UNIVERSIDAD AUTONOMA DE MADRID

QCD tests in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC

Javier Llorente Merino

Departamento de Física Teórica

Universidad Autónoma de Madrid

A thesis submitted for the degree of Doctor in Physics

Supervisor: Fernando Barreiro Alonso

September 2015

A mis padres.

If you want to know the way nature works, we looked at it carefully, and that's the way it looks. You don't like it..., go somewhere else! To another universe, where the rules are simpler.

Richard Feynman

Introducción

El gran colisionador de hadrones LHC instalado en el CERN, y en particular el detector ATLAS, han ampliado el campo de estudio de la física de altas energías a un nivel inexplorado en experimentos anteriores. No sólo han permitido el aumento de la energía en el centro de masas hasta el valor de $\sqrt{s} = 7$ TeV, sino que han logrado una excelente luminosidad a lo largo de estos años de toma de datos. Esto ha sido especialmente importante para el descubrimiento del bosón de Higgs, además de permitir búsquedas de nueva física más allá del modelo estándar, así como estudios de las propiedades de estados finales conocidos, como la producción de pares top-antitop.

En esta tesis se han analizado datos tomados a lo largo del año 2011 para testar el sector fuerte $SU(3)_C$ del modelo estándar.

Los modelos de duchas partónicas y de hadronización de quarks han sido testados en una medida del flujo de energía transversa de los jets producidos en estados finales $t\bar{t}$. Considerando los decaimientos del quark top $t \to bW$ y los subsiguientes decaimientos hadrónicos del bosón W, se han logrado aislar muestras de muy alta pureza de jets inducidos por quarks b y ligeros. Esto, junto con la fina granularidad del sistema calorimétrico de ATLAS, nos ha permitido comparar los flujos transversos correspondientes a los b-jets con los de los jets ligeros (ATLAS Collaboration, Eur. Phys. J. C **73** 2676 (2013)), así como determinar la masa del quark bottom a partir de los llamados efectos de apantallamiento angular, claramente apreciables en estos datos (J. Llorente and J. C., Nucl. Phys. B **889** 401 (2014)). Además, se observa que los b-jets tienen una distribución más ancha de la energía transversa que los jets ligeros, por lo que este observable podría ser utilizado para la discriminación de jets atendiendo al quark que los origina.

El valor de la constante de acoplo fuerte α_s en la escala del polo del Z^0 también ha sido medido en la segunda parte de esta tesis. Esta constante tiene una im-

portancia central en nuestro entendimiento de la naturaleza a escala subatómica, puesto que juega un papel fundamental en la radiación de gluones, además de ser responsable de la estabilidad de la materia nuclear.

Hemos seleccionado una muestra de sucesos con varios jets imponiendo que la suma escalar de momentos transversos de los dos jets con mayor p_{T} sea mayor de 500 GeV, mientras que el momento transverso de los jets adicionales debe ser mayor de 50 GeV. Los jets seleccionados deben ser detectados en la parte central del detector $|\eta| < 2.5$. En estas condiciones, hemos realizado la medida de las correlaciones transversas energía-energía, que son la extensión natural de la función de correlación energía-energía, anteriormente usada en colisionadores $e^{+}e-$ a las energías de PETRA-PEP y LEP-SLC.

Se ha realizado una comparación de estos datos (J. Llorente, F. B. ATL-PHYS-COM-2013-884, enviado a Phys. Lett. B) con las predicciones a segundo orden en QCD perturbativa (A. A., F. B., J. Llorente y W. W., Phys Rev. D **86** 114017 (2012)), que permite la determinación de la constante de acoplo fuerte $\alpha_s(m_Z) = 0.1173 \pm 0.0010\ (\mathrm{exp.})^{+0.0065}_{-0.0026}\ (\mathrm{theo.})$.

Introduction

The Large Hadron Collider facility at CERN, and in particular the ATLAS detector, have provided a testing ground for high energy physics at a level which couldn't be explored in previous experiments. It has not only increased the centre-of-mass energy to the up to now highest value of $\sqrt{s} = 7$ TeV but also has had an excellent luminosity performance during these years of data taking. This has been especially important for the discovery of the Higgs boson, as well as for allowing searches of new physics and studies of the properties of, up to now, poorly known final states as $t\bar{t}$ production.

In this thesis, data taken during the year 2011 have been analyzed to probe the strong $SU(3)_C$ sector of the Standard Model.

The parton shower and hadronisation models have been tested in this thesis by means of the measurement of jet shapes in $t\bar{t}$ final states. Relying on the top-quark decays $t \to bW$ and the subsequent W boson hadronic decays $W \to q\bar{q}'$, samples with very high purity of b- and light-quark induced jets have been selected. This fact, together with the fine granularity of the ATLAS calorimeter system, has allowed us to compare b-jet shapes with those of light jets (ATLAS Collaboration, Eur. Phys. J. C **73** 2676 (2013)) and to determine the b-quark mass from the so-called angular screening effects, clearly visible in these data (J. Llorente and J. C., Nucl. Phys. B **889** 401 (2014)). Furthermore, b-jets are found to be broader than light jets and therefore b-jet shapes may be used for tagging purposes.

The value of the strong coupling constant α_s at the Z^0-pole scale has also been measured in this thesis. This constant plays a key role in our understanding of nature at the subatomic scale, as it determines the strength of gluon radiation, as well as being responsible for the stability of nuclear matter.

We have selected a sample of multijet events by demanding the scalar sum of

transverse momenta of the two leading jets to be greater than 500 GeV while the transverse momenta of the subleading jets is required to be above 50 GeV. All jets are required to lie in the central part of the calorimeter, i.e. $|\eta| < 2.5$. We have performed a measurement of transverse energy-energy correlations, which are the natural extension of the energy-energy correlation function which was popular in e^+e^- colliders at PETRA-PEP and LEP-SLC energies. We have performed a comparison between the data (J. Llorente, F. B. ATL-PHYS-COM-2013-884, submitted to Phys. Lett. B) with NLO pQCD calculations (A. A., F. B., J. Llorente and W. W., Phys. Rev. D **86** 114017 (2012)) which allows us to determine $\alpha_s(m_Z) = 0.1173 \pm 0.0010\ (\text{exp.})^{+0.0065}_{-0.0026}\ (\text{theo.})$.

Contents

Chapter 1

Theoretical framework

In this chapter, the main features of the Standard Model of Particle Physics (SM) and in particular on Quantum Chromodynamics (QCD) are described in some detail.

1.1 The Standard Model: General features

The Standard Model of Particle Physics [1, 2] is a quantum field theory [3] summarising our current knowledge on the interactions of subatomic particles, with the exception of gravity. It is based on the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The first part, $SU(3)_C$, correctly describes a wide variety of data regarding the strong interaction, while $SU(2)_L \times U(1)_Y$ describes the electroweak sector of the theory.

The elementary particles in the SM are divided in two main subgroups, fermions and bosons. Fermions are the building blocks from which matter is constructed, and have a semi-integer value of the spin, while bosons are considered to be the force carriers of the interactions between fermions and have integer spin. In figure 1.1, the elementary particles and their interactions are shown.

Fermions are classified into quarks and leptons depending on the way they interact. Quarks have fractional electric charge, colour charge, as well as weak hypercharge. Thus, they are subject to the strong, weak and electromagnetic interactions. Furthermore, they are the constituents of hadrons, in particular

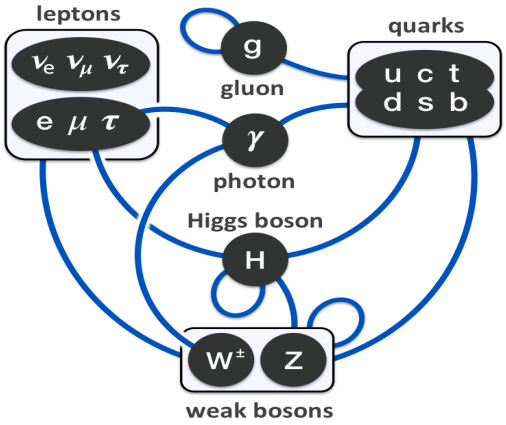

Figure 1.1: The fundamental constituents and interactions of the Standard Model.

nucleons. Leptons have integer electric charge as well as weak hypercharge, and therefore they are subject to the electroweak force.

The massive bosons and fermions in the Standard Model acquire mass according to the Higgs mechanism, where the electroweak symmetry $SU(2)_L \times U(1)_Y$ is broken in order to allow for mass terms. Dirac mass terms of the form $\bar{\Psi}_L m \Psi_R$ are not $SU(2)$ invariant due to the chiral symmetry.

A new particle (the so-called Higgs boson), whose couplings to the rest of particles define their masses, arises from this mechanism as the vibrational modes of the new field Φ, which transforms as an $SU(2)$ doublet and has a non-vanishing vacuum expectation value $v = \langle 0|\Phi|0\rangle \neq 0$.

1.1.1 Symmetry and invariance

The most fundamental feature of the Standard Model is its invariance under the gauge group of transformations $SU(3)_C \times SU(2)_L \times U(1)_Y$. These transformations can be represented mathematically in terms of the generators of the symmetry group. As the simplest example, let us consider the electromagnetic group $U(1)_Y$, whose Lie algebra has just one generator, and its action on a Dirac spinor Ψ.

$$\Psi(x) \mapsto \Psi'(x) = e^{ig\omega(x)}\Psi(x) \tag{1.1}$$

Let us consider the free fermion Lagrangian as well

$$\mathcal{L}_\Psi = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi \tag{1.2}$$

Clearly, the Lagrangian 1.2 is not invariant under the transformation 1.1 due to the presence of terms on the derivatives of the space-time dependent function $\omega(x)$. However, one can consider the covariant derivative D_μ defined as

$$D_\mu = \partial_\mu - igA_\mu \tag{1.3}$$

where A_μ is a gauge field and g is the QED coupling (the electron charge). Then, the prospect of invariance leads us to write the transformation law for A_μ

$$A_\mu \mapsto A'_\mu = A_\mu + \partial_\mu \omega(x) \tag{1.4}$$

Equation 1.4 can also be obtained from the gauge transformations of the classical electromagnetic 4-potential which leave the Maxwell equations invariant

$$\left.\begin{array}{l} \phi' = \phi + \partial_t \Omega(\vec{x}, t) \\ \vec{A}' = \vec{A} - \vec{\nabla}\Omega(\vec{x}, t) \end{array}\right\} \Rightarrow A'_\mu = A_\mu + \partial_\mu \Omega(x) \tag{1.5}$$

By a simple identification of the equations 1.4 and 1.5, it is seen that there is a clear relationship between the Lagrangian 1.6 and electromagnetism. In fact, the QED Lagrangian is the quantum version of the classical equations of electrodynamics, where the field A_μ, introduced by the requirement of invariance in Eq. 1.3, represents the quanta of the electromagnetic fields: the photons.
Defining the field strength tensor in the same fashion as the Faraday tensor in Classical Electrodynamics, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, allows us to write the full Lagrangian of Quantum Electrodynamics (QED) as

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\Psi}(i\gamma^\mu D_\mu - m)\Psi \tag{1.6}$$

The term in $F^{\mu\nu}F_{\mu\nu}$ in Eq. 1.6 encodes the dynamics of the photon field. However, for non-commutative groups such as $SU(N)$, this term has more fundamental physical consequences. The coupling of the gauge bosons with themselves, as

we will see in the next section, is encoded in the new terms arising from the non-trivial structure constants of the group.

1.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the quantum field theory describing the strong interactions [4] and it is based on the gauge group $SU(3)$. The fundamental difference with respect to QED, whose Lagrangian has been described before, is the non-commutativity of the Lie algebra $\mathfrak{su}(3)$ associated with the gauge group. The fundamental representation of this algebra is given by the Gell-Mann matrices:

$$\lambda^1 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^2 = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\lambda^4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda^5 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda^6 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\lambda^7 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda^8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{1.7}$$

These eight matrices provide the group structure for the special unitary group $SU(3)$ via the exponential map. The gauge transformation for a fermion field Ψ_j can therefore be written as

$$\Psi_j(x) \mapsto \Psi'_j(x) = \sum_k \left(e^{ig_s\Gamma(x)}\right)_{jk} \Psi_k(x) \tag{1.8}$$

where $\Gamma(x) \equiv G_a(x)\lambda^a$ is an element of $\mathfrak{su}(3)$ and the indices j, k cover all quark flavours. By convention, the normalization of the generators is chosen to be

$$\mathrm{Tr}\left(\lambda^a\lambda^b\right) = T_R\delta^{ab}; \quad T_R = \frac{1}{2} \tag{1.9}$$

Thus, with this specification, the generators of the Lie group $SU(3)$ fulfill

$$\sum_a \lambda^a_{ij}\lambda^a_{jk} = C_F\delta_{ik}; \quad C_F = \frac{4}{3} \tag{1.10}$$

$$\mathrm{Tr}\left(T^aT^b\right) = \sum_{c,d} f^{cda}f^{cdb} = C_A\delta^{ab}; \quad C_A = 3 \tag{1.11}$$

where T^a are the matrices in the adjoint representation

$$\mathrm{ad}_X(Y) = [X, Y] \quad \forall X, Y \in \mathfrak{su}(3) \tag{1.12}$$

The Lagrangian of the theory, which is left invariant by the transformations in Eq. 1.8, can be written as

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4}F^a_{\mu\nu}F^{\mu\nu}_a + \sum_{j,k}\bar{\Psi}_j\left(i\gamma^\mu D_\mu - m\right)_{jk}\Psi_k \tag{1.13}$$

where the indices j, k cover all quark flavours. In this case, the covariant derivative is constructed by considering the 8 generators of $SU(3)$, which are given by the Gell-Mann matrices 1.7.

$$D_\mu = \partial_\mu - ig_s\sum_a G_a(x)\lambda^a \tag{1.14}$$

and the field strength tensor is defined as the curvature of the gauge connection

$$F_{\mu\nu} = \frac{i}{g_s}[D_\mu, D_\nu] = \partial_\mu G_\nu - \partial_\nu G_\mu + g_s f_{abc}G^b_\mu G^c_\nu\lambda^a \tag{1.15}$$

Here, the vector fields $G^a_\mu(x)$ represent the gluons, which are the bosons responsible for carrying the strong interaction and have $n = 8$ degrees of freedom. It is now clear that the first term in 1.13 contains the self-couplings of the gluon fields arising from the non-Abelian nature of the theory. In Eq. 1.15, f_{abc} are the so-called structure constants of the group $SU(3)$, defined from the commutation relations of the Gell-Mann matrices

$$[\lambda^a, \lambda^b] = if_{abc}\lambda^c \tag{1.16}$$

The strong coupling constant α_s is the fundamental parameter of the theory, giving the interaction strength in the vertices of the gluon-gluon and quark-gluon couplings. This constant will be measured in this thesis using the LHC data, and it can be written in terms of the coupling g_s in equation 1.14 as

$$\alpha_s = \frac{g_s^2}{4\pi} \tag{1.17}$$

1.2.1 Running coupling constant

The value of α_s is not constant, but rather depends on the energy scale Q involved in the interaction process. This is understood to be responsible for the properties of QCD known as confinement (the fact that quarks and gluons cannot be observed as isolated states) and asymptotic freedom (the decrease of the strong interaction strength at short distances).

Let us consider a dimensionless observable R which depends on the energy scale Q. Being R dimensionless, the dependence on Q can only be achieved via the ratio Q^2/μ^2, where μ is an arbitrary parameter introduced by the renormalisation procedure. Because the QCD lagrangian 1.13 does not make explicit reference on the parameter μ, neither should any physical observable such as R. Thus, the independence of $R(Q^2/\mu^2, \alpha_s)$ on μ can be expressed as

$$\mu^2 \frac{d}{d\mu^2} R(Q^2/\mu^2, \alpha_s) \equiv \left[\mu^2 \frac{\partial}{\partial \mu^2} + \mu^2 \frac{\partial \alpha_s}{\partial \mu^2} \frac{\partial}{\partial \alpha_s} \right] R = 0 \tag{1.18}$$

We can transform this equation into a more compact form by introducing the notations

$$t = \log\left(\frac{Q^2}{\mu^2}\right); \quad \beta(\alpha_s) = \mu^2 \frac{\partial \alpha_s}{\partial \mu^2} \tag{1.19}$$

With these definitions, Eq. 1.18 can be rewritten as

$$\left[-\frac{\partial}{\partial t} + \beta(\alpha_s) \frac{\partial}{\partial \alpha_s} \right] R(e^t, \alpha_s) = 0 \tag{1.20}$$

The running of the strong coupling constant α_s is thus determined by the renormalisation group equation

$$\beta(\alpha_s) = Q^2 \frac{\partial \alpha_s}{\partial Q^2} = \frac{\partial \alpha_s}{\partial \log Q^2} \tag{1.21}$$

Here, the function $\beta(\alpha_s)$ can be expanded in perturbation theory as a power series in α_s

$$\beta(\alpha_s) = -\alpha_s^2(\beta_0 + \beta_1 \alpha_s + \beta_2 \alpha_s^2 + \mathcal{O}(\alpha_s^3)) \tag{1.22}$$

The coefficients β_i in Eq. 1.22 are extracted from the higher order corrections to the vertices of the theory. Fig. 1.2 shows the one-loop diagrams contributing to the leading-order (LO) β function

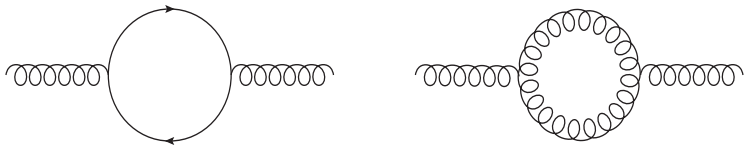

Figure 1.2: Graphs contributing to the β function in the one-loop approximation.

The result for the coefficients up to two loops is given by [5]

$$\beta_0 = \frac{1}{4\pi}\left[11 - \frac{2}{3}n_f\right]; \quad \beta_1 = \frac{1}{(4\pi)^2}\left[102 - \frac{38}{3}n_f\right] \tag{1.23}$$

In the one-loop approximation, i.e. keeping only the first term in the power expansion 1.22, the differential equation 1.21 can be integrated out, leading to the solution

$$\alpha_s(Q^2) = \frac{\alpha_s(Q_0^2)}{1 + \beta_0 \alpha_s(Q_0^2) \log\left(\frac{Q^2}{Q_0^2}\right)} \tag{1.24}$$

which expresses the value of $\alpha_s(Q)$ at a scale Q as a function of its value at some other scale Q_0. Let us define the QCD scale Λ_{QCD} as

$$\Lambda_{\text{QCD}}^2 = Q_0^2 \exp\left[-\frac{1}{\beta_0 \alpha_s(Q_0^2)}\right] \tag{1.25}$$

Then, equation 1.24 can be rewritten in terms of Λ_{QCD} as

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \log\left(\frac{Q^2}{\Lambda^2_{\mathrm{QCD}}}\right)} \tag{1.26}$$

For documentation purposes, the two-loop expression for $\alpha_s(Q)$ in terms of Λ_{QCD} is also quoted [6, 5]

$$\alpha_s(Q^2) = \frac{4\pi}{\beta_0 \log z}\left[1 - \frac{2\beta_1}{\beta_0^2}\frac{\log(\log z)}{\log z}\right]; \quad z = \frac{Q^2}{\Lambda^2_{\mathrm{QCD}}} \tag{1.27}$$

The running of α_s has been studied in several experimental conditions, including e^+e^-, ep and $p\bar{p}$ collisions, and all the available results agree to the renormalisation group equation 1.21, whose solutions have been calculated to higher orders. Figure 1.3 shows a comparison between the theoretical predictions for the 2012 world average of α_s [7] and the presently available data from various experimental sources.

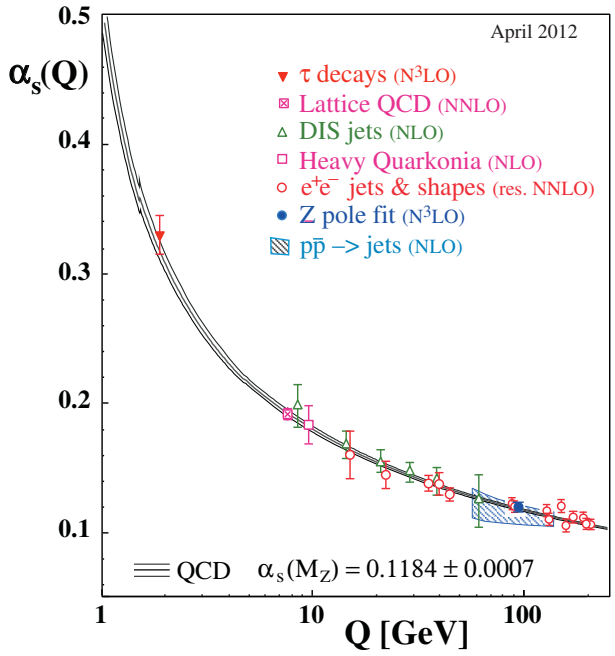

Figure 1.3: The running of the strong coupling α_s as a function of the energy scale Q. The shaded band represents the theoretical calculation, while the data points represent several measurements used as input for the 2012 α_s world average [6]

1.2.2 Proton structure: Deep inelastic scattering

The internal structure of protons has been studied in detail using neutral current deep inelastic scattering (NC DIS), [8], where an electron with momentum k is scattered by a proton with momentum P. In figure 1.4, a typical DIS process is illustrated.

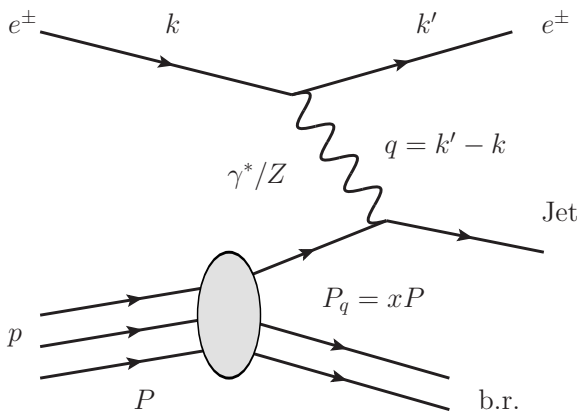

Figure 1.4: Electron-proton deep inelastic scattering

The process occurs via the exchange of a virtual photon γ^* or a Z boson with momentum q between the electron and one of the partons inside the proton. After the interaction, the parton, which carries a fraction x of the momentum of the proton, is separated from it and forms a hadronic jet whose properties are measured by the detector. The scattered electron can also be detected, measuring its final energy and momentum, thus allowing us to know the momentum scale $Q^2 \equiv -q^2$ relevant for the process and to probe the structure of the proton. The data recorded in NC DIS experiments at SLAC brought the first evidence that protons are in fact composite particles, which lead Feynman to develop the quark-parton model in the 1960s [9].

The DIS cross-section can be expressed in terms of a set of three Lorentz-invariant observables

$$Q^2 = -q^2 = (k' - k)^2; \quad x = \frac{Q^2}{2(P \cdot q)}; \quad y = \frac{Q^2}{sx} \tag{1.28}$$

where s is the squared center-of-mass energy of the electron-proton system. In terms of these 3 variables, the double-differential NC DIS cross section takes the

form

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{xQ^4}\left[xy^2F_1 + (1-y)F_2 + \left(y - \frac{y^2}{2}\right)F_3\right] \tag{1.29}$$

where $F_i(x, Q^2)$ are referred to as the structure functions of the proton. These are related to the probability that a quark interacts with the neutral current boson and are given by

$$F_1(x, Q^2) = \frac{1}{2}\sum_i \left[f_{q_i}(x, Q^2) + f_{\bar{q}_i}(x, Q^2)\right] C_{q_i}(Q^2) \tag{1.30}$$

$$F_2(x, Q^2) = x\sum_i \left[f_{q_i}(x, Q^2) + f_{\bar{q}_i}(x, Q^2)\right] C_{q_i}(Q^2) \tag{1.31}$$

$$F_3(x, Q^2) = \sum_i \left[f_{q_i}(x, Q^2) - f_{\bar{q}_i}(x, Q^2)\right] D_{q_i}(Q^2) \tag{1.32}$$

The coefficients C_{q_i}, D_{q_i} are functions of the weak angle θ_W, the weak isospin T_3 and the propagator $P(Q^2)$ [4]. Sometimes, the cross section 1.29 is expressed using the function

$$F_L(x, Q^2) = \left(1 + \frac{4M^2x^2}{Q^2}\right)F_2(x, Q^2) - 2xF_1(x, Q^2) \tag{1.33}$$

which is called the longitudinal structure function and vanishes for $Q^2 \to \infty$. The fact that $F_2 - 2xF_1 = 0$, known as the Callan-Gross relation, is a consequence of the impossibility for spin-$\frac{1}{2}$ quarks to absorb longitudinally polarised vector bosons. The functions $f_i(x, Q^2)$ are the so-called parton distribution functions and parametrise the probability of finding a parton of flavour i, quark or gluon, carrying a fraction x of the momentum of the proton, being Q^2 the interaction scale of the hard process.

1.2.3 Parton distribution functions

The parton distribution functions, hereafter simply PDFs, are not directly predicted in QCD. However, the DGLAP equations which will be discussed later provide an evolution scheme for their Q^2 dependence. To parametrise their analytical form, fits are performed to the DIS data at a given scale Q_0^2, which in the case of HERAPDF discussed here is taken to be $Q_0^2 = 1.9$ GeV^2. This ensures

that the starting scale is below the charm mass threshold. The parametrical form of the PDFs at the fixed scale Q_0^2 is chosen to be [10]

$$xf(x) = Ax^B(1-x)^C(1+Ex^2) \tag{1.34}$$

The initial choice of these parameters for the gluon $xf_g(x)$ and valence quarks $xf_{q_v}(x)$ yields

$$xf_g(x) = A_g x^{B_g}(1-x)^{C_g} \tag{1.35}$$
$$xf_{u_v}(x) = A_{u_v} x^{B_{u_v}}(1-x)^{C_{u_v}}(1+E_{u_v}x^2) \tag{1.36}$$
$$xf_{d_v}(x) = A_{d_v} x^{B_{d_v}}(1-x)^{C_{d_v}} \tag{1.37}$$

Once the PDFs are determined at the scale Q_0^2, their form at an arbitrary scale $Q^2 > Q_0^2$ is derived by solving the DGLAP equations at next-to-leading order (NLO) or next-to-next-to-leading order (NNLO). Figure 1.5 (left) shows the HERAPDF1.0 fits at $Q^2 = 10$ GeV^2, together with the experimental, model and parametrisation uncertainties. In the right part of Figure 1.5, the PDFs are evolved at $Q^2 = 10000$ GeV^2, i.e. at the energy scale relevant for hadron colliders as Tevatron or the LHC, using the DGLAP equations up to NNLO [$\mathcal{O}(\alpha_s^6)$].

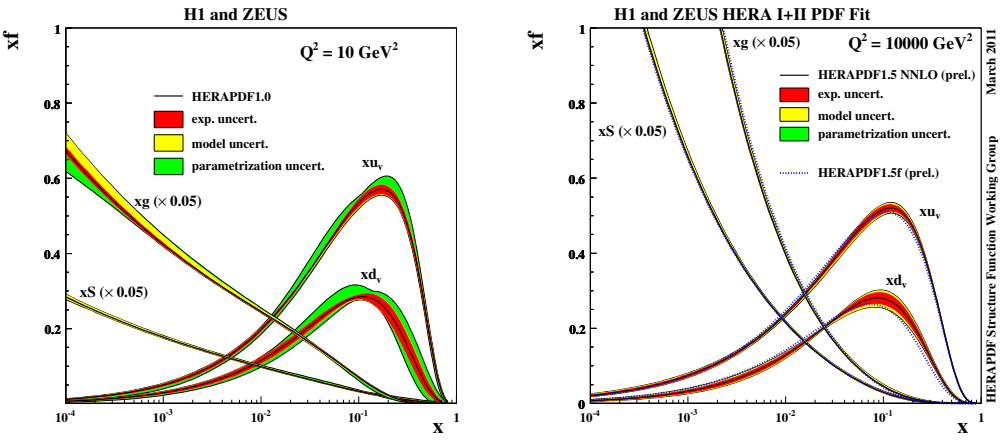

Figure 1.5: The HERAPDF parton distribution functions $xf(x, Q^2)$ as a function of x for valence quarks, gluons and sea quarks obtained for different values of the interaction scale Q^2. The gluon and sea PDFs have been scaled down by a factor of 20 [10]

The parton distribution functions are relevant in the calculation of hadron-hadron cross sections $\sigma(pp \to X)$ from the partonic process $\hat{\sigma}(ij \to Y)$. As it will be stated later, according to the factorisation theorem, the pp cross sections can be expressed as a convolution of the partonic cross sections with the PDFs.

1.2.4 The DGLAP evolution equations

The analytical form of the parton distribution functions is not predicted by QCD. However, differential equations can be derived for the evolution of the PDFs with the energy scale Q^2. The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [11] describe this scale evolution of the PDFs for both quarks and gluons as a convolution $f \otimes P$ of the PDFs at the present scale configuration with the DGLAP kernels $P_{ab}(z)$. They can be written as

$$\frac{\partial f_{q_i}(x,Q^2)}{\partial \log Q^2} = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{d\xi}{\xi} \left[f_{q_i}(\xi,Q^2) P_{qq}\left(\frac{x}{\xi}\right) + f_g(\xi,Q^2) P_{qg}\left(\frac{x}{\xi}\right) \right] \quad (1.38)$$

$$\frac{\partial f_g(x,Q^2)}{\partial \log Q^2} = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{d\xi}{\xi} \left[\sum_i f_{q_i}(\xi,Q^2) P_{gq}\left(\frac{x}{\xi}\right) + f_g(\xi,Q^2) P_{gg}\left(\frac{x}{\xi}\right) \right] \quad (1.39)$$

Equation 1.38 describes the scale evolution of the i-th active flavour quark density inside the proton, while 1.39 describes the running of the gluon PDF. The splitting functions $P_{ab}(z)$ parametrise the probability of a parton b emitting a parton a with momentum fraction z via the processes $q \to qg$ (gluon radiation), $g \to q\bar{q}$ (pair production) or $g \to gg$ (gluon splitting). The LO diagrams for these processes are illustrated in Figure 1.6.

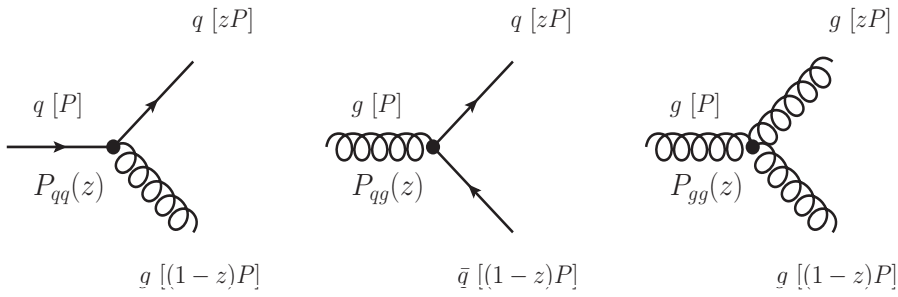

Figure 1.6: Leading-order diagrams contributing to the parton splitting functions.

The parton splitting functions can be expressed as a power series in the strong coupling constant α_s

$$P_{ij}(z) = P_{ij}^{(0)}(z) + \left(\frac{\alpha_s}{2\pi}\right) P_{ij}^{(1)}(z) + \mathcal{O}(\alpha_s^2) \tag{1.40}$$

The leading order contributions were calculated in Ref. [11], and are given by

$$P_{qq}^{(0)}(z) = C_F \left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right] \tag{1.41}$$

$$P_{gq}^{(0)}(z) = C_F \left[\frac{1+(1-z)^2}{z}\right] \tag{1.42}$$

$$P_{qg}^{(0)}(z) = T_R \left[z^2 + (1-z)^2\right] \tag{1.43}$$

$$P_{gg}^{(0)}(z) = 2C_A \left[\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z)\right] + \delta(1-z)\frac{11C_A - 4N_f T_R}{6} \tag{1.44}$$

where factors of the form $(1-z)^{-1}$ have been regularised by reinterpreting them as distributions defined in the following way

$$\int_0^1 \frac{f(z)dz}{(1-z)_+} \equiv \int_0^1 \frac{f(z)-f(1)}{1-z}dz = \int_0^1 \log(1-z)\frac{df}{dz}dz \tag{1.45}$$

1.2.5 The factorisation theorem

One of the most succesful features of QCD is the fact that the perturbative and non-perturbative contributions to the physical cross sections in hadron-hadron collisions can be separated [12], allowing for theoretical predictions which can be testable in a hadron collider such as the LHC. The perturbative part of the interaction can be calculated in terms of the partonic cross section $\hat{\sigma}_{ij}$, for any pair of partons i and j in the initial state, using the standard Feynman rules [1], while the non-perturbative part is absorbed into the PDFs described above.
The differential cross-section for a given process in a hadron-hadron collision can be calculated as the convolution of the partonic cross section $\hat{\sigma}_{ij}$ with the parton distribution functions, after summing over all possible initial-state configurations:

$$\frac{d\sigma}{dX} = \sum_{i,j} \int_0^1 dx_1 \int_0^1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \frac{d\hat{\sigma}_{ij}}{dX}\left(x_1, x_2, \alpha_s(\mu_R^2), \frac{Q^2}{\mu_R^2}\right) \tag{1.46}$$

where X is any physical observable. The parameters μ_R and μ_F are the renormalisation and the factorisation scales, introduced as a result of the procedures of redefinition of the parameters in the theory to cancel the divergences appearing in the calculation of certain diagrams.

1.3 Monte Carlo predictions

A very convenient way of comparing the experimental data with the theoretical predictions is to use a Monte Carlo (MC) generator which simulates events of a given physics process. This simulation starts on the calculation of the parton level cross section $\hat{\sigma}$ at a fixed order in perturbation theory, and follows with the simulation of the parton cascades and non-perturbative effects, such as the hadronisation and multiple parton scatterings. In this section, a general overview of these techniques is given. The structure of a proton-proton collision is shown in Figure 1.7

1.3.1 Parton showers

The parton shower algorithm mimics the remaining terms of the perturbative expansion in α_s by emitting gluons which will eventually split into more partons. This is parametrised by the parton splitting functions 1.41 to 1.44 and implemented via Sudakov factors $\Delta(t_2, t_1)$ [15]. This quantity represents the non-splitting probability of a parton between two scales t_1 and t_2.
Consider the following simpified example, where the scale t is identified as the time axis: If the splitting probability at a time t is given by a function $f(t)$ and one requires that the splitting can only happen at the instant t if it did not happen for $t' < t$, whose probability is $\mathcal{N}(t')$, then the splitting probability $\mathcal{P}(t)$ satisfies

$$\mathcal{P}(t) = \frac{d}{dt}\left(1 - \mathcal{N}(t)\right) = f(t)\mathcal{N}(t) \tag{1.47}$$

Therefore, if the process starts at $t = 0$ with $\mathcal{N}(0) = 1$, the solution of 1.47 is

$$\mathcal{P}(t) = f(t)\exp\left[-\int_0^t f(s)ds\right] \tag{1.48}$$

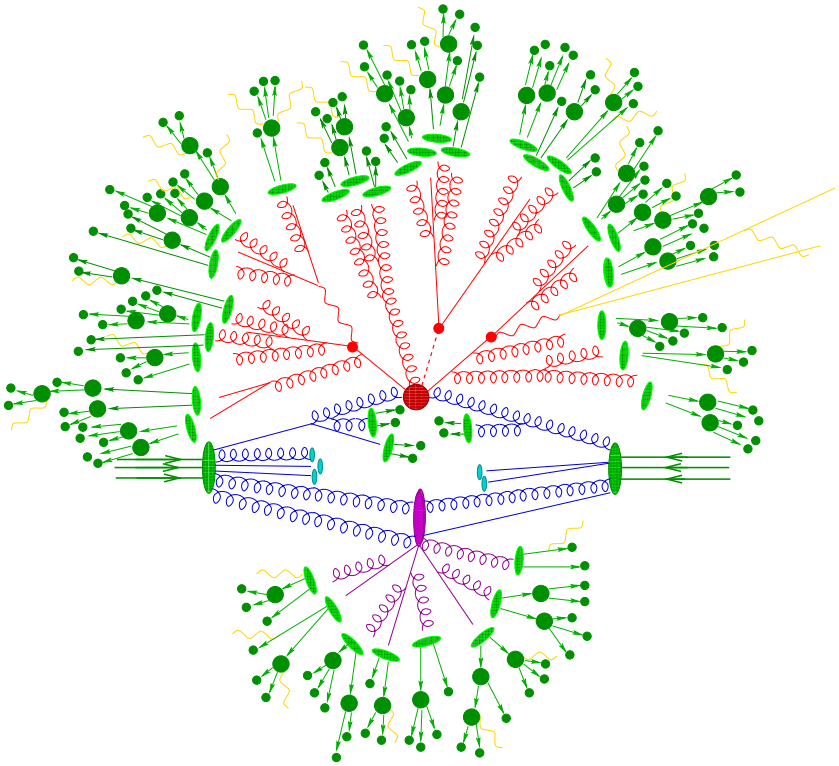

Figure 1.7: A sketch of the structure of a proton-proton collision, including the initial state partons (blue), the underlying event (violet), the hard scattering process and subsequent parton shower (red), and the final state hadrons (green). Figure taken from [14]

And therefore, the Sudakov factor $\Delta(t_2, t_1)$ with $t_1 \leq t_2$ is

$$\Delta(t_2, t_1) = 1 - \int_{t_1}^{t_2} \mathcal{P}(s)ds = 1 + \exp\left(-\int_0^t f(s)ds\right)\Big|_{t=t_1}^{t_2} \leq 1 \tag{1.49}$$

The parton showers can be angle or p_{T}-ordered, depending on the chosen scale t. In the first case, the angle of emission with respect to the incoming parton is decreased in each step, while in the second, the ordering variable in time is the parton transverse momentum. The sucessive branching is stopped when a certain cut-off scale is reached. For the initial-state radiation, the scale variation is reversed (angles and momenta are increased), while for final-state radiation, the scales are forced to decrease in each step.

1.3.2 Hadronisation

The hadronisation is the process of recombination of the final state partons into hadrons, which can be viewed as colour singlet quantum mechanical states. This process is based on the parton-hadron duality hypothesis [16], which establishes that, as hadronisation is a long-distance process involving only small momentum transfers, the flows of momentum and quantum numbers at the hadron level must follow those for the parton level. This implies that partons are recombined with other partons to form hadrons if they are close in phase space. As perturbation theory works well down to low scales $Q \sim 1$ GeV, the assumption is made that α_s can be defined non-perturbatively for arbitrary low scales Q. These assumptions are supported by heavy quark spectra and event shapes data, but they do not provide a model for the way hadronisation actually happens.
The mechanism by which hadrons are formed from partons is simulated using two models, namely the Lund string model [17] and the cluster model [18], which are schematically described in Figure 1.8

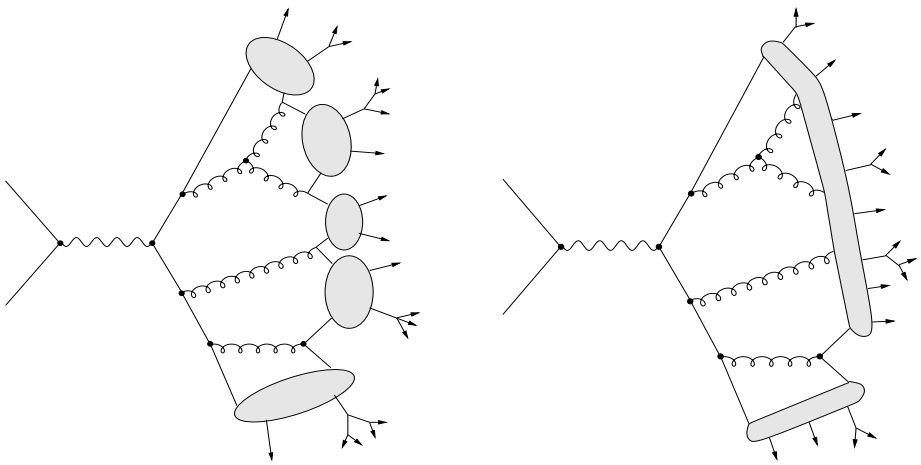

Figure 1.8: Schemes for the cluster (left) and the string (right) hadronisation models [19]

- **The Lund string model:** The string model describes the colour dynamics between quarks in terms of strings, assuming a linear confinement potential. When the quarks separate in the phase space, this confinement potential increases the energy of the string up to the mass threshold of a new $q\bar{q}$ pair.

At this point, the string is broken and the new $q\bar{q}$ pair gives raise to the formation of hadrons. When a gluon splits perturbatively, an additional string segment is created, whereas the remaining gluons at the end of the parton shower lead to kinks in the string segment that connects them.

- **The cluster model:** The cluster model is based on the preconfinement property of the shower, by which neighbouring colour connected partons have an asymptotic mass distribution that falls steeply at high masses and is aymptotically independent of Q^2 and universal. The method starts with a non-perturbative splitting $g \to q\bar{q}$, and follows with the association of $q\bar{q}$ pairs into colour singlet combinations, which are assumed to form clusters. These clusters decay into pairs of hadrons following an isotropic pattern.

1.3.3 Underlying Event

The underlying event (UE) is a result of the interaction between the remnants of the incoming hadrons. It contributes to the final state hadronic activity by interacting with the colour structure of the hard scattering outcoming partons, as well as creating new hadrons from the hadronisation of its own outcoming partons (see Figure 1.7). The parameters involved in the modelling of the UE have to be tuned using experimental data such as the jet shapes [20] or other sensitive observables [104]

1.3.4 Monte Carlo Programs

The general features described above are used by several MC generator programs to describe the experimental data. Here, a description is given for the most commonly used and, in particular, the ones used in the analyses carried out in this thesis.

1.3.4.1 PYTHIA

Pythia [22] is a LO event generator implementing the calculations of $2 \to 1$ and $2 \to 2$ matrix elements. It is matched to a p_{T}-ordered parton shower accounting for the initial and final state radiation. Thus, the relevant scale is $Q^2 = -p_{\mathrm{T}}^2$,

which increases its value as the shower evolves. For the hadronisation modelling, PYTHIA relies on the Lund string model described in the previous section. For the initial-state radiation, PYTHIA parton showers are space-like; i.e. the involved partons have a negative value of $m^2 = E^2 - \vec{p}^2$. In contrast, the value of m^2 is positive for the final-state radiation, thus developing a time-like shower.
The UE modelling is based on different tunes [23] depending on the values of the QCD parameters such as Λ_{QCD} used for this purpose, and they differ from version 6 to version 8, written in FORTRAN and C++, respectively.

1.3.4.2 HERWIG

The HERWIG program [24] implements a wide variety of QCD processes. The parton showers are angular-ordered. This means that, for a branching $a \to bc$ the relevant scale is given by $Q^2 = 2E_a^2(1 - \cos\theta_{bc})$, where E_a is the energy of the particle a and θ_{bc} is the angle between the branching products b and c. The so-defined scale Q^2 increases with the evolution of the shower, thus the angle between the products of sucessive branchings decreases as the shower evolves.
The hadronisation in HERWIG is implemented using the cluster model, while the underlying event is modelled using minimum-bias interactions. It also implements an interface to JIMMY [25], which uses multiple parton interactions to model the underlying event effects.

1.3.4.3 MC@NLO

The MC@NLO program [26] matches exact NLO calculations to a parton shower. The parton shower Monte Carlo programs already implement some of the NLO corrections, so the problem consists in avoiding the double counting between the NLO matrix element and the extra activity arising from the parton shower. This has been implemented by subtracting the NLO effects already accounted for from the matrix element calculation.
Events generated using MC@NLO are associated with a weight which might be $+1$ or -1. The negative weighting accounts for the cancellation of interfering diagrams. The smoothness of the distributions obtained using this scheme is guaranteed since the number of events with negative weights is reasonably small.

MC@NLO is commonly matched to the HERWIG parton showers interfaced with JIMMY.

1.3.4.4 POWHEG

In the POWHEG formalism [27], each event is built by producing the $2 \to 2$ or $2 \to 3$ hard scattering matrix element. The renormalisation and factorisation scales μ_R and μ_F are set to be equal to the transverse momentum of the hard partons ($p_{\mathrm{T}}^{\mathrm{Born}}$). The cut on the minimum transverse momentum of the generated hard partons, $p_{\mathrm{T}}^{\mathrm{Born}}$, may affect the value of the cross section due to the low-p_{T} divergence of the $2 \to 2$ cross section. POWHEG is matched either to PYTHIA or HERWIG for the parton shower step of the event generation. In case of using HERWIG, the UE modelling is tipically performed by the JIMMY program.

1.3.4.5 ALPGEN

The ALPGEN generator [28] is used for the simulation of $2 \to n$ multiparton events. The combination of the different multiplicities is performed by weighting each sample, with k partons in the final state, attending to its cross section to produce the inclusive sample. Then, the generated events at the matrix element level are matched to a parton shower algorithm such as PYTHIA or HERWIG.
The addition of the parton shower on the top of the n-parton final state can lead to an overlap of the event with the $(n+1)$-parton sample if the shower produces an additional jet. To avoid this, ALPGEN uses MLM matching to identify the overlaps. This algorithm is based on the following steps steps: First, the n-parton contribution is generated and matched to the parton shower. Then, a jet algorithm is applied to the outcoming parton shower output and if the number of jets is the same as the generated number of hard partons, the event is kept. If the parton shower has produced one or more additional hard jets, the event is tagged as an overlap and rejected.

1.4 Jets and jet algorithms

The collimated showers of particles formed by the fragmentation of partons produced in hard scattering processes are known as jets. These jets are the experimental counterpart to the hard scattering partons. Thus, the reconstruction of the kinematic properties of partons is possible using jet algorithms. Their performance is normally dependent on a clustering parameter R, which can be regarded as the jet radius in the η–φ[1] space.
In general terms, a jet algorithm needs to fulfill two conditions stating the invariance of the jet topology under the arbitrary addition of soft particles:

- **Collinear safety:** The jet topology in a given event should not depend on the collinear radiation of partons. If an initial parton radiates a soft particle with a small angle, both of them should be reconstructed in the same jet.
- **Infrared safety:** The parton radiation at large angles should not affect the jet topology. Thus, if a given event presents two reconstructed jets, the addition of soft radiation between them should not result in the merging of both jets.

These two requirements have been successfuly implemented into two different types of jet reconstruction algorithms. The cone algorithms, such as SISCone [29] are based on the maximisation of the energy density inside cones of a given radius R, followed by a merging-splitting step where overlapping configurations of two well-defined jets are removed. Other approach, using sequential recombination algorithms use pseudo-particles with defined four-momenta as seeds.

1.4.1 Sequential recombination algorithms

The sequential recombination algorithms are based on the iterative merging of pairs of constituents (i, j) in a same object attending to both their transverse momentum k_t and their spatial distance $\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\varphi_i - \varphi_j)^2}$.

[1] Pseudorapidity is defined as $\eta = -\log\left(\tan\frac{\theta}{2}\right)$ with θ the polar angle.

This is done using a metric d_{ij} defined as

$$d_{ij} = \min\left(k_{t,i}^{2p}, k_{t,j}^{2p}\right) \frac{\Delta R_{ij}^2}{R^2} \tag{1.50}$$

$$d_{i,B} = k_{t,i}^{2p} \tag{1.51}$$

The algorithm works as follows: for a pair of input particles (i, j), the value of d_{ij} is calculated and compared to the value of $d_{i,B}$. If $d_{ij} < d_{i,B}$, the inputs i and j are merged into a single input following a given recombination method. If this condition is not fulfilled, i is identified as a stable jet. Different values of the parameter p lead to different algorithms, such as k_t [30, 31] with $p = 1$, anti-k_t [32] with $p = -1$ or the Cambridge-Aachen [33] algorithm for $p = 0$. Figure 1.9 shows the behaviour of different jet algorithms. All these methods have been implemented in the FastJet program [34].

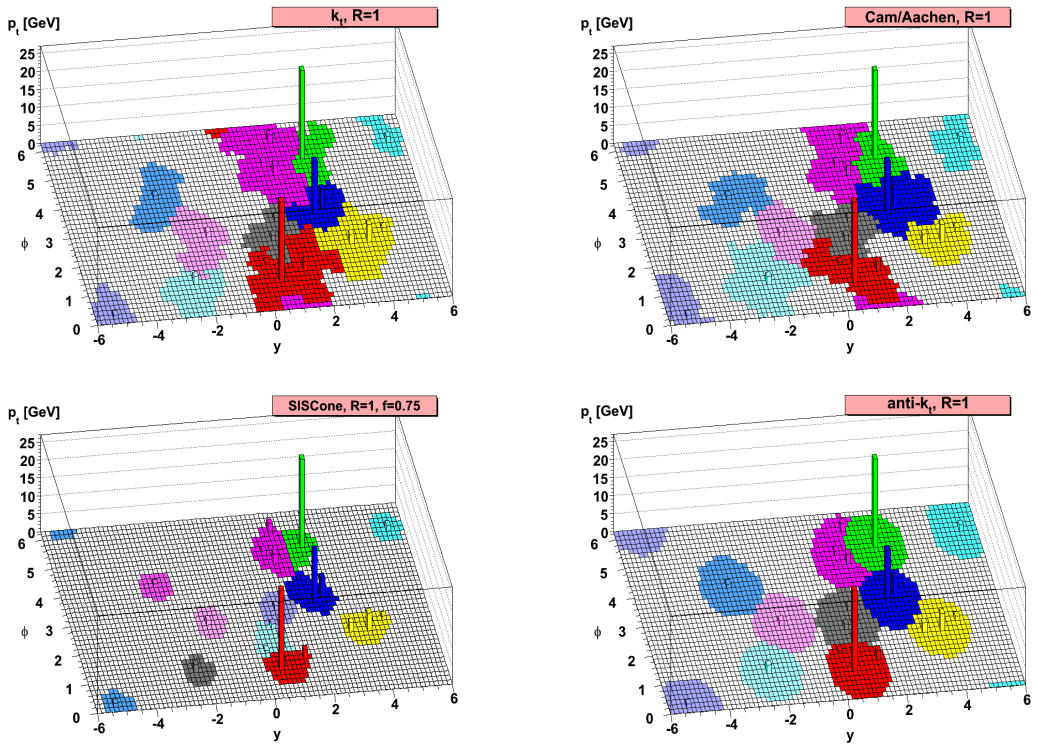

Figure 1.9: Sketch of the transverse profile of the jet areas attending to different algorithms. Figure taken from [32].

1.4.2 Recombination schemes

During the application of equation 1.51, there are different ways of combining the particles i and j. The most common is to simply add their four-momenta combining the inputs 1 and 2 into a new input with $p^\mu = p_1^\mu + p_2^\mu$. This is called the E-scheme and it is the one used for all measurements presented in this thesis, but there are several others, depending on the main observable being used for the recombination. These are the p_T and p_T^2 schemes for the transverse momentum and the E_T and E_T^2 schemes for the transverse energy.
In the E-scheme, the algorithm results in a massive object with massless constituents as inputs. However, in the rest of the methods, the inputs are combined into a massless object. The observable p resulting from the recombination of p_i and p_j, as well as the new object angular coordinates are defined as

$$p_R = p_i + p_j \tag{1.52}$$

$$\eta_R = \frac{w_i\eta_i + w_j\eta_j}{w_i + w_j} \tag{1.53}$$

$$\varphi_R = \frac{w_i\varphi_i + w_j\varphi_j}{w_i + w_j} \tag{1.54}$$

Here, the weights w_k are the transverse momentum p_T for the p_T and p_T^2 schemes and E_T for the E_T and E_T^2 schemes.

1.5 Transverse Energy-Energy Correlations

The energy-energy correlation (EEC) function was proposed in electron-positron annihilation $e^+e^- \to X$ to study the relative energy distribution of charged hadrons in a given event [35]. Recently, Ref. [36] has provided a theoretical calculation of the transverse EEC function in proton-proton collisions at the LHC. In the transverse plane, one can define the TEEC function as

$$\frac{1}{\sigma'}\frac{d\Sigma'}{d\phi} \equiv \frac{1}{N\Delta\cos\phi}\sum_{A=1}^{N}\sum_{ij}\frac{E_{\mathrm{T}i}^A E_{\mathrm{T}j}^A}{\left(\sum_k E_{\mathrm{T}k}^A\right)^2}\delta(\cos\phi - \cos\phi_{ij}) \tag{1.55}$$

The index A runs over a sample of N hard-scattering events, while the indices i and j run over the jets in the event labelled by the index A. Equation 1.55 shows that the TEEC function is nothing else than the distribution of the angle between pairs of jets in a given event weighted by the fraction between the product of their transverse energies and the squared total transverse energy in the event.
At the LO accuracy, the calculation of the TEEC function involves the convolution of the PDFs with the $2 \to 2$ partonic cross section. However, the $2 \to 2$ contributions will only appear in the end-points $\cos\phi = \pm 1$ due to the momentum conservation in the transverse plane. Away from the end-points, the contributions due to the $2 \to 3$ diagrams such as $gg \to ggg$ can be schematically expressed as

$$\frac{1}{\sigma'}\frac{d\Sigma'}{d\phi} = \frac{\sum_{a_i,b_i} f_{a_1}(x_1) f_{a_2}(x_2) \otimes \hat{\Sigma}(a_1 a_2 \to b_1 b_2 b_3)}{\sum_{a_i,b_i} f_{a_1}(x_1) f_{a_2}(x_2) \otimes \hat{\sigma}(a_1 a_2 \to b_1 b_2)} \tag{1.56}$$

where $\hat{\Sigma}(a_1 a_2 \to b_1 b_2 b_3)$ denotes the transverse energy-energy weighted partonic cross section, $f_{a_i}(x_i)$ are the PDFs of the i-th flavour parton inside the proton and $\otimes$ denotes a convolution over the appropriate variables.
In the LO approximation, the TEEC function described above is approximately independent of the PDFs due to the fact that certain normalised distributions for the $2 \to 3$ subprocesses are similar and that the same PDF combination $f_{a_1}(x_1) f_{a_2}(x_2)$ enters both the $2 \to 2$ and the $2 \to 3$ cross sections. Therefore, to a good level of approximation, the TEEC function reads

$$\frac{1}{\sigma'}\frac{d\Sigma'}{d\phi} \sim \frac{\alpha_s(\mu)}{\pi} F(\phi) \tag{1.57}$$

In the NLO accuracy, one can express the TEEC function as the product of the so-called K-factor and the LO result, namely

$$\frac{1}{\sigma'}\frac{d\Sigma'}{d\phi} \sim \frac{\alpha_s(\mu)}{\pi} F(\phi) \left[1 + \frac{\alpha_s(\mu)}{\pi} G(\phi)\right] \equiv \frac{\alpha_s(\mu)}{\pi} F(\phi) K(\phi) \tag{1.58}$$

One can also define the forward-backward asymmetry of the TEEC function (ATEEC) as the difference between the $\cos\phi < 0$ and $\cos\phi > 0$ halves of the

distribution

$$\frac{1}{\sigma'}\frac{d\Sigma'^{\mathrm{asym}}}{d\phi} \equiv \frac{1}{\sigma'}\frac{d\Sigma'}{d\phi}\bigg|_{\pi-\phi} - \frac{1}{\sigma'}\frac{d\Sigma'}{d\phi}\bigg|_{\phi} \tag{1.59}$$

The importance of both the next-to-leading order TEEC and ATEEC functions resides in the fact that, being both to a good extent independent of the PDFs and of the renormalisation and factorisation scales μ_R and μ_F, as it has been shown in Ref. [36], they show a marked dependence on the strong coupling constant $\alpha_s(m_Z)$ and can therefore be used for a measurement of α_s at the LHC. Fig. 1.10 illustrates the behaviour of the TEEC and ATEEC functions with the strong coupling constant.

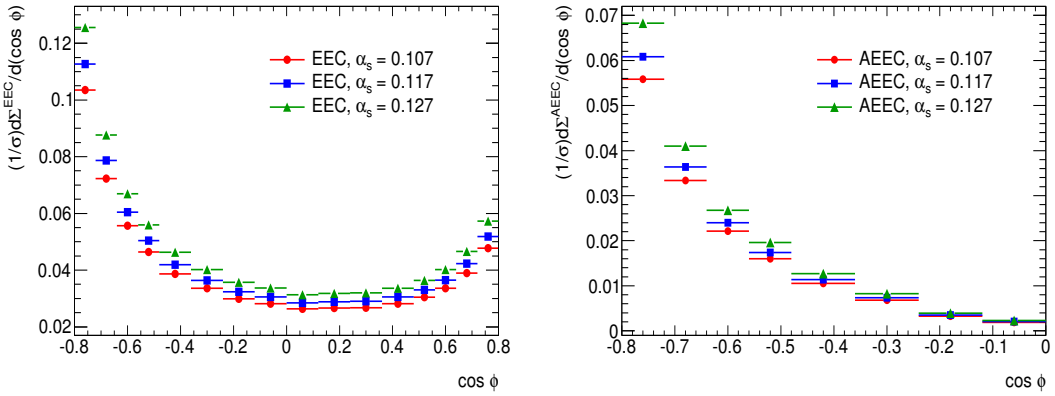

Figure 1.10: Dependence of the TEEC (left) and ATEEC (right) functions with the strong coupling constant $\alpha_s(m_Z)$

The prediction shown in Fig. 1.10 has been ellaborated using the NLOJET++ [37] program. Kinematic cuts on the jets have been set so that their transverse momenta is greater than 50 GeV and their pseudorapidity range corresponds to the central region $|\eta| < 2.5$. The scalar sum of transverse momenta of the two leading jets has to be $p_{\mathrm{T},1} + p_{\mathrm{T},2} > 500$ GeV.

1.6 Jet shapes

Jet shapes [38, 39, 40] are generally accepted as the best-suited observable to study the internal structure of a jet. For jets defined using sequential recombination algorithms with a radius R, jet shapes parametrise the dependence of this

energy spread with the distance $r < R$ of a given jet constituent to the jet axis. The integrated jet shape $\Psi(r)$ is defined as

$$\Psi(r) = \frac{\sum_i p_{\mathrm{T},i}\Theta(r - r_i)}{\sum_i p_{\mathrm{T},i}} \tag{1.60}$$

The index i in Equation 1.60 runs over the jet constituents, which are only entering the sum in the numerator if $r_i < R$. On the other hand, the denominator express the total scalar sum of the jet constituents transverse momenta. Therefore, by definition, the integrated jet shape fulfills $\Psi(R) = 1$.
From the integrated jet shape, one can also define the differential jet shape as its derivative with respect to the clustering distance r

$$\rho(r) = \frac{d\Psi(r)}{dr} \tag{1.61}$$

At leading order in α_s, as 1.61 describes the energy flow $\sim z$ for a jet initiated by a parton of type a, one can write

$$\rho_a(r) = \sum_b \frac{\alpha_s}{\pi r} \int_0^{1-z_0} z P_{ba}(z) dz; \quad z_0 = \begin{cases} \frac{r}{r+R} & \text{if} \quad r < (R_{\mathrm{sep}} - 1)R \\ \frac{r}{R_{\mathrm{sep}} R} & \text{if} \quad r > (R_{\mathrm{sep}} - 1)R \end{cases} \tag{1.62}$$

Here, R_{sep} is a parameter, depending on the jet kinematics, to be fitted from the data [39], and R is the jet radius. Making use of the expressions 1.41 to 1.44 for the leading order parton splitting functions, one finds the LO expressions for the jet shape of quark and gluon -induced jets

$$\rho_q(r) = \frac{\alpha_s C_F}{\pi r} \left[2 \log\left(\frac{1}{z_0}\right) - \frac{3}{2}(1 - z_0)^2 \right] \tag{1.63}$$

$$\begin{aligned} \rho_g(r) = \; & \frac{\alpha_s C_A}{\pi r} \left[2 \log\left(\frac{1}{z_0}\right) - \left(\frac{11}{6} - \frac{z_0}{3} + \frac{z_0^2}{2}\right)(1 - z_0)^2 \right] \\ & + \frac{\alpha_s N_f T_R}{\pi r} \left(\frac{2}{3} - \frac{2z_0}{3} + z_0^2\right)(1 - z_0)^2 \end{aligned} \tag{1.64}$$

1.7 Gluon radiation from heavy quarks

In this section, the angular radiation pattern of a gluon emitted by a quark is studied as a function of the quark mass m_q. Ref. [41] shows that the angular distribution of the radiation off a heavy quark depends drastically on the quark mass. In an approximate approach, for a quark branching $q \to \tilde{q}g$, the invariant mass can be written as $m_q^2 \simeq 2E_qE_g(1-\cos\theta)$ in the regime where the masses are negligible with respect to the energy scale and θ is the angle between the final-state quark and the radiated gluon. In the limit where θ is small, which holds for the jet cone region, one can expand the cosine as the Taylor series $\cos\theta = \sum_{n=0}^{\infty} \frac{(-1)^n\theta^{2n}}{(2n)!}$ and obtain the relation

$$\theta \simeq \frac{m_q}{\sqrt{E_{\tilde{q}}E_g}} = \frac{1}{\sqrt{z(1-z)}}\frac{m_q}{E_q} \tag{1.65}$$

where z is the fraction of energy carried by the radiated gluon ($E_g = zE_q$). For light-quark jets, the dominant effect on the opening angle described by Eq. 1.65 arises from the gluon energy fraction $0 < z < 1$. On the other hand, the opening angle in b-jets is controlled by the heavier mass of the b-quark.
Defining $\theta_0 = m_q/E_q$, the probability of a gluon emission at a small opening angle $\theta < \theta_0 \ll 1$ is given by [42]

$$\frac{d\sigma}{d\omega} = \frac{\alpha_s C_F}{\pi\omega}\frac{(2\sin\theta/2)^2 d(2\sin\theta/2)^2}{[(2\sin\theta/2)^2+\theta_0^2]^2}\left[1+\mathcal{O}(\theta_0,\omega)\right] \sim \frac{1}{\omega}\frac{\theta^2 d\theta^2}{[\theta^2+\theta_0^2]^2} \tag{1.66}$$

In Eq. 1.66, ω corresponds to the energy of the radiated gluon. From here one can infer that for the kinematical region with $\theta < \theta_0$ the amount of radiation is highly suppressed. This effect is known as angular screening, and the region $\theta < \theta_0$ is known as the 'dead cone'.
Defining $\xi = (2\sin\theta/2)^2$ and neglecting terms of order $\mathcal{O}(\theta_0,\omega)$, one can easily infer that the differential cross section as a function of the angle θ has the form

$$\frac{d\sigma}{d\theta} = \frac{d\xi}{d\theta}\frac{d\sigma}{d\xi} = \frac{2\alpha_s C_F}{\pi}\frac{(2\sin\theta/2)^2\sin\theta}{\left[(2\sin\theta/2)^2+\theta_0^2\right]^2}\frac{dz}{z} \tag{1.67}$$

The cross section in Eq. 1.67 is shown in Fig. 1.11 as a function of the quark mass m_q and the quark-gluon angle θ_{qg} for two different values of the energy of the heavy quark, namely 30 and 60 GeV. In this figure, the suppression of the radiation at small angles $\theta < \theta_0$ from the quark direction is clearly seen, as well as the fact that the radiation is less likely to occur for heavier quarks. Also, the effect of the jet energy is clearly seen, showing that for larger energies, the radiation around the quark direction is more collimated and the effect of the dead cone is less pronounced.

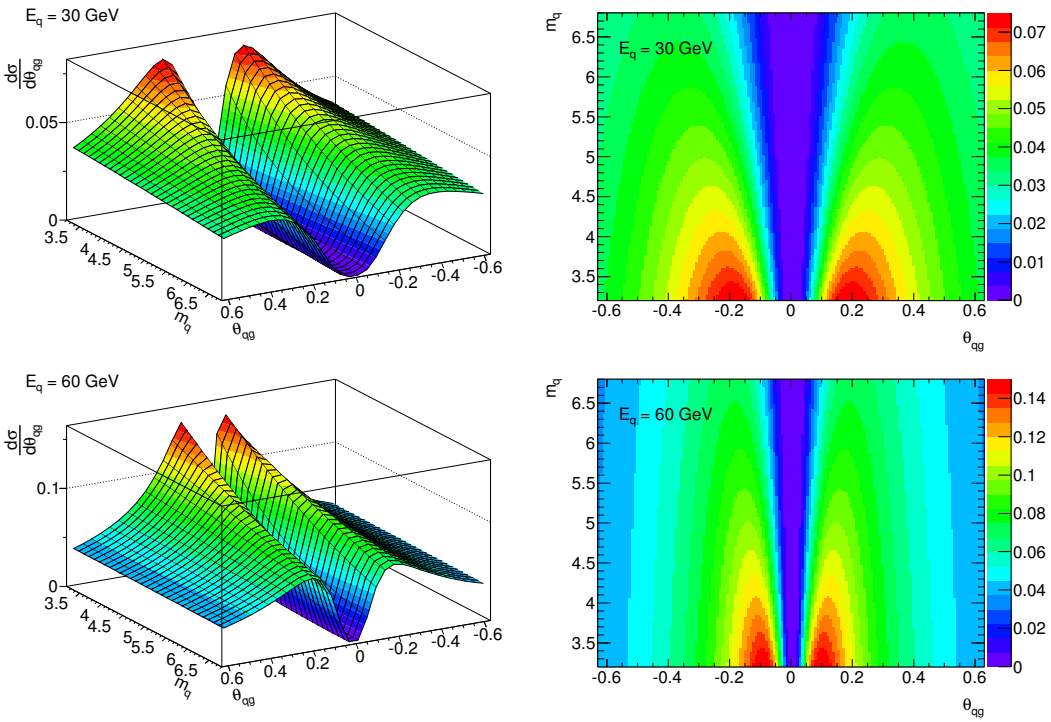

Figure 1.11: The differential cross section of the gluon radiation off a massive quark as a function of its mass m_q and the angle between the quark and the radiated gluon θ_{qg}.

It is then expected that the differential jet shape defined before will be a sensitive observable to the mass of the initiating parton, as the radiated gluons inside the jet will be screened in a cone around the jet axis, and therefore the fraction of transverse momentum in this region will be severely suppressed.

Chapter 2

The LHC and the ATLAS experiment

In this chapter, a description of the experimental details is given, with special attention to the CERN accelerator complex and the ATLAS detector, where the two proton beams collide and the physics processes object of these measurements are produced.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC), see Ref. [43], is the world's largest particle accelerator, buried between 50 and 175 metres under the franco-swiss border in Geneva, Switzerland and spanning a 27 km circumference. Its purpose is to accelerate two proton beams up to a nominal centre-of-mass energy taking the value $\sqrt{s} = 7$ TeV for the analyses presented in this thesis, but which may vary from run to run. Once the nominal energy is reached, both beams are made to collide inside the 4 main experimental detectors ATLAS, ALICE, CMS and LHCb, which are distributed along the accelerator circumference, as shown in Fig. 2.1.

The proton beams are extracted from Hydrogen gas and drift in a smaller system of accelerators containing linear accelerators (LINAC), the Proton Booster and the Proton Synchrotrons (PS, SPS). When the beams reach the LHC, they move around the ring inside vacuum chambers passing through 1232 superconducting

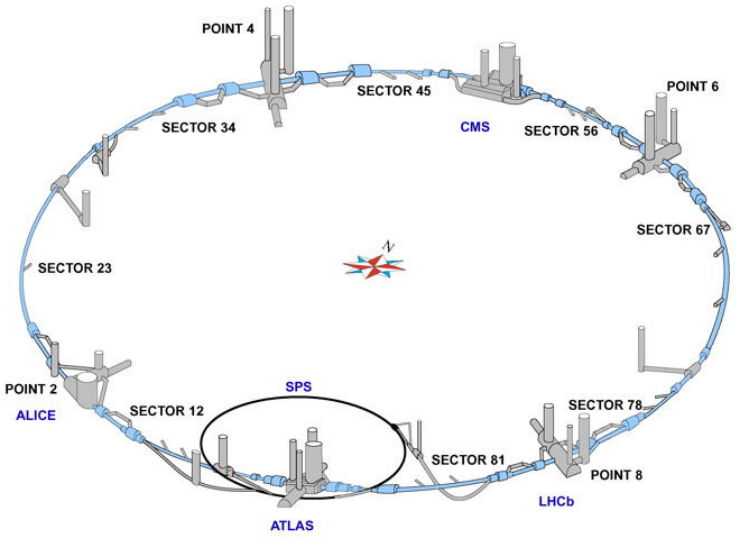

Figure 2.1: The distribution of the LHC experiments ATLAS, ALICE, CMS and LHCb along the LHC ring.

dipole magnets distributed around the circumference. These dipole modules produce a vertical magnetic field which bends the trajectory of the protons via the Lorentz force. At 7 TeV, the magnitude of the magnetic field is around 8.4 Tesla with a current of 11700 A. To avoid excessive resistive losses due to Joule effect with such a high current, a large cryogenics system is needed to provide the liquid helium which is then injected into the dipoles to keep the magnets cold. A standard dipole module is depicted in Fig.2.2

2.1.1 Physics event rate. Luminosity

The motivation for the construction of the LHC is to study the physical processes at the TeV scale, where new physics beyond the Standard Model, such as Supersymmetry (SUSY) could be found, as well as to give answer to some open questions of the Standard Model, as the existence of the Higgs boson.
Every physics process has a probability to occur in a proton-proton collision, which is parametrised by the process cross-section σ. The LHC provides millions of pp collisions per second, giving an expected event rate (the number of events

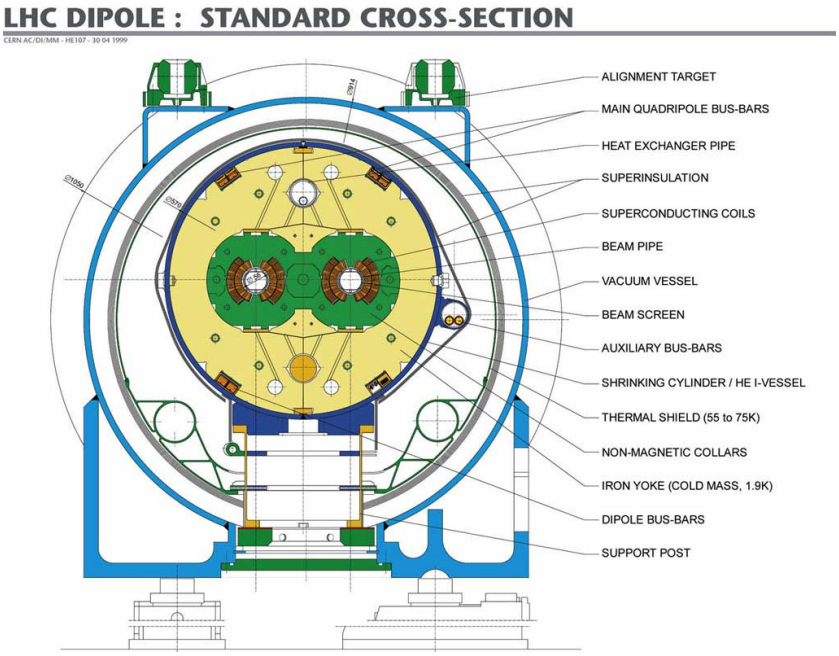

Figure 2.2: The cross section of a dipole module at the LHC

per second) for each process which is proportional to its cross-section [6]

$$R = \frac{dN_{\text{ev}}}{dt} = \sigma\mathcal{L} \Rightarrow N_{\text{ev}}(t_1, t_2) = \sigma \int_{t_1}^{t_2} \mathcal{L}dt = \sigma L \tag{2.1}$$

The proportionality factor $\mathcal{L}$ is referred to as the differential luminosity. It is defined as the number of protons crossing the unit area per unit time, and can be calculated from the beam parameters assuming a gaussian distribution of the proton density inside the beam. If two proton bunches with numbers of particles n_1 and n_2 collide with a frequency f, the differential luminosity is given by

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\Sigma_x\Sigma_y} \tag{2.2}$$

The parameters Σ_x and Σ_y carachterise the beam transverse profiles assuming a 2-dimensional gaussian distribution.

The integrated luminosity L is a key quantity in cross-section measurements and therefore needs to be precisely measured by both the LHC (delivered luminosity) and each experimental detector (recorded luminosity) when the data taking is

ready. In Fig. 2.3, the integrated luminosity delivered to ATLAS as a function of the month in the year for 2010, 2011 and 2012 runs.

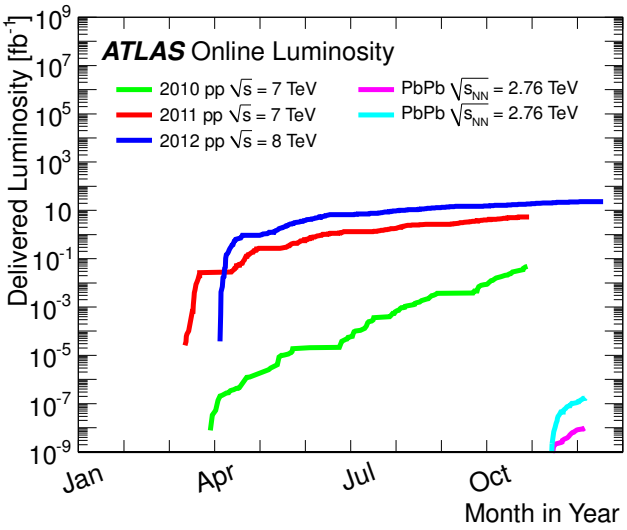

Figure 2.3: Cumulative luminosity versus day delivered to ATLAS during stable beams and for *pp* and Pb-Pb collisions. This is shown for 2010 (green for *pp*, magenta for Pb-Pb), 2011 (red for *pp*, turquoise for Pb-Pb) and 2012 (blue) running. The online luminosity is shown.

2.1.2 Bunch crossing and pileup

The proton beams are not continuous, but have a structure of bunch train. Each bunch contains a high number of protons to maximize the probability of interaction for each bunch crossing. As we have seen, the rate of inelastic events produced in a collider is given by $R_{\text{inel}} = L \times \sigma_{\text{inel}}$, while the average bunch crossing rate is given by the product of the number of bunches per beam and the revolution frequency, $R_{\text{cross}} = N_{\text{bunch}} \times f_{\text{LHC}}$. Therefore, one can define the number of interactions per bunch crossing μ as

$$\mu = \frac{R_{\text{inel}}}{R_{\text{cross}}} = \frac{L \times \sigma_{\text{inel}}}{N_{\text{bunch}} \times f_{\text{LHC}}} \tag{2.3}$$

Attending to the time difference between interactions, one can distinguish two types of pileup collisions. First, the in-time pileup occurs when the pileup signal in

the detector corresponds to the same bunch crossing as the hard-scattering signal, and therefore both are recorded in the same data taking window. On the other hand, the out-of-time pileup is the effect of collisions which have produced in a different bunch crossing has the hard-scattering signal recorded in the detector. In the data taken in 2010, the amount of out-of-time pileup interactions is negligible, as the bunch spacing was smaller than in 2011, when it was rised to the nominal value of 25 ns. The average number of interactions per bunch crossing can be up to $\langle\mu\rangle = 25$ for the data presented here.

2.2 The ATLAS experiment

The ATLAS detector [44] is a multi-purpose particle detector installed at the LHC Point 1, near the CERN Meyrin site. It is a cylindrical apparatus designed to detect the products of the hadron collisions occuring at its center. From the particle interaction point outwards, it contains an Inner Detector (ID), designed to precisely measure the momentum of charged particle tracks, an electromagnetic calorimeter for measuring the energy deposition of electrons and photons, a hadronic calorimeter, used for the measurement of strong-interacting particles and a muon spectrometer for measuring the energy and momentum of muons. A schematic representation of the detector can be seen in Fig. 2.4.

2.2.1 Coordinate system

The ATLAS experiment uses a right-handed coordinate system, which has its center at the nominal interaction point and the z-axis in the direction of the incoming protons. The positive x-axis direction points to the center of the LHC ring and the positive y-axis direction points upwards. Cylindrical coordinates (r, φ) are used to describe the positions in the transverse plane, being φ the azimuthal angle around the beam pipe. The polar angle θ is defined as the angle between the particle direction and the beam axis.
As one does not know the momentum fraction of the partons inside the proton, the collision products can be boosted in the direction of the beam axis. For this reason, one needs to define observables which transforms in a desirable way under

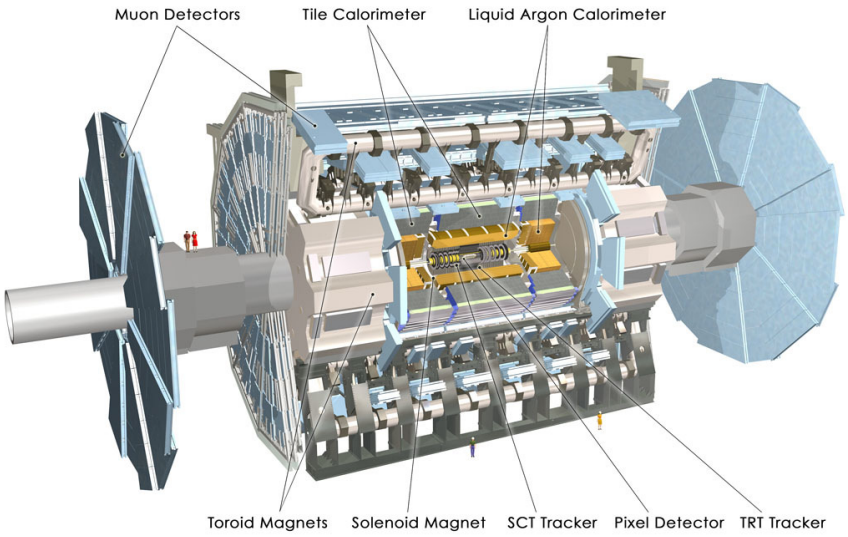

Figure 2.4: General view of the ATLAS detector

z-axis Lorentz boosts. The pseudorapidity η is defined in terms of the polar angle θ as

$$\eta = -\log\left(\tan\frac{\theta}{2}\right) = \frac{1}{2}\log\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right) \tag{2.4}$$

Under a z-axis Lorentz transformation parametrised by a factor γ, it transforms as

$$\eta \to \eta - \log\gamma \tag{2.5}$$

Therefore, the difference in pseudorapidity for two particles i and j is invariant under beam-axis Lorentz transformations. One can also define the rapidity as

$$y = \frac{1}{2}\log\left(\frac{E + p_z}{E - p_z}\right) \tag{2.6}$$

This variable is equivalent to the pseudorapidity for massless objects, where $E = |\vec{p}|$. For the hadron-hadron collisions studied here, once the pseudorapidity is defined, one can define a boost-invariant distance in the η–φ space, ΔR as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2} \tag{2.7}$$

2.2.2 The Inner Detector

The ID is described in detail in Ref. [45]. It is designed to precisely measure the momentum of charged particles produced in the collisions, as well as primary and secondary vertex identification. It covers the pseudorapidity region $|\eta| < 2.5$ and the full azimuthal range and has tracking capabilities for particles with transverse momentum above 500 MeV, but a lower threshold of 100 MeV has also been used for some minimum bias studies. The ID is immersed in a 2 Tesla axial magnetic field which bends the trajectory of charged particles and allows for a measurement of the charge. It consists in three different subsystems, being the innermost the Pixel detector followed by the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). Fig. 2.5 provides a sketch of its structure.

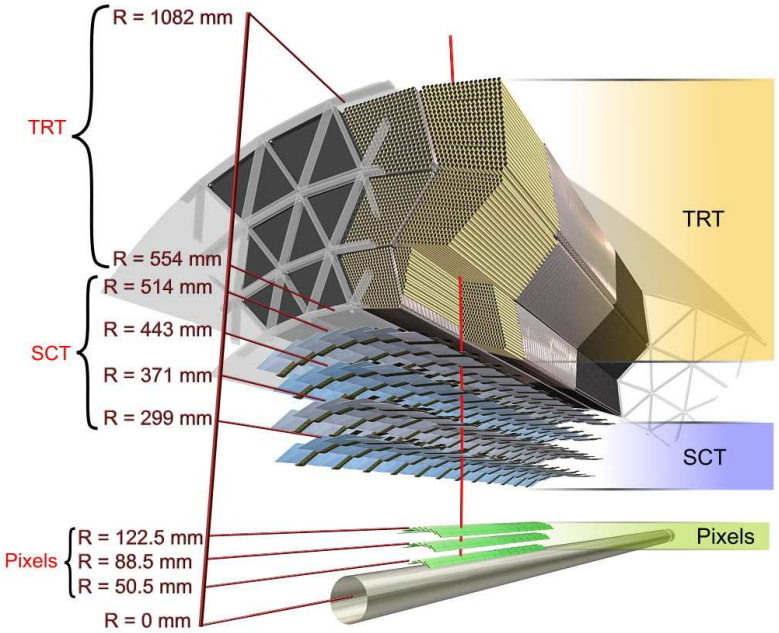

Figure 2.5: Transverse view of the ATLAS Inner Detector subsystems being traversed by a 10 GeV track with $\eta = 0.3$. Each track crosses 3 pixel layers, 4 SCT layers and approximately 35 axial straws in the TRT.

The ID is expected to give an excellent tracking performance. The overall transverse momentum resolution of the ID, resulting from the combined performance

of all subsystems, is required to be

$$\frac{\sigma_{p_{\mathrm{T}}}}{p_{\mathrm{T}}} = 0.05\% \times p_{\mathrm{T}} \oplus 1\% \tag{2.8}$$

The main parameters of the subsystems composing the Inner Detector are given in Table 2.1.

Subsystem	**Position**	**Area** (m^2)	**Resolution** σ (μ**m**)	**Channels** ($\times 10^6$)	η **coverage**
Pixels	1 barrel layer	0.2	$R\varphi = 12$, $z = 66$	16	± 2.5
	2 barrel layers	1.4	$R\varphi = 12$, $z = 66$	81	± 1.7
	4 end-cap disks on each side	0.7	$R\varphi = 12$, $R = 77$	43	1.7-2.5
Silicon strips	4 barrel layers	34.4	$R\varphi = 16$, $z = 580$	3.2	± 1.4
	9 end-cap wheels on each side	26.7	$R\varphi = 16$, $R = 580$	3.0	1.4-2.5
TRT	Barrel straws		170 (per straw)	0.1	± 0.7
	End-cap straws 36 straws per track		170 (per straw)	0.32	0.7-2.5

Table 2.1: The main parameters of the inner detector subsystems. The resolutions quoted are typical values (the actual resolution depends on $|\eta|$).

2.2.2.1 The Pixel Detector

The pixel detector is designed to provide a high granularity as close to the interaction point as possible. It determines the performance of the ID at finding short-lived particles such as b-quarks and τ-leptons.

The system is designed to be highly modular, with three barrels containing approximately 1500 identical barrel modules and four disks on each side containing 1000 disk modules. The whole system uses just one type of support structure in the barrel and one in the disks.

The pixel modules are designed very similarly for the disks and barrel modules. Each barrel module contains 61440 pixel elements, with a readout system consisting in 16 chips. Each of these modules are 62.4 mm long and 22.4 mm wide.

2.2.2.2 The Semiconductor Tracker

The SCT is designed to provide four independent measurements per track in the intermediate radial range, contributing to the measurement of the track momentum, impact parameter and vertex position.
The barrel SCT contains four layers of silicon microstrip detectors, which provide precision measurements of the track coordinates. Each of them has a surface of 6.36×6.40 cm^2 and contain 768 readout strips parallel to the beam direction, providing measurements of the radial and azimuthal coordinates. The readout chain consists on a front-end amplifier and discriminator, followed by a pipeline which stores the hits above a given threshold until the first level of trigger decision.

2.2.2.3 The Transition Radiation Tracker

The TRT consists on gaseous straw detectors, measuring the track radial and azimuthal coordinates up to $|\eta| = 2.0$ with a typical resolution of about 130 μm. Each straw is 4 mm in diameter, providing fast response and good mechanical properties for a maximum straw length of 150 cm.
The barrel is formed by 144 cm long straws, parallel to the beam direction and with their wires divided in two halves approximately at $\eta = 0$. In the end-cap region, the 37 cm long straws are arranged radially in wheels. The number of readout channels is approximately 351000.

2.2.3 The Calorimeter System

The ATLAS calorimeter system is divided into two subdetectors: A Liquid Argon (LAr) electromagnetic calorimeter and a hadronic calorimeter. Its main parameters are detailed in Table 2.2. The calorimeters cover the pseudorapidity region $|\eta| < 4.9$ and are designed for a good containment of electromagnetic and hadronic showers, also providing punch-through containment into the muon system.
Within the $|\eta|$ range matched to the inner detector, the fine granularity of the EM calorimeter is ideally suited for precision measurements of electrons and photons. The coarser granularity of the rest of the calorimeter is sufficient to satisfy the

physics requrements for jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. Fig. 2.6 presents a view of the subsystems composing the ATLAS calorimeter.

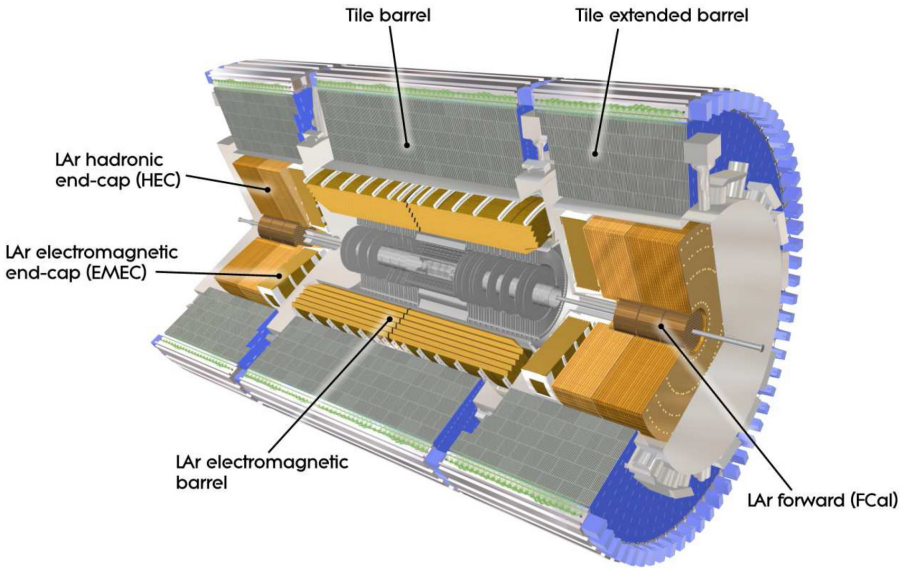

Figure 2.6: Cut-away view of the ATLAS calorimeter system.

2.2.3.1 The LAr electromagnetic calorimeter

The ATLAS electromagnetic calorimeter is divided into a barrel part, with $|\eta| < 1.475$, and two end-caps with $1.375 < |\eta| < 3.2$. The barrel calorimeter consists of two identical half-barrels, separated by a small gap (4 mm) at $z = 0$. Each end-cap is divided into two coaxial wheels: an outer wheel covering the range $1.375 < |\eta| < 2.5$ and an inner one covering $2.5 < |\eta| < 3.2$.

The EM calorimeter has several layers of active material depending on the pseudorapidity region (see Table 2.2). The granularity $\Delta\eta \times \Delta\varphi$ also depends on the η range, being the finest in the central region of the barrel, where maximum precision is required. A sketch of a barrel module is shown in Fig. 2.7, where the granularity is specified for cells in the central region.

In the region with $|\eta| < 1.8$, a presampler detector is used to correct for the energy lost by electrons and photons upstream of the calorimeter. The presampler consists of an active LAr layer of 1.1 cm (0.5 cm) thick in the barrel (end-cap).

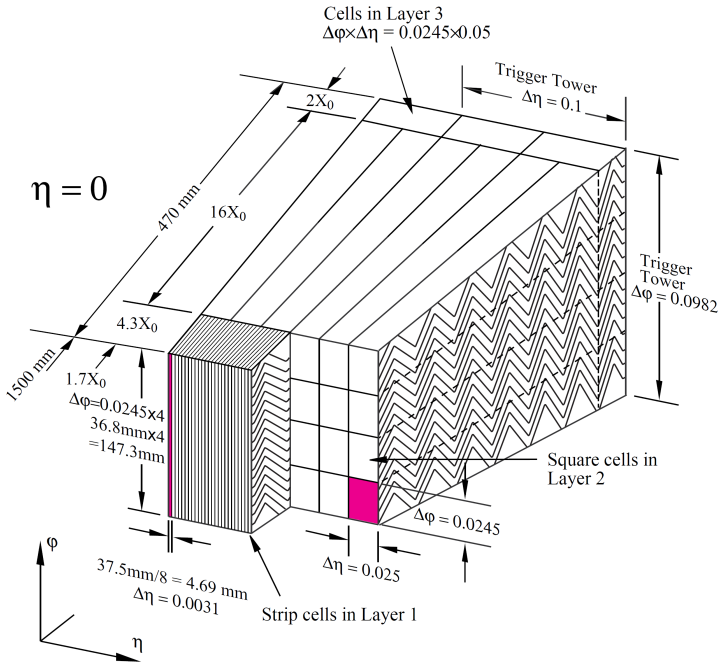

Figure 2.7: Sketch of a barrel module where the different layers are clearly visible with the ganging of electrodes in φ. The granularity in η and φ of the cells of each of the three layers and of the trigger towers is also shown.

The EM calorimeter is a lead-LAr detector with accordion-shaped kapton electrodes and lead absorber plates over its full coverage. The accordion geometry provides a complete azimuthal symmetry without cracks, while the lead thickness in the absorber plates is optimised in terms of energy resolution performance. The resolution is required to be

$$\frac{\sigma_E}{E} = \frac{0.1 \text{ GeV}^{\frac{1}{2}}}{\sqrt{E}} \oplus 0.7\% \tag{2.9}$$

2.2.3.2 The hadronic calorimeters

The hadronic part of the ATLAS calorimeter is composed of several subsystems, namely the tile calorimeter, the LAr hadronic end-cap calorimeter (HEC) and the LAr forward calorimeter (FCal).

	Barrel		End-cap	
EM calorimeter				
Number of layers and $\|\eta\|$ coverage				
Presampler	1	$\|\eta\| < 1.52$	1	$1.5 < \|\eta\| < 1.8$
Calorimeter	3	$\|\eta\| < 1.35$	2	$1.375 < \|\eta\| < 1.5$
	2	$1.35 < \|\eta\| < 1.475$	3	$1.5 < \|\eta\| < 2.5$
			2	$2.5 < \|\eta\| < 3.2$
Granularity $\Delta\eta \times \Delta\varphi$ versus $\|\eta\|$				
Presampler	0.025×0.1	$\|\eta\| < 1.52$	0.025×0.1	$1.5 < \|\eta\| < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$\|\eta\| < 1.40$	0.050×0.1	$1.375 < \|\eta\| < 1.425$
	0.025×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.1	$1.425 < \|\eta\| < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \|\eta\| < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \|\eta\| < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \|\eta\| < 2.4$
			0.025×0.1	$2.4 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 2nd layer	0.025×0.025	$\|\eta\| < 1.40$	0.050×0.025	$1.375 < \|\eta\| < 1.425$
	0.075×0.025	$1.40 < \|\eta\| < 1.475$	0.025×0.025	$1.425 < \|\eta\| < 2.5$
			0.1×0.1	$2.5 < \|\eta\| < 3.2$
Calorimeter 3rd layer	0.050×0.025	$\|\eta\| < 1.35$	0.050×0.025	$1.5 < \|\eta\| < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	
LAr hadronic end-cap				
$\|\eta\|$ coverage			$1.5 \|\eta\| < 3.2$	
Number of layers			4	
Granularity $\Delta\eta \times \Delta\varphi$			0.1×0.1	$1.5 < \|\eta\| < 2.5$
			0.2×0.2	$2.5 < \|\eta\| < 3.2$
Readout channels			5632 (both sides)	
LAr forward calorimeter				
$\|\eta\|$ coverage			$3.1 < \|\eta\| < 4.9$	
Number of layers			3	
Granularity $\Delta x \times \Delta y$ (cm)			FCal1: 3.0×2.6	$3.15 < \|\eta\| < 4.30$
			FCal1: $\sim$ four times finer	$3.10 < \|\eta\| < 3.15$
				$4.30 < \|\eta\| < 4.83$
			FCal2: 3.3×4.2	$3.24 < \|\eta\| < 4.50$
			FCal2: $\sim$ four times finer	$3.20 < \|\eta\| < 3.24$
				$4.50 < \|\eta\| < 4.81$
			FCal3: 5.4×4.7	$3.32 < \|\eta\| < 4.60$
			FCal3: $\sim$ four times finer	$3.29 < \|\eta\| < 3.32$
				$4.60 < \|\eta\| < 4.75$
Readout channels			3524 (both sides)	
Scintillator tile calorimeter				
	Barrel		Extended barrel	
$\|\eta\|$ coverage	$\|\eta\| < 1.0$		$0.8 < \|\eta\| < 1.7$	
Number of layers	3		3	
Granularity $\Delta\eta \times \Delta\varphi$	0.1×0.1		0.1×0.1	
Last layer	0.2×0.1		0.2×0.1	
Readout channels	5760		4092 (both sides)	

Table 2.2: The main parameters of the calorimeter subsystems.

- **Tile calorimeter:** The tile calorimeter barrel covers the region $|\eta| < 1.0$ and its two extended barrels the range $0.8 < |\eta| < 1.7$. It is a sampling calorimeter using steel as the absorber and scintillating tiles as the active material. The barrel and extended barrels are divided into 64 modules extended azimuthally. Radially, it extends from an inner radius of 2.28 m to an outer radius of 4.25 m and is segmented in depth in three layers.

- **LAr hadronic end-cap:** The HEC consists of two independent wheels per end-cap, located directly behind the end-cap electromagnetic calorimeter and sharing the same LAr cryostats. The HEC overlaps with the forward calorimeter to reduce the drop in the material density at the transition region around $|\eta| = 3.1$, extending up to $|\eta| = 3.2$. Each wheel is built from 32 identical wedge-shaped modules, assembled with fixtures at the periphery and at the central bored. Each wheel is divided into two segments in depth for a total of four layers per end-cap.

- **LAr forward calorimeter:** The FCal is integrated into the end-cap cryostats, as this provides clear benefits in terms of uniformity of the coverage as well as reduced radiation levels in the muon spectrometer. The FCal is approximately 10 interaction lengths deep, and consists of three modules in each end-cap: the first, made of copper, is optimised for electromagnetic measurements, while the other two, made of tungsten, measure predominantly the energy of hadronic interactions.

The energy resolution of the hadronic calorimeters is given by

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b \tag{2.10}$$

For the barrel and end-cap hadronic calorimeters, the values of a and b are $a = 0.5$ GeV$^{\frac{1}{2}}$ and $b = 3\%$, while the forward calorimeter has a poorer resolution with $a = 1$ GeV$^{\frac{1}{2}}$ and $b = 10\%$.
The structure of a cryostat, showing the relative positions of the three forward calorimeters is depicted in Fig. 2.8. The outer radius of the cylindrical cryostat vessel is 2.25 m and the length of the cryostat is 3.17 m. The feed-throughs and front-end crates containing the readout electronics are also shown.

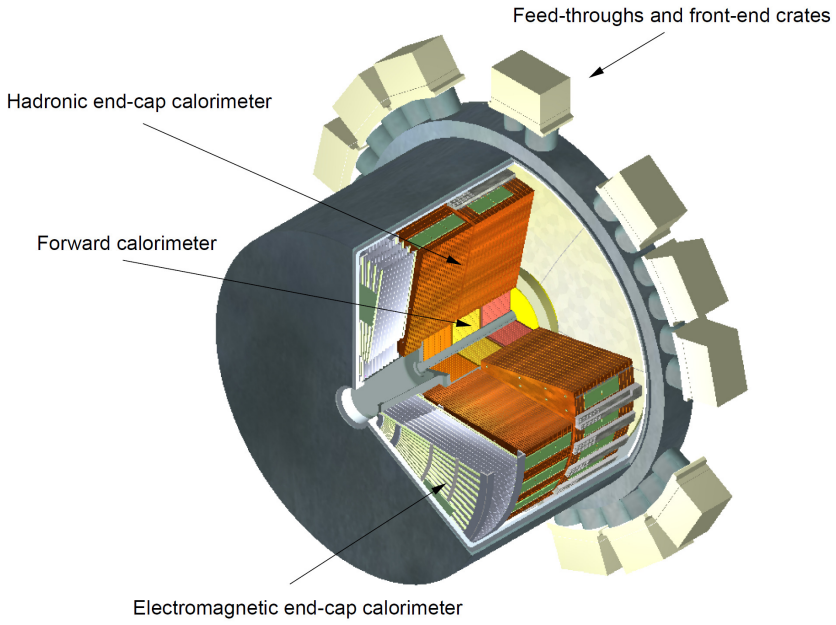

Figure 2.8: Cut-away view of an end-cap cryostat, showing the positions of the three end-cap calorimeters.

2.2.4 The Muon Spectrometer

The ATLAS muon system is based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets. Over the range $|\eta| < 1.4$, the bending is provided by the large barrel toriod, while for $1.6 < |\eta| < 2.7$, the muon trajectories are bent by two smaller end-cap magnets inserted into both ends of the barrel toroid. Over $1.4 < |\eta| < 1.6$, usually referred to as the transition region, the deflection is provided by a combination of barrel and end-cap fields. This configuration provides a field which is mostly orthogonal to the muon trajectories, while minimising the degradation of resolution due to multiple scattering.

In the barrel region, tracks are measured in chambers arranged in three cylindrical layers around the beam axis; in the transition and end-cap regions, the chambers are installed in planes perpendicular to the beam, also in three layers. Fig. 2.9 shows a general view of the muon spectrometer and its components.

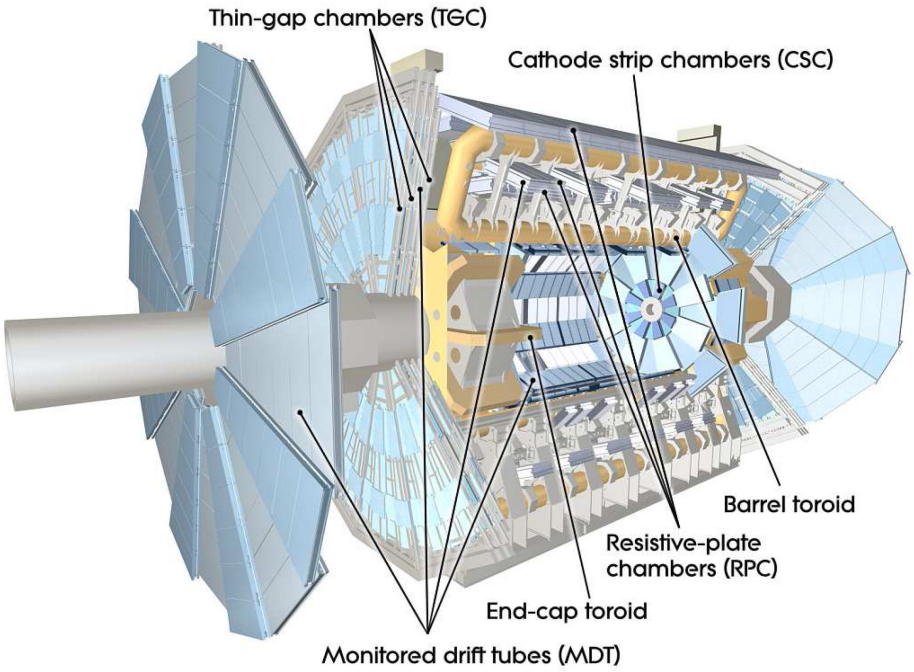

Figure 2.9: Cut-away view of the ATLAS muon system.

The muon spectrometer is divided into four subsystems, namely the Monitored Drift Tubes (MDT), the Cathode Strip Chambers (CSC), the Resistive Plate Chambers (RPC) and the Thin Gap Chambers (TGC). The two first are used for precission tracking purposes, while the two latter provide triggering capabilities. The main parameters of each subsystem, in terms of coverage, number of chambers and number of channels are shown in Table 2.3

System	Coverage	Number of Chambers	Number of Channels	Function
MDT	$\|\eta\| < 2.7$	1150	354000	Precision tracking
CSC	$2.0 < \|\eta\| < 2.7$	32	31000	Precision tracking
RPC	$\|\eta\| < 1.05$	606	373000	Triggering, second coordinate
TGC	$1.05 < \|\eta\| < 2.7$	3588	318000	Triggering, second coordinate

Table 2.3: The main parameters of the subsystems conforming the muon spectrometer.

2.2.4.1 Toroid magnets

The magnetid field in the muon spectrometer is generated by three large air-core toroids. Their performance in terms of bending power is characterised by the line integral $\int \vec{B} \cdot \vec{dl}$ of the magnetic field along an infinite-momentum muon trajectory between the innermost and the outermost muon-chamber planes. The barrel toroid provides 1.5 to 5.5 Tm of bending power in the pseudorapidity range $0 < |\eta| < 1.4$, and the end-cap toroids approximately 1 to 7.5 Tm in the region $1.6 < |\eta| < 2.7$. The bending power is lower in the region $1.4 < |\eta| < 1.6$, where the two magnets overlap.

2.2.4.2 Precision tracking chambers. MDTs and CSCs.

The MDTs perform the precision momentum measurement. They cover the pseudorapidity range $|\eta| < 2.7$ (except in the innermost end-cap layer, where their coverage is limited to $|\eta| < 2.0$). These chambers consist of three to eight layers of drift tubes, operated at an absolute pressure of 3 bar, which achieve an average resolution of 80 μm per tube, or about 35 μm per chamber.
The CSCs are located in the forward region $2 < |\eta| < 2.7$. They are used in the innermost tracking layer due to their higher rate capability and time resolution. The CSCs are multiwire proportional chambers with cathode planes segmented into strips in orthogonal directions. This allows both coordinates to be measured from the induced-charge distribution. The resolution of a chamber is 40 μm in the bending plane and about 5 mm in the transverse plane.

2.2.4.3 Triggering chambers. RPCs and TGCs.

The precision-tracking chambers are complemented by a system of fast trigger chambers capable of delivering track information within a few tens of nanosecond. In the barrel region $|\eta| < 1.05$, the RPCs were selected for this purpose, while in the end-cap $1.05 < |\eta| < 2.4$, the TGCs were chosen. Both chamber types deliver signals with a spread of 15-25 ns. They also provide measurements of both coordinates of the track, one in the bending plane in the η direction and one in the non-bending azimuthal plane.
The purpose of the precision-tracking chambers is to determine the coordinate of

the track in the bending plane. After matching of the MDT and trigger chamber hits in the bending plane, the trigger chamber's coordinate in the non-bending plane is adopted as the second coordinate of the MDT measurement. This method assumes that in any MDT/trigger chamber pair a maximum of one track per event be present, since with two or more tracks the η and φ hits cannot be combined in an unambiguous way.

2.2.5 The Forward Detectors

Three smaller detector systems cover the ATLAS forward region. At ± 17 m from the interaction point lies LUCID (LUminosity measurement using Cerenkov Integrating Detector). It detects inelastic pp scattering in the forward region, providing online relative-luminosity monitoring for ATLAS. The second detector is ALFA (Absolute Luminosity For ATLAS), which is located at ± 240 m. It consists of scintillating fibre trackers located inside Roman pots which are designed to approach as close as 1 mm to the beam. The third system is the Zero-Degree Calorimeter (ZDC), which plays a key role in determining the centrality of heavy-ion collisions. It is located at ± 140 m from the interaction point, just beyond the point where the vacuum beam-pipe divides into two independent pipes. The ZDC modules consist of layers of alternating quartz rods and tungsten plates which will measure neutral particles in the high pseudorapidity region $|\eta| \geq 8.2$. Figure 2.10 shows an schematic view of the placement of these three detectors with respect to the interaction point.

2.2.5.1 The LUCID detector

LUCID consists on twenty 1.5 m long aluminium tubes with a diameter of 15 mm which surround the beam-pipe and point towards the interaction point. The tubes are filled with C_4F_{10} at a constant pressure of $1.2 - 1.4$ bar, providing a Cerenkov threshold of 2.8 GeV for pions and 10 MeV for electrons.
The Cerenkov light emitted by a particle traversing the tube has a half-angle of 3 degrees and is reflected on average three times before the light is measured by photomultiplier tubes (PMTs) which match the size of the Cerenkov tubes. The signal amplitude from these PMTs can be used to distinguish the number of par-

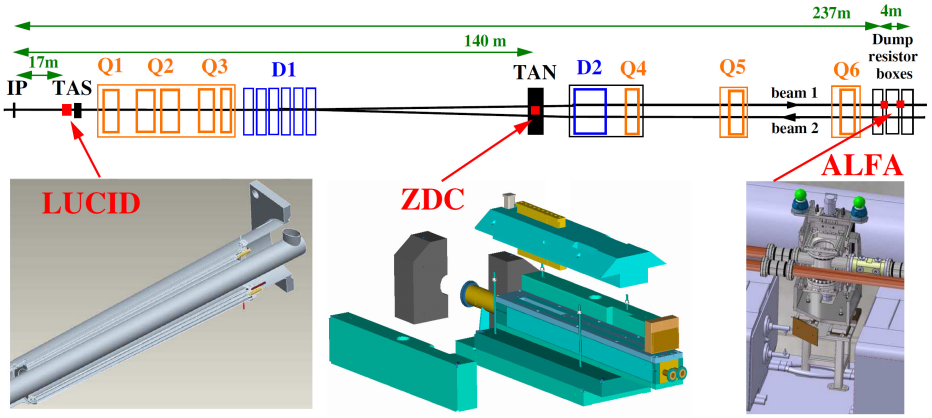

Figure 2.10: Placement of the forward detectors along the beam-line around the ATLAS interaction point (IP).

ticles per tube, and the fast timing response provides unambiguous measurements of individual bunch-crossings.

2.2.5.2 The ZDC detector

The main role of the ZDC detector is to detect very forward neutrons in heavy ion collisions, in which the centrality is strongly correlated with the number of forward neutrons. The ZDCs consist in two arms, located symmetrically with respect to the interaction point. Each arm contains four modules: one electromagnetic and three hadronic.

The EM modules consist of 11 tungsten plates, with their faces perpendicular to the beam direction. A total of 96 quartz rods of 1.0 mm diameter penetrate the tungsten plates, and multi-anode phototubes capture the Cerenkov light from the products of incident particles. The hadronic modules are similar, but instead of mapping the 96 quartz rods onto the pixel of the phototube, they map clusters of four rods into individual phototubes and have only one light-guide.

2.2.5.3 The ALFA detector

The goal of the ALFA detector is the determination of the absolute luminosity via elastic scattering at small angles. The optical theorem connects the elastic-

scattering amplitude in the forward direction to the total cross section, and therefore can be used to extract the absolute luminosity. The ALFA detector consists on ten double-sided modules, each with 64 fibres arranged in stereo u–v geometry on both sides. The square fibres have a width of 0.5 mm. The modules are staggered in depth by multiples of a tenth of the effective fibre pitch, i.e. 70 μm. The fibres are aligned and glued on a precisely machined support structure made of titanium. Before assembly, the fibres are aluminised to reduce light losses and optical cross-talk. The fibre positions are measured by means of optical metrology at various stages of the manufacturing, resulting in a set of equations describing the location of each fibre in a given detector.

2.2.6 The ATLAS Trigger System

The ATLAS Trigger system is composed of three levels of event selection: L1, L2 and Event Filter (EF), collectively referred to as the High Level Trigger (HLT). It is designed to reduce the event rate to about 200 Hz from the nominal bunch crossing rate of 40 MHz. Fig. 2.11 shows a flow diagram of the 3 trigger levels.

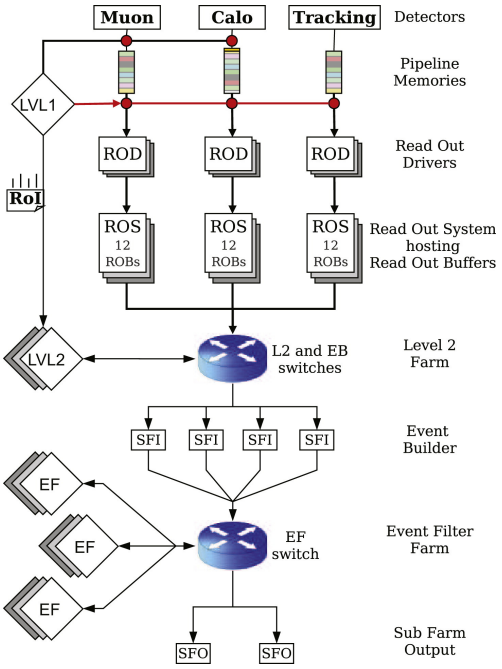

Figure 2.11: Flow diagram of the ATLAS Trigger system [46]

2.2.6.1 Level 1

The L1 trigger performs the initial event selection based on the information from the calorimeters and muon detectors. The calorimeter selection (L1Calo) is based on the information provided by all the calorimeter subsystems, being consistent with high-E_{T} objects such as electrons, photons, hadronically decaying τ-leptons or jet clusters, as well as events with large $E_{\mathrm{T}}^{\mathrm{miss}}$. The L1Calo is designed to work with 7000 analoge trigger towers of reduced granularity ($\Delta\eta \times \Delta\varphi = 0.1 \times 0.1$ in the central region, but larger at high $|\eta|$), see Fig. 2.7. It sends the signal to the L1 computer farm approximately 1.5 μs after the event occurs.
The L1 muon trigger is based on the information provided by the RPC and TGC, the muon chambers with triggering capabilities in the barrel and end-cap regions, respectively. The algorithm is based on the coincidence of aligned hits in different trigger stations, tracking the trajectory of a muon. The low-p_{T} trigger is capable of tracking muons with a threshold range of about 6–9 GeV, while the high-p_{T} trigger has a threshold which varies from 9 to 35 GeV.

2.2.6.2 Level 2 and Event Filter

The L2 is mainly based on regions of interest (RoI) identified at the L1 level. A seed is constructed for each L1 accepted object, that consists of a p_{T} threshold and an $\eta - \varphi$ position for each object. The L2 algorithms make use of information which is not available at the L1 level, such as the Inner Detector tracks. The L2 trigger reduces the event ratio below 3.5 kHz, while takes the decission of keeping or rejecting the event in approximately 40 ms.
The Event Filter must provide additional rejection to reach the desired 200 Hz rate. It is typically based on very similar algorithms to those from the offline selection, as the EF has access to the complete data from a given event, because the EF selection is made after the event building step. The average event processing time is about 4 s.

Chapter 3

Object reconstruction and identification

In this chapter, the performance of the ATLAS detector in terms of physics object reconstruction and identification is discussed, including tracks and vertices, electrons, muons, jets, b-jets and missing transverse energy. These objects will be later used in the physics analyses presented in this work.

3.1 Tracks and vertices

Track reconstruction is essential for electron and muon identification, as well as for track-jet studies. It is done using all subsystems in the inner detector, including the pixel detector, the SCT and the TRT. The track reconstruction software includes global-χ^2 and Kalman-filter techniques [47], and divides the processing into three stages:

- **Pre-processing:** The raw data from the pixel and SCT detectors are converted into clusters, while the TRT raw timing information is calibrated. The SCT clusters are transformed into space-points, combining the cluster information from opposite sides of the SCT modules.
- **Track-finding:** Track seeds are formed by combining space-points in the pixel detector and the first SCT layer. These seeds are then extended

throughout the SCT to form track candidates, which are then fitted and subject to quality cuts. The selected tracks are then extended into the TRT to associate the drift-circle information. Finally, the extended tracks are refitted including the information from all three subsystems.

- **Post-processing:** A vertex-finding algorithm is used to reconstruct primary vertices, followed by the reconstruction of photon conversions and secondary vertices. The vertex candidates are selected by minimsing a χ^2 function which depends on the (x, y, z) position as well as on the track momenta $\vec{p}_i$. The primary vertex is defined as the one maximising the quantity $\sum_i (p_{\mathrm{T}}^i)^2$, where the index i runs over all tracks originating from it.

3.2 Electrons

ATLAS requires a high efficiency electron reconstruction to achieve physics goals such as the Higgs to four lepton $H \to 4\ell$ reconstruction or the measurement of top-quark properties. The electron reconstruction starts with a clustering algorithm maximising the energy containment and follows with differents sets of cuts to separate the high jet background from the electron signal. This section aims to describe each step of this process.

3.2.1 Calorimeter-seeded reconstruction

Electron (and photon) reconstruction starts with a three-step clustering algorithm, referred to as the sliding-window method, described in detail in Ref. [48].

- **Tower building:** The η–φ space in the calorimeter is divided into a grid of towers with size $\Delta\eta \times \Delta\varphi$, which depends on the calorimeter system involved. For the electromagnetic case described here, both the EM Barrel and Endcap cells, with granularity $\Delta\eta \times \Delta\varphi = 0.025 \times 0.025$ are used as towers for $|\eta| < 2.5$.

- **Seed finding:** A window of 5×5 EM cells is moved across the tower grid defined in the previous step. The total transverse energy of the window (defined as the sum of the transverse energy of each cell within the window)

is required to be greater than 3 GeV and to be a local maximum. If these two conditions are fulfilled, then the window is selected as a precluster with η and φ calculated as the energy-weighted average of the position of all cells within a new 3×3 window around the center of the sliding window. Precluster duplicates are removed.

- **Cluster formation:** Once the seeds are defined, a layer-by-layer clustering algorithm is ran by considering all 3×7 (Barrel) and 5×5 (Endcap) rectangles centred in a given seed position. The algorithm starts with the middle layer, using as seed the precluster position computed in the previous step. The strip layer is done in the second place, taking as input the middle-layer barycenter. Finally, the Presampler and back layers are included using the strip and middle layer barycenters as seeds.

Once the electron cluster has been calculated, it is required to match a track from the inner detector. This is done by extrapolating the track position to the LAr calorimeter and requiring it to be within a $\Delta\eta \times \Delta\varphi = 0.05 \times 0.10$ rectangle centred at the cluster position. If the ratio $E/|\vec{p}|$ between the cluster energy and the track momentum is smaller than 10, both the track and the cluster are combined to form the electron candidate. The energy is calculated as the weighted average between the cluster energy and the track momentum, while the angular coordinates η and φ correspond to those from the track.
The electron reconstruction efficiency is high: Approximately 93% of true isolated electrons with $E_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$ are selected as candidates. The 7% inefficiency is understood to arise from the large amount of material in the inner detector, and therefore is η-dependent.

3.2.2 Electron identification

The major goal of the ATLAS electron identification algorithm is to separate the electron signal from the huge multijet background arising from QCD mediated processes. This is achieved by considering different sets of cuts based on several variables. Three electron identification categories are built using this information, referred to as Loose, Medium and Tight [49]

3.2.2.1 Loose cuts

The loose set of cuts is based on the different shower shapes between jets and electrons, as well as in their different level of penetration in the hadronic calorimeter. They involve a cut in the lateral shower width, defined as

$$w_{\eta,2} = \sqrt{\frac{\sum_i E_i \eta_i^2}{\sum_i E_i} - \left(\frac{\sum_i E_i \eta_i}{\sum_i E_i}\right)^2} \tag{3.1}$$

The sum is calculated in a window of 3×5 cells centred at the cluster position, being E_i the cell energy and η_i its pseudorapidity. The loose cuts provide a background rejection factor of about 500 and a high efficiency selection.

3.2.2.2 Medium cuts

The medium cuts provide a jet rejection factor of 5000. They include information from both the calorimeter and the inner detector. The calorimetric cuts are designed to provide a good rejection of the neutral pion decays to two photons, $\pi^0 \to \gamma\gamma$. In the case of π^0 decays, the calorimeter is capable of resolving a second maximum with energy $E_{\mathrm{max},2}$ in a window of size $\Delta\eta \times \Delta\varphi = 0.125 \times 0.2$ around the cell with maximum energy deposit. The calorimetric variables also include the total shower width, defined as

$$w_{\mathrm{stot}} = \sqrt{\left[\sum_i E_i (i - i_{\mathrm{max}})^2\right] \sum_i E_i} \tag{3.2}$$

In this case, the index i runs over all strips in a window of $\Delta\eta \times \Delta\varphi = 0.0625 \times 0.2$, and i_{max} is the index of the maximum-energy strip. The tracking variables include the number of pixel and SCT hits as well as the transverse impact parameter, defined as the transverse distance of closest approach to the primary vertex. A track-cluster matching within $|\Delta\eta_{\mathrm{ct}}| < 0.01$ is also applied.

3.2.2.3 Tight cuts

The tight cuts provide a rejection factor of 50000. They are based on the tightening of the track-cluster spatial matching ($|\Delta\varphi| < 0.02$, $|\Delta\eta| < 0.005$) as well

as in the coincidence of the cluster energy with the track momentum through the ratio E/p. The transverse impact parameter cut is also tightened and the TRT hit information is used for providing further rejection. Electrons from photon conversions are discarded.

3.3 Muons

The ATLAS muon system is designed to precisely measure the momentum of muons in an energy range which tipically goes from a few GeV to the TeV scale. This is important for the reconstruction of many physics processes such as the $t\bar{t}$ production treated in this work or new physics searches such as $Z' \to \mu\mu$ or multilepton SUSY searches.
The muon algorithms available within ATLAS include three schemes for the reconstruction, depending on how the informations from the inner detector and the muon spectrometer are used for the muon reconstruction: standalone muons, combined muons and tagged muons. The standalone muon reconstruction starts from the muon spectrometer information and extrapolates it to the beam line. The spectrometer tracks are defined in terms of five parameters, tipically evaluated at the perigee, i.e. the point of maximum approach to the beam axis. These parameters include the transverse and longitudinal impact parameters d_0 and z_0; the polar and azimuthal angles θ_0 and φ_0 and the ratio of the charge to the track momentum, q/p. The matching of the spectrometer information with the inner detector track results in the combined muon, whereas the tagged muon reconstruction is done from the inner detector to the muon spectrometer. For each of these strategies, two algorithms are available: Staco [51] and Muid [52].

3.3.1 Standalone muons

The standalone muon algorithm starts by building three track segments in the spectrometer chambers. These segments are then linked with each other to form a track, which is then extrapolated to the perigee by the Staco or Muid algorithms. In the Staco case, the extrapolation algorithm is called Muonboy [51], while the Moore [53] algorithm is used on the Muid side.

This extrapolation is done in a way in which energy losses in the calorimeter are taken into account. The Muonboy algorithm does this correction by estimating the amount of material crossed by the muon depending on its pseudorapidity, while Moore also includes calorimetric measurements, in case they are larger than the typical value for the energy loss.
Muons produced in hadron decays in the calorimeter, such as π and K decays, are normally reconstructed in the standalone scheme. They give an estimation of these kind of backgrounds when the matching to the inner detector case is performed.

3.3.2 Combined muons

The matching of the standalone muons to the inner detector tracks is done by the Staco and Muid algorithms by using the five-parameter vector $(d_0, z_0, \theta_0, \varphi_0, q/p)$ described before. The χ^2 of this matching provides a measure of the quality of the matching procedure and it is defined as

$$\chi^2_{\text{match}} = (\mathbf{T_{MS}} - \mathbf{T_{ID}})^{\text{T}}(\mathbf{C_{ID}} + \mathbf{C_{MS}})^{-1}(\mathbf{T_{MS}} - \mathbf{T_{ID}}) \tag{3.3}$$

Here, $\mathbf{T}$ denotes the five-parameter vector expressed at the perigee, while $\mathbf{C}$ is its covariance matrix. The subscripts ID and MS stand for 'inner detector' and 'muon spectrometer' respectively. While Staco does a statistical combination of the ID and MS measurements to obtain the combined muon parameters:

$$\mathbf{T} = (\mathbf{C_{ID}}^{-1} + \mathbf{C_{MS}}^{-1})(\mathbf{C_{ID}}^{-1}\mathbf{T_{ID}} + \mathbf{C_{MS}}^{-1}\mathbf{T_{MS}}) \tag{3.4}$$

The Muid fit accounts for the calorimetric energy losses as well as for the magnetic field in both the calorimeter and the muon spectrometer.

3.3.3 Tagged muons

The tagged muon algorithms MuTag [51] and MuGirl [54] do the extrapolation in opposite direction as the combined muon algorithms. Starting from all inner detector tracks over a certain p_{T} threshold, MuTag defines a χ^2 function using the

extrapolation of the inner detector track and any spectrometer hits in its vicinity, while MuGirl uses a neural-network based algorithm to define a discriminant function. If muon segments are found in the spectrometer in the surrounding region of the extrapolated track, the tagged muon is stored.
The tagged muon reconstruction is optimised for low p_{T} muons. Since muons appearing in $t\bar{t}$ decays have a high-p_{T} threshold, Muid combined muons will be used for the analysis presented here.

3.4 Jets

Jets are the main object used in the analyses presented in this work. As already discussed in Section 1.4, jets are collimated showers of hadrons, defined in an attempt to reconstruct the kinematic properties of the partons from which they originated. The primary objects used in ATLAS for jet building are topological clusters, constructed from calorimeter cells and used as input four-momenta for the anti-k_t algorithm. Then, further calibrations are needed to correct the jet energy measurement for calorimetric non-compensation and leakage, together with other detector effects.

3.4.1 Topological clusters

The basic idea of the topological clustering algorithms [48] is to combine groups of cells with a significant amount of energy, well above the noise threshold. It consists on two steps: cluster making and cluster splitting.

3.4.1.1 Cluster making

The procedure to construct clusters from calorimeter cells is illustrated in Fig. 3.1, and it consists on the following steps:

- **Seeding:** In the first step, the seeds for the clustering algorithm are constructed by selecting calorimeter cells with a signal energy which is greater than four times the RMS of the noise of that cell, $|E| > 4\sigma$, where σ takes into account both electronic and pileup noise.

- **Adding neighbouring cells:** After the seeds are selected, their neighbouring cells are considered. If a neighbouring cell has not been marked as a seed and its signal energy is greater than two times the RMS of the noise, $|E| > 2\sigma$, they are added to the proto-cluster. If the cell is neighbouring two proto-clusters at the same time, the two proto-clusters are merged. In the final step, all the perimeter cells are added to the cluster, regardless of their signal to noise ratio threshold.

- **Finalize:** The cluster making algorithm is finalized by removing clusters whose total energy is smaller than a given threshold.

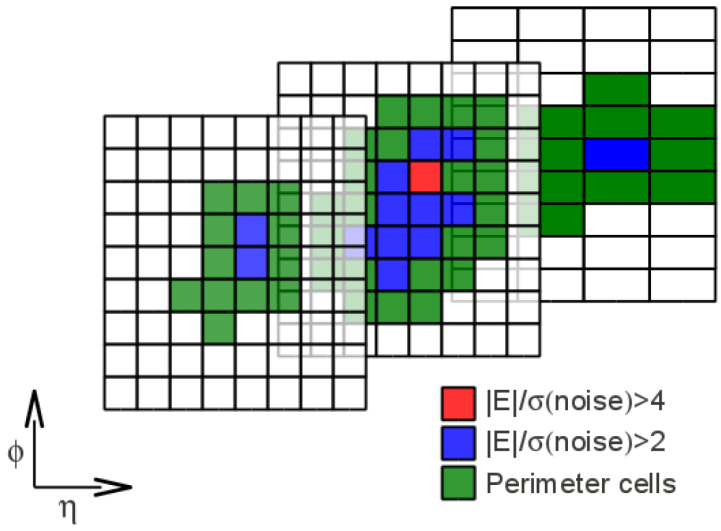

Figure 3.1: Sketch of the formation of a 3-dimensional topological cluster attending to the signal to noise ratio of the constituent calorimeter cells.

3.4.1.2 Cluster splitting

The second part of the clustering algorithm is designed to avoid common situations where topological clusters grow to cover large parts of the calorimeter system. It consists on the following steps:

- **Finding local maxima:** Groups of cells from the topological clusters formed following the cluster making algorithm are selected attending to the following criteria:

- The group of cells has $E > 500$ MeV.
- The energy of the group is greater than that of any other neighbouring cell.
- The number of neighbouring cells inside the initial cluster is above a given threshold (usually 4).

- **Finding neighbours:** Once a local maximum is found, a seed list containing them is constructed. The originally clustered cells surrounding each seed are listed in a neighbour seed list and included to adjacent proto-clusters. If a cell is included into two different proto-clusters, those two whose neighbours have the largest energy will share the cell. Those sharing cells are removed from the list of neighbours and added to a shared cell list to be processed later. In the next iteration, the neighbour list becomes the new seed list. The algorithm finishes when the seed list is empty.

- **Shared cells:** The shared cell list produced in the previous step is expanded by adding neighbours that, being in the original cell set, have not yet been assigned to a proto-cluster. The cells in the expanded list is then added to the two proto-clusters which were to share the cell, with weights given by

$$w_1 = \frac{E_1}{E_1 + rE_2}; \quad w_2 = 1 - w_1; \quad r = e^{d_1 - d_2} \tag{3.5}$$

  Here E_1 and E_2 are the energies of the close-by proto-clusters and d_i are the distances of the shared cells to the centroids of each proto-cluster. These weights (which in practice are usually close to either 0 or 1) give an estimate of the probability of the shared cell to belong to either proto-cluster.

- **Finalize:** Once the previous steps have been accomplished, a proto-cluster has been formed around each local maximum. All parent clusters without a local maxima are also added to the proto-cluster list and converted to the final list of clusters.

3.4.1.3 Local cluster weighting calibration (LCW)

Once the topological clusters are formed using the procedure described in the previous sections, two possibilities exist: to use these clusters as inputs to the jet algorithms described in Sect. 1.4.1 and then apply the jet calibration procedure (EM clusters), or to calibrate their energy individually before any attempt to reconstruct the jets (LCW clusters). The second procedure is called the local cluster weighting (LCW) calibration [55, 56] and it starts with the classification of clusters in electromagnetic or hadronic attending to their energy, depth in the calorimeter and cell energy density. Three weights are then applied separately to account for different effects:

- The calorimeter cells in the clusters are weighted according to the cluster energy and the energy density of the cell. This takes into account the response to hadrons in the calorimeter.
- The cluster is weighted according to the energy deposits in the cluster neighbourhood and to the longitudinal depth of the cluster barycenter λ_c. This accounts for energy deposits not contained within the cluster.
- Finally, the cluster is weighted according to its energy and the fraction of energy in each layer of the calorimeter. This accounts for energy deposited in the calorimeter dead material.

The depth of the cluster barycenter is characterized by the distance λ_c from the front of the calorimeter to the center of the shower. The center of the shower has spatial coordinates given by

$$\langle x_i \rangle = \frac{\sum_k E_k x_i^{(k)}}{\sum_k E_k} \tag{3.6}$$

The index i runs over the three spatial coordinates (x, y, z), and the index k runs over all cells within the clusters with positive energy $E_k > 0$. The shower axis is then determined from the spatial correlation matrix whose elements C_{ij} are given by

$$C_{ij} = \frac{\sum_k E_k^2 (x_i^{(k)} - \langle x_i \rangle)(x_j^{(k)} - \langle x_j \rangle)}{\sum_k E_k^2} \tag{3.7}$$

The shower axis is the eigenvector of this matrix closest to the vector joining the interaction point and the shower center. As before, the index k runs over all cells with positive energy. After these weighting techniques, the resulting energy of the clusters is compared to the energy before the weighting (EM scale). Fig. 3.2 shows the cluster average response $\langle E_{\text{calib}}/E_{\text{EM}}\rangle$ as a function of the cluster energy and pseudorapidity for the three different weighting procedures described above to correct for the low hadronic response, energy deposits outside the topological cluster and non-active material in the calorimeter. The jets used for these studies have transverse momenta above 20 GeV.

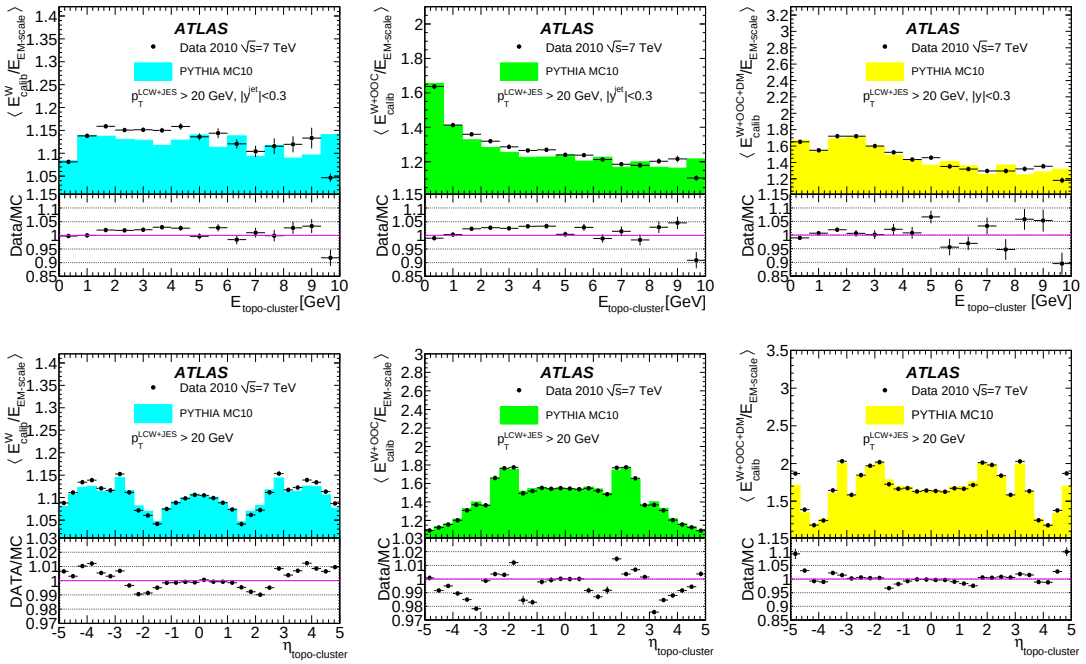

Figure 3.2: Average value of the topocluster response $\langle E_{\text{calib}}/E_{\text{EM}}\rangle$ for hadronic response (left column), out-of-cluster weights (center column) and dead material weights (right column). The top row presents the results as a function of the cluster energy for jets with $|y| < 0.3$, while in the bottom row results are presented as a function of the cluster pseudorapidity [56].

The results above show that the agreement between the data and the Monte Carlo simulation for the average cluster responses to each weighting procedure is within 5%, and it is better for low energy and central pseudorapidity clusters.

3.4.2 Jet calibration

Jets seeded from the EM or LCW topological clusters described above are then reconstructed using the anti-k_t sequential recombination algorithm described in Sect. 1.4.1. Once the jets four-momenta $(E, p_{\mathrm{T}}, \eta, \varphi)$ are reconstructed, a jet energy calibration algorithm restores the jet energy to that of the corresponding particle level jet. This is done following a four-step process [57]

3.4.2.1 Pileup correction

Pileup interactions, understood as multiple pp scatterings in a single bunch crossing, cause the reconstructed jet to have an energy offset with respect to the particle level. A correction is derived from MC simulation as a function of the number of primary interaction vertices N_{PV}, which is a measure of the actual number of collisions in a given event or in-time pileup, and the average number of interactions per bunch crossing μ, which is sensitive to the so-called out-of-time pileup corresponding to crossed signals from previous events. This observable was defined in Eq. 2.3.

The pileup offset correction $\mathcal{O}$ is derived from a given pileup condition $(N_{\mathrm{PV}}^{\mathrm{ref}}, \mu^{\mathrm{ref}})$ so that $\mathcal{O}(N_{\mathrm{PV}}^{\mathrm{ref}}, \mu^{\mathrm{ref}}) = 0$. One can therefore expand to first order in a Taylor series around this point and write

$$\begin{aligned} p_{\mathrm{T}}^{\mathrm{reco}} &= p_{\mathrm{T}}^{\mathrm{truth}} + \mathcal{O}(N_{\mathrm{PV}}^{\mathrm{ref}}, \mu^{\mathrm{ref}}) = \\ p_{\mathrm{T}}^{\mathrm{truth}} + \left(\frac{\partial \mathcal{O}}{\partial N_{\mathrm{PV}}}(\eta_{\mathrm{det}})(N_{\mathrm{PV}} - N_{\mathrm{PV}}^{\mathrm{ref}}) + \frac{\partial \mathcal{O}}{\partial \mu}(\eta_{\mathrm{det}})(\mu - \mu^{\mathrm{ref}}) \right) &= \\ p_{\mathrm{T}}^{\mathrm{truth}} + \alpha(\eta_{\mathrm{det}})(N_{\mathrm{PV}} - N_{\mathrm{PV}}^{\mathrm{ref}}) + \beta(\eta_{\mathrm{det}})(\mu - \mu^{\mathrm{ref}}) & \end{aligned} \tag{3.8}$$

Here η_{det} is the pseudorapidity of the reconstructed jet. The coefficient $\alpha(\eta_{\mathrm{det}})$ is extracted from the MC simulation as the slope of the relation between $p_{\mathrm{T}}^{\mathrm{reco}}$ and N_{PV} in bins of η_{det} and $p_{\mathrm{T}}^{\mathrm{truth}}$. Figure 3.3 shows these relations for two different values of the jet radius parameter R in the Monte Carlo simulation. The slope is found to be independent of the truth jet transverse momentum due to the fact that the pileup interactions do not depend on the hard scattering process.

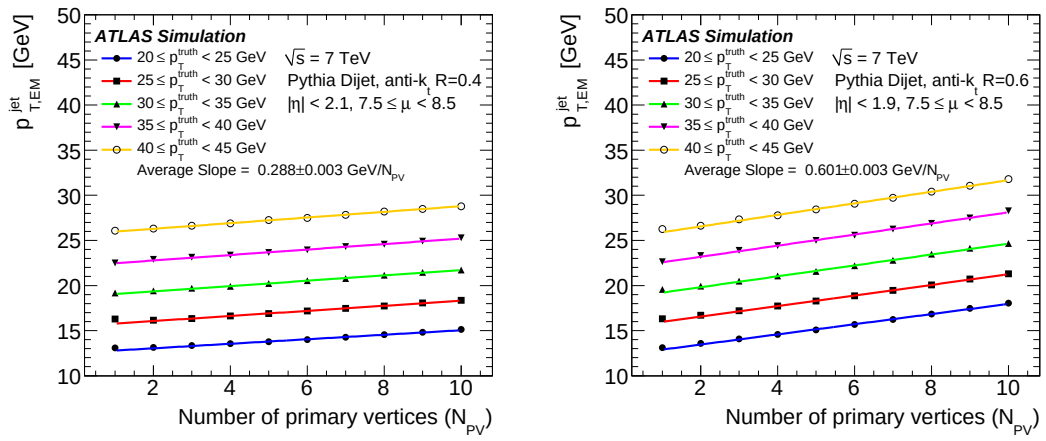

Figure 3.3: Average value of the reconstructed jet transverse momentum $p_{\mathrm{T}}^{\mathrm{reco}}$ as a function of the number of primary vertices N_{PV} for several bins of the truth-jet transverse momentum $p_{\mathrm{T}}^{\mathrm{truth}}$ and $7.5 \leq \mu \leq 8.5$. This is shown for $R = 0.4$ jets (left) and for $R = 0.6$ jets (right) at the EM scale [57].

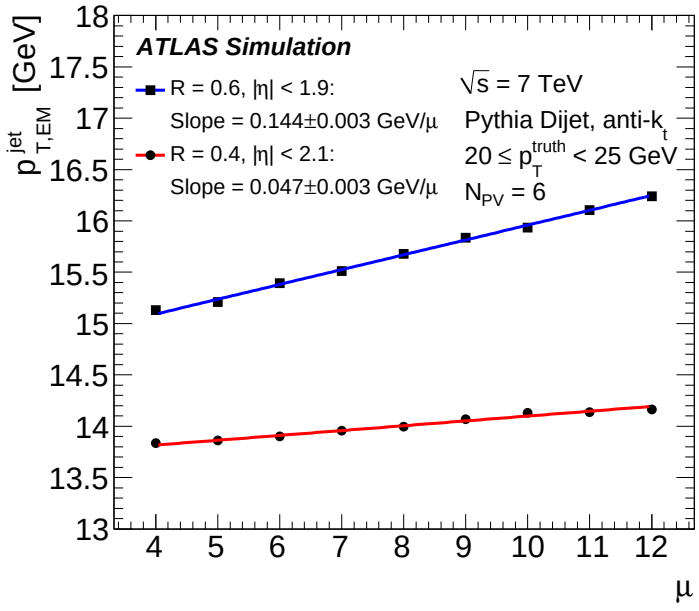

Figure 3.4: Average value of the reconstructed jet transverse momentum $p_{\mathrm{T}}^{\mathrm{reco}}$ as a function of the average number of interactions per bunch crossing μ for $20\ \mathrm{GeV} < p_{\mathrm{T}}^{\mathrm{truth}} < 25\ \mathrm{GeV}$ and $N_{\mathrm{PV}} = 6$ [57].

Similarly, the coefficient $\beta(\eta_{\text{det}})$ is extracted from the MC simulation as the slope of the linear relation between the reconstructed jet transverse momentum and the value of μ for given values of N_{PV} and in bins of the truth jet p_{T}. Figure 3.4 shows the dependence of $p_{\text{T}}^{\text{reco}}$ for EM calibrated jets. Once the pileup offset correction $\mathcal{O}^{\text{EM (LCW)}}$ has been calculated for the given pileup configuration (N_{PV}, μ) in each event and the pseudorapidity η_{det} for each jet, the jet transverse momentum is corrected according to

$$p_{\text{T,EM}}^{\text{corr}} = p_{\text{T,EM}}^{\text{jet}} - \mathcal{O}^{\text{EM}}(N_{\text{PV}}, \mu, \eta_{\text{det}}) \tag{3.9}$$

$$p_{\text{T,LCW}}^{\text{corr}} = p_{\text{T,LCW}}^{\text{jet}} - \mathcal{O}^{\text{LCW}}(N_{\text{PV}}, \mu, \eta_{\text{det}}) \tag{3.10}$$

3.4.2.2 Origin correction

After the pileup offset correction has been applied to a jet, its direction in the calorimeter is corrected so that the jet points to the primary interaction vertex instead of to the geometrical center of the detector. The jet energy remains unchanged in this step.

3.4.2.3 Calibration from MC simulations

The calibration of the energy of a given jet is performed by multiplying it by simple correction factors. To this end, it is useful to define the jet response function as

$$\mathfrak{R}^{\text{EM (LCW)}} = \frac{E_{\text{reco}}^{\text{EM (LCW)}}}{E_{\text{truth}}} \tag{3.11}$$

This response function is the inverse of the correction factors, calculated for different values of the jet energy and pseudorapidity. Figure 3.5 shows the jet response for jets constructed from both EM and LCW clusters in the full η_{det} range of the detector.

3.4.2.4 In situ corrections

The last step of the jet calibration is performed by deriving an in situ correction using relative calibrations between the central and forward rapidity regions. This is referred to as the η intercalibration and it is described in Ref. [57]. The

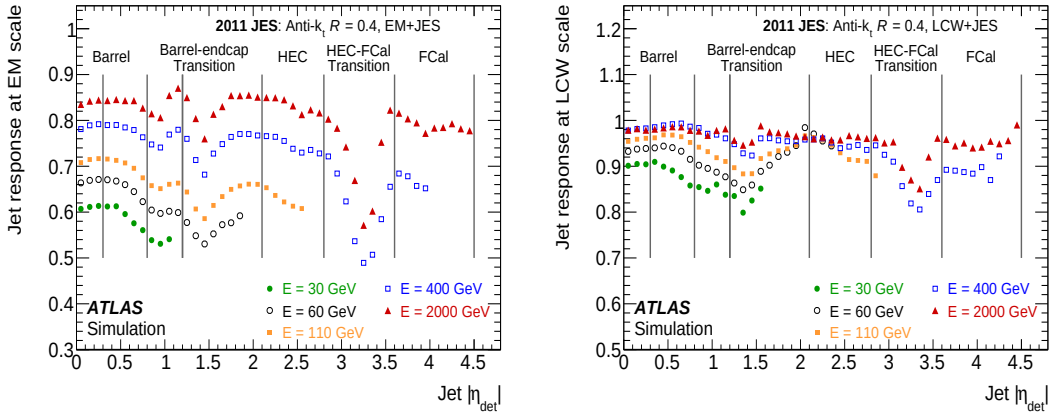

Figure 3.5: Average value of the jet response as a function of the pseudorapidity $|\eta_{\text{det}}|$ for several values of the jet energy E for both the EM (left) and the LCW (right) calibration schemes [57].

jet calibration is also tested in data using the jet momentum balance in several physics processes such as $Z/\gamma^* + 1$ jet, $\gamma +$ jet and the balance between a high-p_{T} jet and the recoiling hadronic system.

3.4.3 Uncertainties in the JES calibration procedure

In this section, the main components of the uncertainty due to the calibration procedure, referred to as the JES uncertainty, are described. Mainly, the uncertainty is derived using in situ measurements of the jet balance in event samples for $Z/\gamma^* +$ jet, $\gamma +$ jet and inclusive jet production. For jets with $p_{\text{T}} > 1$ TeV, the uncertainty is estimated using the calorimeter response to single pions [58]. These, together with the uncertainty in the η-intercalibration mentioned above define the baseline JES uncertainty, which is shown in Fig. 3.6 as a function of the jet p_{T} for both the EM+JES and the LCW+JES calibrations.
Additional uncertainties due to the different response of the calorimeter to gluon and quark jets, and the limited knowledge of the gluon fraction in the jet sample are also included in the total JES uncertainty. For inclusive jet events, the nominal gluon fraction is taken to be the one in the Pythia AUET2B sample, while the uncertainty is taken to be the average difference between the gluon fractions in the Powheg + Pythia and the Herwig++ samples. This uncertainty has

been also derived in top quark pair events in the semileptonic decay mode, with one W boson decaying hadronically while the other one decays to a charged lepton and a neutrino. In this case, the baseline MC sample used for the evaluation of the jet response and the gluon fraction is MC@NLO, while its uncertainty is estimated as the difference between the gluon fractions in the AcerMC and Powheg samples. Figure 3.7 shows the flavour uncertainties derived in both the inclusive jet sample and in the semileptonic $t\bar{t}$ event sample.

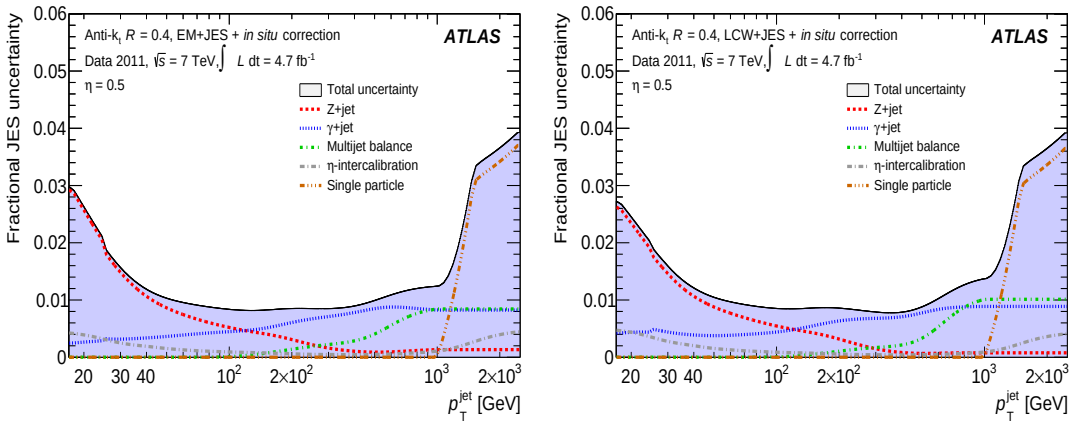

Figure 3.6: The baseline JES uncertainty as a function of the jet p_{T} for EM (left) and LCW (right) jets, derived from in situ methods in jets with $\eta = 0.5$ selected in various physics processes [57]

The uncertainties in both terms of the pileup offset correction described in Sect. 3.4.2.1 are also included in the total JES uncertainty, as well as an additional term due to effects related to nearby jets. This is derived from the comparison of the jet transverse momenta of a calorimeter jet and the matching track-jet for isolated and non-isolated jets, defined as those having a jet with $p_{\mathrm{T}} > 7$ GeV within $\Delta R < 2.5 \times R$, being R the radius parameter used in the anti-k_t clustering. The uncertainty is defined from the differences in the Data and MC distributions on the p_{T} ratios for isolated and non-isolated, cluster-seeded and track-seeded jets (see Ref. [57] for details).
The flavour uncertainties are derived separately for light-quark jets and b-quark jets due to different responses in the calorimeter. Figure 3.8 shows the relative JES uncertainty for b-jets with $\eta = 0.5$ for the EM and LCW calibrations.

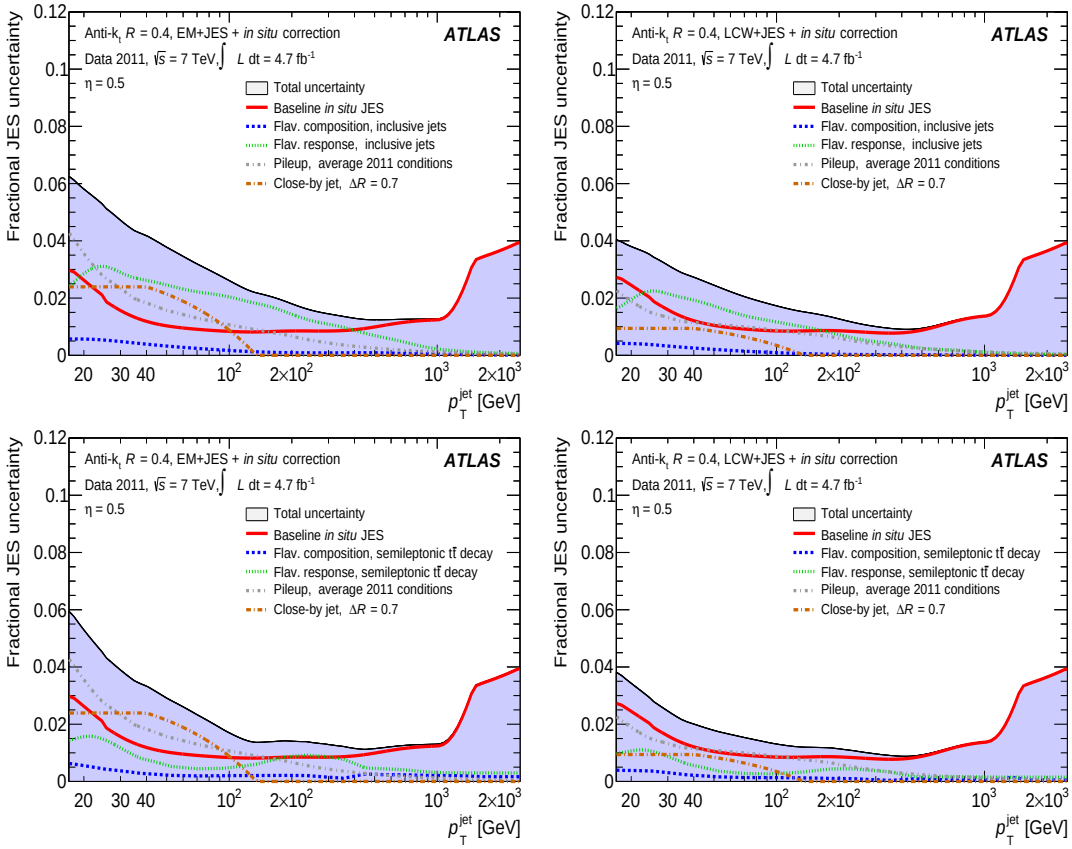

Figure 3.7: The flavour, pileup and close-by JES uncertainties for EM (left) and LCW (right) jets with $\eta = 0.5$ in inclusive jet (top) and $t\bar{t}$ event samples as a function of the jet p_{T} [57]

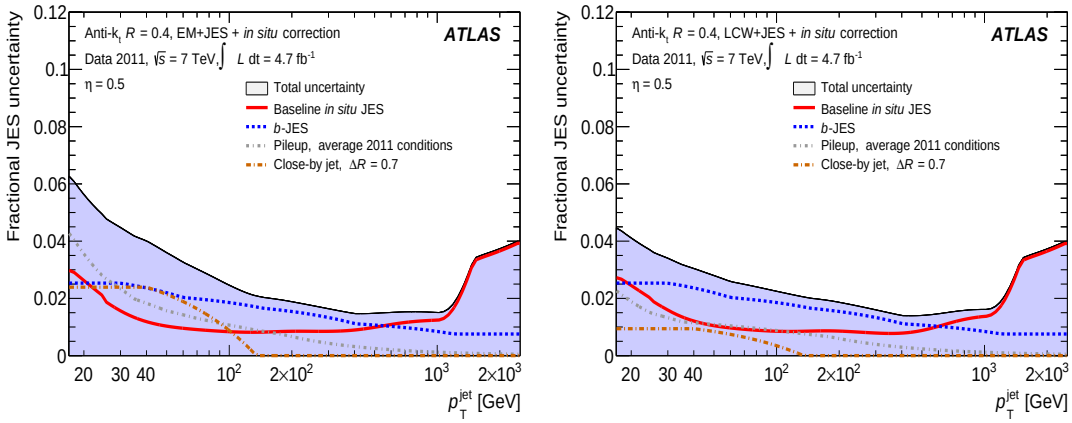

Figure 3.8: The JES uncertainty for EM (left) and LCW (right) b-quark jets with $\eta = 0.5$ as a function of the jet transverse momentum [57]

3.4.4 Jet quality criteria: the good, the bad and the ugly

Jets in ATLAS do not always arise from physics interaction processes, but sometimes they can be reconstructed from detector problems such as calorimeter noise spikes. To avoid the jets arising from undesirable detector effects, a cleaning algorithm has been developed based on the following variables

- f_{EM}: The fraction of energy in the electromagnetic calorimeter.
- f_{max}: Maximum energy fraction in one calorimeter layer.
- f_{HEC}: The fraction of energy in the HEC.
- Q_{LAr}: The fraction of energy corresponding to LAr cells with a cell Q-factor greater than 4000. The cell Q-factor measures the difference between the measured pulse shape ($a_i^{(m)}$) and the predicted pulse shape ($a_i^{(p)}$) that is used to reconstruct the cell energy. It is computed summing over samples as $Q = \sum_i \left(a_i^{(m)} - a_i^{(p)}\right)^2$ and it is stored as 16-bit integer.
- Q_{HEC}: The same as Q_{LAr}, but calculated only for the HEC.
- E_{neg}: The negative energy in the jet.
- t_{jet}: The jet time, computed by averaging the cell time over energy-squared weighted cells, i.e. $t_{\mathrm{jet}} = \frac{\sum_i E_i^2 t_i}{\sum_i E_i^2}$.
- η_{EM}: The pseudorapidity at the EM scale, i.e. before the JES calibration.
- f_{ch}: The jet charged fraction, defined as the sum of the p_{T} for tracks associated to the jet divided by the calibrated jet p_{T}.

Attending to the variables described above, jets are classified into three categories: good, bad and ugly jets.

3.4.4.1 Bad jets

Bad jets are jets not associated to real energy deposits in the calorimeter, but rather to hardware problems such as HEC spikes or noise bursts in the calorimeter; LHC beam conditions or cosmic ray showers. They come in three categories: Loose, Medium and Tight, defined in Table 3.1

	Loose	Medium	Tight
HEC spikes	($f_{\mathrm{HEC}} > 0.5$ and $\lvert Q_{\mathrm{HEC}}\rvert > 0.5$) or $\lvert E_{\mathrm{neg}}\rvert > 60$ GeV	Loose or $f_{\mathrm{HEC}} > 1 - \lvert Q_{\mathrm{HEC}}\rvert$	
Coherent EM noise	$f_{\mathrm{EM}} > 0.95$ and $\lvert Q_{\mathrm{LAr}}\rvert > 0.8$ and $\lvert\eta_{\mathrm{EM}}\rvert < 2.8$	Loose or $f_{\mathrm{EM}} > 0.9$ and $\lvert Q_{\mathrm{LAr}}\rvert > 0.8$ and $\lvert\eta_{\mathrm{EM}}\rvert < 2.8$	Medium or $\lvert Q_{\mathrm{LAr}}\rvert > 0.95$ or ($f_{\mathrm{EM}} > 0.98$ and $\lvert Q_{\mathrm{LAr}}\rvert > 0.05$)
Non-collision background and cosmics	$\lvert t_{\mathrm{jet}}\rvert > 25$ ns or ($f_{\mathrm{EM}} < 0.05$ and $f_{\mathrm{ch}} < 0.05$ and $\lvert\eta_{\mathrm{EM}}\rvert < 2$) or ($f_{\mathrm{EM}} < 0.05$ and $\lvert\eta_{\mathrm{EM}}\rvert \geq 2$) or ($f_{\mathrm{max}} > 0.99$ and $\lvert\eta_{\mathrm{EM}}\rvert < 2$)	Loose or $\lvert t_{\mathrm{jet}}\rvert > 10$ ns or ($f_{\mathrm{EM}} < 0.05$ and $f_{\mathrm{ch}} < 0.1$ and $\lvert\eta_{\mathrm{EM}}\rvert < 2$) or ($f_{\mathrm{EM}} > 0.95$ and $f_{\mathrm{ch}} < 0.05$ and $\lvert\eta_{\mathrm{EM}}\rvert < 2$)	Medium or ($f_{\mathrm{EM}} < 0.1$ and $f_{\mathrm{ch}} < 0.2$ and $\lvert\eta_{\mathrm{EM}}\rvert < 2$) or ($f_{\mathrm{EM}} < 0.1$ and $\lvert\eta_{\mathrm{EM}}\rvert \geq 2$) or ($f_{\mathrm{EM}} > 0.9$ and $f_{\mathrm{ch}} < 0.02$ and $\lvert\eta_{\mathrm{EM}}\rvert < 2$)

Table 3.1: Selection criteria used to reject fake jets arising from HEC spikes, sporadic noise bursts, rare coherent noise and beam or cosmic backgrounds.

3.4.4.2 Ugly jets

Ugly jets correspond to real energy depositions in regions of the calorimeter where the energy reconstruction is not accurate. For example, jets reconstructed in the transition region between the barrel and the end-cap or problematic calorimeter regions would be marked as ugly jets. More precisely, if the fraction of energy in the tile calorimeter gap is greater than 0.5 or the energy fraction corresponding to dead cells or receiving large corrections is greater than 0.5, the jet is marked as ugly.

3.4.4.3 Good jets

Good jets are defined as those jets not marked as bad neither as ugly jets. There are three different definitions of good jets, depending on wether the rejection of bad jets is done using the loose, medium or tight sets of cuts.

3.4.5 The JVF algorithm

In order to reject jets stemming from pileup interactions, an algorithm based on the amount of transverse momentum associated to the primary vertex has been developed. This is called the 'Jet Vertex Fraction' algorithm or JVF and is defined for each jet i and vertex j in a given event. This variable is defined as

$$\mathrm{JVF}(\mathrm{jet_i}, \mathrm{vtx_j}) = \frac{\sum_k p_T(\mathrm{trk_k^{jet_i}}, \mathrm{vtx_j})}{\sum_n \sum_l p_T(\mathrm{trk_l^{jet_i}}, \mathrm{vtx_n})} \tag{3.12}$$

Jets arising from pileup interactions typically have low values of this variable with respect to the primary vertex, while jets produced from the hard scattering process point towards the primary interaction vertex and thus have high values of JVF.

3.5 Identification of jets containing B-hadrons

A crucial goal within ATLAS is the identification of jets containing hadrons composed by at least one b-quark, or simply b-tagging. This is so because of the great

physical importance of the identification of decays such as $H \to b\bar{b}$, where a Higgs boson decays into a b-quark pair, or the top quark decays into one W boson and a b-quark, $t \to Wb$.

The identification of jets containing B-hadrons relies on several properties of these jets. Due to the large lifetime τ of these hadrons, their flight distance with respect to the primary interaction vertex can be resolved in a displaced vertex. Additionally, the average impact parameter of the tracks within one of these jets is greater than that of the tracks within a light jet. Figure 3.9 illustrates these properties. Note that there is a large variety of b-tagging algorithms, each of them with its own particularities. For this reason, this work only covers those which will be later used for the analysis of jet shapes in $t\bar{t}$ events presented in the subsequent chapters. Figure 3.10 shows the light-jet rejection factor versus the efficiency for some of these taggers, together with the efficiency as a function of the jet p_{T}.

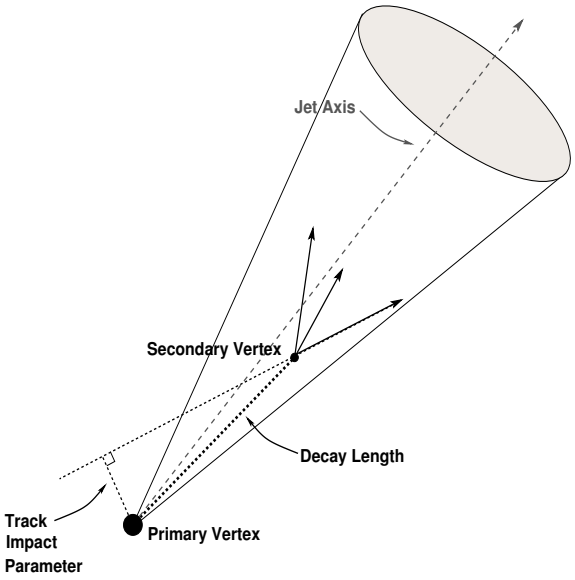

Figure 3.9: Sketch of a jet containing a B-hadron, showing the largest impact parameter of the tracks within the jet as well as the displaced secondary vertex from the B-hadron decay.

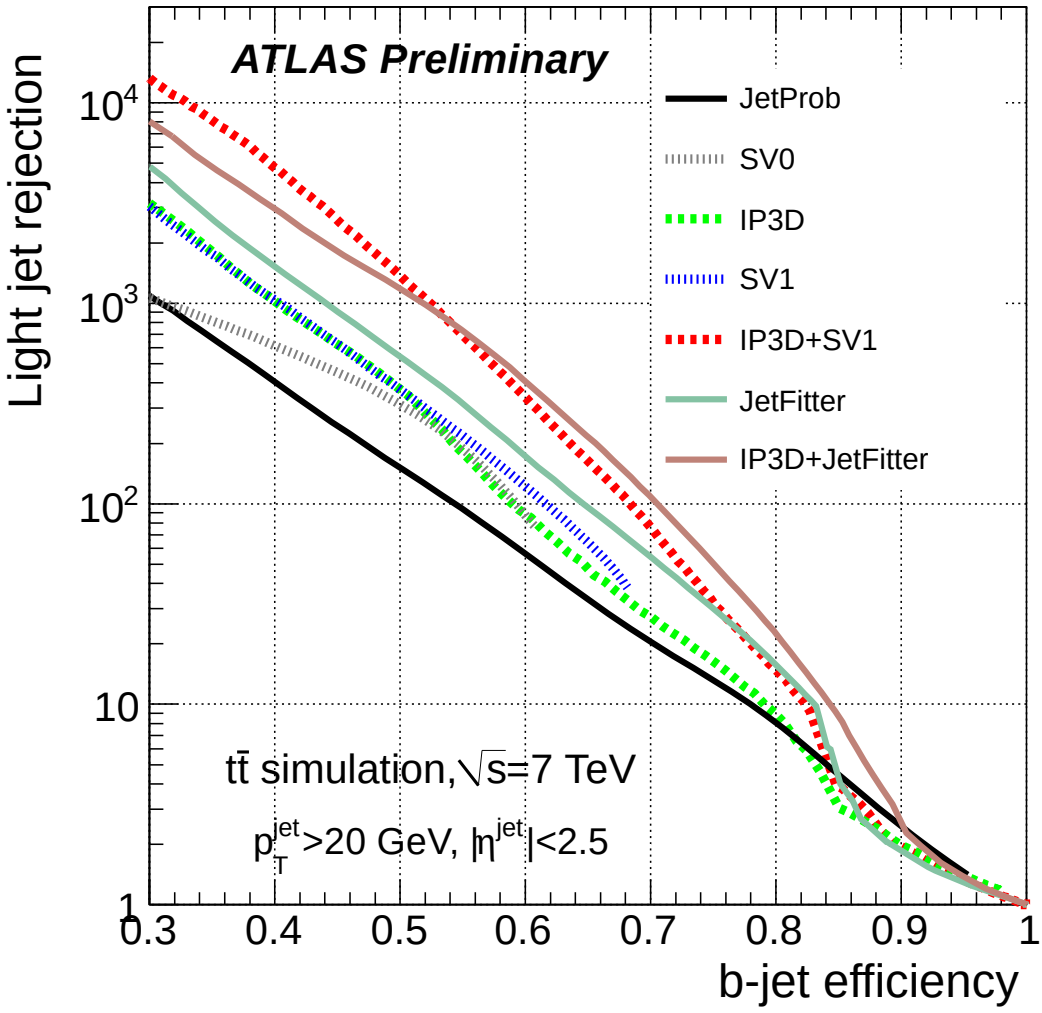

Figure 3.10: Light-jet rejection as a function of the b-tagging efficiency (left) and efficiency versus jet p_{T} (right) for several b-tagging algorithms [59].

3.5.1 Track selection

The properties of the tracks used for b-tagging are descibed in Ref. [59]. This selection is based on the following quality requirements for these tracks:

- At least seven precision hits (pixel or silicon micro-strip hits) on the track.
- At least two hits in the pixel detector, one or more of which must be in the so-called B-layer, the innermost layer of the pixel tracker.
- The transverse momentum of the track must be above 1 GeV.
- The transverse impact parameter has to fulfill $|d_0| < 1$ mm.
- The longitudinal impact parameter has to fulfill $|z_0| \sin\theta < 1.5$ mm.

The tracks selected following the above criteria are then associated to jets depending on their distance ΔR to the axis of a given jet. The association cut is varied depending on the jet p_{T}, because high-p_{T} are more collimated, in order to reduce tracks stemming from the underlying event or pileup interactions which would reduce the discrimination. The cut varies from $\Delta R = 0.45$ at 20 GeV to 0.25 for more energetic jets around 150 GeV. To avoid ambiguities, one track is associated univocally to one jet, the one closest in ΔR.
The transverse and longitudinal impact parameters d_0 and z_0 of the selected tracks are often signed. If the extrapolation of the track crosses the jet axis in front of the primary interaction vertex, its impact parameter is considered as positive, and otherwise negative. The discriminating variable used for b-tagging purposes is the impact parameter significance, defined as the signed impact parameter divided by its error. Figure 3.11 shows the significances for both the transverse and longitudinal impact parameters of tracks associated to jets.

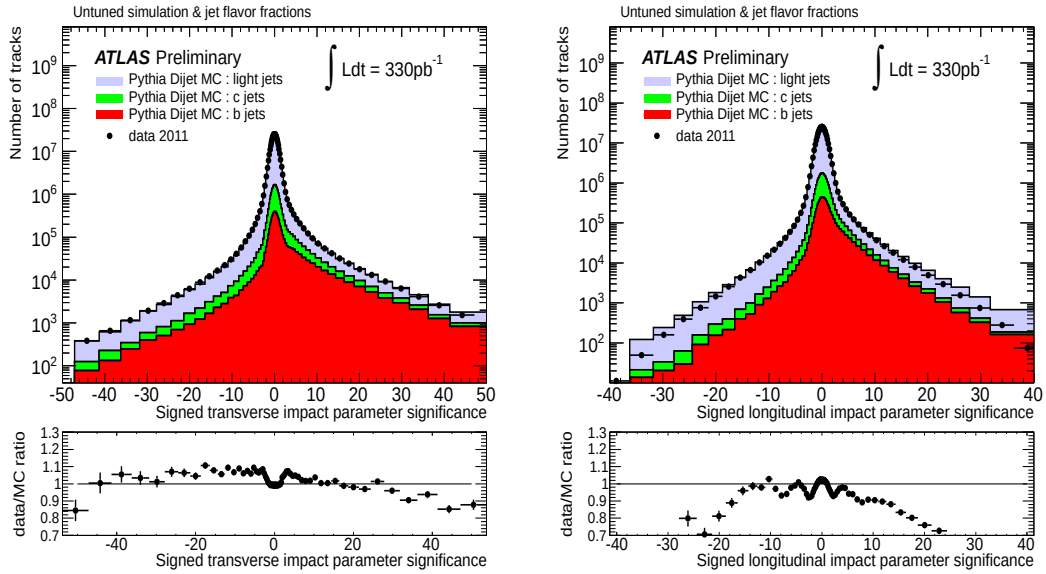

Figure 3.11: Data to MC comparison of the transverse impact parameter significance (left) and the longitudinal impact parameter significance (right) for tracks fulfilling the b-tagging quality criteria and associated to jets. The Monte Carlo predictions are separated in jets arising from the hadronisation of light quarks and gluons (blue), charm quarks (green) and bottom quarks (red) [59].

3.5.2 The IP3D algorithm

The IP3D algorithm uses the impact parameter information in the form of two-dimensional histograms of the transverse significance d_0/σ_{d_0} with respect to the longitudinal significance z_0/σ_{z_0}, taking advantage of the correlations between both variables. These distributions, evaluated for both b and light-quark jets are then smoothed and normalized to be used as inputs for the weight W_{jet} defined as

$$W_{\text{jet}} = \sum_{i=1}^{N_C} \sum_{j=1}^{N_T^i} \log W_{ij} = \sum_{i=1}^{N_C} \sum_{j=1}^{N_T^i} \log \left(\frac{b_i(S_j)}{u_i(S_j)} \right) \tag{3.13}$$

Here, the index i runs over two different categories of tracks depending on wether they have a shared hit with other track or not. For each track, the measured value S_j of the impact parameter significances is compared to the probability densities $b_i(S_j)$ and $u_i(S_j)$ for the track to be associated to a b-jet or to a light

jet, respectively for both categories. Then, the likelihood ratio $b_i(S_j)/u_i(S_j)$ is used to define the per-jet weight in Eq. 3.13, where the index j runs over all tracks associated to a given jet which fall in the i-th quality category. Figure 3.12 shows the resulting weight for the IP3D together with the fraction of jets tagged by the algorithm as a function of the jet p_{T}, for a fixed working point at which the b-tagging efficiency is $\epsilon_b = 60\%$.

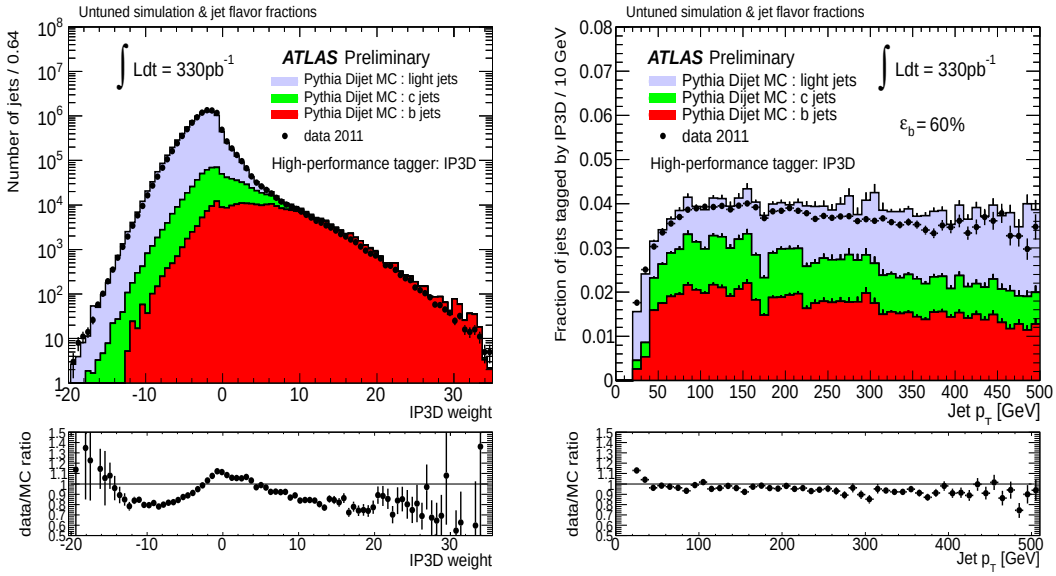

Figure 3.12: Data to MC comparison of the IP3D weight as computed in Eq. 3.13 (left) and the fraction of jets tagged by the 60% efficiency working point of the algorithm as a function of the jet p_{T} (right). The Monte Carlo predictions are separated in jets arising from the hadronisation of light quarks and gluons (blue), charm quarks (green) and bottom quarks (red) [59].

3.5.3 The JetFitter algorithm

A different approach is followed by the JetFitter tagger [60]. This algorithm exploits the fact that hadrons containing b-quarks usually decay via the electroweak interaction to hadrons containing c-quarks rather than to light hadrons, due to the fact that the corresponding CKM matrix elements fulfill $|V_{cb}| \gg |V_{ub}|$. The algorithm makes use of a Kalman filter [47] to reconstruct the full decay chain under the assumption of both secondary vertices arising from the decays of the

B and C hadrons are aligned with respect to the position of the primary vertex. Once the direction of the decay chain is reconstructed by the Kalman filter, the N tracks intersecting the flight direction of the B-hadron are selected, the distances $L_1, \ldots, L_N$ of the intersection points to the primary vertex are determined and the following properties are measured:

- m_{dec}: The invariant mass of the tracks attached to the decay chain.
- x_E: The energy fraction of these tracks with respect to all the charged particles associated to the jet.
- L/σ_L: The flight length significance of the weighted average vertex.

The invariant mass m_{dec} and the energy fraction f_E are shown in Fig. 3.13

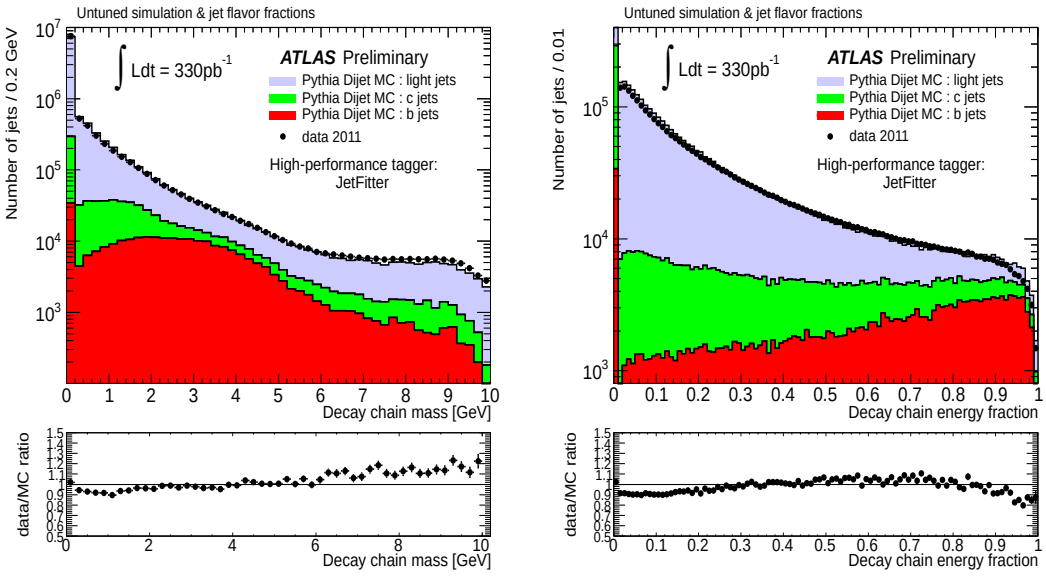

Figure 3.13: Data to MC comparison of the invariant mass of the JetFitter decay chain (left) and of the energy fraction of the tracks intersecting the flight direction of the B-hadron with respect to all charged particles within the jet [59].

The three variables described above are then used to define a per-jet likelihood function of the form

$$W_{\text{jet}} = \sum_{\alpha} C_\alpha \times \mathcal{P}_\alpha(m_{\text{dec}}) \times \mathcal{P}_\alpha(x_E) \times \mathcal{P}_\alpha\left(\frac{L}{\sigma_L}\right) \tag{3.14}$$

where the index α runs over the three possible jet categories (light, charm and bottom). The functions $\mathcal{P}_\alpha(x)$ are the probability density functions of the jet to be classified in these three categories given the variable x, and the coefficient C_α represents how probable it is to find a certain topology for the flavour α. Figure 3.14 shows the resulting tagging weight for JetFitter, together with the fraction of tagged jets as a function of the jet p_{T} for the 60% efficiency working point.

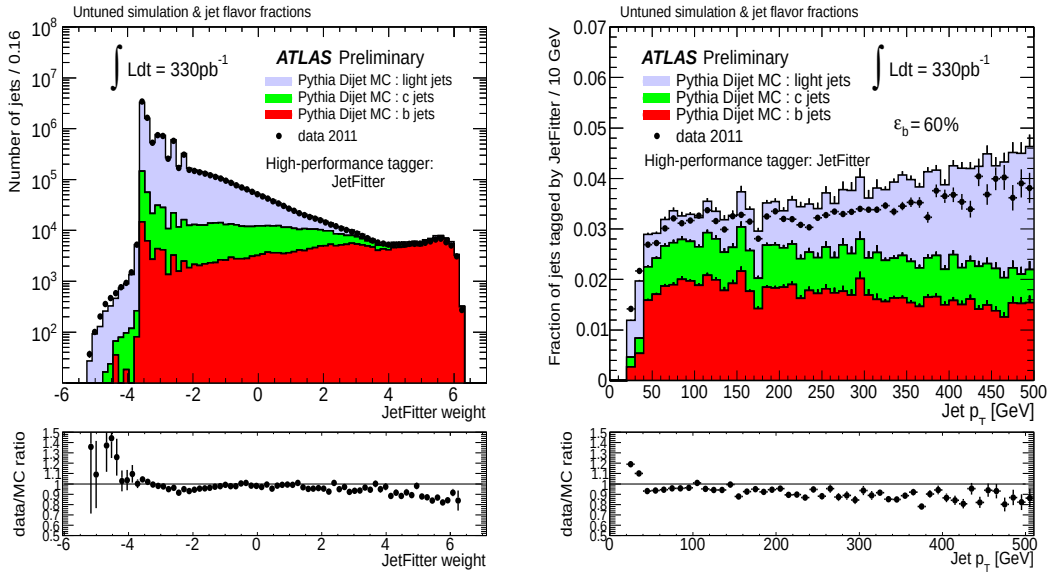

Figure 3.14: Data to MC comparison of the JetFitter weight (left) and the fraction of jets tagged by the 60% efficiency working point of the algorithm as a function of the jet p_{T} (right). The Monte Carlo predictions are separated in jets arising from the hadronisation of light quarks and gluons (blue), charm quarks (green) and bottom quarks (red) [59].

3.5.4 The JetFitterCOMBNN algorithm

The JetFitterCOMBNN algorithm, sometimes called IP3D+JetFitter [59] is based on the combination of both the IP3D and the JetFitter taggers, together with other variables related to the decay chain using an artificial neural network approach trained using Monte Carlo simulated samples for jets of a given flavour. Figure 3.15 presents the tagging weight for this algorithm as well as the tagging rate for the 70% efficiency working point as a function of the jet transverse

momentum.

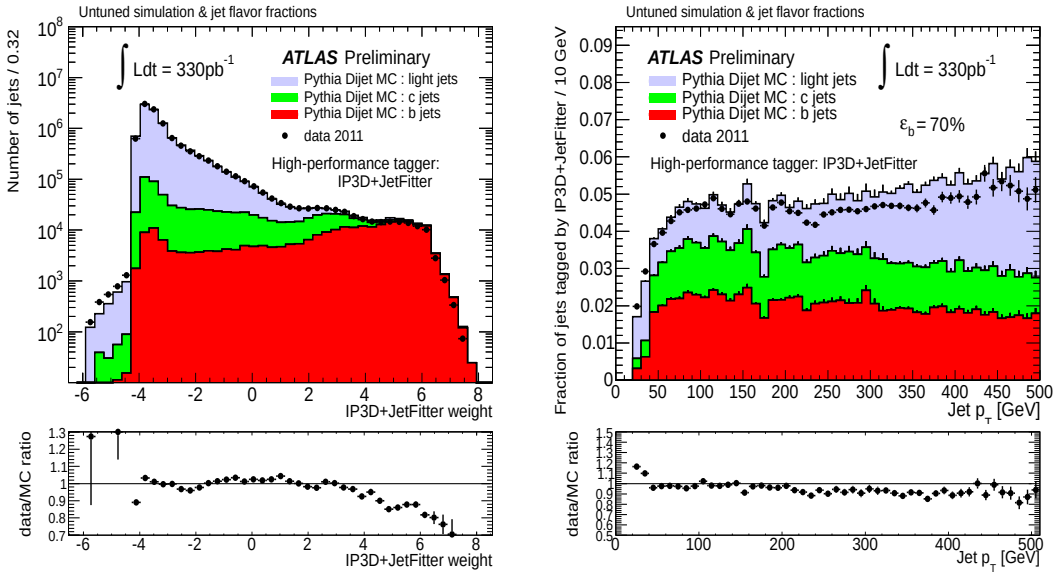

Figure 3.15: Data to MC comparison of the JetFitterCOMBNN weight (left) and the fraction of jets tagged by the 60% efficiency working point of the algorithm as a function of the jet p_{T} (right). The Monte Carlo predictions are separated in jets arising from the hadronisation of light quarks and gluons (blue), charm quarks (green) and bottom quarks (red) [59].

3.6 Missing transverse energy

Some SM particles such as neutrinos or some SUSY-predicted eigenstates such as the neutralino are not expected to leave a physics signature in any of ATLAS subsystems, because they interact very weakly with the matter of the detector. However, their presence in physics events can be deduced from the fact that the total momentum in the transverse plane has to be null due to the momentum conservation. This can be used to define the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ [61] in events where the transverse momentum balance does not compensate. The x and y components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ can be expressed as the sum of two terms, one from the calorimeter and the other from the muon system

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss,calo}} + E_{x(y)}^{\mathrm{miss},\mu} \tag{3.15}$$

The missing transverse energy and its azimuthal angle are then defined as the usual modulus and argument of a 2-dimensional vector

$$E_{\mathrm{T}}^{\mathrm{miss}} = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2} \tag{3.16}$$

$$\varphi^{\mathrm{miss}} = \arctan\left(\frac{E_y^{\mathrm{miss}}}{E_x^{\mathrm{miss}}}\right) \tag{3.17}$$

3.6.1 The calorimeter term

The calorimetric term in Eq. 3.15 is calculated from the cells forming all objects reconstructed in the calorimeter. Therefore, it can be expressed as

$$\begin{aligned} E_{x(y)}^{\mathrm{miss,calo}} = \; & E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss},\gamma} + E_{x(y)}^{\mathrm{miss},\tau} + E_{x(y)}^{\mathrm{miss,jets}} + \\ & E_{x(y)}^{\mathrm{miss,soft}} + E_{x(y)}^{\mathrm{miss,calo},\mu} + E_{x(y)}^{\mathrm{miss,CellOut}} \end{aligned} \tag{3.18}$$

where each term α in Eq. 3.18 is calculated as the negative sum of the x (y) component of the calibrated cell energy inside the corresponding calorimeter objects

$$E_x^{\mathrm{miss},\alpha} = -\sum_{i=1}^{N_{\mathrm{cells}}^{\alpha}} E_i \sin\theta_i \cos\varphi_i; \quad E_y^{\mathrm{miss},\alpha} = -\sum_{i=1}^{N_{\mathrm{cells}}^{\alpha}} E_i \sin\theta_i \sin\varphi_i \tag{3.19}$$

Here, E_i is the cell energy, while θ_i and φ_i represent its polar and azimuthal angles, respectively. All terms in Eq. 3.18 are defined as follows:

- $E_{x(y)}^{\mathrm{miss},e}$, $E_{x(y)}^{\mathrm{miss},\gamma}$ and $E_{x(y)}^{\mathrm{miss},\tau}$ are calculated from cells inside clusters associated to electrons, photons and hadronically decaying τ leptons, respectively.
- $E_{x(y)}^{\mathrm{miss,jets}}$ is calculated from cells inside clusters associated to jets with transverse momentum $p_{\mathrm{T}} > 20$ GeV.
- $E_{x(y)}^{\mathrm{miss,soft}}$ is calculated from cells inside clusters associated to jets with transverse momentum between 7 GeV and 20 GeV.
- $E_{x(y)}^{\mathrm{miss},\mu}$ is calculated from cells inside the cluster left by a muon in the calorimeter (see Sect. 3.6.2).

- $E_{x(y)}^{\text{miss,CellOut}}$ is calculated from cells inside topological clusters which are not associated to any other object.

3.6.2 The muon term

The calculation of the muon term in Eq. 3.15 is performed using the momenta of muons within $|\eta| < 2.7$ in the following way

$$E_{x(y)}^{\text{miss},\mu} = -\sum_{i=1}^{N_\mu} p_{x(y)}^{\mu_i} \tag{3.20}$$

where the sum runs over well reconstructed muons. In the region $|\eta| < 2.5$, this means that the spectrometer track has to be matched with an inner detector track (combined muons, see Sect. 3.3.2). This drastically reduces the number of fake muons (reconstructed muons not matching to a truth muon). The calorimeter muon term in Eq. 3.18 is defined differently depending on wether the muon is isolated or not, attending to the distance ΔR between the muon and a jet. If $\Delta R < 0.3$, the muon is considered isolated. Otherwise, the muon is considered as non-isolated.

- For isolated muons, the transverse momentum is determined by the combination of the two measurements by the muon spectrometer and the inner detector. In this case, the energy lost by the muon in the calorimeter is not added to the calorimetric term in Eq. 3.18, to avoid double counting in the energy.

- For non-isolated muons, the energy lost in the calorimeter cannot be distinguished from the energy deposits from the particles inside the jet. Therefore, the muon spectrometer track is used and the calorimeter term is added to Eq. 3.18 [62]

For the region outside the inner detector range, $2.5 < |\eta| < 2.7$, the muon spectrometer track is used and the calorimeter term is added to Eq. 3.18 accordingly.

3.6.3 Performance in $W \to \ell\nu$ events

In Ref. [61], the performance of this $E_{\mathrm{T}}^{\mathrm{miss}}$ algorithm was investigated in terms of Data / MC agreement in $W \to \ell\nu$ events with one W boson decaying into a charged lepton and a neutrino in the 2010 data sample. Figures 3.16 to 3.18 shows the results for each of the involved terms in the missing transverse energy definition. The Monte Carlo expectations have been superimposed and normalised so they have the same area as the histogram for data, after the expectation for each of the signal and background samples is weighted according to its corresponding cross section.

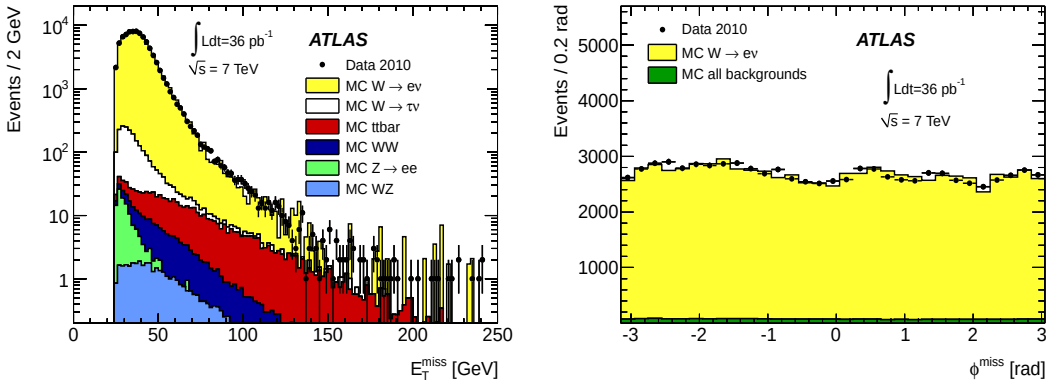

Figure 3.16: The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ (left) and its azimuthal angle φ^{miss} (right) as evaluated in $W \to e\nu$ events [61].

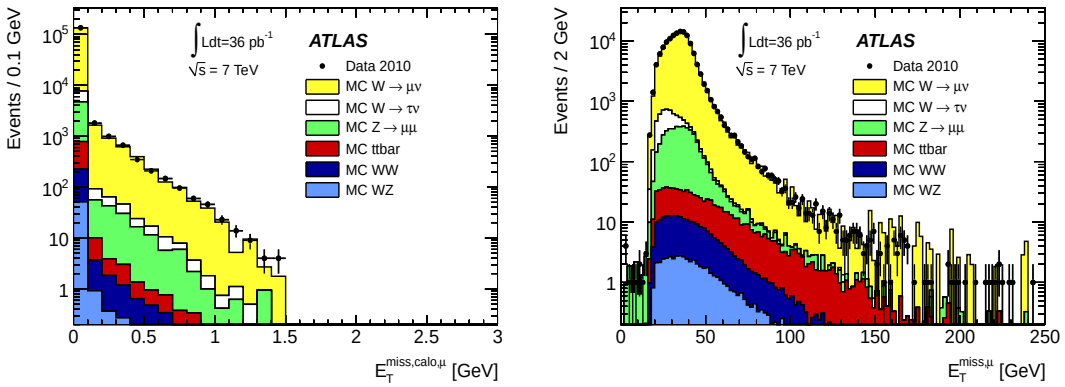

Figure 3.17: Muon terms of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in $W \to \mu\nu$ events. Energy left in the calorimeter by the muon (left) and the reconstructed muon term (right) [61].

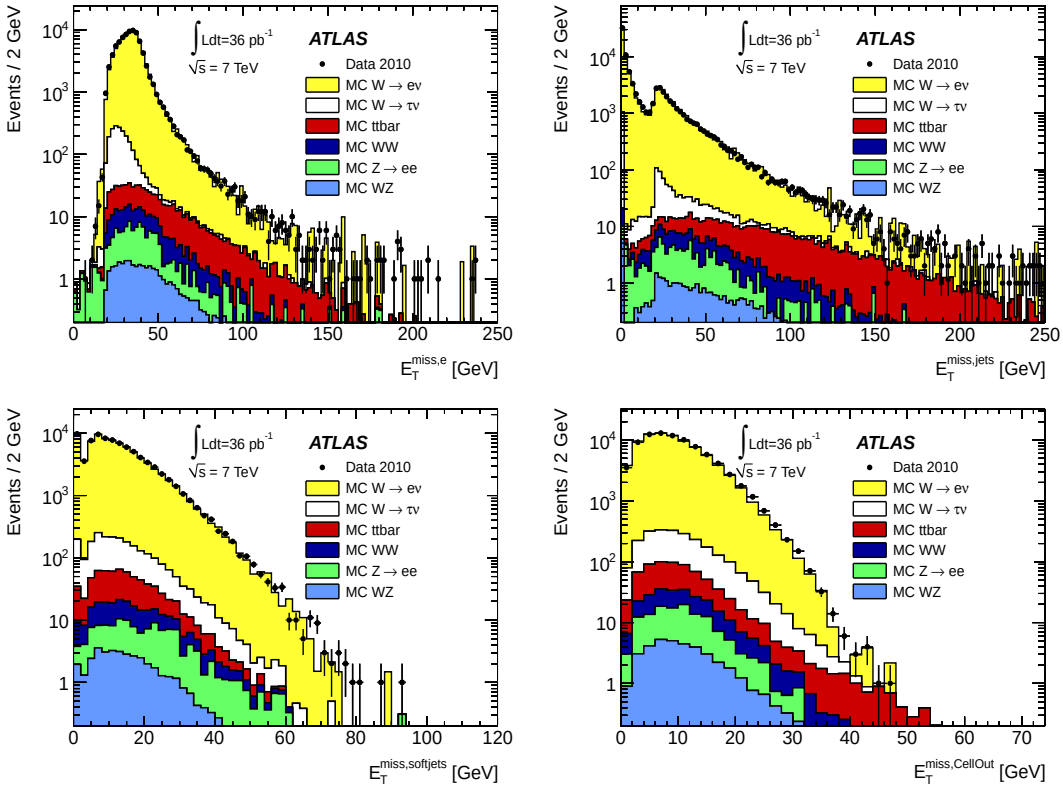

Figure 3.18: Performance in $W \rightarrow e\nu$ events of some of the calorimeter terms involved in the definition of the missing transverse energy, calculated from cells inside electron clusters (top left), jet clusters (top right), soft jet clusters (bottom left) and clusters not associated to any object [61].

Chapter 4

Jet shapes in $t\bar{t}$ events

In this chapter, a measurement of the jet internal structure in top-quark pair events is presented in terms of the density of transverse energy inside the jet as a function of the distance to the jet axis, i.e. the so-called jet shapes. A comparison is made between the jet shapes of jets originating from a b-quark and those originating from the lighter u, d, c and s-quarks. The results obtained in this chapter have been published by the ATLAS Collaboration in Ref. [96]

4.1 Introduction and motivation

Quarks and gluons generated in large momentum transfer interactions evolve into hadron jets in a two step process. The first one is of a perturbative nature and gives rise to a parton shower, while the second one is non-perturbative and is responsible for the hadronisation. The internal structure of a jet is expected to depend primarily on the type of parton it originated from and, to a lesser extent, on the fragmentation. In particular, due to the different colour factors in gg and qg QCD vertices, gluon jets are expected to be broader than quark jets. Also, due to the dead cone effect described in section 1.7, jets induced by the hadronization of a b-quark are expected to be broader than jets originating from lighter quarks. For jets defined using of cone algorithms, jet shapes (see Sect. 1.6) have been traditionally [63, 64, 65, 66, 67, 68, 20] used as a means of understanding the metamorphosis of partons into hadrons in e^+e^-, ep and hadron-hadron colliders.

Experimentally one finds that jets in e^+e^- and ep are narrower than those observed in pp collisions and this is understood as a result of the different admixtures of quark and gluon jets present in these different types of interactions. Furthermore, at high energies, where fragmentation effects become negligible, jet shapes have been found to be in qualitative agreement with NLO QCD predictions and in quantitative agreement with those including leading log corrections. Jet shapes have been also proposed for top-tagging purposes [69] and also as a means for new physics searches in final states with highly boosted particles [71, 72, 73].

The expectation of b-quark jets being broader than light-quark jets is supported by observations by the CDF collaboration [70], where a comparison is presented between jet shapes in a b-enriched sample with a purity between 20% and 30% and an all inclusive sample where no distinction is made about its flavour or color composition. We would like to remark that the b-jets selected in this sample come from gluon splitting, $g \to b\bar{b}$ and populate a region in p_{T} above 80 GeV, where the b-tagger is more efficient.

The approach taken in this analysis is different, since here $t\bar{t}$ final states are considered. This final state provides a major source of b-jets, as the top quark decays dominantly via $t \to Wb$. While the dileptonic sample is a clean and copious source of b-jets, the semileptonic one contains both types of b and light jets, the latter ones being constrained by the W mass.

A comparison of the light and b-jet shapes measured in the $t\bar{t}$ semileptonic decays, is interesting per se in as much improves the CDF measurement to which we referred above. Also, these data can have several other practical applications. One of them is the determination of the b-quark mass which will be presented in the next chapters. The analysis of this particular final state can also provide an additional handle with which to improve the s and b-jet separation for a future determination of the CKM matrix element $|V_{ts}|$ [74]. On the top of that, this will help in tuning presently available MC for top quark pair production in a new kinematic regime. Thus, jet shape measurements will deepen our understanding of the $t\bar{t}$ hadronic final states. This is the aim of this study.

4.2 Monte Carlo samples

For this analysis, two main Monte Carlo samples are used. They implement different hadronization and parton showering models. The generated events are then processed with the ATLAS full detector simulation, based on the GEANT4 software [75]. The baseline signal MC samples used here are generated with MC@NLO [26] or POWHEG [27] for the matrix element calculation, and the parton shower and hadronization processes are generated with HERWIG [24], using the parton distribution functions CTEQ6.6 [76]. Multiparton interactions have been taken into account according to reference [25].

Additional MC samples are used to check the dependence of the jet shapes on the hadronization model. These are generated using POWHEG + PYTHIA [22, 27] convoluted with the MRST2007LO* PDFs [77]. In addition, the ACERMC [78] program interfaced to PYTHIA with the PERUGIA 2010 tune [23] is also used. Table 4.1 summarizes the properties of these signal samples.

For the electroweak background studies, we have considered single and double vector boson production (W, Z, WW, WZ, ZZ) simulated in the ALPGEN samples [28] with the CTEQ6L PDFs [79]. The samples in which the gauge bosons are produced in association with heavy flavour jets are also considered, removing the overlaps between the heavy flavour samples and the standard ones. In addition, the single top s and t-channel processes have been simulated using MC@NLO. See tables 4.2 to 4.5 for a summary of the cross sections, filter efficiencies and numbers of events.

Channel number	Sample	$\sigma \cdot \epsilon_f$ (nb)	Number of events
106206	TTbar_McAtNlo_Jimmy_175GeV	0.74×10^{-1}	1305569
105860	TTbar_PowHeg_Jimmy	5.48×10^{-1}	2997878
105861	TTbar_PowHeg_Pythia	0.79×10^{-1}	2994490
117240	AcerMCttbar_Perugia2010	4.19×10^{-2}	1998585

Table 4.1: Monte Carlo signal samples used for the analysis

Channel number	Sample	$\sigma \cdot \epsilon_f$ (nb)	Number of events
107680	AlpgenJimmyWenuNp0_pt20	6.92	3455037
107681	AlpgenJimmyWenuNp1_pt20	1.31	2499513
107681	AlpgenJimmyWenuNp2_pt20	3.78×10^{-1}	3768265
107682	AlpgenJimmyWenuNp3_pt20	1.02×10^{-1}	1009641
107683	AlpgenJimmyWenuNp4_pt20	2.57×10^{-2}	249869
107684	AlpgenJimmyWenuNp5_pt20	7.02×10^{-3}	69953
107685	AlpgenJimmyWmunuNp0_pt20	6.92	3466523
107690	AlpgenJimmyWmunuNp1_pt20	1.30	641867
107691	AlpgenJimmyWmunuNp2_pt20	3.78×10^{-1}	3768893
107692	AlpgenJimmyWmunuNp3_pt20	1.02×10^{-1}	1009589
107693	AlpgenJimmyWmunuNp4_pt20	2.56×10^{-2}	254879
107694	AlpgenJimmyWmunuNp5_pt20	6.99×10^{-3}	69958
107695	AlpgenJimmyWtaunuNp0_pt20	6.92	3416438
107700	AlpgenJimmyWtaunuNp1_pt20	1.30	641809
107701	AlpgenJimmyWtaunuNp2_pt20	3.78×10^{-1}	3768750
107702	AlpgenJimmyWtaunuNp3_pt20	1.02×10^{-1}	1009548
107703	AlpgenJimmyWtaunuNp4_pt20	2.57×10^{-2}	249853
107704	AlpgenJimmyWtaunuNp5_pt20	6.99×10^{-3}	63692
107280	AlpgenJimmyWbbFullNp0_pt20	4.73×10^{-2}	474933
107281	AlpgenJimmyWbbFullNp1_pt20	3.58×10^{-2}	204933
107282	AlpgenJimmyWbbFullNp2_pt20	1.74×10^{-2}	174942
107283	AlpgenJimmyWbbFullNp3_pt20	7.61×10^{-3}	69969
117293	AlpgenWcNp0_pt20	6.44×10^{-1}	6483825
117294	AlpgenWcNp1_pt20	2.05×10^{-1}	2069456
117295	AlpgenWcNp2_pt20	5.09×10^{-2}	517833
117296	AlpgenWcNp3_pt20	1.14×10^{-2}	114936
117297	AlpgenWcNp4_pt20	2.77×10^{-3}	29977
117284	AlpgenWccFullNp0_pt20	1.28×10^{-1}	1274737
117285	AlpgenWccFullNp1_pt20	1.05×10^{-1}	1049726
117286	AlpgenWccFullNp2_pt20	5.22×10^{-2}	524808
117287	AlpgenWccFullNp3_pt20	1.70×10^{-2}	33984

Table 4.2: Monte Carlo W + jets samples used for the analysis, including heavy flavours

Channel number	Sample	$\sigma \cdot \epsilon_f$ (nb)	Number of events
117910	st_tchan_enu_McAtNlo_Jimmy_175GeV	6.90×10^{-3}	89970
117911	st_tchan_munu_McAtNlo_Jimmy_175GeV	6.88×10^{-3}	89970
117912	st_tchan_taunu_McAtNlo_Jimmy_175GeV	6.96×10^{-3}	89973
117913	st_schan_enu_McAtNlo_Jimmy_175GeV	4.46×10^{-4}	10000
117914	st_schan_munu_McAtNlo_Jimmy_175GeV	4.44×10^{-4}	10000
117915	st_schan_taunu_McAtNlo_Jimmy_175GeV	4.43×10^{-4}	10000
117916	st_Wt_McAtNlo_Jimmy_175GeV	1.39×10^{-2}	184876

Table 4.3: Monte Carlo single top samples used for the analysis, including Wt, t and s-channels

Channel number	Sample	$\sigma \cdot \epsilon_f$ (nb)	Number of events
107650	AlpgenJimmyZeeNp0_pt20	6.70×10^{-1}	6612265
107651	AlpgenJimmyZeeNp1_pt20	1.34×10^{-1}	1333745
107652	AlpgenJimmyZeeNp2_pt20	4.07×10^{-2}	404873
107653	AlpgenJimmyZeeNp3_pt20	1.13×10^{-2}	109942
107654	AlpgenJimmyZeeNp4_pt20	2.86×10^{-3}	29992
107655	AlpgenJimmyZeeNp5_pt20	7.59×10^{-4}	8992
107660	AlpgenJimmyZmumuNp0_pt20	6.70×10^{-1}	6619010
107661	AlpgenJimmyZmumuNp1_pt20	1.35×10^{-1}	1334723
107662	AlpgenJimmyZmumuNp2_pt20	4.07×10^{2}	403886
107663	AlpgenJimmyZmumuNp3_pt20	1.12×10^{-2}	109954
107664	AlpgenJimmyZmumuNp4_pt20	2.85×10^{-3}	29978
107665	AlpgenJimmyZmumuNp5_pt20	7.63×10^{-4}	9993
107670	AlpgenJimmyZtautauNp0_pt20	6.70×10^{-1}	6618801
107671	AlpgenJimmyZtautauNp1_pt20	1.35×10^{-1}	1334664
107672	AlpgenJimmyZtautauNp2_pt20	4.08×10^{-2}	404853
107673	AlpgenJimmyZtautauNp3_pt20	1.13×10^{-2}	109944
107674	AlpgenJimmyZtautauNp4_pt20	2.84×10^{-3}	29982
107675	AlpgenJimmyZtautauNp5_pt20	7.61×10^{-4}	9993
109300	AlpgenJimmyZeebbNp0_nofilter	6.57×10^{-3}	149971
109301	AlpgenJimmyZeebbNp1_nofilter	2.48×10^{-3}	99977
109302	AlpgenJimmyZeebbNp2_nofilter	8.85×10^{-4}	38985
109303	AlpgenJimmyZeebbNp3_nofilter	3.92×10^{-4}	9990
109305	AlpgenJimmyZmumubbNp0_nofilter	6.56×10^{-3}	149971
109306	AlpgenJimmyZmumubbNp1_nofilter	2.47×10^{-3}	99967
109307	AlpgenJimmyZmumubbNp2_nofilter	8.87×10^{-4}	39980
109308	AlpgenJimmyZmumubbNp3_nofilter	3.90×10^{-4}	9994
109310	AlpgenJimmyZtautaubbNp0_nofilter	6.57×10^{-3}	149968
109311	AlpgenJimmyZtautaubbNp1_nofilter	2.49×10^{-3}	98960
109312	AlpgenJimmyZtautaubbNp2_nofilter	8.93×10^{-4}	39978
109313	AlpgenJimmyZtautaubbNp3_nofilter	3.89×10^{-4}	9995

Table 4.4: Monte Carlo Z + jets samples used for the analysis, including heavy flavours

Channel number	Sample	$\sigma \cdot \epsilon_f$ (nb)	Number of events
107100	AlpgenJimmyWWlnulnuNp0	2.10×10^{-3}	49992
107101	AlpgenJimmyWWlnulnuNp1	9.96×10^{-4}	24997
107102	AlpgenJimmyWWlnulnuNp2	4.55×10^{-4}	14996
107103	AlpgenJimmyWWlnulnuNp3	1.76×10^{-4}	9995
107104	AlpgenJimmyWZincllNp0	6.72×10^{-4}	14994
107105	AlpgenJimmyWZincllNp1	4.14×10^{-4}	9998
107106	AlpgenJimmyWZincllNp2	2.25×10^{-4}	4999
107107	AlpgenJimmyWZincllNp3	9.50×10^{-5}	4997
107108	AlpgenJimmyZZincllNp0	5.10×10^{-4}	39989
107109	AlpgenJimmyZZincllNp1	2.31×10^{-4}	19989
107110	AlpgenJimmyZZincllNp2	8.70×10^{-5}	19984
107111	AlpgenJimmyZZincllNp3	3.91×10^{-5}	9995

Table 4.5: Monte Carlo diboson samples used for the analysis

4.3 Event selection

For this study, two samples of $t\bar{t}$ events are selected. A dileptonic one, with both W bosons decaying into leptons (e,μ), and a semileptonic one, where one W decays into leptons and the other one into a $q\bar{q}'$ pair, forming two jets (see figure 4.1). The selection criteria follow those in [80] for the dileptonic sample and [81] for the semileptonic one. The dataset used for the analysis corresponds to 2011 ATLAS data, with a center of mass energy $\sqrt{s} = 7$ TeV and has an integrated luminosity of 1.8 fb^{-1}(up to period J). Release 16 recommendations have been used for the analysis, and the top GoodRunsList has been used to filter the LumiBlocks with detector or data integrity errors.

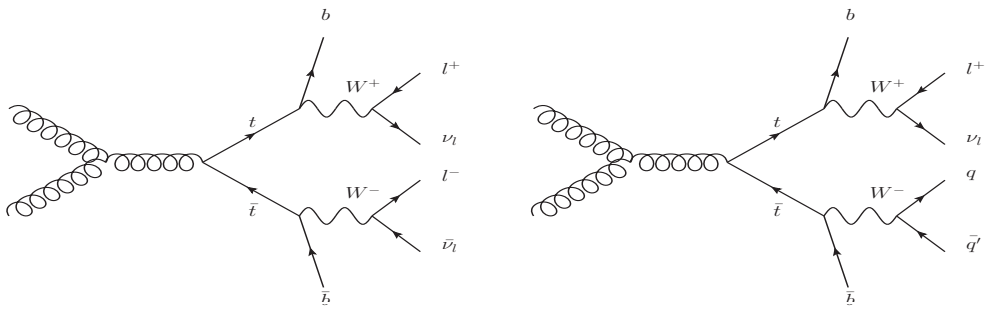

Figure 4.1: LO Feynman diagrams for $gg \to t\bar{t}$ in the dileptonic (left) and semileptonic (right) decay modes

4.3.1 Dileptonic sample

The dileptonic sample is selected as follows. First of all, the event is required to be triggered by the HLT with the streams `EF_e20_medium` (e-channel) or `EF_mu18` (μ-channel), and then the event is required to have two isolated leptons and missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ from the leptonic W boson decays to an electron (muon) and a neutrino.

To reject the non-collision background, the primary vertex is required to have at least four tracks originating from it. The offline selection requires two isolated leptons (e,μ) with $E_T(e) > 25$ GeV and $p_{\mathrm{T}}(\mu) > 20$ GeV), where $E_T = E_{\mathrm{cluster}} \sin(\theta_{\mathrm{track}})$. For electrons, clusters in the region $|\eta| < 2.47$ are selected, avoiding the crack $1.37 < |\eta| < 1.52$, and the dead region in the calorimeter due to the LAr hardware problem affecting the 2011 data. This is done by cutting off all electrons within this region. For muons, a certain number of hits in the inner detector are required to have a track matched to the muon spectrometer, which has to lie in the pseudorapidity range $|\eta| < 2.5$. At least one of the selected leptons has to match the corresponding trigger object.

The jet selection is done by requiring at least 2 jets in the event, selected with the anti-k_t algorithm [32] with a radius $R = 0.4$. Their transverse momenta have to be greater than 25 GeV and their pseudorapidity such that $|\eta| < 2.5$. In addition, at least one of the selected jets has to be tagged with the JetFitterCombNN algorithm (see Sect. 3.5.4), with a cut placed in 2.4, leading to a b-tagging efficiency of 57%.

Jets within the dead LAr region are removed, and the event is rejected if there is a jet identified as calorimeter noise or out-of-time signal (`LooseBad` jets, see Sect. 3.4.4) with $p_{\mathrm{T}} > 20$ GeV. Jets overlapping with a selected electron are removed if they are closer than $\Delta R = 0.2$, while if a jet is closer than $\Delta R = 0.4$ to a muon, the muon is removed. For the missing transverse energy, a cut on $E_{\mathrm{T}}^{\mathrm{miss}} > 60$ GeV is imposed for the ee and $\mu\mu$ channels. For the $e\mu$ channel, H_T is required to be greater than 130 GeV, where H_T is the scalar sum of p_{T} of all muons and jets plus the scalar sum of the E_T of all electrons in the case of the electron channels. To reject the J/ψ and $Z \to \ell\ell$ background, the lepton pair is required to have an invariant mass m_{ee}, $m_{\mu\mu}$ greater than 15 GeV, and to lie outside of the Z mass

window, rejecting all events where the two lepton invariant mass (for the ee and $\mu\mu$ channels) satisfies

$$|m_{\ell\ell} - m_Z| < 10 \text{ GeV}; \quad m_Z = 91 \text{ GeV} \tag{4.1}$$

4.3.1.1 Sample composition

The selected sample does not only contain dileptonic top events, but a set of other background processes as $W+$ jets or $Z+$ jets, with the electroweak gauge bosons decaying into leptons. To estimate the amount of these contaminations, Monte Carlo datasets of all these processes are used, and the above event selection is applied on them. The results are summarized in the following table, where 'Other EW' gives account for the W + jets and diboson (WW, WZ and ZZ) contributions.

Process	Expected number of events	Percentage
$t\bar{t}$ dileptonic	2098	94.9%
$Z \to \ell^+\ell^-$ + jets	9	0.4%
$Z \to \ell^+\ell^- + b\bar{b}$	5	0.2%
Other EW (W, WW, WZ, ZZ)	4	0.2%
Single Top	95	4.3%
Total Expected	2211	100%
Total Observed	2067	

Table 4.6: The expected composition of the $t\bar{t}$ (dileptonic) sample

The dileptonic $t\bar{t}$ events can be divided in six main groups according to the leptons arising from the W decays, namely ee, $e\mu$, $e\tau$, $\mu\mu$, $\mu\tau$, $\tau\tau$. The τ channels represent a non negligible fraction of the total number of events, due to the decays $\tau^+ \to \ell^+\nu_l\bar{\nu_\tau}$. Table 4.7 summarizes the situation for the background subtracted sample.

Process	Expected number of events	Percentage
$t\bar{t} \to e^+e^-$	198	9.5%
$t\bar{t} \to e^\pm\mu^\mp$	1307	62.3%
$t\bar{t} \to \mu^+\mu^-$	365	17.4%
$t\bar{t} \to e^\pm + (\tau^\mp \to \ell^\mp)$	108	5.1%
$t\bar{t} \to \mu^\pm + (\tau^\mp \to \ell^\mp)$	112	5.3%
$t\bar{t} \to \tau^+\tau^- \to \ell^+\ell^-$	8	0.4%
Total	2098	100%

Table 4.7: The expected classification of the dileptonic sample

It can be noticed that these expected yields do not fit the expectations from the branching fractions $\mathcal{B}(W^\pm \to \ell^\pm\nu) \simeq 11\%$ alone. However, they respond to the facts that the reconstruction efficiency is higher for muons than for electrons and that the selection cuts are also different for the $e\mu$ channel than for the others, not requiring the lepton pair to lay outside the Z mass range and not cutting on $E_{\mathrm{T}}^{\mathrm{miss}}$, thus having a bigger acceptance for this particular channel.

4.3.2 Semileptonic sample

The selection criteria for the semileptonic sample is as follows. As above, the event is required to be triggered by the corresponding EF streams for electrons and muons. The offline object selection is done exactly in the same way, but now, the event is required to have only one isolated lepton with $E_T > 25$ GeV for electrons and $p_{\mathrm{T}} > 20$ GeV for muons. To account for the neutrino in the leptonic W decay, a cut is imposed on the missing transverse energy of $E_{\mathrm{T}}^{\mathrm{miss}} > 35$ GeV in the electron channel and $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV in the muon channel. Moreover, if the W transverse mass is defined as

$$m_T^W = \sqrt{2p_{\mathrm{T}}^l E_{\mathrm{T}}^{\mathrm{miss}}(1 - \cos\Delta\varphi_{\ell\nu})} \tag{4.2}$$

cuts of $m_T^W > 25$ GeV are imposed in the e-channel and $E_{\mathrm{T}}^{\mathrm{miss}} + m_T^W > 60$ GeV in the μ-channel, where $\Delta\varphi_{\ell\nu}$ is the angle in the transverse plane between the selected lepton and the neutrino direction, which is obtained from the components E_x^{miss} and E_y^{miss} of the missing transverse energy.

The jet selection is done in this case by requiring at least 4 selected jets ($p_{\mathrm{T}} >$ 25 GeV and $|\eta| < 2.5$) in the final state, and at least one of them has to be tagged as a b-jet using the JetFitterCombNN algorithm.

4.3.2.1 Sample composition

The main background contributions for the semileptonic channel are summarized below. In this case, 'Other EW' includes Z + jets and diboson processes. All backgrounds, with the exception of QCD, have been estimated using MC samples. The multijet background has been estimated using the jet-electron method, which quantifies the probability of a signal lepton (electron or muon) being faked by a hadronic jet. This is done by selecting jets with a high fraction of electromagnetic energy, which are then treated as leptons [86].

Process	Expected number of events	Percentage
$t\bar{t}$ semileptonic	13963	77.4%
$W^{\pm} \to \ell^{\pm}\nu_l$ + jets	559	3.1%
$W^{\pm} \to \ell^{\pm}\nu_l + b\bar{b}$	595	3.3%
$W^{\pm} \to \ell^{\pm}\nu_l + c\bar{c}$	794	4.4%
$W^{\pm} \to \ell^{\pm}\nu_l + c(\bar{c})$	361	2.0%
QCD multijet	902	5.0%
Other EW (Z, WW, WZ, ZZ)	198	1.1%
Single Top	668	3.7%
Total Expected	18040	100%
Total Observed	17019	

Table 4.8: The expected composition of the $t\bar{t}$ (semileptonic) sample

As in the previous case, the signal events can be divided into three main groups according to the main lepton family including the leptonic τ decays. Table 4.9 summarizes the situation.

Process	Expected number of events	Percentage
$t\bar{t} \to e^{\pm}$ + jets	6074	43.5%
$t\bar{t} \to \mu^{\pm}$ + jets	6283	45.0%
$t\bar{t} \to \tau^{\pm}(\to \ell^{\pm})$ + jets	1606	11.5%
Total	13963	100%

Table 4.9: The expected classification of the semileptonic sample

The transverse momentum of the corresponding charged leptons, electrons and muons, is shown in Fig. 4.2, compared to the MC expectations.

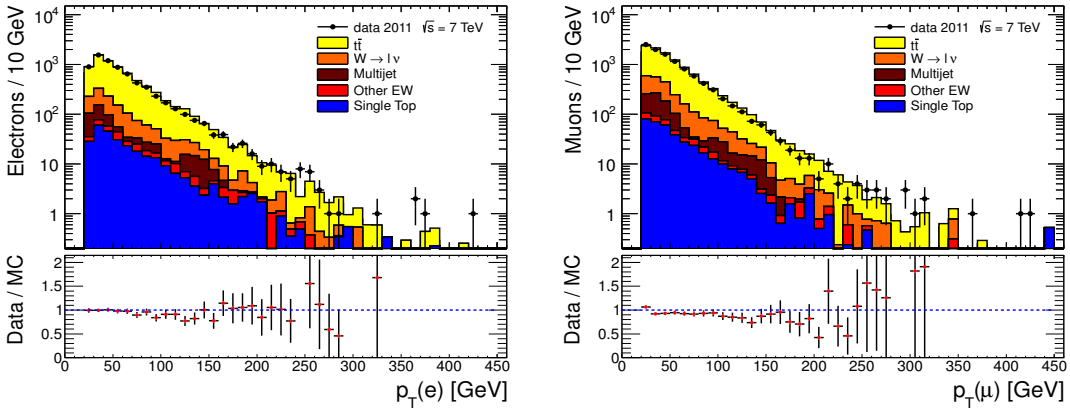

Figure 4.2: The p_{T} distributions for electrons and muons in the semileptonic channel along with MC expectations

4.4 Jet selection

As discussed in previous sections, jets have been defined using the anti-k_t algorithm [32] with radius parameter $R = 0.4$, see [82] for more details. The input objects to the jet algorithm, both for data and detector level simulation, are topological clusters in the calorimeter (see Sect. 3.4.1). These clusters are seeded by calorimeter cells with $|E_{cell}| > 4\sigma$, with σ the RMS of the noise. Neighbouring cells are added and clusters are formed following an iterative procedure.
Jets are calibrated using the EM+JES calibration procedure (see Sect. 3.4.2 and Refs. [56, 57]). This calibration procedure is different for b- and light jets in as much they have different particle composition. See [82] and [84] for more

details and for a discussion of the associated uncertainties.
In order to avoid pileup contamination in the jet samples, the jet vertex fraction, defined as the fraction of tracks in the jet coming from the primary vertex, is required to be $|\mathrm{JVF}| > 0.75$ (see Sect. 3.4.5). This makes the average jet multiplicity independent of the number of vertices. The jet multiplicity obtained after the JVF cleaning cut is shown in figure 4.3 for the semileptonic and dileptonic channels.

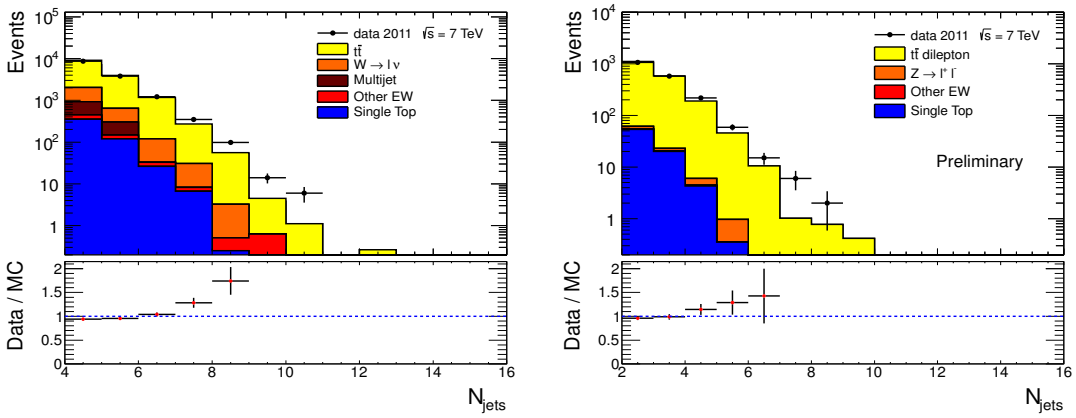

Figure 4.3: The jet multiplicities in the semileptonic (left) and dileptonic sample (right)

Since the shape measurement is based on the energy deposits in the calorimeter, jets which are not separated in $\eta - \varphi$ by more than $\Delta R = 0.8$, which is double the jet radius, are discarded. This is done to take into account possible overlaps between the jet cones, which would bias the shape measurement. These overlapped configurations are normally encountered in boosted W bosons, leading to light jets which are not well separated. We would like to stress that this cut does not affect much the b-jets, which coming from a massive top quark are well separated from other jet activity, but strongly suppresses the light jet sample.
To select the b-jets, the JetFitterCombNN tagger is used with a working point of 2.4. This leads to a b-tagging efficiency of 57%, while significantly reducing the light jet background. All the Monte Carlo distributions shown in this chapter are pileup reweighted, in order to fit the distributions for the average number of interactions per bunch crossing in data and MC.
Additional scale factors (SF) are applied to make the Monte Carlo b-tagging effi-

ciency agree with that of the data, and they are also applied to light-jets (which are required to be anti-tagged) to correct for the inefficiencies. The scale factors are provided by the b-tagging group depending on each jet pseudorapidity and transverse momentum such that for each event, a weight w is calculated as the product of all the b-tagged jets weights times the anti-tagged weights.

$$w = \prod_{i=1}^{N_{jets}} w_i^{\text{jet}} \tag{4.3}$$

The jet-by-jet scale factors, as provided by the b-tagging working group for the configuration used here (JetFitterCombNN with $\epsilon = 57\%$), together with their uncertainties, are summarized in table 4.10 for b-jets (efficiency SF) and tables 4.11 and 4.12 for light jets (inefficiency SF) in the central and forward regions of the detector.

p_{T} range	Scale Factor	Uncertainty
20 GeV $< p_{\text{T}} <$ 30 GeV	0.870	0.081
30 GeV $< p_{\text{T}} <$ 60 GeV	0.938	0.077
60 GeV $< p_{\text{T}} <$ 90 GeV	0.941	0.138
90 GeV $< p_{\text{T}} <$ 140 GeV	0.868	0.010
140 GeV $< p_{\text{T}} <$ 200 GeV	0.868	0.162

Table 4.10: b-tagging efficiency scale factors for b-jets

p_{T} range	Scale Factor	Uncertainty
20 GeV $< p_{\text{T}} <$ 25 GeV	1.338	0.281
25 GeV $< p_{\text{T}} <$ 40 GeV	0.976	0.227
40 GeV $< p_{\text{T}} <$ 60 GeV	1.096	0.225
60 GeV $< p_{\text{T}} <$ 90 GeV	1.119	0.305
90 GeV $< p_{\text{T}} <$ 140 GeV	0.944	0.306
140 GeV $< p_{\text{T}} <$ 200 GeV	1.291	0.582
200 GeV $< p_{\text{T}} <$ 300 GeV	0.967	0.538
300 GeV $< p_{\text{T}} <$ 500 GeV	1.108	0.517

Table 4.11: b-tagging inefficiency scale factors for light-jets in the central region $|\eta| < 1.2$

p_T range	Scale Factor	Uncertainty
20 GeV $< p_T <$ 25 GeV	1.046	0.230
25 GeV $< p_T <$ 40 GeV	1.295	0.311
40 GeV $< p_T <$ 60 GeV	1.102	0.377
60 GeV $< p_T <$ 90 GeV	0.774	0.237
90 GeV $< p_T <$ 140 GeV	1.098	0.322
140 GeV $< p_T <$ 200 GeV	1.182	0.574
200 GeV $< p_T <$ 300 GeV	1.282	0.577
300 GeV $< p_T <$ 500 GeV	1.039	0.299

Table 4.12: b-tagging inefficiency scale factors for light-jets in the forward region $1.2 < |\eta| < 2.5$

The JVF and JetFitterCombNN distributions are shown in figure 4.4.

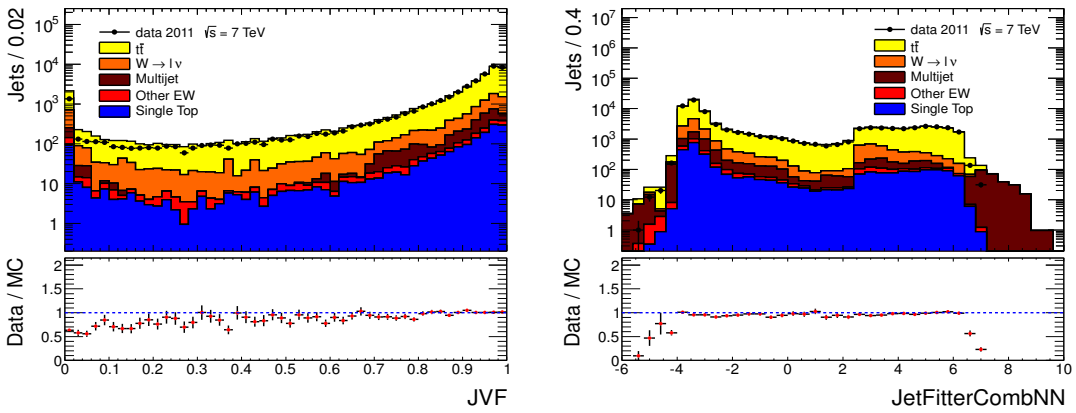

Figure 4.4: JVF and JetFitterCombNN distributions for the semileptonic sample

4.4.1 b-jet samples

To select the b-jet samples, the recommended cut JetFitterCombNN > 2.4 is used. This cut is tuned to achieve an efficiency $\epsilon_b = 57\%$ and a rejection factor of about 500 [85, 89]. The b-tagging algorithm has additional operating points tuned at b-tagging efficiencies of 70% and 80%, leading to light quark rejection factors of 100 and 30 approx. Choosing an operating point other than that at 2.4 results in slightly different purities for both the b- and light-jet samples. This is

discussed further in Sect. 4.8.3

The purity of the b-jet samples is defined as the number of selected jets which are really b-jets over the total number of jets in the sample. It will be shown in the following sections that the purity achieved using this b-tagging algorithm in this b-enriched final state is nearly 90% for the semileptonic sample and is well over this number for the dileptonic one.

The resulting number of b-jets selected in the dileptonic (semileptonic) sample is 2279 (resp. 16735). Figure 4.5 shows the b-jet transverse momentum distributions for the dileptonic and semileptonic channels.

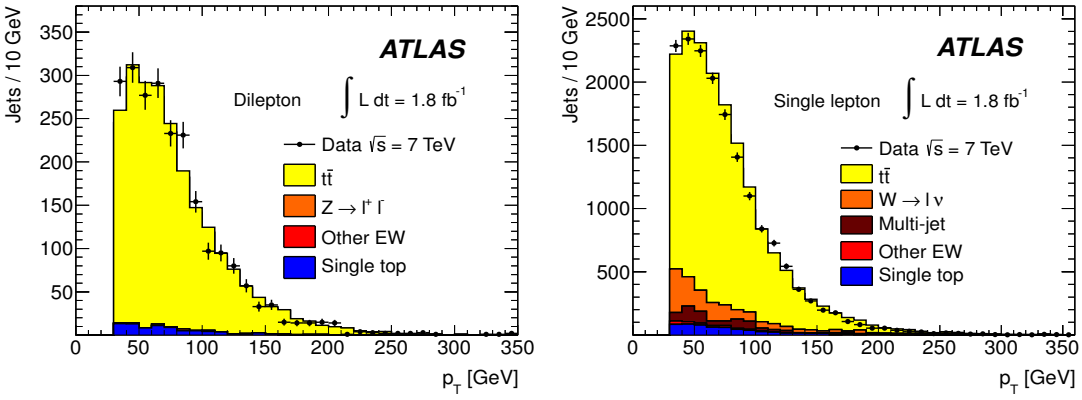

Figure 4.5: The p_{T} distributions for b-jets in the dileptonic (left) and the semileptonic (right) samples along with MC expectations

The following comment is in order: The p_{T} distributions for the b-jets in both the dileptonic and semileptonic samples show a similar behaviour, as expected since they all come from top quark decays. In the kinematic range relevant for this study, they show a fall-off of about 2 orders of magnitude. This is in good agreement with the Monte Carlo expectations as obtained from the MC@NLO generator coupled to the HERWIG fragmentation, as also shown in figures 4.5. We also would like to note that the dileptonic sample is less affected by the background contamination than the semileptonic one, as it has been quantitatively shown in tables 4.6 and 4.8. In particular, the QCD background is expected to

be negligible.

4.4.2 Light jet sample

The hadronic decays $W \to q\bar{q}'$ are a clean source of light quark induced jets, as gluons and b-jets are very supressed. The former because gluons would have to come from radiative corrections of order $\mathcal{O}(\alpha_s)$, and the latter because of the smallness of the CKM matrix elements $|V_{ub}|$ and $|V_{cb}|$. To define the light jet sample, the jet pair in the event which has the closest invariant mass to the nominal mass $m_W = 80.4$ GeV is selected. This pair of jets is also required to be anti-tagged by the JetFitterCombNN algorithm. The number of jets satisfying these criteria is 8376. Figure 4.6 shows the transverse momentum distribution of these jets and the invariant mass of the light-jet pair.

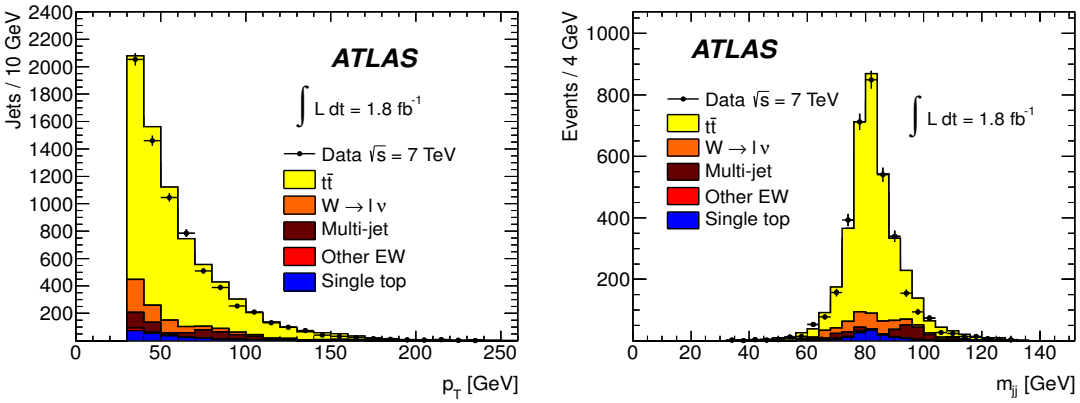

Figure 4.6: The p_{T} and invariant mass distributions for the light-jet pair along with MC expectations

The p_{T} distribution of the light jets coming from W decays exhibit a stronger fall-off than those for the b-jets which were discussed in subsection 4.4.1. This dependence is again in very good agreement with the Monte Carlo expectations. Analogously to the b-jet sample, the light jet purity is defined as the fraction of truth light jets (including u, d, c, s) selected in the light-jet sample. The result is $p_s^{(l)} = 66.16\%$. This purity can be raised to 83% by demanding two b-tagged jets instead of one, but at the price of reducing the statistics by a factor of 4.

4.4.3 Jet purities

To estimate the number of fake light or b-jets which have been selected using the above criteria, the truth Monte Carlo information is used. For b-jets, the truth flavor is known via a $\Delta R < 0.2$ matching with a B-hadron. For light jets, it is demanded that the parton with highest p_{T} within the jet is a light (u, d, c, s) quark. The purity is then defined as

$$p = \sum_k \alpha_k p_k; \quad p_k = 1 - \frac{N_f^{(k)}}{N_T^{(k)}} \tag{4.4}$$

where α_k is the proportion of events in the k-th MC sample (signal or background), given in tables 4.6 and 4.8 and $N_f^{(k)}$, $N_T^{(k)}$ are the number of fakes and the total number of jets in a given sample, respectively.
The calculation for the dileptonic sample is summarized in table 4.13. The contamination in the b-jet sample, as shown in table 4.14, is dominated by light-jet fakes, with the gluon contamination being below 1%. For the light jet sample, the fraction of gluon fakes amounts to 19.42%, while the b-jet fakes correspond to 15%.

As stated in the tables below, the result for the dileptonic channel is $p_d^{(b)} = 99.27\%$, whereas for the semileptonic channel the purity is found to be $p_s^{(b)} = 88.5\%$. It is important to note that the purity achieved using top events is much larger than that in inclusive measurements at Tevatron and the LHC [70, 84].

Process	α_k	p_k (b-jets)
$t\bar{t}$ dileptonic	0.9490	0.9971
$Z \to \ell^+\ell^-$ + jets	0.0039	0.2467
$Z \to \ell^+\ell^- + b\bar{b}$	0.0022	0.9913
Other (W, WW, WZ, ZZ)	0.0016	0.3752
Single Top	0.0433	0.9870
Weighted total	-	99.27%

Table 4.13: Coefficients for the purity calculation in the dileptonic channel using equation 4.4

Process	α_k	p_k (b-jets)	p_k (light jets)
$t\bar{t}$ semileptonic	0.7741	0.9606	0.7248
$W^{\pm} \to \ell^{\pm}\nu_l$ + jets	0.0309	0.4803	0.3626
$W^{\pm} \to \ell^{\pm}\nu_l + b\bar{b}$	0.0326	0.9577	0.2989
$W^{\pm} \to \ell^{\pm}\nu_l + c\bar{c}$	0.0438	0.1389	0.4437
$W^{\pm} \to \ell^{\pm}\nu_l + c(\bar{c})$	0.0204	0.1301	0.2746
QCD multijet	0.0502	0.8869	0.4853
$Z \to \ell^+\ell^-$ + jets	0.0052	0.2326	0.3313
$Z \to \ell^+\ell^- + b\bar{b}$	0.0055	0.9776	0.3429
Diboson (WW, WZ, ZZ)	0.0005	0.5181	0.4470
Single Top	0.0368	0.9577	0.7158
Weighted total	-	88.50%	66.16%

Table 4.14: Coefficients for the purity calculation in the semileptonic channel using equation 4.4

4.5 Jet shapes in the semileptonic channel

Jets are based on calorimetric measurements. The electromagnetic read-out cells have a granularity of typically $\Delta\eta \times \Delta\varphi = 0.025 \times 0.025$ at shower maximum, while hadronic cells have a granularity of $\Delta\eta \times \Delta\varphi = 0.1 \times 0.1$. Topological clusters are built around seed calorimeter cells with energy deposition larger than four times the r.m.s. of the noise energy distribution. In the iterative procees underlying the jet definition, cells with energy larger than twice the r.m.s. of the noise are considered. For the jet samples studied in this work, the jet axis resolution is 0.02 in η and 0.025 in φ.
The differential jet shape $\rho(r)$ in an annulus of inner radius $r - \Delta r/2$ and outer radius $r + \Delta r/2$ from the axis of a given jet is defined as

$$\rho(r) = \frac{1}{\Delta r} \frac{p_{\mathrm{T}}(r - \Delta r/2, r + \Delta r/2)}{p_{\mathrm{T}}(0, R)}; \quad r \leq R \tag{4.5}$$

where $p_{\mathrm{T}}(r_1, r_2)$ is the scalar sum of the p_{T} of the clusters in the annulus with radii r_1 and r_2. The distributions for $\rho(r)$ for several values of the jet internal radius are shown in figure 4.7 for the b-jet sample.

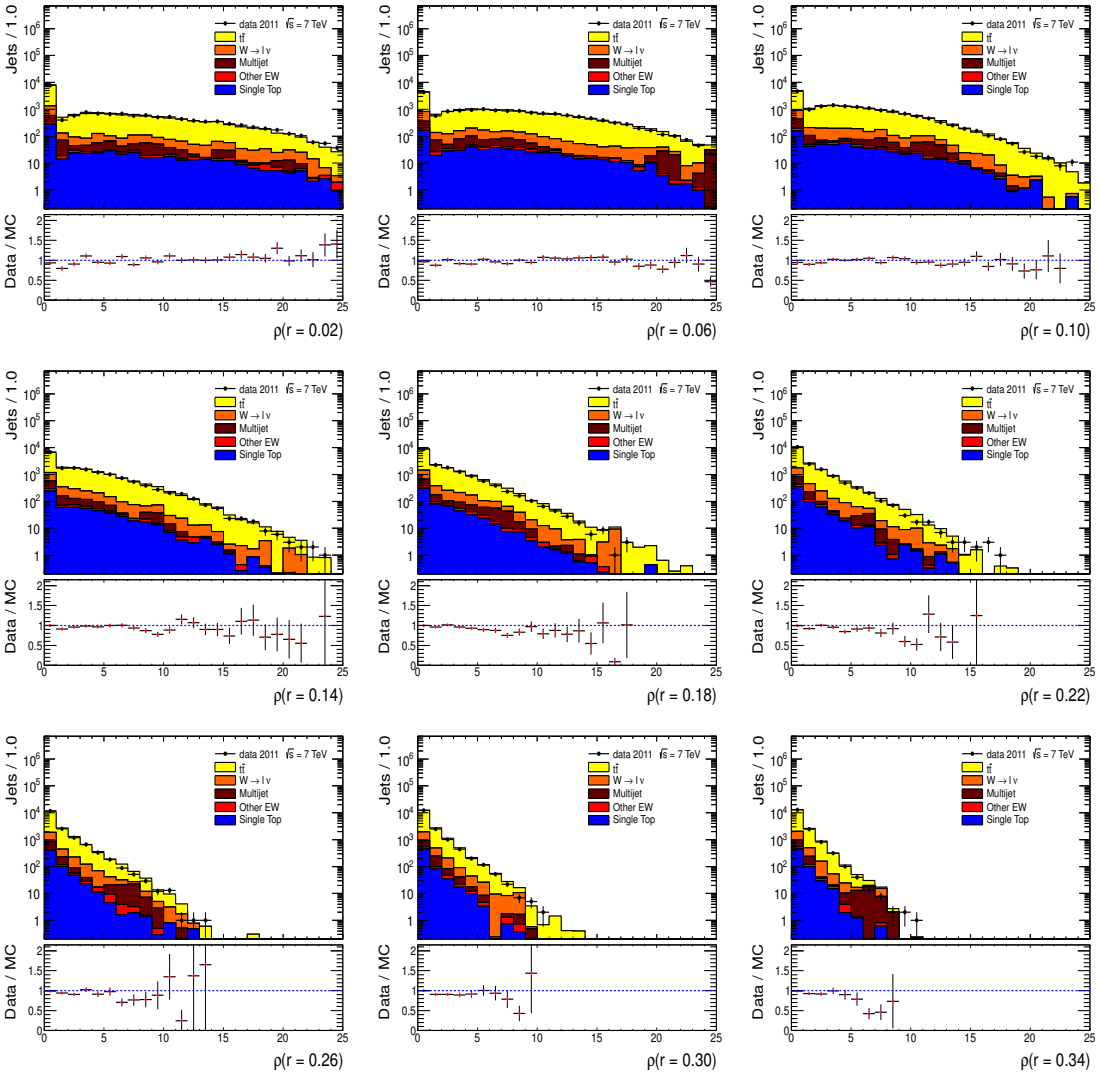

Figure 4.7: $\rho(r)$ distributions for b-jets along with MC expectations

As it is seen, the values for $\rho(r)$ decrease as r is increased, being well peaked at 0, which indicates that the majority of the energy is deposited at low values of r, that is, in the internal parts of the jet cone. In a similar manner, the distributions for $\rho(r)$ in the light-jet sample are shown in figure 4.8

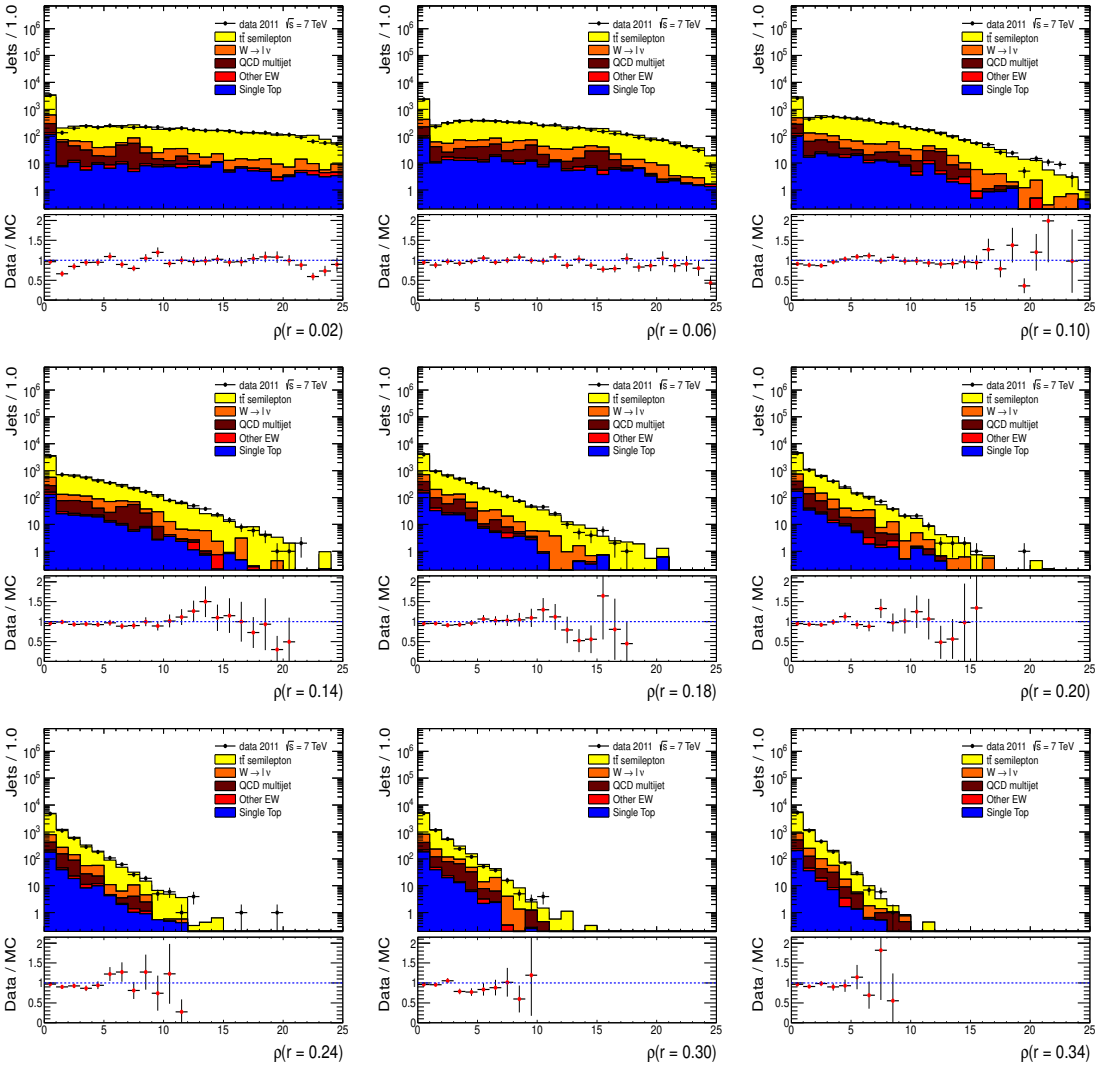

Figure 4.8: $\rho(r)$ distributions for light jets along with MC expectations

Alternatively, the integrated jet shape in a cone of radius $r \leq R$ (here $R = 0.4$ is used) around the jet axis is defined as the cumulative distribution of the differential jet shape, i.e.

$$\Psi(r) = \frac{p_{\mathrm{T}}(0, r)}{p_{\mathrm{T}}(0, R)}; \quad r \leq R \tag{4.6}$$

which satisfies $\Psi(r = R) = 1$. The distributions for $\Psi(r)$ in steps of $\delta r = 0.04$ are shown in figure 4.9 for the b-jet sample selected in the semileptonic channel

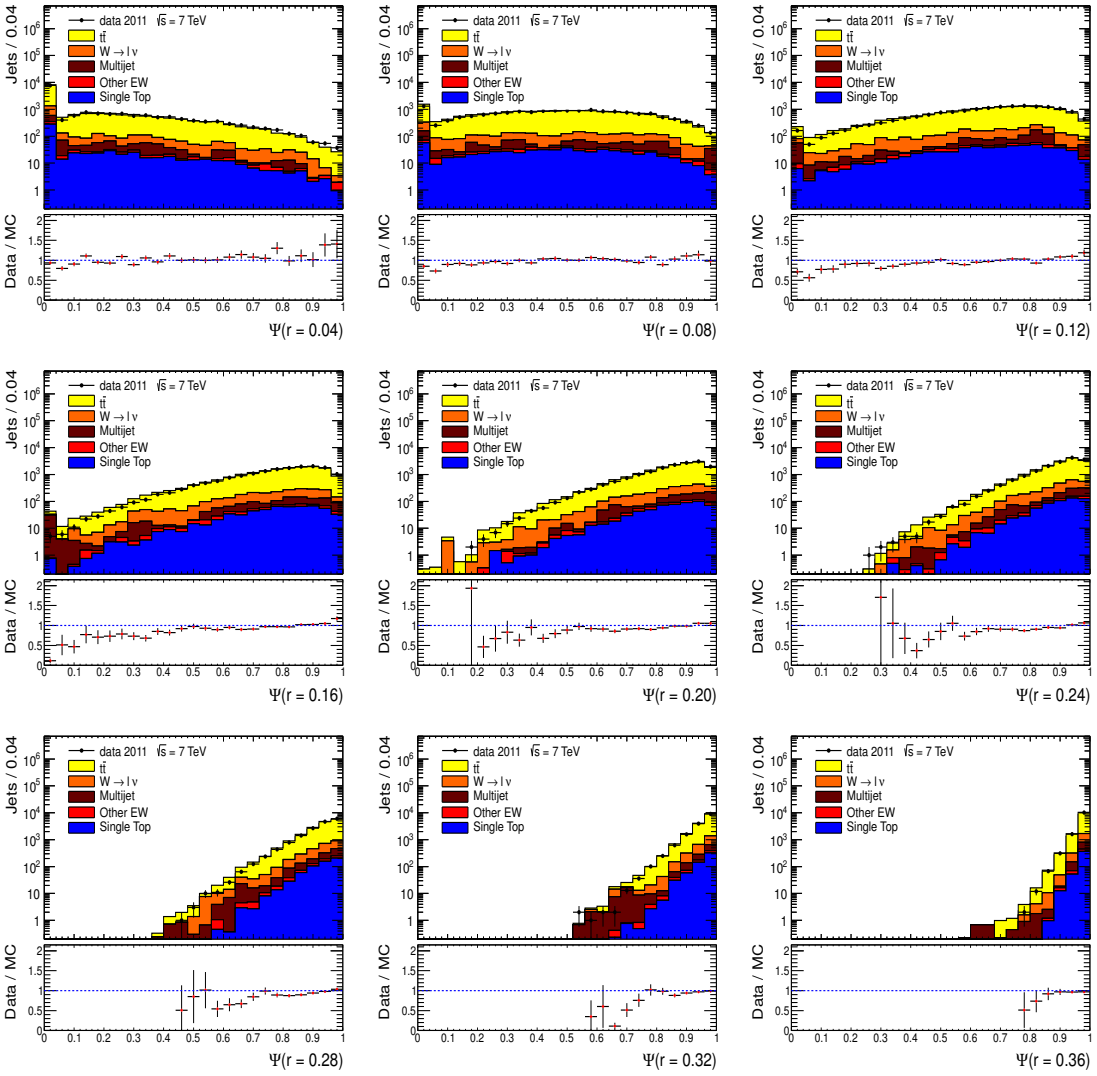

Figure 4.9: $\Psi(r)$ distributions for b-jets along with MC expectations

The following comments are in order: As r increases, the distributions for $\Psi(r)$ shrinks to values closer to the upper limit of 1, and for low values of r, the distributions show a very strong peak at the lower limit of 0 (note the logarithmic scale). Both of these effects are fairly well reproduced by the Monte Carlo, though some small discrepancies are observed, mostly at the kinematical ends.
Analogously, the $\Psi(r)$ distributions for light jets are shown in figure 4.10.

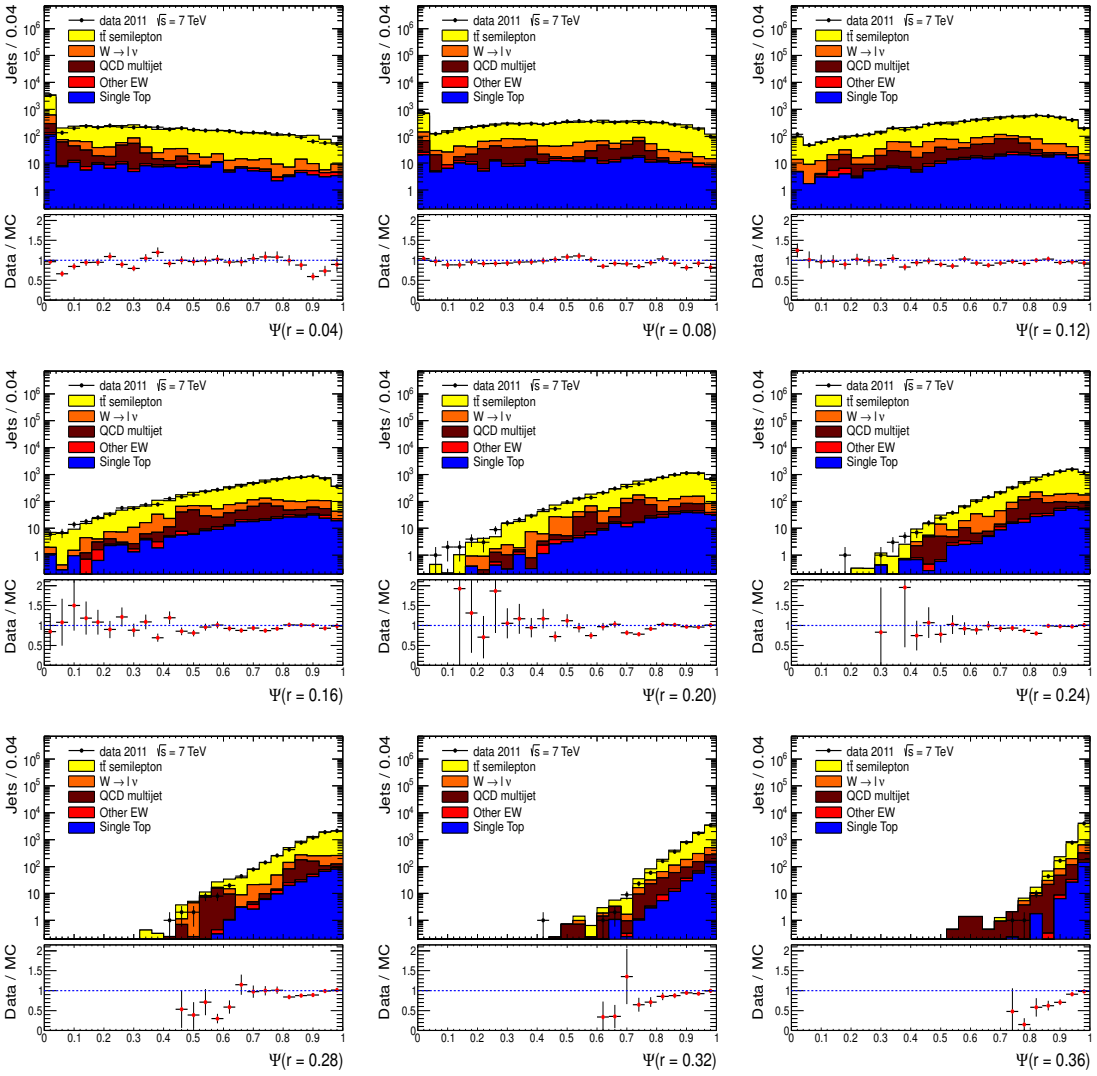

Figure 4.10: $\Psi(r)$ distributions for light jets along with MC expectations

In the following, the mean values of these distributions will be used

$$\langle \rho(r) \rangle = \frac{1}{\Delta r} \frac{1}{N_{jet}} \sum_{jets} \frac{p_{\mathrm{T}}(r - \Delta r/2, r + \Delta r/2)}{p_{\mathrm{T}}(0, R)}; \quad \langle \Psi(r) \rangle = \frac{1}{N_{jet}} \sum_{jets} \frac{p_{\mathrm{T}}(0, r)}{p_{\mathrm{T}}(0, R)} \tag{4.7}$$

where the sum is performed over all jets of a given sample, light (l) or b-jets (b) and N_{jet} is its number of elements. The crucial observation will be that $\langle \Psi(r) \rangle_b < \langle \Psi(r) \rangle_l$, as will be seen in more detail in the following sections.

4.5.1 Shape dependence on pseudorapidity

In this section the dependence of the integrated shape with the jet pseudorapidity is studied. To do this, the jets are classified in pseudorapidity bins, and for each of them the mean value of $\Psi(r)$ is plotted against r. From now on, to avoid fluctuations due to limited statistics in the tail of the p_{T} distributions, the range 30 GeV $< p_{\mathrm{T}} <$ 200 GeV will be considered. The results are shown in figure 4.11.

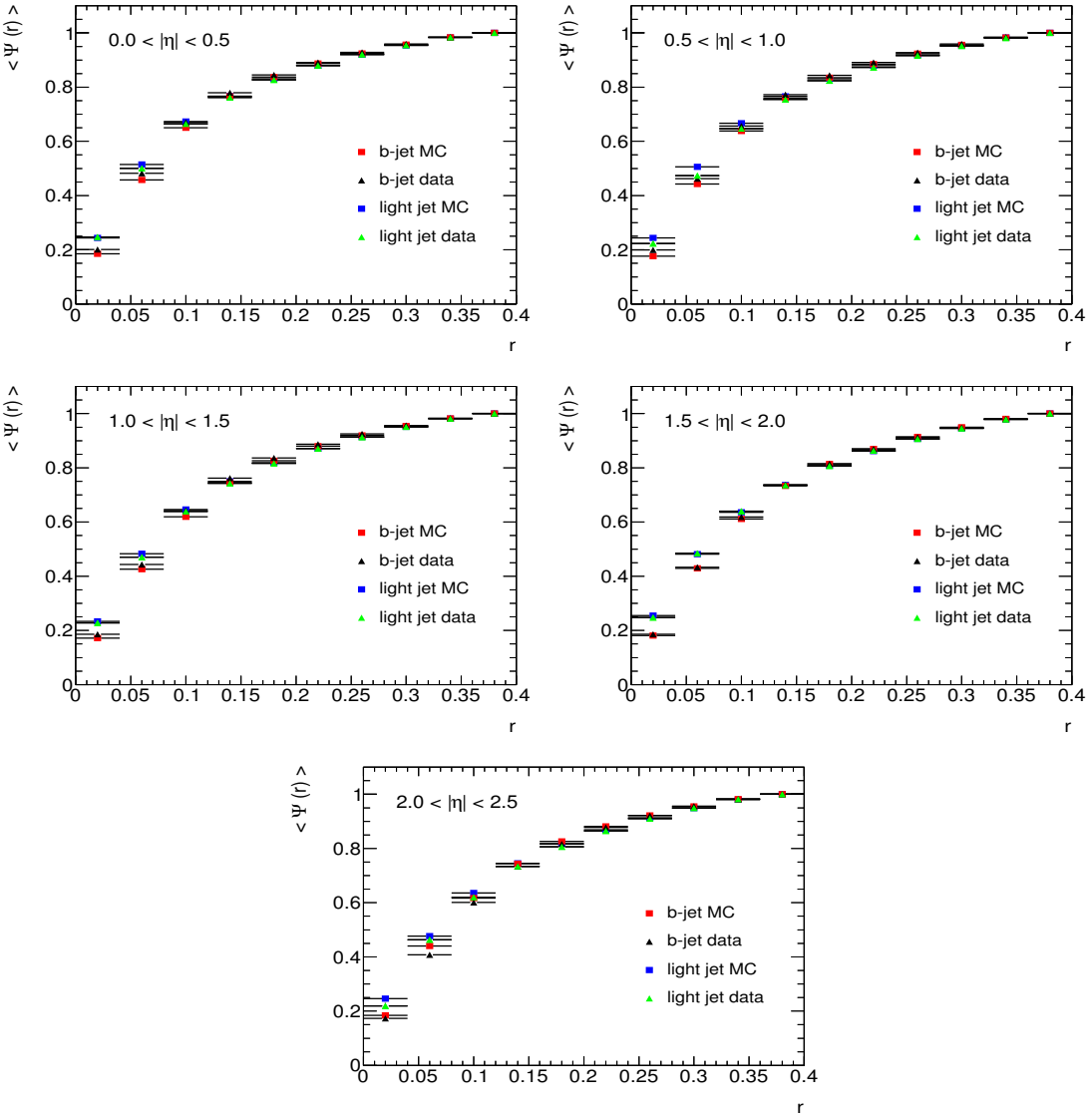

Figure 4.11: $\Psi(r)$ dependence with radius for $|\eta| < 2.5$ along with MC expectations

We want to remark that for every bin in pseudorapidity, the b-jet shapes are wider than the light jets, as discussed in the previous subsection. It is also important to note that the differences between b-jets and light jets are less pronounced when integrated over p_{T} due to the fact that the transverse spectra are different for both samples.

Moreover, these plots are very similar to each other, which leads us to think that the jet shape dependence on pseudorapidity is very weak. To quantify this dependence, the mean value of the variable $1 - \Psi(r = 0.2)$, which is the fraction of energy in the outer half cone of the jet, is plotted versus the absolute value of the pseudorapidity. Figure 4.12 shows the dependence of this variable with $|\eta|$, which is found to be flat, in agreement with the Monte Carlo expectations. This has also been observed in [63, 64, 65, 66, 67, 68, 20, 70].

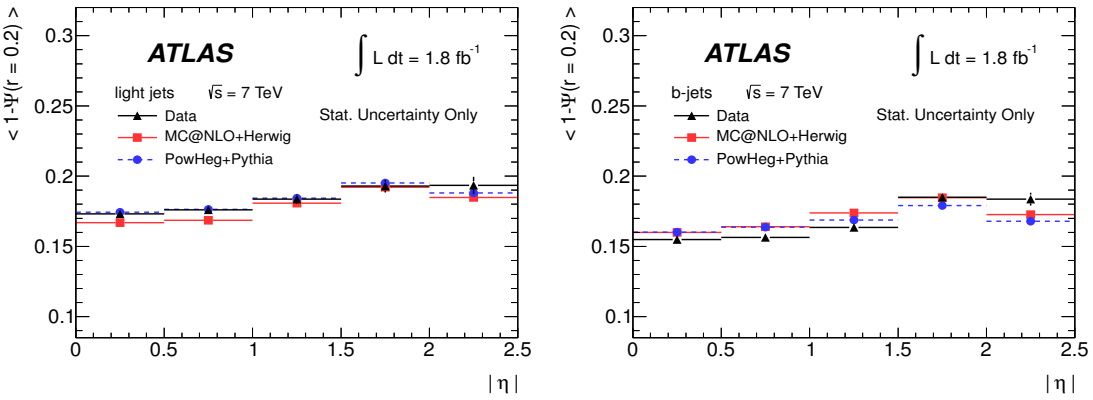

Figure 4.12: Shape dependence on pseudorapidity along with MC expectations

4.5.2 Shape dependence on transverse momentum

Once the dependence with pseudorapidity is clear, one can study the dependence of $\Psi(r)$ with the jet transverse momentum. To do this, the same procedure as in the previous subsection is followed, dividing the jet p_{T} spectra into bins and plotting the average value of $\Psi(r)$ versus r, separately for each bin, as shown in figure 4.13.

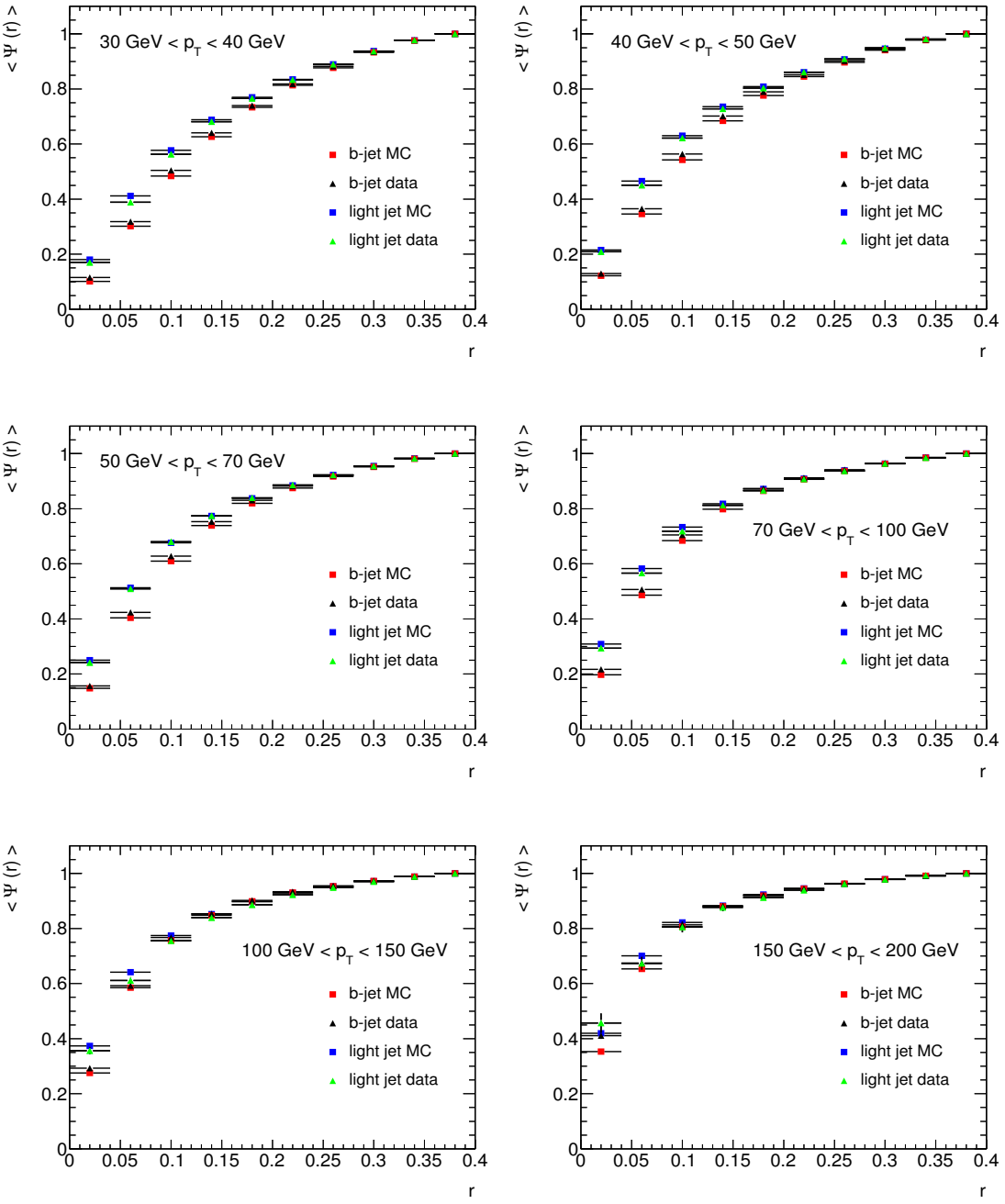

Figure 4.13: $\Psi(r)$ dependence with the jet internal radius in p_{T} bins along with MC expectations

Again, we would like to stress that $\langle\Psi(r)\rangle_b < \langle\Psi(r)\rangle_l$ for each bin in p_{T} [70]. Of course, the difference is more clearly seen for r values smaller than 0.12.

Moreover, it is also observed that the jets are becoming more collimated as p_{T} increases. To quantify this effect, the variable $1 - \Psi(r = 0.2)$ is again used. The result is that $\Psi(r)$ is very sensitive to the jet momentum scale, as also reported in [63, 64, 65, 66, 67, 68, 20, 70], as it can be seen in figure 4.14 for the light and b-jet samples. The agreement with the Monte Carlo expectations is good for the light jets, and fair for the b-jets.

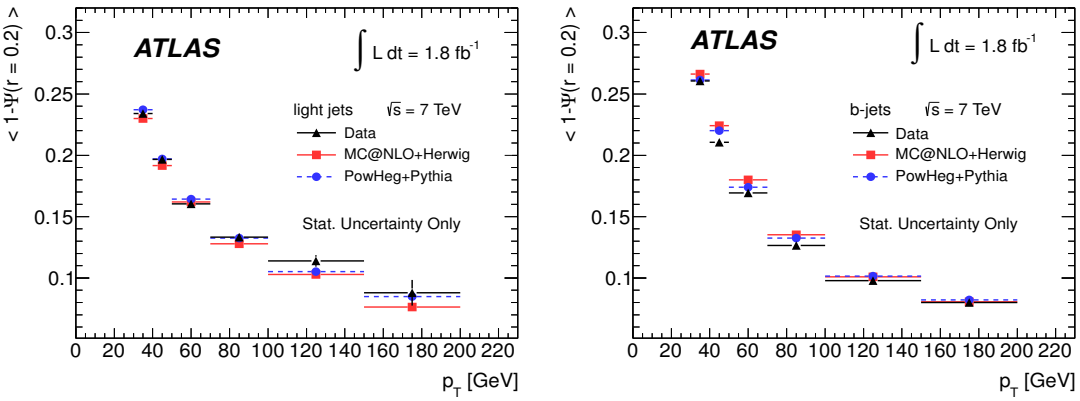

Figure 4.14: The p_{T} dependence of the jet shapes of the light and b-jet shapes along with MC expectations

In order to make clearer the statement that b-jets have wider energy distributions inside the jet cones, the radius $r_{1/2}$ is calculated. This is the value of r at which the energy flow is 50% of the total, i.e. $\Psi(r_{1/2}) = 0.5$. To do this, the plots above are interpolated using the cubic splines method, dividing the full range in 10^4 subranges and finding the point in this grid for which the function $f(r) = \Psi(r) - 1/2$ changes sign. The result is, as expected, that $r_{1/2}$ is greater for b-jets than for light jets. Figure 4.15 shows the running of $r_{1/2}$ with the jet transverse momentum and pseudorapidity.

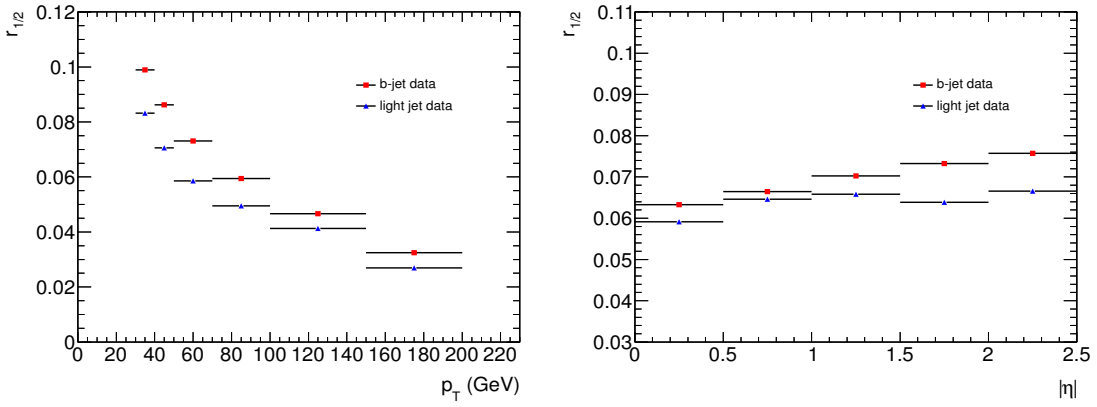

Figure 4.15: The radius $r_{1/2}$ as a function of the jet transverse momentum (left) and pseudorapidity (right)

Results show that $r_{1/2}$ is very dependent with the jet p_T and not much dependent with η. However, the dependence of $r_{1/2}$ with the pseudorapidity is not negligible in this case, as it was for $\Psi(r)$ versus η.

4.6 Jet shapes in the dileptonic channel

In order to cross check the results discussed above, a sample of $t\bar{t}$ events in the dileptonic channel is selected according to the method described in section 4.3.1. Here both W bosons decay in the $W \to \ell\nu$ channel. The jet sample is selected following the same procedure as in the semileptonic channel: jets are required to have $p_{\mathrm{T}} > 30$ GeV and those with an overlap below $\Delta R = 0.8$ are removed. Figure 4.16 shows the profiles of the b-jet shapes for the dileptonic sample compared to those for the light jets in the semileptonic channel.

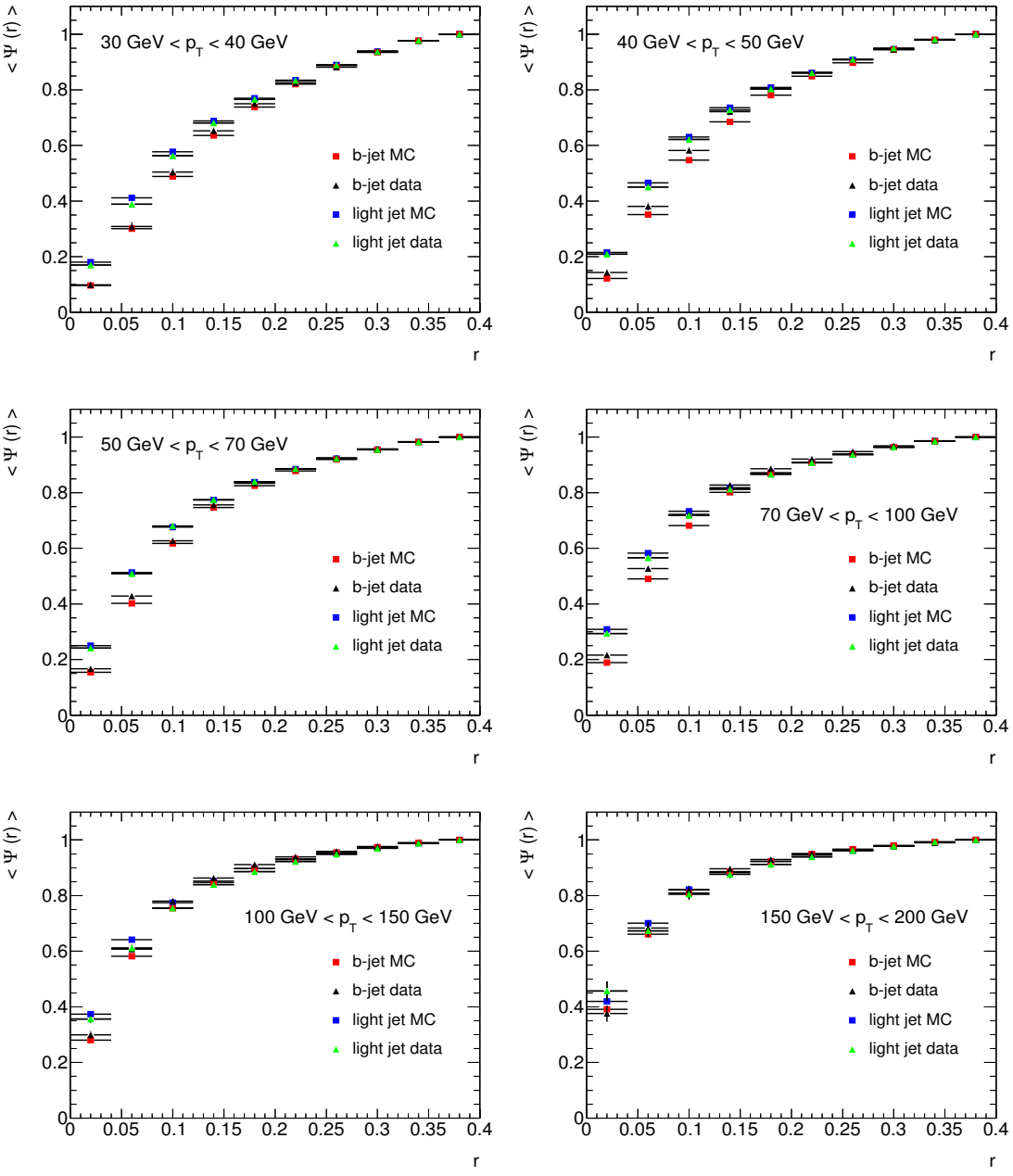

Figure 4.16: Jet shapes for 30 GeV $< p_{\mathrm{T}} <$ 200 GeV in the dileptonic channel along with MC expectations

The results obtained are compatible with the ones discussed in the previous section, thus confirming the observation that b-jets are wider than light ones. In order to compare the shapes between the selected b-jets in both samples, the

ratio α_{ds} is defined between the average b-jet shapes in the dileptonic (d) and the semileptonic (s) samples, i.e.

$$\alpha_{ds}(r) = \frac{\langle \Psi_b(r) \rangle_d}{\langle \Psi_b(r) \rangle_s} \tag{4.8}$$

Figure 4.17 shows the ratio α_{ds} as a function of the jet internal radius

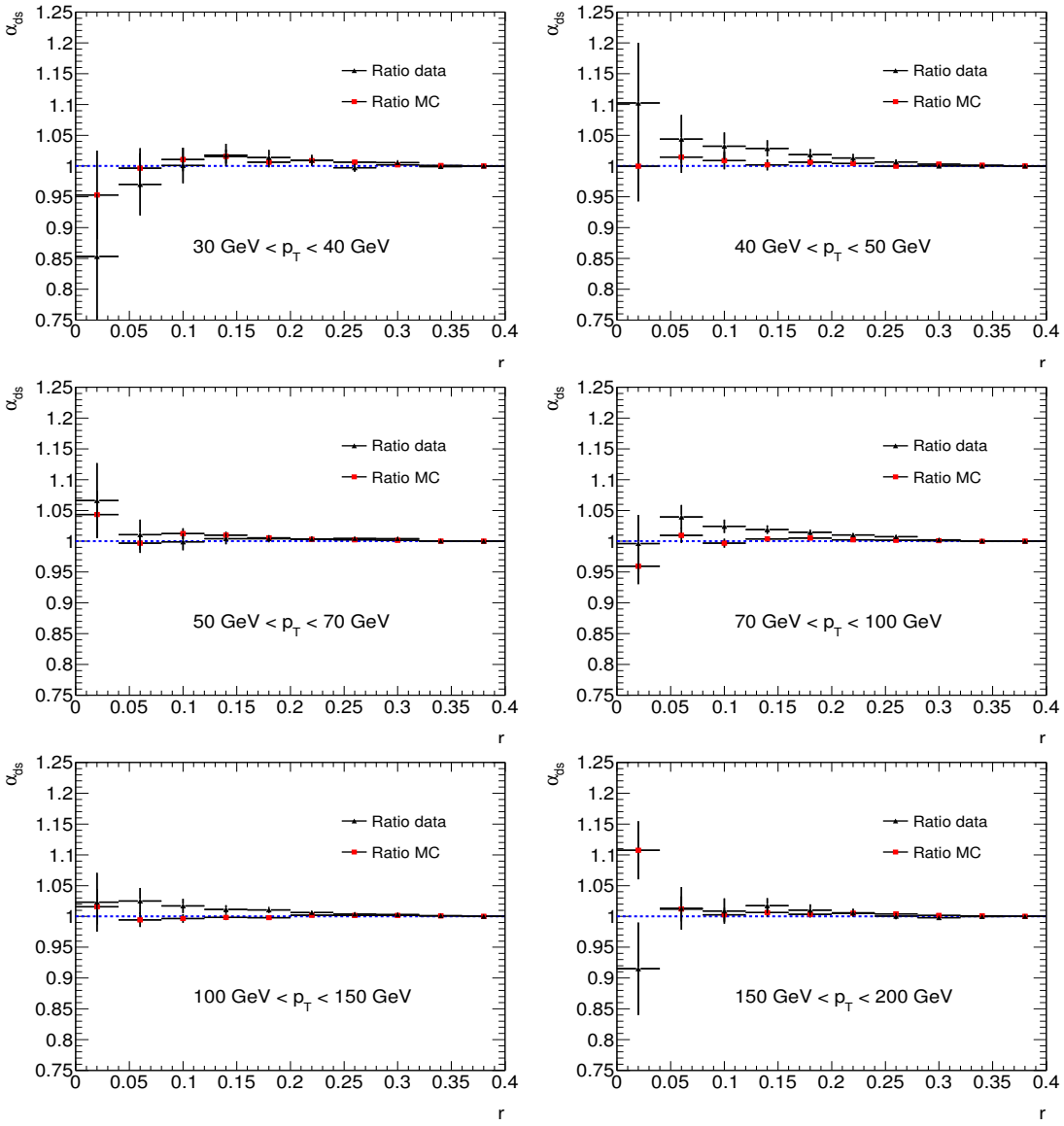

Figure 4.17: Comparison between b-jet samples from dileptonic and semileptonic $t\bar{t}$ events along with MC expectations

4.7 Results at the detector level

Once it has been proved that the b-jets from the dileptonic and the semileptonic channel are compatible, they can be added in a single sample. The results for the differential jet shape at the detector level are shown in figure 4.18, together with the MC expectations from the reference generators discussed before. As it is seen, the differential jet shapes are fairly well described by the MC samples.

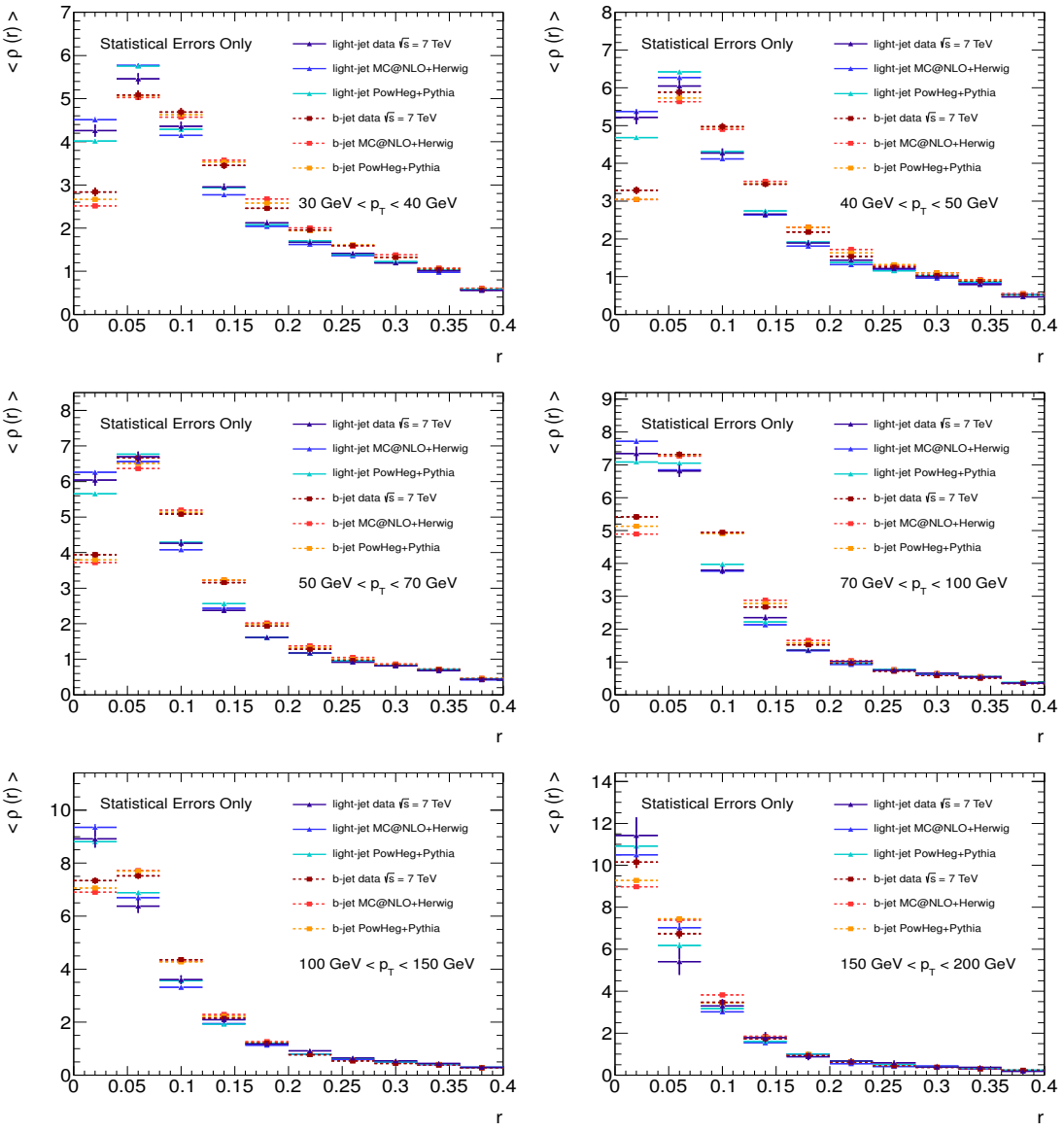

Figure 4.18: Differential jet shapes at the detector level.

The results for the integrated jet shapes at the detector level are summarized in figure 4.19.

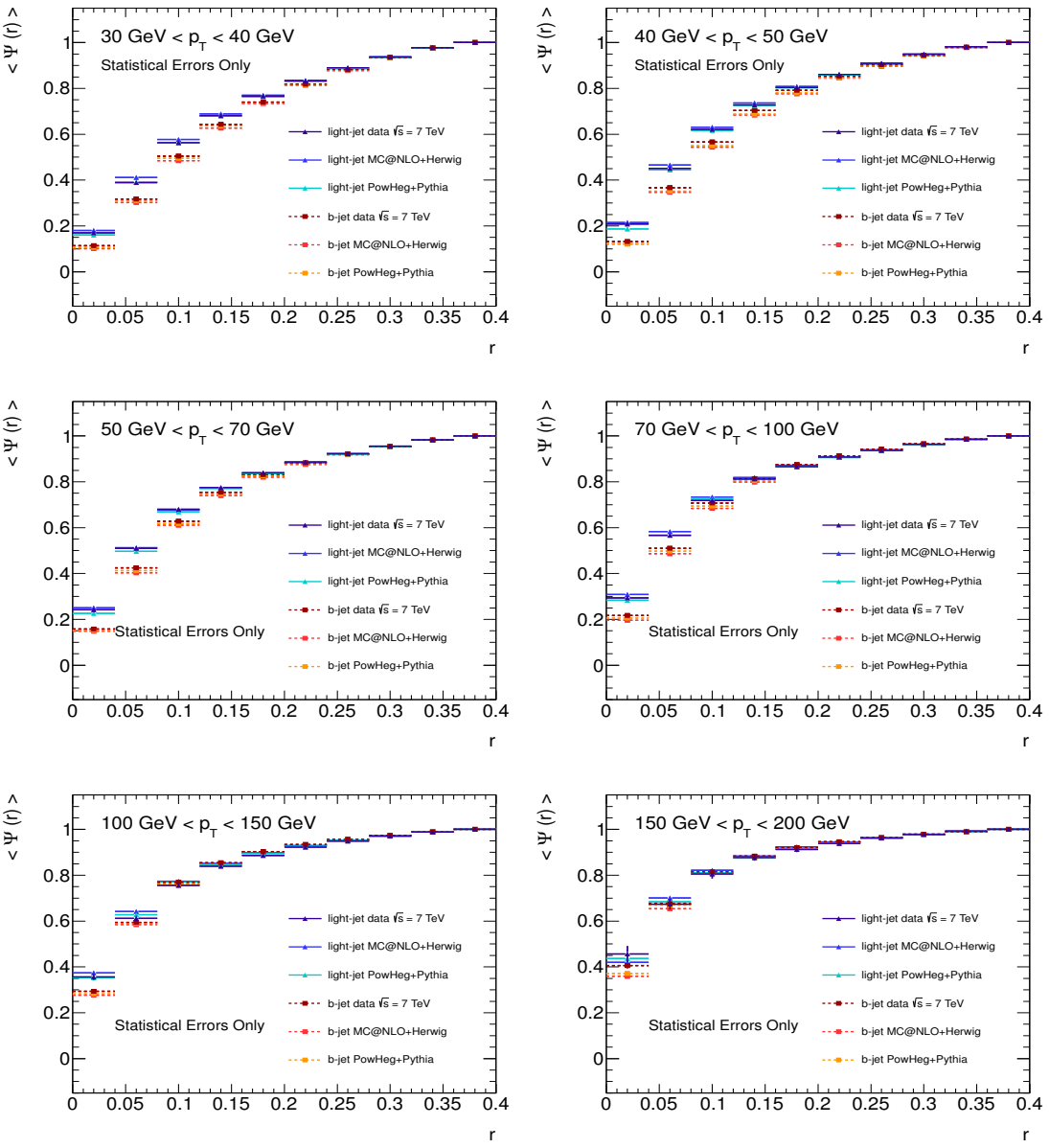

Figure 4.19: Integrated jet shapes at the detector level.

4.8 Systematic uncertainties and stability checks

In this section, an estimation of some systematic uncertainties due to biases induced by b-tagging cuts or jet purity is presented. It is observed that the dominant effects (Jet energy scale & resolution) do not go beyond 10% in most bins.

4.8.1 Jet Energy Scale

The propagation of the jet energy scale uncertainty [56, 57] to the jet shape systematic uncertainty is discussed in this section. To study this effect, the jet energy and transverse momentum are shifted by quantities Δ_+ and Δ_- in the MC samples and before the event selection. Figure 4.20 shows the shifted p_T distributions for light and b-jets. Just in this case and for illustrative reasons, jets with $p_{\mathrm{T}} > 200$ GeV are also considered.

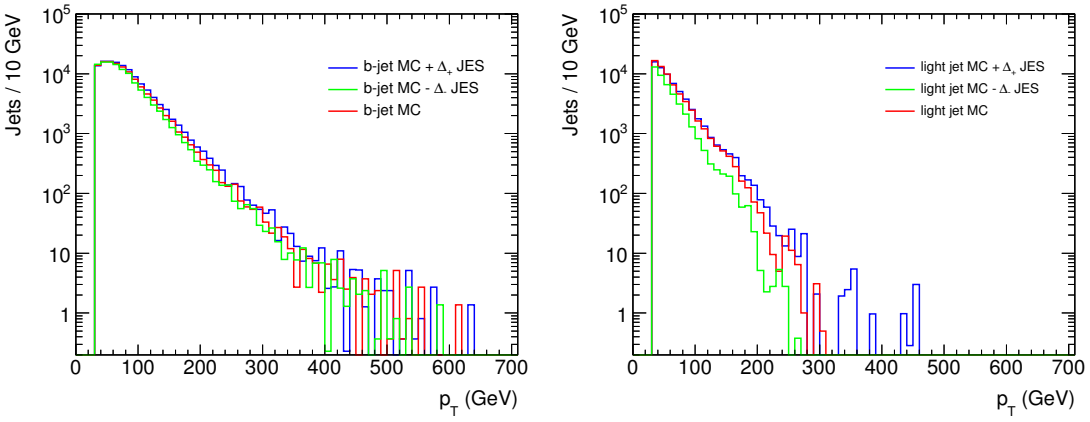

Figure 4.20: JES-shifted p_T distributions for light and b-jets

This shift leads to asymetric uncertainties in the jet shape $\Psi(r)$ due to the fact that the number of jets is increased (decreased) differently as the jet p_T is shifted up (down) in to the jet selection criteria. An additional JES uncertainty affecting only b-jets (bJES) is added in quadrature to the nominal JES uncertainty for jets with true b-flavour. JES is the dominant systematic uncertainty for b-jets, not representing more than 10% in any case.

4.8.2 Jet Energy Resolution

The jet energy resolution in the ATLAS calorimeter is given by

$$\frac{\sigma_{p_T}}{p_T} = \frac{N}{p_T} \oplus \frac{S}{\sqrt{p_T}} \oplus C \tag{4.9}$$

Where N, S and C are the noise, stochastic and constant terms. To compute the uncertainty due to the effect of the jet resolution, the jet energy and transverse momentum are smeared by convoluting these distributions with a gaussian $N(1, \sigma)$ with σ a standard deviation depending on the jet p_T and pseudorapidity [83]. The effect on the transverse momentum distributions can be seen in figure 4.21

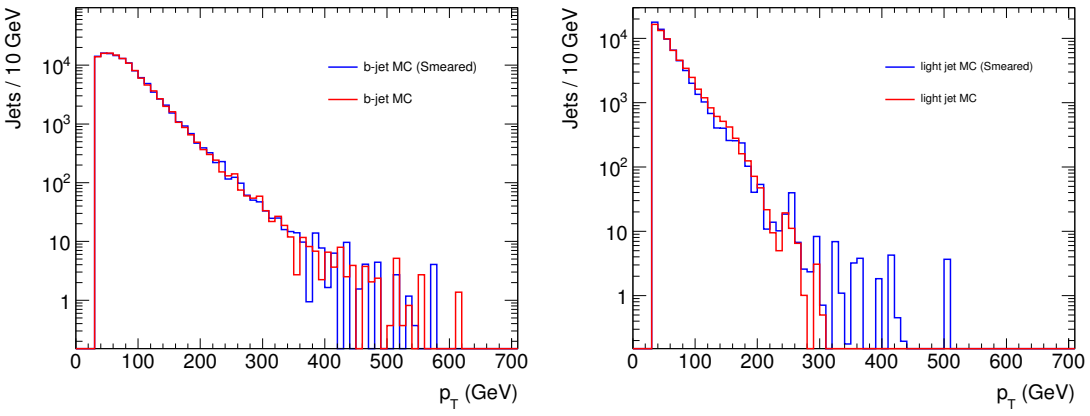

Figure 4.21: JER-smeared p_T distributions for light and b-jets

This smearing is then propagated to the jet shapes, and added in quadrature to the JES uncertainty in order to present both of them as a unique source of systematic error.

4.8.3 Bias due to jet purity

4.8.3.1 *b*-jets

This section contains an estimation of the bias induced in the *b*-jet sample by cutting on JetFitterCombNN > 2.4. To study this effect, Monte Carlo *b*-jets, for which the truth flavour is known, are selected. The differences with the jets tagged by the *b*-tagging algorithm are studied. The JetFitterCombNN variable does not depend much on the jet p_T, as can be seen in figure 4.22, as well as its dependence with the number of primary vertices.

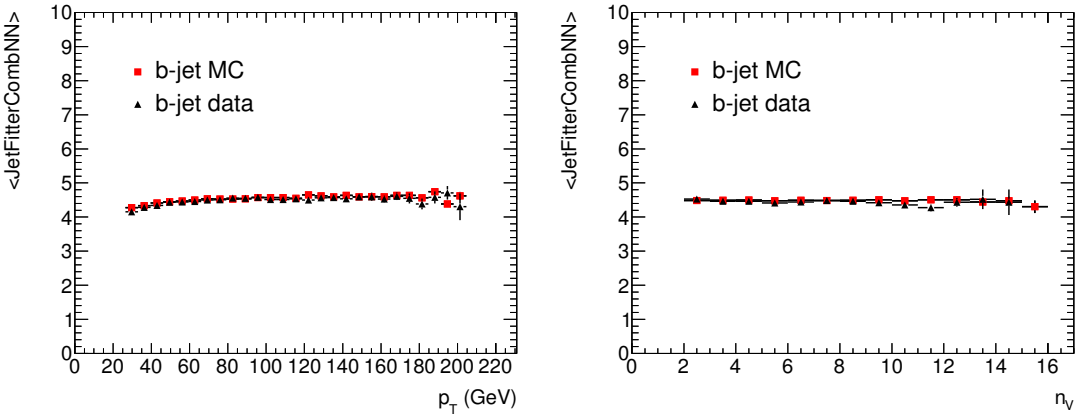

Figure 4.22: The JetFitterCombNN variable dependence with the jet p_T (left) and the number of primary vertices (right) for the *b*-jet sample

The analysis has been repeated using the working point JetFitterCombNN > 0.35, yielding a *b*-tagging efficiency of 70% and purities of 82% for *b*-jets and 67% for light jets. The relative difference $\gamma_{l,b}^{(\rho)}(r)$ defined, for light and *b*-jet differential jet shapes as

$$\gamma_{l,b}^{(\rho)}(r) = \frac{\langle \rho_{l,b}(\epsilon = 70\%) \rangle - \langle \rho_{l,b}(\epsilon = 57\%) \rangle}{\langle \rho_{l,b}(\epsilon = 57\%) \rangle} \tag{4.10}$$

is shown in figure 4.23 for *b*-jets.

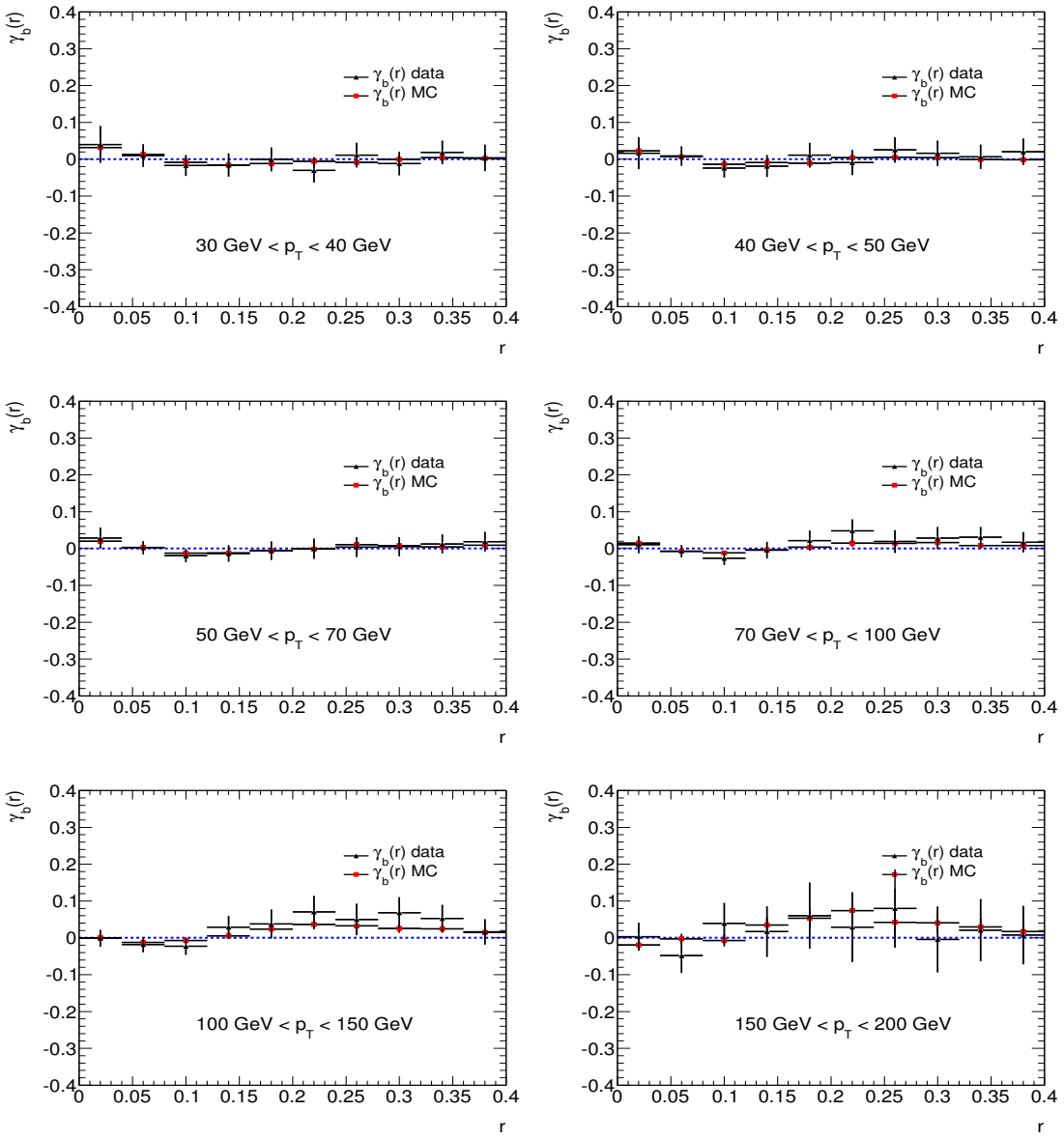

Figure 4.23: Relative differences on differential b-jet shapes due to the variation of the b-tagger working point

As one can see, the values of $\gamma(r)$ are compatible with 0, mod. the statistical error. Analogously, one can study the differences due to this effect in the integrated jet shapes by defining $\gamma^{(\Psi)}(r)$ in an equivalent way. Figure 4.24 shows this relative difference for b-jets.

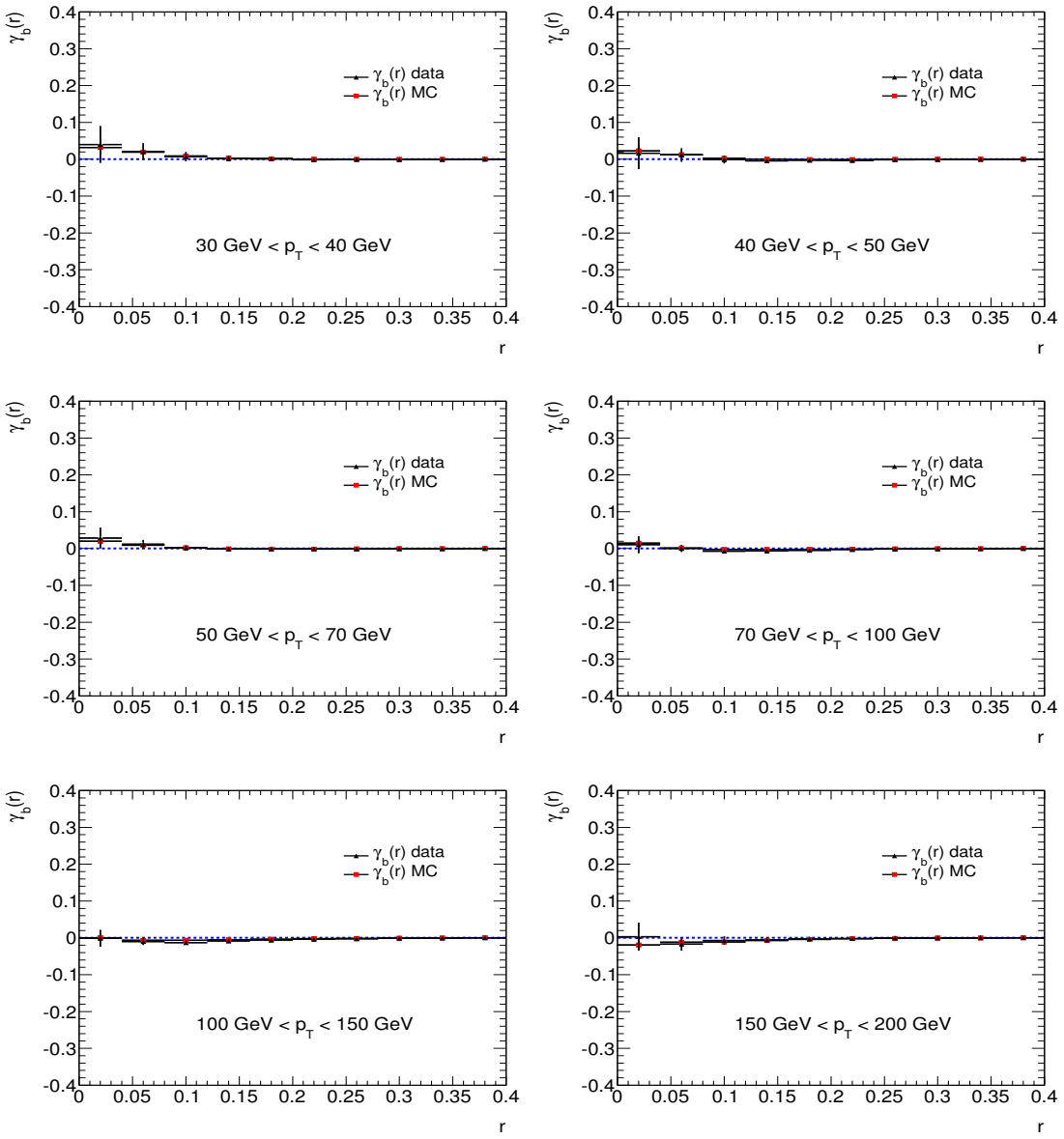

Figure 4.24: Relative differences on integrated b-jet shapes due to the variation of the b-tagger working point

As for the differential case, the values of $\gamma^{(\Psi)}(r)$ are compatible with 0 in each and every bin, mod. the statistical error. This effect is taken into account by the unfolding, as the effects due to purity are corrected as a detector effect.

4.8.3.2 Light jets

For light jets, the bias associated to the purity comes from two sources, either b or gluon jets. The first is due to inefficiencies in the anti-tagging requirement, while gluon jets are produced from initial and final state radiation. A simple way to estimate the uncertainty associated to the light jet purity is to compare the light jet shapes obtained in the sample with one b-tagged jet, with those obtained when demanding two b-tagged jets, which increases the purity of the light jet sample to 82.5%. Figure 4.25 shows the relative difference $\beta(r) = \frac{\Psi_l^{(p_2)}(r) - \Psi_l^{(p_1)}(r)}{\Psi_l^{(p_1)}(r)}$

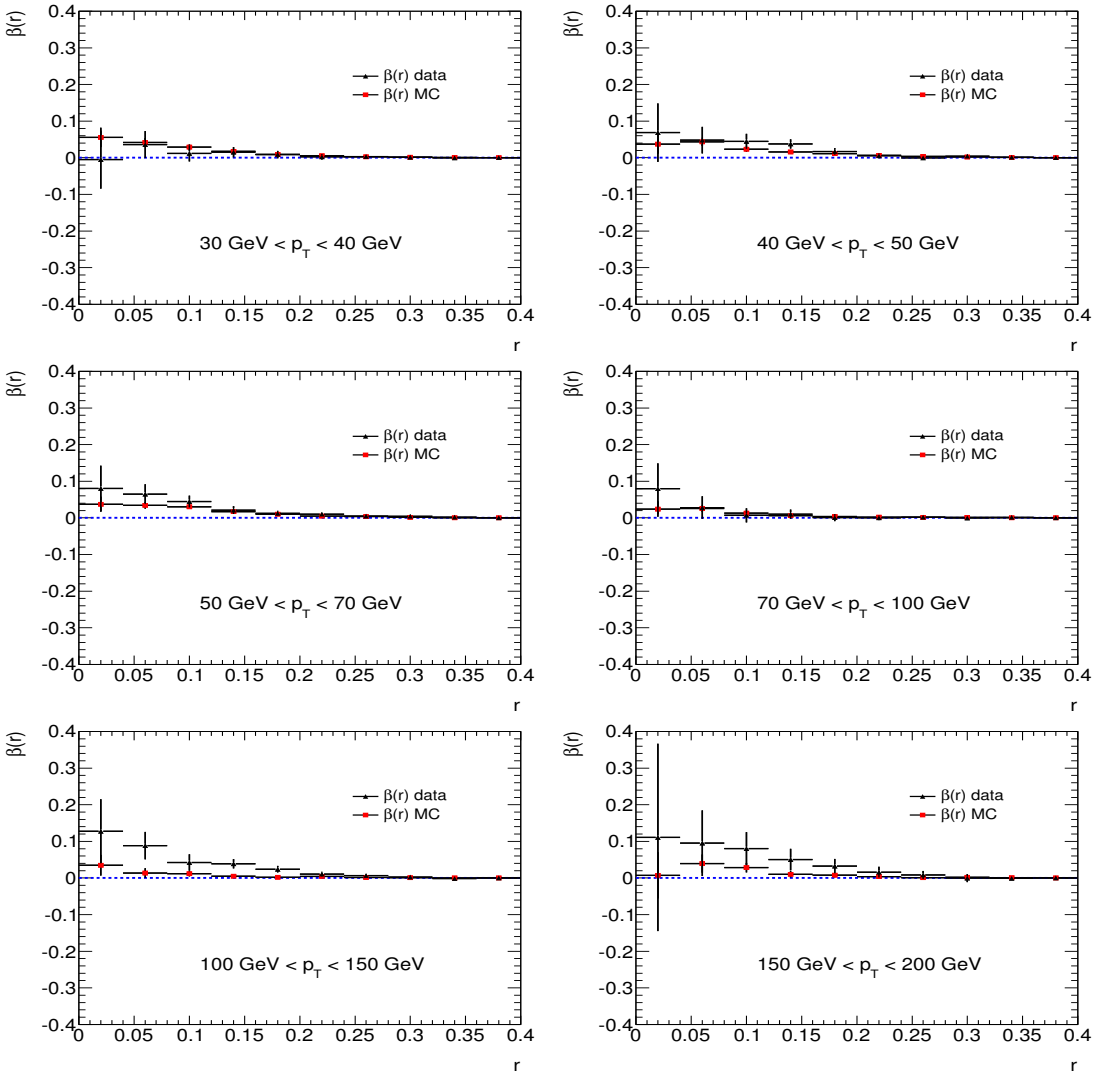

Figure 4.25: Comparison between light jet samples with one and with two b-tagged jets

It can be seen that in all cases $\beta(r) > 0$, which is consistent with the fact that as the light jet purity is increased, the shape is increased too. Unfortunately demanding two b-tagged jets reduces the statistics by roughly a factor of four and $\beta(r)$ shows large statistical errors at small r values. Analogously to the case of b-jets, the differences in the light jet shapes coming from the variation of the b-tagger working point can be estimated. Figure 4.26 and 4.27 show the relative variations (eq. 4.4) for the differential and integrated jet shapes due to this effect.

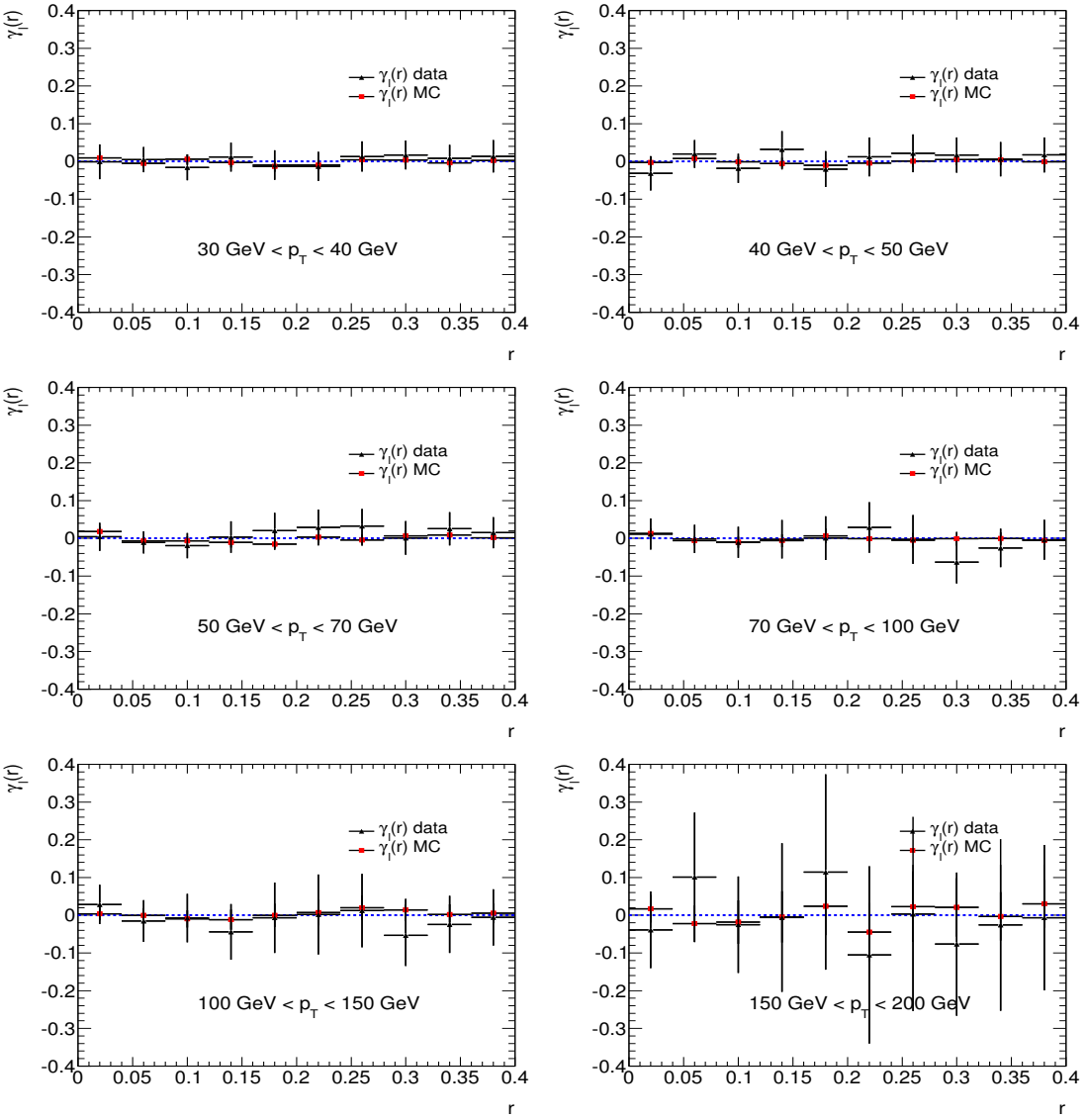

Figure 4.26: Relative differences for differential light jet shapes due to the variation of the b-tagger working point

Similarly, the relative differences for Ψ are shown below

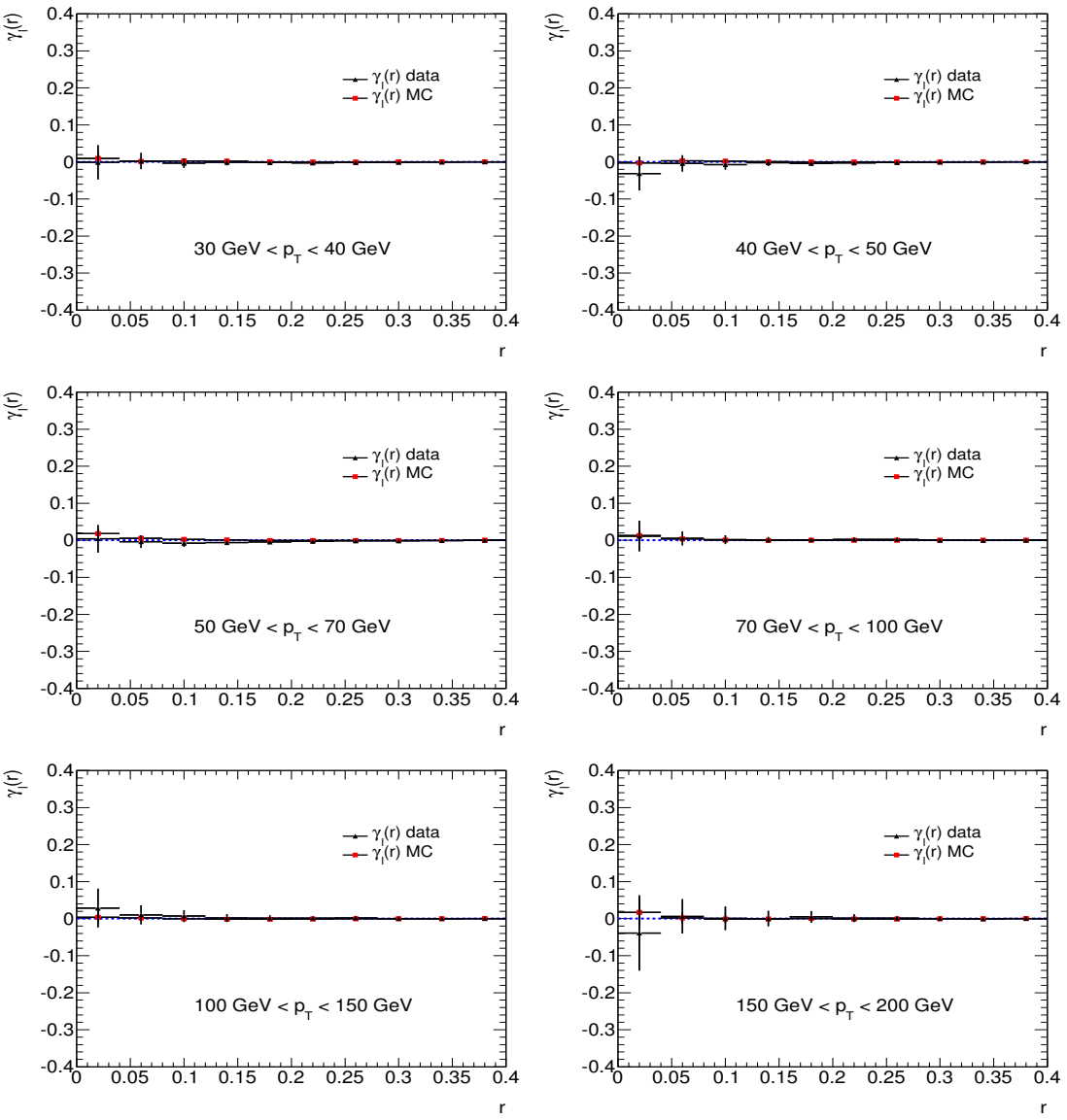

Figure 4.27: Relative differences for integrated light jet shapes due to the variation of the b-tagger working point

4.8.4 Bias due to the cut in ΔR

In this section, the possible bias due to the cut $\Delta R > 0.8$, between the selected jets and any other jet, is studied. To do this, the analysis is repeated using a slightly different cut, $\Delta R > 1.0$, and check the effect of this change. As done before, the

relative difference between both, defined as $\lambda_{l,b}(r) = \frac{\langle\Psi_{l,b}(\Delta R_m{=}1.0)\rangle - \langle\Psi_{l,b}(\Delta R_m{=}0.8)\rangle}{\langle\Psi_{l,b}(\Delta R_m{=}0.8)\rangle}$, is shown in figure 4.28 for b-jets.

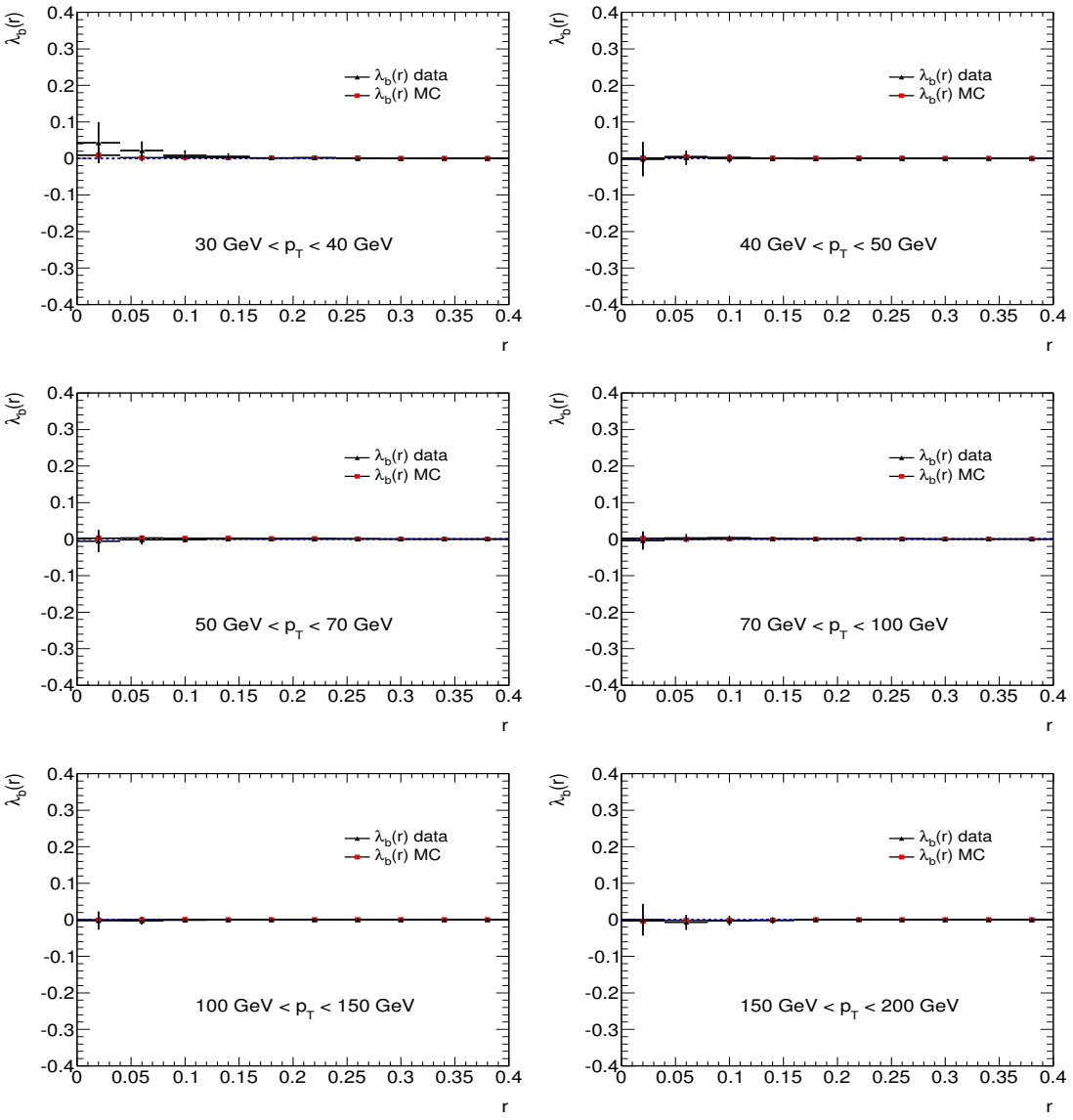

Figure 4.28: Relative differences on b-jet shapes due to the variation of the ΔR cut between jets

In the same manner, the results for light jets are shown in figure 4.29

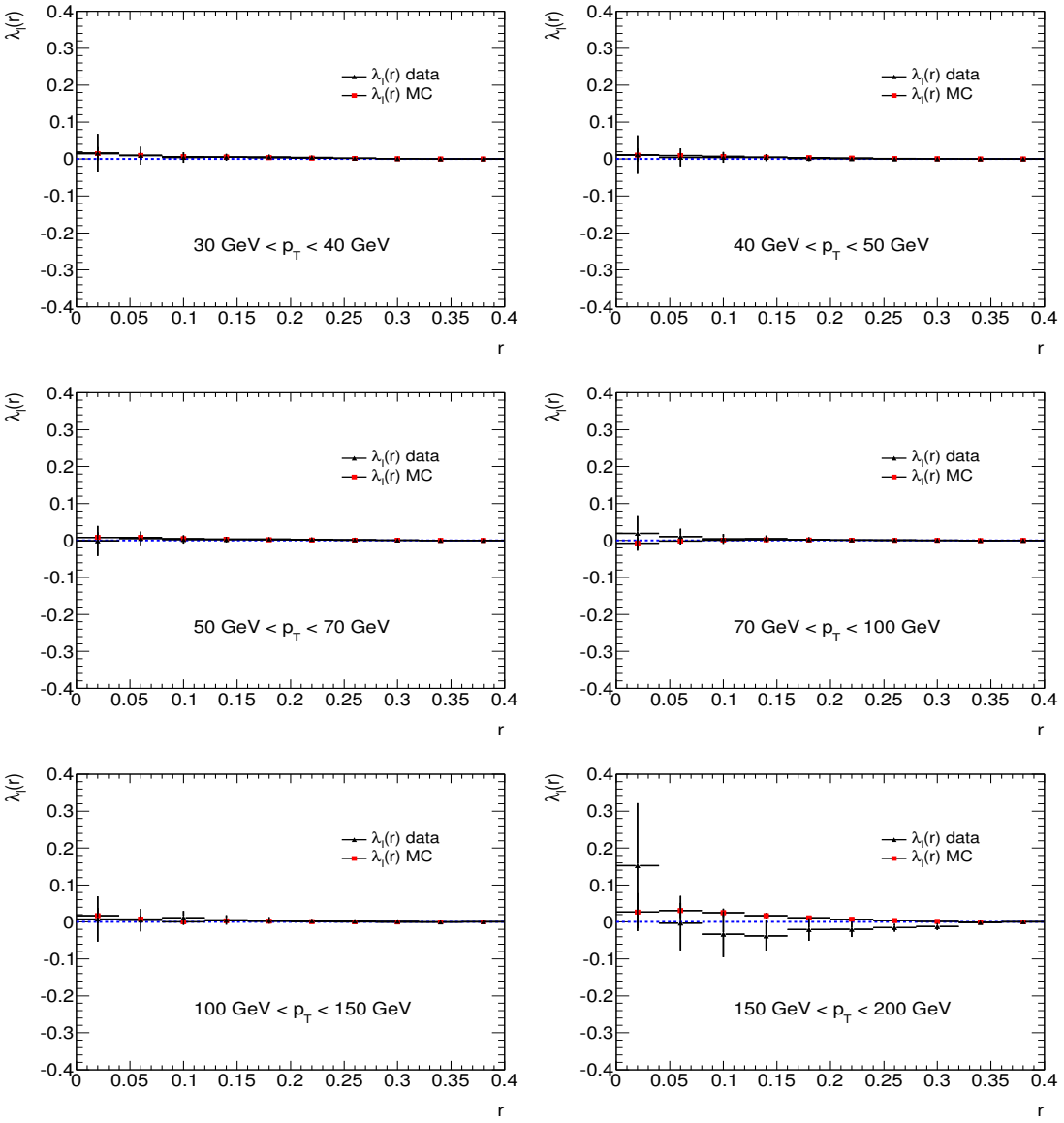

Figure 4.29: Relative differences on light jet shapes due to the variation of the ΔR cut between jets

4.8.5 In-time pileup dependence

In order to estimate the dependence of $\Psi(r)$ with the number of in-time pileup interactions, the mean value of the variable $1 - \Psi(r)$ is plotted for $r = 0.04$

and $r = 0.20$, for both the light and b-jet samples as a function of the number of primary vertices. The result is that jets get wider as the number of primary vertices increases, due to the fact that pileup introduces a constant energy density term in the $\{\eta, \varphi\}$ plane. Figures 4.30 and 4.31 show the dependence of $1-\Psi(r = 0.04)$ and $1-\Psi(r = 0.20)$ with the number of primary vertices, respectively.

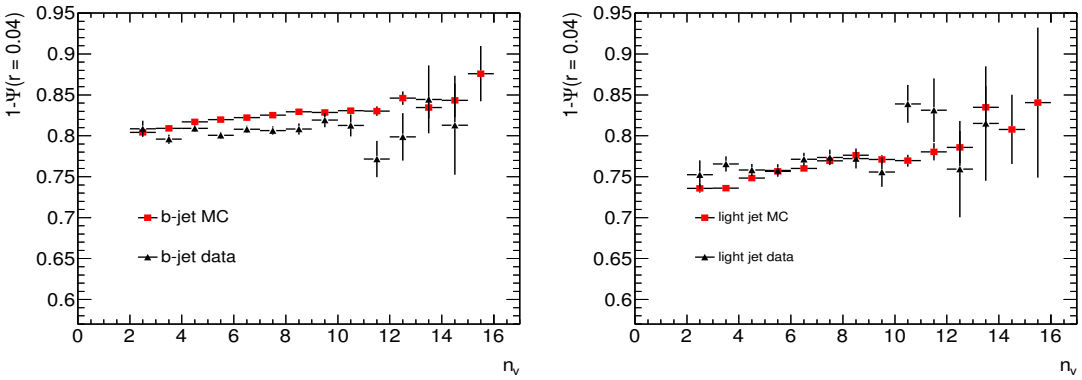

Figure 4.30: Shape ($r = 0.04$) dependence with the number of primary vertices along with MC expectations

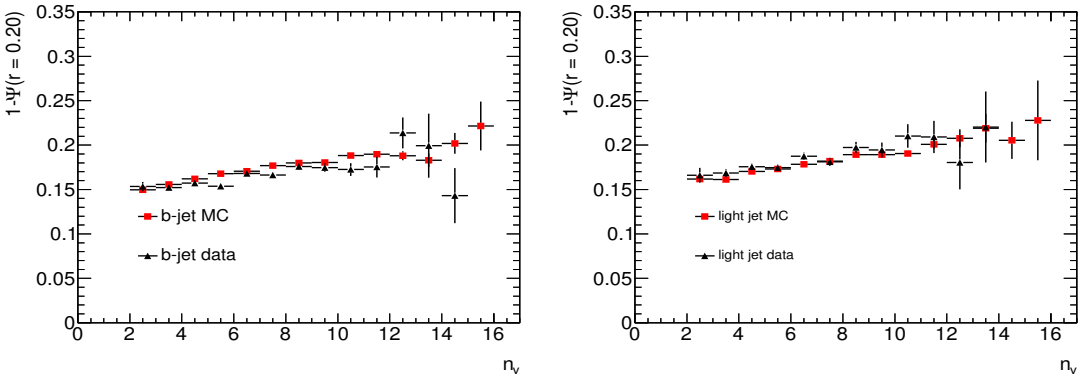

Figure 4.31: Shape ($r = 0.20$) dependence with the number of primary vertices along with MC expectations

This is fairly well reproduced by the Monte Carlo expectations. Notice that the dependence on the number of primary vertices is rather small for small values of r, and very similar for both light and b-jets, thus not affecting the main conclusions of this study.

4.8.6 Bias induced by JVF cuts

The bias induced by the cut $|\text{JVF}| > 0.75$ is estimated in this section. Plots in figure 4.32 show the dependence of $1 - \Psi(r = 0.20)$ with the JVF variable, which is weak and the same for light and b-jets.

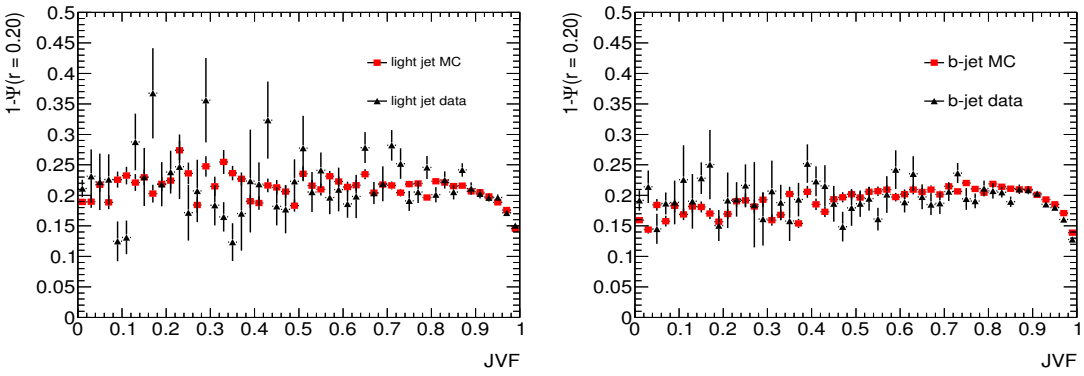

Figure 4.32: Shape ($r = 0.20$) dependence with the jet JVF for light and b-jets along with MC expectations

To estimate the JVF uncertainty, the JVF cut is switched ON and OFF and the differences are studied. They are not greater than 5%.

4.8.7 Cluster Energy Scale

Jet shapes are calculated using locally calibrated TopoClusters associated to jets. These clusters have an uncertainty in both energy and transverse momentum which depends on the transverse momentum. To take this into account [20] the transverse momentum has been varied up and down using the formula

$$p'_{\mathrm{T}} = p_{\mathrm{T}} \left[1 \pm 0.05 \left(1 + \frac{1.5 \text{ GeV}}{p_{\mathrm{T}}} \right) \right] \tag{4.11}$$

For illustrative reasons, figure 4.33 (left) shows the shifted and unshifted transverse momentum of the individual cluster constituents of b-jets. The impact of these shifts on the measured jet shapes varies from 2% to 10% as one approaches the edge of the jets.

4.8.8 Cluster Angular Resolution

The angular coordinates η and φ of individual clusters have been individually smeared using a gaussian with an RMS of 5 mrad, according to the studies performed in [93] using track to cluster extrapolation.

$$\eta' = \eta + N(0, 0.005) \tag{4.12}$$

$$\varphi' = \varphi + N(0, 0.005) \tag{4.13}$$

The impact on the jet shapes has been evaluated, and the influence has been found to be smaller than 1% in most cases. In figure 4.33 (right) shows the gaussian used for the smearing.

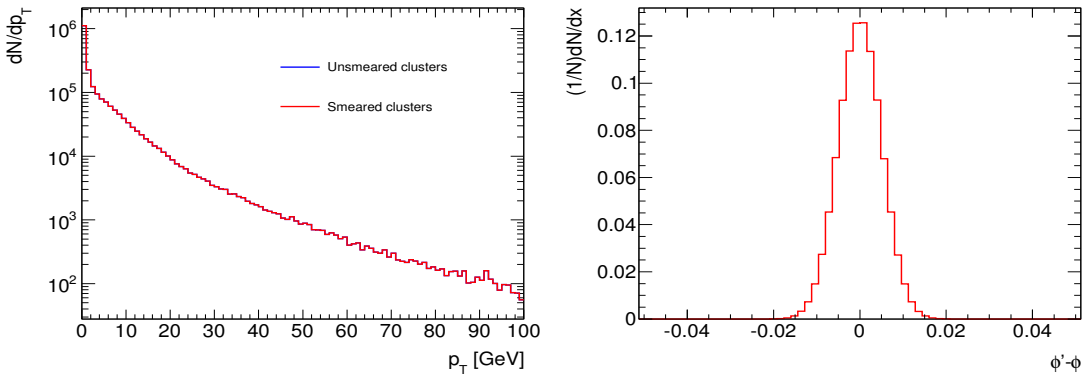

Figure 4.33: p_T distributions of the individual clusters before and after the scaling (left) and gaussian used for angular smearing (right)

4.8.9 Detector Dead Material

The uncertainty associated to the probability of energy losses due to dead material in the calorimeter is taken into account by discarding a fraction of low energy clusters ($E < 2.5$ GeV) using the algorithm described below as a result of the studies carried out in [94]. Clusters inside the jets are discarded if the condition

$$r \leq \mathcal{P}(E = 0) \times e^{-2E} \tag{4.14}$$

is fulfilled, where $r \in [0, 1]$ is a random number satisfying a uniform distribution, $\mathcal{P}(E = 0)$ is the measured probability (28%) of a particle leaving a track and zero energy in the calorimeter, and E is the cluster energy. As a result, approximately 6% of the total number of clusters is dropped from the reconstruction. The impact of this cluster-dropping algorithm in the measured jet shapes is smaller than 2%, and increases as r is increased, being more important for low energy clusters close to the edge of the jets. This uncertainty, added in quadrature to the cluster energy scale and angular resolution is presented as the cluster reconstruction uncertainty in further discussions.

4.9 Unfolding method

In order to correct the data for acceptance and detector effects, thus enabling comparisons with different models, two alternative correction procedures are followed: a bin-by-bin unfolding for the average differential and integrated jet shapes i.e. $\langle\rho(r)\rangle$ and $\langle\Psi(r)\rangle$, and a Bayesian approach where the $\rho(r)$ and $\Psi(r)$ distributions themselves are unfolded, and from these unfolded distributions their average values are determined.

In the first approach, correction factors $F(r)$ are calculated separately for differential, $\langle\rho(r)\rangle$, and integrated, $\langle\Psi(r)\rangle$, jet shapes in both light- and b-jet samples. They are defined as the ratio between the particle and detector level quantities as described by the MC generators discussed in section 4.2, i.e.

$$F^{\rho}_{l,b}(r) = \frac{\langle\rho(r)_{l,b}\rangle_{\mathrm{MC,part}}}{\langle\rho(r)_{l,b}\rangle_{\mathrm{MC,det}}} \tag{4.15}$$

$$F^{\Psi}_{l,b}(r) = \frac{\langle\Psi(r)_{l,b}\rangle_{\mathrm{MC,part}}}{\langle\Psi(r)_{l,b}\rangle_{\mathrm{MC,det}}} \tag{4.16}$$

The particle level values are calculated using particle level jets with the kinematic requirements described in Sec. 4.4. These jets are built using all particles with a lifetime above $10^{-11}s$, excluding muons and neutrinos. For particle level b-jets, a B-hadron with $p_{\mathrm{T}} > 5$ GeV is required to be closer than $\Delta R < 0.3$ from the jet axis, while for light-jets, an equivalent selection to that of the detector level jets is applied, selecting the non-b-jet pair with closest invariant mass to m_W.

4.9.1 Correction factors

The correction factors for the differential b and light-jet shapes as a function of the internal radius r are shown in Fig. 4.34 and 4.35 respectively. They depend mildly on the fragmentation model used.

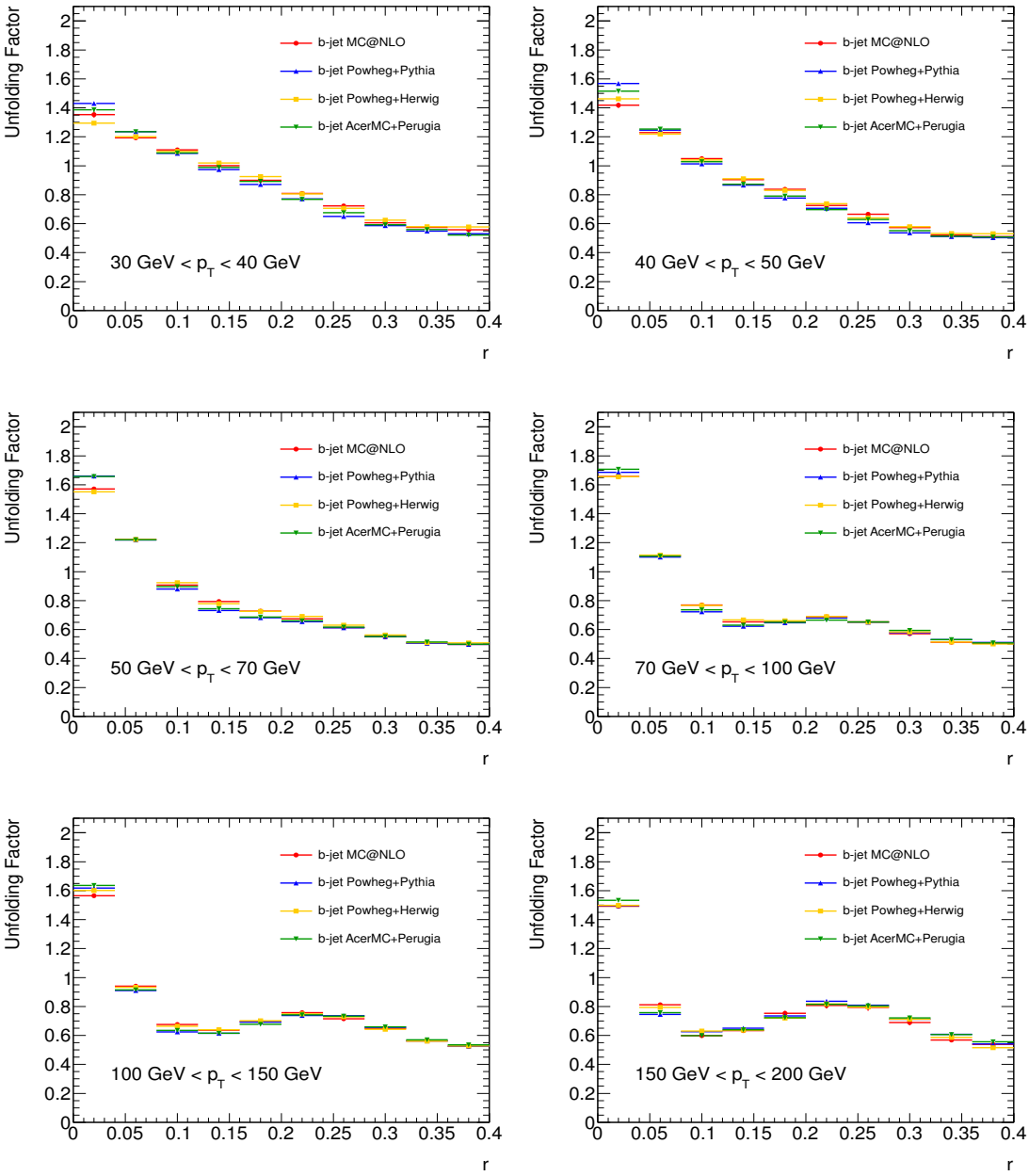

Figure 4.34: Correction factors for differential jet shapes in the b-jet sample

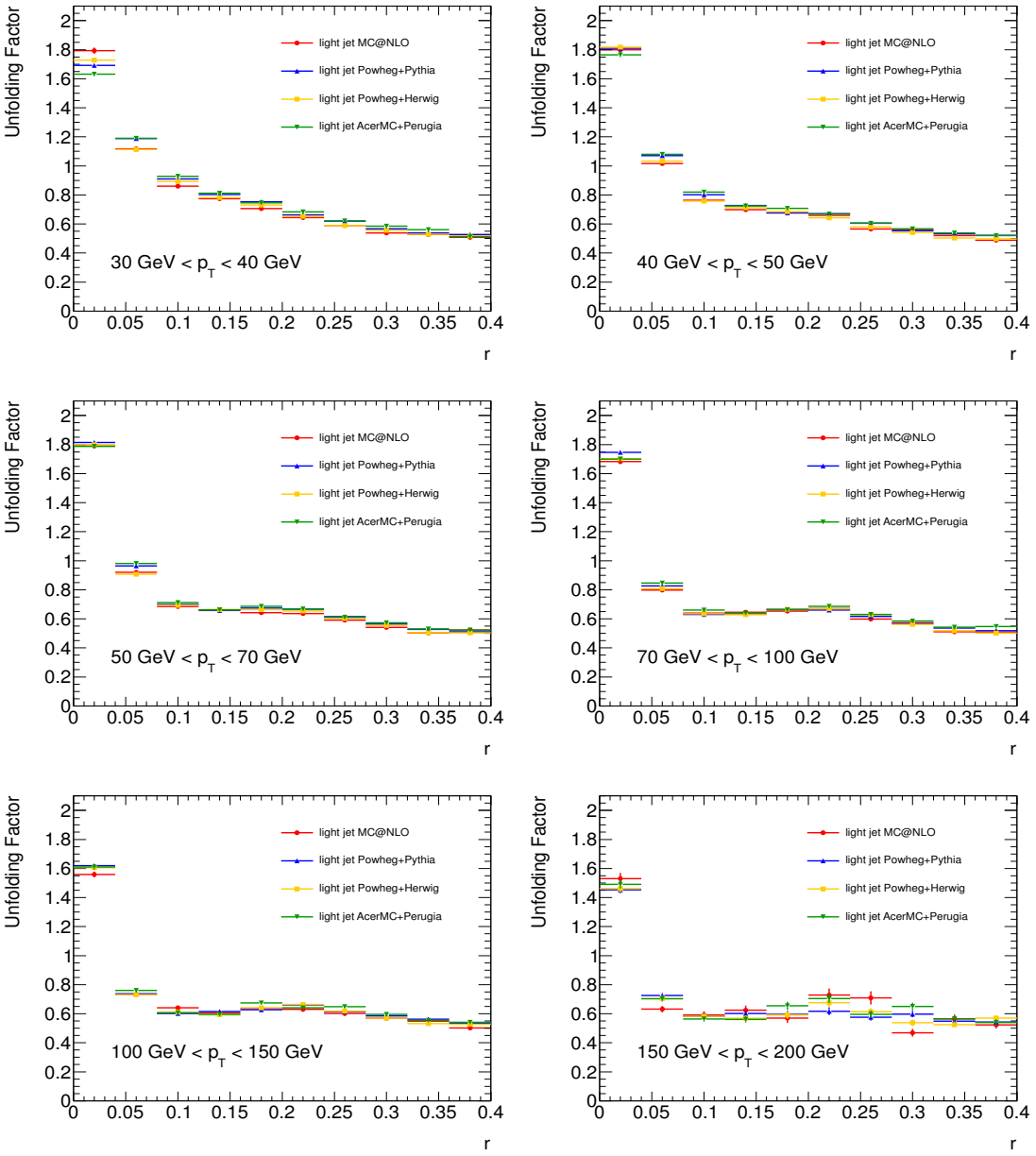

Figure 4.35: Correction factors for differential jet shapes in the light-jet sample

In a similar manner, the correction factors for the integrated jet shapes in the b and light-jet samples are shown here. In this case, they are defined as the ratio between the average values of $\Psi(r)$ at hadron level and detector level. Figures 4.36 and 4.37 show these coefficients for the b and light-jet sample respectively.

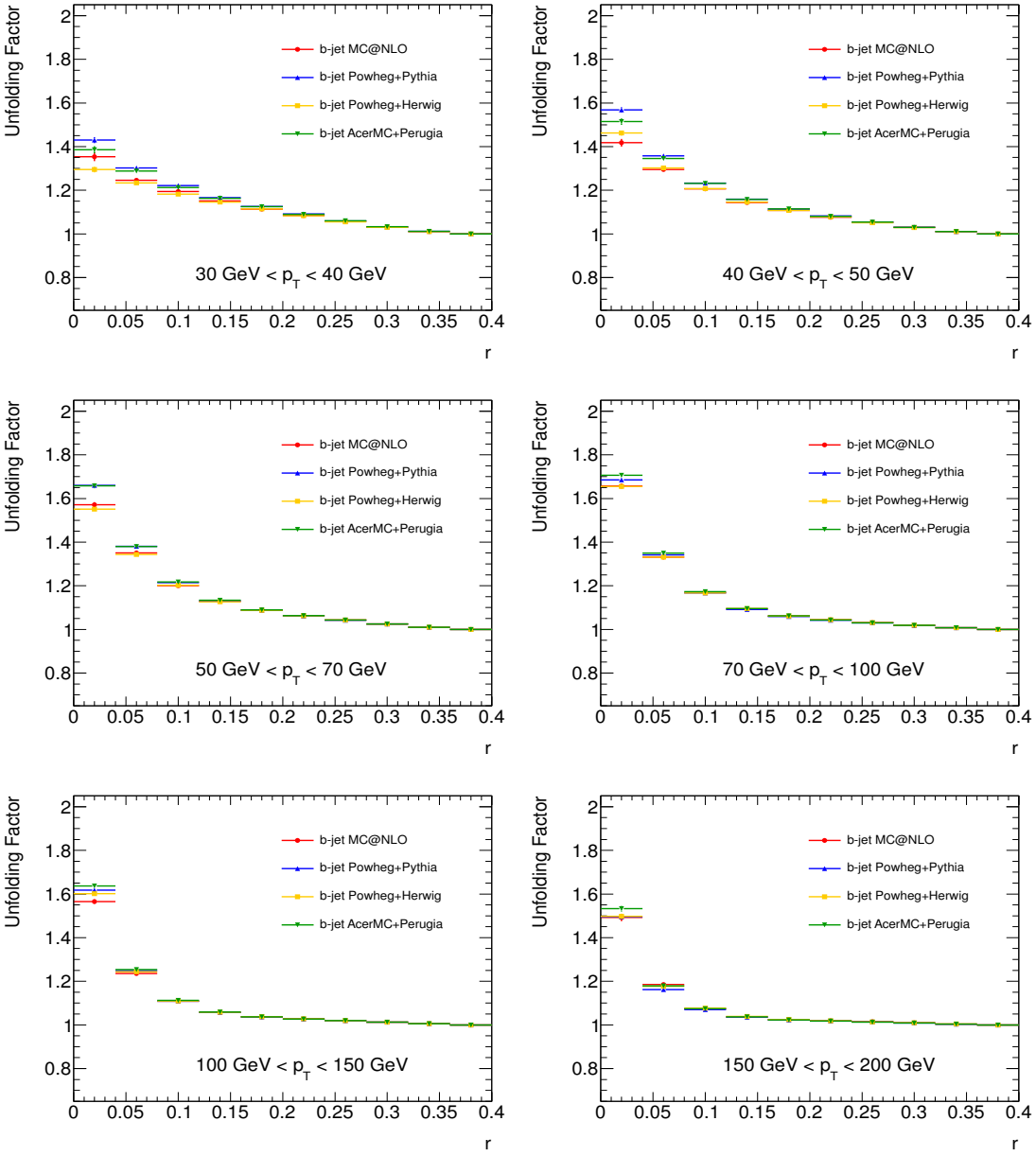

Figure 4.36: Correction factors for integrated jet shapes in the b-jet sample

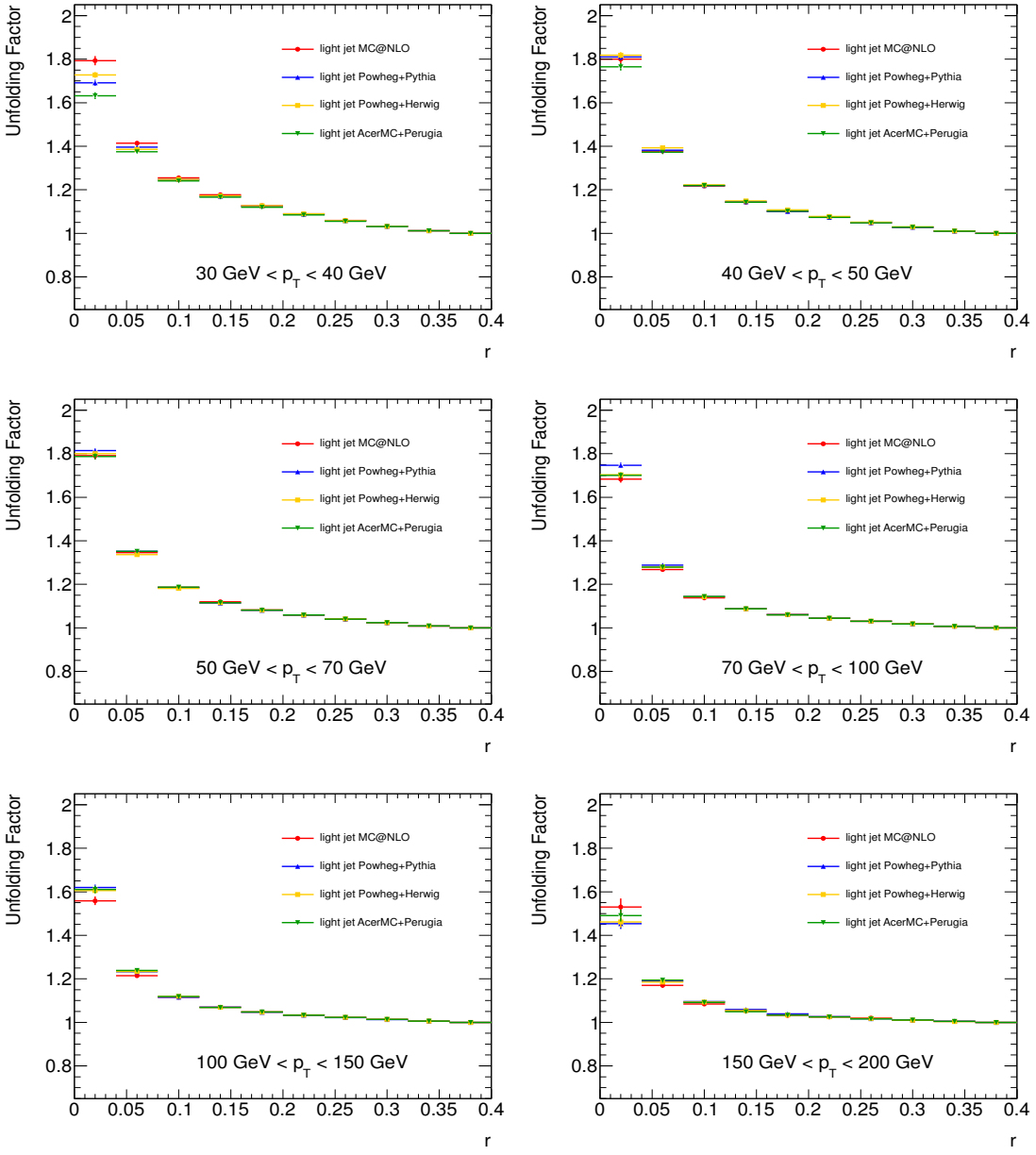

Figure 4.37: Correction factors for integrated jet shapes in the light jet sample

4.9.2 Unfolding of uncertainties

The main systematic uncertainties studied in the previous section are now propagated to the hadron level distributions via the unfolding procedure. This is done by shifting the nominal Monte Carlo distributions at the detector level by one standard deviation σ_i for each source of uncertainty and correcting the shifted distributions using the same factors used to correct the detector data. This is done for each source separately, except for the bias due to the purity, as this is already taken into account by the unfolding method. An additional source of systematic uncertainty for the bin-by-bin unfolding procedure arises from the different hadronisation models used for the calculation of the correction factors.

4.9.3 Total uncertainties

Every source of systematic uncertainty discussed in this and the preceding sections, is then unfolded and added in quadrature to obtain the total systematic uncertainty at hadron level. Figure 4.38 shows the main sources of systematic uncertainties for b-jet differential shapes, while figure 4.39 shows the analogous for light jets. The relative systematic uncertainties for differential jet shapes in the b-jet sample are summarized in tables 4.15 to 4.20 for each p_{T} bin. Similarly, the uncertainties for the differential shapes in the light jet sample are shown in tables 4.21 to 4.26

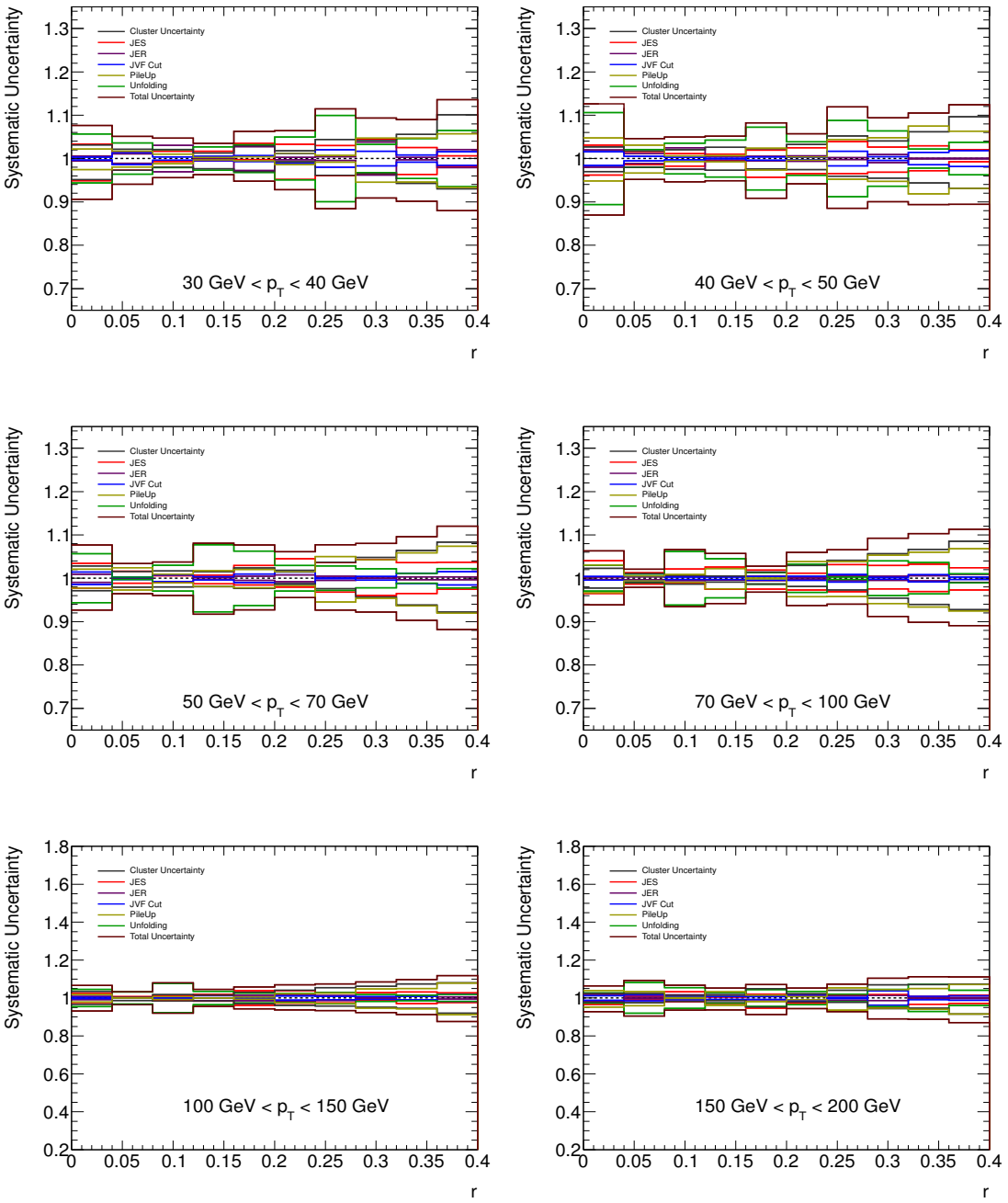

Figure 4.38: Main sources of systematic uncertainties for $\rho_b(r)$

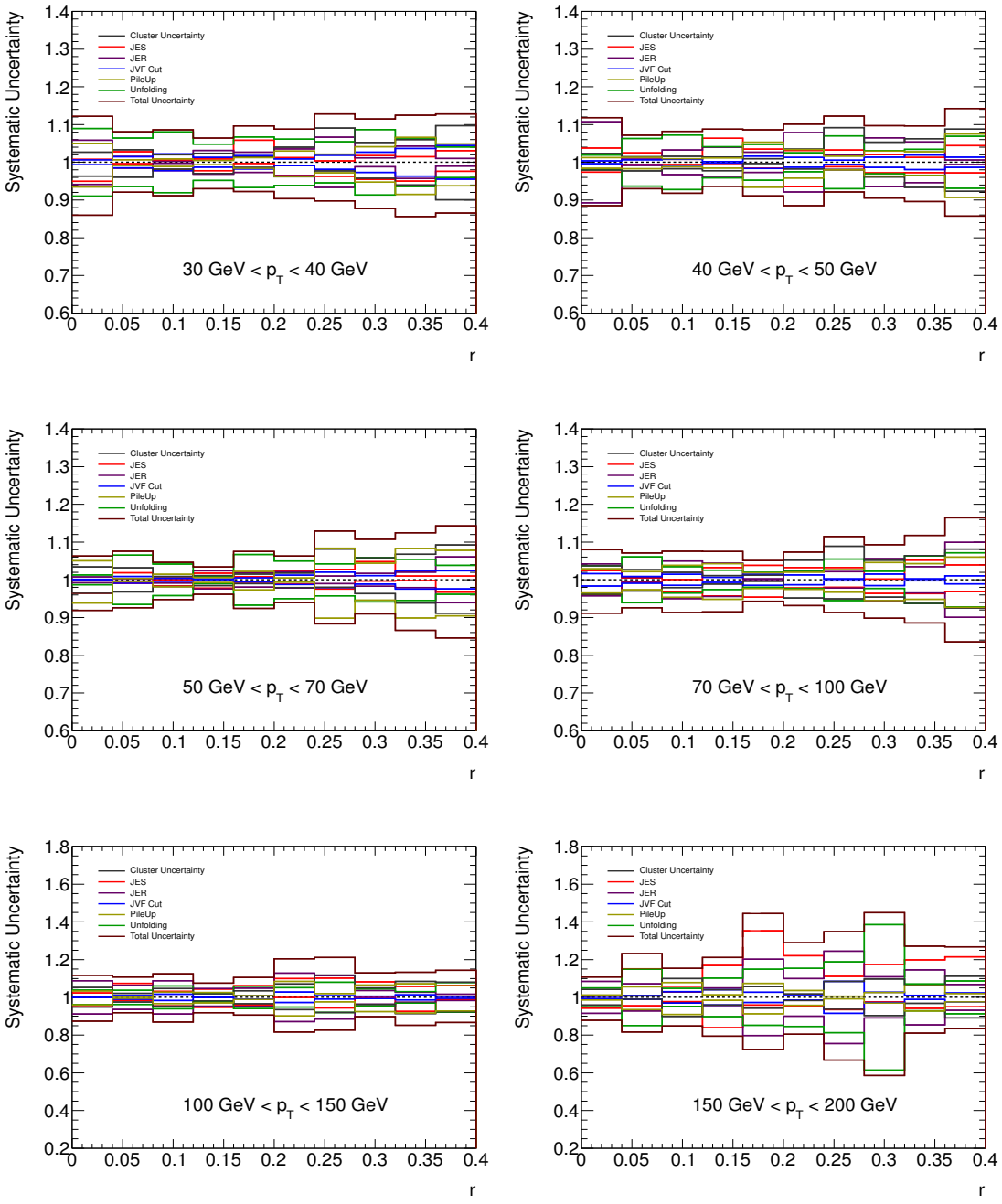

Figure 4.39: Main sources of systematic uncertainties for $\rho_l(r)$

The systematic uncertainties for the integrated jet shapes have also been calculated. The relative values are shown in figures 4.40 and 4.41 for b and light-jets, respectively.

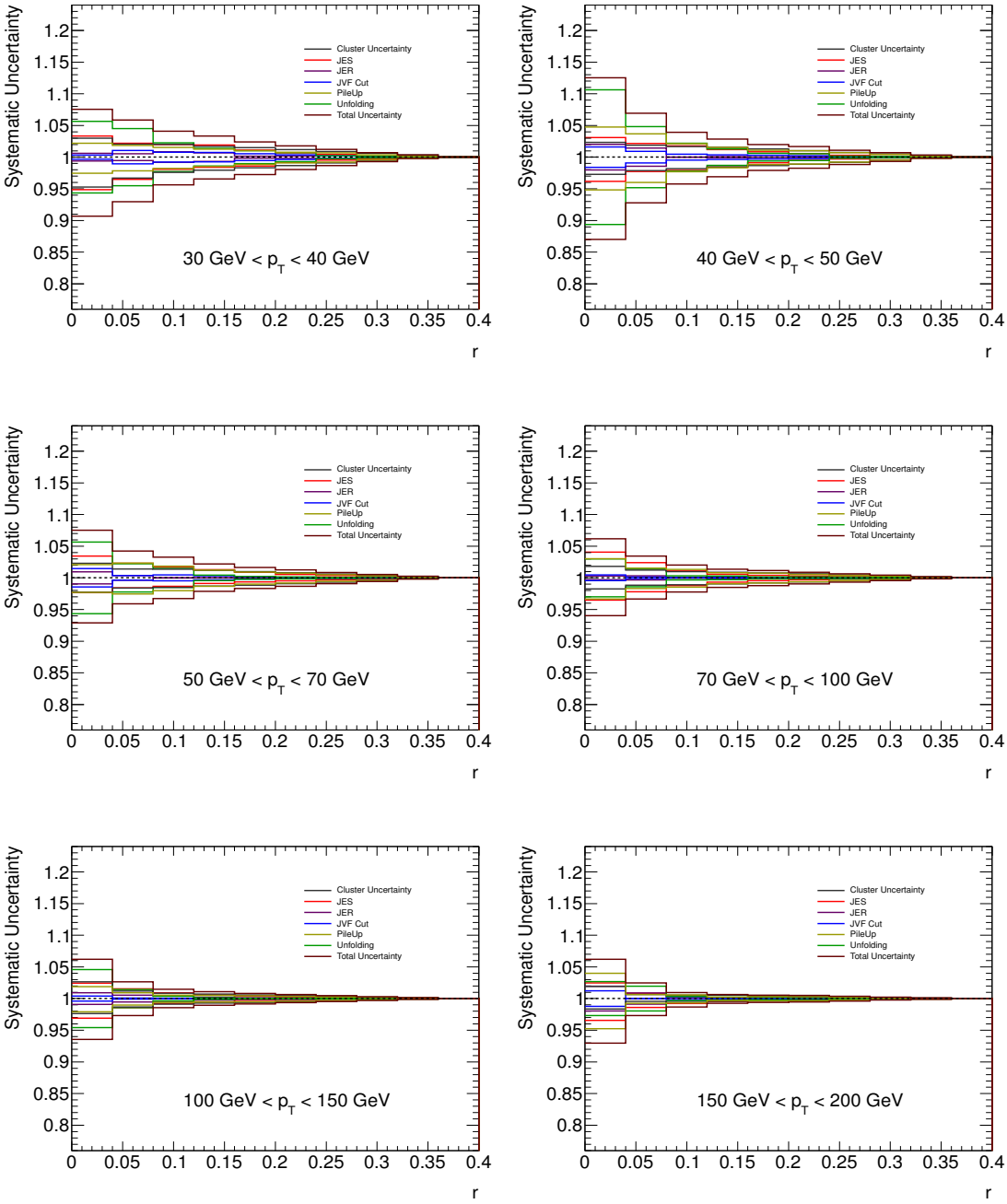

Figure 4.40: Main sources of systematic uncertainties for b-jets

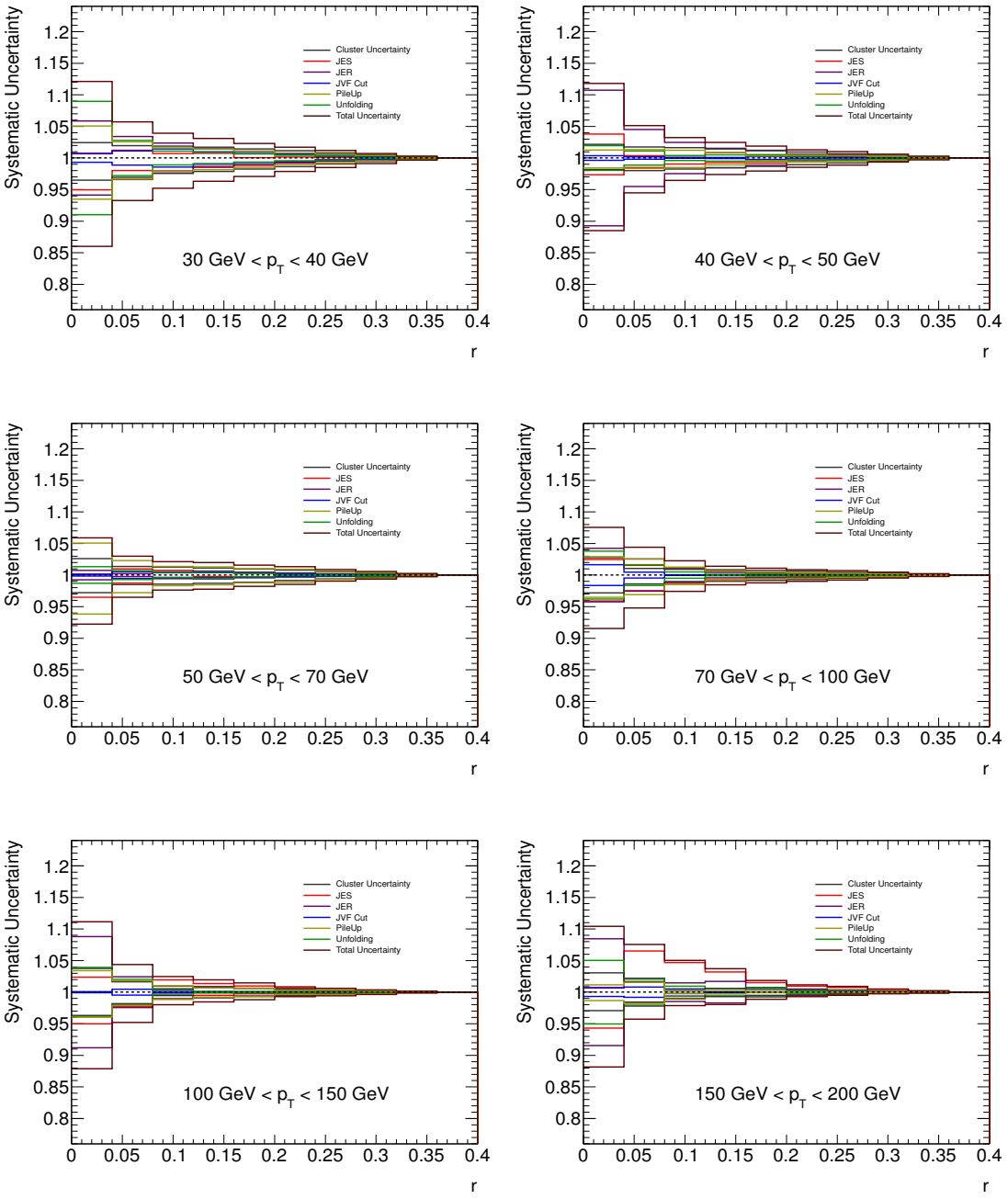

Figure 4.41: Main sources of systematic uncertainties for light jets

These plots show that the dominant sources of uncertainty are the jet energy scale and resolution. The latter uncertainties could go beyond 5% at large transverse momenta. In any case, these systematic uncertainties are not large enough so

as to preempt the observation that b-jets are wider than light jets over a large kinematical range. The numerical values for the relative systematic uncertainties on the integrated jet shapes for the b-jet sample are shown in tables 4.27 to 4.32. Finally, the systematic uncertainties on the integrated jet shapes for the light jet sample are shown in tables 4.33 to 4.38.

b-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 30 GeV $< p_{\mathrm{T}} <$ 40 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.98	4.75	3.34	5.12	0.59	0.33	2.19	2.53	5.65	7.56	9.35
0.08	2.19	2.76	1.60	2.72	1.08	1.41	1.70	1.96	3.62	5.14	5.92
0.12	2.11	1.97	1.67	0.70	3.06	0.37	1.04	1.20	2.09	4.71	4.44
0.16	1.19	2.31	1.66	0.54	0.09	0.51	0.29	0.33	2.71	3.45	3.66
0.20	2.11	2.20	3.48	0.22	2.75	0.68	0.45	0.52	3.18	5.91	4.83
0.24	1.78	0.70	3.30	4.81	0.29	1.11	1.20	1.39	4.94	6.42	7.16
0.28	3.92	3.29	3.00	3.97	0.42	2.02	0.73	0.84	9.95	11.32	11.43
0.32	4.41	3.73	4.46	3.59	3.81	1.69	4.71	5.46	3.25	9.46	9.19
0.36	6.15	6.30	2.57	3.66	0.27	0.81	4.70	5.43	4.60	9.40	10.22
0.40	10.25	7.19	0.64	1.79	2.11	1.59	5.72	6.63	6.50	13.69	12.17

Table 4.15: Systematic uncertainties on $\langle\rho(r)\rangle$ for b-jets in both channels for 30 GeV $< p_T <$ 40 GeV

b-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 40 GeV $< p_{\mathrm{T}} <$ 50 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.33	2.71	3.08	3.81	2.01	1.61	4.77	5.18	10.63	12.54	12.98
0.08	1.80	2.03	1.66	1.44	1.13	0.55	3.08	3.35	2.07	4.62	4.83
0.12	1.81	2.31	1.13	1.76	2.48	0.30	0.62	0.68	3.56	4.88	5.28
0.16	2.00	2.11	1.04	0.43	0.51	0.06	0.69	0.75	4.25	4.89	4.85
0.20	1.76	1.90	1.93	4.29	0.54	0.39	2.41	2.62	7.23	8.09	9.04
0.24	2.75	1.86	2.51	3.48	0.63	0.28	0.36	0.39	3.86	5.42	5.58
0.28	4.68	3.36	3.90	3.48	0.24	1.67	4.40	4.78	8.79	11.68	11.24
0.32	4.56	4.99	2.65	3.13	0.95	0.44	4.83	5.26	6.37	9.63	10.20
0.36	6.55	6.03	2.96	2.78	0.06	1.42	7.49	8.15	2.17	10.70	10.83
0.40	9.81	7.02	2.02	0.75	0.06	1.82	6.33	6.88	3.75	12.56	10.70

Table 4.16: Systematic uncertainties on $\langle\rho(r)\rangle$ for b-jets in both channels for 40 GeV $< p_T <$ 50 GeV

b-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 50 GeV $< p_{\mathrm{T}} <$ 70 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.29	2.27	3.44	2.33	0.94	1.45	2.09	2.28	5.65	7.50	7.12
0.08	1.50	1.94	1.65	1.12	0.04	0.30	2.46	2.68	0.29	3.35	3.52
0.12	1.45	1.80	0.58	0.91	0.78	0.71	0.95	1.03	2.97	3.65	3.89
0.16	1.15	1.67	0.72	1.29	0.23	0.37	1.73	1.89	7.72	8.04	8.23
0.20	1.81	1.72	2.94	1.52	0.50	0.97	2.02	2.21	6.28	7.52	7.13
0.24	1.94	1.37	4.53	2.05	0.76	1.46	1.51	1.65	2.93	6.15	4.49
0.28	3.84	2.59	3.60	3.19	0.47	0.09	5.01	5.48	2.82	7.81	7.42
0.32	5.19	4.71	4.30	3.96	0.05	0.43	4.23	4.63	2.19	8.27	8.01
0.36	6.83	6.59	3.62	3.51	0.32	1.19	5.86	6.40	1.14	9.84	9.97
0.40	9.01	8.52	3.71	2.54	0.32	1.52	7.37	8.05	2.25	12.52	12.30

Table 4.17: Systematic uncertainties on $\langle\rho(r)\rangle$ for b-jets in both channels for 50 GeV $< p_T <$ 70 GeV

b-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 70 GeV $< p_{\mathrm{T}} <$ 100 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	1.82	1.76	4.06	3.50	0.28	0.45	3.03	3.33	2.98	6.17	5.96
0.08	1.07	0.98	1.30	1.32	0.35	0.53	0.51	0.56	0.99	2.12	2.10
0.12	0.80	1.42	2.12	1.19	0.55	0.25	0.87	0.96	6.15	6.64	6.53
0.16	0.48	0.83	2.61	2.55	0.41	0.26	2.24	2.46	4.53	5.73	5.84
0.20	1.09	1.02	1.87	2.48	0.52	0.44	0.10	0.10	1.33	2.63	3.07
0.24	2.70	1.48	1.20	2.68	0.15	0.49	3.88	4.26	3.24	5.88	6.19
0.28	4.27	2.84	3.15	3.09	0.51	0.88	3.83	4.21	0.30	6.63	6.04
0.32	5.73	4.64	2.95	2.45	0.46	0.24	5.31	5.84	3.99	9.27	8.83
0.36	7.02	6.38	3.26	3.05	0.58	0.82	6.02	6.62	3.63	10.50	10.39
0.40	9.05	7.86	2.45	2.70	0.93	0.28	6.85	7.53	1.01	11.70	11.31

Table 4.18: Systematic uncertainties on $\langle\rho(r)\rangle$ for b-jets in both channels for 70 GeV $< p_T <$ 100 GeV

b-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 100 GeV $< p_{\mathrm{T}} <$ 150 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.65	2.38	2.44	3.08	0.90	0.40	1.87	2.10	4.57	6.19	6.44
0.08	0.86	1.59	0.75	0.07	0.20	0.20	0.05	0.06	3.20	3.41	3.58
0.12	1.43	1.15	1.20	1.40	0.03	0.05	0.71	0.80	7.67	7.93	7.92
0.16	1.10	1.17	1.94	1.65	1.43	0.15	0.08	0.09	3.47	4.37	4.26
0.20	2.36	2.11	3.87	3.85	1.42	0.43	0.53	0.60	2.59	5.45	5.34
0.24	3.30	3.01	3.89	2.69	1.68	1.14	2.31	2.59	2.46	6.45	5.76
0.28	4.88	3.51	2.95	2.64	1.31	0.86	2.53	2.85	2.95	7.08	6.21
0.32	6.08	4.86	2.83	1.30	0.17	0.71	4.77	5.36	1.82	8.46	7.61
0.36	7.60	5.79	3.12	2.95	1.15	0.91	5.03	5.65	1.47	9.86	8.86
0.40	8.76	8.83	2.74	2.36	0.57	1.82	7.83	8.80	1.83	12.36	12.96

Table 4.19: Systematic uncertainties on $\langle\rho(r)\rangle$ for b-jets in both channels for 100 GeV $< p_T <$ 150 GeV

b-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 150 GeV $< p_{\mathrm{T}} <$ 200 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	1.90	1.63	2.49	3.43	1.94	1.24	4.00	4.76	2.70	6.19	7.05
0.08	1.70	1.85	1.14	1.09	0.40	1.50	3.45	4.11	8.07	9.14	9.43
0.12	1.40	1.81	3.24	1.87	1.72	0.84	0.24	0.29	5.46	6.79	6.35
0.16	1.73	1.51	0.49	2.80	2.56	0.74	3.26	3.88	2.74	5.33	6.31
0.20	3.95	3.89	1.96	5.41	0.84	0.79	1.58	1.88	4.40	6.53	8.28
0.24	2.13	1.86	1.56	2.74	2.04	1.48	0.72	0.86	3.56	5.15	5.54
0.28	4.11	2.27	1.63	0.17	1.17	0.24	5.36	6.38	2.01	7.33	7.17
0.32	6.35	4.78	1.77	5.17	1.28	3.70	4.46	5.31	4.54	9.96	10.66
0.36	7.17	5.50	0.80	3.31	0.09	0.89	5.01	5.96	7.09	11.32	11.31
0.40	8.15	9.22	1.50	2.83	0.33	1.23	7.09	8.45	4.03	11.70	13.50

Table 4.20: Systematic uncertainties on $\langle\rho(r)\rangle$ for b-jets in both channels for 150 GeV $< p_T <$ 200 GeV

light-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 30 GeV $< p_{\mathrm{T}} <$ 40 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.43	3.51	0.77	5.02	5.85	0.66	5.06	6.52	8.97	12.13	13.97
0.08	2.77	3.49	2.81	0.36	1.52	1.46	0.66	0.85	6.45	7.87	7.68
0.12	1.97	1.81	0.54	2.13	0.28	2.21	0.86	1.11	8.03	8.62	8.86
0.16	1.75	2.49	0.90	2.24	3.15	1.30	0.51	0.66	4.81	6.24	6.82
0.20	1.11	1.09	5.86	0.31	2.69	1.76	1.25	1.61	6.72	9.62	7.70
0.24	3.29	3.04	1.22	3.56	3.54	0.88	2.95	3.80	6.15	8.49	9.35
0.28	9.08	2.41	0.39	3.69	6.65	1.89	2.15	2.77	5.45	12.83	10.23
0.32	5.33	4.33	1.75	4.54	1.17	2.69	4.11	5.30	8.69	11.51	12.31
0.36	6.59	6.58	1.48	5.00	4.31	3.64	6.60	8.50	6.44	12.74	14.63
0.40	9.87	10.06	3.05	2.40	1.03	4.44	4.85	6.26	4.08	12.95	13.55

Table 4.21: Systematic uncertainties on $\langle\rho(r)\rangle$ for light-jets in the semileptonic channel for 30 GeV $< p_T <$ 40 GeV

light-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 40 GeV $< p_{\mathrm{T}} <$ 50 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.15	1.81	3.77	2.63	10.73	0.37	1.26	1.56	1.94	11.81	11.48
0.08	1.62	2.36	2.53	0.56	0.85	0.66	1.38	1.71	6.31	7.20	7.05
0.12	1.48	1.66	0.80	0.59	3.23	0.79	0.95	1.18	7.21	8.17	8.22
0.16	1.15	2.06	6.36	0.69	4.03	0.22	1.21	1.50	4.10	8.74	6.32
0.20	0.97	0.38	3.47	1.58	2.69	1.65	5.31	6.59	4.78	8.59	8.88
0.24	3.34	1.53	3.12	6.48	7.86	1.28	3.37	4.18	2.48	10.09	11.47
0.28	9.28	2.30	3.27	1.18	1.93	0.55	1.61	2.00	6.98	12.33	7.96
0.32	5.52	4.15	2.02	2.79	6.49	1.38	2.96	3.68	3.05	9.83	9.59
0.36	6.84	7.21	1.32	2.74	5.41	1.93	2.82	3.50	3.40	10.05	10.79
0.40	9.18	8.15	4.39	2.79	0.62	1.36	7.50	9.31	6.89	14.47	14.51

Table 4.22: Systematic uncertainties on $\langle\rho(r)\rangle$ for light-jets in the semileptonic channel for 40 GeV $< p_T <$ 50 GeV

light-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 50 GeV $< p_{\mathrm{T}} <$ 70 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.61	2.78	0.08	3.52	0.75	0.14	5.06	6.16	1.31	5.90	7.77
0.08	2.47	2.53	1.86	0.87	0.29	0.92	0.37	0.45	6.54	7.31	7.15
0.12	1.01	1.80	0.01	1.64	0.84	0.36	1.49	1.81	4.18	4.64	5.24
0.16	0.86	1.32	1.76	2.25	2.39	0.19	0.61	0.75	1.02	3.32	3.76
0.20	1.36	0.77	0.81	0.60	1.95	0.56	2.19	2.66	6.71	7.51	7.56
0.24	2.26	2.45	2.39	1.11	2.15	1.23	0.49	0.60	4.96	6.47	6.19
0.28	8.35	2.75	2.81	2.46	1.03	2.04	8.36	10.17	4.24	13.07	11.84
0.32	5.96	3.83	4.83	0.35	1.73	1.15	4.42	5.37	5.86	10.81	9.07
0.36	7.06	6.49	1.02	0.19	2.07	2.47	8.31	10.10	5.47	12.65	13.58
0.40	9.75	9.39	0.99	3.24	6.08	2.41	7.83	9.52	3.86	14.66	15.71

Table 4.23: Systematic uncertainties on $\langle\rho(r)\rangle$ for light-jets in the semileptonic channel for 50 GeV $< p_T <$ 70 GeV

light-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 70 GeV $< p_{\mathrm{T}} <$ 100 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.76	2.81	2.46	4.09	4.23	1.64	2.91	3.51	3.77	7.55	8.47
0.08	2.05	2.39	0.48	0.66	0.67	0.88	2.17	2.62	6.10	6.90	7.17
0.12	1.56	1.65	0.11	3.20	4.98	1.46	3.83	4.61	3.49	7.50	8.56
0.16	0.84	0.88	3.23	4.22	4.41	0.64	4.30	5.18	2.59	7.49	8.48
0.20	1.24	0.33	3.82	4.58	0.23	1.46	1.89	2.27	1.98	5.08	5.69
0.24	4.18	3.74	3.28	2.17	2.33	1.29	2.06	2.48	2.25	6.68	6.09
0.28	8.39	4.07	3.23	2.00	2.17	0.27	2.69	3.24	5.50	11.10	8.13
0.32	5.63	5.05	0.20	3.61	5.62	1.53	4.57	5.50	2.29	9.58	10.39
0.36	6.95	6.78	5.22	3.63	3.50	0.27	4.27	5.14	6.33	12.09	11.75
0.40	8.90	8.24	3.94	3.17	9.94	1.05	6.04	7.28	7.20	16.82	16.81

Table 4.24: Systematic uncertainties on $\langle\rho(r)\rangle$ for light-jets in the semileptonic channel for 70 GeV $< p_T <$ 100 GeV

light-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 100 GeV $< p_{\mathrm{T}} <$ 150 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
.04	3.76	3.68	2.37	5.00	8.80	0.09	3.42	3.97	3.91	11.14	12.12
0.08	1.66	2.16	7.39	1.64	6.42	1.01	0.44	0.51	3.74	10.66	7.99
0.12	3.37	3.31	3.15	4.57	8.79	1.69	2.88	3.34	6.08	12.11	12.66
0.16	1.47	1.39	4.45	5.24	4.62	0.01	2.57	2.99	2.28	7.42	8.05
0.20	2.56	2.40	6.17	4.02	4.84	0.97	0.99	1.15	5.98	10.29	9.13
0.24	5.37	4.60	9.91	0.08	12.80	2.76	8.47	9.84	5.12	19.91	17.77
0.28	10.38	5.85	10.24	5.64	11.52	0.88	2.06	2.39	8.02	20.36	16.42
0.32	5.26	4.06	8.22	0.02	3.63	0.68	6.46	7.50	4.51	13.08	10.33
0.36	7.90	8.02	5.84	7.41	2.65	1.30	7.24	8.41	3.01	12.91	14.42
0.40	8.30	8.03	6.28	1.15	1.63	0.54	6.20	7.20	7.88	14.55	13.51

Table 4.25: Systematic uncertainties on $\langle\rho(r)\rangle$ for light-jets in the semileptonic channel for 100 GeV $< p_T <$ 150 GeV

light-jet Systematic Uncertainties for $\langle\rho(r)\rangle$ (%); 150 GeV $< p_{\mathrm{T}} <$ 200 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	3.04	2.92	0.74	5.70	8.45	0.68	1.16	1.34	5.03	10.41	11.83
0.08	0.67	0.92	15.11	4.44	7.16	0.99	5.58	6.47	14.97	23.16	18.41
0.12	7.14	7.35	5.88	2.26	2.82	3.01	7.85	9.09	4.74	13.67	13.47
0.16	2.99	2.95	16.83	16.04	5.00	4.12	1.38	1.59	10.22	20.99	20.37
0.20	4.10	4.10	35.26	4.14	20.24	2.71	7.48	8.66	14.86	44.20	27.33
0.24	2.10	2.01	22.06	4.76	9.99	1.40	3.64	4.22	15.36	29.01	19.55
0.28	7.67	5.13	11.08	6.56	24.44	8.52	0.67	0.77	18.78	34.71	33.05
0.32	7.79	7.62	17.40	2.41	10.87	2.73	2.64	3.05	38.53	44.50	41.02
0.36	6.42	3.23	19.86	5.81	14.61	1.14	6.19	7.17	7.05	27.18	18.98
0.40	10.20	9.83	21.48	5.01	6.87	2.38	2.00	2.31	8.73	26.43	16.00

Table 4.26: Systematic uncertainties on $\langle\rho(r)\rangle$ for light-jets in the semileptonic channel for 150 GeV $< p_T <$ 200 GeV

b-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 30 GeV $< p_{\mathrm{T}} <$ 40 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.98	4.75	3.34	5.12	0.59	0.33	2.19	2.53	5.65	7.56	9.35
0.08	2.07	3.29	2.18	3.52	0.52	1.05	1.86	2.15	4.51	5.85	7.04
0.12	2.01	2.41	1.98	1.93	0.83	0.79	1.55	1.79	2.25	4.09	4.37
0.16	1.73	2.06	1.91	1.61	0.66	0.73	1.26	1.46	1.39	3.33	3.45
0.20	1.49	1.67	1.13	1.41	0.16	0.52	1.01	1.17	1.01	2.42	2.73
0.24	1.18	1.32	0.69	0.80	0.18	0.36	0.79	0.92	0.70	1.77	1.97
0.28	0.87	0.90	0.42	0.45	0.19	0.19	0.68	0.79	0.21	1.23	1.33
0.32	0.59	0.50	0.13	0.21	0.04	0.08	0.36	0.42	0.14	0.72	0.71
0.36	0.33	0.12	0.02	0.04	0.05	0.04	0.14	0.16	0.08	0.37	0.23
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.27: Systematic uncertainties on $\langle\Psi(r)\rangle$ for b-jets in both channels for 30 GeV $< p_T <$ 40 GeV

b-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 40 GeV $< p_{\mathrm{T}} <$ 50 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.33	2.71	3.08	3.81	2.01	1.61	4.77	5.18	10.63	12.54	12.98
0.08	1.87	2.12	2.16	2.27	1.44	0.92	3.67	4.00	4.82	6.92	7.20
0.12	1.69	1.80	1.78	2.09	0.03	0.48	2.12	2.31	2.18	3.94	4.23
0.16	1.57	1.55	1.20	1.57	0.08	0.37	1.54	1.68	1.30	2.85	3.08
0.20	1.28	1.34	0.83	0.87	0.01	0.28	1.07	1.16	0.59	1.98	2.08
0.24	1.03	1.11	0.56	0.52	0.06	0.23	1.01	1.10	0.54	1.66	1.75
0.28	0.74	0.80	0.30	0.29	0.04	0.12	0.70	0.76	0.15	1.08	1.16
0.32	0.51	0.42	0.16	0.13	0.00	0.10	0.44	0.48	0.10	0.71	0.66
0.36	0.29	0.11	0.05	0.02	0.00	0.04	0.14	0.16	0.06	0.33	0.21
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.28: Systematic uncertainties on $\langle\Psi(r)\rangle$ for b-jets in both channels for 40 GeV $< p_T <$ 50 GeV

b-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 50 GeV $< p_{\mathrm{T}} <$ 70 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.29	2.27	3.44	2.33	0.94	1.45	2.09	2.28	5.65	7.50	7.12
0.08	1.39	1.60	2.31	1.57	0.37	0.35	2.32	2.54	2.25	4.24	4.09
0.12	1.36	1.50	1.72	1.34	0.02	0.47	1.85	2.03	1.49	3.27	3.25
0.16	1.18	1.31	1.30	0.89	0.02	0.32	1.23	1.34	0.44	2.21	2.14
0.20	0.99	1.15	0.88	0.65	0.07	0.20	0.91	0.99	0.24	1.64	1.68
0.24	0.81	0.91	0.54	0.48	0.02	0.09	0.76	0.83	0.17	1.25	1.33
0.28	0.60	0.67	0.35	0.31	0.01	0.08	0.49	0.54	0.08	0.86	0.92
0.32	0.40	0.40	0.18	0.16	0.00	0.07	0.32	0.35	0.02	0.55	0.56
0.36	0.21	0.13	0.07	0.05	0.01	0.03	0.14	0.15	0.02	0.27	0.21
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.29: Systematic uncertainties on $\langle\Psi(r)\rangle$ for b-jets in both channels for 50 GeV $< p_T <$ 70 GeV

b-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 70 GeV $< p_{\mathrm{T}} <$ 100 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	1.82	1.76	4.06	3.50	0.28	0.45	3.03	3.33	2.98	6.17	5.96
0.08	1.24	1.15	2.41	2.20	0.32	0.14	1.52	1.68	1.40	3.43	3.32
0.12	1.00	1.08	1.10	1.22	0.07	0.02	1.33	1.47	0.31	2.02	2.22
0.16	0.90	0.95	0.57	0.68	0.12	0.06	0.82	0.90	0.28	1.38	1.51
0.20	0.77	0.79	0.38	0.43	0.07	0.09	0.75	0.82	0.17	1.16	1.24
0.24	0.62	0.65	0.31	0.29	0.06	0.06	0.54	0.59	0.14	0.89	0.94
0.28	0.47	0.48	0.20	0.18	0.04	0.03	0.39	0.43	0.12	0.65	0.68
0.32	0.31	0.29	0.11	0.11	0.03	0.02	0.24	0.27	0.07	0.42	0.42
0.36	0.17	0.09	0.04	0.04	0.01	0.00	0.10	0.11	0.02	0.20	0.15
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.30: Systematic uncertainties on $\langle\Psi(r)\rangle$ for b-jets in both channels for 70 GeV $< p_T <$ 100 GeV

b-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 100 GeV $< p_{\mathrm{T}} <$ 150 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.65	2.38	2.44	3.08	0.90	0.40	1.87	2.10	4.57	6.19	6.44
0.08	1.18	1.34	1.55	1.42	0.53	0.08	0.91	1.02	1.44	2.64	2.68
0.12	0.85	0.87	0.93	0.78	0.42	0.05	0.54	0.61	0.36	1.47	1.43
0.16	0.72	0.71	0.62	0.52	0.22	0.06	0.49	0.55	0.12	1.10	1.07
0.20	0.58	0.61	0.36	0.27	0.12	0.04	0.43	0.49	0.06	0.82	0.84
0.24	0.47	0.47	0.22	0.17	0.06	0.07	0.34	0.38	0.06	0.63	0.64
0.28	0.35	0.35	0.14	0.10	0.03	0.05	0.27	0.30	0.04	0.47	0.48
0.32	0.23	0.22	0.08	0.08	0.03	0.04	0.17	0.19	0.02	0.30	0.30
0.36	0.11	0.08	0.03	0.03	0.01	0.02	0.09	0.10	0.01	0.15	0.13
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.31: Systematic uncertainties on $\langle\Psi(r)\rangle$ for b-jets in both channels for 100 GeV $< p_T <$ 150 GeV

b-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 150 GeV $< p_{\mathrm{T}} <$ 200 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	1.90	1.63	2.49	3.43	1.94	1.24	4.00	4.76	2.70	6.19	7.05
0.08	0.65	0.42	0.85	1.39	0.88	0.00	0.64	0.76	1.94	2.47	2.69
0.12	0.54	0.63	0.07	0.77	0.39	0.16	0.56	0.67	0.34	0.95	1.32
0.16	0.53	0.44	0.11	0.47	0.14	0.08	0.24	0.28	0.21	0.65	0.75
0.20	0.39	0.41	0.02	0.23	0.10	0.05	0.29	0.35	0.18	0.54	0.62
0.24	0.33	0.33	0.06	0.15	0.04	0.09	0.31	0.36	0.14	0.49	0.54
0.28	0.25	0.26	0.03	0.15	0.02	0.08	0.20	0.23	0.13	0.35	0.41
0.32	0.16	0.16	0.00	0.07	0.00	0.02	0.12	0.15	0.07	0.22	0.24
0.36	0.09	0.07	0.01	0.02	0.00	0.01	0.06	0.07	0.03	0.11	0.11
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.32: Systematic uncertainties on $\langle\Psi(r)\rangle$ for b-jets in both channels for 150 GeV $< p_T <$ 200 GeV

light-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 30 GeV $< p_{\mathrm{T}} <$ 40 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.43	3.51	0.77	5.02	5.85	0.66	5.06	6.52	8.97	12.13	13.97
0.08	1.94	3.05	1.24	2.00	3.42	1.11	2.59	3.34	2.79	5.72	6.72
0.12	1.87	2.43	0.72	2.04	2.36	1.43	1.60	2.06	1.10	3.92	4.80
0.16	1.72	2.13	0.75	1.35	1.47	0.99	1.42	1.83	0.89	3.08	3.69
0.20	1.45	1.76	0.05	1.17	1.03	0.70	1.14	1.47	0.65	2.31	2.93
0.24	1.14	1.35	0.14	0.80	0.67	0.57	0.82	1.06	0.44	1.72	2.14
0.28	0.82	0.97	0.11	0.53	0.23	0.42	0.64	0.83	0.34	1.20	1.50
0.32	0.51	0.57	0.01	0.27	0.15	0.26	0.40	0.51	0.10	0.72	0.87
0.36	0.24	0.24	0.07	0.06	0.02	0.11	0.12	0.15	0.05	0.30	0.31
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.33: Systematic uncertainties on $\langle\Psi(r)\rangle$ for light-jets in the semileptonic channels for 30 GeV $< p_T <$ 40 GeV

light-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 40 GeV $< p_{\mathrm{T}} <$ 50 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.15	1.81	3.77	2.63	10.73	0.37	1.26	1.56	1.94	11.81	11.48
0.08	1.73	1.99	0.38	1.52	4.50	0.18	1.32	1.64	1.13	5.14	5.52
0.12	1.65	1.76	0.07	0.97	2.48	0.07	1.22	1.52	0.40	3.25	3.56
0.16	1.42	1.54	0.85	0.73	1.55	0.09	0.88	1.09	0.50	2.48	2.60
0.20	1.22	1.32	0.46	0.81	1.17	0.23	0.32	0.40	0.55	1.88	2.07
0.24	0.95	1.07	0.24	0.36	0.62	0.14	0.51	0.63	0.29	1.31	1.47
0.28	0.68	0.77	0.06	0.28	0.48	0.16	0.40	0.50	0.34	1.00	1.13
0.32	0.43	0.46	0.14	0.16	0.20	0.10	0.26	0.33	0.15	0.59	0.64
0.36	0.20	0.17	0.09	0.06	0.01	0.03	0.16	0.20	0.09	0.29	0.29
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.34: Systematic uncertainties on $\langle\Psi(r)\rangle$ for light-jets in the semileptonic channels for 40 GeV $< p_T <$ 50 GeV

light-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 50 GeV $< p_{\mathrm{T}} <$ 70 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.61	2.78	0.08	3.52	0.75	0.14	5.06	6.16	1.31	5.90	7.77
0.08	1.37	1.49	0.99	1.27	0.22	0.54	2.28	2.78	0.70	2.98	3.52
0.12	1.26	1.42	0.75	0.57	0.37	0.50	1.37	1.67	0.42	2.14	2.39
0.16	1.13	1.30	0.43	0.22	0.62	0.41	1.28	1.55	0.49	1.97	2.22
0.20	0.95	1.12	0.46	0.24	0.42	0.34	1.01	1.23	0.25	1.58	1.79
0.24	0.79	0.86	0.31	0.17	0.29	0.25	0.93	1.13	0.09	1.32	1.49
0.28	0.59	0.61	0.18	0.07	0.23	0.16	0.55	0.67	0.07	0.88	0.95
0.32	0.37	0.39	0.01	0.05	0.17	0.12	0.38	0.46	0.12	0.58	0.65
0.36	0.17	0.16	0.02	0.06	0.10	0.04	0.13	0.16	0.09	0.26	0.28
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.35: Systematic uncertainties on $\langle\Psi(r)\rangle$ for light-jets in the semileptonic channels for 50 GeV $< p_T <$ 70 GeV

light-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 70 GeV $< p_{\mathrm{T}} <$ 100 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	2.76	2.81	2.46	4.09	4.23	1.64	2.91	3.51	3.77	7.55	8.47
0.08	1.06	1.41	1.53	2.47	2.56	0.46	2.56	3.09	1.64	4.41	5.21
0.12	0.93	1.16	1.19	1.31	1.01	0.06	1.25	1.51	0.52	2.27	2.57
0.16	0.85	0.98	0.73	0.73	0.44	0.01	0.67	0.81	0.12	1.39	1.54
0.20	0.73	0.86	0.45	0.41	0.43	0.10	0.51	0.62	0.16	1.10	1.23
0.24	0.61	0.71	0.30	0.30	0.32	0.04	0.41	0.49	0.07	0.86	0.97
0.28	0.44	0.51	0.19	0.23	0.38	0.05	0.31	0.37	0.13	0.70	0.78
0.32	0.29	0.31	0.18	0.13	0.23	0.01	0.19	0.22	0.06	0.45	0.47
0.36	0.13	0.12	0.06	0.05	0.14	0.02	0.09	0.10	0.04	0.22	0.22
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.36: Systematic uncertainties on $\langle\Psi(r)\rangle$ for light-jets in the semileptonic channels for 70 GeV $< p_T <$ 100 GeV

light-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 100 GeV $< p_{\mathrm{T}} <$ 150 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	3.76	3.68	2.37	5.00	8.80	0.09	3.42	3.97	3.91	11.14	12.12
0.08	1.65	1.81	1.70	2.23	2.44	0.47	1.81	2.10	2.02	4.38	4.79
0.12	0.95	1.04	1.95	1.07	0.52	0.11	1.01	1.17	0.34	2.47	2.00
0.16	0.80	0.83	1.37	0.49	0.90	0.10	0.68	0.79	0.12	1.95	1.54
0.20	0.65	0.69	0.99	0.26	0.61	0.04	0.60	0.69	0.12	1.46	1.19
0.24	0.47	0.53	0.62	0.26	0.15	0.05	0.29	0.34	0.06	0.85	0.70
0.28	0.33	0.37	0.35	0.11	0.14	0.03	0.33	0.39	0.05	0.60	0.57
0.32	0.21	0.23	0.17	0.11	0.06	0.02	0.19	0.23	0.03	0.35	0.35
0.36	0.10	0.10	0.07	0.01	0.02	0.01	0.07	0.09	0.04	0.15	0.14
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.37: Systematic uncertainties on $\langle\Psi(r)\rangle$ for light-jets in the semileptonic channels for 100 GeV $< p_T <$ 150 GeV

light-jet Systematic Uncertainties for $\langle\Psi(r)\rangle$ (%); 150 GeV $< p_{\mathrm{T}} <$ 200 GeV											
r	Clust (Up)	Clust (Down)	JES (Up)	JES (Down)	JER	JVF	PU (Up)	PU (Down)	Unf	Total (Up)	Total (Down)
0.04	3.04	2.92	0.74	5.70	8.45	0.68	1.16	1.34	5.03	10.41	11.83
0.08	1.67	1.59	6.50	1.64	2.19	0.80	1.55	1.79	2.07	7.56	4.26
0.12	0.54	0.61	4.68	1.07	1.45	0.24	0.16	0.19	0.94	5.03	2.14
0.16	0.61	0.70	3.18	0.13	1.70	0.06	0.25	0.29	0.58	3.71	1.95
0.20	0.45	0.51	1.51	0.06	0.75	0.06	0.09	0.10	0.66	1.86	1.13
0.24	0.39	0.44	0.96	0.05	0.50	0.09	0.00	0.00	0.18	1.17	0.69
0.28	0.29	0.31	0.76	0.16	0.08	0.06	0.01	0.01	0.37	0.90	0.52
0.32	0.17	0.16	0.44	0.12	0.12	0.01	0.06	0.07	0.07	0.49	0.25
0.36	0.09	0.09	0.19	0.05	0.06	0.02	0.02	0.02	0.07	0.24	0.14
0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4.38: Systematic uncertainties on $\langle\Psi(r)\rangle$ for light-jets in the semileptonic channels for 150 GeV $< p_T <$ 200 GeV

4.9.4 Cross-checks to the unfolding method

In order to study the stability of the unfolding method, two additional checks have been performed. On the one hand, an iterative bayesian unfolding is carried out, and on the other, the directly unfolded integrated jet shapes are compared with the ones obtained from the unfolded differential distributions.

4.9.4.1 Bayesian unfolding

An iterative Bayesian unfolding has been performed directly on the p_{T}-binned $\rho(r)$ and $\Psi(r)$ distributions for all values of r using RooUnfold [92]. This is done to take into account possible migrations between the bins of these distributions. As an example, figure 4.42 shows the unfolded data distributions, together with the particle level MC, of $\rho(r)$ in the first r bin and for the two first p_{T} bins for both b- and light-jets.

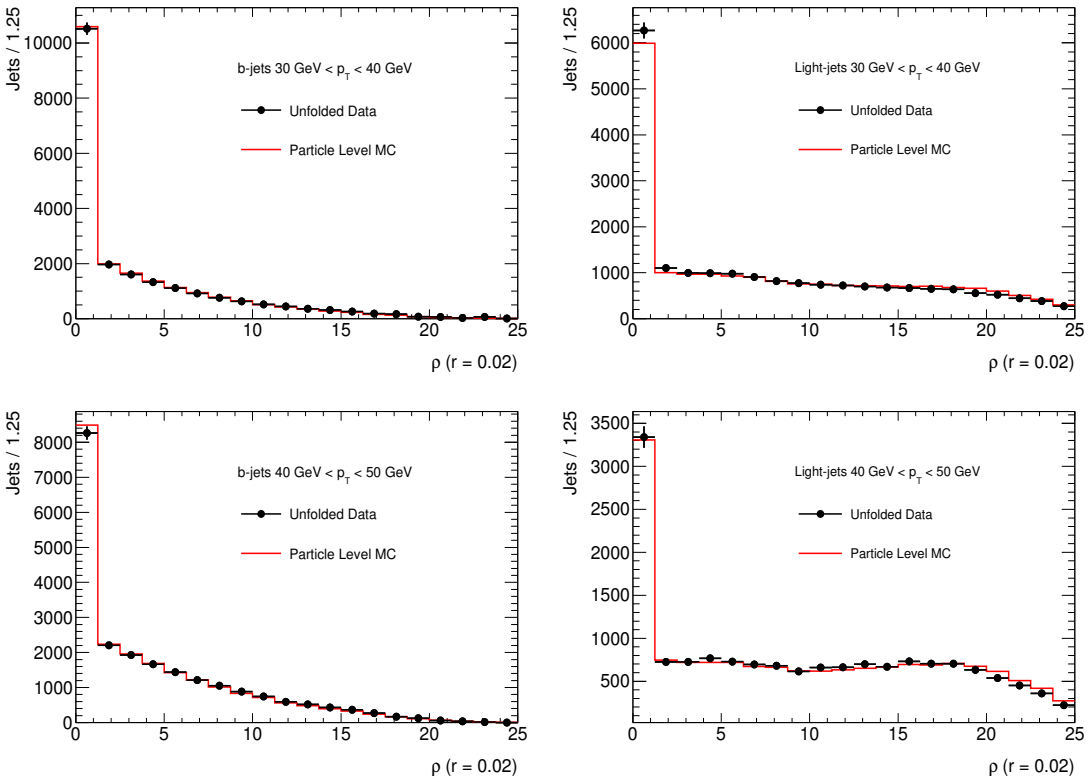

Figure 4.42: Unfolded $\rho(r = 0.02)$ distributions for b- and light-jets with 30 GeV $< p_T <$ 50 GeV using the Bayesian approach.

After the 120 distributions for b- and light-jets for each r and p_{T} bin are unfolded, their mean values $\langle \rho(r_i) \rangle$ are extracted. A comparison between the mean values obtained in the bin-by-bin and the Bayesian methods is presented in figure 4.43.

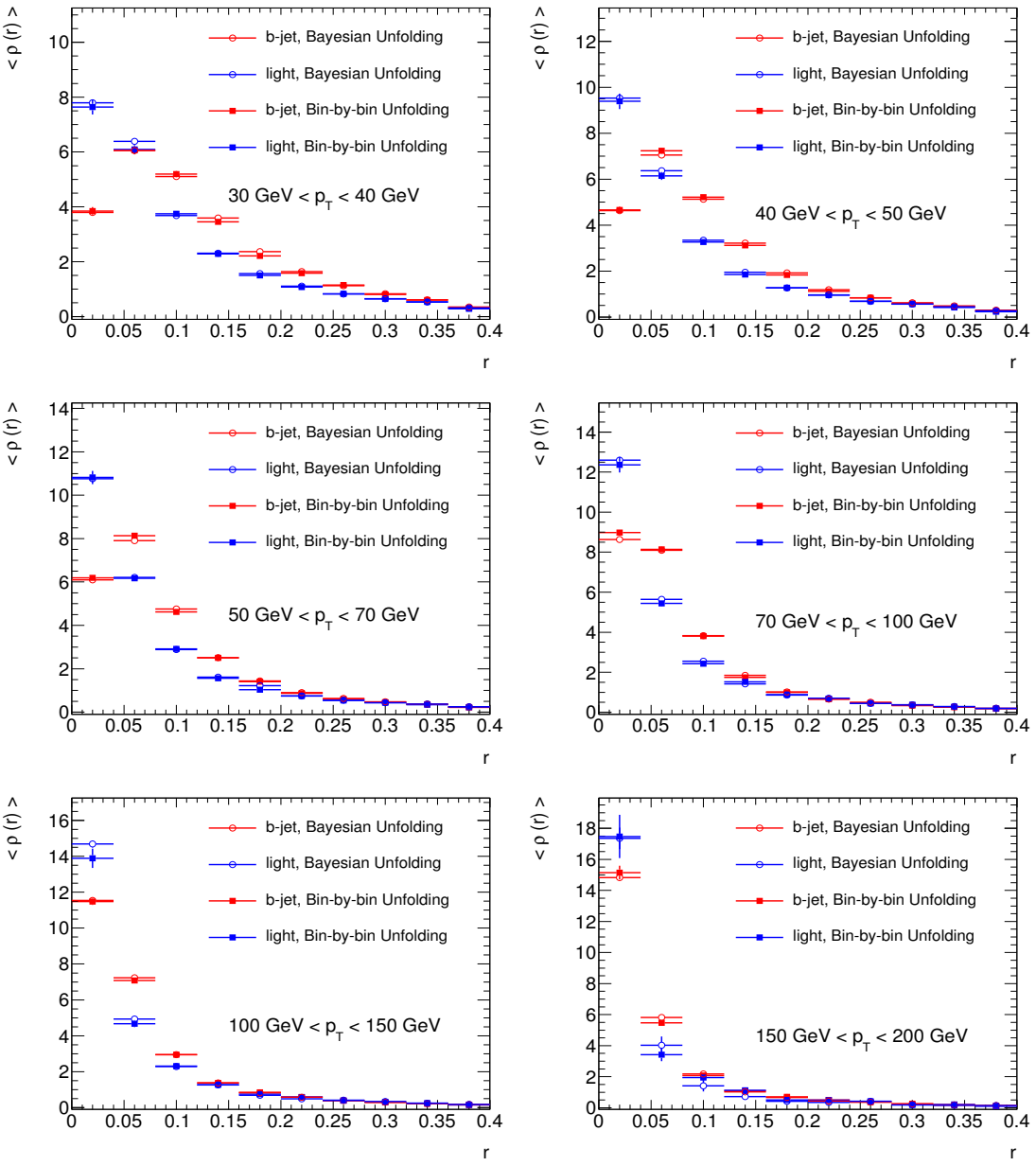

Figure 4.43: Comparison of the bin-by-bin to the bayesian unfolding for differential jet shapes

Analogously, the comparison between the bin-by-bin unfolding and the Bayesian unfolding for the integrated jet shapes is shown in figure 4.44

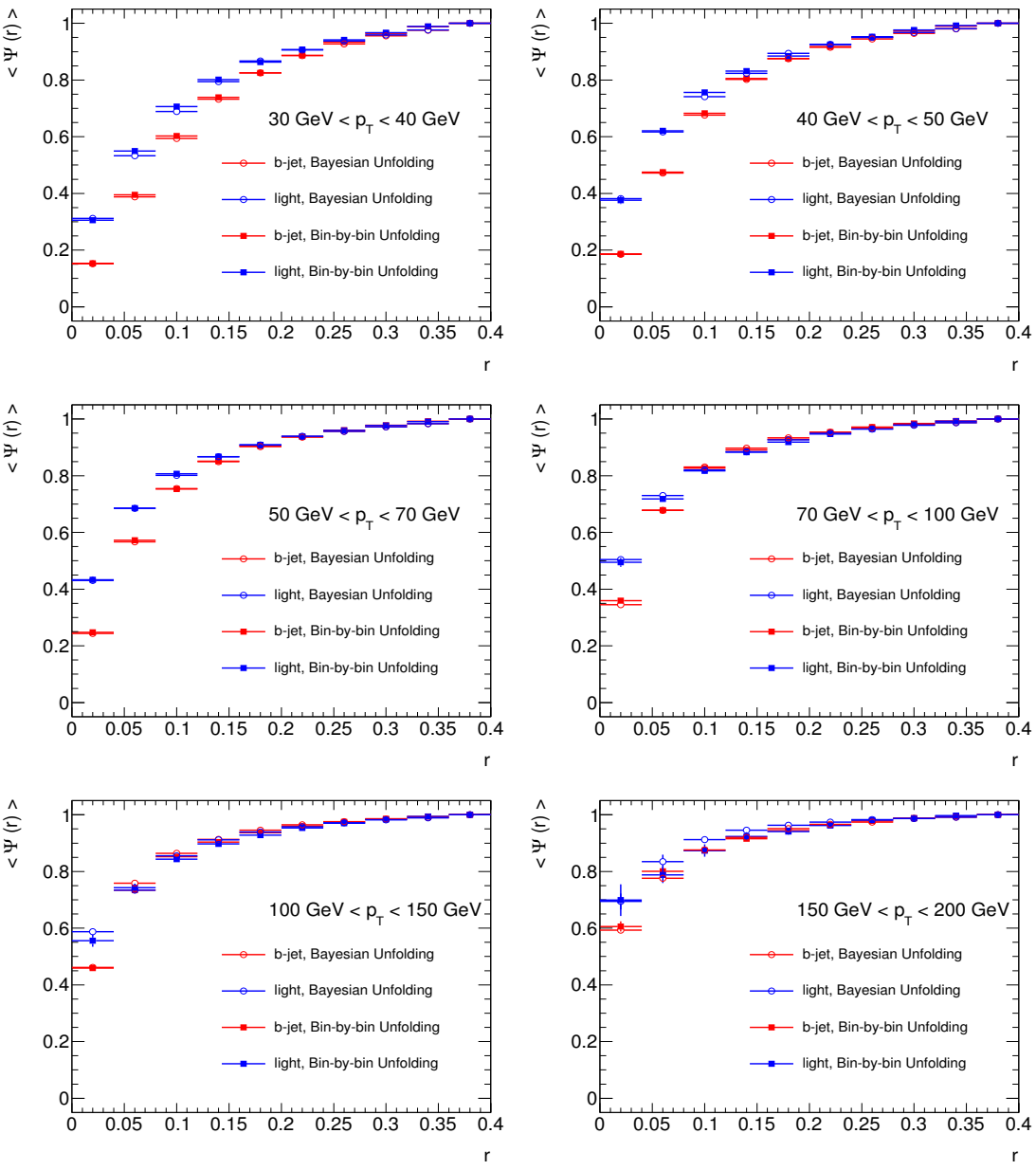

Figure 4.44: Comparison of the bin-by-bin to the bayesian unfolding for integrated jet shapes

4.9.4.2 Integrated shape recalculation

As an additional check of the stability of the unfolding procedure, the directly unfolded integrated jet shapes are compared with those obtained from integrating the unfolded differential distributions. The results agree to better than 1%, see Figure 4.45.

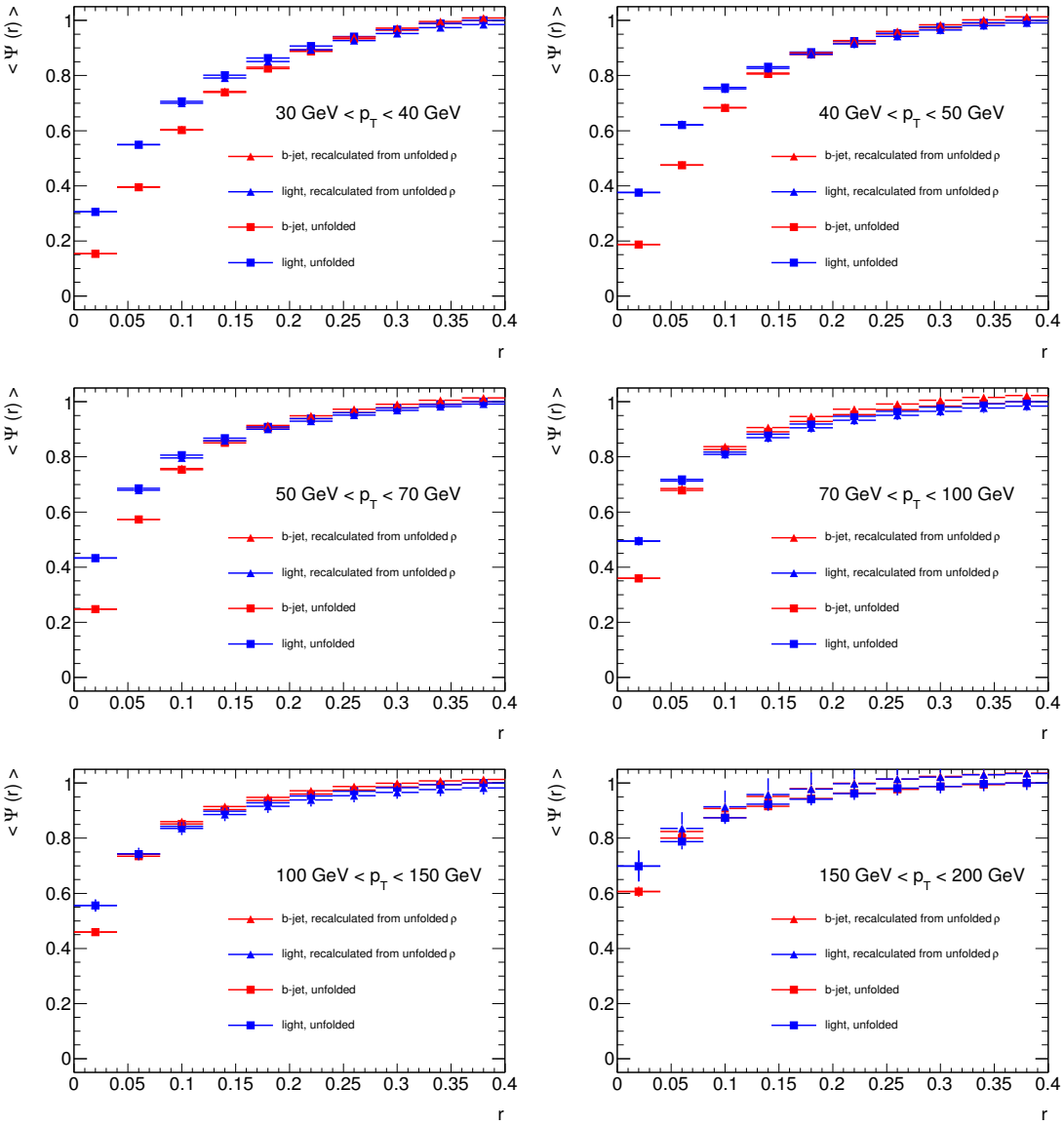

Figure 4.45: Comparison of $\Psi(r)$ obtained from direct unfolding with the cumulative distribution of the unfolded $\rho(r)$ values.

As the unfolding methods for the differential and integrated jet shapes are independent and the values for the comparison above agree at the 2% level, it can be concluded that the bin-by-bin unfolding method used here is stable and, to a large extent, independent of the bin-by-bin underlying correlations in the $\Psi(r)$ distributions.

4.10 Summary and conclusions

The jet structure of $t\bar{t}$ final states has been studied in both the dileptonic and semileptonic modes. The former proves to be a very clean and copious source of b-jets, as the top decays predominantly via $t \to Wb$. The latter is also a clean source of light jets coming from the hadronic decays of one of the W's in the final state. The differences between the b- and light quark jets obtained in this environment have been studied in terms of the differential $\rho(r)$ and integrated jet shapes $\Psi(r)$. These variables have been found to be very dependent on the jet transverse momentum, while its dependence on pseudorapidity is very weak.

The results have shown that the mean value $\langle\Psi(r)\rangle$ is smaller for b- than for light-jets in the region where it is possible to distinguish them, i.e. at low values of the jet internal radius r. This means that b-jets are broader than light quark induced jets, and therefore the cores of light jets are more energetically dense than those of b-jets, which can be regarded as more donut-shaped jets.

The effects due to pileup are found to be small and affect both b- and light-jets in a very similar manner.

The features discussed above, namely independence (strong dependence) of $\Psi(r)$ with $|\eta|$ (resp. p_{T}) and the differences between light and b-jets, have already been reported in several experiments, but we would like to remark that this is the first time that they are studied in $t\bar{t}$ final states. The purity of the light and b-jet samples is considerably higher than in previous inclusive jet based analysis [70], as the fake probability is smaller due to the known presence of b-jets in $t\bar{t}$ events.

In order to summarize the observations made above, in figures 4.46 to 4.51, a comparison is shown between the unfolded values for $\langle\rho(r)\rangle$ evaluated for both, light and b-jets, while in figures 4.52 to 4.57, the comparison between the unfolded values for $\langle\Psi_l(r)\rangle$ and $\langle\Psi_b(r)\rangle$ is shown. For the sake of comparison, the expectations from state of the art Monte Carlos have also been included. Here, fixed order matrix elements calculated up to NLO accuracy are matched to p_T ordered parton showers and to two fragmentation schemes, i.e. MC@NLO+Herwig as baseline and Powheg coupled to Pythia. The agreement between these reference Monte Carlo samples and data is reasonable for both the light and b-jet samples, except for small ($\simeq 10\%$) discrepancies for $r < 0.04$. Although the differential and integrated shape variables are self-normalized, they are known to be primarily sensitive to the matching of matrix elements and parton showers, and to a lesser extent, to details of the fragmentation. Thus these data will help to better constraint current Monte Carlo models and theoretical calculations.
Numerical results for the differential (integrated) jet shapes are given in tables 4.39 to 4.44 (resp. 4.45 to 4.50) for completeness. The uncertainties shown include statistical errors as well as all the systematics discussed in the previous section, including those arising from the unfolding procedure.

4.10.1 Differential jet shapes

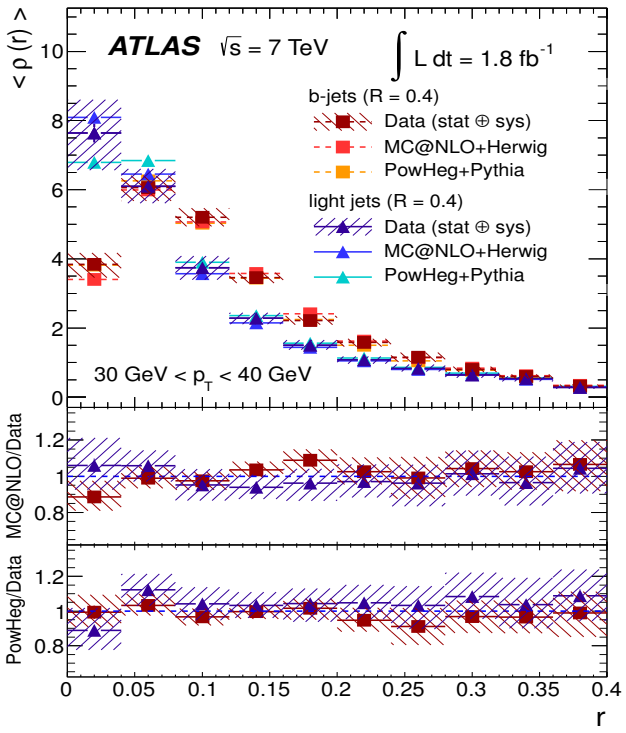

Figure 4.46: Unfolded values for $\langle\rho(r)\rangle$, with statistical and systematic uncertainties for 30 GeV $< p_{\mathrm{T}} <$ 40 GeV.

Collision data $\sqrt{s} = 7$ TeV; 30 GeV $< p_{\mathrm{T}} <$ 40 GeV		
r	$\langle\rho_b(r)\rangle$ [b-jets]	$\langle\rho_l(r)\rangle$ [light jets]
0.02	3.842 ± 0.147 (stat.) $^{+0.291}_{-0.359}$ (sys.)	7.639 ± 0.269 (stat.) $^{+0.927}_{-1.067}$ (sys.)
0.06	6.060 ± 0.144 (stat.) $^{+0.312}_{-0.359}$ (sys.)	6.097 ± 0.156 (stat.) $^{+0.480}_{-0.469}$ (sys.)
0.10	5.199 ± 0.114 (stat.) $^{+0.245}_{-0.231}$ (sys.)	3.745 ± 0.099 (stat.) $^{+0.323}_{-0.332}$ (sys.)
0.14	3.453 ± 0.085 (stat.) $^{+0.119}_{-0.126}$ (sys.)	2.284 ± 0.067 (stat.) $^{+0.142}_{-0.156}$ (sys.)
0.18	2.211 ± 0.055 (stat.) $^{+0.131}_{-0.107}$ (sys.)	1.497 ± 0.046 (stat.) $^{+0.144}_{-0.115}$ (sys.)
0.22	1.580 ± 0.041 (stat.) $^{+0.101}_{-0.113}$ (sys.)	1.077 ± 0.033 (stat.) $^{+0.091}_{-0.101}$ (sys.)
0.26	1.152 ± 0.030 (stat.) $^{+0.130}_{-0.132}$ (sys.)	0.833 ± 0.026 (stat.) $^{+0.107}_{-0.085}$ (sys.)
0.30	0.802 ± 0.020 (stat.) $^{+0.076}_{-0.074}$ (sys.)	0.642 ± 0.019 (stat.) $^{+0.074}_{-0.079}$ (sys.)
0.34	0.600 ± 0.015 (stat.) $^{+0.056}_{-0.061}$ (sys.)	0.533 ± 0.015 (stat.) $^{+0.068}_{-0.078}$ (sys.)
0.38	0.317 ± 0.009 (stat.) $^{+0.043}_{-0.039}$ (sys.)	0.284 ± 0.009 (stat.) $^{+0.037}_{-0.038}$ (sys.)

Table 4.39: Differential shapes $\langle\rho(r)\rangle$ for light and b-jets in both channels for 30 GeV $< p_T <$ 40 GeV

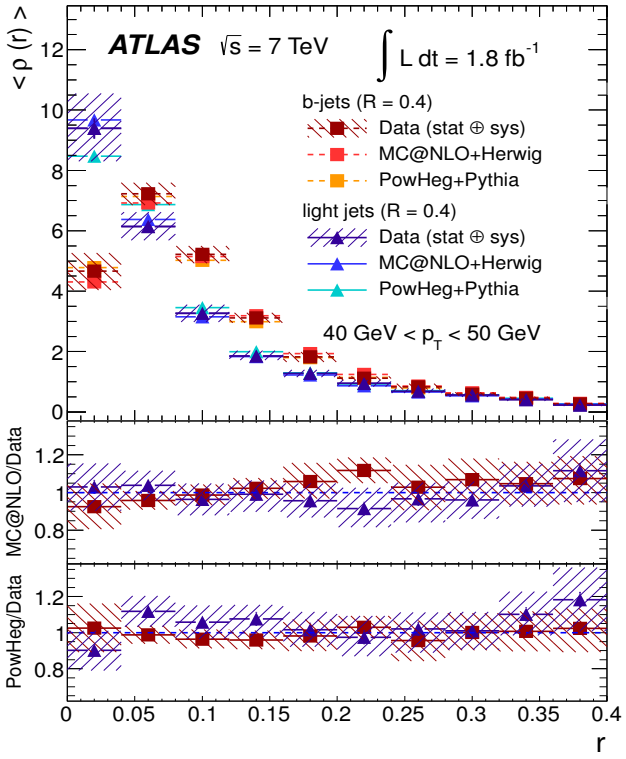

Figure 4.47: Unfolded values for $\langle\rho(r)\rangle$, with statistical and systematic uncertainties for 40 GeV $< p_{\mathrm{T}} <$ 50 GeV.

Collision data $\sqrt{s} = 7$ TeV; 40 GeV $< p_{\mathrm{T}} <$ 50 GeV		
r	$\langle\rho_b(r)\rangle$ [b-jets]	$\langle\rho_l(r)\rangle$ [light jets]
0.02	4.663 ± 0.151 (stat.) $^{+0.585}_{-0.605}$ (sys.)	9.391 ± 0.337 (stat.) $^{+1.109}_{-1.078}$ (sys.)
0.06	7.231 ± 0.145 (stat.) $^{+0.334}_{-0.349}$ (sys.)	6.141 ± 0.171 (stat.) $^{+0.442}_{-0.433}$ (sys.)
0.10	5.218 ± 0.105 (stat.) $^{+0.255}_{-0.275}$ (sys.)	3.266 ± 0.099 (stat.) $^{+0.267}_{-0.268}$ (sys.)
0.14	3.117 ± 0.071 (stat.) $^{+0.152}_{-0.151}$ (sys.)	1.853 ± 0.068 (stat.) $^{+0.162}_{-0.117}$ (sys.)
0.18	1.829 ± 0.046 (stat.) $^{+0.148}_{-0.165}$ (sys.)	1.277 ± 0.048 (stat.) $^{+0.110}_{-0.113}$ (sys.)
0.22	1.115 ± 0.030 (stat.) $^{+0.060}_{-0.062}$ (sys.)	0.949 ± 0.038 (stat.) $^{+0.096}_{-0.109}$ (sys.)
0.26	0.835 ± 0.022 (stat.) $^{+0.097}_{-0.094}$ (sys.)	0.687 ± 0.026 (stat.) $^{+0.085}_{-0.055}$ (sys.)
0.30	0.593 ± 0.015 (stat.) $^{+0.057}_{-0.060}$ (sys.)	0.559 ± 0.020 (stat.) $^{+0.055}_{-0.054}$ (sys.)
0.34	0.458 ± 0.011 (stat.) $^{+0.049}_{-0.050}$ (sys.)	0.413 ± 0.015 (stat.) $^{+0.041}_{-0.045}$ (sys.)
0.38	0.263 ± 0.007 (stat.) $^{+0.033}_{-0.028}$ (sys.)	0.228 ± 0.008 (stat.) $^{+0.033}_{-0.033}$ (sys.)

Table 4.40: Differential shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 40 GeV $< p_{\mathrm{T}} <$ 50 GeV

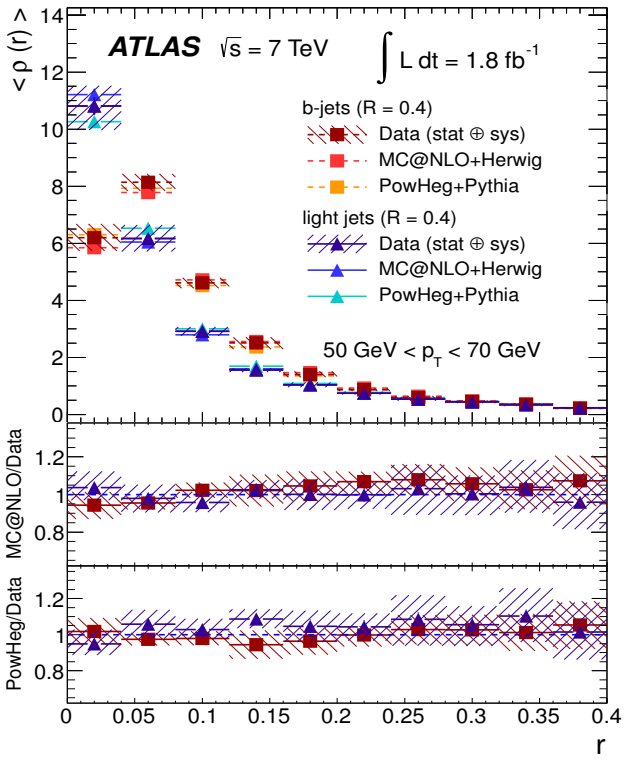

Figure 4.48: Unfolded values for $\langle\rho(r)\rangle$, with statistical and systematic uncertainties for 50 GeV $< p_{\mathrm{T}} <$ 70 GeV.

Collision data $\sqrt{s} = 7$ TeV; 50 GeV $< p_{\mathrm{T}} <$ 70 GeV		
r	$\langle\rho_b(r)\rangle$ [b-jets]	$\langle\rho_l(r)\rangle$ [light jets]
0.02	6.193 ± 0.129 (stat.) $^{+0.465}_{-0.441}$ (sys.)	10.817 ± 0.306 (stat.) $^{+0.638}_{-0.840}$ (sys.)
0.06	8.139 ± 0.107 (stat.) $^{+0.272}_{-0.286}$ (sys.)	6.170 ± 0.141 (stat.) $^{+0.451}_{-0.441}$ (sys.)
0.10	4.615 ± 0.064 (stat.) $^{+0.168}_{-0.179}$ (sys.)	2.916 ± 0.077 (stat.) $^{+0.135}_{-0.153}$ (sys.)
0.14	2.504 ± 0.042 (stat.) $^{+0.201}_{-0.206}$ (sys.)	1.559 ± 0.050 (stat.) $^{+0.052}_{-0.059}$ (sys.)
0.18	1.403 ± 0.027 (stat.) $^{+0.106}_{-0.100}$ (sys.)	1.039 ± 0.037 (stat.) $^{+0.078}_{-0.079}$ (sys.)
0.22	0.867 ± 0.018 (stat.) $^{+0.053}_{-0.039}$ (sys.)	0.745 ± 0.027 (stat.) $^{+0.048}_{-0.046}$ (sys.)
0.26	0.597 ± 0.012 (stat.) $^{+0.047}_{-0.044}$ (sys.)	0.540 ± 0.018 (stat.) $^{+0.071}_{-0.064}$ (sys.)
0.30	0.453 ± 0.009 (stat.) $^{+0.037}_{-0.036}$ (sys.)	0.439 ± 0.015 (stat.) $^{+0.047}_{-0.040}$ (sys.)
0.34	0.359 ± 0.007 (stat.) $^{+0.035}_{-0.036}$ (sys.)	0.341 ± 0.011 (stat.) $^{+0.043}_{-0.046}$ (sys.)
0.38	0.215 ± 0.004 (stat.) $^{+0.027}_{-0.026}$ (sys.)	0.229 ± 0.007 (stat.) $^{+0.034}_{-0.036}$ (sys.)

Table 4.41: Differential shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 50 GeV $< p_{\mathrm{T}} <$ 70 GeV

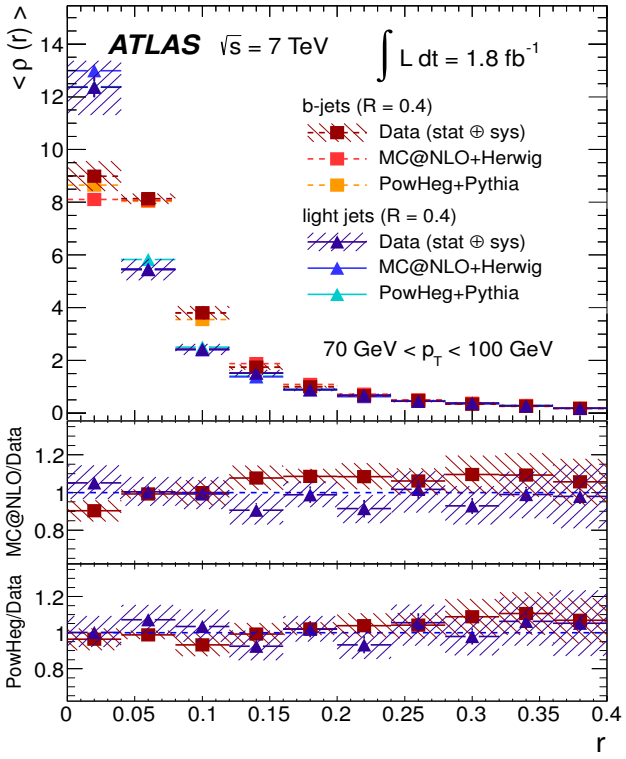

Figure 4.49: Unfolded values for $\langle\rho(r)\rangle$, with statistical and systematic uncertainties for 70 GeV $< p_{\mathrm{T}} <$ 100 GeV.

Collision data $\sqrt{s} = 7$ TeV; 70 GeV $< p_{\mathrm{T}} <$ 100 GeV		
r	$\langle\rho_b(r)\rangle$ [b-jets]	$\langle\rho_l(r)\rangle$ [light jets]
0.02	8.984 ± 0.151 (stat.) $^{+0.554}_{-0.535}$ (sys.)	12.366 ± 0.382 (stat.) $^{+0.933}_{-1.047}$ (sys.)
0.06	8.135 ± 0.102 (stat.) $^{+0.173}_{-0.171}$ (sys.)	5.444 ± 0.158 (stat.) $^{+0.376}_{-0.390}$ (sys.)
0.10	3.802 ± 0.052 (stat.) $^{+0.253}_{-0.248}$ (sys.)	2.424 ± 0.078 (stat.) $^{+0.182}_{-0.208}$ (sys.)
0.14	1.744 ± 0.030 (stat.) $^{+0.100}_{-0.102}$ (sys.)	1.516 ± 0.060 (stat.) $^{+0.114}_{-0.129}$ (sys.)
0.18	1.000 ± 0.020 (stat.) $^{+0.026}_{-0.031}$ (sys.)	0.888 ± 0.039 (stat.) $^{+0.045}_{-0.051}$ (sys.)
0.22	0.658 ± 0.015 (stat.) $^{+0.039}_{-0.041}$ (sys.)	0.681 ± 0.034 (stat.) $^{+0.046}_{-0.041}$ (sys.)
0.26	0.471 ± 0.011 (stat.) $^{+0.031}_{-0.028}$ (sys.)	0.447 ± 0.022 (stat.) $^{+0.050}_{-0.036}$ (sys.)
0.30	0.337 ± 0.007 (stat.) $^{+0.031}_{-0.030}$ (sys.)	0.378 ± 0.017 (stat.) $^{+0.036}_{-0.039}$ (sys.)
0.34	0.259 ± 0.005 (stat.) $^{+0.027}_{-0.027}$ (sys.)	0.280 ± 0.011 (stat.) $^{+0.034}_{-0.033}$ (sys.)
0.38	0.174 ± 0.004 (stat.) $^{+0.020}_{-0.020}$ (sys.)	0.185 ± 0.007 (stat.) $^{+0.031}_{-0.031}$ (sys.)

Table 4.42: Differential shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 70 GeV $< p_{\mathrm{T}} <$ 100 GeV

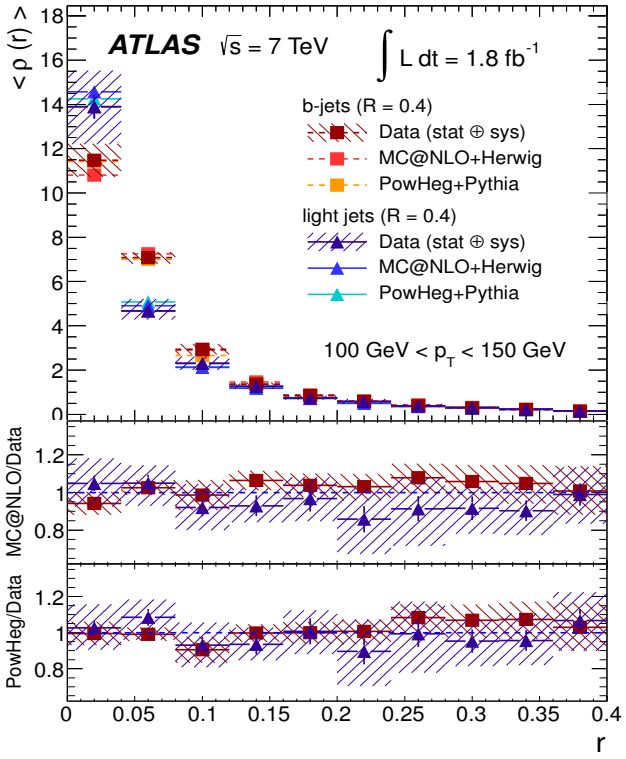

Figure 4.50: Unfolded values for $\langle\rho(r)\rangle$, with statistical and systematic uncertainties for 100 GeV $< p_{\mathrm{T}} <$ 150 GeV.

Collision data $\sqrt{s} = 7$ TeV; 100 GeV $< p_{\mathrm{T}} <$ 150 GeV		
r	$\langle\rho_b(r)\rangle$ [b-jets]	$\langle\rho_l(r)\rangle$ [light jets]
0.02	11.483 ± 0.195 (stat.) $^{+0.711}_{-0.740}$ (sys.)	13.894 ± 0.538 (stat.) $^{+1.547}_{-1.684}$ (sys.)
0.06	7.076 ± 0.109 (stat.) $^{+0.241}_{-0.254}$ (sys.)	4.676 ± 0.199 (stat.) $^{+0.499}_{-0.374}$ (sys.)
0.10	2.943 ± 0.051 (stat.) $^{+0.233}_{-0.233}$ (sys.)	2.307 ± 0.115 (stat.) $^{+0.279}_{-0.292}$ (sys.)
0.14	1.370 ± 0.031 (stat.) $^{+0.060}_{-0.058}$ (sys.)	1.266 ± 0.075 (stat.) $^{+0.094}_{-0.102}$ (sys.)
0.18	0.845 ± 0.024 (stat.) $^{+0.046}_{-0.045}$ (sys.)	0.744 ± 0.053 (stat.) $^{+0.077}_{-0.068}$ (sys.)
0.22	0.584 ± 0.018 (stat.) $^{+0.038}_{-0.034}$ (sys.)	0.581 ± 0.047 (stat.) $^{+0.116}_{-0.103}$ (sys.)
0.26	0.389 ± 0.012 (stat.) $^{+0.028}_{-0.024}$ (sys.)	0.389 ± 0.028 (stat.) $^{+0.079}_{-0.064}$ (sys.)
0.30	0.288 ± 0.009 (stat.) $^{+0.024}_{-0.022}$ (sys.)	0.307 ± 0.021 (stat.) $^{+0.040}_{-0.032}$ (sys.)
0.34	0.214 ± 0.006 (stat.) $^{+0.021}_{-0.019}$ (sys.)	0.244 ± 0.015 (stat.) $^{+0.031}_{-0.035}$ (sys.)
0.38	0.145 ± 0.004 (stat.) $^{+0.018}_{-0.019}$ (sys.)	0.149 ± 0.009 (stat.) $^{+0.022}_{-0.020}$ (sys.)

Table 4.43: Differential shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 100 GeV $< p_{\mathrm{T}} <$ 150 GeV

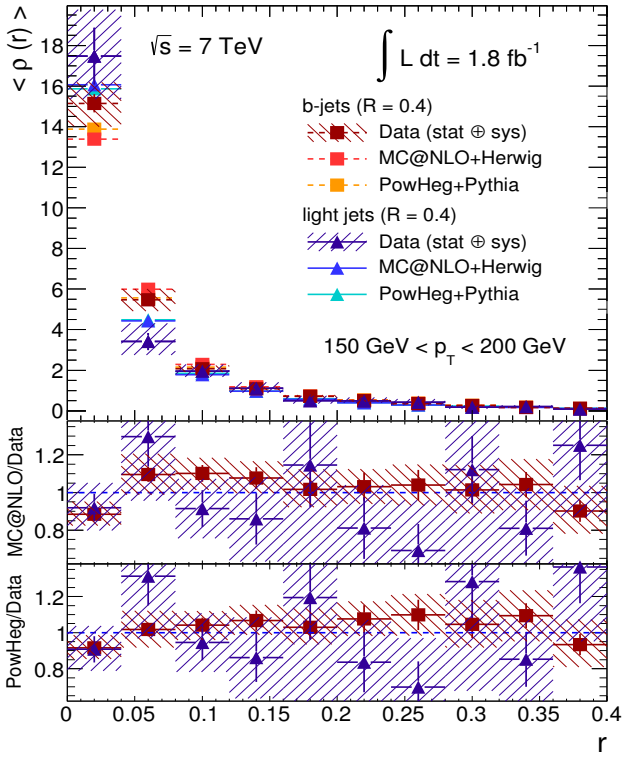

Figure 4.51: Unfolded values for $\langle\rho(r)\rangle$, with statistical and systematic uncertainties for 150 GeV $< p_{\mathrm{T}} <$ 200 GeV.

Collision data $\sqrt{s} = 7$ TeV; 150 GeV $< p_{\mathrm{T}} <$ 200 GeV		
r	$\langle\rho_b(r)\rangle$ [b-jets]	$\langle\rho_l(r)\rangle$ [light jets]
0.02	15.151 ± 0.445 (stat.) $^{+0.938}_{-1.068}$ (sys.)	17.476 ± 1.396 (stat.) $^{+1.818}_{-2.068}$ (sys.)
0.06	5.468 ± 0.197 (stat.) $^{+0.500}_{-0.516}$ (sys.)	3.418 ± 0.419 (stat.) $^{+0.792}_{-0.629}$ (sys.)
0.10	2.074 ± 0.083 (stat.) $^{+0.141}_{-0.132}$ (sys.)	1.952 ± 0.204 (stat.) $^{+0.267}_{-0.263}$ (sys.)
0.14	1.091 ± 0.057 (stat.) $^{+0.058}_{-0.069}$ (sys.)	1.118 ± 0.175 (stat.) $^{+0.235}_{-0.228}$ (sys.)
0.18	0.711 ± 0.045 (stat.) $^{+0.046}_{-0.059}$ (sys.)	0.499 ± 0.097 (stat.) $^{+0.221}_{-0.136}$ (sys.)
0.22	0.499 ± 0.035 (stat.) $^{+0.026}_{-0.028}$ (sys.)	0.491 ± 0.098 (stat.) $^{+0.142}_{-0.096}$ (sys.)
0.26	0.353 ± 0.027 (stat.) $^{+0.026}_{-0.025}$ (sys.)	0.419 ± 0.085 (stat.) $^{+0.146}_{-0.139}$ (sys.)
0.30	0.258 ± 0.019 (stat.) $^{+0.026}_{-0.028}$ (sys.)	0.179 ± 0.028 (stat.) $^{+0.080}_{-0.074}$ (sys.)
0.34	0.167 ± 0.010 (stat.) $^{+0.019}_{-0.019}$ (sys.)	0.210 ± 0.037 (stat.) $^{+0.057}_{-0.040}$ (sys.)
0.38	0.128 ± 0.008 (stat.) $^{+0.015}_{-0.017}$ (sys.)	0.094 ± 0.014 (stat.) $^{+0.025}_{-0.015}$ (sys.)

Table 4.44: Differential shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 150 GeV $< p_{\mathrm{T}} <$ 200 GeV

4.10.2 Integrated jet shapes

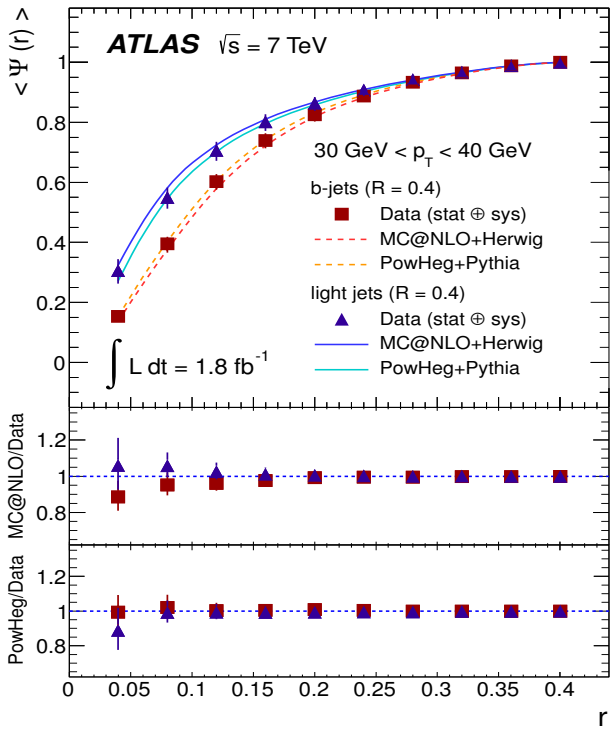

Figure 4.52: Unfolded values for $\langle\Psi(r)\rangle$, with statistical and systematic uncertainties for 30 GeV $< p_{\mathrm{T}} <$ 40 GeV.

Collision data $\sqrt{s} = 7$ TeV; 30 GeV $< p_{\mathrm{T}} <$ 40 GeV		
r	$\langle\Psi_b(r)\rangle$ [b-jets]	$\langle\Psi_l(r)\rangle$ [light jets]
0.04	0.154 ± 0.006 (stat.) $^{+0.012}_{-0.014}$ (sys.)	0.306 ± 0.011 (stat.) $^{+0.037}_{-0.043}$ (sys.)
0.08	0.395 ± 0.007 (stat.) $^{+0.023}_{-0.028}$ (sys.)	0.550 ± 0.009 (stat.) $^{+0.031}_{-0.037}$ (sys.)
0.12	0.602 ± 0.006 (stat.) $^{+0.025}_{-0.026}$ (sys.)	0.706 ± 0.007 (stat.) $^{+0.028}_{-0.034}$ (sys.)
0.16	0.739 ± 0.004 (stat.) $^{+0.025}_{-0.025}$ (sys.)	0.802 ± 0.005 (stat.) $^{+0.025}_{-0.030}$ (sys.)
0.20	0.825 ± 0.003 (stat.) $^{+0.020}_{-0.023}$ (sys.)	0.863 ± 0.004 (stat.) $^{+0.020}_{-0.025}$ (sys.)
0.24	0.887 ± 0.003 (stat.) $^{+0.016}_{-0.017}$ (sys.)	0.907 ± 0.003 (stat.) $^{+0.016}_{-0.019}$ (sys.)
0.28	0.934 ± 0.002 (stat.) $^{+0.012}_{-0.012}$ (sys.)	0.942 ± 0.002 (stat.) $^{+0.011}_{-0.014}$ (sys.)
0.32	0.964 ± 0.001 (stat.) $^{+0.007}_{-0.007}$ (sys.)	0.967 ± 0.001 (stat.) $^{+0.007}_{-0.008}$ (sys.)
0.36	0.988 ± 0.001 (stat.) $^{+0.004}_{-0.002}$ (sys.)	0.989 ± 0.001 (stat.) $^{+0.003}_{-0.003}$ (sys.)
0.40	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)

Table 4.45: Integrated shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 30 GeV $< p_T <$ 40 GeV

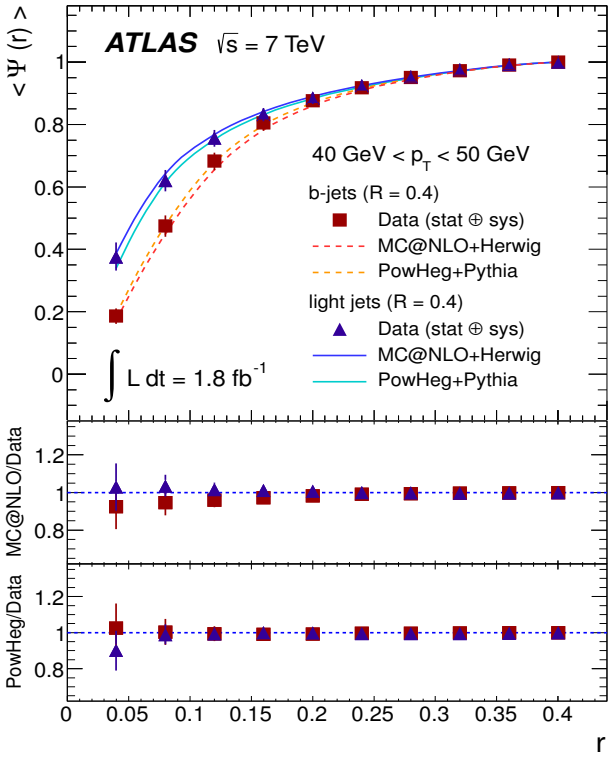

Figure 4.53: Unfolded values for $\langle\Psi(r)\rangle$, with statistical and systematic uncertainties for 40 GeV $< p_{\mathrm{T}} <$ 50 GeV.

Collision data $\sqrt{s} = 7$ TeV; 40 GeV $< p_{\mathrm{T}} <$ 50 GeV		
r	$\langle\Psi_b(r)\rangle$ [b-jets]	$\langle\Psi_l(r)\rangle$ [light jets]
0.04	0.187 ± 0.006 (stat.) $^{+0.023}_{-0.024}$ (sys.)	0.376 ± 0.013 (stat.) $^{+0.044}_{-0.043}$ (sys.)
0.08	0.475 ± 0.007 (stat.) $^{+0.033}_{-0.034}$ (sys.)	0.621 ± 0.011 (stat.) $^{+0.032}_{-0.034}$ (sys.)
0.12	0.683 ± 0.005 (stat.) $^{+0.027}_{-0.029}$ (sys.)	0.757 ± 0.008 (stat.) $^{+0.025}_{-0.027}$ (sys.)
0.16	0.805 ± 0.004 (stat.) $^{+0.023}_{-0.025}$ (sys.)	0.832 ± 0.006 (stat.) $^{+0.021}_{-0.022}$ (sys.)
0.20	0.876 ± 0.003 (stat.) $^{+0.017}_{-0.018}$ (sys.)	0.885 ± 0.004 (stat.) $^{+0.017}_{-0.018}$ (sys.)
0.24	0.918 ± 0.002 (stat.) $^{+0.015}_{-0.016}$ (sys.)	0.925 ± 0.003 (stat.) $^{+0.012}_{-0.014}$ (sys.)
0.28	0.950 ± 0.002 (stat.) $^{+0.010}_{-0.011}$ (sys.)	0.953 ± 0.002 (stat.) $^{+0.010}_{-0.011}$ (sys.)
0.32	0.973 ± 0.001 (stat.) $^{+0.007}_{-0.006}$ (sys.)	0.976 ± 0.001 (stat.) $^{+0.006}_{-0.006}$ (sys.)
0.36	0.990 ± 0.001 (stat.) $^{+0.003}_{-0.002}$ (sys.)	0.992 ± 0.001 (stat.) $^{+0.003}_{-0.003}$ (sys.)
0.40	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)

Table 4.46: Integrated shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 40 GeV $< p_{\mathrm{T}} <$ 50 GeV

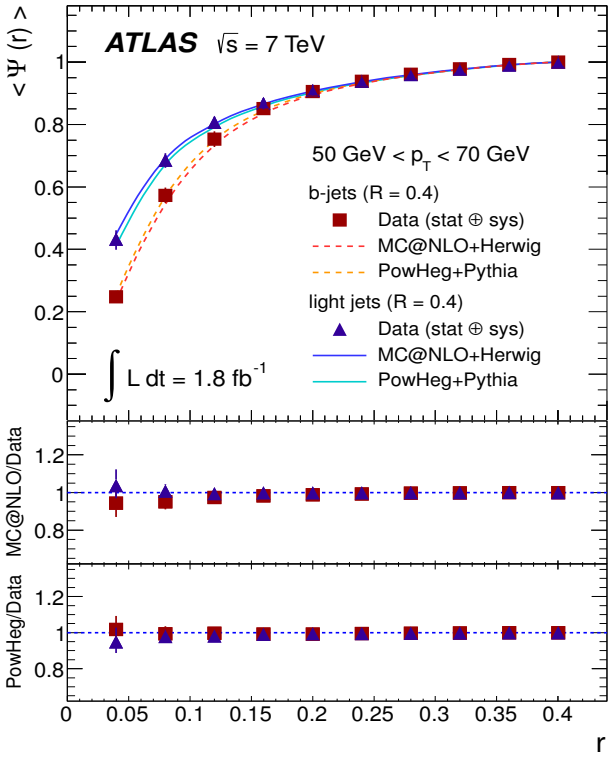

Figure 4.54: Unfolded values for $\langle\Psi(r)\rangle$, with statistical and systematic uncertainties for 50 GeV $< p_{\mathrm{T}} <$ 70 GeV.

Collision data $\sqrt{s} = 7$ TeV; 50 GeV $< p_{\mathrm{T}} <$ 70 GeV		
r	$\langle\Psi_b(r)\rangle$ [b-jets]	$\langle\Psi_l(r)\rangle$ [light jets]
0.04	0.248 ± 0.005 (stat.) $^{+0.019}_{-0.018}$ (sys.)	0.433 ± 0.012 (stat.) $^{+0.026}_{-0.034}$ (sys.)
0.08	0.573 ± 0.005 (stat.) $^{+0.024}_{-0.023}$ (sys.)	0.686 ± 0.009 (stat.) $^{+0.020}_{-0.024}$ (sys.)
0.12	0.753 ± 0.004 (stat.) $^{+0.025}_{-0.025}$ (sys.)	0.807 ± 0.006 (stat.) $^{+0.017}_{-0.019}$ (sys.)
0.16	0.851 ± 0.003 (stat.) $^{+0.019}_{-0.018}$ (sys.)	0.868 ± 0.004 (stat.) $^{+0.017}_{-0.019}$ (sys.)
0.20	0.905 ± 0.002 (stat.) $^{+0.015}_{-0.015}$ (sys.)	0.909 ± 0.003 (stat.) $^{+0.014}_{-0.016}$ (sys.)
0.24	0.938 ± 0.001 (stat.) $^{+0.012}_{-0.013}$ (sys.)	0.939 ± 0.002 (stat.) $^{+0.012}_{-0.014}$ (sys.)
0.28	0.961 ± 0.001 (stat.) $^{+0.008}_{-0.009}$ (sys.)	0.960 ± 0.002 (stat.) $^{+0.008}_{-0.009}$ (sys.)
0.32	0.978 ± 0.001 (stat.) $^{+0.005}_{-0.005}$ (sys.)	0.977 ± 0.001 (stat.) $^{+0.006}_{-0.006}$ (sys.)
0.36	0.992 ± 0.000 (stat.) $^{+0.003}_{-0.002}$ (sys.)	0.990 ± 0.001 (stat.) $^{+0.003}_{-0.003}$ (sys.)
0.40	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)

Table 4.47: Integrated shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 50 GeV $< p_{\mathrm{T}} <$ 70 GeV

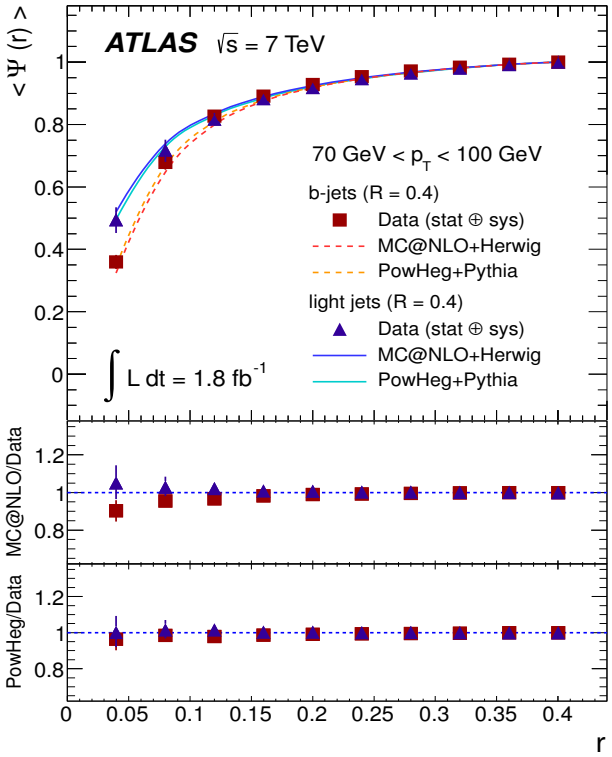

Figure 4.55: Unfolded values for $\langle\Psi(r)\rangle$, with statistical and systematic uncertainties for 70 GeV $< p_{\mathrm{T}} <$ 100 GeV.

Collision data $\sqrt{s} = 7$ TeV; 70 GeV $< p_{\mathrm{T}} <$ 100 GeV		
r	$\langle\Psi_b(r)\rangle$ [b-jets]	$\langle\Psi_l(r)\rangle$ [light jets]
0.04	0.359 ± 0.006 (stat.) $^{+0.022}_{-0.021}$ (sys.)	0.495 ± 0.015 (stat.) $^{+0.037}_{-0.042}$ (sys.)
0.08	0.678 ± 0.005 (stat.) $^{+0.023}_{-0.023}$ (sys.)	0.718 ± 0.010 (stat.) $^{+0.032}_{-0.037}$ (sys.)
0.12	0.827 ± 0.003 (stat.) $^{+0.017}_{-0.018}$ (sys.)	0.818 ± 0.007 (stat.) $^{+0.019}_{-0.021}$ (sys.)
0.16	0.891 ± 0.002 (stat.) $^{+0.012}_{-0.013}$ (sys.)	0.883 ± 0.005 (stat.) $^{+0.012}_{-0.014}$ (sys.)
0.20	0.928 ± 0.002 (stat.) $^{+0.011}_{-0.012}$ (sys.)	0.919 ± 0.004 (stat.) $^{+0.010}_{-0.011}$ (sys.)
0.24	0.954 ± 0.001 (stat.) $^{+0.009}_{-0.009}$ (sys.)	0.947 ± 0.003 (stat.) $^{+0.008}_{-0.009}$ (sys.)
0.28	0.972 ± 0.001 (stat.) $^{+0.006}_{-0.007}$ (sys.)	0.965 ± 0.002 (stat.) $^{+0.007}_{-0.008}$ (sys.)
0.32	0.984 ± 0.001 (stat.) $^{+0.004}_{-0.004}$ (sys.)	0.981 ± 0.001 (stat.) $^{+0.004}_{-0.005}$ (sys.)
0.36	0.993 ± 0.000 (stat.) $^{+0.002}_{-0.002}$ (sys.)	0.992 ± 0.001 (stat.) $^{+0.002}_{-0.002}$ (sys.)
0.40	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)

Table 4.48: Integrated shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 70 GeV $< p_{\mathrm{T}} <$ 100 GeV

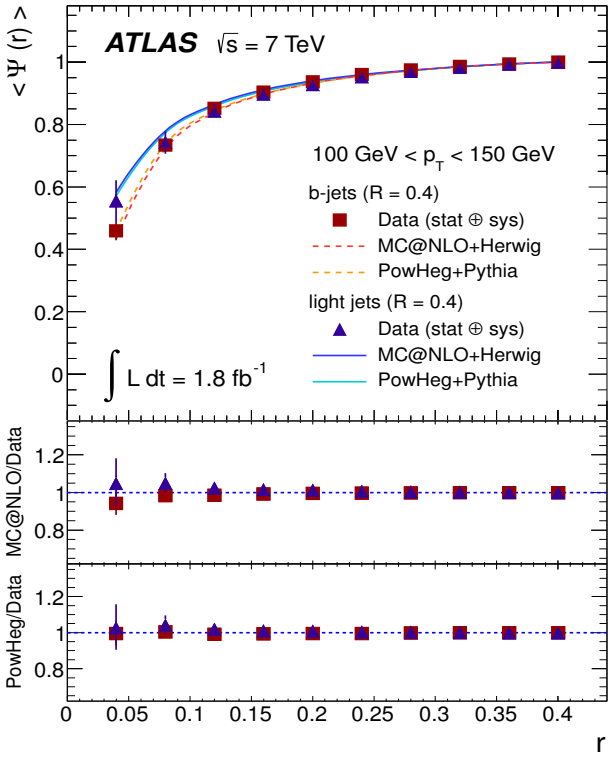

Figure 4.56: Unfolded values for $\langle\Psi(r)\rangle$, with statistical and systematic uncertainties for 100 GeV $< p_{\mathrm{T}} <$ 150 GeV.

Collision data $\sqrt{s} = 7$ TeV; 100 GeV $< p_{\mathrm{T}} <$ 150 GeV		
r	$\langle\Psi_b(r)\rangle$ [all]	$\langle\Psi_l(r)\rangle$ [light]
0.04	0.459 ± 0.008 (stat.) $^{+0.028}_{-0.030}$ (sys.)	0.556 ± 0.022 (stat.) $^{+0.062}_{-0.067}$ (sys.)
0.08	0.734 ± 0.005 (stat.) $^{+0.019}_{-0.020}$ (sys.)	0.743 ± 0.014 (stat.) $^{+0.033}_{-0.036}$ (sys.)
0.12	0.852 ± 0.004 (stat.) $^{+0.013}_{-0.012}$ (sys.)	0.843 ± 0.010 (stat.) $^{+0.021}_{-0.017}$ (sys.)
0.16	0.904 ± 0.002 (stat.) $^{+0.010}_{-0.010}$ (sys.)	0.898 ± 0.007 (stat.) $^{+0.017}_{-0.014}$ (sys.)
0.20	0.937 ± 0.002 (stat.) $^{+0.008}_{-0.008}$ (sys.)	0.928 ± 0.005 (stat.) $^{+0.014}_{-0.011}$ (sys.)
0.24	0.960 ± 0.001 (stat.) $^{+0.006}_{-0.006}$ (sys.)	0.954 ± 0.003 (stat.) $^{+0.008}_{-0.007}$ (sys.)
0.28	0.975 ± 0.001 (stat.) $^{+0.005}_{-0.005}$ (sys.)	0.970 ± 0.002 (stat.) $^{+0.006}_{-0.006}$ (sys.)
0.32	0.986 ± 0.001 (stat.) $^{+0.003}_{-0.003}$ (sys.)	0.983 ± 0.001 (stat.) $^{+0.003}_{-0.003}$ (sys.)
0.36	0.994 ± 0.000 (stat.) $^{+0.001}_{-0.001}$ (sys.)	0.994 ± 0.001 (stat.) $^{+0.001}_{-0.001}$ (sys.)
0.40	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)

Table 4.49: Integrated shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 100 GeV $< p_{\mathrm{T}} <$ 150 GeV

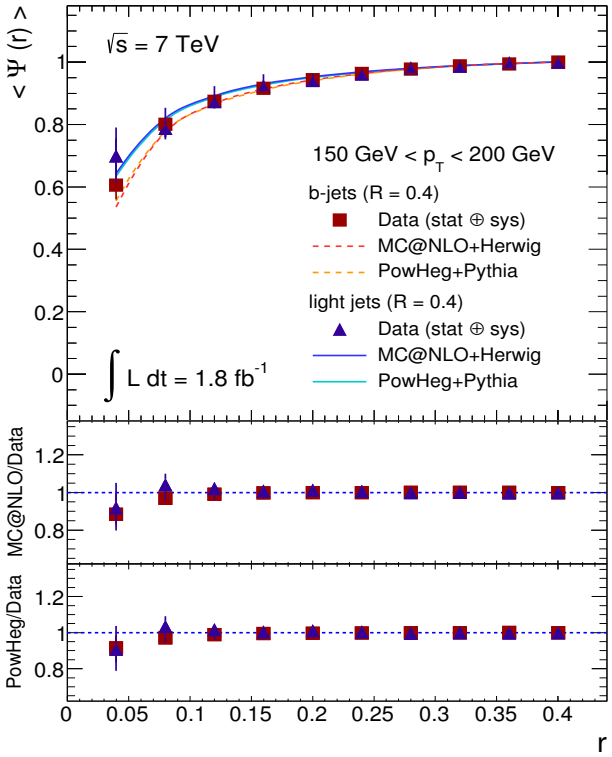

Figure 4.57: Unfolded values for $\langle\Psi(r)\rangle$, with statistical and systematic uncertainties for 150 GeV $< p_{\mathrm{T}} <$ 200 GeV.

Collision data $\sqrt{s} = 7$ TeV; 150 GeV $< p_{\mathrm{T}} <$ 200 GeV		
r	$\langle\Psi_b(r)\rangle$ [b-jets]	$\langle\Psi_l(r)\rangle$ [light jets]
0.04	0.606 ± 0.018 (stat.) $^{+0.038}_{-0.043}$ (sys.)	0.699 ± 0.056 (stat.) $^{+0.073}_{-0.083}$ (sys.)
0.08	0.801 ± 0.011 (stat.) $^{+0.020}_{-0.022}$ (sys.)	0.788 ± 0.029 (stat.) $^{+0.060}_{-0.034}$ (sys.)
0.12	0.874 ± 0.007 (stat.) $^{+0.008}_{-0.012}$ (sys.)	0.874 ± 0.022 (stat.) $^{+0.044}_{-0.019}$ (sys.)
0.16	0.916 ± 0.005 (stat.) $^{+0.006}_{-0.007}$ (sys.)	0.923 ± 0.015 (stat.) $^{+0.034}_{-0.018}$ (sys.)
0.20	0.944 ± 0.003 (stat.) $^{+0.005}_{-0.006}$ (sys.)	0.941 ± 0.011 (stat.) $^{+0.018}_{-0.011}$ (sys.)
0.24	0.964 ± 0.002 (stat.) $^{+0.005}_{-0.005}$ (sys.)	0.962 ± 0.008 (stat.) $^{+0.011}_{-0.007}$ (sys.)
0.28	0.978 ± 0.002 (stat.) $^{+0.003}_{-0.004}$ (sys.)	0.981 ± 0.004 (stat.) $^{+0.009}_{-0.005}$ (sys.)
0.32	0.988 ± 0.001 (stat.) $^{+0.002}_{-0.002}$ (sys.)	0.987 ± 0.003 (stat.) $^{+0.005}_{-0.002}$ (sys.)
0.36	0.994 ± 0.001 (stat.) $^{+0.001}_{-0.001}$ (sys.)	0.997 ± 0.001 (stat.) $^{+0.002}_{-0.001}$ (sys.)
0.40	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)	1.000 ± 0.000 (stat.) $^{+0.000}_{-0.000}$ (sys.)

Table 4.50: Integrated shapes $\langle\Psi(r)\rangle$ for light and b-jets in both channels for 150 GeV $< p_{\mathrm{T}} <$ 200 GeV

Chapter 5

Determination of the b-quark mass from the jet shape data.

This chapter is dedicated to the precise determination of the b-quark mass using the jet shape data described in the previous chapter. The results obtained in this chapter have been published in Ref. [97]

5.1 Introduction and motivation

The theory of angular screening effects described in Sect. 1.7 suggests that the heavier mass of the b-quark is the key feature of the observations made in the previous chapter, namely the wider energy distribution of b-quark jets over light-quark jets. The greater the b-quark mass, the wider the region spanned under the angle $\theta_0 = m/E$, where the radiation is highly suppressed and therefore, the particle content. This discussion makes the jet shape data suitable for a determination of the b-quark mass parameter for the splitting functions in shower Monte Carlos, as described in Ref. [98]. Additionally, jet shapes do also depend on the coupling strength of the strong interaction, which will also be studied in the following sections.

5.2 Monte Carlo predictions

Top-quark pair events have been generated using the PYTHIA 6.4 program. Additionally, the MSTJ(42)=3 switch has been used to take into account the larger mass of the b-quark on the angular distribution of the decay products [98]. Also, the switch MSTJ(43)=3 has been used to set the fragmentation variable z as the fraction of energy in the centre-of-mass frame of the showering partons [22]. Jet shapes naturally depend on the strong coupling constant α_s, as it controls the radiation emitted by strongly-interacting partons, and have been in fact a precise way to determine its value in Ref. [67]. Therefore, one needs to take this effect into account for a precise determination of the b-quark mass. At the one-loop order, the scale dependence of the strong coupling can be parametrised by [6]

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \log\left(\frac{Q^2}{\Lambda^2}\right)}; \quad \beta_0 = \frac{1}{4\pi}\left(11 - \frac{2}{3}n_f\right) \tag{5.1}$$

Eq. 5.1 incorporates the QCD scale Λ, which can be varied for the PYTHIA time-like parton showers arising from a resonant decay using the PARJ(81) switch. Finally, the b-quark mass m_b is varied around its nominal value $m_b = 4.8$ GeV using the PMAS(5) and PARF(105) switches, which control the kinematical mass of the b-quark and its constituent mass, respectively. Additionally, $t\bar{t}$ samples have been generated using the HERWIG++ Monte Carlo program [99]. The differences between the value of m_b obtained in HERWIG++ and that obtained using PYTHIA will be discussed later, and assigned as a theoretical uncertainty.

5.3 Jet selection and jet shape calculation

The final-state particles from the PYTHIA simulation are clustered using the anti-k_t algorithm [32] as implemented in FASTJET [34], with a radius parameter $R = 0.4$. As specified before, muons and neutrinos are left out of the clustering algorithm.

All jets with transverse momentum $p_{\mathrm{T}} > 30$ GeV are pre-selected. To select the jets induced by b-quarks from the top decays, a matching procedure is used between the clustered jets and any hadron containing b-quarks. If one of these

hadrons with $p_{\mathrm{T}} > 5$ GeV is found at a distance $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ from the axis of a given jet, this jet is selected as a b-jet. Alternatively, light-quark jets are selected as the pair of jets which, not containing B-hadrons closer than $\Delta R = 0.3$ to the jet axis, have the closest invariant mass to the nominal W boson mass $m_W = 80.4$ GeV.
The differential jet shape is then calculated for both samples following the formula

$$\langle \rho(r) \rangle = \frac{1}{\Delta r} \frac{1}{N_{\mathrm{jets}}} \sum_{\mathrm{jets}} \frac{p_{\mathrm{T}}(r - \Delta r/2, r + \Delta r/2)}{p_{\mathrm{T}}(0, R)} \tag{5.2}$$

5.4 Analysis procedure

As b-jet shapes depend on both the parton shower QCD scale Λ_s and the b-quark mass m_b, both need to be determined for a precise result. A simultaneous determination of both parameters is not possible because a variation of one of them can be compensated by an opposite variation of the other one, leading to a set of degenerate minima in the plane (m_b, Λ_s). However, it is expected that the light-jet shapes depend only in Λ_s and not in m_b. Therefore, one can determine the parameter Λ_s from the light-jet shapes and use it for the extraction of m_b from the b-jet data.
The method used for the extraction of a physical parameter $\beta = \Lambda_s, m_b$ from a theoretical distribution scan relies on the minimisation of a standard χ^2 for each p_{T} bin using MINUIT [102]. The χ^2 function is defined in a way which takes into account the correlations between the experimental uncertainties via a set of nuisance parameters $\{\lambda_i\}$. In terms of the parameter β to be extracted and the nuisance parameter vector $\vec{\lambda}$, it can be written as

$$\chi^2(\beta; \vec{\lambda}) = \sum_k \frac{(x_k - F_k(\beta; \vec{\lambda}))^2}{\Delta x_k^2 + \Delta \tau_k^2} + \sum_i \lambda_i^2 \tag{5.3}$$

$$F_k(\beta; \vec{\lambda}) = \phi_k(\beta) \left(1 + \sum_i \lambda_i \sigma_{ik}\right) \tag{5.4}$$

In Eq. 5.3, the index k runs over all r bins in a given p_{T} bin, with a given value x_k of the jet shape and with statistical uncertainty Δx_k. Here, $\Delta\tau_k$ represents the

statistical uncertainty on the theoretical predictions. The nuisance parameters λ_i, one for each source of uncertainty, are also involved in Eq. 5.4, where the functions $\phi_k(\beta)$ correspond to the nominal dependence of the jet shape with the parameter β in bin k. They are parametrised in terms of a parabola throughout this paper. Finally, σ_{ik} are the relative uncertainties for source i in the bin k, discussed in Sect. 4.9.3.

Each nuisance parameter corresponds to a different uncertainty on the data. Table 5.1 shows the identification of each λ_i with the corresponding source, ordered from larger to smaller impact.

Nuisance parameter	Source of uncertainty	Impact on data
λ_1	Pileup	$2\% - 10\%$
λ_2	Cluster systematics	$2\% - 10\%$
λ_3	Unfolding-modelling	$1\% - 8\%$
λ_4	Jet energy scale	$\simeq 5\%$
λ_5	Jet energy resolution	$\simeq 5\%$
λ_6	JVF	$< 1\%$

Table 5.1: Identification of the nuisance parameters λ_i with the sources of experimental uncertainty.

5.5 Determination of the parton shower scale Λ_s

In order to determine the QCD scale of the parton shower Monte Carlo which best fits the jet shape data, the dependence of the light-quark jet shapes on Λ_s is studied. Figure 5.1 shows the comparison of the light-jet shape data and the Pythia expectations for several values of Λ_s. The dependence of the jet shapes on Λ_s is clearly seen from the figure.

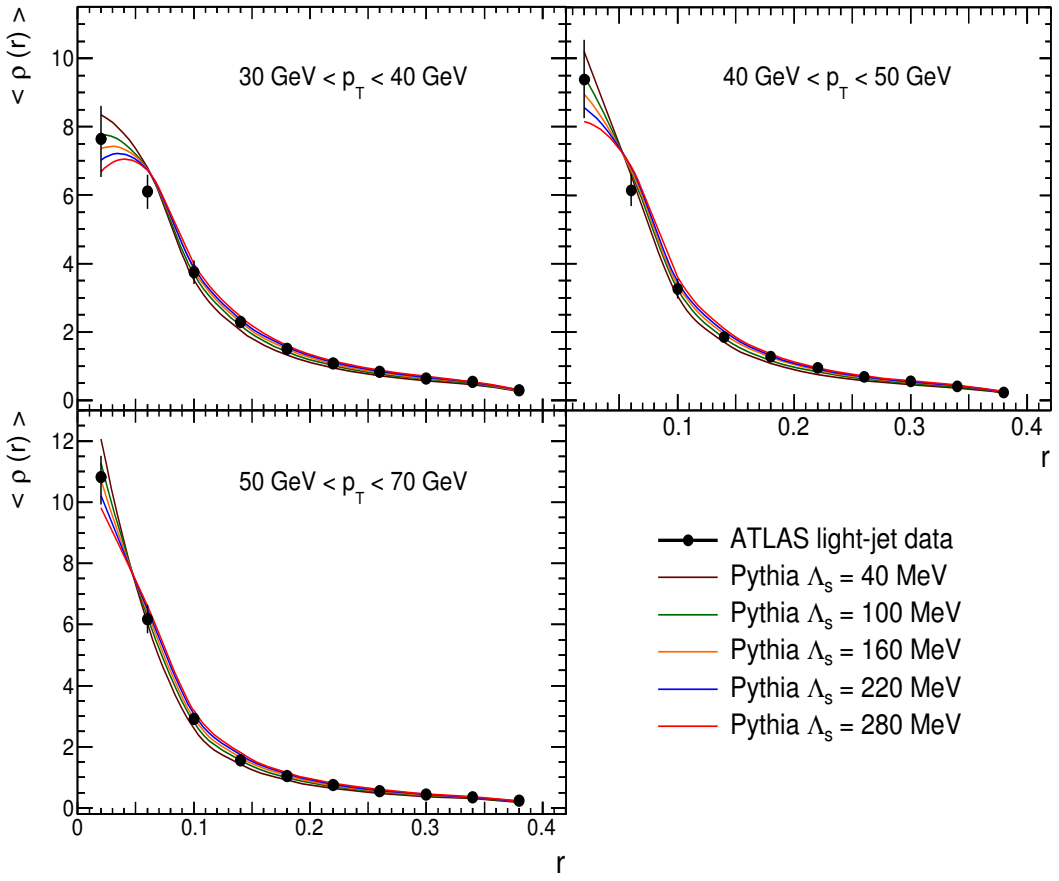

Figure 5.1: Results of the Λ_s scan compared to the light-jet data.

In order to parametrise this dependence and obtain the interpolating functions $\phi_k(\Lambda_s)$ in Eq. 5.4, samples with Λ_s varying from 20 MeV to 300 MeV in steps of 20 MeV have been generated. To illustrate this dependence, Figure 5.2 shows the points obtained from this scan together with the fitted functions $\phi_k(\Lambda_s)$ for $r = 0.02$ in each p_{T} bin.

The fits using Eqs. 5.3 and 5.4 have been performed for every p_{T} bin separately, and finally all of them are combined into a global fit to the three bins with 30 GeV $< p_{\mathrm{T}} <$ 70 GeV. Fig. 5.3 shows the values of the nuisance parameters $\{\lambda_i\}$ involved in the fit, as well as the correlation matrix between them. The

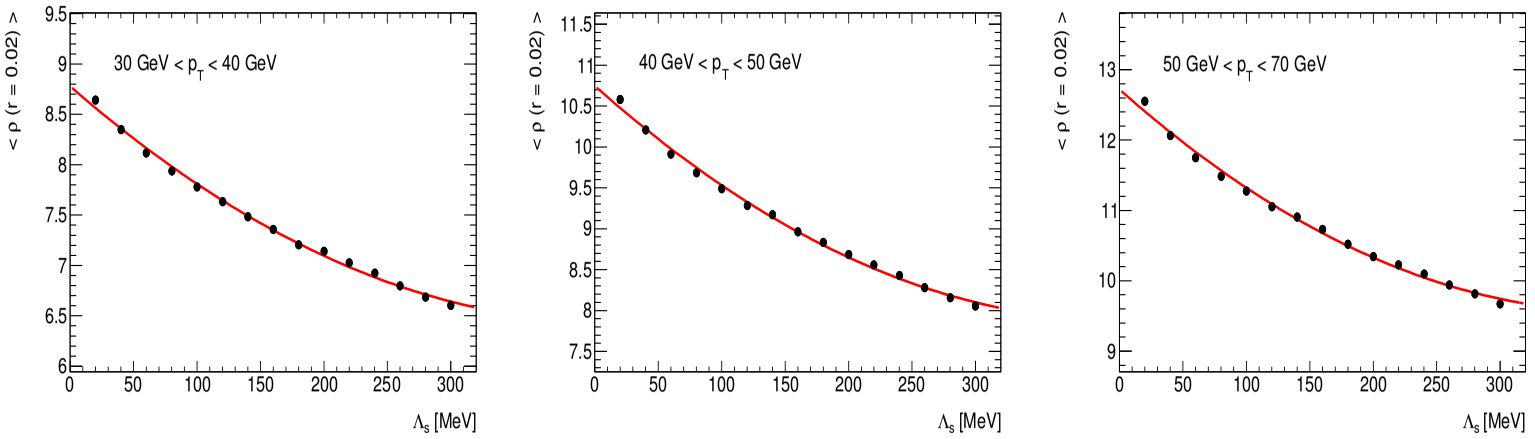

Figure 5.2: Dependence of the light-quark jet shape $\langle\rho(r = 0.02)\rangle$ with the parton shower scale Λ_s for the p_{T} intervals 30 GeV $< p_{\mathrm{T}} <$ 40 GeV (left), 40 GeV $< p_{\mathrm{T}} <$ 50 GeV (middle) and 50 GeV $< p_{\mathrm{T}} <$ 70 GeV (right), together with the interpolating functions $\phi_k(\Lambda_s)$.

values of the nuisance parameters are always compatible with the $\pm 1\sigma$ band, fact which gives us confidence on the quality of the fit. The results of the fits to Λ_s are summarised in Table 5.2, together with the fit uncertainties and the values of χ^2/N_{dof}.

Bin	Λ_s value (MeV)	Fit error (MeV)	χ^2/N_{dof}
30 GeV $< p_{\text{T}} <$ 40 GeV	187.5	24.0	10.6 / 9
40 GeV $< p_{\text{T}} <$ 50 GeV	193.5	24.2	11.0 / 9
50 GeV $< p_{\text{T}} <$ 70 GeV	137.7	17.3	7.8 / 9
Global fit	162.1	9.6	39.0 / 29

Table 5.2: Summary of the results of the fit for Λ_s using the light-jet shape data.

The nominal results obtained here have been derived using the one-loop solution to the renormalisation group equation (RGE) for the PYTHIA parton shower. In addition, the values of Λ_s have also been extracted using HERWIG++ with the solutions to the RGE implemented up to two loops. The resulting value at one loop is $\Lambda_s = 160.7 \pm 15.3$ MeV, in good agreement with the nominal value quoted above. For the two-loop case, the expression for the running strong coupling is [6]

$$\alpha_s(Q^2) = \frac{1}{\beta_0 \log x}\left[1 - \frac{\beta_1}{\beta_0^2}\frac{\log(\log x)}{\log x}\right]; \quad x = \frac{Q^2}{\Lambda^2} \tag{5.5}$$

where β_0 is given in Eq. 5.1 and $\beta_1 = \frac{1}{(4\pi)^2}\left(102 - \frac{38}{3}n_f\right)$. In this case, a value of $\Lambda_s = 276.1 \pm 17.3$ MeV is obtained, which is compatible with the value quoted in Ref. [103] within the uncertainties. The effect of the two-loop running of the shower α_s on the b-quark mass will be explained in Sect. 5.7.5.

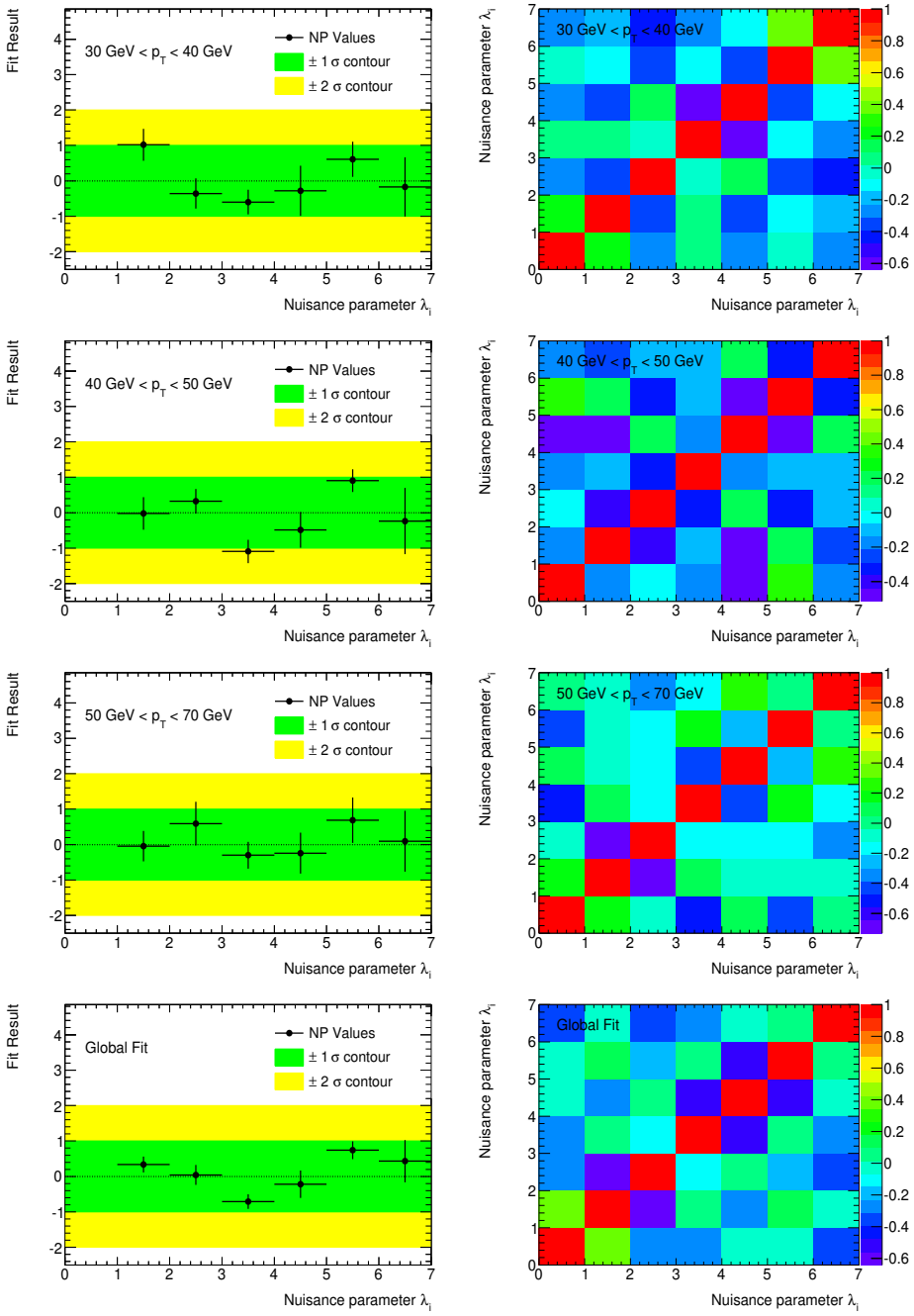

Figure 5.3: Results for the nuisance parameters involved in the Λ_s extraction (left column) and correlation matrices between them (right column) for each p_{T} bin considered. The results obtained for the global fit are shown at the bottom row.

5.6 Determination of the b-quark mass m_b

Once the parton shower scale Λ_s is determined, the result is used to generate Pythia samples with several different values of the b-quark mass. The scan is performed in this case by varying this parameter from 4.0 GeV to 6.0 GeV in steps of 250 MeV. Figure 5.4 shows the comparison of the b-jet shape data with the expectations from Pythia for several values of m_b, including those for $m_b = 3.0$ and 7.0 GeV. The value of the parton shower scale for the predictions shown in Fig. 5.4 is the one corresponding to the global fit to the light-jet shape data,

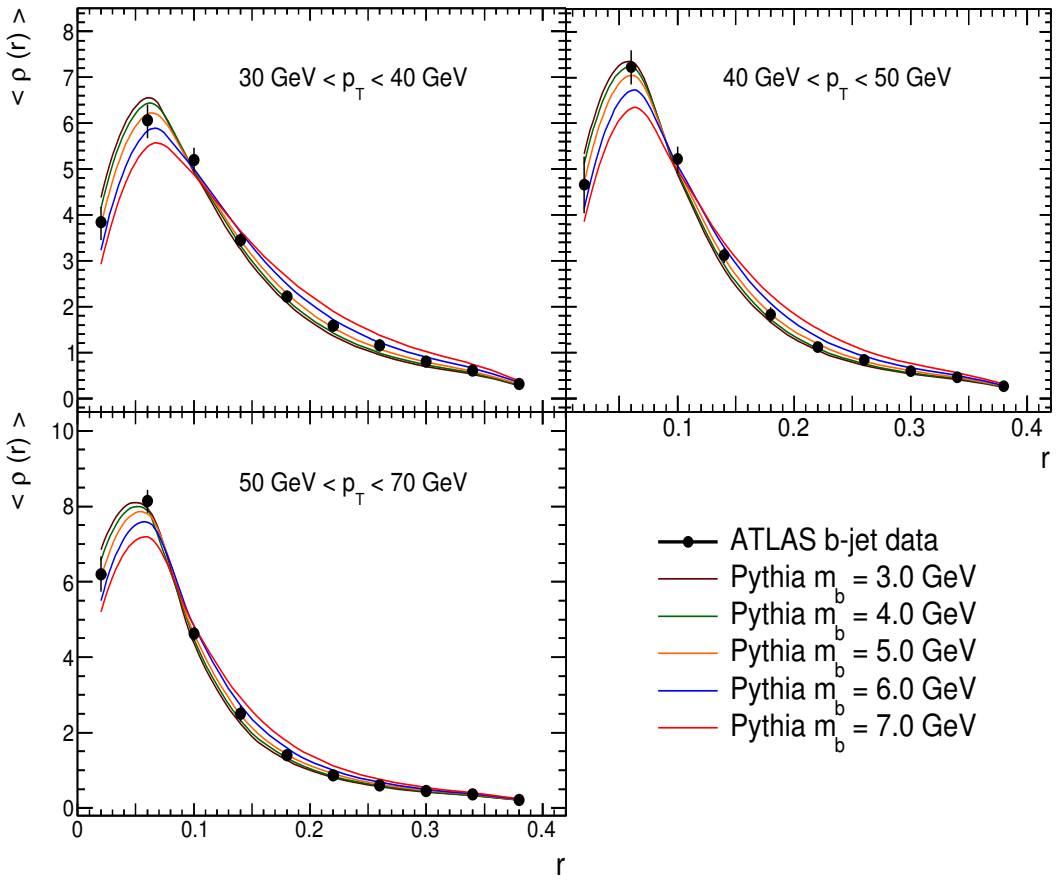

Figure 5.4: Results of the m_b scan compared to the b-jet data. The QCD scale involved in the parton shower has been taken to be the one corresponding to the global fit to the light-jet data, $\Lambda_s = 162.1 \pm 9.6$ MeV

The b-jet shapes show a turn-over close to the jet cores due to the angular screening caused by the heavy mass of the b-quark. The larger m_b is, the wider is the jet in the sense that the inner core has smaller energy deposits. The description of all bins provided by PYTHIA is excellent, showing that it is possible to perform a safe fit to the data. The parametrisation of the interpolating functions $\phi_k(m_b)$ describing the dependence of the differential b-jet shapes with the b-quark mass is done using second-order polynomials as in the case of the shower scale Λ_s. Figure 5.5 shows the dependence for $r = 0.02$ in each p_{T} bin, as predicted by PYTHIA The global fit has been performed including all p_{T} bins and using the global value of the parton shower scale, $\Lambda_s = 162.1 \pm 9.6$ MeV. As a cross-check, for each p_{T} bin, extra samples have been generated using the partial values of Λ_s shown in Table 5.2. The agreement between all the extracted values of m_b is excellent, as can be seen in Table 5.3.

Bin	m_b value (GeV)	Fit error (GeV)	χ^2/N_{dof}
30 GeV $< p_{\mathrm{T}} <$ 40 GeV	5.00	0.14	8.28 / 9
40 GeV $< p_{\mathrm{T}} <$ 50 GeV	4.82	0.19	10.41 / 9
50 GeV $< p_{\mathrm{T}} <$ 70 GeV	4.82	0.13	11.99 / 9
Global fit	4.86	0.08	43.04 / 29

Table 5.3: Summary of the results of the fits for m_b using the b-jet shape data and the corresponding value of Λ_s for each bin listed in table 5.2. The global fit is performed using the globally extracted value of the parton shower scale $\Lambda_s = 162.1$ MeV.

As before, the values of the nuisance parameters and the correlation matrices between them are shown in Figure 5.6 for the fits performed using each extracted value of the shower scale. As in the previous case, we find that the nuisance parameters are well behaved, being always compatible with the $\pm 1\sigma$ contour band. This is specially important for the global fit, as its result will be taken as the central value for our determination. As can be seen in the lower part of Figure 5.6, the behaviour of the fit parameters is very good.

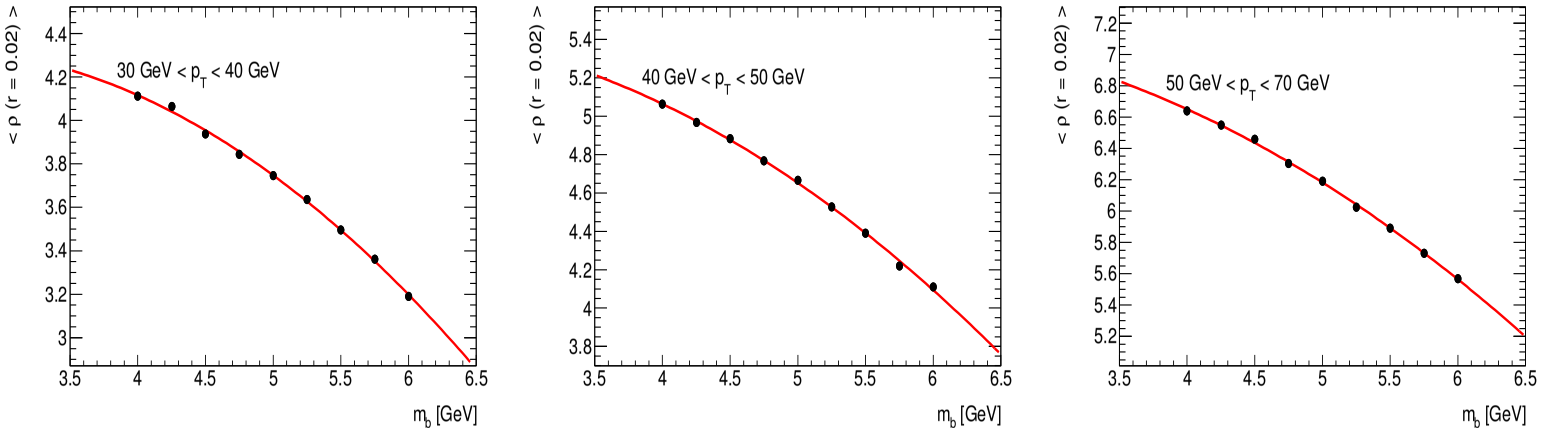

Figure 5.5: Dependence of the b-quark jet shape $\langle\rho(r = 0.02)\rangle$ with the b-quark mass m_b for the p_{T} intervals 30 GeV $< p_{\mathrm{T}} <$ 40 GeV (left), 40 GeV $< p_{\mathrm{T}} <$ 50 GeV (middle) and 50 GeV $< p_{\mathrm{T}} <$ 70 GeV (right), together with the interpolating functions $\phi_k(m_b)$.

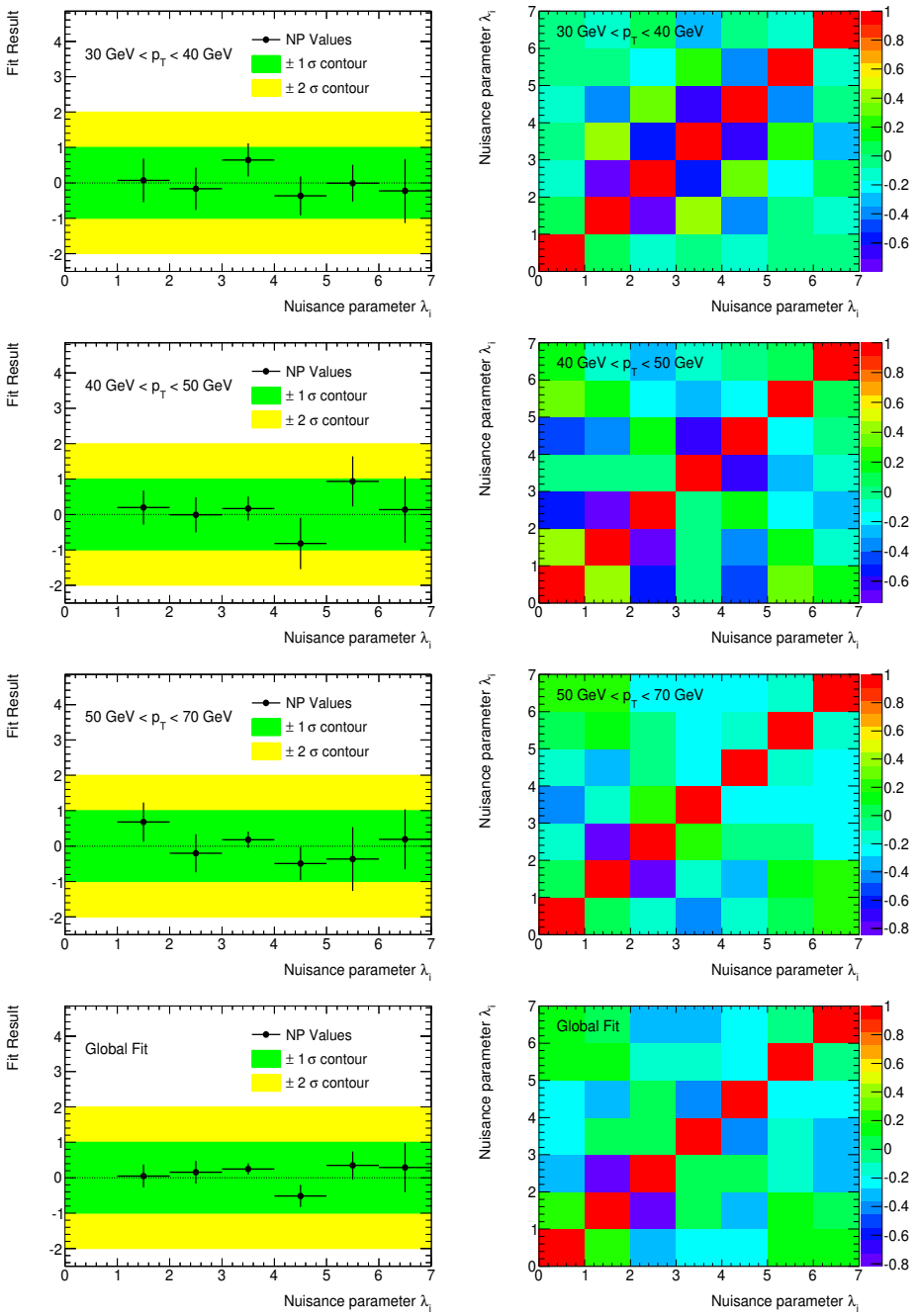

Figure 5.6: Results for the nuisance parameters involved in the m_b extraction (left column) and correlation matrices between them (right column) for each p_{T} bin considered. The results obtained for the global fit are shown at the bottom row. For each p_{T} bin, the corresponding value of Λ_s listed in Table 5.2 has been used. For the global fit, the globally extracted value is $\Lambda_s = 162.1$ MeV.

5.7 Theoretical uncertainties

In this section, the uncertainties on the theory are discussed. They come from several sources, including the modelling of the parton shower, hadronisation and multiple parton interactions. Other effects such as the amount of initial and final-state radiation, the colour reconnection model and the error on the determination of the parton shower scale Λ_s are also studied. The generator modelling uncertainty is the main source of uncertainty for these analysis, not being greater than 400 MeV in terms of the extracted b-quark mass. All variations of the theoretical distributions are performed with respect to the nominal sample, produced using the fitted values of Λ_s and m_b.

5.7.1 Generator modelling

The Pythia predictions use virtuality-ordered parton showers and the Lund string model for the hadronisation. In order to study the impact of this choice on the extraction of m_b, a sample of $t\bar{t}$ events has been generated using the Herwig++ Monte Carlo program [99], which incorporates angular-ordered parton showers as well as the cluster hadronisation model. The modelling of the underlying event (multiparton interactions) is also different between both approaches. For Herwig++, the LHC-UE7-2 tune has been chosen. This is based on the ATLAS measurements of the underlying event using charged particles [104]. On the other hand, Pythia uses the so-called Tune A as default [105], which is based on the correct description of many Tevatron measurements.
In Figure 5.7, the nominal prediction by Pythia is compared to the nominal predictions by Herwig++. In order to study the impact of these differences on the determination of m_b, the full analysis has been repeated using Herwig++. In this case, the b-quark mass is scanned by varying both the NominalMass and the ConstituentMass flags for /Herwig/Particles/b and /Herwig/Particles/bbar. The value of Λ_s has been set to 160.7 ± 15.3 MeV, which is the value obtained in Section 5.5 for the Herwig++ approach. The result for the b-quark mass is $m_b = 5.25 \pm 0.09$ GeV, and the difference with respect to the nominal value is symmetrised and ascribed as a theoretical uncertainty.

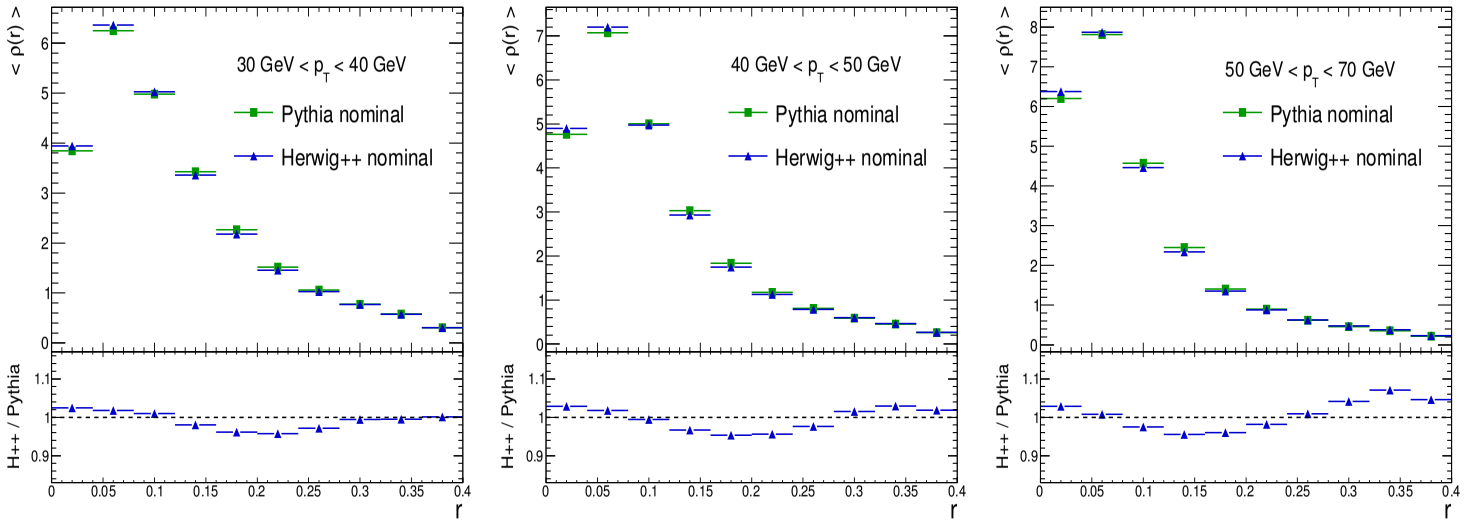

Figure 5.7: The difference between the PYTHIA and HERWIG++ predictions for the fitted values $m_b = 4.86$ GeV and $\Lambda_s = 162.1$ MeV. This difference, arising from the parton shower and hadronisation, is the source of the theoretical uncertainty due to the generator modelling.

5.7.2 Initial-state radiation

The amount of initial-state radiation (ISR) can lead to differences in the jet shapes. To test this effect, two additional samples with reduced and enhanced levels of ISR are generated. The ISR is controlled in Pythia using the parameters PARP(67) and PARP(64). To decrease the ISR, the parameters are set to 0.5 and 4.0 respectively. To increase ISR, they are set to 6.0 and 0.25, respectively. These specifications have been widely used in several ATLAS analyses such as the study of $t\bar{t}$ production with a veto on central jet activity [100]. The effects of these changes on the prediction and the comparison to the nominal Pythia sample are shown in Fig. 5.8. The effect of these variations on the b-quark mass is around 20 MeV, which is negligible for the final result, compared to the generator uncertainty.

5.7.3 Final-state radiation

The effect of the amount of final-state radiation (FSR) on the b-jet shape distributions is studied by varying the parameters PARP(72) and PARJ(82). These two parameters represent the value of Λ_{QCD} in the time-like showers responsible of the FSR (not arising from a resonant decay), and the infrared invariant mass cutoff, below which partons are not assumed to radiate. To increase the levels of FSR, these values are set to 0.384 and 0.5, respectively. To decrease the FSR activity, they are set to 0.096 and 2.0, respectively. This represents a change of a factor of 2 with respect to their nominal values 0.192 and 1.0. Fig. 5.9 shows the effect of these variations on the b-jet shapes, as well as the ratio to the nominal Pythia prediction. The impact of the FSR on the extracted b-quark mass is around 180 MeV.

5.7.4 Colour reconnection

The effect of the modelling of the colour reconnection (CR) between final-state partons is studied by using the ACR [101] tuning of the Pythia Monte Carlo. This tune incorporates a new colour reconnection model, which assumes an enhanced amount of colour connections between partons with respect to the nominal

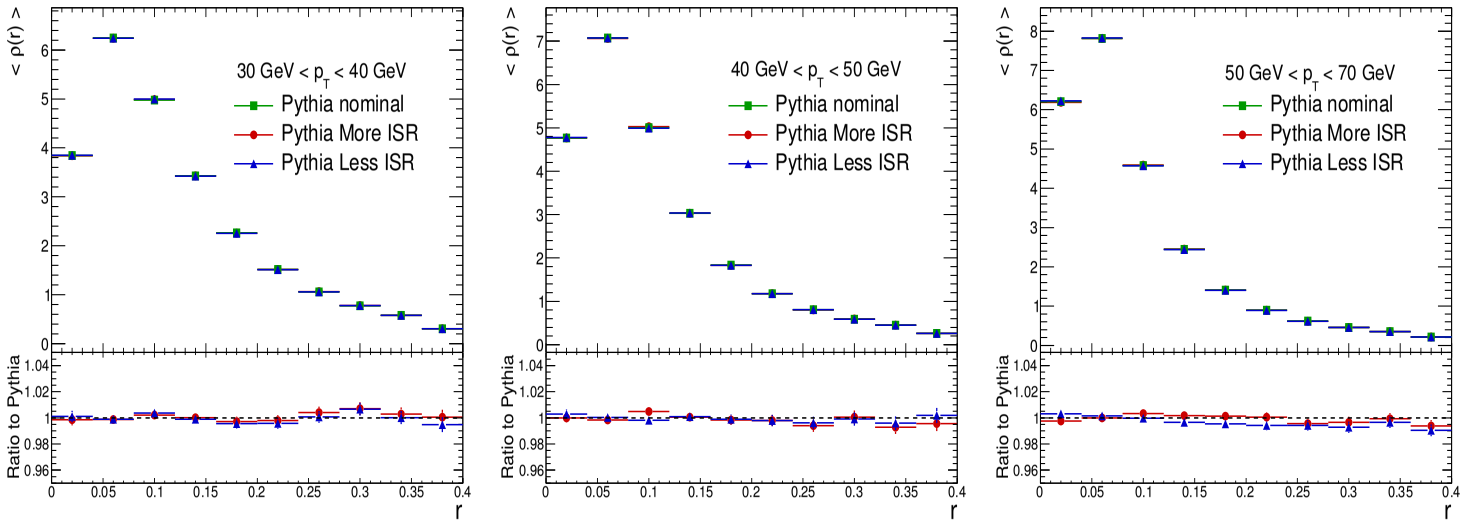

Figure 5.8: The effects of the initial-state radiation on the b-jet shapes.

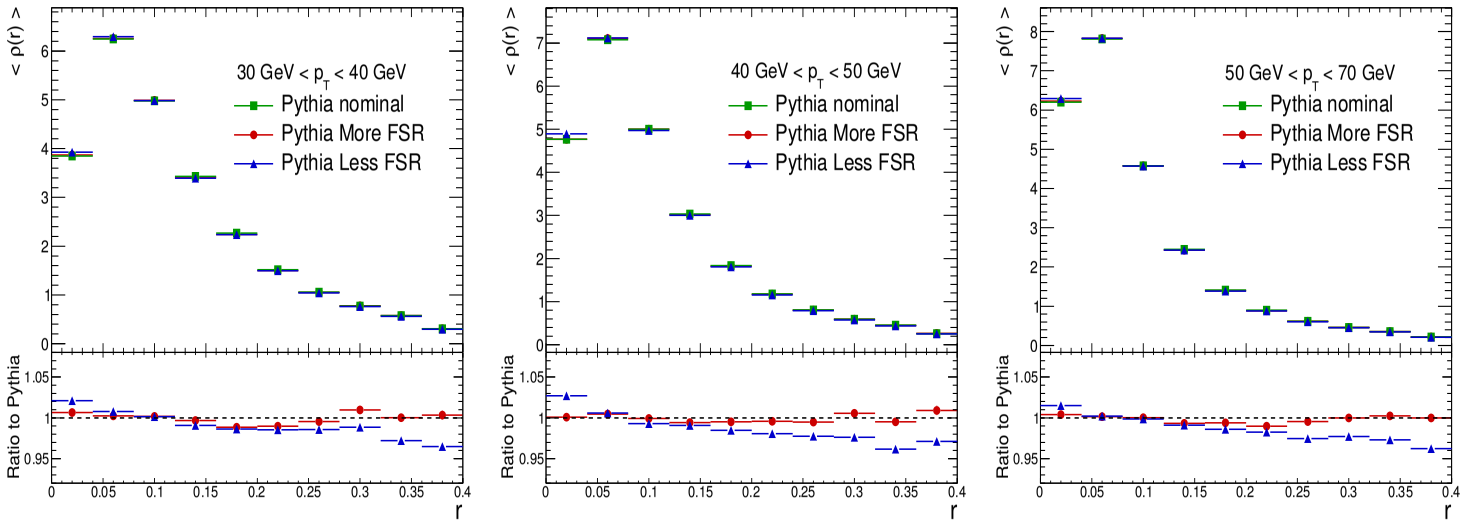

Figure 5.9: The effects of the final-state radiation on the b-jet shapes.

Tune A sample. Figure 5.10 shows that the effect of the new CR modelling is to increase the energy deposit on the jet cores on about 2%. The impact on the b-quark mass is estimated by multiplying the nominal predictions by the ratio Tune ACR/Tune A, and it has an effect of around 170 MeV.

5.7.5 Uncertainty on the Λ_s determination.

The effects of the uncertainties in the determination of the parton shower scale have been also studied. To this end, the full set of m_b variations have been generated again using the values of Λ_s which define the envelope of its determination. Because the value obtained in section 5.5 was $\Lambda_s = 162.1 \pm 9.6$ MeV, the full scan on m_b variations has been repeated using the values $\Lambda_s = 152.5$ MeV and $\Lambda_s = 171.7$ MeV, which define the endpoints of the interval in which Λ_s can vary due to its experimental uncertainty. This is done in this way, instead of simply shifting each theoretical prediction by the nominal variation on the jet shapes due to this effect because the jet shapes are highly dependent on both parameters m_b and Λ_s at the same time. To keep track of this correlation, the full set of theoretical predictions has to be recalculated.
The fits with the varied values of Λ_s are then repeated, and the differences between both of them and the central value are taken as the systematic uncertainties on the b-quark mass, which are in principle asymmetric. It is found that the impact on m_b of the determination of the parton shower scale is around 60 MeV at maximum, which represents the 1.2% of the b-quark mass.
Another source of uncertainty related to the way in which Λ_s is determined arises from the fixed order at which the running of α_s is evaluated. To estimate this uncertainty, the value of m_b has been extracted using the running of $\alpha_s(Q^2)$ up to two loops, which is implemented for the Herwig++ parton shower. The value of the two-loop shower scale was determined to be $\Lambda_s = 276.1 \pm 17.3$ MeV, and the corresponding value of the b-quark mass is $m_b = 5.39 \pm 0.08$ GeV. This value is to be compared with the value obtained for the one-loop running coupling in Herwig++, which was $m_b = 5.25 \pm 0.09$ GeV, and therefore gives a relative uncertainty of 2.7%. For the nominal value of $m_b = 4.86$ GeV, this represents an additional uncertainty of 0.13 GeV, to be added in quadrature to the result of the

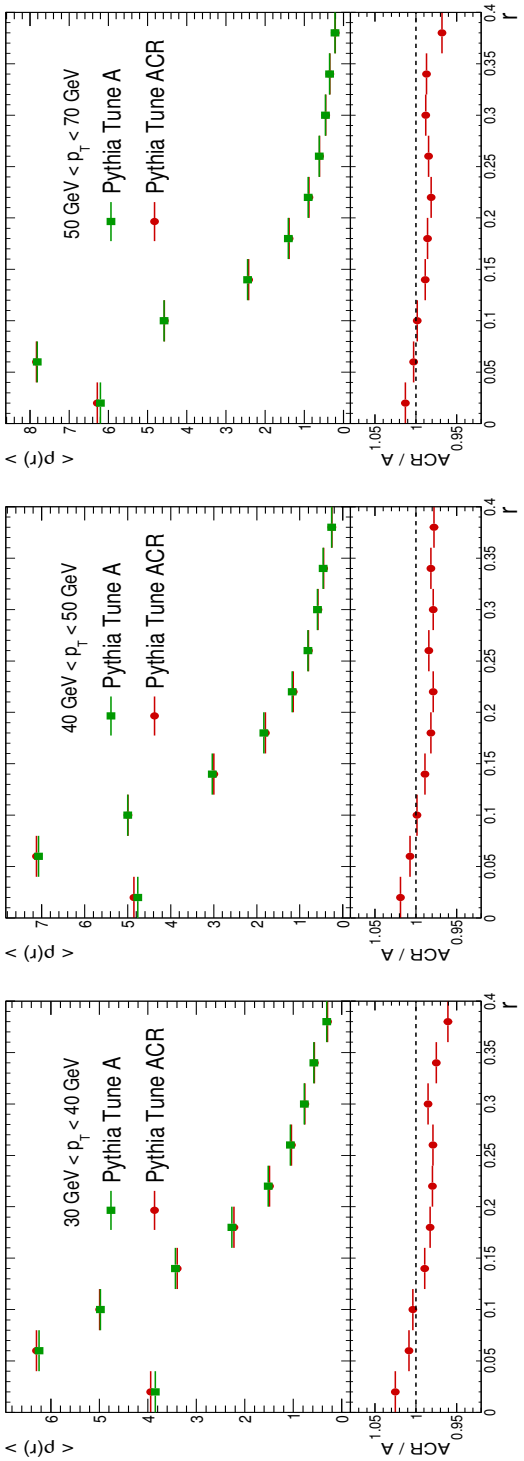

Figure 5.10: The effects of the colour-reconnection modelling on the b-jet shapes.

propagation of the experimental uncertainty in Λ_s, and therefore has a maximum value of 0.14 GeV.

Discrepancies between data and MC on the description of the transverse momentum of light jets can lead to a biased result on the value of Λ_s. In order to check such effect, the light jet shapes were weighted and the fits were redone. These weights were estimated, in a very conservative way, matching the shape of the p_{T} distributions of light and b-jets. The differences on Λ_s with respect to the nominal value were found to be small ($\sim 3\%$). This difference is perfectly covered by the error on the fit procedure ($\sim 6\%$), thus ensuring the robustness of the Λ_s determination.

As a further cross-check on the way in which Λ_s is propagated throughout the analysis, the parton shower scale has been determined using the b-jet shapes obtained with the fitted value of m_b. The results are found to be fully compatible with the previous results obtained in Table 5.2, which reassures us on the extrapolation of Λ_s from light-jets to b-jets.

After the evaluation of the theoretical uncertainties, the final value of the b-quark mass obtained in this analysis can be expressed as

$$m_b = 4.86 \pm 0.08\ (\mathrm{exp.})\ \ \pm 0.39\ (\mathrm{Gen.})\ {}^{+0.02}_{-0.01}\ (\mathrm{ISR})\ {}^{+0.18}_{-0.00}\ (\mathrm{FSR}) \\ {}^{+0.17}_{-0.00}\ (\mathrm{CR})\ {}^{+0.14}_{-0.13}\ (\mathrm{PS\ scale})\ . \tag{5.6}$$

5.8 Summary and conclusions

This study presents a determination of the mass of the b-quark using jet substructure techniques. It is found that the angular screening effects which were predicted in [41, 42] are confirmed and consistent with a reasonable value of the b-quark mass parameter. The dead cone effect was similarly exploited in [106] to determine the c-quark mass in ep collisions at HERA.

Experimental uncertainties have been propagated using nuisance parameters for each source of uncertainty. This ensures that the correlations between all sources are explicitly taken into account. Systematic effects on the theoretical distributions have also been studied. The modelling of the jet shapes by different Monte Carlo generators is the main uncertainty on this analysis, accounting for an 8%

impact on the final value for m_b. Other systematic effects on the theoretical predictions have been studied, such as the amount of initial and final-state radiation, the colour reconnections and the uncertainty on the determination of the parton shower scale Λ_s. Our final result reads

$$m_b = 4.86 \pm 0.08\ (\text{exp.})\ \ \pm 0.39\ (\text{Gen.})\ {}^{+0.02}_{-0.01}\ (\text{ISR})\ {}^{+0.18}_{-0.00}\ (\text{FSR})$$
$${}^{+0.17}_{-0.00}\ (\text{CR})\ {}^{+0.14}_{-0.13}\ (\text{PS scale})\ . \tag{5.7}$$

Figure 5.11 shows the projection of the χ^2 function on the m_b axis, normalized to the number of degrees of freedom, for the central value of the determined mass.

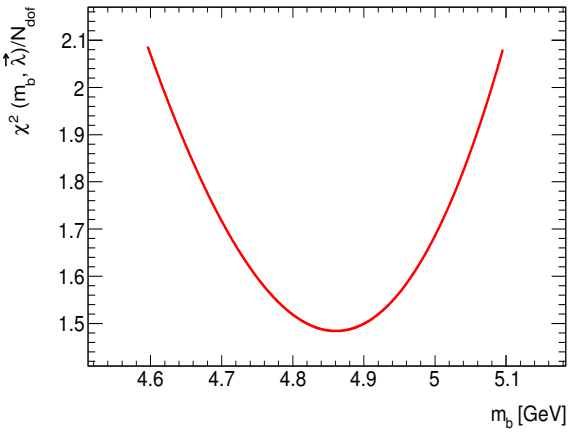

Figure 5.11: The m_b-projection of the $\chi^2(m_b, \vec{\lambda})$ function normalized to the number of degrees of freedom N_{dof}.

Although there is a significant numerical similarity of this value with the value of the pole mass quoted by the Particle Data Group in [6] and also with the values obtained by the LEP Collaborations in Refs. [107] and [108], the value extracted here should not be confused with the QCD pole mass of the b-quark. It should rather be regarded as the on-shell mass parameter affecting the parton shower kinematics, as calculated in Ref. [98]. In any case, it would be theoretically very interesting to define a way in which the Monte Carlo masses for hadronising quarks can be related to the poles of their respective fermionic propagators.

Chapter 6

Measurement of transverse energy-energy correlations and determination of $\alpha_s(m_Z)$

This chapter is dedicated to the measurement of transverse energy-energy correlations in multijet events and the determination of the strong coupling constant from a χ^2 fit of these data to next-to-leading order QCD predictions worked out in Ref. [36]. The results obtained in this chapter have been submitted for publication to Phys. Lett. B at the time of writing this thesis.

6.1 Introduction and motivation

Transverse energy-energy correlations (TEEC, see Sect. 1.5) provide a precise test of perturbative QCD and a way to determine the strong coupling constant α_s in ATLAS. This is so because of the high precision of the calorimetric measurement of the jet angular coordinates η, φ and the cancellation of the systematic uncertainties associated to the jet energy measurement due to the $x_{\mathrm{T}i}x_{\mathrm{T}j}$ weighting. The asymmetry ATEEC, also defined in Sect 1.5 has the experimental advantage of reducing uncertainties which are constant in ϕ, i.e. isotropic effects and also reducing theoretical uncertainties due to the choice of the renormalisation and factorisation scales. This makes these tools excellent allies for a precise determi-

nation of the strong coupling constant.

6.2 Monte Carlo samples

For this analysis, two different Monte Carlo approaches are used, depending on whether the underlying hard process is considered to be $2 \to 2$ or $2 \to n$. The generated events are then processed with the ATLAS full detector simulation, based on the GEANT4 software (see [75]).

The baseline MC samples used here are generated with PYTHIA [22], with the amplitudes calculated at leading order using the MRST2007LO* PDFs (see [77]) and matched to p_{T} ordered parton showers with the AUET2B tune [138, 139] for the modelling of the underlying event. The hadronisation follows the Lund string model. Since the $\hat{p}_{\mathrm{T}}$ distribution of the basic $2 \to 2$ underlying QCD process is known to be a very steeply falling function, the events are generated in eight $\hat{p}_{\mathrm{T}}$ bins and each sample is weighted according to the corresponding cross-sections. For the sake of completeness, the files used are summarised in Table 6.1. These samples have been thoroughly used by the ATLAS Collaboration for jet calibration and production purposes [57, 140]

Channel	Sample	$\hat{p_{\mathrm{T}}}$(GeV)	σ (nb)	Events
207011	J4b	220 - 280	7.31×10^{0}	2744973
105014	J5	280 - 560	2.60×10^{0}	2796405
105015	J6	560 - 1120	3.55×10^{-2}	2791826
105016	J7	1120 - 2240	1.33×10^{-4}	1398937
105017	J8	2240 - ∞	5.68×10^{-9}	1397377

Table 6.1: The PYTHIA samples used for the analysis

Samples generated with HERWIG++ (see [24]) have also been used, using the parton distribution functions CTEQ6.6 (see [76]). Multiparton interactions have been taken into account according to [25]. Here the overall approach is similar, the main subtle differences come from the fact that HERWIG++ uses angular ordered parton showers, a cluster hadronisation scheme and its own underlying

event parametrisation. The HERWIG++ samples are generated in the same $\hat{p}_{\mathrm{T}}$ bins as in the PYTHIA case, and they are known to describe the gross features of the multihadronic final states both at LEP and the Tevatron. As in the previous case, the samples are enumerated in table 6.2. These Monte Carlo samples have been used by the ATLAS Collaboration together with the ones discussed in the preceeding paragraph to estimate systematic uncertainties in the modelling of the hadronic final states.

Channel	Sample	$\hat{p_{\mathrm{T}}}$(GeV)	σ (nb)	Events
207010	J4b	220 - 280	7.00×10^{0}	1999480
113209	J5	280 - 560	2.48×10^{0}	999689
113210	J6	560 - 1120	3.41×10^{-2}	982772
113211	J7	1120 - 2240	1.28×10^{-4}	998356
113212	J8	2240 - ∞	4.55×10^{-9}	949940

Table 6.2: The HERWIG samples used for the analysis

Another different approach to simulate multijet final states is followed by ALPGEN [28]. This approach is based on exact matrix element calculations for $2 \to n$ multi-parton final states, with $n \leq 6$, interfaced with HERWIG+JIMMY to provide the parton shower, hadronisation and underlying event model. In order to properly sample the available phase space, event samples for a fixed value of n are also generated in several p_{T} bins. ALPGEN is known to provide a good description of the multijet final states as measured by ATLAS [141].

6.3 Data sample, trigger and event selection

The data sample used in this analysis corresponds to the full 2011 dataset, having a total integrated luminosity of $\mathcal{L}_d = 4.6$ fb^{-1}. The data have been collected by the EF_J135_A4TC_EFFS trigger, which requires 135 GeV of transverse energy deposit in the EF jet object. The total luminosity recorded by this trigger, taking the prescales into account, is $\mathcal{L}_r = 157.74$ pb^{-1}. It has been shown in Ref. [142] that this trigger is fully efficient when the transverse energy of the leading jet is above 240 GeV, while the cut on $p_{\mathrm{T1}} + p_{\mathrm{T2}} > 500$ GeV, also used in previous

event shape analyses [114], ensures that the p_{T} (and therefore the $E_{\mathrm{T}} \geq p_{\mathrm{T}}$) of the leading jet is always above 250 GeV. Figure 6.1 shows the efficiency curves for the EF_J100, EF_J135 and EF_J180 triggers.

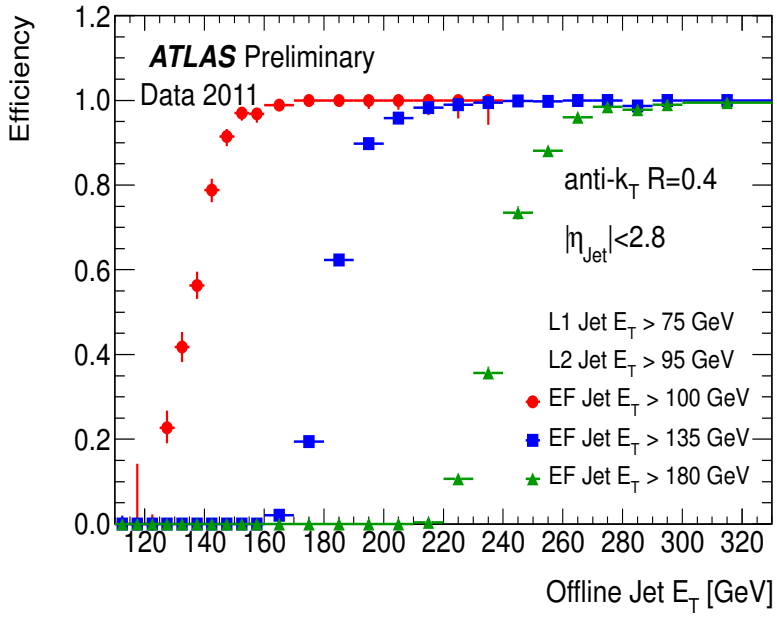

Figure 6.1: The trigger efficiency curves for the EF_J100, EF_J135 and EF_J180 triggers. The cut on $p_{\mathrm{T1}} + p_{\mathrm{T2}} > 500$ GeV ensures that the EF_J135 trigger is fully efficient in the kinematical regime for this analysis. Figure taken from [142]

Events are required to have at least one primary vertex, PV, with five or more tracks with $p_T^{track} \geq 400$ MeV. If there are more than one primary vertex, the one maximizing $\sum p_T^2$ is chosen. To take into account the fact that pileup conditions are different in the Monte Carlo and in the Data samples, a reweighting algorithm has been applied to the MC. This consists on applying a weight to each event depending on the average number of interactions per bunch crossing, $\langle\mu\rangle$, and has been implemented using PileupReweighting-00-02-12. Fig. 6.2 shows the $\langle\mu\rangle$ distributions before and after the reweighting, where the agreement has been forced to be perfect.

Jets are reconstructed with the anti-k_T algorithm [32] with radius parameter $R = 0.4$. The input objects to the jet algorithm, both for data and detector level

simulation, are topological clusters in the calorimeter. These clusters are seeded by calorimeter cells with $|E_{cell}| > 4\sigma$, with σ the RMS of the noise. Neighbouring cells are added and clusters are formed following an iterative procedure.

The baseline calibration for these clusters corrects their energy using local hadronic calibration [143]. Effects due to non-compensation, energy losses in the dead material, shower leakage, as well as inefficiencies in energy clustering and jet reconstruction have to be taken into account. This is done by matching calorimeter jets with MC particle jets in bins of η and p_{T}. This jet calibration also subtracts pileup effects. This is called the jet energy scale, thoroughly discussed in [56, 57]. See [82] for more details and for a discussion of its uncertainties.

The selected events should have at least two jets. The two leading jets are required to fulfill $p_{\mathrm{T1}} + p_{\mathrm{T2}} > 500$ GeV. All jets in the event are required to have a transverse momentum $p_{\mathrm{T}} > 50$ GeV. All selected jets should lie in the central region $|\eta| < 2.5$. These two cuts ensure that the selected jets are well reconstructed and have an optimal energy calibration. The resulting number of selected events is 380000 with an average jet multiplicity $\langle N_{\mathrm{jet}} \rangle = 2.6$. Both BAD and UGLY jets are removed (see Sect. 3.4.4), removing all jets with isBadTight == 1 and isUgly == 1. In addition jets are required to satisfy $|\mathrm{JVF}| > 0.75$ (see Sect. 3.4.5) in order to minimize pile-up effects. The resulting fraction of jets removed by these cuts is 0.16% and 0.28%, respectively.

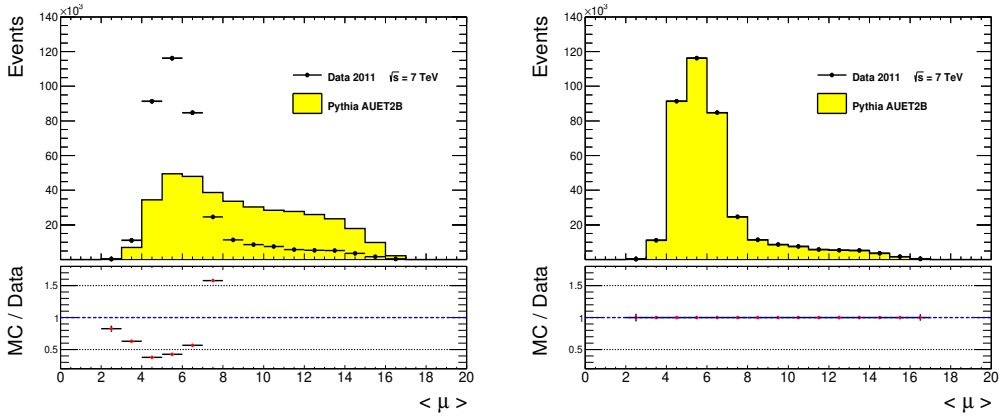

Figure 6.2: Comparison of the $\langle\mu\rangle$ distributions in data and MC before (left) and after the reweighting (right)

6.3.1 Control plots

The resulting distributions in the transverse momentum p_T and the rapidity η of the selected jets are shown in Fig. 6.3 and 6.4, respectively, along with a comparison with Pythia and Alpgen.

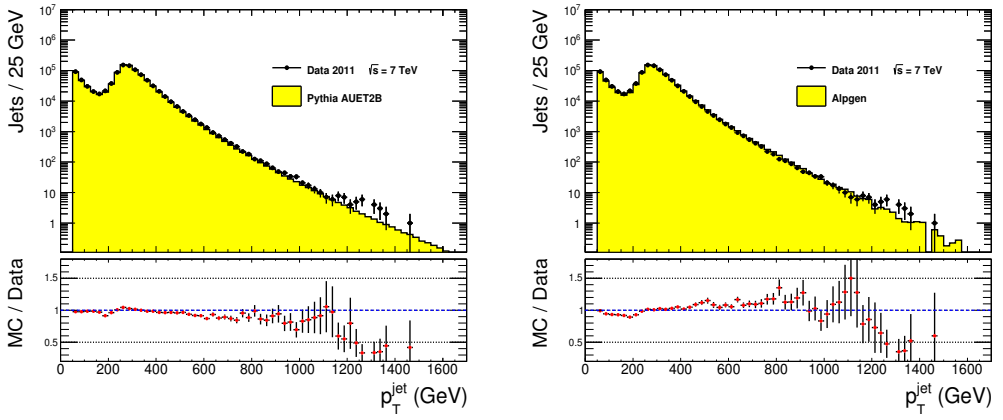

Figure 6.3: Transverse momentum distributions for inclusive jets compared to Pythia (left) and Alpgen (right).

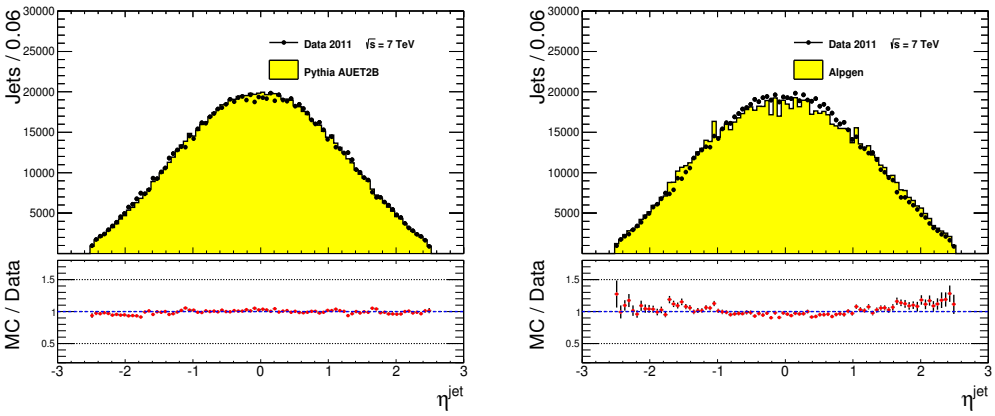

Figure 6.4: Pseudorapidity distributions for inclusive jets compared to PYTHIA (left) and ALPGEN (right).

The leading and subleading jet transverse momenta are shown in Fig. 6.5

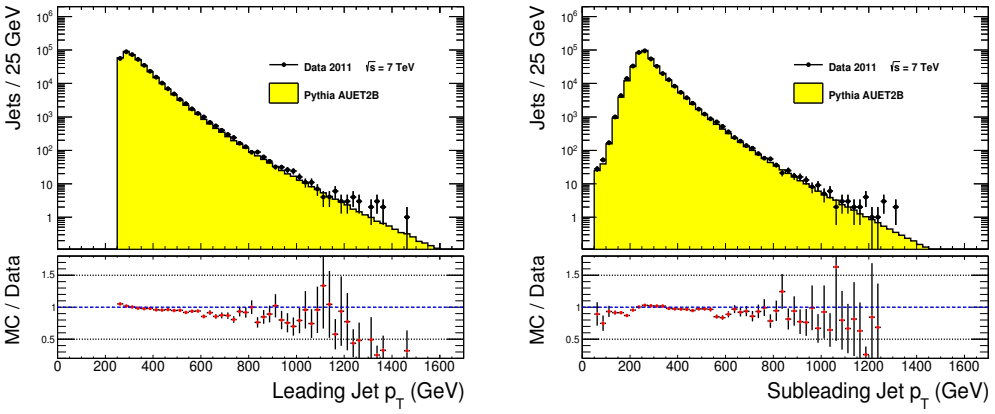

Figure 6.5: Transverse momentum of the leading (left) and subleading (right) jets compared to PYTHIA expectations.

The resulting jet multiplicity as well as the total transverse energy, calculated as the scalar sum of the selected jet transverse energies, are shown in Figs. 6.6 and 6.7 along with a comparison with the PYTHIA expectations. The agreement between data and MC is good for the jet multiplicity and pseudorapidity, while the transverse spectra are harder for the data than for the MC. These differences are well-known and do not represent a problem for the analysis, as the PYTHIA AUET2B is the one describing best the main observables of this analysis, namely

the TEEC and the ATEEC, with less statistical fluctuations. Therefore, this is the Monte Carlo sample used in the unfolding, uncertainty propagation, etc.

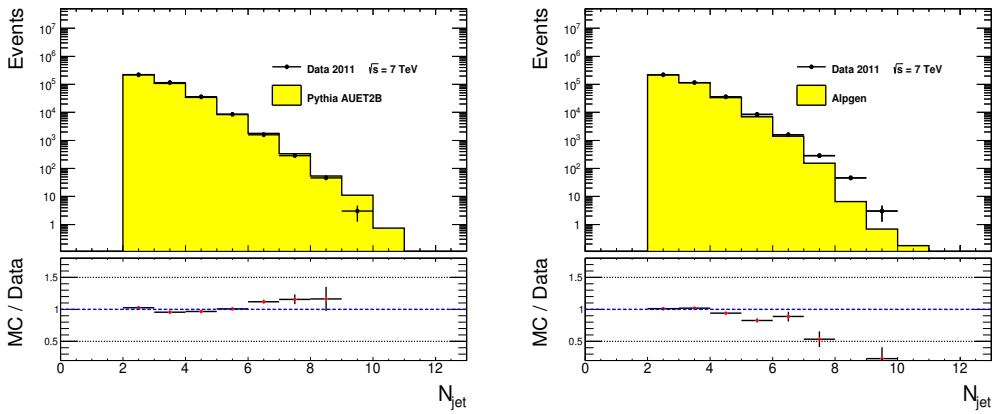

Figure 6.6: Distribution of the jet multiplicity compared to PYTHIA (left) and ALPGEN (right).

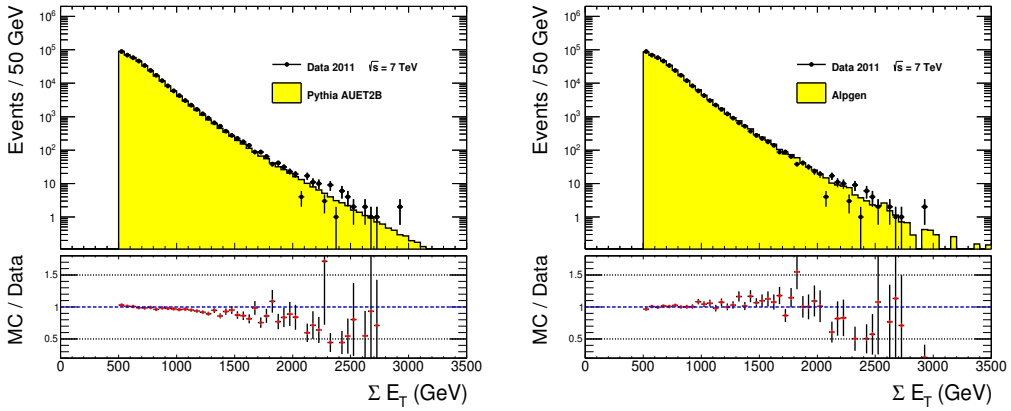

Figure 6.7: Scalar sum of transverse energy of all jets compared to PYTHIA (left) and ALPGEN (right).

Finally, the distribution of the difference in azimuthal angle between any pair of jets in the event is shown in Fig. 6.8

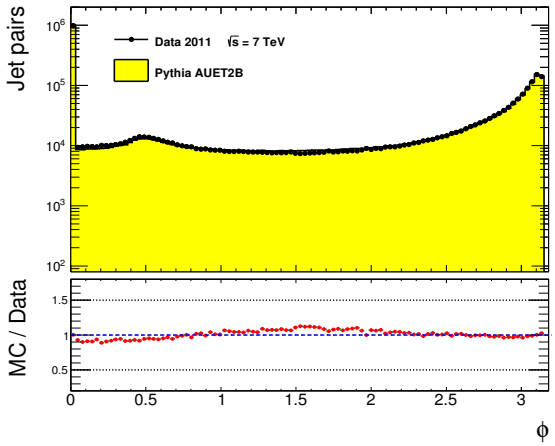

Figure 6.8: Distribution of the differences in azimuthal angle between pairs of jets

6.3.2 Transverse energy versus transverse momentum

Due to the fact that jets have non vanishing masses, the energy and the modulus of the momentum do not necessarily coincide. This is illustrated in Fig. 6.9, where a correlation plot is shown between the transverse energy $E_{\mathrm{T}} = E \sin\theta$ and the jet transverse momentum $p_{\mathrm{T}} = |\vec{p}| \sin\theta$, with θ the polar angle of the jet axis. The distribution of the relative difference $(E_{\mathrm{T}} - p_{\mathrm{T}})/E_{\mathrm{T}}$ is also shown. One can clearly see that the effect of the mass is marginal.

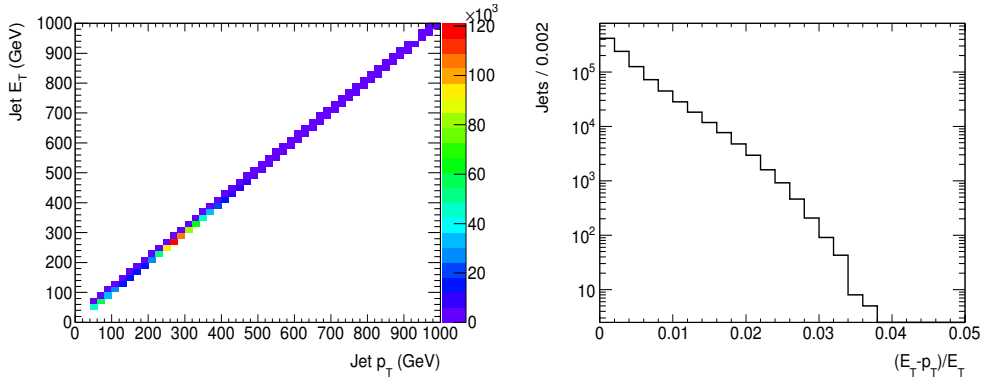

Figure 6.9: Correlation matrix between transverse momentum and transverse energy (left) and the relative difference between both variables (right)

6.4 Results at the detector level

The TEEC and its associated assymetry ATEEC, as defined in equations 1.55 and 1.59 have been calculated using anti-k_t jets with $p_{\mathrm{T}} > 50$ GeV in events with $p_{\mathrm{T1}} + p_{\mathrm{T2}} > 500$ GeV. Effectively, the TEEC distribution is measured by calculating the cosine of the angle in the transverse plane between all possible pairs of jets in the event. Every pair (i, j) represents an entry on the histogram, which is then weighted with the normalised product of the transverse energies defined in equation 6.1

$$w_{ij} = \frac{E_{\mathrm{T}i} E_{\mathrm{T}j}}{\left(\sum_k E_{\mathrm{T}k}\right)^2} \tag{6.1}$$

These weights are defined in such a way that every event is equally weighted, as the sum of these over all possible pairs is always unity (see Eq. 6.2).

$$\sum_{i=1}^{N_{\mathrm{jets}}} \sum_{j=1}^{N_{\mathrm{jets}}} \frac{E_{\mathrm{T}i} E_{\mathrm{T}j}}{\left(\sum_k E_{\mathrm{T}k}\right)^2} = \frac{1}{\left(\sum_k E_{\mathrm{T}k}\right)^2} \left(\sum_{i=1}^{N_{\mathrm{jets}}} E_{\mathrm{T}i}^2 + 2 \sum_{i>j} E_{\mathrm{T}i} E_{\mathrm{T}j} \right) = 1 \tag{6.2}$$

The resulting distribution is then normalised to unit area.

This weighting procedure has many experimental advantages, as reducing the sensitivity to the jet energy scale and resolution. It has also theoretical advantages such as reducing the sensitivity to soft divergencies. This is so because amplitudes for bremsstrahlung processes like $q \to qg$ have a singularity in the mass of the qg pair, which is given by the scalar product of their four-momenta, which for massless partons is $p_i \cdot p_j = 2E_i E_j (1 - \cos\theta_{ij})$. Since the energies are not invariant under Lorentz boosts along the beam direction, the appropriate generalisation to hadron colliders is to use transverse energies.

For the calculation of the statistical error two approaches were followed. In the first approach, if X is the sum of weights per event in a given $\cos\phi$ bin, since the TEEC function is nothing but $\langle X \rangle$ in that bin, the statistical error was calculated as usual i.e.

$$\Delta\langle X \rangle = \frac{\sigma}{\sqrt{N}} \tag{6.3}$$

with N the total number of events in the sample and

$$\sigma = \sqrt{\langle X^2 \rangle - \langle X \rangle^2} \tag{6.4}$$

This underestimates the error slightly because it neglects correlations as the sum of the weights is one and the sum of the angles 2π. To take these effects into account, for the nominal result a bootstrap method was followed, generating per bin and event one thousand replicas obtained from a Poisson distribution.

Figs. 6.10 and 6.11 show the TEEC and ATEEC distributions along with comparisons with reco level Pythia, Herwig and Alpgen expectations. The binning has been chosen as the best compromise between having a fine-binned measurement and avoiding fluctuations. The TEEC exhibits two peaks at $\phi = 0$ (self correlations) and near $\phi = \pi$ (back-to-back region) with a rather flat central plateau. These features are similar to those observed in e^+e^- annihilation [126]. The central plateau is expected to be dominated by hard radiation processes while soft multiple radiation is expected to be important in the small and large angular regions. The ATEEC exhibits a steep fall-off, being poorly determined for $\cos\phi \simeq 0$.

Pythia is the Monte Carlo which describes the data best. The Pythia samples have aproximately ten times the statistics of the Herwig ones and this is the reason why they show smaller fluctuations. This is also the reason why Pythia is chosen as the reference Monte Carlo generator for unfolding purposes.

Alpgen also describes the data well, although it has limited statistics. The description of the TEEC by Pythia and Alpgen is good but for the very forward regions, where discrepancies can reach to about 15%.

The description by Herwig++ is significantly poorer. The ATEEC is fairly described by Pythia and Alpgen with Herwig++ showing discrepancies of the order of 20%.

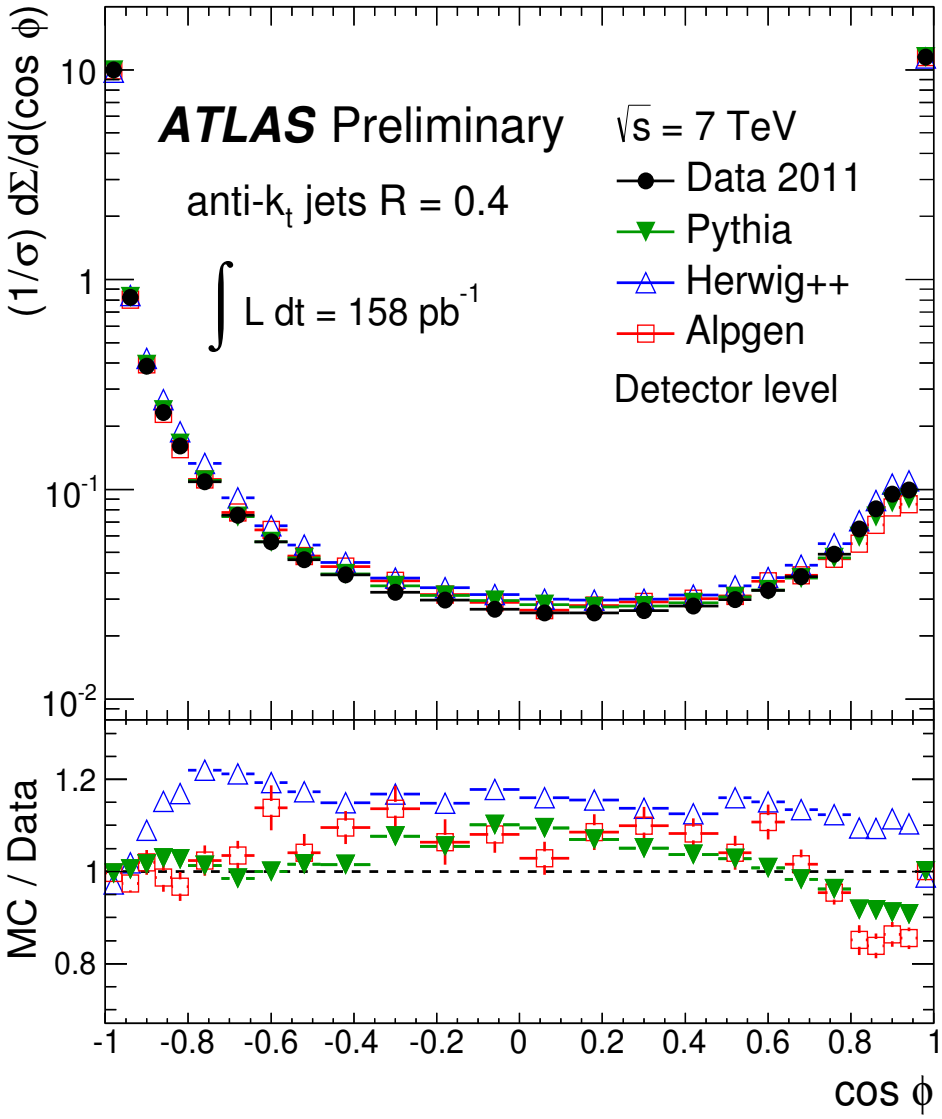

Figure 6.10: The detector level distribution for transverse energy-energy correlation along with comparisons with current Monte Carlo expectations. The uncertainties shown are only statistical.

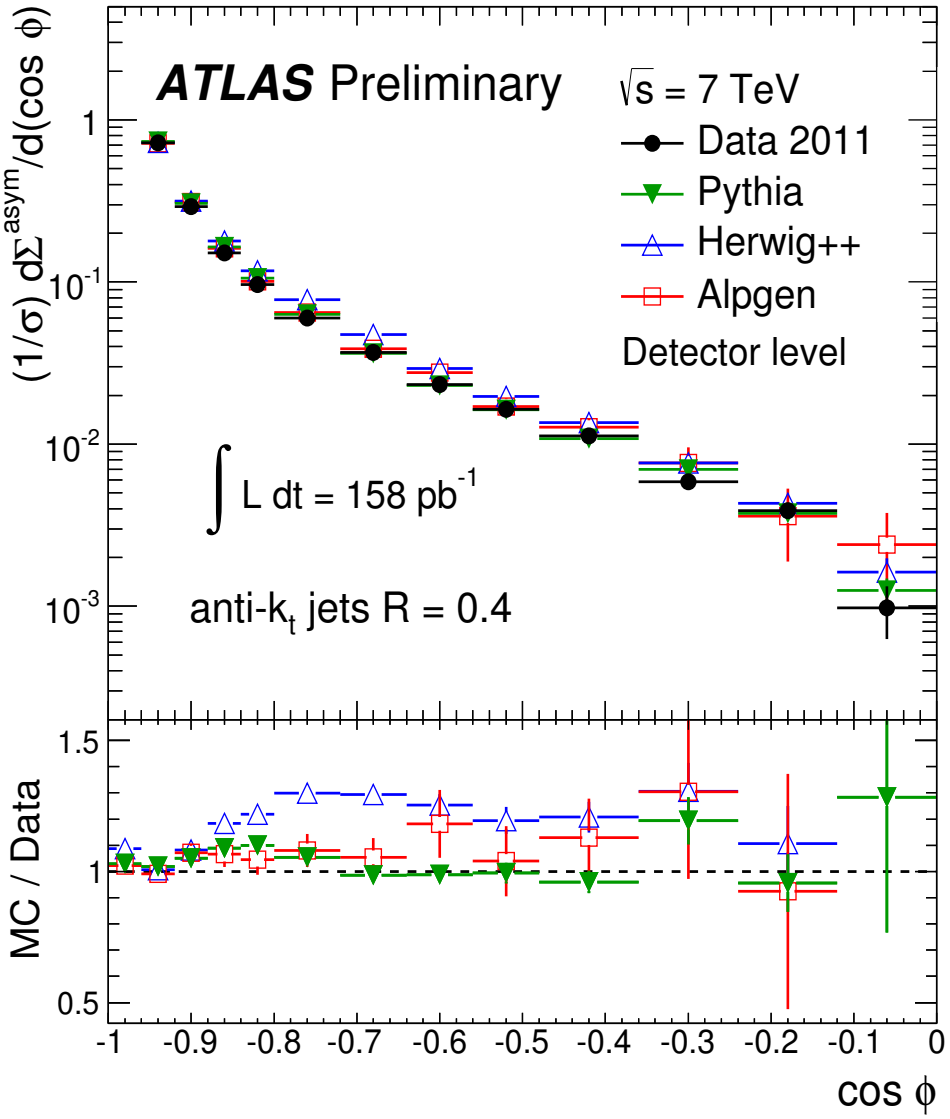

Figure 6.11: The detector level distribution for transverse energy-energy correlation asymmetry along with comparisons with current Monte Carlo expectations. The uncertainties shown are only statistical.

6.5 Unfolding to particle level

In this section, the correction of the data to the hadron level to take detector effects into account is discussed. This allows a direct comparison with existing and future theoretical calculations, as well as with other experiments. To this end, two different approaches are followed: The bin-by-bin and the Bayesian procedures. The hadron-level jets used here contain both muons and non-interacting particles (neutrinos). Therefore, the unfolding procedure fully corrects for energy losses due to partial or total inefficiencies in the detection of these particles by the calorimeter system. The detector level Monte Carlo also accounts for problems related to the LAr hole in some periods of the 2011 data. This is automatically taken into account in the pileup reweighting of the detector level, by weighting the MC events simulating this problem with the fraction of luminosity suffering from this problem in the real data.

6.5.1 Efficiency and Purity

Here, the efficiency and purity in the $\cos(\phi)$ reconstruction is discussed. Efficiency and purity are defined as

$$\varepsilon = \frac{\text{TEEC(Reco, Matched)}}{\text{TEEC(Truth, No matching)}} \tag{6.5}$$

$$\varpi = \frac{\text{TEEC(Reco, Matched)}}{\text{TEEC(Reco, No matching)}} \tag{6.6}$$

Here 'matched' means that there is a one-to-one correspondence between the hadron-level and the reconstructed-level jet pairs. This matching is performed using the distance in the $\eta\varphi$ plane, $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}$. A truth jet pair is considered as 'matched' if both truth jets are matched to reco jets. A truth jet is considered to be matched if the closest reconstructed level jet is within $\Delta R < 0.4$. 'No matching' means that the distribution is calculated using all jet pairs irrespective of whether they fulfill the ΔR condition discussed above or not.

Equation 6.5 illustrates how the efficiency is defined as the ratio between the reconstructed jets and the actual number of generated truth jets per $\cos\phi$ bin.

Equation 6.6 defines the purity as the fraction between the number of reconstructed jets matched to truth jets and the total number of reconstructed jets on a given bin. Figure 6.12 shows that both variables are close to 90% in the full kinematical range for the TEEC measurement.

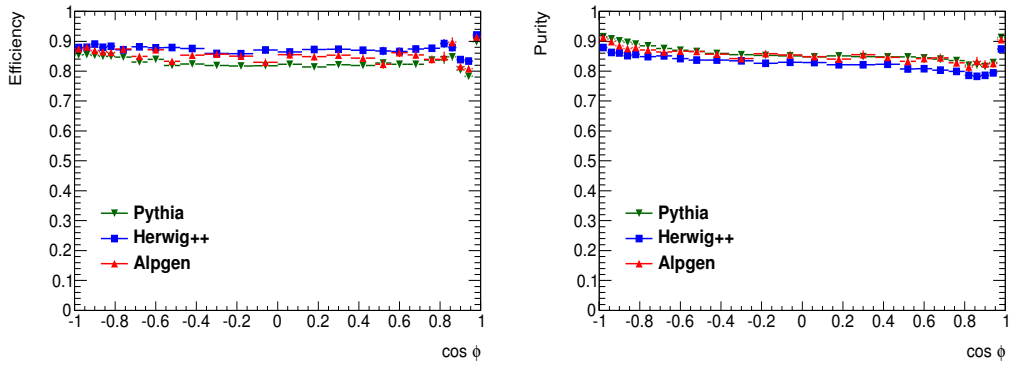

Figure 6.12: Efficiency (left) and Purity (right) of the reconstruction of the differences in azimuth as predicted by the three different Monte Carlos investigated.

6.5.2 Bin-by-bin correlations

Here possible statistical correlations are discussed between bins of the TEEC and ATEEC distributions at the detector level. To extract the correlation coefficients between the values of the bins, the bootstrap method has been used. This method consists in generating N replicas $\{h^{(\alpha)}\}_{\alpha=1}^{N}$ of the measured distribution $h^{(0)}$ (for our case, let $N = 1000$), weighting each entry of the histogram with N different weights distributed following a Poisson distribution centered in 1. The covariance matrix elements between the bin i and the bin j is then defined as

$$V_{ij} = \sum_{\alpha=1}^{N} \left[h_i^{(\alpha)} - h_i^{(0)}\right] \left[h_j^{(\alpha)} - h_j^{(0)}\right] \tag{6.7}$$

The correlation matrix is the covariance matrix normalised to unity, taking into account the statistical uncertainties σ_i for each bin

$$C_{ij} = \sum_{\alpha=1}^{N} \frac{\left[h_i^{(\alpha)} - h_i^{(0)}\right] \left[h_j^{(\alpha)} - h_j^{(0)}\right]}{\sigma_i \sigma_j} \tag{6.8}$$

Ideally, in the case of no bin-by-bin correlations, the correlation matrix is the identity matrix, *i.e* $C_{ij} = \delta_{ij}$. In our case, as it is shown in Fig. 6.13, the crossed bin-by-bin correlations are small except for the endpoints of the TEEC distribution, which is correlated with the rest of the distribution, and the first point of the ATEEC distribution, which also has non-negligible correlations with their neighbouring bins. As it will be seen later, these points do not represent a problem for the extraction of the strong coupling as they will be excluded from the fit for theoretical reasons regarding the dominance of leading-log effects over the fixed order perturbative calculation. Figure 6.13 shows the correlation matrices for the TEEC and ATEEC.

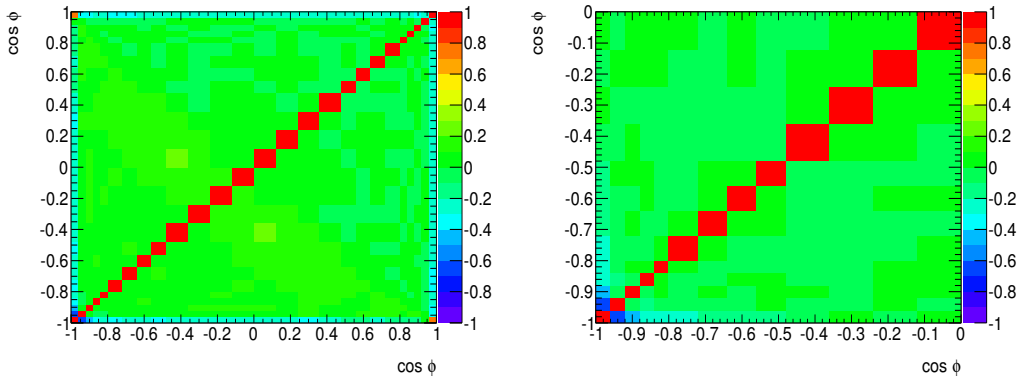

Figure 6.13: Bin-by-bin correlation matrices of the TEEC (left) and ATEEC (right). It is observed that in the central region, away from the endpoints $\cos\phi \sim \pm 1$, the correlations are negligible, while they could have an important effect on the back-to-back and forward regions.

6.5.3 Bin-by-bin unfolding

The bin-by-bin method is based on the calculation of Hadron/Reco correction factors of the form

$$F_{\mathrm{TEEC}}(\cos\phi) = \left[\frac{1}{\sigma'}\frac{d\Sigma'}{d\cos\phi}\right]^{\mathrm{MC}}_{\mathrm{HAD}} \Bigg/ \left[\frac{1}{\sigma'}\frac{d\Sigma'}{d\cos\phi}\right]^{\mathrm{MC}}_{\mathrm{DET}} \tag{6.9}$$

These correction factors differ from unity in just a few percent. Then, the data is corrected to the hadron level by multiplying bin by bin the full distributions

by these factors:

$$\left[\frac{1}{\sigma'}\frac{d\Sigma'}{d\cos\phi}\right]^{\text{DATA}}_{\text{HAD}} = F_{\text{TEEC}}(\cos\phi)\left[\frac{1}{\sigma'}\frac{d\Sigma'}{d\cos\phi}\right]^{\text{DATA}}_{\text{DET}} \tag{6.10}$$

The hadron and reconstructed level TEEC distributions, as well as the correction factors are shown in Fig. 6.14.

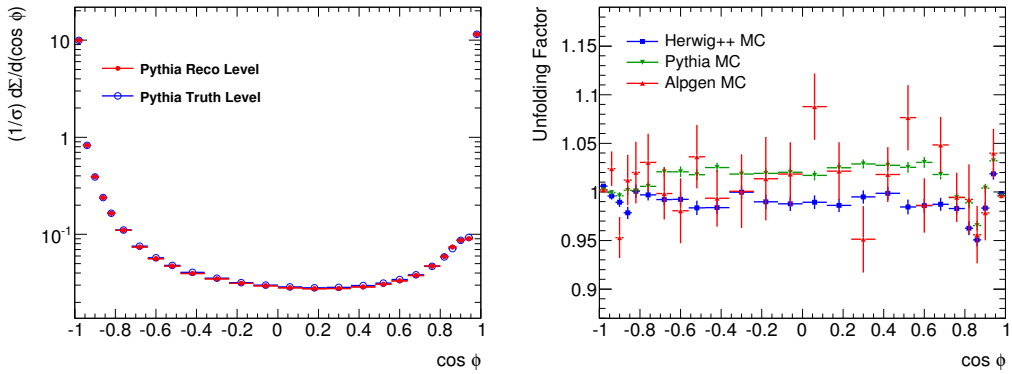

Figure 6.14: PYTHIA expectations for the TEEC at reco and truth level (left) and the correction factors obtained with different Monte Carlo generators (right).

6.5.3.1 Propagation of the statistical uncertainty.

In this section, the propagation of the statistical uncertainty on the unfolded distributions is discussed. It will be assumed that the numerator and the denominator in Eq. 6.9 are fully correlated. This could be justified because they correspond to the same distributions in the same set of events, with the only difference that the denominator is smeared by the detector effects on the GEANT 4 simulation. To take this correlations into account, the bootstrap method has been used to generate $N = 1000$ replicas of the MC sample. This is done by weighting each replica with a random number following a Poisson distribution. The relative difference between the nominal measurement and each of these replicas is shown as a function of $\cos\phi$ in Fig. 6.15 for the truth and reco level distributions, while the corresponding relative differences in the unfolding factors are shown in Fig. 6.16. The statistical uncertainty in the reconstructed level data is derived in the same way.

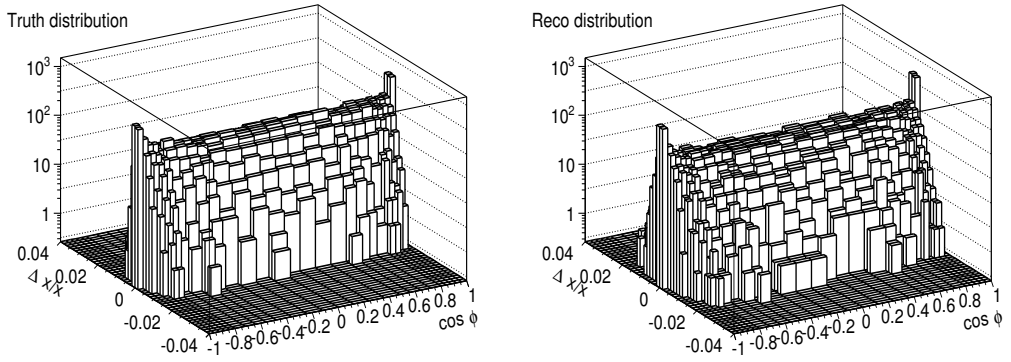

Figure 6.15: Distribution of the relative differences of each bootstrap sample with respect to the nominal for the truth (left) and the reco (right) TEEC distributions.

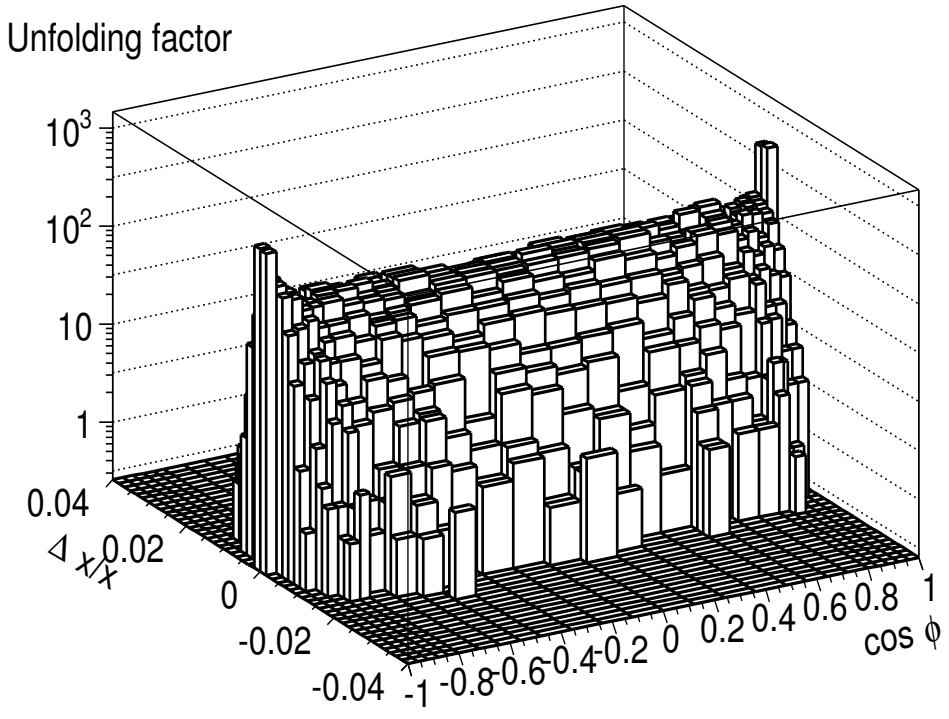

Figure 6.16: Distribution of the relative differences of each bootstrap sample with respect to the nominal for the unfolding factor.

The statistical error in the unfolding factor is then calculated as the RMS of the resulting distribution of relative differences in each $\cos\phi$ bin. Figure 6.17 shows these distributions for high and low statistics bins.

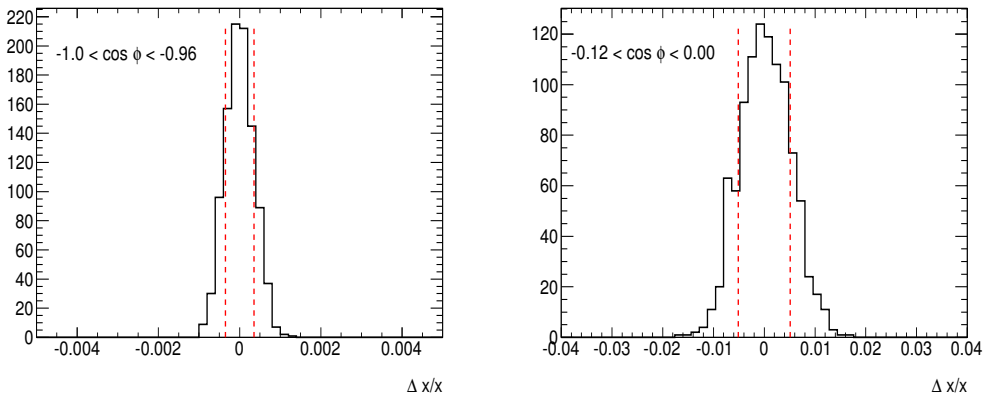

Figure 6.17: Distributions of the relative differences of each bootstrap sample with respect to the nominal unfolding factor for high (left) and low (right) statistics bins of the TEEC distribution. The RMS of the distributions (the statistical errors in each bin) are indicated with red vertical lines.

Once the statistical uncertainty is obtained for the unfolding factor F, the statistical error in the unfolded distribution U is calculated as

$$\frac{\Delta U}{U} = \sqrt{\left(\frac{\Delta F}{F}\right)^2 + \left(\frac{\Delta D}{D}\right)^2} \tag{6.11}$$

where D is the reco-level data distribution. Figure 6.18 shows each component of the uncertainty, together with the total one for each bin of the TEEC distribution. The total statistical uncertainty is at the level of 1%.

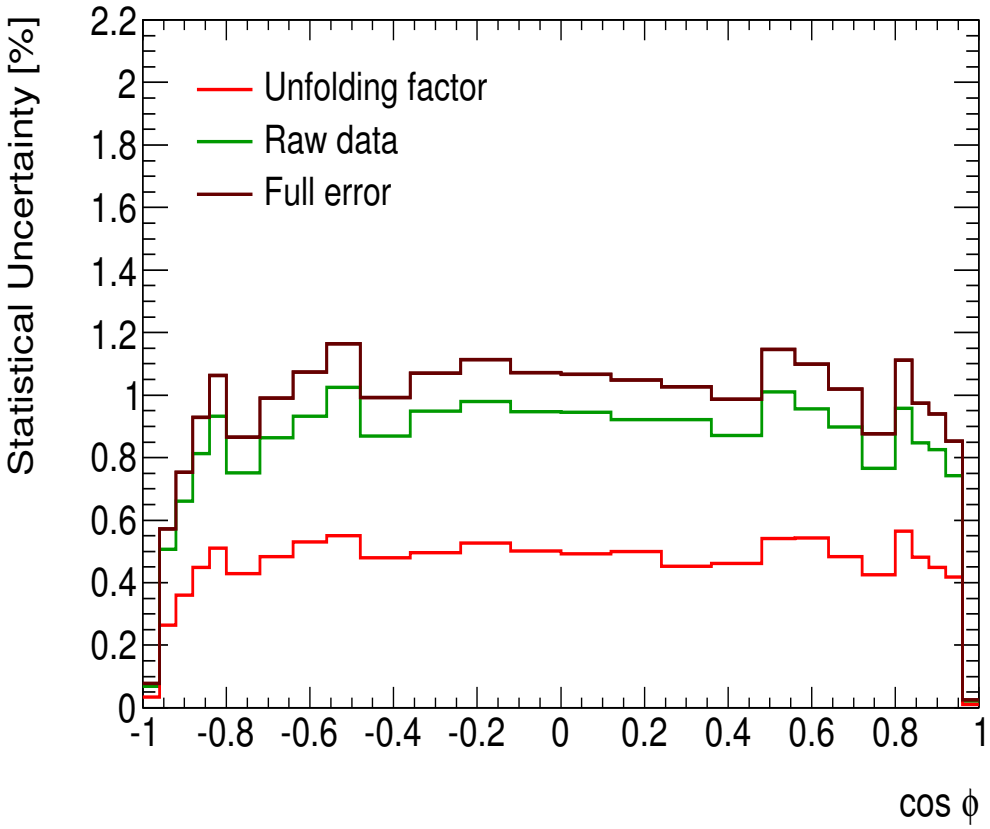

Figure 6.18: The breakdown of the statistical error on the TEEC distribution after the unfolding, showing the MC fraction corresponding to the unfolding factor (in red) and the data fraction corresponding to the raw data (in green).

6.5.4 Bayesian Unfolding

In order to cross-check the bin-by-bin unfolding procedure for any possible effect of migrations in ϕ, a Bayesian unfolding is performed using RooUnfold [92]. Figure 6.19 shows the evolution of the unfolding procedure with the number of iterations, while Figure 6.20 shows the comparison of the Bayesian unfolding output after 5 iterations, where it stabilises, with the bin-by-bin procedure described above. The transfer matrix is shown in Fig. 6.31.

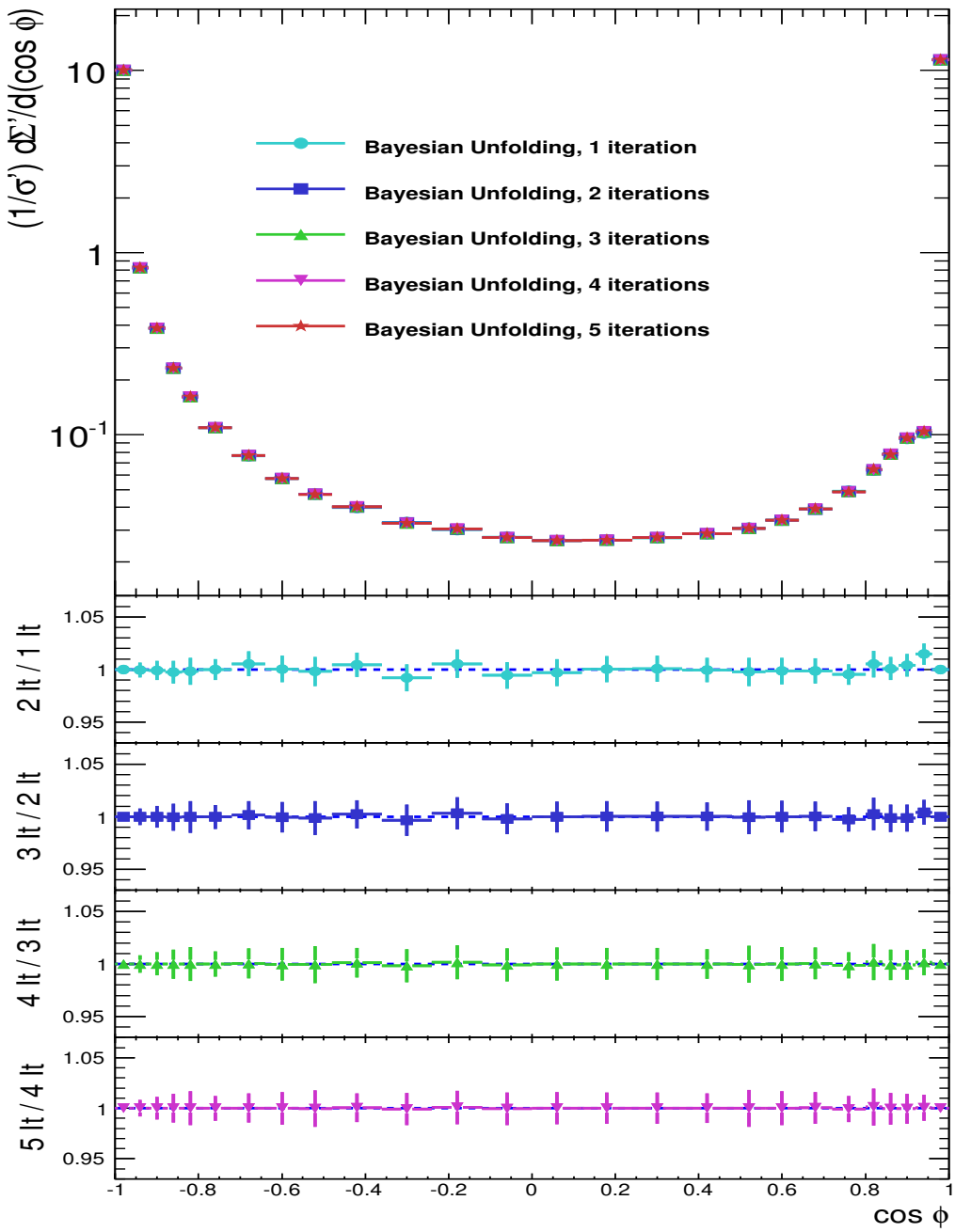

Figure 6.19: Comparison of the Bayesian unfolding output as a function of the number of iterations.

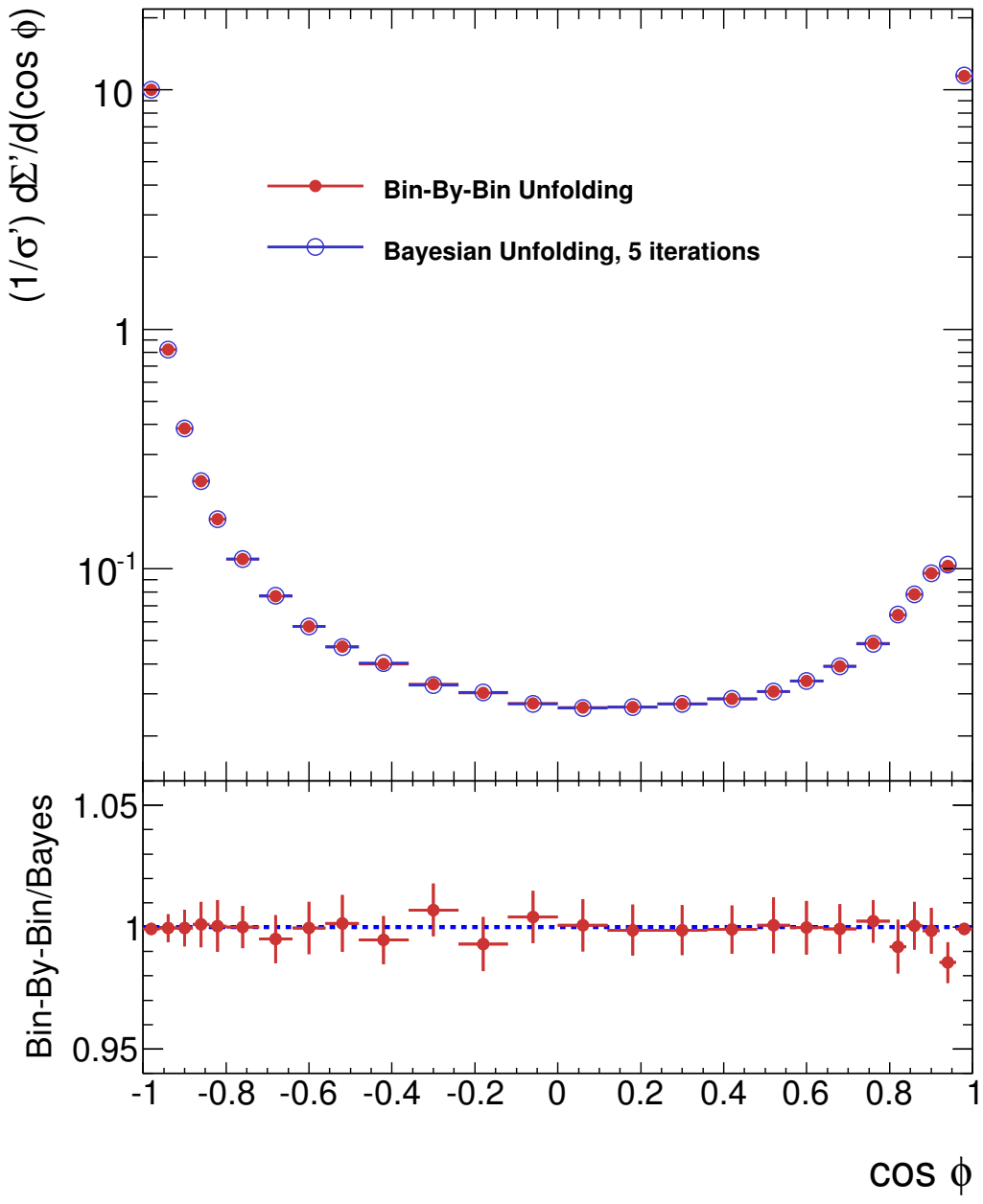

Figure 6.20: Comparison of the Bayesian unfolding output with the bin-by-bin procedure. The uncertainty on the ratio is calculated assuming that the uncertainty in the Bayesian unfolded distribution is negligible.

6.6 Systematic uncertainties

In this section, the experimental sources of uncertainty for the measurement are discussed. They include the Jet Energy Scale and Resolution and the uncertainties associated with the JVF cut, unfolding, the shower modelling and Pileup.

6.6.1 Jet Energy Scale

The Jet Energy Scale (JES) uncertainty [56, 57] is calculated in Monte Carlo by shifting up and down each jet's energy and momentum by 1 sigma for each 63 nuisance parameters depending on the jet transverse momentum and pseudorapidity. The total JES uncertainty for each bin i is calculated as the sum in quadratures of every independent source of uncertainty.

$$\Delta_{\text{JES}} X = \left(\sum_{\alpha \in \mathcal{N}} (\Delta_\alpha X)^2 \right)^{\frac{1}{2}} \tag{6.12}$$

The full set of nuisance parameters for both the TEEC and ATEEC are described in detail in appendix B. The total uncertainty is approximately constant around 4% and is the dominant uncertainty in the analysis.
In order to investigate the effect of the correlations between JES sources in the analysis, two alternative scenarios have been considered, with weaker and stronger correlations. As it will be shown later, the impact of the change of configurations in the value of $\alpha_s(m_Z)$ and its experimental uncertainty is negligible.

6.6.1.1 Flavour composition of the sample

In order to better constrain the flavour-dependent components of the JES uncertainty, the fraction of gluon-induced jets has been estimated and used as an input for the calculation of the relevant JES nuisance parameters. Figure 6.21 shows the gluon fraction of the sample as a function of the jet transverse momentum and pseudorapidity, obtained from both the Pythia and Alpgen samples. The differences between these two predictions has been used as the uncertainty on this parameter.

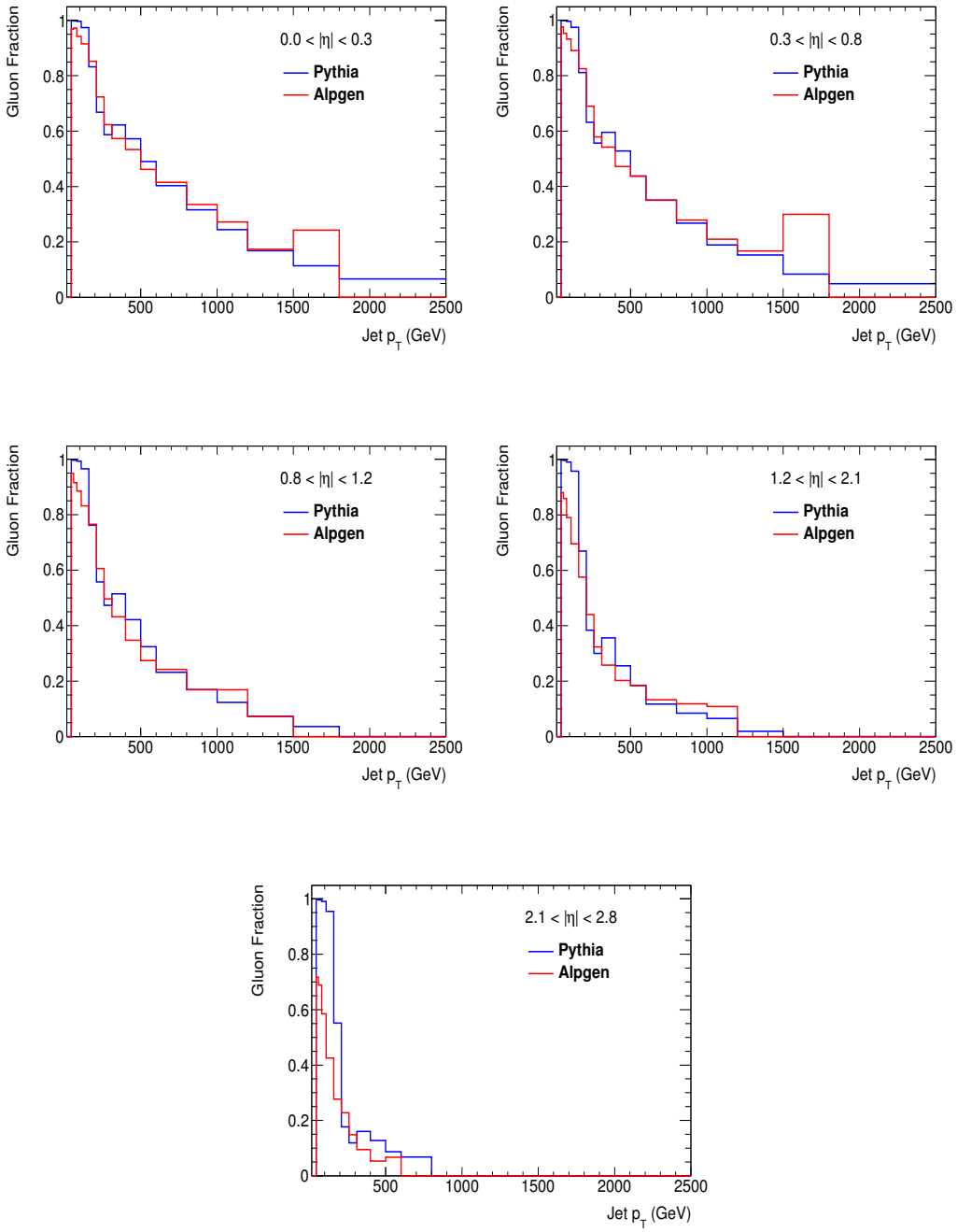

Figure 6.21: The fraction of gluon-induced jets as a function of the jet transverse momentum for the pseudorapidity bins relevant for this analysis

6.6.2 Jet Energy Resolution

The jet energy resolution uncertainty [83] is calculated by smearing each jet's energy and momentum by the p_{T} and η-dependent resolution factor σ_r. The size of this uncertainty on the TEEC function is below 1%.

6.6.3 Uncertainty due to the JVF cut

The uncertainty related to the JVF cut is estimated by analysing the modelling of the JVF cut inefficiencies in Monte Carlo. This is done using the double ratio.

$$\mathcal{R}_{\mathrm{JVF}} = \frac{\left(\dfrac{\mathrm{TEEC(Without\ JVF)}}{\mathrm{TEEC(With\ JVF)}}\right)_{\mathrm{Data}}}{\left(\dfrac{\mathrm{TEEC(Without\ JVF)}}{\mathrm{TEEC(With\ JVF)}}\right)_{\mathrm{MC}}} \tag{6.13}$$

The corresponding uncertainty is of the order of a few per mille in the full TEEC range and therefore it is neglected on the final result. Fig. 6.22 shows the relative value of this uncertainty.

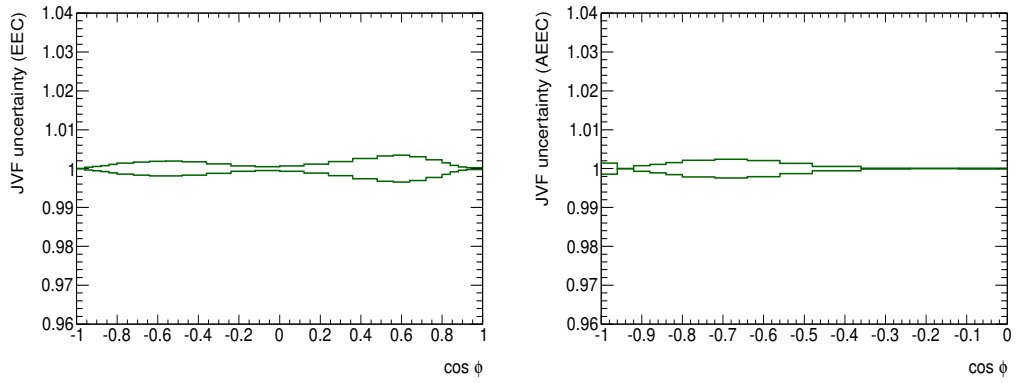

Figure 6.22: The uncertainty due to the JVF cut for the TEEC (left) and the ATEEC (right)

6.6.4 Pileup uncertainty

The pileup uncertainty is estimated by evaluating the description of the pileup dependence of the TEEC and ATEEC functions made by the MC. The estimation

is based on high and low pileup samples selected in both the data and the MC samples. These pileup-varied samples are selected by placing a cut on the average number of interactions per bunch crossing, $\langle\mu\rangle$. The low pileup samples are defined to have $\langle\mu\rangle < 6$, while the high pileup samples are selected by cutting on $\langle\mu\rangle \geq 6$. Figure 6.23 shows the comparison of the nominal TEEC and ATEEC distributions with those obtained in the pileup-varied samples in the Monte Carlo, while Fig. 6.24 shows this comparison for the data.

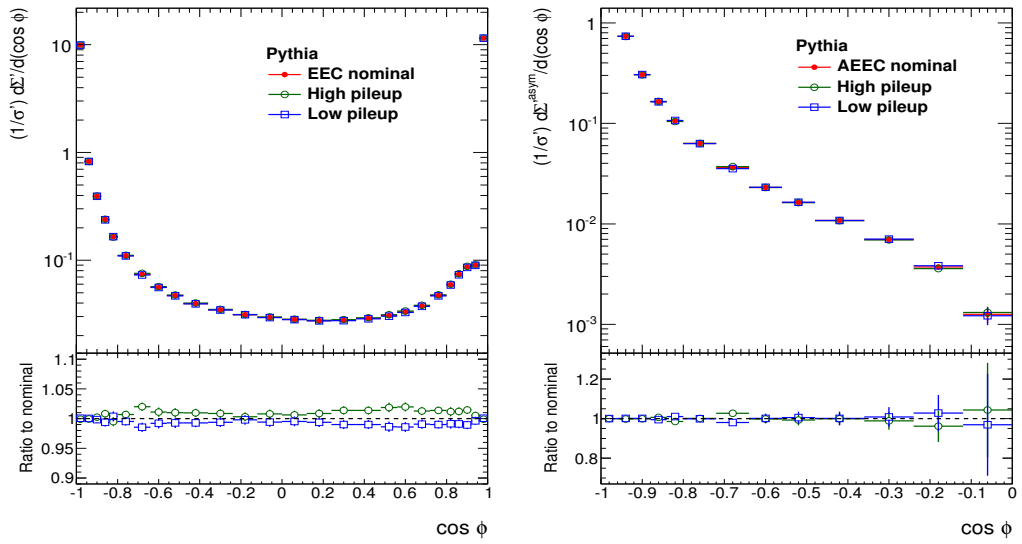

Figure 6.23: The impact of the variation of the average number of interactions per bunch crossing $\langle\mu\rangle$ on the TEEC (left) and ATEEC (right) distributions in the Monte Carlo.

The description of the $\langle\mu\rangle$ dependence of both observables is then estimated using the double ratios

$$\mathcal{R}^{\uparrow}_{\mathrm{PU}} = \frac{\left(\dfrac{\mathrm{TEEC(High\ PU)}}{\mathrm{TEEC(Nominal)}}\right)_{\mathrm{Data}}}{\left(\dfrac{\mathrm{TEEC(High\ PU)}}{\mathrm{TEEC(Nominal)}}\right)_{\mathrm{MC}}}; \quad \mathcal{R}^{\downarrow}_{\mathrm{PU}} = \frac{\left(\dfrac{\mathrm{TEEC(Low\ PU)}}{\mathrm{TEEC(Nominal)}}\right)_{\mathrm{Data}}}{\left(\dfrac{\mathrm{TEEC(Low\ PU)}}{\mathrm{TEEC(Nominal)}}\right)_{\mathrm{MC}}} \tag{6.14}$$

and the uncertainty is defined as the envelope of these two, i.e.

$$\Delta_{\mathrm{PU}} = \max\left\{\left|\mathcal{R}^{\uparrow}_{\mathrm{PU}} - 1\right|, \left|\mathcal{R}^{\downarrow}_{\mathrm{PU}} - 1\right|\right\} \tag{6.15}$$

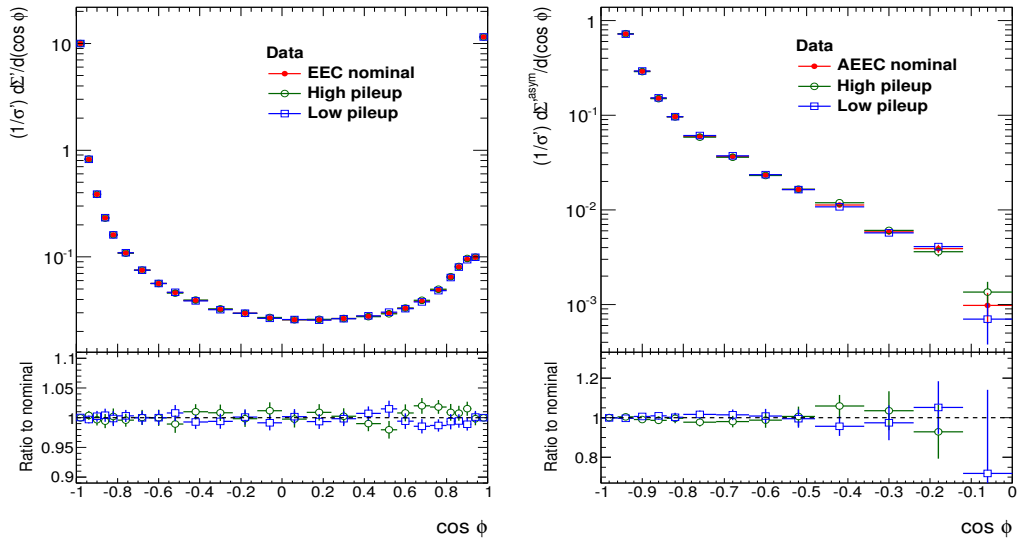

Figure 6.24: The impact of the variation of the average number of interactions per bunch crossing $\langle\mu\rangle$ on the TEEC (left) and ATEEC (right) distributions in the data.

The relative value of this uncertainty is shown in Fig. 6.25

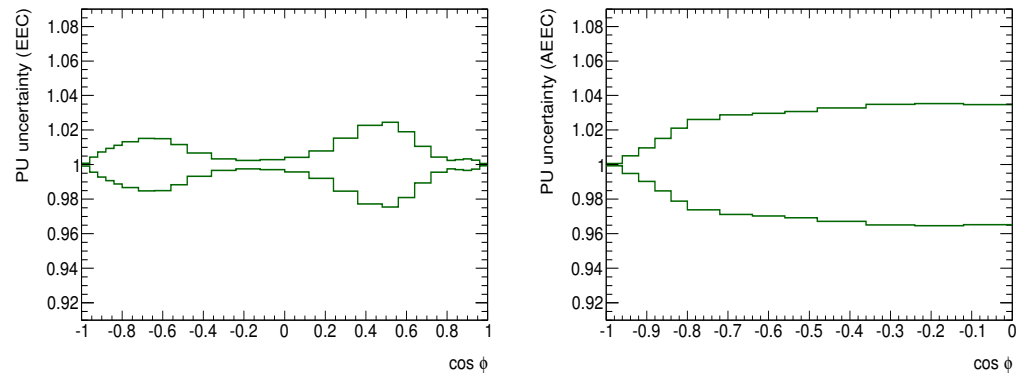

Figure 6.25: The pileup uncertainty for the TEEC (left) and the ATEEC (right)

Due to the fact that this uncertainty is larger than the sum in quadratures of the two pileup terms in JES, these two terms are removed from the jet energy scale uncertainty in favour of this better estimation of the pileup impact.

6.6.5 Jet angular resolution

The jet resolution in the azimuthal angle φ is obtained using the Pythia simulation. Figure 6.26 shows this angular resolution $\varphi^{\text{reco}} - \varphi^{\text{truth}}$ together with the best gaussian fit of this distribution. The obtained resolution is $\sigma_{\varphi} = 10$ mrad. This is much smaller than the bin width used for the TEEC measurement, thus explaining why the transfer matrix shown in Fig. 6.31 is so close to diagonal.

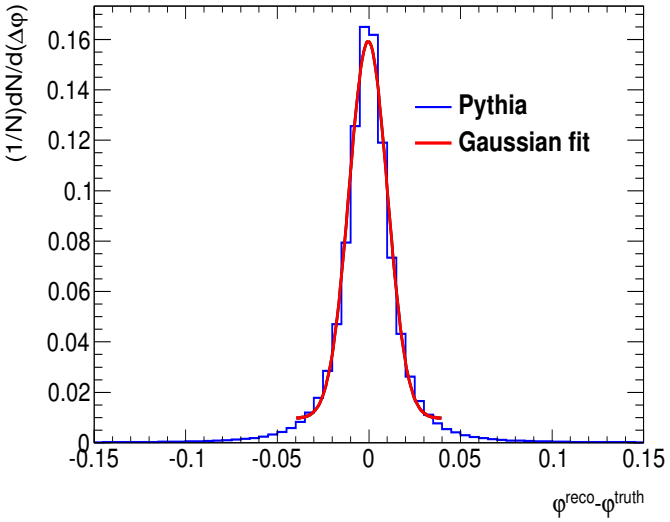

Figure 6.26: The jet angular resolution on the azimuthal angle φ, as obtained from the Pythia AUET2B sample.

The jet angular component φ is then smeared in the Monte Carlo as of 10% of the obtained resolution $\sigma_S = 1$ mrad. This is motivated by track to cluster matching studies done for the azimuthal decorrelation analysis in Ref. [93]. The resulting distributions for the TEEC differ from the nominal in less than 2 per mille and therefore are not taken as a systematic uncertainty. Figure 6.27 shows the comparison of the nominal TEEC and ATEEC distributions with the ones obtained by applying the angular smearing.

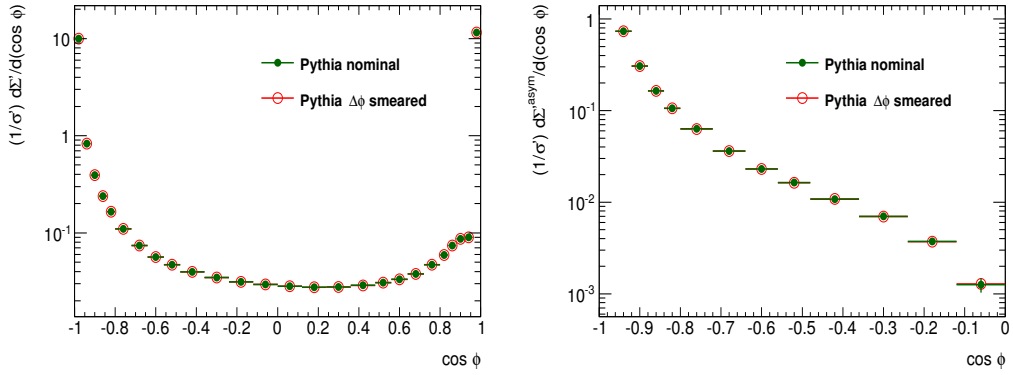

Figure 6.27: The impact of the azimuthal smearing on the TEEC (left) and ATEEC (right) distributions.

Figure 6.28 shows the jet angular resolution uncertainty for both TEEC and ATEEC.

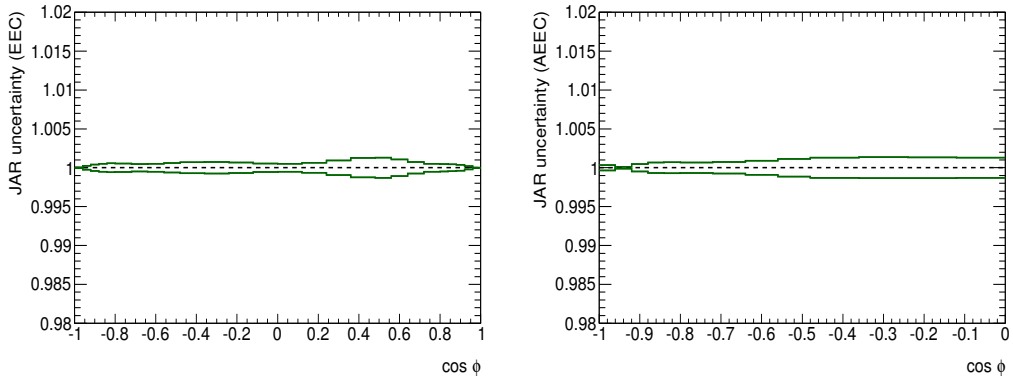

Figure 6.28: Relative uncertainty arising from the jet angular resolution for the TEEC (left) and ATEEC (right) distributions.

6.6.6 Parton shower modelling

To estimate the uncertainty due to the shower modelling, the unfolded data is compared in the cases where Pythia and Herwig++ are used in the unfolding procedure. The shower and hadronisation models in both samples are different, as Pythia sample uses the string hadronisation model, whereas Herwig++ uses the cluster hadronisation model. Moreover, the tuning of the underlying event

in both samples is also different, as multiparton interactions in PYTHIA have been modelled using the AUET2B tune, whereas HERWIG++ uses the UE7000 model. Figure 6.29 shows the differences in both TEEC and ATEEC when using these two different samples for the unfolding. The impact of this uncertainty is smaller than 3.5% in the full $\cos\phi$ range for the TEEC.
Since HERWIG++ gives a poorer description of the TEEC and ATTEC than PYTHIA it has been checked that the systematic differences between data unfolded with these two generators are the same after reweighting the MC samples to fully match the data at reco level following the prescription discussed in subsection 7.8. This is expected since the unfolding factors relate to differences between generator and reco level.

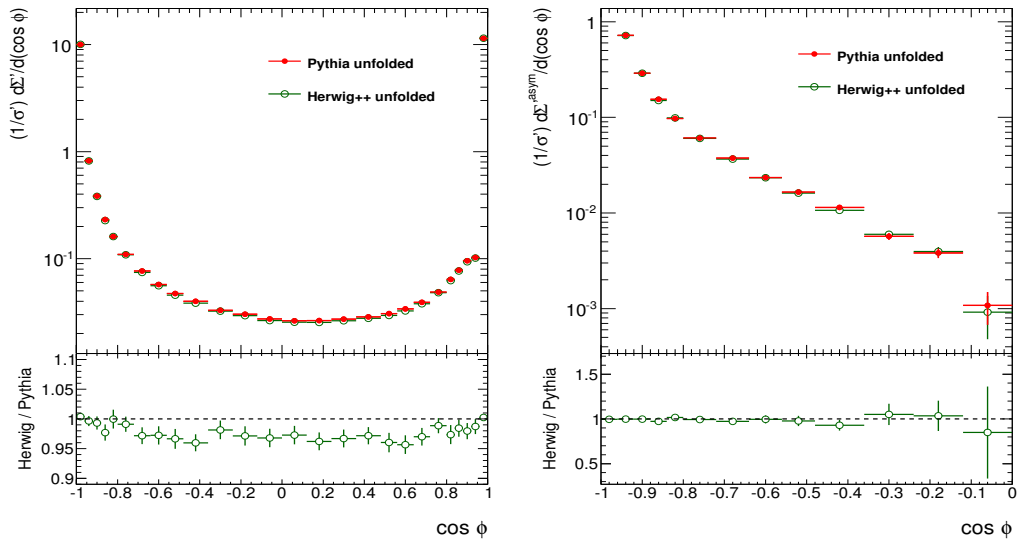

Figure 6.29: The differences in TEEC (left) and ATEEC (right) arising from the parton shower and UE modelling, investigated using PYTHIA and HERWIG++ in the unfolding procedure.

6.6.7 Jet cleaning

In a similar fashion to the uncertainty on the JVF cut, the uncertainty arising from the jet cleaning cuts is estimated using the double ratio between the TEEC distribution obtained by lowering the cleaning cut to MEDIUM while keeping the

removal of UGLY jets in data and MC.

$$\mathcal{R}_{\text{clean}} = \frac{\left(\frac{\text{TEEC(Tight cut)}}{\text{TEEC(Medium cut)}}\right)_{\text{Data}}}{\left(\frac{\text{TEEC(Tight cut)}}{\text{TEEC(Medium cut)}}\right)_{\text{MC}}} \tag{6.16}$$

Figure 6.30 shows that the impact of the jet cleaning mismodelling is negligible in terms of $\mathcal{R}_{\text{clean}}$. This uncertainty is smaller than the statistical uncertainty, and therefore it is not considered as a final systematic uncertainty.

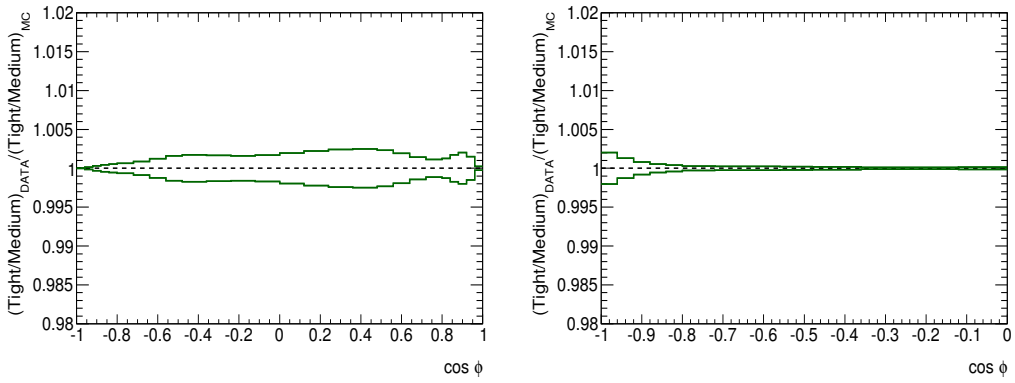

Figure 6.30: The uncertainty due to the jet cleaning cuts for the TEEC (left) and ATEEC (right) distributions.

6.6.8 Uncertainty due to the unfolding procedure

To estimate the uncertainty associated to the unfolding approach, a data-driven closure test is used. This is done by reweighting the transfer matrix between the truth (Y axis) and the Reco (X axis) as obtained using PYTHIA AUET2B following the schema below

- 1. The transfer matrix is reweighted row by row (truth distribution), such that the agreement between Data and MC Reco is enhanced.
- 2. The reweighted matrix is projected on the X axis to obtain the so-modified Reco level distribution.

- 3. The modified Reco level distribution is unfolded using the correction factors calculated using equation 6.9.
- 4. The modified truth level distribution obtained in (1) is compared with the unfolded modified reco distribution obtained in (3). The relative difference between both is taken as the systematic uncertainty.

Figure 6.31 shows the ratio between Data and the Monte Carlo expectations before and after the reweighting, so that an enhancement on the agreement is clearly seen. Figure 6.32 shows the impact of this uncertainty in both TEEC and ATEEC, which is very small due to the fact that the transfer matrix is, to a good level of approximation, diagonal.

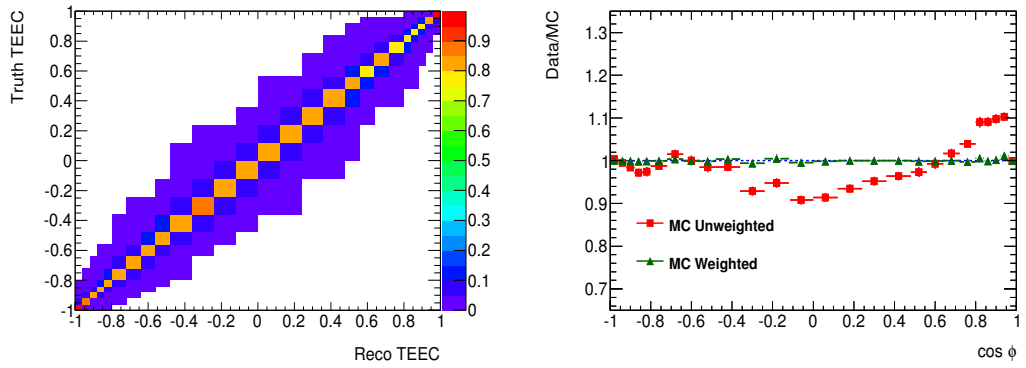

Figure 6.31: Transfer matrix for the TEEC (left), and ratio between Data and MC before and after the reweighting of the truth-level MC (right).

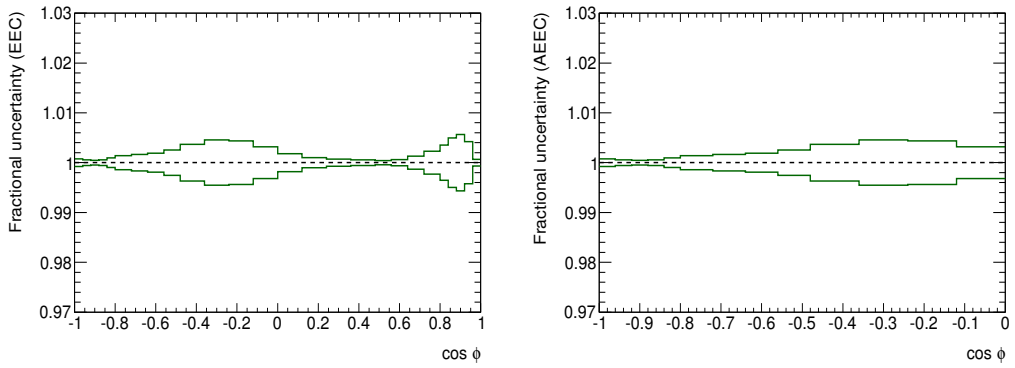

Figure 6.32: The uncertainty due to the unfolding procedure for the TEEC (left) and ATEEC (right) distributions.

6.6.9 Total uncertainty

The uncertainty on the TEEC function is then obtained by adding every independent source in quadratures. The uncertainties are then propagated to the ATEEC function by calculating the forward-backward asymmetry of each of the distributions obtained by variating the TEEC by each independent source of uncertainty. Figures 6.33 and 6.34 show the breakdown of the uncertainties for both the TEEC and ATEEC functions. To avoid statistical fluctuations in the tail of the asymmetry, the five last bins of the ATEEC distributions have been combined for the estimation of the uncertainty. Also, each source of uncertainty has been separately smoothed to avoid statistical fluctuations for both the TEEC and ATEEC.

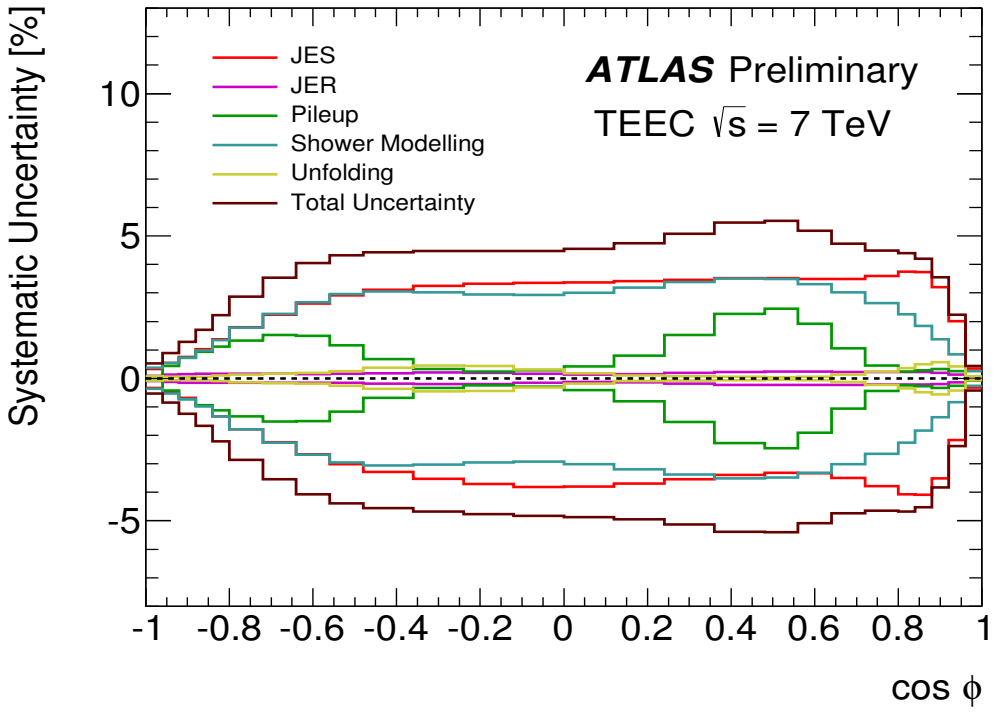

Figure 6.33: Sources of experimental uncertainty for the TEEC function.

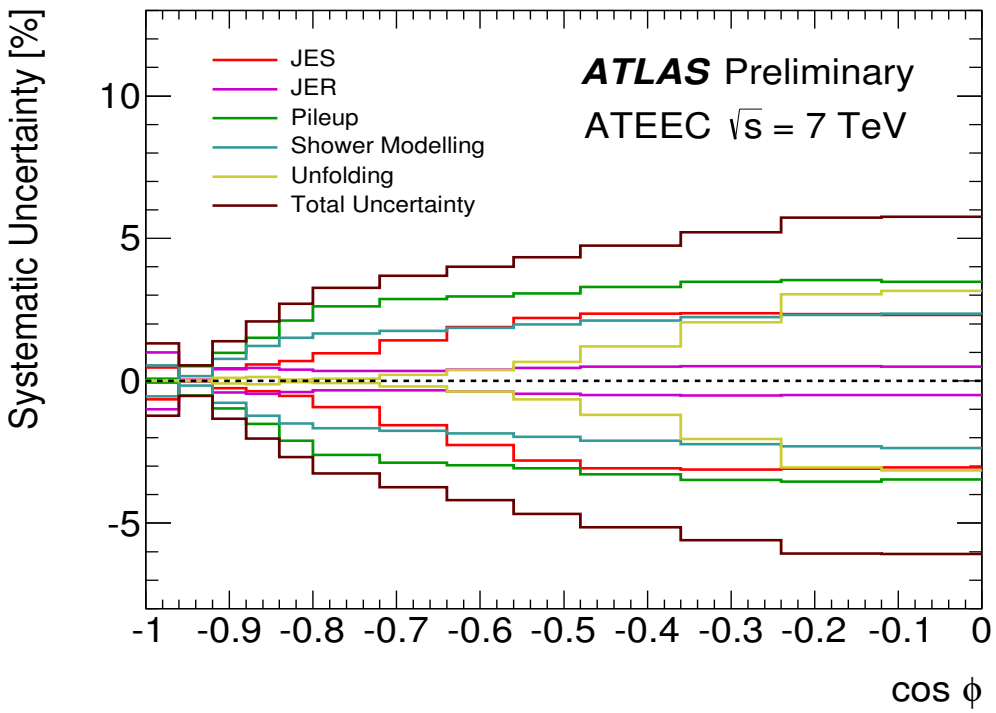

Figure 6.34: Sources of experimental uncertainty for the ATEEC function.

The total uncertainty does not go beyond 5% for the TEEC, being approximately constant. For the ATEEC, it increases as $\cos\phi$ is increased, varying from 1% to about 5%. A detailed summary of the numerical values of the relative uncertainties is presented in Tables 6.3 and 6.4 for TEEC and ATEEC, respectively

$\cos\phi$	JES (↑)	JES (↓)	JER	Shower	PU	Unfolding	Total (↑)	Total (↓)
(-1.00 , -0.96)	0.33 %	0.34 %	0.09 %	0.37 %	0.07 %	0.08 %	0.52 %	0.53 %
(-0.96 , -0.92)	0.54 %	0.49 %	0.13 %	0.53 %	0.43 %	0.06 %	0.89 %	0.86 %
(-0.92 , -0.88)	0.76 %	0.68 %	0.15 %	0.73 %	0.73 %	0.05 %	1.29 %	1.25 %
(-0.88 , -0.84)	1.02 %	0.96 %	0.16 %	0.99 %	0.93 %	0.06 %	1.71 %	1.67 %
(-0.84 , -0.80)	1.37 %	1.34 %	0.16 %	1.34 %	1.12 %	0.10 %	2.23 %	2.21 %
(-0.80 , -0.72)	1.80 %	1.78 %	0.16 %	1.78 %	1.33 %	0.14 %	2.87 %	2.86 %
(-0.72 , -0.64)	2.24 %	2.25 %	0.16 %	2.26 %	1.52 %	0.17 %	3.54 %	3.54 %
(-0.64 , -0.56)	2.62 %	2.66 %	0.16 %	2.68 %	1.50 %	0.19 %	4.04 %	4.07 %
(-0.56 , -0.48)	2.91 %	3.01 %	0.16 %	2.95 %	1.16 %	0.25 %	4.32 %	4.38 %
(-0.48 , -0.36)	3.11 %	3.29 %	0.18 %	3.06 %	0.68 %	0.37 %	4.43 %	4.56 %
(-0.36 , -0.24)	3.24 %	3.53 %	0.20 %	3.02 %	0.34 %	0.45 %	4.47 %	4.68 %
(-0.24 , -0.12)	3.32 %	3.71 %	0.20 %	2.95 %	0.24 %	0.44 %	4.47 %	4.77 %
(-0.12 , 0.00)	3.36 %	3.81 %	0.16 %	2.93 %	0.29 %	0.32 %	4.48 %	4.83 %
(0.00 , 0.12)	3.38 %	3.80 %	0.13 %	3.01 %	0.42 %	0.18 %	4.55 %	4.88 %
(0.12 , 0.24)	3.41 %	3.70 %	0.14 %	3.19 %	0.80 %	0.10 %	4.74 %	4.95 %
(0.24 , 0.36)	3.47 %	3.54 %	0.19 %	3.38 %	1.53 %	0.07 %	5.08 %	5.14 %
(0.36 , 0.48)	3.52 %	3.40 %	0.23 %	3.51 %	2.27 %	0.06 %	5.46 %	5.39 %
(0.48 , 0.56)	3.52 %	3.31 %	0.24 %	3.49 %	2.45 %	0.04 %	5.53 %	5.40 %
(0.56 , 0.64)	3.48 %	3.33 %	0.23 %	3.31 %	1.91 %	0.06 %	5.18 %	5.08 %
(0.64 , 0.72)	3.48 %	3.49 %	0.23 %	3.02 %	1.06 %	0.13 %	4.74 %	4.74 %
(0.72 , 0.80)	3.59 %	3.78 %	0.22 %	2.65 %	0.44 %	0.23 %	4.50 %	4.65 %
(0.80 , 0.84)	3.75 %	4.07 %	0.22 %	2.26 %	0.24 %	0.36 %	4.40 %	4.68 %
(0.84 , 0.88)	3.73 %	4.09 %	0.22 %	1.84 %	0.28 %	0.49 %	4.20 %	4.53 %
(0.88 , 0.92)	3.21 %	3.50 %	0.20 %	1.37 %	0.33 %	0.57 %	3.55 %	3.82 %
(0.92 , 0.96)	2.01 %	2.17 %	0.14 %	0.84 %	0.26 %	0.42 %	2.24 %	2.38 %
(0.96 , 1.00)	0.34 %	0.32 %	0.05 %	0.26 %	0.07 %	0.07 %	0.44 %	0.42 %

Table 6.3: Relative systematic uncertainties for the TEEC function

$\cos\phi$	JES (↑)	JES (↓)	JER	Shower	PU	Unfolding	Total (↑)	Total (↓)
(-1.00 , -0.96)	0.66 %	0.46 %	1.00 %	0.54 %	0.07 %	0.01 %	1.32 %	1.23 %
(-0.96 , -0.92)	0.02 %	0.00 %	0.03 %	0.16 %	0.51 %	0.01 %	0.54 %	0.54 %
(-0.92 , -0.88)	0.43 %	0.27 %	0.41 %	0.77 %	0.98 %	0.10 %	1.38 %	1.34 %
(-0.88 , -0.84)	0.57 %	0.37 %	0.45 %	1.23 %	1.51 %	0.13 %	2.08 %	2.04 %
(-0.84 , -0.80)	0.69 %	0.53 %	0.39 %	1.51 %	2.11 %	0.05 %	2.71 %	2.67 %
(-0.80 , -0.72)	0.97 %	0.92 %	0.34 %	1.66 %	2.61 %	0.07 %	3.26 %	3.25 %
(-0.72 , -0.64)	1.42 %	1.56 %	0.34 %	1.76 %	2.88 %	0.20 %	3.68 %	3.74 %
(-0.64 , -0.56)	1.89 %	2.26 %	0.39 %	1.85 %	2.96 %	0.38 %	4.01 %	4.20 %
(-0.56 , -0.48)	2.21 %	2.81 %	0.45 %	1.97 %	3.07 %	0.66 %	4.34 %	4.68 %
(-0.48 , -0.36)	2.35 %	3.08 %	0.50 %	2.11 %	3.29 %	1.20 %	4.74 %	5.14 %
(-0.36 , -0.24)	2.37 %	3.13 %	0.51 %	2.23 %	3.48 %	2.05 %	5.21 %	5.60 %
(-0.24 , -0.12)	2.34 %	3.09 %	0.51 %	2.31 %	3.54 %	3.04 %	5.73 %	6.07 %
(-0.12 , 0.00)	2.32 %	3.04 %	0.50 %	2.36 %	3.47 %	3.15 %	5.76 %	6.09 %

Table 6.4: Relative systematic uncertainties for the ATEEC function.

6.7 Hadron level results

To obtain the hadron-level results, the bin-by-bin correction factors discussed in Sect. 6.5.3 have been applied. The systematic uncertainties described in the previous section are propagated through the unfolding procedure. The TEEC and ATEEC distributions, once corrected for detector effects, together with their total uncertainties are shown in Figures 6.35 and 6.36. For completeness, the numerical values are also given in Tables 6.5 and 6.6.

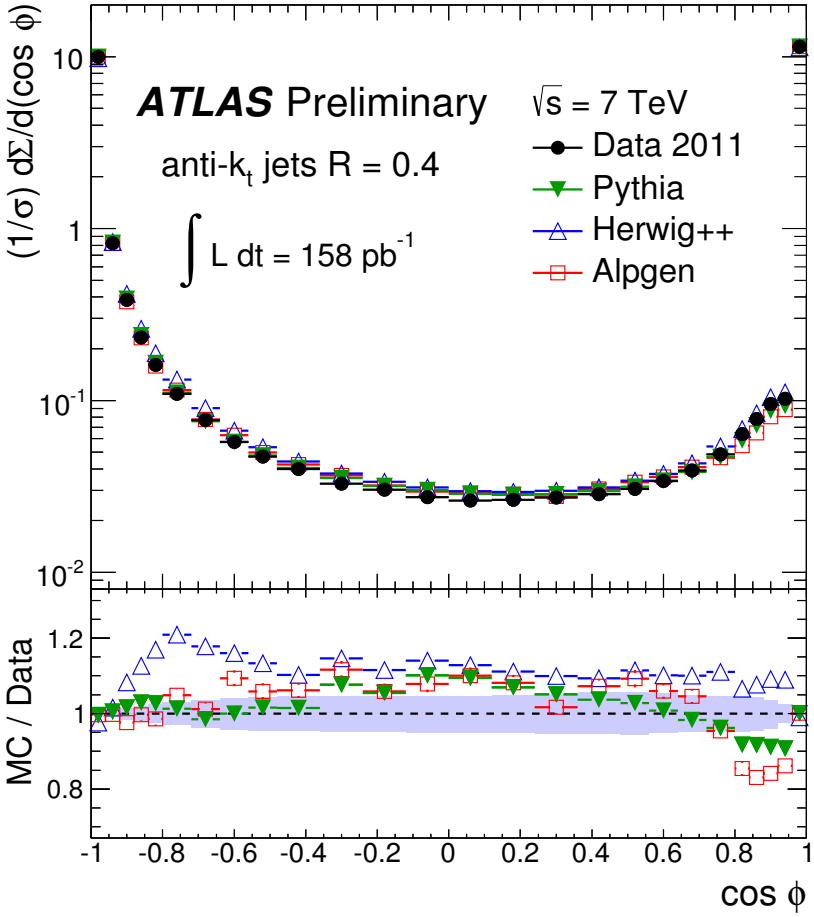

Figure 6.35: The corrected distribution for transverse energy-energy correlation along with comparisons with current Monte Carlo expectations. The statistical uncertainties are shown with error bars, while systematic uncertainties are included into the shaded band. The light blue band on the ratio panel includes both statistical and systematic sources.

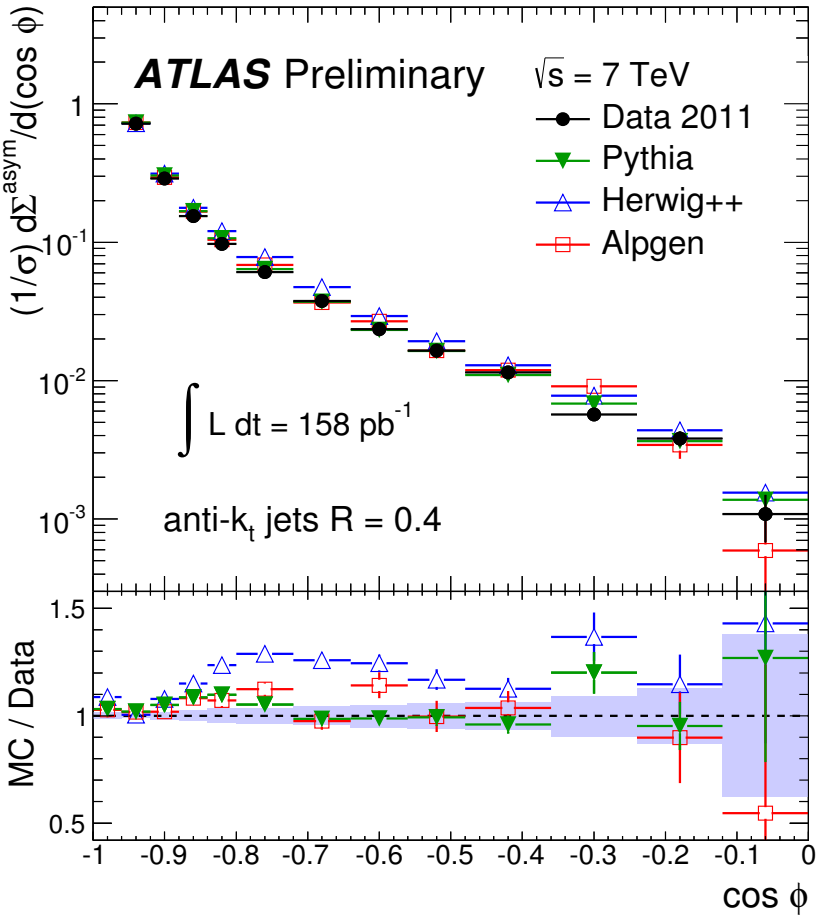

Figure 6.36: The corrected distribution for transverse energy-energy correlation asymmetry along with comparisons with current Monte Carlo expectations. The statistical uncertainties are shown with error bars, while systematic uncertainties are included into the shaded bands. The light blue band on the ratio panel includes both statistical and systematic sources.

$\cos\phi$	TEEC	Stat.	JES	JER	Shower	PU	Unf.
(-1.00 , -0.96)	10.0078	0.0078	$^{+0.0334}_{-0.0342}$	0.0094	0.0374	0.0075	0.0078
(-0.96 , -0.92)	0.8218	0.0047	$^{+0.0044}_{-0.0040}$	0.0011	0.0044	0.0036	0.0005
(-0.92 , -0.88)	0.3848	0.0029	$^{+0.0029}_{-0.0026}$	0.0006	0.0028	0.0028	0.0002
(-0.88 , -0.84)	0.2324	0.0022	$^{+0.0024}_{-0.0022}$	0.0004	0.0023	0.0022	0.0001
(-0.84 , -0.80)	0.1612	0.0017	$^{+0.0022}_{-0.0022}$	0.0003	0.0022	0.0018	0.0002
(-0.80 , -0.72)	0.1095	0.0009	$^{+0.0020}_{-0.0020}$	0.0002	0.0020	0.0015	0.0002
(-0.72 , -0.64)	0.0767	0.0008	$^{+0.0017}_{-0.0017}$	0.0001	0.0017	0.0012	0.0001
(-0.64 , -0.56)	0.0574	0.0006	$^{+0.0015}_{-0.0015}$	0.0001	0.0015	0.0009	0.0001
(-0.56 , -0.48)	0.0472	0.0005	$^{+0.0014}_{-0.0014}$	0.0001	0.0014	0.0005	0.0001
(-0.48 , -0.36)	0.0400	0.0004	$^{+0.0012}_{-0.0013}$	0.0001	0.0012	0.0003	0.0001
(-0.36 , -0.24)	0.0329	0.0004	$^{+0.0011}_{-0.0012}$	0.0001	0.0010	0.0001	0.0001
(-0.24 , -0.12)	0.0302	0.0003	$^{+0.0010}_{-0.0011}$	0.0001	0.0009	0.0001	0.0001
(-0.12 , 0.00)	0.0273	0.0003	$^{+0.0009}_{-0.0010}$	0.0000	0.0008	0.0001	0.0001
(0.00 , 0.12)	0.0262	0.0003	$^{+0.0009}_{-0.0010}$	0.0000	0.0008	0.0001	0.0000
(0.12 , 0.24)	0.0264	0.0003	$^{+0.0009}_{-0.0010}$	0.0000	0.0008	0.0002	0.0000
(0.24 , 0.36)	0.0272	0.0003	$^{+0.0009}_{-0.0010}$	0.0001	0.0009	0.0004	0.0000
(0.36 , 0.48)	0.0286	0.0003	$^{+0.0010}_{-0.0010}$	0.0001	0.0010	0.0006	0.0000
(0.48 , 0.56)	0.0306	0.0004	$^{+0.0011}_{-0.0010}$	0.0001	0.0011	0.0008	0.0000
(0.56 , 0.64)	0.0340	0.0004	$^{+0.0012}_{-0.0011}$	0.0001	0.0011	0.0006	0.0000
(0.64 , 0.72)	0.0391	0.0004	$^{+0.0014}_{-0.0014}$	0.0001	0.0012	0.0004	0.0001
(0.72 , 0.80)	0.0487	0.0004	$^{+0.0017}_{-0.0018}$	0.0001	0.0013	0.0002	0.0001
(0.80 , 0.84)	0.0639	0.0007	$^{+0.0024}_{-0.0026}$	0.0001	0.0014	0.0002	0.0002
(0.84 , 0.88)	0.0780	0.0008	$^{+0.0029}_{-0.0032}$	0.0002	0.0014	0.0002	0.0004
(0.88 , 0.92)	0.0955	0.0009	$^{+0.0031}_{-0.0033}$	0.0002	0.0013	0.0003	0.0005
(0.92 , 0.96)	0.1025	0.0009	$^{+0.0021}_{-0.0022}$	0.0001	0.0009	0.0003	0.0004
(0.96 , 1.00)	11.4484	0.0029	$^{+0.0391}_{-0.0364}$	0.0055	0.0295	0.0084	0.0077

Table 6.5: Corrected values for the energy-energy correlation function (TEEC), together with the absolute statistical and systematic uncertainties.

$\cos\phi$	ATEEC	Stat.	JES	JER	Shower	PU	Unf.
(-1.00 , -0.96)	-1.4406	0.0083	$^{+0.0094}_{-0.0066}$	0.0144	0.0078	0.0010	0.0001
(-0.96 , -0.92)	0.7193	0.0048	$^{+0.0002}_{-0.0000}$	0.0002	0.0012	0.0037	0.0001
(-0.92 , -0.88)	0.2893	0.0030	$^{+0.0012}_{-0.0008}$	0.0012	0.0022	0.0028	0.0003
(-0.88 , -0.84)	0.1544	0.0023	$^{+0.0009}_{-0.0006}$	0.0007	0.0019	0.0023	0.0002
(-0.84 , -0.80)	0.0973	0.0019	$^{+0.0007}_{-0.0005}$	0.0004	0.0015	0.0020	0.0000
(-0.80 , -0.72)	0.0608	0.0010	$^{+0.0006}_{-0.0006}$	0.0002	0.0010	0.0016	0.0000
(-0.72 , -0.64)	0.0376	0.0009	$^{+0.0005}_{-0.0006}$	0.0001	0.0007	0.0011	0.0001
(-0.64 , -0.56)	0.0235	0.0007	$^{+0.0004}_{-0.0005}$	0.0001	0.0004	0.0007	0.0001
(-0.56 , -0.48)	0.0165	0.0007	$^{+0.0004}_{-0.0005}$	0.0001	0.0003	0.0005	0.0001
(-0.48 , -0.36)	0.0115	0.0005	$^{+0.0003}_{-0.0004}$	0.0001	0.0002	0.0004	0.0001
(-0.36 , -0.24)	0.0057	0.0004	$^{+0.0001}_{-0.0002}$	0.0000	0.0001	0.0002	0.0001
(-0.24 , -0.12)	0.0038	0.0004	$^{+0.0001}_{-0.0001}$	0.0000	0.0001	0.0001	0.0001
(-0.12 , 0.00)	0.0011	0.0004	$^{+0.0000}_{-0.0000}$	0.0000	0.0000	0.0000	0.0000

Table 6.6: Corrected values for the asymmetry on the energy-energy correlation function (ATEEC), together with the absolute statistical and systematic uncertainties.

6.8 Theoretical predictions

The pQCD next-to-leading order calculations of the TEEC and ATEEC functions are performed using NLOJet++ [37] interfaced with the MSTW 2008 [13], CT10 [134], NNPDF 2.3 [135] and HERAPDF 1.5 [10] parton distribution functions at NNLO. They have been obtained generating $\mathcal{O}(10^{10})$ events. The results of this programme have been checked by a number of independent NLO jet calculations [36, 133] and have been thoroughly compared with experimental data at ep and hadron colliders. Schematically, this entails the calculation of the $2 \to 3$ partonic subprocesses in the NLO accuracy and of the $2 \to 4$ partonic processes in the LO in $\alpha_s(\mu)$, which contribute to the numerator of the r.h.s of Eq. 3. The azimuthal angle range is restricted to $(-0.92, 0.92)$, which somehow expands the region originally investigated in [36], which was $(-0.8, 0.8)$. Excluding the endpoints removes the self-correlations and frees us from calculating the two-loop virtual corrections to the $2 \to 2$ processes. Thus, the denominator in Eq. 1.56 includes the $2 \to 2$ and $2 \to 3$ processes, which are calculated up to and including $\mathcal{O}(\alpha_s^3)$ corrections.

The calculations have been reported in [36], where furthermore they have been shown to exhibit only a little sensitivity to the PDF choice, MSTW 2008 versus CT10.
For both the numerator and the denominator, the nominal renormalisation and factorisation scales in jet studies, which are inherent in any pQCD calculation, are usually taken to reflect the typical transverse momentum of the process under investigation. For the TEEC and ATEEC measurements, they are taken to be

$$\mu_R = \mu_F \equiv \langle p_{\mathrm{T}12} \rangle = \frac{p_{\mathrm{T},1} + p_{\mathrm{T},2}}{2} \tag{6.17}$$

where $p_{\mathrm{T},1}$ and $p_{\mathrm{T},2}$ are the transverse momenta of the leading and the subleading jet, respectively. This is also the choice in [136].

The theoretical uncertainties on the renormalisation and factorisation scales, as well as on the PDFs are also calculated using NLOJet++. Example distribu-

tions of the NLO theory predictions for the TEEC and ATEEC were shown in Fig. 6.55.

6.8.1 Theoretical uncertainties

The theoretical uncertainties for this analysis are divided into three classes: those corresponding to the renormalisation and factorisation scale variations, the ones corresponding to the PDF eigenvectors, and the ones on the non-perturbative corrections. It has been shown in [36] that the TEEC being defined as the ratio between the 3-jet and the 2-jet inclusive cross sections is not very sensitive to the PDF choice.

6.8.1.1 Scale uncertainties

For calculating the theoretical uncertainty on the renormalisation and factorisation scales, the TEEC and ATEEC distributions have been computed varying independently each of them by a factor of two, with the additional requirement that $1/2 \leq \mu_R/\mu_F \leq 2$. The combinations for the scale variations are therefore

$$\left(\frac{\mu_R}{\mu_0}, \frac{\mu_F}{\mu_0}\right) \in \left\{\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, 1\right), \left(1, \frac{1}{2}\right), (1,2), (2,1), (2,2)\right\} \tag{6.18}$$

where μ_0 is the nominal scale choice. From those, the ones giving the highest uncertainty are those where the scales are shifted in the same direction simultaneously, this is

$$\left\{\left(\frac{1}{2}, \frac{1}{2}\right), (2,2)\right\} \tag{6.19}$$

The impact of the variations in Eq. 6.18 on the TEEC and ATEEC functions are summarised in Figs. 6.37 to 6.42. The two combinations in Eq. 6.19 define the envelope of the scale uncertainty for both the TEEC and ATEEC functions, and therefore, the combination of both will be taken as the total scale uncertainty on these distributions. The total scale uncertainty has been smoothed to avoid statistical fluctuations. Several other choices for the factorization and renormalization scales have been tried, namely the p_{T} of the leading jet $p_{\mathrm{T},1}$, the sum $p_{\mathrm{T},1} + p_{\mathrm{T},2}$ or the total E_{T} of the event. They all give larger uncertainties.

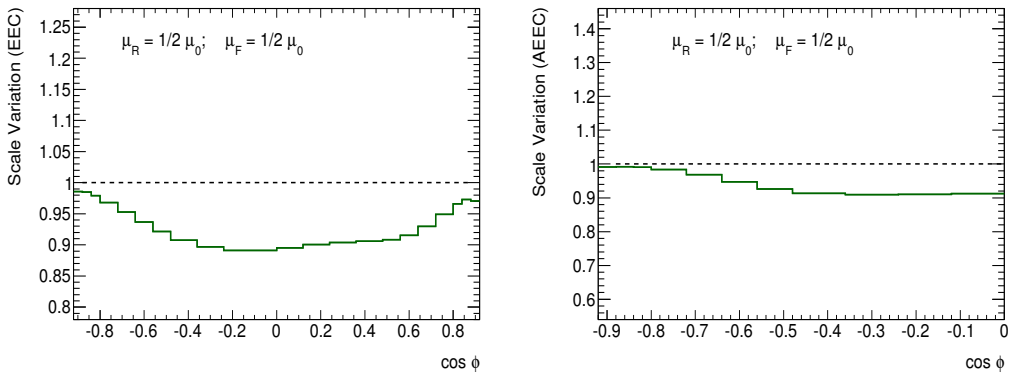

Figure 6.37: The impact on the TEEC (left) and ATEEC (right) of the $(1/2, 1/2)$ scale variation.

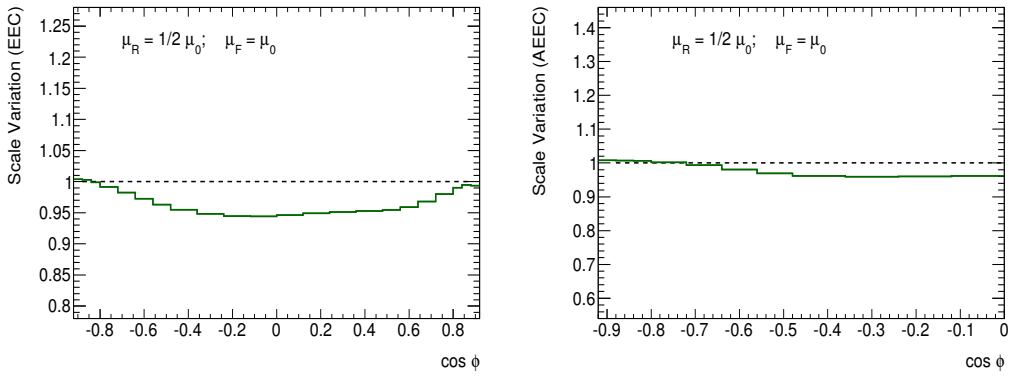

Figure 6.38: The impact on the TEEC (left) and ATEEC (right) of the $(1/2, 1)$ scale variation.

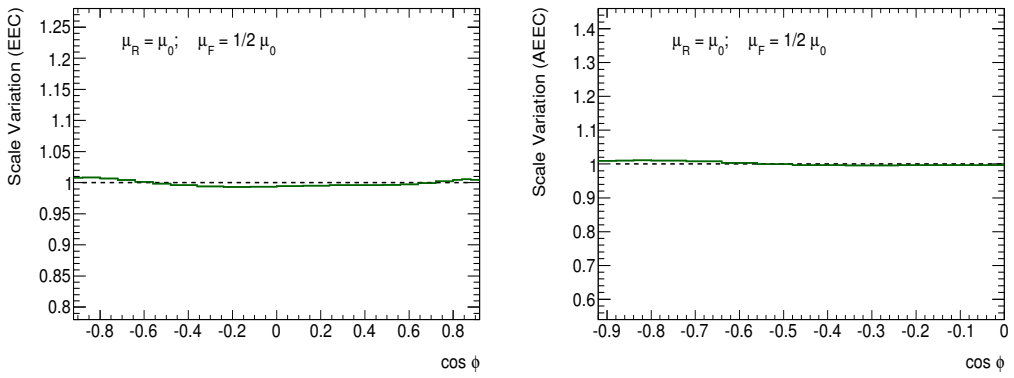

Figure 6.39: The impact on the TEEC (left) and ATEEC (right) of the $(1, 1/2)$ scale variation.

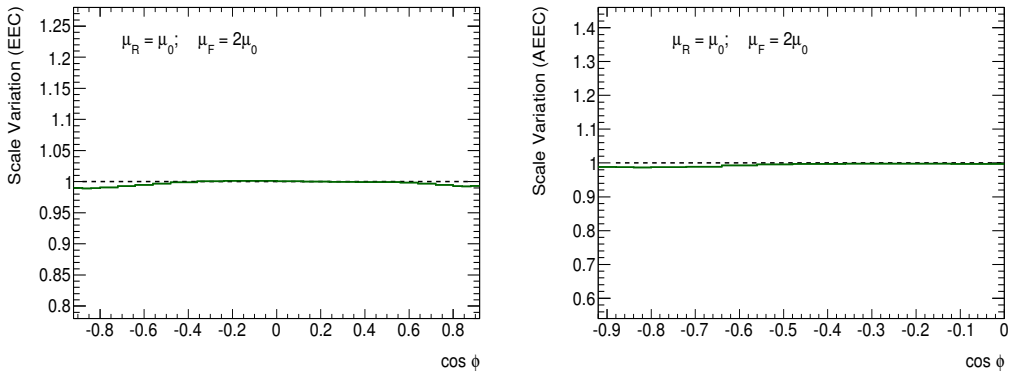

Figure 6.40: The impact on the TEEC (left) and ATEEC (right) of the $(1, 2)$ scale variation.

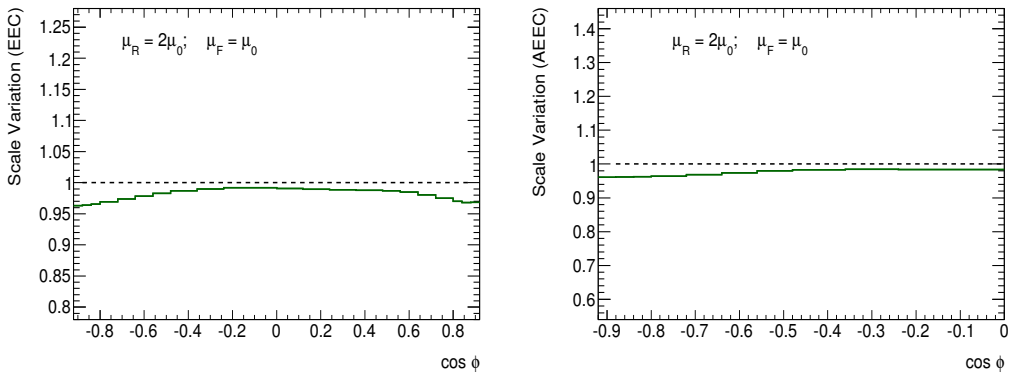

Figure 6.41: The impact on the TEEC (left) and ATEEC (right) of the $(2, 1)$ scale variation.

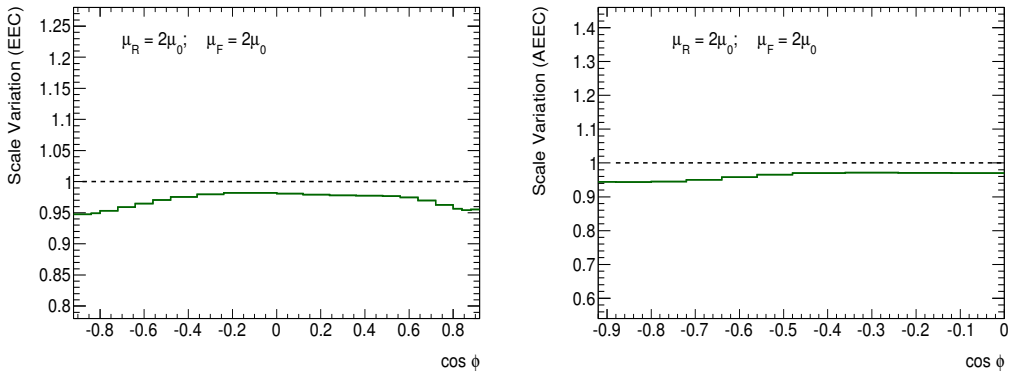

Figure 6.42: The impact on the TEEC (left) and ATEEC (right) of the $(2, 2)$ scale variation.

6.8.1.2 PDF eigenvectors

The CT10 parton density functions provide 50 variations for the 25 fitted parameters at the 90% confidence level. Each of the 25 varied parameters are shifted up and down following the CT10 recommendations in [134], and are combined for each bin of the TEEC and ATEEC functions by using the master formula

$$\Delta X = \frac{1}{2}\sqrt{\sum_{k=1}^{n}\left[X(S_k^{\uparrow}) - X(S_k^{\downarrow})\right]^2} \tag{6.20}$$

6.8.1.3 Non-perturbative corrections

In order to compare the theoretical parton-level predictions with the hadron-level data, the non-perturbative corrections are needed for taking into account effects due to the hadronisation and the underlying event modelling. To this end, Pythia6 AUET2B, Pythia6 AMBT2B and Herwig++ UE7000 samples, as well as Pythia8 AU2 and Pythia8 4C, each with different hadronisation and underlying event models, have been generated with hadronisation and UE turned OFF, and correction factors have been computed for these effects, combined into a total correction factor $\mathcal{F}$.
In Pythia, the switches controlling the hadronisation and the underlying event are given by MSTP(111) = 0 and MSTP(81) = 20, respectively. In Herwig++, the hadronisation switch is given by

```
set /Herwig/EventHandlers/LHCHandler:HadronisationHandler NULL
```

While the switch corresponding to the underlying event is given by

```
set /Herwig/Shower/ShowerHandler:MPIHandler NULL
```

The correction factors are then calculated as the ratio between the TEEC with hadronisation and UE and the TEEC with these effects turned off.

$$\mathcal{F} = \frac{\text{TEEC}(\text{Had} = \text{ON}, \text{UE} = \text{ON})}{\text{TEEC}(\text{Had} = \text{OFF}, \text{UE} = \text{OFF})} \tag{6.21}$$

Figures 6.43 shows the correction factors to be applied to the theoretical predictions.

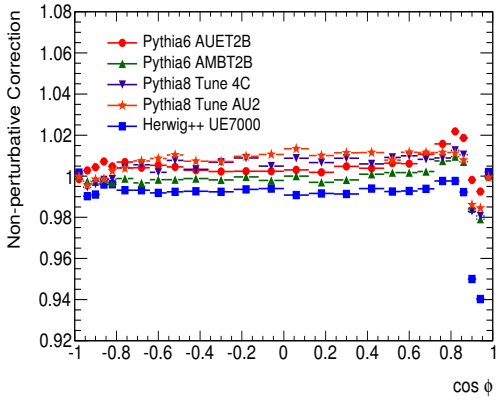

Figure 6.43: The non-perturbative correction factors

The central value used for the correction is chosen to be Pythia AUET2B. The theoretical uncertainty due to the non-perturbative correction is then calculated as the maximum difference between Pythia AUET2B and any of the other choice. This defines the envelope of the uncertainty, as shown in Fig. 6.43.

6.8.1.4 α_s uncertainty

The theoretical uncertainty due to the determination of α_s is estimated, following the PDF4LHC recommendations, by varying the value of the strong coupling in 0.0012. The result is a plain 1% uncertainty on the TEEC. Fig. 6.44 shows the results of this estimation

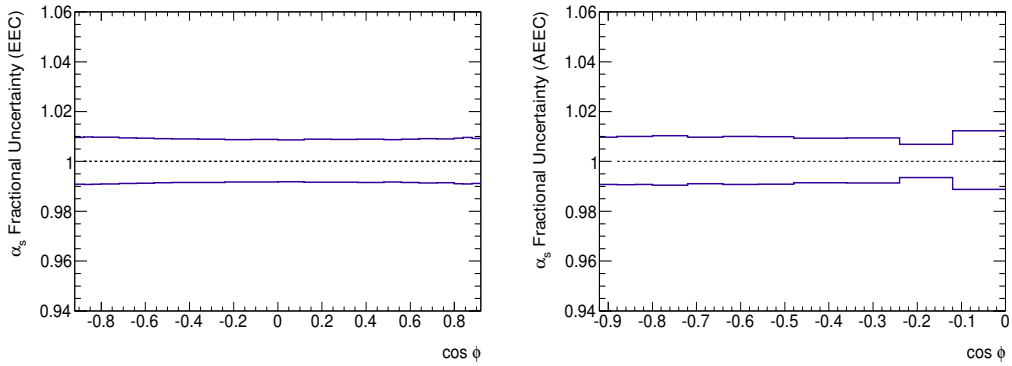

Figure 6.44: The uncertainty due to α_s for the TEEC (left) and ATEEC (right) distributions

6.8.2 Comparison with data

In this section, the nominal theoretical predictions are compared with the unfolded data. Figures 6.45 to 6.48 show these comparisons for both the TEEC and ATEEC. The agreement is good for the TEEC, being always within the uncertainties, and very good for the ATEEC.

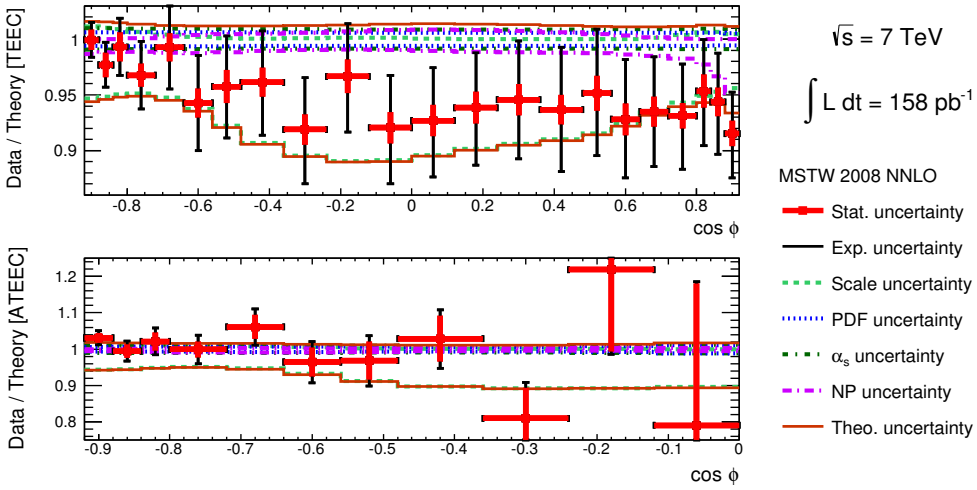

Figure 6.45: Ratios of the measured TEEC and ATEEC functions with respect to the theoretical predictions using NLO pQCD convoluted with the MSTW PDF.

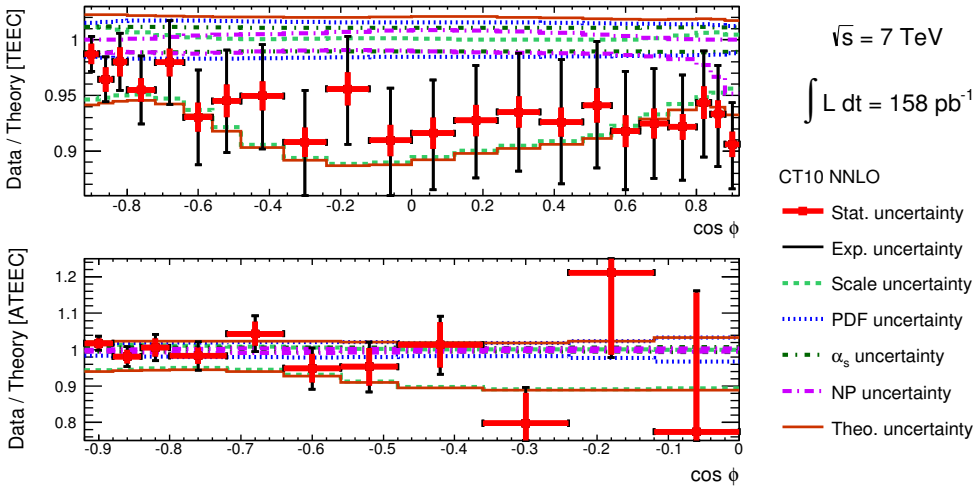

Figure 6.46: Ratios of the measured TEEC and ATEEC functions with respect to the theoretical predictions using NLO pQCD convoluted with the CT10 PDF.

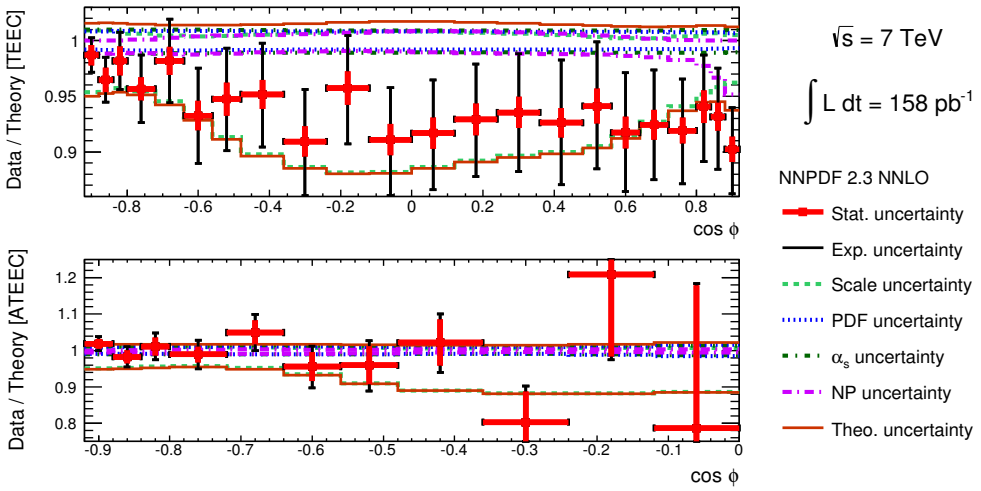

Figure 6.47: Ratios of the measured TEEC and ATEEC functions with respect to the theoretical predictions using NLO pQCD convoluted with the NNPDF 2.3 PDF.

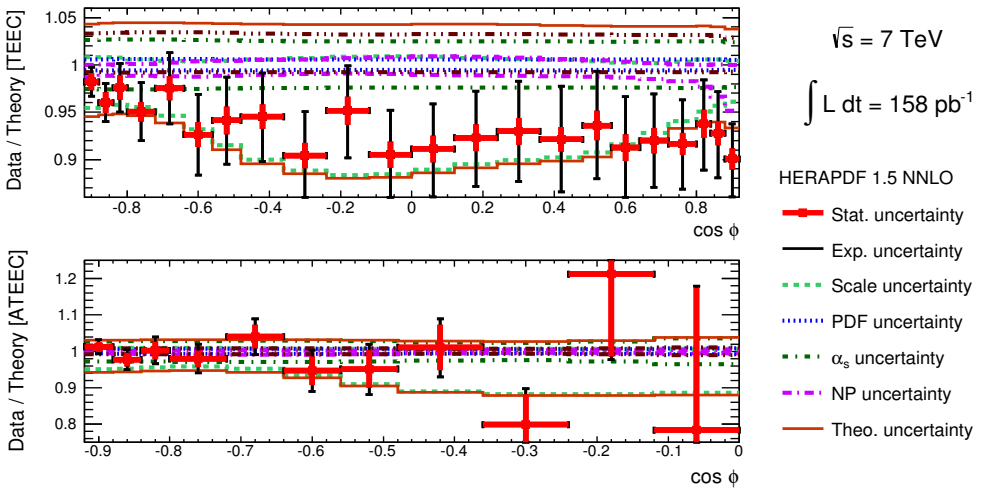

Figure 6.48: Ratios of the measured TEEC and ATEEC functions with respect to the theoretical predictions using NLO pQCD convoluted with the HERAPDF 1.5 PDF.

6.9 Determination of the strong coupling

The evaluation of the strong coupling constant $\alpha_s(m_Z)$ is made by minimizing a standard χ^2 function with correlations between the systematic uncertainties, which are parametrised using the nuisance parameters λ_k categorised in tables 6.7 and 6.8, ordered as a function of increasing impact on the TEEC measurement. The minimum of the χ^2 function is found in a 66-dimensional space, one dimension corresponding to $\alpha_s(m_Z)$ and the rest to the nuisance parameters associated with the experimental errors. The function to be minimised is defined as

$$\chi^2(\alpha_s, \vec{\lambda}) = \sum_{\text{bins}} \frac{(x_i - F_i(\alpha_s, \vec{\lambda}))^2}{\Delta x_i^2 + \Delta \xi_i^2} + \sum_k \lambda_k^2 \tag{6.22}$$

where the theoretical predictions are varied according to

$$F_i(\alpha_s, \vec{\lambda}) = \psi_i(\alpha_s) \left(1 + \sum_k \lambda_k \sigma_k^{(i)}\right) \tag{6.23}$$

In this expressions, x_i correspond to the data points in each distribution (TEEC or ATEEC), and Δx_i are their statistical uncertainties. $\Delta \xi_i$ are the statistical errors on the theoretical predictions, while $\sigma_k^{(i)}$ correspond to the k-th source of systematic uncertainty in the bin i. In a conservative way, the asymmetric systematic uncertainties are symmetrised to their maximum value for the χ^2 definition.

$$\sigma_k^{(i)} = \max\left\{\sigma_k^{(i)\uparrow}, \sigma_k^{(i)\downarrow}\right\} \tag{6.24}$$

The functions $\psi_i(\alpha_s)$ are analytical expressions parametrizing the dependence of each observable (TEEC or ATEEC) on the strong coupling constant. They are obtained by fitting the predictions for each bin as a function of $\alpha_s(m_Z)$. This function is chosen to be a parabola of the form

$$\psi_i(\alpha_s) = a_i \alpha_s^2 + b_i \alpha_s + c_i \tag{6.25}$$

As an example, Fig. 6.49 shows these parameterizations for the first and last bins of the TEEC function, while Fig. 6.50 shows the parabolas $\psi(\alpha_s)$ for the first two bins of the ATEEC function

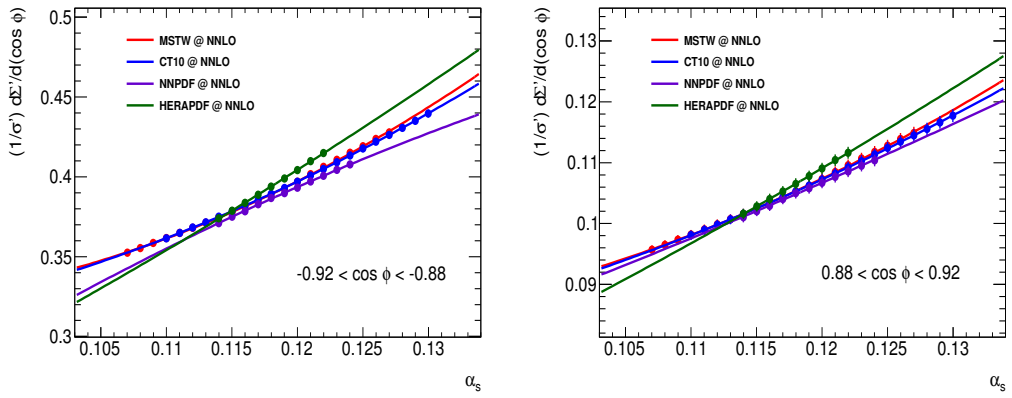

Figure 6.49: Dependence of the TEEC function on the strong coupling constant at the Z mass pole for $-0.92 < \cos\phi < -0.88$ (left) and $0.88 < \cos\phi < 0.92$ (right)

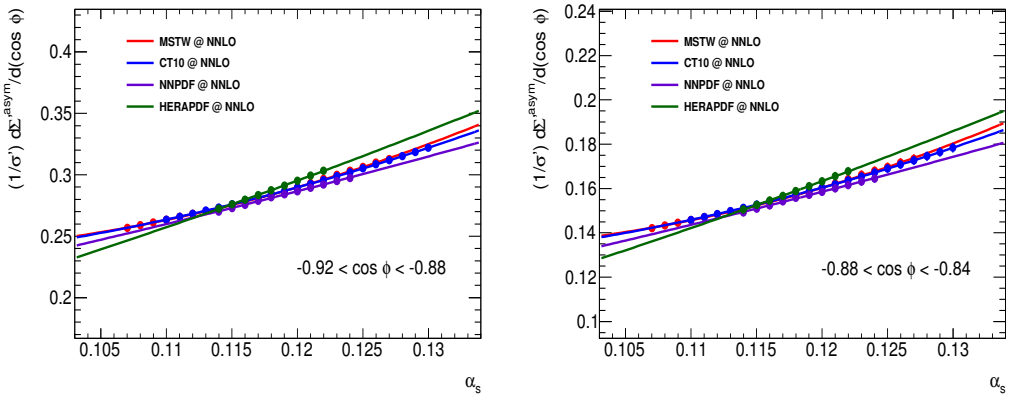

Figure 6.50: Dependence of the ATEEC function on the strong coupling constant at the Z mass pole for $-0.92 < \cos\phi < -0.88$ (left) and $-0.88 < \cos\phi < -0.84$ (right)

The quality of the fit to the NLO predictions is found to be excellent for each bin on both TEEC and ATEEC. This is expected, as the dependence on $\alpha_s(m_Z)$ on an NLO calculation is a polynomial of order 2.

The theoretical uncertainties on the predictions are treated by shifting the theoretical distributions by each independent source of uncertainty (scale variations, all independent PDF uncertainties and non-perturbative corrections) and repeating the fit using the modified theoretical inputs. For the scale uncertainty, the fits are repeated for each different variation of (μ_R, μ_F), and the two values of $\alpha_s(m_Z)$ which separate most from the nominal value are taken as the ones defining the envelope of the uncertainty.

Nuisance parameter number λ_k	Associated uncertainty σ
1	Insitu_Stat32
2	Insitu_Stat33
3	SingleParticle_HighPt
4	RelativeNonClosure_MC11b
5	Insitu_Stat31
6	Insitu_Stat28
7	Insitu_Stat29
8	Insitu_Stat30
9	Insitu_Stat27
10	Insitu_Stat23
11	Insitu_Stat13
12	Insitu_Stat15
13	Insitu_Stat25
14	Insitu_Stat22
15	Insitu_Stat14
16	Insitu_Stat12
17	Insitu_Stat26
18	Insitu_Stat16
19	Insitu_Stat24
20	Insitu_Stat17
21	MJB_Beta
22	MJB_Alpha
23	InSitu_Stat18
24	InSitu_Stat21
25	MPF_JER
26	InSitu_Stat3
27	InSitu_Stat20
28	InSitu_Stat2
29	InSitu_Stat19
30	InSitu_Stat10
31	Zjet_JVF
32	Zjet_Kterm

Table 6.7: The nuisance parameters used in the α_s extraction, ordered as a a function of increasing impact on the TEEC measurement (I)

Nuisance parameter number λ_k	Associated uncertainty σ
33	MJB_ISRFSR
34	InSitu_Stat4
35	MJB_CBJR
36	Insitu_Stat11
37	MJB_Threshold
38	InSitu_Stat8
39	InSitu_Stat5
40	InSitu_Stat9
41	InSitu_Stat1
42	MPF_J2
43	MJB_Intercalibration
44	JER
45	MJB_Fragmentation
46	InSitu_Stat7
47	InSitu_Stat6
48	Zjet_Extrapolation
49	MJB_Asym
50	Zjet_Width
51	MPF_Generator
52	MPF_Pileup
53	EtaIntercalibration_TotalStat
54	MPF_Purity
55	MPF_OOC
56	Zjet_Veto
57	Zjet_MC
58	Unfolding (Closure Test)
59	InSitu_LArEMScale
60	EtaIntercalibration_Modelling
61	Close_By
62	Pileup
63	Shower modelling
64	Flavor_Comp
65	Flavor_Response

Table 6.8: The nuisance parameters used in the α_s extraction ordered as a function of increasing impact on the TEEC measurement (II)

6.9.1 TEEC fits

In this section, the fits for the energy-energy correlation function are presented using four different PDFs, namely MSTW, CT10, NNPDF and HERAPDF, al of them at the NNLO level of precision. Figure 6.51 shows the fit result for each nuisance parameter, together with the correlation coefficients between nuisance parameters for the different choices of the PDF.

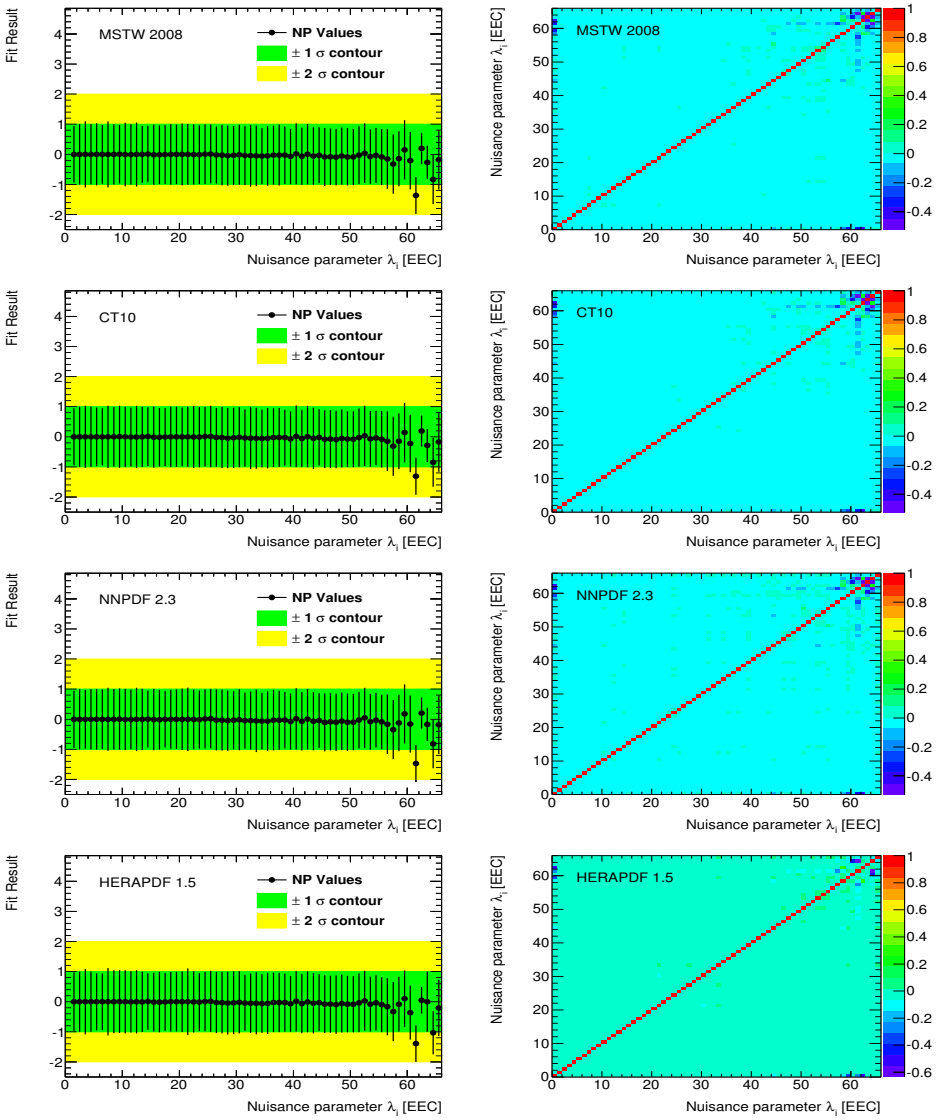

Figure 6.51: Extracted values of the nuisance parameters (left) and correlation matrices (right) for the TEEC fits using different PDFs.

The left hand plots on Fig. 6.51 show that most nuisance parameters remain stable at values close to 0. However, some of them deviate from this value, specially those accounting for the topological and flavor JES sources (Close_By, Flavor_Comp and Flavor_Response), together with the Jet Energy Resolution. In any case, all of these deviations are within 1σ. On the right-hand side of Fig. 6.51, one can see that those parameters that are deviating from 0 show some anticorrelations between them. These fits were repeated using the reduced set of JES nuisance parameters, with very similar results. The results for the value of the strong coupling constant are summarised in Table 6.9

PDF	α_s value	χ^2/N_{dof}
MSTW 2008	0.1175 ± 0.0010 (exp.) $^{+0.0059}_{-0.0019}$ (scale) $^{+0.0006}_{-0.0006}$ (PDF) $^{+0.0002}_{-0.0002}$ (NPC)	29.0 / 21
CT10	0.1173 ± 0.0010 (exp.) $^{+0.0063}_{-0.0020}$ (scale) $^{+0.0017}_{-0.0017}$ (PDF) $^{+0.0002}_{-0.0002}$ (NPC)	28.4 / 21
NNPDF 2.3	0.1183 ± 0.0010 (exp.) $^{+0.0059}_{-0.0013}$ (scale) $^{+0.0009}_{-0.0009}$ (PDF) $^{+0.0002}_{-0.0002}$ (NPC)	29.3 / 21
HERAPDF 1.5	0.1167 ± 0.0007 (exp.) $^{+0.0040}_{-0.0008}$ (scale) $^{+0.0004}_{-0.0004}$ (PDF) $^{+0.0001}_{-0.0001}$ (NPC) $^{+0.0006}_{-0.0024}$ (MOD)	28.7 / 21

Table 6.9: The results for α_s from fits to the TEEC using different PDFs

All values are compatible within the PDF uncertainties obtained using the eigenvector variations, as discussed in Sect. 6.8.1.2. The uncertainties for the case of HERAPDF are significantly smaller because of the different shape of the parabolic behaviour of the TEEC function with α_s.
The χ^2 profile for the nominal value of α_s is shown in Fig 6.52 for the case of the MSTW parton densities.

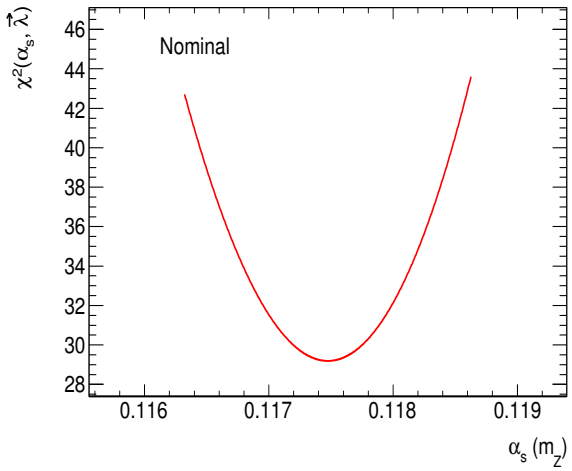

Figure 6.52: The α_s-projection of the χ^2 function for the nominal TEEC fit.

Table 6.10 shows the values of α_s from the scale variations, while Fig. 6.53 shows their χ^2 profiles

Scale Variation	$\alpha_s(m_Z)$	Exp. error	χ^2/N_{dof}
$(1/2, 1/2)$	0.1164	0.0010	49.5 / 21
$(1/2, 1)$	0.1156	0.0010	33.2 / 21
$(1, 1/2)$	0.1163	0.0010	27.9 / 21
$(1, 2)$	0.1188	0.0009	30.6 / 21
$(2, 1)$	0.1216	0.0009	32.6 / 21
$(2, 2)$	0.1233	0.0009	33.9 / 21

Table 6.10: The α_s values for the scale variations

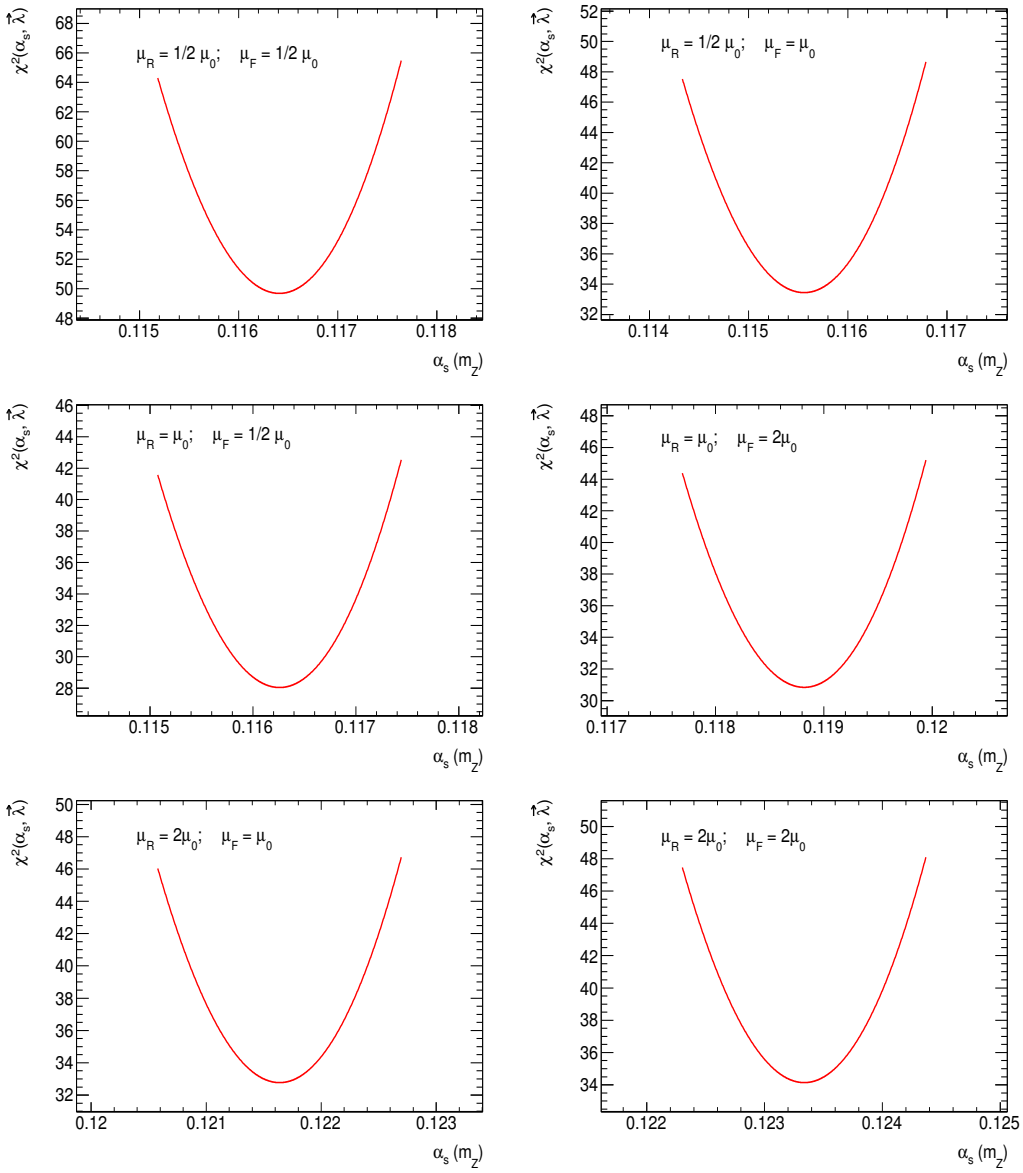

Figure 6.53: The α_s-projections of the $\chi^2(\alpha_s, \vec{\lambda})$ functions for the scale variations in the TEEC fit

Table 6.11 shows the values of α_s from NP variations, while Fig. 6.54 shows their χ^2 profiles

NP Variation	$\alpha_s(m_Z)$	Exp. error	χ^2/N_{dof}
Up	0.1177	0.0010	29.7 / 21
Down	0.1173	0.0010	29.2 / 21

Table 6.11: The α_s values for the scale variations

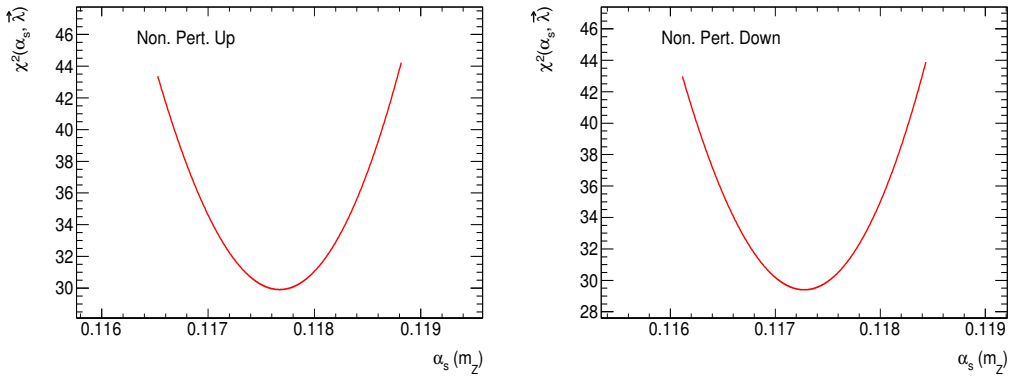

Figure 6.54: The α_s-projections of the $\chi^2(\alpha_s, \vec{\lambda})$ functions for the NP variations in the TEEC fit

To illustrate the meaning of tables 6.10 and 6.11 let us look at table 6.11. The difference between $\alpha_s(m_Z)$ with the Up variation for the NP correction in Table 6.11, 0.1178, and the nominal value in Eq. 6.9 is ascribed to the negative systematic error due to the NP correction, 0.0002. Similarly, the maximum differences between the nominal value and the scale variations obtained in Tab. 6.10 are identified as the scale uncertainties. The numerical values for the Data/Theory comparison after the fit are shown in Table 6.12, together with the partial values of the χ^2 function. From this table, it is clear that the quality of the fit is good overall, with the exception of the points near $\cos\phi \sim 1$, i.e. the region where self-correlations are important, though no clear systematic effects are seen.

$\cos\phi$	Data	Data Stat.	Theory	Theory Stat.	Partial χ^2
(-0.92 , -0.88)	0.3848	0.0029	0.3826	0.0020	0.3948
(-0.88 , -0.84)	0.2324	0.0022	0.2353	0.0016	1.1816
(-0.84 , -0.80)	0.1612	0.0017	0.1594	0.0013	0.6784
(-0.80 , -0.72)	0.1095	0.0009	0.1102	0.0007	0.3480
(-0.72 , -0.64)	0.0767	0.0008	0.0746	0.0006	4.6344
(-0.64 , -0.56)	0.0574	0.0006	0.0583	0.0005	1.1605
(-0.56 , -0.48)	0.0472	0.0005	0.0468	0.0005	0.1997
(-0.48 , -0.36)	0.0400	0.0004	0.0394	0.0004	1.4070
(-0.36 , -0.24)	0.0329	0.0004	0.0337	0.0004	2.5587
(-0.24 , -0.12)	0.0302	0.0003	0.0293	0.0003	3.4666
(-0.12 , 0.00)	0.0273	0.0003	0.0278	0.0003	1.1616
(0.00 , 0.12)	0.0262	0.0003	0.0265	0.0003	0.3654
(0.12 , 0.24)	0.0264	0.0003	0.0264	0.0003	0.0036
(0.24 , 0.36)	0.0272	0.0003	0.0270	0.0003	0.1453
(0.36 , 0.48)	0.0286	0.0003	0.0287	0.0003	0.1426
(0.48 , 0.56)	0.0306	0.0004	0.0303	0.0004	0.3150
(0.56 , 0.64)	0.0340	0.0004	0.0345	0.0004	0.9510
(0.64 , 0.72)	0.0391	0.0004	0.0394	0.0004	0.2150
(0.72 , 0.80)	0.0487	0.0004	0.0490	0.0005	0.2396
(0.80 , 0.84)	0.0639	0.0007	0.0625	0.0007	1.8771
(0.84 , 0.88)	0.0780	0.0008	0.0769	0.0008	1.0198
(0.88 , 0.92)	0.0955	0.0009	0.0978	0.0009	3.5268

Table 6.12: The numerical comparison of the TEEC data versus the theory after the fit, together with the partial χ^2 values.

The λ-term in the χ^2 function (the quadratic sum of all the nuisance parameters) amounts up to 3.01 for the TEEC fit. Figure 6.55 shows the comparison of the data with the theory after the fit, where the systematic uncertainty is calculated taking into account the correlations between the systematic uncertainties, i.e.

$$\Delta x^2 = \sum_{ij} C_{ij}\sigma_i\sigma_j = \sum_i \sigma_i^2 + 2\sum_{i<j} C_{ij}\sigma_i\sigma_j \tag{6.26}$$

Additionally, to estimate the impact of the statistical uncertainty on the total experimental uncertainty, the fit was redone setting all nuisance parameters to 0. The result is that the statistical uncertainty represents a 0.3% of the value of $\alpha_s(m_Z)$, in contrast with the total 0.7%.

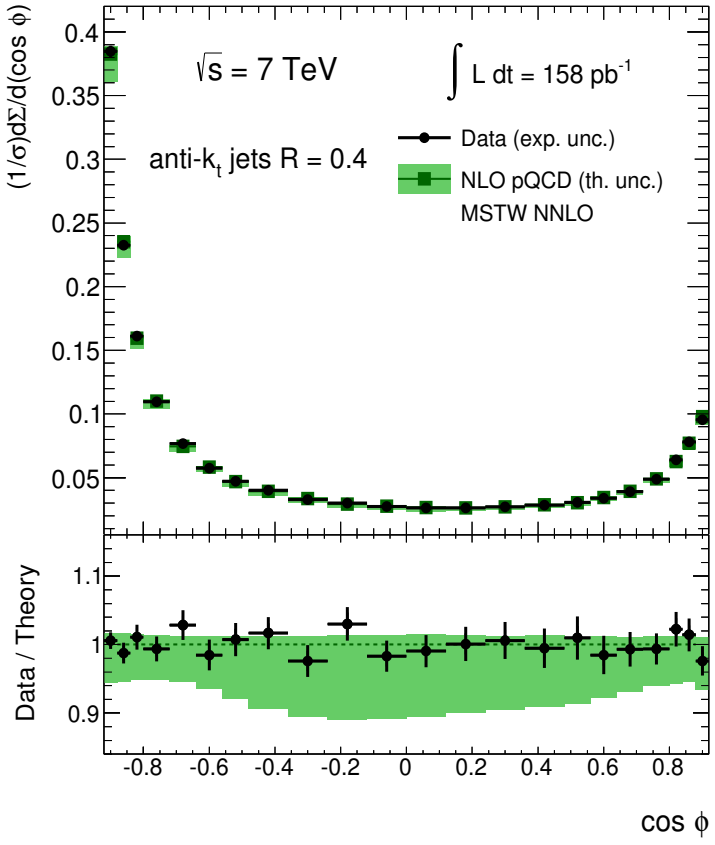

Figure 6.55: Comparison of the TEEC data and the theoretical predictions (modified using the optimal nuisance parameters) for the TEEC-fitted value of α_s. The green band shows the theoretical uncertainties, including the α_s uncertainty.

6.9.2 ATEEC fits

In this section, the fits for the energy-energy correlation asymmetry function are presented. Figure 6.56 shows the fit results for each nuisance parameter, together with the correlation coefficients between nuisance parameters for the different PDFs used in the analysis.

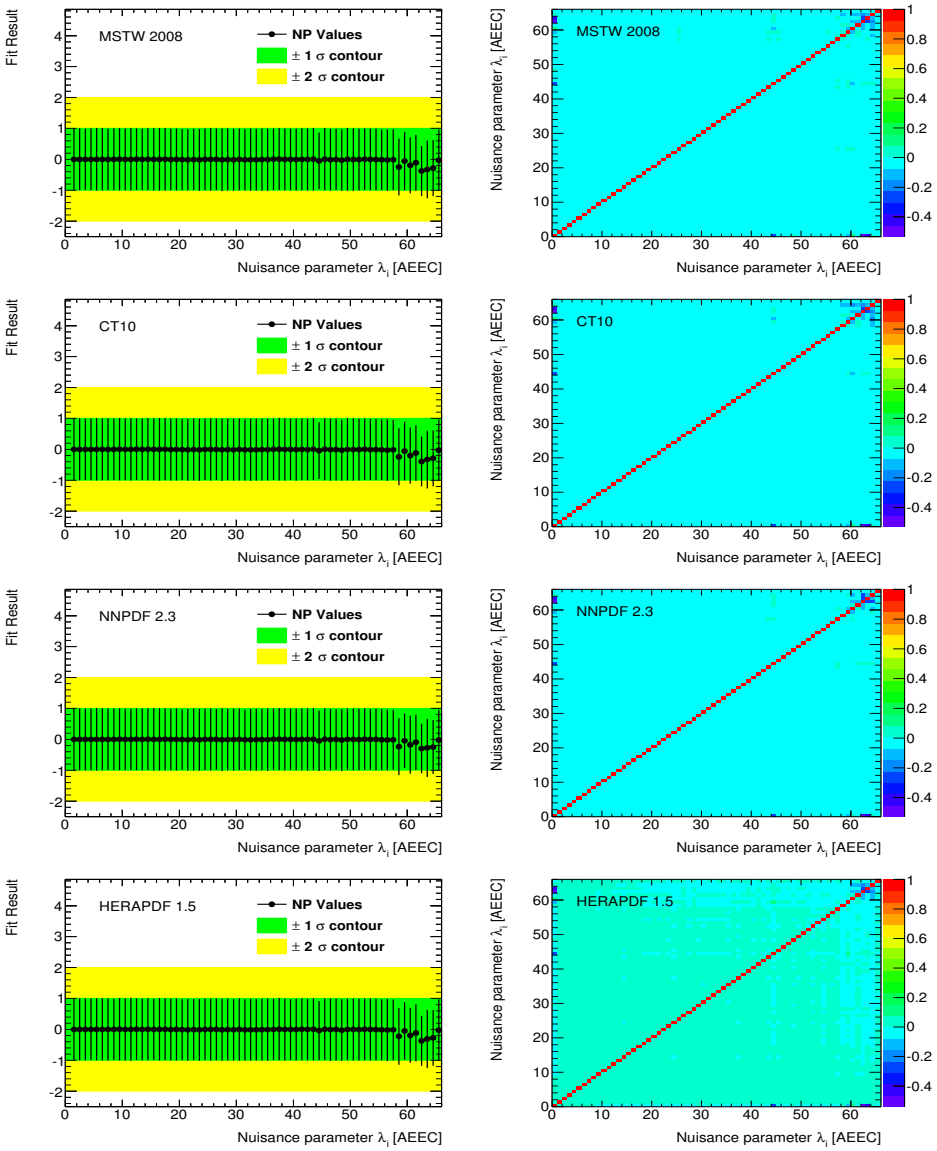

Figure 6.56: Extracted values of the nuisance parameters (left) and correlation matrices (right) for the ATEEC fits using different PDFs.

In this case, as one can see in the left-hand side of Fig. 6.56, the values of the nuisance parameters always remain well behaved, as their values are very close to 0 and therefore the quality of the fits in terms of the values of the χ^2 are better than in the TEEC case. Here again, the correlation matrices shown in the right-hand side of Fig. 6.56 are pretty diagonal and exhibits small anticorrelations.

The results for the value of the strong coupling constant are summarised in Table 6.13

PDF	α_s value	χ^2/N_{dof}
MSTW 2008	0.1195 ± 0.0017 (exp.) $^{+0.0055}_{-0.0015}$ (scale) $^{+0.0006}_{-0.0006}$ (PDF) $^{+0.0000}_{-0.0000}$ (NPC)	12.7 / 10
CT10	0.1195 ± 0.0018 (exp.) $^{+0.0060}_{-0.0015}$ (scale) $^{+0.0016}_{-0.0016}$ (PDF) $^{+0.0000}_{-0.0000}$ (NPC)	12.6 / 10
NNPDF 2.3	0.1206 ± 0.0018 (exp.) $^{+0.0057}_{-0.0013}$ (scale) $^{+0.0009}_{-0.0009}$ (PDF) $^{+0.0000}_{-0.0000}$ (NPC)	12.2 / 10
HERAPDF 1.5	0.1182 ± 0.0013 (exp.) $^{+0.0041}_{-0.0008}$ (scale) $^{+0.0004}_{-0.0004}$ (PDF) $^{+0.0000}_{-0.0000}$ (NPC) $^{+0.0006}_{-0.0025}$ (MOD)	12.1 / 10

Table 6.13: The results for α_s from fits to the ATEEC using different PDFs

Here, once again all values are compatible within the uncertainties. The uncertainties for the case of HERAPDF are significantly smaller because of the different shape of the parabolic behaviour of the ATEEC function with α_s.
The χ^2 profile for the nominal value of α_s is shown in Fig 6.52 for the extraction using the MSTW 2008 PDFs

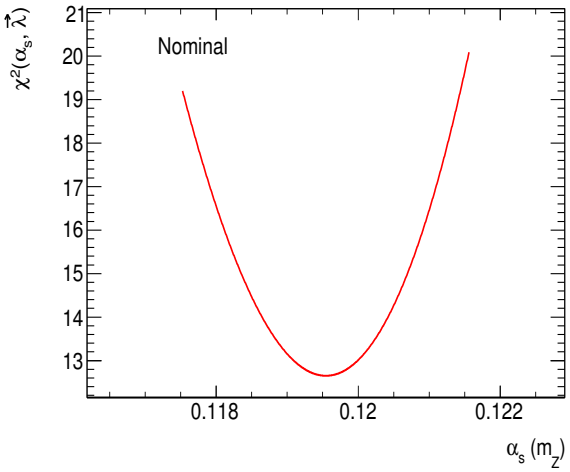

Figure 6.57: The α_s-projection of the $\chi^2(\alpha_s, \vec{\lambda})$ functions for the nominal ATEEC fit.

Table 6.14 shows the values of α_s from the scale variations, while Fig. 6.58 shows their χ^2 profiles

Scale Variation	$\alpha_s(m_Z)$	Exp. error	χ^2/N_{dof}
$(1/2, 1/2)$	0.1191	0.0017	11.9 / 10
$(1/2, 1)$	0.1181	0.0017	11.3 / 10
$(1, 1/2)$	0.1185	0.0017	12.1 / 10
$(1, 2)$	0.1207	0.0016	13.2 / 10
$(2, 1)$	0.1235	0.0016	14.3 / 10
$(2, 2)$	0.1251	0.0016	14.7 / 10

Table 6.14: The α_s values for the scale variations in the ATEEC fit.

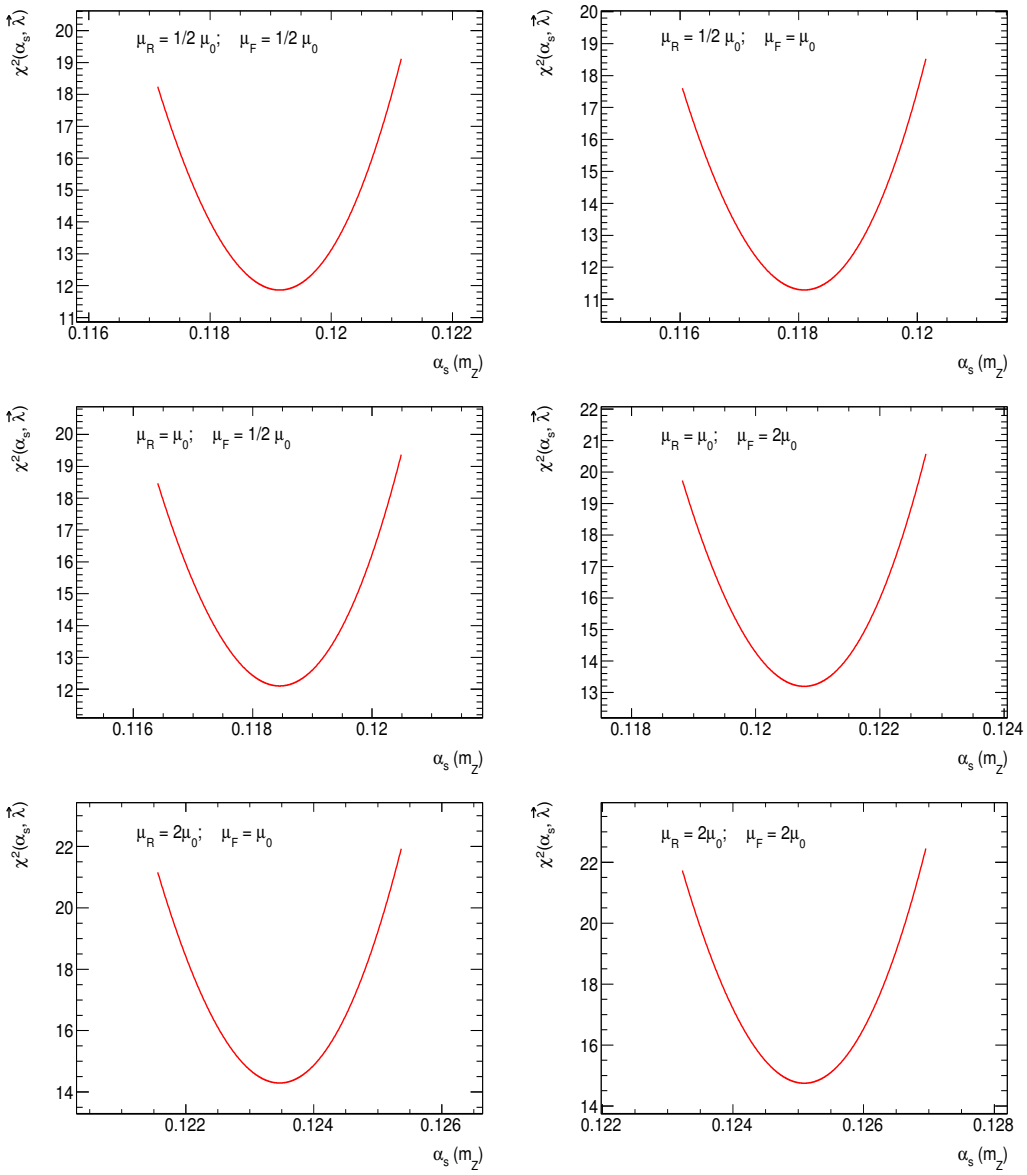

Figure 6.58: The α_s-projections of the $\chi^2(\alpha_s, \vec{\lambda})$ functions for the scale variations in the ATEEC fit.

Table 6.15 shows the values of α_s from NP variations, while Fig. 6.59 shows their χ^2 profiles

NP Variation	$\alpha_s(m_Z)$	Exp. error	χ^2/N_{dof}
Up	0.1196	0.0017	12.1 / 10
Down	0.1195	0.0017	13.3 / 10

Table 6.15: The α_s values for the scale variations

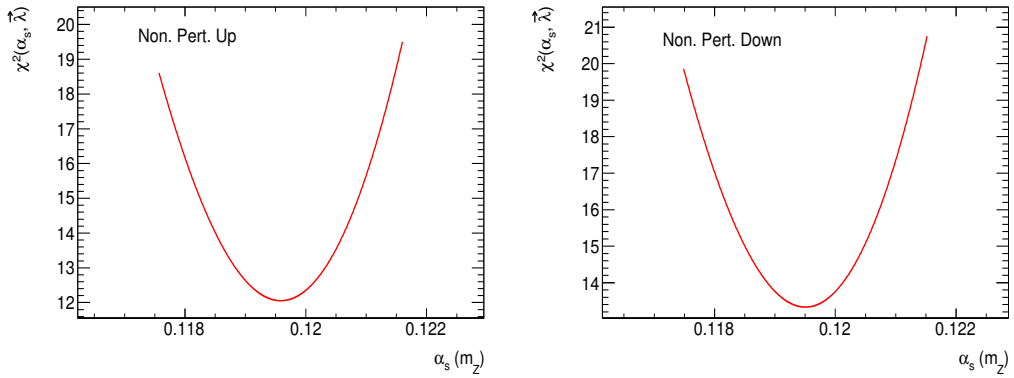

Figure 6.59: The α_s-projections of the $\chi^2(\alpha_s, \vec{\lambda})$ functions for the NP variations in the ATEEC fit.

The numerical values for the Data/Theory comparison after the fit is shown in Table 6.16, together with the partial values of the χ^2 function. In this case, the goodness of the fit extends to the whole fitted region, with partial χ^2 values below 1 in most of the bins.

The λ-term in the χ^2 function (the quadratic sum of all the nuisance parameters) amounts up to 0.45 for the ATEEC fit. Figure 6.60 shows the comparison of the data with the theory after the fit, with the systematic uncertainties estimated according to Eq. 6.26. The statistical uncertainty in the fit was estimated in the same way as for the TEEC fit, by setting all nuisance parameters to 0. The result is that the statistical uncertainty represents a 0.6% of the value of $\alpha_s(m_Z)$, to be compared with the 1.2% of the total experimental uncertainty.

$\cos\phi$	Data	Data Stat.	Theory	Theory Stat.	Partial χ^2
(-0.92 , -0.88)	0.2893	0.0030	0.2862	0.0022	0.6970
(-0.88 , -0.84)	0.1544	0.0023	0.1578	0.0017	1.3875
(-0.84 , -0.80)	0.0973	0.0019	0.0965	0.0015	0.1002
(-0.80 , -0.72)	0.0608	0.0010	0.0615	0.0009	0.2570
(-0.72 , -0.64)	0.0376	0.0009	0.0357	0.0008	2.9757
(-0.64 , -0.56)	0.0235	0.0007	0.0244	0.0007	0.8482
(-0.56 , -0.48)	0.0165	0.0007	0.0170	0.0006	0.2838
(-0.48 , -0.36)	0.0115	0.0005	0.0111	0.0005	0.3269
(-0.36 , -0.24)	0.0057	0.0004	0.0070	0.0005	3.7010
(-0.24 , -0.12)	0.0038	0.0004	0.0031	0.0005	1.4017
(-0.12 , 0.00)	0.0011	0.0004	0.0014	0.0004	0.2207

Table 6.16: The numerical comparison of the data versus the theory after the fit, together with the partial χ^2 values in the ATEEC fit.

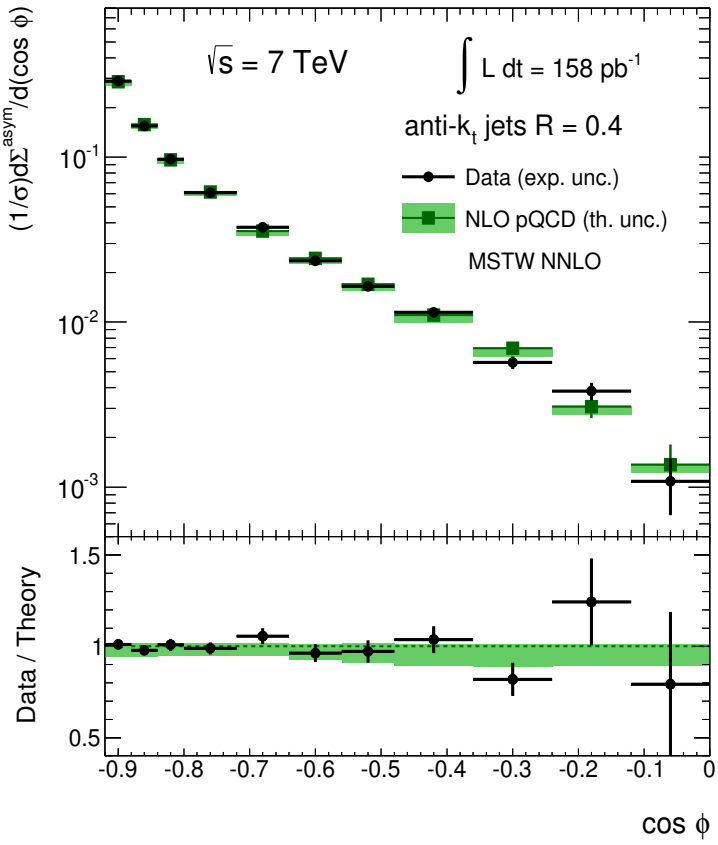

Figure 6.60: Comparison of the ATEEC data and the theoretical predictions (modified using the optimal nuisance parameters) for the ATEEC-fitted value of α_s. The green band shows the theoretical uncertainties, including the α_s uncertainty.

6.9.3 Cross checks: Alternative JES configurations

As it was mentioned in section 6.6.1, two alternative JES scenarios have been investigated: one with stronger correlations between the JES nuisance parameters and one with weaker ones with respect to the default. The stronger correlation configuration contains a total of 54 parameters, including those accounting for pileup, flavor response and composition and event topology. On the other hand, the weaker correlation scenario has a total of 65 nuisance parameters, also including those for pileup, flavor and topology.

6.9.3.1 Fits with stronger correlation configuration

The fits described in the previous sections have been repeated using the stronger correlation configuration. The results for both α_s and its uncertainties agree very well with the default ones. In Figure 6.61, the values of the nuisance parameters and their correlation matrix are shown for the fit to the TEEC function, while Figure 6.62 shows the same for the ATEEC fit.

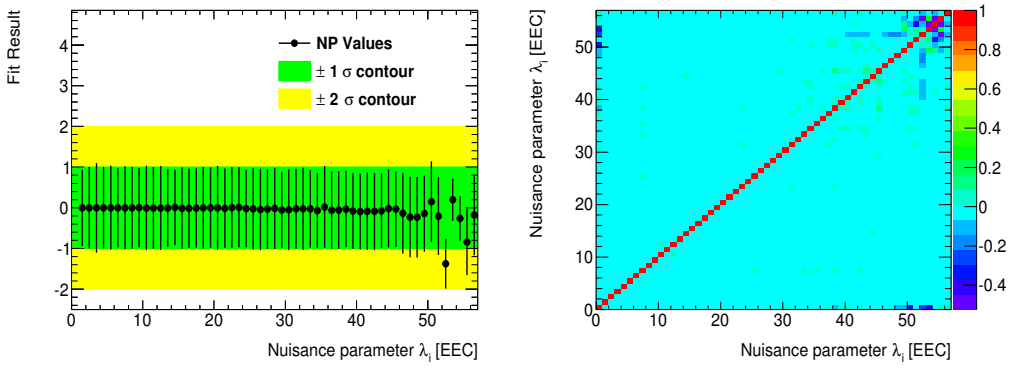

Figure 6.61: Extracted value of the nuisance parameters (left) and correlation matrix (right) for the TEEC fit in the JES stronger correlation scenario.

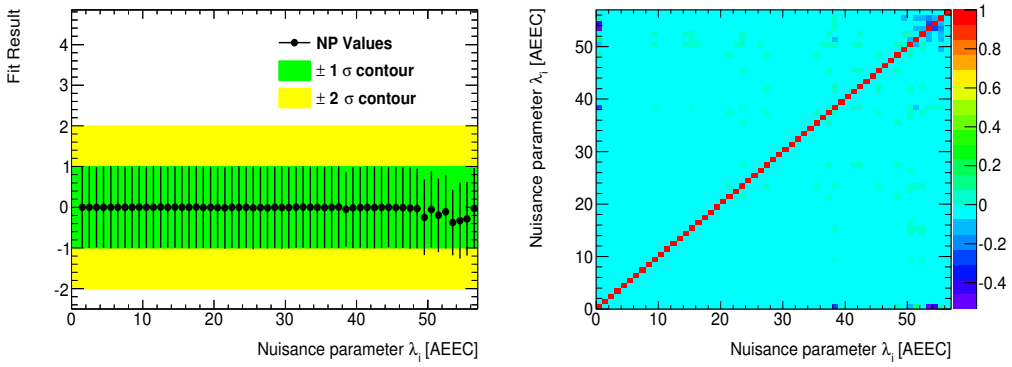

Figure 6.62: Extracted value of the nuisance parameters (left) and correlation matrix (right) for the ATEEC fit in the JES stronger correlation scenario.

The α_s value obtained for the TEEC fit is

$$\alpha_s(m_Z) = 0.1175 \pm 0.0010\ (\text{exp.})\ ^{+0.0058}_{-0.0019}\ (\text{scale})$$
$$^{+0.0006}_{-0.0006}\ (\text{PDF})\ ^{+0.0002}_{-0.0002}\ (\text{NPC}) \tag{6.27}$$

While the value obtained for the ATEEC fit is

$$\alpha_s(m_Z) = 0.1195 \pm 0.0017\ (\text{exp.})\ ^{+0.0055}_{-0.0014}\ (\text{scale})$$
$$^{+0.0006}_{-0.0006}\ (\text{PDF})\ ^{+0.0000}_{-0.0000}\ (\text{NPC}) \tag{6.28}$$

The χ^2/n.d.o.f values are 29.0/21 for the TEEC and 12.6/10 for the ATEEC, very close to those obtained in the default JES scenario. Comparing these values with those in Tables 6.9 and 6.13 for MSTW leads to the conclusion that the use of stronger correlations for JES has no significative impact in either the values of α_s or their uncertainties.

6.9.3.2 Fits with weaker correlation configuration

The fits are now repeated using the weaker correlation configuration. In Figure 6.63, the values of the nuisance parameters and their correlation matrix are shown for the fit to the TEEC function, while Figure 6.64 shows the same for the ATEEC fit.

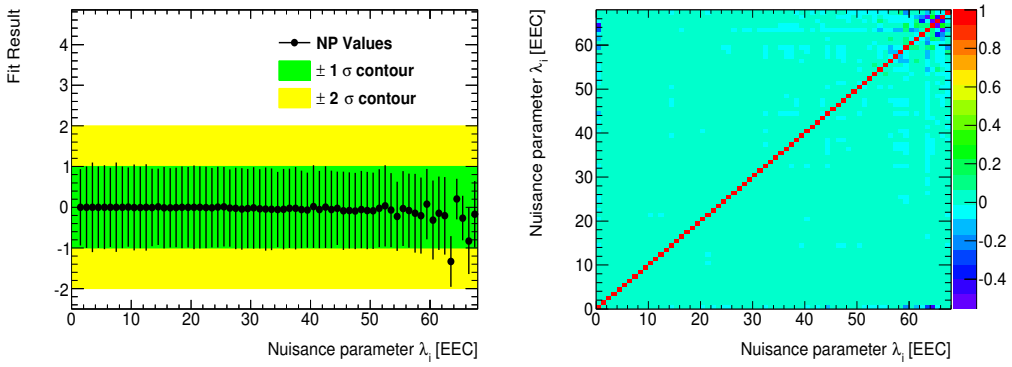

Figure 6.63: Extracted value of the nuisance parameters (left) and correlation matrix (right) for the TEEC fit in the JES weaker correlation scenario.

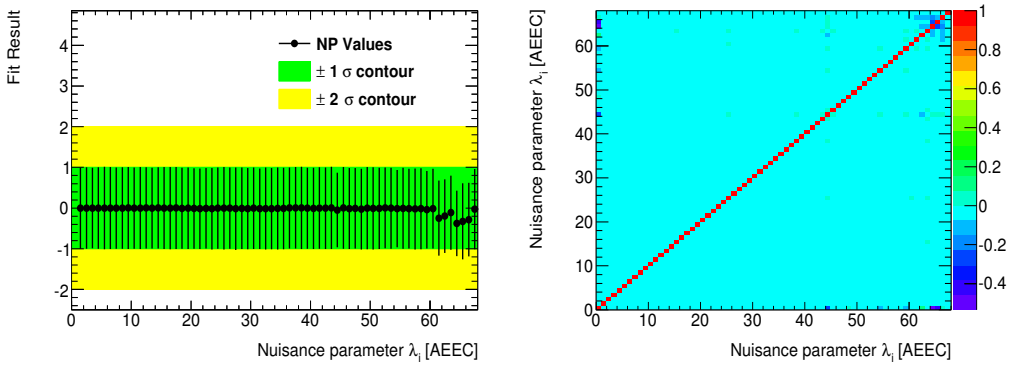

Figure 6.64: Extracted value of the nuisance parameters (left) and correlation matrix (right) for the ATEEC fit in the JES weaker correlation scenario.

The α_s value obtained for the TEEC fit is

$$\alpha_s(m_Z) = 0.1175 \pm 0.0009\ (\text{exp.})\ ^{+0.0058}_{-0.0019}\ (\text{scale})$$
$$^{+0.0006}_{-0.0006}\ (\text{PDF})\ ^{+0.0002}_{-0.0002}\ (\text{NPC}) \tag{6.29}$$

While the value obtained for the ATEEC fit is

$$\alpha_s(m_Z) = 0.1195 \pm 0.0017\ (\text{exp.})\ ^{+0.0055}_{-0.0015}\ (\text{scale})$$
$$^{+0.0006}_{-0.0006}\ (\text{PDF})\ ^{+0.0000}_{-0.0000}\ (\text{NPC}) \tag{6.30}$$

The χ^2/n.d.o.f values are 28.9/21 for the TEEC and 12.6/10 for the ATEEC,

which are also very close to those obtained in the default JES scenario. The comparison of these values with those in Tables 6.9 and 6.13 for the MSTW PDFs leads to the conclusion that the use of weaker correlations for JES has no significative impact in either the values of α_s or their uncertainties.

6.10 Summary and conclusions

The jet-wise energy-energy correlation measurement at the LHC constitutes a high-precision test of QCD at the TeV scale, and a sensitive way to extract the strong coupling constant. Its reduced dependence on the parton densities inside the proton makes it ideal for this purpose. The experimental uncertainties are also reduced due to the fact that the TEEC and its asymmetry exhibit a very small sensitivity to both migration effects and jet energy uncertainties. This analysis is the first look at the TEEC and its asymmetry in terms of jets and at energy scales much higher than those available from past colliders such as LEP or HERA.

Figure 6.65 shows the ATEEC distribution obtained in this study together with the ones obtained in TASSO for e^+e^- center-of-mass energies of 34.8 and 45.3 GeV (see Ref. [124]). Note that the shape of both distributions are very similar, with the ATLAS data, which is at a higher scale, lying consistently below the curves from TASSO. This comparison has to be taken with a grain of salt since the TASSO (ATLAS) data are measured at the hadron (jet) level, and the TASSO (ATLAS) data refers to angles measured in full space (the transverse plane). However since the fragmentation effects for the asymmetry at large angles are very small, the comparison is justified.

The results show that both the TEEC and its forward-backward asymmetry ATEEC are stable with respect to changes in the jet energy scale and resolution and they exhibit a small sensitivity to pileup activity.

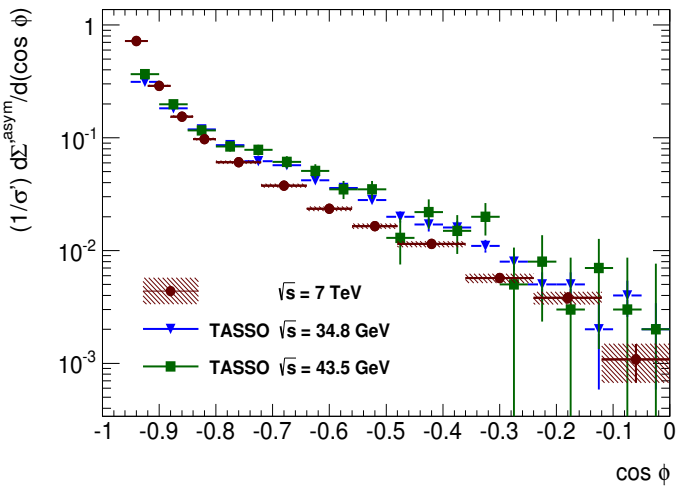

Figure 6.65: Comparison of the data obtained in this analysis with those from TASSO in Ref. [124]

The final result for $\alpha_s(m_Z)$ determined using the TEEC is

$$\alpha_s(m_Z) = 0.1173 \pm 0.0010\ (\text{exp.})\ ^{+0.0063}_{-0.0020}\ (\text{scale})\ ^{+0.0017}_{-0.0017}\ (\text{PDF})\ ^{+0.0002}_{-0.0002}\ (\text{NPC}) \tag{6.31}$$

while the value obtained using the ATEEC is

$$\alpha_s(m_Z) = 0.1195 \pm 0.0018\ (\text{exp.})\ ^{+0.0060}_{-0.0015}\ (\text{scale})\ ^{+0.0016}_{-0.0016}\ (\text{PDF})\ ^{+0.0000}_{-0.0000}\ (\text{NPC}) \tag{6.32}$$

These values are in very good agreement with the world average $\alpha_s(m_Z) = 0.1185 \pm 0.0006$, as well as with several other α_s extractions from the data collected at the LHC [136, 137]. These two values are obtained using the CT10 PDF. Similar results have been obtained with MSTW 2008, NNPDF 2.3 and HERAPDF 1.5 as discussed in the previous sections. The present determination of $\alpha_s(m_Z)$ is limited by the theoretical uncertainties due to the choice of the renormalisation and factorisation scales, which reflect our lack of knowledge about NNLO corrections. This limitation has also been found in other jet studies at LHC energies [136]. This is in contrast with processes like top-quark pair production or Higgs production, for which corrections beyond NLO do exist [144].

A Data to MC comparisons for jet shapes

In this appendix, several Monte Carlo tunes are compared to the jet shape data in $t\bar{t}$ events presented in Chapter 4. The investigated samples include ALPGEN coupled to PYTHIA P2011 and HERWIG + JIMMY, POWHEG+PYTHIA P2011C and ACERMC coupled to PYTHIA P2010, P2011, TUNEA PRO and TUNE A CR PRO. Figures 66 to 70 show this comparison for the differential jet shapes, while Fig. 71 to 75 includes the comparison for the integrated jet shape. It is found that the PYTHIA parton shower with tune A with different models of colour reconnection give the best description of the data, while the PERUGIA tunes are found to be slightly disfavoured.

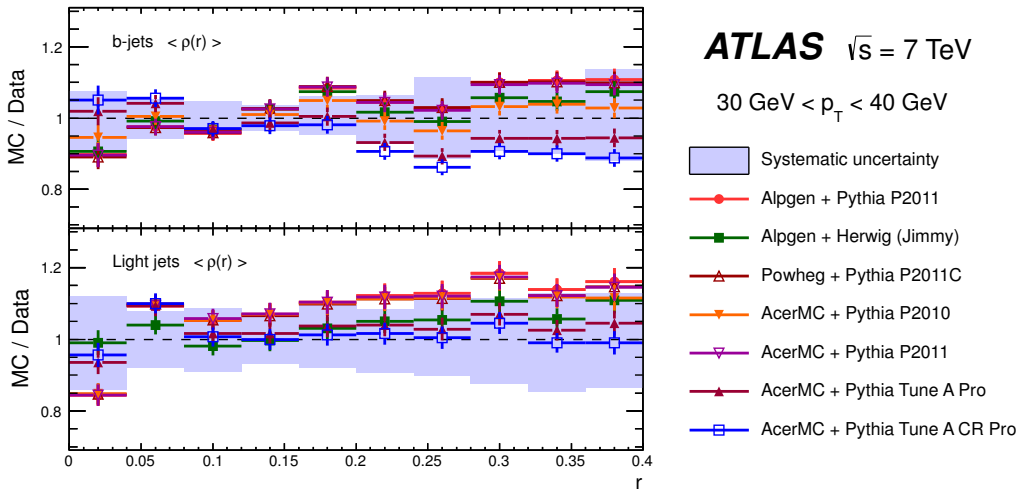

Figure 66: Comparison of the $t\bar{t}$ differential jet shape data for 30 GeV $< p_{\mathrm{T}} <$ 40 GeV with several MC event generators.

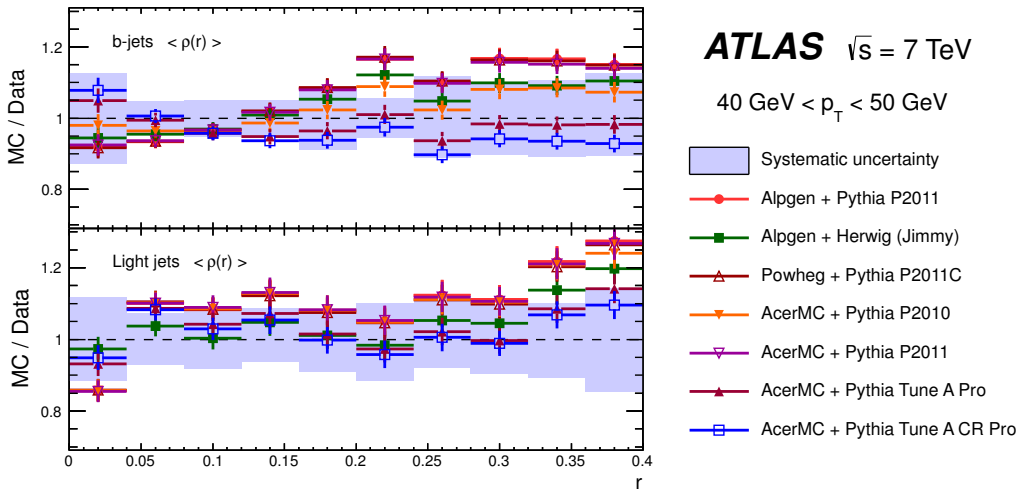

Figure 67: Comparison of the $t\bar{t}$ differential jet shape data for 40 GeV $< p_{\mathrm{T}} <$ 50 GeV with several MC event generators.

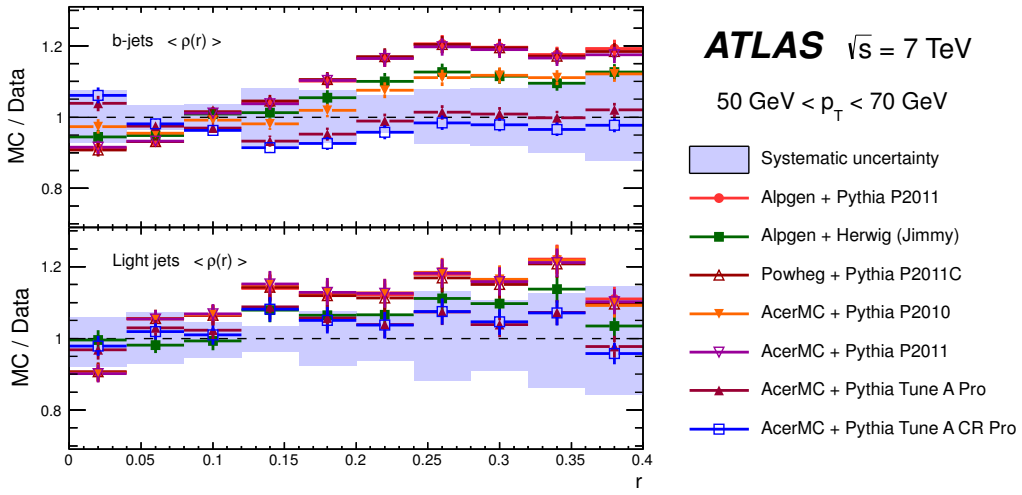

Figure 68: Comparison of the $t\bar{t}$ differential jet shape data for 50 GeV $< p_{\mathrm{T}} <$ 70 GeV with several MC event generators.

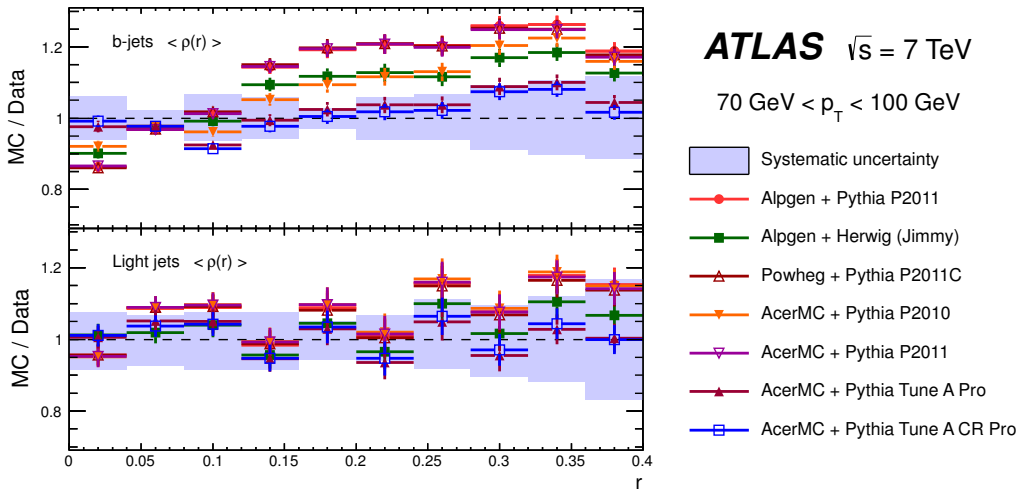

Figure 69: Comparison of the $t\bar{t}$ differential jet shape data for 70 GeV $< p_{\mathrm{T}} <$ 100 GeV with several MC event generators.

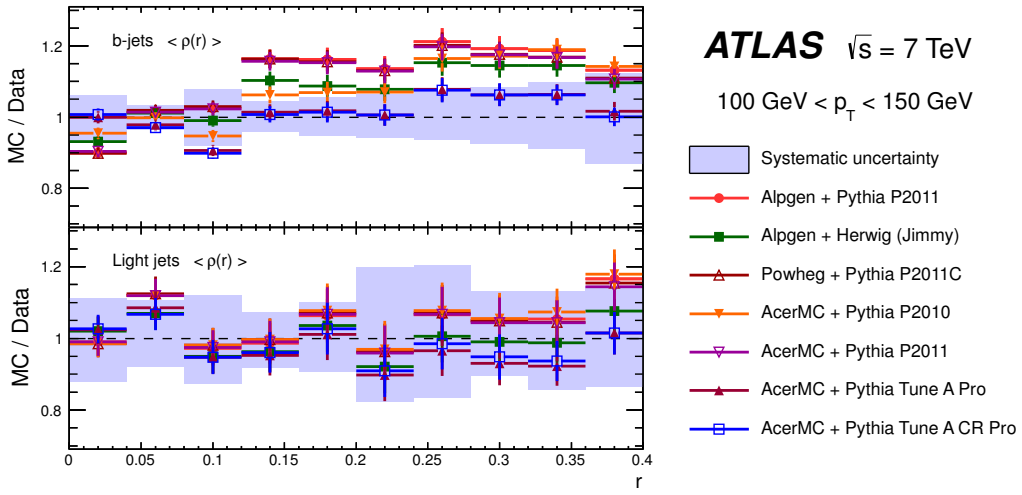

Figure 70: Comparison of the $t\bar{t}$ differential jet shape data for 100 GeV $< p_{\mathrm{T}} <$ 150 GeV with several MC event generators.

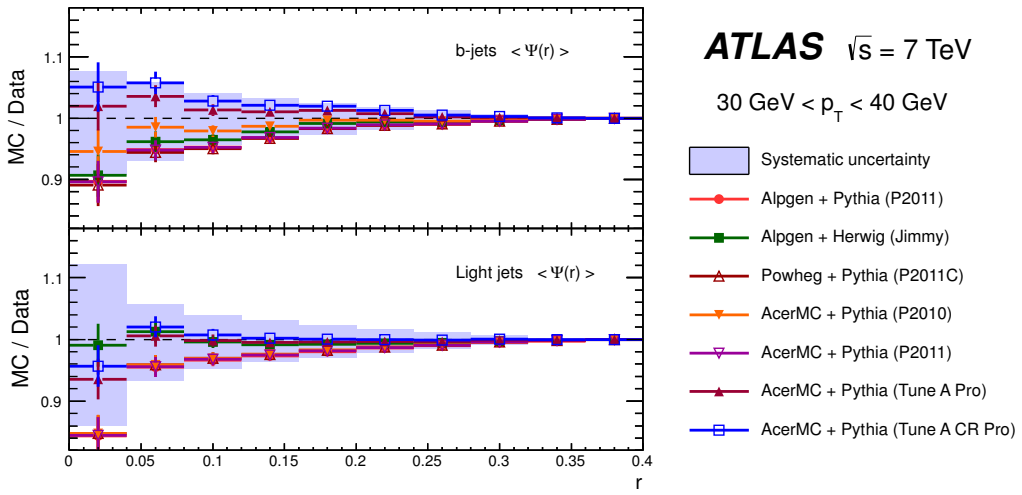

Figure 71: Comparison of the $t\bar{t}$ integrated jet shape data for 30 GeV $< p_{\mathrm{T}} <$ 40 GeV with several MC event generators.

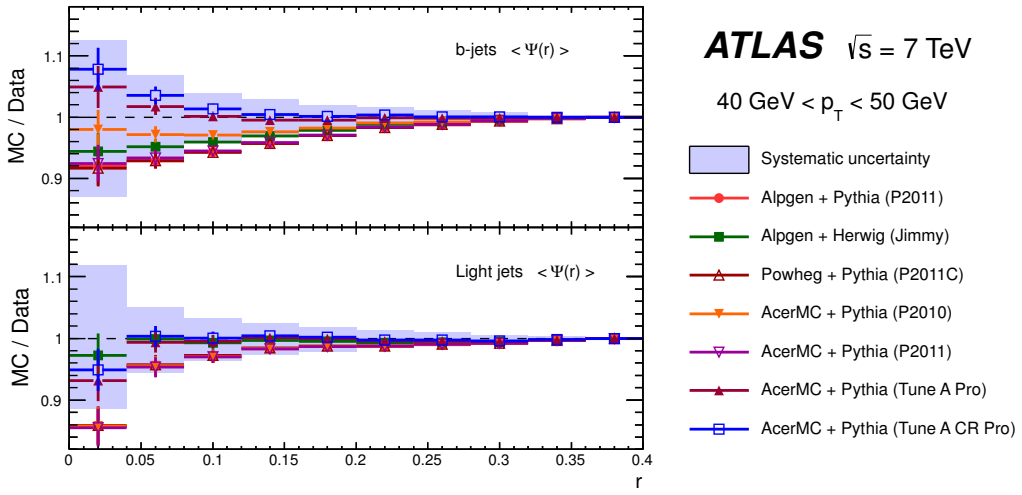

Figure 72: Comparison of the $t\bar{t}$ integrated jet shape data for 40 GeV $< p_{\mathrm{T}} <$ 50 GeV with several MC event generators.

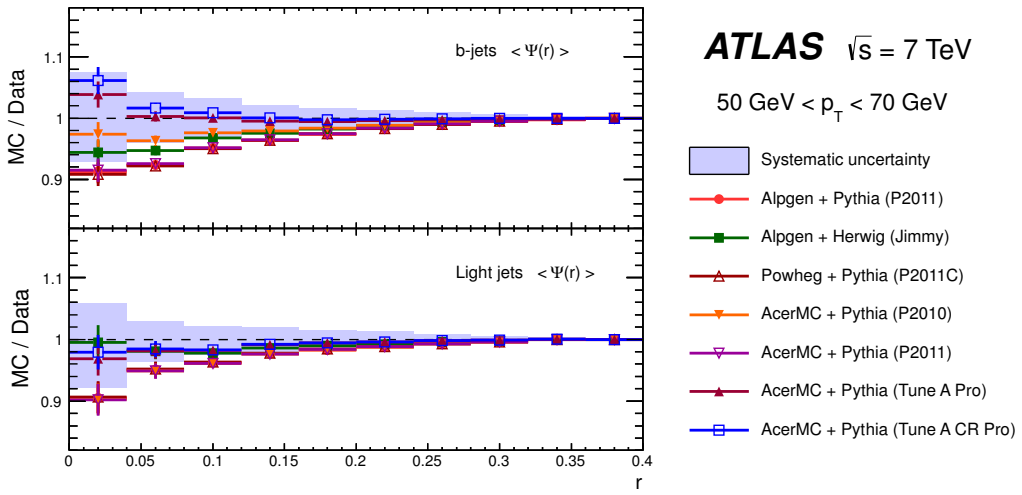

Figure 73: Comparison of the $t\bar{t}$ integrated jet shape data for 50 GeV $< p_{\mathrm{T}} <$ 70 GeV with several MC event generators.

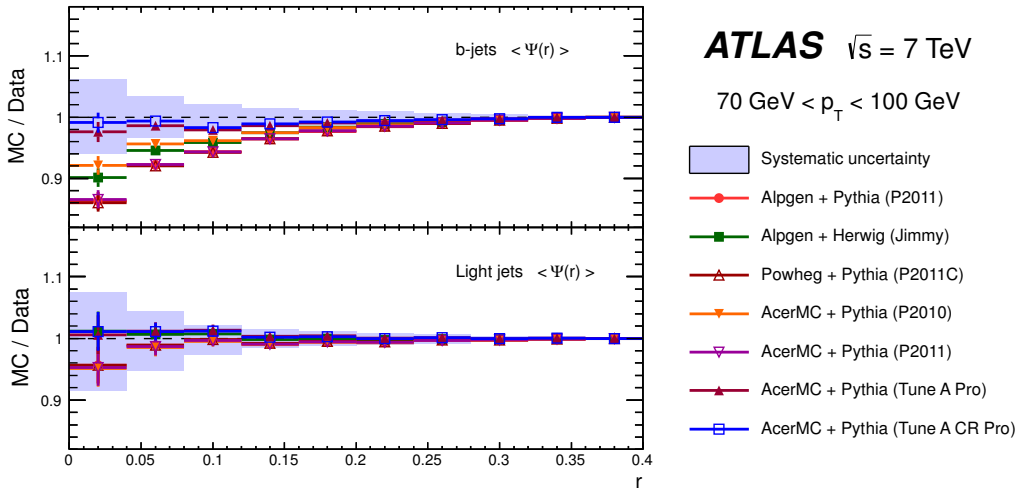

Figure 74: Comparison of the $t\bar{t}$ integrated jet shape data for 70 GeV $< p_{\mathrm{T}} <$ 100 GeV with several MC event generators.

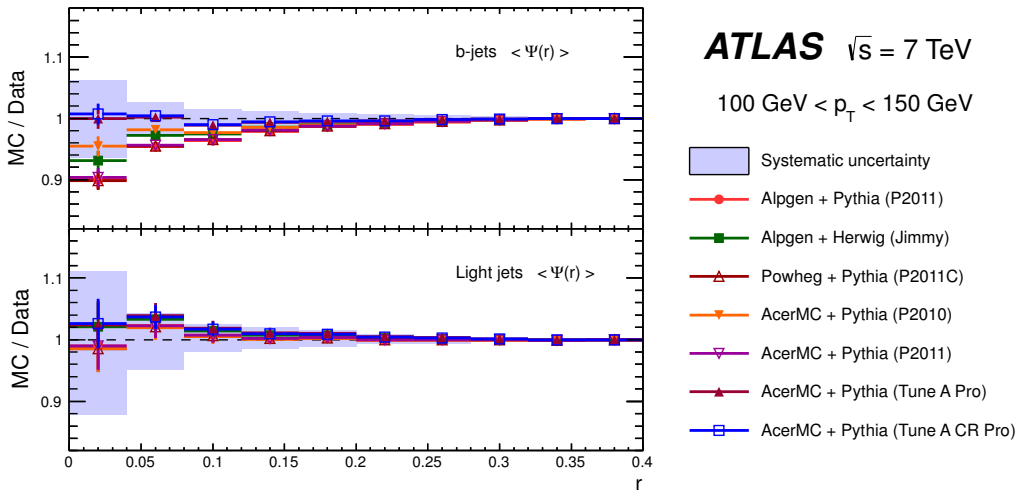

Figure 75: Comparison of the $t\bar{t}$ integrated jet shape data for 100 GeV $< p_{\mathrm{T}} <$ 150 GeV with several MC event generators.

B Independent sources of JES uncertainty for TEEC

This appendix is dedicated to the most important sources of the jet energy scale uncertainty described in Sect. 6.6.1. Figures 76 to 88 show these sources for both the TEEC and ATEEC functions.

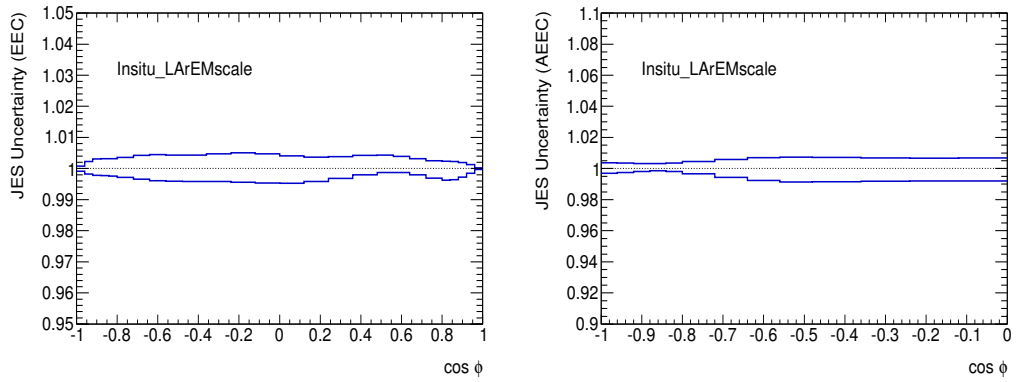

Figure 76: The impact of the INSITU_LAREMSCALE JES parameter in the TEEC (left) and ATEEC (right)

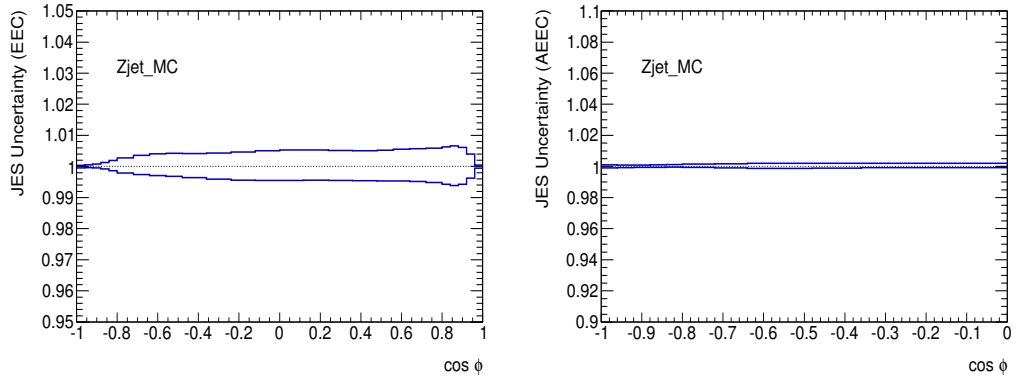

Figure 77: The impact of the ZJET_MC JES parameter in the TEEC (left) and ATEEC (right)

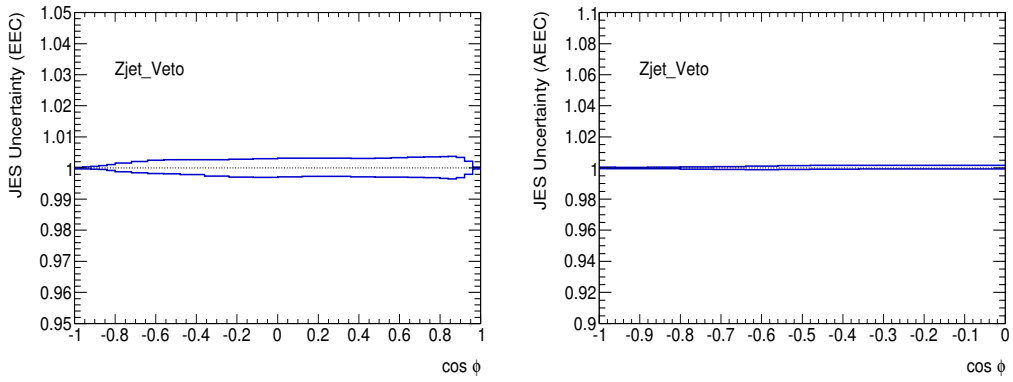

Figure 78: The impact of the ZJET_VETO JES parameter in the TEEC (left) and ATEEC (right)

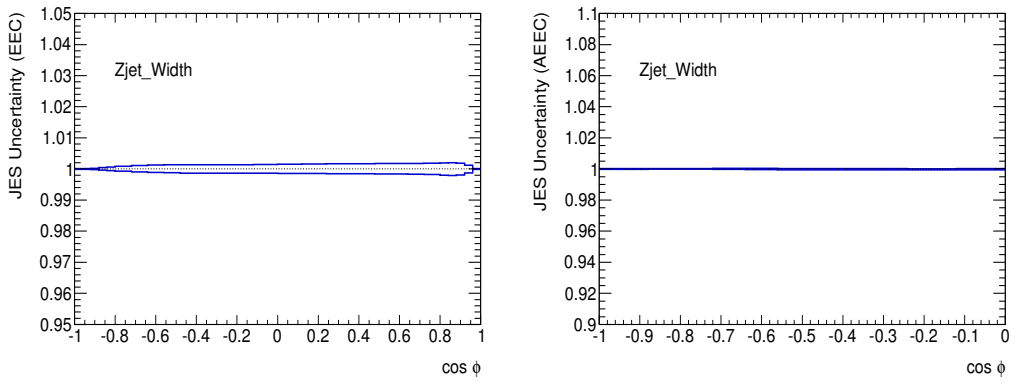

Figure 79: The impact of the ZJET_WIDTH JES parameter in the TEEC (left) and ATEEC (right)

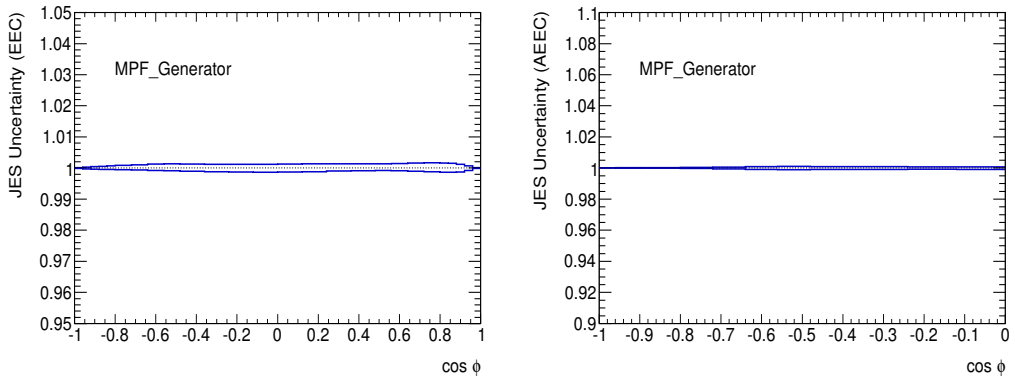

Figure 80: The impact of the MPF_Generator JES parameter in the TEEC (left) and ATEEC (right)

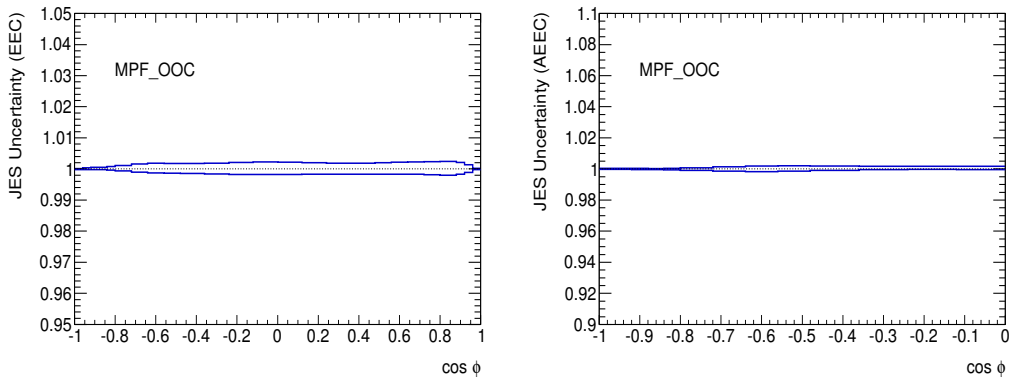

Figure 81: The impact of the MPF_OOC JES parameter in the TEEC (left) and ATEEC (right)

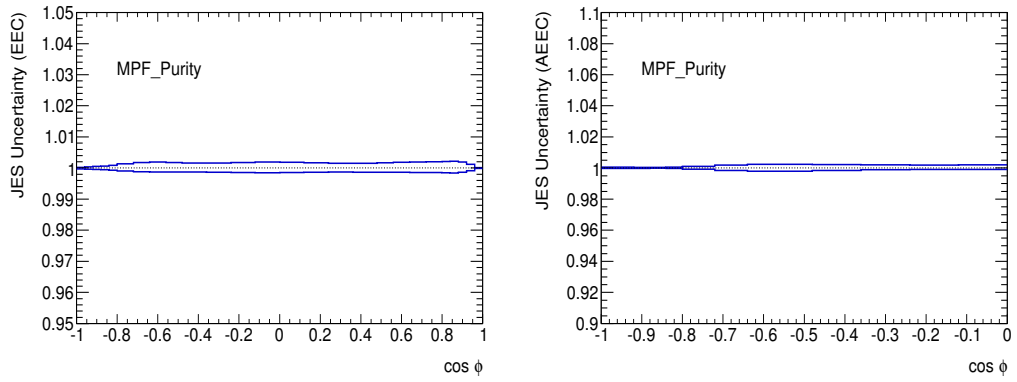

Figure 82: The impact of the MPF_Purity JES parameter in the TEEC (left) and ATEEC (right)

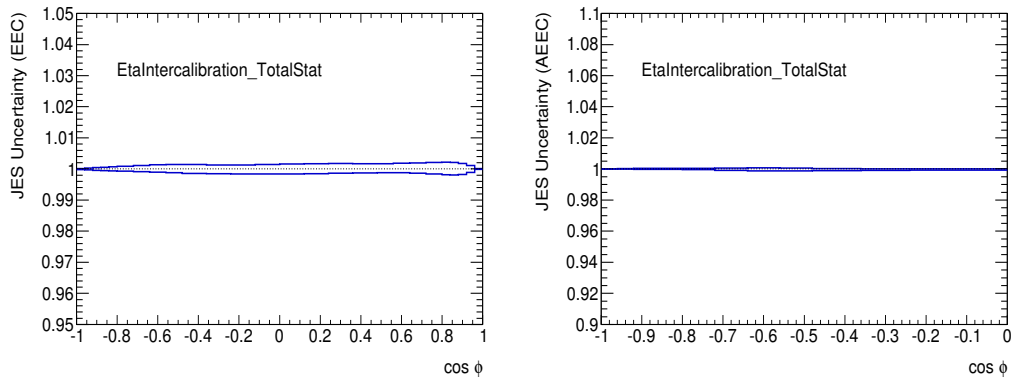

Figure 83: The impact of the EtaIntercalibration_TotalStat JES parameter in the TEEC (left) and ATEEC (right)

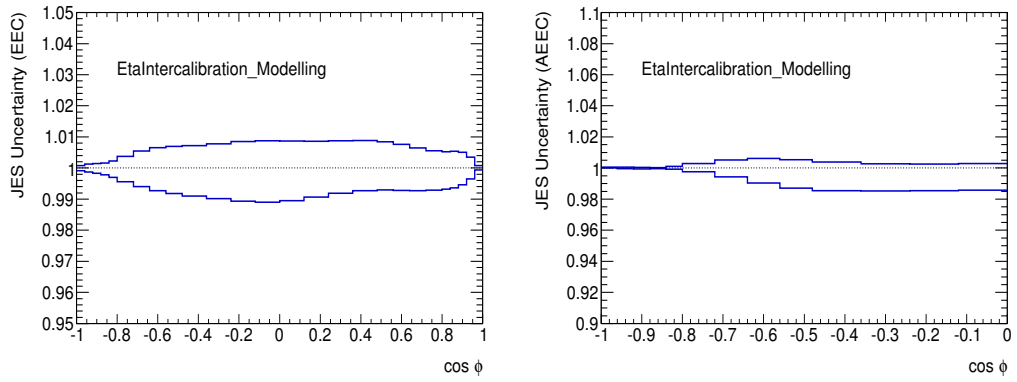

Figure 84: The impact of the EtaIntercalibration_Modelling JES parameter in the TEEC (left) and ATEEC (right)

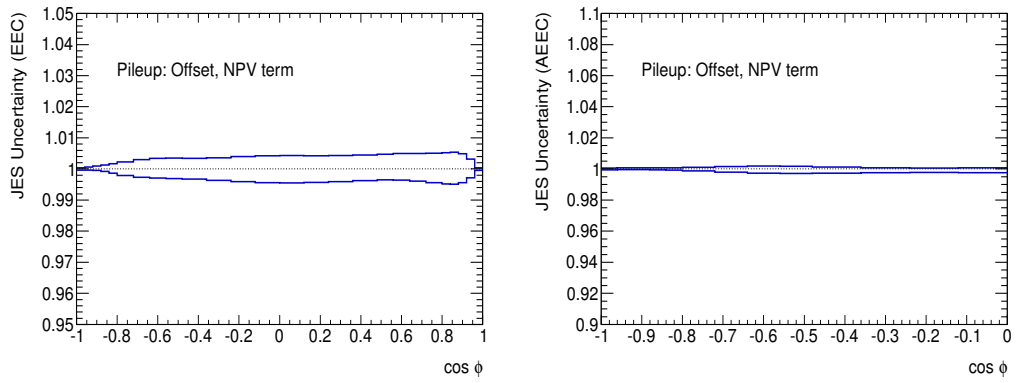

Figure 85: The impact of the Pileup: Offset, NPV term JES parameter in the TEEC (left) and ATEEC (right)

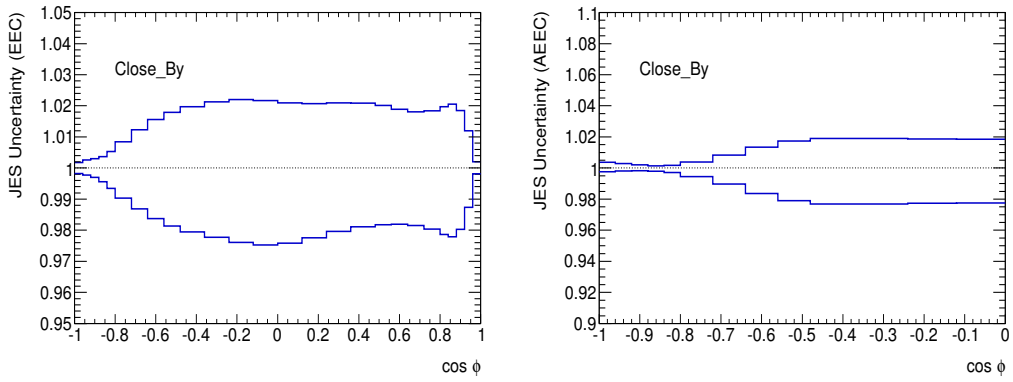

Figure 86: The impact of the Flavor_Comp JES parameter in the TEEC (left) and ATEEC (right)

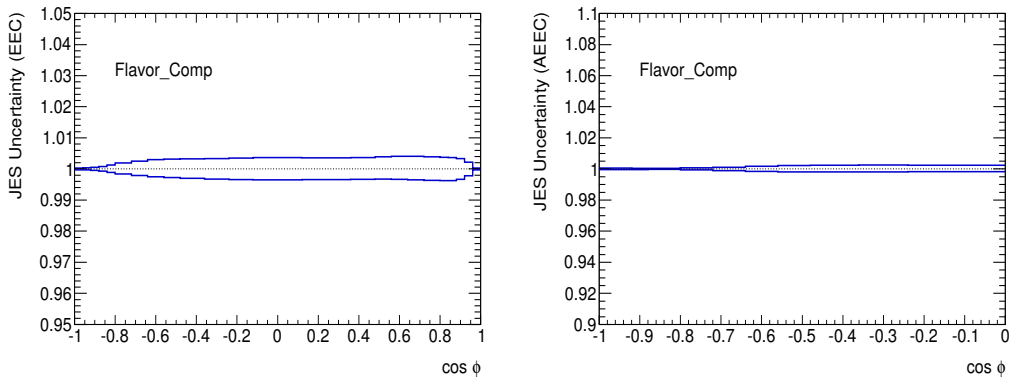

Figure 87: The impact of the Flavor_Response JES parameter in the TEEC (left) and ATEEC (right)

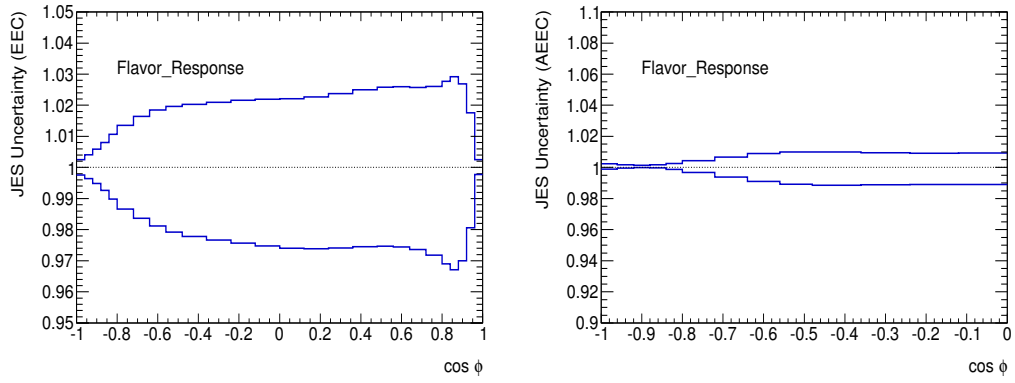

Figure 88: The impact of the CLOSE_BY JES parameter in the TEEC (left) and ATEEC (right)

References

[1] C. P. Burgess, Guy D. Moore. The Standard Model: A Primer (Cambridge University Press, 2007).

[2] F. Halzen, A. D. Martin. Quarks and Leptons: An Introductory Course in Modern Particle Physics (John Wiley and Sons, 1984).

[3] S. Weinberg. The quantum theory of fields (Cambridge University Press, 2005).

[4] R. K. Ellis, W. J. Stirling and B. R. Webber. QCD and Collider Physics (Cambridge University Press, 1996).

[5] G. M. Prosperi, M. Raciti and C. Simolo, [Prog. Part. Nucl. Phys. **58**, 387-438 (2007)].

[6] J. Beringer *et al.* (Particle Data Group), [Phys. Rev. D **86**, 010001 (2012)].

[7] S. Bethke, [arXiv:1210:0325 (hep-ex)].

[8] The ZEUS Collaboration, [Eur. Phys. J. C **70** 945 (2010)].

[9] R. P. Feynman. Photon Hadron Interactions. (W. A. Benjamin, 1971).

[10] F. D. Aaron *et al.*, JHEP **01** (2010) 109.

[11] G. Altarelli, G. Parisi [Nucl. Phys. B **126** 298 (1977)].

[12] J. C. Collins, D. E. Soper and G. Sterman [Nucl. Phys. B **261** 104 (1985)]

[13] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, [Eur. Phys. J. C **63** 189-285 (2009)].

[14] A. Schaelicke *et al.*, [Prog. Part. Nucl. Phys. **53**, 329-338 (2004)].

[15] V.V. Sudakov, [Sov. Phys. JETP **3**, 65 (1956)].

[16] Y. I. Azimov, Y. L. Dokshitzer, V. A. Khoze and S. I. Troian, [Z. Phys. C **27**, 65 (1985)].

[17] B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, [Phys. Rep. **97** 31 (1983)].

[18] R. D. Field, S. Wolfram, [Nucl. Phys. B **213** 65 (1983)].

[19] B. R. Webber, [Int. J. Mod. Phys. **A15S1**, 577-606 (2000)].

[20] The ATLAS Collaboration, [Phys. Rev. D **83**, 052003 (2011)].

[21] The ATLAS Collaboration, [Phys. Rev. D **83**, 112001 (2011)].

[22] T. Sjöstrand *et al.*, [Comput. Phys. Commun. **135**, 238 (2001)].

[23] P. Z. Skands, [Phys. Rev. D **82**, 074018 (2010)].

[24] M. Bahr *et al.*, [Eur. Phys. J. C **58**, 639 (2008)].

[25] J. Butterworth, J. Forshaw and M. Seymour, [Z. Phys. C **72**, 637 (1996)].

[26] S. Frixione *et al.*, [J. High Energy Phys. **01**, 053 (2011)].

[27] S. Frixione, P. Nason and C. Oleari, [J. High Energy Phys. **0711**, 070 (2007)].

[28] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A.D. Polosa, [J. High Energy Phys. **0307**, 001 (2003)]

[29] G. P. Salam and G. Soyez, [J. High Energy Phys. **0705**, 086 (2007)].

[30] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, [Nucl. Phys. B **406**, 187 (1993)].

[31] S. D. Ellis and D. E. Soper, [Phys. Rev. D **48**, 3160 (1993)].

[32] M. Cacciari, G. P. Salam and G. Soyez, [J. High Energy Phys. **0804**, 063 (2008)].

[33] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, [J. High Energy Phys. **9708**, 001 (1997)].

[34] M. Cacciari, G.P. Salam and G. Soyez, [Eur. Phys. J. C **72** 1896 (2012)].

[35] C. L. Basham, L. S. Brown, S. D. Ellis and S. T. Love, [Phys. Rev. Lett. **41**, 1585 (1978); Phys. Rev. D **19**, 2018 (1979)].

[36] A. Ali, F. Barreiro, J. Llorente and W. Wang, [Phys. Rev. D **86**, 114017 (2012)].

[37] Z. Nagy, [Phys. Rev. Lett. **88**, 122003 (2002); Phys. Rev. D **68**, 094002 (2003)].

[38] S.D. Ellis, Z. Kunszt and D. Soper, [Phys. Rev. Lett. **69**, 3615 (1992)].

[39] M. H. Seymour, [Nucl. Phys. B **513**, 269 (1998)].

[40] I. Vitev, S. Wicks and B-W. Zhang, [J. High Energy Phys. **0811**, 093 (2008)]

[41] G. Marchesini and B. R. Webber. Nucl. Phys. **B330** 261-283 (1990).

[42] Yu L. Dokshitzer, V. A. Khoze and S. I. Troyan. J. Phys. **G17** 1602 (1991).

[43] L. Evans and P. Bryant, [J. Instrum. **3** S08001 (2008)].

[44] The ATLAS Collaboration, [J. Instrum. **3** S08003 (2008)].

[45] The ATLAS Inner Detector Community, [ATLAS TDR 4, CERN/LHCC 97-16].

[46] A. Negri on behalf of the ATLAS Collaboration, [Nucl. Instrum. Meth. A **623**, 519 (2010)]

[47] R. Fruehwirth, [Nucl. Instrum. Meth. A **262**, 444 (1987)].

[48] W. Lampl *et al.*, [ATLAS-LARG-PUB-2008-002 (2008)].

[49] The ATLAS Collaboration, [Eur. Phys. J. C **72** 1909 (2012)].

[50] The ATLAS Collaboration, [CERN-OPEN-2008-020].

[51] S. Hassini *et al.*, [Nucl. Instrum. Meth. A **572**, 77 (2007)].

[52] Th. Lagouri *et al.*, [IEEE Trans. Nucl. Sci. **51**, 3030 (2004)].

[53] D. Adams *et al.*, [ATL-SOFT-2003-007 (2003)].

[54] S. Tarem *et al.*, [IEEE **1** 617 (2007)].

[55] The ATLAS Collaboration, [ATLAS-CONF-2010-053]

[56] The ATLAS Collaboration, [Eur. Phys. J. C **73** 2304 (2013)]

[57] The ATLAS Collaboration, [Eur. Phys. J. C **75** 17 (2015)]

[58] The ATLAS Collaboration, [Eur. Phys. J. C **73** 2305 (2013)].

[59] The ATLAS Collaboration, [ATLAS-CONF-2011-102].

[60] G. Piacquadio and C. Weiser, [J. Phys. Conf. Ser. **119** 032032 (2008)].

[61] The ATLAS Collaboration, [Eur.Phys.J. C **72** 1844 (2012)].

[62] The ATLAS Collaboration, [ATLAS-CONF-2010-064].

[63] The CDF Collaboration, [Phys. Rev. D **71** 112002 (2005)]

[64] The D0 Collaboration, [Phys. Lett. B **357** 500 (1995)]

[65] The OPAL Collaboration, [Z. Phys. C **63** 197 (1994)]

[66] The ZEUS Collaboration, [Eur. Phys. J. C **2** 61 (1998)] and [Eur. Phys. J. C **8** 367 (1999)]

[67] The ZEUS Collaboration, [Nucl. Phys. B, **700** 3 (2004)];

[68] The H1 Collaboration, [Nucl. Phys. B **545** 3 (1999)].

[69] G. Soyez *et al.*, [Phys. Rev. Lett. **110** 162001 (2013)]

[70] The CDF Collaboration, [Phys. Rev. D **78** 072005 (2008)]

[71] J.M. Butterworth *et al.* [Phys. Rev. Lett. **100** 242001 (2008)]

[72] D. Kaplan *et al.* [Phys. Rev. Lett. **101** 142001 (2008)];

[73] G. Salam. [Eur. Phys. J. C **67** 637 (2010)]

[74] A. Ali, F. Barreiro and T. Lagouri [Phys. Lett. B **693** 44 (2010)]

[75] S. Agostinelli *et al.* [Nucl. Instr. Meth. A **506** 250 (2003)]

[76] J. Pumplin *et al.* [J. High Energy Phys. **0207** 012 (2002)]

[77] A. Sherstnev and R.S. Thorne [Eur. Phys. J. C **55** (2008) 553-575]

[78] B. Kersevan and E. Richter-Was [Comput. Phys. Commun. **149** 142 (2003)]

[79] J. Pumplin *et al.* [J. High Energy Phys. **02** 032 (2006)]

[80] The ATLAS Collaboration [Phys. Lett. B **707** 459 (2012)]

[81] The ATLAS Collaboration [Phys. Lett. B **711** 244 (2012)]

[82] The ATLAS Collaboration [Eur. Phys. J. C **71** 1512 (2011)]

[83] The ATLAS Collaboration, [Eur. Phys. J. C **73** 3 2306 (2013)]

[84] The Atlas Collaboration [Eur. Phys. J. C **71** 1846 (2011)]

[85] The ATLAS Collaboration [ATLAS-CONF-2011-102]

[86] The ATLAS Collaboration [Phys. Lett. B **717** 330 (2012)]

[87] M. Aleksa *et al.* [ATL-LARG-PUB-2006-003]

[88] M. Aharrouche *et al.* [Eur. Phys. J. C **70** 755 (2010)] and [Eur. Phys. J. C **70** 1193 (2010)]

[89] G. Piacquadio [CERN-THESIS-2010-027]

[90] The ATLAS Collaboration [Eur. Phys. J. C **71** 1795 (2011)]

[91] G. D'Agostini [Nucl. Instrum. Meth. A **362** 487 (1995)]

[92] T. Adye [arXiv:1105.1160 (physics.data-an)]

[93] The ATLAS Collaboration, [Phys. Rev. Lett. **106** 172002 (2011)]

[94] The ATLAS Collaboration [Eur. Phys. J. C, **73** 3 2305 (2013)]

[95] The ATLAS Collaboration [CERN-PH-EP-2011-146]

[96] The ATLAS Collaboration, [Eur. Phys. J. C **73** 2676 (2013)].

[97] J. Llorente and J. Cantero, [Nucl. Phys. B **889** 401 (2014).]

[98] E. Norrbin and T. Sjöstrand, [Nucl. Phys. B **603** 297 (2001).]

[99] M. Bahr *et al.* [Eur. Phys. J. C **58** 639 (2008).]

[100] The ATLAS Collaboration. [Eur. Phys. J. C **72**, 2043 (2012).]

[101] P. Skands and D. Wicke, [Eur. Phys. J. C **52**, 133 (2007).]

[102] M. Fischler, D. Sachs. [arXiv:hep-ph/0306054.]

[103] J. M. Campbell, E. W. N. Glover and C. J. Maxwell, [Phys. Rev. Lett. **81** 1568 (1998).]

[104] The ATLAS Collaboration, [Phys. Rev. D **83** 112001 (2011)].

[105] R.D. Field, CDF Note 6403. [arXiv:hep-ph/0201192.]

[106] A. Perieanu, [DESY-THESIS-2006-002 (2006).]

[107] The DELPHI Collaboration, [Eur. Phys. J. C **55** 525 (2008).]

[108] The ALEPH Collaboration, [Eur. Phys. J. C **18**, 1 (2000).]

[109] The ATLAS Collaboration, [ATL-COM-PHYS-2013-884].

[110] S. Catani and M. H. Seymour, [Phys. Lett. B **378**, 287 (1996)].

[111] The CMS Collaboration, [arXiv:1204.0696 (hep-ex)].

[112] A. Banfi, G. P. Salam and G. Zanderighi, [J. High Energy Phys. **0408**, 062 (2004); J. High Energy Phys. **1006**, 038 (2010)].

[113] The CDF Collaboration [Phys. Rev. D **83**, 112007 (2011)].

[114] The ATLAS Collaboration [Eur. Phys. J. **C72**, 2211 (2012)].

[115] The CMS Collaboration [Phys. Lett. B **699**, 48 (2011)].

[116] A. Ali and F. Barreiro, [Phys. Lett. B **118**, 155 (1982); Nucl. Phys. B **236**, 269 (1984)].

[117] D. G. Richards, W. J. Stirling and S. D. Ellis, [Phys. Lett. B **119**, 193 (1982)].

[118] The PLUTO Collaboration, Ch. Berger *et al.*, [Phys. Lett. B **99** (1981) 292].

[119] The MARKII Collaboration, D. Schlatter *et al.*, [Phys. Rev. Lett. C **49** (1982) 521].

[120] The MARKJ Collaboration, B. Adeva *et al.*, [Phys. Rev. Lett. C **50** (1983) 2051].

[121] The CELLO Collaboration, H. J. Behrend *et al.*, [Phys. Lett. B **138** (1984) 311].

[122] The JADE Collaboration, W. Bartel *et al.*, [Zeit. Phys. C **25** (1984) 231].

[123] The MAC Collaboration, E. Fernández *et al.*, [Phys. Rev. D **31** (1985) 2724].

[124] The TASSO Collaboration, W. Braunschweig *et al.*, [Zeit. Phys. C **36** (1987) 349].

[125] The TOPAZ Collaboration, I. Adachi *et al.*, [Phys. Lett. B **227** (1989) 495].

[126] The DELPHI Collaboration, P. Abreu *et al.*, [Phys. Lett. B **252** (1990) 149].

[127] The OPAL Collaboration, M. Z. Akrawy *et al.*, [Phys. Lett. B **252** (1990) 159].

[128] The ALEPH Collaboration, D. Decamp *et al.*, [Phys. Lett. B **257** (1991) 479].

[129] The L3 Collaboration, B. Adeva *et al.*, [Phys. Lett. B **257** (1991) 469].

[130] The SLD Collaboration, K. Abe *et al.*, [Phys. Rev. D **51** (1995) 962].

[131] A. Ali and G. Kramer, [Eur. Phys. J. **H36**, 245 (2011)]

[132] A. Ali, E. Pietarinen and W. J. Stirling, [Phys. Lett. B **141**, 447 (1984)].

[133] Z. Bern *et al.*, [Phys. Rev. Lett. **109**, 042001 (2012)]

[134] H. L. Lai *et al.*, [Phys. Rev. D **82** 074024 (2010)].

[135] R. D. Ball *et al.*, [Nucl. Phys. B **867** (2013) 244].

[136] The CMS Collaboration, [Eur. Phys. J. **C73** (2013) 2604]

[137] B. Malaescu and P. Starovoitov [Eur. Phys. J. **C72** (2012) 2041]

[138] The ATLAS Collaboration [ATL-PHYS-PUB-2011-009]

[139] The ATLAS Collaboration [ATL-PHYS-PUB-2011-014]

[140] The ATLAS Collaboration [J. High Energy Phys. **05** 059 (2014)]

[141] The ATLAS Collaboration [Eur. Phys. J. **C71**, 1763 (2011)]

[142] The ATLAS Collaboration [ATL-COM-DAQ-2011-031]

[143] C. Issever, K. Borras and D. Wegener, [Nucl. Instrum. Methods A **545**, 803 (2005)]

[144] N. Kidonakis, [Phys. Rev. D **82**, 114030 (2010)].