

Measurement of differential cross sections in $t\bar{t}$ and $t\bar{t}$ +jets production in the lepton+jets decay mode in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data including the extraction of α_s and *in situ* calibration of jets

Miguel Ángel Príncipe Martín,

Programa de Doctorado en Física Teórica
Departamento de Física Teórica
Facultad de Ciencias

Madrid, 2024



Universidad Autónoma
de Madrid



Measurement of differential cross sections in $t\bar{t}$ and $t\bar{t}+\text{jets}$ production in the lepton+jets decay mode in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data including the extraction of α_s and *in situ* calibration of jets

Tesis Doctoral realizada por
Miguel Ángel Príncipe Martín
presentada ante el Departamento de Física Teórica
de la Facultad de Ciencias
de la Universidad Autónoma de Madrid
para optar al Título de Doctor en Física Teórica

Tesis Doctoral dirigida por
Claudia Beatriz Glasman Kuguel
Juan Terrón Cuadrado
Profesores titulares de la Universidad Autónoma de Madrid

Departamento de Física Teórica
Facultad de Ciencias
Universidad Autónoma de Madrid

Madrid, Mayo de 2024

A mi familia y amigos

Abstract

This thesis presents measurements of absolute and normalised differential cross sections for $t\bar{t}$ production in the ℓ +jets channel in proton–proton collisions at a centre-of-mass energy of 13 TeV with 140 fb^{-1} of integrated luminosity recorded by the ATLAS detector at the LHC. Taking these measurements as input, an extraction of the strong coupling at the Z-boson mass scale and a test of the running of this coupling with the scale are performed. In addition, a study of the *in situ* calibration of the jet energy scale for jets using photon-plus-jets events in proton–proton collisions is also included.

The differential cross section is measured using the data sample including the full Run 2 of the LHC. The cross sections are measured as functions of twenty different observables at particle level. The measurements focus on jet observables for the first time with an inclusive $t\bar{t}$ selection in addition to two exclusive channels with at least one or two extra jets. For the first time, next-to-next-to-leading-order simulations are compared with differential cross sections measured by ATLAS at particle level. These measurements are useful for testing the different parton shower and hadronisation models used in Monte Carlo event generators, in addition to reducing the uncertainties in the gluon parton density functions for protons. Additionally, these measurements are sensitive to the mass of the top quarks and the strong coupling. Effective field theory fits focused on top Wilson coefficients can use the normalised cross sections as input for searches beyond the Standard Model.

The strong coupling at the scale of the Z-boson mass is measured and a test of the running of this fundamental parameter of Quantum Chromodynamics is performed. These measurements use three different absolute differential cross sections as input. The measurement of the running constitutes a test of the underlying gauge symmetry of Quantum Chromodynamics along with the number of different active flavours in a wide range of the scale in a single experiment.

The *in situ* jet energy scale is measured using photon-plus-jets events with a topology where the photon and the most energetic jet move in opposite directions in the transverse plane. The data sample used includes the full Run 2 data-taking period of the LHC. This measurement will be used to improve the jet calibration applied in ATLAS analyses at a centre-of-mass energy of 13.6 TeV during Run 3. The *in situ* jet energy scale correction are below 2% in most of the measured range of the photon transverse momentum. The uncertainty in these measurements is below 1% in most of the measured range of the photon transverse momentum.

Contents

Abstract	v
1 Introduction	1
2 Theoretical framework	9
2.1 The Standard Model of particle physics	9
2.2 Quantum Chromodynamics	10
2.2.1 The renormalisation group equation	12
2.2.2 The β function	13
2.2.3 Asymptotic freedom and confinement	14
2.2.4 Parton model and parton distribution functions	14
2.3 Electroweak Theory	18
2.3.1 The Englert–Brout–Higgs mechanism	19
2.3.2 The Cabibbo–Kobayashi–Maskawa matrix	22
2.4 Jet algorithms	23
2.5 Top-quark physics	24
2.6 Monte Carlo simulations	26
2.6.1 Hard processes	28
2.6.2 Parton density functions	29
2.6.3 Initial and final-state radiation	29
2.6.4 Beam remnants and multiple interactions	30
2.6.5 Hadronisation	30
2.6.6 Detector simulation	32
3 Experimental set-up	35
3.1 The LHC accelerator	35
3.1.1 The LHC performance	36
3.1.2 The LHC layout	37
3.1.3 The LHC operations timeline	38
3.2 The ATLAS experiment	39
3.2.1 The inner detector	41
3.2.2 The calorimeters	42
3.2.3 The muon spectrometer	45
3.2.4 The ATLAS forward detectors	47

3.2.5	The trigger system	47
3.2.6	The ATLAS experiment Run 2 timeline	48
4	Electron and photon reconstruction, calibration, identification and isolation	53
4.1	Electron and photon reconstruction	53
4.2	Electron and photon energy calibration	56
4.2.1	Electron and photon energy estimate and expected resolution from the simulation	57
4.2.2	Corrections to account for differences between data and simulations	57
4.2.3	Energy scale and resolution corrections	61
4.3	Electron identification	61
4.4	Photon identification	64
4.5	Electron and photon isolation	67
4.5.1	Electron isolation criteria	68
4.5.2	Photon isolation criteria	68
5	Muon reconstruction, calibration, identification and isolation	73
5.1	Muon reconstruction	73
5.2	Muon calibration	74
5.3	Muon identification	75
5.4	Muon isolation	80
6	Jet reconstruction, calibration and b-tagging	85
6.1	Jet reconstruction	85
6.2	Jet calibration	87
6.2.1	Jet energy scale calibration	87
6.2.2	Jet energy resolution	97
6.3	Pile-up jet suppression from tracking	102
6.4	Jet flavour-tagging	105
7	Missing transverse momentum reconstruction and performance	111
7.1	Missing transverse momentum reconstruction	111
7.2	Missing transverse momentum performance	113
8	<i>In situ</i> jet energy scale calibration using γ+jets events	117
8.1	Data and simulated samples	118
8.2	Data and simulation selection	119
8.3	<i>In situ</i> calibration	120
8.4	The MPF method	121
8.5	Fitting procedure	122
8.6	Signal purity corrections	122
8.7	Systematic uncertainties	128
8.8	Results	133

9	Measurements of top-quark pair differential cross sections in the ℓ+jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data	135
9.1	Data and Monte Carlo samples	136
9.1.1	Signal simulation samples	137
9.1.2	Background simulation samples	139
9.1.3	NNLO signal simulated samples	140
9.1.4	Theoretical uncertainties of the signal predictions	140
9.2	Object reconstruction and event selection	141
9.2.1	Detector-level object reconstruction, identification, calibration and selection	141
9.2.2	Particle-level object definition	143
9.2.3	Particle- and detector-level event selection	143
9.3	Background determination	144
9.3.1	Background determination from Monte Carlo samples	144
9.3.2	Background determination from fake leptons	145
9.4	Pseudo-top reconstruction algorithm	145
9.5	Event yields, observables and detector-level distributions	146
9.5.1	Event yields	146
9.5.2	Observables and detector-level distributions	146
9.6	Cross-section measurement procedure	148
9.6.1	Closure and stress tests	154
9.7	Systematic uncertainties	156
9.7.1	Total systematic uncertainty	159
9.8	Results	169
9.8.1	Differential cross sections for the $t\bar{t}$ inclusive channel	169
9.8.2	Differential cross sections for the $t\bar{t}+1\text{jet}$ channel	175
9.8.3	Differential cross sections for the $t\bar{t}+2\text{jets}$ channel	180
9.8.4	χ^2 and p -values	189
10	Extraction of α_s using top-quark pair differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS	191
10.1	Data and next-to-leading-order calculations	191
10.2	$\alpha_s(m_Z)$ extraction method	192
10.3	Systematics uncertainties	194
10.4	Results	194
11	Summary and conclusions	199
	Appendices	203
A	Background response fits for the <i>in situ</i> measurements	203

B	Uncertainties arising from the purity correction in the <i>in situ</i> measurements	205
C	Studies on R^{bg} for the <i>in situ</i> measurements	211
D	b -tagging optimisation	213
E	Control Regions for $t\bar{t}$ production	217
	Bibliography	240
	Agradecimientos	241

1 Introduction

The top quark is the heaviest known elementary particle, with a mass of about 172.5 GeV, and was discovered at the Tevatron proton–antiproton collider at Fermilab in 1995 [1, 2]. Almost 30 years later, the Large Hadron Collider (LHC), considered as a top factory, has produced a vast sample of top quarks allowing scientists to study in deep detail its properties. This colossal sample of top quarks produced by the LHC and recorded by ATLAS adds up to about 100 million top-quark pairs, making the top quark the subject of a large number of measurements during the Run 1 with proton–proton (pp) collisions at centre-of-mass energies of $\sqrt{s} = 7$ and 8 TeV, and the Run 2 with pp collisions at centre-of-mass energies of $\sqrt{s} = 13$ TeV. These measurements include a wide range of topics from total and differential $t\bar{t}$ cross section, to properties such as the top-quark mass, in addition to searches of new phenomena beyond the Standard Model (SM).

The production of top-quark pairs in pp collisions serves as a test for the SM at the TeV scale. Differential cross sections provides measurements that serve to perform indirect measurements of the top-quark mass or extractions of α_s . Additionally, these measures are sensitive to the existence of new physics beyond the SM at this scale. In this thesis, an analysis measuring differential cross sections focusing on jet observables is presented. These measurements are useful also for tuning the parton shower of Monte Carlo (MC) simulations.

Measurements of α_s provide an ideal testing ground for perturbative Quantum Chromodynamics (QCD) by determining α_s as a function of the scale, testing its running predicted by the renormalisation group equation. Additionally, these measurements are sensitive to new physics as the existence of new coloured fermions would imply modifications to the β -function. A precise measurement of α_s is a fundamental building block for understanding one of the interactions of the SM.

In these measurements, jets play a prominent role as part of the final state. These collimated showers of hadrons are reconstructed using clustering algorithms, which sum the four-momenta of nearby hadrons. To achieve measurements with the highest possible precision, jets undergo a calibration process that takes into account various factors. These include energy losses before the detection, the busy data-taking environment with multiple pp interactions, known as pile-up, and features of the jet reconstruction algorithm. A final step in the calibration chain involves measuring the difference in jet response and resolution between

1 Introduction

data and simulation. This information is then used to apply corresponding corrections to account for these differences.

The structure of this thesis is organised as follows. In Chapter 2, an overview of the theoretical framework underlying the high-energy process involved in the measurements in this thesis is presented. In chapter 3, a description of the experimental set-up is presented, including a brief description of the LHC and a deeper one of the ATLAS detector. In the next chapters a description of the reconstruction techniques for physical objects is presented. Electrons and photons are presented in Chapter 4, muons in Chapter 5, jets in Chapter 6, and the missing transverse momentum in Chapter 7. The next chapters of this thesis present various analysis performed using data from the ATLAS Run 2. The first of these studies is the *in situ* jet energy scale calibration using γ +jets events presented in Chapter 8. This is followed by a measurement of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 139 fb^{-1} of ATLAS data presented in Chapter 9. A third analysis covering the extraction of α_s using top-quark pair differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS is presented in Chapter 10. Finally, a summary and conclusions are presented in Chapter 11.

Some of the state-of-the-art measurements in top physics performed by the ATLAS collaboration are presented in the next paragraphs. They include total and differential $t\bar{t}$ cross section at different centre-of-mass energies, measurements of the top quarks mass and other properties, and searches of new physics.

Total $t\bar{t}$ cross section

The total $t\bar{t}$ cross section ($\sigma_{t\bar{t}}$) was measured by ATLAS at different centre-of-mass energies ($\sqrt{s}$). The increase in luminosity and in each running period allow the measurements of $\sigma_{t\bar{t}}$ to be performed with increased precision. Figure 1.1 shows $\sigma_{t\bar{t}}$ as a function of $\sqrt{s}$ and Table 1.1 shows the values of $\sigma_{t\bar{t}}$ used in this figure [3–10]. The measured $\sigma_{t\bar{t}}$ increases with increasing $\sqrt{s}$. The most precise measurements are obtained in the dileptonic channel, especially by selecting events with opposite-charge electron-muon pairs. However, measurements in other channels, such as the semi-leptonic and all-hadronic decay channels, probe the $t\bar{t}$ final state in extreme kinematic regimes by taking advantage of a higher branching ratio and a larger data sample. Theoretical predictions for $\sigma_{t\bar{t}}$ at next-to-next-to-leading-order (NNLO) accuracy in the strong coupling constant α_s , including the resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [11–15] are available from the TOP++ 2.0 program [16]. These calculations depend on $\sqrt{s}$, the mass of the top quark (m_t) and the proton parton distribution functions (PDF). For $\sqrt{s} = 13$ TeV, $m_t = 172.5$ GeV and the PDF4LHC21 PDFs [17], $\sigma_{t\bar{t}} = 832^{+20}_{-29}$ (scale) ± 35 (PDF+ α_s) ± 23 (m_t) pb, where uncertainties from the factorisation and renormalisation scale variations, the impact of uncertainties on the PDFs and α_s and the effect of the uncertainty on the value of m_t (± 1 GeV) are quoted separately. This prediction is included in Figure 1.1 and shows a good agreement with the

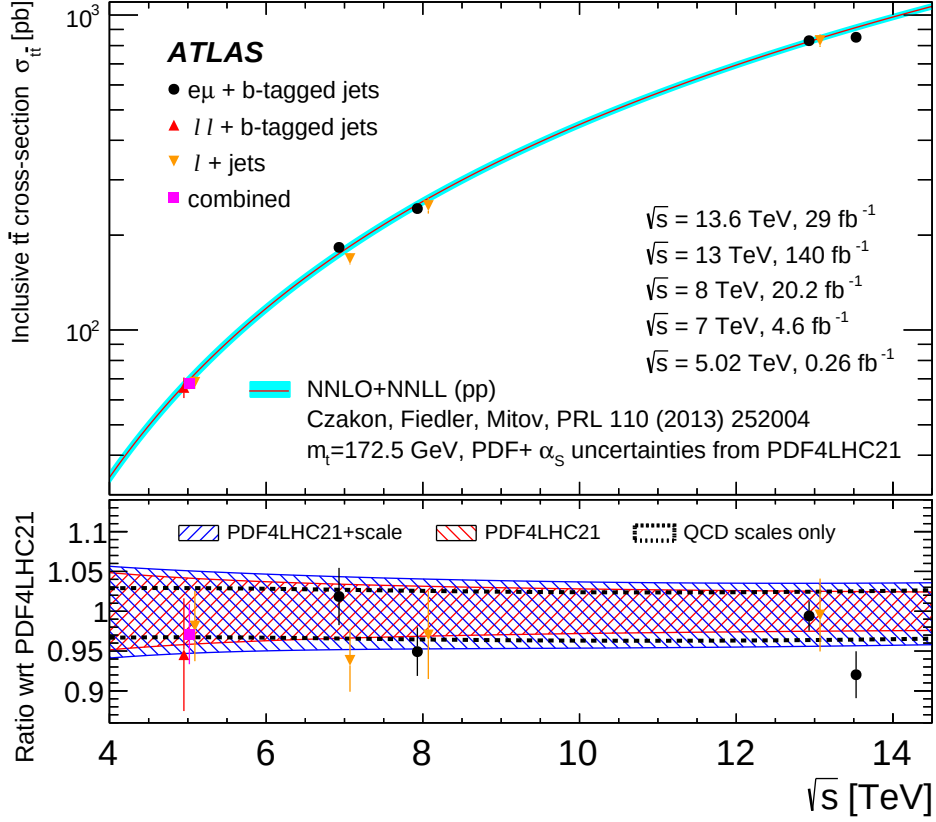


Figure 1.1: Summary of ATLAS $\sigma_{t\bar{t}}$ measurements as functions of $\sqrt{s}$. For comparison, the NNLO QCD calculation complemented with NNLL resummation from TOP++ 2.0 is also included. The theory band represents uncertainties due to renormalisation and factorisation scales, parton density functions and the strong coupling. The measurements and the theory calculation are quoted at $m_t = 172.5 \text{ GeV}$. Measurements made at the same $\sqrt{s}$ are slightly offset for clarity. From reference [10].

measurements at different $\sqrt{s}$. The measured $\sigma_{t\bar{t}}$ has reached a precision of 1.8%, exceeding the precision of the NNLO+NNLL prediction.

Top-quark mass

Since the large value of m_t is close to the electroweak (EW) scale, it plays a significant role in the production of many particles via loop diagrams. In particular, the top quark is involved in the radiative corrections to both the Higgs-boson and W -boson masses, providing a relationship that can be used for precision tests of the consistency of the SM. Measurements of m_t are categorised into direct and indirect measurements. Direct measurements include the reconstruction of the top decay products and their invariant mass, whereas in indirect

1 Introduction

Table 1.1: Values and references of the ATLAS measurements used in Figure 1.1.

Channel	$\sqrt{s}$ [TeV]	$\int \mathcal{L} dt$ [fb^{-1}]	$\sigma_{t\bar{t}}$ [pb]	Reference
Dilepton, ℓ +jets	5	0.257	67.5 ± 2.7	[3]
$e\mu$	7	4.6	183 ± 7	[4]
ℓ +jets	7	4.7	169 ± 7	[5]
$e\mu$	8	20.2	242 ± 9	[4]
ℓ +jets	8	20.2	248 ± 14	[6]
$e\mu$	13	140	829 ± 15	[7]
ℓ +jets	13	139	830 ± 39	[8]
all-jets	13	36.1	864 ± 127	[9]
$e\mu$	13.6	29	850 ± 27	[10]

measurements, m_t rely on measurements of total or differential cross sections and a comparison to the predictions to extract a value of m_t . The direct measurements usually are affected by smaller uncertainties than the indirect measurements, though they are model dependent. The advantage of the indirect measurements is that m_t is extracted from a well-defined theoretical framework, though, as mentioned above, they suffer from larger experimental and theoretical uncertainties. Thus, both types of measurements are complementary. Figure 1.2 shows a summary of the ATLAS measurements of m_t from direct and indirect measurements for different $\sqrt{s}$ and different $t\bar{t}$ channels [4, 18–22] and Figure 1.3 shows a summary of m_t measurements from ATLAS and CMS with $\sqrt{s} = 7 - 13$ TeV. The LHC combined value from the measurements available up to September 2023 is: $m_t = 172.52 \pm 0.33$ GeV, reaching a precision of 0.19%.

Searches

The top-quark, being the most massive particle observed so far, provides a fruitful ground for tests of models beyond the SM. ATLAS has made searches for new phenomena in $t\bar{t}$ events and set limits on a variety of models beyond the SM. The production of $t\bar{t}$ accompanied with large missing transverse momentum allowed the study of new top-quark partners decaying into an on-shell top (or antitop) quark and a long-lived undetected neutral particle [24]. Limits on the cross section of new sources of same-sign top-quark pair production and on the pair-production cross section of new heavy down-type quarks were obtained [25]. Production of resonances in $t\bar{t}$ final states, such as narrow Z' bosons, leptophobic topcolour Z' bosons, Kaluza-Klein gluons or broad color-octet resonances, were performed [26–30]. Limits on the production of supersymmetric top squarks were also set [4].

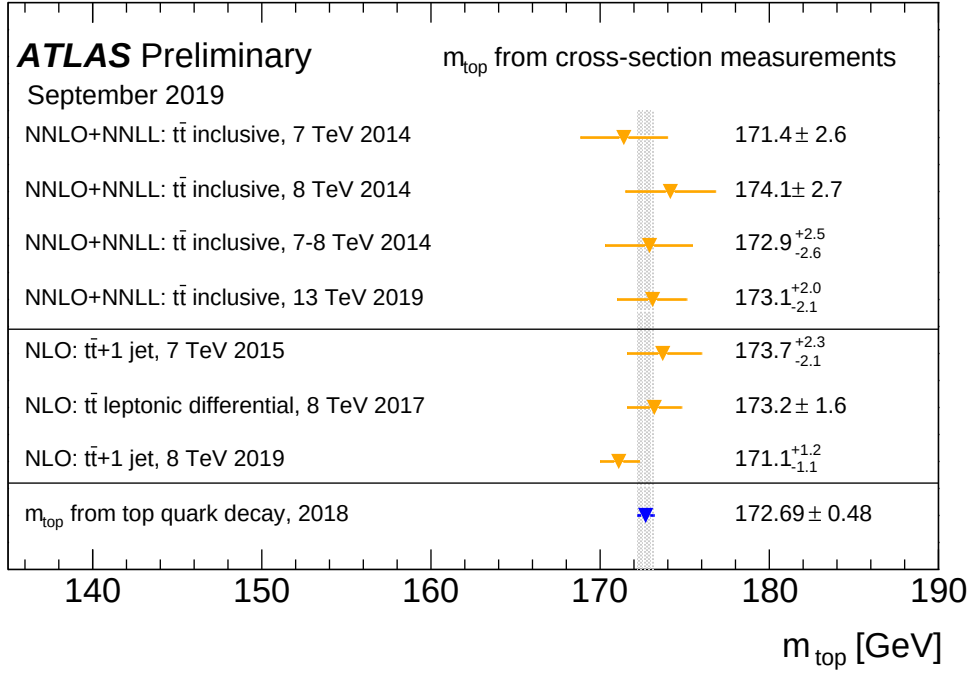


Figure 1.2: Summary of the measurements of m_t from direct and indirect measurements for different $\sqrt{s}$ and different $t\bar{t}$ channels. From reference [23].

Top-quark properties

In addition to its very large mass, the top quark has also other unique properties. Its decay width ($\Gamma_t \approx 1.4$ GeV) is larger than the QCD hadronisation scale $\Lambda \approx 250$ MeV, so the top quark decays before it can form bound states unlike other quarks. This unique feature allows its properties to be studied through its decay products. In Run 2, the ATLAS Collaboration benefitted from the larger data sample. The decay products of the top quark can be used to probe its properties or to test other SM predictions with the top quark as a tool, like checking lepton universality in W -boson decays.

The W -boson polarisation was measured in dileptonic final states [31]. The normalised differential distribution of the angle between the momentum of the charged lepton from the W boson decay and the reversed momentum of the b -quark from the top quark decay, both calculated in the W boson rest frame, was used for this measurement.

Top quark pair production at the LHC is predominantly mediated by the parity-invariant strong interaction, so the top and anti-top quarks are predicted to be produced unpolarised in the SM, while the spin of the top and anti-top quarks are expected to be correlated. As the spin correlation is transferred to its decay products, the study can be performed with the

1 Introduction

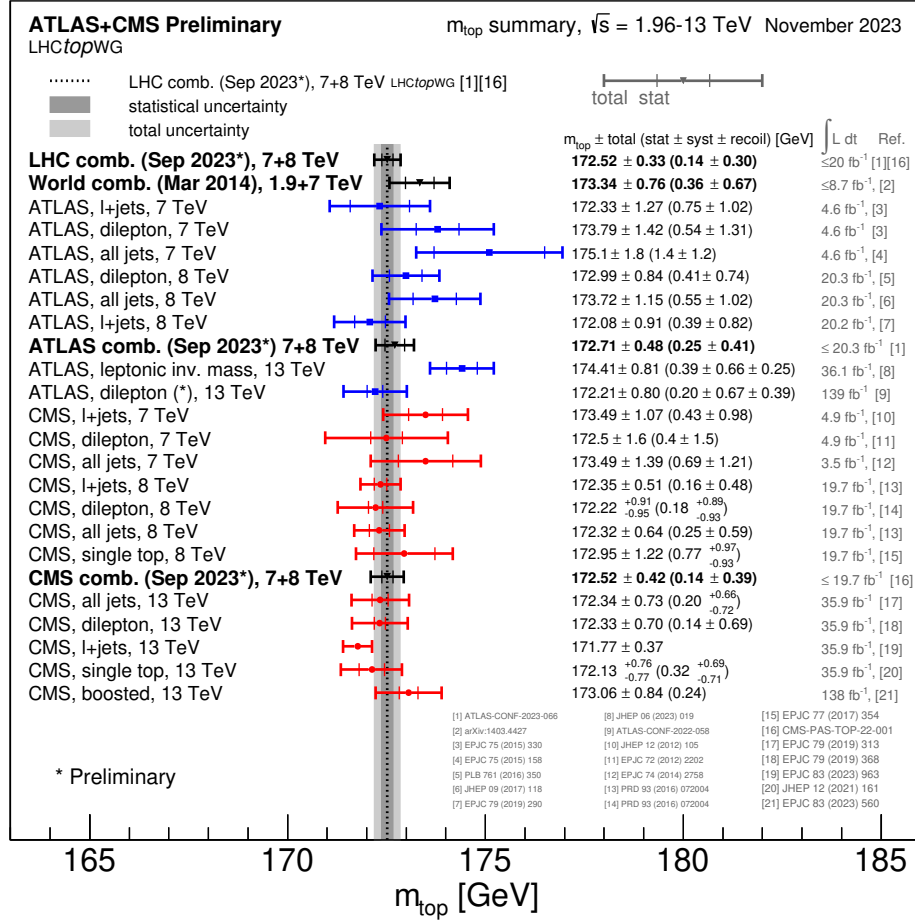


Figure 1.3: Summary of the ATLAS and CMS direct m_t measurements. For comparison, the combinations from LHC and Tevatron+LHC measurements are also included. From reference [23].

leptons from the top quark and antiquark decays, since their spin analysing power is close to unity. Hence the dilepton channel is particularly relevant to measure spin correlations. This correlation has been observed experimentally at the LHC using both $\sqrt{s} = 7$ TeV [32, 33] and 8 TeV [34, 35] collisions.

The full spin information of the top quarks is encoded in the spin density matrix. The precise measurements of the $t\bar{t}$ spin density matrix in a very restricted phase space were recently proposed as a new laboratory to study quantum entanglement. Entanglement is a feature of quantum mechanics, where two entangled particles cannot be described independently from each other. This was observed already in many systems but only at low energy scales. The LHC brings the opportunity to study this effect at high energies and in systems composed of new elementary particles. The $t\bar{t}$ pair is a two-qubit system of two spin-1/2 fermions. Entanglement in top-quark pairs can be studied via the spin correlation between the produced top-quark and anti-top-quark, using leptons from their decays as

spin analysing particles. When produced close to their production threshold, the $t\bar{t}$ pairs produced through gluon-gluon fusion are in a spin singlet state. In this case, maximum entanglement between the top quark and antiquark is expected. When the $t\bar{t}$ system has a larger mass, the entanglement is washed out (but restored at very high $t\bar{t}$ mass above around 1 TeV). A simple observable can signal the presence of entanglement: $D = -3\langle\cos\phi\rangle$, where $\langle\cos\phi\rangle$ is the average of the angle between the charged lepton directions in each of the parent top quark and antiquark rest frames. The existence of an entangled state is then demonstrated by $D < -1/3$. ATLAS performed the first study of quantum entanglement at high energy using $t\bar{t}$ events in the channel with one electron and one muon using the full Run 2 dataset at $\sqrt{s} = 13$ TeV [36].

Top quark production and decay can also be used as a tool to study generic properties of the SM, such as the test of lepton-flavour universality within top quark decays that means that the couplings of the EW gauge bosons to the three charged leptons are equal. A large and unbiased sample of W bosons can be selected in $t\bar{t}$ events and used to test this axiom by measuring the ratio of W -boson decay rates into muons and τ -leptons. The ATLAS analysis [37] measured a value that is in agreement with the hypothesis of universal lepton couplings as postulated in the SM. This is the only such measurement from the LHC so far and yielded twice the precision of previous measurements. Final states with top quarks and jets provide a testing ground of colour flow and reconnection; the distribution and orientation of energy inside jets is predicted to be an experimental handle on colour connections between the hard-scatter quarks and gluons initiating the jets, such as the jet-pull angle [38], the jet-pull magnitude [39] and distributions as functions of the charged-particle multiplicity [40].

Other measurements, such as charge [41–48], charge–parity (CP) [48] and energy asymmetries [49], were performed to test the SM. These measurements are also sensitive to physics beyond the SM and were interpreted by the ATLAS Collaboration within the framework of effective field theories and limits on some of the Wilson coefficients were set. The branching ratio and decay width of the top quarks were measured at $\sqrt{s} = 7$ TeV [50] and 8 TeV [51].

Differential top-quark pair cross-section measurements

Measurements of differential cross sections for $t\bar{t}$ production can be used to extract fundamental parameters of the SM, such as the top-quark mass and the strong coupling constant α_s , and to derive constraints on the parton distribution functions (PDF). These measurements are also sensitive to the existence of new physics beyond the SM, which would appear as deviations of the measurements with respect to the SM predictions. In addition, for all searches for new physics in which $t\bar{t}$ production constitutes a background, the differential cross sections are particularly useful to highlight the differences in MC models and can be used to tune their parameters. ATLAS measured a large variety of differential cross sections in different final states, kinematic regimes and as functions of different observables,

1 Introduction

both at particle and parton levels at $\sqrt{s} = 7$ TeV [52–54], 8 TeV [21, 55–57] and 13 TeV [9, 18, 58–63]. Additional jet activity and heavy-flavour production in $t\bar{t}$ final states was also studied [64–73].

All these studies are sensitive to different approaches for QCD radiation and for models of hadronisation and parton shower. Different MC predictions show different levels of agreement with the measurements of observables related to the production of additional jets, and therefore these measurements turn out to be very useful for testing MC models of $t\bar{t}$ production. In particular, the measurements of additional jet activity in the lepton+jets channel as presented in this thesis, despite the worsening of the precision compared to dilepton channels, provide a larger sample and the ability to fully reconstruct the top-quark kinematics; in addition, the tails of the differential distributions as functions of transverse momenta of the jets and their invariant masses provide sensitivity to EW effects.

2 Theoretical framework

A review of the most relevant aspects of the theoretical framework related to top physics is presented in this chapter. The first part of the chapter introduces the SM, including its building blocks, QCD [74] and EW theory [75–77]. Following this, a description of jet algorithms used in high energy-physics is presented [78]. Finally, an overview of top-quark physics and a presentation of the MC event generators used in the analysis is given [79].

2.1 The Standard Model of particle physics

The SM of particle physics is the theory describing all known elementary particles and their interactions. Developed in the early 1970s, it combines QCD, which describes the strong interaction between quarks mediated by gluons, and the EW theory, which unifies the electromagnetic and weak interactions and incorporates the mechanism of spontaneous symmetry breaking [80, 81]. Since then, it has become a theory capable of predicting a wide variety of phenomena, with many experimental tests successfully validating its predictions.

According to the SM, elementary particles are divided into two types: fermions and bosons. Fermions have spin $\frac{1}{2}$, while bosons have integer spin. Fermions are further divided into quarks and leptons, each organised into three families or generations. Leptons are classified according to their charge (Q) and three flavour numbers: electron (L_e), muon (L_μ), and tau (L_τ) numbers, as presented in Table 2.1.

Table 2.1: Classification of leptons organised into three generations with their quantum numbers.

	ℓ	Q	L_e	L_μ	L_τ
First generation	e	-1	1	0	0
	ν_e	0	1	0	0
Second generation	μ	-1	0	1	0
	ν_μ	0	0	1	0
Third generation	τ	-1	0	0	1
	ν_τ	0	0	0	1

2 Theoretical framework

Table 2.2: Classification of quarks organised into three generations with their quantum numbers.

	q	Q	S	C	B	T
First generation	u	2/3	0	0	0	0
	d	-1/3	0	0	0	0
Second generation	c	2/3	0	1	0	0
	s	-1/3	1	0	0	0
Third generation	t	2/3	0	0	0	1
	b	-1/3	0	0	1	0

Similarly, quarks are classified according to their charge and four flavour numbers named strangeness (S), charm (C), beauty or bottomness (B) and truth or topness (T). Table 2.2 shows their classification along with their quantum numbers.

Additionally, each quark and lepton has an antiparticle partner with opposite quantum numbers. Quarks also come in three different colours due to the strong interaction.

The bosonic content of the SM is mainly determined by the local gauge symmetry that defines the theory. The symmetry group that defines the SM is:

$$SU(3)_C \times SU(2)_L \times U(1)_Y.$$

The $SU(3)$ group is responsible for the strong interaction between quarks, with its generators forming an octet that manifest as the eight gluons of the SM. Similarly, the EW interaction, described by the $SU(2)_L \times U(1)_Y$ group, is mediated by the W bosons, the Z boson, and the photon. Gluons, W bosons, the Z boson and the photon are vector bosons with spin 1. Additionally, the EW symmetry breaking mechanism introduces another boson, the scalar Higgs boson, with spin 0.

Figure 2.1 shows the table of elementary particles of the SM and the interactions affecting each group. The strong interaction, mediated by gluons, affects colour charged particles, i. e. gluons and quarks. The electromagnetic force, mediated by photons, affects charged particles such as quarks, charged leptons, and the W boson. The weak interaction, carried by W and Z bosons, affects all fermions. The Higgs boson interacts with massive fermions and bosons. Outside the SM, the graviton is hypothesised to be responsible for gravity.

2.2 Quantum Chromodynamics

QCD is the theory that describes strong gauge interactions and is one of the fundamental components of the SM. It is a quantum field theory that describes the interaction between quarks and gluons, based on local gauge invariance represented by the $SU(3)_C$ group.

2.2 Quantum Chromodynamics

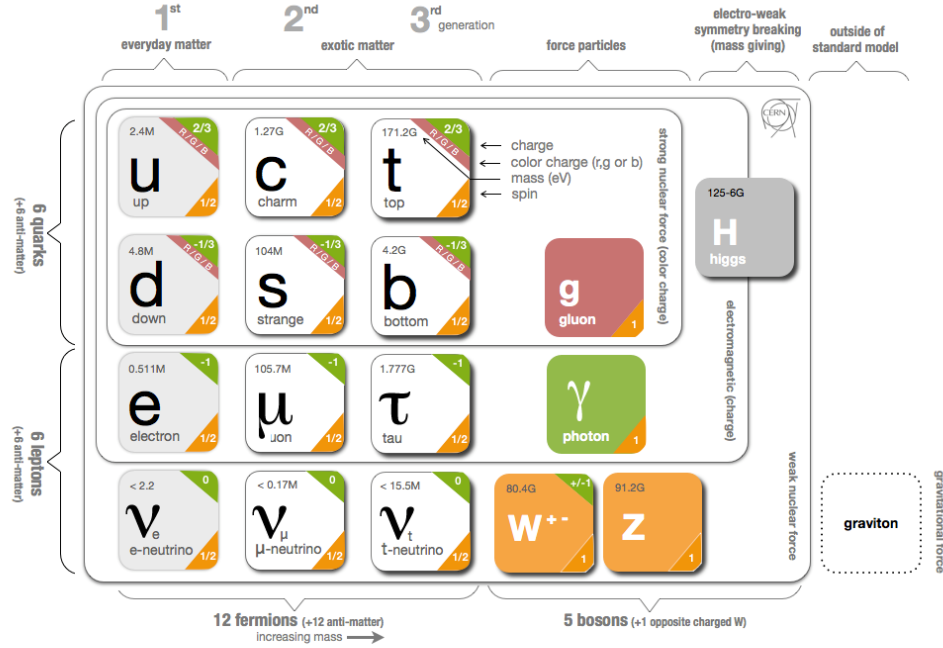


Figure 2.1: The elementary particles of the SM and the interactions affecting each group. From reference [82].

The development of QCD stemmed from the need to explain the large number of hadrons discovered in the preceding decades. In 1964, the quark model [83, 84] was proposed as a solution to disentangle the cause behind such a large number of particles that did not seem fundamental. This model introduces three fundamental particles, quarks, organised into a $SU(3)_F$ flavour group. Quarks combine to form hadrons in pairs and triplets. Later that year, a fourth flavour was added to the model [85]. Finally, in 1975, a model with six quark flavours was proposed to explain the complete spectrum of heavy mesons and baryons [86].

However, the quark model initially posed a challenge due to its contradiction with the Pauli exclusion principle. Some hadrons, such as Ω^- (composed of three strange quarks) and Δ^{++} (composed of three up quarks), have three identical quarks with parallel spins, so every quark should have every quantum number identical, seemingly violating the principle. The solution to this came between 1964 and 1965 with the introduction of an additional quantum degree of freedom for quarks based on the $SU(3)_C$ colour gauge symmetry [87, 88].

This led to the formulation of a gauge theory based on the $SU(3)_C$ symmetry, known as

2 Theoretical framework

QCD, in 1973 [89]. The QCD Lagrangian¹ is given by:

$$\mathcal{L}_{\text{QCD}} = \sum_{\text{flavours}} \bar{q}_i \left(i\gamma^\mu (D_\mu)_{ij} - m\delta_{ij} \right) q_j - \frac{1}{4} G_{\mu\nu}^a G_a^{\mu\nu}. \quad (2.1)$$

In the first term of Equation 2.1, q represents the quark field spinors with indices i and j running from 1 to 3 for the three colours of the fundamental representation of $SU(3)_C$. γ^μ are the Dirac matrices, δ_{ij} is the Kronecker function, and $(D_\mu)_{ij}$ is the gauge covariant derivative defined as:

$$(D_\mu)_{ij} = \partial_\mu \delta_{ij} + ig \frac{1}{2} (\lambda_a)_{ij} \mathcal{A}_\mu^a,$$

where λ_a are the Gell-Mann matrices and $\mathcal{A}_\mu^a$ are the gluon fields in the adjoint representation of $SU(3)_C$ with a running from 1 to 8.

In the second term, $G_{\mu\nu}^a$ represents the gluon field strength tensor that is defined as:

$$G_{\mu\nu}^a = \partial_\mu \mathcal{A}_\nu^a - \partial_\nu \mathcal{A}_\mu^a + gf^{abc} \mathcal{A}_\mu^b \mathcal{A}_\nu^c,$$

where f^{abc} are the structure constants of $SU(3)_C$.

The constant m corresponds to the quark mass and g is the coupling constant of the theory.

2.2.1 The renormalisation group equation

When dealing with QCD using a perturbative approach, including Feynman diagrams with loops lead to integrals over the particle momenta in the loop that diverge. To obtain finite results, a process called renormalisation is performed, absorbing these divergences into physical quantities like coupling constants.

This process involves introducing a new dimensional scale parameter, μ_R , through a specified (though arbitrary) prescription. This leads to renormalised quantities explicitly dependant on μ_R . The transformations of renormalised operators as μ_R varies form a Lie group known as the Renormalisation Group [90]. The equation expressing physics invariance under changes of the parameter μ is called the renormalisation group equation (RGE).

Considering a dimensionless physical observable R , it can be calculated as a perturbation series in the coupling $\alpha_s = g^2/4\pi$ introducing a renormalisation scale μ_R (the energy at which subtractions removing the ultraviolet divergences are performed). Generally, R

¹ The classical term of the Lagrangian is shown. For perturbative analysis of QCD, the gauge-fixing and ghost terms must be considered.

depends only on the ratio Q^2/μ_R^2 and the renormalised coupling, but because μ_R is an arbitrary parameter and the Lagrangian makes no mention of it, R must be independent of μ_R . Therefore,

$$\mu_R^2 \frac{d}{d\mu_R^2} R\left(\frac{Q^2}{\mu_R^2}, \alpha_s\right) \equiv \left[\mu_R^2 \frac{\partial}{\partial \mu_R^2} + \mu_R^2 \frac{\partial \alpha_s}{\partial \mu_R^2} \frac{\partial}{\partial \alpha_s} \right] R = 0. \quad (2.2)$$

Equation 2.2 is the RGE assuming massless quarks and may be rewritten by taking $t = \ln(Q^2/\mu_R^2)$ and $\beta(\alpha_s) = \mu_R^2 \partial \alpha_s / \partial \mu_R^2$ as

$$\left[-\frac{\partial}{\partial t} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} \right] R(e^t, \alpha_s) = 0.$$

This first order partial differential equation is solved by implicitly defining the running coupling, $\alpha_s(Q^2)$, as follows:

$$t = \int_{\alpha_s(\mu_R^2)}^{\alpha_s(Q^2)} \frac{dx}{\beta(x)},$$

which, after differentiation (with $\alpha_s(\mu_R^2)$ simplified as α_s), yields

$$\frac{\partial \alpha_s(Q^2)}{\partial t} = \beta(\alpha_s(Q^2)) \quad \text{or} \quad \frac{\partial \alpha_s(Q^2)}{\partial \alpha_s} = \frac{\beta(\alpha_s(Q^2))}{\beta(\alpha_s)},$$

where $R(1, \alpha_s(Q^2))$ is a solution of Equation 2.2, and all the scale dependence in R is absorbed by the running of the coupling constant $\alpha_s(Q^2)$.

2.2.2 The β function

The function $\beta(\alpha_s)$ introduced in Section 2.2.1 defines the behaviour of the running coupling at different scales via the RGE. In QCD, the β function can be perturbatively expanded as:

$$\beta(\alpha_s) = -b\alpha_s^2 (1 + b'\alpha_s + \mathcal{O}(\alpha_s^2)),$$

where

$$b = \frac{33 - 2n_f}{12\pi}, \quad \text{and} \quad b' = \frac{153 - 19n_f}{2\pi(33 - 2n_f)},$$

2 Theoretical framework

with n_f being the number of quark flavours contributing in the loops. This result contrasts with the β function in QED, which is positive, and is a direct consequence of the non-Abelian nature of QCD.

In the perturbative region, considering only the 1-loop term of the expansion, i. e. neglecting b' and higher-order terms, into the RGE:

$$Q^2 \frac{\partial \alpha_s(Q^2)}{\partial Q^2} = b \alpha_s^2(Q^2),$$

this can be solved as:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \alpha_s(\mu_R^2) b t}.$$

In this expression, as t becomes very large, the running coupling decreases towards zero. This property is called asymptotic freedom, and is caused by the negative sign of the β function.

2.2.3 Asymptotic freedom and confinement

As discussed in Section 2.2.2, at high energy, the coupling constant, α_s , becomes very small, allowing quarks and gluons to behave as free particles. Conversely, at low energies or large distances, α_s becomes very large, making the perturbation theory unusable. This is the reason why quarks and gluons are not observed isolated, leading to the phenomenon known as confinement. Figure 2.2 illustrates the running of the coupling computed at five loops compared to several experimental measurements. Due to the running behaviour of the coupling constant, a specific scale is chosen as a reference for measurements, typically $Q^2 = m_Z^2$, which is large enough to remain in the perturbative regime. The world average value of $\alpha_s(m_Z^2)$ from the latest PDG review [91] is $\alpha_s(m_Z^2) = 0.1180 \pm 0.0009$.

2.2.4 Parton model and parton distribution functions

Simultaneously with the development of the quark model another paradigm regarding hadron structure, known as the parton model [92], was devised. According to the parton model, hadron interactions are caused by interactions with partons, the constituents of the hadron. This model was successfully applied to electron-proton deep inelastic scattering (DIS) [93] and experimentally tested in colliders such as HERA.

The proton structure can be determined by colliding electrons and protons at much higher energies than the proton mass, known as DIS. Figure 2.3 represents an example of a Feynman

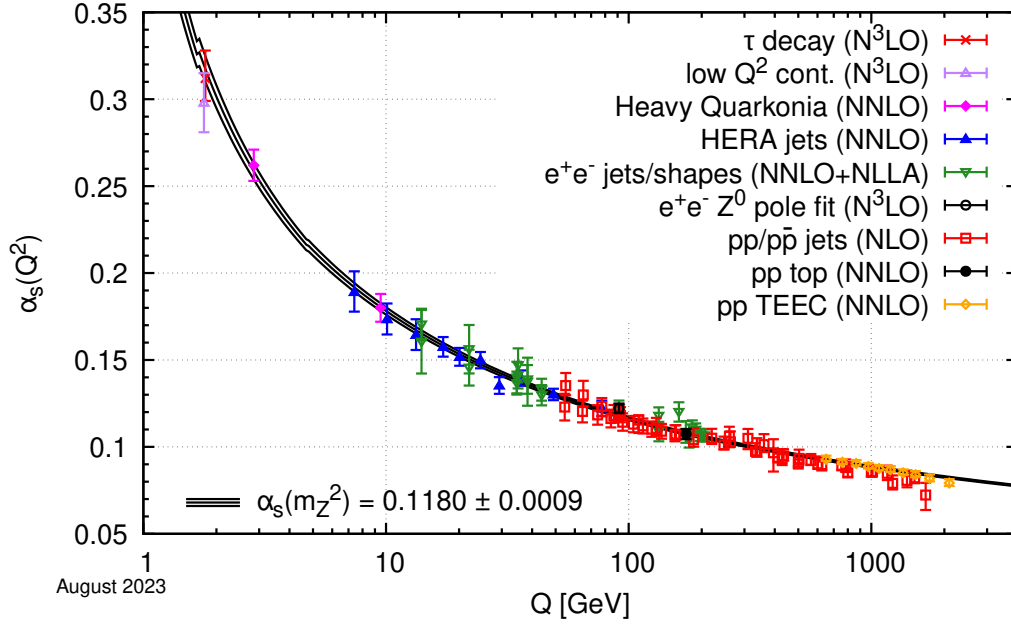


Figure 2.2: Summary of determinations of α_s as a function of the energy scale Q compared to the running of the coupling computed at five loops, using the current PDG average as input. From reference [91].

diagram for such processes. The cross section of electron-proton collisions can be parametrised in terms of the proton structure. Neglecting the proton mass, the differential cross section is written as:

$$\frac{d^2\sigma}{dx dQ^2} = \frac{4\pi\alpha_s}{Q^4} \left\{ \left[1 + (1-y)^2 \right] F_1 + \frac{1-y}{x} (F_2 - 2xF_1) \right\}$$

where $Q^2 = -q^2$ with q as the four-momentum of the virtual photon, $x = Q^2 / (2p \cdot q)$ with p as the four momentum of the proton, $y = q \cdot p / (k \cdot p)$ with k as the four-momentum of the colliding electron, and F_i are the structure functions characterising the proton structure. These functions depend on x and Q^2 , but in the Bjorken limit, i. e. as $Q^2 \rightarrow \infty$, the dependence remains only on x . This implies that the virtual photon interacts with pointlike constituents, as the dependence on Q^2 can be interpreted as a length scale characterising the size of the constituents. Both structure functions are related by the Callan-Gross equation, which arises from the fermionic nature of partons:

$$2xF_1 = F_2$$

and can be computed as a combination of the probability distribution for each quark flavour as:

2 Theoretical framework

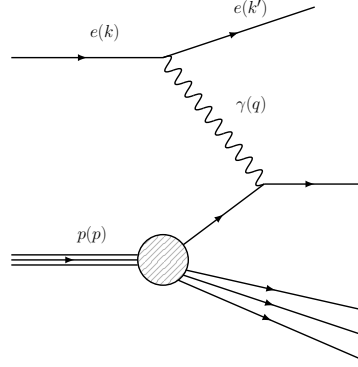


Figure 2.3: Feynman diagram of deep inelastic scattering of an electron with momentum k and a proton with momentum p , mediated by a photon with momentum q . The final state of the scattering consists of the electron with momentum k' , the parton scattered by the photon, and the remaining proton fragments.

$$F_2(x) = 2xF_1(x) = \sum_{q,\bar{q}} e_q^2 x q(x),$$

where e_q is the quark charge and $q(x)$ are the quark probability distributions, which represent the probability of a quark q carrying a momentum fraction between x and $x + dx$ from the proton. These distributions are commonly referred to as parton distribution functions (PDFs). However, DIS experiments show that only approximately 50% of the proton momentum is carried by these charged partons. The presence of neutral partons is resolved in QCD with the introduction of gluons as mediators of the strong force.

Despite the prediction in the Bjorken limit, when including higher-order terms in the perturbative QCD (pQCD) expansion, contributions proportional to $\alpha_s \ln Q^2$ from gluons emitted with high transverse momentum appear, and the scaling is broken. Additionally, other divergences from collinear emissions also contribute. The latter type of contribution is not subject to theorems for cancellation due to its origin from non-perturbative effects. Expanding the PDFs similarly to the renormalisation of the coupling constant, the renormalised PDFs are given by:

$$q(x, \mu_F^2) = q_0(x) + \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\zeta}{\zeta} q_0(\zeta) \left\{ P_{qq}\left(\frac{x}{\zeta}\right) \ln\left(\frac{\mu_F^2}{\kappa^2}\right) + C_q\left(\frac{x}{\zeta}\right) \right\} + \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\zeta}{\zeta} g_0(\zeta) \left\{ P_{qg}\left(\frac{x}{\zeta}\right) \ln\left(\frac{\mu_F^2}{\kappa^2}\right) + C_g\left(\frac{x}{\zeta}\right) \right\} + \dots,$$

2.2 Quantum Chromodynamics

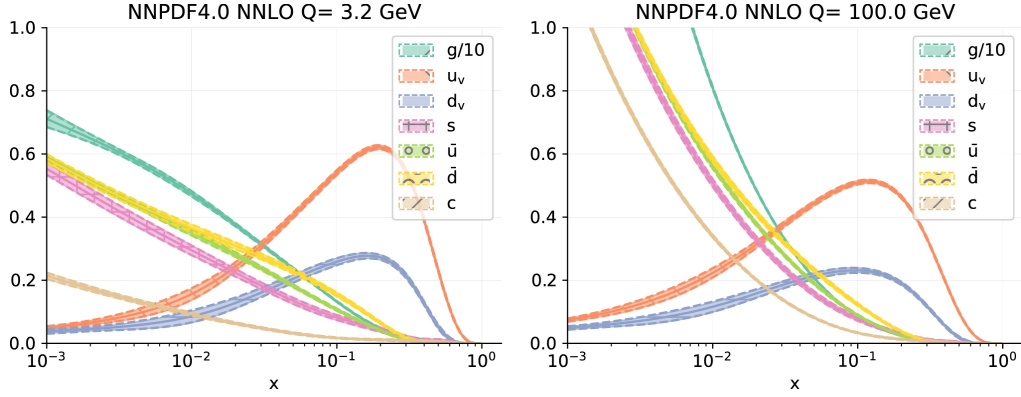


Figure 2.4: Parton distribution functions obtained by the NNPDF4.0 collaboration at NNLO with $Q = 3.2$ GeV (left) and $Q = 100$ GeV (right). From reference [94].

where q_0 and g_0 are the bare distributions for quarks and gluons, μ_F is the factorisation scale analogous to the renormalisation scale of α_s , P_{ij} are the splitting functions accounting for the probability of emitting another parton, κ is an energy scale arising from the cut-off in the loop momentum and C is a calculable function depending on the factorisation scheme.

From this expression, F_2 can be derived. Similarly to obtaining the running of the coupling constant, by imposing the invariance of F_2 with respect to μ_F , the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations are derived. The DGLAP equations can be condensed into a matrix equation in the space of quarks, antiquarks, and gluons as:

$$\mu_F^2 \frac{\partial}{\partial \mu_F^2} \begin{pmatrix} q_i(x, \mu_F^2) \\ g(x, \mu_F^2) \end{pmatrix} = \frac{\alpha_s(t)}{2\pi} \sum_{q_j, \bar{q}_j} \int_x^1 \frac{d\zeta}{\zeta} \begin{pmatrix} P_{q_i q_j} \left(\frac{x}{\zeta}, \alpha_s(\mu_F^2) \right) & P_{q_i g} \left(\frac{x}{\zeta}, \alpha_s(\mu_F^2) \right) \\ P_{g q_j} \left(\frac{x}{\zeta}, \alpha_s(\mu_F^2) \right) & P_{g g} \left(\frac{x}{\zeta}, \alpha_s(\mu_F^2) \right) \end{pmatrix} \begin{pmatrix} q_j(\zeta, \mu_F^2) \\ g(\zeta, \mu_F^2) \end{pmatrix}$$

Similarly to the coupling constant, the DGLAP equations provide information about the evolution of the PDFs with the energy scale, but their absolute values must be obtained from experimental measurements. Several collaborations work to obtain PDF parameterisations using results measured by experiments. Some examples are NNPDF [94] and MMHT [95]. Figure 2.4 shows the PDFs for gluons and up, down, charm and strange quarks obtained by NNPDF at NNLO for two different scales.

This method of characterising the structure of the proton, which cannot be achieved by pQCD alone, allows the separation of cross-section computations into two factors. One accounting for the hard scattering, where pQCD can be applied, and another part given by the PDFs. Both parts are combined through convolution. This approach, known as factorisation, leads to the following expression for the cross section in proton-proton collisions:

2 Theoretical framework

Table 2.3: Charge, third component of weak isospin, and weak hypercharge for left-handed and right-handed fermions of the SM.

left-handed fermions				right-handed fermions			
Fermions	Q	T_3	Y_W	Fermions	Q	T_3	Y_W
$u \ c \ t$	$+\frac{2}{3}$	$+\frac{1}{2}$	$+\frac{1}{3}$	$u \ c \ t$	$+\frac{2}{3}$	0	$+\frac{4}{3}$
$d \ s \ b$	$-\frac{1}{3}$	$-\frac{1}{2}$	$+\frac{1}{3}$	$d \ s \ b$	$-\frac{1}{3}$	0	$-\frac{2}{3}$
$\nu_e \ \nu_\mu \ \nu_\tau$	0	$+\frac{1}{2}$	-1				
$e \ \mu \ \tau$	-1	$-\frac{1}{2}$	-1	$e \ \mu \ \tau$	-1	0	-2

$$d\sigma(p_1, p_2) = \sum_{i,j} \int_0^1 \int_0^1 dx_1 dx_2 f_1(x_1, \mu_F^2) f_2(x_2, \mu_F^2) \hat{\sigma}_{ij}\left(x_1, p_1, x_2, p_2, \alpha_s(\mu_R^2), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2}\right).$$

2.3 Electroweak Theory

The EW theory unifies the description of both the electromagnetic and weak interactions. Together with QCD, it forms one of the foundational blocks of the SM of particle physics. This theory is based on a local gauge invariance represented by the $SU(2)_L \times U(1)_Y$ group.

The EW gauge group consists of the $SU(2)_L$ sector, where the conserved charge is the weak isospin, T_3 , and the $U(1)_Y$ sector, where the conserved quantity is the weak hypercharge, Y_W . However, neither of these quantities is truly conserved in a scenario of spontaneously broken symmetry, and only a specific non-trivial linear combination of them, $Q = T_3 + \frac{1}{2}Y_W$, remains conserved. According to weak isospin, fermions are arranged in $SU(2)_L$ doublets, known as left-handed fermions, or in $SU(2)_L$ singlets, referred to as right-handed fermions. For fermions, right-handed antiparticles form doublets with the opposite T_3 with respect to their left-handed counterparts, while left-handed antiparticles are singlets.

Tables 2.3 and 2.4 summarise the charge, third component of weak isospin, and hypercharge of fermions and bosons respectively.

Similarly to QCD, a Lagrangian density composed of two terms — the fermionic kinetic term and the Yang-Mills term corresponding to the gauge symmetry group $SU(2) \times U(1)$ — is written as:

$$\mathcal{L} = \bar{Q}_j i \gamma^\mu D_\mu Q_j + \bar{u}_j i \gamma^\mu D_\mu u_j + \bar{d}_j i \gamma^\mu D_\mu d_j + \bar{L}_j i \gamma^\mu D_\mu L_j + \bar{\ell}_j i \gamma^\mu D_\mu \ell_j - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}.$$

In the first five terms, the j -index runs over the three fermion generations; Q_j and L_j are the $SU(2)_L$ doublets for quarks and leptons, respectively. Similarly, u_j , d_j , and ℓ_j are the

Table 2.4: Charge, third component of weak isospin, and weak hypercharge for vector and scalar bosons of the SM.

Bosons	Q	T_3	Y_W
W^+	+1	+1	0
W^-	-1	-1	0
Z	0	0	0
γ	0	0	0
g	0	0	0
H	0	$-\frac{1}{2}$	+1

$SU(2)_L$ singlets corresponding to up-type quarks, down-type quarks, and charged leptons, respectively. γ^μ represents the Dirac matrices, and D_μ is the gauge covariant derivative defined as:

$$D_\mu = \partial_\mu - i\frac{g}{2}T_i W_\mu^i - i\frac{g'}{2}Y_W B_\mu, \quad (2.3)$$

where g and g' are the gauge constants representing the coupling to weak isospin and weak hypercharge, respectively. T_i are the Pauli matrices, with $i = 1, 2, 3$, Y_W is the weak hypercharge, and W_μ^i and B_μ are the four EW vector boson fields.

In the last two terms, $W_{\mu\nu}^i$ and $B_{\mu\nu}$ represent the bosonic field strengths defined as:

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk}W_\mu^j W_\nu^k,$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$$

where ϵ^{ijk} is the Levy-Civita symbol, representing the structure constants of $SU(2)_L$.

2.3.1 The Englert–Brout–Higgs mechanism

The EW Lagrangian does not include mass terms for either fermions or vector bosons because their addition would violate the gauge symmetry of the theory. However, experiments show that only one of the EW bosons is massless, the photon, while the other three are massive. To incorporate mass terms in the Lagrangian without violating gauge invariance, the Englert–Brout–Higgs mechanism of spontaneous symmetry breaking is introduced in the EW theory.

This mechanism introduces a scalar field:

2 Theoretical framework

$$h = \begin{pmatrix} h^+ \\ h^0 \end{pmatrix}.$$

This field is a weak isospin doublet with weak hypercharge $Y_W = 1$. Therefore, the upper component can be identified as having electric charge one, and the lower component is neutral. Using this scalar field, additional terms can be added to the Lagrangian:

$$\mathcal{L}_h = |D_\mu h|^2 - V(h^\dagger h) - y_u^{ij} \bar{Q}_i T_2 h^* u_j - y_d^{ij} \bar{Q}_i h d_j - y_\ell^{ij} \bar{L}_i h \ell_j + \text{h.c.}$$

In the first term, the covariant derivative is the same as in Equation 2.3, as $SU(2)_L \times U(1)_Y$ symmetry must be preserved.

The second term in this Lagrangian is the Higgs potential, given by:

$$V(h^\dagger h) = -\mu^2 h^\dagger h + \lambda (h^\dagger h)^2. \quad (2.4)$$

The last three explicit terms correspond to the Yukawa couplings with y_u^{ij} , y_d^{ij} , and y_ℓ^{ij} representing the Yukawa couplings matrices. The hermitian conjugate refers only to these three terms.

For positive values of μ^2 and λ , the potential in Equation 2.4 possesses a degenerate minimum corresponding to:

$$|h| = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}},$$

where v is the vacuum expectation value of the Higgs field; re-parameterising this field around the minimum:

$$h = \exp\left(-iT_i \frac{\theta^i}{v}\right) \begin{pmatrix} 0 \\ \frac{v+H}{\sqrt{2}} \end{pmatrix},$$

where θ^i and H are the four real degrees of freedom equivalent to the original complex doublet. The three θ^i fields correspond to Goldstone bosons resulting from the symmetry breaking, and the H field represent the physical Higgs boson. The re-parameterisation redefines the Higgs sector Lagrangian in terms of the new field as:

$$\mathcal{L}_h = \left| D_\mu \begin{pmatrix} 0 \\ \frac{v+H}{\sqrt{2}} \end{pmatrix} \right|^2 - V(H) - \left[y_u^{ij} \bar{u}_i^L u_j^R + y_d^{ij} \bar{d}_i^L d_j^R + y_\ell^{ij} \bar{\ell}_i^L \ell_j^R + \text{h.c.} \right] \frac{v+H}{\sqrt{2}}, \quad (2.5)$$

where u_i^L and d_i^R are the positive and negative components of the weak isospin Q_i doublet, and ℓ_i^L is the negative component of the weak isospin L_i doublet. The u_i^R , d_i^R and ℓ_i^R are the weak isospin singlets. The re-parameterised potential can be expressed as:

$$V(H) = \mu^2 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4. \quad (2.6)$$

This new Lagrangian gives rise to a series of new terms. From the potential in Equation 2.6, the first term, $-\mu^2 H^2$, represents the mass term of the Higgs boson with a mass $m_H = \sqrt{2}\mu$. Similarly, from the Yukawa terms in Equation 2.5, the combination of the Yukawa couplings and the vacuum expectation value generates terms bilinear in the fermion fields proportional to mass matrices defined as:

$$M_u^{ij} = \frac{v}{\sqrt{2}} y_u^{ij}, \quad M_d^{ij} = \frac{v}{\sqrt{2}} y_d^{ij}, \quad \text{and} \quad M_\ell^{ij} = \frac{v}{\sqrt{2}} y_\ell^{ij}.$$

Diagonalising these matrices gives the mass eigenstates, corresponding to physical particles. However, the rotation necessary for diagonalising also involves a rotation in flavour eigenstates, as explained in Section 2.3.2.

From the covariant derivative term, the vacuum expectation value produces a term of the form:

$$\mathcal{L}_M = \frac{v^2}{8} \left\{ g^2 \left[(W_\mu^1)^2 + (W_\mu^2)^2 \right] + (g W_\mu^3 - g' B_\mu)^2 \right\},$$

where the first term can be rearranged as:

$$\frac{g^2 v^2}{8} \left[(W_\mu^1)^2 + (W_\mu^2)^2 \right] = M_W^2 W_\mu^+ W^{-\mu}, \quad \text{with } W_\mu^\pm = \frac{W_\mu^1 \mp i W_\mu^2}{\sqrt{2}} \quad \text{and} \quad M_W^2 = \frac{g^2 v^2}{4},$$

where the $W^\pm$ fields are the physical massive charged vector bosons that carry the EW force.

The second term can be rearranged by defining the following combination of the W_μ^3 and B_μ fields as:

$$\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \quad \text{with} \quad \tan \theta_W = \frac{g'}{g}.$$

This transformation results in the mass term of the Z field, the massive neutral vector boson that mediates the EW force. The mass of the Z boson is given by:

2 Theoretical framework

$$M_Z^2 = \frac{v^2 (g^2 + g'^2)}{4}$$

and the photon remains massless.

2.3.2 The Cabibbo–Kobayashi–Maskawa matrix

Since the fermions in the Lagrangian are expressed in terms of mass eigenstates resulting from the diagonalisation of the mass matrices, the fermion doublets that appear in the coupling terms of quarks with charged currents are no longer mass eigenstates. This leads to flavour-changing charged current terms for quarks in the form:

$$\mathcal{L}_{CC}^q = \frac{g}{\sqrt{2}} (\bar{u} \quad \bar{c} \quad \bar{t})_L \gamma^\mu V \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L W_\mu^+ + \text{h.c.},$$

where V is a unitary matrix known as the Cabibbo–Kobayashi–Maskawa (CKM) matrix:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}.$$

The unitarity constraint, $V^\dagger V = 1$ imposes the following nine conditions:

$$\sum_{j=1}^3 |V_{ij}|^2 = 1, \quad \sum_{i=1}^3 |V_{ij}|^2 = 1, \quad \text{and} \quad \sum_{j=1}^3 V_{ki}^* V_{kj} = 0 \text{ for } (i \neq j).$$

The remaining nine free parameters can be divided into three real mixing angles and six phases. Five of the six phases can be absorbed into arbitrary phase rotations of the quark fields. The standard parameterization of the CKM matrix is:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix},$$

where $s_{ij} = \sin \theta_{ij}$ and $c_{ij} = \cos \theta_{ij}$, and the phase δ is the parameter responsible of the violation of the CP symmetry in the SM.

2.4 Jet algorithms

A consequence of QCD confinement is that quarks and gluons cannot exist as free particles. Therefore, when produced in particle colliders, these particles undergo processes of fragmentation or hadronisation, resulting in jets, which are collimated spray of stable particles. Jets are the signature of the presence of quarks and gluons.

The process of reconstructing these composite objects is called a jet algorithm. Jet algorithms must satisfy certain properties to behave correctly from a theoretical and experimental standpoints. These properties mainly include infrared safety, collinear safety, and independence on input objects. An infrared-safe jet algorithm should not be influenced by the presence of additional particles with infinitesimal energy radiated by the primary partons. A collinear-safe algorithm should not be affected by particles radiated at very small angles relative to the original parton. Independence on input objects means that the algorithm should produce consistent results regardless of whether the input provides information on partons, particles, or detector information from tracks and calorimeter cells.

Jet algorithms are generally classified into two main groups: cone algorithms and sequential clustering algorithms. Cone algorithms assume that particles forming jets appear in conical regions in the η - ϕ plane, imposing rigid circular boundaries on jets. While these algorithms are easy to implement, they often face theoretical challenges such as being infrared or collinear unsafe or failing to resolve multijet scenarios correctly. Examples include the iterative cone with progressive removal (IC-PR) [96], iterative cone with split-merge procedure (IC-SM) [97], and seedless infrared safe cone (SIScone) [98] algorithms.

Sequential clustering algorithms assume that particles within jets have similar transverse momenta, resulting in more flexible shapes in the η - ϕ plane. Although these algorithms offer better theoretical properties compared to cone algorithms, they tend to be more computationally expensive. The core principle behind every sequential clustering algorithm involves defining a pair of distances as follows:

$$d_{ij} = \min \left(k_{t,i}^{2p}, k_{t,j}^{2p} \right) \frac{\Delta_{ij}^2}{R^2}, \text{ and } d_{iB} = k_{t,i}^{2p}.$$

Here $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, R is the radius parameter of the algorithm and $k_{t,i}$, y_i and ϕ_i are the transverse momentum, rapidity, and azimuth of particle i . The parameter p is characteristic of each algorithm. Examples of sequential clustering algorithms include the k_t algorithm (with $p = 1$) [99], the Cambridge/Aachen algorithm (with $p = 0$) [100], and the anti- k_t algorithm (with $p = -1$) [101].

The analyses presented in this thesis used the anti- k_t algorithm for jet reconstruction. For jet calibration, in addition, the k_t algorithm is employed for pile-up corrections.

2 Theoretical framework

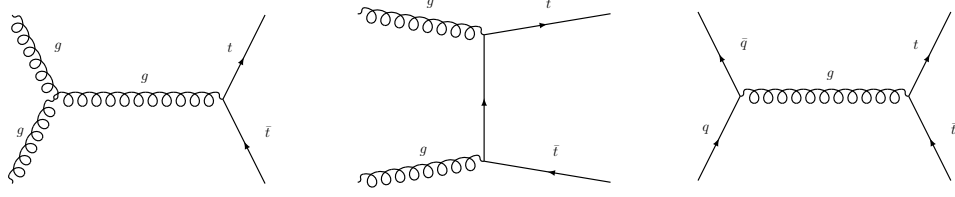


Figure 2.5: Feynman diagrams illustrating the dominant processes for $t\bar{t}$ production in hadron colliders: gluon-gluon fusion (left), t -channel (centre), and quark-antiquark annihilation (right) processes.

2.5 Top-quark physics

In the SM, the top quark is an elementary particle of the third generation known for its large mass, $m_t = 172.69 \pm 0.30$ GeV, and width, $\Gamma_t = 1.42^{+0.19}_{-0.15}$ GeV [91]. In terms of QCD, the top quark is a colour triplet. In the EW sector, the charge of the top quark is $Q = 2/3$, the left-handed top quark has a weak-isospin of $T_3 = +1/2$ forming a doublet with the left-handed bottom quark, resulting in a hypercharge of $Y_W = +1/3$. On the other hand, the right-handed top quark is a singlet under $SU(2)_L$, so $T_3 = 0$ and $Y_W = +4/3$.

The large width of the top quark is primarily due to its mass being well above the Wb threshold, and the fact that $|V_{tb}| \gg |V_{td}|$, leading to its dominant decay mode being $t \rightarrow Wb$. The SM prediction for the top-quark width at next-to-leading order (NLO) (neglecting terms of order m_b^2/m_t^2 , α_s^2 and $(\alpha_s/\pi) M_W^2/m_t^2$) is given by [102]:

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_s}{3} \left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right].$$

This decay width corresponds to a very short lifetime of about $0.5 \cdot 10^{-24}$ s, resulting in the top quark decaying before it can form bound states.

Top quarks are predominantly produced in pairs at hadron colliders. The dominant processes at leading order in QCD for top-quark production include $gg \rightarrow t\bar{t}$ and $q\bar{q} \rightarrow t\bar{t}$, represented by the Feynman diagrams in Figure 2.5. At $p\bar{p}$ colliders like the Tevatron, the dominant process is quark-antiquark annihilation, whereas at pp colliders like the LHC, gluon-gluon fusion is the primary production mechanism.

Predictions for the total cross sections of $t\bar{t}$ production are available at NNLO+NNLL accuracy in QCD [15], as well as differential QCD predictions at NNLO at the particle level [103]. Assuming a top-quark mass of 172.5 GeV, the theoretical NNLO+NNLL cross section in pp collisions at $\sqrt{s} = 13$ TeV is $\sigma_{t\bar{t}} = 832^{+20}_{-29}$ (scale) ± 35 (PDF, α_s) ± 23 (m_t) pb.

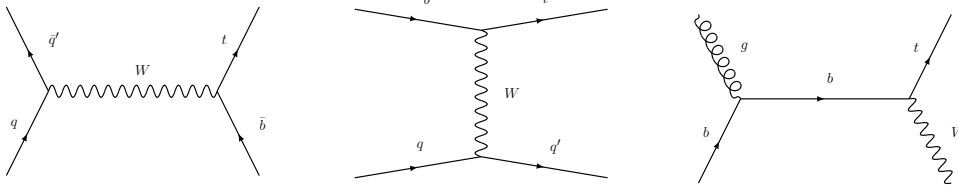


Figure 2.6: Feynman diagrams illustrating the dominant processes for single top-quark production in hadron colliders: s -channel (left), t -channel (centre), and tW -channel (right).

Due to the rapid decay of the top quark, only its decay products are observed in the detector. The final states from $t\bar{t}$ production are categorised into three classes: all-hadronic, lepton+jets (ℓ + jets), and dilepton ($\ell\ell$) channels. In the all-hadronic channel, $t\bar{t} \rightarrow W^+ (\rightarrow q\bar{q}') b W^- (\rightarrow q''\bar{q}''') \bar{b}$, both W bosons decay hadronically into quark-antiquark pairs. In the lepton+jets channel, $t\bar{t} \rightarrow W^+ (\rightarrow \ell^+\nu_\ell) b W^- (\rightarrow q\bar{q}') \bar{b}$ or $t\bar{t} \rightarrow W^+ (\rightarrow q\bar{q}') b W^- (\rightarrow \ell^-\bar{\nu}_\ell) \bar{b}$, one W boson decays hadronically and the other decays leptonically in a charged lepton and a neutrino. In the dilepton channel, $t\bar{t} \rightarrow W^+ (\rightarrow \ell^+\nu_\ell) b W^- (\rightarrow \ell'^-\bar{\nu}_{\ell'}) \bar{b}$, both W bosons decay leptonically.

Another mode of top-quark production in hadron colliders is the single top-quark production, an EW process mediated by a W boson. Figure 2.6 illustrate the leading-order (LO) Feynman diagrams for the three contributing channels: s -channel, t -channel, and tW -channel. In the s - and t -channel processes, a virtual W boson is exchanged to produce a top quark in association with another quark, whereas in the tW -channel, the top quark is produced in association with a W boson.

Predictions for the total cross section of single top quark and top antiquark production are available separately for each channel. For instance, at NNLO in QCD, the t -channel total cross section assuming a top-quark mass of 172.5 GeV, is $\sigma_t = 54.6^{+0.7}_{-0.4} (\text{scale})^{+0.9}_{-0.7} (\text{PDF}, \alpha_s)^{+0.6}_{-0.4} (m_t)$ pb for top quarks and $\sigma_{\bar{t}} = 29.7^{+0.4}_{-0.3} (\text{scale})^{+0.6}_{-0.5} (\text{PDF}, \alpha_s) \pm 0.3 (m_t)$ pb for top antiquarks in pp collisions at $\sqrt{s} = 13$ TeV [104]. Similarly, for the tW -channel, the total cross section at NLO+NNLL with the same conditions is $\sigma = 79.3^{+1.9}_{-1.8} (\text{scale}) \pm 2.2 (\text{PDF}, \alpha_s) \pm 1.2 (m_t)$ pb [105]. For the s -channel, NNLO+NNLL calculations yield a total cross section of $\sigma = 6.839^{+0.058}_{-0.034} (\text{scale})$ pb [106].

In addition to $t\bar{t}$ and single top-quark production, hadron colliders also allow access to processes involving top quarks in association with other particles, such as $t\bar{t}V$, $t\bar{t}VV$ (where $V = \gamma, W, Z$), $t\bar{t}H$ and $t\bar{t}t\bar{t}$. Detailed calculations including NLO in EW and QCD with soft gluon emission correction resummation to NNLL accuracy, are available for $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$ processes [107]. For $t\bar{t}t\bar{t}$, the accuracy extends to NLO in EW and QCD with soft gluon corrections at next-to-leading-logarithmic (NLL) accuracy [108]. Additionally, approximate NNLO predictions are available for $t\bar{t}W$ [109] and $t\bar{t}H$ [110].

2.6 Monte Carlo simulations

MC event generators are key tools in high-energy collider physics. They rely on the MC method to compute integrals that are challenging to solve using conventional numerical interpolation methods. In high-energy collisions involving elementary particles, the resulting final states are typically complex, comprising numerous hadrons, leptons, photons, and neutrinos. Understanding the underlying physics description of these final states is complicated due to limited comprehension of the physics involved and the impracticality of analytical approaches due to the high particle multiplicities involved. Additionally, simulating intricate geometric factors representing detectors is a common requirement for experimenters.

MC methods enable the generation of complete events by breaking down the problem into simpler steps. These steps include simulating key aspects of events such as the hard process, initial-state radiation (ISR), final-state radiation (FSR), beam remnants, hadronisation, and decays. Consequently, the generated events can be compared with experimental observations. MC event simulators play a crucial role in extracting physics insights from comparisons with experimental data and in studying physics for future experiments.

To a first approximation, all processes can be reduced to interactions among fundamental particles such as quarks, leptons, and gauge bosons. However, corrections to this simplified picture are needed to provide a more precise description of reality. These corrections can be categorised into three main types: bremsstrahlung-type modifications, higher-order corrections, and non-perturbative corrections.

The first type of corrections involves the emission of additional final-state particles through processes like $e \rightarrow e\gamma$ or $q \rightarrow qg$. Due to the value of the strong coupling constant and the presence of the triple gluon vertex, emissions in QCD of quarks and gluons are particularly abundant. This leads to phenomena known as parton showers, where a single initial coloured particle can give rise to multiple new particles in the final state. Bremsstrahlung corrections are universal, meaning they do not depend on the specific process but rather on key parameters such as the momentum transfer scale. These corrections can be incorporated to arbitrarily high orders using probabilistic methods, although exact calculations become increasingly complex beyond lower orders.

The higher-order corrections involve combinations of loop diagrams and the soft parts of bremsstrahlung graphs. This combination of loops and bremsstrahlung diagrams is necessary to cancel certain divergences. Calculating these corrections is typically challenging as higher orders are included, with few results available beyond one order higher than the non-trivial one.

Non-perturbative corrections are caused by the confinement predicted in QCD for quarks and gluons. This requires complementing perturbative descriptions with models of hadronisation, where coloured partons transform into jets of colourless hadrons, photons, and

leptons. Hadronisation can be subdivided into fragmentation and decays. In the former, quark-antiquark pairs are created to form hadrons, which significantly contributes to the overall high multiplicity of particles in these events. While a detailed understanding of this process is lacking, various models are used to simulate it.

With these corrections, the simple initial structure of processes becomes significantly more complex, and handled by MC event generators. This process is not completed in a single step but rather involves splitting the process into several manageable components. The output of an event generator should resemble real-world events, exhibiting both average behaviour and fluctuations akin to those observed in actual data.

To describe a typical high-energy event in a hadron collider such as the LHC, a MC event generator must incorporate simulations of various physics processes. Figure 2.7 provides a graphic decomposition of the collision simulation, which can be organised in a sequential manner as follows:

1. Initially, two protons approach each other from opposite directions. Each proton is characterised by a PDF defining the partonic substructure in terms of flavour composition and energy distributions. This is represented as the triple incoming green arrows and the big green blob in Figure 2.7.
2. From each proton, a parton initiates a branching sequence that builds up the ISR, involving processes such as $q \rightarrow qg$, $g \rightarrow qq$, and $g \rightarrow gg$. This is represented in dark blue in Figure 2.7.
3. Next, the hard process is computed with one parton from each of the two showers. This is represented by a large red blob in Figure 2.7.
4. The hard process may produce short-lived particles, such as top quarks or massive gauge bosons, whose decay products are considered in association with the products of the hard process. The decay of short-lived particles is displayed as small red blobs in Figure 2.7.
5. Concurrently with the hard process, other partons from the remnants of the incoming protons can interact, constituting the underlying event. The secondary interaction is displayed with a purple blob in Figure 2.7.
6. Similarly to the ISR, outgoing particles may branch building up the FSR. FSR is represented in red for the hard interaction and in purple for the secondary scattering in Figure 2.7.
7. After removing the initiating partons from the colliding protons, the remnants are left behind with a net colour charge that contributes to the final state. This is represented by light blue blobs in Figure 2.7.

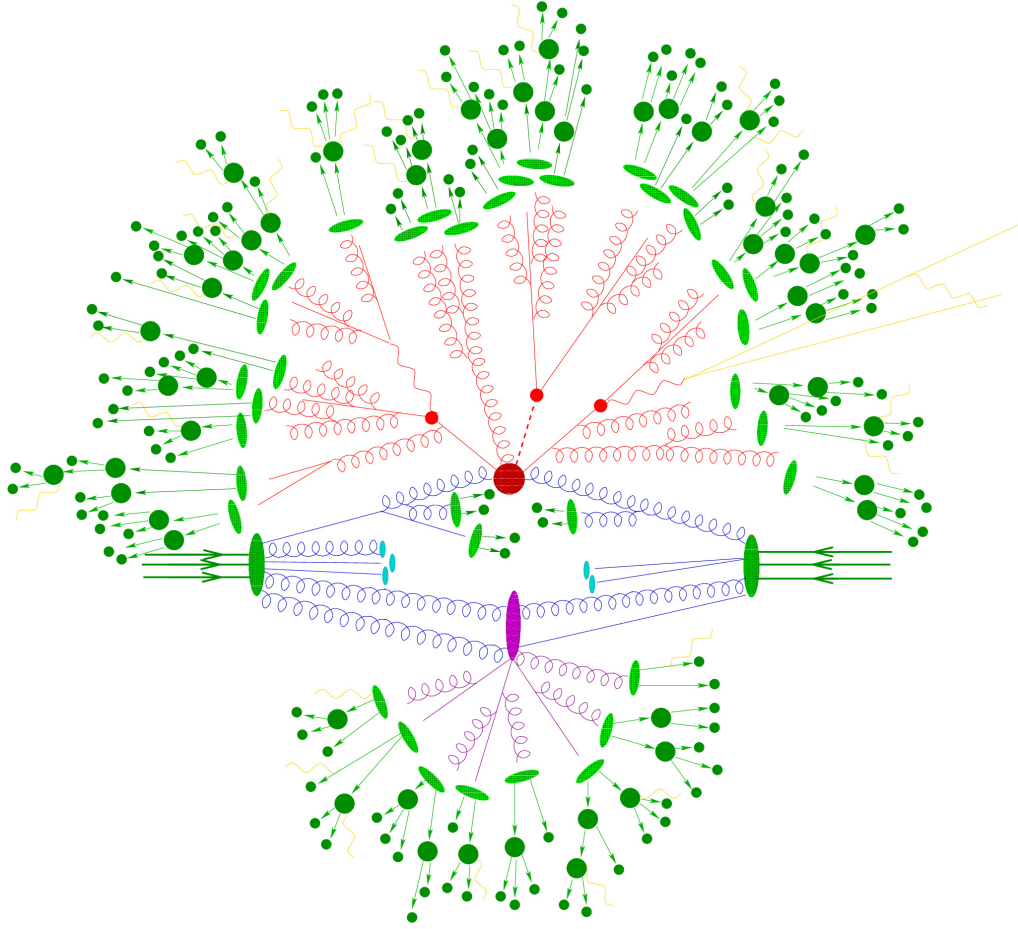


Figure 2.7: Representation of a generated event. From reference [111].

8. The outgoing coloured partons, due to confinement, fragment into colour-neutral hadrons. This is represented as light green blobs in Figure 2.7.
9. Many of the produced hadrons are unstable and decay into stable states. This is displayed as dark green blobs in Figure 2.7.
10. Finally, the produced particles interact with the detector and the detector response is simulated.

2.6.1 Hard processes

The event simulation of a hadron collision begins with the hard scattering, which is the subprocess where particles from each beam scatter. This interaction involves the highest momentum transfer of the event, and is typically classified according to the number of

final-state objects, such as $2 \rightarrow 1$, $2 \rightarrow 2$ or $2 \rightarrow 3$ processes. The main challenge arises when the number of particles in the final state increases, and currently, there is no generic treatment of processes with three or more final-state particles. MC event generators are thus restricted to modelling hard processes with a small number of final-state particles. Another classification can be made according to the type of interactions or particles produced.

While the low number of final-state particles might seem limiting for event generators, subsequent processes like parton showers enhance the final state by adding additional particles, emulating the types of events observed in hadron collider experiments. Additionally, the production of short-lived particles (resonances) such as t , W , Z , H or particles beyond the SM leads to decay chains that increase the number of final-state particles.

2.6.2 Parton density functions

Another essential ingredient in event simulation processes are PDFs, which describe the non-perturbative nature of hadrons as QCD bound states. PDFs cannot be derived from first principles, so parameterisations at lower scales are fitted to experimental data and then extrapolated to high energies by the DGLAP equations provided by the theory.

PDFs are primarily used to describe composite object such as protons or antiprotons, but similar principles can be applied to fundamental particles such as electrons. In the case of electrons, PDFs are useful for describing radiation, mainly photons, produced by the electron. Moreover, detailed electron PDF include other leptons, quarks, and gluons.

2.6.3 Initial and final-state radiation

Processes involving coloured or charged objects either in the initial or final states produce a significant radiation in the form of photons and gluons. This radiation leads to event signatures that are topologically much more complex than the simulated by $2 \rightarrow 2$ processes. Two approaches are commonly used to model this radiation: the matrix-element method and the parton-shower method.

The matrix-element method involves computing Feynman diagrams order by order. While theoretically more correct, calculations become increasingly challenging at higher orders, limiting precise corrections to only a few processes with better than one-loop precision.

The parton-shower method accounts for an arbitrary number of sequential branchings of one parton into two, using approximations that simplify kinematics, processes interference, and helicity structure. This provides a limited prediction of the rate of well-separated jets. In practice, this method is often used with a matching scheme to include hard-gluon emission in the hard process.

2 Theoretical framework

Due to the computational limitations of the matrix-element approach and the simplicity and flexibility of the parton shower method, the latter is usually preferred as the primary choice.

2.6.4 Beam remnants and multiple interactions

Another aspect of event generation is the modelling of beam remnants and multiple interactions, known as the underlying event. This process is particularly significant in hadron colliders because incoming particles are composite states that, after losing a parton, no longer have neutral colour charge. However, as mentioned earlier, fundamental particles can also be modelled by PDFs, which can lead to an underlying event, although less significant compared to hadrons.

These corrections account for secondary partons from each colliding particle participating in interactions, either in the hadronisation, as they are colour-connected to the hard interaction, or generating secondary interactions.

2.6.5 Hadronisation

The non-perturbative nature of QCD at long distances causes quarks and gluons, as coloured particles, to transform into colour-neutral hadrons. Due to the high value of the strong coupling constant in this regime, a perturbative approach is not feasible and other phenomenological models are employed. These models include independent fragmentation, string fragmentation, and cluster fragmentation. While the independent fragmentation is the simplest scheme for generating hadron distributions, the other two models, represented in Figure 2.8, are commonly used in event generators.

The independent fragmentation model was originally designed to reproduce the jets formed in e^+e^- collisions. This model approaches the fragmentation of a quark by combining it with the antiquark of a $q\bar{q}$ pair created from the vacuum. This results in a meson with energy fraction z and the remaining quark with an energy fraction $1 - z$. The remaining quark then undergoes the same fragmentation process until its energy falls below a cut-off. Gluons are treated by splitting into a $q\bar{q}$ pair using the $g \rightarrow q\bar{q}$ splitting functions of the DGLAP equations, or by assigning the full momentum of the gluon to one of the quarks in the pair.

The simplicity of this model comes with certain weaknesses, such as the assumption of the parton being on-shell, the independence from the parton energy, the neglect of the coloured remnants, and the inability to merge close jets into a single jet.

In the string model of fragmentation, the produced $q\bar{q}$ pairs move in opposite directions, transferring part of their energy into colour fields. represented as a string with uniform energy per unit length. This concept is consistent with quarkonium spectroscopy. When

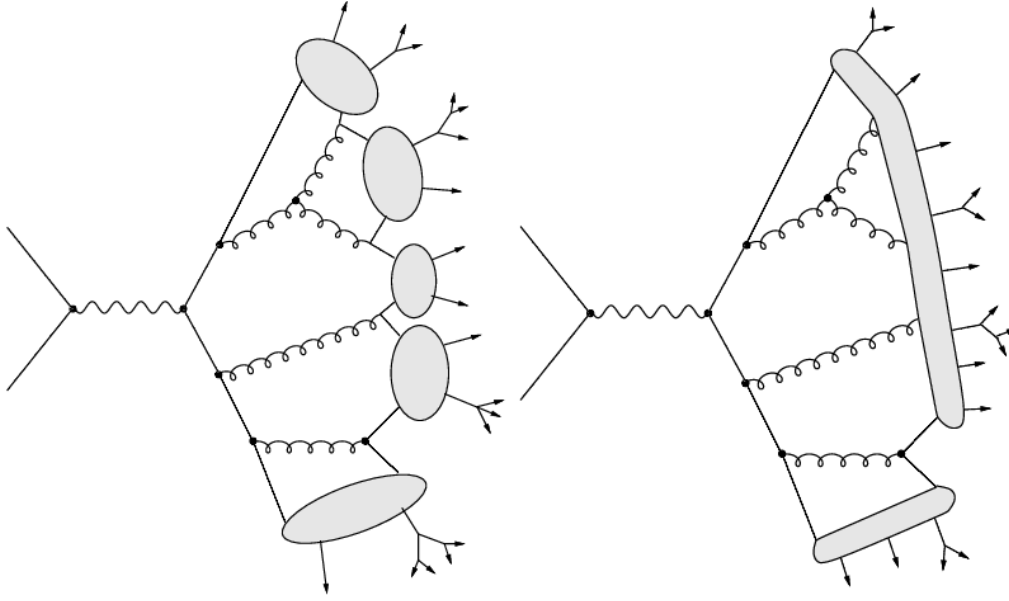


Figure 2.8: Representation of the cluster fragmentation (left) and string fragmentation (right) models. From reference [74].

enough energy is stored in the string, it spontaneously collapses, producing $q\bar{q}$ pairs in the form of mesons, similar to the independent fragmentation model. For gluon production, the string is deformed with kinks in the direction of gluons, representing localised energy and momentum carried by the gluon. For final states with multiple partons, the ambiguity of how $q\bar{q}$ pairs are assigned to the string ends and how gluons are included in one string or another is resolved by tracking the colour charge flow, stretching each string segment between colour-connected neighbours.

The string fragmentation model treatment of final-state gluons provides better agreement with experiments for three-jet final states in e^+e^- collisions compared to the independent scheme.

In the cluster fragmentation model, final-state gluons are non-perturbatively split into $q\bar{q}$ pairs. Subsequently, neighbouring quarks not originating from the same gluon are combined into colour singlets or clusters. These clusters can be treated as superpositions of meson resonances that decay into a number of hadrons, depending on the model.

Following fragmentation, many of the hadrons produced are unstable, and their decays are simulated using information on the branching ratios of the created hadrons.

2 Theoretical framework

2.6.6 Detector simulation

The final step of event simulation corresponds to simulating the detector effects. For this task, the GEANT4 simulation toolkit is employed [112]. Additionally, to emulate the collisions produced by the LHC, multiple proton-proton collisions must be simulated per bunch crossing to recreate the pile-up conditions [113].

GEANT4

The detector simulation employed by ATLAS collaboration is based on GEANT4, a versatile toolkit designed for simulating particle interactions within matter, featuring extensive capabilities in tracking, geometry, physics models, and hit generation. It covers a wide range of physics processes, including electromagnetic, hadronic, and optical interactions, accommodating various particles, materials, and energies spanning from 250 eV to the TeV scale. GEANT4 employs object-oriented programming techniques in C++ to facilitate adaptation across diverse applications in particle physics, nuclear physics, accelerator design, space engineering, and medical physics.

GEANT4 is developed to meet the software demands of modern experiments, which typically involve components like event generators, detector simulations, reconstruction, and analysis tools. GEANT4 serves as the foundation for simulation components, necessitating well-defined interfaces for interaction with other software components and providing reusable parts for integration.

The design requirements for GEANT4 include modularity, flexibility, and transparent physics implementation that allows user validation. It is designed to be comprehensible, customizable, and extendable across all domains, enabling users to selectively utilise required components.

Its key parts in the simulation of particle-matter interactions include geometry and materials, particle interactions, tracking management, digitization and hit management, event and track management, visualization and visualization frameworks, and user interface.

Pile-up simulation in the ATLAS experiment

The high luminosity of the LHC leads to a significant background known as pile-up, which poses challenges for physics analyses at ATLAS. In proton-proton collisions, in addition to the main collision of interest, secondary interactions between other protons of each bunch can be created. This affects every physics object reconstruction from adding energy to jets to generating extra muons. The number of collisions per bunch crossing has increased over time, due to the higher luminosities, reaching high values that significantly impact analyses. Understanding and accurately modelling this background is crucial for ATLAS analyses.

Pile-up can be divided in five components:

1. In-time pile-up: additional proton-proton collisions occurring in the same bunch-crossing as the collision of interest.
2. Out-of-time pile-up: additional proton-proton collisions occurring in bunch-crossings just before and after the collision of interest.
3. Cavern background: the gas of neutrons and photons inundating the cavern during a typical run of the LHC.
4. Beam halo events: the effect of protons from a bunch scraping against an up-stream collimator.
5. Beam gas events: collisions between the proton bunch and residual gas inside the beam-pipe.

ATLAS employs two methods to model pile-up and its effects. The first method involves a detailed simulation of the detector to replicate each component of the pile-up background. The second method uses specially recorded data events to emulate pile-up. Events obtained from both methods are mixed and added to the hard scattering.

3 Experimental set-up

In this chapter, a description of the experimental set-up used for the data acquisition is presented including an overview of the LHC accelerator [114–117], as well as a characterisation of the ATLAS detector [118].

3.1 The LHC accelerator

The LHC is a particle-particle accelerator located in the 26.7 km tunnel that previously housed the Large Electron-Positron collider [119, 120]. The tunnel consists of eight straight sections connected by eight arcs in a ring-like shape. Additionally, two transfer tunnels connect this to the rest of the CERN accelerator complex, as shown in Figure 3.1.

During the LHC Run 2, protons arrived at the LHC through these injection tunnels after being accelerated by a chain of pre-accelerators starting with the Linear accelerator 2 (Linac2), followed by the Proton Synchrotron Booster (BOOSTER), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS). This acceleration process began with a bottle of hydrogen gas as the source of protons. The hydrogen was ionised with an electric field to separate the electrons, and then introduced into Linac2, a linear accelerator that uses radiofrequency cavities to charge cylindrical conductors. The conductors behind the protons charged positively pushed the protons while the conductors negatively charged located ahead pulled them. Additionally, quadrupole magnets kept the beam collimated. At the end of Linac2, protons were accelerated to 50 MeV. Subsequently, protons were introduced to BOOSTER, a synchrotron with a circumference of 157 m, where the energy is increased to 1.4 GeV. The protons then were introduced into the PS, another synchrotron with a circumference of 628 m and 277 room-temperature electromagnets. After cycling inside the PS, the protons reached an energy of 26 GeV. From the PS, the protons were sent to the SPS, the last step before entering the LHC. The SPS is another synchrotron with a circumference of nearly 7 km and 1317 room-temperature electromagnets. The output energy of the protons from the SPS is 450 GeV. After the pre-acceleration process, the protons were injected into the LHC, where they reached an energy of 6.5 TeV to achieve the nominal centre-of-mass energy of 13 TeV.

3 Experimental set-up

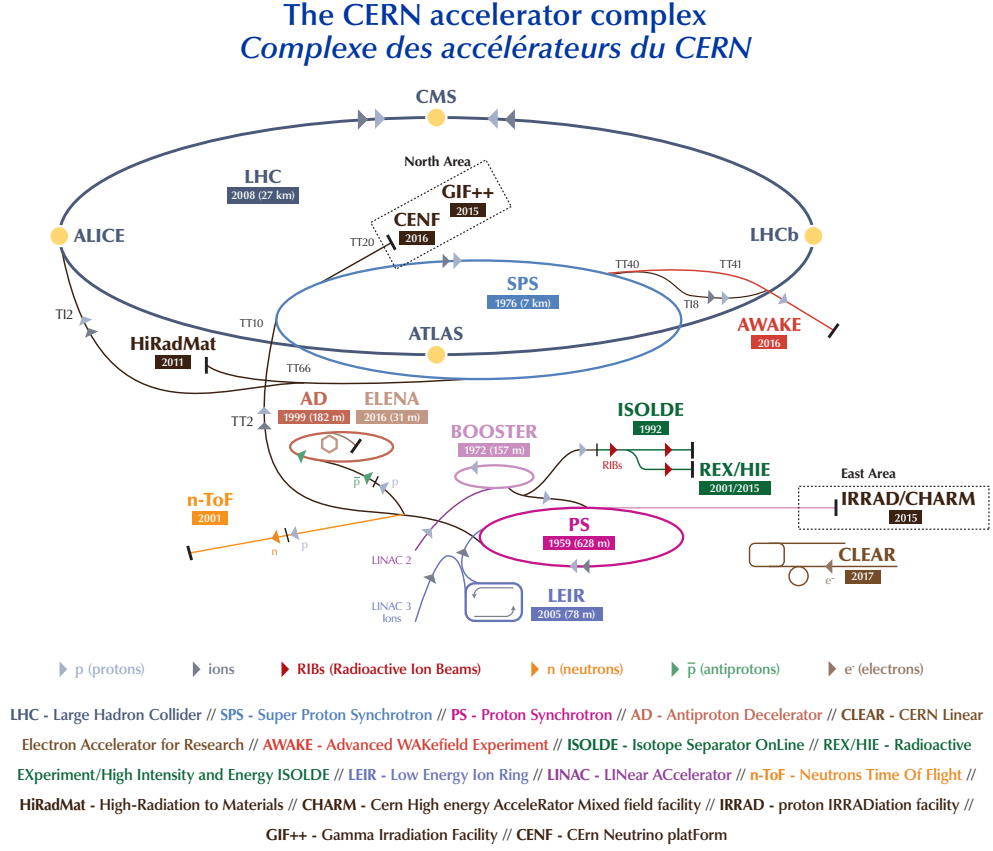


Figure 3.1: The CERN accelerator complex including the LHC, its four major experiments (dots), the pre-accelerators (Linac2, BOOSTER, PS, and SPS), and other facilities. From reference [121].

The LHC hosts four major experiments: two general-purpose experiments, ATLAS and CMS [122], one experiment focused on b -quark physics, LHCb [123], and one dedicated ion experiment, ALICE [124] (see Figure 3.1).

3.1.1 The LHC performance

The primary goal of the LHC is to provide collisions to its detectors for precision measurements of the SM and searches for physics beyond the SM. The design centre-of-mass energy is set to 14 TeV, enabling measurements to probe the TeV scale.

Another crucial parameter of an accelerator as the LHC is the luminosity, defined as the number of collisions per unit of area and time. It is related to the number of events produced per second by:

$$N_{\text{event}} = \mathcal{L}\sigma_{\text{event}},$$

where σ_{event} is the cross section for the event studied and $\mathcal{L}$ is the instantaneous luminosity. This luminosity depends on various beam parameters and can be expressed for a Gaussian beam distribution as:

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma}{4\pi \varepsilon_n \beta^*} F,$$

where N_b is the number of protons per bunch, n_b is the number of crossing bunches per beam, f_{rev} is the revolution frequency, γ is the relativistic gamma factor, ε_n is the normalised transverse beam emittance, β^* is the beta function at the collision point, and F is the geometric luminosity reduction factor accounting for the crossing angle at the interaction point (IP) defined as:

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*} \right)^2 \right)^{-1/2},$$

where θ_c is the full crossing angle at the IP, σ_z is the RMS bunch length, and σ^* is the transverse RMS beam size at the IP.

The goal for the peak instantaneous luminosity for ATLAS and CMS is $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, while for LHCb, the peak luminosity aims to be $\mathcal{L} = 10^{32} \text{ cm}^{-2}\text{s}^{-1}$, and $\mathcal{L} = 10^{27} \text{ cm}^{-2}\text{s}^{-1}$ for ALICE [124]. These high luminosities preclude the use of antiprotons, and hence a particle-anti-particle collider with a common vacuum and magnet system. Therefore, the LHC is designed as a proton-proton collider with separate magnet fields and vacuum chambers. Since there is not enough room in the tunnel for two separate rings, the LHC uses twin-bore magnets consisting of two sets of coils and beam channels within the same mechanical structure and cryostat.

3.1.2 The LHC layout

The fundamental structure of the LHC consists of 8 octants. Each octant has a straight section of approximately 528 m in length at the centre, connected by arcs to the neighbouring sections. These straight sections can function as either experimental or utility insertions. Notably, the two high luminosity experimental insertions, hosting the ATLAS experiment at Point 1 and the CMS experiment at Point 5, are strategically placed at diametrically opposite straight sections. Additional experimental insertions are found at Points 2 and 8, hosting the ALICE and LHCb experiments, respectively, in addition to the injection systems for Beam 1 and Beam 2. The opposite insertions at Points 3 and 7 each contain two collimation systems.

3 Experimental set-up

The radiofrequency (RF) systems are located at Point 4, with one independent system for each beam. These RF systems are responsible for increasing the energy of the protons with each revolution. Finally, Point 6 contains the beam dump system where the two beams are extracted from the accelerator when necessary.

The arc sections of the LHC have a lattice layout consisting of 23 regular arc cells, each measuring 106.9 m. These cells are constructed from two 53.45 m half cells, each containing a 5.355-meter-long cold mass (within a 6.63-meter-long cryostat) short straight section (SSS) assembly, and three 14.3 m dipole magnets. The design optimises the arc cell for a maximum integrated dipole field with minimal interconnections and compact beam envelopes. The apertures for each beam are separated by 194 mm.

Dispersion suppressors are strategically positioned at the transition points between LHC arcs and straight sections. These suppressors, totalling 16 sections, serve to adapt the LHC reference orbit to the tunnel geometry, cancel horizontal dispersion generated in the arc, and facilitate matching the insertion optics to the periodic optics of the arc. Their design involves standard arc cells with missing dipole magnets to achieve these objectives.

3.1.3 The LHC operations timeline

The construction of the LHC was completed in 2008, with expectations to operate at a centre-of-mass energy of 10 TeV by the end of the same year. However, on September 19th, 2008, a magnet quench occurred in sectors 3 and 4, resulting in a loss of around 6 tonnes of liquid helium and breaking the vacuum conditions of the beam pipe.

Following this incident, the LHC operated during Run 1 at 7 TeV in 2010 and 2011, and later at 8 TeV in 2012 [125]. In October 2010, the luminosity reached $2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. In September 2011, after increasing the number of bunches, the number of protons per bunch, and reducing the bunch size, the maximum LHC luminosity reached $3.6 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$. During 2012, a peak intensity of $7.7 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ was reached, 25% of the design target. Figure 3.2 shows the peak and integrated luminosity between 2010 and 2012. The mean number of interactions per crossing was lower than 4 in 2010 and it was increased in the following years up to 9.1 in 2011 and 20.7 in 2012.

A long shutdown took place during 2013 and 2014 for maintenance and upgrades including the superconducting magnet interconnections and to increase the centre-of-mass energy to 13 TeV.

The LHC resumed operations in 2015. This run, called Run 2, lasted until the end of 2018. This data-taking period, which corresponds to the dataset used in the analyses presented in this thesis, is summarised in Section 3.2.6. The mean number of interactions per crossing during this period was 33.7.

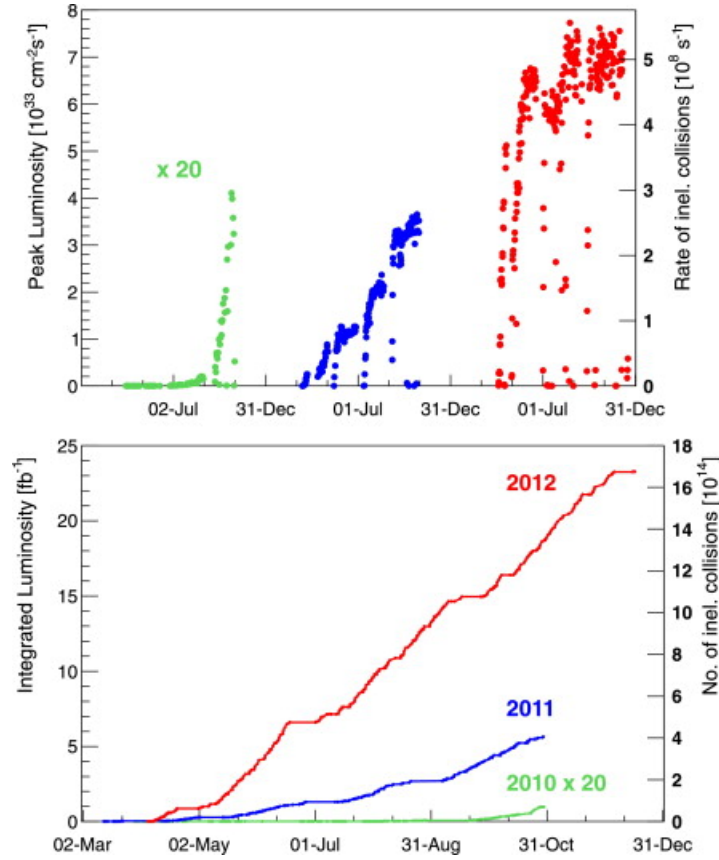


Figure 3.2: Peak (top) and integrated (bottom) luminosity between 2010 and 2012 for proton collisions. The 2010 luminosity values have been multiplied by a factor 20. From reference [125].

Following this period, a second long shutdown took place until 2022. During this period, the centre-of-mass energy was increased to 13.6 TeV and Linac4 replaced Linac2 as the source of proton beams for the LHC. Since July 4th, 2022, the Run 3 data-taking period has been ongoing.

3.2 The ATLAS experiment

The ATLAS (A Toroidal LHC ApparatuS) experiment is one of the two general-purpose experiments at the LHC, together with CMS (Compact Muon Solenoid). ATLAS is designed to exploit the high energy and luminosity conditions provided by the LHC, enabling more precise measurements of the SM, such as tests of QCD and EW theories, investigations into the properties of the top quark, and the search for the Higgs boson. Additionally, it conducts searches for physics beyond the SM, such as new heavy gauge bosons, supersymmetric particles or extra dimensions.

3 Experimental set-up

To achieve these goals, the experiment must meet several crucial requirements. It requires fast and radiation-resistant electronics. High detector granularity is imperative to manage the particle fluxes and minimise overlapping events. A large acceptance in pseudorapidity, with almost full azimuthal angle coverage, is essential. The charged-particle momentum resolution and reconstruction efficiency in the inner tracker must be high-quality, and the vertex detectors are needed to resolve secondary vertices from τ -leptons and b -jets. The calorimetry must be excellent, both in the electromagnetic part for electron and photon identification and energy measurements, and in the hadronic part for accurate jet and missing transverse energy measurements. The muon identification, momentum resolution, and charge determination must be precise. Highly efficient triggering on low transverse-momentum objects with sufficient background rejection is a prerequisite.

The coordinate system of the experiment is defined with the nominal interaction point as the origin. The z -axis is defined along the beam direction, with the positive x -axis pointing from the interaction point to the centre of the LHC ring and the positive y -axis pointing upwards. The azimuthal angle ϕ is measured around the beam axis, and the polar angle θ is the angle from the beam axis. Pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$ and for massive objects, rapidity is defined as $y = 1/2 \ln[(E + p_z)/(E - p_z)]$. Transverse momentum p_T , transverse energy E_T , and missing transverse energy E_T^{miss} are defined in the x - y plane. The distance in pseudorapidity-azimuthal angle space is defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

The layout of the ATLAS experiment forms a cylindrical structure, as shown in Figure 3.3, with forward-backward symmetry with respect to the interaction point. Two different magnet systems are used. The first is a superconducting solenoid surrounding the inner detector with a field of 2 T, and the other is the superconducting toroid system around the calorimeters. The latter is formed by three subsystems: one barrel with a field of around 0.5 T and two endcaps with approximately 1 T, each arranged with eight-fold azimuthal symmetry. The ATLAS experiment consists of several different detectors, such as the inner detector (ID) responsible for tracks and vertex identification, allowing the measurement of momenta of charged particles and electron identification. This detector is divided into three different systems: the pixel detector, the semiconductor tracker (SCT), and the transition radiation tracker (TRT). Outside the solenoid, the calorimeters are divided into the liquid-argon (LAr) electromagnetic calorimeter, the hadronic scintillator-tile calorimeter, and the LAr forward calorimeter, which measures both electromagnetic and hadronic energy deposits. Surrounding the calorimeter is the muon spectrometer (MS), composed of monitored drift tubes (MDTs), cathode strip chambers (CSCs), resistive plate chambers (RPCs), and thin gap chambers (TGCs). In addition to the detectors mentioned, the trigger system selects events of interest for physics studies from the raw interaction rate of 1 GHz to the data recording rate of 200 Hz during Run 1 and 1000 Hz during Run 2. A description of each of these detectors is presented in the following sections.

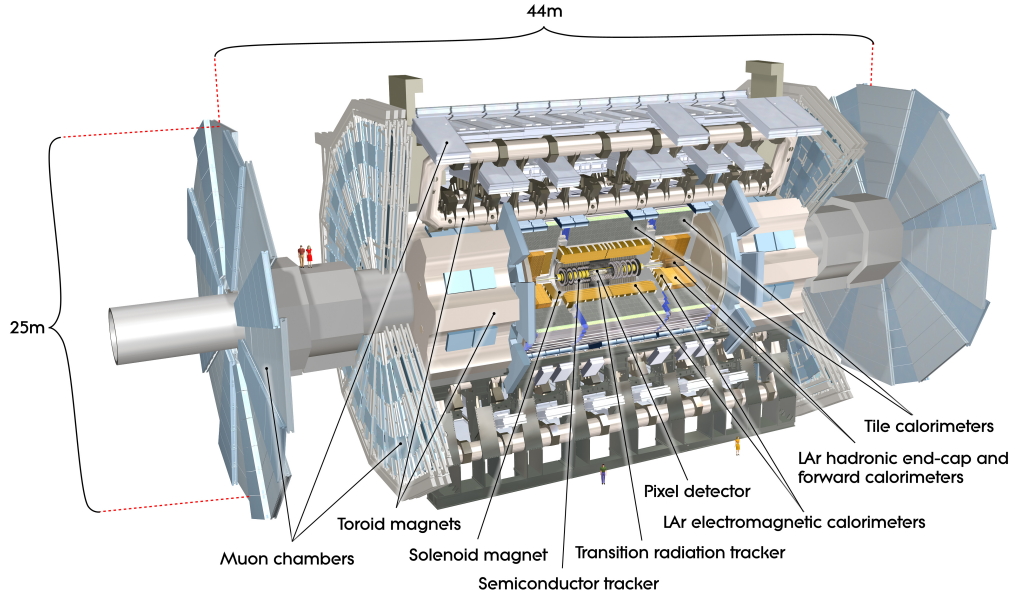


Figure 3.3: Representation of the ATLAS experiment layout with the main detectors and magnet systems labelled. From reference [118].

3.2.1 The inner detector

The ID is responsible for detecting tracks produced by charged particles in the ATLAS experiment. The high-energy collisions at the LHC result in approximately 1000 particles produced every 25 ns from the collision point, leading to a high track density. To meet the momentum and vertex resolution requirements, precise measurements with fine detector granularity are essential. These features are provided by the combination of the pixel detector, the silicon microstrips of the SCT and the straw tubes of the TRT.

The layout of the ID, depicted in Figure 3.4, covers the region of $|\eta| < 2.5$. The precision tracking detectors, namely the pixel detector and SCT, are arranged on concentric cylinders in the barrel region and on disks perpendicular to the beam axis in the endcap regions. The pixel detector, segmented in $R - \phi$ and z , consists of three cylindrical layers in the barrel and three disks in each endcap. Each track typically crosses three pixel layers, providing an intrinsic accuracy of $10 \mu\text{m}$ in $R - \phi$ and $115 \mu\text{m}$ in z both in the barrel and in the endcaps. Similarly, the SCT consists of four cylindrical layers in the barrel and nine disks in each endcap. Each layer comprises two 6.4 cm long sensors forming a 40 mrad angle with a strip pitch of $80 \mu\text{m}$. The SCT typically features eight strip layers (four space points) crossed by each track, achieving an intrinsic accuracy of $17 \mu\text{m}$ in $R - \phi$ and $580 \mu\text{m}$ in z . Both the pixel detector and SCT are made of silicon sensors.

3 Experimental set-up

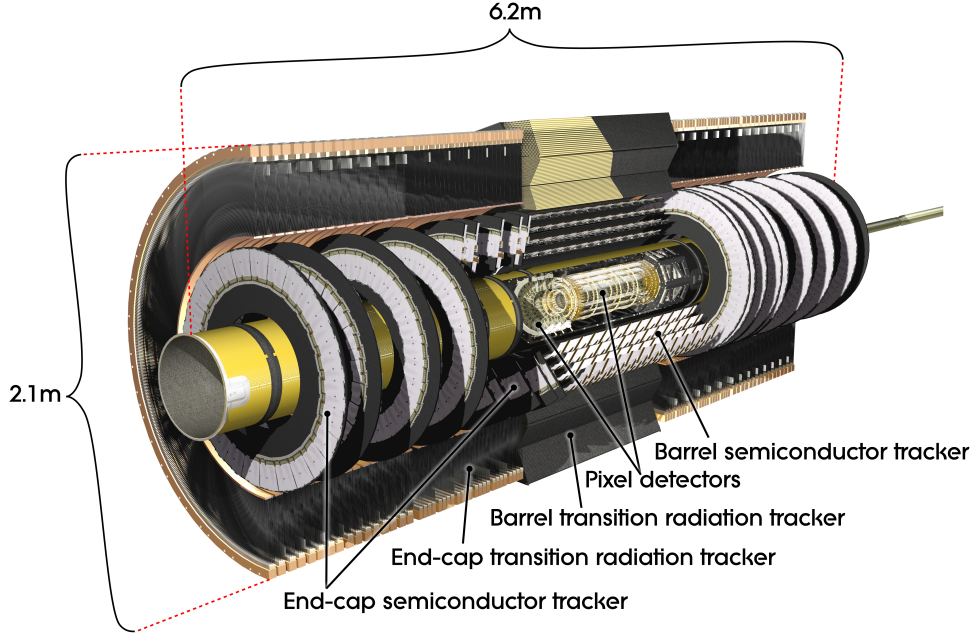


Figure 3.4: Representation of the inner detector of ATLAS with the main systems labelled. From reference [118].

The TRT, composed of 4 mm diameter straw tubes, is placed in the barrel parallel to the beam with a length of 144 cm and in the endcaps radially with a length of 37 cm. It provides around 36 hits per track, significantly contributing to momentum measurements, with an intrinsic accuracy of $130 \mu\text{m}$ per straw. The TRT straws are polyimide drift tubes filled with a mixture of xenon, carbon dioxide and oxygen.

The combination of precision trackers at small radii with the TRT at a larger radius ensures robust pattern recognition and high precision in both $R - \phi$ and z coordinates. The tracking measurements of the ID system complement the precision measurements of the electromagnetic calorimeter. Additionally, electron identification capabilities are enhanced by detecting transition-radiation photons in the xenon-based gas mixture of the straw tubes. The semiconductor trackers facilitate impact-parameter and vertex measurements for heavy-flavour and τ -lepton tagging, with the innermost layer of pixels, at a radius of about 5 cm, enhancing secondary-vertex measurement performance.

3.2.2 The calorimeters

The calorimeters are the components of the detector responsible for collecting the energy carried by electromagnetic and hadronic particles. Figure 3.5 provides an overview of the layout of the calorimeters. These systems employ diverse methodologies suited to the

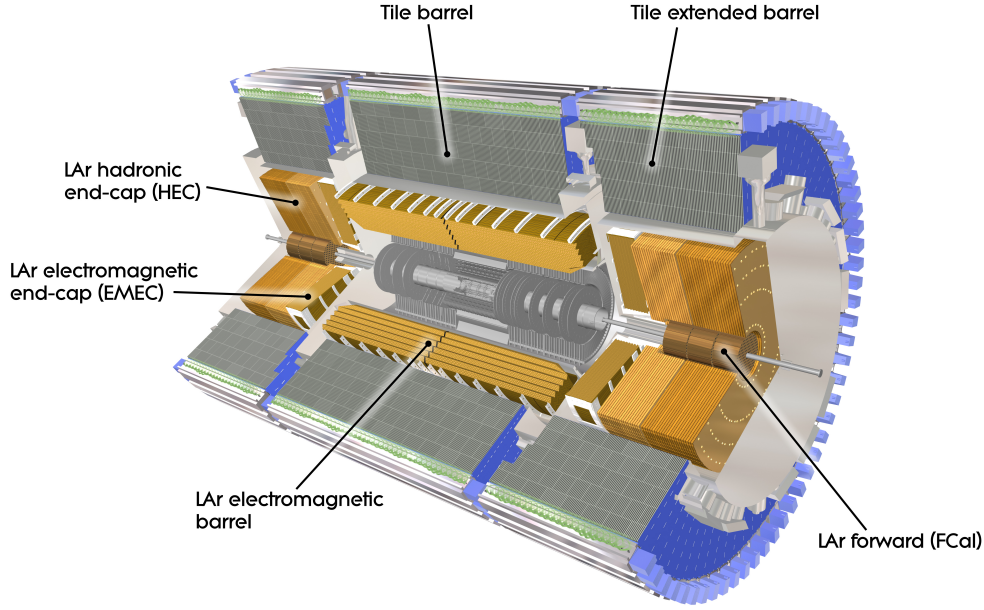


Figure 3.5: Representation of the ATLAS calorimeters layout with the main systems labelled. From reference [118].

varied requirements of relevant physics processes and cover a wide range of pseudorapidity $|\eta| < 4.9$. Within the η region overlapping with the ID, the fine granularity of the electromagnetic (EM) calorimeter facilitates precision measurements of electrons and photons. In the rest of the calorimeter, a coarser granularity suffices to fulfil the physics requisites for jet reconstruction and E_T^{miss} measurements.

Effective containment of electromagnetic and hadronic showers, alongside the limitation of punch-through into the muon system, is imperative for calorimeters. This consideration constrains the calorimeter depth in the design process. The EM calorimeter has a total thickness of over 22 radiation lengths (X_0) in the barrel and over 24 X_0 in the endcaps. With approximately 9.7 interaction lengths (λ) of active calorimeter in the barrel and 10 λ in the endcaps, it delivers robust resolution for high-energy jets. Moreover, the total thickness, incorporating 1.3 λ from the supportive structure, amounts to 11 λ at $\eta = 0$ and has been experimentally and computationally demonstrated to effectively mitigate punch-through well below the irreducible level of prompt or decay muons. This thickness, combined with extensive η -coverage, ensures reliable E_T^{miss} measurement, which is important for numerous physics signatures, particularly in SUSY searches.

The EM calorimeter consists of a barrel in the region of $|\eta| < 1.475$ and two endcap systems in the region of $1.375 < |\eta| < 3.2$, each housed within its dedicated cryostat. Each endcap wheel is divided in octants. Eight out of the 16 modules of the EM endcap calorimeter were

3 Experimental set-up

assembled at Universidad Autónoma de Madrid. To reduce the amount of material in front of the calorimeter, the central solenoid and the LAr calorimeter share a common vacuum vessel, eliminating two additional vacuum walls. The barrel calorimeter consists of two identical halves, separated by a small 4 mm gap at $z = 0$. Each endcap calorimeter is mechanically divided into two coaxial wheels: an outer wheel covering the region $1.375 < |\eta| < 2.5$, and an inner wheel covering the region $2.5 < |\eta| < 3.2$. Constructed as a lead-LAr detector with accordion-shaped kapton electrodes and lead absorber plates, the EM calorimeter ensures complete ϕ symmetry without azimuthal cracks. The lead thickness in the absorber plates is optimised as a function of η to enhance EM calorimeter performance in energy resolution. Within the precision physics region, $|\eta| < 2.5$, the EM calorimeter is segmented into three depth sections, while for the endcap inner wheel, the calorimeter is divided into two depth sections, featuring coarser lateral granularity compared to the rest of the acceptance. The $|\eta| < 1.8$ region is covered by a presampler detector employed to correct the energy lost by electrons and photons upstream of the calorimeter. The presampler comprises an active LAr layer, with a thickness of 1.1 cm in the barrel and 0.5 cm in the endcap region.

The tile calorimeter is positioned immediately outside the EM calorimeter. It consists of the barrel section flanked by two extended barrels. The barrel section covers the $|\eta| < 1.0$ region, while the extended barrels cover the range $0.8 < |\eta| < 1.7$. Utilising steel as the absorber and scintillating tiles as the active material, it operates as a sampling calorimeter. Azimuthally, both the barrel and extended barrels are divided into 64 modules. Radially, the tile calorimeter spans from an inner radius of 2.28 m to an outer radius of 4.25 m. Structurally, it is segmented in depth into three layers of approximately 1.5 , 4.1 , and 1.8λ in the barrel, and 1.5 , 2.6 , and 3.3λ for the extended barrel. At $\eta = 0$, the total detector thickness at the outer edge of the tile-instrumented region is 9.7λ .

The hadronic endcap calorimeter (HEC) comprises two independent wheels per endcap, positioned directly behind the EM endcap calorimeter and sharing the same cryostats. To mitigate the decrease in material density at the transition from the endcap to the forward calorimeter around $|\eta| = 3.1$, the HEC extends to $|\eta| = 3.2$, thereby overlapping with the forward calorimeter. Similarly, the η range of the HEC also slightly intersects with that of the tile calorimeter in the $1.5 < |\eta| < 1.7$ region. Each wheel is constructed from 32 identical wedge-shaped modules, assembled with fixtures at both the periphery and central bore. In addition, each wheel is divided into two segments in depth, resulting in four layers per endcap. The wheels closest to the interaction point consist of 25 mm parallel copper plates, while those farther away employ 50 mm copper plates (with the first plate being half-thickness in each wheel). The outer radius of the copper plates is 2.03 m, while the inner radius is 0.475 m (except in the overlap region with the forward calorimeter, where this radius is reduced to 0.372 m). These copper plates are located with a pitch of 8.5 mm filled with LAr that serves as the active medium for this sampling calorimeter.

The Forward Calorimeter (FCal) covers the $|\eta| > 3.1$ region and is integrated into the endcap cryostats, offering uniform coverage of the calorimeter and reducing radiation background

levels in the muon spectrometer. To mitigate neutron albedo within the inner detector cavity, the front face of the FCal is recessed approximately 1.2 m relative to the EM calorimeter front face. This design limits the depth of the calorimeter, making a high-density configuration necessary. With a depth of approximately 10λ , the FCal comprises three modules in each endcap: the first, crafted from copper, is suited for electromagnetic measurements, while the remaining two, composed of tungsten, predominantly measure energy from hadronic interactions. Each module features a metal matrix housing regularly spaced longitudinal channels filled with an electrode structure consisting of concentric rods and tubes aligned parallel to the beam axis. The sensitive medium within the gap between the rod and the tube is LAr. This geometric arrangement facilitates precise gap control, with gaps as small as 0.25 mm in the initial section to mitigate issues arising from ion buildup.

3.2.3 The muon spectrometer

The MS is the component responsible for measuring the momentum of muons. Figure 3.6 provides an overview of the muon detector layout. This spectrometer relies on the magnetic deflection of muon trajectories within large superconducting air-core toroid magnets, equipped with both trigger and high-precision tracking chambers. For $|\eta| < 1.4$, magnetic bending is provided by the barrel toroid. For $1.6 < |\eta| < 2.7$, two smaller endcap magnets within both ends of the barrel toroid bend the muon trajectories. In the transition region, $1.4 < |\eta| < 1.6$, magnetic deflection involves a combination of barrel and endcap fields. This magnetic setup yields predominantly a field orthogonal to the muon trajectories, while mitigating resolution degradation due to multiple scattering. Chamber placement involves three cylindrical layers encircling the beam axis within the barrel region, whereas in the transition and endcap regions, chambers are installed in planes perpendicular to the beam, also organised in three layers.

To address the two different tasks of the MS, namely precision tracking and triggering, the muon chambers are divided into monitored drift tubes (MDTs) and cathode strip chambers (CSCs) for the former; and resistive-plate chambers (RPCs) and thin-gap chambers (TGCs) for the latter. These chambers are arranged both in the barrel and in the endcaps. In the barrel, chambers are distributed in three concentric cylindrical shells around the z -axis at radii of approximately 5 m, 7.5 m, and 10 m. The innermost layer, located before the toroid magnet, is composed only of MDT chambers. The middle layer, placed in the gap left by the toroid magnets, consists of two RPC layers sandwiching a layer of MDT chambers. The outermost layer that forms the shell of the ATLAS barrel comprises MDT chambers and RPCs. In the two endcap regions, muon chambers are arranged in wheels perpendicular to the beam axis at distances of approximately $|z| \approx 7.4$ m, 10.8 m, 14 m, and 21.5 m. In the innermost wheel, located in front of the endcap toroid magnets, the $2 < |\eta| < 2.7$ region is covered by CSCs, while the $|\eta| < 2$ region is composed of MDT chambers preceded by TGCs. This is followed by the extra layer covering approximately the $1.5 < |\eta| < 2$ region,

3 Experimental set-up

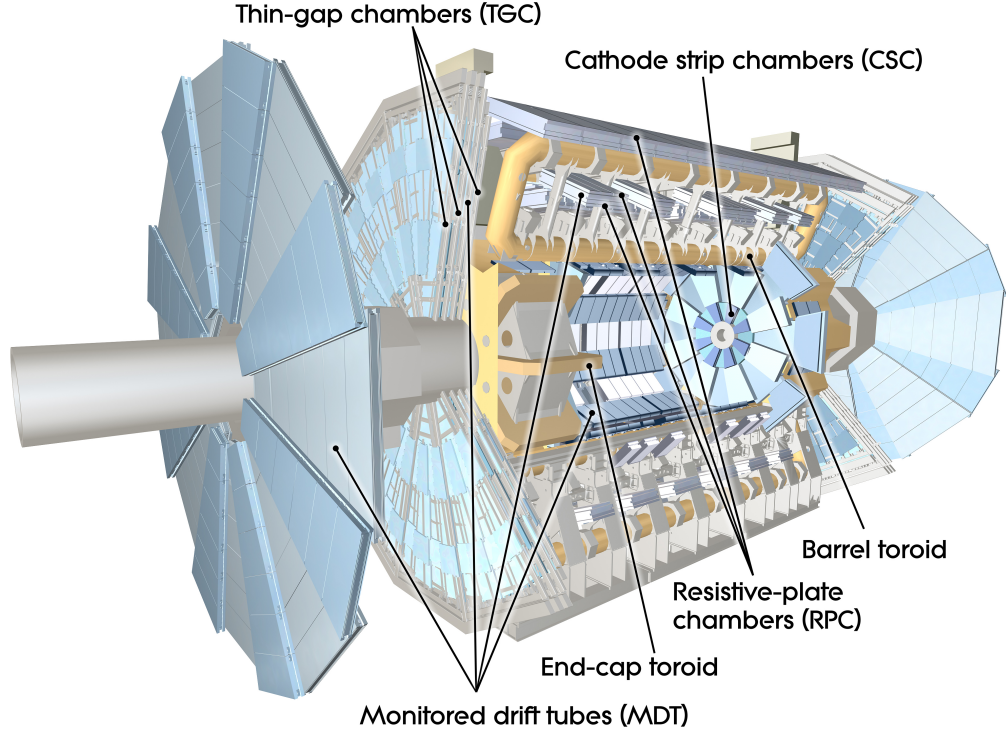


Figure 3.6: Representation of the muon spectrometer of ATLAS with the main systems labelled. From reference [118].

composed only of MDT chambers. The middle layer, located outside the endcap toroids, is composed of three layers of TGCs. The MDT chambers are located in between the first and second of these layers. The outermost layer, that cover the $2 < |\eta| < 3$ region, is composed only of MDT chambers.

The MDT chambers have an overall coverage range of $|\eta| < 2.7$. These chambers are formed by several layers of drift tubes, between three and eight, operating at an absolute pressure of 3 bar. The resolution achieved is $80 \mu\text{m}$ per tube, or about $35 \mu\text{m}$ per chamber.

The CSCs are located in the forward region, $2 < |\eta| < 2.7$, due to their higher rate capability and time resolution. These chambers are made of cathode planes segmented into strips in orthogonal directions allowing both coordinates to be measured. The resolution of these chambers is $40 \mu\text{m}$ in the bending plane and about 5 mm in the transverse plane.

Two different types of chambers are used for the muon trigger system. In the barrel region, $|\eta| < 1.05$, RPCs are used due to their good spatial and time resolution. In the endcap region, $1.05 < |\eta| < 2.4$, TGCs are selected due to their good time resolution and high rate capability. Additionally, these chambers provide the azimuthal coordinate that is not resolved by MDT chambers.

3.2.4 The ATLAS forward detectors

Three smaller detector systems are deployed to cover the ATLAS forward region. The primary purpose of two of them is to measure the luminosity delivered to ATLAS. Positioned at 17 m in each direction from the interaction point, LUCID [126] (LUminosity measurement using Cerenkov Integrating Detector) detects inelastic proton-proton scattering in the forward direction, serving as the principal online relative-luminosity monitor for ATLAS. The second system, ALFA [127] (Absolute Luminosity For ATLAS), situated at 240 m, is made of scintillating fibre trackers housed within Roman pots designed to approach up to 1 mm of the beam. The third system, the Zero-Degree Calorimeter [128] (ZDC) whose main role is to measure the centrality of heavy-ion collisions. Located at 140 m from the interaction point, just beyond the juncture where the shared straight-section vacuum-pipe bifurcates into two autonomous beam-pipes, ZDC modules consist of layers of alternating quartz rods and tungsten plates engineered to detect neutral particles for $|\eta| \geq 8.2$. Additionally, in 2017, a new forward detector, called ATLAS forward proton (AFP), was installed at approximately 200 m from the interaction point.

For the measurement of the luminosity, the ATLAS experiment relies on multiple redundant and complementary luminosity detectors and algorithms [129]. The primary system dedicated to luminosity measurements is the LUCID detector. Additionally, the beam conditions monitor system, the ID, and the calorimeter take complementary measurements of the luminosity. LUCID measures luminosity by detecting Cherenkov light produced in the quartz window of the 16 photomultiplier tubes. The beam conditions monitor, located in the ID, is a system used in low luminosity collisions as a consistency cross-check for LUCID measurements. The ID and the calorimeter uses the track activity and the energy deposits to estimate the luminosity. Once a year, a van der Meer absolute calibration is taken by varying the beam separation in the horizontal plane using LUCID and the beam conditions monitor system. Additionally, a relative calibration using a ^{207}Bi source is performed by LUCID in every inter-fill.

3.2.5 The trigger system

The LHC collision conditions imply an event rate of approximately 40 MHz. In order to reduce the rate of recorded events a selection procedure is needed. The ATLAS trigger system reduces the original rate from 40 MHz to a final recording rate of up to 1 kHz.

The trigger system operates across two distinct tiers, level-1 (L1) and high-level trigger (HLT), with each level refining decisions from the preceding one and applying additional selection criteria.

The L1 trigger is a hardware-based system. At this level, high-transverse-momentum particles like muons, electrons, photons, jets, τ -leptons decaying into hadrons, and significant

3 Experimental set-up

missing or total transverse energy are identified utilizing data from a subset of detectors. High-transverse-momentum muons are discerned through trigger chambers in both the barrel and endcap regions of the MS, while calorimeter selections rely on reduced granularity information from all calorimeters. The central trigger processor manages a trigger menu comprising various trigger selections. Selections with a high acceptance rate are defined with a pre-scaling factor to randomly reject a fraction of events passing the selection. The pre-scaling factors are chosen based on bandwidth utilization as luminosity and background conditions fluctuate. Events meeting the L1 trigger criteria progress to subsequent stages of detector-specific electronics and onward to data acquisition via point-to-point links. Additionally, the L1 trigger defines one or more regions of interest in each event, representing areas within the detector where interesting features are identified. This first selection makes a decision in less than $2.5 \mu\text{s}$, and reduce the acceptance rate to 75 kHz.

The HLT is a software-based system and employs full detector data within the regions of interest with complete granularity and precision. The HLT menu includes selections with one or a few particles such as electron, photon, muon, τ -leptons, jets, and missing transverse energy. The HLT selection reduces the event rate to about 1 kHz. This selection is executed using offline analysis procedures within an average event processing time of approximately 4 s.

3.2.6 The ATLAS experiment Run 2 timeline

From 2015 to 2018, the ATLAS experiment recorded collisions delivered by the LHC during Run 2. The overall conditions of this period included a centre-of-mass energy of 13 TeV and a bunch spacing of 25 ns. The luminosity reached a peak of around $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, resulting in up to 60 proton-proton collisions per bunch crossing.

In 2015, the LHC Run 2 began. On the morning of May 5th, the first collisions were recorded by ATLAS with an energy of 450 GeV per beam. The first collisions at 13 TeV were delivered on May 20th, and the experiment started recording under stable physics conditions on June 3rd. The number of bunches per beam was increased to 2244, each containing $1.15 \cdot 10^{11}$ protons. As shown in Figure 3.7, the peak luminosity reached $5 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$. A total integrated luminosity of 3.2 fb^{-1} was recorded by ATLAS this year.

In 2016, after repairing the endcap magnet bellows, collisions resumed in the LHC. The number of protons per bunch was increased to $1.2 \cdot 10^{11}$, and the design luminosity of $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ was obtained, with a maximum of $1.38 \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ achieved by the end of the year, as shown in Figure 3.8. A total integrated luminosity of 33.4 fb^{-1} was recorded by ATLAS this year.

In 2017, after the replacement of a magnet in the LHC, collisions returned to ATLAS. This year, the number of bunches per beam was initially increased to 2556 but later reduced to 1916. A new luminosity record of $2.09 \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ was achieved during this year as

3.2 The ATLAS experiment

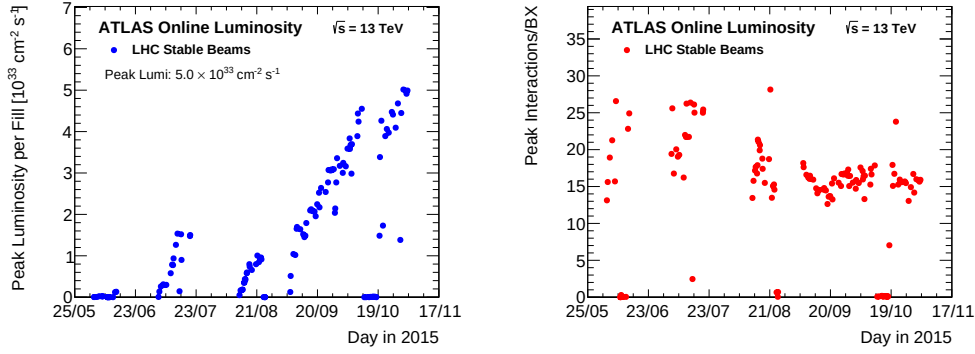


Figure 3.7: Measured peak luminosity delivered to ATLAS (left) and maximum number of inelastic collisions per beam crossing (right) during stable beams for pp collisions at a centre-of-mass energy of 13 TeV for each fill in 2015. From reference [130].

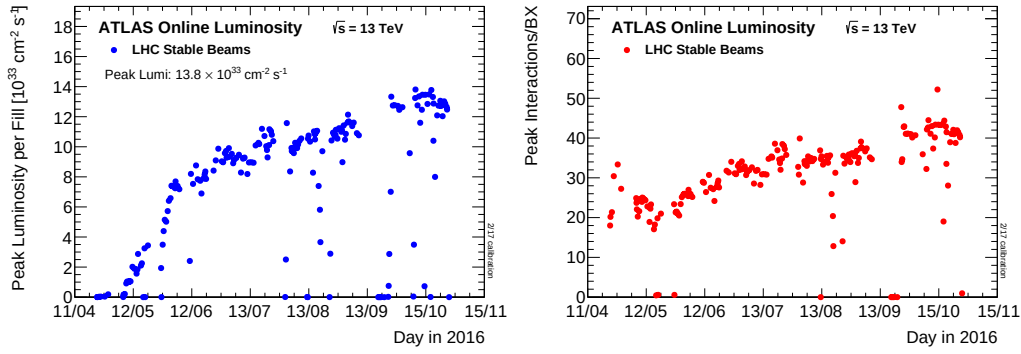


Figure 3.8: Measured peak luminosity delivered to ATLAS (left) and maximum number of inelastic collisions per beam crossing (right) during stable beams for pp collisions at a centre-of-mass energy of 13 TeV for each fill in 2016. From reference [130].

shown in Figure 3.9. A total integrated luminosity of 44.6 fb^{-1} was recorded by ATLAS this year.

In 2018, collisions resumed with the number of bunches per beam returning to 2556. This year the luminosity was maintained approximately constant at $1.9 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, as can be observed in Figure 3.10. A total integrated luminosity of 58.8 fb^{-1} was recorded by ATLAS this year.

The full Run 2 data-taking period achieved a total integrated luminosity of 140 fb^{-1} [129]. The cumulative luminosity as a function of time for the complete Run 2 is shown in Figure 3.11. The mean number of interactions per crossing was 13.4 in 2015 and was increased year after year up to 36.1 in 2018. The recorded luminosity as a function of the mean number of interactions per crossing during the LHC Run 2 is shown in Figure 3.12.

3 Experimental set-up

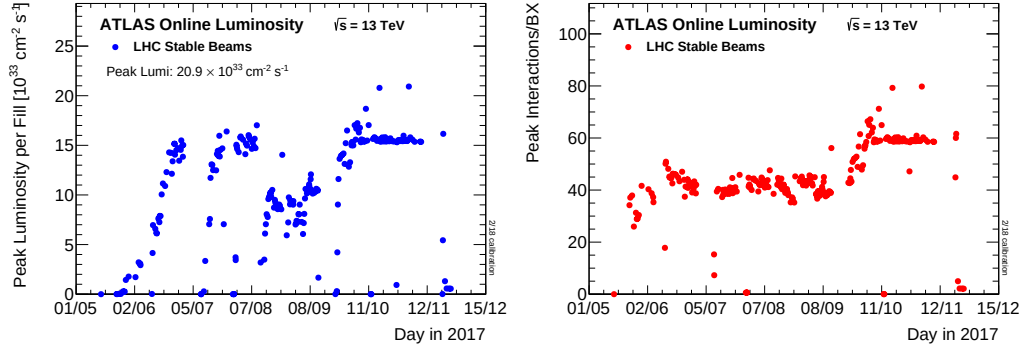


Figure 3.9: Measured peak luminosity delivered to ATLAS (left) and maximum number of inelastic collisions per beam crossing (right) during stable beams for pp collisions at a centre-of-mass energy of 13 TeV for each fill in 2017. From reference [130].

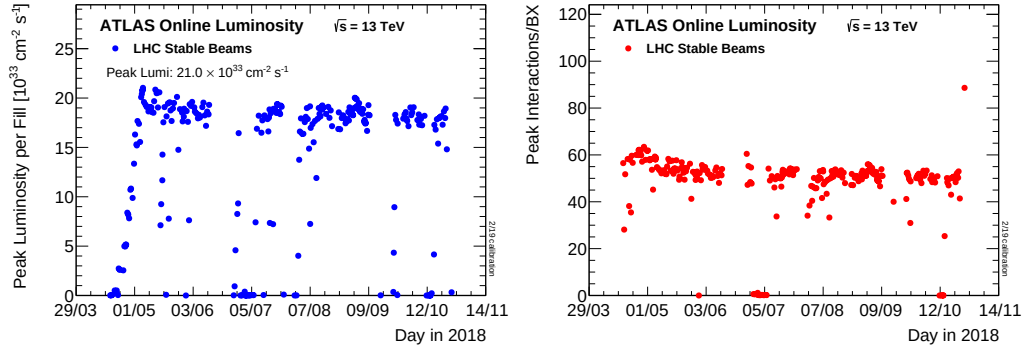


Figure 3.10: Measured peak luminosity delivered to ATLAS (left) and maximum number of inelastic collisions per beam crossing (right) during stable beams for pp collisions at a centre-of-mass energy of 13 TeV for each fill in 2018. From reference [130].

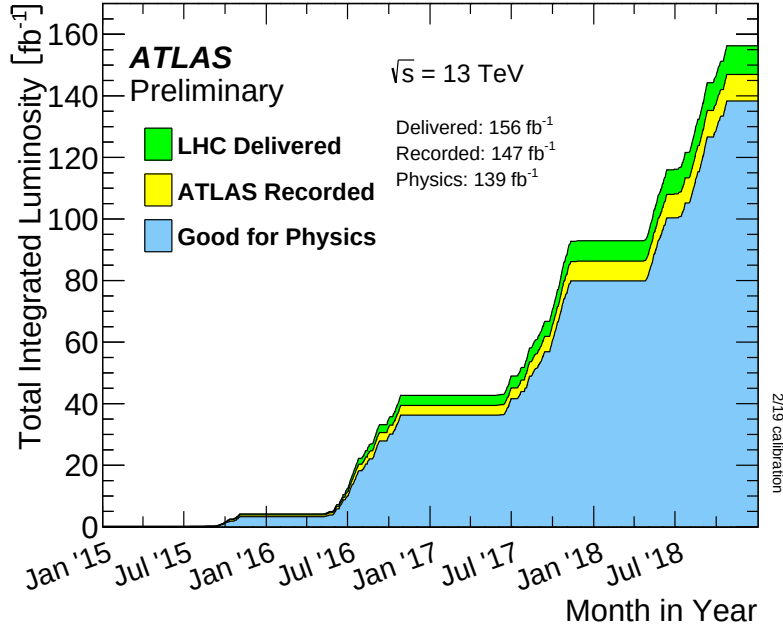


Figure 3.11: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at a centre-of-mass energy of 13 TeV in Run 2. From reference [130].

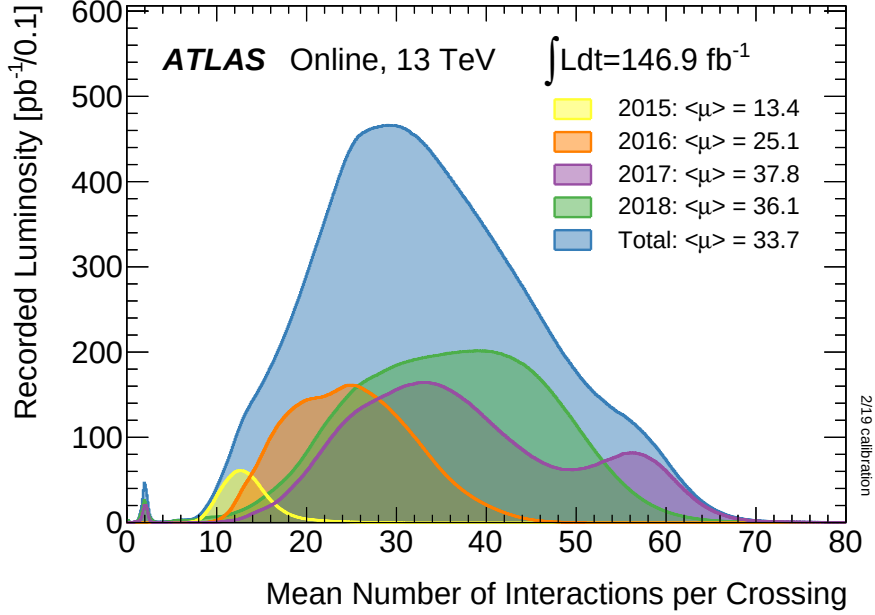


Figure 3.12: The recorded luminosity distribution as a function of the mean number of interactions per crossing for Run 2 pp collision data at a centre-of-mass energy of 13 TeV. The full Run 2 distribution is also shown separately for each year campaign. From reference [130].

4 Electron and photon reconstruction, calibration, identification and isolation

This chapter aims to describe the process of building electron objects used in the analysis, following the methodology employed by the ATLAS Collaboration during Run 2 [131, 132].

4.1 Electron and photon reconstruction

The electron reconstruction process is performed together with photons and proceeds as shown in Figure 4.1. The reconstruction of electrons and photons in ATLAS is achieved by implementing dynamic, variable-size clusters called superclusters [132]. Superclusters adapt their size to capture energy from bremsstrahlung photons and electron conversions, providing better energy recovery than fixed-size clusters. An electron is defined as an energy deposit in the calorimeter matched with a track, while a converted photon has a cluster matched to a conversion vertex, and an unconverted photon lacks such matches. The reconstruction process is carried out for electrons and photons with $|\eta| < 2.5$. The first part of the algorithm involves selecting topo-clusters [133, 134] and tracks, matching them together and building conversion vertices to be matched with topo-clusters. Then the supercluster-building process is done independently for photons and electrons. Finally, the electron and photon objects used for analysis are built.

The topo-cluster reconstruction algorithm is initiated by forming proto-clusters in the EM and hadronic calorimeters based on noise thresholds. Only the LAr EM calorimeter, excluding the presampler and first layer, is utilised for electron and photon reconstruction. In the transition region, between $1.37 < |\eta| < 1.63$, the energy measured by the presampler and the scintillator between the calorimeter cryostats is also taken into account. These cells are considered as seeds if they have a significance $|\zeta_{\text{cell}}^{\text{EM}}| \geq 4$, where

$$\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}},$$

$E_{\text{cell}}^{\text{EM}}$ is the cell energy accounting deposits from electromagnetic showers and $\sigma_{\text{noise,cell}}^{\text{EM}}$ is the expected cell noise coming from electronic noise and pile-up. Neighbouring cells meeting

4 Electron and photon reconstruction, calibration, identification and isolation

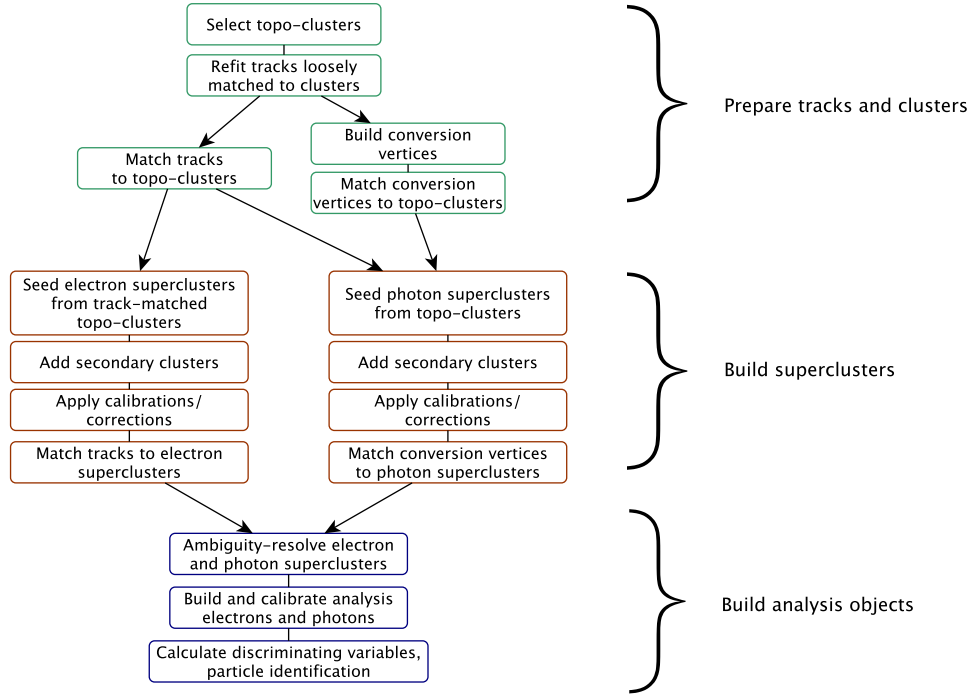


Figure 4.1: Algorithm flow diagram for the electron and photon reconstruction. From reference [132].

the significance criterion, $|\zeta_{\text{cell}}^{\text{EM}}| \geq 2$, are added to the proto-cluster iteratively. Proto-clusters with overlapping cells are merged and a crown of nearest-neighbour cells is added to the cluster. Proto-clusters with multiple local maxima are split into separate clusters. Local maxima are considered if the cell has $E_{\text{cell}}^{\text{EM}} > 500$ MeV, at least 4 neighbouring cells, and no neighbouring cell with higher energy. For these proto-clusters, the energy from the EM calorimeter must be larger than 400 MeV and a cut on the EM fraction (f_{EM}), the ratio between the EM and the total energy of the cluster, is applied to remove some topo-clusters coming from pile-up. The requirement of $f_{\text{EM}} > 0.5$ rejects around 60% of pile-up clusters without affecting true electrons.

The track reconstruction process starts with track hits in either the pixel detector or SCT. In the pixel detector, one hit equates to one space-point while in the SCT a space-point is built from a hit in a pair of stereo strips. Track seeds are formed from three of these space-points and are extended with pattern-recognition algorithms to form full tracks with at least 7 hits. Additionally, if the previous fitting fails, tracks with at least 4 hits loosely matched with an EM fixed-size cluster are fitted with an optimised method accounting for energy loss of charged particles in detector material. Once a full track is formed with the pixel detector and SCT hits, the TRT extension is applied including the hits in this detector that matched

the extrapolation of the track.

In the next step, topo-clusters and tracks are matched. This is done when they fulfil $|\Delta\eta| < 0.05$ and $-0.10 < q \cdot (\phi_{\text{track}} - \phi_{\text{clus}}) < 0.05$, where q refers to the charge of the track. If multiple tracks are matched to a single cluster, an ambiguity criteria is applied, as described in what follows. Tracks with hits in the pixel detector are preferred to tracks with hits only in the SCT. Within each category, tracks closer in ΔR are preferred. If the difference in ΔR is less than 0.01, tracks with more hits in the pixel detector are preferred. The last rule to resolve the ambiguity is choosing tracks with a hit in the innermost layer of the pixel detector.

Conversion vertices are reconstructed from tracks loosely matched to EM fixed-size clusters. Both tracks with hits in the pixel detector and SCT and tracks reconstructed only in the TRT are used for the conversion reconstruction. Both two-track and single-track conversion vertices are formed. Two-track vertices should have opposite-charge tracks and a vertex consistent with a massless particle. The TRT is used to determine if tracks are compatible with an electron.

The conversion vertices are then matched to the EM topo-clusters. For a conversion vertex with a track from the pixel detector and SCT, the extrapolation of the track to the calorimeter should be within $|\Delta\eta| < 0.05$ and $|\Delta\phi| < 0.05$. For TRT barrel tracks, the match requires $|\Delta\eta| < 0.35$ and $|\Delta\phi| < 0.02$, while in the endcap the requirement is changed to $|\Delta\eta| < 0.2$ and $|\Delta\phi| < 0.02$. In case multiple conversion vertices are matched to the same topo-cluster, two-track conversion vertices with pixel detector and SCT tracks are preferred to those with TRT tracks, followed by single-track vertices. In case of ambiguity, the vertex with the smallest conversion radius is chosen.

Once tracks and clusters are selected, the reconstruction of electron and photon superclusters follows. The process begins by ordering the EM topo-clusters according to descending E_T . These clusters are tested in order — separately for electrons and photons — and selected as electron supercluster seeds if $E_T > 1$ GeV and they are matched to a track with at least 4 hits in the pixel detector and SCT. The photon supercluster seeds are selected if $E_T > 1.5$ GeV.

If a certain topo-cluster is selected as a supercluster seed, satellite clusters are searched. For both photon and electron supercluster seeds, other neighbouring cluster are considered satellites if they fall inside a window centred on the seed. For photon supercluster seeds a $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ window is used; while for electron supercluster seeds a wider window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$ is used. For electrons, the satellite cluster is required to share the matched track with the seed. For converted photons, topo-clusters are considered satellites if they share a conversion vertex with the seed. Once a topo-cluster is selected as a satellite, it cannot be used as a seed cluster. The combination of the seed and the satellite topo-clusters creates a supercluster.

Once superclusters are built, calorimeter cells are assigned to them. For this step, the presampler and the first three layers of the LAr Calorimeter are used. In the transition region — $1.4 < |\eta| < 1.6$ — energy from the scintillator between the calorimeter cryostats is also

4 Electron and photon reconstruction, calibration, identification and isolation

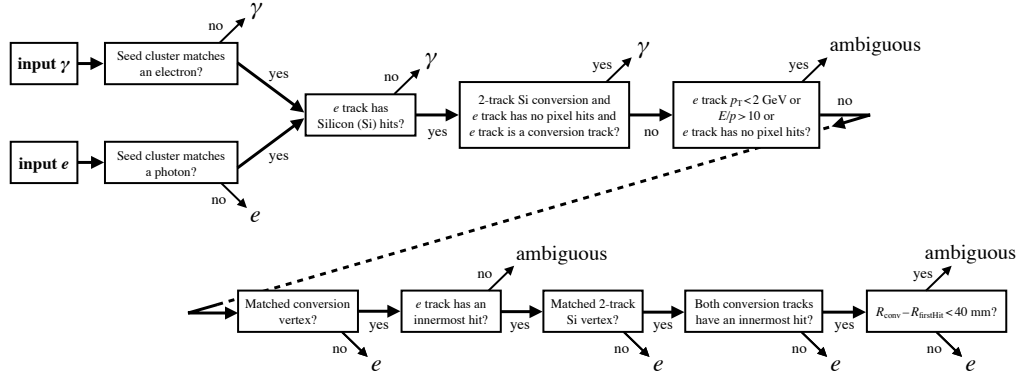


Figure 4.2: Algorithm flow diagram for solving the ambiguity between electron and photon objects. From reference [132].

added. To reduce the pile-up noise, the topo-cluster maximum width is restricted in the barrel (endcap) region to 0.075 (0.125) in the η region. After selecting the cells forming the supercluster, their energy is calibrated and their position corrected.

The final step of the supercluster algorithm is the track matching. It is done in a similar way to the matching to the topo-cluster.

Once the supercluster has been built, calibrated, and matched with tracks, the analysis objects for electrons and photons are created. The first step of this process is to resolve, when possible, the ambiguity produced by building independently electron and photon superclusters. This process is summarised in Figure 4.2.

Finally, a second calibration is performed with the information provided by the tracks matched to the superclusters and several discriminating variables involved in the identification process are computed. More details in the calibration process are described in Section 4.2 and a full description of the identification variables is given in Sections 4.3 and 4.4.

4.2 Electron and photon energy calibration

After reconstructing the electron and photon objects, it is necessary to perform an energy calibration for various effects [135]. These corrections account for energy losses in the material between the interaction point and the calorimeter, energy deposits in neighbouring cells of the supercluster, and energy losses beyond the LAr calorimeter.

4.2.1 Electron and photon energy estimate and expected resolution from the simulation

A correction considering all the effects mentioned above is computed using multivariate regression algorithms separately for electrons, converted photons and unconverted photons. The regression is carried out in intervals of $|\eta|$ and E_T . The variables used in the multivariate regression include the energy deposited in the calorimeter, the energy deposited in the presampler, the ratio of energies deposited in the first and second layers of the EM calorimeter, the pseudorapidity impact point of the shower in the calorimeter, and the distances $\Delta\eta$ and $\Delta\phi$ between the impact point of the shower and the centre of the closest cell in the second calorimeter layer. The impact point of the shower is computed as the energy-weighted average of the positions of the cells in the cluster. Additionally, for converted photons, the conversion radius is used, and for cases where both associated tracks have at least one hit in either the pixel detector or SCT, two additional variables are used. These variables are the ratio of the vectorial sum of the conversion track momenta and the energy measured in the calorimeter and the fraction of the conversion momentum carried by the highest- p_T track. For those objects within the transition region between the barrel and endcap; the energy deposited in the scintillators installed in the transition region and the difference between the cluster position and the centre of the scintillator cell are also used.

The energy resolution as a function of $|\eta|$ after applying the regression algorithm in the simulation is shown in Figure 4.3 for electrons and photons, both converted and unconverted. The resolution is measured as the interquartile range in the $E_{\text{calib}}/E_{\text{gen}}$ distribution divided by 1.35 to convert to the equivalent standard deviation of a Gaussian distribution. Here, E_{calib} and E_{gen} refer to the energy after and before applying the regression algorithm.

4.2.2 Corrections to account for differences between data and simulations

To account for residual differences between data and simulation, four corrections are applied. The first correction involves the intercalibration of the first and second calorimeter layers. It uses energy deposits left by muons from $Z \rightarrow \mu\mu$ events in the calorimeter. These energy deposits, unlike those created by electrons and photons, are not affected by the passive material in front of the EM calorimeter. The relative calibration of the two layers is computed as:

$$\alpha_{1/2} = \frac{\langle E_1 \rangle^{\text{data}} / \langle E_1 \rangle^{\text{MC}}}{\langle E_2 \rangle^{\text{data}} / \langle E_2 \rangle^{\text{MC}}},$$

where $\langle E_i \rangle$ denotes the most probable value of the deposited muon energy in the i -th layer of the calorimeter. The most probable value is computed following two different methods. The first one, referred to as the fit method, is based on fitting both data and simulation

4 Electron and photon reconstruction, calibration, identification and isolation

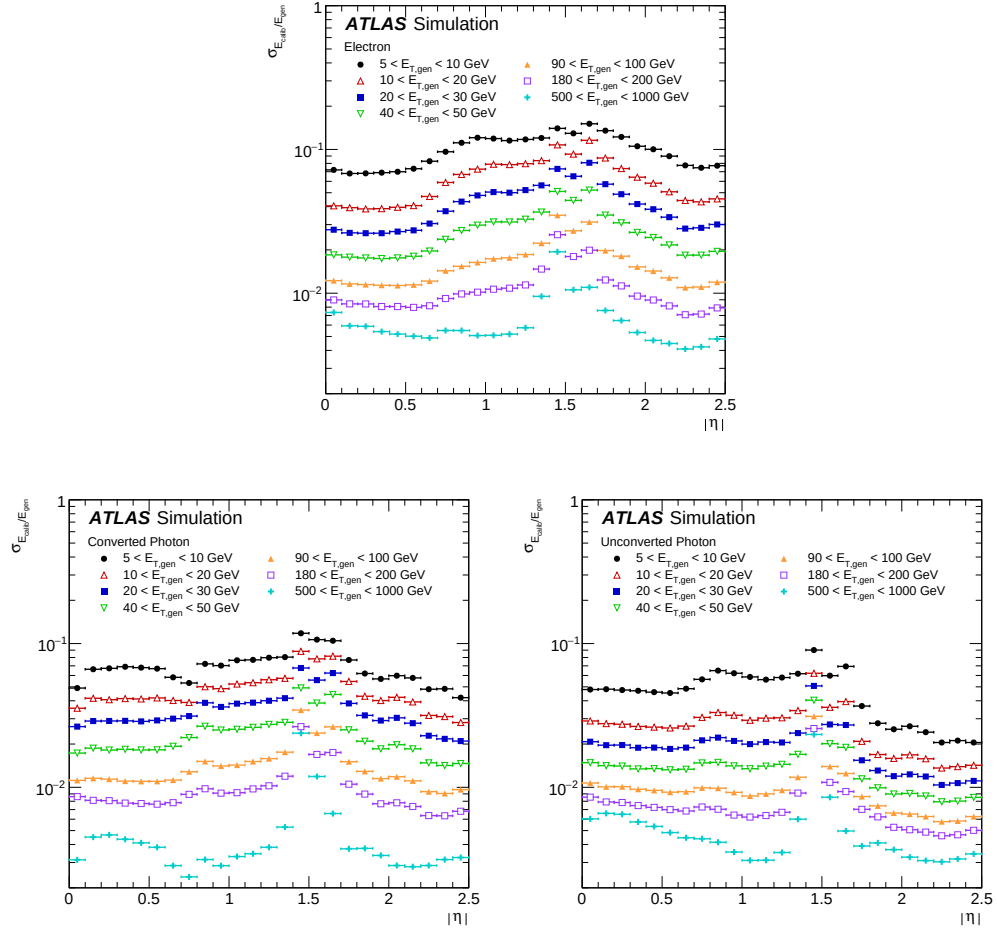


Figure 4.3: Measurement of the energy resolution $\sigma_{E_{\text{calib}}}/E_{\text{gen}}$ for electrons (top) and photons (bottom). From reference [135].

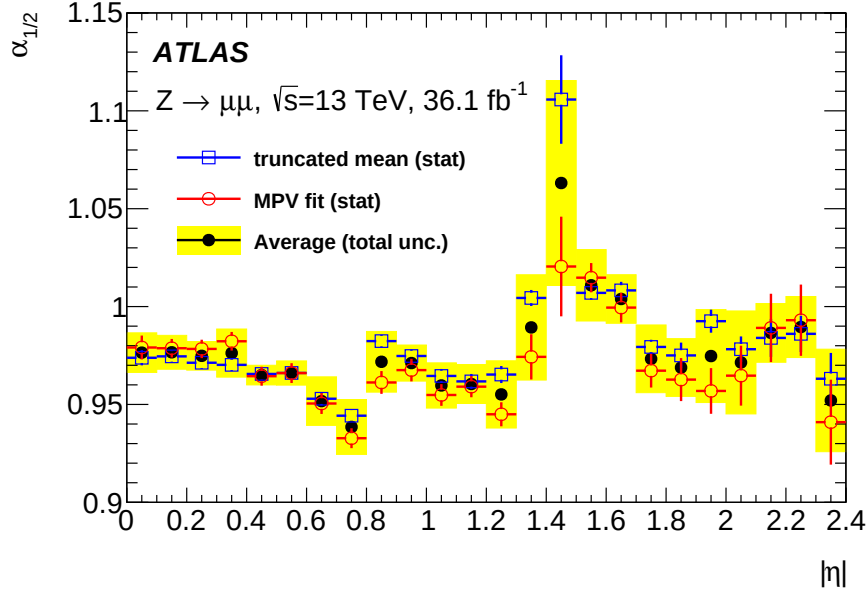


Figure 4.4: Measurement of the relative calibration between the first and second calorimeter layer $\alpha_{1/2}$ as a function of $|\eta|$ for the two different methods and the average of both. The uncertainty bars correspond to the statistical component in both methods and for the average the systematic uncertainty is included. From reference [135].

energy distributions using the convolution of a Landau and a noise distributions. A second method, called the truncated mean, computes the mean value of the energy distributions in a window centred around the mean of the whole distribution. The width of the window is taken between 3σ and 4σ , with σ being the RMS of the complete distribution. This is computed as a function of $|\eta|$ for both methods, as shown in Figure 4.4.

The second correction focuses on the presampler layer of the calorimeter. The presampler energy scale, α_{PS} , uses electrons produced in $Z \rightarrow ee$ events. In this case, these events are affected by the passive material in front of the presampler. To minimise the dependence of α_{PS} on material mismodelling, an additional correction factor is taken into account. For this reason, the expression for the presampler energy scale is:

$$\alpha_{\text{PS}} = \frac{E_0^{\text{data}}}{E_0^{\text{MC}}} \times \frac{1}{1 + A \left(\frac{E_{1/2}^{\text{data}}}{E_{1/2}^{\text{MC}} b_{1/2}} - 1 \right)},$$

where

- E_0^{data} and E_0^{MC} are the average energies deposited in the presampler in data and simulation, respectively.

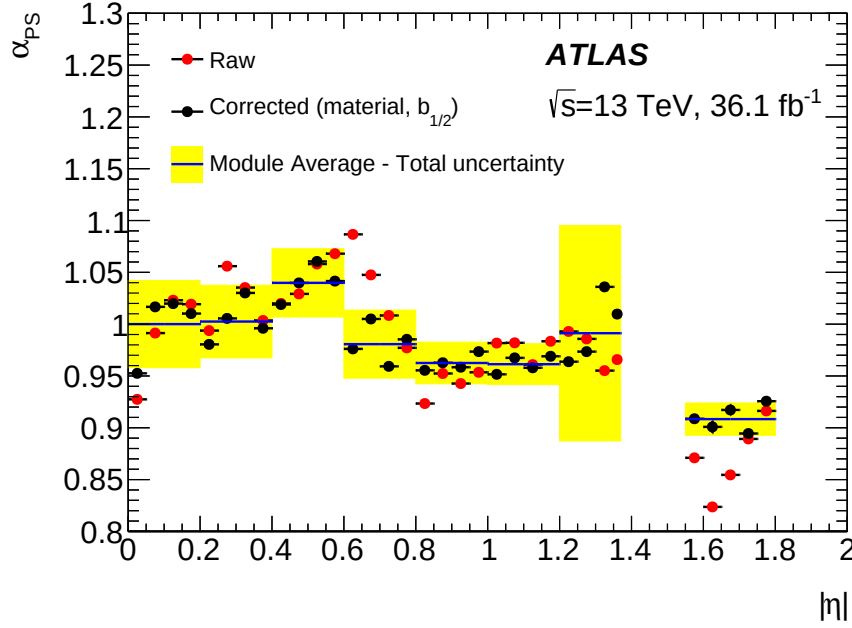


Figure 4.5: Measurement of the presampler energy scale, α_{PS} , as a function of $|\eta|$ before and after the correction to account for the material in front of the presampler. The average of the corrected values is shown as solid lines with uncertainty bars corresponding to the statistical and systematic uncertainties. From reference [135].

- $E_{1/2}^{\text{data}}$ and $E_{1/2}^{\text{MC}}$ are the averages of the ratio between the energy deposited in the first and second layers of the calorimeter in data and simulation, respectively.
- $b_{1/2}$ is the ratio of $E_{1/2}$ in data and simulation estimated from unconverted photons. It measures the dependence on the material in front of the presampler and after being applied to the ratio of $E_{1/2}$, it is only proportional to the mismodelling of the material in front of the presampler.
- A models the dependence of $E_{1/2}$ and E_0 on the material in front of the presampler.

The presampler energy scale is measured in intervals of $|\eta|$, excluding the region between the barrel and endcap ($1.37 < |\eta| < 1.52$). The measurements are done in intervals of 0.05 and then averaged in bins of 0.2 (0.3) pseudorapidity units in size for the barrel (endcap) region. Figure 4.5 shows the presampler energy scale before and after the correction to account for the influence of the material in front of the presampler.

Another calibration is done to reduce the dependence on the pile-up due to the relative position of the crossing bunches in the bunch train and the luminosity of the bunches that varies with time. To reduce this effect, the average energy induced by these effects is computed using simulations as a function of the bunch luminosity and bunch position.

An additional correction is done to reduce the pile-up energy shifts separately for each calorimeter layer. This is done as a function of the pseudorapidity and the average number of interactions per bunch crossing.

To reduce the effects coming from energy loss between the calorimeter modules, a correction is parameterised using the ratio of the calorimeter energy to the track momentum as a function of ϕ . For the inhomogeneities caused by sectors of the calorimeter where the high voltage is set to a non-nominal value to mitigate the effect of short circuits, a similar approach is used.

4.2.3 Energy scale and resolution corrections

A measurement of the differences between data and MC simulations in the energy scale, α_i , and energy resolution, c_i , is performed. These quantities are defined as:

$$E^{\text{data}} = E^{\text{MC}} (1 + \alpha_i) \text{ and } \left(\frac{\sigma_E}{E} \right)^{\text{data}} = \left(\frac{\sigma_E}{E} \right)^{\text{MC}} \oplus c_i,$$

where i corresponds to the region in pseudorapidity, and the symbol $\oplus$ denotes a sum in quadrature. To perform this measurement, invariant mass distributions of dielectron systems in $Z \rightarrow ee$ events are used. In these distributions, the difference in the average dielectron invariant mass and the mass resolution is parameterised as:

$$m_{ij}^{\text{data}} = m_{ij}^{\text{MC}} (1 + \alpha_{ij}) \text{ and } \left(\frac{\sigma_m}{m} \right)_{ij}^{\text{data}} = \left(\frac{\sigma_m}{m} \right)_{ij}^{\text{MC}} \oplus c_{ij},$$

where i and j are the regions in pseudorapidity of the electrons; and the parameters α_{ij} and c_{ij} are given by $\alpha_{ij} = (\alpha_i + \alpha_j) / 2$, and $c_{ij} = (c_i + c_j) / 2$. These variables are extracted by minimising the shape difference between data and MC simulations to later obtain the α_i and c_i parameters by solving the overconstrained set of equations. The difference in energy scale is applied to the data while the energy resolution is used to correct the simulation.

The measurement of the energy scale factor α_i and the energy resolution constant term c_i in regions of pseudorapidity is shown in Figure 4.6.

4.3 Electron identification

A set of quality criteria is applied to electron candidates to enhance the purity of the selected electron objects. This is achieved through a selection process that reduces the number of electron candidates originating from hadronic jets, converted photons or electrons produced in heavy-flavour hadron decays. This process relies on a set of variables summarised in

4 Electron and photon reconstruction, calibration, identification and isolation

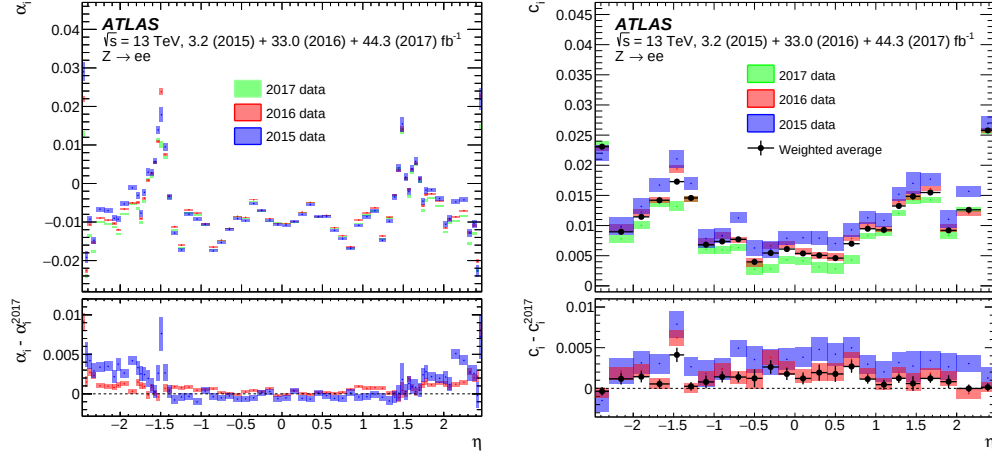


Figure 4.6: Measurement of the energy scale factor α_i (left) and energy resolution constant term c_i (right). The difference of each campaign with respect to 2017 data is shown at the bottom panel of each plot. The shaded areas correspond to the statistical uncertainties. From reference [132].

Table 4.1. These variables provide information about the electron track, the lateral and longitudinal shape of the electromagnetic shower, and the spatial compatibility of the electron track with the reconstructed cluster. Some variables serve as cut-based selection criteria for the electron candidates while others are used to construct a likelihood discriminant.

The variables used as requirements for electron candidates include the number of hits in the innermost pixel layer, the number of hits in the pixel detector, the total number of hits in the pixel and SCT detector, the shower width w_{stot} , and the ratio of the cluster energy to the track momentum.

The remaining variables are used to build an electron likelihood for both signal, L_S , and background, L_B , as follows:

$$L_S(\mathbf{x}) = \prod_{i=1}^n P_{S,i}(x_i),$$

where $\mathbf{x}$ is the vector of the electron candidate variables used to build the likelihood, and $P_{S,i}$ is the value of the probability density function for the i -th variable with a value x_i . The background likelihood is constructed analogously.

Both likelihoods are used to compute a discriminant defined as:

$$d_L = \frac{L_S}{L_S + L_B}.$$

Table 4.1: Discriminating variables used in the electron identification.

Variable	Description
R_{had_1}	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the ranges $ \eta < 0.8$ and $ \eta > 1.37$)
R_{had}	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta < 1.37$)
f_3	Ratio of the energy in the third layer to the total energy in the EM calorimeter
R_η	Ratio of the sum of the energies of the cells contained in a $3 \times 7 \eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 7×7 rectangle, both centred around the most energetic cell
$w_{\eta 2}$	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2) / (\sum E_i) - ((\sum E_i \eta_i) / (\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells
R_ϕ	Ratio of the sum of the energies of the cells contained in a $3 \times 3 \eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 3×7 rectangle, both centred around the most energetic cell
w_{stot}	Total lateral shower width, $\sqrt{(\sum E_i (i - i_{\text{max}})^2) / (\sum E_i)}$, where i runs over all cells in a window of $\Delta\eta \approx 0.0625$ and i_{max} is the index of the highest-energy cell
E_{ratio}	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies
f_1	Ratio of the energy measured in the first layer of the electromagnetic calorimeter to the total energy of the EM cluster
$n_{\text{innermost}}$	Number of hits in the innermost pixel layer
n_{Pixel}	Number of hits in the pixel detector
n_{Si}	Total number of hits in the pixel and SCT detectors
d_0	Transverse impact parameter relative to the beam-line
$ d_0 / \sigma(d_0) $	Significance of transverse impact parameter defined as the ratio of d_0 to its uncertainty
$\Delta p / p$	Momentum lost by the track between the perigee and the last measurement point divided by the momentum at perigee
eProbabilityHT	Likelihood probability based on transition radiation in the TRT
$\Delta\eta_1$	$\Delta\eta$ between the cluster position in the first layer of the EM calorimeter and the extrapolated track
$\Delta\phi_{\text{res}}$	$\Delta\phi$ between the cluster position in the second layer of calorimeter and the momentum-rescaled track, extrapolated from the perigee, times the charge q
E / p	Ratio of the cluster energy to the measured track momentum

4 Electron and photon reconstruction, calibration, identification and isolation

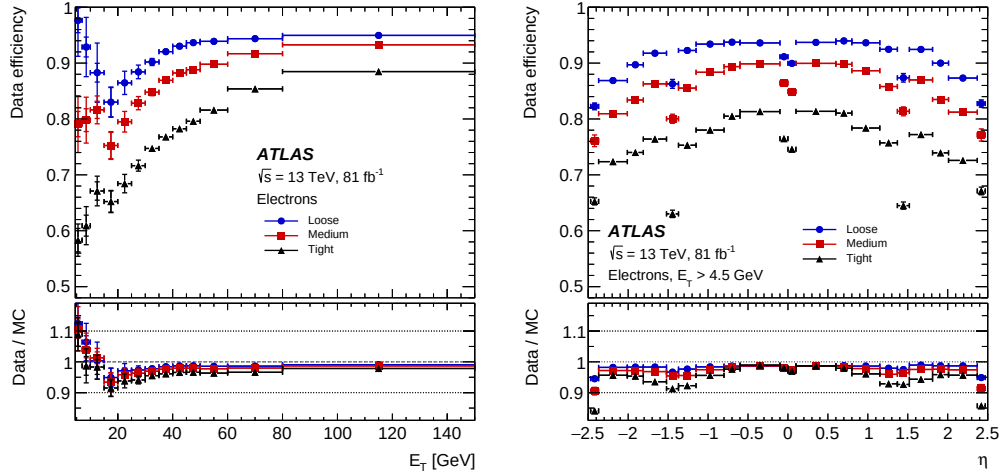


Figure 4.7: Efficiencies of the electron identification operating points measured using $Z \rightarrow ee$ events in data as a function of electron transverse energy (left) and electron pseudorapidity (right). The data and simulation ratio is shown at the bottom panel of each plot. The uncertainties include both statistical and systematic components. From reference [132].

This discriminant, d_L , which peaks around 1 for signal and 0 for background, is transformed by an inverse sigmoid function.

The probability density functions are measured using $Z \rightarrow ee$ and $J/\psi \rightarrow ee$ events in bins of the electron transverse energy and pseudorapidity, and smoothed to avoid non-physical fluctuations.

Three operating points, denoted as *Loose*, *Medium* and *Tight*, are defined to achieve different background rejection factors or signal efficiencies. These operating points are set to have signal efficiencies of 93%, 88%, and 80%, each with its corresponding cuts in the variables serving as selection criteria and a limit in the likelihood discriminant to achieve the specified efficiency.

The measurements of the efficiency of the operating points as functions of electron transverse energy and pseudorapidity are shown in Figure 4.7.

4.4 Photon identification

Photon candidates undergo an identification process to enhance their purity by rejecting photon candidates reconstructed from deposits originated by hadronic jets. This process uses several variables in a cut-based selection. The complete list of the variables involved in the selection, along with their descriptions, is presented in Table 4.2.

Table 4.2: Discriminating variables used in the photon identification.

Variable	Description
R_{had_1}	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the ranges $ \eta < 0.8$ and $ \eta > 1.37$)
R_{had}	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \eta < 1.37$)
R_η	Ratio of the sum of the energies of the cells contained in a $3 \times 7 \eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 7×7 rectangle, both centred around the most energetic cell
$w_{\eta 2}$	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2) / (\sum E_i) - ((\sum E_i \eta_i) / (\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells
R_ϕ	Ratio of the sum of the energies of the cells contained in a $3 \times 3 \eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 3×7 rectangle, both centred around the most energetic cell
w_{stot}	Total lateral shower width, $\sqrt{(\sum E_i (i - i_{\text{max}})^2) / (\sum E_i)}$, where i runs over all cells in a window of $\Delta\eta \approx 0.0625$ and i_{max} is the index of the highest-energy cell
w_{s3}	Lateral shower width, $\sqrt{(\sum E_i (i - i_{\text{max}})^2) / (\sum E_i)}$, where i runs over all cells in a window of 3 cells around the highest-energy cell
f_{side}	Energy fraction outside core of three central cells, within seven cells
ΔE_s	Difference between the energy of the cell associated with the second maximum, and the energy reconstructed in the cell with the smallest value found between the first and second maxima
E_{ratio}	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies
f_1	Ratio of the energy measured in the first layer of the electromagnetic calorimeter to the total energy of the EM cluster

4 Electron and photon reconstruction, calibration, identification and isolation

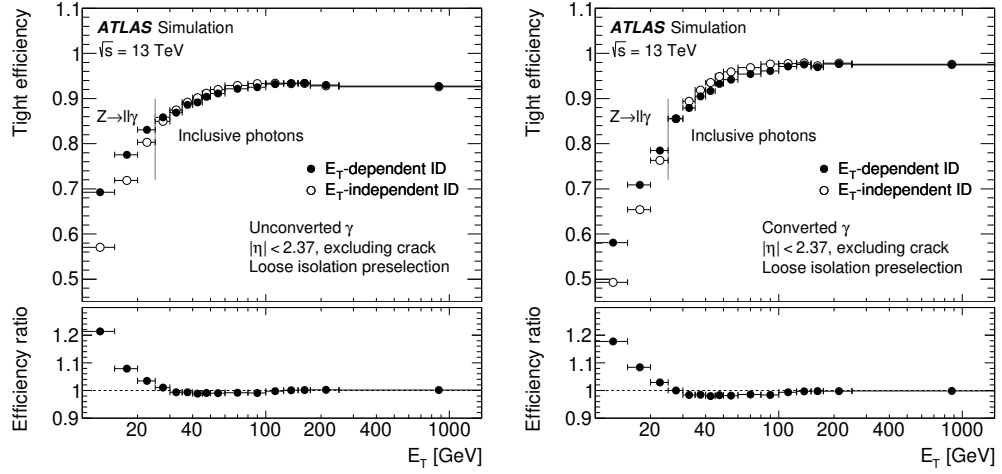


Figure 4.8: Efficiency of the *Tight* photon identification operating point measured using $Z \rightarrow \ell\ell\gamma$ events for unconverted (left) and converted photons (right). In each case, the E_T -dependent and E_T -independent optimisations are compared. The ratio of both optimisations is shown at the bottom panel of each plot. From reference [132].

The primary identification selection, denoted as *Tight*, includes cuts on eleven variables. The *Medium* and *Loose* selections, used in the triggering process, are less restrictive and optimised for different regions of pseudorapidity. The *Loose* selection uses the R_{had} , $R_{\text{had}1}$, R_η , and $w_{\eta2}$ shower-shape variables. The *Medium* selection adds a loose cut on E_{ratio} . Both *Loose* and *Medium* identifications do not differentiate between converted and unconverted photons and are optimised in regions of pseudorapidity. These identifications serve as the base for the triggering of photons.

The *Tight* selection distinguishes between converted and unconverted photons and is optimised in regions of $|\eta|$ and transverse energy. It includes cuts on eleven variables: R_{had} , $R_{\text{had}1}$, R_η , $w_{\eta2}$, R_ϕ , w_{stot} , w_{s3} , f_{side} , ΔE_s , E_{ratio} , and f_1 . This identification is optimised using simulations. For photons with $10 < E_T < 25$ GeV, a $Z \rightarrow \ell\ell\gamma$ simulated sample is used as signal and data from $Z + \text{jets}$ events are used as background. Above $E_T = 25$ GeV, simulations of inclusive photons are used as signal and a dijet simulation is used as background.

The efficiency of the *Tight* photon identification operating point after the optimisation is measured separately for converted and unconverted photons. The efficiency optimised in E_T is referred as E_T -dependent and compared with a reference identification efficiency that is not optimised. The measurement is shown in Figure 4.8.

4.5 Electron and photon isolation

The activity in the vicinity of electrons and photons can be quantified from energy deposits in the calorimeter and the momentum tracks. This leads to two different classes of isolation variables.

The variable based on calorimeter deposits is called E_T^{coneXX} , representing the transverse energy around an electron or photon. To compute this variable, the raw transverse energy, $E_{T,\text{raw}}^{\text{isolXX}}$, is computed by summing the transverse energy of positive-energy topo-clusters whose barycentre is inside a cone centred on the electron or photon barycentre. The cone used is denoted as XX in the variable, where XX refers to its size as $\Delta R = \text{XX}/100$. The typical cones used are $\Delta R = 0.2$ for electrons and $\Delta R = 0.2$ and 0.4 for photons. Since the raw transverse energy includes the energy of the particle, the energy of the electromagnetic calorimeter cells in a $\Delta\eta \times \Delta\phi = 5 \times 7$ window around the barycentre of the particle cluster, $E_{T,\text{core}}$, is subtracted. The usage of a fixed-size window is advantageous because of the stable subtraction for real or fake/non-prompt objects. The disadvantage is that the energy outside the window is not subtracted. To correct for this effect, the leakage is parameterised as a function of the transverse energy and $|\eta|$ using simulations without pile-up. Additionally, another correction to account for the pile-up and underlaying-event contributions to the isolation cone is estimated in regions of $|\eta|$. This correction uses the jet-area method [136], which consist in computing and ambient transverse energy density produced by low-energy jets that is multiplied by the isolation cone area. After applying these corrections, the calorimeter isolation variable is computed as:

$$E_T^{\text{coneXX}} = E_{T,\text{raw}}^{\text{isolXX}} - E_{T,\text{core}} - E_{T,\text{leakage}}(E_T, |\eta|, \Delta R) - E_{T,\text{pile-up}}(|\eta|, \Delta R).$$

The variable based on the tracks around the electron or photon, p_T^{coneXX} , is computed by adding the transverse momentum of tracks around the electron track or photon cluster. For electrons and converted photons, the associated track is removed. In the case of electrons produced in the decay of heavy particles, other decay products are closer when the transverse momentum is high. For this reason, electrons use a variable cone size to define the track isolation variable $p_T^{\text{varconeXX}}$ whose cone depends on the transverse momentum as:

$$\Delta R = \min\left(\frac{10 \text{ GeV}}{p_T}, \Delta R_{\text{max}}\right),$$

where ΔR_{max} is the maximum cone size, typically 0.2. To compute p_T^{coneXX} and $p_T^{\text{varconeXX}}$, tracks are required to have $p_T > 1 \text{ GeV}$ and $|\eta| < 2.5$, at least seven hits in the pixel detector and SCT, at most one shared hit – defined as $(n_{\text{Pixel}}^{\text{sh}} + n_{\text{SCT}}^{\text{sh}})/2$, where $n_{\text{Pixel}}^{\text{sh}}$ and $n_{\text{SCT}}^{\text{sh}}$ are the numbers of hits shared by several tracks in the pixel detector SCT –, at most two layers without hits in the pixel detector and SCT and at most one layer without hit in the pixel

4 Electron and photon reconstruction, calibration, identification and isolation

Table 4.3: Definitions of the electron isolation criteria. In the Gradient working point definition, the unit of p_T is GeV.

Working point	Calorimeter isolation	Track isolation
Gradient	$\epsilon = 0.1143 \times p_T + 92.14\%$ (with E_T^{cone20})	$\epsilon = 0.1143 \times p_T + 92.14\%$ (with $p_T^{\text{varcone20}}$)
HighPtCaloOnly	$E_T^{\text{cone20}} < \max(0.015 \times p_T, 3.5 \text{ GeV})$	-
Loose	$E_T^{\text{cone20}} / p_T < 0.20$	$p_T^{\text{varcone20}} / p_T < 0.15$
Tight	$E_T^{\text{cone20}} / p_T < 0.06$	$p_T^{\text{varcone20}} / p_T < 0.06$

detector. Additionally, for electrons, the track must be used in the primary vertex fit, or if it was not used in any vertex fit, it must satisfy $|\Delta z_0 \sin \theta| < 3 \text{ mm}$, where $|\Delta z_0|$ is the longitudinal impact parameter relative to the chosen primary vertex. For photons, tracks are required to satisfy $|\Delta z_0 \sin \theta| < 3 \text{ mm}$.

4.5.1 Electron isolation criteria

ATLAS uses a set of four different electron-isolation working points. They can be defined in two different ways: setting a cut on the isolation variables or fixing a target efficiency. A summary of the set of working points is shown in Table 4.3. The *Gradient* working point is set to give an efficiency of 90% at $p_T = 25 \text{ GeV}$ and 99% at $p_T = 60 \text{ GeV}$, uniform in pseudorapidity. Simulations of $J/\psi \rightarrow ee$ for $E_T < 15 \text{ GeV}$ and $Z \rightarrow ee$ for $E_T > 15 \text{ GeV}$ are used to define the cuts in E_T^{cone20} and $p_T^{\text{varcone20}}$. For *HighPtCaloOnly* only a cut in the calorimeter isolation variable is used, and for *Tight* and *Loose*, both calorimeter and track isolation variable fixed cuts are used.

The measurement of the efficiency of the electron isolation working points is computed using data recorded in 2017 in a sample enriched with $Z \rightarrow ee$ events. The efficiencies are measured as a function of electron transverse energy and pseudorapidity as shown in Figure 4.9. In the region below 15 GeV, a correction using $J/\psi \rightarrow ee$ events is applied.

4.5.2 Photon isolation criteria

Three photon-isolation working points are used in ATLAS: *Loose*, *Tight*, and *TightCaloOnly*. All of them use fixed cuts in the isolation variables as summarised in Table 4.4. For *TightCaloOnly*, only the calorimeter isolation variable is used; for *Tight* and *Loose*, both calorimeter and track isolation variable fixed cuts are used.

A discrepancy for the calorimeter isolation variables between data and simulation has been observed since Run 1 [137]. This discrepancy is caused by a mismodelling in the simulation of the lateral profile development of the electromagnetic showers. This causes a shift in the simulated distribution of the calorimeter photon isolation variables from data and is

4.5 Electron and photon isolation

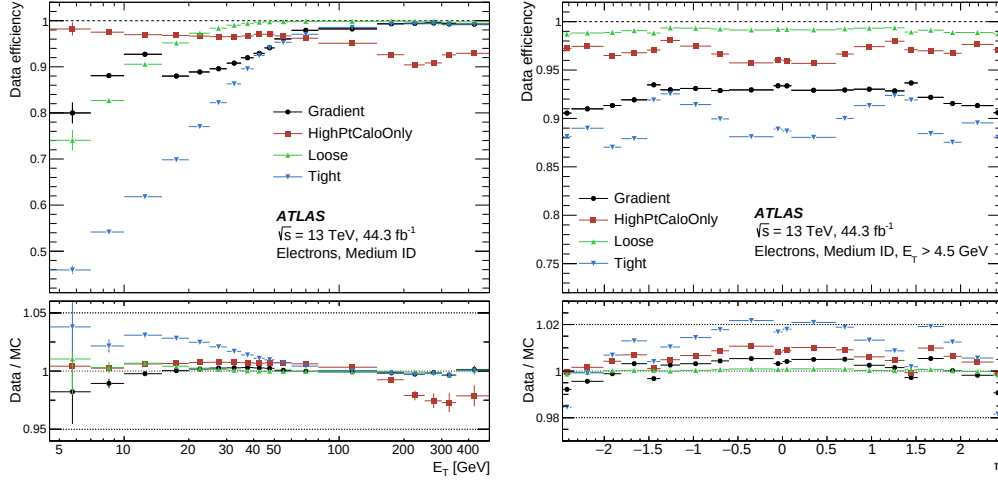


Figure 4.9: Efficiencies of the electron isolation working points measured using $Z \rightarrow ee$ events as a function of electron transverse energy (left) and pseudorapidity (right). The data and simulation ratio is shown in the bottom panel of each plot. The uncertainties include the statistical and systematic components. From reference [132].

Table 4.4: Definitions of the photon isolation criteria.

Working point	Calorimeter isolation	Track isolation
Loose	$E_T^{\text{cone20}} < 0.065 \times E_T$	$p_T^{\text{cone20}} / E_T < 0.05$
Tight	$E_T^{\text{cone40}} < 0.022 \times E_T + 2.45 \text{ GeV}$	$p_T^{\text{cone20}} / E_T < 0.05$
TightCaloOnly	$E_T^{\text{cone40}} < 0.022 \times E_T + 2.45 \text{ GeV}$	-

corrected by a data-driven method. This method measures the difference between the peaks in data and simulation in bins of transverse energy and pseudorapidity differentiating converted and unconverted photons. Then the measured shift is applied to correct the simulation.

The photon isolation efficiency is measured using two main signatures: radiative Z decays, namely $Z \rightarrow \ell\ell\gamma$ events for $10 < E_T < 100 \text{ GeV}$, and inclusive photons for $25 \text{ GeV} < E_T < 1.5 \text{ TeV}$. In both cases, measurements are done as functions of photon pseudorapidity and transverse energy and separately for converted and unconverted photons with data taken in 2017. The measurement of the efficiencies in the low- E_T range with $Z \rightarrow \ell\ell\gamma$ events is shown in Figure 4.10. For the inclusive photon measurements, it is done separately for the calorimeter and track isolation variables. The efficiency measurement for the calorimeter isolation variable is shown in Figure 4.11 and applies the track isolation variable in the preselection. For the track isolation variable efficiency shown in Figure 4.12, the calorimeter isolation variable is applied in the preselection.

4 Electron and photon reconstruction, calibration, identification and isolation

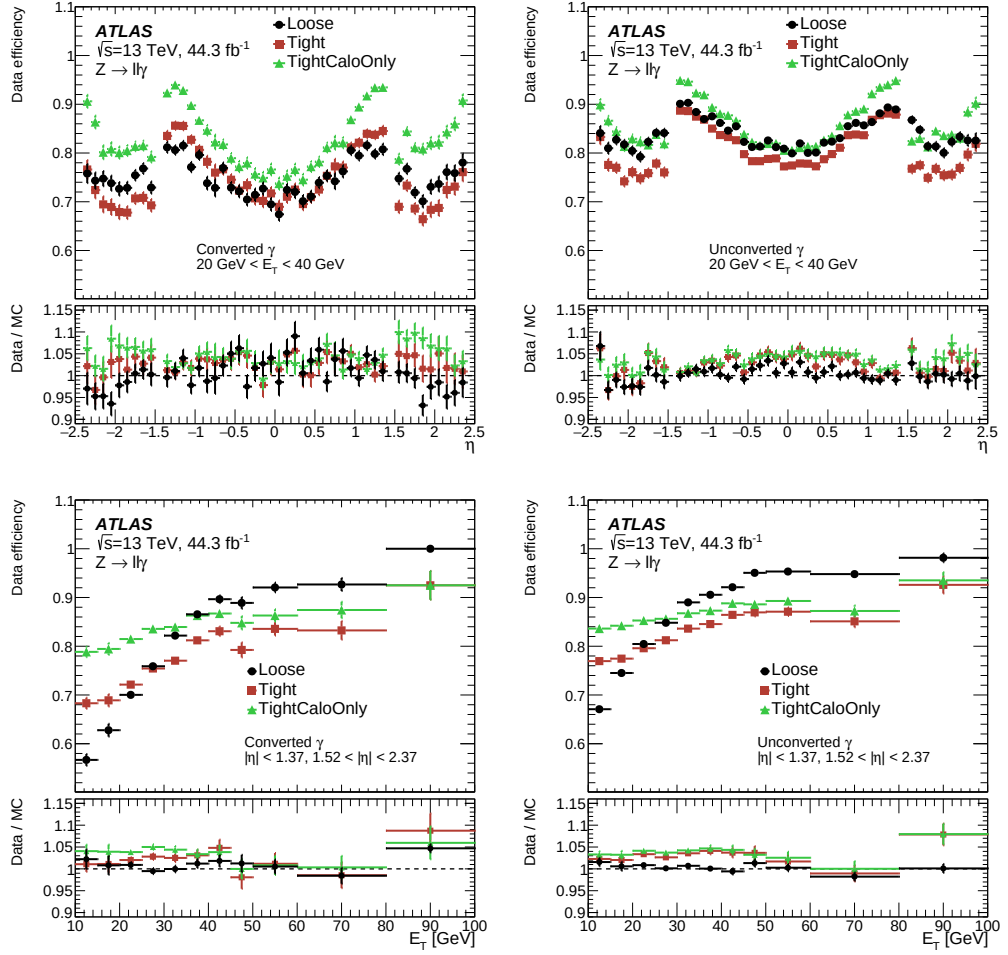


Figure 4.10: Efficiencies of the photon isolation working points measured using $Z \rightarrow \ell\ell\gamma$ events as a function of photon pseudorapidity (top) and transverse energy (bottom) for converted (left) and unconverted photons (right). The data and simulation ratio is shown in the bottom panel of each plot. The uncertainties include the statistical and systematic components. From reference [132].

4.5 Electron and photon isolation

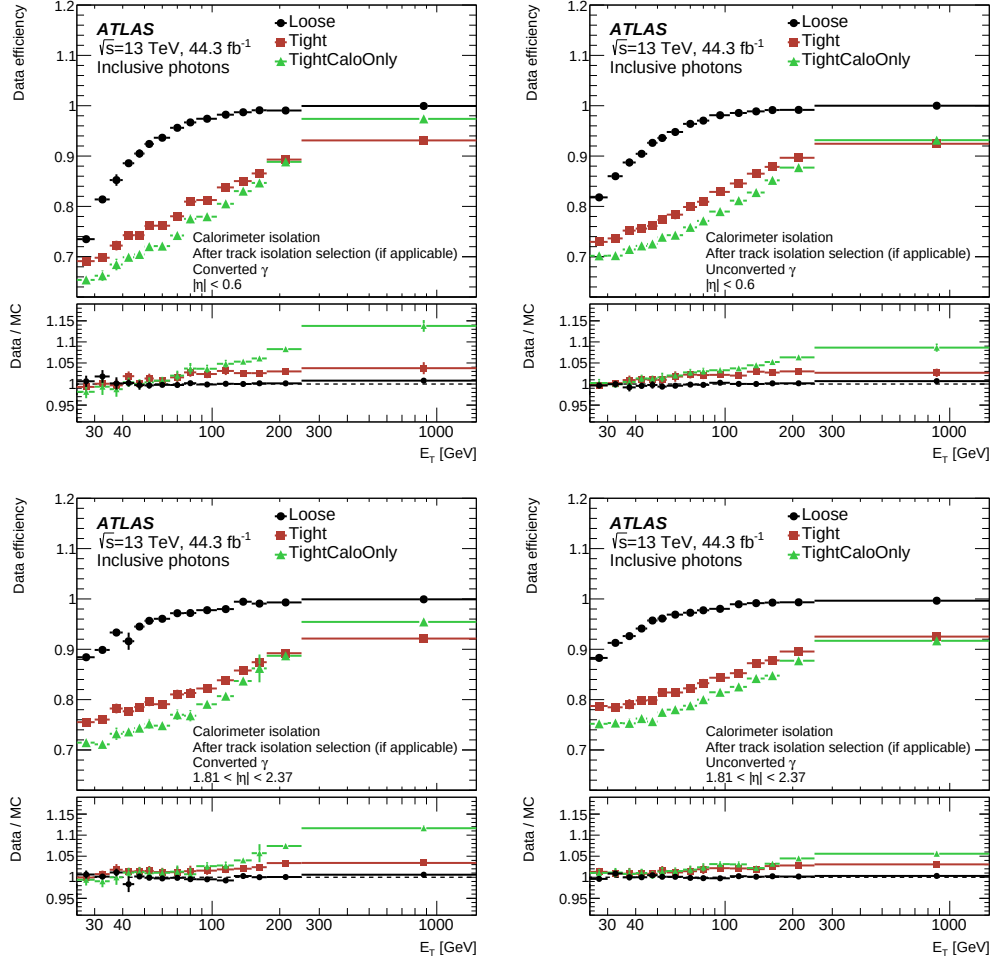
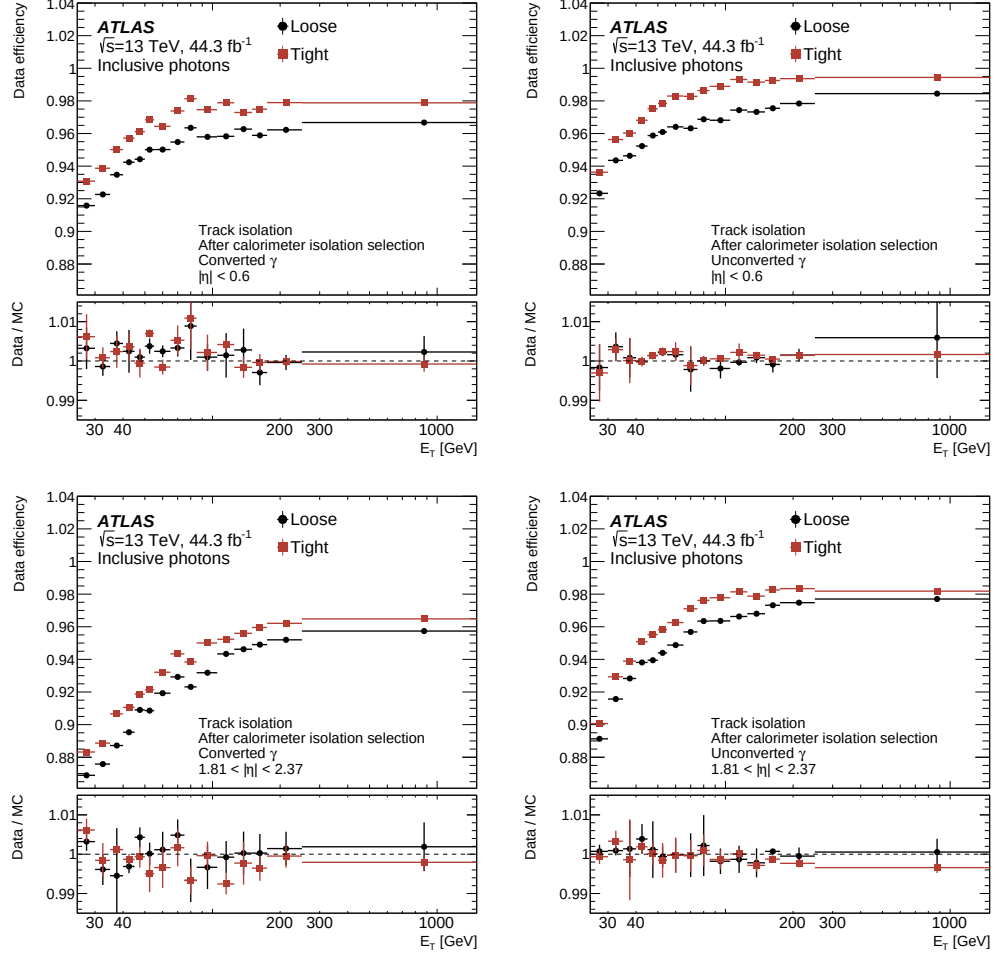


Figure 4.11: Efficiencies of the calorimeter isolation working points for photons measured using inclusive photon events as a functions of photon transverse energy for $|\eta| < 0.6$ (top) and $1.81 < |\eta| < 2.37$ (bottom) for converted (left) and unconverted photons (right). The data and simulation ratio is shown in the bottom panel of each plot. The uncertainties include the statistical and systematic components. From reference [132].

4 Electron and photon reconstruction, calibration, identification and isolation



5 Muon reconstruction, calibration, identification and isolation

In this chapter, a description of the procedure to build muon objects used in the analysis following the methodology employed by the ATLAS Collaboration during Run 2 [138, 139] is presented.

5.1 Muon reconstruction

Muon reconstruction in ATLAS is primarily based on the tracking provided by the MS. Additionally, information from the ID and calorimeters is used. This gives rise to several muon types: combined, inside-out combined, MS extrapolated, segment-tagged and calorimeter-tagged reconstruction.

In the MS track reconstruction the first step is the identification of short straight-line track segments from an individual station. Segments are combined using the IP and a parabolic trajectory — the first-order approximation to the muon bending in the magnetic field. Then, these candidates are fitted to a muon trajectory accounting for interactions with the detector material. After the fit, outlier hits are removed, and missing hits within the trajectory are added to perform the fit again. In the case of tracks sharing a large fraction of hits, the lower-quality track is removed. An exception is made for tracks with identical hits in two stations and no shared hits in the third one. Then, a final fit is performed constrained by the IP and accounting for the energy loss in the calorimeters. The track is extrapolated to the beam line to correlate the p_T to the IP.

The combined reconstruction matches MS and ID tracks and performs a fit, taking into account the energy loss in the calorimeter. This trajectory is used to remove outlier hits and include missed ones that fall within the trajectory. After this step, the fit is repeated. In the $|\eta| > 2.5$ region, MS tracks are combined with track segments from the pixel detector and SCT. These types of combined muons are referred to as silicon-associated forward muons.

The inside-out combined muons are reconstructed by extrapolating ID tracks to the MS. At least three hits in the MS are required, and a combined fit is performed, accounting for energy loss. This increases the efficiency for low- p_T muons that may not reach the middle station of the MS.

5 Muon reconstruction, calibration, identification and isolation

The MS extrapolated reconstruction is applied when the MS track does not match any ID track. This increases the coverage region outside the ID range up to $|\eta| = 2.7$.

Segment-tagged muons are those in which the ID track extrapolated to the MS is tightly matched in ΔR with a reconstructed MS segment. In this case, the muon parameters are obtained only from the ID track information.

Calorimeter-tagged muons are reconstructed from ID tracks extrapolated to the calorimeter. If the calorimeter deposits are compatible with a minimum-ionising particle, the ID track is tagged as a muon candidate.

The ID tracks used in the combined, inside-out and segment-tagged reconstruction are required to have a $p_T > 2$ GeV, while calorimeter-tagged muons use a threshold of 5 GeV.

5.2 Muon calibration

A calibration process is applied to muons to correct their transverse momentum. This correction is performed separately for the ID and the MS and is given by:

$$p_T^{\text{Cor,Det}} = \frac{p_T^{\text{MC,Det}} + \sum_{n=0}^1 s_n^{\text{Det}}(\eta, \phi) \left(p_T^{\text{MC,Det}}\right)^n}{1 + \sum_{m=0}^2 \Delta r_m^{\text{Det}}(\eta, \phi) \left(p_T^{\text{MC,Det}}\right)^{m-1} g_m},$$

where $p_T^{\text{Cor,Det}}$ is the corrected transverse momentum for each detector, Det = ID, MS, and $p_T^{\text{MC,Det}}$ is the simulated uncorrected transverse momentum. The s_n^{Det} and Δr_m^{Det} are the momentum scale and resolution as functions of η and ϕ . The $\eta - \phi$ regions are selected to cover homogeneous detector areas.

The scale correction depends on two terms. The s_0^{Det} term corrects the inaccuracy in the simulation of the energy loss in the calorimeter and other materials between the interaction point and the detector. For the ID detector, s_0^{ID} is set to zero as the energy loss is negligible. The s_1^{Det} term models inaccuracy in the description of the magnetic field integral and the dimension of the detector in the direction perpendicular to the magnetic field.

The resolution correction is characterised by three parameters assuming that the relative resolution can be modelled as:

$$\frac{\sigma(p_T)}{p_T} = \frac{\Delta r_0}{p_T} \oplus \Delta r_1 \oplus \Delta r_2 \cdot p_T,$$

where $\oplus$ denotes the sum in quadrature. The Δr_0 term describes the fluctuations of the energy loss in the material and is set to zero for both ID and MS as they are negligible. The

second term, Δr_1 , models the local magnetic field inhomogeneities, multiple scattering and displacement of the hits in the transverse direction. The third term, Δr_2 , accounts for the spatial resolution of the hit measurements and misalignment of the MS.

These seven ($\Delta r_1^{\text{ID}}, \Delta r_2^{\text{ID}}, s_1^{\text{ID}}, \Delta r_1^{\text{MS}}, \Delta r_2^{\text{MS}}, s_0^{\text{MS}},$ and s_1^{MS}) calibration constants are extracted using $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events. For the ID correction, $\Delta r_1^{\text{ID}}, \Delta r_2^{\text{ID}},$ and s_1^{ID} are fitted in regions of $\eta - \phi$ using $m_{\mu\mu}^{\text{ID}}$. Similarly, for the MS constants, the $m_{\mu\mu}^{\text{MS}}$ distribution is used for the fits. In the case of the $Z \rightarrow \mu\mu$ events, an additional variable,

$$\rho = \frac{p_{\text{T}}^{\text{MS}} - p_{\text{T}}^{\text{Cor,ID}}}{p_{\text{T}}^{\text{Cor,ID}}},$$

which measures the p_{T} imbalance between the measurement in the MS and the corrected transverse momentum in the ID, is used.

From the corrected transverse momenta of the ID and MS, the combined p_{T} is computed by performing a weighted average:

$$p_{\text{T}}^{\text{Cor,CB}} = f \cdot p_{\text{T}}^{\text{Cor,ID}} + (1 - f) p_{\text{T}}^{\text{Cor,MS}},$$

where the weight f is derived from simulations by solving the linear equation

$$p_{\text{T}}^{\text{MC,CB}} = f \cdot p_{\text{T}}^{\text{MC,ID}} + (1 - f) p_{\text{T}}^{\text{MC,MS}}.$$

After this calibration, the invariant mass of the dimuon system and its mass resolution are shown in Figures 5.1 and 5.2. The measurements are extracted as a function of the leading muon pseudorapidity with $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events.

5.3 Muon identification

Muon candidates go through an identification procedure to enhance their purity. This procedure separates prompt muons — those produced in the hard scattering — and non-prompt muons, which are the product of the hadron decay. A set of variables is used to establish a selection, including the number of precision stations, precision hole stations, q/p compatibility and ρ' , as detailed in the following:

- The number of precision stations is the count of MS stations with at least three hits in the MDT or CSC detectors.
- The number of precision hole stations is the quantity of MS stations with fewer than three hits and is missing at least three hits expected given its trajectory.

5 Muon reconstruction, calibration, identification and isolation

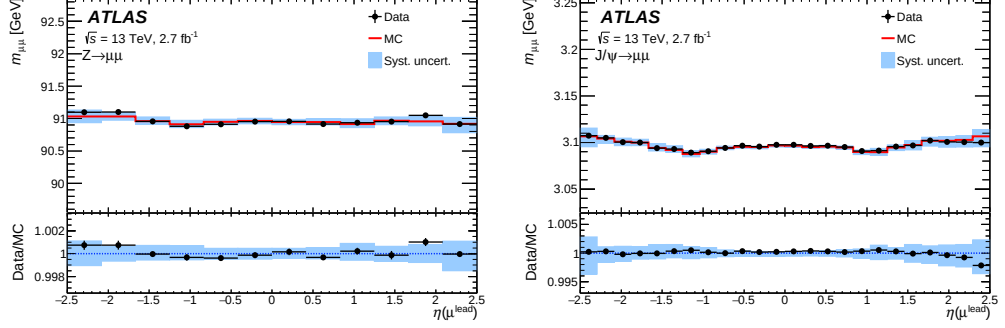


Figure 5.1: Mean invariant mass of the dimuon system for combined muons as a function of the leading muon pseudorapidity, using $Z \rightarrow \mu\mu$ events (left) and $J/\psi \rightarrow \mu\mu$ events (right). The measurements are represented as markers, and the simulation is shown as a line. The bottom panel shows the Data and MC ratio. The error bars represent the statistical uncertainty, and the shaded band represents the systematic uncertainty. From reference [139].

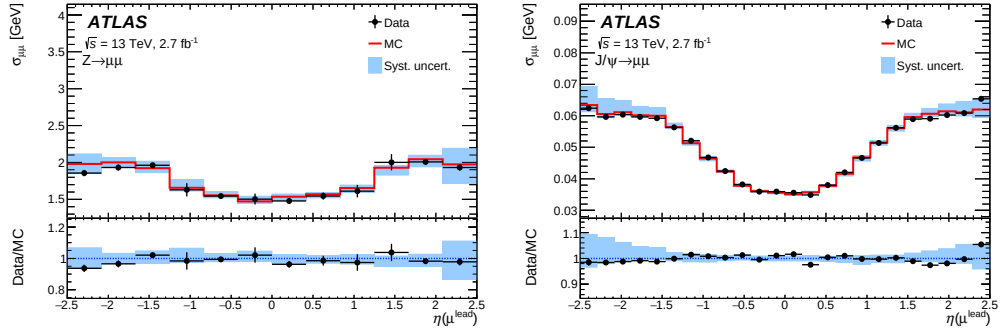


Figure 5.2: Invariant mass resolution of the dimuon system for combined muons as a function of the leading muon pseudorapidity, using $Z \rightarrow \mu\mu$ events (left) and $J/\psi \rightarrow \mu\mu$ events (right). The measurements are represented as markers, and the simulation is shown as a line. The bottom panel shows the Data and MC ratio. The error bars represent the statistical uncertainty, and the shaded band represents the systematic uncertainty. From reference [139].

- q/p compatibility is the agreement between the ID and MS in the charge-momentum ratio, defined for combined and inside-out muons as:

$$q/p \text{ compatibility} = \frac{|q/p_{\text{ID}} - q/p_{\text{MS}}|}{\sqrt{\sigma^2(q/p_{\text{ID}}) + \sigma^2(p/q_{\text{MS}})}},$$

where q/p_{ID} and q/p_{MS} are charge-momentum ratios in the ID and MS, and σ is the corresponding uncertainty.

- ρ' is the ratio of the difference in p_T measured by the ID and the MS divided by the combined track p_T , defined for combined and inside-out muons as:

$$\rho' = \frac{|p_{T,\text{ID}} - p_{T,\text{MS}}|}{p_{T,\text{CB}}},$$

where $p_{T,\text{ID}}$, $p_{T,\text{MS}}$, and $p_{T,\text{CB}}$ are the muon p_T measured by the ID, the MS, and the one obtained after combining hits from both ID and MS.

- The normalised χ^2 of the combined track fit.

Muons from combined, inside-out, segment-tagged and calorimeter-tagged reconstructions are required to have at least one hit in the pixel detector and at least five hits in the SCT, with at most two missing hits in the pixel detector and SCT. An exception is applied to silicon-associated forward muons, where the required hits in the SCT is reduced to four.

Three main working points are defined for muon identification: *Loose*, *Medium* and *Tight*. The *Medium* identification, in the $|\eta| < 2.5$ region, accepts only combined and inside-out muons. These are required to have two precision stations, except for muons in the $|\eta| < 0.1$ region, which only need one precision station if they have at most one precision hole station. A q/p compatibility smaller than seven is also required. In the $2.5 < |\eta| < 2.7$ region, MS extrapolated and silicon-associated forward muons are accepted if they have at least three precision stations. The *Loose* identification includes, apart from the *Medium* muons, calorimeter-tagged and segment-tagged muons in the $|\eta| < 0.1$ region. Additionally, inside-out muons with $p_T < 7$ GeV and $|\eta| < 1.3$ are accepted with one precision station if they are reconstructed as segment-tagged muons. For *Tight* muons, candidates are selected from *Medium* muons. Only combined and inside-out muons with at least two precision stations and the normalised $\chi^2 < 8$ are accepted. Additional requirements in q/p compatibility and ρ' are applied. These cuts are optimised in regions of p_T and $|\eta|$ to maximise the rejection of non-prompt muons.

Figure 5.3 shows the measured efficiencies for the three working points as functions of p_T using $J/\psi \rightarrow \mu\mu$ events and as functions of η using $Z \rightarrow \mu\mu$ events.

5 Muon reconstruction, calibration, identification and isolation

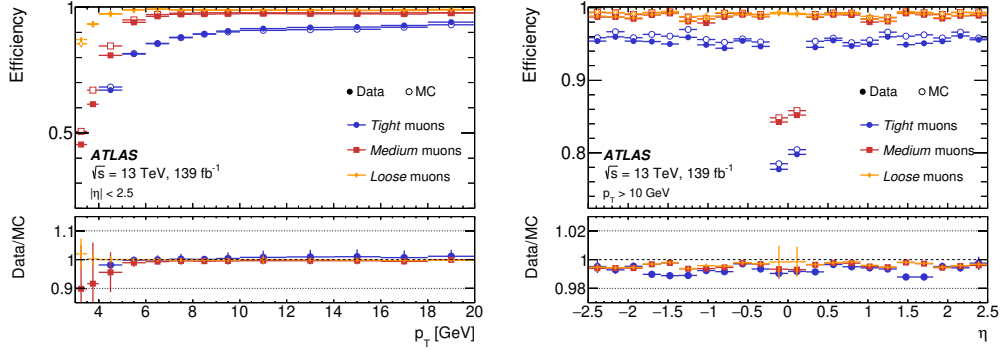


Figure 5.3: Efficiencies of the muon identification *Loose*, *Medium* and *Tight* working points. The left plot displays the efficiencies measured using $J/\psi \rightarrow \mu\mu$ events as a function of p_T . The right plot shows the efficiency measured using $Z \rightarrow \mu\mu$ events as a function of η . The measurements are displayed as filled markers and the predictions are displayed as open markers. The data and simulation ratio is shown in the bottom panel of each plot. The error bars include the statistical uncertainty in the top panel and both statistical and systematic uncertainties in the bottom panels of each plot. From reference [138].

For very high- p_T muons, a different working point is defined. This working point, named *High- p_T* , is designed to improve the momentum resolution of muons, which worsens rapidly with p_T . The *High- p_T* working point selects only combined and inside-out *Medium* muons. They are required to have at least three precision stations, except for muons crossing the magnetic field inversion zones which are equipped with additional chambers and require four precision stations. Another exception is made for muons in the $|\eta| < 1.3$ region with two precision stations if the missing hits are in the inner station. Muons from regions where the relative alignment of the traversed chambers is not known with sufficient precision are rejected. An additional cut is applied to the q/p relative uncertainty of the combined track. The threshold is optimised as a function of p_T and $|\eta|$.

Figure 5.4 shows the measured efficiency of the *High- p_T* working point as a function of η and ϕ using $Z \rightarrow \mu\mu$ events.

Similarly to the *High- p_T* identification, for the low- p_T region, a pair of working points is defined: the *Low- p_T* cut-based and the *Low- p_T* multivariate. For these working points, only combined and inside-out muons are used. The inside-out muons are required to be independently reconstructed as segment-tagged. Both combined and inside-out muons use at least one precision station, with the exception of the region $|\eta| > 1.3$, where two precision stations are required. For $p_T > 10$ GeV, the working point is defined as the *Medium* identification.

For these working points, three additional variables are used:

- The momentum balance significance (MBS) defined as:

5.3 Muon identification

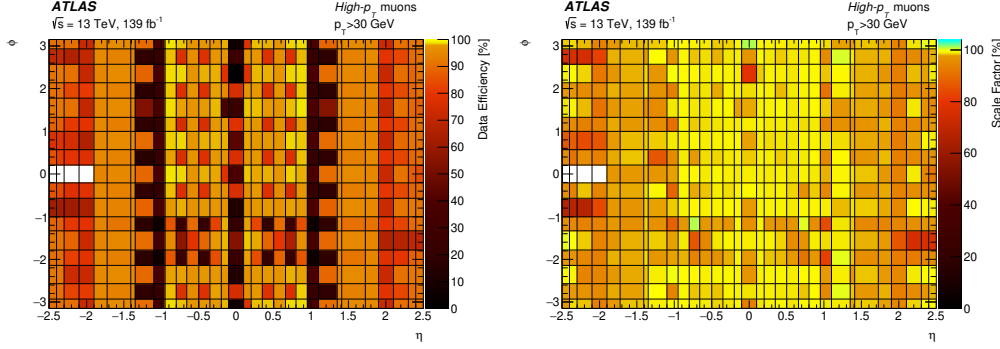


Figure 5.4: The left plot displays the measurement of the efficiency of the muon identification *High-p_T* working point as a function of η and ϕ using $Z \rightarrow \mu\mu$ events. The right plot shows the data/MC scale factor of the working point. From reference [138].

$$\text{MBS} = \frac{|p_{\text{ID}} - \hat{p}_{\text{MS}} - E_{\text{loss}}|}{\sigma(E_{\text{loss}})},$$

where p_{ID} and $\hat{p}_{\text{MS}}$ are the momentum measured in the ID and at the entrance of the MS, E_{loss} is the energy lost in the calorimeter system and $\sigma(E_{\text{loss}})$ is its uncertainty. If no momentum is measured in the MS, MBS is set to 0.

- The scattering neighbour significance (SNS) is defined as the largest value of the scattering angle significance over the entire track. It is evaluated considering pairs of adjacent hits and computed as the angular difference between the two half tracks divided by its uncertainty.
- The scattering curvature significance (SCS) is defined as the integral of scattering angle significances before and after the assumed decay vertex. It examines all potential pairs of partial tracks originating or concluding at each hit. The SCS value is calculated as the maximum absolute difference, among all pairs, between the two sums of significances along each partial track. This value is then normalised to the square root of the total number of pairs.

For the *Low-p_T* cut-based working point, each of the three significance variables needs to be lower than three. In the case of the *Low-p_T* multivariate, a gradient BDT is used separately for combined and inside-out muons. In addition to the three significance variables, the BDT uses the energy lost in the calorimeters, the number of MS segments associated with the muon and their direction relative to the track in the ID, and the number of missing precision hits in the middle MS station.

The efficiencies for the cut-based and multivariate *Low-p_T* working points are shown in Figure 5.5. The efficiencies are measured using $J/\psi \rightarrow \mu\mu$ events as functions of p_{T} and

5 Muon reconstruction, calibration, identification and isolation

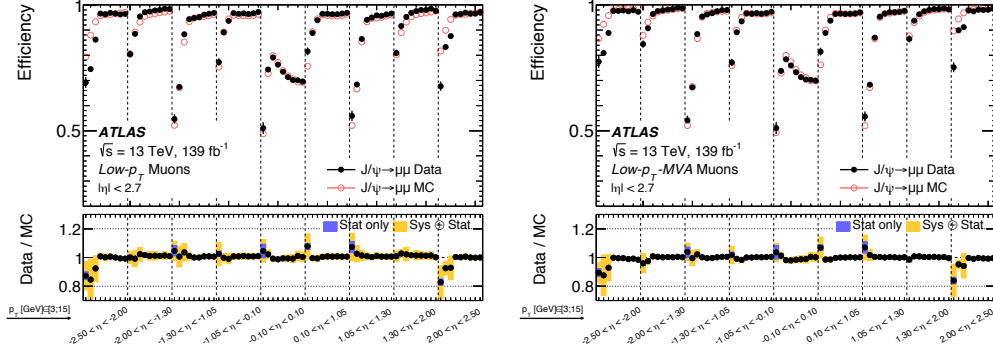


Figure 5.5: Muon identification efficiencies for the $Low-p_T$ working points. The cut-based (multivariate) identification is shown in the left (right) plot. The efficiency measurements, derived from $J/\psi \rightarrow \mu\mu$ events as functions of p_T and η are represented as filled markers, while the simulation is represented as open markers. Statistical uncertainties are represented by the error bars. The data and prediction ratio is displayed in the bottom panel of each plot, including both statistical and systematic uncertainties. From reference [138].

η . The p_T ranges from 3 to 15 GeV. The multivariate criteria show a larger efficiency and smaller uncertainties in the SFs.

5.4 Muon isolation

Muon isolation is crucial for discriminating muons originating from hard scatterings and those produced in hadron decays. Isolation is defined as the transverse energy within a cone around a muon divided by its transverse momentum. The computation employs information from either ID tracks, calorimeter topological cell clusters, or a combination of both.

Track-based isolation is the scalar sum of transverse momenta of ID tracks associated with the primary vertex in a cone in the $\eta - \phi$ plane of a certain size ΔR around the muon, subtracting the muon track. Two variables are defined with different ΔR . The first variable, p_T^{cone20} , sets $\Delta R = 0.2$. For cases with topologies where activity near the muon is expected, $p_T^{\text{varcone30}}$ is defined with $\Delta R = \min(10 \text{ GeV}/p_T^\mu, 0.3)$. To compute these variables, the minimum transverse momentum of a track being either 500 MeV or 1 GeV, depending on the working point.

Calorimeter-based isolation is the scalar sum of transverse energy of topological cell clusters within a cone in the $\eta - \phi$ plane of a certain size ΔR ; the contributions from the muon energy deposits and pile-up effects are subtracted. The calorimeter-based isolation variable is defined with $\Delta R = 0.2$, labelled $E_T^{\text{topocone20}}$.

A combination of track-based and calorimeter-based isolation variables is done to define the particle-flow-based isolation variable. This variable is defined as the weighted sum of the

Table 5.1: Definitions of the muon isolation criteria.

Working point	Isolation definition	Track p_T threshold
<i>PflowLoose</i> <i>PflowTight</i>	$(p_T^{\text{varcone30}} + 0.4 \cdot E_T^{\text{neflow20}}) < 0.16 \cdot p_T^\mu$ $(p_T^{\text{varcone30}} + 0.4 \cdot E_T^{\text{neflow20}}) < 0.045 \cdot p_T^\mu$	500 MeV
<i>Loose</i> <i>Tight</i>	$p_T^{\text{varcone30}} < 0.15 \cdot p_T^\mu, E_T^{\text{topocone20}} < 0.3 \cdot p_T^\mu$ $p_T^{\text{varcone30}} < 0.04 \cdot p_T^\mu, E_T^{\text{topocone20}} < 0.15 \cdot p_T^\mu$	1 GeV
<i>HighPtTrackOnly</i> <i>TightTrackOnly</i>	$p_T^{\text{cone20}} < 1.25 \text{ GeV}$ $p_T^{\text{varcone30}} < 0.06 \cdot p_T^\mu$	1 GeV
<i>PLBDTLoose</i> <i>PLBDTTight</i>	$p_T^{\text{varcone30}} < \max(1.8 \text{ GeV}, 0.15 \cdot p_T^\mu)$ BDT cut to mimic <i>TightTrackOnly</i> efficiency $p_T^{\text{varcone30}} < \max(1.8 \text{ GeV}, 0.15 \cdot p_T^\mu)$ BDT cut to mimic <i>Tight</i> efficiency	1 GeV

track-based isolation, $p_T^{\text{varcone30}}$, and the transverse energy of neutral particle-flow objects in a cone of size $\Delta R = 0.2$, E_T^{neflow20} . The weight for track-based isolation is set to 1, while for E_T^{neflow20} is 0.4. Contributions from muon energy deposit and pile-up are subtracted from E_T^{neflow20} .

A multivariate discriminant based on BDT, labelled *prompt lepton BDT*, is defined using eight different variables: calorimeter and track isolation, tracks information within a cone of size $\Delta R = 0.4$ around the muon (including track multiplicity), and the likelihood of originating within a jet from a b -hadron decay.

Eight isolation working points are defined, including *PflowLoose* and *PflowTight* using the particle-flow-based isolation variable, *Loose* and *Tight* using both track-based and calorimeter-based isolation variables, *HighPtTrackOnly* and *TightTrackOnly* using only track-based isolation variables and *PLBDTLoose* and *PLBDTTight* using the multivariate discriminant. For some of them (*PflowLoose*, *PflowTight*, *Loose*, *Tight*, *TightTrackOnly*) there is a second variation. In this variation, $p_T^{\text{varcone30}}$ is used in the range where $p_T^\mu < 50 \text{ GeV}$ and it is substituted by p_T^{cone20} for $p_T^\mu > 50 \text{ GeV}$. Table 5.1 summarises the definitions of these working points.

Figures 5.6 and 5.7 present the measured isolation efficiencies for *Loose*, *PflowLoose* and *PLBDTLoose* selections, as well as for *Tight*, *PflowTight* and *PLBDTTight* selections, respectively. The efficiencies are measured using $Z \rightarrow \mu\mu$ as functions of transverse momentum and ΔR (jet, μ).

5 Muon reconstruction, calibration, identification and isolation

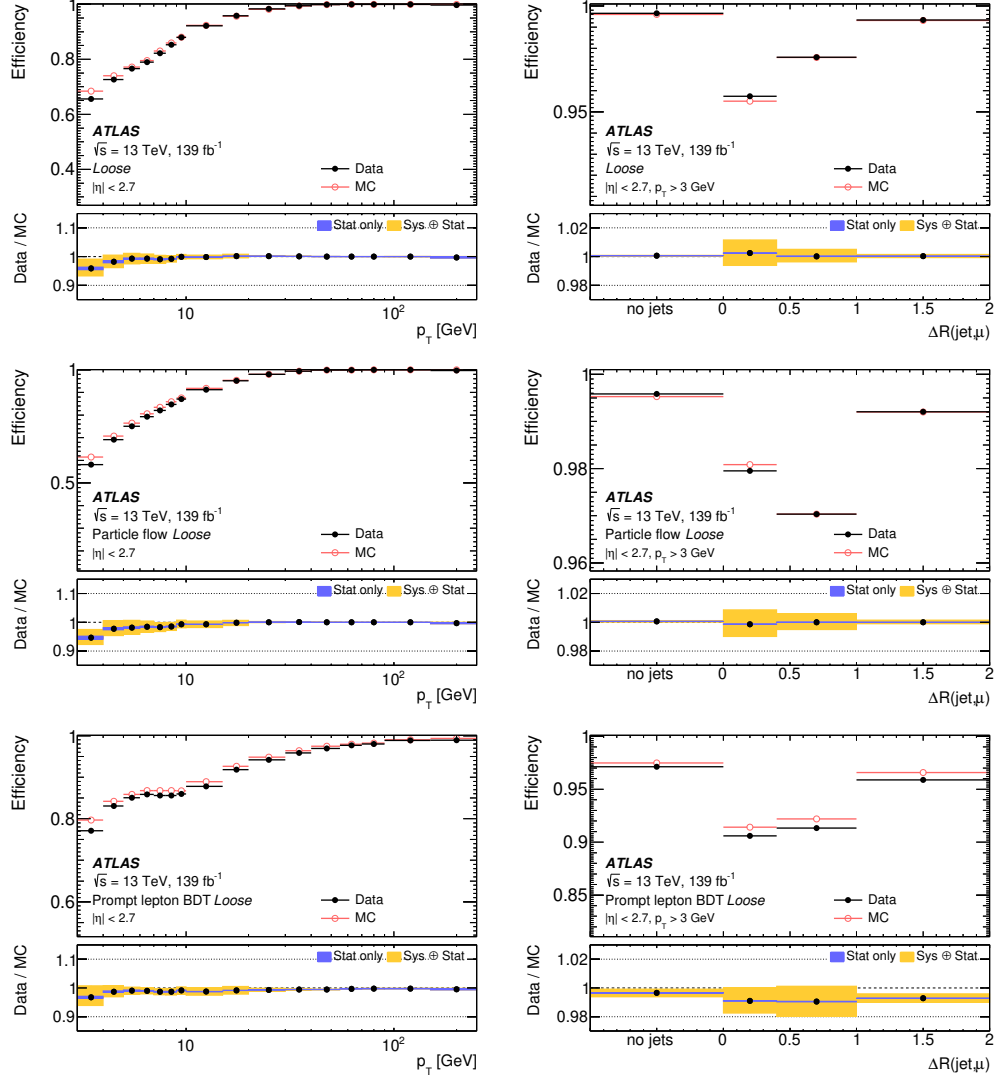


Figure 5.6: Efficiencies of muon isolation measured in $Z \rightarrow \mu\mu$ for Loose working points. The *Loose* working point is displayed in the top row, *PFlowLoose* in the middle row, and *PLBDTLoose* in the bottom row. The efficiency measurements are shown as functions of transverse momentum on the left column and as functions of $\Delta R(\text{jet}, \mu)$ on the right column. Measurements are represented as filled markers, and the simulation is represented as open markers. The data and prediction ratio is presented at the bottom panel of each plot, with statistical and systematic uncertainties. From reference [138].

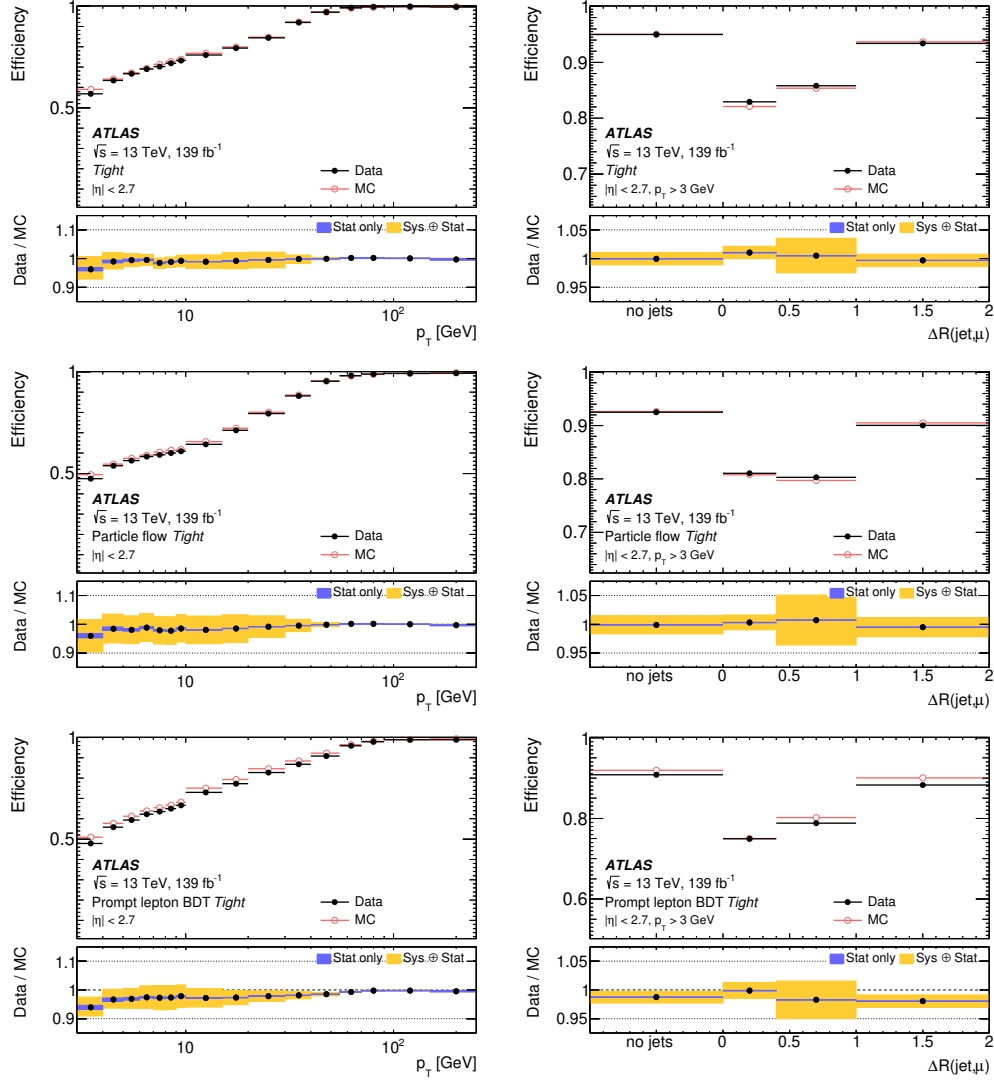


Figure 5.7: Efficiencies of muon isolation measured in $Z \rightarrow \mu\mu$ for Tight working points. The *Tight* working point is displayed in the top row, *PFlowTight* in the middle row, and *PLBDTTight* in the bottom row. The efficiency measurements are shown as functions of transverse momentum on the left column and as functions of $\Delta R(\text{jet}, \mu)$ on the right column. Measurements are represented as filled markers, and the simulation is represented as open markers. The data and prediction ratio is presented at the bottom panel of each plot, with statistical and systematic uncertainties. From reference [138].

6 Jet reconstruction, calibration and b -tagging

In this chapter, the reconstruction and calibration procedure for jets used by ATLAS in Run 2 is presented [140, 141]. Additionally, Section 6.3 provides a description of the pile-up jet suppression techniques using track information [142]. Finally, in Section 6.4, the b -jet identification method is summarised [143].

6.1 Jet reconstruction

The jet reconstruction process begins by constructing topo-clusters using the same algorithm employed for electrons and photons, as described in Section 4.1. These topo-clusters, along with the ID tracks, serve as input for the particle flow algorithm.

The particle flow algorithm, summarised in Figure 6.1, starts by applying a selection over the tracks. Subsequently, an attempt to match the tracks and a single topo-cluster is performed and the expected energy deposit in the calorimeter by the particle that created the track is computed. If the topo-cluster is not sufficiently energetic, additional topo-clusters are added. Finally, the expected energy deposited by the particle that produced the track is subtracted from the matched topo-clusters. If the remaining energy deposits are compatible with shower fluctuations, the remnants are removed.

The track selection step involves several cuts. Firstly, tracks are required to have at least nine hits in the pixel detector and SCT, with no missing hit when a hit would be expected. Additionally, tracks must be in the $|\eta| < 2.5$ region and have $p_T > 0.5$ GeV. Also, tracks with $p_T > 40$ GeV are not considered, as such energetic particles are poorly isolated, compromising the energy subtraction. Finally, any track matched with an electron or muon identified with medium quality criteria is rejected.

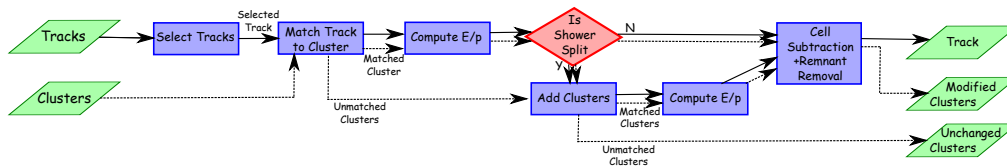


Figure 6.1: Flow diagram of the particle flow algorithm. From reference [140].

6 Jet reconstruction, calibration and b -tagging

Once topo-clusters are built and tracks are selected, the next step is matching them. For this process, tracks are sorted in descending p_T -order. For each track, topo-clusters are ranked by the distance defined as:

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

where $\Delta\phi$ and $\Delta\eta$ are the distances between the track extrapolated to the second layer of the calorimeter and the barycentre of the topo-cluster. Additionally, σ_ϕ and σ_η are the topo-cluster width computed as the standard deviation of the topo-cluster cells. A preliminary selection is applied by requiring $E^{\text{clus}}/p^{\text{trk}} > 0.1$, where E^{clus} is the energy of the topo-cluster and p^{trk} is the track momentum. After this preselection, the topo-cluster closest in $\Delta R'$ is selected. If no topo-cluster with $\Delta R' < 1.64$ is found, the track is kept unmatched.

Once topo-clusters are matched with tracks, the average energy deposited in the calorimeter by a particle with a certain p^{trk} is given by:

$$\langle E_{\text{dep}} \rangle = p^{\text{trk}} \left\langle \frac{E_{\text{ref}}^{\text{clus}}}{p_{\text{ref}}^{\text{trk}}} \right\rangle,$$

where $\langle E_{\text{ref}}^{\text{clus}}/p_{\text{ref}}^{\text{trk}} \rangle$ is determined using single-pion event simulated samples without pile-up. This involves summing the energies of topo-clusters inside a $\Delta R = 0.4$ cone around the track. It is measured in regions of the p_T and η of the track as well as a function of the layer of the highest energy density. This layer is defined as the one with the largest ratio between the increase in energy density and the increase in depth with respect to the previous layer. This ratio is computed as:

$$\Delta\rho'_i = \frac{\langle\rho'\rangle_i - \langle\rho'\rangle_{i-1}}{d_i - d_{i-1}},$$

where i denotes the layer, $\langle\rho'\rangle$ is the energy density, and d is the depth of the layer in interaction lengths. The values for $\langle\rho'\rangle_0$ and d_0 are taken as 0. For the EM calorimeter, the presampler is considered together with the first layer, and the Tile and HEC calorimeters are treated as single layers. The energy density of a layer is defined as the weighted average of the energy density of every cell in the layer. The weight of each cell is computed as the integral over the cell area in the $\eta - \phi$ plane of a Gaussian distribution centred on the extrapolated track with a width of 0.035. Additionally, the spread of the expected energy deposit, $\sigma(E_{\text{dep}})$, is computed as the standard deviation of the $E_{\text{ref}}^{\text{clus}}/p_{\text{ref}}^{\text{trk}}$ distribution in single-pion simulations.

The next step is to include satellite topo-clusters with the matched ones. For this stage, a discriminant distinguishes between cases where the particle deposits its energy in a single or multiple topo-clusters. This discriminant is defined as:

$$S(E^{\text{clus}}) = \frac{E^{\text{clus}} - \langle E_{\text{dep}} \rangle}{\sigma(E_{\text{dep}})}.$$

In the case when $S(E^{\text{clus}}) > -1$, no additional topo-clusters are added. Otherwise, topo-clusters within a cone of $\Delta R = 0.2$ around the track position extrapolated to the second layer of the calorimeter are considered to be matched to the track.

Once additional topo-clusters are matched when needed, the subtraction step is applied. In cases where $\langle E_{\text{dep}} \rangle$ is larger than the total energy of the topo-clusters, these are removed completely. Otherwise, a more sophisticated algorithm is applied cell by cell. For this subtraction, the most likely energy density profile is used. It is obtained from simulated single-pion samples as functions of the track momentum and pseudorapidity and the layer of the highest energy density. Then, density profiles are divided into rings in the $\eta - \phi$ plane for each layer, and the subtraction is applied starting with the more energetic rings. If the energy of a ring exceeds the energy of the cells in this region, cells are completely removed. In the other case, the energy of the cells inside the ring is scaled to compensate $\langle E_{\text{dep}} \rangle$.

After the subtraction, the remnant removal is applied when the remaining energy of the topo-cluster is smaller than $1.5\sigma(E_{\text{dep}})$. Then, a final cut removes tracks with $|\Delta z_0 \sin \theta| > 2$ mm to suppress pile-up effects. Finally, the anti- k_t algorithm is applied over the selected tracks and the remaining clusters with $R = 0.4$ to reconstruct the particle flow jets.

6.2 Jet calibration

Following the reconstruction process, the four-momentum of jets is corrected. This correction is carried out in two distinct steps: the jet energy scale calibration and the jet energy resolution.

6.2.1 Jet energy scale calibration

The jet energy scale calibration aims to restore the four-momentum of the jets to that at the particle level. This calibration is conducted through several steps as illustrated in Figure 6.2, starting with pile-up corrections to account for energy from additional proton-proton interactions. It then proceeds with the absolute calibration, global sequential calibration, and residual *in situ* calibration.

6 Jet reconstruction, calibration and b -tagging

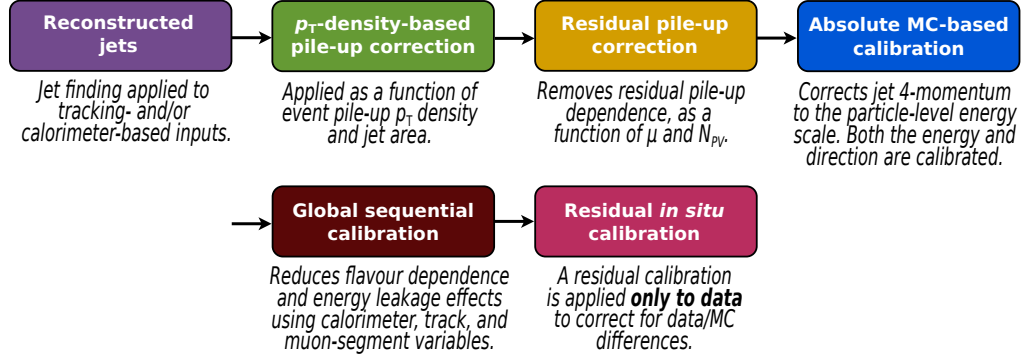


Figure 6.2: Flow diagram of the jet energy calibration process. From reference [141].

Pile-up corrections

Additional proton-proton interactions, resulting from the high luminosity of the LHC may increase the measured jet energy during the reconstructing process. To correct for this excess, a pile-up correction is applied. The first step involves a per-event p_T -density-based subtraction, computed as $\rho \times A$, where A represents the area of the jet in the $\eta - \phi$ plane, and ρ is the median jet p_T density ($\langle p_T/A \rangle$). The calculation of ρ uses jets with $|\eta| < 2$, reconstructed using the k_t algorithm with radius $R = 0.4$.

The cut in $|\eta|$ exploits the fine segmentation region of the calorimeter to characterise the pile-up activity. However, this description is not completely accurate in the forward region and a residual pile-up correction is performed by adding two of additional terms to the correction given by:

$$p_T^{\text{corr}} = p_T^{\text{reco}} - \rho \times A - \alpha \times (N_{PV} - 1) - \beta \times \mu,$$

where N_{PV} is the number of reconstructed primary vertices and μ is the mean number of interactions per bunch crossing, sensitive to in-time and out-of-time pile-up, respectively. To determine the dependence on these variables, α and β are computed using the difference between the p_T of the reconstructed and true jets in simulations, measured in regions of $|\eta|$ and p_T^{true} , the jet p_T at true level.

Absolute calibration

The absolute calibration, including jet energy scale and η calibration, corrects for energy losses in passive material, out-of-cone effects, effects on the η -dependent granularity of the

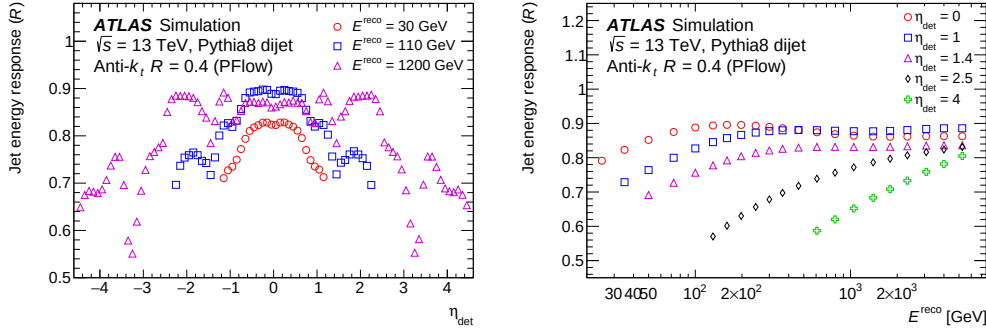


Figure 6.3: Jet energy response as functions of the reconstructed jet η (left) and E^{reco} (right). From reference [141].

calorimeter and non-compensating calorimeter response. This calibration is derived from the response obtained from simulations of dijet events for anti- k_t jets with radius $R = 0.4$.

The jet energy response $\mathcal{R}$ is computed using simulations as the mean of a Gaussian fit to the $E^{\text{reco}}/E^{\text{true}}$ distribution. It is measured in bins of E^{true} and η_{det} . The fit is performed as a function of E^{true} since the response distribution is Gaussian, while for E^{reco} this is not the case. The response is then parameterised as a function of E^{reco} using a numerical inversion procedure, and the jet calibration factor is taken as the inverse of the response. Figure 6.3 shows the jet energy response for PFlow jets as a function of E^{reco} and η_{det} in bins of η_{det} and E^{reco} , respectively.

A second correction, the η calibration, is then applied to reduce the deviation from zero in the signed difference between η^{reco} and η^{truth} , as shown in Figure 6.4. This correction is obtained as the difference between η^{reco} and η^{truth} as a function of E^{true} and η_{det} . Similarly to the jet energy scale response, a numerical inversion procedure is used to derive the correction as a function of E^{reco} .

Global sequential calibration

Following the absolute calibration, a series of corrections are implemented to address the response effects caused by sources such as flavour, energy distribution of constituents, and the transverse distribution. This correction is divided into six stages, with each stage correcting the four-momentum of the jet using different observables as functions of p_T^{true} and η_{det} . The six variables used in the global sequential calibration encompass the longitudinal structure of energy deposits observables, information on tracks associated with the jet, and information from muon chambers behind a jet. These variables are defined as:

- f_{charged} , the fraction of p_T measured from ghost-associated tracks with $p_T > 500$ MeV (for jets with $|\eta_{\text{det}}| < 2.5$);

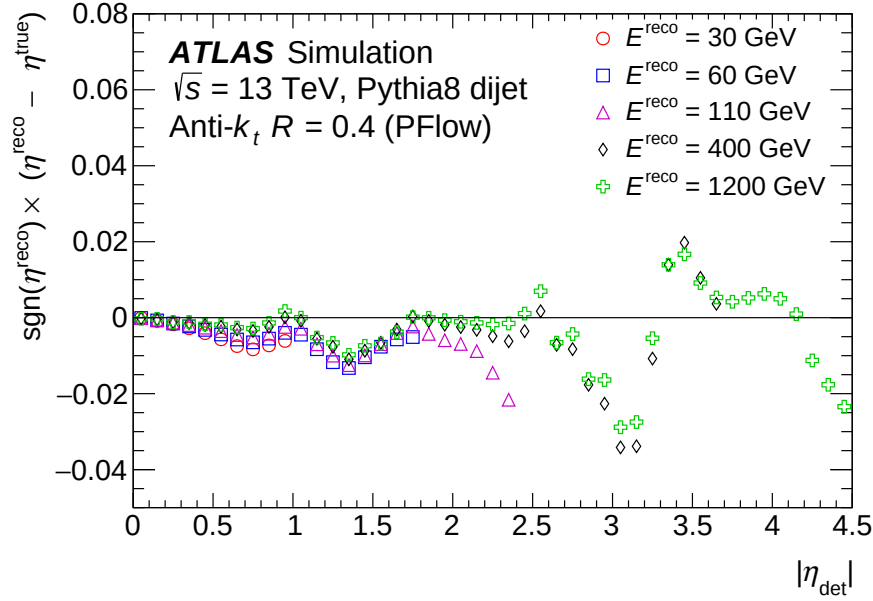


Figure 6.4: The signed difference between η^{reco} and η^{truth} as a function of $|\eta_{\text{det}}|$. From reference [141].

- f_{Tile0} , the fraction of jet energy measured in the first layer of the hadronic Tile calorimeter (for jets with $|\eta_{\text{det}}| < 1.7$);
- f_{LAr3} , the fraction of jet energy measured in the third layer of the electromagnetic LAr calorimeter (for jets with $|\eta_{\text{det}}| < 3.5$);
- n_{trk} , the number of tracks with $p_T > 1$ GeV ghost-associated with the jet (for jets with $|\eta_{\text{det}}| < 2.5$);
- w_{trk} , also known as track width, the average p_T -weighted transverse distance in the $\eta - \phi$ plane between the jet axis and all tracks of $p_T > 1$ GeV ghost-associated with the jet (for jets with $|\eta_{\text{det}}| < 2.5$);
- n_{segments} , the number of muon track segments ghost-associated with the jet (for jets with $|\eta_{\text{det}}| < 2.7$). In this step, the correction is obtained as a function of the jet energy instead of the p_T .

The jet p_T response, defined as the average of the $p_T^{\text{reco}}/p_T^{\text{true}}$ distribution as function of each global sequential calibration observable, is shown in Figure 6.5 for various p_T regions, except for the n_{segments} distribution, which uses energy regions. After applying the complete sequence of corrections, these dependencies are reduced to less than 2%, with the remaining deviation from unity resulting from unaccounted correlations between observables.

The fractional jet resolution, $\sigma_{\mathcal{R}}/\mathcal{R}$, where $\sigma_{\mathcal{R}}$ is the resolution of the response distribution computed as the standard deviation of a Gaussian distribution, is used to determine the size

6.2 Jet calibration

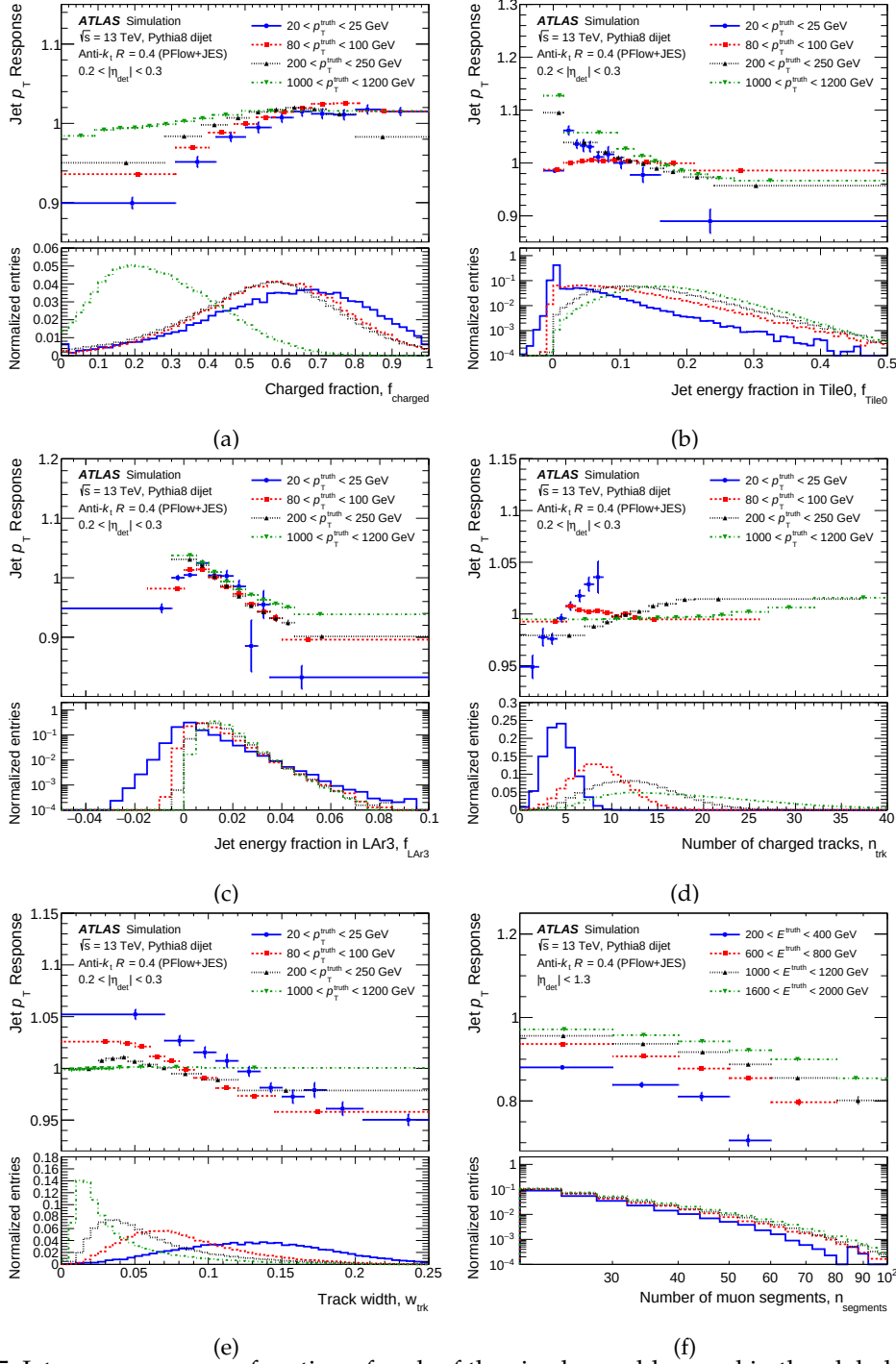


Figure 6.5: Jet p_T response as a function of each of the six observables used in the global sequential calibration (a) fraction of the jet p_T carried by charged particles, (b) fraction of energy in the first layer of the Tile calorimeter, (c) fraction of energy in the third layer of the electromagnetic calorimeter, (d) number of tracks, (e) track width, and (f) number of muon spectrometer track segments associated with the jet. The measurements are presented in four ranges of p_T^{true} except for the number of muon segments distribution that uses E^{true} regions. The error bars represent only the statistical uncertainty. From reference [141].

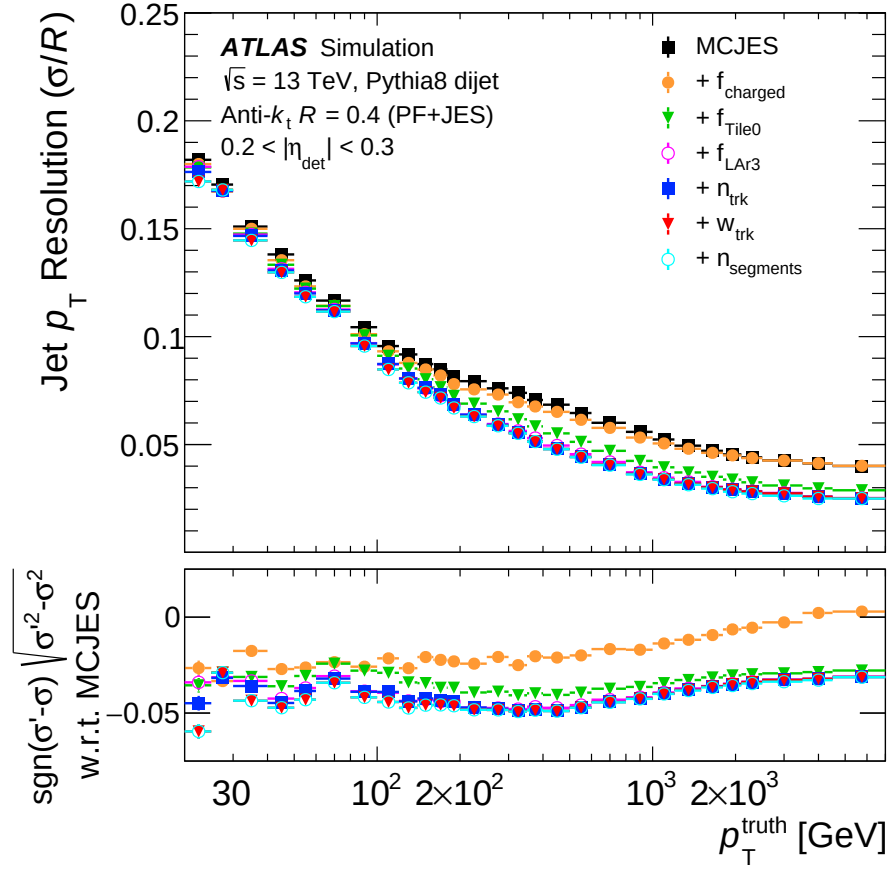


Figure 6.6: Fractional jet resolution after each step of the global sequential calibration. The black squared markers represent the resolution after the absolute calibration with none of the stages of the global sequential calibration. The other markers represent the resolution after each step of the calibration. The lower panel shows the difference in quadrature between the resolution before any stage is applied and after the corresponding step. From reference [141].

of fluctuations in the jet energy reconstruction. It is shown in Figure 6.6 and is reduced after each step of the calibration.

In situ calibration

Once the jets are calibrated, residual differences between data and simulation jet responses persist. These differences arise from imperfect simulation of detector material, hard scattering, underlying events, jet formation and pile-up. The *in situ* calibration addresses these issues by measuring the jet response in both data and simulation and correcting for the differences between data and simulations.

The jet response is computed as the ratio of the jet p_T and the p_T of a well-calibrated object

balancing the jet. This process involves three stages. Firstly, a relative calibration, the η intercalibration, uses central jets to correct the jet energy scale of forward jets. Secondly, the Z+jet and γ +jet analyses use Z bosons and photons as reference objects to correct central jets. Finally, the multijet balance analysis uses a set of well-calibrated low- p_T jets to calibrate a single high- p_T jet in the central regions balanced to them.

The η intercalibration takes jets from the central region of the detector ($|\eta_{\text{det}}| < 0.8$), which are measured with more precision, as well-calibrated objects to correct jets in the forward region ($0.8 \leq |\eta_{\text{det}}| < 4.5$). Events must have exactly two jets in different η regions of the detector. For these events the response ratio is computed as:

$$\mathcal{R} = \frac{2 + \langle \mathcal{A} \rangle}{2 - \langle \mathcal{A} \rangle},$$

where $\langle \mathcal{A} \rangle$ is the mean value of the asymmetry of the two jets, defined as:

$$\mathcal{A} = \frac{p_T^{\text{left}} - p_T^{\text{right}}}{p_T^{\text{avg}}},$$

where p_T^{right} (p_T^{left}) are the transverse momentum for the jet with larger (smaller) η and p_T^{avg} is the average of both p_T . The average response $\langle R_{ijx} \rangle$ is measured for each bin i of η^{left} , j of η^{right} , and x of p_T^{avg} . Additionally, the response ratio defines the calibration factor of the jets as $\mathcal{R} = c^{\text{left}}/c^{\text{right}}$. These correction factors are obtained by the minimisation of a χ^2 and are normalised such that the average of the central regions is unity.

Figure 6.7 shows the extracted relative jet response and the calibration factor as functions of η_{det} and p_T^{jet} in regions of p_T^{avg} and η_{det} , respectively.

The second stage corrects the energy scale using the p_T balance between the jet and the well-calibrated photon or Z boson. It is extracted for jets of the central region, and the η intercalibration ensures it is extrapolated to the forward region. The combination of the Z+jets and the γ +jets measurements provides a wider range in p_T than achieved separately. While Z bosons are measured with high precision at low- p_T from 17 GeV, precision is reduced at moderate and high p_T where statistics are lower. In comparison, γ +jets precision at $p_T < 100$ GeV is limited, but it reaches the high- p_T range with better precision. The *in situ* calibration using γ +jets events is explained in detail in Chapter 8 and briefly summarised here.

This stage uses the missing- E_T projection fraction (MPF) technique, which balances the reference object p_T against the full hadronic recoil in an event. In a Z/ γ +jets event at the particle level, transverse momentum conservation implies:

6 Jet reconstruction, calibration and b -tagging

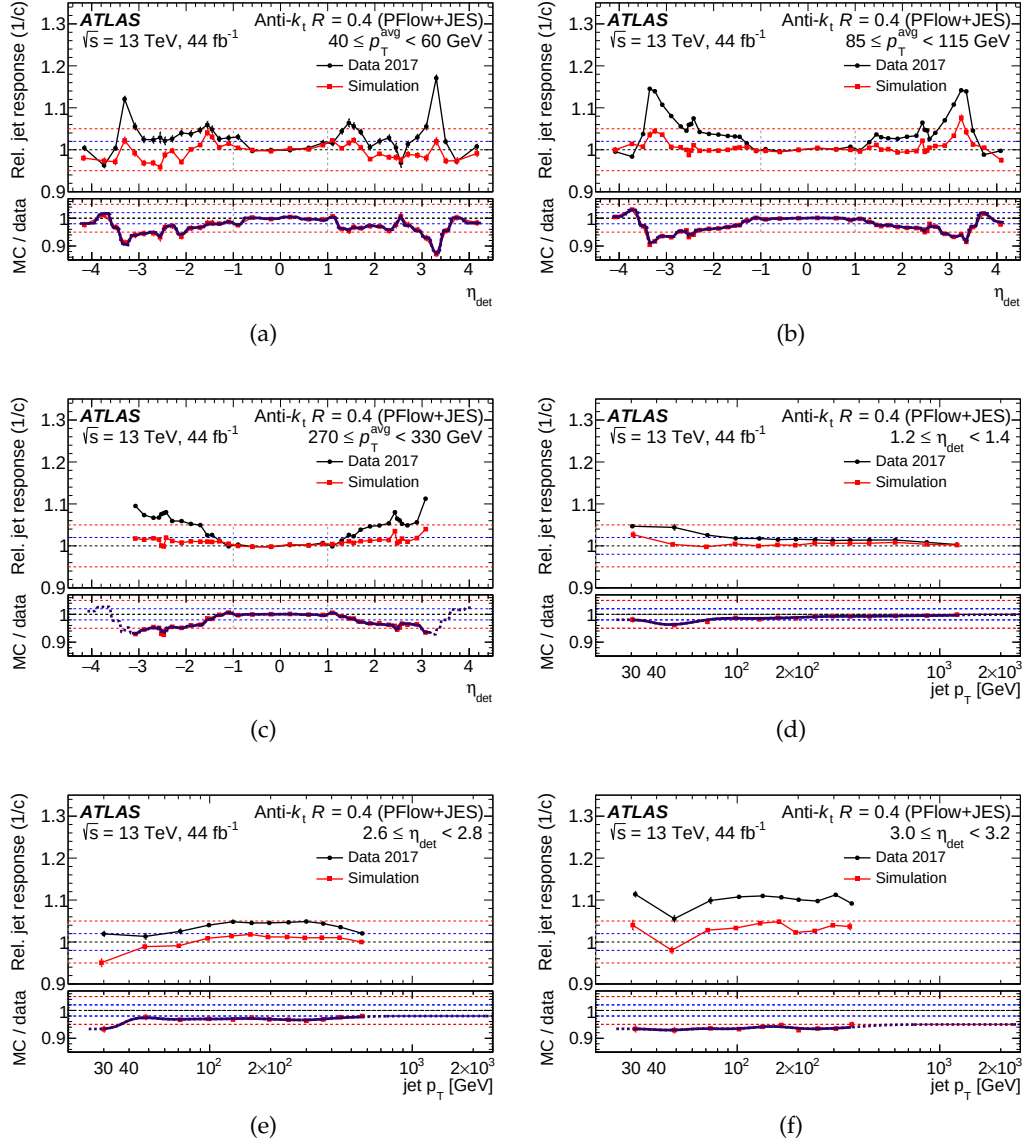


Figure 6.7: Relative jet response for data and simulation as a function of η_{det} for jets of (a) $40 \text{ GeV} \leq p_T^{\text{avg}} < 60 \text{ GeV}$, (b) $85 \text{ GeV} \leq p_T^{\text{avg}} < 115 \text{ GeV}$ and (c) $270 \text{ GeV} \leq p_T^{\text{avg}} < 330 \text{ GeV}$; and as function of p_T for jets with (d) $1.2 \leq \eta_{\text{det}} < 1.4$, (e) $2.6 \leq \eta_{\text{det}} < 2.8$, and (f) $3.0 \leq \eta_{\text{det}} < 3.2$. The lower panel shows the ratio of simulation and data response for the measurements (markers) as well as the interpolation used as calibration factor (solid line) and the extrapolation to regions without measurements (dotted line). From reference [141].

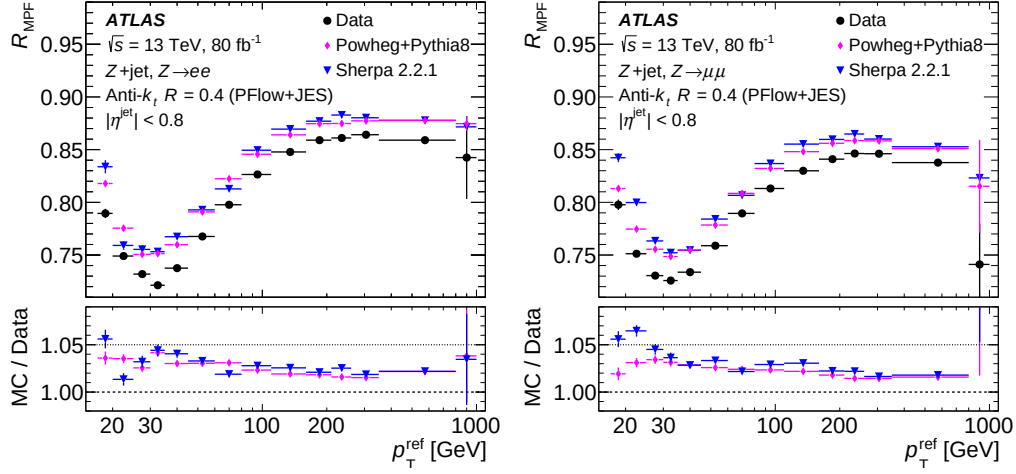


Figure 6.8: MPF jet response as a function of reference p_T using Z+jets events where the Z boson decays into electrons (left) and muons (right). Black markers represent the data response for the 2015-2017 period and are compared with two different simulations. The bottom panel shows the ratio of the simulation and measurement that defines the *in situ* correction to be applied. The error bars represent only the statistical uncertainties. From reference [141].

$$\vec{p}_{T,\text{truth}}^{\text{ref}} + \vec{p}_{T,\text{truth}}^{\text{recoil}} = 0.$$

At the detector level, with a calorimeter response, r_{MPF} , different than unity and assuming the only source of missing energy is this response, the conservation relationship becomes:

$$\vec{p}_T^{\text{ref}} + r_{\text{MPF}} \vec{p}_T^{\text{recoil}} = -\vec{E}_T^{\text{miss}}.$$

After taking the projection of each term in the direction of the reference object, $\hat{n}_{\text{ref}}$, the MPF response, $\mathcal{R}_{\text{MPF}}$ is defined as:

$$\mathcal{R}_{\text{MPF}} = \left\langle 1 + \frac{\hat{n}_{\text{ref}} \cdot \vec{E}_T^{\text{miss}}}{p_T^{\text{ref}}} \right\rangle.$$

The MPF response is shown in Figures 6.8 and 6.9 for Z+jets and γ +jets respectively, comparing measurements and simulations. The bottom panels show the ratio of simulation and data, which is taken as the correction factor. As these measurements are functions of p_T^{ref} , a mapping is applied to get them as a function of p_T^{jet} .

Finally, the last step of the *in situ* calibration uses the multijet balance (MJB) technique to correct high- p_T jets outside the range of the Z/ γ +jet analysis. To compute MJB response,

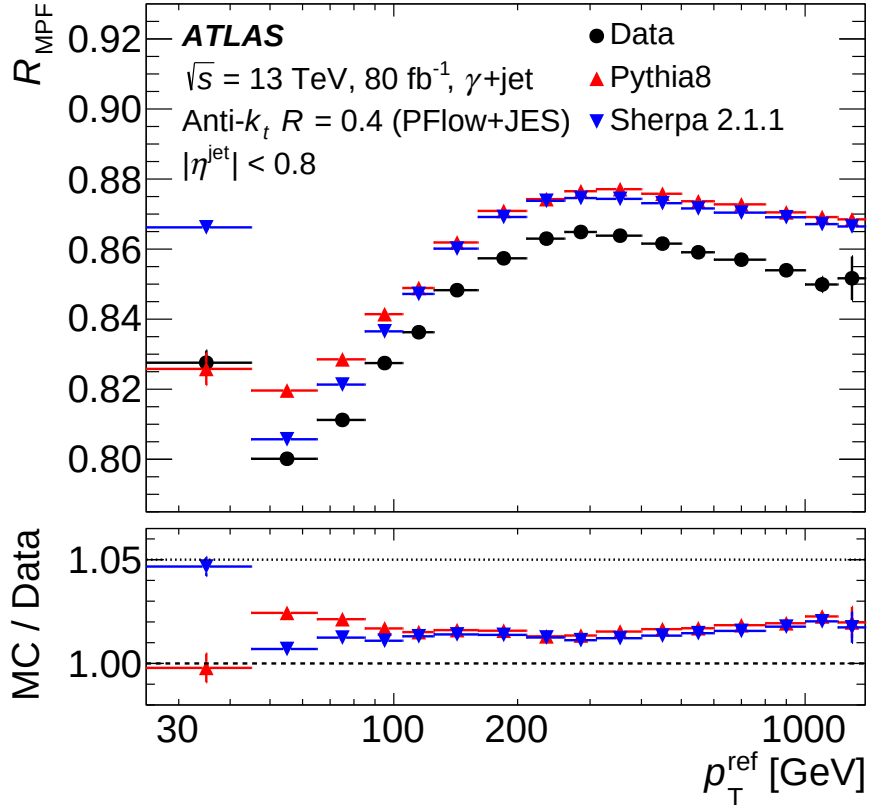


Figure 6.9: MPF jet response as a function of reference p_T using γ +jets events. Black markers represent the data response for the 2015-2017 period and are compared with two different simulations. The bottom panel shows the ratio of the simulation and measurement that defines the *in situ* correction to be applied. The error bars represent only the statistical uncertainties. From reference [141].

events are required to have a single high- p_T jet balanced against a system of well-calibrated jets by the Z/γ +jet correction. The response is defined as:

$$\mathcal{R}_{\text{MJB}} = \left\langle \frac{p_T^{\text{lead}}}{p_T^{\text{ref}}} \right\rangle,$$

where p_T^{ref} is taken as the vector sum of all jets in the recoil system. This process is done twice, with the first iteration only selecting events with subleading jets with $p_T < 1.2$ TeV, and the second raising the threshold to 1.8 TeV. This approach allows the calibration of events with a subleading jet not calibrated by the Z/γ +jet correction.

Figure 6.10 shows the MJB response as a function of p_T^{ref} . The measurement is compared with four simulations, and in the bottom panel, the ratio that defines the correction is shown. This ratio is then mapped from p_T^{ref} to p_T^{jet} to be correctly applied to jets.

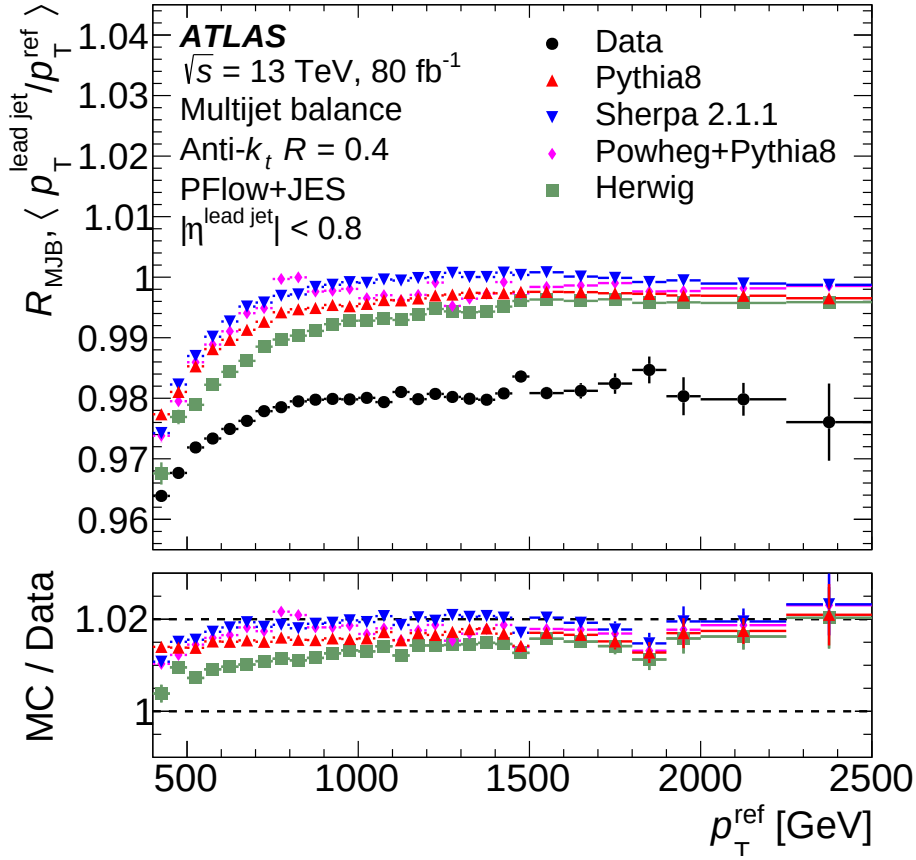


Figure 6.10: MJB response as a function of reference p_T . Black markers represent the data response for the 2015-2017 period and are compared with four different simulation,. The bottom panel shows the ratio of the simulation and measurement that defines the *in situ* correction to be applied. The error bars represent only the statistical uncertainties. From reference [141].

With the results from the Z/γ +jet analyses and the MJB, a combination is performed to produce a single calibration covering the full range of p_T^{jet} . This combination parameterises each of the four corrections ($Z(\rightarrow ee)$ +jets, $Z(\rightarrow \mu\mu)$ +jets, γ +jets and MJB). After they are parameterised in bins of 1 GeV using second-order polynomial splines, a χ^2 minimisation is performed at each bin. Figure 6.11 shows the result of the combination for the *in situ* correction as a function of p_T^{jet} .

6.2.2 Jet energy resolution

The jet energy resolution, which quantifies the precision with which the energy of a jet is measured, can be parameterised by three different components: noise (N), stochastic (S) and

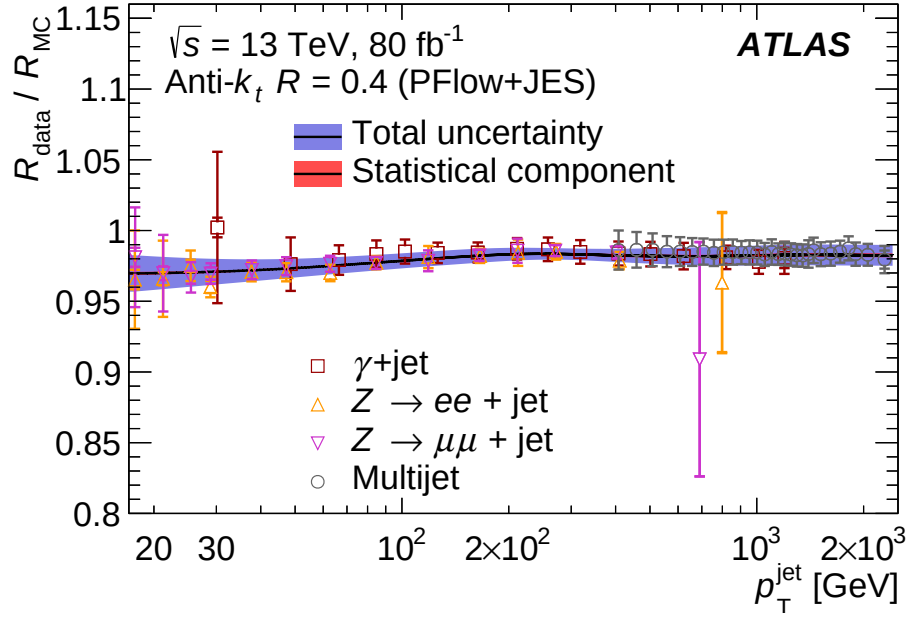


Figure 6.11: Jet response ratio of data and simulation as a function of p_T^{jet} . Measurements for $Z(\rightarrow ee)+\text{jets}$, $Z(\rightarrow \mu\mu)+\text{jets}$, $\gamma+\text{jets}$ and MJB are represented by markers, while the combined result is shown as a solid line. The inner horizontal ticks on the error bars represent statistical uncertainties while the outer horizontal ticks include the systematic uncertainties. Statistical and systematic uncertainties of the correction are depicted as bands. From reference [141].

constant (C). These terms contribute to the relative jet energy resolution as expressed by the following equation:

$$\frac{\sigma(p_T)}{p_T} = \frac{N}{p_T} \oplus \frac{S}{\sqrt{p_T}} \oplus C,$$

where $\oplus$ denotes the sum in quadrature. The noise term is caused by pile-up and electronic noise from the front-end electronics, dominating at low p_T . The stochastic term corresponds to statistical fluctuations in the deposited energy, and the constant arises from energy depositions in passive material, non-uniformities of response across the calorimeter, and the starting point of hadron showers.

The jet energy resolution is computed by measuring the balance in dijet systems. Additionally, an independent estimation of the noise term is performed. Both measurements are combined and then applied to reconstructed jets.

Resolution measurement using dijet events

The measurement of the jet energy resolution is performed using dijet events. After calibrating the jets, deviations from exact balance are only caused by experimental resolution. To measure this difference, the asymmetry ($\mathcal{A}$) is computed as:

$$\mathcal{A} = \frac{p_T^{\text{probe}} - p_T^{\text{ref}}}{p_T^{\text{avg}}},$$

where p_T^{ref} is the transverse momentum of a jet selected in the $0.2 \leq |\eta_{\text{det}}^{\text{ref}}| \leq 0.7$ region, considered well calibrated, excluding a seam around $\eta_{\text{det}} = 0$ to measure the energy as clean as possible; p_T^{probe} is the transverse momentum of a jet from $|\eta_{\text{det}}^{\text{ref}}| \leq 4.5$ and p_T^{avg} is their average. The standard deviation of $\mathcal{A}$ ($\sigma_{\mathcal{A}}$) is measured as a function of p_T^{avg} and $\eta_{\text{det}}^{\text{probe}}$ and may be expressed as:

$$\sigma_{\mathcal{A}} = \frac{\sigma_{p_T}^{\text{probe}} \oplus \sigma_{p_T}^{\text{ref}}}{\langle p_T^{\text{avg}} \rangle},$$

where $\sigma_{p_T}^{\text{probe}}$ and $\sigma_{p_T}^{\text{ref}}$ are the jet energy responses for the respective jets. For calibrated jets, the average transverse momenta must be equal, allowing the energy resolution of the probe jet to be expressed as:

$$\left\langle \frac{\sigma_{p_T}}{p_T} \right\rangle^{\text{probe}} = \sigma_{\mathcal{A}} \ominus \left\langle \frac{\sigma_{p_T}}{p_T} \right\rangle^{\text{ref}}.$$

The relative jet energy resolution for the reference jet is obtained from events where both jets fall in the central reference region, simplifying the expression to:

$$\left\langle \frac{\sigma_{p_T}}{p_T} \right\rangle^{\text{ref}} = \frac{\sigma_{\mathcal{A}}}{\sqrt{2}}.$$

Additionally, a correction to $\sigma_{\mathcal{A}}$ is applied to remove effects not coming from the detector, such as additional radiation and non-perturbative processes, expressed as:

$$\sigma_{\mathcal{A}}^{\text{det}} = \sigma_{\mathcal{A}}^{\text{meas}} \ominus \sigma_{\mathcal{A}}^{\text{truth}}.$$

The measurement of the relative jet energy resolution as a function of p_T^{jet} in the central region of $\eta_{\text{det}}^{\text{probe}}$ is shown in Figure 6.12.

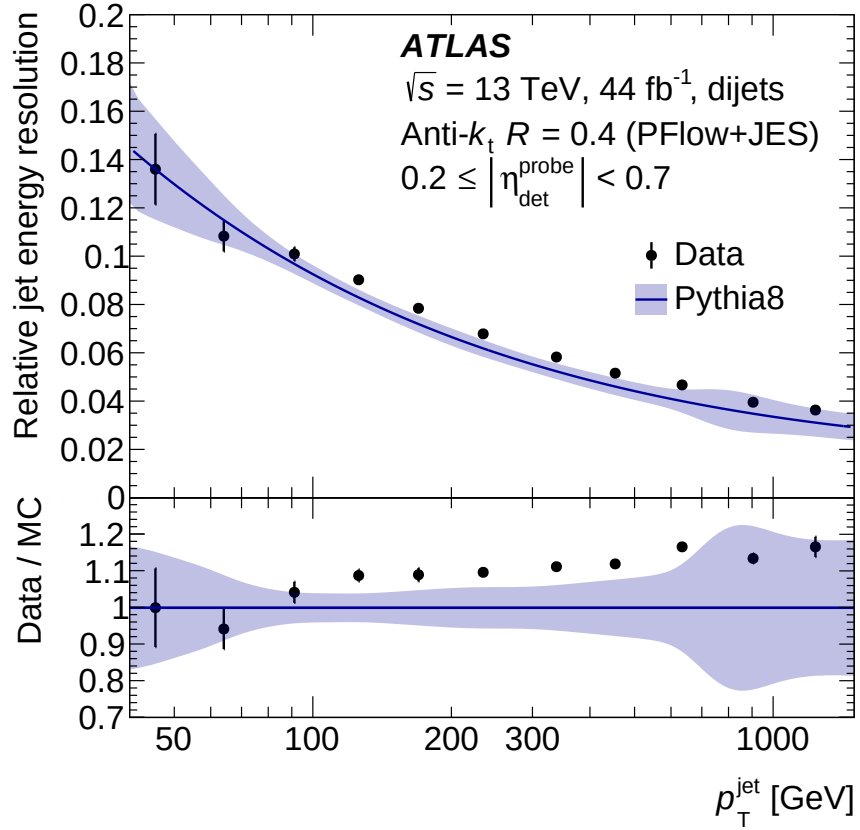


Figure 6.12: Relative jet energy resolution as a function of p_T^{jet} in the central region of the detector measured using dijet events. The data resolution is shown as black markers and is compared with the simulated resolution displayed as a solid line. The error bars represent the statistical uncertainty of the measurement, and the shaded band represents the systematic uncertainty of the simulation. The bottom panel shows the ratio of the data and simulation. From reference [141].

Noise measurement using random cones

A direct estimation of the noise term is obtained separately for the contributions caused by pile-up and electronic noise.

For the pile-up contribution, the random cones method is used. In this method, a couple of cone regions of the $\eta - \phi$ plane are selected and the energy deposits are added. The cones are selected randomly in ϕ and in opposite regions of η . From the transverse momentum measured in each cone, the quantity Δp_T^{RC} is measured as:

$$\Delta p_T^{\text{RC}} = p_T^{c1} - p_T^{c2},$$

where p_T^{c1} and p_T^{c2} are the transverse momenta measured in each cone. Measuring the central 68% confidence interval of the distribution of Δp_T^{RC} (σ_{RC}) as a function of pseudorapidity and the expected number of interactions per bunch crossing, the pile-up noise term (N^{PU}) can be computed as:

$$N^{\text{PU}} = \frac{\sigma_{\text{RC}}}{2\sqrt{2}}.$$

To estimate this electronic contribution, a simulation sample with $\mu = 0$ is used to measure the jet energy resolution. Finally, the total noise term is obtained as a sum in quadrature of both contributions and shown as a function of $|\eta|$ in Figure 6.13.

Combination and application of jet energy resolution

A combination of the dijet measurement constrained with the noise term obtained independently is performed. This result as a function of p_T^{jet} is shown in Figure 6.14.

The combined resolution is applied to simulation to ensure that the jet energy resolution for simulation and data match. To achieve this, in regions of jet p_T where the resolution in data is larger than in simulation, the simulation should be smeared until its average resolution matches that of data. In other regions, no smearing is performed since data should remain unaltered.

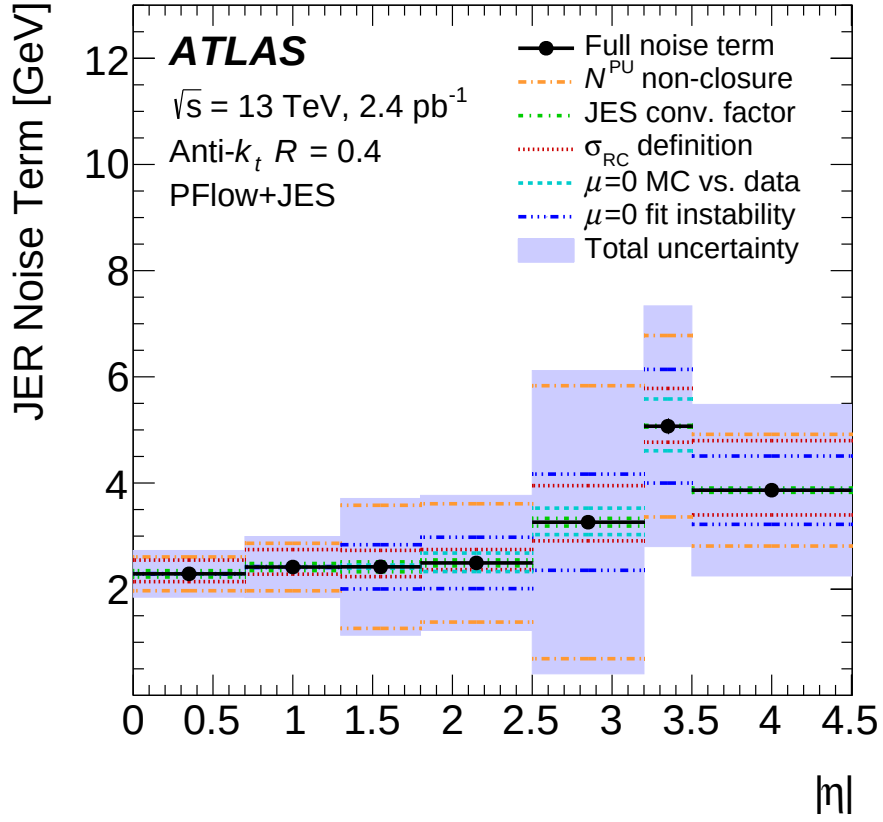


Figure 6.13: Jet energy resolution noise term as a function of $|\eta|$. The error band represents the total uncertainty of the measurement. From reference [141].

6.3 Pile-up jet suppression from tracking

Even after calibration, where pile-up subtraction takes place, some pile-up remains. It is caused by local fluctuations in the pile-up activity. To achieve further cleaning of the pile-up, information from the tracks matched to each jet is used. A discriminant called jet vertex tagger (JVT) is employed in this process.

The JVT discriminant is constructed as a two-dimensional likelihood based on a k -nearest neighbour algorithm. The variables used are R_{pT} and corrJVF , defined as:

$$R_{pT} = \frac{\sum_k p_{T,k}^{\text{track}}(\text{PV}_0)}{p_T^{\text{jet}}},$$

where the scalar sum of the p_T runs over the tracks associated with the jets originating in the hard-scatter vertex, and the corrected jet vertex fraction is defined as:

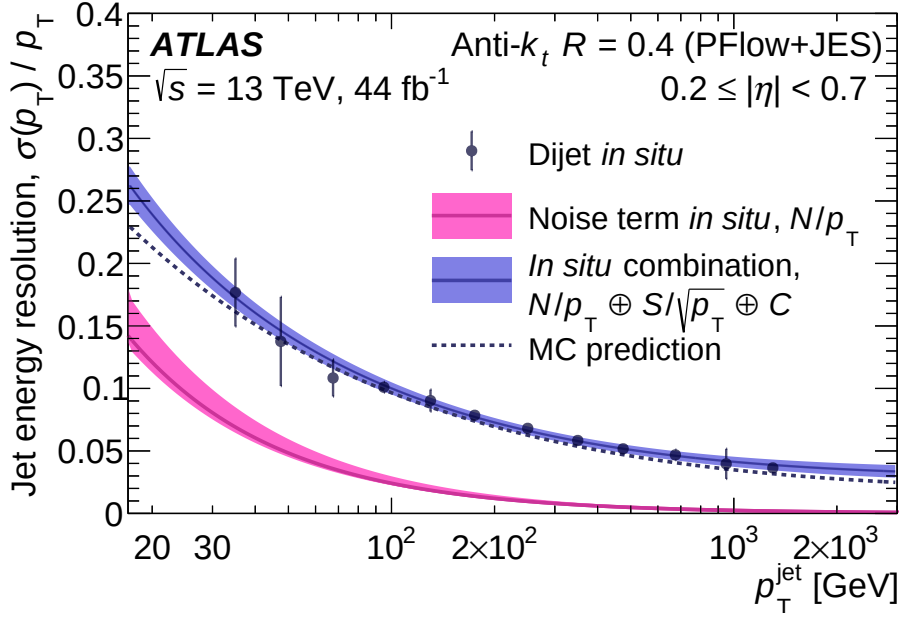


Figure 6.14: Combined jet energy resolution and noise term as a function of p_T^{jet} . The error band represent the total uncertainty of the measurement. From reference [141].

$$\text{corrJVF} = \frac{\sum_m p_{T,m}^{\text{track}}(\text{PV}_0)}{\sum_l p_{T,l}^{\text{track}}(\text{PV}_0) + \frac{\sum_{n \geq 1} \sum_l p_{T,l}^{\text{track}}(\text{PV}_n)}{k \cdot n_{\text{track}}^{\text{PU}}}},$$

where $\sum_{n \geq 1} \sum_l p_{T,l}^{\text{track}}(\text{PV}_n)$ is the scalar sum of p_T of the tracks associated with the jet originating from a primary vertex different from the hard-scatter vertex, $n_{\text{track}}^{\text{PU}}$ is the total number of pile-up tracks per event and $k = 0.01$. The probability defining the JVT discriminant for each point of the corrJVF- R_{pT} plane is computed as the ratio of the number of hard-scatter jets to the number of hard-scatter plus pile-up jets in a local neighbourhood. This neighbourhood is defined as the 100 nearest neighbours produced by simulated samples of signal and pile-up. Jets selected from a dijet sample with $20 < p_T < 50$ GeV and $|\eta| < 2.4$ are used for the training of the k-nearest neighbour algorithm.

In Figure 6.15, the efficiency of the JVT cut is measured using $Z(\rightarrow \mu\mu)$ +jets events for three different JVT working points: $\text{JVT} > 0.2$, $\text{JVT} > 0.4$ and $\text{JVT} > 0.7$, as functions of p_T^{ref} defined as the transverse momentum of the Z boson.

6 Jet reconstruction, calibration and b -tagging

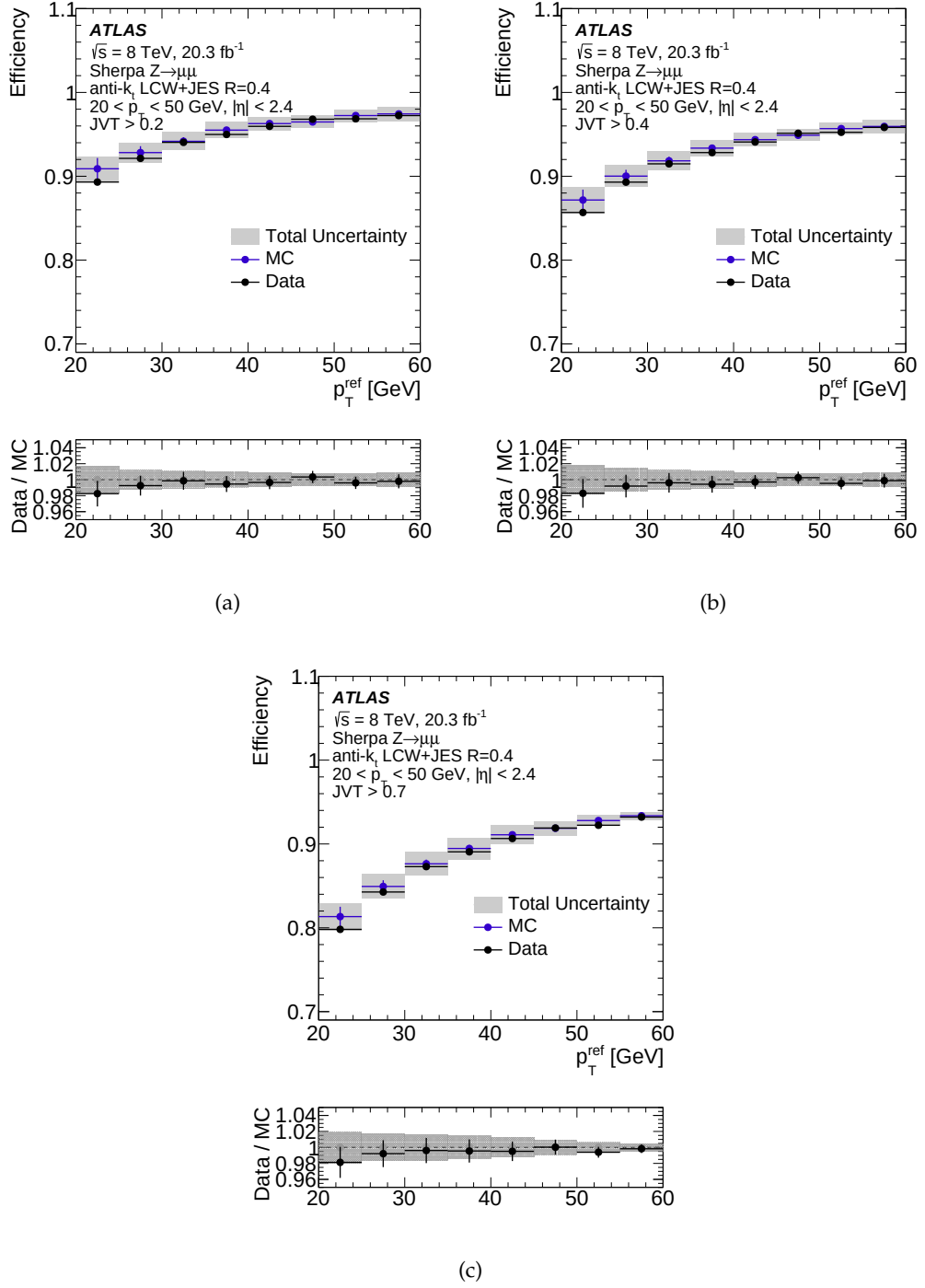


Figure 6.15: Efficiency as a function of p_T^{ref} for (a) $\text{JVT} > 0.2$, (b) $\text{JVT} > 0.4$, and (c) $\text{JVT} > 0.7$ measured using $Z (\rightarrow \mu\mu) + \text{jets}$ events. The simulation is compared to the measured efficiency and the ratio is shown in the lower panels. The total uncertainty of the simulation is shown as a grey band. From reference [142].

6.4 Jet flavour-tagging

The selection of jets containing b -, c -hadrons and jets originating only from light quarks is an important selection criterion in some analyses. This process, referred to as flavour tagging, is based on information such as charged-particle tracks, primary vertices, and jets. Additionally, information about hadrons contained within jets is taken into account in the simulations. This information is used to design low-level algorithms to identify heavy quarks, which are later used as input for the high-level flavour tagger.

The low-level algorithms can be split into two categories. The first one includes algorithms based on impact parameters and the second uses information from reconstructing displaced vertices. Two algorithms based on impact parameters are defined: IP2D and IP3D . The first one uses only the signed transverse impact parameter significance, defined as the ratio of the the impact parameter and its uncertainty, whereas the second also uses the signed longitudinal impact parameter significance. These methods define three log-likelihood ratios for each algorithm: one separating b -jets and light-flavour jets, a second differentiating b -jets and c -jets and a last one discriminating c -jets and light-flavour jets. These log-likelihood ratios are defined as:

$$\text{LLR} = \sum_{i=1}^N \ln (p_b (S_i) / p_{\text{light}} (S_i))$$

for the b -jets against light-flavour jets discriminant and analogously for the other two. Here, the sum runs over the tracks associated with the jets and $p_b (S_i)$ ($p_{\text{light}} (S_i)$) are the probability density functions for the b -jet (light-flavour jet) hypothesis as a function of the track signed impact parameter significance. These algorithms assume the different track contributions to be independent of each other, and the addition of more input variables becomes difficult. To overcome this situation, a recurrent neural network algorithm is used, taking as inputs the signed transverse and longitudinal impact parameter significances, the fraction of transverse momentum carried by the track relative to the jet p_T , the angular distance between the track and the jet axis, and the hit multiplicity of the track in the different ID layers of each track. The tracks are ordered by the absolute value of the transverse impact parameter significances, and the outputs from the recurrent neural network are the probabilities of being a b -jet, c -jet, or light-flavoured jet. These probabilities are combined in a discriminant defined as:

$$D_{\text{RNNIP}} = \ln \left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) p_{\text{light}}} \right),$$

where p_b , p_c , and p_{light} are the three flavour probabilities, and f_c is the c -jet fraction set as 0.07.

6 Jet reconstruction, calibration and b -tagging

Two more low-level flavour taggers are used based on the displaced vertex approach. Firstly, the `SV1` algorithm that reconstructs a displaced secondary vertex. This algorithm identifies possible two-track vertices reconstructed with the tracks associated with the jet and rejects those compatible with K_S^0 or Λ decays, hadronic interactions with the detector, or photon conversions. Then it runs over each possible two-track vertex to fit its position. In case the invariant mass of the vertex is larger than 6 GeV, a χ^2 is computed for every track to test the track-vertex association, and the track with a larger χ^2 is removed from the fit. This process is repeated until the invariant mass of the vertex is under 6 GeV, and 8 discriminant variables are used as input for the high-level algorithm. Secondly, a topological multi-vertex finding algorithm, called `JetFitter`, takes advantage of the topological structure of the decays of b - and c -hadrons and tries to reconstruct the complete decay chain. This approach tries to find the flight path of b -hadrons and vertex positions even for vertices with only one track. Once the decay chains are traced, 17 variables are computed and used as input for the high-level flavour tagger.

To maximise the performance of these flavour taggers, the `DL1` algorithm series is designed based on fully connected multi-layer feed-forward neural networks. The first version of the series, called `DL1`, uses as input the outputs from the `IP2D`, `IP3D`, `SV1`, and `JetFitter` algorithms. Additionally, another variation, called `DL1r`, was designed including the `RNNIP` output. In Table 6.1, the discriminant variables used as input are listed. These neural networks have as output the probabilities for a jet to be a b -jet, c -jet, or a light-flavour jet. Finally, a discriminant for the b -tagging is defined as:

$$D_{\text{DL1/DL1r}} = \ln \left(\frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\text{light}}} \right),$$

where p_b , p_c and p_{light} are the outputs of the neural network, and f_c is set to 0.018 to optimise the performance. Analogously, a discriminant for c -tagging is defined as:

$$D_{\text{DL1/DL1r}}^c = \ln \left(\frac{p_c}{f_b \cdot p_b + (1 - f_b) \cdot p_{\text{light}}} \right),$$

where f_b is set to 0.2. The working points for these algorithms are defined according to the efficiency of b -jets passing a fixed cut in the discriminant.

The performance of b -tagging is characterised by rejection factors for light-flavoured and c -jets. These rejection factors are computed as the inverse of the efficiency for each kind of background jet. Figure 6.16 presents the rejection factors for the `DL1r` algorithm in comparison with the `RNNIP`, `SVKine`, and `JFKine` algorithms as functions of the efficiency of b -tagging. The `SVKine` and `JFKine` taggers are shown in this figure solely to illustrate the performance of the `SV1` and `JetFitter` algorithms. The `DL1r` tagger performs better than all other low-level algorithms. Figure 6.17 provides a similar comparison for the `DL1`

Table 6.1: Discriminating variables used as input in the high-level flavour tagger algorithms.

Algorithm	Variable	Description
IP2D	p_T	Jet p_T .
	$ \eta $	Jet $ \eta $.
	$\ln(P_b/P_{\text{light}})$	Likelihood ratio of the b -jet to light-flavour jet hypotheses.
	$\ln(P_b/P_c)$	Likelihood ratio of the b -jet to c -jet hypotheses.
RNNIP	$\ln(P_c/P_{\text{light}})$	Likelihood ratio of the c -jet to light-flavour jet hypotheses.
	P_b	b -jet probability.
	P_c	c -jet probability.
SV1	P_{light}	Light-flavour jet probability.
	$m(2^{\text{nd}})$ (SV)	Invariant mass of tracks at the secondary vertex assuming pion mass.
	$f_E(2^{\text{nd}})$ (SV)	Jet energy fraction of the tracks associated with the secondary vertex.
	$N_{\text{TrkAtVtx}}(2^{\text{nd}})$ (SV)	Number of tracks used in the secondary vertex.
	$N_{2\text{TrkVtx}}(\text{SV})$	Number of two-track vertex candidates.
	$L_{xy}(2^{\text{nd}})$ (SV)	Transverse distance between the primary and secondary vertices.
	$L_{xyz}(2^{\text{nd}})$ (SV)	Distance between the primary and secondary vertices.
	$S_{xyz}(2^{\text{nd}})$ (SV)	Distance between the primary and secondary vertices divided by its uncertainty.
	$\Delta R(\vec{p}_{\text{jet}}, \vec{p}_{\text{vtx}})$ (SV)	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex.
JetFitter	$m(2^{\text{nd}})$ (JF)	Invariant mass of tracks at the secondary vertex assuming pion mass.
	$f_E(2^{\text{nd}})$ (JF)	Jet energy fraction of the tracks associated with the secondary vertex.
	$N_{\text{TrkAtVtx}}(2^{\text{nd}})$ (JF)	Number of tracks used in the secondary vertex.
	$N_{2\text{TrkVtx}}(\text{JF})$	Number of two-track vertex candidates.
	$L_{xy}(2^{\text{nd}})$ (JF)	Transverse distance between the primary and secondary vertices.
	$L_{xyz}(2^{\text{nd}})$ (JF)	Distance between the primary and secondary vertices.
	$S_{xyz}(2^{\text{nd}})$ (JF)	Distance between the primary and secondary vertices divided by its uncertainty.
	$\Delta R(\vec{p}_{\text{jet}}, \vec{p}_{\text{vtx}})$ (JF)	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex.
	$N_{1\text{-trkvertices}}$	Number of single-prong displaced vertices.
	$N_{\geq 2\text{-trkvertices}}$	Number of multi-prong displaced vertices.
	m_{Trk}	Invariant mass of tracks from displaced vertices.
	$E(2^{\text{nd}})$	Energy of the tracks associated with the secondary vertex.
	f_E	Jet energy fraction of the tracks associated with the displaced vertices.
	N_{TrkAtVtx}	Number of tracks from multi-prong displaced vertices.
	$\eta_{\text{trk}}^{\text{min, max, avg}}(2^{\text{nd}})$	Min., max. and avg. pseudorapidity of tracks at the secondary vertex.

6 Jet reconstruction, calibration and b -tagging

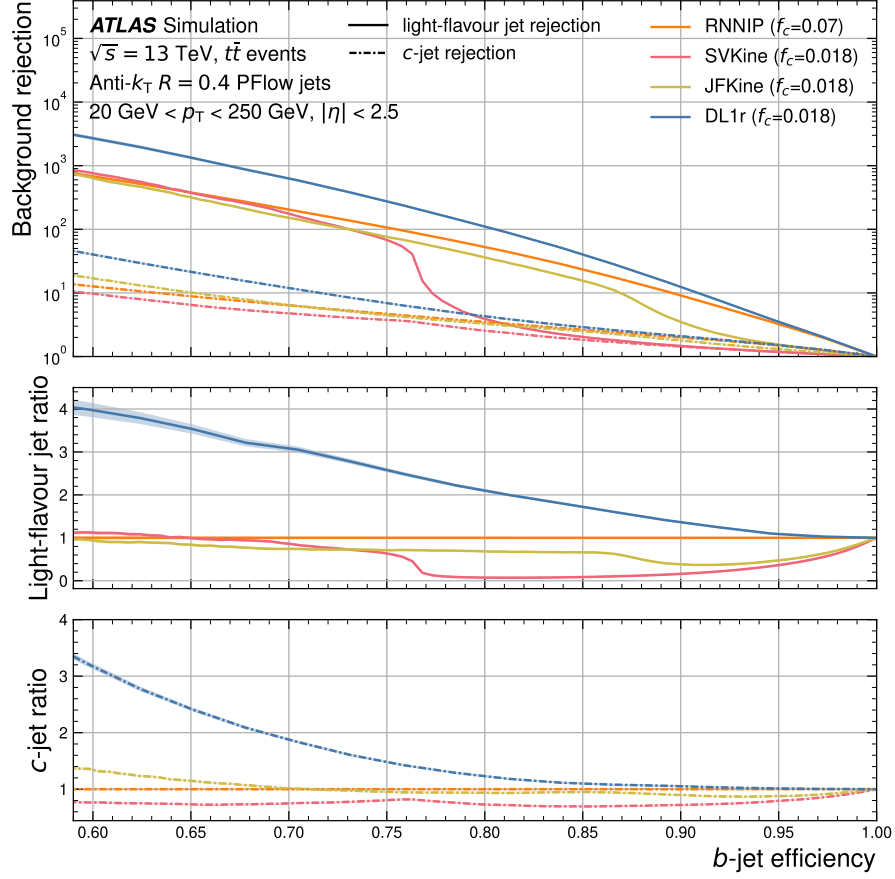


Figure 6.16: The rejection factor for light-flavour jet and c -jet as functions of the efficiency of the DL1r algorithm, compared with RNNIP, SVKine and JFKine algorithms. The lower panels show the ratio of rejection factors of light-flavour and c -jet for these algorithms over the RNNIP rejection factor. The coloured bands represent statistical uncertainties. From reference [143].

and MV2c10 taggers. The MV2c10 algorithm is a previous ATLAS tagger [144] based on boosted decision trees. Both the DL1 and DL1r taggers exhibit a larger rejection factor than MV2c10 for all efficiencies, demonstrating that the DL1 series has a superior performance.

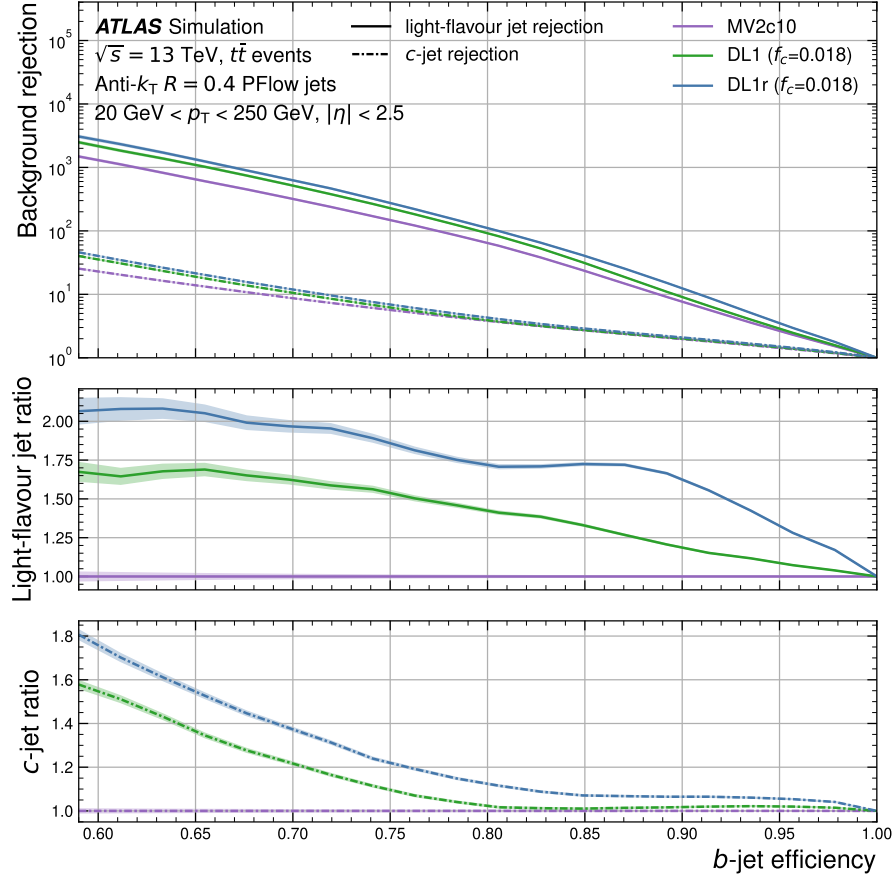


Figure 6.17: The rejection factor for light-flavour jet and c -jet as functions of the efficiency of the `DL1r` algorithm, compared with `DL1` and `MV2c10` algorithms. The lower panels show the ratio of the rejection factors of light-flavour and c -jet for these algorithms over the `MV2c10` rejection factor. The coloured bands represent the statistical uncertainties. From reference [143].

7 Missing transverse momentum reconstruction and performance

In this chapter, a description of the method used to reconstruct the missing transverse momentum in ATLAS is presented [145, 146].

7.1 Missing transverse momentum reconstruction

The missing transverse momentum in the event, whose magnitude is denoted as E_T^{miss} , is defined as the negative vector sum of the p_T of the reconstructed and calibrated objects in the event.

The reconstruction of E_T^{miss} is performed using the transverse momenta of the particles present in an event, including muons, electrons, photons, hadronically decaying τ -leptons, and jets. This is calculated following the expression:

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss},\mu} + E_{x(y)}^{\text{miss},e} + E_{x(y)}^{\text{miss},\gamma} + E_{x(y)}^{\text{miss},\tau} + E_{x(y)}^{\text{miss,jets}} + E_{x(y)}^{\text{miss,soft}},$$

where each term is computed as the negative vectorial sum of transverse momenta of every candidate of each kind. The first five terms are commonly referred to as the E_T^{miss} hard term, while the last one, called soft term, accounts for transverse momenta of energy deposits or tracks of charged particles that are not associated with any candidate.

Additionally, another important magnitude related to E_T^{miss} , the total transverse energy in the detector, $\sum E_T$, is computed. This quantity measures the total event activity and is related to the resolution of E_T^{miss} . It is defined as the scalar sum of the transverse momenta of reconstructed objects and the soft term that contributes to the reconstruction of E_T^{miss} :

$$\sum E_T = \sum p_T^\mu + \sum p_T^e + \sum p_T^\gamma + \sum p_T^\tau + \sum p_T^{\text{jets}} + \sum p_T^{\text{soft}}.$$

To minimise double counting in the E_T^{miss} hard term, different objects are treated sequentially. Firstly, muons and electrons are selected following the same criteria used for the selection in the analysis. Photons and hadronically decaying τ -leptons are either selected following the selection in the analysis or their corresponding detector signals are treated as if arising

7 Missing transverse momentum reconstruction and performance

from hadronic jets. Finally, the jet selection is applied analogously to the selection in the analysis.

To resolve the ambiguity between particles reconstructed with calorimeter deposit information such as electrons, photons, hadronically decaying τ -leptons, and jets, a hierarchy is set. Electrons are the higher-priority objects followed by photons. Hadronically decaying τ -leptons come later, and jets are the lower-priority objects. Finally, the energy deposits associated with jets are accounted for in either the hard or soft term. This ordering is based on the variables defined in the following. Firstly, the ratio of electron, photon, or hadronically decaying τ -lepton energy and the jet energy is computed:

$$f_{\text{overlap}} = \frac{E_i^{\text{EM}}}{E_{\text{jet}}^{\text{EM}}},$$

where i represents the electron, photon or hadronically decaying τ -lepton; the energies are calibrated at the EM scale. The second variable is the difference in transverse momentum at the EM scale of the jet and object i :

$$\Delta p_T^{\text{EM},i,\text{jet}} = p_T^{\text{EM},\text{jet}} - p_T^{\text{EM},i}.$$

Jets with $f_{\text{overlap}} < 1.0$ and $\Delta p_T^{\text{EM},i,\text{jet}} > 20$ GeV are treated in the jet term and, otherwise, contribute to the soft term.

Additionally, calorimeter showers from muons may be reconstructed as jets and another ambiguity resolution is performed. To remove these jets, the following four conditions must be met:

- the ID track of the muon is associated with the jet calorimeter deposits.
- $p_T^{\mu-\text{ID}} / p_T^{\text{jet-ID}} > 0.8$. The p_T of the ID track associated with the muon represents a significant fraction of the transverse momentum of the sum over all ID tracks associated with the jet.
- $p_T^{\text{jet}} / p_T^{\mu-\text{ID}} < 2$. The overall transverse momentum of the jet is not too large compared to the transverse momentum of the muon ID track.
- $N_{\text{track}}^{\text{PV}} < 5$. The total number of tracks associated with the jet and emerging from the hard-scatter PV is small.

Similarly, muons may produce FSR in the form of photons that are not reconstructed as photons or electrons. As the transverse momentum of these photons reconstructed as jets is not recovered in muon reconstruction, jets falling in the following selection are included in the hard term of E_T^{miss} :

7.2 Missing transverse momentum performance

- the ID track of the muon is associated with the jet calorimeter deposits.
- $N_{\text{track}}^{\text{PV}} < 3$. The total number of tracks associated with the jet and emerging from the hard-scatter PV is very small.
- $f_{\text{EM}} < 0.9$, where $f_{\text{EM}} = E_{\text{jet}}^{\text{EM}} / E_{\text{jet}}$. The jet energy is largely deposited in the electromagnetic calorimeter.
- $p_{\text{T}}^{\text{jetPS}} > 2.5$ GeV. The transverse momentum contribution of the jet in the presampler calorimeter is large.
- $w_{\text{jet}} < 0.1$, where $w_{\text{jet}} = \sum_i \Delta R_i p_{\text{T}}^i / \sum_i p_{\text{T}}^i$ and ΔR_i is the angular distance of topocluster i from the jet axis. The jet is narrow with a width comparable to a dense electromagnetic shower.
- $p_{\text{T}}^{\mu-\text{ID}} / p_{\text{T}}^{\text{jet-ID}} > 0.8$. The p_{T} of the ID track associated with the muon represents a significant fraction of the transverse momentum of the sum over all ID tracks associated with the jet.

The soft term is composed of all detector tracks and calorimeter energy deposits not matched to the reconstructed objects defined previously and may contain contributions from the hard scatter as well as the underlying event and pile-up interactions. To reconstruct this term, the particle flow reconstruction is used. Tracks associated to charged particle flow constituents not associated to any object are selected if they have $p_{\text{T}} > 0.5$ GeV, are in the region of $|\eta| < 2.5$, and have a longitudinal impact parameter $|\Delta z_0 \sin \theta| < 2.0$ mm.

7.2 Missing transverse momentum performance

The missing transverse momentum performance is characterised by a high level of complexity, due to the composite character of the observable. The resolution for $E_{\text{T}}^{\text{miss}}$ is a convolution of the contribution from multiple objects. To assess the quality of the $E_{\text{T}}^{\text{miss}}$ reconstruction its scale and resolution are measured and compared to simulations using $Z \rightarrow \mu\mu$ events, where no genuine $E_{\text{T}}^{\text{miss}}$ is expected.

The projection of the $\mathbf{E}_{\text{T}}^{\text{miss}}$ vector in the direction of the $\mathbf{p}_{\text{T}}^Z$ is used to measure the $E_{\text{T}}^{\text{miss}}$ scale. This observable is sensitive to the contribution from the hadronic recoil in the opposite direction to the Z boson. It is measured as a function of the transverse momentum of the Z boson as an indicator of the hardness of the interaction.

Figure 7.1 shows the average projection of $\mathbf{E}_{\text{T}}^{\text{miss}}$ vector onto the direction of the Z boson transverse momentum vector using $Z \rightarrow \mu\mu$ events. The residual offsets for data and simulations agree within the uncertainties.

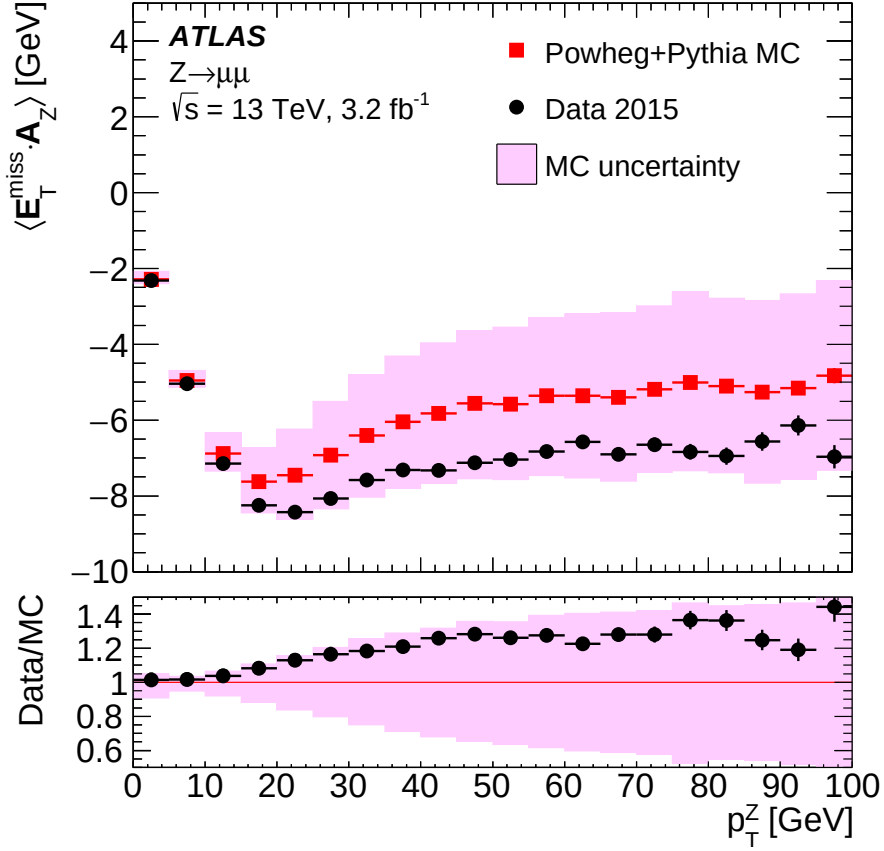


Figure 7.1: The average projection of E_T^{miss} onto the direction $\mathbf{A}_Z$ of the Z boson transverse momentum vector as a function of p_T^Z measured using $Z \rightarrow \mu\mu$ events. The simulation is compared to data and the ratio is shown in the lower panel. The total uncertainty is represented as a shaded area. From reference [145].

The E_T^{miss} resolution is estimated by the width of the combined $E_{x,y}^{\text{miss}}$ distributions. The width is computed in terms of the RMS.

Figure 7.2 shows the RMS of the $E_{x,y}^{\text{miss}}$ resolution as functions of $\sum E_T$ and the number of primary vertices. The $\sum E_T$ distribution shows two different behaviours. For the $50 \text{ GeV} < \sum E_T < 200 \text{ GeV}$ range, where the two muons are the dominant hard objects contributing to the missing transverse momentum, the resolution quickly rises from 5 GeV to about 20 GeV. Beyond this range, as the transverse momentum of the individual jets rises and the number of contributing jets slowly increases, the RMS is driven by the resolution from jets and the slope decreases. For the distribution as function of the number of primary vertices, data shows a less steep slope than the simulation. However, in both cases data and simulation agree within the uncertainties.

7.2 Missing transverse momentum performance

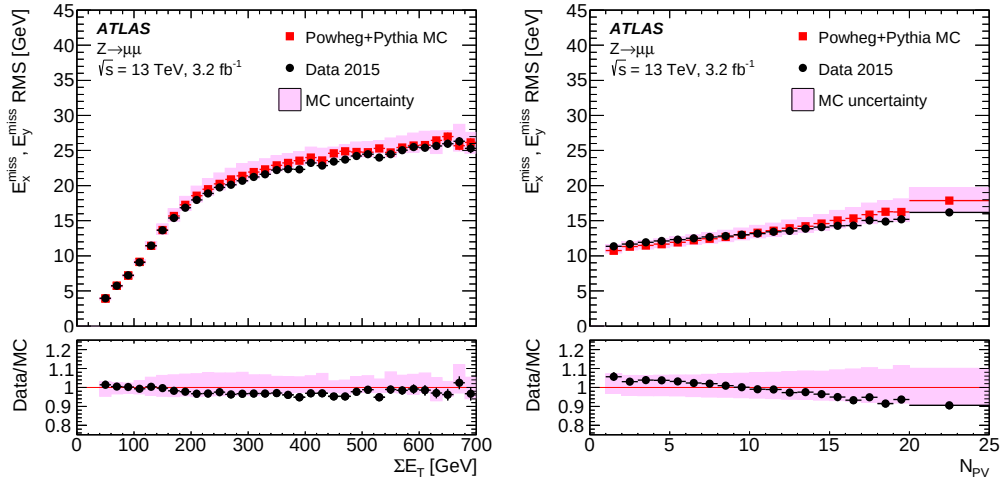


Figure 7.2: The RMS width of the $E_{x,y}^{\text{miss}}$ distributions as functions of ΣE_T (left) and the number of primary vertices measured using $Z \rightarrow \mu\mu$ events. The simulation is compared to data and the ratio is shown in the lower panel. The total uncertainty is represented as a shaded area. From reference [145].

8 *In situ* jet energy scale calibration using γ +jets events

The energetic pp collisions produced by the LHC yield final states predominantly characterised by jets, which are collimated sprays of charged and neutral hadrons along with their decay products. Jets are crucial for the physics programme carried out using the ATLAS detector, and a precise understanding of jet reconstruction is vital for a wide range of processes. Measurements of both the jet energy scale (JES) and resolution (JER) of these complex objects are therefore essential for precision measurements of the SM and for searches for new phenomena. Improved measurements of the jet energy scale calibration and an evaluation of their performance in simulation and data pave the way to achieve better precision on the JES for Run 3 and beyond. These techniques were developed for jets reconstructed with the anti- k_t algorithm [101, 147] with a radius parameter $R = 0.4$ using particle flow inputs [140, 141]. Previous calibration strategies by the ATLAS Collaboration, using up to 81 fb^{-1} data [141, 148–152], are extended and improved by taking advantage of the full Run 2 data sample of 140 fb^{-1} .

The jet energy scale calibration consists of a series of steps. The first stage of the calibration uses simulation to derive corrections to the jet energy scale, reducing the impact of pile-up, detector effects, and other parameters. The second stage of the calibration is an *in situ* calibration, correcting for remaining differences between data and MC simulation, derived using well-measured reference objects, including photons and Z bosons.

This chapter presents the measurements of the *in situ* jet energy scale calibration using γ +jets events, as detailed in Ref. [153]. The structure of this chapter is as follows. Section 8.1 summarises the data and simulation samples used to measure the *in situ* jet energy scale, and Section 8.2 describes the selection criteria employed in the analysis. Section 8.3 briefly describes the *in situ* calibration method in the jet calibration process. Section 8.4 presents the MPF method used to measure the jet response, Section 8.5 details the fitting algorithm used to estimate the MPF response, and Section 8.6 describes the correction method applied to the measured MPF response to reduce the effects from the multijet background. Section 8.7 details the uncertainties taken into account in the analysis, and the results are presented in Section 8.8.

8.1 Data and simulated samples

The data used for the calibration study were collected by ATLAS in pp collisions at $\sqrt{s} = 13$ TeV from 2015 to 2018 with all subdetectors operational, corresponding to an integrated luminosity of 140 fb^{-1} . The proton bunch crossing interval was 25 ns during this period of data taking. The average number of interactions per bunch crossing, μ , for the Run 2 data is 34. These conditions, along with the detector configuration and reconstruction thresholds, are taken into account in producing and reconstructing simulated events [154].

The *in situ* calibration uses simulated MC γ +jets events. PYTHIA 8.230 [155] is used as the nominal MC sample. Events were generated using tree-level matrix elements for photon-plus-jet final states as well as LO QCD dijet events, with the inclusion of initial- and final-state parton showers. The fragmentation component was modelled by final-state QED radiation arising from calculations of all $2 \rightarrow 2$ QCD processes. The NNPDF2.3LO [156] PDF set was used in the matrix-element calculation, the parton shower, and the simulation of the multi-parton interactions. The samples include a simulation of the underlying event with parameters set according to the A14 tune [157]. The Lund string model [158, 159] was used for the description of the fragmentation into hadrons.

An alternative MC sample generated with SHERPA 2.2 [160] was also used. For this sample, NLO-accurate matrix elements for up to two partons, and LO-accurate matrix elements for up to four partons were calculated with the Comix [161] and OPENLOOPS [162–164] libraries. The default SHERPA parton shower [165] based on Catani–Seymour dipole factorisation and the cluster hadronisation model [166] were used. The samples used the dedicated set of tuned parameters developed by the SHERPA authors for this generator version and the NNPDF3.0NNLO PDF set [167]. The NLO matrix elements for a given jet multiplicity were matched to the parton shower using a colour-exact variant of the MC@NLO algorithm [168]. Different jet multiplicities were then merged into an inclusive sample using an improved CKKW matching procedure [169, 170] which was extended to NLO accuracy using the MEPS@NLO prescription [171]. The merging cut was set dynamically at a scale of 20 GeV according to the prescription in Ref. [172].

For the SHERPA samples, the renormalisation and factorisation scales for the photon-plus-jet core process were set to the transverse energy of the photon, E_T^γ . The strong coupling constant was set to $\alpha_s(m_Z) = 0.118$ and the QED coupling constant was evaluated in the Thomson limit. Photons from the matrix elements were required to be within the rapidity range $|y_\gamma| < 2.7$, and isolated according to a smooth-cone isolation criterion [173] with $\delta_0 = 0.1$, $\epsilon_\gamma = 0.1$, and $n = 2$. The effects of QCD scale uncertainties were evaluated [174] using seven-point variations of the factorisation and renormalisation scales in the matrix elements. The scales were varied independently by factors of 0.5 and 2, avoiding variations in opposite directions. The PDF uncertainties for the nominal PDF set were evaluated using the 100 variation replicas, as well as ± 0.001 shifts of α_s . Additionally, the results were

Table 8.1: List of single-photon triggers used in the γ +jet *in situ* calibration. Each trigger is shown with its corresponding E_T^γ range.

E_T^γ range [GeV]	HLT trigger
17 – 20	HLT_g10_loose
20 – 25	HLT_g15_loose
25 – 30	HLT_g20_loose
30 – 40	HLT_g25_loose
40 – 45	HLT_g35_loose
45 – 50	HLT_g40_loose
50 – 55	HLT_g45_loose
55 – 65	HLT_g50_loose
65 – 75	HLT_g60_loose
75 – 85	HLT_g70_loose
85 – 105	HLT_g80_loose
105 – 125	HLT_g100_loose
125 – 145	HLT_g120_loose
145 – inf	HLT_g140_loose

cross-checked using the central values of the CT14NNLO [175] and MMHT2014NNLO [176] PDF sets.

All samples are reconstructed using a full detector simulation and minimum-bias interactions simulated using PYTHIA 8 with the A3 set of tuned parameters [177] and NNPDF2.3LO PDF were superimposed set to represent multiple pp interactions during the same or nearby bunch crossings. The distribution of the average number of pile-up interactions in simulation is reweighted during data analysis to match that observed in the Run 2 data.

8.2 Data and simulation selection

The event sample used in this analysis is selected using a collection of single-photon triggers. The lowest unprescaled trigger for single-photon events during the complete Run 2 is the HLT_g140_loose, which selects events compatible with having a *Loose* photon with $p_T^\gamma > 140$ GeV and reaches an efficiency near 100% in the $E_T^\gamma > 145$ GeV range. To measure the energy scale for jets below this threshold, it is necessary to use prescaled triggers and weight the events selected according to the trigger efficiency. The trigger menu used for event selection is shown in Table 8.1 with the corresponding E_T^γ range for each individual trigger.

8 *In situ* jet energy scale calibration using γ +jets events

Events must satisfy common selection requirements in the *in situ* analysis. Each event is required to have at least one reconstructed primary vertex with at least two matched tracks of $p_T > 500$ MeV. Jets arising from cosmic rays, non-collision background and calorimeter noise are vetoed by applying data-quality requirements [178]. In addition, jets with $20 < p_T < 60$ GeV and $|\eta_{\text{det}}| < 2.4$ are required to satisfy the criteria of JVT [142, 179]. The JVT criteria reject jets from pile-up interactions by matching jets with the primary vertex; it has a selection efficiency of 97% for hard scatter jets at the nominal operating point.

Photon candidates entering the analysis are required to have $E_T^\gamma > 25$ GeV and $|\eta^\gamma| < 1.37$ and to satisfy the *Tight* identification and isolation selection criteria [132]. The jet is removed if it falls within $\Delta R = 0.4$ (0.35) of a photon (lepton).

Events must contain a jet with p_T greater than 10 GeV that falls within $|\eta| = 0.8$. To suppress effects from additional parton radiation, further requirements are imposed on the azimuthal angle between the photon and the leading jet $\Delta\phi^{\text{ref,jet}} > 2.9$ and p_T of the subleading jet $p_T < \max(0.3 \times p_T^{\text{ref}}, 12)$ GeV, where the subleading jet falls within $|\eta| = 4.5$.

8.3 *In situ* calibration

The *in situ* correction, as the final calibration step, accounts for differences in the jet response between simulation and data. Such differences arise due to the imperfect simulation of detector response and the amount of detector material, and the modelling of physics processes involved: hard scattering, underlying event, pile-up, jet formation and particle interactions with the detector material. The *in situ* calibration also provides important validation of the MC calibration of jets by comparing the data-to-MC difference between the p_T balance of a jet recoiling from a well-calibrated object or system.

The *in situ* calibration response, $\mathcal{R}_{in\ situ}$, is defined as the average ratio of the jet p_T to the transverse momentum of the reference object p_T^{ref} , derived as a function of p_T^{ref} . The $\mathcal{R}_{in\ situ}$ response is susceptible to effects such as the radiation of additional partons or the loss of energy outside the reconstructed jet cone. Dedicated event selections are applied to mitigate these effects. A double ratio, insensitive to these secondary effects provided they are well-modelled in simulations, is defined as

$$\mathcal{C} = \frac{\mathcal{R}_{in\ situ}^{\text{data}}}{\mathcal{R}_{in\ situ}^{\text{MC}}}.$$

The calibration factor applied to the jet four-momentum in data is obtained by a numerical inversion of this double ratio as a function of jet p_T , and as a function of η_{det} in η intercalibration (see below).

Two stages of *in situ* analyses are done sequentially to assess the performance of MC calibrations. First, a relative *in situ* calibration referred to as the η intercalibration is done, which corrects the energy scale of forward jets ($0.8 < |\eta_{\text{det}}| < 4.5$) to match that of the central jets ($|\eta_{\text{det}}| < 0.8$) using the p_T balance in a dijet system. Second, an absolute calibration is done by measuring the p_T balance of a central jet against a well-calibrated Z boson or a photon. The missing- E_T projection fraction (MPF) method [180] is used in Z/γ +jet events to calculate the p_T balance between the full hadronic recoil and a Z boson or a photon. The method is less susceptible to effects of pile-up and the threshold of the jet reconstruction than the direct balance method, allowing a reliable measurement of the low- p_T jet response below 100 GeV.

For each *in situ* analysis, the main sources of systematic uncertainties arise from the simulation model of physics processes, the measurement of the reference object, and the p_T balance due to the selected event topology. Uncertainties related to the simulation model of physics effects are addressed by taking the difference between the predictions of two distinct MC event generators. The difference between jet response in simulations depends on hadronisation models that lead to different jet composition [181]. Uncertainties in the reference object are estimated by propagating the $\pm 1\sigma$ calibration uncertainties through the analysis. Uncertainties due to the selected event topology are evaluated by varying the event selection criteria and comparing the impact on the response ratios between data and MC simulation. To reduce the statistical fluctuations when applying the systematic variations, a rebinning procedure similar to that used in previous publications [141] is employed to obtain statistically significant results using pseudo-experiments. This rebinning procedure is only performed in regions where no sharp variations in p_T response are observed to ensure that no real physics effects are removed.

8.4 The MPF method

The MPF method measures the p_T balance between the reference objects and the full hadronic recoil in Z/γ +jet events. This technique allows for the direct computation of the calorimeter response to hadronic showers. It is minimally affected by pile-up and underlying event, which are uniform across the detector and thus cancelled out in the MPF method. According to the conservation of transverse momentum, the transverse momentum of all the hadronic activity in a Z/γ +jet event, p_T^{recoil} , should be equal and opposite to the transverse momentum of the reference boson, p_T^{ref} , at the particle level, such that

$$\vec{p}_{T,\text{truth}}^{\text{ref}} + \vec{p}_{T,\text{truth}}^{\text{recoil}} = 0. \quad (8.1)$$

At the detector level, well-calibrated objects have a response of unity while the calorimeter response to the hadronic recoil, r_{MPF} , is lower than unity, resulting in possible missing

8 *In situ* jet energy scale calibration using γ +jets events

transverse momentum $\vec{E}_T^{\text{miss}}$ in the event. Therefore, Equation 8.1 can be written as:

$$\vec{p}_T^{\text{ref}} + r_{\text{MPF}} \cdot \vec{p}_T^{\text{recoil}} = -\vec{E}_T^{\text{miss}}.$$

Projecting the vector terms along the direction of the reference boson using a unit vector $\hat{n}_{\text{ref}}$ in the transverse plane, r_{MPF} depends only on the missing transverse momentum and the transverse momentum of the reference boson. The average of r_{MPF} , $\mathcal{R}_{\text{MPF}}$, is measured as a function of p_T^{ref} , the p_T of the reference Z/γ boson:

$$\mathcal{R}_{\text{MPF}} = \left\langle 1 + \frac{\hat{n}_{\text{ref}} \cdot \vec{E}_T^{\text{miss}}}{p_T^{\text{ref}}} \right\rangle,$$

where $\vec{E}_T^{\text{miss}}$ is computed using particle-flow objects calibrated at the EM scale.

8.5 Fitting procedure

The MPF response is computed in bins of p_T^{ref} . In each bin, the distribution of r_{MPF} is fitted to a Gaussian distribution. The fit is performed in a window centred on the mean of the complete distribution, whose width is between 2.4σ and 4σ , where σ represents the RMS of the complete distribution. This limitation in the fitting range aims to reduce the contribution from mismodelling in the tails of the distribution. The mean value of $\mathcal{R}_{\text{MPF}}$ is taken as the mean of the Gaussian distribution obtained in the fit.

The results of the fits of the MPF response in different bins of p_T^{ref} for the data sample are shown in Figure 8.1. Figures 8.2 and 8.3 present the results of the fits for the simulated samples generated with PYTHIA and SHERPA, respectively.

8.6 Signal purity corrections

Even after requiring photons to be *Tight* and isolated, a remaining amount of background events is still present in the sample, as shown in Ref. [182]. Figure 8.4 shows how the distribution of *Tight* photon candidates is well described by a fitted linear combination of the signal simulation and the non-*Tight* photon candidate distributions. This background comes from jets misidentified as photons and introduces a variation in the MPF response. A correction in the response is needed to take into account the contribution of the background. For this purpose, a data-driven method is used.

Four regions are defined in the $\gamma_{\text{ID}} - E_T^{\text{iso}}$ plane:

8.6 Signal purity corrections

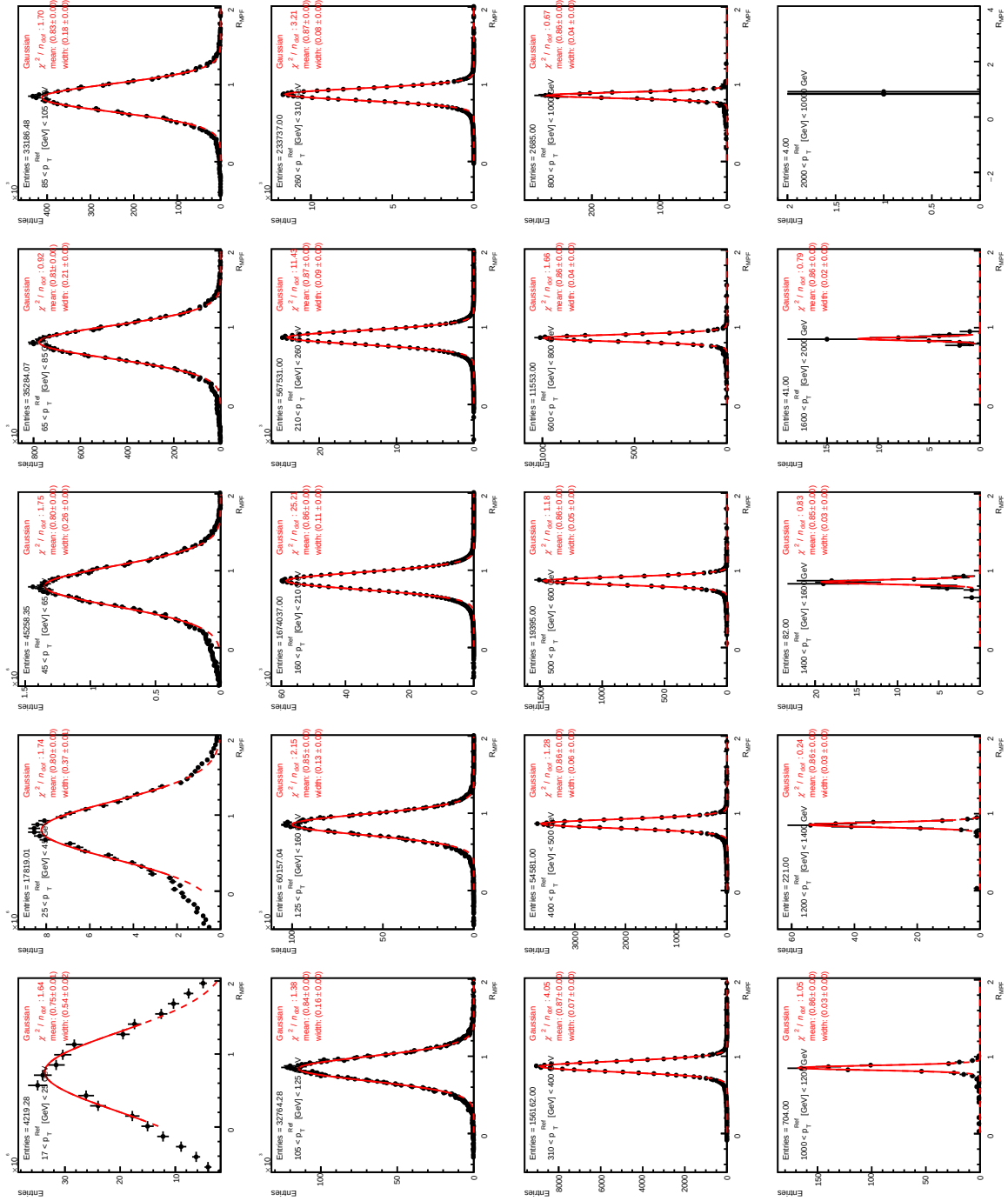


Figure 8.1: Distributions of r_{MPF} in different bins of $p_{\text{T}}^{\text{ref}}$ for 2015-2018 data. The data distributions are represented as dots, with the Gaussian fitted distribution displayed as a solid line in the fitting range and dashed outside.

8 In situ jet energy scale calibration using γ +jets events

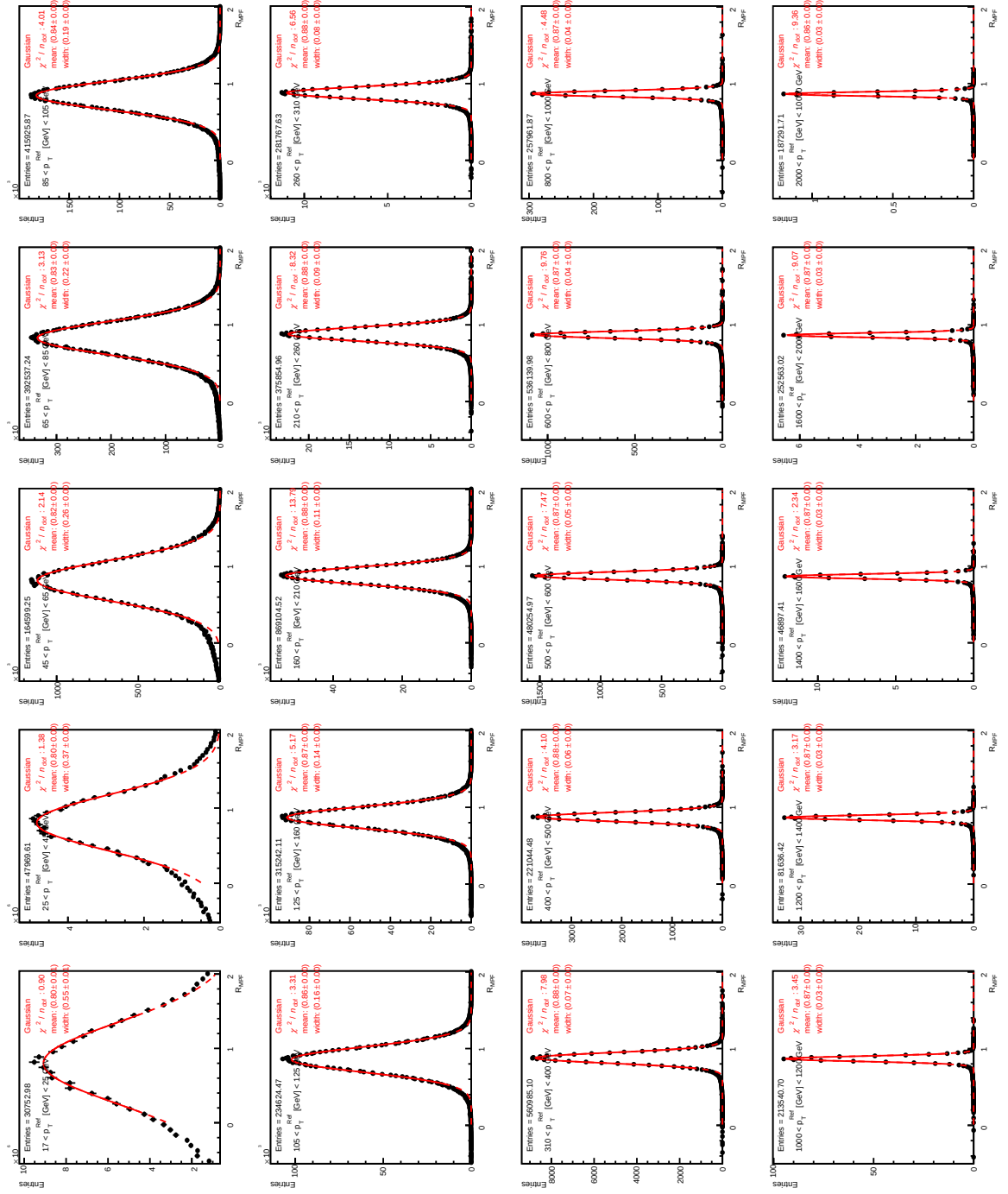


Figure 8.2: Distributions of r_{MPF} in different bins of $p_{\text{T}}^{\text{ref}}$ for the PYTHIA simulation. The simulated distributions are represented as dots, with the Gaussian fitted distribution displayed as a solid line in the fitting range and dashed outside.

8.6 Signal purity corrections

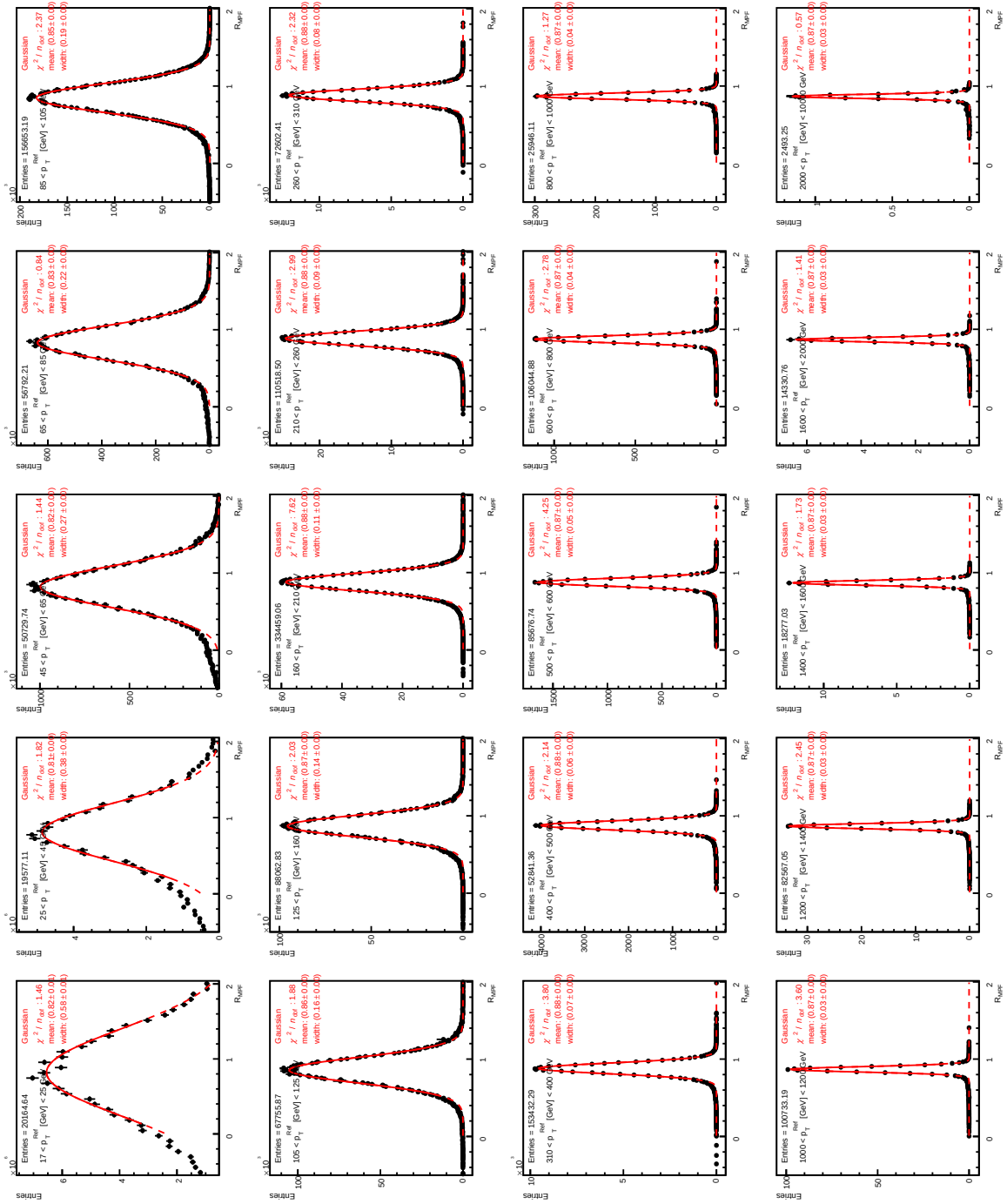


Figure 8.3: Distributions of r_{MPF} in different bins of p_T^{ref} for the SHERPA simulation. The simulated distributions are represented as dots, with the Gaussian fitted distribution displayed as a solid line in the fitting range and dashed outside.

8 In situ jet energy scale calibration using γ +jets events

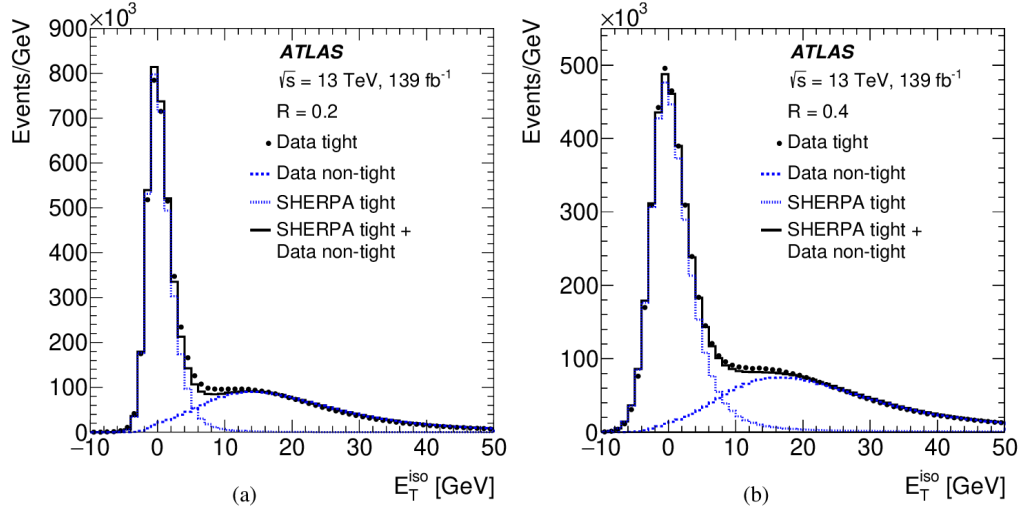


Figure 8.4: The E_T^{iso} distributions for *Tight* photon candidates (dots) for (a) $R = 0.2$ and (b) $R = 0.4$ isolation cone. The data distribution is shown together with the signal simulation (dotted line). Additionally, the result of fitting the distribution (solid line) with a linear combination of the signal simulation and the non-*Tight* photon candidates (the distribution of the non-*Tight* contribution is shown separately as a dashed line). From reference [182].

- Region A: the signal region. It contains *Tight* and isolated photon candidates.
- Region B: the control region with *Tight* and non-isolated photon candidates.
- Region C: the control region with non-*Tight* and isolated photon candidates.
- Region D: the control region with non-*Tight* and non-isolated photon candidates.

For the non-*Tight* photon candidate sample, photons failing the *Tight* identification criteria are required to pass the *LoosePrime4* identification, which adds to the *Loose* identification the requirements on R_ϕ , w_{stot} , and f_1 from the *Tight* working point. Analogously, for the non-isolated photon candidates, the *Loose* isolation WP is required.

The correction applied to the response is computed assuming the measured response, i. e., the response in Region A, as a linear combination of the signal response and the response in the control region that describes best the background in Region A. This is given by:

$$\mathcal{R}_A = P \cdot \mathcal{R}^{\text{corr}} + (1 - P) \cdot \mathcal{R}_C,$$

where P is the signal purity in Region A, the fraction of signal events present in this region, $\mathcal{R}^{\text{corr}}$ is the corrected response, and $\mathcal{R}_R$ with $R = A, C$ is the response measured in each region. Figure 8.5 shows the response for regions A and C together with their ratio. The response in the background region is around 2% higher than in the signal region at low

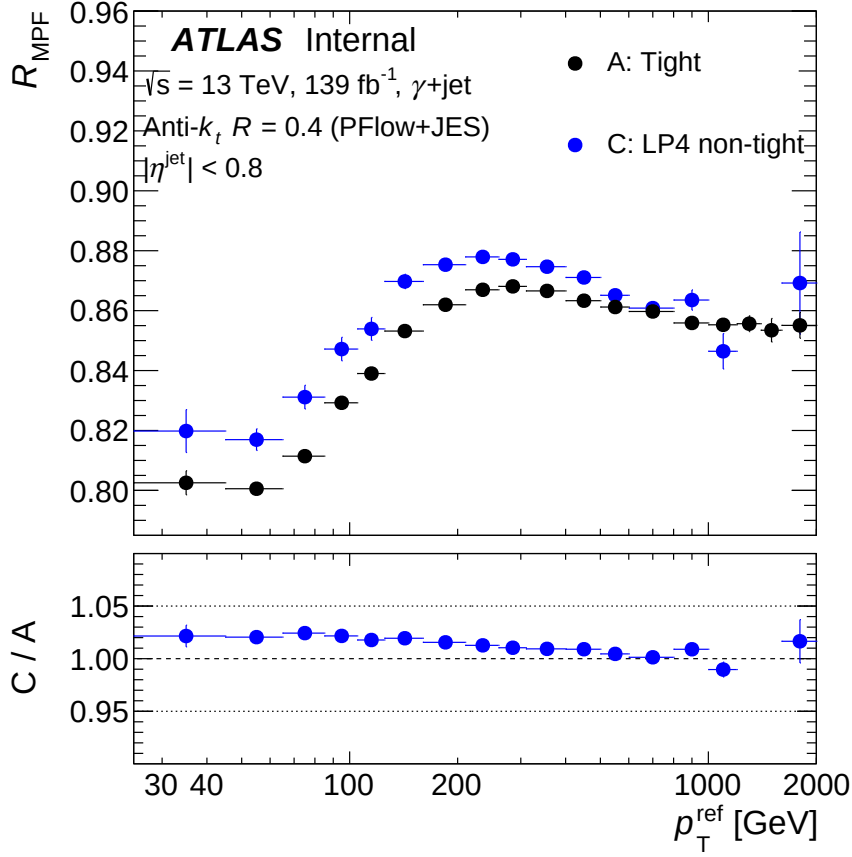


Figure 8.5: MPF response as a function of $p_{\text{T}}^{\text{ref}}$ for γ +jets events for both the signal region (isolated and *Tight* photon candidates) and the background region with isolated and non-*Tight* photon candidates. The bottom plot shows the ratio of the MPF response in both regions. Error bars represent the statistical uncertainty in the fit.

$p_{\text{T}}^{\text{ref}}$ and becomes more compatible as $p_{\text{T}}^{\text{ref}}$ increases. The reduced number of events in the high- $p_{\text{T}}^{\text{ref}}$ range makes the statistical uncertainty increase. In Appendix A, the plots of the fits for Region C are shown.

To compute the signal purity needed for the response correction, the same data-driven method employed in previous publications [183–185] is used. This method uses a counting technique based on observed events in the three control regions previously defined. The signal purity is defined as:

$$P = \frac{N_{\text{A}}^{\text{sig}}}{N_{\text{A}}},$$

8 *In situ* jet energy scale calibration using γ +jets events

where N_A is the total measured number of events in Region A and N_A^{sig} is the number of signal events in Region A. To compute the number of signal events, the number of events for the background is considered to follow:

$$\frac{N_A^{\text{bg}}}{N_C^{\text{bg}}} = R^{\text{bg}} \frac{N_B^{\text{bg}}}{N_D^{\text{bg}}},$$

where N_R^{bg} with $R = A, B, C, D$ is the number of background events in each region and R^{bg} represents the correlation factor between γ_{ID} and E_T^{iso} in background events. This correlation factor characterises the deviation from the uncorrelated case for which it is set to be unity. With this relationship, the number of signal events in Region A is given by:

$$N_A^{\text{sig}} = N_A - N_A^{\text{bg}} = N_A - R^{\text{bg}} \cdot \frac{N_B^{\text{bg}} \cdot N_C^{\text{bg}}}{N_D^{\text{bg}}}.$$

At the same time, N_R^{bg} with $R = B, C, D$ is computed as:

$$N_R^{\text{bg}} = N_R - N_R^{\text{sig}} = N_R - \varepsilon_R \cdot N_A^{\text{sig}},$$

where ε_R is the signal leakage fractions on Region R. The signal leakage fractions are computed as $\varepsilon_R = N_R^{\text{sig,MC}} / N_A^{\text{sig,MC}}$. Figure 8.6 shows the signal leakage fractions of each background region for PYTHIA and SHERPA. Once the signal leakage fractions are computed, setting R^{bg} to unity is the only assumption needed to have a solvable quadratic equation in N_A^{sig} . Taking the positive solution from the equation, or the smaller in the case of two positive solutions, as N_A^{sig} , the purity, shown in Figure 8.7, is computed. The signal purity computed with both samples behaves similarly. In the $p_T^{\text{ref}} > 1200$ GeV range, as unity is covered by the statistical uncertainty, the purity is set to 100%.

The nominal correction is obtained with the signal leakage fractions computed with the PYTHIA sample. The SHERPA sample correction is computed to estimate a source of uncertainty coming from the signal leakage fractions. Both correction factors are shown in Figure 8.8. The correction factors obtained from both samples range from around 99.3% in the first bin and approximate to unity as p_T^{ref} increases.

8.7 Systematic uncertainties

Several sources of systematic uncertainties are considered. Uncertainties due to the energy scale and resolution of the photons are derived from existing calibrations for each object

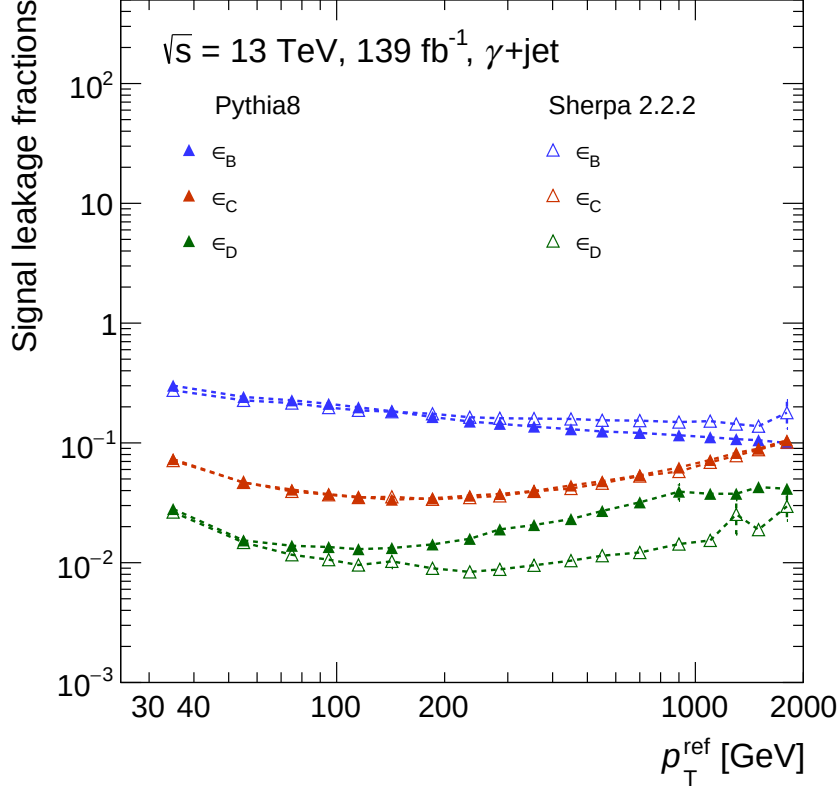


Figure 8.6: Signal leakage fractions as a function of p_T^{ref} in γ +jets events used for the purity correction of the jet response. Solid markers represent the signal leakage fractions computed with PYTHIA while SHERPA is displayed as empty markers. Error bars represent the statistical uncertainty.

and propagated through the corresponding analysis. The impact of additional parton radiation on the response measurement is evaluated by varying the selection criteria for the subleading jet veto and $\Delta\phi^{\text{ref, jet}}$. Uncertainties arising from pile-up suppression are estimated by comparing the response measurement between tighter and looser JVT working points. The statistical uncertainty dominates at the highest and lowest p_T bins. Pseudo-experiments are implemented in the estimation of uncertainties to reduce statistical fluctuations.

Photon energy scale and resolution

The uncertainty on the MPF response originating from the energy scale and resolution of the photon is estimated by propagating a positive and a negative component for the energy scale and the corresponding one for the energy resolution. The systematic uncertainty in

8 *In situ* jet energy scale calibration using γ +jets events

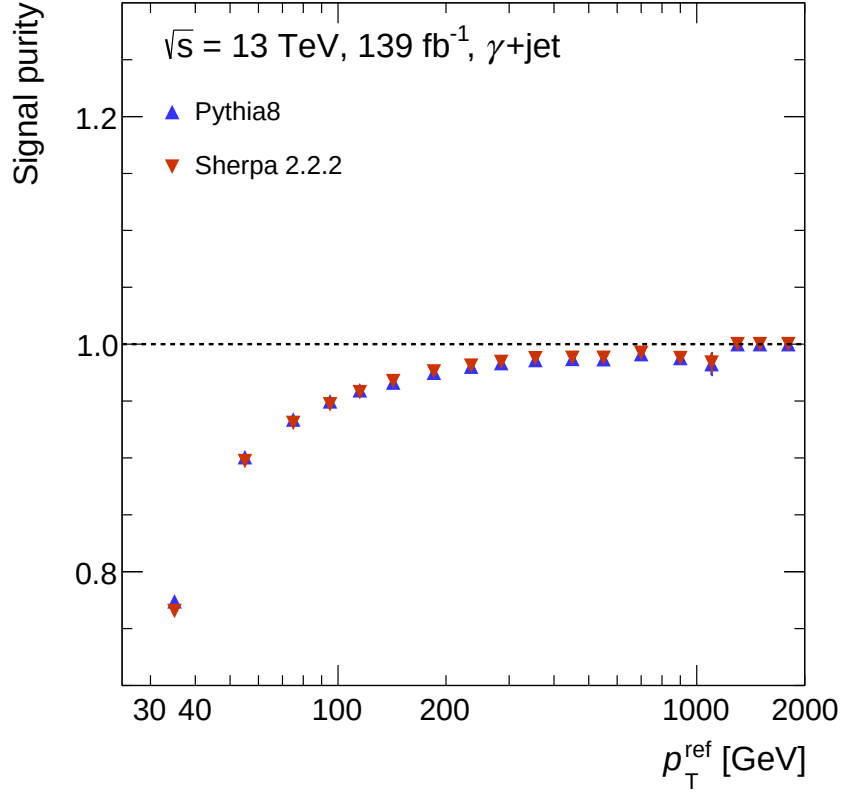


Figure 8.7: Signal purity in Region A as a function of p_T^{ref} in γ +jets events used for the correction of the jet response. The purity computed using PYTHIA (SHERPA) is represented as triangles pointing upward (downward). Error bars represent the statistical uncertainty.

the MC-to-data ratio coming from the energy resolution is found to be negligible. The uncertainty originated from the energy scale is the dominant contribution above 100 GeV, increasing from 0.3% in the low- p_T^{ref} region to 0.8%. A bootstrap method is used to compute this uncertainty together with rebinning to reduce the effect of statistical fluctuations.

Signal modelling

Two different simulated samples were used to compute the MC response for jets, PYTHIA and SHERPA. These samples differ in the hard interaction approximation order, the parton shower, and the fragmentation into hadrons. The signal modelling uncertainty is taken as the difference in the MC-to-data ratio. This contribution is dominant in the p_T^{ref} under

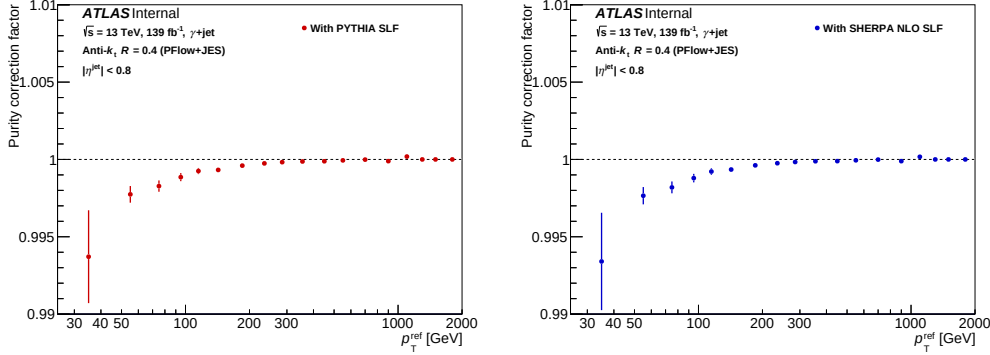


Figure 8.8: The purity correction factor for the jet MPF response computed as a function of p_T^{ref} in γ +jets events. The left plot displays the correction factor computed with PYTHIA and the one computed with SHERPA is presented in the plot on the right. Error bars represent the statistical uncertainty.

100 GeV. In this region, the uncertainty reaches 0.7% while above this threshold it is typically smaller than 0.2%.

Purity Correction

Uncertainties arising from photon purity are assessed using the same methodology documented in [186], in which one of the jets is misreconstructed as a photon. Three different contributions are taken into consideration to compute this uncertainty. They account for the effects on the purity correction coming from the sample used to compute the signal leakage fractions, the identification working point defining the non-*Tight* regions, and the deviation from unity of the $\gamma_{\text{ID}} - E_T^{\text{iso}}$ correlation. A detailed breakdown of each component of the uncertainty arising from the purity correction is presented in Appendix B.

Suppression of additional radiation

For the measurement of the jet MPF response, requirements on the topology of the γ +jets events are made. These requirements include a minimum azimuthal distance between the photon and the leading jet and a restriction on the p_T of the subleading jet. These criteria reduce the impact from additional QCD radiation ensuring a well-balanced topology. A systematic uncertainty is taken into account by varying each of these criteria. The $\Delta\phi$ selection is modified by ± 0.1 around the nominal value, $\Delta\phi > 2.9$. In the case of the veto to the subleading jet, the selection is changed from $p_T < \max(0.3 \times p_T^{\text{ref}}, 12)$ GeV to $p_T < \max(0.2 \times p_T^{\text{ref}}, 10)$ GeV and $p_T < \max(0.4 \times p_T^{\text{ref}}, 12)$ GeV. The estimated uncertainties for

8 *In situ* jet energy scale calibration using γ +jets events

both contributions are found to be negligible. A bootstrap method is used to compute this uncertainty together with rebinning to reduce the effect of statistical fluctuations.

Suppression of jets coming from pile-up

Events selected for the measurement of the MPF response are required to pass a JVT requirement to remove jets coming from pile-up collisions. An uncertainty accounting for the effect of this selection is taken into account by varying the JVT requirement. The nominal selection, which uses the $\text{JVT} > 0.2$ working point, is tightened to $\text{JVT} > 0.4$ and loosened by removing the JVT selection. This uncertainty is found to be negligible. A bootstrap method is used to compute this uncertainty together with rebinning to reduce the effect of statistical fluctuations.

Statistical uncertainty

The statistical uncertainty due to the limited luminosity affects both the fits and the purity corrections. The statistical uncertainty is significant in the low- p_T^{ref} range, where the trigger prescales limit the number of events recorded, reaching 0.8% and in the high- p_T^{ref} region, where the uncertainty reaches around 0.5%.

Total uncertainty

The total uncertainty for the calibration is presented for the γ +jet measurement in Figure 8.9 together with each of the individual components discussed above. Uncertainties are dominated by the modelling in MC simulations in the low- p_T regions and by the energy scale of the photon for $p_T > 100$ GeV. Apart from the first bin, the total uncertainty is smaller than 1%.

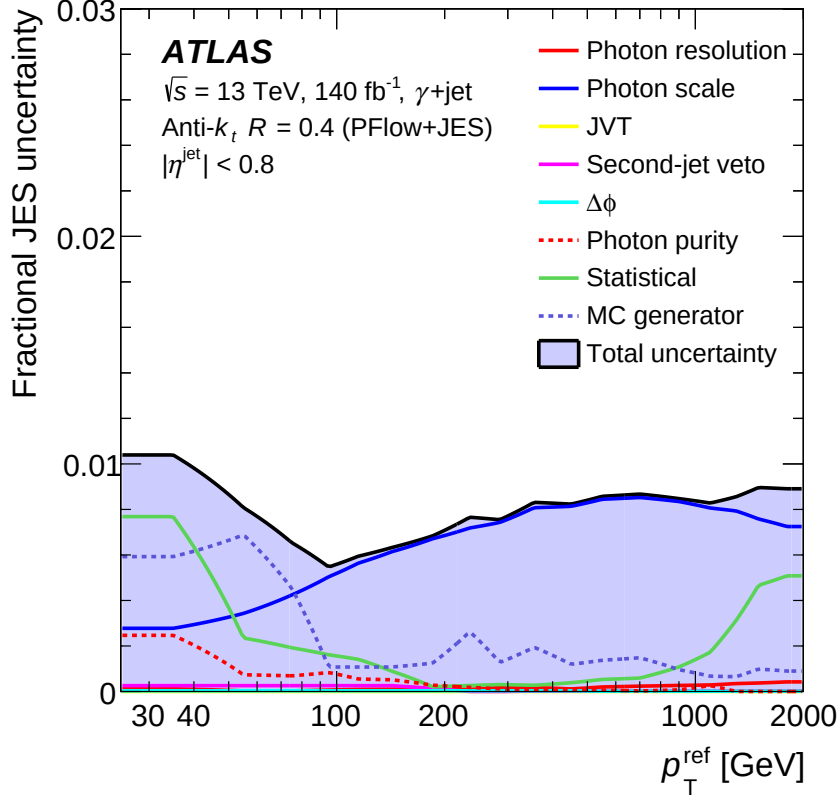


Figure 8.9: Total systematic uncertainty in the MPF response ratios as a function of p_T^{ref} for jets calibrated up to, and including, the η intercalibration in $\gamma + \text{jet}$ events together with the individual components.

8.8 Results

The MPF response as a function of the reference photon p_T is shown in Figure 8.10 using $\gamma + \text{jet}$ events for data and two different MC samples. The MC sample used to derive the nominal calibration for $\gamma + \text{jet}$ corresponds to PYTHIA 8. The alternative MC sample corresponds to SHERPA 2.2 to determine the uncertainty from the MC event modelling. The dip in the MPF response at low p_T^{ref} arises due to two opposing effects: the jet reconstruction threshold which tends to increase the response at the lowest jet p_T values between 25 GeV and 45 GeV, and the apparent rise in MPF response as a function of p_T . The MC-to-data response ratio for $\gamma + \text{jet}$ deviates from unity by less than 3% and decreases to just above 1% as p_T^{ref} increases for both simulations. The ratio is consistent with the results using Z+jet in Ref. [153].

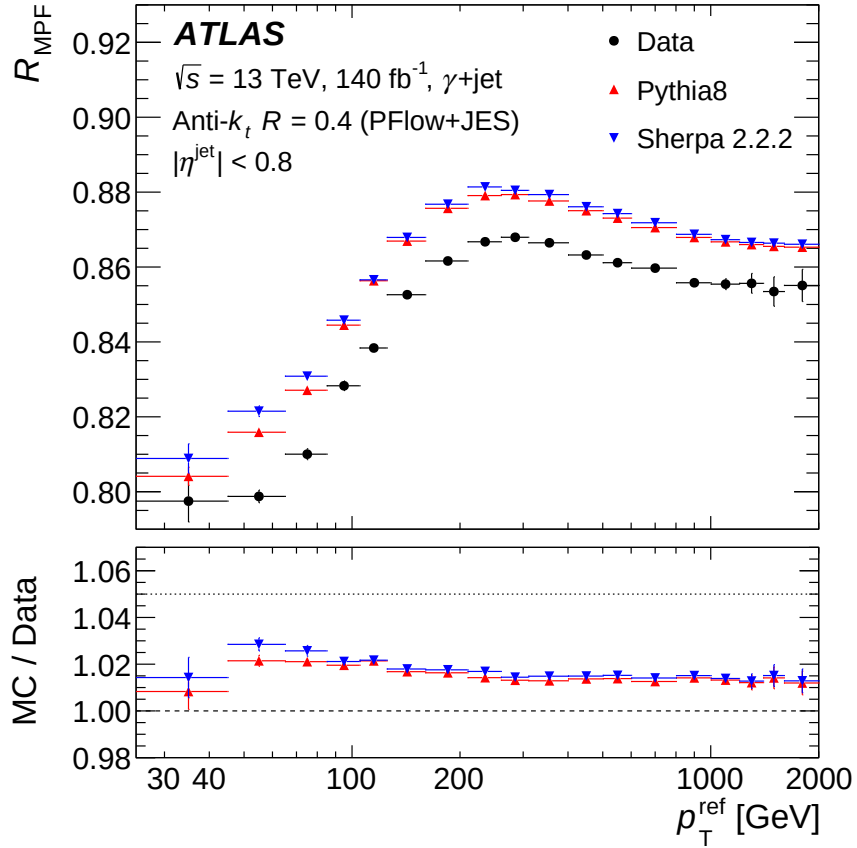


Figure 8.10: The MPF response as a function of p_T^{ref} measured in data and simulations for γ +jet. The data are represented by the black dots. The PYTHIA 8 predictions are represented by the triangles while the SHERPA 2.2 predictions are represented by the inverted triangles. The MC-to-data response ratios are shown in the bottom panel. The error bars correspond to the statistical uncertainties.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

The production of top-quark pairs ($t\bar{t}$) in pp collisions provides a testing ground for the SM at the TeV scale. Differential cross sections can be used to extract fundamental parameters of the SM, such as the top-quark mass [187, 188] and the strong coupling constant α_s [189], and to derive constraints on the PDFs [189]. These measurements are also sensitive to the existence of new physics beyond the SM [62, 190], which would appear as deviations of the measurements with respect to the SM predictions. In addition, for all searches for new physics in which $t\bar{t}$ production constitutes a background, the $t\bar{t}$ differential cross sections are particularly useful to highlight the differences in MC models and can be used to tune the MC parameters.

The ATLAS [9, 52–59, 61, 71, 72] and CMS [189, 191–201] Collaborations have published measurements of $t\bar{t}$ differential cross sections at centre-of-mass energies $\sqrt{s} = 7, 8$ and 13 TeV in pp collisions using final states containing a pair of leptons or a single lepton plus jets, both in the full phase space using parton-level variables and in fiducial phase-space regions using observables constructed from final-state particles (particle level). These results have been used extensively to improve the modelling of MC generators [202–206] and to reduce the uncertainties in the gluon PDF [167].

The analysis presented here focuses on the study of jet observables, which have not been explored in previous analyses, in the $t\bar{t}$ and $t\bar{t}$ +jets systems. In the SM, the top quark decays almost exclusively into a W boson and a b -quark. The signature of a $t\bar{t}$ decay is therefore determined by the W -boson decay modes. This analysis makes use of the ℓ +jets ($\ell = e, \mu$) $t\bar{t}$ decay mode, also called the semileptonic channel, where one W -boson decays into an electron or a muon and a neutrino, and the other W -boson decays into a quark–antiquark pair¹. Since the reconstruction of the top quark depends on its decay products, in the following the top quark is referred to as *hadronically* or *leptonically decaying top quark* (or alternatively *hadronic/leptonic top*), depending on the W -boson decay mode.

¹ Events in which the W -boson decays into an electron or muon through a τ -lepton decay may also meet the selection criteria.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

Absolute and normalised differential cross sections for the $t\bar{t}$ and $t\bar{t}$ +jets systems are measured as functions of various jet observables to characterise the kinematics, topology and dynamics of the $t\bar{t}$ system and are also sensitive to initial- and final-state QCD radiation effects, represented by the production of additional jets. State-of-the-art NLO and NNLO QCD predictions interfaced to parton-shower algorithms are compared to the measurements. The measured particle-level differential cross sections as functions of jet observables presented here benefit from the increase in integrated luminosity and the reduced detector uncertainties compared to previous publications. Therefore, more stringent tests of pQCD are achieved.

The structure of this chapter is as follows. Section 9.1 presents the data and MC samples used in the analysis. The object reconstruction and event selection is described in Section 9.2. Section 9.3 describes the background determination procedure. The pseudo-top algorithm is described in Section 9.4. Section 9.5 presents the event yields, and the distributions at the detector level of simulation. Section 9.6 describes the cross-section measurement procedure and the systematic uncertainties are presented in Section 9.7. Section 9.8 presents the results of the analysis.

9.1 Data and Monte Carlo samples

The data used in this analysis were collected with the ATLAS detector during the pp collision running periods of 2015–2018, when the LHC operated at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. Only events taken during stable beam conditions and satisfying detector- and data- quality requirements [154], which include the calorimeters, muon system and inner tracking detectors being in nominal operation, are considered. The total integrated luminosity of the collected sample is 140 fb^{-1} . The uncertainty in the combined 2015–2018 integrated luminosity is 0.83% [129], obtained using the LUCID-2 detector [126] for the primary luminosity measurements.

Samples of MC events are generated to study the characteristics of the signal events. These samples are also used to perform the unfolding procedure to obtain the measured cross sections. In addition, several MC samples are used to determine the background to the signal process, except for the background from misidentified or non-prompt leptons (called *fakes* background), which relies instead on a data-driven technique (see Section 9.3). All the MC samples are described below and summarised in Table 9.1.

Most of the samples of generated events are passed through a full detector simulation [207] based on the GEANT 4 [112] program. Other samples, used for estimating some of the systematic uncertainties, are produced using the fast simulation software AtIfast-II [208]. All MC samples are reconstructed and analysed by the same algorithms as the data.

9.1 Data and Monte Carlo samples

Table 9.1: Summary of signal and background MC samples, showing the event generator for the hard-scattering process, the order in pQCD of the cross section used for normalisation together with its value and total uncertainty, the PDF choice, as well as the parton-shower generator and the corresponding tune used.

Physics process	Generator & PDF set for hard process	Parton shower	Tune	Cross section normalisation
signal samples:				832 ± 51 pb
– PWG+PY8	POWHEG-BOX v2 & NNPDF3.0NLO	PYTHIA 8.230	A14	
– PWG+HW7	POWHEG-BOX v2 & NNPDF3.0NLO	HERWIG 7.1.3	H7.1-Default	
– aMC@NLO+HW7	MADGRAPH5_aMC@NLO 2.3.3 & NNPDF3.0NLO	HERWIG 7.1.3	H7.1-Default	
– SHERPA 2.2.12	SHERPA 2.2.12 & NNPDF3.0NNLO	SHERPA	SHERPA	
– PWG+PY8 MINNLOPS	POWHEG-BOX v2 & NNPDF3.0NNLO	PYTHIA + MINNLOPS	A14	
background samples:				
– single-top quark samples:				299 ± 8 pb
* PWG+PY8	POWHEG-BOX v2 & NNPDF3.0NLO	PYTHIA 8.230	A14	
* PWG+HW7	POWHEG-BOX v2 & NNPDF3.0NLO	HERWIG 7.0.4	H7.1-Default	
– W+jets	SHERPA 2.2.1 & NNPDF3.0NNLO	SHERPA	SHERPA	20100 ± 1000 pb
– Z+jets	SHERPA 2.2.1 & NNPDF3.0NNLO	SHERPA	SHERPA	1906 ± 95 pb
– diboson	SHERPA 2.2.2 & NNPDF3.0NNLO	SHERPA	SHERPA	211 ± 8 pb
– $t\bar{t}V$	MADGRAPH5_aMC@NLO & NNPDF3.0NLO	PYTHIA 8.230	A14	1.5 ± 0.1 pb
– $t\bar{t}H$	POWHEG-BOX v2 & NNPDF3.0NLO	PYTHIA 8.230	A14	0.51 ± 0.05 pb

Pile-up from additional pp collisions in the same and neighbouring bunch crossings is simulated by overlaying each MC event with a variable number of simulated inelastic pp collisions generated using PYTHIA 8.186 with the ATLAS set of tuned parameters for minimum-bias events (A3 tune) [209]. The MC events are weighted (pile-up reweighting) so that the distribution of the average number of interactions per bunch crossing matches the one observed in the data.

9.1.1 Signal simulation samples

In this section, the MC generators used for the simulation of the $t\bar{t}$ event samples are described for the nominal sample, the alternative samples used to estimate systematic uncertainties and the other samples used in the comparisons of the measured differential cross sections. The top-quark mass (m_t) and width are set to 172.5 GeV and 1.32 GeV [210], respectively, in all samples; additional samples, in which m_t and the width of the top quark are varied, are used to study the effect of these parameters (see below).

Nominal $t\bar{t}$ sample. The nominal signal $t\bar{t}$ events are generated using the NLO matrix elements (ME) implemented in the POWHEG-BOX v2 [211–214] generator with the NNPDF3.0NLO [167] PDF. Events where both top quarks decayed hadronically are not included. The parton shower, fragmentation, and the underlying events are simulated using PYTHIA 8.230 [215] with the NNPDF2.3LO PDF set and the A14 tune [157]. The h_{damp} parameter, which controls the p_T of the first gluon emission beyond the Born configuration

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

in POWHEG-BOX v2, is set to $1.5m_t$ [203]. The main effect of this parameter is to regulate the high- p_T emission against which the $t\bar{t}$ system recoils. Signal $t\bar{t}$ events generated with these settings are referred to as the nominal signal sample. In all the following figures and tables the predictions based on this MC sample are referred to as PWG+PY8.

Additional signal samples are generated for estimating modelling uncertainties (see Section 9.7). These samples are:

- events are generated using the same POWHEG-BOX v2 set-up but interfaced to HERWIG 7.1.3 [216, 217], which includes an angular-ordered parton-shower model. This sample is generated using the H7.1-Default tune [218] and the MMHT2014LO PDF set [176] in the parton shower. This MC sample is referred to as PWG+HW7;
- events are generated with MADGRAPH5_aMC@NLO 2.3.3 [219] and the NNPDF3.0NNLO PDF set. The parton shower, fragmentation, and the underlying events are simulated using HERWIG 7.1.3 with the MMHT2014LO PDF set and the H7.1-Default tune. This MC sample is referred to as aMC@NLO+HW7;
- events are generated with SHERPA 2.2.12 [160], which models the zero and one additional-parton process at NLO accuracy and up to four additional partons at LO accuracy using the MEPS@NLO prescription [171], with the NNPDF3.0NNLO PDF set [167]. The calculation uses its own parton-shower tune and hadronisation model. These MC samples are referred to as SHERPA 2.2.12;
- a sample is produced with the same settings as the nominal PWG+PY8 sample but with the h_{damp} parameter set to $3m_t$. This MC sample is referred to as PWG+PY8 h_{damp} ;
- two other PWG+PY8 samples with $m_t = 169 \text{ GeV}$ or $m_t = 176 \text{ GeV}$ are also generated. All other parameters remain the same as in the nominal PWG+PY8 sample, except for the width of the top quark, which is also adjusted in these two additional samples. These MC samples are referred to as PWG+PY8 m_t .

All the samples mentioned above, except for that used to unfold the data (see Section 9.6), were processed using fast simulation. All these MC predictions are normalised to NNLO plus NNLL calculations in pQCD using a k -factor. The cross section used to evaluate the k -factor is $\sigma_{t\bar{t}} = 832_{-29}^{+20}$ (scale) ± 35 (PDF, α_s) ± 23 (m_t) pb, as calculated with the TOP++ 2.0 [11–16, 220] program to NNLO in pQCD, including soft-gluon resummation to NNLL, and assuming $m_t = 172.5 \text{ GeV}$. The first uncertainty comes from the independent variation of the factorisation (μ_F) and renormalisation (μ_R) scales, the second one is associated with variations in the PDFs and α_s , following the PDF4LHC [221] prescription with the MSTW2008 68% CL NNLO [222], CT10 NNLO [223] and NNPDF2.35f FFN [156] PDF sets, and the third uncertainty arises from the assumption on the value of m_t .

9.1.2 Background simulation samples

Several processes can produce the same final state as the $t\bar{t} \ell + \text{jets}$ decay mode. The events produced by these backgrounds need to be estimated and subtracted from the data to determine the top-quark pair cross sections. The MC simulation used for these background processes are described below. The background processes simulated using MC samples are:

- Single top-quark production in the t - and s -channels and in association with a W -boson processes are modelled at NLO in QCD using POWHEG-BOX v2 with the NNPDF3.0NLO PDF set. The parton shower and hadronisation effects are simulated with PYTHIA 8.230, using the A14 tune and the NNPDF2.3LO PDF set. The t -channel sample is generated with the four-flavour scheme. The diagram-removal scheme (DR) [224] is employed in the tW simulation to handle the interference with $t\bar{t}$ production [203]. The single-top-quark cross section for the tW channel is normalised using its NLO+NNLL prediction, while the t - and s -channels are normalised using their NLO predictions [106, 225–228]. Events generated using the diagram-subtraction method [229] (DS) instead of the DR technique are considered as a systematic uncertainty on the modelling of this background (see Section 9.7). In addition, events generated with POWHEG-BOX v2 and interfaced to HERWIG 7.0.4 are also generated using the H7.1-Default tune and the MMHT2014LO PDF.
- Events containing a W or Z boson in association with jets ($W + \text{jets}$ and $Z + \text{jets}$) are simulated with the SHERPA 2.2.1 [160] generator using NLO-accurate MEs for up to two jets, and MEs accurate to LO for up to four jets calculated with the COMIX [161] and OPENLOOPS [163] libraries. They are matched with the SHERPA parton-shower [165] using the ME+PS@NLO prescription [171] and the tune developed by the SHERPA authors with the NNPDF3.0NNLO PDF set. The $W + \text{jets}$ and $Z + \text{jets}$ predictions are normalised to the NNLO cross sections [230, 231].
- Diboson production (WW , ZZ and WZ), with one of the bosons decaying hadronically and the other leptonically, is generated using SHERPA 2.2.2 with MEs computed at NLO accuracy in QCD with up to one additional parton and at LO accuracy with up to three additional partons. The NNPDF3.0NNLO PDF set is used. These predictions are normalised to the NLO cross sections evaluated by the generator.
- The production of $t\bar{t}V$, with $V = W, Z$, events is modelled at NLO using the MADGRAPH5_aMC@NLO 2.3.3 generator with the NNPDF3.0NLO PDF set. PYTHIA 8.230 with the A14 tune and the NNPDF2.3LO PDF is used to simulate the parton showers. These predictions are normalised to the NLO cross sections evaluated by the generator.
- The production of $t\bar{t}H$ events is modelled at NLO using the POWHEG-BOX v2 generator with the NNPDF3.0NLO PDF set. PYTHIA 8.230 with the A14 tune and the

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

NNPDF2.3LO PDF is used to simulate the parton showers. These predictions are normalised to the fixed-order calculation including NLO QCD and EW corrections [232].

9.1.3 NNLO signal simulated samples

The first matched computation of top-quark pair production at NNLO QCD, including all-order radiative corrections as implemented via parton-shower simulations, is reported in Ref. [233] and referred to as the MINNLOPS procedure. This procedure involves three steps. In the first step, $t\bar{t}$ pairs plus one light parton are generated using the POWHEG-BOX v2 program, which provides a NLO calculation and includes a second light parton. The second step involves the calculation of the limit in which the light partons generated in the previous step become collinear to the $t\bar{t}$ pair; appropriate Sudakov form factors and higher-order terms are used to guarantee that the calculation remains finite and accurate at NNLO for inclusive $t\bar{t}$ production. The third step includes the generation of the kinematics of the second radiated parton using POWHEG-BOX v2, which was accounted for only inclusively in the first step; in this way, the NLO accuracy of the $t\bar{t}$ +jet cross section is preserved. After this third step, subsequent radiation is generated using the parton-shower method of PYTHIA. For the generation of this sample, the NNPDF3.0 NNLO PDF set was used. In addition to the parton-shower step, PYTHIA was also used for the fragmentation and underlying-event simulation with the NNPDF23LO PDF sets and the A14 tune.

This prediction was also normalised to $\mathcal{O}(\text{NNLO}+\text{NNLL})$ in pQCD using the cross section $\sigma_{t\bar{t}} = 832 \text{ pb}$, as calculated with the TOP++ 2.0 program, as mentioned above. In all the following figures and tables the predictions based on this sample are referred to as PWG+PY8 MINNLOPS.

9.1.4 Theoretical uncertainties of the signal predictions

The theoretical uncertainties for the differential cross-section predictions are estimated in the following way:

- The uncertainties in the QCD predictions from PWG+PY8 and PWG+PY8 MINNLOPS due to higher-order terms are estimated by repeating the calculations using values of μ_R and μ_F scaled by factors 0.5 and 2.
- The uncertainty in the NLO QCD predictions from PWG+PY8 due to the uncertainty in the proton PDFs is estimated using the 100 replicas from the NNPDF3.0 analysis [167].
- The uncertainty in the NLO QCD predictions from PWG+PY8 due to the uncertainty in α_s is estimated by repeating the calculations using two additional sets of proton PDFs from the NNPDF3.0 analysis, for which different values of α_s at the mass of the

Z boson are assumed in the fits, namely 0.117 and 0.119; in this way, the correlation between α_s and the PDFs is preserved.

The total theoretical uncertainty for the NLO differential cross-section predictions of POWHEG is obtained by adding in quadrature the uncertainties listed above and it is dominated by the contribution from the scale variations. The total theoretical uncertainty is approximately 10% to 15% for the theoretical predictions from POWHEG as functions of the angular observables, whereas for the transverse-momentum and invariant-mass observables the uncertainty rises up to approximately 35% to 40% in the tail of the distributions. The same estimation of the theoretical uncertainties as described above is performed for the predictions of SHERPA 2.2.12 and aMC@NLO. The uncertainty from the scale variations in the aMC@NLO predictions is very similar to that of POWHEG. For the SHERPA 2.2.12 predictions, it is a factor between approximately 1.1 and 3 larger than for POWHEG. This increase in the uncertainty for the SHERPA calculation is expected due to the presence of a matrix element calculation with up to four additional partons at LO. These LO terms lead to larger uncertainties when varying the scales. For POWHEG and aMC@NLO these LO terms are absent and handled by the parton shower at lower accuracy.

For PWG+PY8 MINNLOPS, only scale variations are available; in these calculations at NNLO, a significant reduction of the scale uncertainty with respect to those for the NLO calculations is obtained. The uncertainties due to the PDFs and α_s are expected to be similar to, or smaller than, those affecting the NLO prediction [234]; the uncertainty due to the scale variations in the NNLO prediction is still larger than those arising from the PDFs and α_s uncertainties. For the comparison with the data in Section 9.8, the total uncertainty associated with the PWG+PY8 MINNLOPS prediction is taken as the sum in quadrature of the scale uncertainties of the NNLO calculation and, as an approximation, the uncertainties due to the PDFs and α_s as estimated from PWG+PY8 at NLO.

9.2 Object reconstruction and event selection

The following sections describe the detector- and particle-level objects used to characterise the final-state event topology and to define the fiducial phase-space regions for the measurements.

9.2.1 Detector-level object reconstruction, identification, calibration and selection

Primary vertices. Primary vertices are formed from reconstructed tracks that are spatially compatible with the interaction region. The hard-scatter primary vertex is chosen to be the one with at least two associated tracks and the highest $\sum p_T^2$, where the sum extends over all tracks with $p_T > 0.5$ GeV matched to the given vertex [235].

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

Electrons. Electron candidates are reconstructed as explained in Chapter 4 and are required to have a transverse momentum $p_T > 27 \text{ GeV}$ and a pseudorapidity $|\eta| < 2.47$, excluding the transition region between the barrel and endcap calorimeters ($1.37 < |\eta| < 1.52$). The longitudinal impact parameter $|\Delta z_0|$ of the associated track is required to satisfy $|\Delta z_0 \sin \theta| < 0.5 \text{ mm}$ and the transverse impact parameter significance $|d_0|/\sigma(d_0) < 5$. The electron candidates must satisfy a *Tight* likelihood-based identification criterion based on shower shape and track selection requirements (see Section 4.3). The candidates must also pass p_T - and η -dependent isolation requirements based on tracks and energy clusters, defined for the *Tight* isolation working point (see Section 4.5); these isolation requirements are used to reduce the background from jets misidentified as prompt leptons (fake leptons) or due to semileptonic decays of heavy-flavour hadrons (non-prompt real leptons) [236].

Muons. Muon candidates are reconstructed as explained in Chapter 5. The track associated with the muon candidate must have $|z_0 \sin \theta| < 0.5 \text{ mm}$ and $|d_0|/\sigma(d_0) < 3$. Candidates are required to have $p_T > 27 \text{ GeV}$ and $|\eta| < 2.5$. To reduce the background from muons originating from heavy-flavour decays inside jets, muons are required to be isolated using track-quality and isolation criteria similar to those applied to electrons (see Section 5.4).

Jets. Jets are reconstructed using the anti- k_t algorithm [101] with radius parameter $R = 0.4$ as implemented in the FASTJET package [237]. Jet reconstruction is performed on PFlow objects (see Chapter 6). Their four-momentum is calibrated as explained in Section 6.2. To reduce the number of jets originating from pile-up, the JVT *Tight* WP is used (see Section 6.3). The jets are selected if they have $p_T^{\text{jet}} > 25 \text{ GeV}$ and $|y^{\text{jet}}| < 2.5$.

b -tagging. To identify jets containing b -hadrons, the DL1r flavour tagging algorithm is used (see Section 6.4). Jets are considered as b -tagged if the value of the DL1r-algorithm output is larger than a certain threshold that provides a 70% b -jet tagging efficiency. The working point of 70% is selected after studying the number of events obtained in data, contribution from the different types of background processes and systematic uncertainties (see Appendix D).

Overlap removal. To remove ambiguities in the classification of the objects mentioned above, a procedure called *overlap removal* is applied to assign a unique hypothesis to each object. If a muon shares a track with an electron, it is likely to have undergone bremsstrahlung and hence the electron is not selected. To prevent double-counting of electron energy deposits as jets, the jet closest to a reconstructed electron is discarded if $\Delta R(\text{jet}, e) < 0.2$. Subsequently, to reduce the impact of non-prompt electrons, if $\Delta R(\text{jet}, e) < 0.4$, then that electron is removed. In cases where a jet is within $\Delta R(\text{jet}, \mu) = 0.4$ of a muon, if the jet has fewer than three tracks the jet is removed whereas if the jet has at least three tracks the muon is removed.

Missing E_T . The missing transverse momentum, E_T^{miss} , in the event is reconstructed as explained in Chapter 7 and is used to compute the transverse mass of the reconstructed leptonically decaying W boson (See Section 9.4). No requirement in E_T^{miss} is imposed in the event selection.

9.2.2 Particle-level object definition

Particle-level objects are defined in simulated events using only stable particles, which are defined as those with a mean lifetime $\tau > 30$ ps. The fiducial phase space used for the measurements in this analysis is defined using a series of requirements applied to particle-level objects analogous to those used in the selection of the detector-level objects.

Leptons. Electrons and muons are required to not originate from a generated hadron in the MC event, either directly or through a τ -lepton decay. This ensures that the lepton comes from an EW decay without requiring a direct match to a W boson. Events where the W -boson decays into a leptonically decaying τ -lepton are accepted. The four-momenta of the bare leptons are then modified by adding the four-momenta of all photons, not originating from hadron decay, within a cone of size $\Delta R = 0.1$, to take into account final-state photon radiation. Such *dressed leptons* are then required to have $p_T > 27$ GeV and $|\eta| < 2.5$.

Jets. Particle-level jets are reconstructed using the same anti- k_t algorithm used at the detector level with $R = 0.4$. The jet-reconstruction procedure takes as input all stable particles, except for charged leptons and neutrinos not from hadron decay as described above. Particle-level jets are required to have $p_T > 25$ GeV and $|\eta| < 2.5$.

b -tagging. A particle-level jet is identified as a b -jet if a hadron with $p_T > 5$ GeV containing a b -quark is matched to the jet through the ghost-matching technique described in Ref. [238].

Overlap removal. Particle-level objects are subject to different overlap removal criteria than reconstructed objects. After lepton dressing and jet clustering, muons and electrons with separation $\Delta R < 0.4$ from a jet are excluded. Since the electron–muon overlap removal at detector level depends on the detector-level reconstruction of these objects, it is not applied at particle level.

Missing E_T . Neutrinos from hadron decays, either directly or via a τ -lepton decay, are not considered. The particle-level missing transverse momentum is calculated from the four-vector sum of the selected neutrinos.

9.2.3 Particle- and detector-level event selection

The event selection comprises a set of requirements based on the general event quality and on the reconstructed objects, defined above, that characterise the final-state event topology. The event selection follows closely that of the previous analysis [61]. In addition, in this analysis, three channels are selected according to the number of jets in the final state. The same selection, described below, is applied to the reconstructed-level objects in data and MC samples and a similar one to the particle-level objects in the MC.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

Events were selected online during data taking by single-electron [239] or single-muon [240] triggers. Multiple triggers were combined in a logical OR to increase the selection efficiency. The lowest-threshold triggers utilised isolation requirements for reducing the trigger rate. The isolated-lepton triggers had p_T thresholds of 20 GeV for muons and 24 GeV for electrons in 2015 data, and 26 GeV for both lepton types in 2016, 2017 and 2018 data. They were complemented by other triggers with higher p_T thresholds but no isolation requirements to increase the trigger efficiency.

Offline, events are required to have a reconstructed primary vertex with two or more associated tracks and to contain exactly one reconstructed charged-lepton (electron or muon) candidate with $p_T > 27$ GeV. Each event is also required to contain at least four jets with $p_T > 25$ GeV and $|y| < 2.5$, of which at least two must be tagged as b -jets. This is the first channel studied, called $t\bar{t}$ inclusive. If in addition to these four jets there is a fifth jet, the second channel is selected and called $t\bar{t}+1\text{jet}$. If in addition to these five jets there is a sixth jet, the third channel is selected and called $t\bar{t}+2\text{jets}$. The $t\bar{t}+2\text{jets}$ sample is a subset of the $t\bar{t}+1\text{jet}$. The $t\bar{t}+1\text{jet}$ is a subset of the $t\bar{t}$ inclusive sample.

9.3 Background determination

After the event selection explained in Section 9.2.3, various backgrounds contribute to the event yields. Data-driven techniques are used to estimate the background that derives from events containing jets mimicking the signature of charged leptons or leptons from hadron decays, for which precise enough simulations are not available. Background from other physics processes, which include leptons in the final state, are taken into account using the MC samples described in Section 9.1.2.

Dilepton top-quark pair events (including decays into τ -leptons) can satisfy the event selection and are considered in the analysis as signal at both the detector and particle levels. All the $t\bar{t}$ semileptonic events that pass the fiducial selection described in Section 9.2.3 are considered as signal. For this reason, the leptonic top-quark decays into τ -leptons are considered as signal only if the τ -lepton decays leptonically. Cases where both top quarks decay into a τ -lepton, which in turn decays into hadrons, are accounted for in the multijet background.

9.3.1 Background determination from Monte Carlo samples

The single-top-quark background, comprising t -channel, s -channel and tW production modes, constitutes the largest contribution, amounting to 4.3%, 3.6% and 3.1% for the $t\bar{t}$ inclusive, $t\bar{t}+1\text{jet}$ and $t\bar{t}+2\text{jets}$ channels, respectively, of the total event yield in each channel. The W +jets background represents the second largest contribution to the background,

amounting to 2.3%, 1.9% and 1.8% for the three channels, respectively. The background contributions from each of the Z +jets, diboson, $t\bar{t}V$ and $t\bar{t}H$ samples of events to the total yield is smaller than 1% for all three channels.

9.3.2 Background determination from fake leptons

Multijet production processes, including production of hadronically decaying $t\bar{t}$ pairs, have a large cross section and mimic the ℓ +jets signature due to fake leptons or non-prompt real leptons. The multijet background is estimated directly from data using a matrix method [241]. The estimation is based on the introduction of a *Loose* lepton definition, obtained by removing the isolation requirement and loosening the identification criteria given in Section 9.2.1. The number of fake and non-prompt leptons contained in the signal region is evaluated by inverting the matrix that relates the number of *Loose* and *Tight* leptons to the number of real and fake leptons. This matrix is built using the efficiencies for fake and real leptons to pass the *Tight* selection. The fake-lepton efficiency is measured using data in control regions dominated by multijet background with the real-lepton contribution subtracted using MC simulation. The real-lepton efficiency is extracted by applying a tag-and-probe technique using leptons from Z -boson decays. The multijet background (called fakes) contributes 1.4%, 1.3% and 1.3% to the total event yield of each of the three channels, respectively.

9.4 Pseudo-top reconstruction algorithm

The pseudo-top algorithm [54] is used to reconstruct the leptonic and hadronic top quarks. This algorithm reconstructs the four-momenta of the top quarks and their complete decay chain from final-state objects, namely the charged lepton (electron or muon), missing transverse momentum, and four jets, two of which are b -tagged, using the same method both at detector and particle levels and described briefly below.

First, the two b -jet candidates are chosen. In events with more than two b -jet candidates, only the two with the highest p_T values are considered as the b -jets from the decay of the top quarks, and called *jet-b1* and *jet-b2*, with $p_T^{\text{jet-b1}} > p_T^{\text{jet-b2}}$. The rest of the jets are considered as non b -jet candidates.

The two non- b -tagged jets to form the hadronically-decaying W are chosen from the available jets in each event (excluding those already tagged as coming from a b -quark) such that the invariant mass of the pair is closest to m_W and called *jet-W1* and *jet-W2*, with $p_T^{\text{jet-W1}} > p_T^{\text{jet-W2}}$.

For the neutrino, the x and y components of its transverse momentum are taken from the corresponding components of the missing transverse momentum, whereas the z component

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

is calculated by imposing the W -boson mass constraint on the invariant mass of the charged-lepton–neutrino system. If the resulting quadratic equation has two real solutions, the one with the smaller value of $|p_z|$ is chosen. If the discriminant is negative, only the real part is considered.

The leptonically-decaying W boson is reconstructed from the charged lepton and the neutrino. The leptonic top quark is reconstructed from the leptonic W and the b -tagged jet closest (jet-b1 or jet-b2) in ΔR to the charged lepton. The hadronically-decaying W boson is reconstructed from the two non- b -tagged jets (jet-W1 and jet-W2) whose invariant mass is closest to the mass of the W boson, as mentioned above, and the hadronic top quark is reconstructed from the hadronic W boson and the other b -tagged jet. This choice yields the best performance of the algorithm in terms of the correspondence between the detector and particle levels.

9.5 Event yields, observables and detector-level distributions

9.5.1 Event yields

The individual e +jets and μ +jets channels have very similar distributions and give consistent results at detector level, thus they are combined by summing the distributions. Table 9.2 shows the final yields from the three channels for the combined ℓ +jets decay mode for data, signal and background events. The selection leads to a sample with an expected background of 9%, 8% and 8% for the three channels, respectively. The overall difference between data and prediction is 2%, 2% and 1% for the three channels, respectively, which means a good agreement of the SM with the data, after taking into account the systematic uncertainties (see Section 9.7).

9.5.2 Observables and detector-level distributions

Measurements of differential cross sections in the $t\bar{t}$ inclusive, $t\bar{t}+1\text{jet}$ and $t\bar{t}+2\text{jets}$ channels, as defined in Section 9.2.3, are performed as functions of different jet observables.

In the following, the two jets used to reconstruct the hadronically-decaying W are called *jet-W1* and *jet-W2*, with $p_T^{\text{jet-W1}} > p_T^{\text{jet-W2}}$ (as defined in Section 9.4). For the $t\bar{t}+1\text{jet}$ and $t\bar{t}+2\text{jets}$ channels, the additional jets, called *jet-rad1* and *jet-rad2*, are chosen as follows: *jet-rad1* is the highest p_T jet of those which are not part of the $t\bar{t}$ system; and *jet-rad2* is the second-highest p_T jet of those which are not part of the $t\bar{t}$ system.

The following observables are measured in the $t\bar{t}$ inclusive channel:

9.5 Event yields, observables and detector-level distributions

Table 9.2: Observed and expected number of events in the ℓ +jets decay mode after the full event selection. The systematic uncertainties (see Section 9.7) are also included. Due to correlations, the sum of the different impacts of systematic uncertainties might not add up to the total impact of all systematic uncertainties.

Process	Yield $t\bar{t}$ inclusive	Yield $t\bar{t}$ +1jet	Yield $t\bar{t}$ +2jets
$t\bar{t}$	$4\,120\,000^{+690\,000}_{-670\,000}$	$2\,110\,000 \pm 420\,000$	$860\,000 \pm 240\,000$
single-top quark	$194\,000 \pm 33\,000$	$82\,000^{+19\,000}_{-18\,000}$	$28\,900^{+8\,300}_{-8\,100}$
W +jets	$103\,000^{+53\,000}_{-52\,000}$	$44\,000 \pm 23\,000$	$17\,200 \pm 8\,900$
fakes	$66\,000 \pm 33\,000$	$30\,000 \pm 15\,000$	$12\,000 \pm 6\,000$
Z +jets	$37\,000 \pm 19\,000$	$14\,600 \pm 7\,600$	$5\,400 \pm 2\,800$
$t\bar{t}V$	$13\,100 \pm 1\,100$	$10\,300 \pm 900$	$6\,470^{+660}_{-650}$
diboson	$6\,000 \pm 600$	$2\,820^{+350}_{-340}$	$1\,220 \pm 180$
$t\bar{t}H$	$5\,460 \pm 320$	$4\,710 \pm 310$	$3\,260 \pm 270$
Total SM prediction	$4\,540\,000^{+710\,000}_{-700\,000}$	$2\,300\,000^{+440\,000}_{-430\,000}$	$940\,000 \pm 240\,000$
Data	4 445 113	2 248 410	924 751

- The kinematics of the two jets forming the W boson is characterised by measuring the transverse momenta of jet-W1 ($p_T^{\text{jet-W1}}$) and jet-W2 ($p_T^{\text{jet-W2}}$) and the absolute value of their rapidity ($|y^{\text{jet-W1}}|$ and $|y^{\text{jet-W2}}|$, respectively).
- The topology of the two jets forming the W boson is characterised by measuring $|\Delta y^{\text{jet-W1-jet-W2}}|$, defined as the absolute value of the difference in rapidity between jet-W1 and jet-W2, and $|\Delta\phi^{\text{jet-W1-jet-W2}}|$, defined as the absolute value of the difference in azimuth between jet-W1 and jet-W2.

Figure 9.1 shows the data, backgrounds and signal MC distributions at detector level for these observables. Some differences in shape between the distributions for data and SM predictions are observed, especially at high values of ($p_T^{\text{jet-W1}}$), which are covered by the systematic uncertainties.

The following observables are measured in the $t\bar{t}$ +1jet channel:

- The kinematics of the first QCD emission is characterised by measuring the transverse momentum of jet-rad1 ($p_T^{\text{jet-rad1}}$) and the absolute value of its rapidity ($|y^{\text{jet-rad1}}|$);

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

- The topology of the first QCD emission is characterised by measuring $|\Delta\phi^{\text{toplep}} - \text{jet-rad1}|$, defined as the absolute value of the difference in azimuth between the leptonic top (toplep) and jet-rad1, and $|\Delta\phi^{\text{tophad}} - \text{jet-rad1}|$, defined as the absolute value of the difference in azimuth between the hadronic top (tophad) and jet-rad1. In addition, the absolute value of the difference in azimuth between jet-W1 and jet-rad1 ($|\Delta\phi^{\text{jet-W1}} - \text{jet-rad1}|$) provides more insight into the topology of these events.
- The dynamics of the $t\bar{t}$ +1jet events is characterised by measuring the invariant mass of the $t\bar{t}$ – jet-rad1 system ($m^{t\bar{t} - \text{jet-rad1}}$).

Figure 9.2 shows the data, backgrounds and signal MC distributions at detector level for these observables. There is an adequate description of the measured distributions by the SM prediction; the $m^{t\bar{t} - \text{jet-rad1}}$ measured distribution is described by the SM prediction, except for the $m^{t\bar{t} - \text{jet-rad1}} > 4 \text{ TeV}$ region.

The following observables are measured in the $t\bar{t}$ +2jets channel:

- The kinematics of the second QCD emission is characterised by measuring the transverse momentum of jet-rad2 ($p_T^{\text{jet-rad2}}$) and the absolute value of its rapidity ($|y^{\text{jet-rad2}}|$);
- The topology of the second QCD emission is characterised by measuring $|\Delta y^{\text{jet-rad1} - \text{jet-rad2}}|$, defined as the absolute value of the difference in rapidity between jet-rad1 and jet-rad2, and $|\Delta\phi^{\text{jet-rad1} - \text{jet-rad2}}|$, defined as the absolute value of the difference in azimuth between jet-rad1 and jet-rad2. Other angular correlations provide more insight into the topology of the $t\bar{t}$ +2jets events: $|\Delta\phi^{\text{toplep}} - \text{jet-rad2}|$, defined as the absolute value of the difference in azimuth between the leptonic top and jet-rad2, $|\Delta\phi^{\text{tophad}} - \text{jet-rad2}|$, defined as the absolute value of the difference in azimuth between the hadronic top and jet-rad2, and $|\Delta\phi^{\text{jet-W1}} - \text{jet-rad2}|$, defined as the absolute value of the difference in azimuth between jet-W1 and jet-rad2.
- The dynamics of the $t\bar{t}$ +2jets events is characterised by measuring the invariant mass of the jet-rad1 and jet-rad2 system ($m^{\text{jet-rad1} - \text{jet-rad2}}$).

Figures 9.3 and 9.4 show the data, backgrounds and signal MC distributions at detector level for these observables. There is an adequate description of the measured observables by the SM prediction.

9.6 Cross-section measurement procedure

Differential cross sections are measured at particle level as functions of the observables defined in Section 9.5 for the three channels. The data distributions, after background subtraction (see Section 9.3), are unfolded to the particle level using the iterative application

9.6 Cross-section measurement procedure

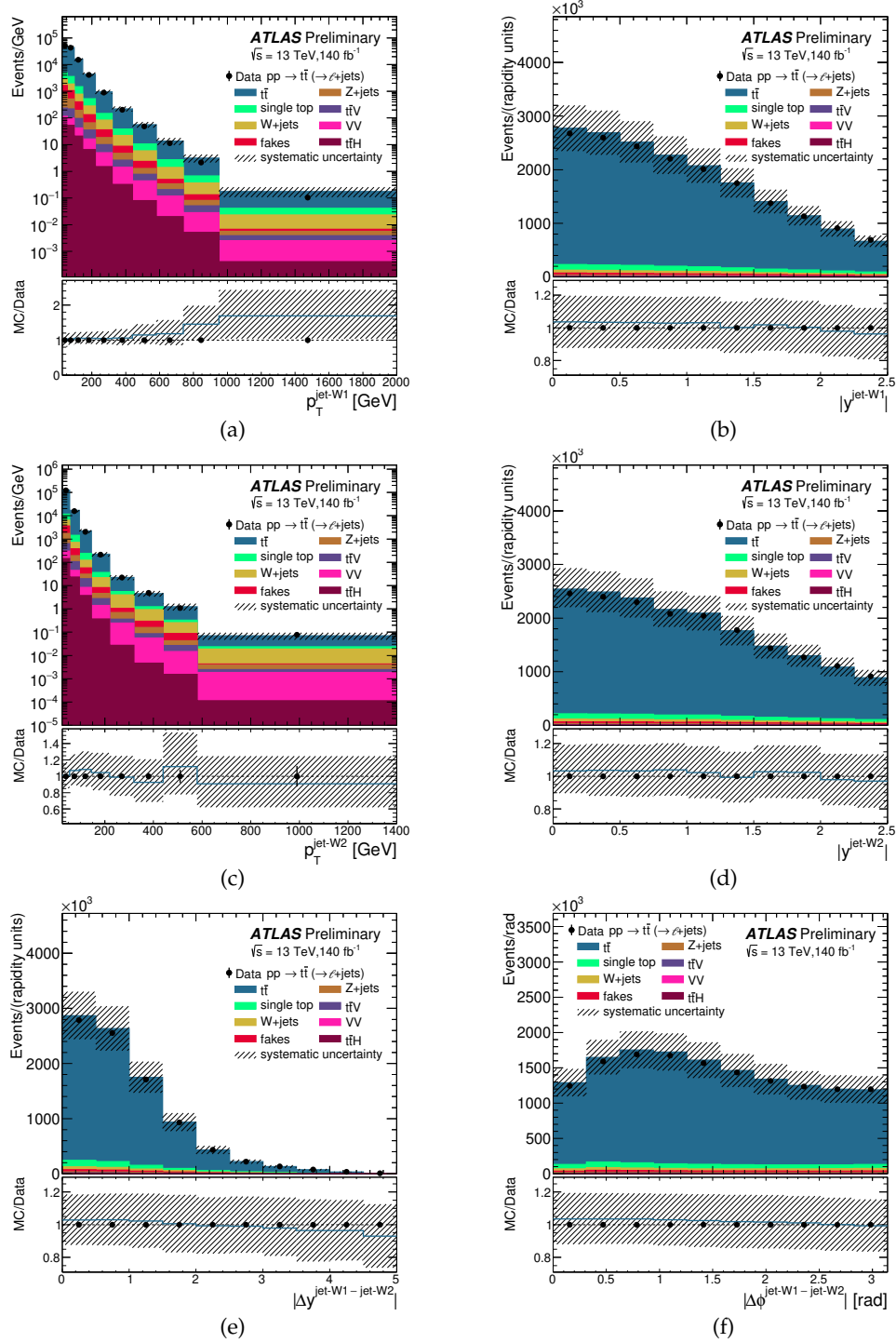


Figure 9.1: The measured (a) $p_T^{\text{jet-W1}}$, (b) $|y^{\text{jet-W1}}|$, (c) $p_T^{\text{jet-W2}}$, (d) $|y^{\text{jet-W2}}|$, (e) $|\Delta y^{\text{jet-W1-jet-W2}}|$, and (f) $|\Delta\phi^{\text{jet-W1-jet-W2}}|$ distributions divided by the bin width (dots) in the $t\bar{t}$ inclusive channel. The error bars represent the data statistical uncertainty; for many points, the uncertainty is smaller than the marker size and, thus, not visible. The hatched bands represent the systematic uncertainties (see Section 9.7). The MC simulation of the signal (blue histograms) and various backgrounds (represented by histograms of different colours) are also included. The lower part of each figure shows the ratio of the total SM and the data distributions.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

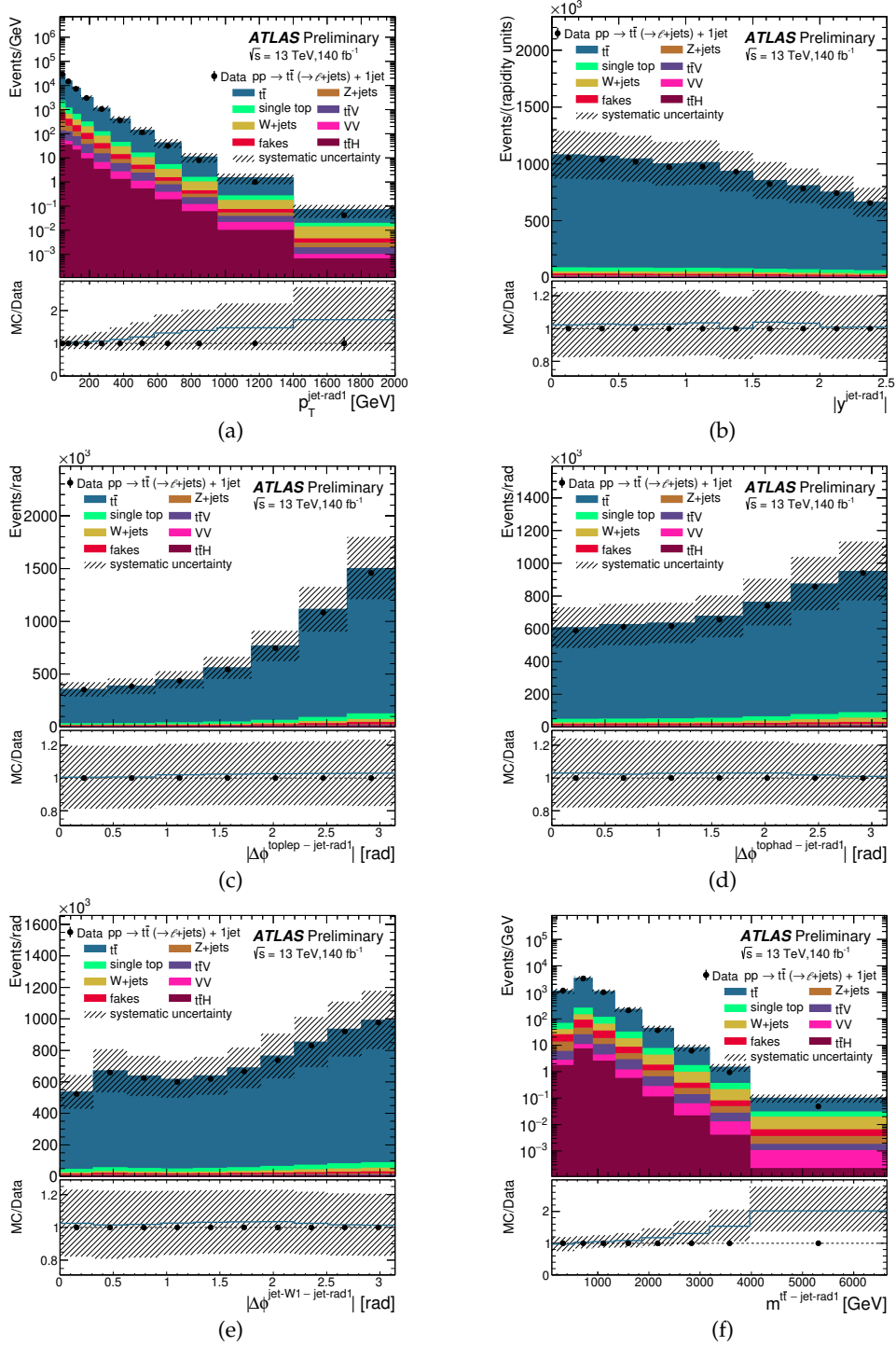


Figure 9.2: The measured (a) $p_T^{\text{jet-rad1}}$, (b) $|y^{\text{jet-rad1}}|$, (c) $|\Delta\phi^{\text{toplep-jet-rad1}}|$, (d) $|\Delta\phi^{\text{tophad-jet-rad1}}|$, (e) $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$, and (f) $m^{\text{tt-jet-rad1}}$ distributions divided by the bin width (dots) in the $t\bar{t} + 1\text{jet}$ channel. The error bars represent the data statistical uncertainty; for many points, the uncertainty is smaller than the marker size and, thus, not visible. The hatched bands represent the systematic uncertainties (see Section 9.7). The MC simulation of the signal (blue histograms) and various backgrounds (represented by histograms of different colours) are also included. The lower part of each figure shows the ratio of the total SM and the data distributions.

9.6 Cross-section measurement procedure

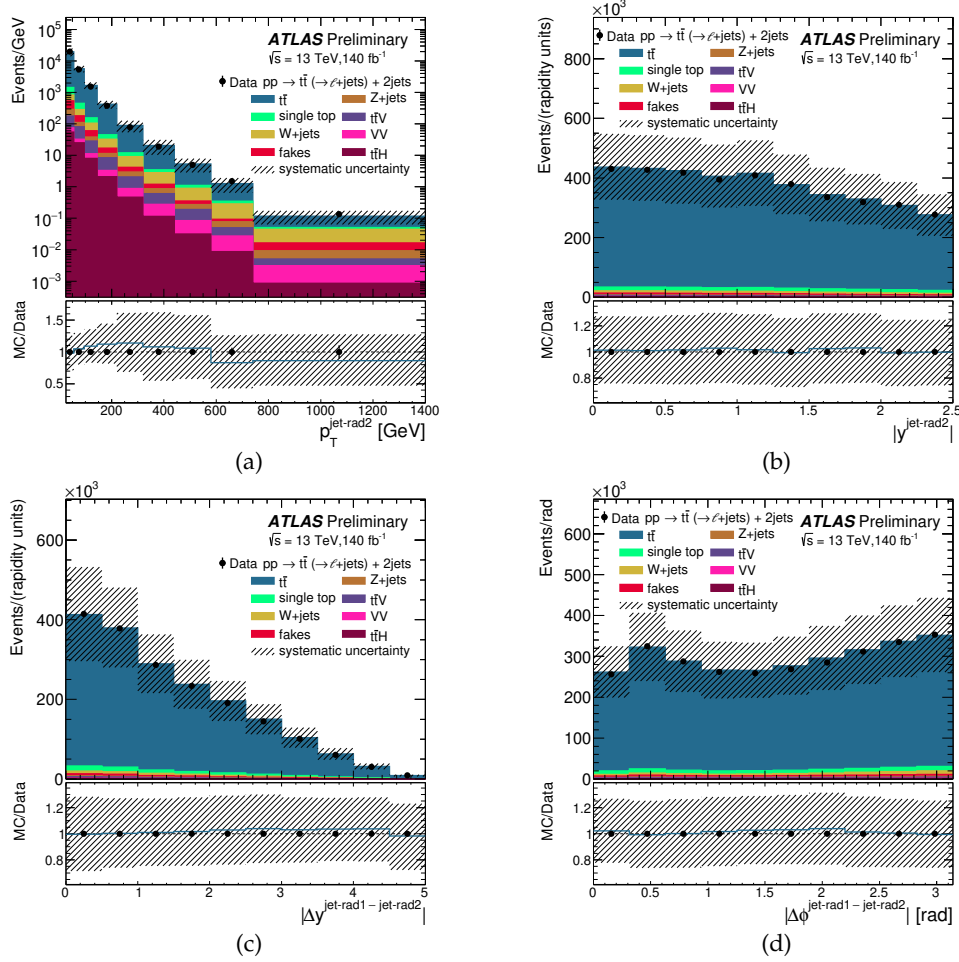


Figure 9.3: The measured (a) $p_T^{\text{jet-rad2}}$, (b) $|y^{\text{jet-rad2}}|$, (c) $|\Delta y^{\text{jet-rad1} - \text{jet-rad2}}|$, and (d) $|\Delta\phi^{\text{jet-rad1} - \text{jet-rad2}}|$ distributions divided by the bin width (dots) in the $t\bar{t} + 2\text{jets}$ channel. The error bars represent the data statistical uncertainty; for many points, the uncertainty is smaller than the marker size and, thus, not visible. The hatched bands represent the systematic uncertainties (see Section 9.7). The MC simulation of the signal (blue histograms) and various backgrounds (represented by histograms of different colours) are also included. The lower part of each figure shows the ratio of the total SM and the data distributions.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

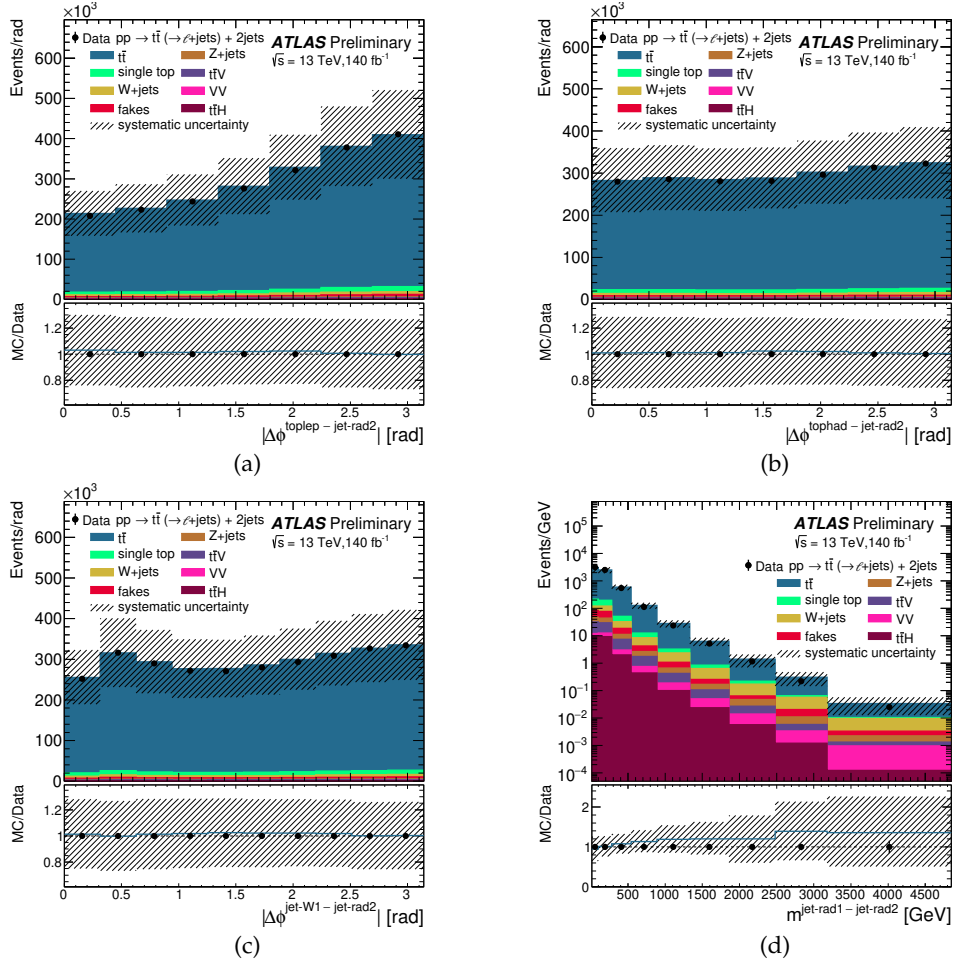


Figure 9.4: The measured (a) $|\Delta\phi^{\text{toptlep}} - \text{jet-rad2}|$, (b) $|\Delta\phi^{\text{tophad}} - \text{jet-rad2}|$, (c) $|\Delta\phi^{\text{jet-W1}} - \text{jet-rad2}|$, and (d) $m^{\text{jet-rad1} - \text{jet-rad2}}$ distributions divided by the bin width (dots) in the $t\bar{t} + 2\text{jets}$ channel. The error bars represent the data statistical uncertainty; for many points, the uncertainty is smaller than the marker size and, thus, not visible. The hatched bands represent the systematic uncertainties (see Section 9.7). The MC simulation of the signal (blue histograms) and various backgrounds (represented by histograms of different colours) are also included. The lower part of each figure shows the ratio of the total SM and the data distributions.

of Bayes' theorem to obtain the measured differential cross sections. The Bayesian unfolding [242] method as implemented in RooUnfold [243] is used. In this method, repeated application of Bayes' theorem is used to invert the unfolding matrix. The regularisation parameter is the number of iterations (N_{iter}), therefore regularisation is achieved by stopping the iterative procedure at a given value of N_{iter} . The results are found to be fairly insensitive to N_{iter} for most of the phase space measured; two iterations, i.e. $N_{\text{iter}} = 2$, are used in this analysis. Uncertainties related to this choice of number of iterations are considered (see Sections 9.6.1 and 9.7).

The differential cross section as a function of observable X in bin i is computed using the equation

$$\frac{d\sigma}{dX_i} \equiv \frac{1}{\mathcal{L} \cdot \Delta X_i} \cdot N_i^{\text{unf}},$$

where the index i iterates over bins of observable X , ΔX is the bin width, $\mathcal{L}$ is the integrated luminosity and N^{unf} represents the unfolded data value in each bin (overflow and underflow events are not considered when evaluating N^{unf}) and is calculated as

$$N_i^{\text{unf}} \equiv \frac{1}{\varepsilon_i} \cdot \sum_j f_{\text{match}}^j \cdot \mathcal{M}_{ij}^{-1} \cdot f_{\text{acc}}^j \cdot \left(N_{\text{data}}^j - N_{\text{bkg}}^j \right),$$

where N_{data} (N_{bkg}) is the number of data (background) events in each bin of observable X . The acceptance fraction f_{acc} is defined as

$$f_{\text{acc}}^j = \frac{N_{\text{particle} \wedge \text{detector}}^j}{N_{\text{detector}}^j},$$

where $N_{\text{particle} \wedge \text{detector}}$ is the number of detector-level signal MC events that satisfy the particle-level selection in each bin; this factor corrects for events that are generated outside the fiducial phase-space but satisfy the detector-level selection. The $\mathcal{M}$ is the unfolding matrix and is constructed from the correlation matrix, which maps the generated particle-level signal MC events to the detector-level events in their overlap region. Thus, the probability for particle-level events to be reconstructed in the same bin is represented by the diagonal elements, and the off-diagonal elements describe the fraction of particle-level events that migrate into other bins. To suppress combinatorial background in $\mathcal{M}$, additional requirements of angularly well matched reconstructed objects are applied. For all events, the reconstructed top quarks are required to satisfy $\Delta R(\text{top-}x_{\text{part}}, \text{top-}x_{\text{det}}) < 0.8$, where top- x represents the leptonic or hadronic top. In the $t\bar{t}$ inclusive channel, matching of the jets is performed by requiring $\Delta R(\text{jet-}y_{\text{part}}, \text{jet-}y_{\text{det}}) < 0.3$, where jet- y represents jet-W1, jet-W2. In the $t\bar{t}+1\text{jet}$ and $t\bar{t}+2\text{jets}$ channels, this requirement is applied to jet-rad1, jet-rad2 or both,

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

where necessary. In addition, the reconstructed objects (top or jets) which are not matched to a truth object as well as reconstruction inefficiencies (truth objects which are not matched to a reconstructed object) are taken into account in the Bayesian method. In addition, the inclusion of a matching scheme ensures the stability of the iterative process, so that the results are insensitive to N_{iter} , except for statistical fluctuations, as mentioned above. The effect of the matching procedure in the unfolding is accounted for by the matching fraction f_{match} , which is defined as

$$f_{\text{match}}^j = \frac{N_{\text{particle} \wedge \text{detector} \wedge \text{match}}^j}{N_{\text{particle} \wedge \text{detector}}^j},$$

where $N_{\text{particle} \wedge \text{detector} \wedge \text{match}}$ is the number of signal MC detector-level events that satisfy the particle-level selection and matching requirements and $N_{\text{particle} \wedge \text{detector}}$ is the number of particle-level events that satisfy the detector-level selection without the matching requirements. Finally, the efficiency ε is defined as

$$\varepsilon^i = \frac{N_{\text{particle} \wedge \text{detector}}^i}{N_{\text{particle}}^i},$$

where N_{particle} is the number of signal MC events at particle level in each bin; this factor corrects for events that satisfy the particle-level selection but are not selected at the detector level.

Figure 9.5 shows the ε , f_{match} and f_{acc} for $|y^{\text{jet-W1}}|$, $|y^{\text{jet-rad1}}|$ and $|y^{\text{jet-rad2}}|$ and the unfolding matrices for the same observables. The signal samples of PWG+PY8 are used in this figure.

The normalised differential cross sections are measured by dividing the differential cross sections by the total cross section in the region of phase space considered for each channel.

9.6.1 Closure and stress tests

The unfolding procedure is checked by performing several closure and stress tests, which validate the consistency of the procedure.

The first test is a cross-check of the reconstructed-truth level matching procedure used: the cross sections are unfolded with (nominal cross sections) and without matching in the transfer matrices and the ratio of these two cross sections is performed; these ratios are within the statistical uncertainty of the measurement, and typically consistent with unity within 1%. Thus, it is verified that the matching procedure does not affect the unfolding.

A second test is performed to check that the unfolding procedure recovers the generated result in a statistically-independent sample. For this test, a MC signal sample is divided into two

9.6 Cross-section measurement procedure

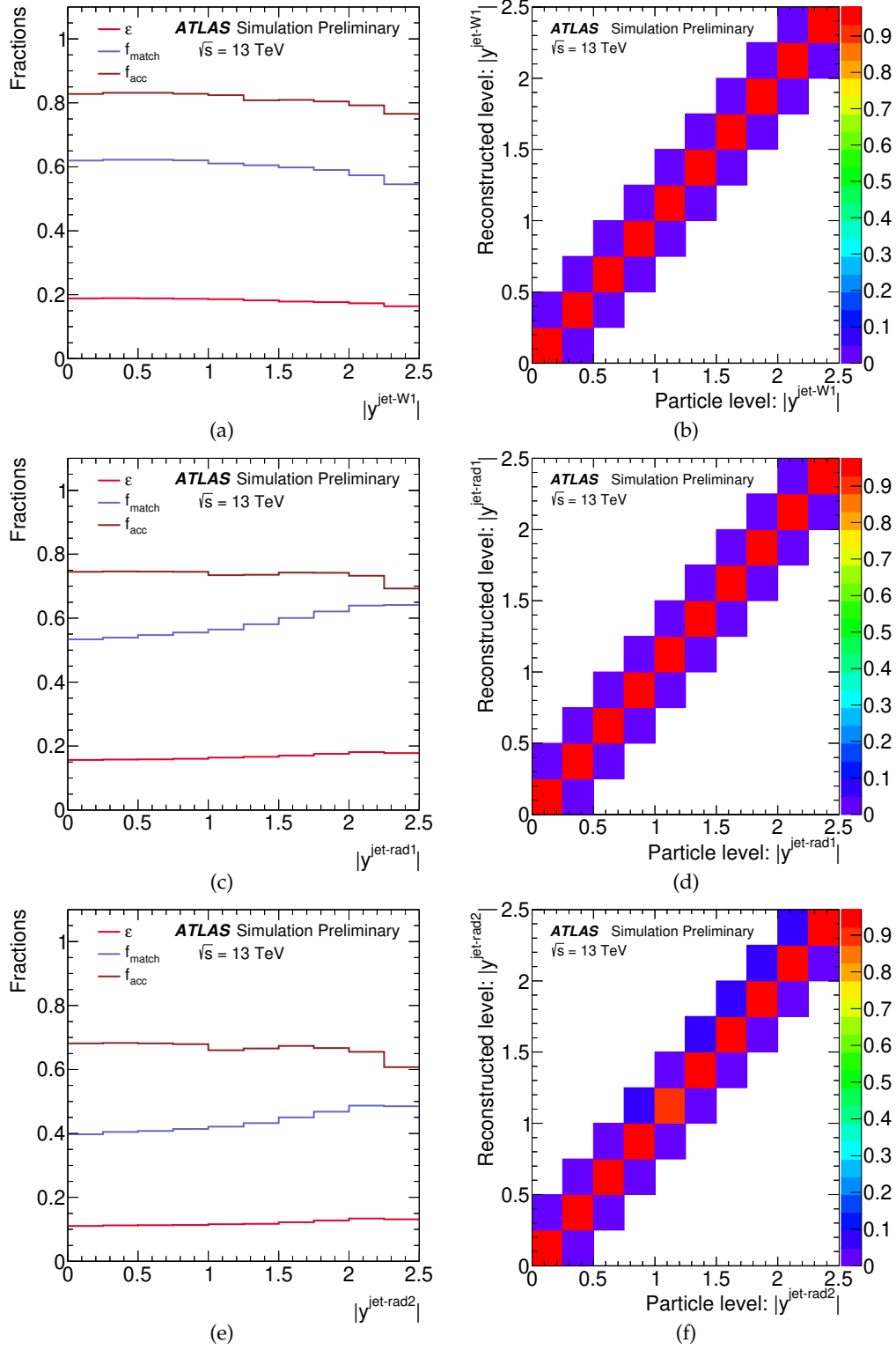


Figure 9.5: The efficiency ε (red lines), matching fraction f_{match} (blue lines) and acceptance fraction f_{acc} (brown lines) for (a) $|y^{\text{jet-W1}}|$, (c) $|y^{\text{jet-rad1}}|$, and (e) $|y^{\text{jet-rad2}}|$. The unfolding matrix $\mathcal{M}_{ij}$ for (b) $|y^{\text{jet-W1}}|$, (d) $|y^{\text{jet-rad1}}|$, and (f) $|y^{\text{jet-rad2}}|$; the distributions are normalised such that the sum of the entries in each column is equal to unity. The signal samples of PWG+PY8 are used.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

subsamples A and B ; sample A is used as pseudo-data and unfolded to particle level using sample B (pseudo-data cross section). The pseudo-data cross sections are compared with the generated results of sample A , B or $A + B$; the differences observed in these comparisons are typically less than 1%. Thus, the generated result is recovered in a statistically-independent sample.

A stress test is performed by weighting the nominal MC signal sample so that it matches the data distributions after background subtraction at reconstruction level. This weighted MC sample is used to unfold the data (weighted cross sections); the ratios of the weighted and nominal cross sections are within the statistical uncertainty of the measurement, and typically consistent with unity within 1%. The differences observed in this test are considered as a non-closure systematic uncertainty (see Section 9.7).

A second stress test is performed by using the weighted MC sample as pseudo-data and unfolding it with the nominal MC sample (weighted pseudo-data cross section); the ratio of the weighted pseudo-data cross section and the generated result gives back, as expected, the input fitting function for each observable. Thus, it is concluded that differences between data and MC at reconstruction level do not affect the unfolded measured cross sections.

A third stress test is performed to ascertain the level of recoverability of the signal for the possible presence of new physics in the data. The existence of a $t\bar{t}$ resonance is hypothesised using a gaussian-shape bump which is introduced at particle level and propagated to the reconstruction level using a weighting procedure (stressed pseudo-data). To determine whether the observables studied are sensitive to this hypothetical resonance and whether the unfolding procedure is able to recover the stressed shape and normalisation, the stressed pseudo-data is unfolded using the nominal MC sample and compared to the stressed generated result. The ratio of these cross sections shows that the shape and normalisation of the effects due to the resonance are recovered for all observables.

9.7 Systematic uncertainties

The sources of systematic uncertainties that affect the measurements of the differential cross sections include the b -tagging calibration and efficiency, the detector energy scale and resolution, the reconstruction efficiency, the background and signal modelling, the statistics of the MC samples and the uncertainty arising from the non-closure. Each source of uncertainty is separated into independent contributions and evaluated by varying the background and the signal MC samples in a correlated way, where necessary. The varied background samples are used to subtract the background and the varied signal samples are used to unfold the background-subtracted data distributions, as explained in Section 9.6. Systematic uncertainties are evaluated with respect to the nominal measured cross sections separately for the upward and downward variations (or in the case of a single variation by symmetrising the single deviation) for each bin of each observable. For the normalised cross

sections, each contribution to the systematic uncertainty is assumed to be fully correlated. Each source of systematic uncertainty is discussed separately in the following.

***b*-tagging calibration and efficiency.** The uncertainties in the *b*-tagging calibration and efficiency are determined separately for *b*-jets, *c*-jets and light-flavour-jets [244–246] using a component breakdown in 19 independent contributions. They depend on the p_T for *b*- and *c*-jets, and on the p_T and η for light-flavour jets, and they account for differences between data and simulation. The dominant contribution arises from the *b*-quark eigenvectors and amounts to typically 3% for the absolute cross sections and decreases to typically below 0.5% for the normalised cross sections.

Detector energy scale and resolution. The JES is calibrated using a combination of test beam data, simulation and *in situ* techniques (see Section 6.2.1). Its uncertainty is decomposed into a set of 30 uncorrelated components, with contributions from pile-up, jet flavour composition and response, single-particle response, and effects of jets not contained within the calorimeter. The uncertainty of the JER is represented by eight components accounting for jet- p_T and η -dependent differences between simulation and data (see Section 6.2.2). The uncertainty in E_T^{miss} due to a possible miscalibration of its soft-track component is derived from data and simulation comparisons of the p_T balance between the hard and the soft E_T^{miss} components (see Chapter 7). To account for the difference in pile-up distributions between the simulation and data, the pile-up profile in the simulation is corrected to match the one in the data. The uncertainty associated with the correction factor is applied. The uncertainty on the electron and muon energy scale and resolution is also included. The resulting 49 up and down independent contributions are added in quadrature and the dominant contribution comes from the effect of the pile-up on the JES. This uncertainty is typically in the range between 3 to 10%, depending on the observable, for the absolute cross sections and decreases to the range between 0.3 to 9% for the normalised cross sections.

Reconstruction efficiency. Reconstruction, identification, isolation and trigger performance for electrons and muons differ between data and simulations. Scale factors are applied to simulated events to correct for these differences. These scale factors are assessed using $Z \rightarrow \ell^+ \ell^-$ events in simulation and data. The associated systematic uncertainties are propagated to the measured cross sections. The JVT requirement and pile-up systematics are also included. The resulting 16 up and down independent contributions are added in quadrature and the dominant contribution comes from the JVT contribution. This uncertainty is in the range between 0.5 to 2% for both the absolute and normalised cross sections.

Background modelling. The uncertainties on the modelling of the background are estimated by changing the normalisation of the MC samples according to the uncertainty of the predicted cross section used for the normalisation of these samples (see Table 9.1). For the W +jets, Z +jets and fakes background, an uncertainty of 50% is included for the normalisation. This value for the uncertainty of the normalisation is determined to cover the differences observed between data and simulations in control regions with exactly three jets and at most one *b*-tagged jet or at least four jets and at most one *b*-tagged jet (see Appendix E);

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

these control regions are away from the signal region and enriched in these background processes. In addition, for the single-top-quark background, since the tW contribution is the largest among the three single-top-quark production channels, the uncertainty arising from the method used to handle the interference between tW and $t\bar{t}$ production is determined by comparing the tW simulated sample that uses the DS method with the nominal one based on the DR technique. The resulting uncertainties are symmetrised and added in quadrature. This uncertainty is typically in the range between 1.5 to 15%, depending on the observable, for the absolute cross sections and decreases to the range between 0.1 to 10% for the normalised cross sections.

Signal modelling. The uncertainties that affect the signal modelling and the method of estimation are discussed below. The effects of the parton-shower and hadronisation models are estimated by unfolding the data after background subtraction using the PWG+HW7 sample and performing the difference with respect to the nominal cross section (typically $< 2\%$). Since the single-top-quark processes are the dominant background, PWG+HW7 samples were also used to model this process when estimating this uncertainty. The impact of the matching procedure between the ME calculation and the parton-shower model is evaluated using the aMC@NLO+HW7 sample to unfold the data after background subtraction and performing the difference with respect to the cross section unfolded using the PWG+HW7 samples, both for signal and single-top-quark background (typically $\approx 1\%$). The uncertainties affecting the description of the hard gluon radiation are evaluated by unfolding the data after background subtraction using the PWG+PY8 h_{damp} sample and performing the difference with respect to the nominal cross section (typically $< 0.5\%$). The effect of the assumed value of m_t (172.5 GeV) is estimated using samples of PWG+PY8 generated with $m_t = 169$ GeV or $m_t = 176$ GeV to unfold the data after background subtraction and performing the difference with respect to the nominal cross section; the value of the uncertainty obtained in this way is then scaled by $1/3.5$, to account for the current uncertainty in m_t (1 GeV) and amounts to typically $< 0.3\%$. The first three contributions listed above are symmetrised. The dominant contribution arises from the parton-shower and hadronisation models. The resulting uncertainty is typically in the range between 0.5 to 10%, depending on the observable, for the absolute cross sections and decreases to the range between 0.2 to 8% for the normalised cross sections. The theoretical uncertainties that affect the unfolding procedure and the method of estimation are discussed below. These uncertainties affect the $t\bar{t}$ signal and the single-top-quark background samples and are taken as correlated in the two MC samples. The uncertainty due to missing higher-order QCD corrections in the ME computation is estimated by independently varying μ_R and μ_F by factors of 2 and 0.5 with respect to the central value (typically $< 1\%$). The uncertainties in the amount of initial- and FSR from the parton shower are assessed by, respectively varying the corresponding parameter of the A14 parton shower tune (Var3c) [157] (typically $< 0.5\%$) and by varying by factors of 2 and 0.5 the scale μ_R^{FSR} (typically $\approx 2\%$). The uncertainties due to the PDFs are estimated following the PDF4LHC prescription [221, 247] (typically $< 0.5\%$). The dominant contribution to the systematic uncertainty arising from the theoretical uncertainties on the

unfolding comes from the variation of the μ_R^{FSR} scale. The resulting uncertainty is typically in the range between 1 to 5%, depending on the observable, for the absolute cross sections and decreases to the range between 0.5 to 4% for the normalised cross sections.

Non-closure. As explained in Section 9.6.1, the differences observed in the stress test are considered as a non-closure systematic uncertainty. This uncertainty is typically below 0.5%.

9.7.1 Total systematic uncertainty

The total systematic uncertainty is computed by adding in quadrature the sources of uncertainty listed in the previous section. Figures 9.6 to 9.9 (9.10 to 9.13) show the resulting total systematic uncertainty, together with the contribution from each source, the statistical uncertainty of the MC samples and the data statistical uncertainty for the absolute (normalised) cross sections. The statistical uncertainties in data and MC samples are evaluated using *toy experiments* (bootstrap technique) with 1000 statistically-independent replicas. In the figures, the systematic contributions are stacked, except for the data statistical uncertainty. In addition, there is an uncertainty from the combined 2015–2018 integrated luminosity, which amounts to 0.83% [129] (not included in the figures).

In the case of the angular observables, for the $t\bar{t}$ inclusive channel, the total relative systematic uncertainty is $\approx 7\%$ and dominated by the b -tagging calibration uncertainty. In the case of the $t\bar{t}+1\text{jet}$ channel, the total relative systematic uncertainty amounts to $\approx 10\%$ and is dominated by the detector energy scale and resolution. The $t\bar{t}+2\text{jets}$ channel has a total relative systematic uncertainty of $\approx 13\%$ and is also dominated by the detector energy scale and resolution. For the normalised cross sections, the total relative systematic uncertainty is $\approx 1\%$, $\approx 1.5\%$ and $\approx 2\%$ for the $t\bar{t}$ inclusive, $t\bar{t}+1\text{jet}$ and $t\bar{t}+2\text{jets}$ channels.

In the case of the transverse momentum and invariant-mass observables, the values for the uncertainties quoted above are only valid at low p_T or invariant-mass. The total relative systematic uncertainty in the absolute (normalised) cross sections increases up to $\approx 38\%$ ($\approx 37\%$) for $p_T^{\text{jet-W1}}$, $\approx 38\%$ ($\approx 37\%$) for $p_T^{\text{jet-W2}}$, $\approx 57\%$ ($\approx 55\%$) for $p_T^{\text{jet-rad1}}$, $\approx 77\%$ ($\approx 74\%$) for $m^{t\bar{t}-\text{jet-rad1}}$, $\approx 50\%$ ($\approx 52\%$) for $p_T^{\text{jet-rad2}}$ and $\approx 48\%$ ($\approx 46\%$) for $m^{\text{jet-rad1}-\text{jet-rad2}}$ in the last bins of these distributions. In these regions of phase space, the background modelling and statistical uncertainties dominate.

The data statistical uncertainty is below 0.3% for the angular observables in the $t\bar{t}$ inclusive channel; in the tails of the $p_T^{\text{jet-W1}}$ and $p_T^{\text{jet-W2}}$ cross sections, it increases up to 17% and 19%, respectively. For the $t\bar{t}+1\text{jet}$ channel, the data statistical uncertainty is $\approx 0.3\%$ for the angular observables and increases up to 36% and 23% in the last bin of the $p_T^{\text{jet-rad1}}$ and $m^{t\bar{t}-\text{jet-rad1}}$ cross sections, respectively. In the $t\bar{t}+2\text{jets}$ channel, the data statistical uncertainty is $\approx 0.4\%$

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

for the angular observables and increases up to 19% and 28% in the last bin of the $p_{\text{T}}^{\text{jet-rad2}}$ and $m^{\text{jet-rad1} - \text{jet-rad2}}$ cross sections, respectively.

9.7 Systematic uncertainties

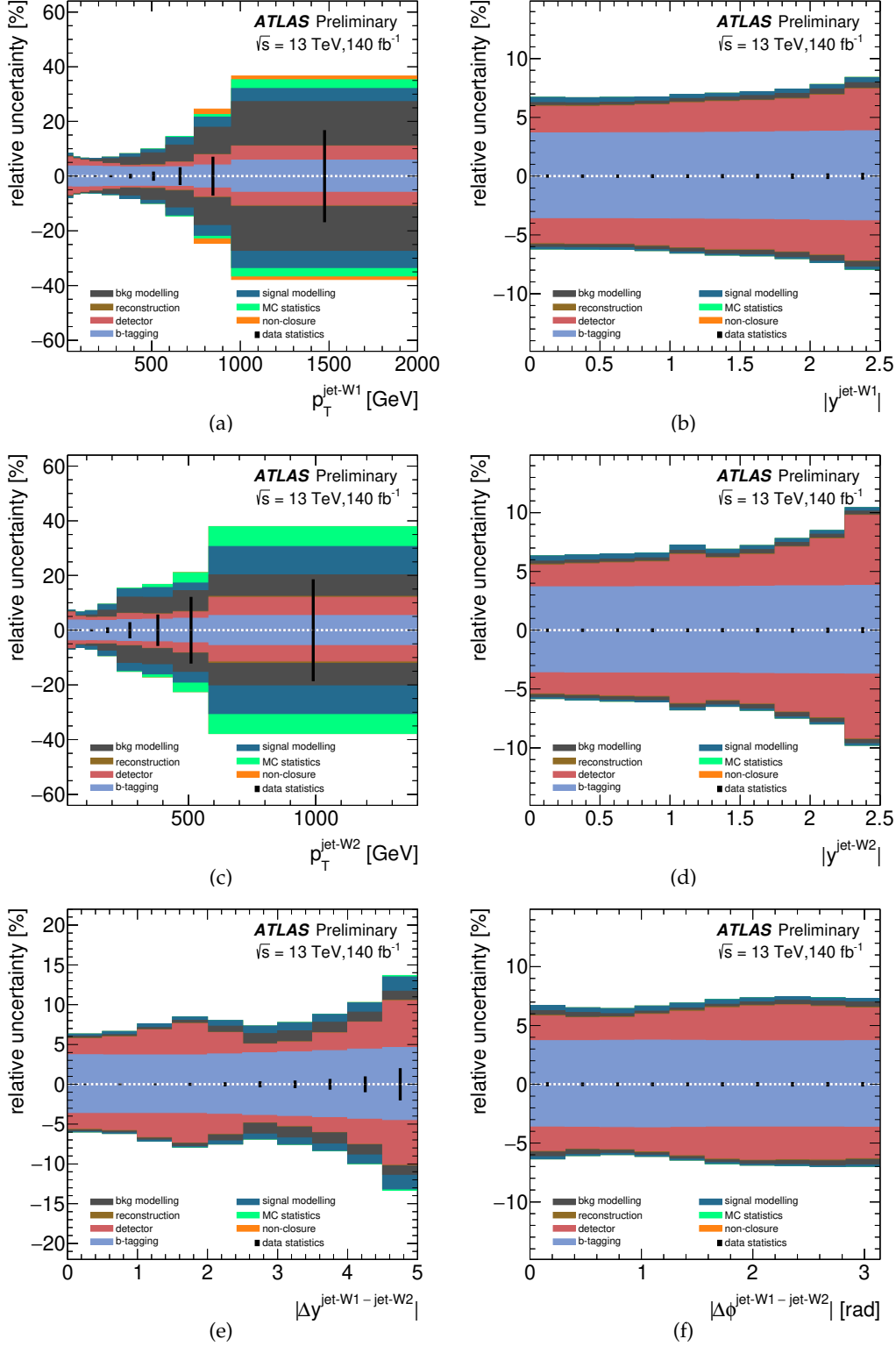


Figure 9.6: Total relative systematic uncertainties in percentage on the measured cross sections for (a) $p_T^{\text{jet-W1}}$, (b) $|y^{\text{jet-W1}}|$, (c) $p_T^{\text{jet-W2}}$, (d) $|y^{\text{jet-W2}}|$, (e) $|\Delta y^{\text{jet-W1} - \text{jet-W2}}|$, and (f) $|\Delta\phi^{\text{jet-W1} - \text{jet-W2}}|$ in the $t\bar{t}$ inclusive channel. The individual sources of systematic uncertainty (plotted in stacked form) and the data statistical uncertainty are also included.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

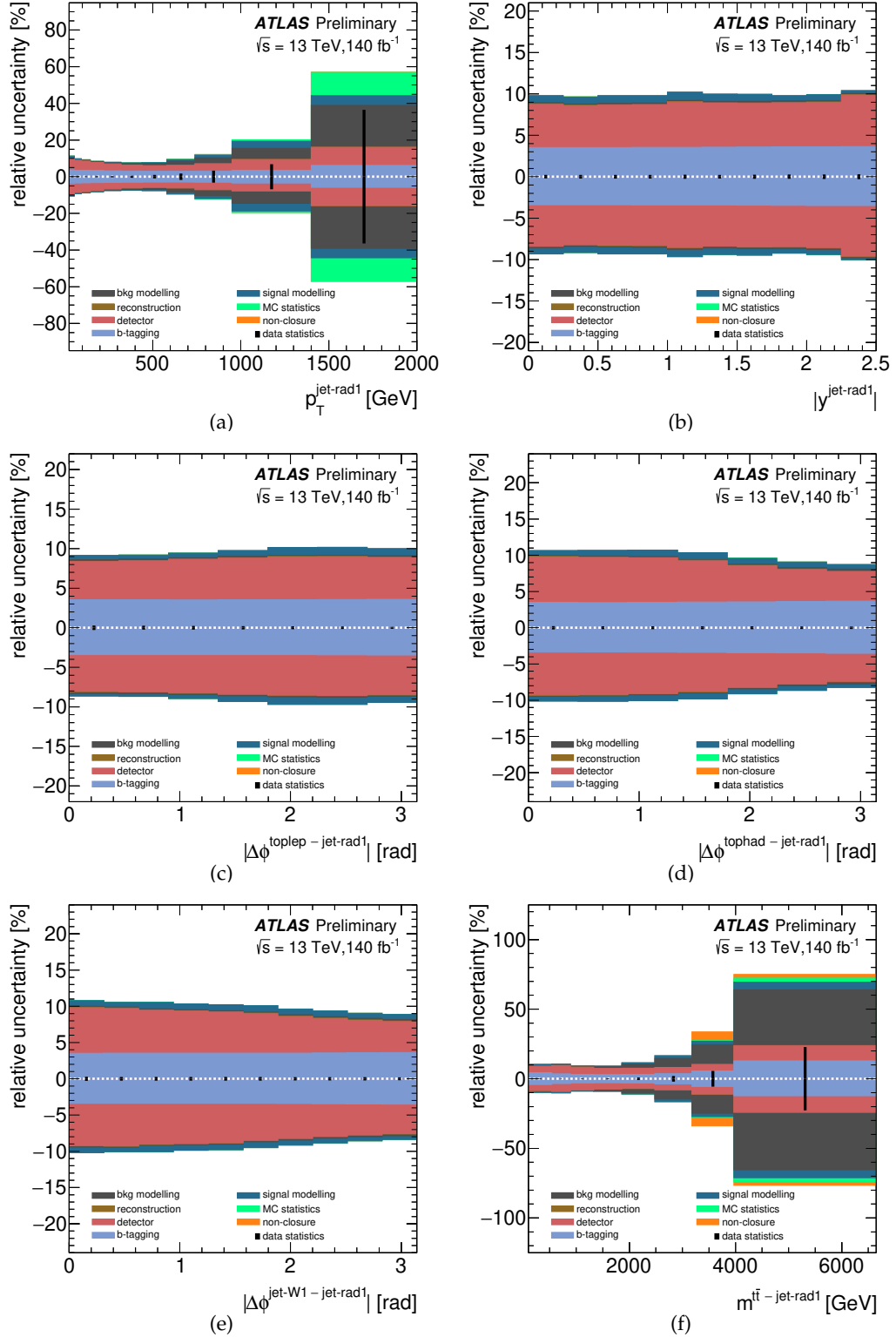


Figure 9.7: Total relative systematic uncertainties in percentage on the measured cross sections for (a) $p_T^{\text{jet-rad1}}$, (b) $|y^{\text{jet-rad1}}|$, (c) $|\Delta\phi^{\text{tolep-jet-rad1}}|$, (d) $|\Delta\phi^{\text{tophad-jet-rad1}}|$, (e) $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$, and (f) $m^{\text{t}\bar{\text{t}}-\text{jet-rad1}}$ in the $t\bar{t}+1\text{jet}$ channel. The individual sources of systematic uncertainty (plotted in stacked form) and the data statistical uncertainty are also included.

9.7 Systematic uncertainties

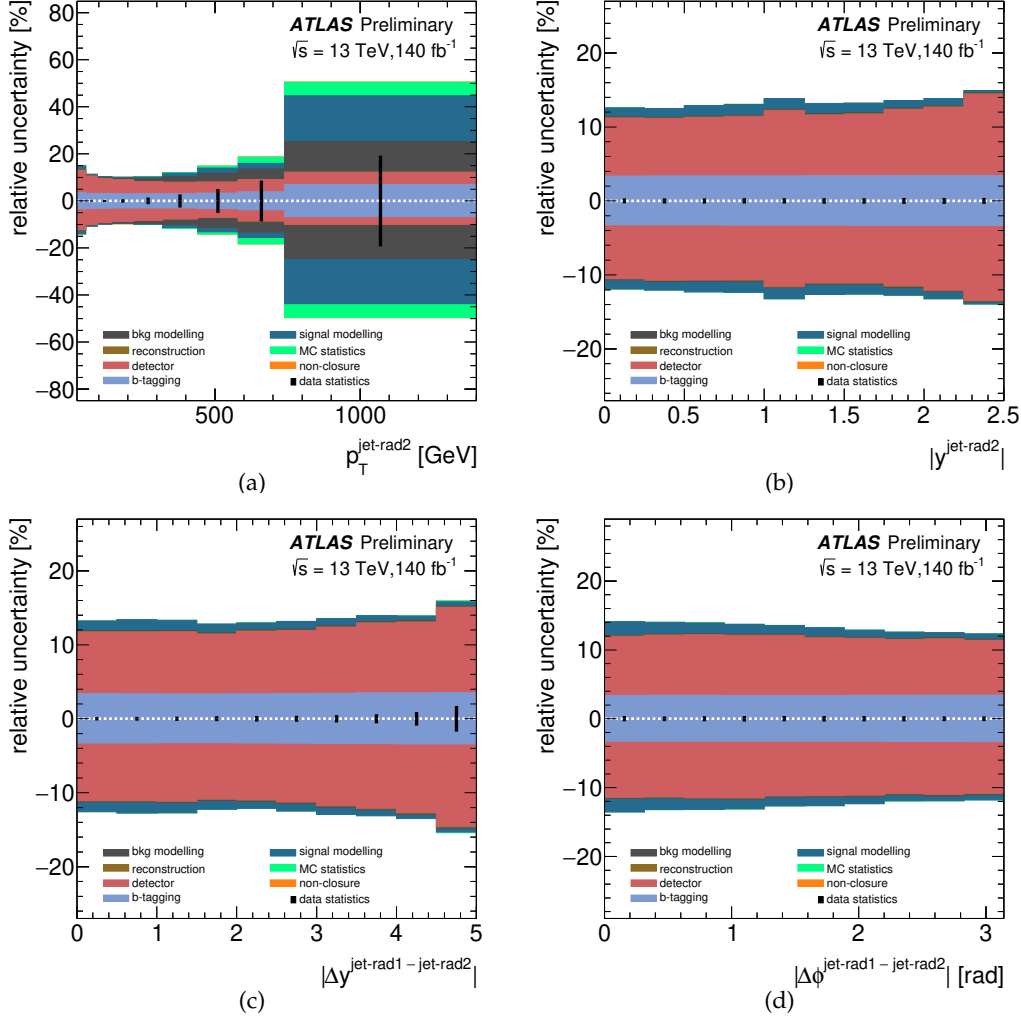


Figure 9.8: Total relative systematic uncertainties in percentage on the measured cross sections for (a) $p_T^{\text{jet-rad2}}$, (b) $|y^{\text{jet-rad2}}|$, (c) $|\Delta y^{\text{jet-rad1} - \text{jet-rad2}}|$, and (d) $|\Delta\phi^{\text{jet-rad1} - \text{jet-rad2}}|$ in the $t\bar{t}+2\text{jets}$ channel. The individual sources of systematic uncertainty (plotted in stacked form) and the data statistical uncertainty are also included.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

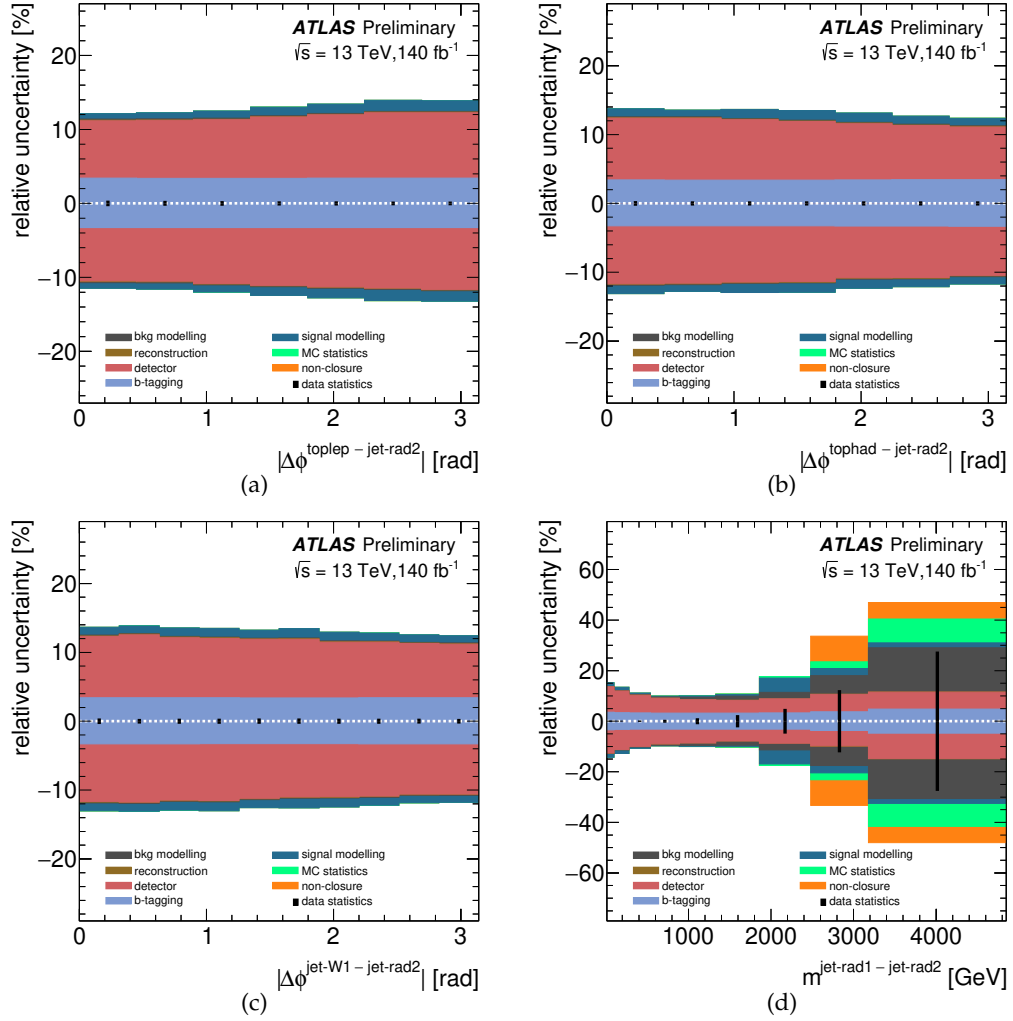


Figure 9.9: Total relative systematic uncertainties in percentage on the measured cross sections for (a) $|\Delta\phi^{\text{toptop}} - \text{jet-rad2}|$, (b) $|\Delta\phi^{\text{tophad}} - \text{jet-rad2}|$, (c) $|\Delta\phi^{\text{jet-W1}} - \text{jet-rad2}|$, and (d) $m^{\text{jet-rad1}} - \text{jet-rad2}$ in the $t\bar{t}+2\text{jets}$ channel. The individual sources of systematic uncertainty (plotted in stacked form) and the data statistical uncertainty are also included.

9.7 Systematic uncertainties

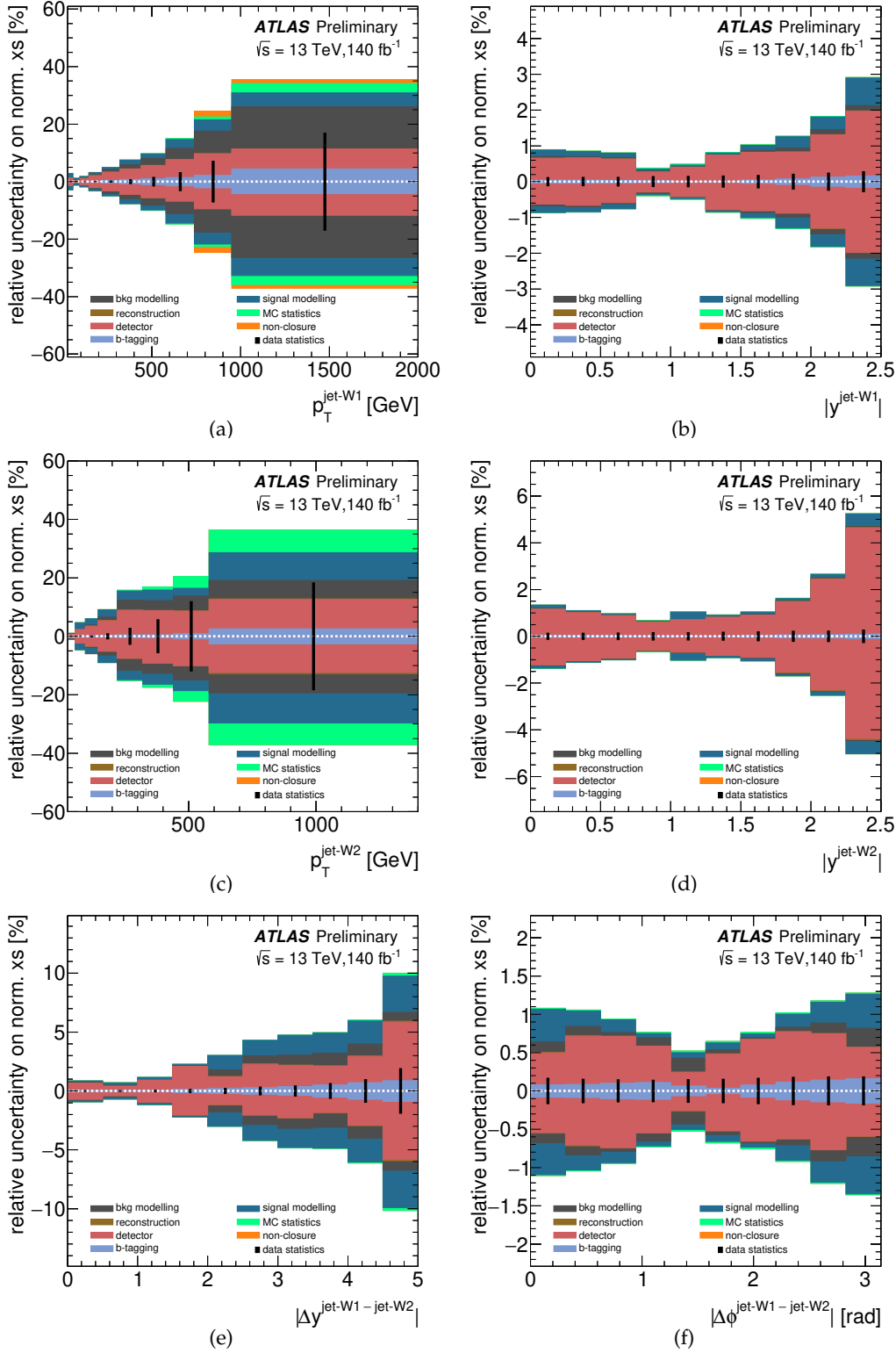


Figure 9.10: Total relative systematic uncertainties in percentage on the measured normalised cross sections for (a) $p_T^{\text{jet-W1}}$, (b) $|y^{\text{jet-W1}}|$, (c) $p_T^{\text{jet-W2}}$, (d) $|y^{\text{jet-W2}}|$, (e) $|\Delta y^{\text{jet-W1} - \text{jet-W2}}|$, and (f) $|\Delta\phi^{\text{jet-W1} - \text{jet-W2}}|$ in the $t\bar{t}$ inclusive channel. The individual sources of systematic uncertainty (plotted in stacked form) and the data statistical uncertainty are also included.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in $\nu\nu$ collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

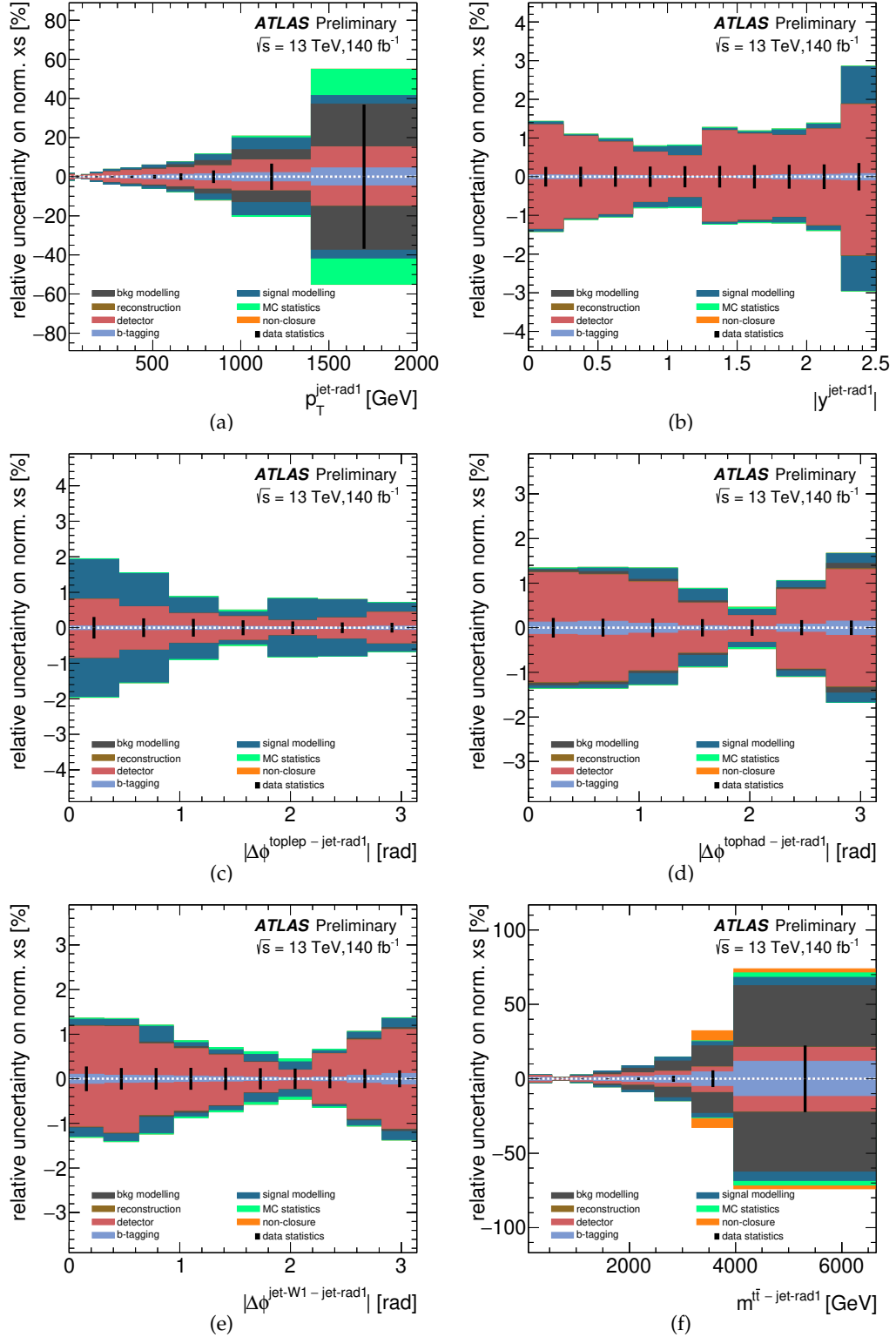


Figure 9.11: Total relative systematic uncertainties in percentage on the measured normalised cross sections for (a) $p_T^{\text{jet-rad1}}$, (b) $|y^{\text{jet-rad1}}|$, (c) $|\Delta\phi^{\text{tolep-jet-rad1}}|$, (d) $|\Delta\phi^{\text{tophad-jet-rad1}}|$, (e) $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$, and (f) $m^{t\bar{t}-\text{jet-rad1}}$ in the $t\bar{t}+1\text{jet}$ channel. The individual sources of systematic uncertainty (plotted in stacked form) and the data statistical uncertainty are also included.

9.7 Systematic uncertainties

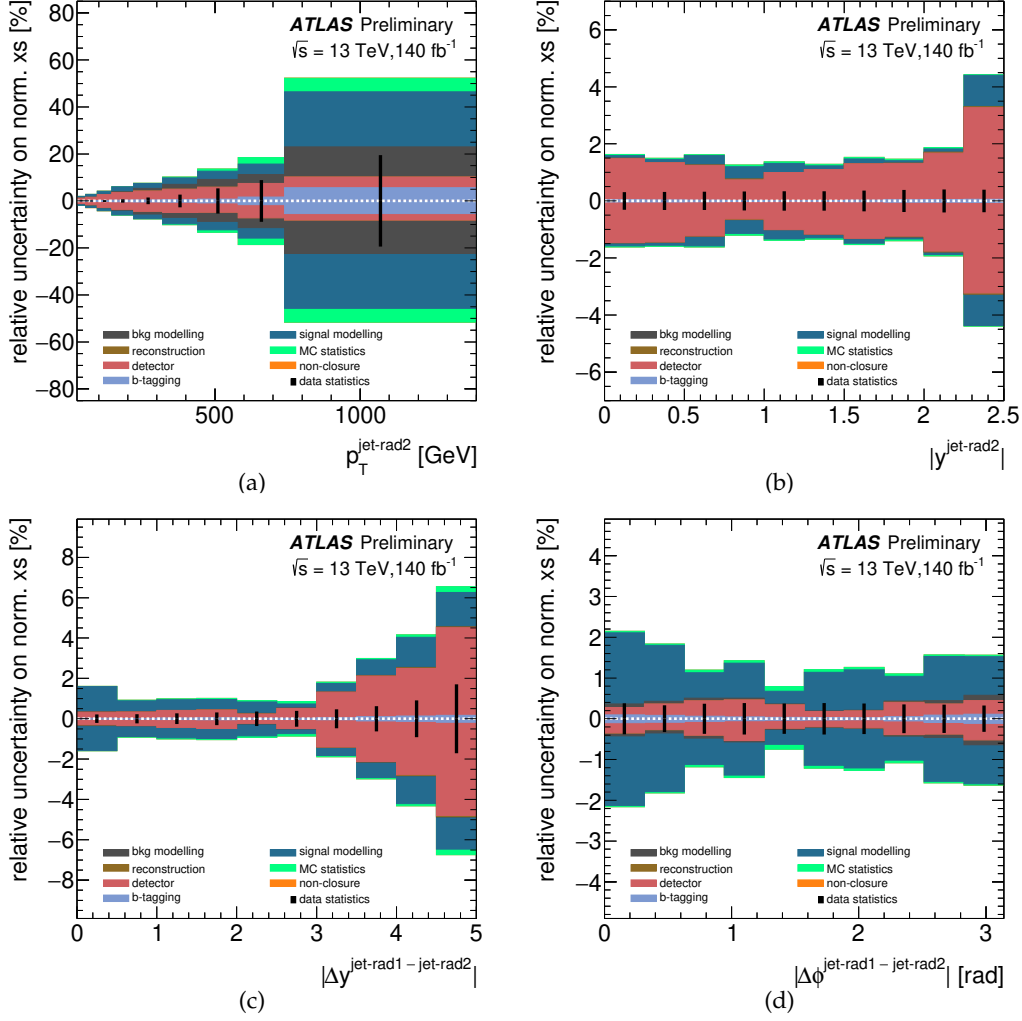


Figure 9.12: Total relative systematic uncertainties in percentage on the measured normalised cross sections for (a) $p_T^{\text{jet-rad2}}$, (b) $|y^{\text{jet-rad2}}|$, (c) $|\Delta y^{\text{jet-rad1-jet-rad2}}|$, and (d) $|\Delta\phi^{\text{jet-rad1-jet-rad2}}|$ in the $t\bar{t}+2\text{jets}$ channel. The individual sources of systematic uncertainty (plotted in stacked form) and the data statistical uncertainty are also included.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

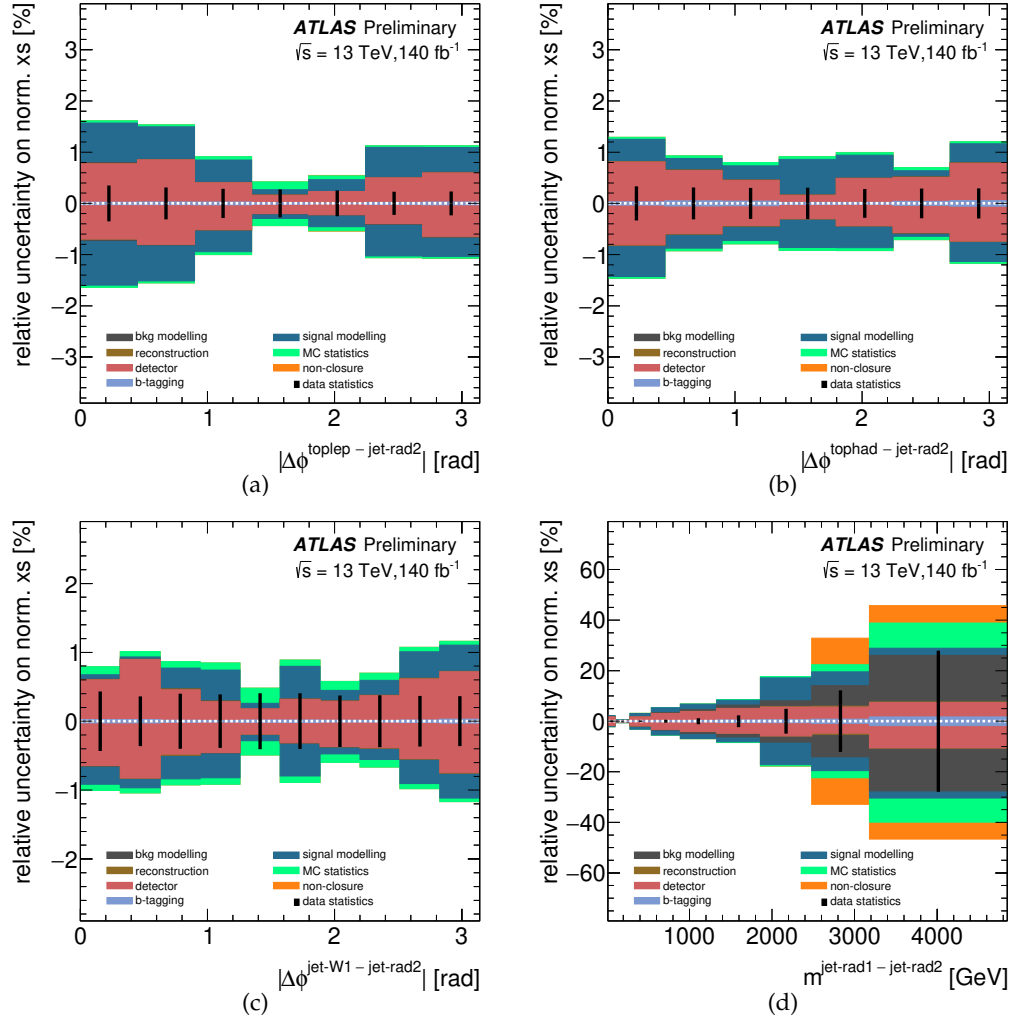


Figure 9.13: Total relative systematic uncertainties in percentage on the measured normalised cross sections for (a) $|\Delta\phi^{\text{toptop} - \text{jet-rad2}}|$, (b) $|\Delta\phi^{\text{tophad} - \text{jet-rad2}}|$, (c) $|\Delta\phi^{\text{jet-W1} - \text{jet-rad2}}|$, and (d) $m^{\text{jet-rad1} - \text{jet-rad2}}$ in the $t\bar{t}+2\text{jets}$ channel. The individual sources of systematic uncertainty (plotted in stacked form) and the data statistical uncertainty are also included.

9.8 Results

The measured absolute and normalised differential cross sections for $t\bar{t}$ production at particle level in the $t\bar{t}$ inclusive, the $t\bar{t}+1\text{jet}$ and the $t\bar{t}+2\text{jets}$ channels are presented in Sections 9.8.1, 9.8.2 and 9.8.3, respectively. NLO QCD predictions from PWG+PY8, PWG+HW7, aMC@NLO+HW7 and SHERPA 2.2.12 and NNLO QCD predictions from PWG+PY8 MiNNLOPS are compared to the measurements.

9.8.1 Differential cross sections for the $t\bar{t}$ inclusive channel

Figures 9.14 and 9.15 show the $t\bar{t}$ absolute and normalised, respectively, differential cross sections in the $\ell+\text{jets}$ decay mode for the $t\bar{t}$ inclusive channel as functions of $p_T^{\text{jet-W1}}$, $|y^{\text{jet-W1}}|$, $p_T^{\text{jet-W2}}$, $|y^{\text{jet-W2}}|$, $|\Delta y^{\text{jet-W1-jet-W2}}|$ and $|\Delta\phi^{\text{jet-W1-jet-W2}}|$. Values of $p_T^{\text{jet-W1}}$ and $p_T^{\text{jet-W2}}$ up to ~ 1.5 TeV and ~ 1.1 TeV, respectively, are measured. The measured cross section as a function of $p_T^{\text{jet-W1}}$ ($p_T^{\text{jet-W2}}$) decreases by more than five (six) orders of magnitude within the measured range. The measured cross section as function of $p_T^{\text{jet-W1}}$ has a harder spectrum than that of $p_T^{\text{jet-W2}}$, which decreases more rapidly with increasing $p_T^{\text{jet-W2}}$. The measured cross sections as functions of $|y^{\text{jet-W1}}|$ and $|y^{\text{jet-W2}}|$ decrease as the rapidity increases and have very similar shape and normalisation. The measured cross section as a function of $|\Delta y^{\text{jet-W1-jet-W2}}|$ ($|\Delta\phi^{\text{jet-W1-jet-W2}}|$) exhibits a peak around $|\Delta y^{\text{jet-W1-jet-W2}}| = 0$ ($|\Delta\phi^{\text{jet-W1-jet-W2}}| = 1$ rad).

The predictions of PWG+PY8, aMC@NLO+HW7, PWG+HW7 and SHERPA 2.2.12 are compared to the measurements in Figure 9.14. As explained in Section 9.1.1, these predictions are calculated at NLO, however, the normalisation is performed at $\mathcal{O}(\text{NNLO}+\text{NNLL})$. The $|y^{\text{jet-W1}}|$, $|y^{\text{jet-W2}}|$, $|\Delta y^{\text{jet-W1-jet-W2}}|$ and $|\Delta\phi^{\text{jet-W1-jet-W2}}|$ measured cross sections are well described in shape and normalisation by all the predictions, though the prediction from aMC@NLO+HW7 has a tendency to be above the data for $|\Delta y^{\text{jet-W1-jet-W2}}| > 2$. The measured cross section as a function of $p_T^{\text{jet-W1}}$ is described by the predictions of PWG+PY8 and PWG+HW7 only for $p_T^{\text{jet-W1}} < 500$ GeV, whereas aMC@NLO+HW7 and SHERPA 2.2.12 provide a good description of the data in the whole measured range. The measured cross section as a function of $p_T^{\text{jet-W2}}$ is well described by the NLO predictions; the differences between the measurement and the prediction from aMC@NLO+HW7 are within the theoretical and experimental uncertainties. From the comparison of the normalised differential predictions to the measurements (see Figure 9.15), the same conclusions are drawn.

The NNLO predictions of PWG+PY8 MiNNLOPS are compared to the measured absolute and normalised differential cross sections in Figures 9.16 and 9.17, respectively. For comparison, the predictions of PWG+PY8 are also included in these figures. For all observables, the theoretical uncertainty in the NNLO calculations is smaller than for the NLO predictions. The angular observables are well described by the NNLO predictions within the experimental and theoretical uncertainties, though there is a tendency of the NNLO predictions

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

to be below the data for $|\Delta y^{\text{jet-W1-jet-W2}}|$, which is clearly observed in the comparison of the normalised cross section (Figure 9.17(e)). A significant improvement in the shape and normalisation of the differential cross section as a function of $p_{\text{T}}^{\text{jet-W1}}$ is observed.

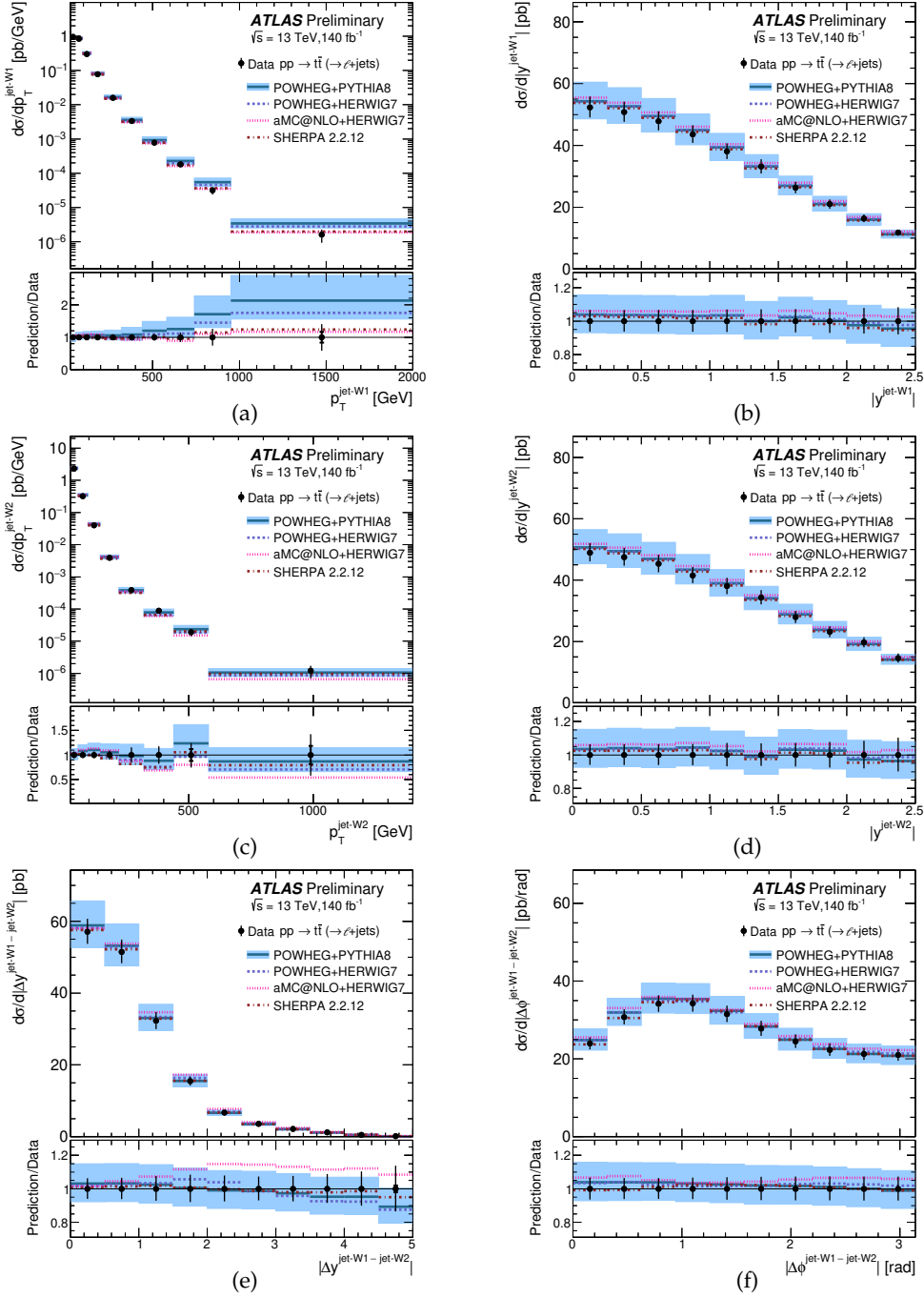


Figure 9.14: The measured differential cross sections for $t\bar{t}$ production in the $\ell + \text{jets}$ decay mode as functions of (a) $p_T^{\text{jet-W1}}$, (b) $|y|^{\text{jet-W1}}$, (c) $p_T^{\text{jet-W2}}$, (d) $|y|^{\text{jet-W2}}$, (e) $|\Delta y|^{\text{jet-W1-jet-W2}}$, and (f) $|\Delta\phi|^{\text{jet-W1-jet-W2}}$ (dots) in the $t\bar{t}$ inclusive channel. The NLO QCD predictions from PWG+PY8 (solid lines), PWG+HW7 (dashed lines), aMC@NLO+HW7 (dotted lines) and SHERPA 2.2.12 (dot-dashed lines) normalised to $\mathcal{O}(\text{NNLO}+\text{NNLL})$ are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded bands represent the theoretical uncertainty on the PWG+PY8 prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

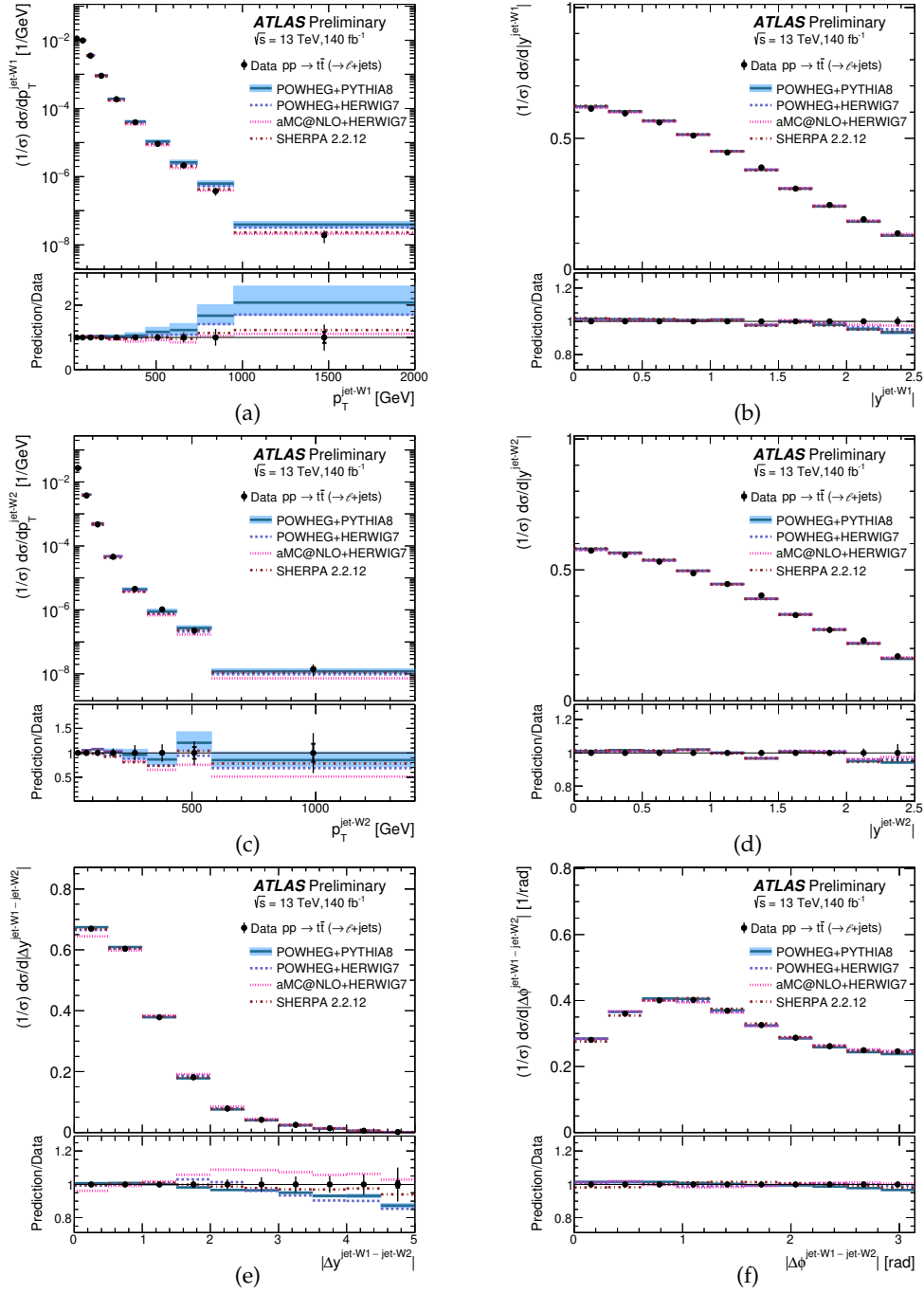


Figure 9.15: The measured normalised differential cross sections for $t\bar{t}$ production in the ℓ +jets channel as functions of (a) $p_T^{\text{jet-W1}}$, (b) $|y^{\text{jet-W1}}|$, (c) $p_T^{\text{jet-W2}}$, (d) $|y^{\text{jet-W2}}|$, (e) $|\Delta y^{\text{jet-W1} - \text{jet-W2}}|$, and (f) $|\Delta\phi^{\text{jet-W1} - \text{jet-W2}}|$ (dots) in the $t\bar{t}$ inclusive channel. The NLO QCD predictions from PWG+PY8 (solid lines), PWG+HW7 (dashed lines), aMC@NLO+HW7 (dotted lines) and SHERPA 2.2.12 (dot-dashed lines) are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded bands represent the theoretical uncertainty on the PWG+PY8 prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

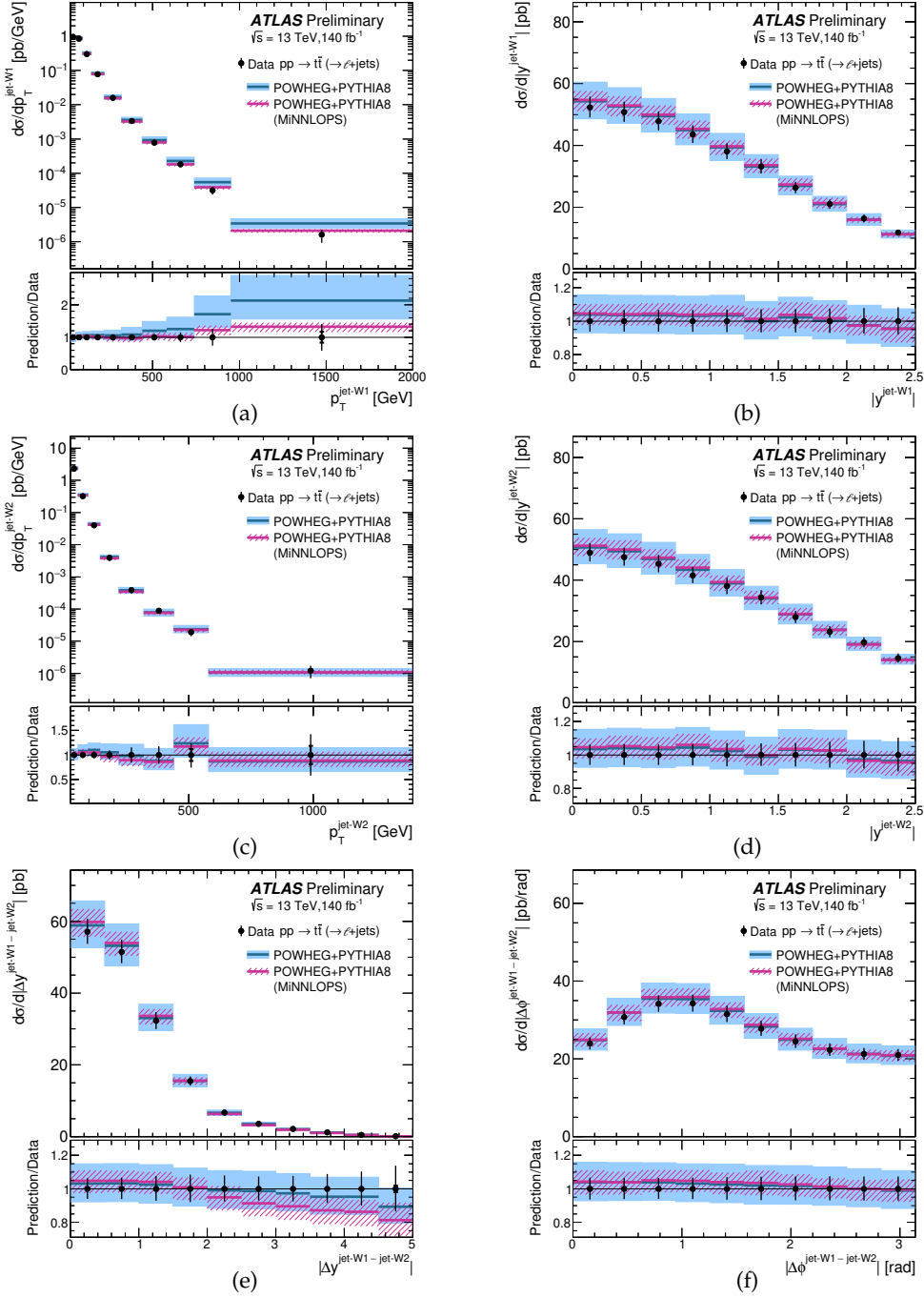


Figure 9.16: The measured differential cross sections for $t\bar{t}$ production in the $\ell + \text{jets}$ decay mode as functions of (a) $p_T^{\text{jet-W1}}$, (b) $|y^{\text{jet-W1}}|$, (c) $p_T^{\text{jet-W2}}$, (d) $|y^{\text{jet-W2}}|$, (e) $|\Delta y^{\text{jet-W1-jet-W2}}|$, and (f) $|\Delta\phi^{\text{jet-W1-jet-W2}}|$ (dots) in the $t\bar{t}$ inclusive channel. The NLO QCD predictions from PWG+PY8 (blue lines) and the NNLO QCD predictions from PWG+PY8 MINNLOPS (purple lines) are also shown. Both predictions are normalised to the $\mathcal{O}(\text{NNLO}+\text{NNLL})$ total cross section. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded (hatched) bands represent the theoretical uncertainty on the PWG+PY8 (PWG+PY8 MINNLOPS) prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

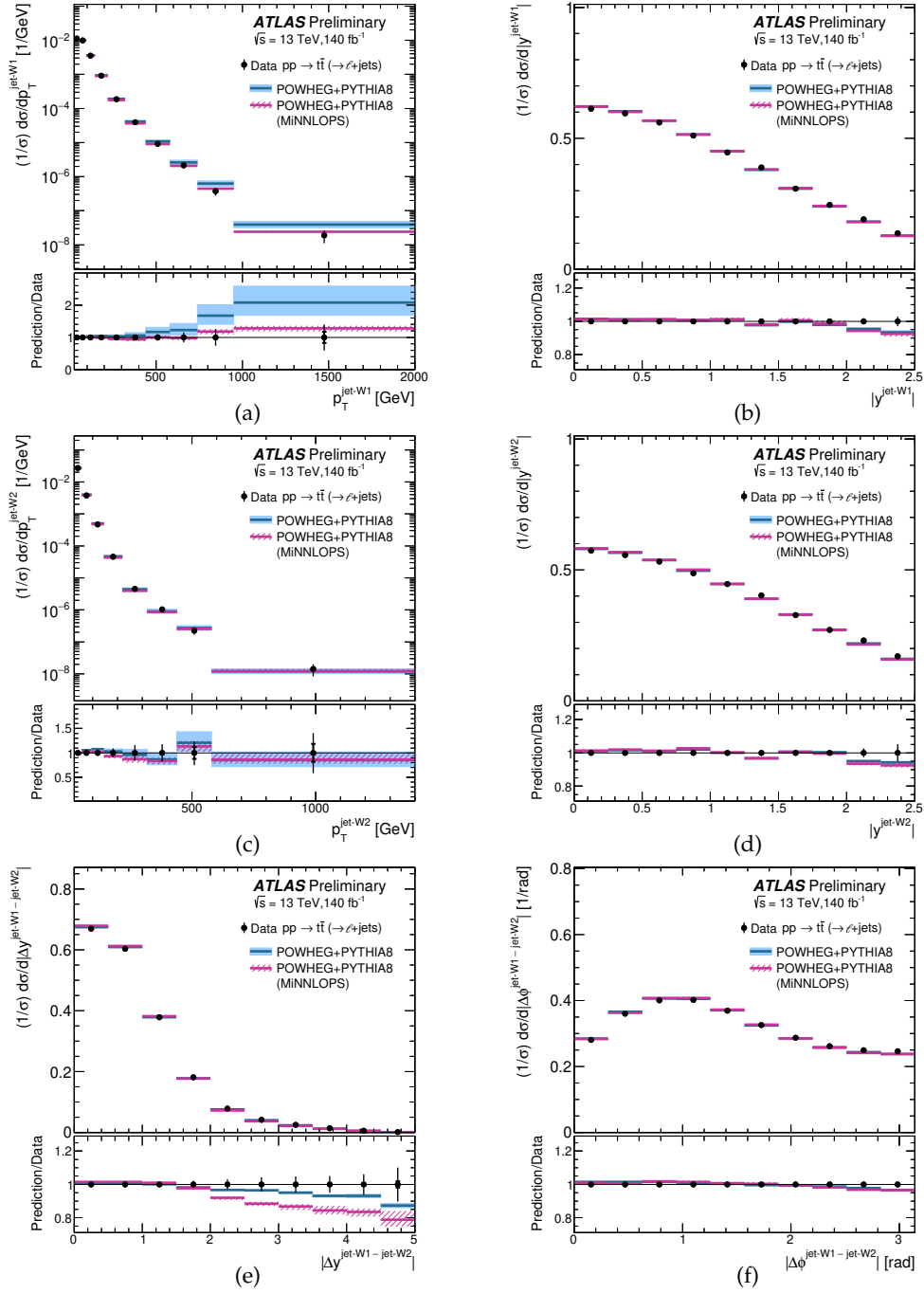


Figure 9.17: The measured normalised differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $p_T^{\text{jet-W1}}$, (b) $|y^{\text{jet-W1}}|$, (c) $p_T^{\text{jet-W2}}$, (d) $|y^{\text{jet-W2}}|$, (e) $|\Delta y^{\text{jet-W1-jet-W2}}|$, and (f) $|\Delta\phi^{\text{jet-W1-jet-W2}}|$ (dots) in the $t\bar{t}$ inclusive channel. The NLO QCD predictions from PWG+PY8 (blue lines) and the NNLO QCD predictions from PWG+PY8 MINNLOPS (purple lines) are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded (hatched) bands represent the theoretical uncertainty on the PWG+PY8 (PWG+PY8 MINNLOPS) prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

9.8.2 Differential cross sections for the $t\bar{t}+1\text{jet}$ channel

Figures 9.18 and 9.19 show the $t\bar{t}$ differential absolute and normalised, respectively, cross sections in the $\ell+\text{jets}$ decay mode for the $t\bar{t}+1\text{jet}$ channel as functions of $p_T^{\text{jet-rad1}}$, $|y^{\text{jet-rad1}}|$, $|\Delta\phi^{\text{toplep-jet-rad1}}|$, $|\Delta\phi^{\text{tophad-jet-rad1}}|$, $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$ and $m^{t\bar{t}-\text{jet-rad1}}$. Values of $p_T^{\text{jet-rad1}}$ and $m^{t\bar{t}-\text{jet-rad1}}$ up to ~ 1.7 TeV and ~ 5.3 TeV, respectively, are measured. The measured cross section as a function of $p_T^{\text{jet-rad1}}$ ($m^{t\bar{t}-\text{jet-rad1}}$) decreases by more than six (five) orders of magnitude within the measured range. The measured cross section as a function of $m^{t\bar{t}-\text{jet-rad1}}$ decreases for $m^{t\bar{t}-\text{jet-rad1}} < 0.5$ TeV due to the kinematic constraints. The shape of the measured cross section as a function of $|y^{\text{jet-rad1}}|$ is different than that of $|y^{\text{jet-W1}}|$: the value of the cross section at rapidity of 2.5 with respect to the value at 0 is higher for $|y^{\text{jet-rad1}}|$ than for $|y^{\text{jet-W1}}|$, which indicates that the first emission tends to be more isotropic in rapidity than the jets forming the hadronically-decaying top. This is investigated further by measuring the cross sections as functions of $|\Delta\phi^{\text{toplep-jet-rad1}}|$ and $|\Delta\phi^{\text{tophad-jet-rad1}}|$: the measured cross sections have very different shape, which indicates that the first emission has a higher probability of arising far from the leptonically-decaying top. In addition, the measured cross section as a function of $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$ shows a peak at $|\Delta\phi^{\text{jet-W1-jet-rad1}}| \approx 0.5$ rad and at π rad, which indicates that there are more events with a gluon emission back-to-back with jet-W1, though a fraction of the events has an emission close to jet-W1.

The NLO QCD predictions describe well the measurements as functions of $|y^{\text{jet-rad1}}|$, $|\Delta\phi^{\text{toplep-jet-rad1}}|$, $|\Delta\phi^{\text{tophad-jet-rad1}}|$ and $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$. As is the case for the measurements in the $t\bar{t}$ inclusive channel, the cross section as a function of $p_T^{\text{jet-rad1}}$ is described by the predictions from PWG+PY8 and PWG+HW7 only for $p_T^{\text{jet-rad1}} < 500$ GeV, whereas the predictions from aMC@NLO+HW7 and SHERPA 2.2.12 provide a good description of the data in the whole measured range. The cross section as a function of $m^{t\bar{t}-\text{jet-rad1}}$ is well described by all predictions for $m^{t\bar{t}-\text{jet-rad1}} < 3$ TeV. The same conclusions can be drawn from the normalised differential cross sections in Figure 9.19.

A significant improvement by the NNLO predictions of PWG+PY8 MINNLOPS is observed for the description of the measured differential cross sections as functions of $p_T^{\text{jet-rad1}}$ and $m^{t\bar{t}-\text{jet-rad1}}$ with respect to the predictions of PWG+PY8, as seen in Figures 9.20 and 9.21.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

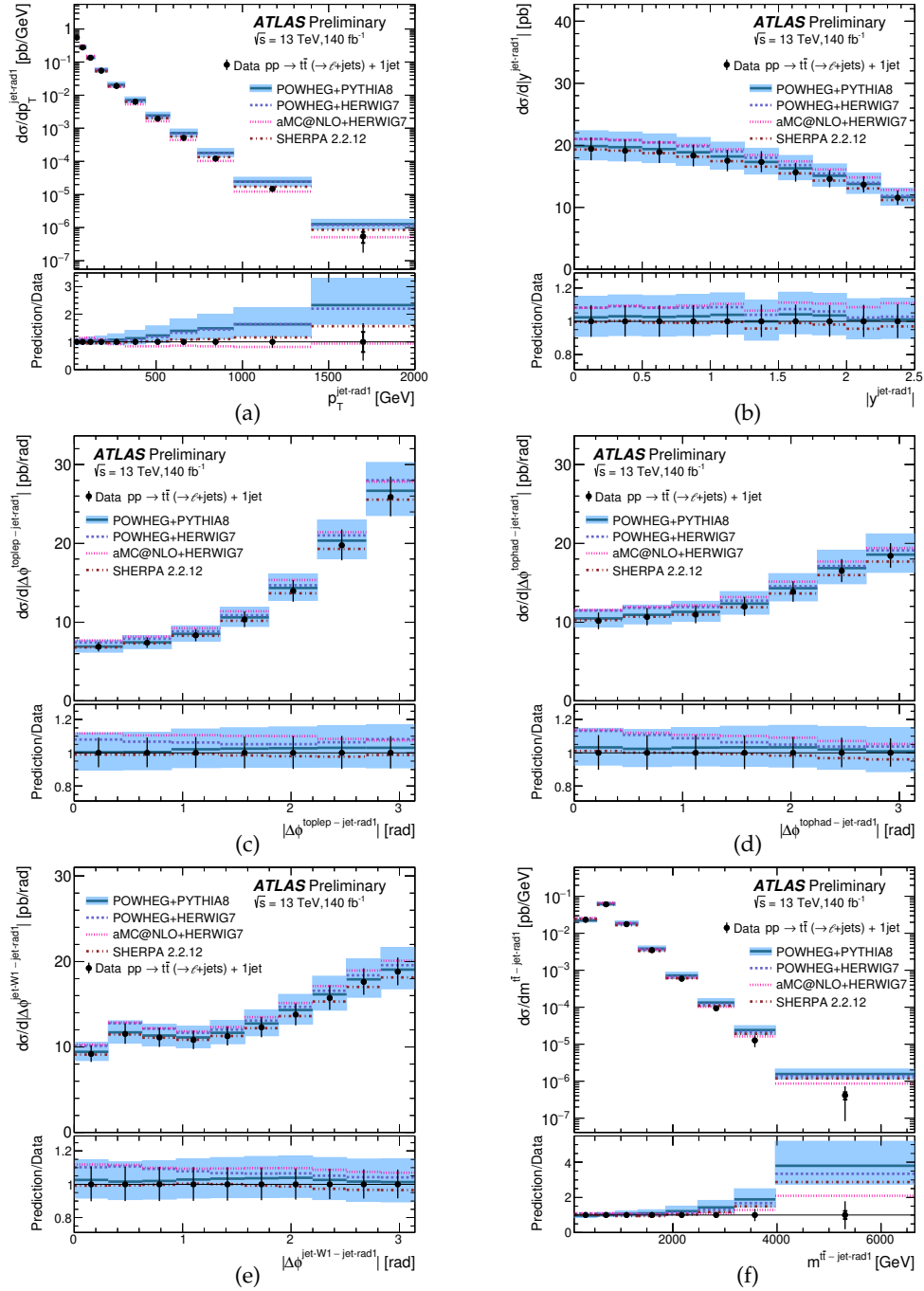


Figure 9.18: The measured differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $p_T^{\text{jet-rad1}}$, (b) $|y^{\text{jet-rad1}}|$, (c) $|\Delta\phi^{\text{toplep-jet-rad1}}|$, (d) $|\Delta\phi^{\text{tophad-jet-rad1}}|$, (e) $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$, and (f) $m^{\text{tt-jet-rad1}}$ (dots) in the $t\bar{t}$ +1jet channel. The NLO QCD predictions from PWG+PY8 (solid lines), PWG+HW7 (dashed lines), aMC@NLO+HW7 (dotted lines) and SHERPA 2.2.12 (dot-dashed lines) normalised to $\mathcal{O}(\text{NNLO}+\text{NNLL})$ are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded bands represent the theoretical uncertainty on the PWG+PY8 prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

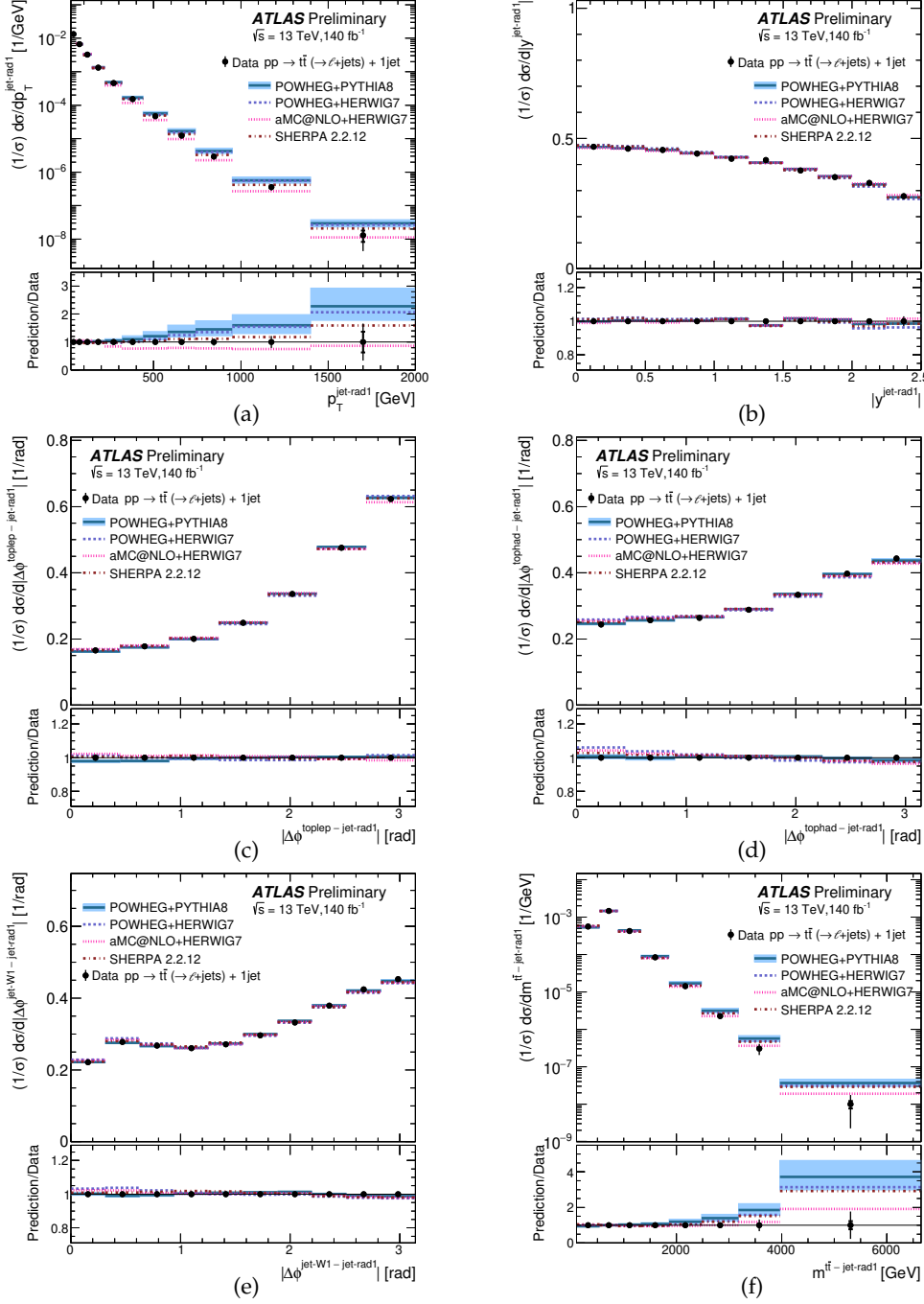


Figure 9.19: The measured normalised differential cross sections for $t\bar{t}$ production in the ℓ +jets channel as functions of (a) $p_T^{\text{jet-rad1}}$, (b) $|y^{\text{jet-rad1}}|$, (c) $|\Delta\phi^{\text{toplep-jet-rad1}}|$, (d) $|\Delta\phi^{\text{tophad-jet-rad1}}|$, (e) $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$, and (f) $m^{t\bar{t}-\text{jet-rad1}}$ (dots) in the $t\bar{t}$ +1jet channel. The NLO QCD predictions from PWG+PY8 (solid lines), PWG+HW7 (dashed lines), aMC@NLO+HW7 (dotted lines) and SHERPA 2.2.12 (dot-dashed lines) are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded bands represent the theoretical uncertainty on the PWG+PY8 prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

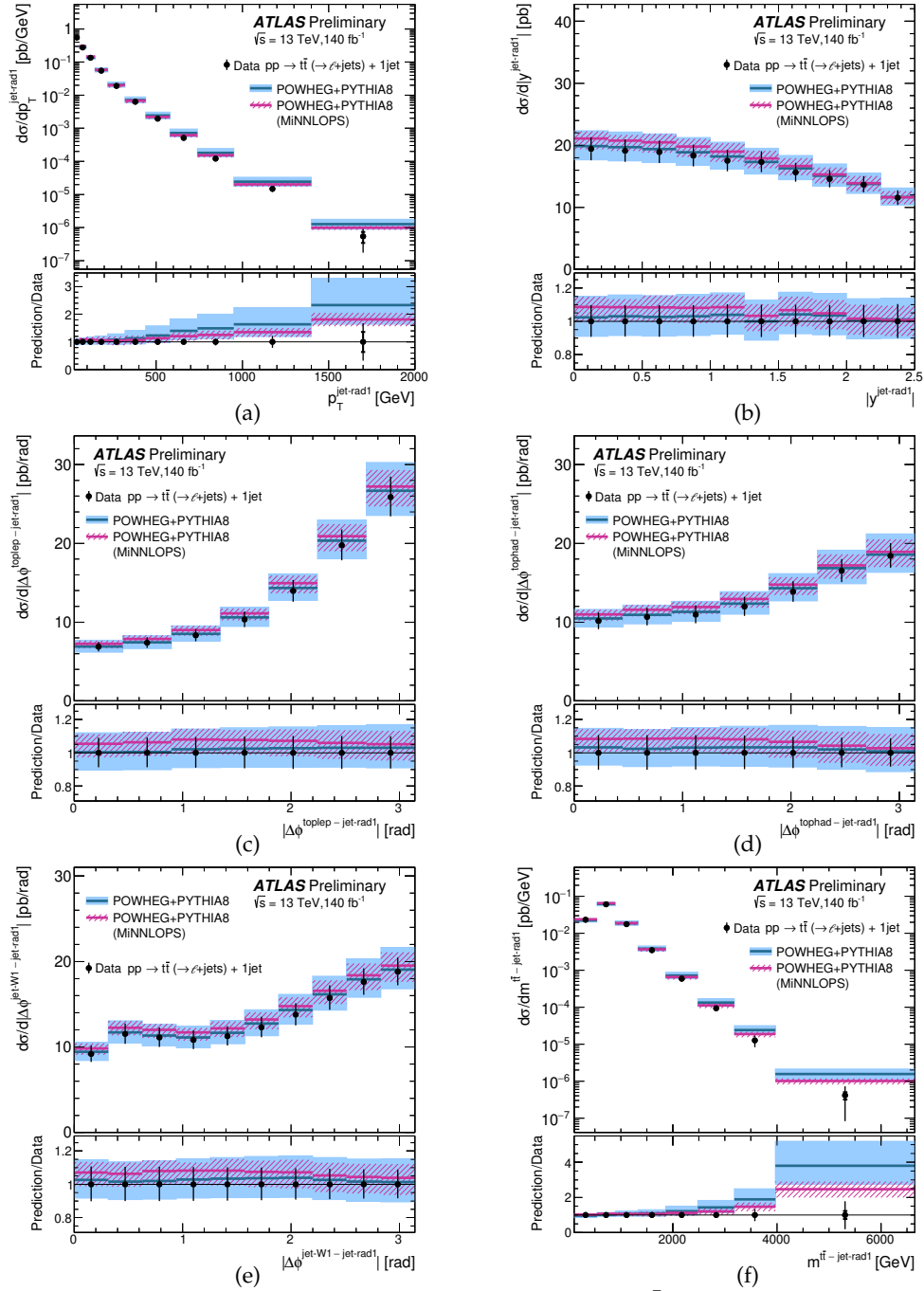


Figure 9.20: The measured differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $p_T^{\text{jet-rad1}}$, (b) $|y^{\text{jet-rad1}}|$, (c) $|\Delta\phi^{\text{toplep-jet-rad1}}|$, (d) $|\Delta\phi^{\text{tophad-jet-rad1}}|$, (e) $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$, and (f) $m^{\text{t}\bar{\text{t}}-\text{jet-rad1}}$ (dots) in the $t\bar{t}$ +1jet channel. The NLO QCD predictions from PWG+PY8 (blue lines) and the NNLO QCD predictions from PWG+PY8 MINNLOPS (purple lines) are also shown. Both predictions are normalised to the $\mathcal{O}(\text{NNLO}+\text{NNLL})$ total cross section. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded (hatched) bands represent the theoretical uncertainty on the POWHEG+PY8 (POWHEG+PY8 MINNLOPS) prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

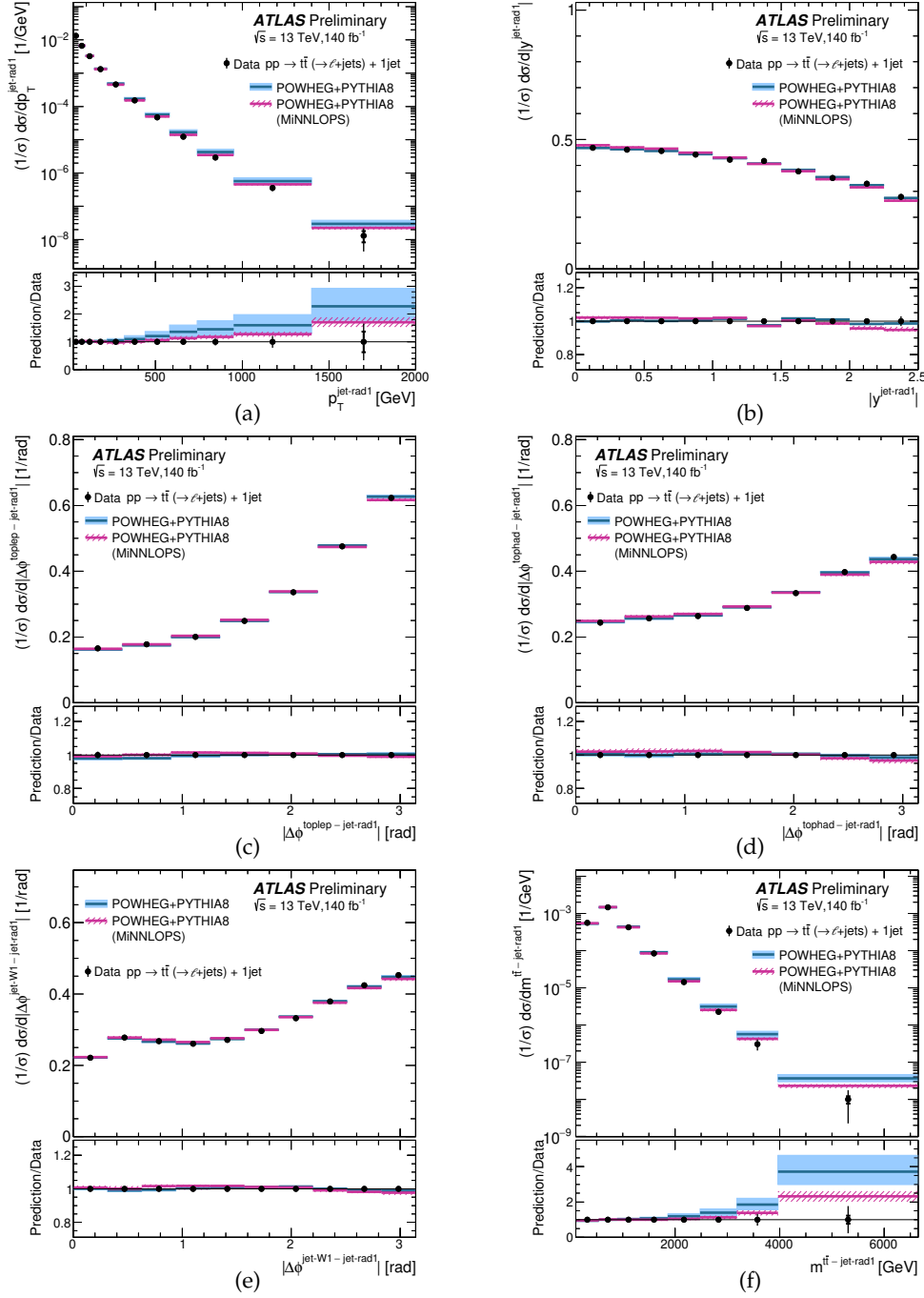


Figure 9.21: The measured normalised differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $p_T^{\text{jet-rad1}}$, (b) $|y^{\text{jet-rad1}}|$, (c) $|\Delta\phi^{\text{tolep-jet-rad1}}|$, (d) $|\Delta\phi^{\text{tophad-jet-rad1}}|$, (e) $|\Delta\phi^{\text{jet-W1-jet-rad1}}|$, and (f) $m^{\ell+jets-jet-rad1}$ (dots) in the $t\bar{t}$ +1jet channel. The NLO QCD predictions from PWG+PY8 (blue lines) and the NNLO QCD predictions from PWG+PY8 MINNLOPS (purple lines) are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded (hatched) bands represent the theoretical uncertainty on the PWG+PY8 (PWG+PY8 MINNLOPS) prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

9.8.3 Differential cross sections for the $t\bar{t}+2\text{jets}$ channel

Figures 9.22 and 9.23 (Figures 9.24 and 9.25) show the $t\bar{t}$ absolute (normalised) differential cross sections in the ℓ +jets decay mode for the $t\bar{t}+2\text{jets}$ channel as functions of $p_T^{\text{jet-rad2}}$, $|y^{\text{jet-rad2}}|$, $|\Delta y^{\text{jet-rad1-jet-rad2}}|$, $|\Delta\phi^{\text{jet-rad1-jet-rad2}}|$, $|\Delta\phi^{\text{toplep-jet-rad2}}|$, $|\Delta\phi^{\text{tophad-jet-rad2}}|$, $|\Delta\phi^{\text{jet-W1-jet-rad2}}|$ and $m^{\text{jet-rad1-jet-rad2}}$. Values of $p_T^{\text{jet-rad2}}$ and $m^{\text{jet-rad1-jet-rad2}}$ up to ~ 1.2 TeV and ~ 4 TeV, respectively, are measured. The measured cross sections as functions of $p_T^{\text{jet-rad2}}$ and $m^{\text{jet-rad1-jet-rad2}}$ decrease by more than five orders of magnitude within the measured range. The measured cross section as a function of $|y^{\text{jet-rad2}}|$ decreases as the rapidity increases and has a very similar shape to the one for $|y^{\text{jet-rad1}}|$. The measured cross section as a function of $|\Delta y^{\text{jet-rad1-jet-rad2}}|$ decreases rapidly with increasing $|\Delta y^{\text{jet-rad1-jet-rad2}}|$ and the $|\Delta\phi^{\text{jet-rad1-jet-rad2}}|$ distribution exhibits peaks at 0.5 rad and π rad, with approximately the same rate; these features indicate that the second hard emission tends to be close in rapidity to the first emission and that there is a weak dependence of the difference in ϕ between the first and second emissions. The decrease of the cross section at $|\Delta\phi^{\text{jet-rad1-jet-rad2}}| \approx 0$ is due to the fact that since the jets are reconstructed with $R = 0.4$, they cannot be closer than that value in ϕ if they have the same rapidity. The measured cross sections as functions of $|\Delta\phi^{\text{toplep-jet-rad2}}|$, $|\Delta\phi^{\text{tophad-jet-rad2}}|$ and $|\Delta\phi^{\text{jet-W1-jet-rad2}}|$ show similar features to those measured for jet-rad1; thus, the second emission also has a higher probability of arising far from the leptonically-decaying top.

The NLO QCD predictions describe well the measurements as functions of $|y^{\text{jet-rad2}}|$, $|\Delta y^{\text{jet-rad1-jet-rad2}}|$, $|\Delta\phi^{\text{jet-rad1-jet-rad2}}|$, $|\Delta\phi^{\text{toplep-jet-rad2}}|$, $|\Delta\phi^{\text{tophad-jet-rad2}}|$, and $|\Delta\phi^{\text{jet-W1-jet-rad2}}|$, though the prediction from aMC@NLO+HW7 has a somewhat different shape than the data for $|\Delta y^{\text{jet-rad1-jet-rad2}}|$ (see Figures 9.22(c) and 9.24(c)). The measurement as a function of $p_T^{\text{jet-rad2}}$ is described adequately by the predictions of PWG+PY8, PWG+HW7 and SHERPA 2.2.12, whereas the prediction from aMC@NLO+HW7 fails. For $m^{\text{jet-rad1-jet-rad2}}$, the predictions describe the data only for $m^{\text{jet-rad1-jet-rad2}} < 1$ TeV; for larger $m^{\text{jet-rad1-jet-rad2}}$ values, the predictions of PWG+PY8, PWG+HW7 and SHERPA 2.2.12 (aMC@NLO+HW7) overestimate (underestimate) the data.

The comparison of the NNLO predictions to the measurements in the $t\bar{t}+2\text{jets}$ channel is shown in Figures 9.26 to 9.29. The central value of the NNLO predictions for the angular variables in this channel has a tendency to be above the data, though the PWG+PY8 MINNLOPS predictions are compatible with the measurements within the experimental and theoretical uncertainties, except for the normalised cross section as a function of $|\Delta y^{\text{jet-rad1-jet-rad2}}|$, for which the PWG+PY8 MINNLOPS prediction is below the data for $|\Delta y^{\text{jet-rad1-jet-rad2}}| > 3$. The description of the data by the NNLO prediction is somewhat worse than that at NLO for the $p_T^{\text{jet-rad2}}$ observable (see Figures 9.26(a) and 9.28(a)), especially at high $p_T^{\text{jet-rad2}}$ values. This effect might be attributed to the fact that these calculations are only lowest order in QCD for the second emission; thus, higher-order corrections are modelled with the parton

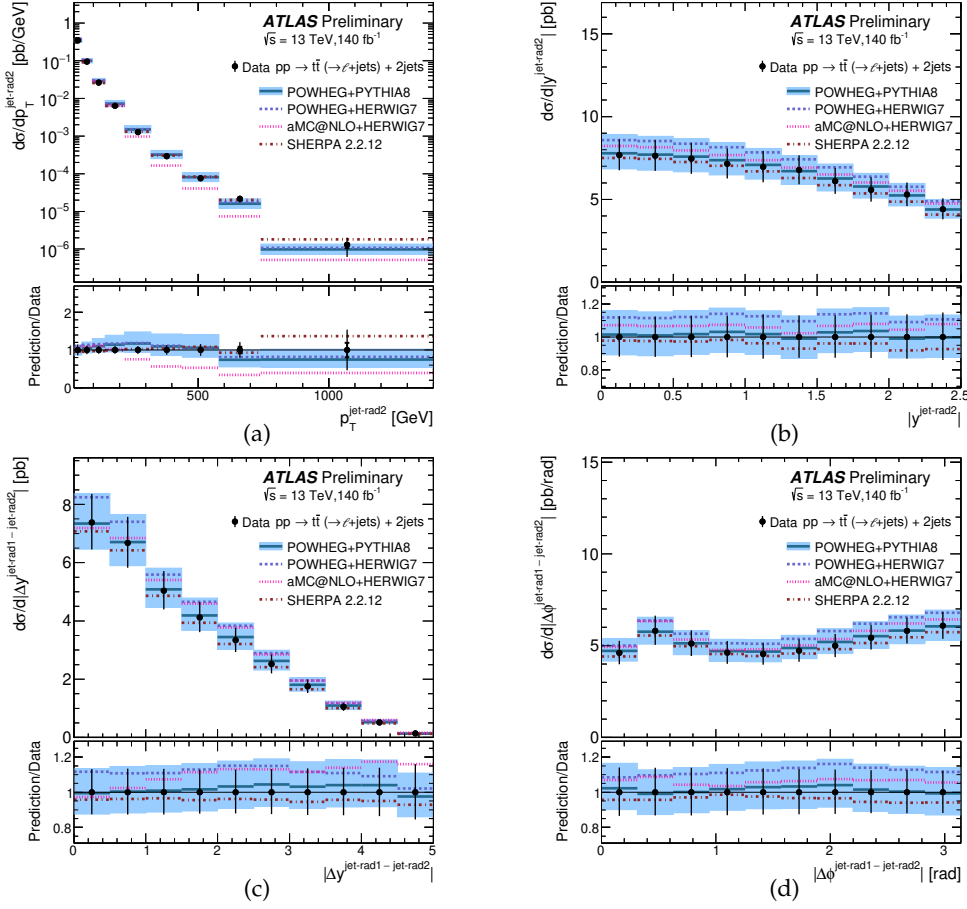


Figure 9.22: The measured differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $p_T^{\text{jet-rad2}}$, (b) $|y^{\text{jet-rad2}}|$, (c) $|\Delta y^{\text{jet-rad1}} - \text{jet-rad2}|$, and (d) $|\Delta\phi^{\text{jet-rad1}} - \text{jet-rad2}|$ (dots) in the $t\bar{t}$ +2jet channel. The NLO QCD predictions from PWG+PY8 (solid lines), PWG+HW7 (dashed lines), aMC@NLO+HW7 (dotted lines) and SHERPA 2.2.12 (dot-dashed lines) normalised to $\mathcal{O}(\text{NNLO}+\text{NNLL})$ are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded bands represent the theoretical uncertainty on the PWG+PY8 prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

shower and the theoretical uncertainty due to the scale variations might be underestimated since it does not account for the scale variations in the parton shower.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

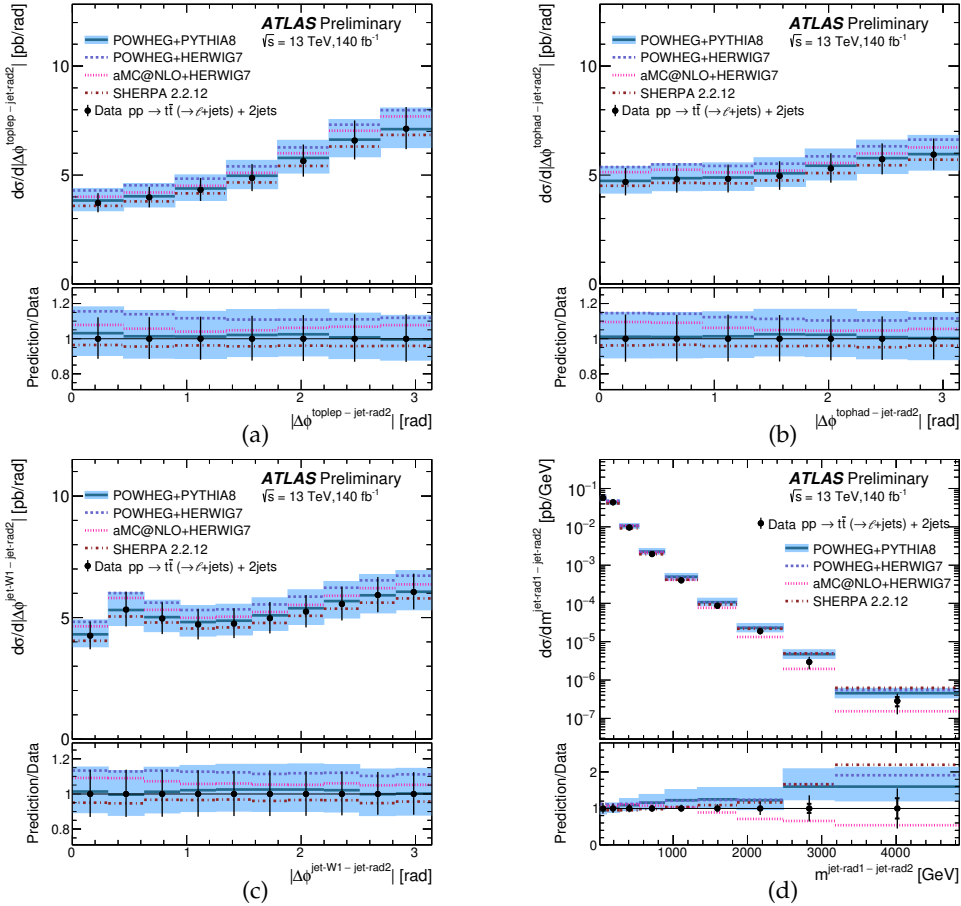


Figure 9.23: The measured differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $|\Delta\phi^{\text{toplep}} - \text{jet-rad2}|$, (b) $|\Delta\phi^{\text{tophad}} - \text{jet-rad2}|$, (c) $|\Delta\phi^{\text{jet-W1}} - \text{jet-rad2}|$, and (d) $m^{\text{jet-rad1}} - \text{jet-rad2}$ (dots) in the $t\bar{t}+2\text{jet}$ channel. The NLO QCD predictions from PWG+PY8 (solid lines), PWG+HW7 (dashed lines), aMC@NLO+HW7 (dotted lines) and SHERPA 2.2.12 (dot-dashed lines) normalised to $\mathcal{O}(\text{NNLO}+\text{NNLL})$ are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded bands represent the theoretical uncertainty on the PWG+PY8 prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

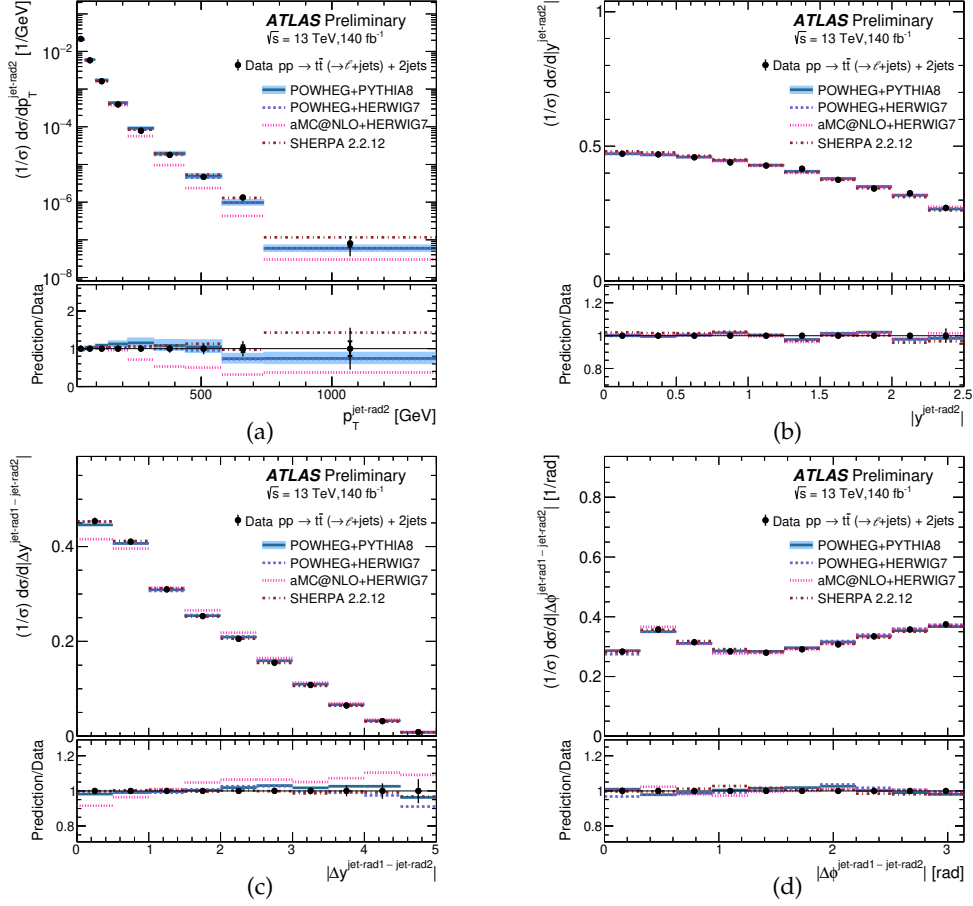


Figure 9.24: The measured normalised differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $p_T^{\text{jet-rad2}}$, (b) $|y^{\text{jet-rad2}}|$, (c) $|\Delta y^{\text{jet-rad1-jet-rad2}}|$, and (d) $|\Delta\phi^{\text{jet-rad1-jet-rad2}}|$ (dots) in the $t\bar{t}$ +2jet channel. The NLO QCD predictions from PWG+PY8 (solid lines), PWG+HW7 (dashed lines), aMC@NLO+HW7 (dotted lines) and SHERPA 2.2.12 (dot-dashed lines) normalised to $\mathcal{O}(\text{NNLO}+\text{NNLL})$ are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded bands represent the theoretical uncertainty on the PWG+PY8 prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

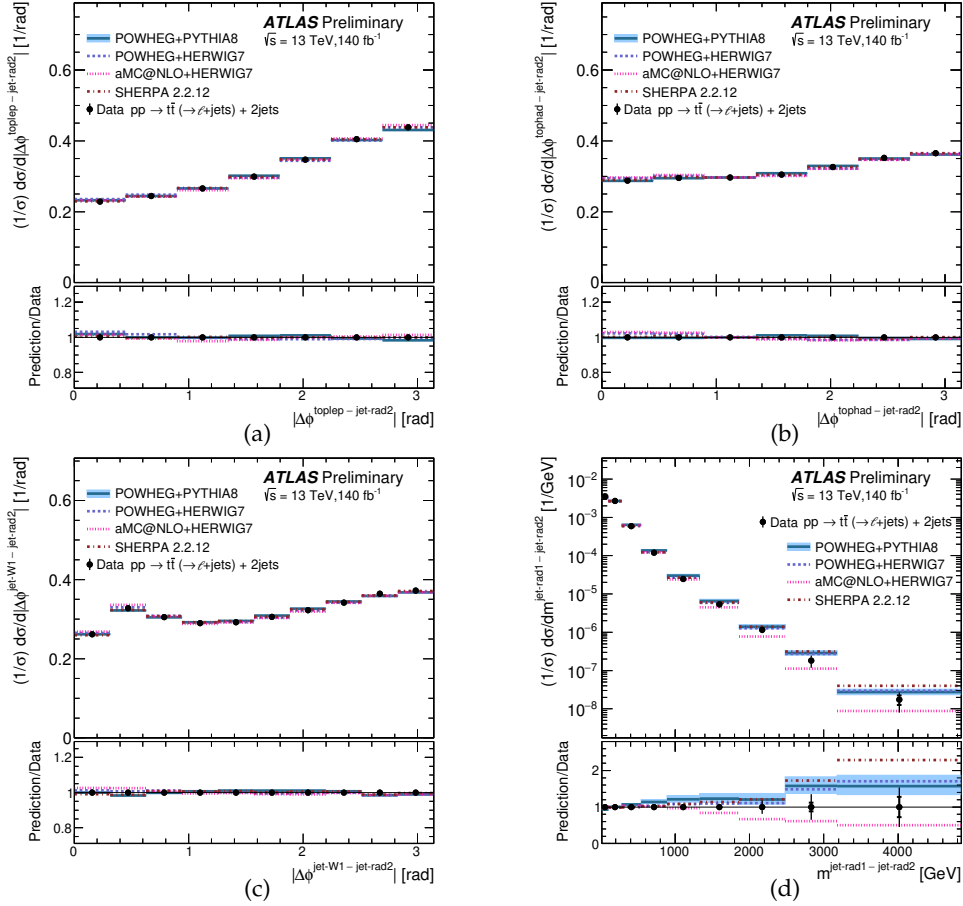


Figure 9.25: The measured normalised differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $|\Delta\phi^{\text{toplep-jet-rad2}}|$, (b) $|\Delta\phi^{\text{tophad-jet-rad2}}|$, (c) $|\Delta\phi^{\text{jet-W1-jet-rad2}}|$, and (d) $m^{\text{jet-rad1-jet-rad2}}$ (dots) in the $t\bar{t}+2\text{jet}$ channel. The NLO QCD predictions from PWG+PY8 (solid lines), PWG+HW7 (dashed lines), aMC@NLO+HW7 (dotted lines) and SHERPA 2.2.12 (dot-dashed lines) normalised to $\mathcal{O}(\text{NNLO}+\text{NNLL})$ are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded bands represent the theoretical uncertainty on the PWG+PY8 prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

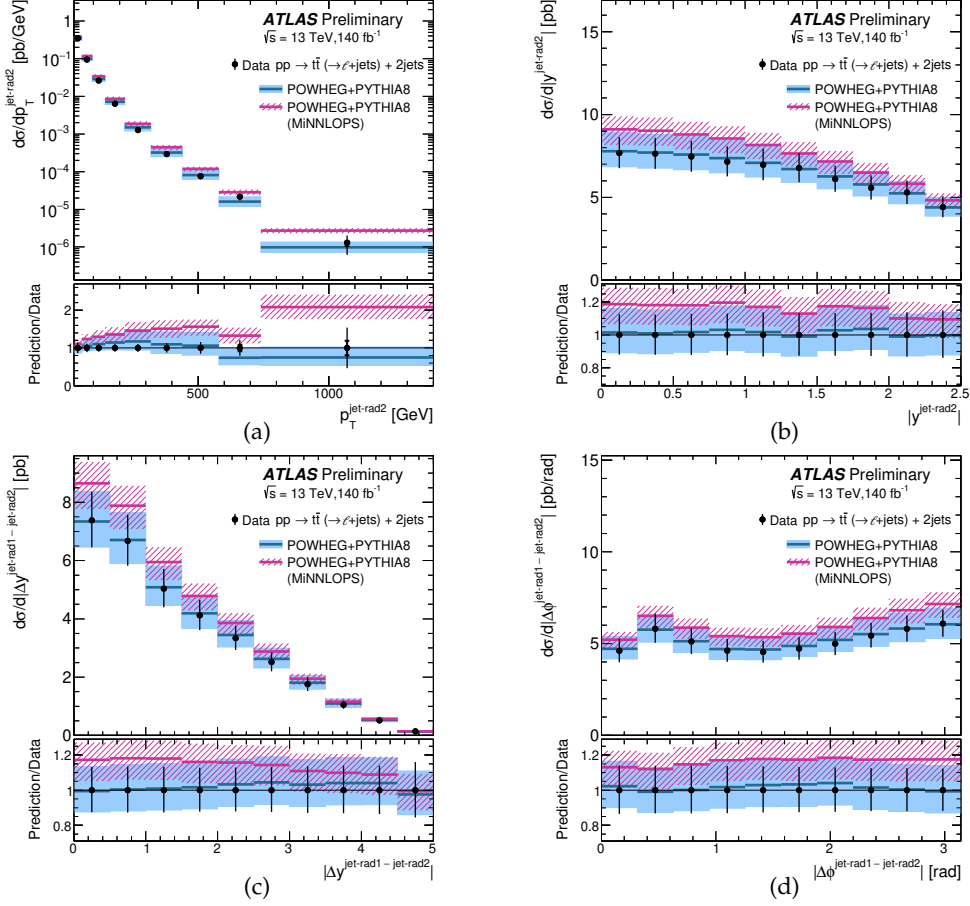


Figure 9.26: The measured differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $p_T^{\text{jet-rad2}}$, (b) $|y^{\text{jet-rad2}}|$, (c) $|\Delta y^{\text{jet-rad1} - \text{jet-rad2}}|$, and (d) $|\Delta\phi^{\text{jet-rad1} - \text{jet-rad2}}|$ (dots) in the $t\bar{t}$ +2jet channel. The NLO QCD predictions from PWG+PY8 (blue lines) and the NNLO QCD predictions from PWG+PY8 MINNLOPS (purple lines) are also shown. Both predictions are normalised to the $\mathcal{O}(\text{NNLO}+\text{NNLL})$ total cross section. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded (hatched) bands represent the theoretical uncertainty on the PWG+PY8 (PWG+PY8 MINNLOPS) prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

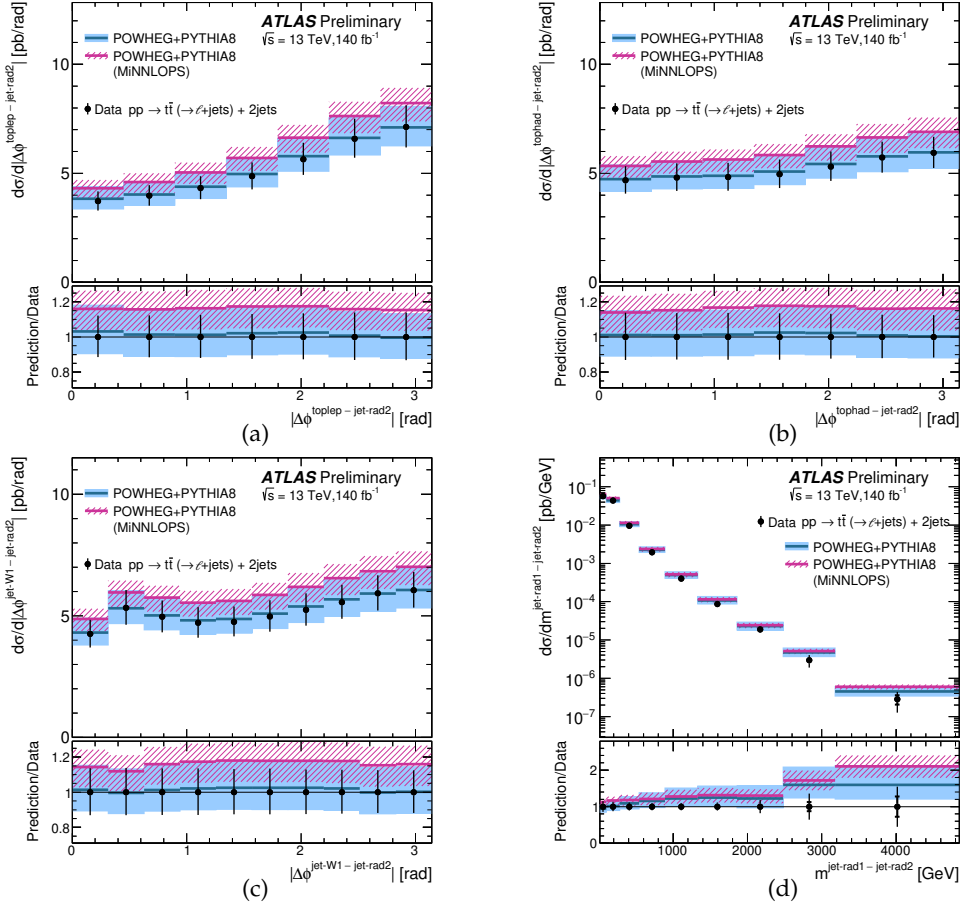


Figure 9.27: The measured differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $|\Delta\phi^{\text{toplep}} - \text{jet-rad2}|$, (b) $|\Delta\phi^{\text{tophad}} - \text{jet-rad2}|$, (c) $|\Delta\phi^{\text{jet-W1}} - \text{jet-rad2}|$, and (d) $m^{\text{jet-rad1} - \text{jet-rad2}}$ (dots) in the $t\bar{t}+2\text{jet}$ channel. The NLO QCD predictions from PWG+PY8 (blue lines) and the NNLO QCD predictions from PWG+PY8 MINNLOPS (purple lines) are also shown. Both predictions are normalised to the $\mathcal{O}(\text{NNLO}+\text{NNLL})$ total cross section. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded (hatched) bands represent the theoretical uncertainty on the PWG+PY8 (PWG+PY8 MINNLOPS) prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

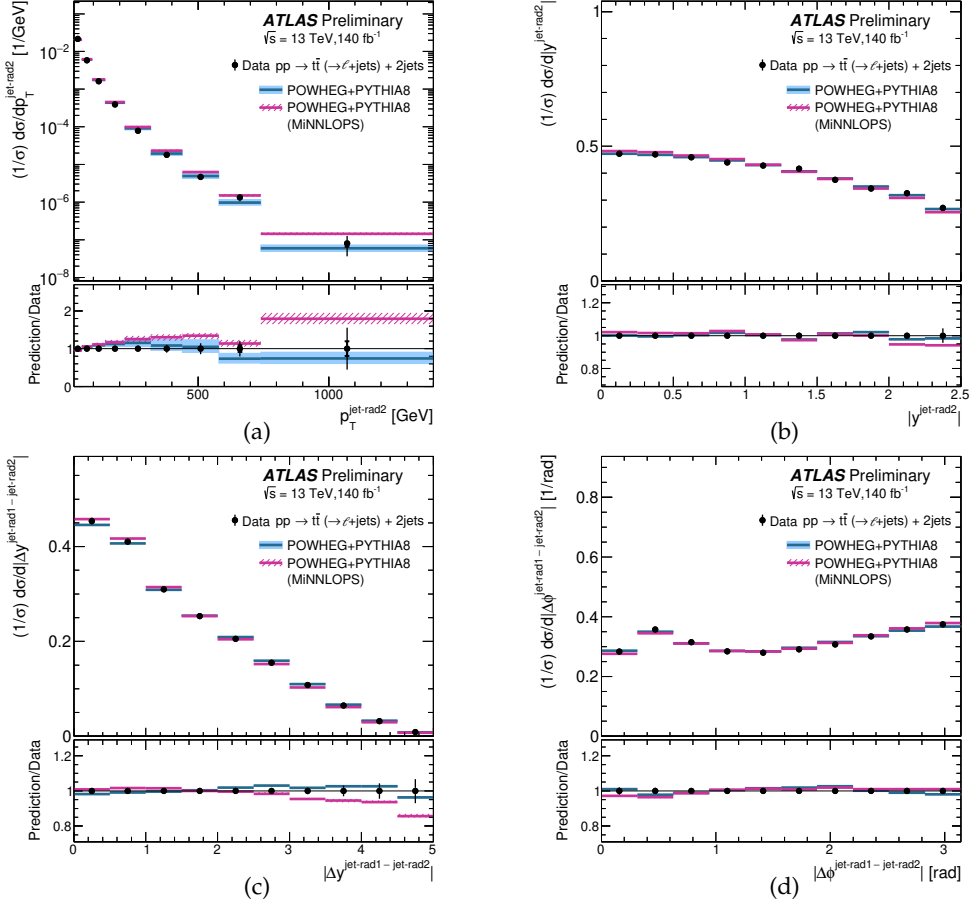


Figure 9.28: The measured normalised differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $p_T^{\text{jet-rad2}}$, (b) $|y^{\text{jet-rad2}}|$, (c) $|\Delta y^{\text{jet-rad1-jet-rad2}}|$, and (d) $|\Delta\phi^{\text{jet-rad1-jet-rad2}}|$ (dots) in the $t\bar{t}$ +2jet channel. The NLO QCD predictions from PWG+PY8 (blue lines) and the NNLO QCD predictions from PWG+PY8 MINNLOPS (purple lines) are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded (hatched) bands represent the theoretical uncertainty on the PWG+PY8 (PWG+PY8 MINNLOPS) prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

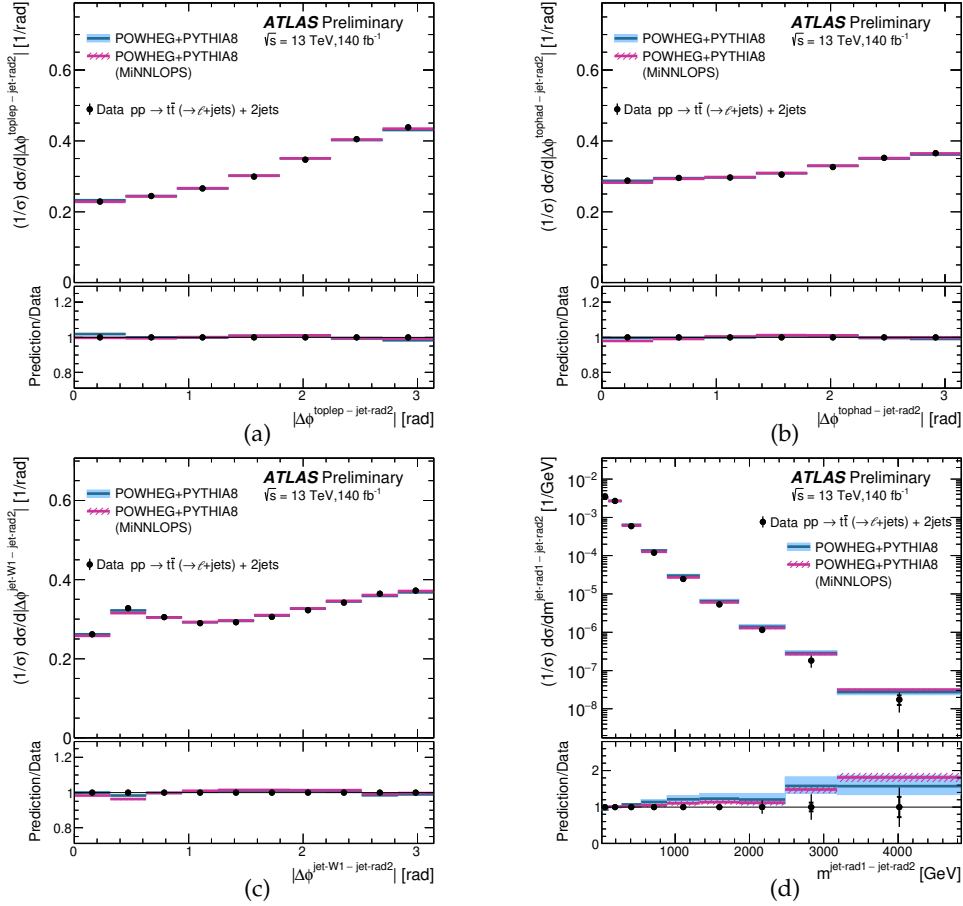


Figure 9.29: The measured normalised differential cross sections for $t\bar{t}$ production in the ℓ +jets decay mode as functions of (a) $|\Delta\phi^{\text{toplep}} - \text{jet-rad2}|$, (b) $|\Delta\phi^{\text{tophad}} - \text{jet-rad2}|$, (c) $|\Delta\phi^{\text{jet-W1}} - \text{jet-rad2}|$, and (d) $m^{\text{jet-rad1} - \text{jet-rad2}}$ (dots) in the $t\bar{t}$ +2jet channel. The NLO QCD predictions from PWG+PY8 (blue lines) and the NNLO QCD predictions from PWG+PY8 MINNLOPS (purple lines) are also shown. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible. The shaded (hatched) bands represent the theoretical uncertainty on the PWG+PY8 (PWG+PY8 MINNLOPS) prediction. The lower part of the figures show the ratio of the predicted and the measured cross sections.

9.8.4 χ^2 and p -values

To quantify the level of agreement between the measured differential cross sections and the theoretical predictions, χ^2 values are evaluated using the following equation [248]

$$\chi^2 = \sum_{i,k} \left[D_i - T_i \left(1 - \sum_j \gamma_{ij} b_j \right) \right] C_{\text{stat,uncor},ik}^{-1} \left[D_k - T_k \left(1 - \sum_j \gamma_{kj} b_j \right) \right] + \sum_i \log \frac{\delta_{i,\text{uncor}}^2 T_i^2 + \delta_{i,\text{stat}}^2 D_i T_i}{\delta_{i,\text{uncor}}^2 D_i^2 + \delta_{i,\text{stat}}^2 D_i^2} + \sum_j b_j^2, \quad (9.1)$$

known as the Hessian method [249].

In Equation 9.1, indices i and k represent different bins in a differential cross section and index j represents each source of uncertainty. The symbol D represents the measured differential cross section in a given bin as a function of each observable. The statistical uncertainties are incorporated in the covariance matrix $C_{\text{stat,uncor}}$. This matrix contains also the systematic uncertainty due to non-closure (see Section 9.7), which is the only experimental uncertainty considered uncorrelated between bins. All the other systematic uncertainties, denoted by the symbol γ in Equation 9.1, are considered as correlated between bins and are accounted for by using nuisance parameters (symbols b in Equation 9.1). The parameters b are obtained from a minimisation using the MINUIT [250] program.

The symbol T represents the theoretical prediction in each bin of each differential cross section. The uncertainties that affect the theoretical predictions (see Section 9.1.4) are taken into account: the uncertainties arising from μ_R , μ_F and the amount of initial- and final-state radiation in the parton shower are considered as uncorrelated whereas those arising from the PDFs and the value of α_s are considered as correlated bin to bin. As for the experimental uncertainties, the uncorrelated uncertainties are incorporated into $C_{\text{stat,uncor}}$ and the correlated uncertainties are accounted for by using the nuisance parameters b .

The first term in Equation 9.1 gives the main contribution to the χ^2 and the second term is a small bias correction term, referred to as the *log penalty* [249], which arises because the diagonal term of $C_{\text{stat,uncor}}$ is given by $C_{\text{stat,uncor},ii} = \delta_{i,\text{uncor}}^2 T_i^2 + \delta_{i,\text{stat}}^2 D_i T_i$, with different weights for the statistical ($\delta_{i,\text{stat}}$) and uncorrelated ($\delta_{i,\text{uncor}}$) uncertainties.

Table 9.3 shows the χ^2 and p -values for each differential cross section and for each theoretical prediction tested separately. The p -values obtained confirm the qualitative conclusions discussed in the previous sections.

9 Measurements of top-quark pair differential cross sections in the ℓ +jets channel
in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data

Table 9.3: χ^2 values, number of degrees of freedom and p -values for each observable and each prediction for the absolute differential cross sections.

Prediction	PWG+PY8		PWG+HW7		aMC@NLO+HW7		SHERPA 2.2.12		PWG+PY8 MiNNLOPS	
Observable	χ^2/NDF	p -value	χ^2/NDF	p -value	χ^2/NDF	p -value	χ^2/NDF	p -value	χ^2/NDF	p -value
$p_{\text{T}}^{\text{jet-W1}}$	6.2/10	0.79	4.1/10	0.94	2.7/10	0.99	1.8/10	1.0	3.4/10	0.97
$ y^{\text{jet-W1}} $	1.8/10	1.0	1.6/10	1.0	2.3/10	0.99	1.1/10	1.0	2.7/10	0.99
$p_{\text{T}}^{\text{jet-W2}}$	2.3/8	0.97	0.45/8	1.0	2.8/8	0.94	0.53/8	1.0	2.6/8	0.96
$ y^{\text{jet-W2}} $	1.9/10	1.0	1.7/10	1.0	2.4/10	0.99	1.1/10	1.0	2.7/10	0.99
$ \Delta y^{\text{jet-W1} - \text{jet-W2}} $	1.6/10	1.0	2.7/10	0.99	4.3/10	0.93	0.9/10	1.0	6.0/10	0.81
$ \Delta \phi^{\text{jet-W1} - \text{jet-W2}} $	1.7/10	1.0	1.6/10	1.0	2.5/10	0.99	1.2/10	1.0	2.1/10	1.0
$p_{\text{T}}^{\text{jet-rad1}}$	7.7/11	0.74	7.1/11	0.79	3.3/11	0.99	2.6/11	1.0	6.4/11	0.85
$ y^{\text{jet-rad1}} $	1.6/10	1.0	2.8/10	0.99	3.4/10	0.97	0.69/10	1.0	3.0/10	0.98
$ \Delta \phi^{\text{tolep} - \text{jet-rad1}} $	1.4/7	0.99	2.2/7	0.95	2.9/7	0.89	0.81/7	1.0	2.3/7	0.94
$ \Delta \phi^{\text{tophad} - \text{jet-rad1}} $	1.4/7	0.98	2.6/7	0.92	2.9/7	0.89	0.85/7	1.0	2.4/7	0.93
$ \Delta \phi^{\text{jet-W1} - \text{jet-rad1}} $	1.6/10	1.0	2.9/10	0.98	3.4/10	0.97	0.77/10	1.0	2.7/10	0.99
$m^{\text{t}\bar{\text{t}}} - \text{jet-rad1}$	7.7/8	0.46	6.5/8	0.59	4.4/8	0.81	4.4/8	0.82	5.5/8	0.71
$p_{\text{T}}^{\text{jet-rad2}}$	3.2/9	0.96	4.0/9	0.91	37/9	< 0.01	1.7/9	1.0	11/9	0.28
$ y^{\text{jet-rad2}} $	1.4/10	1.0	3.9/10	0.95	2.4/10	0.99	0.2/10	1.0	5.3/10	0.87
$ \Delta y^{\text{jet-rad1} - \text{jet-rad2}} $	1.6/10	1.0	4.0/10	0.95	4.0/10	0.95	0.13/10	1.0	5.4/10	0.86
$ \Delta \phi^{\text{jet-rad1} - \text{jet-rad2}} $	1.4/10	1.0	4.0/10	0.95	2.4/10	0.99	0.25/10	1.0	5.2/10	0.88
$ \Delta \phi^{\text{tolep} - \text{jet-rad2}} $	1.3/7	0.99	3.2/7	0.86	2.0/7	0.96	0.46/7	1.0	4.1/7	0.77
$ \Delta \phi^{\text{tophad} - \text{jet-rad2}} $	1.2/7	0.99	3.2/7	0.87	2.1/7	0.96	0.46/7	1.0	4.1/7	0.77
$ \Delta \phi^{\text{jet-W1} - \text{jet-rad2}} $	1.3/10	1.0	3.8/10	0.95	2.4/10	0.99	0.22/10	1.0	5.1/10	0.88
$m^{\text{jet-rad1} - \text{jet-rad2}}$	6.2/9	0.72	6.7/9	0.66	2.6/9	0.98	4.2/9	0.9	8.8/9	0.45

10 Extraction of α_s using top-quark pair differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS

The strong coupling constant, α_s is a fundamental parameter of the theory that describes the colour interactions. The value of this parameter is not given by the theory; it must be extracted from experimental measurements. However, the theory provides the behaviour of α_s as a function of the scale of the interaction (running). Measuring the running of α_s is a test of QCD, as the running depends on the gauge symmetry group underlying the theory and the number of active flavours.

The calculations of cross sections for $t\bar{t}$ production are sensitive to α_s as the main production channels in hadron colliders, such as the LHC, are mediated by the strong interaction. The measurements from the analysis presented in Chapter 9 are used for the extraction of $\alpha_s(m_Z)$ and the test of the running, which are presented in this chapter.

The structure of this chapter is as follows. Section 10.1 presents the measurements and simulations used in the extraction of α_s . The method used for the extraction is described in Section 10.2. Section 10.3 details the systematic uncertainties taken into account in the measurement, and the results are presented in Section 10.4.

10.1 Data and next-to-leading-order calculations

Three of the $t\bar{t}$ absolute differential cross sections presented in Chapter 9 have been taken as input for this analysis. The distributions chosen are the cross sections as functions of $p_T^{\text{jet-W-lead}}$, $p_T^{\text{jet-rad-lead}}$, and $m^{t\bar{t}-\text{jet-rad-lead}}$. These cross sections are measured with events recorded by ATLAS with 140 fb^{-1} of Run 2 at $\sqrt{s} = 13$ TeV using the $t\bar{t} \ell + \text{jets}$ channel. Distributions as functions of transverse momenta and invariant masses have been preferred for this analysis over the angular distributions and angular correlations since these distributions have a strong dependence on α_s . The $p_T^{\text{jet-rad-sublead}}$ distribution has not been used as the predictions for this observable are only NLO and the second QCD radiation originates exclusively from the parton showers.

10 Extraction of α_s using top-quark pair differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS

Three predictions with different values of $\alpha_s(m_Z)$ are used in this analysis. These samples are modelled using the POWHEG BOX v2 [211–214] generator, which provides matrix elements at NLO in α_s , and the NNPDF3.0NLO [167] PDF. The h_{damp} parameter, which controls the matching in POWHEG and effectively regulates the high- p_T radiation against which the $t\bar{t}$ system recoils, was set to $1.5m_t$ [203]. The functional form of the renormalisation and factorisation scales was set to the default scale $\sqrt{m_t^2 + p_T^2}$. The events were interfaced with PYTHIA 8.230 [155] for the parton shower and hadronisation, using the A14 set of tuned parameters [157] and the NNPDF2.3LO set of PDFs [156]. The decays of bottom and charm hadrons were simulated using the EVTGEN 1.6.0 program [251].

The $t\bar{t}$ sample was normalised to the cross-section prediction at NNLO in QCD including the resummation of NNLL soft-gluon terms calculated using TOP++ 2.0 [11–16, 220]. For pp collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, this cross section corresponds to $\sigma_{t\bar{t}}^{\text{NNLO+NNLL}} = 832 \pm 51$ fb using a top-quark mass of $m_t = 172.5$ GeV.

The different values used for $\alpha_s(m_Z)$ are 0.117, 0.118, and 0.119. Figure 10.1 shows the ratios of the predictions for the three observables computed with the three values of $\alpha_s(m_Z)$.

The scale uncertainty is computed by varying separately the renormalisation and factorisation scales. Both scales are increased and decreased by a factor of two. The NNPDF3.0NLO replicas were used to evaluate the PDF uncertainties for the nominal PDF.

10.2 $\alpha_s(m_Z)$ extraction method

The method to extract $\alpha_s(m_Z)$ follows that of Refs. [252, 253]. In this method, the differential cross section in bin i is parameterised as a function of $\alpha_s(m_Z)$ as:

$$\left[\frac{d\sigma}{dA}(\alpha_s(m_Z)) \right]_i = C_1^i \cdot \alpha_s(m_Z) + C_2^i \cdot \alpha_s^2(m_Z), \quad (10.1)$$

where A is the corresponding observable and C_j^i with $j = 1, 2$ are two parameters characterising the dependence on $\alpha_s(m_Z)$ for each bin. A pair of parameters C_1^i and C_2^i is extracted for each bin. They are determined by a χ^2 -fit method of Equation 10.1 for each bin using the three predictions at three different values of $\alpha_s(m_Z)$. The extracted $\alpha_s(m_Z)$ value is computed as the solution to the quadratic Equation 10.1 given the pair of C_j^i parameters and the measured cross section.

Equation 10.1 has been chosen to be quadratic since it provides a good approximation to the dependence on $\alpha_s(m_Z)$ of the NLO predictions. To account for the correlation of $\alpha_s(m_Z)$ on the PDFs, together with the variation of $\alpha_s(m_Z)$ in the hard process, the PDF is varied accordingly.

10.2 $\alpha_s(m_Z)$ extraction method

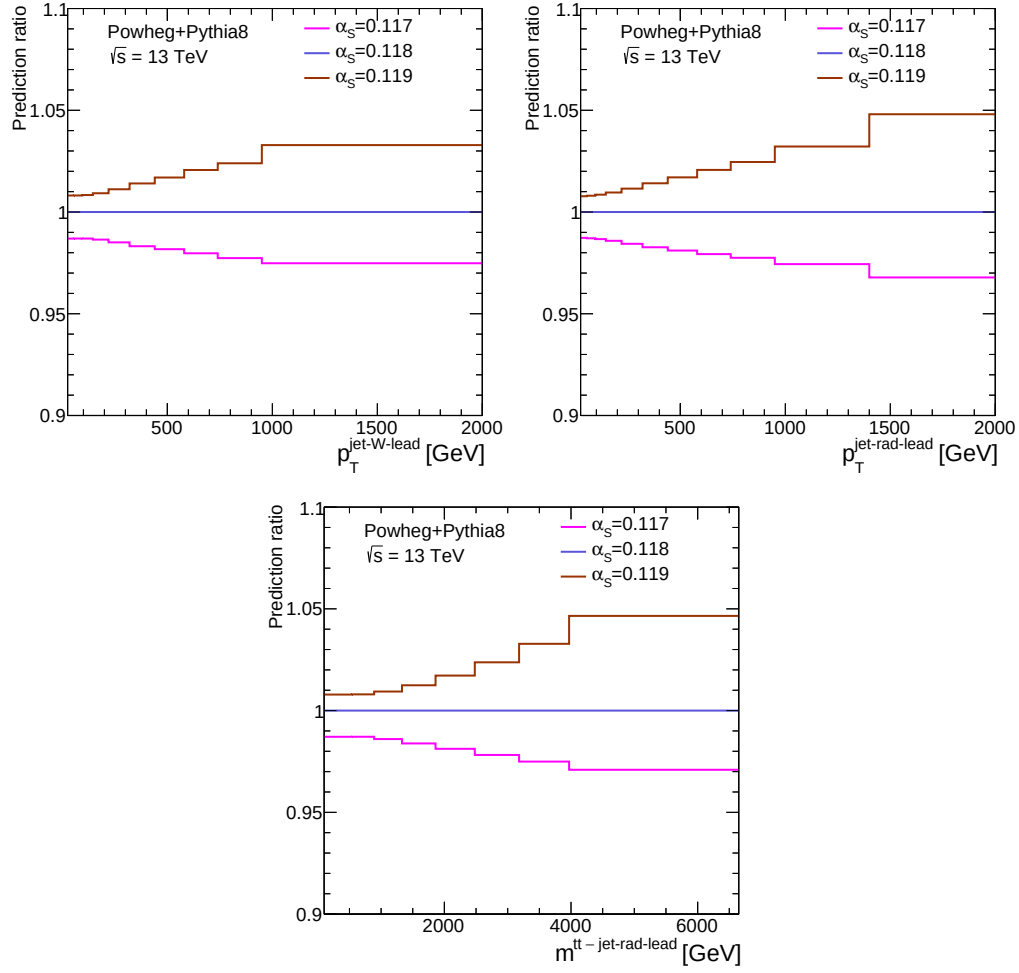


Figure 10.1: Ratios of the predictions computed with the three different $\alpha_s(m_Z)$ and the prediction computed with the nominal $\alpha_s(m_Z) = 0.118$ value. In the top row, the ratio as a function of $p_T^{\text{jet-W-lead}}$ (left) and as a function of $p_T^{\text{jet-rad-lead}}$ (right) are presented. In the bottom plot, the ratio as a function of $m^{\text{tt-jet-rad-lead}}$ is shown.

10 Extraction of α_s using top-quark pair differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS

A combined value of $\alpha_s(m_Z)$ is computed for each of the three observables. Another χ^2 -fit method of Equation 10.1 is applied simultaneously to every bin using as inputs the measured cross sections and the pairs of C_j^i parameters as input.

Additionally, an extraction of α_s as a function of the scale is computed. For this extraction, the scale in each bin is taken as the transverse momentum or the invariant mass of the corresponding observable at the bin centre. For each prediction, the α_s is computed at the corresponding scale of each bin. In this case, Equation 10.1 is modified using $\alpha_s(Q)$ instead of $\alpha_s(m_Z)$. After this change, the same process is repeated.

10.3 Systematics uncertainties

The sources of systematic uncertainties affecting the extraction of $\alpha_s(m_Z)$ and the test of the running of α_s are propagated from those affecting the measurement of the differential cross sections presented in Chapter 9. These uncertainties account for the b -tagging calibration and efficiency, the detector energy scale and resolution, the reconstruction efficiency, the background and signal modelling, the statistics of the MC samples, and the uncertainty arising from non-closure. Each source of uncertainty is separated into independent contributions and evaluated by varying the measured cross section and extracting a new value of $\alpha_s(m_Z)$. Systematic uncertainties are evaluated with respect to the nominal value of $\alpha_s(m_Z)$ or $\alpha_s(Q)$ separately for the upward and downward variations (or in the case of a single variation by symmetrising the single deviation) for each bin of each observable or for the combined extraction.

The total systematic uncertainty is computed by adding in quadrature the sources of uncertainty listed in the previous sections together with the statistical uncertainty.

10.4 Results

The values of $\alpha_s(m_Z)$ extracted for each bin using the absolute differential $t\bar{t}$ cross sections are shown in Figure 10.2. The values extracted using the three cross sections, i. e. $p_T^{\text{jet-W-lead}}$, $p_T^{\text{jet-rad-lead}}$, and $m^{t\bar{t}-\text{jet-rad-lead}}$, exhibit a similar behaviour. The extracted values typically agree with the world average [91], with some exceptions when the transverse momentum and invariant mass increase. Additionally, a trend is observed in the three cases. The value of the extracted $\alpha_s(m_Z)$ decreases as the transverse momentum and invariant mass increase. However, the values extracted from each bin are typically in agreement with each other and with the world average within uncertainties.

Three values of $\alpha_s(m_Z)$ using the combination of every bin in each cross section have been extracted. The results are:

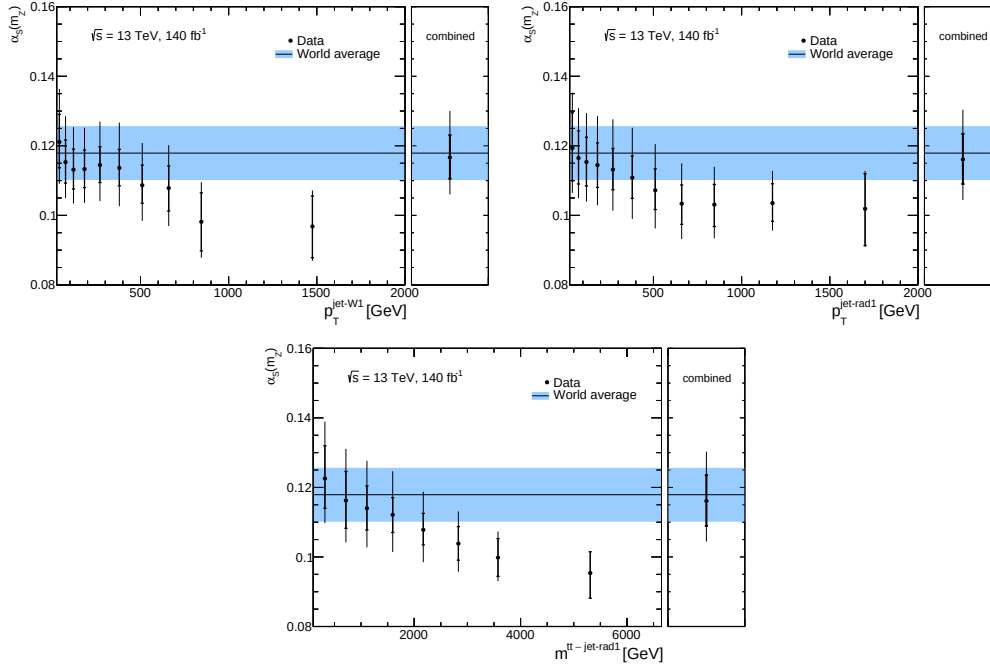


Figure 10.2: Extracted values of $\alpha_s(m_Z)$ obtained for each bin of the measured differential cross sections. In the first row, the $p_T^{\text{jet-W-lead}}$ ($p_T^{\text{jet-rad-lead}}$) cross section is used for the extraction in the left (right) plot. The bottom plot presents the $\alpha_s(m_Z)$ extracted values using the $m^{t\bar{t}}\text{-jet-rad-lead}$ cross section. The right side of each plot displays the combined value for $\alpha_s(m_Z)$. The inner error bars represent the experimental uncertainty in the extraction of $\alpha_s(m_Z)$, and the outer error bars represent the sum in quadrature of the experimental and theoretical uncertainties. The world average [91] is shown as a solid line with a shaded area representing the uncertainty.

- From $p_T^{\text{jet-W-lead}}$: $\alpha_s(m_Z) = 0.1167^{+0.0064}_{-0.0061} (\text{exp.})^{+0.0117}_{-0.0087} (\text{theo.})$
- From $p_T^{\text{jet-rad-lead}}$: $\alpha_s(m_Z) = 0.1161^{+0.0073}_{-0.0071} (\text{exp.})^{+0.0122}_{-0.0093} (\text{theo.})$
- From $m^{t\bar{t}}\text{-jet-rad-lead}$: $\alpha_s(m_Z) = 0.1161^{+0.0074}_{-0.0072} (\text{exp.})^{+0.0121}_{-0.0092} (\text{theo.})$

Figure 10.3 shows the running of α_s . The dependence of α_s on the scale extracted from each observable is in agreement with the running predicted by QCD. Similarly to the extraction of $\alpha_s(m_Z)$, the region at high scale shows greater differences between the measurements from individual bins and the running of the combined $\alpha_s(m_Z)$, but it is in agreement with the predicted behaviour within the uncertainties.

The comparison of the three combined values for $\alpha_s(m_Z)$ is shown in Figure 10.4 together with the extracted values obtained by the ATLAS Collaboration using multijets transverse energy–energy correlations (TEEC) [254], and the results obtained by the ZEUS Collaboration in DIS [255] and in photoproduction [256]. Every measurement is in agreement with the

10 Extraction of α_s using top-quark pair differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS

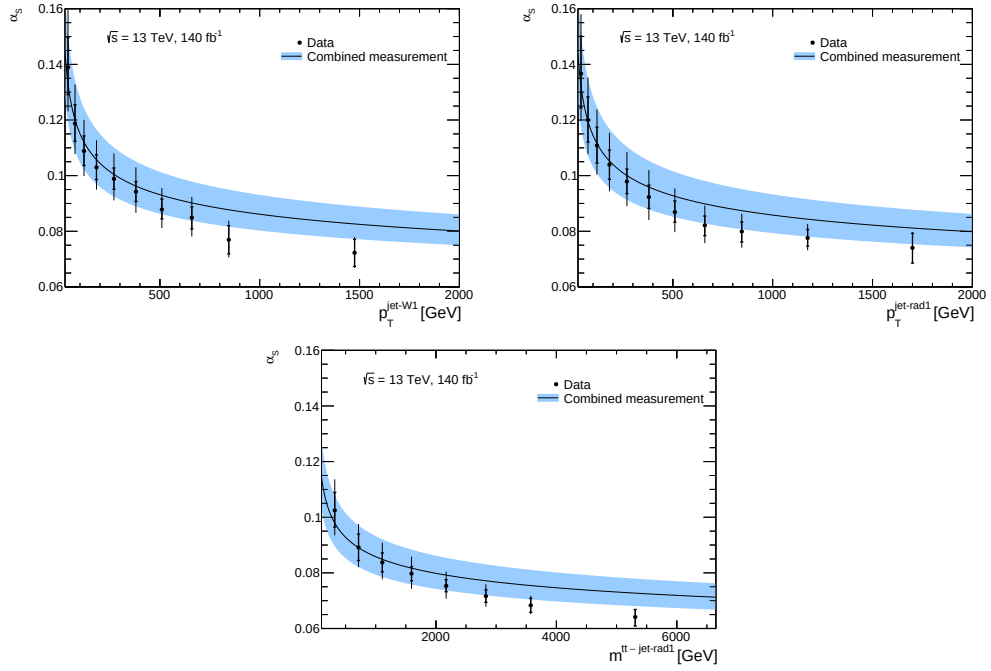


Figure 10.3: Running of $\alpha_s(Q)$ obtained with three different measured differential cross sections. In the first row, the $p_T^{\text{jet-W-lead}}$ ($p_T^{\text{jet-rad-lead}}$) cross sections is used for determining the values of $\alpha_s(Q)$ in the left (right) plot. The bottom plot presents the $\alpha_s(Q)$ running obtained using the $m^{t\bar{t}} - \text{jet-rad-lead}$ cross section. The running extracted from the combined measurement is shown as a solid line surrounded by a shaded area representing the sum in quadrature of the experimental and theoretical uncertainties.

world average, within the uncertainties. The uncertainties for the values obtained in this analysis are larger than in the other studies. This is due to the fact that extraction of α_s in the TEEC analysis used ratios of cross sections, for which the correlated uncertainties cancel, and is based on NNLO predictions. The analyses from ZEUS yield smaller uncertainties due to the fact that a careful choice of observables and region of phase space was made so that experimental and theoretical uncertainties were minimised.

Figure 10.5 shows the running of $\alpha_s(Q)$ as a function of the scale obtained with three different measured differential cross sections. In this figure, they are compared to the running obtained by the ATLAS Collaboration using multijets transverse energy–energy correlations (TEEC) [254], and the results obtained by the ZEUS Collaboration in DIS [257] and in photoproduction [256]. The running of α_s for the extracted value of $\alpha_s(m_Z)$ using $p_T^{\text{jet-rad-lead}}$ is shown for comparison together with the running of the world average. The predicted running of α_s is in good agreement with the extracted values at different scales.

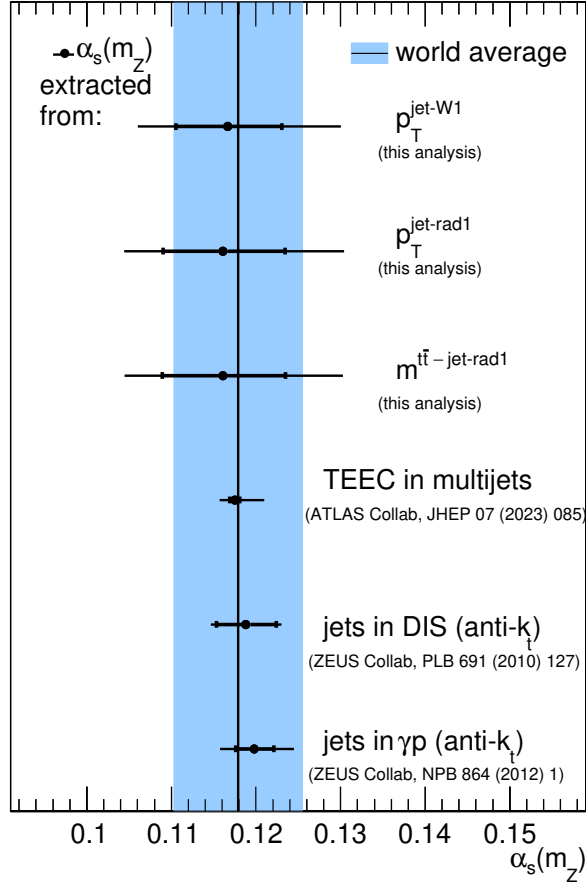


Figure 10.4: Comparison of the combined values of $\alpha_s(m_Z)$ extracted using the $t\bar{t}$ differential cross sections as functions of $p_T^{\text{jet-W-lead}}$, $p_T^{\text{jet-rad-lead}}$, and $m^{t\bar{t}-\text{jet-rad-lead}}$ together with the extracted values obtained by the ATLAS Collaboration using multijets transverse energy–energy correlations (TEEC) [254], and the results obtained by the ZEUS Collaboration in DIS [255] and in photoproduction [256]. Experimental uncertainties are represented by the inner error bars and the outer error bars represent the sum in quadrature of the experimental and theoretical uncertainties. The world average [91] is shown as a solid line with a shaded area representing the uncertainty.

10 Extraction of α_s using top-quark pair differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS

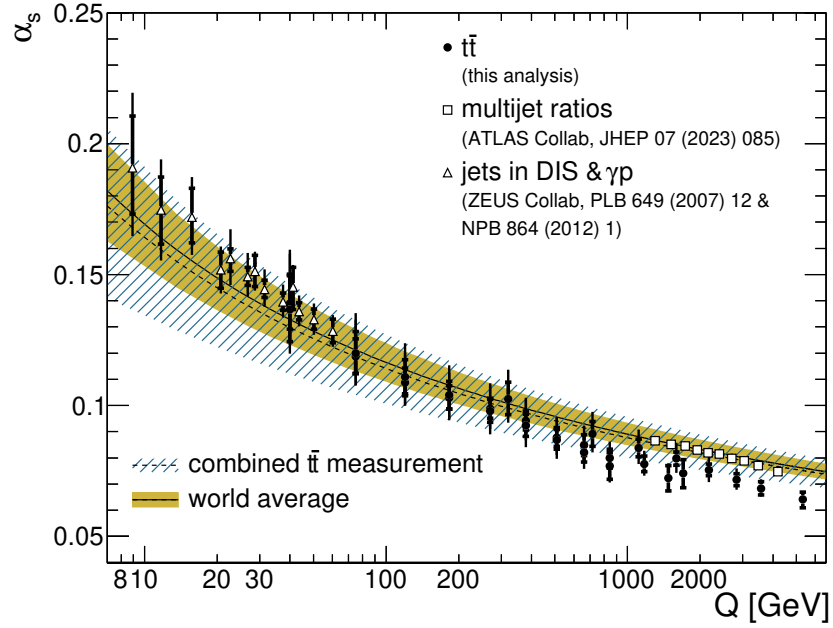


Figure 10.5: Running of $\alpha_s(Q)$ obtained from the measured differential $t\bar{t}$ cross sections (dots), the running obtained by the ATLAS Collaboration using multijets transverse energy–energy correlations (TEEC) [254] (open squares), and the results obtained by the ZEUS Collaboration in DIS [257] and in photoproduction [256] (open triangles). Experimental uncertainties are represented by the inner error bars and the outer error bars represent the sum in quadrature of the experimental and theoretical uncertainties. The running of the combined value of $\alpha_s(m_Z)$ extracted using the $t\bar{t}$ differential cross sections as a function of $p_T^{\text{jet-rad-lead}}$ is represented by a dashed line surrounded by a hatched area representing the sum in quadrature of the experimental and theoretical uncertainties. The running of the world average [91] of $\alpha_s(Q)$ is represented by a solid line surrounded by a shaded area representing the uncertainty.

11 Summary and conclusions

Three different measurements were presented in this thesis. The analysis included the *in situ* jet energy scale calibration using γ +jets events, the measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data, and the extraction of α_s using top-quark pair differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS.

The *in situ* jet energy scale calibration using γ +jets events using data recorded by ATLAS during the 2015-2018 period was measured. The JES was measured as a function of the photon transverse momentum for photons with $p_T^\gamma > 25$ GeV and $|\eta^\gamma| < 1.37$. The leading jet, reconstructed using the PFlow method, was required to have $p_T^{\text{jet}} > 10$ GeV and $|\eta^{\text{jet}}| < 0.8$. The JES was measured with an uncertainty below 1% in the $p_T^{\text{ref}} > 45$ GeV range. This was achieved thanks to the complete luminosity of Run 2, the MPF method, and PFlow jets, which improved the statistics and the techniques used for the measurements with respect to previous calibrations. The *in situ* correction deviates from unity by less than 3% and decreases to just above 1% as p_T^{ref} increases. Between 25 and 45 GeV, this trend is not observed due to limited statistics. The JES measured with both PYTHIA and SHERPA simulations exhibit similar behaviours. The measurement range reaches up to 2000 GeV.

The measurements of top-quark pair differential cross sections in the ℓ +jets channel in pp collisions at $\sqrt{s} = 13$ TeV using 140 fb^{-1} of ATLAS data recorded between 2015 and 2018 was presented. The $t\bar{t}$ events were selected in the ℓ +jets decay mode. The measurements presented benefit from the large statistics available from the full Run 2 sample and a reduction of the systematic uncertainties compared to previous analyses. Absolute and normalised differential cross sections as functions of jet observables, including jet transverse momenta and rapidities, angular correlations and invariant masses, were measured in three channels: $t\bar{t}$ inclusive with $N^{\text{jets}} \geq 4$, $t\bar{t}+1\text{jet}$ with $N^{\text{jets}} \geq 5$ and $t\bar{t}+2\text{jets}$ with $N^{\text{jets}} \geq 6$. The observables, either constructed from the jets forming the $t\bar{t}$ system or arising from hard QCD radiation, were measured and used to characterise the kinematics, topology and dynamics of the $t\bar{t}$ system; they also proved to be sensitive to initial- and final-state QCD radiation effects, represented by the production of additional jets. The measured cross section as a function of $p_T^{\text{jet-W1}}$ has a harder spectrum than that of $p_T^{\text{jet-W2}}$. The measured cross sections as functions of $|y^{\text{jet-W1}}|$ and $|y^{\text{jet-W2}}|$ have very similar shape and normalisation. The measured cross section as a function of $|\Delta y^{\text{jet-W1-jet-W2}}|$ ($|\Delta\phi^{\text{jet-W1-jet-W2}}|$) exhibits a peak around $|\Delta y^{\text{jet-W1-jet-W2}}| = 0$ ($|\Delta\phi^{\text{jet-W1-jet-W2}}| = 1$ rad). These observables characterise the kinematics of the two light jets forming the hadronically-decaying top. Values of $m^{t\bar{t}\text{-jet-rad1}}$

11 Summary and conclusions

($m^{\text{jet-rad1} - \text{jet-rad2}}$) up to ~ 5.3 TeV (~ 4 TeV) were measured. The measured cross sections as functions of $m^{t\bar{t} - \text{jet-rad1}}$ and $m^{\text{jet-rad1} - \text{jet-rad2}}$ decrease by more than five orders of magnitude within the measured range. These observables characterise the dynamics of the first and second emissions. From the event topology point of view, differences in the shape of the measured cross sections as functions of the rapidities and angular correlations for the $t\bar{t}+1\text{jet}$ and $t\bar{t}+2\text{jets}$ channels with respect to those measured in the $t\bar{t}$ inclusive channel indicate that the first and second emissions tend to be more isotropic in rapidity than the jets forming the hadronically-decaying top. This is supported by the measurements as functions of $|\Delta\phi^{\text{tolep} - \text{jet-rad1}}|$, $|\Delta\phi^{\text{tophad} - \text{jet-rad1}}|$, $|\Delta\phi^{\text{tolep} - \text{jet-rad2}}|$ and $|\Delta\phi^{\text{tophad} - \text{jet-rad2}}|$, which show that the first and second emissions have a higher probability of arising far from the leptonically-decaying top and that there are more events with a gluon emission back-to-back with the jets forming the leptonically-decaying top. In addition, the measured features indicate that the second hard emission tends to be close in rapidity to the first emission and that there is a weak dependence on the difference in ϕ between the first and second emissions. NLO QCD predictions normalised to $\mathcal{O}(\text{NNLO}+\text{NNLL})$ describe well the shape and normalisation of the measured absolute differential cross sections as functions of angular observables, but the transverse momentum and invariant mass observables are, in general, described well only at low values. Recently, predictions for $t\bar{t}$ production at NNLO QCD have become available and the interface with the parton shower provided by PYTHIA via the MiNNLOPS matching scheme allows the computation of differential cross sections for jet observables, as those presented here. The NNLO predictions are compared to the measured cross sections and provide a good description of the angular observables, though there is a tendency of the NNLO predictions to be below the data for $|\Delta y^{\text{jet-W1} - \text{jet-W2}}| > 3$. The description of the data by the NNLO prediction is somewhat worse than that at NLO for the $p_T^{\text{jet-rad2}}$. For the transverse momentum and invariant mass observables, the NNLO QCD predictions provide a significant improvement in the description of the $p_T^{\text{jet-W1}}$, $p_T^{\text{jet-W2}}$, $p_T^{\text{jet-rad1}}$, $m^{t\bar{t} - \text{jet-rad1}}$ and $m^{\text{jet-rad1} - \text{jet-rad2}}$ observables with respect to that of the NLO predictions.

The **extraction of α_s using top-quark pair differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV measured by ATLAS** was presented. The extraction was performed using the differential cross sections as functions of $p_T^{\text{jet-W-lead}}$, $p_T^{\text{jet-rad-lead}}$, and $m^{t\bar{t} - \text{jet-rad-lead}}$. The strong coupling at the Z-boson mass scale was measured for each bin of the cross-section distributions. The extracted values typically agree with the world average, except for very large values of the transverse momenta and invariant mass. From each distribution a combined value was extracted. All of them agree with each other within the uncertainties and there is good agreement with the world average. The running of α_s was measured for the same three distributions taking the transverse momentum and the invariant mass of each bin as the scale. In the three cases, $\alpha_s(Q)$ decreases as Q increases and their dependence is in agreement with the running of the combined $\alpha_s(m_Z)$ within the uncertainties as predicted by QCD.

Appendices

A Background response fits for the *in situ* measurements

Figure A.1 shows the r_{MPF} distributions and the fitted Gaussians to extract the mean response in Region C. The fitting method is the same as that used for the *Tight* and isolated sample. In the 1200 GeV – 1600 GeV range, a couple of bins do not have enough points to perform the fit. However, this is not a problem because the purity approaches unity as $p_{\text{T}}^{\text{ref}}$ increases. In the $p_{\text{T}}^{\text{ref}} > 1200$ GeV region, the purity is compatible with unity within the uncertainties and no background response is needed for the correction.

A Background response fits for the *in situ* measurements

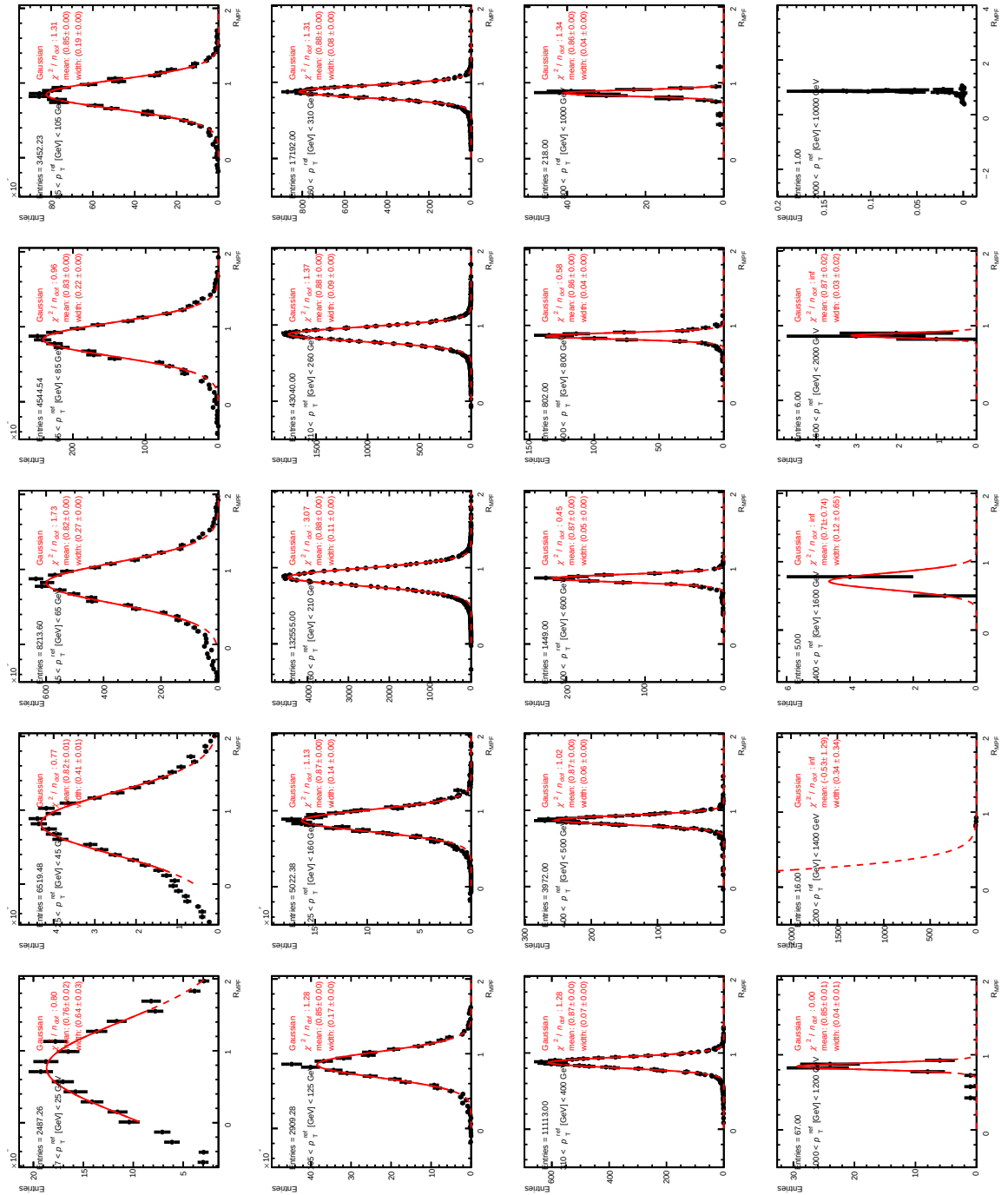


Figure A.1: Distributions of r_{MPF} in different bins of $p_{\text{T}}^{\text{ref}}$ for 2015-2018 data in Region C. The data distributions are represented as dots together with the Gaussian-fitted distribution displayed as a solid line in the fitting range and dashed outside.

B Uncertainties arising from the purity correction in the *in situ* measurements

The uncertainty of the *in situ* jet energy scale arising from the purity correction is computed as the sum in quadrature of three different components:

- The uncertainty arising from using the SHERPA sample to compute the signal leakage fraction in the purity estimation.
- The uncertainty from varying the photon identification working point used to define background regions C and D.
- The uncertainty from deviations from unity of the R^{bg} correlation.

For the uncertainty originating from the simulation sample used to compute the signal leakage fractions, the difference in purity from using the SHERPA sample instead of the PYTHIA sample is taken as the uncertainty. Figure B.1 shows this uncertainty contribution. This uncertainty varies from around 0.03% in the first bin and decreases as $p_{\text{T}}^{\text{ref}}$ increases. This uncertainty is symmetrised.

For the uncertainty arising from the identification criteria used for photon candidates in regions C and D, three alternative working points are used to compute three different nuisance parameters:

- *LoosePrime2*: adds to the *Loose* identification the requirements on R_{ϕ} , w_{stot} , f_1 , E_{ratio} , and ΔE_s from the *Tight* working point. This is the tightest selection criterion used as an alternative to *LoosePrime4*.
- *LoosePrime3*: adds to the *Loose* identification the requirements on R_{ϕ} , w_{stot} , f_1 , and E_{ratio} from the *Tight* working point.
- *LoosePrime5*: adds to the *Loose* identification the requirements on R_{ϕ} and f_1 from the *Tight* working point. This is the loosest selection criterion used as an alternative to *LoosePrime4*.

Figure B.2 shows the three nuisance parameters for the three identification working points. All of them are negligible except in the first bin where *LoosePrime2* and *LoosePrime3* are in absolute value slightly higher than 0.5%, while *LoosePrime5* is much lower.

B Uncertainties arising from the purity correction in the *in situ* measurements

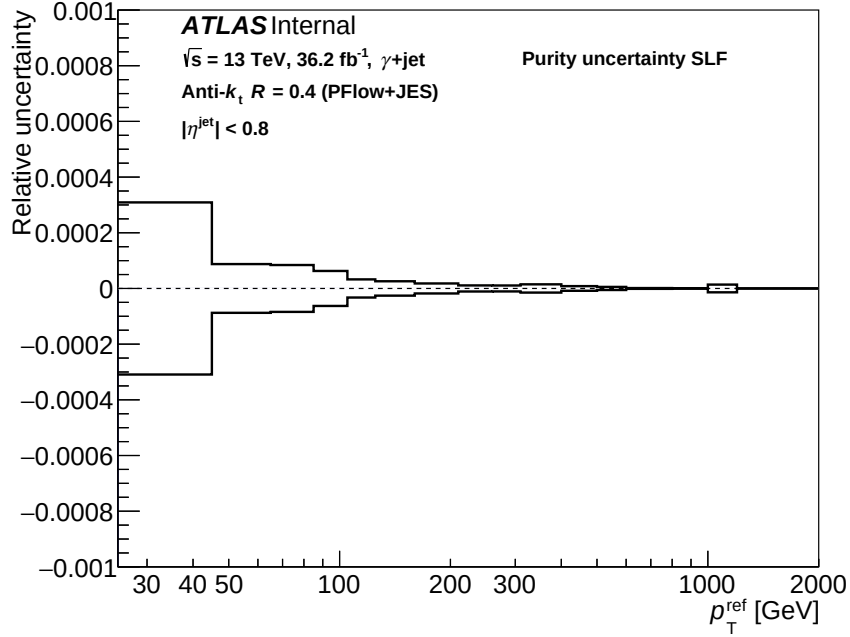


Figure B.1: Relative uncertainty of the purity arising from the simulation used to compute the signal leakage fractions as a function of p_T^{ref} .

The three nuisance parameters are combined with the envelope distribution of all of them. Figure B.3 shows the uncertainty of the purity arising from the identification working point used to define the background regions. The positive component reaches above 0.1% in the first bin, while the negative reaches -0.5% . In both cases, they are negligible for higher p_T^{ref} .

For the uncertainty arising from deviations from unity in the R^{bg} correlation, R^{bg} has been varied from ± 0.2 in the low- p_T^{ref} region to ± 0.6 as p_T^{ref} increases. This variation is taken as a conservative modification in the value of R^{bg} given by the studies presented in Appendix C. This uncertainty goes from around 0.2% at low- p_T^{ref} and reduces for higher values of p_T^{ref} (see Figure B.4).

The combination in quadrature of the three components of the purity uncertainty is shown in Figure B.5. The positive component decreases from above 0.2% in the first bin as p_T^{ref} increases. A similar behaviour is observed for the negative component starting above -0.6% .

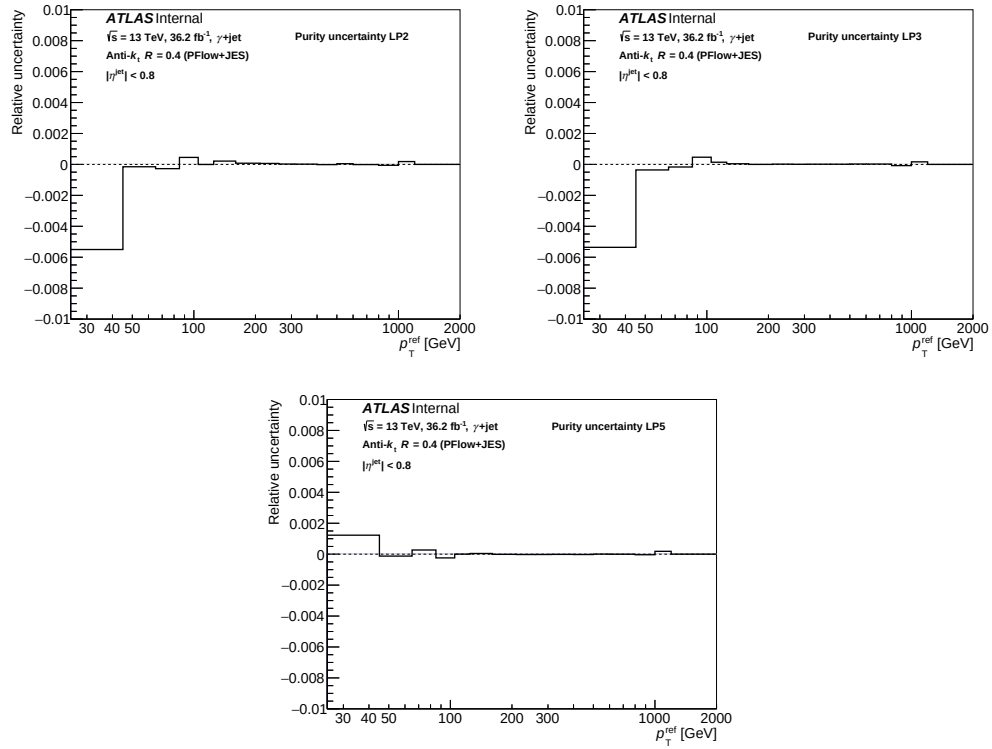


Figure B.2: Relative uncertainty of the purity arising from the three different identification working points used to define background regions as a function of p_T^{ref} . In the first row, the nuisance parameter of the *LoosePrime2* (left) and, the *LoosePrime3* (right) are presented. In the bottom plot, the *LoosePrime5* uncertainty is shown.

B Uncertainties arising from the purity correction in the *in situ* measurements

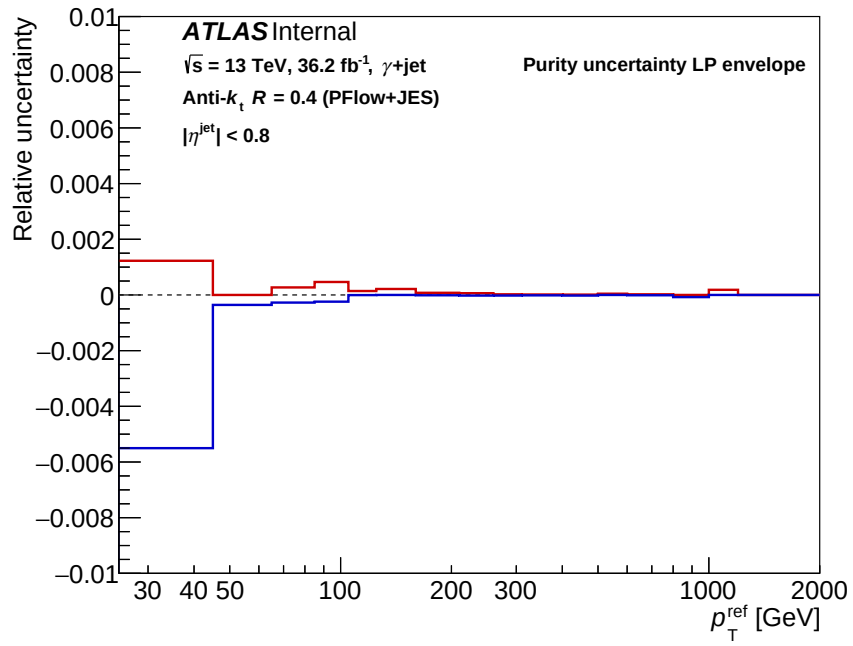


Figure B.3: Relative uncertainty of the purity accounting for the identification working point used to define background regions as a function of p_T^{ref} .

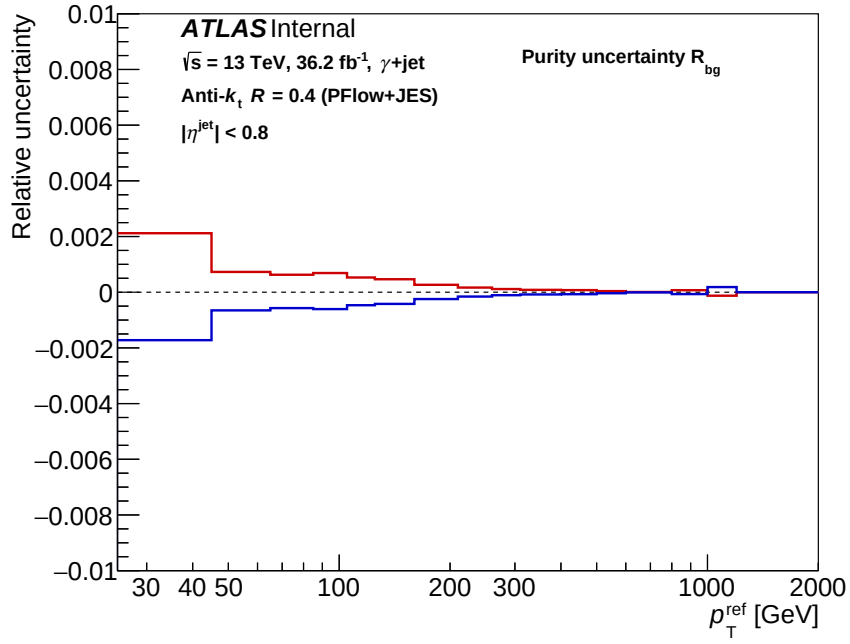


Figure B.4: Relative uncertainty of the purity accounting for deviations from unity in the R^{bg} correlation as a function of p_T^{ref} .

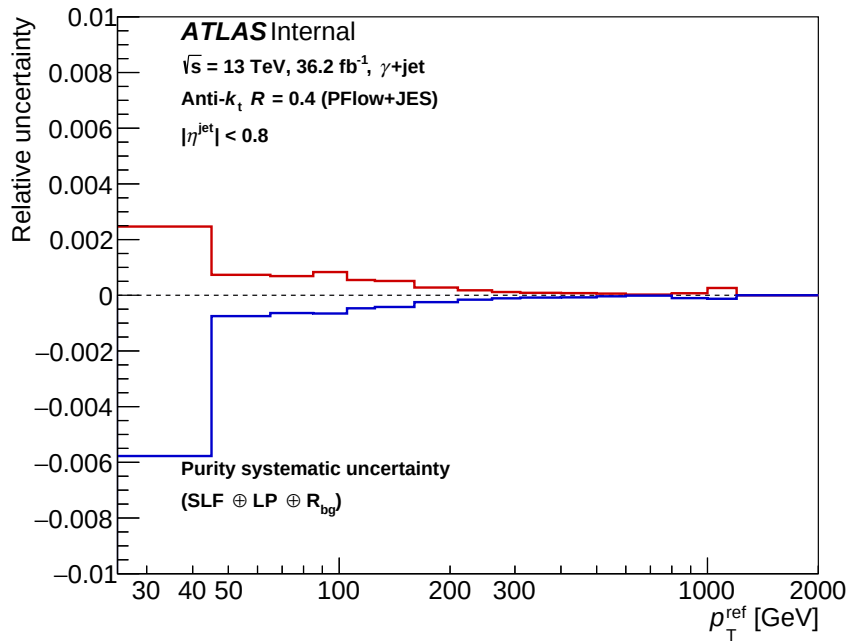


Figure B.5: Relative total uncertainty of the purity as a function of p_T^{ref} .

C Studies on R^{bg} for the *in situ* measurements

The isolation and identification photon variables used to define the plane in the data-driven method employed in the correction (see Section 8.6) are assumed as uncorrelated ($R^{\text{bg}} = 1$). To account for possible deviations from unity caused by a degree of correlation, an uncertainty is estimated by varying R^{bg} (see Appendix B). To set the variation range for R^{bg} , the same method used in Ref. [182] is applied.

R^{bg} is computed as the correlation between four control regions:

$$R^{\text{bg}} = \frac{N_{B'} \cdot N_{D''}}{N_{B''} \cdot N_{D'}}$$

where N_R with $R = B', B'', D', \text{ and } D''$ are the number of events in the background control regions defined as follows:

- Region B' : *Tight* photon candidates with $E_T^{\text{cone40}} > 0.022 \cdot E_T^\gamma + 2.45 \text{ GeV}$ and $E_T^{\text{cone40}} < 0.022 \cdot E_T^\gamma + 12.45 \text{ GeV}$.
- Region B'' : *Tight* photon candidates with $E_T^{\text{cone40}} > 0.022 \cdot E_T^\gamma + 12.45 \text{ GeV}$.
- Region D' : non-*Tight* photon candidates with $E_T^{\text{cone40}} > 0.022 \cdot E_T^\gamma + 2.45 \text{ GeV}$ and $E_T^{\text{cone40}} < 0.022 \cdot E_T^\gamma + 12.45 \text{ GeV}$.
- Region D'' : non-*Tight* photon candidates with $E_T^{\text{cone40}} > 0.022 \cdot E_T^\gamma + 12.45 \text{ GeV}$.

The R^{bg} computation is repeated six times by shifting the E_T^{iso} thresholds by 0, 1, 2, 3, 4, and 5 GeV. Figure C.1 shows the distribution of R^{bg} for the six different thresholds in E_T^{iso} .

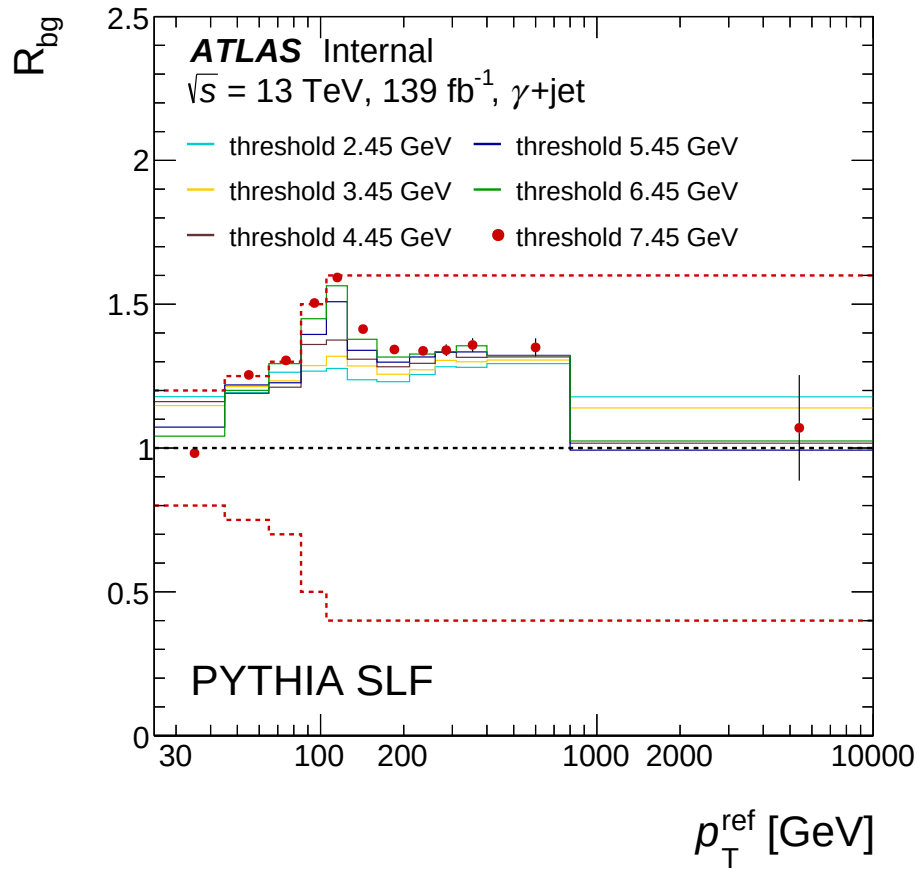


Figure C.1: R^{bg} for different $E_{\text{T}}^{\text{iso}}$ thresholds as a function of $p_{\text{T}}^{\text{ref}}$. The dashed line represents the value of R^{bg} used for the uncertainty variations.

D b -tagging optimisation

The final state of the $t\bar{t}$ system decay contains two b quarks and consequently the choice of the b -tagging WP has an important role in the efficiency of the selection used to identify $t\bar{t}$ events and on the capability to reject background events. The increase in data statistics given by basing this analysis on the full Run 2 dataset and the use of the new b -tagging algorithm DL1r allows the investigation of a tighter WP than used previously. The advantage of a tighter b -tagging WP is a reduction of the amount of remaining background in the selected sample and the reduction of its related uncertainty. The disadvantages of a tighter WP are the increase in the b -tagging uncertainty and the reduction in data statistics. Four different WPs are available for DL1r: 85%, 77%, 70% and 60%, which correspond to the efficiency for tagging the b -jet.

Data statistics. The effect of the tagging efficiency on the data statistics was investigated by counting the number of events in data, background and signal MC. Figure D.1(a) shows the integrated number of events as a function of the WP and Table D.1 shows the event yields in the $t\bar{t}$ inclusive channel for each WP. The number of data events decreases rapidly with decreasing WP. There is a good description of the data by the SM prediction for all WPs.

Background contamination. The effect of the tagging efficiency on the background contribution was investigated by studying the contribution of each background process as a function of WP. Figure D.1(b) shows the percentage of each background to the total SM prediction, together with the systematic uncertainties associated to each process. The contribution of each background process decreases as the WP decreases, except for the $t\bar{t}V$ and $t\bar{t}H$ contributions which are approximately constant. The background from W +jets dominates for the 85% WP and its contribution to this WP is much higher than for the other WPs; also, its uncertainty is higher for the 85% WP. For the other WPs, the background from single-top processes dominates. Figure D.1(c) shows the percentage of the ratio of data and SM prediction; this ratio decreases slightly with decreasing WP. In the same figure, the contribution from the signal MC and the background are shown. The contribution from the signal increases as the WP decreases, demonstrating that the contamination of the background decreases, being maximal for the 85% WP and decreasing rapidly as the WP decreases.

Systematic uncertainties. The effect of the tagging efficiency on the systematic uncertainties was investigated by estimating the contribution of each systematic nuisance parameter (NP) as a function of the WP. See Section 9.7 for full details on the systematic uncertainties. Figure D.2 shows the relative systematic uncertainty in the measured cross section as a

D b -tagging optimisation

Table D.1: Observed and expected number of events in the ℓ +jets channel for different DL1r WPs in the $t\bar{t}$ inclusive channel.

Process	Yield 85%	Yield 77%	Yield 70%	Yield 60%
$t\bar{t}$	6537684 (86%)	5088726 (90%)	4088358 (91%)	2971083 (91%)
single top	341275 (4.5%)	248561 (4.4%)	192840 (4.3%)	136233 (4.2%)
W +jets	383440 (5.0%)	159852 (2.8%)	101786 (2.3%)	64442 (2.0%)
Fakes	174952 (2.3%)	95017 (1.7%)	65491 (1.5%)	41674 (1.3%)
Z +jets	119868 (1.6%)	53483 (0.94%)	36567 (0.81%)	24275 (0.75%)
$t\bar{t}V$	21491 (0.28%)	16406 (0.29%)	13136 (0.29%)	9430 (0.29%)
diboson	19592 (0.26%)	8911 (0.17%)	5918 (0.13%)	3882 (0.12%)
$t\bar{t}H$	7183 (0.09%)	6205 (0.10%)	5423 (0.12%)	4389 (0.13%)
Total prediction	7605485	5677161	4509522	3255408
Data	7541003	5617401	4445113	3189801
Data/Prediction	0.99	0.99	0.99	0.98

function of $|y^{\text{jet}-W1}|$ for each DL1r WP. The largest contribution, not surprisingly, arises from the b -tagging NPs and the second-largest contribution comes from the detector NPs for all WPs. The background and detector (b -tagging) systematics are larger for the 85% (60%) WP than for the other WPs.

As a conclusion from all these studies, the 85% WP is discarded because it produces the largest background contamination and largest background systematic; the 60% WP is discarded because it produces the largest b -tagging and statistical uncertainties. From the two remaining WPs, the 70% WP was selected as WP in this analysis because the statistical uncertainties are similar but the background contamination is slightly smaller and is the WP that yields the same efficiency as the previous analysis [61].

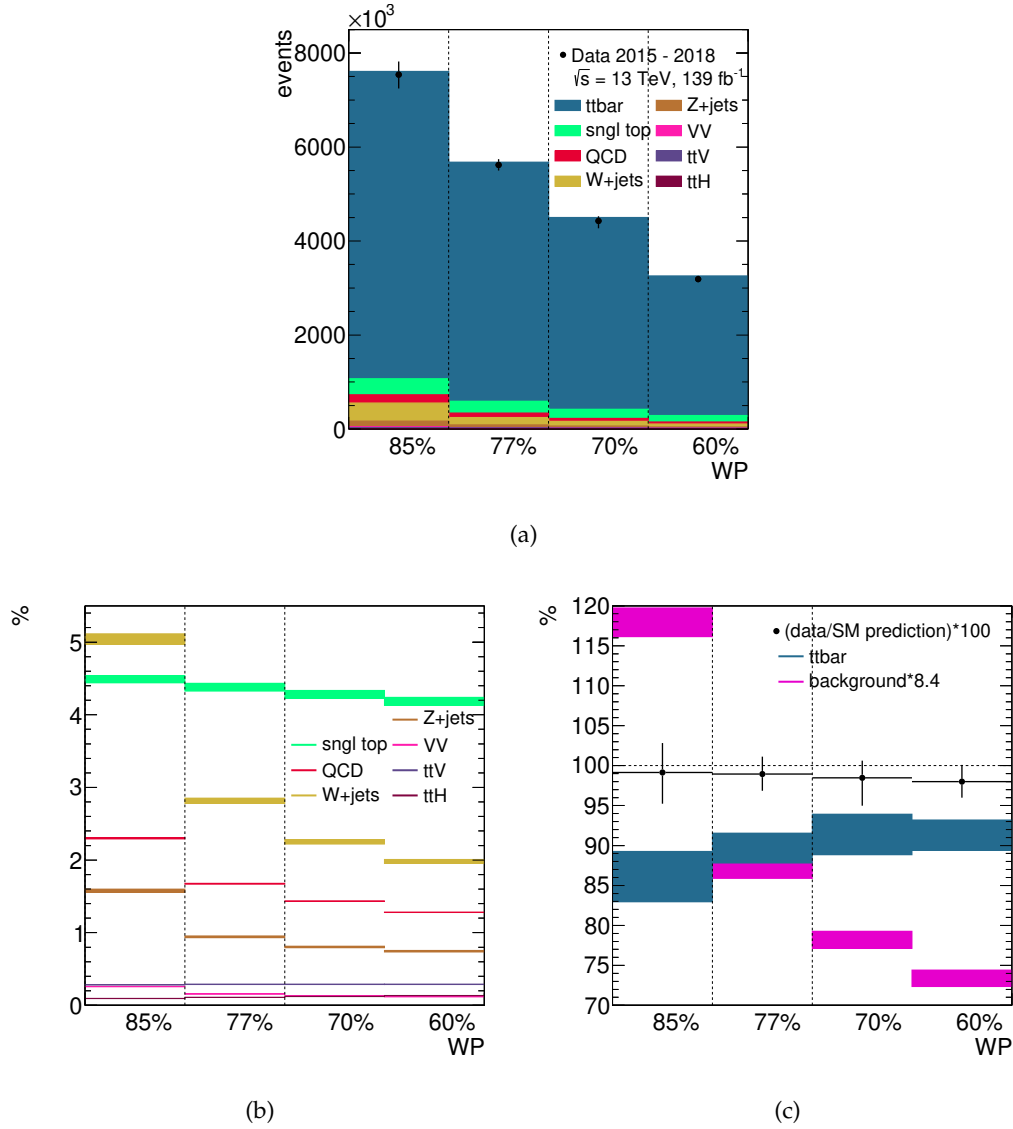


Figure D.1: The measured (a) integrated number of events (dots) for the $t\bar{t}$ inclusive channel as a function of DL1r WP. The error bars represent the statistical and systematic uncertainties added in quadrature. The MC simulation of the signal (blue histograms) and various backgrounds are also included. (b) The percentage with respect to the total SM prediction for the different background processes as a function of the DL1r WP. The width of the bands represent the systematic uncertainty associated to each background type. (c) The percentage of the ratio of the data to the SM prediction (dots) as a function of DL1r WP. The error bars represent the statistical and systematic uncertainties added in quadrature. The blue bands represent the contribution and uncertainty of the signal MC to the SM prediction. The magenta bands represent the contribution and uncertainty of the background processes to the SM prediction. The background contribution has been multiplied by a factor 8.4 to aid visibility.

D b -tagging optimisation

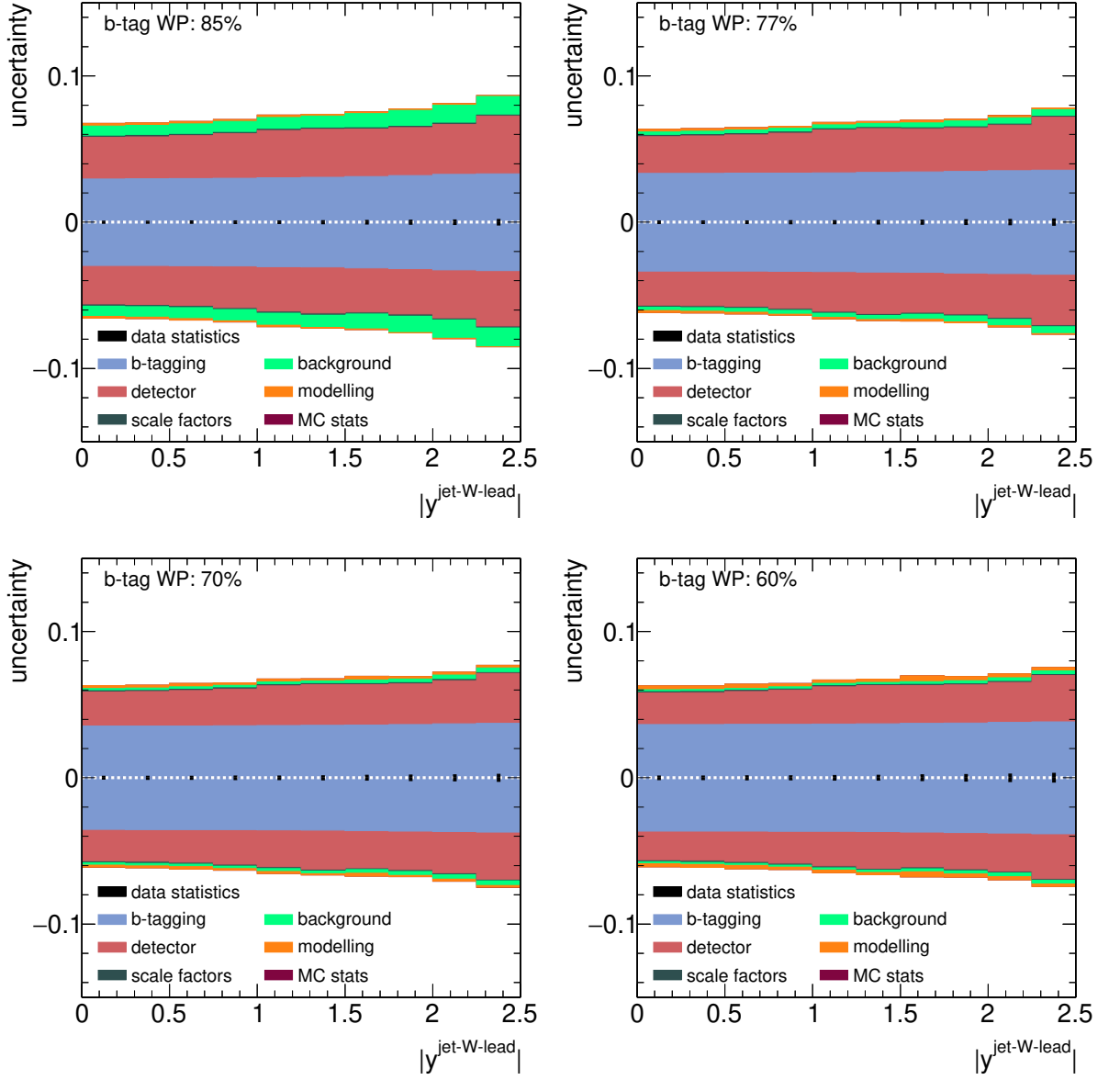


Figure D.2: The relative uncertainty in the measured cross section as a function of $|y^{\text{jet-W-lead}}|$ in the $t\bar{t}$ inclusive channel for different DL1r WPs. The contributions from the b -tagging, detector, scale factors, background and modelling NPs and data and MC statistics are shown separately.

E Control Regions for $t\bar{t}$ production

Figures 9.1 to 9.4 show the data and MC comparison for the signal region (SR), which is defined as having at least four jets and at least two b -tagged jets. Three control regions (CR) were studied, which are dominated by background and depleted of signal contribution, to estimate an uncertainty on the normalisation of the background contributions arising from the W +jets, Z +jets and fakes processes. These CRs are defined as far away as possible to the SR and taken as:

- CR1: exactly 3 jets and at least 2 b -tagged jets;
- CR2: exactly 3 jets and at most 1 b -tagged jet;
- CR3: at least 4 jets and at most 1 b -tagged jet.

Figures E.1 to E.3 show the data compared to the SM predictions at the detector level in each CR. The distributions shown in these figures are as functions of the transverse momentum of the lepton (p_T^{lep}), the pseudorapidity of the lepton (η^{lep}), the missing transverse momentum (E_T^{miss}), the transverse mass of the W boson reconstructed from the lepton and the E_T^{miss} (m_T^W), the transverse momentum of the jets (p_T^{jet}), and the transverse momentum of the b -jets ($p_T^{b\text{-jet}}$). It is seen in this figures that the normalisation of the SM prediction in the control regions is similar to that of the signal region. In addition, control regions 2 and 3 are dominated by the background contributions from W +jets, Z +jets and fakes. On the other hand, CR1, which is closest to the definition of the SR, has a higher signal and lower background contributions; still, compared to the SR, the background is high.

An uncertainty of 50% on the W +jets, Z +jets and fakes contributions covers all deviations observed between data and SM prediction for all observables in CRs 2 and 3; CR 1, which as mentioned above has a lower background contamination from W +jets, Z +jets and fakes, shows deviations in the tails of the p_T distributions that are larger than the systematic uncertainty on the normalisation of these background processes (see Section 9.7 for the full list of systematic uncertainties).

E Control Regions for $t\bar{t}$ production

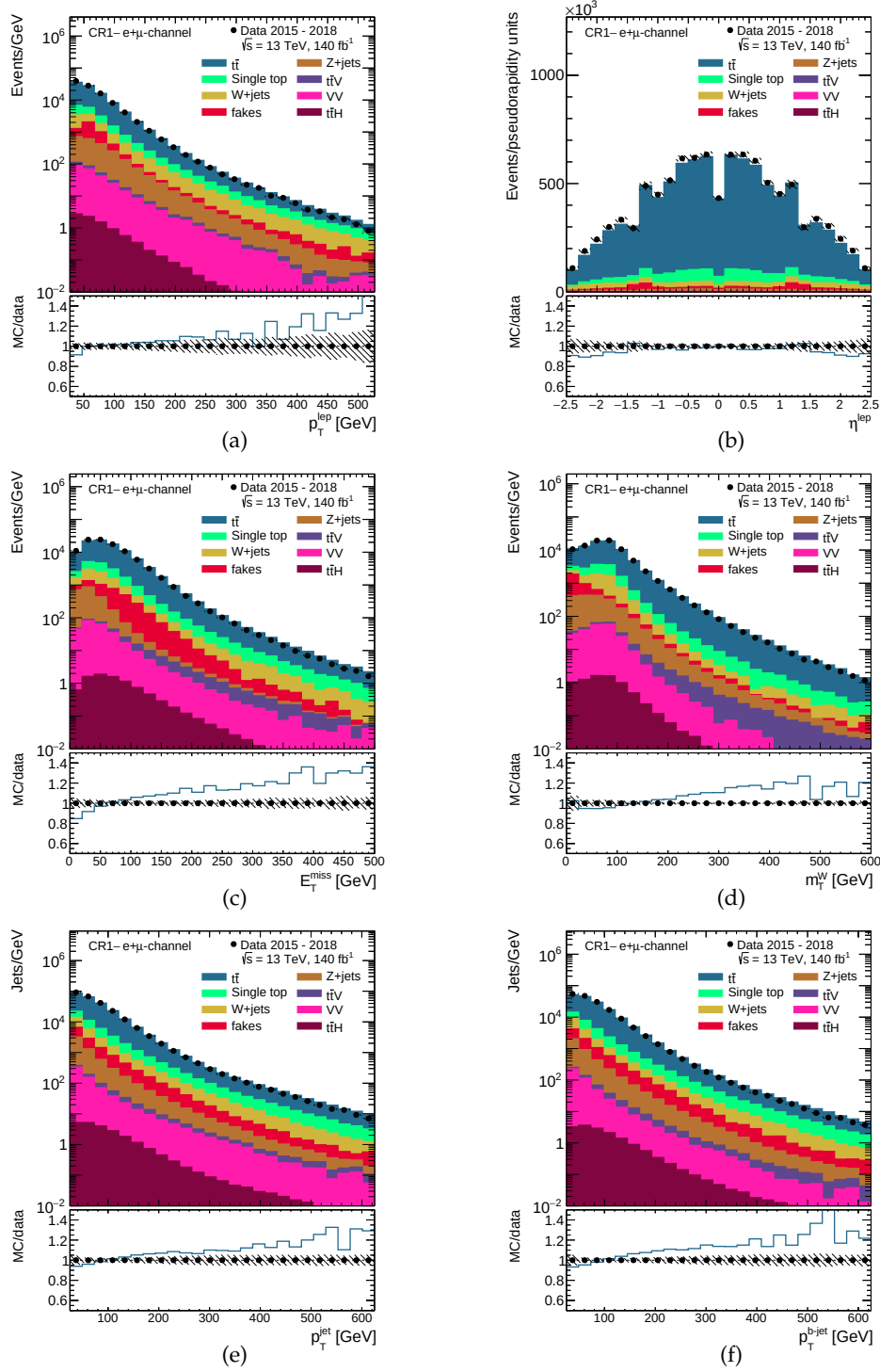


Figure E.1: The measured (a) p_T^{lep} , (b) η^{lep} , (c) E_T^{miss} , (d) m_T^W , (e) p_T^{jet} , and (f) $p_T^{\text{b-jet}}$ distributions (dots) in the $t\bar{t}$ inclusive channel for control region 1. The MC simulation of the signal (blue histograms) and various backgrounds are also included. The lower part of each figure shows the ratio of the total SM distribution to the data. The error bars represent the statistical uncertainty of the data and the hatched bands represent the uncertainty on the normalisation of the W+jets, Z+jets and fakes contributions.

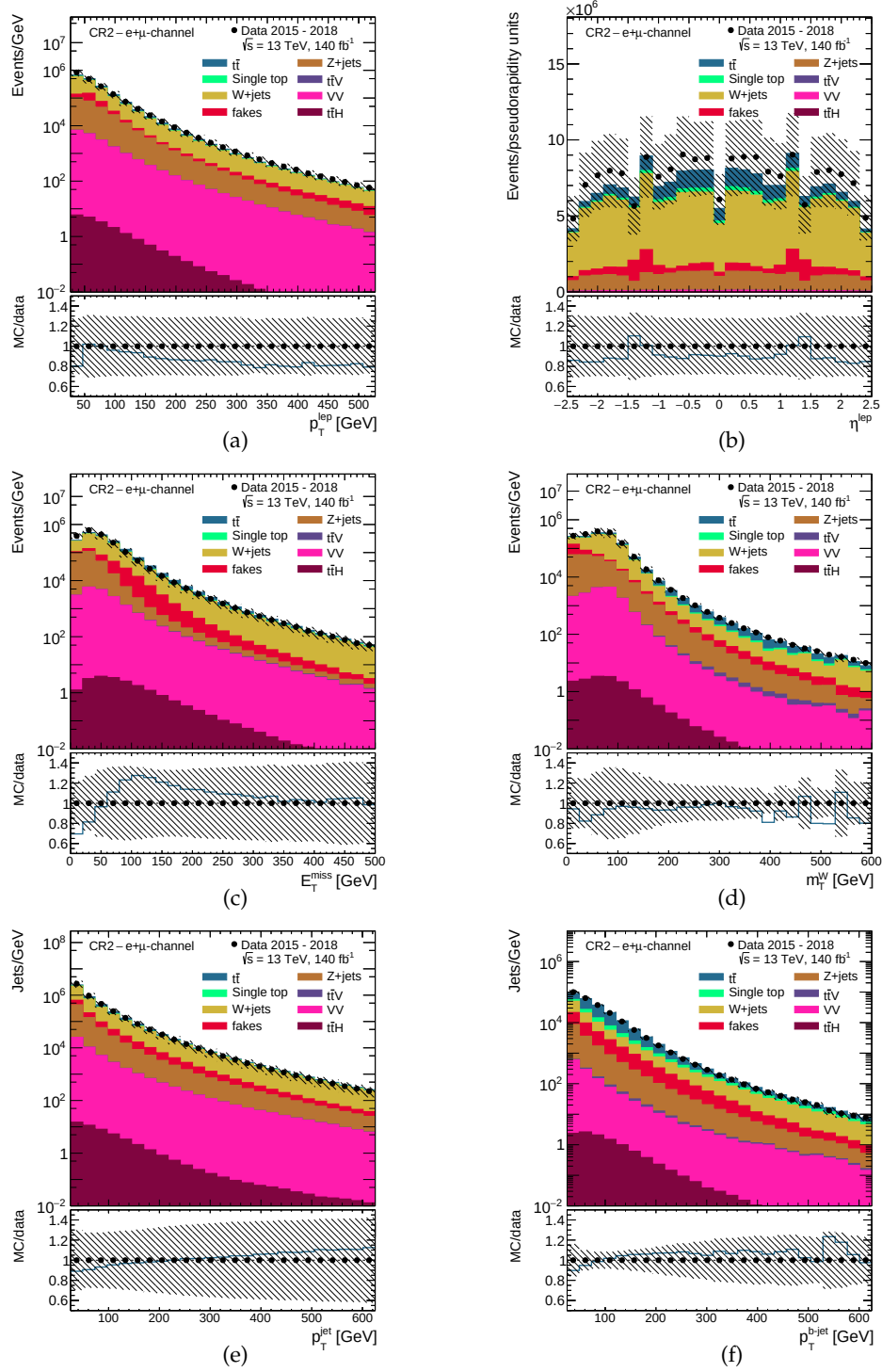


Figure E.2: The measured (a) p_T^{lep} , (b) η^{lep} , (c) E_T^{miss} , (d) m_T^W , (e) p_T^{jet} , and (f) $p_T^{b\text{-jet}}$ distributions (dots) in the $t\bar{t}$ inclusive channel for control region 2. The MC simulation of the signal (blue histograms) and various backgrounds are also included. The lower part of each figure shows the ratio of the total SM distribution to the data. The error bars represent the statistical uncertainty of the data and the hatched bands represent the uncertainty on the normalisation of the W +jets, Z +jets and fakes contributions.

E Control Regions for $t\bar{t}$ production

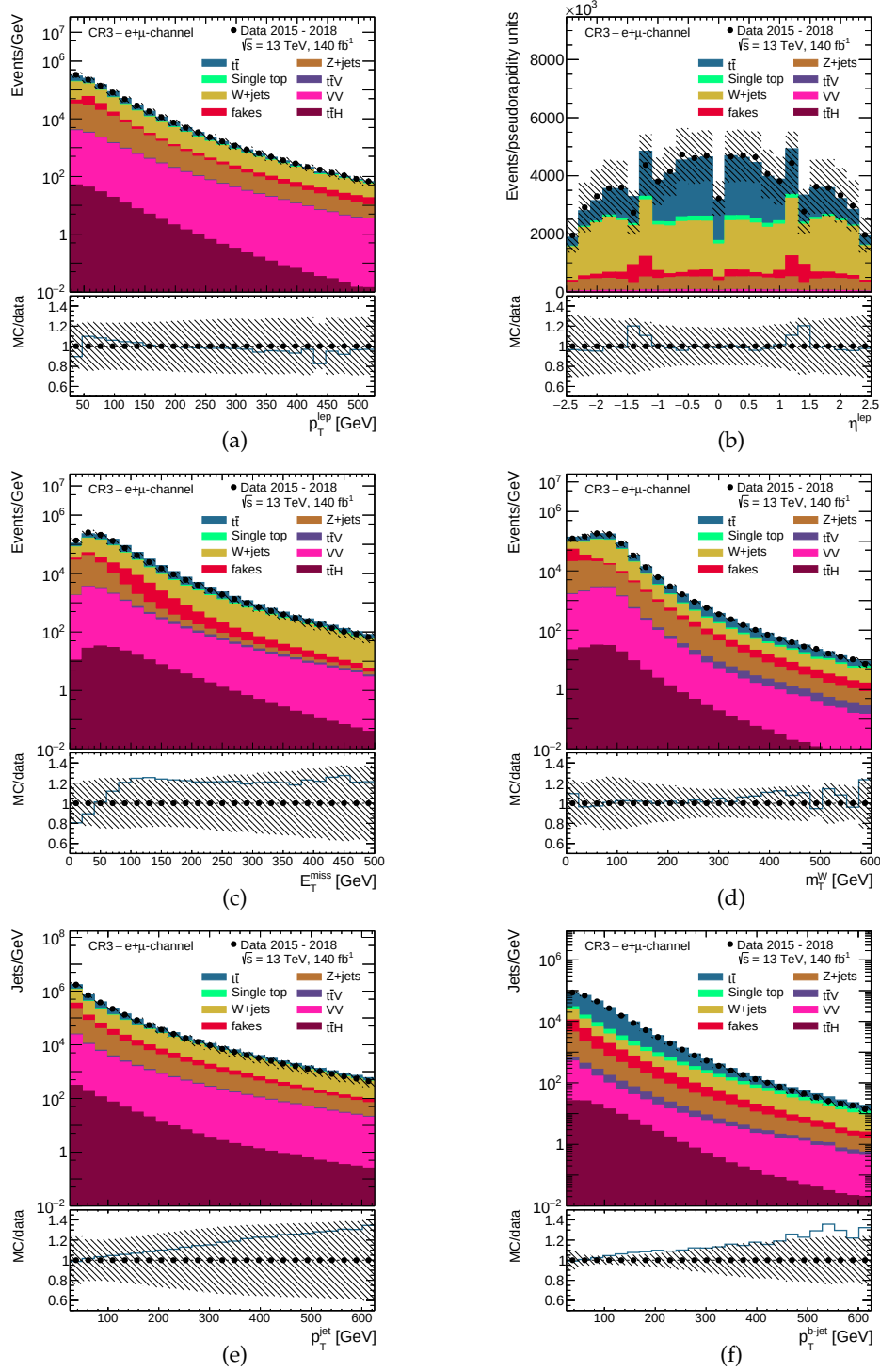


Figure E.3: The measured (a) p_T^{lep} , (b) η^{lep} , (c) E_T^{miss} , (d) m_T^W , (e) p_T^{jet} , and (f) $p_T^{\text{b-jet}}$ distributions (dots) in the $t\bar{t}$ inclusive channel for control region 3. The MC simulation of the signal (blue histograms) and various backgrounds are also included. The lower part of each figure shows the ratio of the total SM distribution to the data. The error bars represent the statistical uncertainty of the data and the hatched bands represent the uncertainty on the normalisation of the W+jets, Z+jets and fakes contributions.

Bibliography

- [1] D0 Collaboration, *Observation of the top quark*, *Phys. Rev. Lett.* **74** (1995) 2632, arXiv: [hep-ex/9503003](#).
- [2] CDF collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, *Phys. Rev. Lett.* **74** (1995) 2626, arXiv: [hep-ex/9503002](#).
- [3] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section in pp collisions at $\sqrt{s} = 5.02$ TeV with the ATLAS detector*, *JHEP* **06** (2023) 138, arXiv: [2207.01354 \[hep-ex\]](#).
- [4] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b -tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector*, *Eur. Phys. J. C* **74** (2014) 3109, arXiv: [1406.5375 \[hep-ex\]](#), Addendum: *Eur. Phys. J. C* **76** (2016) 642.
- [5] ATLAS Collaboration, *Measurement of the inclusive $t\bar{t}$ production cross section in the lepton + jets channel in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using support vector machines*, *Phys. Rev. D* **108** (2023) 032014, arXiv: [2212.00571 \[hep-ex\]](#).
- [6] ATLAS Collaboration, *Measurement of the inclusive and fiducial $t\bar{t}$ production cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **78** (2018) 487, arXiv: [1712.06857 \[hep-ex\]](#).
- [7] ATLAS Collaboration, *Inclusive and differential cross-sections for dilepton $t\bar{t}$ production measured in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, *JHEP* **07** (2023) 141, arXiv: [2303.15340 \[hep-ex\]](#).
- [8] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, *Phys. Lett. B* **810** (2020) 135797, arXiv: [2006.13076 \[hep-ex\]](#).
- [9] ATLAS Collaboration, *Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *JHEP* **01** (2021) 033, arXiv: [2006.09274 \[hep-ex\]](#).
- [10] ATLAS Collaboration, *Measurement of the $t\bar{t}$ cross section and its ratio to the Z production cross section using pp collisions at $\sqrt{s} = 13.6$ TeV with the ATLAS detector*, *Phys. Lett. B* **848** (2024) 138376, arXiv: [2308.09529 \[hep-ex\]](#).

Bibliography

- [11] 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*, [*Phys. Lett. B* **710** \(2012\) 612](#), arXiv: [1111.5869 \[hep-ph\]](#).
- [12] P. Bärnreuther, M. Czakon and A. Mitov, *Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \rightarrow t\bar{t} + X$* , [*Phys. Rev. Lett.* **109** \(2012\) 132001](#), arXiv: [1204.5201 \[hep-ph\]](#).
- [13] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, [*JHEP* **12** \(2012\) 054](#), arXiv: [1207.0236 \[hep-ph\]](#).
- [14] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, [*JHEP* **01** \(2013\) 080](#), arXiv: [1210.6832 \[hep-ph\]](#).
- [15] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_s^4)$* , [*Phys. Rev. Lett.* **110** \(2013\) 252004](#), arXiv: [1303.6254 \[hep-ph\]](#).
- [16] M. Czakon and A. Mitov, *Top++: A program for the calculation of the top-pair cross-section at hadron colliders*, [*Comput. Phys. Commun.* **185** \(2014\) 2930](#), arXiv: [1112.5675 \[hep-ph\]](#).
- [17] PDF4LHC Working Group, *The PDF4LHC21 combination of global PDF fits for the LHC Run III*, [*J. Phys. G* **49** \(2022\) 080501](#), arXiv: [2203.05506 \[hep-ph\]](#).
- [18] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section and lepton differential distributions in $e\mu$ dilepton events from pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [*Eur. Phys. J. C* **80** \(2020\) 528](#), arXiv: [1910.08819 \[hep-ex\]](#).
- [19] ATLAS Collaboration, *Determination of the top-quark pole mass using $t\bar{t} + 1$ -jet events collected with the ATLAS experiment in 7 TeV pp collisions*, [*JHEP* **10** \(2015\) 121](#), arXiv: [1507.01769 \[hep-ex\]](#).
- [20] ATLAS Collaboration, *Measurement of the top-quark mass in $t\bar{t} + 1$ -jet events collected with the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV*, [*JHEP* **11** \(2019\) 150](#), arXiv: [1905.02302 \[hep-ex\]](#).
- [21] ATLAS Collaboration, *Measurement of lepton differential distributions and the top quark mass in $t\bar{t}$ production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, [*Eur. Phys. J. C* **77** \(2017\) 804](#), arXiv: [1709.09407 \[hep-ex\]](#).
- [22] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \rightarrow \text{lepton+jets}$ channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results*, [*Eur. Phys. J. C* **79** \(2019\) 290](#), arXiv: [1810.01772 \[hep-ex\]](#).
- [23] ATLAS Collaboration, *Top quark mass and properties summary plots November 2023*, ATL-PHYS-PUB-2023-036, 2023, URL: <https://cds.cern.ch/record/2882248>.

- [24] ATLAS Collaboration, *Search for New Phenomena in $t\bar{t}$ Events With Large Missing Transverse Momentum in Proton–Proton Collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector*, [Phys. Rev. Lett. **108** \(2012\) 041805](#), arXiv: [1109.4725 \[hep-ex\]](#).
- [25] ATLAS Collaboration, *Search for same-sign top-quark production and fourth-generation down-type quarks in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, [JHEP **04** \(2012\) 069](#), arXiv: [1202.5520 \[hep-ex\]](#).
- [26] ATLAS Collaboration, *A search for $t\bar{t}$ resonances with the ATLAS detector in 2.05fb^{-1} of proton–proton collisions at $\sqrt{s} = 7$ TeV*, [Eur. Phys. J. C **72** \(2012\) 2083](#), arXiv: [1205.5371 \[hep-ex\]](#).
- [27] ATLAS Collaboration, *A search for $t\bar{t}$ resonances in lepton+jets events with highly boosted top quarks collected in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, [JHEP **09** \(2012\) 041](#), arXiv: [1207.2409 \[hep-ex\]](#).
- [28] ATLAS Collaboration, *Search for resonances decaying into top-quark pairs using fully hadronic decays in pp collisions with ATLAS at $\sqrt{s} = 7$ TeV*, [JHEP **01** \(2013\) 116](#), arXiv: [1211.2202 \[hep-ex\]](#).
- [29] ATLAS Collaboration, *Search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 4.7fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV*, [Phys. Rev. D **88** \(2013\) 012004](#), arXiv: [1305.2756 \[hep-ex\]](#).
- [30] ATLAS Collaboration, *A search for $t\bar{t}$ resonances using lepton-plus-jets events in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, [JHEP **08** \(2015\) 148](#), arXiv: [1505.07018 \[hep-ex\]](#).
- [31] ATLAS Collaboration, *Measurement of the polarisation of W bosons produced in top-quark decays using dilepton events at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, [Phys. Lett. B **843** \(2023\) 137829](#), arXiv: [2209.14903 \[hep-ex\]](#).
- [32] ATLAS Collaboration, *Observation of spin correlation in $t\bar{t}$ events from pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, [Phys. Rev. Lett. **108** \(2012\) 212001](#), arXiv: [1203.4081 \[hep-ex\]](#).
- [33] ATLAS Collaboration, *Measurements of spin correlation in top–antitop quark events from proton–proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, [Phys. Rev. D **90** \(2014\) 112016](#), arXiv: [1407.4314 \[hep-ex\]](#).
- [34] ATLAS Collaboration, *Measurement of Spin Correlation in Top–Antitop Quark Events and Search for Top Squark Pair Production in pp Collisions at $\sqrt{s} = 8$ TeV Using the ATLAS Detector*, [Phys. Rev. Lett. **114** \(2015\) 142001](#), arXiv: [1412.4742 \[hep-ex\]](#).
- [35] ATLAS Collaboration, *Measurements of top quark spin observables in $t\bar{t}$ events using dilepton final states in $\sqrt{s} = 8$ TeV pp collisions with the ATLAS detector*, [JHEP **03** \(2017\) 113](#), arXiv: [1612.07004 \[hep-ex\]](#).

Bibliography

- [36] ATLAS Collaboration, *Observation of quantum entanglement in top-quark pairs using the ATLAS detector*, (2023), arXiv: [2311.07288 \[hep-ex\]](#).
- [37] ATLAS Collaboration, *Test of the universality of τ and μ lepton couplings in W-boson decays with the ATLAS detector*, [Nature Phys. **17** \(2021\) 813](#), arXiv: [2007.14040 \[hep-ex\]](#).
- [38] ATLAS Collaboration, *Measurement of colour flow with the jet pull angle in $t\bar{t}$ events using the ATLAS detector at $\sqrt{s} = 8$ TeV*, [Phys. Lett. B **750** \(2015\) 475](#), arXiv: [1506.05629 \[hep-ex\]](#).
- [39] ATLAS Collaboration, *Measurement of colour flow using jet-pull observables in $t\bar{t}$ events with the ATLAS experiment at $\sqrt{s} = 13$ TeV*, [Eur. Phys. J. C **78** \(2018\) 847](#), arXiv: [1805.02935 \[hep-ex\]](#).
- [40] ATLAS Collaboration, *Measurements of observables sensitive to colour reconnection in $t\bar{t}$ events with the ATLAS detector at $\sqrt{s} = 13$ TeV*, [Eur. Phys. J. C **83** \(2023\) 518](#), arXiv: [2209.07874 \[hep-ex\]](#).
- [41] ATLAS Collaboration, *Measurement of the charge asymmetry in top quark pair production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, [Eur. Phys. J. C **72** \(2012\) 2039](#), arXiv: [1203.4211 \[hep-ex\]](#).
- [42] ATLAS Collaboration, *Measurement of the top quark charge in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, [JHEP **11** \(2013\) 031](#), arXiv: [1307.4568 \[hep-ex\]](#).
- [43] ATLAS Collaboration, *Measurement of the top quark pair production charge asymmetry in proton–proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, [JHEP **02** \(2014\) 107](#), arXiv: [1311.6724 \[hep-ex\]](#).
- [44] ATLAS Collaboration, *Measurement of the charge asymmetry in dileptonic decays of top quark pairs in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, [JHEP **05** \(2015\) 061](#), arXiv: [1501.07383 \[hep-ex\]](#).
- [45] ATLAS Collaboration, *Measurement of the charge asymmetry in top-quark pair production in the lepton-plus-jets final state in pp collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector*, [Eur. Phys. J. C **76** \(2016\) 87](#), arXiv: [1509.02358 \[hep-ex\]](#), Erratum: [Eur. Phys. J. C **77** \(2017\) 564](#).
- [46] ATLAS Collaboration, *Measurement of the charge asymmetry in highly boosted top-quark pair production in $\sqrt{s} = 8$ TeV pp collision data collected by the ATLAS experiment*, [Phys. Lett. B **756** \(2016\) 52](#), arXiv: [1512.06092 \[hep-ex\]](#).
- [47] ATLAS Collaboration, *Measurements of the charge asymmetry in top-quark pair production in the dilepton final state at $\sqrt{s} = 8$ TeV with the ATLAS detector*, [Phys. Rev. D **94** \(2016\) 032006](#), arXiv: [1604.05538 \[hep-ex\]](#).

- [48] ATLAS Collaboration, *Evidence for the charge asymmetry in $pp \rightarrow t\bar{t}$ production at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [JHEP **08** \(2023\) 077](#), [arXiv: 2208.12095 \[hep-ex\]](#).
- [49] ATLAS Collaboration, *Measurement of the energy asymmetry in $t\bar{t}j$ production at 13 TeV with the ATLAS experiment and interpretation in the SMEFT framework*, [Eur. Phys. J. C **82** \(2022\) 374](#), [arXiv: 2110.05453 \[hep-ex\]](#).
- [50] ATLAS Collaboration, *Measurement of the top quark branching ratios into channels with leptons and quarks with the ATLAS detector*, [Phys. Rev. D **92** \(2015\) 072005](#), [arXiv: 1506.05074 \[hep-ex\]](#).
- [51] ATLAS Collaboration, *Direct top-quark decay width measurement in the $t\bar{t}$ lepton+jets channel at $\sqrt{s} = 8$ TeV with the ATLAS experiment*, [Eur. Phys. J. C **78** \(2018\) 129](#), [arXiv: 1709.04207 \[hep-ex\]](#).
- [52] ATLAS Collaboration, *Measurements of top quark pair relative differential cross-sections with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV*, [Eur. Phys. J. C **73** \(2013\) 2261](#), [arXiv: 1207.5644 \[hep-ex\]](#).
- [53] ATLAS Collaboration, *Measurements of normalized differential cross-sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, [Phys. Rev. D **90** \(2014\) 072004](#), [arXiv: 1407.0371 \[hep-ex\]](#).
- [54] ATLAS Collaboration, *Differential top-antitop cross-section measurements as a function of observables constructed from final-state particles using pp collisions at $\sqrt{s} = 7$ TeV in the ATLAS detector*, [JHEP **06** \(2015\) 100](#), [arXiv: 1502.05923 \[hep-ex\]](#).
- [55] ATLAS Collaboration, *Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in $\sqrt{s} = 8$ TeV proton–proton collisions using the ATLAS detector*, [Phys. Rev. D **93** \(2016\) 032009](#), [arXiv: 1510.03818 \[hep-ex\]](#).
- [56] ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, [Eur. Phys. J. C **76** \(2016\) 538](#), [arXiv: 1511.04716 \[hep-ex\]](#).
- [57] ATLAS Collaboration, *Measurement of top quark pair differential cross sections in the dilepton channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV with ATLAS*, [Phys. Rev. D **94** \(2016\) 092003](#), [arXiv: 1607.07281 \[hep-ex\]](#).
- [58] ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the $e\mu$ channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, [Eur. Phys. J. C **77** \(2017\) 292](#), [arXiv: 1612.05220 \[hep-ex\]](#).
- [59] ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, [JHEP **11** \(2017\) 191](#), [arXiv: 1708.00727 \[hep-ex\]](#).

Bibliography

- [60] ATLAS Collaboration, *Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, [*Phys. Rev. D* **98** \(2018\) 012003](#), arXiv: [1801.02052 \[hep-ex\]](#).
- [61] 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*, [*Eur. Phys. J. C* **79** \(2019\) 1028](#), arXiv: [1908.07305 \[hep-ex\]](#),
Erratum: [*Eur. Phys. J. C* **80** \(2020\) 1092](#).
- [62] 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 $\sqrt{s} = 13$ TeV*, [*JHEP* **06** \(2022\) 063](#), arXiv: [2202.12134 \[hep-ex\]](#).
- [63] ATLAS Collaboration, *Differential $t\bar{t}$ cross-section measurements using boosted top quarks in the all-hadronic final state with 139 fb^{-1} of ATLAS data*, [*JHEP* **04** \(2023\) 080](#), arXiv: [2205.02817 \[hep-ex\]](#).
- [64] ATLAS Collaboration, *Measurement of $t\bar{t}$ production with a veto on additional central jet activity in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, [*Eur. Phys. J. C* **72** \(2012\) 2043](#), arXiv: [1203.5015 \[hep-ex\]](#).
- [65] ATLAS Collaboration, *Measurement of jet substructure in boosted $t\bar{t}$ events with the ATLAS detector using 140 fb^{-1} of 13 TeV pp collisions*, (2023), arXiv: [2312.03797 \[hep-ex\]](#).
- [66] ATLAS Collaboration, *Study of heavy flavor quarks produced in association with top quark pairs at $\sqrt{s} = 7$ TeV using the ATLAS detector*, [*Phys. Rev. D* **89** \(2014\) 072012](#), arXiv: [1304.6386 \[hep-ex\]](#).
- [67] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section as a function of jet multiplicity and jet transverse momentum in 7 TeV proton–proton collisions with the ATLAS detector*, [*JHEP* **01** \(2015\) 020](#), arXiv: [1407.0891 \[hep-ex\]](#).
- [68] ATLAS Collaboration, *Measurements of fiducial cross-sections for $t\bar{t}$ production with one or two additional b -jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, [*Eur. Phys. J. C* **76** \(2016\) 11](#), arXiv: [1508.06868 \[hep-ex\]](#).
- [69] ATLAS Collaboration, *Measurement of jet activity in top quark events using the $e\mu$ final state with two b -tagged jets in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, [*JHEP* **09** \(2016\) 074](#), arXiv: [1606.09490 \[hep-ex\]](#).
- [70] ATLAS Collaboration, *Measurement of jet activity produced in top-quark events with an electron, a muon and two b -tagged jets in the final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [*Eur. Phys. J. C* **77** \(2017\) 220](#), arXiv: [1610.09978 \[hep-ex\]](#).

- [71] ATLAS Collaboration, *Measurements of differential cross sections of top quark pair production in association with jets in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, [JHEP **10** \(2018\) 159](#), arXiv: [1802.06572 \[hep-ex\]](#).
- [72] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [JHEP **04** \(2019\) 046](#), arXiv: [1811.12113 \[hep-ex\]](#).
- [73] ATLAS Collaboration, *Measurements of jet observables sensitive to b -quark fragmentation in $t\bar{t}$ events at the LHC with the ATLAS detector*, [Phys. Rev. D **106** \(2022\) 032008](#), arXiv: [2202.13901 \[hep-ex\]](#).
- [74] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and collider physics*, Cambridge University Press, 1996.
- [75] S. L. Glashow, *Partial Symmetries of Weak Interactions*, [Nucl. Phys. **22** \(1961\) 579](#).
- [76] S. Weinberg, *A Model of Leptons*, [Phys. Rev. Lett. **19** \(1967\) 1264](#).
- [77] A. Salam, *Weak and Electromagnetic Interactions*, [Conf. Proc. C **680519** \(1968\) 367](#).
- [78] R. Atkin, *Review of jet reconstruction algorithms*, [J. Phys. Conf. Ser. **645** \(2015\) 012008](#), ed. by A. S. Cornell and B. Mellado.
- [79] T. Sjöstrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 physics and manual*, [JHEP **05** \(2006\) 026](#), arXiv: [hep-ph/0603175](#).
- [80] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, [Phys. Rev. Lett. **13** \(1964\) 321](#).
- [81] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, [Phys. Rev. Lett. **13** \(1964\) 508](#), ed. by J. C. Taylor.
- [82] A. Purcell, *Go on a particle quest at the first CERN webfest. Le premier webfest du CERN se lance à la conquête des particules*, (2012) 10, URL: <https://cds.cern.ch/record/1473657>.
- [83] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, [Phys. Lett. **8** \(1964\) 214](#).
- [84] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 1*, (1964), URL: <https://cds.cern.ch/record/352337>.
- [85] J. D. Bjorken and S. L. Glashow, *Elementary Particles and $SU(4)$* , [Phys. Lett. **11** \(1964\) 255](#).
- [86] H. Harari, *A New Quark Model for Hadrons*, [Phys. Lett. B **57** \(1975\) 265](#).
- [87] O. W. Greenberg, *Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons*, [Phys. Rev. Lett. **13** \(1964\) 598](#).
- [88] M. Y. Han and Y. Nambu, *Three Triplet Model with Double $SU(3)$ Symmetry*, [Phys. Rev. **139** \(1965\) B1006](#), ed. by T. Eguchi.

Bibliography

- [89] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, *Phys. Lett. B* **47** (1973) 365.
- [90] E. C. G. Stueckelberg and A. Petermann, *The normalization group in quantum theory*, *Helv. Phys. Acta* **24** (1951) 317, URL: <https://www.e-periodica.ch/digbib/view?pid=hpa-001:1951:24::317#319>.
- [91] R. L. Workman et al., *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.
- [92] R. P. Feynman, *The behavior of hadron collisions at extreme energies*, *Conf. Proc. C* **690905** (1969) 237.
- [93] J. D. Bjorken and E. A. Paschos, *Inelastic Electron Proton and gamma Proton Scattering, and the Structure of the Nucleon*, *Phys. Rev.* **185** (1969) 1975.
- [94] R. D. Ball et al., *The path to proton structure at 1% accuracy*, *Eur. Phys. J. C* **82** (2022) 428, arXiv: 2109.02653 [hep-ph].
- [95] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin and R. S. Thorne, *Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs*, *Eur. Phys. J. C* **81** (2021) 341, arXiv: 2012.04684 [hep-ph].
- [96] G. Arnison et al., *Hadronic Jet Production at the CERN Proton - anti-Proton Collider*, *Phys. Lett. B* **132** (1983) 214.
- [97] G. C. Blazey et al., 'Run II jet physics', *Physics at Run II: QCD and Weak Boson Physics Workshop: Final General Meeting*, 2000 47, arXiv: [hep-ex/0005012](https://arxiv.org/abs/hep-ex/0005012).
- [98] G. P. Salam and G. Soyez, *A Practical Seedless Infrared-Safe Cone jet algorithm*, *JHEP* **05** (2007) 086, arXiv: 0704.0292 [hep-ph].
- [99] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, *Nucl. Phys. B* **406** (1993) 187.
- [100] CMS Collaboration, *A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging*, (2009), URL: <https://cds.cern.ch/record/1194489>.
- [101] M. Cacciari, G. P. Salam and G. Soyez, *The anti- k_t jet clustering algorithm*, *JHEP* **04** (2008) 063, arXiv: 0802.1189 [hep-ph].
- [102] M. Jezabek and J. H. Kuhn, *QCD Corrections to Semileptonic Decays of Heavy Quarks*, *Nucl. Phys. B* **314** (1989) 1.
- [103] S. Catani, S. Devoto, M. Grazzini, S. Kallweit and J. Mazzitelli, *Top-quark pair production at the LHC: Fully differential QCD predictions at NNLO*, *JHEP* **07** (2019) 100, arXiv: 1906.06535 [hep-ph].

- [104] J. Campbell, T. Neumann and Z. Sullivan,
Single-top-quark production in the t -channel at NNLO, **JHEP** **02** (2021) 040,
arXiv: [2012.01574 \[hep-ph\]](#).
- [105] N. Kidonakis and N. Yamanaka,
Higher-order corrections for tW production at high-energy hadron colliders,
JHEP **05** (2021) 278, arXiv: [2102.11300 \[hep-ph\]](#).
- [106] N. Kidonakis, *NNLL resummation for s -channel single top quark production*,
Phys. Rev. D **81** (2010) 054028, arXiv: [1001.5034 \[hep-ph\]](#).
- [107] A. Broggio et al., *Top-quark pair hadroproduction in association with a heavy boson at NLO+NNLL including EW corrections*, **JHEP** **08** (2019) 039,
arXiv: [1907.04343 \[hep-ph\]](#).
- [108] M. van Beekveld, A. Kulesza and L. M. Valero,
Threshold Resummation for the Production of Four Top Quarks at the LHC,
Phys. Rev. Lett. **131** (2023) 211901, arXiv: [2212.03259 \[hep-ph\]](#).
- [109] L. Buonocore et al., *Precise Predictions for the Associated Production of a W Boson with a Top-Antitop Quark Pair at the LHC*, **Phys. Rev. Lett.** **131** (2023) 231901,
arXiv: [2306.16311 \[hep-ph\]](#).
- [110] S. Catani et al., *Higgs Boson Production in Association with a Top-Antitop Quark Pair in Next-to-Next-to-Leading Order QCD*, **Phys. Rev. Lett.** **130** (2023) 111902,
arXiv: [2210.07846 \[hep-ph\]](#).
- [111] T. Gleisberg et al., *Event generation with SHERPA 1.1*, **JHEP** **02** (2009) 007,
arXiv: [0811.4622 \[hep-ph\]](#).
- [112] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4 – a simulation toolkit*,
Nucl. Instrum. Meth. A **506** (2003) 250.
- [113] Z. Marshall and the ATLAS Collaboration,
Simulation of Pile-up in the ATLAS Experiment,
J. Phys. Conf. Ser. **513** (2014) 022024, ed. by D. L. Groep and D. Bonacorsi.
- [114] O. S. Brüning et al., *LHC Design Report*, vol. 1: the LHC Main Ring,
CERN Yellow Reports: Monographs, Geneva: CERN, 2004,
URL: <https://cds.cern.ch/record/782076>.
- [115] O. S. Brüning et al., *LHC Design Report*,
vol. 2: the LHC Infrastructure and General Services,
CERN Yellow Reports: Monographs, Geneva: CERN, 2004,
URL: <https://cds.cern.ch/record/815187>.
- [116] M. Benedikt, P. Collier, V. Mertens, J. Poole and K. Schindl, *LHC Design Report*,
vol. 3: the LHC Injector Chain, CERN Yellow Reports: Monographs,
Geneva: CERN, 2004, URL: <https://cds.cern.ch/record/823808>.
- [117] L. Evans and P. Bryant, *LHC Machine*, **JINST** **3** (2008) S08001.

Bibliography

- [118] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, **JINST 3 (2008) S08003**.
- [119] *LEP design report*, vol. 1: The LEP injector chain, Report, Geneva: CERN, 1983, URL: <https://cds.cern.ch/record/98881>.
- [120] *LEP design report*, vol. 2: The LEP main ring, Report, Geneva: CERN, 1984, URL: <https://cds.cern.ch/record/102083>.
- [121] E. Mobs, *The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018*, 2018, URL: <https://cds.cern.ch/record/2636343>.
- [122] CMS Collaboration, *The CMS Experiment at the CERN LHC*, **JINST 3 (2008) S08004**.
- [123] LHCb Collaboration, *The LHCb Detector at the LHC*, **JINST 3 (2008) S08005**.
- [124] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, **JINST 3 (2008) S08002**.
- [125] J. Wenninger, *The LHC collider*, **Comptes Rendus Physique 16 (2015) 347**.
- [126] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, **JINST 13 (2018) P07017**.
- [127] S. Abdel Khalek et al., *The ALFA Roman Pot Detectors of ATLAS*, **JINST 11 (2016) P11013**, arXiv: [1609.00249 \[physics.ins-det\]](https://arxiv.org/abs/1609.00249).
- [128] ATLAS Collaboration, *Zero Degree Calorimeters for ATLAS*, (2007), URL: <https://cds.cern.ch/record/1009649>.
- [129] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, **Eur. Phys. J. C 83 (2023) 982**, arXiv: [2212.09379 \[hep-ex\]](https://arxiv.org/abs/2212.09379).
- [130] ATLAS Collaboration, *Public ATLAS Luminosity Results for Run-2 of the LHC*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Integrated_and_Instantaneous_Lum.
- [131] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, **Eur. Phys. J. C 79 (2019) 639**, arXiv: [1902.04655 \[physics.ins-det\]](https://arxiv.org/abs/1902.04655).
- [132] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, **JINST 14 (2019) P12006**, arXiv: [1908.00005 \[hep-ex\]](https://arxiv.org/abs/1908.00005).
- [133] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, **Eur. Phys. J. C 77 (2017) 490**, arXiv: [1603.02934 \[hep-ex\]](https://arxiv.org/abs/1603.02934).
- [134] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002, 2008, URL: <https://cds.cern.ch/record/1099735>.

- [135] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data*, *JINST* **14** (2019) P03017, arXiv: 1812.03848 [hep-ex].
- [136] M. Cacciari, G. P. Salam and S. Sapeta, *On the characterisation of the underlying event*, *JHEP* **04** (2010) 065, arXiv: 0912.4926 [hep-ph].
- [137] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton–proton collision data*, *Eur. Phys. J. C* **72** (2012) 1909, arXiv: 1110.3174 [hep-ex].
- [138] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **81** (2021) 578, arXiv: 2012.00578 [hep-ex].
- [139] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **76** (2016) 292, arXiv: 1603.05598 [hep-ex].
- [140] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) 466, arXiv: 1703.10485 [hep-ex].
- [141] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) 689, arXiv: 2007.02645 [hep-ex].
- [142] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) 581, arXiv: 1510.03823 [hep-ex].
- [143] ATLAS Collaboration, *ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset*, *Eur. Phys. J. C* **83** (2023) 681, arXiv: 2211.16345 [physics.data-an].
- [144] ATLAS Collaboration, *Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, 2016, URL: <https://cds.cern.ch/record/2160731>.
- [145] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **78** (2018) 903, arXiv: 1802.08168 [hep-ex].
- [146] ATLAS Collaboration, *E_T^{miss} performance in the ATLAS detector using 2015–2016 LHC pp collisions*, ATLAS-CONF-2018-023, 2018, URL: <https://cds.cern.ch/record/2625233>.
- [147] M. Cacciari and G. P. Salam, *Dispelling the N^3 myth for the k_t jet-finder*, *Phys. Lett. B* **641** (2006) 57, arXiv: hep-ph/0512210.

Bibliography

- [148] ATLAS Collaboration, *Jet energy measurement with the ATLAS detector in proton–proton collisions at $\sqrt{s} = 7$ TeV*, [*Eur. Phys. J. C* **73** \(2013\) 2304](#), arXiv: [1112.6426 \[hep-ex\]](#).
- [149] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, [*Eur. Phys. J. C* **75** \(2015\) 17](#), arXiv: [1406.0076 \[hep-ex\]](#).
- [150] ATLAS Collaboration, *Determination of jet calibration and energy resolution in proton–proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, [*Eur. Phys. J. C* **80** \(2020\) 1104](#), arXiv: [1910.04482 \[hep-ex\]](#).
- [151] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [*Phys. Rev. D* **96** \(2017\) 072002](#), arXiv: [1703.09665 \[hep-ex\]](#).
- [152] ATLAS Collaboration, *In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector*, [*Eur. Phys. J. C* **79** \(2019\) 135](#), arXiv: [1807.09477 \[hep-ex\]](#).
- [153] ATLAS Collaboration, *New techniques for jet calibration with the ATLAS detector*, [*Eur. Phys. J. C* **83** \(2023\) 761](#), arXiv: [2303.17312 \[hep-ex\]](#).
- [154] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, [*JINST* **15** \(2020\) P04003](#), arXiv: [1911.04632 \[physics.ins-det\]](#).
- [155] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, [*Comput. Phys. Commun.* **191** \(2015\) 159](#), arXiv: [1410.3012 \[hep-ph\]](#).
- [156] R. D. Ball et al., *Parton distributions with LHC data*, [*Nucl. Phys. B* **867** \(2013\) 244](#), arXiv: [1207.1303 \[hep-ph\]](#).
- [157] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: <https://cds.cern.ch/record/1966419>.
- [158] B. Andersson, G. Gustafson, G. Ingelman and T. Sjöstrand, *Parton fragmentation and string dynamics*, [*Phys. Rept.* **97** \(1983\) 31](#).
- [159] T. Sjöstrand, *Jet fragmentation of multiparton configurations in a string framework*, [*Nucl. Phys. B* **248** \(1984\) 469](#).
- [160] E. Bothmann et al., *Event generation with Sherpa 2.2*, [*SciPost Phys.* **7** \(2019\) 034](#), arXiv: [1905.09127 \[hep-ph\]](#).
- [161] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, [*JHEP* **12** \(2008\) 039](#), arXiv: [0808.3674 \[hep-ph\]](#).
- [162] F. Buccioni et al., *OpenLoops 2*, [*Eur. Phys. J. C* **79** \(2019\) 866](#), arXiv: [1907.13071 \[hep-ph\]](#).
- [163] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, [*Phys. Rev. Lett.* **108** \(2012\) 111601](#), arXiv: [1111.5206 \[hep-ph\]](#).

- [164] A. Denner, S. Dittmaier and L. Hofer,
COLLIER: *A fortran-based complex one-loop library in extended regularizations*,
[Comput. Phys. Commun.](#) **212** (2017) 220, arXiv: 1604.06792 [hep-ph].
- [165] S. Schumann and F. Krauss,
A parton shower algorithm based on Catani–Seymour dipole factorisation,
[JHEP](#) **03** (2008) 038, arXiv: 0709.1027 [hep-ph].
- [166] J.-C. Winter, F. Krauss and G. Soff, *A modified cluster-hadronization model*,
[Eur. Phys. J. C](#) **36** (2004) 381, arXiv: hep-ph/0311085.
- [167] R. D. Ball et al., *Parton distributions for the LHC run II*, [JHEP](#) **04** (2015) 040,
arXiv: 1410.8849 [hep-ph].
- [168] S. Höche, F. Krauss, M. Schönherr and F. Siegert,
A critical appraisal of NLO+PS matching methods, [JHEP](#) **09** (2012) 049,
arXiv: 1111.1220 [hep-ph].
- [169] S. Catani, F. Krauss, B. R. Webber and R. Kuhn,
QCD Matrix Elements + Parton Showers, [JHEP](#) **11** (2001) 063,
arXiv: hep-ph/0109231.
- [170] S. Höche, F. Krauss, S. Schumann and F. Siegert,
QCD matrix elements and truncated showers, [JHEP](#) **05** (2009) 053,
arXiv: 0903.1219 [hep-ph].
- [171] S. Höche, F. Krauss, M. Schönherr and F. Siegert,
QCD matrix elements + parton showers. The NLO case, [JHEP](#) **04** (2013) 027,
arXiv: 1207.5030 [hep-ph].
- [172] F. Siegert, *A practical guide to event generation for prompt photon production with Sherpa*,
[J. Phys. G](#) **44** (2017) 044007, arXiv: 1611.07226 [hep-ph].
- [173] S. Frixione, *Isolated photons in perturbative QCD*, [Phys. Lett. B](#) **429** (1998) 369,
arXiv: hep-ph/9801442.
- [174] E. Bothmann, M. Schönherr and S. Schumann,
Reweighting QCD matrix-element and parton-shower calculations,
[Eur. Phys. J. C](#) **76** (2016) 590, arXiv: 1606.08753 [hep-ph].
- [175] S. Dulat et al.,
New parton distribution functions from a global analysis of quantum chromodynamics,
[Phys. Rev. D](#) **93** (2016) 033006, arXiv: 1506.07443 [hep-ph].
- [176] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne,
Parton distributions in the LHC era: MMHT 2014 PDFs, [Eur. Phys. J. C](#) **75** (2015) 204,
arXiv: 1412.3989 [hep-ph].

Bibliography

- [177] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, 2016, URL: <https://cds.cern.ch/record/2206965>.
- [178] ATLAS Collaboration, *Selection of jets produced in 13 TeV proton–proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029, 2015, URL: <https://cds.cern.ch/record/2037702>.
- [179] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: <https://cds.cern.ch/record/1700870>.
- [180] ATLAS Collaboration, *Determination of the jet energy scale and resolution at ATLAS using Z/γ -jet events in data at $\sqrt{s} = 8$ TeV*, ATLAS-CONF-2015-057, 2015, URL: <https://cds.cern.ch/record/2059846>.
- [181] ATLAS Collaboration, *Dependence of the Jet Energy Scale on the Particle Content of Hadronic Jets in the ATLAS Detector Simulation*, ATL-PHYS-PUB-2022-021, 2022, URL: <https://cds.cern.ch/record/2808016>.
- [182] ATLAS Collaboration, *Inclusive-photon production and its dependence on photon isolation in pp collisions at $\sqrt{s} = 13$ TeV using 139fb^{-1} of ATLAS data*, *JHEP* **07** (2023) 086, arXiv: 2302.00510 [hep-ex].
- [183] ATLAS Collaboration, *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **08** (2016) 005, arXiv: 1605.03495 [hep-ex].
- [184] ATLAS Collaboration, *Measurement of the cross section for inclusive isolated-photon production in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, *Phys. Lett. B* **770** (2017) 473, arXiv: 1701.06882 [hep-ex].
- [185] ATLAS Collaboration, *Measurement of the inclusive isolated-photon cross section in pp collisions at $\sqrt{s} = 13$ TeV using 36fb^{-1} of ATLAS data*, *JHEP* **10** (2019) 203, arXiv: 1908.02746 [hep-ex].
- [186] ATLAS Collaboration, *Measurement of the ratio of cross sections for inclusive isolated-photon production in pp collisions at $\sqrt{s} = 13$ and 8 TeV with the ATLAS detector*, *JHEP* **04** (2019) 093, arXiv: 1901.10075 [hep-ex].
- [187] ATLAS Collaboration, *Measurement of the top-quark mass using a leptonic invariant mass in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **06** (2023) 019, arXiv: 2209.00583 [hep-ex].
- [188] CMS Collaboration, *Measurement of the top quark mass using a profile likelihood approach with the lepton+jets final states in proton–proton collisions at $\sqrt{s} = 13$ TeV*, *Eur. Phys. J. C* **83** (2023) 963, arXiv: 2302.01967 [hep-ex].

- [189] CMS Collaboration, *Measurement of $t\bar{t}$ normalised multi-differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV, and simultaneous determination of the strong coupling strength, top quark pole mass, and parton distribution functions*, *Eur. Phys. J. C* **80** (2020) 658, arXiv: 1904.05237 [hep-ex].
- [190] CMS Collaboration, *Search for physics beyond the standard model in top quark production with additional leptons in the context of effective field theory*, *JHEP* **12** (2023) 068, arXiv: 2307.15761 [hep-ex].
- [191] CMS Collaboration, *Measurement of differential cross sections for the production of top quark pairs and of additional jets in lepton+jets events from pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **97** (2018) 112003, arXiv: 1803.08856 [hep-ex].
- [192] CMS Collaboration, *Measurements of $t\bar{t}$ differential cross sections in proton–proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons*, *JHEP* **02** (2019) 149, arXiv: 1811.06625 [hep-ex].
- [193] 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*, *Phys. Rev. D* **103** (2021) 052008, arXiv: 2008.07860 [hep-ex].
- [194] CMS Collaboration, *Measurement of differential top-quark-pair production cross sections in pp collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J. C* **73** (2013) 2339, arXiv: 1211.2220 [hep-ex].
- [195] CMS Collaboration, *Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV*, *Eur. Phys. J. C* **75** (2015) 542, arXiv: 1505.04480 [hep-ex].
- [196] CMS Collaboration, *Measurement of the differential cross sections for top quark pair production as a function of kinematic event variables in pp collisions at $\sqrt{s} = 7$ and 8 TeV*, *Phys. Rev. D* **94** (2016) 052006, arXiv: 1607.00837 [hep-ex].
- [197] CMS Collaboration, *Measurement of the integrated and differential $t\bar{t}$ production cross sections for high- p_T top quarks in pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Rev. D* **94** (2016) 072002, arXiv: 1605.00116 [hep-ex].
- [198] CMS Collaboration, *Measurement of double-differential cross sections for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV and impact on parton distribution functions*, *Eur. Phys. J. C* **77** (2017) 459, arXiv: 1703.01630 [hep-ex].
- [199] CMS Collaboration, *Measurement of normalized differential $t\bar{t}$ cross sections in the dilepton channel from pp collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **04** (2018) 060, arXiv: 1708.07638 [hep-ex].
- [200] CMS Collaboration, *Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton–proton collisions at 13 TeV*, *Phys. Rev. D* **95** (2017) 092001, arXiv: 1610.04191 [hep-ex].

Bibliography

- [201] CMS Collaboration,
Measurements of differential cross sections of top quark pair production as a function of kinematic event variables in proton–proton collisions at $\sqrt{s} = 13$ TeV, *JHEP* **06** (2018) 002, arXiv: 1803.03991 [hep-ex].
- [202] ATLAS Collaboration, *Further studies on simulation of top-quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2016-016, 2016, URL: <https://cds.cern.ch/record/2205262>.
- [203] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, 2016, URL: <https://cds.cern.ch/record/2216168>.
- [204] ATLAS Collaboration,
Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO, ATL-PHYS-PUB-2017-007, 2017, URL: <https://cds.cern.ch/record/2261938>.
- [205] ATLAS Collaboration,
Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run 2, ATL-PHYS-PUB-2018-009, 2018, URL: <https://cds.cern.ch/record/2630327>.
- [206] CMS Collaboration, *Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements*, *Eur. Phys. J. C* **80** (2020) 4, arXiv: 1903.12179 [hep-ex].
- [207] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J. C* **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det].
- [208] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, 2010, URL: <https://cds.cern.ch/record/1300517>.
- [209] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012, URL: <https://cds.cern.ch/record/1474107>.
- [210] J. Gao, C. S. Li and H. X. Zhu,
Top Quark Decay at Next-to-Next-to Leading Order in QCD, *Phys. Rev. Lett.* **110** (2013) 042001, arXiv: 1210.2808 [hep-ph].
- [211] S. Frixione, G. Ridolfi and P. Nason,
A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction, *JHEP* **09** (2007) 126, arXiv: 0707.3088 [hep-ph].
- [212] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, arXiv: hep-ph/0409146.
- [213] S. Frixione, P. Nason and C. Oleari,
Matching NLO QCD computations with parton shower simulations: the POWHEG method, *JHEP* **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

- [214] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, **JHEP** **06** (2010) 043, arXiv: 1002.2581 [hep-ph].
- [215] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, **Comput. Phys. Commun.** **178** (2008) 852, arXiv: 0710.3820 [hep-ph].
- [216] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, **Eur. Phys. J. C** **76** (2016) 196, arXiv: 1512.01178 [hep-ph].
- [217] M. Bähr et al., *Herwig++ physics and manual*, **Eur. Phys. J. C** **58** (2008) 639, arXiv: 0803.0883 [hep-ph].
- [218] M. H. Seymour and A. Siodmok, *Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data*, **JHEP** **10** (2013) 113, arXiv: 1307.5015 [hep-ph].
- [219] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, **JHEP** **07** (2014) 079, arXiv: 1405.0301 [hep-ph].
- [220] M. Beneke, P. Falgari, S. Klein and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, **Nucl. Phys. B** **855** (2012) 695, arXiv: 1109.1536 [hep-ph].
- [221] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, **J. Phys. G** **43** (2016) 023001, arXiv: 1510.03865 [hep-ph].
- [222] 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*, **Eur. Phys. J. C** **64** (2009) 653, arXiv: 0905.3531 [hep-ph].
- [223] H.-L. Lai et al., *New parton distributions for collider physics*, **Phys. Rev. D** **82** (2010) 074024, arXiv: 1007.2241 [hep-ph].
- [224] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, **JHEP** **07** (2008) 029, arXiv: 0805.3067 [hep-ph].
- [225] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, **Phys. Rev. D** **83** (2011) 091503, arXiv: 1103.2792 [hep-ph].
- [226] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-* , **Phys. Rev. D** **82** (2010) 054018, arXiv: 1005.4451 [hep-ph].
- [227] M. Aliev et al., *HATHOR – HAdronic Top and Heavy quarks crOss section calculator*, **Comput. Phys. Commun.** **182** (2011) 1034, arXiv: 1007.1327 [hep-ph].

Bibliography

- [228] P. Kant et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, [*Comput. Phys. Commun.* **191** \(2015\) 74](#), arXiv: 1406.4403 [hep-ph].
- [229] E. Re, *Single-top Wt -channel production matched with parton showers using the POWHEG method*, [*Eur. Phys. J. C* **71** \(2011\) 1547](#), arXiv: 1009.2450 [hep-ph].
- [230] ATLAS Collaboration, *ATLAS simulation of boson plus jets processes in Run 2*, ATL-PHYS-PUB-2017-006, 2017, URL: <https://cds.cern.ch/record/2261937>.
- [231] 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*, [*Phys. Rev. D* **69** \(2004\) 094008](#), arXiv: hep-ph/0312266.
- [232] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph].
- [233] J. Mazzitelli et al., *Next-to-Next-to-Leading Order Event Generation for Top-Quark Pair Production*, [*Phys. Rev. Lett.* **127** \(2021\) 062001](#), arXiv: 2012.14267 [hep-ph].
- [234] A. Ablat et al., *Exploring the impact of high-precision top-quark pair production data on the structure of the proton at the LHC*, [*Phys. Rev. D* **109** \(2024\) 054027](#), arXiv: 2307.11153 [hep-ph].
- [235] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-026, 2015, URL: <https://cds.cern.ch/record/2037717>.
- [236] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b -tagged jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [*Phys. Lett. B* **761** \(2016\) 136](#), arXiv: 1606.02699 [hep-ex], Erratum: [*Phys. Lett. B* **772** \(2017\) 879](#).
- [237] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, [*Eur. Phys. J. C* **72** \(2012\) 1896](#), arXiv: 1111.6097 [hep-ph].
- [238] M. Cacciari, G. P. Salam and G. Soyez, *The Catchment Area of Jets*, [*JHEP* **04** \(2008\) 005](#), arXiv: 0802.1188 [hep-ph].
- [239] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, [*Eur. Phys. J. C* **80** \(2020\) 47](#), arXiv: 1909.00761 [hep-ex].
- [240] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, [*JINST* **15** \(2020\) P09015](#), arXiv: 2004.13447 [physics.ins-det].

- [241] ATLAS Collaboration, *Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, **JINST** **18** (2023) T11004, arXiv: 2211.16178 [hep-ex].
- [242] G. D’Agostini, *A multidimensional unfolding method based on Bayes’ theorem*, **Nucl. Instrum. Meth. A** **362** (1995) 487.
- [243] T. Adye, ‘Unfolding algorithms and tests using RooUnfold’, *Proceedings, 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding (PHYSTAT 2011)* (CERN, Geneva, Switzerland, 17th–20th Jan. 2011) 313, arXiv: 1105.1160 [physics.data-an].
- [244] ATLAS Collaboration, *ATLAS b -jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, **Eur. Phys. J. C** **79** (2019) 970, arXiv: 1907.05120 [hep-ex].
- [245] ATLAS Collaboration, *Measurement of b -tagging efficiency of c -jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001, 2018, URL: <https://cds.cern.ch/record/2306649>.
- [246] ATLAS Collaboration, *Calibration of light-flavour b -jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2018-006, 2018, URL: <https://cds.cern.ch/record/2314418>.
- [247] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, (2011), arXiv: 1101.0538 [hep-ph].
- [248] ATLAS Collaboration, *Determination of the parton distribution functions of the proton using diverse ATLAS data from pp collisions at $\sqrt{s} = 7, 8$ and 13 TeV*, **Eur. Phys. J. C** **82** (2022) 438, arXiv: 2112.11266 [hep-ex].
- [249] J. Pumplin et al., *Uncertainties of predictions from parton distribution functions. 2. The Hessian method*, **Phys. Rev. D** **65** (2001) 014013, arXiv: hep-ph/0101032.
- [250] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, **Comput. Phys. Commun.** **10** (1975) 343.
- [251] D. J. Lange, *The EvtGen particle decay simulation package*, **Nucl. Instrum. Meth. A** **462** (2001) 152.
- [252] ZEUS Collaboration, *Measurement of dijet production in neutral current deep inelastic scattering at high Q^2 and determination of $\alpha(s)$* , **Phys. Lett. B** **507** (2001) 70, arXiv: hep-ex/0102042.
- [253] ZEUS Collaboration, *Inclusive jet cross-sections in the Breit frame in neutral current deep inelastic scattering at HERA and determination of $\alpha(s)$* , **Phys. Lett. B** **547** (2002) 164, arXiv: hep-ex/0208037.

- [254] ATLAS Collaboration, *Determination of the strong coupling constant from transverse energy–energy correlations in multijet events at $\sqrt{s} = 13$ TeV with the ATLAS detector*, [JHEP 07 \(2023\) 085](#), arXiv: [2301.09351 \[hep-ex\]](#).
- [255] H. Abramowicz et al., *Inclusive-jet cross sections in NC DIS at HERA and a comparison of the k_T , anti- k_T and SIScone jet algorithms*, [Phys. Lett. B 691 \(2010\) 127](#), arXiv: [1003.2923 \[hep-ex\]](#).
- [256] H. Abramowicz et al., *Inclusive-jet photoproduction at HERA and determination of α_s* , [Nucl. Phys. B 864 \(2012\) 1](#), arXiv: [1205.6153 \[hep-ex\]](#).
- [257] S. Chekanov et al., *Jet-radius dependence of inclusive-jet cross-sections in deep inelastic scattering at HERA*, [Phys. Lett. B 649 \(2007\) 12](#), arXiv: [hep-ex/0701039](#).

Agradecimientos

Terminar de redactar esta tesis supone la finalización de un periodo muy importante de mi vida. Importante por los más de cinco años que ha durado, pero también por todo lo que he vivido tanto en lo relativo a esta tesis como fuera de ella. Para mi, este final supone un logro muy importante que afronto con mucha alegría y emoción. También con gran alivio. Es por esto, que no puedo terminar sin agradecer a todas las personas que me han acompañado en este proceso.

En primer lugar, me gustaría agradecer esta tesis a Claudia y Juan. Podíamos decir que esta aventura comenzó en 2016, hace casi ocho años, cuando accediste a ser mi tutora del Trabajo Final de Grado y continuó durante el Máster los años siguientes. En 2019, con el inicio de esta tesis, se nos unió Juan. Estos años he aprendido, primero de fotones y finalmente con pares top-antitop, mucho con vosotros. Pero también quisiera destacar el cómo; y es que a pesar de la dificultad de los métodos, o de las ininteligibles instrucciones de las twikis de ATLAS, vosotros habéis sido capaces de hacerlos comprensibles. También me llevo de estos años trabajando con vosotros el trato. Os agradezco toda la formación que me habéis dado. Mi criterio científico no sería el mismo de no haber tenido dos referentes como vosotros.

También quiero agradecerse a mi familia, en especial a mi madre y a mi tía. Me lleváis acompañando toda la vida, y desde siempre me habéis transmitido la importancia que tiene la formación, gracias a la cual he llegado hasta aquí. Gracias por hacer todo lo posible para que pudiese alargar mi formación académica. Gracias por apoyarme en las decisiones que me llevaron a convertirme en Físico. Gracias por acompañarme en la elección de empezar este doctorado. Y gracias por ayudarme tanto durante estos años y permitirme dedicar todo el tiempo que necesitaba la elaboración de esta tesis. Os quiero.

Gracias a las compañeras de la Asociación de Estudiantes de Ciencias Manuela Malasaña, tanto a las que hicisteis el Grado en Física conmigo, como a las que estuvisteis mientras realizaba el doctorado, por todas las aventuras que hemos compartido haciendo de la UAM un lugar mejor. Bravo, Richi, Paula, Julia, Josemi, Amaya, Laura y Hugo, gracias por esas tardes después de trabajar en el local. Bravo, gracias por ser mi compañero de mus tantas veces, y perdón a las parejas derrotadas, pobre Richi. A Paula, Julia, Josemi y Amaya os deseo mucha suerte en vuestras carreras en la ciencia y espero ser, al menos, de tanto apoyo en ellas como vosotros en la mía.

Gracias a María, Nagore, Paloma, Sergio, Ángel, Gala, Ramón, Noelia, Jacobo y Adrián por haber compartido tantos años haciendo que la universidad no sea el lugar irrespirable que

algunos quieren. A Sergio por ser el ejemplo para empezar este doctorado, pero también un amigo. A Ángel por ser un buen referente y amigo. A Nagore por esas comidas en el césped de la universidad en los días más estresantes de esta tesis. A María y a Paloma por estar ahí cuando legítimamente me preguntaba si merecía la pena todo esto y por ser tan buenas amigas. Gracias María por haber sido mi puente con todo lo demás durante las estancias, por no dejarme solo. También a los más nuevos que os habeis incorporado más recientemente a esta pequeña familia como Marijose, Juanma por continuar lo que empezamos. Y a Julia que, tras su paso en Malasaña, también se nos ha unido.

Gracias a Belén, Miguel, Diego, Marina, Paloma, Goncho, Lucía, Andrea y Julia por estos años que hemos pasado juntos haciendo el barrio un lugar mejor. También por todas las cervezas y alguna que otra noche en el *Chaquetño*. A Belén y Miguel por ser los mejores vecinos que uno puede pedir y por estar ahí siempre que lo he necesitado. A Diego y Marina por tantos ratos en el *Mamá Teresa* y el *Achuri*, por estar ahí y por ser tan buenos amigos.

Gracias a Carlos, Jose, Diana, Andrés, Josemi y Oscar por tantas cenas en el *Alma Cheli* cada vez que Jose hace una visita a la capital. A Santi, Carmen, Emilia y Jose Miguel por ser una referencia. Gracias a todas vosotras por llevar un mundo nuevo en vuestros corazones.

Gracias a Daniel, Ana, Guillermo, Javi, Mercedes, Pili y Carlos por hacerme un hueco en vuestras vidas en Ginebra. Por esos momentos de relax a la hora de comer y por hacer más alegre un lugar demasiado nublado. A Daniel por toda tu ayuda durante mi cualificación como autor de ATLAS y después mientras realizaba la calibración *in situ* de los jets. Y por mantener tan bien las cosas de trabajo fuera de los momentos de ocio. A Ana por tener tiempo para echarme una mano cuando lo he necesitado.

También quiero agradecer a los miembros de la Colaboración ATLAS que han hecho posible que unos átomos almacenados en una bombona de hidrógeno en la frontera franco-suiza se conviertan en las muestras de datos usadas en los análisis de esta tesis.

Por último, me gustaría agradecer a las instituciones que han financiado económicamente este trabajo. Al proyecto RTI2018-095791-B-I00 del Ministerio de Ciencia, Innovación y Universidades que financió mi contrato durante los primeros años de doctorado de octubre de 2019 a diciembre de 2021. Al proyecto PID2021-122738OB-I00 del Ministerio de Ciencia e Innovación que ha financiado ambas estancias en el CERN y varias participaciones en congresos. Al Departamento de Física Teórica por el contrato de Ayudante que he tenido desde septiembre de 2022. Y al Servicio Público de Empleo Estatal por la prestación cobrada entre dichos contratos.

Termino estos agradecimientos esperando que hagan justicia a tantas personas que habéis aportado durante esta tesis y espero no haberme dejado a muchos fuera. En ese caso, disculpad.