UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Física

Measurement of inclusive jet cross-sections and search for resonances decaying into a photon and a jet

Tesis para optar por el título de Doctor de la Universidad de Buenos Aires en el área Ciencias Físicas

Jonathan David Bossio Sola

Director de tesis: Dr. Ricardo Piegaia
Consejero de estudios: Dr. Rodolfo Sassot
Lugar de trabajo: Departamento de Física, FCEyN, UBA

Buenos Aires, Junio 2018

Measurement of inclusive jet cross-sections and search for resonances decaying into a photon and a jet

Abstract

Inclusive jet cross-sections are measured in proton-proton collisions at a centre-of-mass energy of 13 TeV. The measurement uses a dataset with an integrated luminosity of 3.2 fb^{-1} recorded in 2015 with the ATLAS detector at the Large Hadron Collider. Jets are identified using the anti-kt algorithm with a radius parameter value of $R = 0.4$. The inclusive jet cross-sections are measured double-differentially as a function of the jet transverse momentum, covering the range from 100 GeV to 3.5 TeV, and the absolute jet rapidity up to $|y| = 3$. Next-to-leading-order and next-to-next-to-leading-order perturbative quantum chromodynamics calculations corrected for non-perturbative and electroweak effects are compared to the measured cross-sections. In addition, a search is performed for new phenomena in events having a photon with high transverse momentum and a jet collected in 36.7 fb^{-1} of proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV recorded with the ATLAS detector at the Large Hadron Collider. The invariant mass distribution of the leading photon and jet is examined to look for the resonant production of new particles or the presence of new high-mass states beyond the Standard Model. No significant deviation from the background-only hypothesis is observed and cross-section limits for generic Gaussian-shaped resonances are extracted. Excited quarks hypothesised in quark compositeness models and high-mass states predicted in quantum black hole models with extra dimensions are also examined in the analysis. The observed data exclude, at 95% confidence level, the mass range below 5.3 TeV for excited quarks and 7.1 TeV (4.4 TeV) for quantum black holes in the Arkani-Hamed-Dimopoulos-Dvali (Randall-Sundrum) model with six (one) extra dimensions.

Keywords: ATLAS, LHC, QCD, inclusive jet cross section, BSM, resonance production

Acknowledgements

In the first place, I want to thanks to my advisor Ricardo Piegaia, not only for giving me the chance to work in the Experimental High Energy Physics group of Universidad de Buenos Aires, but also for the advices and the so many conversations we had throughout all these years. The achievements and knowledgment gained are, without doubt, due to his great capability as a mentor. I also want to acknowledge Gustavo Otero y Garzón for also giving me the chance to work in this group as well as for all the help recieved. In addition, I want to thanks to the rest of the group, to the past members, Hernan and Sabri, and to those who remain, Gino, Flor, Roberta and Mariana. All of you contributed to some extend, either with work discussions or of any other kind, as well as in work projects. I would also like to thanks to all members of the ATLAS experiment with whom I shared studies and analyses as well as all those who have been very helpful in transmitting their knowledge and endured my great amount of questions. At the risk of forgetting someone, I want to thanks to Bogdan, David Lopez, Jeff, TJ, Dimitris, Fabrice, Chris Meyer, Chris Young, Steven, Steve, Pavel, Zdenek, Esteban, Tancredi, Gabriel, Max, Kate, Sasha, Santiago and Oleg. I would also like to acknowledge the great work (and willigness to help) of all, past and current, (sub-)conveners from the SM, JetEtmiss and Exotics ATLAS groups. Thank you Aime for your support and patience thoughtout all these years of long days at work, and for being the best life partner I could ever wish to have. To my family, thanks for the support and your trust in me.

Contents

Introduction

This thesis presents a measurement of the inclusive jet cross-sections and a search for new phenomena in high-mass final states with a photon and a jet in proton-proton (pp) collisions at a centre-of-mass energy of 13 TeV using data collected by the ATLAS detector. ATLAS ("A Toroidal LHC ApparatuS") is one of the two general-purpose detectors at the Large Hadron Collider (LHC) and the biggest multi-purpose particle detector ever built. It is used to investigate a wide range of physics, from Standard Model measurements, such as precision tests of quantum chromodynamics or study of the properties of the Higgs boson, to the search of new phenomena like extra dimensions and dark matter candidates. The LHC, built by the European Organization for Nuclear Research (CERN) and installed in a 27-kilometre circular tunnel, is the world's largest and most powerful particle collider. This machine is capable of colliding energetic beams of protons (or heavier nuclei) at rates upward of millions per second. The precision and high beam energy of the LHC allow to explore the tera-electronvolt scale, an energy range never before achieved in a particle collider.

Precise measurements of jet cross-sections are crucial in understanding physics at hadron colliders. They probe quantum chromodynamics (QCD), the theory of strong interactions, where jets are interpreted as resulting from the fragmentation of quarks and gluons produced in a short-distance scattering process. Inclusive jet events represent a background to many other processes at hadron colliders. The predictive power of fixed-order QCD calculations is therefore relevant in many searches for new physics. The measurement of the inclusive jet cross-sections presented here includes for the first time a comparison to the next-to-next-to-leading-order (NNLO) perturbative QCD prediction.

Prompt photons in association with jets are copiously produced at the LHC, mainly through quark-gluon scattering ($qg \to q\gamma$). The γ + jet(s) final state provides a sensitive probe for a class of phenomena beyond the Standard Model that could manifest themselves in the high end of the invariant mass spectrum of the γ+jet system. In this thesis, a search is presented that covers the mass range 1.5–5 TeV. Results are interpreted in the context of several models, providing limits on production cross-sections and masses of new phenomena objects that supersede previous ATLAS exclusion regions.

As part of the graduate student activities, a large number of infrastracture developments were carried on within the jet performance group. Among the most important ones, the derivation and study of the global sequential calibration, a key step in the jet energy hadronic calibration, is presented as an example in this thesis.

This thesis is organised as follows. Chapter 1 begins with an introduction of the Standard Model of particle physics, a motivation for new physics, and a description of Monte Carlo simulations. The LHC and the ATLAS detector are introduced in Chapter 2. The global sequential calibration is described in Chapter 3. The measurement of the inclusive jet cross-sections is presented in Chapter 4, while the search for photon+jet resonances is detailed in Chapter 5. Finally, the conclusions are given in Chapter 5.6.

Chapter 1

Theoretical introduction

1.1 The Standard Model

The Standard Model (SM) is a quantum field theory that describes the behaviour of all experimentally-observed particles under the influence of the electromagnetic, weak and strong forces.[1] The SM describes for instance the forces that hold together protons and neutrons in the atomic nuclei associated to strong interactions, the binding of electrons to nuclei in atoms or of atoms in molecules caused by electromagnetism, and the energy production in the sun and the other stars which occurs through nuclear reactions induced by weak interactions.

The description of electroweak and strong interactions is introduced in terms of symmetries, using the formalism of gauge theories, which play a fundamental role in particle physics. The SM is a renormalizable field theory based on local symmetries (i.e., separately valid at each space-time point) with a set of conserved currents and charges, hypercharge Y, weak isospin T, and colour C. The symmetry groups respectively associated to these charges are $U(1)_Y$, $SU(2)_L$, and $SU(3)_C$. The subindex L is used to denote the $SU(2)_L$ group because only left-handed fermions (see Section 1.1.2) have weak isospin. The subgroups $SU(2)_L \times U(1)_Y$ and $SU(3)_C$ represent the electroweak and the quantum chromodynamic sectors, respectively. The electric charge Q is related to weak isospin and hypercharge through the Gellmann-Nishijima formula $Q = T_3 + \frac{Y}{2}$ [1, 2], with the third component of weak isospin written as T_3. A complete discussion of the SM is broadly available in the literature [3, 4], a brief description is given in the next sections.

1.1.1 Particle content and interactions

The Standard Model describes all elementary particles[2] known up to now, and their interactions. Particles are categorised into two classes: "bosons" and "fermions". Bosons have integer spin and obey the Bose-Einstein statistics, whereas fermions have half-integer spin and follow the Fermi-Dirac statistics. Fermions and charged bosons have a corresponding antiparticle with the same mass and whose quantum numbers are opposite in sign.

The interactions are pictured in terms of exchanges of elementary particles and the

[1] In principle the gravitational force should also be included in the list of fundamental interactions but their impact is fortunately negligible at the distance and energy scales usually considered in particle physics experiments.

[2] An elementary particle is considered point-like, i.e. there is no evidence of any internal structure, when the position and spin are its only degrees of freedom.

couplings between them, which are mathematically characterised in the Lagrangian by quadratic and higher order terms, respectively. The probability of a certain process is obtained by considering all relevant Feynman diagrams[3], which are pictorial representations of the terms in the perturbative expansion (see Section 1.1.6). The Feynman diagrams connect the incoming and outgoing particles via the allowed vertices (couplings) and propagators (particle exchanges, known as force mediating particles).

The SM includes twelve elementary particles of spin $1/2$. These are classified according to how they interact (i.e. charges they carry) and they are divided in two families: "quarks" and "leptons". There are six flavours of quarks: "up" (u), "down" (d), "charm" (c), "strange" (s), "top" (t) and "bottom" (b). The up, down and strange quarks are frequently referred to as the light quarks, while the bottom and charm quarks as the heavy quarks. The heaviest quark of the SM, the quark top, was the last to be found [5, 6]. The quarks interact via the strong interaction and they carry an internal quantum number denoted as colour (C), which represents the charge associated to the strong interaction. The colour charge is of three types: "red" (r), "green" (g) and "blue" (b). Quarks also carry fractionary electric charge ($2/3$ or $-1/3$)[4] and weak isospin, hence, quarks interact with other fermions both electromagnetically and via the weak interaction. Quarks are only observed forming colour-neutral composite particles, called "hadrons". Each hadron is composed by a set of "valence" quarks, as well as a "sea" of gluons and quark-antiquark pairs. Valence quarks are responsible for the hadron quantum numbers. Both valence and sea quarks, along with the gluons, share the total momentum of the hadron. There are two types of hadrons depending on the composition of valence quarks: "baryons" (composed of three quarks) and "mesons" (constituted of a quark and an antiquark).

Fermions that do not carry colour charge are called leptons, and there are six flavours of them: "electron" (e^-), "muon" (μ^-), "tau" (τ^-) and their corresponding neutrinos, denoted by ν_e, ν_μ and ν_τ, respectively. The neutrino does not carry electric charge either, so its dynamics is driven by the weak force only. The electron, muon and tau leptons interact both electromagnetically and weakly. Fermions are classified in pairs, following the weak isospin doublets of their left-handed components. These pairs are themselves organised in three generations:

$$\begin{array}{ccc} \begin{pmatrix} \nu_e \\ e^- \end{pmatrix} & \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} & \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix} \\ \begin{pmatrix} u \\ d' \end{pmatrix} & \begin{pmatrix} c \\ s' \end{pmatrix} & \begin{pmatrix} t \\ b' \end{pmatrix} \end{array} \tag{1.1}$$

The weak eigenstates (primed quarks, ν_e, ν_μ and ν_τ from Equation 1.1) of the quark and neutrino fields are not identical to their mass eigenstates (d, s, b, ν_1, ν_2 and ν_3). Instead, the primed quarks (ν_e, ν_μ and ν_τ) are linear combinations of d, s and b (ν_1, ν_2 and ν_3) parametrised by the CKM (PMNS) matrix, proposed by Cabibbo, Kobayashi and Maskawa [7] (Pontecorvo, Maki, Nakagawa and Sakata [8, 9]). The coupling between the neutrino (quark) mass eigenstates from different generations is thus proportional to the (very small) off-diagonal elements of the PMNS (CKM) matrix. The fermionic sector of the Standard Model is summarised in Table 1.1.

Spin 1 particles, the "gauge bosons" (also called vector bosons) are defined as force carries that mediate the strong, weak and electromagnetic fundamental interactions. The

[3]Feynman diagrams are named after its inventor, Richard Feynman.

[4]The electric charge is given in units of the elementary charge, e, which is the charge carried by the positron.

Table 1.1: Fermionic sector of the Standard Model. For the neutrinos, the limits correspond to the effective masses, $\mathrm{m}^2_{\nu_f} = \sum_{i=1}^{3} |U_{fi}|\mathrm{m}^2_{\nu_i}$, where U is the PMNS matrix and f=e,μ,τ. Values are taken from the Particle Data Group (PDG) [10].

1st Generation			2nd Generation			3rd Generation		
QUARKS (spin = 1/2)								
Particle	mc^2	$Q[e]$	Particle	mc^2	$Q[e]$	Particle	mc^2	$Q[e]$
u	2.2 MeV	2/3	c	1.28 GeV	2/3	t	173.1 GeV	2/3
d	4.7 MeV	−1/3	s	96 MeV	−1/3	b	4.18 GeV	−1/3
LEPTONS (spin = 1/2)								
Particle	mc^2	$Q[e]$	Particle	mc^2	$Q[e]$	Particle	mc^2	$Q[e]$
ν_e	$<$ 2 eV	0	ν_μ	$<$ 0.19 MeV	0	ν_τ	$<$ 18.2 MeV	0
e^-	0.511 MeV	-1	μ^-	105.7 MeV	-1	τ^-	1776.9 MeV	-1

electromagnetic force between electrically charged particles is mediated by the photon, denoted by γ, which is massless and well-described by the Quantum electrodynamics theory (see Section 1.1.2). The weak interactions between quarks and leptons are mediated by the W^+, W^- and Z bosons, which are all massive.

The $W^\pm$ carries an electric charge of ±1, and couples to the electromagnetic interaction. The weak interactions involving the $W^\pm$ exclusively act on the left-handed particles and right-handed antiparticles only. The electrically neutral Z boson interacts with right- and left-handed particles and antiparticles. The $W^\pm$ bosons may also interact among themselves, with Z and γ, as the W^+ and W^- are both weakly and electrically charged. The three weak gauge bosons along with the photon are grouped together, as collectively mediating the electroweak interaction. The remaining gauge bosons are the eight gluons (g). They are massless and mediate the strong interaction between colour charged particles. Since gluons have an effective colour charge they can also interact among themselves. Quarks, gluons and their interactions are described by the theory of Quantum Chromodynamics, and it is further discussed in Section 1.1.3. The remaining boson H, the "Higgs boson", is discussed briefly in Section 1.1.5. The SM bosons are summarised in Table 1.2.

Table 1.2: Standard Model bosons. Masses are taken from the Particle Data Group (PDG) [10].

Particle	mc^2	$Q[e]$	Spin
g	0	0	1
$W^\pm$	80.385 GeV	±1	1
Z	91.1876 GeV	0	1
γ	0	0	1
H	125.09 GeV	0	0

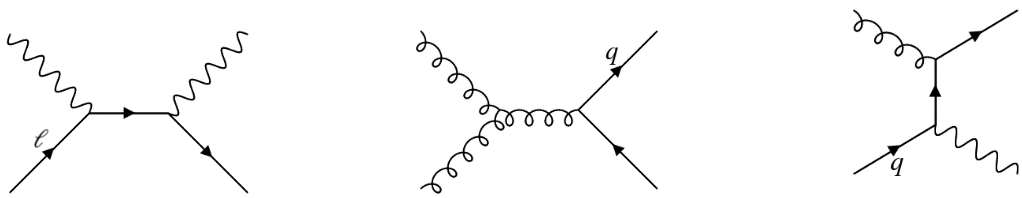
The couplings allowed by the Standard Model, and the kind of interaction they describe, are detailed in Table 1.3. It is customary to attribute the interaction responsible for a process according to the vertices present in the corresponding tree level diagrams, although there are processes that cannot be ascribed to a unique interaction. This is illustrated in Figure 1.1, which shows (a) a pure EM process (in this case mediated by a

Table 1.3: Couplings allowed by the Standard Model, and the kind of interaction they describe [10, 11].

Electromagnetic	$\gamma q\bar{q}$ $\gamma\ell\bar{\ell} \quad \ell = (e, \mu, \tau)$
Electroweak	$Wq\bar{q}' \quad W\ell\bar{\nu} \quad Zq\bar{q} \quad Z\ell\bar{\ell}$ $W^+W^-\gamma \quad W^+W^-Z$ $W^+W^-\gamma\gamma \quad W^+W^-ZZ \quad W^+W^-Z\gamma \quad W^+W^-W^+W^-$
Strong	$gq\bar{q}$ $ggg \quad gggg$
Higgs interactions	$H\ell\bar{\ell} \quad Hq\bar{q}$ $HW^+W^- \quad HZZ \quad HHZZ \quad HHW^+W^-$ $HHH \quad HHHH$

lepton) contributing to Compton scattering, (b) a pure strong process (here mediated by a gluon) contributing to the jet cross-section measurement discussed in Chapter 4 and (c) a mixed EM-strong process (in this example mediated by a quark) which corresponds to the main source of background for the analysis presented in Chapter 5.

Figure 1.1: Tree level Feynman diagrams for (a) a pure EM process (mediated by a lepton), (b) a pure strong process (mediated by a gluon) and (c) a mixed EM-strong process (mediated by a quark).



(a) (b) (c)

1.1.2 The electroweak interactions

Quantum electrodynamics (QED) is the relativistic quantum field theory based on the symmetry group $U(1)$ that describes the interactions of charged particles via the exchange of one (or more) photon. The full theory of QED was developed by Feynman, Schwinger and Tomonaga throughout the 1940s [12]. The structure of the SM is, in a sense, a generalisation of this theory, extending the gauge invariance of electrodynamics to a larger set of conserved currents and charges. The coupling of charged fermion fields ψ to the photon field A^μ is described by the QED Lagragian density, which is given by

$$\mathcal{L}_{QED} = \bar{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}.$$

The covariant derivative D_μ and the field strength tensor $F_{\mu\nu}$ are given by

$$D_\mu = \partial_\mu - ieA_\mu$$

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu,$$

such that the Lagrangian is invariant under local $U(1)$ gauge transformations. The γ^μ are the Dirac matrices, which satisfy $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$. The strength of the interaction is characterised by the coupling $\alpha = e^2/4\pi$. In addition to electromagnetic interactions, fermions are subject to weak interactions. Both are manifestations of the unified electroweak theory, which is described by the gauge symmetry $SU(2)_L \times U(1)_Y$ as proposed by Glashow, Weinberg and Salam [13–15].

The fermion fields are expressed by Dirac spinors which can be decomposed into a left- (L) and a right-handed (R) component, which respectively correspond to the -1 and +1 eigenvalues of the matrix operator $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. Consequently, the L and R components are obtained by applying the chirality projector operators

$$P_L = \frac{1-\gamma^5}{2} \text{ and } P_R = \frac{1+\gamma^5}{2}, \tag{1.2}$$

respectively. The L fermion fields $\psi_i = \binom{\nu_i}{l_i}_L$ and $\binom{u_i}{d'_i}_L$ of the i^{th} generation transform as doublets under the $SU(2)_L$ symmetry group. The conserved quantum number under $SU(2)_L$ transformations is the third component of the weak isospin, T_3, which is equal to $+1/2$ for the upper component in each doublet and $-1/2$ for its isospin partner. The R fermion fields are invariant under $SU(2)_L$. The weak interaction is the only fundamental interaction that breaks parity-symmetry.

The gauge fields corresponding to the generators of the gauge symmetry are W^i_μ with $i = 1, 2, 3$ for $SU(2)_L$, and B_μ for $U(1)_Y$. The respective coupling strengths are denoted g and g' and the field strength tensors are given by

$$W^i_{\mu\nu} = \partial_\mu W^i_\mu + g\epsilon_{ijk} W^j_\mu W^k_\nu$$

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

Analogous to $\mathcal{L}_{QED}$, the interactions between the gauge fields and fermions are described by the Lagrangian density

$$\mathcal{L}_{EW} = i\sum_f \bar{\psi}_f \gamma^\mu D_\mu \psi_f - \frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu},$$

which is invariant under local $SU(2)_L \times U(1)_Y$ gauge transformations when the covariant derivative is given by

$$D_\mu = \partial_\mu + \frac{1}{2} ig\tau^i W^i_\mu - \frac{1}{2} ig'YB_\mu.$$

The generators associated with the $SU(2)_L$ symmetry group are the Pauli matrices τ_i and the generator of the $U(1)_Y$ symmetry is the hypercharge Y.

Initially, the proposed unification failed because it predicted massless gauge fields associated to the generators of the $SU(2)_L$ symmetry group, analogous to the photon in QED, which were not observed. Instead there was indirect evidence for the massive charged $W^\pm$ and neutral Z^0 bosons, which have masses close to 80 and 90 GeV, respectively [16, 17]. A mechanism was required for the weak bosons to acquire mass. The proposed solution involves spontaneous symmetry breaking through the Higgs mechanism (see Section 1.1.5).

1.1.3 Quantum chromodynamics

The huge effort to describe the rich spectrum of mesons and baryons resonances that were discovered during the 1950s, prompted Gell-Mann and Zweig to propose in 1964 the quark model [18, 19], which asserts that hadrons are in fact composites of smaller constituents. This framework could rationalise the vast hadron spectroscopy already observed in terms of smaller particles; the quarks, in the fundamental representation of $SU(3)$. The most prominent properties of QCD are "asymptotic freedom" and "confinement", and they are introduced in Section 1.1.4

The quark model was formalised into the theory of Quantum Chromodynamics with quarks carrying an additional quantum number called colour charge. Without colour charge, it would seem that the quarks inside some hadrons exist in symmetric quantum states, in violation of the Pauli exclusion principle (this was indeed the problem of the quark model as proposed by Gell-Mann and Zweig). The colour theory extends the electroweak Lagrangian to the symmetry under $SU(3)_C$ transformations, which introduces eight new physical gauge fields, the gluons.

The Quantum Chromodynamics Lagrangian density is given by

$$\mathcal{L}_{QCD} = \sum_q \bar{\psi}_{q,a} \left(i\gamma^\mu (D_\mu)_{ab} - m_q \delta_{ab} \right) \psi_{q,b} - \frac{1}{4} G^A_{\mu\nu} G^{A\mu\nu}.$$

The $\psi_{q,a}$ are the quark fields for flavour q and colour index a, which runs from 1 to $N_c = 3$. The covariant derivative D_μ and the gluon field strength tensor $G^A_{\mu\nu}$ are given by

$$D_\mu = \partial_\mu + i g_s t^A \mathcal{A}^A_\mu,$$

$$G^A_{\mu\nu} = \partial_\mu \mathcal{A}^A_\nu - \partial_\nu \mathcal{A}^\nu_\mu + g_s f^{ABC} \mathcal{A}^B_\mu \mathcal{A}^C_\nu,$$

where $\mathcal{A}^A_\mu$ are the gluon fields with index A, B, C running from 1 to $N_c^2 - 1 = 8$. The 3×3 matrices t^A are the generators of the $SU(3)$ group and satisfy $[t^A, t^B] = i f^{ABC} t^C$. The strong coupling strength g_s is usually replaced by $\alpha_s = g_s^2/4\pi$.

1.1.4 Asymptotic freedom and confinement

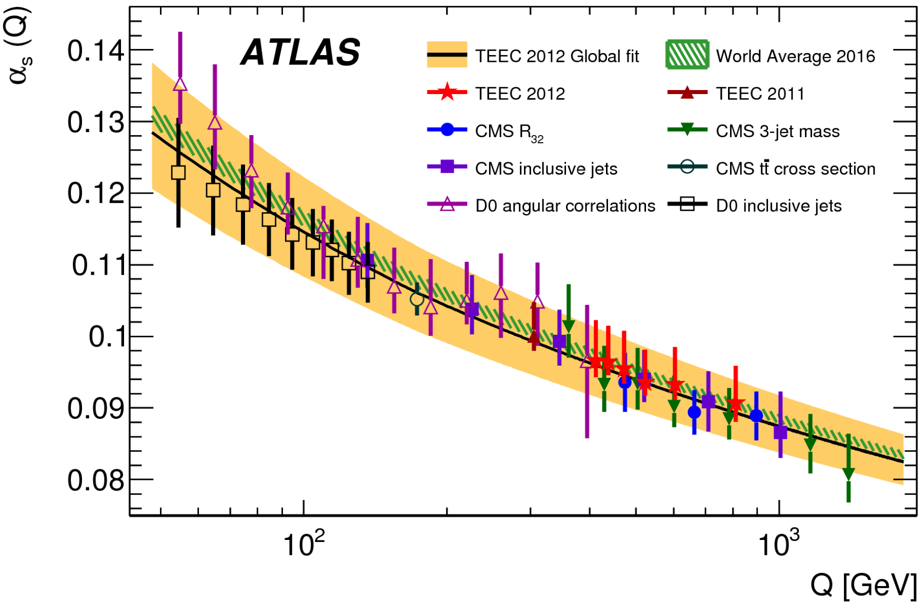
One of the approaches to solving QCD is referred to as Perturbative QCD [20] (see Section 1.1.6), which essentially relies on the idea of an order-by-order expansion of a given observable in terms of a small coupling $\alpha_s \equiv g_s^2/4\pi$, where g_s denotes the QCD coupling constant.

The coupling α_s decreases for increasing Q^2 (i.e. the scale of the process) and vanishes asymptotically, where large Q^2 corresponds to small distances. Figure 1.2 shows the running of α_s at various scales, illustrating the good consistency between the measurements.

The QCD interaction becomes weak in processes with large Q^2, therefore the quarks and gluons (collectively called "partons") behave as essentially free in such limit. This phenomenon is referred to as "asymptotic freedom", and it was discovered by Gross, Politzer and Wilczek [22, 23] in 1973. On the contrary, the interaction strength becomes large at small transferred momenta (i.e. large distances), of order $Q \lesssim \Lambda_{\text{QCD}}$.[5] The increasing force either binds the quarks together or it breaks when the energy density of the colour field between the quarks is great enough to create from the vacuum quark-antiquark pairs. As a consequence, quarks do not exist in isolation, but rather hadronise to form tightly bound composite states, with compensating colour charges so that they are overall neutral in colour (hadrons). The impossibility of separating colour charges as

[5]The scale at which the coupling diverges is referred to as Λ_{QCD}.

Figure 1.2: The QCD coupling determined in different experiments, together with the (green) band from the world average. Image from Ref. [21].



individual quarks and gluons is known as "confinement" [24], and explains why quarks and gluons are only observed, at low energies, trapped together into colour-neutral hadrons.[6]

1.1.5 Higgs mechanism

The photon and the gluons have zero masses as a consequence of the exact conservation of the corresponding symmetry generators: the electric charge and the eight colour charges, respectively. On the other hand, the weak bosons have large masses signalling that the corresponding symmetries are largely broken. The Standard Model predicts in principle massless fermions and gauge bosons, in contradiction with the observations. The idea behind the underlying mechanism of generating non-zero masses while preserving the renormalisability of the theory was initially studied by Nambu, Goldstone [25, 26] and Anderson [27], and developed into a full relativistic model in 1964 independently and almost simultaneously by three groups of physicists: Higgs [28, 29], Englert and Brout [30], and Guralnik, Hagen and Kibble [31]. Furthermore, Higgs proposed the existence of a hypothetical massive scalar elementary particle as a proof for this idea [32]. This particle has no spin, and therefore is classified as a boson.

The so-called Higgs mechanism consists of inducing a spontaneous breaking of the electroweak gauge symmetry:

$$SU(3)_C \times SU(2)_L \times U(1)_Y \rightarrow SU(3)_C \times U(1)_Q,$$

generating massive $W^{\pm}$ and Z gauge bosons, predicting also the presence in the physical spectrum of one spin-0 particle, "the Higgs boson" [3, 4]. Weinberg and Salam were the first to apply the Higgs mechanism to the breaking of the electroweak symmetry [15, 33] and showed how a Higgs mechanism could be incorporated into Glashow's electroweak theory [13], setting thus the foundations of the Standard Model of particle physics.

[6]In very-high energy enviroments, such as the universe shortly after the Big Bang, quarks and gluons are only weakly linked by the strong force, forming what is called a quark-gluon plasma.

The relationship between the masses of $W^{\pm}$ and Z gauge bosons predicted by the SM is given by

$$m_W = m_Z \cos\theta_W, \quad \text{with} \quad \tan\theta_W = g'/g, \tag{1.3}$$

where the weak missing angle θ_W relates the hypercharge and weak coupling constants, denoted by g and g', respectively. The $W^{\pm}$ and Z vector bosons were discovered in 1982 by the UA1 and UA2 collaborations [17, 34], where the prediction given by Equation 1.3 was successfully verified. Furthermore, not only the bosons acquire mass through the Higgs mechanism but also the fermions. The masses are not predicted by the SM so they are just parameters of this theory. After the discovery of the top quark in 1995 by D0 and CDF collaborations [5, 6], the masses of the fermions have been all measured experimentally.[7]

On 4th July 2012, the ATLAS [35] and the CMS [36] collaborations at the LHC simultaneously claimed the discovery of a previously unknown boson of mass between 125 and 127 GeV/c^2, whose behaviour has been found (so far) to be consistent with the Standard Model Higgs boson [37, 38]. Currently, a tremendous experimental effort is underway aimed to study the properties of the discovered Higgs-like particle.

1.1.6 Perturbative Quantum Chromodynamics

The experimental consequence of asymptotic freedom is that the hard interactions of quarks and gluons at the energy scale probed by hadron colliders can be described by perturbative QCD (pQCD), although their presence in the final state can only be inferred indirectly, as they appear confined in colourless hadrons. Each order of the perturbative expansion corresponds to an additional power in the coupling constant. This power is related to the number of vertices in the matrix element QCD Feynman diagrams, with $\sqrt{\alpha_s}$ per vertex, with the exception of the 4-gluon vertex which contributes with α_s. Each increasing order in α_s of the perturbative expansion simply corresponds to a set of diagrams with the correct combination of vertices. By drawing all possible Feynman diagrams for a given order of perturbation theory, all the terms in the calculation can be read off. In this context, leading-order diagrams are also known as "tree-level" diagrams (with no internal loops). Since the value of α_s varies with energy, it must be evaluated at the energy scale of the interaction. In particular, at the electroweak scale, $\alpha_s(M_Z) \sim 0.117$, the perturbative expansion converges relatively fast, allowing all except the lower order terms to be ignored. The complexity of the process determines the order to which the calculation can be performed.

Using this formalism, the two parton hard interaction cross-section can be computed to some fixed order in perturbation theory. However, hadron colliders such as the LHC do not produce simple parton-parton interactions, but instead collisions of hadrons that consist of multiple partons. The factorisation theorem [39, 40] allows the perturbative calculations for parton interactions to be extended to proton-proton collisions as well as to separate the short-distance component of the scattering process described by pQCD from the non-perturbative long-distance component. This theorem states that the total cross-section for two hadrons can be obtained by weighting and combining the cross-sections for two particular partons. This weighting is done using $f_i(x, Q^2)$, the parton distribution functions (PDFs), which describe the parton density for a parton of species i in a hadron, with a fraction x of the hadron energy-momentum when the hadron is

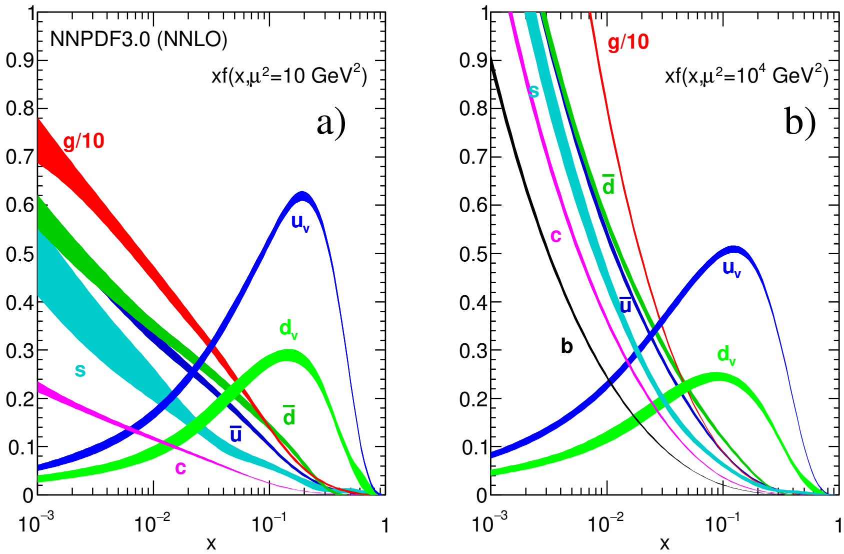
[7] For the neutrinos, only the difference between the square masses are known ($\Delta \mathrm{m}^2_{ij} = \mathrm{m}^2_i - \mathrm{m}^2_j$).

probed at a resolution scale Q^2. The cross-section for a hard scattering process $pp \to X$, initiated by two hadrons with four-momenta P_1 and P_2 can be written as:

$$\sigma_{pp\to X} = \sum_{i,j} \int_0^1 dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \hat{\sigma}_{ij}(p_1, p_2, \alpha_s(\mu_R^2), Q^2/\mu_R^2, Q^2/\mu_F^2). \tag{1.4}$$

Here, x_1 and x_2 are the momentum fractions carried by the interacting partons, and $p_1 = x_1 P_1$ and $p_2 = x_2 P_2$ are the interacting parton momenta. The partonic cross-section $\hat{\sigma}_{ij}$, corresponding to the interaction of partons i and j, is calculated at a fixed order in α_s, which is evaluated at some renormalisation scale, μ_R, and factorisation scale, μ_F. The total cross-section is obtained by summing over all possible parton flavours and integrating over all possible momentum fractions. The parton distribution functions, f_i and f_j, are evaluated at a factorisation scale, μ_F, which can be thought of as the scale that separates short-distance, perturbative physics, from long-distance, non-perturbative physics (i.e., separates hard and soft processes). Figure 1.3 shows the parton distribution function evaluated at two different factorisation scales for all possible partons.

Figure 1.3: The bands are x times the parton distribution functions $f_i(x, Q^2)$ (where $i = u_\mathrm{v} = u - \bar{u}, d_\mathrm{v} = d - \bar{d}, \bar{u}, \bar{d}, s \simeq \bar{s}, c \simeq \bar{c}, b \simeq \bar{b}, g$) obtained in an NNLO NNPDF3.0 global analysis [41] at scales $\mu_F^2 = 10\,\mathrm{GeV}^2$ (left) and $\mu_F^2 = 10^4\,\mathrm{GeV}^2$ (right), with $\alpha_s\left(M_Z{}^2\right) = 0.118$. Images from Ref. [10].



If the perturbative expansion were carried to all orders, the cross-section $\sigma_{pp\to X}$ in Equation 1.4 would be independent of μ_F and μ_R. In actual finite order calculations this is not true. They are usually both taken to be equal, $\mu_F = \mu_R = \mu$, chosen at the typical scale Q^2 of the process, in order to minimise the contribution of (uncalculated) higher order terms which appear as logarithmic terms of the form $\log(Q^2/\mu_R^2)$ and $\log(Q^2/\mu_F^2)$. The dependence of the prediction on μ_F and μ_R is assigned as a theoretical uncertainty. The fact that the cross-section of a process should be independent of the factorisation scale μ_F led to the DGLAP equations, published separately in the 1970s by Yuri Dokshitzer, Vladimir Gribov and Lev Lipatov, and Guido Altarelli and Giorgio Parisi [42]. These equations determine the evolution of the PDFs with Q^2.[8]

On the other hand, the dependence on x must be obtained by fitting possible cross-section predictions to data from hard scattering experiments. The understanding of the

[8]The DGLAP evolution equations describe the change with Q^2 of the quark densities due to gluon radiation and gluon splitting, and of the gluon density due to radiation from quarks and gluons.

PDFs plays a key role for interpreting the data at hadron colliders in terms of the SM predictions and possible deviations. Dedicated groups perform the parametrisation of PDFs using data from different experiments and processes.

1.1.7 Particle physics in hadron colliders

The variety of physics processes ocurring in a hadron-hadron collision is highlighted next in an inwards-outwards flow (i.e., from short-distance to long-distance processes).

Initially two hadrons are coming in on a collision course, where each hadron can be thought as a group of essentially collinear partons quantitatively characterised by the parton distributions (see Section 1.1.6). In a collision scenario with accelerated particles carrying electromagnetic and colour charges, bremsstrahlung can occur, e.g. as gluon radiation such as $q \to qg$. A collision between two partons, one from each side, takes place producing the hard process of interest, that can be calculated by a perturbative approach to some order in α_s, which corresponds to the number of outgoing partons. Emissions that are started off from the two incoming colliding partons are called "initial-state radiation" (ISR). After the collision, outgoing partons can also radiate. This emission is referred to as "final-state radiation" (FSR). With the parton shower development, the colour field strength increases as partons recede and they can break up by the production of quark-antiquark pairs. Thus, quarks and antiquarks may combine to produce a primary hadron. The creation of hadrons as a consequence of the confinement phenomenon is referred to as "hadronisation". Each of the incoming hadrons is made up of a multitude of additional partons, which may also interact within the same hadron-hadron collision. These semi-hard secondary collisions are referred to as "multiple parton interactions" (MPI).[9] Each of these further collisions also may be associated with its radiation. The additional products of the collision that are not explicitely related to the hard process (radiation, hadron remnants, products of multiple parton interactions, etc.), are generally grouped altogether and called "underlying event" (UE).

Overall, the complex picture of a hadron-hadron collision introduced in this section is illustrated in Figure 1.4.

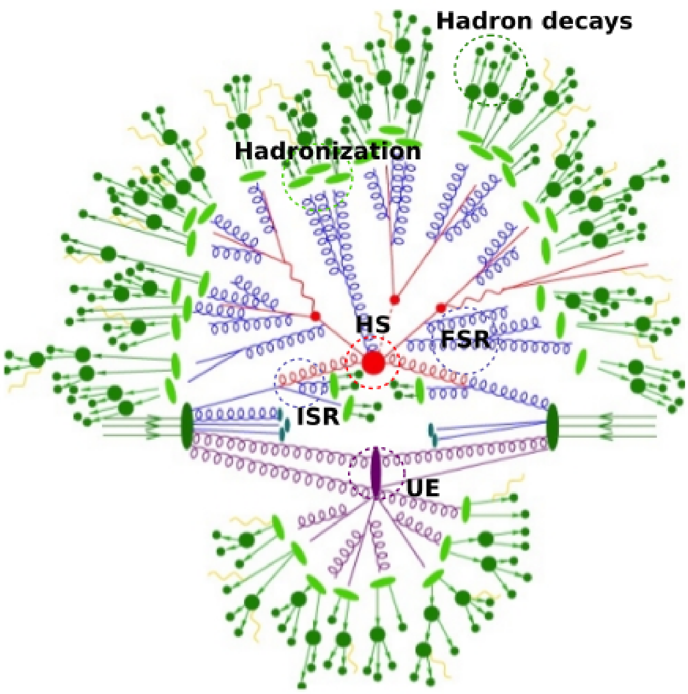
The processes that take place during a hadron-hadron collision cannot be completely calculated through pQCD. Currently, there are different Monte Carlo (MC) simulations available, involving various suitable approximations, that have been developed to address some of the phenomena mentioned above (see [44, 45] and references therein). The key features of the MC simulations required for the understanding of the analyses presented in this thesis are described in Section 1.2.

1.1.8 Beyond the Standard Model

The SM provides a strong theoretical foundation to our understanding of high energy physics and agrees remarkably well with experimental results. However, theoretical discrepancies point to the possibility of new physics beyond the standard model (BSM). An example is the hierarchy problem [46–48], described simply as a theoretical inconsistency between the relative strengths of the gravitational and electroweak forces. The differences can only be accounted for through very fine-tuned cancellations of force couplings against quantum corrections, an overly arbitrary and unlikely scenario.

[9] The multiple parton interactions are a completely separate effect from the multiple proton interactions that may occur in each bunch collision event in a hadron collider.

Figure 1.4: Illustration of the stages of a hadron-hadron collision: the hard interaction (HS), the initial- and final-state radiation (ISR and FSR, respectively), the hadronisation, the subsequent hadron decays, and the underlying event. Image from Ref. [43].



Experimental evidence for BSM physics is also piling up. Analysis of galactic rotation curves [49] and gravitational lensing [50] point to discrepancies between the mass of galaxies as calculated through gravitational predictions and through extrapolations of the aggregate stellar luminosity. This discrepancy leads to the existence of a hypothetical type of matter distinct from ordinary matter known as dark matter. Many cosmological observations, including the measurement of the cosmic microwave background [51], indicates a universal composition of 5% matter, 27% dark matter, and 68% dark energy, another unknown form of energy.

Proton-proton collisions at a new energy frontier of 13 TeV have the potential to create new, heavy particles that are not predicted by the SM. If a new particle couples to the SM it may be created by parton collisions at the LHC and can therefore subsequently decay into SM particles, leaving evidence of its existence. This potential evidence of BSM physics may be found in precision measurements of cross-sections as deviations from the SM, or in searches of non-standard-model resonance signals decaying into SM particles.

1.2 Monte Carlo simulations

Having QCD predictions is crucial in precision measurements as well as in searches for new physics. The "fixed-order predictions" include the first terms in the QCD perturbative expansion for a given cross-section. As more terms are involved in the expansion, an improvement in the accuracy of the prediction is expected. The complexity of the pQCD calculations increases significantly with the number of outgoing partons. Furthermore, pQCD calculations do not provide the full hadronic level event description. An alternative approach is provided by the so-called Monte Carlo programs.

1.2.1 Hard interaction and parton shower

In order to describe a $2 \to n$ process from the Lagrangian of the theory (where n represents a given number of partons in the final state), a set of Feynman diagrams has to be lay down and evaluated following precise rules, so that matrix elements (ME) can be calculated in powers of the strong coupling constant α_s. The factorial increase of Feynman diagrams with the number of outgoing partons makes it hard to have calculations at high order. Nevertheless, a complex $2 \to n$ process can be factorised into a simple core process, e.g. $2 \to 2$, convoluted with parton splitting probabilities to mimic a higher order process. Simulation programs implementing this approach are for instance PYTHIA [52] and HERWIG++ [53, 54]. These use leading-order (LO) perturbative calculations of matrix elements of $2 \to 2$ processes and implement higher-order QCD processes approximately via the so-called initial- and final-state "parton showers" (PS) [45, 55] to produce the equivalent of multi-parton final states.

In a hard process with virtuality Q^2 (i.e. hardness), the incoming (outgoing) partons are expected to radiate a succession of harder (softer) gluons while approaching (retreating from) the collision. The emission ratio for a branching such as $q \to qg$ diverges when the gluon either becomes collinear with the quark or when the gluon energy vanishes. Furthermore, the non-Abelian character of QCD[10] leads to $g \to gg$ branchings with similar divergences. The third main branching $g \to q\bar{q}$ does not have the soft divergence. For next-to-leading-order (NLO) perturbative QCD MC programs, like SHERPA [43] and POWHEG [56, 57], the parton shower needs to be matched to the ME calculation[11,12] to avoid double counting of radiation generally modelled by a sequence of emissions that, starting from the scale where confinement becomes important, increase the virtuality in each emission until it matches the Q^2 of the hard process. Similarly, the final-state radiation is constituted by a sequence of emissions that decrease the virtuality of the partons until a lower cut-off is reached ($Q_0^2 \equiv \Lambda_{\mathrm{QCD}} \lesssim 1\,\mathrm{GeV}$). Thus, the whole phase space is expected to be covered with a smooth transition from ME to PS. Below Q_0^2, no further branchings are simulated. Perturbation theory ceases to be meaningful, and confinement effects and hadronisation phenomena take over. Different MC programs control the coherence of succesive emissions by ordering them in terms of their transverse momenta ($p_\perp$) or angle with respect to the parton direction. In particular, PYTHIA and HERWIG++ use shower models that are $p_\perp^2$- and angular-ordered, respectively.

1.2.2 Underlying event

In addition to the hard interaction that is generated by the Monte Carlo simulation, it is also necessary to account for the interactions between the incoming proton remnants (MPIs). This is usually modelled through multiple extra $2 \to 2$ scattering, occurring at a scale of a few GeV. The modelling of the underlying event is crucial in order to give an accurate reproduction of the energy flow that accompanies hard scatterings in hadron

[10] The transformations of the symmetry group do not commute in the case of the Quantum Chromodynamics and weak groups.

[11] POWHEG is interfaced with other programs, PYTHIA or HERWIG++, to provide the parton shower and the hadronisation model (see Section 1.2.3).

[12] The interfacing of matrix element calculations to PS is generally done using a matching procedure (MLM [58] or CKKW [59, 60]), that essentially relies on a slicing of the phase space where some region is constrained to be generated by the parton shower, whereas the rest is covered by the ME calculation (i.e., where the PS approximation has limitations).

colliders. The UE can include additional hard interactions and soft processes which can not be calculated perturbatively. These are modelled with adjustable parameters which are tuned to experimental data.

1.2.3 Hadronisation

As the evolution reaches $Q_0^2 \equiv \Lambda_{\mathrm{QCD}}$, the parton shower phase is truncated since the coupling forces become significant and confinement takes place. This phenomenon cannot still be described from first principles, and therefore, it involves some modelling to transform all the outgoing coloured partons into colourless hadrons of a typical 1 GeV mass scale. The dynamics of this evolution is generally absorbed in fragmentation functions that represents the probability of a parton to fragment into a certain hadron of the final state. Many of these primary hadrons are unstable and decay further at various timescales. Those that are sufficiently long-lived have their decays visible in the detector, or they are stable. There are several models of the hadronisation process, that attempt to connect the results of the parton shower and the final particle spectrum observed. These models can be complemented and tuned using experimental observations. The hadronisation is commonly described by either the Lund string fragmentation model [61] (as implemented in Pythia), or the cluster fragmentation model [62] (as implemented in Herwig++ and Sherpa[13]). Essentially, the Lund string fragmentation model asummes a linear confinement, where the energy stored in the colour field between quarks and antiquarks is assumed to increase linearly with the separation of colour charges. Thus, it depicts the colour force by means of a linearly rising potential as charges separate. The potential energy stored increases as partons recede, so it may break up by the production of new quark-antiquark pairs that screen the endpoint colours. Then, quarks and antiquarks may combine to produce hadrons. The cluster fragmentation model is based on the colour preconfinement property of the branching processes, which assumes that the separation of the colour charges forming a singlet are inhibited. After the perturbative parton branching process, the remaining gluons are splitted into light $q\bar{q}$ pairs, and then neighbouring quarks and antiquarks can be combined into colour singlets (colourless "clusters"), with masses distributions peaking at low values and asymptotically independent of the hard subprocess scale.

1.2.4 Tunes

Due to the non-perturbative, and therefore incalculable, nature of much of the soft physics processes, like the shower approximations, hadronisation and underlying event, MC generators inevitably contain a number of free parameters. These different parameters are usually tuned with data from colliders. A specific set of chosen parameters for a MC generator is referred to as a "tune".

Two ATLAS Pythia tunes are used for event simulations in this thesis: the A14 tune [63] for the hard interaction and the AU2 tune [64] for pile-up collisions. The A14 tune is based on the MONASH tune [65] of the Pythia authors which uses e^+e^- collision data for the hadronisation parameters, and minimum bias pp collision data at LHC to constrain parameters sensitive to initial state radiation and the underlying event. The A14 tune uses in addition a large variety of ATLAS data sensitive to multiple parton interactions and ISR/FSR, and includes jets built from tracks and variables sensitive to

[13] Sherpa can use its own hadronisation model or be interfaced to Pythia.

the internal jet structure. The AU2 tune targets the PYTHIA parameters sensitive to the modelling of minimum bias pp collisions and is based on tracks and calorimeter clusters information.

1.2.5 Detector simulation

In order to use events[14] produced by Monte Carlo generators to model particle collisions that one might observe with the detector, the output of these generators is passed through a detector simulation [66]. ATLAS uses the GEANT 4 [67] toolkit. GEANT 4 is an extensive particle simulation package that governs all aspects of the propagation of particles through detectors, based on a description of the geometry of the detector components and the magnetic field. The physics processes include, among others, ionisation, Bremsstrahlung, photon conversions, multiple scattering, scintillation, absorption and transition radiation. The last step involves the digitalisation, which simulates the detector outputs in the same format as the actual raw data.

The detector is described in terms of almost 30 million volumes with properties, which in case of the ATLAS detector (described in Section 2.2) are constructed based on two databases: the geometry database and the conditions database. The former contains all basic constants, e.g. dimensions, positions and material properties of each volume. The latter is updated according to the circumstances at a given time and contains for instance dead channels, temperatures and misalignments. As a result, several layouts of the detector are available. Test beam data taken with components of the ATLAS detector before completion have aided the validation and further improvement of the detector simulation. Due to the detailed and complicated geometry of ATLAS and the diversity and complexity of the physics processes involved, the consumed computing time per event is large ($\mathcal{O}(1\,\text{hour})$).

[14] An event refers to the results just after a particle collision, occurring in a very short time span, at a well-localised region of space. Here, the meaning of an event is different from the one used in the theory of relativity, in which an "event" is a spacetime coordinate.

Chapter 2

The ATLAS experiment

2.1 The Large Hadron Collider

The Large Hadron Collider [68] is the world's largest, most advanced, and powerful particle collider ever built. The LHC is a proton–proton synchrotron built by the European Organization for Nuclear Research, installed in the 27-kilometre circular tunnel previously occupied by the Large Electron Positron (LEP) collider at approximately 100 metres underground below the border between France and Switzerland. It is a precision instrument capable of colliding energetic beams of protons (or heavier nuclei) at rates upward of millions per second in interaction "points" (IPs) of the size of microns. The precision and high beam energy make of the LHC an extraordinary tool to explore the tera-electronvolt (TeV) scale, an energy range never before achieved in controlled collider experiments for elementary particle physics. ATLAS [35] is one of two general-purpose detectors at the LHC and is described in Section 2.2. It investigates a wide range of physics, from the study of the Higgs boson to the search of extra dimensions and dark matter candidates.

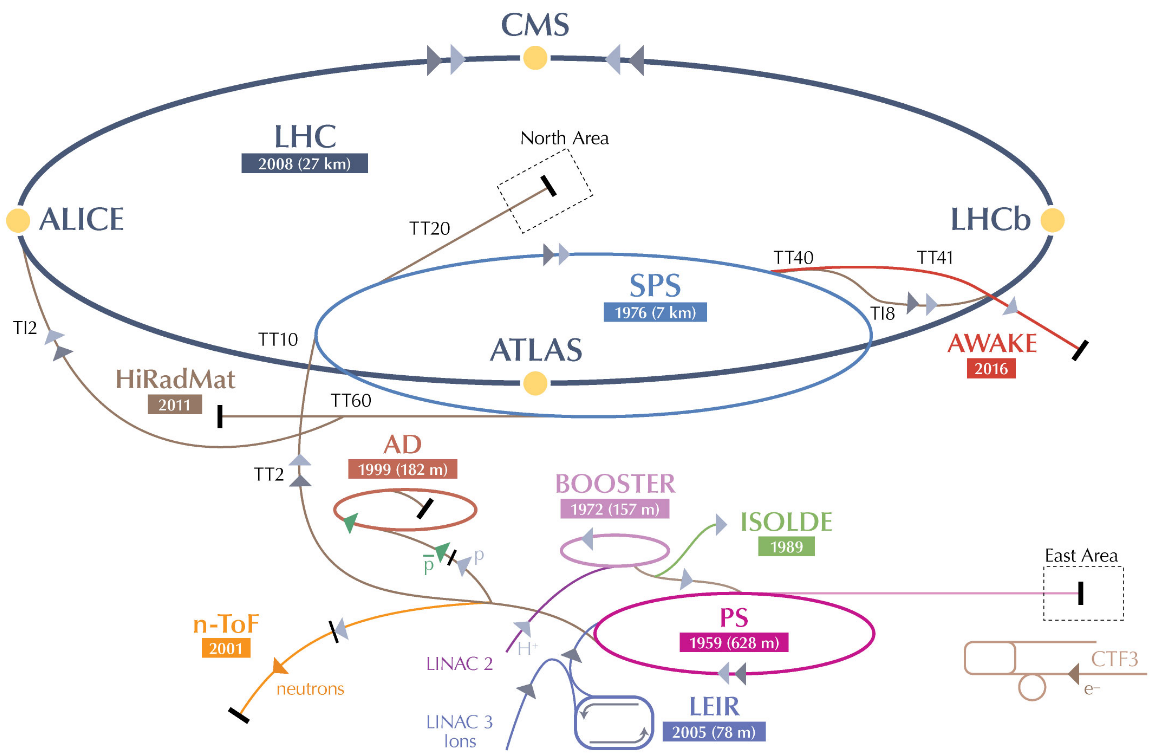
2.1.1 Machine design

The LHC had been originally designed to produce collisions between beams of protons at a centre-of-mass energy ($\sqrt{s}$) of 14 TeV. An unexpected technical incident occurred on September 2008[1] [69] delaying the LHC's physics program and forcing to operate below the design specifications at $\sqrt{s} = 7$ TeV during the first years (2010-2011) and at $\sqrt{s} = 8$ TeV in the 2012 period. The period of data-taking from 2010 to 2012 is commonly known as "Run 1". During 2013 and 2014, the LHC entered a period named "Long Shutdown", the first one of its life (LS1). Although there were no collisions, the whole CERN site was a hive of activity, with large-scale work under way to modernise the infrastructure and prepare the LHC for operation at higher energy. In 2015, the LHC started its second run of data-taking, known as "Run 2", operating at $\sqrt{s} = 13$ TeV. The decision to begin the LHC's second run at 13 TeV was taken in order to optimise the delivery of particle collisions for physics research, and thereby speed the route to potential new physics. Technical details motivating this decision can be found in Ref. [70].

The LHC ring is the last shackle in a series of intermediate accelerators, some of which have been operating for decades, that successively increase the energy of the protons. The whole chain is illustrated in Figure 2.1. The protons are extracted from ionised Hydrogen

[1] A short-circuit in a connection between superconductors in the tunnel burned a hole in a vessel containing liquid helium, with catastrophic consequences for the LHC machine.

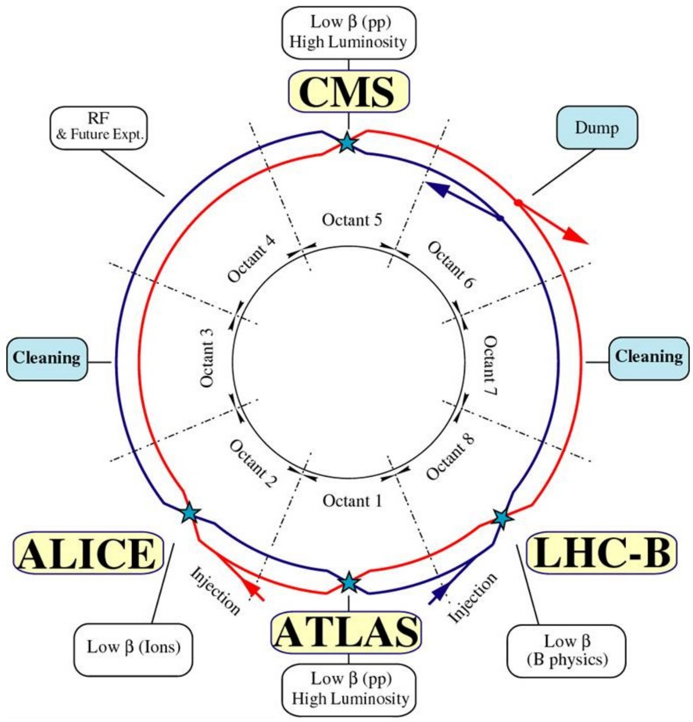
Figure 2.1: The CERN complex encompassing several accelarators (the length of each is included between parenthesis). Image from Ref. [71].



atoms and are feed to the linear particle accelerator Linac 2 where proton energy is raised to 50 MeV. These are then inserted in the Proton Synchrotron Booster, where protons are further boosted to 1.4 GeV and injected then into the Proton Synchrotron (PS) which increases their energy up to 25 GeV. The last accelerator on the chain before injecting the protons to the LHC ring, is the Super Proton Synchrotron (SPS) which further increases proton energy to 450 GeV. Up to this stage, all protons flow in the same direction, but the injection to the LHC ring occurs in two points resulting in two contra-rotating beams, as schematically shown in Figure 2.2.

The LHC contains two adjacent, high-vacuum beam pipes (maintained to an impresive pressure of 10^{-13} atm) running in parallel throughout all the ring, except for some discontinuities in the vicinity of the interaction points where the beams cross. The bending of the proton beam into a nearly circular trajectory is primarily driven by 1232 dipole superconducting magnets designed to operate at 11.08 kA, generating a magnetic field of 8.3 T necessary for 6.5 TeV proton beams. These magnets are cooled down to 1.9 K (−271.3 °C, colder than the 2.7 K of outer space) by liquid-helium vessels for proper operation at the superconducting state. A large series of smaller quadrupole, sextupole, octupole, and decapole magnets provide beam shaping to counteract Coulomb repulsion of the proton aggregate by squeezing the beam in its transverse plane. Beam acceleration within the LHC occurs in eight radiofrequency (RF) resonant cavities per beam where the electromagnetic field oscillates at a frequency that is a multiple of the protons revolution frequency (f_{rev}). Charged particles passing through the cavity are accelerated or decelerated by the electromagnetic force until becoming synchronised to the oscillating field

Figure 2.2: LHC layout with four interaction points at the site of the detectors, the beam injection points and the dump site. Image from Ref. [72].



in the cavity. The net effect is that protons tend to be grouped in tight bunches of tens of centimetres long each. Each bunch contains approximately 1.15×10^{11} protons, that ensure high instantaneous luminosity at the collision points and hence maximise the number of collisions. Not all stationary regions of the RF potential, known as “buckets”, are occupied, only every 10^{th} bucket is treated as a possible bunch slot. There are 3564 possible bunch slots.[2] In Run 2 (Run 1) the bunches were separated by a time $\tau_{\text{bunch}} = 25\,\text{ns}$ ($\tau_{\text{bunch}} = 50\,\text{ns}$). As a consequence of the SPS and LHC injection strategy, the bunches are organised in “bunch-trains”. The number of bunches per bunch-train and the separation (empty bunches) between bunch-trains is different for each data-taking period and tend to vary during data-taking. In addition to organise protons longitudinally along the beam, once a stable time structure has been achieved the frequency of the cavities is slowly increased to transfer power to push the protons forward along the accelerator. This is the process called synchrotron acceleration. Each cavity delivers 2 MV (an accelerating field of 5 MV/m) at 400 MHz, and is kept at 4.5 K. Plateau energy is reached in around 15 minutes. Both beams traveling in opposite directions are squeezed and focused at four intersection points where the detectors are placed: the multipurpose ATLAS and CMS [36] experiments, and two other dedicated detectors ALICE [73] and LHCb [74] for heavy ion and B-physics experiments, respectively (see Figure 2.2).

[2] The exact number of possible bunch slots is obtained when considering the exact LHC perimeter, 26.659 km, and 24.95 ns as bunch time separation, usually approximated to 25 ns in non-technical litterature.

2.1.2 Machine operation during 2015 and 2016

The rate of production of a certain process i is directly related to the cross-section $\sigma_{pp} \to i$ and the LHC capability to put the protons into close contact, as encoded by the instantaneous luminosity L_0. Formally the rate is expressed by

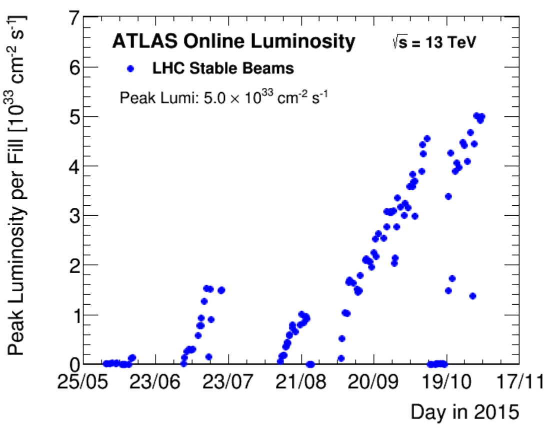
$$\mathcal{R}_i = L_0 \sigma_{pp \to i}, \tag{2.1}$$

where the cross-section is a purely physical magnitude (calculable from scattering amplitudes), whereas the luminosity depends exclusively on the beams optics. It can be expressed by

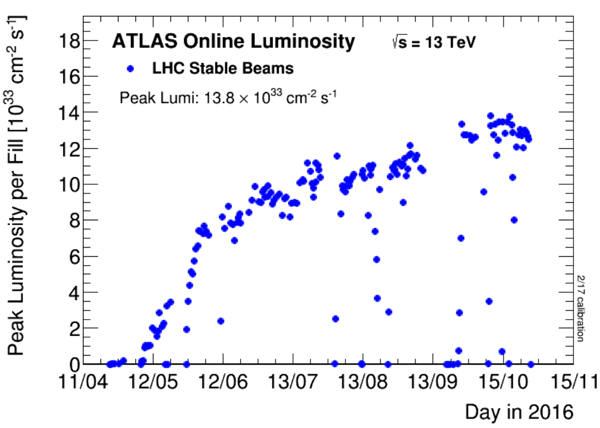
$$L_0 = n_c f_{rev} \frac{N_1 N_2}{4\pi \sigma_x^* \sigma_y^*},$$

where n_c is the number of colliding bunch pairs, N_b is the number of protons per bunch in the beam b, and $\sigma_{x,y}^*$ are the transverse sizes of the beam at the interaction point. Clearly the luminosity -and by inheritance the production rate- increases as bunches have more protons, more bunches crossings occur (either because there are more filled bunches in the pipe or the crossings are more frequent) or if the beams are more collimated at the IP reducing the transverse spread of protons. As a consequence, the observation of unlikely scattering processes with small cross-sections requires to maximise the delivered instantaneous luminosity, which is in turn limited by the machine performance and protection thresholds. The timeline of peak luminosity as measured online using the luminosity detectors (see Section 2.2.5) during 2015 and 2016 data taking periods is shown in Figure 2.3. The time integrated luminosity $L_0^{\text{int}} = \int L_0 \, \mathrm{d}t$ is shown in Figure 2.4.

Figure 2.3: The peak instantaneous luminosity delivered to ATLAS during stable beams for pp collisions at 13 TeV centre-of-mass energy is shown for each LHC fill as a function of time for 2015 and 2016 data taking periods ((a) and (b) respectively). The luminosity is determined using counting rates measured by the luminosity detectors, and is based on a preliminary 13 TeV calibration determined using the van-der-Meer beam-separation method [75], where the two beams are scanned against each other in the horizontal and vertical planes to measure their overlap function. Images from Ref. [76].



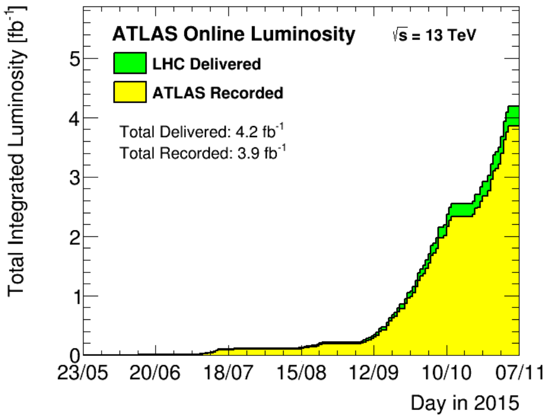
(a)



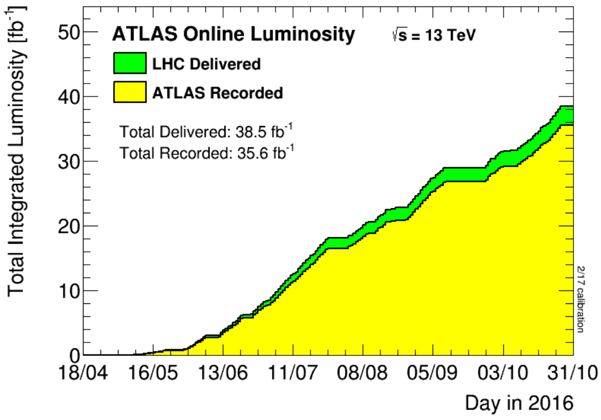
(b)

A pp collision is in essence a stochastic process. As such the high luminosity not only enhances the probability of a single pp scattering but also increases the likelihood for multiple pp collisions per bunch crossing. The number of inelastic collisions per bunch crossing is a random variable that follows a Poisson distribution with parameter μ estimated from

Figure 2.4: Cumulative luminosity versus time delivered to (green) and recorded by ATLAS (yellow) during stable beams for pp collisions at 13 TeV centre-of-mass energy in 2015 and 2016 ((a) and (b) respectively). This accounts for the luminosity delivered from the start of stable beams until the LHC requests ATLAS to put the detector in a safe standby mode to allow for a beam dump or beam studies. Images from Ref. [76].



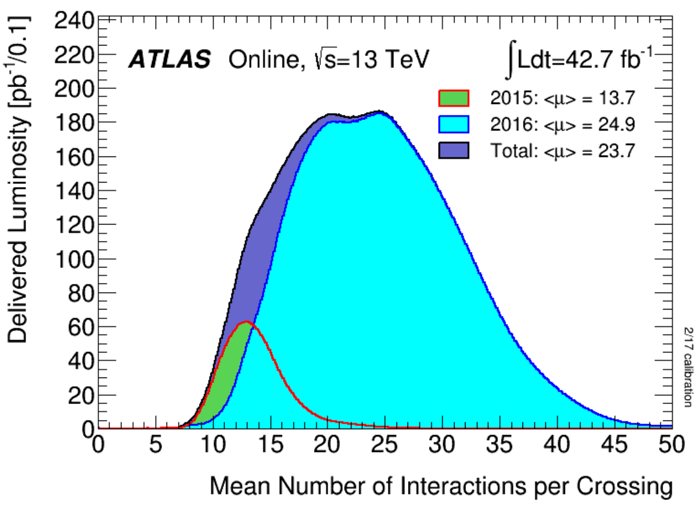
(a)



(b)

the luminosity and the inelastic cross-section. Beam optics and bunch parameters at the IP can change across different periods of data taking resulting in clear changes of the luminosity of the machine. Therefore, the underlying Poisson distribution can be shifted towards higher or lower values of $\langle\mu\rangle$ as shown in Figure 2.5 for 2015 and 2016 data.

Figure 2.5: Luminosity-weighted distribution of the mean number of interactions per crossing for the 2015 and 2016 pp collision data at 13 TeV centre-of-mass energy. All data delivered to ATLAS during stable beams is shown. The integrated luminosity and the mean μ value are given. The mean number of interactions per crossing corresponds to the mean of the Poisson distribution of the number of interactions per crossing calculated for each bunch. It is calculated from the instantaneous per bunch luminosity as $\mu = L_{\text{bunch}} \times \sigma_{\text{inel}}/f_{rev}$ where L_{bunch} is the per bunch instantaneous luminosity, σ_{inel} is the inelastic cross-section which we take to be 80 mb for 13 TeV collisions, and f_{rev} is the LHC revolution frequency. Image from Ref. [76].



In each bunch crossing there are, along with the hard-scatter process, inelastic interactions which contribute with soft additional particles in the detector that blur the resolution of the hard process of interest. This phenomenon is called pile-up and can occur in two forms. One contribution, known as “in-time” pile-up, arises from particle emissions produced by the additional proton-proton collisions occurring in the same bunch crossing as the hard-scatter interaction. The second form, named “out-of-time” pile-up, is a consequence of the short time separation between subsequent bunches ($\tau_{\text{bunch}} = 25\,\text{ns}$ in Run 2). Due to the long time taken to produce and read out the electronic signal in some sub-detectors (see Section 2.2), previous crossings affect the signal shape of the interaction of interest.

The 2015 *pp* program started effectively on 3 June when the LHC shift crew declared “stable beams” as two 6.5 TeV proton beams were brought into collision at the LHC’s four interaction points. The *pp* data-taking continued until 3 November. The 2016 data-taking period occurred from 22 April to 26 October. The instantaneous luminosity peaked in 2015 at $5 \times 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$ (see Figure 2.3a) and in 2016 at $13.8 \times 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$ (see Figure 2.3b), both close to the design value $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$. The average number of *pp* collisions per bunch crossing was 13.7 in 2015 and 24.9 in 2016, peaking at 50 simultaneous hard-scatter interactions in 2016 (see Figure 2.5).

The average efficiency of the 2015 and 2016 ATLAS data taking periods was of 92.1% and 92.4% respectively (in rough terms, it corresponds to the yellow area divided by the green area in Figure 2.4). The loss of efficiency can be due to the turn-on of the high voltage of some sub-detectors at the beginning of an LHC fill, to deadtime or to individual problems with a given sub-detector that prevented the ATLAS data taking to proceed.

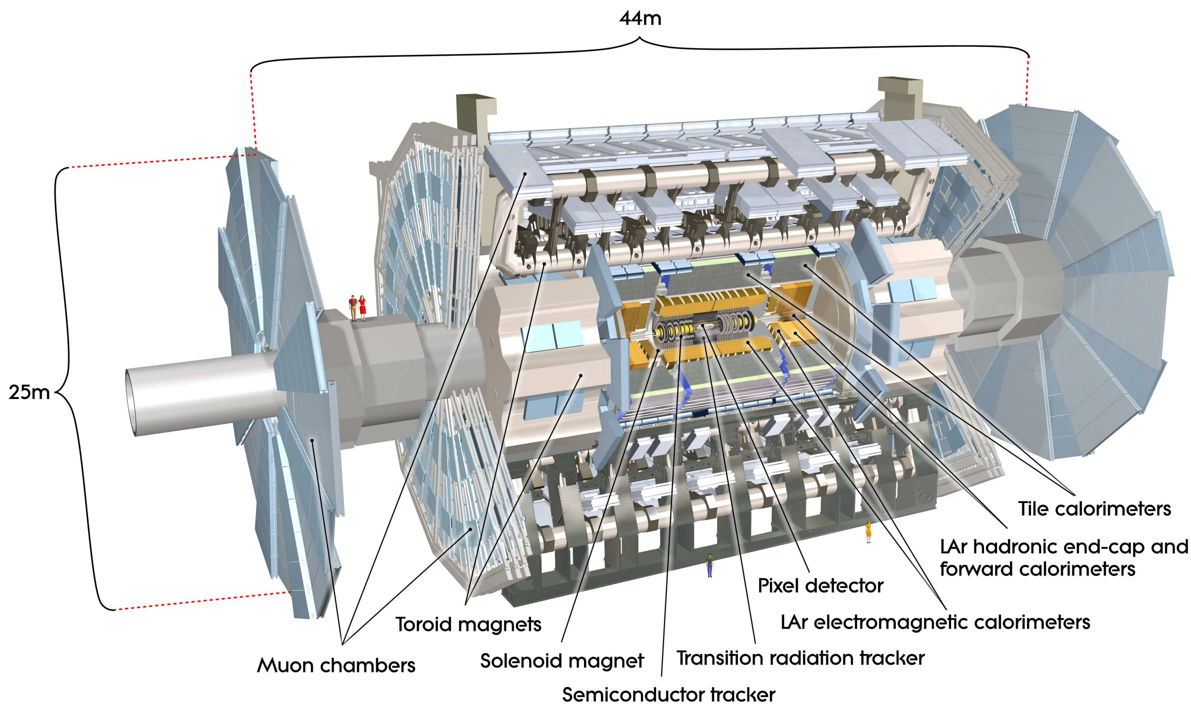
The measurement of the inclusive jet cross-sections presented in this thesis is based on the data collected in the 2015 data taking period. The search for new phenomena in high-mass final states with a photon and a jet is based on the 2015 and 2016 samples. Only data collected during stable beam periods in which all sub-detectors were fully operational are used. A brief description of the ATLAS detector and its sub-systems is given in the next section.

2.2 The ATLAS detector

ATLAS is one of the multi-purpose detectors at the LHC. It spans over 44 m in length and 25 m in height, is located at Interaction Point 1 (see Figure 2.2) and is composed of many sub-systems as illustrated in Figure 2.6. The ATLAS layout is nominally forward-backward symmetric with respect to the interaction point, and consists of three primary detection systems layered radially upon each other: a central inner tracker, a calorimeter system, and a muon spectrometer. The inner tracker (see Section 2.2.2) is embedded in a 2 T solenoidal magnetic field (see Section 2.2.6) and is used to measure the position and momentum of charged particles. The calorimeter system determines the energy of both neutral and electrically charged particles (see Section 2.2.3). The muon spectrometer is located within a large toroidal magnetic field and allows to measure the position and momentum of muons (see Section 2.2.4).

The principle behind this organisation of sub-systems can be described by simple concepts. Vertex identification (discussed in Section 2.3.1.2) requires high resolution track reconstruction; to minimise large extrapolations the tracking system is inserted as close as possible to the interaction point. The energy of charged and neutral particles is measured in the second layer, the calorimeter system. This has a large volume and is made of materials that guarantee a large number of radiation and interaction lengths. Ideally, particles are completely absorbed within the calorimeter(s) volume and their energy is thus measured as the sum of the energy deposited by the showered products of the original particles. This is the opposite case of the tracking system which is designed to minimise modifying the original trajectory of the particles. Muons usually transverse the calorimeter loosing of the order of 2 GeV, and then their momentum can be measured with improved precision using the outermost muon tracking system (in general in combination with the inner tracker). A common factor to all sub-systems forming ATLAS is that they have to resolve physics signals and correctly identify different fundamental particles emerging from the collision spot at enormous rates, and be radiation resistant to survive the harsh enviroment [77]. Each sub-system is uniquely designed to detect the energy signature of passing leptons, photons, and hadrons, as visualised in Figure 2.7.

Figure 2.6: Cut-away view of the ATLAS detector and its sub-systems. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tonnes. Image from Ref. [35].

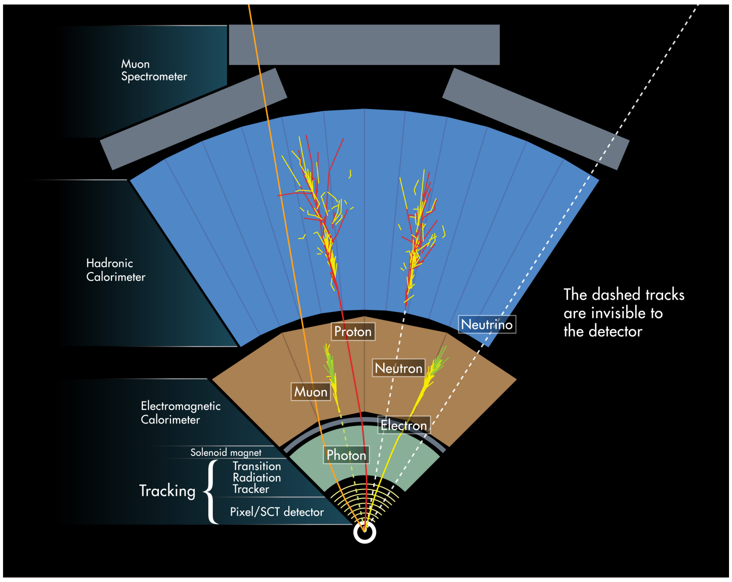


Another difficulty, consequence of the huge amount of collisions per second, is the computing power and network speeds needed to transfer the information collected at the detector, convert the measured electric pulses into physical entities (reconstructed electrons, photons, jets, etc.) and store the information permanently in disks. The following sections are intended to provide a brief overview of the ATLAS detector and of its trigger and acquisition system. For further details the reader is kindly deferred to the original source of all this information in Ref. [35].

2.2.1 Coordinate system and kinematic variables

The coordinate system used in ATLAS is a right-handed Cartesian system centred at the geometric centre of the detector. The z-axis lies along the beam line, while the

Figure 2.7: A cross-section of the ATLAS detector showing sub-detectors and their interaction with basic particles. Solid lines signify energy deposited by a single charged particle, groups of lines represent particle showers, and dotted lines illustrate no interaction with the sub-detector. Image from Ref. [78].



transverse x-y plane is defined with x pointing toward the centre of the LHC ring and y points upwards. The azimuthal angle ϕ is measured around the beam axis, and the polar angle θ is obtained from the beam axis. Note that while $\Delta\phi$ is invariant under longitudinal Lorentz boosts, the polar angle is not a convenient angular variable because the separation between two objects, $\Delta\theta$, depends on the boost of the event. In its stead, two variables are introduced: the rapidity y and pseudorapidity η, defined in equations 2.2 and 2.3, respectively.

$$y = \frac{1}{2}\ln\left(\frac{E + p_{\mathrm{z}}}{E - p_{\mathrm{z}}}\right) \tag{2.2}$$

$$\eta = -\ln\tan\left(\theta/2\right) \tag{2.3}$$

The pseudorapidity is equal to the rapidity for massless particles. Thus, for relativistic objects y is closely related to θ, with the added advantage that Δy is boost invariant. This is very important for instance for jet reconstruction and track association. These relay on the distance between two points (either calorimeter cells, or reconstructed objects) in the rapidity-azimuthal angle space defined as

$$\Delta R(1,2) = \sqrt{(y_1 - y_2)^2 + (\phi_1 - \phi_2)^2}, \tag{2.4}$$

which is boost invariant.

The conservation of momentum in the plane transverse to the beam axis implies that the transverse momentum of the collision products should sum to zero. Any imbalance is known as "missing transverse momentum", or E_T^{miss},[3] and may be indicative of weakly-interacting, stable particles in the final state. Within the Standard Model, this arises from

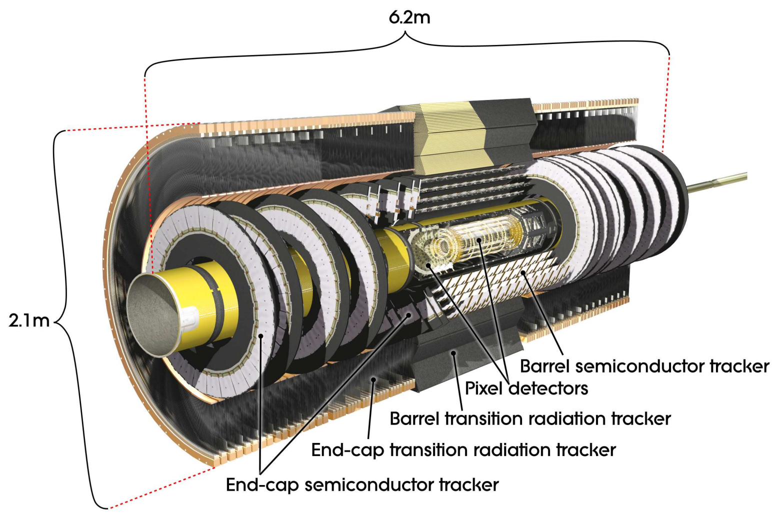
[3]The quantity E_T^{miss} is a transverse vector and is fully determined by its magnitude and its azimuthal angle. The subscript T is used to indicate the component of a vector in the transverse plane: $A_T = \sqrt{A_x^2 + A_y^2}$.

neutrinos. There are also prospects for such particles in theories beyond the Standard Model, making E_T^{miss} an important variable in searches for exotic signatures. Fake E_T^{miss} can also result from interacting Standard Model particles which escape the acceptance of the detector, are badly reconstructed, fail to be reconstructed altogether, or simply because of the finite resolution of the calorimeter energy measurements. E_T^{miss} thus can also serve as an important measure of the overall event reconstruction performance.

2.2.2 Tracking system

The inner detector (ID), or central tracking system, is responsible of reconstructing the path of charged particles via high-resolution position measurements. Tracks are used, among other things, for vertex reconstruction and identification (see Section 2.3.1.2). Combined with a solenoidal magnetic field of 2 T aligned with the beam line, tracks are used to measure the direction and momentum for particles having nominally $p_{\mathrm{T}} \gtrsim 500\,\mathrm{MeV}$, where p_{T} is the transverse momentum, within the acceptance pseudorapidity range $|\eta| < 2.5$. The measurements are conducted by the combined use of three different technologies: the Pixel detector, the SemiConductor Tracker (SCT) and the Transition Radiation Tracker (TRT). Figures 2.8 and 2.9 show diagrams of the inner detector components, which are described next.

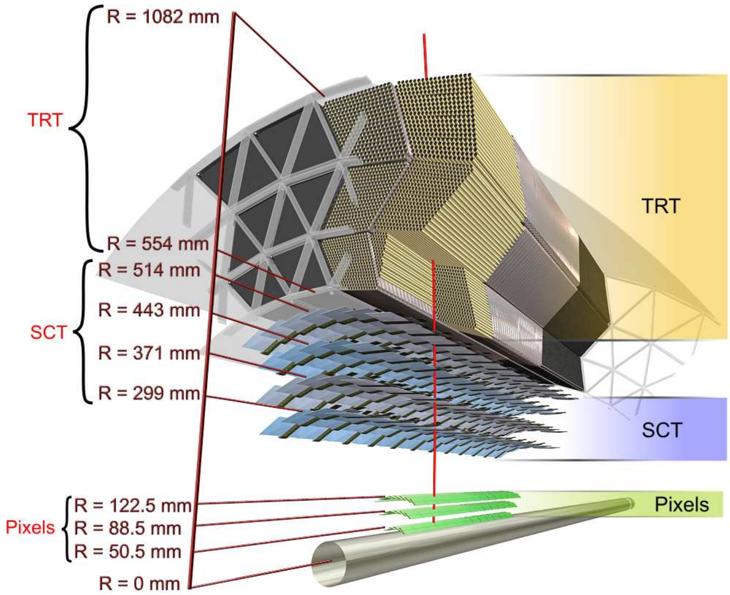
Figure 2.8: Cut-away view of the ATLAS inner detector. Image from Ref. [35].



2.2.2.1 Silicon pixel detector

The silicon pixel tracker is the innermost part of the tracking system. It is composed of three (four) concentric barrel layers in Run 1 (2). Three additional disks at each end-cap are settled at distances of 495, 580, and 650 mm from the IP along the beam direction. A new pixel layer (insertable b-layer, or IBL) was inserted closer to the collision point during LS1 [79]. This was a significant technical challenge because the beam pipe had to be removed and replaced with a smaller radius pipe upon which the IBL was mounted and inserted with all of its services into the small space between the beam pipe and the original pixel detector. Each one of the original three barrel layers span an area of 50 μm

Figure 2.9: Drawing showing the sensors and structural elements traversed by a charged track with $p_{\mathrm{T}} = 10\,\mathrm{GeV}$ in the barrel inner detector. The track traverses successively the beryllium beam pipe, the three cylindrical silicon-pixel layers with individual sensors elements of $50 \times 400\,\mu\mathrm{m}^2$, the four cylindrical double layers (one axial and one with a stereo angle of 40 mrad) of silicon-microstrip sensors of pitch $80\,\mu$m, and approximately 36 axial straws of 4 mm diameter contained in the transition-radiation tracker modules. Image from Ref. [35].



by $400\,\mu$m. In order to cope with a higher radiation dose, the IBL sensors are smaller ($50 \times 250\,\mu\mathrm{m}^2$). This results in a total of approximately 92 million pixels.

2.2.2.2 Silicon strip tracker

The SemiConductor Tracker (SCT) implements long, narrow silicon strips making the coverage of a larger area more economical in terms of the number of channels. The physical principle is the same as for the Pixel detector; ionising particles traversing the semiconductor produce electron-hole pairs whose currents are read out by the front-end chips, providing a binary response: a “hit” is registered only if the pulse height in a channel exceeds a preset threshold. There are 4088 SCT modules, each composed of two single-sided 64 mm silicon strip sensors attached to one another with a 40-mrad angular pitch to provide two-dimensional hit localisation. The SCT operates in the same enviroment conditions as the pixel detector at approximately $-10\,^\circ$C to reduce the leakage current. The SCT is divided into four concentric barrel layers ($|\eta| < 1.1$) and nine end-cap disks ($1.1 < |\eta| < 2.5$) at each side. The disks are placed at a distance between 934 and 2720 mm from the interaction point along the beam direction. This system provides a total of approximately 6.3 million channels.

2.2.2.3 Transition radiation tracker

The Transition Radiation Tracker surrounds the silicon detectors providing measurements of charged particles up to $|\eta| = 2.0$. It consists of proportional drift gaseous tubes (straws) 4 mm in diameter arranged longitudinally in up to 73 layers in the barrel ($|\eta| < 1.0$) and 160 radial straw planes in each end-cap ($1.0 < |\eta| < 2.0$). The barrel straws span

1440 mm in length and are further arranged into three cylindrical groups divided in 32 ϕ sectors. There are 52544 axially oriented tubes in the barrel. End-cap TRTs have 122880 straws, each of 370 mm long. In total, there are approximately 350000 readout channels. Drift tubes are filled with a Xenon-CO_2-O_2 gas mixture that is ionised upon the passage of charged particles. Ions within the gas drift to the straw wall (acting as a cathode) and towards a tungsten (gold-coated) wire running inside the tube (that acts as an anode) due to a large electric potential of -1530 V applied between the wall and the wire. Under normal operating conditions, the maximum electron collection time is ~ 48 ns. In addition, the barrel drift tubes are embedded in a matrix of polypropylene fibers that aid in electron identification via transition X-rays produced when charged particles cross inhomogeneous materials. In the end caps, the transition radiation is generated by foil interleaved between the straws.

2.2.3 Calorimeter system

In the LHC (or any other hadronic collider) the value of $\sqrt{s}$ is the initial energy at the hadronic level but the energy of the actual partonic process is unknown beforehand. The energy of a given event is determined *ex post facto* by the calorimeter system. The ATLAS calorimeter system is a non-compensating sampling calorimeter composed of different sub-detectors spanning full ϕ-symmetry and pseudorapidity range ($|\eta| < 4.9$, $0.5° \lesssim \theta \lesssim 179.5°$). Broad η coverage is mandatory for an adequate event reconstruction, specially for accurate determination of the missing energy in physics processes where weakly interacting particles are expected. The term "sampling" denotes the design approach that utilises intercalated layers of absorbing and active materal which induce particle showers and measure their energy, respectively [80]. The readout signal is proportional to the energy deposited in each layer of the active media. "Non-compensation" means that part of the energy of nuclear collisions between the high-energy particles from the event and the detector material nuclei remains "invisible" to the active read out as is lost in the form of nuclear recoil or fission.[4] A consequence of non-compensation is that the energy response for hadronic particles (interacting strongly with the nuclei) is smaller than for particles of the same energy that interact predominantly by electromagnetic (EM) forces. The particle showers produced from EM particles and those produced from hadronic particles involve different physics processes and therefore develop differently in the detector [10]; thus, two types of calorimeters (EM and hadronic) are designed to cater to each shower type. They are segmented into towers in η and ϕ which point towards the centre of the detector.

Electrons, positrons, and photons produce EM showers: energetic electrons radiate Bremsstrahlung photons, while energetic photons convert to electron-positron pairs when traversing dense material. To characterise the thickness of a material as seen by the EM radiation, the radiation lenght (X_0) measures both the average distance required to lower high energy electrons to $1/e$ times its initial energy by Bremsstrahlung, and also 7/9 the mean free path for high energy photons to undergo pair production. For hadronic calorimeters, the nuclear interaction length (λ) measures the average distance traveled by a hadron before interacting with a nucleus.

The longitudinal depth of an EM shower is expressed in terms of the radiation length

[4] A compensating calorimeter like the D0 calorimeter at the Tevatron, had a fissionable material as absorber such that the energy lost due to nuclear interaction is "compensated" by the energy released by fissioned atoms.

and the critical energy (E_c[5]), as shown in Equation 2.5.

$$X = X_0 \frac{\ln(E_0/E_c)}{\ln 2} \tag{2.5}$$

Hadronic showers involve processes including hadron production from nuclei interactions, hadron decays and de-exciting nuclei. Primarily, charged particles are produced by the hadronisation after the parton shower and hadron decay. Showers initiated by hadrons produce primarily pions and kaons. About one third of the energy dissipated in the hadronic shower is via neutral pions which decay -within less than 30 nm- via $\pi^0 \to \gamma\gamma$ and produce an EM component of the shower characterised by electron Bremsstrahlung and e^+e^- pair production. Additionally, particles may interact with the nuclei in the absorber in a variety of nuclear processes, which then produce additional particles near the MeV scale. The longitudinal depth of the hadronic shower scales with the nuclear interaction length, where $\lambda \propto A^{\frac{1}{3}}$.

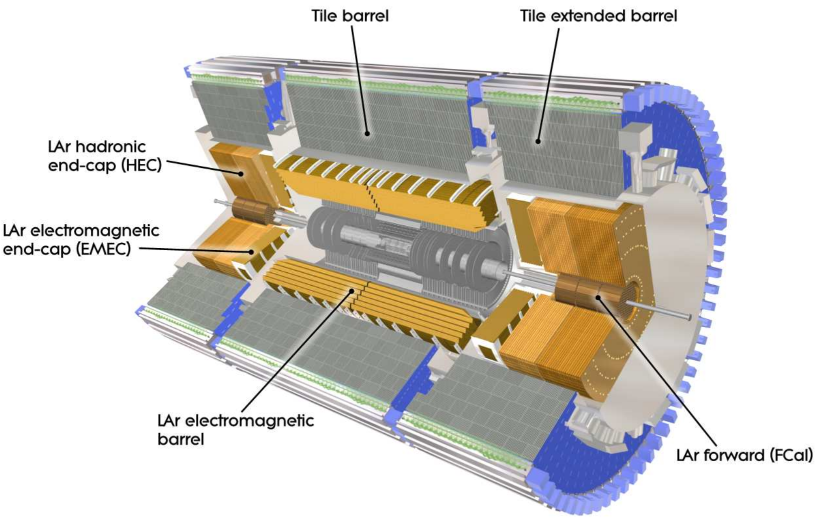
In ATLAS, two calorimeter technologies are utilised. The innermost calorimeters use sampling technology with liquid argon (LAr) as the active detector medium. These are housed in three cryostats, one barrel and two end-caps, held at 80 K using liquid nitrogen. The barrel cryostat contains the Pb/LAr electromagnetic barrel calorimeter (EMB), whereas the two end-cap cryostats each contain the Pb/LAr electromagnetic end-cap calorimeter (EMEC), the Cu/LAr hadronic end-cap calorimeter (HEC), and the copper-tungsten/LAr forward calorimeter (FCal), as illustrated in Figure 2.10. The outermost sampling detector uses steel as absorber and plastic scintillator-tiles as the active component, and is cooled with water. The hadronic tile calorimeters are divided in three parts: the central barrel and two extended barrels, shown in Figure 2.10. This second construction approach provides high opacity to hadrons (9.7 nuclear interaction lengths at $\eta = 0$) at a lower cost as compared to liquid argon calorimeters. LAr EM calorimeters cover $|\eta| < 3.2$, LAr hadronic calorimeters cover $1.5 < |\eta| < 4.9$, and Tile hadronic calorimeters cover $|\eta| < 1.7$.

The active material reacts to passing charge particles and consists of either plastic scintillators or liquid argon. In plastic scintillators the charged particles produce light via excitation followed by scintillation. The number of photons produced is proportional to the energy deposited by the incident particle and is guided to photomultiplier tubes (PMTs) by optical fibers. In the liquid argon medium the incident charged particles ionise the liquid, causing electrons and positive ions to drift (driven by a bias voltage) towards electrodes which measure the deposited charge.

Passive material consists of heavy absorber material which interacts with both charged and neutral particles but which does not measure the deposited energy. Incident particles may interact with the material through several mechanisms. Charged particles can radiate photons through Bremsstrahlung, which may pair-produce electrons and positrons that may themselves radiate further photons. Hadron-nucleon scattering can produce strongly interacting particles which can further interact with both the active and passive materials, again causing a shower of particles. Through this process the energy of charged and neutral particles may be measured and contained within the calorimeter. Muons deposit a relatively small amount of energy in the calorimeters as they are not strongly interacting and are more massive than electrons. Neutrinos do not interact with the calorimeters

[5]The critical energy ($E_c \propto \frac{1}{Z}$) measures the energy at which losses from ionisation (logarithmic with energy) are comparable to losses from Bremsstrahlung (linear with energy), above which Bremsstrahlung dominates. The radiation length $X_0 \propto Z(Z+1)$ [10].

Figure 2.10: Cut-away view of the ATLAS calorimeter system. Image from Ref. [35].



at all, appearing only as a momentum imbalance inside the detector as discussed in Section 2.2.1.

Minimising the resolution of the energy response is essential to providing an accurate measurement of particle energy. The fractional calorimeter resolution as a function of energy is described as:

$$\frac{\sigma_E}{E} = \frac{N}{E} \oplus \frac{S}{\sqrt{E}} \oplus C$$

where N is a measurement of the noise, dominant at low energies, arising from background and electronics, S parametrises the stochastic uncertainty due to the random sampling of the active and passive materials, and C is a constant term, dominant at higher energies arising from non-uniformities in the detector. These terms are added in quadrature ($\oplus$) to derive the fractional resolution.

Construction details for each calorimeter system are described next.

2.2.3.1 Electromagnetic calorimetry

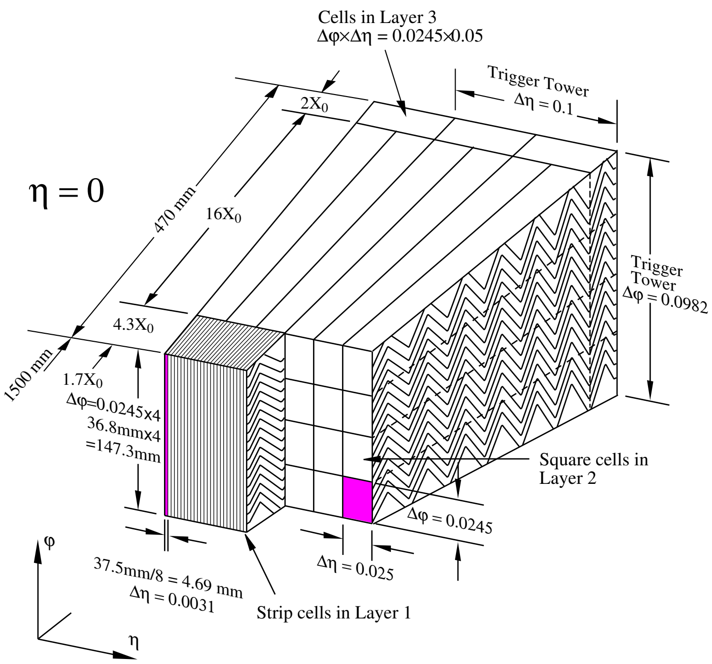
The lead (laminated with stainless steel)/LAr electromagnetic calorimeter system (EMB and EMEC modules) is the closest one to the interaction point, requiring a high radiation tolerance and fine granularity in η and ϕ to measure the narrow showers of photons and electrons. The ATLAS EM calorimeters use liquid argon as the active material due to its high radiation tolerance and uniformity. Argon is held in a liquid state at 89 K through the use of cryostats.

The whole volume of the EMB covers the region $|\eta| < 1.475$, and consists of two identical half-barrels each of 3.2 m in length and an inner (outer) diameter of 2.8 m (4 m) and formed by 1024 absorbers. The EMB is housed inside the same cryostat that holds the solenoid magnet. The primary layer of the EMB has a fine lateral and longitudinal segmentation ($\Delta\eta \times \Delta\phi = 0.003 \times 0.1$) that permits a discrimination between single or two close by photon showers produced respectively by prompt photons[6] and neutral pions

[6] "Prompt" photons are photons produced in the collision before hadronisation.

decays, for showers within $|\eta| < 2.5$, as seen in Figure 2.11. A second layer collects the largest fraction of the energy of the electromagnetic shower (around 80%), and is finely segmented in square towers of $\Delta\phi \times \Delta\eta = 0.025 \times 0.0245$ in the transverse plane ($\sim 4 \times 4\,\text{cm}^2$ for $\eta = 0$) which permits a detailed reconstruction of the showers. The third layer has half the granularity of the previous layer and measures the tail of the electromagnetic shower. A presampler (not shown in figures) covers the region $|\eta| < 1.8$ to improve the energy measurement for particles that start showering before entering the calorimeter. In the two electromagnetic end-cap components (EMEC) the accordion waves are parallel to the radial direction and run axially covering a region between $1.375 < |\eta| < 3.2$.

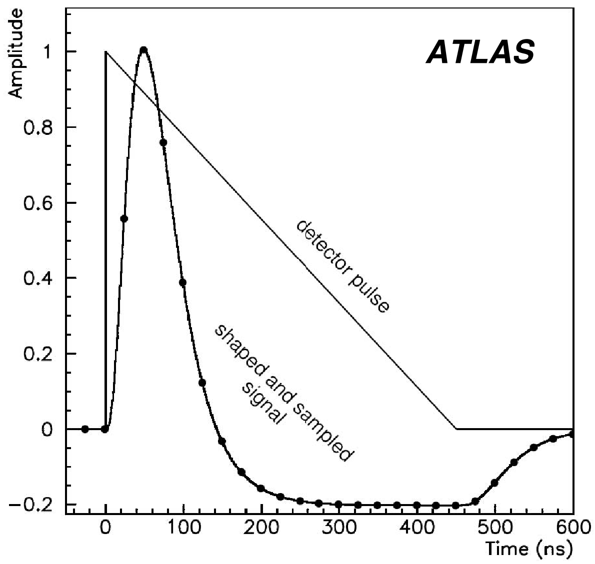
Figure 2.11: Sketch of a barrel module where the different layers are clearly visible with the ganging of electrodes in ϕ. The granularity in η and ϕ of the cells of each of the three layers and of the trigger towers is also shown. Image from Ref. [35].



The readout electrodes are located in the gaps between the absorbers and consist of three conductive copper layers separated by insulating polyimide sheets. The two outer cooper layers are at the high-voltage potential and the inner one is used to read out the signal via capacitive coupling. Both the EMB and EMEC are segmented radially into three readout layers to allow for measurement of longitudinal shower profiles. Within each layer are lead plates organised into an accordion structure that allows for uniform sampling across the full ϕ coverage and eliminating the need for servicing gaps. The average drift time of electrons is roughly 450 ns, leading to long analog signal pulses shown by the solid line in Figure 2.12. The signal is shaped to have a long negative tail used to restore the signal baseline and to reduce the sensitivity to pile-up [81]. The pulses are sampled digitally every 25 ns with four samples used to extract the signal pulse shape for amplitude and timing measurements. Five samples had been used to extract the pulse shape during the 2012 data-taking period but offered no substantial improvement over four samples. The four signal samples are converted to an energy measurement using constants calculated through test-beam data and dedicated calibration runs [35, 82].

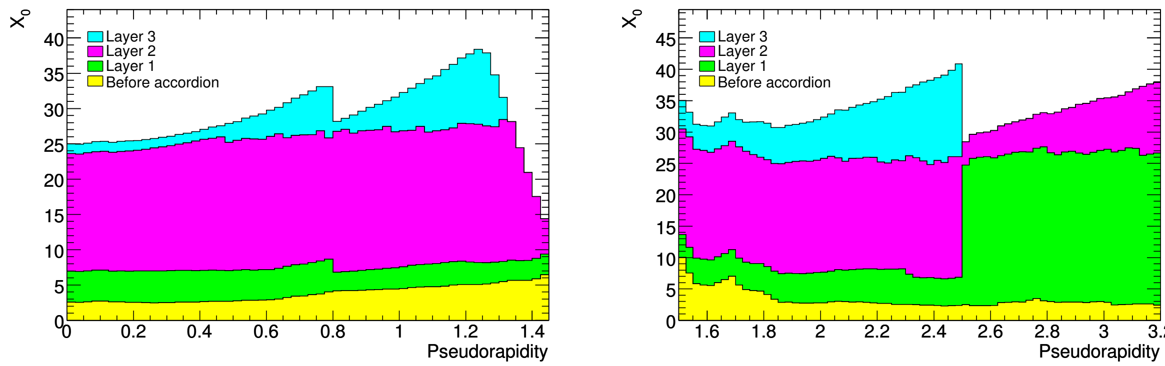
The effective thickness for electrons and photons in each layer of the EM calorimeters

Figure 2.12: The triangular signal shape of LAr signal amplitude versus time (solid line) compared to the shaped bi-polar pulse of the readout electronics (circles). The dots indicate an ideal position of samples separated by 25 ns. Four samplings are used to measure the pulse amplitude and time in contrast to five in Run 1. Image from Ref. [82].



is shown in Figure 2.13. It is measured in terms of radiation lenghts (X_0) and varies as a function of η. The total radiation length ensures that the energy of electrons and photons is mostly contained within the EM calorimeters.

Figure 2.13: Cumulative amounts of material, in units of radiation length X_0. The thicknesses of each accordion layer as well as the amount of material in front of the accordion are shown, separately, for (a) the barrel and (b) the end-cap, as a function of $|\eta|$. Images from Ref. [35].



(a) (b)

Geometrical groups of cells called "trigger towers" defined by the analog sum of the readouts from channels within $\Delta\phi \times \Delta\eta = 0.1 \times 0.1$ (see Figure 2.11) are practical as inputs for rapid trigger evaluation (but without resolution along the direction in which the shower develops), as described in Section 2.2.7.

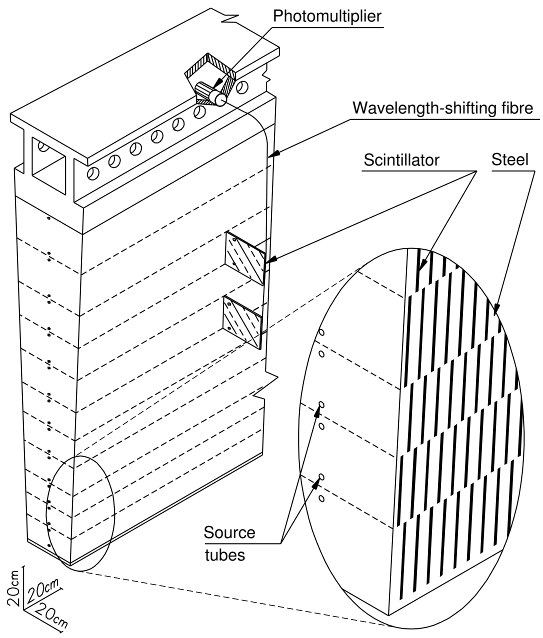
2.2.3.2 Hadronic calorimeters

This section describes the tile calorimeter, the LAr hadronic end-cap calorimeter, and the LAr forward calorimeter. The first two calorimeters cover $|\eta| < 3.2$, while the latter one covers the region $3.1 < |\eta| < 4.9$ where no dedicated EM calorimeter exists. Hadron showers will deposit a fraction of their energy in the EM calorimeter but usually extend past its 2 nuclear interaction lengths (λ). The hadronic calorimeters extend this by up to $10\,\lambda$, fully containing most high energy hadron showers.

2.2.3.2.1 Tile calorimeter

The Tile calorimeter (TileCal) consists of plastic polystyrene scintillator tiles, serving as active material, with steel absorbers. It is segmented in 64 wedge-shaped modules, each spanning $\Delta\phi \sim 0.1$, for full 2π azimuthal coverage, and containing the scintillator, steel, and readout electronics. A scheme of one of these modules is shown in Figure 2.14. The

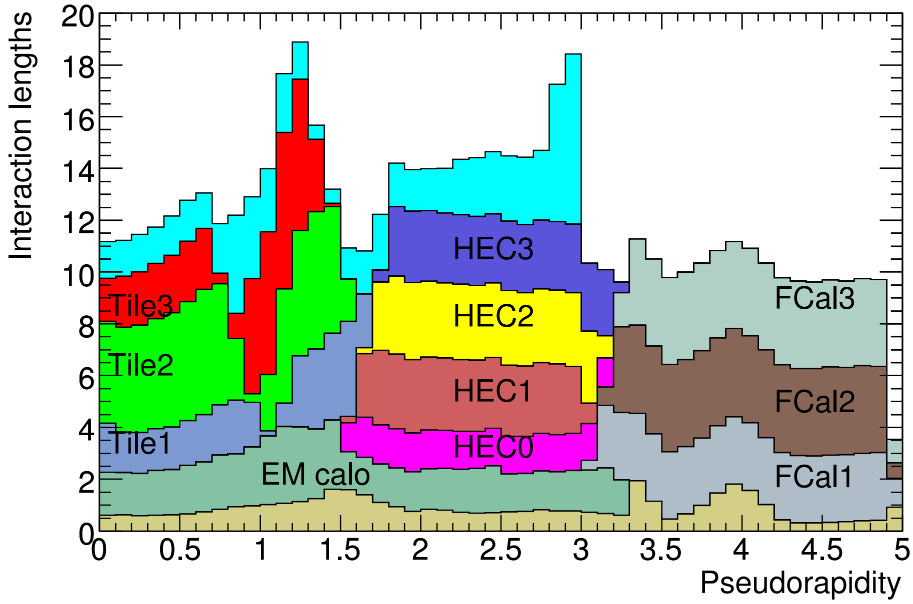
Figure 2.14: Schematic of the mechanical assembly of a single module of the Tile calorimeter showing the tile scintillators, the wavelength-shifting fibers and the photomultipliers for optical readout. Image from Ref. [35].



electronics are housed in steel support structures furthest from the beam line to reduce radiation exposure. Alternating steel and polystyrene are segmented radially into three readout layers, allowing for the measurement of longitudinal shower profiles. The layers thickness, in terms of λ, for the different components of the calorimeter system are shown in Figure 2.15.

Wavelength-shifting 1 mm-diameter fibers connect the polystyrene plates to photomultiplier tubes. The PMTs convert the ultraviolet scintillation light produced in the active material due to passing charged particles into an electrical signal. The scintillators are grouped into cells of 0.1×0.1 in $\eta - \phi$ in the first two layers and 0.2×0.1 in $\eta - \phi$ in the third layer.

Figure 2.15: Cumulative amount of material, in units of interaction length λ, as a function of $|\eta|$, in front of the electromagnetic calorimeters, in the electromagnetic calorimeters themselves, in each hadronic layer, and the total amount at the end of the active calorimetry. Also shown for completeness is the total amount of material in front of the first active layer of the muon spectrometer (up to $|\eta| = 3.0$). Image from Ref. [35].



Each cell is read out on both sides by two PMTs, requiring a total of 9852 PMTs to service the whole detector, to provide redundancy and sufficient information to partially equalise signals produced by particles entering the scintillating tiles at different impact positions. As it occurs with the EMB, the tile calorimeter is also capable to provide with analog sums of subsets of its different electronic channels, to form trigger towers used at the Level 1 trigger (see Section 2.2.7).

Readout electronics located in steel girders are used to shape, amplify, and sum PMT signals. Narrow signals are shaped into pulses with a full width half maximum of 50 ns. The shaped pulses are amplified at two complementary gain scales (1:64) to fully cover cell energies from roughly 220 MeV to 1.5 TeV and are subsequently digitally sampled at 40 MHz. Seven samples of the shaped pulse are fit with optimal filter coefficients [83] to extract the signal magnitude and timing. The steel girders also house low-voltage power supplies and infrastructure used to distribute high voltage to the PMTs.

Several calibration procedures are used to correct the measured energy to the EM scale [84–87]. Calibrations are used to derive initial conversion factors and to account for changes in the scintillator response over time or in the gain of individual PMTs over time, caused by light and radiation exposure as well as temperature fluctuations.

2.2.3.2.2 Hadronic end-cap calorimeter

Hadronic end-cap calorimeters extend the range of the hadronic calorimeter at $1.5 < |\eta| < 3.2$ (see Figure 2.10). Each end-cap is composed of two wheels with copper plate absorbers oriented perpendicular to the beam line. The absorbers provide a depth of ~ 10 nuclear interaction lenghts (see Figure 2.15). The design is influenced by their proximity to the beam where the high energy detritus from the hard scatter can be harmful for the detectors. Hence, LAr is used as the active material with readout electrodes gathering ionisation charge. The liquid argon is recycled so the effects of radiation damage are attenuated over time. The HEC has a granularity of $0.1 \times 0.1\,(\eta - \phi)$ in the first wheel and $0.2 \times 0.2\,(\eta - \phi)$ in the second.

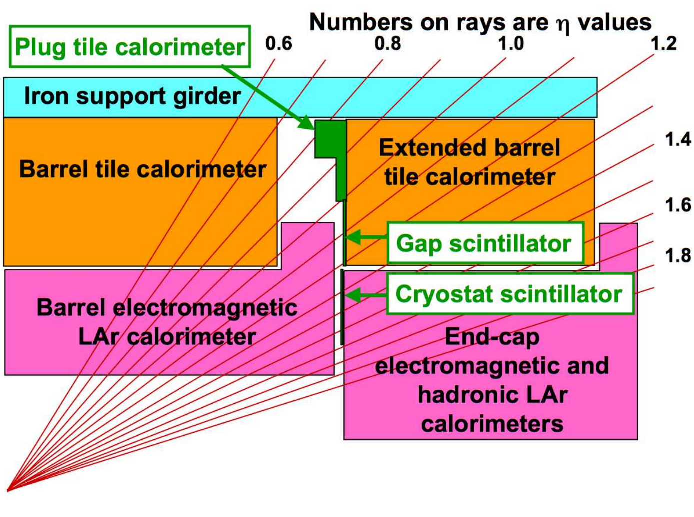
2.2.3.2.3 Forward calorimeters

The forward calorimeters [88] are located at each end-cap cryostat covering the region $3.1 < |\eta| < 4.9$ and are instrumented with LAr (as active material) to handle the high radiation. The forward calorimeters at each side is split into three modules, in a flat-plate design: an innermost electromagnetic layer is composed of copper/LAr modules (FCal1), and two outermost hadronic layers are made of cooper-tungsten/LAr modules. Cooper plates facilitate heat interchange, while the tungsten absorbers minimise the lateral spread of hadronic showers. The overall depth of these detectors is approximately 11 nuclear interaction lengths (see Figure 2.15).

2.2.3.3 Instrumentation in gap between calorimeters

The gaps between the barrel and end-cap calorimeters, known as transition regions, used for cables and services for the inner detector as well as power supplies and services for the barrel LAr calorimeter. At the outer radius of the detector, a reduced section of a standard TileCal sub-module, the plug, provides additional coverage in this region and significantly reduces the neutron flux from the inner-detector volume into the muon system. Gap and cryostat scintillators, covering respectively the regions $1.0 < |\eta| < 1.2$ and $1.2 < |\eta| < 1.6$, provide signals which can be used to correct for energy losses in the inactive material in the transition regions. Both gap and cryostat scintillators have full coverage in ϕ, with a segmentation of $\Delta\phi = 0.1$. The transition regions together with the added instrumentation are shown in Figure 2.16.

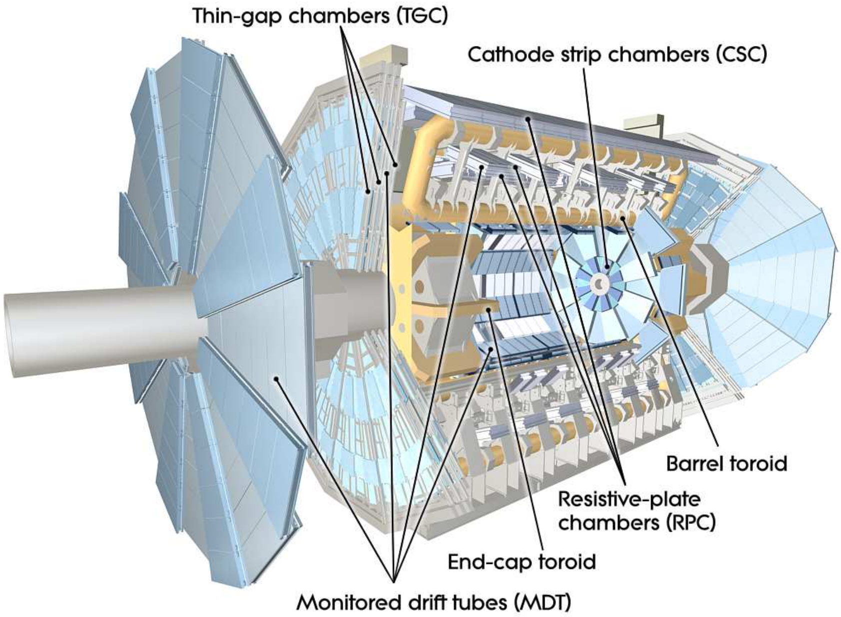
Figure 2.16: Schematic of the transition regions between the barrel and end-cap calorimeters, where additional scintillator elements are installed to provide corrections for energy lost in the inactive material.



2.2.4 Muon spectrometer

The muon spectrometer (MS) is responsible for the momentum measurement of charged particles exiting the barrel and end-cap calorimeters in the pseudorapidity range $|\eta| < 2.7$ and select events of interest in the region $|\eta| < 2.4$. The magnetic field for the muon spectrometer is produced by the barrel and end-cap air-core toroid magnets. The air-core magnet concept for the muon spectrometer minimises the amount of material traversed by the muons after exiting the calorimeters. The layout of the muon system and the magnets is shown in Figure 2.17. The muon spectrometer consists of four primary sub-systems, two precision muon trackers: the monitored drift tubes (MDT) and cathode strip chambers (CSC), and two triggering sub-systems: resistive plate chambers (RPC) and thin gap chambers (TGC), that are briefly described next.

Figure 2.17: Cut-away view of the ATLAS muon system. Image from Ref. [35].



2.2.4.1 Precision sub-systems

The MDT chambers cover the region $|\eta| < 2.7$, except in the innermost end-cap layer where their coverage is limited to $|\eta| < 2.0$. The basic element is a pressurised drift tube with a diameter of $\sim 30\,\text{mm}$, operating with Ar/CO_2 gas (93/7%) at 3 bar pressure (this mixture was selected due to its aging properties and small likelihood of forming deposits within the tube). The electrons resulting from ionisation due to crossing muons are collected at the central gold plated tungsten-rhenium wire acting as an anode. The maximum drift time from the wall to the wire is of about 700 ns, when operated at a nominal potential of 3080 V. The direction of the tubes in the barrel and end-caps is along ϕ. Signal transmission to the electronics and connection to the HV supply system are at opposite ends.

The limit for safe operation of the MDTs is at counting rates of about $150\,\text{Hz/cm}^2$, which is exceeded in the region $|\eta| < 2$ for the first layer of the end-cap. In this region, the

MDTs are replaced by CSCs which are safe up to counting rates of about 1000 Hz/cm^2, sufficient up to the forward boundary of the barrel muon system at $|\eta| = 2.7$. The CSCs are multiwire proportional chambers with the wires oriented in the radial direction. To either side of the wire plane are cathode strip planes separated from the anode by 2.5 mm. Both cathodes are segmented, one with the strips perpendicular to the wires, and the other cathode has strips parallel to the wires. The readout is performed on the cathodes. Both sub-systems are located between the eight coils of the superconducting barrel toroid magnet, while the end-cap chambers are in front and behind the two end-cap toroid magnets.

2.2.4.2 Trigger chambers

An essential design criterion of the muon system was the capability to trigger on muon tracks. The precision-tracking chambers have therefore been complemented by a system of fast trigger chambers capable of delivering track information within a few tens of nanoseconds after the passage of the particle. In the barrel region ($|\eta| < 1.05$), Resistive Plate Chambers (RPC) were selected for this purpose, while in the end-cap ($1.05 < |\eta| < 2.4$) Thin Gap Chambers (TGC). Both chamber types deliver signals in less than 25 ns, thus providing the ability to tag the beam-crossing and allowing the Level 1 trigger logic (see Section 2.2.7) to recognise muon multiplicity and approximate energy range. The trigger chambers measure both coordinates of the track, one in the bending (η) plane and one in the non-bending (ϕ) plane.

The RPC is a gaseous parallel electrode-plate (i.e. no wire) detector, formed by two resistive plates, kept parallel at a distance of 2 mm by insulating spacers. The electric field between the plates of about 4.9 kV/mm allows avalanches to form along the ionising tracks towards the anode.

TGCs provide two functions in the end-cap muon spectrometer: the muon trigger capability and the determination of the second azimuthal coordinate to complement the measurement of the MDTs in the bending (radial) direction. TGCs are multi-wire proportional chambers with the characteristic that the wire-to-cathode distance of 1.4 mm is smaller than the wire-to-wire distance of 1.8 mm. The high electric field around the TGC wires and the small wire-to-wire distance lead to very good time resolution of 4 ns. Only tracks at normal incidence passing midway between two wires have much longer drift times due to the vanishing drift field in this region. This high electric field necessitates a quenching gas mixture of 55% carbon dioxide and 45% n-pentane. The TGCs, forming circular disks, are mounted in two concentric rings, an outer (end-cap) one covering the rapidity range $1.05 < |\eta| < 1.92$ and an inner (forward) one covering the rapidity range $1.92 < |\eta| < 2.4$.

2.2.5 Forward detectors

The forward detectors are responsible for the measurement of luminosity and beam conditions, are located at the limits or outside the main detector volume (but form part of the ATLAS sub-systems), and are discribed next.

2.2.5.1 The Beam Condition Monitor

One of the worst-case scenarios during LHC operation arises if several proton bunches hit the collimators in front of the detectors. While the accumulated radiation dose from such unlikely accidents corresponds to that acquired during a few days of normal operation, and as such provides no major contribution to the integrated dose, the enormous instantaneous rate might cause detector damage. The Beam Condition Monitor (BCM) detects such incidents and trigger an abort (beam dump) in time to prevent serious damage to the detector. It also helps determining the luminosity.

BCM, consists of two stations, each with four modules, placed symmetrically around the interaction point at $z = \pm 184\,\text{cm}$ and 5.5 cm away from the beam line, which corresponds to $|\eta| = 4.2$. Each module, includes two radiation-hard diamond sensors. The difference in time-of-flight between the two stations distinguishes particles from normal collisions ($\Delta t = 0, 25, 50\,\text{ns}$, etc.) from those arising from stray protons or beam-gas background which frequently initiate charged particle showers, originated well upstream (or downstream) from the ATLAS interaction point. Due to their very fast response time and intrinsically very high resistance to radiation, the BCM detectors will be used throughout the lifetime of the experiment to distinguish these stray beam particles from those originating from proton-proton interactions.

In Run 2 the Diamond Beam Monitor (DBM) [89] was added to complement the BCM, which can saturate at instantaneous luminosity of $10^{34}\,\text{cm}^{-2}\text{s}^{-1}$. The DBM, located approximately one metre from the IP, adds tracking capabilities with chemical vapor deposition diamond sensors.

2.2.5.2 LUCID detector

LUCID (LUminosity measurement using Cherenkov Integrating Detector) is designed to detect inelastic pp scattering in the forward direction, in order to both measure the integrated luminosity and to provide online monitoring of the instantaneous luminosity and beam conditions. LUCID is placed at $\pm 17\,\text{m}$ from the interaction point and 10 cm apart from the beam line, covering the range $|\eta| \sim 5.8$. LUCID consists of 20 aluminum cones directed at the IP filled with C_4F_{10} at 1.2–1.4 bar, in order to measure Cherenkov radiation[7] from incident particles. By assuming the number of particles detected is proportional to the number of interactions in the bunch crossing, LUCID is able to measure μ in order to calculate the luminosity.

In Run 2, the original LUCID detector was upgraded to LUCID-2 [90] to deal with several challenges: fast aging from a large total integrated luminosity, constrains on electronics from 25 ns bunch crossings, increased particle fluxes saturating PMTs and counting algoritms, and a new beam pipe material.[8] The new design no longer uses a gas, but rather quartz windows as the Cherenkov medium.

2.2.6 The solenoidal and toroidal magnet system

The magnetic field responsible for bending the trajectory of charged particles is provided by the ATLAS magnet system. Thus, the momentum of particles can be measured via the radius of curvature of the tracks left within the detector systems. This magnetic system is 22 m in diameter and 26 m in length, with a stored energy of 1.6 GJ. ATLAS features a unique hybrid system of four large superconducting magnets, as shown in Figure 2.18.

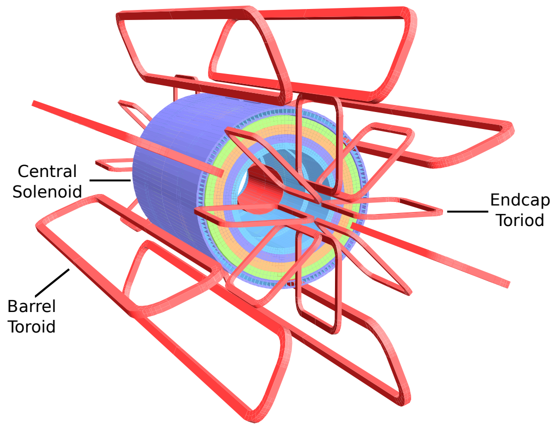
[7] Radiation emitted when charged particles travel through a dielectric medium faster than the velocity of light in that medium.

[8] Moving from stainless steel to aluminum approximately increases the flux of particles through LUCID by a factor of four.

The central solenoid has an inner radius of 1.23 m, a total length of 5.8 m, and sits just outside and surrounding the inner detector cavity. It shares the cryostat with the LAr EMB calorimeter, but in a separated vessel as the superconducting solenoid is kept at 4.5 K. A 2 T axial field is generated with 7.73 kA flowing through a single-layer aluminum coil wound by a niobium-titanium (NbTi) conductor. The flux is returned by the steel of the ATLAS hadronic calorimeter and its girder structure.

Three large superconducting toroids (one barrel and two end-caps), arranged with an eight-fold azimuthal symmetry around the calorimeters, provide bending power for the muon spectrometer. The barrel toroid extends 25.3 m along the z axis whereas the end-cap toroids have 5 m in lenght. The toroidal magnets implement coils consisting of a conductor made of a mixture of aluminum, copper and an alloy of NbTi at 20.5 kA and provide 0.2–2.5 T in the barrel and 0.2–3.5 T in the endcap.

Figure 2.18: Geometry of magnet windings and tile calorimeter steel. The eight barrel toroid coils, with the end-cap coils interleaved are visible. The solenoid windings lie inside the calorimeter volume. The tile calorimeter is modelled by four layers with different magnetic properties, plus an outside return yoke. For the sake of clarity the forward shielding disk is not displayed. Image from Ref. [35].

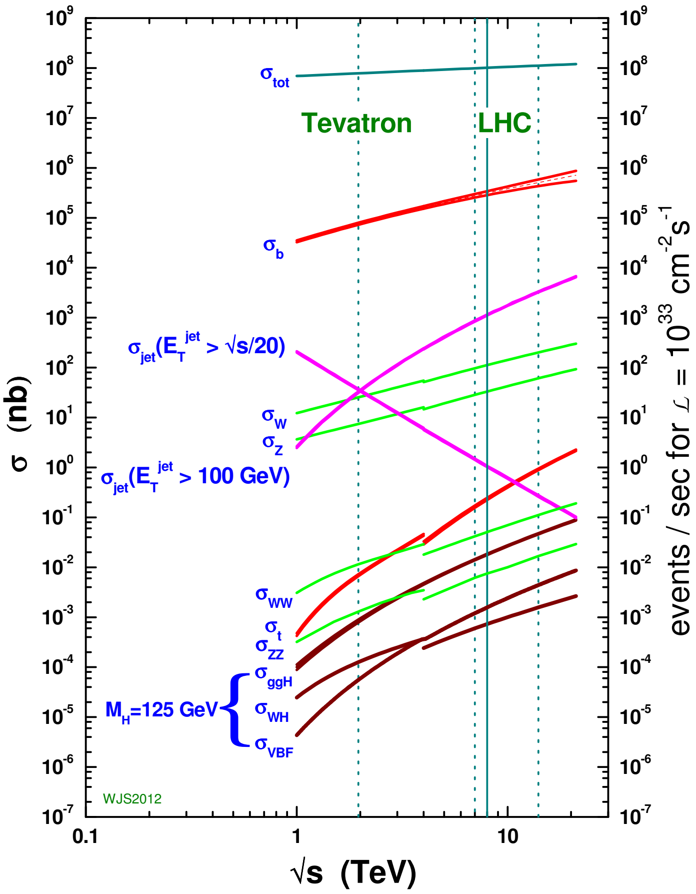


2.2.7 Trigger and data adquisition in Run 2

The proton-proton interaction rate at the LHC is 40 MHz (proton bunches collide every 25 ns). Since the information coming from the $\sim$ 100 million readout channels from all subsystems constitute 1.6 MB of data per event, if no filtering is applied, this corresponds to a total recording rate near 60 TB/s. At this rate and due to limitations in processing power and storage in the data acquisition system[9], a continuous extraction of collision data is not feasible. In addition, the vast majority of processes of interest have smaller cross-sections than the inclusive collisions, dominated by low energy interactions, and therefore there is no practical reason to save the information of events which do not provide interesting physics. In addition, the cross-section is massively dominated by QCD interactions such as multijet production, whereas processes like Higgs boson production only contribute few *parts per trillion* to the total cross-section. This can be seen in Figure 2.19, which shows cross-section predictions for different physics processes, obtained for $p\bar{p}$ and pp collisions. Hence, a sophisticated system is needed for selecting and reconstructing events online

[9] The event data recording rate, based on technology and resource limitations, is of the order of 1 kHz [91].

Figure 2.19: Cross sections from proton collisions as a function of centre-of-mass energy. The discontinuity between Tevatron [92] and LHC energies reflects the switch from $p\bar{p}$ to pp collisions. Image from Ref. [93].



efficiently (maximising the rejection of unwanted events while minimising the rejection of interesting events).

The ATLAS Trigger and Data Acquisition (TDAQ) [94] system is a highly optimised network of hardware and software providing a solution to this problem. This trigger system is an essential component of the experiment as it is responsible to make (fast) the decision to whether or not select given bunch-crossing for later study. In Run 2, the TDAQ system consists of a hardwared-based first-level trigger (L1) and a software-based high-level trigger (HLT) with a total rate of approximately 1.6 GB/s. The original design of the TDAQ system, utilised in Run 1, consisted on a three level triggering system with a final recording rate of 960 MB/s, where the HLT consisted of separate Level-2 (L2) and Event Filter (EF) triggers. In Run 2, nearly all components received an upgrade, to maintain acceptably low p_{T} thresholds while dealing with increasingly challenging demands: increased centre-of-mass energy from 8 TeV to 13 TeV, increased pile-up, decreased bunch spacing interval from 50 ns to 25 ns, and increased instantaneous luminosity [91, 95]. In addition, L2 and EF triggers were merged into a single homogeneous system, the HLT, simplifying development and reducing the CPU and network consumption [96]. A comparison of LHC and TDAQ parameters between Runs 1 and 2 is deprecated in Table 2.1. A comprehensive description of the Run 1 Trigger and Data Acquisition system of the ATLAS experiment can be found in Ref. [97], while the performance of the Run 2 TDAQ system is detailed in Ref. [91].

Table 2.1: Comparison of LHC and TDAQ parameters between Run 1 (8 TeV) and Run 2 (13 TeV, up to 2016). The increased luminosity, centre-of-mass energy, and pile-up required updated algorithms and architecture to address the increased production rate and to keep the output rate reasonable. Values from Refs. [76, 95, 98, 99].

Property	Run 1	Run 2
$\sqrt{s}$ [TeV]	8	13
L_{peak} [10^{33}cm^{-2}s^{-1}]	7.73	13.8
Mean number of interactions per crossing	20.7	23.7
Bunch spacing [ns]	50	25
L1 rate [kHz]	70	100
Output rate [kHz (GB/s)]	0.6 (0.96)	~ 1 (~ 1.6)

The L1 trigger consists of fast, custom-made electronics that can perform event selections using complex quantities such as missing transverse energy. It is capable of reducing the data rate from 40 MHz down to 100 kHz; however, it is built for speed rather than precision measurements.

The L1 trigger receives coarse measurements from the calorimeters and muon systems. Sums of calorimeter energy deposits from 7000 trigger towers are collected by the "L1Calo" trigger, while RPC and TGC track information is collected by the "L1Muon" trigger. Event data is buffered on detector-specific front-end electronics while the L1 trigger decision is being formed. Muon triggers at L1 are based on a coincidence of hits among several layers of the trigger muon chambers. Electrons and photons are triggered on energy deposits in the EM calorimeter. Jet candidates are constructed at L1 from coarse calorimeter towers made of trigger-elements using a sliding window algorithm [100]. A trigger-element is defined as a sum of cells in a $0.2 \times 0.2\,(\eta - \phi)$ region, and a sliding window algorithm examines the total E_{T} against a trigger threshold value in a 4×4 region. Various criteria are used to decide which events pass the L1 trigger and the HLT, such as the p_{T} and η of reconstructed objects or the calculated missing E_{T} of an event (E_T^{miss}). For each set of criteria, unique trigger exists. A "trigger menu" consisting of 512 "trigger items", logical combinations of trigger inputs from L1Calo and L1Muon, is processed by the Central Trigger Processor (CTP) [101, 102] which produces a trigger decision within 2.5 μs. The items in the trigger menu can be "prescaled" so that a predetermined fraction of events passing the trigger are randomly ignored, thereby reducing the overall rate. Triggers that are more stringent in their selection are unprescaled, and are able to save all passing events. The CTP is also responsible for limiting the time between L1 accepts (simple dead time) and a leaky bucket algorithm to limit the number of L1 accepts in a given number of bunch crossings (complex dead time) [103]. A detector may issue a "busy" signal to the CTP in order to signal that its buffers are full, preventing further L1 accepts until the busy signal is cleared. After the L1 trigger acceptance, the events are buffered in the ReadOut System (ROS) and processed by the HLT. The L1 trigger sends the η and ϕ coordinates which caused the L1 trigger to fire, the so-called "region of interest" (RoI), to the HLT.

The second level of the triggering system is the software-based HLT, which reduces the 100 kHz L1 rate to a final ~ 1 kHz event output rate, with an average latency (decision time) of 200 ms. The HLT has access to fine-granularity calorimeter information, precision measurements from the MS, and tracking information from the ID, which are not available at L1. The HLT uses reconstruction algorithms and calibrations similar to those used offline. As needed, the HLT reconstruction can either be executed within RoIs (approximately 2% of the total event data) identified at L1 or for the full detector. In both cases the data is retrieved on demand from the readout system. Several (~ 2500) triggers are provided in a trigger menu, based on the number of objects, their transverse momentum, missing transverse energy, and certain identification and isolation criteria. As in Run 1, in order to reduce the processing time, most HLT triggers use a two-stage approach with a fast first-pass reconstruction to reject the majority of events and a slower precision reconstruction for the remaining events. However, with the merging of the previously separate L2 and EF triggers, there is no longer a fixed bandwidth or rate limitation between the two steps. Figure 2.20 shows different triggers and their rates for data taking in October 2015 and July 2016. The triggers which are used in this thesis are discussed in detail later.

If an event is accepted, the data is stored to perform full offline event reconstruction (see next section). The luminosity of the proton beams decreases steadily as collisions progress, causing data collection conditions to vary as a function of time. Data is therefore collected in time intervals of one minute, referred to as lumi-blocks. Each one is assigned an integrated luminosity depending upon detector and beam conditions over that lumi-block. They are also useful for isolating problematic data which may have been taken while a sub-detector was inoperative or malfunctioning, allowing their easy rejection during data analysis. ATLAS provides lists of lumi-blocks of "good" data for each data taking period known as Good Run Lists (GRL). Most events passing the HLT are written to the "main" physics stream, while a few events which require a larger processing time are saved to a "debug" stream for future reprocessing.

2.3 Object reconstruction and identification

The ATLAS reconstruction software uses a wide variety of different algorithms to identify individual particles and jets from tracks and energy depositions recorded by the detector. As these algorithms greatly influence the efficiency and performance of the reconstruction process, they are under constant development. This section gives a short overview on some of the algorithms used to reconstruct events recorded by ATLAS during LHC Run 2 (up to 2016). The E_T^{miss} and hadronically decaying τ leptons reconstruction is not discussed here but are described in detail in Refs. [105] and [106], respectively.

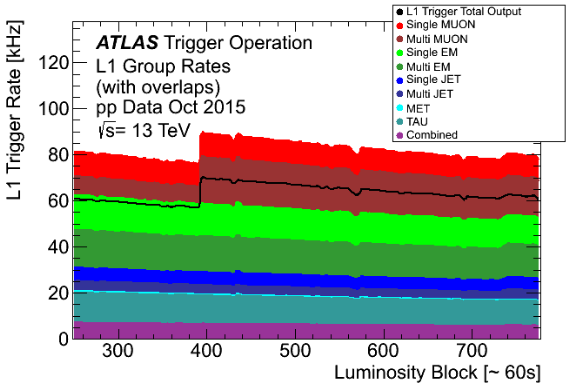
2.3.1 Track and vertex reconstruction

2.3.1.1 Track reconstruction

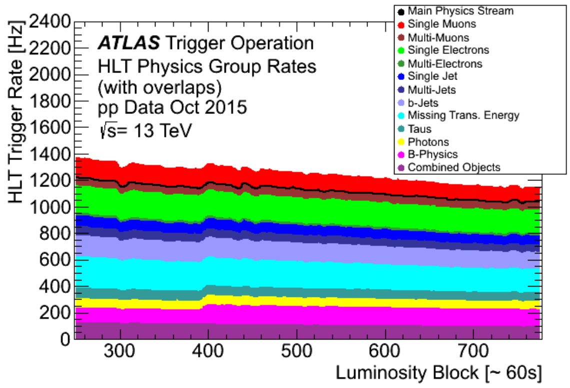
Track reconstruction is the process of identifying individual charged particles emerging from the interaction and measuring their traces in the inner detector. Such charged particles can be measured up to $|\eta| = 2.5$, which is the η coverage of the tracking system (see Section 2.2.2). With the curvature radius and the known magnetic field, it is ultimately possible to calculate the transverse momenta of particles. This procedure is sub-divided in four stages [107]:

- Individual cells from the silicon detector are clustered by connecting adjacent hits. Cluster splitting is performed using a neural network. Timing information from the TRT is converted into drift-circles.

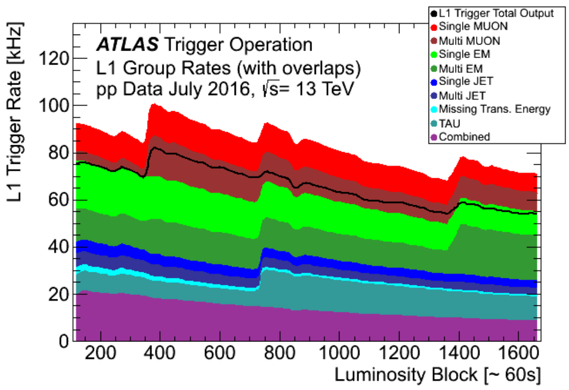
Figure 2.20: Physics trigger group rates for the L1 and HLT triggers are presented ((a,c) and (b,d) respectively). The rates displayed in panels (a,b) ((c,d)) are shown as a function of the number of luminosity blocks which correspond to an average 60 s per luminosity block in a fill taken in October 2015 (July 2016) with a peak luminosity of $L = 4.6\,10^{33}\text{cm}^{-2}\text{s}^{-1}$ ($L = 1.02\,10^{34}\text{cm}^{-2}\text{s}^{-1}$) and an average pile-up of $\mu = 14.7$ ($\mu = 24.2$). In panels (a,c), the rate of the individual L1 trigger groups specific for trigger physics objects at L1 are presented. Overlaps are only accounted for in the L1 trigger total output rate. Panels (b,d) shows the rate of the individual HLT trigger groups specific for trigger physics objects. Overlaps are only accounted for in the total Main Physics Stream rate. In the b-jet (i.e. jet initiated by a quark b) as well in the tau group, the multi object triggers are contained. The B-physics triggers are mainly muon based trigger algorithms. The combined group represents multiple triggers of different objects, as combinations of electrons, muons, taus, jets and missing transverse energy. Common features to all rates are their exponential decay with decreasing luminosity during an LHC fill. The rates periodically increase due to change of prescales to optimise the bandwidth usage, dips are due to deadtime and spikes are caused by detector noise. Images from Ref. [104].



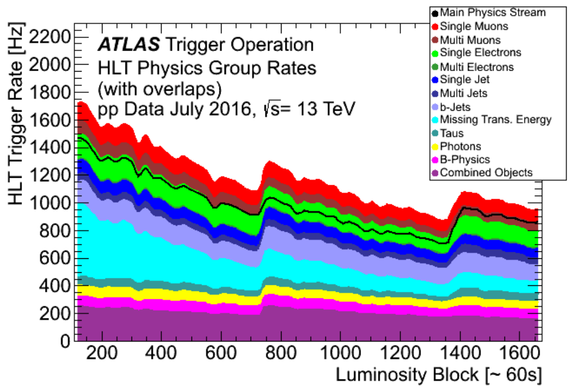
(a)



(b)



(c)



(d)

- Track seeds are identified from a combination of pixel and strip layer hits, requiring a total of three hits. They are processed by applying a set of thresholds as well as a combinatorial Kalman filter to form a combined particle track.
- Ambiguities are resolved using a scoring mechanism taking into account double usage of hits in several tracks of the same event as well as layers missing a hit for a specific track.
- Tracks within the TRT coverage that survive the ambiguity solver are extended and combined with the TRT tracking information. This significantly increases the momentum resolution due to the increased lever.

After tracks have been identified, splines are fit and combined with data about the magnetic field strength to measure the momentum of the particle associated to the track.

2.3.1.2 Vertex reconstruction

A "primary vertex" (PV) is the location where two protons collide. Primary vertices are reconstructed by matching up intersecting tracks, which proceeds in three main steps: seeding, track assigment, and fitting. A rough outline of the procedure is as follows:

- All reconstructed tracks satisfying certain criteria are extrapolated to the beam axis. The impact parameter (z_0), defined as the longitudinal distance at the point of closest approach, is computed for every track with respect to the nominal beam spot. This information is used to generate seeds, for which an iterative likelihood maximisation method is used.
- Tracks compatible with a seed are grouped together for fitting.
- An adaptive vertex fitting algorithm [108] is used to estimate the position and uncertainty of the vertex.
- Incompatible tracks not used in a previous vertex fit are used to repeat the procedure starting from the creation of a new seed.

The vertex with the largest $\sum p_{\mathrm{T}}^2$ for all associated tracks is labeled as the hard-scatter vertex. All other reconstructed vertices are considered to be pile-up. The identification of vertices becomes more challenging as the average number of interactions per bunch crossing (the pile-up) increases, and thus new algorithms are being developed to increase the resilience of ATLAS vertexing [109].

2.3.2 Electron reconstruction

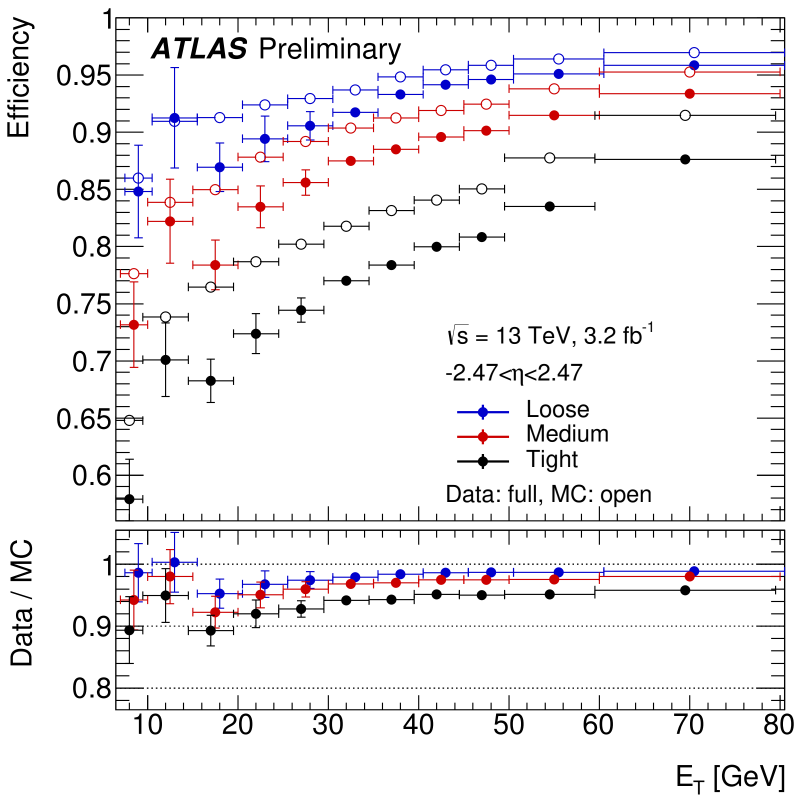
Electrons are reconstructed by matching reconstructed tracks in the ID to clusters of EM calorimeter energy deposits, whose shower profile is consistent with an EM shower [110]. To build the EM cluster, a sliding window algorithm is used [100]. The window is a fixed-size grid of cells, $N_\eta \times N_\phi = 3 \times 5$, in the middle layer of the LAr calorimeter where approximately 80% of the energy of an EM shower is deposited. For each cell, the energy is summed across all the longitudinal layers, forming a tower. If the transverse energy of the towers in the window is above 2.5 GeV and is a local maximum, a seed cluster is formed.

Pattern recognition and track fitting algorithms are then used to match ID tracks to seed clusters. Track seeds from the ID with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$ are tested against pion and electron pattern recognition algorithms, and are required to reside in a cluster seed. A track fit is performed twice to extrapolate the ID track to the calorimeter, using the track momentum or the cluster momentum, and is required to be near the cluster. If either succeds, a final optimised track reconstruction is performed using a Gaussian Sum Fitter (GSF) [112].

An electron candidate is formed if there is at least one ID track matched to the seed cluster. The EM cluster is then rebuilt by summing energy in a grid of 3×7 (5×5) cells in each layer of the barrel (end-cap) region, starting in the middle layer. After corrections are applied[10], the four-momentum of the electron candidate is calculated using the energy measurement of the EM cluster and the η and ϕ measurements from the track.

Background objects like hadronic jets, non-prompt electrons from photon conversions, and semi-leptonic decays of hadrons with heavy quarks can "fake" an electron signature and be reconstructed at this stage as an electron candidate [111]. A multivariate analysis technique (MVA) [113] takes into account several cluster and track variables to create a likelihood identification for each candidate as either signal or background.[11] Different working points with different levels of electron efficiency and background rejection are provided. These are "Loose", "Medium", and "Tight", which correspond respectively to approximately 95%, 90% and 80% identification efficiency for an electron with $E_T >$ 40 GeV. The reconstruction and identification efficiencies for all working points are show in Figure 2.21.

Figure 2.21: Reconstruction and identification efficiencies for signal electrons as a function of the transverse energy E_T, integrated over the full pseudorapidity range, are shown for several working points. Image from Ref. [110].



Further suppresion of background electrons is provided by isolation criteria [110]. This allows to select, as much as possible, prompt electrons, removing those coming from hadron decays and contained within jets. For the isolation requirements, both tracking

[10] A correction to the clusters is applied taking into account the radiative losses suffered by the electrons while traversing the silicon tracker by virtue of a Gaussian-Sum-Filter [112].

[11] Multivariate analysis techniques are used extensively in physics analyses to separate signal from background, since they allow, in contrast to cut-based methods, the simultaneous evaluation of several properties when making a selection decision.

and calorimeter information is used. Working points using calorimeter and/or track based isolation are provided for either a specific efficiency with variable cuts or fixed cut values. The calorimeter isolation uses the sum of the transverse energy (E_{T}) in a cone of $\Delta R < 0.2$ around the electron centre ($E_{\mathrm{T}}^{\mathrm{cone0.2}}$), substracting the 5×7 cell window contribution from the electron. For the track isolation, the sum of the transverse momenta for quality tracks in a variable cone of size $\Delta R < \min(0.2, 10\,\mathrm{GeV}/E_{\mathrm{T}})$ centred around and excluding the electron track, $p_{\mathrm{T}}^{\mathrm{varcone0.2}}$, is used.

2.3.3 Photon reconstruction

At photon energies above 1 GeV, the interaction of the photons with the tracker is completely dominated by e^+e^- pair production, also known as "photon conversion". All other interactions between the photons and the tracker material, such as Compton or Rayleigh scattering, have cross-sections which are orders of magnitude below photon conversion, and can thus be ignored.

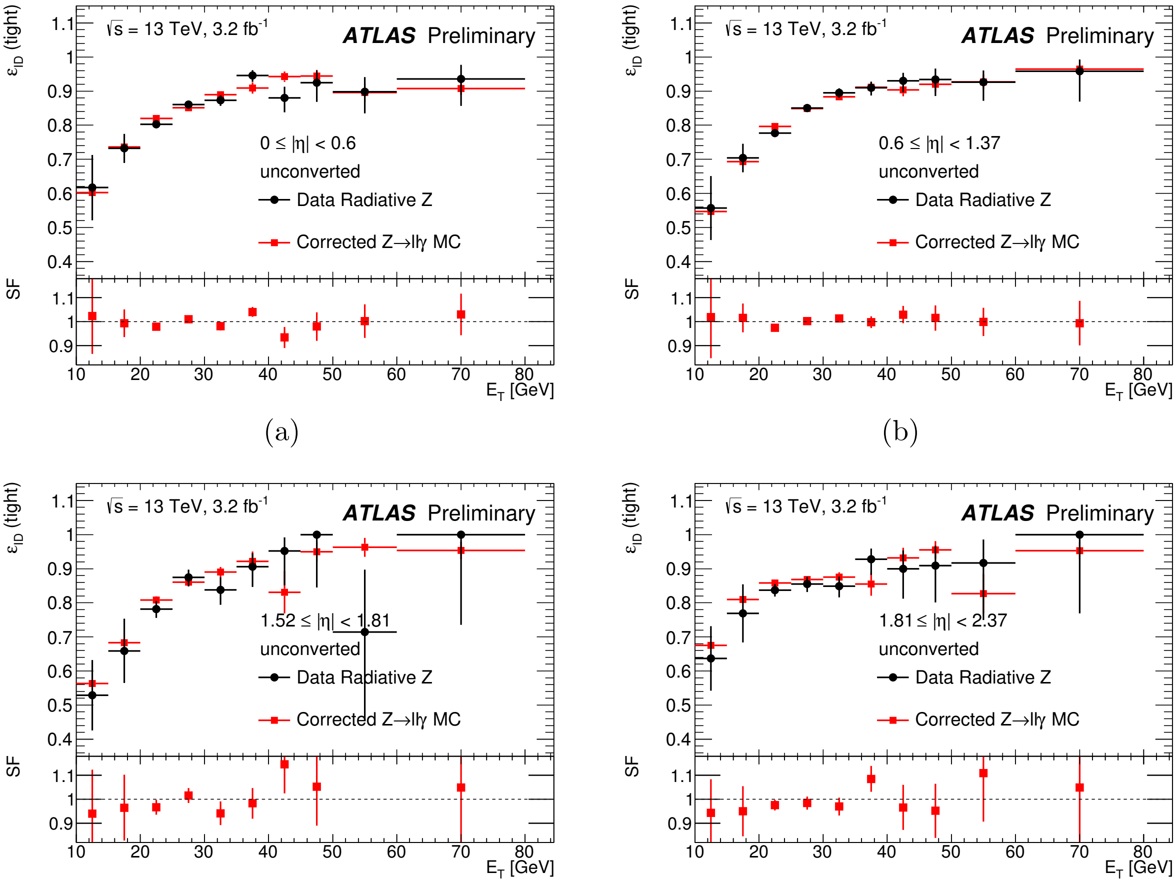
Photon candidates are reconstructed from clusters of energy, deposited in the EM calorimeter [114, 115]. Those candidates that are matched to either a reconstructed conversion vertex formed from two tracks constrained to originate from a massless particle or a single track with its first hit after the innermost layer of the pixel detector, are classified as "converted photons". Candidates without a matching track or reconstructed conversion vertex in the ID are classified as "unconverted photons". Those matched to a track consistent with a track originating from an electron from the hard-scatter vertex are kept as electrons.

As it is the case of electrons, also photon identification criteria are based on shower shapes in the calorimeter system as well as discriminating variables (DVs). A loose selection is derived using only the information from the hadronic calorimeter and the lateral shower shape in the second layer of the electromagnetic calorimeter, which contains most of the photon energy. The final tight selection applies tighter criteria to these variables, based on the shower shape in the finely segmented first calorimeter layer to ensure the compatibility of the measured shower profile with that originating from a single photon impacting the calorimeter. In all cases, the requirements on the DVs are tuned separately for unconverted and converted photons, in several pseudorapidity regions. To further reject the background from jets misidentified as photons, the photon candidates are required to be isolated using both calorimeter and tracking detector information.

The photon efficiency is shown in Figure 2.22 for data and MC simulations for different pseudorapidity intervals. For the photon efficiency measurements, a data sample enriched in $Z \to ll\gamma$ events is exploited. No significant differences are observed between data-driven measurements and simulations. This is because, in order to improve the description of the photon DVs, corrections are applied to the simulated values, by applying a shift to each of them, whose value is optimised separately for unconverted and converted photons, and as a function of the pseudorapidity. The small differences in the efficiency between the simulation and the data are taken into account by computing data-to-MC efficiency ratios, also referred to as scale factors, to correct the identification efficiency in MC.

The energy of each photon used in this thesis is corrected using MC simulation and data as described in Ref. [116]. The EM energy clusters are calibrated separately for converted and unconverted photons, based on their properties including the longitudinal shower development. The energy scale and resolution of the photon candidates after the MC-based calibration are further adjusted based on a correction derived using $Z \to e^+e^-$

Figure 2.22: Comparison of the radiative Z boson data-driven efficiency measurements of unconverted photons to the $Z \rightarrow ll\gamma$ simulation as function of E_T in the region $10 \leq E_T \leq 80\,\text{GeV}$, in the four pseudorapidity intervals. The bottom panels show the ratio of the data-driven results to the MC predictions ("scale factor", SF). The uncertainty bars represent the sum in quadrature of the statistical and systematic uncertainties. Images from Ref. [114].



(a) (b) (c) (d)

events from data and MC simulation.

2.3.4 Muon reconstruction

Muon reconstruction in ATLAS is performed exploiting the information from the ID and the MS [117]. Until they reach the MS, muons traverse typically 100 radiation lenghts, mostly instrumented by the EM and hadronic calorimeters. Muons produced in the hard interactions in the LHC tipically have energies between 10 GeV and 1 TeV. At these energies, muons can be treated as Minimum Ionising Particles (MIP). They deposit an amount of energy in the calorimeter system of about 3–4 GeV [118].

In the ID, muons are reconstructed like any other charged particle. A Hough transform [119] is used to search for hits aligned on a trajectory in the bending plane of the detector. Muon track candidates are then built by fitting together hits from segments in different layers. The segments are selected using criteria based on hit multiplicity and fit quality and are matched using their relative positions and angles. The information on energy deposit in the calorimeter system is also used and corrections are applied to take into account the energy loss in the material.

Four different types of muons are defined in ATLAS, depending on which sub-detectors are used in reconstruction. They are listed below:

- **Combined (CB) muons:** track reconstruction is performed independently in the ID and MS, and a combined track is formed with a global fit that uses the hits from both sub-detectors. Muons are firstly reconstructed in the MS, where the track density is much smaller, and then extrapolated inward and matched to an ID track. They are the most commonly used muons in physics analyses since they have the highest purity and the best kinematic resolution.

- **Segment-tagged (ST) muons:** a muon is classified as ST if a track in the ID, once extrapolated to the MS, is associated with at least one local track segment in the MDT or CSC chambers. ST muons are used when muons cross only one layer of MS chambers, either because of their low p_{T} or because they pass through regions with reduced MS acceptance.

- **Calorimeter-tagged (CT) muons:** a track in the ID is identified as a muon if it can be matched to an energy deposit in the calorimeter compatible with a MIP. This kind of muons have the lowest purity with respect to all the other muon types but they allow to recover the acceptance in regions where MS is partially instrumented (to host cabling and services), close to $\eta = 0$.

- **Extrapolated (ME) muons:** the muon trajectory is reconstructed based only on the MS track and a loose requirement on compatibility with originating from the IP. ME muons are mainly used to extend the acceptance for muon reconstruction into the region $2.5 < |\eta| < 2.7$, which is not covered by the ID.

In the same way as for electrons, muon candidates are defined as "Loose", "Medium", and "Tight", with increasing purity, according to the quality of the reconstruction and identification[12].

Muon reconstruction efficiency as a function of η is shown in Figure 2.23 for Loose, Medium, and Tight muons. The ATLAS muon reconstruction efficiency is very high (close to 100%).

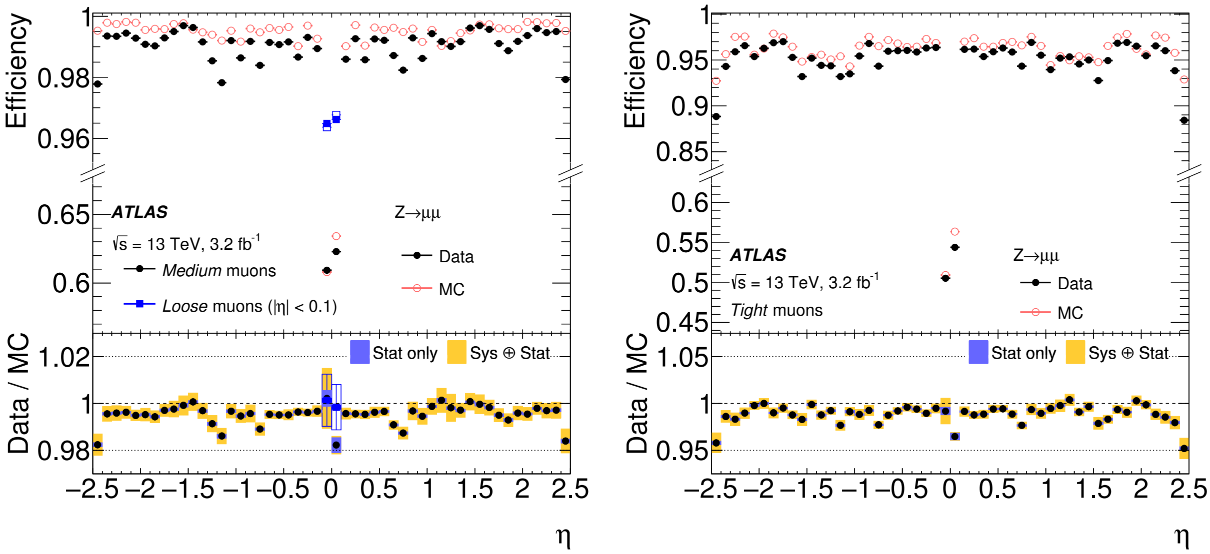
Concerning the isolation requirements, similar considerations, as for electrons, are made for muons: track-based and calorimeter-based isolation measurements are used to create working points with either 1) efficiency based bariable cuts or 2) fixed cuts. The track-based muon isolation, $p_{\mathrm{T}}^{\mathrm{varcone0.3}}$, uses a slightly larger cone, $\Delta R < 0.3$, with respect to the electron case.

2.3.5 Jets

Jets are the experimental realisation of the partonic showering and subsequent hadronisation of particles carrying colour charge. In this section, the jet algorithms used to reconstruct jets are described. A brief characterisation is also given of what are the inputs used to build the experimental jets.

[12] In fact, for muons there is another identification criteria known as "High-pT", which is optimised for searches for high-mass Z' and W' resonances.

Figure 2.23: Panels (a) and (b) show the muon reconstruction efficiency as a function of η measured in $Z \to \mu\mu$ events for muons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ for respectively the Medium and Tight muon selections. Panel (a) also shows the efficiency of the Loose selection (squares) in the region $|\eta| < 0.1$ where the Loose and Medium selections differ significantly. The error bars on the efficiencies indicate the statistical uncertainty. Panels at the bottom show the ratio of the measured to predicted efficiencies, with statistical and systematic uncertainties. Images from Ref. [117].



(a) (b)

2.3.5.1 Jet algorithms

There is no fundamental, quantitatively unique definition of a jet. The jet definition may be adapted for a specific application (i.e. particular signatures) or experimental condition (high/low luminosity, high/low centre-of-mass energy, etc.) in order to reconstruct more closely the hard scattering process. This flexibility is an accepted practice in many analyses, as long as the same definition is consistently used when comparing data to theoretical calculations or Monte Carlo simulations. Nonetheless, to ensure stable definitions, jet algorithms must satisfy reasonable criteria:

- Infrared and collinear (IRC) “safe” (see below)
- Minimal sensitivity to hadronisation, underlying event and pile-up
- Applicability in both experimental analyses and theoretical calculations
- Reliability of (finite) results at all orders in perturbation theory

The above list is an extension of the “Snowmass accord” [120] originally aimed to facilitate the comparison among jet physics results, that still provides a common ground for the development of modern jet algorithms.

Commonly, ATLAS uses jet definitions from the family of sequential recombination algorithms [121–123]. These were specifically designed to be insensitive to the presence of soft (infrared) and collinear gluon emission, and are thus usable for calculations done to any order in perturbation theory. Recombination algorithms try to reverse the pattern of QCD parton showers and multi-gluon radiations characterised by softer emissions as the shower evolves [124]. For this, the algorithm defines a distance d_{ij} between entities (particles or energy deposits in the calorimeter) and between entity i and the beam d_{iB}:

$$\begin{aligned} d_{ij} &= \min\left(p_{T_i}^{2p}, p_{T_j}^{2p}\right)\frac{\Delta R_{ij}^2}{R^2} \\ d_{iB} &= p_{T_i}^{2p}, \end{aligned} \tag{2.6}$$

where ΔR_{ij}^2 is defined in Equation 2.4, p_{Ti} is the transverse momentum of entity i, and parameters p and R regulate the evolution of the clustering. The algorithm identifies the smallest of the distances and if it is a d_{ij} then combines entities i and j, while if it is d_{iB} defines entity i as a jet and removes it from the list of entities. The new four-moment (p^μ) of the combined pair is defined as $p^\mu_{i+j} = p^\mu_i + p^\mu_j$. Then, distances for the set of remaining entities are calculated and the procedure is repeated until no entities are left.

Two examples of sequential recombination algorithms are the k_t algorithm ($p = 1$) [121], extensively used at LEP, and the Cambridge-Aachen (C/A) algorithm ($p = 0$) [122, 123]. In the k_t algorithm the soft particles are merged first, which introduces a strong sensitivity to small fluctuations of the energy density of the parton shower. The C/A algorithm relies solely on angular ordering of emissions in order to reconstruct the shower by omitting the transverse momentum from d_{ij}, which again results in a strong dependence to the experimental conditions. A case of particular interest corresponds to $p = -1$, referred to as the "anti-k_t" jet-clustering algorithm [125]. In this, the soft particles will tend to cluster with hard ones long before they cluster among themselves, effectively reducing the sensitivity to fluctuations of the parton shower or energy contributions from UE. If a hard particle has no hard neighbours within a distance $2R$, then it will simply accumulate all the soft particles within a circle of radius R, resulting in a perfectly circular cone-shape jet in the $(y-\phi)$ plane. More complex boundaries can occur for busy environments composed of many close-by showers.

The other free parameter in Equation 2.6 is the radius R that sets a characteristic scale in the $(y-\phi)$ space, and generally speaking defines the area of the jet (πR^2) for highly energetic and isolated jets. There is no strict recommendation for the choice of R, but it is commonly advised to use values between 0.4 and 1.2, where effects of hadronisation and the influence of the underlying event are minimised [126] (see also Figure 38 of Ref. [127]). As the radius increases less energy is "splashed out" of the jet, a phenomena known as out-of-cone energy, but the influence of underlying event also increases as the area of the jet growths.

The most common values for anti-k_t jets in ATLAS are $R = 0.4, 1.0$ [128–130]. The so-called large-R jets ($R = 1.0$) are used (among other things) to identify boosted hadronically decaying vector bosons. Jets from two-body decays of boosted vector bosons have a substructure typically absent from decays of gluons and light quarks [129]. To elucidate the substructure differences, grooming techniques [131–133] remove soft contributions to the jet. In particular, ATLAS commonly uses a grooming technique called trimmig [129, 133]. Jet trimming further aims to improve the jet mass resolution by removing contamination from initial state radiation, multiple parton interactions, and pile-up interactions.

2.3.5.2 Topological clusters

A possible choice for inputs to the jet reconstruction algorithm is the set of high-granularity calorimeter cells. The cell signals are measured at the electromagnetic (EM) energy scale. This scale reconstructs the energy deposited by electrons and photons correctly but does not include any corrections for the loss of signal for hadrons due to the non-compensating nature of the ATLAS calorimeter. Another choice for inputs to the jet algorithm is the set of towers of cells running longitudinally along the calorimeter (like the sliding window used for electron reconstruction, 2.3.2). However, none of those inputs reproduce the actual shape of the particle cascade and are more susceptible to electronic noise and other sources of fluctuations such as pile-up. For that reason, the input objects to the jet algorithm are three-dimensional topological clusters (topoclusters) [134] built from individual calorimeter cells.

The topocluster formation follows cell signal-significance patterns generated by electromagnetic and hadronic showers. In this, the clustering algorithm implicitly performs a topological noise suppression by removing cells with insignificant signals which are not in close proximity to cells with significant signals. This approach takes advantage of the high granularity of the calorimeter system. Finely segmented lateral readout together with longitudinal sampling layers allows the resolution of energy-flow structures generating these spatial signal patterns, thus retaining only signals important for particle and jet reconstruction while efficiently removing insignificant signals induced by noise. As the calorimeters are subject to a significant amount of background noise, the mean noise level is constantly measured and subtracted from measurements.

Topological clusters are formed through an iterative procedure [134] which identifies the most significant energy deposits, the seed cells, and then clusters neighboring cells into a single topocluster. Seed cells are first identified as the calorimeter cells with an energy significantly above a predefined noise threshold $|E_{\text{cell}}| > 4\sigma_{\text{noise}}$. The seed cell forms a protocluster and neighbouring cells are iteratively added to it if they have an energy $|E_{\text{cell}}| > 2\sigma_{\text{noise}}$. Once the iterative process ends and a stable protocluster is formed, all cells adjacent to the protocluster are added, independent of the magnitude of their signal. Through this method a topocluster is formed by a core of cells with significant energy surrounded by an envelope of cells containing any residual or leaked energy.

The jet-finding algorithms treat the topoclusters as massless four-vectors of magnitude $E = \sum E_{\text{cell}}$ which point from the centre of the detector to the energy-weighted barycentre of the topocluster. The resulting topological cell clusters have shape and location information, which can be exploited to improve the energy resolution. A local cell signal weighting (LCW) [134, 135] method can be applied to each topocluster, bringing clusters from the EM to the LCW scale, including corrections for dead material, out-of-cluster losses for pions and calorimeter response for hadronic clusters (identified using their topology and energy density).

2.3.5.3 Particle flow algorithm

An alternative approach is taken in the particle flow algorithm, in which measurements from both the tracker and the calorimeter are combined to form the signals, which ideally represent individual particles. The energy deposited in the calorimeter by all the charged particles is removed. Jet reconstruction is then performed on an ensemble of "particle flow objects" consisting of the remaining calorimeter energy and tracks which are matched to the hard interaction. The benefits of reconstructing jets with the particle flow algorithm (detailed in Ref. [136]) are:

- For low-energy charged particles, the momentum resolution of the tracker is significantly better than the energy resolution of the calorimeter.

- The acceptance of the detector is extended to softer particles, as tracks are reconstructed for charged particles with a minimum transverse momentum $p_{\mathrm{T}} > 400\,\mathrm{MeV}$, whose energy deposits often do not pass the noise thresholds required to seed topo-clusters [136].

- The angular resolution of a single charged particle, reconstructed using the tracker is much better than that of the calorimeter.

- Low-p_{T} charged particles originating within a hadronic jet are swept out of the jet cone by the magnetic field by the time they reach the calorimeter. By using the tracks azimuthal coordinate at the perigee, these particles are clustered into the jet

- When a track is reconstructed, one can ascertain whether it is associated with a vertex. Therefore, in the presence of multiple in-time pile-up interactions, the effect of additional particles on the hard-scatter interaction signal can be mitigated by rejecting signals originating from pile-up vertices.

However, particle flow introduces a complication. For any particle whose track measurement ought to be used, it is necessary to correctly identify and subtract its signal in the calorimeter to avoid double-counting. In the particle flow algorithm, a Boolean decision is made as to whether to use the tracker or calorimeter measurement. The ability to accurately subtract all of a single particle energy, without removing energy deposited by other particles, forms the key performance criterion upon which the algorithm is optimised.

2.3.5.4 Jets considered in this thesis

From now on, jets are reconstructed using the anti-k_t algorithm with a radius parameter value of $R = 0.4$ using the FastJet software package [137]. The inputs to the jet algorithm are either stable simulated "truth" particles or energy deposits in the calorimeter (particle flow objects for p-flow jets), and the resulting jets are referred to as "truth jets" or "particle jets" and "reconstructed jets", respectively. The inputs to truth jets are simulated particles with a lifetime τ, defined by $c\tau > 10\,\mathrm{mm}$, that are (unless stated otherwise) neither final-state muons nor neutrinos. Muons and neutrinos are ignored as they leave little or no visible energy in the calorimeters.

For the physics analyses discussed in Chapters 4 and 5, jets reconstructed from EM topoclusters (AntiKt4EMTopo) are used. In addition, jets reconstructed from LCW topoclusters (AntiKt4LCTopo) and with the particle flow (p-flow) algorithm (AntiKt4EMPFlow) are used for the derivation of the global sequential calibration presented in Chapter 3.

Chapter 3

Global sequential calibration

3.1 Preface

The determination of the jet energy scale, its uncertainty, and the achievement of an optimal jet energy resolution are the major tasks of the ATLAS jet calibration program. Jets in data and simulation are calibrated following the procedure described in detail in Ref. [130]. The global sequential (GS) calibration [138] is part of the jet energy scale calibration [130]. This chapter details the derivation of the GS calibration to be used in 13 TeV collision and simulated data. The GS calibration was derived for AntiKt4EMTopo, AntiKt4LCTopo and AntiKt4EMPFlow jets, the latter for the first time in the ATLAS collaboration. This calibration was used by all analyses using jets in 2015 and 2016 data.

Section 3.2 summarises the jet calibration scheme used in ATLAS in Run 2. The MC simulation sample and the object selection used to derive the GS calibration are described in Sections 3.3 and 3.4, respectively. The GS calibration technique is introduced in Section 3.5 and the results are presented in Section 3.6.

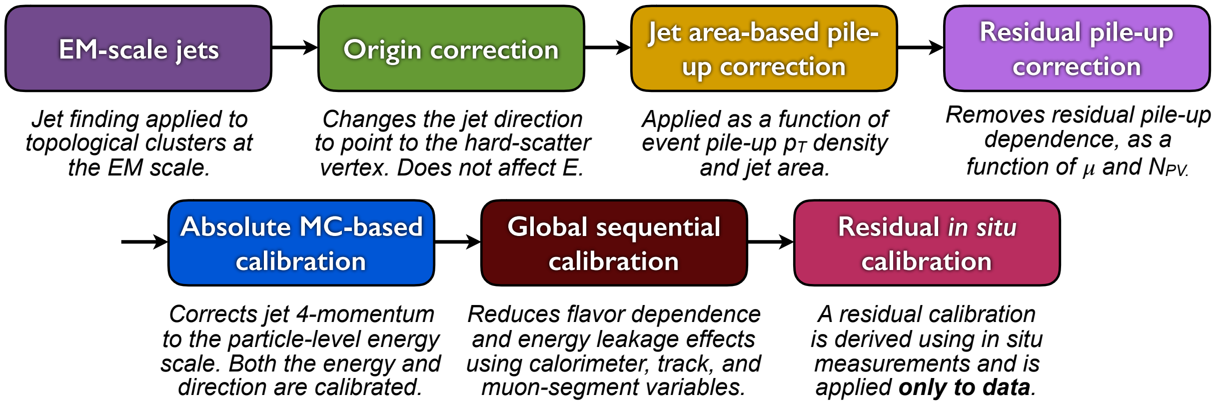
3.2 Introduction to the jet calibrations

The jet energy calibration restores the jet energy to that of jets reconstructed at the particle level. The full chain of corrections is illustrated in Figure 3.1. All stages correct the full four-momentum unless otherwise stated, scaling the jet p_{T}, energy, and mass. Each stage is briefly discussed next.

3.2.1 Origin correction

Calorimeter jets are reconstructed assuming they point back to the centre of the detector, which usually does not match to the hard-scatter vertex. In the origin correction [139], the four-momenta of the jets are recalculated to originate from the hard-scatter vertex rather than from the centre of the detector, while keeping the jet energy constant, and consenquently improving the η resolution of jets. This correction is not applied to p-flow jets, since the origin correction is already applied to the particle flow objects, given that track information is already available when reconstructing them.

Figure 3.1: Illustration of the calibration stages. Each one is applied to the four-momentum of the jet, except for the origin correction. Image from Ref. [130]



3.2.2 Pile-up correction

Jets are corrected for the contributions from additional proton-proton interactions within the same (in-time) or nearby (out-of-time) bunch crossings. First, a correction based on the jet area (A) and the average transverse energy density (ρ) of the event is applied [140], $\Delta p_{\mathrm{T}} = -\rho \times A$. The jet area is a measure of the susceptibility of the jet to pile-up and is determined jet by jet. The average energy density serves as a measure of the pile-up activity and is calculated event by event as the median energy density of jets reconstructed with the k_t algorithm with a radius parameter value of $R = 0.4$. After this correction, there remains some dependence of the average jet p_{T} on pile-up activity. An additional correction is therefore derived by comparing reconstructed jets to matched particle jets in simulated inclusive jet events [141]. The correction is parametrised as a function of the mean number of interactions per bunch crossing (μ) and the number of reconstructed primary vertices in the event (N_{PV}), such that both the out-of-time and in-time effects are taken into account.

3.2.3 Absolute Monte Carlo-based calibration

A jet energy- and η-dependent correction, known as JES, is applied to reconstructed jets in data and Monte Carlo simulation [139]. It is derived from simulated samples of inclusive jet events and is designed to lead to agreement in energy and direction between MC reconstructed and particle jets on average, correcting for instrumental effects, non-compensating calorimeter response, energy losses in dead material and out-of-cone effects.

3.2.4 Global sequential calibration

The Monte Carlo based procedure described in the previous section corrects the calibration of jets depending on their momentum and rapidity. However, for a given p_{T}–η bin, the substructure and thus the response can vary from jet to jet depending on the softness and the nature of the particles it is composed of, its transverse distribution, and the fluctuations of its development in the calorimeter [130, 138]. Furthermore, the average particle composition and shower shape of a jet varies for quark- and gluon-initiated jets.

A quark-initiated jet will often include hadrons with a higher fraction of the jet p_{T} that penetrate further into the calorimeter, while a gluon-initiated jet will typically contain more particles of softer p_{T}, leading to a lower calorimeter response and a wider transverse profile. The GS calibration is a series of multiplicative corrections to reduce the effects from these fluctuations and improve the jet resolution. It is based on global jet observables such as the longitudinal structure of the energy depositions within the calorimeters, tracking information associated to the jet, and information related to the activity in the muon chambers behind a jet. The GS corrections are designed to improve performance without changing the average jet energy scale in the QCD inclusive jet sample used as a reference, and are discussed in detail throughout this chapter.

3.2.5 In situ techniques

At this point, the calibration of MC jets is finished. However, a residual "in situ" calibration is derived to correct for remaining differences between the jet energy response in data and simulation.

This correction is calculated using γ+jet, Z+jet, dijet and multijet p_{T}-balance techniques [130, 142, 143]. Up to a jet p_{T} of about 950 GeV(400 GeV), the p_{T} balance between a photon (Z boson) and a jet is exploited. The multijet p_{T}-balance technique calibrates a high-p_{T} jet ($300 < p_{\mathrm{T}} < 2000$ GeV) recoiling against a collection of lower-p_{T} jets. Beyond 2000 GeV the response is considered constant. All these corrections are derived for the central jets, with $|\eta| < 1.2$ and are combined following the procedure described in Refs. [130, 144]. The relative response of all detector regions is equalised using a p_{T}-balance method exploiting dijet events (η-intercalibration) where the two leading jets are in different η-regions.

3.3 Monte Carlo simulation samples

MC event generators are used to simulate the energy and direction of particles produced in pp collisions. An overview of the MC event generators used in ATLAS can be found in Ref. [145]. The baseline simulation samples used for the GS corrections are inclusive jet events produced using Pythia 8 (version 8.160) [146] with the NNPDF2.3LO [147] LO PDF set and the A14 underlying-event tune. The EvtGen 1.2.0 program [148] was used to model bottom and charm hadron decays.

Pile-up events, i.e additional pp interactions that are not correlated with the hard-scatter event of interest, are simulated as minimum bias events produced with Pythia 8 using the A2 tune [64] and the MSTW2008LO [149] PDF set. The simulated detector signals from these events are overlaid on the detector signal from the hard-scatter event, in order to reproduce the pile-up conditions of the different periods of data taking in 2015 and 2016.

Generated events are propagated through a full simulation of the ATLAS detector based on Geant 4 that simulates the interactions of the particles produced by the event generators with the detector material. Hadronic showers are simulated with the FTFP [150] and BERT [151] models.

3.4 Object selection

3.4.1 Jets

The inputs to reconstructed jets are topoclusters or particle flow objects, discussed in Sections 2.3.5.2 and 2.3.5.3, respectively. The topoclusters are initially reconstructed at the EM scale, which correctly measures the energy deposited in the calorimeter by particles produced in EM showers. A second topocluster collection is built by calibrating the calorimeter cells such that the response of the calorimeters to both electromagnetic and hadronic showers is correctly accounted for. This calibration uses the local cell signal weighting (LCW) method, that improves the resolution compared to the EM scale by correcting the cluster energy according to its degree of compatibility with an EM or hadronic depositions, and thus reduces fluctuations due to the non-compensating response of the ATLAS calorimeters [135].

Reconstructed jets are calibrated using the multi-step procedure outlined in Section 3.2, up to the JES calibration. After this, the energy scale of the calorimeter jets built from EM or LCW-scale topoclusters are referred to as EM+JES or LCW+JES, respectively. On the other hand, the energy scale of the p-flow jets after the JES calibration is labeled as EMPFlow+JES.

3.4.2 Tracks

Tracks are required to have transverse momentum of at least 500 MeV, in addition to quality criteria based on impact parameters and number of hits in the different ID subdetectors [152].

3.4.3 Association of tracks and truth jets to reconstructed jets

It is useful to associate particles, tracks, and truth jets to reconstructed jets. There are two standard methods of association that are used for different purposes: ΔR matching and ghost association [140].

The method known as ΔR matching associates reconstructed jets to a truth (particle) jet when both are closer to each other than to any other jet and lie within a radius of $R = 0.3$.

The ghost association procedure is used to associate hadrons, partons or tracks to reconstructed jets. It treats the objects as four-vectors of infinitesimal magnitude during the jet reconstruction and assigns them to the jet with which they are clustered.

For composition studies, the flavour of a reconstructed jet in simulation is the PDG ID[1] [106] of the ghost-associated parton with highest energy. Only the two jets with highest p_{T} are used in the flavour studies presented in Section 3.6.3.

3.4.4 Muon segments

Muon track segments are used in the jet calibration as a proxy for the uncaptured jet energy carried by energetic particles leaking past the calorimeters without being fully absorbed. The segments are partial tracks reconstructed within an MS module. Segments

[1]The PDG ID is a numbering convention of known and theorised particles to facilitate information exchange between different Monte Carlo event generators and detector simulations.

are assigned to jets using the method of ghost association described above, with each segment treated as an input four-vector of infinitesimal magnitude to the jet reconstruction.

3.4.5 Jet structure properties

Several effects can impact the calorimeter response to jets:

- The number of particles that share the energy of the jet
- The fraction of the jet energy carried by electromagnetic particles (i.e., π^0, η, γ, $e^{\pm}$)
- The transverse distribution in $\eta - \phi$ of the particle jet constituents
- Fluctuations in the development of the calorimeter shower

The response of a jet of a given p_{T} is higher if the jet is composed of a few high p_{T} particles rather than many softer particles, because the response of a single charge hadron increases with particle p_{T}. This is due to the nature of the hadronic shower development. After each nuclear interaction, on average, 1/3 of the momentum is carried by π^0s (which have a response equal to one given that $\pi^0 \to \gamma\gamma$), and 2/3 by charged pions. Thus, after n steps in the shower development, $1 - \left(\frac{2}{3}\right)^n$ of the energy is transferred to π^0s. On average, the higher the particle p_{T}, the higher the value of n, the larger the EM component of the hadronic shower, and thus the larger the jet response.

Wider jets have on average lower response than narrow jets. This is due to three reasons. A wider shower can deposit a fraction of the jet energy beyond the radius of the jet, an effect known as out-of-cone radiation. Wider jets have on average lower energy density (energy per unit area), and could thus fail the threshold for cluster reconstruction. Finally, wider jets have larger probability to hit poorly instrumented transition regions.

Due to fluctuations inherent in the development of hadronic showers, the fraction of energy deposited in dead material can vary from jet to jet and, for high energy jets, a shower can deposit a fraction of its energy beyond the calorimeter, affecting consequently the jet response.

There are several jet structure observables that are sensitive to the shower attributes that impact the calorimeter response and help to implicity correct in average for those effects. For each jet definition, the variables that provide the largest improvement in performance have been selected in an empirical way [136, 138]. These observables are described next.

The longitudinal structure of the jet is characterised by the fractional energy deposited in the different longitudinal layers of the calorimeters before any jet calibration is applied ("layer fractions" or f_{layer}):

$$f_{\mathrm{layer}} = \frac{E_{\mathrm{EM}}^{\mathrm{layer}}}{E_{\mathrm{constit}}^{\mathrm{jet}}},$$

where $E_{\mathrm{constit}}^{\mathrm{jet}}$ is the energy of the jet before the jet energy scale calibration and $E_{\mathrm{EM}}^{\mathrm{layer}}$ is the energy deposited in the layer of interest at the EM scale. Two layer fractions are used, f_{Tile0} and f_{LAr3}. f_{Tile0} is the fraction of jet energy measured in the first layer of the hadronic Tile calorimeter ($|\eta_{\mathrm{det}}| < 1.7$), while f_{LAr3} is the energy fraction in the third layer of the electromagnetic LAr calorimeter ($|\eta_{\mathrm{det}}| < 3.5$). These fractions are sensitive to the energy deposited in the cryostat vessel and to the hadronic fraction of the shower.

Tracking information is included by using three observables, depending on the jet definition. These are the number of tracks (with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$) ghost-associated to the jets (n_{trk}), the jet track width ($width_{\mathrm{trk}}$), and the charged fraction (f_{charged}). The jet track width carries information about the transverse structure of the jet, and is defined as the average angular distance between the tracks (with $p_{\mathrm{T}} > 1\,\mathrm{GeV}$) ghost-associated to the jet and the calorimeter jet axis, weighted by the track p_{T}:

$$width_{\mathrm{trk}} = \frac{\sum_i p_{\mathrm{T}}^i \Delta R(i, \mathrm{jet})}{\sum_i p_{\mathrm{T}}^i},$$

where i refers to the tracks. The charged fraction, the fraction of the jet p_{T} measured from ghost-associated tracks (with $p_{\mathrm{T}} > 500\,\mathrm{MeV}$), allows to determine the degree of under-calibrated signal due to the lower energy deposit of hadrons in the hadronic calorimeter, and is defined as:

$$f_{\mathrm{charged}} = \frac{\sum_i p_{\mathrm{T}}^i}{p_{\mathrm{T}}^{\mathrm{jet,constit}}},$$

where i refers to the tracks and $p_{\mathrm{T}}^{\mathrm{jet,constit}}$ is the p_{T} of the jet before the jet energy scale calibration.

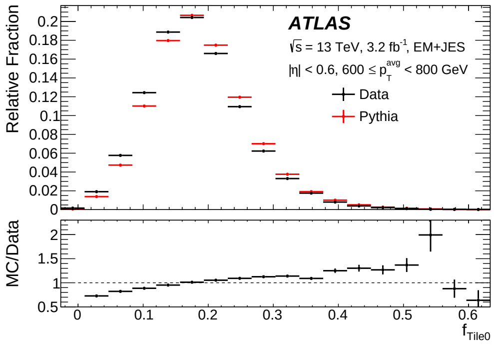
Activity in the muon chambers is characterised using the number of muon segments ghost-associated to the jet (N_{segments}). The N_{segments} correction, also known as punch-through correction, reduces the tails of the response distribution caused by high-p_{T} jets that are not fully contained in the calorimeter. All the corrections are derived as a function of jet p_{T}, except for the punch-through correction which is derived as a function of jet energy, being more correlated with the energy escaping the calorimeters. These corrections are performed such that the jet energy scale is unaltered on average. As a consequence, the jet energy resolution is improved and the sensitivity to jet fragmentation effects such as differences between quark- or gluon-induced jets is reduced.

The GS corrections are determined in jet $|\eta_{\mathrm{det}}|$ (the pseudorapidity of the jet before the jet direction is corrected to point to the hard-scatter vertex) bins of width 0.1 from $|\eta_{\mathrm{det}}| = 0$ to $|\eta_{\mathrm{det}}| = 3.5$. Corrections are derived for reconstructed jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ at the EM+JES, LCW+JES and EMPFlow+JES scales. The variables used, as well as the order in which the corrections are applied, are summarised in Table 3.1. The improvement in performance does not depend on the sequence of the corrections [138].

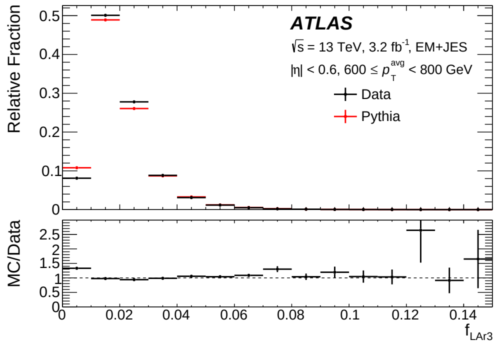
For jets at the LCW+JES scale, only the tracking and uncontained calorimeter jet corrections are applied, since the LCW calibration already takes into account information about the fractional energy deposited in each layer of the calorimeter, and the energy measured around the jet. No further improvement in resolution is achieved through the use of such variables for jets at the LCW+JES scale. For p-flow jets, no punch-through correction was derived due to a technical problem with the calculation of N_{segments}.

Figures 3.2 and 3.3 show the distributions of the f_{Tile0}, f_{LAr3}, n_{trk}, $width_{\mathrm{trk}}$, N_{segments}, and f_{charged} variables in data and MC events. The underlying distributions of these observables are fairly well modelled by MC simulation. Slight differences with data have a negligible impact on the GS corrections as long as the dependence of the average jet response on the observables is well modelled in MC simulation. A complete study, using data and MC simulations, of the average jet response dependence on the GS observables is presented in Refs. [136, 153], confirming the good modelling of the MC simulation.

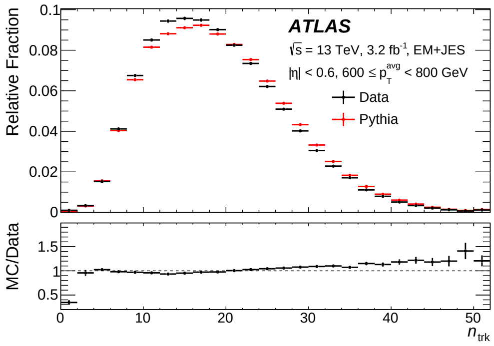
Figure 3.2: The distributions of observables used by the GS calibration in inclusive jet events for anti-k_t $R = 0.4$ jets within $|\eta| < 0.6$ in data (black) and Pythia MC simulation (red). The variables shown are (a) the fractional energy in the first Tile calorimeter layer, (b) the fractional energy in the third LAr calorimeter layer, (c) the number of tracks per jet, (d) the p_{T}-weighted track width, and (e) the number of muon track segments per jet. Jets are selected only from events in which $600 < p_{\mathrm{T}}^{\mathrm{avg}} < 800$ GeV, where $p_{\mathrm{T}}^{\mathrm{avg}}$ is the average p_{T} of the two leading p_{T} jets. Jets are calibrated up to the EM+JES scheme without the GS calibrations applied. Distributions are individually normalised to unit area. The ratio of Pythia to data is shown in the bottom panels.



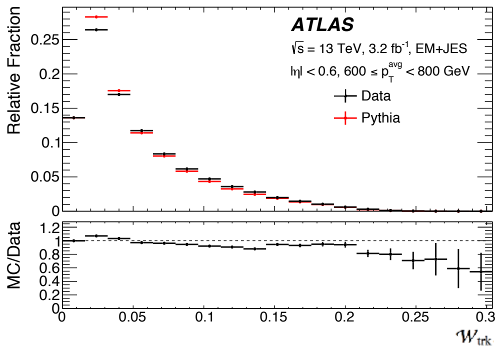
(a)



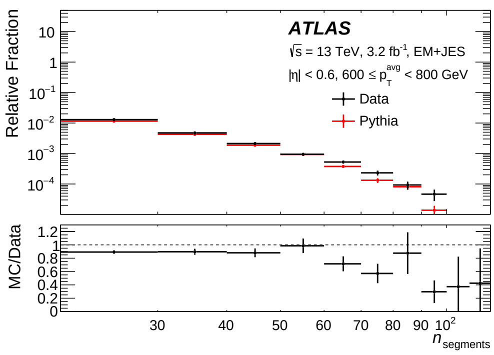
(b)



(c)



(d)



(e)

Table 3.1: Sequence of global sequential corrections used to improve the jet performance in each $|\eta_{\mathrm{det}}|$ region for AntiKt4EMTopo, AntiKt4LCTopo, and AntiKt4EMPFlow jets.

AntiKt4EMTopo

$|\eta_{\mathrm{det}}|$ region	Corrections				
$[0, 1.7]$	f_{Tile0}	f_{LAr3}	n_{trk}	$width_{\mathrm{trk}}$	N_{segments}
$[1.7, 2.5]$		f_{LAr3}	n_{trk}	$width_{\mathrm{trk}}$	N_{segments}
$[2.5, 2.7]$		f_{LAr3}			N_{segments}
$[2.7, 3.5]$		f_{LAr3}			

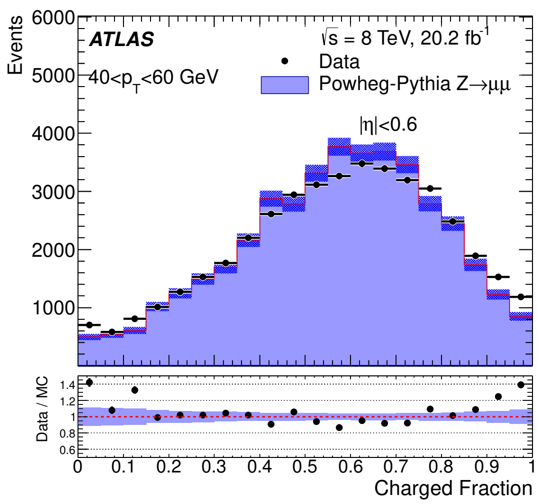
AntiKt4LCTopo

$|\eta_{\mathrm{det}}|$ region	Corrections		
$[0, 2.5]$	n_{trk}	$width_{\mathrm{trk}}$	N_{segments}
$[2.5, 3.5]$			N_{segments}

AntiKt4EMPFlow

$|\eta_{\mathrm{det}}|$ region	Corrections		
$[0, 1.7]$	f_{charged}	f_{Tile0}	f_{LAr3}
$[1.7, 2.7]$	f_{charged}		f_{LAr3}
$[2.7, 3.5]$			f_{LAr3}

Figure 3.3: Comparison of the charged fraction (fractional jet p_{T} carried by reconstructed tracks) for a selection of p-flow jets with $40 < p_{\mathrm{T}} < 60\,\mathrm{GeV}$ and $|\eta| < 0.6$, selected in $Z \to \mu\mu$ events from collision data and MC simulation. The simulated samples are normalised to the number of events in data. Jets are calibrated up to the EMPFlow+JES scale without the GS calibrations applied. The ratio of data to the MC simulation is shown in the bottom panel. Image from Ref. [136].



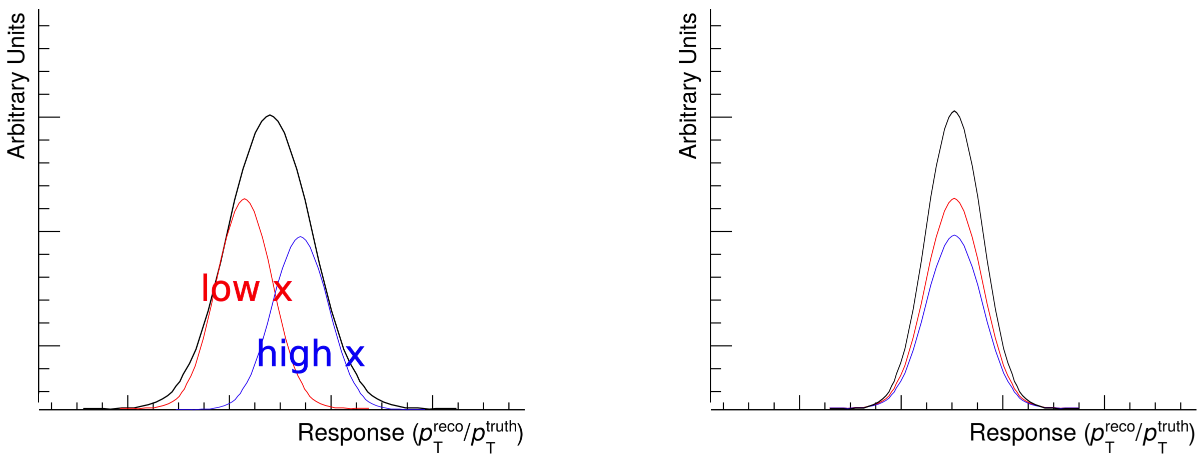
3.5 Global sequential calibration technique

The average calorimeter response to jets, or average jet response, for a given population of reconstructed jets is defined as $\mathcal{R} = \langle p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{truth}} \rangle$, where $p_{\mathrm{T}}^{\mathrm{reco}}$ is the reconstructed jet p_{T} and $p_{\mathrm{T}}^{\mathrm{truth}}$ is the p_{T} of the matching (within $\Delta R < 0.3$) truth jet. The mean is taken from a Gaussian fit with a standard range of $1.6\,\sigma$ to the $p_{\mathrm{T}}^{\mathrm{reco}}/p_{\mathrm{T}}^{\mathrm{truth}}$ distribution. In the case of the correction for uncontained jets, the jet energy is used instead of the p_{T}. The jet resolution ($\sigma_{\mathcal{R}}$) is given by the standard deviation of the Gaussian fit to the jet response distribution. The fractional jet resolution $\left(\sigma_{\mathcal{R}}/\mathcal{R} = \sigma_{p_{\mathrm{T}}^{\mathrm{reco}}}/p_{\mathrm{T}}^{\mathrm{reco}}\right)$ is used to determine the size of the fluctuations in the jet energy reconstruction. All jets used to determine the average jet response are required to be "isolated" in order to avoid ambiguity in the matching between reconstructed and truth jets. The isolation requirement for reconstructed jets is that there should be no other reconstructed jet within $\Delta R = 2.5R$, and only one truth jet with $p_{\mathrm{T}}^{\mathrm{truth}} > 7\,\mathrm{GeV}$ within a cone of $\Delta R = 1.5R$, where $R = 0.4$ for all jet definitions used in this analysis.

For each observable x, a jet four-momentum correction is derived in MC events as a function of x in bins of $p_{\mathrm{T}}^{\mathrm{truth}}$ and $|\eta_{\mathrm{det}}|$ by inverting the reconstructed average jet response: $C(x) = \mathcal{R}^{-1}(x)$. The punch-through correction is an exception, since it is constructed as a function of the jet energy (E^{truth}) and $\log N_{\mathrm{segments}}$. The punch-through correction is only derived for jets with $N_{\mathrm{segments}} > 20$.

The average jet response is then scaled back to the initial value, such that after this correction the dependence of the average jet response on the variable x is removed without changing the average jet energy. This results in a reduction of the spread of the reconstructed calorimeter jet energy and, thus, an improvement in resolution. Corrections for each observable are applied independently and sequentially to the jet four-momentum, neglecting correlations between observables. No improvement in resolution was found from including such correlations or altering the sequence of the corrections [138]. Figure 3.4 illustrate how the GS technique improves the resolution.

Figure 3.4: Illustration of how the technique described in the text improves the jet resolution. Response distributions before (left) and after (right) the GS calibration are presented. For fixed $p_{\mathrm{T}}^{\mathrm{truth}}$, the response is Gaussian but depends on x. A correction as a function of x is derived to shift the average response for fixed x to match the full (i.e. average jet response for the full x range) average jet response, thus improving the resolution.



First, the average jet response is evaluated in bins of $|\eta_{\mathrm{det}}|$, $p_{\mathrm{T}}^{\mathrm{truth}}$ (or E^{truth}) and jet property x, then the function F is constructed in $|\eta_{\mathrm{det}}|$ regions using a two dimensional

Gaussian kernel:

$$F\left(p_{\mathrm{T}}^{\mathrm{truth}}, x\right) = \frac{\sum_{j=1}^{N_{\mathrm{bins}}} R_j w_j}{\sum_{j=1}^{N_{\mathrm{bins}}} w_j},$$

where:

$$w_j = \frac{1}{\Delta R_j^2} \times \ \mathrm{Gauss}\left(\frac{\log p_{\mathrm{T}}^{\mathrm{truth}} - \log \left\langle p_{\mathrm{T}}^{\mathrm{truth}}\right\rangle_j}{\sigma_{p_{\mathrm{T}}}} \oplus \frac{x - \langle x\rangle_j}{\sigma_x}\right),$$

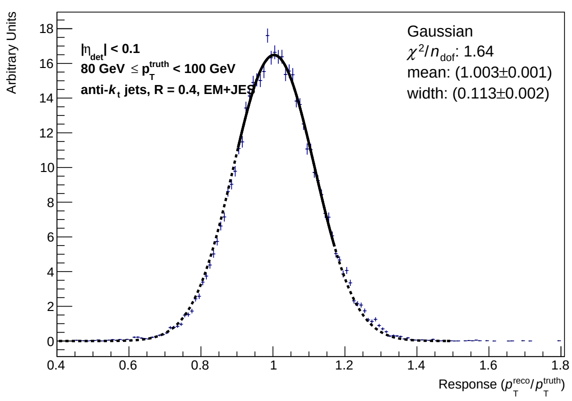
and j denotes the index of a $\left(p_{\mathrm{T}}^{\mathrm{truth}}, x\right)$-bin, R_j is the average jet response, ΔR_j is the statistical uncertainty of R_j, $\left\langle p_{\mathrm{T}}^{\mathrm{truth}}\right\rangle_j$ and $\langle x\rangle_j$ are the average p_{T} and x of the jets in each bin, Gauss(x) is the amplitude of a Gaussian function with $\mu = 0$ and $\sigma = 1$, $\sigma_{p_{\mathrm{T}}}$ and σ_x are width-parameters of the Gaussian kernel and $\oplus$ denotes addition in quadrature. This parametrisation is used to smooth out statistical fluctuations in the response curve. The kernel-width parameters used were selected to capture the shape of the average jet response across x and $p_{\mathrm{T}}^{\mathrm{truth}}$, and at the same time provide stability against statistical fluctuations. This procedure requires that the correction for variable $x_i\,(C(x_i))$ is calculated using jets to which the correction for variable $x_{i-1}\,(C(x_{i-1}))$ has already been applied. The jet transverse momentum after correction number i is given by:

$$p_{\mathrm{T}}^{i} = C(x_i) \times p_{\mathrm{T}}^{i-1} = C(x_i) \times C(x_{i-1}) \times p_{\mathrm{T}}^{i-2} = \ldots$$

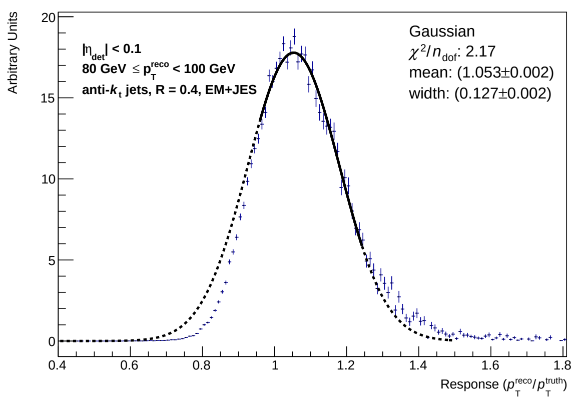
From now on, $F\left(p_{\mathrm{T}}^{\mathrm{truth}}, x\right)$ will be referred to as $\mathcal{R}(p_{\mathrm{T}}^{\mathrm{truth}})$ for simplicity.

The choice of deriving the corrections as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ instead of $p_{\mathrm{T}}^{\mathrm{reco}}$ is motivated by the fact that for fixed $p_{\mathrm{T}}^{\mathrm{truth}}$ ($p_{\mathrm{T}}^{\mathrm{reco}}$) bins the response distribution is (not) Gaussian. The shape of the distribution for bins of $p_{\mathrm{T}}^{\mathrm{reco}}$ is determined by the underlying falling p_{T} spectrum of the jet distribution. Since the cross-section is higher for low-$p_{\mathrm{T}}^{\mathrm{truth}}$ jets, for any fixed $p_{\mathrm{T}}^{\mathrm{reco}}$ bin, there will be more jets with low $p_{\mathrm{T}}^{\mathrm{truth}}$ (i.e.: high response). This causes a tail towards the high-response end of the response distribution, as illustrated in Figure 3.5.

Figure 3.5: Panels (a) and (b) show the response distribution for AntiKt4EMTopo jets at the EM+JES scale within $|\eta_{\mathrm{det}}| < 0.1$ with $80 < p_{\mathrm{T}}^{\mathrm{truth}} < 100$ GeV and $80 < p_{\mathrm{T}}^{\mathrm{reco}} < 100$ GeV, respectively. A Gaussian fit is overlaid in both distributions to illustrate that the response distribution is not Gaussian when plotted in $p_{\mathrm{T}}^{\mathrm{reco}}$ ranges.



(a)



(b)

This choice introduces a complication, the correction factors are derived as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ while the correction needs to be applied to reconstructed jets in MC and data, where in the latter no “truth” information is available. The “numerical inversion” technique solves this problem. In essence, it allows to apply p_{T}-dependent calibrations to reconstructed jets when the average response of these jets has a dependence on $p_{\mathrm{T}}^{\mathrm{truth}}$. There are two steps in this technique:

1. Calculation of $\mathcal{R}\left(p_{\mathrm{T}}^{\mathrm{truth}}\right)$ from a Gaussian fit to the response distributions in different $p_{\mathrm{T}}^{\mathrm{truth}}$ bins.
2. Estimation of $\mathcal{R}\left(p_{\mathrm{T}}^{\mathrm{reco}}\right)$ approximating $p_{\mathrm{T}}^{\mathrm{reco}}$ by $p_{\mathrm{T,est}}^{\mathrm{reco}} = \mathcal{R}\left(p_{\mathrm{T}}^{\mathrm{truth}}\right) \cdot p_{\mathrm{T}}^{\mathrm{truth}}$, expression which holds on average.

Note that $\mathcal{R}\left(p_{\mathrm{T}}^{\mathrm{reco}}\right)$ is, in fact, calculated as a function of $p_{\mathrm{T,est}}^{\mathrm{reco}}$ by transforming the x-axis of the response calculated in step 1. However, when it is used as part of the calibration it is evaluated on a jet-by-jet basis as a function of the jet $p_{\mathrm{T}}^{\mathrm{reco}}$. For this reason, in the rest of this chapter, $p_{\mathrm{T}}^{\mathrm{reco}}$ and $p_{\mathrm{T,est}}^{\mathrm{reco}}$ may be used interchangeably in this context, depending on whether the emphasis is on how the $\mathcal{R}$ function is calculated or how it is used. With an expression for $\mathcal{R}\left(p_{\mathrm{T}}^{\mathrm{reco}}\right)$, the jet can be now calibrated through the inversion of the response as $p_{\mathrm{T}}^{\mathrm{calib}} = p_{\mathrm{T}}^{\mathrm{reco}}/\mathcal{R}\left(p_{\mathrm{T}}^{\mathrm{reco}}\right)$. The jet four-momentum is calibrated in an equivalent way: using $1/\mathcal{R}\left(p_{\mathrm{T}}^{\mathrm{reco}}\right)$ as a scale factor for each of its components.

3.6 Performance of the global sequential calibration

In this section, the most relevant results regarding the calorimeter response to jets, the estimation of the jet resolution improvement, and the sensitivity to jet flavour are presented.

3.6.1 Jet response dependence on jet structure observables

The average jet p_{T} response in MC simulation as a function of each corresponding observable for several intervals of $p_{\mathrm{T}}^{\mathrm{truth}}$ is shown in Figure 3.6. All the variables are shown for AntiKt4EMTopo jets, with the exception of f_{charged} which is shown for AntiKt4EMPFlow jets. Jets are respectively calibrated at the EM+JES and the EMPFlow+JES scale without GS corrections. For all of the jet structure variables, a strong dependence of the average jet response as a function of the studied property is observed. The strongest dependence is observed for the n_{trk} and $width_{\mathrm{trk}}$ variables. All figures satisfy that the average jet response is equal to one. Consequently, the response is higher (smaller) than one for EM (hadronic) jets since they were overcorrected (undercorrected) by the JES. The observed dependence of the response on each observable is the expected:

- f_{charged}: The higher the charged fraction, the higher is the response, since the clusters are replaced by the p-flow objects with improved response. The lower the charged fraction, the smaller is the benefit from using tracks, and consequently the lower is the response.
- n_{trk}: The response decreases with the number of associated tracks, since the higher the number of particles sharing the jet energy, the smaller the response of the jet.
- $width_{\mathrm{trk}}$: Wider jets has higher $width_{\mathrm{trk}}$ and, as explained in Section 3.4.5, lower response with respect narrow jets.

- f_{Tile0}: The higher the f_{Tile0} fraction, the lower the EM component of the shower, and consequently the lower the response.

- f_{LAr3}: The higher the f_{LAr3} fraction, the higher the energy deposited in the non-instrumented area between the Tile and EM calorimeters, and hence the lower the response.

- N_{segments}: The higher the number of associated segments, the higher the energy uncontained in the calorimeter, and consequently, the lower the response.

The distribution of each observable in MC simulation for each $p_{\text{T}}^{\text{truth}}$ interval is shown in the bottom panels of those figures. The spike at zero in the f_{Tile0} distribution of Figure 3.6 at low $p_{\text{T}}^{\text{truth}}$ reflects jets that are fully contained in the electromagnetic calorimeter and do not deposit energy in the Tile calorimeter. The negative tail in the f_{LAr3} distribution (and, to a lesser extent, in the f_{Tile0} distribution) at low $p_{\text{T}}^{\text{truth}}$ reflects calorimeter noise fluctuations. AntiKt4LCTopo jets are not shown since they present a similar behaviour as AntiKt4EMTopo jets for n_{trk}, $width_{\text{trk}}$ and N_{segments} observables, while f_{Tile0}, f_{LAr3} are not not shown either for AntiKt4EMPFlow jets since the behaviour is analogous to the AntiKt4EMTopo case.

The dependence of the jet response on each observable is significantly reduced after the full GS calibration is applied, as shown in Figure 3.7. All the variables are shown for AntiKt4EMTopo jets, and only f_{charged} for AntiKt4EMPFlow jets. Deviations of the order of 2% (except for few cases at low p_{T}) are observed, reflecting the small correlations between observables that are unaccounted for in the corrections and small discrepancies between the actual response curve and the smoothed one. Similar results are observed for AntiKt4LCTopo and AntiKt4EMPFlow jets.

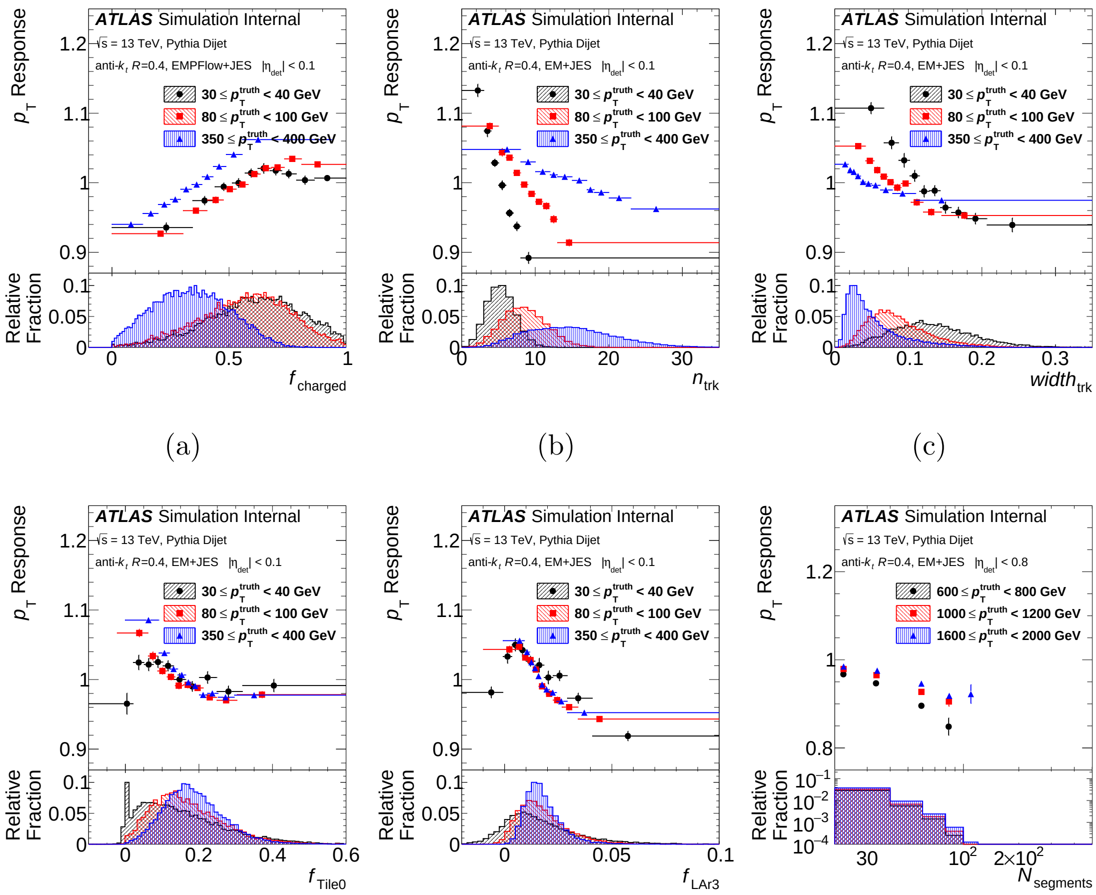
3.6.2 Jet response and resolution after the GS calibration

Figures 3.8 and 3.9 show the average jet response for AntiKt4EMTopo and AntiKt4EMPFlow jets, respectively, as a function of $p_{\text{T}}^{\text{truth}}$ for the Pythia 8 MC samples where the corrections were derived. The average jet response after each correction remains almost unchanged, as desired, except for few bins at low $p_{\text{T}}^{\text{truth}}$ where differences of $1-2\%$ are observed. Comparable results are obtained for AntiKt4LCTopo jets.

Figures 3.10 and 3.11 show the jet resolution for AntiKt4EMTopo and AntiKt4EMPFlow jets, respectively, as a function of $p_{\text{T}}^{\text{truth}}$ for the same MC samples. The jet resolution improves as more corrections are applied. The improvement observed for AntiKt4LCTopo jets, for which only the tracking and N_{segments}-based corrections are applied, was found to be smaller. This is expected since no calorimeter-based GS corrections are applied because this is already accounted for in the LCW calibration.

Figure 3.10 seems to indicate no improvement is achieved due to the punch-through correction, but note that all jets are used regarless of their number of associated segments, and this correction is only applied to jets with $N_{\text{segments}} > 20$. The fraction of jets with $N_{\text{segments}} > 20$ is negligible, except at very high E^{true}, as shown in Figure 3.12. Figure 3.13 shows the improvement in the RMS/Mean of the response as a function of the true jet energy for jets not fully contained in the calorimeters, i.e. jets with $N_{\text{segments}} > 20$, after applying the GS calibration. The mean corresponds to the mean value of the jet response distribution. The RMS was used instead of the resolution since it gives information about the reduction in the low jet response tail after applying the punch-through jet correction.

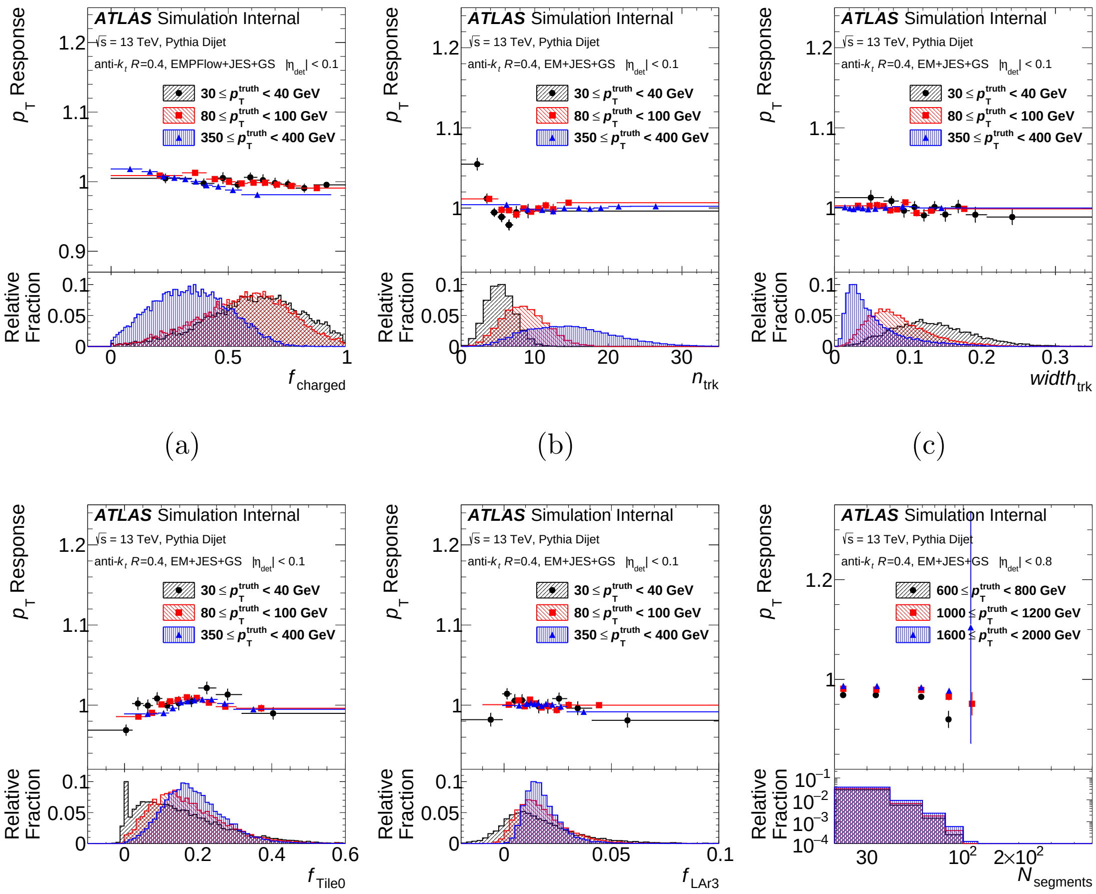
Figure 3.6: Average jet response as a function of (a) f_{charged}, (b) n_{trk}, (c) $width_{\text{trk}}$, (d) f_{Tile0}, (e) f_{LAr3}, and (f) N_{segments} variables for several $p_{\text{T}}^{\text{truth}}$ intervals for AntiKt4EMPFlow jets in panel (a) and for AntiKt4EMTopo jets in panels (b,c,d,e,f) with $|\eta_{\text{det}}| < 0.1$ ($|\eta_{\text{det}}| < 0.8$ for the N_{segments} variable). Markers are placed at the mean of the jet property distribution in that bin. Calorimeter (p-flow) jets are at the EM+JES (EMPFlow+JES) scale without GS corrections. The underlying distributions of the jet structure variables, normalised to their integral for each $p_{\text{T}}^{\text{truth}}$ bin, are shown in the lower part of each subfigure.



(a) (b) (c)

(d) (e) (f)

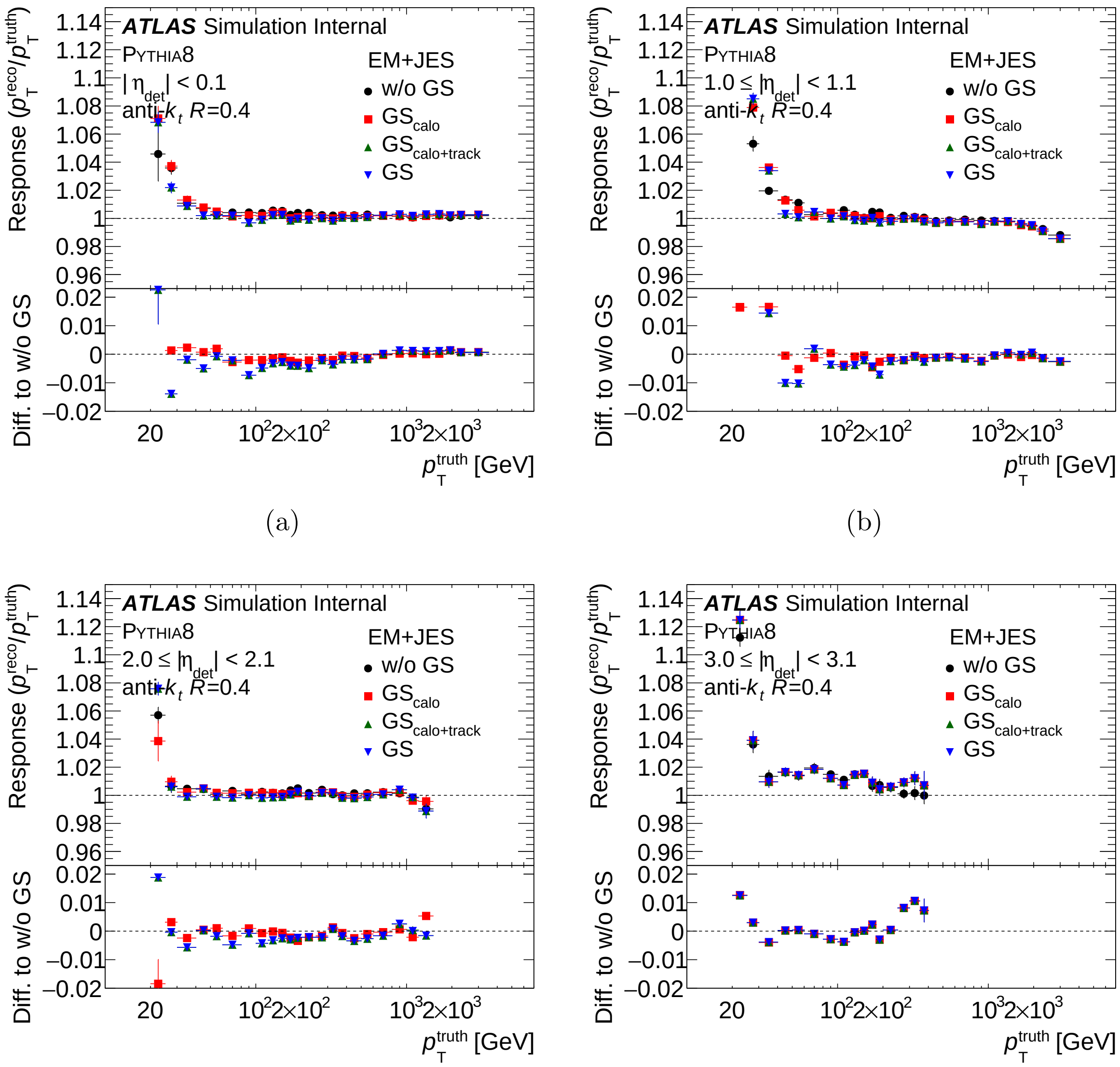
Figure 3.7: Similar to Figure 3.6, but after the GS corrections are applied.



(a) (b) (c)

(d) (e) (f)

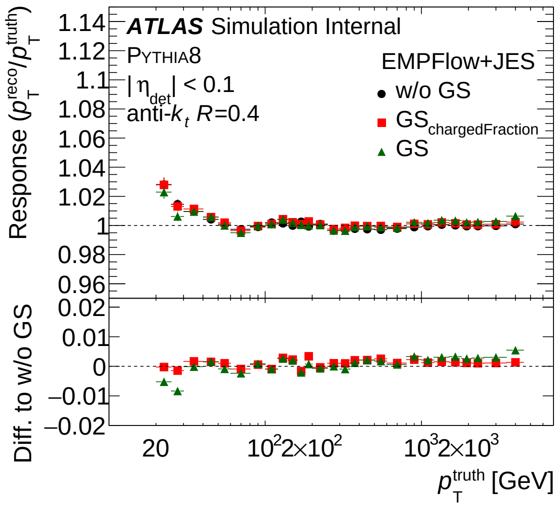
Figure 3.8: Average jet response as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ in MC simulation for AntiKt4EMTopo jets with (a) $|\eta_{\mathrm{det}}| < 0.1$, (b) $1.0 < |\eta_{\mathrm{det}}| < 1.1$, (c) $2.0 < |\eta_{\mathrm{det}}| < 2.1$, and (d) $3.0 < |\eta_{\mathrm{det}}| < 3.1$. Markers show the response at the EM+JES scale without GS corrections (black circles), with calorimeter-based GS corrections only (red squares), with calorimeter- and tracking-based GS corrections only (green upper triangles), and including all GS corrections (blue lower triangles). The lower part of each subfigure show the difference between the average jet response with the GS corrections and the average jet response at the EM+JES scale without the GS calibration.



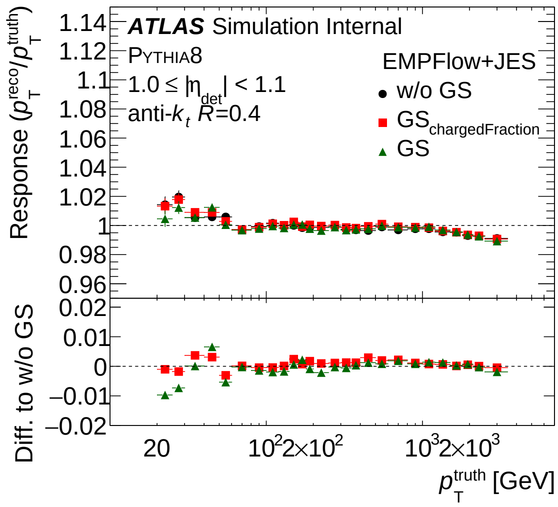
(a) (b) (c) (d)

The observed improvement decreases with increasing jet energy, as the fraction of energy not contained in the calorimeter for these jets becomes smaller relative to their total energy.

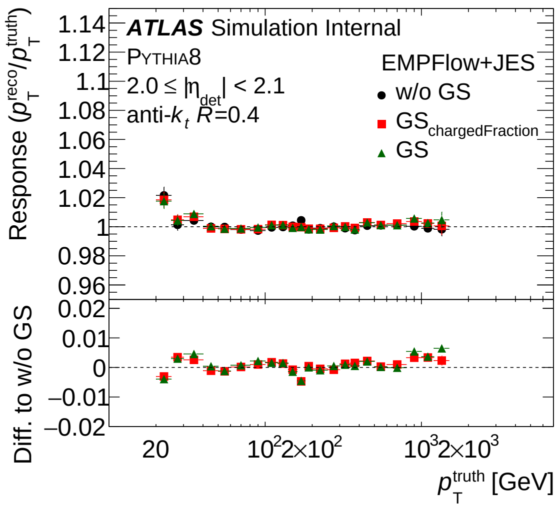
Figure 3.9: Average jet response as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ in MC simulation for AntiKt4EMPFlow jets with (a) $|\eta_{\mathrm{det}}| < 0.1$, (b) $1.0 < |\eta_{\mathrm{det}}| < 1.1$, (c) $2.0 < |\eta_{\mathrm{det}}| < 2.1$, and (d) $3.0 < |\eta_{\mathrm{det}}| < 3.1$. Markers show the response at the EMPFlow+JES scale without GS corrections (black circles), with charged fraction based GS correction only (red squares), and including all GS corrections (green upper triangles). The lower part of each subfigure show the difference between the average jet response with the GS corrections and the average jet response at the EMPFlow+JES scale without the GS calibration.



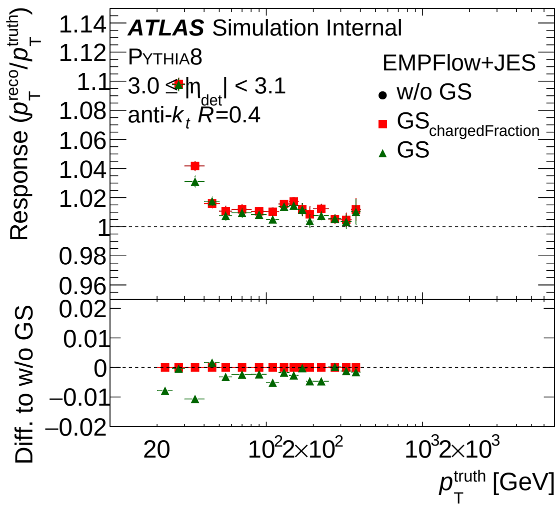
(a)



(b)



(c)



(d)

Figure 3.10: Jet transverse momentum resolution as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ in the Pythia 8 MC sample for (a) $|\eta_{\mathrm{det}}| < 0.1$, (b) $1.0 < |\eta_{\mathrm{det}}| < 1.1$, (c) $2.0 < |\eta_{\mathrm{det}}| < 2.1$, and (d) $3.0 < |\eta_{\mathrm{det}}| < 3.1$, for AntiKt4EMTopo jets. Markers show resolution at the EM+JES scale without GS corrections (black circles), with calorimeter-based GS corrections only (red squares), with calorimeter- and tracking-based GS corrections only (green upper triangles), and including all the GS corrections (blue lower triangles). The lower panel of each subfigure show the difference in quadrature between the resolution with the GS corrections (σ') and the resolution at the EM+JES scale without GS corrections (σ).

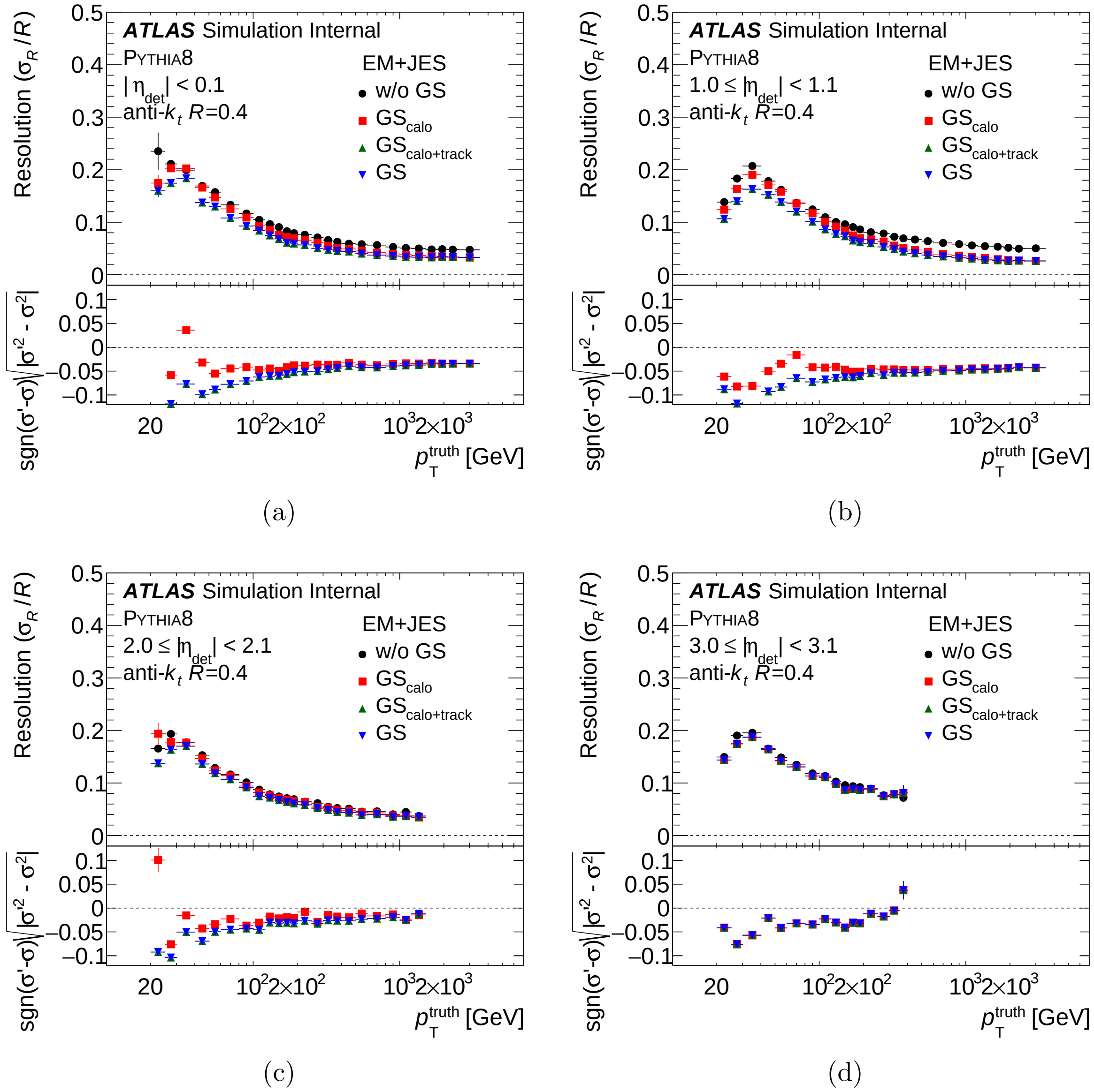
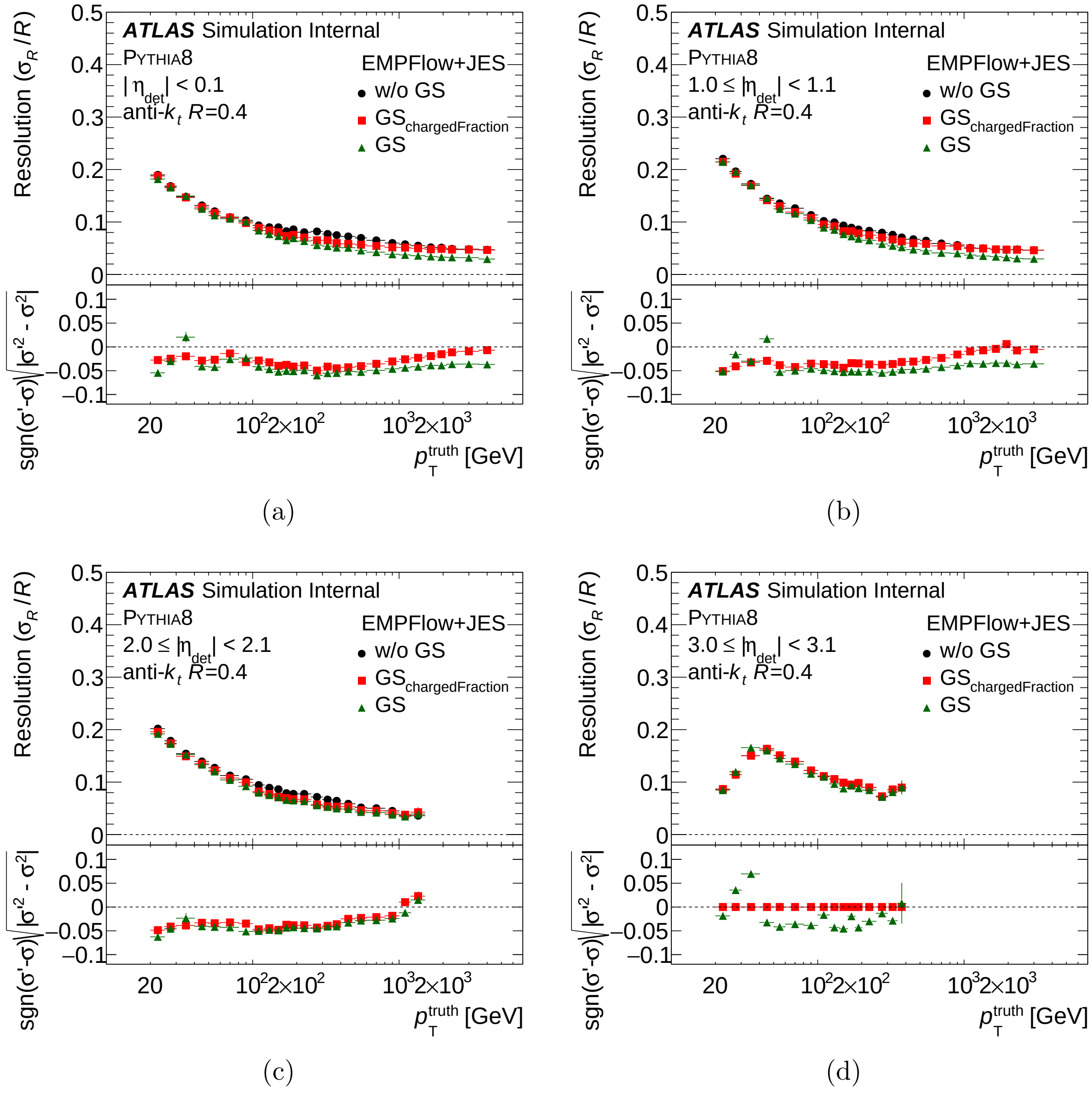


Figure 3.11: Jet transverse momentum resolution as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ in the Pythia 8 MC sample for (a) $|\eta_{\mathrm{det}}| < 0.1$, (b) $1.0 < |\eta_{\mathrm{det}}| < 1.1$, (c) $2.0 < |\eta_{\mathrm{det}}| < 2.1$, and (d) $3.0 < |\eta_{\mathrm{det}}| < 3.1$, for AntiKt4EMPFlow jets. Markers show resolution at the EMPFlow+JES scale without GS corrections (black circles), with f_{charged}-based GS correction only (red squares), and including all the GS corrections (green upper triangles). The lower panel of each subfigure show the difference in quadrature between the resolution with the GS corrections (σ') and the resolution at the EMPFlow+JES scale without GS corrections (σ).



(a) (b) (c) (d)

Figure 3.12: Fraction of jets with $N_{\text{segments}} > 20$ as a function of E^{true} for AntiKt4EMTopo jets with $|\eta_{\text{det}}| < 0.8$.

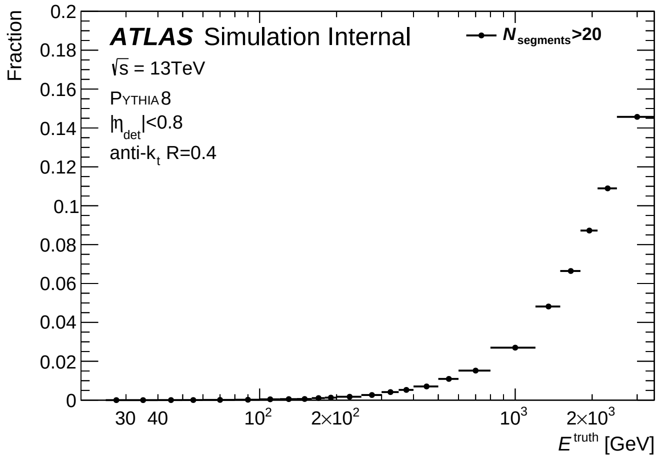
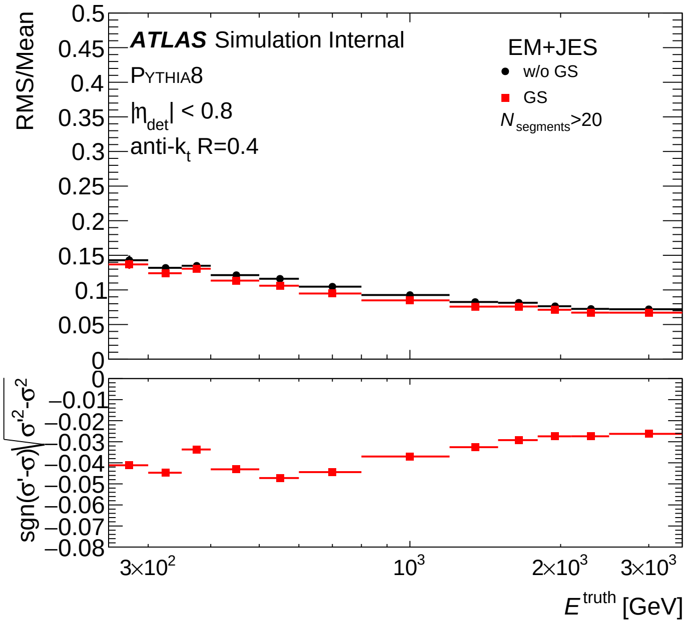


Figure 3.13: RMS/Mean of the response as a function of the truth jet energy for AntiKt4EMTopo jets with $N_{\text{segments}} > 20$ and $|\eta_{\text{det}}| < 0.8$, with (black circles) and without (red squares) the GS calibration. The lower panel show the difference in quadrature between the RMS/Mean with (σ') and without (σ) the GS calibration.



3.6.3 Flavour dependence of the jet response in simulation

The GS calibration preserves the average jet response in the specific Pythia 8 MC sample used to derive the corrections. If the GS calibration is applied to a sample with jets produced in different physics processes the average jet response can be different. This is because the jet response is sensitive to the flavour of a jet, and the flavour composition of a sample depends on the physics process. The jet structure variables and thus the calorimeter response to jets are sensitive to differences in fragmentation between light- (LQ-) and gluon-initiated jets produced in different physics processes. The difference in the response between jets initiated by quarks or gluons and a lack of knowledge of the flavour composition of the analysed data are a source of uncertainty in physics analyses involving jets. This is briefly discussed in Section 4.5.1 and in more detail in Ref. [130]. It is desirable for the flavour dependence to be as small as possible, because it limits the impact of these effects on the systematics uncertainty of the jet energy scale.

Studies of the flavour dependence of the calorimeter response to jets calibrated with the GS calibration are presented in this section. The flavour dependence is studied using jets from inclusive jet MC events processed through the flavour labelling procedure explained in Section 3.4.3. The flavour dependence of the jet response is in part a result of the differences in particle level properties of the different types of jets. A gluon-initiated jet tends to have more particles, and those particles tend to be softer than in the case of a LQ-initiated jet. Also, a gluon-initiated jet tends to be wider (i.e. with lower energy density in the core of the jet) before reaching the calorimeter. The magnetic field in the inner detector amplifies the broadness of the gluon-initiated jet, since their low p_{T} charged particles tend to bend more than the higher p_{T} particles in a LQ-initiated jet. The harder particles in a LQ-initiated jet tend to penetrate further into the calorimeter.

Figure 3.14 shows the jet flavour fractions as a function of p_{T}. At low p_{T} the sample is dominated by gluon-initiated jets. Between 300 GeV and 400 GeV, the contributions are of the same order. Above approximately 400 GeV, LQ-initiated jets increasingly dominate the sample. In this MC sample, the contribution of jets initiated by a quark b (b-jets) is negligible.

Figure 3.15 shows the average jet response for LQ-initiated jets in inclusive jet MC events in two η_{det} regions before and after applying the GS corrections. After the GS corrections, the average jet response of LQ-initiated jets gets closer to 1, mainly due to the tracking-based corrections. The impact on the average jet response of the uncontained calorimeter jet correction with respect to the previous ($width_{\mathrm{trk}}$) correction is negligible.

Figure 3.16 shows the average jet response for gluon-initiated jets in inclusive jet MC events in two η_{det} regions. Gluon-initiated jets have an average jet response that is higher than unity for $p_{\mathrm{T}}^{\mathrm{truth}} < 30$ GeV and smaller than unity in the complementary $p_{\mathrm{T}}^{\mathrm{truth}}$ range. This is due to an overcorrection from the JES at low p_{T} (the average response after the JES correction is higher that one), as seen in Figure 3.8. The GS calibration brings the response closer to unity, mainly due to the tracking-based corrections.

Figure 3.17 shows the average jet response for b-jets in inclusive jet MC events in two η_{det} regions. All b-jets, whether decaying semileptonically or hadronically, are included. The average jet response for b-jets at the EM+JES scale is larger than one, resembling the average jet response for LQ-initiated jets. The tracking corrections bring the average jet response closer to 1.

Figure 3.18 shows the difference in average jet response between LQ- and gluon-initiated jets with and without the GS calibration for AntiKt4EMTopo jets in two different η_{det} regions. The difference between the LQ- and gluon-initiated average jet response

Figure 3.14: Fraction of LQ-initiated jets (black squares), gluon-initiated jets (red circles), and b-jets (green upper triangles) as a function of $p_{\mathrm{T}}^{\mathrm{reco}}$ for AntiKt4EMTopo jets with $|\eta_{\mathrm{det}}| < 3.5$ and at the EM+JES with the GS corrections applied.

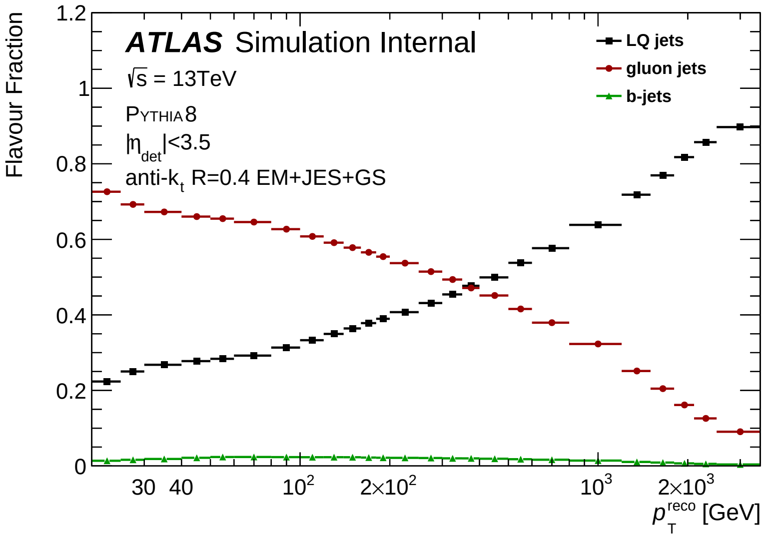
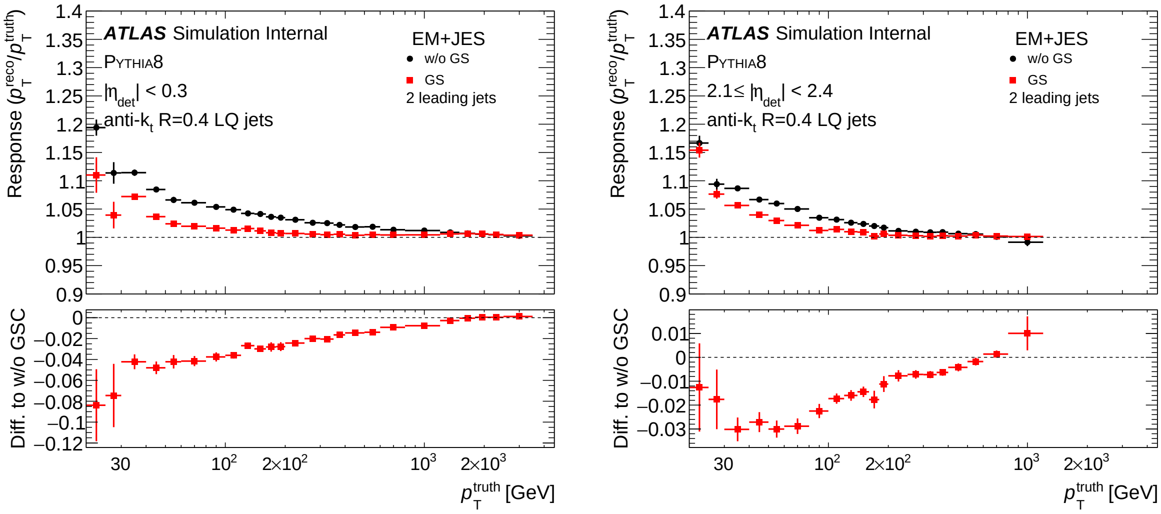


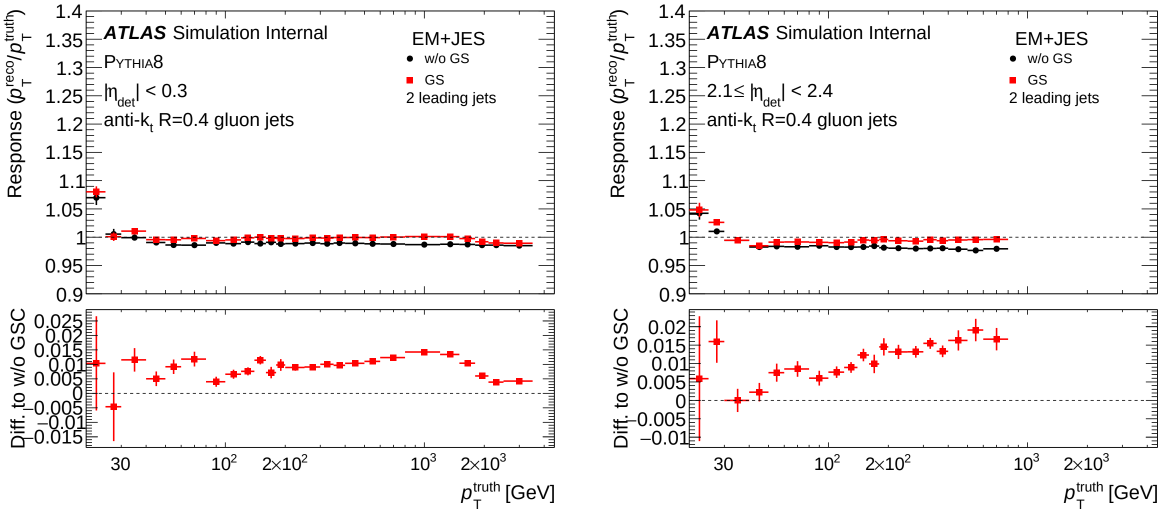
Figure 3.15: Average jet response as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ for LQ-initiated AntiKt4EMTopo jets for (a) $|\eta_{\mathrm{det}}| < 0.3$ and (b) $2.1 < |\eta_{\mathrm{det}}| < 2.4$ in the PYTHIA 8 MC sample. The average jet response at the EM+JES scale without GS corrections (black circles) and including all the GS corrections (red squares) are shown. Only the two highest p_{T} jets in each event are used. The lower part of each subfigure show the difference between the average jet response with the GS corrections and the average jet response at the EM+JES scale without GS corrections.



(a) (b)

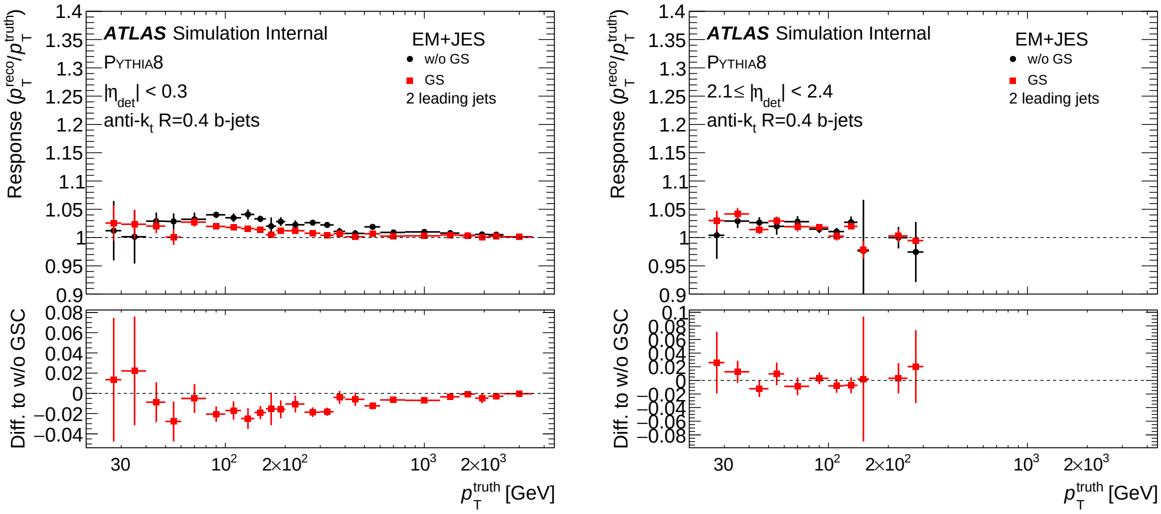
is considerably reduced after applying the GS corrections. A reduction of the flavour dependence on the average jet response is also obtained for AntiKt4LCTopo and AntiKt4EMPFlow jets.

Figure 3.16: Average jet response as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ for gluon-initiated AntiKt4EMTopo jets for (a) $|\eta_{\mathrm{det}}| < 0.3$ and (b) $2.1 < |\eta_{\mathrm{det}}| < 2.4$ in the Pythia 8 MC sample. The average jet response at the EM+JES scale without GS corrections (black circles) and including all the GS corrections (red squares) are shown. Only the two highest p_{T} jets in each event are used. The lower part of each subfigure show the difference between the average jet response with the GS corrections and the average jet response at the EM+JES scale without GS corrections.



(a) (b)

Figure 3.17: Average jet response as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ for AntiKt4EMTopo b-jets for (a) $|\eta_{\mathrm{det}}| < 0.3$ and (b) $2.1 < |\eta_{\mathrm{det}}| < 2.4$ in the Pythia 8 MC sample. The average jet response at the EM+JES scale without GS corrections (black circles) and including all the GS corrections (red squares) are shown. Only the two highest p_{T} jets in each event are used. The lower part of each subfigure show the difference between the average jet response with the GS corrections and the average jet response at the EM+JES scale without GS corrections.

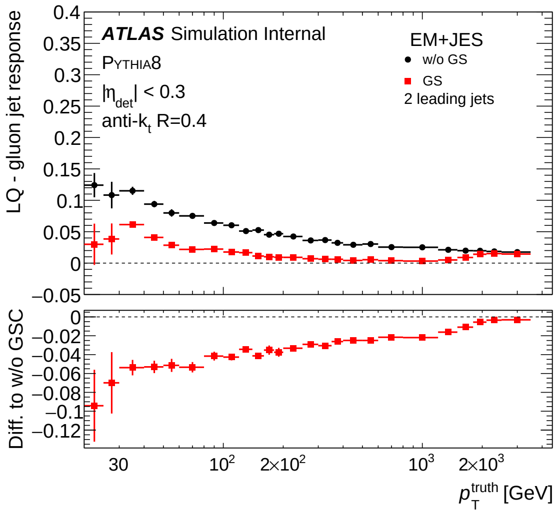


(a) (b)

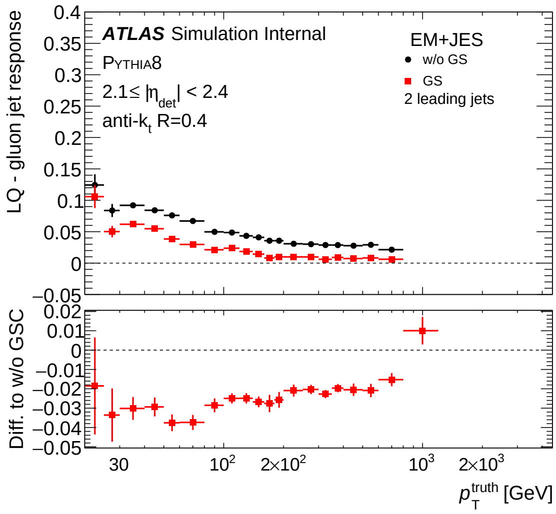
3.6.4 Global sequential calibration in ATLAS jet triggers

The GS calibration discussed in this chapter was also applied online in all the small-R ($R = 0.4$) jet HLT triggers in the data collected in 2017. Up to 2016, only the pile-up

Figure 3.18: Difference in the average response of LQ- and gluon-initiated jets as a function of $p_{\mathrm{T}}^{\mathrm{truth}}$ for AntiKt4EMTopo jets with (a) $|\eta_{\mathrm{det}}| < 0.3$ and (b) $2.1 < |\eta_{\mathrm{det}}| < 2.4$. The difference after at EM+JES without the GS corrections (black circles), and including all the GS corrections (red squares) are shown. Only the two highest p_{T} reconstructed jets are used. The lower part of each subfigure show the difference between the average LQ − gluon jet response with the GS corrections and the average LQ − gluon jet response at the EM+JES scale without the GS corrections.



(a)

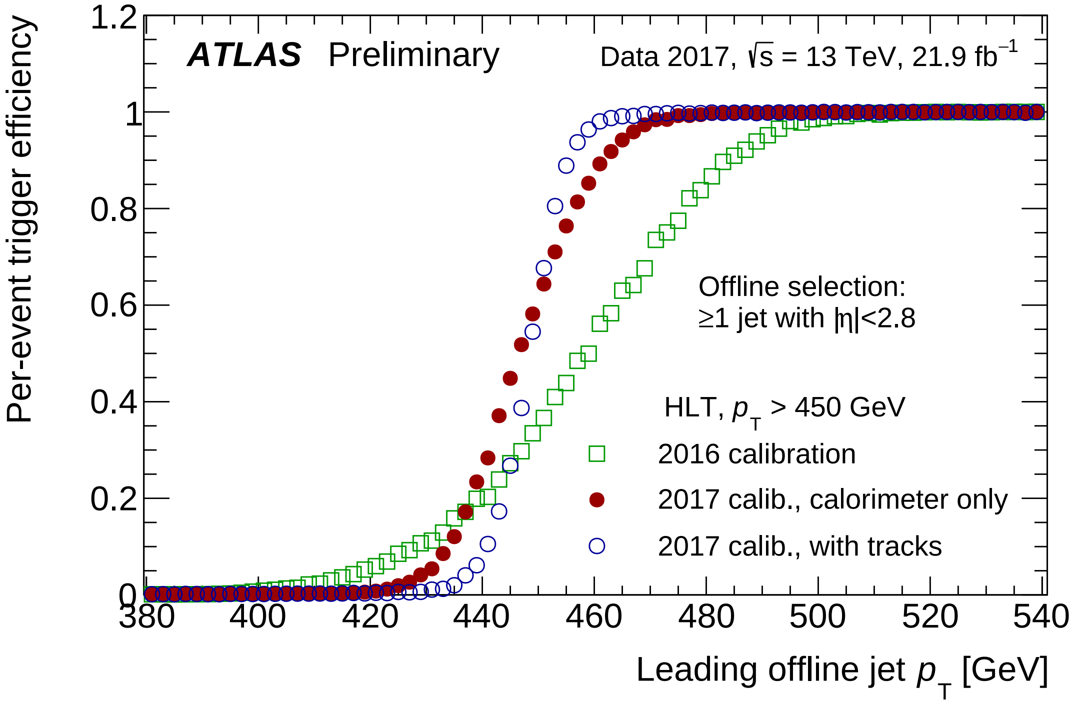


(b)

jet area and absolute MC-based calibrations had been applied online at the HLT level. In 2017, two different approaches were taken. In all the small-R jet HLT triggers, the calorimeter-based GS corrections and in situ calibrations were applied. For some small-R jet HLT triggers, the tracking-based GS calibration was also applied. The tracking-based GS corrections are only used in a small set of HLT triggers since a full tracking reconstruction is very CPU intensive at the HLT level. The inclusion of the GS calibration, together with the in situ calibrations, improves significantly the performance of small-R jet HLT triggers, allowing to use lower offline thresholds and making the trigger more efficient, as shown in Figure 3.19.

The punch-through correction is not applied, since this correction would cost additional CPU with negligible improvement in the HLT jet triggers performance. The vast majority of uncontained calorimeter jets have a high energy, and those jets are nevertheless recorded as they are above the lowest unprescaled trigger threshold. This does indeed miss some extreme tails of lower energy jets that still are uncontained in the calorimeter, but usually the trigger efficiencies are defined at the $\sim 99.5\%$ efficient point, making their contribution negligible.

Figure 3.19: Efficiencies are shown for a single-jet trigger [154] with a 450 GeV threshold and three different calibrations applied to jets in the ATLAS high-level trigger. Offline jets are selected with $|\eta| < 2.8$. In green (open squares) the calibration applied in 2016 data, in red (closed circles) the updated calibration applied in 2017, utilising only calorimeter information, and in blue (open circles) this updated calibration additionally with track information. The extra calibration steps include the GS calibration and the application of in situ corrections. Since tracking is not guaranteed to be available for all jet thresholds, options are provided with and without the tracking-based corrections. These additional corrections allow for improved agreement between the scale of trigger and offline jets as a function of both η and p_{T}, and thus the trigger efficiency rises much more rapidly. Image from Ref. [155].



Chapter 4

Measurement of inclusive jet cross-sections

4.1 Preface

Precise measurements of jet cross-sections are crucial in understanding physics at hadron colliders. They probe quantum chromodynamics, the theory of strong interactions, where jets are interpreted as resulting from the fragmentation of quarks and gluons produced in a short-distance scattering process. Jet cross-sections provide valuable information about the strong coupling constant, α_s, and the structure of the proton. Also, inclusive jet events represent a background to many other processes at hadron colliders. The predictive power of fixed-order QCD calculations is therefore relevant in many searches for new physics.

Inclusive jet production has been measured in hadron collisions at the Sp$\bar{\mathrm{p}}$S and Tevatron colliders at various centre-of-mass energies. The latest and most precise results at $\sqrt{s} = 1.96$ TeV are detailed in Refs. [156, 157]. At the LHC, in Run 1, the ALICE, ATLAS and CMS collaborations measured inclusive jet cross-sections in proton-proton collisions at centre-of-mass energies of $\sqrt{s} = 2.76$ TeV [158–160] and $\sqrt{s} = 7$ TeV [161–164] and, recently, ATLAS and CMS at $\sqrt{s} = 8$ TeV [165, 166]. Finally, in Run 2, CMS published a low luminosity, $71\,\mathrm{pb}^{-1}$, result at $\sqrt{s} = 13$ TeV [167].

This chapter presents measurements of the inclusive jet cross-sections in proton-proton collisions at $\sqrt{s} = 13$ TeV centre-of-mass energy by the ATLAS Collaboration using data collected in 2015 and corresponding to an integrated luminosity of $3.2\,\mathrm{fb}^{-1}$. The inclusive jet cross-sections are measured double-differentially as a function of the jet transverse momentum, p_T, and absolute jet rapidity, $|y|$. Jets are reconstructed from EM topoclusters using the anti-k_t jet clustering algorithm with a radius parameter value of $R = 0.4$. The measurements cover the kinematic region of 100 GeV $< p_\mathrm{T} <$ 3.5 TeV and $|y| < 3$.

Next-to-leading-order perturbative QCD predictions calculated using several parton distribution function sets, corrected for electroweak and non-perturbative effects, are quantitatively compared to the measurement results. In addition, the measured cross-sections are compared to the recently published complete next-to-next-to-leading-order pQCD calculation [168, 169].

This chapter is organise as follows. Inclusive jet cross-sections are introduced in Section 4.2. The data and MC simulation samples used in this analysis are discussed in Section 4.3, while Section 4.4 summarises the event and jet selection. The jet energy scale and resolution, and their uncertainties are detailed in Section 4.5. Other systematic uncertainties are discussed in Section 4.6. Detector level jet kinematic distributions are presented in Section 4.7. The unfolding method used to deconvolute detector effects and the propagation of the systematic uncertainties are described in Section 4.8 and 4.9, respectively. The theoretical predictions are detailed in Section 4.10 and the results are presented and discussed in Section 4.11.

4.2 Cross-section definitions

In this measurement, jet cross-sections are determined for the so-called particle jets. These jets are built from stable particles, i.e. those fulfilling $c\tau > 10\,\text{mm}$, where τ is the proper lifetime. This definition includes muons and neutrinos, which is not the case for the particle jets used to derive the MC-based jet calibrations detailed in Section 3.2. pQCD at NLO and NNLO, on the other hand, predict cross-sections for jets at parton level. Non-perturbative corrections are applied to obtain theoretical predictions at particle level, as discussed in Section 4.10.2.

Jets are identified using the anti-k_t jet algorithm as implemented in the FastJet package with radius parameter $R = 0.4$. The use of the anti-k_t algorithm is well motivated since it is infrared- and collinear-safe, and produces geometrically well-defined ("cone-like") jets, as already discussed in Section 2.3.5.

Inclusive jet double-differential cross-sections are measured as a function of jet p_{T} in six equal-size bins of the absolute jet rapidity ($|y|$). Only jets in the kinematic range $p_{\mathrm{T}} > 100\,\text{GeV}$ and $|y| < 3.0$ are considered, to ensure that the jet energy scale is well understood, as described in Section 3.2. The inclusive jet production cross-section can be expressed as a ratio of the number of jets in data after correcting for detector effects, N_{jets}, to the integrated luminosity of the data, $\mathcal{L}$, in a given interval of momentum and rapidity, Δp_{T} and Δy respectively:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}p_{\mathrm{T}}\mathrm{d}y} = \frac{N_{\text{jets}}}{\mathcal{L}\Delta p_{\mathrm{T}}\Delta y}.$$

The p_{T} binning is chosen according to the detector p_{T} resolution, such that the bin width is approximately twice the p_{T} resolution, with the exception of the highest p_{T} bins in each rapidity range where bins are merged to avoid large statistical fluctuations and non-Gaussian statistical uncertainties due to a low number of entries. The criterion followed is that bins were merged when the expected number of events, as predicted by MC simulation (see Section 4.3), was less or equal than 10.

4.3 Dataset and Monte Carlo simulations

The measurement uses proton-proton collision data at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ collected by the ATLAS detector during the 2015 data-taking period of the LHC. The integrated collected luminosity is $3.2\,\text{fb}^{-1}$ with an uncertainty of 2.1%. The uncertainty in the luminosity is derived following a methodology similar to that detailed in Ref. [170], from a calibration of the luminosity scale using x-y beam-separation scans performed in August 2015.

Simulated jet events were produced using four different Monte Carlo event generators for comparisons to data and to derive corrections. The Pythia 8 program (version 8.186) was used for the baseline comparisons, the deconvolution of detector effects and the propagation of systematic uncertainties. It uses LO pQCD matrix elements for $2 \to 2$ processes, along with a leading-logarithmic p_{T}-ordered parton shower [171] including photon radiation, underlying event simulation with multiple parton interactions [172], and hadronisation with the Lund string model. The samples were created using a set of tuned parameters called the A14 tune and the NNPDF2.3LO LO PDF set. The EvtGen 1.2.0 program was used to model bottom and charm hadron decays. NLO samples of simulated events were produced using Powheg and showered with Pythia 8 for systematic studies as discussed in Section 4.5.1, and for optimising bin widths. The A14 tune and the CT10 [173] PDF set were used. In addition, NLO samples of simulated events were produced using Sherpa 2.1.1 [43] event generator together with CT10 PDF set for comparisons to data and for unfolding studies, as discussed in Section 4.7 and 4.8, respectively. For the evaluation of non-perturbative effects, the LO Pythia 8 and the Herwig++ (v2.7.1 [174]) event generators were also employed as described in Section 4.10.2.

In all the MC samples, the effects of multiple proton-proton interactions in the same and neighbouring bunch crossings were included by overlaying inelastic minimum-bias events generated with Pythia 8. Events were further weighted to reproduce the observed distribution of the average number of collisions per bunch crossing in data, as explained in Section 4.3.1. The stable particles from the generated events were passed through the ATLAS detector simulation based on Geant 4 and were reconstructed with the same version of the ATLAS software as was used to process the data.

4.3.1 Pile-up reweighting

The presence of pile-up is included in the MC, allowing to properly reproduce the effects of pile-up in the simulation. However, the distributions of pile-up collisions in data and MC do not match, as shown for instance in Figure 4.1, where the $\langle\mu\rangle$ distributions in MC and data are compared[1]. This is because the average pile-up distribution in MC is chosen before data was collected or before data taking ended. To match the distribution of pile-up collisions in MC and data, an event by event reweighting technique is applied to MC using the official ATLAS pile-up reweighting (PRW) tool. This is a two-step process. First, the integrated luminosity from the data is plotted in bins of average pile-up $\langle\mu\rangle$, with a binning chosen to match the discrete values used in the reference pile-up distributions of the MC. Then, the pile-up weight for a MC event corresponding to the i^{th} $\langle\mu\rangle$-bin is given by

$$w_i = \frac{L_i/L}{N_i/N},$$

where L is the total integrated luminosity of the data, L_i is the integrated luminosity of all data in bin i, N is the sum of generator-weights of the whole MC sample, and N_i is the sum of generator-weights of the events in bin i.

The results are shown in Figures 4.2 and 4.3. Figure 4.2 compares the N_{PV} distribution in data and MC before (left) and after (right) pile-up reweighting. The comparison between the $\langle\mu\rangle$ distributions in data and MC after pile-up reweighting is shown in Figure 4.3. The agreement in both distributions is clearly enhanced and it is overall very good.

[1]Studies of the number of vertices shows that the value of $\langle\mu\rangle$ in data derived from the luminosity was overestimated and it had to be scaled down by a factor of $1/1.16$.

Figure 4.1: Comparison of the pile-up conditions in MC (shaded area) and data (points). Both MC and data distributions are area normalised.

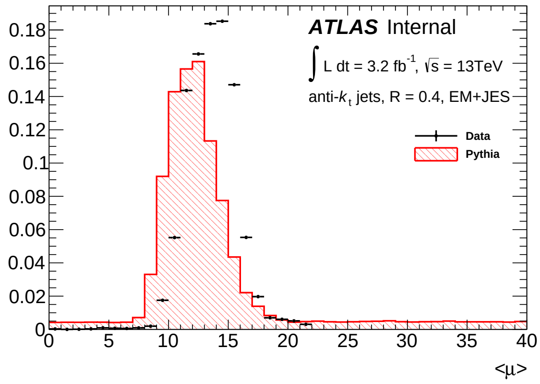
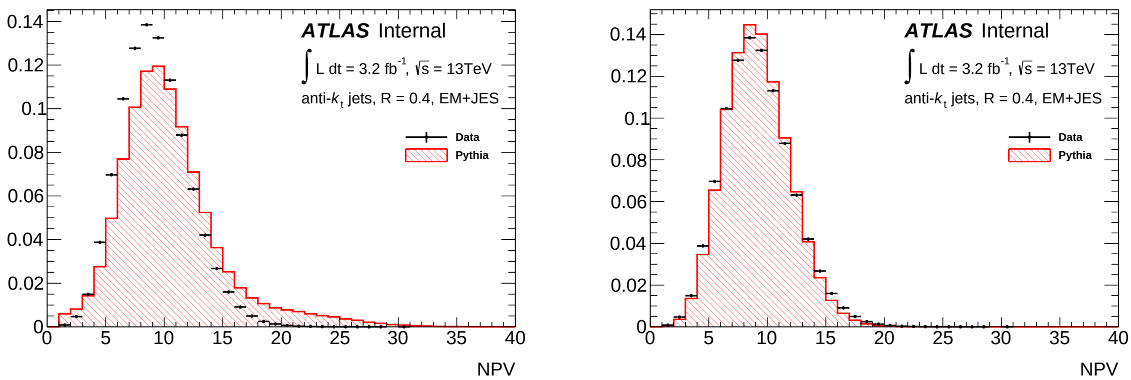
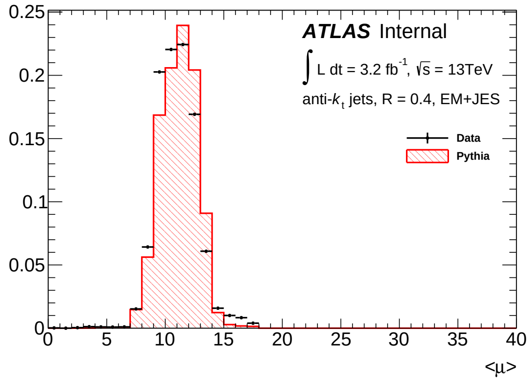


Figure 4.2: Comparison of the N_{PV} distribution in MC (shaded area) and data (points) before (a) and after (b) pile-up reweighting. Both MC and data distributions were normalised to the same area. The agreement is clearly enhanced.



(a) (b)

Figure 4.3: Comparison of $\langle\mu\rangle$ distribution in MC (shaded area) and data (points) after pile-up reweighting. Both MC and data distributions were normalised to the same area.



4.4 Event and jet selection

A suite of single-jet triggers [154] with thresholds varying from 55 GeV to 360 GeV are used to record events with at least one jet with transverse energy above the threshold in the region $|\eta| < 3.2$. To keep the trigger rate to an acceptable level, the triggers with lower E_{T} thresholds are prescaled by recording only a predefined fraction of events. The highest-threshold trigger accepts all events passing the threshold.

A p_{T}-dependent trigger strategy for selecting jets is adopted in order to optimise the statistical power of the measurement. For each event, a jet of a given p_{T} is used if the event fired the trigger with the lowest prescale[2] that is fully efficient at that p_{T}. The effective luminosities range from $81\,\mathrm{nb}^{-1}$ for $75 < p_{\mathrm{T}} < 100$ GeV, where the trigger prescaling is largest, to $3.2\,\mathrm{fb}^{-1}$ for $p_{\mathrm{T}} > 442$ GeV, where an unprescaled trigger is used. The 75–100 GeV p_{T} bin is used just to reduce the bias from the unfolding procedure (see Section 4.8) for $p_{\mathrm{T}} > 100$ GeV. The corresponding trigger thresholds and luminosities for each p_{T} bin of the measurement are shown in Table 4.1. The trigger efficiencies are studied offline in data as a function of p_{T} and rapidity, by emulating the online trigger decision. An event passes the trigger emulation (i.e., the trigger is considered fired) if there is at least one L1 and one HLT object that overstep the corresponding E_{T} threshold. The trigger efficiency for a given p_{T} bin is obtained as the fraction of events that pass the trigger emulation with the leading jet p_{T} falling in that bin in a sample obtained by requiring at least one online jet satisfying $E_{\mathrm{T}} > 15$ GeV. The result is shown for two representative $|y|$ bins in Figure 4.4. The trigger efficiency is always larger than 99.9% in the p_{T} range where it is used.

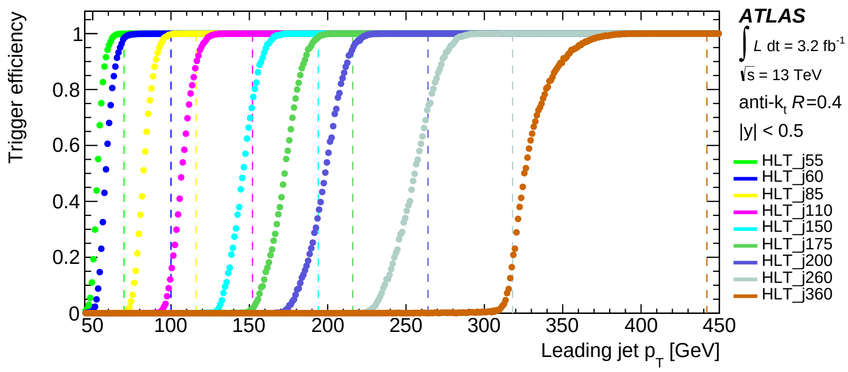
Table 4.1: Trigger strategy of the measurement. The table shows the p_{T} ranges filled by the respective HLT triggers and their corresponding luminosity. No dependence on the rapidity was found, and consequently, the same trigger thresholds are used in the full rapidity range.

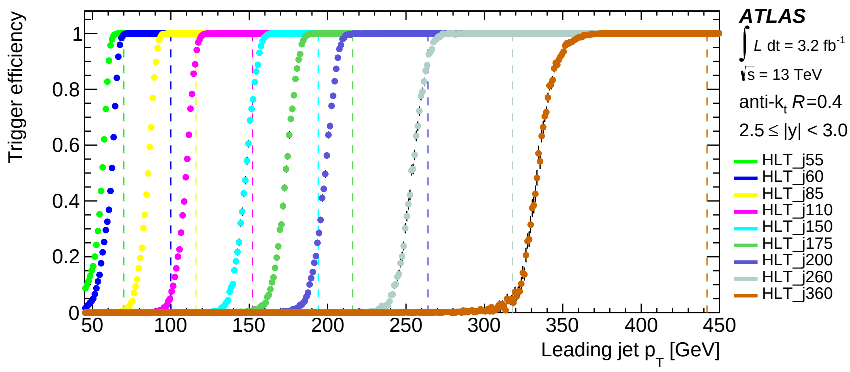
p_{T} range, [GeV]	99.9% efficiency trigger	Luminosity [pb^{-1}]
75–100	HLT_55	0.0807252
100–134	HLT_60	0.118673
134–152	HLT_85	0.483386
152–194	HLT_110	1.37096
194–240	HLT_150	5.09517
240–264	HLT_175	10.1739
264–318	HLT_200	18.7346
318–442	HLT_260	65.0237
442–∞	HLT_360	3158.32

The corresponding GRL (see Section 2.2.7) for the 2015 data taking period is used to select data. Events are further required to have at least one primary vertex with at least two associated well-reconstructed tracks satisfying $p_{\mathrm{T}} > 400$ MeV to reject events due to cosmic-ray muons and other non-collision backgrounds. Quality criteria are applied, using the `LooseBad` jet definition defined in Ref. [175], to reject events with jets from beam-induced background due to proton losses upstream of the interaction point, cosmic-ray air

[2]That is, with highest effective luminosity.

Figure 4.4: Turn-on curves of the trigger efficiency as a function of the leading jet p_{T} for the first (top) and last (bottom) rapidity bins, obtained by emulating the single-jet triggers in data. The dashed lines represent the trigger thresholds used in the analysis. Similar curves are obtained for the rest of the rapidity bins.



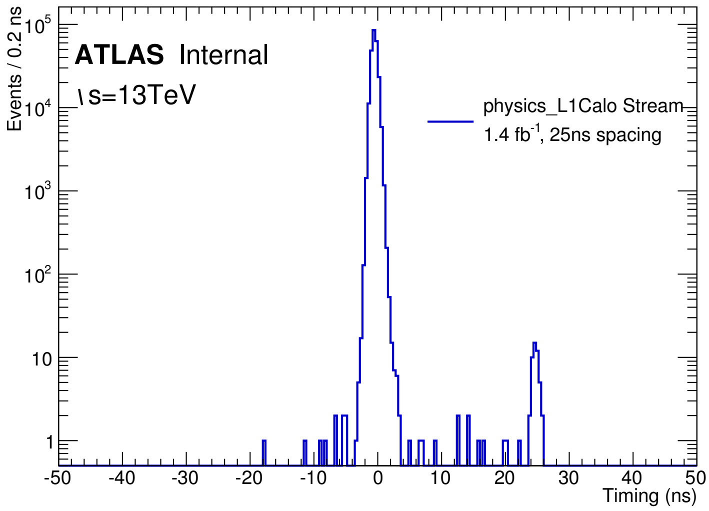


showers overlapping with collision events and calorimeter noise from large-scale coherent noise or isolated pathological cells. In addition, the non-collision background contribution is further reduced applying the `TightBad` quality criteria [175] to leading p_{T} jets.

4.4.1 L1 trigger mistimed events

A small number of events (110) were found to be mistimed in the full 2015 dataset (see Figure 4.5). These mistimed events are due to an artifact of the Finite Input Response (FIR) filters used in the regular Bunch Crossing Identification (BCID) algorithm [82]. The BCID algorithm runs on a stream of data from each trigger tower that is digitised at 40 MHz. The regular BCID algorithm is based on the peak position of a FIR filter which is also used to estimate the tower transverse energy. The FIR filter is formed using successive weighted sums of four time samples.

Figure 4.5: Median tower timing for the $\sim 1.4\,\mathrm{fb}^{-1}$ of data showing the mistimed events.



The early BCID of the regular algorithm that appeared in Run 2, occurred due to an unanticipated effect of the negative coefficients used in the FIR filter. These negative coefficients cause an incorrect BCID assignment if the ADC value of the first sample before saturation is high enough and if it is followed by at least four saturated samples. The Run 1 configuration only used positive FIR coefficients, and thus this issue did not affect the Run 1 data.

A small number of events were mistimed to the bunch crossing, and thus, were not accepted. However, these events were still saved in the so-called "physics_L1Calo" stream and hence it was possible to recover them from there. These events were reconstructed shifting the LAr and Tile pulses by 25 ns to calculate the correct calorimeter energies. However, it was not possible to recover the tracking information and muon segments. Hence, the `TightBad` jet quality criteria can not be used, and only `LooseBad` is used instead.

The performance of the jets after the reprocessing and the impact of the missing information was extensively studied for the analysis presented in Ref. [176] and it was concluded that the possible sources of miscalibration are small compared to the statistical uncertainty. Additional studies have been made in order to ensure the usability of the mistimed events in this measurement.

All the events have $\Delta\Phi$, between the two leading p_T jets, close to π, as can be seen in Figure 4.6, and the second leading p_T jet has reasonable p_T fraction of the leading p_T jet, as shown in Figure 4.7. Finally, the maximum fractional energy deposition in a single calorimeter layer (f_{max}) distribution of leading p_T jets is shown in Figure 4.8. One single jet event was found with f_{max} of 0.94 and p_T of 1.2 TeV, but there is no evidence suggesting to drop this event. The potential removal of the these events would affect the cross-section in about 0.01%. All these studies suggest that the mistimed events are safe to be included in the analysis, and they were.

Figure 4.6: Distribution of $\Delta\Phi$ between the two leading p_T jets in mistimed events.

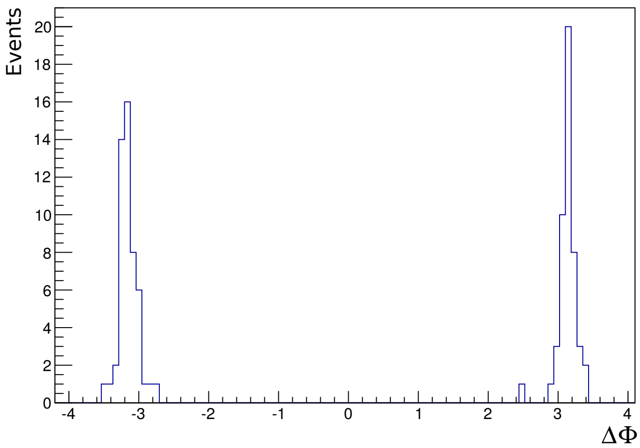


Figure 4.7: Distribution of the fraction of second leading jet p_T with respect the leading jet p_T in mistimed events.

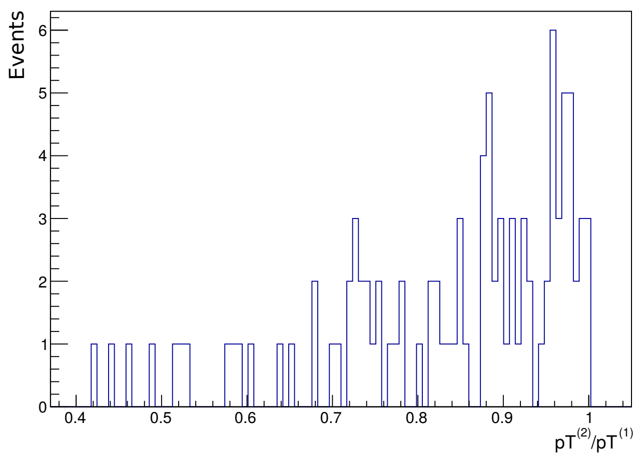
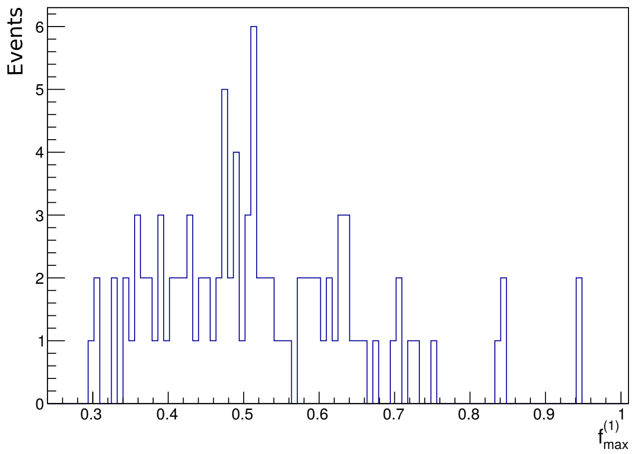


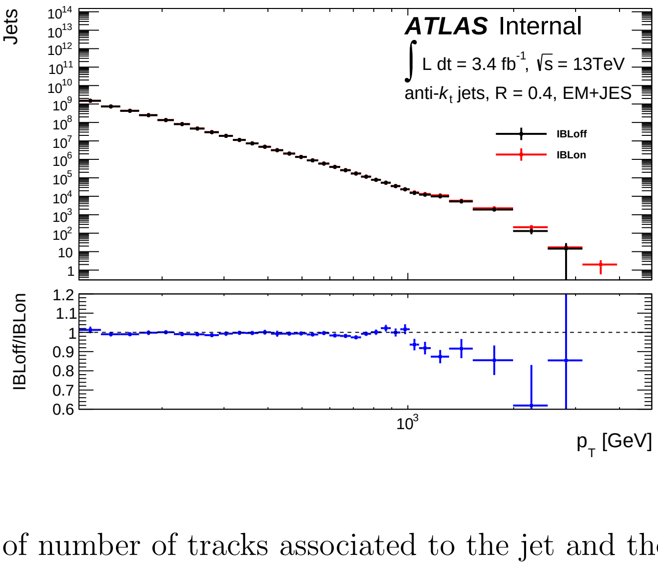
Figure 4.8: Distribution of the maximum fractional energy deposition in a single calorimeter layer (f_{max}) of the leading p_T jet in mistimed events.



4.4.2 Data recorded with the IBL sub-detector not functioning

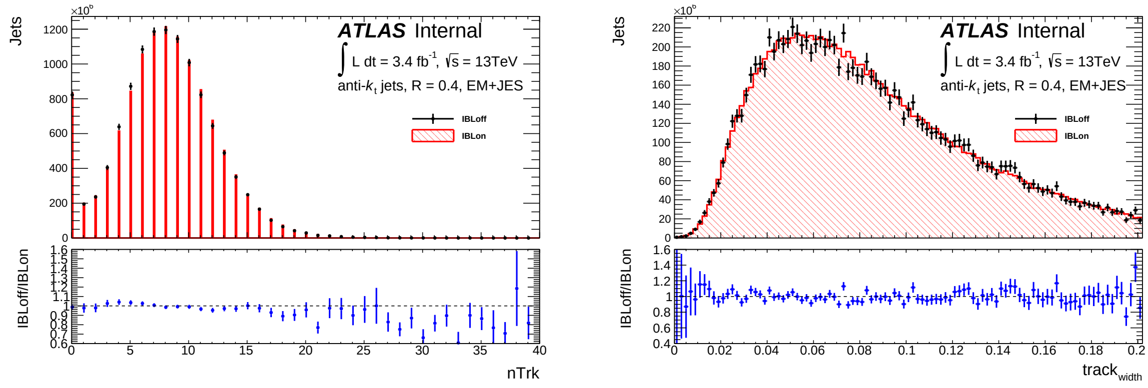
A small subset of the data collected in 2015, with a luminosity of 217 pb^{-1}, was recorded with the IBL sub-detector off or in standby. A comparison of the inclusive p_T distribution of jets between IBL off and IBL on data is depicted in Figure 4.9. A 10% level of disagreement at high p_T is observed.

Figure 4.9: The p_T distribution at the detector level of jets with $|y| < 3.0$ in IBL off data (black points) and in IBL on data (red points) normalised to the total area. The ratio of the two (blue points) is shown in the bottom panel. At high p_T a difference of 10% is seen.



The distribution of number of tracks associated to the jet and the distribution of the width of the jet measured with the tracks associated to it were compared between IBL off and IBL on data (see Figure 4.10). No significant differences were found in order to explain the 10% level disagreement found in the jet p_T distribution. Since the analysis is not dominated by statistics, it was decided not to include the IBL off data in the measurement.

Figure 4.10: Comparison between IBL off and IBL on data on the distribution of: (a) number of tracks associated to the jet, (b) the width of the jet measured with the tracks associated to it.

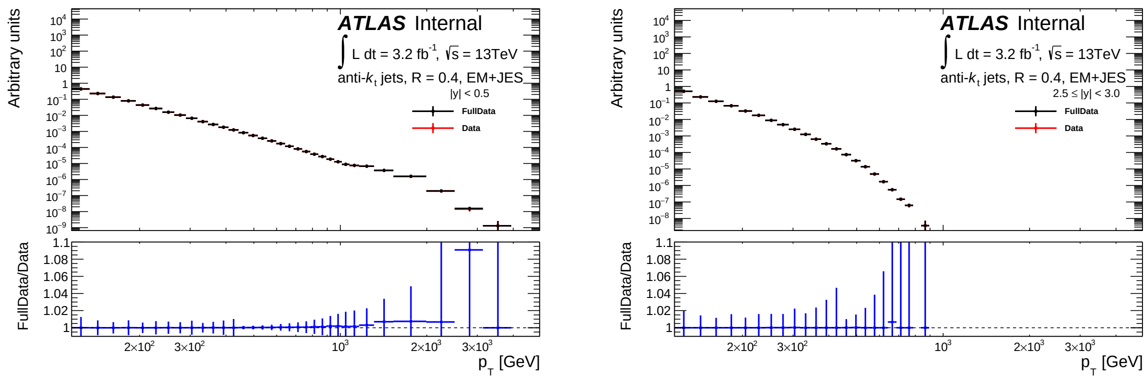


(a) (b)

4.4.3 Data recorded in the debug stream

There are events where the HLT does not give an answer because of a timeout or a crash. These are saved to the debug stream. The HLT is run offline on these events, and for events which would have passed the HLT these are reconstructed as for the main physics stream. These events were added to the analysis. The difference of adding this stream is shown in Figure 4.11 for representative $|y|$ bins. The impact of adding the debug stream on the jet p_{T} distributions is small, except at high p_{T} and the central region where it represents and increment of 10% in the number of jets with p_{T} of the order of 2 TeV.

Figure 4.11: Jet p_{T} distribution at the detector level in different bins of the absolute jet rapidity in data, including the debug stream (black points), and not including it (red points). The ratio of the two (blue points) is shown in the bottom panels.



(a) (b)

4.4.4 Cut Flow summary

The cut flow for the inclusive jet cross-section measurement as a function of various requirements that are sequentially applied is shown in Table 4.2 for anti-k_t jets with $R = 0.4$.

Table 4.2: Cut flow of the inclusive jet cross-section measurement as a function of sequentially applied requirements, excluding 3,199,968 events with the IBL not functioning, including 122,937 events from the debug stream and 110 recovered mistimed events .

Event requirements (cumulative)	Number of Events
Total number of events passing any jet trigger	54,634,195
After GRL	52,530,455
After requiring $N_{\mathrm{trk@PV}} > 1$	52,487,179
After requiring at least one central single-jet trigger	28,640,275
After applying jet cleaning criteria	25,372,957
Passing jet p_{T} and rapidity selection	19,552,836

4.5 Jet energy calibration and resolution

This measurement uses AntiKt4EMTopo jets, calibrated with the full jet energy scale (JES[3]) calibration following the procedure introduced in Section 3.2. The systematic uncertainties associated to the JES, as well as the jet resolution and its uncertainties are briefly described in this section.

4.5.1 Jet energy scale uncertainties

The measurement presented in this chapter uses the most detailed description of the systematic uncertainties considered in ATLAS. There are, in total, 76 independent sources of systematic uncertainty treated as being uncorrelated among each other [130]. These are summarised in Table 4.3 and introduced next.

The systematic and statistical uncertainties of each of the in situ calibration techniques contribute to the total JES uncertainty as independent systematic components.

Differences between the calorimeter responses to jets initiated by light quarks or gluons and a lack of knowledge of the flavour composition of the analysed data lead to additional uncertainties, known as flavour response and flavour composition uncertainties, respectively. This is because the uncertainty on the response of LQ-initiated jets is covered by the rest of the JES uncertainties while the uncertainty on the gluon response requires additional uncertainty components. The flavour response uncertainty is $f_{\mathrm{g}} \cdot \Delta\mathrm{R}_{\mathrm{g}}$, while the flavour composition uncertainty is $\frac{\Delta f_{\mathrm{g}} \cdot |\mathrm{R}_{\mathrm{LQ}} - \mathrm{R}_{\mathrm{g}}|}{f_{\mathrm{g}} \cdot \mathrm{R}_{\mathrm{g}} + f_{\mathrm{LQ}} \cdot \mathrm{R}_{\mathrm{LQ}}}$, where f_{g} and f_{LQ} are the fraction of gluon- and LQ-initiated jets in the sample, respectively. Δf_{g} is the uncertainty on the gluon fraction. R_{g} and R_{LQ} are the response of gluon- and LQ-initiated jets in Pythia 8, respectively, and $\Delta\mathrm{R}_{\mathrm{g}}$ is the difference between the responses of gluon-initiated jets predicted by Pythia 8 and Herwig++. For instance, if $f_{\mathrm{LQ}} = 1$ and $\Delta f_{\mathrm{g}} = 0$, then both flavour uncertainties are zero. On the contrary, if $f_{\mathrm{g}} = 1$ and $\Delta f_{\mathrm{g}} = 0$, then the flavour response uncertainty accounts for the difference on the gluon response between two different MC generators, and the flavour composition uncertainty is zero. In this case, the response is different than in the samples used for the calibrations. Still, that applies for both data and MC simulation used in the unfolding procedure and hence should have no effect. If $f_{\mathrm{g}} = 1$ but $\Delta f_{\mathrm{g}} \neq 0$, then the flavour composition uncertainty is the uncertainty on the

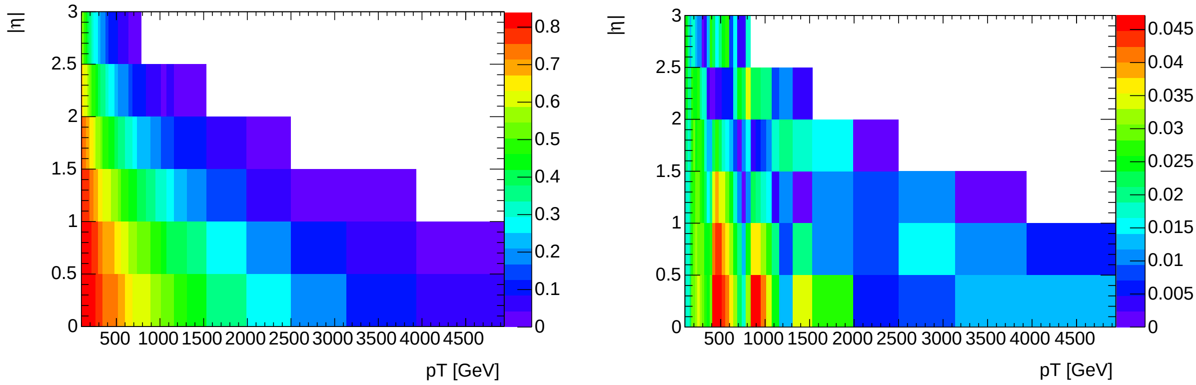
[3] Despite being confusing, the absolute MC-based calibration and the full jet energy scale calibration chain are both known as JES. It should be clear from the context.

Table 4.3: Summary of the systematic and statistical uncertainties in the JES [130].

Name	Description
Z+jet	10 systematic uncertainty components Statistical uncertainty over 13 regions of jet p_{T}
γ+jet	5 systematic uncertainty components Statistical uncertainty over 15 regions of jet p_{T}
Multijet balance	5 systematic uncertainty components Statistical uncertainty over 16 regions of $p_{\mathrm{T}}^{\mathrm{leading}}$
η-intercalibration	3 systematic uncertainty components Statistical uncertainty
Pile-up	Uncertainty of the μ modelling in MC simulation Uncertainty of the N_{PV} modelling in MC simulation Uncertainty of the ρ modelling in MC simulation Uncertainty in the residual p_{T} dependence
Jet flavor	Uncertainty in the jet composition b/w quarks and gluons Uncertainty in the jet response of gluon-initiated jets
Punch-through	Uncertainty in GSC punch-through correction
Single-particle response	Uncertainty from single-particle and test-beam analyses

gluon fraction times the relative response difference between LQ- and gluon-initiated jets in Pythia 8. Pythia 8 and Powheg+Pythia 8 Monte Carlo simulations are used to estimate the flavour composition of the sample as a function of p_{T} and rapidity. The result from Pythia 8 is taken as the nominal quark/gluon composition, and the difference between the two simulations as an estimate of the fractions uncertainty. The gluon fractions and their uncertainty are shown in Figure 4.12. $\mathrm{R_g}, \mathrm{R_{LQ}}$ and $\Delta\mathrm{R_g}$ are provided by ATLAS. As a cross-check, the fraction of jets with no assigned flavour was investigated and was found to be negligible, as shown in Figure 4.13, and hence no uncertainty was considered due to this effect.

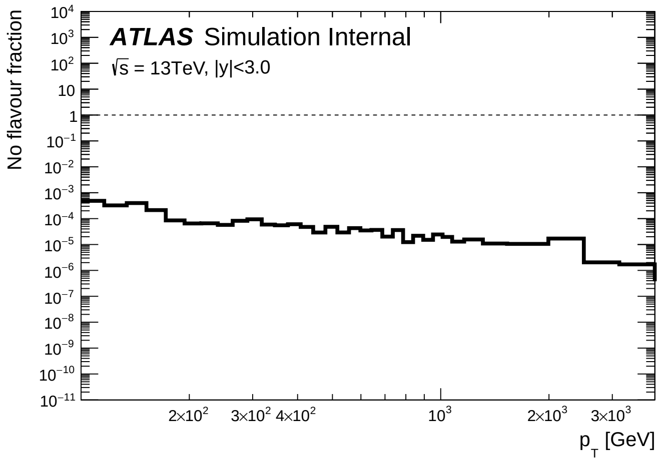
Figure 4.12: (a) the fraction of gluon-induced jets and (b) the uncertainty on the gluon fraction, in bins of jet pseudorapidity and p_{T} for anti-k_t jets with $R = 0.4$.



(a) (b)

A systematic uncertainty on the punch-through correction, applied only to jets with $N_{\mathrm{segments}} > 20$, is derived as the difference in the jet response between data and MC in dijet events as a function of the number of muon segments [138].

Figure 4.13: Fraction of jets in Pythia 8 with no flavour as a function of jet p_{T} for jets with $|y| < 3.0$.



An uncertainty in the jet energy scale at high p_{T} ($p_{\mathrm{T}} > 2000\,\mathrm{GeV}$), where in situ methods cannot be used, is derived from single-particle response measurements [177], and is known simply as E/P uncertainty.

Four uncertainties are included to account for potential mismodelling of pile-up in the MC simulation: the number of reconstructed primary vertices, the average number of interactions per bunch crossing, the energy density in jets and the residual dependence of the jet p_{T} on pile-up. The description and evaluation of the pile-up uncertainties are detailed in Refs. [130, 141].

All of these uncertainties are treated as being fully correlated across p_{T} and η, with the exception of the statistical uncertainty of the η-intercalibration which is propagated as being uncorrelated between the 245 different η and p_{T} bins in which it was derived [130, 142]. The JES uncertainty is 1% in the 200–600 GeV range of jet p_{T}, 2% at 2 TeV, and reaches 3% above 3 TeV. The uncertainty is fairly constant as a function of η and reaches 2.5% at 80 GeV for the most forward jets [130].

4.5.2 Jet energy resolution and its uncertainties

The jet p_{T} resolution (JER) is derived using the data collected during 2012. It is obtained in situ from the standard deviation of the ratio of the p_{T} of a jet to the p_{T} of other well-measured objects (a photon or a Z boson [142, 143]) in an event, following techniques similar to those used to determine the JES uncertainty. The p_{T}-balance technique in dijet events (η-intercalibration) [142] allows a measurement of the JER at high jet rapidities and for a wide range of transverse momenta. Noise from the calorimeter electronics and pile-up forms a significant component of the JER at low p_{T}. A study in zero-bias data[4] allows this contribution to be constrained. In addition, a MC simulation is used in each in situ JER measurement to correct for fluctuations present at particle level due to the underlying event and out-of-cone contributions from QCD radiation and hadronisation. The results from all these methods are combined in a way similar to that for the JES [144].

The JER uncertainty has in total 11 components. Eight of these components are obtained by combining the systematic uncertainties associated to the in situ methods. One

[4]The zero-bias sample contains data collected by recording events exactly one accelerator turn after a high-p_{T} first-level calorimeter trigger. These events will thus be contained in a random filled bunch collision with a rate proportional to the instantaneous luminosity [144].

component is the uncertainty due to the electronic and pile-up noise measurement. Another is the absolute JER difference between data and MC simulation as determined with in situ methods. Finally, the JER uncertainties are completed with an extra component to account for the differences between the 2012 and 2015 data-taking conditions [178]. Each JER systematic component describes an uncertainty that is taken to be fully correlated in jet p_{T} and η. The 11 JER components are treated as fully uncorrelated with each other.

4.5.3 Jet angular resolution and its uncertainties

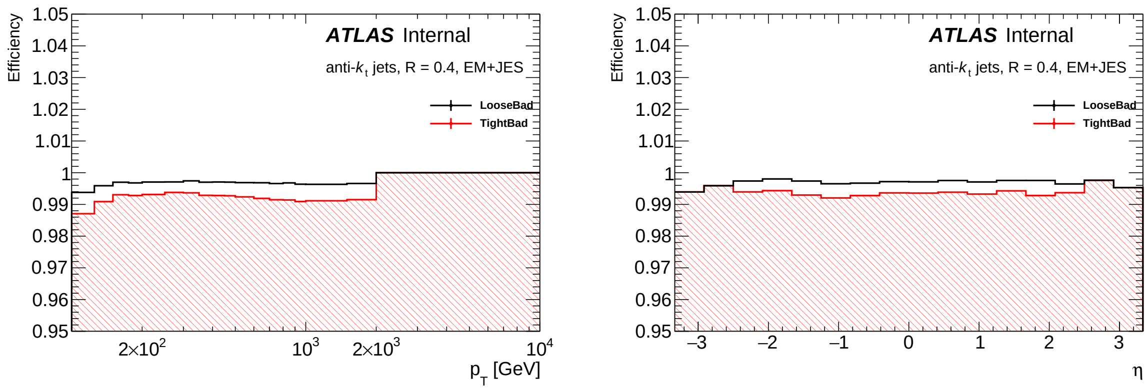
The jet angular resolution (JAR) is estimated in MC simulation from the differences in rapidity and azimuthal angle between reconstructed jets and matching particle jets. This estimate is validated by comparing the standard jets built from calorimeter energy deposits to those built from tracks in the inner detector [139, 179]. From these studies, the JAR is assigned an uncertainty of 10% to account for possible differences between data and MC simulation.

4.6 Other systematic uncertainties

4.6.1 Jet cleaning uncertainties

The efficiency for selecting jets from proton-proton collisions using the jet quality criteria (`LooseBad` and `TightBad`) was studied by an ATLAS dedicated group [175]. These efficiencies, as a function of p_{T} and η, are shown in Figure 4.14, and are of the order of 99%. In Figure 4.15, these efficiencies are shown simultaneously as a function of p_{T} and η. Each value was taken as the minimum efficiency for the given p_{T} and η bins shown in Figure 4.14.

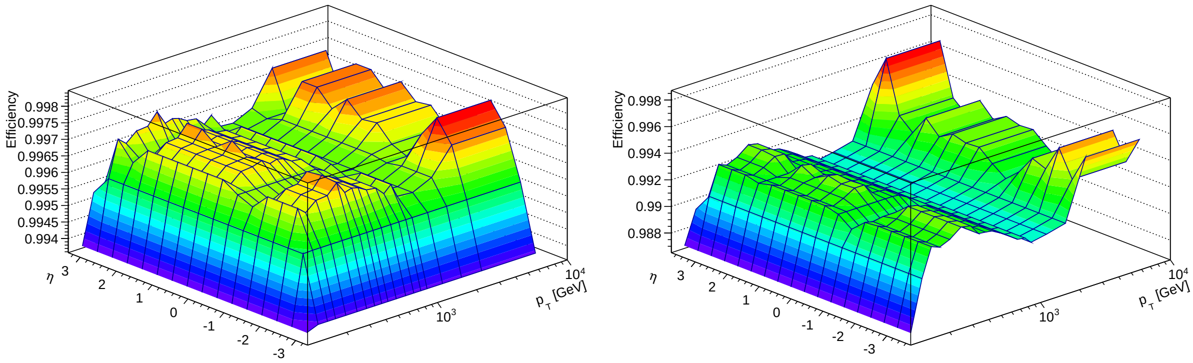
Figure 4.14: Efficiency of selecting jets from proton-proton collisions as a function of p_{T} (a) and η (b) for the `LooseBad` (black) and `TightBad` (red) jet quality criteria.



(a) $|y| < 3.0$ (b) $p_{\mathrm{T}} > 100\,\mathrm{GeV}$

Given that no uncertainty is assigned to these efficiencies, it was decided not to apply them as corrections (1/efficiency), but instead to consider them as an asymmetric uncertainty.

Figure 4.15: Efficiency of selecting jets from proton-proton collisions as a function of p_{T} and η for the `LooseBad` (a) and `TightBad` (b) jet quality criteria.



(a) `LooseBad` (b) `TightBad`

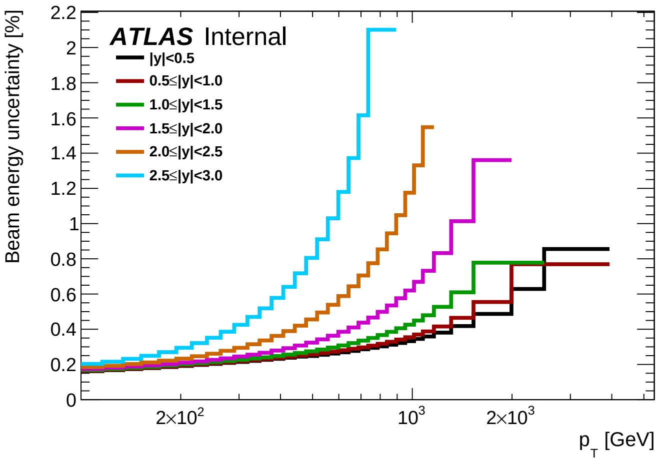
4.6.2 Jet reconstruction uncertainty

In the measurement performed at $\sqrt{s} = 7$ TeV, jet reconstruction efficiency was evaluated and an inefficiency was found only for jets with $p_{\mathrm{T}} < 30\,\mathrm{GeV}$. Hence, no inefficiency in jet reconstruction is expected in this measurement and no uncertainty is assigned for this efficiency.

4.6.3 Beam energy uncertainty

An uncertainty in the beam energy of 0.1% [180] is considered when comparing data with the theory prediction (discussed in detail in Section 4.10) at a fixed beam energy. The induced uncertainty at the cross-section level is evaluated by comparing the theory predictions at the nominal and shifted beam energies. It amounts for 0.2% at low p_{T} and 0.9% at high p_{T} in the central region and rises to 2% at the highest p_{T} and high rapidity, as shown in Figure 4.16.

Figure 4.16: Beam energy uncertainty contribution to the theoretical cross-section as a function of the jet p_{T} for each $|y|$ of the measurement.



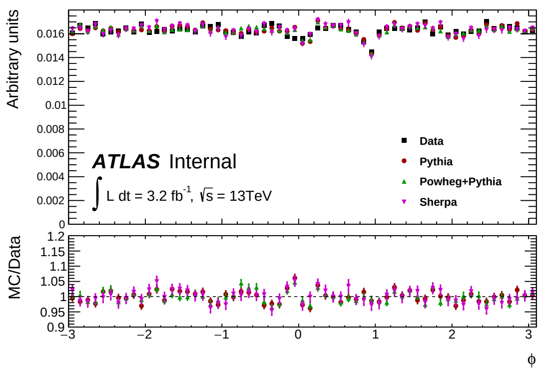
4.7 Detector level jet kinematic distributions

The Pythia MC sample described in Section 4.3 is used to correct for detector effects in the unfolding procedure described in Section 4.8. It is thus important that the MC simulation used to unfold shows reasonable agreement with data in the basic detector level jet kinematic distributions. In the following sub-sections the jet kinematic distributions relevant for the measurement are compared between data and different MC simulation samples (Pythia 8, Powheg+Pythia 8 and Sherpa 2.1.1).

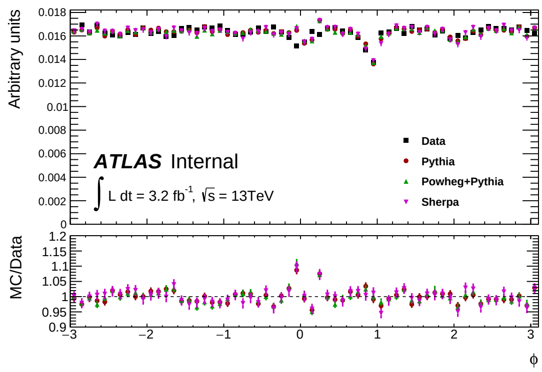
4.7.1 ϕ distributions

The jet ϕ distribution for different p_{T} ranges is presented in Figure 4.17. All the MC simulations describe data to within few percent accuracy. The dip around $\phi \sim 1$ is related to the dead Tile modules during the 2015 data taking period, and are very well described by the simulations.

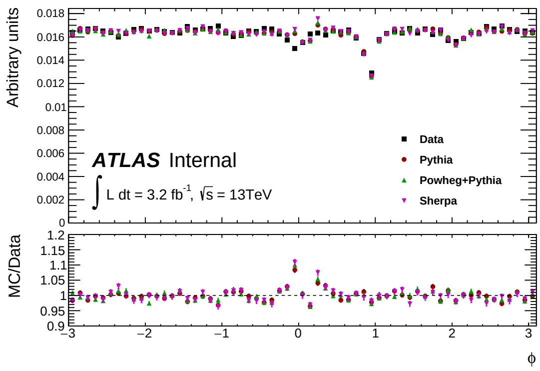
Figure 4.17: The jet ϕ distribution at the detector level in four different p_{T} ranges in data (black squares), Pythia (red circles), Powheg+Pythia (green upper triangles), and Sherpa (pink lower triangles), all normalised to the total area. The ratio of the MC simulations to the data is shown in the bottom panels.



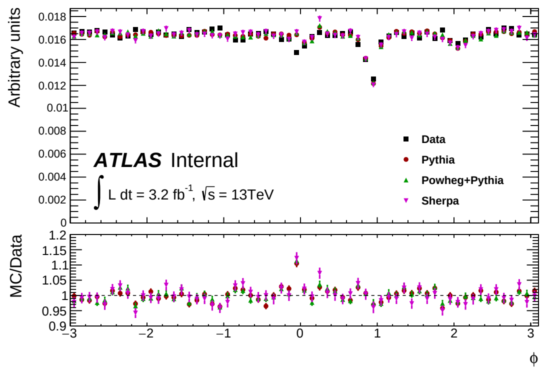
(a) $p_{\mathrm{T}} \geq 150$ GeV



(b) $p_{\mathrm{T}} \geq 250$ GeV



(c) $p_{\mathrm{T}} \geq 500$ GeV

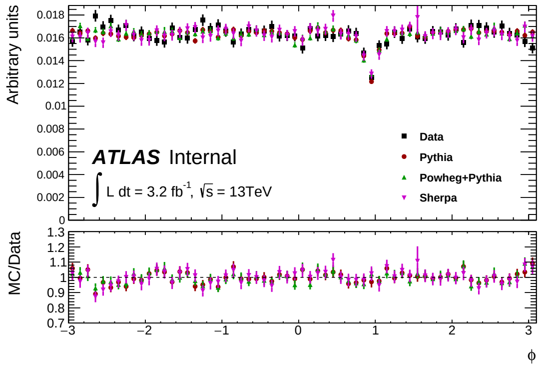


(d) $p_{\mathrm{T}} \geq 642$ GeV

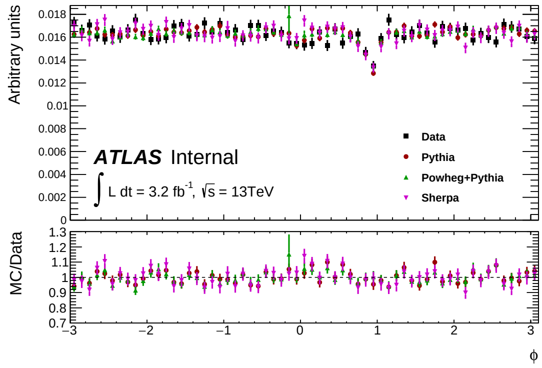
Since the measurement is done in bins of $|y|$, the ϕ distributions in data and MC simulations are compared in the jet rapidity bins of the measurement and are shown in

Figure 4.18. All MC simulation samples present similar (and good) agreement with the data.

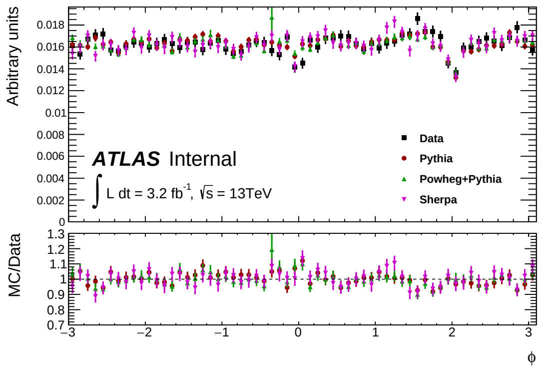
Figure 4.18: The jet ϕ distribution at the detector level in the six jet rapidity bins of the measurement in data (black squares), Pythia (red circles), Powheg +Pythia (green upper triangles), and Sherpa (pink lower triangles), all normalised to the total area. The ratio of the MC simulations to the data is shown in the bottom panels.



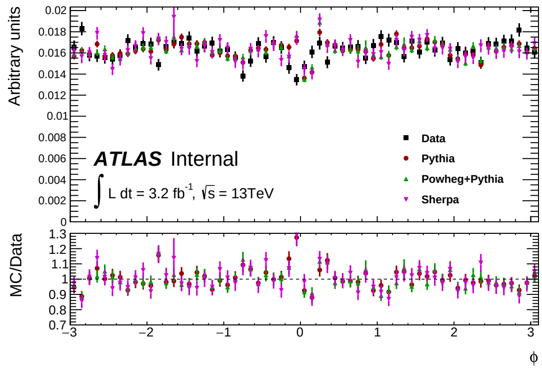
(a) $|y| < 0.5$



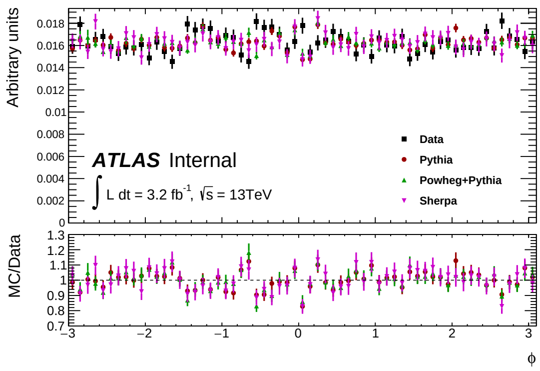
(b) $0.5 < |y| < 1.0$



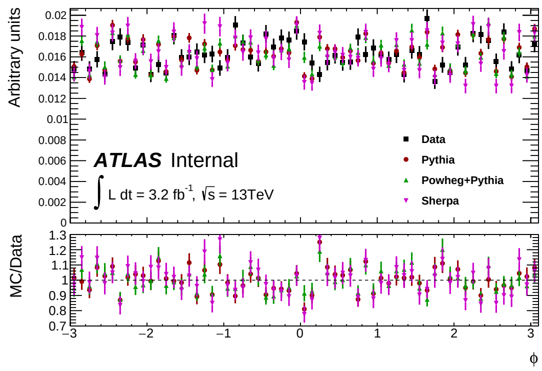
(c) $1.0 < |y| < 1.5$



(d) $1.5 < |y| < 2.0$



(e) $2.0 < |y| < 2.5$

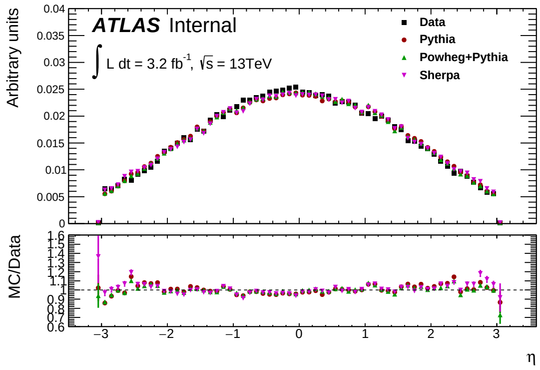


(f) $2.5 < |y| < 3.0$

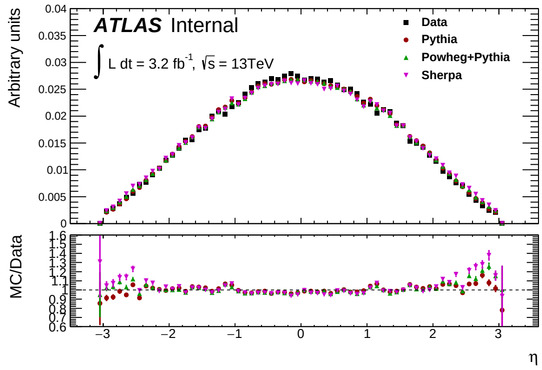
4.7.2 η distributions

The jet pseudorapidity distribution for different p_{T} ranges in data is compared to three MC simulations in Figure 4.19. The accuracy to model the η distribution of the simulations is satisfactory and comparable, although Pythia 8 performs better in the forward region with respect to the NLO simulations.

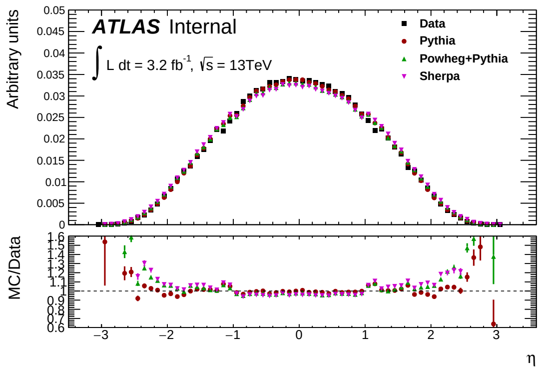
Figure 4.19: The jet η distribution at the detector level in six different p_{T} ranges in data (black squares), Pythia (red circles), Powheg +Pythia (green upper triangles), and Sherpa (pink lower triangles), all normalised to the total area. The ratio of the MC simulations to the data is shown in the bottom panels.



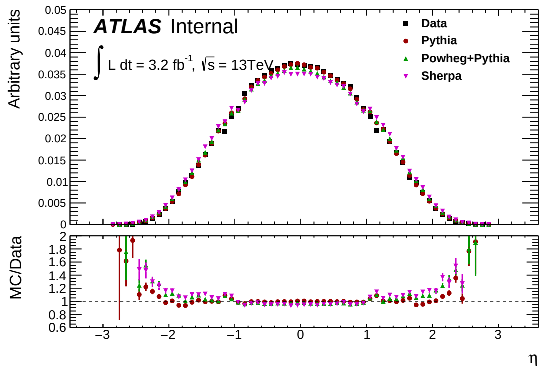
(a) $p_{\mathrm{T}} \geq 150$ GeV



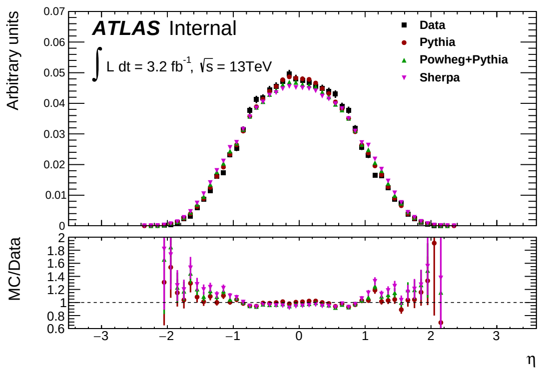
(b) $p_{\mathrm{T}} \geq 250$ GeV



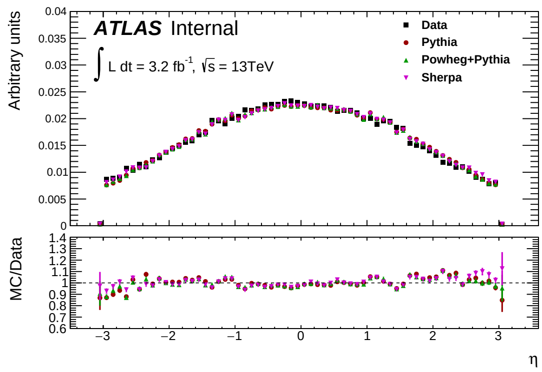
(c) $p_{\mathrm{T}} \geq 500$ GeV



(d) $p_{\mathrm{T}} \geq 642$ GeV



(e) $p_{\mathrm{T}} \geq 1100$ GeV



(f) inclusive

4.7.3 p_T distributions

Figure 4.20 shows the inclusive jet p_T distribution in bins of $|y|$. In the first three $|y|$ bins the MC/data ratios are flat over the whole p_T range within 10%. It is important to remark that the value itself of the ratio is not determinant. In effect, the area normalisation of data and MC distributions are dominated by the first few bins, where the bulk of the cross-section lies. That is to say, regardless of the distributions, the ratio of normalised plots will always agree at low p_T. As a consequence, the plots should not be regarded on a quantitative basis, but only as shape comparions. In the fourth $|y|$ bin, the MC/data ratio for PYTHIA deviates from a flat behaviour with increasing p_T, indicating that PYTHIA underestimates the yield of high p_T jets. At $2.0 < |y| < 2.5$, all MC simulations present comparable deviations (in absolute values) to the data, although PYTHIA presents a better modelling of the high p_T range. In the most forward $|y|$ bin, the NLO MC simulations considerably overestimate the number of high p_T jets.

The separated p_T distributions for leading, second leading, and third leading p_T jets are presented in Figure 4.21, the distribution including all jets in the events is also shown. The level of agreement is again satisfactory, except for the POWHEG +PYTHIA MC generator which shows a considerable mismodelling of the third leading p_T distribution, as shown in Figure 4.21d.

Due to the substantial mismodelling of the p_T distribution at large rapidity shown by the NLO MC generators (and of the third leading p_T distribution observed for the POWHEG +PYTHIA MC simulation sample), it was decided to use PYTHIA 8 as the baseline MC simulation. A systematic uncertainty on the unfolding procedure due to the difference in shape between data and baseline MC simulation is introduced in Section 4.8.

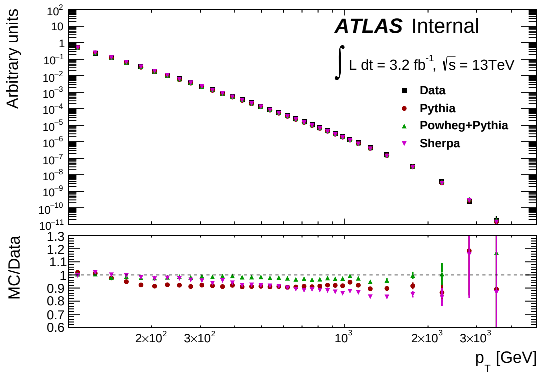
4.7.4 η–ϕ maps

Figure 4.22 shows the η–ϕ plots for both PYTHIA (left) and data (right) at intermediate p_T (top) and high p_T (bottom). In PYTHIA, there is an unpopulated η–ϕ region ($\eta \sim 2, \phi \sim 0$) which is not apparent in data. In order to investigate this region in detail, the ϕ distribution for data and MC in the η range $1.75-2.25$ is shown in Figure 4.23a. It is clear that although the data is flat in ϕ, as expected from the basically azimuthal symmetry of the detector, the MC displays a different behavior. The two central bins around $\phi = 0$ lie below the mean of the distribution, while their adjacent bins are above it. This fluctuation essentially cancels to agreement with data when averaged (see Figure. 4.23b). This unusual situation, in which the MC simulates a problem not present in data, could be a leftover correction for a module that is not dead anymore. This is not considered a problem for this measurement given the two uses of the MC in this analysis, the unfolding due to the energy resolution and the estimation of systematic uncertainties, which are not affected by this redistribution of events between adjacent ϕ bins.

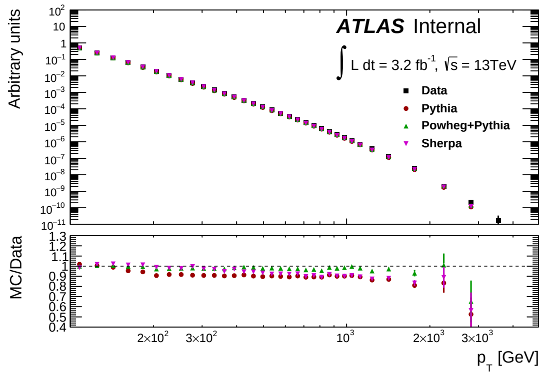
4.7.5 Impact of the tight cleaning criteria on the data

Figure 4.24 shows the inclusive p_T distribution at the detector level in data when using the `TightBad` cleaning criteria for the leading p_T jets and `LooseBad` for the rest of the jets, and when using `LooseBad` for all the jets. The change in the inclusive jet cross-sections measured in data by using the `TightBad` cleaning criteria in the leading p_T jets is 1% at low p_T and 15% at high p_T for the first $|y|$ bin of the measurement, while it is negligible across the full p_T range for the last $|y|$ bin. As shown in Figure 4.25, the MC/Data ratio

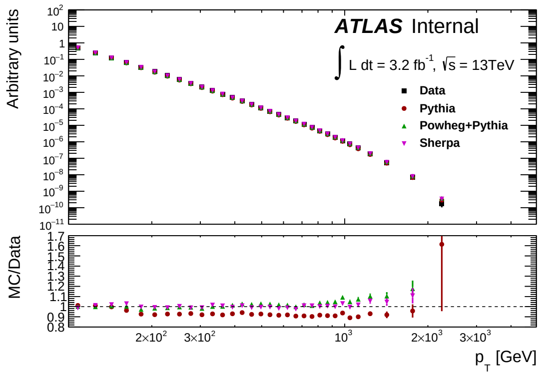
Figure 4.20: The p_{T} distribution at the detector level in different bins of the absolute jet rapidity in data (black squares), PYTHIA (red circles), POWHEG +PYTHIA (green upper triangles), and SHERPA (pink lower triangles), all normalised to the total area. The ratio of the MC simulations to the data is shown in the bottom plots.



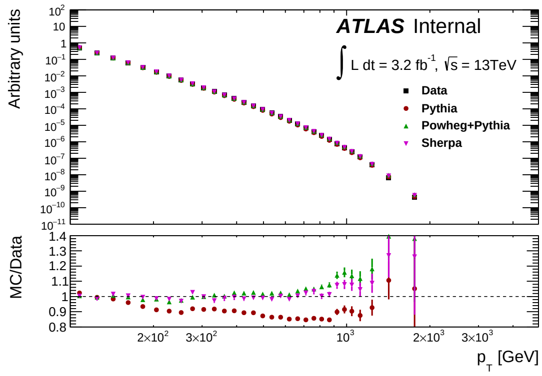
(a) $|y| < 0.5$



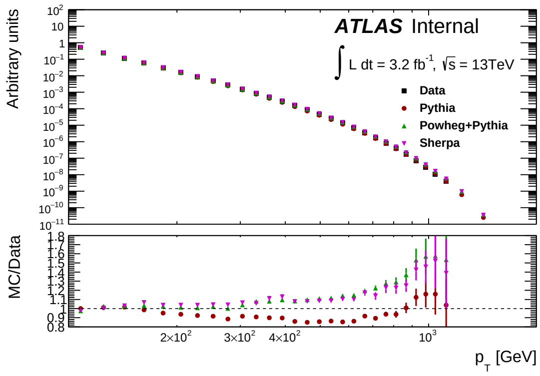
(b) $0.5 < |y| < 1.0$



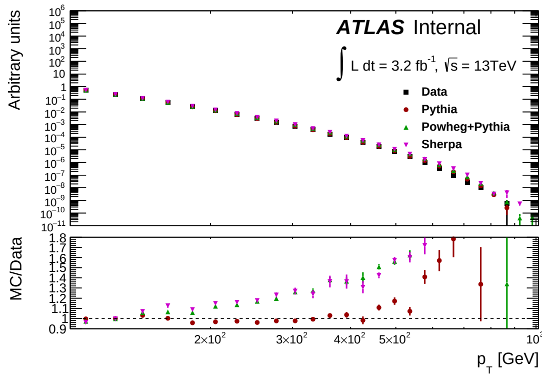
(c) $1.0 < |y| < 1.5$



(d) $1.5 < |y| < 2.0$



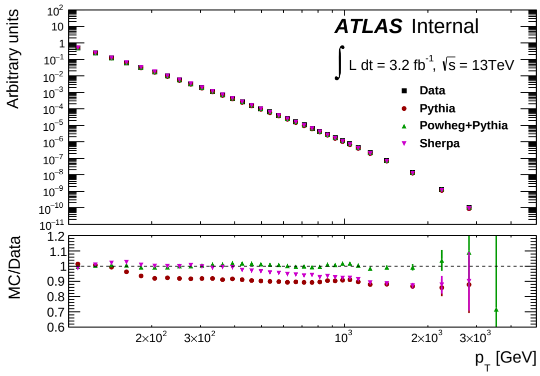
(e) $2.0 < |y| < 2.5$



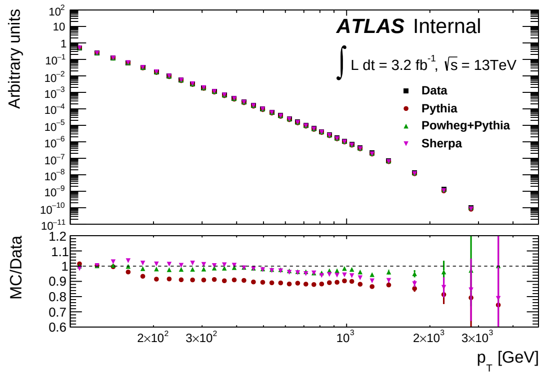
(f) $2.5 < |y| < 3.0$

of the p_{T} distribution for the first rapidity bin is slightly improved when using `TightBad` for leading p_{T} jets instead of using `LooseBad`.

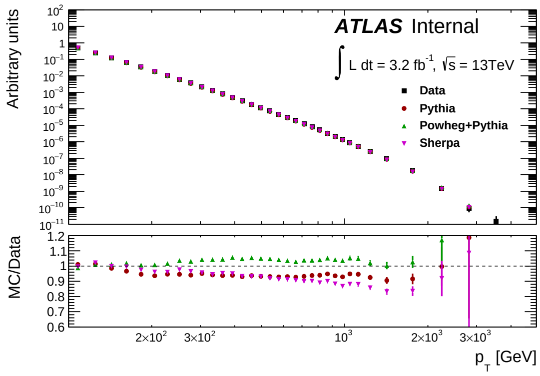
Figure 4.21: The p_{T} distribution for $|y| < 3.0$ at the detector level of leading p_{T} jets in data (black squares), Pythia (red circles), Powheg +Pythia (green upper triangles), and Sherpa (pink lower triangles), all normalised to the total area. The ratio of the MC simulations to the data is shown in the bottom panel.



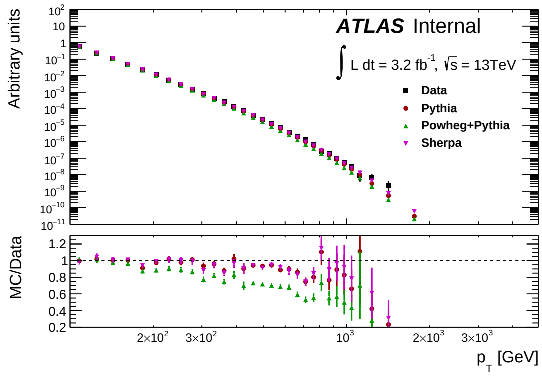
(a) inclusive



(b) first leading jet



(c) second leading jet

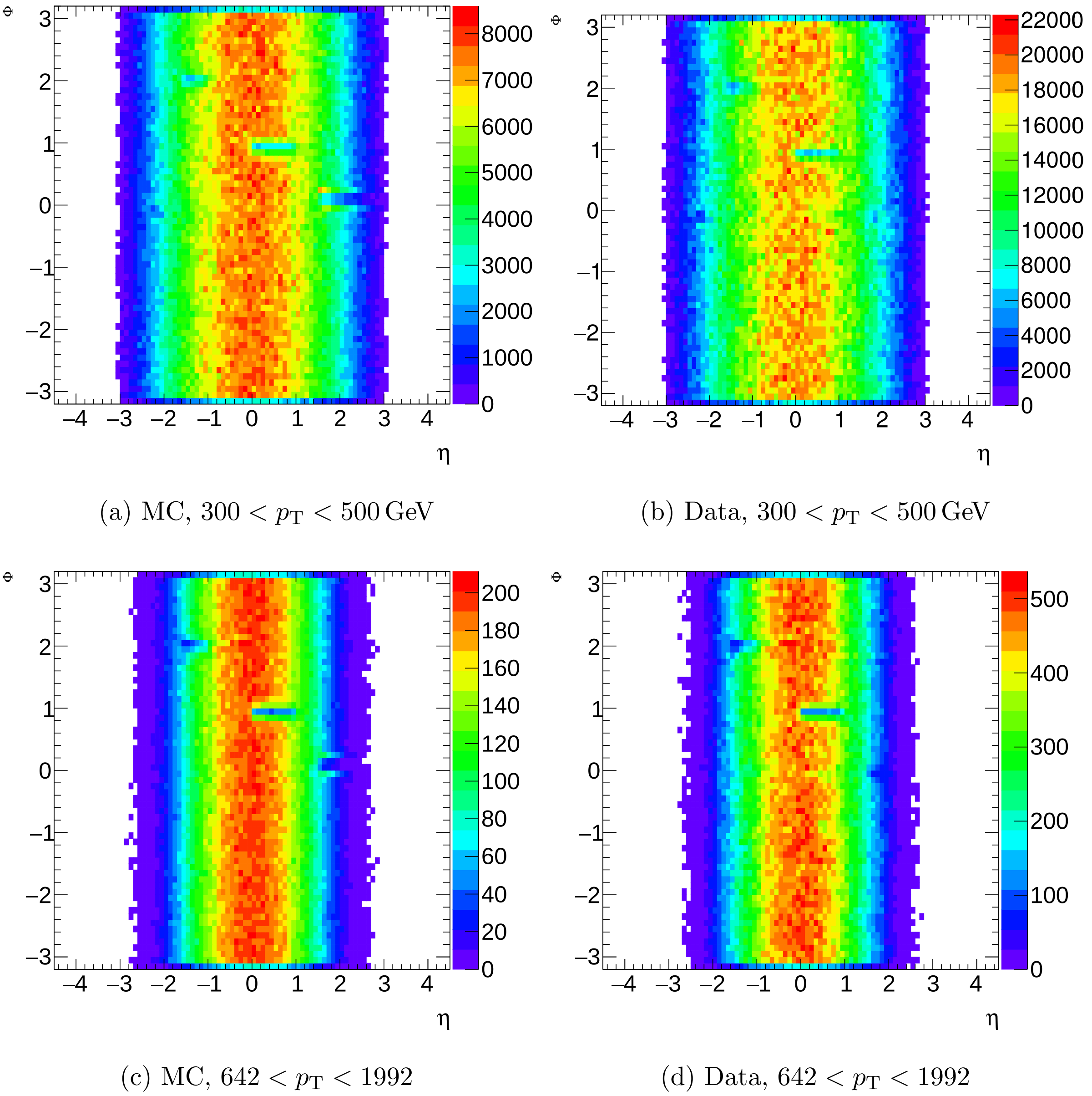


(d) third leading jet

4.7.6 Leading jet fractions

It is interesting to see the level of contribution of leading p_{T} jets to the inclusive jet cross-sections in data. Figure 4.26 shows that the cross-section is dominated by leading and second leading jets, as expected. The contribution of third leading jets is small, of the order of a few percents at low p_{T} and of a few per mille at high p_{T}. The contribution from fourth leading jets is negligible (and not show in the figure).

Figure 4.22: η–ϕ maps in both Pythia 8 (a and c) and data (b and d) at intermediate p_{T} (300–500 GeV, panels (a) and (b)) and at high p_{T} (642–1992 GeV, panels (c) and (d)). The MC shows an unpopulated η–ϕ region ($\eta \sim 2, \phi \sim 0$) which is not apparent in data.



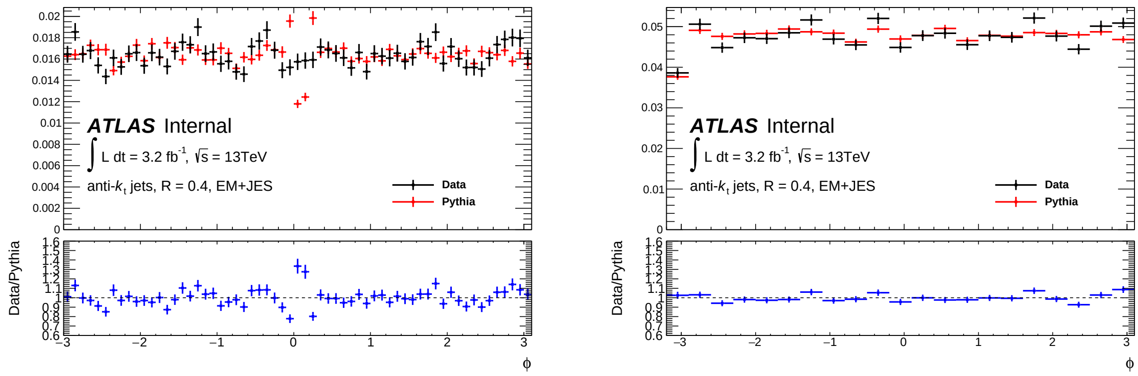
(a) MC, $300 < p_{\mathrm{T}} < 500$ GeV

(b) Data, $300 < p_{\mathrm{T}} < 500$ GeV

(c) MC, $642 < p_{\mathrm{T}} < 1992$

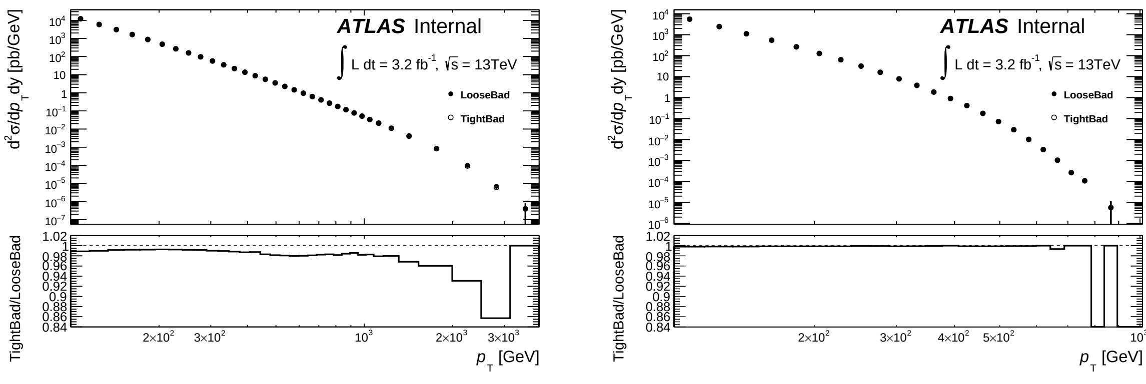
(d) Data, $642 < p_{\mathrm{T}} < 1992$

Figure 4.23: The jet ϕ distribution at the detector level in the $1.75 \leq \eta < 2.25$ region in data (black points) and PYTHIA 8 (red points) normalised to the total area is shown in (a). The ratio of the two (blue points) is shown in the bottom panel. (b) The same distribution but doubling the size of the ϕ bins.



(a) (b)

Figure 4.24: The jet p_{T} distribution at the detector level in data using the `LooseBad` (full circles) and the `TightBad` (empty circles) cleaning criteria for the leading p_{T} jets for the first (a) and last (b) $|y|$ bins of the measurement. The ratio of the two is shown in the bottom panels.



(a) $|y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 4.25: Ratio of the p_{T} distribution at the detector level in data to the one in MC simulation when using the `LooseBad` (full circles) and the `TightBad` (empty circles) cleaning criteria for the leading p_{T} jets for the first $|y|$ bin of the measurement.

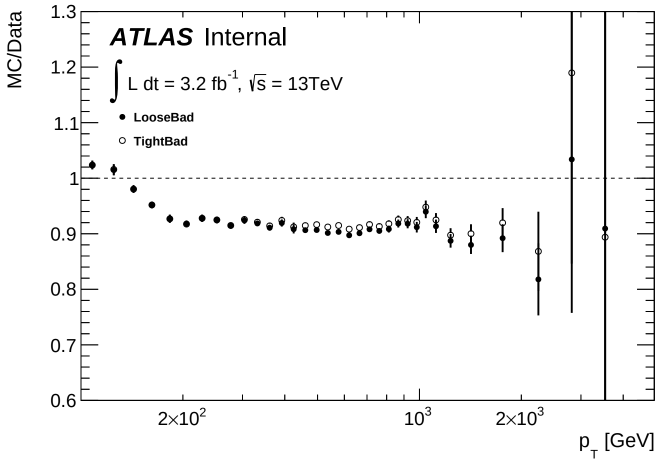
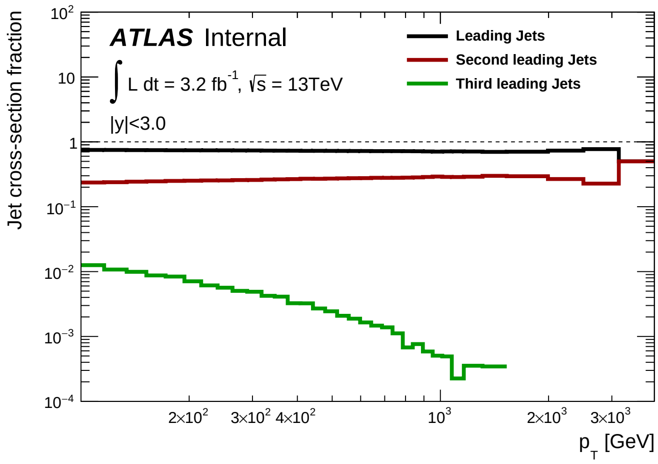


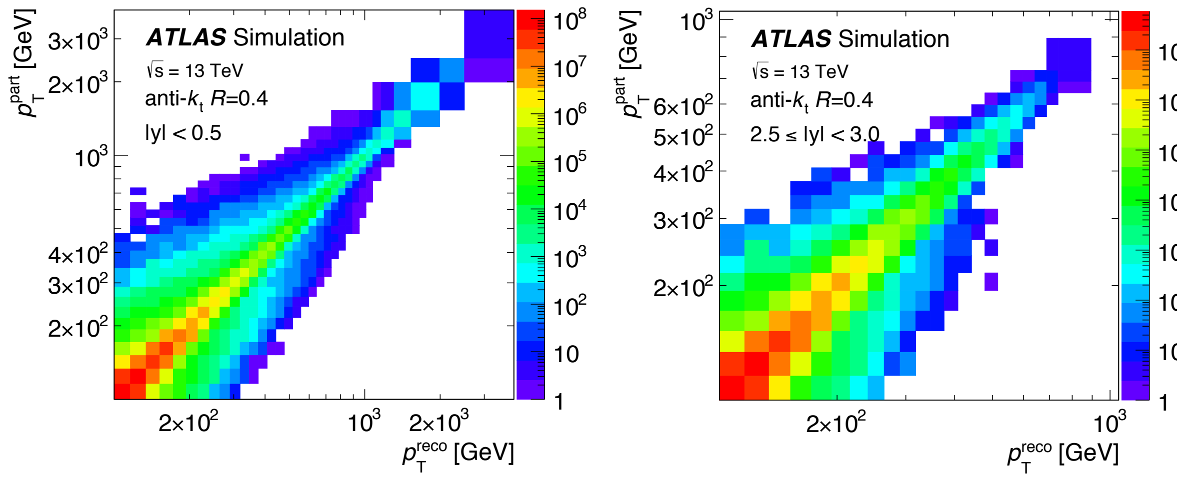
Figure 4.26: Contributed fraction to the p_{T} distribution at reconstructed level in data from leading jets (black), second leading jets (red), and third leading jets (green). Fourth leading jets are not shown since their contribution is negligible.



4.8 Unfolding of detector effects

The reconstructed jet spectra in data are corrected for detector inefficiencies and resolution effects to obtain inclusive jet cross-sections that refer to the stable particles entering the detector. The unfolding of the detector resolution in jet p_{T} is based on a modified Bayesian technique, the iterative dynamically stabilised (IDS) method [181]. This unfolding method uses a transfer matrix constructed using samples of simulated events, which describes the migrations of jets across p_{T} bins between particle level jets and reconstructed level jets. The transfer matrix is filled jet by jet by matching a particle jet with a reconstruction level jet, when both are closer to each other than to any other jet, lie within a radius of $R = 0.3$, have $p_{\mathrm{T}} > 75\,\mathrm{GeV}$ and belong to the same rapidity bin. It is shown in Figure 4.27 for two representive $|y|$ bins. The looser jet p_{T} requirement helps to reduce the unfolding bias (discussed below) at low p_{T}.

Figure 4.27: The transfer matrix for the most central and forward $|y|$ bins. The z-axis corresponds to the number of predicted events for a luminosity of $1\,\mathrm{fb}^{-1}$.



(a) (b)

The unfolding technique is performed in three steps, correcting for the matching impurity at the reconstruction level, the smearing of matched jets between p_{T} bins, and the matching inefficiency at the particle level,

$$N_i^{\mathrm{part}} = \sum_k N_k^{\mathrm{reco}} \cdot \mathcal{P}_k \cdot \mathcal{U}_{ik} \,/\, \mathcal{E}_i \,,$$

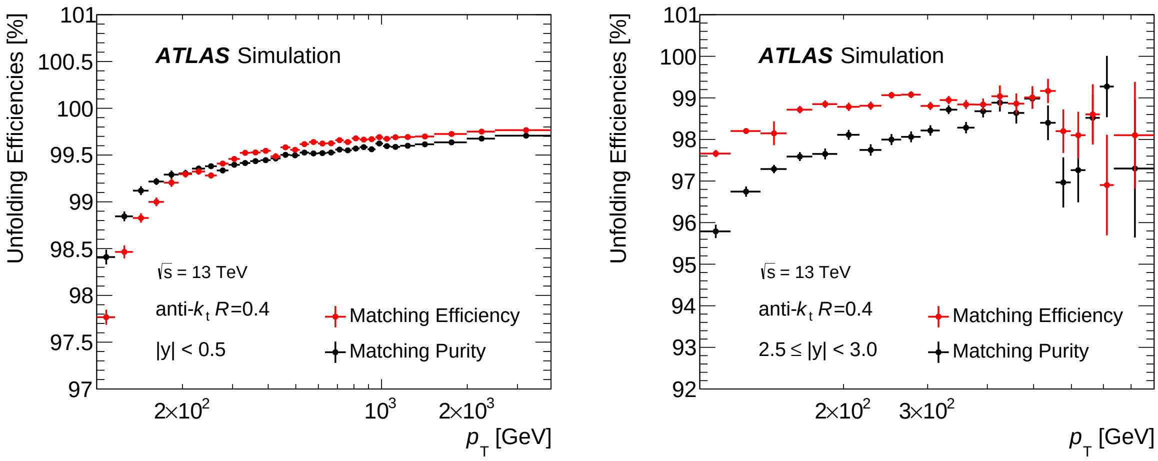
where i and k are the p_{T} bin indices of the jets at the particle and reconstruction levels and N^{part} and N^{reco} are the numbers of particle level and reconstruction level jets in a given bin, respectively. The symbols $\mathcal{P}$ and $\mathcal{E}$ denote respectively the matching purity and the matching efficiency. The symbol $\mathcal{U}$ denotes the unfolding matrix, where $\mathcal{U}_{ik}$ describes the probability for a jet at reconstruction level in p_{T} bin k to originate from the particle level in p_{T} bin i. The unfolding matrix is obtained from the transfer matrix (A) in the following way:

$$\mathcal{U}_{ik} = \frac{A_{ik}}{\sum_{l=1}^{n} A_{il}},$$

where the sum runs over all the p_{T} bins of the measurement.

The matching purity ($\mathcal{P}_k$) is defined as the fraction of reconstruction level jets that are matched to a particle level jet for a given p_{T} bin k. The matching efficiency ($\mathcal{E}_i$) is defined as the fraction of particle level jets that are matched to a reconstruction level jet for a given p_{T} bin i. If matched particle and reconstructed jets are in different rapidity bins then they are reassigned as being unmatched. In this way, the migrations across jet $|y|$ bins are effectively taken into account by bin-to-bin corrections. The jet matching efficiency and purity are shown in Figure 4.28. The jet matching efficiency is 97.8% (97.6%) at $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ for low (high) jet rapidity, and reaches 99.7% at high p_{T}. On the other hand, the jet matching purity is 98.4% (95.8%) at $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ for low (high) jet rapidity, and reaches (99.6%) at high p_{T}.

Figure 4.28: The matching efficiency and purity in the most central and forward $|y|$ bins.

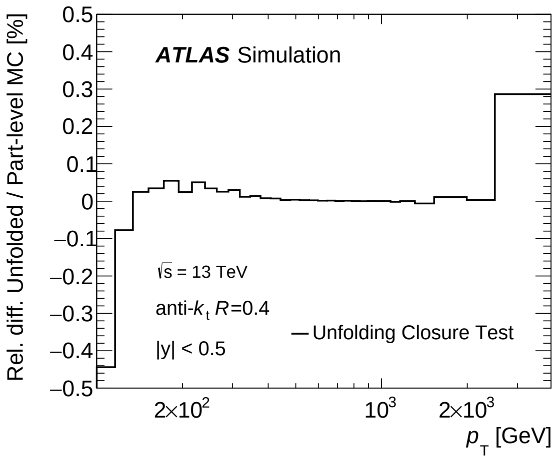


(a)

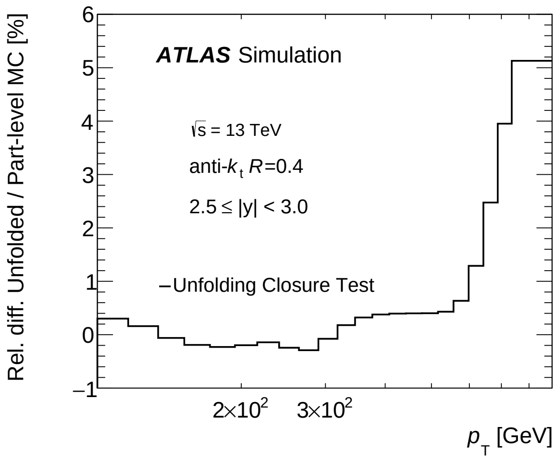
(b)

The unfolding matrix $\mathcal{U}$ depends on the details of the MC model, given that the transfer matrix is used to build it. This model improves when iterated, where the number of iterations is chosen such that the residual bias is within a tolerance of 1% in the bins with less than 10% statistical uncertainty. The residual bias is evaluated through a data-driven closure test [181, 182], in which the particle level spectrum in the MC simulation is reweighted to improve agreement between data (after applying the matching efficiency at reconstruction level) and reweighted MC events in the reconstruction level spectra. The weights are obtained by fitting with a fourth order logarithmic polynomial the ratio of the data to reconstruction level MC spectra. The reweighted reconstruction level MC spectra is unfolded using the nominal unfolding matrix, and the result is compared with the reweighted particle level spectrum. To obtain results that are statistically significant, this procedure is performed using a bootstrapping technique [183], in which pseudo-experiments are generated from the MC simulation by sampling each event with a weight taken from a Poisson distribution with a mean of one. Each pseudo-experiment therefore emphasises a unique subset of the MC simulation while maintaining statistical correlations between the nominal and varied samples. The resulting bias for each pseudo-experiment is evaluated, i.e., the difference between the unfolded result of the reweighted reconstruction level MC spectra and the expected result (i.e., the reweighted particle level spectra). The systematic uncertainty is then calculated as the average of all the biases obtained with the bootstraping technique. This systematic uncertainty accounts for the mismodelling of the shape of the p_{T} distribution of the MC simulation. Using one iteration, the uncertainty bias is of the order of a few per mille, except at high p_{T} ($\sim$ 1TeV) and high rapidity where it increases to 5%, as shown in Figure 4.29. The data unfolded using two iterations is in agreement (within the IDS bias uncertainty) with the data unfolded using one iteration. Since the bias due to shape differences between data and MC simulation is small after one iteration and since increasing the number of iterations enhances the statistical uncertainty, it was decided to use in this measurement a single iteration.

Figure 4.29: Relative difference 100 $\times$ (unfolded reconstruction level MC - particle level MC) / particle level MC, for the inclusive jet p_{T} cross-section for (a) the $|y| < 0.5$ bin and (b) the $2.5 \leq |y| < 3.0$ bin. The smoothing procedure described in Section 4.9 was applied to the 2.5–3.0 $|y|$ bin because of large statistical fluctuations due to the small number of events at high p_{T}.



(a)

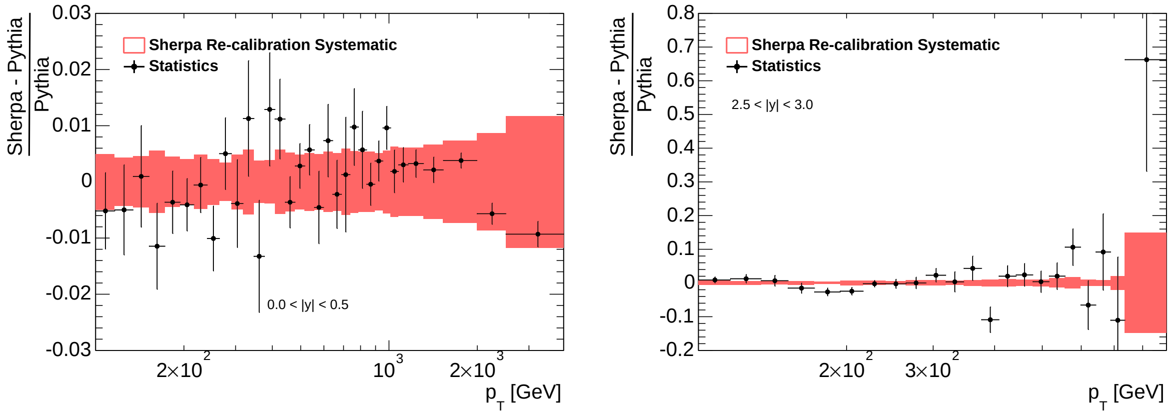


(b)

The dependence of the unfolding procedure on pile-up was studied. The data was unfolded using two MC simulation samples, consisting on Pythia 8 with and without pile-up reweighting technique being applied (see Section 4.3.1), and the smoothed (see next Section) relative difference between the two was taken as an extra systematic uncertainty. This uncertainty is smaller that 1%, except at high p_{T} and at large rapidity where it grows up to 1–2%.

The impact on the unfolding technique of the choice of the MC generator was investigated using the Sherpa sample. In order not to consider effects due to differences in the p_{T} response between Sherpa and Pythia 8, the response of Sherpa is corrected to have the same one as in Pythia 8. Since a perfect agreement between the response of the recalibrated Sherpa and the response of Pythia 8 was not achieved, a systematic uncertainty was assigned to account for differences in the responses at the $^1/_{1000}$ level. Sherpa was also reweighted by the ratio of the p_{T} distributions at truth level between Pythia 8 and Sherpa to avert effects due to shape differences. The relative difference between the unfolded Pythia spectra at reconstruction level using the recalibrated/reweighted Sherpa sample and the Pythia spectra at truth level is shown in Figure 4.30 for representative $|y|$ bins. Since the difference agrees in all $|y|$ bins and lies within the uncertainty on the recalibration of Sherpa, no uncertainty on the MC generator choice for the unfolding procedure is considered in this analysis.

Figure 4.30: Relative difference between unfolded PYTHIA 8 spectra at reconstruction level with recalibrated/reweighted SHERPA and PYTHIA spectra at truth level for (a) the first $|y|$ bin and (b) the last $|y|$ bin of the measurement. The red band shows the systematic uncertainty due to differences in the responses at $^1/_{1000}$ level.



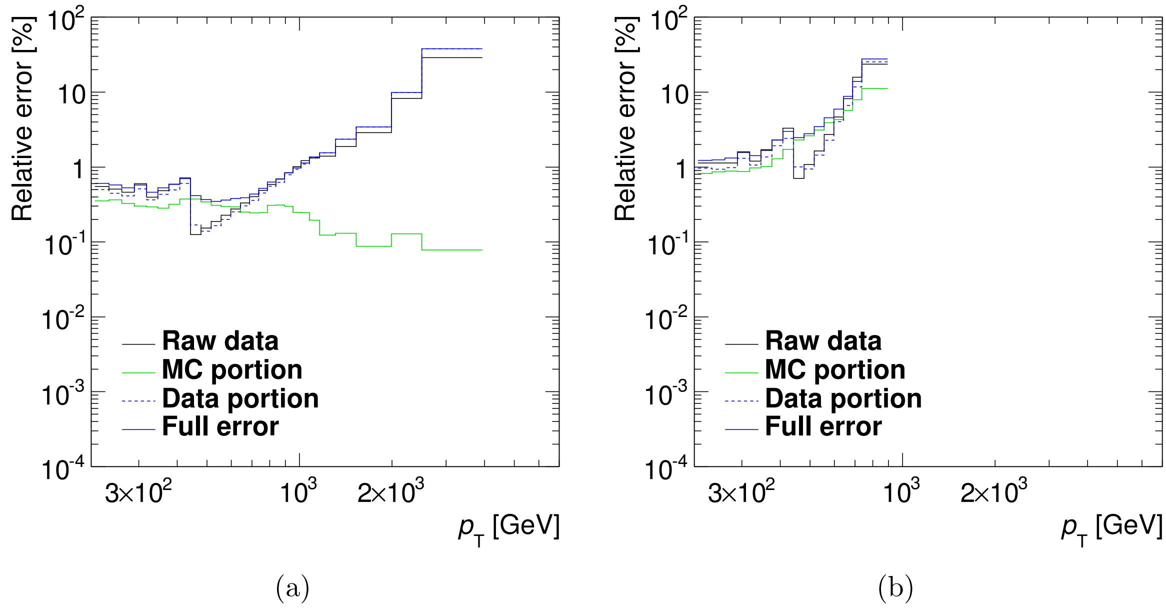
(a) (b)

4.9 Propagation of the uncertainties to the cross-sections

The statistical uncertainties are propagated through the unfolding procedure using a bootstrapping technique based on an ensemble of 1000 pseudo-experiments. Each pseudo-experiment is constructed by reweighting each event in data and simulation according to a Poisson distribution with expectation value equal to one. This procedure preserves the correlations between jets produced in the same event. The unfolding is performed for each pseudo-experiment and a covariance matrix is constructed for the cross-section in each $|y|$ bin. The total statistical uncertainty is obtained from the covariance matrix, where bin-to-bin correlations are also encoded. The separate contributions from the data and from the MC statistics are obtained through the same procedure by reweighting either the data or the simulated events. The total statistical uncertainty and its separate contributions are shown in Figure 4.31. The total statistical uncertainty is of the order of a few per mille (percent) at low p_{T} and reaches 40% (30%) at the last p_{T} bin for the first (last) $|y|$ of the measurement. The contribution to the total statistical uncertainty from the MC simulation is negligible, except at high p_{T} and large $|y|$.

All components of the JES uncertainty (see Section 4.5) are propagated through the unfolding procedure using pseudo-data (MC simulations) to avoid the impact of the larger statistical fluctuations in data. The jet p_{T} in pseudo-data is scaled up and down by one standard deviation of each component. This procedure takes into account the correlations between various phase-space regions. The resulting pseudo-data spectra are unfolded for detector effects using the nominal unfolding matrix. The difference between the nominal unfolded cross-section and the systematically shifted unfolded cross-section is taken as a systematic uncertainty. The jet energy scale is the dominant uncertainty for $p_{\mathrm{T}} < 2500\,\mathrm{GeV}$ ($p_{\mathrm{T}} < 700\,\mathrm{GeV}$) in the first (last) rapidity bin. In the complementary regions, the statistical uncertainty prevails. Figure 4.32 shows the break-down of the JES systematic uncertainties in representative jet rapidity bins. At low p_{T}, the dominant systematic uncertainties are the jet flavour uncertainties, while at high p_{T}, they are the

Figure 4.31: Statistical uncertainty on the unfolded spectra with contributions from both data and MC simulation samples for anti-k_t jets with $R = 0.4$ for (a) the first $|y|$ bin and (b) the last $|y|$ bin of the measurement. The black solid line shows the relative uncertainty due to data before unfolding ("Raw data"), the solid green line shows the relative uncertainty due to the available MC statistics ("MC portion"), the dash blue line shows the relative uncertainty due to data after unfolding ("Data portion"), and finally the solid blue line is the total relative statistical uncertainty ("Full error").



E/P and the punch-through (jet flavour and $\eta-$intercalibration) uncertainties in the first (last) rapidity bin.

The uncertainty in the JER is the second largest individual source of systematic uncertainty. There are 11 components, some of which can involve a JER degradation in part of p_T–η phase-space and a JER improvement in the complementary part, which allows (anti-)correlations to be accounted for. The effect of each of the components ($\Delta\sigma$) is evaluated by smearing the energy of the reconstructed jets by a factor σ_{smear}, calculated as:

$$\sigma^2_{\mathrm{smear}} + \sigma^2_{\mathrm{nominal}} = (\sigma_{\mathrm{nominal}} + \Delta\sigma)^2 \, ,$$

where $\sigma_{\mathrm{nominal}}$ is the nominal fractional jet energy resolution. In another words, the smearing consists of a degradation of the jet energy resolution by one standard deviation of its uncertainty component. The degradation of the JER is achieved by smearing the reconstructed jets in the relevant phase-space region in the MC simulation used as pseudo-data. On the other hand, an effective improvement of the JER is achieved by smearing the energy of the jets in the MC simulation used in constructing the transfer matrix. The difference between the modified spectrum unfolded with the systematically varied transfer matrix to the nominal spectrum unfolded with the nominal transfer matrix is taken as a systematic uncertainty.

The effect of the uncertainty in the JAR is evaluated by smearing the η and ϕ of the reconstructred jets in a similar way as done for the JER, and it was found to be negligible.

The uncertainty on the jet cleaning procedure (introduced in Section 4.6) is also propagated through the unfolding procedure using pseudo-data. This systematic uncertainty, for the first and last $|y|$ bin of the measurement, is shown in Figure 4.33.

The uncertainty in the luminosity measurement of 2.1% is propagated as being cor-

Figure 4.32: Panels (a) and (b) ((c) and (d)) show the relative JES systematic uncertainties for the inclusive jet cross-sections in the first (last) jet rapidity bin.

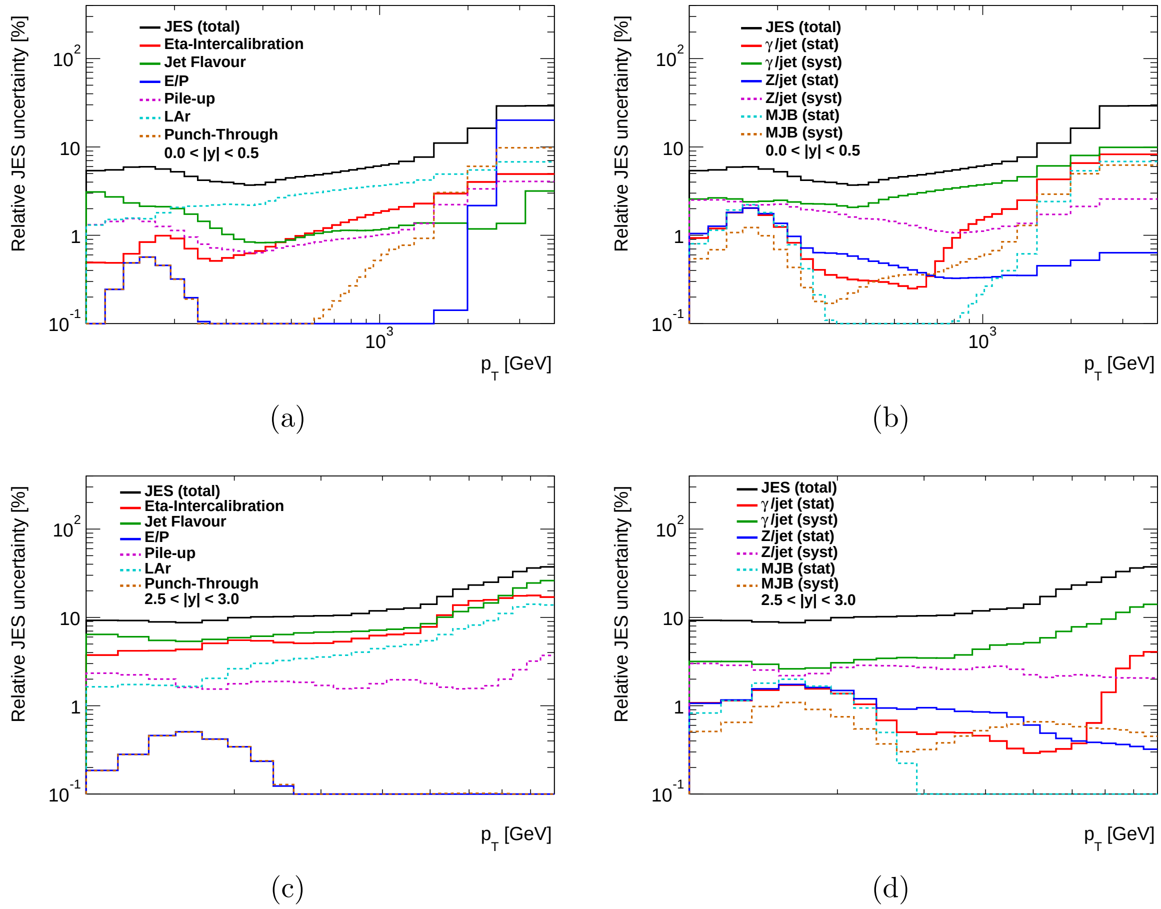
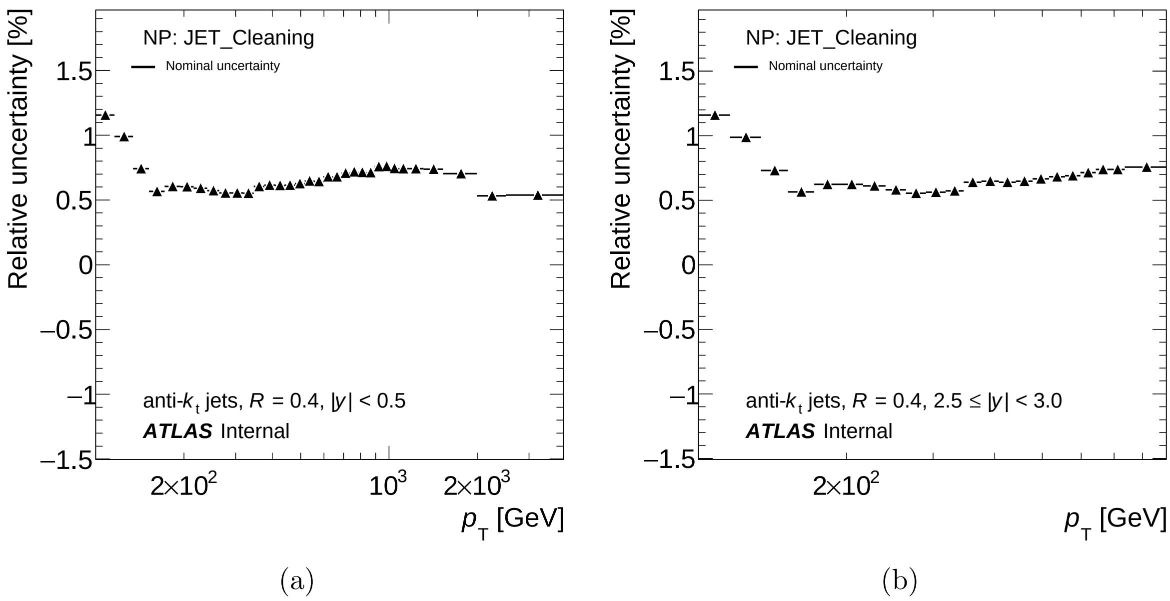


Figure 4.33: Systematic jet cleaning uncertainty for the first (a) and last (b) $|y|$ bin of the measurement.

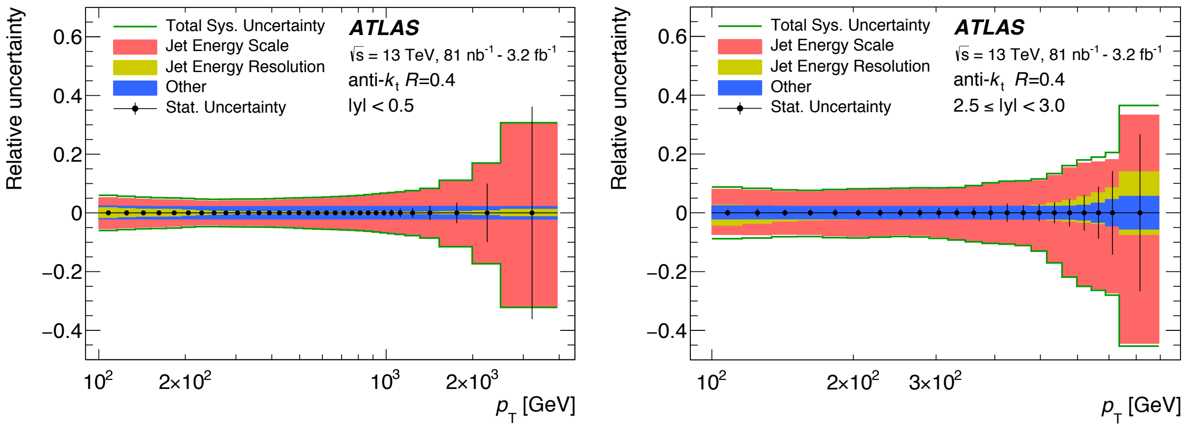


related across all measurement bins.

In order to assess the statistical precision of the systematic uncertainty estimates, each component is re-evaluated using a set of pseudo-experiments. The statistical fluctuations of the systematic uncertainty estimates are minimised using a smoothing procedure. To achieve this, for each component, the p_{T} bins are combined until the propagated uncertainty value in the bin has a Gaussian statistical significance larger than two standard deviations. A Gaussian kernel smoothing [139] is used to obtain the values in the original fine bins.

Figure 4.34 shows the individual components of the systematic uncertainties added in quadrature in representative phase-space regions. In the central (forward) region the total uncertainty is about 5% (8%) at medium p_{T} of 300–600 GeV. The uncertainty increases towards both lower and higher pT reaching 6% (10%) at low p_{T} and 30% ([-45%,+40%]) at high p_{T}.

Figure 4.34: Relative systematic uncertainty for the inclusive jet cross-section as a function of the jet p_{T} for the first and last rapidity bins ((a) and (b) respectively). The individual uncertainties are shown in different colours: the jet energy scale, jet energy resolution and the other uncertainties (jet cleaning, luminosity and unfolding bias). The total systematic uncertainty, calculated by adding the individual uncertainties in quadrature, is shown as a green line. The statistical uncertainty is shown as vertical black lines.



(a) (b)

4.10 Theoretical predictions

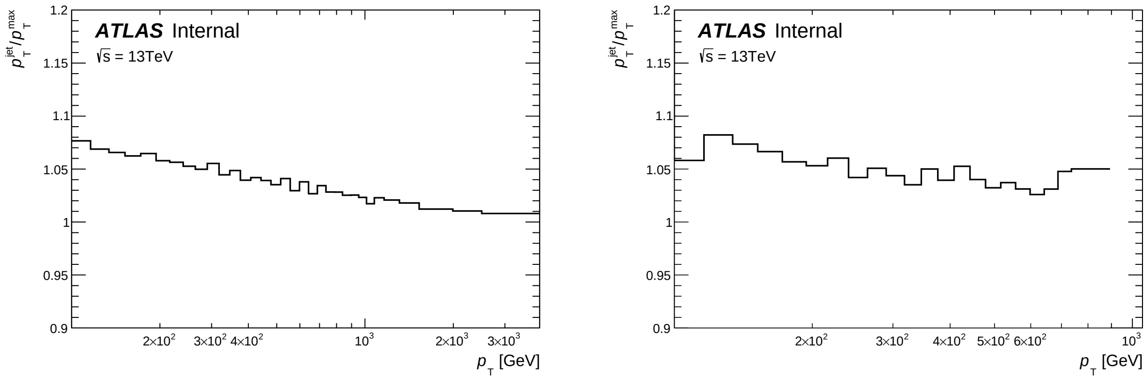
Theoretical predictions of the cross-sections are obtained using NLO and NNLO pQCD calculations with corrections for non-perturbative and electroweak effects.

4.10.1 Next-to-leading-order pQCD calculations

The NLO pQCD predictions are calculated using NLOJET++ 4.1.3 [184] interfaced to APPLGRID [185] for fast and flexible calculations with various PDF sets and various values of the renormalisation and factorisation scales. The inclusive jet cross-section prediction is calculated using $p_{\mathrm{T}}^{\mathrm{max}}$, the transverse momentum of the leading jet in the event, as the renormalisation scale (μ_{R}) and the factorisation scale (μ_{F}). An alternative scale choice, $\mu_{\mathrm{R}} = \mu_{\mathrm{F}} = p_{\mathrm{T}}^{\mathrm{jet}}$, the p_{T} of each individual jet that enters the cross-section calculation, is also considered. This scale choice is proposed in Ref. [186]. Both scale choices were used in the previous ATLAS analysis at $\sqrt{s} = 8$ TeV [165]. The predictions are calculated using several PDFs provided by the LHAPDF6 [187] library: the NLO CT14 [188], MMHT 2014 [189], NNPDF 3.0 [190], and HERAPDF 2.0 [191] sets, and the NNLO ABMP16 [192] set. The value of the strong coupling constant (α_{s}) is taken from the corresponding PDF set.

The main uncertainties in the NLO predictions come from uncertainties associated with the PDFs, the choice of renormalisation and factorisation scales, and the uncertainty in the value of α_{s}. PDF uncertainties are defined at the 68% confidence level and propagated through the calculations following the prescription given for each PDF set, as recommended by the PDF4LHC group for PDF-sensitive analyses [193]. Calculations are redone with varied renormalisation and factorisation scales to estimate the uncertainty due to missing higher-order terms in the pQCD expansion. The nominal scales are independently varied up or down by a factor of two in both directions excluding opposite variations of μ_{R} and μ_{F}. The envelope of the resulting variations is taken as the scale uncertainty. The difference between the predictions obtained with the $p_{\mathrm{T}}^{\mathrm{jet}}$ and the $p_{\mathrm{T}}^{\mathrm{max}}$ scale choice, shown by the ratio of the two in Figure 4.35 for representative phase-space regions, is treated as an additional uncertainty. This uncertainty is 8% at low p_{T} ($p_{\mathrm{T}} \sim 100$ GeV), decreases with p_{T} and is negligible at high p_{T}. The uncertainty from α_{s} is evaluated by calculating the cross-sections using two PDF sets that differ only in the value of α_{s} used and then scaling the cross-section difference corresponding to an α_{s} uncertainty $\Delta\alpha_{\mathrm{s}} = 0.0015$ as recommended in Ref. [193].

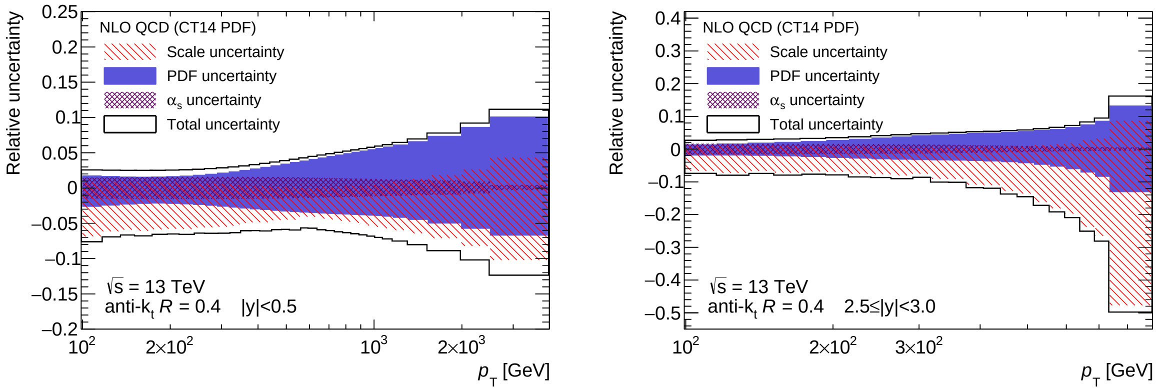
Figure 4.35: Ratio of the NLO QCD prediction calculated using the CT14 PDF set using the scale choice $p_{\mathrm{T}}^{\mathrm{jet}}$ to the one calculated with the scale choice $p_{\mathrm{T}}^{\mathrm{max}}$, for the first (a) and last (b) $|y|$ bin of the measurement.



(a) $|y| < 0.5$ (b) $2.5 < |y| < 3.0$

All the uncertainties in the NLO QCD cross-section predictions, except the one due to the difference between $p_{\mathrm{T}}^{\mathrm{jet}}$ and $p_{\mathrm{T}}^{\mathrm{max}}$ QCD scale choices, obtained with the CT14 PDF set are shown in Figure 4.36 for representative phase-space regions. The uncertainty due to the choice of renormalisation and factorisation scale is dominant in most phase-space regions, rising from 10% at about $p_{\mathrm{T}} = 100$ GeV in the central rapidity bin to about 50% in the highest p_{T} bins in the most forward rapidity region. The PDF uncertainties vary from 2% to 12% depending on the jet p_{T} and rapidity. The contribution from the α_{s} uncertainty is about 2% at low p_{T} and negligible for the highest p_{T} bin in each rapidity range.

Figure 4.36: Relative NLO QCD uncertainties in the inclusive jet cross-sections calculated using the CT14 PDF set. Panels (a) and (b) correspond respectively to the first and last $|y|$ bins for the measurement. The uncertainties due to the renormalisation and factorisation scale, the α_{s}, the PDF and the total uncertainty are shown. The total uncertainty, calculated by adding the individual uncertainties in quadrature, is shown as a black line.



(a) (b)

4.10.2 Non-perturbative corrections

Non-perturbative corrections $\mathrm{k_{NP}}$ are applied to the parton-level cross-sections. The correction factors are calculated using LO MC event generators, as the bin-by-bin ratio of the nominal particle-level MC cross-sections to the MC cross-section derived from jets reconstructed from the partons remaining after showering, when the modelling of hadronisation and the underlying event are switched off. The correction factors are evaluated using two event generators, Pythia 8 and Herwig++ with different tunes and PDF sets, which are listed in Table 4.4. The main difference between estimated non-perturbative corrections arises from using Pythia 8 or Herwig++. The baseline correction is taken from Pythia 8 using the A14 tune with the NNPDF2.3 LO PDF set, which corresponds to the tune yielding the closest result to Herwig++. The envelope of all corrections is considered as a systematic uncertainty.

Table 4.4: Summary of the soft-physics model tunes used for the evaluation of the non-perturbative corrections for each event generator and PDF set.

	CTEQ6L1 [194]	CTEQ6L1	MSTW2008LO	CT10	NNPDF2.3LO	NNPDF2.3LO	CTEQ6L1
Pythia 8	4C [195]	AU2	A14	AU2	MONASH	A14	A14
Herwig++	UE-EE-5 [196, 197]	UE-EE-4 [196, 197]	UE-EE-5				

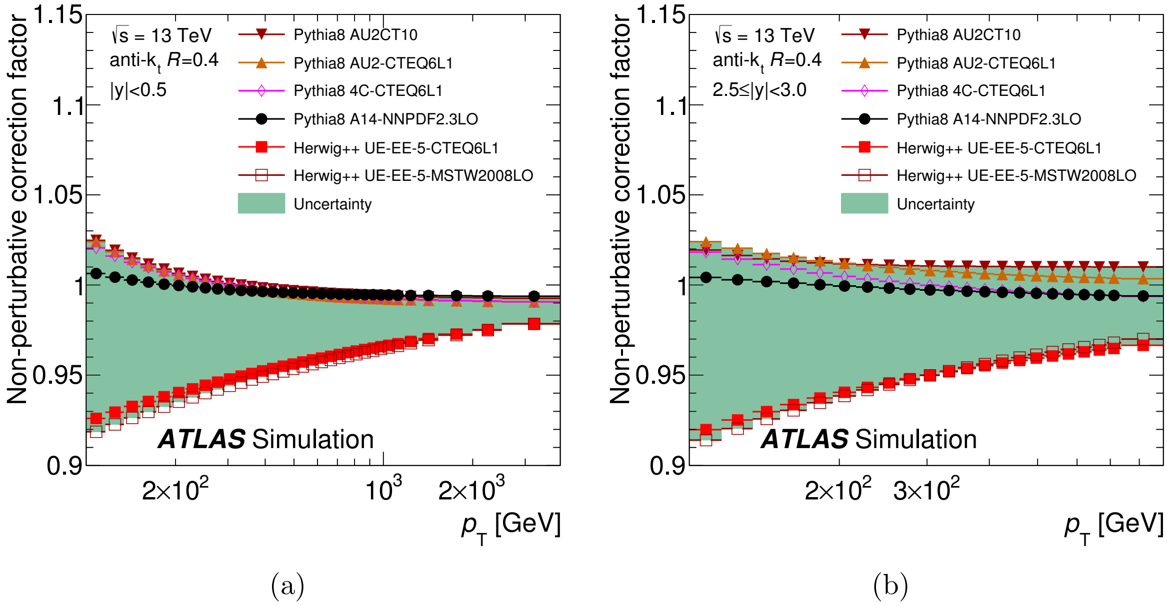
In order to avoid statistical fluctuations, the following function is used to fit the non-perturbative correction factors as a function of the jet p_{T} for each rapidity bin of the measurement:

$$\mathrm{k_{NP}}(p_{\mathrm{T}}) = a + \frac{b}{p_{\mathrm{T}}^{c}},$$

where a, b and c are free fit parameters. The fit is performed in the range $25 \leq p_{\mathrm{T}} < 3500$ GeV. The final non-perturbative corrections correspond to the fitted functions calculated at the central values of the p_{T} bins. In most cases, the difference between the fitted function and the points lies within 1%. For a few cases, in particular for Herwig++, the difference can go up to 2%.

The correction factors for a set of representative event generators and tunes are shown in Figure 4.37 in illustrative $|y|$ bins as a function of $p_{\rm T}$. The values of the correction are in the range 0.92–1.03 at low $p_{\rm T}$ and 0.98–0.99 (0.97–1.01) at high $p_{\rm T}$ for the first (last) rapidity bin.

Figure 4.37: Non-perturbative correction factors for the inclusive jet NLO pQCD prediction as a function of jet $p_{\rm T}$ for the first and last rapidity bin, (a) and (b) respectively. The corrections are derived using Pythia 8 with the A14 tune with the NNPDF2.3 LO PDF set. The envelope of all MC configuration variations is shown as a band, and considered as a systematic uncertainty.



4.10.3 Electroweak corrections

The NLO pQCD predictions are corrected for the effects of γ and $W^{\pm}/Z$ interactions at tree and one-loop level. They are derived using an NLO calculation of electroweak (EW) contributions to the LO pQCD process. The correction is defined as the ratio of a $2 \to 2$ calculation including tree-level effects of order $\alpha_{\rm s}^2$, α^2, and $\alpha_{\rm s}\alpha$ (from interference of QCD and EW diagrams), plus weak loop corrections of order $\alpha_{\rm s}^2\alpha$ to the LO QCD $2 \to 2$ calculation.

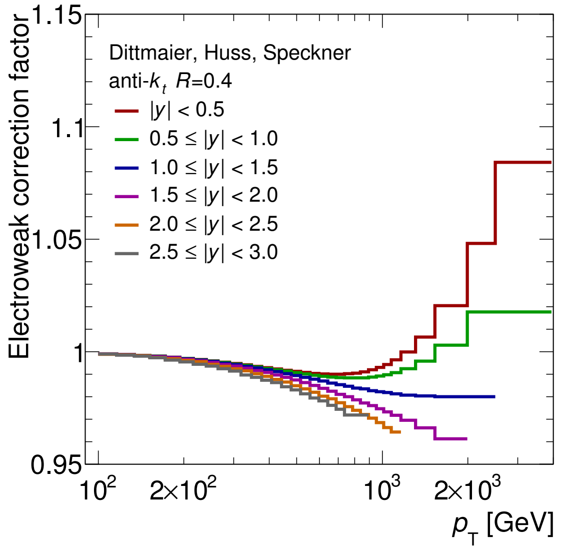
The correction factors are derived in the phase space considered for the measurement presented here and were provided by the authors of Ref. [198]. No uncertainty associated with these corrections is presently estimated.

The electroweak correction factors as a function of the jet $p_{\rm T}$ in bins of $|y|$ are shown in Figure 4.38. The electroweak correction is small for low jet transverse momenta, reaches 8% at the highest $p_{\rm T}$ (3 TeV) for the central $|y|$ bin, and is less than 4% for the rest of the $|y|$ bins.

4.10.4 Next-to-next-to-leading-order pQCD calculations

The recently developed NNLO pQCD predictions were provided by the authors of Refs. [168, 169] using the NNLOJET program and the MMHT 2014 NNLO PDF set for two different choices of the $\mu_{\rm R}$ and $\mu_{\rm F}$ scales, respectively $p_{\rm T}^{\rm jet}$ and $p_{\rm T}^{\rm max}$. The non-perturbative and

Figure 4.38: Electroweak correction factors as a function of the jet p_{T} for all $|y|$ bins.



electroweak corrections described in Sections 4.10.2 and 4.10.3, respectively, are applied to the predictions. In addition to the statistical uncertainties on the calculations, which are larger for higher p_{T} and high rapidities, two sources of uncertainty are considered in this NNLO calculation: the scale uncertainty and the systematic uncertainty in the non-perturbative correction. To obtain the scale uncertainty, both scales (renormalisation and factorisation) are varied simultaneously by a factor of 0.5 or 2.[5] If both variations yield changes with the same sign, the scale uncertainty is obtained from the larger change. No PDF uncertainty is considered for the NNLO pQCD predictions due to computing time limitations, since NNLOJET is yet not interfaced to APPLGRID.

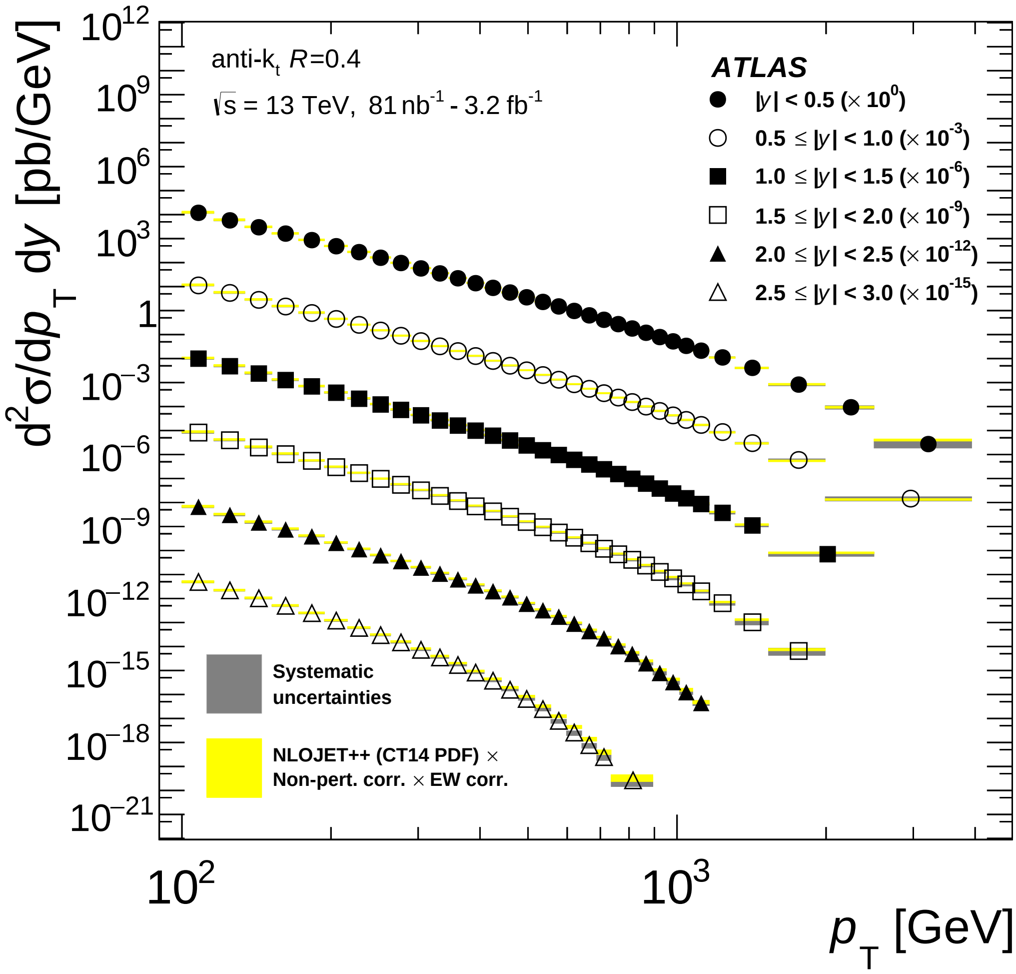
4.11 Results

The measured double-differential inclusive jet cross-sections are shown in Figure 4.39 as a function of p_{T} for the six jet rapidity bins. The measurements cover the jet p_{T} range from 100 GeV to 3.5 TeV for $|y| < 3.0$, thus reaching significantly higher than the previous ATLAS measurements [165, 199, 200]. The NLO pQCD predictions using the CT14 PDF set, corrected for non-perturbative and electroweak effects, are also shown in the figure.

The ratios of the NLO pQCD predictions to the measured inclusive jet cross-sections as a function of p_{T} in the six jet rapidity bins are shown in Figure 4.40 for the CT14, MMHT 2014, NNPDF 3.0, HERAPDF 2.0 and ABMP16 PDF sets. No significant deviation of the data points from the predictions is seen; the NLO pQCD predictions and data agree within uncertainties. This behaviour is compatible with the results of the comparison between data and the pQCD predictions in the previous ATLAS measurement at $\sqrt{s} = 8$ TeV [165]. In the forward region ($|y| > 2$) there is a tendency for the NLO pQCD prediction using the CT14, MMHT 2014, NNPDF 3.0, and AMBP16 PDF sets to overestimate the measured

[5] A different approach to estimate the scale uncertainty was used for NNLO due to computing time limitations. At NLO the simultaneous variations are not always the dominant ones, although they are at high p_{T}.

Figure 4.39: Inclusive jet cross-sections as a function of p_T and $|y|$. The statistical uncertainties are smaller than the size of the symbols used to plot the cross-section values. The dark gray shaded areas indicate the experimental systematic uncertainties. The data are compared to NLO pQCD predictions calculated using NLOJET++ with $p_\mathrm{T}^\mathrm{max}$ as the QCD scale and the CT14 NLO PDF set, to which non-perturbative and electroweak corrections are applied. The yellow shaded areas indicate the predictions with their uncertainties. At low and intermediate p_T bins the experimental systematic uncertainties are comparable to the theory uncertainties (drawn on top) and therefore are barely visible.



cross-section in the high p_T range, although the difference from data does not exceed significantly the range covered by the experimental and theoretical uncertainties.

The ratios of the NNLO pQCD predictions to the measured inclusive jet cross-sections as a function of p_T in the six jet rapidity bins are shown in Figure 4.41 for the two different scale choices, p_T^{jet} and p_T^{max}, together with the NLO case for comparison. When using p_T^{jet} as a scale, the NNLO pQCD predictions describe the data within uncertainties, with the exception of the forward ($|y| > 2$) high p_T range where it tends to overestimate the measured cross-section. The predictions using p_T^{max} as the scale overestimate the measured cross-sections.

The NLO pQCD predictions, corrected for non-perturbative and electroweak effects, are quantitatively compared to the measurement using the method described in Ref. [200]. The χ^2 value and the corresponding observed p-value (p_{obs}) are computed taking into account the asymmetries and the (anti-)correlations of the experimental and theoretical uncertainties. The individual experimental and theoretical uncertainty components are assumed to be uncorrelated among each other and fully correlated across the p_T and $|y|$ bins. The correlations of the statistical uncertainties across different phase-space regions are taken into account using covariance matrices derived from 1000 pseudo-experiments obtained by fluctuating the data and the MC simulation (see Section 4.9).

For the theoretical prediction and separately for each scale choice (p_T^{max} and p_T^{jet}), the uncertainties related to the scale variations, the PDF eigenvectors, the non-perturbative corrections and the strong coupling constant are treated as additional uncertainty components. In the case of the NNPDF 3.0 PDF set, the replicas [190] are used to evaluate a covariance matrix, from which the eigenvectors are then determined.

Table 4.5 shows the summary of the p_{obs} values for each individual rapidity bin of the measurement. Table 4.6 reports the results obtained from a global fit to all the p_T and rapidity bins of the measurement. Given that in this case the p_{obs} values are very small, the results are directly presented in terms of the χ^2 per degree of freedom (dof).

Fair agreement is seen (with p-values in the percent range) when considering jet cross-sections in individual jet rapidity bins treated independently, with some tension present in the 1.5–2.5 rapidity region. Comparable results are obtained for PDF sets determined with similar data. Strong tension between data and theory is observed when simultaneously considering data points from all jet transverse momentum and rapidity regions (Table 4.6), a behaviour already observed in the previous ATLAS measurement at $\sqrt{s} = 8$ TeV [165].

Consideration of all data points together requires a good understanding of the correlations of the experimental and theoretical systematic uncertainties in jet p_T and rapidity. Although the correlations of most uncertainties are generally well known, the systematic uncertainties that are based on simple comparisons between two options (two-point uncertainties) are not well defined. This is the case for instance for the in situ multijet balance uncertainties due to different fragmentation models and the theoretical uncertainty related to the alternative scale choice. In these cases, alternative decorrelation scenarios can in principle be used instead of the default full correlation model. To this end, systematic uncertainties are split into sub-components whose size varies with jet rapidity and p_T, keeping their sum in quadrature equal to the original uncertainty.

Reference [165] presents a detailed discussion on reasonable alternative correlation scenarios. Following the proposal therein, decorrelation scenarios were applied simultaneously to the largest sources of two-point experimental uncertainties (the JES flavour response, the JES multijet p_T-balance fragmentation, and the pile-up energy density in jets) as well as the theoretical uncertainties (the scale variations, the alternative scale choice

Figure 4.40: Comparison of the measured inclusive jet cross-sections and the NLO pQCD predictions shown as the ratios of predictions to the measured cross-sections. The ratios are shown as a function of the jet p_{T} in six $|y|$ bins. The predictions are calculated using NLOJET++ with $p_{\mathrm{T}}^{\mathrm{max}}$ as the QCD scale and five different PDF sets: CT14, MMHT 2014, NNPDF 3.0, HERAPDF 2.0 and ABMP16. Non-perturbative and electroweak corrections are applied to the theoretical predictions. The first three PDF sets are shown on top and the last two together with CT14 are shown in the bottom. The CT14 case is repeated to serve as a reference for comparison. The uncertainties of the predictions, shown by the coloured lines, include all the uncertainties discussed in Section 4.10. The grey bands show the total data uncertainty including both the systematic (JES, JER, unfolding, jet cleaning, luminosity) and statistical uncertainties.

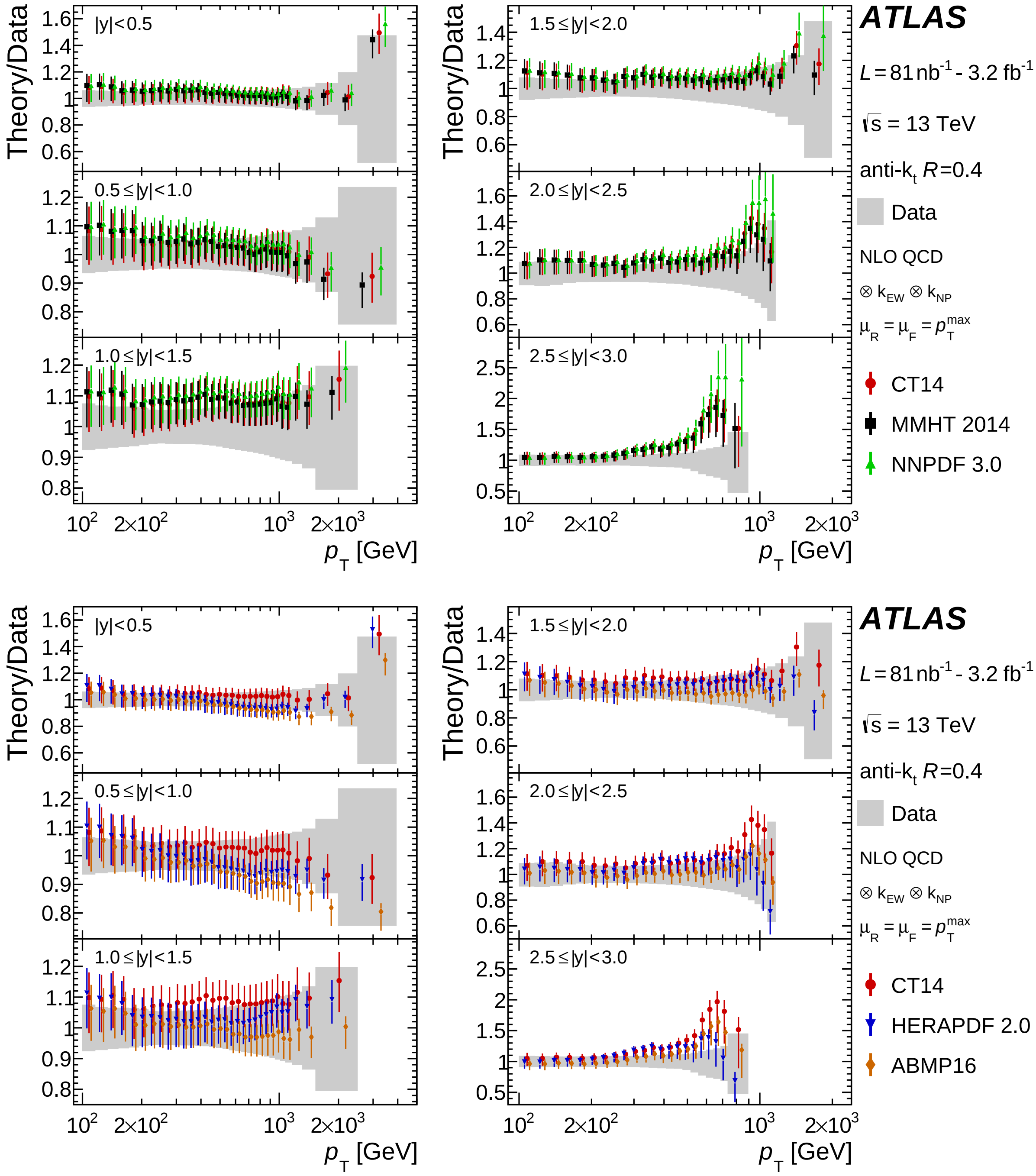
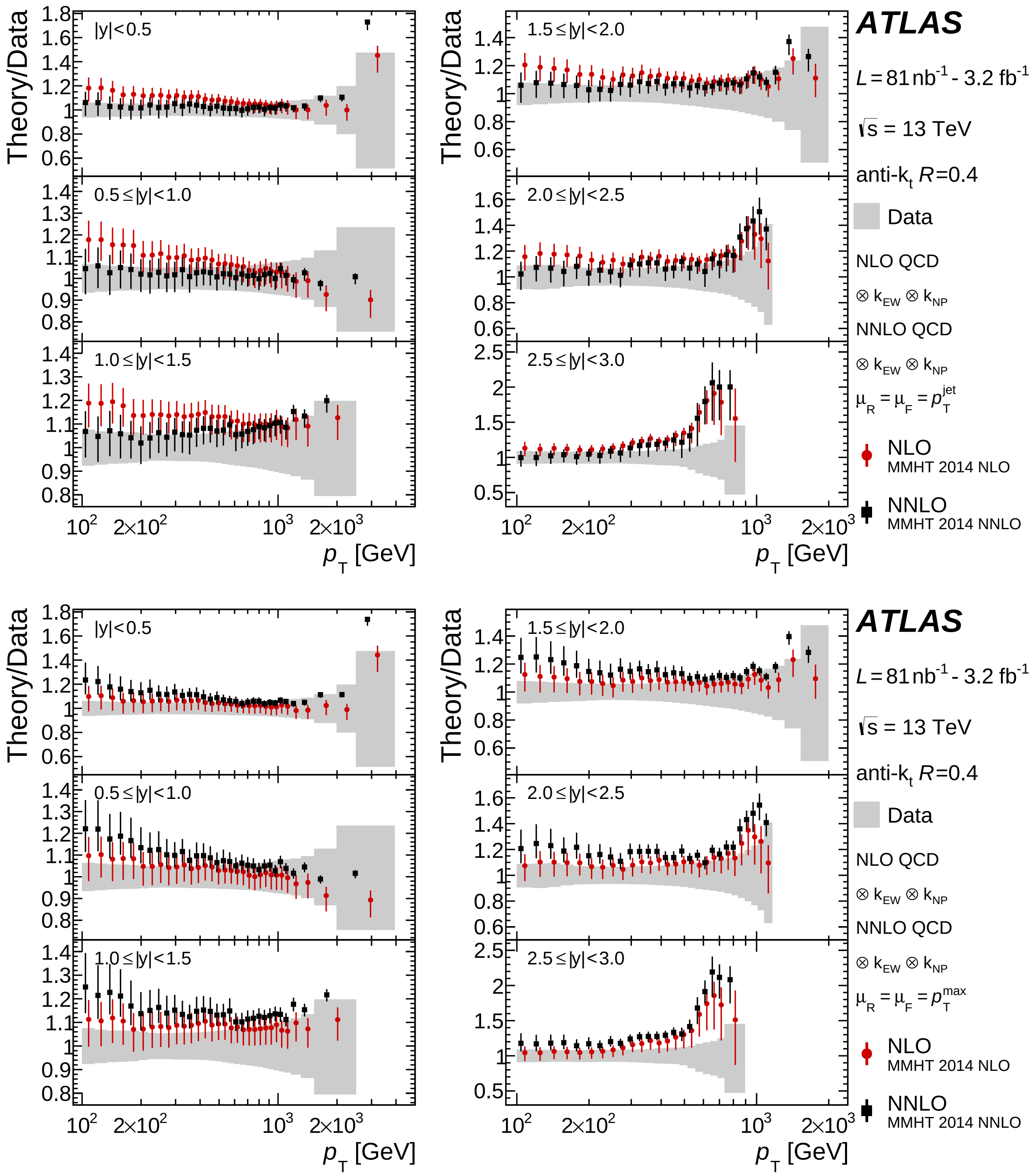


Figure 4.41: Ratios of the NLO and NNLO pQCD predictions to the measured inclusive jet cross-sections, shown as a function of the jet p_{T} in six $|y|$ bins. The NLO predictions are calculated using NLOJET++ with the MMHT 2014 NLO PDF set. The NNLO predictions are provided by the authors of Refs. [168, 169] using NNLOJET with the MMHT 2014 NNLO PDF set. The predictions calculated using $p_{\mathrm{T}}^{\mathrm{jet}}$ ($p_{\mathrm{T}}^{\mathrm{max}}$) as the QCD scale are shown in the top (bottom). Non-perturbative and electroweak corrections are applied to the predictions. The NLO and NNLO uncertainties are shown by the coloured lines, including all the uncertainties discussed in Section 4.10. The grey bands show the total data uncertainty including both the systematic (JES, JER, unfolding, jet cleaning, luminosity) and statistical uncertainties.



and the non-perturbative corrections) using the splitting options that yielded the largest χ^2 reduction for each single component in Ref. [165]. The χ^2 using the CT14 PDF set and the $p_{\mathrm{T}}^{\mathrm{max}}$ scale choice is found to be reduced by 58 units ($\chi^2/\mathrm{dof} = {}^{361}/_{177}$) compared to the nominal configuration, but the corresponding p-value is still $\ll 10^{-3}$, in agreement with the conclusions of the previous ATLAS measurement at $\sqrt{s} = 8\,\mathrm{TeV}$ [165]. It should be noted that, for some uncertainty components, the notion of a standard deviation (i.e., the size of the uncertainty itself) is not well defined. Furthermore, the phase-space dependence of the size of the uncertainties is a key ingredient in the χ^2 evaluation, and for some sources, it is particularly difficult to determine in detail. Both could explain part of the observed tension between the measurement and the theory. The remaining tension could be due either to the break-down of the assumptions made in the treatment of two-point systematic uncertainty components, or to an incomplete theoretical description, such as missing higher-order corrections.

Table 4.5: Summary of observed p-values (p_{obs}) from the comparison of the inclusive jet cross-section and the NLO pQCD prediction corrected for non-perturbative and electroweak effects for various PDF sets, separately for the two scale choices and for each rapidity bin of the measurement.

	p_{obs}				
Rapidity ranges	CT14	MMHT 2014	NNPDF 3.0	HERAPDF 2.0	ABMP16
$p_{\mathrm{T}}^{\mathrm{max}}$					
$\|y\| < 0.5$	67%	65%	62%	31%	50%
$0.5 \leq \|y\| < 1.0$	5.8%	6.3%	6.0%	3.0%	2.0%
$1.0 \leq \|y\| < 1.5$	65%	61%	67%	50%	55%
$1.5 \leq \|y\| < 2.0$	0.7%	0.8%	0.8%	0.1%	0.4%
$2.0 \leq \|y\| < 2.5$	2.3%	2.3%	2.8%	0.7%	1.5%
$2.5 \leq \|y\| < 3.0$	62%	71%	69%	25%	55%
$p_{\mathrm{T}}^{\mathrm{jet}}$					
$\|y\| < 0.5$	69%	67%	66%	30%	46%
$0.5 \leq \|y\| < 1.0$	7.4%	8.9%	8.6%	3.4%	2.0%
$1.0 \leq \|y\| < 1.5$	69%	62%	68%	45%	54%
$1.5 \leq \|y\| < 2.0$	1.3%	1.6%	1.4%	0.1%	0.5%
$2.0 \leq \|y\| < 2.5$	8.7%	6.6%	7.4%	1.0%	3.6%
$2.5 \leq \|y\| < 3.0$	65%	72%	72%	28%	59%

Table 4.6: Summary of χ^2/dof values obtained from a global fit using all p_{T} and rapidity bins, comparing the inclusive jet cross-section and the NLO pQCD prediction corrected for non-perturbative and electroweak effects for several PDF sets and for the two scale choices. All the corresponding p-values are $\ll 10^{-3}$.

χ^2/dof all $\|y\|$ bins	CT14	MMHT 2014	NNPDF 3.0	HERAPDF 2.0	ABMP16
$p_{\mathrm{T}}^{\mathrm{max}}$	${}^{419}/_{177}$	${}^{431}/_{177}$	${}^{404}/_{177}$	${}^{432}/_{177}$	${}^{475}/_{177}$
$p_{\mathrm{T}}^{\mathrm{jet}}$	${}^{399}/_{177}$	${}^{405}/_{177}$	${}^{384}/_{177}$	${}^{428}/_{177}$	${}^{455}/_{177}$

Since the uncertainties in the NNLO pQCD predictions do not yet include the contributions from the PDF and α_{s} uncertainties, it is not possible to perform a quantitative comparison to the measurements. However, one can conclude from Figure 4.41 that the differences between data and the theoretical predictions at NNLO are smaller (larger) than at NLO for the $p_{\mathrm{T}}^{\mathrm{jet}}$ ($p_{\mathrm{T}}^{\mathrm{max}}$) scale choice.

Chapter 5

Search for new phenomena in high-mass photon+jet final states

5.1 Preface

Prompt photons in association with jets are copiously produced at the LHC, mainly through quark-gluon scattering ($qg \to q\gamma$). The dominant process is illustrated in Figure 5.1a, and corresponds to a t-channel. The γ + jet(s) final state provides also a sensitive probe for a class of phenomena beyond the Standard Model that could manifest themselves in the high end of the invariant mass spectrum of the γ+jet system ($m_{\gamma j}$). The search is performed by looking for localised excesses of events in the $m_{\gamma j}$ distribution with respect to the SM prediction. Two classes of benchmark signal models are considered. The first class consists of specific BSM models and the second one is based on generic Gaussian-shaped mass distributions.

The BSM signals are implemented in Monte Carlo simulation and appear as broad peaks in the $m_{\gamma j}$ spectrum. In this search, two scenarios are considered: quarks as composite particles and extra spatial dimensions. In the first case, if quarks are composed of more fundamental constituents bound together by some unknown interaction, new effects should appear depending on the value of the compositeness scale Λ. A quark substructure could explain the origin of the three generations of quarks as well as their various masses and behaviours under weak interactions. In particular, if Λ is sufficiently smaller than the centre-of-mass energy, excited quark (q^*) states may be produced in high-energy pp collisions at the LHC [201–203]. The q^* production at the LHC could result in a resonant peak at the mass of the q^* (m_{q^*}) in the $m_{\gamma j}$ distribution if the q^* can decay into a photon and a quark ($qg \to q^* \to q\gamma$). This present search considers excited up (u^*) and down (d^*) quarks, and only the SM gauge interactions for q^* production. In the second scenario, the existence of extra spatial dimensions (EDs) is assumed to provide a solution to the hierarchy problem [46–48]. Certain types of ED models predict the fundamental Planck scale M^* in the $4+n$ dimensions (n being the number of extra spatial dimensions) to be at the TeV scale, and thus accessible in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ at the LHC. In such a TeV-scale M^* scenario of the extra dimensions, quantum black holes (QBHs) may be produced at the LHC as a continuum above the threshold mass (M_{th}) and then decay into a small number of final-state particles including photon-quark/gluon pairs before they are able to thermalise [204–207]. In this case a broad resonance-like structure could be observed just above M_{th} on top of the SM $m_{\gamma j}$ distribution. The M_{th} value for QBH production is taken to be equal to M^* while the maximum allowed QBH mass is set to either $3M^*$ or the LHC pp centre-of-mass energy of 13 TeV, whichever is smaller. The upper bound on the mass ensures that the QBH production is far from the "thermal" regime, where the classical description of the black hole and its decay into high-multiplicity final states should be used. In this analysis, the extra-dimensions model proposed by Arkani-Hamed, Dimopoulous and Dvali (ADD) [208] with $n = 6$ flat EDs, and the one by Randall and Sundrum (RS1) [209] with $n = 1$ warped ED are considered. Only the following six non-thermal black hole states and their decays to a photon and a parton are considered in this search:

$$\begin{array}{lcllcl} u+g & \rightarrow & \mathrm{QBH}^{2/3} & \rightarrow & u+\gamma \\ \bar{d}+g & \rightarrow & \mathrm{QBH}^{1/3} & \rightarrow & \bar{d}+\gamma \\ q+\bar{q} & \rightarrow & \mathrm{QBH}^{0} & \rightarrow & g+\gamma \\ g+g & \rightarrow & \mathrm{QBH}^{0} & \rightarrow & g+\gamma \\ d+g & \rightarrow & \mathrm{QBH}^{-1/3} & \rightarrow & d+\gamma \\ \bar{u}+g & \rightarrow & \mathrm{QBH}^{-2/3} & \rightarrow & \bar{u}+\gamma \end{array}$$

where u represents all up-type quarks, d all down-type quarks, and q all quark flavours.

The second class of benchmark models, based on a generic Gaussian-shaped mass distribution with different values of its mean and standard deviation, provides a generic interpretation for the presence of signals with different Gaussian widths, ranging from a resonance with a width similar to the reconstructed $m_{\gamma j}$ resolution of $\sim 2\%$ to wide resonances with a width up to 15%. This approach is considered since new physics is not only limited to particular models but may appear in unexpected places with an unexpected form. Hence, the search presented here is ultimately model independent and is sensitive to any localised excess which may be caused by any BSM process with a γ + jet final state.

The ATLAS and CMS experiments at the LHC have performed searches for excited quarks in the γ + jet final state using pp collision data recorded at $\sqrt{s} = 7$ TeV [210], 8 TeV [211, 212] and 13 TeV [213]. In the ATLAS searches, limits for generic Gaussian-shaped resonances were obtained at 7, 8 and 13 TeV while a limit for QBHs in the ADD model ($n = 6$) was first obtained at 8 TeV. The ATLAS search at 13 TeV with data taken in 2015 was further extended to constrain QBHs in the RS1 model ($n = 1$). The dijet resonance searches at ATLAS [176, 214] and CMS [215] using pp collisions at $\sqrt{s} = 13$ TeV also set limits on the production cross-sections of excited quarks and QBHs.

This chapter presents the search based on the full 2015 and 2016 dataset recorded with the ATLAS detector, corresponding to 36.7 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV. The analysis strategy is unchanged from the one reported in Ref. [213], focusing on the region where the γ + jet system has a high invariant mass. The search here presented complements the dijet results and provides an independent check for the presence of these signals in different decay channels.

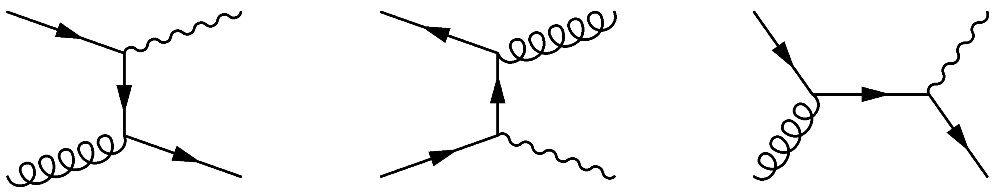
This chapter is organised as follows. The SM background γ+jet final state processes are discussed in Section 5.2. Section 5.3 summarises the data and simulation samples used in this study. The event selection is presented in Section 5.4. The signal and background modelling are introduced in Section 5.5 together with the signal search and limit-setting strategies. Finally the results are discussed in Section 5.6.

5.2 Background processes

Photon plus jet events may be produced at tree level in the SM through either Compton scattering of a quark and a gluon or through quark-antiquark annihilation, as shown in Figure 5.1. The quark-gluon diagrams account for most of $\gamma+\text{jet}$ production since those are favored over the quark-antiquark diagrams due to the PDFs. $\gamma+\text{jet}$ events can also be produced at the LHC through gluon annihilation, but this is an NNLO process (Figure 5.2 shows the lowest order Feynman diagram).

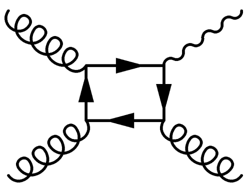
Events with a real high-p_{T} photon and one or more jets can also be produced if a parton shower quark rediates off a photon, or when a hadron in a multijet event decays producing a photon. While such photons tend to appear near or inside jets and thus fail isolation criteria, the much larger multijet cross-sections (e.g. the ratio of dijet to $\gamma+\text{jet}$ cross-sections is of the order of α_s/α) imply that such fragmentation production can be a non-negligible contribution to isolated $\gamma+\text{jet}$ signatures.

Figure 5.1: Standard Model diagrams contributing to $\gamma+\text{jet}$ production at tree level.



(a) (b) (c)

Figure 5.2: Lowest order $gg \rightarrow \gamma+\text{jet}$ diagram in the Standard Model.



Events without a photon at parton level can also pass the event selection, and are called "fakes". This is a smaller component of background photon candidates and is due to hadrons producing in the detector energy deposits that have characteristics similar to those of real photons. By far, the dominant process is from dijet events, in which one of the jets pass the photon selection.

5.3 Data and Monte Carlo simulations

The data sample used in this analysis was collected from pp collisions in the 2015–2016 LHC run at a centre-of-mass energy of 13 TeV, and corresponds to an integrated luminosity of $36.7 \pm 1.2\ \text{fb}^{-1}$. The uncertainty was derived, following a methodology similar to that detailed in Ref. [170], from a preliminary calibration of the luminosity scale using x–y beam-separation scans performed in August 2015 and May 2016. The data are required to satisfy a number of quality criteria ensuring that the relevant detectors were operational while the data were recorded.

Monte Carlo samples of simulated events are used to study the background modelling for the dominant γ+jet processes, to optimise the selection criteria and to evaluate the acceptance and selection efficiencies for the signals considered in the search. Events from SM processes containing a photon with associated jets were simulated using the SHERPA 2.1.1 event generator, requiring a photon transverse energy E_{T}^{γ} above 70 GeV at the generator level. The matrix elements were calculated with up to four final state partons at leading order in quantum chromodynamics and merged with the parton shower [216] using the ME+PS@LO prescription [217]. The CT10 PDF set was used in conjunction with a dedicated parton shower tuning developed by the SHERPA authors. A second sample of SM γ + jet events was generated using the LO PYTHIA 8.186 event generator with the LO NNPDF 2.3 PDF set and the A14 tuning of the underlying-event parameters. The PYTHIA simulation includes leading-order γ + jet events from both the direct processes (the hard subprocesses $qg \to q\gamma$ and $q\bar{q} \to g\gamma$) and bremsstrahlung photons in QCD dijet events. The samples have been separately generated in E_T^{γ} bins to cover the full spectrum relevant for this analysis. In each slice, three samples are generated depending on the flavor of the jet : b-jet, c-jet and light jet. To estimate the systematic uncertainty associated with the background modelling, a large sample of events was generated with the next-to-leading-order JETPHOX v1.3.1_2 [218] program. Events were generated at parton level for both the direct and fragmentation photon contributions, using the NLO photon fragmentation functions [219] and the NLO NNPDF 2.3 PDF set, and setting the renormalisation, factorisation and fragmentation scales to the photon E_{T}^{γ}. Jets of partons are reconstructed using the anti-k_t algorithm with a radius parameter of $R = 0.4$. The generated photon is required to be isolated by ensuring that the total transverse energy of partons inside a cone of size $\Delta R = 0.4$ around the photon is smaller than 7.07 GeV + $0.03 \times E_{\mathrm{T}}^{\gamma}$, equivalent[1] to the photon selection in data described in Section 5.4.

Samples of excited quark events were produced using PYTHIA 8.186 with the LO NNPDF 2.3 PDF set and the A14 set of tuned parameters for the underlying event. The Standard Model gauge interactions and the magnetic-transition type couplings [201–203] to gauge bosons were considered in the production processes of the excited states of the first-generation quarks (u^*, d^*) with degenerate masses. The compositeness scale Λ was taken to be equal to the mass m_{q^*} of the excited quark, and the coefficients f_{s}, f and f' of magnetic-transition type couplings to the respective SU(3), SU(2) and U(1) gauge bosons were chosen to be unity. The q^* samples were generated with m_{q^*} values between 0.5 and 6.0 TeV in steps of 0.5 TeV.

The QBH samples were generated using the QBH 2.02 [220] event generator with the CTEQ6L1 PDF set and PYTHIA 8.186 for the parton shower and underlying event tuned with the A14 parameter set. The M_{th} values were chosen to vary between 3.0 (1.0) and 9.0 (7.0) TeV in steps of 0.5 TeV for the QBH signals in the ADD (RS1) model. All the qg, $\bar{q}g$, gg and $q\bar{q}$ processes were included in the QBH signal production while only final states with a photon and a quark/gluon were considered for the decay. All six quark flavours were included together with their antiquark counterparts in both the production and decay processes.

Apart from the sample generated with JETPHOX which is a parton-level calculator, all the simulated samples include the effects of multiple pp interactions in the same and neighbouring bunch crossings (pile-up) and were processed through the ATLAS detector

[1] The parton-level isolation requirement takes into account the correlation between reconstruction-level isolation energies and particle-level isolation energies, as a proxy for the parton-level isolation, as evaluated using γ + jet events simulated by PYTHIA 8.186.

simulation model based on GEANT 4. Pile-up effects were emulated by overlaying simulated minimum-bias events from PYTHIA 8.186, generated with the AU2 tune for the underlying event and the MSTW2008LO PDF set. The number of overlaid minimum-bias events was adjusted to match the one observed in data (see Section 4.3.1). All the MC samples except for the JETPHOX sample were reconstructed with the same software as that used for collision data.

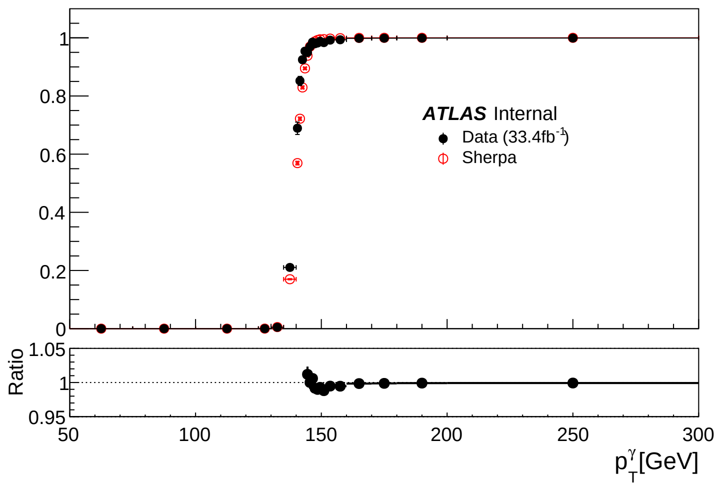
5.4 Event selection

Photons are reconstructed from clusters of energy deposits in the EM calorimeter as described in Section 2.3.3. Both the converted and unconverted photon candidates are used in the analysis. Photon candidates are required to have $E_{\mathrm{T}}^{\gamma} > 25\,\mathrm{GeV}$ and $|\eta^{\gamma}| < 2.37$ and satisfy the "tight" identification criteria, designed to lower significantly the contribution from background, defined in Ref. [115]. Photons are identified based on the profile of the energy deposits in the first two layers of the EM calorimeter and the energy leakage into the hadronic calorimeter. To further reduce the contamination from $\pi^0 \to \gamma\gamma$ or other neutral hadrons decaying into photons, the photon candidates are required to be isolated from other energy deposits in an event. The calorimeter isolation variable $E_{\mathrm{T,\,iso}}$ is defined as the sum of the E_{T} of all positive-energy topological clusters reconstructed within a cone of $\Delta R = 0.4$ around the photon direction excluding the energy deposits in an area of size $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ centred on the photon cluster. The photon energy expected outside the excluded area is subtracted from the isolation energy while the contributions from pile-up and the underlying event are subtracted event by event [221]. The photon candidates are required to have $E_{\mathrm{T,\,iso}}^{\gamma} = E_{\mathrm{T,\,iso}} - 0.022 \times E_{\mathrm{T}}^{\gamma}$ less than 2.45 GeV. This E_{T}^{γ}-dependent selection requirement is used to guarantee an efficiency greater than 90% for signal photons in the E_{T}^{γ} range relevant for this analysis. The efficiency for the signal photon selection varies from $(90 \pm 1)\%$ to $(83 \pm 1)\%$ for signal events with masses from 1 to 6 TeV. The dependency on the signal mass is mainly from the efficiency of the tight identification requirement while the isolation selection efficiency is approximately $(99 \pm 1)\%$ over the full mass range.

Jets are reconstructed from topological clusters calibrated at the electromagnetic scale using the anti-k_t algorithm with a radius parameter $R = 0.4$. The jets are calibrated to the hadronic energy scale by the procedure detailed in Section 3.2. Jets from pile-up interactions are suppressed by applying the jet vertex tagger [222], using information about tracks associated with the hard-scatter and pile-up vertices, to jets with $p_{\mathrm{T}}^{\mathrm{jet}} < 60\,\mathrm{GeV}$ and $|\eta^{\mathrm{jet}}| < 2.4$. In order to remove jets due to calorimeter noise or non-collision backgrounds, events containing at least one jet failing to satisfy the loose quality criteria (introduced already in Section 4.4) are discarded. Jets passing all the requirements and with $p_{\mathrm{T}}^{\mathrm{jet}} > 20\,\mathrm{GeV}$ and $|\eta^{\mathrm{jet}}| < 4.5$ are considered in the rest of the analysis. Since a photon is also reconstructed as a jet, jet candidates in a cone of $\Delta R = 0.4$ around a photon are not considered.

This analysis selects events based on an unprescaled single-photon trigger requiring at least one photon candidate with $E_{\mathrm{T}}^{\gamma} > 140$ GeV which satisfies loose identification conditions [115] based on the shower shape in the second sampling layer of the EM calorimeter and the energy leakage into the hadronic calorimeter. Selected events are required to contain at least one primary vertex with two or more tracks with $p_{\mathrm{T}} > 400\,\mathrm{MeV}$, to further reject events due to non-collision backgrounds. Photon candidates are required to satisfy the "tight" identification and isolation conditions discussed above. The kinematic requirements for the highest-E_{T} photon in the events are tightened to $E_{\mathrm{T}}^{\gamma} > 150\,\mathrm{GeV}$ and $|\eta^{\gamma}| < 1.37$. The E_{T}^{γ} requirement is used to select events with nearly 100% trigger efficiency (see Figure 5.3). The η^{γ} requirement is imposed to enhance the signal-to-background ratio, as shown in Figure 5.4.

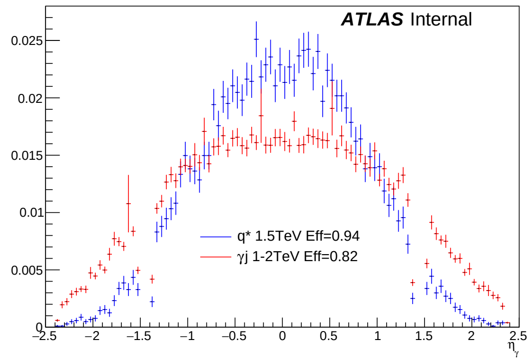
Figure 5.3: Measured efficiency of the single-photon trigger used in the analysis as a function of offline photon p_{T}. The measurement has been performed with data collected in 2016 with a di-photon trigger requiring at least two photon candidates which satisfy loose identification conditions and have $E_{\mathrm{T}}^{\gamma} > 35\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{\gamma} > 25\,\mathrm{GeV}$, respectively. The measured efficiency curve is compared with the MC prediction, evaluated with the same method.



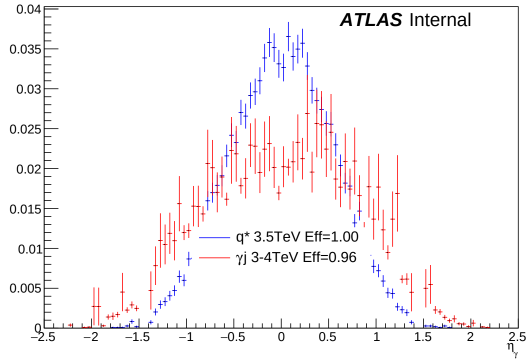
Moreover, an event is rejected if there is any jet with $p_{\mathrm{T}}^{\mathrm{jet}} > 30\,\mathrm{GeV}$ within $\Delta R < 0.8$ around the photon. This is an extended photon isolation requirement to further avoid possible contamination from nearby jets to the photon isolation cone, which might lead to a bias in the calibration of the photon. The presence of additional tight and isolated photons with $E_{\mathrm{T}}^{\gamma} > 150\,\mathrm{GeV}$ in events is negligible for both signal and background events, and therefore allowed. The $\gamma+\mathrm{jet}$ system is formed from the highest-E_{T} photon and the highest-p_{T} jet in the event. Finally, the highest-p_{T} jet in the event is required to have $p_{\mathrm{T}}^{\mathrm{jet}} > 60\,\mathrm{GeV}$ and the pseudorapidity difference between the photon and the jet ($\Delta\eta_{\gamma j} \equiv |\eta^{\gamma} - \eta^{\mathrm{jet}}|$) must be less than 1.6 to enhance signals over the $\gamma+\mathrm{jet}$ background. Background events with low E_{T}^{γ} and high $m_{\gamma j}$ are likely to have large $\Delta\eta_{\gamma j}$. This can be understood as a kinematic constraint since events with small transverse activity need to have large separation in longitudinal direction between photon and jet to attain a large $m_{\gamma j}$. Since the fraction of events with low p_{T} photons in high $m_{\gamma j}$ region is higher in background than signal because of the t-channel production contribution only present for background, the $\Delta\eta_{\gamma j}$ cut improves the signal-background separation in the high $m_{\gamma j}$ region. Figure 5.5 shows the $\Delta\eta_{\gamma j}$ distributions for some signal models. The shapes are similar for various mass points and for both the QBH and excited quark models. It was found that the optimal $\Delta\eta_{\gamma j}$ cut does not depend on the signal models.

Table 5.1 shows the cut flow for this search. 3.75×10^{6} events are selected by the criteria introduced above. 6.34×10^{5} events with an invariant mass ($m_{\gamma j}$) of the selected $\gamma+\mathrm{jet}$ system greater than 500 GeV remain in the data sample, while 2.10×10^{4} events are in the high mass regions ($m_{\gamma j} > 1.1$ TeV) that are used in the fit on the invariant mass.

Figure 5.4: η^{γ} distributions for excited quark signals after all selection criteria applied but $|\eta^{\gamma}| < 1.37$ cut. For the SM $\gamma + \mathrm{j}$ sample, a mass cut is applied to account for the effect on the mass fit. The $|\eta^{\gamma}| < 1.37$ cut improves the signal-to-noise ratio in the selected samples. The drop around $1.37 < |\eta^{\gamma}| < 1.52$ corresponds to the crack region of the EM calorimeter.

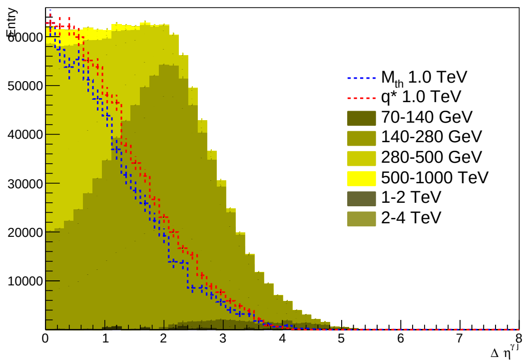


(a) $M_{q*} = 1.5\,\mathrm{TeV}$

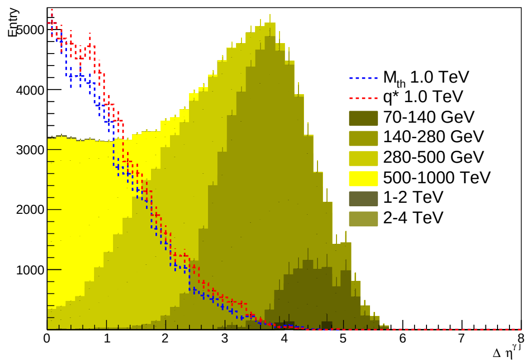


(b) $M_{q*} = 3.5\,\mathrm{TeV}$

Figure 5.5: $\Delta\eta_{\gamma j}$ distribution for signal ($M_{q^*} = 1.0\,\mathrm{TeV}$, and QBH with $M_{\mathrm{th}} = 1.0\,\mathrm{TeV}$) and backgrounds. Panels (a) and (b) shows signal and background events which satisfy $0.5 < m_{\gamma j} < 1.5$ TeV and $1.0 < m_{\gamma j} < 2.0$ TeV, respectively. For both (a) and (b) the signal normalisation is arbitrary. The different colours show the different E_{T}^{γ} slices of SHERPA samples.



(a) $0.5 < m_{\gamma j} < 1.5\,\mathrm{TeV}$



(b) $1.0 < m_{\gamma j} < 2.0\,\mathrm{TeV}$

Table 5.1: Cut flow of the search for new phenomena in high-mass final states with a photon and a jet as a function of sequentially applied requirements.

Events requirements (cumulative)	N_{event}	ε_{abs}	ε_{rel}
Total	140824880	2.19×10^{-2}	2.19×10^{-2}
GRL	137084448	2.14×10^{-2}	0.97
$N_{\text{trk@PV}}$	137082976	2.14×10^{-2}	1.00
Event Cleaning	136708272	2.13×10^{-2}	1.00
Trigger	93311040	1.45×10^{-2}	6.83×10^{-1}
γ loose	58937192	9.18×10^{-3}	6.32×10^{-1}
γ tight	17956998	2.80×10^{-3}	3.05×10^{-1}
γ $p_T > 150$ GeV	13597482	2.12×10^{-3}	7.57×10^{-1}
γ $\lvert\eta\rvert < 1.37$	8696047	1.36×10^{-3}	6.40×10^{-1}
γ isolation	5067771	7.90×10^{-4}	5.83×10^{-1}
Photon-jet correlation	5067669	7.90×10^{-4}	1.00
At least one jet after overlap removal	5063919	7.89×10^{-4}	1.00
Extended isolation	4826424	7.52×10^{-4}	0.95
Jet cleaning	4815590	7.50×10^{-4}	1.00
Jet $p_T > 60\ GeV$	4772917	7.44×10^{-4}	0.99
$\Delta\eta(\gamma, j) < 1.6$	3750954	5.85×10^{-4}	7.86×10^{-1}

5.5 Statistical analysis

The data are examined for the presence of a significant deviation from the SM prediction using a test statistic based on the profile likelihood ratio technique [223]. Limits on the visible cross-section for generic Gaussian-shaped signals and limits on the cross-section times branching ratio for specific benchmark models are computed using the CL_s prescription [224]. The details of the signal and background modelling used for the likelihood function construction are discussed in Sections 5.5.1 and 5.5.2, respectively, while a summary of the statistical procedures used to establish the presence of a signal or set limits on the production cross-sections for new phenomena is given in Section 5.5.3. The model (signal and background modelling) has been implemented in the RooFit package [225], while the limits have been set with the RooStats package [226].

5.5.1 Signal modelling

The signal model is built starting from the probability density function (pdf) of the $m_{\gamma j}$ distribution $f_{\text{sig}}(m_{\gamma j})$ at the reconstruction level. For a Gaussian-shaped resonance with mass m_{G}, the $m_{\gamma j}$ pdf is modelled by a normalised Gaussian distribution with the mean located at $m_{\gamma j} = m_{\text{G}}$. The standard deviation of the Gaussian distribution is chosen to be 2%, 7% or 15% of m_{G}, where 2% approximately corresponds to the effect of the detector resolution on the reconstruction of the photon–jet invariant mass. For the q^* and QBH signals, the $m_{\gamma j}$ pdfs are created from the normalised reconstructed $m_{\gamma j}$ distributions after applying the selection requirements described in Section 5.4 using simulated MC events, and a kernel density estimation technique [227] is applied to smooth the distributions. The signal pdfs for intermediate mass points at which signal events were not generated are obtained from the simulated samples by using a moment-morphing method [228]. The signal template for the q^* and QBH signals is then constructed as $f_{\text{sig}}(m_{\gamma j}) \times (\sigma_{\text{S}} \cdot B \cdot A \cdot \varepsilon) \times \mathcal{L}_{\text{int}}$, where f_{sig} is scaled by the product of the cross-section times branching ratio to a photon and a quark or gluon ($\sigma_{\text{S}} \cdot B$), acceptance ($A$), selection efficiency (ε) and the integrated luminosity ($\mathcal{L}_{\text{int}}$) for the data sample. The product of the acceptance times efficiency ($A \cdot \varepsilon$) is found to be about 50% for all the q^* and QBH models, varying only by a few percent with m_{q^*} or M_{th}. This dependence is accounted for in the model by interpolating between the generated mass points using a third-order spline. For the q^* and QBH signals, limits are set on $\sigma_{\text{S}} \cdot B$ after correcting for the acceptance and efficiency $A \cdot \varepsilon$ of the selection criteria.

Experimental uncertainties in the signal yield arise from uncertainties in the luminosity (±3.2%), photon identification efficiency (±2%), trigger efficiency (±1% as measured in Ref. [229]) and pile-up dependence (±1%). The pile-up uncertainty is derived by using alternative scale factors in the pile-up reweighting tool (see Section 4.3.1). The impact of the uncertainties in the photon isolation efficiency, photon and jet energy scales is negligible. A 1% uncertainty in the signal yield is included to account for the statistical error in the acceptance and selection efficiency estimates due to the limited size of the MC signal samples. The impact of the PDF uncertainties on the signal acceptance is found to be negligible compared to the other uncertainties. The photon and jet energy resolution uncertainties (±2% of the mass) are accounted for as a variation of the width for the Gaussian-shaped signals. The impact of the resolution uncertainty on intrinsically large width signals is found to be negligible and thus not included in the signal models for the q^* and QBH. The typical difference between the peaks of the reconstructed and generator-level $m_{\gamma j}$ distributions for the excited-quark signals is well below 1%.

A summary of systematic uncertainties in the signal yield and shape included in the statistical analysis is given in Table 5.2.

Table 5.2: Summary of systematic uncertainties in the signal event yield and shape included in the fit model. The signal mass resolution uncertainty affects the generic Gaussian signal shape, while the other uncertainties affect the event yield.

Uncertainty	q^* and QBH	Generic Gaussian
Signal mass resolution	N/A	$\pm 2\% \cdot m_{\text{G}}$
Photon identification	±2%	N/A
Trigger efficiency	±1%	N/A
Pile-up dependence	±1%	N/A
MC event statistics	±1%	N/A
Luminosity	±3.2%	

In order to facilitate the reinterpretation of the present results in alternative physics models, the fiducial acceptance and efficiency for events with the invariant mass of the γ + jet system around m_{q^*} or M_{th} (referred to as “on-shell” events hereafter) are also provided. The chosen $m_{\gamma j}$ ranges are $0.6 m_{q^*} < m_{\gamma j} < 1.2 m_{q^*}$ for the q^* signal and $0.8 M_{\text{th}} < m_{\gamma j} < 3.0 M_{\text{th}}$ for the QBH signal. The fiducial region at particle level, as summarised in Table 5.3, is chosen to be close to the one used in the event selection at reconstruction level.

The fiducial acceptance A_{f}, defined as the fraction of generated on-shell signal events falling into the fiducial region, increases from 56% to 63% with increasing signal mass M_{th} from 1.0 to 6.5 (9.0) TeV for the QBH in the RS1 (ADD) model. The A_{f} value for

Table 5.3: Requirements on the photon and jet at particle level to define the fiducial region and on the detector-level quantities for the selection efficiency.

Particle-level selection for fiducial region
Photon : $E_{\mathrm{T}}^{\gamma} > 150\,\mathrm{GeV}$, $|\eta^{\gamma}| < 1.37$
Jet : $p_{\mathrm{T}}^{\mathrm{jet}} > 60\,\mathrm{GeV}$, $|\eta^{\mathrm{jet}}| < 4.5$
Photon–Jet η separation : $|\Delta\eta_{\gamma j}| < 1.6$
No jet with $p_{\mathrm{T}}^{\mathrm{jet}} > 30\,\mathrm{GeV}$ within $\Delta R < 0.8$ around the photon
Detector-level selection for selection efficiency
Tight photon identification
Photon isolation
Jet identification including quality and pile-up rejection requirements

the q^* model varies very similarly to that for the RS1 QBH signal. The rise in the fiducial acceptance as a function of M_{th} (m_{q^*}) is driven mainly by the increase of the efficiency for the photon η requirement since the photons tend to be more central as M_{th} (m_{q^*}) becomes larger.

The fiducial selection efficiency ε_{f} is defined as the ratio of the number of on-shell events in the particle-level fiducial region passing the selection at the reconstruction level, including photon identification, isolation and jet quality criteria, to the number of generated on-shell events in the particle-level fiducial region. The migration of generated on-shell events outside the particle-level fiducial region into the selected sample at the reconstruction level is found to be negligible. The fiducial selection efficiency decreases from 88 (86) to 82 (80)% within the same M_{th} ranges as above for the RS1 (ADD) QBH model and is not highly dependent on the kinematics of the assumed signal production processes. The ε_{f} for the q^* model behaves very similarly to that for the RS1 QBH model. The reduction in the fiducial selection efficiency is caused mainly by the inefficiency of the shower shape requirements used in the photon identification for high-E_{T} photons. The fiducial acceptance and selection efficiencies for the three benchmark signal models are shown in Figure 5.6 as functions of m_{q^*} or M_{th}.

5.5.2 Background modelling

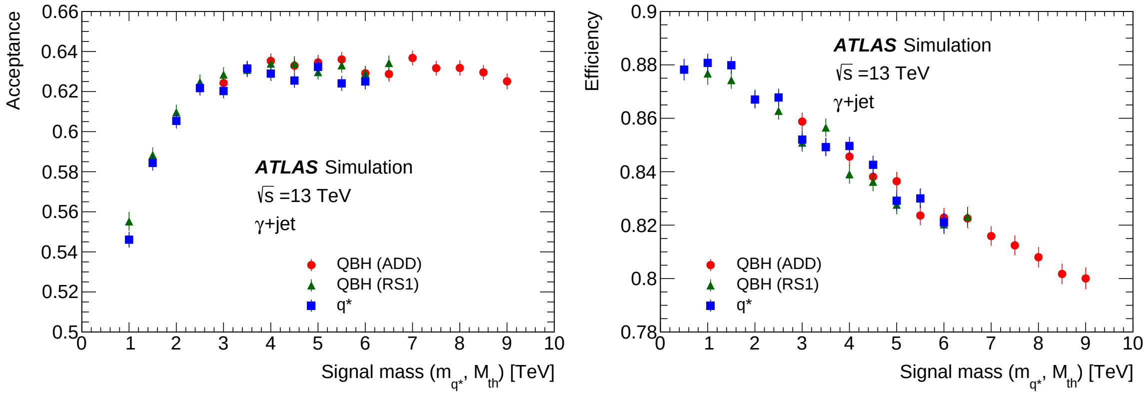
The $m_{\gamma j}$ distribution of the background is modelled using a functional form. A family of functions, similar to the ones used in the previous searches for $\gamma+$jet [210, 211, 213] and $\gamma\gamma$ resonances [230] as well as dijet resonances [176] is considered:

$$f_{\mathrm{bg}}(x) = N(1-x)^p x^{\sum_{i=0}^{k} a_i (\log x)^i}\,, \tag{5.1}$$

where x is defined as $m_{\gamma j}/\sqrt{s}$, p and a_i are free parameters, and N is a normalisation factor. The number of free parameters describing the normalised mass distribution is thus $k+2$ with a fixed N. The normalisation N as well as the shape parameters p and a_i are simultaneously determined by the final fit of the signal plus background model to data.

The goodness of a given functional form in describing the background is quantified based on the potential bias introduced in the fitted number of signal events. To quantify this bias the functional form under test is used to perform a signal + background fit to a large sample of background events built from the JETPHOX prediction. The large

Figure 5.6: (a) Fiducial acceptance and (b) selection efficiencies for the three signal models considered in the analysis as a function of the excited-quark mass m_{q^*} or the QBH threshold mass M_{th}. The fiducial region definition is detailed in Table 5.3. The description of the selection criteria can be found in the text.



(a) (b)

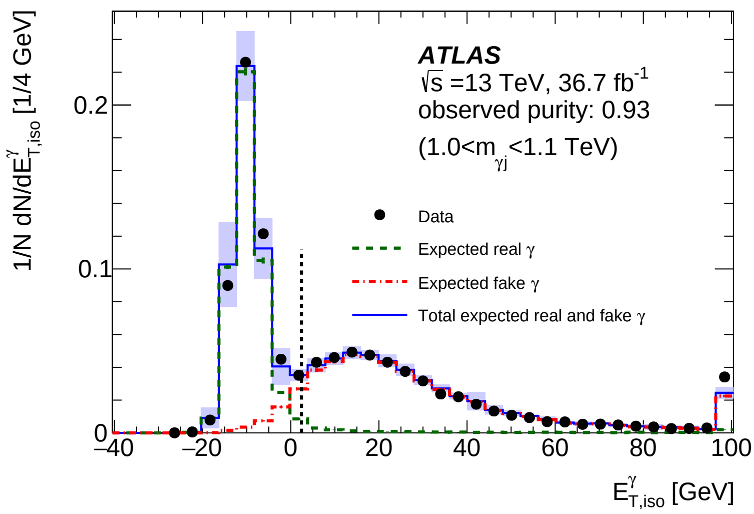
JETPHOX event sample is used for this purpose as the shape of the background prediction can be extracted with sufficiently small statistical uncertainty.

The parton-level JETPHOX calculations do not account for effects from hadronisation, the underlying event and the detector resolution. Therefore, the nominal JETPHOX prediction is corrected by calculating the ratio of reconstructed jet p_{T} to parton p_{T} in the SHERPA $\gamma+$jet sample and applying the parametrised ratio to the JETPHOX parton p_{T}.

In addition, an $m_{\gamma j}$-dependent correction is applied to the JETPHOX prediction to account for the contribution from multijet events where one of the jets is misidentified as a photon (fake photon events). This correction is estimated from data as the inverse of the purity, defined as the fraction of real $\gamma+$jet events in the selected sample. The purity is measured in bins of $m_{\gamma j}$ by exploiting the difference between the shapes of the $E^{\gamma}_{\mathrm{T,\,iso}}$ distributions of real photons and jets faking photons; the latter typically have a larger $E^{\gamma}_{\mathrm{T,\,iso}}$ value due to nearby particles produced in the jet fragmentation. The purity is estimated by performing a two-component template fit to the $E^{\gamma}_{\mathrm{T,\,iso}}$ distribution in bins of $m_{\gamma j}$. The templates of real- and fake-photon isolation distributions are obtained from MC (SHERPA) simulation and from data control samples, respectively. The $E^{\gamma}_{\mathrm{T,\,iso}}$ variable for real photons from SHERPA simulation is corrected to account for the observed mismodelling in the description of isolation profiles between data and MC events in a separate control sample. The template for fake photons is derived in a data sample where the photon candidate fails to satisfy the tight identification criteria but fulfills a looser set of identification criteria. Details about the correction to the real-photon template and the derivation of the fake-photon template are given in Ref. [231]. To reduce the bias in the $E^{\gamma}_{\mathrm{T,\,iso}}$ shape due to different kinematics, both the real- and fake-photon templates are obtained by applying the same set of kinematic requirements used in the main analysis. As an example, Figure 5.7 shows the $E^{\gamma}_{\mathrm{T,\,iso}}$ distribution of events within the range $1.0 < m_{\gamma j} < 1.1$ TeV, superimposed on the best-fit result. This procedure is repeated in every bin of the $m_{\gamma j}$ distribution and the resulting estimate of the purity is shown as a function of $m_{\gamma j}$ in Figure 5.8. The uncertainty in the measured purity includes both statistical and systematic components. The latter are estimated by recomputing the purity using different data control samples for the fake-photon template or alternative templates for real photons obtained from PYTHIA simulation or removing the data-to-MC corrections applied to $E^{\gamma}_{\mathrm{T,iso}}$ in the SHERPA sample and by symmetrising the variations. The variation from different data control samples for the fake-photon template has the largest effect on the purity (0.04 at $1.0 < m_{\gamma j} < 1.1$ TeV). The measured purity is approximately constant at 93% over the $m_{\gamma j}$ range above 500 GeV, indicating that the fake-photon contribution does not depend significantly on $m_{\gamma j}$.

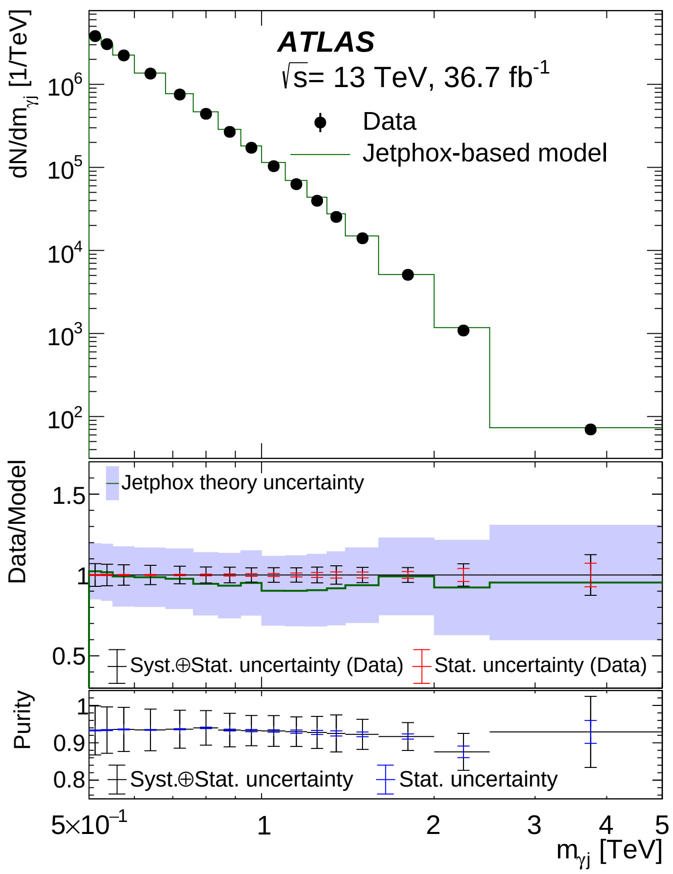
The $m_{\gamma j}$ distribution in data compared to the corrected JETPHOX $\gamma+$jet prediction normalised to data in the $m_{\gamma j} > 500$ GeV region is shown in Figure 5.8. Theoretical uncertainties in the JETPHOX prediction are computed by considering the variations induced by $\pm 1\sigma$ of the NNPDF 2.3 PDF uncertainties, by switching between the nominal NNPDF 2.3 and CT10 or MSTW2008 PDF sets, by the variation of the value of the strong coupling constant by ± 0.002 around the nominal value of 0.118 and by the variation of the renormalisation, factorisation and fragmentation scales between half and twice the photon transverse momentum. The differences between data and the corrected JETPHOX prediction shown in Figure 5.8 are well within the uncertainties associated with the perturbative QCD prediction.

Figure 5.7: Distribution of $E^{\gamma}_{\mathrm{T,iso}} = E_{\mathrm{T,iso}} - 0.022 \times E^{\gamma}_{\mathrm{T}}$ for the photon candidates in events with $1.0 < m_{\gamma j} < 1.1$ TeV, and the comparison with the result of the template fit. Real- and fake-photon components determined by the fit are shown by the green dashed and red dot-dashed histograms, respectively, and the sum of the two components is shown as the solid blue histogram. The blue band shows the systematic uncertainties in the real- plus fake-photon template. The last bin of the distribution includes overflow events. The vertical dashed line corresponds to the isolation requirement used in the analysis. The photon purity determined from the fit for the selected sample in the $1.0 < m_{\gamma j} < 1.1$ TeV mass range is 93%.



The number of signal events extracted by the signal + background fit to the pure background model described above is called the “spurious signal” [37] and it is used to select the optimal functional form and the $m_{\gamma j}$ range of the fit. In order to account for the assumption that the corrected JETPHOX prediction itself is a good representation of the data, the fit is repeated on modified samples obtained by changing the nominal shape to account for several effects: firstly, the nominal distribution is corrected to follow the envelope of the changes induced by $\pm 1\sigma$ variations of the NNPDF 2.3 PDF uncertainty, the variations between the nominal NNPDF 2.3 and CT10 or MSTW2008 PDF sets, the variation of the value of the strong coupling constant by ± 0.002 around the nominal value of 0.118 and the variation of the renormalisation, factorisation and fragmentation scales between half and twice the photon transverse momentum; secondly the corrections

Figure 5.8: Distribution of the invariant mass of the γ + jet system as measured in the $\gamma +$ jet data (dots), compared with the JETPHOX (green histogram) $\gamma +$ jet predictions. The JETPHOX distribution is obtained after correcting the parton-level spectrum for showering, hadronisation and detector resolution effects as described in the text. The JETPHOX spectrum is normalised to the data in the $m_{\gamma j}$ range above 500 GeV. The ratio of the data to JETPHOX prediction as a function of $m_{\gamma j}$ is shown in the middle panel (green histogram): the theoretical uncertainty is shown as a shaded band. The statistical uncertainty from the data sample and the sum of the statistical uncertainty plus the systematic uncertainty from the background subtraction are shown as inner and outer bars respectively. The measured $\gamma +$ jet purity as a function of $m_{\gamma j}$ is presented in the bottom panel (black histogram): the statistical uncertainty of the purity measurement is reported as the inner error bar while the total uncertainty is shown as the outer error bar.



for the hadronisation, underlying event and detector effects are removed; and finally the corrections for the photon purity are changed within their estimated uncertainty. The largest absolute fitted signal from all variations of the nominal background sample discussed above is taken to be the spurious signal.

The spurious signal is evaluated at a number of hypothetical masses over a large search range. It is required to be less than 40% of the background statistical uncertainty, as quantified by the statistical uncertainty of the fitted spurious signal, anywhere in the investigated search range. In this way the impact of the systematic uncertainties due to background modelling on the analysis sensitivity is expected to be subdominant with respect to the statistical uncertainty. Functional forms that cannot meet this requirement are rejected. For different signal models, the functional form and fit range are determined separately. All considered functions with k up to two (four parameters) are found to fulfill the spurious-signal requirement when fitting in the range $1.1 < m_{\gamma j} < 6.0$ TeV for the q^* signal and $1.5\ (2.5) < m_{\gamma j} < 6.0\ (8.0)$ TeV for the RS1 (ADD) QBH signal. To further consolidate the choice of nominal background functional form, an F-test [232] is performed to determine if the change in the χ^2 value obtained by fitting the JETPHOX sample with an additional parameter is significant. The test indicates that the $k = 0$ (1) functional form with two (three) parameters can describe the present data sufficiently well over the entire fit range for the QBH (q^*) signal search, and there is no improvement by adding more parameters to the background fit function.

Given the fit range determined by the spurious signal test, the search is performed for the q^* (RS1 and ADD QBH) signal within the $m_{\gamma j}$ range above 1.5 (2.0 and 3.0) TeV, to account for the width of the expected signal. The estimated spurious signal for the selected functional form is converted into a spurious-signal cross-section (σ_{spur}), which is included as the uncertainty due to background modelling in the statistical analysis. The spurious-signal cross-section, and the ratio of the spurious-signal cross-section to its uncertainty ($\delta\sigma_{\text{spur}}$) and to the signal cross-section (σ_{S}) for the three benchmark models under investigation are given in Table 5.4 in the different search ranges. While both σ_{spur} and $\sigma_{\text{spur}}/\delta\sigma_{\text{spur}}$ decrease with the hypothesised signal mass, the ratio $\sigma_{\text{spur}}/\sigma_{\text{S}}$ increases with m_{q^*} or M_{th}, becoming as large as 15% in the case of excited quarks with $m_{q^*} = 6$ TeV.

Table 5.4: Spurious-signal cross-sections (σ_{spur}), and the ratio of the spurious-signal cross-sections to their uncertainties ($\delta\sigma_{\text{spur}}$) and to the signal cross-sections (σ_{S}) for the three benchmark models. The values of these quantities are given at the boundaries of the search range reported in the first row.

	q^*		RS1 QBH		ADD QBH	
Search boundaries [TeV]	1.5	6.0	2.0	6.0	3.0	8.0
σ_{spur} [fb]	3.9	1.1×10^{-2}	4.0	6.6×10^{-4}	8.7×10^{-2}	5.0×10^{-5}
$\sigma_{\text{spur}}/\delta\sigma_{\text{spur}}$ [%]	37	14	39	8	20	3
$\sigma_{\text{spur}}/\sigma_{\text{S}}$ [%]	0.16	15	1.0	7.5	0.0017	0.037

A similar test is performed to determine the functional form and fit ranges for the Gaussian-shaped signal with a 15% width. The test indicates that the same functional form and fit range as those used for the q^* signal are optimal for a wide-width Gaussian signal. The same functional form and mass range is used for all the Gaussian signals.

5.5.3 Statistical test

In this section, the procedure used to search for a new phenomenon in the context of a frequentist statistical test is presented. For purposes of discovering a new signal process, the null hypothesis (H_0) is defined as describing only known processes, here designated as background. This is to be tested against the alternative hypothesis (H_1), which includes both background and the sought after signal. When setting limits, the model with signal plus background plays the role of H_0, which is tested against the background-only hypothesis H_1.

To summarise the outcome of such a search the level of agreement of the observed data with a given hypothesis H is quantified by computing an observed p-value (p_{obs}), i.e. the probability, under assumption of H, of finding data of equal or greater incompatibility with the predictions of H. One can regard the hypothesis as excluded if its p_{obs} value is observed below a specific threshold. In this search, a profile-likelihood-ratio test statistic is used to quantify the level of compatibility between the data and either the SM background or the signal prediction, and to set limits on the presence of possible signal contributions in the $m_{\gamma j}$ distribution. The likelihood function $\mathcal{L}$ is built from a Poisson probability for the number of observed events (n) and expected events (N) in the selected sample:

$$\mathcal{L}(M, \sigma_{\mathrm{S}} \cdot B, \boldsymbol{\theta} | \{m_{\gamma j}\}_{i=1}^{n}) = \mathrm{Pois}(n | N(\sigma_{\mathrm{S}} \cdot B, \boldsymbol{\theta})) \times \left(\prod_{i=1}^{n} f(m_{\gamma j}^{i}, \boldsymbol{\theta}) \right) \times G(\boldsymbol{\theta}),$$

where M is the parameter to specify the MC signal channels (representing M_{th}, m_{q^*} or m_{G}), $\sigma_{\mathrm{S}} \cdot B$ is the signal cross-section multiplied by the branching ratio to a photon and a quark or a gluon, $\{m_{\gamma j}\}_{i=1}^{n}$ are the $\gamma +$jet system invariant masses in all observed events, $f(m_{\gamma j}^{i}, \boldsymbol{\theta})$ is the value of the probability density function of the invariant mass distribution evaluated for each candidate event i, and $\boldsymbol{\theta}$ are "nuisance parameters", whose values are known within systematic uncertainties. The $G(\boldsymbol{\theta})$ term collects the set of constraints on the nuisance parameters associated with the systematic uncertainties in the signal yield, in the spurious signal and in the resolution (only for Gaussian signals) and it is represented by normal distributions centred at zero and with unit variance.

The pdf of the $m_{\gamma j}$ distribution is given as the normalised sum of the signal and background pdfs:

$$f(m_{\gamma j}^{i}, \boldsymbol{\theta}) = \frac{1}{N} [N_{\mathrm{sig}}(\boldsymbol{\theta}_{\mathrm{yield}}) f_{\mathrm{sig}}(m_{\gamma j}^{i}) + N_{\mathrm{bg}} f_{\mathrm{bg}}(m_{\gamma j}^{i}, \boldsymbol{\theta}_{\mathrm{bkg}})],$$

where f_{sig} and f_{bg} are the normalised signal and background $m_{\gamma j}$ distributions described in the previous sections. The $\boldsymbol{\theta}_{\mathrm{yield}}$ are nuisance parameters associated with the signal yield uncertainties while $\boldsymbol{\theta}_{\mathrm{bkg}}$ are the nuisance parameters of the background shape. The expected number of events N is given by the sum of the expected numbers of signal events (N_{sig}) and background events (N_{bg}). The number of background events is extracted from a background-only fit to data. The N_{sig} term can be expressed as

$$N_{\mathrm{sig}}(\boldsymbol{\theta}_{\mathrm{yield}}) = N_{\mathrm{sig}}^{\mathrm{model}} + N_{\mathrm{sig}}^{\mathrm{spur}} = (\sigma_{\mathrm{S}} \cdot B \cdot A \cdot \varepsilon \cdot F(\boldsymbol{\delta}_{\varepsilon}, \boldsymbol{\theta}_{\varepsilon}) + \sigma_{\mathrm{spur}} \cdot \theta_{\mathrm{spur}}) \times \mathcal{L}_{\mathrm{int}} \times F(\delta_{\mathcal{L}}, \theta_{\mathcal{L}}),$$

where σ_{spur} and θ_{spur} are the spurious-signal cross-section described in Section 5.5.2 and its nuisance parameter while $\mathcal{L}_{\mathrm{int}}$ and $F(\delta_{\mathcal{L}}, \theta_{\mathcal{L}})$ are the integrated luminosity and its uncertainty. Apart from the spurious signal, which has a Gaussian uncertainty, all other systematic uncertainties are treated as log-normal. These uncertainties are incorporated into the likelihood (as described in Appendix A) by multiplying the relevant parameter of the statistical model by a factor $F(\delta_X, \theta_X) = \mathrm{e}^{\delta_X \theta_X}$, where δ_X is the size of the relative uncertainty. The parameter of interest in the fit to Gaussian-shaped resonances is the visible cross-section $\sigma_{\mathrm{S}} \cdot B \cdot A \cdot \varepsilon$ while that in the fit to q^* and QBH signals is $\sigma_{\mathrm{S}} \cdot B$. For the latter case, the additional nuisance parameters for the signal efficiency uncertainties $F(\boldsymbol{\delta}_\varepsilon, \boldsymbol{\theta}_\varepsilon)$ are included.

Table 5.5 summarises all the quantities entering in the likelihood. Table 5.6 summarises how the systematic uncertainties are implemented.

Table 5.5: List of all the parameters entering in the likelihood.

Parameter	Description
$\sigma_{\mathrm{S}} \cdot B$	parameter of interest
M_{th}, m_{q^*} or m_{G}	parameter of the model
$A \times \varepsilon$	total efficiency and acceptance
δ_{ID}	uncertainty on the photon-ID
$\delta_{\mathrm{trigger}}$	uncertainty on the trigger efficiency
δ_{pileup}	uncertainty due to pile-up reweighting effect on signals efficiency
$\delta_{\mathrm{isolation}}$	uncertainty on the isolation cut efficiency
δ_{MC}	uncertainty due to the statistical error on the efficiency from MC
σ_{spur}	spurious cross-section
$\mathcal{L}_{\mathrm{int}}$	luminosity after GRL
$\delta_{\mathcal{L}}$	uncertainty on L
p_1	background shape parameter
p_2	background shape parameter
p_3	background shape parameter
θ_i	pull variation of the auxiliary measurement

Table 5.6: List of the implementation of all the systematic uncertainties.

Quantity	Effective term
$\mathcal{L}_{\mathrm{int}}$	$\mathcal{L}_{\mathrm{int}} \times e^{(\delta_{\mathcal{L}}\theta_{\mathcal{L}})}$
ε	$\varepsilon \times e^{\delta_{\mathrm{trigger}}\theta_{\mathrm{trigger}} + \delta_{\mathrm{ID}}\theta_{\mathrm{ID}} + \delta_{\mathrm{isolation}}\theta_{\mathrm{isolation}} + \delta_{\mathrm{pileup}}\theta_{\mathrm{pileup}} + \delta_{\mathrm{MC}}\theta_{\mathrm{MC}}}$
N_S	$N_S^{model} + N_S^{spur} = (\sigma_{\mathrm{S}} \cdot B \times A \times \varepsilon + \sigma_{\mathrm{spur}} \cdot \theta_{\mathrm{spur}}) \times \mathcal{L}_{\mathrm{int}}$

The profile likelihood ratio is built as follows:

$$\lambda(\sigma) = \frac{\mathcal{L}(\sigma, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\sigma, \hat{\boldsymbol{\theta}})} \tag{5.2}$$

where $\hat{\sigma}$ and $\hat{\boldsymbol{\theta}}$ are respectively the values of $\sigma = \sigma_{\mathrm{S}} \cdot B$ (or $\sigma = \sigma_{\mathrm{S}} \cdot B \cdot A \cdot \epsilon$) and nuisance parameters that unconditionally maximise the likelihood, while $\hat{\hat{\boldsymbol{\theta}}}(\sigma)$ are the values of the nuisance parameters that maximise the likelihood on the condition that σ is held fixed to a given value. From Equation 5.2, one can see that $0 \leq \lambda \geq 1$, with λ near 1 implying good agreement between the data and the hypothesised value of σ.

The significance of a possible deviation from the SM prediction is estimated by computing the p_{obs} value, defined as the probability to observe, under the background model hypothesis, an excess at least as large as the one observed in data.

If the observed σ is consistent with zero for a particular signal hypothesis, a limit setting is performed with the following procedure. Upper limits on the visible cross-section ($\sigma_{\mathrm{S}} \cdot B \cdot A \cdot \varepsilon$) for the Gaussian-shaped resonances or on the signal cross-section times branching ratio to a photon and a quark or a gluon ($\sigma_{\mathrm{S}} \cdot B$) for the q^* and QBH signals are set at 95% confidence level (CL) using a modified frequentist method (CL$_s$) with a test statistic q constructed from the profile likelihood ratio λ and defined as follows:

$$q(\sigma) = \begin{cases} -2\log \frac{\mathcal{L}(\sigma,\hat{\hat{\boldsymbol{\theta}}}(\sigma))}{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \text{if } \hat{\sigma} < 0 \\ -2\log \frac{\mathcal{L}(\sigma,\hat{\hat{\boldsymbol{\theta}}}(\sigma))}{\mathcal{L}(\hat{\sigma},\hat{\boldsymbol{\theta}})} & \text{if } 0 \leq \hat{\sigma} \leq \sigma, \\ 0 & \text{if } \hat{\sigma} > \sigma \end{cases} \tag{5.3}$$

The reason for setting $q = 0$ for $\hat{\sigma} > \sigma$ is that when setting an upper limit, one would not regard data with $\hat{\sigma} > \sigma$ as representing less compatibility with σ than the data obtained, and therefore this is not taken as part of the rejection region of the test. Since $\sigma \geq 0$, if data is found such that $\hat{\sigma} < 0$, then the best level of agreement between the data and any physical value of σ occurs for $\sigma = 0$. This is the reason for using $\mathcal{L}(0, \hat{\hat{\boldsymbol{\theta}}}(0))$ instead of $\mathcal{L}(\hat{\sigma}, \hat{\boldsymbol{\theta}})$ in the denominator when $\hat{\sigma} < 0$.

In order to test a hypothesis of signal strength σ, the confidence level on the signal, $\mathrm{CL}_s(\sigma)$, is determined as

$$\mathrm{CL}_s(\sigma) = \frac{\mathrm{CL}_{s+b}}{\mathrm{CL}_b} = \frac{p_{s+b}(\sigma)}{1 - p_b(\sigma)},$$

where $p_{s+b}(\sigma)$ $(1 - p_b(\sigma))$ is the probability to obtain a value of $q(\sigma)$ larger than the observed value, $q_{\mathrm{obs}}(\sigma)$, when the data contains the tested signal (no signal). The value of $q_{\mathrm{obs}}(\sigma)$ is calculated from data using Equation 5.3.

The observed 95% CL upper limit on the signal strength, σ_{obs}, is then found by requiring

$$\mathrm{CL}_s(\sigma_{\mathrm{obs}}) = 0.05.$$

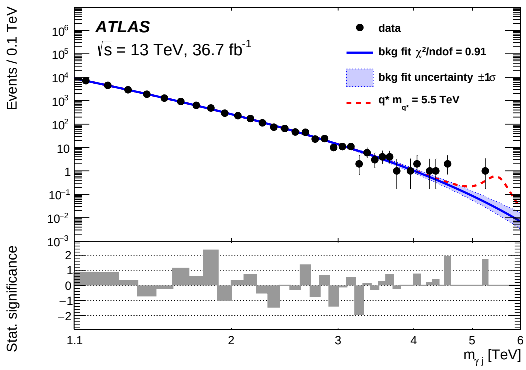
5.6 Results

The photon–jet invariant mass distributions obtained from the selected data are shown in Figure 5.9, together with the background-only fits using the model described in Section 5.5.2 and examples of the expected distributions from the background plus signal models under test. No significant deviation from the background-only prediction is observed in any of the distributions. The most significant excess is observed at 1.8 TeV for the assumption of a 2%-width Gaussian model for a local significance of 2.1 standard deviations.

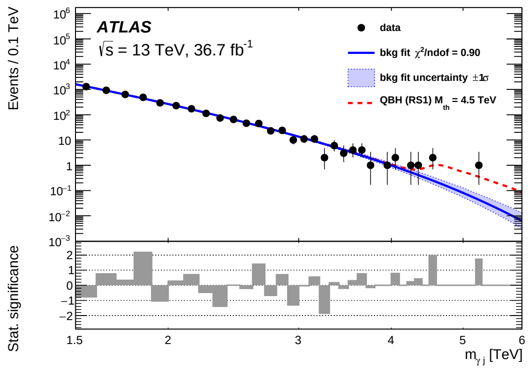
Limits are placed at 95% CL on the visible cross-section in the case of generic Gaussian-shaped resonances and on the production cross-section times branching ratio to a photon and a quark or gluon for the excited-quark and QBH signals. More detail on the procedure to set the limits, with respect to the previous section, is given next.

For each model, hypothetical mass, and assumed signal strength σ, pseudo-experiments are generated to obtain the pdf of q. Each pseudo-experiment consists in generating pseudo-data following the S+B or B-only models and calculating the ensuing values q_{SB} and q_{B} of the statistic q, taking into account not only the statistical but also the systematic

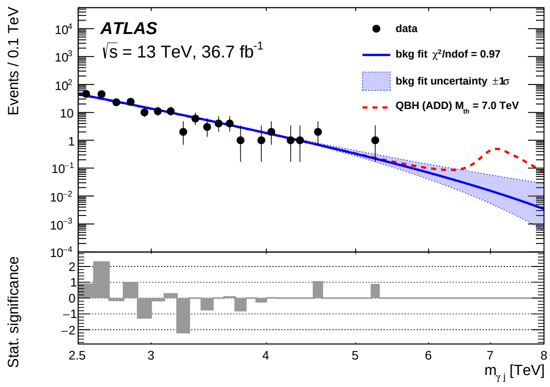
Figure 5.9: Distributions of the invariant mass of the γ + jet system of the observed events (dots) in 36.7 fb^{-1} of data at $\sqrt{s} = 13$ TeV and fits to the data (solid lines) under the background-only hypothesis for searches in the (a) excited quarks, (b) QBH (RS1) with $n = 1$ and (c) QBH (ADD) with $n = 6$ models. The $\pm 1\sigma$ uncertainty in the background prediction originating from the uncertainties in the fit function parameter values is shown as a shaded band around the fit. The predicted signal distributions (dashed lines) for the q^* model with $m_{q^*} = 5.5$ TeV and the QBH model with $M_{\text{th}} = 4.5$ (7.0) TeV based on RS1 (ADD) are shown on top of the background predictions. The bottom panels show the bin-by-bin significances of the data–fit differences, considering only statistical uncertainties.



(a)

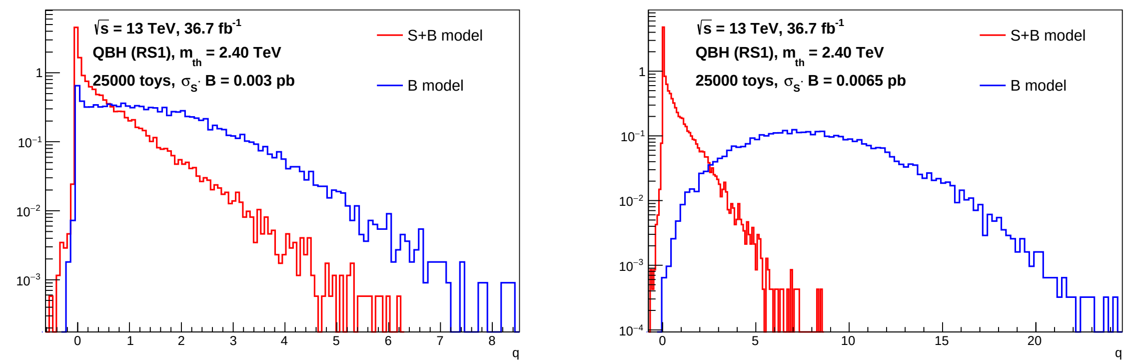


(b)



(c)

Figure 5.10: Sampling distribution of q under the background-only hypothesis (blue) and under the signal+background hypothesis (red) for (a) $\sigma_{\mathrm{S}} \cdot B = 0.003\,\mathrm{pb}$ and (b) $\sigma_{\mathrm{S}} \cdot B = 0.0065\,\mathrm{pb}$. The pdfs were generated by pseudo-experiments for the QBH RS1 model with $M_{\mathrm{th}} = 2.4\,\mathrm{TeV}$.



(a) (b)

fluctuations described by the distribution function of the respective nuisance parameters. Samples of 12500 and 25000 events are generated to construct respectively the q_{B} and q_{SB} distributions. More pseudo-experiments are needed for the S+B case because the analysis is sensitive to the tail of its pdf. As an example, the result obtained for the QBH RS1 model with $\mathrm{M_{th}} = 2.4$ TeV for two values of σ, $0.003\,\mathrm{pb}$ and $0.0065\,\mathrm{pb}$, is shown in Figure 5.10. It can be observed that q serves to disentangle data compatible with the background-only hypothesis from data compatible with new physics, and that the distinction between hypotheses is improved with increasing signal strength.

The observed signal confidence level, $\mathrm{CL}_s(\sigma)$, is obtained by calculating $q_{\mathrm{obs}}(\sigma)$ from the data and evaluating p_{s+b}^{obs} and p_b^{obs}:

$$\begin{aligned} p_{s+b}^{\mathrm{obs}}(\sigma) &= \int_{q_{\mathrm{obs}}}^{\infty} f\left(q_{SB}|\sigma\right) \mathrm{d}q_{SB} \\ p_{b}^{\mathrm{obs}}(\sigma) &= \int_{-\infty}^{q_{\mathrm{obs}}} f\left(q_{B}|\sigma\right) \mathrm{d}q_{B} \end{aligned} \tag{5.4}$$

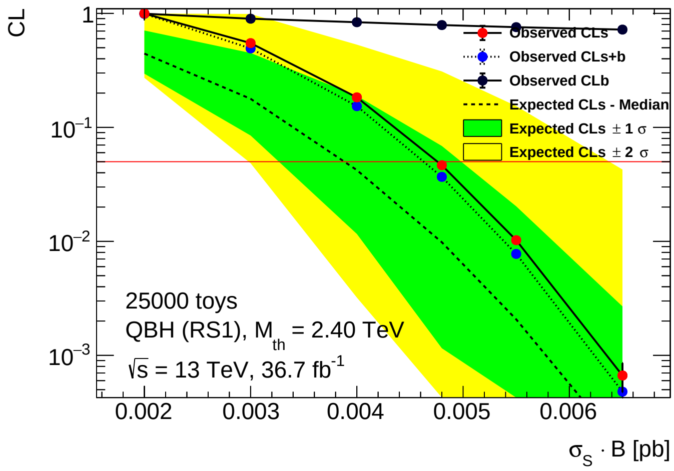
This is illustrated in Figure 5.11, which plots in red circles the results for the observed CL_s as a function of the signal strength for the QBH RS1 model with $\mathrm{M_{th}} = 2.4\,\mathrm{TeV}$.

The observed 95% CL upper limit for a given model and mass, σ_{obs}, is defined as the value of the signal strength for which $\mathrm{CL}_s(\sigma_{\mathrm{obs}}) = 0.05$. This result is obtained numerically by interpolating $\mathrm{CL}_s(\sigma)$, as shown by the solid line in Figure 5.11. The results of σ_{obs} as a function of the signal mass for the three models and the generic Gaussian-shaped resonances considered in this analysis are shown in Figures 5.13 and 5.14, respectively.

In addition to the observed values, it is important to determine the sensitivity of the experiment. This is quantified by σ_{exp}, the expected 95% CL upper limit on the signal cross-section under the assumption that data does not contain signal. It is also relevant to establish by how much σ_{exp} can fluctuate due to the statistical and systematic uncertainties, in particular in what range it should lie with 68.3% and 95.5% probabilities, the so-called $\pm 1\,\sigma$ and $\pm 2\,\sigma$ ranges.

In order to summarise quantitatively the background-only distribution of q one evaluates the sample quantiles, $q_{\text{-}2}$, $q_{\text{-}1}$, q_0, q_1, q_2, corresponding respectively to the probabilities $P_{\text{-}2} = 0.023$, $P_{\text{-}1} = 0.159$, $P_0 = 0.5$, $P_1 = 0.841$, $P_2 = 0.977$, rounded to a precision of three

Figure 5.11: Observed and expected (median, $\pm 1\sigma$, and $\pm 2\sigma$) CL$_s$ values as a function of $\sigma_{\mathrm{S}} \cdot B$ for QBH RS1 model with $\mathrm{M_{th}} = 2.4\,\mathrm{TeV}$. The crossings of the horizontal line with the curves representing the different cases indicate the respective cross-section bounds at 95% CL.



decimal digits.[2] With this definition, q_0 corresponds to the median (denoted also as q_{exp}) and $q_{\text{-}1}$–q_1 ($q_{\text{-}2}$–q_2) to the $1\,\sigma$ ($2\,\sigma$) range around the median. The respective p-values are calculated with Equation 5.4 replacing q_{obs} by q_i with $i = \text{-}2, \text{-}1, 0, 1, 2$. Figure 5.12 illustrates the determination of p_{s+b} for q_{exp} and q_1.

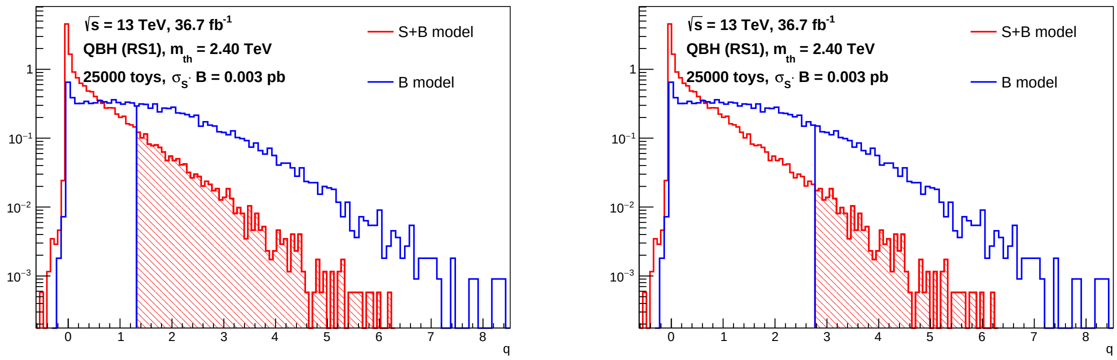
Figure 5.11 illustrates the CL$_s$ values calculated for q_0, $q_{\text{-}1}$, q_1, $q_{\text{-}2}$, q_2 for 6 assumed signal strengths for the QBH RS1 model with $\mathrm{M_{th}} = 2.4\,\mathrm{TeV}$. These values are logarithmically interpolated and respectively indicated with a dotted line, the borders of the green band and the outer borders of the yellow band. The expected 95% CL upper limit $\sigma_{\mathrm{exp}} = \sigma_0$, as well as $\sigma_1, \sigma_{\text{-}1}$ ($\sigma_2, \sigma_{\text{-}2}$), the $\pm 1\,\sigma$ ($\pm 2\,\sigma$) confidence interval for σ_{exp}, are numerically found from the intersections of the respective CL$_s(\sigma)$ curves and the horizontal CL$_s = 0.05$ line. Similar plots are obtained for each model and each assumed signal mass. The results for the three theoretical models and the generic Gaussian-shaped resonances are summarised on Figures 5.13 and 5.14, respectively, where the results of σ_{obs}, σ_{exp} and the $\pm 1\,\sigma$ and $\pm 2\,\sigma$ ranges of σ_{exp} are plotted as a function of the signal mass.

The observed and expected lower limits on the signal mass are obtained from the intersection of the signal strength predicted by the respective model (red solid lines) and the σ_{obs} and σ_{exp} curves, as shown in Figure 5.14.

The Gaussian signals are excluded for visible cross-sections above 0.25–1.1 fb (0.08–0.2 fb), depending on the width, at a mass m_{G} of 3 TeV (5 TeV). In the case of the benchmark signal models considered in this analysis, the presence of a signal with a mass below 5.3, 4.4 and 7.1 TeV for the excited quarks, RS1 and ADD QBHs, respectively, can be excluded at 95% CL. The limits improve on those in Ref. [213] by about 0.9, 0.6 and 0.9 TeV for the excited quarks, RS1 and ADD QBHs, respectively.

[2] $P_i = \int_{-\infty}^{q_i} f(q_{\mathrm{B}}|\sigma)\,\mathrm{d}q_{\mathrm{B}}$.

Figure 5.12: Sampling distribution of q under the background-only hypothesis (blue) and under the signal+background hypothesis (red) for the QBH RS1 model with $M_{\text{th}} = 2.4\,\text{TeV}$ and $\sigma_{\text{S}} \cdot B = 0.003\,\text{pb}$. The pdfs were generated by pseudo-experiments. The shaded area illustrates the p_{s+b} values using (a) q_{exp} and (b) $q_{\text{exp}} + 1\sigma$ (q_1). The two q values are shown by vertical blue lines.



(a) (b)

Figure 5.13: Observed (solid lines) and expected (dotted lines) 95% CL upper limits on the visible cross-sections $\sigma_{\text{S}} \cdot B \cdot A \cdot \varepsilon$ in 36.7 fb^{-1} of data at $\sqrt{s} = 13\,\text{TeV}$ as a function of the mass m_{G} of the Gaussian resonances with three different Gaussian widths between 2% and 15%. The calculation is performed using ensemble tests at mass points separated by 100 GeV over the search range.

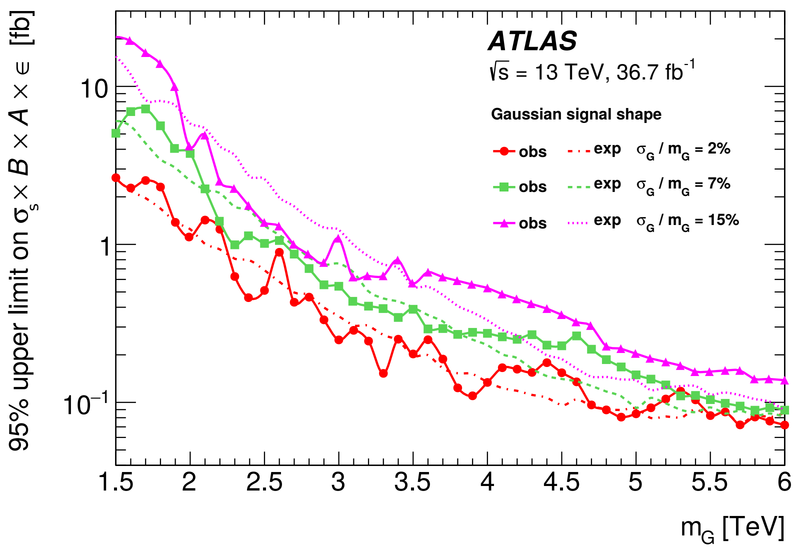
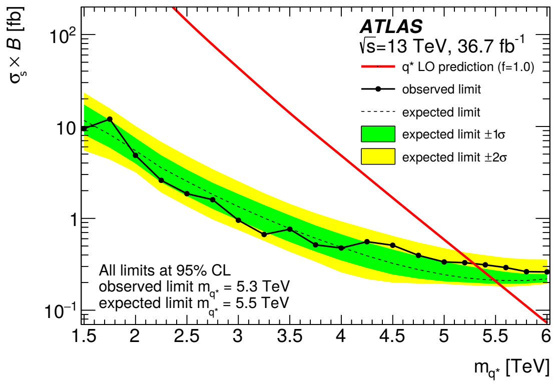
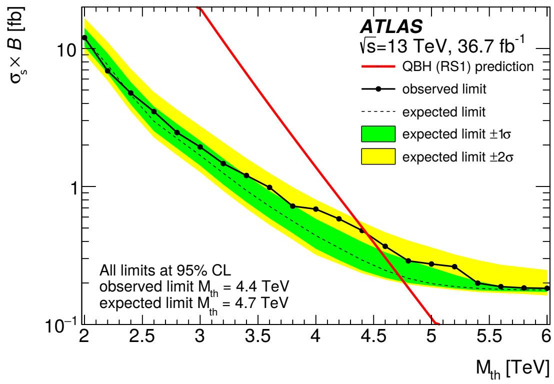


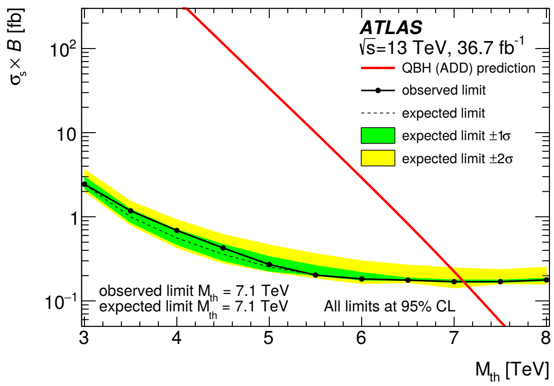
Figure 5.14: Observed 95% CL upper limits (solid line with dots) on the production cross-section times branching ratio $\sigma_S \cdot B$ to a photon and a quark or gluon in 36.7 fb^{-1} of data at $\sqrt{s}$ = 13 TeV for the (a) excited-quarks, (b) QBH (RS1) with $n = 1$ and (c) QBH (ADD) with $n = 6$ models. The limits are placed as a function of m_{q^*} for the excited quarks and $M_{\rm th}$ for the QBH signals. The calculation is performed using ensemble tests at mass points separated by 200 (500) GeV for the RS1 (ADD) model over the search range. For the q^* model the step size is 250 GeV up to 5 TeV and then 200 GeV up to 6 TeV. The limits expected if a signal is absent (dashed lines) are shown together with the $\pm 1\sigma$ and $\pm 2\sigma$ intervals represented by the green and yellow bands, respectively. The theoretical predictions of $\sigma_S \cdot B$ for the respective benchmark signals are shown by the red solid lines. The intersection of the theoretical prediction line and the observed (expected) upper limit yields the observed (expected) lower limit on the mass of the excited quark or $M_{\rm th}$ for QBH signals.



(a)



(b)



(c)

Conclusions

The GS corrections are part of the jet calibrations used in ATLAS called jet energy scale. The GS corrections are derived in MC for jets reconstructed with the anti-k_t algorithm with a jet radius parameter value of $R = 0.4$ using different type of inputs to the jet algorithm, and use information about how the jet deposits energy in the calorimeter as well as jet fragmentation properties that are characterised based on tracking variables to improve the performance. The jet resolution improves as more corrections are applied. The improvement observed for AntiKt4LCTopo jets, for which only the tracking and N_{segments}-based corrections are applied, was found to be smaller. The GS calibration achieves a reduction of the flavour dependence on the jet response. The average jet response remains almost unchanged, except for few bins at low $p_{\text{T}}^{\text{truth}}$ where differences of 1–2% are observed.

Inclusive jet cross-sections in proton-proton collisions at $\sqrt{s} = 13$ TeV are measured for jets reconstructed with the anti-k_t algorithm with a jet radius parameter value of $R = 0.4$. The measurement uses data collected at the LHC with the ATLAS detector during 2015 corresponding to an integrated luminosity of 3.2 fb^{-1}. The inclusive jet cross-sections are measured double-differentially in the jet transverse momentum and jet rapidity in a kinematic region between 100 GeV and 3.5 TeV with $|y| < 3$. The dominant systematic uncertainty arises from the jet energy calibration. A quantitative comparison of the measurement to fixed-order NLO QCD calculations, corrected for non-perturbative and electroweak effects, shows overall fair agreement (with p-values in the percent range) when considering jet cross-sections in individual jet rapidity bins independently. A significant tension (with p-values $\ll 10^{-3}$) between data and theory is observed when considering data points from all jet transverse momentum and rapidity regions. No significant differences between the inclusive jet cross-sections and the fixed-order NNLO QCD calculations corrected for non-perturbative and electroweak effects are observed when using $p_{\text{T}}^{\text{jet}}$ as the QCD scale. The NNLO pQCD predictions using $p_{\text{T}}^{\text{max}}$ as the scale overestimate the measured inclusive jet cross-sections.

A search is performed for new phenomena in events having a photon with high transverse momentum and a jet collected in 36.7 fb^{-1} of pp collision data at a centre-of-mass energy of $\sqrt{s} = 13$ TeV recorded with the ATLAS detector at the LHC. The invariant mass distribution of the γ + jet system above 1.1 TeV is used in the search for localised excesses of events. No significant deviation is found. Limits are set on the visible cross-section for generic Gaussian-shaped resonances and on the production cross-section times branching ratio for signals predicted in models of excited quarks or quantum black holes. The data exclude, at 95% CL, the mass range below 5.3 TeV for the excited quarks and 7.1 (4.4) TeV for the quantum black holes with six (one) extra dimensions in the Arkani-Hamed–Dimopoulos–Dvali (Randall–Sundrum) model. These limits supersede the previous ATLAS exclusion limits for excited quarks and quantum black holes in the γ + jet final state.

Appendices

Appendix A

How to incorporate systematic uncertainties in statistical models

In order to incorporate a systematic uncertainty of relative size δ to a parameter p, which the likelihood depends on, the parameter p is replaced by:

$$p \cdot \mathrm{F}\,(\delta, \theta) \cdot \mathrm{G}(\theta),$$

where $\mathrm{F}\,(\delta, \theta)$ depends on the probability density function of the parameter p and θ is a nuisance parameter, a random variable, whose probability density function $\mathrm{G}(\theta)$ is the standard normal distribution $\mathrm{N}(0, 1)$:

$$\mathrm{G}(\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-\theta^2}{2}\right)$$

In this analysis there only appear Gaussian and log-normal uncertainties. In this appendix, it is shown that the corresponding $\mathrm{F}\,(\delta, \theta)$ functions are:

$$\begin{aligned} \mathrm{F}_{\mathrm{G}}\,(\delta, \theta) &= 1 + \delta \cdot \theta \\ \mathrm{F}_{\mathrm{LN}}\,(\delta, \theta) &= \exp\left(\sqrt{\log(1+\delta^2)}\,\theta\right) \end{aligned}$$

A.1 Gaussian case

If θ has a standard normal distribution then it is easily shown that the random variable x, defined as $x \equiv p\,(1 + \delta \cdot \theta)$, has a Gaussian distribution with mean p and width $\delta \cdot p$:

$$\begin{aligned} \theta &\sim \mathrm{N}(0, 1) \\ \delta \cdot \theta &\sim \mathrm{N}(0, \delta) \\ 1 + \delta \cdot \theta &\sim \mathrm{N}(1, \delta) \\ p\,(1 + \delta \cdot \theta) &\sim \mathrm{N}(p, \delta \cdot p) \end{aligned}$$

A.2 Log-normal case

To describe the systematic uncertainty on a parameter that is positive by definition, it is common to use a log-normal distribution $\mathrm{LN}(\mu, \sigma^2)$:

$$\mathrm{LN}(\mu, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right),$$

whose median and variance are:

$$\begin{aligned}\text{Median}(x) &= \exp(\mu)\\ \text{Variance}(x) &= [\exp\left(\sigma^2\right) - 1]\exp\left(2\mu + \sigma^2\right)\end{aligned}$$

It is shown next that if θ has a standard normal distribution, then the random variable x, defined as $x \equiv p\exp(S\cdot\theta)$, has a log-normal distribution $\text{LN}(\ln p, S^2)$ with:

$$\begin{aligned}\text{Median}(x) &= \exp(\ln p) = p\\ \text{Variance}(x) &= [\exp\left(S^2\right) - 1]\exp\left(2\ln p + S^2\right) = [\exp\left(S^2\right) - 1]p^2\exp\left(S^2\right)\end{aligned}$$

The probability density function of x ($\text{f}_x(t)$) satisfies $\text{f}_x(t) = \frac{d\text{F}_x}{dt}$, where F_x is the cumulative distribution function, defined as $\text{F}_x(t) = \text{P}(x \leq t)$. Therefore:

$$\begin{aligned}\text{F}_x &= \text{P}\left(x \leq t\right)\\ &= \text{P}\left(p\exp\left(S\cdot\theta\right) \leq t\right)\\ &= \text{P}\left(\exp\left(S\cdot\theta\right) \leq \frac{t}{p}\right)\\ &= \text{P}\left(S\cdot\theta \leq \ln\frac{t}{p}\right)\\ &= \text{P}\left(\theta \leq \frac{1}{S}\ln\frac{t}{p}\right) = \text{F}_\theta\left(\frac{1}{S}\ln\frac{t}{p}\right)\end{aligned}$$

Then:

$$\begin{aligned}\text{f}_x(t) = \frac{d\text{F}_x}{dt} &= \text{f}_\theta\left(\frac{1}{S}\ln\frac{t}{p}\right)\frac{d}{dt}\left[\frac{1}{S}\ln\frac{t}{p}\right]\\ &= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{1}{S}\ln\frac{t}{p}\right]^2\right)\cdot\frac{1}{S\cdot t}\\ &= \frac{1}{\sqrt{2\pi}\cdot S\cdot t}\exp\left(-\frac{1}{2}\left(\frac{\ln t - \ln p}{S}\right)^2\right) = \text{LN}(\ln p, S^2)\end{aligned}$$

For the case $S = \sqrt{\ln(1+\delta^2)}$:

$$\begin{aligned}\text{Variance}(x) = \sigma^2(x) &= [(1+\delta^2) - 1]p^2(1+\delta^2)\\ &= \delta^2 p^2(1+\delta^2)\end{aligned}$$

Since the size of the relative uncertainties appearing in this thesis are small ($\delta \ll 1$), two approximations are frequently done in ATLAS:

$$\begin{aligned}x &= p\ \exp\left(\sqrt{\log(1+\delta^2)}\,\theta\right) \cong p\ \exp\left(\delta\cdot\theta\right) + \mathcal{O}(\delta^3)\\ \text{Variance}(x) &= \delta^2p^2(1+\delta^2) \cong (\delta\cdot p)^2 \Rightarrow \sigma(x) = \delta\cdot p\end{aligned}$$

In summary, $x = p\exp(\delta\cdot\theta)$ with $\theta \sim \text{N}(0,1)$ is a good approximation of, as long as the errors are small, a log-normal distributed random variable with:

$$\begin{aligned}\text{Median}(x) &= p\\ \sigma(x) &= \delta\cdot p\end{aligned}$$

Bibliography

[1] K. Nishijima, *Charge Independence Theory of V Particles*, Progress of Theoretical Physics **13** (1955) 285.

[2] M. Gell-Mann, *The interpretation of the new particles as displaced charge multiplets*, Il Nuovo Cimento (1955-1965) **4** (1956) 848.

[3] F. Halzen and A. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, Wiley Text Books, 1984.

[4] M. Peskin and D. Schroeder, *Quantum field theory*, Addison-Wesley Advance Book Program, 1996.

[5] CDF Collaboration, *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74** (1995) 2626.

[6] D0 Collaboration, *Search for High Mass Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8\,TeV$*, Phys. Rev. Lett. **74** (1995) 2422, arXiv: `hep-ex/9411001`.

[7] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[8] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Sov. Phys. JETP **7** (1958) 172, [Zh. Eksp. Teor. Fiz.34,247(1957)].

[9] Z. Maki, M. Nakagawa, S. Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870, [,34(1962)].

[10] Particle Data Group, *Review of Particle Physics*, Chin. Phys. C **40** (2016) 100001.

[11] D. J. Griffiths, *Introduction to elementary particles*, 2008, ISBN: 9783527406012.

[12] F. J. Dyson, *The Radiation Theories of Tomonaga, Schwinger, and Feynman*, Phys. Rev. **75** (1949) 486.

[13] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[14] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13** (1964) 168.

[15] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[16] UA2 Collaboration,
Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider,
Phys. Lett. B **122** (1983) 476.

[17] UA1 Collaboration, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c^2 at the CERN SPS Collider*, Phys. Lett. B **126** (1983) 398.

[18] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*,
Phys. Lett. **8** (1964) 214.

[19] J. J. J. Kokkedee, *The Quark Model*, Benjamin, 1969.

[20] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and collider physics*,
Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **8** (1996) 1.

[21] ATLAS Collaboration,
Determination of the strong coupling constant α_s from transverse energy-energy correlations in multijet events at $\sqrt{s} = 8$ TeV using the ATLAS detector,
Eur. Phys. J. C **77** (2017) 872, arXiv: `1707.02562 [hep-ex]`.

[22] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*,
Phys. Rev. Lett. **30** (1973) 1343.

[23] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*,
Phys. Rev. Lett. **30** (1973) 1346.

[24] H. Georgi and S. L. Glashow, *Unity of All Elementary-Particle Forces*,
Phys. Rev. Lett. **32** (1974) 438.

[25] Y. Nambu,
Quasi-Particles and Gauge Invariance in the Theory of Superconductivity,
Phys. Rev. **117** (1960) 648.

[26] J. Goldstone, *Field theories with superconductor solutions*,
Il Nuovo Cimento (1955-1965) **19** (1961) 154.

[27] P. W. Anderson, *Plasmons, Gauge Invariance, and Mass*,
Phys. Rev. **130** (1963) 439.

[28] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*,
Phys. Rev. Lett. **13** (1964) 508.

[29] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*,
Phys. Lett. **12** (1964) 132.

[30] F. Englert and R. Brout,
Broken Symmetry and the Mass of Gauge Vector Mesons,
Phys. Rev. Lett. **13** (1964) 321.

[31] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble,
Global Conservation Laws and Massless Particles,
Phys. Rev. Lett. **13** (1964) 585.

[32] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*,
Phys. Rev. **145** (1966) 1156.

[33] A. Salam, *Elementary Particle Physics: Relativistic Groups and Analyticity*,
Eight Nobel Symposium, 1968.

[34] *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s}$ *= 540-GeV*, Phys. Lett. B **122** (1983) 103.

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

[36] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[37] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: `1207.7214 [hep-ex]`.

[38] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: `1207.7235 [hep-ex]`.

[39] J. C. Collins and D. E. Soper, *The Theorems of Perturbative QCD*, Annual Review of Nuclear and Particle Science, 1987.

[40] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1, arXiv: `hep-ph/0409313 [hep-ph]`.

[41] J. Rojo et al., *The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II*, J. Phys. G **42** (2015) 103103, arXiv: `1507.00556 [hep-ph]`.

[42] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, Nucl. Phys. B **126** (1977) 298.

[43] T. H. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **02** (2009) 007, arXiv: `0811.4622 [hep-ph]`.

[44] M. A. Dobbs et al., "Les Houches guidebook to Monte Carlo generators for hadron collider physics", 2004 411, arXiv: `hep-ph/0403045 [hep-ph]`.

[45] T. Sjostrand, "Monte Carlo Generators", 2006 51, arXiv: `hep-ph/0611247 [hep-ph]`.

[46] S. Weinberg, *Gauge Hierarchies*, Phys. Lett. B **82** (1979) 387.

[47] M. J. G. Veltman, *The Infrared - Ultraviolet Connection*, Acta Phys. Polon. B **12** (1981) 437.

[48] C.H. Llewellyn Smith, G. G. Ross, *The Real Gauge Hierarchy Problem*, Phys. Lett. B **105** (1981) 38.

[49] T. S. van Albada, J. N. Bahcall, K. Begeman, and R. Sancisi, *Distribution of dark matter in the spiral galaxy NGC 3198*, ApJ **295:305-313** (1985).

[50] N. Kaiser and G. Squires, *Mapping the dark matter with weak gravitational lensing*, ApJ **404:441-450** (1993).

[51] Adam, R. et al., *Planck 2015 results. I. Overview of products and scientific results*, Astron. Astrophys. **594** (2016) A1, arXiv: 1502.01582 [astro-ph.CO].

[52] T. Sjostrand, S. Mrenna and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, arXiv: hep-ph/0603175 [hep-ph].

[53] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv: hep-ph/0011363 [hep-ph].

[54] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C **58** (2008) 639, arXiv: 0803.0883 [hep-ph].

[55] S. Höche, *Introduction to parton-shower event generators*, SLAC-PUB-16160, 2015 235, arXiv: 1411.4085 [hep-ph].

[56] S. Alioli, K. Hamilton, P. Nason, C. Oleari, E. Re, *Jet pair production in POWHEG*, JHEP **04** (2011) 081, arXiv: 1012.3380 [hep-ex].

[57] Paolo Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[58] S. Hoeche et al., "Matching parton showers and matrix elements", 2005 288, arXiv: hep-ph/0602031 [hep-ph].

[59] S. Catani et al., *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, arXiv: hep-ph/0109231 [hep-ph].

[60] F. Krauss, *Matrix elements and parton showers in hadronic interactions*, JHEP **08** (2002) 015, arXiv: hep-ph/0205283 [hep-ph].

[61] B. Andersson et al., *Parton Fragmentation and String Dynamics*, Phys. Rept. **97** (1983) 31.

[62] B. R. Webber, *A QCD Model for Jet Fragmentation Including Soft Gluon Interference*, Nucl. Phys. B **238** (1984) 492.

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

[64] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, ATL-PHYS-PUB-2012-003, 2012, URL: https://cds.cern.ch/record/1474107.

[65] Peter Skands, Stefano Carrazza and Juan Rojo, *Tuning PYTHIA 8.1: the Monash 2013 tune*, Eur. Phys. J. C **74** (2014) 3024, arXiv: 1404.5630 [hep-ph].

[66] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823, arXiv: 1005.4568 [physics.ins-det].

[67] S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[68] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[69] *Summary of the analysis of the 19 September 2008 incident at the LHC*, 2008, URL: https://cds.cern.ch/record/1135729.

[70] *Restarting the LHC: Why 13 Tev?*, (2014),
URL: http://cds.cern.ch/record/1998739.

[71] F. Marcastel,
CERN's Accelerator Complex. La chaîne des accélérateurs du CERN, 2013,
URL: https://cds.cern.ch/record/1621583.

[72] J. Caron, *LHC Layout.. Schema general du LHC.*, LHC-PHO-1997-060, 1997,
URL: http://cds.cern.ch/record/841573.

[73] ALICE Collaboration, *The ALICE experiment at the CERN LHC*,
JINST **3** (2008) S08002.

[74] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[75] S. van der Meer, *Calibration of the effective beam height in the ISR*,
CERN-ISR-PO-68-31. ISR-PO-68-31, 1968,
URL: http://cds.cern.ch/record/296752.

[76] ATLAS Collaboration, *ATLAS TWiki*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[77] ATLAS Collaboration, *ATLAS Policy on Radiation Tolerant Electronics: ATLAS Radiation Tolerance Criteria*, ATC-TE-QA-0001, 1997.

[78] J. Pequenao and P. Schaffner,
An computer generated image representing how ATLAS detects particles,
CERN-EX-1301009, 2013, URL: https://cds.cern.ch/record/1505342.

[79] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*,
CERN-LHCC-201-013, ATLAS-TDR-19, 2010,
URL: http://cds.cern.ch/record/1291633.

[80] G. F. Knoll, *Radiation detection and measurement*, 4th ed.,
Wiley, New Work, NY, 2010.

[81] W. E. Cleland and E. G. Stern, *Signal processing considerations for liquid ionization calorimeters in a high rate environment*,
Nucl. Instrum. Meth. A **338** (1994) 467.

[82] H. Abreu et al.,
Performance of the electronic readout of the ATLAS liquid argon calorimeters,
JINST **5** (2010) P09003.

[83] ATLAS Tile Calorimeter Collaboration, *Signal reconstruction of the ATLAS Hadronic Tile Calorimeter: implementation and performance*,
J. Phys. Conf. Ser. **293:012056** (2011).

[84] M. Aharrouche et al., *Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam*,
Nucl. Instrum. Meth. A **614** (2010) 400.

[85] J. Colas et al.,
Response uniformity of the ATLAS liquid argon electromagnetic calorimeter,
Nucl. Instrum. Meth. A **582** (2007) 429, arXiv: 0709.1094 [physics.ins-det].

[86] M. Aharrouche et al., *Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam*,
Nucl. Instrum. Meth. A **568** (2006) 601, arXiv: physics/0608012v1.

[87] 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].

[88] A. Artamonov et al., *The ATLAS Forward Calorimeter*, JINST **3** (2008) P02010, URL: http://stacks.iop.org/1748-0221/3/i=02/a=P02010.

[89] ATLAS Collaboration, *The ATLAS Diamond Beam Monitor: Luminosity detector at the LHC*, Nucl. Instrum. Meth. A **824** (2016) 459.

[90] ATLAS LUCID Collaboration, *ATLAS LUCID detector upgrade for LHC Run 2*, PoS **EPS-HEP2015** (2015) 275.

[91] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[92] R. R. Wilson, *The Tevatron*, Phys. Today **30N10** (1977) 23.

[93] W.J. Stirling, private communication.

[94] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, (2009), arXiv: 0901.0512 [hep-ex].

[95] ATLAS Collaboration, *The evolution of the Trigger and Data Acquisition System in the ATLAS experiment*, J. Phys. Conf. Ser. **523** (2014) 012019.

[96] ATLAS Collaboration, *ATLAS High-Level Trigger algorithms for Run-2 data-taking*, ATL-DAQ-PROC-2015-018, 2015, URL: https://cds.cern.ch/record/2016651.

[97] ATLAS TDAQ Collaboration, *The ATLAS Data Acquisition and High Level Trigger system*, JINST **11** (2016) P06008.

[98] ATLAS Collaboration, *The ATLAS Data Acquisition System: from Run 1 to Run 2*, Nucl. Part. Phys. Proc. **273-275** (2016) 939.

[99] ATLAS Collaboration, *ATLAS TWiki*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[100] W. Lampl et al., *Calorimeter clustering algorithms: Description and performance*, 2008.

[101] S. Ask et al., *The ATLAS Central Level-1 Trigger Logic and TTC System*, JINST **3** (2008).

[102] G. Anders et al., *The upgrade of the ATLAS Level-1 Central Trigger Processor*, JINST **8** (2013) C01049.

[103] S. Ask et al., *The ATLAS central level-1 trigger logic and TTC system*, JINST **3** (2008) P08002.

[104] ATLAS Collaboration, *ATLAS TWiki*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults.

[105] ATLAS Collaboration,
Performance of missing transverse momentum reconstruction with the ATLAS detector in the first proton–proton collisions at $\sqrt{s} = 13$ *TeV*,
ATL-PHYS-PUB-2015-027, 2015, URL: https://cds.cern.ch/record/2037904.

[106] Particle Data Group, *Review of Particle Physics*, Phys. Rev. D **86** (2012) 010001.

[107] A. Salzburger on behalf of the ATLAS Collaboration,
Optimisation of the ATLAS Track Reconstruction Software for Run-2,
J. Phys. Conf. Ser. **664** (2015) 072042.

[108] CMS Collaboration, *Adaptive Vertex Fitting*, CMS-NOTE-2007-008, 2007,
URL: https://cds.cern.ch/record/1027031.

[109] G. Borissov et al. on behalf of the ATLAS Collaboration,
ATLAS strategy for primary vertex reconstruction during Run-2 of the LHC,
J. Phys. Conf. Ser. **664** (2015) 072041.

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

[111] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton–proton collision data*,
ATLAS-CONF-2014-032, 2014, URL: https://cds.cern.ch/record/1706245.

[112] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, URL: https://cds.cern.ch/record/1449796.

[113] A. Hoecker et al., *TMVA: Toolkit for Multivariate Data Analysis*,
PoS ACAT **040** (2007), arXiv: physics/0703039,
URL: http://tmva.sourceforge.net/.

[114] ATLAS Collaboration, *Photon identification in 2015 ATLAS data*,
ATL-PHYS-PUB-2016-014, 2016, URL: https://cds.cern.ch/record/2203125.

[115] ATLAS Collaboration, *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-1 data*, (2016),
arXiv: 1606.01813 [hep-ex].

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

[117] 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].

[118] B. Lenzi, R. Nicolaidou, and S. Hassani, *TrackInCaloTools: A package for measuring muon energy loss and calorimetric isolation in ATLAS*,
J. Phys. Conf. Ser. **219** (2010) 032049.

[119] J. Illingworth and J. Kittler, *A survey of the Hough transform*,
Computer vision, graphics, and image processing, 1988.

[120] J. E. Huth et al., “Toward a standardization of jet definitions”, *1990 DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90) Snowmass, Colorado, June 25-July 13, 1990*, 1990 0134, URL: http://lss.fnal.gov/cgi-bin/find_paper.pl?conf-90-249.

[121] S. Catani et al., *Longitudinally invariant K_t clustering algorithms for hadron hadron collisions*, Nucl. Phys. B **406** (1993) 187.

[122] Y. L. Dokshitzer, *Better jet clustering algorithms*, JHEP **08** (1997) 001, arXiv: hep-ph/9707323 [hep-ph].

[123] M. Wobisch and T. Wengler, “Hadronization corrections to jet cross-sections in deep inelastic scattering”, *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, 1998 270, arXiv: hep-ph/9907280 [hep-ph], URL: https://inspirehep.net/record/484872/files/arXiv:hep-ph_9907280.pdf.

[124] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (7 1993) 3160, URL: https://link.aps.org/doi/10.1103/PhysRevD.48.3160.

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

[126] M. Dasgupta, M. Lorenzo, and G. P. Salam, *Non-perturbative QCD effects in jets at hadron colliders*, JHEP **02** (2008) 055, arXiv: 0712.3014 [hep-ph].

[127] C. Buttar et al., “Standard Model Handles and Candles Working Group: Tools and Jets Summary Report”, *Physics at TeV colliders, La physique du TeV aux collisionneurs, Les Houches 2007 : 11-29 June 2007*, 2008 121, arXiv: 0803.0678 [hep-ph], URL: https://inspirehep.net/record/780755/files/arXiv:0803.0678.pdf.

[128] ATLAS Collaboration, *Performance of Jet Algorithms in the ATLAS Detector*, ATL-PHYS-INT-2010-129, 2010, URL: https://cds.cern.ch/record/1311867.

[129] ATLAS Collaboration, *Performance of jet substructure techniques for large-R jets in proton–proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, JHEP **09** (2013) 076, arXiv: 1306.4945 [hep-ex].

[130] 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].

[131] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. **100** (2008) 242001, arXiv: 0802.2470 [hep-ph].

[132] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Techniques for improved heavy particle searches with jet substructure*, Phys. Rev. D **80** (2009) 051501, arXiv: 0903.5081 [hep-ph].

[133] D. Kroihn, J. Thaler and L. Wang, *Jet Trimming*, JHEP **02** (2010) 084, arXiv: 0912.1342 [hep-ph].

[134] 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].

[135] ATLAS Collaboration, *Local Hadronic Calibration*, ATL-LARG-PUB-2009-001-2, 2008, URL: https://cds.cern.ch/record/1112035.

[136] 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].

[137] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2017) 1896, arXiv: 1111.6097 [hep-ex].

[138] ATLAS Collaboration, *Jet global sequential corrections with the ATLAS detector in proton–proton collisions at $\sqrt{s} = 8$ TeV*, ATLAS-CONF-2015-002, 2015, URL: https://cds.cern.ch/record/2001682.

[139] 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].

[140] M. Cacciari and G. Salam, *Pileup subtraction using jet areas*, Phys. Lett. B **659** (2008), arXiv: 0707.1378 [hep-ph].

[141] 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].

[142] ATLAS Collaboration, *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8$ TeV*, ATLAS-CONF-2015-017, 2015, URL: https://cds.cern.ch/record/2008678.

[143] 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.

[144] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS*, ATLAS-CONF-2015-037, 2015, URL: https://cds.cern.ch/record/2044941.

[145] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145, arXiv: 1101.2599 [hep-ph].

[146] 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].

[147] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **10** (2012) 003, arXiv: 1207.1303 [hep-ex].

[148] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

[149] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189, arXiv: 0901.0002 [hep-ph].

[150] B. Andersson, G. Gustafson, and H. Pi, *The FRITIOF model for very high-energy hadronic collisions*, Z. Phys. C **57** (1993) 485.

[151] H. W. Bertini, *Low-Energy Intranuclear Cascade Calculation*, Phys. Rev. **131** (4 1963) 1801.

[152] ATLAS Collaboration, *Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in High Pile-Up LHC Environment*, ATLAS-CONF-2012-042, 2012, URL: https://cds.cern.ch/record/1435196.

[153] ATLAS Collaboration, *Properties of jets and inputs to jet reconstruction and calibration with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-036, 2015, URL: https://cds.cern.ch/record/2044564.

[154] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[155] ATLAS Collaboration, *ATLAS TWiki*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetTriggerPublicResults.

[156] CDF Collaboration, *Measurement of the inclusive jet cross section at the Fermilab Tevatron $p\bar{p}$ collider using a cone-based jet algorithm*, Phys. Rev. D **78** (2008) 052006, arXiv: 0807.2204 [hep-ex].

[157] D0 Collaboration, *Measurement of the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. D **85** (2012) 052006, arXiv: 1110.3771 [hep-ph].

[158] ALICE Collaboration, *Measurement of the inclusive differential jet cross section in pp collisions at $\sqrt{s} = 2.76$ TeV*, Phys. Lett. B **722** (2013), arXiv: 1301.3475 [hep-ph].

[159] ATLAS Collaboration, *Measurement of the inclusive jet cross section in pp collisions at $\sqrt{s} = 7$ TeV and comparison to the inclusive jet cross section at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Eur. Phys. J. C **73** (2013) 2509, arXiv: 1304.4739 [hep-ex].

[160] CMS Collaboration, *Measurement of the inclusive jet cross section in pp collisions at $\sqrt{s} = 2.76$ TeV*, Eur. Phys. J. C **76** (2016) 265, arXiv: 1512.06212 [hep-ex].

[161] ATLAS Collaboration, *Measurement of inclusive jet and dijet cross sections in proton–proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector*, Eur. Phys. J. C **71** (2011) 1512, arXiv: 1009.5908 [hep-ex].

[162] CMS Collaboration, *Measurements of differential jet cross sections in proton–proton collisions at $\sqrt{s} = 7$ TeV with the CMS detector*, Phys. Rev. D **87** (2013) 112002, arXiv: 1212.6660 [hep-ex].

[163] CMS Collaboration,
Measurement of the ratio of inclusive jet cross sections using the anti-k_T algorithm with radius parameters $R = 0.5$ and 0.7 in pp collisions at $\sqrt{s} = 7$ TeV,
Phys. Rev. D **90** (2014) 072006, arXiv: 1406.0324 [hep-ex].

[164] CMS Collaboration,
Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV,
Phys. Rev. Lett. **107** (2011) 132001, arXiv: 1106.0208 [hep-ex].

[165] ATLAS Collaboration, *Measurement of the inclusive jet cross-sections in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*,
JHEP **09** (2017) 020, arXiv: 1706.03192 [hep-ex].

[166] CMS Collaboration,
Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s} = 8$ TeV and cross section ratios to 2.76 and 7 TeV,
JHEP **03** (2017) 156, arXiv: 1609.05331 [hep-ex].

[167] CMS Collaboration, *Measurement of the double-differential inclusive jet cross section in proton–proton collisions at $\sqrt{s} = 13$ TeV*,
Eur. Phys. J. C **76** (2016) 451, arXiv: 1605.04436 [hep-ex].

[168] J. Currie, E.W.N. Glover and J. Pires,
NNLO QCD predictions for single jet inclusive production at the LHC,
Phys. Rev. Lett. **118** (2017) 072002, arXiv: 1611.01460 [hep-ex].

[169] J. Currie et al.,
Single jet inclusive production for the individual jet p_T scale choice at the LHC,
Acta Phys. Polon. B **48** (2017) 955, arXiv: 1704.00923 [hep-ex].

[170] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76** (2016) 653,
arXiv: 1608.03953 [hep-ex].

[171] T. Sjöstrand and P. Z. Skands,
Transverse-momentum-ordered showers and interleaved multiple interactions,
Eur. Phys. J. C **39** (2005) 129.

[172] Torbjörn Sjöstrand and Peter Zeiler Skands,
Multiple interactions and the structure of beam remnants, JHEP **03** (2004) 053,
arXiv: hep-ph/0402078.

[173] Hung-Liang Lai et al., *New parton distributions for collider physics*,
Phys. Rev. D **82** (2010) 074024, arXiv: 1007.2241 [hep-ph].

[174] J. Bellm, S. Gieseke, D. Grellscheid, A. Papaefstathiou, S. Platzer et al.,
Herwig++ 2.7 Release Note, 2013, arXiv: 1310.6877 [hep-ph].

[175] 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.

[176] ATLAS Collaboration, *Search for New Phenomena in Dijet Mass and Angular Distributions from pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector*,
Phys. Lett. B **754** (2016) 302, arXiv: 1512.01530 [hep-ex].

[177] ATLAS Collaboration, *A measurement of the calorimeter response to single hadrons and determination of the jet energy scale uncertainty using LHC Run-1 pp-collision data with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 26, arXiv: 1607.08842 [hep-ex].

[178] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2015-015, 2015, URL: https://cds.cern.ch/record/2037613.

[179] 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].

[180] J. Wenninger and E. Todesco, *Large Hadron Collider momentum calibration and accuracy*, Phys. Rev. Accel. Beams **20** (2017) 081003.

[181] Bogdan Malaescu, *An Iterative, Dynamically Stabilized (IDS) Method of Data Unfolding*, (2011), arXiv: 1106.3107 [physics.data-an].

[182] ATLAS Collaboration, *Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. D **86** (2012) 014022, arXiv: 1112.6297 [hep-ex].

[183] Gerhard Bohm, *Introduction to statistics and data analysis for physicists*, Verl. Dt. Elektronen-Synchrotron, 2010 336, ISBN: 978-3-935702-41-6.

[184] Zoltan Nagy, *Next-to-leading order calculation of three jet observables in hadron hadron collision*, Phys. Rev. D **68** (2003) 094002, arXiv: hep-ph/0307268.

[185] T. Carli et al., *A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project*, Eur. Phys. J. C **66** (2010) 503, arXiv: 0911.2985 [hep-ph].

[186] S. Carrazza and J. Pires, *Perturbative QCD description of jet data from LHC Run-I and Tevatron Run-II*, JHEP **10** (2014) 145, arXiv: 1407.7031 [hep-ex].

[187] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, Eur. Phys. J. C **75** (2015), arXiv: 1412.7420 [hep-ph].

[188] 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].

[189] 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].

[190] Richard D. Ball et al., *Parton distributions for the LHC Run II*, JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph].

[191] ZEUS and H1 Collaborations, *Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data*, Eur. Phys. J. C **75** (2015) 580, arXiv: 1506.06042 [hep-ex].

[192] S. Alekhin, J. Blümlein, S. Moch and R. Placakyte, *Parton Distribution Functions, α_s and Heavy-Quark Masses for LHC Run II*, Phys. Rev. D **96** (2017) 014011, arXiv: 1701.05838 [hep-ex].

[193] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001, arXiv: 1510.03865 [hep-ph].

[194] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv: hep-ph/0201195.

[195] Richard Corke and Torbjorn Sjöstrand, *Interleaved parton showers and tuning prospects*, JHEP **03** (2011) 032, arXiv: 1011.1759 [hep-ph].

[196] Stefan Gieseke, Christian Rohr and Andrzej Siodmok, *Colour reconnections in Herwig++*, Eur. Phys. J. C **72** (2012) 2225, arXiv: 1206.0041 [hep-ph].

[197] Michael H. Seymour and Andrzej Siodmok, *Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data*, JHEP **10** (2013) 113, arXiv: 1307.5015 [hep-ph].

[198] Stefan Dittmaier, Alexander Huss and Christian Speckner, *Weak radiative corrections to dijet production at hadron colliders*, JHEP **11** (2012) 095, arXiv: 1210.0438 [hep-ph].

[199] ATLAS Collaboration, *Measurement of the inclusive jet cross-section in proton–proton collisions at $\sqrt{s} = 7$ TeV using 4.5 fb^{-1} of data with the ATLAS detector*, JHEP **02** (2015) 153, arXiv: 1410.8857 [hep-ex], Erratum: JHEP **09** (2015) 141.

[200] ATLAS Collaboration, *Measurement of dijet cross sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector*, JHEP **05** (2014) 059, arXiv: 1312.3524 [hep-ex].

[201] U. Baur, I. Hinchliffe and D. Zeppenfeld, *Excited Quark Production at Hadron Colliders*, Int. J. Mod. Phys. A **2** (1987) 1285.

[202] U. Baur, M. Spira and P. M. Zerwas, *Excited-quark and -lepton production at hadron colliders*, Phys. Rev. D **42** (1990) 815.

[203] Satyaki Bhattacharya, Sushil Singh Chauhan, Brajesh Chandra Choudhary and Debajyoti Choudhury, *Quark excitations through the prism of direct photon plus jet at the LHC*, Phys. Rev. D **80** (2009) 015014, arXiv: 0901.3927 [hep-ph].

[204] Savas Dimopoulos and Greg L. Landsberg, *Black holes at the Large Hadron Collider*, Phys. Rev. Lett. **87** (2001) 161602, arXiv: hep-ph/0106295.

[205] Steven B. Giddings and Scott D. Thomas, *High energy colliders as black hole factories: The end of short distance physics*, Phys. Rev. D **65** (2002) 056010, arXiv: hep-ph/0106219.

[206] Douglas M. Gingrich,
Quantum black holes with charge, colour, and spin at the LHC,
J. Phys. G **37** (2010) 105008, arXiv: 0912.0826 [hep-ph].

[207] Xavier Calmet, Wei Gong and Stephen D.H. Hsu,
Colorful quantum black holes at the LHC, Phys. Lett. B **668** (2008) 20,
arXiv: 0806.4605 [hep-ph].

[208] Nima Arkani-Hamed, Savas Dimopoulos and G.R. Dvali,
The Hierarchy problem and new dimensions at a millimeter,
Phys. Lett. B **429** (1998) 263, arXiv: hep-ph/9803315.

[209] Lisa Randall and Raman Sundrum,
A Large mass hierarchy from a small extra dimension,
Phys. Rev. Lett. **83** (1999) 3370, arXiv: hep-ph/9905221.

[210] ATLAS Collaboration,
Search for Production of Resonant States in the Photon-Jet Mass Distribution Using pp Collisions at at $\sqrt{s} = 7$ TeV Collected by the ATLAS Detector,
Phys. Rev. Lett. **108** (2012) 211802, arXiv: 1112.3580 [hep-ex].

[211] ATLAS Collaboration, *Search for new phenomena in photon+jet events collected in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*,
Phys. Lett. B **728** (2014) 562, arXiv: 1309.3230 [hep-ex].

[212] CMS Collaboration, *Search for excited quarks in the γ+jet final state in proton–proton collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B **738** (2014) 274,
arXiv: 1406.5171 [hep-ex].

[213] ATLAS Collaboration, *Search for new phenomena with photon+jet events in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
JHEP **03** (2016) 041, arXiv: 1512.05910 [hep-ex].

[214] ATLAS Collaboration, *Search for new phenomena in dijet events using 37 fb^{-1} of pp collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
Phys. Rev. D **96** (2017) 052004, arXiv: 1703.09127 [hep-ex].

[215] CMS Collaboration, *Search for dijet resonances in proton–proton collisions at $\sqrt{s} = 13$ TeV and constraints on dark matter and other models*,
Phys. Lett. B **769** (2017) 520, arXiv: 1611.03568 [hep-ex].

[216] Steffen Schumann and Frank Krauss,
A Parton shower algorithm based on Catani-Seymour dipole factorisation,
JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph].

[217] Stefan Höche, Frank Krauss, Steffen Schumann and Frank Siegert,
QCD matrix elements and truncated showers, JHEP **05** (2009) 053,
arXiv: 0903.1219 [hep-ph].

[218] S. Catani, M. Fontannaz, J. P. Guillet and E. Pilon,
Cross-section of isolated prompt photons in hadron hadron collisions,
JHEP **05** (2002) 028, arXiv: hep-ph/0204023.

[219] L. Bourhis, M. Fontannaz, J.Ph. Guillet and M. Werlen,
Next-to-leading order determination of fragmentation functions,
Eur. Phys. J. C **19** (2001) 89, arXiv: hep-ph/0009101.

[220] D. M. Gingrich, *Monte Carlo event generator for black hole production and decay in proton-proton collisions*, Comput. Phys. Commun. **181** (2010) 1917, arXiv: 0911.5370 [hep-ph].

[221] ATLAS Collaboration, *Measurement of the inclusive isolated prompt photon cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. D **83** (2011) 052005, arXiv: 1012.4389 [hep-ex].

[222] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018, 2014, URL: https://cds.cern.ch/record/1700870.

[223] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, [Erratum: Eur. Phys. J.C73,2501(2013)], arXiv: 1007.1727 [physics.data-an].

[224] Alexander L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693.

[225] W. Verkerke, and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv: physics/0306116 [physics].

[226] L. Moneta et al., *The RooStats Project*, PoS **ACAT2010** (2010) 057, arXiv: 1009.1003 [physics.data-an].

[227] Kyle S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198, arXiv: hep-ex/0011057.

[228] Max Baak, Stefan Gadatsch, Robert Harrington and Wouter Verkerke, *Interpolation between multi-dimensional histograms using a new non-linear moment morphing method*, Nucl. Instrum. Meth. A **771** (2015) 39, arXiv: 1410.7388 [physics.data-an].

[229] ATLAS Collaboration, *Search for dark matter at $\sqrt{s} = 13$ TeV in final states containing an energetic photon and large missing transverse momentum with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 393, arXiv: 1704.03848 [hep-ex].

[230] ATLAS Collaboration, *Search for resonances in diphoton events at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **09** (2016) 001, arXiv: 1606.03833 [hep-ex].

[231] ATLAS Collaboration, *High-E_{T} isolated-photon plus jets production in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Nucl. Phys. B **918** (2017) 257, arXiv: 1611.06586 [hep-ex].

[232] R. A. Fisher, *On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P*, J. Royal Statistical Society **85** (1922) 87.