Jet calibration, cross section measurements and New Physics searches with the ATLAS experiment within the Run 2 data

Robert Hankache

Sorbonne Université
Step'UP doctoral school
Laboratoire de Physique Nucléaire et de Hautes Energies

PhD thesis in PHYSICS

Jet calibration, cross-section measurements and New Physics searches with the ATLAS experiment within the Run 2 data

by: **Robert HANKACHE**

Jury:

Pr.	Paolo	Azzurri,	*referee*
Pr.	Emanuela	Barberis,	*referee*
Pr.	Tancredi	Carli,	*president of the jury*
Pr.	Matthew	Charles,	*examiner*
Dr.	Bogdan	Malaescu,	*thesis co-director*
Pr.	Mélissa	Ridel,	*thesis director*

Presented and defended September 24^{th}, 2019

Acknowledgment

"Gratitude is when memory is stored in the heart and not in the mind."

The success in my academic path could not be achieved without the different teachers and supervisors I encountered during all my path.
Mainly, I wish to thank my thesis director and co-director, Mélissa Ridel and Bogdan Malaescu, who gave me all their attention and knowledge, always pushed me to give my best and provided me the solid ground to improve, succeed and finish my PhD.
I also want to thank my Masters 2 responsible, Delphine Hardin et Sophie Trincaz-Duvoid, who believed in me and through them I was able to come to France.

Beyond the academic support, my success was not possible without the moral and emotional supports. I thank my family for supporting me from the moment I decided to leave for France to continue my education and always being there for me.
My special gratitude is for my fiancée, *Rana Bsaibes*, who, although being more than 3000 km apart, never lost faith in our relation, gave me the needed support to flourish and sacrificed a lot for us. I owe her my love and our beautiful relation.
The final gratitude is for *God Father* who, through *Jesus Christ*, in the unity of the *Holy Spirit*, showed us his endless love and mercy.

Contents

Introduction

The Standard Model of particle physics is very successful in describing the properties of elementary particles and the electro-magnetism, weak and strong fundamental interactions. Its success is due to the successful description of many observables that are measured precisely, but also to its predictions of particles and properties that were verified later by experiments. Nevertheless, the Standard Model cannot be the final theory of nature. For example, it still does not include the gravitational interaction, explain the dark matter ... Many theories beyond the Standard Model exist, but their validity is still to be verified. Hence, the need to keep pushing further in particle physics experiments, whether collecting more data to be sensitive to small signals or reaching higher energies to explore new territories.

The ATLAS experiment, to which I contribute, is one of the biggest physics collaborations that comprises about 3000 scientific authors from 183 institutions around the world, representing 38 countries from all the world's populated continents. Its detector is the largest volume detector ever constructed for a particle collider; it was built and is upgraded using the latest technological advances in detector systems. The detector is built around an interaction point of the large hadron collider (LHC). The Run I data taking period happened between the years 2010 and 2012 with collision energies up to 8 TeV. The Run II data taking period happened between the years 2015 and 2018 with collision energies up to 13 TeV, the highest energies reached by a collider to date. In my thesis, I use the proton-proton collisions data collected during Run II at 13 TeV energies.

This manuscript aims to summarize the most important parts of my contributions within the ATLAS collaboration. It is organized as follows. First, chapter 1 introduces the Standard Model, some Beyond Standard Model theories and the different approaches and techniques used to calculate the predictions of those models. The LHC and the ATLAS detector are introduced in chapter 2. Jet objects reconstruction and calibration are presented in details, since they are the main objects used in this thesis. The next three chapters summarize my contributions in performance and physics analyses. My first contribution, which is also my qualification task to become an ATLAS author, is detailed in chapter 3 and corresponds to the study of one method of jet calibration, namely the eta-intercalibration. Chapter 4 details the direct search for new physics using the invariant mass of two jets with an emphasis on the folding technique that I implemented. Chapter 5 details a new jet cross-section measurement that is done for the first time, the leading jet double differential cross-section, with details on both the experimental measurement and the theoretical predictions aspects, where I contributed in both. Finally, the conclusion is presented.

1 Standard Model and Beyond, with a focus on QCD and its predictions

In this chapter[1], I describe the Standard Model of particle physics (SM) in the first section, then, I detail more the Quantum Chromo-Dynamics (QCD) theory related directly to my thesis. Next, I present different methods used to calculate theoretical predictions. Finally, I introduce some Beyond Standard Model (BSM) theories used in the search for new physics.

1.1 The Standard Model (SM)

The Standard Model of particle physics (SM) describes the properties of elementary particles and three of their four fundamental interactions, electro-magnetism, the weak and the strong interactions in a consistent relativistic quantum field theory. By elementary particle, we mean a quantum particle with no evidence of internal structure. We distinguish two groups of elementary particles: fermions, the building blocks of the visible matter[2], and bosons, the force carrying particles. Anti-matter is also described by the SM which has the same properties as matter but with opposite quantum numbers (see section 1.1.1). Figure 1.1 shows an info-graphic of the SM particles, fermions and bosons, with their masses, spins, electric and color quantum numbers. The contents in this info-graphic are developed in the belows sections.

The Standard Model is one of the most successful physical models. In fact, it provides calculations with high precision of many independent observables, the most precise one being the electromagnetic fine structure constant known to more than 10 orders of precision. The model made also a lot of predictions which have been verified experimentally. Moreover, the model is minimal: it is constructed using only fields, interactions, and parameters which are necessary for consistency and/or observed experimentally.

The SM is formulated in the Lagrangian formalism, in terms of a Lorentz-invariant local function $\mathcal{L}$, the Lagrangian density. The fermions obey the Pauli exclusion principle, follow Fermi-Dirac statistics and can be characterized by a real positive quantity m and a half-integer number s which can be identified with the mass and spin of the associated particles, respectively. The corresponding

[1] The materials in this chapter are based on my "NPAC Masters 2" lecture notes (namely the "QFT" course by Matteo Cacciari and Sebastien Descotes Genon, the "particles physics" course by Mélissa Ridel and Patrick Robbe), my CERN 2016 summer school lecture notes (namely the "QFT and SM" course by Andrej Arbuzov, the "QCD" course by Kirill Melnikov), the lecture notes of the "QCD" course by Gavin Salam at the CERN 2009 summer school, Alexander Huss thesis dissertation, the "Review of Particle Physics" 2018 book by the PDG collaboration [1], the "An introduction to quantum field theory" book by Peskin and Schroeder [2], and the various references cited later.

[2] Visible matter constitute less than 5% of the total energy of our universe.

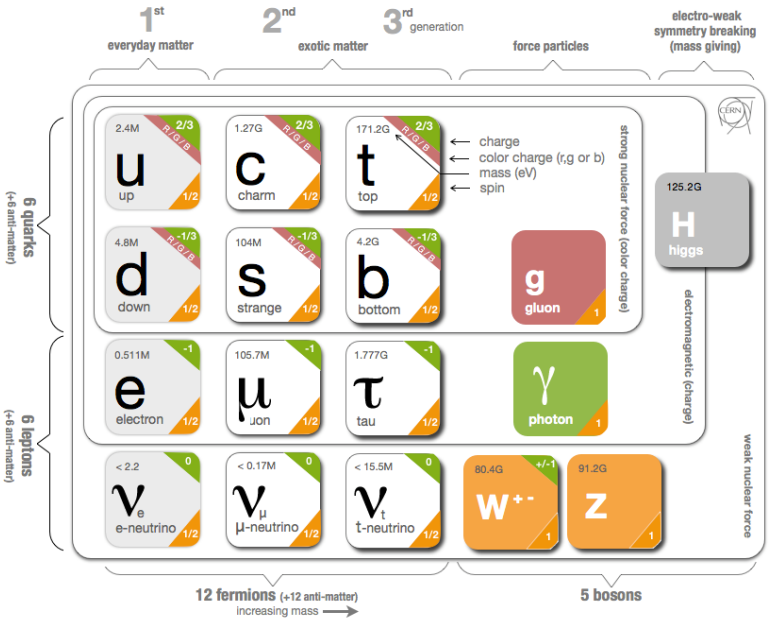


Figure 1.1: Info-graphic of the SM particles, fermions and bosons, showing their masses, spins, electric and color quantum numbers. Figure from CERN website.

force fields obey Bose-Einstein statistics and the equations of motion given by the Euler-Lagrange equations, which determine the dynamics of the theory. The SM is formulated as a spontaneously broken non-Abelian gauge[3] theory with the underlying gauge group

$$\mathrm{SU(3)_C \times SU(2)_W \times U(1)_Y}. \tag{1.1}$$

The electroweak (EW) sector is described by the Glashow–Salam–Weinberg [3–5] model of electroweak interactions with the associated $\mathrm{SU(2)_W \times U(1)_Y}$ gauge group. The strong interaction is described by quantum chromodynamics [6–9] (QCD) with the associated gauge group $\mathrm{SU(3)_C}$.

1.1.1 Lagrangian formulation

The SM Lagrangian can be divided into the following parts:

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{Yang-Mills}} + \mathcal{L}_{\mathrm{Fermion}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}}. \tag{1.2}$$

Let us examine successively each of these parts.

Gauge theories and the Yang-Mills Lagrangian

The SM is constructed on the principle of gauge (local) symmetries. Let us considering the Lagrangian

[3] The terms "spontaneously broken" and "gauge" are explained in section 1.1.1.

of a field Φ which is invariant under the global symmetry of the SM group:

$$\Phi(x) \rightarrow \Phi'(x) = U(\theta_C, \theta_W, \theta_Y)\Phi(x), \tag{1.3}$$

$$U(\theta_C, \theta_W, \theta_Y) = \exp[-ig_s\theta_C^a t_C^a + ig\theta_W^i t_W^i - ig'\theta_Y t_Y], \tag{1.4}$$

where the different t denote the generators of the Lie algebra of the different SM groups θ, and the different g parameters denote the corresponding coupling constants.

For a local transformation where the group parameters depend on the space-time coordinate x, $\theta \rightarrow \theta(x)$, the covariant transformation property of the derivation of the filed, $\partial_\mu\Phi$, is destroyed. The gauge-covariant derivative D_μ is a generalization of the ordinary derivative $\partial_\mu\Phi$ which replaces the latter in the Lagrangian to preserve its invariance under gauge transformations. It is defined as:

$$D_\mu = \partial_\mu + ig_s G_\mu^a t_C^a + igW_\mu^i t_W^i + ig'B_\mu t_Y, \tag{1.5}$$

where the added massless gauge fields, G_μ^a ($\times 8$), W_μ^i ($\times 3$) and B_μ ($\times 1$) ensure the covariant transformation of D_μ. This replacement is called the method of minimal substitution.

The requirement of gauge invariance automatically induces interaction terms, where the gauge fields act as the mediators.

The dynamics for the gauge fields are incorporate into the theory by adding a gauge-invariant kinetic term, given by the Yang–Mills Lagrangian:

$$\mathcal{L}_{\text{Yang–Mills}} = -\frac{1}{4}G_{\mu\nu}^a G^{a,\mu\nu} - \frac{1}{4}W_{\mu\nu}^i W^{i,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.6}$$

where the field strength tensors are defined as

$$\begin{aligned} G_{\mu\nu}^a &= \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c \\ W_{\mu\nu}^i &= \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon^{ijk} W_\mu^j W_\nu^k \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \end{aligned}$$

where f^{abc} and ϵ^{ijk} are the structure constants of the SU(3) and SU(2) algebras, respectively. The bilinear terms in the Lagrangian $\mathcal{L}_{\text{Yang–Mills}}$ describe the free motion of the fields, the cubic and quartic terms in the gauge fields describes their interaction among themselves. Furthermore, any explicit mass term for the gauge fields is forbidden, as it would break gauge invariance.

For QCD, the number of colors $N_c = 3$. $N_c^2 - 1 = 8$ gauge fields exist, corresponding to the 8 kinds of gluons. $t_A^{ab}t_A^{bc} = C_F\delta_{ac}$ with $C_F = 4/3$ is the color-factor associated with gluons emission from a quark. $f^{acd}f^{bcd} = C_A\delta_{ab}$ where $C_A = N_c = 3$ is the color-factor associated with gluon emission from a gluon. $t_A^{ab}t_B^{ab} = T_R\delta_{AB}$ where $T_R = 1/2$ is the color-factor for a gluon to split to a $q\bar{q}$ pair.

Fermion Lagrangian

The fermions, spin 1/2 matter content of the SM, can be classified into the leptons (l,ν_l), which are only subject to the electromagnetic and weak interactions, and the quarks (q_u , q_d), which also interact strongly (they come in 3 color flavors). The fermions come in three generations, which have

Table 1.1: The fermionic matter content of the SM with the 3 generations shown. The third component of the weak isospin, I_W^3, and the hypercharge, Y, quantum numbers are shown. For the quarks, three color quantum numbers also exist which are not shown here. Subscript letters R and L denotes right and left chiralities. Superscript letters L and Q denotes the lepton and quark doublets.

		1^{st}	2^{nd}	3^{rd}	I_W^3	Y
leptons	$\Phi_L^{L'}$	$\begin{pmatrix}\nu_e'\\ e'\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu'\\ \mu'\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau'\\ \tau'\end{pmatrix}_L$	$\begin{pmatrix}1/2\\ -1/2\end{pmatrix}$	-1
	$\Phi_R^{l'}$	e_R'	μ_R'	τ_R'	0	-2
quarks	$\Phi_L^{Q'}$	$\begin{pmatrix}u'\\ d'\end{pmatrix}_L$	$\begin{pmatrix}c'\\ s'\end{pmatrix}_L$	$\begin{pmatrix}t'\\ b'\end{pmatrix}_L$	$\begin{pmatrix}1/2\\ -1/2\end{pmatrix}$	1/3
	$\Phi_R^{u'}$	u_R'	c_R'	t_R'	0	4/3
	$\Phi_R^{d'}$	d_R'	s_R'	b_R'	0	-2/3

identical properties under the gauge interactions, but differ in their mass and flavor. Fermions are also assigned left-handed and right-handed chirality, a representation that comes from the nature of the weak interaction. For massless particles, chirality is equivalent to the helicity, given by the projection of its spin in the direction of its momentum. Leptons come in left-handed doublets and charged right-handed singlets. Quarks come in left-handed doublets, down right-handed singlets and up right-handed singlets. As a consequence, any explicit mass terms for the fermions are forbidden, as they mix the two chiralities and, as such, would break gauge invariance. Table 1.1 lists the singlets and doublets fermions for the three generations along their field notation. The prime notation denotes that the fields are eigenstates with respect to the gauge interaction. Subscript letters R and L denotes right and left chiralities. Superscript letters L and Q denotes the lepton and quark doublets.

The Lagrangian for the fermionic fields is obtained by applying the method of minimal substitution to the free Dirac theory of massless fermions, leading to:

$$\mathcal{L}_{\text{Fermion}} = \sum_{i=1}^{3} \left(\bar{\Phi}_L^{L_i'} \gamma^\mu i D_\mu \Phi_L^{L_i'} + \bar{\Phi}_R^{l_i'} \gamma^\mu i D_\mu \Phi_R^{l_i'} + \bar{\Phi}_L^{Q_i'} \gamma^\mu i D_\mu \Phi_L^{Q_i'} + \bar{\Phi}_R^{u_i'} \gamma^\mu i D_\mu \Phi_R^{u_i'} + \bar{\Phi}_R^{d_i'} \gamma^\mu i D_\mu \Phi_R^{d_i'} \right) \tag{1.7}$$

The gauge-covariant derivative D_μ, which includes gauge fields terms, gives rise to the interaction terms between fermionic and gauge fields. The fields indicated here are eigenstates with respect to the gauge interactions (as opposed to mass eigenstates).

Quarks transform under SU(3) where the generators are given by $t_C^a = \frac{\lambda^a}{2}$ (λ^a denotes the Gell-Mann matrices). Left-handed fermion doublets transform under SU(2) where the generators are given by $t_W^i = I_W^i = \frac{\sigma^i}{2}$ (σ^i denotes the Pauli matrices), where I_W^i denotes the isospin. Charged fermions transform under U(1) where the generator is given by $t_Y = \frac{Y}{2}$, where Y denotes the hypercharge. The electric charge Q of the particle is given by:

$$Q = I_W^3 + \frac{Y}{2}. \tag{1.8}$$

Higgs mechanism

Any explicit mass terms for the gauge bosons are forbidden, as they would break gauge invariance.

However, they can be accommodated in a gauge theory by the mechanism of spontaneous symmetry breaking (SSB) [10–12]. This is accomplished by allowing a field to develop a non-vanishing vacuum expectation value (vev), which has to be a scalar in order to preserve Lorentz invariance. A scalar weak isospin doublet is introduced with a hypercharge Y_ϕ and parametrized as follows:

$$\phi(x) = \begin{pmatrix} \varphi^+(x) \\ \varphi^0(x) \end{pmatrix}. \tag{1.9}$$

The most general Lagrangian of ϕ which satisfies the restrictions imposed by gauge invariance and the stability of the vacuum is:

$$\begin{aligned} &\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \\ &V(\phi) \equiv \text{Higgs potential} = -\mu^2(\phi^\dagger \phi) + \frac{\lambda}{4}(\phi^\dagger \phi)^2, \quad \lambda > 0. \end{aligned} \tag{1.10}$$

For $\mu^2 > 0$, the field acquires a non-vanishing vev ϕ_0 with:

$$|\phi_0|^2 = \frac{2\mu^2}{\lambda} \equiv \frac{v^2}{2}. \tag{1.11}$$

The potential V have a Mexican hat shape which is illustrated in figure 1.2. In order to preserve the electromagnetic gauge symmetry, only the neutral component of φ^0 can develop the vev. The new parametrization is:

$$\phi(x) = \begin{pmatrix} \varphi^+(x) \\ \frac{1}{\sqrt{2}}(v + H(x)) \end{pmatrix}. \tag{1.12}$$

The real field H(x) corresponds to the massive mode associated with the physical Higgs boson, discovered in 2012 [13, 14], a neutral spin-0 particle with mass

$$M_H = \sqrt{2}\mu. \tag{1.13}$$

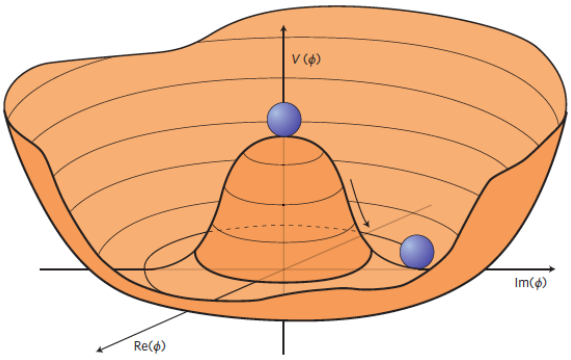


Figure 1.2: An illustration of the Higgs potential with $\mu^2 > 0$, in which case the minimum (vev) is at $|\phi_0|^2 = 2\mu^2/\lambda$. The shift of the minimum from the center is the origin of the SSB.

Mass terms for the EW gauge fields are generated through the SSB and originates from the vev part of the Higgs potential. This mechanism is known as the Higgs mechanism. The Higgs mechanism induces mixing terms among the EW gauge fields. We will redefine the fields to make appear the physical ones. The mass matrix of the EW gauge bosons is diagonalized as follows:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W^3_\mu \end{pmatrix} \tag{1.14}$$

with the weak mixing angle θ_W given by:

$$\cos\theta_W \equiv c_W = \frac{g}{\sqrt{g^2 + g'^2}}. \tag{1.15}$$

The massless gauge field A_μ corresponds to the photon and the massive field Z_μ corresponds to the electrically neutral Z boson. The elementary charged is:

$$e = \frac{gg'}{\sqrt{g^2 + g'^2}}, \tag{1.16}$$

and the EW coupling (fine-structure) constant is:

$$\alpha = \frac{e^2}{4\pi}. \tag{1.17}$$

The two massive charged bosons W^+ and W^- are described by the fields:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp W^2_\mu). \tag{1.18}$$

The masses of the bosons are:

$$\begin{aligned} M_W &= \frac{vg}{2} \\ M_Z &= \frac{v\sqrt{g^2 + g'^2}}{2} \end{aligned} \tag{1.19}$$

giving

$$\cos\theta_W = \frac{M_W}{M_Z}.$$

The strong gauge field is not affected by the Higgs mechanism. The strong boson, gluon, stays therefore massless. The strong coupling constant is:

$$\alpha_S = \frac{g_S^2}{4\pi}.$$

Yukawa couplings

Explicit fermion mass term is forbidden in the SM due to the distinct transformation properties of the two fermion chiralities. The solution to fermion mass terms is also provided by the Higgs field. Defining the charge conjugate as

$$\phi^c = i\sigma^2\phi^* = ((\varphi^0)^*, -\varphi^-)^T,$$

we construct the general gauge-invariant interaction between the fermions and the Higgs filed:

$$\mathcal{L}_{\text{Yukawa}} = -\sum_{i,j=1}^{3}\left(\bar{\Phi}_L^{L'_i}\Lambda_{ij}^{l}\Phi_R^{l'_j}\phi + \bar{\Phi}_L^{Q'_i}\Lambda_{ij}^{u}\Phi_R^{u'_j}\phi^c + \bar{\Phi}_L^{Q'_i}\Lambda_{ij}^{d}\Phi_R^{d'_j}\phi + \text{h.c.}\right), \tag{1.20}$$

where "i" and "j" represent the three fermion generations and "h.c." stands for hermitian conjugate of the terms preceding it. The Yukawa couplings Λ_{ij}^{f} are complex 3×3 matrices in generation space. The vev part of the Higgs field generates mass terms of the fermions that can be diagonalized by a bi-unitary transformation:

$$m_{f,i} = \frac{v}{\sqrt{2}}\sum_{k,l=1}^{3} U_{L,ik}^{f}\Lambda_{kl}^{f}U_{R,li}^{f\dagger}\,, \text{ where } f = l, u, d. \tag{1.21}$$

The transition of the eigenstates of the gauge interactions, $\Phi^{f'}$, to the mass eigenstates, Φ^{f}, is given by:

$$\Phi_L^{f,i} = \sum_{j=1}^{3} U_{L,ij}^{f}\Phi_L^{f',i} \text{ and } \Phi_R^{f,i} = \sum_{j=1}^{3} U_{R,ij}^{f}\Phi_R^{f',i}. \tag{1.22}$$

In the quark sector, charged-current interaction terms like $\bar{\Phi}_L^{u'_i}\gamma^\mu W_\mu \Phi_L^{d'_i}$ in gauge basis become $\bar{\Phi}_L^{u_i}\gamma^\mu W_\mu (U_L^{u\dagger}U_L^{d})\Phi_L^{d_j}$ in mass basis, giving interactions between mass eigenstates from different families. The quark mixing matrix, parameterizing the transition between different generations of the mass eigenstates, is given by the Cabibbo-Kobayashi-Maskawa [15, 16] (CKM) matrix defined as follows:

$$V_{CKM} = U_L^{u\dagger}U_L^{d} = U_L^{u}U_L^{d\dagger} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.23}$$

It appears only in the charged-current interactions and can be parametrized in terms of three angles and a complex phase, where the latter is the only source of CP [4] violation in the SM. The CP violation phenomenon could explain why the universe is dominate by matter over anti-matter.

In the original formulation of the SM, all neutrinos are assumed to be massless. This is not true anymore after the observation of neutrino oscillations. Neutrino masses can be added in a way similar to the other fermions or through other models like seesaw mechanism.

[4] In a CP transformation, particles are switched with their anti-particles (charge conjugate) and the physical system is mirrored, i.e. sign flip of the space coordinates (parity).

Independent parameters
The SM Lagrangian depends on a set of 18 independent parameters (if considering massless neutrinos) which can be chosen as:

- 2 couplings: g_S and e,
- 3 boson masses: M_W, M_Z and M_H,
- 9 fermion masses: 3 m_l and 6 m_q,
- 4 parameters of the CKM matrix.

1.1.2 Perturbative approach

If the interactions are not too strong, the SM prediction can be evaluated perturbatively about the free field theory. The free theory is described by the bilinear terms of the Lagrangian and the solutions of the corresponding field equations are given by the one-particle wave functions of the respective field. In the so-called weak asymptotic limit the interactions are assumed to be negligible long before and after the scattering reaction ($t \to \pm\infty$), where the particles are considered to be far apart. As a consequence, the particles entering and exiting the reaction can be described by the free one-particle states. The S-matrix, calculated from the Lagrangian, transforms the incoming configuration for $t \to -\infty$ into the outgoing configuration for $t \to +\infty$ and encapsulates the information of the scattering reaction. The cross section is then calculated by:

$$\sigma_p = \langle f|S|i\rangle^2,$$

where i and f denotes the initial and final states respectively.

The interaction is pictorially represented in terms of so-called Feynman diagrams, which are assembled from two types of building blocks: the propagators and the vertices. The former are represented by lines and are derived from the free, bilinear part of the Lagrangian, and the latter emerge from the remaining interaction terms and are represented by vertices that join multiple lines. The associated analytic expressions can be directly derived from the Lagrangian 1.2 and gives rise to the Feynman rules. Lastly, the external legs translate into the one-particle wave functions. Examples of Feynman diagrams describing the interaction between an electron and a positron by the intermediate of a photon are shown in the figure 1.3. On the left is shown a diagram with the lowest order possible for the interaction with no loops, called the tree level. On the right is shown a one loop diagram. When performing calculations, all the possible diagrams should be calculated, and interferences should be taken into account when the same initial and final states exist for multiple diagrams.

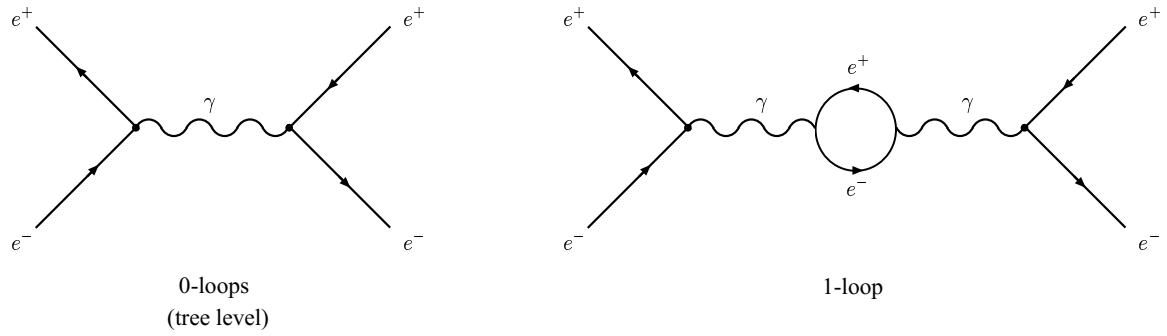


Figure 1.3: Examples of Feynman diagrams for the interaction between an electron and a positron by the intermediate of a photon. On the left is shown a diagram with the lowest order possible for the interaction with no loops, called the tree level. On the right is shown a 1 loop diagram.

1.1.3 Renormalization and running coupling

As we have seen, the SM Lagrangian depends on a set of independent parameters which are identified with the couplings and the masses of the particles. In order to make theoretical predictions, it is necessary to relate these parameters to independent measurable quantities. The step of re-parameterizing the theory in terms of physically measured quantities is known as renormalization. In fact, in the calculation of higher-order corrections, divergences arise which forbid any direct physical interpretation of the (bare) parameters appearing the Lagrangian. In a renormalizable theory, all divergences drop out in relations that express physical observables in terms of measurable (renormalized) quantities. The renormalization moves the divergences into the bare parameters and introduces an (unphysical) renormalization scale μ_R at which the regularization is performed and the physical observable is calculated. When summing over all perturbative orders, the renormalization requires the prediction to be independent on the choice of the scale μ_R. The renormalizability of a theory is therefore a crucial property in order to obtain predictions using perturbation theory. Exceptions are made for non-renormalizable theories in their phase-space region where the predictions are still reliable.

The self-interacting nature of the fields influences the free part of the theory by a shift of the bare parameters to the renormalized ones. Predictions for observables are expressed in terms of renormalized couplings $\alpha(\mu_R^2)$, as a function of a renormalization scale μ_R. When one takes μ_R close to the scale of the momentum transfer Q in a given process, then the renormalized coupling is indicative of the effective strength of the interaction in that process.

The change in the value of α as a function of the renormalization scale is measured using higher order corrections to the interaction propagator (resummation of self-energy). The resulting equation is called the renormalization group equation (RGE). For the strong interaction, the RGE satisfies:

$$\mu_R^2 \frac{d\alpha_S}{d\mu_R^2} = \beta(\alpha_S) = -(b_0\alpha_S^2 + b_1\alpha_S^3 + \cdots) < 0, \tag{1.24}$$

where "b" coefficients are calculated with a corresponding number of loops for the higher order corrections, and $b_0 = (11C_A - 4n_f T_R)/12\pi$. Since we have 3 colors and the number of light quark flavors ($m_q << \mu_R$), n_f, is at most equal to 6, b_0 is positive which implies that β is negative [1].

The fact that the right-hand side of the RGE is negative for α_S, meaning the coupling value decreases with the renormalization scale (higher scale ≡ smaller distances), has an important consequences on the strong interaction. It is the origin of quarks and gluons confinement and asymptotic freedom described in the next section. Figure 1.4 shows the running of α_S as measured by multiple experiments. The value of α_S is also shown at a scale $\mu_R = M_Z$.

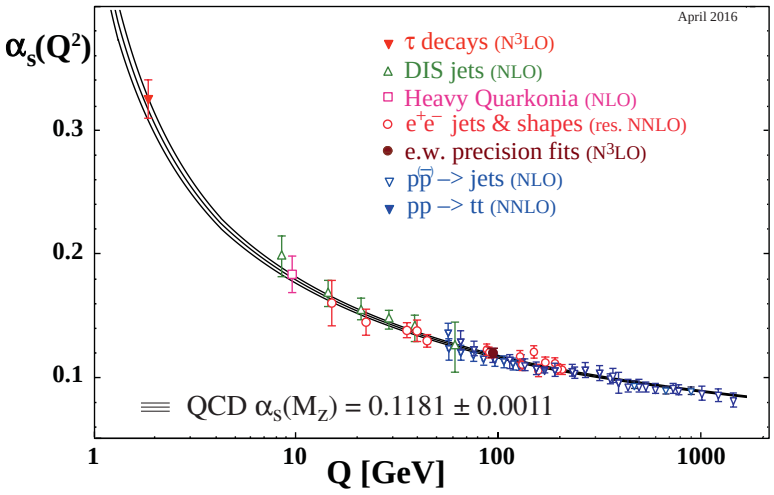


Figure 1.4: Summary of measurements of α_S as a function of the energy scale Q. The respective degree of QCD perturbation theory used in the extraction of α_S is indicated for each experiment where some reach a precision of N3LO. Figure from PDG book [1].

In figure 1.5, the value of $\alpha_S(M_Z^2)$ is shown for multiple experiments and the world average measured is:

$$\alpha_S(M_Z^2) = 0.1181 \pm 0.0011. \tag{1.25}$$

At last, it is worth noting that for the weak and electro-magnetic couplings, the β term is positive and the couplings increase with the renormalization scale.

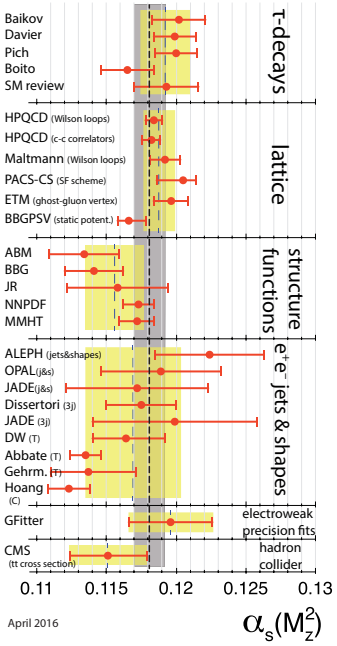


Figure 1.5: Summary of determinations of $\alpha_S(M_Z^2)$ from six sub-fields. The dotted line and grey (dark shaded) band represent the world average value. Figure from PDG book [1].

1.2 Deeper into Quantum Chromo-Dynamics (QCD)

1.2.1 Asymptotic freedom

As we have seen in the previous section, the strong coupling constant α_S becomes smaller at high energy scales. If we let the coupling run until infinite energy using the RGE, α_S become zero. This phenomenon is referred to as "asymptotic freedom" and implies that at high energies, quarks and gluons behave as if they were free.

For energies higher than few GeV, the strong coupling constant α_S becomes significantly smaller than 1. The strong interaction can then be describes in a perturbative way. Figure 1.6 shows the interaction vertices for QCD, while figure 1.7 shows some of the Feynman diagrams for QCD interaction at tree level when the perturbative approach is applicable.

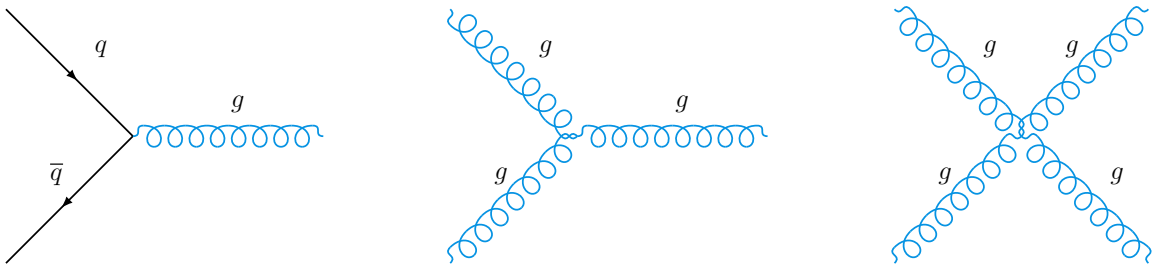


Figure 1.6: Interaction vertices for QCD.

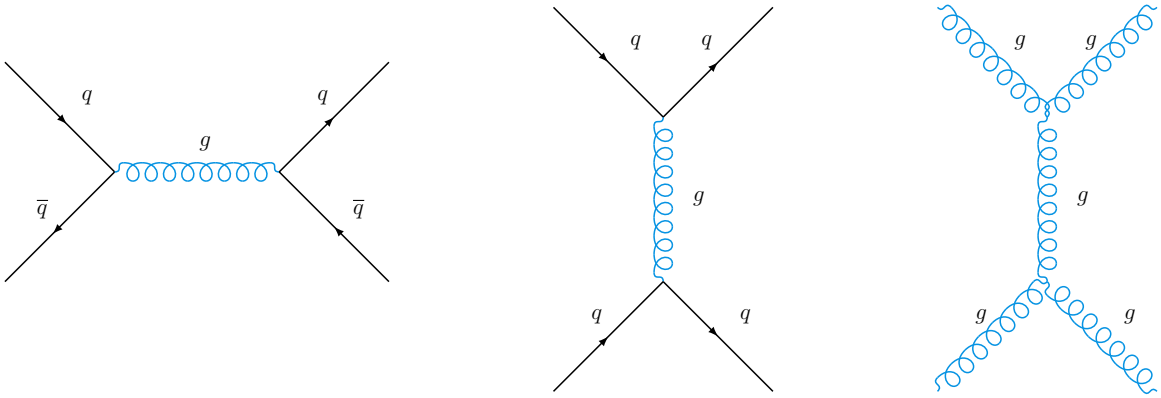


Figure 1.7: Feynman diagrams of some of the QCD interaction at tree order (LO).

1.2.2 Showering, hadronization and confinement

At high distances (≡ low energies), the strong coupling constant α_S have higher values. As a first consequence, after quarks and gluons are produced and begin to separate, they radiate/shower lots of gluons and exhibit $g \rightarrow q\bar{q}$ splittings. This process is named showering (also Bremsstrahlung).

A second important consequence, high α_S values at high distance leads to what is called the confinement of quark and gluons, collectively called partons. As an example showed in figure 1.8, if a pair of quark-antiquark are pulling away from each another, the strong interaction between them becomes stronger and stronger which either keeps them bound together or, if the two quarks were too energetic, the energy of the color field connecting the two quarks becomes high enough and creates an additional quark-antiquark pair. The confinement means that we do not find/detect partons in isolation. They are always confined in bound systems, called hadrons, which are color neutral, e.g. protons and neutrons.

Although in the perturbative approach we treat partons as free at small distances, we do in a second step consider the transit of outgoing particles from partons to color neutral hadrons, a process known as hadronization.

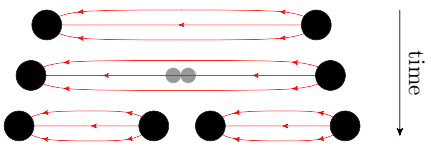


Figure 1.8: Illustration of pair of quark-antiquark creation (using the string model). Figure from PDG book [1].

Figure 1.9 shows an example of a hard perturbative QCD interaction followed by showering then hadronization processes.

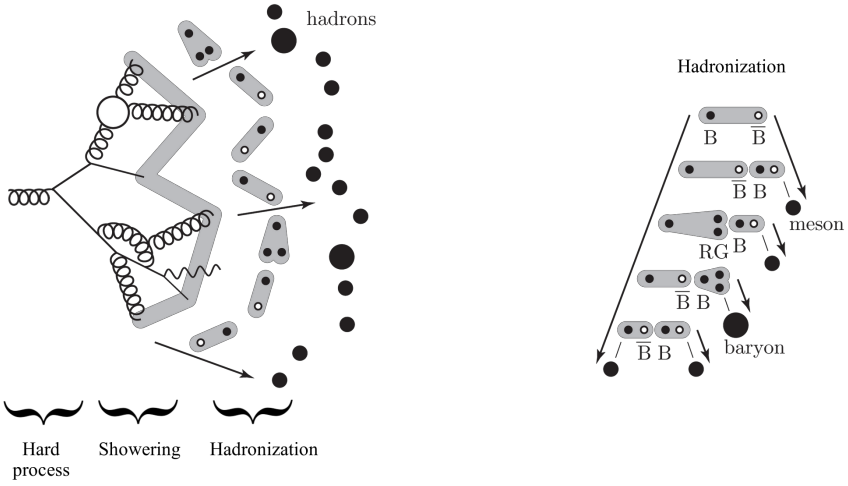


Figure 1.9: An example of a hard perturbative QCD interaction followed by showering then hadronization processes. The figure on the right shows a zoom on hadronization.

For small energies where the perturbative approach fails, lattice QCD method is sometimes used specially for describing the properties of hadrons which are used to extract the strong coupling constant and quark masses. In this method, space-time is discretized on a lattice/grid with a spacing "a". The lattice spacing plays the role of the ultraviolet regulator at the order of 1/a [17]. Quark fields are placed on the nodes and gauge fields on the links between nodes. Numerical solution methods including Monte Carlo sampling are used to solve the theory and compute the probabilities of field configurations. The precision of lattice calculations depends on its spacing, and the higher the energies, the smaller the spacing is needed. Hence due to computational resources, lattice QCD is limited to low energy region.

1.2.3 Jet definition

Following the showering and hadronization of a hard parton, we end up with a spray of particles around the direction of the original parton. Those particles can be stable particles in simulations or tracks and energy deposits in detectors. One would like to regroup those particles into one object, called a "jet", that characterizes the initial parton that seeded them. Jets are therefore a tool to look at the first instants after a collision and are proof of the asymptotic freedom property of QCD: jets are seeded by partons that are free at high energies.

Jet algorithms, also called jet finders, are a set of rules with some parameters, for grouping the particles produced in the final state of a high energy collision into jets. The algorithm runs over all the particles in an event and returns the jets. Two main classes of jet algorithms are the cone-type and the clustering-type classes.
For cone-type algorithms, the position and number of jets, which are expected to be cone-like agglomerations of particles, are found iteratively by measuring the flow of energy within a cone over all possible placements of the cones until reaching a stable one [18]. They are a kind of top-down algorithms: they begin with a predefined structure and search where it fits in the event.
On the other hand, clustering-type algorithms are of bottom-up kind: they start with the constituents and assemble them sequentially two particles on each step. Clusters are formed until reaching a set of stopping rules and form the jets.

An important property to check in jet algorithms is if their output is infrared and collinear (IRC) safe. As we will see in more details in section 1.3.1, QCD develops divergences when emitting an infinitely soft (infrared) or collinear particles. Luckily, the divergence is canceled by higher order contributions (additional loops). For the cancellation to work, the algorithm output should not be sensible to IRC emissions such that the contributions from the different perturbative orders are correctly summed.

Pre-LHC era, cone-type algorithms were used mainly in hadronic collisions since such collisions produces many particles and one needs an algorithm that remains reasonably fast at large multiplicities. An inconvenience is that cone-type algorithms available at that time were IRC unsafe. In e^+e^- collisions, clustering-type algorithms were used because they were IRC safe and provided more accurate predictions [19].

In the following, I will only detail the latest development in the clustering-type class. Those algorithms use particle-particle and particle-beam distances defined respectively as:

$$\begin{aligned} d_{ij} &= \min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{(y_i - y_j)^2 + (\phi_i - \phi_j)^2}{R^2}, \\ d_{iB} &= p_{Ti}^{2p}, \end{aligned} \tag{1.26}$$

where the jet radius R and the exponent p are parameters of the algorithm. The two distances are calculated for all constituents. If the smallest is an inter-particle one, the two particles are combined by summing their four-momenta and replaced by the resulting particle. On the other hand, if the smallest is a beam-distance, the considered particle is labeled a jet and excluded from further iterations.

Setting p to 1 corresponds to the k_t algorithm [20] where the clustering distance is proportional to the inverse of the emission probability. Setting p to 0 corresponds to the Cambridge/Aachen algorithm

[21] where the clustering is based exclusively on the angular distance between the particles. The distance measures of these two algorithms are modeled on the physical behavior of the QCD emission probability. They are expected to walk backward the parton-hadrons processes beginning by the combination of soft and collinear fragments up to reconstruct the original hard parton.

A new approach abandons the connection between clustering and emission hierarchy. The algorithm is named "anti-k_t" [22] and sets p=-1. The algorithm begins to cluster soft particles around the hardest one. Soft particles do not modify the shape of the jet, giving the algorithm resilience to soft radiation. This algorithm turned out having many advantages and is adopted as the default algorithm by LHC experiments. In addition to being IRC safe as the others, it produces jets with very regular borders (circular if isolated with an area of πR^2) and that do not usually extend beyond a distance R from the hard particle(s) that seeds them. Also, the fluctuation of the jet area caused by soft particles is almost zero. These properties reduce the sensitivity to the background noise and facilitate experimental corrections for these and detector-related effects. The jet momenta smearing is thus reduced.

Figure 1.10 shows an example of jet clustering using two algorithms: k_t (left) and anti-k_t (right). The regular jet shapes of the later are visible.

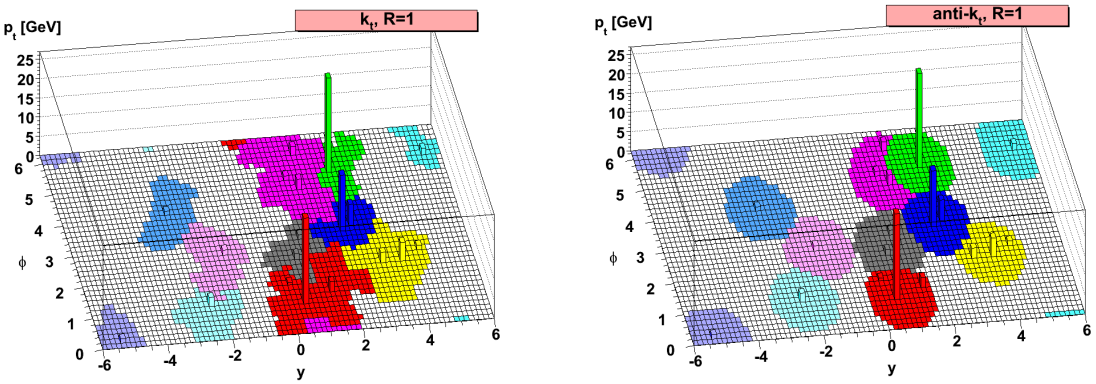


Figure 1.10: An example of jet clustering using two algorithms: k_t and anti-k_t. Figure from [22].

1.2.4 The contents of the proton: Parton Distribution Functions (PDF)

Due to confinement, scattering experiments containing quarks and gluons in their initial states can only be performed using colorless bound states hadrons, like protons. Hadrons cannot be described using perturbation theory and their structure cannot be determined by the first principles of QCD[5]. As a consequence, the determination of their properties relies on experimental input. The parton model

[5] Still, lattice QCD allows the hadron structure description from the first principles, but the approach applicability is limited to low energies.

describes the hadron as a composite object made up of quarks and gluons. In fact, in addition to the valance quarks that constitute the hadron (e.g. uud for a proton), the hadron also contains gluons and pairs of $q\bar{q}$ (sea quarks) that are part of the strong interaction binding the hadron. The distribution of the momentum of the hadron among its constituents is described by the parton distribution functions (PDF) $f^{(0)}_{a/H}(x_a)$, the probability density to find a parton "a" carrying the longitudinal momentum fraction x_i of the parent hadron H.

The PDFs are mainly measured using the data from deep inelastic scattering experiments (DIS) of $e^{\pm}$ and hadrons [1]. The scattering involves EW processes by the exchange of photons and Z for neutral-current interactions and $W^{\pm}$ for charged-current interactions. PDFs of q and $\bar{q}$ can be directly extracted. For the gluon which does not interact electroweakly, we use the momentum sum rule:

$$\sum_a \int x_a \, f_a(x_a) \, dx = 1. \tag{1.27}$$

For the valence distributions and taking the proton as an example, the following equations apply:

$$\int (f_u(x) - f_{\bar{u}}(x)) \, dx = 2 \, , \quad \int (f_d(x) - f_{\bar{d}}(x)) \, dx = 1.$$

In addition to DIS data, PDFs measurements also use p-p and p-$\bar{p}$ data to further improve their precision. For example [23], jet production (inclusive jets, dijet ...) allows to better constrain quarks and gluons at medium and large x, a region where the constraints from DIS data are only indirect, while inclusive production of W and Z bosons provides constrains on quark flavors separately. PDFs depend on the energy scale at which they are probed. Figure 1.11 shows the momentum probability distribution $x_a f_a(x_a)$ for the different constituents "a" of the protons for two energy scales. At high x values, valence quarks dominates while at low x values the gluons do. Also, it is worth noting that some PDFs now do include photons distribution [24].

PDFs are universal, they do not depend on the actual scattering reaction that probes them. They can be extracted from one type of experiment, e.g. e-p scattering, and used in another one, e.g. hadron-hadron collisions.

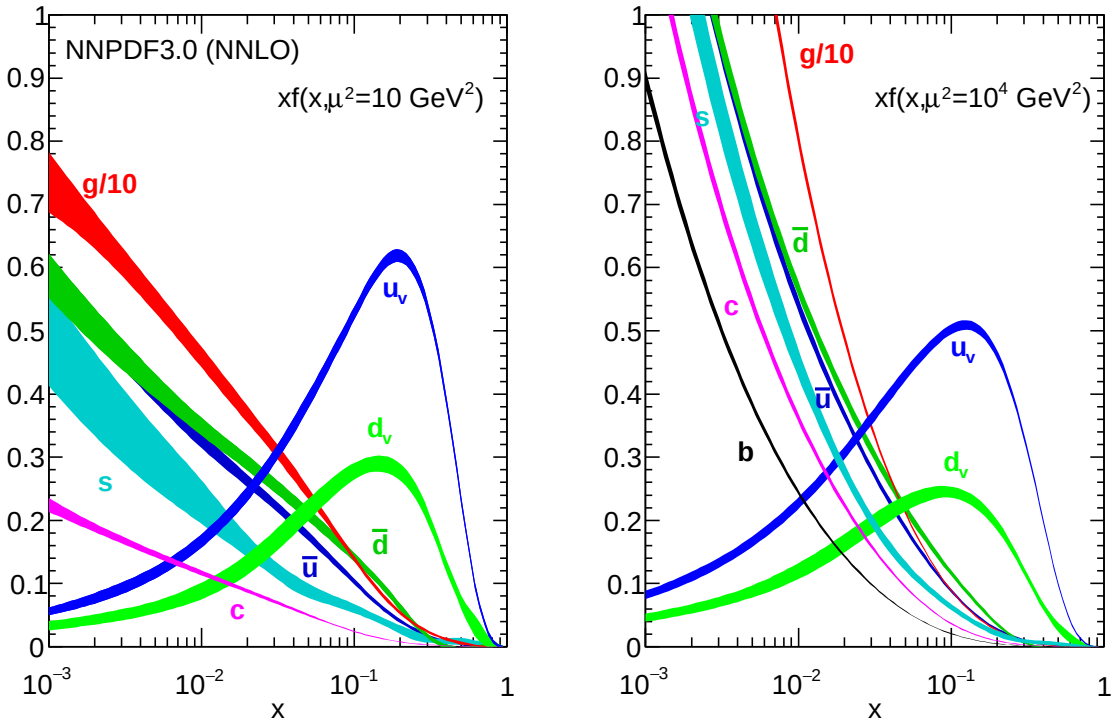


Figure 1.11: Momentum probability distribution $x_a f_a$ for the different constituents "a" of the protons are shown for two energy scales. Figure from PDG book [1].

1.2.5 p-p collisions

In p-p (same for any hadron-hadron) collisions, the total cross section for a scattering process is factorized into a hard perturbative partonic sub-processes $\hat{\sigma}_{ab}$ convoluted with proton PDFs:

$$\sigma_{pp\to X} = \sum_{a,b} \int_0^1 dx_a \int_0^1 dx_b \; f_{a/p}(x_a) f_{b/p}(x_b) \hat{\sigma}_{ab\to X}(p_a, p_b, \mu_R^2), \tag{1.28}$$

with $p_{a,b} = x_{a,b}\, p(\text{proton})$. This approach works for leading order calculations. When considering collinear initial-state splitting of partons, divergences arise. The cancellation of divergences requires the inclusive treatment of degenerate states by summing with the virtual contributions. In the case of initial-state splitting, the cancellation does not work because collinear splitting modifies the momentum that enters the hard scattering process, whereas for virtual loops the momentum is unchanged, as is shown in figure 1.12.

The initial-state collinear singularities correspond to non-perturbative effects. Therefore, the workaround to the singularities is to attribute them to the description of the proton and thus to the PDFs. This is

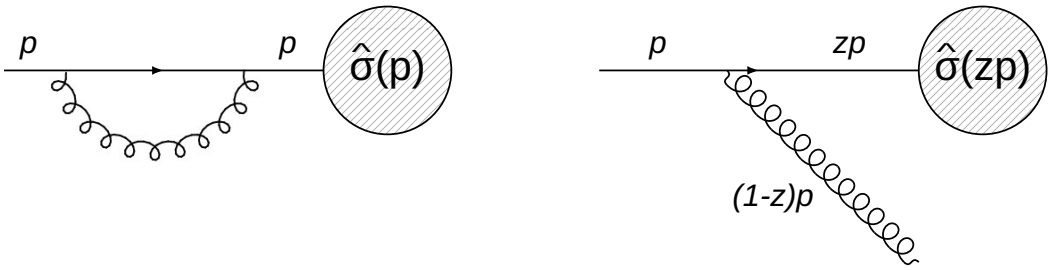


Figure 1.12: Feynman diagrams for initial-state parton splitting at the right and virtual contribution at the left, where the shaded circle represents the hard partonic scattering with its corresponding cross-section which depends on the entering parton momentum.

done, using similar procedures to renormalization, by substituting the PDFs by:

$$f_{a/p}(x) \rightarrow f_{a/p}(x, \mu_F^2) - \text{constant} * \sum_b \int_x^1 \frac{dz}{z} f_{b/p}(\frac{x}{z}, \mu_F^2) P_{ab}(z), \tag{1.29}$$

where P_{ab} are Altarelli-Parisi splitting functions [6] [25], describing the probability density of the splitting process $a \rightarrow b + c$. This procedure introduces a new scale μ_F, called the factorization scale, which separates the long-distance from the short-distance interactions effectively moving the collinear singularities from the hard process to the PDFs. The new total cross section is now:

$$\sigma_{pp\rightarrow X} = \sum_{a,b} \int_0^1 dx_a \int_0^1 dx_b \; f_{a/p}(x_a, \mu_F^2) f_{b/p}(x_b, \mu_F^2) \hat{\sigma}_{ab\rightarrow X}(p_a, p_b, \mu_R^2, \mu_F^2), \tag{1.30}$$

The additional dependence of the PDFs on the factorization scale μ_F must vanish if all orders in perturbation theory are taken into account. This condition leads to the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi [25–27] (DGLAP) evolution equations for the PDFs:

$$\frac{\partial}{\partial(\ln \mu_F^2)} f_{a/p}(x, \mu_F^2) = \frac{\alpha_S}{2\pi} \sum_b \int_x^1 \frac{dz}{z} f_{b/p}(\frac{x}{z}, \mu_F^2) P_{ab}(z). \tag{1.31}$$

Using the DGLAP evolution equations, on can be extrapolate PDFs from the scales where they are measured at to the desired one in order to make theoretical predictions.

In the event of a hard scattering, the proton disintegrates and its constituents are sprayed in all directions after forming colorless states. In addition, secondary soft scatterings can happen between the partons from the two protons. Color connections can also exist between the hard outgoing partons and the protons remnant. Those effects are labeled under the underlying event (UE).

Figure 1.13 sketch the different effects, hard and soft, in a hadron-hadron scattering event.

[6] Considering quarks and gluons, we have 4 splitting functions: P_{qq}, P_{qg}, P_{gq} and P_{gg}.

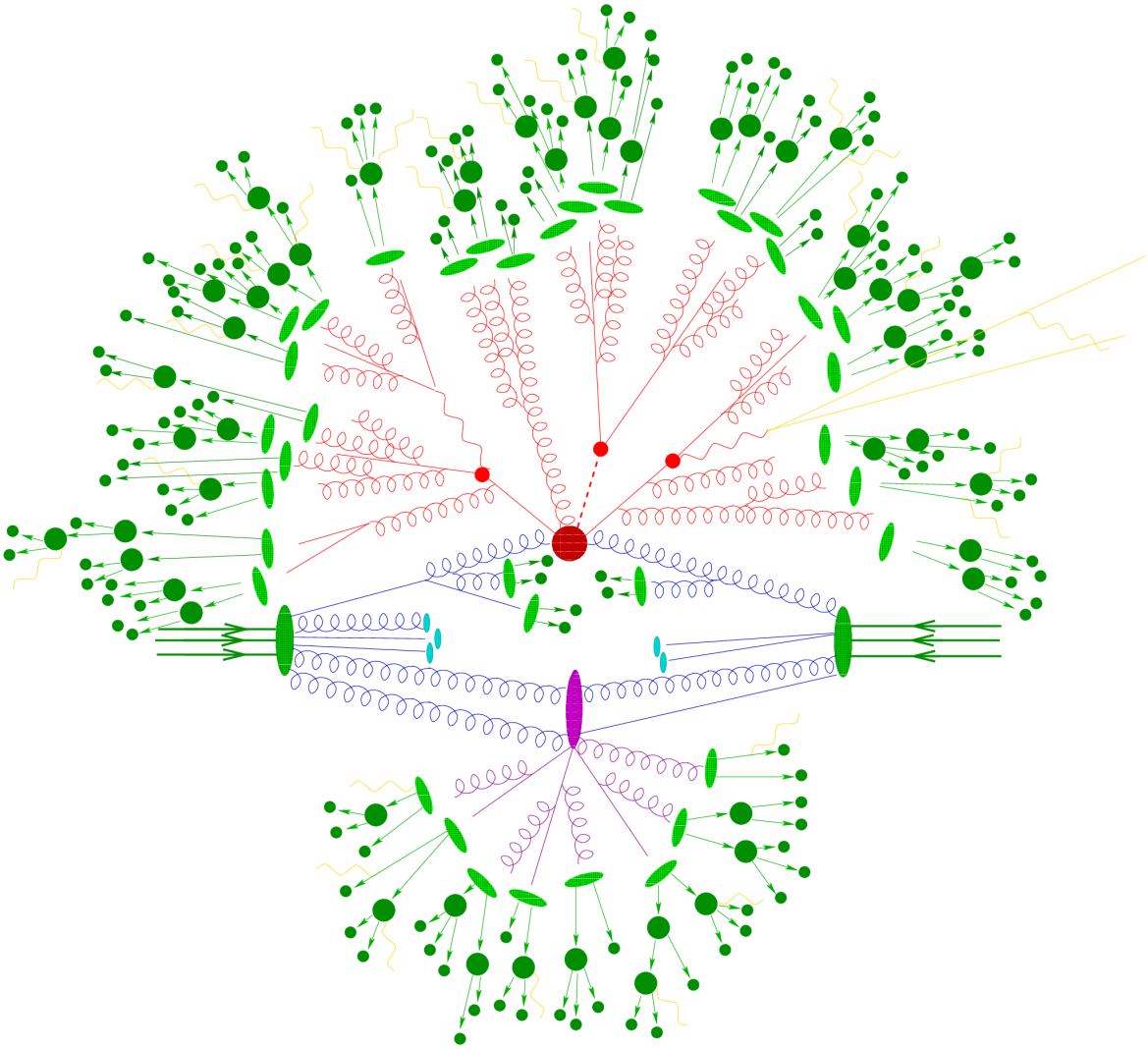


Figure 1.13: Sketch of a hadron-hadron collision. The red blob in the center represents the hard collision, surrounded by a tree-like structure representing Bremsstrahlung as simulated by parton showers. Parton-to-hadron transitions are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation. The purple blob indicates a secondary softer scattering event. Blue blobs represent the hadronization of the rest of the protons partons. Purple and blue blobs represent the underlying event. Figure from [28].

1.3 Precision predictions

Now that the SM is introduced with a focus on QCD, we have the necessary building blocks to make theoretical predictions. I show next some of the different approaches that are relevant to this thesis.

1.3.1 Partonic predictions

The first type of theoretical predictions uses partons in the final state to calculate desired observables. Jets can then be formed using the partons as inputs. Event are generated by:

- randomly choosing a phase-space configurations (e.g. 2 partons for dijet events at LO),
- calculate the event weight from the amplitude of S-matrix elements convoluted with the PDFs.

When trying to include higher order corrections, one would have to deal with divergences in the calculation. Taking for example a q-$\bar{\text{q}}$ event and trying to calculate the NLO corrections (NLO diagrams shown in figure 1.14), the cross section of the scattering with a soft gluon emissions is:

$$\mathrm{d}\sigma_{\mathrm{q\bar{q}g}} = \mathrm{d}\sigma_{\mathrm{q\bar{q}}} \times \frac{\alpha_{\mathrm{S}}}{\pi} \frac{\mathrm{d}\theta}{\theta} \frac{\mathrm{d}\phi}{2\pi} \underbrace{2C_F \frac{dE}{E}}_{\equiv \mathrm{P_{qq}}}, \tag{1.32}$$

where E is the energy of the gluon and θ the angle between the gluon and the quark emitting it. The cross section with soft gluon emission factorizes into a product of the cross section of the hard scattering times the probability density to emit a gluon. Two divergences arises from the latter, corresponding to the integration of the real-emission corrections over the unresolved regions:

- infrared (soft) divergence when $\mathrm{E} \to 0$,
- collinear divergence when $\theta \to 0$ or π.

These singularities are also present in the virtual correction with an opposite sign. The inclusive treatment of such degenerate states leads to the exact cancellation of all IR singularities between the virtual and the real contributions in physical observables. This was proven in the Kinoshita-Lee-Nauenberg (KLN) theorem [29, 30].

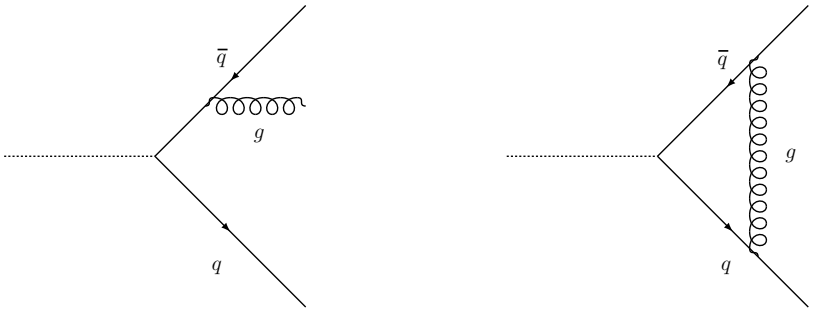


Figure 1.14: NLO Feynman diagrams for a $q\bar{q}$ event with real emission on the left and virtual contribution on the right.

For higher order perturbative calculations, proper summation of virtual and real contributions that have the same power in α_{S} is crucial. These are called fixed-order predictions. Needless to say, the observable should also be IRC safe. A sketch summarizing the different diagrams that need to be included in each calculation order for QCD events is shown in figure 1.15. Positive and negative

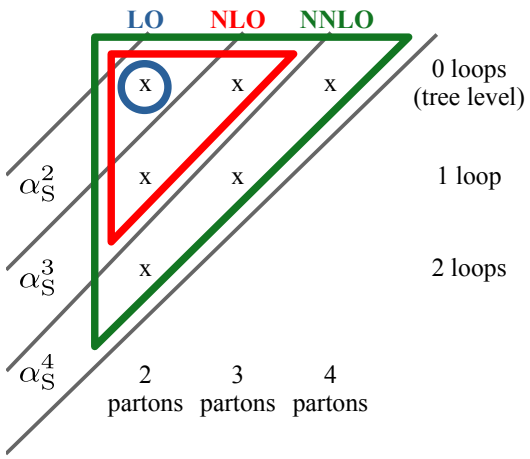


Figure 1.15: A sketch summarizing the different diagram types (number of partons and loops) that need to be included in each calculation order for QCD events.

event weights for NLO and higher orders will be generated, but with the total sum being positive for cross sections.

In addition, when the observable is measured close to the boundary of its phase-space described by a distance $y << 1$, the perturbative power expansion in α_S involves terms of the type $\alpha_S^n L^k$ ($k \leq 2n$), where $L = -\ln y$. The most dominant terms are the leading logarithm (LL) for $k = 2n$, next comes the next-to-leading logarithm (NLL) for $k = 2n - 1$ and so on. These double logarithmic terms are for example in our case due to final-state radiation of soft and collinear partons. For $y << 1$, the logarithm L becomes much larger than 1 which spoils the convergence of the series. Those terms are summed to all orders in α_S ($n \to \infty$) to recover the convergence of the prediction. This procedure is called resummation [31].

The renormalization and factorization scales, μ_R and μ_F, should be chosen for the predictions calculation. Several choices can be done, for example p_T of the leading jet, p_T of each jet, invariant dijet mass m_{jj}, mean of the two leading jets $H_{T2}/2$, mean of the two leading partons $\hat{H}_{T2}/2$, etc. Each scale choice has its own motivations. Usually μ_R is taken equal to μ_F, to have more stable predictions since terms like $\ln(\mu_R^2/\mu_F^2)$ are found in the calculations which can lead to large variations if the functional forms of the two scales differ by a non multiplicative factor.

As we have seen before, since we are not including all perturbative orders in our calculations, a residual dependence on the scale exists for α_s and for the PDFs. To try to evaluate the effect of this dependence, each of μ_R and μ_F scales is conventionally varied by factors of 0.5 and 2. The change on the observable is taken as an uncertainty. Usually, this method covers well the effect of missing orders except when new production channels are introduced in higher orders.

A widely used algorithms with NLO accuracy is the NLOJET++ [32, 33] package which uses Catani-Seymour dipole subtraction method [34] to account for the singularities in real emission and virtual

contributions. Algorithms with NNLO accuracy have become available recently, like the NNLOJET [35, 36] package.

1.3.2 Non-perturbative MC simulation

Partonic predictions, although very powerful tools, lack the non-perturbative (NP) treatment of the final state particles. Monte Carlo (MC) event generators implement multiple techniques to take into account the NP effects. The main steps for generating an event are:

1. perturbative processes:
    - hard partonic scattering,
    - resonance decays,
    - parton showering.
2. non-perturbative effects:
    - underlying event: multiple particles soft interaction, protons remnant,
    - hadronization,
    - hadron decays,
    - Bose-Einstein and color reconnection.

Let us detail some of these steps.

Showering
As a first step, hard partonic scattering is generated from matrix-elements (ME) calculated using the Feynman rules. Let us first consider generators that uses only LO ME (2→2 partons processes). The two most used are PYTHIA [37, 38] and HERWIG [39, 40]. In soft-collinear limits, it is hard/impossible to calculate parton splitting/radiating effects using ME. For that, parton showering algorithms are used.

First let us express the probability of a parton not radiating above a scale k_T:

$$\begin{aligned}\Delta(k_T, Q) \equiv P(\text{no emission above } k_T) &= 1 - \frac{\alpha_S}{\pi} \int\int \frac{d\theta}{\theta} \sum_{a,b=q,g} P_{ab}\, \Theta(E\theta - k_T), \\ &\approx \exp\left(-\frac{\alpha_S}{\pi} \int\int \frac{d\theta}{\theta} \sum_{a,b=q,g} P_{ab}\, \Theta(E\theta - k_T)\right),\end{aligned} \tag{1.33}$$

where "Q" is the energy of the initial parton, "E" and "θ" are the energy and angle of the emitted parton with respect to the initial one, and the Θ function ensure to have an emission above the scale k_T. $\Delta(k_T, Q)$ is a Sudakov form factor [41]. The relation between the Sudakov form factor and the probability of a radiation P_R is:

$$\frac{dP_R}{dk_T} = \frac{d\Delta(k_T, Q)}{dk_T}. \tag{1.34}$$

Using MC methods, this distribution can be generated and we get the first emission. The procedure is repeated with smaller and smaller scales until reaching a cutoff. This showering method is based on p_{T} ordering and is used by PYTHIA. On the other hand, HERWIG uses this procedure but with an angle ordering.

Parton showering allows to pass from two- to multi-partons final state. This recovers to a certain degree the missing higher orders for the soft-collinear limits but it fails to describe well the hard large-angle radiation. NLO-ME generators try to improve the description of those emissions, like POWHEG [42–44] generator which is used along with PYTHIA or HERWIG for parton showering, or SHERPA [45]. Additional partons can now be produces either from the ME or from the showering. Matching between the two production mechanism is used to avoid double counting. The improvement of those generators are mostly important when events with widely separated jets need to be well described. For example, it is important for the description of dijet systems where the third jet plays an important role.

Hadronization

Hadronization denotes the non-perturbative transition process of colored partons to colorless hadrons. This step follows the parton showering where also the color assignment of the partons has been done. The two most common models used are the string model, used by PYTHIA, and the cluster model, used by HERWIG.

The string model was briefly introduced in section 1.2.2 A pair of q-$\bar{\mathrm{q}}$ are connected through a color potential which increases linearly at high distances, proportional to a factor k. The potential can be thought of as a string with a tension k. As the string grows, the non-perturbative creation of q-$\bar{\mathrm{q}}$ pairs can break the string as illustrated in figure 1.8. Gluons are also treated in a similar way except that, since they have two colors, they connect to two string pieces. They create a transverse structure in the originally one-dimensional object. Also for gluons, the rate of hadron production is twice larger than the one of quarks. The model steps are as follows. At first, it maps color-connected pairs of partons to string pieces. Next, strings evolve and break independently forming new q-$\bar{\mathrm{q}}$ pair. Selecting randomly between the left and right sides of the broken string, one hadron is generated and the other leftover string enters the next iteration of the process. Baryon production can also be incorporated. The process is finished by decaying the low energy string to two hadrons directly.

The cluster hadronization model is based on pre-confinement, assuming that color-singlet subsystems of partons, labeled clusters, are suppressed at large masses. Gluons are forced to split to light q-$\bar{\mathrm{q}}$ pairs before the hadronization process begins. Therefore, clusters are only formed by q-$\bar{\mathrm{q}}$ pairs. If a cluster has an invariant mass above some cutoff value, it is broken into two. Low mass clusters decay directly to a single hadron, while the more massive ones decays into two hadrons with relative probabilities proportional to the available phase space for each possible two-hadron combination.

1.3.3 Non-perturbative corrections to partonic predictions

As we saw above, partonic predictions are precise and flexible, while MC generators include NP effects. In order to benefit the most from the two, partonic predictions times NP correction factors $\mathcal{K}_{\mathrm{NP}}$ are used. The factors $\mathcal{K}_{\mathrm{NP}}$ are calculated from the MC simulation events as follows:

- calculate the desired observable using only hard processes (partonic level): O(ME+showering),
- recalculate the same observable now adding the NP effects (hadronic level): O(ME+showering+hadronization+UE),
- the correction factor in each observable bin i is:

$$\mathcal{K}^{i}_{NP} = \frac{O^{i}(\text{ME + showering + hadronization + UE})}{O^{i}(\text{ME + showering})}. \tag{1.35}$$

The factors are then multiplied by the partonic predictions in each bin:

$$O^{i}(\text{partonic + NP effects}) = O^{i}(\text{partonic}) \times \mathcal{K}^{i}_{NP}. \tag{1.36}$$

Multiplication factors are one way to include the NP effects into the theoretical predictions. The other would be to build a 2D migration matrix, $\mathcal{M}$(partons $\rightarrow$ hadrons), between the partonic and hadronic levels, and then convolute the partonic predictions with this matrix:

$$O^{i}(\text{partonic + NP effects}) = \sum_{j} O^{j}(\text{partonic}) \times \mathcal{M}^{ji}(\text{partons} \rightarrow \text{hadrons}) \tag{1.37}$$

Figure 1.16 shows an example of the correction factors for the case on inclusive jet cross-section using multiple MC generators and tunes. A tune consists of a PDF set choice and an optimization of the MC generator parameters (affecting the showering, the hadronization, the UE ...) using information hadron collision data. The first thing to notice is that the factors calculated using PYTHIA and HERWIG generators are systematically different. This is expected since both use different models for NP effects (different hadronization models, ...). The difference is taken as a systematic uncertainty. The second thing is that the factors are only a few percent away from unity. In fact the hadronization and UE effects work in opposite directions and compensate to some degree: the hadronization tends to lower the energy of the jet due to the hadron creation outside the jet area, while the UE adds energy in all of the phase space and thus also inside the jet.

An improved method to add the NP effects is to use instead a transfer matrix relating the two calculation levels (partonic and hadronic) described above. For the same event, the observable is calculated at both levels. Next, a (geometrical) matching is used to match jets between both levels, the observables associated to which are then used to fill a 2D histogram. Figure 1.17 shows an example of a work-in-progress transfer matrix for a jet p_{T} observable. The advantage of this method is its lower dependence on the underlying perturbative calculation and observable spectrum shape of the MC generator used to calculate the NP effects (LO for Pythia and Herwig). The method also takes into account the unmatched observable and the total matrix used is called the folding matrix. An existing use-case of the folding method is describe in chapter 4, section 4.3, where more details are given for the construction and usage of the matrix.

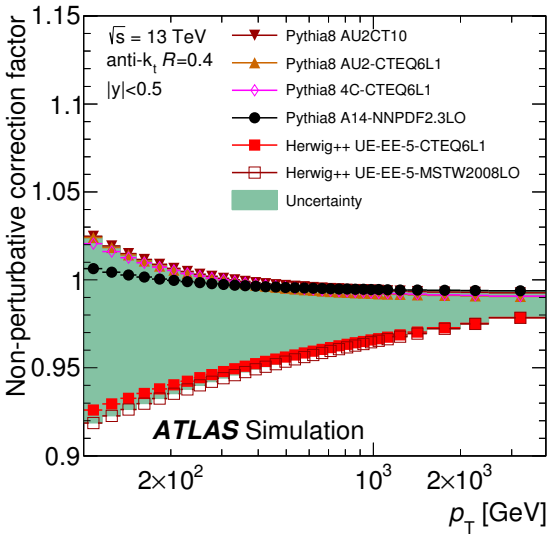


Figure 1.16: An example of NP correction factors for inclusive jet cross section using multiple MC generators and tunes.

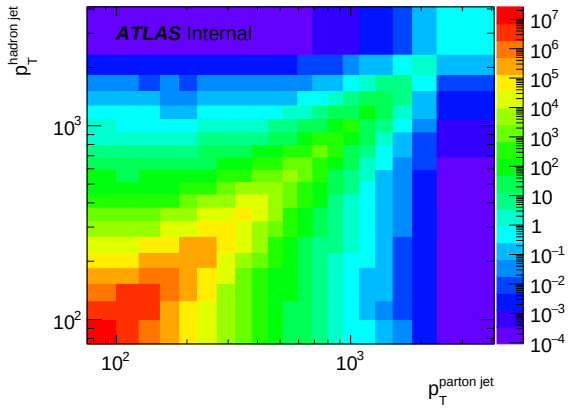


Figure 1.17: An example of a transfer matrix used to correct the NP effects for a jet p_T observable.

1.3.4 EW corrections to QCD prediction

The total jet cross-section is dominated by QCD contributions due to the much larger coupling constant of the strong force. Nevertheless, given the accuracy of the current perturbative QCD calculations, EW contributions at NLO precision are added. The EW corrections can exceed 10% at high p_T due to an increasing logarithmic terms of Sudakov origin [46]. The EW contributions include Born contributions, virtual and photon radiative corrections. Figure 1.18 shows the inclusive jets differential cross-section at NLO precision including both QCD and EW contributions as a function of p_T. In addition to the total one, all the individual contribution terms in the perturbative series that factorize the coupling constant combinations $\alpha_S^n \alpha^m$ are shown (where "n" and "m" are integers),

with n + m = 2 for LO and n + m = 3 for NLO contributions. The EW correction factors are hence defined as the ratio of the full QCD+EW contributions to only QCD ones by:

$$\mathcal{K}_{\mathrm{EW}} = \frac{\mathcal{O}(\alpha_{\mathrm{S}}^{\mathrm{n}}\alpha^{\mathrm{m}})}{\mathcal{O}(\alpha_{\mathrm{S}}^{\mathrm{k}})}, \tag{1.38}$$

where for our case of NLO precision corrections k = n + m = 2 & 3.

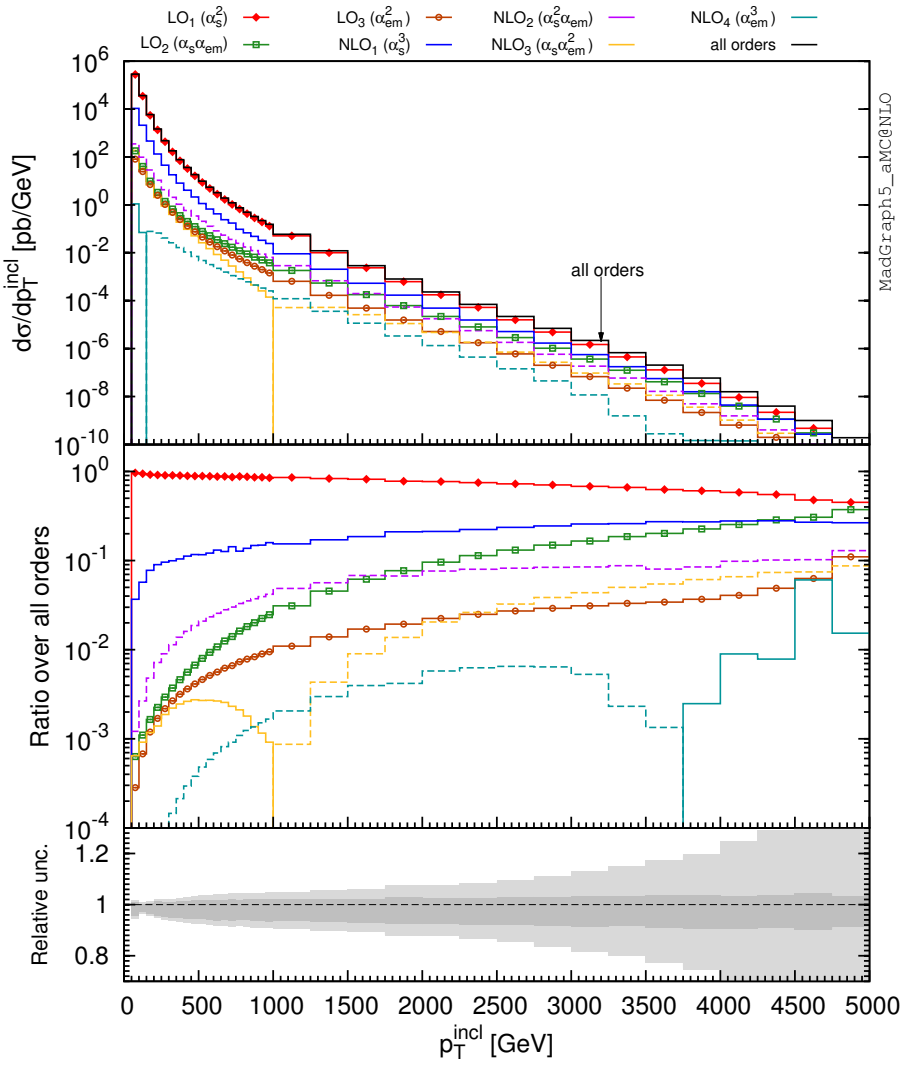


Figure 1.18: Inclusive jets differential cross-section at NLO precision including both QCD and EW contributions as a function of p_{T}. In addition to the total one, all the individual contribution terms in the perturbative series that factorize the coupling constant combinations $\alpha_{\mathrm{S}}^{\mathrm{n}}\alpha^{\mathrm{m}}$ are shown, with n + m = 2 for LO and n + m = 3 for NLO contributions. Figure from [46].

1.4 Beyond Standard Model (BSM)

During my PhD, I participated to search analyses of new BSM particles (see chapter 4). In this section, I present an initial context of the motivation of such searches and some of the BSM models available.

In spite of the nice theoretical features and successful experimental verification of the SM for many years, we have many reasons to think that the SM is not the final theory of Nature. First, there are some phenomenological issues, many coming from cosmology:

- the SM does not include the gravitational force: in fact, we don't know how to quantize the General Relativity,
- the SM does not explain the origin of the baryon asymmetry of the universe,
- the SM does not explain the dark matter [47],
- the SM does not explain the dark energy [48],
- no unification between EW and strong sectors,
- etc.

On the other hand, many questions are still open like:

- the origin of the symmetries,
- the origin of the three fermion generations,
- the origin of the neutrino masses,
- the mass hierarchy problem,
- the naturalness problem,
- etc.

For those reasons, we think that the SM is not the "theory of everything" and that there should be some new physics (at higher energies), the SM being a sub-section of a bigger theory. The SM can also be regarded as an effective low-energy approximation of a more general theory (Grand Unified Theory (GUT) [49], . . .).

When searching for new physics beyond the SM, two possibilities exist. The first is that the new physics scale is reachable by the experiment energies and therefore the new physics gives a resonance when probed at that scale. Specific model describes the interaction of these new physics. On the other hand if the scale is higher than what the experiment can reach, the new physics can still affect the number and topology of events. For those cases, the description of the new physics is usually done using effective field theories. I will give examples for both of those approaches, focusing more on the decay modes that contains quarks and gluons and hence producing events with jets, the final states of the search analyses described in chapter 4.

1.4.1 Resonant models

The models introduced below add new particles to the ones of the SM to cover some of the problems or questions listed in the previous paragraph. The details can be found in the references given for each model.

Excited quark q*
The idea of particle substructure is known as compositeness. According to composite models, SM quarks and leptons are predicted as composite particles, therefore excited states of those particles do exist. The q^* model [50] assumes an excited state of spin 1/2 and quark-like SM coupling constants. It couples to SM quarks and gluons via the strong gauge interactions, with a production and decay of the form: $q* \rightarrow qg$. The compositeness scale, Λ, is set to the q^* mass.

Quantum black hole (QBH)
The LHC should be able to produce QBH under the condition that the universe contains sufficiently large extra dimensions [51]. The quantum-gravity energy scale M_D at which micro black holes are produced decreases as the number, n, of these large extra dimensions increases. Through gravitational interactions, resonance-like production of predominantly two-body final states, mainly jets, M_D is predicted.

W′
A heavy charged boson, W′, is assumed in this model [52]. It can decay to $q\bar{q}$ producing dijet events. In this thesis, W′ is assumed to have the SM couplings. The only difference with respect to SM weak bosons is in their masses.

Z′
Models describing possible interactions between dark matter (DM) and the SM are also considered. In the Z′ model [53], the new spin-1 particle arises from a simple extension of the Standard Model (SM) with an additional U(1) gauge symmetry. DM particles are assumed to be Dirac fermions that have charges only under this new group. Assuming that some SM particles are also charged under this group, the Z′ can mediate interactions between the SM and DM. In this thesis, the Z′ is assumed to be lepto-phobic[7] which increases the sensitivity in dijet searches.

W*
The compositeness of the W boson is considered here, with W^* being the excitation state [54]. The W^* couples to $q\bar{q}$ decaying into dijet events. The angular distribution of the decays is more focused toward the forward region for this model. The W^* is also assumed to be lepto-phobic here.

[7] Lepto-phobic means no interaction with the lepton sector

Figure 1.19 shows some distributions of the dijet invariant mass produced by the described BSM models. The same mass is used for the BSM particles to show the different shapes that the models produce.

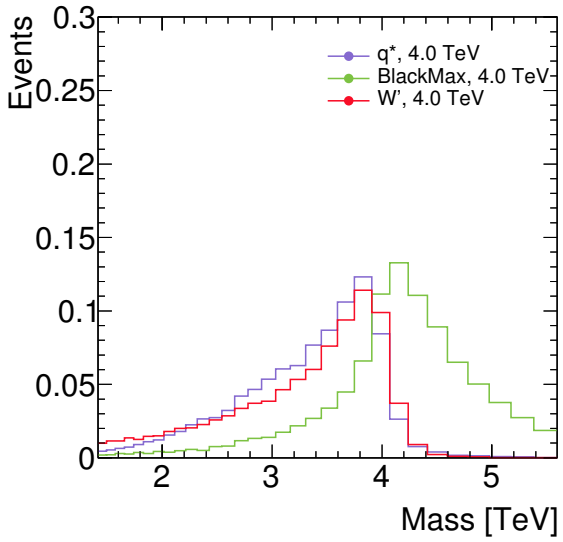


Figure 1.19: Invariant mass of the two leading jets, mjj, distributions for some BSM models. The same BSM particle mass is used in all the models.

1.4.2 Effective field model

Even if I don't describe the search analyses that use effective field models in this manuscript, I briefly present one model for completeness and since it is used in one occasion in chapter 4 to compare with the SM.

Contact interaction

Not all new physics phenomena need to appear as a localized excess. When the energy is not high enough to resolve the details of the interaction, non-resonant signals, described using effective field theories, are used as a probe. Here, an effective Fermi operator, a four-point interaction term or contact interaction (CI), describes an effective interaction between four quarks [55]. The Lagrangian of the interaction is:

$$\mathcal{L}_{qqqq} = \eta \frac{g^2}{2\Lambda^2} \bar{\Phi}_L^{Q_i} \gamma^\mu \Phi_L^{Q_i} \bar{\Phi}_L^{Q_j} \gamma_\mu \Phi_L^{Q_j}, \tag{1.39}$$

where "i" and "j" are quark flavor indices. $\eta = \pm 1$ represents the type of the interference between CI and QCD production modes: + for destructive and - for constructive interferences. Figure 1.20 shows the feynman diagram of the CI interaction. An analogy can be done with Fermi's theory ($\mathcal{L}_{\text{Fermi}} = G_F \bar{\Phi}^p \gamma^\mu \Phi^n \bar{\Phi}^e \gamma_\mu \Phi^\nu$) describing the β-decay ($n \rightarrow p e^- \bar{\nu}_e$) where the weak interaction is described by a 4-point contact term. In the SM, the interaction is rather described by the exchange

of a W^- boson, but still Fermi's theory is enough to describe the weak interaction at energies much lower than the W boson mass. The same principle is used here.

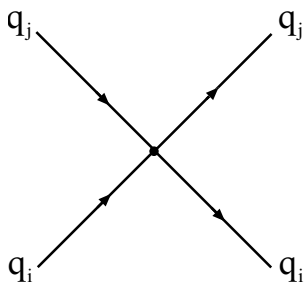


Figure 1.20: Feynman diagram for the CI model described in equation 1.39. "i" and "j" are quark flavor indices.

Figure 1.21 shows the dijet invariant mass distribution produced via QCD interactions with and without CI terms.

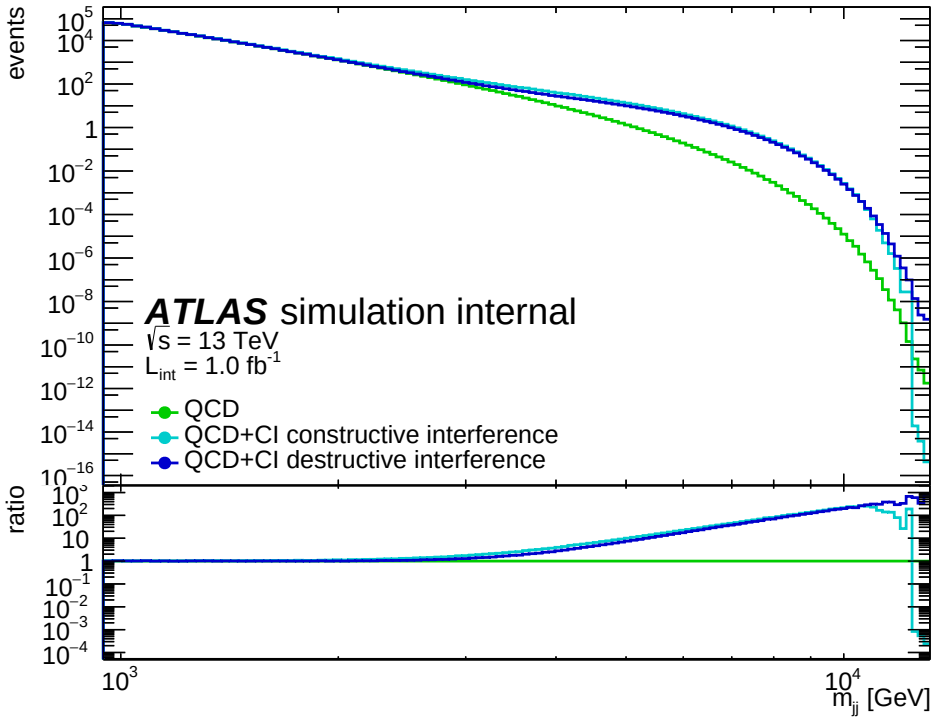


Figure 1.21: mjj distributions for dijet events produced via QCD interactions with and without CI terms, using Pythia8 generator (with showering and hadronization). The CI scale is $\Lambda = 7$ TeV.

2 The Large Hadron Collider (LHC) and the ATLAS experiment

In this chapter, I describe the ATLAS experiment. After a brief introduction of the LHC, I detail the ATLAS detector with an emphasis on the detector parts relevant to the jet reconstruction since they are the main objects used in all the analyses described in the following chapters. Last, I detail the jet reconstruction and calibration.

2.1 LHC

The Large Hadron Collider (LHC) is a superconducting hadron accelerator and collider installed in a 26.7 km tunnel[1] underground between 45 and 170 m below the surface. The "machine" is described in [56]. The LHC contains two rings with counter-rotating hadrons beams, accelerated with a center of mass energy up to 13 TeV and planned to go up to 14 TeV in the coming years. The delivered beams have 50 and 25 ns bunch spacing, where the latter is the nominal configuration for the Run II period during the years 2015 until 2018. The LHC benefits from the CERN accelerator complex that acts as injector. Four interaction points exist where the following detectors are found: two high-luminosity general detectors ATLAS and CMS, two lower-luminosity detectors, LHCb dedicated for B-physics and ALICE dedicated for ion collisions for the identification of the quark-gluon plasma [57]. Figure 2.1 shows the CERN accelerator complex which includes the different beam acceleration and injection stages from LINAC to PS to SPS to LHC.

[1] The tunnel previously hosted the Large Electron Positron (LEP) collider.

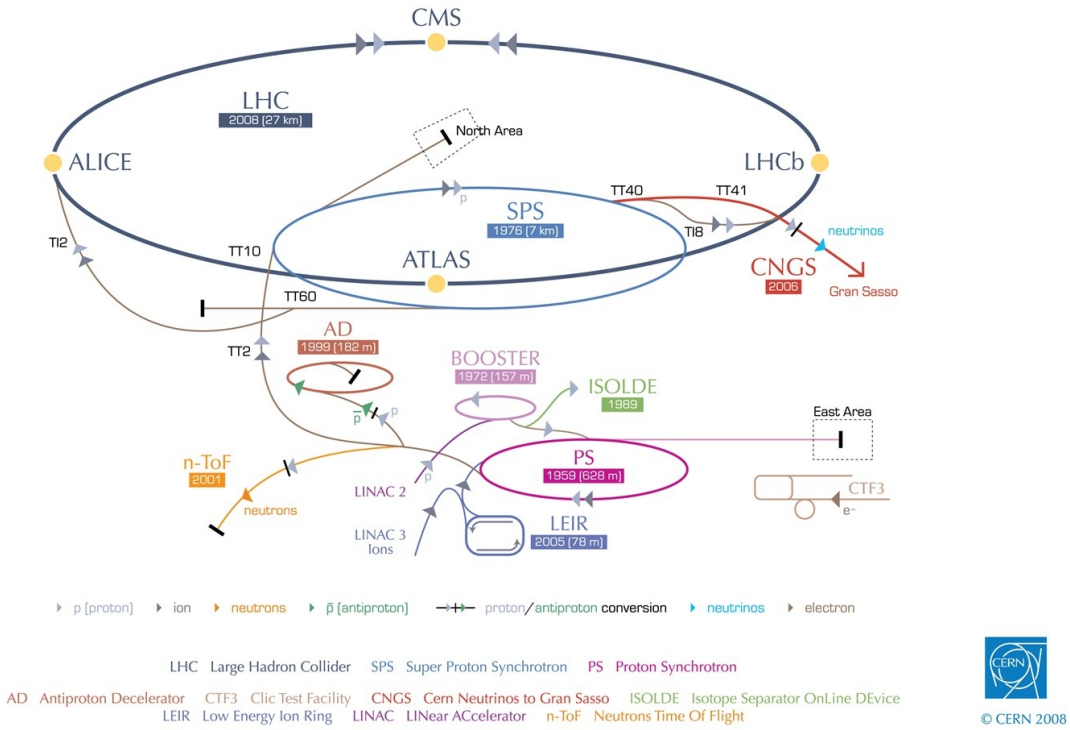


Figure 2.1: The LHC is the largest ring of the CERN chain of particle accelerator complex. The smaller machines are used in a chain to help boost the particles to their final energies. Figure from CERN website.

2.2 ATLAS: A Toroidal LHC ApparatuS

The ATLAS detector is one of the two general purpose detectors (along with CMS) that uses the LHC beam to collide Hadrons at the highest energy reached by any accelerator. The detector is described in [58]. The dimensions of the detector are 25 m in height and 44 m in length. The overall weight of the detector is approximately 7000 tonnes. It is used to probe p-p, p-Pb, Pb-Pb and other hadrons collisions. In this thesis, I will only focus on p-p collisions for both performance and physics studies.

One of the main goals of the ATLAS experiment is to prove the existence of a particle compatible with the predictions of the spontaneous symmetry breaking, the Higgs boson (introduced in chapter 1.1.1). This was achieved in 2012 and detailed in this paper [13]. Another goal is to search for new physics signals and test the many BSM models that exist.

The detector is composed of multiple sub-systems to achieve the highest detecting capabilities for various ranges of particle types and physics signals. The three main (and biggest) parts are the inner tracking detector, the electromagnetic and hadronic calorimeter and the muon spectrometer. They are built one on top of the other in a cylindrical onion arrangement. Each system is divided into a central barrel and two end-caps for the two forward regions. The calorimeters have also more

forward components. Figure 2.2 shows a cut-away view of the ATLAS detector with a labeling of the different components. Figure 2.3 shows how the detector reconstructs and identifies the different type of particles, like electrons, muons, photons, hadrons . . . Details for each of the main sub-systems of the detector are given in the next sections.

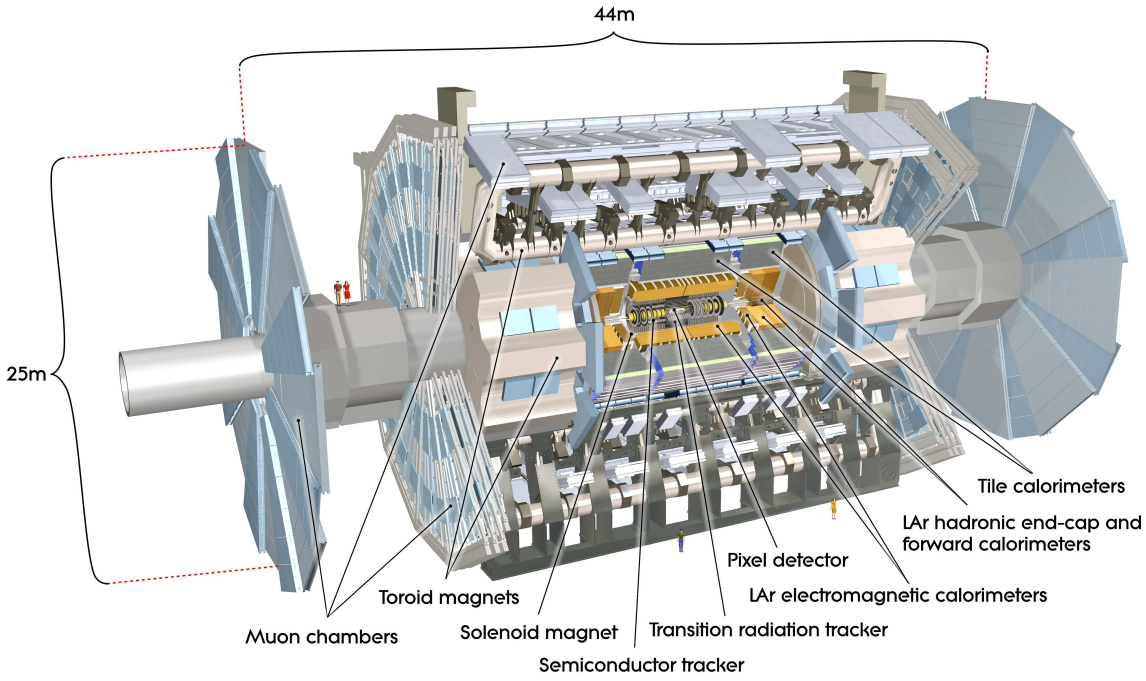


Figure 2.2: Cut-away view of the ATLAS detector with a labeling of the different components. Figure from [58].

The nominal interaction point is defined as the origin of the coordinate system, while the beam direction defines the z-axis and the x-y plane is transverse to the beam direction. The positive x-axis is defined as pointing from the interaction point to the center of the LHC ring and the positive y-axis is defined as pointing upwards. The azimuthal angle ϕ is the angle around the beam axis, and the polar angle θ is the angle from the beam axis. Instead of θ, the pseudo-rapidity η is used to express the polar position and is defined as:

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

The pseudo-rapidity is used in performance studies since it is directly related to a geometrical direction. In p-p collision, the two interacting partons have their momenta along the z-axis and which are different. The scattering system is thus boosted in the z-direction. The difference $\Delta\eta$ ($\Delta\theta$) is not Lorentz invariant. A new variable is defined, the rapidity y, which is Lorentz invariant under a z-direction boost:

$$y = \frac{1}{2}\ln\frac{E + p_z}{E - p_z},$$

where E is the jet energy and p_z is its momentum along the z-axis. As such, the rapidity is used in physics analysis. For a massless particle, y and η are identical.

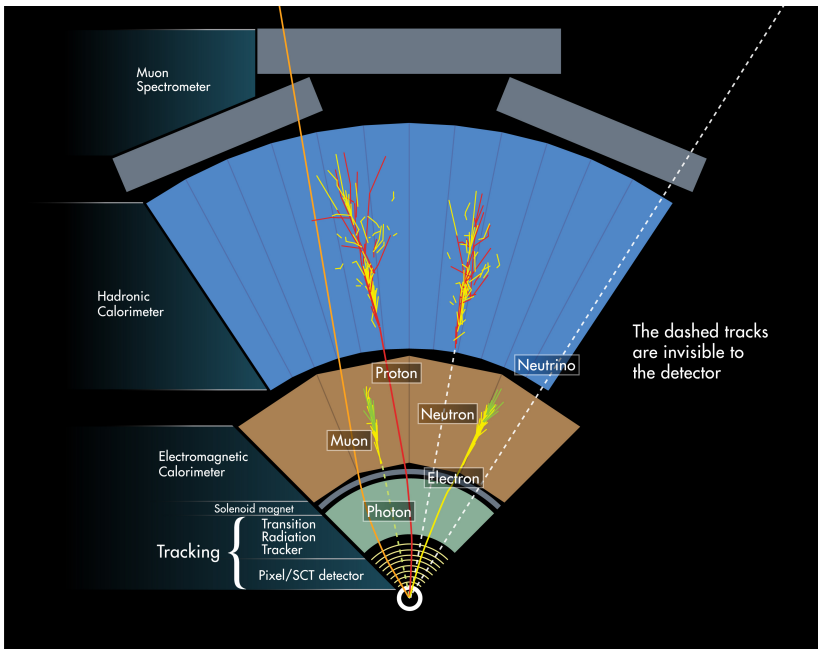


Figure 2.3: A transverse section of the ATLAS detector showing the tree main sub-systems: the tracker, the calorimeter and the muon spectrometer. The paths and interactions of different particles with the detector are shown. Figure from [59].

2.2.1 Inner Detector

The inner detector (ID) is used to reconstruct paths of charged particles (tracks). It is immersed in a 2 T magnetic field generated by the central solenoid, bending the charged particles while they pass through. The ID is able to give, among others, the following informations:

- the path of charged particles,
- the sign of the charge,
- the momentum of the particle,
- primary and secondary vertices related to the p-p collisions and the outgoing particle decays respectively,
- electron and photon identification (with the combined input of the electromagnetic calorimeter).

To meet those demands, the ID is built with high resolution capabilities and installed as close as possible to the beam pipe. The ID consists of three independent but complementary sub-detectors: silicon pixel, semi-conductor tracker (SCT) and transition radiation tracker (TRT).

The layout of the ID is illustrated in figure 2.4.

The precision tracking detectors (pixels and SCT) with high-resolution pattern recognition capabilities cover the region $|\eta| < 2.5$. They rely on the fact that a charged particle when passing through a semi-conductor produces electron-hole pairs which are read out and provide a “hit” information. In

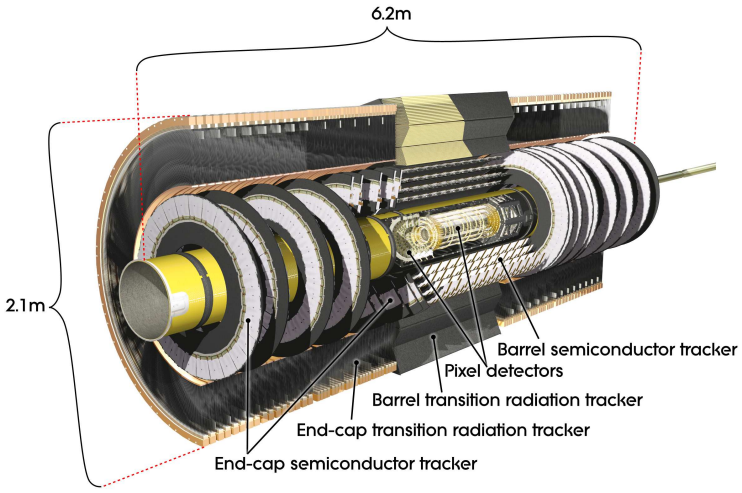


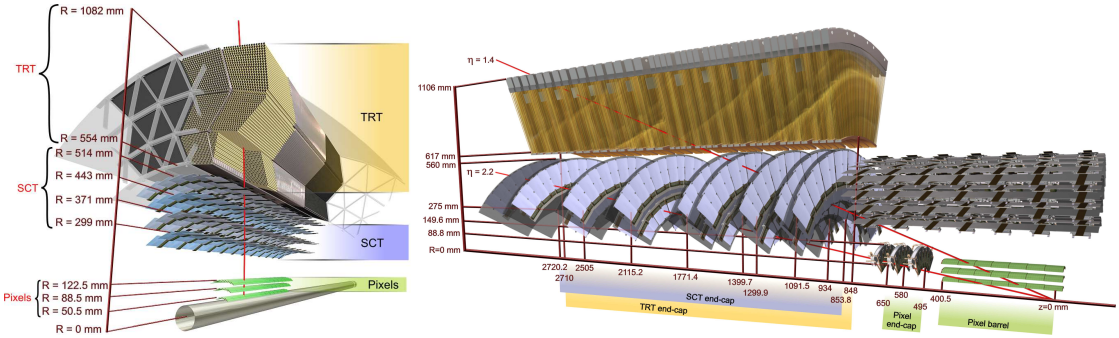
Figure 2.4: Cut-away view of the ATLAS Inner Detector. Figure from [58].

the barrel region, they are arranged on concentric cylinders around the beam axis, while in the end-cap regions they are located on disks perpendicular to the beam axis. To contain the semiconductors annealing and to reduce the leakage current, the sensors are operated in the temperature range -10° to -5°. The highest granularity is achieved around the vertex region using silicon pixel detectors. In addition to the three layers, a fourth layer, the IBL, was added before the beginning of Run II in an innermost position, improving tracking performance. Pixel sensors involve a double-sided processing and operate under voltages up to 600 V. Each track typically crosses three pixel layers. The nominal pixel size is 50×400 μm^2 with 250 μm thickness. For the IBL, the pixel size is 50×250 μm^2. The intrinsic accuracies in the barrel are 10×115 μm in ϕ-z and in the disks are 10×115 μm in ϕ-R.

For reasons of cost and reliability, SCT uses silicon micro-strip with classic single-sided p-in-n technology. Their operating voltages are between 250 and 350 V. For the SCT, typically eight strip layers are crossed by each track, equivalent to four space points. The sensor thickness is 285 μm. In the barrel region, this detector uses small-angle stereo strips to measure both coordinates, with one set of strips in each layer parallel to the beam direction, measuring ϕ. They consist of two 6.4 cm long daisy-chained sensors with a strip pitch of 80 μm. In the end-cap region, the detectors have a set of strips running radially and a set of stereo strips at an angle of 40 mrad. The mean pitch of the strips is also approximately 80 μm. The intrinsic accuracies per module in the barrel are 17×580 μm in ϕ-z and in the disks are 17×580 μm in ϕ-R.

The TRT is composed of layers of gaseous drift tubes (straw) interleaved with transition radiation material and covers the region $|\eta| < 2$. The gas mixture, Xe/CO2/O2, is ionized by the passage of charged particles and the charges are read out by tungsten anodes. The TRT only provides ϕ information, for which it has an intrinsic accuracy of 130 μm per straw. The tubes have 4 mm diameter. In the barrel region, the straws are parallel to the beam axis and are 144 cm long. In the end-cap region, the 37 cm long straws are arranged radially in wheels. Under normal operating conditions, the maximum electron collection time is $\approx$ 48 ns.

Figure 2.5 shows the sensors and structural elements traversed by a charged particle in the barrel and



(a) (b)

Figure 2.5: Drawing showing the sensors and structural elements traversed by a charged particle in the barrel ID in a) and in the end-cap ID in b). Figure from [58].

in the end-cap. The combination of precision trackers at small radii with the TRT at a larger radius gives very robust pattern recognition and high precision in both ϕ and z coordinates. The straw hits at the outer radius contribute significantly to the momentum measurement, since the lower precision per point compared to the silicon is compensated by the large number of measurements and longer measured track length.

2.2.2 Calorimeter

A calorimeter measures the energy a particle loses as it passes through. It is designed to stop/absorb most of the particles coming from a collision (all SM particles except muons and neutrinos), forcing them to deposit most of the time all of their energy within the detector. In dense material and at high energies, electrons and positrons predominantly lose energy by radiating photons (bremsstrahlung), while high energy photons exhibit e^+e^- pair conversions. The successive radiation and pair production is called electromagnetic shower. The characteristic amount of matter traversed is given in units of radiation length, X_0, which is the mean distance over which the electron energy becomes $1/e$ of its initial value and equal to 7/9 of the mean free path for pair production by a high energy photon. On the other hand, hadronic showers begin by interaction through the strong force between incident hadrons[2] and nuclei of the detector producing new hadrons, mainly pions, and exciting the nuclei (which can lose nucleons). The new hadrons, along with the initial ones, exhibit further interactions and thus building the shower. About 1/3 of the produced pions are π^0 which decay into 2γ and induce electromagnetic showers. The characteristic amount of matter traversed by hadrons is given in units of interaction length, λ, which is the mean distance traveled by a hadronic particle before undergoing an inelastic nuclear interaction. A detector usually does not have the same response to electromagnetic and hadronic showers. This is due to the existence of an invisible energy in the hadronic shower: the energy used to release protons and neutrons from calorimeter nuclei, and the kinetic energy carried by recoil nuclei do not lead to a calorimeter signal. Some detectors use compensation techniques to obtain a similar response for the two showers. The ATLAS calorimeter is a non-compensating one.

[2] It is useful to note that charged hadrons also induce ionizations in the detector material, but the fraction of energy lost in this process is much smaller than in the hadronic interactions.

The ATLAS calorimeters consist of a number of sampling detectors with full ϕ-symmetry and coverage. These calorimeters cover the range $|\eta| < 4.9$ and have enough thickness to contain electromagnetic and hadronic showers. This is needed to satisfy the physics requirements for electron, photon and jet reconstructions and E_T^{miss}[3] measurements. They also reduce punch-through[4] into the muon system well below the level of prompt or decay muons.

The ATLAS calorimeters are designed into sub-systems providing maximum absorption of electromagnetic or hadronic interactions. Also, they are split into 3 regions: the first consisting of a central electromagnetic and a hadronic tile barrels, the second of an electromagnetic end-cap calorimeter (EMEC) and a hadronic end-cap calorimeter (HEC), and the last of the forward detectors (FCal). All these calorimeters, except the tile, use liquid argon (LAr) as the active detector medium: charged particles ionize the LAr and the produced electrons-ions propagate to the electrodes and cathodes respectively due to the high electric field. Liquid argon has been chosen for its intrinsic linear behavior, its stability of response over time and its intrinsic radiation-hardness. A view of the calorimeters is presented in figure 2.6.

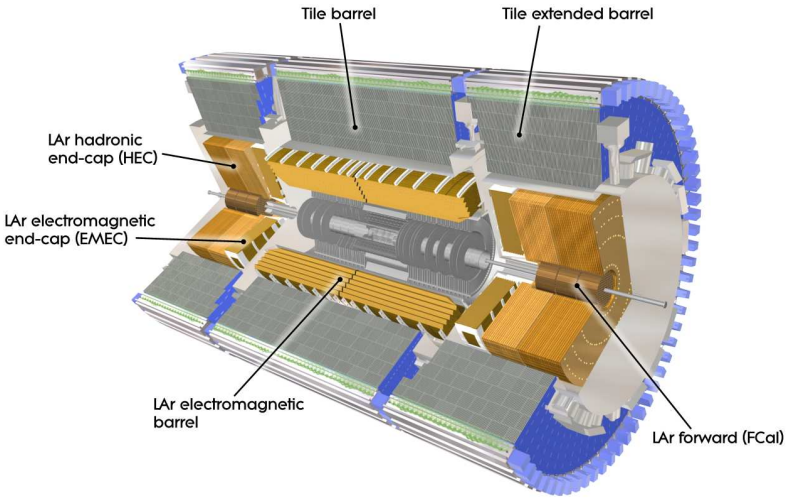


Figure 2.6: Cut-away view of the ATLAS calorimeter system. Figure from [58].

Electromagnetic barrel and end-cap calorimeters

The barrel part covers the range $|\eta| < 1.475$ while the two end-cap components cover the range $1.375 < |\eta| < 3.2$. Each end-cap calorimeter is mechanically divided into two coaxial wheels: outer and inner wheels split at $|\eta| = 2.5$. The EM calorimeter is a lead-LAr detector with accordion-shaped electrodes and lead absorber plates over its full coverage. The accordion geometry provides complete *phi* symmetry and coverage without azimuthal cracks and a fast extraction of the signal at the end of the electrodes.

[3] E_T^{miss} represents the unbalance in the sum of the measured transverse momenta of all detected particles in an event.
[4] Punch-through is the shower energy that is not fully contained in the calorimeters and reaches the muon spectrometer.

Over the η region matched to the inner detector, $|\eta| < 2.5$, the fine granularity of the EM barrel and the end-cap outer wheel is ideally suited for precision measurements of electrons and photons. They are segmented in three active layers in depth (two layers for the EMEC between $1.375 < |\eta| < 1.5$). An accurate position measurement is obtained by finely segmenting the first layer in η direction ($\Delta\eta = 0.0031$). The second layer collects the largest fraction of the energy of the electromagnetic shower with cell dimensions of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The third layer collects only the tail of the electromagnetic shower and is therefore less segmented in η. Figure 2.7 shows the three layers and their granularities in the barrel, with the accordion geometry also visible. For the end-cap inner wheel, the calorimeter is segmented in two layers in depth and has a coarser lateral granularity. Furthermore in the region $|\eta| < 1.8$, the electromagnetic calorimeters are complemented by presamplers, an instrumented argon layer, which provides a measurement of the energy lost in front of the electromagnetic calorimeters.

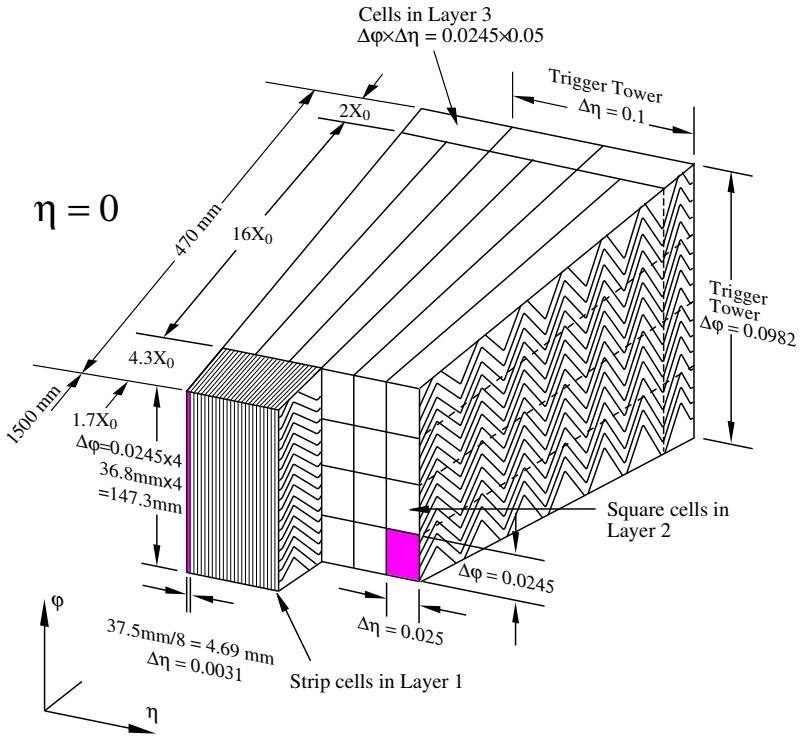
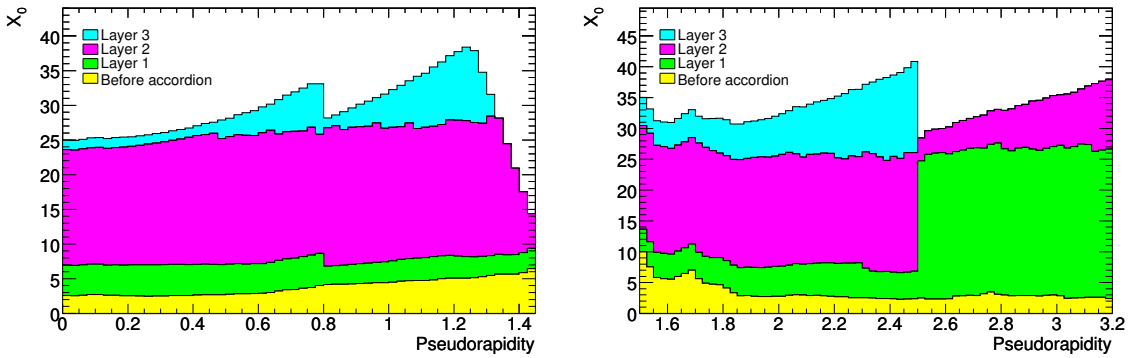


Figure 2.7: Sketch of a barrel module where the different layers and the accordion shape are visible. The granularity in η and ϕ of the cells of each of the three layers and of the trigger towers is also shown. Figure from [58].

The total active thickness of the barrel ranges between 22-33 X_0, whereas the thickness of the EMEC ranges between 24-38 X_0. Figure 2.8 shows the radiation length X_0 of the different layers in function of η separately for the barrel and the end-cap. In the barrel, the size of the drift gap on each side of the electrode is 2.1 mm, which corresponds to a total drift time of about 450 ns for its nominal operating voltage of 2000 V. The drift gap on each side of the electrodes is not constant for the EMEC. To obtain a uniform η-independent detector response, the high voltage varies in steps with η. Figure 2.9 shows an example of signal shapes produced in the detector and on the output of the barrel electromagnetic

module.



(a) (b)

Figure 2.8: Cumulative amounts of material, in units of radiation length X_0 as a function of $|\eta|$, in front of and in the electromagnetic calorimeters. The figures show separately the barrel (left) and end-cap (right) regions. Figures from [58].

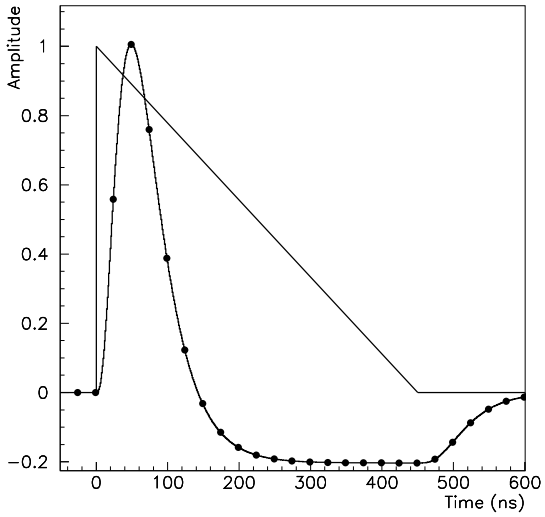


Figure 2.9: Amplitude versus time for triangular pulse of the current in a LAr barrel electromagnetic cell and of the output signal after bi-polar shaping. Also indicated are the sampling points every 25ns (which also correspond to the successive bunch crossing). Figure from [58].

Tile and HEC hadronic calorimeters

The tile calorimeter is a sampling calorimeter using steel as the absorber and plastic scintillator as the active medium. The choice of this technology provides maximum radial depth for the least cost for ATLAS. Ionizing particles crossing the tiles induce the production of ultraviolet scintillation

light in the base material (polystyrene) and this light is subsequently converted to visible light by wavelength-shifting fluors. The tiles are inserted into a plastic sleeve, which both protects the tile and improves the scintillation light yield due to its high reflectivity of 95%. Wavelength-shifting fibers placed in contact with the tile edges collect the scintillation light produced in the scintillators, convert it to a longer wavelength and transmit it to the photomultiplier (PMT). Fiber grouping is used to define a three-dimensional cell structure in such a way as to form three radial sampling layers. The fibers are read out by two different PMT's to provide redundancy and sufficient information to partially equalize signals produced by particles entering the scintillating tiles at different impact positions. Figure 2.10 shows a schematic of a tile module where the absorbers, the scintillators and the optical readout components are shown.

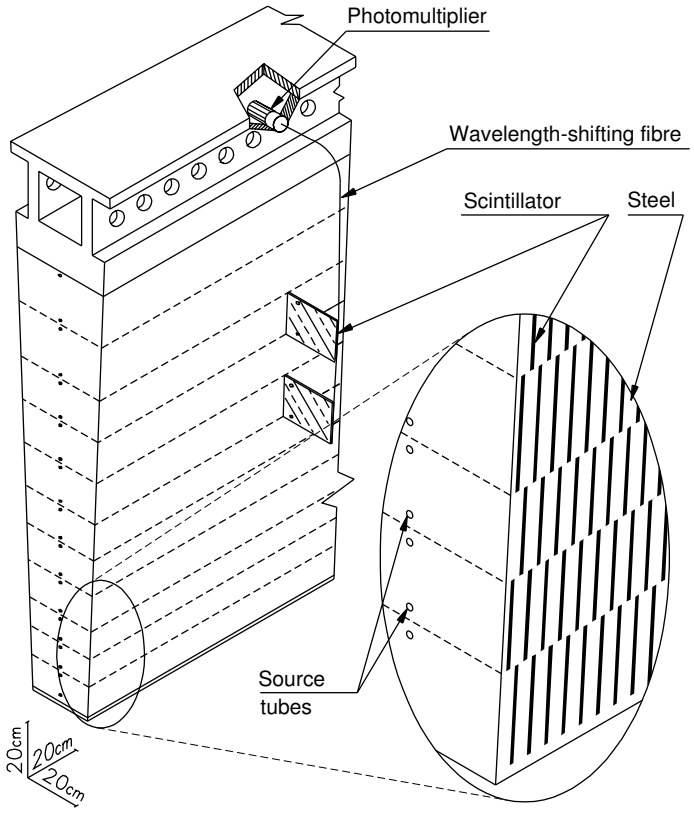


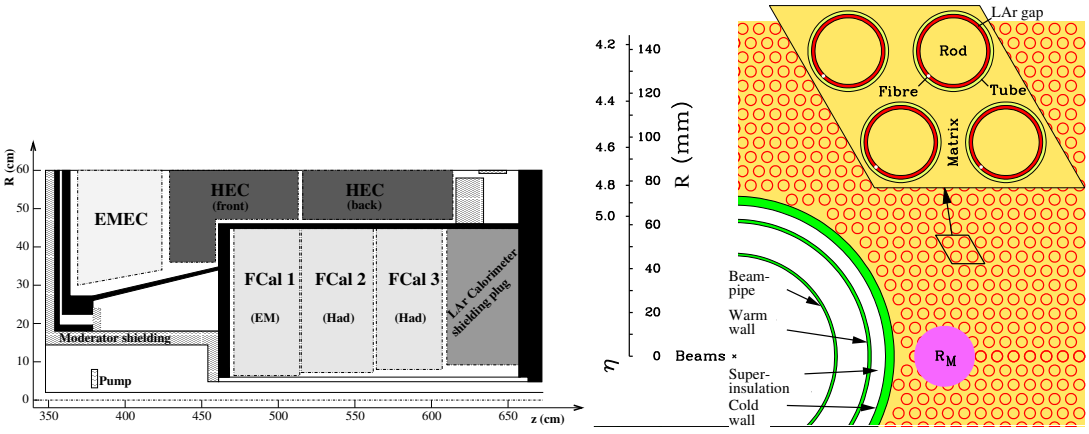
Figure 2.10: Schematic of a tile module where the steel absorbers, the plastic scintillators and the optical readout components (the fibers and the photomultipliers) are shown. Figure from [58].

The tile is subdivided into a barrel which covers the region $|\eta| < 1$, and two extended barrels in the range $0.8 < |\eta| < 1.7$. The radial depth of the tile calorimeter is approximately 7.4 λ. The assembled module forms an almost-periodic steel-scintillator structure with a ratio by volume of approximately 4.7:1. The orientation of the scintillator tiles radially and normal to the beam line allows for almost seamless azimuthal calorimeter coverage. Finally, the calorimeter is equipped with three calibration systems: charge injection, laser and a ^{137}Cs radioactive γ-source. These systems test the optical and digitized signals at various stages and are used to set the PMT gains to a uniformity of $\pm 3\%$.

The HEC is a copper-LAr sampling calorimeter with a flat-plate design covering the range $1.5 < |\eta| < 3.2$. It is subdivided into two wheels per end-cap. It extends out to $|\eta| = 1.5$ to overlap with the tile calorimeter and to $|\eta| = 3.2$ to overlap with the forward calorimeter. For the nominal high voltage of 1800 V, the typical drift time for electrons in the drift zone is 430 ns.

Forward calorimeter

The Fcal is also a sampling calorimeter using LAr as the active medium. It is split into three 45 cm deep modules: one electromagnetic module (FCal1) and two hadronic modules (FCal2 and FCal3). As the FCal modules are located at high η, they are exposed to high particle fluxes. This has resulted in a design with very small LAr gaps, which have been obtained by using an electrode structure of small-diameter rods, centered in tubes which are oriented parallel to the beam direction. The liquid-argon gaps are smaller than the usual 2 mm gap of the electromagnetic barrel calorimeter to avoid ion build-up problems and to provide at the same time the highest possible density needed due to their smaller length. These smaller gaps also lead to a faster signal with roughly the same instantaneous current but smaller integrated current. In the electromagnetic layer (FCal1), the triangular current pulse at the electrode has a full drift time of 60 ns. For FCal2 and FCal3, the full drift time scales with the gap size. To optimise the resolution and the heat removal, copper was chosen as the absorber for FCal1, while mainly tungsten was used in FCal2 and FCal3, to provide high absorption length to increase containment and minimize the lateral spread of hadronic showers. Figure 2.11 shows a schematic diagram of the three FCal modules on the left and the electrode structure of FCal1 on the right.



(a) (b)

Figure 2.11: a) Schematic diagram showing the three FCal modules. b) Electrode structure of FCal1 with the matrix of copper plates and the copper tubes and rods with the LAr gap for the electrodes. Figures from [58].

Figure 2.12 shows the interaction length λ in the different regions of the electromagnetic and hadronic calorimeters. The figure shows that hadrons deposit a fraction of their energies in the EM calorimeter but are not fully contained by them (except for π^0 which decays predominantly to 2γ).

In addition to reconstructing the deposited energies, the calorimeter is able, along with the information from the ID, to identify electrons, photons, charged and neutral hadrons.

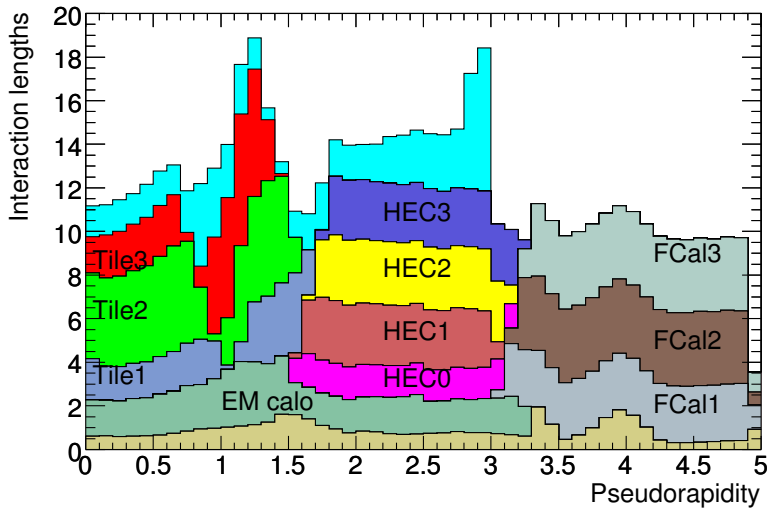


Figure 2.12: Cumulative amount of material, in units of interaction length λ, as a function of $|\eta|$, in front of the electromagnetic calorimeters, in the electromagnetic calorimeters themselves, in each hadronic sub-detector, and the total amount at the end of the active calorimetry. Also shown for completeness is the total amount of material in front of the first active layer of the muon spectrometer (up to $|\eta| < 3.0$). Figure from [58].

Pile-up sensitivity

Pile-up is the energy contribution from additional p-p collisions to the one we are studying. The additional collisions come predominantly from QCD interactions at low energies. Two type of pile-up can be identified. The first is the in-time pile-up which is due to multiple p-p collisions in the same bunch-crossing. The second is the out-of-time pile-up due to the long charge collection time in the LAr calorimeter (400-500 of ns) compared to the 25 ns bunch-crossing interval. The LAr calorimeter is more sensitive to the pile-up compared to the tile calorimeter mainly because it absorbs most of the low energy particles coming from the additional low energy collisions. A bipolar shaped signal with net zero integral over time is used to read the calorimeters which reduces the sensitivity to pile-up. The jet calibration, discussed in section 2.3.3, corrects for the remaining sensitivity.

2.2.3 Muon spectrometer

The muon spectrometer (MS) forms the outer part of the ATLAS detector and is designed to reconstruct tracks of charged particles exiting the barrel and end-cap calorimeters, to identify muons and to measure their momentum in the range $|\eta| < 2.7$. Is is also used to detect punch-through signals: showers inside calorimeters are sometimes not fully contained in the calorimeter and the charged particles that escape it induce hits in the MS. In this case, the energy detected by the MS is used to complement the ones from the calorimeter. The MS is also designed to trigger on muons in the region $|\eta| < 2.4$. It uses separate instrumentation for the high-precision tracking and trigger chambers. The barrel toroid provides magnetic bending over the range $|\eta| < 1.6$, whereas end-cap magnets provide it over the range $1.4 < |\eta| < 2.7$ where the overlap region is $1.4 < |\eta| < 1.6$. The driving performance goal is a stand-alone transverse momentum resolution of approximately 10% for 1 TeV tracks. The

barrel region is divided into eight octants symmetric in ϕ. Each octant is subdivided in the azimuthal direction in two sectors, a large and a small one overlap in ϕ. This overlap of the chamber boundaries minimizes gaps in detector coverage and also allows for the relative alignment of adjacent sectors using tracks recorded by both a large and a small chambers. The chambers in the barrel are arranged in three concentric cylindrical shells around the beam axis. In the two end-cap regions, muon chambers form large wheels, installed in planes perpendicular to the beam to the z-axis and also arranged in three layers. A view of the muon spectrometer is presented in figure 2.13.

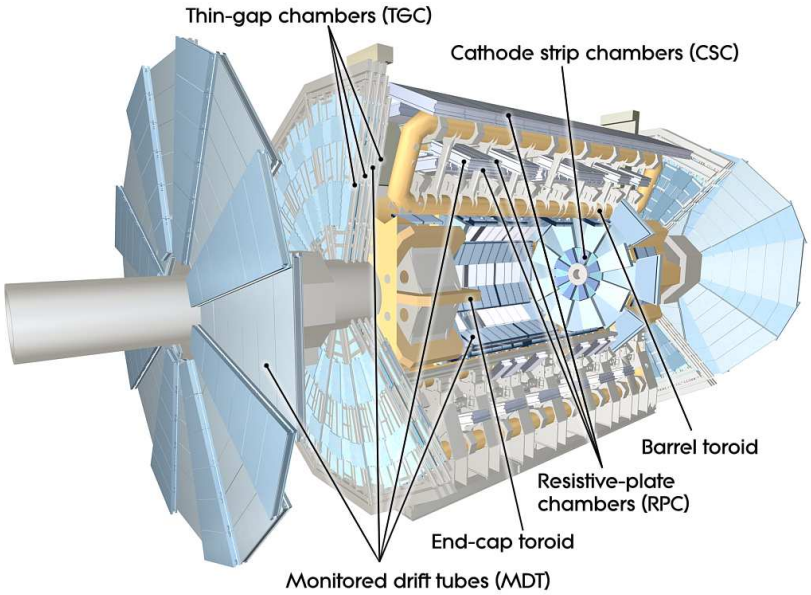
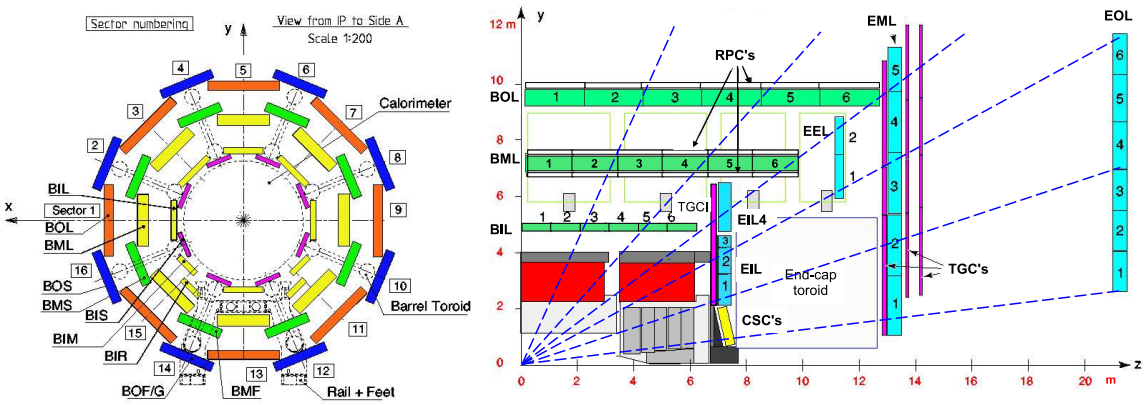


Figure 2.13: Cut-away view of the muon spectrometer. Figure from [58].

The precision momentum measurement is performed by the monitored drift tube (MDT) chambers. They cover the range $|\eta| < 2.7$, except in the innermost end-cap layer where their coverage is limited to $|\eta| < 2.0$. The MDTs are pressurized drift tubes operating with Ar/CO2 gas mixture (93/7) at 3 bar. The electrons resulting from ionization are collected at the central tungsten-rhenium wire at a potential of 3080 V. One advantage of the cylindrical geometry is that the measurement accuracy depends only weakly on the angle of incidence of the track onto the chamber plane. A disadvantage is the long pulse train. The maximum drift time from the wall to the wire is about 700 ns. These chambers consist of three to eight layers of MDTs, which achieve an average resolution of 80 μm per tube, or about 35 μm per chamber. The particle fluxes and muon-track density are highest in the forward direction. Therefor, in the region $2 < |\eta| < 2.7$, cathode-strip chambers (CSC) are used in the inner-most tracking layer due to their higher rate capability and time resolution. The CSCs are multi-wire proportional chambers with cathode planes segmented into strips in orthogonal directions. This allows both coordinates to be measured from the induced-charge distribution on neighboring cathode strips. The resolution of a chamber is 40 μm in the bending (η) plane and about 5 mm in the transverse (ϕ) plane. The locations of MDT wires and CSC strips along a muon trajectory must be known to better than 30 μm. A high-precision optical alignment system monitors the positions and internal deformations of the MDT chambers, complemented by track-based alignment algorithms

using high-momentum muons. Figure 2.14 shows a transverse and a longitudinal sections, where the three layers arrangement in the barrel and the end-cap can be seen.

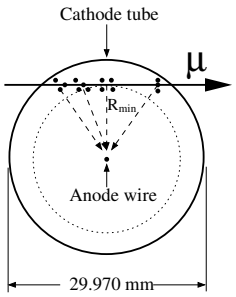


(a) (b)

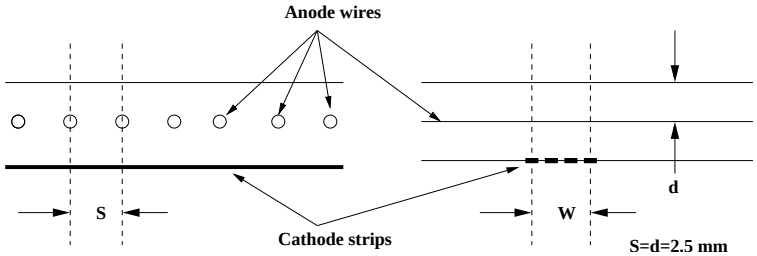
Figure 2.14: a) Section of the barrel muon system perpendicular to the beam axis (non-bending ϕplane), showing the three concentric cylindrical layers composed of eight large and eight small chambers symmetric in ϕ. b) Section of the muon system in a plane containing the beam axis (bending η plane) showing the layers in the barrel and end-cap regions. MDTs are shown in green in the barrel and cyan in the end-cap, while CSC are shown in yellow. Figures from [58].

Another essential property of the MS is the capability to trigger on muon tracks. The precision-tracking chambers have therefore been complemented by a system of fast trigger chambers capable of delivering track information within a few tens of nanoseconds after the passage of the particle. In addition to muons multiplicity and approximate energies, the trigger also provides bunch-crossing identification and a second ϕ coordinate measurement to complement the MDT one. In the barrel region ($|\eta| < 1.05$), resistive plate chambers (RPC) are used due to good spatial and time resolution as well as adequate rate capability. The RPC is a gaseous parallel electrode-plate (no wire) detector. The two resistive plates are kept parallel to each other at a distance of 2 mm. The electric field between the plates of about 4.9 kV/mm allows avalanches to form along the ionizing tracks towards the anode. At the nominal operating voltage of 9.8 kV, a signal with a width of about 5 ns is generated by the track. In the end-cap region ($1.05 < |\eta| < 2.4$), thin gap chambers (TGC) are chosen providing good time resolution, high rate capability and robustness. TGCs are multi-wire proportional chambers with the characteristic that the wire-to-cathode distance of 1.4 mm is smaller than the wire-to-wire distance of 1.8 mm. With a highly quenching gas mixture of CO2 and n-pentane, this cell geometry allows for operation in a quasi-saturated mode. The high electric field around the TGC wires and the small wire-to-wire distance lead to very good time resolution for the large majority of the tracks, with signals arriving inside a time window of 25 ns more than 99% of the time.

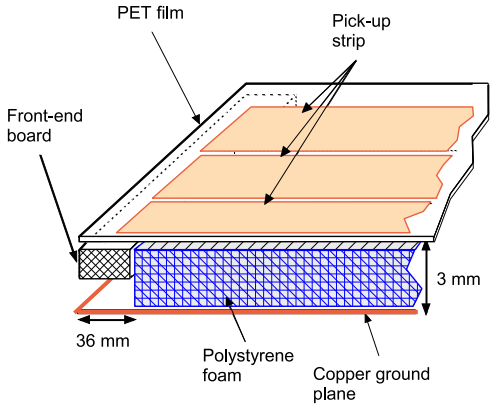
Figure 2.15 shows schematics of the different technologies used in the different MS chambers described above: MDT, CSC, RPC and TGC.



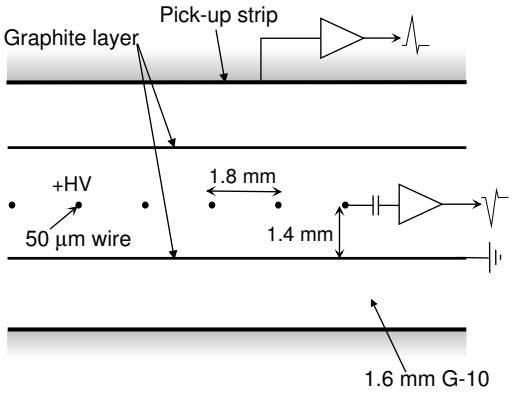
(a) Monitored drift tube (MDT)



(b) Cathode-strip chambers (CSC)



(c) Resistive plate chambers (RPC)



(d) Thin gap chambers (TGC)

Figure 2.15: Schematics of the different technologies used in the different MS chambers. Figures from [58].

2.2.4 Forward detectors

The ATLAS detector contains, in addition to the three main detectors described in the previous sections, multiple detectors in the very forward region. I focus here on some of these detectors that are used to measure the luminosity. A precise measurement of the integrated luminosity is a key component of physics analyses, whether for cross-section measurements or searches of new physics. To convert between the instantaneous luminosity and the instantaneous mean number of interactions per bunch crossing μ^{inst}, a reference inelastic cross-section of $\sigma_{\text{inel}} = 80$ mb is used. The average quantity, $\mu =< \mu^{\text{inst}} >$, over lumi-blocks (usually 1 minute duration) is generally used to characterize the pile-up condition.

LUCID detector

LUCID stands for LUminosity measurement using Cerenkov Integrating Detector. It is located at about 17 m from the ATLAS interaction point. Its main purpose is to detect inelastic p-p scattering in the forward direction, in order to both measure the integrated luminosity and to provide online

monitoring of the instantaneous luminosity and beam conditions. LUCID is based on the principle that the number of interactions in a bunch-crossing is proportional to the number of particles detected in this detector. Cerenkov light is emitted by a charged particle traversing the detector and is then detected by a PMT. During Run I, aluminum tubes filled with C4F10 gas were used to induce the Cerenkov light. For Run II, the quartz windows of the PMTs induce the Cerenkov light. The signal amplitude from these PMTs can be used to distinguish the number of particles per tube, and the fast timing response provides measurements of individual bunch crossings.

ALFA detector

ALFA stands for Absolute Luminosity For ATLAS. It is located at about 240 m from the ATLAS interaction point. For luminosity measurements, the detector uses the optical theorem which connects the total cross-section to the elastic-scattering amplitude at small angles. Special runs with high β^*[5] are used to get extremely small scattering angles (3 μrad) needed for these measurements. ALFA uses a scintillating-fiber trackers connected to PMTs. It is useful to note that the detector is also used for other physics measurements.

Other methods are also used to further minimize the luminosity uncertainties, for example from offline track counting measurements. After typical data-quality selections, the full Run II p-p collisions corresponds to an integrated luminosity of 139 fb^{-1}, with an uncertainty of 1.7%. Figure 2.16 shows the total integrated luminosity during Run II. It also shows the luminosity delivered by LHC, which shows that the ATLAS experiment is very efficient at exploiting the stable beam time.

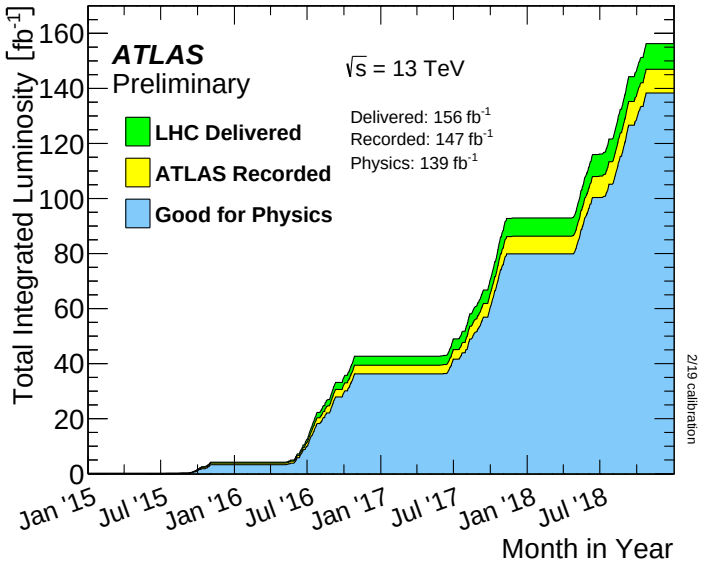


Figure 2.16: Cumulative luminosity versus time delivered to ATLAS (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams for pp collisions at 13 TeV center-of-mass energy in LHC Run II. Figures from the ATLAS public plots page [60].

[5] β^* is a parameter which describes how much the beams are squeezed: larger values mean less squeezed.

Figure 2.17 shows the luminosity-weighted distribution of the mean number of interactions per bunch crossing μ for the different data taking years and for the total Run II.

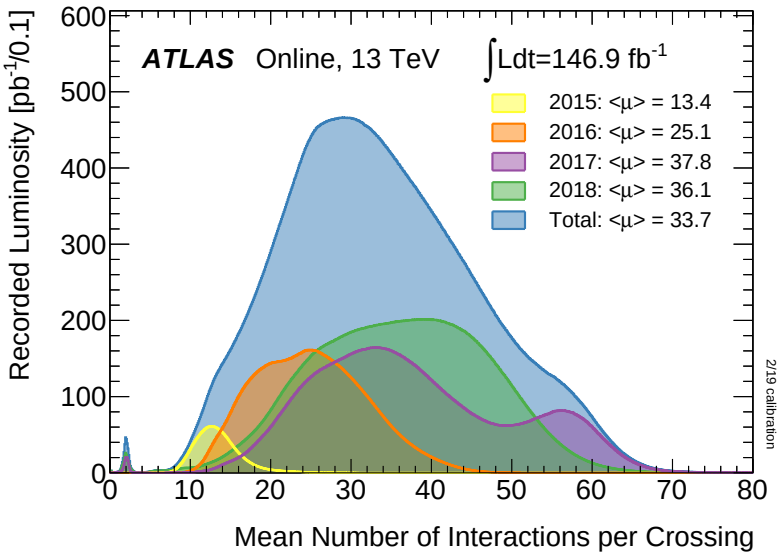


Figure 2.17: Luminosity-weighted distribution of the mean number of interactions per bunch crossing μ for the different data taking years and for the total Run II. The average μ for each data taking year is also given. Figures from the ATLAS public plots page [60].

2.2.5 Trigger

Bunch crosses the ATLAS detector at a rate of 40 MHz yielding about 1 billion collision per second (for $< \mu >= 25$). If for each crossing the event will be saved and since each event stores more than 1.5 MB of raw data, the data recording rate is more than 60TB/s which is not possible for the available processing, networking and data storing systems. In addition, most events come from QCD interactions and give low p_T jets. The high p_T range or the other scattering processes, which have much lower cross-sections, are more of interest for us. For example at 13 TeV, Higgs boson production happens less than 1 per ten billion events, particles from new physics are expected to be produced at an even smaller rate. Therefore, a trigger system is an essential component of the detector selecting interesting events and prescaling[6] others. The ATLAS trigger system has two levels: a hardware-based level 1 (L1) and a software-based high level trigger (HLT) which reduces the rate from 40 MHz to about 100 kHz and then to about 1 kHz respectively. Figure 2.18 shows a schematic of the ATLAS trigger.

The L1 trigger performs the initial event selection using a limited amount of the total detector information to make a decision in ≈ 2.5 μs. It searches for high transverse-momentum electrons, muons, taus (decaying hadronically), photons and jets, as well as large missing transverse energy

[6] Select only a defined percentage of events with same criteria.

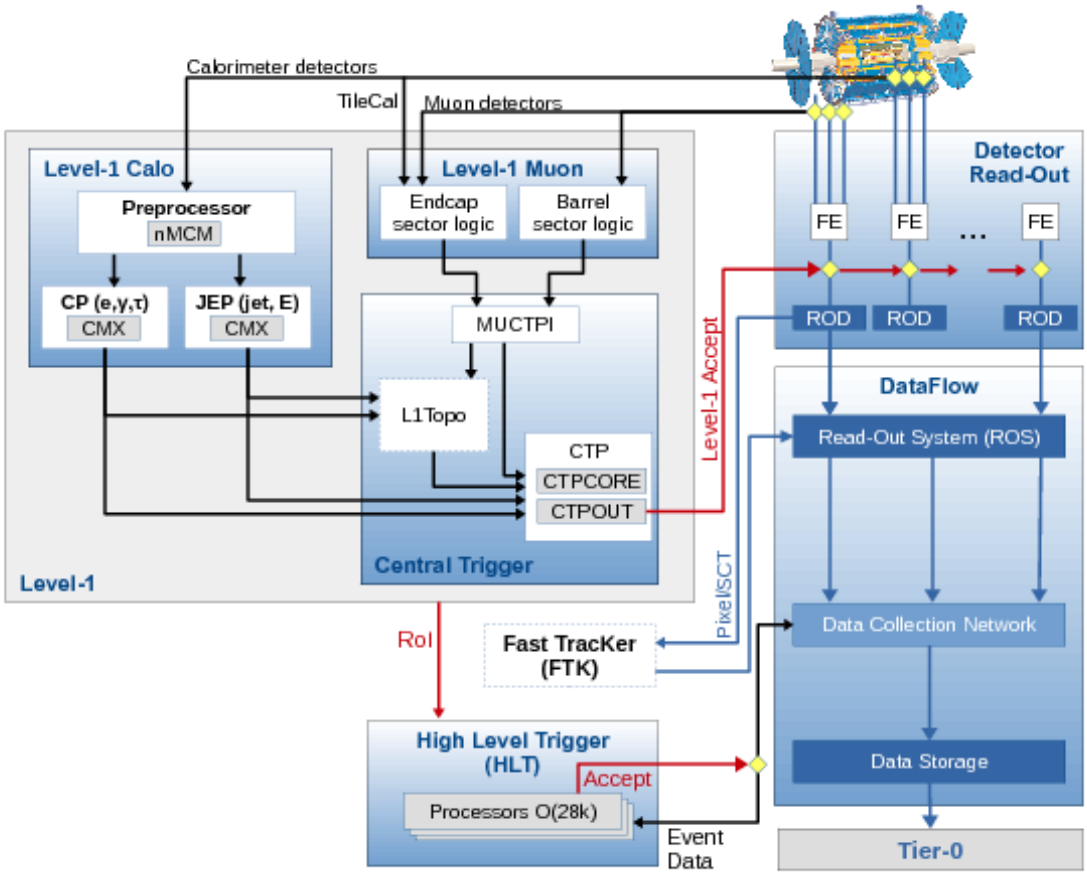


Figure 2.18: The ATLAS trigger system in Run II (also shown part of the data acquisition system). Figure from [61].

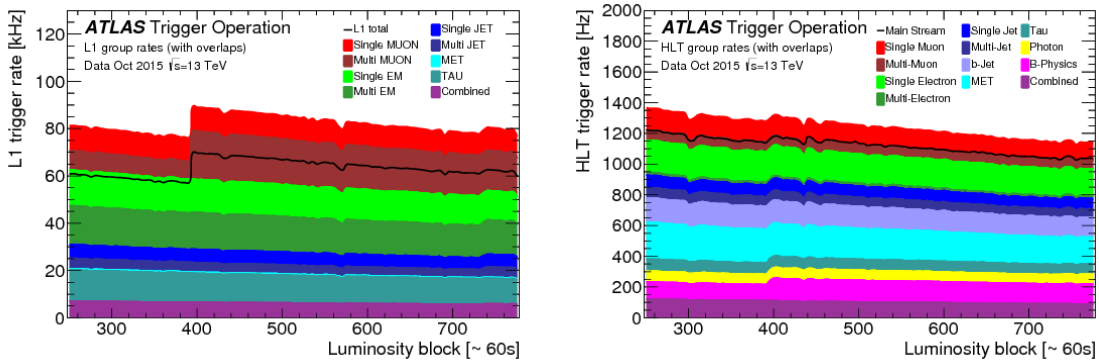
(E_T^{miss}) and large total transverse momentum. An essential function of the L1 trigger is a precise identification of the bunch-crossing of interest, overcoming the challenges coming from the physical size of the muon spectrometer (which implies times-of-flight exceeding the bunch-crossing interval) and from the fact that the width of the LAr calorimeter signal extends over many (typically four) bunch-crossings. From the muon spectrometer, the L1 trigger uses only the information from the trigger chambers. In addition, informations from all the calorimeters are used but with a reduced granularity. A trigger menu made up of combinations of trigger selections is used to select interesting processes with particular particle multiplicities, minimum transverse momentum, etc. The L1 trigger defines one or more regions-of-interest (RoI) (η and ϕ coordinates) where interesting features have been identified. This information is then passed and used by the HLT trigger. For jets, trigger towers are used which have coarser granularity than the calorimeter (0.1×0.1 in $\Delta\eta \times \Delta\phi$ mostly, but wider at high $|\eta|$) (see figure 2.7). Jet trigger elements are defined using 2×2 trigger towers. The jet sliding window algorithm identifies E_T sums within overlapping windows consisting of 2×2, 3×3, or 4×4 jet elements. These sums are then compared to predefined jet energy thresholds and the location of

a 2×2 local maximum defines the coordinates of the jet RoI.

The HLT trigger uses most of the information from the detector with finer granularities than L1 level. It also uses information from the tracker using the fast tracker algorithm (FTK). It minimize the differences with the offline analysis which improves the selection efficiency. Depending on the trigger selection, the HLT trigger can check for the required objects only within regions selected by the RoI or in all the detector regions. The latency of the HLT trigger is in the order of 200 ms. For jets, the reconstruction used is the same as the offline one (see next section). The calibration also includes most of the steps used offline. It is useful to note that the HLT calo trigger integrates detector information from multiple bunch-crossing same as offline objects (since the detector response is slower than the time separating two bunch-crossing).

Figure 2.19 shows L1 and HLT trigger rates grouped by trigger signature during an LHC fill. In this thesis, single jet triggers are used which, using few examples, have the following nomenclatures and requirements:

- HLT_j15 (HLT_j25,HLT_j35): trigger which is random at L1 level (only needs to pass the prescaling) and has at least one HLT jet with $p_{\mathrm{T}} > 15$ (25, 35) GeV and $\eta < 3.2$,
- HLT_j45: trigger passing the L1 requirement of at least one jet RoI with $\mathrm{E}_{\mathrm{T}} > 15$GeV and $\eta < 3.1$, and passing the HLT requirement of at least one jet with $p_{\mathrm{T}} > 45$GeV and $\eta < 3.2$,
- HLT_j45_320eta490: same as HLT_j45 but with the region requirement $3.1 < \eta < 4.9$ for L1 level and $3.2 < \eta < 4.9$ for HLT level.



(a) L1 trigger

(b) HLT trigger

Figure 2.19: L1 and HLT trigger rates grouped by trigger signature during an LHC fill in October 2015 with a peak luminosity of $4.5 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Due to overlaps, the sum of the individual groups is higher than the total rates which are shown as black lines. Figures from [61].

2.2.6 Object reconstruction

Figure 2.20 shows an example of an event produced in ATLAS, where tracks in the ID are shown in red and orange, EM and hadronic calorimeter energy deposits are shown in green and yellow

respectively, with the magnitude of the deposits reflected in the size of the colored shapes. Although the event is very busy, two high-energy back-to-back objects can be identified.

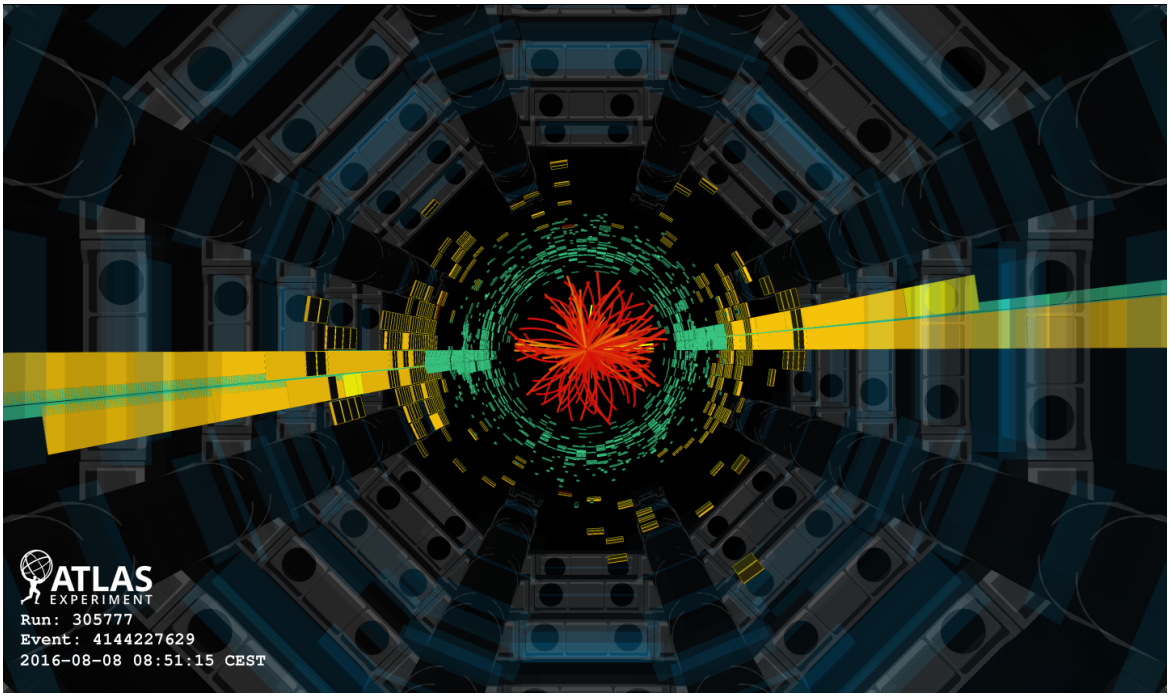


Figure 2.20: An example of an event produced in ATLAS, where tracks in the ID are shown in red and orange, EM and hadronic calorimeter energy deposits are shown in green and yellow respectively, with the magnitude of the deposits reflected in the size of the colored shapes.

The ATLAS sub-systems, ID, calorimeter and MS, are used individually or together to define several interesting objects to be used in physics analysis. Let me detail some of those objects.

Primary vertices

A primary vertex (PV) is the point of the collision of two protons. As already mentioned, multiple p-p collisions are produced in the same bunch-crossing. Our goal is to identify the hardest scattering, to be used in analyses, from the rest of the soft collisions which form the in-time pile-up. For jets analyses, tracks with p_{T}>0.5 GeV and passing quality criteria are associated to different primary vertices for which the positions are determined by tracks fitting. The vertex with the highest $\sum_{\mathrm{tracks}} p_{\mathrm{T}}^2$ is considered the hard-scatter primary vertex (PV_0) and the others as primary vertices due to in-time pileup.

Electrons and photons

Electron and photon reconstruction starts by building clusters out of the energy deposits in the EM calorimeter. The EM calorimeter, using all three layers in depth, is divided into towers (summing the energies in all layers) of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. A sliding-window algorithm with windows of

size 3 in $\eta - \phi$ space is used to search for EM clusters the shower shapes of which are consistent with an EM shower.

For an electron, a matching between at least one track in the ID and the EM cluster is required. Tracks are also tested for pattern recognition to distinguish pion and electron tracks. Electrons and positrons are distinguished by the curvature direction of their tracks.

Photons can interact with detector material before reaching the calorimeter and dominantly decay into a pair of e^+e^-. A converted photon is identified from an EM cluster matched to double tracks coming from a conversion vertex consistent with a massless particle, or matched to a single conversion track not having hits in the innermost sensitive layers. EM clusters with no matched track are identified as unconverted photons.

Muons

Although muons escape the detector, they do leave traces of their passage in the different sub-detectors. In addition to tracks is the ID and MS, high-energy muons deposit energies of few GeV in the calorimeter. Hence, to identify a muon, a track in the MS can be used alone or combined with a track in the ID and also with energy deposits in the calorimeter. MS tracks are also extrapolated inward and tested on compatibility with originating from the interaction point.

In this thesis, jets are the main objects used for performance studies and physics analysis. Their reconstruction and calibration are presented into detail in the next section.

2.3 Jet reconstruction and calibration

2.3.1 Jet constituents and reconstructions

As already mentioned in section 1.2.3, the most widely algorithm used in high energy physics, and hence ATLAS, is the anti-k_t algorithm. The algorithm can form jets using various type of inputs, where the only needed constituent variables are the four-vector momenta. Let us detail some of the inputs used in ATLAS.

Topological clusters
The lateral and longitudinal segmentation of the calorimeters allows for a three-dimensional reconstruction of particle showers. To take advantage of this segmentation, topo-clusters are built from topologically connected calorimeter cells. Topo-clusters are seeded by cells whose signals exceed the expected noise by four times its standard deviation, $S > 4\sigma_{\text{noise}}$. Neighboring cells with $S > 2\sigma_{\text{noise}}$ are then added iteratively. Finally, all cells neighboring the formed topo-cluster are added. Hence, the topo-cluster algorithm separates continuous energy showers rather than energy deposits from different particles. It efficiently suppresses the calorimeter noise which originates from both electronic and pile-up sources.

The cells used are initially calibrated to the electromagnetic scale (EM scale) which correctly measure the response of electromagnetic shower. The EM scale is derived from test beams measurements and from MC simulations. The jets built from topo-clusters at EM scale are called EMTopo jets. Hadronic showers produce responses that are lower than the EM scale due to the non-compensating nature of the ATLAS detector. A second topo-cluster collection tries to correct for the hadronic response by classifying clusters as either electromagnetic or hadronic (primarily based on the energy density and the longitudinal shower depth), and applying local cluster weights accordingly. The LCW clusters are not detailed further more since they are not used in this thesis.

Particle flow objects
This algorithm tries to improve on top of the EM topo-cluster algorithm by including the information from tracks [62]. A cell-based energy subtraction algorithm is employed to substitute the energy measured in the detector by the momentum of the overlapping track. The different steps of the subtraction algorithm, showed in figure 2.21, are as follows:

- first, well-measured tracks are selected following a quality criteria (minimum number of hits, $0.5 < p_{\text{T}} < 40$ GeV, track not matched as electron or muon),
- the algorithm then attempts to match each track to a single topo-cluster in the calorimeter: a geometrical distance weighted by the inverse of cluster dimensions, $\Delta R' = \sqrt{(\Delta\phi/\sigma_\phi)^2 + (\Delta\eta/\sigma_\eta)^2}$ where σ_ϕ and σ_η represent the angular topo-cluster widths, is used to search for the closest cluster to each track (ordered in a descending $p_{\text{T}}^{\text{track}}$ order) with a requirement on the cluster to have $E^{\text{cluster}}/p^{\text{track}} > 0.1$; if no matching is found with $\Delta R' < 1.64$, no subtraction is performed,

- the expected energy to be deposited in the calorimeter, $\langle E_{dep} \rangle$, by the particle that created the track is computed based on the topo-cluster position and the track momentum,
- the algorithm also evaluates the probability that the particle energy was deposited in more than one cluster and decides if more topo-clusters are to be added to the matched one,
- then the energy subtraction is performed on the set of matched clusters: if $\langle E_{dep} \rangle$ exceeds the total energy of the set, the whole clusters are removed, otherwise the subtraction is performed cell by cell,
- finally, if the remaining energy in the system is consistent with the expected shower fluctuations of a single particle's signal, the topo-cluster remnants are removed.

Together, the set of selected tracks and the remaining topo-clusters in the calorimeter are fed to the jet finding algorithm to construct what is called PFlow jets.

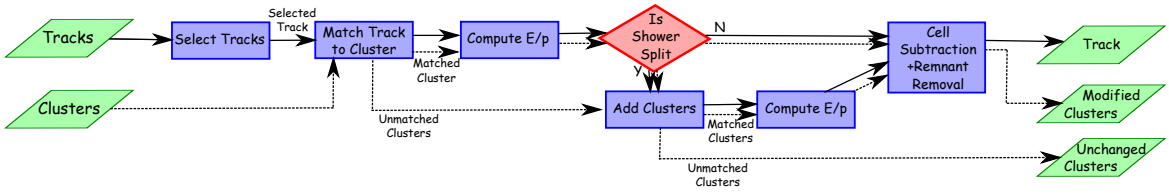


Figure 2.21: A flow chart of the energy subtraction algorithm. Figure from [62].

Figure 2.22 illustrates how the subtraction procedure is designed to deal with different cases, for example when the energy deposited by one particle can be splited into two clusters or when two particles deposited energies are constructed in the same clusters.

Truth jets

Truth jets are reconstructed from stable simulated particles: stable hadrons, electrons, photons, muons and neutrinos. When truth jets are used in physics analyses, muons and neutrinos are also considered among the jet constituents since they are products of particle decays and our aim is to recover the particles level. On the other hand, when doing calorimeter performance studies, muons and neutrinos are excluded since they don't (or very weakly) interact with the calorimeter.

With the exception of truth jets, the different jet collections need further calibration to correct their energies and momenta from multiple effects that are detailed later. The results are shown only for EMTopo and EMPFlow jets reconstructed using the anti-k_t algorithm with the radius parameter R=0.4 (those two collections are the only collections used in this thesis). But first, let us look at some jet quality selection criteria.

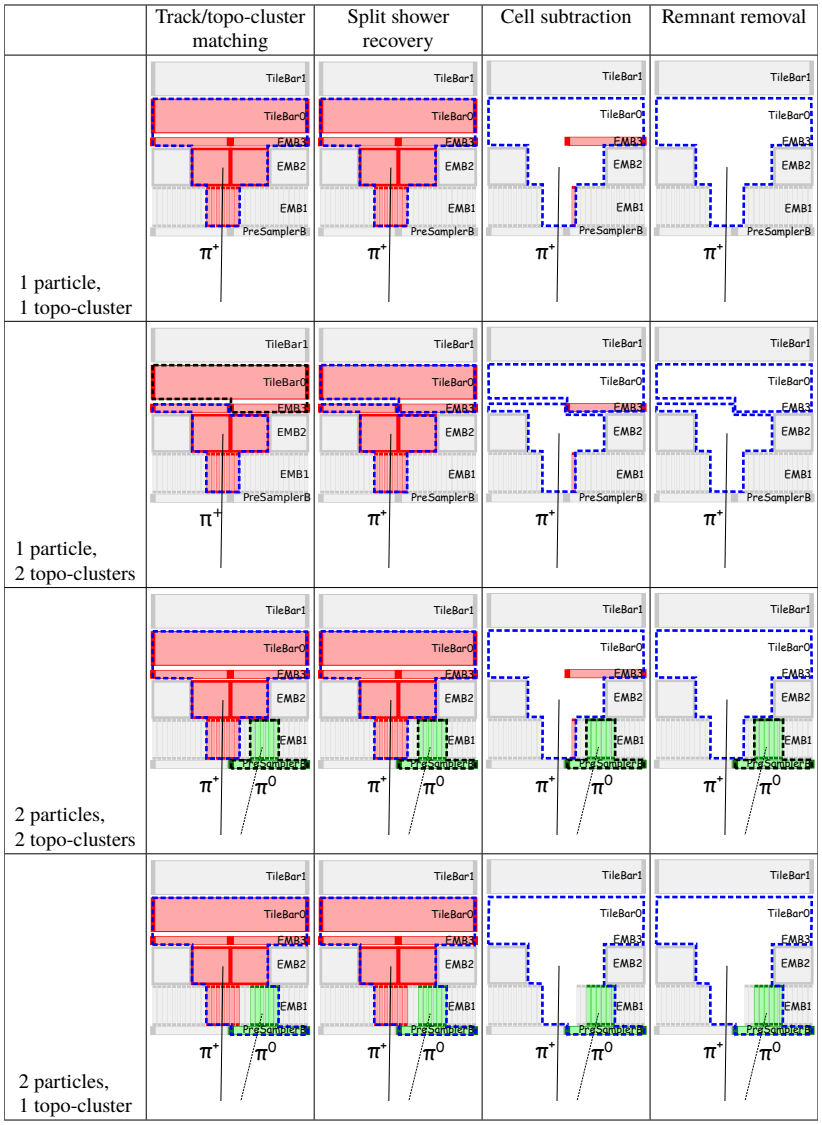


Figure 2.22: Examples of how the particle flow algorithm is designed to deal with several different clusters/track cases. Figure from [62].

2.3.2 Quality selection

Energy contributions from pile-up, in-time and out-of-time, are subtracted on average from hard-scatter jets as is detailed in section 2.3.3.1. However, local fluctuations in the pile-up activity can produce pile-up jets. The jet-vertex-tagger [63] (JVT) is used to discriminate if a jet originates from the hard-scattering (signal) or from pile-up. The JVT is a 2-dimensional likelihood constructed from

the variables corrJVF and R_{p_T}, giving the relative probability for a jet to be of signal type, where:

$$\begin{aligned} \text{corrJVF} &= \frac{\sum p_T^{trk}(PV_0)}{\sum p_T^{trk}(PV_0) + \frac{\sum_{(n\geq 1)} \sum p_T^{trk}(PV_n)}{k \times n_{trk}^{PU}}} \\ R_{p_T} &= \frac{\sum p_T^{trk}(PV_0)}{p_T^{jet}} \end{aligned} \tag{2.1}$$

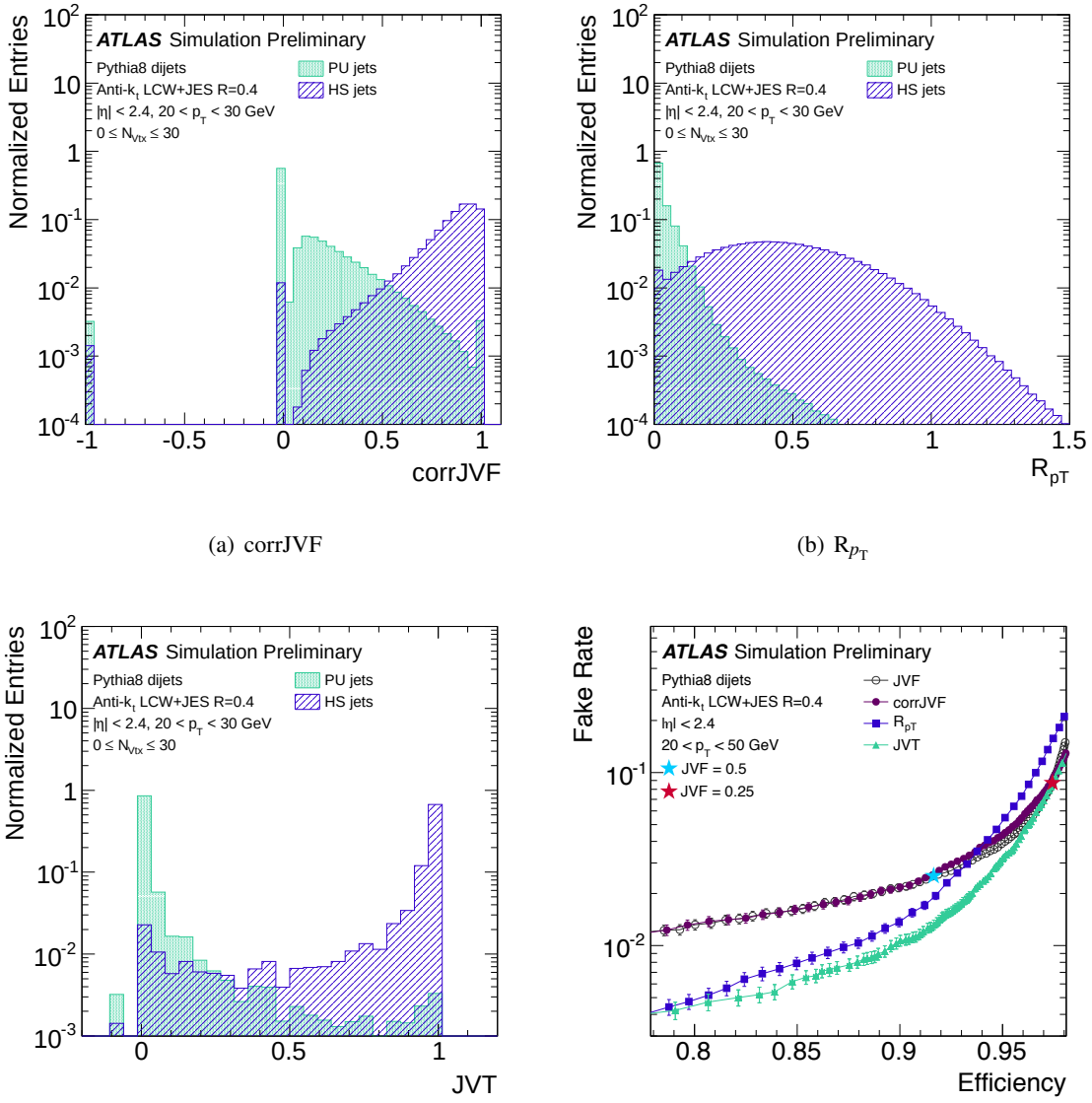
where $\sum p_T^{trk}$ is the scalar p_T sum of the tracks that are associated with the jet, n_{trk}^{PU} is the number of pile-up tracks in an event and the scaling factor "k" taken roughly as the slope of $\langle p_T^{PU} \rangle$ with n_{trk}^{PU} (k=0.01). The corrJVF only includes informations from the tracker, while the R_{p_T} also includes informations from the calorimeter. A value corrJVF = -1 is assigned to jets with no associated tracks. In simulation, the two jet samples, hard-scatter (HS) and pile-up (PU), are selected respectively by whether the jet is or is not matched (geometrically using $\Delta R = 0.4$) to a hard-scatter truth jet. Figure 2.23 shows the discrimination power of corrJVF, R_{p_T} and JVT. The efficiency versus fake rate is also shown for the three variable, where the gain using JVT is visible.

Jets from non-collision origins (bad/fake jets) need to be identified and excluded (procedure also known as cleaning) [64]. These fake jets can come from one of these sources:

- beam induced background: protons deviating from proton bunches upstream of the interaction point and scattering with the accelerator materials giving muons that reach the ATLAS detector,
- muons from cosmic-ray showers produced in the atmosphere overlapping with collision events,
- noisy LAr calorimeter cells, not detected by data quality inspection.

Quality discriminating variables used are:

- using the quadratic difference between the actual and expected pulse shapes in a cell, Q_{cell}^{LAr}, as a measure of a pulse quality, we define the following variables:
  - $\langle Q \rangle$: the energy-squared weighted average of Q_{cell}^{LAr} of the cells in a jet,
  - f_Q^{LAr} (f_Q^{HEC}): fraction of the energy in the LAr (HEC) calorimeter cells with poor pulse quality (Q_{cell}^{LAr}>4000).
- E_{neg}: sum of the energies of all cells with negative energy,
- f_{EM} (f_{HEC}): the ratio of the energy deposited in the electromagnetic (HEC) calorimeter to the total energy of the jet,
- f_{max}: the ratio of the maximum energy deposited in a single layer to the total energy of the jet,
- f_{ch}: the scalar sum of the p_T of the tracks coming from the primary vertex associated to the jet divided by the jet p_T.



(a) corrJVF

(b) R_{p_T}

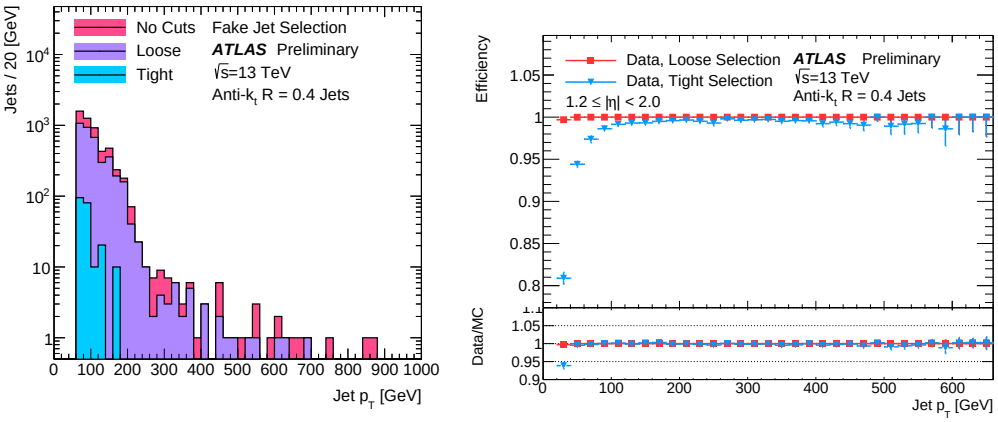
(c) JVT

(d) Efficiency vs fake rate

Figure 2.23: Distributions of corrJVF(a), R_{p_T} (b) and JVT (c) for pileup and hard-scatter jets. d) Efficiency versus fake rate for the different variables. Figures from [63].

Fake jets tend to be more localized longitudinally and characterized by very high or very low values for f_{EM} and f_{HEC}, high values of f_{max} and low values of f_{ch}. Fake jets from noisy cells are also discriminated using pulse shape quality and are characterized by high values of $\langle Q \rangle$ and f_Q^{LAr} or f_Q^{HEC}. Also, they tend to produce high absolute values of E_{neg}. Two sets of criteria, BadLoose and BadTight cleaning, are defined where the former is optimized for high good jets selection efficiency and the latter for high fake jet discrimination (≡ high purity). Using two samples one enriched with

fake and the other with good jets, the fake jets discriminating power and the good jets selection efficiency can be measured respectively as shown in figure 2.24. The good jets sample selects dijet events where the two leading jets are required to be back-to-back ($\Delta\phi_{jj} > 3$ rad) and to be balanced ($(p_T^1 - p_T^2)/(p_T^1 + p_T^2) < 0.3$). The fake jets sample select jets in events characterized by an unbalanced total transverse momentum satisfying $|\overrightarrow{H}_T^{miss} \equiv -\sum \overrightarrow{p_T}| > 70$ GeV (the sum uses all jets with $p_T > 20$ GeV) and where the fake jet is required to have a direction opposite to the one of $\overrightarrow{H}_T^{miss}$ ($\Delta\phi_{\overrightarrow{H}_T^{miss},jet} > 3$ rad). In addition, the fake jet is required to have a time (see next paragraph for time definition) higher than 6 ns to reject contribution from good physics signals like Z$\rightarrow \nu\nu$+jet. The loose cleaning has an efficiency above 99.5% (99.9%) for $p_T > 20$ (100) GeV. The tight cleaning has an efficiency above 95% (99.5%) for $p_T > 20$ (100) GeV.



(a) Fake jets enriched sample

(b) Good jets enriched sample

Figure 2.24: a) Fake jet transverse momentum distribution before and after applying cleaning loose and tight criteria. b) Jet cleaning efficiency in function of p_T for $1.2 < |\eta| < 2$. Figures from [64].

Jet time [65] is also used to identify jets coming from energy deposited from collisions of neighboring bunch-crossings, due to the large charge collection time of the calorimeter (hundreds of ns) compared to the bunch-crossing interval of 25 ns. Those jets are called out-of-time pile-up jets. In addition, jet time helps rejecting part of the non-collision background. A LAr calorimeter cell time, defined with respect to the event time recorded by the trigger which is synchronized with the bunch-crossing time, is determined from the signal shape of the LAr calorimeter response. The jet time is computed as the energy-square-weighted time average of the LAr cells reconstructed within the jet.

2.3.3 Jet energy scale (JES) calibration

JES calibration consists of multiple correction steps aimed at correcting the energy scale of the jets reconstructed at detector level (reco jets) to that of truth jets at particle-level [66, 67]. Figure 2.25 shows the different steps of the calibration. Regardless of the type of calibration of the jet inputs (EM-scale, PFlow-scale or any other), the calibration steps are the same but are implemented for each different jet reconstruction.

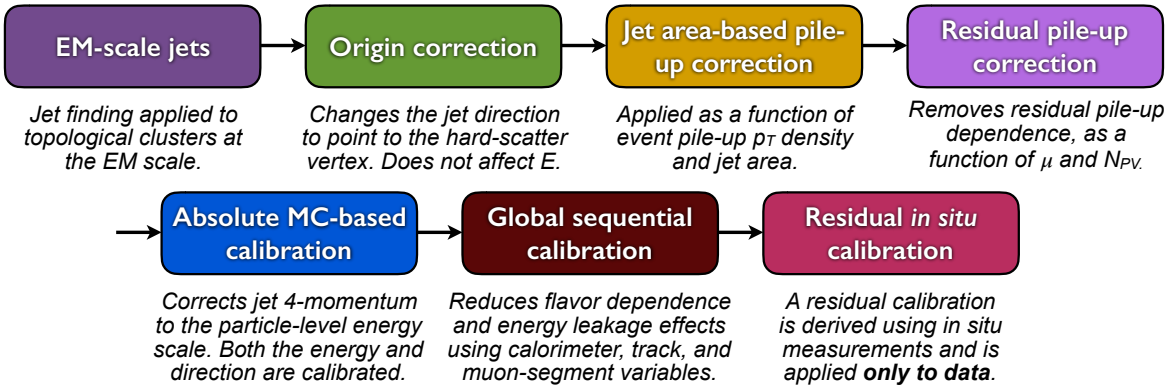


Figure 2.25: Jet energy scale various correction steps. Figure from [67].

First, the origin correction corrects the jet direction from pointing from the geometrical center of the detector to the hard-scatter primary vertex PV_0, but without changing the jet energy. As a result, the jet η resolution is improved. Recently, the step of correcting the origin was moved to the jet reconstruction step. Jet constituents entering the jet finding algorithm are now corrected to point to the PV_0.
Next, the pile-up contribution to the jet energy and momentum is corrected using two components: an area based subtraction and an additional residual correction derived from MC simulation. Then the absolute calibration, based also on MC simulation, is applied which corrects both energy and η direction of jets to the ones from truth jets. The global sequential calibration further improves the calibration and minimizes the differences between jet flavors using additional information from the ID, the MS and the calorimeter. At last, to catch any mis-modeling of the detector simulation which makes the correction not perfect for the data, a calibration using in-situ/real events is derived by comparing jets to a well measured reference object. For these corrections, η_{det}, the jet η pointing from the geometrical center of the detector, is used to remove any ambiguity as to which region of the detector is probed by the jet. Let me detail each of the mentioned steps.

2.3.3.1 Pile-up corrections

In-time and out-of-time pile-up contribute to the jet energy measured by the calorimeter. The first correction uses the pile-up energy density to subtract its contribution in jets according to the jet area. Each jet has a defined area, A, measured by the jet finding algorithm using ghost association. Infinitesimal momentum ghost particles are added uniformly in solid angle to the event before jet reconstruction. The ghost particles are then clustered with the real particles into a jet, with their infinitesimal momentum not affecting the clustering output. The jet area is then the number of ghost particles associated with it divided by their area density. Next, we still need, for the first correction, to measure the pile-up p_{T} density, ρ, and subtract it from the jet p_{T}. To measure ρ, new jets are clustered using kt algorithm with a radius of 0.4 reconstructed only from positive-energy topo-clusters with $|\eta| < 2$ and with no minimum p_{T} threshold. The kt algorithm is used due to its sensitivity to soft

radiation. Only the central region is used due to the higher occupancy of the forward detectors. The p_T density of each jet is thus p_T/A. ρ is taken as the median of the p_T density distribution (the median is used to reduce the bias from hard-scatter jets populating the high tails of the distribution). Figure 2.26 shows the ρ distribution for a given N_{PV} and μ.

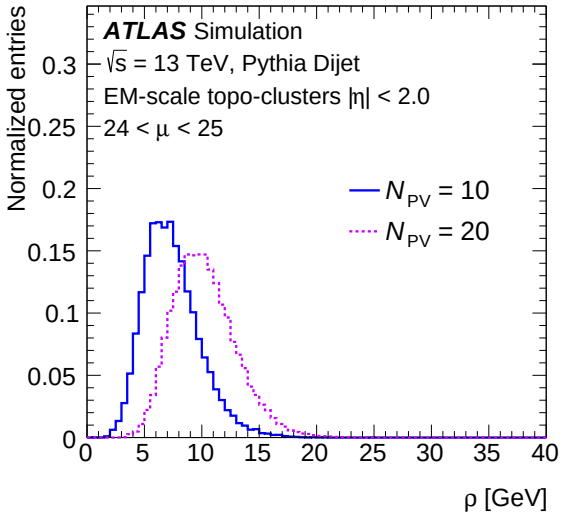


Figure 2.26: Per-event median p_T density, ρ, at NPV = 10 (solid) and NPV = 20 (dotted) for $24 < \mu < 25$ as found in MC simulation. Figure from [67].

The ρ calculation does not fully describe the pile-up contribution in the forward calorimeter region. A residual calibration is applied in addition. To characterize the in-time pile-up activity, the number of primary vertices (N_{PV}) in an event provides a good indication of the number of p-p collision in the same event. The average number of interactions μ, which is calculated from several bunch-crossings, provides a good estimation of the out-of-time pile activity. MC simulation is used to measure the difference between the reconstructed jet p_T and truth jet p_T matched geometrically within a distance $\Delta R = 0.3$. The difference is measured as a function of each of N_{PV} and μ variables, and in bins of η_{det}. The dependence is found to be linear, and the slopes of the differences define the correction coefficients:

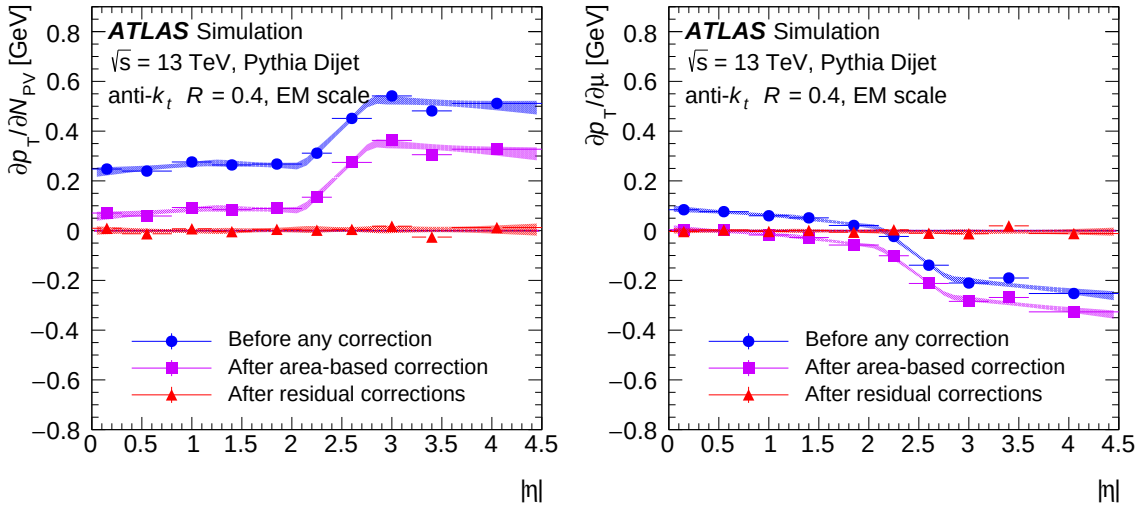
$$\begin{aligned} \alpha(\eta_{det}) &= \frac{\partial p_T}{\partial N_{PV}}(\eta_{det}), \\ \beta(\eta_{det}) &= \frac{\partial p_T}{\partial \mu}(\eta_{det}). \end{aligned} \tag{2.2}$$

No significant evidence is found for cross-terms in the sensitivity of the jet p_T to in-time and out-of-time pile-up. This was tested by measuring the α (β) slopes of the jet p_T dependence on NPV (μ) for different bins of μ (NPV) and finding them to be fairly similar.
The total pile-up corrected p_T is therefore:

$$p_T^{corr} = p_T^{reco} - \rho \times A - \alpha \times (N_{PV} - 1) - \beta \times \mu. \tag{2.3}$$

The ratio of the p_T^{corr} to the uncorrected p_T^{reco} is taken as a correction factor and applied to the jet four-momentum, without affecting the jet η direction. Figure 2.27 shows the dependence of the jet

p_{T} on in-time and out-of-time pile-up contribution before and after the two correction steps. The central values of the purple band, derived from linear fits in 4 $|\eta|$ regions ($|\eta| < 1.2$, $1.2 < |\eta| < 2.2$, $2.2 < |\eta| < 2.8$, $2.8 < |\eta| < 4.5$), gives the residual correction coefficients α and β. After the total correction, the dependence is compatible with zero.



(a) In-time pile-up dependence

(b) Out-of-time pile-up dependence

Figure 2.27: Dependence of the jet p_{T} on in-time (a) and out-of-time (b) pile-up contribution before and after the two correction steps. The full markers are the results for each $|\eta_{\mathrm{det}}|$ bin. The bands are the 68% confidence intervals of the linear fits in 4 $|\eta_{\mathrm{det}}|$ regions ($|\eta| < 1.2$, $1.2 < |\eta| < 2.2$, $2.2 < |\eta| < 2.8$, $2.8 < |\eta| < 4.5$). The central values of the purple band gives the residual correction coefficients α and β. Figures from [67].

Four systematic uncertainties are introduced with the calibration. They account for the mis-modeling of $\mathrm{N_{PV}}$ and μ in MC simulation, for the ρ calculation which contains contribution from the underlying event (UE), and for the p_{T}-dependence of the corrections coefficients α and β (they have a logarithmic dependence on $p_{\mathrm{T}}^{\mathrm{truth}}$).

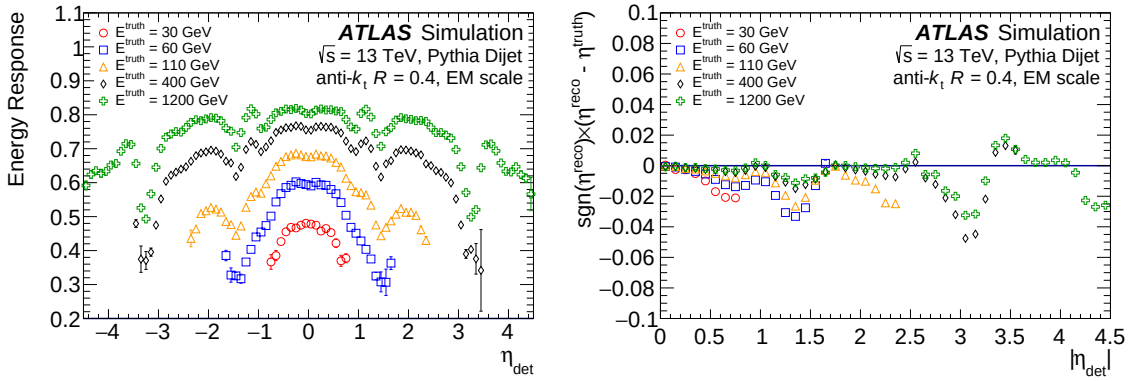
2.3.3.2 Absolute MC calibration (MC-JES)

After the pile-up calibration is applied, absolute JES and η corrections correct the reconstructed jet four-momentum to the particle-level energy scale. First, the JES is derived. Reconstructed jets in the MC simulation are geometrically matched to truth jets within $\Delta\mathrm{R} = 0.3$. An isolation criteria is also applied to both reco and truth jets. The average energy response, $\mathcal{R}$, is calculated as the mean of a Gaussian fit to the distribution $\mathrm{E}^{\mathrm{reco}}/\mathrm{E}^{\mathrm{truth}}$. $\mathcal{R}$ is calculated in bins of $\mathrm{E}^{\mathrm{truth}}$ and η_{det}. Through numerical inversion, the average response is parametrized from $\mathrm{E}^{\mathrm{truth}}$ as a function of $\mathrm{E}^{\mathrm{reco}}$, and the calibration factor is taken as the inverse of this response.

Secondly, the jet η needs to be corrected. In fact, a bias to the jet η direction is caused by an artificial increase of the energy in one side of the jet with respect to the other. The biases are largest

in the transition between different calorimeter sub-detectors or when the granularity changes. The difference between the reconstructed, η^{reco}, and the truth, η^{truth}, directions is calculated for matched and isolated jets. The difference, calculated in bins of E^{truth} and η_{det}, is numerically inversed and parametrized as function of E^{reco}. Unlike the other calibration stages, the η calibration corrects only the jet p_{T} and η, not the full four-momentum.

Figure 2.28 shows the average energy response $\mathcal{R}$ in the left and the η difference between reco and truth jets on the right as a function of η_{det} and for the different E^{truth} bins. Notable sharp changes in the energy response and hence giving the highest η differences can be seen in the barrel-endcap ($|\eta|_{\text{det}} \approx 1.4$) and endcap-forward ($|\eta|_{\text{det}} \approx 3.1$) transition regions.



(a) Energy response

(b) η difference

Figure 2.28: Average energy response $\mathcal{R}$ (a) and (signed) η difference between reco and truth jets (b) as a function of η_{det} and for the different E^{truth} bins. Figures from [67].

2.3.3.3 Global sequential calibration

Even after applying the previous described calibrations, the calorimeter is still sensitive to the jet particle composition and the distribution of energy within the jet. The average particle composition and shower shape of a jet varies depending on the partons initiating the jet, most notably between quark- and gluon-initiated jets. A quark-initiated jet will often include hadrons with a higher fraction of the jet p_{T}, penetrates further into the calorimeter and its energy is more concentrated in its center. On the other hand, a gluon-initiated jet will typically contain a higher multiplicity of hadrons ($\approx$ double the quark-initiated multiplicity), more particles of softer p_{T} and larger angle from the center, leading to a lower calorimeter response and a wider transverse profile. Hence, a residual dependence of the JES on longitudinal and transverse features of the jet exists. Using multiple observables describing these features, the global sequential calibration (GSC) applies independent jet four-momentum corrections independently for each observable in successive order. As a consequence of removing the response sensitivity to the longitudinal and transverse features, the jet resolution is improved by the GSC and the difference between the different jet flavor responses is reduced. An important note is that the GSC does not change the average energy response in the dijet sample used to derive it, but only reduces its fluctuation. The five observables used for the correction are (in the order of their application):

- f_{Tile0}: the fraction of jet energy in the first layer of the hadronic tile calorimeter ($|\eta|_{det} < 1.7$),
- f_{LAr3}: the fraction of jet energy in the third layer of the electromagnetic LAr calorimeter ($|\eta|_{det} < 3.5$),
- n_{trk}: the number of tracks with $p_T > 1$GeV associated to the jet ($|\eta|_{det} < 2.5$),
- W_{trk}: the average p_T-weighted transverse distance from the jet center of tracks with $p_T > 1$GeV associated to the jet ($|\eta|_{det} < 2.5$),
- $n_{segments}$: the number of muon track segments associated to the jet ($|\eta|_{det} < 2.7$).

The first two observables give information on the penetration of the jet, while the second two give information on the particle multiplicity of the jet and their transverse profile. The last one is aimed at correcting the response of high-p_T jets that are not fully contained in the calorimeter by using the punch-through signal in the muon spectrometer. The corrections are derived as a function of η_{det}, and p_T^{truth} for the first four observables and E^{truth} for the last one since it is more correlated with the energy escaping the calorimeter. The same methods for geometrical matching between reconstructed and truth jets and for numerical inversion are used. In addition, an overall constant is multiplied to each numerical inversion to ensure the average energy is unchanged at each of the five corrections. Figure 2.29 shows the average jet response as a function of the five GSC variables.

For PFlow jets, which are reconstructed from tracks in addition to energy deposits in the calorimeter, an additional observable is used for the GSC correction and is first to be applied: $f_{charged}$, the fraction of the jet p_T measured from associated tracks with $p_T > 0.5$GeV ($|\eta|_{det} < 2.5$).

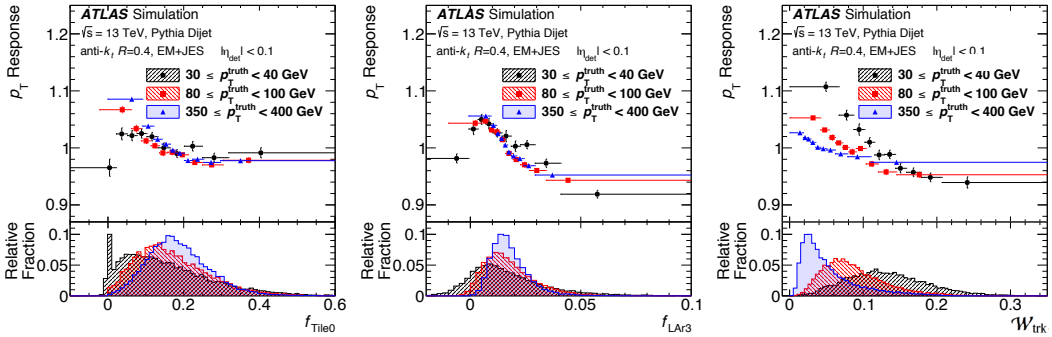
Flavor systematic uncertainty

Even after the GSC calibration, a residual flavor dependence of the response exists. Hence, a systematic uncertainty should be evaluated to account for that. First, using a nominal MC simulation, the fraction of gluon-induced jets, f_g is obtained as a function of p_T and η for the specific analysis under-study. In addition, from multiple simulation generators, the uncertainty on this fraction, Δf_g, can also be computed [7]. The total flavor uncertainty is thus:

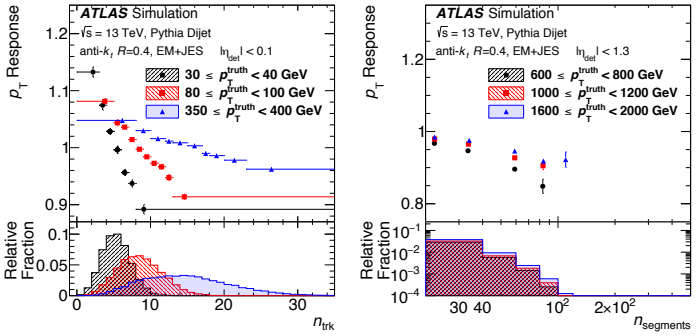
$$\Delta\mathcal{R}_{flavor} = \Delta f_g \times (\mathcal{R}_q - \mathcal{R}_g) \oplus f_g \times \Delta\mathcal{R}_g, \tag{2.4}$$

where $\mathcal{R}_q$ and $\mathcal{R}_g$ are the quark- and gluon-initiated jet responses respectively. The first term, called the flavor composition uncertainty, gives the uncertainty from the difference in the responses of quark- and gluon-initiated jets. The second term, called the flavor response uncertainty, accounts for the differences between the gluon-initiated jet responses obtained from the two showering and hadronization models implemented in Pythia and Herwig 1.3.2.

[7] For inclusive jets, Δf_g is in the order of few percent and not exceeding 10% for the majority of the phase space.



(a) f_{Tile0} (b) f_{LAr3} (c) n_{trk}



(d) W_{trk} (e) $n_{segments}$

Figure 2.29: The average jet response in MC simulation as a function of the five GSC variables for three ranges of pTtruth. The calorimeter distributions (a) and (b) are shown with no GSC corrections applied, the track-based distributions (c) and (d) are shown with both preceding calorimeter corrections applied, and the punch-through distribution (e) is shown with the four calorimeter and track-based corrections applied. The distributions of the underlying observables in MC simulation are shown in the lower panels for each pTtruth region, normalized to unity. Figures from [67].

2.3.3.4 In-situ calibration

As we have seen for all the previous calibration methods, the jet response measurement relies on the simulation of ATLAS active detectors and dead materials, as well as on the simulation of the hard scatter, underlying event, pile-up, particle showers in the detectors, jet particles multiplicity and transverse shape. The last calibration methods capture the residual differences between simulation and data, and measure correction factors to be applied only on data. For in-situ methods, the jet response is calculated by balancing the p_T of a probe jet against that of a well-calibrated reference object or system, where the probe and reference objects are fairly back-to-back in the transverse plane. The response is defined as the mean (or mean of a fit) of the ratio of the two p_T:

$$\mathcal{R} = \langle \frac{p_T^{\text{probe}}}{p_T^{\text{reference}}} \rangle. \tag{2.5}$$

The ratio of $\mathcal{R}$ between data and MC simulation is a good estimate of the additional JES correction factor that needs to be applied on data:

$$C_{\text{in-situ}} = \frac{\mathcal{R}^{\text{MC}}_{\text{in-situ}}}{\mathcal{R}^{\text{data}}_{\text{in-situ}}}. \tag{2.6}$$

The balance and hence the response $\mathcal{R}$ are sensitive to additional jets in the system and on energy flow into or out of the jet cone. On the other hand, the double ratio C is more robust to those secondary effects as long as they are well modeled in simulation. In addition, event selections are designed to reduce the impact of such secondary effects. The following in-situ methods are used:

- eta-intercalibration: using dijet events, it corrects the energy scale of forward jets using well-measured central jets ($|\eta| < 0.8$),
- Z(γ)+jet: corrects the energy scale of central jets using the well-measured Z(γ) objects,
- multi-jets balance (MJB): corrects the energy scale of a central high p_{T} jet using a system of well-calibrated low p_{T} multi-jets.

Figure 2.30 shows the systems used to exploit their balance for the different calibrations. The corrections are measured in bins on $p_{\text{T}}^{\text{ref}}$ ($p_{\text{T}}^{\text{average}}$ and η_{det} in the case of eta-intercalibration). Numerical inversions are used to transform the corrections into a function of probe jet p_{T}.

Z+jet, γ+jet and MJB corrections, mostly relevant in low, mid and high p_{T} ranges respectively, are statistically combined into one calibration covering the full p_{T} range. Although this calibration is computed using only central jets, it is also applied to forward jets since the eta-intercalibration, the first in-situ calibration to be applied, has already equalized the central and forward energy scales. Since in-situ calibrations are performed sequentially, systematic uncertainties are propagated from each to the next one. The event selection cuts are a source of systematic uncertainties and are evaluated by varying the cut to looser or tighter values in both data and MC. The uncertainty from the measurement of the reference object is also propagated. The last systematic uncertainty is related to the MC physics modeling (affecting the balance) and is calculated by measuring the correction factors C using two MC generators and taking the difference as the systematic. Smoothing or rebinning is applied to systematic uncertainties to ensure that the values of the uncertainties are statistically significant and not a result of fluctuations. The statistical correlation is taken into account during those steps.

Common quality criteria are applied to all three in situ analyses. Each event must have a reconstructed vertex with at least two associated tracks of $p_{\text{T}} > 0.5$ GeV. All jets must satisfy cleaning quality criteria as described in 2.3.2. Furthermore, the JVT tool is used to reject pile-up jets.

The eta-intercalibration, one of the main studies I performed thoroughly, is described in the next chapter with all my contributions also detailed. Let me here describe the other in-situ methods and the in-situ combination.

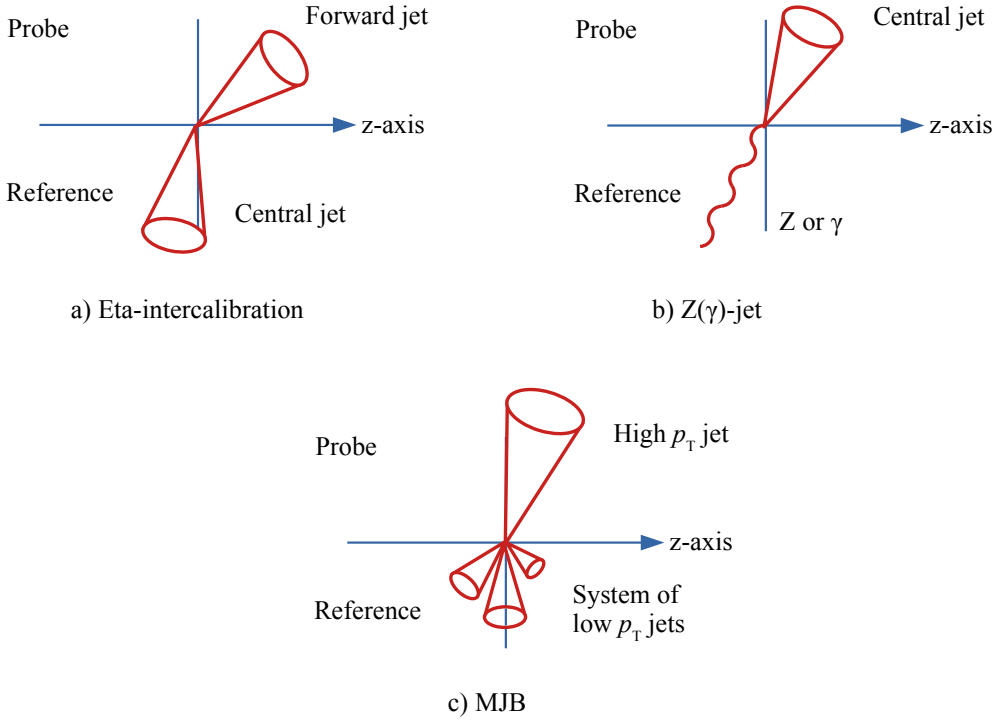


Figure 2.30: Reference and probe objects used in the different in-situ calibrations.

Z/γ+jet

The balance between a central jet ($|\eta| < 0.8$) and a well-measured Z or photon is used to measure the jet response and compare it between MC simulation and data. The Z boson is reconstructed through the decay channels $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$. The correction is evaluated separately for each channel. Photons, electrons and muons are precisely measured in the ATLAS detector, which explains why they are used as reference objects for JES measurements.

Two techniques to derive the response are used. The first is the direct balance (DB) which uses the p_T of the fully reconstructed jet and compares it to the reference p_T. The correction is sensitive to additional jets in the event affecting the balance between the leading jet and the reference. To minimize that, a p_T cut on additional radiation is used and a minimum azimuthal separation $\Delta\phi$ between the jet and the reference is imposed. Furthermore, an improvement is achieved by only considering the boson p_T parallel to the jet axis, and hence we define the reference p_T and the response as:

$$\begin{aligned} p_T^{\text{ref}}(\text{DB}) &= p_T^{Z/\gamma} \times \cos(\Delta\phi), \\ \mathcal{R}_{\text{DB}} &= \left\langle \frac{p_T^{\text{jet}}}{p_T^{\text{ref}}(\text{DB})} \right\rangle. \end{aligned} \tag{2.7}$$

Jet energies are affected by out-of-cone radiation (OOC, the energy radiated at an angle larger than

the jet cone) and underlying event. This is true also for truth jets in simulation, which means that the balance with the Z boson (which is not affected by OOC and underlying event (UE) effects) is not perfect even at truth level. Correction factors k for the effects of OOC and UE are estimated by measuring the p_{T} profile of tracks around the jet. Those factors are calculated for data and simulation and are applied to the Z p_{T}, and the variated response measurement is performed. The comparison of the final double ratio of responses between data and MC simulation for the variated measurement and the nominal one (without k factors) gives the systematic uncertainty of the effects of OOC and UE.

A second technique, the missing-$\mathrm{E_T}$ projection fraction (MPF), uses the full hadronic recoil in an event as a probe instead of just the leading jet p_{T}. It measures the response of the calorimeter to hadronic showers. From transverse momentum conservation, we have:

$$\overrightarrow{p_{\mathrm{T}}}^{\mathrm{ref}} + \overrightarrow{p_{\mathrm{T}}}^{\mathrm{recoil}} = \overrightarrow{0}. \tag{2.8}$$

The missing transverse momentum in the event, $\mathrm{E_T}^{\mathrm{miss}}$, is calculated directly from calorimeter topo-clusters in the case the EMTopo jet calibration and from particle flow objects in the case of the PFlow jet calibration, ensuring that the energy scale is consistent. Assuming that any missing transverse momentum is due only to the response of the hadronic recoil (i.e. the response of the reference object is considered equal to 1), we have:

$$\overrightarrow{p_{\mathrm{T}}}^{\mathrm{ref}} + \mathcal{R}_{\mathrm{MPF}} \times \overrightarrow{p_{\mathrm{T}}}^{\mathrm{recoil}} = -\overrightarrow{E_T}^{\mathrm{miss}}. \tag{2.9}$$

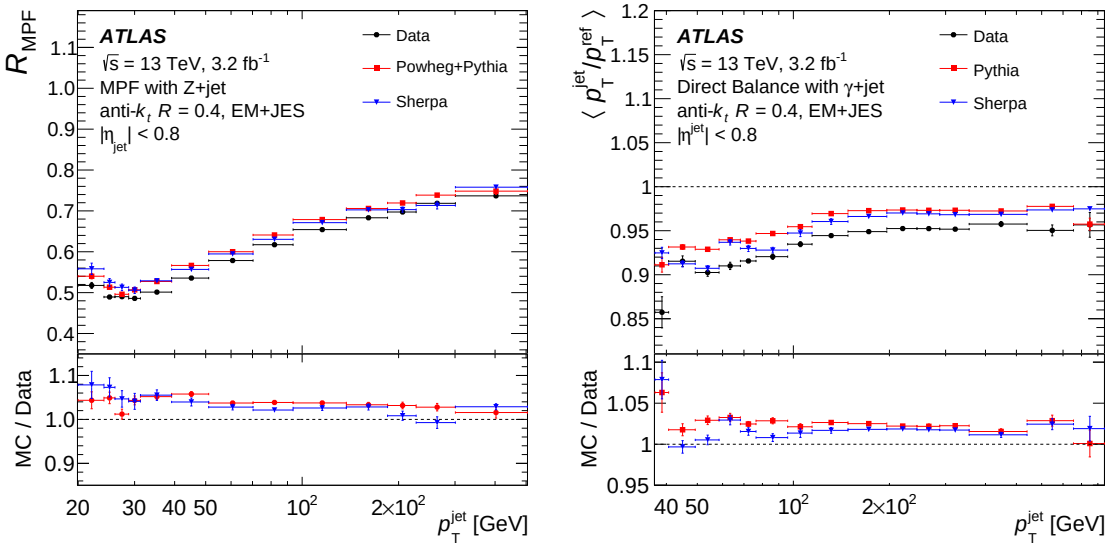
Combining the two equations and projecting in the direction of the reference object $\hat{\mathrm{n}}_{\mathrm{ref}}$, we get:

$$\mathcal{R}_{\mathrm{MPF}} = \left\langle 1 + \frac{\hat{\mathrm{n}}_{\mathrm{ref}} \cdot \overrightarrow{E_T}^{\mathrm{miss}}}{p_{\mathrm{T}}^{\mathrm{ref}}} \right\rangle. \tag{2.10}$$

The average response in both DB and MPF techniques is computed from a fit using a modified Poisson distribution. The MPF technique is less sensitive to additional particle activity that is symmetric in the transverse plane, like pile-up and underlying event. Sensitivity on additional radiation still exists and a cut is used to put a threshold on additional jets p_{T}. The out-of-cone and jet reconstruction effects are reduced compared to DB technique, since MPF technique only uses jets for the event selection. The systematic uncertainty of the effects of OOC and UE are taken from the DB technique. Numerical inversion is derived using the average jet p_{T}, before the current calibration, within each reference p_{T} bin.

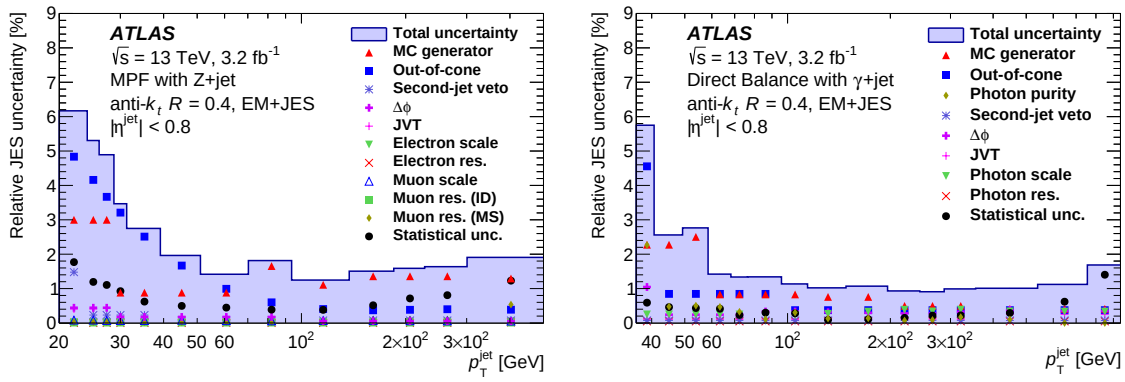
Both techniques are used to measure the final correction as a cross-check. Recently, the correction from MPF technique is the one used in the final combination. Figure 2.31 shows the response measurement, using data and 2 MC simulations, in Z+jet in the left using the MPF technique, and for γ+jet on the right using the DB technique. In the bottom plot, the ratios between simulations and data are shown, which correspond to the in-situ correction factors, where one is for the nominal values and the other is for measuring the modeling systematic uncertainty.

Figure 2.32 shows the difference systematic uncertainties of the calibration. They come from simulation modeling, event selection, OOC and UE effects, and the effect of the uncertainties on the reference object measurement (γ, e and μ energy/momentum scale and resolution).



(a) Z+jet (b) γ+jet

Figure 2.31: Average p_{T} response for a) Z+jet events using the MPF technique and b) γ+jet events using the DB technique as a function of jet p_{T}. The jets used to measure the response are calibrated up to the eta-intercalibration stage. The response is given for data and two distinct MC samples, and the MC-to-data ratio plots in the bottom panels reflect the derived in situ corrections. Figures from [67].



(a) Z+jet (b) γ+jet

Figure 2.32: Systematic uncertainties for a) Z+jet events using the MPF technique and b) γ+jet events using the DB technique as a function of jet p_{T}. The uncertainties account for event selection, OOC and UE effects, modeling differences and effects of the reference object energy/momentum scale and resolution. Figures from [67].

MJB

The last stage of in situ calibration extends the reach of the correction to high-p_{T} jets above the range of the Z/γ+jet calibration using the multi-jet balance (MJB) technique. The balance of a single high-p_{T} jet against a recoil system of two or more lower-p_{T} jets is used to define the MJB response

as:

$$\mathcal{R}_{\mathrm{MJB}} = \left\langle \frac{p_{\mathrm{T}}^{\mathrm{lead}}}{p_{\mathrm{T}}^{\mathrm{recoil}}} \right\rangle, \tag{2.11}$$

where $p_{\mathrm{T}}^{\mathrm{recoil}}$ is the four-momentum sum of the recoil jets. The leading jet is calibrated only up to the eta-intercalibration stage, while the recoil system uses in addition the Z/γ+jet calibration. The jets of the recoil system are selected to ensure they are well calibrated. They are limited to the region $|\eta| < 2.8$, their p_{T} is limited to the range of the Z/γ+jet calibration, and the recoil system is not allowed to have one jet with more than 80% of its p_{T} to suppress events with dijet topology from the selection. Isolation of the leading jet from contamination by the recoil system is ensured by requiring that the azimuthal separation $\Delta\phi$ between the leading jet and the direction of the recoil system is at least 0.3 rad, and that the $\Delta\phi$ between the leading jet and any individual jet in the recoil system with a $p_{\mathrm{T}} > 0.25 p_{\mathrm{T}}^{\mathrm{lead}}$ is at least 1.0 rad. The leading jets are limited to the region $|\eta| < 1.2$ for the 2015 calibration but is later reduced to $|\eta| < 0.8$ for the full Run II calibration when more statistics were collected.

The constrains of the MJB analysis on the recoil system to be well calibrated lead to the exclusion of events with very high $p_{\mathrm{T}}^{\mathrm{lead}}$ as their second leading jet has a momentum outside the range of calibration by the Z/γ+jet analyses. To address this, MJB proceeds via multiple iterations (two are found to be enough). After the first iteration, the recoil system is also calibrated by MJB results from the previous iteration extending the p_{T} range of the well-calibrated jets. For the full Run II calibration, the iteration procedure extended the MJB calibration to $p_{\mathrm{T}}^{\mathrm{lead}} = 2.5$TeV.

The response $\mathcal{R}_{\mathrm{MJB}}$ is shown for data and MC simulation in figure 2.33. As expected, an offset is seen between data and MC simulation, reflecting that the recoil system in data is fully calibrated to Z/γ+jet stage while the leading jet is only partially calibrated. The response is below unity even in MC simulation, particularly at low p_{T}, reflecting the effect of the selection and isolation requirements of the leading jet and the recoil system on $\mathcal{R}_{\mathrm{MJB}}$. The MC-to-data ratio is shown in the bottom panel and a fairly constant correction of 2% is derived.

Systematic uncertainties are shown in figure 2.34. They account for the event selection, the MC modeling and recoil jets calibration. For the event selection, the uncertainties are measured by varying separately each cut to looser or tighter values. The uncertainty due to MC modeling is taken as usual from the difference in the MJB correction between different generators. The JES uncertainties, including the in-situ ones, on the recoil jets are also propagated by varying the calibration by $\pm 1\sigma$ for each component individually. The jet flavor uncertainty dominates at low p_{T} while γ+jet calibration uncertainty dominates at high p_{T}.

Combination of in-situ results

From each of the previously described in-situ methods which compares the response (using a probe and reference objects) between data and simulation $\mathcal{R}_{\mathrm{in-situ}} = \mathcal{R}^{\mathrm{data}}/\mathcal{R}^{\mathrm{MC}}$, a correction factor $C_{\mathrm{in-situ}} = 1/\mathcal{R}_{\mathrm{in-situ}}$ (equation 2.6) is measured and is to be applied to data. The first correction applied is the eta-intercalibration to remove the η_{det} dependence of the jet response. Next, the four absolute corrections, Z $\rightarrow$ ee+jet, Z $\rightarrow$ $\mu\mu$+jet, γ+jet and MJB, must be combined to produce a single calibration covering the full range of jet p_{T}, from 17 GeV to 2.5 TeV for the full Run II calibration. The combination takes into consideration the overlapping of the four corrections by accounting for

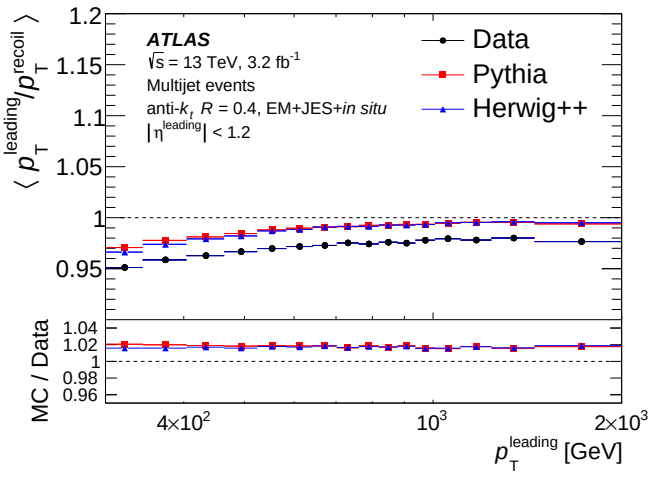


Figure 2.33: Average p_{T} response of the MJB analysis in multi-jet events as a function of leading jet p_{T}. The leading jets used to measure the response are calibrated up to the eta-intercalibration stage while for the recoil system they are fully calibrated. Figures from [67].

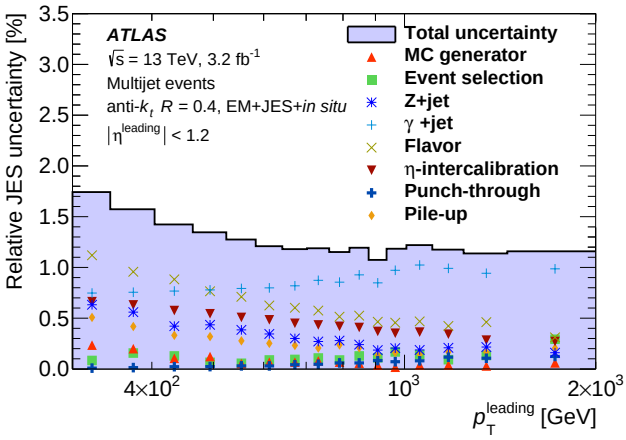


Figure 2.34: Systematic uncertainties from the event selection, the MC modeling and recoil jets calibration are shown for the MJB analysis as a function of leading jet p_{T}. The leading jets used to measure the response are calibrated up to the eta-intercalibration stage while for the recoil system they are fully calibrated. Figures from [67].

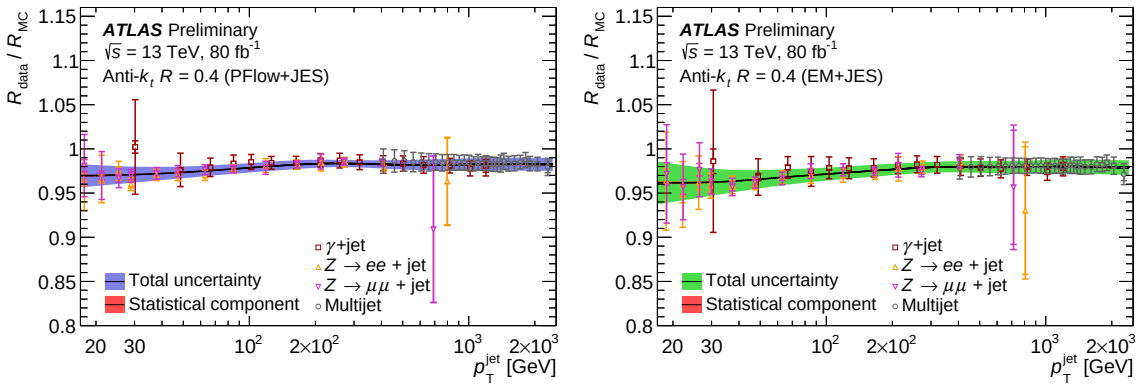
their relative statistical power as well as any tension between different response ratio measurements in the same p_{T} range.

First, for each of the four absolute corrections, the ratio $\mathcal{R}_{\mathrm{in-situ}}$, initially evaluated at the barycenter $\langle p_{\mathrm{T}}^{\mathrm{jet}} \rangle$ of each $p_{\mathrm{T}}^{\mathrm{jet}}$ bin, is determined in a fine 1 GeV binning using interpolating second-order polynomial splines. Next, a χ^2 minimization is performed in each 1 GeV bin for the available absolute measurement in that p_{T} range, and taking into account their statistical and systematic uncertainties to determine a weight for each measurement in each bin. The smaller the uncertainties

for a measurement, the higher its weight and its effect on the combination. The local χ^2 is used to quantify the agreement between the in-situ methods. If a tension is found in a bin, expressed by the factor $\sqrt{\chi^2/\mathrm{N_{dof}}}$ being larger than 1 where $\mathrm{N_{dof}}$ is the number of degree of freedom, the uncertainties of the measurement inside this bin are multiplied by this factor. The final calibration is obtained by smoothing the combined inputs from the fine binning using a Gaussian kernel.

Each uncertainty source of the in-situ methods is treated as fully correlated across p_T and η ranges, while it is treated as independent of other sources within and between calibration methods. Therefore, for each uncertainty source, the corresponding in-situ correction is shifted by $\pm 1\sigma$ of the uncertainty, then the finer binning interpolation, the combination (≡ averaging) and the smoothing procedures are repeated, while keeping the relative weights of the different in-situ methods the same as the nominal one. The difference between the combined calibration curve with the systematically shifted input and the nominal calibration curve is taken as the 1σ variation for each uncertainty source. For the uncertainty sources of the Z/γ+jet calibrations that are propagated to the MJB calibration, one-to-one correlations are considered and each of those uncertainties shifts coherently the corrections of the two methods, Z+ or γ+jet and MJB.

Figure 2.35 shows the response ratio $\mathcal{R}^\mathrm{data}/\mathcal{R}^\mathrm{MC}$ for EMPFlow and EMTopo jets from the four absolute in-situ calibration methods individually, Z → ee+jet, Z → $\mu\mu$+jet, γ+jet and MJB, and the final smoothed combination with the statistical and systematic uncertainties shown. Smaller uncertainties for EMPFlow calibration are visible at low p_T (due to the smaller jet energy resolution as is shown in the next section). On the other hand, at high p_T, the central values and their uncertainties become similar between EMPFlow and EMTopo calibrations which is expected. The residual correction to be applied only to data is the inverse of this ratio.



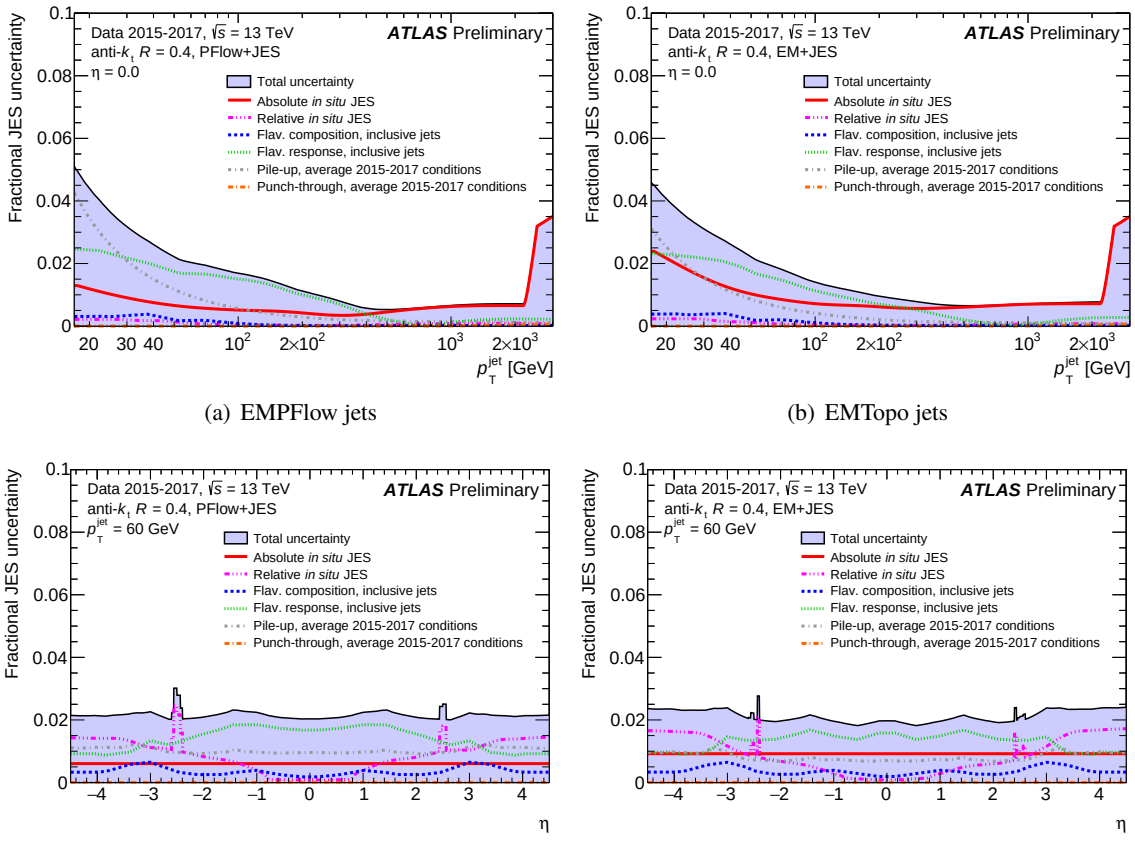
(a) EMPFlow jets

(b) EMTopo jets

Figure 2.35: Data-to-simulation ratio of the average jet p_T response as a function of jet p_T. The combined result is based on four in-situ measurement using Z → ee+jet, Z → $\mu\mu$+jet, γ+jet and MJB techniques. The errors represent the statistical (inner error bars and small inner band) and the total uncertainty (statistical and systematic uncertainties added in quadrature, outer error bars and outer band). These results are for anti-kt jets with R=0.4 reconstructed from particle flow objects in a) and from topo-clusters in b). The residual correction to be applied only to data is the inverse of this ratio. Figures from the ATLAS public plots page [68].

At the end of this JES calibration section, let me show 1D slices of the full 2D (p_T, η) uncertainties of

the full chain of the calibration. Figure 2.36 shows the full uncertainties for EMPFlow and EMTopo calibrations for $\eta = 0$ as a function of p_{T}, and for $p_{\mathrm{T}} = 60$ GeV as a function of η.



(a) EMPFlow jets

(b) EMTopo jets

(c) EMPFlow jets

(d) EMTopo jets

Figure 2.36: Fractional jet energy scale systematic uncertainty components for $\eta = 0$ as a function of p_{T} in a) and b), and for $p_{\mathrm{T}} = 60$GeVas a function of η in c) and d). These results are for anti-kt jets with R=0.4 reconstructed from particle flow objects in a) and c), and from topo-clusters in b) and d). The total uncertainty (all components summed in quadrature) is shown as a filled region topped by a solid black line. Topology-dependent components are shown under the assumption of a dijet flavor composition. Figures from the ATLAS public plots page [68].

2.3.4 Jet energy resolution (JER)

Knowing the resolution of our detector is a crucial step in any physics analysis. The resolution in data is used to validate the simulation of the detector and to correct it where it is possible. The resolution is used also to connect the particle/truth and reconstructed levels; passing from the first to the second through folding and from the second to the first through unfolding of the detector effects. Hence, knowing precisely the jet energy resolution (JER) in data is very important to validate the detector simulation which is used in folding and unfolding techniques. In addition, systematic uncertainties

impacting the JER determination are evaluated and propagated to the physics measurements. The JER determination is done after applying all the calibration steps that impact the resolution, as well as (a preliminary version of) the in-situ calibration (which only impacts the average response).

For JES determination using in-situ methods, average responses from several event topologies are used defined as $\mathcal{R} = \langle p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{ref}} \rangle$ (equation 2.5). On the other hand, the width of this distribution, $\sigma(p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{ref}})$, is used to probe the detector jet energy resolution (JER). $\sigma(p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{ref}})$ is not equal to the JER since even at particle/truth level, the balance of the jet and the reference object is not perfect due to physics effects: the balance is affected by additional radiation which is not clustered in the same jet, by the hadronization which dissipate particles out of the jet cone, and by the UE effects which adds particles to the jet. The balance is only perfect at parton level for exclusively $2 \rightarrow 2$ scattering.

The contribution of physics effects to the width of $p_{\mathrm{T}}^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{ref}}$ distribution is estimated using MC simulation, by measuring the width of the balance distribution between a truth jet, which have a matching with jets at reconstructed level, and the reference object. By subtracting the physics effect, only the detector effects remain; the in-situ JER estimation, is then:

$$\mathrm{JER}_{\mathrm{in-situ}} \equiv \left(\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} \right)_{\mathrm{in-situ}} = \sigma(p_{\mathrm{T}}^{\text{reco jet}}/p_{\mathrm{T}}^{\mathrm{ref}}) \ominus \sigma(p_{\mathrm{T}}^{\text{truth jet}}/p_{\mathrm{T}}^{\mathrm{ref}}). \tag{2.12}$$

The in-situ JER estimation can be performed for data and for simulation. Dijet and Z/γ+jet methods can be used and combined for this matter.

Another method of estimating the JER in simulation is by directly comparing the p_{T} of matched truth jets and reco jets. In this method, the MC JER is estimated in bins of $p_{\mathrm{T}}^{\mathrm{truth}}$ by:

$$\mathrm{JER}^{\mathrm{MC}} \equiv \left(\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} \right)^{\mathrm{MC}} = \sigma(p_{\mathrm{T}}^{\text{reco jet}}/p_{\mathrm{T}}^{\text{truth jet}}). \tag{2.13}$$

The two methods of JER determination in simulation are cross-checked and give compatible results.

The JER is parametrized as the standard functional form expected for calorimeter-based resolutions [69], with three independent contributions, as:

$$\mathrm{JER} \equiv \frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{\mathrm{N}}{p_{\mathrm{T}}} \oplus \frac{\mathrm{S}}{\sqrt{p_{\mathrm{T}}}} \oplus \mathrm{C}. \tag{2.14}$$

The term "N" parametrizes the effect of electronic and pile-up noise and is mostly relevant at low p_{T}. The term "S" parametrizes stochastic effects, such as the sampling nature of the detector, and is mostly relevant at mid p_{T}. The term "C" parametrizes constant effects, such as losses in dead material, and is mostly relevant at high p_{T}. The goal is then to measure these terms.

The term "N" is split into its two contributing effects: $\mathrm{N} = \mathrm{N}_{\mathrm{pile-up}} \oplus \mathrm{N}_{\mathrm{electronic},\ \mu=0}$. The term $\mathrm{N}_{\mathrm{electronic},\ \mu=0}$ is measured using MC simulation where no pile-up is added. For the term $\mathrm{N}_{\mathrm{pile-up}}$, it should be determined using the nominal pile-up profile from data. It is found that the term $\mathrm{N}_{\mathrm{pile-up}}$ is difficult to be determined from in-situ measurements since it is relevant at low p_{T} where the uncertainties are high. Instead, the random cones method is used. Two random cones of a radius 0.4 are projected at a random value of η^1, η^2, ϕ^1 and $\phi^2 = \phi^1 + \pi$. The difference, $\Delta p_{\mathrm{T}} = p_{\mathrm{T}}^1 - p_{\mathrm{T}}^2$, is centered around zero since the pile-up noise in these cones is expected to balance. 68% confidence

interval is used to define the width, σ_{RC}, of the distribution since it cannot be well fitted by a Gaussian function. The term $N_{pile-up}$ is equal to $\sigma_{RC}/\sqrt{2}$. Lately, it was found that an additional term $N2/p_T^{1.5}$, a p_T dependent pile-up term, improved the description of the pile-up effect on the JER; more studies are currently being performed.

Fixing the term "N" found in previous methods, the two other terms, "S" and "C", are found by fitting the equation 2.14 to the in-situ distribution $JER_{in-situ}$. Figure 2.37 shows, for 8 TeV data, the JER as measured using the three in-situ methods displaying the compatibility between the measurements. The final fit using the function in Equation 2.14 is also included. The JER in data and simulation are then compared. If the JER in MC simulation is smaller than the JER in data, the resolution in simulation is deteriorated to match the one in data using Gaussian random fluctuation. If the JER in MC simulation is larger than the JER in data, no action is taken for the nominal MC samples, instead the difference is taken into account by the systematic uncertainties.

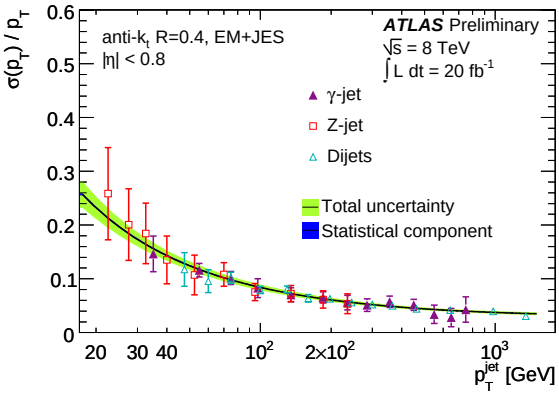
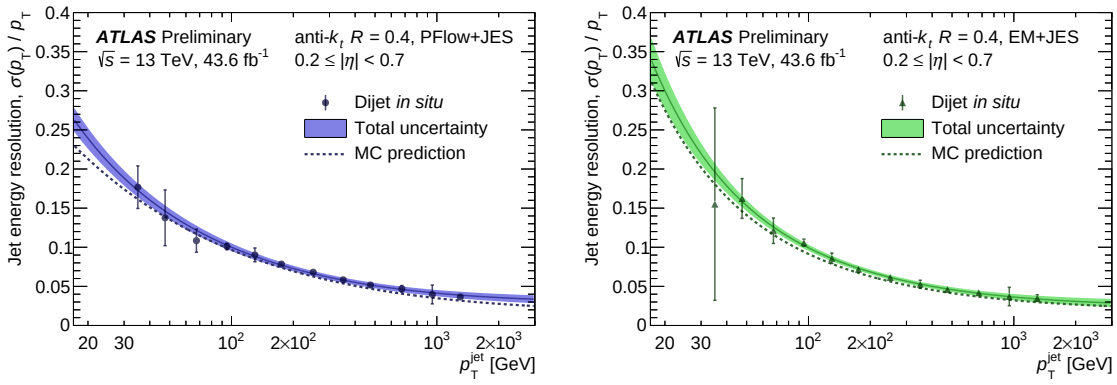


Figure 2.37: The jet resolution as a function of p_T for the 8 TeV data. The three in-situ inputs to the measurement are shown displaying the compatibility between the measurements. The final fit using the function in Equation 2.14 is included with its associated statistical and total uncertainty. Figure from [70].

Figure 2.38 shows, for 13 TeV data, the jet energy resolution $\sigma(p_T)/p_T$ as parametrized in equation 2.14, using random cones to evaluate the pile-up noise term and fitted to dijet in-situ results. Figure 2.39 compares the JER functions of EMPFlow and EMTopo when the former is smaller at low p_T and then becomes similar to the latter at mid and high p_T as expected.

Individual and total (summed in quadrature) systematic uncertainties on the evaluation of the JER are shown in figure 2.40 for $\eta = 0.2$.



(a) EMPFlow jets (b) EMTopo jets

Figure 2.38: Jet energy resolution, for 13 TeV Run II data, as a function of jet p_{T} parametrized as $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}} = \mathrm{N}/p_{\mathrm{T}} \oplus \mathrm{S}/\sqrt{p_{\mathrm{T}}} \oplus \mathrm{C}$. The pile-up noise term is evaluate using random cones. The dijet in-situ measurements, used to fit the JER function, are also shown. These results are for anti-kt jets with R=0.4 reconstructed from particle flow objects in a) and from topo-clusters in b). Figures from the ATLAS public plots page [71].

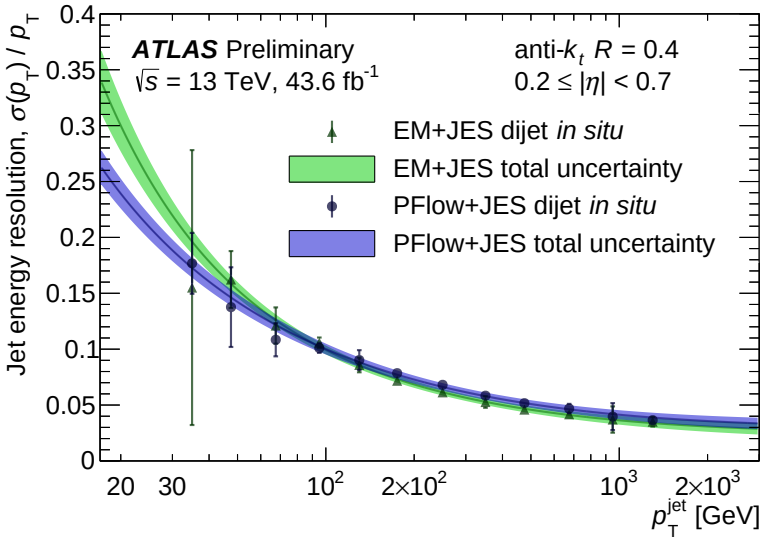
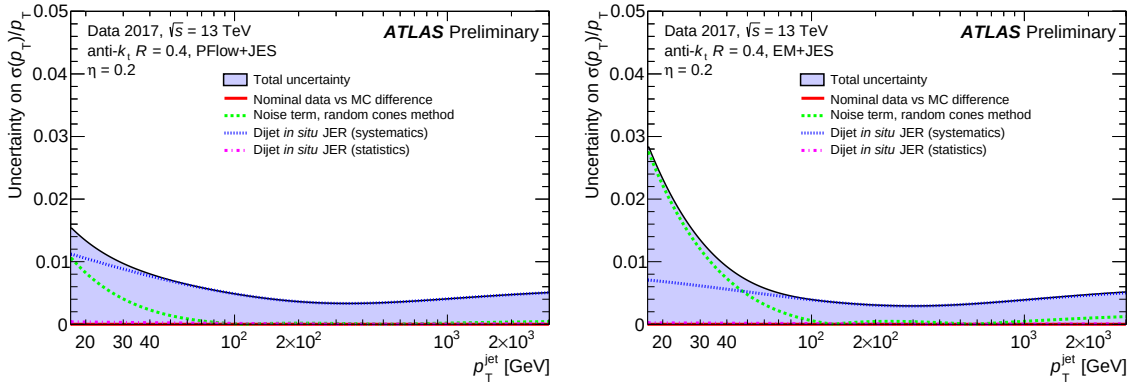


Figure 2.39: Jet energy resolution comparison between EMPFlow and EMTopo jets as a function of jet p_{T}. The JER is parametrized as $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}} = \mathrm{N}/p_{\mathrm{T}} \oplus \mathrm{S}/\sqrt{p_{\mathrm{T}}} \oplus \mathrm{C}$. Figure from the ATLAS public plots page [71].



(a) EMPFlow jets

(b) EMTopo jets

Figure 2.40: Systematic uncertainties on the jet energy resolution $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}}$ as a function of jet p_{T} for $\eta = 0.2$. These results are for EMPFlow jets in a) and for EMTopo jets in b). Figures from the ATLAS public plots page [71].

3 Eta-intercalibration

The very first step in term of in-situ jet calibration, using a comparison between the transverse momenta of two objects which are supposed to be balanced, is the calibration of the jets as a function of where they are in the detector, in the central part or in the forward regions. This calibration as a function of η and p_{T} of the jet is described in detail in this chapter.
In the first section, I present in general the eta-intercalibration correction and the different steps for measuring it.
In the following sections, I present the results of the different improvements, checks and tests I developed and implemented.
In the last section, I present the eta-intercalibration correction results I measured and that are used in the official ATLAS calibration for Run II.

3.1 Introduction

The first calibration to be applied among the in-situ methods is the eta-intercalibration [66, 67]. The aim, as said, is to calibrate jets in the forward region of the detector relative to jets in the central regions (reference objects). This is done by exploiting the transverse momentum balance of the two leading (highest p_{T}) jets (referred to as dijet system below) in multi-jets events. By comparing data to MC simulation, a 2D correction function is derived as a function of p_{T} and η_{det} of jets. η_{det} is the pseudo-rapidity measured from the detector center and not from the hard primary vertex (PV_0), i.e. before applying the origin correction (refer to section 2.3.3). Figure 3.1 shows the 2D binning.

Jets in the central region, defined as $|\eta_{\mathrm{det}}| < 0.8$, receive an average correction of 1 at this step of the calibration because they are the reference objects. Events are selected to contain hard back-to-back dijet topologies by requiring a minimum azimuthal separation between the two leading jets ($\Delta\phi^{12}$>2.5 rad) and by suppressing additional jets ($p_{\mathrm{T}}^{\mathrm{jet}\ 3}/p_{\mathrm{T}}^{\mathrm{avg}} < 0.25$).

The η-binning of the correction is chosen to follow the detector structure. Two strategies exists to derive this correction function. The simpler method, called the central reference method, uses dijet systems where one of the jets is required to be in the central region and the other is in the forward region. The central region is taken as one η bin and only the forward region is divided into multiple η bins. The second method, called the matrix method, uses all region combinations of dijet systems, central-central, central-forward or forward-forward dijet systems. Both central and forward regions are divided into multiple η bins. Both methods are briefly described in the following even if the central reference method is only used as a cross check of the matrix method.

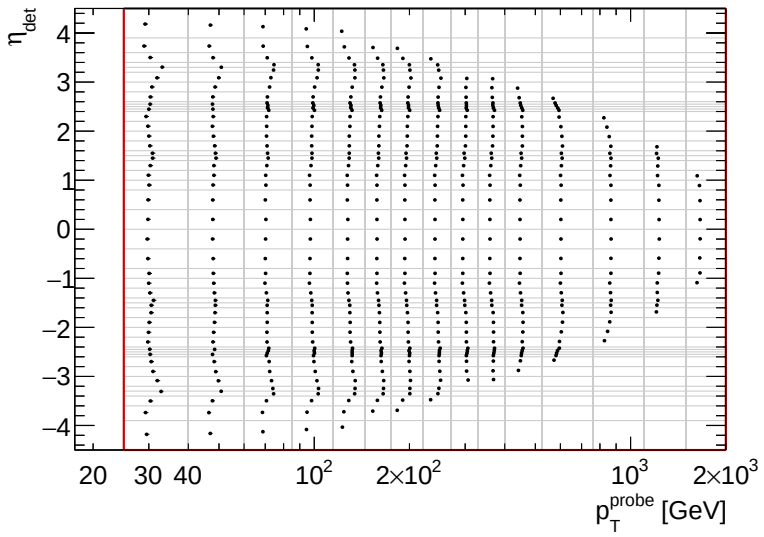


Figure 3.1: Eta-intercalibration 2D binning in p_T and η_{det} of jets. Bin edges are shown in grey lines. Black dots represent p_T and η averages in each bin.

3.1.1 Central reference method

In this method, we study the transverse momentum balance of dijet systems where one of the two leading jets is in the central region (with p_T^{ref}) and the other is in the forward region (with p_T^{probe}). The p_T balance can be expressed in terms of the asymmetry $\mathcal{A}$,

$$\mathcal{A} = \frac{p_T^{probe} - p_T^{ref}}{p_T^{avg}}, \tag{3.1}$$

$$\text{where } p_T^{avg} = (p_T^{probe} + p_T^{ref})/2. \tag{3.2}$$

The intercalibration factor c of the probe jet is then:

$$\frac{p_T^{probe}}{p_T^{ref}} = \frac{2+\mathcal{A}}{2-\mathcal{A}} = \frac{1}{c} \equiv \text{relative jet response.} \tag{3.3}$$

Dividing the measurements in bins of jet η_{det} and p_T^{avg}, with one η bin in the central region, the asymmetry distribution $\mathcal{A}_{ik}$ is evaluated for each probe jet η_{det} bin i and p_T^{avg} bin k. An asymmetry distribution example is shown in figure 3.2. The intercalibration factor is calculated using the formula:

$$c_{ik} = \frac{2+ < \mathcal{A}_{ik} >}{2- < \mathcal{A}_{ik} >},$$

where $< \mathcal{A}_{ik} >$ is the mean value of the asymmetry distribution in each bin.

Due to the constraints imposed on dijet systems, this method suffers from reduced statistics. It is only used as a cross check for the matrix method described in the following section.

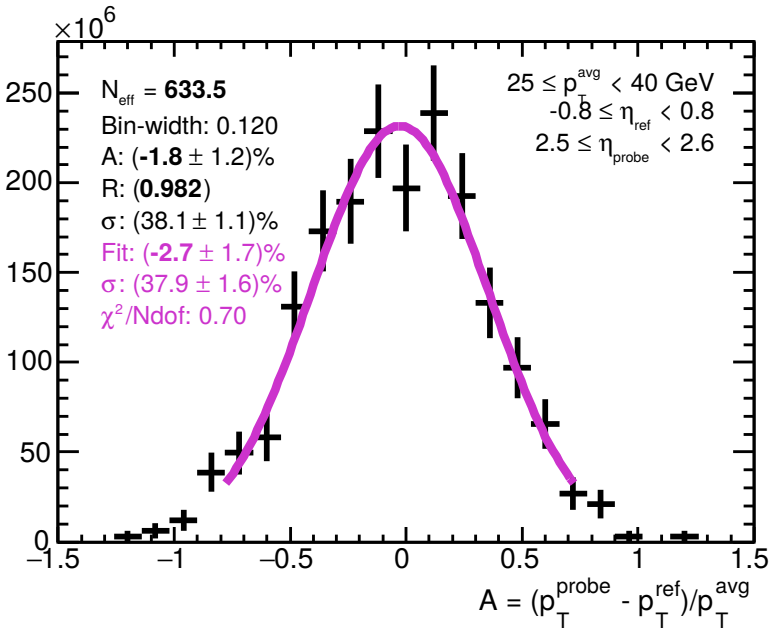


Figure 3.2: Asymmetry $\mathcal{A}$ distribution shown in black points for one bin of $p_{\mathrm{T}}^{\mathrm{avg}}$, η_{ref} and η_{probe}. The Gaussian fit in magenta line is used to extract the asymmetry mean $< \mathcal{A}_{\mathrm{ik}} >$.

3.1.2 Matrix method

The matrix method uses all dijet systems regardless of the η regions of its 2 jets. To do so, each dijet event is assigned a "left" and "right" jets, defined by $\eta_{\mathrm{det}}^{\mathrm{left}} < \eta_{\mathrm{det}}^{\mathrm{right}}$. The new equations are:

$$\mathcal{A} = \frac{p_{\mathrm{T}}^{\mathrm{left}} - p_{\mathrm{T}}^{\mathrm{right}}}{p_{\mathrm{T}}^{\mathrm{avg}}} \tag{3.4}$$

$$\mathcal{R} = \frac{\mathrm{c}^{\mathrm{right}}}{\mathrm{c}^{\mathrm{left}}} = \frac{p_{\mathrm{T}}^{\mathrm{left}}}{p_{\mathrm{T}}^{\mathrm{right}}} = \frac{2 + \mathcal{A}}{2 - \mathcal{A}}. \tag{3.5}$$

where $\mathcal{R}$ is the ratio of responses. Using the binning system as before but with several bins inside the reference region, $\mathcal{R}_{\mathrm{ijk}}$ is obtained for each $\eta_{\mathrm{det}}^{\mathrm{left}}$ bin i, $\eta_{\mathrm{det}}^{\mathrm{right}}$ bin j, and $p_{\mathrm{T}}^{\mathrm{avg}}$ bin k. An example of η bins combination is shown in figure 3.3. For each fixed $p_{\mathrm{T}}^{\mathrm{avg}}$ bin k with N η bins, the corresponding intercalibration factors $\mathrm{c_{ik}}$, with i = 1...N, are obtained simultaneously by minimizing the following function [66]:

$$\mathrm{S}(\mathrm{c_{1k}}, ..., \mathrm{c_{Nk}}) = \sum_{\mathrm{j=1}}^{\mathrm{N}} \sum_{\mathrm{i=1}}^{\mathrm{j-1}} \left(\frac{1}{\Delta < \mathcal{R}_{\mathrm{ijk}} >} (\mathrm{c_{ik}} < \mathcal{R}_{\mathrm{ijk}} > - \mathrm{c_{jk}}) \right)^2 + \mathrm{X}(\mathrm{c_{1k}}, ..., \mathrm{c_{Nk}}), \tag{3.6}$$

where $\Delta < \mathcal{R} >$ is the statistical uncertainty of $< \mathcal{R} >$. The X function defined by,

$$\mathrm{X}(\mathrm{c_{1k}}, ..., \mathrm{c_{Nk}}) = \lambda \left(\frac{1}{\mathrm{N}} \sum_{\mathrm{i=1}}^{\mathrm{N}} \mathrm{c_{ik}} - 1 \right)^2, \tag{3.7}$$

is used to fix the value of the average of c_{ik}, since the normalization of c_{ik} factors is not constrained by the minimization of the first term of the function S. In fact, multiplying all c_{ik} by the same factor is also a solution to the minimization of this first term. Also, the X function prevents the minimizing algorithm to choose the trivial solution, which is all $c_{ik} = 0$. The λ factor is a Lagrange multiplier, an arbitrary chosen constant that does not affect the solution (as long as it is sufficiently large for numerical stability, e.g. $\lambda \approx N$). The resulting factors c_{ik} are scaled afterward such that the simple average of the factors in the central region is equal to 1.

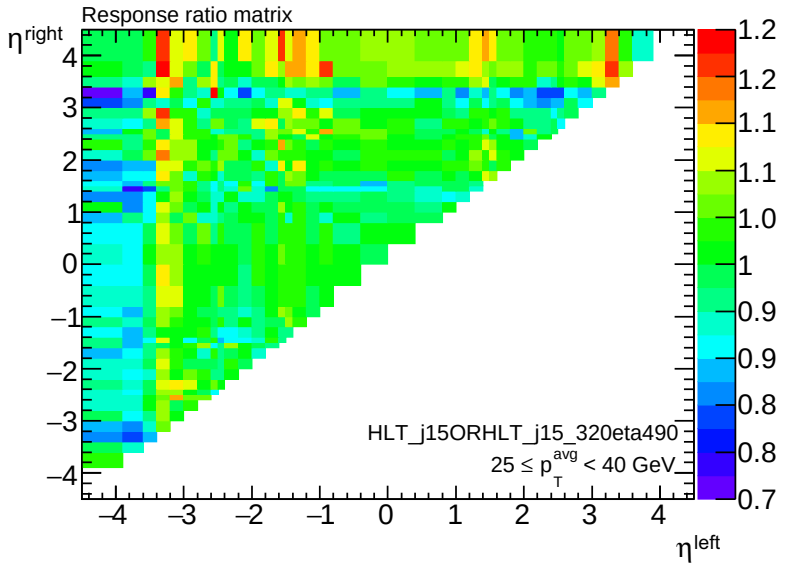


Figure 3.3: Response ratios $\mathcal{R}_{ijk}$ as a function of η_{det}^{left} and η_{det}^{right} showing the η bins combination used in the matrix method.

The minimization of the function S gives the central values of the factors c_i. In order to evaluate the uncertainties on these values, 1000 random fluctuations ("toys") of $\mathcal{R}$ are used to propagate their uncertainties to the factors c_i in the following way:

- for each toy, replace:
$$\mathcal{R} \rightarrow \mathcal{R} + \Delta\mathcal{R} \times G(0, 1),$$
where $G(0, 1)$ is a random Gaussian number with mean = 0 and sigma = 1,
- minimize the function S using the modified values of $\mathcal{R}$,
- denoting $c_i^{(t)}$ the result from each toy t, the uncertainty of each c_i factor is:
$$\Delta c_i = \mathrm{RMS}(c_i^{(t)}).$$

Figure 3.4 shows in black points the factors c_i as a function of η for one bin of p_T^{avg}. For each c_i, the results from the fluctuations of $\mathcal{R}$ are shown in colored lines.

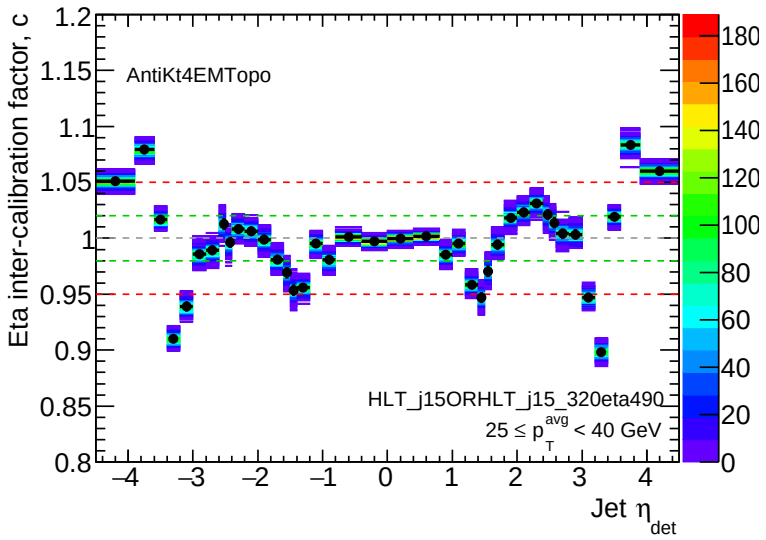


Figure 3.4: Eta-intercalibration factors c_i as a function of η for one bin of p_T^{avg}. The results of each c_i from the fluctuations of $\mathcal{R}$, which are used to measure the statistical uncertainty Δc_i, are shown in colored lines.

Previously, the MINUIT algorithm [72], a numerical minimizer, was used to minimize the function S. Two draw backs of numerical methods are:

- the minimization becomes very slow with large number of variables (η bins),
- the minimization does not always converge.

For those reasons, I developed and implemented a better analytic solution. The details are discussed in section 3.2.

3.1.3 Residual correction

For each (p_T^{avg}, η_{det})-bin denoted by the index i, the residual correction is derived from the data/MC simulation ratio:

$$C_i = \frac{c_i^{Data}}{c_i^{MC}}. \tag{3.8}$$

In order to smooth statistical fluctuations and to have a continuous calibration, the correction function F_{rel} is defined as a two-dimensional Gaussian kernel by:

$$F_{rel}(p_T, \eta_{det}) = \frac{\sum_{i=1}^{N_{bins}} C_i w_i}{\sum_{i=1}^{N_{bins}} w_i}, \tag{3.9}$$

$$\text{with} \quad w_i = \frac{1}{\Delta C_i^2} \times \text{Gaus}\left(\frac{\log p_T - \log < p_T^{probe} >_i}{\sigma_{\log p_T}} \oplus \frac{\eta_{det} - < \eta_{det} >_i}{\sigma_{\eta_{det}}} \right), \tag{3.10}$$

where ΔC_i is the statistical uncertainty of C_i, $< p_T^{probe} >_i$ and $< \eta_{det} >_i$ are the averages of p_T and η in the bin i. The Gaussian has a zero mean value. Its width is controlled by $\sigma_{\log p_T}$ and $\sigma_{\eta_{det}}$, where these parameters are chosen to capture the shape of the ratio data to MC simulation, and to provide stability against statistical fluctuations (more details in the results section 3.6.3). Those parameters are tuned using the closure of the calibration (defined in section 3.1.5). The weights w_i of the kernel are inversely proportional to the variance of the C_i ratios, such that more precise ratio bins have more power in the function F_{rel}.

Although residual correction points are calculated in bins of p_T^{avg}, the calibration is applied to jets in bins of p_T^{probe}. To take this difference into account, a numerical inversion method is used. Using equations 3.2 and 3.3, we get:

$$< p_T^{probe} >= 2 < p_T^{avg} > *R/(R+1), \tag{3.11}$$

where R is the relative jet response.

An example of the calibration is shown in figure 3.5.

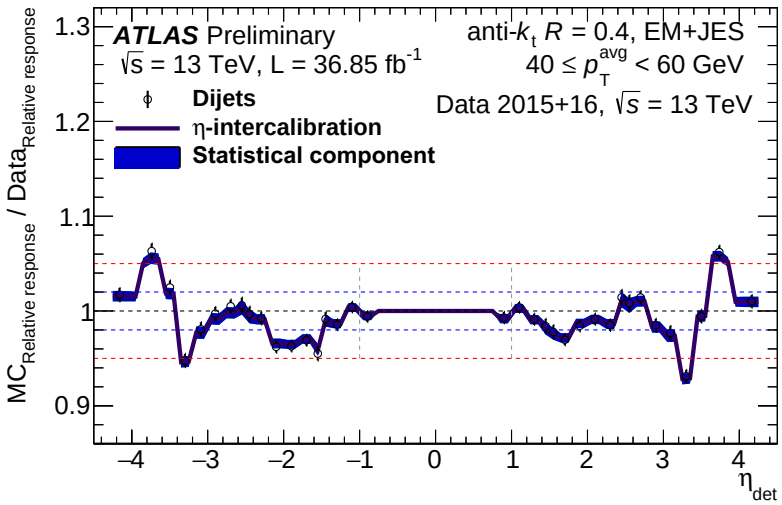


Figure 3.5: The eta-intercalibration correction function F_{rel} as a function of η is shown in magenta line with its statistical uncertainty shown in blue shade. The individual correction points C_i are shown as black open circles.

3.1.4 Systematic uncertainties

The difference in the relative jet response between data and MC simulation comes from:

- mis-modeling in detector simulation,
- mis-modeling of physics,

- event selection criteria.

Except for the first effect, which is what we want to measure and correct for, we need to evaluate systematic uncertainties of all the other effects on the calibration results.

Different MC generators have different event topologies and jet sub-structure. They differ especially on the parton production mechanism (like production order) and on the showering model which mostly impact the additional jet kinematics. Since there is a priori no reason to trust one generator more than the other, we use two generators, one for nominal calibration and one for systematic uncertainty evaluation. The calibration is derived using each of the two generators and the difference between the two calibrations is taken as a systematic uncertainty.

To select hard back-to-back dijet systems, event selection cuts are applied on $\Delta\phi^{12}$, $p_{\mathrm{T}}^{\mathrm{jet}\ 3}/p_{\mathrm{T}}^{\mathrm{avg}}$ and JVT. To evaluate the calibration dependence on those variables, we do the following for each of the selection cuts:

- vary the nominal cut value up and down (loose and tight) simultaneously for data and MC simulation,
- derive the variated residual calibrations,
- compare the up and down variations results to the nominal one, and take the maximum difference as a systematic uncertainty.

An important thing we want to avoid is that our systematic uncertainty evaluation is significantly affected by statistical fluctuations. That is why we verify that the systematic uncertainty values are significantly larger than their statistical uncertainties.
The nominal and variated selection have a high proportion of common events and their calibration results are thus highly correlated. To correctly calculate the statistical uncertainties of the systematic uncertainties we use the bootstrap method described below.

Bootstrap method
The bootstrap method [73] is used to correctly propagate uncertainties where the existence of complex correlations make it difficult/impossible to do it analytically. Correlations exist when the same events are used in multiple places (in systematic uncertainty evaluation for example), when doing a smoothing, when folding or unfolding, etc.
The method works as follows:

- generate N sample toys: for each event and for each toy, generate a random integer Poisson fluctuation with the mean parameter equal to one, n = Poisson(1), and fill the event "n" times,
- for each sample toy t: calculate the variable in question $x^{(t)}$,
  for example for the systematic evaluation case, calculate for each sample toy the nominal and variated calibrations and take the difference,
- the uncertainty of the variable is the RMS of the toys results or the sigma of the Gaussian fit:

$$\Delta x = \mathrm{RMS}(x^{(t)}) \text{ or } \sigma(\text{Gaussian fit}(x^{(t)})),$$

- the number of toys N should be big enough to reduce the uncertainty on the value of Δx which is proportional to $1/\sqrt{N}$, N=1000 is a reasonable choice.

An important step is the choice of the seed of the random number generator. The seed should be different for the different events but should be the same when the same event is used in different places or analyses. For that, the event number and run number (and MC channel number in case of a simulation) are used to correctly define the seed.

Figure 3.6 shows an example of uncertainties measurement using the bootstrap method. The results from 1000 sample toys are filled in a histogram. The RMS and the sigma of the Gaussian fit are shown.

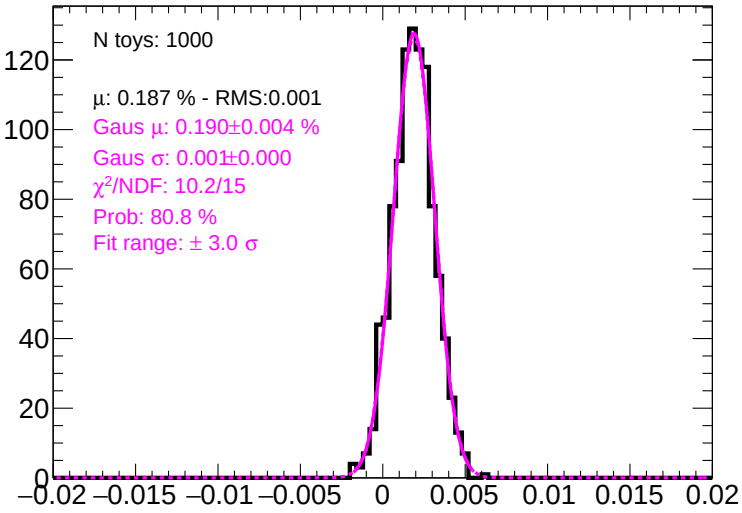


Figure 3.6: Example of measuring uncertainties using bootstrap method.

3.1.5 Closure test

To check whether the calibration have successfully detected and corrected all the differences in the jet responses, the calibration closure is tested. For that, we:

- derive the residual calibration,
- apply it to data only,
- re-derive a second residual calibration using the corrected data,
- a good closure is obtained when this second calibration is compatible with unity,
- any deviation is taken as an additional non-closure systematic uncertainty.

Figure 3.7 shows an example of this closure as a function of η for one $p_{\mathrm{T}}^{\mathrm{avg}}$ bin. Only small deviations from unity are seen. More details and results are discussed in results section 3.6.
The closure test is also used to choose the parameters of the smoothing kernel, $\sigma_{\log p_{\mathrm{T}}}$ and $\sigma_{\eta_{\mathrm{det}}}$ (see 3.1.3). Multiple choices for these parameters are tested and the combination that gives the smallest non-closure, $\sigma_{\log p_{\mathrm{T}}} = 0.18$ and $\sigma_{\eta_{det}} = (\eta$ bin width)/6, is chosen (more details in section 3.6.3).

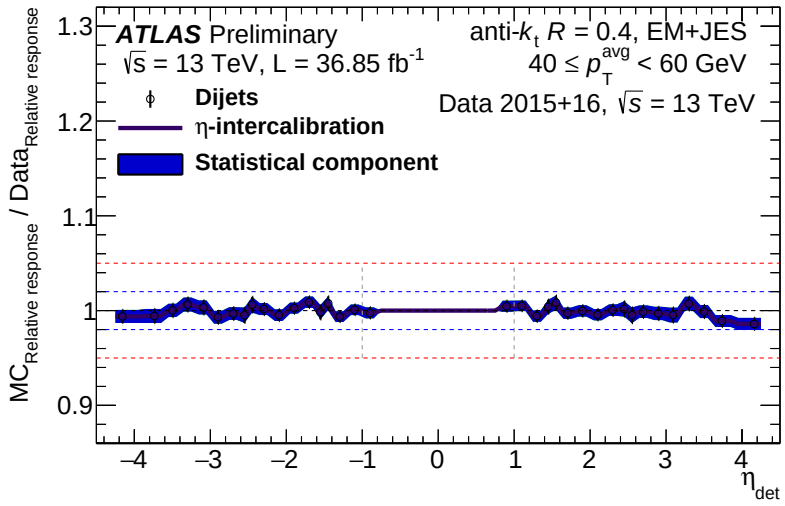


Figure 3.7: Calibration closure as a function of η for one $p_{\mathrm{T}}^{\mathrm{avg}}$ bin.

3.2 Analytic solution

Previous calibrations are found to not have sufficiently fine η binning to be able to detect narrow peaks in the jet response. This is visible in figure 3.8 where the ratio between jet η distributions from data and MC simulation shows a narrow and significant peak at $2.4 < |\eta| < 2.5$. This localized disagreement cannot be explained by physics causes and is due to a calibration with wide binning where jet responses are averaged. Finer η binning is to be used. As previously mentioned, the numerical method for minimizing the equation S (3.6) has the inconvenience of being very slow with large number of variables (η bins) and not converging for some cases.

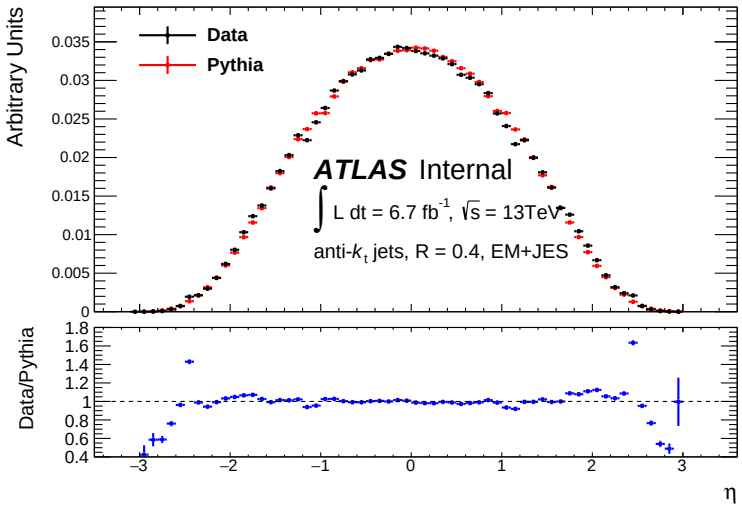


Figure 3.8: Comparison of jet η distributions between data and simulation for old calibration.

To address those inconveniences, I developed and implemented a better analytic solution as follows. The equation S, which was introduced in the matrix method paragraph, depends on N variables c_i. The minimum of S satisfies the conditions:

$$\frac{\partial S}{\partial c_i} = 0 \text{ , for i = 1 ,.., N .} \tag{3.12}$$

Expending the derivative with respect to c_α, the following is found (the p_T^{avg} index "k" in $\mathcal{R}_{ijk}$ is omitted hereafter for clarity):

$$\sum_{i=1}^{\alpha-1}\left(\left(\frac{-\mathcal{R}_{i\alpha}}{\Delta^2\mathcal{R}_{i\alpha}}+\frac{\lambda}{N^2}\right)c_i\right)+\left(\sum_{i=1}^{\alpha-1}\frac{1}{\Delta^2\mathcal{R}_{i\alpha}}+\sum_{i=\alpha+1}^{N}\frac{\mathcal{R}_{\alpha i}^2}{\Delta^2\mathcal{R}_{\alpha i}}+\frac{\lambda}{N^2}\right)c_\alpha+\sum_{i=\alpha+1}^{N}\left(\left(\frac{-\mathcal{R}_{\alpha i}}{\Delta^2\mathcal{R}_{\alpha i}}+\frac{\lambda}{N^2}\right)c_i\right)-\frac{\lambda}{N}=0\,. \tag{3.13}$$

Taking α = 1 ,.., N and writing the system of equations in matrix form, we get:

$$\underbrace{\begin{pmatrix} \sum_{i=2}^{N} \frac{\mathcal{R}_{1i}^2}{\Delta^2 \mathcal{R}_{1i}} + \frac{\lambda}{N^2} & \cdots & \frac{-\mathcal{R}_{1\alpha}}{\Delta^2 \mathcal{R}_{1\alpha}} + \frac{\lambda}{N^2} & \cdots & \frac{-\mathcal{R}_{1N}}{\Delta^2 \mathcal{R}_{1N}} + \frac{\lambda}{N^2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{-\mathcal{R}_{1\alpha}}{\Delta^2 \mathcal{R}_{1\alpha}} + \frac{\lambda}{N^2} & \cdots & \sum_{i=1}^{\alpha-1} \frac{1}{\Delta^2 \mathcal{R}_{i\alpha}} + \sum_{i=\alpha+1}^{N} \frac{\mathcal{R}_{\alpha i}^2}{\Delta^2 \mathcal{R}_{\alpha i}} + \frac{\lambda}{N^2} & \cdots & \frac{-\mathcal{R}_{\alpha N}}{\Delta^2 \mathcal{R}_{\alpha N}} + \frac{\lambda}{N^2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{-\mathcal{R}_{1N}}{\Delta^2 \mathcal{R}_{1N}} + \frac{\lambda}{N^2} & \cdots & \frac{-\mathcal{R}_{\alpha N}}{\Delta^2 \mathcal{R}_{\alpha N}} + \frac{\lambda}{N^2} & \cdots & \sum_{i=1}^{N-1} \frac{1}{\Delta^2 \mathcal{R}_{iN}} + \frac{\lambda}{N^2} \end{pmatrix}}_{U} \times \underbrace{\begin{pmatrix} c_1 \\ \vdots \\ c_\alpha \\ \vdots \\ c_N \end{pmatrix}}_{C} = \underbrace{\begin{pmatrix} \frac{\lambda}{N} \\ \vdots \\ \frac{\lambda}{N} \\ \vdots \\ \frac{\lambda}{N} \end{pmatrix}}_{A} \tag{3.14}$$

The solution to the vector C of the correction factors c_i is obtained through a matrix inversion:

$$U \times C = A \quad \rightarrow \quad C = U^{-1} \times A \tag{3.15}$$

This analytic method gives similar results for c_i factors as the minimization method. Figure 3.9 shows the level of compatibility between the two results which for most of the bins are not distinguishable. Still, it has the advantage of always converging and most importantly that it is much faster than the numerical minimization-based method. In fact for the number of c_i variables that we will typically have (30-50 variables), the analytic method is more than one thousand times faster (see figure 3.10).

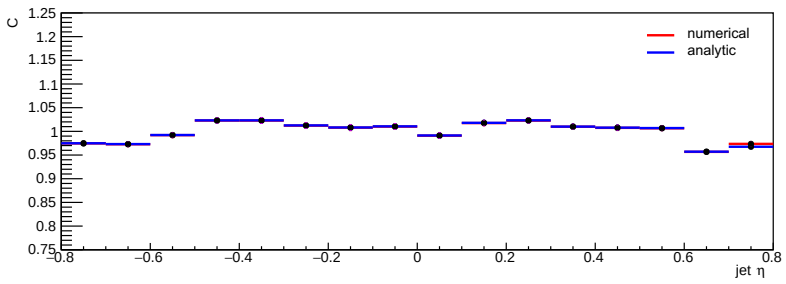


Figure 3.9: Comparison of c_i results from the two methods of minimization: numerical and analytic.

An additional check I did is a closure test on the calculation of intercalibration factors. c_i values are pre-defined and R_{ij} factors are calculated accordingly. The equation S defined using those R_{ij} as inputs is then minimized and c_i factors are computed. The difference between $c_i^{computed}$ and $c_i^{injected}$ is shown in figure 3.11. The two methods are compatible with each other and give a non-closure smaller than 1 per 10 thousand. The non-closure is much smaller than the statistical uncertainties.

It is worth noting that I tried other methods to define the equation S. In particular, I tested a new approach to get rid of the need of the function X (which is used to impose a normalization) by changing

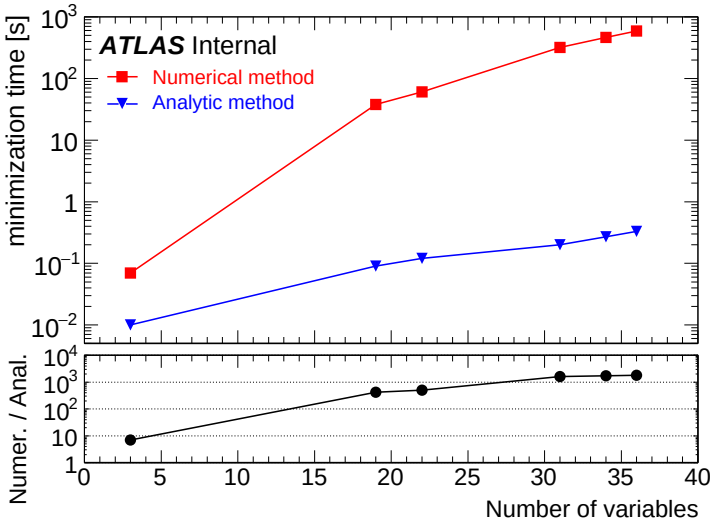


Figure 3.10: Minimization speed comparison between numerical and analytic minimization-based methods of the function S (3.6) as a function of the number of c_i variables (noting that the time plotted is the total time to minimize 1000 times the function S corresponding to the 1000 random fluctuations needed to propagate the uncertainties).

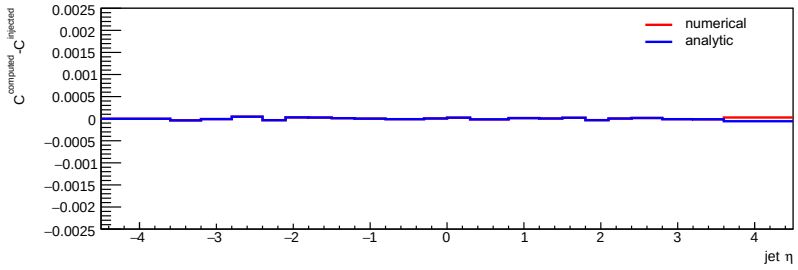


Figure 3.11: Closure test on the calculation of intercalibration factors by comparing $c_i^{computed}$ (for each of the 2 minimization methods) and $c_i^{injected}$.

the equation variables to ratios between intercalibration factors c_i with respect to a reference c_{ref}:

$$S'(\frac{c_1}{c_{ref}}, ..., \frac{c_N}{c_{ref}}) = \sum_{j=1}^{N} \sum_{i=1}^{j-1} \left(\frac{1}{\Delta < \mathcal{R}_{ij} >} (\frac{c_i}{c_{ref}} < \mathcal{R}_{ij} > - \frac{c_j}{c_{ref}}) \right)^2 . \tag{3.16}$$

The S' function minimization is found to be dependent on the choice of the c_{ref}. Looping over all N factors c_i, taking each one as a c_{ref} and minimizing S', the average of the N results for each c_i is found to be compatible with the result of S minimization. The drawback is having to repeat the minimization N times which degrades the speed of the process which is the main issue I was trying to reach.

3.3 Choice of Monte-Carlo generator

The choice of MC generators (one for the nominal calibration, and one for the evaluation of a systematic uncertainty) is based on two criteria:

- the modeling of the two leading jets which are used to define the asymmetry of the p_{T} balance,
- the modeling of the third jet that is used in the event selection and thus has an impact on the event topologies.

For the modeling of the two leading jets, NLO generators (Powheg-Pythia, Powheg-Herwig, Sherpa) are superior to the LO generators (Pythia, Herwig). To backup this claim, one case study is shown in section 3.3.2. The choice using this criteria is straightforward.

For LO generators, the third jet comes from the showering of the two leading partons.
For NLO generators, in addition we have third jets coming directly from the matrix elements. Therefore, a p_{T} cut is defined to split between the two production mechanisms to avoid double counting. One would assume that NLO generators would be better at modeling the third jet production, but as we will see a good tuning for the mentioned cut is needed otherwise the modeling will give worse predictions.

The following MC generators (refer to chapter 1.3.2) are used to compare to data:

- Pythia8 (LO)
- Herwig++ (LO)
- PowhegPythia8 (NLO)
- Sherpa (NLO)
- PowhegHerwig (NLO)

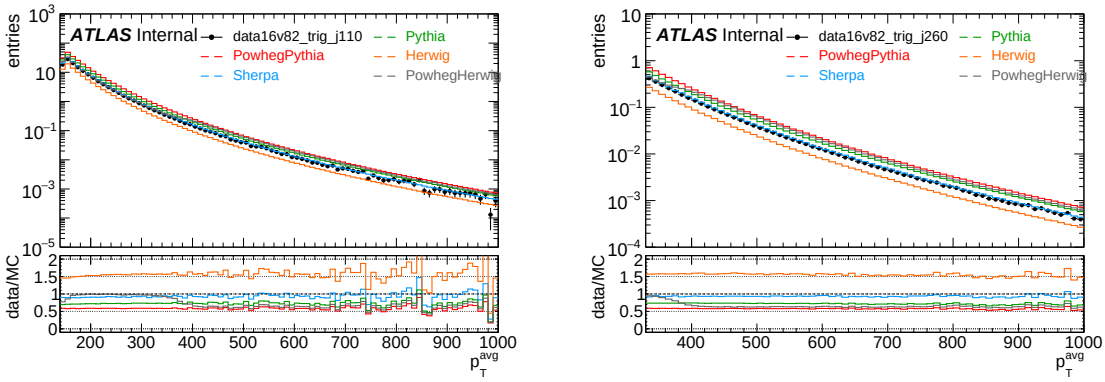
3.3.1 $p_{\mathrm{T}}^{\mathrm{avg}}$ distributions

First, the $p_{\mathrm{T}}^{\mathrm{avg}}$ distributions should be checked to make sure that they are smooth and that the MC simulation ones have a similar shape as the data one. A global shift has no effect on eta-intercalibration results. On the other hand, if the shapes are different, meaning the distribution of $p_{\mathrm{T}}^{\mathrm{avg}}$ inside each bin is different, and since the calibration is p_{T} dependent, the final results will include a bias from this difference.

Figure 3.12 shows, for data/MC simulation, the following $p_{\mathrm{T}}^{\mathrm{avg}}$ distributions ratio:

- Pythia, PowhegPythia and Sherpa: the ratio is stable within fluctuations,
- Herwig: the ratio is stable except for $p_{\mathrm{T}} < 80\mathrm{GeV}$,
- PowhegHerwig: the ratio is stable except a drop for $350 < p_{\mathrm{T}} < 400\mathrm{GeV}$.

The two drops in the ratios are due to a slice normalization problems, but overall the shapes are compatible with the one from the data.



(a) Trigger HLT_j110

(b) Trigger HLT_j260

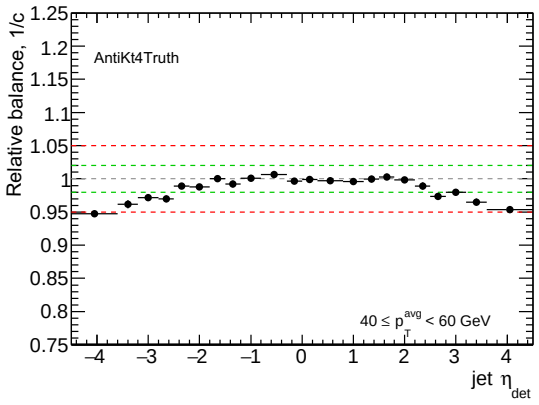
Figure 3.12: p_T^{avg} distributions for various MC generators and comparison to data. Single jet triggers are used, one firing at 110 GeV and the other at 260 GeV.

3.3.2 Truth level relative jet balance

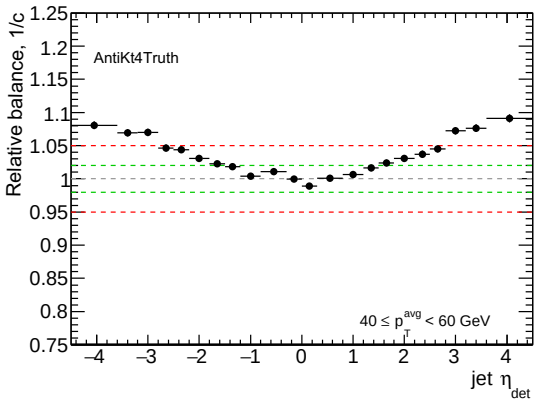
Even if the truth information is not available for data, it is useful to compare the truth-level relative jet balance for the different generators. This is shown in figure 3.13 as function of jet η.
Without specific studies, we cannot say that a flat relative balance at truth level is the correct one. In fact, physics effects could induce a deviation from unity (for example color connection between outgoing partons).
Nonetheless, we see that LO generators, Pythia and Herwig, deviate significantly from unity and in opposing direction. The differences between the two generators are more than 10% in forward bins. Contrarily, NLO generators have a closer behavior.

In the past when LO generators were used to derive eta-intercalibration corrections, high uncertainties were observed in the forward region because of this differences of balances at truth level.

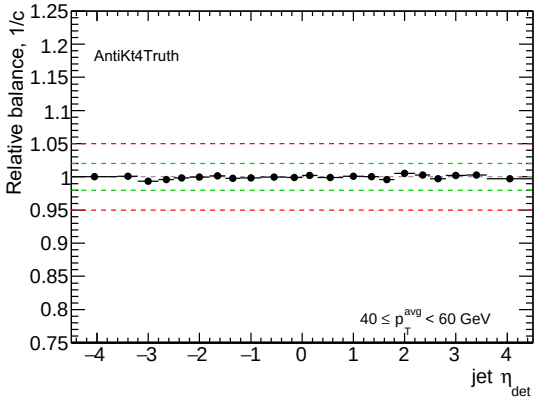
From figure 3.13, we can also naively guess that forward jets radiate harder third jets in Pythia, whereas in Herwig central jets radiate harder third jets.



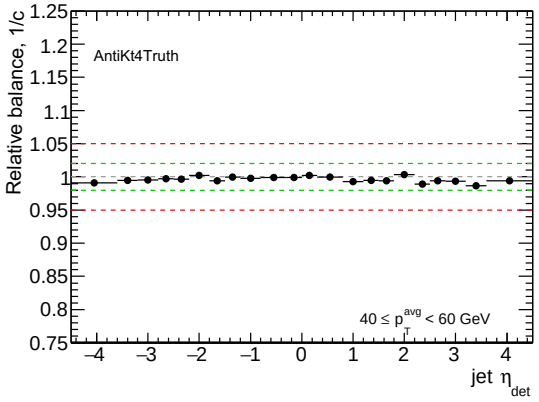
(a) Pythia



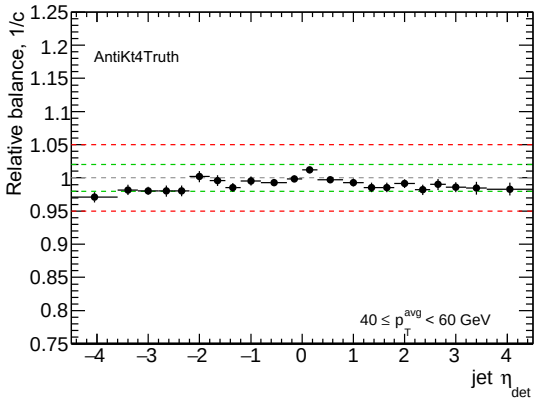
(b) Herwig



(c) PowhegPythia



(d) PowhegHerwig

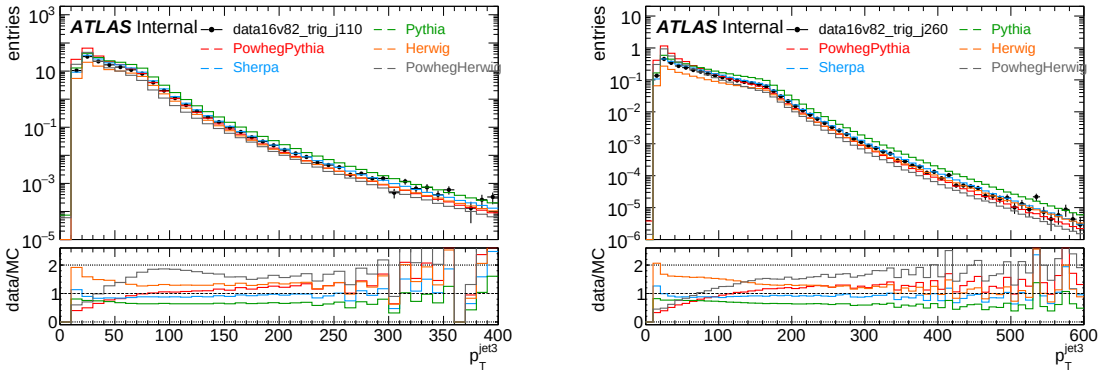


(e) Sherpa

Figure 3.13: Truth jets Relative balances, for $40 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 60$ GeV.

3.3.3 $p_{\mathrm{T}}^{\mathrm{j}\,3}$ distributions

$p_{\mathrm{T}}^{\mathrm{j}\,3}$ distributions give us a first global evaluation of how well the third jet is modeled.



(a) Trigger HLT_j110, $p_{\mathrm{T}}^{\mathrm{avg}} > 145$GeV

(b) Trigger HLT_j260, $p_{\mathrm{T}}^{\mathrm{avg}} > 330$GeV

Figure 3.14: $p_{\mathrm{T}}^{\mathrm{j}\,3}$ distributions for various MC generators and comparison to data. Single jet triggers are used, one firing at 110 GeV and the other at 260 GeV.

Figure 3.14 shows, for data/MC simulation, the following $p_{\mathrm{T}}^{\mathrm{j}\,3}$ distributions ratio:

- Sherpa: the ratio is stable within fluctuations,
- Pythia and Herwig: the ratio slowly decreasing toward high p_{T},
- PowhegPythia: the ratio varies more significantly, decreasing toward low p_{T},
- PowhegHerwig: same behaviour as PowhegPythia but with larger variations.

Already here we can see that the Powheg generator has a mis-modeling for the third jet, giving more soft jets compared to data (third jet distributions are higher at low p_{T} for this generator).

3.3.4 The dependence of the asymmetry on $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$

The calibration takes as input the asymmetry of the balance between the two leading jets. Hence, the more direct question of what is the effect of any mis-modeling of the third jet on the calibration result should be asked. For that, I plot the asymmetry mean (using the matrix method definition) as a function of $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$ and compare the MC simulation and the data.

Taking one range of $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$ and plotting the asymmetry distribution, figure 3.15 shows the first complication with a distribution with a double peak. This is expected since we request a hard third jet which impacts the balance of the two leading jets through the global momentum conservation

and since most third jets originate from a parton radiated by one of the two leading jets, thus the asymmetry value is shifted to negative values in one case and to positive values in the other.

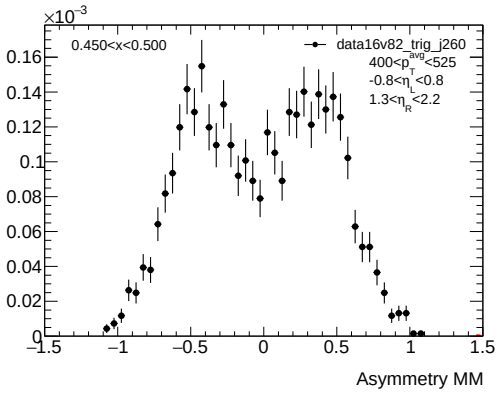


Figure 3.15: Asymmetry distribution of the two leading jets with a requirement of $0.45 \leq p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}} < 0.5$.

This is why it is important to distinguish those two cases. This is done by comparing the azimuthal separation $\Delta\phi$ between the third jet and each of the two leading jets. The third jet is considered coming from the closest leading jet. Recalling that the two leading jets are assigned to be left or right based on their pseudo-rapidities ($\eta^{\mathrm{L}} < \eta^{\mathrm{R}}$), we split the events into two groups:

- L tag for the third jet: $\Delta\phi$(third jet, Left leading jet) $<$ $\Delta\phi$(third jet, Right leading jet),
- R tag for the third jet: $\Delta\phi$(third jet, Left leading jet) $>$ $\Delta\phi$(third jet, Right leading jet).

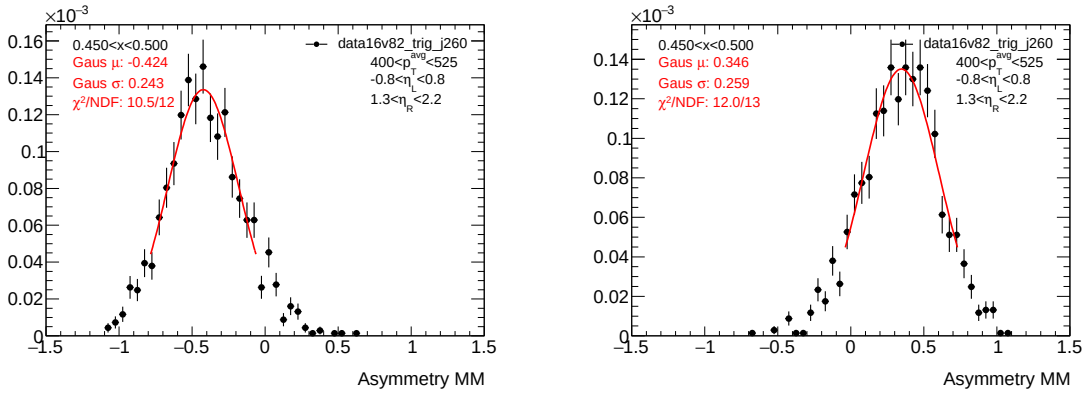
Figure 3.16 shows that this splitting works very well into separating the two peaks which are well fitted by two Gaussian functions. As a result, the asymmetry mean can now be extracted for each of the tagged groups.

The plot of the asymmetry mean as a function of $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$ for one of the tags is shown in figure 3.17(a). Note that the negative $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$ bins correspond to the case where we have no third jet information in our event. This happens for events where the third jet does not exist or most likely that its p_{T} is smaller than the minimum requirement to save the jet (10 GeV). The asymmetry differences between data and MC simulations seen across all the bins come from two sources:

- mis-modeling of the detector simulation affecting the balance between the two leading jets, which is what I want to correct in eta-intercalibration,
- mis-modeling of the third jet production, which is what I want to evaluate in this test.

To isolate the mis-modeling of the third jet, I do the following for each MC simulation:

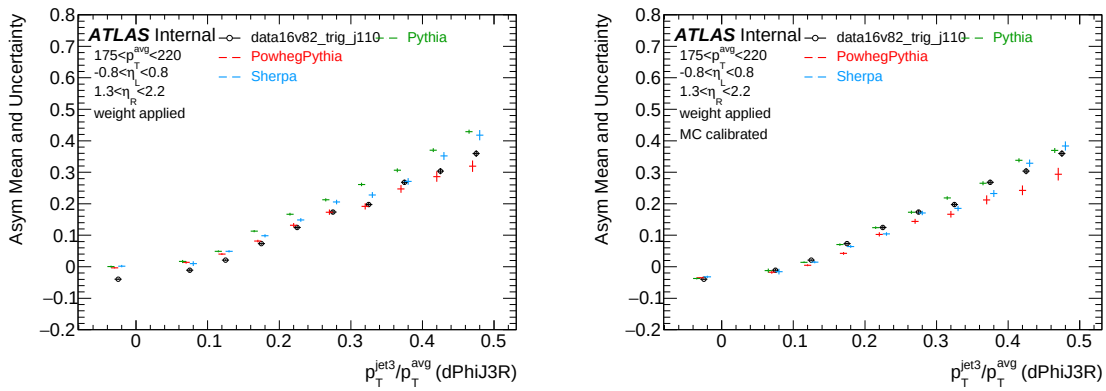
- select events with a very tight cut on third jet: no third jet in the case of PowhegPythia, Pythia and PowhegHerwig, $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}} < 0.1$ for Sherpa and Herwig (since they have low statistics),
- compare to data and derive eta-intercalibration correction,



(a) L tag (b) R tag

Figure 3.16: Asymmetry distributions of the two leading jets with a requirement of $0.45 \leq p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}} < 0.5$, same as the distribution in figure 3.15 but, split into 2 groups depending whether the third jet is closer in ϕ to left or right leading jet.

- calibrate the MC simulation and re-derive the asymmetry distributions.



(a) Before calibrating MC simulation (b) After calibrating MC simulation

Figure 3.17: Asymmetry mean as a function of $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$. In 3.17(b), the MC simulation is calibrated to isolate the mis-modeling of the third jet production from the mis-modeling of the detector simulation. Negative $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$ bins correspond to the case with no third jets ($p_{\mathrm{T}}^{\mathrm{j}\,3}$ smaller than the threshold).

Figure 3.17(b) shows the result after applying this correction. As expected, the MC simulation asymmetry values in the case of no third jet (negative $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$) become similar to the one of data.

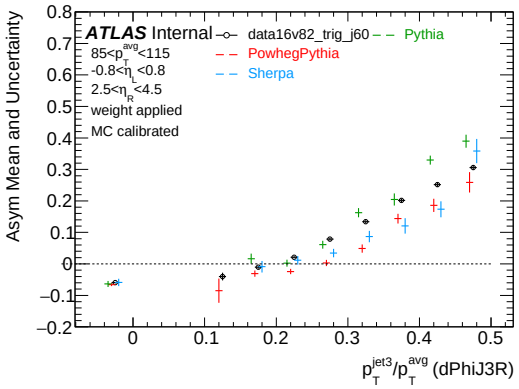
Figures 3.18 and 3.19 show plots of the asymmetry as a function of $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$, after calibrating the MC

simulation to match the data in the case of no third jet. The results shown are for three $p_{\mathrm{T}}^{\mathrm{avg}}$ regions, for both R and L tags. It is important to note that in those plots η left, η^{L}, region is central and η right, η^{R}, is forward.

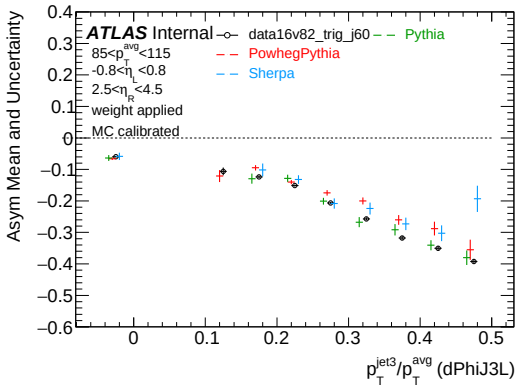
The plots show the following for the different generators:

1. Pythia:
   - figures 3.18(b), 3.18(d) and 3.18(f): a good agreement with data for L tag third jet,
   - figures 3.18(a), 3.18(c) and 3.18(e): higher asymmetries for R tag,
2. Herwig:
   - figures 3.19(a), 3.19(c) and 3.19(e): a good agreement with data for R tag third jet,
   - figures 3.19(b) and 3.19(d): higher absolute asymmetries for L tag,
3. Sherpa:
   - figures 3.18(c) to 3.18(f): good agreement with data for medium and high $p_{\mathrm{T}}^{\mathrm{avg}}$,
   - figures 3.18(a) and 3.18(b): smaller absolute asymmetries for low $p_{\mathrm{T}}^{\mathrm{avg}}$,
4. PowhegPythia: figures 3.18(a) to 3.18(f): smaller absolute asymmetries compared to data,
5. PowhegHerwig: figures 3.19(a) to 3.19(f): smaller absolute asymmetries compared to data,

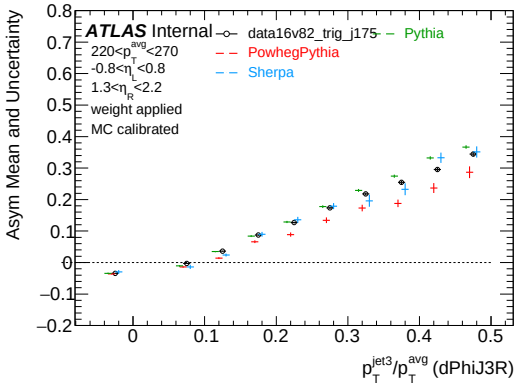
Comparing the different generators to data, Pythia and Herwig generators show moderate performance with a trend to give a harder third jet when it originates from a jet in the forward region in the case of Pythia and from a jet in the central region in the case of Herwig.
Sherpa generator shows a good performance except at low $p_{\mathrm{T}}^{\mathrm{avg}}$ with softer third jets.
Powheg generators have the worse mis-modeling of the third jet production, giving softer third jets. Nonetheless, they have the same trend for R and L tags which reduces the final effect on the eta-intercalibration result since we use the two groups together to derive the calibration.



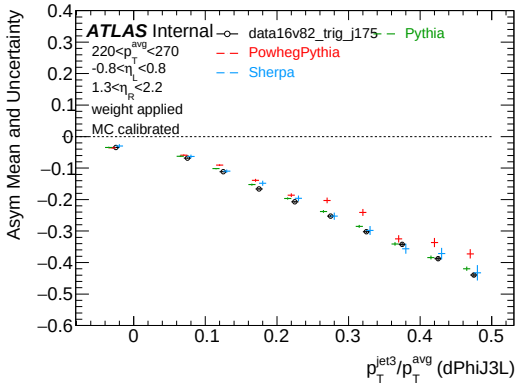
(a) $85 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 115$, R tag



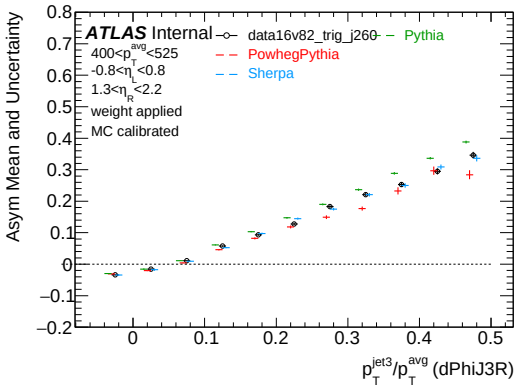
(b) $85 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 115$, L tag



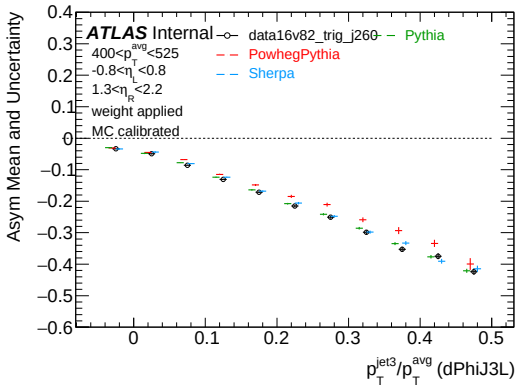
(c) $220 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 270$, R tag



(d) $220 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 270$, L tag

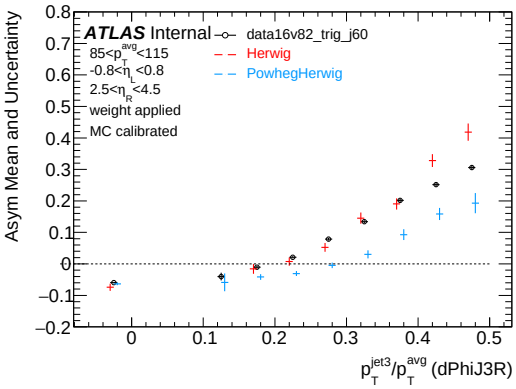


(e) $400 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 525$, R tag

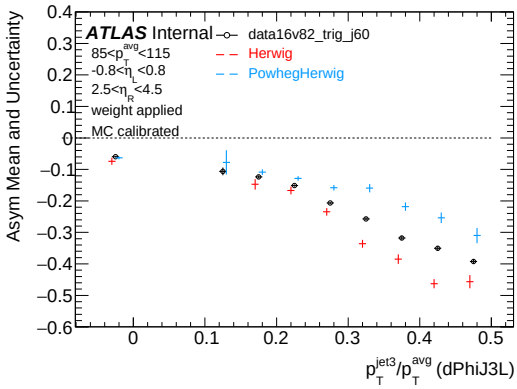


(f) $400 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 525$, L tag

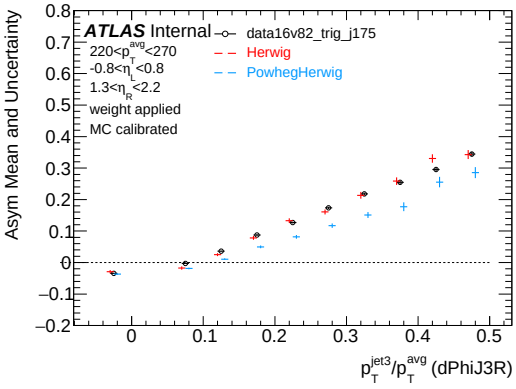
Figure 3.18: Asymmetry mean distribution as a function of $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}}$. MC simulation is calibrated to data using events with no third jet.



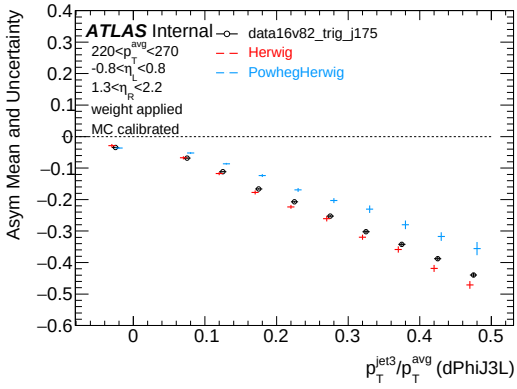
(a) $85 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 115$, R tag



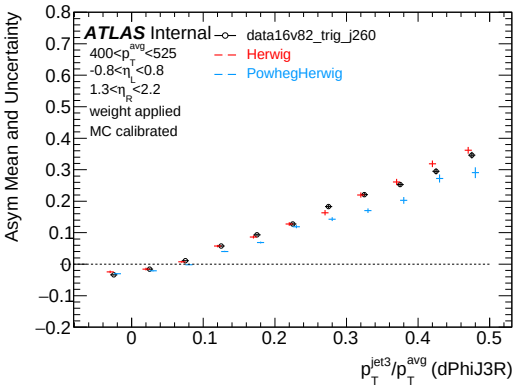
(b) $85 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 115$, L tag



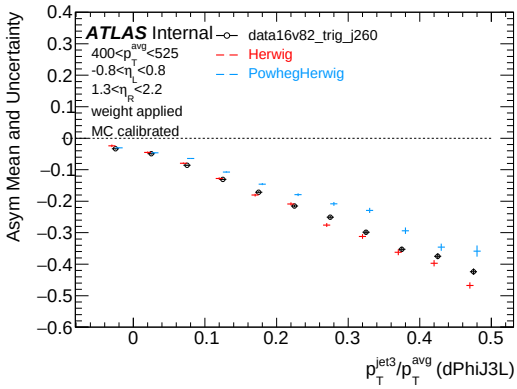
(c) $220 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 270$, R tag



(d) $220 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 270$, L tag



(e) $400 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 525$, R tag



(f) $400 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 525$, L tag

Figure 3.19: Asymmetry mean distribution as a function of $p_{\mathrm{T}}^{\mathrm{j\,3}}/p_{\mathrm{T}}^{\mathrm{avg}}$. Same plots as in figure 3.18 but using Herwig and PowhegHerwig MC generators.

3.3.5 Conclusion

Relying on the previous results from the various sections, I arrived to the following conclusions. First, it is confirmed that the study will use the following MC generators for eta-intercalibration:

- PowhegPythia (nominal),
- Sherpa.

Second, a tighter cut on the third jet p_{T} will be used to constrain the eta-intercalibration to the region where the agreement between data and MC simulations is good:

- old cut [67]: $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}} < 0.4$,
- new cut: $p_{\mathrm{T}}^{\mathrm{j}\,3}/p_{\mathrm{T}}^{\mathrm{avg}} < 0.25$.

On the other hand, a new version of PowhegPythia generator is addressing the mis-modeling of the third jet. A preliminary result is shown in figure 3.20. The plot of the new version, shown in cyan color, is closer to the results from the other generators like Pythia, in contrast with the plot of the current version, shown in orange color, which is far away from all the others and giving softer third jets.

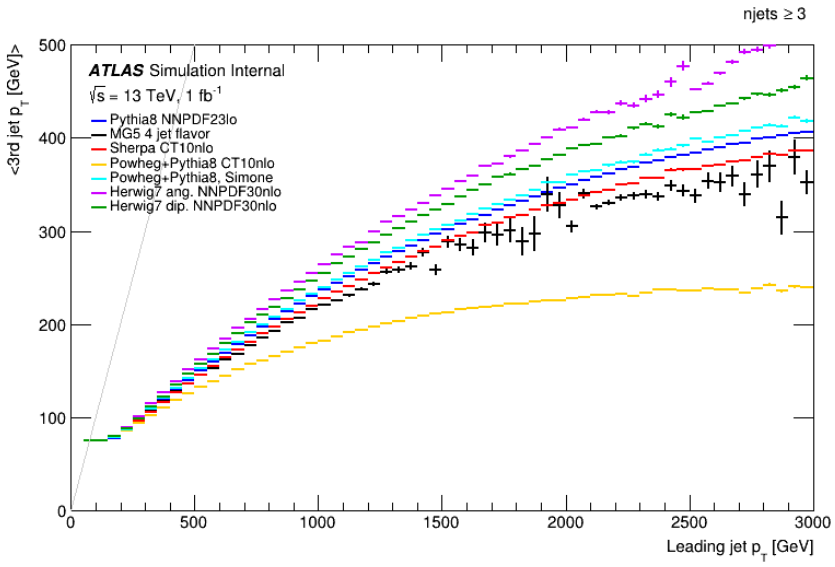


Figure 3.20: Average $p_{\mathrm{T}}^{\mathrm{j}\,3}$ distribution as a function of $p_{\mathrm{T}}^{\mathrm{j}\,1}$ for different MC generators. The current version of PowhegPythia generator is shown in orange color. The new version of PowhegPythia generator is shown in cyan color.

3.4 Dependence of the calibration results on the pile-up profile

The event sample used to derive the eta-intercalibration may have a different pile-up profile (refer to chapter 2.2.4) than the samples used for the different physics analyses. In fact, one of the techniques used to increase the sample size available for the eta-intercalibration is to increase its dedicated trigger rate at the end-of-fill where the pile-up is lower. The same is not done for the other triggers which are used for physics analyses. Before the eta-intercalibration step, jet energies have already been corrected for pile-up contributions as described in chapter 2.3.3.1. In principle, any remaining residual dependence should be covered by the systematic uncertainties on this correction. Nevertheless, it is important to check whether the variation of the calibration results for different pile-up profiles is well covered by these uncertainties.

Pile-up profiles (μ) for different analysis are mostly overlapping. However, I perform an extreme test where I split each of our samples for data and MC simulation into two sets of non-overlapping μ profiles:

- cut 1: $0 < \mu < 20$,
- cut 2: $20 < \mu < 50$.

The average μ in each set is 16.8 and 25.4 respectively. Then for each set, I derive the corresponding eta-intercalibration using the same μ selection for both data and MC simulation. The resulting calibrations are referred to as $\mathcal{R}^{\text{cut 1}}$ and $\mathcal{R}^{\text{cut 2}}$. The difference between the two calibrations is given by:

$$\Delta\mathcal{R} = \mathcal{R}^{\text{cut 1}} - \mathcal{R}^{\text{cut 2}}.$$

The next step is to measure the uncertainties on this difference. I begin, first, with the systematic ones. All the calibration factors are the same for the two sets except the μ- and NPV-dependent components of the pile-up calibration as defined in equation 2.3 (where NPV is the number of reconstructed primary vertices). In fact, changing the μ profile also changes the NPV one. At the same time, the μ- and NPV-dependent components of the pile-up calibration are fairly independent. This is understood by the fact that μ (which is an average over one lumi-block of 1 minute usually) characterizes the out-of-time pile-up activity whereas NPV characterizes the in-time pile-up activity which are fairly independent. Accordingly, the following procedure is applied:

- measure for each set the means $\overline{\mu}$ and $\overline{\text{NPV}}$,
- measure the corresponding uncertainties σ_μ and σ_{NPV} in each (p_{T}, η) bin,
- since the uncertainties for each component are fully correlated between the 2 samples, the uncertainty propagation yields:

$$\sigma_\mu^{\Delta\mathcal{R}} = \sigma_\mu^{\mathcal{R}^{\text{cut 1}}} - \sigma_\mu^{\mathcal{R}^{\text{cut 2}}},$$

$$\sigma_{\text{NPV}}^{\Delta\mathcal{R}} = \sigma_{\text{NPV}}^{\mathcal{R}^{\text{cut 1}}} - \sigma_{\text{NPV}}^{\mathcal{R}^{\text{cut 2}}}.$$

- since the uncertainties from different components are not correlated, the full systematic uncertainty on $\Delta\mathcal{R}$ is then:

$$\sigma^{\Delta\mathcal{R}}_{\text{syst.}} = \sqrt{(\sigma^{\Delta\mathcal{R}}_{\mu})^2 + (\sigma^{\Delta\mathcal{R}}_{\text{NPV}})^2}.$$

For the statistical uncertainties, since the 2 sets have no common events, I simply have:

$$\sigma^{\Delta\mathcal{R}}_{\text{stat.}} = \sqrt{(\sigma^{\mathcal{R}^{\text{cut 1}}}_{\text{stat.}})^2 + (\sigma^{\mathcal{R}^{\text{cut 2}}}_{\text{stat.}})^2}.$$

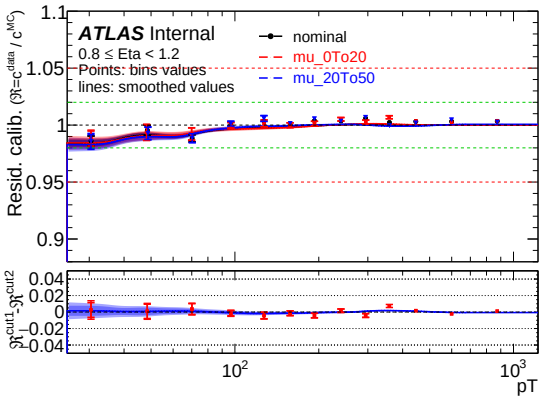
The total uncertainties on the calibration difference is then:

$$\sigma^{\Delta\mathcal{R}}_{\text{total}} = \sqrt{(\sigma^{\Delta\mathcal{R}}_{\text{stat.}})^2 + (\sigma^{\Delta\mathcal{R}}_{\text{syst.}})^2}.$$

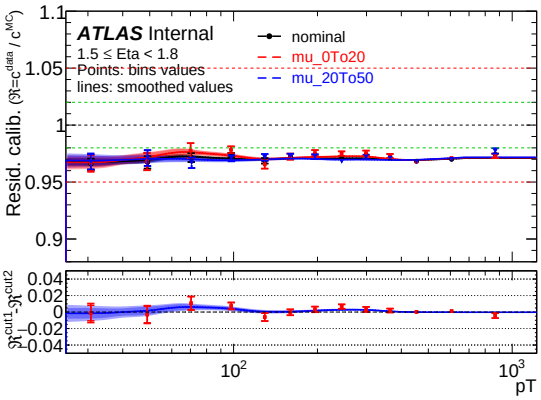
In the case of a non-dependence of the calibration results on the pile-up profile, $\Delta\mathcal{R}$ should be compatible with 0 (the full uncertainties should cover the 0 value).

Figure 3.21 shows the calibration results for the two sets of μ ranges for different η regions. The plots in the bottom pads show the calibrations difference $\Delta\mathcal{R}$. The inner error bars correspond to the statistical uncertainties, while the outer bars correspond to the full uncertainties. As it can be noticed, $\Delta\mathcal{R}$ is compatible with zero in most of the phase-space. For intermediate p_{T} and $2.1 < \eta < 2.5$, $\Delta\mathcal{R}$ uncertainties barely fails to cover the zero value for some points, but also we notice that $\Delta\mathcal{R}$ fluctuates around zero which suggests that the tension we are seeing is just a statistical fluctuation. On the other hand, for high p_{T} and $3.6 < \eta < 4.5$, the constant and systematic deviation of $\Delta\mathcal{R}$ from zero is just an artifact of the extrapolation of the calibration to this region where, due to kinematic limits, no jets can be found.

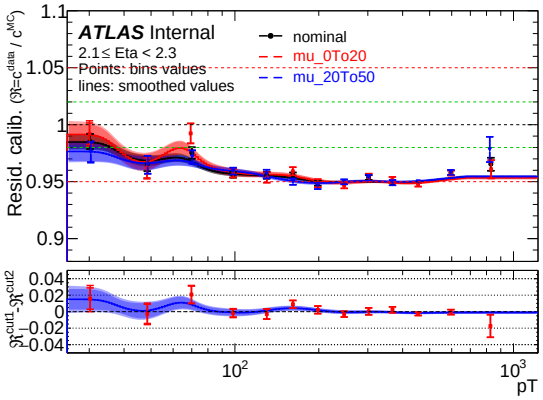
In conclusion, no strong dependence of the calibration results on the pile-up profile is found for this extreme test. Consequently, using the calibration derived with one pile-up profile and applying it to an analysis using a slightly different profile are compatible. The residual dependence is covered by the uncertainties.



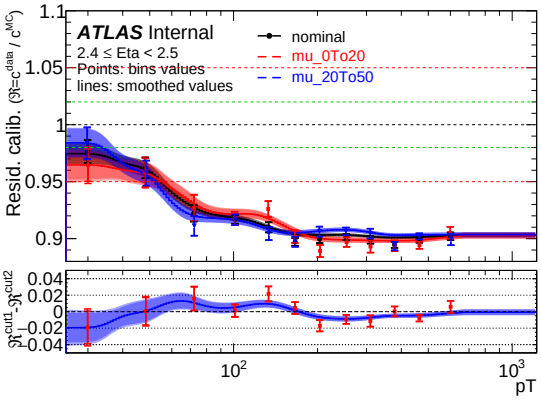
(a) $0.8 < \eta < 1.2$



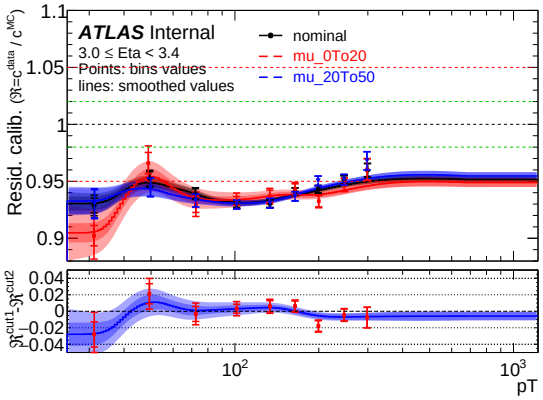
(b) $1.5 < \eta < 1.8$



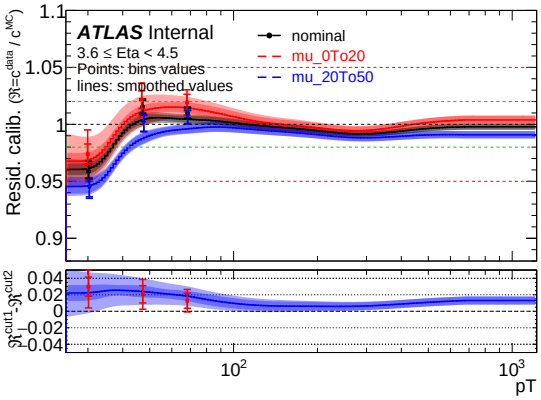
(c) $2.1 < \eta < 2.3$



(d) $2.4 < \eta < 2.5$



(e) $3.0 < \eta < 3.4$



(f) $3.6 < \eta < 4.5$

Figure 3.21: Residual calibration factors as a function of jets p_{T} for different η regions. The black, red and blue dots correspond respectively to the nominal events selection, additional cuts of $0 < \mu < 20$ and $20 < \mu < 50$. The continuous lines show the smoothed calibration functions. The plots in the bottom pads show the calibrations difference between the two regions with different μ cuts (red - blue) and which, in case of non-dependence of the calibration results on the pile-up profile, should be compatible with 0. The inner error bars correspond to the statistical uncertainties, while the outer bars correspond to the full uncertainties.

3.5 Forward and central triggers efficiencies and combination strategy

When selecting events using triggers, one would want to check that the triggers are used in the fiducial region where they are fully efficient. Moreover, the single jet triggers in ATLAS are divided between a central region ($|\eta| < 3.2$) and a forward one ($3.2 < |\eta| < 4.9$). When combining central and forward triggers for a given p_{T} bin, only the combination is required to be fully efficient. The fulfillment of the latter condition does not imply that each individual trigger is fully efficient.

This is illustrated in figure 3.22, where the asymmetry distributions of the same fiducial region but using only central or forward triggers are plotted. It is clear that the fits from the two distributions are not compatible with each other. This is not a problem of fitting since the ratio of χ^2/Ndf is close to one for both fits. In fact, this can be understood by the inefficiencies of the individual triggers. Here I am using the standard method defined in section 3.1. Recalling the definition of the asymmetry $\mathcal{A}^{\text{standard m.}} = (p_{\mathrm{T}}^{\text{probe}} - p_{\mathrm{T}}^{\text{ref}})/p_{\mathrm{T}}^{\text{avg}}$, in our case the probe jet belongs to the forward trigger region. The reference jet always belongs to the central trigger region in the standard method. I get the following:

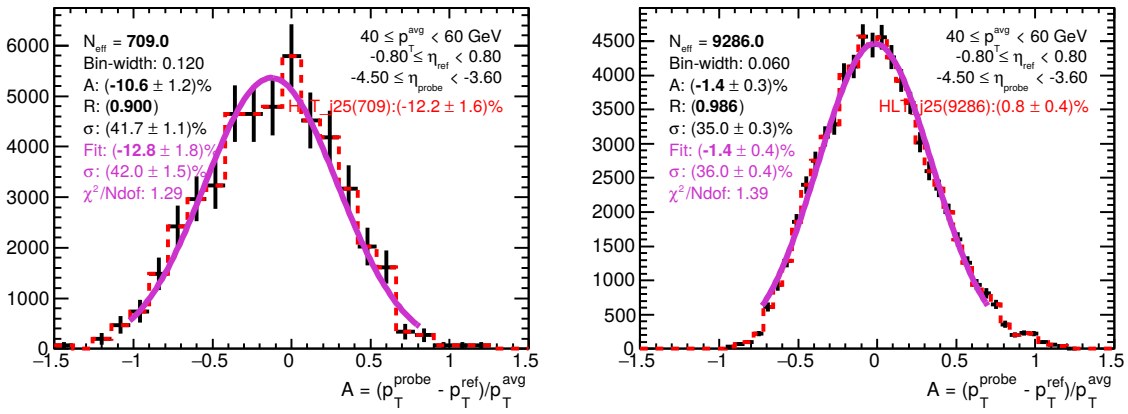
- central trigger: the inefficiency leads to a bias toward higher reference jet p_{T} resulting in a bias toward lower $\mathcal{A}$ values,
- forward trigger: the inefficiency leads to a bias toward higher probe jet p_{T} resulting in a bias toward higher $\mathcal{A}$ values.

This is exactly what is seen in the plots. In addition, I can note an asymmetry in the distributions, specially in the one using the forward trigger which, as we will see later, is less efficient than the central. This is a direct effect of the inefficiency, which changes the shape of the distributions, since it increasingly affects one side of $\mathcal{A}$. In our case for the forward trigger distribution, the lower the values of $\mathcal{A}$ are, the more they are affected by the inefficiency and the lower the distribution will be for the corresponding bins. This means the bins to the left side are suppressed in comparison to the right side, which is what we see. The same differences between the relative response results of central and forward triggers can also be seen in figure 3.23 with the same conclusion on the inefficiencies of individual triggers.

One way to calculate the efficiency of a trigger, called hereafter probe trigger, is through an emulation which tells, for a reconstructed event, if it would have passed the emulated trigger [74]. In this method, a reference trigger is used with the condition that it should be fully efficient in the region where we want to study the efficiency of the probe trigger. Using the events that actually fired the reference trigger (also called the actual trigger information), I emulate if each event would have fired the probe trigger or not: I test if the event passes the kinematic threshold of the probe trigger (also called the raw trigger information). The efficiency is then just the ratio between the number of events that passed the emulation and the total number of events:

$$\text{Eff} = \frac{\mathrm{N}^{\text{events}}(\text{fire reference trigger \& pass probe emulation})}{\mathrm{N}^{\text{events}}(\text{fire reference trigger})}. \tag{3.17}$$

Efficiency studies are done in kinematic variables that the analysis depends on. For eta-intercalibration, the main variable is $p_{\mathrm{T}}^{\text{avg}}$. Additional binning in η can also be used. A trigger is considered as fully efficient in the region where its efficiency is higher than 99%. The p_{T} of the transition between the



(a) Central trigger

(b) Forward trigger

Figure 3.22: Standard method asymmetry distributions and fits of the same fiducial region using only central or forward triggers.

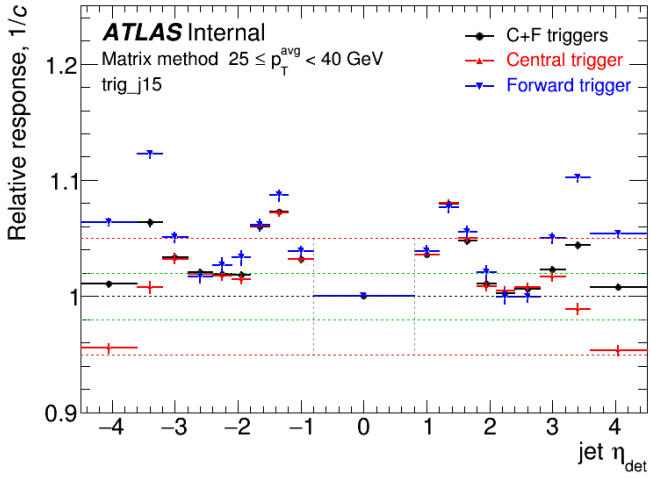


Figure 3.23: Relative jet responses from matrix method using only central (red), forward (blue) or combined (black) triggers.

two regimes is called the turn-on p_{T}. Usually, we use the trigger 5-10 GeV higher than it's turn-on p_{T} to take into account the uncertainties on our efficiency calculations.

Figure 3.24 shows the efficiencies for trigger HLT_j25 (a jet above 25 GeV at the high level trigger reconstruction) using HLT_j15 (a jet above 15 GeV at the high level trigger reconstruction) as a reference as a function of $p_{\mathrm{T}}^{\mathrm{avg}}$ and also $p_{\mathrm{T}}^{\text{leaging jet}}$ for comparison. The efficiencies are calculated for central, forward and combined triggers (when the efficiency is calculated for an individual trigger, events for which the 2 leading jets are not in the studied region are excluded). The HLT_j25 trigger is used in our analysis for $p_{\mathrm{T}}^{\mathrm{avg}} \geq 40$ GeV. It is clear that the individual triggers have worse efficiencies

compared to their combination but also they are not fully efficient in the $p_{\mathrm{T}}^{\mathrm{avg}}$ interval used in the analysis, in contrast with the combination which is fully efficient. It is also worth noting that the forward trigger performs worse than central one.

The emulation method, which uses the raw trigger information, does not include the effect of the triggers prescales. That is why a proper combination that correctly takes into account prescales is needed, or otherwise this leads to a new source of inefficiencies for the trigger combinations. I tested two methods: the simple and the inclusion combination methods.

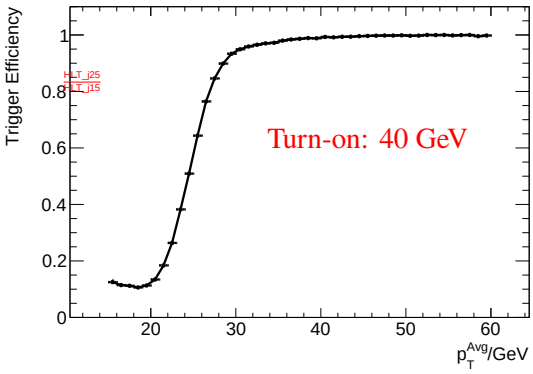
The simple method works as follows. Using only the information of the actual (fired) triggers, events are split into two groups:

1. F events:
    - where the actual forward trigger is passed, including when the actual central trigger is also passed,
    - and with the weight of the event which is given by the prescale of the forward trigger.
2. C events:
    - where only the actual central trigger is passed,
    - here, the weight of the event is given by the prescale of the central trigger.

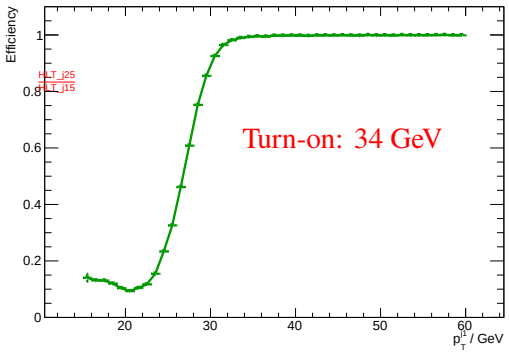
We can already see the problem with this method. For low p_{T}, the prescales for both central and forward triggers are very high. This means that the two groups of events, one passing the central prescale selection and the other passing the forward one, almost do not intersect. Since in the current method only the information of actual trigger is used (which requires passing the prescale), the two groups are filled almost independently. This leads an inefficiency of one trigger that cannot be compensated by the other trigger.

In contrast, the inclusion method [75] tries to overcome this problem by also taking into account the information of raw triggers. Events that fire the actual central or forward trigger are split into three groups:

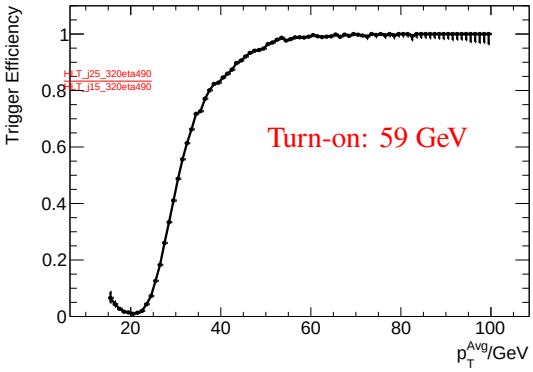
1. F events:
    - where only the raw forward trigger is passed,
    - and where the weight is given by the prescale of the forward trigger ($\mathrm{d_F}$).
2. C events:
    - where only the raw central trigger is passed,
    - and where the weight is given by the prescale of the central trigger ($\mathrm{d_C}$).
3. C&F events:
    - where both the raw central and forward triggers are passed,



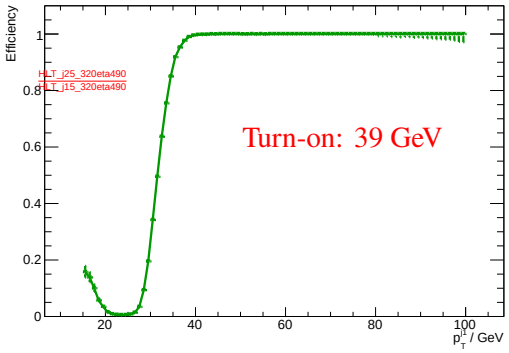
(a) Central trigger



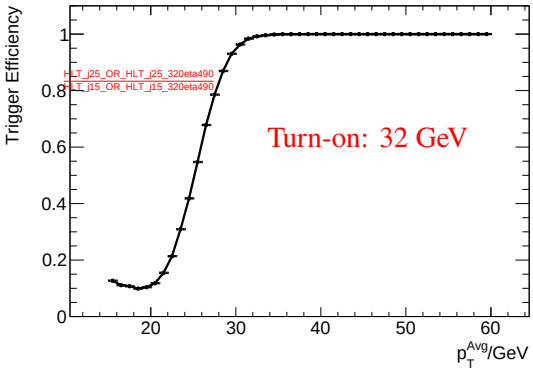
(b) Central trigger



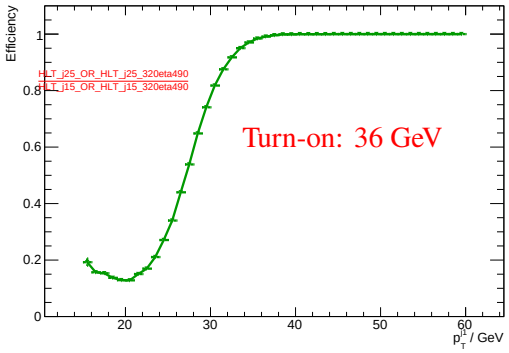
(c) Forward trigger



(d) Forward trigger



(e) Central and forward triggers combination



(f) Central and forward triggers combination

Figure 3.24: Efficiencies for trigger HLT_j25 using HLT_j15 as a reference as a function of $p_{\mathrm{T}}^{\mathrm{avg}}$ in left plots and $p_{\mathrm{T}}^{\text{leading jet}}$ in right plots. From top to bottom, the plots are for central, forward and combined triggers. The HLT_j25 trigger is used for $p_{\mathrm{T}}^{\mathrm{avg}} \geq 40$ GeV and the turn-on p_{T} is required to be at least 5 GeV lower.

- and where the weight is 1/P, where P is the probability that an event passes at least one trigger prescaling selection:

$$P = \frac{1}{d_C} + \frac{1}{d_F} - \frac{1}{d_C * d_F}.$$

I simulated events and triggers decisions to test those two methods. Different distributions are filled in the following procedure:

1. The reference distribution:
   - uses all events without prescaling,
   - if the event fires the trigger (here actual and raw decision are the same since there is no prescaling), a weight of 1 is used when filling the distribution.
2. The tested distributions:
   - apply first a prescaling,
   - for each of the combination methods described before, each passing event is then used to fill the distribution with the proper weight.

The reference distribution is an unbiased distribution given that the emulation verifies the full efficiency of the triggers combination.
The tested distributions are compared to the reference one by two means:

- median of the distribution,
- mean of a Gaussian fit.

For the Gaussian fits, different ranges are used:

- fixed range:
  $\pm 1\sigma$ and $\pm 2\sigma$ from mean: this helps to see how the bias changes with the interval, also the range $\pm 2\sigma$ is the one used in the eta-intercalibration fitting procedures,
- variable range:
  the largest symmetric range around the mean with a criteria to have the p-value of the fit ≥ 0.05: this range is used to check that there is no bias due to bad fits.

Taking as an example one fiducial region where one of the two leading jets is central and the other is forward, the resulting histograms are shown in figure 3.25 with the colors black, red and blue corresponding to the reference, simple method and inclusion method distributions. Several observations can be drawn comparing the individual plots before comparing the medians and means:

- the total number of entries of the inclusion method is very close to the reference one, whereas for the simple method it is much different (noting that for this analysis, the normalization has no effect on the final result),
- the effective number of entries of the inclusion method is significantly higher than that of the simple method (≈ 1.75 times): this means that the inclusion method has more statistical power and yields smaller uncertainties,

- figure 3.25(d) shows a good overlap between the inclusion method and the reference distributions, whereas the simple method has a systematic shift to the left side of the distribution.

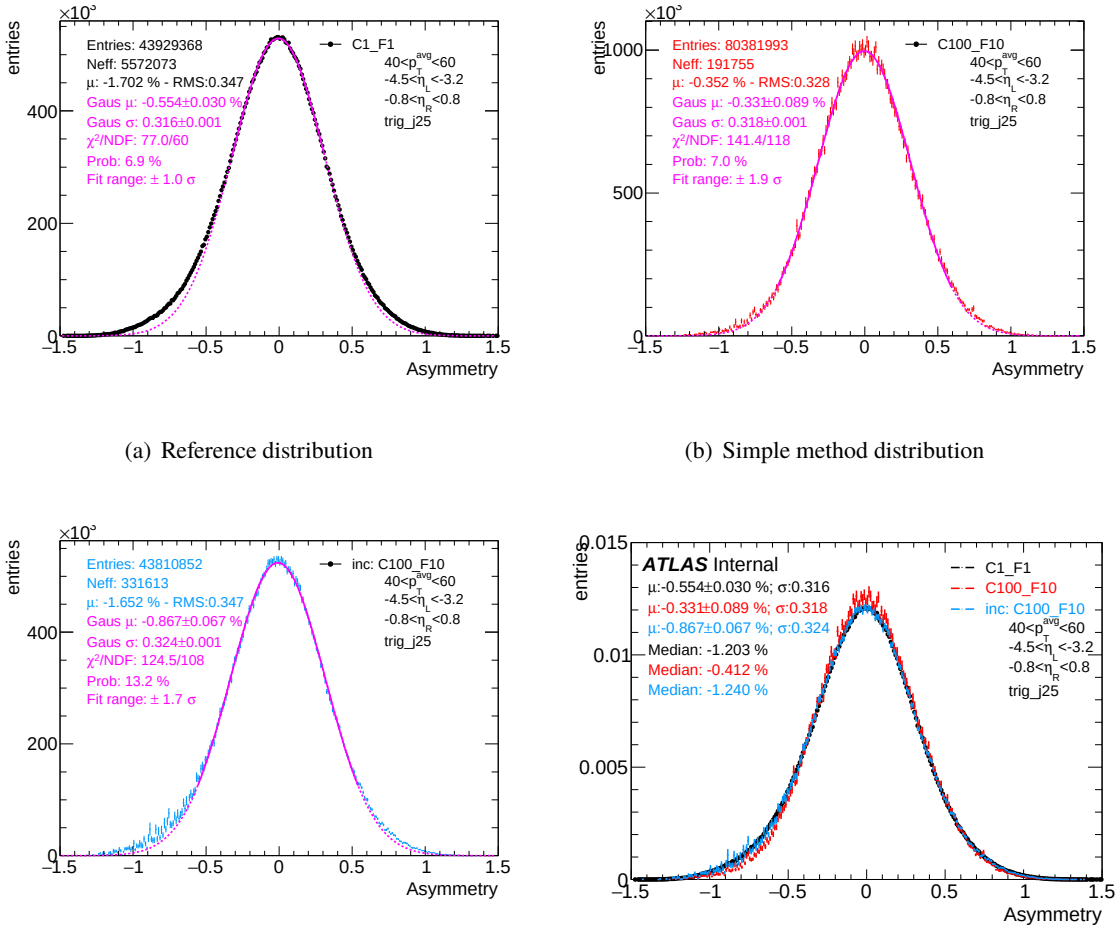
It is important also to note that the inclusion method has asymmetric uncertainties (between the left and right parts of the distribution), which is due to the way its three groups of events are split and weighted. The reason can be explained as follows:

- first, the central trigger in our simulation has a prescale 10 times the one of forward trigger (this is close to the real experimental settings),
- this means the weight for the "C" only events group is 10 times bigger and so do the uncertainties,
- the forward trigger is less efficient than the central and the inefficiency is for negative asymmetry values,
- the "C" only distribution will recover this inefficiency by including more events (which have less statistical power),
- this gives bigger uncertainties to the left side of the final distribution (all groups added).

What interests me is measuring the deviations of the medians and the means from fits of the tested distributions (of the different trigger combination methods) from the reference distribution ones and if they are significant. But since the same events are used, the values of medians and means are correlated between the reference and the tested distributions. To be able to properly measure the uncertainties on the differences, I use the bootstrap method (see 3.1.4).

Table 3.1 shows the deviations for three different kinematic space regions. The red values correspond to the deviations of the simple method, while the blue ones of the inclusion method. The results can be described as follow:

- median:
  - simple method: the deviations are statistically significant and are up to 1% which is the order of the effects we are trying to measure and to correct in eta-intercalibration; this shows that one side of the distribution is attenuated by the inefficiencies (at least more than the other side) which shifts significantly the median to the opposite side: positively when the forward region is left and negatively when the forward region is right as expected,
  - inclusion method: the deviations are small and within uncertainties; they are compatible with 0.
- fits:
  - simple method: the deviations increase from a per-mille level for $\pm 1\sigma$ range to a per-cent level for $\pm 2\sigma$ range; this shows that the more I fit the tails of the distributions where the inefficiencies are the highest, the more the fit mean deviates from the reference value,
  - inclusion method: a residual significant deviation exists for some fits specially for $\pm 2\sigma$ range, but the deviations are small ($< 0.3\%$): this is due to the fact that inclusion method distributions have non-symmetric uncertainties which constrains the fit more in one side and gives this residual shift in the fit mean.



(a) Reference distribution

(b) Simple method distribution

(c) Inclusion method distribution

(d) Superposition of normalized distributions

Figure 3.25: Matrix method asymmetry distributions where the colors black, red and blue correspond to the reference, simple method and inclusion method distributions. Normalization, shape but specifically median and fit mean are used to compare the two trigger combination methods to the reference and test for biases due to inefficiencies.

From the comparison of total numbers of events and medians, the inclusion method is verified to measure well the number of entries in different bins of the asymmetry and thus correcting well for the inefficiencies of individual triggers. The residual deviations of the means are due to non-symmetric uncertainties.

For the simple method, biases are found in the total number of events, median and fit mean. The method is not able to correct these inefficiencies.

Finally, I compare the results using these two methods for real data. Figure 3.26 shows the comparison of relative jet responses for two $p_{\mathrm{T}}^{\mathrm{avg}}$ intervals using the central reference method. The bottom pads show the absolute difference between the two combination methods. Significant differences are seen

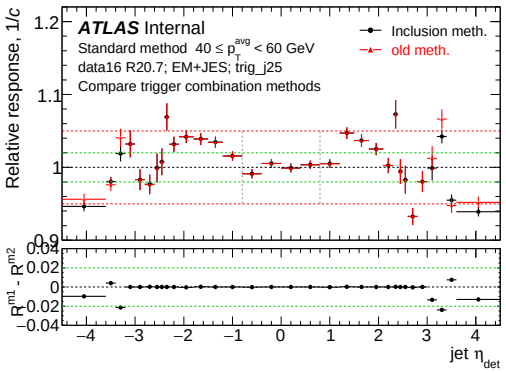
Table 3.1: Deviations of medians and means from fits between tested combination methods and reference distributions. The red values correspond to comparison of the simple method, while the blue ones to the inclusion method. Results from three different kinematic regions are shown for $40 < p_T < 60$ GeV and triggers HLT_j25. The uncertainties are calculated using bootstrap method to take into account correlations.

Differences to the reference distribution (in ‰)			
	$-0.8 < \eta^R < 0.8$ $-4.5 < \eta^L < -3.2$	$-3 < \eta^R < 3$ $-4.5 < \eta^L < -3.2$	$-0.8 < \eta^L < 0.8$ $3.2 < \eta^R < 4.5$
Δ Median	0.788 ± 0.092	0.795 ± 0.046	- 1.020 ± 0.088
	- 0.038 ± 0.069	- 0.046 ± 0.036	- 0.047 ± 0.076
Δ Fit mean	0.231 ± 0.088	0.166 ± 0.053	- 0.608 ± 0.089
(variable range)	- 0.284 ± 0.069	- 0.137 ± 0.038	0.034 ± 0.070
Δ Fit mean	- 0.175 ± 0.161	0.098 ± 0.086	- 0.112 ± 0.140
(±1 σ range)	- 0.202 ± 0.070	- 0.048 ± 0.053	- 0.019 ± 0.096
Δ Fit mean	0.738 ± 0.083	0.670 ± 0.046	- 0.942 ± 0.087
(±2 σ range)	0.163 ± 0.067	0.130 ± 0.033	- 0.241 ± 0.066

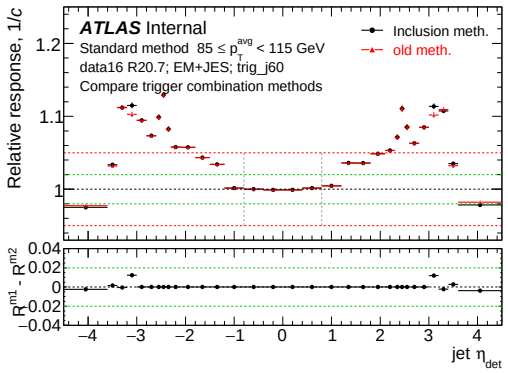
in the forward region ($|\eta| > 3.2$). The differences can go as high as 4% with an average between $1 - 2$‰.
For the bins with $|\eta| < 3$, the difference is 0 as expected since for the central reference method, both probe and reference jets are in the central region. For the bin with $3.0 \leq |\eta| < 3.2$, in principle there should be no difference since it is also central. But, a significant difference exists and this is caused by the fact that since this bin is close to the transition between central-forward regions, several entries for this bin come from events firing the forward trigger: a jet which is forward at trigger level with η close to 3.2 can be reconstructed as a central jet after off-line calibration. This is verified in the figure 3.27 where we see that for both combination methods, the bin $3.0 \leq |\eta_{\text{probe}}| < 3.2$ has entries from the forward trigger.

For the matrix method, the results are shown in figure 3.28. As can be seen, the bias spreads from forward to central bins, due to the correlation between all the relative jet responses (or the intercalibration factors c_i) inside equation 3.6. The differences between relative response are smaller than the central reference method, but spread to more bins.

As a result of the studies and tests shown before, I changed the triggers combination method used for eta-intercalibration from the old simple method to the new inclusion method.

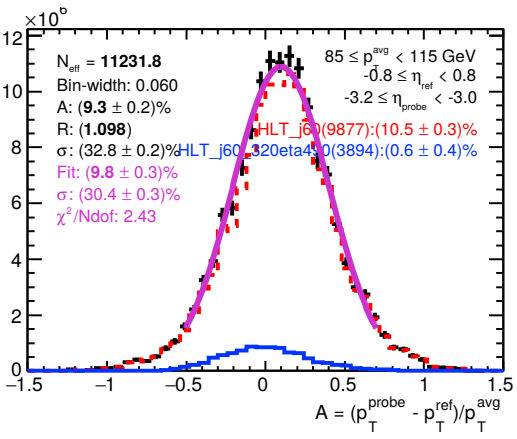


(a) $40 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 60$ GeV, central reference method

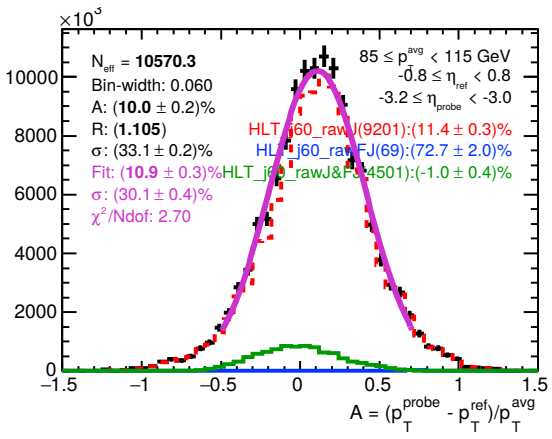


(b) $85 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 115$ GeV, central reference method

Figure 3.26: Using the central reference method, comparison of relative jet responses for two triggers combination methods, the inclusion in black and the simple in red. The bottom pads show the absolute difference between the two.

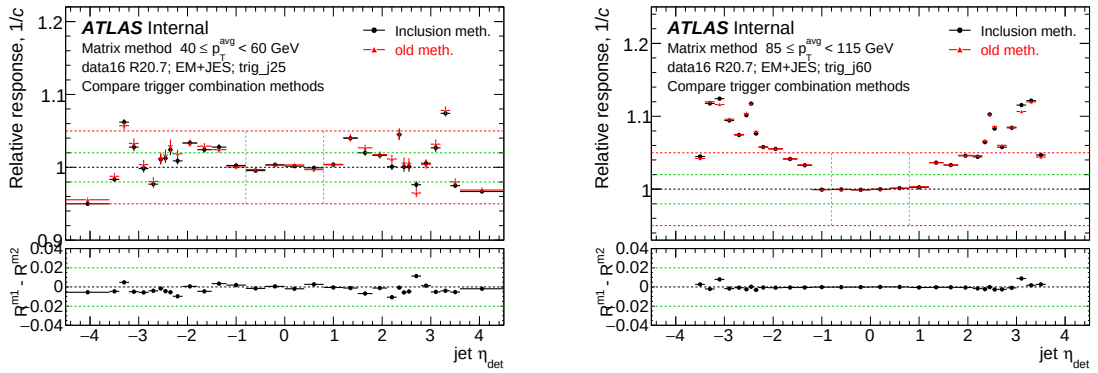


(a) Simple combination method



(b) Inclusion combination method

Figure 3.27: Asymmetry distributions where the two jets are in the central region, but where one of the jets is close to the central-forward transition at $|\eta| = 3.2$. The distributions have entries from events firing the forward trigger (blue histogram for simple method, blue and green histograms for inclusion method). This is caused by a jet which is forward at trigger level with η close to 3.2 that is reconstructed as a central jet after off-line calibration.



(a) $40 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 60$, matrix method

(b) $85 \leq p_{\mathrm{T}}^{\mathrm{avg}} < 115$, matrix method

Figure 3.28: Using the matrix method, comparison of relative jet responses for two triggers combination methods, the inclusion in black and the simple in red. The bottom pads show the absolute difference between the two.

3.6 Calibration results

3.6.1 Data selection

36.85 fb^{-1} of data collected in years 2015 and 2016 are used to measure the eta-intercalibration for the first data taking period. For the second period, 43.8 fb^{-1} of data collected in year 2017 are used. The calibration is measured separately for the two data taking periods since the detector conditions (aging, dead modules ...) change between years and as a cross-ckeck of the evolution of the detector response with time.
MC simulation for each period is generated with a pileup profile matched to the corresponding one of data.

The calibration is measured for the two jet reconstruction collections: EMTopo and EMPFlow (see 2.3.1).

Nominal selection cuts and their variations for systematic uncertainty evaluation are shown in table 3.2.

Trigger efficiencies
I studied the trigger efficiencies using the emulation method as defined in equation 3.17. The trigger is considered fully efficient if its efficiency is higher than 99%. This condition is verified in bins of η. Figure 3.29 shows, on the left plot, the efficiency curves inclusive in η for three triggers and, on the right plots, the efficiency for one trigger binned in η and $p_{\mathrm{T}}^{\mathrm{avg}}$. The efficiency threshold is defined as the $p_{\mathrm{T}}^{\mathrm{avg}}$ beyond which the trigger is fully efficient for each of the different bins of η. A summary of the efficiency thresholds is shown in table 3.3.

$p_{\mathrm{T}}^{\mathrm{avg}}$ binning and trigger selection
Concerning the p_{T} binning, the change I did compared to the past version of the eta-intercalibration, is to extend the calibration to higher values from 1200 to 2000 GeV. The binning at lower $p_{\mathrm{T}}^{\mathrm{avg}}$ is kept the same.
$p_{\mathrm{T}}^{\mathrm{avg}}$ bins = [25, 40, 60, 85, 115, 145, 175, 220, 270, 330, 400, 525, 760, 1100, 1500, 2000] GeV.
The trigger combination selection for each $p_{\mathrm{T}}^{\mathrm{avg}}$ bin taking into account efficiency thresholds are shown in table 3.4.

Table 3.2: Nominal selection cuts and their variations for systematic uncertainty evaluation. The indices 1, 2 and 3 refer respectively to the leading, sub-leading, and sub-sub-leading jet.

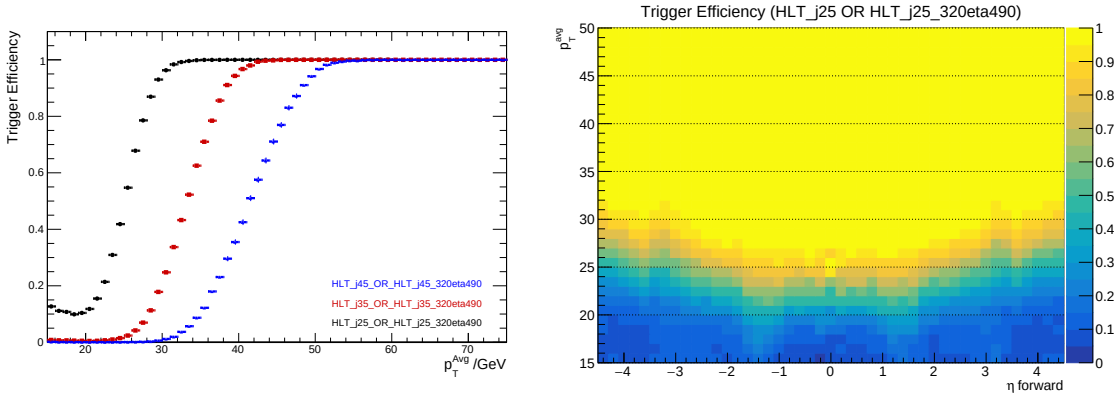
Cut	Loose	Nominal	Tight
$\Delta\phi^{12}$	2.3	2.5	2.8
$p_{\mathrm{T}}^{\mathrm{jet\ 3}}/p_{\mathrm{T}}^{\mathrm{avg}}$	0.35	0.25	0.15
JVT: EMTopo	0.11	0.59	0.91
JVT: PFlow	0	0.2	0.5

Table 3.3: 99% efficiency $p_{\mathrm{T}}^{\mathrm{avg}}$ thresholds for combination of central and forward triggers.

central and forward triggers combination	99% efficiency thresholds [GeV]
HLT_j15 OR HLT_j15_320eta490	25
HLT_j25 OR HLT_j25_320eta490	34
HLT_j35 OR HLT_j35_320eta490	45
HLT_j45 OR HLT_j45_320eta490	55
HLT_j60 OR HLT_j60_320eta490	80
HLT_j110 OR HLT_j110_320eta490	127
HLT_j175 OR HLT_j175_320eta490	197
HLT_j260 OR HLT_j26_320eta490	189
HLT_j360 OR HLT_j36_320eta490	396

Table 3.4: Trigger combination selection for each $p_{\mathrm{T}}^{\mathrm{avg}}$ bin.

$p_{\mathrm{T}}^{\mathrm{avg}}$ bins [GeV]	Triggers combination selected
25 - 40	HLT_j15 OR HLT_j15_320eta490
40 - 60	HLT_j25 OR HLT_j25_320eta490
60 - 85	HLT_j35 OR HLT_j35_320eta490
85 - 115	HLT_j60 OR HLT_j60_320eta490
115 - 145	
145 - 175	HLT_j110 OR HLT_j110_320eta490
175 - 220	
220 - 270	HLT_j175 OR HLT_j175_320eta490
270 - 330	
330 - 400	HLT_j260 OR HLT_j260_320eta490
400 - 525	HLT_j360 OR HLT_j360_320eta490
525 - 760	2015+16 data:
760 - 1100	HLT_j360 OR HLT_j360_320eta490
1100 - 1500	2017 data:
1500 - 2000	HLT_j400 OR HLT_j400_320eta490



(a) (b)

Figure 3.29: Trigger efficiencies using emulation method, inclusive in η in a) and binned in η in (b).

3.6.2 η binning optimization

With the new minimization methods (section 3.2), the number of η bins can be increased as needed. It was the aim of this development. The only limitation is the statistics that exist. In our analysis, the statistics limitation comes mostly from MC simulation at low p_{T}. The optimization procedure I used is the following:

- plot the jet response in data using very fine bins,
- optimization for peaks: combine bins only where there is no sharp variations in response variation (no peaks),
- optimization for statistics: combine additional bins where there is not enough statistics (specially due to MC simulation).

Using this procedure, I get the following optimized η binning with 46 bins (vs 31 bins for the binning of the old eta intercalibration results [67]):
$|\eta|$ bins = [0.0, 0.4, 0.8, 1.0, 1.2, 1.4, 1.5, 1.6, 1.8, 2.0, 2.2, 2.4, 2.45, 2.5, 2.55, 2.6, 2.8, 3.0, 3.2, 3.3, 3.4, 3.6, 3.9, 4.5].

For the first two $p_{\mathrm{T}}^{\mathrm{avg}}$ bins, due to low statistics, the following two changes on $|\eta|$ bins are applied:

- [2.4, 2.45, 2.5, 2.55, 2.6] is changed to [2.4, 2.5, 2.6],
- [3.2, 3.3, 3.4] is changed to [3.2, 3.4].

The improvement using this optimized binning can be seen in figure 3.30 where jet response measurement results using a very fine binning (in black), the new optimized binning (in red) and the old binning (in blue) are shown. Taking for example the region of $-3.4 < \eta < -3$ where the old binning uses only one bin whereas the new optimized binning uses 3, it is clear that the old binning does not describe well the existing peak. The difference is of the order of several percent which is significant for the level of precision we aim to have. The improvement is also visible in many other η regions,

including the region of $2.4 < |\eta| < 2.5$ where the problem in jet distributions was first seen (figure 3.8).

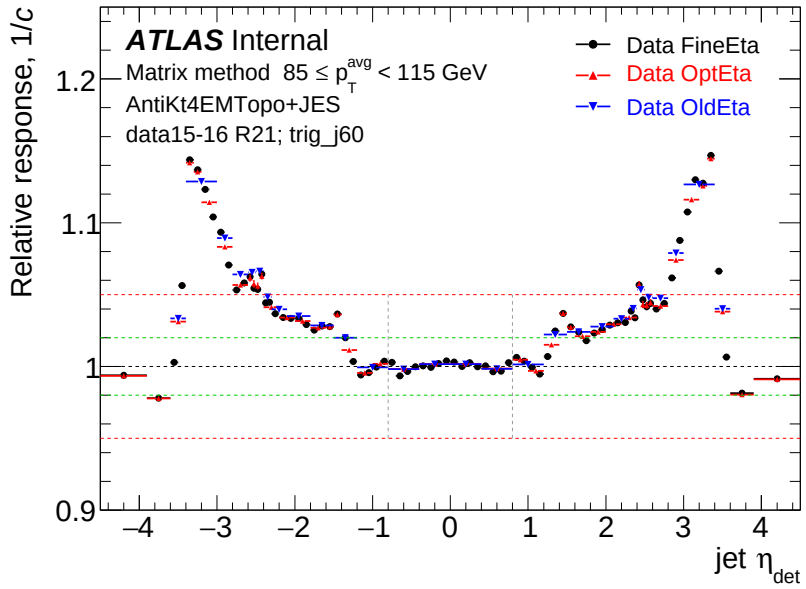


Figure 3.30: Comparison of jet response measurement results between a very fine binning (in black), the new optimized binning (in red) and the old binning (in blue).

Another change I did is to extend the binning to more forward regions (higher $|\eta|$ values) when there is enough statistics.

3.6.3 Calibration central values

Figure 3.31 shows different steps of the calibration calculation I performed. In the upper pad, relative jet responses as a function of η for one bin of $p_{\mathrm{T}}^{\mathrm{avg}}$ are shown for data as black points and for MC simulation as red points. Those values are measured using the matrix method (section 3.1.2). Spikes in data that are not well described by MC simulation are visible. In the bottom pad, the response ratios of MC simulation to data are shown as red points. Those points are the calibration points used as inputs to the smoothing kernel. The magenta line in the bottom pad represents the smoothed calibration in one slice of $p_{\mathrm{T}}^{\mathrm{jet}}$.

The standard minimization method, which suffers from low statistics specially at low p_{T}, is used as a cross check on the results of the matrix method. This is shown in figure 3.32 where the calibration points and smoothed calibration are shown as a function of η with their statistical uncertainties for the two methods. As can be seen for $40 < p_{\mathrm{T}} < 60$ GeV, the statistical uncertainties for the standard method are much larger than the ones for the matrix method but the two calibrations show the same shape within fluctuations. For $85 < p_{\mathrm{T}} < 115$ GeV, the uncertainties are smaller and the compatibility between the two methods is visible more clearly.
Also in this figure, a comparison between the calibrations of EMTopo and EMPFlow jets is shown.

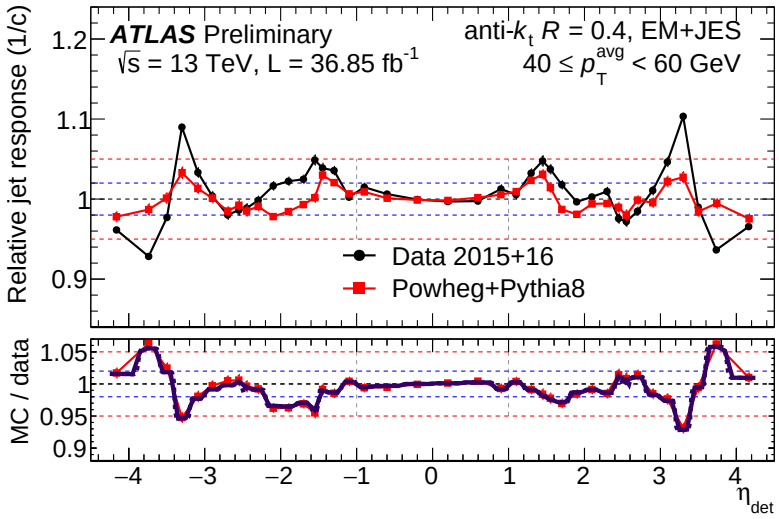


Figure 3.31: The upper pad shows the relative jet responses as a function of η for one bin of $p_{\mathrm{T}}^{\mathrm{avg}}$ for data as black points and for MC simulation as red points. The bottom pad shows the ration between the responses as red points and the smoothed calibration as magneta line.

The shapes are close but a significant difference also exists, for example for $3 < |\eta| < 3.6$. The difference is expected and the two jet reconstructions need separate calibrations.

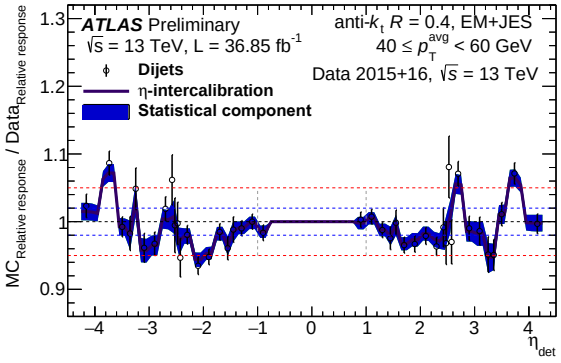
In the previous plots, only 1D sections of the calibration are shown and only for 2015+16 data. In figures 3.33 and 3.34, I show respectively 2D plots (p_{T}, η) of calibration central values and statistical uncertainties for 2015+16 and 2017 data, EMTopo and EMPFlow jets. A kinematic region corresponding to an energy of 4.5 TeV is shown in red lines. This is roughly the reach of the bins I use and where beyond that very few jets are found.
It is worth noting that at low p_{T}, 2017 data statistical uncertainties are significantly higher than the ones of 2015+16 data. I checked and found two main sources:

- the prescales of the low triggers, HLT_j15, HLT_j25 and HLT_j35, were increased by a factor of about two,
- due to higher pile-up conditions, the fraction of the rejection of events due to the selection criteria also increased by about a double at low p_{T}, with the $p_{\mathrm{T}}^{\mathrm{jet\ 3}}/p_{\mathrm{T}}^{\mathrm{avg}}$ cut being the cut that had the most impact on this increase.

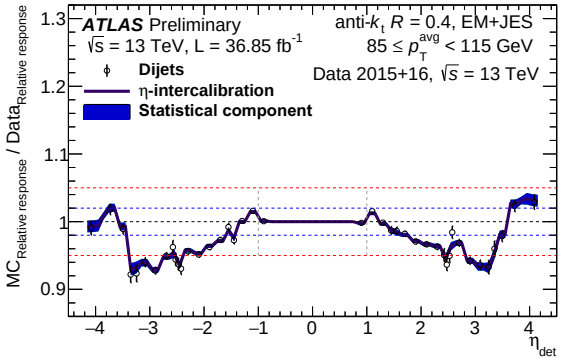
Figure 3.35 shows the calibration difference significance between 2015+16 and 2017 data. A significance of more than 4 and up to about 8 can be seen in some of the phase space. This proves the need of separate calibration for the two run periods.

For the smoothing kernel parameters, $\sigma_{\log p_{\mathrm{T}}}$ and $\sigma_{\eta_{\mathrm{det}}}$ (eq. 3.10), I tested combinations of the following values:

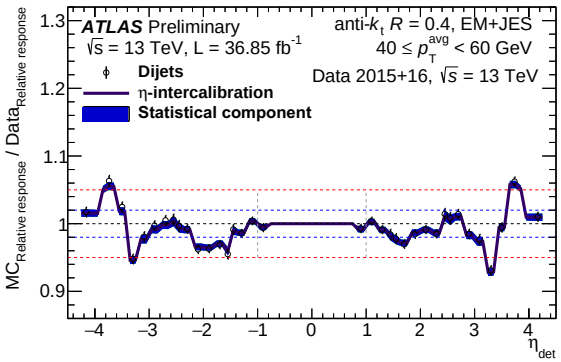
- $\sigma_{\log p_{\mathrm{T}}} = 0.12$, 0.15, 0.18 or 0.2,
- $\sigma_{\eta_{\mathrm{det}}} = (\eta$ bin width)/f with f = 1,2, 3, 4 or 6.



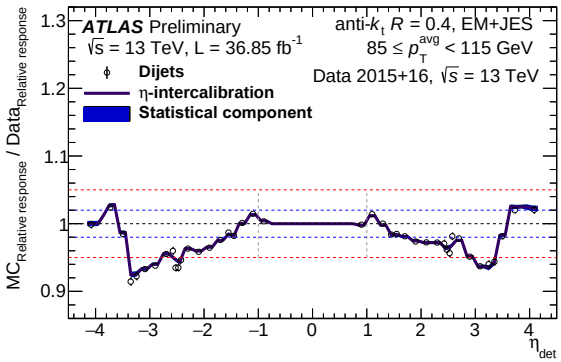
(a) EMTopo jets, Standard method, $40 < p_T < 60$ GeV



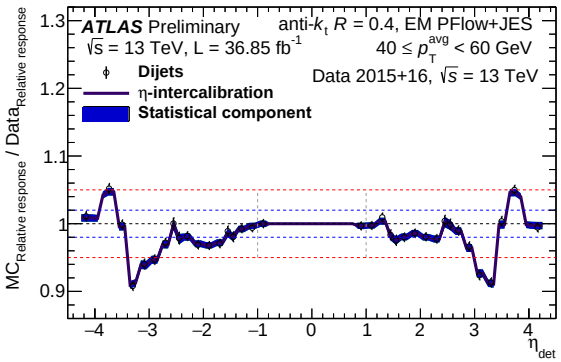
(b) EMTopo jets, Standard method, $85 < p_T < 115$ GeV



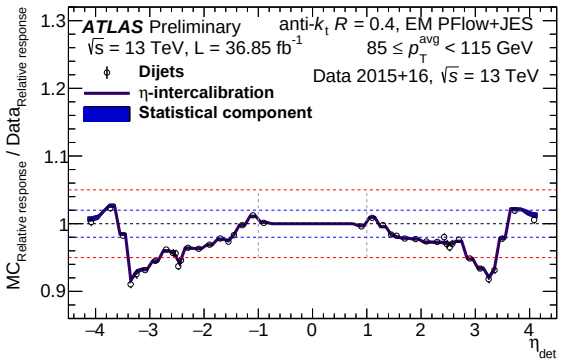
(c) EMTopo jets, Matrix method, $40 < p_T < 60$ GeV



(d) EMTopo jets, Matrix method, $85 < p_T < 115$ GeV



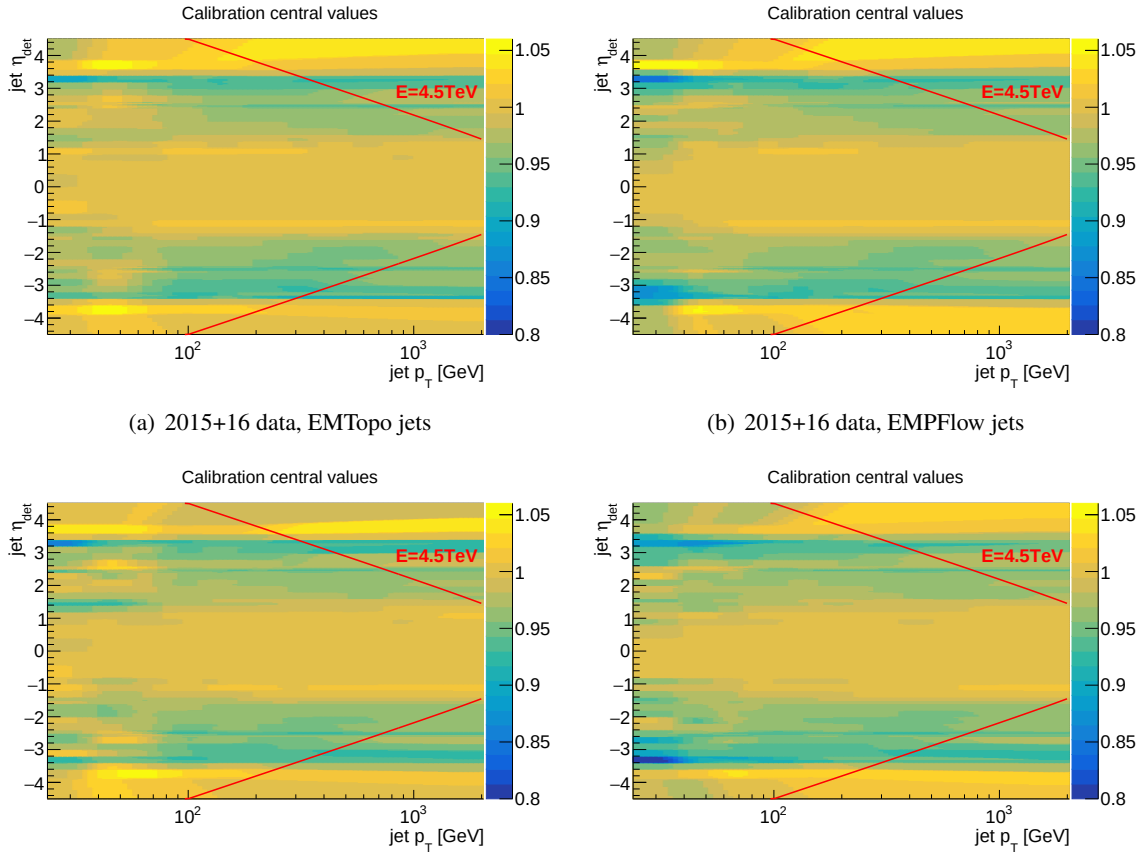
(e) EMPFlow jets, Matrix method, $40 < p_T < 60$ GeV



(f) EMPFlow jets, Matrix method, $85 < p_T < 115$ GeV

Figure 3.32: Comparison of smoothed calibration as a function of η for EMTopo using standard method, EMTopo and EMPFlow using matrix method. Statistical uncertainties are shown.

By comparing the closure of the calibrations, I found the best combination to be $\sigma_{\log p_T} = 0.18$ and $\sigma_{\eta_{det}} = (\eta \text{ bin width})/6$.



(a) 2015+16 data, EMTopo jets

(b) 2015+16 data, EMPFlow jets

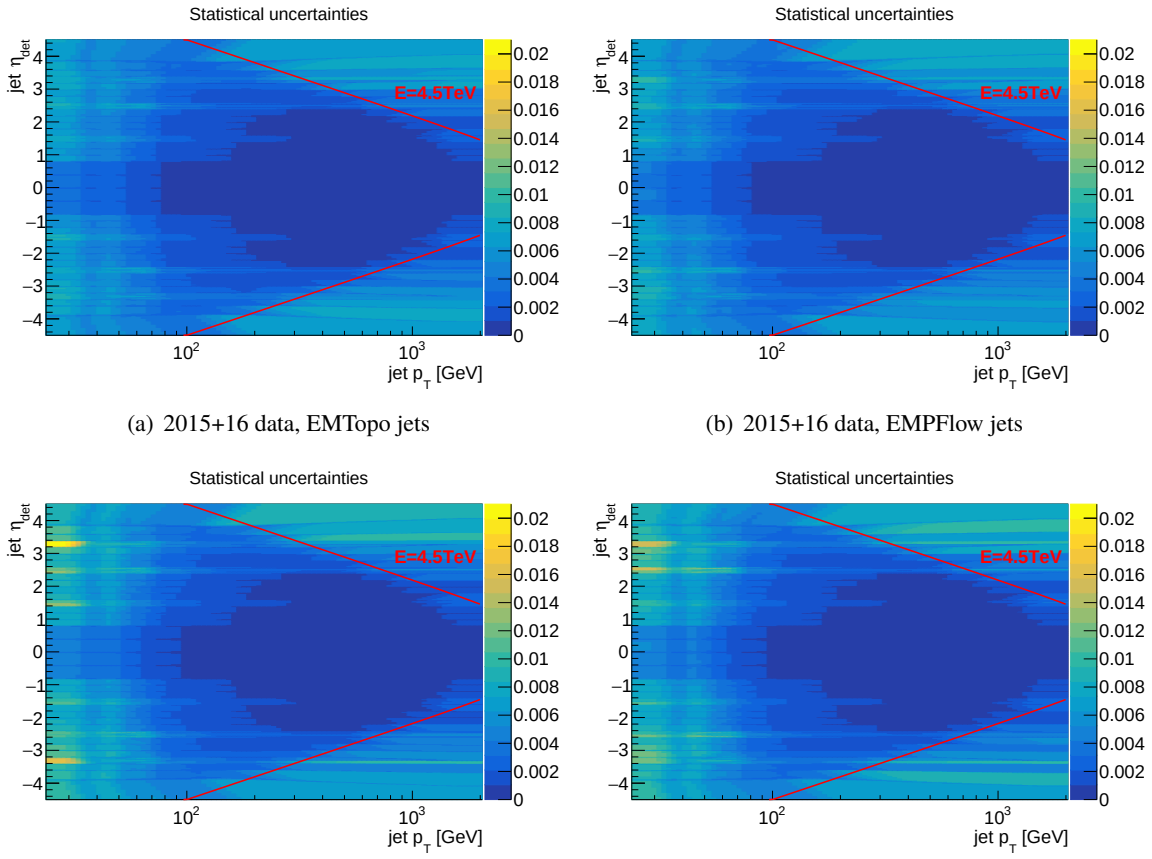
(c) 2017 data, EMTopo jets

(d) 2017 data, EMPFlow jets

Figure 3.33: Calibration central values 2D plots (p_T, η) for 2015+16 and 2017 data, EMTopo and EMPFlow jets.

Noting that the bigger the parameter is the more smoothing is done in the corresponding direction, we notice that the smoothing is higher in the p_T direction where we expect the calibration to change slowly and in the opposite the smoothing is smaller in the η direction where detector effects (cracks, changes in the detector technology ...) can induce fast changes in the response.

2D plots of the χ^2 values between each calibration point and the smoothed calibration value are used to verify that the smoothing is following well the input points. One of those check plots is shown in figure 3.36. The majority of bins have χ^2 values smaller than 1, some have values higher than 2 and few higher than 4. Also, the total sum of χ^2 values is smaller than the number of input points. The smoothing is following well the calibration points.



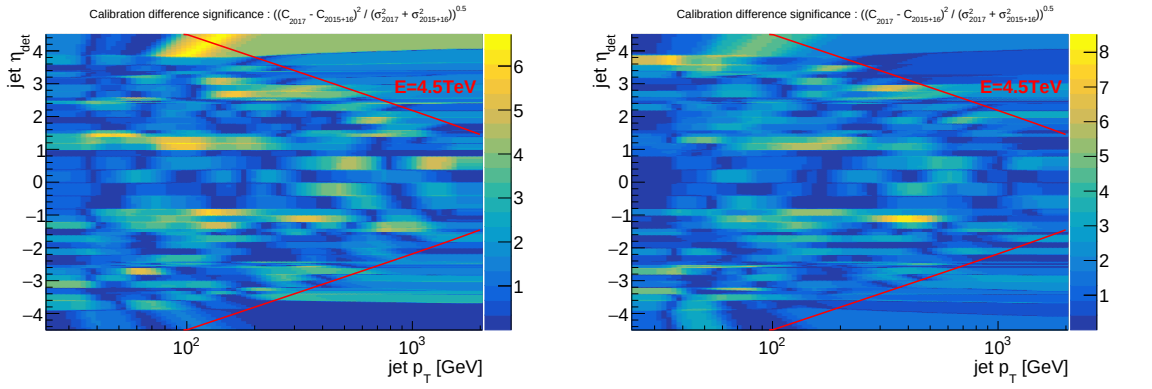
(a) 2015+16 data, EMTopo jets

(b) 2015+16 data, EMPFlow jets

(c) 2017 data, EMTopo jets

(d) 2017 data, EMPFlow jets

Figure 3.34: Calibration statistical uncertainties 2D plots (p_T, η) for 2015+16 and 2017 data, EMTopo and EMPFlow jets.



(a) EMTopo jets

(b) EMPFlow jets

Figure 3.35: Calibration difference significance 2D plots (p_T, η) between 2015+16 and 2017 data for EMTopo and EMPFlow jets.

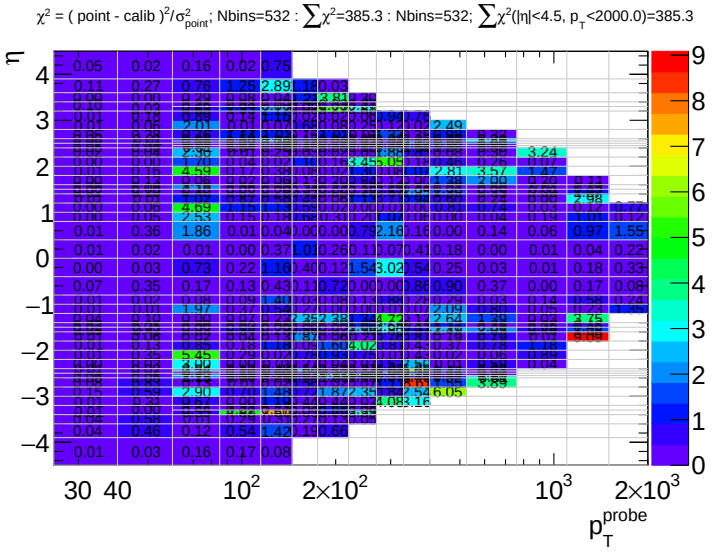
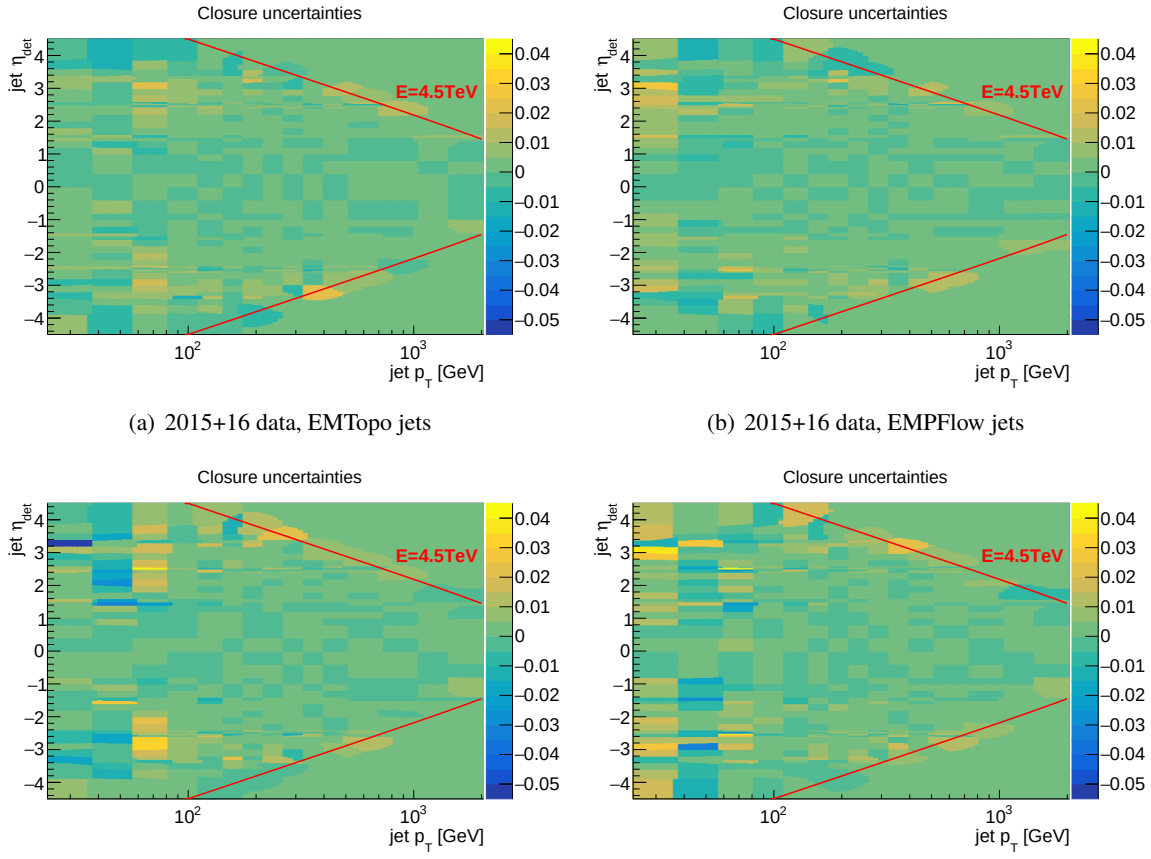


Figure 3.36: χ^2 values between calibration points and the smoothed calibration.

3.6.4 Closure test

The closure test, presented in section 3.1.5, is used to check if the calibration is able to detect and correct the differences in responses between data and MC simulation, and to eventually add a systematic uncertainty in the phase space region where the closure is not good enough.
Figure 3.37 shows the calibration closure uncertainties 2D plots (p_{T}, η) for 2015+16 and 2017 data, EMTopo and EMPFlow jets. For the majority of points, the non-closure is smaller than 1%. For 2015+16 data, the maximum closure uncertainty (in absolute values) is about 2.5%, whereas it is about 5% for 2017 data.
I checked that for some of the points at low p_{T}, the high closure uncertainties come from the fact that their statistical uncertainties is higher than their neighboring points in the p_{T} direction, which means that the smoothing will be more constrained by those neighboring points since more smoothing is done in the p_{T} direction (as explained in the previous section).

Although the closure uncertainty by itself is useful as a cross check, the significance of the closure uncertainty is also important to check since we do not want to double count the uncertainties from statistical fluctuations (the same events are used for the calibration and the closure test measurements). As expected, the significance of the closure uncertainty is small for the points with maximum closure uncertainty at low p_{T}. The regions where the significance was systematically higher than 2 are for $2.4 < |\eta| < 2.6$ or $\mathrm{E} > 4.5$ TeV (where E is the jet energy). Only for those regions, the closure uncertainties are propagated to other analysis.



(a) 2015+16 data, EMTopo jets

(b) 2015+16 data, EMPFlow jets

(c) 2017 data, EMTopo jets

(d) 2017 data, EMPFlow jets

Figure 3.37: Calibration closure uncertainties 2D plots (p_T, η) for 2015+16 and 2017 data, EMTopo and EMPFlow jets.

3.6.5 Systematic uncertainties

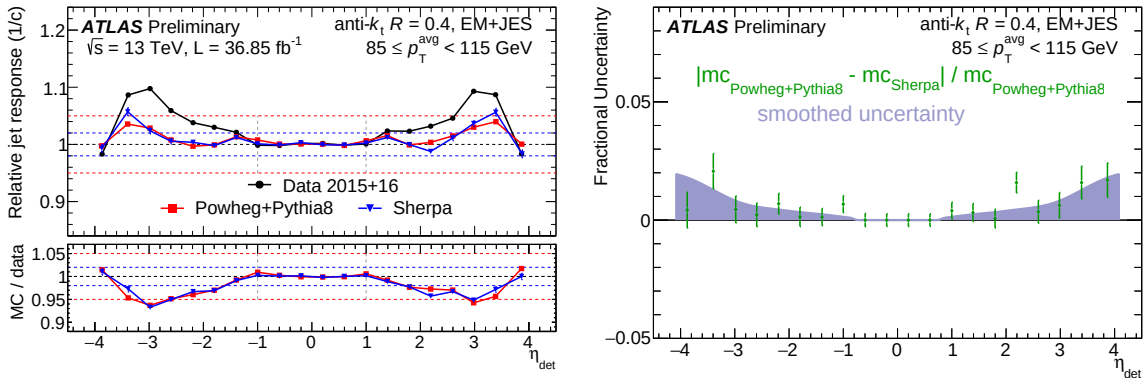
As described in section 3.1.4, systematic uncertainties related to MC simulation modeling and to event selection need to be measured. Those physics effects are not dependent on the data taking period. That is why I will only measure them using 2015+16 data and apply them to all data from various years. The choice of 2015+16 data is due to the higher statistics available at low p_T as mentioned before.
In addition, physics effects are expected to change smoothly with p_T and η. That is why I will use wider bins and increase the smoothing. This is needed to obtain enough statistical significance of the systematic uncertainties: statistical uncertainty of the systematic uncertainty should be smaller than the systematic uncertainty itself.

The new η binning and smoothing parameters that I found to give enough statistical significance are:

- $|\eta|$ bins = [0.0, 0.4, 0.8, 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.6, 4.5],
- $\sigma_{\log p_{\mathrm{T}}} = 0.4$ for modeling uncertainty and 0.25 for selection uncertainties,
- $\sigma_{\eta_{\mathrm{det}}} = \eta$ bin width.

For the modeling systematic uncertainty, two MC generators, PowhegPythia (nominal) and Sherpa (variation), are used to measure the relative jet responses. The relative difference of the responses is taken as the uncertainty. The modeling uncertainty is symmetrized since there is no reason for it to be different at $\pm\eta$. In addition, the uncertainty is constrained to always be increasing since this is the expected behavior from the modeling differences.

Figure 3.38 shows in a) the relative jet responses of the two MC generators in addition to the ones of data as a function of η. In b), the relative response difference between the two MC generators is shown in green points and the symmetrized smoothing is shown in blue-gray shading.



(a) Relative responses

(b) Modeling systematic uncertainty

Figure 3.38: Modeling systematic uncertainty, defined as the relative difference between the responses of two MC generators PowhegPythia and Sherpa, as a function of η for one bin of $p_{\mathrm{T}}^{\mathrm{avg}}$.

For the selection systematic uncertainties, the nominal and variated calibration uses the same input events and thus are highly correlated. I use the bootstrap method (3.1.4) to correctly calculate the statistical uncertainties.
For the comparison between the nominal and variated calibration, the relative difference can be calculated using the individual calibration points and then the smoothing is applied, or for each calibration points (nominal and variated) the smoothing is applied and then the relative difference is computed. I tested both methods and found that they give similar results modulo some minor effects. The latter method is used for the final uncertainties measurement.

Figure 3.39 shows the uncertainty corresponding to the up variation of the cut on $p_{\mathrm{T}}^{\mathrm{jet\ 3}}/p_{\mathrm{T}}^{\mathrm{avg}}$ as a function of η for one bin of $p_{\mathrm{T}}^{\mathrm{avg}}$. The relative difference of the calibration points is shown as green points and the two methods of uncertainty measurement and smoothing are shown in blue and red lines.

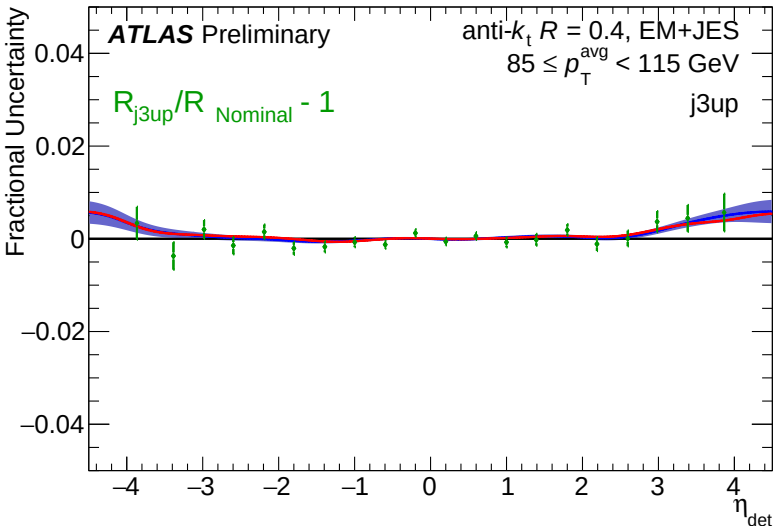
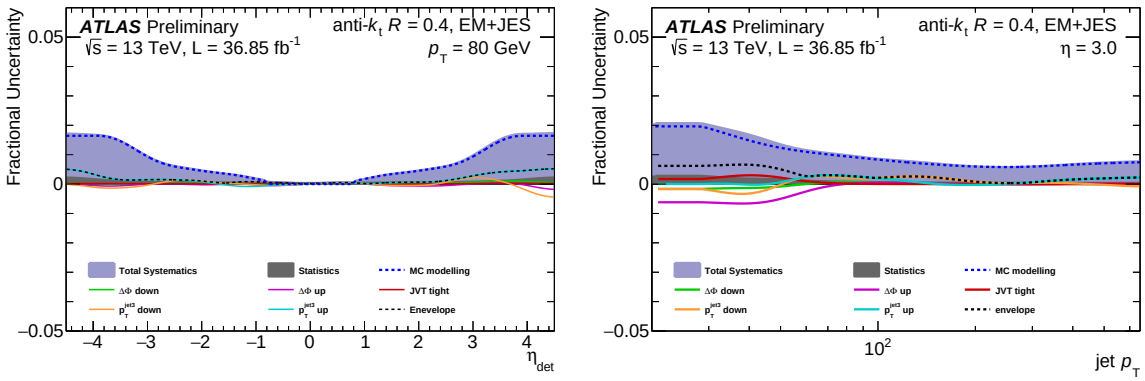


Figure 3.39: Selection systematic uncertainty corresponding to the up variation of the cut on $p_T^{\text{jet 3}}/p_T^{\text{avg}}$ as a function of η for one bin of p_T^{avg}. The two methods of uncertainty measurement and smoothing are shown in red and blue lines.

Figure 3.40 shows the individual and total systematic uncertainties for $p_T = 80$ GeV in one plot and $\eta = 3$ in the other [1]. For both, the dominating systematic uncertainty is the one coming from the modeling. The total uncertainty increases in the direction of low p_T or high $|\eta|$ as expected.



(a) $p_T = 80$ GeV

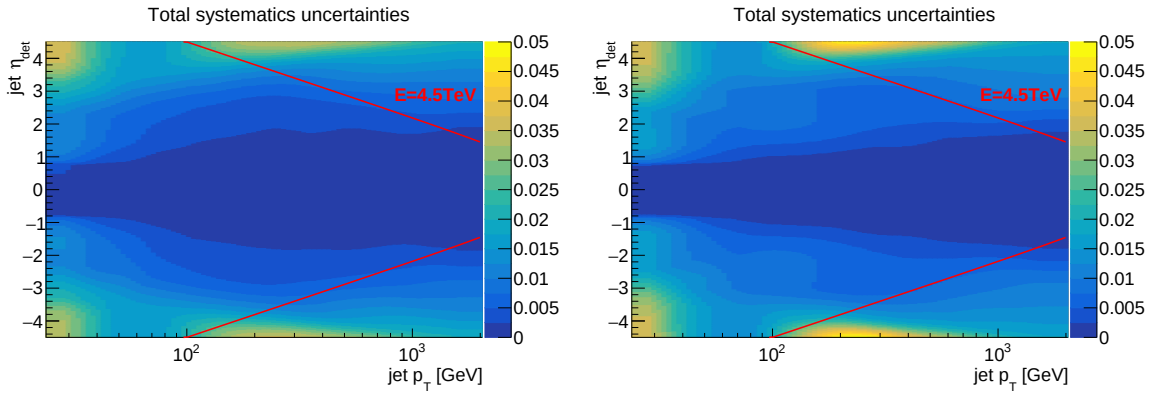
(b) $\eta = 3$

Figure 3.40: Individual and total systematic uncertainties for $p_T = 80$ GeV in a) and $\eta = 3$ in b).

Figure 3.41 shows the total systematic uncertainties 2D plots for EMTopo and EMPFlow. The uncertainties are close for both jet reconstructions. We can also see that the uncertainty increases in

[1] The statistical uncertainty shown here is not the one of the nominal calibration since the η binning is different; it is shown only for comparison reasons

the direction of low p_{T} or high $|\eta|$. For some regions with E > 4.5 TeV, the high uncertainties are just an artifact of the extrapolation of the calibration but no extra care is made since there are no jets in those regions. The maximum uncertainty for the region with E < 4.5 TeV is about 3.75% for both EMTopo and EMPFlow jets.



(a) EMTopo jets

(b) EMPFlow jets

Figure 3.41: Total systematic uncertainties 2D plots (p_{T}, η) for EMTopo and EMPFlow jets using 2015+16 data.

The eta-intercalibration results (central values, statistical and systematic uncertainties) shown before are now used in the official ATLAS calibration for the Run II period.

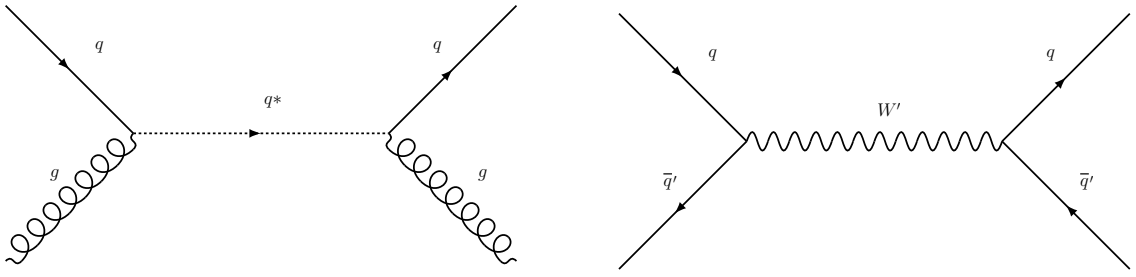
4 Direct search for new phenomena in dijet events

Switching topics to physics analyses, I present in this chapter a direct search for Beyond Standard Model signals in dijet events. After a brief introduction, I describe the analysis from the observable definition and the background estimation, to the search and limits setting techniques. Next, I detail the new folding technique I developed and implemented, and used when setting limits on model-independent signals. Last, I present the results of the analysis using 37.0 fb^{-1} of combined 2016 and 2015 data, published in the paper [76].

4.1 Introduction

As already mentioned in chapter 1, the Standard Model (SM) cannot fully explain our universe. Many questions have no answers within the SM. That is why searching new physics beyond the Standard Model (BSM) is a primary goal of research at the LHC and ATLAS.

The ATLAS experiment is used to directly search for a BSM signal. If BSM particles can interact with the proton partons (quarks, gluons), they can be produced directly in proton-proton collision and consequently produce partons when they decay. Examples are shown in figure 4.1. The decay product partons shower, hadronize and are reconstructed as jets. The production rates for BSM signals decaying to two jets can be large (as can be seen later in the results section where BSM signals can be of the same order as the data background in some phase-space regions), allowing us to directly search for a deviation from the SM predictions.



(a) q* (b) W'

Figure 4.1: Scattering examples via particles from beyond the Standard Model.

The dominant background is the production of hadronic jet pairs from $2 \to 2$ parton scattering process via strong interactions described by QCD. Two important properties of this background are:

- at high masses, the QCD dijet invariant mass m_{jj} distribution is smooth and monotonically decreasing,
- most high mass dijets production occurs in the forward direction due to the dominant t-channel poles scattering processes.

These two properties are exploited when searching for new signals. More details are given in the following section.

The most important thing in these searches is to estimate the background that comes from SM interactions. Two techniques exist:

- background estimated directly from data,
- background estimated from MC simulation.

After estimating the background, we search for a significant excess which is a direct indication of the existence of a signal beyond the standard model. If no significant signal is found, we set limits on benchmark BSM models or on generic signals.

4.2 Analysis overview

4.2.1 Observable

Resonance peak search in the dijet invariant mass, m_{jj}, distribution is a first method of searches. If the new particle has a mass within the range we are searching, its decay to two jets introduces an excess in m_{jj} distribution, localized near the mass of this resonance. Since the QCD dijet m_{jj} distribution at high mass is smooth, background is estimated directly from data through a fit. A localized excess doesn't bias the fit in a significant way.

The other property of this QCD background is that it peaks in the forward direction. Many BSM theories predict a more isotropic additional dijet production. If the new particle mass is higher than our search range to be able to see a peak but still has a smooth effect on m_{jj} distribution, the dijet angular production property is exploited for searches. This method uses the dijet angular separation defined as: $\chi = \exp^{|y_1 - y_2|}$ where y_1 and y_2 are the rapidity of each jet. Since no resonance is supposed in the probed mass range, angular search uses MC simulation to estimate the background.

Hereafter, I will only detail the resonance search method which is the one I contributed to, published in the paper [76].

Binning
The choice of binning of the dijet mass distribution is defined here. The line-shape of a resonant signal is the convolution of its intrinsic width with the parton distribution functions and the detector mass resolution. In the limit of an intrinsically narrow signal, the binning should be narrow enough to sample the line-shape over several bins while also providing numerous bins to serve as inputs for the background fit. However, the binning should not be too narrow to limit the effects of event migration from bin-to-bin due to detector resolution and avoid the signal being swamped by statistical fluctuations.

For that, the bin boundaries are chosen such that the width is equal to the resolution at the given mass (and the boundaries are forced to integer values). The dijet mass resolution is calculated as the width of a Gaussian fit to the ratio of the reconstructed dijet mass over the truth dijet mass, $m_{jj}^{Reco}/m_{jj}^{Truth}$. Figure 4.2 shows the dijet mass resolution.

The efficiency (purity) is calculated as the fraction of events in a truth (reconstructed) m_{jj} bin that falls in the same corresponding reconstructed (truth) m_{jj} bin. The two observables are used as an indicator that the properties of the events are changing smoothly over the mass range. Figure 4.3 shows both the smooth efficiency and purity evolutions.

The m_{jj} binning used for the background fit, the resonance search and the limits setting contains 91 bins, ranging between 1.1 and 8.2 TeVwith a relative bin width decreasing from 30% to 19%.

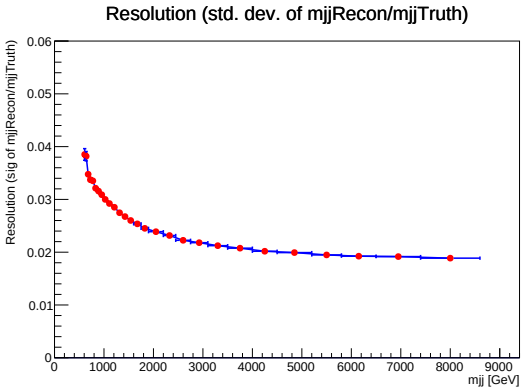
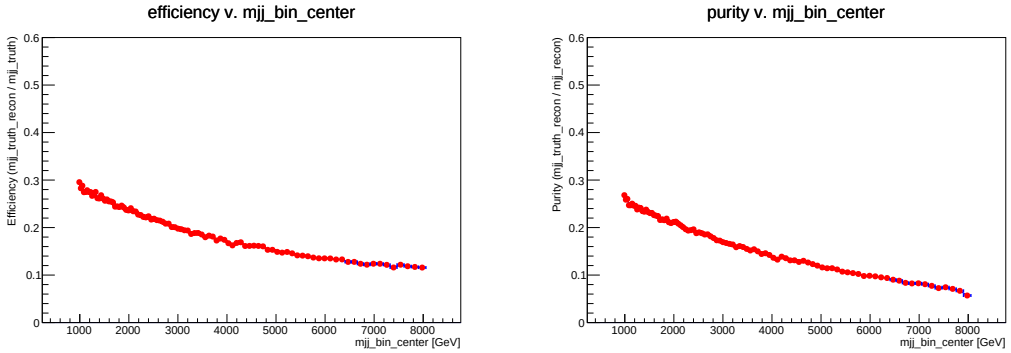


Figure 4.2: The dijet mass m_{jj} resolution as a function of m_{jj}^{Reco}. The resolution is defined as the width of a Gaussian fit to $m_{jj}^{Reco}/m_{jj}^{Truth}$.



(a) Efficiency

(b) Purity

Figure 4.3: Efficiency(a) and purity(b) for the chosen bin boundaries. Efficiency is defined as the fraction of events in a given truth m_{jj} bin where the reconstructed m_{jj} falls in the same range. Purity is defined as the fraction of events in given reconstructed m_{jj} bin where the truth m_{jj} falls in the same range.

Event selection

The event selection uses the following cuts. First all jets are required to have $p_T > 60$ GeV. This requirement allows removal most of pile-up jets. Since we use only one non-prescaled trigger, HLT_j380, an additional cut on the leading jet is set to $p_T^{j1} > 440$ GeV. This requirement ensures that we use the trigger with an efficiency higher than 99.5%, since the trigger selection asks for a high level trigger jet above 380 GeV.

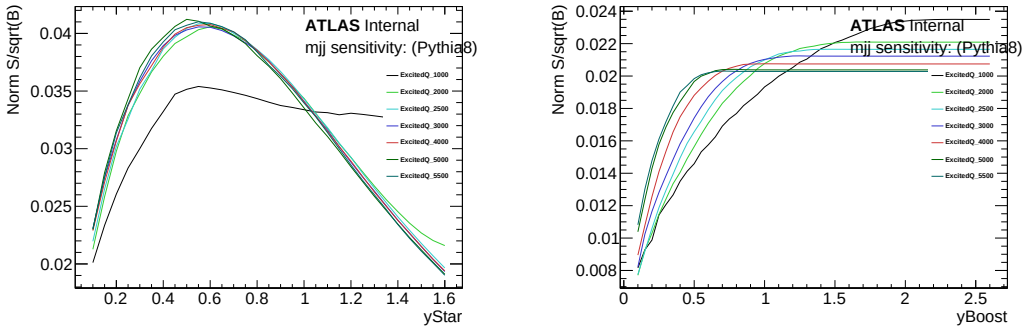
Additional cuts are used to exploit the angular distribution difference between background and signal. As already mentioned, the QCD dijet background is dominated at high mass by t-channel processes where the production is proportional to $(1 - \cos\theta)^{-2}$ where θ is the angle between the 2 jets. In contrast, most BSM models productions follow a polynomial in $\cos\theta$ (for example q^* model, defined in chapter 1.4, is expected to be flat in $\cos\theta$). Cuts on dijet angular variables are studied to optimize

the sensitivity. These variables are:

$$\text{angular separation: } y^* = \frac{y_1 - y_2}{2}, \tag{4.1}$$

$$\text{angular boost: } y_B = \frac{y_1 + y_2}{2}. \tag{4.2}$$

Figure 4.4 shows the binned significance (sum of $S/\sqrt{B}$ in quadrature for each bin, where S and B are respectively the number of signal and background events) for the m_{jj} distribution as a function of the upper cut value of y^* and y_B for various q^* signals. The cuts are chosen to maximize this significance. A cut of $|y^*| < 0.6$ is optimal. For y_B, no cut is chosen.



(a) y^* (b) y_B

Figure 4.4: Optimization of resonance search cuts for q^* signals with different masses: 1 TeV (black), 2 TeV (light green), 2.5 TeV (cyan), 3 TeV (blue), 4 TeV (red), 5 TeV (dark green), 5.5 TeV (turquoise). The following variables are shown here: (a) y^*, (b) y_B.

The significance was also studied as a function of other variables but no additional cut was found to be beneficial.

4.2.2 Background estimation

In the search for resonances in the m_{jj} spectrum, a data-driven estimate is used for the SM contribution. The m_{jj} spectrum is fit by a smooth functional form,

$$f(x) = p_1(1-x)^{p_2} x^{p_3 + p_4 \ln x + p_5 (\ln x)^2}, \tag{4.3}$$

where the p_i are fit parameters (p_4 and p_5 are forced initially to zero), and $x \equiv m_{jj}/\sqrt{s}$. Previous studies [77] have found this ansatz to provide a satisfactory fit to the dijet mass distribution predicted by leading and next-to-leading-order QCD MC simulation. It is also able to describe the dijet mass distributions observed in the data from all prior hadron collider experiments. As the m_{jj} distribution predicted by QCD is a complicated convolution of t- and s-channel contributions, quark and gluon PDFs, detector resolution, and kinematic selections, it is a remarkable empirical observation that it can be approximated by a single analytic function with only a handful of parameters. The x^p

term is motivated by the LO QCD matrix element while $(1-x)^p$ is a common parametrization for the behavior of parton distributions with the nice property of vanishing as x approaches unity. As searches have applied the technique to larger and larger data-sets, spanning wider and wider ranges in dijet mass, factors of the form $x^{\ln(x)^n}$ where added to the fit function to be able to properly describe the full m_{jj} spectrum.

A feature of the functional form used in the fitting is that it allows for expected background variations but does not accommodate localized excesses that could indicate the presence of new physics signals. However, the effects of smooth deviations from QCD, such as contact interactions, could be absorbed by the background fitting function, and therefore this background estimation technique is used only to search for resonant effects.

Previously, the full mass range was fitted with this functional form. At a first step, only three-parameters are used. Using likelihood ratio with Wilk's approximation as a test statistics, if the current function is judged to not fit the data well, it is dropped in favor of the next higher order function and re-tested.

In an effort to prevent the possible breakdown of our fit function at high integrated luminosity, the global function fit has been replaced by the Sliding Window Fit method (SWiFt), replacing a fit on the full spectrum with a sliding localized fit on smaller m_{jj} ranges where we expect the function in Equation 4.3 to properly model the QCD background contribution even with very high statistics. SWiFt produces a non-parametric global background model.

In smaller mass windows, the function is fitted to the data and the evaluation of the function at the window's central bin is taken to be the background estimation for the bin. By sliding over the entire mass range, the background is estimated this way bin-by-bin. The window is not able to be centered at every point in the spectrum, and so in those edge cases the bins below (or above) are filled using the result from the lowest (highest) window fit.

Figure 4.5 shows the SWiFt background made using the three-parameter version of the analytical function in Eq. 4.3. It also shows the background prediction obtained from a fit to the whole spectrum with a four-parameter version. The bottom panel shows the difference in the background prediction of the two fitting methods. The differences we see are much smaller than the statistical uncertainties on the data σ_{data}.

The window size, defined in terms of number of mass bins to the left and right of a window center, is chosen by examining several statistical tests. The largest window size for which these tests are reasonable is taken to be the nominal window size. The statistical tests are the χ^2/ndf test (measures in general the agreement between two distributions), the Kolmogorov–Smirnov (KS) test (more sensitive to differences on the edges - picks out lopsided distributions, etc) and the Wilks test (measures nominal vs alternative fit's agreement with data). Specifically, we require the χ^2/ndf to be below 2, the KS p-value and the Wilks p-value to be above 0.05 (for detailed discussions about the statistical tests, refer to [73, 78, 79]). The difference between the nominal SWiFt background (using the three-parameter function) and the alternate SWiFt background (using four-parameter function) are assessed as a systematic uncertainty (see next section).

Multiple window combinations of left and right sizes were tested. The statistical measures of a subset of window sizes are shown in figure 4.6. The window size with 30 bins to left and 20 bins to the right

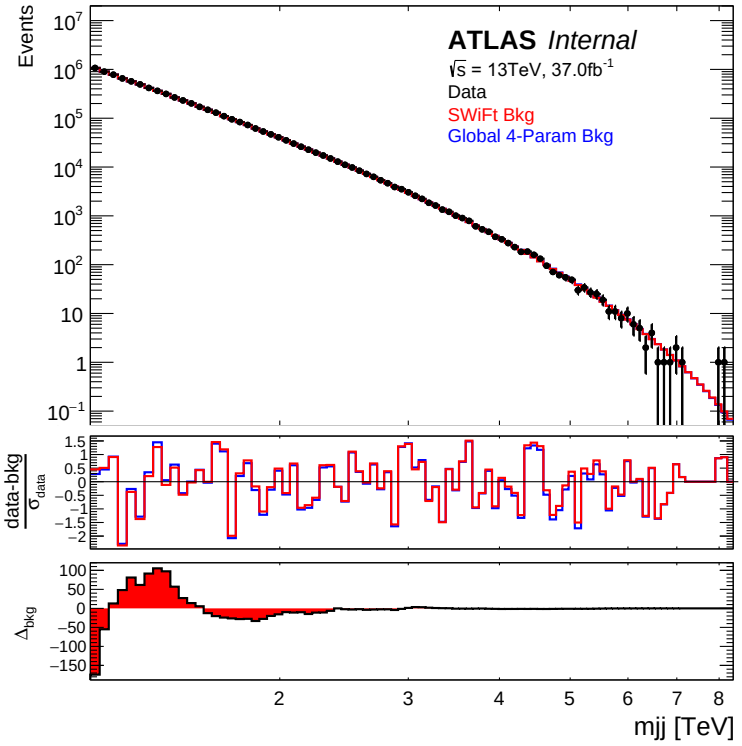
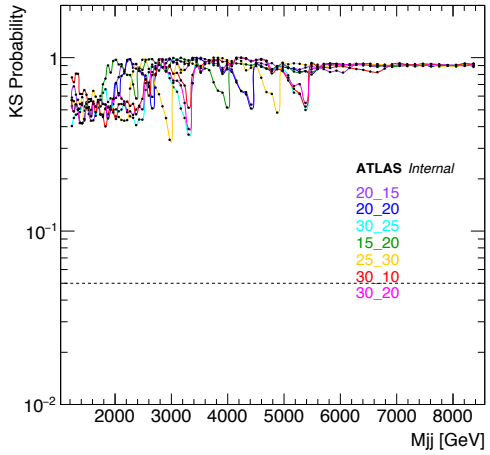


Figure 4.5: Data driven SWiFt background, made using the three-parameter dijet function, shown with the resonance selection dijet data for 37.0 fb^{-1} of Run II data. Also shown is the four-parameter function fit to the data using the old global fitting method. The second panel shows the residuals of data with respect to the SWiFt and the global fitting method derived background. The third panel shows the difference in the predicted events in each m_{jj} bin from the SWiFt and the global fitting method.

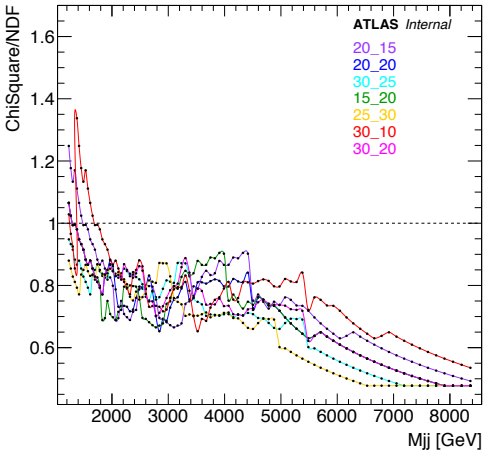
of the center is chosen. Figure 4.6(d) also shows the difference between the background made with the chosen window size and the background made with the other window sizes. The differences are small with respect to the number of events in each bin, smaller than the statistical uncertainties of the background.

Two sets of backgrounds are created under different assumptions:

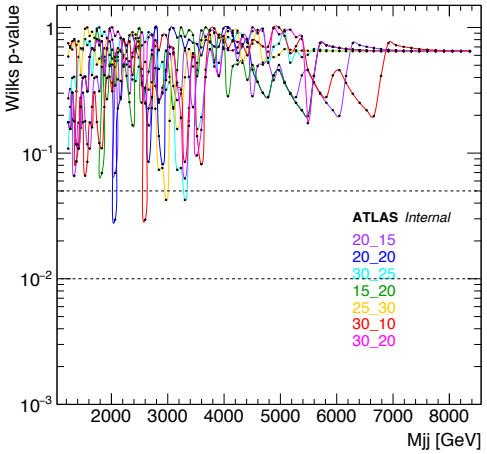
- No signal subtraction: this is used as the nominal background in plots and is used as an input to search algorithm BumpHunter (see later for more details), which identifies if there is a significant excess in the data. The background is estimated by smoothing the data with the analytic function (equation 4.3).
- Subtraction of the signal at a specific mass point: this is used as the background for the Bayesian limit-setting machinery. Backgrounds are made by smoothing the data after signals are subtracted at the mass for which the limits are being set. This produces a separate global background estimate for each mass point and each benchmark signal considered.



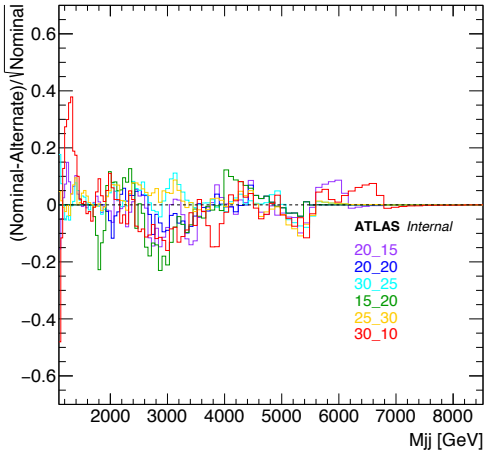
(a) KS p-value



(b) χ^2/ndf



(c) Wilks p-value



(d) Difference between chosen and other background from different window sizes

Figure 4.6: Statistical tests examined window-by-window to pick background window size for the 37.0 fb^{-1} of dijet data. a) KS p-value comparing fit to data in each window, b) χ^2/ndf comparing fit to data in each window, c) Wilks p-value comparing the SWiFt three-parameter nominal and SWiFt four-parameter alternate backgrounds. Figure d) shows the difference between the background made with the chosen window and the backgrounds made with the other window sizes.

4.2.3 Systematic uncertainties

Systematic uncertainties are evaluated for the background and for the signal.

Systematic uncertainties on background

The systematic uncertainties affecting the data-driven background estimation are:

- the uncertainty on the choice of the fit function, since the forms selected for testing are somewhat arbitrary and other parameterizations could also have been suitable;
- the statistical uncertainty on the fit result which depends on the statistical precision of the data, affecting the precision of the determination of the parameters of the fitted function.

The following paragraphs explain how they are estimated.

Uncertainties on the choice of fit function:
As a nominal fit function we consider the three-parameter version of Eq. 4.3 while as the alternative we consider the four-parameter version. The nominal and alternative background are compared and the difference between the two is used to derive an uncertainty. A collection of pseudo-data are thrown from the nominal background result, and from each pseudo-experiment both a nominal and alternative backgrounds are derived. The mean of the difference between the nominal and alternative background is recorded in each bin and is used to define the size of the uncertainty on the function choice.

The uncertainty is not symmetrized, but rather a direction for the uncertainty is chosen in each bin using the direction from the nominal to the alternative fit in data. The maintenance of this directionality allows for realistic variation of the parameterization in the limit setting phase. A single nuisance parameter is defined to correspond to the function choice; its variation is used to scale linearly between the nominal choice and an alternative representing the average distance to a second valid parameterization. This is done under the assumption that a variation in the opposite direction would be unreasonable as it would fall outside the envelope of trustworthy descriptions parameterized by the two functions.

Uncertainties on the parameters of the fitting function:
The second uncertainty is the one associated with the quality of the fit itself. Under ideal circumstances this would be derived as a confidence band on the function determined by the covariance matrix of the fitted parameters. However, in cases where the parameters of the function are strongly correlated, or whenever the likelihood function has a badly-behaved maximum, it is not possible for numerical algorithms to accurately compute this covariance matrix.

Since the confidence interval on a function is meant to represent the 1σ region within which the fit would fall in the large-number limit of repeated trials, it can also be found by throwing pseudo-experiments and fitting each. This method does not need accurate estimation of the parameter uncertainties; instead, the pseudo-experiments are generated using Poisson statistics based on the nominal background model after a fit to data. Each pseudo-data is fitted using the same starting conditions as the observed data, and the uncertainty on the fit in each bin is defined to be the RMS of the function value in that bin for all the pseudo-experiments.

Figure 4.7 shows both of these uncertainties for 37.0 fb^{-1} of Run II data.

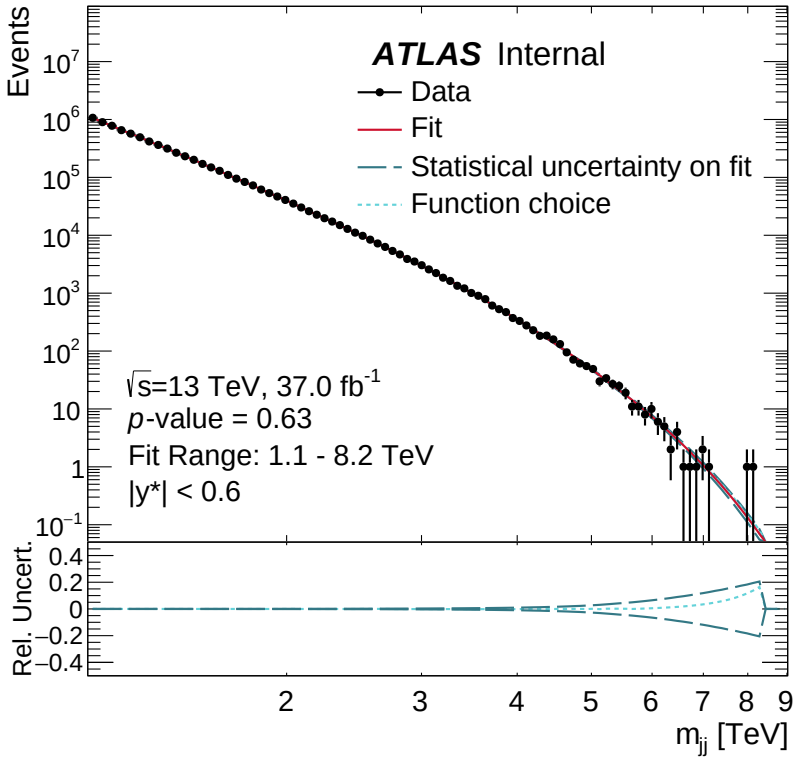


Figure 4.7: The nominal fit function (SWiFt three-parameter), its statistical uncertainty and the alternative fit function (SWiFt four-parameter) are shown. In the bottom pad, the cyan line shows the one-sided uncertainty on the choice of the fit function, while the blue line shows the statistical uncertainty on the fit.

Systematic uncertainties on signal

The following uncertainties are considered on the signal when setting limits:

- luminosity uncertainty,
- jet energy scale, JES, uncertainty,
- PDF and scale uncertainties.

As previously, a description of the estimation of each uncertainty is given below.

Luminosity uncertainty:
A luminosity uncertainty is applied as a scale factor to the normalization of the signal samples. The uncertainty in the combined 2015+2016 37.0 fb^{-1} integrated luminosity is 3.2% (for the luminosity and its uncertainty determination, refer to [80]).

JES uncertainty:
The jet energy scale uncertainty is applied to the signal, using a reduced set of nuisance parameters, by shifting the four-vectors of all jets by 1σ of the uncertainty and fitting the result to determine the shift in peak location as a percentage of signal mass. The uncertainty ranges from 1.5% at the lowest

masses to 3% for masses above 4.5 TeV.
For the generic signals, in our case the Gaussian signal, the estimate of the uncertainty is derived using the q^* signal template that is closest in mass.

The jet energy resolution and angular resolution uncertainties has been found to be negligible. As such, they are ignored.

PDF and scale uncertainties:
They mostly affect the normalization of the signal. A 1% flat systematic is applied, evaluated in previous search analyses and found to be conservative.

4.2.4 Search and Limits setting techniques

Search technique

The main statistical test employed in the dijet resonance search is based on the BumpHunter algorithm [81, 82] and is used to establish the presence or absence of a resonance in the dijet mass spectrum. The algorithm operates on the binned invariant mass distribution, m_{jj}, obtaind from the data. The bins are treated as independent and the data are assumed to be Poisson-distributed around a theoretical value given by the hypothesis coming from the background fit. The algorithm compares the background estimate with the data in mass intervals of varying widths formed by combining neighboring bins. Starting with a two-bin signal window, the algorithm scans across the entire distribution, then steps through successively larger signal windows up to half of the whole fit range. For each point in the scan, it computes the p-value of the difference between the data and the background. The p-value is defined as the probability of finding a difference at least as large as the difference observed between the data and the hypothesis. The most significant part away from the smooth spectrum, the so-called "bump", is defined by the set of bins that have the smallest probability of arising from a Poisson background fluctuation. During this procedure, the background model is not changed or refit to the data outside of the excluded region.

The BumpHunter algorithm accounts for the so-called "look-elsewhere effect" [83, 84], by performing a series of pseudo-experiments drawn from the background estimate to determine the probability that random fluctuations in the background-only hypothesis would create an excess anywhere in the spectrum at least as significant as the one observed.

One must also ensure the background estimate is not biased by a potential signal. If the most significant local excess from the background fit has a p-value larger than 0.01, no statistically significant excess has been observed and the analysis proceeds to the limit-setting stage. If it has a p-value smaller than 0.01, this region is excluded and a new background fit is performed. Then the result of this fit is used for a second run stage of the BumpHunter algorithm where an unbiased estimate of the global significance of any excess is obtained. This exclusion was not needed for the current analysis as will be seen later on.

Limits setting technique

For each mass point, m_R, of a studied (benchmark or generic) signal (see chapter 1.4), a Bayesian method is applied to data at the same mass to calculate a posterior likelihood and set an observed 95% CL (credibility level) upper limit on the cross section times acceptance times branching ratio, $\sigma \times \mathcal{A} \times \mathrm{BR}$, for the new physics resonant signal as a function of m_R. The limit on $\sigma \times \mathcal{A} \times \mathrm{BR}$ from data is interpolated logarithmically between mass points to create a continuous curve in m_{jj}.

This form of analysis is applicable to all resonant phenomena where the new physics resonance couplings are strong compared to the scale of perturbative QCD at the signal mass, so that interference with QCD terms can be neglected. The acceptance calculation includes all reconstruction steps and analysis cuts.

In the Bayesian marginalization, the parameter of interest corresponds to the normalization of the signal template which is given a constant prior. One nuisance parameter with a Gaussian prior is introduced for each of the systematic uncertainties on the data-driven background and on the signal described in section 4.2.3. These uncertainties are incorporated into the marginalization by convolving them with the posterior probability distribution. Credibility intervals are calculated numerically from the resulting convolutions. The posterior probability is then integrated to determine the 95% CL interval.

Expected 95% upper limits are also calculated using a profile likelihood method, where the nuisance parameters are allowed to float while the signal yield is scanned upwards. The 95% upper limit is defined as the number of events N such that LLH(N) = LLH(base) + 1.92, approximately 2σ worse than the base case. For positive signals, LLH(Base) is the likelihood of the best-fit signal. Otherwise, it is the likelihood corresponding to zero signal events. At each mass point, a large number of pseudo-experiments are thrown based on a nominal template defined from the best fit values of all nuisance parameters with signal fixed to zero. For each pseudo-experiment, the limit is calculated. The central value of the expected limit is taken to be the median of the pseudo-experiment distribution. The 1 and 2 sigma bands are taken to be median ± 34% and median ± 47.5% respectively.

In previous iterations, limits on generic signal shapes were provided on the reconstructed level which includes the detector effects. In the following section, I present a new technique I developed and implemented which improves the way we put limits on generic signal shapes by calculating them on the particle (hadron) level, allowing for easier re-interpretation after their publication.

4.3 Folding technique

In this section, I present the description, tests and results of a new technique I developed and implemented which was added to the dijet search analysis as one of the two main improvements of the analysis (the second improvement being the SWiFt technique described in section 4.2.2).

4.3.1 Motivation

In past studies [85–87], limits were set on signal models that can contribute to the reconstructed m_{jj} distribution [1]. Often, these studies considered signals with some given shape (for example Gaussian) and amplitude at detector level. Doing so, the description of the corresponding distribution folds the actual physical signal and detector effects (acceptance and resolution). Here I describe a method for folding a truth-level model with the detector response, allowing factorization of physics and detector effects. One goal is to enable the publication of limits on a given signal model, which can be interpreted from the phenomenological point of view at particle (hadron) level, without further need of information on the detector response. This method should also allow for a better understanding of various features in the distributions, yield potentially more stable fits, facilitate the study of constraints on the binning due to the jet energy resolution (JER) and facilitate the interpolation between various parameter points (for example mass, width) for which full simulation samples of a given signal model were generated.

4.3.2 Method description

The method described here consists in defining a model for the "truth" distribution at particle level, within a set of fiducial cuts. This model is then folded with the detector effects. One can compare the result of this convolution with the data distribution, for example when performing fits or when setting limits.

The detector effects are described through a Monte Carlo-based transfer matrix, relating the truth and reconstructed observables. Events passing both the fiducial selection at truth level and the same selection as data at reconstructed level, called "matched" events in the following, are used to fill a 2D distribution: a matrix N_{ij}, where i is the bin number corresponding to the truth quantity and j the bin number corresponding to the reconstructed one. I then derive a folding matrix for the "matched" events, giving the probability for an event generated in a bin i at truth level to be reconstructed in a bin j:

$$A_{ij}^{\text{matched}} = \frac{N_{ij}^{\text{matched}}}{\sum_k N_{ik}^{\text{matched}}}. \tag{4.4}$$

[1] Here I study the folding method in the context of the m_{jj} distribution, but the same procedure can be applied to other observables.

I also define, in a given truth (reconstructed) bin, the fraction of events that pass both the truth and reconstructed level selections, among the total number of events at truth (reconstructed) level:

$$\epsilon_i^T = \frac{\sum_k N_{ik}^{\text{matched}}}{\sum_k N_{ik}^{\text{matched}} + N_i^{T\ \text{only}}}, \tag{4.5}$$

$$\epsilon_j^R = \frac{\sum_k N_{kj}^{\text{matched}}}{\sum_k N_{kj}^{\text{matched}} + N_j^{R\ \text{only}}}. \tag{4.6}$$

These fractions of events are also called "matching efficiencies".

The folding of a truth model $f(m_{jj}^T)$ with the detector effects is done as follows:

$$f_i \longmapsto F_j = \sum_i f_i * \epsilon_i^T * A_{ij}/\epsilon_j^R \equiv \sum_i f_i * \tilde{A}_{ij}, \tag{4.7}$$

in order to obtain the reconstructed level model $F(m_{jj}^R)$, where A and ϵ's are the folding matrix and the efficiencies defined previously. The global folding matrix

$$\tilde{A}_{ij} \equiv \epsilon_i^T * A_{ij}/\epsilon_j^R \tag{4.8}$$

accounts for the event migrations and matching (in)efficiencies caused by detector effects.

In order to avoid large simulation-based (model dependent) extrapolations, it is preferable to use truth level fiducial cuts that are close to the ones applied at reconstructed level. I use here the same cuts for both levels.

The statistical uncertainties affecting the result of the folding are also calculated. These uncertainties are caused by the statistical fluctuations affecting the global folding matrix ($\tilde{A}$) and by the ones impacting the original truth distribution (when the latter is not obtained from a smooth function, but rather from a sample of events). When evaluating these statistical uncertainties, the correlations among the various inputs (correlations among the bins of the A_{ij}, ϵ_i^T and ϵ_j^R) must be accounted for. This is achieved using the Bootstrap method [88], where pseudo-experiments are generated fluctuating the weight of each event based on a Poisson distribution with the mean parameter equal to 1. The seed of the random number generator is defined independently for each event, based on the event number. This approach guarantees that the pseudo-experiments are generated coherently when the same events play a role at several levels in the analysis (for example for efficiencies and for the folding matrix), allowing hence to take into account the corresponding correlations.

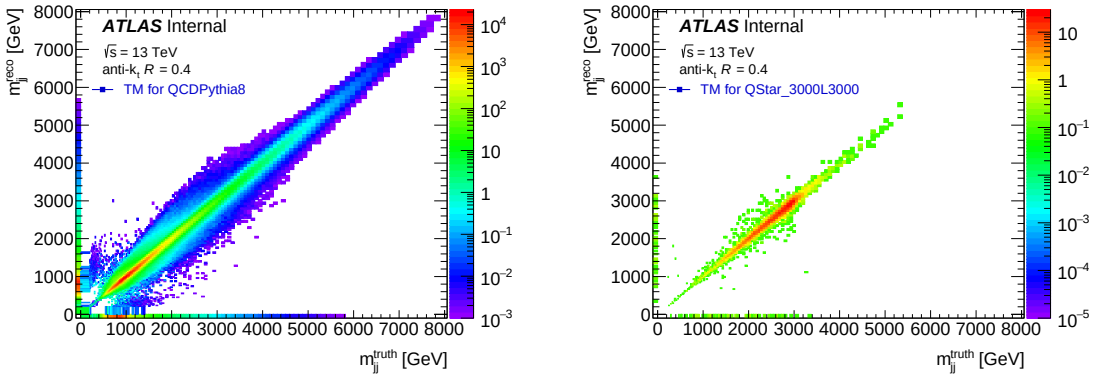
A geometrical matching criteria can optionally be applied when building the transfer matrix. This can be done by requiring for a given event that the truth and reconstructed leading and sub-leading jets are geometrically matched within $\Delta R(\text{jet}_{\text{reco}}, \text{jet}_{\text{truth}}) < 0.3$ [2]. The matching criteria is considered as being satisfied if either the (leading-leading; subleading-subleading) or (subleading-leading; leading-subleading) pairs of jets satisfy the ΔR condition. Such events are included in the category of "matched events" of the transfer matrix. The entries from the other events are added both to the unmatched truth and reconstructed categories. The effect of the geometrical matching criteria is evaluated in the following section and this requirement is applied only when explicitly stated.

[2] $\Delta R = \sqrt{\Delta\phi^2 + \Delta y^2}$

4.3.3 Folding matrices from different MC samples

The characteristics of the events at hadron level (event topology, quark/gluon composition impacting the hadron content of the jets, etc...) impact the average detector response and resolution, hence the folding matrix. It is therefore interesting to compare the inputs of the folding procedure (matching efficiencies, the folding matrix for matched events, as well as the global folding matrix) for various signal and background MC samples.

Figure 4.8 shows the transfer matrices, including both matched events and events passing only the truth and respectively reconstructed level selections, obtained from two MC samples. Each event contributes with the corresponding weight to this matrix. These transfer matrices are used to derive the efficiencies and folding matrices discussed below.



(a) QCD Pythia 8

(b) q*(3 TeV)

Figure 4.8: Transfer matrix between the truth and the reconstructed level m_{jj} based on two MC samples: one from QCD and one from q*(3 TeV). The horizontal line (vertical column) at 0 GeV correspond to events that pass the selection only at truth (reconstructed) level.

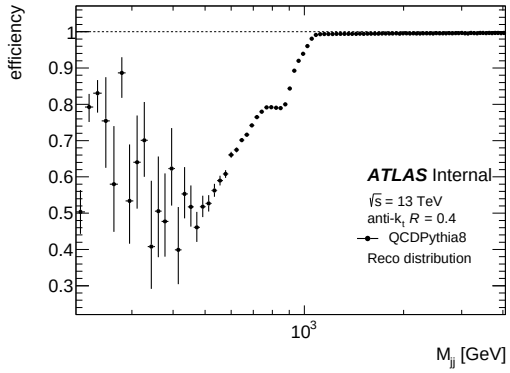
Figure 4.9 shows the matching efficiencies at truth and reconstructed level. While these efficiencies can be as low as ≈ 40% at low mass, they are ≈ 100% for $m_{jj} > 1$ TeV.

In figure 4.10, I show the folding matrix for matched events, while in figure 4.11, I show the global folding matrix, obtained by combining the folding matrices of the various samples with the corresponding matching efficiencies, following the procedure indicated in equation 4.8.

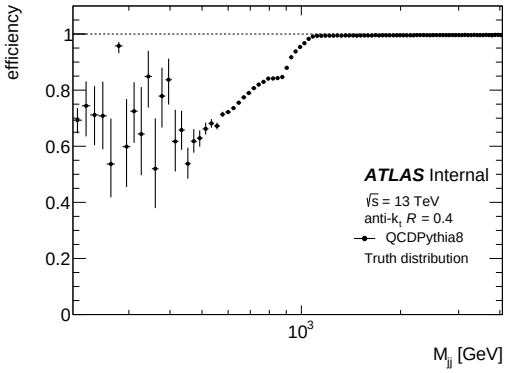
Figure 4.12 shows the transfer and global folding matrices obtained from the QCD Pythia8 samples, with the geometrical matching criteria applied. The matrices with and without this criteria are very similar, except a small difference for $m_{jj}^{truth} < 400$ GeV.

Those $\tilde{A}$ matrices are used in the next sections to assess their effect on the folded signal distributions and on the limits setting levels.

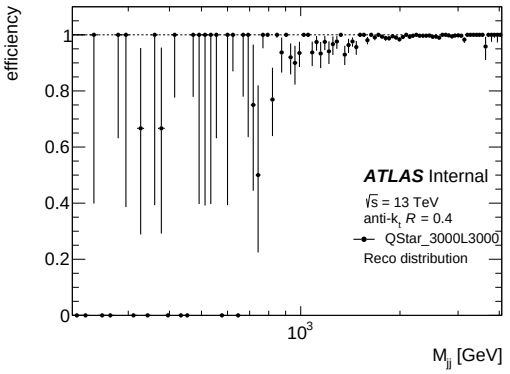
Figure 4.13 shows the m_{jj} detector response distribution in three MC samples, QCD, q*(3 TeV) and W′(3 TeV). Comparing the Gaussian fit σ, the core of the resolution are very similar between



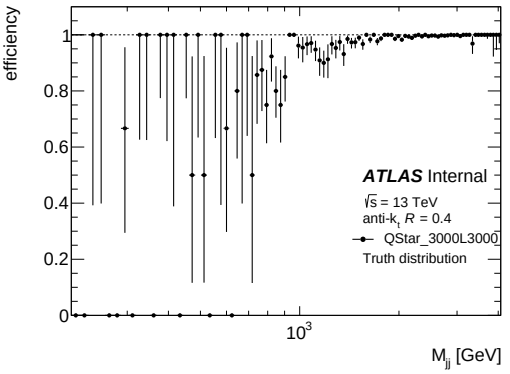
(a) ϵ^{R} for QCD Pythia 8



(b) ϵ^{T} for QCD Pythia 8



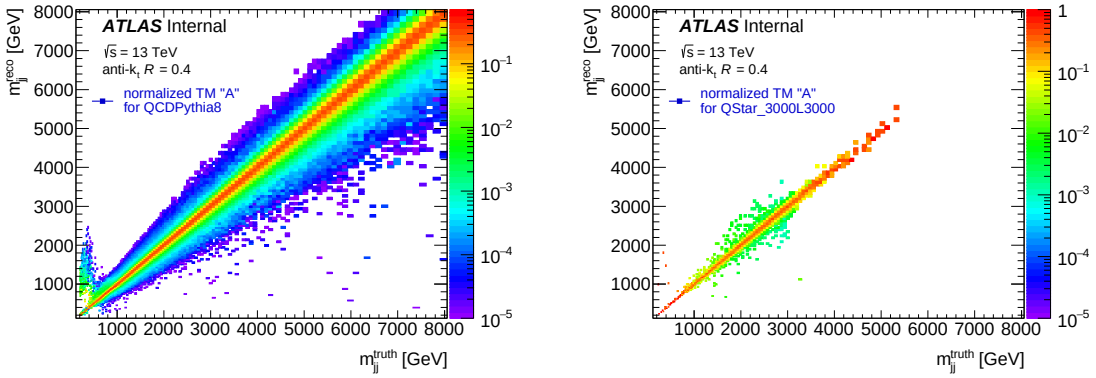
(c) ϵ^{R} for q*(3 TeV)



(d) ϵ^{T} for q*(3 TeV)

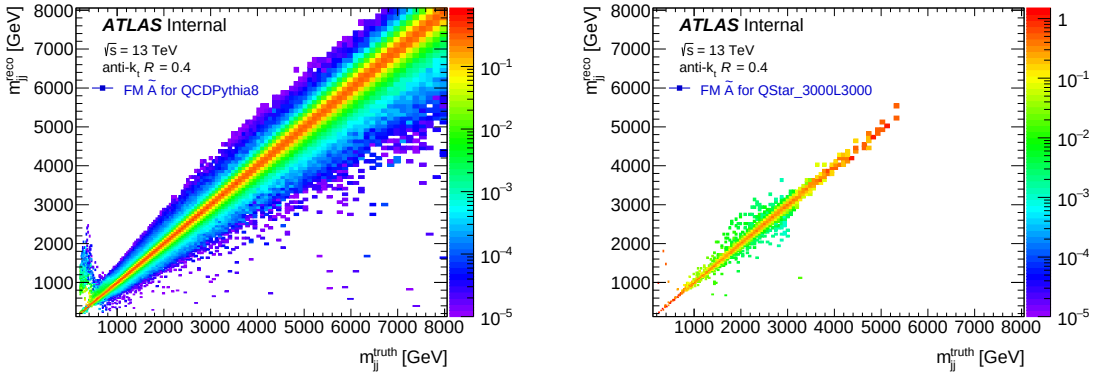
Figure 4.9: Matching efficiencies at truth and reconstructed level for two MC samples: one from QCD and one from q*(3 TeV).

QCD and q*, whereas a small difference is visible between QCD and W′(3 TeV). The different BSM scenarios used here are explained in details in chapter 1.4. Let me note that this comparison does not provide any information about the tails of the resolution.



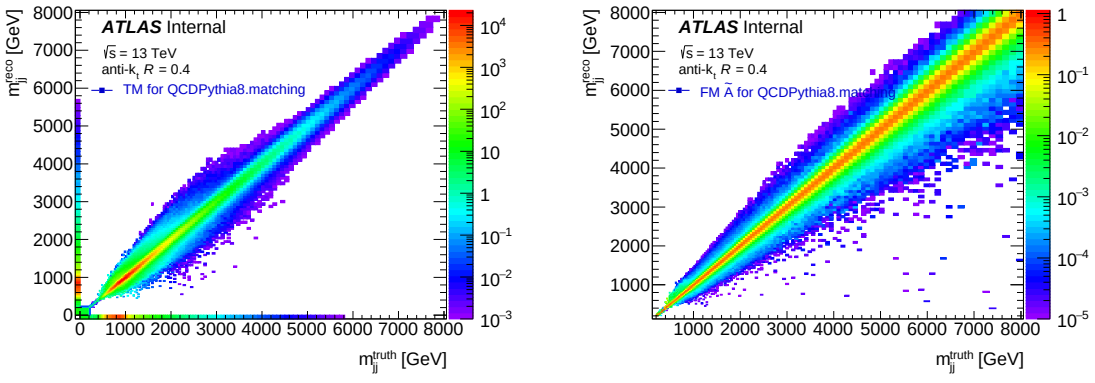
(a) QCD Pythia 8 (b) q*(3 TeV)

Figure 4.10: Folding matrix for the "matched" events ($A^{matched}$) for two MC samples: one from QCD and one from q*(3 TeV).



(a) QCD Pythia 8 (b) q*(3 TeV)

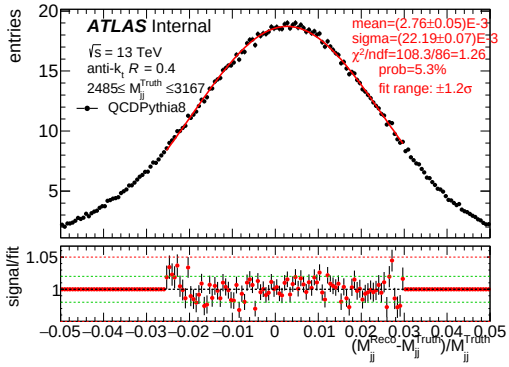
Figure 4.11: Global folding matrix ($\tilde{A}$) for two MC samples: one from QCD and one from q*(3 TeV).



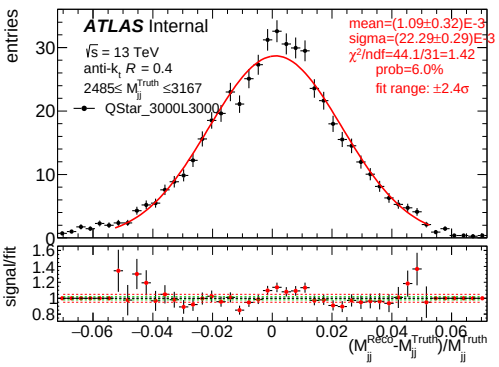
(a) Transfer matrix

(b) Ã matrix

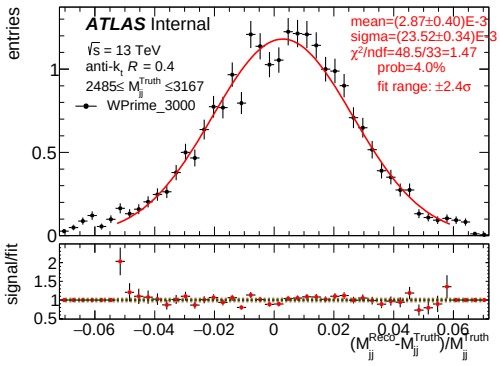
Figure 4.12: Transfer and global folding matrix (Ã) for the QCD Pythia8 sample, with the geometrical matching criteria applied.



(a) QCD Pythia 8



(b) q*(3 TeV)



(c) W′(3 TeV)

Figure 4.13: m_{jj} detector response distribution in three MC samples, QCD, q*(3 TeV) and W′(3 TeV), for the truth m_{jj} range [2.485; 3.167] TeV. The result of a Gaussian fit of this distribution is also shown.

4.3.4 Tests on the folding procedure

In this section, I show the tests done to check the compatibility of the folding technique with the full simulation of events (detector level).

A first test of the procedure I did consists in folding (following the procedure in equation 4.7) the truth distribution of a given MC sample with the detector effects described by a global folding matrix derived from the same MC sample. Figure 4.14(a) shows the truth m_{jj} distribution, the reconstructed distribution obtained from the sample of fully simulated events and the distribution resulting from the folding procedure. As expected, in this case, a perfect agreement is achieved between the two reconstructed distributions.

Next, I test the folding of the truth distribution of a given signal MC sample using a global folding matrix derived from a QCD sample. The truth m_{jj} distribution of the signal, the reconstructed distribution obtained from the sample of fully simulated events and the distribution resulting from the folding procedure are shown in figure 4.14(b) and figure 4.15. In this case, some differences are observed between the two reconstructed distributions, mainly in the peak region. Such differences could be due to the change in the average detector response and resolution (see figure 4.13) for various event topologies and various parton contents, although the statistical significance of the effects observed here is rather limited.

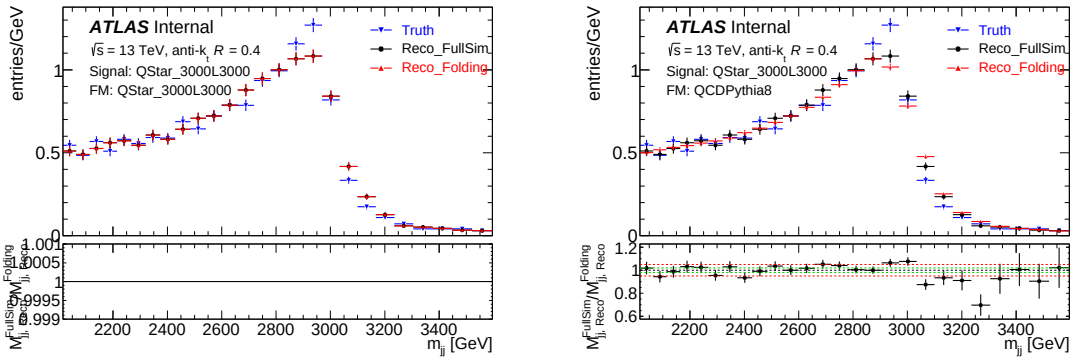
For the contact interaction (CI 1.4) signal (with $\Lambda = 7$ TeV) (figure 4.15(e) and figure 4.15(f)), which changes the ratio between quark and gluon jets, we see a systematic shift of the ratio between the fully simulated and folded distribution of about 2% for $m_{jj} > 4$ TeV (where the CI signal becomes non negligible). I use in the following sections $\tilde{A}$ matrices from CI samples to probe the effect of the choice of the signal samples for building the $\tilde{A}$ matrices, when setting limits using the folding method.

Figure 4.16(a) shows the relative uncertainties of the m_{jj} distributions obtained from the simulation and from the folding procedure respectively for a $q^*(3\text{ TeV})$ signal folded with $\tilde{A}$ matrices from QCD. For the truth and the fully simulated reconstructed distributions, the relative uncertainties are very similar, as expected. The difference due to the effect of the detector resolution on the fully simulated reconstructed distribution is clearly visible between 3 and 3.2 TeV. On the other hand, the reconstructed distribution from the folding procedure has smaller relative uncertainties. This is not due to a gain in statistics, but rather to the correlation between bins induced by the folding procedure. These correlations are evaluated using pseudo-experiments (pe) where all the inputs (m_{jj}^{truth} and $\tilde{A}$) used to derive the folding result are fluctuated, applying the following formulas:

$$\begin{aligned} \text{correlation}(M_i, M_j) &= \text{cov}(M_i, M_j) / \sqrt{\text{cov}(M_i, M_i) \times \text{cov}(M_j, M_j)}, \\ \text{covariance} \equiv \text{cov}(M_i, M_j) &= \sum_{pe=1}^{N_{pe}} \left[\left(Y^{pe}(M_i) - Y^0(M_i) \right) \left(Y^{pe}(M_j) - Y^0(M_j) \right) \right] / N_{pe}, \end{aligned} \tag{4.9}$$

where M_i is the bin index, N_{pe} is the number of pseudo-experiments, Y^{pe} is the bin entries in the corresponding pseudo-experiment and Y^0 is the bin entries in the nominal distribution. The full correlations are shown in figure 4.16(b). To distinguish the different sources of uncertainties and correlations, the correlation results of pseudo-experiments performed for each input alone are also

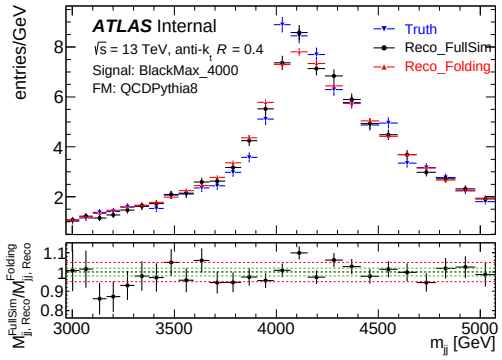
shown. The fluctuations of the folding matrix induce negative correlations between adjacent bins (see figure 4.16(c)), due to the constraints that the total probability in each column of the normalized A_{ij} is equal to one. This matrix accounts for migrations of events between the bins, preserving the total number of events. The fluctuations of the truth m_{jj} distribution induce positive correlations between adjacent bins, since each truth m_{jj} bin contributes through the folding to several reconstructed bins (see figure 4.16(d)).



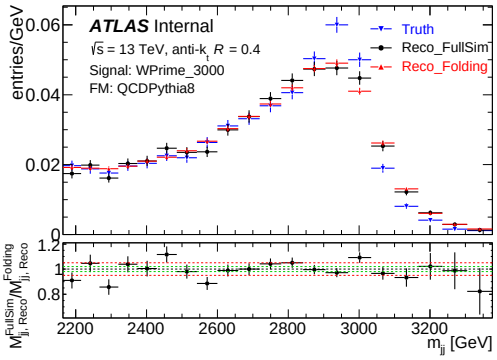
(a) q^*(3 TeV) folded with $\tilde{A}(q^*(3\text{ TeV}))$

(b) q^*(3 TeV) folded with $\tilde{A}$(QCD Pythia 8)

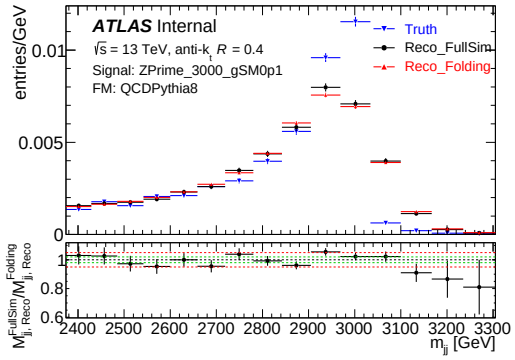
Figure 4.14: q^*(3 TeV) signal model folded with $\tilde{A}$ matrices from various samples (red points). The truth distribution (blue points) and the one obtained through the full simulation (black points) are also indicated. The ratio between the distribution obtained through the full simulation and the folding result is also shown.



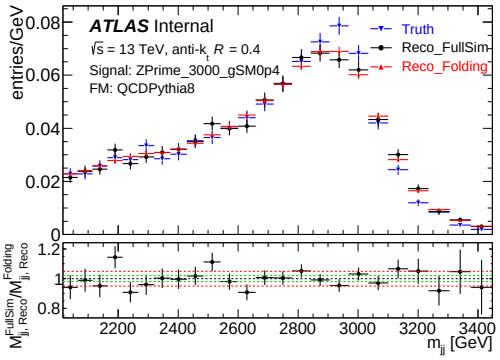
(a) QBH(4TeV) folded with $\tilde{A}$(QCD)



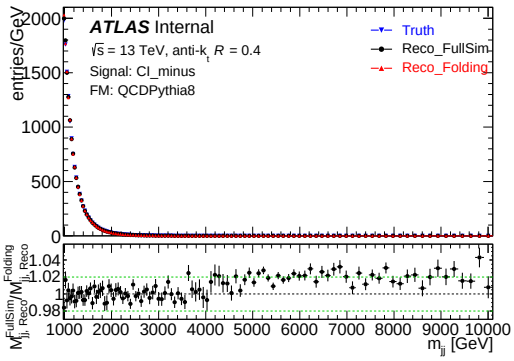
(b) W′(3TeV) folded with $\tilde{A}$(QCD)



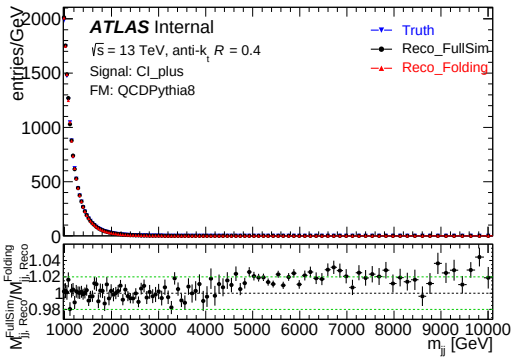
(c) Z′(3TeV, g_q=0.1) folded with $\tilde{A}$(QCD)



(d) Z′(3TeV, g_q=0.1) folded with $\tilde{A}$(QCD)

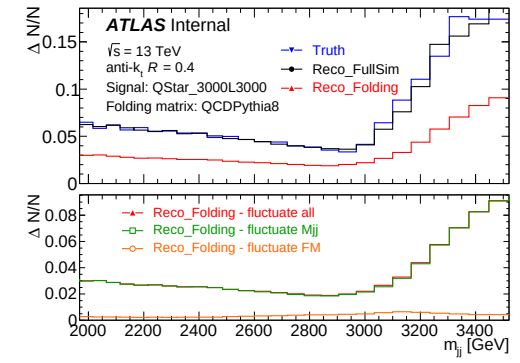


(e) CI (Λ = 7 TeV, η=-1) folded with $\tilde{A}$(QCD)

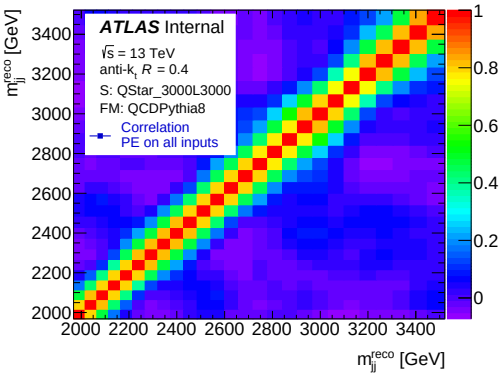


(f) CI (Λ = 7 TeV, η=+1) folded with $\tilde{A}$(QCD)

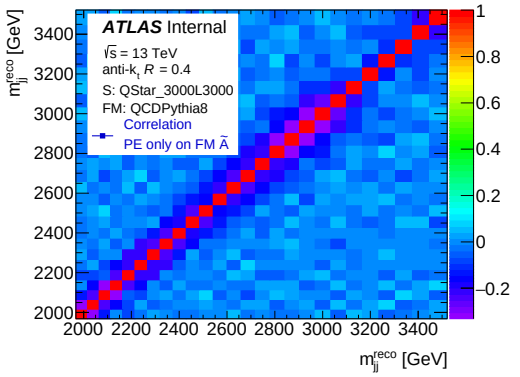
Figure 4.15: Various signal models (1.4) folded with $\tilde{A}$ matrices from QCD sample are shown in red points. The full simulation distribution is shown in black points. The ratio between the distribution obtained through the full simulation and the folding result is also shown.



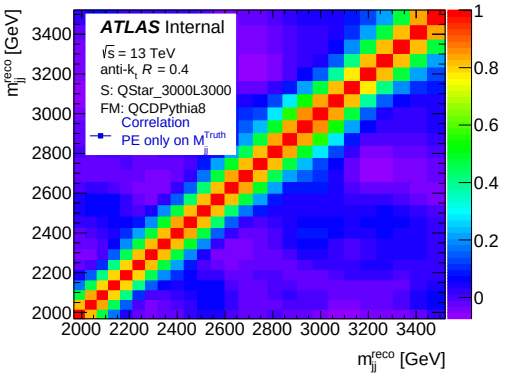
(a) Relative uncertainties



(b) Full correlation matrix



(c) Correlations when fluctuating only the folding matrix



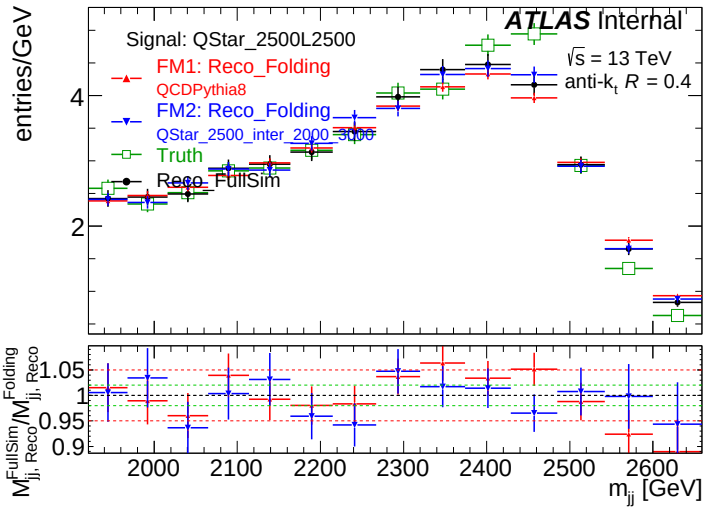
(d) Correlations when fluctuating only the truth m_{jj} distribution

Figure 4.16: (a) Relative uncertainties of the truth distribution, fully simulated reconstructed distribution and reconstructed distribution obtained through the folding procedure. (b, c, d) Correlations of the statistical uncertainties affecting the result of the folding procedure, when fluctuating the various inputs.

4.3.5 Interpolation tests

When deriving limit bands, a scan using different parameter points is used. In the current method, the full simulation is needed for each parameter point, which limits the number of point that can be scanned since these simulations require lots of computing resources. With the folding procedure, a method to overcome this problem becomes available, using an interpolation procedure. If for example two fully simulated samples for resonance masses M_1 and M_2 are available, only the truth distribution at an intermediate mass M' is needed in order to derive the reconstructed m_{jj} distribution at M'. The truth and reconstructed quantities of each of the two full simulation samples are scaled (event by event) from their corresponding mass to M', effectively preserving the relative resolution. Two folding matrices at the intermediate mass M' are produced and then averaged to form the final folding matrix, used to fold the truth distribution of the parameter point M'. Using this technique, smoother limit bands can be derived.

In figure 4.17, I show the result of the folding of a q^* signal of mass 2.5 TeV using an interpolation of two q^* signals at 2 and 3 TeV respectively. This distribution is compared with the result of the folding using the $\tilde{A}$ matrix derived from QCD sample, as well as with the result from the full simulation. The various distributions are compatible within the uncertainties, proving the reliability of the folding procedure for performing interpolations.



(a)

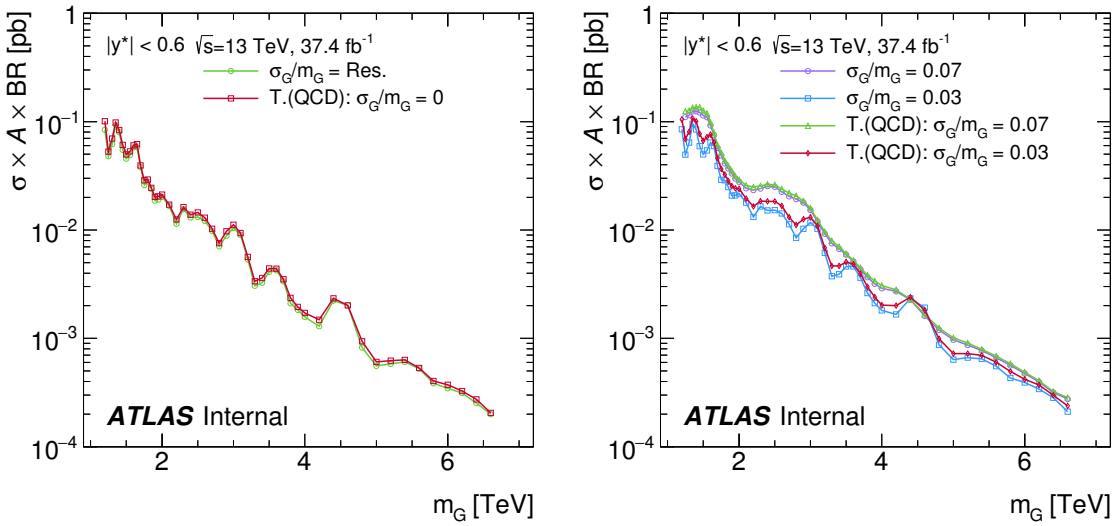
Figure 4.17: Folding of a q^* signal of mass 2.5 TeV using an interpolation of two q^* signals at 2 and 3 TeV respectively is shown in blue points. The result is compared with the folding using the $\tilde{A}$ derived from QCD sample in red points, as well as with the result from the full simulation in black points.

4.3.6 Gaussian limits

The main goal of the folding technique is to be used for setting limits at particle level. Here, I present the tests and comparisons done for that matter.

Figure 4.18 shows the comparison between the limits on Gaussian signals at truth level folded with the detector effects and at reconstructed level respectively. Results are shown for several different ratios between the Gaussian width and mean (σ/M) and are represented as a function of the mass (≡ Gaussian mean) to facilitate and make more precise the use of these limits by theoreticians when comparing their signal models to the Gaussian ones.

In the case of a narrow signal, which corresponds to σ/M=detector resolution at reconstructed level and σ/M=0 at truth level, a small difference between the two methods is observed. This is due to non-Gaussian tails of the resolution which are taken into account by the folding matrix and to an effect related to the binning of the folding matrix. In the case of σ/M=constant, a direct quantitative comparison between the two methods is not possible, since the folded signal includes the resolution effect that is mass dependent. However, for resonance widths comparable to the resolution ($\sigma/\mathrm{M} \approx 0.03$), the difference observed between the results of the two approaches emphasizes the relevance of the folding method. Due to the dependence of the resolution on mass, the folded signal at reconstructed level is also not symmetric and this has an impact on the limits. For large widths, the effect of the detector resolution on the global width is smaller, and the difference between the two methods is reduced.

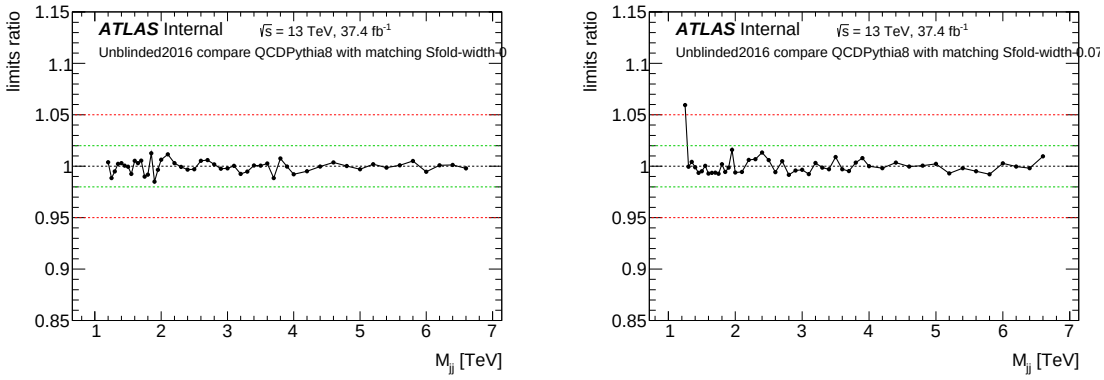


(a) (b)

Figure 4.18: Comparison between the limits on Gaussian signals at truth level folded with the detector effects (red points in a), red and green points in b)) and at reconstructed level respectively. Results are shown for several different ratios between the Gaussian width and mean.

Figure 4.19 shows the ratio of the limits on Gaussian signals at truth level folded with the detector

effects using $\tilde{A}$ matrices from QCD Pythia8 samples with and without the matching criteria applied. For the two signal widths shown, the ratio is close to unity, within small fluctuations. Based on this result, I do not add a systematic uncertainty to account for the impact of the geometrical matching criteria.



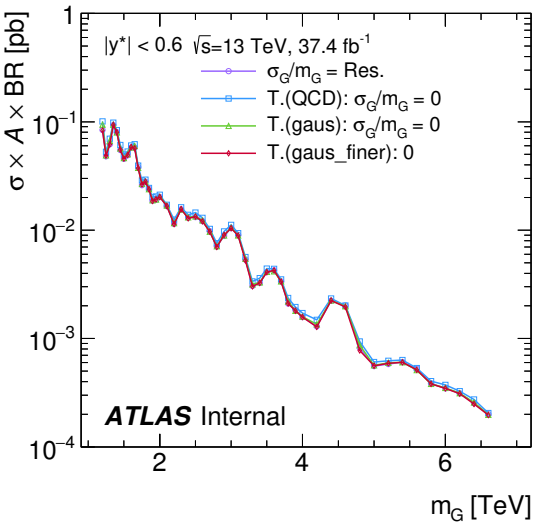
(a) $\sigma/M = 0$ (b) $\sigma/M = 0.07$

Figure 4.19: Ratio of the limits on Gaussian signals at truth level folded with the detector effects using $\tilde{A}$ matrices from QCD Pythia8 samples with the matching criteria applied divided by the one without matching. Results are shown for two different resonance widths. The green dotted lines correspond to a 2% deviation from 1, and the red ones to 5%.

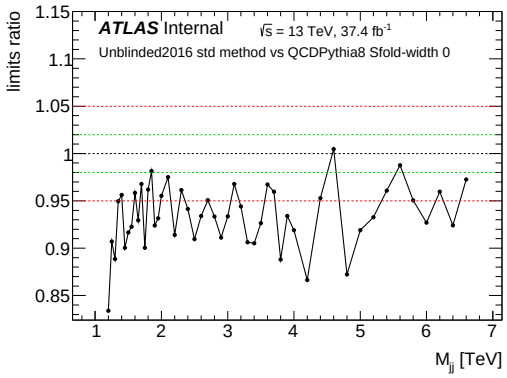
To further study the effects of the non-Gaussian tails of the resolution and of the binning of the truth axis of the folding matrix on the limit calculations, I compare in figure 4.20 the limits on various signals with $\sigma/M = 0$ at truth level folded with the detector effects, and a signal with σ/M =detector resolution at reconstructed level. For the folding method, three folding matrices are considered: one matrix evaluated from a QCD Pythia8 simulation and two matrices using an approximation of a Gaussian resolution effect using "standard" and finer binning respectively for the truth axis. In figure 4.20(b), one can see a systematic shift (more than 5% in average) between the limits obtained for a signal at reconstructed level and one at truth level folded using the QCD Pythia8 sample. In figure 4.20(c), where the folding matrix is an approximation of a Gaussian resolution with "standard" binning (i.e. the same binning for the truth and reconstructed axis), the ratio of the limit to the one obtained for signal at reconstructed level is close to unity, but with a small systematic deviation. This shows that the non-Gaussian resolution tails, which are taken into account by the folding method, have a significant effect on limit calculations.

The motivation to use finer binning for the truth axis in the folding matrix is to have a better description of the mass dependent resolution and of the migrations from the truth signal to the various reconstructed bin (which depend on the position of the signal within the large bins). For a folding matrix with Gaussian resolution and 5 times finer binning for the truth axis (see figure 4.20(d)), the ratio of the limits has smaller fluctuations and is closer to unity compared to the one obtained using a folding matrix with standard binning (see figure 4.20(c)). The fine binning (5 times finer than the "standard" one) is chosen as the default binning for the truth axis of the folding matrix,

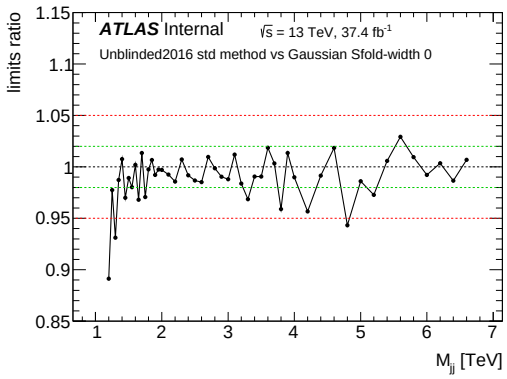
in the evaluation of the main results of this study.



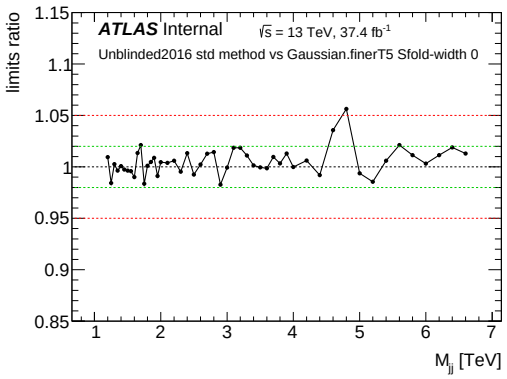
(a) Limits obtained using signals reconstructed with various methods



(b) Ratio between the limit obtained with "σ_G/m_G = Res." divided by the ones for "T.(QCD) : $\sigma_G/m_G = 0$"



(c) Ratio between the limit obtained with "σ_G/m_G = Res." divided by the ones for "T.(gaus) : $\sigma_G/m_G = 0$"

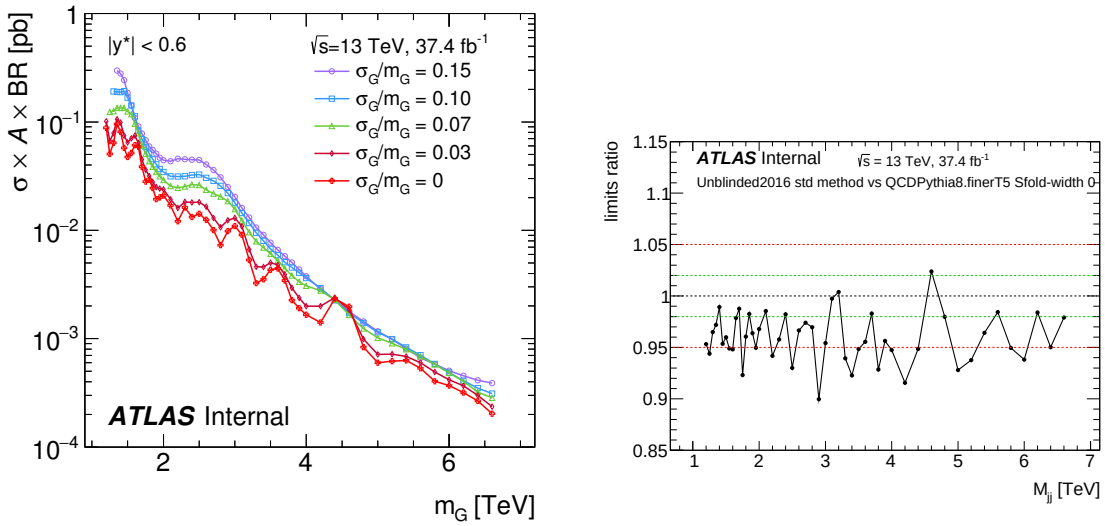


(d) Ratio between the limit obtained with "σ_G/m_G = Res." divided by the ones for "T.(gaus_finer) : $\sigma_G/m_G = 0$"

Figure 4.20: Comparison between the limits on signals with a width at reconstructed level equal to the resolution (σ_G/m_G = Res.) and on narrow signals at truth level folded with the detector effects respectively. Three folding matrices are used, one evaluated from a QCD Pythia8 simulation (T.(QCD) : $\sigma_G/m_G = 0$, blue points) and the other two using an approximation of a Gaussian resolution effect using standard binning (T.(gaus) : $\sigma_G/m_G = 0$, green points) and finer binning (T.(gaus_finer) : $\sigma_G/m_G = 0$, red points) for the truth axis. The green dotted lines correspond to a 2% deviation from 1, and the red ones to 5%.

Figure 4.21(a) shows the limits for various widths of Gaussian signals at truth level folded with an $\tilde{A}$ matrix from QCD Pythia8 samples, having 5 times finer binning for the truth axis compared to the reconstructed one. Figure 4.21(b) shows that the effect of the non-Gaussian resolution tails for

the result obtained using the folding method is of about 3 – 4%, compared to the limits on Gaussian signals at reconstructed level. This constitute an additional point in favor of the folding procedure.

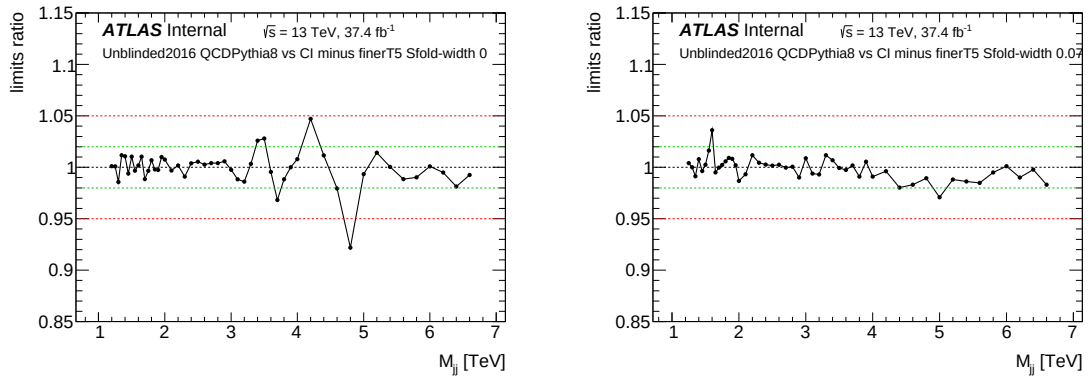


(a) (b)

Figure 4.21: (a) Limits of various widths Gaussian signals at truth level folded with $\tilde{A}$ matrix form QCD Pythia8 samples, using 5 times finer binning for the truth axis compared to the reconstructed one. (b) Ratio of the limit for a signal with a width equal to the resolution at reconstructed level, divided by the limit with a narrow signal at truth level folded using a matrix with fine bins on the truth axis. The green dotted lines correspond to a 2% deviation from 1, and the red ones to 5%.

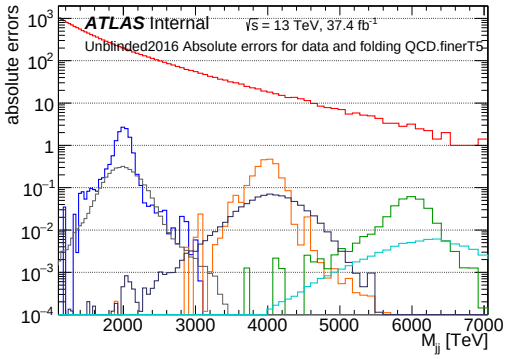
To evaluate the effect of the sample choice for calculating the folding matrix, limits were calculated using an $\tilde{A}$ matrix from CI signal samples (with $\Lambda = 7$ TeV). In figure 4.22, a small change, in the limit of about 1% with respect to the result obtained using the QCD Pythia8 transfer matrix, can be seen for $m_{jj} > 4$ TeV where the CI signal becomes non negligible. Based on this small value of the change, no additional uncertainty is added from the choice of the signal sample for building the $\tilde{A}$ matrix.

I have also tested that the statistical uncertainties induced by the folding procedure (caused by the available MC statistics used to build the $\tilde{A}$ matrix) discussed into some detail in Section 4.3.4, have very little impact on these Gaussian limits studies. Figure 4.23 shows a comparison between the absolute statistical uncertainties of data and of several folded Gaussian signals. The normalization of each Gaussian signal used for this plot corresponds to the value of the limit on the number of signal events. This comparison indicates that the statistical uncertainties induced by the folding method on the signal is negligible and can hence be safely neglected in the limits setting procedure.



(a) $\sigma/M = 0$ (b) $\sigma/M = 0.07$

Figure 4.22: Ratios of limits for signals folded with $\tilde{A}$ matrices from CI samples divided by the ones obtained with $\tilde{A}$ matrices from QCD Pythia8 samples. Results are shown for two different resonance widths. The green dotted lines correspond to a 2% deviation from 1, and the red ones to 5%.



(a)

Figure 4.23: Statistical uncertainties of data (smoothly falling red graph), and for signals folded with $\tilde{A}$ matrices from QCD Pythia8 sample, with finer binning for the truth axis. The signals at truth level are Gaussians having three different means (2, 4 or 6 TeV) and two different widths ($\sigma/M = 0$ or 0.1) respectively. The normalization of each Gaussian signal corresponds to the limit on the number of signal events.

4.4 Results

The results of this analysis, shown hereafter, have been published in the paper [76].

4.4.1 Search results

The search phase result for the resonance analysis using the BumpHunter algorithm (see 4.2.4 for its definition) and 37.0 fb^{-1} of combined 2016 and 2015 data is presented in figure 4.24. The BumpHunter p-value of the most discrepant region is 0.63. As this discrepancy is not significant, no region is excluded in the background fit.

Figure 4.25 shows the p-value of the most significant bump for each of the pseudo-experiments of the background model as a function of its m_{jj} region. The observed BumpHunter test statistic is then compared to the results from pseudo-experiments and a global p-value of the most discrepant region is calculated by integration over the test statistic distribution.

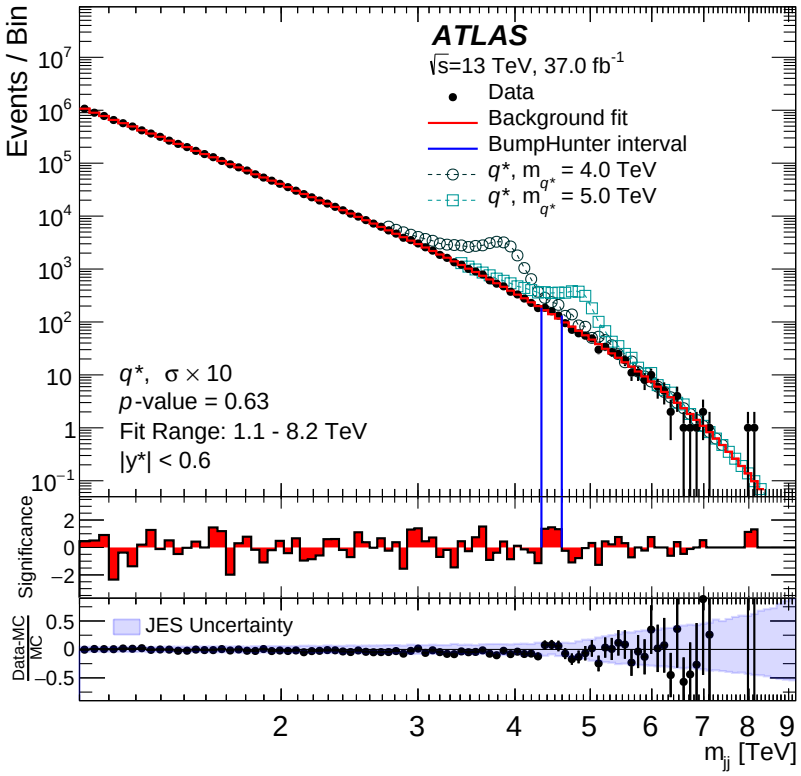
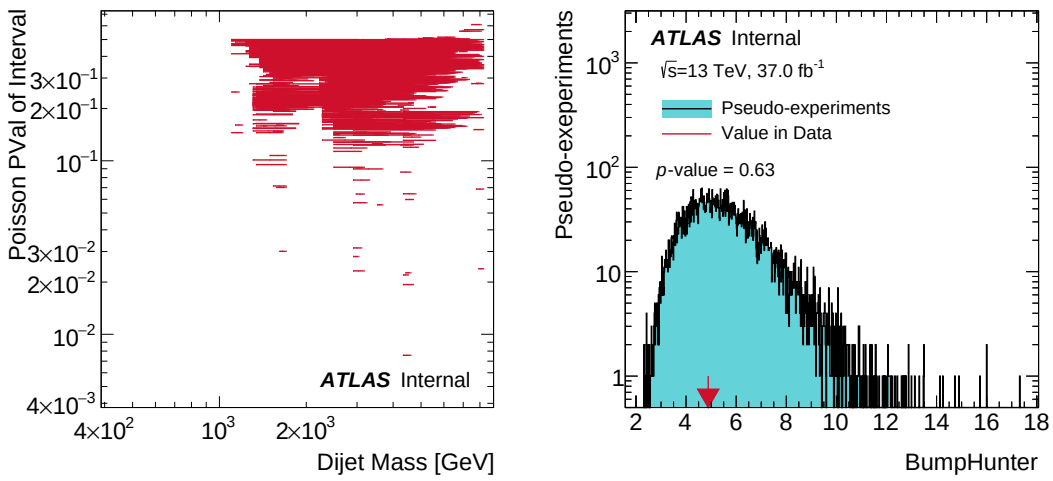


Figure 4.24: The reconstructed dijet mass distribution m_{jj} (filled points) is shown in black points. The solid red line depicts the background prediction from the sliding-window fit. Predictions for benchmark q* signals are normalized to a cross-section large enough to make the shapes distinguishable above the data. The vertical lines indicate the most discrepant interval identified by the BumpHunter algorithm, having a p-value of 0.63. The middle panel shows the bin-by-bin significances of the data–fit differences, considering only statistical uncertainties. The lower panel shows the relative differences between the data and the prediction of Pythia 8 simulation of QCD processes, corrected for NLO and electroweak effects, and is shown purely for comparison. The shaded band denotes the experimental uncertainty in the jet energy scale calibration. Figure from [76].

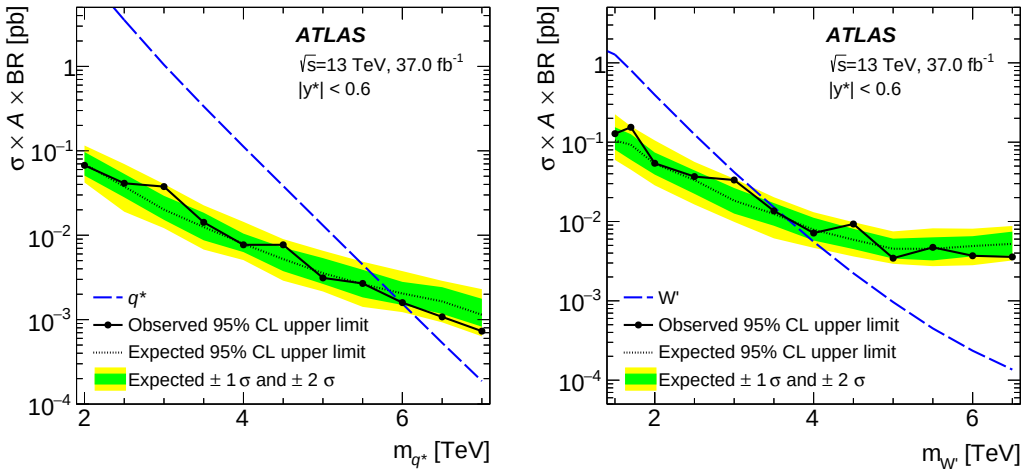


(a) (b)

Figure 4.25: a) p-value of the most significant bump for each of the pseudo-experiments as a function of its m_{jj} region. b) Comparison of the observed BumpHunter test statistic to the results from pseudo-experiments.

4.4.2 Limits setting results

Figure 4.26 shows the 95% CL upper limits on $\sigma \times \mathcal{A} \times$ BR for the models q* and W′, with observed limits of 6 and 3.6 TeV and expected limits of 5.8 and 3.7 TeV respectively.



(a) q* (b) W′

Figure 4.26: The 95% CL upper limits obtained from the dijet invariant mass (m_{jj}) distribution on cross-section times acceptance times branching ratio to two jets, $\sigma \times A \times$ BR, for the models q* and W′. Figures from [76].

Figure 4.27 shows the 95% CL upper limits on $\sigma \times \mathcal{A} \times \mathrm{BR}$ for a generic Gaussian signal with a mean m_G and width σ_G at particle level. Observed limits are shown for five different widths, from a narrow width (0%) to 15% of m_G. The expected limits and the corresponding $\pm 1\sigma$ and $\pm 2\sigma$ bands are also indicated for a narrow-width resonance as described in section 4.2.4.

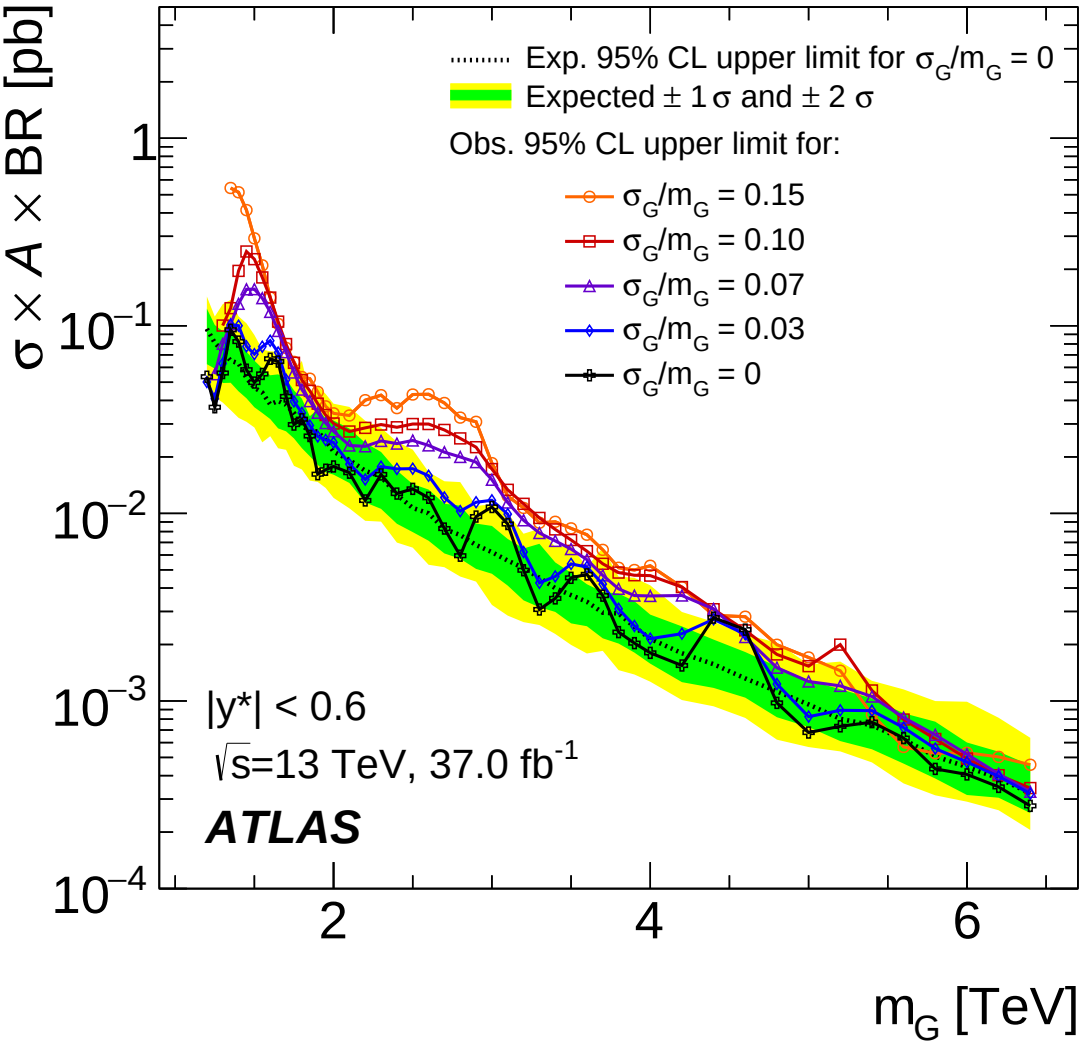


Figure 4.27: The 95% CL upper limits obtained from the dijet invariant mass m_{jj} distribution on cross-section times acceptance times branching ratio to two jets, $\sigma \times A \times \mathrm{BR}$, for a hypothetical signal with a cross-section σ that produces a Gaussian contribution to the particle-level m_{jj} distribution, as a function of the mean of the Gaussian mass distribution m_G. Observed limits are obtained for five different widths, from a narrow width to 15% of m_G. The expected limits and the corresponding $\pm 1\sigma$ and $\pm 2\sigma$ bands are also indicated for a narrow-width resonance. Figure from [76].

5 Precision measurement: leading jet cross-section

In this chapter, I describe a new jet cross-section measurement that is done for the first time, the leading jet double differential cross-section, with details on both the experimental measurement and the theoretical predictions aspects, where I contributed in both.
I first present the motivation of this new analysis. Then, I describe the data selection, followed by the transfer matrix and the unfolding methods used to factorize the detector effects from the measurement. Next, the systematic uncertainties evaluation is presented. The theoretical predictions calculation, challenges and improvements are detailed, in addition to the evaluation of the statistical and systematic uncertainties. Last, the comparison between the data measurement and the theoretical predictions is shown.

5.1 Motivation

In the previous chapter, I presented a direct search approach for new physics, where we look for an additional bump on top of a smooth spectrum produced by the new resonance near its mass. On the other hand, an indirect search approach is also used to search for new physics. In this approach, an observable is measured in data and compared to a "precise" SM prediction where we search for smooth deviations between the two. Many observables are used to capture the various kinematics of the event, for example the differential cross-sections of inclusive jets (as a function of p_{T} and $|\mathrm{y}|$ in ATLAS [89–91] and CMS [92–94] analyses) and dijets (as a function of dijet invariant mass $\mathrm{m_{jj}}$ and $\mathrm{y^*}$ in ATLAS analyses [91, 95], $\mathrm{m_{jj}}$ and $|\mathrm{y}|^{\mathrm{max}}$ or $p_{\mathrm{T}}^{\mathrm{avg}}$, $\mathrm{y_B}$ and $\mathrm{y^*}$ in CMS analyses [96, 97]). Other observables include for example the azimuthal separation angle to catch correlations in this variable. Focusing on the inclusive jets observable, tensions are observed when comparing data with theoretical prediction using the full bin range. This is the case for ATLAS analyses using 8 TeV [90] and 13 TeV [91] data where the χ^2/dof for the latter are shown in table 5.1. When the comparison is done in individual rapidity bins, the agreement between data and prediction is generally good. This points to a possible source of the tension originating from the correlation of the uncertainties between the different rapidity bins. Moreover, the effect of changing the correlations is found to be significant. For example, using uncorrelated statistical uncertainties of the eta-intercalibration for each of its calibration component (each (p_{T},η) bin) reduces the χ^2 by more than 200 units. For CMS analyses, the tensions are also seen for the initial measurement using 7 TeV [98] data, which led the collaboration to modify the correlation model to improve the agreement.

Theorists raised a concern about the inclusive jet observable that it does not conserve the physics correlations between the jets of the event. Events with different number of jets are filled in the same

Table 5.1: Summary of χ^2/dof values obtained from a global fit using all p_T and rapidity bins, comparing the inclusive jet cross-section and the NLO pQCD prediction corrected for non-perturbative and electroweak effects for several PDF sets and for the two scale choices. All the corresponding p-values are $\ll 10^{-3}$. Table from [91].

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

spectrum; the information on the number of jets in an event is lost. In addition, jets in a same bin can come from leading jets, sub-leading jets and so on. During the "ATLAS Standard Model" workshop in 2017 (Thesaloniki), Alexander Huss, a theorist, proposed studying the leading jet cross-section arguing that this observable is more robust from the theoretical point of view. No physics correlations are missed using the leading jet observable, in addition to being the highest p_T in the event and hence reaching the highest energies where we hope to detect new physics. Moreover, as I will discuss later, there is less ambiguity in the theoretical prediction when defining the renormalization and factorization scales for the event.

In this chapter, I present the development of this new analysis for which I did both the data measurement and prepared the theoretical predictions. The observable is defined double-differentially as:

$$\frac{d^2\sigma(\text{Leading jet})}{dp_T dy} = \frac{N^{\text{particle level}}_{\text{Leading jets}}}{\mathcal{L}\Delta p_T \Delta y}, \tag{5.1}$$

where $\mathcal{L}$ denotes the luminosity, Δp_T and Δy the bins widths in transverse momentum and in rapidity. As I detail later, both the data measurement and the theoretical predictions contain additional challenges compared to the inclusive jets case. I first introduce the data measurement steps, then the theoretical prediction ones and at the end the comparison between the two results.

5.2 Data selection and quality

The full Run II data, passing the data quality selection, is used with a total integrated luminosity of 139 fb^{-1}. The nominal MC generator used is Pythia with two statistically independent productions, to match the pile-up profile changes for different data taking periods.

The nominal jet reconstruction collection used is EMPFlow, which has a better resolution at low p_T, and cross checked with EMTopo. Jets are required to be in the region $|y| < 3$, and the final measurement is given for p_T values above 100 GeV.

The measurement is binned in p_T and rapidity. For the rapidity, the following binning is used:
|y| = [0, 0.5, 1, 1.5, 2, 2.5, 3].

Pile-up jets rejection using the JVT tool (refer to chapter 2.3.2) is not used here since the measurement is well above the p_T range where the pile-up activity is significant ($p_T < 60$GeV). On the other hand for the MC simulation, high-p_T pile-up jets, of few hundreds of GeV, are observed which do not match what is observed in data. A cleaning procedure is added for MC simulation events to reject such jets. Comparing the truth and reco jets, an MC event is considered to include badly simulated high-p_T pile-up jets and thus rejected if, before applying any measurement fiducial cuts:

1. the number of truth jets is zero,
2. the average p_T of the two leading jets at reco level is more than 1.4 times higher than the leading truth jet p_T:

$$p_T^{avg}(\text{reco}) \equiv 0.5 \times \left(p_T^{j\,1}(\text{reco}) + p_T^{j\,2}(\text{reco})\right) > 1.4 \times p_T^{j\,1}(\text{truth}).$$

Let me recall that the jet resolution, at $p_T = 100$ GeV and above, is about 10% and goes down with increasing p_T. This means that for a good jet to fail the latter requirement, the p_T of the first two leading jets should fluctuate upward by more than $4 \times \sigma$ each, which have an extremely small probability.

Jet cleaning and time cuts are used to veto background jets, and the optimization of those cuts are detailed in next sections.

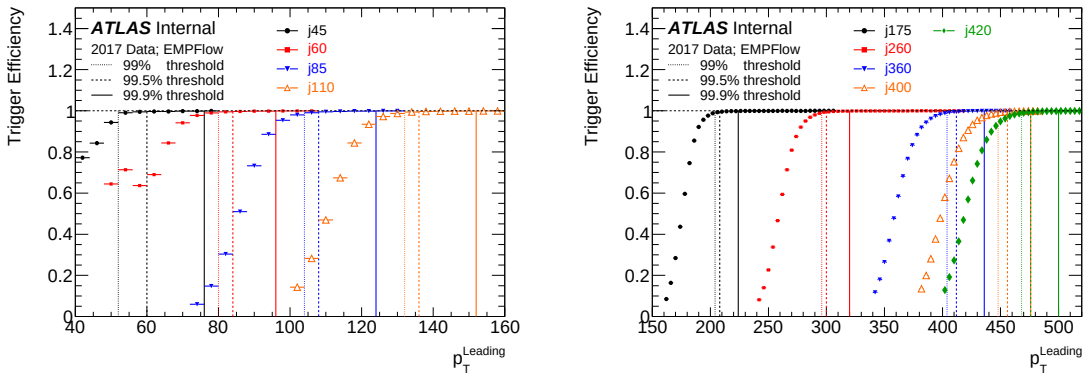
5.2.1 Triggers

For data, we have to select the triggers to be used for each p_T interval. The minimum trigger efficiency requirement for this analysis is 99.5%, tighter than the 99% value for eta-intercalibration since the measurement here depends on the luminosity. In fact, the triggers are used here with an efficiency even higher than 99.5% for most bins. Using the emulation method described in chapter 3.5, the trigger efficiencies are calculated as a function of the leading jet p_T. Since for this measurement only jets with $|y| < 3$ are considered, only central single jet triggers are used. Let me recall that the triggers have two levels: a L1 level which is only sensitive to one bunch-crossing, and a HLT level which integrates on multiple bunch-crossings (refer to chapter 2.2.5). The emulation technique uses

a reference trigger which fires at lower p_{T} compared to the probe trigger. In addition to the difference between HLT levels, the reference and probe triggers don't have the same L1 level in general. What I want to measure here is the efficiency of the HLT level, the L1 level efficiency being already checked by the "JETM" performance group. Hence, the equation 3.17 is changed to the more appropriate form:

$$\mathrm{Eff}(p_{\mathrm{T}}^{\mathrm{leading}}) = \frac{\mathrm{N}^{\mathrm{events}}(\text{fire reference trigger \& pass L1\&HLT probe emulation})}{\mathrm{N}^{\mathrm{events}}(\text{fire reference trigger \& pass L1 probe emulation})}. \tag{5.2}$$

Using this equation, I get the efficiency distributions shown in figure 5.1 for 2017 data and EMPFlow jets where the 99%, 99.5% and 99.9% efficiency tun-on thresholds are indicated.



(a) (b)

Figure 5.1: Trigger efficiency curves for the different central single jet triggers for the 2017 data and EMPFlow jets. 99%, 99.5% and 99.9% efficiency tun-on thresholds are shown as dotted, dashed and full lines respectively. The efficiency is calculated using equation 5.2.

The 99.5% efficiency turn-on p_{T} are summarized in table 5.2 for EMTopo and EMPFlow, 2015+2016 and 2017+2018 data periods. The difference between 2015+2016 and 2017+2018 efficiency turn-on values is mostly due to the improvement of the HLT jet calibration beginning 2017+2018 data taking period making it closer to the offline calibration and leading to lower turn-on p_{T}. On the other hand, EMTopo jets have lower turn-on p_{T} compared to EMPFlow; this is due to the fact that HLT jets are reconstructed from topo-clusters and hence the deviations between the HLT- and EMTopo jets are smaller compared to the differences between HLT- and EMPFlow jets.

An additional verification of the good efficiency is to plot the p_{T} distribution in narrow bins using the selected triggers for each p_{T} range, fit the spectrum by a smooth function, and check that at the low edge of each p_{T} range (triggers transition region) no systematic decrease in the number of entries exists when comparing the fine bin entries with the smooth fit. Figure 5.2 shows the result of this test, where zoomed plots are shown for parts of the ratio between the entries in bins and the smooth function; no systematic drops are visible.

In previous inclusive jet cross-section measurements, the jets in an event were treated independently. For each jet, the trigger, corresponding to the p_{T} range that contains the jet p_{T}, is checked if it is fired,

Table 5.2: 99.5% efficiency tun-on p_{T} thresholds of central single jet triggers separately for the data taking period 2015+2016 and 2017+2018, for EMTopo and EMPFlow jets.

Trigger	2015+2016 period EMTopo jets	2015+2016 period EMPFlow jets	2017+2018 period EMTopo jets	2017+2018 period EMPFlow jets
HLT_j45	56	68	56	60
HLT_j60	76	88	72	83
HLT_j85	104	112	100	108
HLT_j110	132	140	124	136
HLT_j175	208	220	196	208
HLT_j260	300	320	284	300
HLT_j360	416	432	396	412
HLT_j380	440	452	-	-
HLT_j420	-	-	460	476

and if so the jet is used to fill histograms. At low and mid p_{T}, where the triggers are heavily prescaled, a multi-jet event can have only one of its jets used to fill the distribution. For this measurement where the orders of jets is important, as is explained later, the event physics correlations should not be removed. Therefore, the triggering is done on an event level. The trigger corresponding to the leading jet p_{T} is checked if it is fired, and if so all the jets in the event are used (here I only use the first three as I explain later). In addition to preserving the physics correlations, this approach increases the statistics of the second jet and above, since the trigger corresponding to the leading jet p_{T} is less prescaled compared to the triggers corresponding to the other jets p_{T}. I did check that the two approaches are compatible within the statistical fluctuations, and that for the third jet, the increase in statistic can reach a factor of four in some bins.

Entries are corrected for trigger prescales by applying a weight factor. Since the prescale for each trigger changes through the data taking period, within a year and between years, the total luminosity recorded by the trigger for the full Run II data is used to define an equivalent factor to be applied to all events passing this trigger. This method, compared to applying different weight factors, increases the effective statistics as was demonstrated in [75].

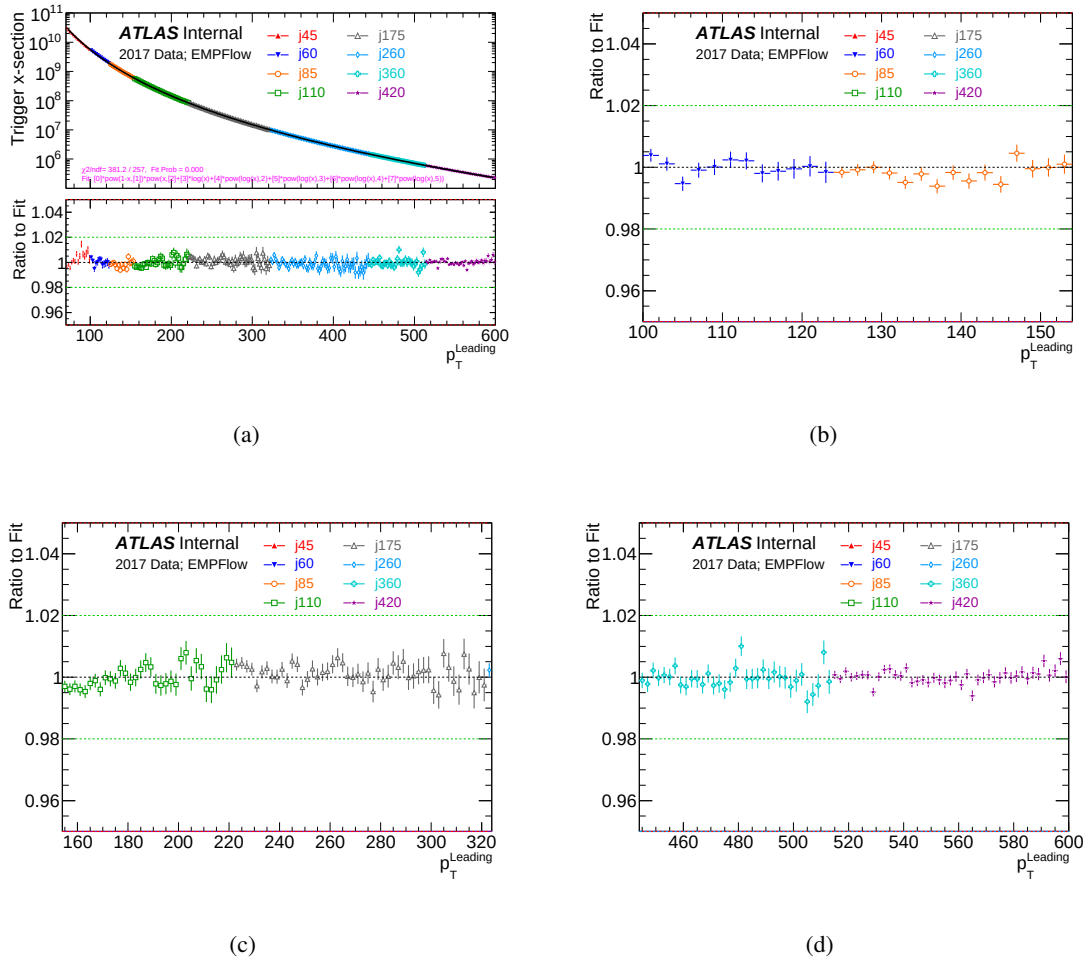
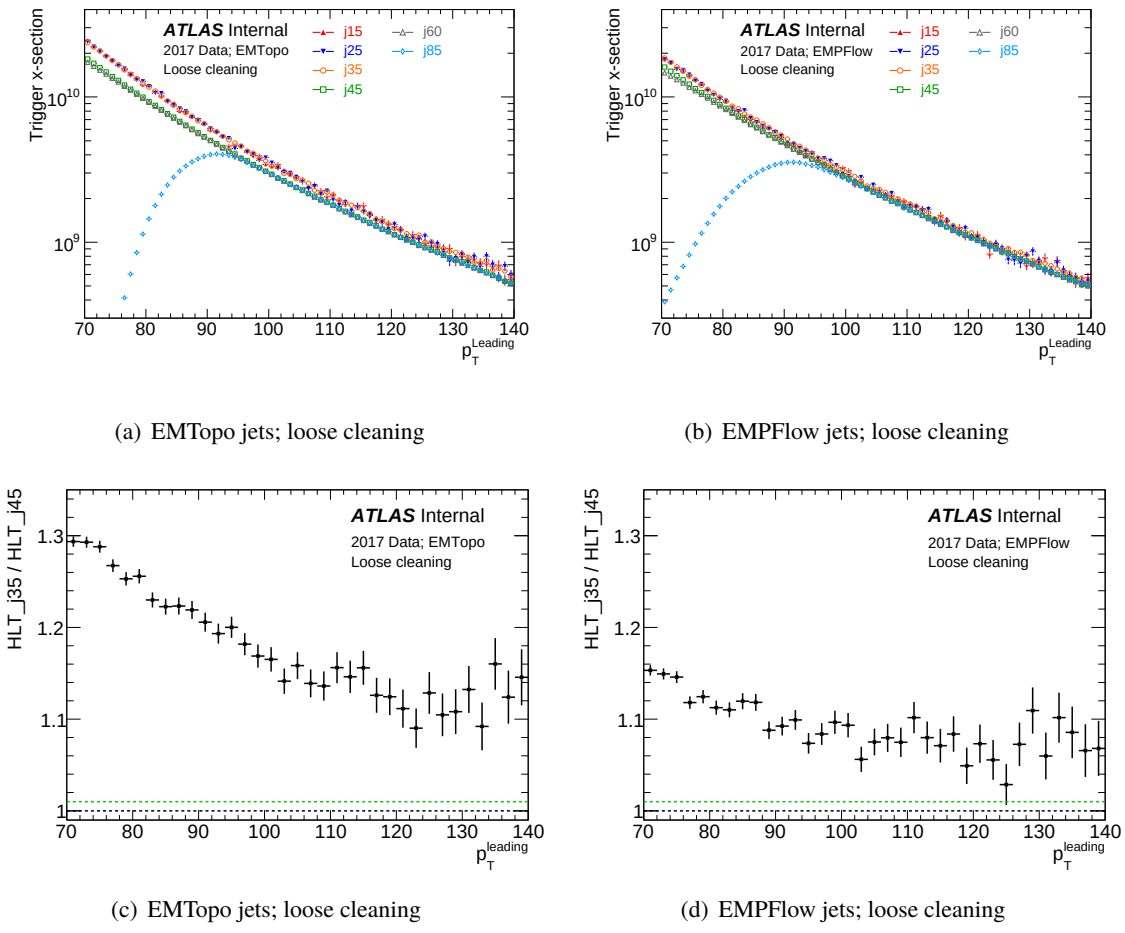


Figure 5.2: p_T distribution in narrow bins using the selected triggers for each p_T range. The distribution is fitted with a smooth function. Ratio plots between the distribution and the fit are shown.

5.2.2 Cleaning criteria and jet time cut

To further check the compatibility between triggers, the leading jet p_T distributions, obtained with the various triggers, are compared. Although the low p_T triggers are not used in this analysis, I also include them in the comparison since they make it more complete, and since they will be used in other jet cross-section measurements. First, a loose cleaning (refer to chapter 2.3.2) is applied to jets. Figure 5.3 shows in the top plots the $p_T^{leading}$ distributions from the different triggers, for EMTopo and EMPFlow jets. The distributions show a significant disagreement between two groups of triggers: the low p_T triggers, HLT_j35 and below, and the other triggers, HLT_j45 and above. Within each group, triggers have compatible distributions, noting that the disagreement for HLT_j85 at low p_T is due to the trigger being below its turn-on p_T threshold and no attention is needed for this part. The only difference between the two groups of trigger, other than the p_T fire thresholds, is that the low p_T

triggers are random at L1 level. A better quantification of the disagreement is shown in the bottom plots of figure 5.3, where the ratio of the distributions from HLT_j35 and HLT_j45 are shown. The difference decreases with p_T, and is about 30% (15%) for EMTopo (EMPFlow) jets at 70 GeV. At higher p_T, the difference reaches a plateau significantly higher than unity: 11% for EMTopo and 4% for EMPFlow jets. It is noticeable that EMPFlow jets, which include tracks information in addition to topo-clusters (refer to chapter 2.3.1), perform better than EMTopo.



(a) EMTopo jets; loose cleaning

(b) EMPFlow jets; loose cleaning

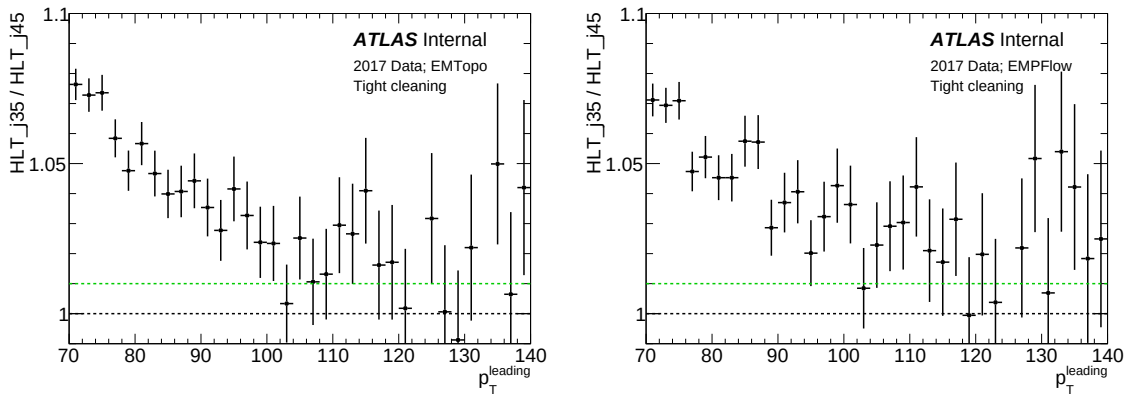
(c) EMTopo jets; loose cleaning

(d) EMPFlow jets; loose cleaning

Figure 5.3: Top plots: $p_T^{leading}$ distributions from six different single jet triggers, for EMTopo (a) and EMPFlow (b) jets. Bottom plots: ratio of the distributions from HLT_j35 and HLT_j45 triggers. The jet cleaning criteria used is BadLoose. The HLT_j35 trigger is random at L1 level, whereas the HLT_j45 trigger is L1 calo-based.

If now I use tight cleaning on jets, which includes an additional cut using tracks (tracks p_T fraction over one layer maximum energy fractions, f_{ch}/f_{max}), a huge improvement is obtained as is visible in figure 5.4. The disagreement at p_T = 70 GeV drops to 6-7%, and at high p_T the plateau is just about 2%.

To better understand the source of the disagreement, let me show some leading jet time distributions. Figure 5.5 shows the leading jet time distributions from the two triggers HLT_j35 and HLT_j45, for



(a) EMTopo jets; tight cleaning

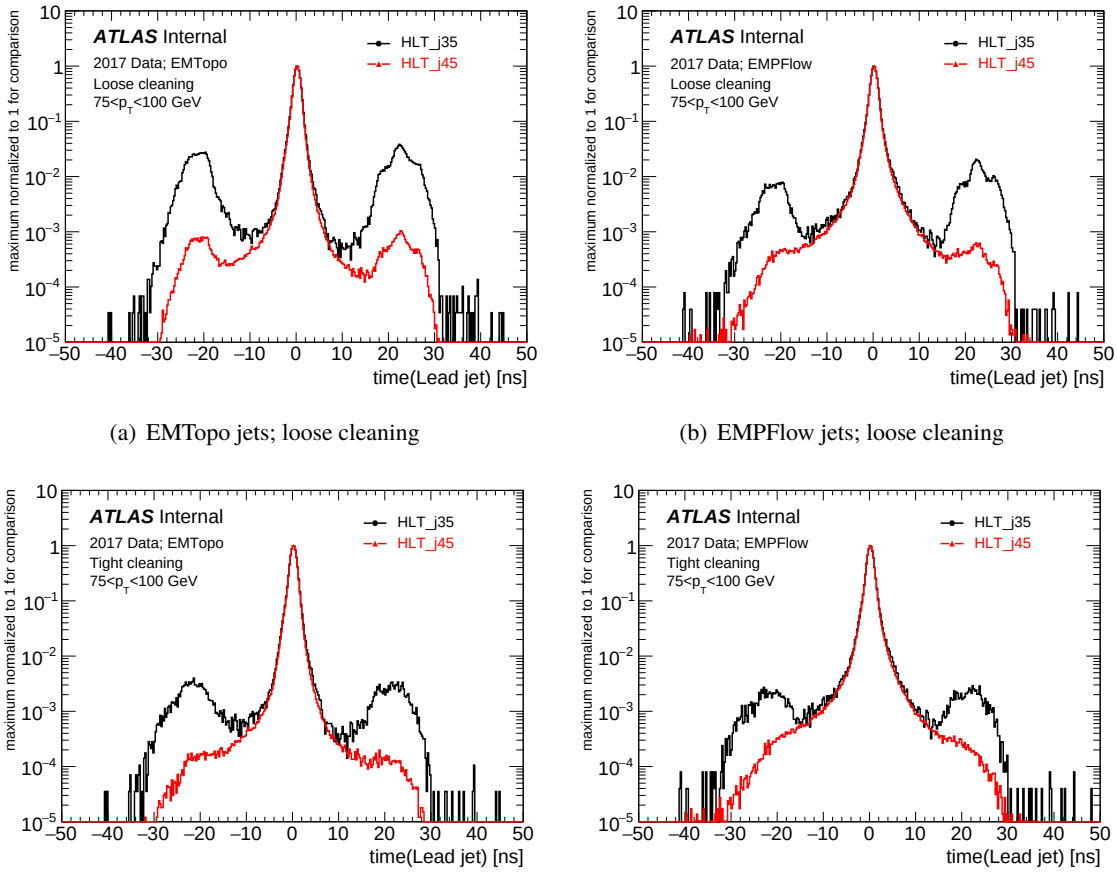
(b) EMPFlow jets; tight cleaning

Figure 5.4: Ratio of the $p_{\mathrm{T}}^{\mathrm{leading}}$ distributions from HLT_j35 and HLT_j45 single jet triggers, for EMTopo (a) and EMPFlow (b) jets. The jet cleaning criteria used is BadTight. The HLT_j35 trigger is random at L1 level, whereas the HLT_j45 trigger is L1 calo-based.

loose (tight) cleaning in top (bottom) plots and for EMTopo (EMPFlow) jets in left (right) plots. The first thing to notice is the additional peaks at ±20-25 ns, comparable with the time interval between bunch-crossings. The jets forming the additional peaks are out-of-time pile-up jets, which are generated from collisions at adjacent bunch-crossings but are also reconstructed in the current event due to the large response time of the LAr calorimeter.
In addition, we notice a reduction of those additional peaks when using EMPFlow instead of EMTopo jets, or when using tight instead of loose cleaning. This is expected since both EMPFlow jets and tight cleaning use information form tracks which have a better time resolution, much smaller than 25 ns, and a tracker with very fine angular widths, and thus no out-of-time tracks are reconstructed in the event.

The other interesting thing to notice in figure 5.5 is that, for all the plots with the different jet reconstructions and cleaning criteria, the leading jet time distributions from HLT_j35 and HLT_j45 triggers agree well in the central peak region (0 ± 10ns) and disagree significantly outside that region specially at the location of the additional peaks. HLT_j35 trigger has more (out-of-time) jets, compatible with what is seen in figures 5.3 and 5.4 where the ratio is higher than unity. As already mentioned, the low p_{T} triggers are random at L1 level. Although the HLT p_{T} thresholds are different between triggers, they don't have an effect if the two triggers being compared are fully efficient in the p_{T} range considered, as is the case here. An important difference between L1 and HLT jets is that for the former only integrates over the current bunch-crossing, whereas the latter integrates over multiple bunches. Hence, the L1 jets are much less sensitive to out-of-time activity, in contrast with HLT jets. When a trigger includes a L1 calo-based level decision, the latter provides an important amount of filtering of out-of-time effects. The low p_{T} triggers, random at L1 level, lack this filtering. If, for example, an event with no high-p_{T} in-time jets is neighbored by an event with high-p_{T} activity, the HLT jets are affected by this neighboring activity and sometimes pass the HLT p_{T} threshold. Hence, a trigger random at L1 level is fired in this case. On the other hand, for the same example, if the trigger requires a L1 level decision, the L1 jets most of the time don't pass the required threshold

and therefore prevent the trigger from firing due to neighboring energetic activity. The tight cleaning helps rejecting some of the out-of-time activity since the out-of-time tracks are not present in the event. The rejection is not perfect since out-of-time deposits in the calorimeter and in-time tracks can be reconstructed in a same jet, and the requirements on track activity is passed.



(a) EMTopo jets; loose cleaning

(b) EMPFlow jets; loose cleaning

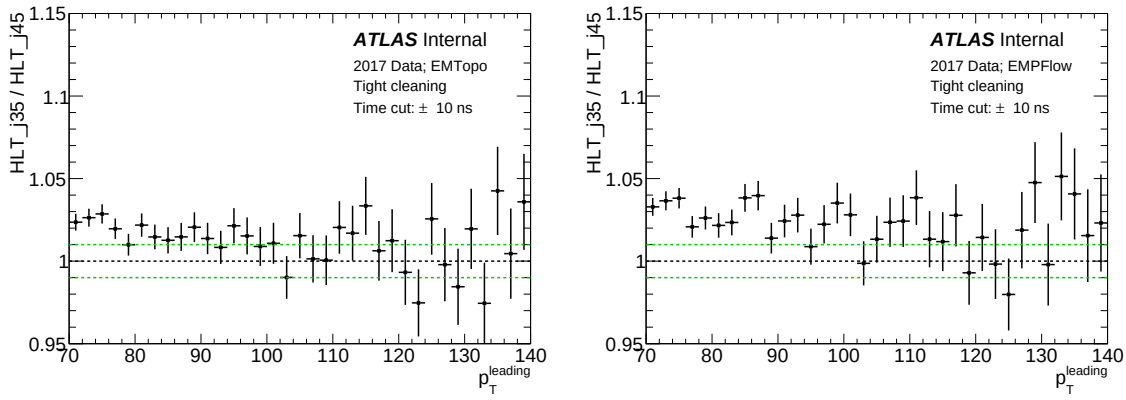
(c) EMTopo jets; tight cleaning

(d) EMPFlow jets; tight cleaning

Figure 5.5: Leading jet time distributions from the two triggers HLT_j35 and HLT_j45, for loose (tight) cleaning in top (bottom) plots and for EMTopo (EMPFlow) jets in left (right) plots. The distributions are for jets with $75 < p_{\mathrm{T}} < 100$ GeV where both triggers are fully efficient. The HLT_j35 trigger is random at L1 level, whereas the HLT_j45 trigger is L1 calo-based.

Applying now a time cut of ± 10 ns in addition to the tight cleaning, the ratio of the $p_{\mathrm{T}}^{\mathrm{leading}}$ distributions from HLT_j35 and HLT_j45 triggers is shown in figure 5.6. The disagreement at $p_{\mathrm{T}} = 70$ GeV drops to 2-3%, and converges to unity at high p_{T}. The time cut improves significantly the agreement between the two triggers, although a residual difference is still present due to the wide jet time resolution, which means out-of-time pile-up jets can sometimes have an absolute time smaller than 10 ns and are not hence removed by the time cut.

Another way to check the disagreement and the improvement is by using the ratio of $p_{\mathrm{T}}^{\mathrm{leading}}$ distributions for loose over tight cleaning criteria from the same trigger. For in-time jets, this ratio is only



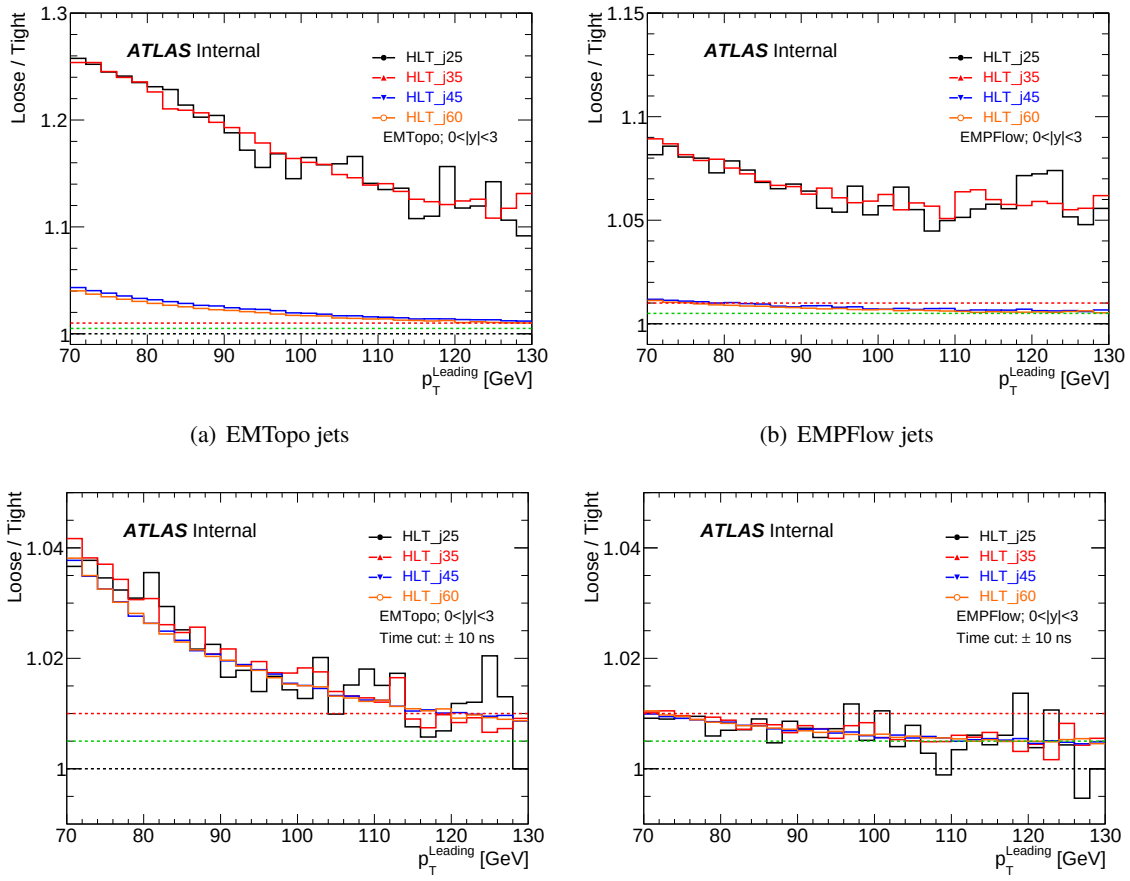
(a) EMTopo jets; tight cleaning; ± 10 ns time cut

(b) EMPFlow jets; tight cleaning; ± 10 ns time cut

Figure 5.6: Ratio of the $p_{\mathrm{T}}^{\mathrm{leading}}$ distributions from HLT_j35 and HLT_j45 single jet triggers, for EMTopo (a) and EMPFlow (b) jets. The jet cleaning criteria used is BadTight. An additional jet time cut of ± 10 ns is applied. The HLT_j35 trigger is random at L1 level, whereas the HLT_j45 trigger is L1 calo-based.

affected by the purity and efficiency of the two cleaning criteria, which are not dependent on the trigger. The only difference between the ratios from the different triggers comes from the level of sensitivity of the triggers to out-of-time activity. Figure 5.7 shows loose over tight cleaning $p_{\mathrm{T}}^{\mathrm{leading}}$ distributions for four triggers, where on the bottom plots an additional time cut of ± 10 ns cut is applied. For both EMTopo and EMPFlow jets, without a time cut, the difference between the ratios is of the order of several percent, whereas when a time cut is applied, the agreement between the ratios is better than a percent level. Again, from these plots, the better performance of EMPFlow jets is visible.

As a consequence of the studies shown in this section, I use a time cut of ± 10 ns for the current measurement. The inefficiency from such a cut is discussed in the systematic uncertainties section 5.5.4. The tight cleaning on the leading jet, which is mainly chosen to get a good purity of jets (rejecting the majority of fake jets), is found also to help rejecting out-of-time jets. The same is true for EMPFlow jets which are mainly chosen for their better resolution at low p_{T}.



(a) EMTopo jets

(b) EMPFlow jets

(c) EMTopo jets; ± 10 ns time cut

(d) EMPFlow jets; ± 10 ns time cut

Figure 5.7: Ratio of $p_T^{leading}$ distributions for loose over tight cleaning criteria from the same trigger. Ratios from four different triggers are shown. The HLT_j25 and HLT_j35 triggers are random at L1 level, whereas the HLT_j45 and HLT_j60 triggers are L1 calo-based.

5.3 Transfer matrix and binning optimization

As with every measurement, one would want to deconvolute the detector effects, resolution and acceptance, from the observable under study. In addition to other reasons, this facilitates the comparison to theoretical model and the combination with other experiments. Using MC simulation of the detector, one is able to relate the truth and reconstructed observables, jet p_T for the current measurement. Events at particles level are generated using Pythia MC generator and then the response of the detector is simulated. Using the same fiducial selection to minimize the reliance on simulation extrapolation, truth and reco jets in each event, passing the selection, are then matched together using a geometrical matching, with a maximum distance criteria between the matched truth-reco jets of $\Delta R = 0.3$[1]. In addition, the matching should be bijective: if a reco jet "i" have the truth jet "j" as the

[1] $\Delta R = \sqrt{\Delta\phi^2 + \Delta y^2}$

closest jet to it, the truth jet "j" should also have the reco jet "i" as its closets jet. Otherwise, the reco jet "i" is considered as having no matching. Using the matched jets, I build the Monte Carlo-based transfer matrix that is used later to deconvolute the detector effects. The unmatched jets, due to geometrical or selection criteria, are accounted for by the matching efficiency corrections defined, for each of the reco and truth levels, as:

$$\begin{aligned} \epsilon^{\text{reco}} &= \sum \text{matched reco jets} \;/\; \sum (\text{matched reco jets} + \text{unmatched reco jets}), \\ \epsilon^{\text{truth}} &= \sum \text{matched truth jets} \;/\; \sum (\text{matched truth jets} + \text{unmatched truth jets}). \end{aligned} \tag{5.3}$$

My goal is to measure the leading jet at truth level which, due to the detector resolution, can become a second[2] jet, a third jet and so on at the reconstructed level. The flip of jet orders between truth and reco jets is an non-negligible effect specially between the first two leading jets. For events with very close p_{T} for the two leading jets, the flip probability is close to 50%. The flip effect is visible in figure 5.8 where is plotted, for each p_{T} bin of the leading truth jets on the Y-axis, the probability of the order of the reco jet which is matched to it on the X-axis. The probability that the leading truth jet flips its order at reco level increases from about 9% to 40% between 100 GeV and 4 TeV. The truth leading jet flips are dominated to the second reco jets, much less to third reco jets and even less to fourth reco jets. Hence for the following studies, I only use the first three jet orders

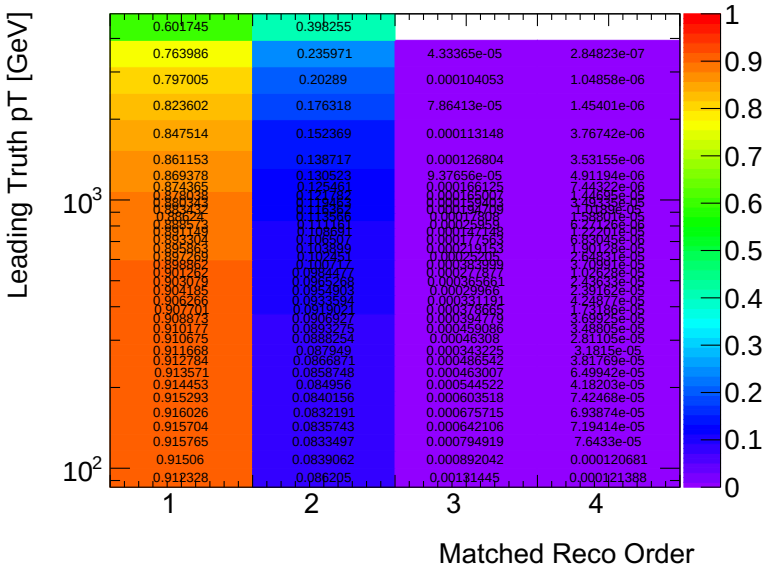


Figure 5.8: The y-axis shows p_{T} of the leading truth jet. the x-axis shows the order of the reco jet matched to it. The z-axis shows the probability of the order at reco level of the leading truth jet. Results using Pythia MC simulation.

Therefore, jet orders and flipping need to be taken into account. Hence, I separate between jet orders when the truth and reco jets are filled in the transfer matrix, which now in addition to including p_{T} bin migrations also includes jet order migration between the truth and reco levels due to the resolution.

[2] Leading (first), sub-leading (second), sub-sub-leading (third) jets correspond to the descending p_{T} order of jets.

Considering one rapidity bin where both the matched truth and reco jets fall inside, the transfer matrix is shown in figure 5.9. The reco level is shown in the x-axis and the truth level on the y-axis. The matrix is split into green boxes corresponding to each combination of one reco jet order and one truth jet order. R1 (T1), R2 (T2) and so on correspond to leading reco (truth) jets and sub-leading reco (truth) jets respectively. Taking for example the boxes in the bottom which correspond to matched leading truth jets (T1), entries from the different reco jet orders (R1, R2 and R3) can be seen. The same is true for all the other truth jet orders. Needless to say, the diagonal boxes, corresponding to no order flip, have higher entries than the non-diagonal ones.

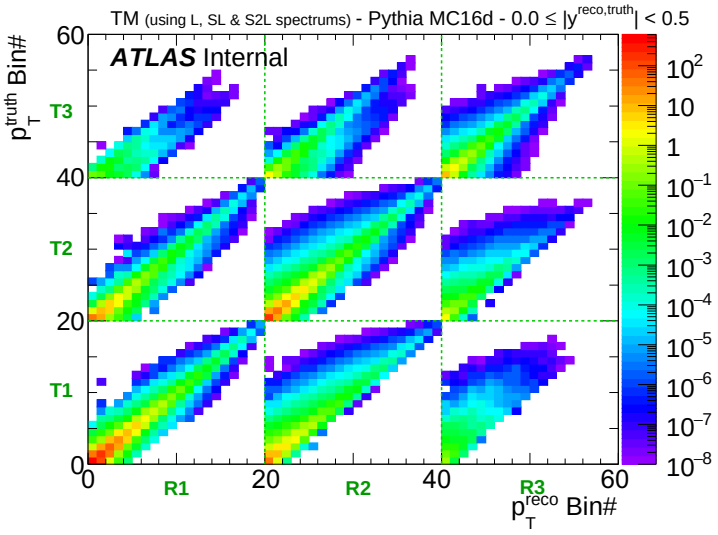


Figure 5.9: Transfer matrix where both jet p_T bin and order migrations between truth and reco levels for matched jets are considered. Reco level is shown in x-axis and truth level on y-axis. Jets are split between green boxes corresponding to one reco jet order and one truth jet order. R1 (T1), R2 (T2) and so on correspond to leading reco (truth) jets and sub-leading reco (truth) jets respectively. Inside each box, p_T bin numbers are shown. Results using Pythia MC simulation.

The probability (≡ folding) matrix "P" is derived from the transfer matrix "T" by normalizing to unity in each truth bin:

$$P_{ij} = \frac{T_{ij}}{\sum_k T_{kj}}, \tag{5.4}$$

where "i" and "j" are the indices of reco and truth bins respectively. Figure 5.10 shows the probability matrix derived from the transfer matrix of figure 5.9.

The geometrical matching requirement of $\Delta R < 0.3$ is comparable to the width of rapidity binning, $\Delta y = 0.5$. Matched truth and reco jets can hence fall in different rapidity bins. For that, in addition to jet p_T and order, I also include in the transfer matrix the migration between the different rapidity bins. Figure 5.11 shows the total transfer matrix that I use, where the new red boxes correspond to each of the truth and reco jet rapidity bins combination (green boxes represent jet order combinations as

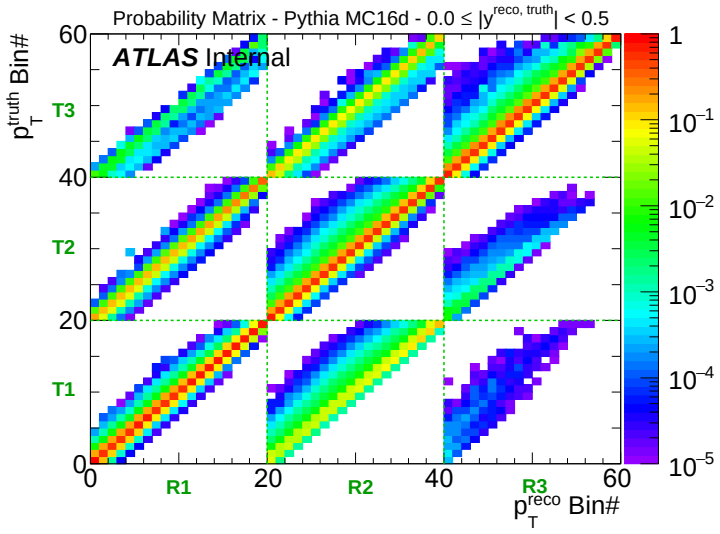


Figure 5.10: Probability matrix normalized to unity in each truth bin. Results using Pythia MC simulation.

before). Migration between only adjacent rapidity bins are visible, reflecting the fact that $\Delta R < \Delta y$.

One would like to optimize the fraction of bin migration taking into consideration the three types of migration included in the transfer matrix: p_T, order and rapidity variables. The transfer matrix migration efficiency and purity describe the fraction of events not undergoing migrations normalized in each truth and reco bin respectively:

$$\begin{aligned} \text{Efficiency}(p_T^{\text{truth, j}}) &= \frac{T_{jj}}{\sum_i T_{ij}}, \\ \text{Purity}(p_T^{\text{reco, i}}) &= \frac{T_{ii}}{\sum_j T_{ij}}, \end{aligned} \tag{5.5}$$

where $T_{ii} \equiv T_{jj}$ are the diagonal bins corresponding to no migration situation in any of the three variables considered. To optimize the binning, I require a minimum efficiency and purity of about 0.4 in each bin for the leading and sub-leading jets (going down to about 0.3 for few bins). If the efficiency and purity are smaller, the bin is enlarged. In addition, for the last bins describing the limits of the phase-space, different p_T bins merging is done for the different rapidity bins giving asymmetric bins toward the limits of the phase-space. This is needed to ensure that the efficiency and purity is acceptable and that there are sufficient entries in those bins, without constraining the p_T bins to be identical in different rapidity bins which have different phase-space limits. Figure 5.12 shows the efficiency and purity for each of the jet order and rapidity bins considered, as a function of p_T.

In addition to the efficiency and purity requirements, some of the bin edges were shifted between 5 GeV and a maximum of 20 GeV to match the trigger turn-on thresholds and thus optimize to the

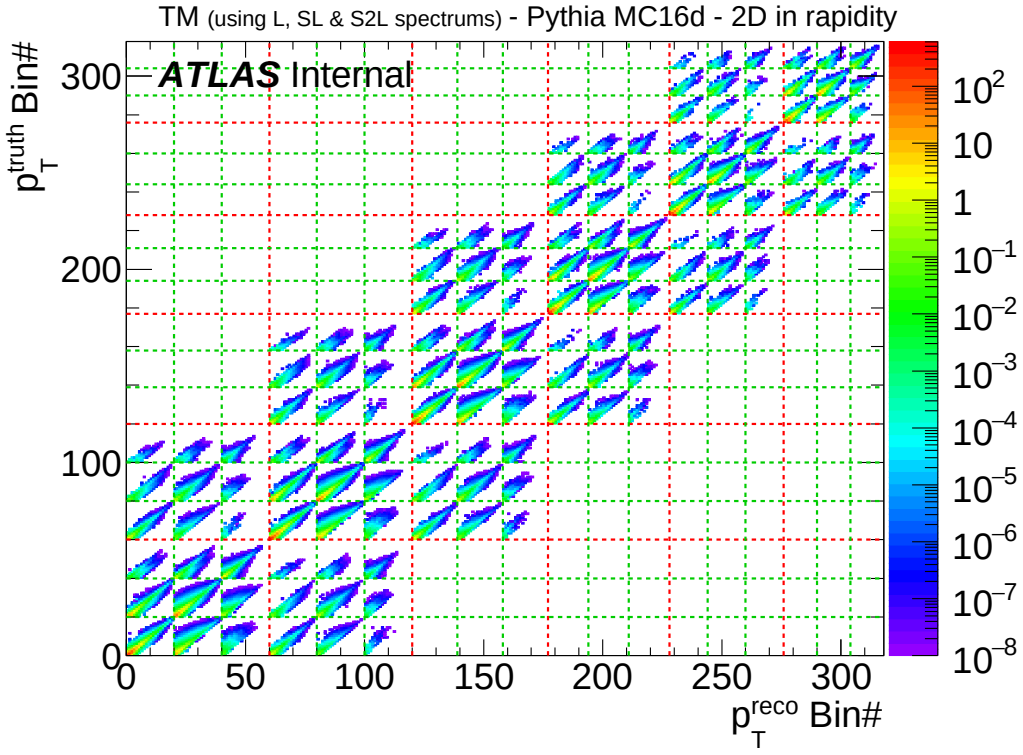
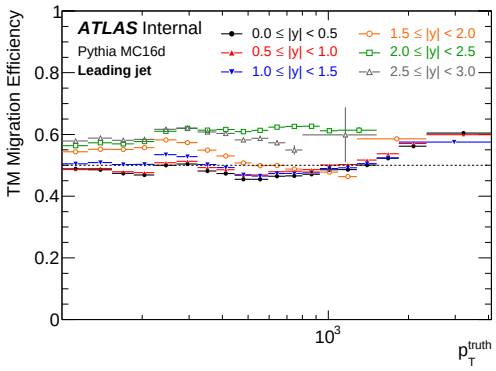


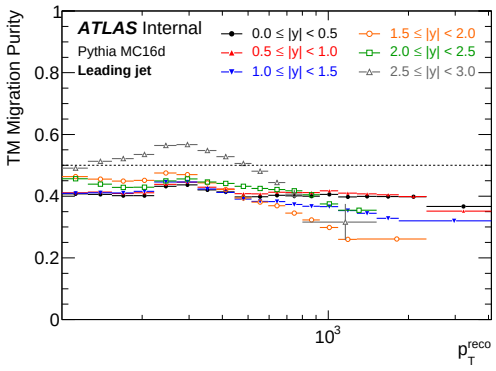
Figure 5.11: Transfer matrix where jet p_T bin, order and rapidity bin migrations between truth and reco levels for matched jets are considered. Reco level is shown in x-axis and truth level on y-axis. Red boxes correspond to truth and reco jet rapidity bins combination. Smaller green boxes correspond to truth and reco jet orders combination. Results using Pythia MC simulation.

maximum statistics possible. Although the final measurement begins at 100 GeV, an additional lower bin should be used to take into account the low p_T migration from and to the additional bin, reducing the reliance on the matching efficiency corrections and thus on simulation. The final asymmetric p_T binning that I obtained is:

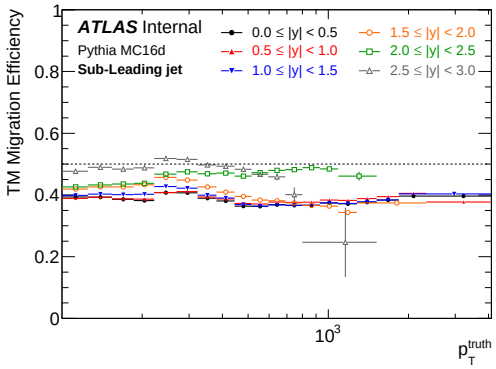
(same for all y bins):	[75,	100,	125,	154,	186,	222,	269,	323,	380,	445,	515,
$0.0 \leq \|y\| < 0.5$:	595,	690,	800,	930,	1090,	1280,	1520,	1840,	2340,	4100]	GeV
$0.5 \leq \|y\| < 1.0$:	595,	690,	800,	930,	1090,	1280,	1520,	1840,	2340,	4100]	GeV
$1.0 \leq \|y\| < 1.5$:	595,	690,	800,	930,	1090,	1280,	1520,	1840,		4100]	GeV
$1.5 \leq \|y\| < 2.0$:	595,	690,	800,	930,	1090,	1280,			2340]	GeV	
$2.0 \leq \|y\| < 2.5$:	595,	690,	800,	930,	1090,		1520]	GeV			
$2.5 \leq \|y\| < 3.0$:	595,	690,	800,				1520]	GeV			



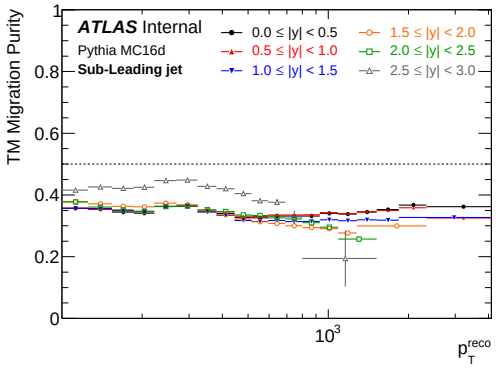
(a) Efficiency - Leading jet



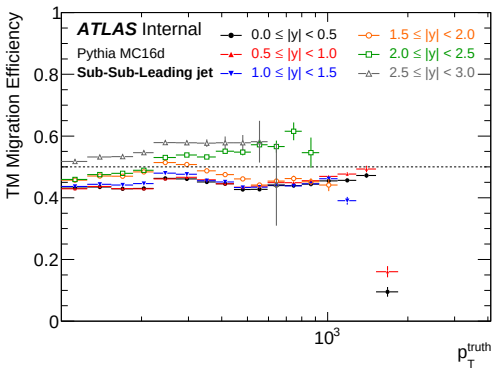
(b) Purity - Leading jet



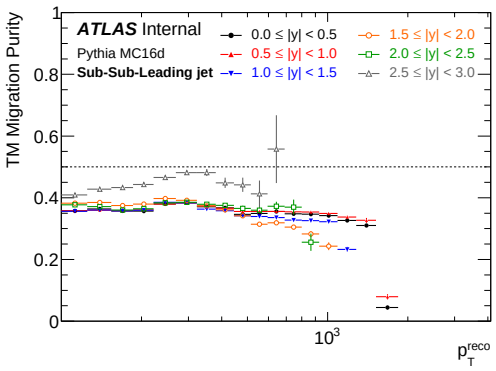
(c) Efficiency - Sub-leading jet



(d) Purity - Sub-leading jet



(e) Efficiency - Sub-sub-leading jet



(f) Purity - Sub-sub-leading jet

Figure 5.12: Transfer matrix migration efficiency and purity for each of the jet order and rapidity bins, as a function of p_T. Results using Pythia MC simulation.

5.4 Unfolding

The transfer matrix obtained in the previous section depends on the particle level MC simulation (here Pythia is used) and the detector response simulation. For the latter, after the application of the full chain of calibration as described in chapter 2.3.3 including the in-situ calibration for data, the jet energy scales and resolutions of both data and simulation at reconstructed level agree with each other, and any difference is covered by dedicated uncertainties. The improvement of the latter is already done in the calibration phase. For the particle level simulation spectrum, we know that it is not perfect, due for example to the missing orders in the simulation. In addition, we are also looking for new physics signal which is not considered in the simulation. Therefore, when unfolding the data, we do not want to be biased by the particle spectrum of the simulation. A naive method consists of first normalizing the transfer matrix in each truth bin yielding what is called the global folding matrix, the matrix containing the reconstruction probability density of a truth/particle level observable as described in the direct search chapter 4.3. Next, the matrix is inverted to get the global unfolding matrix (that is used afterward in the multiplication with the data spectrum); this method does not work since the fluctuations from low entry bins are propagated to the high entry ones destroying the precision of the result. Several unfolding methods exist that regularize the unfolding and improve the particle level input spectrum in the transfer matrix. A very powerful method, called IDS (iterative, dynamically stabilized) method [99, 100], is used extensively in the different jet cross-section measurements and is found to achieve small biases on the unfolded data spectrum. I focus, next, only on this method which is used for the leading jet measurement.

5.4.1 The IDS unfolding method

The unfolding consists of three steps, where the first (last) step is just a matching efficiency factor multiplication (division) to correct for the unmatched reconstructed (truth) jets due to acceptance or matching criteria (the factors are defined in equation 5.3). The second step, the most complex one, uses the transfer matrix already discussed before and hence only the matched truth-reconstructed jet pairs. Let me detail this second step here. The idea of the IDS method is to improve the transfer matrix used for the unfolding (only the particle level spectrum part) through iterations while providing a dynamic stabilization of several relevant variables, for example, those related to statistical fluctuation of data, large narrow new signals, background subtraction ... Many of the stabilization procedures are not necessary in this analysis and are omitted from the next formulas. The folding and unfolding probabilities are defined as:

$$P_{ij} = \frac{T_{ij}}{\sum_k T_{kj}} \text{ and } \tilde{P}_{ij} = \frac{T_{ij}}{\sum_k T_{ik}}, \tag{5.6}$$

where the folding (unfolding) probability corresponds to the probability of an entry in a reconstructed bin “i” (truth bin “j”) to come from an entry in a truth bin “j” (reconstructed bin “i”). Defining the global normalization to be applied to simulation as N_{data}/N_{MC}, the unfolding is given by:

$$u_j = t_j \frac{N_{data}}{N_{MC}} + \sum_k \mathcal{F} \Delta d_k \tilde{P}_{kj} + (1 - \mathcal{F}) \Delta d_k \delta_{kj}, \tag{5.7}$$

where u_j is the unfolded data in a bin "j", t_j is the simulation entries in the same bin, $\mathcal{F}$ is a regularization function, and the difference between data and simulation at reconstructed level is given by $\Delta d_k = d_k - r_k \times N_{data}/N_{MC}$ (d_k and r_k are the data and reconstructed simulation entries in bin "k"). The first contribution in equation 5.7 is just the truth entries in the same bin as the unfolded data multiplied by the global normalization factor. The second and third contributions account for the propagation of differences between data and reconstructed simulation, where the former propagates a fraction $\mathcal{F}$ using the unfolding probabilities, whereas the latter puts all the remaining entries in the same unfolding bin. The regularization function defines the fraction to be propagated and the fraction to be kept in the same bin, stabilizing the unfolding against statistical fluctuations of the data. The function $\mathcal{F}$ needs to be a monotone equation bounded between 0 and 1; the one used here is defined as:

$$\mathcal{F} = 1 - \exp\left[-\left(\frac{\Delta d_k}{\lambda\sigma}\right)^2\right], \tag{5.8}$$

where σ is the data statistical uncertainties and λ is a regularization parameter. The regularization method is local, in the sense that it does not impose global constrains on the shape of the unfolded distribution and corrects the transfer matrix locally in each truth bin for the differences with respect to the unfolded data.

The next step is to improve the transfer matrix which improves the unfolding probability. The folding probability is not changed since it is only related to the detector response simulation which is not changed in the unfolding methods. Defining the difference between data and simulation at truth level as $\Delta u_j = u_j/(N_{data}/N_{MC}) - t_j$, the difference in each truth bin "j" is propagated to each of the reconstructed bins "i" through the folding probabilities by:

$$T'_{ij} = T_{ij} + \Delta u_j P_{ij}. \tag{5.9}$$

After each iteration of the unfolding and the matrix improvement, the differences between data and simulation are reduced as I show later. In some cases, the second term of the equation can also be regularized by a function $\mathcal{F}$ to reduce spurious effects; this is not needed here. One would want to build a test to decide on the number of iterations needed. A simple one can be to evaluate the amount of the change of the unfolded data between successive iterations and stop when the improvement is less than some factor. In my case, I am using a data-driven closure test to estimate the bias induced by the non-perfect simulation spectra and to choose based on the bias the number of iterations needed.

To properly propagate the statistical uncertainties on the final unfolded results and calculate their correlations, the Bootstrap method (refer to chapter 3.1.4) is used where pseudo-experiments are generated fluctuating both the data events used to fill the detector level spectra and the MC simulation events used to fill the transfer matrix. In addition, pseudo-experiments on either data or MC simulation events are also used to compare the individual data or simulation contributions to the statistical uncertainties and their correlations.

5.4.2 Data-driven closure test and bias estimation

The bias comes from the remaining differences, after the iterative transfer matrix improvement, between the unfolded data and the truth spectra which also leads to difference between data and

reconstructed spectra. The bias is estimated using a data-driven closure test. The aim is to use a reconstructed MC spectrum, which is very close to the data and for which the truth spectrum is known, as a pseudo-data and unfold it using the nominal procedure used to unfold the data. For that, the truth MC spectrum is reweighted by a smooth function in such a way that the agreement between the data and the reconstructed MC spectrum (obtained from the reweighted truth spectrum convoluted with the detector response) is improved. As for all the steps of the unfolding, the detector response (≡ folding) matrix is not changed. To make the reweighted reco spectrum more representative of the available data, its relative statistical uncertainties are matched to the ones from data. The reweighted reco spectrum is now used as pseudo-data and unfolded using the same unfolding procedure for the data. The difference between the unfolded reweighted reco spectrum and the reweighted truth spectrum is the bias of the unfolding method, since the reweighted truth spectrum is the true/expected spectrum. The unfolding is iterated until the bias becomes of the same order of the statistical uncertainties, or, as I use next, until the sum of the statistical and bias uncertainties is minimized. The IDS method has shown that it can reduce significantly the bias without deteriorating the statistical precision, which is why it is chosen here as the unfolding method to be used, especially since I use a transfer matrix with jet orders for which the individual spectra are not expected to be well modeled by the simulation (they can not describe the data better than the inclusive spectrum).

5.4.3 Tuning and results

As a first step, let me show how, through iterations, the MC simulation spectra at truth and reconstructed levels are improved. In figure 5.13(a), I show the ratio between the data and the modified MC reco leading jet spectra for multiple numbers of unfolding iterations as a function of the leading jet bin number, where all the rapidity bins are shown consecutively and separated by the dashed vertical lines (from the most central at the left to the most forward at the right). Although not shown here, the unfolding and the improvement of the transfer matrix is done simultaneously for all the three jet orders. 0 iteration means that the transfer matrix is not changed; only the global normalization between data and MC simulation is applied. It is clear that for 0 iteration, the spectra difference at detector level is significantly high, higher than 5% for the majority of bins and up to 20% in some bins. The difference in the shapes of the two spectra also shows some rapidity dependent structure. After 1 iteration (≡ 1 round of improvement), the difference is considerably reduced and becomes lower than 5% except for few bins with large statistical fluctuations. The improvement is not yet perfect, so I tested more iterations. For 5 iterations and more, the difference becomes smaller than 1%.

The same is true for the ratio between the unfolded data and the modified truth MC spectra, where the difference is significantly high for 0 iteration and is reduced to less than 1% fro 5 iterations and more. This can be seen in figure 5.13(b). Let me note that here, both the unfolded data and the truth spectra are affected by the improvement of the transfer matrix and hence the unfolding. In previous analyses like inclusive jet cross-section where no jet orders are considered, only one iteration is found to be enough. In this measurement, the jet orders are taken into consideration in the transfer matrix, resulting in significant entries far from the matrix diagonal (as can be seen in figure 5.11) and also in additional correlations. Therefore, more iterations are needed to improve the transfer matrix. Also

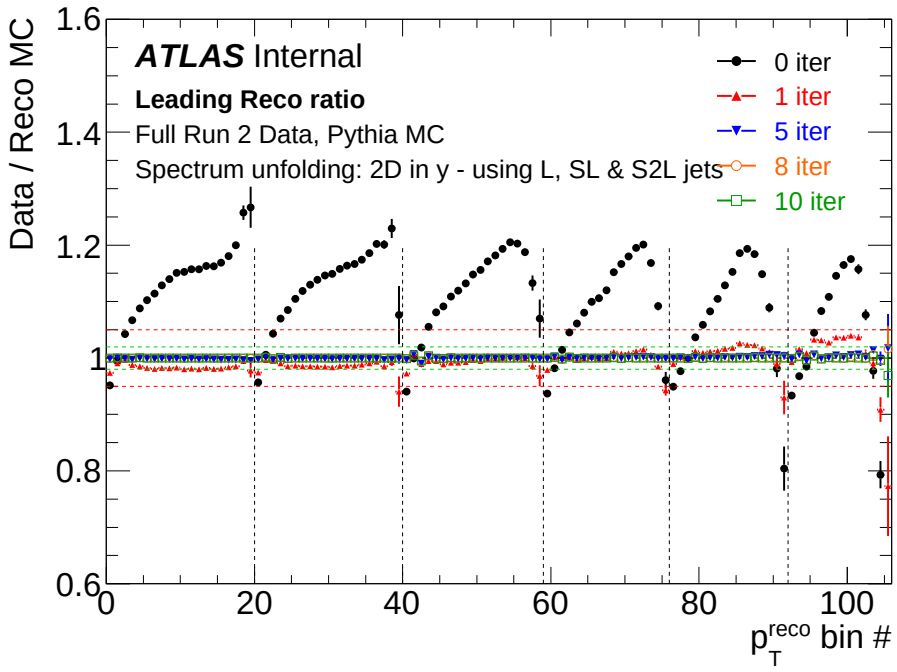
in this measurement, the jet rapidity migrations are taken into account increasing the bin migrations, but their effect on the matrix improvement is found to be small.

The last figure showed the importance and the effect of iterations on correcting the transfer matrix. The real metric to choose the optimum number of iterations is through the data driven closure test. In figure 5.14, I show separately the statistical uncertainties in a) and the unfolding bias in b) of the unfolded leading jet distribution for multiple numbers of unfolding iterations on the top pads, and the ratio to 0 iteration on the bottom pads. As expected, the statistical uncertainties increase the more we iterate due to the anti-correlations that are created by the iterations as I show later. For the bias, the iterations drastically reduce it from few percent to less than 0.5% in the majority of bins after 5 iterations. The more we iterate, the more the bias is reduced; it is less than 0.05% in the majority of bins after 10 iterations. For 0 iteration, the bias is the largest in most of the bins. Since the two uncertainties behave oppositely with respect to the iterations, the optimization of the number of iterations is therefore to minimize the (quadratic) sum of the two.

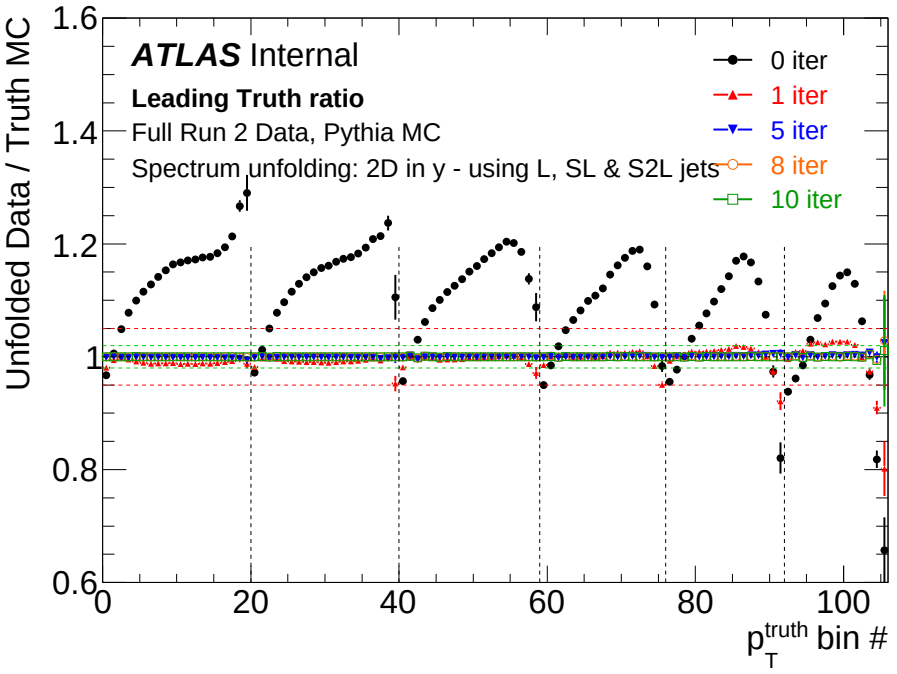
The total statistical and bias uncertainties of the unfolded leading jet are shown in figure 5.14(c). The improvement between 0 or 1 iteration and 5 and more iterations is clear in the majority of bins, except for some fluctuations. Between 5, 8 and 10 iterations, the improvement fluctuates from bin to bin. In some bins, the unfolding with 10 iterations has smaller total uncertainty, while in other bins, it is the unfolding with 5 iterations. Having more than 5 iterations does not show a decisive benefit while increasing the (anti-) correlations and being more prone to irregular spurious fluctuations. For those reasons, I choose to use the unfolding with 5 interactions as the nominal one. Using the same procedure, I also optimize the regularization parameter λ introduced in the IDS section, and I found the optimal value to be equal to 1.

Figure 5.15 shows, in the most central rapidity bin, the leading jet p_{T} distributions for data and for the unfolded data on the top pad, and their ratio on the bottom one. The unfolded data is shifted toward lower p_{T} values, which is an expected effect of the resolution on an increasingly falling spectrum.

The effect of the iterative unfolding on the correlations between the unfolded data bins can be seen by comparing correlation matrices for 0 and 5 iterations. In figure 5.16, I show, only for the most central rapidity bin for better visibility, the full correlation matrices (including both data and simulation effects calculated using the Bootstrap method) of the unfolded results for the three jets orders considered in this analysis for 0 and 5 iterations. For 0 iteration, positive correlations are visible between the leading and sub-leading jets, notably when the two are in the same p_{T} bin. For 5 iterations, those correlations do not exist anymore. In addition, anti-correlations in the bins adjacent to the diagonal now appear.

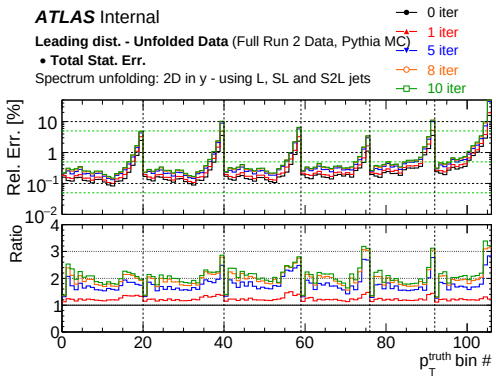


(a) Detector level

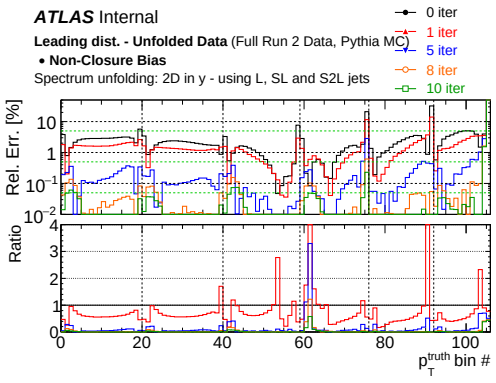


(b) particle level

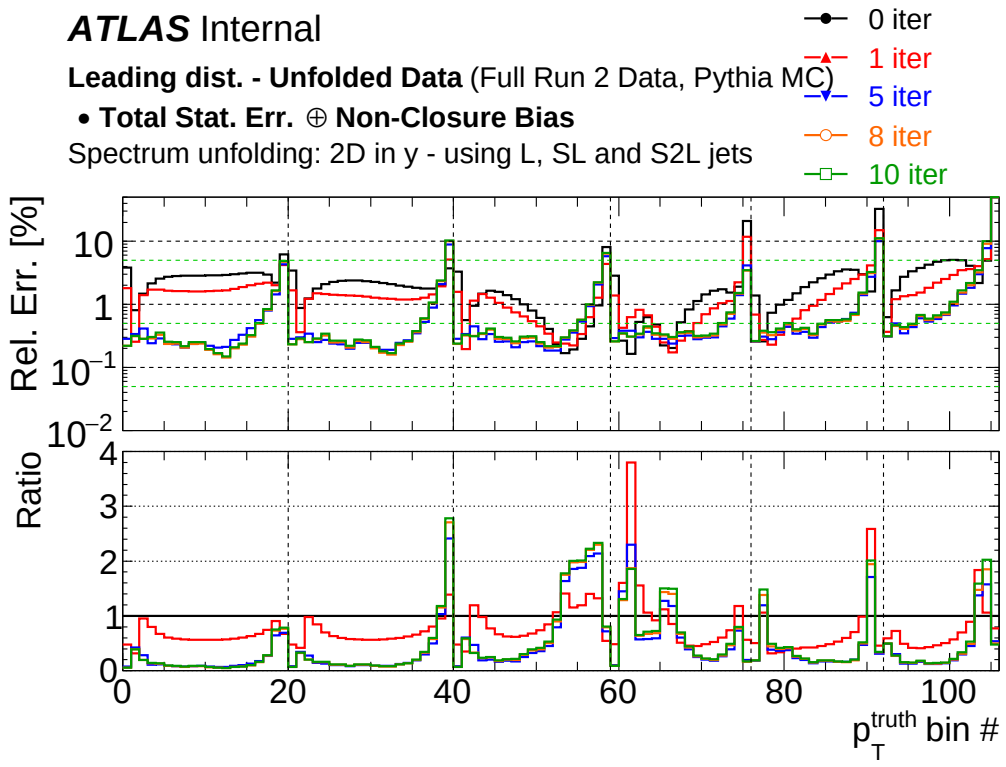
Figure 5.13: Ratio between data and the modified MC reco leading jet spectra in a) and between unfolded data and the modified MC truth leading jet spectra in b) for multiple numbers of unfolding iterations as a function of the leading jet bin number, where all the rapidity bins are shown consecutively and separated by the dashed vertical lines (from the most central at the left to the most forward at the right).



(a) Unfolded leading jet - Statistical uncertainties



(b) Unfolded leading jet - Unfolding bias



(c) Unfolded leading jet - Statistical uncertainties ⊕ unfolding bias

Figure 5.14: Statistical uncertainties in a), unfolding bias in b) and their quadratic sum in c) of the unfolded leading jet distribution for multiple numbers of unfolding iterations on the top pads and the ratio to 0 iteration on the bottom pads, as a function of the leading jet bin number, where all the rapidity bins are shown consecutively and separated by the dashed vertical lines (from the most central at the left to the most forward at the right).

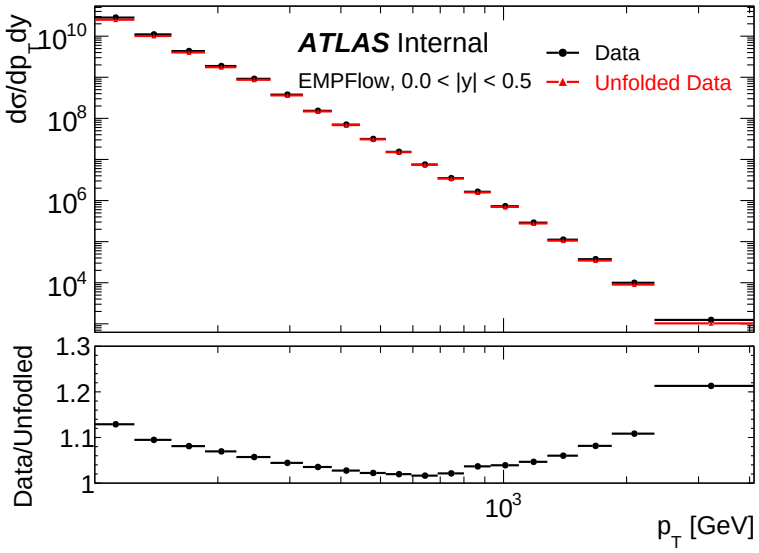
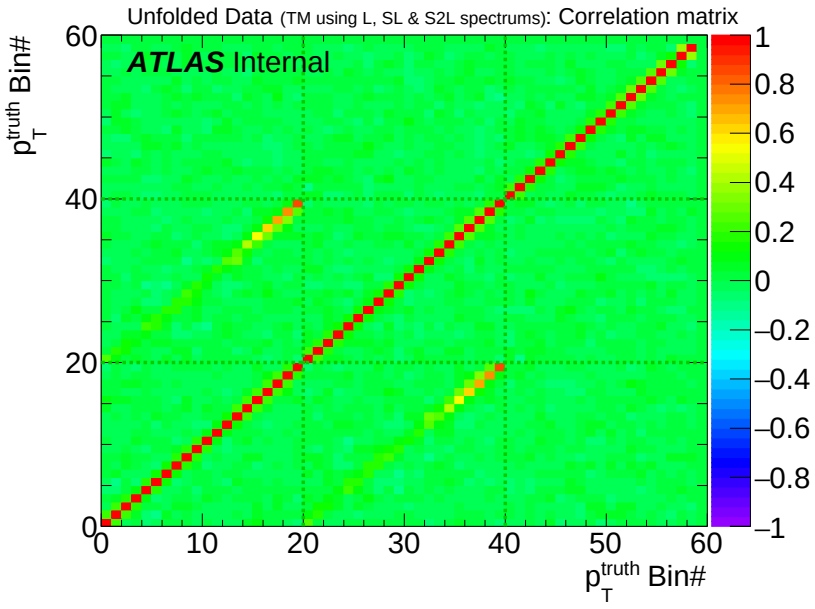
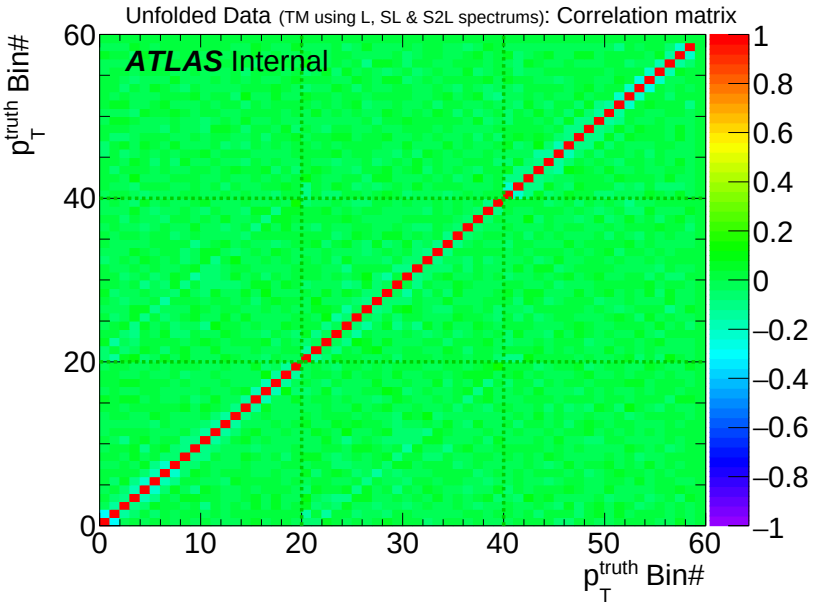


Figure 5.15: Leading jet p_{T} distributions for data and for the unfolded data on the top pad, and their ratio on the bottom one in the most central rapidity bin.



(a) 0 unfolding iteration

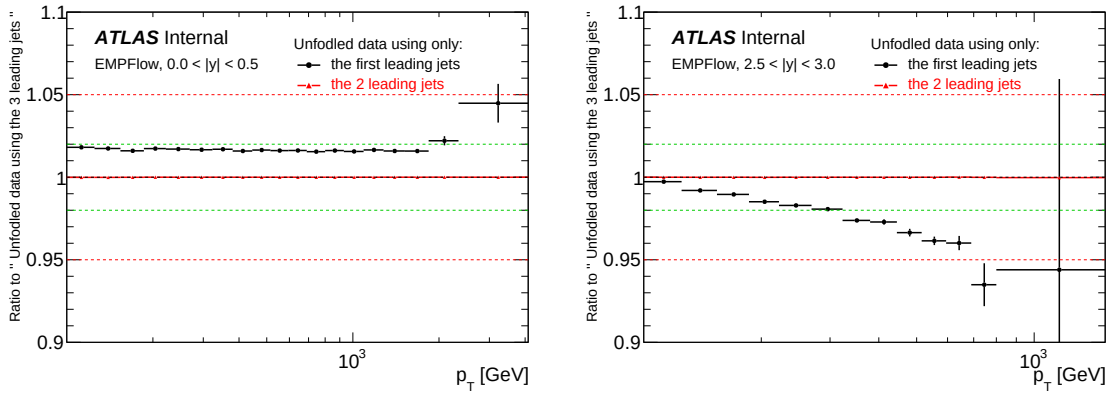


(b) 5 unfolding iterations

Figure 5.16: Full correlation matrices (including both data and simulation effects) of the unfolded results for three jets orders for 0 and 5 iterations.

5.4.4 The effect of jet order flips

It is important to check the effect of the jet order flips on the final unfolded data results. More specifically, I test here the effect of including only one or two leading jets in the definition of the data spectra and the transfer matrix used in the unfolding procedure. Taking the case of using only leading jets as an example, the leading jets at truth or reco levels that have their order flipped for their matched jets are not included in the transfer matrix; instead, they are accounted for by the matching efficiency terms of the unfolding, defined in equation 5.3. The test checks if the higher reliance on efficiency corrections induces a bias on the unfolded results. In figure 5.17, I show the ratio of the unfolded leading jet spectra between an unfolding using only one (in black) or two (in red) leading jet orders over the one using three leading jet orders (the nominal unfolding). The statistical uncertainty on the ratio is shown and computed using the Bootstrap method to properly account for the correlations between the unfolding results of the different cases. For both rapidity bins shown, the unfolding using only leading jets shows a systematic deviation from unity with a difference up to 5%. In this case, a non negligible bias is introduced due to the missing jet order flips not included in the transfer matrix. On the other hand, for the case where the unfolding uses the two leading jet orders, the ratio deviation from unity is less than a permille level. This is expected since the probability of a leading jet at reco or truth level to become a third jet at the other level is very small. In addition, this is assuring that no additional bias exists from not including four or more leading jet orders in the unfolding.



(a) $0.0 < |y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 5.17: The ratio of the unfolded leading jet spectra between an unfolding using only one (in black) or two (in red) leading jet orders over the one using three leading jet orders (the nominal unfolding), in the most central and most forward rapidity bins as a function of p_{T}. The error bars represent the statistical uncertainty computed using the Bootstrap method.

Event by event unfolding

It is worth noting that I tried a different unfolding method which is done event by event, as opposed to the current method which uses the full observable spectrum. The procedure is first to assign an unfolding probability, the probability distribution of the jet p_{T} at particle level, for each jet of the event based on its p_{T} and rapidity. Next, all the probabilities are convoluted to obtain the probability

distribution of the leading jet at particle level and is used to fill the observable spectrum. A second method is to calculate, by integrating the probability distribution obtained from the convolution, the total probability of each jet to be the leading one at particle level, and then each jet is filled with a weight equal to the latter probability. For both methods, a non-closure is observed when performing what is called a technical closure test. This test consists of using the same MC events to build the transfer matrix and hence the unfolding probabilities, and to build the reco spectrum to be unfolded. In this case, the unfolded reco spectrum and the truth spectrum should match perfectly; using this new unfolding method, a few percent non-closure is observed. More development is needed to check the source of this non-closure and whether it can be eliminated or if this method has a non-reducible bias. For the rest of the analysis, the nominal IDS unfolding is used.

5.5 Systematic uncertainties

5.5.1 JES

The various jet energy scale uncertainties from the jet calibration (related to both simulation and in-situ methods) described in chapter 2.3.3 are propagated to the unfolded cross-section results. Using each JES uncertainty, the p_T of all jets in (pseudo-)data are fluctuated coherently by the uncertainty value upward or downward ($\equiv \pm 1\sigma$). This way, each systematic uncertainty is treated as fully correlated across the phase space (p_T, y ...) but independent of the other systematic uncertainties. For each JES uncertainty, two new systematically variated, up and down, spectra are obtained which are then unfolded using the nominal transfer matrix. The deviation of the unfolded variated (pseudo-)data with respect to the unfolded nominal (pseudo-)data is the systematic uncertainty on the cross-section coming from the JES uncertainty. Although JES uncertainties are symmetric, their effect on the cross-section is asymmetric due essentially to the non-linearly falling spectra; that is why both up and down variations need to be evaluated. Both data or reco-level simulation used as pseudo-data can be used to propagate the uncertainties. The preference to use pseudo-data is its higher statistics specially at hight p_T as can be seen in figure 5.18. Hence, I use pseudo-data for the JES and also the JER (see next section) uncertainties propagation in this analysis.

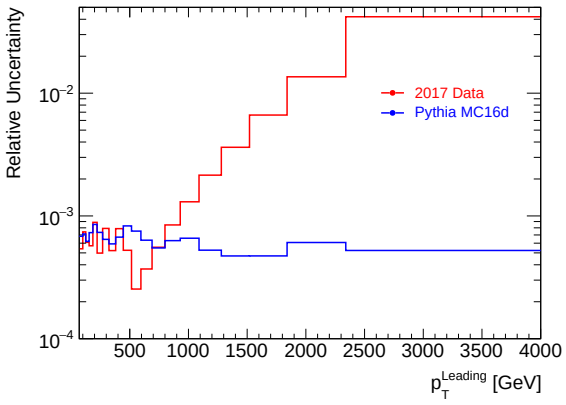


Figure 5.18: Relative statistical uncertainties of the leading jet distribution in data and in simulation as a function of p_T.

5.5.1.1 Flavor uncertainties

For the propagation of the flavor response and composition uncertainties as introduced in chapter 2.3.3.3, the gluon fraction and its uncertainty should be provided. Truth partons from the simulation are matched to reco jets, and each jet is considered initiated by the highest p_T parton matched with it. The nominal gluon fraction is determined using all jets of the nominal Pythia MC simulation. For the fraction uncertainty, I compare the nominal fraction with the result from different generators, here I used Powheg-Pythia and Sherpa. In addition, since jet orders are used for the measurement and the unfolding, I also compare the nominal fraction with the ones measured using

only the leading or the sub-leading jets. The maximum deviation in each (p_T,y) bin between the nominal and all the variated fractions is taken as the uncertainty. Figure 5.19 shows in text the gluon fraction from inclusive jets samples using Pythia, Powheg-Pythia and Sherpa MC's as a function of p_T and y. The gray-scale represent the nominal fraction from Pythia MC. An average difference of about 7-10% can be seen in the majority of bins.

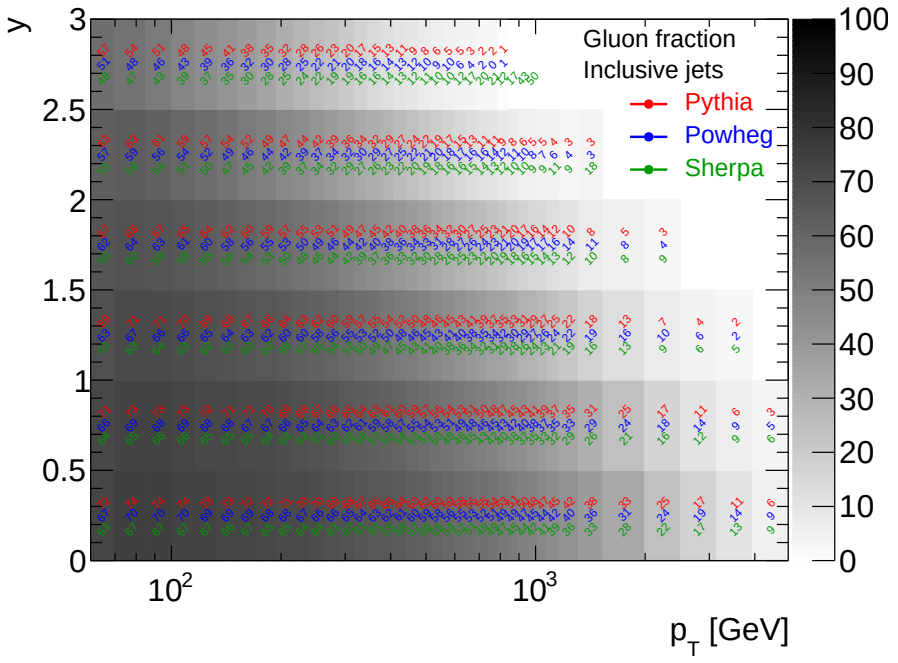


Figure 5.19: Gluon fractions shown in text from inclusive jets samples using Pythia, Powheg-Pythia and Sherpa MC's as a function of p_T and y. The gray-scale represent the nominal fraction from Pythia MC.

5.5.2 JER

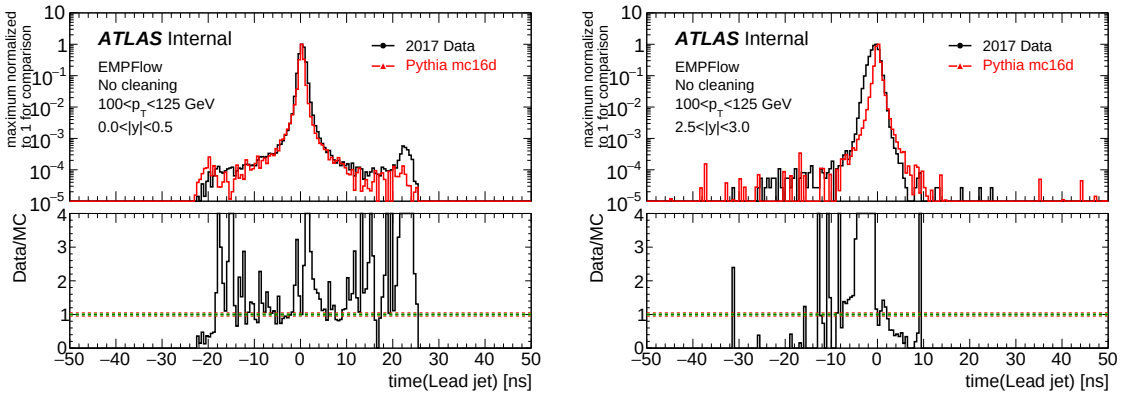
The uncertainties on the jet resolution are propagate by smearing the jet p_T using a Gaussian function with a variance $\sigma^2_{\text{smear}} = (\sigma_{\text{nominal}} + \Delta\sigma_{\text{uncertainty}})^2 - \sigma^2_{\text{nominal}}$. When the resolution needs to be increased, the simulation reco jets used to build the transfer matrix (TM) are smeared. A new TM is obtained. On the other hand, when the resolution needs to be decreased, it is not straightforward how to decrease the resolution of jets in the TM. Instead, the jets in the (pseudo-)data are smeared giving the same effect. Hence, for the JER uncertainty propagation, jets in both (pseudo-)data and transfer matrices are smeared depending on each region, and the unfolding is done coherently using the variated spectrum and the variated TM for each uncertainty.

5.5.3 Luminosity

The total uncertainty on the 139 fb^{-1} full Run II data luminosity is 1.7%. This uncertainty is added to the uncertainties of the unfolded leading jet distribution.

5.5.4 Jet time cut

As already mentioned in section 5.2.2, I apply a jet time cut of ±10 ns to reject out-of-time jets. Since the jet time resolution is not much smaller than the bunch-crossing interval, in-time jets can have their time outside the ±10 ns range. Hence, an inefficiency of this cut needs to be measured and a systematic uncertainty needs to be evaluated. This cut is, as usual, applied both to data and to simulation. Let us first compare the jet time distributions between data and simulation to check if the out-of-time effect is present in simulation. In figure 5.20, I show the data and simulation time distributions in the most central and the most forward rapidity bins, and in the first p_{T} bin in the upper plots and their ratio in the bottom plots. Each distribution has an arbitrary y-axis where its maximum is equal to unity to facilitate the comparison. As can be seen, the simulation has additional bumps at $\pm 20 - 25$ ns coming from out-of-time jets. I checked also that the general shape of the two distributions are close for all rapidity bins, except in the most forward one where the data time distribution is much larger and more asymmetric than the simulation one. The two distributions have differences in their jet time resolution which is reflected by the ratio being different than unity.



(a) $0.0 < |y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 5.20: Data and simulation time distributions in the upper plots and their ratio in the bottom plots in the most central and the most forward rapidity bins, for the $100 < p_{\mathrm{T}} < 125$ GeV range. Each distribution has an arbitrary y-axis where its maximum is equal to unity to facilitate the comparison.

To measure the inefficiency of the cut, I first fit the core of the distribution between ±10 ns to obtain a function $\mathcal{F}$ describing only the in-time jet time, and then use the relation:

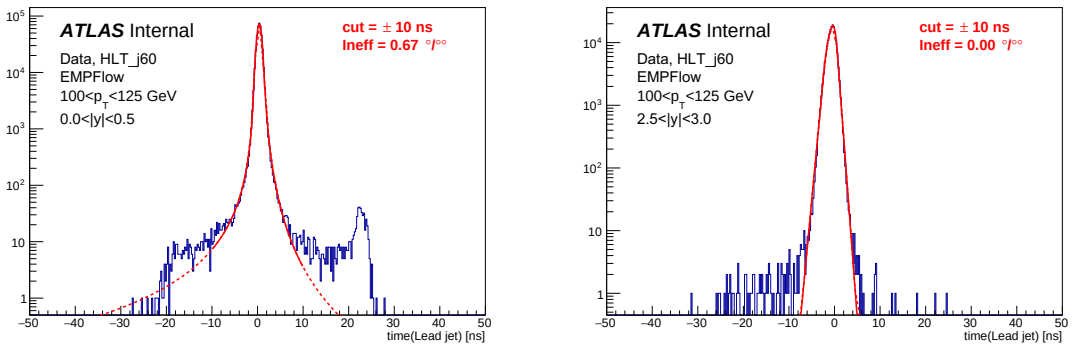
$$\text{inefficiency} = 1 - \frac{\int_{-10}^{+10} \mathcal{F}\,\mathrm{dt}}{\int_{-\infty}^{+\infty} \mathcal{F}\,\mathrm{dt}}. \tag{5.10}$$

Concerning the choice of the function $\mathcal{F}$, I use a double-sided crystal-ball which is flexible enough to describe the asymmetric sides of the distribution and its large tails. The function is a Gaussian in

the core and a power law in the tails with the two sides asymmetric and defined as follows:

$$\mathcal{F} = \mathrm{N} \times \begin{cases} \mathrm{A_L}(\mathrm{B_L} - \frac{\mathrm{x}-\bar{\mathrm{x}}}{\sigma_\mathrm{L}})^{-\mathrm{n_L}}, & \text{for } \frac{\mathrm{x}-\bar{\mathrm{x}}}{\sigma_\mathrm{L}} \leq -\alpha_\mathrm{L}, \\ \exp\left(-\frac{(\mathrm{x}-\bar{\mathrm{x}})^2}{2\sigma_\mathrm{L}^2}\right), & \text{for } -\alpha_\mathrm{L} < \frac{\mathrm{x}-\bar{\mathrm{x}}}{\sigma_\mathrm{L}} \leq 0, \\ \exp\left(-\frac{(\mathrm{x}-\bar{\mathrm{x}})^2}{2\sigma_\mathrm{R}^2}\right), & \text{for } 0 < \frac{\mathrm{x}-\bar{\mathrm{x}}}{\sigma_\mathrm{R}} \leq \alpha_\mathrm{R}, \\ \mathrm{A_R}(\mathrm{B_R} - \frac{\mathrm{x}-\bar{\mathrm{x}}}{\sigma_\mathrm{R}})^{-\mathrm{n_R}}, & \text{for } \alpha_\mathrm{R} \leq \frac{\mathrm{x}-\bar{\mathrm{x}}}{\sigma_\mathrm{L}}. \end{cases} \tag{5.11}$$

The function depends on 8 parameters, $\bar{\mathrm{x}}$, σ_L, σ_R, $\mathrm{n_L}$, $\mathrm{n_R}$, α_L, α_R and N (A and B parameters are functions of n and α). Figure 5.21 shows two data jet time distributions in the most central and the most forward bins, their fits with the function $\mathcal{F}$ and the inefficiencies of a ±10 ns cut. The fit function inside the fit range is drawn in a solid line, whereas the dashed line represents its extrapolation. A noticeable behavior in both data and simulation is that the forward bins have less out-of-time jets and hence higher efficiency. In addition, the higher the p_T regions is, the less out-of-time jets exist. The asymmetric distributions and fits are also visible in the figure.



(a) $0.0 < |\mathrm{y}| < 0.5$ (b) $2.5 < |\mathrm{y}| < 3.0$

Figure 5.21: Data jet time distributions and their fits in the most central and the most forward bins, for the $100 < p_\mathrm{T} < 125$ GeV range. The inefficiencies of a ±10 ns cut are also shown.

Table 5.3 summarize the inefficiencies of a time cut on the leading jets in the lowest two p_T bins and the different rapidity bin in data and simulation. Let me recall that the bin $75 < p_\mathrm{T} < 100$ GeV is only used for the unfolding step and dropped for the final result. The dependence of the inefficiency described before with respect to p_T and rapidity can be clearly seen in the table. The simulation not only shows the same dependence, but also has inefficiency values comparable to the data ones; the differences being at most about 1‰. Therefore, the unfolding will correct the majority of this inefficiency. In addition, for the first bin to be used for the final measurement ($100 < p_\mathrm{T} < 125$ GeV), the maximum inefficiency is just 2.2‰. For the next p_T bin (not shown here), the inefficiencies drop to negligible values. I also checked the inefficiencies on the sub-leading jets time cut. The general behavior is the same as for the leading jet with slightly smaller values.

As the inefficiencies are small and comparable between data and simulation, there is no need to propagate the inefficiency using the unfolding from the detector level to the particle level. I consider

Table 5.3: Inefficiencies of a ±10 ns time cut on the leading jets in the lowest two p_T bins and the different rapidity bins in data and simulation.

y	Data $75 < p_T < 100$	Simulation $75 < p_T < 100$	Data $100 < p_T < 125$	Simulation $100 < p_T < 125$ GeV
$0.0 < \|y\| < 0.5$	4.0 ‰	5.5 ‰	0.7 ‰	0.7 ‰
$0.5 < \|y\| < 1.0$	9.4 ‰	9.4 ‰	1.6 ‰	1.3 ‰
$1.0 < \|y\| < 1.5$	14.4 ‰	13.8 ‰	2.2 ‰	1.6 ‰
$1.5 < \|y\| < 2.0$	4.5 ‰	4.5 ‰	0.6 ‰	0.3 ‰
$2.0 < \|y\| < 2.5$	0.7 ‰	2.0 ‰	0.1 ‰	0.2 ‰
$2.5 < \|y\| < 3.0$	0.4 ‰	0.2 ‰	0.0 ‰	0.0 ‰

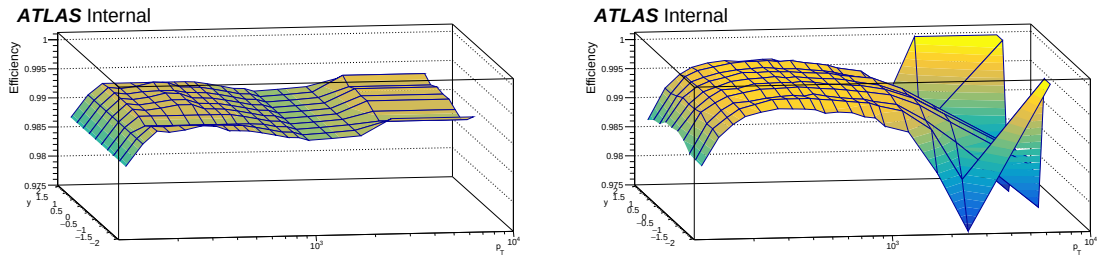
directly the full data inefficiencies shown in the table 5.3 for the $100 < p_T < 125$ GeV range as a time cut systematic uncertainty on the unfolded leading jet result. The choice of the full data inefficiencies and not the difference between data and simulation is a more conservative choice, but it is not an issue due to the small values of the uncertainties.

5.5.5 Jet cleaning

As we already saw in section 5.2.2, a tight cleaning on the leading jet is used to highly reject fake jets but also helps with the rejection of out-of-time pile up. For the additional jets used in the measurement, the second and third jets, only a loose cleaning is applied to reduce the total inefficiency of the selection but on the same time to keep a good rejection efficiency of fake jets. Two veto strategies for cleaning exist. The first, called event veto and used previously, rejects all the event if any jet does not pass the required cleaning level. Its motivation is to reduce measurement biases due to fake jets. The inconvenient is that the total inefficiency, in addition to being increased, is itself dependent on the event kinematic, since the jet cleaning efficiency depends on the jet p_T and rapidity. The new proposal is to use a jet veto strategy, where the jet not passing the cleaning is excluded from the accepted event jets, but the event is kept. After a jet is excluded, the jet orders are rearranged to replace the order of the excluded jet; if for example the highest p_T jet does not pass the tight cleaning to be considered as a leading jet, it is excluded and now I test if the second highest p_T jet passes also the tight cleaning and if so it is considered as the leading jet. The motivation of this veto is to reduce the inefficiencies, to avoid a veto which is dependent on the event kinematic and to rely on the unfolding to correct the inefficiencies.

One would first think that a simple multiplicative factor, equal to the inverse of the inefficiency, on the data or simulation reco spectra corrects the inefficiency effects, but this is not the case due to the kinematic dependences and jet order migrations. In addition, I want to measure a systematic uncertainty to cover the cleaning differences between data and simulation. In fact, the cleaning efficiency is different between data and simulation as can be seen in figure 5.22 where the tight

cleaning efficiency is shown as a function of p_{T} and y. At low and mid p_{T}, the efficiency in simulation is higher, then, after 2 TeV, it becomes smaller compared to the one for the data.



(a) Data (b) Pythia simulation

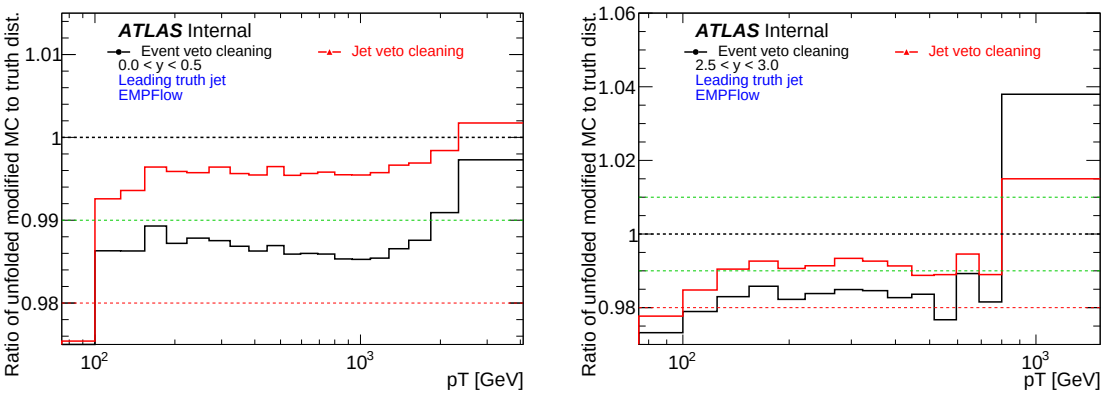
Figure 5.22: Jet tight cleaning efficiencies in data in a) and in simulation in b) as a function of p_{T} and y.

To compare the two veto strategies and evaluate a cleaning systematic, I use the following closure test:

- the transfer matrix (TM) used for the unfolding is constructed using the nominal procedure, applying the simulation cleaning only on the reco jets; two TM are obtained, one for each cleaning veto procedure,
- a pseudo-data spectrum is obtained using simulation reco jets but where the cleaning efficiency is simulated using the data cleaning efficiency histogram of figure 5.22; same here, two TM are obtained, one for each cleaning veto procedure,
- for each veto strategy, the corresponding pseudo-data spectrum is unfolded with the corresponding TM,
- the unfolded pseudo-data spectrum is compared with the simulation truth jets one, and any deviation is due to the difference between data and simulation cleaning.

Let me note that in MC simulation, fake jets are not simulated and thus jets not passing the cleaning are just due to the inefficiency of the cleaning.

In figure 5.23, I show the ratio between the unfolded pseudo-data and the simulation truth spectra for the most central and the most forward rapidity bins. As expected, the closure is not perfect and significant deviation from unity are visible. The ratio is smaller than unity for low and mid p_{T} and then becomes higher than unity at high p_{T}, as is expected from the difference of the efficiencies shown in figure 5.22. What is interesting is the non-closure difference between the two veto strategies. The event veto method has roughly double the non-closure compared to the jet veto method. In the central rapidity bin and mid p_{T}, the non-closure is on average 1.2% for the event veto method compared to just 0.5% for the jet veto method. One additional point in favor of using the jet veto method is that for the event veto cleaning, additional inefficiencies, not accounted for in this closure test, come from events rejected due to the existence of real fake jets, whereas for the jet veto method only the fake jet is removed. The evaluation of those additional inefficiencies is more tricky, but we can fairly say that the inefficiency of event veto cleaning can be even worse.



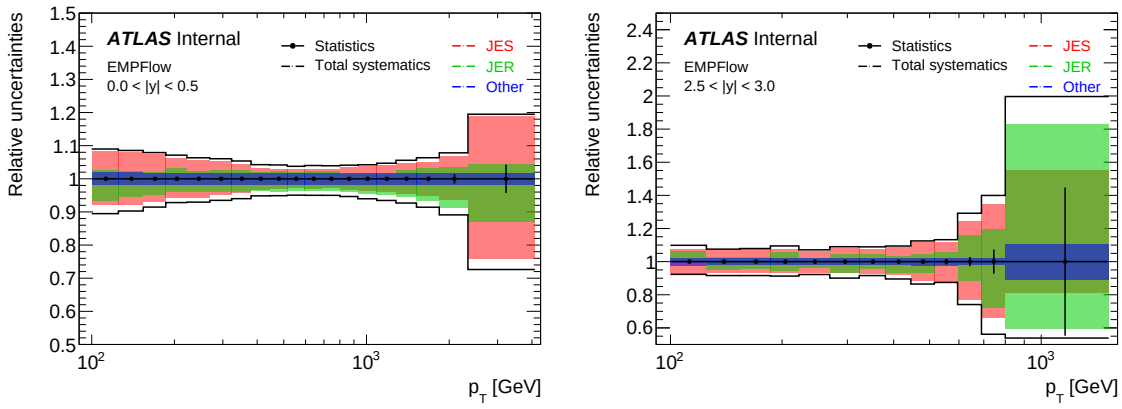
(a) $0.0 < |y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 5.23: Ratio between unfolded pseudo-data and simulation truth spectra for the most central and the most forward rapidity bins as a function of p_{T}. The pseudo-data spectrum is obtained using simulation reco events but where the cleaning efficiency is simulated using the data cleaning efficiency. The results from two cleaning strategies are shown: the event veto in black and the jet veto in red.

From the results in figure 5.23 and the above mentioned reasons, it is clear that the jet veto method has smaller total inefficiencies and a better simulated effect; hence I use this method as the default one. I also use the non-closure values obtained from this closure test as the cleaning systematic uncertainty to cover the difference between data and simulation cleaning.

5.5.6 Total systematic uncertainties

The total systematic uncertainties (black line envelope) are shown in figure 5.24 and in addition separately the JES, the JER and all the other systematic uncertainties. The statistical uncertainties are also shown as black error bars centered at unity. As expected, the JES uncertainties dominates in almost all bins, the total uncertainties are smallest at mid p_{T} ranges ($\approx$ 4% (8%) in the most central (forward) rapidity bin) and are largest at high p_{T}. Asymmetric systematic uncertainties are also visible, due to the non-linearly falling spectra as a function of p_{T}. The statistical uncertainties are much smaller than the systematic ones (<1%) except at the phase-space limits.



(a) $0.0 < |y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 5.24: JES (red band), JER (green band), others (blue band) and total (black line envelope) systematic uncertainties shown for the most central and most forward rapidity bins as a function of p_{T}. The "others" uncertainty group includes the jet cleaning, the jet time cut, the luminosity and the unfolding bias uncertainties. The statistical uncertainties are also shown as black error bars centered at unity.

5.6 Theoretical prediction

5.6.1 Fixed order calculations

For the theoretical prediction, the first step is to perform fixed order calculations. Recalling equation 1.30 (and changing the total cross-section to a differential one),

$$\mathrm{d}\sigma_{\mathrm{pp}\to\mathrm{X}} = \sum_{\mathrm{a,b}} \int_0^1 \mathrm{dx_a} \int_0^1 \mathrm{dx_b}\ \mathrm{f_{a/p}}(\mathrm{x_a}, \mu_{\mathrm{F}}^2)\mathrm{f_{b/p}}(\mathrm{x_b}, \mu_{\mathrm{F}}^2)\mathrm{d}\hat{\sigma}_{\mathrm{ab}}(p_a, p_b, \mu_F^2, \mu_R^2),$$

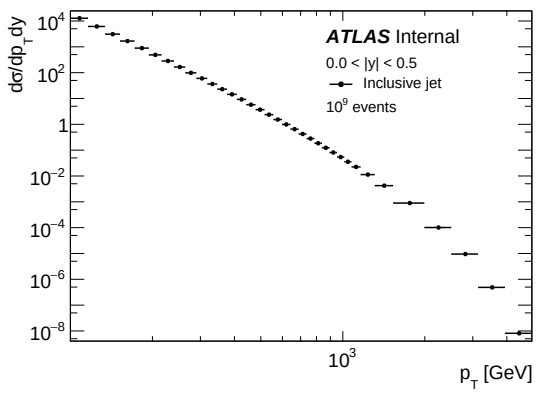
the differential cross-section for a scattering process is factorized into a hard perturbative partonic sub-processes $d\hat{\sigma}_{ab}$ convoluted with proton PDFs. The calculation of $\mathrm{d}\sigma$ hence depends on many parameters such as the PDF inputs, the scales μ_{F} and μ_{R}, the reference value of strong coupling constant α_{S} and so on. If one wants to use many PDF sets from different experiments or to propagate the uncertainties from the different inputs, the calculations become very cpu intensive. To facilitate the predictions, the APPLGRID project [101] was developed. It consists of including a-posteriori the convolution with the PDFs and the dependence on the scales and α_{S}. For that, 3-dimensional grids are used for each observable bin to store the perturbative coefficients (≡ the amplitude of the Feynman diagrams as a function of α_{S}) as a function of a scale and the two Bjorken-x of the two scattering partons (a mapping is used to change the axes to optimize the 3D space). For each perturbative order and for each type of partonic sub-processes, a 3D grid is generated. For the case of QCD interaction, the grids are split into seven partonic sub-processes:

$$\mathrm{gg},\ \mathrm{qg},\ \mathrm{gq},\ \mathrm{qq},\ \mathrm{q\bar{q}},\ \mathrm{qq'} \text{ and } \mathrm{q\bar{q}'},$$

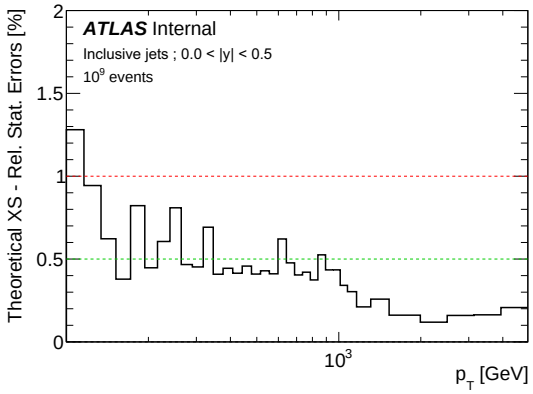
where g, q and q′ denotes respectively a gluon, a quark and a different flavor quark. The perturbative coefficients are calculated using fixed order generators 1.3.1 (with MC sampling of the phase-space);

for this analysis, I use the program NLOJET++. Then, one can a-posteriori convolute with the PDF inputs, choose the reference value of α_S and vary the scales by a multiplicative factor. For the convolution with PDFs, the LHAPDF package [102] is used along HOPPET package [103] for the evolution of the PDF to the required scale.

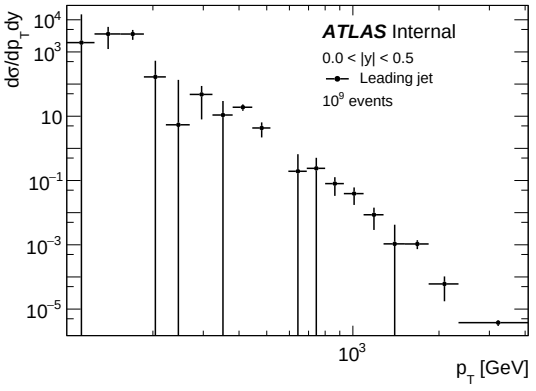
Using the above described method, I produced theoretical predictions for both inclusive jets and leading jet observables. In figure 5.25, I show, for the central rapidity bin, the differential cross-sections and the relative statistical uncertainties for each of the two observables. Both predictions are calculated using 10^9 events. For the statistical uncertainties determination, the prediction in split into N samples of equal number of events and the following relation is used: $\Delta^{\text{stat}} = \text{RMS}/\sqrt{\text{N}}$. The first striking thing to notice is that, even though the same number of events are used, the leading jet observable has much higher uncertainties, sometimes bigger than the cross-section value itself, and that the magnitude of the uncertainty changes a lot between adjacent bins. In addition, its cross-section fluctuates a lot and sometimes becomes even negative. This is a sign that large positive and negative perturbative coefficients are produced in the calculation and are not well canceled.



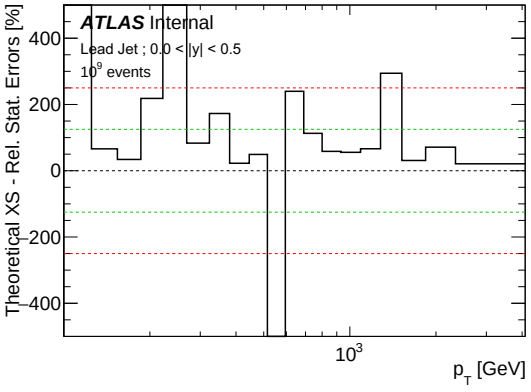
(a) Inclusive jets - differential cross-section



(b) Inclusive jets - relative statistical uncertainties



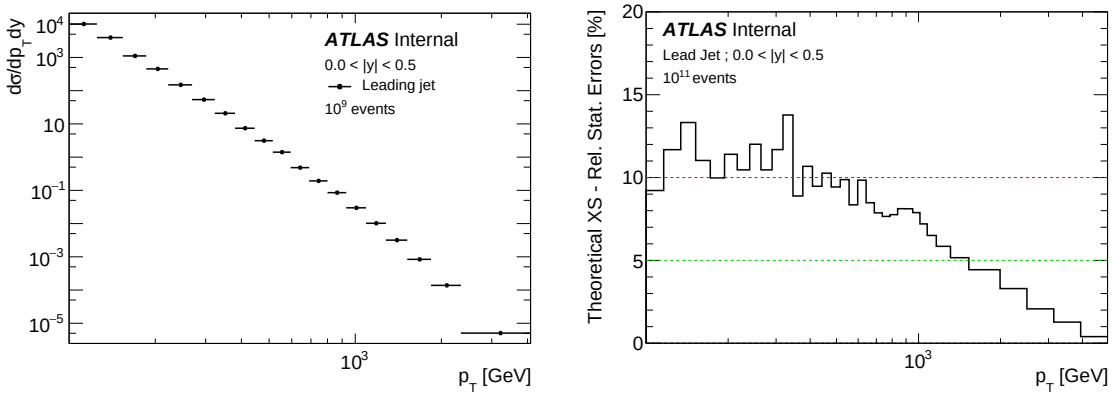
(c) Leading jet - differential cross-section



(d) Leading jet - relative statistical uncertainties

Figure 5.25: Differential cross-sections and relative statistical uncertainties for inclusive jets (a, b) and leading jet (c, d) observables in the central rapidity bin. The number of simulated events for each distribution is 10^9.

Increasing the number of events by 100 fold to 10^{11}, the leading jet cross-section becomes more stable, no more negative values exists and the statistical uncertainties behave more regularly as shown in figure 5.26. Still, the uncertainties are 100 times larger than the ones from inclusive jet observable for the same number of events. This is due to the large positive and negative coefficients that widen the distribution of the samples predictions in each bin leading to a large value of the RMS. In addition, a 10% uncertainty is larger than what we aim, and producing much more events is not feasible. In any case, we should dig deeper into what is the source of this behavior.



(a) Leading jet - differential cross-section

(b) Leading jet - relative statistical uncertainties

Figure 5.26: Differential cross-section and relative statistical uncertainty for leading jet observable in the central rapidity bin. The number of simulated events is 10^{11}.

As it turns out, the leading jet observable is infra-red (IR) sensitive. Let me detail the behavior of the leading jet observable using the example diagrams of the figure 5.27. For both LO and NLO with one virtual loop diagrams, the leading jet observable is degenerate: the event has two same p_{T} partons which are reconstructed as two same p_{T} jets. In this case, I treat both of the jets as leading and fill both of them with a 1/2 weight. For NLO with real emission diagrams, if the emission is within the jet reconstruction cone, the same degenerate case as before is produced. On the other hand, in the case of a real emission at a large angle, the degeneracy is broken and only one leading jet is obtained. Considering now the IR cases of very soft real or virtual contributions, as I already detailed in chapter 1.3.1, the diagram amplitudes are large but opposite; let me denote them $\mathcal{M}$ and $-\mathcal{M}$ respectively. Considering multiple rapidity bins and that the two jets "a" and "b" are in different rapidity bins y^{a} and y^{b}, the sum of the contributions from real and virtual diagrams are $\mathcal{M}^{\mathrm{real}} - 0.5\mathcal{M}^{\mathrm{virtual}}$ in y^{a} bin, and $-0.5\mathcal{M}^{\mathrm{virtual}}$ in y^{b} bins. As can be seen, the cancellation of large contributions is unsatisfied. Of course, the additional case of the parton "a" being the one who emits the radiation should be considered, which after being added the cancellation is satisfied. Hence, if when simulating the phase-space only one emission from one side is simulated, the cancellation is not perfect leading to large positive or negative contributions and thus large fluctuations between the results of the different samples. This is what I meant by IR sensitivity. Let me note that, for the sake of clarity, I did not mention here the counter terms that are used to subtract the divergences in each type of NLO diagrams, and since when adding all the contributions their effect cancels.

To verify my conclusion on the behavior of the leading jet observable and its IR sensitivity, I produced

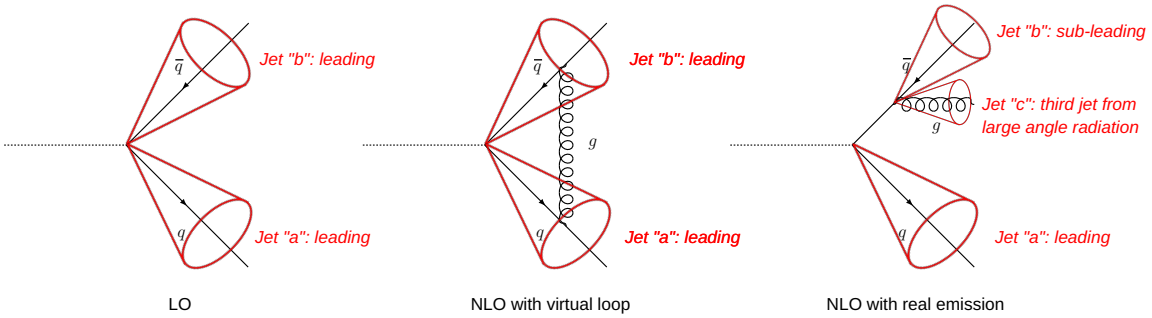
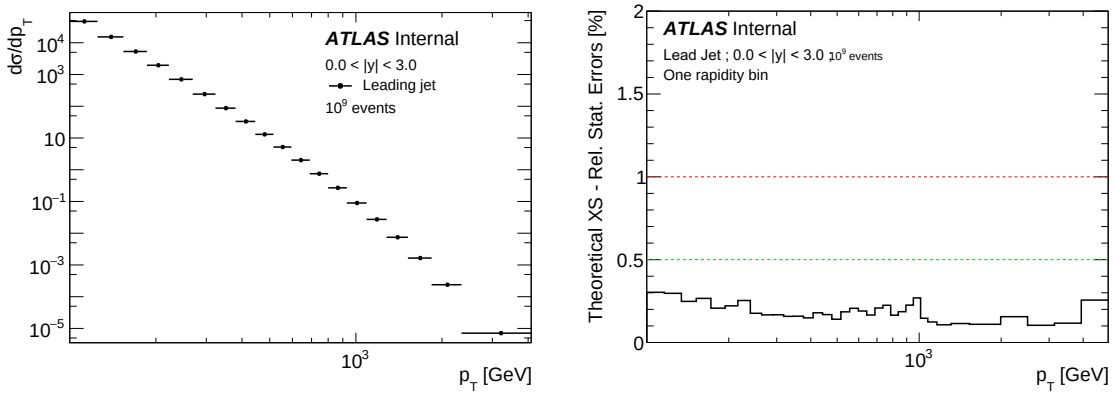


Figure 5.27: Examples of Feynman diagrams for the different fixed order contributions, LO, NLO with virtual loop or real emission. The reconstructed jets in each diagram are shown along their p_{T} orders.

theoretical predictions using an inclusive rapidity bin $0 < |y| < 3$. Figure 5.28 shows the results using 10^9 events. As can be seen, no negative or large fluctuations are observed, and the statistical uncertainties are at a permille level even with moderate number of simulated events. This test agrees with what is expected from the IR sensitivity explanation given before.



(a) Leading jet - differential cross-section

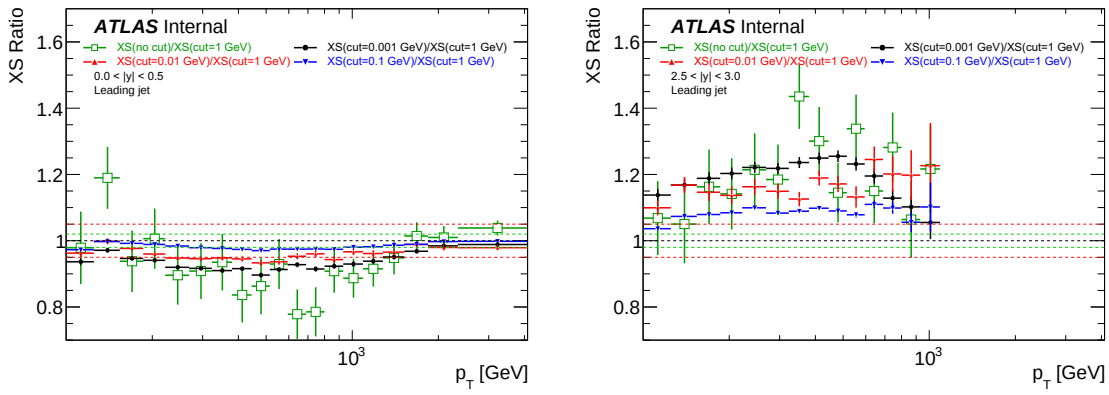
(b) Leading jet - relative statistical uncertainties

Figure 5.28: Differential cross-section and relative statistical uncertainty for leading jet observable using one inclusive rapidity bin. The number of simulated events is 10^9.

Back to the multiple rapidity bins case, to try to stabilize the prediction, I test an approach forcing the large diagram amplitudes from real emissions to cancel out with the virtual ones. For that, I define a cut-off value Λ when comparing the first two leading jets in an event, below which I consider the leading jet is degenerate and consider both jets as leading, same as in the LO case. The regularization cut-off can be defined in absolute or relative values and the leading jet is considered degenerate if the following condition is met:

$$\begin{aligned} &\text{absolute case} : p_{\mathrm{T}}^{\mathrm{j}\,1} - p_{\mathrm{T}}^{\mathrm{j}\,2} < \Lambda, \\ &\text{relative case} : (p_{\mathrm{T}}^{\mathrm{j}\,1} - p_{\mathrm{T}}^{\mathrm{j}\,2})/p_{\mathrm{T}}^{\mathrm{j}\,1} < \Lambda. \end{aligned} \tag{5.12}$$

It is useful to note that for such cases, when the second leading jet is used to fill histograms, its actual p_{T}, $p_{\mathrm{T}}^{\mathrm{j}\,2}$, is used (although is is slightly smalled than $p_{\mathrm{T}}^{\mathrm{j}\,1}$). This cut-off forces diagrams with large contributions from very soft emissions, and hence passing the regularization condition, to be considered degenerate, same as the opposite large contributions from the virtual loop diagrams; thus, both the first two jets are considered as leading jets and each is assigned a weight equal 1/2 when filling histograms. The final result should not depend on the chosen value of Λ. To test that, I vary the value of Λ and try to find the smallest value that does not bias the final result but also reduces the fluctuation of the original un-regularized prediction. In figure 5.29, I show cross-section results from some of the tested Λ values using the absolute cut-off case, divided by a reference distribution to facilitate the comparison. I use as a reference the distribution with the largest value of Λ since it has the lowest statistical uncertainties; a compatibility with unity is hence not what is to be looked for but rather the closeness of the result to the unregularized case within the fluctuations. Comparing the cross-section results using $\Lambda = 1$ GeV with $\Lambda = 0.001$ GeV (black dots), the bias introduced is up to 10% at the central rapidity bin and up to 20% at the forward rapidity bin. In addition, the orientation of the bias flips from central to forward bins, preserving the total number of entries as expected. It is intriguing the high dependence of the result on the value of Λ, but nevertheless we see the convergence of the result toward the unregularized one the smaller we take Λ, which is a good sign. The effect of the spectrum shift to lower p_{T}, due to the fact that for a degenerate leading jet from regularization, the second leading jet has a smaller p_{T} than the first leading, does not explain the differences between the results of the different Λs. $\Lambda = 0.001$ GeV result looks very close to the unregularized one (black dots compared to green squares), have the least biases and still providing some acceptable amount of regularization. Below this value, the regularization is not enough, and above it the biases are significant. Therefore, for the final results, the regularization cut-off $\Lambda = 0.001$ GeV is used.

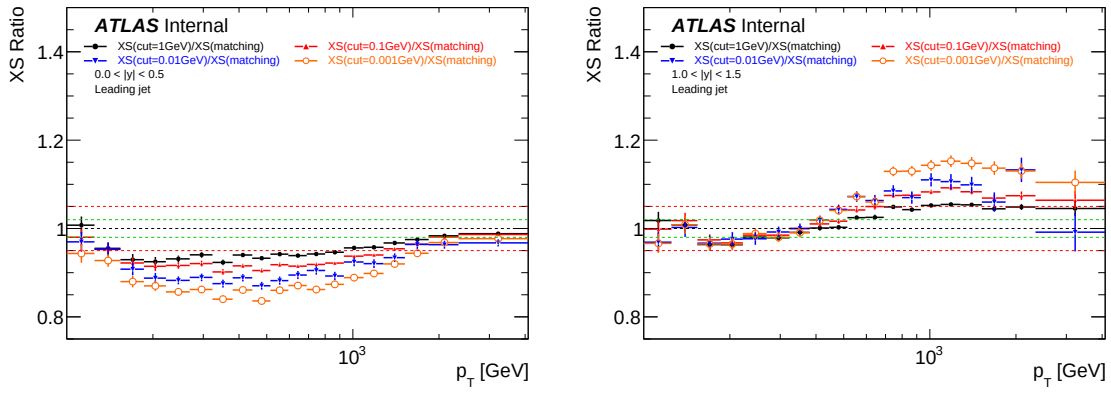


(a) $0.0 < |y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 5.29: Differential cross-section results for the leading jet observable using various cut-off Λ values for regularization, divided by a reference ($\Lambda = 1$ GeV) to facilitate the comparison. In addition, the results using no regularization are also shown. The number of simulated events is 10^{11}, except for the case of $\Lambda = 0.001$ GeV where it is 10^{12}.

I also checked the relative cut-off case which did not provide more stable results and is more aggressive at higher p_{T}, so it is dropped for the time being. The observable high IR sensitivity and high

dependence on the cut-off Λ are currently being investigated deeper by theorists, namely Alexander Huss and his team-mates, checking for example that the logarithmic resummation is behaving well for this observable and that the regularization is not affecting the resummation. Another approach is currently being developed, also by Alexander Huss, where, instead of a cut-off, one matches jets between the virtual loop (or more precisely the counter term) and the real emission diagram and removes the degeneracy from the former by using the p_T order of the latter. Taking the same examples in figure 5.27, when a diagram with a degenerate leading jet is matched to the diagram with a real emission, we consider the leading jet in the former to be the jet "a", same as the latter. The first results of this method are compared to the results using regularization and shown in figure 5.30. The compatibility between the two methods gets worse when using smaller cut-off Λ values, hinting that the matching is introducing fake correlations between jets in different rapidity bins. More investigations are being done. The two methods provide a good strategy for results cross-checking, although the cut-off regularization have an advantage of being more simple to implement specially since we want to eventually generate NNLO predictions where more diagrams with divergences exists and the diagram matching should be done at three different levels (which is not trivial).

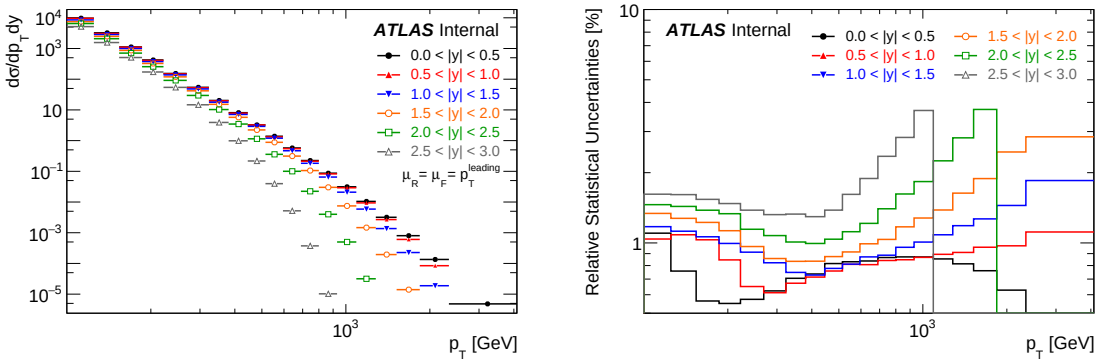


(a) $0.0 < |y| < 0.5$ (b) $1.0 < |y| < 1.5$

Figure 5.30: Ratios of the differential cross-section results for the leading jet observable between various cut-off Λ values and the matching regularization techniques.

Until now, I do not mention the scales that are used. First, the two, factorization and renormalization, scales are generally taken equal to a multiplicative factor to avoid large fluctuation from logarithmic factors of their ratio. In general for inclusive jets, the two widely used scales are p_T of each jet or p_T of the leading jet in the event. In the case of the leading jet observable, the two scales are equal, reducing the ambiguity in making a choice which naturally is p_T^{leading} for this observable. One can also think of including scales related to the whole event, like the simple p_T sum of all jets, H_T, or of all partons, $\hat{H}_T$, but those are not considered for the time being.

The differential cross-sections and the relative statistical uncertainties for the different rapidity bins using $\Lambda = 0.001$ GeV are shown in figure 5.31.



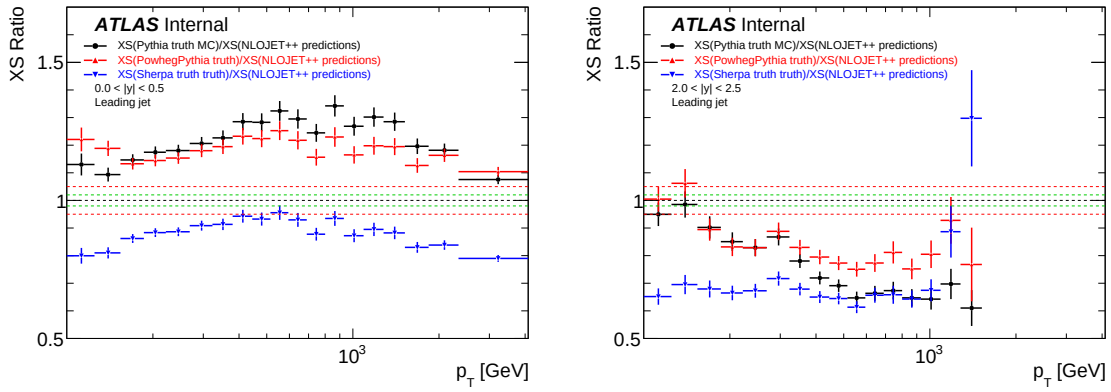
(a) Leading jet - differential cross-section

(b) Leading jet - relative statistical uncertainties

Figure 5.31: Differential cross-sections and relative statistical uncertainties for leading jet observable for the different rapidity bins, using a cut-off regularization $\Lambda = 0.001$ GeV and p_{T} leading scales. The number of simulated events is 10^{12}.

5.6.2 Fixed order vs truth MC simulation

It is useful also to compare the fixed order predictions with the results using truth level MC simulation. In figure 5.32, I show the comparison with three generators, Pythia (LO), PowhegPythia (NLO) and Sherpa (NLO), for two rapidity regions. Differences in the normalization and also in the spectrum shapes are visible and vary between the different rapidity bins. The compatibility of the predictions with LO generator results is not expected. On the other hand, the differences between the predictions and the NLO generator results are larger than one would expect. Part of the difference could be because of still a residual IR sensitivity of the predictions. At the same time, the non-compatibility between the two NLO generators is significant and the theoretical predictions in central rapidity bins lie in between the two NLO simulation results, indicating that the total difference between the predictions and the simulations are not just a bias in the predictions side.



(a) $0.0 < |y| < 0.5$ (b) $2.0 < |y| < 2.5$

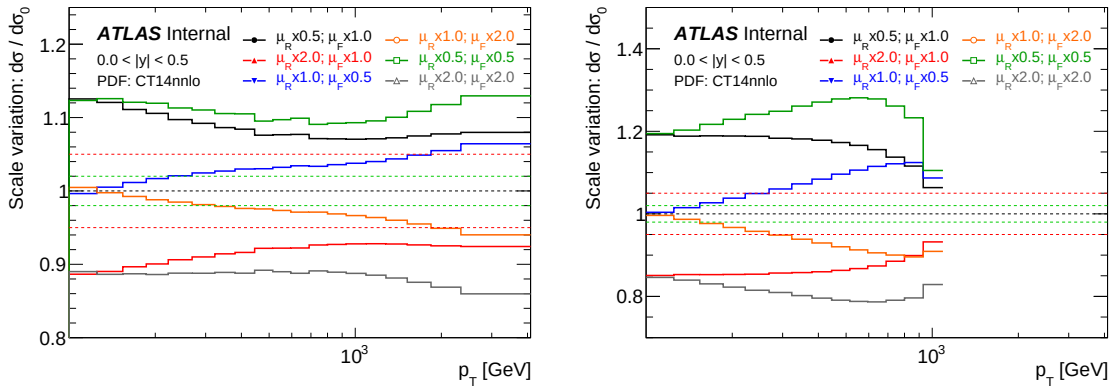
Figure 5.32: Ratios of the leading jet cross-section using truth level MC simulations to the theoretical predictions one for two rapidity bins as a function of p_{T}. Pythia (LO), PowhegPythia (NLO) and Sherpa (NLO) generators are shown in black, red and blue respectively.

5.6.3 Theoretical systematic uncertainties

The three systematic uncertainties considered here on the partonic predictions are coming from missing orders (≡ scales uncertainties) and from the propagation of the reference α_{S} and PDF uncertainties. To evaluate the uncertainties from missing perturbative orders, no exact procedure exists. A convention is widely used to variate the renormalization and factorization scales (independently or together) by factors of two, upward and downward. The envelope of all the variations is taken as the scales uncertainty. This procedure in principle covers the variation of the predictions when including an additional higher perturbative order, except when new production channels are introduced in the added order. In figure 5.33, I show, for all the different scales variation, the ratio of the differential cross-section to the nominal one in the most central and the most forward rapidity bins. The PDF set used is CT14nnlo. For forward rapidity and high p_{T} bins, corresponding to the phase-space limit region, the scales uncertainties decreases. I checked that this behavior is not present in the case of the inclusive jets observable. The last two p_{T} bins of figure 5.33(b), from 800 GeVto 1090 GeV, are combined when comparing afterward with data to have the same binning as data.

The PDF sets provides, in addition to the nominal one, several variated fit functions corresponding to their different uncertainty sources. The variated functions are given in uncorrelated eigenvectors, hence they are considered independently when propagating the uncertainties. The variated cross-sections are calculated using each of the variated PDF, and the uncertainties are added in quadrature. Although usually the variated PDF are given symmetric for upward and downward variations, they are propagated and summed separately since when convoluted with the partonic sub-processes, $\mathrm{d}\hat{\sigma}$, the uncertainties become non-symmetric. One more detail is that for many PDF sets, the variations are given for a C.L. = 90%. In this case, the uncertainties are divided by a factor of 1.645 to shift them back to C.L. = 68% (≡ 1σ variation).

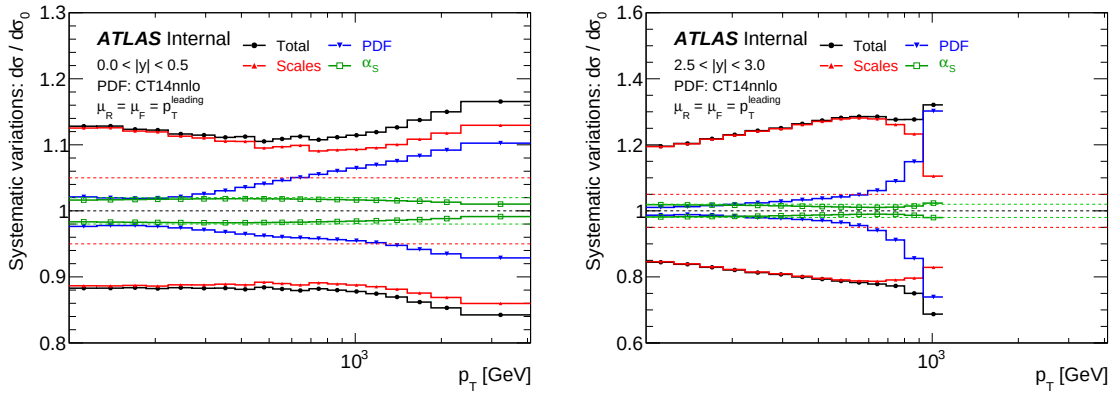
The reference value of α_{S}, usually given at the Z boson mass, is fluctuated by $\pm 1\sigma$ of its full



(a) $0.0 < |y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 5.33: Ratio of the differential cross-sections from the scales variation to the nominal one in the most central and the most forward rapidity bins as a function of p_{T}. The PDF set used is CT14nnlo.

uncertainty. In this thesis, I use the world average from the PDG book [1]: $\alpha_{\mathrm{S}}(\mathrm{M}_{\mathrm{Z}}^2) = 0.118 \pm 0.001$. Needless to say, the value of α_{S} is taken the same in the calculation of the partonic sub-processes, $\mathrm{d}\hat{\sigma}$, and in the PDF fits (PDF fits are given for several α_{S} values). In figure 5.34, I show the systematic deviations of the cross-section for each of the three systematic sources, along with their quadratic sum, for the most central and the most forward bins. It is noticeable that the uncertainties are dominated by the scales one for both rapidity bins.

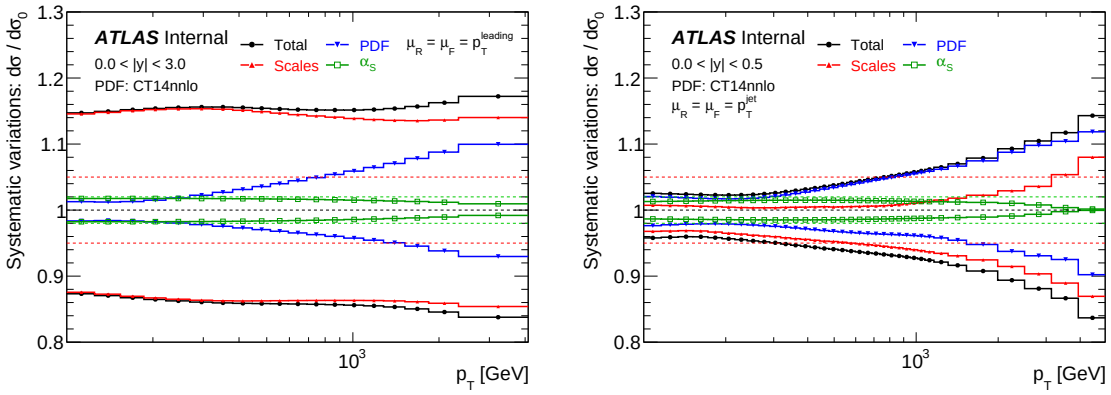


(a) $0.0 < |y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 5.34: Systematic deviations of the double-differential cross-section of the leading jet observable for each of the three systematic sources, scales, PDF and α_{S}, along with their quadratic sum, for the most central and the most forward bins. The PDF set used is CT14nnlo.

For the purpose of comparing with the uncertainties of the leading jet observable, I also show in figure 5.35 the systematic deviations for the leading jet observable inclusive in rapidity, which as said before is not affected by IR effects, and the inclusive jets observable in the central rapidity

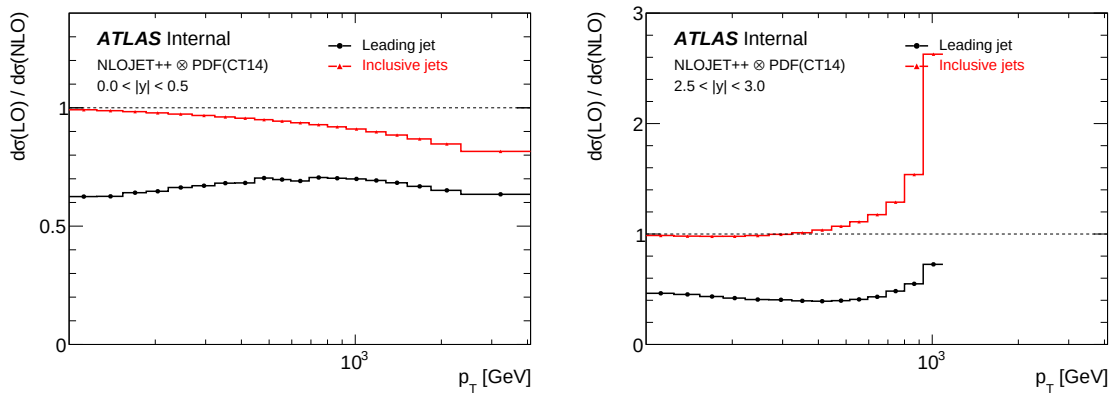
bin. It is interesting to see that the high systematic uncertainties, dominated by the scales one, are also present for the leading jet observable inclusive in rapidity, in contrast with the inclusive jets observable. The dominating scales uncertainties are most likely not related to IR effects but related to the observable itself. It is interesting to see if this behavior comes from the fact that the leading jet observable is much more sensitive to the higher perturbative orders, or if the inclusive jets observable is under-estimating the scales uncertainty due to the non-conservation of physics correlation. Figure 5.36 shows the ratios of the cross-section predictions at LO precision over NLO precision for the leading jet observable and for the inclusive jets observable. For both rapidity bins I show, the ratio for inclusive jets observable is close to unity at low p_{T}, then increases (deceases) at high p_{T} for the central (forward) bin. The behavior of the leading jet observable is very different; for all (p_{T},y) bins, the ratio is significantly lower than unity, ranging roughly between 0.4 and 0.7. The NLO diagram contributions are at the same order of the LO contributions. One would then expect that the missing higher orders still contain important corrections to the predictions, and hence the high scales systematic uncertainties at low p_{T} for the leading jet observable compared to the inclusive jets observable are expected.



(a) Leading jet - 0.0 < |y| < 3.0

(b) Inclusive jets - 0.0 < |y| < 0.5

Figure 5.35: Systematic deviations of the cross-section of the leading jet observable inclusive in rapidity in a) and of the inclusive jets observable in b) for each of the three systematic sources, scales, PDF and α_{S}, along with their quadratic sum, for the most central and most forward bins. The PDF set used is CT14nnlo.



(a) $0.0 < |y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 5.36: Ratio between the cross-section predictions at LO precision over NLO precision for the leading jet observable in black and for the inclusive jets observable in red, for the most central and the most forward rapidity bins as a function of p_T.

5.6.4 Non-perturbative correction factors

As mentioned in chapter 1.3.3, the effects of the two non-perturbative processes, the hadronization and the underlying event, are added to the QCD partonic predictions. As explained, the two methods of including them are through correction factors or through a transfer matrix. Although the latter method has more advantages, for it to be applied for this measurement, the theoretical predictions and the transfer matrix should include the calculations for the first two or three jet orders to take into account the jet order flips between the parton and hadron levels, same as is done for the transfer matrix describing the detector effects. When I tried to calculate the predictions for the sub-leading jet cross-section, it turned out that this observable is much more IR sensitive than the leading jet one. In fact, even with the regularization applied, the cross-section in forward bins is negative. Negative cross-sections for the sub-leading jet observable was also observed and described in the paper [104]. Hence, the use of this observable in the current version of the predictions is not possible. When/if a better regularization of the predictions of the leading jet cross-section is developed, it would then be interesting to calculate again the sub-leading jet cross-section. As for the third jet cross-section, the prediction is stable and not IR sensitive. It is at first counter-intuitive, but the reason for that is straightforward. The only diagrams contributing to the third jet cross-section are the ones with real additional emissions. In addition, since the observable begins at 100 GeV, a natural cut-off exists for the very soft contributions, and hence for the divergent diagrams.

Using the first method, I show in figure 5.37 the NP correction factors in the different rapidity bins as a function of p_T. First, the ratios of the leading jet observable from MC simulation at the levels partons+showering+hadronization+UE over partons+showering are calculated. Next, the NP factors are calculated from the fit results on the ratios to reduce the statistical fluctuations using the following fit function:

$$\mathcal{K}_{NP}(p_T) = 1 + a/p_T^b, \tag{5.13}$$

where a and b are two parameters to be fitted. For the nominal results I show, the MC generator is Pythia with "A14NNPDF" tune. For the systematic uncertainty of this correction, the nominal results are compared with the ones using Pythia with different tunes and the ones using a different generator (Herwig, Sherpa ...) with the various tunes. Since at the time of writing this thesis those additional simulations for the leading jet observable were not ready, the systematic uncertainty evaluation is thus not included.

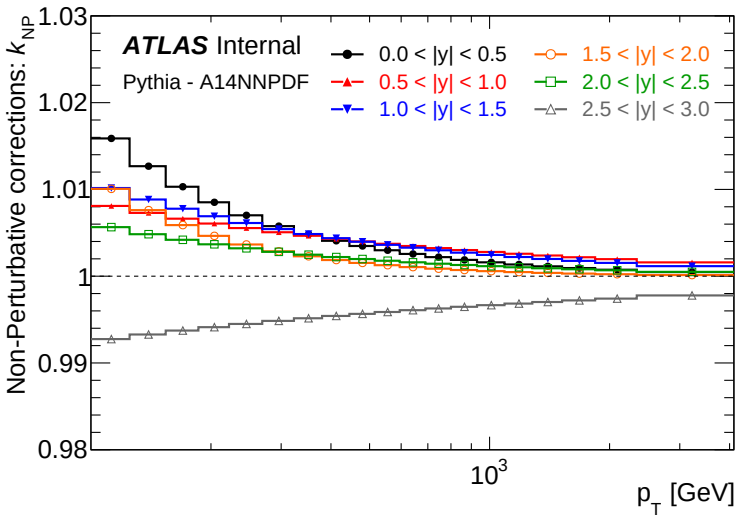


Figure 5.37: The NP correction factors for the leading jet cross-section in the different rapidity bins as a function of p_T, calculated using Pythia generator with "A14NNPDF" tune.

EW correction factors

The EW correction factors (refer to chapter 1.3.4) for the leading jet cross-section are still not available. The same IR sensitivity problem exists in the calculation of these factors, and hence their derivation is postponed to after fully converging on the regularization of the QCD predictions, since the same regularization techniques are to be used. Same as the case for inclusive jets, the corrections are expected to be significant only at high p_T.

5.7 Results

The complete predictions using the QCD partonic cross-sections convoluted with various PDF sets and multiplied by the NP correction factors are compared with the unfolded data in the different rapidity bins as a function of p_T as shown in figure 5.38. The full statistical and systematic uncertainties are shown for the predictions as error bars on the ratios, and the ones of the measurement as a band around unity. In the most central rapidity bin, the ratio is significantly higher than unity at low p_T, but then becomes compatible with unity for the three shown predictions at approximately 400 GeV and stays compatible up to high p_T. The more we go in the forward region, the ratio at the lowest p_T does not change much, but now the ratios in the mid and high p_T becomes less and less compatible with unity. When comparing all (p_T,y) bins, the tension between the predictions and the measurement is clear, specially at low p_T or forward rapidity.

Figure 5.39 shows the comparison between the predictions and the unfolded data for the inclusive in rapidity case, $|y| < 3$. Here also a tension is visible at low p_T. Above about 700 GeV, the predictions using the different PDF sets and the data are compatible. Let me note that for a $p_T > 650$ GeV, all jets have $|y| < 3$ due to kinematic limits (and considering a massless jet). It would be useful to test if expanding the rapidity selection would improve the agreement at low p_T. For that, the forward triggers should also be used and combined with the central ones using the inclusion method introduced in chapter 3.5.

The comparison between the LO precision predictions and the unfolded data is shown in figure 5.40 for the different rapidity bins as a function of p_T. The compatibility in this case is much better, where the deviation of the prediction/data ratio from unity is covered by the total uncertainties in most bins. It is also noticeable that for the majority of bins the ratio is smaller than unity, whereas for the NLO precision prediction case it is higher. It would be very interesting to see if the comparison with the NNLO precision predictions, once they are available, would push the ratio even more upward or if it would flip back closer to unity.

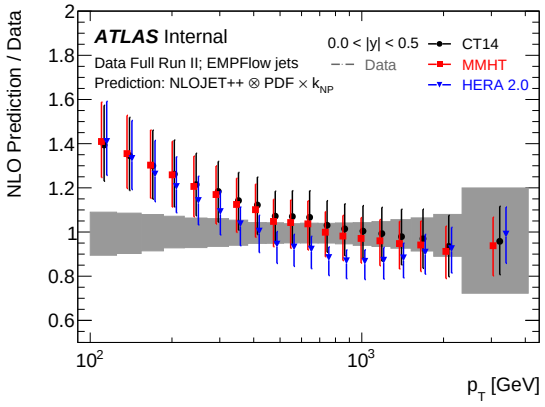
It is important to note that the current prescription to evaluate the scale uncertainties does not seem to cover the missing higher orders for the leading jet observable predictions. Figure 5.41 shows the LO cross-section ratios for the different scales variations to the nominal case. In the most central rapidity bin, the maximum factors range from 1.25 to 1.35 for the different p_T bins, whereas for the most forward rapidity bin they range from 1.25 to 1.5. On the other hand, from figure 5.36, the LO scale uncertainty factors needed to cover the difference between LO and NLO precision predictions are about 1.54 (1/0.6) in the most central rapidity bin, and about 2 (1/0.5) to 2.5 (1/0.4) in the most forward one. This means that in some bins, specially at low p_T, the LO scale uncertainties should be more than double to cover this difference. Usually, an under-coverage of the scale uncertainties is observed when new production channels are introduced in the higher order, which is not the case here. It is interesting to check if this may come from the fact that for LO diagrams, the leading jet observable is degenerate, whereas for NLO diagrams with real emission outside the jet cone, it is not anymore. For that, we can derive the theoretical predictions using a wider jet radius and then check if the coverage of the scale uncertainties is improved. In addition, it would be again also important to derive the NNLO precision predictions and check if the NLO scale uncertainties cover the difference between the NLO and NNLO precision predictions.

Another check I did is to compare the leading jet observable predictions to the inclusive jets ones, which is shown in figure 5.42. At LO precision (the plot on the left), the ratio is 0.5 at hight p_{T} and slightly higher at low p_{T} as one would expect. The factor 0.5 comes from the fact that for LO diagrams, the two same-p_{T} parton are reconstructed as two same-p_{T} jets, and hence we have in general two entries per event in the same p_{T} bin for the inclusive jet observable, versus one entry per event for the leading jet observable. The only exception is if one of the two jets is outside the rapidity selection cuts, then we get only one entry per event also for the inclusive jet observable. At NLO precision, we expect the ratio to be larger, since for NLO diagrams with real emission outside the jet cone, the second and third jets enter lower bins where the total entries are dominated by the leading jet due to the steeply falling p_{T} spectra. What we observe from the right plot of figure 5.42 is that the ratio increase significantly, becoming higher than unity at the forward region which is unrealistic. In fact, the higher-than-unity ratios come from the fact that the sub-leading observable has negative weights at the forward region as described in section 5.6.4. Also, this is consistent with what is observed in figure 5.38 where the prediction at NLO precision is significantly larger than the unfolded data for the leading jet observable. Again, it would be useful to do this comparison once we have also the NNLO precision predictions, since, as already found in the paper [104], negative cross-sections for the sub-leading jet is absent at NNLO precision predictions.

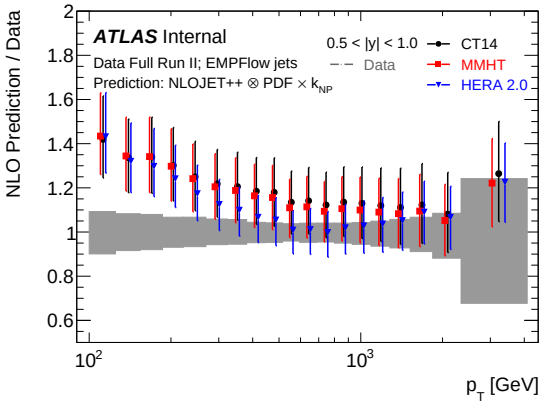
Last, I show the comparison between MC truth distributions and unfolded data in figure 5.43. For the MC distributions, only the statistical uncertainties are considered. Pythia and Powheg+Pythia generators give distributions significantly larger than the unfolded data in all (p_{T},y) bins, ranging from 20% to 60%. For Pythia generator, a good agreement is not expected in the first place since it uses a LO matrix element generator. For Powheg+Pythia generator, the level of disagreement is somehow surprising, but this is the older version of the generator which was also found not to describe well the third jet distribution as already detailed in chapter 3.3. On the other hand, Sherpa generator gives distributions with a good compatibility in the central regions or at low p_{T}; tensions exists only for forward and mid to high p_{T} bins.

In summary, despite the important progress on the theoretical predictions calculation with an extensive collaboration with theorists, tensions between the NLO predictions and the measurement are visible as observed in figure 5.38. Nevertheless, several points for improving the predictions still exist which can reduce those tensions. The following list is a proposal for improvement:

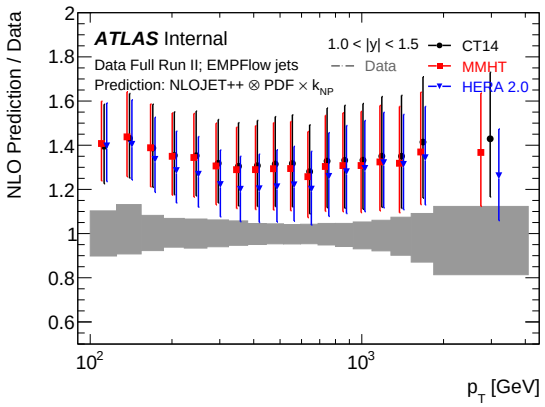
1. compute the NNLO precision predictions using the same regularization method; all the previous test should be re-done to better understand the source of the tensions if still existing,
2. identify if there still exists a residual IR sensitivity of the prediction (counter terms, resummation ...) and better regularize it,
3. with better regularization and NNLO predictions, the sub-leading jet cross-section can also be computed and the non-perturbative corrections can be applied using a transfer matrix; a check with the $\mathcal{K}_{\mathrm{NP}}$ factors method is to be done to verify if the latter introduced a bias,
4. the EW corrections can now be computed applying the same regularization, since the IR sensitivity is also present here,
5. possibly extend the rapidity selection to more forward regions,
6. test the prediction behavior using a wider jet radius.



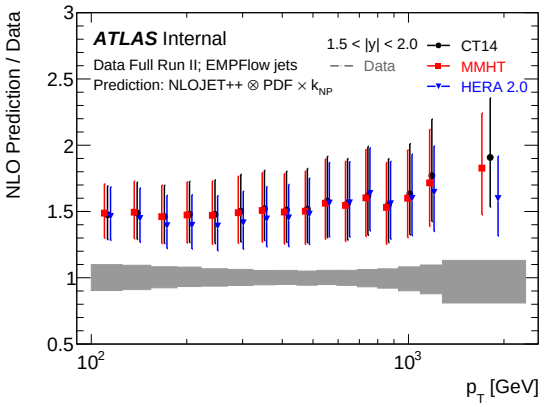
(a) $0.0 < |y| < 0.5$



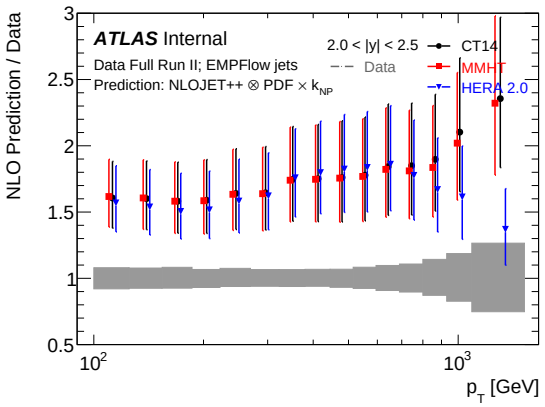
(b) $0.5 < |y| < 1.0$



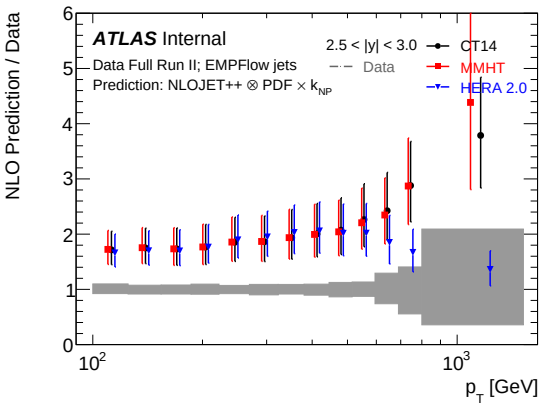
(c) $1.0 < |y| < 1.5$



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



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



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

Figure 5.38: Ratios of the complete theoretical predictions, using the NLO QCD partonic cross-sections convoluted with various PDF sets and multiplied by the NP correction factors, over the unfolded data for the leading jet observable in the different rapidity bins as a function of p_{T}. Three PDF sets are used: CT14 (black), MMHT (red) and HERA2.0 (blue). The full statistical and systematic uncertainties are shown for the predictions as error bars on the ratios, and the ones of the measurement as a gray band around unity.

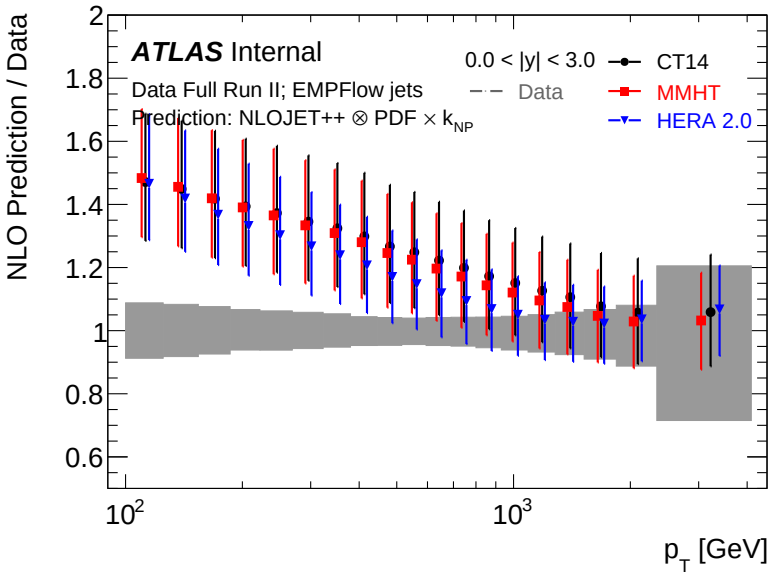
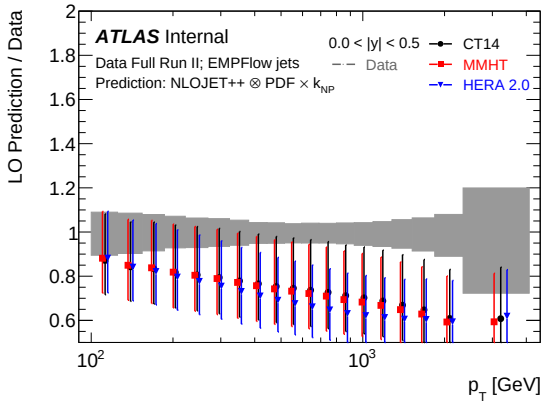
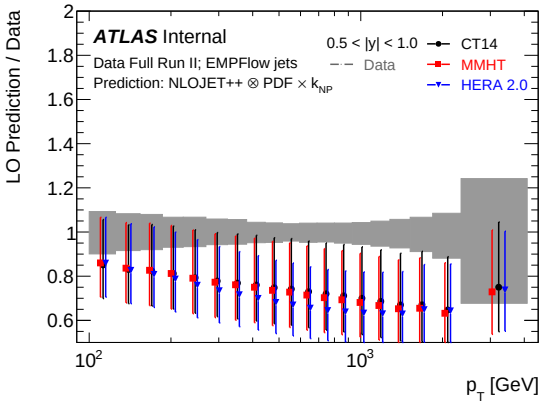


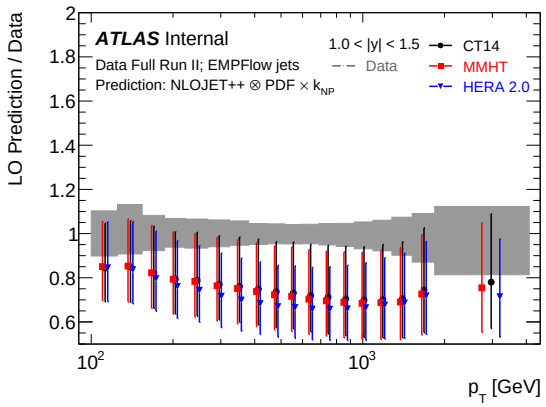
Figure 5.39: Ratios of the complete theoretical predictions, using the NLO QCD partonic cross-sections convoluted with various PDF sets and multiplied by the NP correction factors, over the unfolded data for the leading jet observable inclusive in rapidity as a function of p_{T}. Three PDF sets are used: CT14 (black), MMHT (red) and HERA2.0 (blue). The full statistical and systematic uncertainties are shown for the predictions as error bars on the ratios, and the ones of the measurement as a gray band around unity.



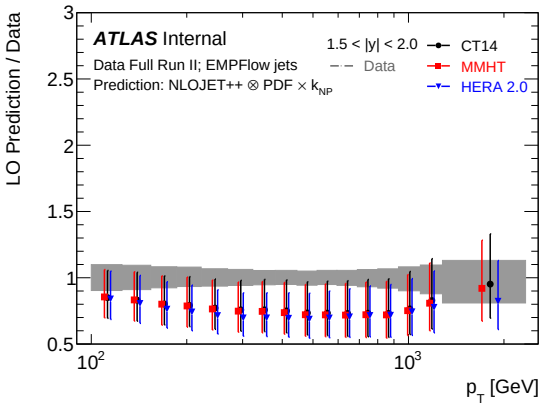
(a) $0.0 < |y| < 0.5$



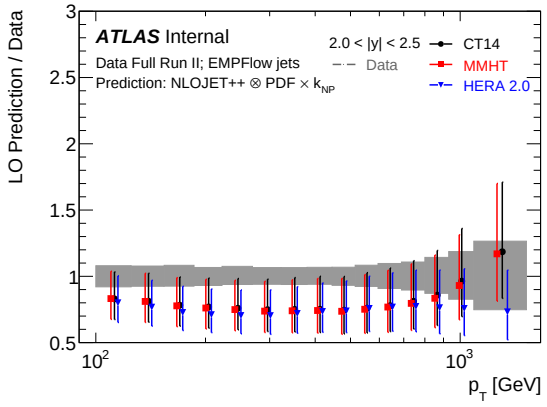
(b) $0.5 < |y| < 1.0$



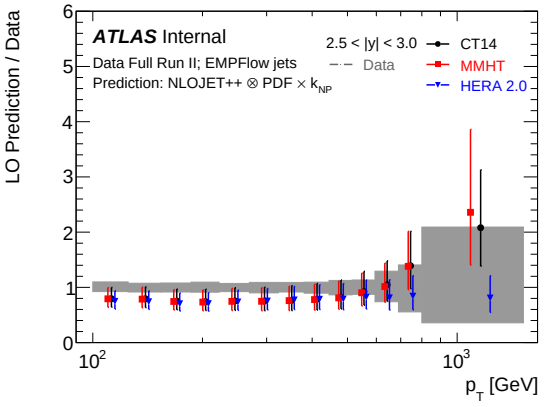
(c) $1.0 < |y| < 1.5$



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

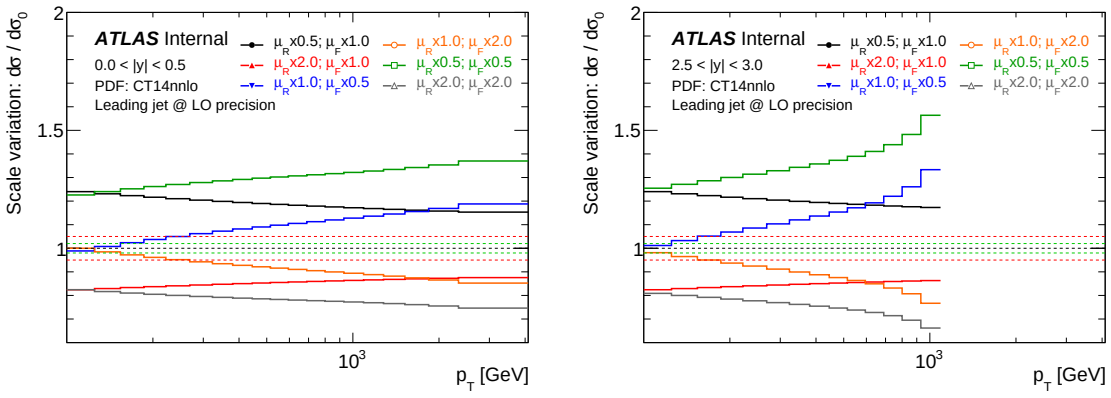


(e) $2.0 < |y| < 2.5$



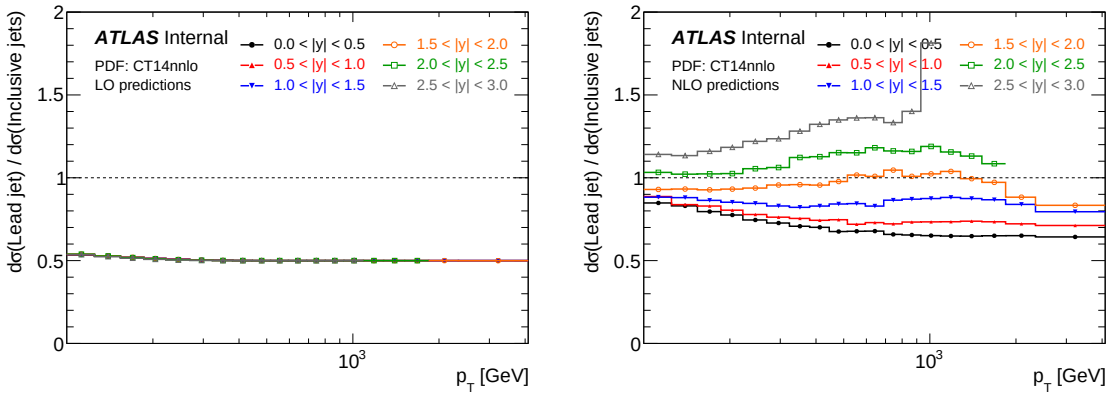
(f) $2.5 < |y| < 3.0$

Figure 5.40: Ratios of the complete theoretical predictions, using the LO QCD partonic cross-sections convoluted with various PDF sets and multiplied by the NP correction factors, over the unfolded data for the leading jet observable in the different rapidity bins as a function of p_{T}. Three PDF sets are used: CT14 (black), MMHT (red) and HERA2.0 (blue). The full statistical and systematic uncertainties are shown for the predictions as error bars on the ratios, and the ones of the measurement as a gray band around unity.



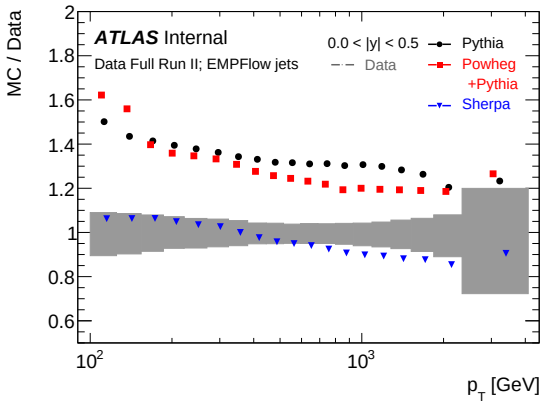
(a) $0.0 < |y| < 0.5$ (b) $2.5 < |y| < 3.0$

Figure 5.41: Ratio of the LO partonic cross-sections from the scales variation to the nominal one in the most central and the most forward rapidity bins as a function of p_{T}. The PDF set used is CT14nnlo.

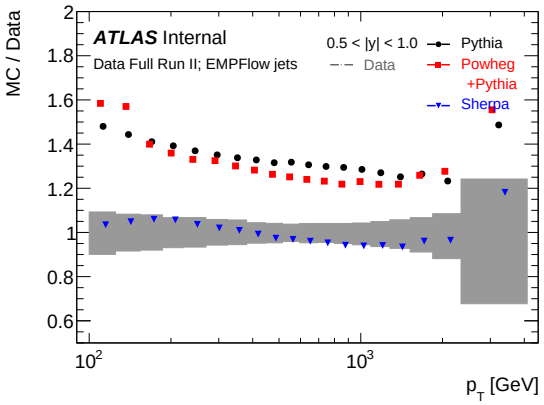


(a) LO precision (b) NLO precision

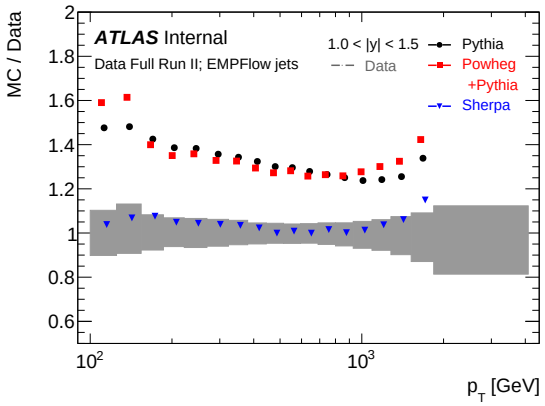
Figure 5.42: Ratio between the QCD partonic cross-sections of the leading jet observable over the inclusive jets observable in the different rapidity bins as a function of p_{T}, at LO prediction precision on the left and NLO one on the right.



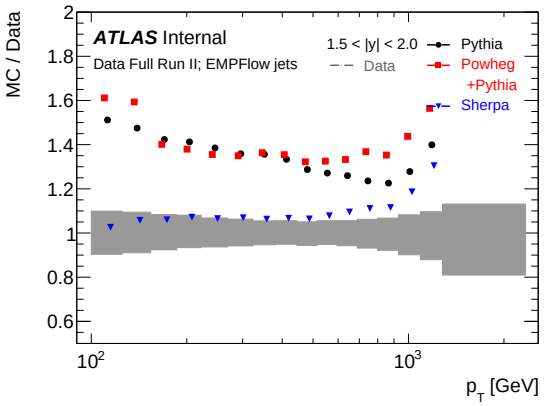
(a) $0.0 < |y| < 0.5$



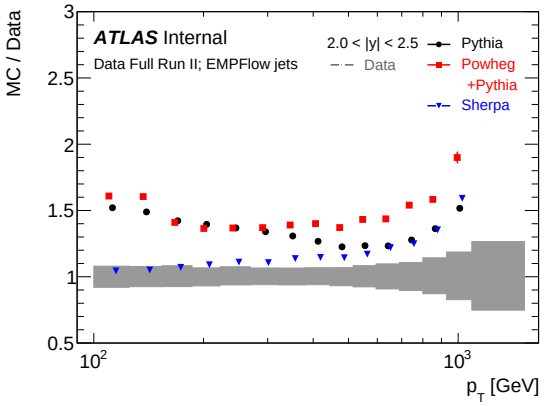
(b) $0.5 < |y| < 1.0$



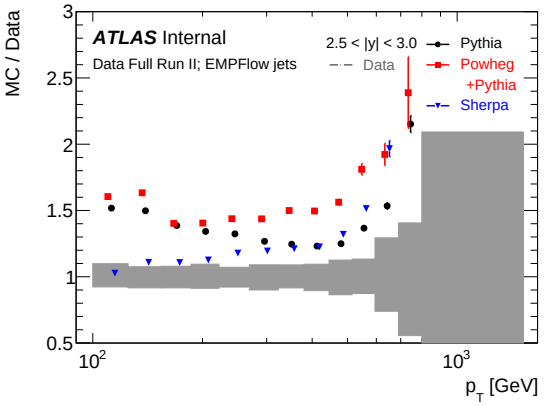
(c) $1.0 < |y| < 1.5$



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



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



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

Figure 5.43: Ratios of the truth MC over the unfolded data for the leading jet observable in the different rapidity bins as a function of p_{T}. Three MC generators are used: Pythia (black), Powheg+Pythia (red) and Sherpa (blue). The full statistical and systematic uncertainties of the measurement are shown as a gray band around unity. For the MC distributions, only the statistical uncertainties are shown as error bars.

Conclusion

This manuscript summarizes my most important work during the three years of PhD.
I developed and implemented an important improvement for the eta-intercalibration, a jet in-situ calibration method aiming at calibrating jets in the forward regions of the detector relative to the jets in the central region, making it much faster using an analytic solution. This allows to use finer η binning and hence improving the description of the peaks in the jet response as well as the closure of the method. MC generators are compared with data to check their level of compatibility and to verify the choice of the generators used in the calibration. In addition, the robustness of the calibration with respect to the pile-up conditions is verified. A better trigger combination method is tested and implemented, replacing the biased older one. After the implementation of all those improvement, I derived the nominal values of the eta-intercalibration correction and the full uncertainties that are used as part of the final Run II jet calibration. Currently, a paper describing all the jet calibration methods used in Run II is being circulated within the collaboration.

On the physics analyses side, direct searches for new phenomena are very important analyses that look for Beyond Standard Model signals. When no significant signal is observed, as it is currently the case, exclusion limits are put on benchmark Beyond Standard Model signals or on generic shape signals. The aim of the latter is to extend the re-interpretation of the results where new theoretical models are compared with the generic signals and limits are evaluated on those models. Previously, the limits on the generic signals were set at detector level, which includes the detector resolution, and hence complicated the usability of those limits. I developed and implemented a folding technique to be used when calculating those limits which are now evaluated at particle level, facilitating the comparison with theoretical models and also the combination of limits obtained in different studies. The method has also additional potential applications, like for example the interpolation between various parameter points for which the full simulation samples were generated. The analysis, including the new folding technique, is already published in the paper [76].

Jet cross-section measurements are the other very important analyses used to test the Standard Model and for indirect searches of Beyond Standard Model contributions. Various observables were and are still used to perform those analyses. A new observable was proposed to be used for the first time: the leading jet cross-section. I had the chance to work on all the aspects of this new analysis, on both data measurement and theoretical predictions development and implementation. From the data selection and quality checks, to the definition of the transfer matrix, the methods previously used had to be re-checked and some new ones to be developed. The observable, although simple in its definition, is much more complex to be measured, compared to inclusive jets or dijet measurements, due to the flips of jet orders between parton, hadron and detector reconstruction levels.
The theoretical predictions are very challenging to be calculated. The IR sensitivity of the observable makes the implementation of regularizations necessary. Multiple checks are done, in addition to comparisons with the predictions of other observables. The full statistical and systematic uncertainties

are calculated for both the measurement and the predictions. The final comparison between the theoretical predictions and the measurement shows some tension in various regions. The tensions are a challenge to improve the study, mainly on the theoretical predictions side.

Bibliography

[1] M. T. et al., *Review of Particle Physics*, Phys. Rev. D **98** (3 2018) 030001, URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[2] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Reading, USA: Addison-Wesley, 1995, ISBN: 9780201503975, 0201503972, URL: http://www.slac.stanford.edu/~mpeskin/QFT.html.

[3] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22**.4 (1961) 579 –588, ISSN: 0029-5582, URL: http://www.sciencedirect.com/science/article/pii/0029558261904692.

[4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (21 1967) 1264–1266, URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[5] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367–377.

[6] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (26 1973) 1343–1346, URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1343.

[7] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?* Phys. Rev. Lett. **30** (26 1973) 1346–1349, URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1346.

[8] H. Fritzsch and M. Gell-Mann, *Current algebra: Quarks and what else?* eConf **C720906V2** (1972) 135–165, arXiv: hep-ph/0208010 [hep-ph].

[9] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the color octet gluon picture*, Physics Letters B **47**.4 (1973) 365 –368, ISSN: 0370-2693, URL: http://www.sciencedirect.com/science/article/pii/0370269373906254.

[10] P. Higgs, *Broken symmetries, massless particles and gauge fields*, Physics Letters **12**.2 (1964) 132 –133, ISSN: 0031-9163, URL: http://www.sciencedirect.com/science/article/pii/0031916364911369.

[11] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16 1964) 508–509, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[12] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9 1964) 321–323, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[13] G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1–29, arXiv: 1207.7214 [hep-ex].

[14] S. Chatrchyan et al.,
Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC,
Phys. Lett. **B716** (2012) 30–61, arXiv: 1207.7235 [hep-ex].

[15] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*,
Phys. Rev. Lett. **10** (12 1963) 531–533,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[16] M. Kobayashi and T. Maskawa,
CP-Violation in the Renormalizable Theory of Weak Interaction,
Progress of Theoretical Physics **49**.2 (1973) 652–657, ISSN: 0033-068X, eprint:
http://oup.prod.sis.lan/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf,
URL: https://doi.org/10.1143/PTP.49.652.

[17] K. G. Wilson, *Confinement of quarks*, Phys. Rev. D **10** (8 1974) 2445–2459,
URL: https://link.aps.org/doi/10.1103/PhysRevD.10.2445.

[18] G. P. Salam and G. Soyez, *A Practical Seedless Infrared-Safe Cone jet algorithm*,
JHEP **05** (2007) 086, arXiv: 0704.0292 [hep-ph].

[19] M. Cacciari, *Phenomenological and theoretical developments in jet physics at the LHC*,
Int. J. Mod. Phys. **A30**.31 (2015) 1546001, arXiv: 1509.02272 [hep-ph].

[20] S. Catani et al.,
Longitudinally invariant K_t clustering algorithms for hadron hadron collisions,
Nucl. Phys. **B406** (1993) 187–224.

[21] Y. L. Dokshitzer et al., *Better jet clustering algorithms*, JHEP **08** (1997) 001,
arXiv: hep-ph/9707323 [hep-ph].

[22] M. Cacciari, G. Salam, and G. Soyez, *The anti-k_T jet clustering algorithm*,
JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[23] J. Rojo et al., *The PDF4LHC report on PDFs and LHC data: Results from Run I and preparation for Run II*, J. Phys. **G42** (2015) 103103, arXiv: 1507.00556 [hep-ph].

[24] K. P. O. Diener, S. Dittmaier, and W. Hollik, *Electroweak higher-order effects and theoretical uncertainties in deep-inelastic neutrino scattering*,
Phys. Rev. **D72** (2005) 093002, arXiv: hep-ph/0509084 [hep-ph].

[25] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*,
Nuclear Physics B **126**.2 (1977) 298 –318, ISSN: 0550-3213,
URL: http://www.sciencedirect.com/science/article/pii/0550321377903844.

[26] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*,
Sov. Phys. JETP **46** (1977) 641–653, [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[27] V. N. Gribov and L. N. Lipatov, *Deep inelastic e p scattering in perturbation theory*,
Sov. J. Nucl. Phys. **15** (1972) 438–450, [Yad. Fiz.15,781(1972)].

[28] S. Höche, "Introduction to parton-shower event generators," *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, 2015 235–295, arXiv: 1411.4085 [hep-ph].

[29] T. Kinoshita, *Mass singularities of Feynman amplitudes*, J. Math. Phys. **3** (1962) 650–677.

[30] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, Phys. Rev. **133** (6B 1964) B1549–B1562, URL: https://link.aps.org/doi/10.1103/PhysRev.133.B1549.

[31] R. Bonciani et al., *Sudakov resummation of multiparton QCD cross-sections*, Phys. Lett. **B575** (2003) 268–278, arXiv: hep-ph/0307035 [hep-ph].

[32] Z. Nagy, *Three-jet cross sections in hadron-hadron collisions at next-to-leading-order*, Phys. Rev. Lett. **88** (2002) 122003, arXiv: hep-ph/0110315 [hep-ph].

[33] Z. Nagy, *Next-to-leading order calculation of three-jet observables in hadron-hadron collision*, Phys. Rev. **D68** (2003) 094002, arXiv: hep-ph/0307268 [hep-ph].

[34] S. Catani and M. H. Seymour, *A general algorithm for calculating jet cross-sections in NLO QCD*, Nucl. Phys. **B485** (1997) 291–419, arXiv: hep-ph/9605323 [hep-ph].

[35] A. Gehrmann-De Ridder et al., *Precise QCD predictions for the production of a Z boson in association with a hadronic jet*, Phys. Rev. Lett. **117**.2 (2016) 022001, arXiv: 1507.02850 [hep-ph].

[36] J. Currie et al., *Single Jet Inclusive Production for the Individual Jet p_T Scale Choice at the LHC*, Acta Phys. Polon. **B48** (2017) 955–967, arXiv: 1704.00923 [hep-ph].

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

[38] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159–177, arXiv: 1410.3012 [hep-ph].

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

[40] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639–707, arXiv: 0803.0883 [hep-ph].

[41] V. V. Sudakov, *Vertex parts at very high-energies in quantum electrodynamics*, Sov. Phys. JETP **3** (1956) 65–71, [Zh. Eksp. Teor. Fiz.30,87(1956)].

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

[43] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[44] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043, arXiv: 1002.2581 [hep-ph].

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

[46] R. Frederix et al., *The complete NLO corrections to dijet hadroproduction*, JHEP **04** (2017) 076, arXiv: 1612.06548 [hep-ph].

[47] V. Rubin, F. W.K. Jr, and N Thonnard, *Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605/R = 4 kpc/ to UGC 2885/R =4 kpc/*, Astrophysical Journal **238** (1980) 471–487.

[48] A. G. Riess et al., *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, Astron. J. **116** (1998) 1009–1038, arXiv: astro-ph/9805201 [astro-ph].

[49] H. Georgi and S. L. Glashow, *Unity of All Elementary Particle Forces*, Phys. Rev. Lett. **32** (1974) 438–441.

[50] U. Baur, I. Hinchliffe, and D. Zeppenfeld, *Excited quark production at hadron colliders*, Int. J. Mod. Phys. **A2** (1987) 1285.

[51] P. Meade and L. Randall, *Black holes and quantum gravity at the LHC*, JHEP05 **2008** (2008) 003, arXiv: 0708.3017 [hep-ph].

[52] G. Altarelli, B. Mele, and M. Ruiz-Altaba, *Searching for New Heavy Vector Bosons in $p\bar{p}$ Colliders*, Z. Phys. **C45** (1989) 109, [Erratum: Z. Phys.C47,676(1990)].

[53] M. Chala et al., *Constraining Dark Sectors with Monojets and Dijets*, JHEP **07** (2015) 089, arXiv: 1503.05916 [hep-ph].

[54] M. V. Chizhov, V. A. Bednyakov, and J. A. Budagov, *A unique signal of excited bosons in dijet data from pp-collisions*, Phys. Atom. Nucl. **75** (2012) 90–96, arXiv: 1010.2648 [hep-ph].

[55] E. Eichten et al., *Supercollider Physics*, Rev. Mod. Phys. **56** (1984) 579–707.

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

[57] H. Bohr and H. Nielsen, *Hadron production from a boiling quark soup: A thermodynamical quark model predicting particle ratios in hadronic collisions*, Nuclear Physics B **128**.2 (1977) 275 –293, ISSN: 0550-3213, URL: http://www.sciencedirect.com/science/article/pii/0550321377900323.

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

[59] J. Pequenao and P. Schaffner, "How ATLAS detects particles: diagram of particle paths in the detector," 2013, URL: https://cds.cern.ch/record/1505342.

[60] ATLAS Collaboration, "Luminosity results in Run 2," URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[61] M. Aaboud et al., *Performance of the ATLAS Trigger System in 2015*, Eur. Phys. J. **C77**.5 (2017) 317, arXiv: 1611.09661 [hep-ex].

[62] M. Aaboud et al., *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. **C77**.7 (2017) 466, arXiv: 1703.10485 [hep-ex].

[63] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, ATLAS-CONF-2014-018 (2014), URL: https://cds.cern.ch/record/1700870.

[64] ATLAS Collaboration, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, ATLAS-CONF-2015-029 (2015), URL: https://cds.cern.ch/record/2037702.

[65] G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Eur. Phys. J. **C73**.3 (2013) 2304, arXiv: 1112.6426 [hep-ex].

[66] G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 17, arXiv: 1406.0076 [hep-ex].

[67] M. Aaboud et al., *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. **D96**.7 (2017) 072002, arXiv: 1703.09665 [hep-ex].

[68] ATLAS Collaboration, "Jet energy scale and uncertainties in 2015-2017 data and simulation," 2018, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2018-006/.

[69] G. Aad et al., *Jet energy resolution in proton-proton collisions at* $\sqrt{s} = 7$ *TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73**.3 (2013) 2306, arXiv: 1210.6210 [hep-ex].

[70] ATLAS Collaboration, *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS* (2015).

[71] ATLAS Collaboration, "Jet energy resolution in 2017 data and simulation," 2018, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/JETM-2018-005/.

[72] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343–367.

[73] G. Bohm and G. Zech, *Introduction to statistics and data analysis for physicists*, Hamburg: DESY, 2010 447 p, ISBN: 9783935702416, URL: http://bib-pubdb1.desy.de/record/339910.

[74] G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics* (2009), arXiv: 0901.0512 [hep-ex].

[75] V. Lendermann et al., *Combining Triggers in HEP Data Analysis*, Nucl. Instrum. Meth. **A604** (2009) 707–718, arXiv: 0901.4118 [hep-ex].

[76] M. Aaboud et al., *Search for new phenomena in dijet events using 37* fb^{-1} *of pp collision data collected at* $\sqrt{s}$ *=13 TeV with the ATLAS detector*, Phys. Rev. **D96**.5 (2017) 052004, arXiv: 1703.09127 [hep-ex].

[77] G. Aad et al., *Search for New Particles in Two-Jet Final States in 7 TeV Proton-Proton Collisions with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **105** (2010) 161801, arXiv: 1008.2461 [hep-ex].

[78] G. Cowan, *Statistical data analysis*, 1998, ISBN: 9780198501565.

[79] K. Cranmer, "Practical Statistics for the LHC," *Proceedings, 2011 European School of High-Energy Physics (ESHEP 2011): Cheile Gradistei, Romania, September 7-20, 2011*, [,247(2015)], 2015 267–308, arXiv: 1503.07622 [physics.data-an].

[80] T. A. collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector at the LHC* (2019), URL: https://cds.cern.ch/record/2677054.

[81] CDF Collaboration, T. Aaltonen et al., *Global search for new physics with 2.0 fb^{-1} at CDF*, Phys. Rev. **D79** (2009) 011101, arXiv: 0809.3781 [hep-ex].

[82] G. Choudalakis, *On hypothesis testing, trials factor, hypertests and the BumpHunter* (2011), arXiv: 1101.0390 [physics.data-an].

[83] L. Lyons, *Open statistical issues in particle physics*, Ann. Appl. Stat. **2** (2008) 887–915.

[84] E. Gross and O. Vitells, *Trial factors or the look elsewhere effect in high energy physics*, Eur. Phys. J. **C70** (2010) 525–530, arXiv: 1005.1891 [physics.data-an].

[85] ATLAS Collaboration, *ATLAS search for new phenomena in dijet mass and angular distributions using pp collisions at $\sqrt{s}$ = 7 TeV*, JHEP **1301** (2013) 029, arXiv: 1210.1718 [hep-ex].

[86] ATLAS Collaboration, *Search for new phenomena in the dijet mass distribution using pp collision data at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, Phys. Rev. D **91** (2015) 052007, arXiv: 1407.1376 [hep-ex].

[87] G. Aad et al., *Search for new phenomena in dijet mass and angular distributions from pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Lett. **B754** (2016) 302–322, arXiv: 1512.01530 [hep-ex].

[88] G. Bohm and G. Zech, *Introduction to statistics and data analysis for physicists*, Hamburg: Verl. Dt. Elektronen-Synchrotron, 2010 336–340, ISBN: 978-3-935702-41-6.

[89] G. Aad et al., *Measurement of the inclusive jet cross-section in proton-proton collisions at $\sqrt{s}$ = 7 TeV using 4.5 fb^{-1} of data with the ATLAS detector*, JHEP **02** (2015) 153, [Erratum: JHEP09,141(2015)], arXiv: 1410.8857 [hep-ex].

[90] M. Aaboud et al., *Measurement of the inclusive jet cross-sections in proton-proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, JHEP **09** (2017) 020, arXiv: 1706.03192 [hep-ex].

[91] M. Aaboud et al., *Measurement of inclusive jet and dijet cross-sections in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **05** (2018) 195, arXiv: 1711.02692 [hep-ex].

[92] S. Chatrchyan et al., *Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s}$ = 7 TeV*, Phys. Rev. Lett. **107** (2011) 132001, arXiv: 1106.0208 [hep-ex].

[93] V. Khachatryan et al., *Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s}$ = 8 TeV and cross section ratios to 2.76 and 7 TeV*, JHEP **03** (2017) 156, arXiv: 1609.05331 [hep-ex].

[94] V. Khachatryan et al., *Measurement of the double-differential inclusive jet cross section in proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. **C76**.8 (2016) 451, arXiv: 1605.04436 [hep-ex].

[95] G. Aad et al., *Measurement of dijet cross sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector*, JHEP **05** (2014) 059, arXiv: 1312.3524 [hep-ex].

[96] S. Chatrchyan et al., *Measurement of the differential dijet production cross section in proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, Phys. Lett. **B700** (2011) 187–206, arXiv: 1104.1693 [hep-ex].

[97] A. M. Sirunyan et al., *Measurement of the triple-differential dijet cross section in proton-proton collisions at* $\sqrt{s} = 8$ *TeV and constraints on parton distribution functions*, Eur. Phys. J. **C77**.11 (2017) 746, arXiv: 1705.02628 [hep-ex].

[98] V. Khachatryan et al., *Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at* $\sqrt{s} = 7$ *TeV*, Eur. Phys. J. **C75**.6 (2015) 288, arXiv: 1410.6765 [hep-ex].

[99] B. Malaescu, *An Iterative, dynamically stabilized method of data unfolding* (2009), arXiv: 0907.3791 [physics.data-an].

[100] B. Malaescu, "An Iterative, Dynamically Stabilized(IDS) Method of Data Unfolding," *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, 2011 271–275, arXiv: 1106.3107 [physics.data-an].

[101] T. Carli et al., *A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project*, Eur. Phys. J. C 66 (2010) **66** (2010) 503–524, arXiv: 0911.2985 [hep-ex].

[102] A. Buckley et al., *LHAPDF6: parton density access in the LHC precision era*, Eur. Phys. J. **C75** (2015) 132, arXiv: 1412.7420 [hep-ph].

[103] G. P. Salam and J. Rojo, *A Higher Order Perturbative Parton Evolution Toolkit (HOPPET)*, Comput. Phys. Commun. **180** (2009) 120–156, arXiv: 0804.3755 [hep-ph].

[104] J. Currie et al., *Infrared sensitivity of single jet inclusive production at hadron colliders*, JHEP **10**.arXiv:1807.03692 (2018) 155. 42 p, 42 pages, 22 figures (version published in JHEP), URL: https://cds.cern.ch/record/2630405.