the matching efficiency (d).

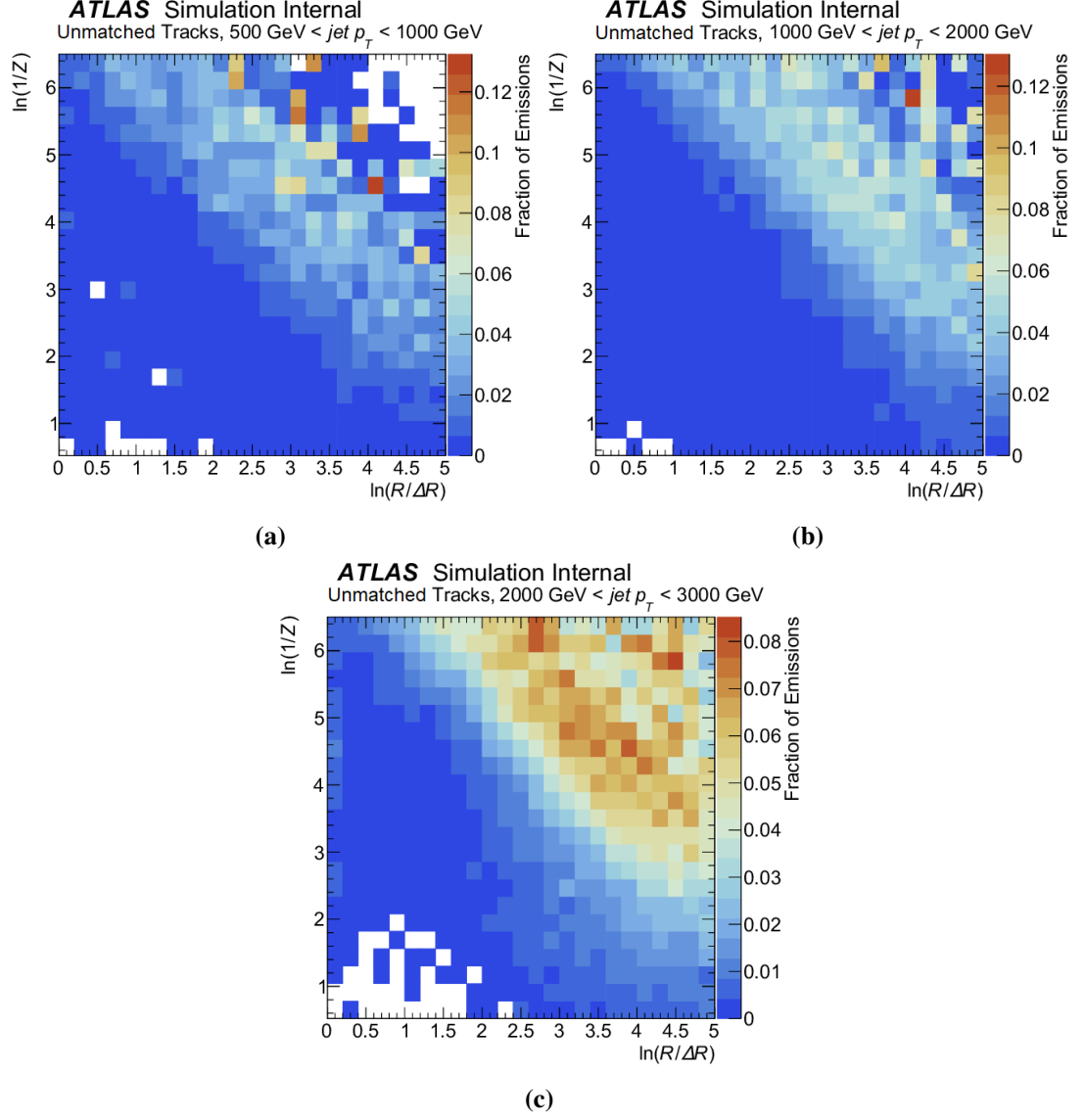


Figure 3.5: Distribution across the Lund Jet plane of the rate of unmatched tracks for three different bins of emission p_T .

3.3 Effect of dead TileCal modules in 2018 data

During some of the data taking periods of Run 2 in 2017 and 2018, a small number of modules of the TileCal had to be turned off. Since this measurement uses calorimeter topocluster jets for jet-level kinematic selections, there is a risk that the missing modules may introduce a bias into the measurement of the Lund Jet Plane.

In 2017, two Tile modules, LBC63 and EBA03, were turned off for the entire year. The $\eta - \phi$ regions corresponding to the two inactive modules can clearly be seen in figure 3.6c. The effect of the dead modules was also replicated in the Monte Carlo simulated events for this data period, mc16d.

During part of the 2018 run, from 27 May until 20 June, a coolant leak occurred in the Tile Calorimeter and another two modules, LBA29 and LBA30, had to be turned off. A third module, LBA32, was turned off from 20 June until the end of Run 2. The corresponding Monte Carlo simulated events, mc16e, also had LBA32 turned off, but not LBA29 and LBA30. The $\eta - \phi$ region covered by the inactive modules is visible in 3.7c, which shows the jet $\eta - \phi$ distribution for the whole of 2018.

The impact on the Lund Jet Plane is studied by comparing planes constructed from jets within $\Delta R < 0.4$ of the dead tile module to jets from a reference region of the same size and with the same η as the region of interest. These Lund Jet Planes are shown in Figures 3.6 and 3.7. The ratio of the reference and probe regions for the 2017 and 2018 periods are shown in figure 3.8.

The difference in the LJP ratio between the probe and reference regions is limited to around 0.5%. The ratio of normalised LJPs between the data periods with faulty Tile modules and the 2018 B,C periods, during which the calorimeter was fully functional, is shown in figure 3.9. The change in the data/data ratio is $< 1\%$ in most of the LJP, except for very small ΔR and z where it is around 2%. Since the 2017 and 2018 data periods together make up 73% of the total Run 2 dataset [49], we can expect the impact on the total measurement to be even somewhat smaller. The systematic uncertainty introduced by the inactive TileCal modules can therefore be safely neglected.

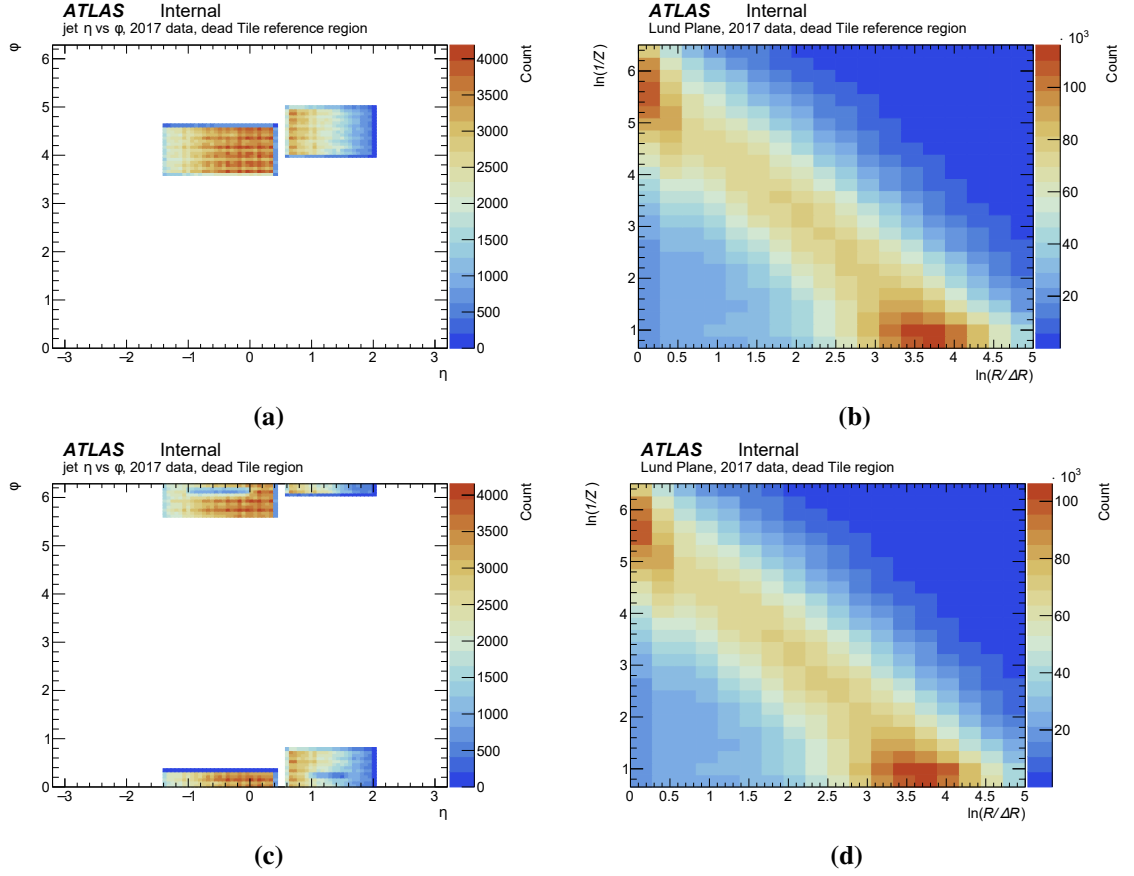


Figure 3.6: Reference and probe regions for investigating the impact of missing Tile modules in the 2017 data on the reconstructed Lund Jet Plane. Shown in (c) are the "probe" regions where $\Delta\eta < 0.4$ and $\Delta\phi < 0.4$ to the missing tile modules, while (a) shows the "reference" clean region with equivalent η and $\eta - \phi$ area. Lund Jet Planes reconstructed from jets in the reference and probe are shown in (b) and (d).

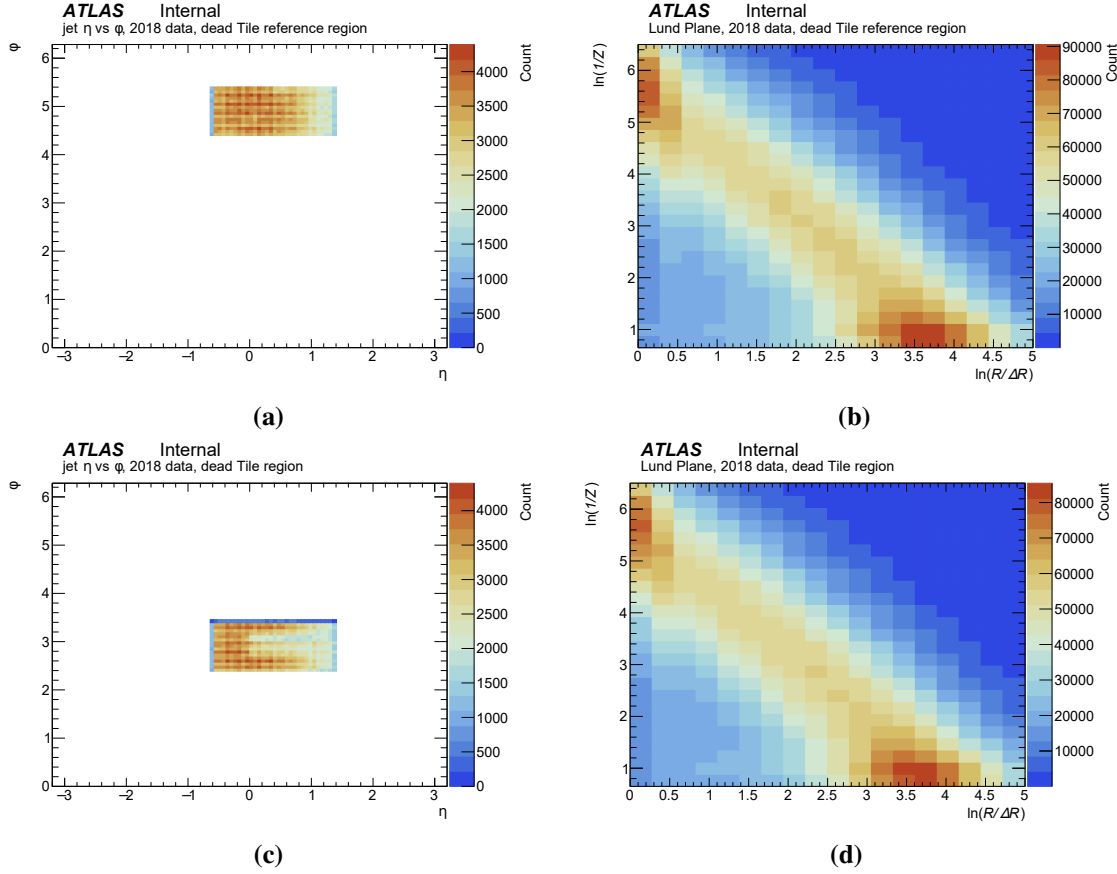


Figure 3.7: Reference and probe regions for investigating the impact of missing Tile modules in the 2018 data on the reconstructed Lund Jet Plane. Shown in (c) are the "probe" regions where $\Delta\eta < 0.4$ and $\Delta\phi < 0.4$ to the missing tile modules, while (a) shows the "reference" clean region with equivalent η and $\eta - \phi$ area. Lund Jet Planes reconstructed from jets in the reference and probe are shown in (b) and (d).

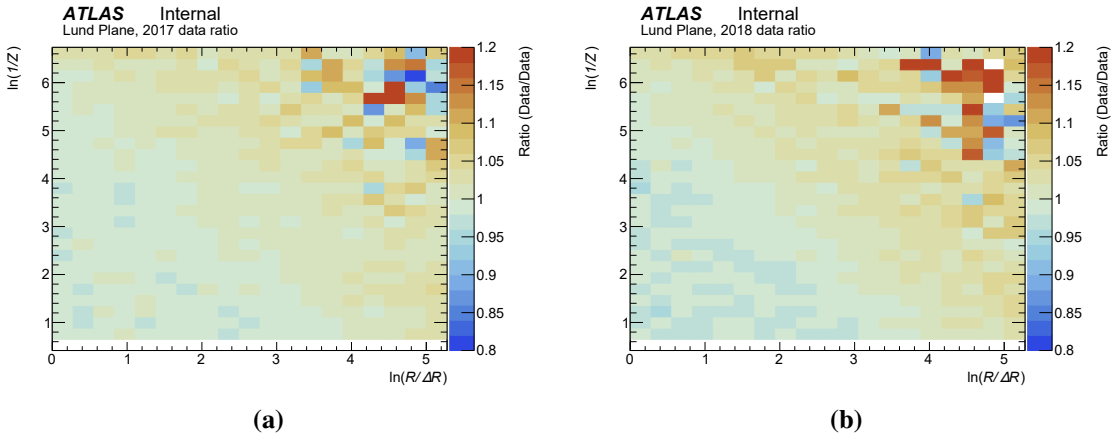


Figure 3.8: Lund Jet Plane ratios between reference and probe regions in 2017 (a) and 2018 (b) data. A definition of the reference and probe regions, as well as the corresponding Lund Jet Planes are shown in Figures 3.6 and 3.7. The large differences observed in the upper right-hand corner of the plane are due to the small amount of statistics in this region.

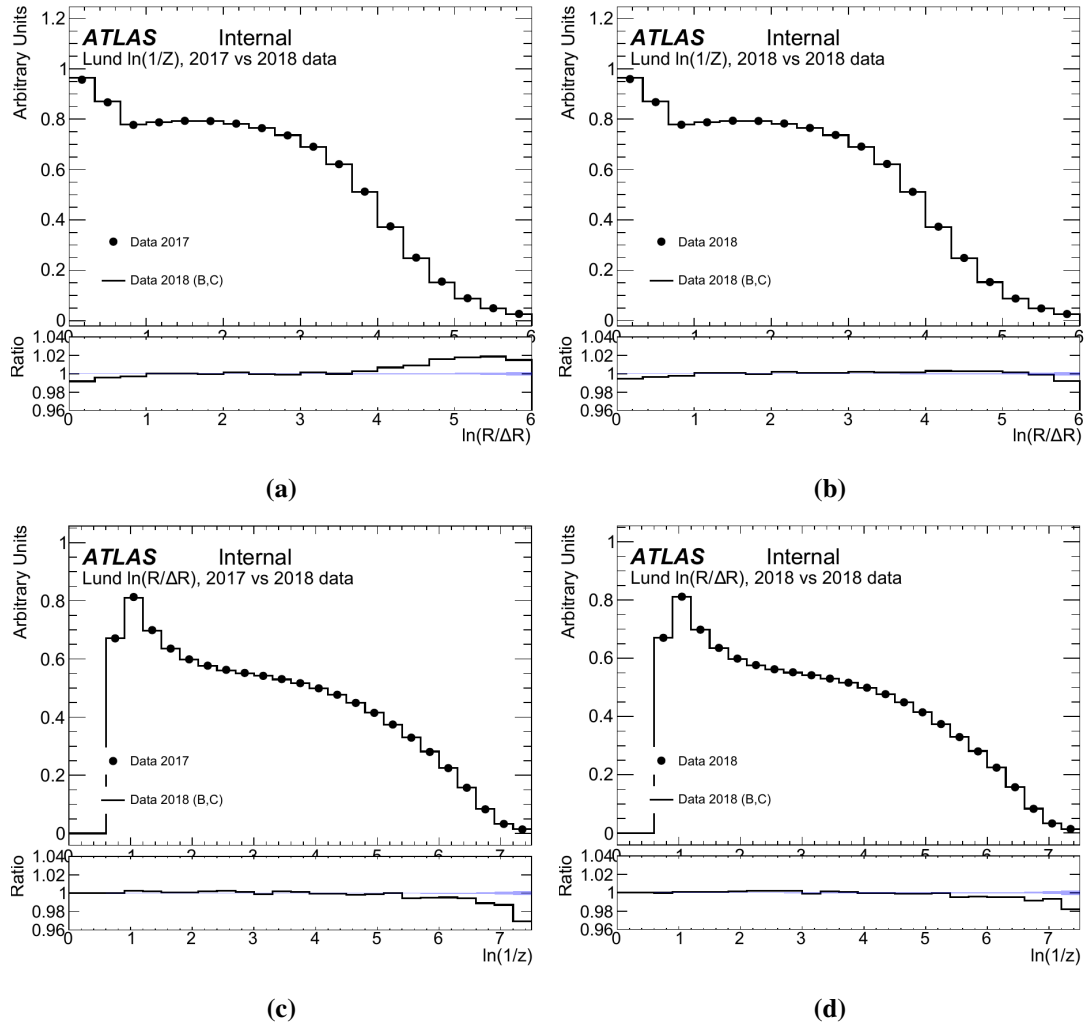


Figure 3.9: The normalized LJP distribution, projected along $\ln(R/\Delta R)$ (a,b) and $\ln(1/z)$ (c,d), with ratios to the 2018 B,C periods (before the cooling leak, when the full TileCal was operational). The 2017 data period is shown in (a,c) and the 2018 period is shown in (b,d).

3.4 Unfolded results

The Unfolded Lund Jet Plane for dijets in ATLAS Run 2 data is shown in figure 3.10. More than 29.5 million jets pass the fiducial selections of the measurement. The 2D histograms of Lund Jet Plane emissions have been normalised by the total number of jets N_{jets} and the bin area to give the Lund Jet Plane density as a double differential cross-section

$$\rho_{LJP}(\Delta R, z) = \frac{1}{N_{jets}} \times \frac{d^2 N_{emissions}}{d \ln(1/z) d \ln(\Delta R/R)}. \quad (3.12)$$

A comparison between slices of the LJP in data and different MC generators is shown in 3.11.

The difference in the Lund Jet Plane density resulting from varying Parton Shower algorithms (Herwig 7 Ang. ord. vs dipole) is most pronounced in the bottom left corner of the Plane. Varying hadronisation models (Sherpa 2.2.5 AHADIC vs. String) gives a larger discrepancy close to the diagonal of the LJP. The slice through figure 3.11c shows how the wide-angle region of the plane is more sensitive to the choice of parton shower models, while at smaller angles the difference between hadronisation models is more pronounced. The effect of using LO vs. NLO matrix elements (Pythia vs. Powheg+Pythia) is most pronounced very close to the bottom left corner of the plane, where hard and wide angle perturbative effects are dominant. The maximum difference in the LJP density resulting from the varying samples ranges from only about 10% for the matrix elements to around 50% for the parton shower and hadronisation models.

The best overall agreement with the data is obtained for the Herwig 7 samples with angle-ordered parton showering. This can be seen in, for example, in the LJP slices at low and high $\ln(R/\Delta R)$ shown in figure 3.11a and figure 3.11b. Here, the MC/Data ratio is closest to 1 for Herwig 7 (angular ordered) than for the other generators. The difference between the Data and Herwig in these regions is usually within 1 or 2 times the total statistical and systematic uncertainty. However, none of the MC generators are able to model the Lund Jet Plane across all its regions and there are some regions in which other generators than Herwig 7 model the data more closely. At small z and small ΔR , the data is more closely modelled by Pythia 8 or Powheg + Pythia 8, while in the region $0.5 \lesssim \ln(R/\Delta R) \lesssim 1.5$, $2.0 \lesssim \ln(1/z) \lesssim 3.0$ it is most closely matched by Sherpa. The

most significant mismodelling for all generators is observed in the top left corner of the LJP, which is populated by soft emissions at wide angles. The origin of these emissions is mostly due to the underlying event. The difference between the data and the Monte Carlo in this region is between approximately 1.5 – 2.5 times the total uncertainty for Herwig 7 (angular ordered) and Sherpa (string or AHADIC hadronisation) to 4.0 – 4.5 times the total uncertainty for Pythia 8, Powheg + Pythia 8 and Herwig 7 (dipole). Differences between the angular ordered and dipole classes of parton shower models can be observed in the LJP by comparing the results for the two Herwig samples. The difference between these two models is most pronounced along diagonals across the plane where z and ΔR are both large (see figures 3.11a and 3.11c), suggesting that hard and wide-angle emissions are generated at different rates between the two parton shower models.

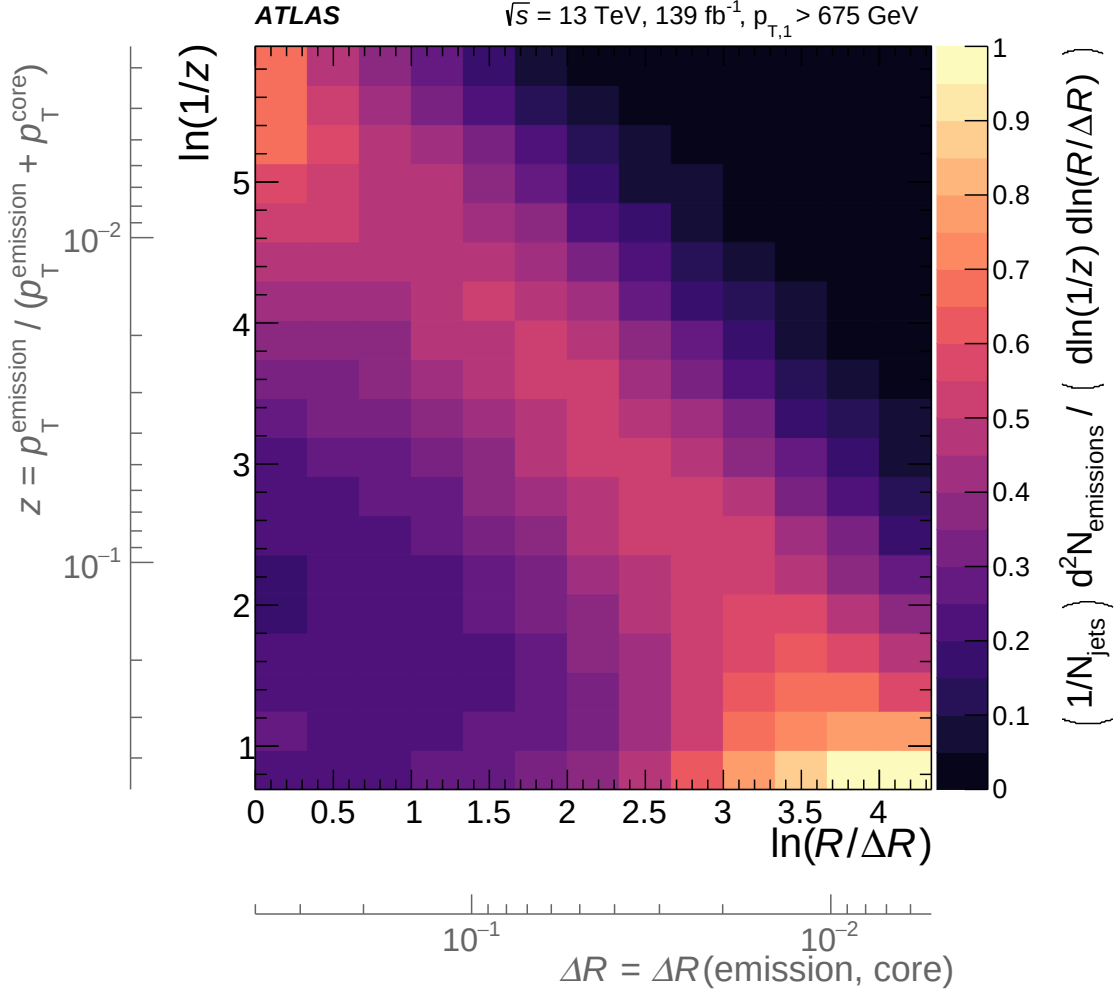


Figure 3.10: The Lund Jet Plane for charged particles in dijets as measured in Run 2 Data. The 2-d distribution has been unfolded to the particle level. From Ref. [115].

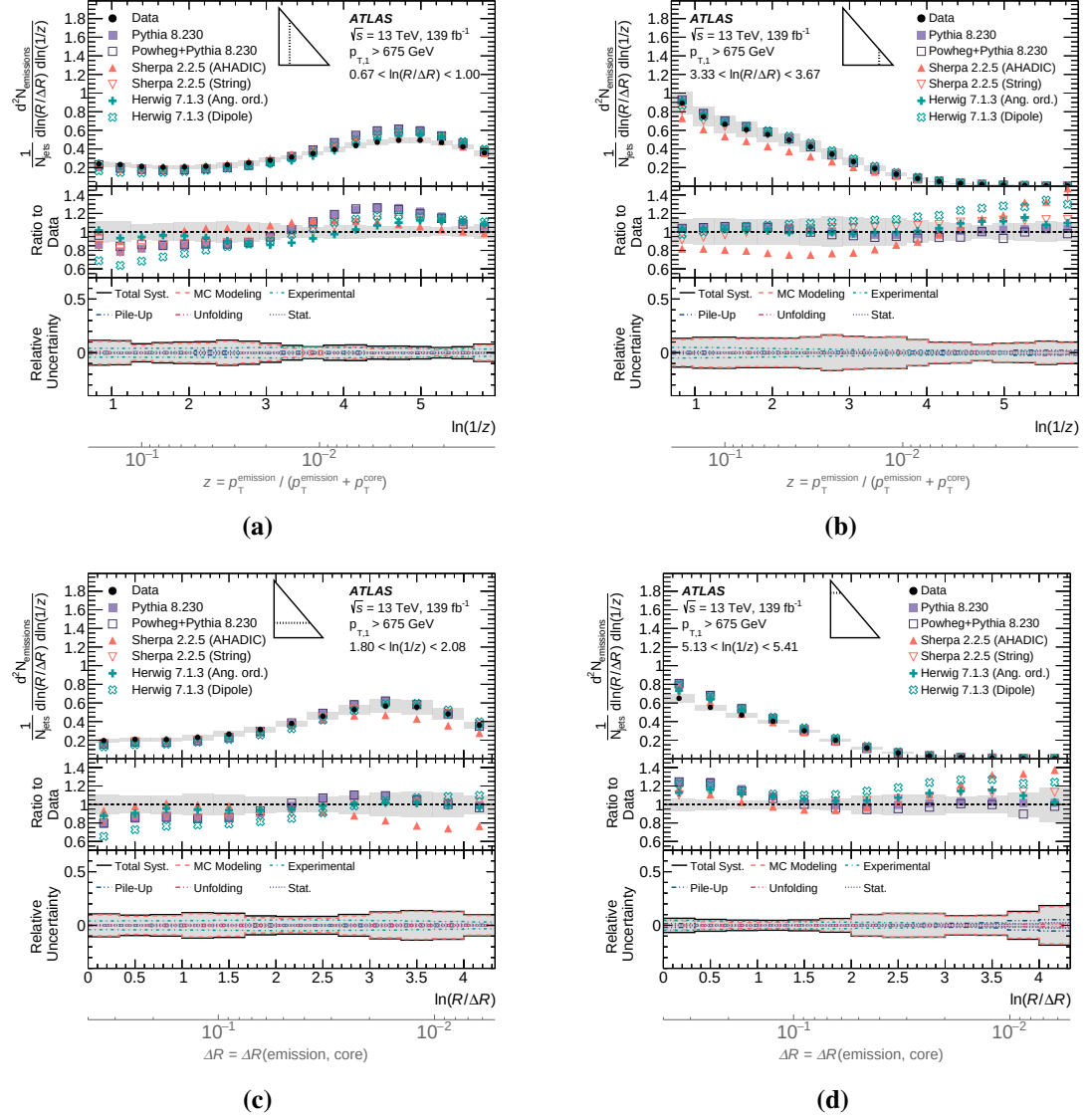


Figure 3.11: Slices through the Lund Jet Plane in dijets data and Monte Carlo simulation. The cut values for the fixed variable in each of the slices are: (a) $0.67 < \ln(R/\Delta R) < 1.00$, (b) $3.33 < \ln(R/\Delta R) < 3.67$, (c) $1.80 < \ln(1/z) < 2.08$, and (d) $5.13 < \ln(1/z) < 5.41$. From Ref. [115].

3.5 Conclusion and subsequent developments

The measurement of Lund Jet Plane in dijets, published in ref. [115], is a pioneering high-precision measurement of the substructure of jets. It provides an opportunity for better modelling of the radiation patterns inside jets through improved MC generator tuning and for testing theoretical predictions.

Kar et al. [149] have used the measurement to produce a new tune for Pythia 8. The modelling of the distribution of the LJP, as well as the Soft Drop mass [150], is improved by the new tune by about 5 – 10% compared to the A14 [141]. In the LJP, modelling is particularly improved in the soft collinear regions where there are significant UE and MPI contributions.

Shortly following the publication of the measurement, a theoretical prediction of the primary Lund Jet Plane density in dijets was calculated by Lifson et al. [151]. This prediction is an analytic re-summation of single logarithms in perturbative QCD with NLO corrections for soft effects and MC-based corrections for contributions from UE and MPI. A sample of comparison plots between the theoretical predictions and the measurement is shown in figure 3.12.

The theoretical predictions are in agreement with the measurement except at large ΔR (the region shaded red in figure 3.12) and at very small z . Lifson et al. note that this a region with large contributions from non-perturbative effects but that there are also some perturbative corrections in this region whose size is difficult to estimate and which would not be required if the measurement had been performed with CA jets rather than CA re-clustered anti- k_T jets. The re-clustering introduces logarithmic terms proportional to $\ln\left(\frac{R}{R-\Delta R}\right)$ called *boundary logarithms* which are not included in the analytic re-summation. The boundary logarithms are caused by soft collinear radiation at the edge of the jet which is not captured by the anti- k_T algorithm. If the jets had instead been clustered using CA from the beginning, these additional terms would not be present. A future measurement using a pure CA jet definition could help shed light on the dominant effects in the wide-angle region by facilitating analytic predictions.

The LJP has also been measured by the ALICE experiment [152] in minimum bias events at $\sqrt{s} = 13\text{ TeV}$ using inclusive charged particle jets in the range $20 \leq p_T \leq$

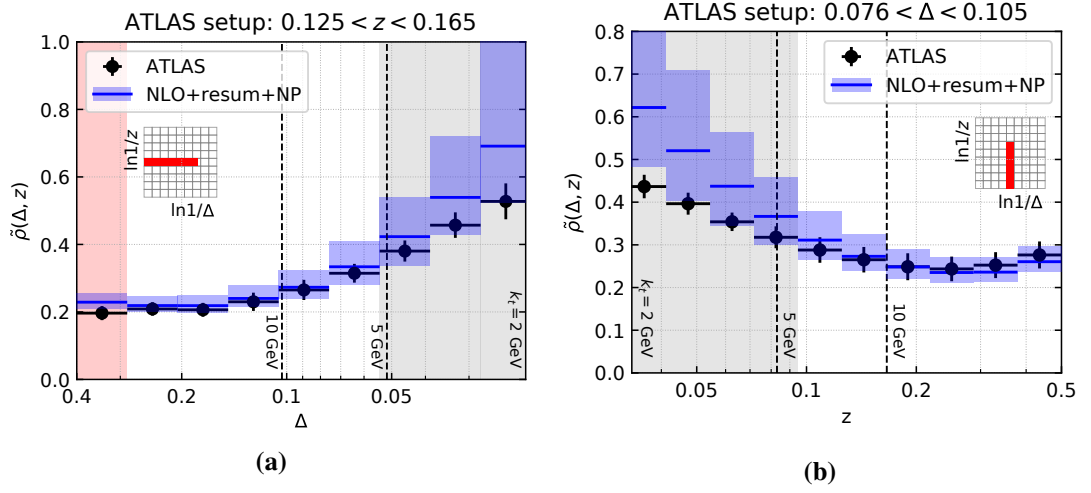


Figure 3.12: Theoretical predictions of the LJP density compared to the ATLAS Run 2 measurement in slices along $R/\Delta R$, here denoted Δ , (a) and z (b). In the region shaded red in (a) there is significant disagreement between theory and measurement. From ref. [151].

120 GeV. This measurement probes a much lower range of p_T than the ATLAS measurement, a region where underlying event and non-perturbative hadronisation effects are dominant. ALICE have also probed even lower- p_T regions, using jets in the range 5 – 36 GeV to make the first direct observation of the *dead cone effect* [153, 33]. The dead cone effect refers to a suppression of the amount of radiation within jets initiated by massive quarks at angles smaller than the critical value m_q/E_q , where m_q and E_q are the mass and energy of the quark (see section 1.1.7, [32]). Charm quark initiated jets are selected by requiring the presence of $D^0 \rightarrow K\pi$ candidates. By computing the ratio of LJPs between c -jets and (predominantly light) inclusive jets, the dead cone effect is made visible as a suppression of emissions at small ΔR .

Chapter 4

Measurement of the Lund Jet Plane in $t\bar{t}$ events

4.1 Introduction

During its second run, the LHC set new records for both centre-of-mass energy and luminosity in proton-proton collisions. Unprecedented quantities of heavy standard model particles, such as top quarks, were produced, enabling measurements of their properties with record precision. By studying the substructure of top-quark initiated jets in this dataset, it is possible to test the validity of Standard Model predictions in phase-spaces that are not accessible in simple differential cross-section measurements using jet kinematics. One technique for exploring such phase-spaces is the Lund Jet Plane.

To investigate the shape of the Lund Jet Plane for jets that are initiated by boosted objects, semi-leptonic $t\bar{t}$ decays are studied. The basic topology of a semi-leptonic $t\bar{t}$ decay is shown in figure 4.1. The events are identified by the signature of the leptonically decaying top quark. Top quarks (almost) always decay to a b -quark and a promptly decaying W boson (the probability, as given by the CKM matrix element $|V_{tb}|^2$, is assumed to be one and has been measured in experiment at $R = 1.014 \pm 0.03$ [154]). As the W decays to a lepton and a ν , the leptonically decaying top, or anti-top, is identified by the presence of one highly energetic lepton, a b -tagged $R = 0.4$ jet and missing E_T . In a semi-leptonic event, the second top, or anti-top, quark decays hadronically, meaning the promptly decaying W produces a pair of light quarks. The hadronic decay products are

reconstructed using jets. The jets to be studied are *large- R* jets with a distance parameter of $R = 1.0$. At lower transverse momenta, it is possible that only parts of the decay products are reconstructed within the large- R jet.

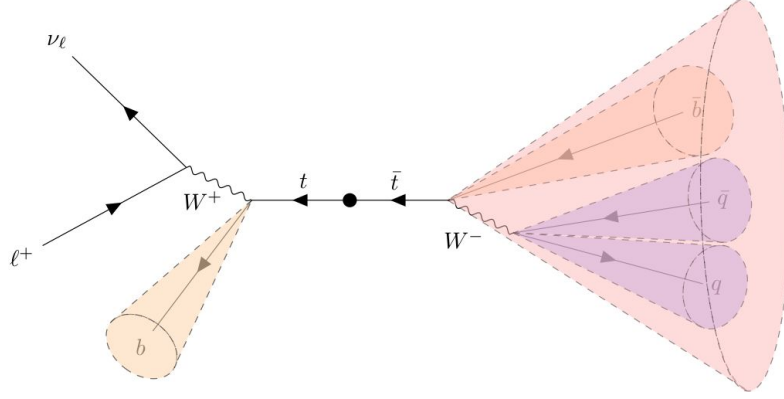


Figure 4.1: A Feynman diagram showing a semi-leptonic $t\bar{t}$ -decay. The leptonically decaying top is reconstructed as an e/μ and an $R = 0.4$ (small R) b -tagged jet, while the decay products of the hadronically decaying top are captured by the $R = 1.0$ (large R) jet.

4.1.1 Construction of the Lund Jet Plane

The construction of the Lund Jet Plane for $t\bar{t}$ jets starts with the calibrated large- R jets. These are trimmed (see Section 1.5.1) anti- k_T jets built from calorimeter topoclusters with a distance parameter $R = 1.0$. The trimming parameters are $R_{trim} = 0.2$ and $f_{trim} < 0.05$. Like for the measurement in dijets using $R = 0.4$ jets (see 3.1.3), the LJP is constructed by re-clustering charged particle tracks within the jet. However, the association of tracks to jets is performed using ghost-association (see Section 1.4.3.4) rather than through a simple geometric matching in ΔR . Trimming is applied to the ghost-associated tracks with the same parameters as for the topoclusters.

At the level of truth Monte Carlo truth particles (particle level), large- R jets are reconstructed from stable particles ($c\tau > 10\text{mm}$) using the anti- k_T algorithm with $R = 1.0$. Trimming is applied to the truth particles with the same parameters as for the topoclusters at the detector level. The charged constituents of the jet are used to construct the particle level Lund Jet Plane. This method was chosen as it was found to provide a closer match between the reconstructed and particle level emissions, in the case of large- R jets, than the geometric matching.

figure 4.2 shows the distance between reconstructed and particle-level emissions ΔR_{PL-RL} for LJPs constructed from ΔR -matched tracks and ghost-associated tracks. A comparison between particle-level LJP definitions, using only charged particles or charged and neutral particles, is also shown. The closest match is found between the ghost-associated track LJP emissions and the charged truth particle constituent LJP emissions. The reason why the ghost tracks are more closely matched to the truth particles than the ΔR -matched tracks, is likely due to the trimming that has been applied to the truth particles. The ghost-associated tracks have also been trimmed, whereas the ΔR matched tracks have not. Re-clustering the untrimmed ΔR matched tracks using Cambridge-Aachen changes the clustering sequence at the detector level, which gives rise to a larger discrepancy in η and ϕ between the detector and particle level emissions.

A comparison between the $\eta - \phi_{var}$ distributions of the emissions for individual jets and the corresponding LJPs is shown in figure 4.3. The angle $\phi_{var} \equiv \arctan((y_{emission} - y_{core})/(\phi_{emission} - \phi_{core}))$ is the azimuth angle of the emission relative to the core redefined such that it is invariant under a Lorentz-boost along the z direction of the detector, and η is evaluated relative to the axis of the jet. The rapidity y of the tracks is estimated from the measured kinematics of the tracks by assuming a mass of $m_{\pi}^{\pm} = 139 \text{ MeV}$. Detector and particle level emissions are matched by distance in $\eta - \phi_{var}$ space when constructing the response matrices required for the iterative Bayesian unfolding procedure. The detector-level LJPs built from the ghost-associated tracks are closely matched to the particle-level constituent LJPs in both $\eta - \phi_{var}$ and LJP space. The matching procedure is described in greater detail in section 4.4.3.

The construction of the LJP from the input tracks and truth particles is performed by re-clustering the particles using the Cambridge-Aachen algorithm and following the procedure outlined in Section 1.7.

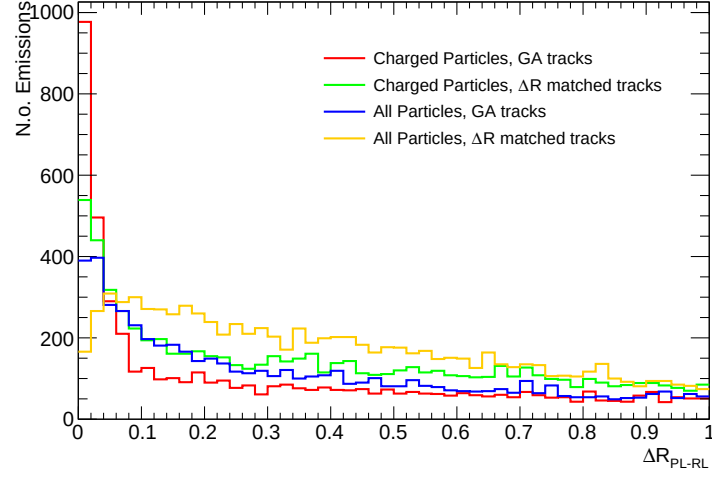


Figure 4.2: Distance between emissions at the particle and detector level for different methods of reconstructing the LJP. The distances are computed from a subsample of $t\bar{t}$ Monte Carlo events.

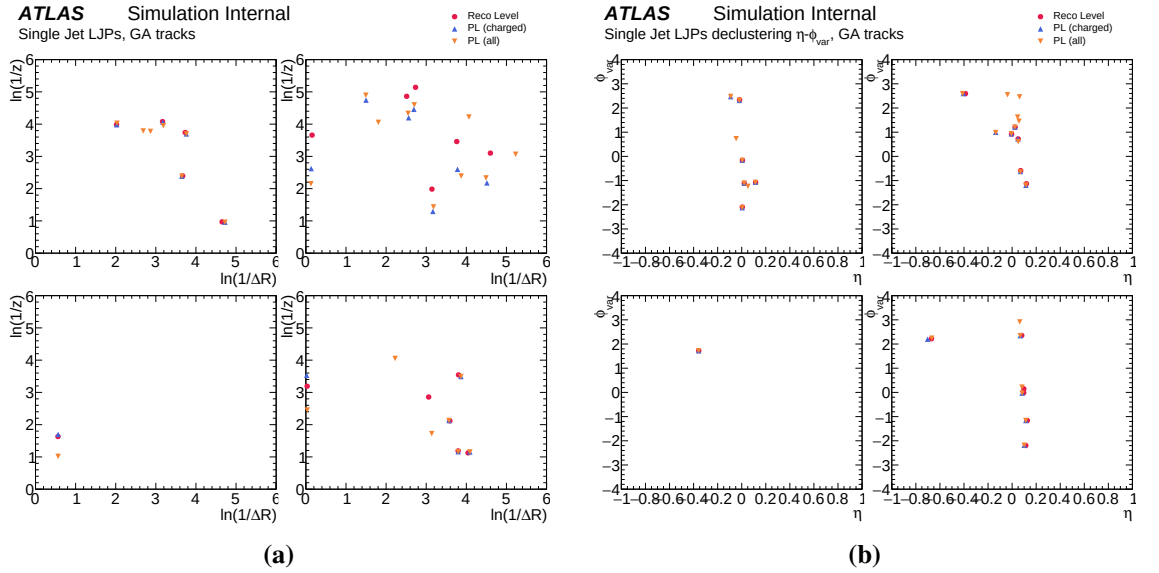


Figure 4.3: Lund Jet Planes for individual jets at the particle and the detector level (a) and distributions in $\eta - \phi_{var}$ for emissions (b).

4.2 Data, objects and selections

4.2.1 Dataset

The measurement is performed using the ATLAS Run 2 dataset collected between 2015 and 2019. The dataset contains proton-proton collisions at $\sqrt{s} = 13$ TeV and has a combined integrated luminosity of 139 fb^{-1} . The uncertainty in the total integrated luminosity is around 1.7% as measured by the LUCID-2 detector [49, 155]. The data is required to pass the ATLAS standard data quality requirements.

4.2.2 Triggers

Events are required to pass at least one single-lepton trigger. The names of the triggers used in each data period are shown in Table 4.1. A full description of the electron and muon triggers can be found in refs. [156, 157].

All of the single-lepton triggers used in this analysis are unprescaled. Each of the high-level single-lepton triggers requires a lepton candidate above a certain p_T . The trigger p_T thresholds are 26, 60 and 300 GeV for electrons, and 26 and 50 GeV for muons, except in the 2015 data period where they are 24, 60 and 120 GeV for electrons and 20 and 50 GeV for muons. The lowest p_T single-lepton triggers also include isolation criteria on the momentum sum of tracks near the lepton candidates. The high-level electron triggers use tracks matched to clusters in the EM-calorimeter as inputs, while the muon triggers use combined muons, i.e. muon candidates that are reconstructed from hits in the ID and MS. For the electron candidates, a multivariate likelihood discriminant called LH is used to quantify the reconstruction quality. Electron triggers whose names include the string `nod0` exclude the track impact parameter d_0 in the calculation of LH to increase efficiency.

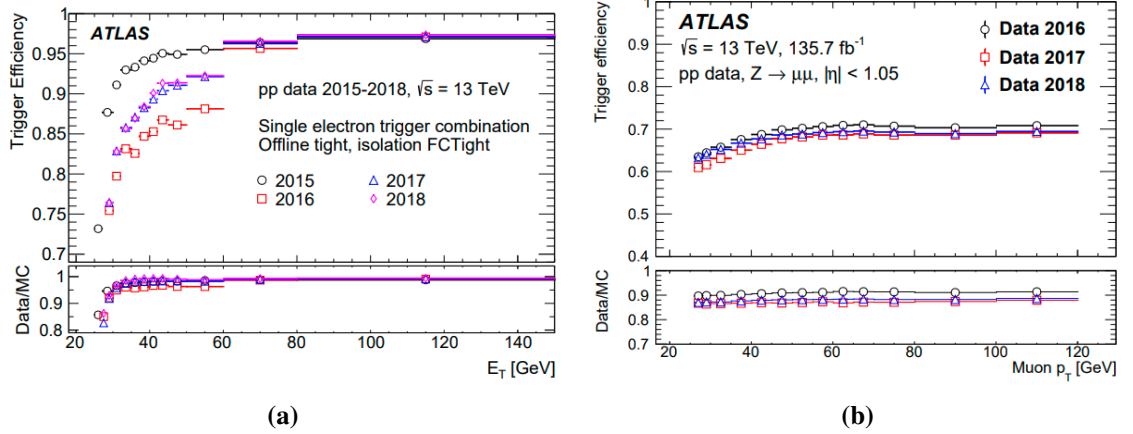
Examples of the typical trigger efficiencies in the Run 2 data for the electron and muon triggers are shown in figure 4.4.

4.2.3 Event selection

A summary of the selections applied to the various objects in the semi-leptonic $t\bar{t}$ events is shown in Table 4.2.

Table 4.1: Lists of single lepton triggers in each data-taking period. A detailed description of each trigger is available in refs. [156, 157].

Period	Triggers
2015	Electrons - HLT_e24_lhmedium_L1EM20VH - HLT_e60_lhmedium - HLT_e120_lhloose Muons - HLT_mu20_iloose_L1MU15 - HLT_mu50
2016-2018	Electrons - HLT_e26_lhtight_nod0_ivarloose - HLT_e60_lhmedium_nod0 - HLT_e300_etcut Muons - HLT_mu26_ivarmedium - HLT_mu50

**Figure 4.4:** Trigger efficiencies in Run 2 for (a) the combined single-electron triggers as a function of the offline-reconstructed electron E_T and (b) combined single-muon triggers as a function of p_T . From refs. [156, 157].

Semi-leptonic $t\bar{t}$ events that pass the triggers described above are required to have one reconstructed lepton with $p_T > 27 \text{ GeV}$. This is a typical p_T threshold that has been used in many ATLAS measurements of semi-leptonic $t\bar{t}$ differential cross-sections in Run 2 [158]. Reconstructed leptons are geometrically matched to trigger-level lepton candidates. A cut on the missing transverse energy $E_T^{\text{miss}} > 20 \text{ GeV}$ and its sum with M_T^W ¹, the transverse component of the invariant mass of the leptonically decaying W -boson $E_T^{\text{miss}} + M_T^W > 60 \text{ GeV}$ is also applied. This accounts for the neutrino produced in the leptonic top decay and reduces the amount of fake leptons and multijet backgrounds.

Events are also required to have at least one large- R trimmed LC-Topo jet. This large- R jet must have $p_T > 350 \text{ GeV}$ and $|\eta| < 1.1$ to ensure that the entire width of the jet is contained within the pseudorapidity region $|\eta| < 2.1$ where there is a high tracking acceptance in the ID. The large- R jet must also be located at a sufficient distance $\Delta R(\ell, l_{\text{jet}}) > 2.3$ from the lepton.

Lastly, the events must contain 2 b -tagged small- R jets with $p_T > 25 \text{ GeV}$. The first of the b -jets is due to the leptonic top decay and is required to be located close to the lepton with $\Delta R(\ell, b - \text{jet}) < 2.3$. The second b -jet, due to the hadronic decay, is required to be located either inside or outside the large- R jet, depending on whether the jets to be studied are due to top or W decays.

Two selections are defined for the large- R jets: a "top"-jet selection and a W -jet selection. The "top" jets are defined as having a mass greater than 140 GeV and a nearby b -jet within $\Delta R < 1.0$ (the clustering of small- R and large- R jets is not exclusive, hence constituent topoclusters of the b -jet are also part of the large- R jet). The " W "-jets must fall within the mass window $60 \text{ GeV} < m < 100 \text{ GeV}$ and must have a b -jet located outside the jet cone at $\Delta R > 1.0$.

¹ $m_T^W = \sqrt{2p_T^\ell E_T^{\text{miss}} (1 - \cos \Delta\phi(p_T^\ell, E_T^{\text{miss}}))}$

Table 4.2: The selections for semi-leptonic $t\bar{t}$ events.

Object type	Detector level selection
Common semi-leptonic $t\bar{t}$ selection	
$R = 1.0$ Jet	• 1 trimmed LC-Topo jet • $p_T^{ljet} > 350$ GeV • $\eta ^{ljet} < 1.1$
$R = 0.4$ Jet	• ≥ 1 EM-PFLow jet • $p_T^{sjet} > 25$ GeV • 1 b-tag (DL1r, $\epsilon_{sig} = 77\%$)
Lepton	• 1 lepton with $p_T > 27$ GeV
Missing E_T	• $E_{miss}^T > 20$ GeV • $E_{miss}^T + M_W^T > 60$ GeV
Lepton + b -jet	• $\Delta R(lepton, b-jet) < 1.5$ • $m_{lepton+bjet} > 120$ GeV
$R = 1.0$ jet + lepton	• $\Delta R(lepton, ljet) > 2.3$
"top"-jet selection	
$R = 1.0$ jet	• $m_{ljet} > 140$ GeV
2^{nd} b -jet	• 1 b -tag (DL1r, $\epsilon_{sig} = 77\%$)
$R = 1.0$ jet + 2^{nd} b -jet	• $\Delta R(b-jet, R = 1.0 jet) < 1.0$)
"W"-jet selection	
$R = 1.0$ jet	• $60 \text{ GeV} < m_{ljet} < 100 \text{ GeV}$
2^{nd} b -jet	• 1 b -tag (DL1r, $\epsilon_{sig} = 77\%$)
$R = 1.0$ jet + 2^{nd} b -jet	• $\Delta R(b-jet, ljet) > 1.0$)

4.2.3.1 Event yields

Event yields for the various datasets evaluated in this analysis are shown in Table 4.3. The number of emissions for each selection in each of the samples is also shown. For the background samples, only the detector level yields are shown, as there are no particle-level selections applied to these samples.

As can be seen in Table 4.4, The ratio of the number of jets passing the detector level selections to those passing the particle level selections is reduced in the W/top jet selections compared to the common jet selection. This is due to the requirement that a second b -jet be present in the event. The factor by which the ratio of detector to particle-level events is reduced is approximately 0.77, which is equal to the efficiency of the b -tagger.

Table 4.3: Detector-Level event yields for the Data and Monte Carlo samples.

Sample Name	Common sel.	top jets	W jets
Number of Jets			
$t\bar{t}$ Powheg+Pythia8	107933	33750	3198
Data	102291	29328	3134
Single top	6786	652	212
$t\bar{t}+V$	1159	333	46
$W+e$ +jets	1820	55	23
$W+\mu$ +jets	1858	51	25
$W+\tau$ +jets	283	8	4
diboson	428	12	6
Z+jets	283	8	4
Fake leptons	2502	226	31
Ratio Data/MC	0.8485	0.8410	0.8902
Signal Purity	0.8953	0.9679	0.9085
Number of Emissions			
$t\bar{t}$ Powheg+Pythia8	770681	240892	21774
Data	730382	212224	21341
Single top	46646	4725	1392
$t\bar{t}+V$	8344	2456	322
$W+e$ +jets	14278	425	184
$W+\mu$ +jets	14961	406	210
$W+\tau$ +jets	2298	67	33
diboson	3161	93	47
Z+jets	2298	67	33
Fake leptons	17914	1634	217
Ratio Data/MC	0.8467	0.8518	0.8892
Signal Purity	0.8934	0.9669	0.9073

Table 4.4: Detector- and Particle-Level event yields for $t\bar{t}$ Monte Carlo samples.

Sample Name	Common sel.	top jets	W jets
Detector-Level Number of Jets			
$t\bar{t}$ Powheg+Pythia8	107933	33750	3198
$t\bar{t}$ Powheg+Herwig7.0	94916	29914	2704
$t\bar{t}$ aMCatNLO+Pythia8	97351	32090	2600
$t\bar{t}$ PP8 hdampVar	111878	34499	3389
$t\bar{t}$ Powheg+Herwig7.2	102734	31028	3086
$t\bar{t}$ Sherpa 2.2.10	72077	22469	2237
$t\bar{t}$ PP8 ME. Corr. Off	104734	32145	2897
Particle-Level Number of Jets			
$t\bar{t}$ Powheg+Pythia8	140468	59288	5754
$t\bar{t}$ Powheg+Herwig7.0	125152	53652	4905
$t\bar{t}$ aMCatNLO+Pythia8	130041	58286	4746
$t\bar{t}$ PP8 hdampVar	145545	60336	5930
$t\bar{t}$ Powheg+Herwig7.2	134501	55044	5475
$t\bar{t}$ Sherpa 2.2.10	91828	38298	3862
$t\bar{t}$ PP8 ME. Corr. Off	138011	57167	5182
Detector-Level Number of Emissions			
$t\bar{t}$ Powheg+Pythia8	770681	240892	21774
$t\bar{t}$ Powheg+Herwig7.0	681027	219380	18326
$t\bar{t}$ aMCatNLO+Pythia8	694091	233208	17239
$t\bar{t}$ PP8 hdampVar	802349	246540	23179
$t\bar{t}$ Powheg+Herwig7.2	748462	225899	21837
$t\bar{t}$ Sherpa 2.2.10	524284	167856	15451
$t\bar{t}$ PP8 ME. Corr. Off	753671	231764	19811
Particle-Level Number of Emissions			
$t\bar{t}$ Powheg+Pythia8	1584573	667105	61656
$t\bar{t}$ Powheg+Herwig7.0	930892	410042	33699
$t\bar{t}$ aMCatNLO+Pythia8	960873	441285	31951
$t\bar{t}$ PP8 hdampVar	1078781	448932	41188
$t\bar{t}$ Powheg+Herwig7.2	1012401	418089	39290
$t\bar{t}$ Sherpa 2.2.10	694773	303146	26826
$t\bar{t}$ PP8 ME. Corr. Off	1037283	433164	36291

4.2.3.2 Detector-level jet kinematic observables

The detector-level distributions of jet mass and p_T after cuts have been applied are shown in 4.5 and 4.6 for the top and W jet selections respectively.

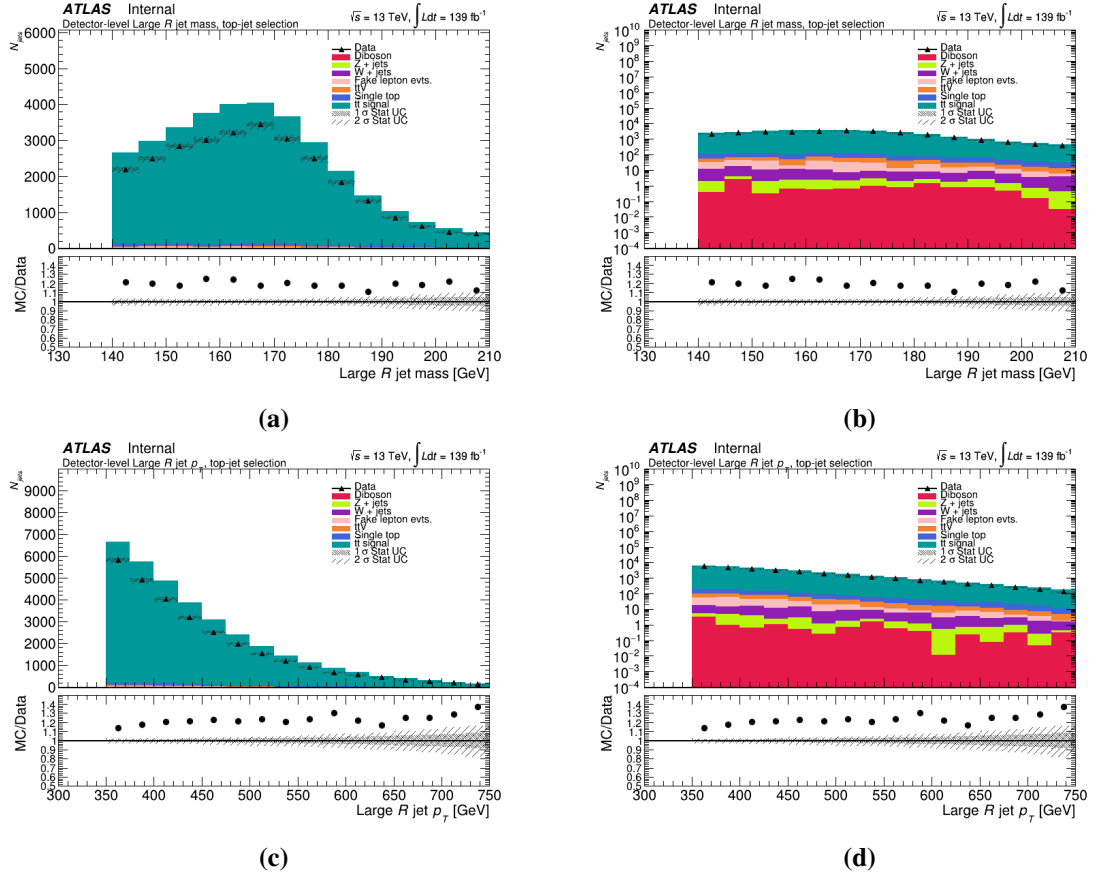


Figure 4.5: Detector-level large- R jet p_T and mass for the "top" jet selection.

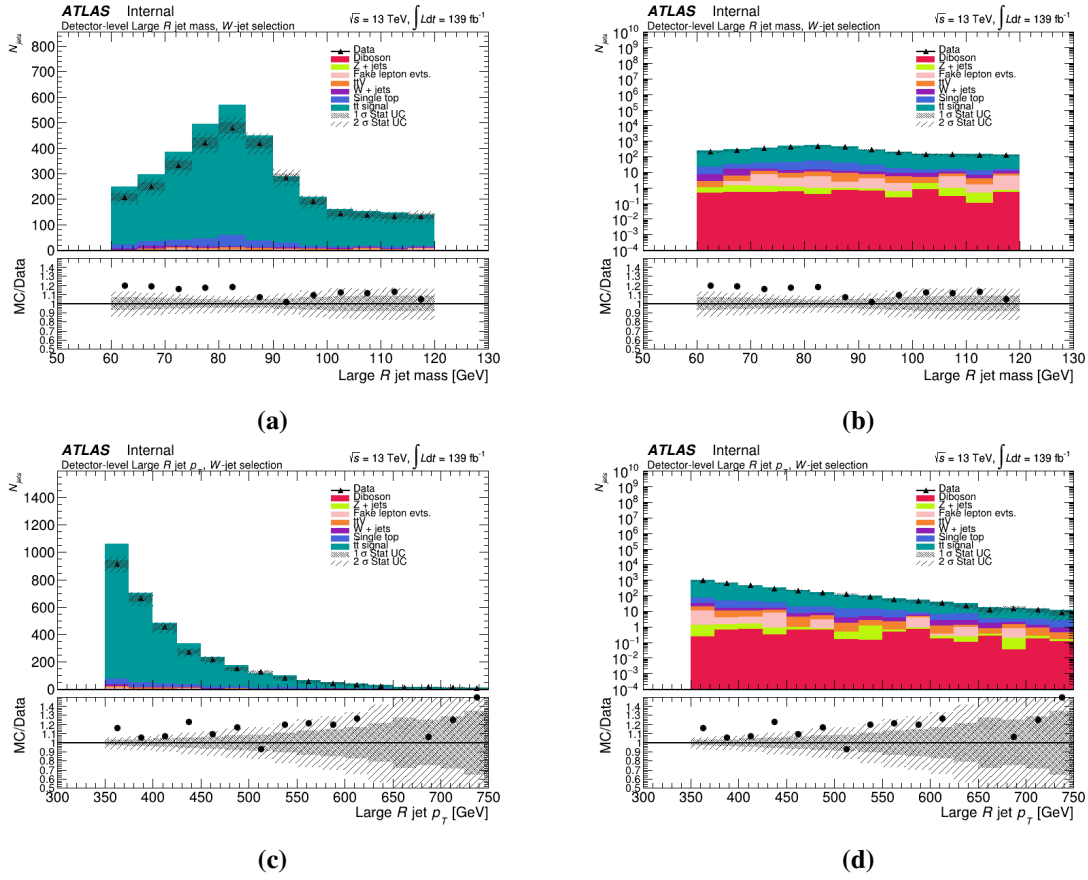


Figure 4.6: Detector-level large- R jet p_T and mass for the W jet selection.

4.2.4 Object definitions

Electrons are reconstructed from inner detector information which is combined with calorimeter (ECAL) information for improved p_T reconstruction [159, 160]. The combined electrons are constructed by matching ID tracks to clusters in the Liquid Argon calorimeter. The quality of the reconstructed electrons must satisfy the `TightLH` likelihood-based working point as well as the `FCTight` isolation working point defined in ref. [161]. Combined electron tracks are associated to the primary vertex by requiring the cuts $|d_0/\sigma_{d_0}| < 0.5$ and $|\Delta z_0| < 0.5$ mm on the transverse and longitudinal impact parameters d_0 and z_0 . Combined electron calorimeter clusters are required to lie within the region $|\eta| < 2.47$ where there is sufficient ID information, as well as outside the barrel-endcap gap $1.37 < |\eta| < 1.52$ of the ECAL. Lastly, the combined electrons are also required to have $p_T > 27$ GeV.

Muons are reconstructed by combining inner detector tracks with muons reconstructed in the spectrometer. Combined muon ID tracks are associated to the primary vertex by requiring $|d_0/\sigma_{d_0}| < 0.5$ and $|\Delta z_0| < 0.5$ mm. The muons are also required to be within the ID acceptance region $|\eta| < 2.47$. The muons must satisfy the `Medium` reconstruction working point and the `FCTight` isolation working point defined in ref. [125]. The combined muons are also required to have $p_T > 27$ GeV.

Tracks are reconstructed using the standard reconstruction procedure (see section 1.3.2) adapted to the dense tracking environment experienced in ATLAS Run 2 [83]. The tracks are required to pass the `TightPrimary` working point as defined in ref. [162]. Track candidates are required to have at least 9 hits in the silicon detector if $|\eta| < 1.65$ or 11 hits in the silicon detector if $|\eta| > 1.65$. No two tracks are allowed to share more than one hit in each module of the Pixel and SCT. Tracks candidates must have no more than two missing hits (holes) in total and none in the Pixel detector or the IBL if it lies in the path of the track candidate. High quality track reconstruction is of great importance to this analysis, since tracks are used to examine the jet substructure and construct the Lund Jet Plane. Tracks are associated to Large-R jets using the ghost-association procedure [97, 127].

Small- R jets are reconstructed from particle-flow objects using the anti- k_T algorithm with distance parameter $R = 0.4$. The effect of pileup is suppressed using the Jet Vertex Tagger (JVT) by requiring jets with $p_T < 60 \text{ GeV}$ to have a JVT discriminant of greater than 0.5. The energy of the particle-flow jets is calibrated so that it matches the expected Particle-Level (EM-scale) energy measurement [127]. Jets are required to fall within the ID acceptance region $|\eta| < 2.5$ and must have $p_T > 25 \text{ GeV}$.

B-tagging is performed on the small- R jets using the DL1r machine-learning based algorithm [129, 163]. A cut is made on the output discriminant to give a b jet identification efficiency of 77% for jets in the region $20 < p_T < 600 \text{ GeV}$. The information that is input into the DL1r machine-learning algorithm includes the jet p_T and $|\eta|$, as well as information from high-impact-parameter tracks associated to the jet, the displaced secondary vertex and other topological features of particle decays that are visible in the ID.

Large- R jets are reconstructed from calorimeter topoclusters using the anti- k_T algorithm with distance parameter $R = 1.0$. The energy of the calorimeter clusters is calibrated for each cluster using the Local Cluster Weighting (LCW) scheme. The Trimming algorithm (see 1.5.1) is applied to remove radiation due to pileup. The parameters used for the trimming are $R_{trim} = 0.2$ and $f_{trim} < 0.05$. The jet p_T , η and mass response are calibrated using the MC-based jet calibrations techniques and data-based in-situ mass and energy calibration [134]. All large- R jets are required to have $|\eta| < 2.0$ and $p_T > 200 \text{ GeV}$ before subsequent selections are applied.

Missing E_T is reconstructed by calculating the negative vector sum of all calibrated objects in the event. This calculation considers only the small- R particle-flow jets, which are allowed to overlap with the large- R jets. The sum also includes a "soft" term which is calculated from the 4-momenta of ID tracks which are not associated to any of the reconstructed objects in the event. A non-zero missing transverse energy is expected for semi-leptonic $t\bar{t}$ events due to the presence of a neutrino from the leptonically decaying top quark.

Overlap removal is performed on the reconstructed leptons and small- R jets to avoid double counting their 4-momenta in the reconstruction of the missing E_T . Small- R jets that are close to electrons with $\Delta R(j, e) < 0.2$ are removed. So are small- R jets close

to muons with $\Delta R(j, \mu) < 0.2$ or share at least one track with the muon if there are fewer than three tracks contained within the jet. Any electrons or muons that satisfy $\Delta R(j, \ell) < 0.4$ for any of the remaining small- R jets are also removed. Small- R jets and large- R jets are constructed from the same collection of ID tracks and Calorimeter clusters and are allowed to overlap.

4.2.5 Particle level definitions

When modelling the properties of top quarks, particle physicists often distinguish between predictions at the level of partons and at the level of hadrons and other colour neutral particles. The top is unique among the quarks as predictions about its properties, such as its mass, can be tested at the parton level. There are therefore two types of MC "truth" in top physics: the parton level truth, and the *particle level* truth. As the Lund Jet Plane is a property of the stable hadrons that make up jets, it only makes sense to model it at the particle level. To avoid confusion with the terminology used in other measurements of top decays we use the term "particle level" to denote MC predictions at the level of hadrons. In the measurement of the LJP in dijet events [115] these were simply referred to as the "truth level".

Particle level event selections are shown in table 4.5. Event yields for these selections are presented in table 4.4. The particle-level phase space is defined by the kinematics of stable particles, meaning truth particles with a decay length of $c\tau > 10$ mm.

Small- R Jets Small- R jets are reconstructed from stable truth particles with $|\eta| < 5.0$ using the anti- k_T algorithm with $R = 0.4$. Leptons and neutrinos are excluded from the jet reconstruction procedure.

Large- R Jets Large- R jets are reconstructed from stable truth particles with $|\eta| < 5.0$ using the anti- k_T algorithm with $R = 1.0$. Leptons and neutrinos are excluded from the jet reconstruction procedure. Trimming [102] is applied to the jet constituents using the same parameters as for the detector-level topocluster jets: $R_{trim} = 0.2$ and $f_{trim} < 0.05$.

Lund Jet Plane At the particle-level, the Lund Jet Plane is reconstructed from charged truth particles. These are selected from the constituents of the trimmed anti- k_T $R = 1.0$ truth jets. The charged jet constituents are re-clustered using the Cambridge-Aachen

algorithm and the LJP is constructed from the clustering sequence. Only stable particles with $p_T > 500 \text{ MeV}$ excluding muons and neutrinos are used.

b -jets Jets originating from b -quark decays are identified at particle-level using ghost-associated B -hadrons. A small- R jet is considered a b -jet if it has at least one ghost-associated B -hadron with $p_T > 5$

Leptons Particle-level electrons and muons are identified in the truth record by requiring that they do not originate from the decays of hadrons or taus. The leptons are dressed, meaning that the four-momentum of photons within $\Delta R < 0.1$ are added to the four-momentum of the truth lepton. The selections applied to the dressed leptons are the same as for the reconstructed leptons.

Missing- E_T The missing transverse energy at the particle-level is calculated from the four-momenta of neutrinos. In the truth record, the neutrinos are required to have originated in the same electroweak decays as any leptons that satisfy the conditions above.

Table 4.5: The particle-level event selections for semi-leptonic $t\bar{t}$ events.

Object type	Particle-level selection
Common semi-leptonic $t\bar{t}$ selection	
$R = 1.0$ Jet	• 1 trimmed anti-k_T truth-particle jet • $p_T^{ljet} > 350 \text{ GeV}$ • $\eta ^{ljet} < 1.1$
$R = 0.4$ Jet	• ≥ 1 anti-k_T truth-particle jet • $p_T^{sjet} > 25 \text{ GeV}$ • 1 ghost associated B-hadron
Lepton	• 1 dressed truth e/μ with $p_T > 27 \text{ GeV}$, $ \eta < 2.5$
Missing E_T	• $E_{miss}^T > 20 \text{ GeV}$ • $E_{miss}^T + M_W^T > 60 \text{ GeV}$
Lepton + b -jet	• $\Delta R(lepton, b-jet) < 1.5$
$R = 1.0$ jet + lepton	• $\Delta R(lepton, ljet) > 2.3$
"top"-jet selection	
$R = 1.0$ jet	• $m_{ljet} > 140 \text{ GeV}$
2^{nd} b -jet	• 1 ghost associated B -hadron
$R = 1.0$ jet + 2^{nd} b -jet	• $\Delta R(b-jet, R = 1.0 jet) < 1.0$
"W"-jet selection	
$R = 1.0$ jet	• $60 \text{ GeV} < m_{ljet} < 100 \text{ GeV}$
2^{nd} b -jet	• 1 ghost associated B -hadron
$R = 1.0$ jet + 2^{nd} b -jet	• $\Delta R(b-jet, ljet) > 1.0$

4.2.6 Monte Carlo simulation

The Lund Jet Plane measurement is compared to predictions from Monte Carlo simulations. Matrix element calculations, parton showering and hadronisation are simulated using a number of different frameworks according to current recommendations by the ATLAS simulations subgroup. Separate Monte Carlo samples are generated for the signal $t\bar{t}$ process, as well as the dominant background processes. Sources of systematic modelling uncertainties are also evaluated from comparisons between the nominal Monte Carlo samples and a number of alternative Monte Carlo samples. The particles simulated in each of these samples are input into a detector simulation using the `Geant 4` software framework [123] in order to simulate the response of the entire ATLAS detector.

The effect of multiple proton-proton interactions per bunch crossing (pileup) is simulated using Minimum-Bias events generated using Pythia 8.186 [121] with the NNPDF2.3LO [140] set of PDFs using the A3 tune [122]. These events are overlaid onto the simulated hard-scatter processes for all signal and background MC samples. An event-by-event pileup-reweighting is also performed using the ATLAS recommended `PileupRewightingTool` in order to match the observed number of average interactions per bunch crossing $\langle\mu\rangle$ observed in Run 2.

4.2.6.1 Nominal $t\bar{t}$ signal sample

The production of $t\bar{t}$ events is modelled using the PowhegBox v2 [118, 43, 119, 120] generator. PowhegBox provides matrix elements at next-to-leading order (NLO) in the strong coupling constant α_s with the NNPDF3.0NLO [164] parton distribution function and the h_{damp} parameter set to $1.5m_t$ [165]. The events are interfaced with Pythia 8.230 [38] for the parton shower and hadronisation, using the A14 set of tuned parameters [141] and the NNPDF2.3LO set of PDFs [140]. The decays of bottom and charm hadrons are simulated using the EvtGen v1.6.0 program [166].

The sample is normalised to the total $t\bar{t}$ cross-section prediction at next-to-next-to-leading order (NNLO) in QCD including the re-summation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms calculated using `Top++ 2.0` [167–173]. For proton–proton collisions at a centre-of-mass energy of $\sqrt{s} = 13\text{ TeV}$, this cross section corresponds to $\sigma(t\bar{t})_{NNLO+NNLL} = 832 \pm 51\text{ fb}$ using a top-quark mass of $m_t = 172.5\text{ GeV}$.

Modelling uncertainties on the total cross-section due to PDF and α_s are calculated following the PDF4LHC prescription [174] using the alternative PDF sets: MSTW2008 68% CL NNLO [175, 176], CT10 NNLO [177, 178] and NNPDF2.3 with fixed flavour number ($n_f = 5$) [140].

4.2.6.2 Alternative $t\bar{t}$ signal samples

In addition to the nominal Powheg + Pythia 8 sample, range of alternative Monte Carlo event generators are compared at the particle level to the unfolded data. The alternative generators include alternative matrix element generators (Powheg vs. Sherpa vs. aMC@NLO) as well as alternative parton shower and hadronisation models (Pythia 8 vs. Herwig vs. Sherpa). As the Lund Jet Plane covers a wide range of energies within the jets, it is expected to be sensitive to differences in the modelling of parton showers and hadronisation as well as the matching of the parton shower to the matrix element. Samples using three different versions (7.0, 7.1 and 7.2) of the Herwig parton shower and hadronisation model are also included as alternatives, which could allow us to identify whether improvements in modelling of the Lund Jet Plane was achieved as the generator was updated from one version to the next. We also compare an alternative configuration of Powheg using a different value of the free parameter h_{damp} . As h_{damp} controls the matching between the parton shower and the matrix element generator, it is expected to have an impact on the LJP distributions. The alternative h_{damp} sample together with the Powheg+Herwig 7.0 and aMC@NLO+Pythia 8 samples are also used at the detector level to estimate modelling systematic uncertainties.

Powheg+Herwig 7.0 The impact of the parton shower and hadronisation model is evaluated by comparing the nominal generator setup with a sample produced with the Powheg-Box v2 [118, 43, 119, 120] generator which provides matrix elements at next-to-leading order (NLO) in the strong coupling constant α_s with the NNPDF3.0nlo [164] parton distribution function (PDF) and the h_{damp} parameter set to $1.5 m_t$ [165]. The functional form of the renormalisation and factorisation scale is set to the default scale $\sqrt{m_{top}^2 + p_T^2}$. The events are interfaced with Herwig 7.04 [34, 35], using the H7UE set of tuned parameters and the MMHT2014LO PDF set [144]. The decays of bottom and charm hadrons are simulated using the EvtGen v1.6.0 program [166].

Powheg+Herwig 7.1.3 Another alternative parton showering model is provided by Herwig version 7.1.3 [143]. This version includes many improvements in over the 7.0 version, including multijet merging at next-to-leading order QCD and a new model for soft radiation in multi-parton-interactions (MPI). PowhegBox v2 with identical parameters to the nominal sample is used to generate the matrix elements.

Powheg+Herwig 7.2 events are interfaced with Herwig version 7.2.1 [179]. This is the newest version of Herwig. Changes to this version include improvements to the modelling of both perturbative and non-perturbative effects. Changes to non-perturbative processes includes corrections to the production of soft $g \rightarrow s\bar{s}$ radiation, which increases the ratio of kaons to pions in the sample. Matrix elements are generated using Powheg-Box v2.

hdampVar The impact of the choice of the h_{damp} parameter in Powheg is evaluated by comparing the nominal sample to a Powheg + Pythia 8 sample with identical settings except $h_{damp} = 3.0m_t$. The h_{damp} parameter controls the matching between Powheg and Pythia 8 events and modifies the amount of high- p_T radiation that the $t\bar{t}$ system recoils against.

Sherpa 2.2.10 Alternative matrix-element and parton-showering+hadronisation modelling is provided by Sherpa version 2.2.10 [47]. Matrix elements are calculated at NLO for up to two partons, and at LO for up to four partons using the Comix [41] and OpenLoops 1 [180–182] libraries. Events are matched to the Sherpa parton shower simulator [36]. The Sherpa parton shower simulations use the MEPS@NLO prescription with a standard set of tunes developed by the Sherpa authors [183–186]. The NNPDF3.0nnlo [164] set of PDFs is used and the samples are normalised to NNLO predictions [187].

aMC@NLO+Pythia 8 An alternative matrix element generator, MadGraph5_aMC@NLO (often referred to using the shorthand aMC@NLO), is used to estimate the uncertainty associated with the hard-scatter component of the Monte Carlo simulations. For the calculation of the hard-scattering, MadGraph5_aMC@NLO v2.6.0 with the NNPDF3.0NLO [164] PDF set is used. The events are interfaced with Pythia 8.230 [38], using the A14 set of tuned parameters [141] and the NNPDF2.3lo set of PDFs [140]. Top quarks are decayed at LO using MadSpin [188, 189] to preserve spin correlations. The decays of

bottom and charm hadrons are simulated using the EvtGen v1.6.0 program [166]. The shower starting scale has the functional form $\mu_q = H_T/2$ [190], where H_T is defined as the scalar sum of the p_T of all outgoing partons. The renormalisation and factorisation scale are chosen to match the nominal Powheg setup.

4.2.6.3 Monte Carlo background modelling

Background events are simulated for different processes using a collection of Monte Carlo generators. The hard processes that give rise to the different sources of background are simulated using different Matrix Element generators which have been chosen separately for each process to maximise modelling accuracy. Matrix element calculations are interfaced with parton showering and hadronisation simulations. For the Single top, $t\bar{t}$ + vector boson, $t\bar{t}H$ and tW events, the interfacing is performed with Pythia 8.230 [38] with the A14 tune [141] and the NNPDF2.3lo [140] set of PDFs. For W +jets, Z +jets and diboson events, where Matrix elements are simulated using Sherpa, the Sherpa parton showering and hadronisation simulations are used instead. The Sherpa parton shower simulations [36] use the MEPS@NLO prescription with a standard set of tunes developed by the Sherpa authors [183–186].

The full list of background processes and the tools used to simulate them is as follows:

Single-top matrix elements are modelled using the PowhegBox v2 [43, 119, 120, 191, 192] generator at NLO in QCD using either the four-flavour (t -channel) or five-flavour (s -channel) scheme and the corresponding NNPDF3.0nlo [164] set of PDFs.

W +jets and Z +jets event matrix elements are simulated with the Sherpa v2.2.1 [47] generator using NLO matrix elements for up to two partons, and LO matrix elements for up to four partons calculated with the Comix [41] and OpenLoops 1 [180–182] libraries. The NNPDF3.0nnlo [164] set of PDFs is used and the samples are normalised to NNLO predictions [187].

$t\bar{t}$ with associated vector boson ($t\bar{t}W$, $t\bar{t}Z$) events are modelled using the MadGraph5_aMC@NLO v2.3.3 [46] generator at NLO with the NNPDF3.0nnlo [164] PDF. Decays of bottom and charm hadrons are simulated using the EvtGen v1.2.0 [166] program.

$t\bar{t}$ with associated Higgs boson ($t\bar{t}H$) events are modelled using the PowhegBox v2 generator at NLO [43, 119, 120, 191, 193] with the NNPDF3.0nlo [164] PDF set. The decays of bottom and charm hadrons from these events are simulated using EvtGen v1.6.0 [166].

Diboson events (WW , ZZ and WZ) are simulated using the Sherpa v2.2.2 [47] matrix element generator, including off-shell effects and Higgs-boson contributions. The matrix elements are calculated to NLO accuracy in QCD for up to one additional parton and to LO accuracy for up to three additional parton emissions. The matrix element calculations are matched and merged with the Sherpa parton shower simulator based on Catani-Seymour dipole factorisation [36, 41]. Virtual QCD corrections are provided by the OpenLoops 1 [180–182] library. The NNPDF3.0nnlo [164] set of PDFs is used along with the dedicated set of tuned parton-shower parameters developed by the Sherpa authors.

Top quarks with associated W boson (tW) events are modelled using the PowhegBox v2 [43, 119, 120, 191, 194] matrix element generator at NLO in QCD using the five-flavour scheme and the NNPDF3.0nlo [164] set of PDFs. The diagram removal (DR) scheme [195] is used to remove interference and overlap with $t\bar{t}$ production. The related uncertainty is estimated by comparing with an alternative sample generated using the diagram subtraction (DS) scheme [165, 195].

4.3 Background estimation

4.3.1 Estimation of "fake" lepton background events

Some of the reconstructed leptons in the selection are expected to be due to events in which no leptons are produced in the hard-scatter process. These processes are either leptonically decaying high- p_T hadrons or hadronic jets which have punched through the calorimeters into the Muon spectrometer so that they are reconstructed as muons. Events in which these misreconstructed leptons are present are referred to as *fakes*.

The number of fakes is estimated from the number of events in data by varying the working point on the lepton reconstruction quality. By comparing the number of events passing the *tight* requirements to the more inclusive *loose* requirements, it is possible to get an estimate of the number of fakes. The estimation is performed using the method

outlined in ref. [196]. The number of fake lepton events passing the tight selection N_f^t is related to the total number of events in the tight selection N^t and the total number of events in the loose selection N^l through the formula

$$N_f^t = \frac{\varepsilon_f}{\varepsilon_r - \varepsilon_f} (\varepsilon_r N^l - N^t), \quad (4.1)$$

where ε_r and ε_f are the fractions of real, respectively fake, leptons that pass both the tight and loose requirements. It is possible to calculate an event-by-event weighting w_i that reshapes the distribution of all the events in the tight selection to the number of fakes:

$$w_i = \frac{\varepsilon_f}{\varepsilon_r - \varepsilon_f} (\varepsilon_r - \delta_{tl}^i) \quad (4.2)$$

where δ_{tl} is equal to 1 if the event passes both the tight and loose lepton requirements and 0 otherwise. Hence, it is only necessary to know the efficiencies ε_r and ε_f in order to estimate the distributions of fake background events. The efficiencies have been measured in experiment by the ATLAS top reconstruction group and the resulting weights are retrieved using a tool in the `AnalysisTop` analysis software framework.

4.3.2 Difference in the total $t\bar{t}$ cross-section between data and Monte Carlo

A difference in the value of the total $t\bar{t}$ cross-section between Data and Monte Carlo simulation is observed. The predicted total $t\bar{t}$ cross-section is about 14% larger than in Data for semi-leptonic $t\bar{t}$ events. This discrepancy is well-known and has been observed in semi-leptonic top cross-section measurements at ATLAS in Run 1 [197], Run 2 [198, 158, 199] and early Run 3 data [200]. The same effect has also been observed in the fully hadronic decay channel [201], where the difference between the predicted and measured cross-section is around 17%. The CMS experiment also observed a difference in the total cross-section in both decay channels [202]. The discrepancy is believed to be due to higher-order corrections to the NLO $t\bar{t}$ matrix element calculations in Powheg. The K -factor, the factor that is used to scale the total cross-section of the events generated at NLO to NNLO predictions, is constant despite there being a jet p_T dependence in the

actual higher-order corrections. This leads to a discrepancy between the predicted and measured cross-sections that is especially pronounced at high p_T . A cross-check has been performed to make sure that a possible similar mismodelling of the total cross-sections in the background samples is captured by the systematic uncertainties (see section 4.3.4). As the LJP is normalised to the total number of jets, differences in the total $t\bar{t}$ cross-section between the measurement and the MC samples do not affect the distributions of the LJP density.

4.3.3 Detector level signal and background

The detector level Lund Jet Plane is shown in figures 4.7 and 4.8 for the top jets and W jets. Signal and background Monte Carlo, as well as the data are shown for slices through the LJP. The LJP is normalised to the total number of jets in the Data and Monte Carlo.

The distribution of signal and background events across the Lund Jet Plane is generally uniform. There are no bins in which background events dominate and the signal purity varies between 96 and 98%, in the case of top jets, and 87 to 98%, in the case of W jets, across the most populated areas of the LJP. It is only at very small ΔR and z , where radiation from non-perturbative effects dominates, that the purity drops to around 90% for the top jets and about 80% for the W jets. The dominant source of background events across the LJP are single top decays, whose contributions to the number of emissions in each bin varies between 1 and 3% for top jets and between 4 and 8% for W jets. The second most dominant source of background are $t\bar{t} + V$ events in the top jet selection, whose contributions is 0.9 – 1.2%, and W +jets in the W jet selection, which contribute 0.6 – 4.8% in the central part of the LJP.

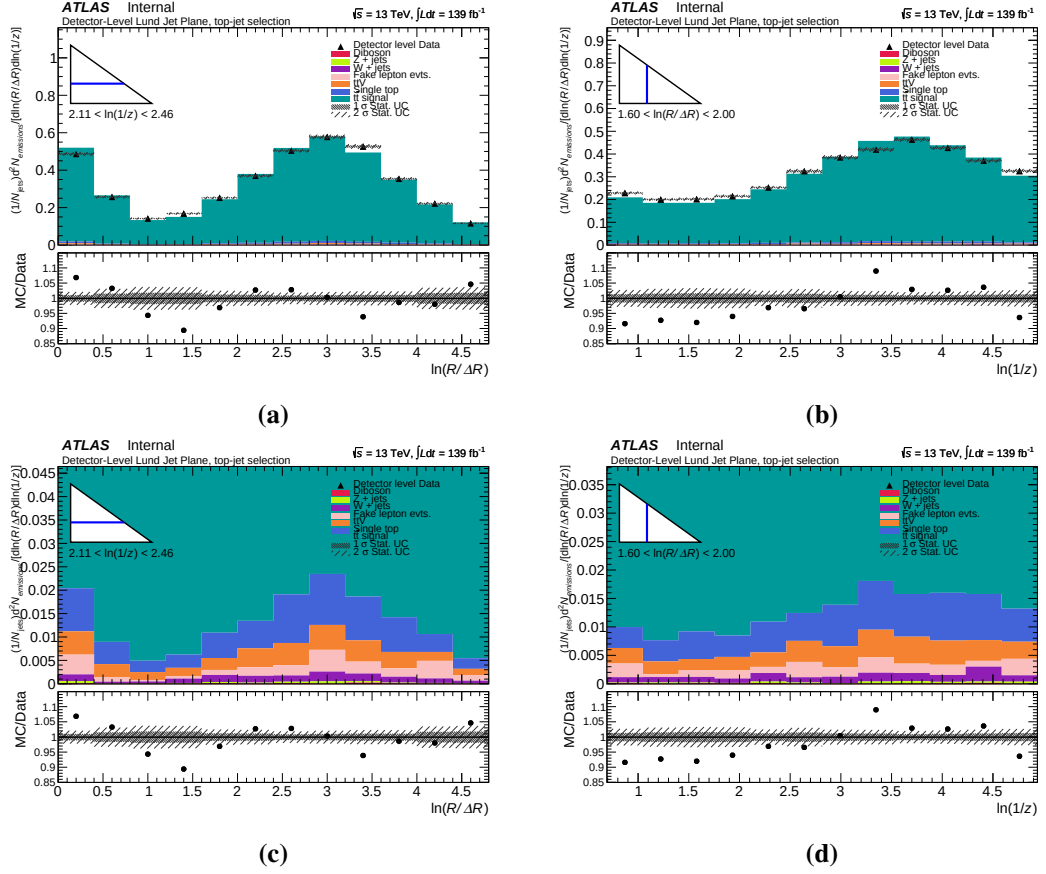


Figure 4.7: Detector-level Lund Jet Plane slices for the top jet selection. One slice along $\ln(R/\Delta R)$ (a,c) and $\ln(1/z)$ (b,d) is shown with the full signal+background stack (a,b) and the background portion of the MC zoomed in (c,d).

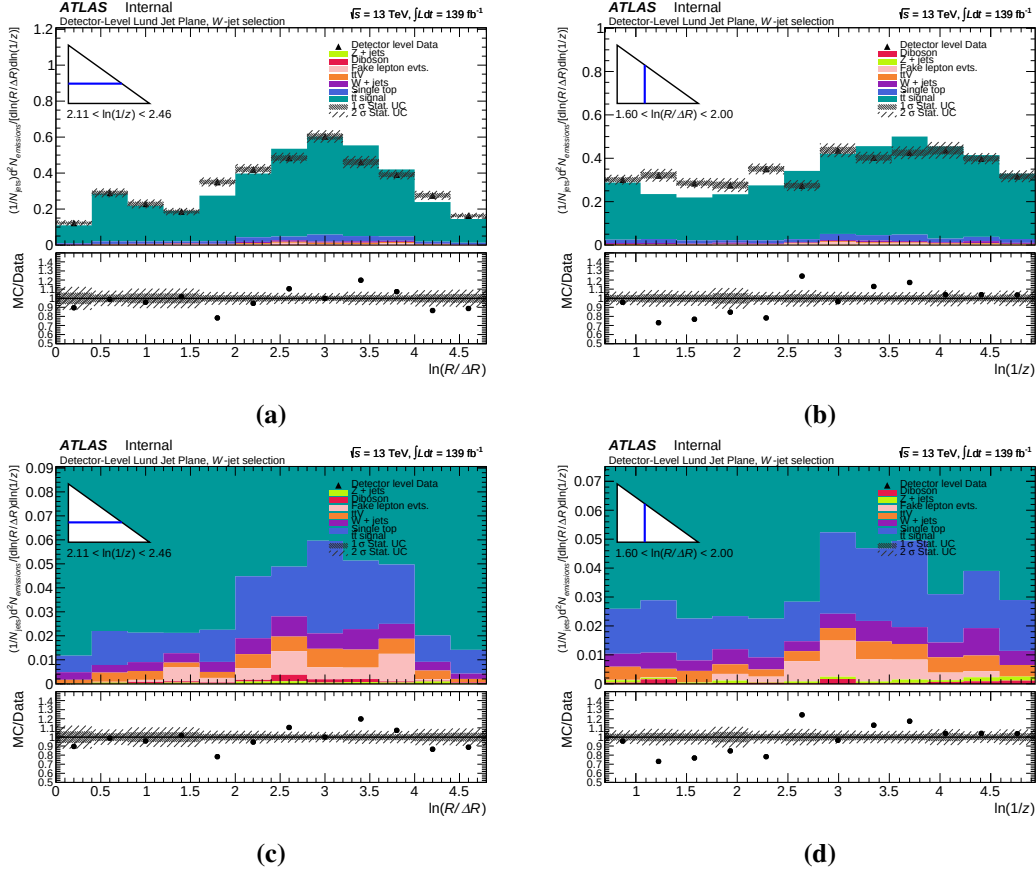


Figure 4.8: Detector-level Lund Jet Plane slices for the W jet selection. One slice along $\ln(R/\Delta R)$ (a,c) and $\ln(1/z)$ (b,d) is shown with the full signal+background stack (a,b) and the background portion of the MC zoomed in (c,d).

4.3.4 Detector-level background cross-check

To gauge the impact of mismodelling in the background samples, a cross-check is performed by selecting jets in a high-background control region. The control region is defined by the following four cuts

$$N_{sjets} \leq 5; \quad N_{bTag} \geq 1; \quad m_{lb} > 150 \text{ GeV}; \quad 60 < m_{ljet} < 120 \text{ GeV}, \quad (4.3)$$

where N_{sjets} is the number of small- R jets, N_{bTag} is the number of b -tagged jets, m_{ljet} is the large- R jet mass and m_{lb} is the combined mass of the lepton- b -jet system. The observables are defined identically to those used in the signal definitions. Note that although m_{lb} is not an observable used in the signal definition, it is useful for selecting a region of high background purity as most $t\bar{t}$ events fall into the region below 120 GeV. The value of m_{lb} is calculated by summation of the lepton and b -jet four vectors.

The detector-level large- R jet p_T and mass distributions in the control region are shown in figure 4.9 along with the associated total systematic uncertainty on the background. A breakdown of the individual background modelling systematics is shown in figure 4.10. A breakdown of the detector systematics associated with the background MC is available in appendix B.4. The purity of the background in the control region is estimated at 50.2%. As can be seen in these figures, the uncertainties on the background are larger than the mismodelling that we observe in the control region. Most of the uncertainties on the background are due to modelling, however we do also observe substantial contributions from detector systematics, particularly those associated with the large- R jets.

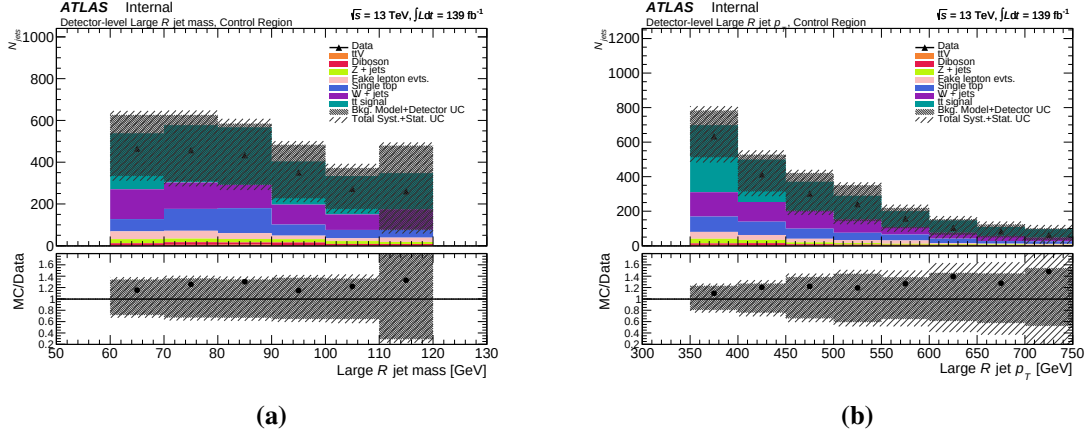


Figure 4.9: Detector-level large- R jet mass (a) and p_T (b) in the control region. The background uncertainties and the total systematic+statistical uncertainty is also shown.

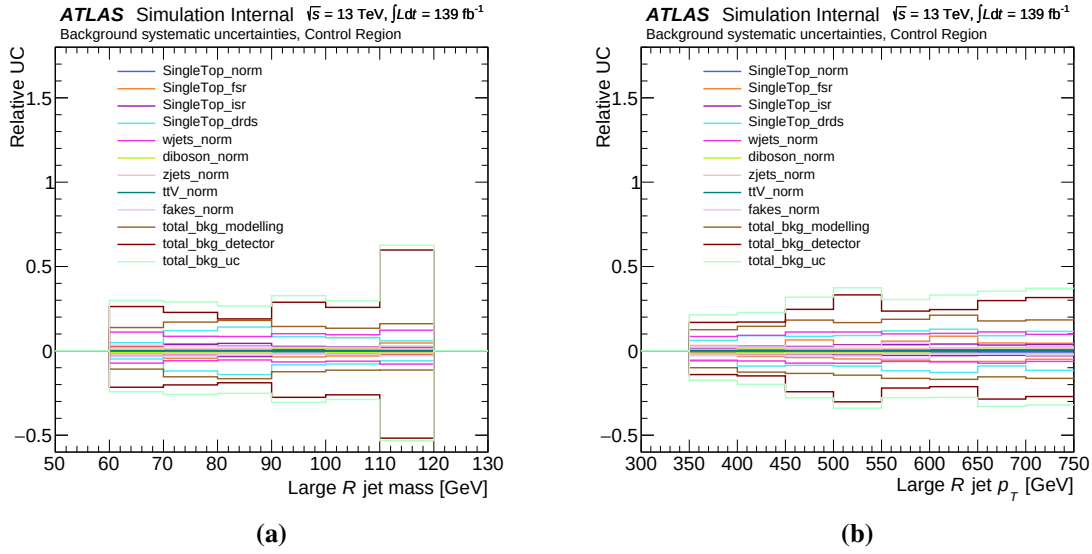


Figure 4.10: Background modelling uncertainties vs. large- R jet mass (a) and p_T (b) in the control region. The uncertainties is evaluated for unnormalised detector-level distributions.

4.4 Unfolding

The unfolding is performed using the implementation of iterative Bayesian unfolding available in the `RoUnfold` [148] framework. The iterative Bayesian unfolding procedure is largely identical to the one used to unfold the LJP in dijets and it is described in detail in section 3.1.4. The detector level histograms that are obtained from the data are unfolded to the particle level using a response matrix derived from the nominal Powheg+Pythia 8 Monte Carlo samples.

4.4.1 The response matrix

The response matrix is constructed by filling a 2D histogram with the detector-level bins of the Lund Jet Plane along one axis and the particle-level bins along the other. Since this must be done for every emission in every event, it is necessary to match detector-level emissions to particle-level emissions. Due to the presence of pileup and fake tracks, the number of tracks contained within the jet – and hence the number of emissions in the LJP – is often different. The LJP coordinates of an emission at detector- and particle-level can also differ if tracks are excluded or added to the declustering. Therefore, a geometric matching procedure must be performed on the detector- and particle level emissions. The matching is performed in $\eta - \phi$ space and is described in greater detail in section 4.4.3.

The response for slices through the LJP are shown in figure 4.11. The fraction of entries on the diagonal is also shown. Typically, the fraction of entries on the diagonal of the response matrix is higher for slices along the $\ln(R/\Delta R)$ axis, where it is around 85 – 92%, than along the $\ln(1/z)$ axis, where it is around 60 – 85%. There is, in other words, a broader distribution of possible particle level z coordinates for a given detector level z than there is for a detector level ΔR .

Bin-to-bin migrations are shown in figure 4.12. The migration of emissions between bins at the detector and particle level is quantified by the conditional probabilities $P_{ij}(DL|PL)$ and $P_{ij}(PL|DL)$ which are calculated from the response matrix. The plots in figure 4.12 show $P_{ij}(DL|PL)$, the probability that an emission is reconstructed in bin i at the detector level given that it falls into bin j at the particle level. Most emissions, across most of the LJP, fall into the diagonal of the migration matrices, meaning they

are reconstructed in the same bins at the detector and particle level. The remaining migrations go into and from adjacent bins in a mostly symmetrical distribution. There is some asymmetry in the migrations at very wide and very small emission angles, approximately $\ln(R/\Delta R) < 1.2$ and $\ln(R/\Delta R) > 4.0$. The small-angle effects can be explained by the limited resolution of the detector, which is also reflected by the size of the tracking systematic uncertainties in this region (see section 4.6.6), causing small angles to be systematically misreconstructed as larger ones. The wide-angle effects could be the result of emissions at wide opening angles, or close to the edge of the jet, being clustered at the particle level but not at the detector level or vice-versa. The acceptance and inefficiency factors, which represent the fraction of detector and particle level emissions respectively that enter into response matrix, are also very small in the wide-angle region of the plane, as is discussed in section 4.4.2.

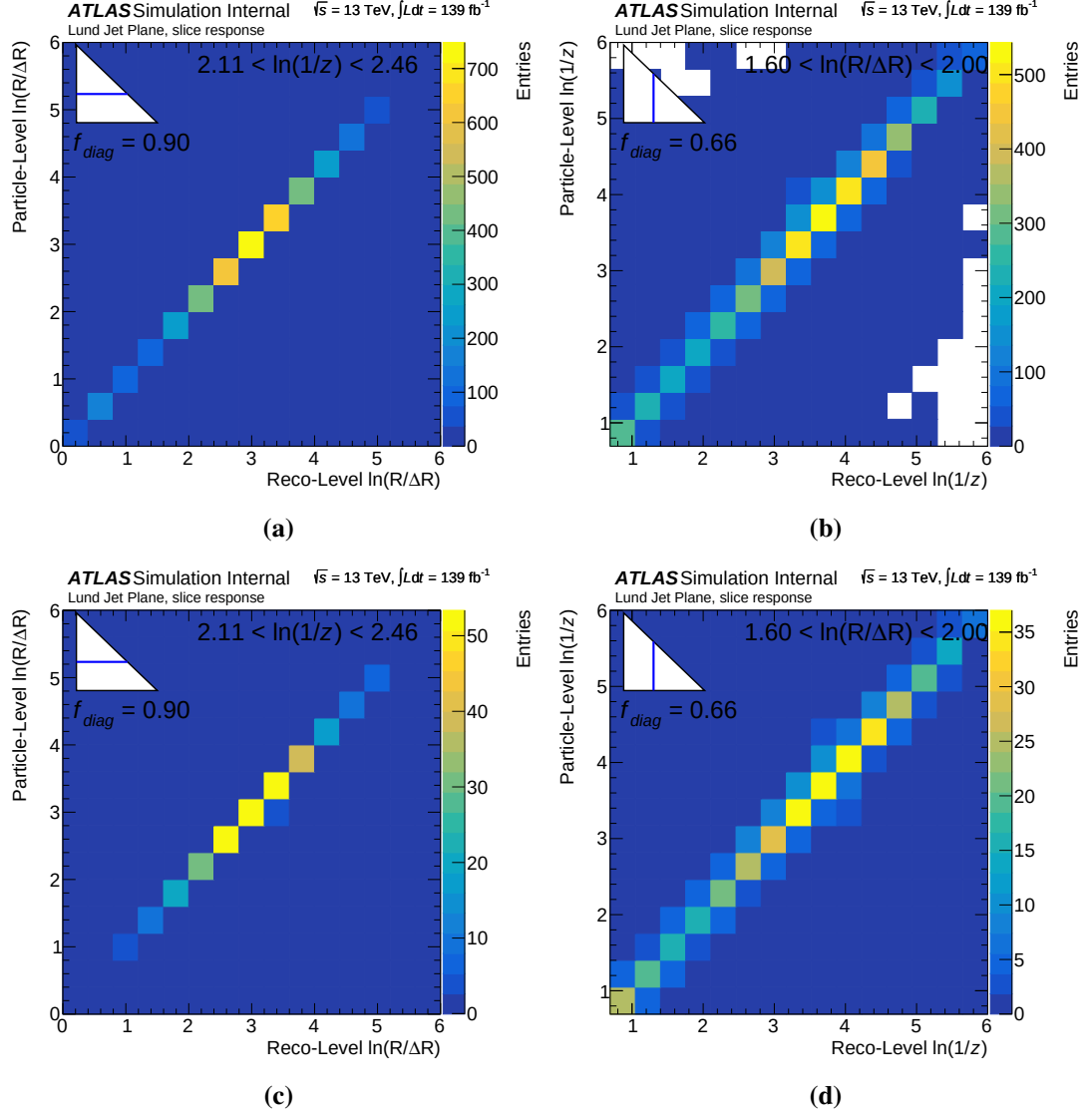


Figure 4.11: The response matrix for a single Lund Jet Plane slice along $\ln(R/\Delta R)$ (a,c) and $\ln(1/z)$ (b,d). The response is shown separately for top jets (a,b) and W jets (c,d).

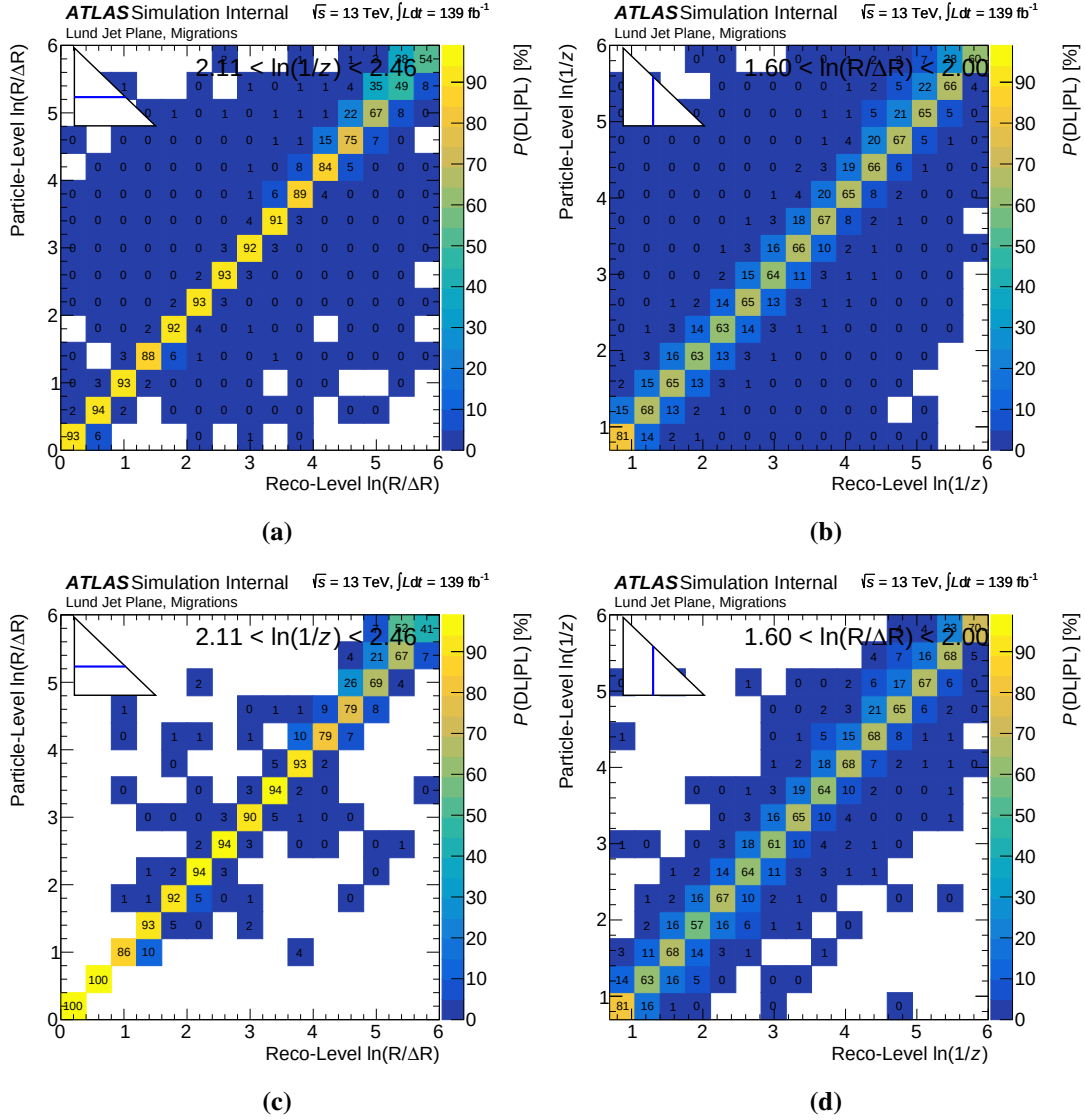


Figure 4.12: Migration matrices for a Lund Jet Plane slice along the $\ln(R/\Delta R)$ (a,c) and $\ln(1/z)$ (b,d) axis for top (a,b) and W jets (c,d). The conditional probability for particle level bins (y-axis) to be reconstructed in the detector level bins (x-axis) is used to quantify the migrations.

4.4.2 Acceptance and inefficiency factors

The acceptance and inefficiency factors f and c , as described in section 3.1.4 and defined in equation (3.4), are shown in Figure 4.13.

The correction factors to the response are significantly lower than the reconstruction efficiency, which from Table 4.4 is around 60%. The additional losses are incurred due to the matching requirement between detector and particle level emissions in the construction of the response matrix. A notable feature of the acceptance and inefficiency factors across the LJP are the low values observed at large angles, approximately $\ln(R/\Delta R) \lesssim 0.4$ or $\Delta R \lesssim 0.67$ for the top jets or $\ln(R/\Delta R) \lesssim 0.8$ or $\Delta R \lesssim 0.449$ for the W jets. This is likely a consequence of the definition of the LJP observable, which involves re-clustering the constituents of anti- k_T jets using the Cambridge-Aachen algorithm. Emissions at wide angles that would ordinarily be captured by the CA algorithm fall outside the $R = 1.0$ cone of the anti- k_T jet. This has some important implications for the theoretical description of the LJP, as has been discussed in section 3.5.

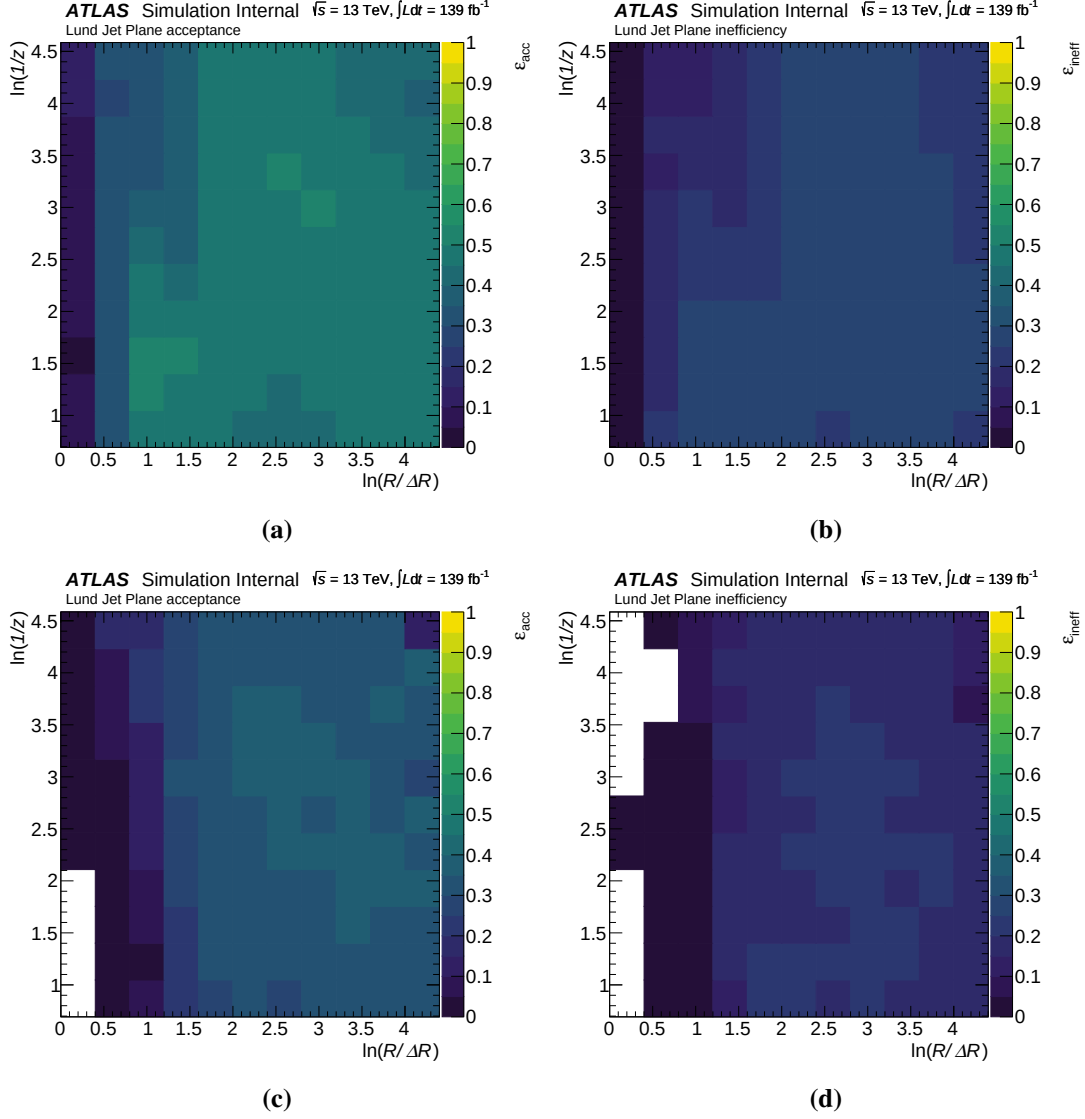


Figure 4.13: The Acceptance (a) and Inefficiency (b) correction factors across the Lund Jet Plane for the top (a) and (b) W jets.

4.4.3 Comparison of emission matching procedures

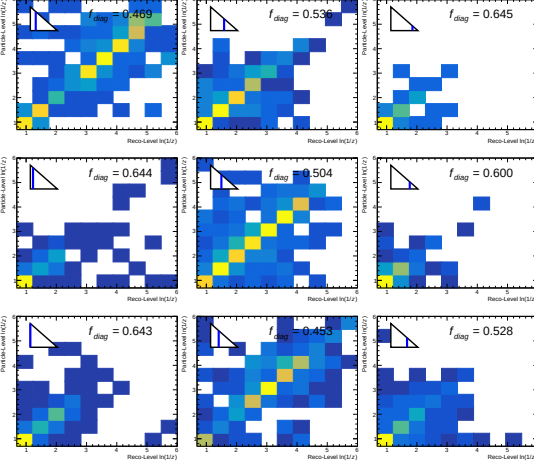
The simplest method is a *sequential matching*, where detector- and particle-level LJP entries are matched according to the order of the emissions. Due to the angular-ordered nature of the Cambridge-Aachen algorithm, this can also be considered an ordering of emissions by ΔR . A more sophisticated method is a *geometric matching* procedure where each detector-level emission is matched to the closest particle-level emission in $\eta - \phi_{var}$. The final matching procedure which is considered is a *unique geometric matching*, where the total distance between pairs of detector- and particle-level emissions is minimised. The minimisation is performed using the combinatorial scheme commonly referred to as the Hungarian Algorithm [203].

Figure 4.14 shows the response matrix for a single slice through the LJP for a collection of different matching procedures. Figure 4.15 shows the fraction of entries on the diagonal of the response matrix for a collection of different matching procedures. These exploratory studies were performed on a combined subsample of 10,000 top and W jets. The sequential matching procedure yields more off-diagonal elements in the response matrix. Due to the presence of pileup and other additional soft radiation within the jet, there are often additional emissions at the detector level which are not present at particle-level. This produces a mismatch between the emission orders unless the number of reconstructed and particle-level emissions is exactly the same. A higher degree of diagonality in the response is observed for the two geometric matching procedure. The unique geometric matching yields slightly fewer off-diagonal entries than the simple geometric matching.

To construct the response matrix for the measurement of the Lund Jet Plane, the simple geometric matching procedure is preferred over the more computationally complex unique geometric matching. Unreconstructed emissions at the particle level and misreconstructed detector level emissions that do not have a close match are excluded by requiring a cut-off value on the $\eta - \phi_{var}$ distance $\Delta R_{RL-PL} < 0.1$ for the geometric matchings. This cut-off value is the same as was used in the LJP measurement in dijets [115]. It was determined in the dijets measurement² that tightening this cut has a very small effect

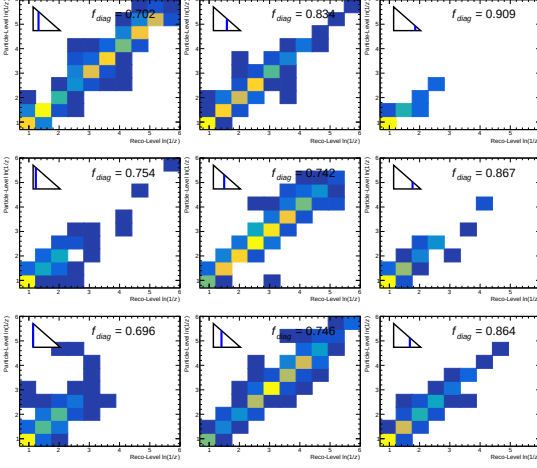
(< 0.5%) on the unfolded LJP distribution, as unmatched emissions are correct for in the unfolding by the acceptance and inefficiency factors.

ATLAS Simulation Internal
LJP $\ln(R/\Delta R)$ slices Response Matrices, sequential matching



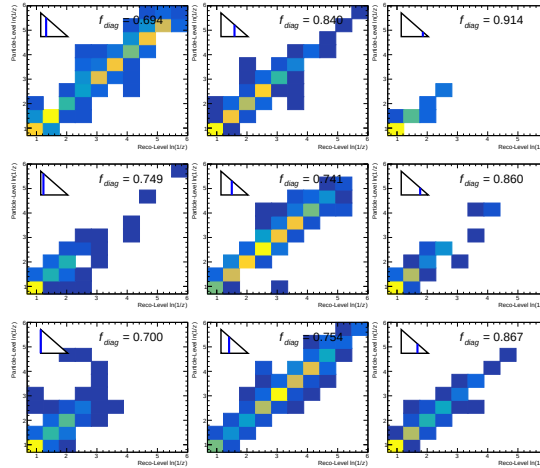
(a)

ATLAS Simulation Internal
LJP $\ln(R/\Delta R)$ slices Response Matrices, $\delta R(\eta, \phi)$ closest match



(b)

ATLAS Simulation Internal
LJP $\ln(R/\Delta R)$ slices Response Matrices, $\delta R(\eta, \phi)$ unique match



(c)

Figure 4.14: Response matrices for $\ln(R/\Delta R)$ slices of the LJP for the sequential (a), simple geometric (b) and unique geometric (c) matching.

²Further details on this study can be found in the ATLAS internal note accompanying [115].

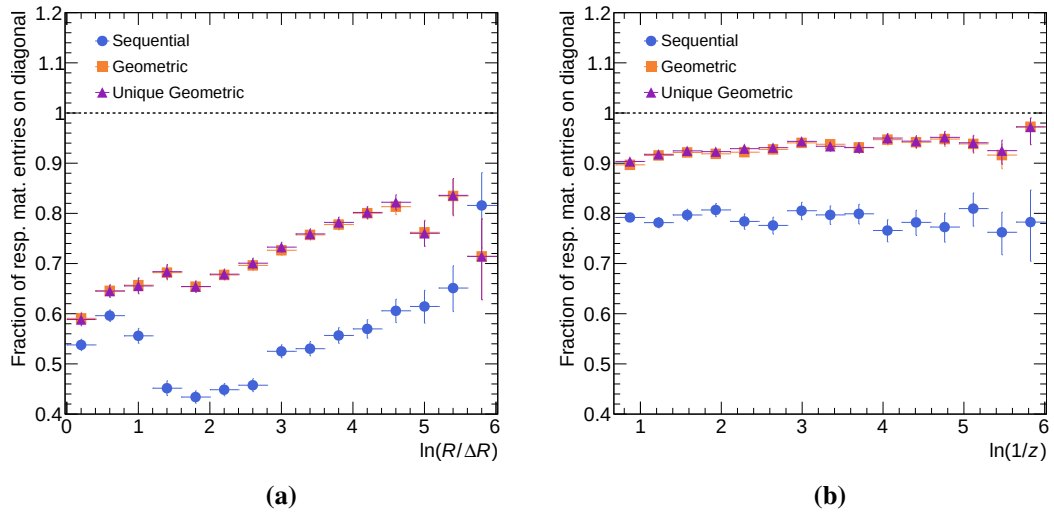


Figure 4.15: Fraction of entries on the diagonal of the response matrix for horizontal and vertical slices through the Lund Jet Plane at different values of $\ln(R/\Delta R)$ (a) and $\ln(1/z)$. Statistical uncertainties are shown as error bars.

4.4.4 Technical closure test

The technical closure of the iterative Bayesian unfolding is tested by applying the unfolding to detector-level $t\bar{t}$ Monte Carlo distributions. Figure 4.16 shows both the unfolded detector-level and particle-level Lund Jet Plane. The particle-level distributions are recovered by applying the unfolding to the detector level Monte Carlo, thus showing that the iterative Bayesian unfolding procedure is implemented correctly and working as expected.

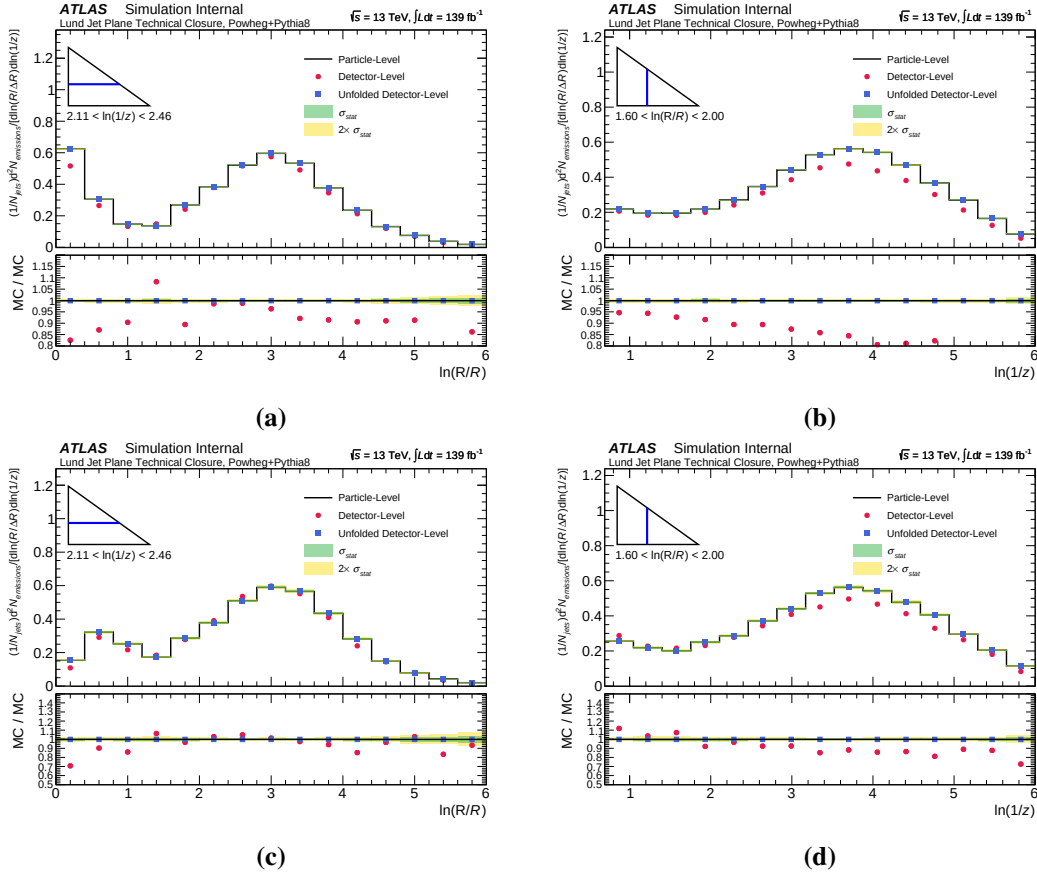


Figure 4.16: Unfolded detector-level Monte-Carlo LJP slices for the top (a,b) and W (c,d) selections. Unmodified detector-level LJPs are also shown. All distributions are normalised.

4.4.5 Data-driven non-closure uncertainty

The uncertainty associated with the iterative Bayesian unfolding is computed according to the ATLAS standard model group prescriptions [204]. The method employed is a *data-driven closure test* [205, 206] which estimates the uncertainty associated with the choice of the prior distributions used in the iterative Bayesian unfolding.

The calculation of the data-driven non-closure starts with the detector and particle level histograms for events that pass both the detector and particle level selections. These are obtained by summing the rows and columns of the response matrix, such that the particle-level histogram is $t_i(DL \wedge PL) = \sum_j \mathbf{R}_{ij}$ and the detector level histogram is $r_i(DL \wedge PL) = \sum_j \mathbf{R}_{ji}$. The detector and particle level histograms passing both selections are related through a normalised response matrix

$$r_i(DL \wedge PL) = \tilde{\mathbf{R}}_{ij} t_j(DL \wedge PL), \quad (4.4)$$

where

$$\tilde{\mathbf{R}}_{ij} = \mathbf{R}_{ij} / \sum_{i'} \mathbf{R}_{i'j}. \quad (4.5)$$

The measured (detector level) data d_i is reweighted to distribution of detector level events passing both detector and particle level selections $r_i(DL \wedge PL)$ using the formula

$$\tilde{d}_i = \varepsilon_i d_j = \frac{r_j(DL \wedge PL)}{r_j(DL)} d_j, \quad (4.6)$$

where $r_j(DL)$ is the histogram for all events passing the detector level selections in Monte Carlo. The reweighted data $\tilde{d}_i$ is then used to estimate a prior from which the uncertainty on the Bayesian unfolding procedure can be estimated. We estimate this prior from the

following re-weighting of the particle-level distribution $t_i(DL \wedge PL)$:

$$\tilde{t}_i = w_i t_i = \frac{\tilde{d}_i}{r_i(DL \wedge PL)} t_i(DL \wedge PL) \quad (4.7)$$

Having defined the quantities above, the non-closure systematic uncertainty is given by

$$\sigma_{nc} = \text{IBU}(\tilde{R}_{ij} \tilde{t}_{ij}) - \tilde{t}_i. \quad (4.8)$$

where IBU denotes the iterative Bayesian unfolding procedure performed with the same response matrix (R_{ij}) as the measurement. By repeatedly updating the prior $\tilde{t}_i$ one obtains the data-driven non-closure uncertainty associated with the number of iterations used. As the prior gets updated, the bias introduced by the choice of $\tilde{t}$ is expected to shrink with each iteration such that $\sigma_{nc} \rightarrow 0$ as $N_{IBU} \rightarrow \infty$.

The size of the data-driven non-closure uncertainty across the Lund Jet Plane is shown in Figure 4.17. The size of the uncertainty is $< 2\%$ across most of the LJP, except in the upper left-hand corner where there are very few events. The response matrix is only very sparsely populated in for these bins, hence there is a larger uncertainty associated with the unfolding. The reason for the absence of some emissions in this region is most likely an effect of the trimming algorithm, which is applied to the jet constituents at both the detector and particle level, but does not necessarily affect all the ghost-associated tracks. The trimming algorithm (see section 1.5.1 for a precise definition) re-clusters the jet constituents with a subjet radius parameter $R_{trim} = 0.2$ and then discards subjets with $p_T^{subjet} / p_T^{jet} \equiv f_{trim} < 0.05$. One can therefore expect that the region of the LJP that is most affected by the trimming algorithm is approximately bounded by $\ln(R/R_{trim}) \lesssim 1.609$, $\ln(1/f_{trim}) \lesssim 2.229$. The region where the non-closure is largest is $\ln(R/\Delta R) \lesssim 1.2$, $\ln(1/z) \lesssim 3.5$, within the region affected by the trimming. The non-closure observed in this region is of consequence for interpreting the unfolded results in section 4.7.

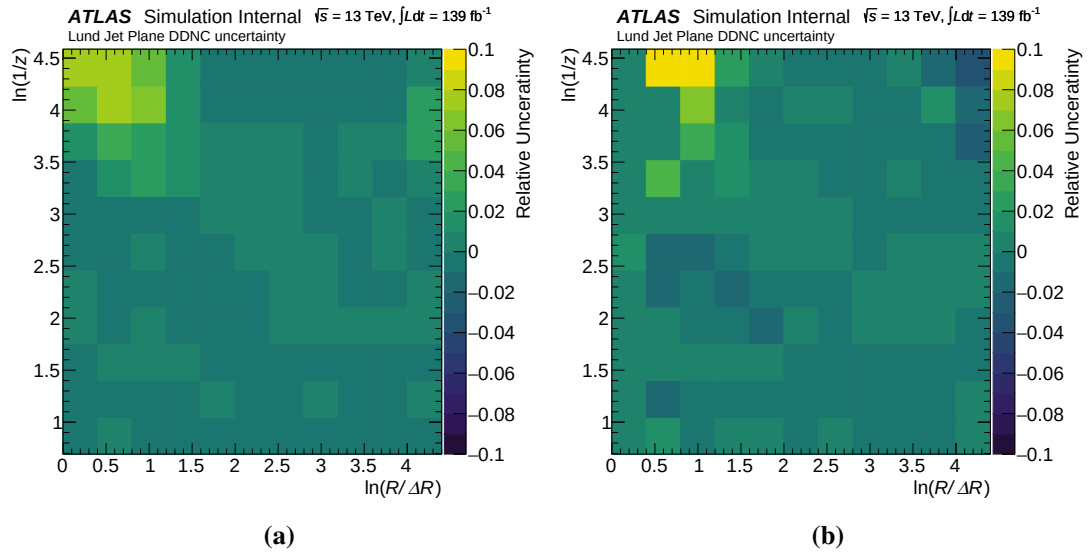
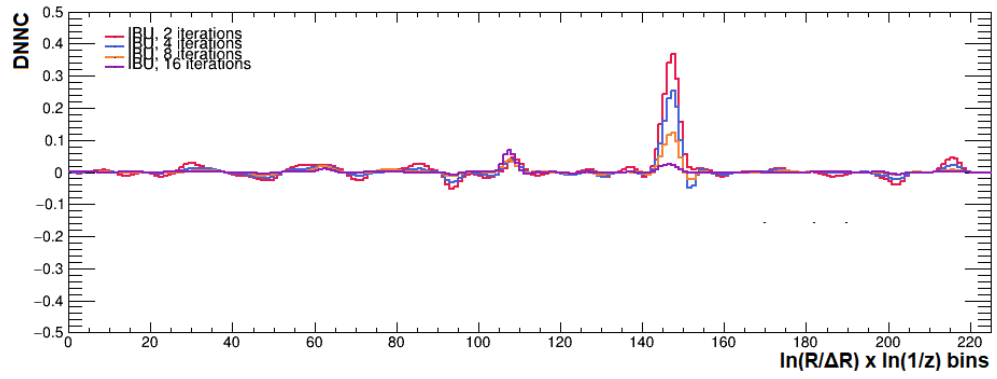


Figure 4.17: The Data-Driven non-closure uncertainty across the Lund Jet Plane for the "top" (a) and "W" (b) selections.

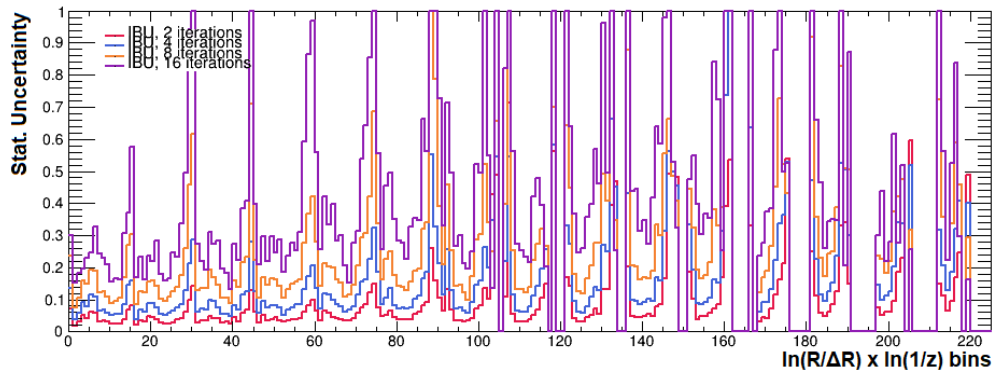
4.4.6 Choice of the number of unfolding iterations

The number of iterations to use in the Bayesian unfolding procedure is chosen such that the data-driven non-closure uncertainty and other sources of uncertainty are minimised. The Data-Driven non-closure uncertainty decreases with the number of iterations, while the unfolded statistical and systematic uncertainties increase. This relationship can be seen in figure 4.18. The DDNC uncertainty across the Lund Jet Plane is evaluated as described in 4.4.5 and is smoothed in 2 dimensions by applying Gaussian kernel smoothing with a 3×3 kernel size. The Statistical uncertainty on the unfolded data is evaluated using the Bootstraps method, as described in Section 4.6.1. The size of DNNC uncertainty is quite small, less than 1% in most bins across the bulk of the plane, even for a small number of unfolding iterations. The statistical uncertainty, on the other hand, grows rapidly with each unfolding step. In order to keep the statistical uncertainty below approximately 10% in the more densely populated regions of the Lund Jet Plane, the number of iterations must be 4 or less.

The sharp peak in the DDNC shown in figure 4.18 corresponds to the region affected by trimming effects, as discussed in section 4.4.5. Also visible in this plot are some periodic changes in the non-closure with a period of roughly 15 bins. These effects are likely to be artefacts of the unfolding of the flattened 2-d LJP distributions which show discontinuities between each row in the LJP.



(a)



(b)

Figure 4.18: Data-driven non-closure (a) and statistical uncertainty (b) for IBU with different iteration numbers.

4.5 Unfolded results

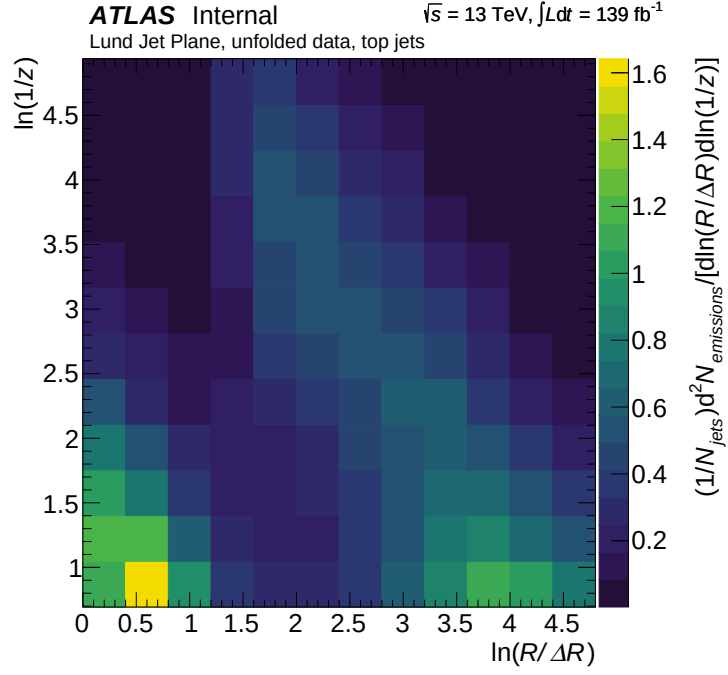
The measurement is unfolded to the Particle Level using the iterative Bayesian unfolding process as was done for the Lund Jet Plane measurement in dijets. The unfolded LJP is normalised by the number of jets to give the Lund Jet Plane density

$$\rho(\Delta R, z) = \frac{1}{N_{jets}} \frac{d^2 N_{emissions}}{d \ln(1/z) d \ln(R/\Delta R)}, \quad (4.9)$$

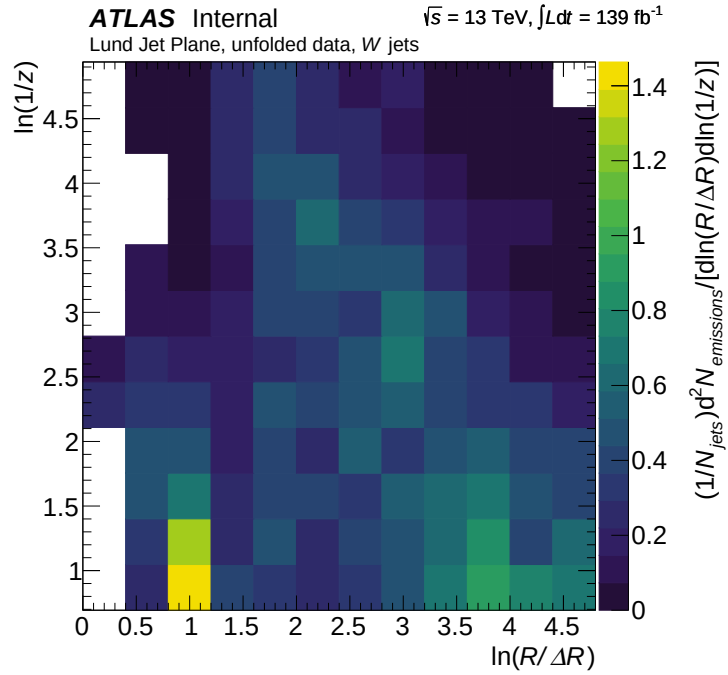
where the total number of jets passing the selections, N_{jets} , is unfolded simultaneously with the LJP distributions³. The ratio $N_{emissions}/N_{jets}$, to which the LJP has been normalised, can be found in appendix B.3 for the unfolded data and particle-level MC with associated uncertainties. The difference in this ratio between the unfolded data and the MC is quite small, $< 1\%$ for the top jets and $< 2\%$ for the W jets.

The unfolded measurement of the Lund Jet Plane in semi-leptonic $t\bar{t}$ events is shown in figure 4.19 for both the top jet and W jet selections. The distinguishing feature of the LJP for W and top jets is the high density of emissions around $\ln(1/z) < 1.5$ and $\ln(R/\Delta R) < 1.5$. Due to the hard and wide-angle nature of these emissions, they are easily identified with the decays $t \rightarrow qq\bar{b}$ and $W \rightarrow qq$ that also give rise to a three- or two-pronged energy distribution in $\eta - \phi$ space. For the top jets, there is a broader distribution of emissions in this area, which is perhaps the result of the presence of a second hard splitting in the decay of the top. However, since the LJP is constructed by only following the hardest branch in the CA-declustering we cannot expect that this splitting is always visible in the LJP. For the W jet selection, there are a number of empty bins, e.g. at $\ln(R/\Delta R) < 0.5$, which is due to the absence of emissions in the corresponding regions at the detector level. The rapidly decreasing emission density at large ΔR for W jets might be explained by the kinematics of the jet constituents. At $p_T > 350 \text{ GeV}$ the decay products of the W are highly boosted along the jet axis, which leads to a lower density of jet constituents at wide angles.

³For an observable with a single bin, like the total number of jets N_{jets} , unfolding is equivalent to scaling the observable in data by the factor $f \times (1/c) = N_{jets}^{particle \text{ level}} / N_{jets}^{detector \text{ level}}$. This follows from the definition of the unfolding matrix in equation (3.3), which is unity for $n_{bins} = 1$.



(a)



(b)

Figure 4.19: The Unfolded Lund Jet Plane in Run 2 Data for the top jet (a) and W jet (b) selections.

4.6 Uncertainties

4.6.1 Statistical uncertainty

The statistical uncertainty on data and Monte Carlo samples are estimated using the *Bootstraps* method [207] implemented in the ATLAS `BootstrapGenerator` package [208–211]. The *Bootstraps* method is a versatile tool for computing statistical. For each event, $N_{bootstrap}$ pseudo-events, or *replicas*, are sampled with a probability following a Poisson distribution. Statistical uncertainties can be estimated for each observable by calculating a statistical error over the $N_{bootstrap}$ random samples. The bootstrap replicas can also be used to calculate other statistical properties, such as the covariance (see appendix B.2), between the bins of one measured distribution or between any two measured observables.

The *Bootstraps* method makes it very easy to propagate statistical uncertainties for unfolded distributions. Replica distributions are generated for the data and MC histograms, as well as the response matrix. For each replica of the data distributions, a replica of the background Monte Carlo distributions is subtracted and the result is unfolded using the a replica of the response matrix. The statistical uncertainty on the unfolded result can then be evaluated from the standard deviation over the $N_{bootstrap}$ unfolded replicas. This measurement uses $N_{bootstrap} = 200$ and the standard deviation over bootstrap samples to estimate the statistical uncertainties.

The statistical uncertainty across the LJP is shown for top jets in figure 4.20 and for W jets in figure 4.21. In the middle of the plane, where there is ample statistics, the statistical uncertainty is around 3 to 5 times greater in the detector level data than in the detector level Monte Carlo. The statistical uncertainty in the unfolded data is between 10 and 30 times greater than for the particle level Monte Carlo. Since we find that the total systematic uncertainty (see section 4.6.6) is greater than the statistical uncertainty on the unfolded measurement, the contributions from MC statistical uncertainties to the total uncertainty are so small as to be negligible.

Statistical covariance and correlation matrices can be found in appendix B.2. These have been evaluated from the bootstraps analogously to the statistical uncertainty. The bin-by-bin statistical correlation across the unfolded LJP for top and W jets is < 0.2

in nearly all off-diagonal elements of the correlation matrix, except at small angles ($\ln(R/\Delta R) > 5.0$) where it increases to around 0.5. This is possibly due to the difficulty of resolving these very small angles.

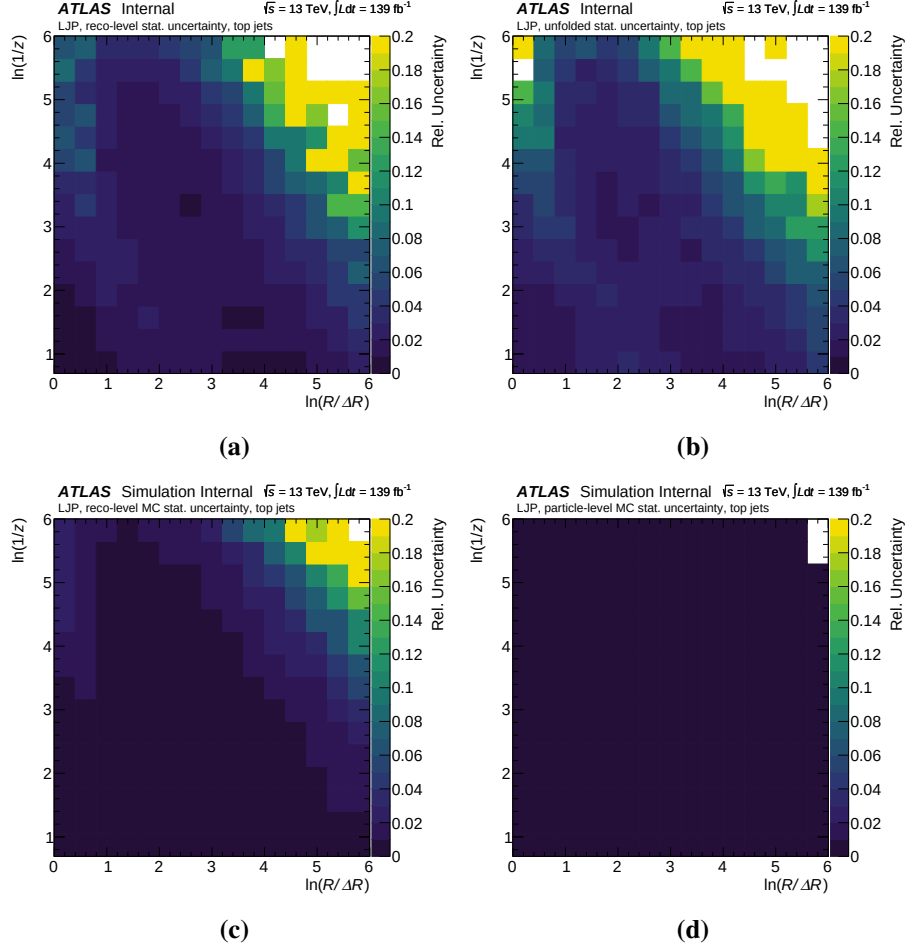


Figure 4.20: Statistical uncertainty across the Lund Jet Plane for the detector level data (a), unfolded data (b), detector level nominal MC (c) and particle level MC (d) in the top-jet selection.

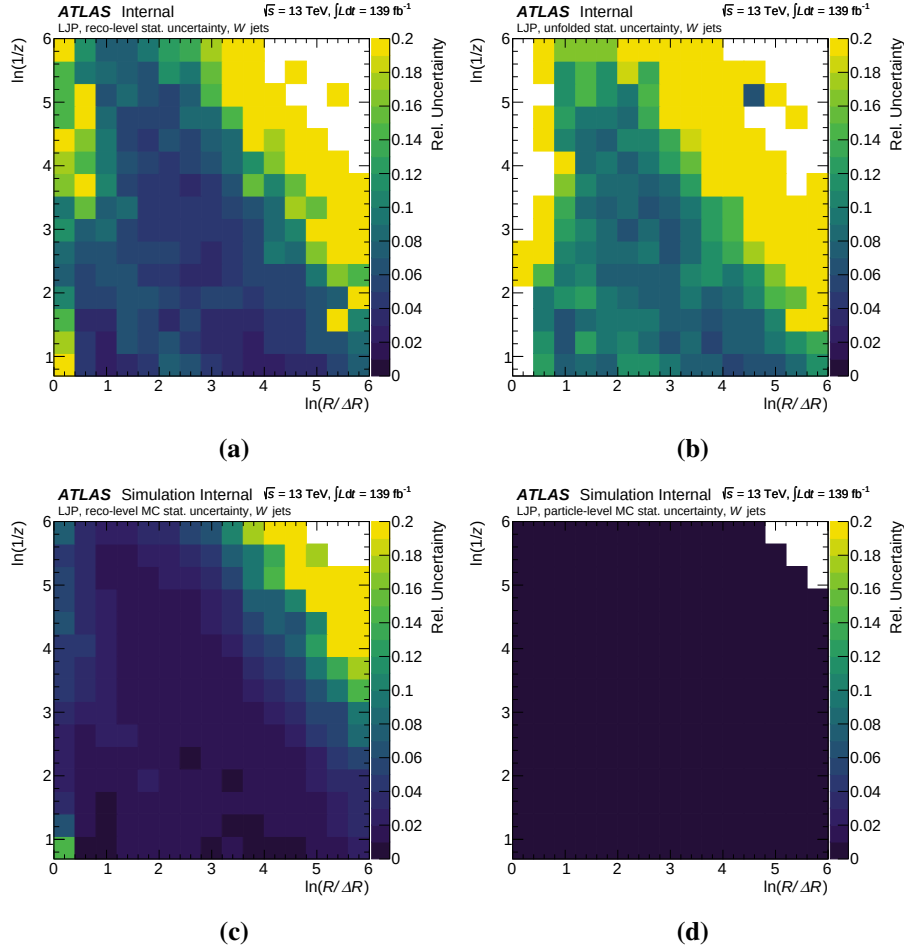


Figure 4.21: Statistical uncertainty across the Lund Jet Plane for the detector level data (a), unfolded data (b), detector level nominal MC (c) and particle level MC (d) in the W -jet selection.

4.6.2 Systematic uncertainties

Systematic uncertainties are considered from a number of different sources relating to the modelling of the Monte Carlo samples and the reconstruction of the different objects considered in the analysis. The following sections provide a description of each source of systematic uncertainty.

4.6.3 $t\bar{t}$ modelling systematics

Systematic uncertainties associated with the modelling of $t\bar{t}$ Monte Carlo simulations are estimated from alternative simulated events obtained by either using a different MC generator or by varying the parameters on the nominal Powheg+Pythia 8 simulation. A detailed description of the alternative Monte Carlo samples is given in section 4.2.6.2. All events from the alternative Monte Carlo event generators are passed through the same Geant 4 detector simulation as the nominal events.

To obtain the modelling systematic uncertainties on the unfolded Lund Jet Plane distributions, the detector-level distributions for each alternative sample are unfolded using the detector response derived from the nominal $t\bar{t}$ events. The difference between the unfolded result and the particle-level distributions is taken to be the systematic uncertainty.

Matrix Element modelling uncertainty is evaluated using the aMC@NLO+Pythia 8 alternative Monte Carlo sample. This sample uses the MadGraph5_aMC@NLO matrix element generator as an alternative to Powheg.

Parton Showering and Hadronisation modelling uncertainty is evaluated using the Powheg+Herwig 7 alternative Monte Carlo sample. The sample uses the Herwig 7 parton shower generator as an alternative to Pythia 8.

Variation on the h_{damp} parameter in the Powheg+Pythia 8 simulations is captured by one modelling uncertainty. The uncertainty is estimated from an alternative Monte Carlo sample in which the value of h_{damp} is doubled from $1.5m_t$ to $3.0m_t$.

FSR Modelling uncertainty is estimated by varying the renormalisation scale that regulates the amount of final state radiation in Pythia 8. A "down" and "up" systematic variation is produced by varying the scale by a factor of 0.5 and 2.0.

ISR Modelling uncertainty is estimated by comparing the nominal $t\bar{t}$ sample with two additional samples [190], obtained reweighting the nominal one. The reweighting is performed by varying the factorisation and renormalisation scales by a factor of 0.5 or 2 while simultaneously changing the Var3c variation from the A14 tune in the parton shower generator. As for the FSR, one obtains a "down" and an "up" systematic variation.

4.6.4 Object-reconstruction systematics

Lepton systematics are estimated from a set of data/MC scale factors that have been derived due to the differences in reconstruction efficiency between data and Monte Carlo. A standard set of recommended scale factors exists for the ATLAS Run 2 dataset for electrons [159, 161, 160] and muons [125] that have been derived by the ATLAS electron and muon subgroups. The scale factors have been derived using leptonic Z , W and J/ψ decays for which the total cross-section and hence the expected number of events is well known. There are also systematic uncertainties relating to the reconstruction of the muon kinematics. These include muon momentum scale and resolution uncertainties as well as track resolution in the Inner Detector and the Muon Spectrometer.

Small- R jet systematics for jet energy scale and resolution are derived according to the standard prescriptions of the ATLAS $\text{Jet}/E_T^{\text{miss}}$ group [135, 127]. Contributions to this uncertainty come from MC mismodelling, event topology, pileup and statistics. For the JES, there is also a number of uncertainties associated with the η inter-calibration. The number of systematic uncertainties has been reduced using an eigenvector decomposition of the correlation matrix of each systematic with jet p_T . The `CategoryReduction` reduced set of systematic variations is used. This comprises 30 systematics for the jet energy scale (JES) and 8 for the jet energy resolution (JER). Lastly, there is also a systematic uncertainty relating to the JVT cut which has been derived from data/MC scale factors.

B-tagging systematic uncertainties are derived from the variations to the data/MC B-tagging scale factors [128, 129]. The number of systematic variations is reduced using the same statistical method used for the small- R JES uncertainty. There are 19 b-tagging systematics in total: 9 relating to the b -jet efficiency scale factor, 4 for the c -jet and light

jet efficiencies respectively and two relating to the extrapolation of the b -tagging scale factors to high- p_T .

Large- R jet systematics are derived from the in situ calibrations of the JES, JMS, JER and JMR for the $R = 1.0$ anti- k_T topocluster jets [134, 212]. The jet energy scale (JES) systematic uncertainty contains contributions due to MC modelling, event selections for deriving the in-situ response, jet flavour response and statistics. For this analysis there is also a topology component which captures differences in energy response between top-quark initiated jets and the QCD jets from which the calibrations are derived. The jet energy resolution (JER) systematic uncertainty also contains contributions from modelling and statistics, as well as a non-closure component derived from the difference between multiple methods used to estimate the JER.

There are two estimates of the jet mass scale (JMS) uncertainty: one from the R_{trk} method [133] and one from the forward-folding in-situ calibrations [134]. For both estimates, there are components relating to MC modelling, in-situ event selection and statistics. There are also three JMS systematics that account for the finite binning used to derive the mass response. Just as for the JES, there is also a systematic uncertainty associated with the top jet topology. The $R = 1.0$ jet mass resolution (JMR) systematic uncertainties are associated with MC modelling, statistics and the choice of $R = 0.4$ and $R = 1.0$ jet calibrations used to derive the JMR. The total JMR systematic uncertainty is set to a conservative estimate of 20% in events that fall outside the phase space of the JMR measurement.

Track reconstruction systematics are evaluated for four main sources of uncertainty [162]. The first systematic is due to the difference in reconstruction efficiency between data and simulation. The second is the uncertainty in the measurement of the track impact parameter. The third is uncertainty due to fake reconstructed tracks which is obtained by comparing the number of reconstructed tracks as a function of the number of bunch crossings μ in data and Monte Carlo. Lastly, there is a bias uncertainty which affects the reconstructed values of d_0, z_0 and Q/p . The tracking uncertainties are propagated for all tracks before they are ghost-associated to the large- R jets. Due to the small emission

angles that are resolved in the LJP, it is expected that the tracking systematics will have an impact the measured LJP where the ΔR coordinate approaches the tracking resolution.

Missing E_T systematics are propagated from the energy scale and resolution uncertainties of leptons and jets. Since the total E_T^{miss} is also affected by the soft term, there are additional energy scale and resolution systematics related to this term.

4.6.5 Background modelling systematics

One or more systematic uncertainties are associated with the modelling of the background Monte Carlo samples. As the signal purity of the semi-leptonic $t\bar{t}$ decay channels is high and the background is mostly dominated by single top decays, the less dominant backgrounds have only one associated modelling uncertainty. Due to the very small contributions from the subdominant backgrounds, the uncertainty estimates on their total cross-section is quite conservative. The uncertainties on the background samples follow the standard top group prescriptions and are identical to those used in other semi-leptonic $t\bar{t}$ cross-section measurements using the entire Run 2 dataset, eg. refs. [213, 199].

The effect of the background systematic uncertainties on the unfolded measurement is estimated by subtracting the modified backgrounds from the Data in turn and unfolding the result. The final uncertainty is the difference to the unfolded background-subtracted Data.

Single Top: There are four sources of background uncertainty associated with the most dominant background process:

1. Uncertainty in the NNLO total cross-section estimate used to normalise the sample.
Varies between the different decay channels of the single top process.
2. ISR uncertainty. The uncertainty in the modelling of ISR in the single top sample is evaluated in Powheg+Pythia 8 simulation, analogously to the method used to estimate this uncertainty in $t\bar{t}$ (see Section 4.6.3).
3. FSR uncertainty. The uncertainty in the modelling of FSR in the single top sample is evaluated by varying the renormalisation scale, which is the same method used to estimate this uncertainty in $t\bar{t}$ (see Section 4.6.3).

4. Uncertainty associated with the DR scheme for removing interference between $t\bar{t}$ and tW production processes. This is calculated using an alternative sample where the interference is removed using the diagram subtraction (DS) scheme is used instead.

W+jets: The uncertainty on the cross-section of W +jets samples is estimated by varying the energy scale parameters μ_r and μ_f in the Sherpa simulation. These are varied simultaneously such that there is an up- and down-component associated with this background uncertainty.

ttV: The uncertainty on the total cross-section of processes where a $t\bar{t}$ pair is produced with an associated vector (or Higgs) boson is set to a constant 13.3%.

Z+jets: The uncertainty on the total cross-section of the Z + jets process is set to a constant 50% due to the small number of events that pass the selections of this analysis.

Diboson: The uncertainty on the total cross-section of the diboson process is set to a constant 50%.

Fake leptons: The uncertainty on the amount of miss-identified leptons is conservatively estimated to 50% due to the small contribution of this background to the total predicted cross-section.

4.6.6 Summary of systematic uncertainties

The uncertainties are evaluated on the normalised LJP distributions. The total systematic uncertainty and its main contributions is shown in Figures 4.22 and 4.23. In the central part of the LJP ($\ln(1/z) \lesssim 3.5$, $0.5 \lesssim \ln(R/\Delta R) \lesssim 3.5$), where the statistical uncertainty is small ($< 20\%$), the total uncertainty is around 20 – 30% in the top jet selection and between 40% and 60% in the W jet selection. For both top jets and W jets, the dominant sources of uncertainty are the modelling systematics. As different sources of modelling uncertainty dominate different parts of the plane, there is much variation in the total uncertainty across the LJP. Among the detector uncertainties the most significant contributions are the tracking uncertainties, followed by the large- R jet calibration uncertainties. The tracking uncertainty across the middle of the LJP is between 2% and 7% in the top jets and between 5% and 10% in the W jets. The contributions from tracking

uncertainties are especially large at small splitting angles where the density of tracks is high and the size of the angles approaches the tracking resolution of the ID. In the region $\ln(1/z) > 3.5, \ln(R/\delta R) > 3.5$ the tracking uncertainty grows as large as 15%, even in the top jet selection where systematic uncertainties are generally smaller. There is also a sizeable contribution from the background modelling uncertainties to the total systematic uncertainty which is more significant for the W jets than for the top jets. Uncertainties that are associated with objects like small- R jets and leptons, which are used to select events but do not play a part in the construction of the LJP, mostly affect the total number of events in the sample and therefore do not contribute significantly to the uncertainty on the normalised LJP.

The modelling systematic uncertainties are shown in figures 4.24 and 4.25. The dominant sources of modelling uncertainty vary across the LJP. The largest source of uncertainty in the middle of the plain is the Parton Showering and hadronisation uncertainty derived from the unfolding of the Powheg+Herwig 7 samples. In the region surrounding $\ln(R/\Delta R) \approx 1.5, \ln(1/z) \approx 1.0$ where there is a large contribution from the PS+Hadronisation uncertainty there is also a large contribution from the Matrix Element uncertainty. This uncertainty, which is derived using MadGraph5_aMC@NLO matched with Pythia 8, quantifies the uncertainties introduced by the alternative ME generator and the matching between it and the Pythia 8 parton shower and hadronisation model. Out of the two radiation systematics, the most dominant source of uncertainty is FSR. As can be seen in figures 4.24a and 4.24e, the FSR uncertainty contributes significantly in the lower left-hand corner of the LJP where there are also large contributions from the PS+Hadronisation and Matrix Element uncertainties. However, the peak value of the FSR uncertainty in this region is more shifted towards the hard and collinear part of the LJP.

The systematic uncertainties associated with the simulated background events are shown in Figures 4.26 and 4.27. The contribution to the total uncertainty from the modelling of the background is comparatively small. In the central region ($\ln(1/z) \lesssim 3.5, 0.5 \lesssim \ln(R/\Delta R) \lesssim 3.5$) it is around 1 – 2% for the top jets and around 5% for the W jets. The largest contributions are associated with the Single top sample which contributes the

majority of background events. The largest contributions to the uncertainty on the modelling of the single top background are from FSR and DR-DS, the difference between the diagram subtraction and diagram removal schemes for modelling interference of $t\bar{t}$ and tW events. The size of both uncertainties these uncertainties is roughly the same across the LJP. In the top jet selection, the FSR uncertainty dominates at small values of z , while the DR-DS uncertainty is more dominant at larger values.

Out of the detector systematic uncertainties that are not mentioned above, the most dominant contributions are from the reconstruction of jets. The uncertainties associated only with the small- R jets is 1 – 3% for the top jets and 4 – 5% for the W jets in the bulk of the plane. Large- R JES, JER, JMS and JMR uncertainties, which are correlated with their corresponding small- R jet uncertainties, each contribute 2 – 3% in the top jets and approximately 4 – 5% in the W jets. The total systematic uncertainty associated with B-tagging is around 1% for top jets and between 2% and 5% for W jets. The remaining uncertainties associated with the reconstruction of electrons, muons and MET each contribute around 1% for the top jets and 1 – 3% for the W jets.

A complete breakdown of the different sources of systematic uncertainty, including individual systematic uncertainties for each category of detector systematics, is available in appendix [B.1](#).

Systematic covariance and correlation matrices can be found in appendix [B.2](#). Correlations between bins are evaluated from the sign of the difference between the systematic variation and the nominal Powheg + Pythia 8 sample. Two bins are considered correlated if the signs of the systematic is the same and anti-correlated if they are opposite. As the systematic uncertainties are quite flat, there are large correlations (> 0.6) between adjacent bins in many regions of the LJP. There are also some regions where the systematic uncertainties are anti-correlated. These regions are typically separated by several bins (as can be seen in eg. figure [B.17d](#)).

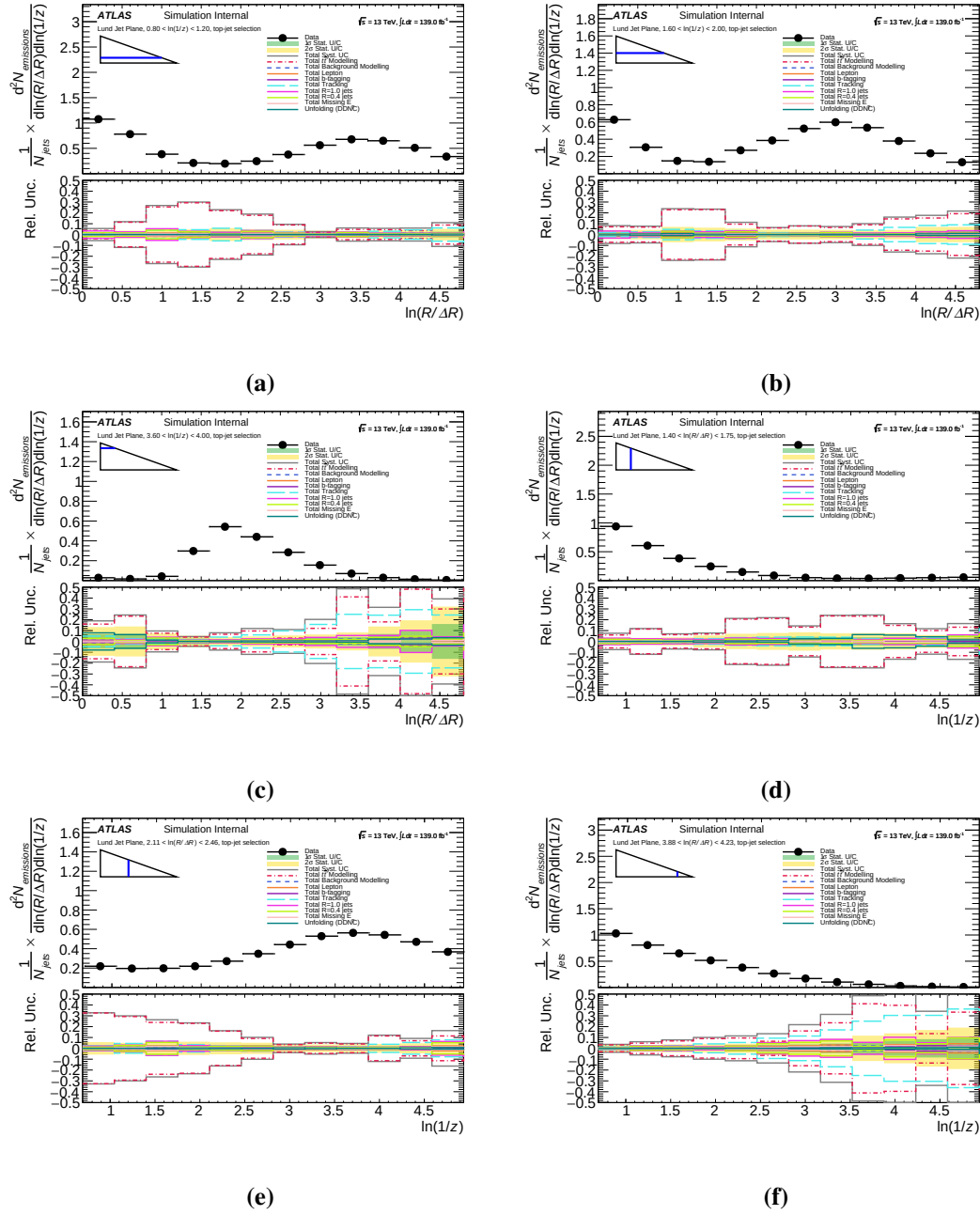


Figure 4.22: Summary of systematic uncertainties in slices through the Lund Jet Plane for the "top" jet selection.

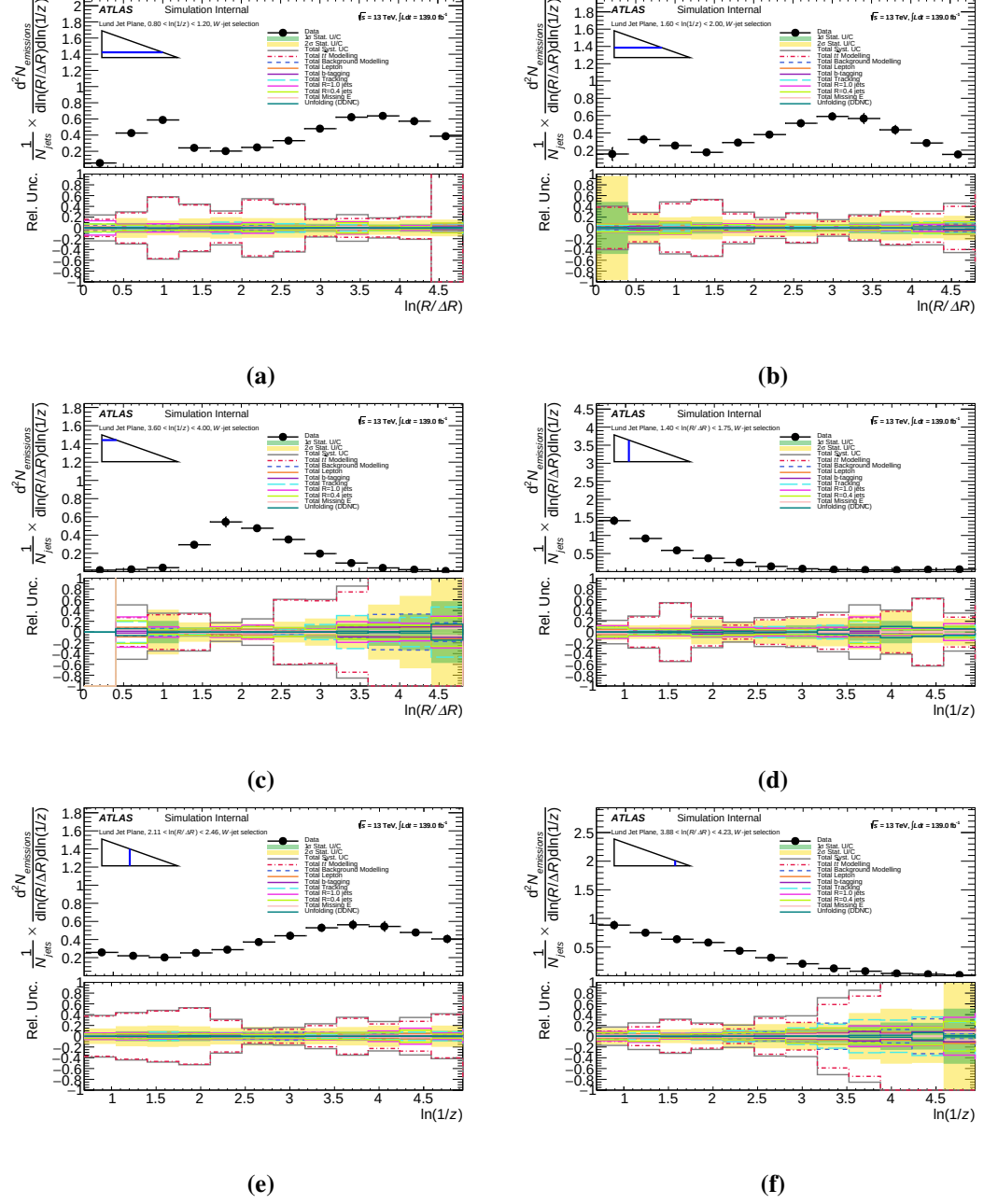


Figure 4.23: Summary of systematic uncertainties in slices through the Lund Jet Plane for the W jet selection.

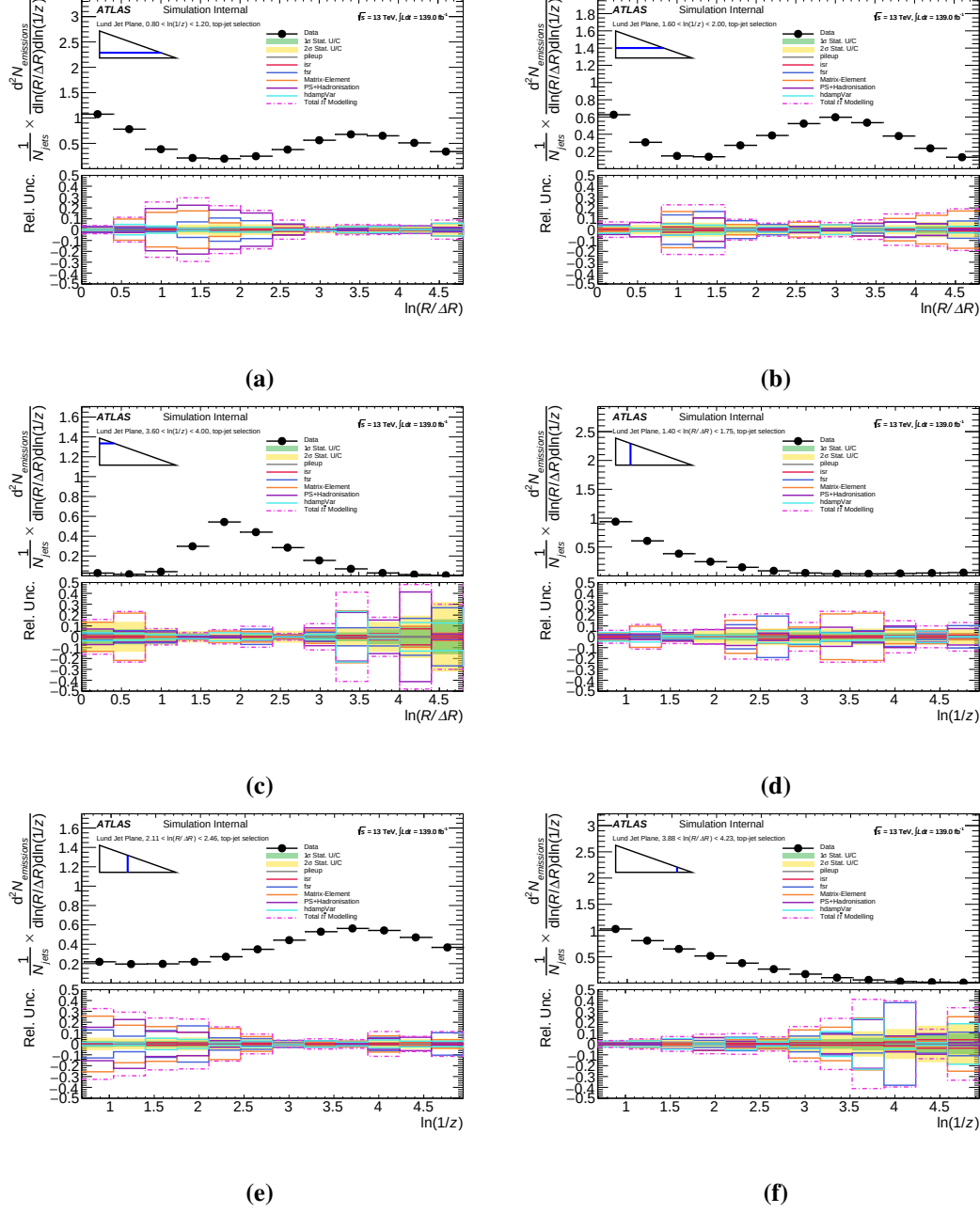


Figure 4.24: Summary of $t\bar{t}$ modelling systematic uncertainties in slices through the Lund Jet Plane for the "top" jet selection.

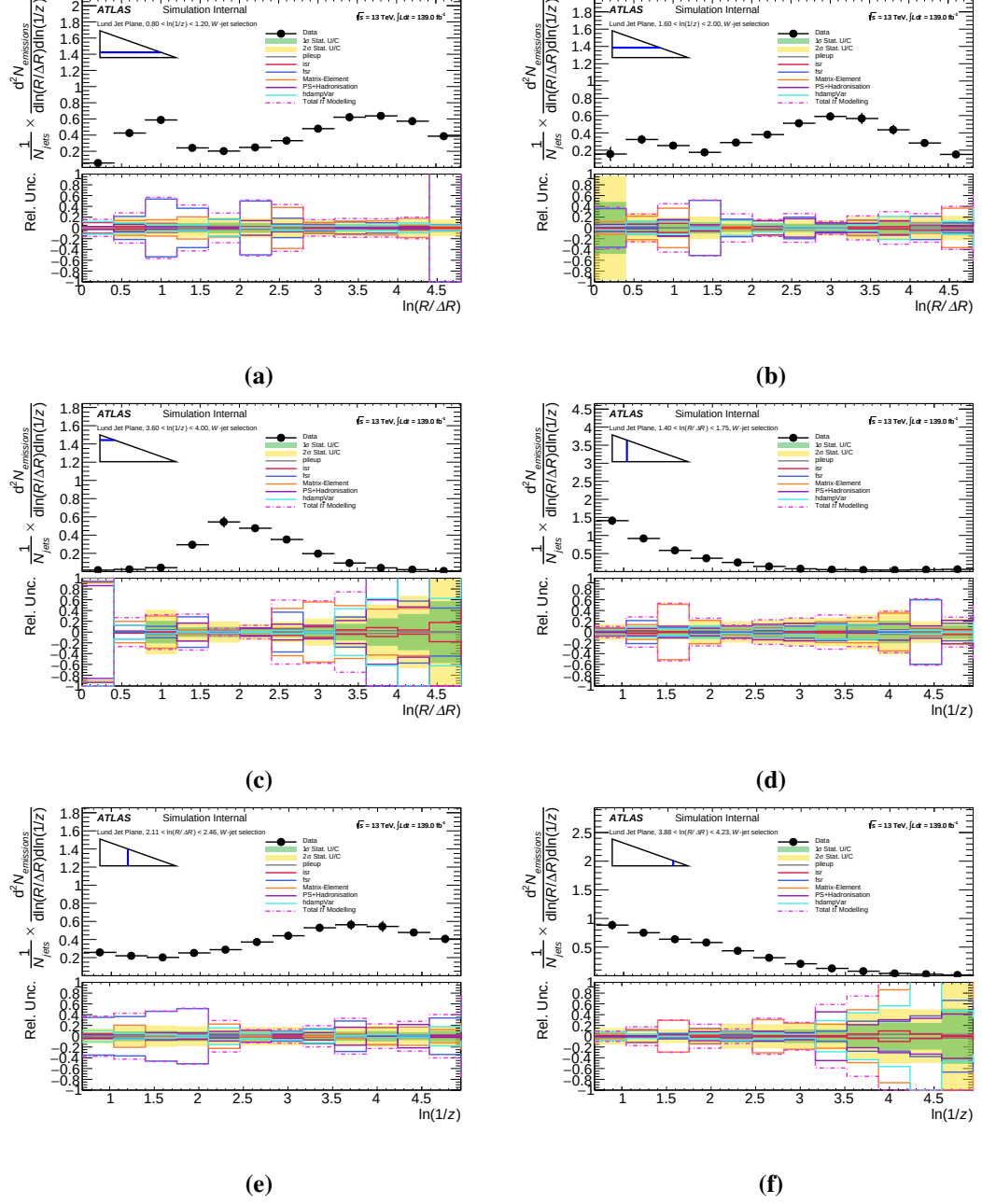


Figure 4.25: Summary of $t\bar{t}$ modelling systematic uncertainties in slices through the Lund Jet Plane for the W jet selection.

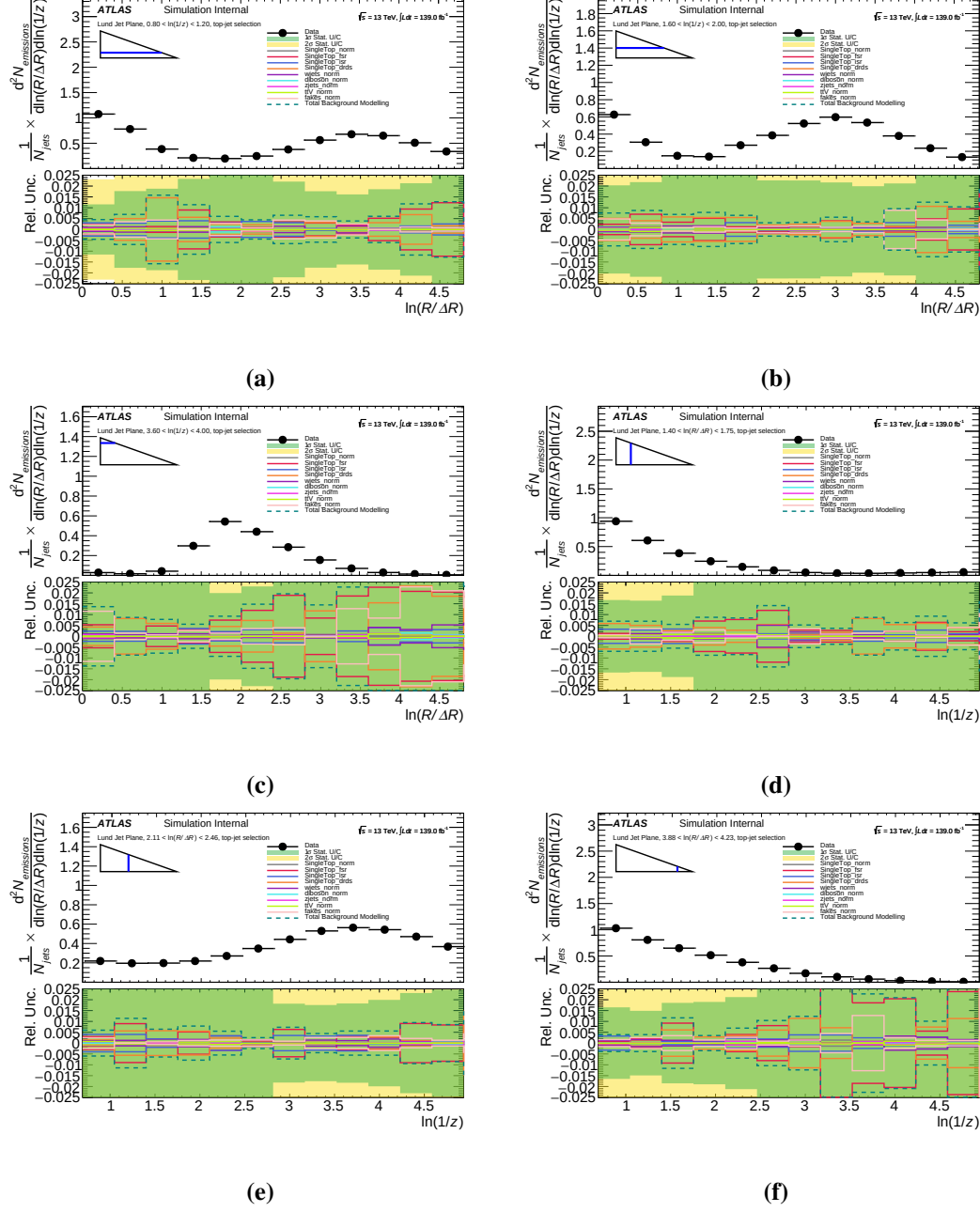


Figure 4.26: Summary of background modelling systematic uncertainties in slices through the Lund Jet Plane for the "top" jet selection.

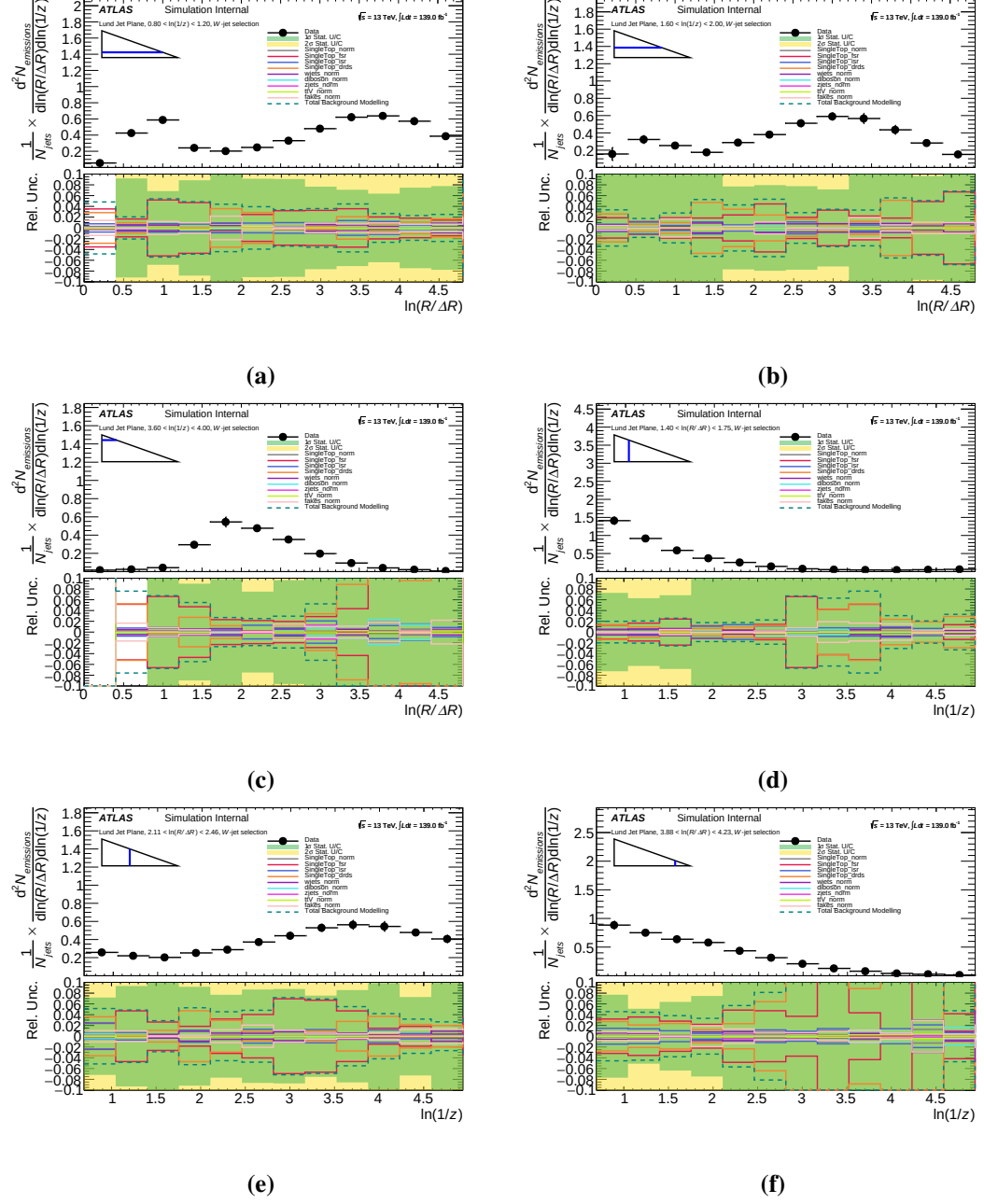


Figure 4.27: Summary of background modelling systematic uncertainties in slices through the Lund Jet Plane for the W jet selection.

4.7 Comparison of the unfolded results to alternative $t\bar{t}$ MC samples

Comparisons of the unfolded measurement and all the different Monte Carlo generators at the particle level are shown for the top jet selection in figures 4.28, 4.30 and 4.32 and for the W jets in figures 4.29, 4.31 and 4.33. The unfolded measurement and the particle level LJP are normalised to the total number of jets in each sample. A comparison of the total normalisation of the LJP, the ratio $N_{\text{emissions}}/N_{\text{jets}}$, between the different MC generators and the measurement is shown in appendix B.3 together with the associated uncertainties.

Figures 4.28 and 4.29 show the unfolded measurement compared to the four main MC samples: the nominal MC sample Powheg+Pythia 8, Powheg+Herwig 7.04, aMC@NLO+Pythia 8 and Sherpa 2.2.10. Powheg+Herwig 7.04 uses an alternative parton shower model to the nominal sample, aMC@NLO+Pythia 8 uses an alternative Matrix Element generator and Sherpa 2.2.10 uses an altogether different procedure to model both simulation steps. Differences between the MC generators and the measurement are mostly within the uncertainty bands in both the top jet and W jet selections. There are a few bins that stick out, however, such as at $0.4 < \ln(R/\Delta R) < 0.8$, $\ln(1/z) > 3.5$ where there is a discrepancy much larger than the total uncertainty for all generators. This bin is located in the area affected by the jet trimming, as can be seen in the plots of the data-driven non-closure in section 4.4.5. Since we observe unusually high non-closure in this region and we know that trimming is treated differently in data and MC events, it is safer to exclude this region of the LJP from the fiducial phase-space than to draw conclusions from the observed disagreement.

The LJP for aMC@NLO+Pythia 8 does not differ very much from either the measurement or the Powheg+Pythia 8 sample in the wide-angle, high- z region where we expect contributions from the hard-scatter process (the lower-left corner of the plane, approximately $0 \lesssim \ln(R/\Delta R) \lesssim 2$, $0.69 \lesssim \ln(1/z) \lesssim 2$). In the most densely populated bin in this region for the top jets, the observed difference between the measurement and aMC@NLO+Pythia 8 is about 2.5% (1% for the W jets), well within the total uncertainty of 5% (25% for W jets) and on the same scale as the differences to Powheg+Herwig

7.04, 1% (5% for W jets), Sherpa 2.2.10, 2.5% (5% for W jets), and Powheg+Pythia 8, 1% (5% for W jets). The largest differences for this sample are instead observed along a diagonal through the middle of the plane (eg. around $1.5 \lesssim \ln(R/\Delta R) \lesssim 3$ in the $1.4 < \ln(1/z) < 1.75$ slice shown in figure 4.28b) where the modelling uncertainty is large and the difference to the Powheg+Heriwig 7 LJP is of similar size. This results suggests that the observed differences are more due to the matrix element matching corrections to the parton shower, rather than the modelling of the hard scatter process in MadGraph5_aMC@NLO.

The LJP for Sherpa 2.2.10 deviates notably from the unfolded measurement and the other generators where both z and ΔR are small. This can be seen in the slice $3.6 > \ln(R/\Delta R) > 4.0$, that is in figure 4.28f and figure 4.29f. Here there are a number of bins (3 in the top jet selection and 5 for the W jets) above $\ln(1/z) = 2$ where the difference between Sherpa and the unfolded data is greater than the total uncertainty. A possible explanation for the difference between Pythia and Sherpa in this region is discussed in section 4.7.1.

Figures 4.30 and 4.31 show three different configurations of Powheg+Pythia 8 against the unfolded data. These three configurations are the nominal Powheg+Pythia 8 sample, the hdampVar sample where the Powheg parameter h_{damp} is set to $3.0m_t$ and a sample with Matrix Element matching corrections in Pythia turned off. The differences between the nominal Powheg+Pythia 8 and the hdampVar configuration are very small across the whole LJP. As the h_{damp} parameter controls the size of the hardest emission in the parton shower, the radiation it controls is almost entirely due to the $t\bar{t}$ matrix element. It can be concluded that the impact of this radiation on the density of LJP is minimal for both top jets and W jets. The Powheg+Pythia 8 sample with the ME corrections turned off exhibits the a similar behaviour at small z to the aMC@NLO+Pythia 8 sample (compare e.g. the distributions in figure 4.33d and figure 4.28d). This suggests that the discrepancy between the Powheg and aMC@NLO samples is mainly due to the parton shower matching in Pythia 8, rather than the modelling of the hard scatter process itself. This is also confirmed by the ratio of particle level LJPs for aMC@NLO+Pythia

8 and Powheg+Pythia 8 without the ME corrections, shown for the whole 2-d LJP in section 4.7.1 and section 4.7.1, which is very close to 1 everywhere.

The three different versions of Powheg+Herwig 7 are shown compared to the unfolded data in figures 4.32 and 4.33. The LJP density varies similarly for all three generators across the LJP. The largest differences with the measurement are observed along the diagonal of the plane where the systematic uncertainty is largest. This is also the region where the differences between the three generators are largest, as can be seen in the ratio of 7.1.3 to 7.04 and 7.2 to 7.04 shown in figure 4.37 and figure 4.37 in the following section.

Maps of the Lund Jet Plane showing the MC generators that are closest (or furthest) from the data are presented in appendix B.5. In the top-jet selection, the Monte Carlo generators that best fit the data in the hard and wide angle region ($\ln(1/z) < 1.5$, $\ln(R/\Delta R) < 0.8$) are Powheg + Herwig 7.1.3 and Powheg + Herwig 7.2. Powheg+Herwig 7.1.3 fits the data best in 3 out of 4 bins in this region with differences between the data and Monte Carlo ranging between 0.02 to 0.38 times the total uncertainty (σ_{tot}).

In the middle region of the plane ($1.5 < \ln(1/z) < 2.5$, $0.8 < \ln(R/\Delta R) < 2.5$), the best match to the data is observed for the Powheg + Pythia 8 sample without ME corrections which shows the smallest difference with the data in 8 out of the 27 bins in this region. The next best match is provided by the aMC@NLO + Pythia 8 sample which matches the data most closely in 5 out of the 27 bins. As is discussed in section 4.7.1 and shown in figures 4.36 and 4.37, the difference between these two generator configurations is small ($< 5\%$) across nearly all of the LJP. Both these configurations handle the matching between matrix-element and parton shower generation in a similar way. Therefore, it is possible that this matching configuration is the most accurate for modelling the LJP. However, other Monte Carlo generators, such as the nominal Powheg + Pythia 8 or Powheg + Herwig 7, also match the data to within the uncertainty bands inside this region. Differences between Powheg + Pythia 8 (ME. corr. off) and the data range between $0.01\sigma_{tot}$ to $0.83\sigma_{tot}$, while the differences between aMC@NLO + Pythia 8 and the data range between $0.01\sigma_{tot}$ and $0.22\sigma_{tot}$.

In the small-angle region ($3.0 < \ln(R/\Delta R) < 4.5$), large differences are observed between Sherpa 2.2.10 and the data. In 25 out of 40 bins in this region, the greatest difference between the measurement and Monte Carlo is observed for Sherpa 2.2.10. The difference is greater than σ_{tot} in 15 of these bins and its maximum value is $2.6\sigma_{tot}$. As is discussed in section 4.7.1 and shown in figure 4.34, there is a considerable difference in the number of emissions in this region between Sherpa 2.2.10 and Powheg + Pythia 8, possibly due to the absence of colour-reconnection in the former generator. However, these differences are still within the total uncertainty.

In the W -jet selection, hard and wide-angle emissions ($\ln(1/z) < 1.5$, $\ln(R/\Delta R) < 1.5$) are best modelled by two of the Powheg + Pythia 8 configurations (nominal and $h_{damp} = 3.0m_t$) and the aMC@NLO + Pythia 8 sample. The difference between these MC samples and the data ranges from $0.07\sigma_{tot}$ to $0.58\sigma_{tot}$. The largest disagreement in this region is observed for Powheg + Herwig 7.1.3 in 3 out of the 4 bins at $0.36\sigma_{tot}$ to $0.69\sigma_{tot}$. In the populated bin with the lowest $\ln(1/z)$ and $\ln(R/\Delta R)$, the largest difference between data and Monte Carlo is with the aMC@NLO + Pythia 8 sample at $0.73\sigma_{tot}$.

In the middle region ($1.5 < \ln(1/z) < 3.0$, $1.5 < \ln(R/\Delta R) < 3.0$), close to the densely populated diagonal of the LJP, the MC generators that most closely match the data are Powheg + Herwig 7.2 and Powheg + Herwig 7.1.3. In 9 out of 27 bins in this region, the smallest difference between data and MC is observed for Powheg + Herwig 7.2 with values in the range $0.22\sigma_{tot}$ to $1.04\sigma_{tot}$, while in 7 bins the smallest values are observed for Powheg + Herwig 7.1.3 ranging from $0.04\sigma_{tot}$ to $0.65\sigma_{tot}$. There are also 4 bins in this region where the closest match to the data is Powheg + Herwig 7.0.4 with differences ranging from $0.001\sigma_{tot}$ to $0.75\sigma_{tot}$. This result suggests that the parton shower and hadronisation effects in W jets are best modelled by the Herwig parton shower and hadronisation simulation and that the newest version tested, Herwig 7.2, models the data better.

Similarly to the top jets, there are large differences between Sherpa 2.2.10 and the data in the small-angle region ($3.0 < \ln(R/\Delta R) < 4.5$). The data-MC difference is greatest for Sherpa 2.2.10 in 16 out of the 39 populated bins in this region. In 5 of these bins, the data-MC difference is greater than σ_{tot} up to a maximum value of $4.4\sigma_{tot}$. These

large differences are due to the increased amount of soft and wide-angle radiation in Sherpa. This effect, which is shown compared to Powheg + Pythia 8 in figure 4.35, is also observed relative to the data, suggesting that it is the result of mismodelling. A possible source of mismodelling, the absence of colour-reconnection modelling in Sherpa, is discussed in section 4.7.1.

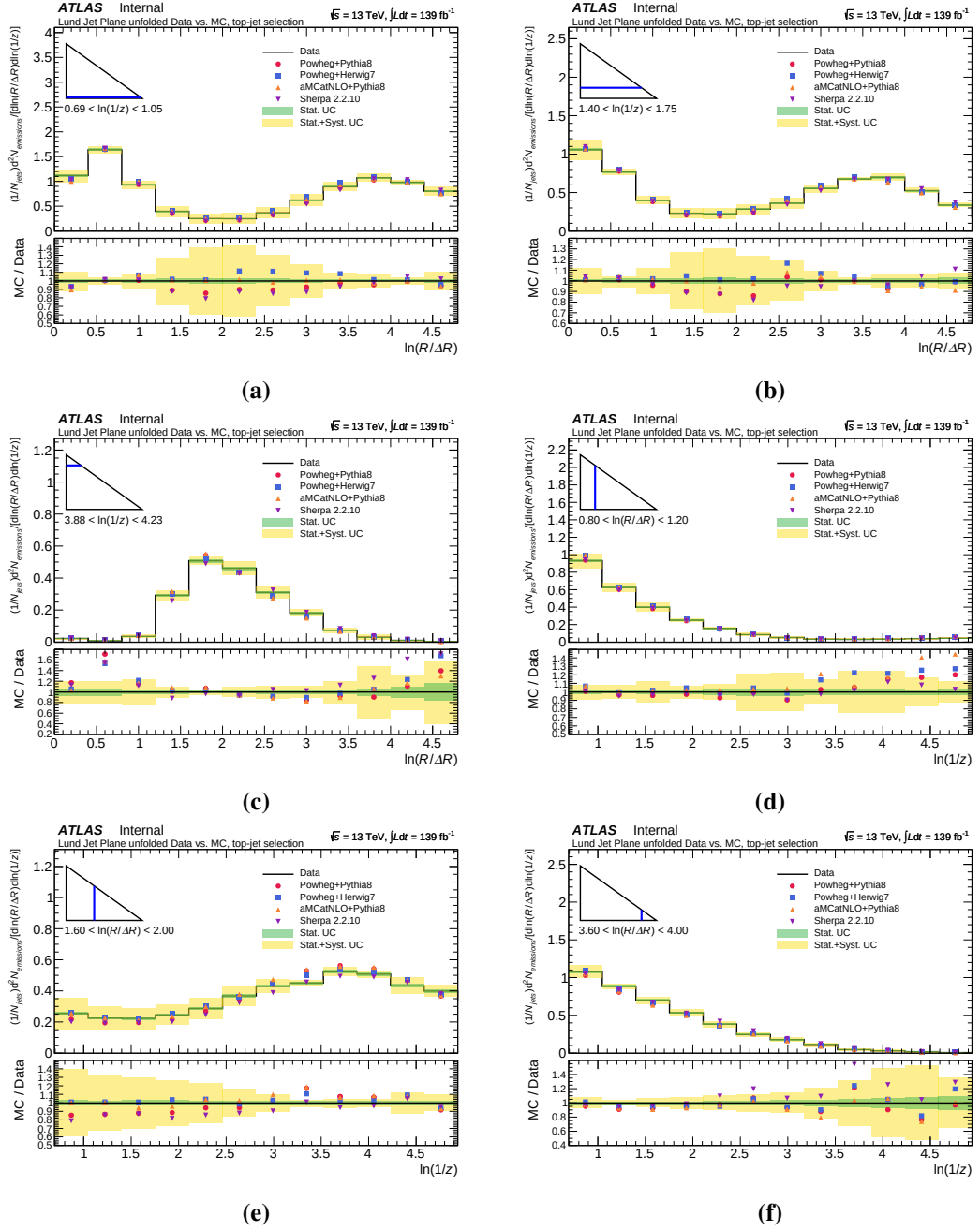


Figure 4.28: Slices through the Unfolded Lund Jet Plane for top jets comparing the Powheg+Pythia 8, Powheg+Herwig 7.04, aMC@NLO+Pythia 8 and Sherpa MC generators at the particle level to the unfolded measurement. The statistical uncertainty and the total uncertainty on the unfolded measurement are also shown.

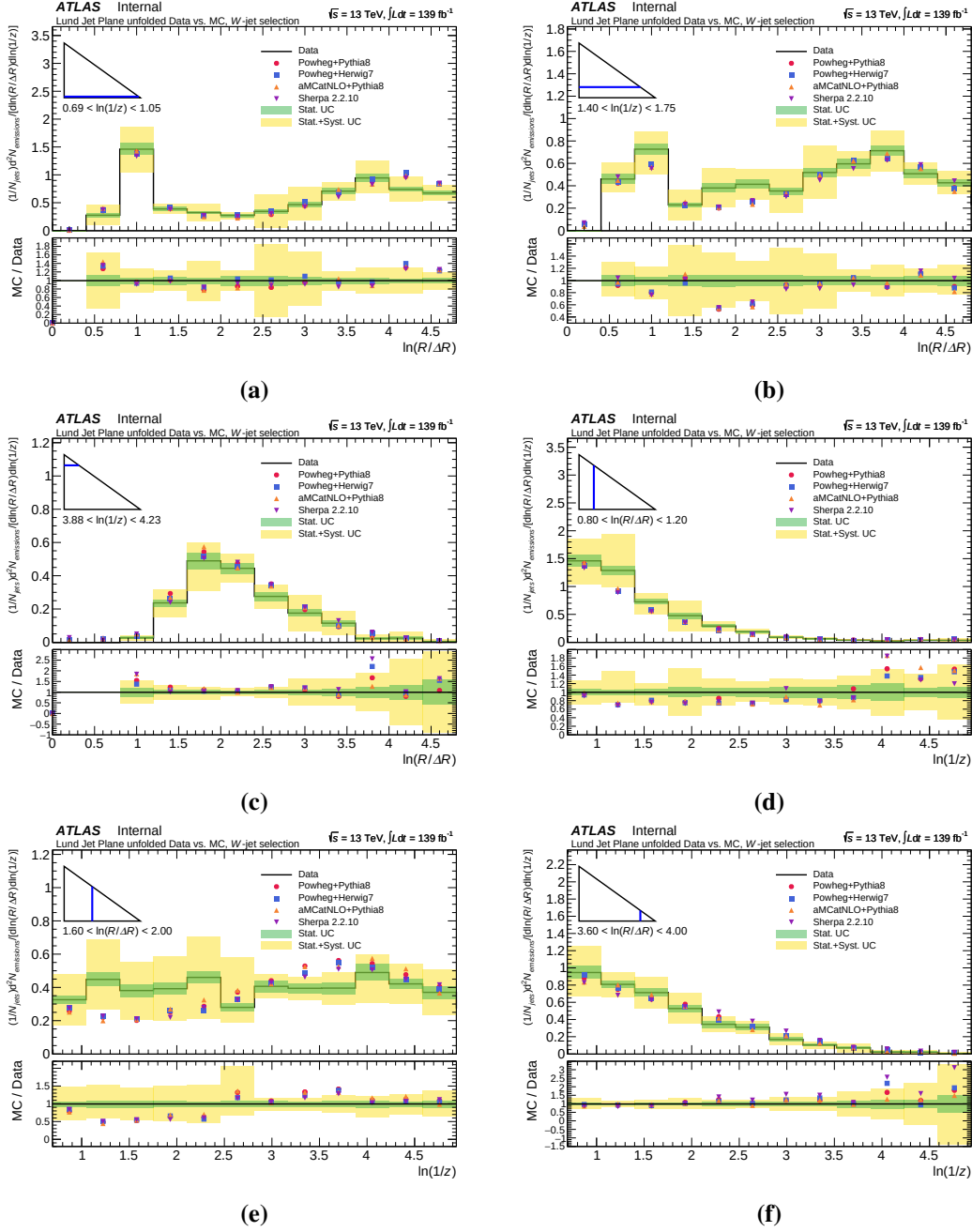


Figure 4.29: Slices through the Unfolded Lund Jet Plane for W jets comparing the Powheg+Pythia 8, Powheg+Herwig 7.04, aMC@NLO+Pythia 8 and Sherpa MC generators at the particle level to the unfolded measurement. The statistical uncertainty and the total uncertainty on the unfolded measurement are also shown.

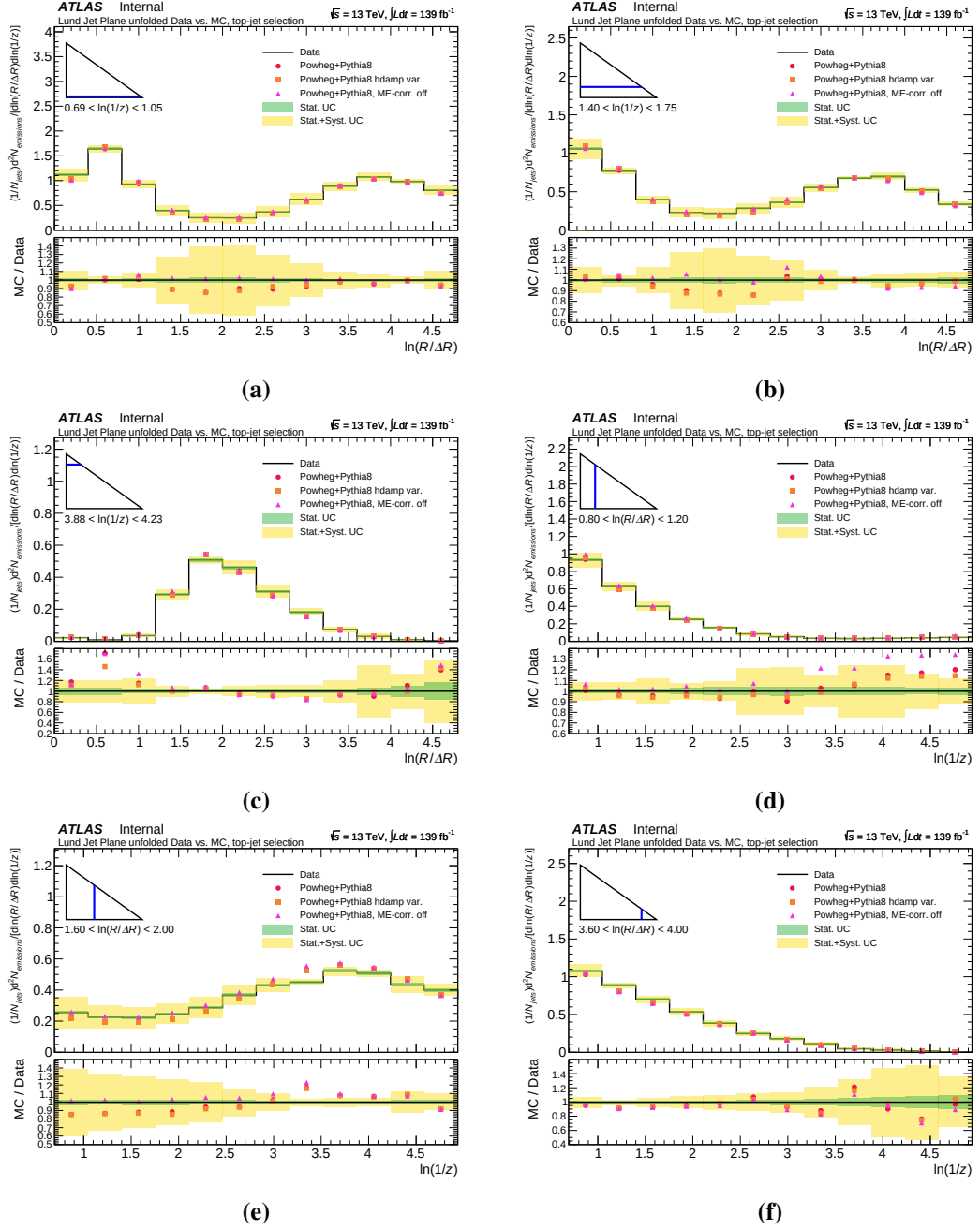


Figure 4.30: Slices through the Unfolded Lund Jet Plane for top jets comparing Powheg+Pythia 8 MCs at the particle level to the unfolded measurement. The samples shown are the nominal PP8 MC, the h_{damp} variation and PP8 without Matrix-Element matching correction. The statistical uncertainty and the total uncertainty on the unfolded measurement are also shown.

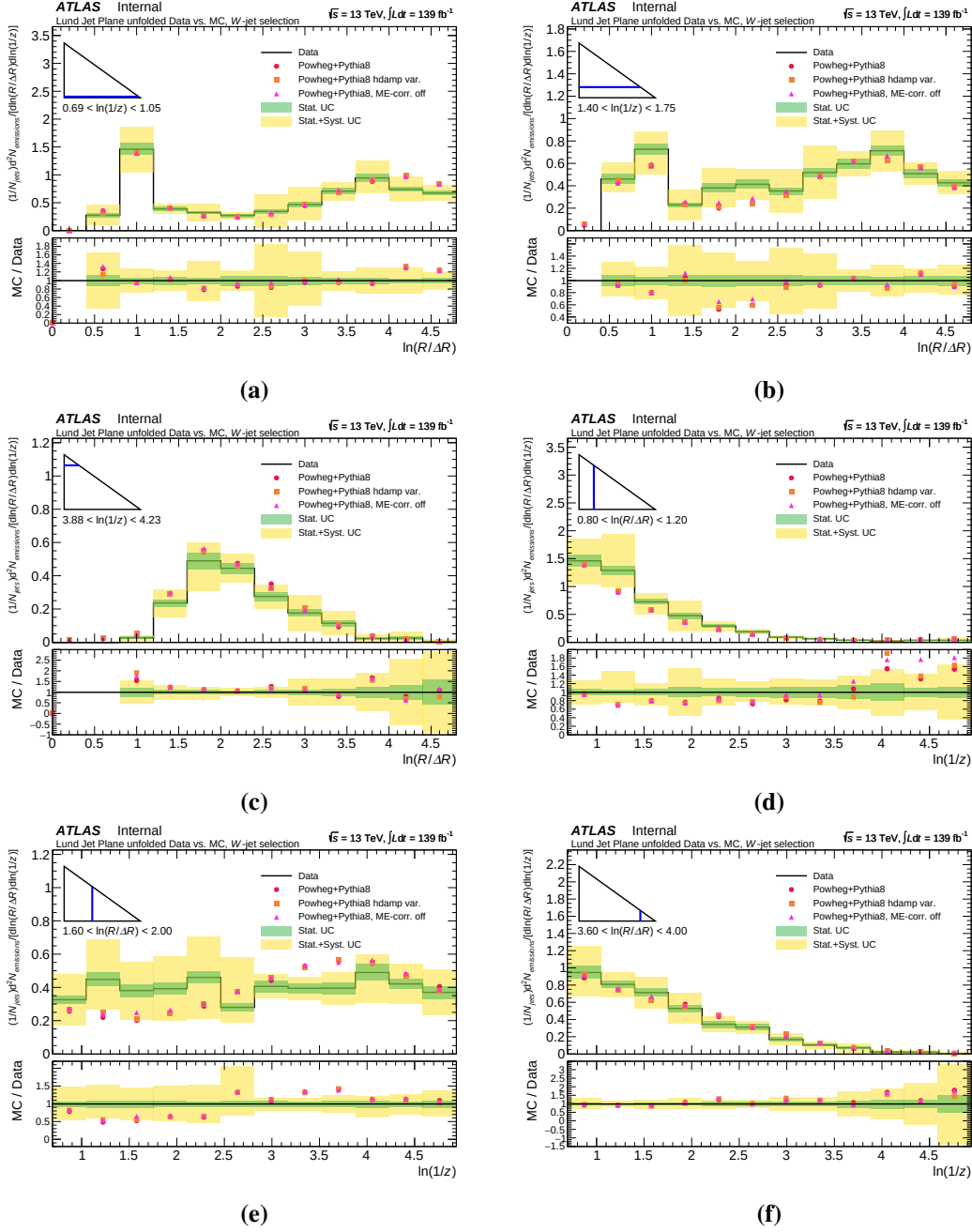


Figure 4.31: Slices through the Unfolded Lund Jet Plane for W jets comparing Powheg+Pythia 8 MCs at the particle level to the unfolded measurement. The samples shown are the nominal PP8 MC, the h_{damp} variation and PP8 without Matrix-Element matching correction. The statistical uncertainty and the total uncertainty on the unfolded measurement are also shown.

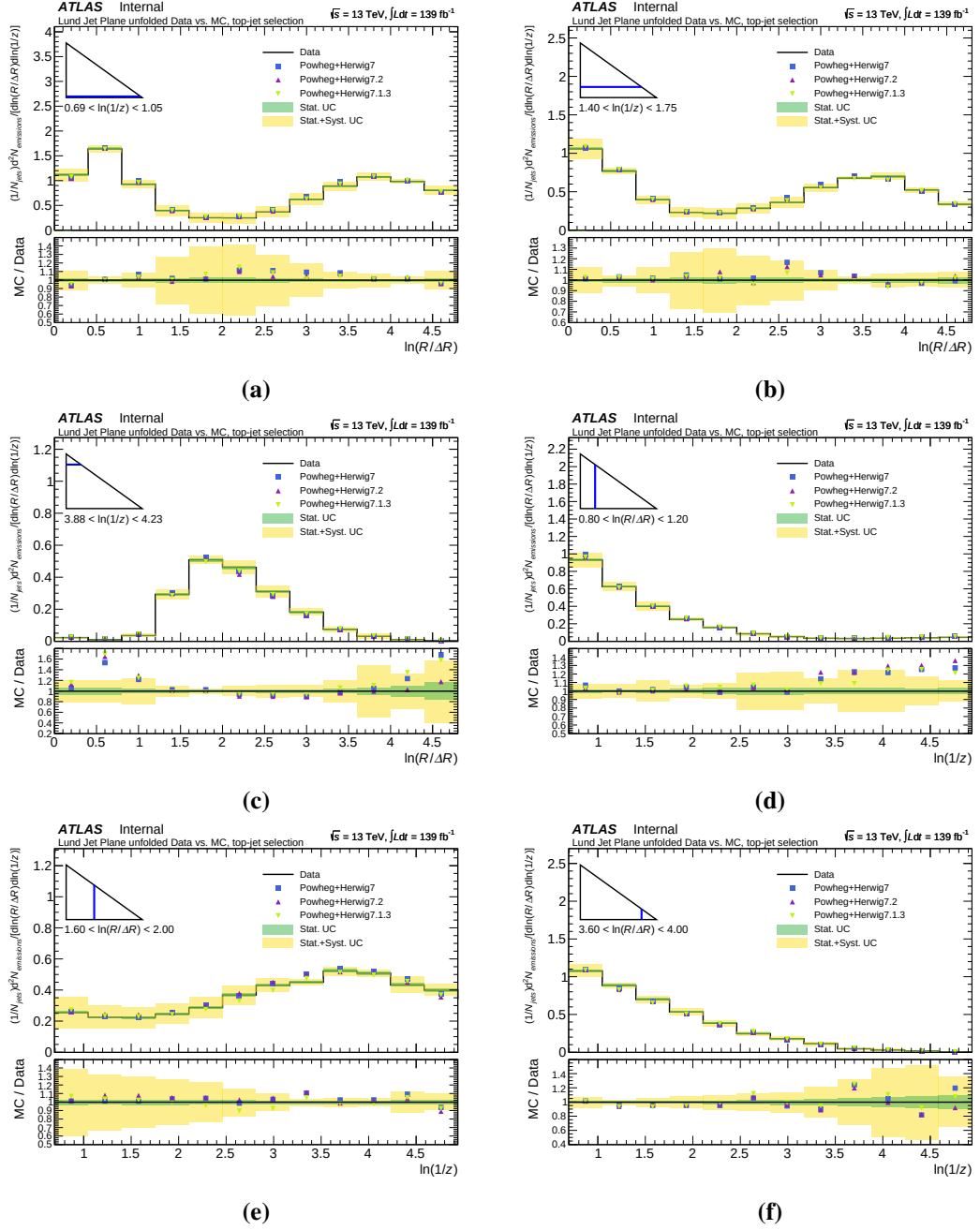


Figure 4.32: Slices through the Unfolded Lund Jet Plane for top jets comparing Powheg+Herwig 7 MCs at the particle level to the unfolded measurement. The versions of Herwig compared are 7.04, 7.1.3 and 7.2. The statistical uncertainty and the total uncertainty on the unfolded measurement are also shown.

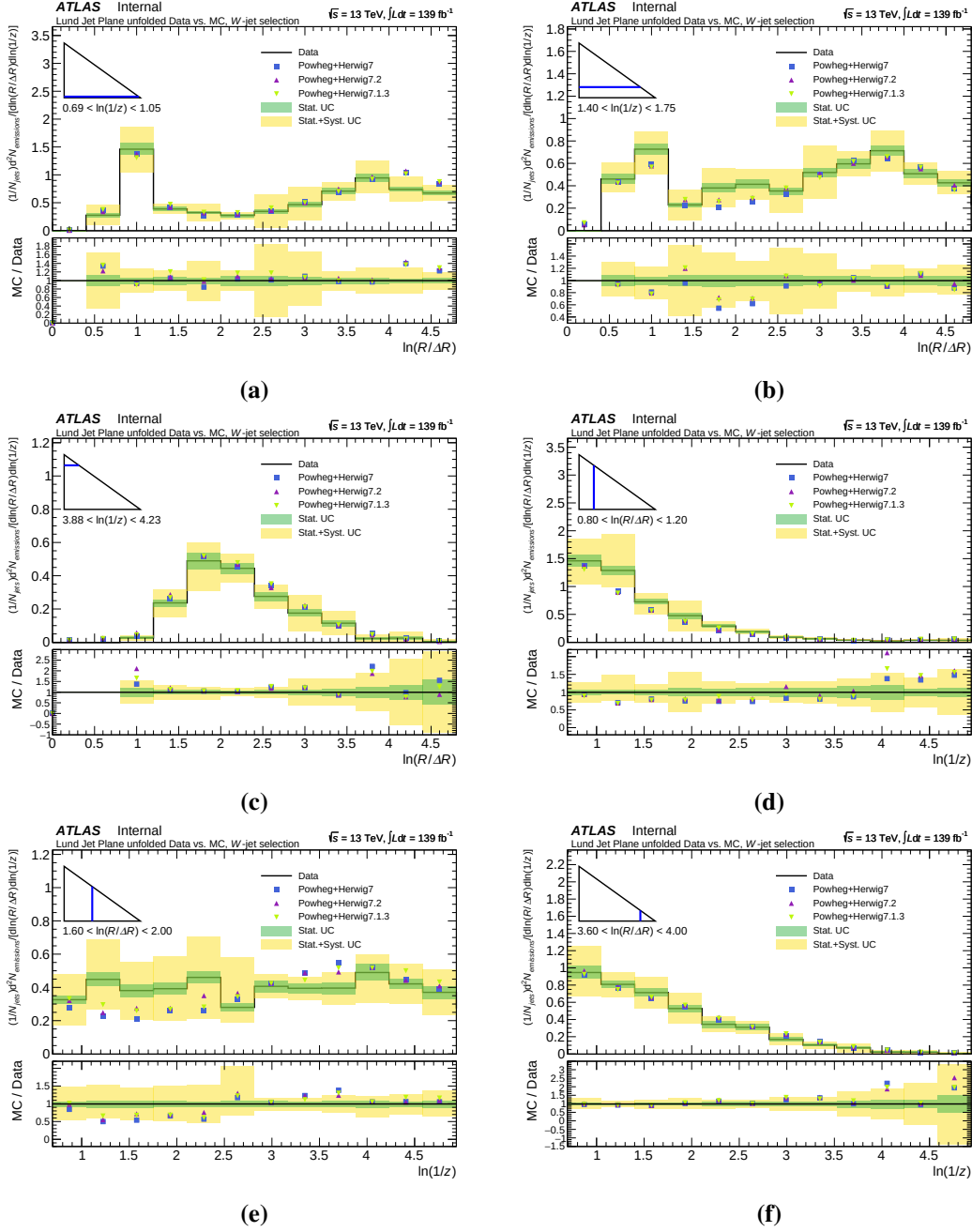


Figure 4.33: Slices through the Unfolded Lund Jet Plane for W jets comparing Powheg+Herwig 7 MCs at the particle level to the unfolded measurement. The versions of Herwig compared are 7.04, 7.1.3 and 7.2. The statistical uncertainty and the total uncertainty on the unfolded measurement are also shown.

4.7.1 Comparison of $t\bar{t}$ MC/MC ratios

To better show the effects of variations in the modelling of $t\bar{t}$ events on the Lund Jet Plane, ratios between the alternative MC generators and the nominal Powheg+Pythia 8 sample at particle-level are shown in figures 4.34 and 4.35. figures 4.36 and 4.37 show other MC/MC ratios that are of interest, namely the ratios of Powheg+Herwig 7.1.3 and Powheg+Herwig 7.2 to Powheg+Herwig 7.04 and aMC@NLO+Pythia 8 to Powheg+Pythia 8 with ME corrections turned off. In the ratios to Powheg+Pythia 8, most of the generators using alternative parton shower models show an increase in emissions in the lower left triangle bounded by $1 \lesssim \ln(R/\Delta R) \lesssim 3$ and $\ln(2) < \ln(1/\Delta R) \lesssim 4$. The ratio of the Powheg+Pythia 8 sample with the varied h_{damp} parameter in Powheg is close to 1 across the whole LJP.

In the ratio of Sherpa 2.2.10 to Powheg+Pythia 8, in section 4.7.1 and section 4.7.1, there is a notable increase in the density of emissions in the upper right-hand corner of the LJP. This increase in soft and collinear emissions might be explained by the absence of colour-reconnection modelling (CR) in this generator. Colour-reconnection is known to be a necessary feature when modelling hadronic top quark decays and for accurately predicting the top quark mass [214]. The effect of CR mismodelling has also been measured by ATLAS using tracks outside $R = 0.4$ jets [215]. The need for CR modelling stems from the leading colour approximation which is utilised by most MC event generators. In this approximation, quarks and gluons are connected to exactly one, respectively two, other partons by objects called colour-strings. This approximation no longer holds in the presence of non-perturbative interactions with a large number of partons, i.e. multi-parton interactions, as these would introduce a multitude of new colour-strings with overlapping phase spaces. This means that there is an over-production of radiation due to MPI, and therefore greater multiplicities of low- p_T tracks both inside and outside jets, in models that do not model interactions with the MPI colour strings. Colour reconnection models attempt to correct this mismodelling by re-assigning colours to partons as multi-parton systems are introduced into the calculations [216]. In ref. [214] it is shown how the inclusion of CR in the modelling of $t\bar{t}$ events increases the average p_T of charged particle tracks at large track multiplicities and, in some cases, produces

higher energy densities at wide angles within jets. The LJP for Sherpa 2.2.10 shows a slight suppression of emissions in the large z (p_t fractions) and large- ΔR regions, and a sizeable increase of emissions at small z and ΔR relative to the nominal Powheg + Pythia 8 sample. Since the LJP is constructed from charged particles, this suggests an increased production of charged particles at low p_T . The absence of CR modelling in Sherpa is a possible explanation for this effect, as it would increase the number of low p_T emissions and produce a shift in the momentum and angular distributions of radiation within the jets towards lower momentum fractions and smaller angles. A way to confirm that the observed differences in Sherpa are indeed due to the absence of CR would be to produce additional LJPs from a Powheg+Pythia 8 MC sample with CR modelling turned off. Although it would be straightforward to carry out such a study, as alternative Powheg+Pythia 8 samples have been produced within the ATLAS top group for the studies in ref. [215], it has not been possible to include it here due to time constraints.

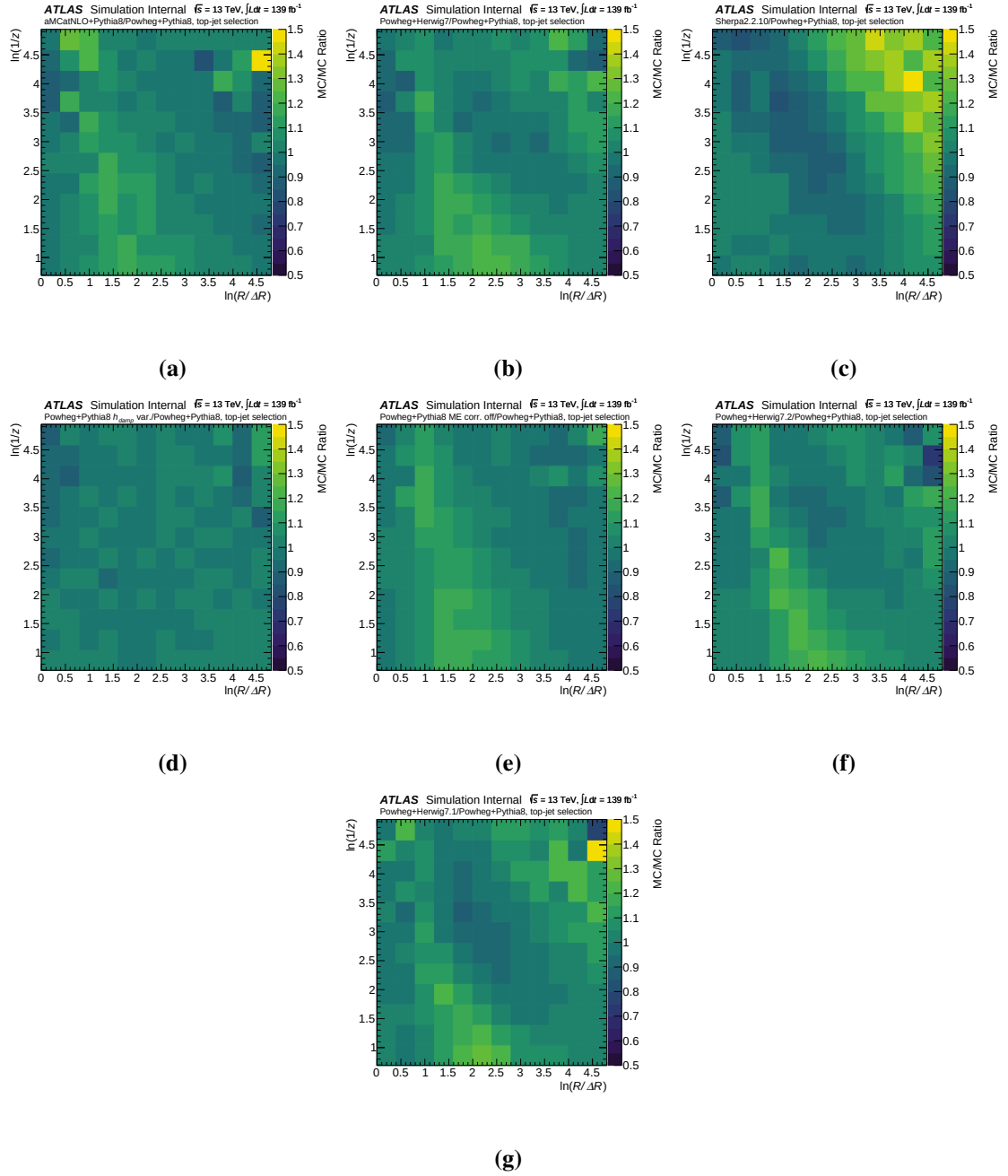


Figure 4.34: Ratio of LJPs for a range of Monte Carlo generators in the top jet selection. Figures (a-g) show the ratios of aMC@NLO+Pythia 8, Powheg+Herwig 7.04, Sherpa 2.2.10, Powheg+Pythia 8 ($h_{\text{damp}} = 3.0m_t$), Powheg+Pythia 8 (ME corrections off), Powheg+Herwig 7.2 and Powheg+Herwig 7.1.3 to the nominal Powheg+Pythia 8 sample.

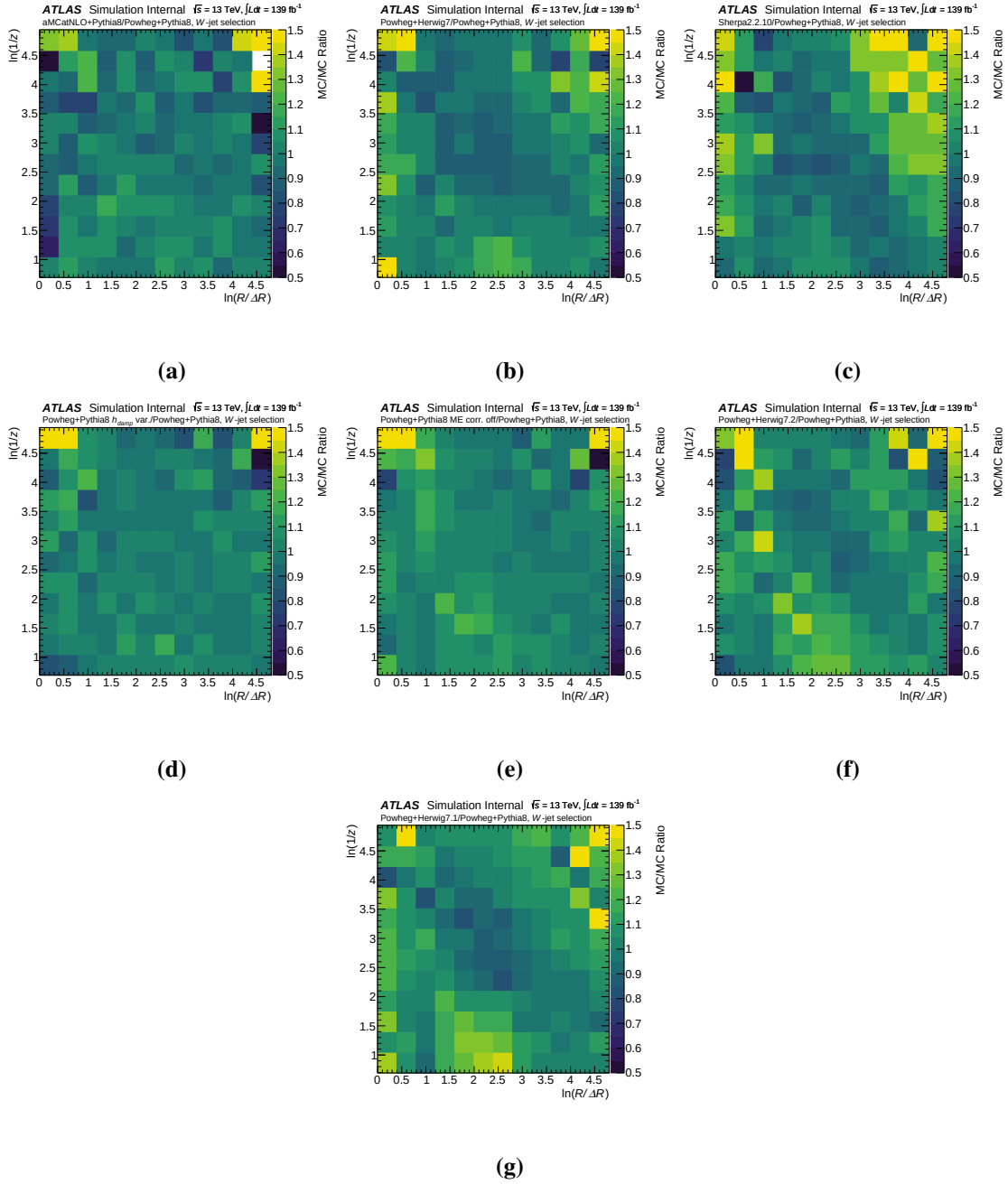


Figure 4.35: Ratio of LJPs for a range of Monte Carlo generators in the W jet selection. Figures (a-g) show the ratios of aMC@NLO+Pythia 8, Powheg+Herwig 7.04, Sherpa 2.2.10, Powheg+Pythia 8 ($h_{\text{damp}} = 3.0m_t$), Powheg+Pythia 8 (ME corrections off), Powheg+Herwig 7.2 and Powheg+Herwig 7.1.3 to the nominal Powheg+Pythia 8 sample.

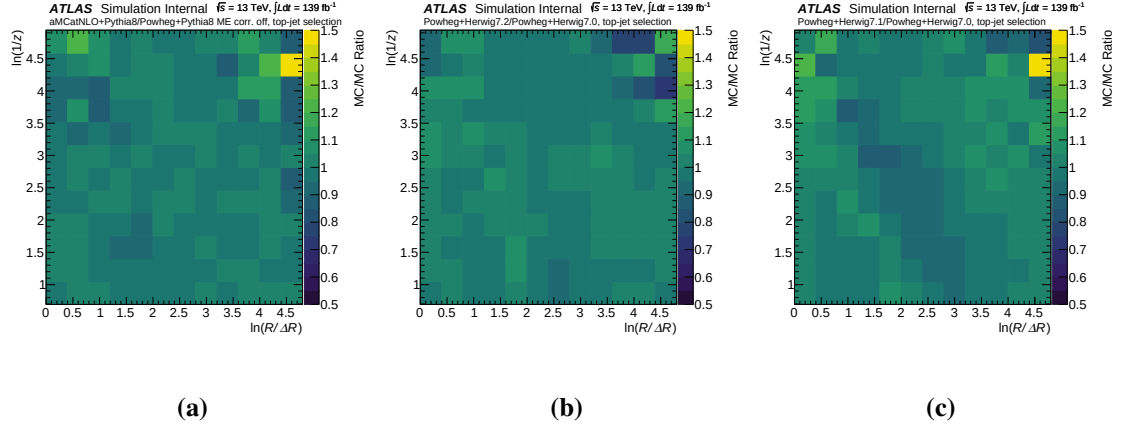


Figure 4.36: Ratio of LJPs for a range of Monte Carlo generators in the top jet selection. Figure (a) shows the ratio of aMC@NLO+Pythia 8 to Powheg+Pythia 8 (ME corrections off). Figures (b) and (c) show the ratios of Powheg+Herwig 7.1.3 and Powheg+Herwig 7.2 to Powheg+Herwig 7.0.

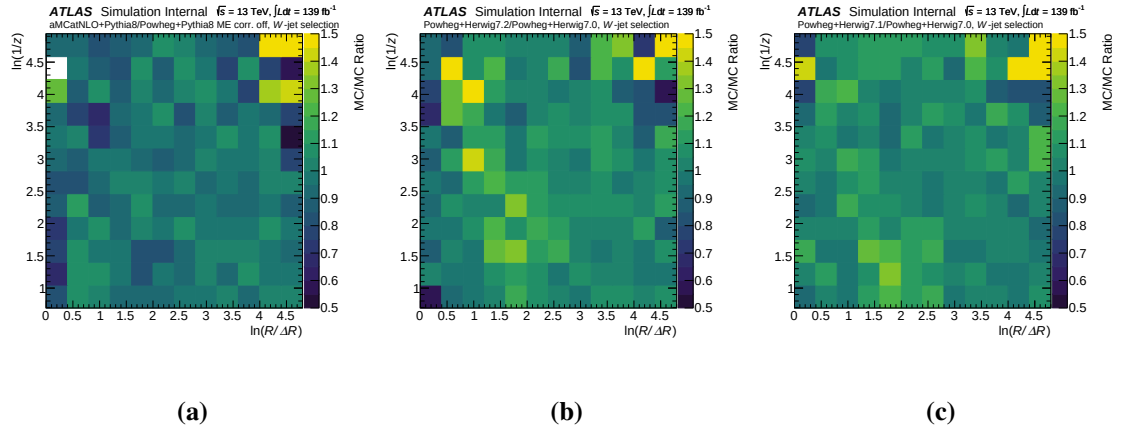


Figure 4.37: Ratio of LJPs for a range of Monte Carlo generators in the W jet selection. Figure (a) shows the ratio of aMC@NLO+Pythia 8 to Powheg+Pythia 8 (ME corrections off). Figures (b) and (c) show the ratios of Powheg+Herwig 7.1.3 and Powheg+Herwig 7.2 to Powheg+Herwig 7.0.

4.8 Conclusion

A measurement of the Lund Jet Plane in $t\bar{t}$ events using the ATLAS Run 2 data has been presented. This is the first measurement of this observable for boosted-object topologies using $R = 1.0$ jets. The measurement complements other measurements of the LJP, such as the dijets measurement in chapter 3 or the measurements by ALICE described in section 3.5, which have only probed small-radius jets predominantly initiated by light quarks and gluons. This measurement is the first to present the LJP for W boson and top quark initiated jets. It is therefore a useful resource to those wishing to construct W and top taggers making use of the LJP. Several studies have proposed the LJP as a powerful discriminant in top and W jet tagging [114, 217].

The measurement is unfolded to the particle level using Iterative Bayesian Unfolding with 4 iterations. Statistical and systematic uncertainties are evaluated on the unfolded measurement. Systematic uncertainties have been propagated by unfolding systematic variations on the nominal Powheg + Pythia 8 Monte Carlo sample using the nominal response. For top jets, the total systematic uncertainty in the centre of the LJP, where it is most densely populated (around $0.5 \lesssim \ln(R/\Delta R) \lesssim 3.5, \ln(1/z) \lesssim 3.5$), is found to be 20 – 30%, while for W jets it is 40 – 60%. The number of unfolding iterations is optimised to minimise the impact of the unfolding procedure on the uncertainties.

The unfolded measurement was compared to a wide range of Monte Carlo samples, chiefly: Powheg + Pythia 8, Powheg + Herwig 7.0.4, Powheg + Herwig 7.1.3, Powheg + Herwig 7.2, MadGraph5_aMC@NLO + Pythia 8 and Sherpa 2.2.10. Differences between the unfolded measurement and the different generators at the particle level were found to be within the estimated uncertainties across most of the LJP. However, some differences in the Monte Carlo LJP distributions are observed which are thought to be due to differences in modelling of various types of physical effects. In some regions of the LJP, the differences between the data and the Monte Carlo samples are sufficiently large to be incompatible with the measurement.

All the Monte Carlo generator configurations tested show discrepancies with the measurement in some parts of the LJP. The largest and smallest discrepancies are observed for different generators in different parts of the LJP, reflecting the fact that the

stages of jet production, from hard scatter to parton showering and hadronisation, are modelled differently in each generator configuration. At wide angles and large momentum fractions, the data is best modelled by Powheg + Herwig 7.1.3 for the top jets and aMC@NLO + Pythia 8 for the W jets. In the middle section of the plane, where it is densely populated and dominated by parton shower and hadronisation effects, different configurations of Pythia 8 give the closest match to the data for top jets, while for W jets this region is best modelled by MCs using different versions of the Herwig 7 parton shower. At small angles and low momentum fractions there is no clear winner. However, it is evident that Sherpa 2.2.10 models this region badly, as discrepancies with the measurement of up to 4.4 times the total uncertainty are observed in some bins. The exact source of mismodelling in this kinematic regime is unknown, however it is possible that the absence of colour-reconnection modelling in the Sherpa parton shower contributes to this effect. This could be further investigated in a future study using different configurations of the other MC generators, such as Pythia 8, with and without colour-reconnection modelling.

Chapter 5

Deep Learning W and top jet taggers

5.1 Introduction

The Lund Jet Plane for W and top quark initiated jets – measured in the previous chapter – shows many distinctive features compared to the LJP for light quark jets. This was already noted by Dreyer et al. in ref. [114] and it was suggested that this property of the LJP could be exploited by machine-learning based jet taggers using LJPs as inputs. The chapter explores a number of different machine learning top and W taggers that make use of LJPs constructed from ATLAS detector level information. The performance of the taggers is evaluated and compared to a simpler machine-learning tagger using jet images and the 3-variable cuts-based W tagger.

5.2 Machine learning in jet tagging

Machine learning methods have played an important role in the development of novel jet tagging techniques over the last few years. The advent of powerful *Deep Neural Networks*¹ that can easily be constructed with user-friendly Python libraries, such as TensorFlow [218], Keras [219] and most recently PyTorch [220], has given rise to a large number of brand new machine-learning based taggers. Thanks to the complex internal structure of jets, and the algorithms used to construct them, there is a wide range of neural network architectures that can be used to exploit different features of the jet substructure in order to classify the origin of a jet.

¹A thorough description of the of the many machine-learning techniques that have perfected in the 2010s and are commonly referred to as "Deep Learning" can be found in ref. [221].

The range of possible applications of deep learning techniques in jet tagging is not just limited to constructing neural-network based classifiers with the aim of attaining the highest possible tagging efficiency. One example is mass-decorrelation using adversarial neural networks [222], which aims to reduce the shaping of the mass distribution of background jets.

5.2.1 Performance metrics

The performance of neural network classifiers is typically evaluated using a *Receiver-Operator Characteristic* or ROC curve. This curve displays information about the classifiers performance both at correctly classifying signal and at rejecting background. As the output of the classifier is a continuous score, there is a range of working points that can be fixed by choosing which scores to accept and which to reject. The ROC curve displays the range of possible working points. There are a number of different metrics from which ROC curves can be constructed. This section summarises the definitions of some of the most common performance metrics and how they relate to the ROC curve.

True positive rate (TPR) The fraction of correctly classified signal samples:

$$TPR = \frac{N_{sig}^{tag}}{N_{sig}^{tag} + N_{sig}^{untag}}. \quad (5.1)$$

An optimal classifier should achieve a TPR as close to 1 as possible while keeping the amount of miss-identified background low.

False positive rate (FPR) The fraction of background samples incorrectly classified as signal:

$$FPR = \frac{N_{bkg}^{tag}}{N_{bkg}^{tag} + N_{bkg}^{untag}}. \quad (5.2)$$

On the ROC curve it is usually plotted against the TPR and should be close to zero for as many values of the TPR as possible.

ROC curve AUC The area under the ROC curve where the FPR is plotted on one axis and TPR on the other. Values for the ROC AUC range from 0.5 for a random classifier to 1.0 for an optimal one. A typical ROC curve using the TPR and FPR metrics is shown in figure 5.3.

Signal efficiency ϵ_{sig} The efficiency with which signal events are correctly classified. Its definition is identical to the true positive rate. The term is more commonly used in the context of taggers in particle physics than in the field of machine learning.

Background Rejection The rate with which background is rejected:

$$1/\epsilon_{bkg} = \frac{N_{bkg}^{tag} + N_{bkg}^{untag}}{N_{bkg}^{tag}}. \quad (5.3)$$

It is typically plotted against the signal efficiency on the ROC curve. For an optimal tagger that rejects all background the background rejection is infinite.

5.2.2 Monte Carlo samples

The taggers are trained on Monte Carlo simulated events. Background jets are obtained from multijet samples, while the W or top signal are taken from samples of $W' \rightarrow WZ \rightarrow qqqq$ and $Z' \rightarrow t\bar{t}$ events respectively. These events are generated using the *Heavy Boson Triplets* [223] framework for simulating BSM particles. BSM heavy gauge bosons, called W' and Z' , are created with a specified mass to produce large quantities of W bosons or top quarks. The masses of the W' and z' are both set to 2 TeV and the events are reweighted to produce a flat p_T distribution of jets in the range 200 – 300 GeV. The remainder of the simulation chain in the signal samples is simulated using Pythia 8 (v 8.235) [38], using the A14 set of tuned parameters [141] and the NNPDF2.310 set of PDFs [140]. The multijet background samples are generated using Pythia 8.230 with the A14 tune and the NNPDF2.310 set of PDFs. The decays of bottom and charm hadrons are simulated using the EvtGen v1.6.0 program [166].

The complete response of the ATLAS detector [224] is simulated using Geant 4 [123]. The effect of multiple proton-proton interactions per bunch crossing (pileup) is simulated by overlaying events with Minimum-Bias samples generated using Pythia 8 with the NNPDF2.310 [140] set of PDFs using the A3 tune [122]. The minimum bias events are reweighted to approximate the pileup conditions in Run 2 data.

The mass of the W'/Z' is varied from 400 GeV to 5 TeV such that a flat p_T distribution is achieved for W/t jets in the range 400 GeV to 2 TeV.

Multijet events are obtained from simulations of the dijet process. In order to obtain a large number of samples across a wide range of jet energies, events are generated in slices of particle-level $R = 0.6$ jet p_T (JX samples). The slices are weighted by cross-section so that a physical, smoothly falling p_T -spectrum is achieved.

5.2.3 Jet definitions and truth-jet labelling

Calibrated large- R jets, typically $R = 1.0$ trimmed anti- k_T topocluster jets, are used to apply kinematic selections on the jet p_T at the detector level. However, it is the untrimmed constituents, or in the case of the Lund Jet Planes the tracks contained within the jet, that are used to construct the jet images and LJPs.

The detector-level jets are matched in $\eta - \phi$ space to anti- k_T truth jets which are labelled as top, W or QCD jets depending on the number of ghost-associated B -hadrons. For W jets there must be no ghost-associated B -hadron, and for top jets there must be ≥ 1 . To ensure that the decay products of the top quarks and W bosons are sufficiently contained within the truth jet, it must also have a jet mass close to the W or top mass. This is the same truth-labelling scheme as was used in the evaluation of other taggers at ATLAS and is further described in section 2.5.2.

5.3 Jet Image taggers

Jet images are a straightforward representation of the kinematics of a jet's constituents. They typically consist of 2D histograms over $\eta - \phi$ or $\eta - y$ coordinates, with the jet-axis at the origin, where the content of each bin represents a measure of the constituent energy or momentum that is invariant under Lorentz boosts.

The advent of powerful neural-network image recognition algorithms in the early 2010s sparked a growing interest among the LHC jet physics community to apply these techniques to boosted-object jet tagging. The first studies to construct images from the constituents of jets [225] used simple linear discriminants, like principal component analysis (PCA) and Fisher's Linear Discriminant (FLD), to distinguish images of W jets from QCD background. Improved tagger performance was soon achieved by substituting these discriminants for neural networks [226, 227] and convolutional neural networks which are specialised for image classification tasks. Jet images have since also been used as

inputs to neural network-based top taggers in independent studies by other groups [228–230]. Both ATLAS and CMS have implemented neural-network based taggers using jet images as inputs and compared their performance to state-of-the-art taggers [231, 232]. Even in more recent years, as sequence or graph-based machine learning algorithms have become more popular in jet tagging applications, performant taggers are still being constructed using jet images in combination with other jet substructure observables [233–235].

The images are constructed from the calorimeter topoclusters that make up the $R = 1.0$ Anti- k_T jets. No grooming is applied on these topoclusters. The x and y -coordinates of the images are the η and ϕ coordinates relative to the jet axis, hence the axis of the jet is always at the centre of the images. A rotation $(\eta, \phi) \mapsto (\alpha_1, \alpha_2)$ is applied to the topoclusters before filling the images such that the highest energy topocluster is always located at $\alpha_1 = 0$, or "12 o'clock" in the plotted image. Each entry into the image histograms is weighted by the transverse topocluster energy $E_T^c = E^c \cosh(\eta^c)$ which is Lorentz invariant and hence unaffected by translations along η .

Average jet images for 10,000 W , top and QCD initiated jets, as they were constructed for our tagging studies, are shown in figure 5.1. There is a visible difference in the distribution of energy across the images for QCD, top quark and W boson initiated jets. For W -jets, there is a large number of entries at $\alpha_1 = 0, \alpha_2 > 0$ where the highest energy jet constituent is located and there is a broader peak in the α_2 profile of the jet image. This is due to the two-pronged topology of the $W \rightarrow q\bar{q}$ decay, a large fraction of the energy of one of the decaying quarks becomes reconstructed as a single topocluster. The typically three-pronged substructure of the top jets is not obviously visible in the sum of the jet images. However, there is a broader tail along the α_2 direction, compared to QCD initiated jets, which is likely due to the presence of highly-boosted top decay products in the top jets. The jet images typically contain between 40 to 130 jet constituents and there is considerable variation in the number of constituents between the different types of jets. For the jet images shown in figure 5.1, the average number of constituents is 84 for QCD jets, 91 for top jets and 74 for W jets.

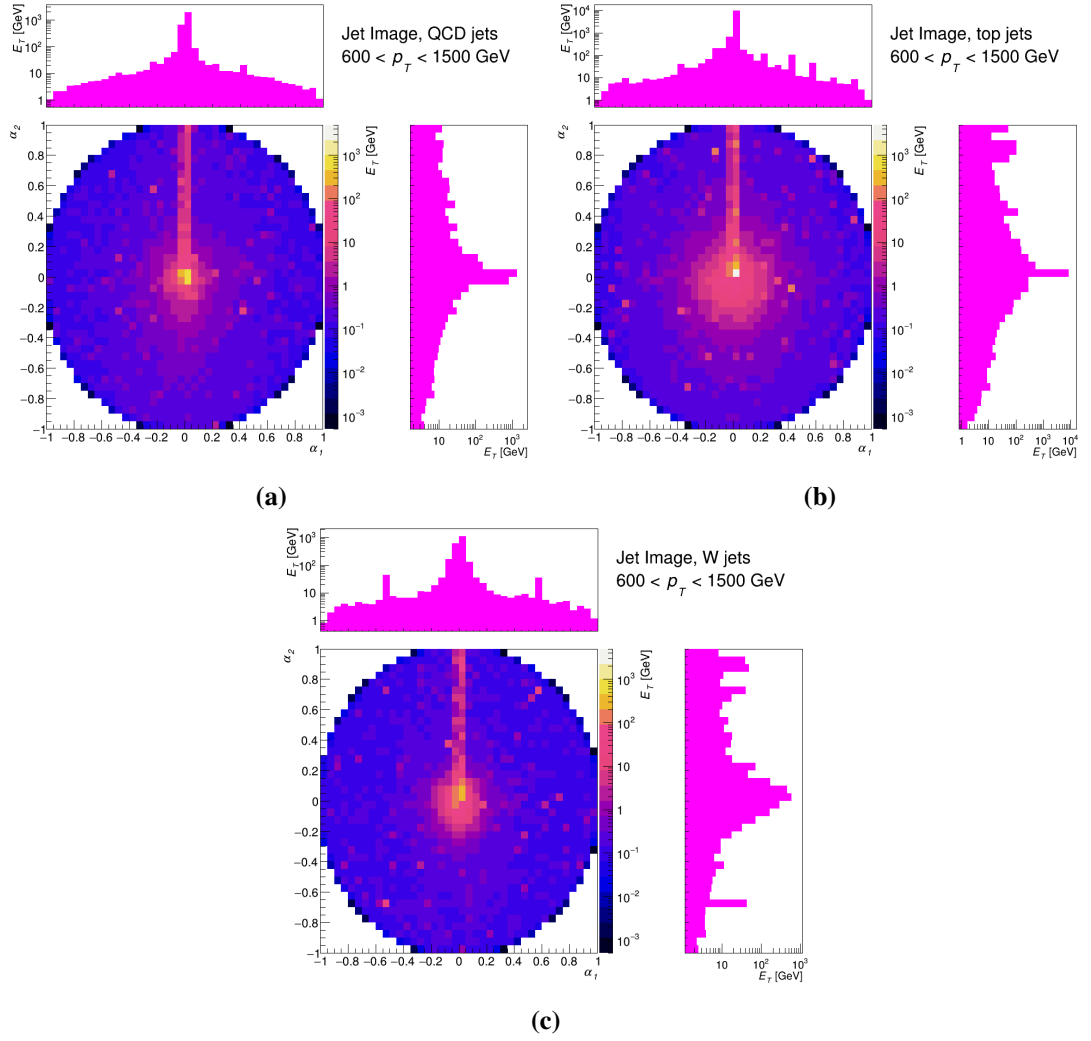


Figure 5.1: Average jet images for QCD (a), top quark (b) and W boson (c) initiated large- R jets. The z -axis shows the Lorentz-invariant transverse energy.

5.3.1 Convolutional Neural Networks

Convolutional neural networks (CNNs) are a class of Deep Neural Networks that have received much attention in recent years thanks to their extraordinary performance in image classification tasks. This makes CNNs a natural candidate for constructing jet-image based taggers.

A schematic of a Convolutional Neural Network is shown in figure 5.2. Convolutional neural networks are built up from a series of *layers* of different types that process the input data in a non-linear way. The two basic types of layers in a CNN are fully connected layers and convolution layers. The fully connected layers combine the outputs of the previous layer through a linear combination $z_j = \sum_i W_{ij}x_i$ followed by a non-linear activation function $f(z)$. The matrix W_{ij} is known as the weights matrix. The weights of each layer are updated during training.

Convolutional layers perform a convolution operation on the input vectors with an $n \times n$ kernel. Typically, the length of the input to the convolution layers is much greater than n . The $n \times n$ entries of the kernel are the trainable parameters, or weights, of the convolution layers. Convolution layers are able to process very large input vectors with a much smaller set of weights than fully connected layers, whose weight vectors must be the same length as the input. This is a very useful property, as fully connected neural networks become . Kernel convolution operations are commonly used to perform image processing operations such as edge detection filters. Convolutional layers are therefore excellent at extracting image-based information, such as shape and texture, so that it can be processed further through fully connected layers.

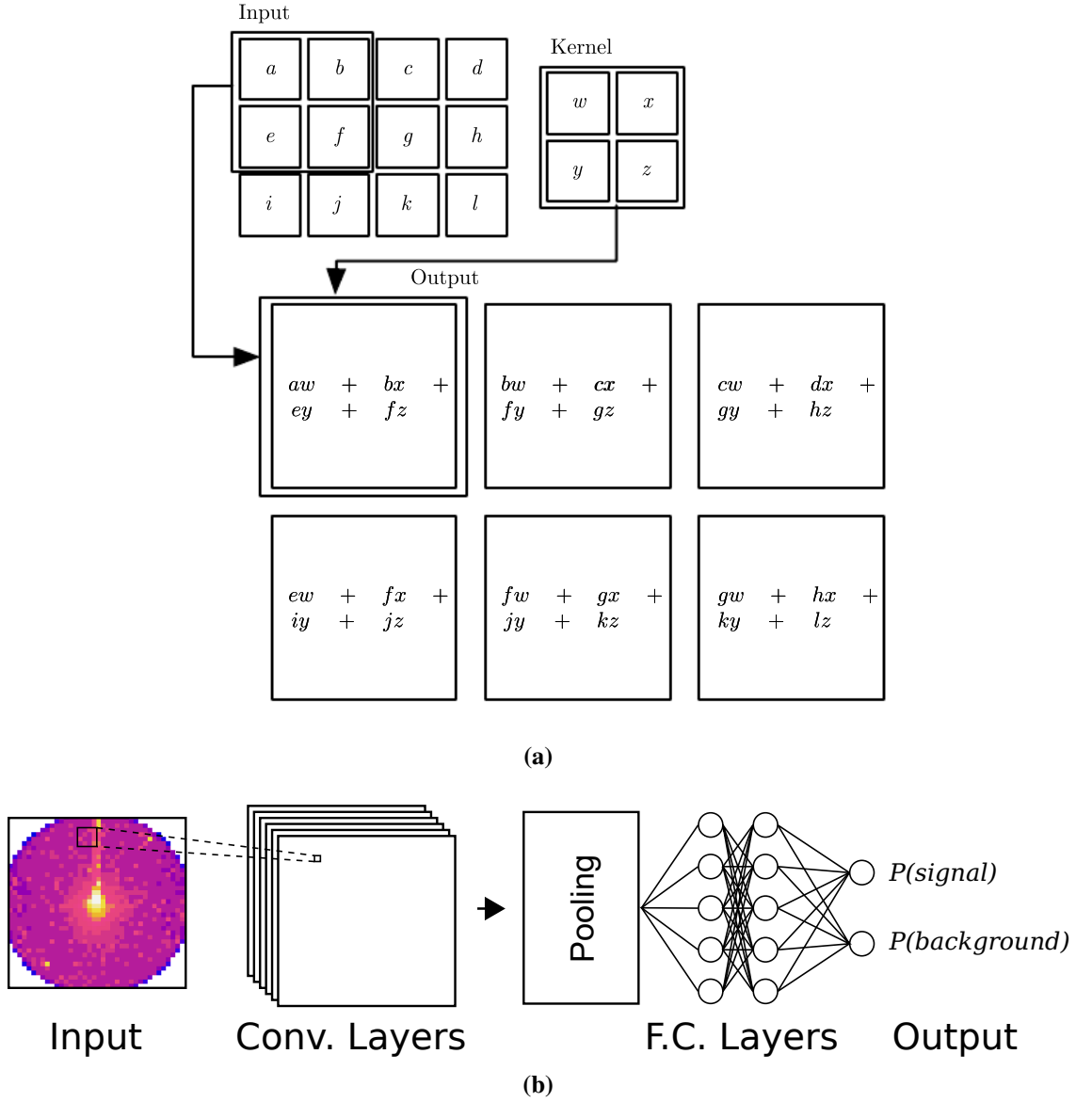


Figure 5.2: Diagrams illustrating the structure of a CNN. Presented in (a) is a figure from ref. [221] showing the operations performed by a kernel image convolution. The pixel values of the convolved image are linear combinations of the input pixels, eg. a, b, e, f , weighted by the kernel weights w, x, y, z . The main components of a typical CNN image classifier are shown in (b). The input image is convolved multiple times in the convolution layers. The output of the convolutions is pooled (using eg. Max pooling) and fed into a series of fully connected layers. The final output layer of the neural network consists of two nodes whose activation represents the classifier probabilities for the output classes, in this case signal and background.

5.3.1.1 CNN network architecture

A detailed description of the model architecture is provided in appendix C.2. The convolutional neural network architecture that we use for our studies is based on the LeNet-5 [236] model. This is a simple CNN model with a long history which has been optimised for low-resolution image inputs. Images with 40×40 pixels are input into the Neural Network. These are then processed by three 2-D convolution layers with 16 nodes each and a kernel size of 3×3 . The convolution outputs are modified by a *parametric rectified linear unit* (PReLU [237]) activation function. The PReLU activation function is a more recent development than the sigmoid functions used in the original LeNet-5 model that improves network performance by allowing its function parameters to change adaptively during training. After down-sampling the outputs of the convolution layers are fed into two fully connected layer with 128 and 2 output nodes respectively. The values of the two output nodes of the final fully connected layer are the network's final output probabilities for the signal and background classes. *Dropout* [238] layers are used to prevent overfitting. The total number of trainable parameters of the convolutional neural network is around 710,000.

The model is implemented using the Keras [219] high-level framework for constructing neural networks with TensorFlow [218] graph-based computation engine as its back-end. The TensorFlow library also enables training on Nvidia graphics cards using the CUDA [239] framework for hardware acceleration. To train our neural networks, we use arrays of up to 4 Nvidia GPUs of the model *Titan X* or *GeForce RTX 2080*. The learning rate of the neural networks is computed adaptively using the Adam [240] optimisation algorithm which is provided by the Keras package.

5.4 Using the Lund Jet Plane for W and top tagging

The high momentum transfer processes that give rise to top and W jets leave distinctive features in the Lund Jet Plane of the jet. There has been a considerable number of efforts exploring the possibility of using the Lund Jet Plane for boosted object tagging, either by applying cuts to the LJP variables themselves or by using the LJP as input to a number of different Machine Learning algorithms. The potential application of LJPs to jet tagging

was already suggested when LJPs were introduced in ref. [114]. This study showed a promising performance for a Lund Jet Plane based tagger with a recurrent neural network architecture. The performance of similar taggers using simulated ATLAS detector data is presented later in this chapter along with comparisons to image-based CNNs and other approaches.

5.4.1 Comparison of CNN performance for Jet Images and LJPs

The image-based convolutional neural network, whose architecture is described in section 5.3.1.1, is trained and evaluated on jet images and Lund Jet Planes. The jet images are constructed from calorimeter topoclusters, as described in section 5.3. An image resolution of 40×40 pixels is used. The LJPs are reconstructed from tracks contained within $\Delta R < 1.0$ of the jets. The LJP images are built by filling a 2-d histogram with 40×40 bins of the coordinates $\ln(R/\delta R)$ and $\ln(1/z)$ between $0 - 7$ and $\ln(2) - 7$. The sum of the LJP images is shown in figure 5.6. Note that for each LJP image, the pixel values are simply the number of entries in each bin which is typically 1.

The performance of the CNN tagger is evaluated by scoring the network performance on a statistically independent training dataset computing the ROC curve. A typical ROC curve for the CNN is shown in figure 5.3. As a combined measure of the performance of the neural network, the area under the curve (ROC AUC) is computed. To get and estimate for the uncertainty on the ROC AUC due to the limited statistics of the training data and possible biases in the choice of the , a technique called k -fold cross-evaluation is used. The cross-evaluation is performed by dividing the training data into k parts of equal size. The network is then re-trained and evaluated k times using one sub-sample for the testing dataset and the remaining $k - 1$ samples for training. The mean and standard deviation on the ROC curve can then be evaluated from the k folds.

The ROC AUC for the image-based CNN tagger using either Jet Images or Lund Jet Planes as inputs is shown in figure 5.4. A consistently higher score is achieved for the Lund Jet Planes across the full range of jet p_T than for the jet images. For both tagger inputs, the tagger performance improves with increasing p_T until about 1800 GeV where it starts to plateau. At very high p_T there is a larger difference in performance between the two types of input images. For the jet images, this may be a resolution effect. The

constituents of the jet become highly collimated at these energies, which causes a high concentration of entries at the centre of the image. The topology of the Lund Jet Plane images is not as affected by these effects as the z and ΔR kinematic variables of the emissions are plotted logarithmically in the LJP. It is also worth noting that, since the Jet Images are built from topoclusters and the LJPs are built from tracks, the constituents exhibit different resolution effects as a function of p_T . With greater p_T , topoclusters tend to become wider and less numerous as the showers in the calorimeter grow and eventually merge. Hence, it becomes more difficult to resolve topoclusters at very high p_T . This is not the case for the tracks, whose angular resolution is near constant as function of p_T . On the question of whether Jet Images or LJPs are the better inputs to CNN jet taggers one might also consider the relative complexity of the observables and the time it takes to construct them. Constructing the LJPs takes a little longer as it requires clustering the tracks within the jet using the Cambridge-Aachen algorithm. However, clustering N jet constituents using the FastJet [241] package takes $\mathcal{O}(N \ln N)$ time, so the time penalty incurred for constructing the LJPs grows relatively slowly with the number of jet constituents.

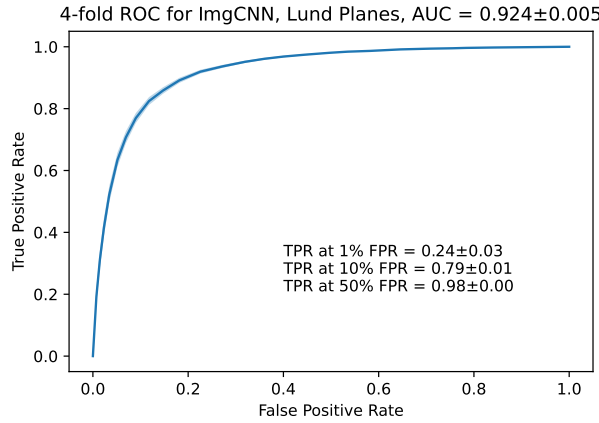


Figure 5.3: ROC curve for the CNN W tagger using Lund Jet Plane inputs for jets with p_T in the $[1800 - 2300]$ GeV range. The uncertainty on the curve is estimated using 4-fold cross-evaluation.

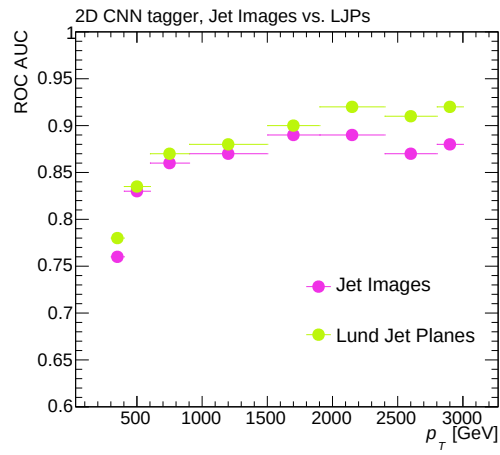


Figure 5.4: ROC AUC scores as a function of p_T for a CNN top tagger using jet images and Lund Jet Planes. Statistical uncertainties in the AUC from k -fold cross-validation are smaller than the size of the markers.

5.4.2 Recurrent Neural Networks

Recurrent neural networks (RNNs) are a class of neural networks that is particularly suited for extracting information from sequential data. The range of applications of RNNs is very broad and includes topics such as machine translation, anomaly detection in time series data or speech recognition. These neural networks contain a memory, or hidden state vector which can retain information as the input vectors are processed in sequence. This allows the network to extract complex correlations between elements of the input, regardless of where in the sequence they are located. One of the frequently used types of recurrent neural networks is the Long Short Term Memory (LSTM) network [242]. The LSTM is built from a collection of non-linear vector functions of the form

$$f = \sigma(Wx + Uh + b) \quad (5.4)$$

where x is the input vector, h is the hidden state vector and σ is a non-linear activation function, such as a sigmoid or hyperbolic tangent. The matrices U, V and the bias vector b contain the trainable weights of the network. Each operation f is referred to as a "gate". By chaining multiple gates together and combining their outputs, the flow of information in and out of the hidden state h can be controlled by the network weights. The names of the four gates that make up the LSTM, the *input gate*, *output gate*, *forget gate* and the *cell state*, reflect the function of each gate in controlling the information flow through the network. A diagram showing the structure of an LSTM cell is presented in figure 5.5.

Gated Recurrent Units (GRUs) [243] are another type of recurrent neural networks, similar to LSTMs, but with fewer trainable parameters. Notably, the GRU does not contain an output gate that produces an output vector y . The output of a GRU cell is instead identical to its hidden vector h .

RNNs and other sequence-based neural networks have found a wide range of applications in jet physics. At ATLAS, jet flavour taggers have been constructed using LSTMs [244]. These taggers use unordered sequences of track features as inputs and take advantage of the LSTMs ability to extract correlations between the elements of sequences of arbitrary length. At CMS, top and W taggers have been developed along a similar princi-

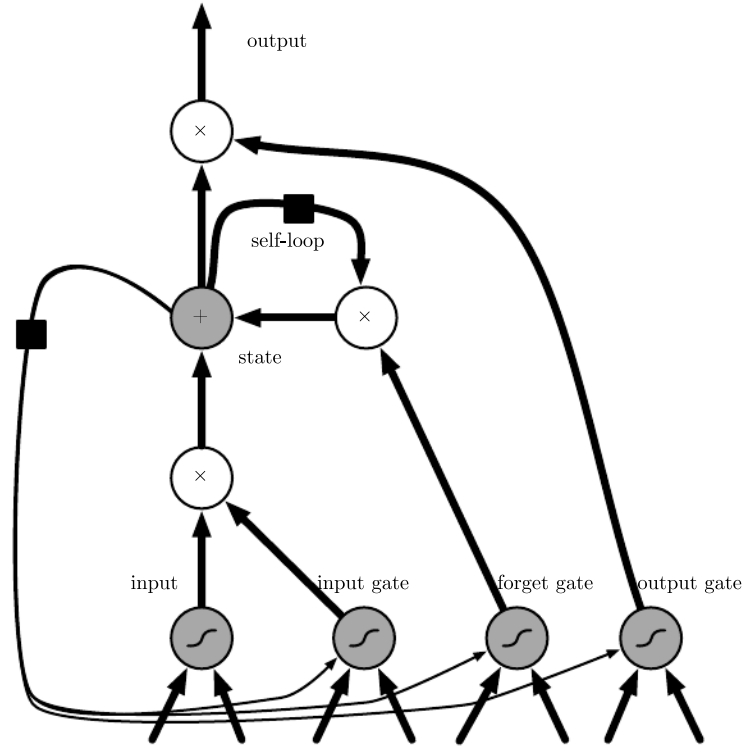


Figure 5.5: A diagram showing the structure of an LSTM cell. The cell input and the state vector of the cell are combined in a series of cells: the input gate, the forget gate and the output gate. Each cell applies a sigmoid activation to produce a non-linear output. The state vector of the cell is passed from one step of the sequence to the next. The forget gate controls the information to be retained in the state vector, while the input and output gates modify the input and output vectors. From ref. [221].

ple using 1-dimensional convolutional neural networks which take sequences of jet constituent kinematics as inputs [245, 231]. As the 1d-CNN architecture is more sensitive to the order of input sequence, the constituents are sorted by p_T .

The Lund Jet Plane can also be used as input to sequence-based DNN taggers. The primary LJP is constructed from a sequence of kinematic variables resulting from the Cambridge-Aachen declustering of the jet constituents. Therefore it may also be conceived of as a sequence of LJP coordinates, which is ordered by the angular coordinate ΔR due to the angular-ordered property of the CA-algorithm. The following section discusses some results of RNN sequence-based W and top taggers, using the Lund Jet Plane, and compares their performance to the CNN image-based approach.

5.4.3 Higher-order LJPs

The primary and secondary Lund Jet Planes are shown figure 5.6 for top, W and QCD initiated jets. The secondary Lund Jet Plane is computed by following the softer branch in the first step of the CA-declustering and then the harder branch in every subsequent step. In the case of W and top jets, the first step of the declustering is likely to correspond to the hard splitting which gives rise to the two- or three pronged substructure of these jets. The secondary Lund Jet Plane might therefore contain additional information about the decay of the lower p_T quark(s) from W or top decay that is not visible in the primary Lund Jet Plane. Our convolutional neural network model allows for the addition of "colour" channels, or additional input images with the same dimensions as the primary Lund Jet Plane, without increasing the number of trainable parameters. It is therefore a simple task to input both the primary and secondary LJPs into the CNN tagger. The LSTM tagger can take an arbitrary number of sequences in its input layer, however this does increase the number of trainable parameters of the neural network.

A comparison of ROC AUC scores for a CNN and LSTM top tagger with and without inclusion of the secondary Lund Jet Plane is shown in figure 5.7. While the LSTM tagger performs significantly better than the CNN tagger across the whole range of p_T , there is little performance to be gained by including the secondary LJP as an input to either tagger. Including the secondary plane only changes the AUC score by up to 0.02. The biggest improvements to the AUC score are at $p_T < 2.3\text{TeV}$ for the CNN. At $p_T > 2.3\text{TeV}$ the CNN performs worse with inclusion of the secondary plane than without. These results suggest that there is only very little information in the secondary LJP to help the neural networks distinguish between top jets and QCD background. As can be seen in figure 5.6, the secondary LJPs for signal and background jets are visually quite similar. It can be concluded that the relevant features of the input vectors, for either the CNN or the LSTM, are unchanged by the inclusion of the secondary LJP.

5.4.4 RNN tagger comparisons

A collection of different RNN taggers are trained and evaluated on the ATLAS Monte Carlo samples. All taggers use sequences of three LJP variables z , ΔR and k_T as inputs. The length of the input vectors is fixed at 30. Since this is typically larger than the

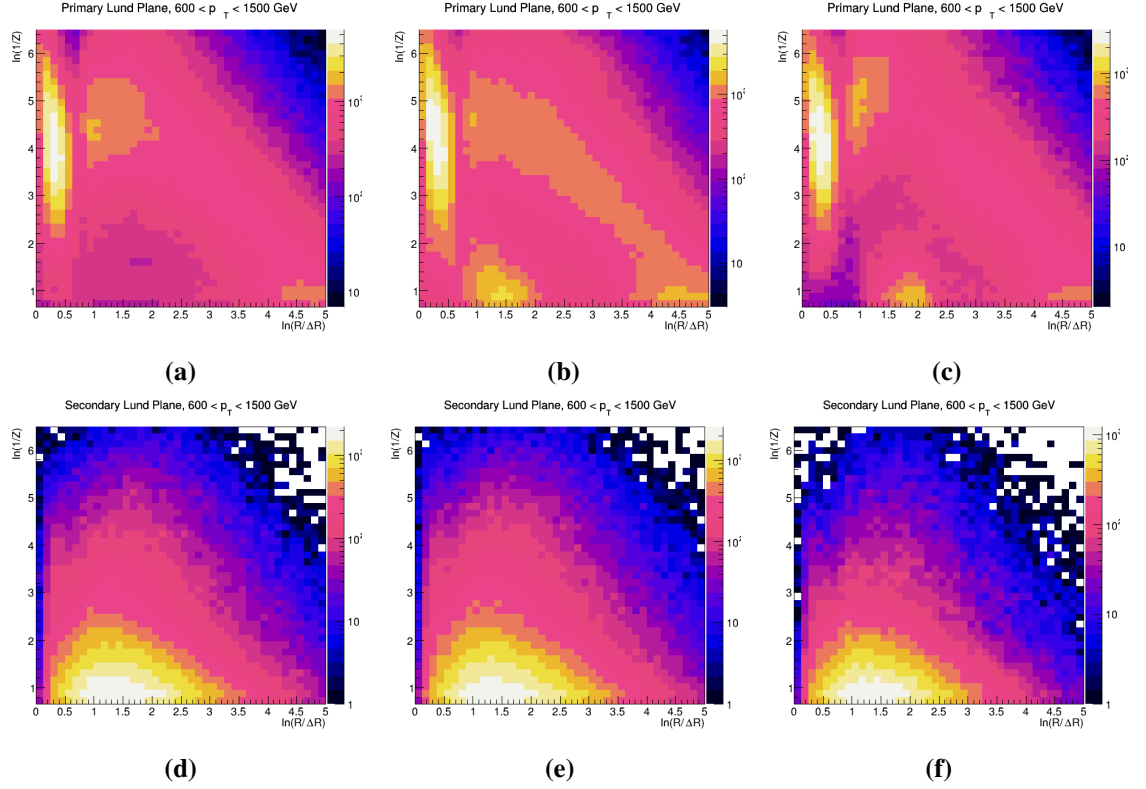


Figure 5.6: Primary (a,b,c) and Secondary (d,e,f) Lund Jet Planes for QCD, top and W initiated large- R jets.

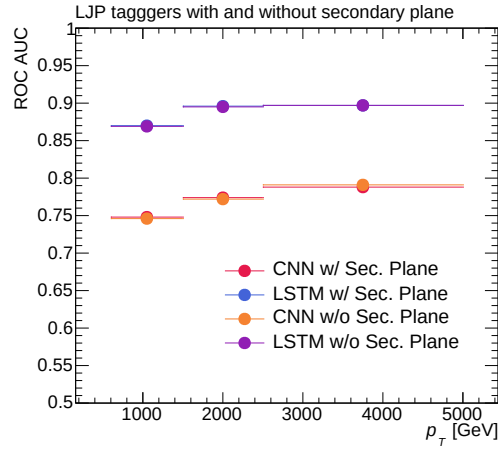


Figure 5.7: ROC AUC scores as a function of p_T for the CNN top tagger using either the primary LJP or the sum of the secondary and primary plane. Statistical uncertainties in the AUC from k -fold cross-validation are smaller than the size of the markers.

number of emissions on the LJP, zeros are added to the LJP coordinate sequences to fill the vectors. A description of each tagger is given below. A more detailed layer-by-layer specification of the model architecture is provided in appendix C.2.

LSTM The LSTM network consists of a single LSTM layer with output size 100, followed by a fully-connected layer. The output of the fully connected layer, which is also the network output, is fed into a sigmoid activation function. The total number of trainable parameters is 41,802.

GRU The GRU network consists of a single GRU layer with output size 100, followed by a fully-connected layer and a sigmoid activation function. Hence, with the exception of the GRU cells, its structure is identical to the LSTM network. Due to the relative simplicity of the GRU cells' structure, the number of trainable parameters is smaller than for the LSTM at 31,402.

1D-CNN A 1-dimensional convolutional neural network is also trained on the sequential LJP input. The architecture is based on the image-CNN, with 3 layers of convolutions followed by a max-pooling layer and two fully-connected layers. ReLu activation functions are applied to the outputs of all layers, except the final fully-connected layer which is fed into a sigmoid function. This CNN architecture has 181,502 trainable parameters in total.

2D-CNN A alternative implementation of a CNN with LJP sequences as inputs. This network uses 2-d convolution layers instead of 1-d convolutions. The width of the convolution kernel is set to 1, so that the network only calculates the convolutions along the length of the input vectors.

A comparison of ROC AUC scores for the RNN and 1D-CNN taggers is shown in Figure 5.8. The Neural networks are trained and evaluated on samples of 64,000 jets containing 50% signal and 50% background. Both to-tagging and W -tagging performance is evaluated. The GRU and LSTM achieve a higher AUC score than the sequence CNNs across the whole p_T range. The performance of the GRU and the LSTM is very similar with slightly higher scores for the GRU.

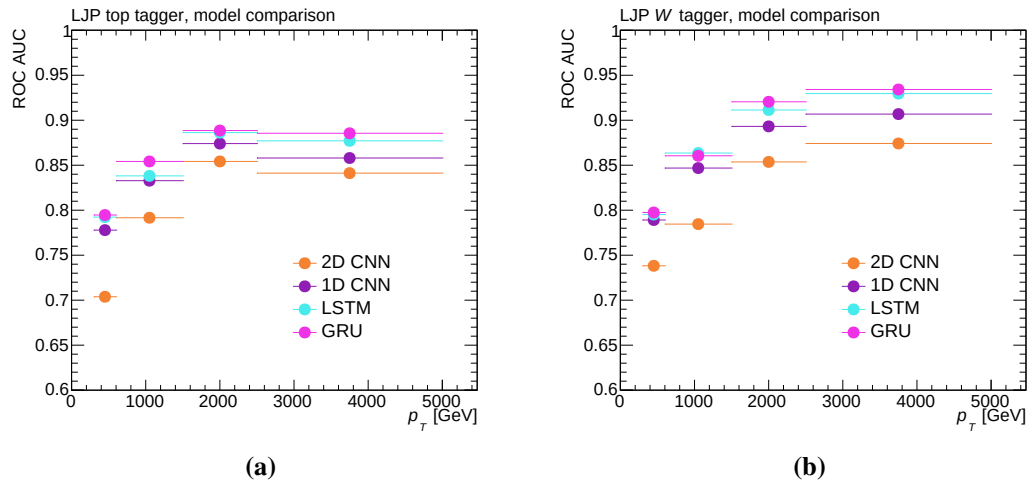


Figure 5.8: ROC AUC scores as a function of p_T for various (a) top and (b) W taggers using sequences as inputs. Statistical uncertainties in the AUC from k -fold cross-validation are smaller than the size of the markers.

5.4.5 Using UFOs as inputs to the LJP

So far, only charged particle tracks have been used in the construction of our LJP-based taggers. One potential way to improve the tagger performance is to include neutral particles in the construction of the LJP. Jets originating from top or W decays tend to contain a lower proportion of neutral hadrons than gluon initiated jets. This additional neutral radiation leads to more complex clustering trees and longer sequences of LJP coordinates, which could improve the power of the RNN taggers.

Unified Flow Objects, described in detail in section 1.3.5, are the latest development of combined track-cluster jet inputs at ATLAS. They combine the high spatial resolution of the ID with the energy measurements of the calorimeter, while also including neutral particle signatures through the inclusion of calorimeter topoclusters without associated tracks. Two pileup-mitigation procedures called Constituent Subtraction [88] and Soft-Killer [89] are applied to the UFOs before the LJPs are constructed. These cleaned objects, called CSSK-UFOs, are then gathered within a radius of 1.0 of the calibrated anti- k_T $R = 1.0$ LC-topocluster jets and re-clustered using the CA-algorithm to construct the LJP.

A comparison of ROC AUC scores for the GRU top and W taggers using either tracks or UFOs as inputs is shown in Figure 5.9. A significant improvement in tagging performance is observed for the UFOs, particularly for jets with lower p_T . The tagger is thus able to benefit from the additional information provided by the UFOs about the neutral jet constituents. The share of neutral UFOs is largest for $p_T < 500\text{ GeV}$, where they make up nearly 40% of the jet constituents. Around $p_T \approx 1 - 1.5\text{ TeV}$ the fraction of neutral UFOs reaches a minimum, which may be why the improvement achieved for the top tagger is smaller in the $500 < p_T < 1500\text{ GeV}$ bin than in the two adjacent bins. The distribution of charged, neutral and combined UFOs inside the jets for different p_T is shown in appendix C.1.

Since these studies have been conducted, new jet calibrations have become available, including for large- R jets constructed from UFOs. This jet definition may be more suitable for constructing UFO LJPs, as they can simply be built from the set of jet con-

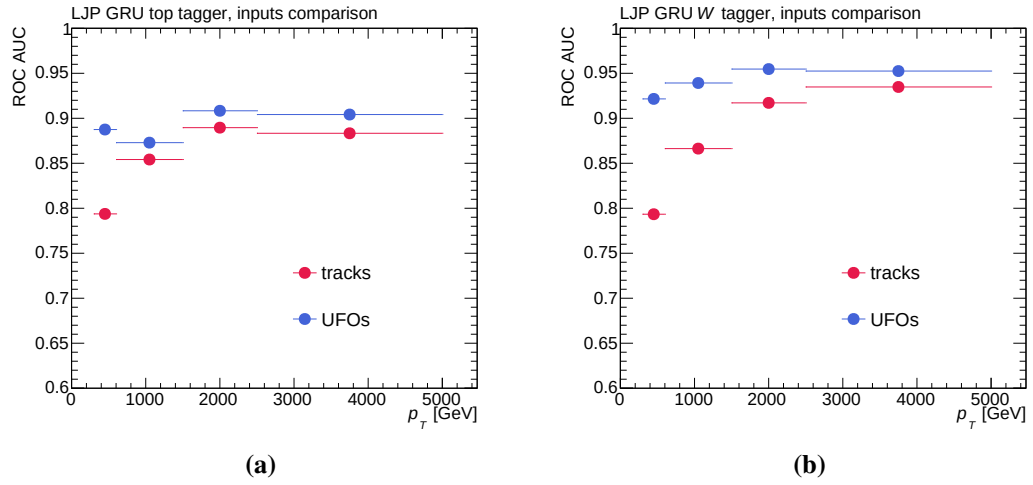


Figure 5.9: ROC AUC scores for the GRU (a) top and (b) W taggers using either tracks or UFOs to construct the Lund Jet Plane. Statistical uncertainties in the AUC from k -fold cross-validation are smaller than the size of the markers.

stituents. A follow-up study using the new jet definition could potentially yield an even better tagging performance than the results presented here.

5.4.6 Tagger performance vs. p_T at fixed working points

In most physics analyses, it is desirable to operate the tagger at a fixed working point with a constant signal efficiency across a wide range of jet p_T . In this way, the purity of the signal can be estimated easily even for p_T inclusive measurements. This section shows how working points with a fixed signal efficiency can be defined for our DNN taggers and compares the performance of the taggers at those points.

The working point of the DNN tagger is fixed by choosing a cut-off value on the score that is output by the neural network classifiers. Each cut-off value yields a different value for the signal- and background-efficiency on the ROC-curve between 0 and 1. By binning the testing dataset in p_T and producing a ROC curve in each p_T bin, we find the cut-off value on the classifier score in each p_T -bin that yields the desired signal efficiency. The relationship between the cut-off values and p_T is shown in figure 5.10 for the GRU W -tagger at the target signal efficiencies of 50% and 80%. A smooth function is fitted to the measured cut-off values. An exponential function of the form $a + be^{cx}$ is chosen for the fit function. The χ^2/NDF goodness-of-fit value is less than 0.1 for the functional fits at both working points.

Figure 5.11 shows the background rejection and signal efficiency as a function of jet p_T for the various RNN W taggers. These graphs have been obtained by producing histograms of p_T for tagged and untagged signal and background events at the fixed working points. By computing ratios of the distributions according to the definitions of the signal efficiency and background rejection given in section 5.2.1, one obtains p_T distributions of the efficiencies. The graphs in figure 5.11 also show a comparison to the 3-variable cuts-based ATLAS W -tagger whose definition can be found in section 2.2.4. At both working points we see that the GRU and LSTM taggers produce the highest background rejection out of the DNN taggers across the whole p_T range. These two taggers achieve particularly good results at $p_T > 1500\text{GeV}$ compared to the 3-variable tagger. The 1D-CNN tagger achieves the worst overall performance out of the sequence-based DNN taggers. For the 1D-CNN, the p_T dependent working point is determined using two different types of functional fits: an exponential function and a 6th order polynomial. As can be seen in figure 5.11c, the polynomial fit provides a flatter signal efficiency distribution than the

exponential, making it a more appropriate choice when fixing the working point of this particular tagger.

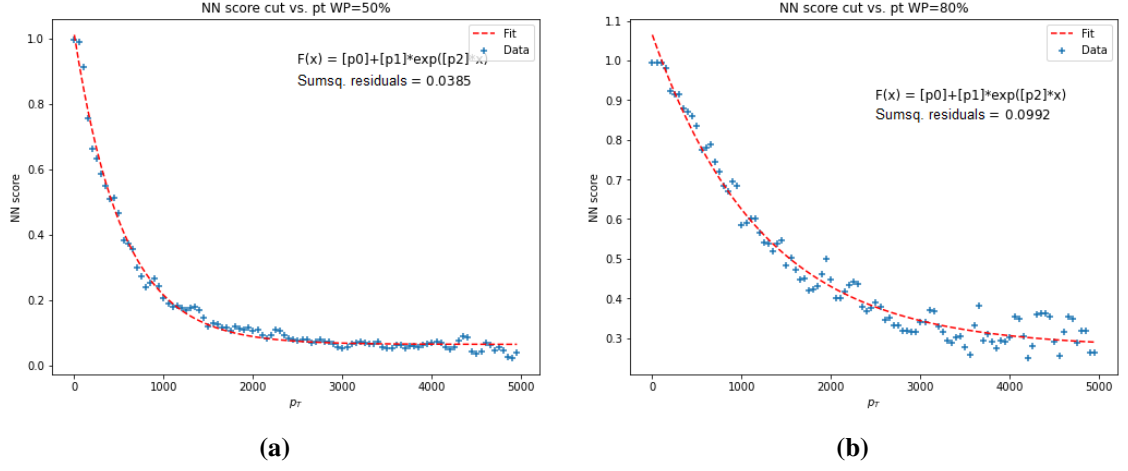


Figure 5.10: Exponential fits of the output variable of the GRU tagger as function of p_T for signal efficiency working points at 50% (a) and 80% (b).

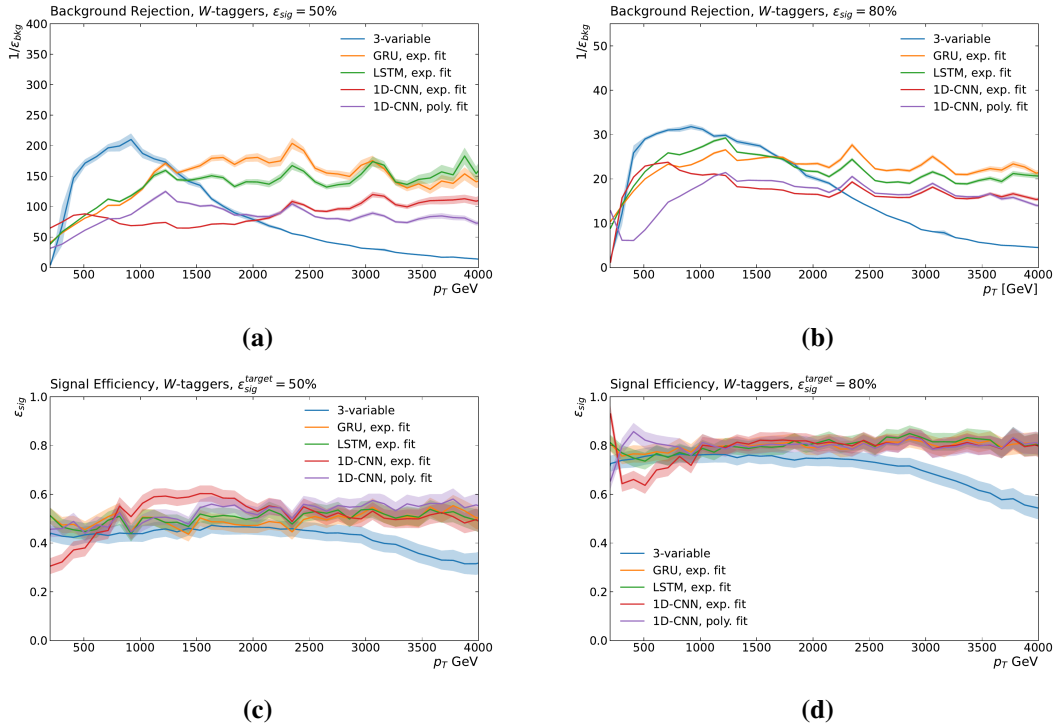


Figure 5.11: Background rejection (a,b) and signal efficiencies (c,d) for the sequence-based DNN W taggers and the 3-variable cuts-based tagger. DNN working points have been fitted to yield a constant signal efficiency of either 50% (a,c) or 80% (b,d). Statistical uncertainties are indicated by the shaded bands.

5.4.7 Graph Neural Networks

5.4.7.1 EdgeConv and GNNs from point clouds

Graph Neural Networks (GNN) are neural networks which take vector representations of graph as inputs and whose structure mimics the structure of the input graphs. Most implementations of GNNs that have been used for jet tagging are built using the *EdgeConv* [246] layer. This type of GNN was originally developed to process point-clouds, collections of points in space such as those that are acquired by 3-d scanning technologies like Lidar. The EdgeConv layers are able to operate on the point clouds directly without first needing to reconstruct 3-d meshes or 2-d images. The point clouds are represented as directed graphs representing the local neighbourhood of each point, rather than an unordered collection of points in space. The EdgeConv layers are able to exploit local symmetries in the geometry of the point clouds by applying convolution operations along the edges of the graph. Given an input graph $\mathcal{G} = (\mathcal{E}, \mathcal{V})$ with $\mathcal{V}$ vertices and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ edges, the EdgeConv layers perform an operation of the form:

$$x'_i = \sum_{j:(i,j) \in \mathcal{E}} \theta_m x_j, \quad (5.5)$$

where θ_m denotes a collection of m different convolution filters that act on the point x_i and its neighbours x_j . The parameters of the convolution filters θ_m constitute the weights of the EdgeConv layers, and are trained using back-propagation like the kernel entries of a CNN layer. The output of an EdgeConv layer is a new set of points. A neighbourhood function can then be applied to the output points so that a new graph is produced. As a graph is propagated through a series of EdgeConv layers, both the vertices and edges of the graph are updated according to the trainable parameters of the network. This feature makes EdgeConv a more versatile architecture than earlier types of GNNs that tended to keep the shape of the input graph fixed.

ParticleNet [247] is an application of the EdgeConv GNN to jet tagging. The jet is represented as an unordered set of particles which are combined into a graph using a k -nearest-neighbour algorithm in detector space. As of 2022, the ParticleNet tagger has

already found an application in a search for resonances of the type $X \rightarrow HY \rightarrow b\bar{b}b\bar{b}$ at CMS [248], where it was used to tag large- R jets containing $b\bar{b}$ pairs.

5.4.7.2 The Lund Jet Plane as a graph

By following every branch of the Cambridge-Aachen declustering sequence until the individual input particles are recovered, it is possible to represent the declustering as a graph. This graph is a binary tree, where each node represents a step in the CA algorithm where two proto-jets are combined into one. The coordinates of the LJP, which can be calculated for every step of the CA-declustering can be treated as node features for a graph neural network. An example of a graph representation of the CA-declustering of a single jet is shown in figure 5.12. As can be seen in this graph, there are typically many additional splittings outside of the sequence from which the primary LJP is constructed. Typically, the number of additional splittings is larger for QCD background jets than for W or top jets.

5.4.7.3 GNNs on Lund Jet Planes: LundNet

The ParticleNet tagger only uses a simple k -nearest-neighbour algorithm to construct a local neighbourhood graph from the particle coordinates. It is however possible to replace this algorithm by any other combinatorial algorithm that uses the kinematics of the jet constituents. A jet-physics aware alternative is to use a jet algorithm. This is precisely the innovation provided by LundNet [217, 249] which uses a Lund Jet Plane graph, created using the CA-algorithm, to construct the input graph.

We construct the LJP graphs from UFOs within $R = 1.0$ LC-topocluster jets. Six kinematic variables are used to construct the node features. The LJP graph is input into a series of EdgeConv layers. The structure of the graph is kept fixed between each layer. A total of six EdgeConv layers with between 32 and 128 filters each transform the node features before they are fed into a pooling layer and a fully connected layer with 256 nodes. A ReLu activation function is applied to the output of all layers except the output layer. A detailed description of the architecture of the LundNet model, including layer parameters, is provided by the PyTorch class definition shown in figure C.6. The model is trained on a total of 1.1 million jets, of which 800,000 are dijet background and 300,000

W signal jets. As for the CNNs and RNNs, the learning rate is optimised dynamically using the Adam algorithm.

ROC curves for the LundNet W -tagger compared to the LSTM RNN tagger and the 1D-CNN are shown in figure 5.13. The LundNet outperforms the LSTM RNN at very high ϵ_{sig} and achieves a marginally higher AUC, however, at lower ϵ_{sig} it rejects significantly fewer background events than the RNN.

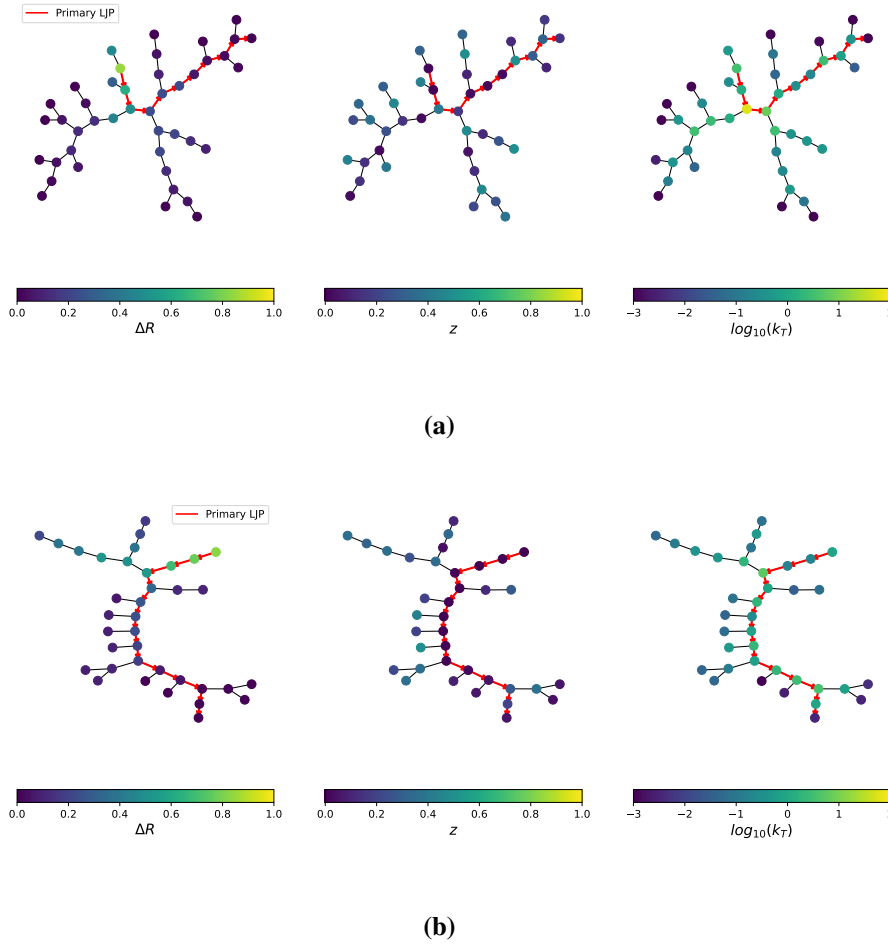


Figure 5.12: A binary tree constructed from the Cambridge-Aachen clustering sequence of a single W jet (a) and a single QCD jet (b) from the training samples. The colour of the nodes in each column represent the values of the LJP coordinates ΔR , z and k_T . The path following the node with greater p_T at each branching, from which the primary LJP is constructed, is indicated by the red arrows.

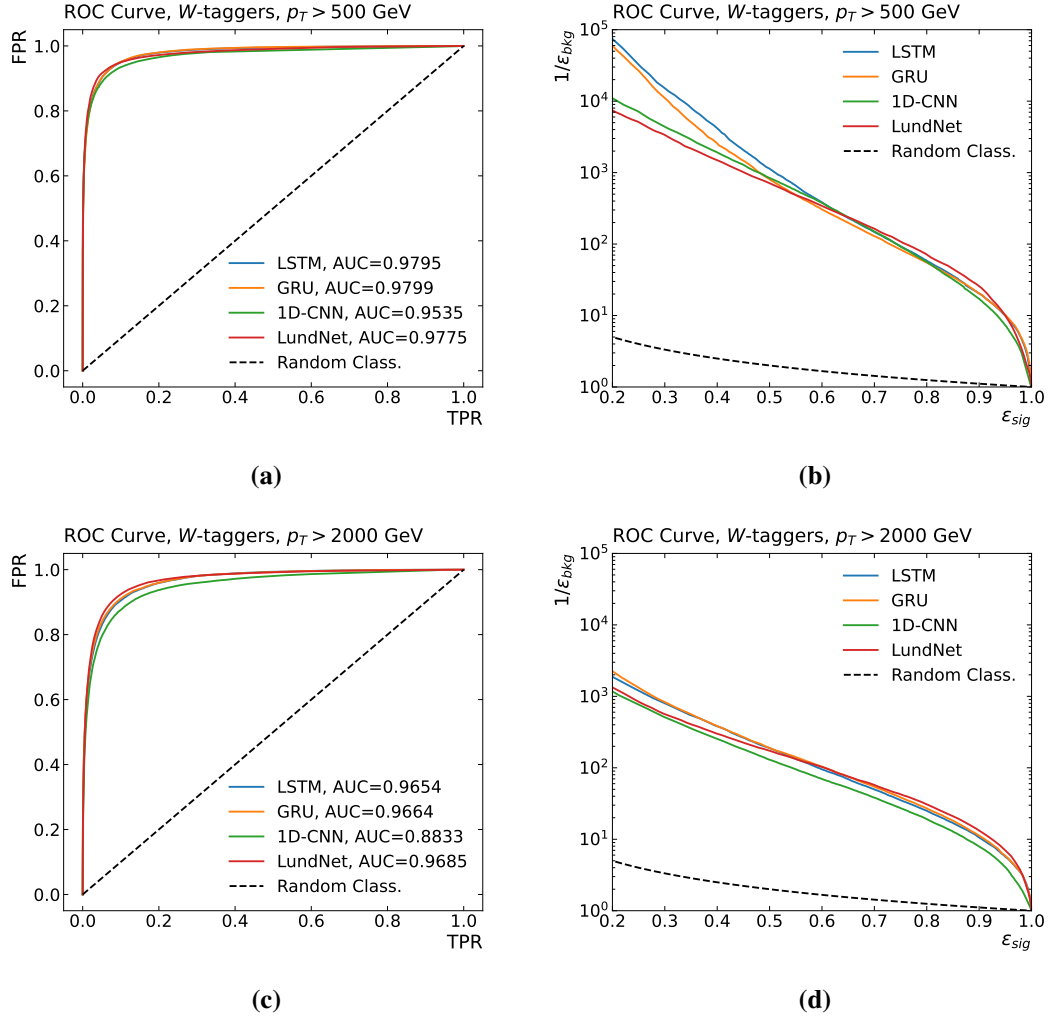


Figure 5.13: ROC curves for the W GNN taggers compared to the RNN taggers for jets with $p_T > 500$ GeV (a,b) and $p_T > 2000$ GeV (c,d). The TPRs vs. FPR with areas under the curve is shown in (a,c). The background rejection in the high-TPR region is shown in (b,d). The dashed line shows ROC curves for a random classifier.

5.5 Conclusion

A large collection of Deep-Learning W and top jet taggers using the Lund Jet Plane has been presented in this chapter. The studies presented here show that the LJP is a versatile representation of jet substructure that can be used to construct a great variety of novel taggers. By parametrising the LJP as an image, sequence or tree, we can employ a wide range of state-of-the-art neural network architectures: CNNs, LSTMs, GRUs and GNNs. More complex neural network architectures, such as the GNN, that are able to take advantage of the full clustering tree perform better than simpler ones. The LJP's have been shown to yield a neural network performance that is equivalent to, or even better than other specialised representations of jets. For example, we show that LJP images outperform $\eta - \phi$ images of jet constituents as CNN inputs, particularly at large jet p_T . Other studies have shown (see ref. [217]) that a LJP-based GNN performs at a similar level to ParticleNet, one of the best Deep-Learning taggers currently available. We have also shown how ATLAS' latest type of jet input objects, UFOs, can improve the the performance of LJP-based neural network taggers. This is an encouraging result for the development of a future LJP-based top or W tagger for ATLAS analyses.

Chapter 6

Thesis Conclusion

This thesis has presented a wide-ranging exploration of the substructure of jets, using the Lund Jet Plane observable as its main tool. A diverse collection of studies involving this novel jet observable were presented. These studies investigate the properties of the LJP in great detail and showcase its wide range of applications. A secondary focus of this thesis has been the development of new W and top jet taggers. As the LJP reveals new information about the inner structure of jets, it can also be powerful tools in the identification of jets initiated by W bosons and top quarks.

In chapter 3 the first-ever measurement of the LJP in LHC data was introduced. Specific contributions and cross-checks on this pioneering measurement were highlighted, including studies of the composition of the Lund Jet Plane at the particle level and an estimation of effects of faulty calorimeter modules on the measurement. The measurement was carried out with a high degree of precision thanks to the great amount of dijet events available in the new ATLAS Run 2 data. The measurement found that none of the available Monte Carlo generators were able to accurately model the whole of the LJP [115]. Analytical calculations of the LJP published after the measurement were found to be in good agreement with the ATLAS data [151].

A brand new measurement of the Lund Jet Plane using $R = 1.0$ jets from semi-leptonic $t\bar{t}$ decays was presented in chapter 4. This is the first measurement of the LJP for jet topologies involving highly boosted objects, namely W bosons and top quarks. Previous measurements, such as the pioneering ATLAS measurement in chapter 3 or the ALICE measurement in minimum-bias events [152], have only studied small-radius

jets initiated by light quarks and gluons. The new measurement was carried out using the ATLAS Run 2 data at $\sqrt{s} = 13\text{TeV}$. Measured distributions of the detector-level Lund Jet Plane were unfolded to the particle level using Iterative Bayesian Unfolding and compared to a range of Monte Carlo event generators, each using different configurations for the modelling of matrix elements, parton showers and hadronisation. Specifically, the MC generator configurations considered were: Powheg + Pythia 8, Powheg + Herwig 7, MadGraph5_aMC@NLO + Pythia 8 and Sherpa. The number of unfolding iterations was optimised to minimise the impact of the unfolding procedure on the measurement uncertainty. The total systematic uncertainty on the measurement was estimated at around 20 – 30% for top jets and 40 – 60% in the centre of the LJP (around $0.5 \lesssim \ln(R/\Delta R) \lesssim 3.5, \ln(1/z) \lesssim 3.5$) is found to be , while for W jets it is . Differences between the unfolded measurement and the different generators at the particle level were found to be within the estimated uncertainties across most of the LJP. However, considerable differences in the Lund Jet Plane distributions were observed for the different generator configurations. These discrepancies are thought to be due to the different approaches used by the generators in modelling parton showers, hadronisation and matching between matrix elements and parton showers. As the MC generators are incompatible with the measurement in some regions of the Lund Jet Plane, the measurement could help improve MC modelling in the future by providing a basis for new generator tunes, as had been done for the measurement in dijets [149].

This thesis has also contributed to the development and evaluation of novel W and top jet taggers. In chapter 2, the performance of current W and top jet taggers is evaluated by measuring signal efficiencies in ATLAS Run 2 data and Monte Carlo. The measured tagging efficiencies were used to calculate data/MC scale-factors which are part of the published recommendations for ATLAS physicists that want to make use of the taggers in Run 2 analyses. During the course of these studies, changes were made to the process by which the tagger performance in data was evaluated. These changes included updated $R = 1.0$ jet mass calibrations and new truth-labelling conventions making use of hadron-level kinematics. Updated results for W and top tagging performance in data were published in ref. [117]. Thanks to the new data and updated recommendations, modelling

uncertainties on the efficiency scale factors were reduced compared to the results that had been published in 2018 [116].

In chapter 5, a collection of completely new W and top jet taggers were introduced. These taggers make use of the technological innovations provided by Deep Neural network architectures and combine these with the many possible representations of the Lund Jet Plane. A range of popular deep neural network architectures, namely CNNs, LSTMs, GRUs and GNNs, were used in the construction of the taggers. Representations of the Lund Jet Plane as an image, sequence or tree were constructed to take advantage of the different neural network architectures. The Lund Jet Plane was successfully used to construct deep learning taggers, giving a neural network performance equivalent to, or better than simpler representations of jets such as $\eta - \phi$ jet images. It was also found that the tagger performance could be further improved by including neutral particle signatures in the construction of the Lund Jet Planes, using a novel class of jet inputs called UFO. Lastly, it was also found that more complex and recently developed types of neural network, such as the GNN, performed better than simpler ones. The GNN was able to use the full clustering tree from which the LJP is constructed to identify W and top jets.

As of April 2023, the Lund Jet Plane remains a popular research subject. In particular, there have been numerous recent publications that have made use of the LJP for various types of jet tagging, including quark-gluon [249], $h \rightarrow b\bar{b}$ [250] and b tagging [251]. All of these studies also make use of Machine Learning methods to process the LJPs. New potential areas of application are also being found, notably it has been proposed that measurements of the LJP can be used to place limits on BSM physics. *Dark sector showers*, in which parton-like dark matter particles shower and decay to SM partons, could produce visible signatures in the LJP [252]. The LJP has also recently been measured by the CMS collaboration for jets with $R = 0.4$ and $R = 0.8$ in events with at least one hadronic jet with $p_T > 700\text{GeV}$ [253]. Like the ATLAS measurement in dijets, this measurement was unfolded to the particle level and compared to state-of-the-art Monte Carlo and analytical predictions.

In conclusion, the Lund Jet Plane is an exceptionally versatile tool in the exploration the properties of jets. This thesis has presented new insights about jets initiated by W bosons and top quarks thanks to this new observable. Thanks to its many applications for jet substructure measurements and tagging, it is likely to remain an important tool for LHC jet physicists in Run 3 and beyond.

Appendix A

Performance of Run 2 W taggers

Presented in this appendix is a comparison of the performance of the W DNN tagger and the cuts-based 3-variable W tagger which were studied in chapter 2. Since the 3-variable tagger achieved a higher background rejection than the DNN tagger in Pythia8 MC at the 50% target signal efficiency, it became the recommended W tagger in the ATLAS jet-tagging subgroup. The 3-variable cuts-based W tagger was studied in detail alongside the recommended top tagger (top DNN) in ref. [117].

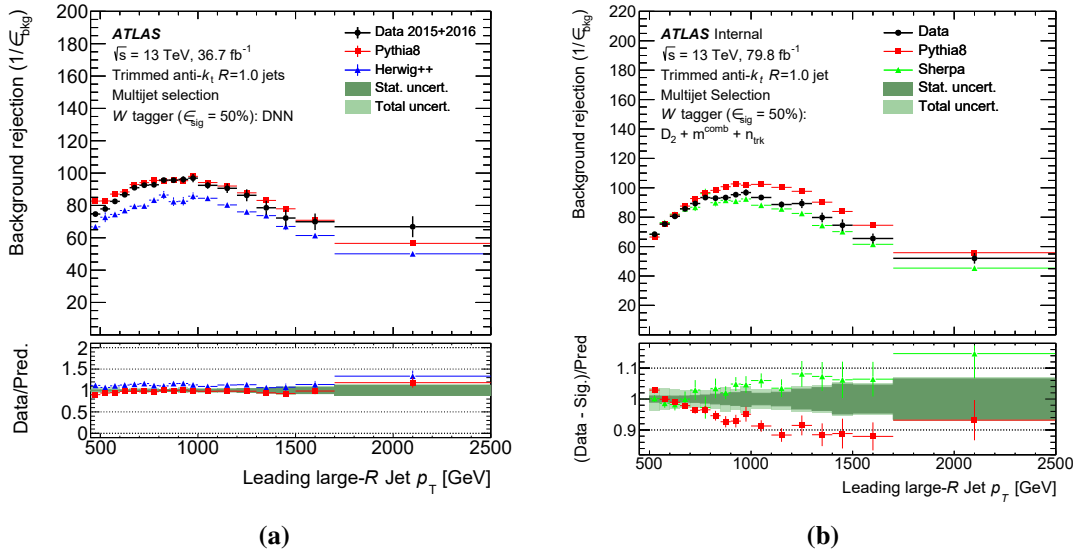


Figure A.1: Background rejection vs. large- R jet p_T for the DNN (a) and 3-variable (b) W taggers. The result for the W DNN tagger was obtained using 2015-2016 data and is taken from [116], while the 3-variable tagger result was evaluated from the larger 2015-2017 dataset and is taken from [117].

Appendix B

Measurement of the Lund Jet Plane in $t\bar{t}$ events

B.1 Complete systematic uncertainties on the LJP

The following appendix contains plots of the complete set of detector systematic uncertainties in the Lund Jet Plane in $t\bar{t}$ measurement described in chapter 4. The uncertainties are shown for slices through the Lund Jet Plane and grouped by the objects they are associated with: jets, missing- E_T , tracks and leptons. All uncertainties are evaluated for normalised LJPs in the top-jet and W -jet selections.

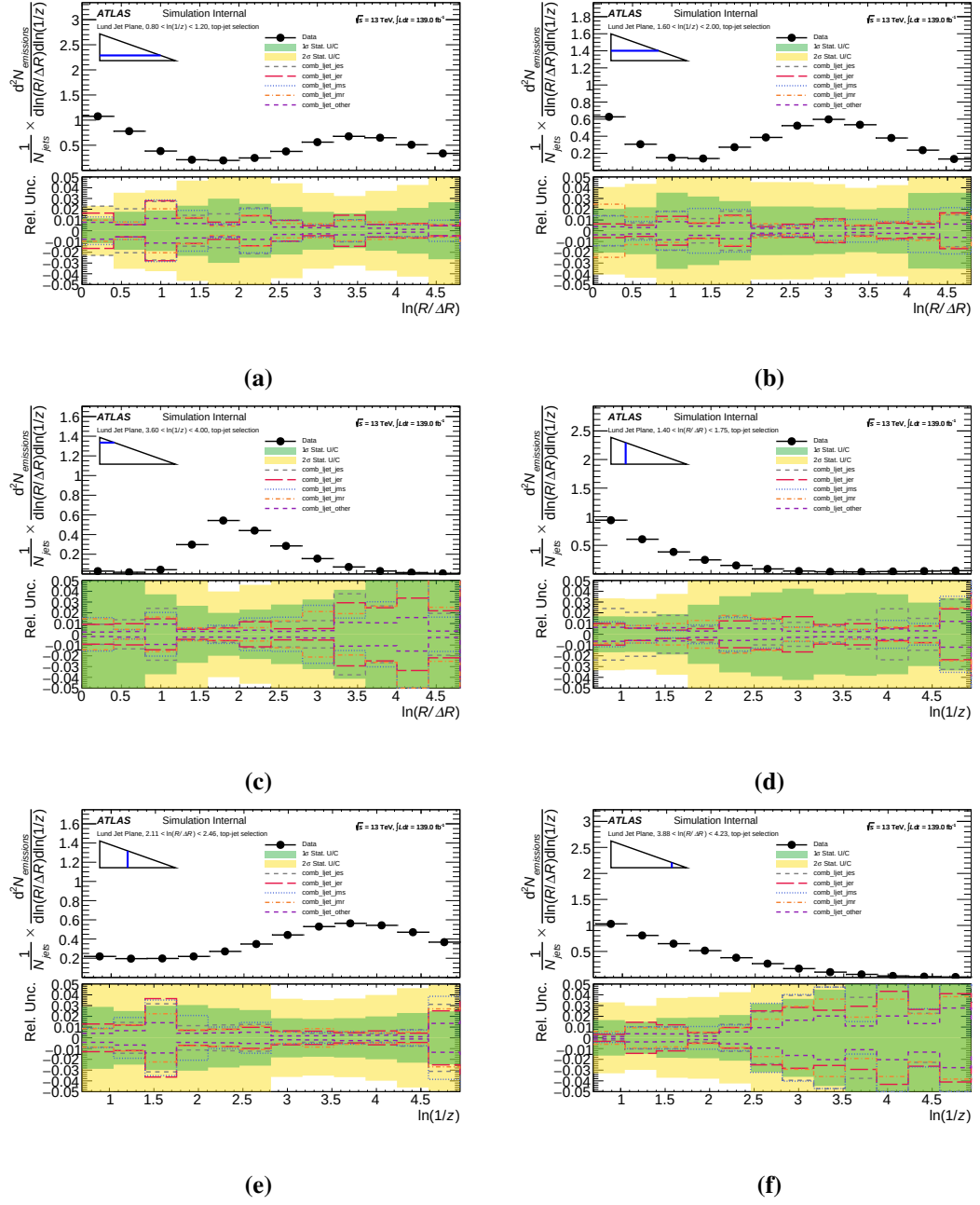


Figure B.1: Unfolded Lund Jet Plane slices for large-R jet systematic uncertainties in the top jet selection.

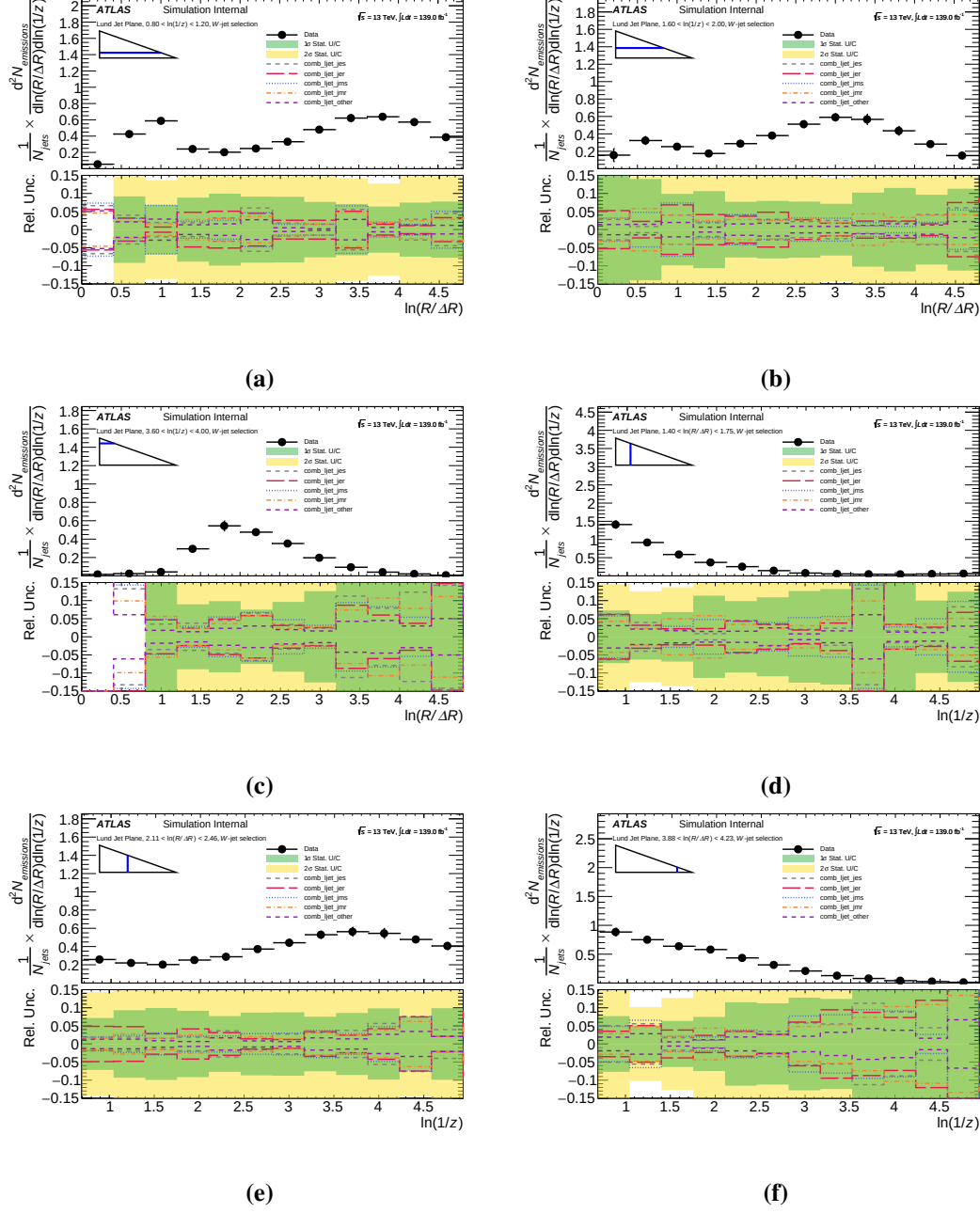


Figure B.2: Unfolded Lund Jet Plane slices for large-R jet systematic uncertainties in the W jet selection.

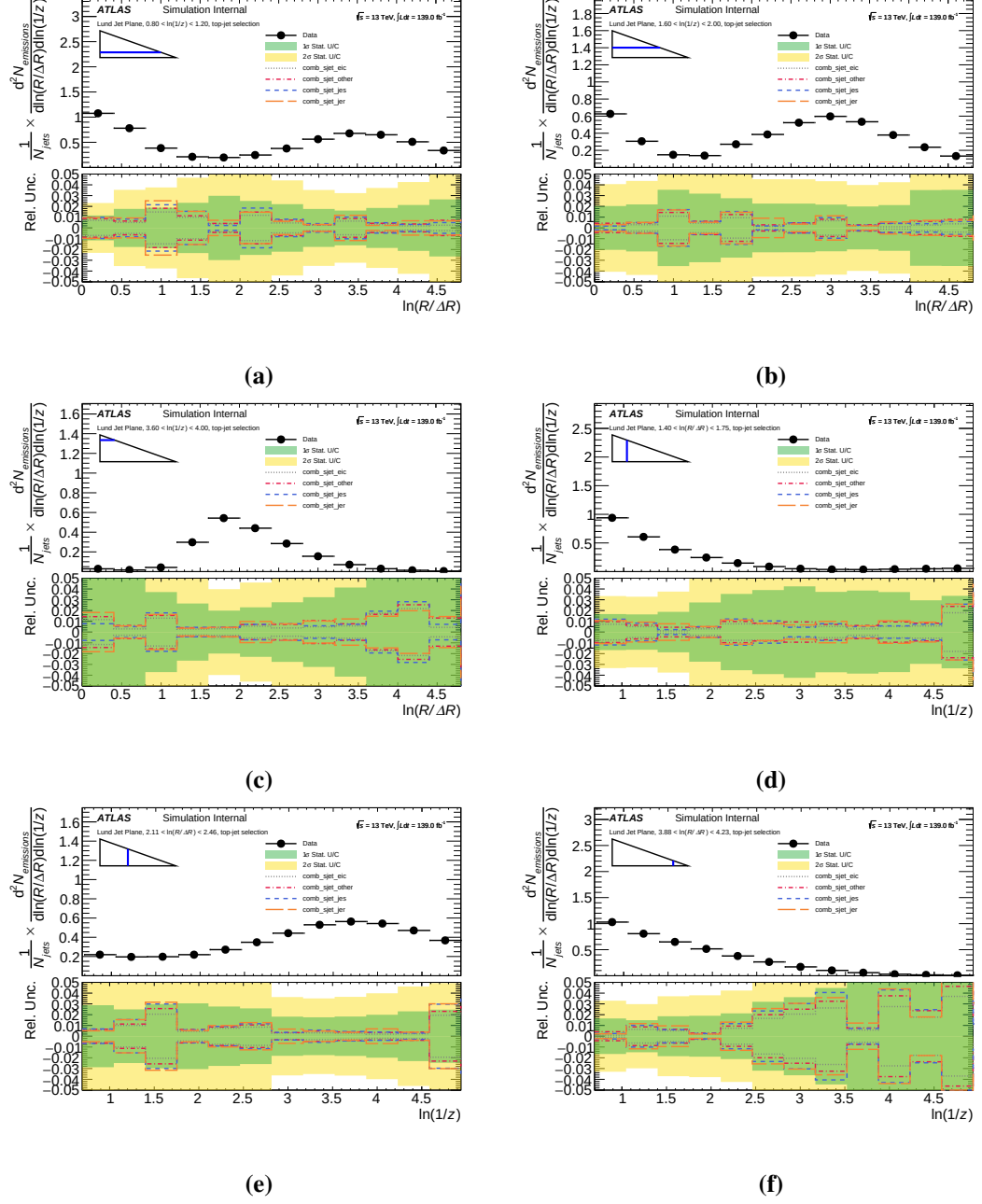


Figure B.3: Unfolded Lund Jet Plane slices for small R jet systematic uncertainties in the top jet selection.

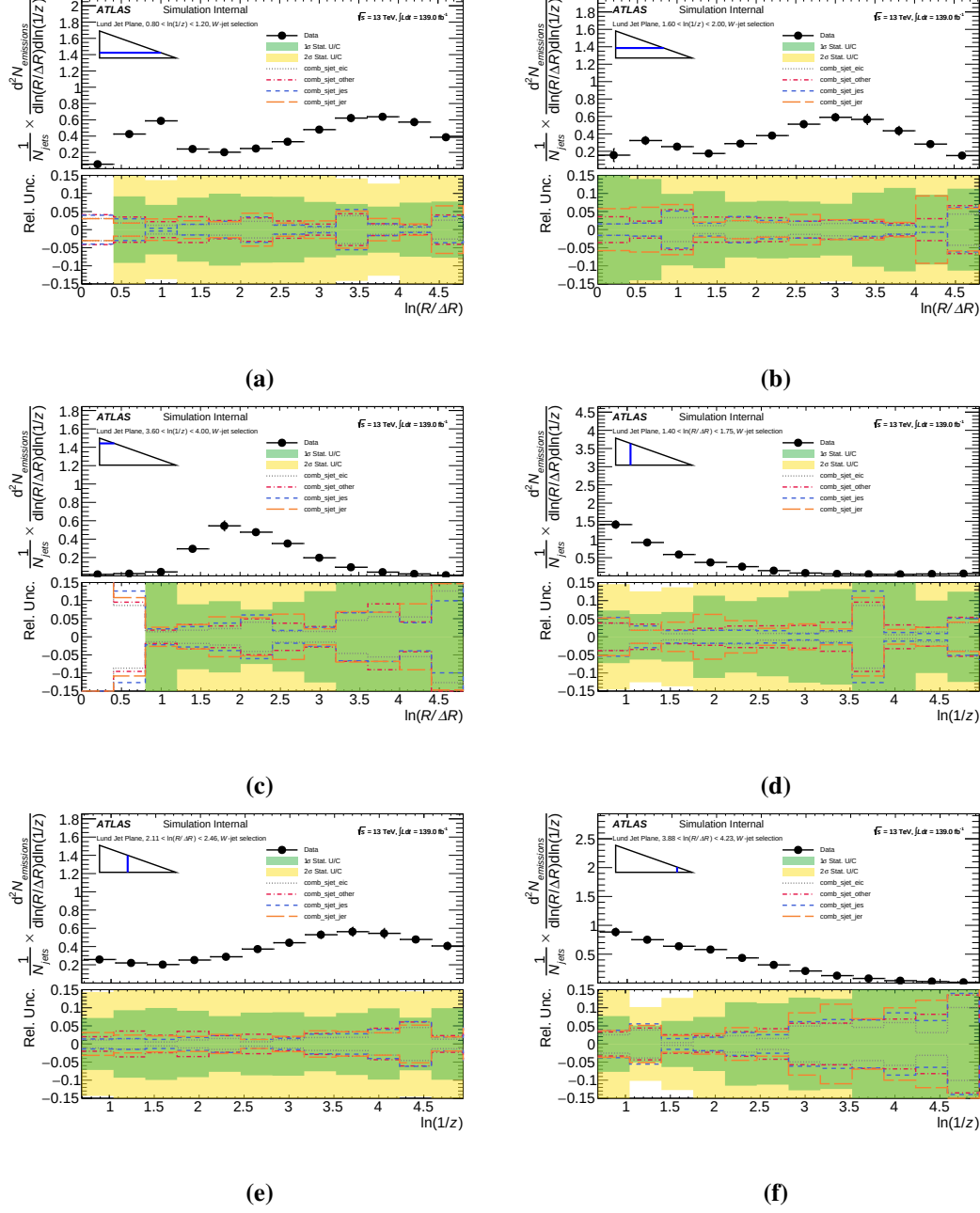


Figure B.4: Unfolded Lund Jet Plane slices for small R jet systematic uncertainties in the W jet selection.

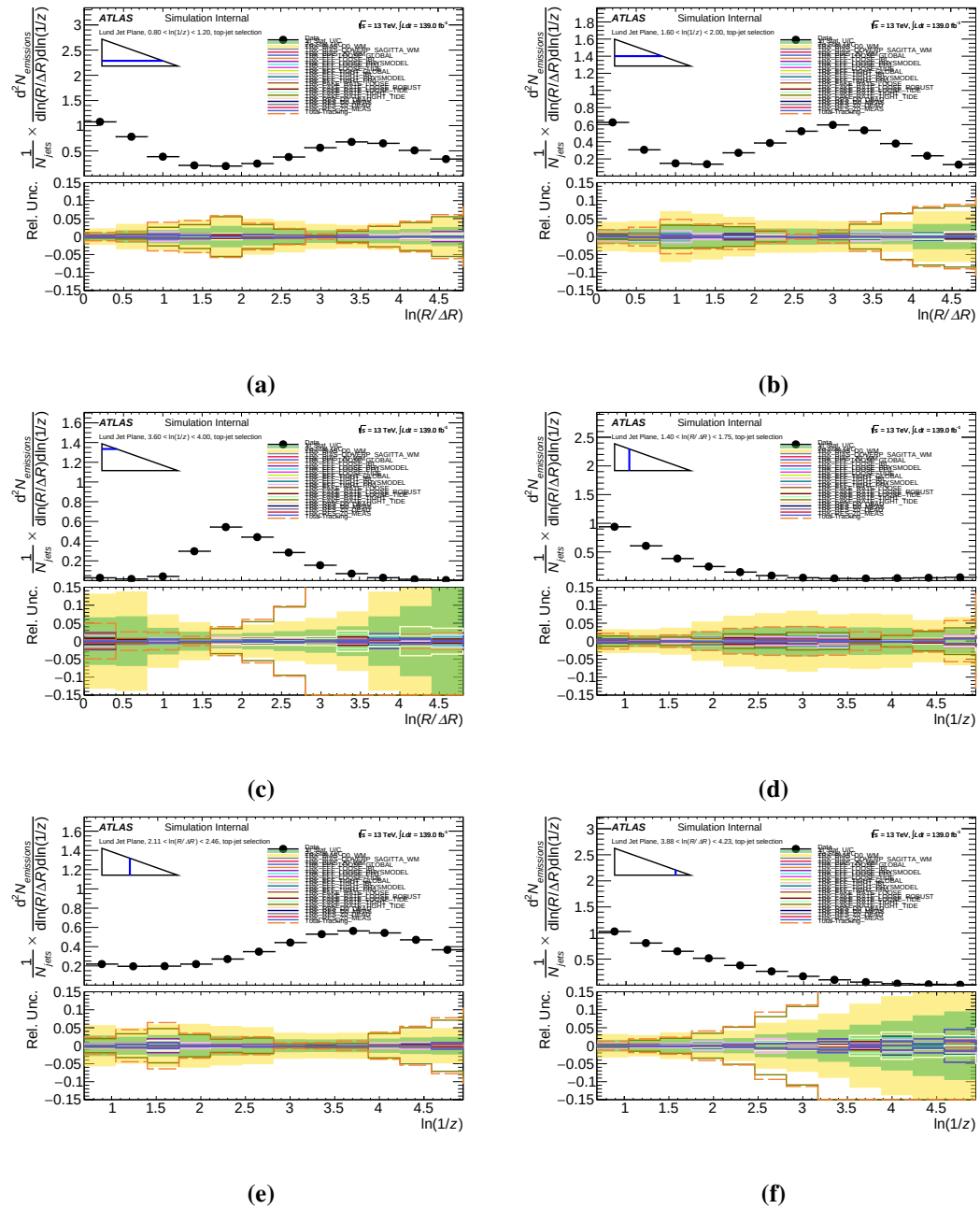


Figure B.5: Unfolded Lund Jet Plane slices for tracking systematic uncertainties in the top jet selection.

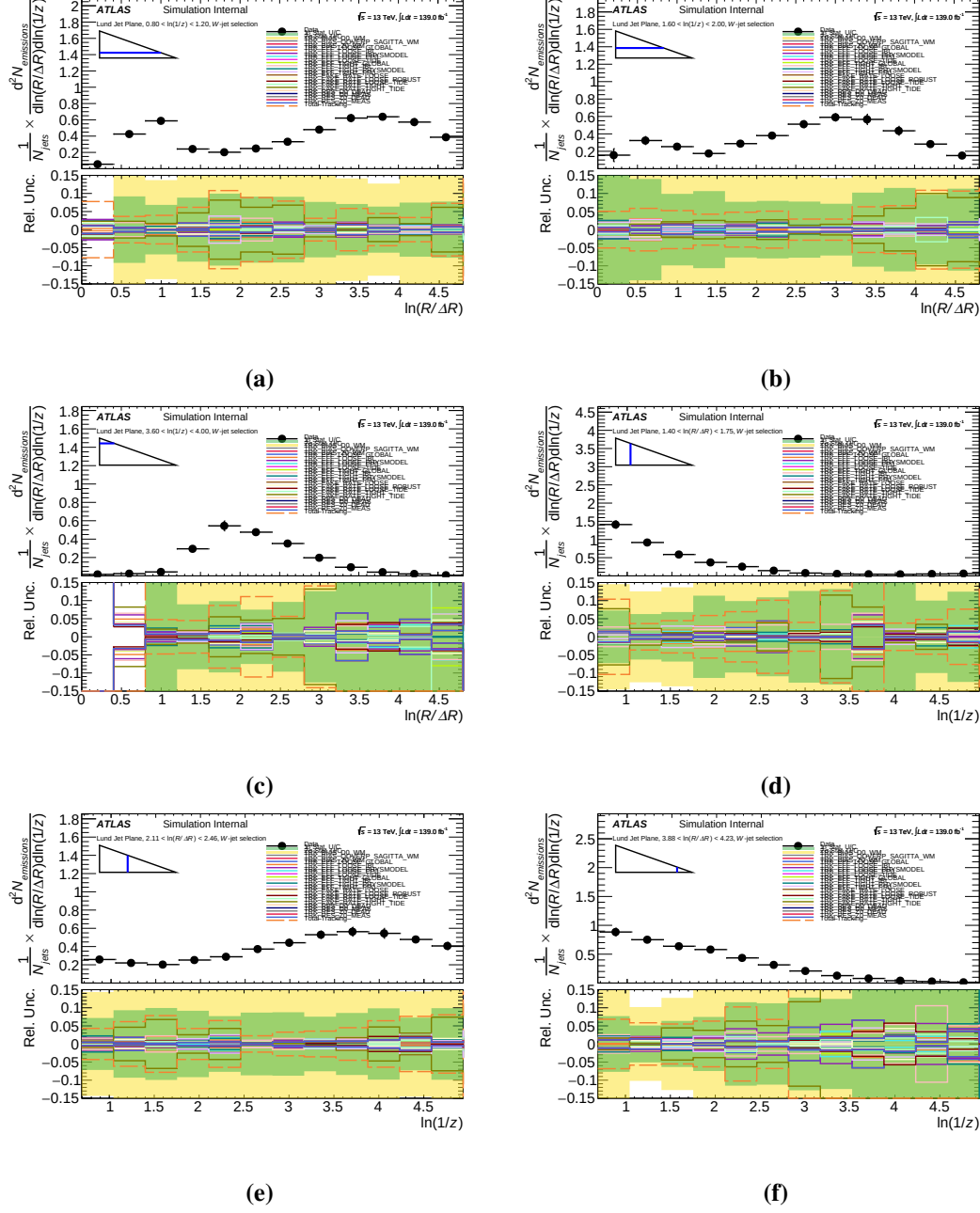


Figure B.6: Unfolded Lund Jet Plane slices for tracking systematic uncertainties in the W jet selection.

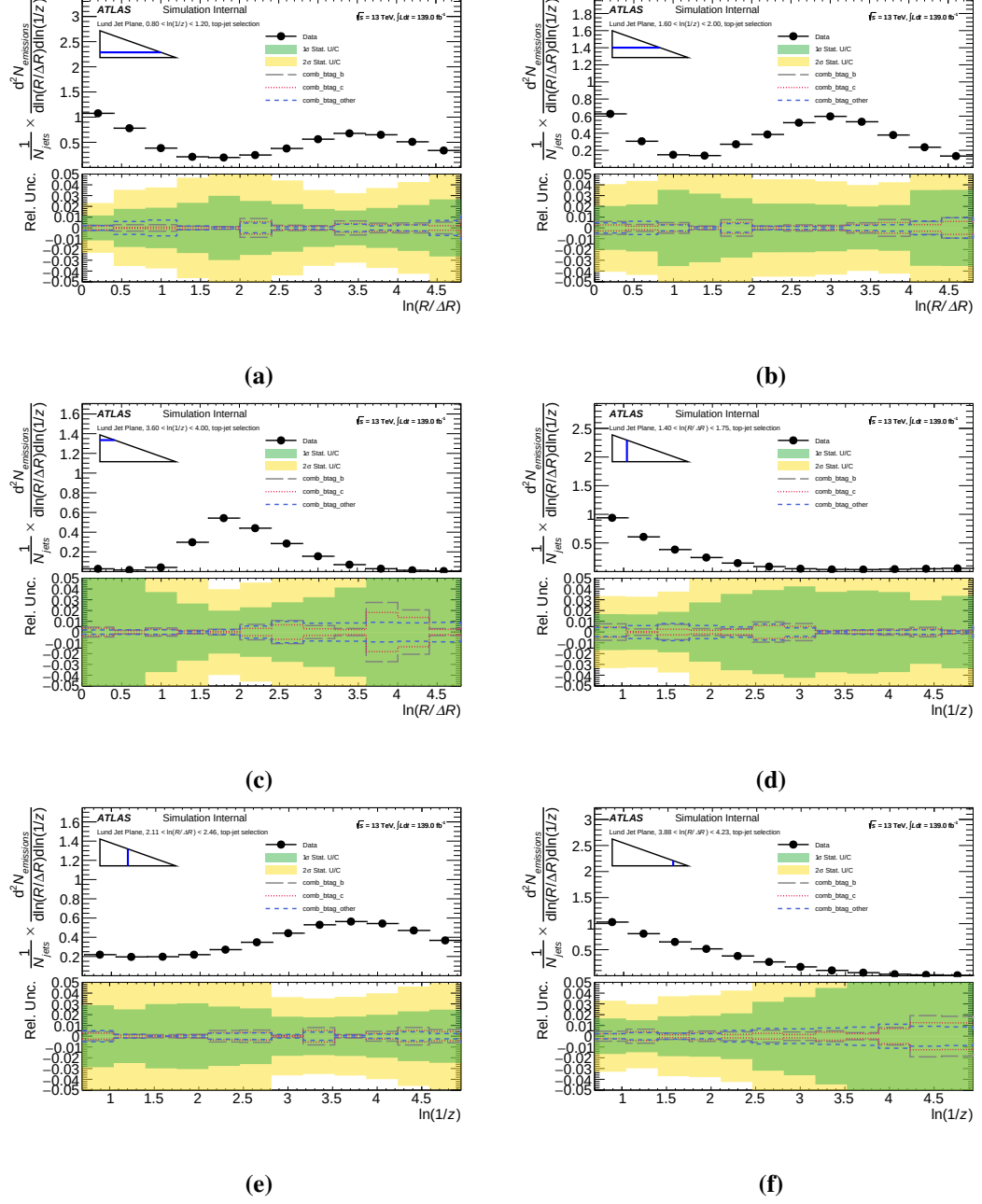


Figure B.7: Unfolded Lund Jet Plane slices for b-tagging systematic uncertainties in the top jet selection.

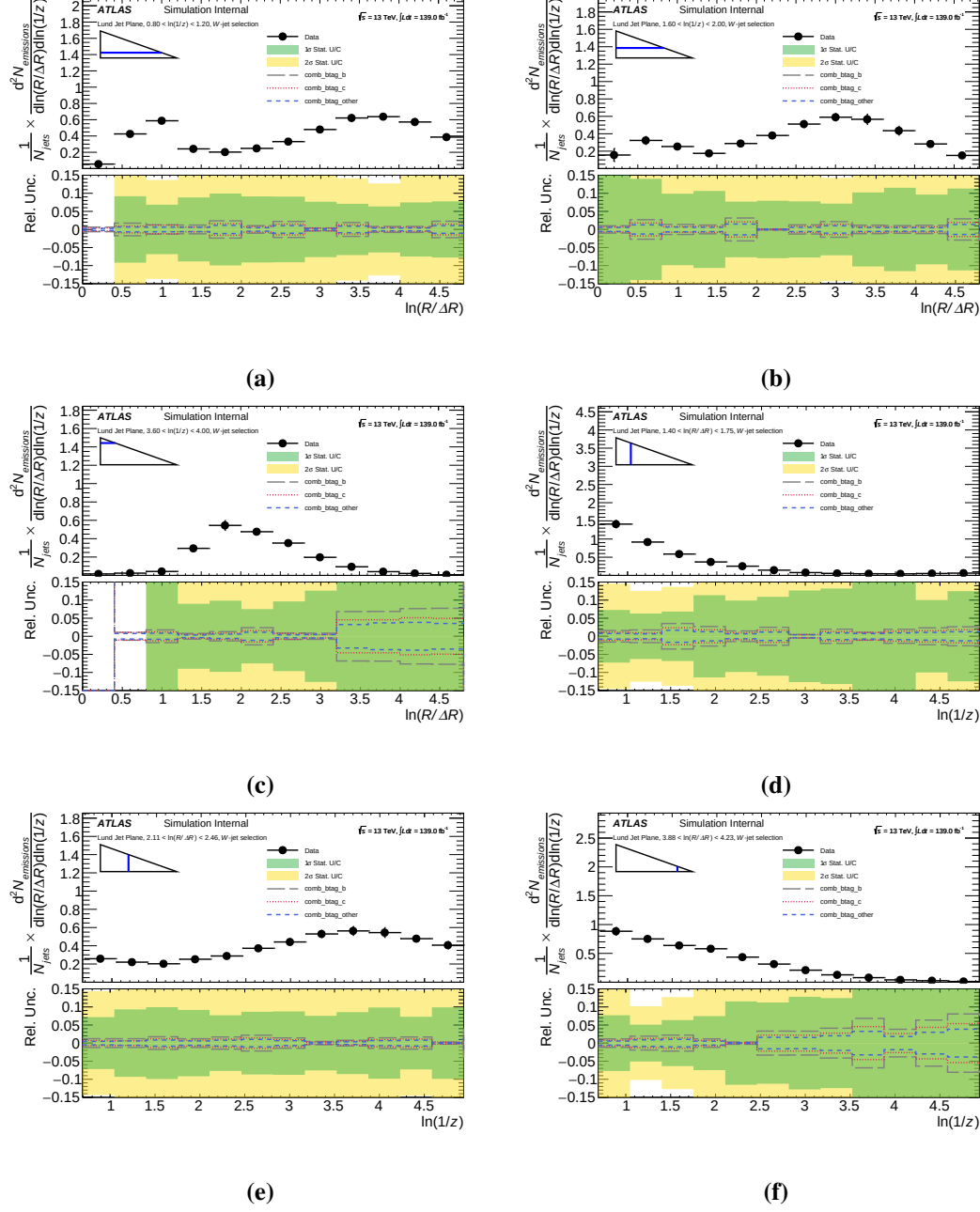


Figure B.8: Unfolded Lund Jet Plane slices for b-tagging systematic uncertainties in the W jet selection.

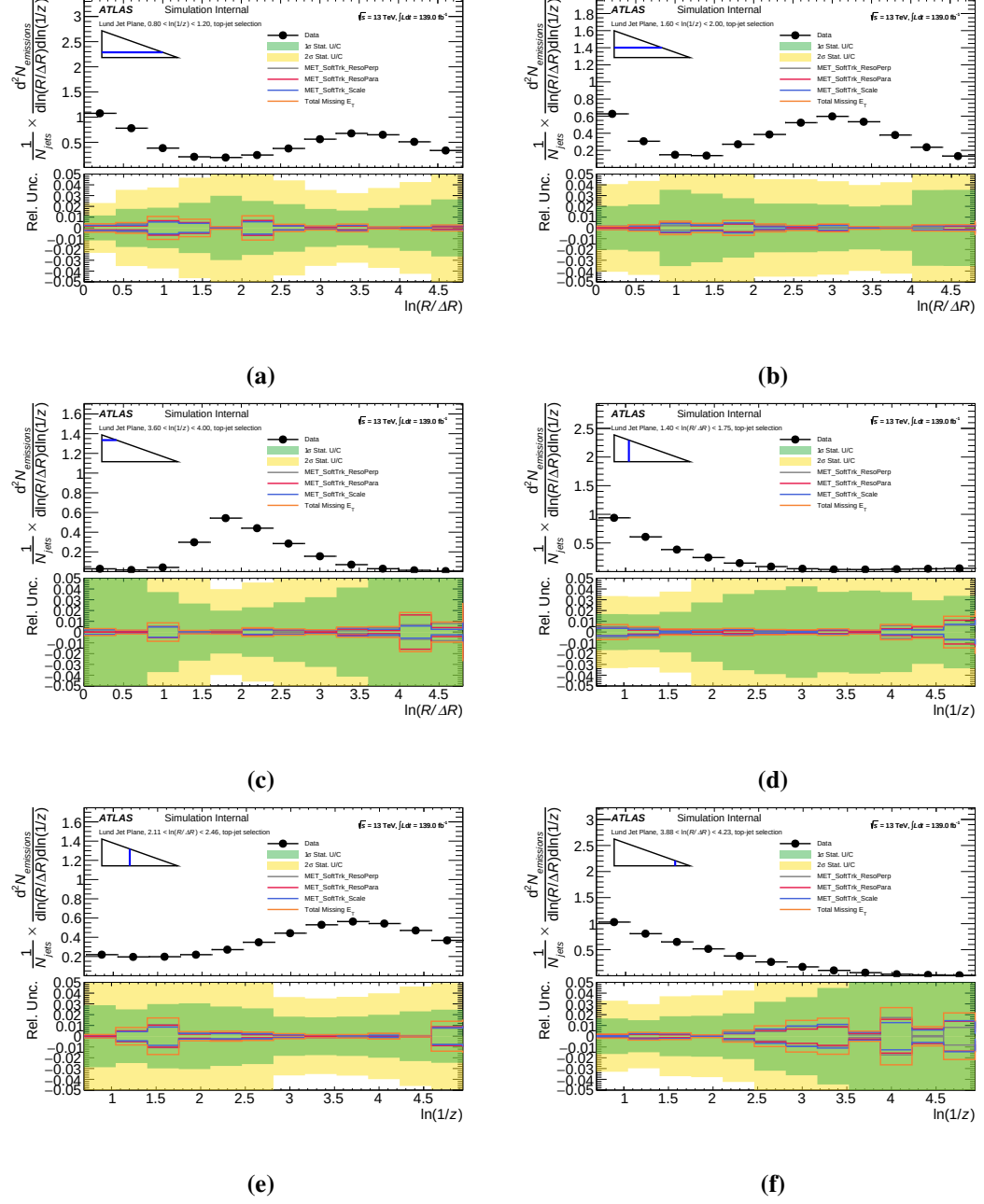


Figure B.9: Unfolded Lund Jet Plane slices for MET systematic uncertainties in the top jet selection.

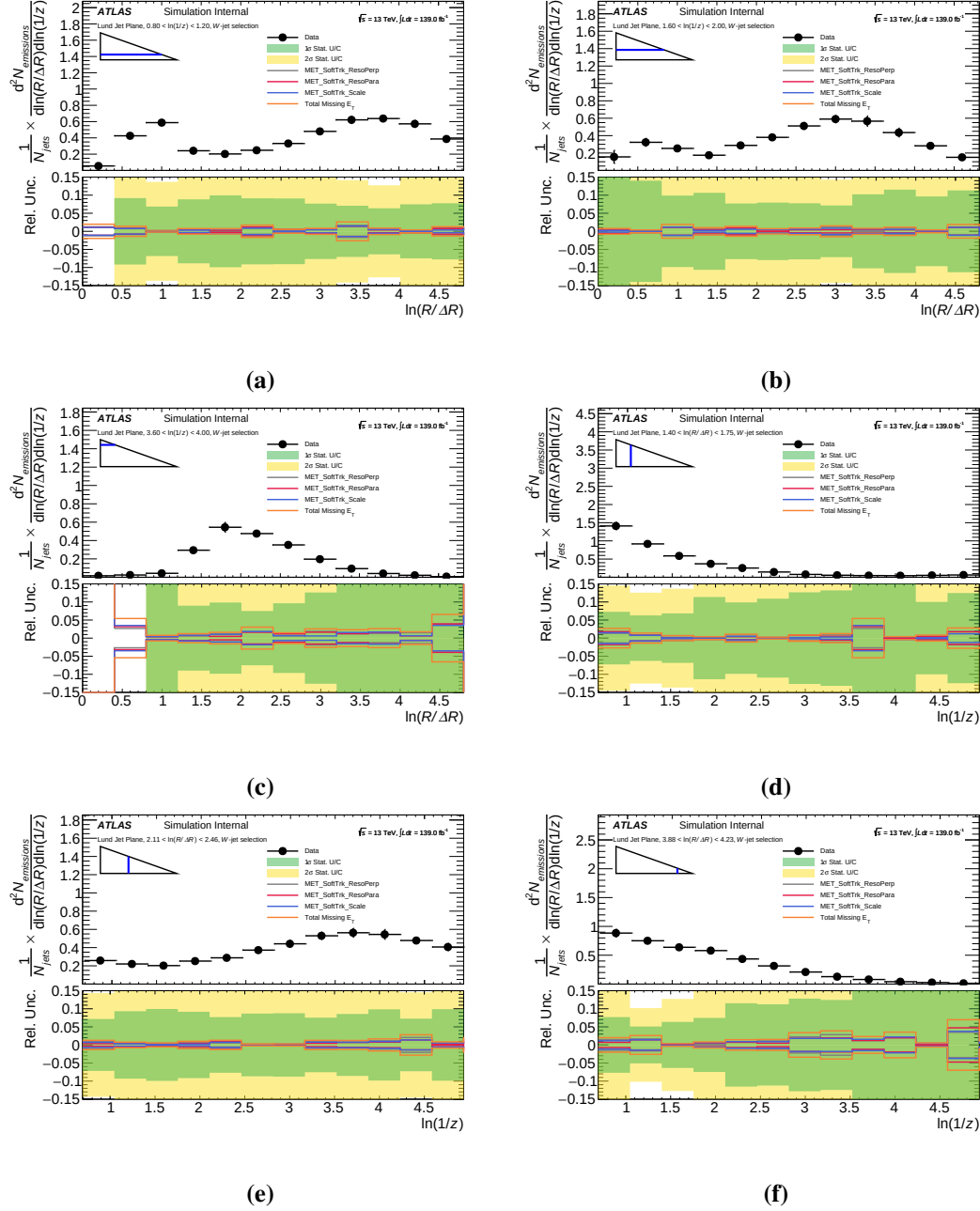


Figure B.10: Unfolded Lund Jet Plane slices for MET systematic uncertainties in the W jet selection.

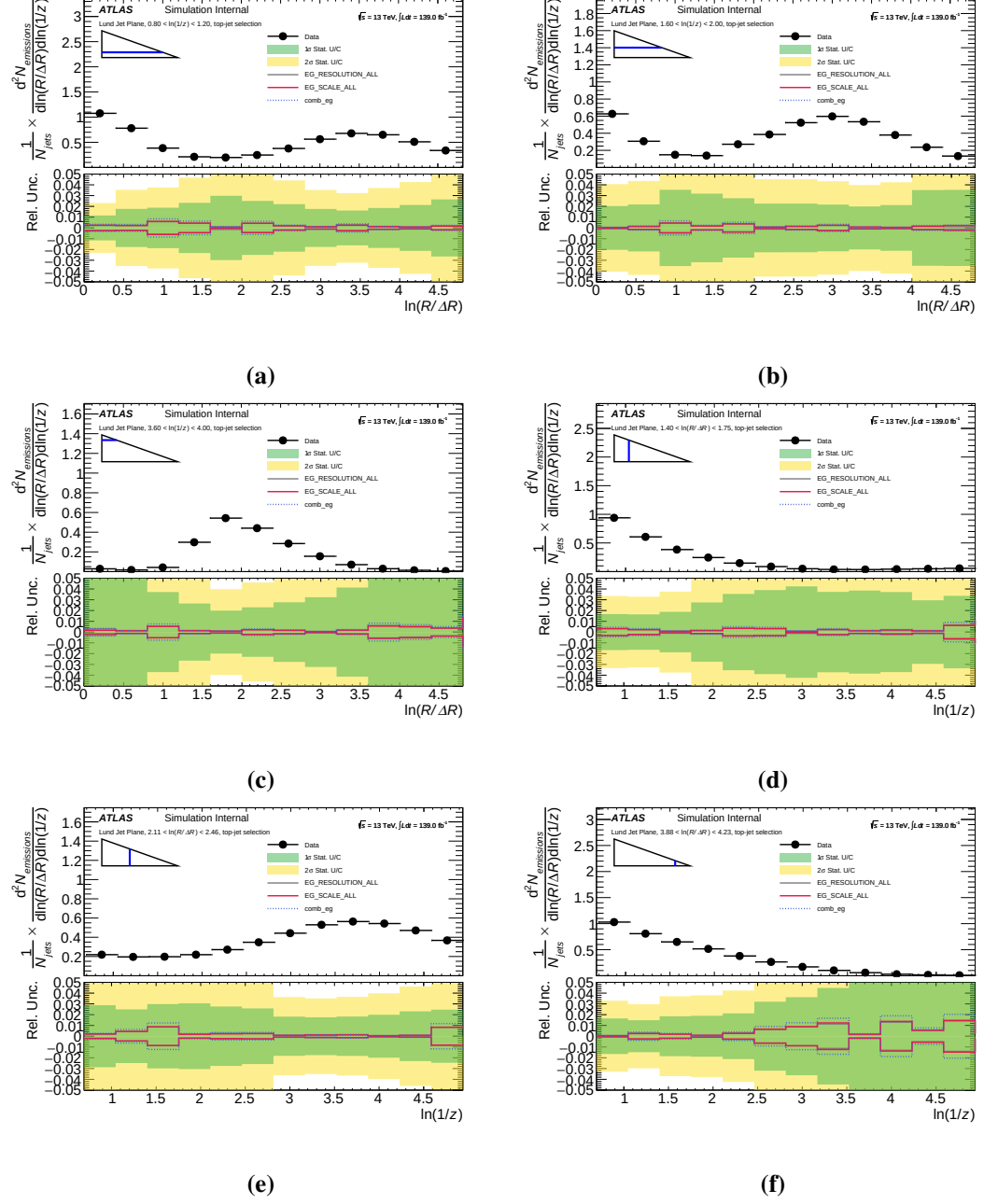


Figure B.11: Unfolded Lund Jet Plane slices for electron systematic uncertainties in the top jet selection.

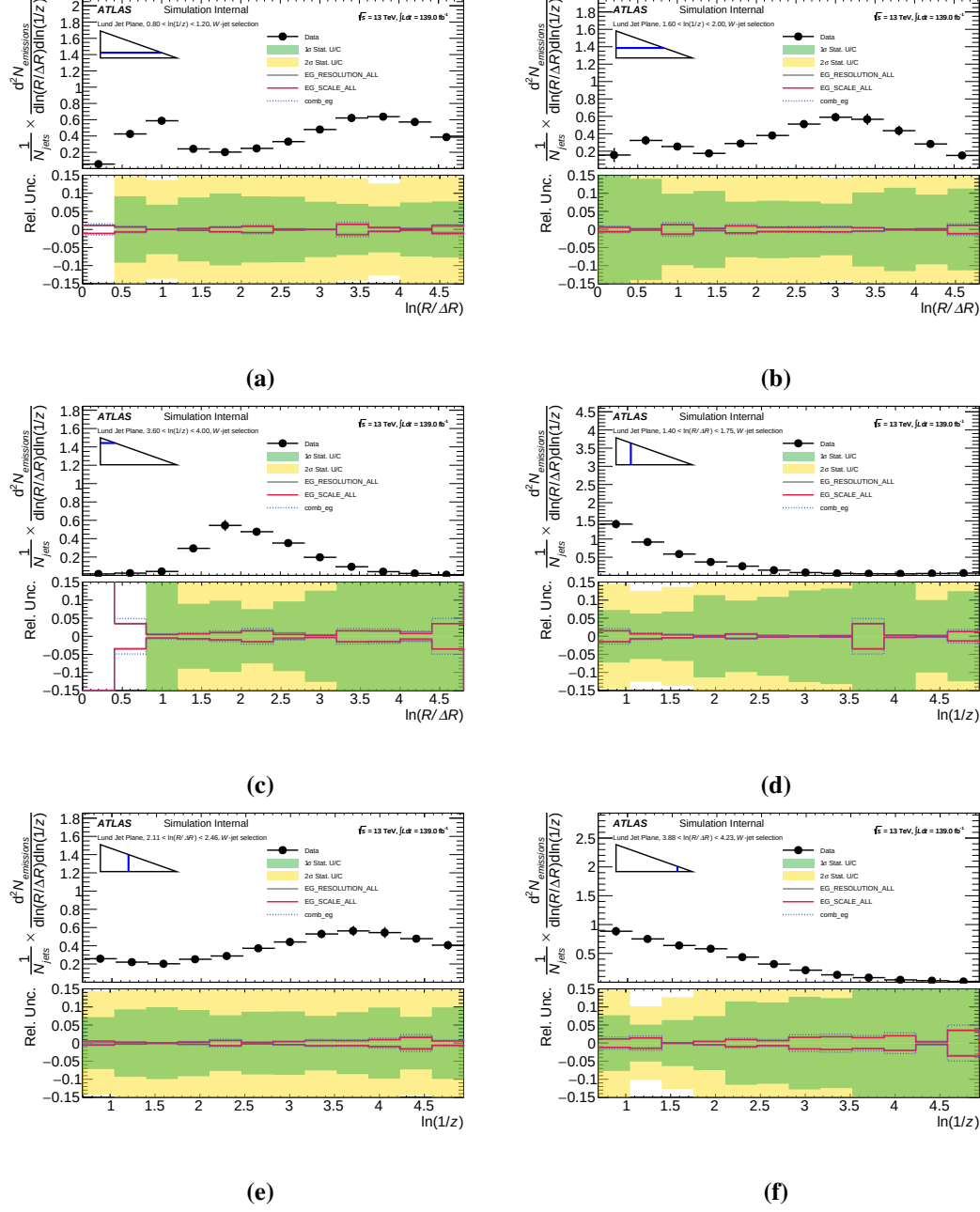


Figure B.12: Unfolded Lund Jet Plane slices for electron systematic uncertainties in the W jet selection.

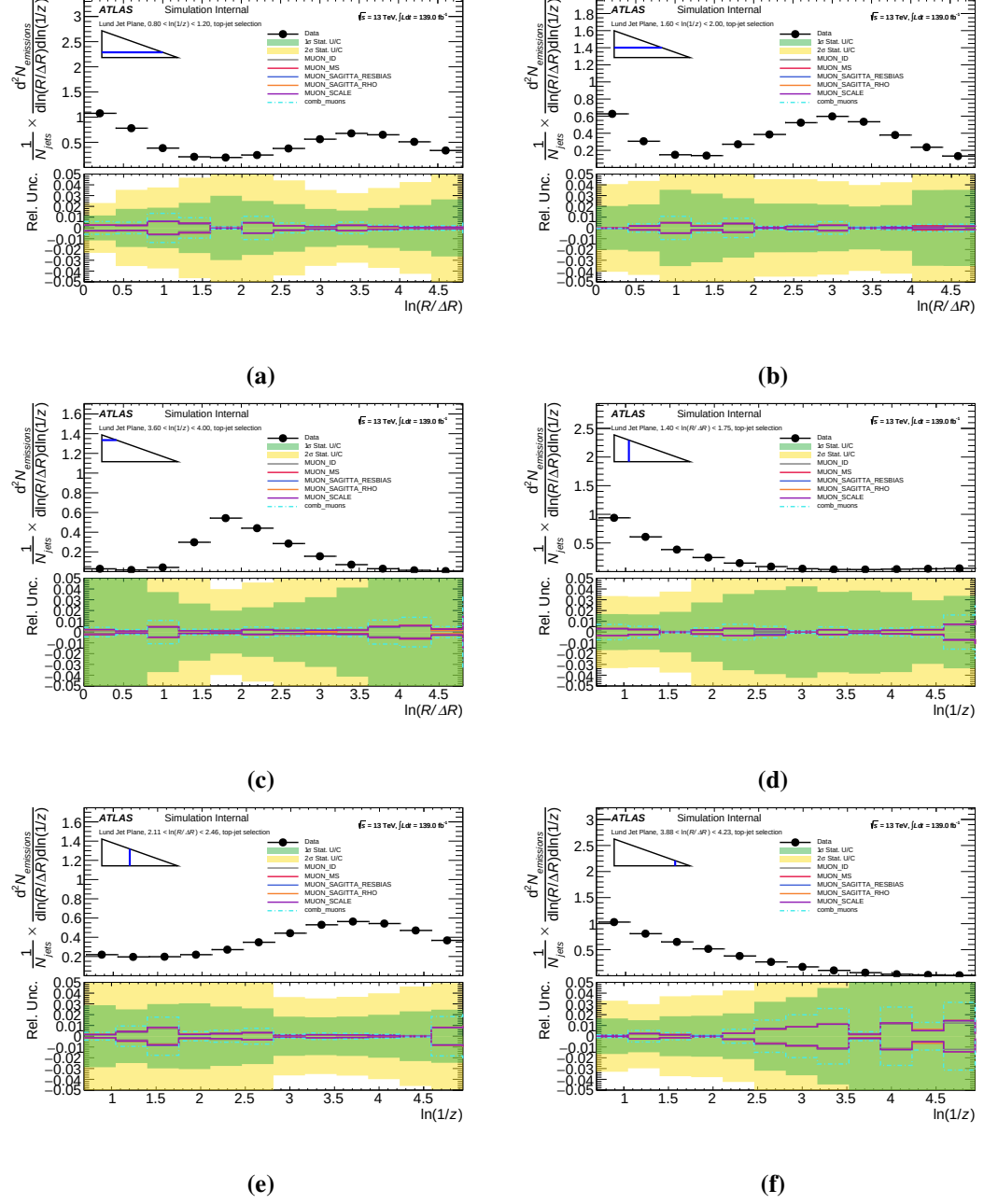


Figure B.13: Unfolded Lund Jet Plane slices for muon systematic uncertainties in the top jet selection.

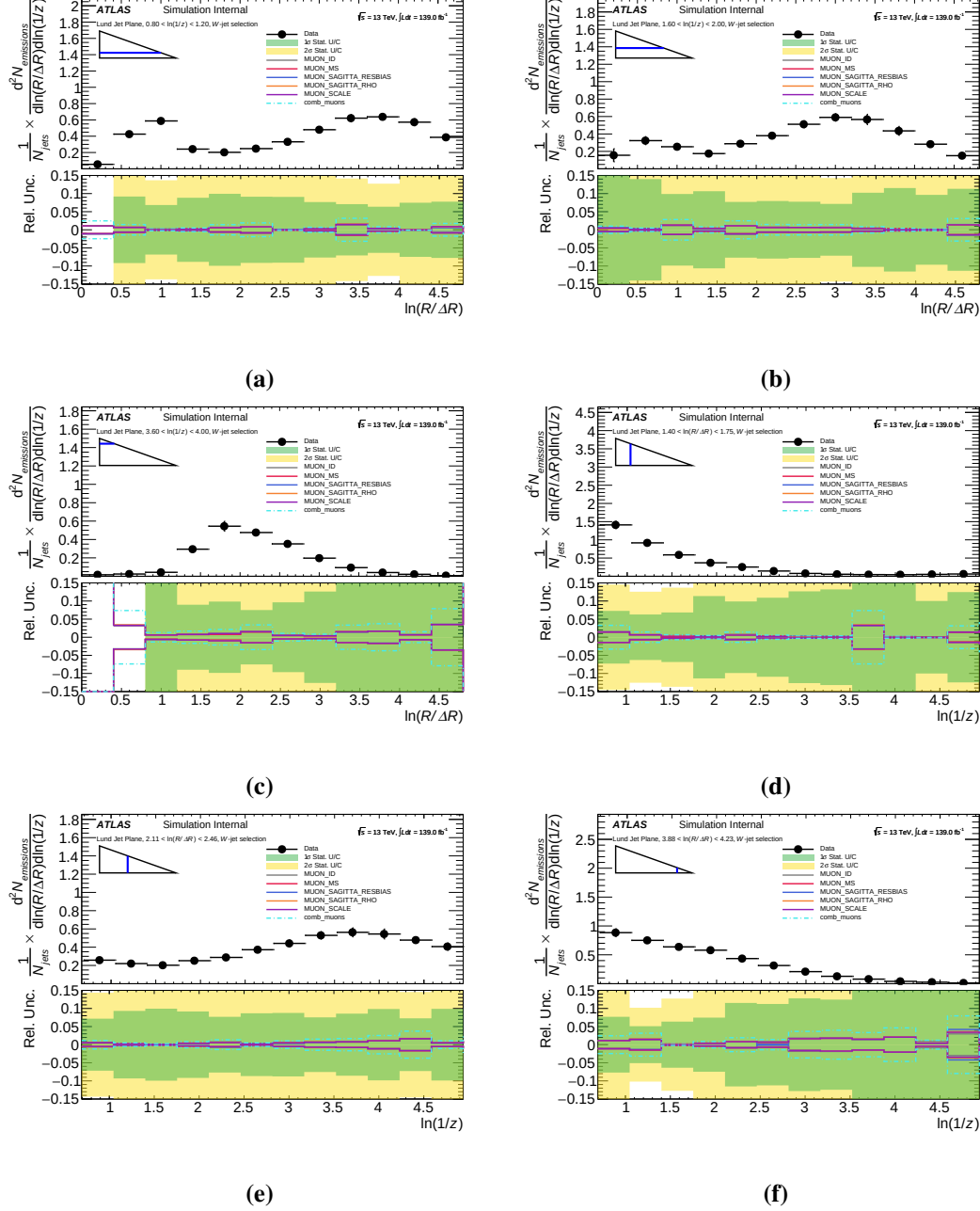


Figure B.14: Unfolded Lund Jet Plane slices for muon systematic uncertainties in the W jet selection.

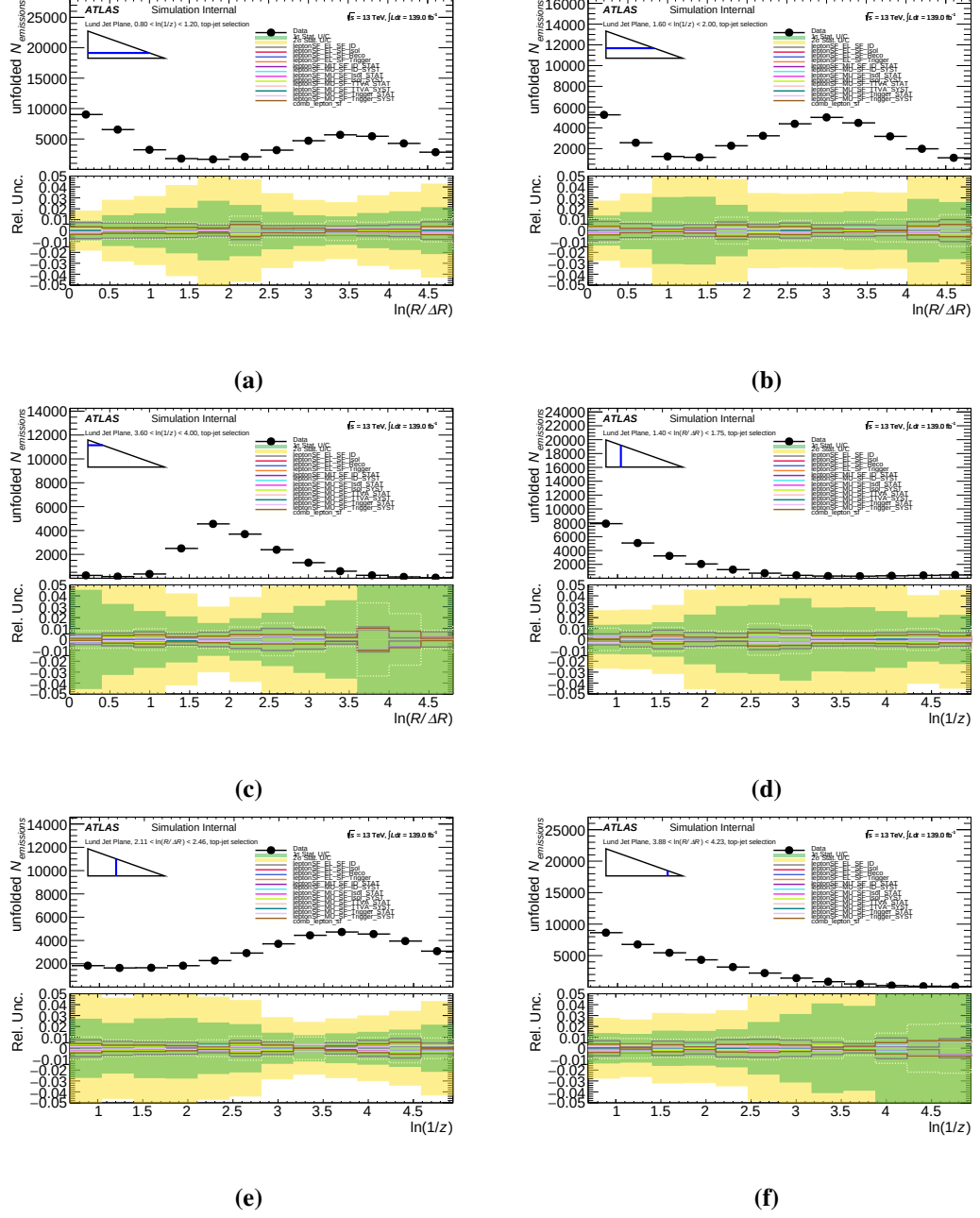


Figure B.15: Unfolded Lund Jet Plane slices for lepton-scalefactor systematic uncertainties in the top jet selection.

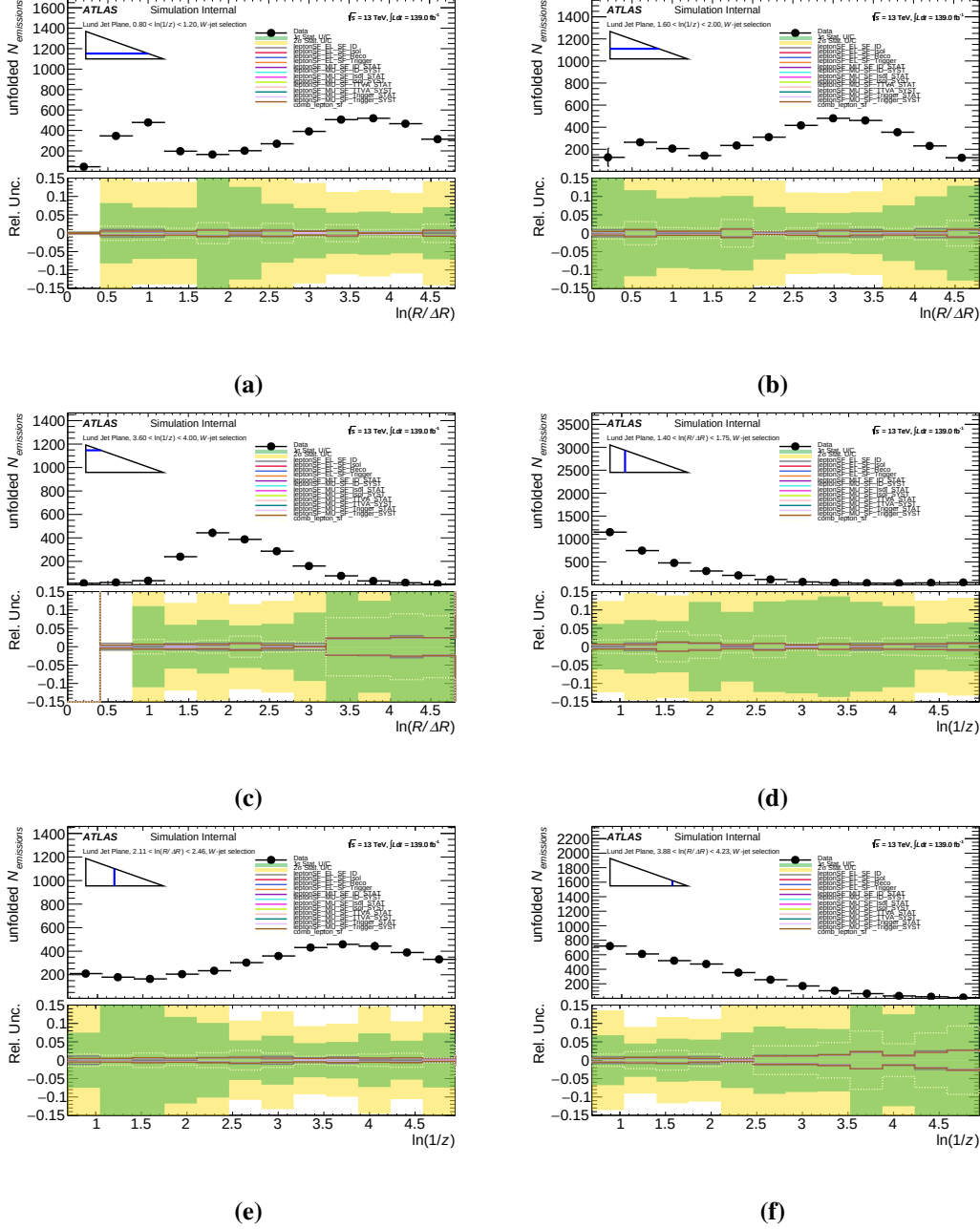


Figure B.16: Unfolded Lund Jet Plane slices for lepton-scalefactor systematic uncertainties in the W jet selection.

B.2 Covariance and correlation matrices

Systematic covariances are evaluated bin-by-bin from the systematic variations. Correlations between bins are evaluated from the sign of the up/down components of the uncertainties. Two bins are considered correlated if the sign of the uncertainties is the same and anti-correlated if they are opposite. The average of the up/down variations is taken for each systematic. In cases where the signs of the up and down components are opposite, the sign of the larger component is used in the calculation of the covariance. Explicitly, the formula for computing the systematic covariance matrix is

$$C_{ij}^{syst} = \frac{1}{4} \sum_{l=0}^{N_{syst}} s_i s_j \left(\left| \sigma_i^{up\,syst,l} \right| + \left| \sigma_i^{down\,syst,l} \right| \right) \left(\left| \sigma_j^{up\,syst,l} \right| + \left| \sigma_j^{down\,syst,l} \right| \right) \quad (B.1)$$

where x^{nom} denotes the bins of the nominal histogram, x^{syst} denotes the histogram bins for the systematic variation and N_{syst} is the total number of systematics. The sign of the systematic is :

$$s_i = \begin{cases} \text{sgn}(\sigma_i^{up\,syst,l}), & \text{if } |\sigma_i^{up\,syst,l}| > |\sigma_i^{down\,syst,l}|. \\ \sigma_i^{down\,syst,l}, & \text{otherwise.} \end{cases} \quad (B.2)$$

The covariance and correlation matrices for top- and W jets are shown in Figure B.17 and Figure B.19.

The statistical covariance matrix is evaluated using the bootstrap replicas in analogy with the statistical uncertainty. The formula for the covariance between bins i and j is

$$C_{ij}^{stat} = \frac{1}{N_{BS}} \sum_{k=1}^{N_{BS}} \left(x_i^k - \tilde{x}_i \right) \left(x_j^k - \tilde{x}_j \right), \quad (B.3)$$

where x^k are the N_{BS} replica histograms and $\tilde{x}$ is their average which is equivalent to the measured histogram. The systematic and statistical correlation matrices are defined as

$$K_{ij}^{stat} = \frac{1}{\sigma_i^{syst} \sigma_j^{syst}} C_{ij}; \quad K_{ij}^{syst} = \frac{1}{\sigma_i^{stat} \sigma_j^{stat}} C_{ij}^{stat}, \quad (B.4)$$

where σ^{syst} and σ^{stat} are the histograms of the systematic and statistical uncertainty.

The total systematic covariance and correlation matrices are shown in Figures B.17 and B.19. The statistical covariance and correlation is shown in Figures B.18 and B.20.

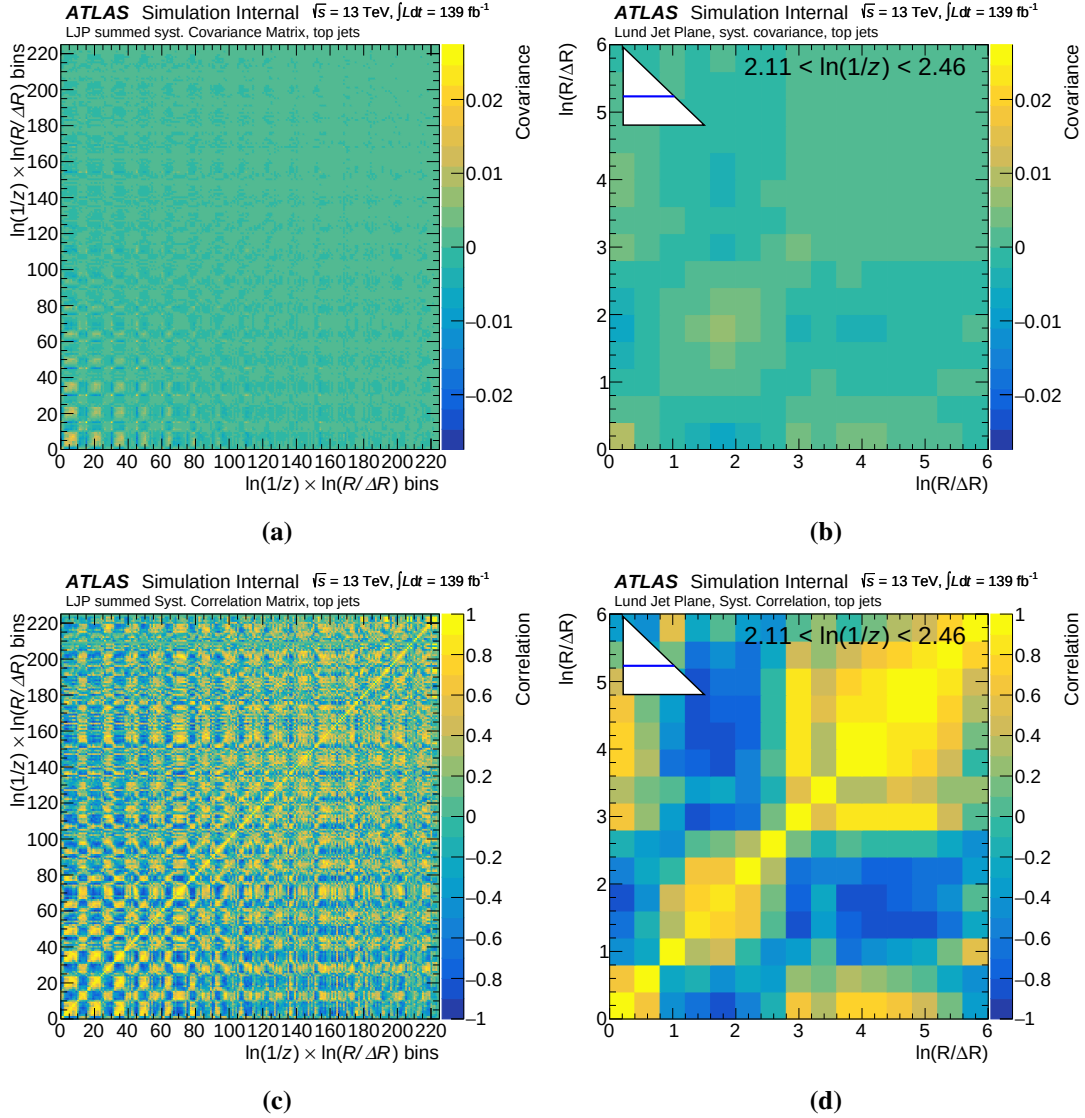


Figure B.17: The total systematic covariance matrix in the top jet selection is shown for the whole LJP in (a) and for a slice along the $\ln(R/\Delta R)$ axis in (c). The corresponding correlation matrices are shown in (b) and (d).

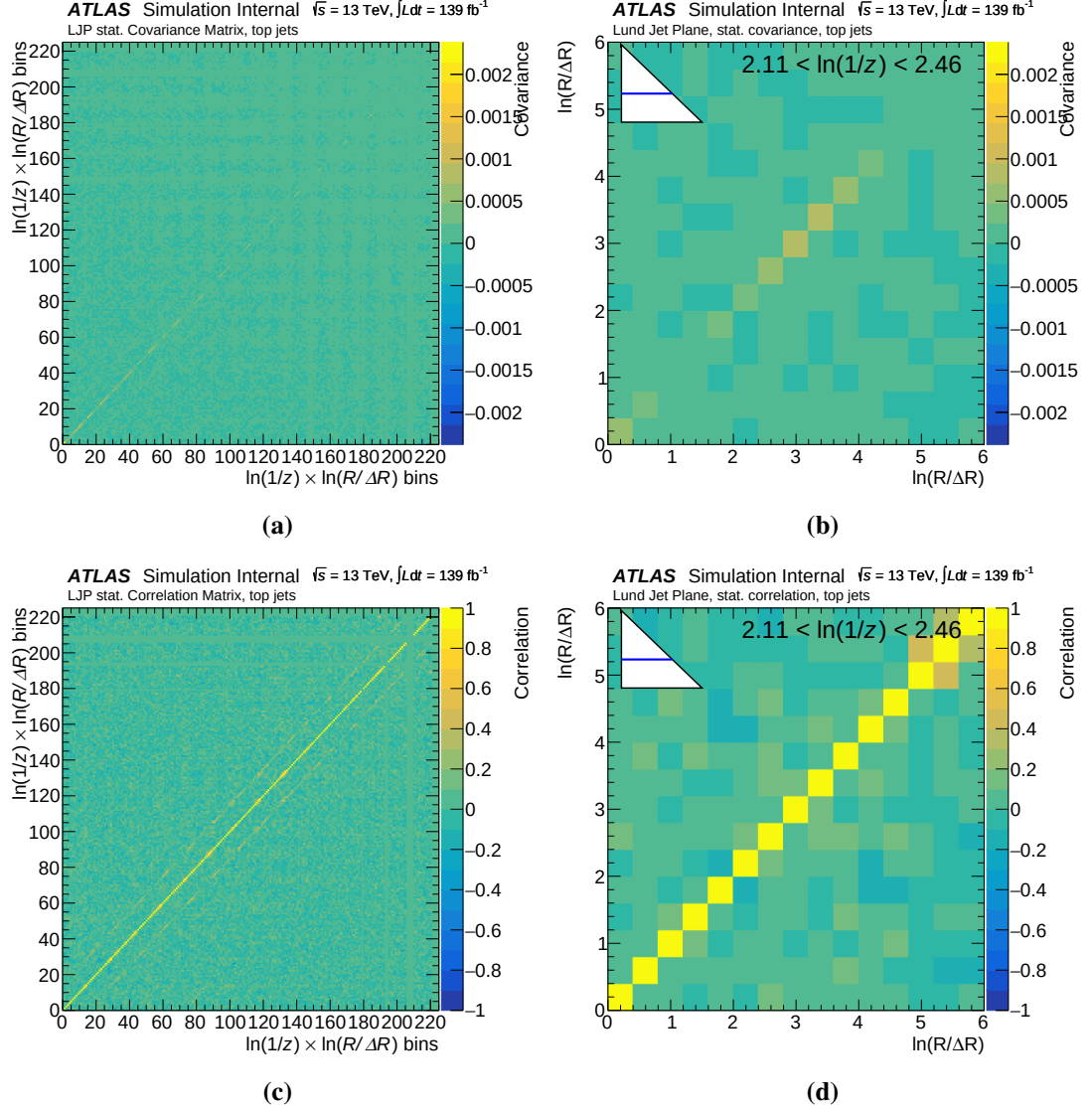


Figure B.18: The statistical covariance matrix in the top jet selection is shown for the whole LJP in (a) and for a slice along the $\ln(R/\Delta R)$ axis in (c). The corresponding correlation matrices are shown in (b) and (d).

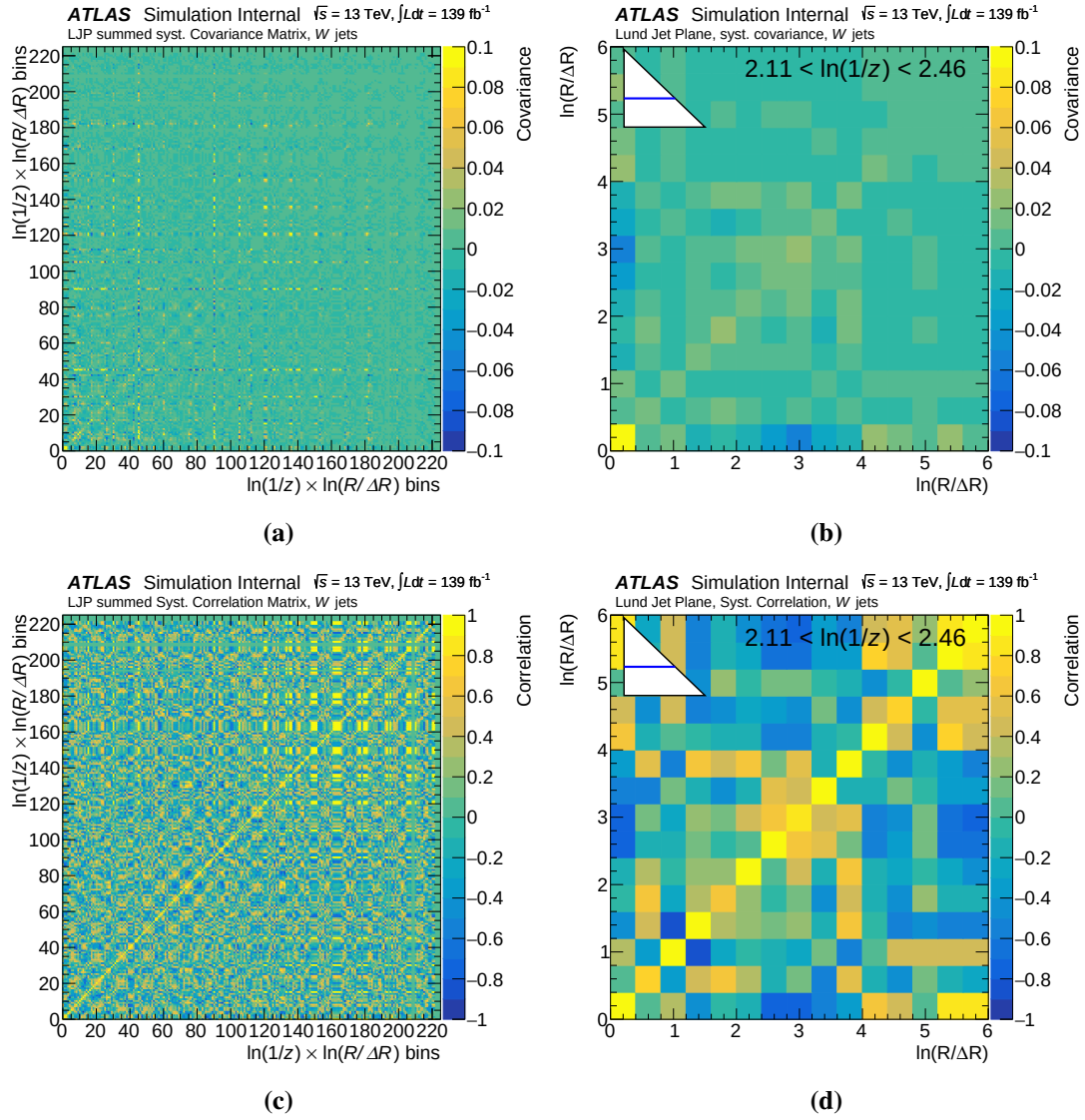


Figure B.19: The total systematic covariance matrix in the W jet selection is shown for the whole LJP in (a) and for a slice along the $\ln(R/\Delta R)$ axis in (c). The corresponding correlation matrices are shown in (b) and (d).

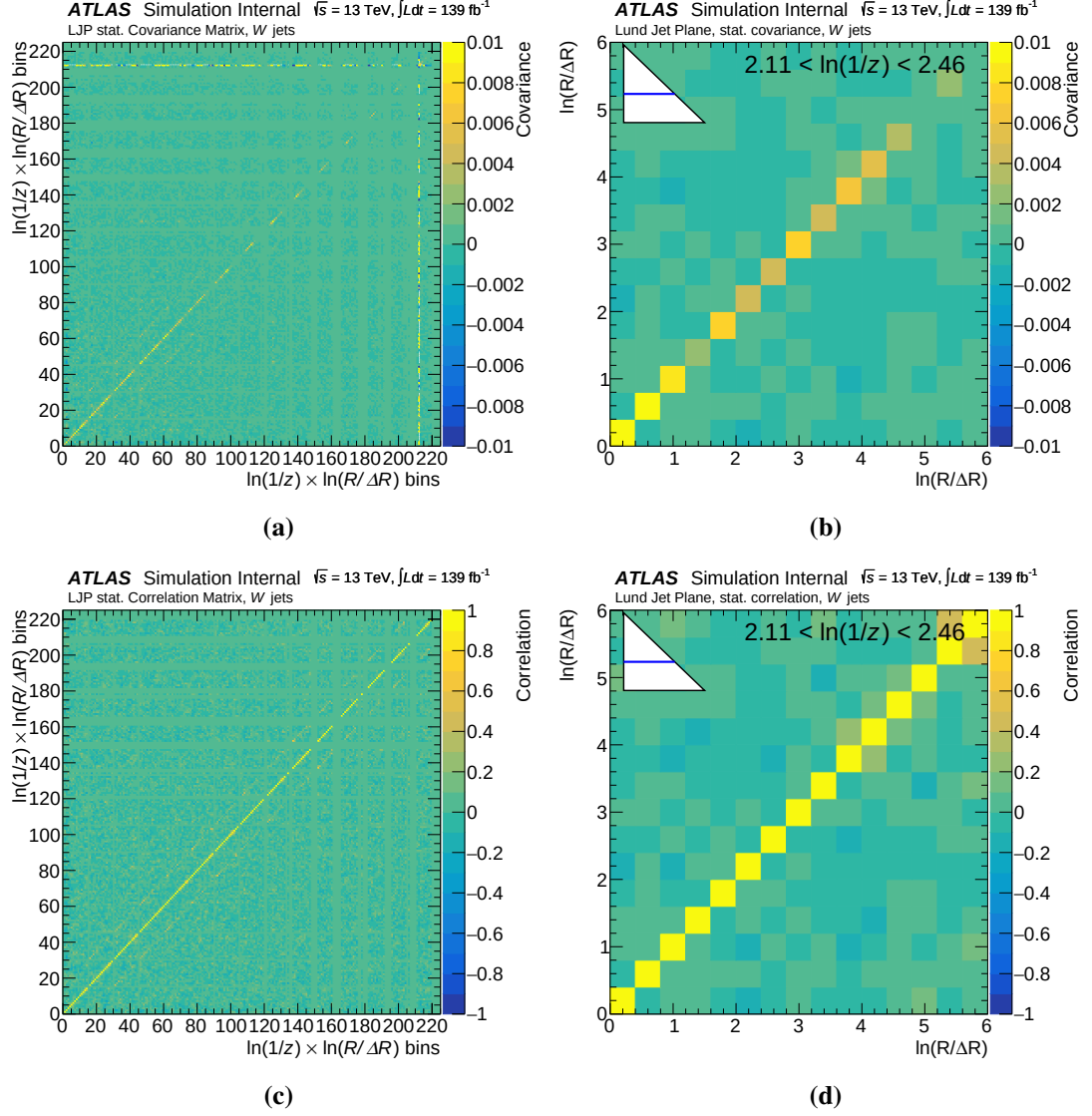


Figure B.20: The statistical covariance matrix in the W jet selection is shown for the whole LJP in (a) and for a slice along the $\ln(R/\Delta R)$ axis in (c). The corresponding correlation matrices are shown in (b) and (d).

B.3 Normalisation of the LJP with uncertainties

This section contains a summary of the average density of emissions across the Lund Jet Plane, given by $N_{\text{emissions}}/N_{\text{jets}}$, to which the measured LJP is normalised. Figure B.21 shows $N_{\text{emissions}}/N_{\text{jets}}$ for the unfolded data, the complete set of MC generators evaluated in the analysis and the different sources of uncertainty associated with the unfolded measurement.

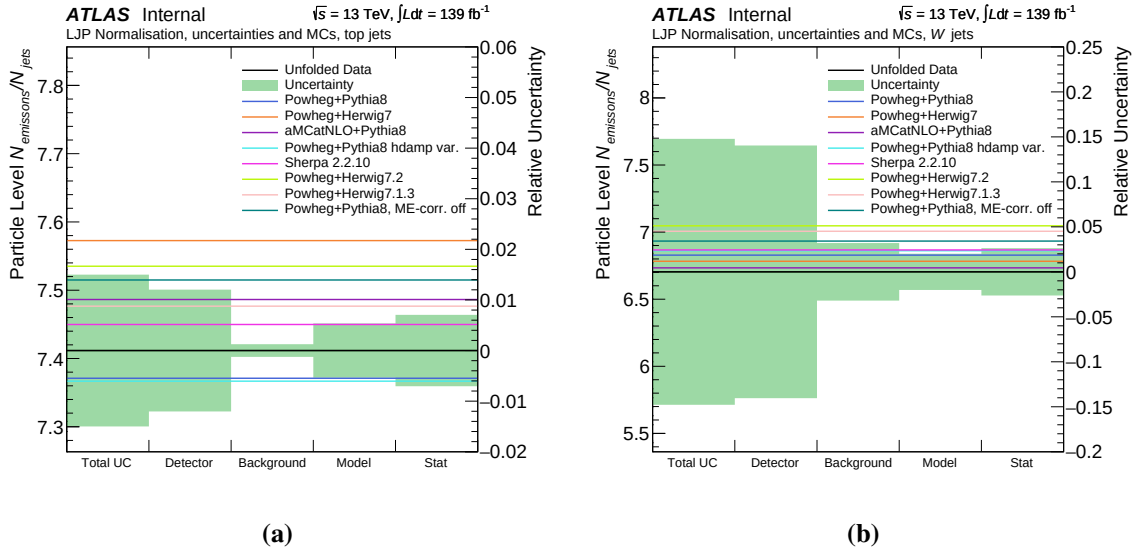


Figure B.21: The average density of emissions across the LJP for the top jets (a) and W jets (b) for unfolded data and particle level MCs. The total uncertainty on the measurement, as well as the contributions from $t\bar{t}$ modelling, background modelling, detector systematics and statistical uncertainty are also shown.

B.4 Background cross-checks

This section contains breakdowns of the uncertainties associated with the background MC samples in the control region defined in section 4.3.4. The uncertainties on the large- R jet mass and p_T distributions are shown in figure B.22, including background and $t\bar{t}$ modelling and the complete set of detector systematic uncertainties.

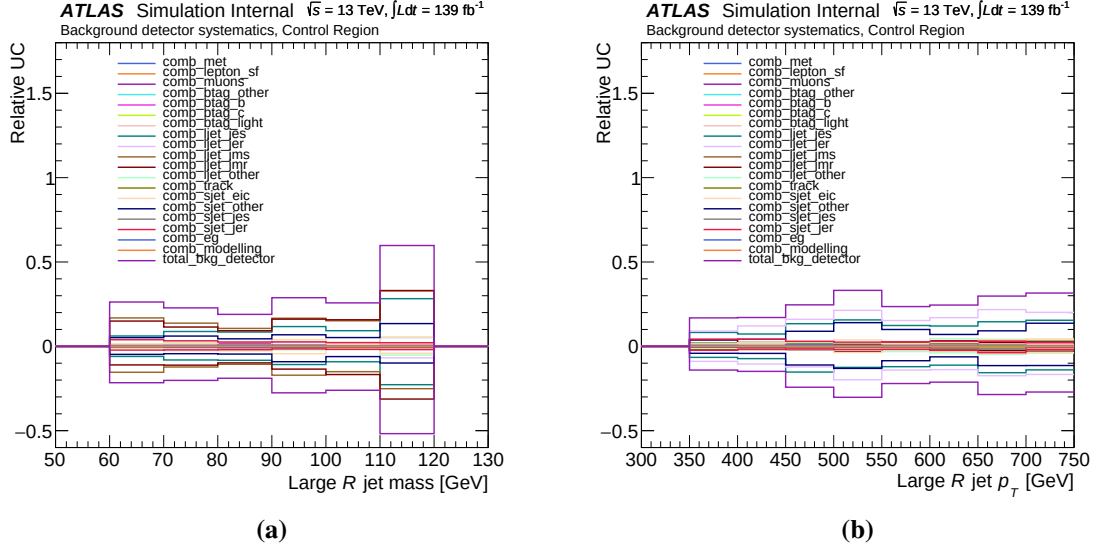


Figure B.22: Background detector uncertainties vs. large- R jet mass (a) and p_T (b) in the control region. The uncertainties is evaluated for unnormalised detector level distributions.

B.5 Smallest & greatest Data-MC LJP for alternative MC samples

This appendix complements the slices through the Lund Jet Plane for data and Monte Carlo samples shown in section 4.7. The samples with the smallest and greatest difference with the data in each bin are shown in figure B.23 for the top jets and figure B.24 for the W jets. Figures B.25 and B.26 shows only the Powheg + Pythia 8 samples and figures B.27 and B.28 show only the Powheg + Herwig 7 samples.

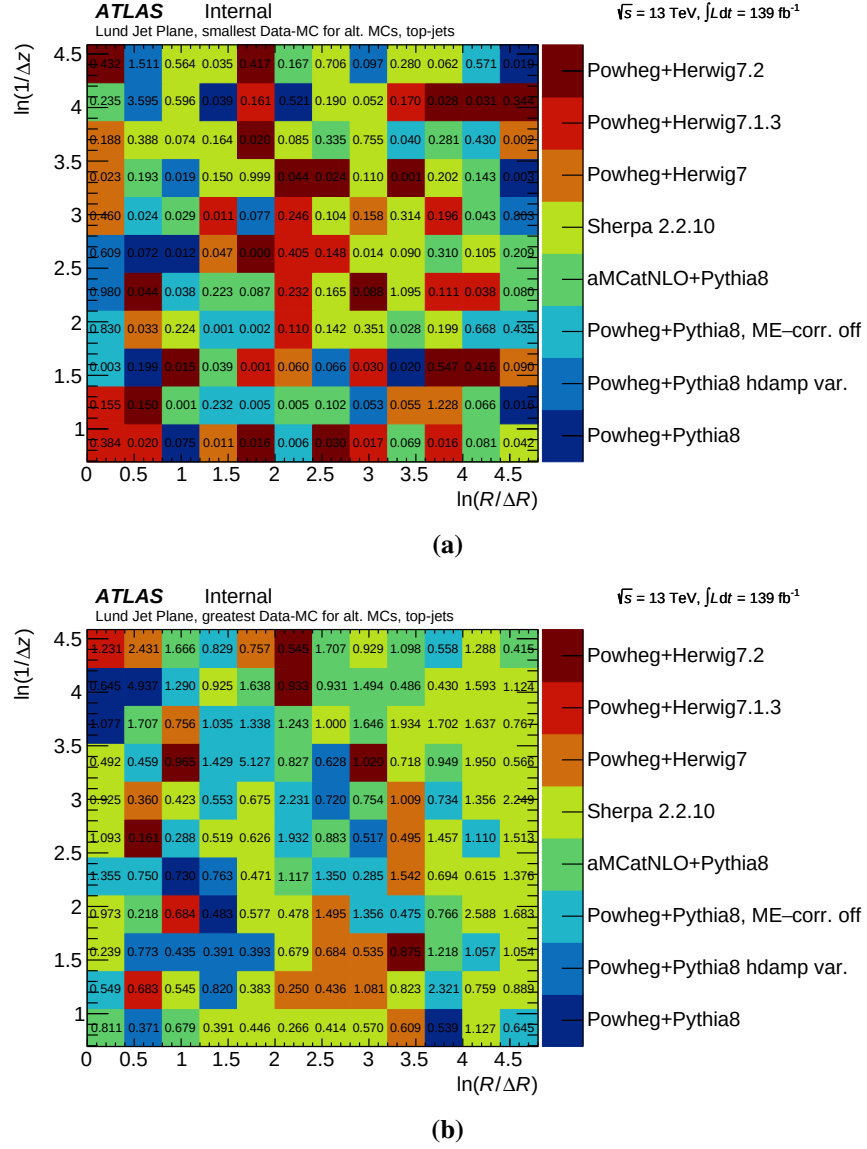


Figure B.23: Maps of the LJP for top jets showing the MC generators with the smallest (a) and greatest (b) difference to the Data. The value of $(Data - MC)/\sigma_{tot}$ for the relevant generator is printed in each bin.

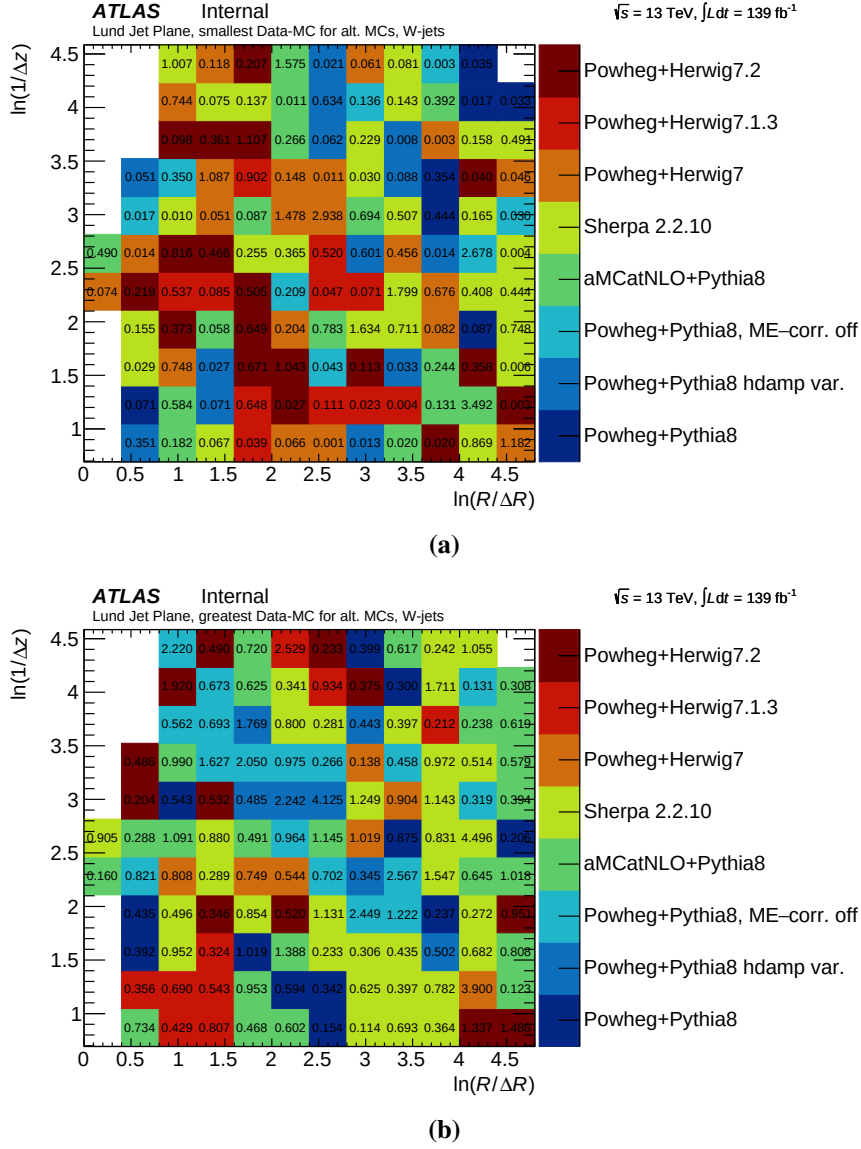


Figure B.24: Maps of the LJP for W jets showing the MC generators with the smallest (a) and greatest (b) difference to the Data. The value of $(Data - MC)/\sigma_{tot}$ for the relevant generator is printed in each bin.

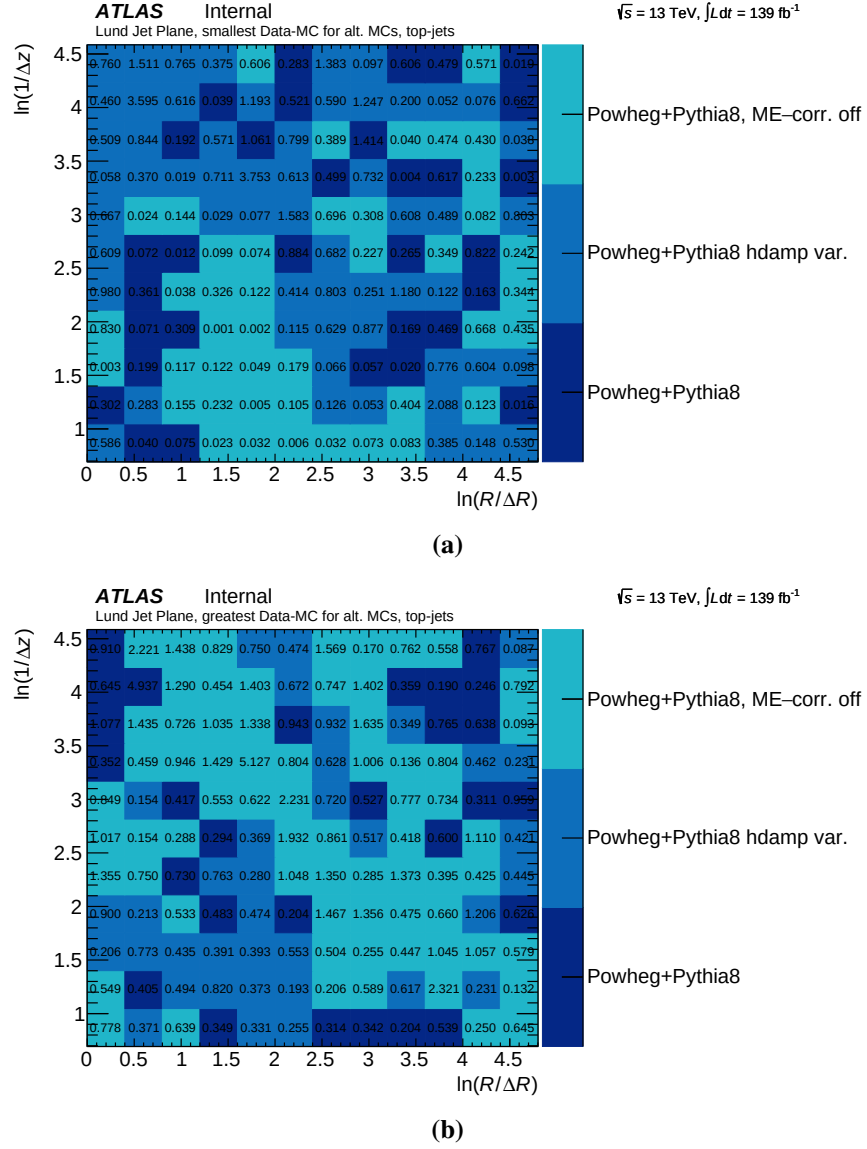


Figure B.25: Maps of the LJP for top jets showing the Powheg + Pythia MC generators with the smallest (a) and greatest (b) difference to the Data. The value of $(Data - MC)/\sigma_{tot}$ for the relevant generator is printed in each bin.

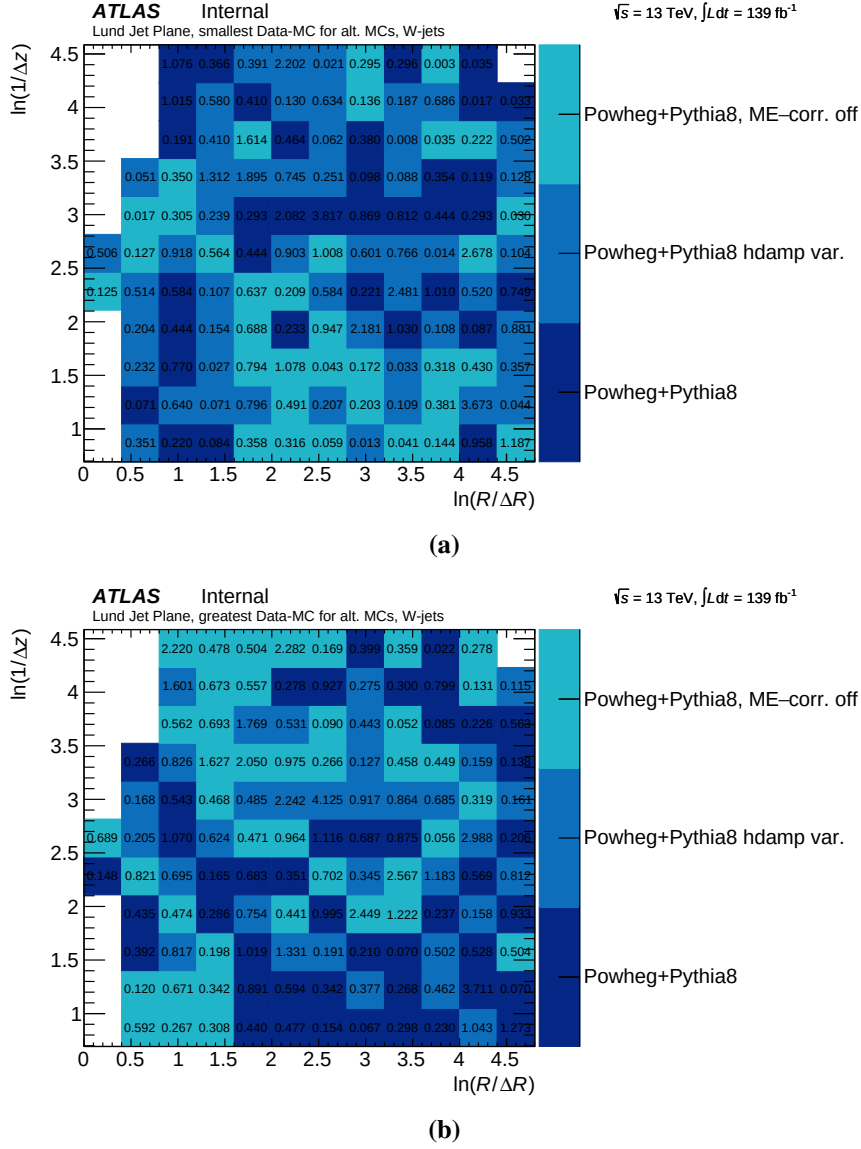


Figure B.26: Maps of the LJP for W jets showing the Powheg + Pythia MC generators with the smallest (a) and greatest (b) difference to the Data. The value of $(Data - MC)/\sigma_{tot}$ for the relevant generator is printed in each bin.

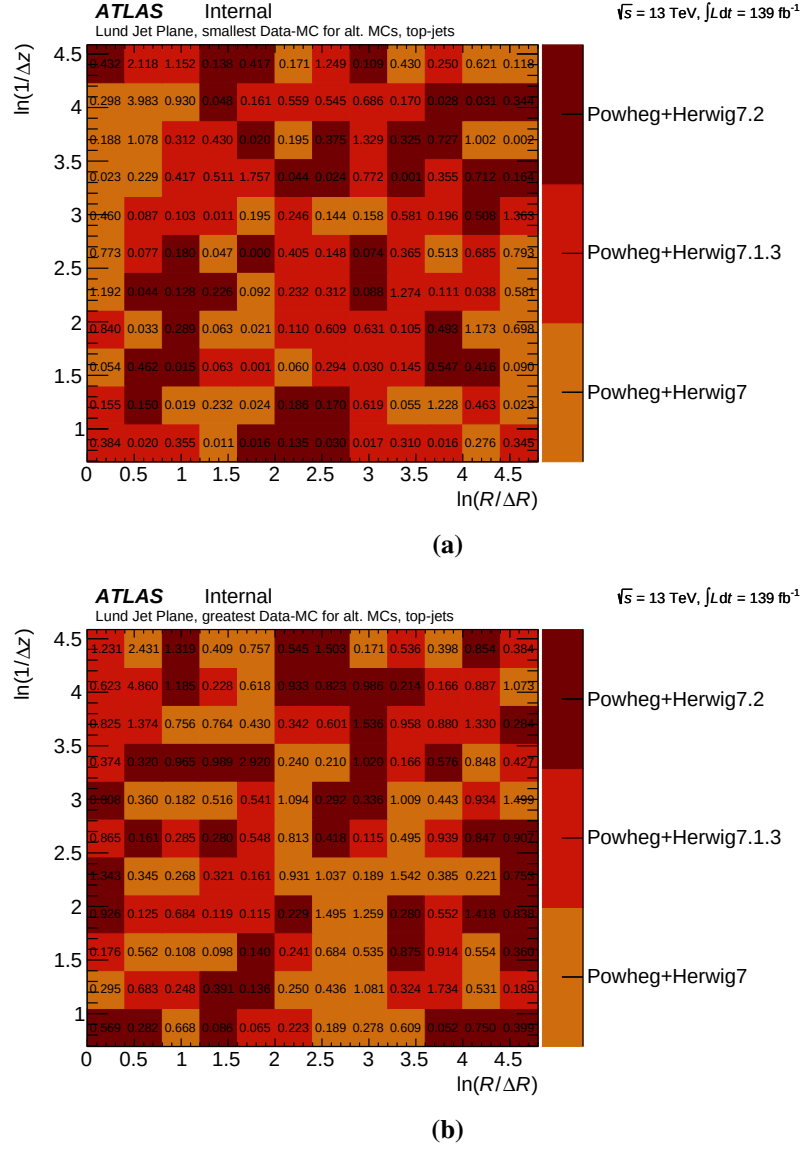


Figure B.27: Maps of the LJP for top jets showing the Powheg + Herwig MC generators with the smallest (a) and greatest (b) difference to the Data. The value of $(Data - MC)/\sigma_{tot}$ for the relevant generator is printed in each bin.

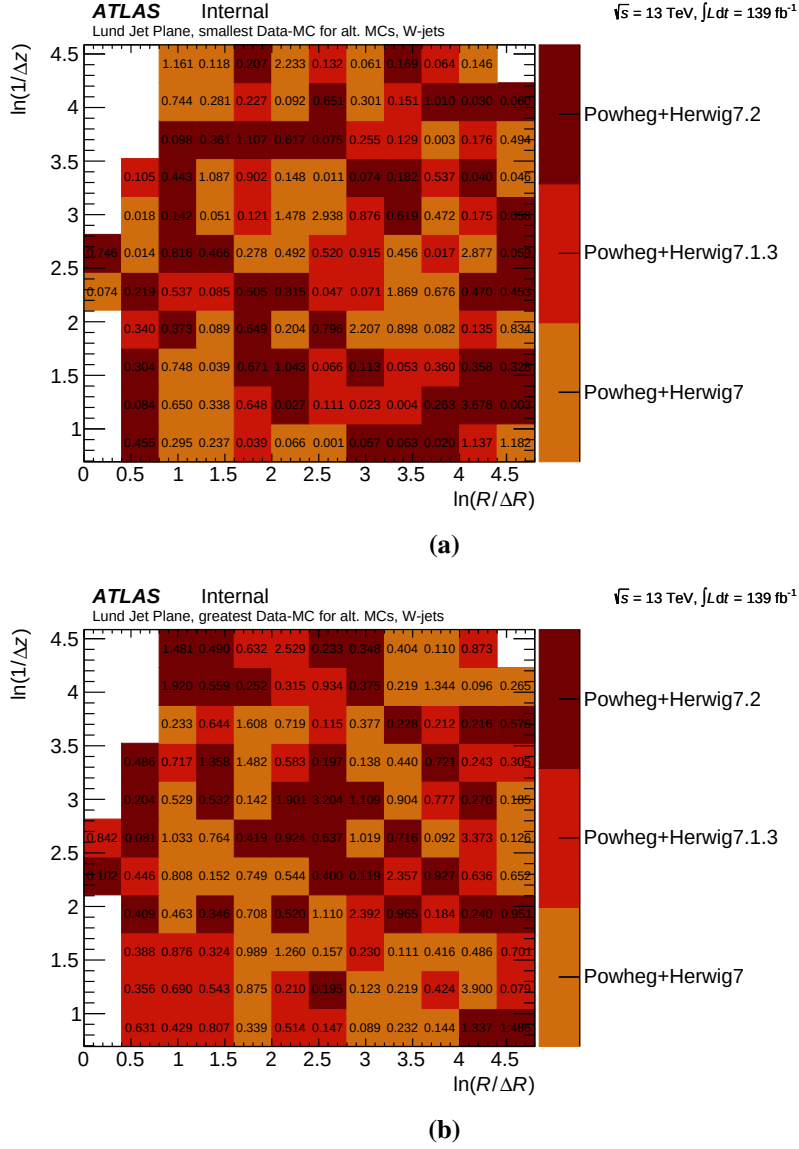


Figure B.28: Maps of the LJP for W jets showing the Powheg + Herwig MC generators with the smallest (a) and greatest (b) difference to the Data. The value of $(Data - MC)/\sigma_{tot}$ for the relevant generator is printed in each bin.

Appendix C

Deep learning W and top taggers

C.1 UFO types within W and top jets

Plots of the distribution of UFOs within W and top jets are shown in figure C.1. These were produced to study where tagger performance can be expected to improve by including neutral UFOs in the tagger inputs.

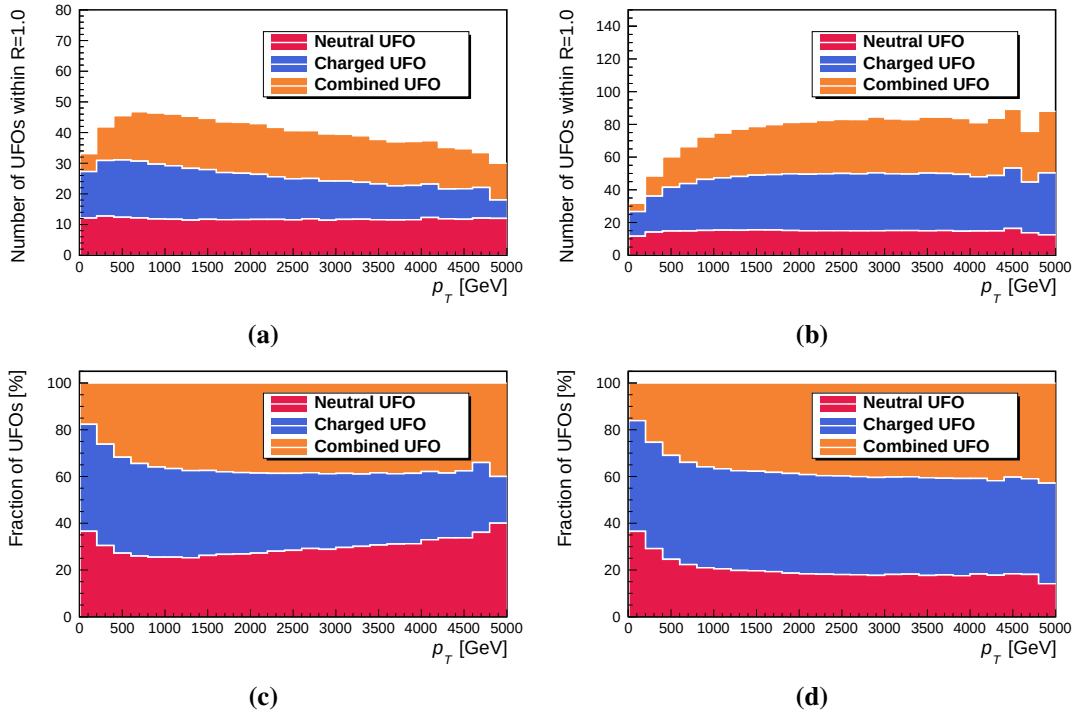


Figure C.1: Number (a,b) and fraction (c,d) of UFOs within $R = 1.0$ of W jets (a,c) and top jets (b,d).

C.2 Model architectures

This appendix contains detailed descriptions of the architectures of the Neural Networks described in chapter 5. Models that were implemented using the TensorFlow and Keras frameworks are described by printouts of the `model.summary` method provided by Keras. This method yields a list containing the layers of the model with the shape of their output and their number of trainable parameters. PyTorch model architectures are described by their class definition, where the list of layers and layer connections are provided in the `__init__` and `forward` methods.

Model: "ImgCNN"

Layer (type)	Output Shape	Param #
input_3 (InputLayer)	[(None, 40, 40)]	0
reshape_2 (Reshape)	(None, 40, 40, 1)	0
conv2d_4 (Conv2D)	(None, 38, 38, 16)	160
p_re_lu_4 (PReLU)	(None, 38, 38, 16)	23104
conv2d_5 (Conv2D)	(None, 36, 36, 16)	2320
p_re_lu_5 (PReLU)	(None, 36, 36, 16)	20736
max_pooling2d_2 (MaxPooling2D)	(None, 18, 18, 16)	0
dropout_4 (Dropout)	(None, 18, 18, 16)	0
flatten_2 (Flatten)	(None, 5184)	0
dense_4 (Dense)	(None, 128)	663680
activation_4 (Activation)	(None, 128)	0
dropout_5 (Dropout)	(None, 128)	0
dense_5 (Dense)	(None, 2)	258
activation_5 (Activation)	(None, 2)	0
Total params: 710,258		
Trainable params: 710,258		
Non-trainable params: 0		

Figure C.2: Architecture of the image CNN model as printed by the `model.summary` method provided by Keras.

Model: "LSTM"

Layer (type)	Output Shape	Param #
input_5 (InputLayer)	[(None, 30, 3)]	0
lstm (LSTM)	(None, 100)	41600
dense_7 (Dense)	(None, 2)	202
activation_7 (Activation)	(None, 2)	0

=====
 Total params: 41,802
 Trainable params: 41,802
 Non-trainable params: 0
 =====

Figure C.3: Architecture of the LSTM RNN model as printed by the `model.summary` method provided by Keras.

Model: "GRU"

Layer (type)	Output Shape	Param #
input_4 (InputLayer)	[(None, 30, 3)]	0
gru (GRU)	(None, 100)	31200
dense_6 (Dense)	(None, 2)	202
activation_6 (Activation)	(None, 2)	0

=====
 Total params: 31,402
 Trainable params: 31,402
 Non-trainable params: 0
 =====

Figure C.4: Architecture of the GRU RNN model as printed by the `model.summary` method provided by Keras.

Model: "1DCNN"

Layer (type)	Output Shape	Param #
input_6 (InputLayer)	[(None, 30, 3)]	0
conv1d (Conv1D)	(None, 28, 100)	1000
conv1d_1 (Conv1D)	(None, 26, 100)	30100
conv1d_2 (Conv1D)	(None, 24, 100)	30100
dropout_6 (Dropout)	(None, 24, 100)	0
max_pooling1d (MaxPooling1D)	(None, 12, 100)	0
flatten_3 (Flatten)	(None, 1200)	0
dense_8 (Dense)	(None, 100)	120100
dense_9 (Dense)	(None, 2)	202
activation_8 (Activation)	(None, 2)	0
Total params: 181,502		
Trainable params: 181,502		
Non-trainable params: 0		

Figure C.5: Architecture of the 1-D CNN model as printed by the `model.summary` method provided by Keras.

```

1 class LundNet(torch.nn.Module):
2     def __init__(self):
3         super(LundNet, self).__init__()
4         #EdgeConv GNN blocks
5         self.conv1 = EdgeConv(nn.Sequential(nn.Linear(6, 32),
6                                             nn.BatchNorm1d(num_features=32),
7                                             nn.ReLU()),
8                                         nn.Linear(32, 32),
9                                         nn.BatchNorm1d(num_features=32),
10                                        nn.ReLU()), aggr='mean')
11         self.conv2 = EdgeConv(nn.Sequential(nn.Linear(64, 32),
12                                             nn.BatchNorm1d(num_features=32),
13                                             nn.ReLU()),
14                                         nn.Linear(32, 32),
15                                         nn.BatchNorm1d(num_features=32),
16                                        nn.ReLU()), aggr='mean')
17         self.conv3 = EdgeConv(nn.Sequential(nn.Linear(64, 64),
18                                             nn.BatchNorm1d(num_features=64),
19                                             nn.ReLU()),
20                                         nn.Linear(64, 64),
21                                         nn.BatchNorm1d(num_features=64),
22                                        nn.ReLU()), aggr='mean')
23         self.conv4 = EdgeConv(nn.Sequential(nn.Linear(128, 64),
24                                             nn.BatchNorm1d(num_features=64),
25                                             nn.ReLU()),
26                                         nn.Linear(64, 64),
27                                         nn.BatchNorm1d(num_features=64),
28                                        nn.ReLU()), aggr='mean')
29         self.conv5 = EdgeConv(nn.Sequential(nn.Linear(128, 128),
30                                             nn.BatchNorm1d(num_features=128),
31                                             nn.ReLU()),
32                                         nn.Linear(128, 128),
33                                         nn.BatchNorm1d(num_features=128),
34                                        nn.ReLU()), aggr='mean')
35         self.conv6 = EdgeConv(nn.Sequential(nn.Linear(256, 128),
36                                             nn.BatchNorm1d(num_features=128),
37                                             nn.ReLU()),
38                                         nn.Linear(128, 128),
39                                         nn.BatchNorm1d(num_features=128),
40                                        nn.ReLU()), aggr='mean')
41         # 'Sequential' layers to combine outputs of EdgeConv blocks
42         self.sc1 = nn.Sequential(nn.Linear(3, 32, bias=False),
43                                 nn.BatchNorm1d(num_features=32))
44         self.sc2 = nn.Sequential(nn.ReLU())
45         self.sc3 = nn.Sequential(nn.Linear(32, 64, bias=False),
46                                 nn.BatchNorm1d(num_features=64))
47         self.sc5 = nn.Sequential(nn.Linear(64, 128, bias=False),
48                                 nn.BatchNorm1d(num_features=128))
49         #Fully connected layer
50         self.seq1 = nn.Sequential(nn.Linear(448, 384),
51                                 nn.BatchNorm1d(num_features=384),
52                                 nn.ReLU())
53         #Output layer
54         self.seq2 = nn.Sequential(nn.Linear(384, 256), nn.ReLU())
55         self.lin = nn.Linear(256, 1)
56
57     def forward(self, data):
58         x, edge_index, batch = data.x, data.edge_index, data.batch
59         x1 = self.conv1(x, edge_index)
60         x1 = self.sc2(self.sc1(x) + x1)
61         x2 = self.conv2(x1, edge_index)
62         x2 = self.sc2(x2 + x1)
63         x3 = self.conv3(x2, edge_index)
64         x3 = self.sc2(self.sc3(x2) + x3)
65         x4 = self.conv4(x3, edge_index)
66         x4 = self.sc2(x4 + x3)
67         x5 = self.conv5(x4, edge_index)
68         x5 = self.sc2(self.sc5(x4) + x5)
69         x6 = self.conv6(x5, edge_index)
70         x6 = self.sc2(x6 + x5)
71         #Concatenate EdgeConv outputs
72         x = torch.cat((x1, x2, x3, x4, x5, x6), dim=1)
73         x = self.seq1(x)
74         #Max Pooling
75         x = global_mean_pool(x, batch)
76         x = self.seq2(x)
77         #Dropout
78         x = F.dropout(x, p=0.1)
79         x = self.lin(x)
80         #Sigmoid activation on classifier output
81         return torch.softmax(x)
82

```

Figure C.6: Architecture of the LundNet GNN, as defined in PyTorch. The layers of the network are declared in the `__init__` method. The forward method defines how the input data is processed by the layers during forward propagation.

Bibliography

- [1] M. Thomson, *Modern Particle Physics*, Modern Particle Physics, Cambridge University Press (2013).
- [2] M. Srednicki, *Quantum field theory*, Cambridge University Press (2007).
- [3] M. Schwartz, *Quantum Field Theory and the Standard Model*, Quantum Field Theory and the Standard Model, Cambridge University Press (2014).
- [4] T. Lancaster and S. Blundell, *Quantum Field Theory for the Gifted Amateur*, OUP Oxford (2014).
- [5] Particle Data Group et al., *Review of particle physics*, *Progress of Theoretical and Experimental Physics* **2020** (2020) .
- [6] S. Haywood, *Symmetries and Conservation Laws in Particle Physics: An Introduction to Group Theory for Particle Physicists*, Imperial College Press (2011).
- [7] SUPER-KAMIOKANDE collaboration, *Evidence for oscillation of atmospheric neutrinos*, *Physical Review Letters* **81** (1998) 1562 [[hep-ex/9807003](#)].
- [8] SNO collaboration, *Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ solar neutrinos at the sudbury neutrino observatory*, *Physical Review Letters* **87** (2001) [[nucl-ex/0106015](#)].
- [9] P.W. Higgs, *Broken symmetries and the masses of gauge bosons*, *Phys. Rev. Lett.* **13** (1964) 508.
- [10] ATLAS collaboration, *Observation of $h \rightarrow b\bar{b}$ decays and vh production with the ATLAS detector*, *Physics Letters B* **786** (2018) 59–86 [[1808.08238](#)].
- [11] R.K. Ellis, W.J. Stirling and B.R. Webber, *QCD and Collider Physics*, Cambridge University Press (1996), [10.1017/CBO9780511628788](#).
- [12] J.H. Schwarz and N. Seiberg, *String theory, supersymmetry, unification, and all that*, *Rev. Mod. Phys.* **71** (1999) S112 [[hep-th/9803179](#)].
- [13] J. Ellis, *Searching for supersymmetry and its avatars*, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **377** (2019) 20190069 [[1910.07489](#)].

- [14] J.R. Ellis, J.S. Hagelin, D.V. Nanopoulos, K.A. Olive and M. Srednicki, *Supersymmetric relics from the big bang*, *Nucl. Phys. B* **238** (1984) 453.
- [15] ATLAS collaboration, *SUSY summary plots march 2022*, ATLAS Note [ATL-PHYS-PUB-2022-013](#), CERN, Geneva (2022).
- [16] CMS collaboration, “SUSY public results: Summary plots for 13 TeV.” <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS> (retrieved May 15, 2023), 2022.
- [17] ATLAS collaboration, *Dark matter summary plots for s-channel mediators*, ATLAS Note [ATL-PHYS-PUB-2020-021](#), CERN, Geneva (2020).
- [18] CMS collaboration, “CMS exotica summary plots for 13 TeV data.” https://twiki.cern.ch/twiki/bin/view/CMSPublic/SummaryPlotsEXO13TeV#CMS_Exotica_Summary_plots_for_13 (retrieved May 15, 2023), 2022.
- [19] ATLAS collaboration, *Search for single production of a vector-like t quark decaying into a higgs boson and top quark with fully hadronic final states using the ATLAS detector*, *Physical Review D* **105** (2022) [2201.07045].
- [20] ATLAS collaboration, *Search for pair-production of vector-like quarks in pp collision events at $\sqrt{s} = 13$ TeV with at least one leptonically decaying z boson and a third-generation quark with the ATLAS detector*, [2210.15413](#).
- [21] CMS collaboration, *Search of a vector-like quark $t' \rightarrow th$ in the di-photon final state*, CMS Physics Analysis Summary [CMS-PAS-B2G-21-007](#), CERN, Geneva (2022).
- [22] CMS Collaboration, *Search for pair production of vector-like quarks in leptonic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, [2209.07327](#).
- [23] ATLAS collaboration, *Search for pair-produced scalar and vector leptoquarks decaying into third-generation quarks and first- or second-generation leptons in pp collisions with the ATLAS detector*, [2210.04517](#).
- [24] ATLAS collaboration, *Summary plots from ATLAS searches for pair-produced leptoquarks*, ATLAS Note [ATL-PHYS-PUB-2022-012](#), CERN, Geneva (2022).
- [25] CMS collaboration, *Search for new physics in the tau plus missing transverse momentum final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, CMS Physics Analysis Summary [CMS-PAS-EXO-21-009](#), CERN, Geneva (2022).
- [26] CMS collaboration, *The search for a third-generation leptoquark coupling to a τ lepton and a b quark through single, pair and nonresonant production at $\sqrt{s} = 13$ TeV*, CMS Physics Analysis Summary [CMS-PAS-EXO-19-016](#), CERN, Geneva (2022).

- [27] ATLAS collaboration, *Search for resonances decaying into photon pairs in 139 fb-1 of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Physics Letters B* **822** (2021) 136651 [[2102.13405](#)].
- [28] ATLAS collaboration, *Search for TeV-scale gravity signatures in high-mass final states with leptons and jets with the ATLAS detector at $\sqrt{s} = 13$ TeV*, *Physics Letters B* **760** (2016) 520 [[1606.02265](#)].
- [29] CMS collaboration, *Search for lepton-flavor violating decays of heavy resonances and quantum black holes to $e\mu$ final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Journal of High Energy Physics* **2018** (2018) [[1802.01122](#)].
- [30] S. Marzani, G. Soyez and M. Spannowsky, *Looking inside jets: an introduction to jet substructure and boosted-object phenomenology*, *Lecture Notes in Physics* (2019) [[1901.10342](#)].
- [31] P. Bartalini, E.L. Berger, B. Blok, G. Calucci, R. Corke, M. Diehl et al., *Multi-parton interactions at the LHC*, [1111.0469](#).
- [32] Y.L. Dokshitzer, V.A. Khoze and S.I. Troyan, *On specific QCD properties of heavy quark fragmentation ('dead cone')*, *Journal of Physics G: Nuclear and Particle Physics* **17** (1991) 1602.
- [33] N. Zardoshti, *First direct observation of the dead-cone effect*, *Nuclear Physics A* **1005** (2021) 121905 [[2004.05968](#)].
- [34] M. Bähr, S. Gieseke, M.A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada et al., *Herwig++ physics and manual*, *The European Physical Journal C* **58** (2008) 639 [[0803.0883](#)].
- [35] J. Bellm, S. Gieseke, D. Grellscheid, S. Plätzer, M. Rauch, C. Reuschle et al., *Herwig 7.0/herwig++ 3.0 release note*, *The European Physical Journal C* **76** (2016) [[1512.01178](#)].
- [36] S. Schumann and F. Krauss, *A parton shower algorithm based on catani-seymour dipole factorisation*, *Journal of High Energy Physics* **2008** (2008) 038 [[0709.1027](#)].
- [37] T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, *Journal of High Energy Physics* **2006** (2006) 026 [[hep-ph/0603175](#)].
- [38] T. Sjöstrand, S. Ask, J.R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Computer Physics Communications* **191** (2015) 159 [[1410.3012](#)].
- [39] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Höche, H. Hoeth et al., *General-purpose event generators for LHC physics*, *Physics Reports* **504** (2011) 145 [[1101.2599](#)].

- [40] T. Stelzer and W. Long, *Automatic generation of tree level helicity amplitudes*, *Computer Physics Communications* **81** (1994) 357 [[hep-ph/9401258](#)].
- [41] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *Journal of High Energy Physics* **2008** (2008) 039 [[0808.3674](#)].
- [42] F. Krauss, R. Kuhn and G. Soff, *AMEGIC++ 1.0, a matrix element generator in C++*, *Journal of High Energy Physics* **2002** (2002) 044 [[hep-ph/0109036](#)].
- [43] P. Nason, *A new method for combining NLO QCD with shower monte carlo algorithms*, *Journal of High Energy Physics* **2004** (2004) 040 [[hep-ph/0409146](#)].
- [44] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, *Journal of High Energy Physics* **2002** (2002) 029 [[hep-ph/0204244](#)].
- [45] S. Frixione and B.R. Webber, *The MC@NLO 3.3 event generator*, [hep-ph/0612272](#).
- [46] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *Journal of High Energy Physics* **2014** (2014) [[1405.0301](#)].
- [47] E. Bothmann, G.S. Chahal, S. Höche, J. Krause, F. Krauss, S. Kuttimalai et al., *Event generation with sherpa 2.2*, *SciPost Physics* **7** (2019) [[1905.09127](#)].
- [48] ATLAS collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, *Eur. Phys. J. C* **76** (2016) 653. 71 p [[1608.03953](#)].
- [49] ATLAS collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS Note [ATLAS-CONF-2019-021](#), CERN, Geneva (Jun, 2019).
- [50] I. Béjar Alonso, O. Brüning, P. Fessia, L. Rossi, L. Tavian and M. Zerlauth, *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, CERN Yellow Reports: Monographs, CERN, Geneva (2020), [10.23731/CYRM-2020-0010](#).
- [51] ATLAS collaboration, *Expected performance of the ATLAS detector at the high-luminosity LHC*, ATLAS Note [ATL-PHYS-PUB-2019-005](#), CERN, Geneva (Jan, 2019).
- [52] E. Mobs, “The CERN accelerator complex - 2019. complexe des accélérateurs du CERN - 2019.” General Photo, <https://cds.cern.ch/record/2684277> (retrieved May 15, 2023), Jul, 2019.

- [53] ATLAS collaboration, *ATLAS experiment at the CERN large hadron collider*, *JINST* **3** (2008) S08003.
- [54] ATLAS collaboration, *Track reconstruction performance of the ATLAS inner detector at $\sqrt{s} = 13$ TeV*, ATLAS Note [ATL-PHYS-PUB-2015-018](#), CERN, Geneva (Jul, 2015).
- [55] A.L. Rosa, *ATLAS insertable b-layer: from construction to operation*, pre-print (2016) [[1610.01994](#)].
- [56] ATLAS collaboration, *ATLAS insertable b-layer technical design report*, Tech. Rep. [CERN-LHCC-2010-013](#), [ATLAS-TDR-19](#) (Sep, 2010).
- [57] ATLAS collaboration, *The ATLAS transition radiation tracker (trt) proportional drift tube: design and performance*, *Journal of Instrumentation* **3** (2008) P02013.
- [58] ATLAS collaboration, *Expected tracking performance of the ATLAS inner tracker at the HL-LHC*, ATLAS Note [ATL-PHYS-PUB-2019-014](#), CERN, Geneva (Mar, 2019).
- [59] ATLAS collaboration, *ATLAS pixel detector electronics and sensors*, *Journal of Instrumentation* **3** (2008) P07007.
- [60] ATLAS COLLABORATION collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC run 1*, *Eur. Phys. J. C* **77** (2017) 490 [[1603.02934](#)].
- [61] ATLAS collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, Technical design report. ATLAS, CERN, Geneva (1996).
- [62] B. Aubert et al., *Development and construction of large size signal electrodes for the ATLAS electromagnetic calorimeter*, *Nucl. Instrum. Methods Phys. Res., A* **539** (2004) 558.
- [63] D. Oliveira Damazio, *Signal processing for the ATLAS liquid argon calorimeter: studies and implementation*, ATLAS Note [ATL-LARG-PROC-2013-015](#), CERN, Geneva (Nov, 2013).
- [64] M. Aleksa et al., *ATLAS liquid argon calorimeter phase-i upgrade: Technical design report*, Tech. Rep. [CERN-LHCC-2013-017](#), [ATLAS-TDR-022](#) (Sep, 2013).
- [65] ATLAS collaboration, *ATLAS liquid argon calorimeter phase-ii upgrade: Technical design report*, Tech. Rep. [CERN-LHCC-2017-018](#), [ATLAS-TDR-027](#), CERN, Geneva (Sep, 2017).
- [66] ATLAS collaboration, *ATLAS tile calorimeter: Technical Design Report*, Technical design report. ATLAS, CERN, Geneva (1996).

- [67] ATLAS collaboration, *Technical design report for the phase-ii upgrade of the ATLAS tile calorimeter*, Tech. Rep. [CERN-LHCC-2017-019](#), [ATLAS-TDR-028](#), CERN, Geneva (Sep, 2017).
- [68] T. Barillari, *ATLAS liquid argon hadronic end-cap calorimeter: construction and selected beam test results*, [Nuclear Physics B - Proceedings Supplements](#) **150** (2006) 102–105 [[physics/0407026](#)].
- [69] A. Artamonov et al., *The ATLAS forward calorimeter*, [Journal of Instrumentation](#) **3** (2008) P02010.
- [70] ATLAS collaboration, *ATLAS magnet system: Technical design report, vol. 1-4*, ATLAS Technical design report [CERN-LHCC-97-018](#), [ATLAS-TDR-6](#), CERN, Geneva (1997).
- [71] ATLAS collaboration, *ATLAS muon spectrometer: Technical design report*, ATLAS Technical design report [CERN-LHCC-97-022](#), [ATLAS-TDR-10](#), CERN, Geneva (1997).
- [72] ATLAS collaboration, *Performance of the ATLAS level-1 topological trigger in run 2*, [The European Physical Journal C](#) **82** (2022) [[2105.01416](#)].
- [73] W. Buttinger, *The ATLAS level-1 trigger system*, ATLAS Note [ATL-DAQ-PROC-2012-024](#), CERN, Geneva (2012).
- [74] ATLAS collaboration, *Operation of the ATLAS trigger system in run 2*, [Journal of Instrumentation](#) **15** (2020) P10004 [[2007.12539](#)].
- [75] K. Bos, N. Brook, D. Duellmann, C. Eck, I. Fisk, D. Foster et al., *LHC computing grid: Technical design report. version 1.06 (20 jun 2005)*, LCG Technical design report [CERN-LHCC-2005-024](#), [LCG-TDR-001](#), CERN, Geneva (2005).
- [76] ATLAS collaboration, *Overview of ATLAS panda workload management*, [Journal of Physics: Conference Series](#) **331** (2011) 072024.
- [77] ATLAS collaboration, *ATLAS computing: technical design report*, ATLAS Technical design report [CERN-LHCC-2005-022](#), [ATLAS-TDR-17](#), CERN, Geneva (2005).
- [78] J. Elmsheuser and A. Girolamo, *Overview of the ATLAS distributed computing system*, [EPJ Web of Conferences](#) **214** (2019) 03010.
- [79] R. Brun and F. Rademakers, *Root: An object oriented data analysis framework*, [Nucl. Instrum. Meth. A](#) **389** (1997) 81.
- [80] P. Speckmayer, *Energy measurement of hadrons with the CERN ATLAS calorimeter*, PhD Thesis [CERN-THESIS-2008-051](#), Vienna, Tech. U. (2008).
- [81] D. Wicke, *A new algorithm for solving tracking ambiguities*, Tech. Rep. [DELPHI-98-163-PROG-236-TRACK-92](#) (1998).

- [82] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse and A. Salzburger, *The new ATLAS track reconstruction (newt)*, *Journal of Physics: Conference Series* **119** (2008) 032014.
- [83] ATLAS collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC run 2*, *The European Physical Journal C* **77** (2017) [1704.07983].
- [84] R.E. Kalman, *A new approach to linear filtering and prediction problems*, *Journal of Basic Engineering* **82** (1960) 35.
- [85] R. Frühwirth, *Application of kalman filtering to track and vertex fitting*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987) 444.
- [86] ATLAS collaboration, *Jet reconstruction and performance using particle flow with the ATLAS detector*, *The European Physical Journal C* **77** (2017) [1703.10485].
- [87] ATLAS collaboration, *Improving jet substructure performance in ATLAS using track-caloclusters*, ATLAS Note [ATL-PHYS-PUB-2017-015](#), CERN, Geneva (Jul, 2017).
- [88] P. Berta, M. Spousta, D.W. Miller and R. Leitner, *Particle-level pileup subtraction for jets and jet shapes*, *Journal of High Energy Physics* **2014** (2014) [1403.3108].
- [89] M. Cacciari, G.P. Salam and G. Soyez, *Softkiller, a particle-level pileup removal method*, *The European Physical Journal C* **75** (2015) [1407.0408].
- [90] ATLAS collaboration, *Optimisation of large-radius jet reconstruction for the ATLAS detector in 13 TeV proton-proton collisions*, ATLAS Note [ATLAS-CONF-2020-021](#), CERN, Geneva (Jul, 2020).
- [91] J.E. Huth et al., *Toward a standardization of jet definitions*, in *1990 DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90)*, pp. 0134–136, 12, 1990, <https://cds.cern.ch/record/217490>.
- [92] G.P. Salam, *Towards jetography*, *Eur. Phys. J.* **C67** (2010) 637 [0906.1833].
- [93] S.D. Ellis and D.E. Soper, *Successive combination jet algorithm for hadron collisions*, *Physical Review D* **48** (1993) 3160–3166 [hep-ph/9305266].
- [94] Y. Dokshitzer, G. Leder, S. Moretti and B. Webber, *Better jet clustering algorithms*, *Journal of High Energy Physics* **1997** (1997) 001–001 [hep-ph/9707323].
- [95] M. Wobisch and T. Wengler, *Hadronization corrections to jet cross sections in deep-inelastic scattering*, *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)* (1998) 270–279 [hep-ph/9907280].

- [96] M. Cacciari, G.P. Salam and G. Soyez, *The anti-ktjet clustering algorithm*, *Journal of High Energy Physics* **2008** (2008) 063–063 [[0802.1189](#)].
- [97] M. Cacciari, G.P. Salam and G. Soyez, *The catchment area of jets*, *JHEP* **04** (2008) 005 [[0802.1188](#)].
- [98] M.H. Seymour, *Tagging a heavy higgs boson*, *ECFA Large Hadron Collider (LHC) Workshop: Physics and Instrumentation* (1991) 557.
- [99] M.H. Seymour, *The average number of subjets in a hadron collider jet*, *Nucl. Phys. B* **421** (1994) 545.
- [100] J.M. Butterworth, A.R. Davison, M. Rubin and G.P. Salam, *Jet substructure as a new higgs-search channel at the large hadron collider*, *Physical Review Letters* **100** (2008) [[0802.2470](#)].
- [101] L. Asquith et al., *Jet substructure at the large hadron collider : Experimental review*, [1803.06991](#).
- [102] D. Krohn, J. Thaler and L.-T. Wang, *Jet trimming*, *JHEP* **02** (2010) 084 [[0912.1342](#)].
- [103] S.D. Ellis, C.K. Vermilion and J.R. Walsh, *Recombination algorithms and jet substructure: Pruning as a tool for heavy particle searches*, *Phys. Rev. D* **81** (2010) 094023 [[0912.0033](#)].
- [104] ATLAS collaboration, *Measurement of jet-substructure observables in top quark, w boson and light jet production in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Journal of High Energy Physics* **2019** (2019) [[1903.02942](#)].
- [105] A.J. Larkoski, G.P. Salam and J. Thaler, *Energy correlation functions for jet substructure*, *Journal of High Energy Physics* **2013** (2013) [[1305.0007](#)].
- [106] A. Larkoski, I. Moult and D. Neill, *Analytic boosted boson discrimination*, *Journal of High Energy Physics* **2016** (2016) [[1507.03018](#)].
- [107] J. Thaler and K. Van Tilburg, *Identifying boosted objects with n-subjettiness*, *Journal of High Energy Physics* **2011** (2011) [[1011.2268](#)].
- [108] A.J. Larkoski, S. Marzani, G. Soyez and J. Thaler, *Soft drop*, *Journal of High Energy Physics* **2014** (2014) [[1402.2657](#)].
- [109] M. Dasgupta, A. Fregoso, S. Marzani and G.P. Salam, *Towards an understanding of jet substructure*, *Journal of High Energy Physics* **2013** (2013) [[1307.0007](#)].
- [110] B. Andersson, G. Gustafson, L. Lönnblad and U. Pettersson, *Coherence effects in deep inelastic scattering*, *Zeitschrift für Physik C Particles and Fields* **43** (1989) 625.

- [111] L. Gribov, E. Levin and M. Ryskin, *Semihard processes in QCD*, *Physics Reports* **100** (1983) 1.
- [112] B. Andersson, *The Lund Model*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press (1998), [10.1017/CBO9780511524363](#).
- [113] A.H. Hoang, S. Mantry, A. Pathak and I.W. Stewart, *Extracting a short distance top mass with light grooming*, *Physical Review D* **100** (2019) [[1708.02586](#)].
- [114] F.A. Dreyer, G.P. Salam and G. Soyez, *The lund jet plane*, *Journal of High Energy Physics* **2018** (2018) [[1807.04758](#)].
- [115] ATLAS collaboration, *Measurement of the lund jet plane using charged particles in 13 TeV proton-proton collisions with the ATLAS detector*, *Physical Review Letters* **124** (2020) [[2004.03540](#)].
- [116] ATLAS collaboration, *Performance of top-quark and w-boson tagging with ATLAS in run 2 of the LHC*, Tech. Rep. [CERN-EP-2018-192](#), CERN, Geneva (Aug, 2018).
- [117] ATLAS collaboration, *Boosted hadronic vector boson and top quark tagging with ATLAS using Run 2 data*, in 12th International Workshop on Boosted Object Phenomenology, Reconstruction and Searches in HEP (BOOST 2020) [ATL-PHYS-PUB-2020-017](#), CERN, Geneva (2020).
- [118] S. Frixione, G. Ridolfi and P. Nason, *A positive-weight next-to-leading-order monte carlo for heavy flavour hadroproduction*, *Journal of High Energy Physics* **2007** (2007) 126 [[0707.3088](#)].
- [119] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, *Journal of High Energy Physics* **2007** (2007) 070 [[0709.2092](#)].
- [120] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower monte carlo programs: the POWHEG BOX*, *Journal of High Energy Physics* **2010** (2010) [[1002.2581](#)].
- [121] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, *Computer Physics Communications* **178** (2008) 852 [[0710.3820](#)].
- [122] ATLAS collaboration, *The pythia 8 a3 tune description of ATLAS minimum bias and inelastic measurements incorporating the donnachie-landshoff diffractive model*, ATLAS Note [ATL-PHYS-PUB-2016-017](#), CERN, Geneva (Aug, 2016).
- [123] S. Agostinelli et al., *Geant4—a simulation toolkit*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003) 250.

- [124] ATLAS collaboration, *Jet mass reconstruction with the ATLAS detector in early run 2 data*, Tech. Rep. [ATLAS-CONF-2016-035](#), CERN, Geneva (Jul, 2016).
- [125] ATLAS collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV*, *The European Physical Journal C* **76** (2016) [[1603.05598](#)].
- [126] ATLAS collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ s = 7 TeV with the ATLAS detector*, *The European Physical Journal C* **75** (2015) [[1406.0076](#)].
- [127] ATLAS collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector.*, *Phys. Rev. D* **96** (2017) 072002. 36 p [[1703.09665](#)].
- [128] ATLAS collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC run*, Tech. Rep. [ATL-PHYS-PUB-2016-012](#), CERN, Geneva (Jun, 2016).
- [129] ATLAS collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, *JHEP* **08** (2018) 089. 49 p [[1805.01845](#)].
- [130] ATLAS collaboration, *Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Journal of High Energy Physics* **2015** (2015) 1 [[1506.00962](#)].
- [131] ATLAS collaboration, *Identification of hadronically-decaying w bosons and top quarks using high-level features as input to boosted decision trees and deep neural networks in ATLAS at $\sqrt{s} = 13$ TeV*, Tech. Rep. [ATL-PHYS-PUB-2017-004](#), CERN, Geneva (Apr, 2017).
- [132] ATLAS collaboration, *Simulation of top quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV*, Tech. Rep. [ATL-PHYS-PUB-2016-004](#), CERN, Geneva (Jan, 2016).
- [133] ATLAS collaboration, *Performance of jet substructure techniques for large-r jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector.*, *JHEP* **09** (2013) 076. 83 p [[1306.4945](#)].
- [134] ATLAS collaboration, *In situ calibration of large-radius jet energy and mass in 13 TeV proton-proton collisions with the ATLAS detector*, *The European Physical Journal C* **79** (2019) [[1807.09477](#)].
- [135] ATLAS collaboration, *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV.*, *Eur. Phys. J. C* **73** (2011) 2304. 100 p [[1112.6426](#)].

- [136] A. Czarnecki, J.G. Körner and J.H. Piclum, *Helicity fractions of W bosons from top quark decays at next-to-next-to-leading order in QCD*, *Physical Review D* **81** (2010) [[1005.2625](#)].
- [137] Y. Cui, Z. Han and M.D. Schwartz, *W-jet tagging: Optimizing the identification of boosted hadronically-decaying W bosons*, *Physical Review D* **83** (2011) [[1012.2077](#)].
- [138] ATLAS collaboration, *Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV*, *The European Physical Journal C* **76** (2016) [[1510.05821](#)].
- [139] ATLAS collaboration, *Performance of the ATLAS trigger system in 2015*, *The European Physical Journal C* **77** (2017) [[1611.09661](#)].
- [140] R.D. Ball, V. Bertone, S. Carrazza, C.S. Deans, L.D. Debbio, S. Forte et al., *Parton distributions with LHC data*, *Nuclear Physics B* **867** (2013) 244 [[1207.1303](#)].
- [141] ATLAS collaboration, *ATLAS pythia 8 tunes to 7 TeV data*, ATLAS Note [ATL-PHYS-PUB-2014-021](#), CERN, Geneva (Nov, 2014).
- [142] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Physical Review D* **93** (2016) [[1506.07443](#)].
- [143] J. Bellm, S. Gieseke, D. Grellscheid, P. Kirchga  er, F. Loshaj, G. Nail et al., *Herwig 7.1 release note*, [1705.06919](#).
- [144] L.A. Harland-Lang, A.D. Martin, P. Motylinski and R.S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, *The European Physical Journal C* **75** (2015) [[1412.3989](#)].
- [145] ATLAS collaboration, *Measurement of soft-drop jet observables in pp collisions with the ATLAS detector at $s=13$ TeV*, *Physical Review D* **101** (2020) [[1912.09837](#)].
- [146] ATLAS collaboration, *Properties of jet fragmentation using charged particles measured with the ATLAS detector in pp collisions at $\sqrt{s}=13$ TeV*, *Physical Review D* **100** (2019) [[1906.09254](#)].
- [147] G. D’Agostini, *A multidimensional unfolding method based on bayes’ theorem*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **362** (1995) 487.
- [148] T. Adye, *Unfolding algorithms and tests using roounfold*, [1105.1160](#).
- [149] D. Kar and P. Sarmah, *Effect of new jet substructure measurements on pythia8 tunes*, [2101.11395](#).

- [150] ATLAS collaboration, *Measurement of the soft-drop jet mass in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Physical Review Letters* **121** (2018) [1711.08341].
- [151] A. Lifson, G.P. Salam and G. Soyez, *Calculating the primary Lund jet plane density*, *Journal of High Energy Physics* **2020** (2020) [2007.06578].
- [152] ALICE collaboration, *Physics preliminary summary: Measurement of the primary Lund plane density in pp collisions at $\sqrt{s} = 13$ TeV with ALICE*, ALICE Note [ALICE-PUBLIC-2021-002](#) (2021).
- [153] L. Cunqueiro and M. Płoskoń, *Searching for the dead cone effects with iterative declustering of heavy-flavor jets*, *Physical Review D* **99** (2019) [1812.00102].
- [154] CMS collaboration, *Measurement of the ratio $b(t \rightarrow wb)/b(t \rightarrow wq)$ in pp collisions at $\sqrt{s} = 8$ TeV*, *Physics Letters B* **736** (2014) 33 [1404.2292].
- [155] G. Avoni, M. Bruschi, G. Cabras, D. Caforio, N. Dehghanian, A. Floderus et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, *Journal of Instrumentation* **13** (2018) P07017.
- [156] ATLAS collaboration, *Performance of electron and photon triggers in ATLAS during LHC run 2*, *The European Physical Journal C* **80** (2020) [1909.00761].
- [157] ATLAS collaboration, *Performance of the ATLAS muon triggers in run 2*, *Journal of Instrumentation* **15** (2020) P09015 [2004.13447].
- [158] ATLAS collaboration, *Measurements of top-quark pair differential and double-differential cross-sections in the ℓ +jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *The European Physical Journal C* **79** (2019) [1908.07305].
- [159] ATLAS collaboration, *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data*, *The European Physical Journal C* **77** (2017) [1612.01456].
- [160] ATLAS collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, *Journal of Instrumentation* **14** (2019) P12006 [1908.00005].
- [161] 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*, *The European Physical Journal C* **79** (2019) [1902.04655].
- [162] ATLAS collaboration, *Early inner detector tracking performance in the 2015 data at $\sqrt{s} = 13$ TeV*, ATLAS Note [ATL-PHYS-PUB-2015-051](#), CERN, Geneva (Dec, 2015).

- [163] ATLAS collaboration, ATLAS, *The European Physical Journal C* [1907.05120].
- [164] R.D. Ball, V. Bertone, S. Carrazza, C.S. Deans, L.D. Debbio, S. Forte et al., *Parton distributions for the LHC run II*, *Journal of High Energy Physics* **2015** (2015) [1410.8849].
- [165] ATLAS collaboration, *Studies on top-quark monte carlo modelling for top2016*, ATLAS Note [ATL-PHYS-PUB-2016-020](#), CERN, Geneva (Sep, 2016).
- [166] D.J. Lange, *The evtgen particle decay simulation package*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **462** (2001) 152.
- [167] M. Beneke, P. Falgari, S. Klein and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, *Nuclear Physics B* **855** (2012) 695 [1109.1536].
- [168] M. Cacciari, M. Czakon, M. Mangano, A. Mitov and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, *Physics Letters B* **710** (2012) 612 [1111.5869].
- [169] P. Bärnreuther, M. Czakon and A. Mitov, *Percent-level-precision physics at the tevatron: Next-to-next-to-leading order QCD corrections to $q\bar{q} \rightarrow t\bar{t} + x$* , *Physical Review Letters* **109** (2012) [1204.5201].
- [170] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, *Journal of High Energy Physics* **2012** (2012) [1207.0236].
- [171] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, *Journal of High Energy Physics* **2013** (2013) [1210.6832].
- [172] M. Czakon, P. Fiedler and A. Mitov, *Total top-quark pair-production cross section at hadron colliders through $\mathcal{O}(\alpha_s^4)$* , *Physical Review Letters* **110** (2013) [1303.6254].
- [173] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, *Computer Physics Communications* **185** (2014) 2930 [1112.5675].
- [174] J. Butterworth, S. Carrazza, A. Cooper-Sarkar, A.D. Roeck, J. Feltesse, S. Forte et al., *Pdf*, *Journal of Physics G: Nuclear and Particle Physics* **023001** [1510.03865].
- [175] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Parton distributions for the LHC*, *The European Physical Journal C* **63** (2009) 189 [0901.0002].

- [176] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Uncertainties on α_s in global PDF analyses and implications for predicted hadronic cross sections*, *The European Physical Journal C* **64** (2009) 653 [[0905.3531](#)].
- [177] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P.M. Nadolsky, J. Pumplin et al., *New parton distributions for collider physics*, *Physical Review D* **82** (2010) [[1007.2241](#)].
- [178] J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, P. Nadolsky et al., *The CT10 next-to-next-to-leading order global analysis of QCD*, *Physical Review D* **89** (2014) [[1302.6246](#)].
- [179] J. Bellm, G. Bewick, S.F. Ravasio, S. Gieseke, D. Grellscheid, P. Kirchga  er et al., *Herwig 7.2 release note*, *The European Physical Journal C* **80** (2020) [[1912.06509](#)].
- [180] F. Cascioli, P. Maierh  fer and S. Pozzorini, *Scattering amplitudes with open loops*, *Physical Review Letters* **108** (2012) [[1111.5206](#)].
- [181] A. Denner, S. Dittmaier and L. Hofer, *Collier: A fortran-based complex one-loop library in extended regularizations*, *Computer Physics Communications* **212** (2017) 220 [[1604.06792](#)].
- [182] F. Buccioni, J.-N. Lang, J.M. Lindert, P. Maierh  fer, S. Pozzorini, H. Zhang et al., *OpenLoops 2*, *The European Physical Journal C* **79** (2019) [[1907.13071](#)].
- [183] S. H  che, F. Krauss, M. Sch  nherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *Journal of High Energy Physics* **2012** (2012) [[1111.1220](#)].
- [184] S. H  che, F. Krauss, M. Sch  nherr and F. Siegert, *QCD matrix elements + parton showers. the NLO case*, *Journal of High Energy Physics* **2013** (2013) [[1207.5030](#)].
- [185] S. Catani, F. Krauss, B.R. Webber and R. Kuhn, *QCD matrix elements + parton showers*, *Journal of High Energy Physics* **2001** (2001) 063 [[hep-ph/0109231](#)].
- [186] S. H  che, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, *Journal of High Energy Physics* **2009** (2009) 053 [[0903.1219](#)].
- [187] C. Anastasiou, L. Dixon, K. Melnikov and F. Petriello, *High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order*, *Physical Review D* **69** (2004) [[hep-ph/0312266](#)].
- [188] S. Frixione, E. Laenen, P. Motylinski and B.R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in monte carlo simulations*, *Journal of High Energy Physics* **2007** (2007) 081 [[hep-ph/0702198](#)].

- [189] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in monte carlo simulations*, *Journal of High Energy Physics* **2013** (2013) [[1212.3460](#)].
- [190] ATLAS collaboration, *Studies on top-quark monte carlo modelling with sherpa and mg5_amc@nlo*, ATLAS Note [ATL-PHYS-PUB-2017-007](#), CERN, Geneva (May, 2017).
- [191] R. Frederix, E. Re and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, *Journal of High Energy Physics* **2012** (2012) [[1207.5391](#)].
- [192] S. Alioli, P. Nason, C. Oleari and E. Re, *Nlo single-top production matched with shower in powheg: s- and t-channel contributions*, *JHEP* **09** (2009) 111 [[0907.4076](#)].
- [193] H.B. Hartanto, B. Jäger, L. Reina and D. Wackerroth, *Higgs boson production in association with top quarks in the POWHEG BOX*, *Physical Review D* **91** (2015) [[1501.04498](#)].
- [194] E. Re, *Single-top wt-channel production matched with parton showers using the POWHEG method*, *The European Physical Journal C* **71** (2011) [[1009.2450](#)].
- [195] S. Frixione, E. Laenen, P. Motylinski, C. White and B.R. Webber, *Single-top hadroproduction in association with a w boson*, *Journal of High Energy Physics* **2008** (2008) 029 [[0805.3067](#)].
- [196] ATLAS collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s} = 8\text{TeV}$ with the ATLAS detector*, ATLAS Note [ATLAS-CONF-2014-058](#), CERN, Geneva (Oct, 2014).
- [197] ATLAS collaboration, *Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in $\sqrt{s} = 8\text{TeV}$ proton-proton collisions using the ATLAS detector*, *Physical Review D* **93** (2016) [[1510.03818](#)].
- [198] ATLAS collaboration, *Measurements of top-quark pair differential cross-sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13\text{TeV}$ using the ATLAS detector*, *Journal of High Energy Physics* **2017** (2017) [[1708.00727](#)].
- [199] ATLAS collaboration, *Measurements of differential cross-sections in top-quark pair events with a high transverse momentum top quark and limits on beyond the standard model contributions to top-quark pair production with the ATLAS detector at $\sqrt{s} = 13\text{TeV}$* , *Journal of High Energy Physics* **2022** (2022) [[2202.12134](#)].
- [200] G. Guerrieri, *First run 3 data/mc plots for the measurement of the top-quark pair production cross-section in pp collisions at centre-of-mass energy of 13.6 tev with*

the atlas experiment at the lhc, in *15th International Workshop on Top Quark Physics*, September, 2022, DOI [2211.09414].

- [201] ATLAS collaboration, *Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *Physical Review D* **98** (2018) [1801.02052].
- [202] The CMS Collaboration, *Measurement of differential $t\bar{t}$ production cross sections using top quarks at large transverse momenta in pp collisions at $\sqrt{s} = 13$ TeV*, *Physical Review D* **103** (2021) [2008.07860].
- [203] H.W. Kuhn, *The hungarian method for the assignment problem*, *Naval Research Logistics Quarterly* **2** (1955) 83–97.
- [204] A. Armbruster, K. Kroeninger, B. Malaescu and F. Spano, *Practical considerations for unfolding*, ATLAS Note ATL-COM-PHYS-2014-277, CERN, Geneva (Apr, 2014).
- [205] B. Malaescu, *An iterative, dynamically stabilized method of data unfolding*, 0907.3791.
- [206] B. Malaescu, *An iterative, dynamically stabilized(ids) method of data unfolding*, in *PHYSTAT 2011*, pp. 271–275, 2011, DOI [1106.3107].
- [207] B. Efron, *Bootstrap methods: Another look at the jackknife*, *The Annals of Statistics* **7** (1979) 1 .
- [208] ATLAS collaboration, *Evaluating statistical uncertainties and correlations using the bootstrap method*, ATLAS Note ATL-PHYS-PUB-2021-011, CERN, Geneva (Apr, 2021).
- [209] ATLAS collaboration, *Measurement of dijet cross sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector*, *JHEP* **05** (2014) 059 [1312.3524].
- [210] ATLAS collaboration, *Measurement of the inclusive jet cross-sections in proton – proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **09** (2017) 020 [1706.03192].
- [211] ATLAS collaboration, *Measurement of inclusive jet and dijet cross -sections in proton –proton collisions at $\sqrt{s} = 13$,TeV with the ATLAS detector*, *JHEP* **05** (2018) 195 [1711.02692].
- [212] ATLAS collaboration, *Measurement of the ATLAS detector jet mass response using forward folding with 80 fb^{-1} of $\sqrt{s} = 13$ TeV pp data*, ATLAS Note ATLAS-CONF-2020-022, CERN, Geneva (Jul, 2020).

- [213] *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, *Physics Letters B* **810** (2020) 135797 [[2006.13076](#)].
- [214] S. Argyropoulos and T. Sjöstrand, *Effects of color reconnection on $t\bar{t}$ final states at the LHC*, *Journal of High Energy Physics* **2014** (2014) [[1407.6653](#)].
- [215] ATLAS collaboration, *Measurements of observables sensitive to colour reconnection in $t\bar{t}$ events with the ATLAS detector at $\sqrt{s} = 13$ TeV*, [2209.07874](#).
- [216] T. Sjöstrand, *Colour reconnection and its effects on precise measurements at the LHC*, in *XLIII International Symposium on Multiparticle Dynamics (ISMD13)*, Sep, 2013 [[1310.8073](#)].
- [217] F.A. Dreyer and H. Qu, *Jet tagging in the lund plane with graph networks*, *Journal of High Energy Physics* **2021** (2021) 52 [[2012.08526](#)].
- [218] M. Abadi et al., “TensorFlow: Large-scale machine learning on heterogeneous systems.” <https://www.tensorflow.org/> (retrieved May 15, 2023), 2015.
- [219] F. Chollet et al., “Keras.” <https://keras.io> (retrieved May 15, 2023), 2015.
- [220] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan et al., *Pytorch: An imperative style, high-performance deep learning library*, [1912.01703](#).
- [221] I. Goodfellow, Y. Bengio and A. Courville, “Deep learning.” MIT Press. Full text available online at <http://www.deeplearningbook.org>., 2016.
- [222] ATLAS collaboration, *Performance of mass-decorrelated jet substructure observables for hadronic two-body decay tagging in ATLAS*, Tech. Rep. [ATL-PHYS-PUB-2018-014](#), CERN, Geneva (Jul, 2018).
- [223] D. Pappadopulo, A. Thamm, R. Torre and A. Wulzer, *Heavy vector triplets: bridging theory and data*, *Journal of High Energy Physics* **2014** (2014) [[1402.4431](#)].
- [224] ATLAS collaboration, *ATLAS simulation infrastructure*, *Eur. Phys. J. C* **70** (2010) 823 [[1005.4568](#)].
- [225] J. Cogan, M. Kagan, E. Strauss and A. Schwartzman, *Jet-images: computer vision inspired techniques for jet tagging*, *Journal of High Energy Physics* **2015** (2015) [[1407.5675](#)].
- [226] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman and A. Schwartzman, *Jet-images — deep learning edition*, *Journal of High Energy Physics* **2016** (2016) [[1511.05190](#)].

- [227] P. Baldi, K. Bauer, C. Eng, P. Sadowski and D. Whiteson, *Jet substructure classification in high-energy physics with deep neural networks*, *Physical Review D* **93** (2016) [[1603.09349](#)].
- [228] L.G. Almeida, M. Backovic, M. Cliche, S.J. Lee and M. Perelstein, *Playing tag with ann: Boosted top identification with pattern recognition*, [1501.05968](#).
- [229] J. Pearkes, W. Fedorko, A. Lister and C. Gay, *Jet constituents for deep neural network based top quark tagging*, [1704.02124](#).
- [230] G. Kasieczka, T. Plehn, M. Russell and T. Schell, *Deep-learning top taggers or the end of QCD?*, *Journal of High Energy Physics* **2017** (2017) [[1701.08784](#)].
- [231] CMS Collaboration, *Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques*, *Journal of Instrumentation* **15** (2020) P06005 [[2004.08262](#)].
- [232] ATLAS collaboration, *Constituent-based top-quark tagging with the ATLAS detector*, Tech. Rep. [ATL-PHYS-PUB-2022-039](#), CERN, Geneva (2022).
- [233] S. Macaluso and D. Shih, *Pulling out all the tops with computer vision and deep learning*, *Journal of High Energy Physics* **2018** (2018) [[1803.00107](#)].
- [234] Y.-C.J. Chen, C.-W. Chiang, G. Cottin and D. Shih, *Boosted w/z tagging with jet charge and deep learning*, *Physical Review D* **101** (2020) [[1908.08256](#)].
- [235] S. Diefenbacher, H. Frost, G. Kasieczka, T. Plehn and J. Thompson, *CapsNets continuing the convolutional quest*, *SciPost Physics* **8** (2020) [[1906.11265](#)].
- [236] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, *Gradient-based learning applied to document recognition*, *Proceedings of the IEEE* **86** (1998) 2278.
- [237] K. He, X. Zhang, S. Ren and J. Sun, *Delving deep into rectifiers: Surpassing human-level performance on imagenet classification*, [1502.01852](#).
- [238] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, *Dropout: A simple way to prevent neural networks from overfitting*, *Journal of Machine Learning Research* **15** (2014) 1929.
- [239] NVIDIA, P. Vingelmann and F.H. Fitzek, “Cuda, release: 11.2.” <https://developer.nvidia.com/cuda-toolkit> (retrieved May 15, 2023), 2021.
- [240] D.P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, in *International Conference on Learning Representations (ICLR)*, 2017, DOI [[1412.6980](#)].
- [241] M. Cacciari, G.P. Salam and G. Soyez, *FastJet user manual*, *The European Physical Journal C* **72** (2012) [[1111.6097](#)].

- [242] S. Hochreiter and J. Schmidhuber, *Long short-term memory*, [*Neural computation* **9** \(1997\) 1735](#).
- [243] K. Cho, B. van Merriënboer, D. Bahdanau and Y. Bengio, *On the properties of neural machine translation: Encoder-decoder approaches*, [1409.1259](#).
- [244] D. Guest, J. Collado, P. Baldi, S.-C. Hsu, G. Urban and D. Whiteson, *Jet flavor classification in high-energy physics with deep neural networks*, [*Phys. Rev. D* **94** \(2016\) 112002](#) [[1607.08633](#)].
- [245] CMS collaboration, *Boosted jet identification using particle candidates and deep neural networks*, Tech. Rep. [CMS-DP-2017-049](#) (2017).
- [246] Y. Wang, Y. Sun, Z. Liu, S.E. Sarma, M.M. Bronstein and J.M. Solomon, *Dynamic graph cnn for learning on point clouds*, [*ACM Trans. Graph.* **38** \(2019\) 1801.07829](#).
- [247] H. Qu and L. Gouskos, *Jet tagging via particle clouds*, [*Physical Review D* **101** \(2020\) 1902.08570](#).
- [248] CMS collaboration, *Search for a massive scalar resonance decaying to a light scalar and a higgs boson in the four b quarks final state with boosted topology*, [*Physics Letters B* \(2022\) 137392](#) [[2204.12413](#)].
- [249] F.A. Dreyer, G. Soyez and A. Takacs, *Quarks and gluons in the lund plane*, [*Journal of High Energy Physics* **2022** \(2022\) 2112.09140](#).
- [250] L. Cavallini, A. Coccaro, C.K. Khosa, G. Manco, S. Marzani, F. Parodi et al., *Tagging the Higgs boson decay to bottom quarks with colour-sensitive observables and the Lund jet plane*, [*Eur. Phys. J. C* **82** \(2022\) 493](#) [[2112.09650](#)].
- [251] O. Fedkevych, C.K. Khosa, S. Marzani and F. Sforza, *Identification of b jets using QCD-inspired observables*, [*Phys. Rev. D* **107** \(2023\) 034032](#) [[2202.05082](#)].
- [252] T. Cohen, J. Roloff and C. Scherb, *Dark sector showers in the lund jet plane*, [2301.07732](#).
- [253] CMS collaboration, *Measurement of the primary lund jet plane density in proton-proton collisions at $\sqrt{s} = 13$ TeV*, CMS Physics Analysis Summary [CMS-PAS-SMP-22-007](#), CERN, Geneva (2023).