Probing Quantum Chromodynamics with the ATLAS Detector: Charged-Particle Event Shape Variables and the Dijet Cross-Section

Dissertation
zur Erlangung des Grades „Doktor der Naturwissenschaften“
am Fachbereich Physik, Mathematik und Informatik
der Johannes Gutenberg-Universität in Mainz

Tobias Hülsing
geboren in New Haven, CT

22.09.2014

Institut für Physik
Staudingerweg 7, 55128 Mainz, Deutschland
Johannes Gutenberg-Universität Mainz

Erstgutachter: Aus Datenschutzgründen entfernt.
Zweitgutachter: Aus Datenschutzgründen entfernt.

Tobias Hülsing: *Probing Quantum Chromodynamics with the ATLAS Detector: Charged-Particle Event Shape Variables and the Dijet Cross-Section*, PhD Thesis © 22.09.2014

BETREUER:
Aus Datenschutzgründen entfernt.

LOCATION:
Mainz

TAG DER MÜNDLICHE PRÜFUNG:
22.01.2015

ABSTRACT

Quantum chromodynamics, *QCD*, the theory of the strong interaction is split into two regimes. Scattering processes of the proton constituents, the partons, with a high momentum transfer Q^2 can be calculated and predicted with perturbative calculations. At low momentum transfers between the scattering particles perturbation theory is not applicable anymore, and phenomenological methods are used to describe the physics in this regime. The ATLAS experiment at the Large Hadron Collider, *LHC*, provides the possibility to analyze QCD processes at both ends of the momentum scale. Two measurements are presented in this thesis, emphasizing one of the two regimes each:

The measurement of charged-particle event shape variables in inelastic proton–proton collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV analyses the transverse momentum flow and structure of hadronic events. Due to the, on average, low momentum transfer, predictions of these events are mainly driven by non-perturbative models. Three event shapes are studied: The transverse thrust, the thrust minor and the transverse sphericity. Data recorded during low pile-up beam conditions in early 2010, corresponding to an integrated luminosity of 168 μb, is analyzed. Charged particles with a transverse momentum above 500 MeV are used to compute the event shape variables. In addition to differential event shape distributions in p_T of the highest-transverse-momentum particle, the evolution of each event shape variable as a function of charged particle multiplicity and summed transverse momentum is presented. Predictions from different Monte Carlo models are compared with data and they show significant deviations from data. Using the results to tune the free parameters of the phenomenological models can improve future Monte Carlo predictions and give additional insight to the non-perturbative part of the QCD.

The measurement of the double-differential dijet cross section as function of the dijet mass and rapidity separation of the two highest-transverse-momentum jets in an event, the leading and the subleading jet, probes the QCD up to the highest scales. Proton–proton collisions at $\sqrt{s} = 8$ TeV recorded during 2012, corresponding to an integrated luminosity of 20.3 fb^{-1}, are analyzed. Jets are identified using the anti-k_t jet clustering algorithm with the radius parameters $R = 0.4$ and $R = 0.6$. The cross section is measured for events with leading jet transverse momentum above 100 GeV and subleading jet transverse momentum above 50 GeV within $|y| < 2.8$. Several NLO predictions are calculated, using four different parton distribution functions, and corrected for non-perturbative effects. The measured results are corrected for detector effects and compared to the four NLO predictions. A good agreement within the systematic and statistical uncertainties between data and theory is observed for dijet masses below ~ 3 TeV. Above that, deviations are observed between the data and perturbative predictions. Miscellaneous behaviour for theory predictions using different PDF sets is visible in the results. Including the results of the measurement into future PDF determinations can improve their performance.

ZUSAMMENFASSUNG

Die Quantenchromodynamik ist die zugrundeliegende Theorie der starken Wechselwirkung und kann in zwei Bereiche aufgeteilt werden. Harte Streuprozesse, wie zum Beispiel die Zwei-Jet-Produktion bei hohen invarianten Massen, können störungstheoretisch behandelt und berechnet werden. Bei Streuprozessen mit niedrigen Impulsüberträgen hingegen ist die Störungstheorie nicht mehr anwendbar und phänemenologische Modelle werden für Vorhersagen benutzt. Das ATLAS Experiment am Large Hadron Collider am CERN ermöglicht es, QCD Prozesse bei hohen sowie niedrigen Impulsüberträgen zu untersuchen. In dieser Arbeit werden zwei Analysen vorgestellt, die jeweils ihren Schwerpunkt auf einen der beiden Regime der QCD legen:
Die Messung von Ereignisformvariablen bei inelastischen Proton–Proton Ereignissen bei einer Schwerpunktsenergie von $\sqrt{s} = 7$ TeV misst den transversalen Energiefluss in hadronischen Ereignissen. Ereignisse mit mindestens sechs geladenen Teilchen mit einem transversalem Impuls von mindestens 500 MeV wurden selektiert. Die Daten wurden Anfang 2010 aufgenommen, während der Beschleuniger bei einer niedrigen instantanen Luminosität betrieben wurde um Ereignisse mit mehreren simultanen Proton–Proton Kollisionen zu vermeiden. Durch den im Durchschnitt niedrigen Impulsübertrag bei diesen Ereignissen nutzen Monte Carlo-Simulationen phänomenologische Modelle um Vorhersagen zu erstellen. Drei verschieden Ereignisformvariablen wurden untersucht: Der transverse Thrust, der Thrust Minor und die transverse Sphericity. Inklusive Verteilungen der Ereignisformvariablen sowie die Distributionen in verschiedenen Bereichen des Transversaimpulses der Spur mit dem höchsten transversalen Impuls wurden erstellt. Zusätzlich wurde die Entwicklung der Ereignisformvariablen in Abhängigkeit von der Multiplizität geladener Teilchen sowie der Summe der Transversalimpulse aller geladenen Teilchen untersucht. Die Ergebnisse wurden mit verschiedenen Monte Carlo Vorhersagen verglichen. Eine signifikante Abweichung zwischen Messung und Theorie konnte beobachtet werden.
Die Messung des zweifachdifferentiellen Zwei-Jet-Wirkungsquerschnittes als Funktion der invarianten Masse sowie der Rapiditätsdifferenz der beiden Jets mit den höchsten Transversalimpulsen kann genutzt werden um Theorievorhersagen zu überprüfen. Proton–Proton Kollisionen bei $\sqrt{s} = 8$ TeV, welche während der Datennahme im Jahr 2012 aufgezeichnet wurden, entsprechend einer integrierten Luminosität von 20.3 fb^{-1}, wurden analysiert. Jets wurden mit Hilfe des anti-k_t Algorithmuses mit zwei verschiedenen Radiusparametern, $R = 0.4$ und $R = 0.6$ rekonstruiert. Der Wirkungsquerschnitt wurde in Events mit führenden Jets mit einem Transversalimpuls $p_T > 100$ GeV und nächst-zu-führenden Jets mit $p_T > 50$ GeV in einem Rapidtätsbereich von $|y| < 2.8$ gemessen. NLO Vorhersagen, gefaltet mit vier verschieden Parton-Dichte-Funktionen, wurden berechnet und auf nicht-perturbative Effekte korrigert. Die Daten wurden auf Detektoreffekte korrigiert und mit den NLO Vorhersagen verglichen. Eine Übereinstimmung im Rahmen der statistischen und systematischen Unsicherheiten bei inverianten Massen unter 3 TeV kann beobachtet werden. Bei höheren invarianten Massen kann die Theorievorhersage die Daten nicht mehr korrekt beschreiben.

CONTENTS

Part I

INTRODUCTION

1

INTRODUCTION

The history of modern particle physics started with Democritus, who believed that everything we conceive consists of smallest particles, the atoms. Even though experiments have taught us that atoms are not *indivisible*, as the Greek word atomos suggests, the idea of smallest particles still exists. The present theory of the smallest particles and how they build up the world surrounding us, is the *Standard Model of particle physics*. Democritus' atoms are replace by 12 fundamental particles in the Standard Model. Three of the four fundamental forces, not including gravity, are contained and well described within the Standard Model. Despite its overwhelming successful, the Standard Model still fails to describe certain phenomena observed, like the existence of dark matter and dark energy. The journey is not over yet, and trying to solve the remaining unknowns is the main undertaking of modern particle physics. The Large Hadron Collider and its associated experiments built at the European Organization for Nuclear Research, *CERN*, in Geneva/Switzerland recently added another piece to the puzzle with the discovery of the *Higgs* boson. The search for new phenomena is not the only goal of the experiments at the LHC. They are also continuously reviewing the existing theory and getting a deeper insight into models used within the theory. This thesis contains two measurements, engaging the latter task: Testing the Standard Model and providing additional information to known approaches.

The LHC, the biggest and one of the most sophisticated machines ever built, is designed to collide protons and heavy ions at high energies. The ATLAS experiment, one of the two multipurpose particle detectors at the LHC, is able to measure the particles produced in the collisions at a very high precision. Studying properties of the particles detected allows us to derive information about how they were produced in the first place. Most new particles predicted by theories are assumed to decay before reaching the detector. Therefore the production mechanism of detected particles is often of great interest. However, the directly observed particles also provide information about the structure of the colliding particles, the protons, and properties of the fundamental interactions used to produce the observed particles.

The quantumchromo dynamic, *QCD*, the theory of the strong interaction, is the dominant interaction at the LHC. But QCD has a weak spot: At low energies, the theory is not calculable, and phenomenological models replace exact predictions. The first measurement described in this thesis, the measurement of charged-particle event shape observables in proton–proton collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV, analyzes inelastic proton–proton collisions with low momentum transfer. These *soft* events fall into the regime of the non-perturbative QCD, the non-calculable regime of the QCD. The models used to describe the physics in soft interactions have many free parameters, which need to be obtained from measurements. Event shape variables measure the energy flow perpendicular to the incoming proton beams. Three event shape variables, the *transverse thrust*, the *thrust minor* and the *transverse sphericity* are used in the analysis as they are sensitive to the soft part of the proton–proton scattering. Data collected at the beginning of 2010, when the LHC was operating at low instantaneous luminosities, were analyzed and the results compared to several Monte Carlo predictions.

Quarks and gluons, also called *partons*, in the final state of an event with high momentum transfer often are detected as sprays of particles flying roughly in the same direction. An approach to connect these particles backwards to the quarks and gluons produced in the scattering process is the concept of *jets*. Jet algorithms recombine particles, presumably originating from the same parton, into a new physics object called jet. The second analysis presented in this thesis uses the jet definition to measure the double-differential dijet cross section as a function of the dijet mass and rapidity separation of the two jets with the highest transverse momentum. Proton-proton collisions recorded in 2012 at a center-of-mass energy of $\sqrt{s} = 8$ TeV were analyzed, corresponding to an integrated luminosity of 20.3 fb^{-1}. Jet cross section measurements are capable of testing perturbative QCD predictions at the highest energy scales. In addition, the dijet cross section is sensitive to the momentum distributions of the partons inside the proton. These *parton distribution functions* are derived by using measurements sensitive to the proton structure, like the dijet cross section presented in this thesis.

The thesis is structured as follows: Chapter 2 gives a brief introduction into the standard model of particle physics, followed by a description of the physics in proton–proton collisions in Chapter 3. The definition of jets and the jet clustering algorithm used in this thesis is presented in Chapter 4. The theoretical overview ends with a short synopsis on theoretical predictions and Monte Carlo simulation in Chapter 5. The experimental setting, namely the LHC and the ATLAS experiment, are outlined in Chapters 6 and 7. The reconstruction of the physics objects used in the two analyses, jets and tracks, are presented in Chapters 8 and 9 respectively. The charged particle event shape analysis is described in Part IV and the double-differential dijet cross section measurement is presented in Part V. A summary of the results and an outlook on future perspectives of the two analyses is given in Part VI.

Part II

THEORY

2

THE STANDARD MODEL OF PARTICLE PHYSICS

The Standard Model of particle physics describes the dynamics and interactions① of all currently known elementary particles and builds the theoretical foundation of the physics used in this thesis.

With the evolving success of quantum field theory and the discovery of many new particles in the 1960s, a new theory was required to explain the emergence of these new particles. Gell-Mann [1] and Zweig [2] postulated the quark model, in which the observed particles were built of smaller constituents, the quarks. At this time, they only predicted three quarks with a charge of $\pm 2/3e$ or $\pm 1/3e$, e being the charge of an electron. Even though this model was not complete, the Ansatz was used to develop the Standard Model together with the work of Glashow, Weinberg and Salam [3–5] on the electroweak unification. The following developed Standard Model not only succeeded in describing the phenomena observed, but also gave precise predictions of particles yet undiscovered. One of them is the Higgs particle, which was found in 2012 by the ATLAS and CMS collaborations [6, 7]. Even though it is very successful, the standard model misses to describe some known phenomena, like dark matter and dark energy. New extensions of the standard model, like supersymmetry, or even completely new theories, e. g. string theory, can solve existing problems but have not yet been proven. Collider experiments like those at the LHC might give evidence for new physics and theories.

The standard model is built of 12 fundamental particles, five force carriers, the recently discovered Higgs boson and 3 forces. An overview of the particles and the forces except gravity is given below. Gravity is the weakest force, even much weaker than the weak force, and therefore does not play a role in particle physics. That makes it hard to analyze gravity in particle physics experiments, but also allows neglecting it.

In the following, natural units are used, which means that the speed of light c and the Plank constant ħ are set to one.

2.1 THE FUNDAMENTAL PARTICLES OF THE STANDARD MODEL

The fundamental particles are split into two groups: The fermions, carrying a half-integral spin, which build up all non-elementary particles, and the bosons, having an integral spin, all acting as mediators of the forces with the exception of the Higgs boson. The bosons are the photon γ, W^+, W^-, Z^0, H and eight gluons g. The γ, $W^\pm$ and the Z^0 are the gauge bosons of the electroweak force. Gluons mediate the strong force. The Higgs boson gives mass to fundamental particles using the Higgs mechanism. The

---

① Except gravity, which is negligible in the processes described here.

non-abelian character of the standard model enables self-interaction to the gauge bosons of the weak and strong force.

The fermions are grouped in three lepton and three quark families/generations. Quarks can interact via the strong force, in contrast to the leptons. Members of the second and third families decay into members of the first family after a short lifetime. To every fermion an antiparticle exists, with inverted charge quantum numbers while conserving all other properties. Neutrinos take a special role within the fermions as they are the only fermions which do not interact via the electromagnetic force. The standard model predicts them to be massless, yet the observation of neutrino oscillation② is only possible with non-zero masses. Up to now, only upper mass limits are known, and the exact mass of the neutrinos is still unknown.

An overview of all fundamental particles and their properties is given in Fig. 2.1.

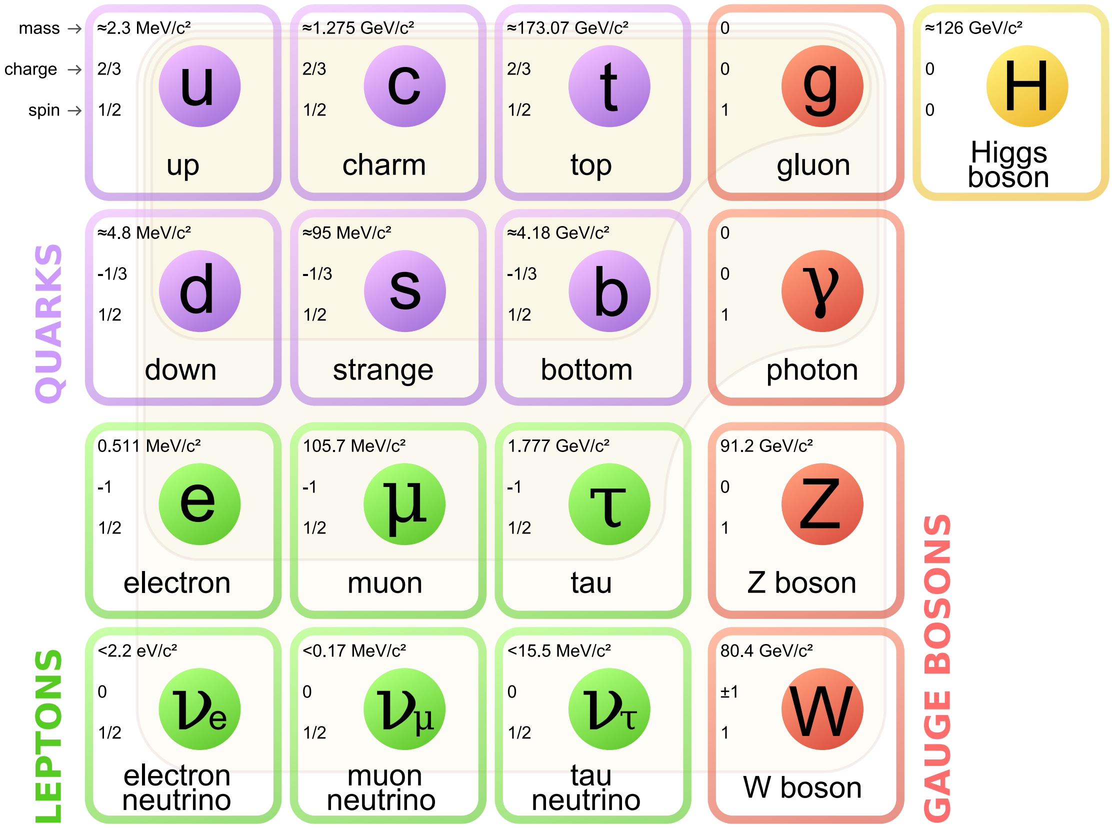


Figure 2.1: Fundamental particles of the standard model. Each of the first three rows represents one family (1. to 3. from left to right) [8].

2.2 THE ELECTROWEAK FORCE

In 1961, Glashow, Weinberg and Salam introduced the unification of the electromagnetic and weak interaction into one underlying force: the electroweak force. The electroweak interaction is described by a non-abelian $SU(2) \times U(1)$ symmetry group. The requirement of local gauge invariance predicts four massless vector bosons. The observed bosons: $W^{\pm}$, Z^0 and γ, are linear combinations of these gauge bosons. As shown

---

② the change of the flavour over time

before, three of the four bosons are not massless ($W^{\pm}$, Z^0) and therefore, the symmetry of the group has to be broken. Introducing a spontaneous symmetry breaking allows heavy bosons. A realization of this symmetry breaking is given by the Brout-Englert-Higgs mechanism [9, 10]. The theory proposes a scalar field to which all massive particles couple. The masses of the W and Z bosons arise from the coupling to this new field. As the Higgs boson does not carry any charge it does not couple to the photon, which therefore remains massless. A consequence of this Higgs mechanism is the existence of (at least) one scalar boson. One main goal of the LHC was to prove or disprove the existence of such a particle. In July 2012, ATLAS and CMS presented the observation of a new particle in the mass range of an expected Higgs boson (around 125 GeV). In March 2013, the compatibility of this new particle with the Higgs boson (high probability of matching spin, charge and parity quantum numbers) was shown.

2.3 THE STRONG FORCE

The key to the development of the standard model was the emergence of new particles. These were predicted to be built of fundamental particles, the quarks. The strong force creates bound states of a quark–antiquark pair (Mesons) or three quarks (Baryons). However, particles like the $\Delta^{++} = \left|u^{\uparrow}u^{\uparrow}u^{\uparrow}\right\rangle$ violated Pauli's exclusion principle that only allows totally antisymmetric wave functions. The introduction of a new quantum number, the color, solved the problem. Each quark carries a color, arbitrarily named after the three base colors red (r), blue (b) and green (g), while anti-quarks carry anti-color ($\bar{r}$, $\bar{g}$ and $\bar{b}$). The resulting quantum chromodynamic (QCD) is based on a non-abelian symmetry group $SU(3)$. Requiring locale gauge invariance, eight fields with the same number of associated bosons are predicted. The bosons of the strong interaction are called gluons and carry color and anti-color of different types at the same time. The Lagrangian of QCD, which describes the dynamics of the strong interaction, is

$$\mathcal{L}_{\text{QCD}} = \mathcal{L}_{\text{Dirac}} + \mathcal{L}_{\text{Int}} \tag{2.1}$$

$$= \sum_{f} \bar{\psi}_{f,a} \left(\imath\gamma^{\mu}\partial_{\mu}\delta_{ab} - m_f\delta_{ab} - g_S\gamma^{\mu}t^{C}_{ab}\mathcal{A}^{C}_{\mu}\right)\psi_{f,a} - \frac{1}{4}F^{C}_{\mu\nu}F^{C\mu\nu}. \tag{2.2}$$

The Einstein summing convention is applied. $\psi_{f,a}$ represents the quark field with flavor f, mass m_f and color index a, $\psi_f = \left(\psi_{qR}, \psi_{qG}, \psi_{qB}\right)^T$ and γ^{μ} are the Dirac γ-matrices. The eight gluon fields mentioned before are described by $\mathcal{A}^{C}_{\mu}$, with $C = 1, ..., 8$ and t^{C}_{ab} being the eight corresponding generators of the $SU(3)$. The coupling constant of the QCD is $\alpha_S = \frac{g_S}{4\pi}$. The first part of Eq. (**??**) in brackets describes free fermions with masses m_f while the last, non-abelian term is responsible for the self-coupling of gluons. The field strength tensors of QCD $F^{C}_{\mu\nu}$ can be expanded to

$$F^{a}_{\mu\nu} = \partial_{\mu}\mathcal{A}^{a}_{\nu} - \partial_{\nu}\mathcal{A}^{a}_{\mu} + g_S f^{abc}\mathcal{A}^{b}_{\mu}\mathcal{A}^{c}_{\nu}, \qquad \left[t^A, t^B\right] = \imath f_{ABC} t^C \tag{2.3}$$

with f_{ABC} being the structure constants of the $SU(3)$.

The fact that no free quarks have been observed yet leads to the assumption that bound states only exists as color singlet states. This is known as *confinement* and is mathematically described by an increasing coupling constant α_S with increasing distance. On the other hand, it can be shown, that quarks and gluons act like free particles inside the proton or neutron. At short distances, quarks and gluons are quasi-free. This phenomenon is called asymptotic freedom. The dependence of α_S on the energy scale μ is described by the renormalization group equation

$$\frac{d\alpha_S(\mu^2)}{d\ln\mu^2} = \beta(\alpha_S(\mu^2)), \qquad \beta(\alpha_S) = -\alpha_S^2(b_0 + b_1\alpha_S + b_2\alpha_S^2 + \ldots), \tag{2.4}$$

where

$$b_0 = \frac{33 - 2n_f}{12\pi}, \quad b_1 = \frac{153 - 19n_f}{24\pi^2} \tag{2.5}$$

with n_f being the number of light quark flavours③. The asymptotic freedom and confinement are direct consequences of the negative sign in Eq. (**??**). At leading order, and ignoring the fact that n_f depends on μ④, a simple solution for α_S is:

$$\alpha_S = \frac{1}{b_0 \ln\frac{\mu^2}{\Lambda^2}}, \tag{2.6}$$

where Λ (also called Λ_{QCD}) is the scale at which the coupling diverges. Only for scales $\mu \gg \Lambda$ perturbation theory is valid. The value of Λ_{QCD} is dependent on a scale choice μ and the order of calculation. However, its order of magnitude, 200 MeV, is constant. This behavior splits the QCD into two regimes. As long as the coupling is sufficiently small ($\mathcal{O}(0.1)$), perturbative calculations can be used. If the coupling becomes larger at lower momentum transfers, perturbation theory fails and phenomenological models are used to describe the physics. The two analyses presented in this thesis belong to different regimes. While the dijet cross section measurement is mainly described by perturbative QCD, with small impact of non-perturbative effects at lower dijet masses, the physics of the charged particle event-shape analysis is purely described by phenomenological models. A brief overview of the two sectors will be discussed in the following.

2.3.1 *Perturbative QCD*

Perturbative calculations of QCD are based on the idea of an order-by-order expansion in $\alpha_S = \frac{g_S^2}{4\pi} \ll 1$. Any observable f can then be predicted as:

$$f = f_1\alpha_S + f_2\alpha_S^2 + f_3\alpha_S^3 + \ldots \tag{2.7}$$

To reach a required precision, one can calculate the series up to a fixed order, assuming the remaining terms to be small. The cross section can be written as

$$\sigma = \sum_i \sigma_i \alpha_S^i \tag{2.8}$$

To compute the unknown coefficients σ_i, QCD Feynman diagrams/rules can be used. Feynman introduced them as a shorthand to visualize and calculate transition amplitudes $|\mathcal{M}|^2$. σ_i is proportional to the sum of all matrix elements contributing to the order i. Integrating over all accessible initial and final states modulo constant coefficients results in the σ_i. All QCD processes can be built from the three fundamental interaction vertices shown in Fig. 2.2. The results are partonic cross sections, which

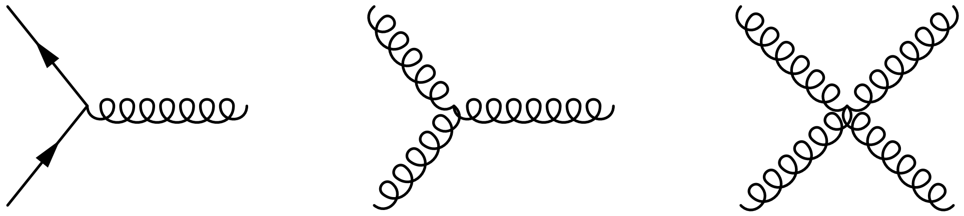
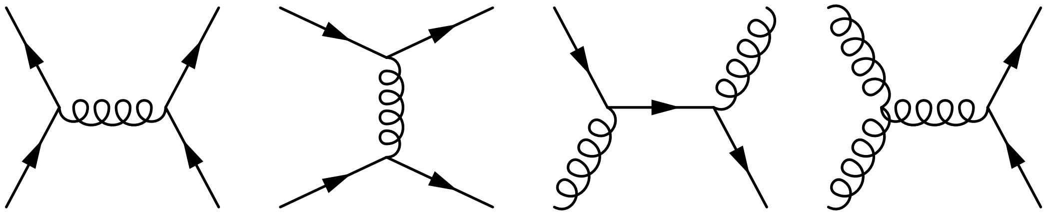
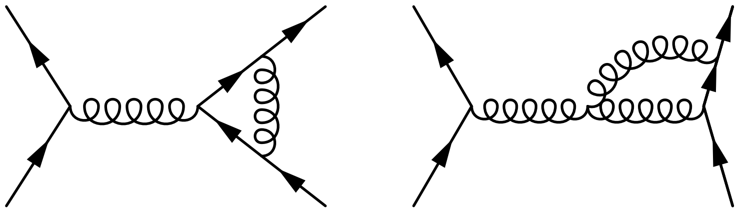


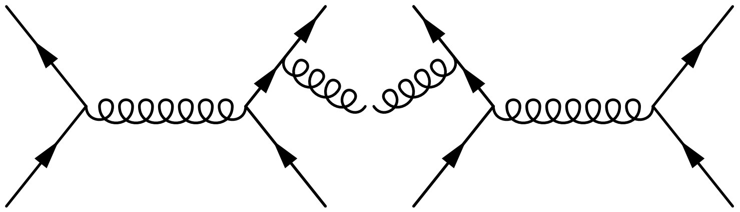
Figure 2.2: Fundamental vertices of QCD represented as Feynman diagrams.



(a) Examples for leading order QCD Feynman diagrams.



(b) Examples for virtual NLO contributions.



(c) Examples for NLO contributions with real emissions.

Figure 2.3: Example for QCD Feynman diagrams, including LO, virtual NLO and real NLO contributions.

means that all ingoing and outgoing particles are either quarks or gluons. Examples for Feynman diagrams for dijet production are shown in Fig. 2.3. Leading order (LO) processes, which are of order $\mathcal{O}(\alpha_S^2)$, are 2 → 2 processes. Contributions from next to leading order (NLO) include 2 → 3 scatterings with a real emission of gluons (initial and final state radiation) and 2 → 2 parton scatterings with gluon loops, which are of order $\mathcal{O}(\alpha_S^3)$. These additional contributions from NLO introduce divergences into the calculation of matrix elements. The probability of a gluon being radiated from a parton is proportional to $\frac{dE}{E}\frac{d\theta}{\theta}$ with the energy of the radiated gluon E and the angle between the parton and the radiated gluon θ. This probability diverges for small energies and/or small angles. The KLN theorem [11, 12] states that these infrared divergences

---

③ flavours are considered light, if the quark mass is below the energy scale μ

④ flavours are considered light, if the quark mass is below the energy scale μ

cancel each other for collinear and infrared (IRC) safe observables. An observable f is IRC safe if:

$$f_{n+1}(p_1,\dots,p_s,\dots,p_n) \to f_n(p_1,\dots,p_n) \qquad \text{for} \quad p_s \to 0 \quad (2.9)$$

$$f_{n+1}(p_1,\dots,p_a,p_b,\dots,p_n) \to f_n(p_1,\dots,p_a+p_b,\dots,p_n) \qquad \text{for} \quad p_a \parallel p_b \quad (2.10)$$

Eq. (**??**) defines infrared safety, the emission of a low energetic gluon does not change the observable, and Eq. (**??**) defines collinear safety, the observable does not change if one parton splits into two collinear partons with the same sum of momenta (or vice versa).

The loop diagrams introduce ultraviolet divergences: The momentum of a gluon inside a loop is not restricted and the gluon therefore can be off-shell. This introduces another divergence when integrating over the gluon momentum. Using a renormalization scale, μ_R, renormalizes the integral and removes the divergences. The cut-off is absorbed in a renormalized strong coupling constant $\alpha_S(\mu_R^2)$ as introduced previously in Eq. (**??**). At the LHC, α_S ranges from $\alpha_S = 0.08$ at a scale of 5 TeV (e.g. heavy dijet systems) up to $\alpha_S \sim 1$ at a scale of 5 GeV (minimum bias events). As a reference point, the strong coupling constant is quoted at a renormalization of $\mu_R = M_Z$, the mass of the Z^0 boson, with a world average value [13] of

$$\alpha_S(M_Z^2) = 0.1184 \pm 0.0007.$$

2.3.2 *Non-perturbative aspects of QCD*

In the previous section, examples of Feynman diagrams in leading and next-to-leading order were shown. They all have in common that the final state consists of partons. At the same time, the concept of confinement has been introduced. The final state, as observed and measured, cannot be made up of color-charged particles. The transition from partons to hadrons, which are observed in the detector, is called *hadronization* and cannot be calculated perturbatively. The vast majority of Monte Carlo generators uses either the Lund-String-Model [14] or the Cluster-Fragmentation [15].

The PYTHIA Monte Carlo generator family [16] uses the Lund-String-Fragmentation. In this model, a color-field between two quarks is assumed. The larger the distance between the two quarks, the higher the energy stored in the color-field. The field between two quarks increases until the field lines break up and a new $q\bar{q}$ pair is produced. The newly produced quarks are connected to the originating quarks via color-strings. The quark-pair-production continues until quark-pairs with on-shell masses of hadrons are reached. A schematic example is shown in Fig. 2.4. Each line represents an (anti-)quark, and the shaded area represents the color-field. V-shaped junctions are quark-pair productions.

The HERWIG Monte Carlo generators [15] use the cluster fragmentation model to predict the hadronization. The cluster model starts with the splitting of all gluons from the parton shower non-perturbatively into quark–antiquark pairs. Color-singlets are formed from all available (anti-)quarks. If the mass of the cluster exceeds

$$M^p > C^p_{\mathrm{max}} + (m_i + m_j)^p, \qquad (2.11)$$

the cluster is fissioned and builds new clusters with a light (u, d and s) quark–antiquark pair drawn from the vacuum. M is the mass of the cluster, C and p are free parameters

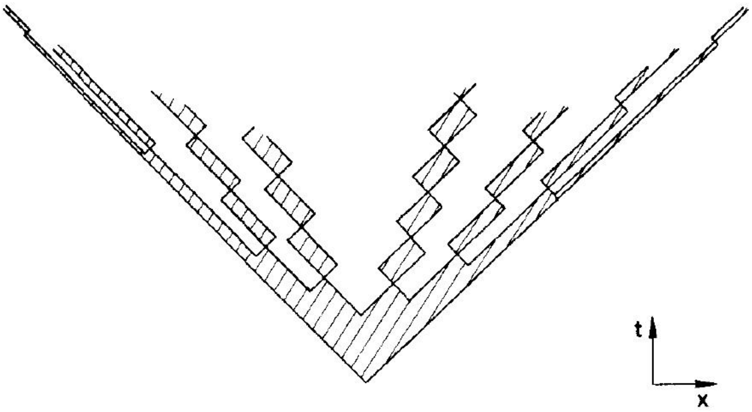


Figure 2.4: Model of the Lund-String-Fragmentation. Every line represents an (anti-)quark and at every crossing, a $q\bar{q}$ pair is produced. Shaded areas represent non-vanishing fields [14].

of the model. The masses of the quarks forming the cluster in the first place are m_i and m_j. This fission process is repeated until every cluster mass fails to fulfill the relation above. Each of the remaining clusters decays into one or two light hadrons.

Both hadronization models have in common that every high-momentum parton in the final state before hadronization results in a bunch of particles flying roughly in the same direction. The higher the parton momentum, the more collimated the particles are. The idea of jets is to cluster all particles from a parton into one new object. A detailed approach to jet clustering and used algorithms is presented in Sec. 4.

3

PROTON-PROTON COLLISIONS

The initial state of the QCD processes shown before are all made up of quarks or gluons. Due to the confinement a pure parton collider cannot exist, and pp collider are used instead. Protons are built of 3 valence quarks (uud) and gluons holding them together. In addition, the gluons split into quark pairs or further gluons, which recombine shortly after again. The quarks produced by splittings are called sea quarks. When colliding protons at energies high enough to resolve the inner structure of the proton, valence quarks, sea quarks and/or gluons from each proton collide and scatter.

3.1 THE HARD SCATTERING PROCESS

A basic sketch of a pp collision is shown in Fig. 3.1. The difficulty is that neither the momentum nor the type of the colliding particles can be experimentally determined. In order to be able to make predictions about proton–proton collisions, probability functions of the momentum distribution of the partons inside the proton are used. The *parton-distribution functions* (PDF) $f_{a/p}(x, \mu_F)$ provide the probability of a parton a carrying the momentum fraction x of the proton ($p_{\text{parton}} = x p_{\text{proton}}$). x is the Bjorken-x variable, also known from deep inelastic scattering. The factorization scale μ_F determines the threshold up to which parton emissions are provided by the PDF and is usually chosen equivalently to the renormalization scale μ_R. Above μ_F, the splitting is provided by the partonic cross section. The factorization theorem shows that PDFs are universal, i. e. independent from the hard scatter. To derive the probability for a process $pp \mapsto X$, the probability of picking randomly two partons from the proton with a certain momentum fraction x and the probability of the scattering of two partons into the desired final state $\hat{\sigma}_{ab \mapsto X}$ are folded. The integrated cross section is achieved by summing over all possible parton types and integrating over the momentum fractions of both partons. Mathematically, this can be expressed as:

$$\sigma(pp \mapsto X) = \sum_{a,b} \int_0^1 dx_a \int_0^1 dx_b \, \hat{\sigma}_{ab \mapsto X}(x_a, x_b, s, \mu_R^2, \mu_F^2) f_{a/p}(x_a, \mu_F) f_{b/p}(x_b, \mu_F) \tag{3.1}$$

with the center-of-mass energy $\sqrt{s}$ of the two protons.

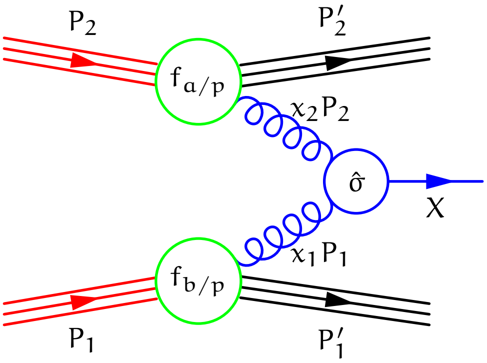


Figure 3.1: Hard scattering in a proton–proton collision. The red lines indicate the incoming protons, built of many partons. The momentum fraction carried by the colliding partons is given by the PDFs $f_{a/p}$. The hard scattering itself is drawn in blue, where the outgoing arrow X is not necessarily a single particle, but could also represent a system of many particles.

3.2 PARTON DENSITY FUNCTIONS

As explained above, the parton density functions provide the probability that a parton of type a carries the portion x of the proton momentum. This section outlines the way how PDFs are obtained. To achieve this, a closer look at the structure of the proton is needed. As mentioned earlier, the proton contains the three valence quarks (uud) and gluons, gluing the proton together. These gluons can split into quark pairs. These quarks are called sea quarks. The proton can be understood as a constant fusion and fission of partons, without changing the protons properties as a whole. The probabilities of these splittings are given by the Altarelli-Parisi splitting functions [18].
Since there is no known prediction for the x-dependency of the PDFs, this information needs to be drawn from measurements. First measurements of PDFs at high x and low Q^2 were performed in deep inelastic scattering (fixed target and ep collisions at HERA [19, 20]). To expand this information to other (x, Q^2) regions, PDF fits are used. The evolution in x is unknown, and needs to be measured while the expansion in Q^2 is performed using the DGLAP formalism [18, 21–23]. At a given (arbitrarily selected) factorization scale and order of $\alpha_S(\mu_R)$, a parametrization in x at a starting scale Q_0 is chosen. These PDFs are evolved to the momentum scales where measurements are provided, using the DGLAP formalism, and folded with partonic cross sections. The result is then compared to the measurements. The free parameters are determined by minimizing a global χ^2 fit to all measurements used in the global fit. The probability for a gluon emission diverges as the momentum of the gluon goes to zero. The capability of resolving these splittings is proportional to the momentum scale of the collision. Therefore, divergences vanish, as long as the momentum scale is finite. On the other hand, the higher the momentum scale, the higher the contribution from sea quarks. Fig. 3.3 shows the parton density function as a function of the momentum fraction x provided by the MSTW PDF group. Two observations could be made. First of all, the contribution from gluons is dominating. In fact, 50% of the proton momentum is carried by the gluons, and not by the valence quarks, as might be assumed naively. Second of all, the increased contributions from sea quarks due to a higher resolution is visible at higher momentum scales and low x. At $x = 10^{-3}$, the PDF increases by about an order of magnitude for the sea quarks, going from $Q^2 = 10\,\mathrm{GeV}^2$ to $Q^2 = 10^4\,\mathrm{GeV}^2$. The colored bands represent the uncertainties on the PDFs. There are many sources

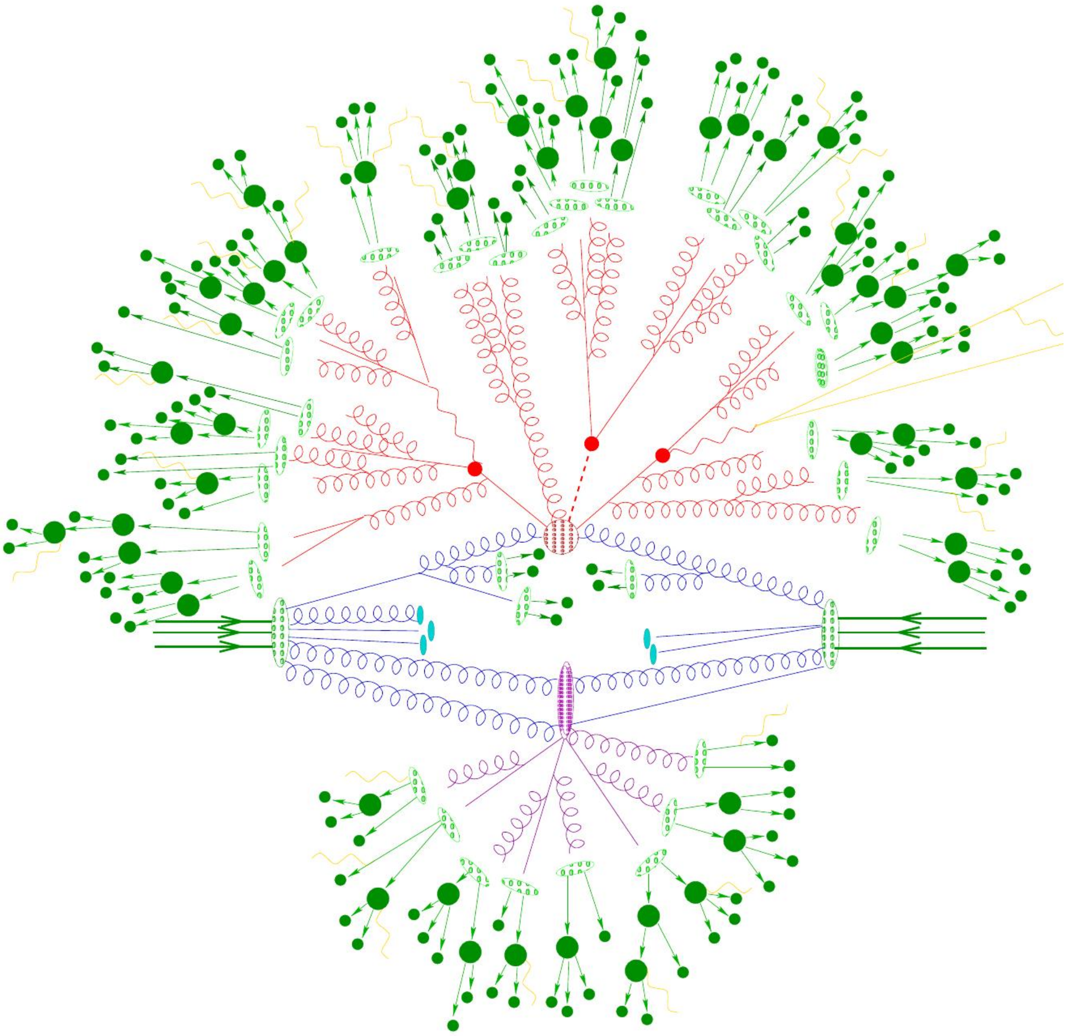


Figure 3.2: Example of a pp collision. The incoming partons are drawn in blue, the hard scattering is colored in red and contributions of the underlying event in purple. The hadronization is indicated by green lines [17].

of uncertainties. First of all, the uncertainty on the data used for the PDF fits. These are propagated to the PDFs. The propagation however is not straightforward since contributing uncertainties might be correlated. A set of uncorrelated parameters is determined using the Hessian method [24]. These new parameters are then used for the error propagation. The parametrization chosen can be altered, and the variations taken as uncertainties. The last main contribution to PDF uncertainties are input parameters: e.g. heavy quark masses or the value of the strong coupling constant are defective values. They can be chosen at fixed values with given uncertainties or are additional fit parameters in the parametrization. Both approaches lead to uncertainties. Even if two PDF groups use the same input data, it is not necessarily true that the predicted PDFs agree. Different choices of parameterizations and handling of input parameters as mentioned above can lead to different results. Therefore, different PDF sets were used in the dijet cross section analysis presented in this thesis, and the uncertainties are propagated through the calculation and are shown in the results.

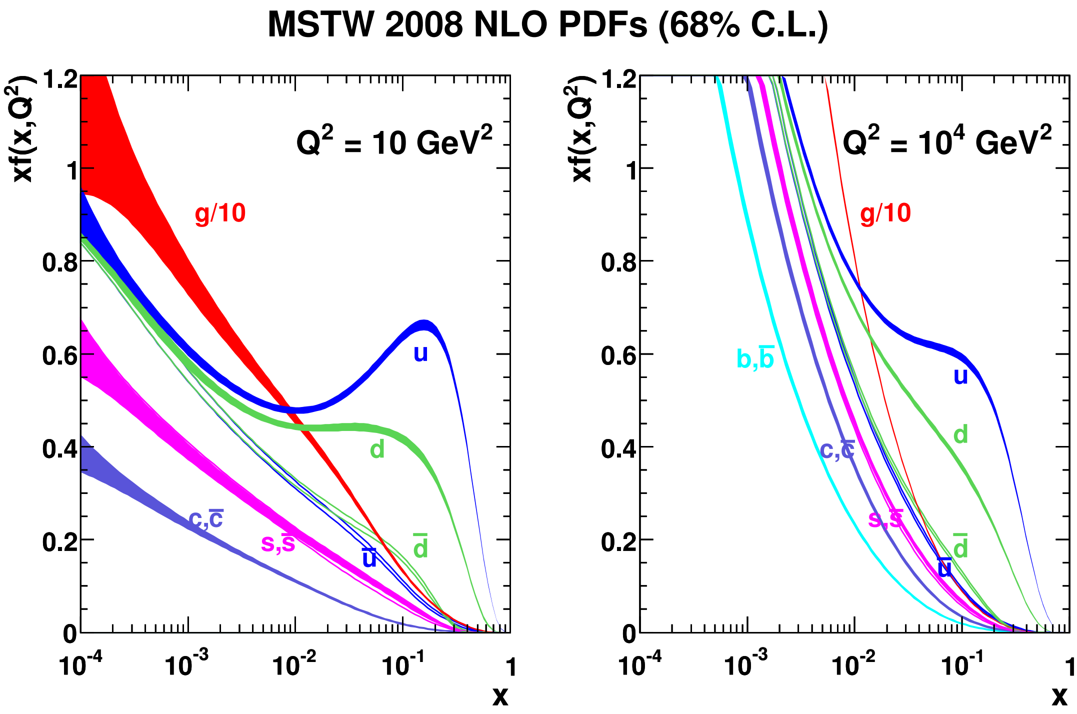


Figure 3.3: Parton density functions at NLO from the MSTW group for two different momentum scales Q^2. The gluon contribution scaled by 0.1 [25].

3.3 MULTIPARTON INTERACTIONS, PARTON SHOWER AND HADRONIZATION

The full picture of a proton–proton collision including the underlying event and hadronization is shown in Fig. 3.2. The hard scattering described above is drawn in red. A complexity is added to the picture by the fact that it cannot be assured that only one pair of partons scatter. These additional, mainly soft $2 \rightarrow 2$ QCD scatterings are donated as the *underlying event* (UE), drawn in purple. The low momentum transfer in these additional parton interactions pushes them into the non-perturbative regime, and phenomenological models are used to describe them. The models are based on the idea that the number of parton interactions is given by the cross section ratio $\langle N_{\text{int}} \rangle = \sigma_{\text{soft}}/\sigma_{pp}$. To keep it finite, σ_{soft} is only evaluated above a minimum transverse momentum $p_{T,\text{min}}$. This cut-off is usually one of many free parameters in the underlying event modeling of the phenomenological models used in the Monte Carlo generators. A good description of the underlying event is needed as it adds energy depositions in the detector to the hard scattering processes, which is mostly the object of interest.

Fig. 3.1 neglects any parton splitting so far. Initial and final state radiation/parton shower are added to the complete picture. The parton showering is based on the so called Sudakov form factors, $\Delta(Q_0, Q_1)$. These factors describe the probability of a parton not splitting into two new partons between the scale Q_0 and Q_1. Starting from a $2 \rightarrow 2$ scattering, Sudakov form factors are applied recursively to each in- and outgoing parton down to a minimum scale Q_0. The time-like final state parton shower is evolved in time and decreases the virtuality Q, while space-like initial state radiation is evolved backwards in time, starting at the hard scattering and decreasing the scale Q until it is matched to the PDFs. The minimum scale Q_0^2 is chosen to be the energy at which emissions cannot be resolved anymore, to avoid divergences. Usually, values of few GeV are used in Monte Carlo simulations. The choice of the scale Q is arbitrary, however the transverse momentum of the radiated gluon or the angle between the parton and the radiated gluon are common choices in Monte Carlo simulations. Since a collinear approximation entered the derivation of the Sudakov form factors, the parton shower describes soft radiations well, while harder/non collinear radiations are described worse. Combinations of NLO predictions and parton shower have the complexity of double counting (contributions enter the calculation via parton shower and NLO matrix elements). A scale barrier at which the parton shower stops and the radiation is predicted by NLO matrix elements needs to be introduced, and a matching has to be applied to avoid double counting.

The particles produced so far are all partons carrying a non-neutral color charge, which undergo the process of hadronization, described earlier. The hadronization is added to Fig. 3.2 in green. The original simplified model of a pp collision from Fig. 3.1 is only a small part. Studying the behaviour of parton shower, the underlying event and hadronization, however, can also provide new insight into non-perturbative QCD for a better understanding of the complete picture.

At the LHC, no single, well-separated protons collide, but series of bunches with in average 1.7×10^{11} protons per bunch. This leads to the effect of pile-up. Pile-up describes the situation where not all detector signals are originated from a single pp collision but may also be caused by additional coinciding collisions and/or signals remaining in the detector from previous collisions (if signal decay time is above the time gap between two consecutive collisions). A more detailed discussion about pile-up is given later.

4

JET DEFINITIONS

The final states of pp collisions have a very complex structure. Fig. 8.1 and Fig. 3.2 show examples for such final states. The underlying feynman diagram on parton level however has a simple structure. Tracing back the final state particles to the partons is one goal of jet algorithms. Reducing the complexity also introduces the possibility to make theoretical predictions, e. g. to jet cross sections. In order to be a reliable and useful tool for physics analyses, the jet algorithm needs to be stable and reproducible. To ensure the cancellation of virtual and real divergences, jet algorithms need to be infrared and collinear safe. The last requirement is resilience. Jets should be resilient to effects that cannot be calculated and the results on different stages of event simulation should be stable. The three stages are:

Parton level Jet finding is applied to the final state before hadronization. Parton momenta predicted by NLO calculations or generated by MC generators are used as input variables.

Particle level The jet algorithm is applied to stable①, color-neutral particles after hadronization.

Detector level In experiments, only detector signals are accessible. Therefore, jets are formed from detector signals. These can be tracks (from charged particles only), calorimeter cells, or more complex topological clusters (see Sec. 8.1).

As mentioned before, partons are detected as a spray of particles roughly flying in the same direction. A simple approach for a jet algorithm would be to use actual cones and recombine all particles within the cone to a jet. Cone algorithms were very popular at recent collider experiments. A common approach for a cone algorithm would be②:

1. Choose a seed (usually the highest p_T input particle).
2. Build a cone with a radius of $R = \sqrt{\Delta\phi^2 + \Delta y^2}$ around the seed, where ϕ is the azimuthal angle and y the rapidity. Combine all particles inside the cone to a proto-jet (usually by summing over the particles four-momenta).
3. Repeat the previous step until the cone direction does not change from one iteration to the next. The proto-jet now is called jet. Then remove the particles attached to the jet from the seed list.
4. Repeat steps 1-3 until all particles are clustered into jets.

---

① Particles with a lifetime $\tau > 10$ ps are considered stable

② using particles as input constituents to simplify the description

More advanced algorithms introduced a *split and merge* step in which overlapping jets were either merged or the energy depositions in the overlap region split among the jets. Many variations of the described cone algorithm have been developed. The main problem however is the seed finding as this step is not infrared and/or collinear safe. Imagine only particles above a certain threshold can serve as seeds, as it was often the case. If a particle, with a p_T above the threshold splits into two or even more particles which are all below the threshold, the seed list would differ and with that the result of the jet finding as well. If, in another case, an infrared radiation is above the seed threshold and enters the seed list, it might cause the merging of two existing jets or form even an additional jet. Both cases would violate the infrared safety. Therefore, most cone algorithms are not in use anymore. The SISCone algorithm [26] is a seedless cone algorithm which is infrared and collinear safe. However, another class of jet algorithms came into favor: The recombination algorithms. The most popular member is the k_t-algorithm family [27]. The k_t algorithm uses the four following steps for jet finding:

1. Calculate the distance $d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta y^2 + \Delta\phi^2}{R^2}$ between all particle pairs i, j.
2. Calculate the beam distance $d_{iB} = k_{ti}^2$.
3. Combine particles with smallest distance d_{ij} or, if d_{iB} is the smallest distance, call it a jet. Remove all associated particles from the list.
4. Repeat steps 1-3 until all particles are associated to jets.

The transverse momentum is described by k_t and R is a free distance parameter, determining the jet radius. The algorithm inside the k_t-family is specified by the parameter p and three choices are used. The original k_t algorithm, using $p = 1$, is hierarchical in relative p_T. Choosing $p = 0$, called Cambridge/Aachen, completely removes the p_T dependency from the algorithm and is purely hierarchical in angle. The algorithm used in this thesis is the *anti*-k_t algorithm, using $p = -1$. All three members of the k_t algorithm family are collinear and infrared safe as they do not rely on seeds. The anti-k_t algorithm however produces nearly perfectly conical hard jets. A comparison of jet active areas for the three k_t and the SISCone jet clustering algorithm are shown in Fig. 4.1. To calculate the active area, *ghost* particles with infinitesimally small transverse momenta are added to the event and clustered within the jets. The IRC safety of the jet algorithms secures that the resulting jets do not change. A more detailed description is given in Sec. 8.2.1.

The combination of regular/conical shaped jets and the IRC safety is the reason why the anti-k_t algorithm is chosen for this analysis. The jet radius is the second parameter to choose. It regularizes the cone size of the jet. This has three major consequences: The larger the cone size, the more contributions originating from the underlying event are collected by the jet. As the jet should be a proxy for partons, small jet radii are favoured. Reducing the jet radius could lead to out of cone losses. Radiated partons could fall outside the cone, meaning not contributing to the jet energy. The hadronization also produces broader particle sprays. Losses could occur if the jet radius is too small. The dependency of Δp_T, the difference in transverse momentum of the parton and the jet at particle level, on the jet radius is shown in Table 4.1. In this analysis, two jet radii have been chosen: $R = 0.4$ and $R = 0.6$. This is consistent with previous jet analyses in ATLAS, for an easier possible combination of the results.

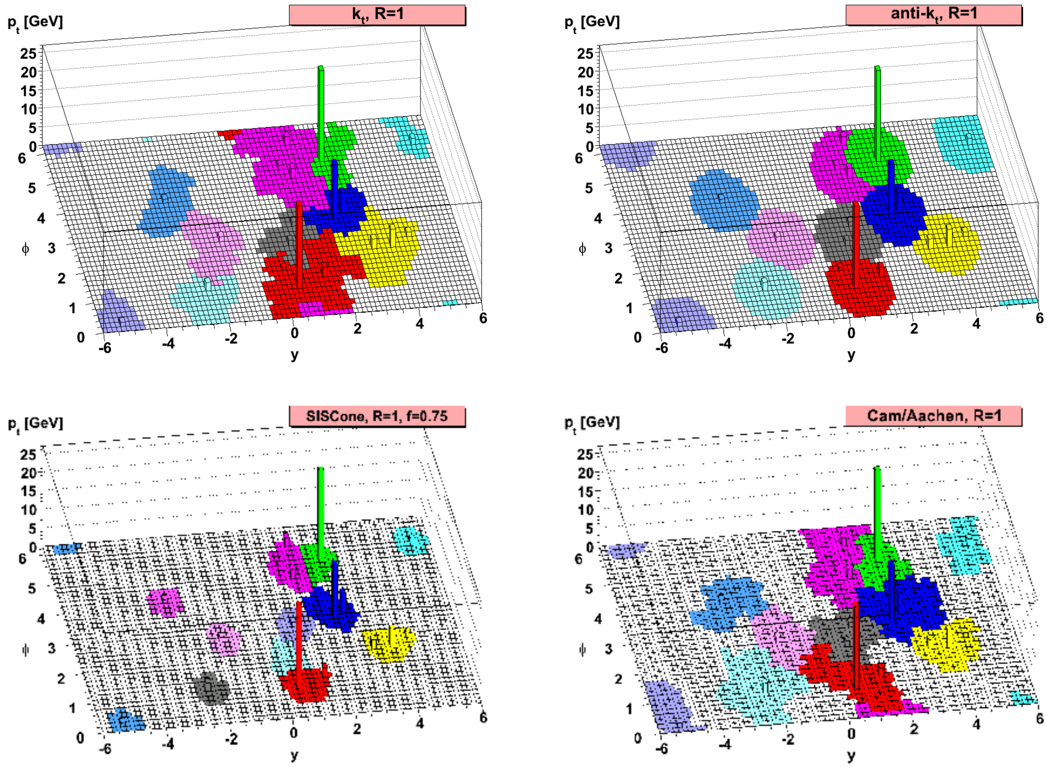


Figure 4.1: Active area for four different jet clustering algorithms in a sample parton level event generated with Herwig. Jet areas are calculated using ghost particles. A detailed description is given in Sec. 8.2.1 [28].

Effect	R-dependency
Perturbative radiation	$\Delta p_T \sim \ln(R)$
Hadronization	$\Delta p_T \sim -\frac{1}{R}$
Underlying event	$\Delta p_T \sim R^2$

Table 4.1: Dependency of the difference between the parton's and the jet's transverse momentum Δp_T on the radius parameter R. Dependencies taken from [29]

5

THEORETICAL PREDICTIONS AND SIMULATION

Monte Carlo (MC) simulations are a widely used tool in particle physics analyses. Especially the non-perturbative part of pp collisions can only be simulated and cannot be calculated. As Monte Carlo simulations also include a detector simulation, they are often used to correct the data back from detector level to particle or parton level. It is always preferred to use data-driven methods instead of Monte Carlo, as the Monte Carlo introduces new biases as it can never perfectly reproduce the data. But in most cases, a proper data-driven correction method is not available. Monte Carlo is also often used for background estimations. Pure data-driven techniques are in favour, but again it is not always possible to develop such a technique. Studying detector effects, developing energy calibrations and many more applications often make use of Monte Carlo simulations.

Depending on the process, the LO predictions provided by Monte Carlo simulations are not able to describe the data well due to missing, non-negligible higher order contributions. In this case, higher order theory calculations are needed. First approaches to combine NLO calculations with parton showers have been made [30–32], but generated events at a high enough number to perform the analysis were not available at the time of publishing this thesis. In order to make a useful comparison, the data and the theory predictions need to be at the same "level". Pure calculations are always provided at parton level. If possible, the data is corrected back to parton level or both are corrected to particle level. The latter is done for the jet analysis presented later as a satisfying correction to parton level of the data is not possible.

This chapter will describe the theoretical predictions used in the two analyses presented in this thesis. Sec. 5.1.1 gives a brief introduction to Monte Carlo simulation and will highlight the differences between the two Monte Carlo generators mainly used in ATLAS: PYTHIA and HERWIG. The second part in Sec. 5.2 introduces the NLOjet++ tool which was used to produce NLO cross section predictions on parton level for the dijet cross section analysis.

5.1 MONTE CARLO SIMULATION

Monte Carlo simulations use fixed order hard-coded matrix elements, here LO $2 \to 2$ processes, provided by perturbative QCD calculations to describe the hard scattering. PDF sets are used when calculating the hard scattering, which inclusively integrate over initial-state radiation. The initial state partons of the hard scattering then gets evolved backwards in time, adding initial state radiation and therefore increasing the scale, until they are matched to the proton PDF. The final state partons undergo parton showering as described before. The resulting partons undergo the hadronization, and the underlying event is simulated and added to the event. The generated particles then pass the detector simulation, developed using the GEANT4 framework [33]. GEANT4 works in two steps. At first, the four vectors of all produced particles in the event generation are fed to GEANT4. The interactions, so called *hits*, between the particles and the detector material is simulated within the framework. In a second step, the hits in active material then are transformed into detector signals. These are stored in the exact same container format as the data and undergo the same reconstruction as data events do. Additional generated events are added according to the expected pile-up situation in data.

5.1.1 *Monte Carlo generator*

ATLAS mainly uses two generator families: PYTHIA and HERWIG. The PYTHIA and Herwig++ generators are widely used in particle physics and different versions/tunes are used in this thesis. They mainly differ in the modeling of hadronization, parton shower and the underlying event. While PYTHIA uses the Lund-String fragmentation, Herwig++ makes use of the cluster fragmentation (see Sec. 2.3.2). The parton shower applied can either be ordered in the momentum of the emitted parton or the angle between the emtting and the emitted parton. Herwig++ implemented the latter while PYTHIA uses the p_T-ordered parton shower. As the two generators are both equally successful in describing measured data, but use different approaches, they are often used to compare the spread of possible MC predictions.

5.1.2 *Monte Carlo tuning*

The final result of the MC however does not solely depend on the generator used, but could also differ while using the exact same approaches. The models used to describe hadronization, parton shower etc. have many free parameters which cannot be determined by first principles. Instead, Monte Carlo tuning is used to regulate these parameters. Data sensitive to these non-perturbative effects is used and the parameters are tuned to give the best description over a wide variety of input data. Different sets of input data, tuning methods and PDFs used in the tuning process lead to so called tunes. The tunes can specialize to describe e. g. certain effects predominantly well, with the effect providing less precise predictions for other parts. To achieve this, input data can be weighted differently (e. g. accent underlying event measurements) or completely ignored. A few parameters (PYTHIA nomenclature used) which predominantly manipulate the predictions of the tunes are described in the following.

Multi parton interaction cut-off The minimal transverse momentum needed to allow $2 \to 2$ processes was already mentioned. The corresponding tuning parameters are PARP(82) and PARP(90):

$$p_T^{\min} = \text{PARP(82)} \left(\frac{E}{1.8\ \text{TeV}} \right)^{\text{PARP(90)}} \tag{5.1}$$

with E being the center-of-mass energy of the collision in TeV.

Proton overlap The impact parameter of the scattering is described by the overlap of two Gaussian matter distributions, parameterized by PARP(83) and PARP(84).

Color reconnection Color reconnection describes the interaction between color strings (Sec. 2.3.2). The probability of a color string not to participate in color reconnection is given as:

$$(1-\mathrm{PARP}(78))^{n_{\mathrm{MPI}}} \tag{5.2}$$

n_{MPI} denotes the number of MPI occurring in the event. The color reconnection of fast moving color strings is suppressed by the factor

$$\frac{1}{1+\mathrm{PARP}(77)^2\cdot p_{\mathrm{avg}}^2} \tag{5.3}$$

with the average momentum p_{avg} of the partons being connected via the color string.

Initial state radiation The lower cut-off scale for ISR is given by PARP(93)

Herwig uses similar parameterizations where applicable. The cut-off scale e.g. is given as the angle between the initial parton and the radiated one.
The current default tune for 2012 in ATLAS is the PYTHIA 8 ATLAS Underlying Event Tune 2 (AU2), with an emphasis on underlying event tuning. The AU2 tune uses the CT10 NLO PDF set. For analyses at $\sqrt{s}=7$ TeV, the PYTHIA 6 ATLAS Minimum Bias Tune 2B, as the name suggests with a focus on minimum bias events, was widely used as reference. Minimum bias data has been selected, using a minimal trigger requirement, to select an unbiased sample of events, which mainly consists of inelastic low-Q^2 proton–proton scatterings.

5.2 NLOJET++

Perturbative QCD predictions at next-to-leading order were produced using the *NLOjet++* [34, 35] software tool, interfaced with *APPLGRID* [36]. APPLGRID stores the coefficients of perturbative calculations provided by NLOjet++ in a (x, Q^2) grid. The grid can then be convoluted with different PDF sets and scale choices, without the need to recalculate the coefficients again. In the case of this analysis, the CT10, HERAPDF 1.5, NNPDF2.3 and MSTW2008 PDF sets have been used. The default scale choices are taken from previous ATLAS analyses:

$$\mu_R=\mu_F=p_{\mathrm{T,lead}}\cdot e^{0.3\cdot y^*} \tag{5.4}$$

with $p_{\mathrm{T,lead}}$ being the highest transverse momentum of any jet in the event and $y^*=|y_1-y_2|/2$, half the rapidity separation between the two jets with the highest p_{T}. The scale choice has been preferred over a more intuitive and basic scale choice such as $p_{\mathrm{T}}^{\mathrm{lead}}$, as those lead to unstable cross section calculations, even negative values at high y^*. These high y^* regions are not included in the analysis presented here, but an extension to higher y^* values is easier using these scale choices.

5.2.1 *Systematic uncertainties on NLO calculations*

The uncertainty on the theory calculation can be split into three sources:

Free scale choices of μ_R and μ_F The calculation used in NLOjet++ include two free scale choice. The scales were chosen to be

$$\mu_R = \mu_F = p_T^{\text{lead}} e^{0.3y^*} \tag{5.5}$$

and varied by a factor of two independently. The differences are taken as the uncertainty due to that free scale choice.

Propagation of the uncertainty on α_S A second source of uncertainty is α_S. The strong coupling constant is varied within its uncertainty. The spread of the results is used as the systematic uncertainty on the calculation.

Propagation of PDF uncertainties The last component of the uncertainty on the NLO calculation are the uncertainties on the different PDFs used. Each PDF used provides a set of systematic uncertainties. The main source of systematic uncertainties are the propagations of experimental inputs to the PDF and uncertainties of the fits used to derive the PDFs. Uncertainties on theoretical values such as quark masses are used when provided. The PDF sets are fluctuated within these given uncertainties and the resulting differences in the calculated cross section is taken as the systematic uncertainty.

The uncertainties for two exemplary bins in y^* are shown in Fig. 5.1. The combined uncertainty is about 5% at the lowest m_{12} bin and grows to above 20%. The contributions from the different sources are equal in size for lower dijet masses. The higher the dijet mass, the more dominant the PDF uncertainty gets. All sources of uncertainties are considered to be uncorrelated and were added in quadrature.

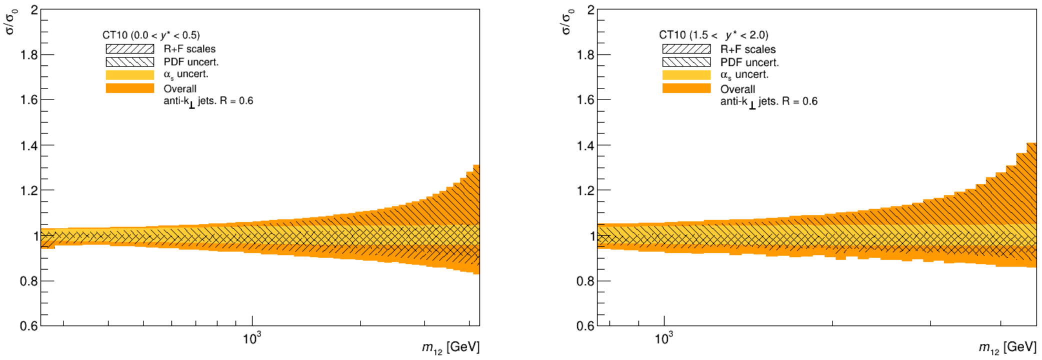


Figure 5.1: Systematic uncertainties of the cross section prediction generated with NLOjet++. Hashed areas correspond to the scale and PDF uncertainties, the light yellow band represents the propagated uncertainty on α_S and the combined uncertainties are drawn in orange. Both plots are drawn with a logarithmic scale on the x-axis. Left side are the uncertainties for the lowest y^* bin from $y^* < 0.5$ and the right side shows the uncertainties for $1.5 < y^* < 2.0$.

5.2.2 *Non-perturbative corrections*

Theory corrections were computed with NLOjet++ as mentioned before. These predictions are made on parton level. To be able to compare it to the data after unfolding (which has been corrected back to particle level), the theory prediction needs to be corrected to particle level as well. These correction are called non-perturbative corrections as the two effects leading from parton to particle level are of non-perturbative nature. The corrections were derived using Monte Carlo simulations. Dijet events were produced on parton level first, and then at particle level with hadronisaition and underlying event turned on. The ratio of the two distributions are used as the non-perturbative corrections, shown in Fig. 5.2. While the calculation itself is error-prone,

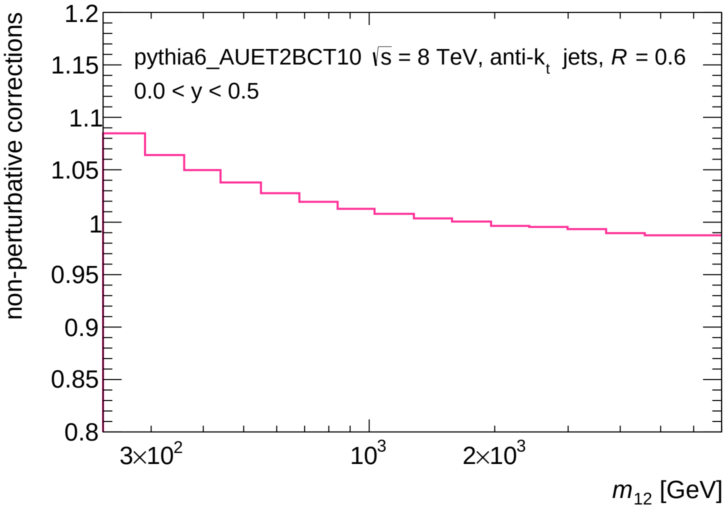


Figure 5.2: Non-perturbative corrections computed with the Pythia8 Monte Carlo generatore using the AUET2B tune with CT10 PDF sets for $0 < y^* < 0.5$.

the non-perturbative corrections add extra uncertainties to the theory prediction. To account for different corrections from different Monte Carlo simulations, a wide variety of PYTHIA 8 and PYTHIA 6 tunes paired with different PDF sets were used to produce non-perturbative corrections. The tune used to derive the unfolding corrections for data and the jet energy scale, PYTHIA 8 AU2 CT10 is used as the nominal correction and the spread of the predictions is then taken as a systematic uncertainty. All available corrections are shown in Fig. 5.3 as well as the resulting uncertainties. A strong m_{12} dependence is visible. The higher the mass of the dijet system, the lower the non-perturbative correction. This is well understood as non-perturbative effects add a constant amount of energy to the event, independent of the hard scattering process. The smaller the jet energies, the higher the relative energy change. In addition, high-p_T jets are more collimated and out-of-cone losses due to non-perturbative effects are smaller. The nominal correction is one of the highest corrections. Therefore, the spread to higher values is small while the uncertainty to lower values is up to 12%. The corrections shown in Fig. 5.3 cluster into three groups. PYTHIA 8 predicts throughout all tunes and PDF sets the highest corrections with up to 15%. The next lower group are most of the PYTHIA 6 tunes, with corrections of up to 10%, decreasing in the same way as the corrections derived from PYTHIA 8. An exception represent the two Perugia tunes. They predict a very small, more or less flat correction. The differences between PYTHIA 8 and PYTHIA 6 are probably caused by the new enhanced MPI model in PYTHIA 8 which changes the underlying event. The small corrections for the Perugia tunes could be explained by a suppressed underlying event activity. A closer look at the tuning parameters could reveal more information about the differences but due to the small impact to this analysis, this is not part of this thesis but could be subject of future studies on dijet measurements.

Currently, only LO and NLO predictions are available for dijet production. Approximative calculations of NNLO contributions predict the additional contributions to be small, hence the comparison of data with NLO predictions is assumed to be justified.

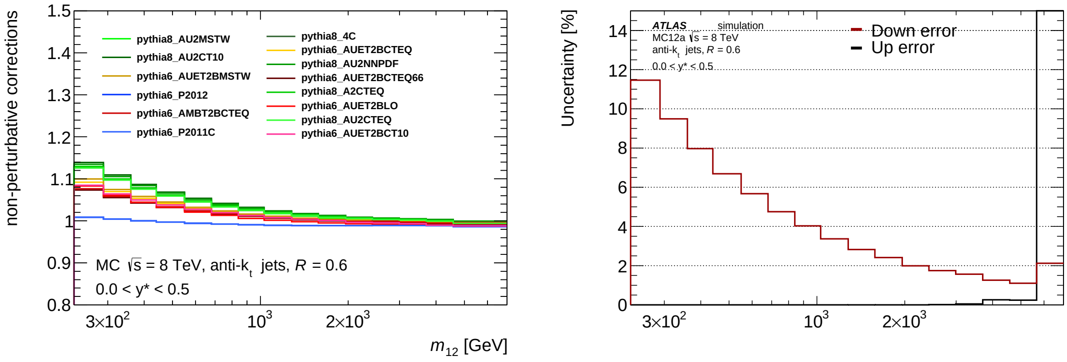


Figure 5.3: Non-perturbative corrections derived from different combinations of Monte Carlo generators, tunes and PDF sets for $0 < y^* < 0.5$ (left) and the resulting relative systematic uncertainty on the non-perturbative corrections (right). The red line represents the uncertainty towards lower values, the red towards higher values.

Part III

THE ATLAS DETECTOR AND THE LHC

6

THE LHC AND ITS EXPERIMENTS

This chapter introduces the *Large Hadron Collider* [37], LHC, and the experiments associated with it.

The LHC is a circular particle collider designed to accelerate protons and heavy ions and is situated at the European Organization for Nuclear Research (Conseil Européen pour la Recherche Nucléaire, CERN) in Geneva, Switzerland. The accelerator was built into the existing tunnel of its ancestor, the Large Electron–Positron Collider LEP. With a circumference of 27 km, at a depth ranging from 50 m to 175 m, it is the largest particle accelerator ever built. The LHC itself is the last step of a 5-level acceleration system. The linear accelerator (LINAC) accelerates the protons to 50 MeV. In three intermediate steps, the Proton Synchrotron Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), the energy of the protons gets further increased to 450 GeV before being injected into the LHC where they get accelerated to the final energy. A schematic overview of the LHC and its associated experiments is shown in Fig. 6.1. Superconducting magnets with magnetic fields of up to 8.3 T keep the protons on their circular path. The proton beams are organized in bunches and bunch trains. Each beam is split into 1380 bunches which consists of in average 1.7×10^{11} protons. Due to limitations of the PS, these bunches are not continuously filled into the LHC but are combined in so called bunch trains. During 2012, every bunch train contained up to 36 bunches with a spacing of 50 ns between each of them. A gap of 345 ns is between two consecutive bunch trains. The complete acceleration process takes about 30 min. After that, another 20 min are needed to test the stability of the beams and prepare them for physics collisions. After roughly 20 h of data taking, the luminosity is decreased by 50% and a beam dump followed by a refill is more efficient for luminosity collection than proceeding with data taking.

Until 2012, the LHC provided proton–proton collisions with center-of-mass energies of $\sqrt{s} = 900$ GeV, 2.76 TeV, 7 TeV and 8 TeV. For Run II in 2015, a center-of-mass energy of 13 TeV is planned while the design energy is 14 TeV. The LHC was designed to reach an instantaneous luminosity of $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1} while the highest luminosity obtained so far during data taking has been $\mathcal{L} = 8 \times 10^{33}$ cm^{-2}s^{-1}. Six detectors are using the proton–proton collisions provided by the LHC: Two multipurpose detectors, ATLAS [38] and CMS [39], ALICE [40], specialized on lead–lead collisions, LHCb [41], designed for high precision measurements in b-physics and two smaller experiments, TOTEM [42] and LHCf [43]. TOTEM predominantly measures the inclusive proton–proton cross section while LHCf studies the behavior of cosmic-rays-like particle cascades.

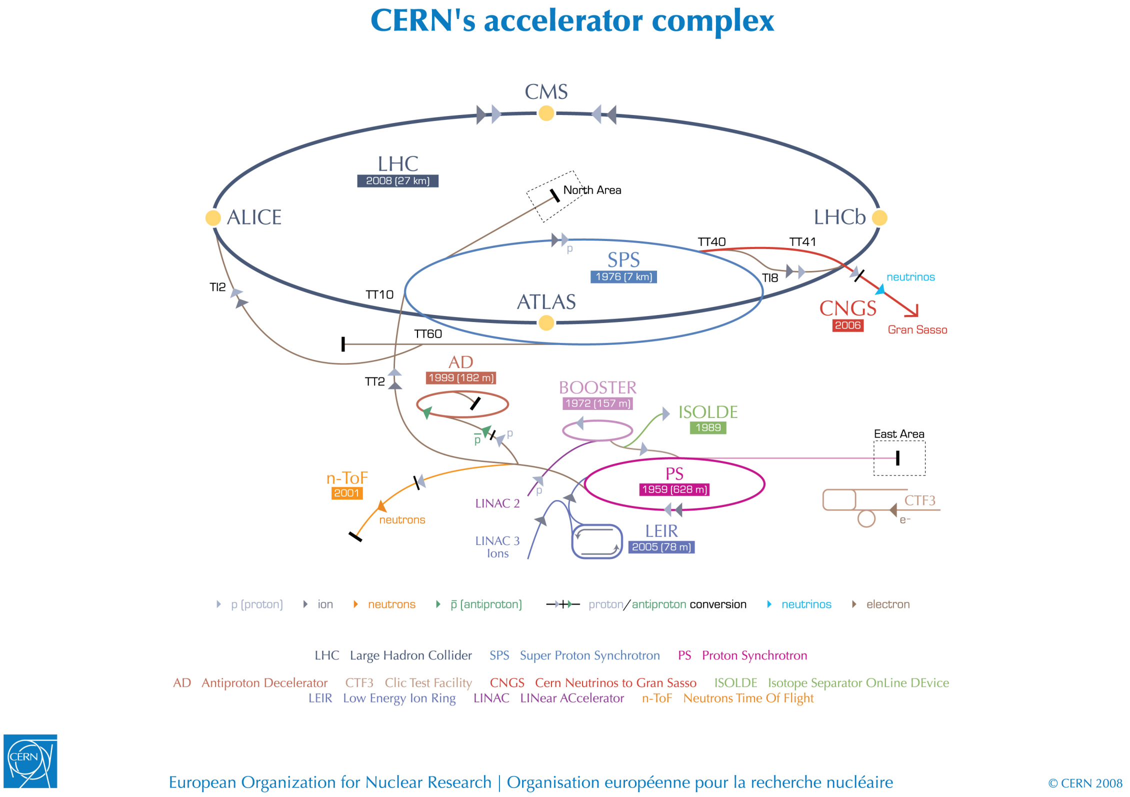


Figure 6.1: Schematic overview of the LHC accelerating complex and its associated experiments.

7

THE ATLAS EXPERIMENT

The ATLAS detector (**A** **T**oroidal **L**HC **A**pparatu**S**) is one of the two multipurpose detectors. With a length of 45 m and a diameter of 25 m, ATLAS is the largest detector at the LHC. Over 3000 physicists from over 38 countries are part of the ATLAS collaboration. One of the main goals of ATLAS, the search for a Higgs-Boson, was accomplished in July 2012. But many other, equally interesting studies were and are performed within the collaboration. End of 2012, Run I successfully finished with a recorded integrated luminosity of $\int \mathcal{L}\,dt \approx 25\ \mathrm{fb}^{-1}$, collected mostly at 7 TeV and 8 TeV. In 2015, Run II will start in which collisions at a center-of-mass energy of $\sqrt{s} = 13$ TeV are going to be recorded and analyzed.
This section describes the ATLAS experiment, including the detector, the trigger system and computing.

7.1 THE DETECTOR

The ATLAS detector covers close to 4π in solid angle around the interaction point and is split into three main subdetectors: The inner (tracking) detector, the calorimeters and the muon system. An overview of the detector is given in Fig. 7.1.

THE COORDINATE SYSTEM OF THE ATLAS EXPERIMENT To describe processes in the detector, a cylindrical coordinate system, fitted to the shape of the detector, has been introduced. The z-axis points along the beam pipe and the (x, y)-plane is perpendicular to it. The positive x-axis points to the center of the accelerate ring and the positive y-axis is directed upwards to the surface. The azimuthal angle ϕ lies inside the (x, y)-plane and the polar angle is θ. Instead of the polar angle, the pseudorapidity η is used and is defined as:

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

The correlation between η and θ is shown in Fig. 7.2. $\eta = 0$ lies within the (ϕ, r)-plane and $\eta = \infty$ points along the z-axis. The rapidity y, is defined as

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right),$$

where p_z is the momentum component along the z-axis. For massless particles, the rapidity is equal to the pseudorapidity. The rapidity difference between two particles is lorentz-invariant under boosts along the z-axis. The pseudorapidity however is often preferred, as particle production is roughly constant as a function of η.

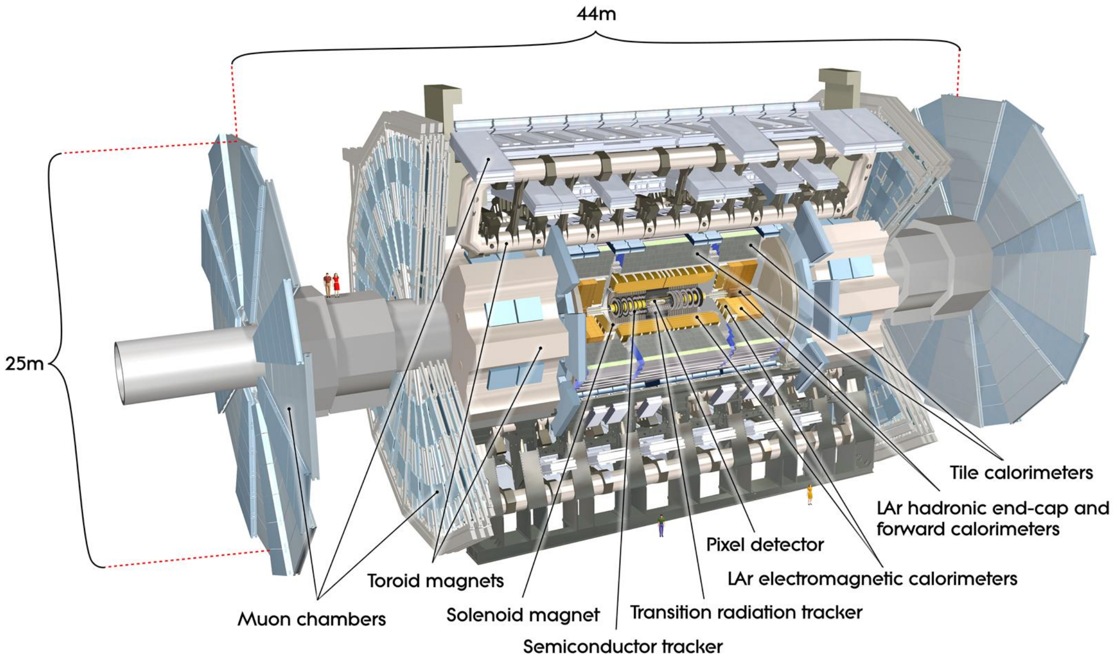


Figure 7.1: A detailed computer-generated image of the ATLAS detector and its subdetectors [44].

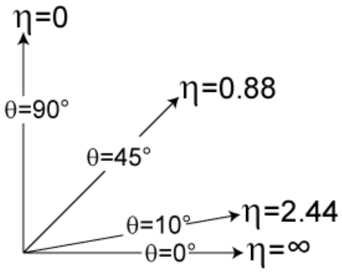


Figure 7.2: Pseudorapidity η for different values of the polar angle θ [45].

7.1.1 *Inner detector*

The most inner part of the detector is the tracking detector, also called inner detector (ID). It is designed to measure charged particles with a high spatial precision and an excellent momentum resolution down to 100 MeV. Tracks above $p_T = 400$ GeV are used to reconstruct the interaction vertices, define the primary vertex among these, and identify secondary vertices from short-lived particles. The inner detector is placed inside the central solenoid, which provides a 2 T-axial magnetic field. Within $|\eta| < 2.5$, the bent trajectories of charged particles are reconstructed and the sign of charge as well as the momentum are measured. A schematic view of the inner detector is shown in Fig. 7.3. The ID is built out of three subdetectors.

The pixel detector The most inner part is the pixel detector, built out of three concentric cylinders, the *barrel*, and three disks of silicon pixel sensors in the more forward region, the *end-caps*. The innermost layer, only 50.5 mm away from the beam line, mainly contributes to the identification of short-lived particles such as B hadrons and therefore is referred to as B-layer. Each module has a position resolution of 10 μm in the $R-\phi$ plane and 115 μm in $z(R)$ for the central (endcap) region.

The SCT detector The second part of the inner detector, the semi-conductor-tracker SCT, is made out of silicon microstrip layers and is organized as the pixel detector in a barrel and end-caps. The barrel is composed of four double-layered concentric cylinders and the end-caps have nine double-layered discs each. The dual layers are used to measure the z coordinate of hits. Therefore the SCT, as well as the pixel detector, provides discrete space-points measurements. The resolution in the $R-\phi$ plane is 17 μm and 580 μm in $z(R)$ for the central (endcap) region.

The transition radiation tracker The transition radiation tracker (TRT) consists of gaseous straw tube elements interleaved with transition radiation material. Charged particles emit transition-radiation photons induced by the transition radiation material. The intensity is proportional to $\sim E/m$. The mass dependency is used to distinguish between electrons (with a $m \approx 0$) and heavier objects. The accuracy is decreased compared to the pixel and SCT detector (130 μm in $R-\phi$, no information in z), but with an average of 36 hits per track, it provides continuous tracking information and improves the track reconstruction. In contrast to the two other subdetectors, the TRT covers only the central region with $|\eta| < 2.0$, split into a barrel ($|\eta| < 1$) and two endcaps covering $1 < |\eta| < 2$.

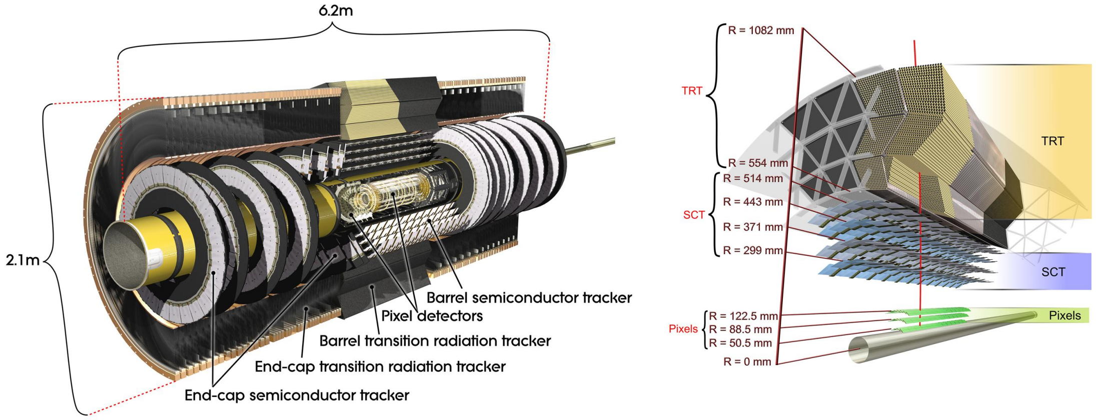


Figure 7.3: Overview of the ATLAS Inner Detector. Cut through the (y, z)-axis on the left and a profile of a segment in (r, ϕ) on the right [44].

7.1.2 *Calorimeter*

The calorimeters are used for energy measurements of incident particles. ATLAS uses two different calorimeter types. An electromagnetic calorimeter is used to detect mainly electromagnetic interacting particles and the hadronic calorimeter measures strong interacting particles. A schematic view of the calorimeters is shown in Fig. 7.4.

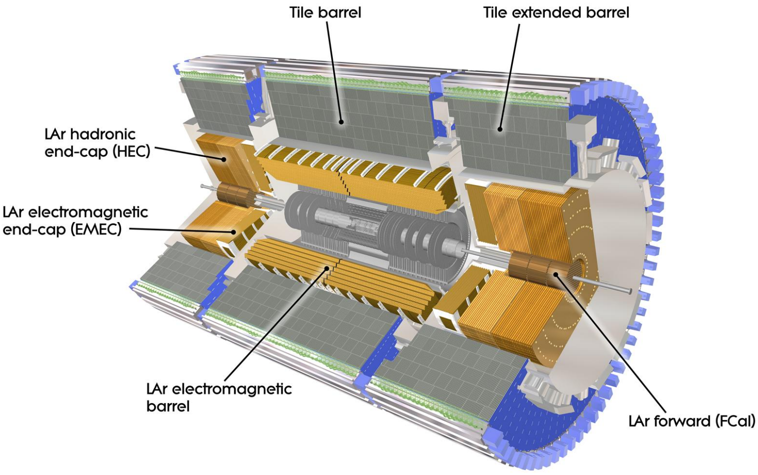


Figure 7.4: Schematic view of the ATLAS calorimeter [44].

The electromagnetic calorimeter

The electromagnetic calorimeter is designed to precisely measure electrons, positrons and photons. It uses liquid argon as active and lead as passive material. Liquid argon was chosen due to its linear behaviour, its stability of response and radiation-hardness. Particles traversing the LAr produce ionization electrons, which drift to electrodes. There, they produce electrical currents proportional to the deposited energy. The original signal has a triangular shape, with a rise time of about 1 ns and a decay time of several hundreds ns. As the next proton–proton happens only 25 ns later, the signals overlap. To compensate this, the pulse is reshaped into a bipolar shape, shown in Fig. 7.5. In the reshaped signal, the rise is followed by a steeper decay into an undershoot (negative voltage). Due to the high bunch-crossing rate, the next pulse arrives, before the previous one is completely decayed. The undershoot is introduced to com-

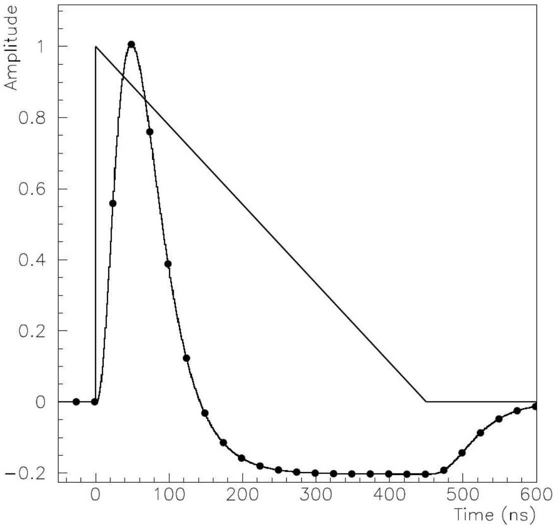


Figure 7.5: Amplitue vs time for the triangular pulse in a LAr barrel cell and the reshaped pulse form [44].

pensate for remaining signals from previous bunch-crossings.
The calorimeter is split into two parts, a central barrel, covering $|\eta| < 1.475$ and two end-caps covering $1.375 < |\eta| < 3.2$. In the most central region, $|\eta| < 1.8$, an instrumented argon layer, the presampler, is installed to measure the energy loss of electrons and photons before reaching the electromagnetic calorimeter. The barrel covers 22 and the end-caps 24 radiation lengths. The radiation length X_0 is defined as the travel distance of a particle, until it lost $1/e$ of it's energy ($E(x) = E_0 e^{-x/X_0}$).
In order to provide a full coverage in ϕ, an accordion geometry has been chosen for the absorbers and electrodes. This geometry also allows to have several active layers in depth. This and a very fine granularity in η and ϕ, e. g. in the second longitudinal layer with 0.025×0.025 (0.025×0.025) in the barrel (end-caps $1.375 < |\eta| < 2.5$) in $\Delta\eta \times \Delta\Phi$, provides good resolution. The innermost layer has a finer granularity in η of 0.003×0.1, e. g. to identify neutral pions decaying into photons. The fractional resolution has been measured in test beams, resulting in $\sigma/E = 10\%/\sqrt{E} \oplus 0.7\%$. The more forward region (up to $|\eta| < 4.9$) is covered by a forward calorimeter, which is described later.

The hadronic calorimeter

The outermost part of the calorimeter is the hadronic calorimeter, designed to measure hadronic particle showers. The end-caps use liquid argon as active material and copper as an absorber and cover the region $1.5 < |\eta| < 3.2$. The central part, split into two barrels, the central barrel and the extended barrel, are tile calorimeters, using steel as passive material and scintillator as the active medium. The central barrel is located in the region $|\eta| < 1.0$, and the extended barrel covers the range $0.8 < |\eta| < 1.7$ and provides a granularity of 0.1×0.1 in $\Delta\eta \times \Delta\phi$. They are composed of three layers in r, 1.5, 4.1 and 1.8 interaction lengths thick in the central barrel and 1.5, 2.6 and 3.3 in the extended barrel. The combined interaction lengths of the calorimeter are shown in Fig. 7.6. The calorimeter is designed to cover at least 10 radiation lengths over the the full detector. This is sufficient to stop most hadrons inside the calorimeter, punch-through (particles escaping the detector), however, is observed. The energy calibration corrects for the losses due to punch through. The hadronic calorimeter again covers the full 2π in ϕ. The designed energy resolution is $\sigma/E = 50\%/\sqrt{E} \oplus 3\%$. The readout of the tile calorimeter is realized via wavelength shifters and photomultiplier tubes (PMT). The power supplies of these PMT are exposed to high radiation. This causes the power supply to trip from time to time, resulting in dropping their voltage to zero. A mechanism to apply corrections for tripped modules or to remove parts of the data, if three consecutive modules tripped simultaneously, was applied.

The forward calorimeter

To cover the region close to the beam pipe ($3.1 < |\eta| < 4.9$), two additional calorimeters have been installed. Due to the high radiation it is excessed to, the forward calorimeter requires to have a high radiation-robustness. Liquid argon is used as active material again, as it provides the needed radiation-hardness. The electromagnetic part of the forward calorimeter uses copper and the hadronic one wolfram as absorber materials.

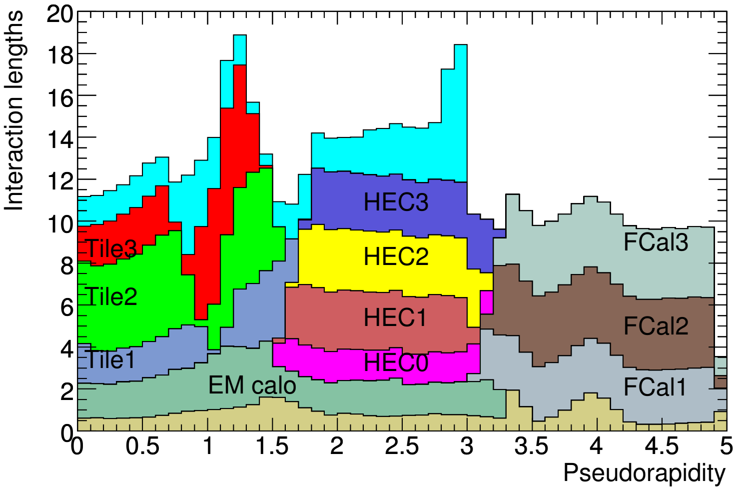


Figure 7.6: Cumulative amount of material in units of interaction length, as a function of $|\eta|$. The uppermost contribution is the muon spectrometer and the brownish part at the bottom represents the total amount of material in front of the first layer. [44].

Detector component	Resolution	η coverage
Tracking	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$	± 2.5
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2
Hadronic calorimetry (jets)		
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < |\eta| < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7

Table 7.1: Resolution and η coverage of the detector components of the ATLAS detector [44].

7.1.3 *The muon system*

The outermost and biggest part of the detector is the muon system. The muon system is placed into a 2.5 T magnetic field, produced by 8 barrel toroid coils. The energy and momentum get measured by monitored drifted tubes (MDT) and Cathode strip chambers (CSC). Particles out of jets which are not fully absorbed by the calorimeters, so called punch through jets, can also be detected with the help of the muon system.

A summary of the coverage and resolution of the subdetectors is given in Table 7.1

7.2 THE TRIGGER SYSTEM

At the design luminosity, the proton–proton interaction rate is about 1 GHz, while event data recording is designed to work at 200 Hz. The rate has to be reduced by a factor of $5 \cdot 10^6$. This is accomplished by the trigger system. Every event has to pass a three level filtering system to be written into the storage. A wide variety of trigger requirements are applied into the trigger system, trying to select jets, electrons, photons, τ etc. The functionality of the trigger system is presented by the example of jet triggers, as used in the analysis presented later.

The L1-trigger, a hardware based trigger, selects events based on calorimeter and muon system information with reduced granularity. Central jet triggers①, like those used in the jet analysis presented later on, require a minimum amount of transverse energy deposited inside the calorimeter in a window of 0.8×0.8 in $\Delta\eta \times \Delta\phi$. Events are stored into buffers until they get rejected or accepted. The L1 trigger reduces the rate to 75 kHz. The L2-trigger applies a basic jet-finding algorithm to the region of interest, identified and passed to L2 by the L1 trigger. The second level trigger is software-based and runs on PC farms to reduce the rate further to $\approx$ 1 kHz. Before finally accepting the event, the event filter (EF) reconstructs the event with the full detector information. An anti-k_t jet algorithm with a radius parameter of $R = 0.4$ searches for jets and a last threshold on transverse energy is applied to reduce the rate to the aimed value of 200 Hz.

The minimum bias filter, as used in the charged-particle event shape analysis, requires at least one hit in the *minimum bias trigger scintillators*, which are placed at $z = \pm 3.56$ m around the beam pipe covering $2.09 < |\eta| < 3.84$.

Even after filtering the events, the rates are too high to store. Therefore, prescales are applied on every trigger level, reducing the number of events accepted by a constant factor. With decreasing luminosity during the run, the prescales are decreased to secure a constant rate of about 200 Hz. An overview over the single trigger levels, the rate and the decision time is shown in Fig. 7.7. The application of jet triggers in the analysis is described in Sec. 15.1.

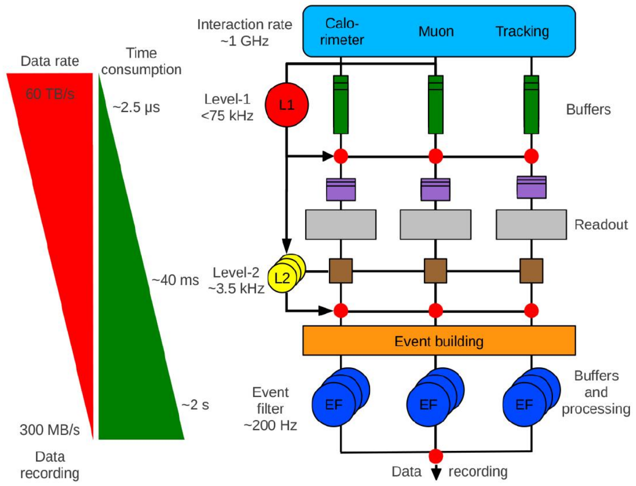


Figure 7.7: Overview of the three level trigger system in ATLAS. The selection proceeds top to bottom in the diagram [44].

7.3 COMPUTING

The amount of data stored by the ATLAS experiment is about 1 million gigabyte per year. The raw data, after being accepted by the trigger system, is written to CERNs

---

① restricted to $|\eta| < 3.2$

own computer cluster (TIER-0) as RDO (Raw Data Objects). These files get distributed into the *GRID*. The grid consists of four tiers of computer clusters. CERNs own computer cluster, the TIER-0, stores all the raw data and starts the reconstruction of the events. It is supported by the TIER-1 computer clusters in event reconstruction and Monte Carlo generation. The output are so called ESDs (Event Summary Data) and AODs (Analysis Object Data). ESDs represent an intermediate step of the reconstruction, AODs already containing information about reconstructed and calibrated physics objects such as jets, electrons and muons and are suitable as input for physics analyses. Smaller derivations of the AODs, the DPDs (Derived Physics Objects) are distributed further to TIER-2 computing centers to be available for analysis. DPDs do not hold all information of the AODs, but selected objects for physics analysis. From there, local copies are downloaded to TIER-3 computer clusters for analysis purposes. Further transformations in other data formats, e. g. D3PDs, being DPD information stored in a ROOT [46] container format, are available for an easier access and readout of the data. All TIER-1 and TIER-2 computer clusters provide spare CPU time for the reprocessing of data and Monte Carlo generation. This system guarantees a broad accessibility to the data and shared CPU performance to accomplish the task of handling these big amounts of data.

8

JETS IN ATLAS

An example of a dijet event inside the detector is shown in Fig. 8.1. Two intersections of the detector are shown on the left side. The high number of energy depositions in the calorimeters and reconstructed tracks disclose the complexity in reconstructing jets. The separation between particles originating from the hard-scattered parton and particles from the underlying event or pile-up is a major issue of jet reconstruction. The top middle plot in Fig. 8.1 shows the reconstructed interaction vertices, represented by colored squares. With 11 vertices, the event is even below the mean value of around 20 simultaneous interactions. The average dijet event therefore is even more complex. Calibrating the detector signals to reconstruct the real jet energy is another big task of the jet calibration. The procedure of obtaining jets from detector signals will be discussed in the following chapters. At first, the concept of topological cluster, the input objects for the jet clustering algorithm, will be presented, followed by a description of the energy calibration. Sec. 8.3 shows how possible fake jets and badly reconstructed jets are identified.

8.1 TOPOLOGICAL CLUSTER

Jets can be built out of many different detector signals. One of the commonly used ones are topological clusters. Calorimeter jets are preferred, as they include neutral particles, in contrast to track jets, and have a better energy resolution at high energies. Using pure cell information would introduce noise into the jet energy, which is not trivial to correct for. The concept of topological cluster, presented in this section, reduces the noise and therefore is used as input for the jet finding.

Particles entering the calorimeter produce particle showers and their energy depositions are spread over many calorimeter cells. One approach to combine all energy depositions originating from the same particle are topological clusters [47]. The topological cluster algorithm is split into two steps. In the *Cluster Maker*, calorimeter cells get combined based on their signal-to-noise ratio (*SNR*). The electromagnetic noise as a function of η is shown in Fig. 8.2. At high luminosities (as in the 2012 data taking), the noise at higher $|\eta|$ is dominated by pile-up. Using the SNR decreases the noise in the topological clusters. After the cluster maker, the *Cluster Splitter* divides overlapping showers and decides how to split the energy of shared cells. Both steps use the same scheme. Cells with a significant energy above the expected noise are used as seed cells. Starting from these seeds, adjacent cells are iteratively added to the cluster, if they are above a lower threshold on the SNR. To catch the tails of the shower, all direct neighbour cells are added in a last step.

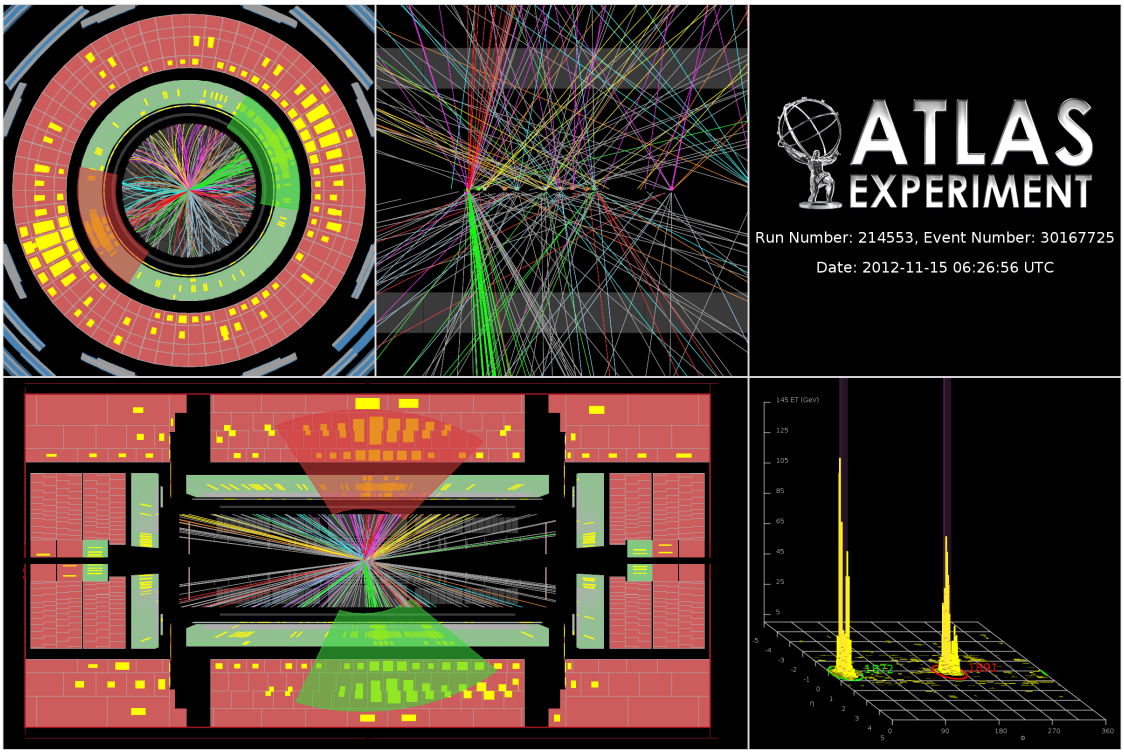


Figure 8.1: Event display of a dijet event. The leading jets have an energy of ~ 1.9 TeV and the dijet mass is 3.76 TeV. The top left plot shows a projection in the $r - \phi$ plane. The top middle plot shows the reconstructed vertices as well as the measured tracks in the event. The bottom parts show the $z - \eta$ projection on the left and energy depositions in a $\eta - \phi$ coordinate system on the right.

The Cluster Maker

The cluster maker performs in three parts:

Finding seeds Search for cells with SNR $> t_{\text{seed}}$ and add them to a *seed list*. Each seed cell forms a *proto-cluster*.

Finding Neighbours The seed list gets ordered by SNR. Starting with the seed with the highest SNR, all neighbour cells are considered. If they are above the neighbour threshold $t_{\text{neighbour}}$ and have not been in the seed list yet, the cell is added to a *neighbour seed list* and included in the adjacent proto-cluster. If the cell connects two proto-clusters, they get merged. Cells with $t_{\text{cell}} <$ SNR $\leqslant t_{\text{neighbour}}$ are added to the proto-cluster which provides the neighbour with the highest SNR, but do not get added to the neighbour seed list. After processing all seed cells, the neighbour seed list becomes the new seed list and the old one is discarded. This procedure is repeated until no more neighbours with a SNR above $t_{\text{neighbour}}$ can be found.

Finalize The resulting proto-clusters are sorted in descending order of E_T and converted to *clusters*.

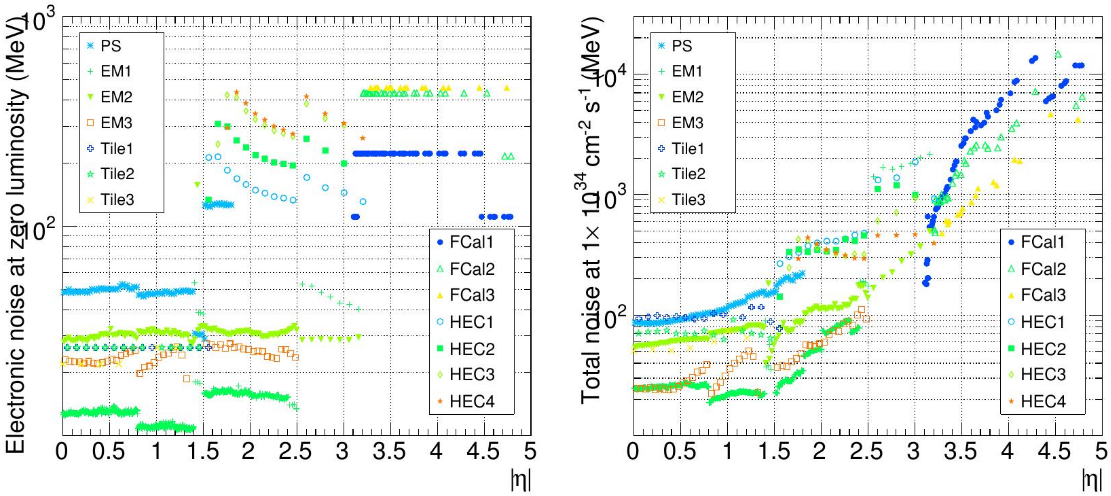


Figure 8.2: Per-cell electronics noise (left) and total noise at high luminosity (right), in MeV, for each calorimeter layer[47].

The Cluster Splitter

The cells in the clusters are used as input for the cluster splitter to separate overlapping particle shower, which operates in the four following steps:

Finding local maxima All cells satisfying the following three criteria are defined as local maxima:

- $E > 500$ MeV
- No adjacent cell with a higher energy
- At least 4 adjacent cells within the parent cluster

Each local maximum will form exactly one topocluster at the end of the cluster splitter, hence the number of topoclusters is determined after this step.

Finding neighbours The finding neighbour step from the cluster maker is performed again with the following changes:

- Only cells clustered in the Cluster Maker are considered
- No thresholds are applied ($t_{\text{neighbour}} = t_{\text{cell}} = 0$)
- Cells which connect at least two proto-clusters are added to a shared cell list and the sharing happens in the next step

Shared Cells Cells from the original cells list which were not added to any proto-cluster during the "finding neighbours" routine are used to iteratively expand the shared cells using the "finding neighbours" recipe. The energy of the so formed shared cluster is split between the two adjacent proto-clusters with the most energetic neighbour (in case the shared cell connects more than two proto-clusters). The energy is not equally distributed among the two clusters but is divided by the following function:

$$w_1 = \frac{E_1}{E_1 + rE_2}, \quad w_2 = 1 - w_1, \quad r = \exp|d_1 - d_2| \tag{8.1}$$

where $E_{1,2}$ are the energies of the two proto-clusters which share the cell and $d_{1,2}$ are the distances from the proto-clusters to the shared cell (in units of EM-shower scales (5 cm)).

Finalize Each local maximum has been converted into a proto-cluster. The remaining parent clusters without a local maximum are transformed into proto-clusters.

The proto-clusters found after all steps have been completed are the final topological clusters.

Even though the topocluster algorithm is designed to suppress noise, clusters can be formed from noise only. The expected number of these clusters is given by Eq. (**??**).

$$N_{\text{clus}} = N_{\text{cells}} \sqrt{\frac{2}{\pi}} \int_{t_{\text{seed}}}^{\infty} e^{-t^2/2} dt \tag{8.2}$$

Fig. 8.3 shows the average cell granularity in η, drawn in blue, and the predicted number of noise-only topological clusters, shown as red dots. The agreement between the two distributions shows the direct correlation between granularity and noise. The topoclusters used in this thesis as input for jet clustering algorithms were produced using $t_{\text{seed}} = 4$, $t_{\text{neighbour}} = 2$ and $t_{\text{cell}} = 0$ for the clustering algorithm.

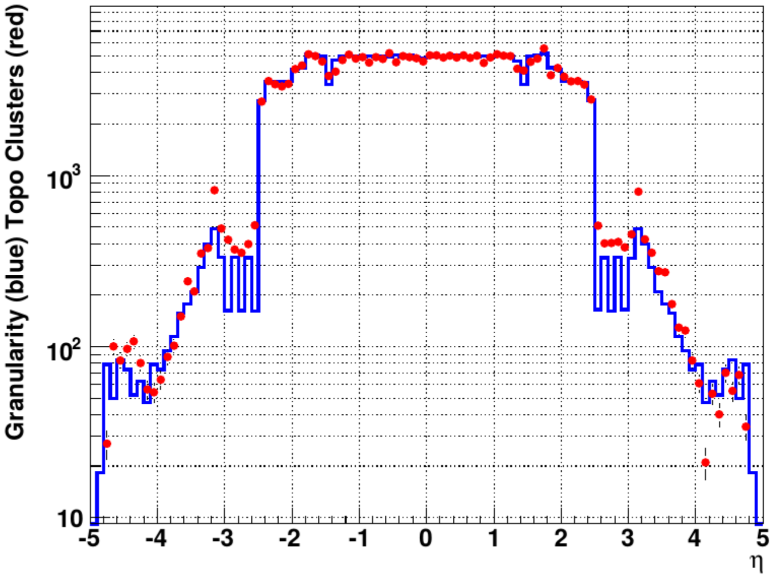


Figure 8.3: Average cell granularity (number of cells per $\Delta\eta = 0.1$) taken from the detector geometry (blue histogram) and as calculated from the distribution of topological clusters in simulated noise-only events (red points).[47]

Topological cluster energy calibration

The translation of a calorimeter signal to an energy deposition was derived to reconstruct electrons at their *true* energy. This calibration is called *EM calibration*, as it is calibrated to electromagnetic energy depositions. As jets consist of a wide variety of particles, using the EM calibration cannot reproduce the jet energy correctly, if built from calorimeter signals. Therefore, a new calibration scheme, the *local cluster weighting calibration, LCW*, was introduced. Based on cluster shape variables like the longitudinal cluster depth and the cluster isolation, a probability p to be hadronic is assigned to every topocluster. Hadronic energy correction weights, w_{HAD}, were derived using a Monte Carlo simulation for charged and neutral pions. The weights are dependent of the calorimeter layer $|\eta|$, topocluster energy, and cell energy density. Each topocluster energy then gets weighted with

$$w_{\mathrm{HAD}} \cdot p + w_{\mathrm{EM}}(1-p) \tag{8.3}$$

where $w_{\mathrm{EM}} = 1$.
In addition, effects like out-of-cluster energy depositions, *true* energy depositions not passing the SNR threshold during topoclustering, and dead material (energy loss in inactive detector material) are applied during the LCW calibration. For each of the three effects, energy correction weights for hadronic and electromagnetic topoclusters are derived using Monte Carlo and a correction weight is applied in analogy to the initial weighting:

$$w_{\mathrm{HAD}} \cdot p + w_{\mathrm{EM}}(1-p)$$

where p is again the probability of the topocluster to be hadronic and $w_{\mathrm{HAD}}/w_{\mathrm{EM}}$ the derived energy correction weights.

8.2 JET ENERGY CALIBRATION

After calibrating the topocluster as described above, the jet finding algorithm runs on the different available jet constituents (simulated particles, tracks and topological clusters) as indicated in the scheme in Fig. 8.4. The reconstructed calorimeter jets need to be calibrated further, as the basic EM or LCW calibration cannot efficiently correct the energy back to their *true* energy. The aim of the jet calibration is to correct the reconstructed jet 4-vector to match the *true* jet. The calibration needs to correct for several effects.

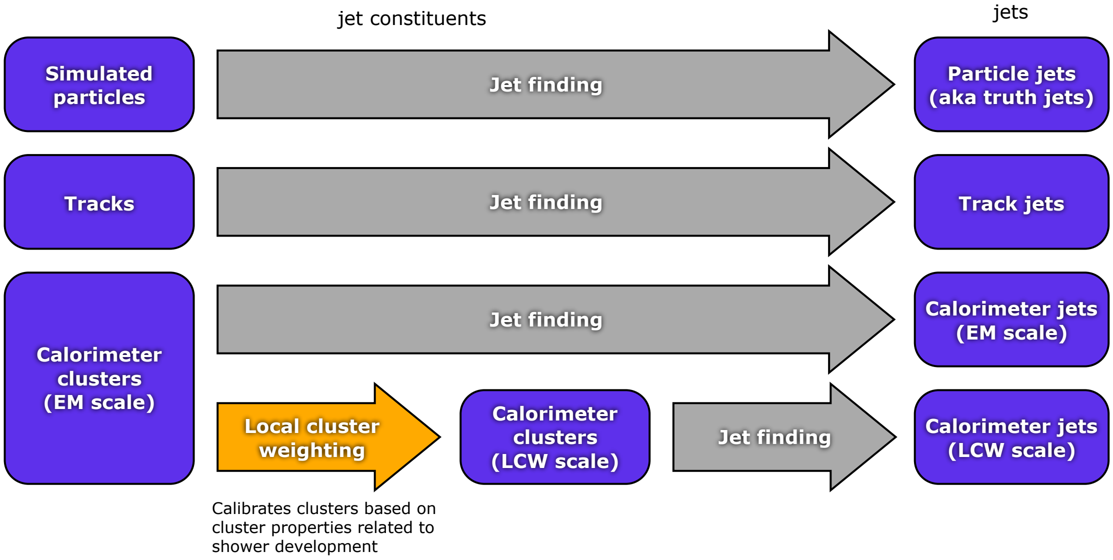


Figure 8.4: Schematic jet finding and calibrating scheme [48].

- **Non-compensating calorimeter:**
  The energy inside the calorimeter is measured due to ionization induced by the incoming particles and the particle showers they produce. However, not all energy is converted into ionization energy. Energy depositions due to binding energy losses from nuclear spallations and recoil cannot be detected.

- **Pile-up and underlying event:**
  In addition to the particles originating from the true jet, particles from the underlying event and pile-up events deposit energy in the calorimeter and add energy to the jet. This additional energy needs to be subtracted from the measured jet energy.

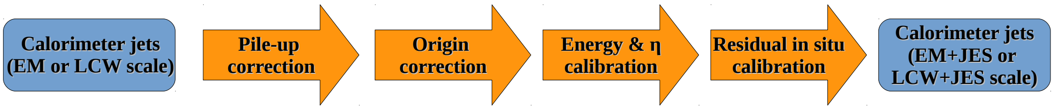
- **Punch through:**
  Particles penetrating the full detector will not be fully reconstructed and therefore, the measured jets need to be corrected for the energy loss due to the punch-through.

- **Origin correction:**
  The direction of the jet cannot be precisely extracted from calorimeter information only. The direction of the jet gets reconstructed by pointing the energy depositions to the primary vertex, i. e. the vertex with the highest scalar sum of squared track transverse momenta $\sum p_{T,trk}^2$, associated with this vertex.

- **Out-of-cone losses:**
  Not all particles originating from the parton jet get included into the reconstructed jet. This is mainly due to hadronization effects, radiations under large angles and the magnetic field inside the inner detector, which bends charged particles trajectories out of the reconstructed jet. The out-of-cone loss is of course highly dependent on the jet size parameter R.

- **Dead material:**
  Energy depositions in inactive detector material.

The correction of the energy is divided into four steps as indicated below:



In a first step, the pile-up contribution to the jet energy gets subtracted by a jet area based correction method. After that, the directions of the jets are corrected to point to the primary vertex. This step does not effect the jet energy. After applying a Monte Carlo based energy correction, an additional data driven correction is applied.

8.2.1 *Pile-up corrections*

One of the biggest challenges for the jet energy calibration for the 2012 data taking is the big amount of pile-up. With a mean of ~ 21 simultaneous pp-interactions (Fig. 15.3,

8.1) and a LHC bunch spacing of 50 ns (time between colliding proton bunches), the jet energy needs to be corrected for energy depositions originating from other pp collisions.
Two effects contribute to the pile-up:

In-time pile-up Additional pp-interactions in the same bunch crossing produce additional signals inside the detector.

Out-of-time pile-up Electronic signals from previous bunch crossing could be still present in the detector. The mean signal length of many detector components is longer than the bunch spacing (Fig. 8.5).

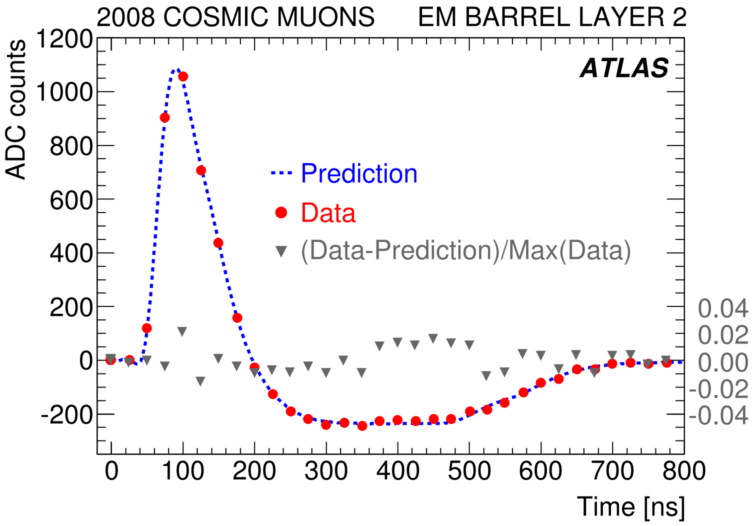


Figure 8.5: Typical single ionization pulse in a cell of layer 2 of the barrel of the LAr calorimeter. Red dots represent data, blue dots show the prediction and the grey triangles the relative difference (data - prediction)/S_{max}, on the scale shown on the right side of the plot (normalized to the data) [49].

Two variables are used to measure pile-up. The number of reconstructed primary vertices (N_{PV}) is a direct measurement of in-time pile-up, while the average number of interactions per bunch crossing ($\langle\mu\rangle$) is a measure of both, in-time and out-of-time pile-up, as it is derived from the instantaneous luminosity L via

$$\langle\mu\rangle = \frac{L \times \sigma_{\text{inel.}}}{N_{\text{bunch}} \times f_{\text{LHC}}} \tag{8.4}$$

with the total inelastic pp cross-section $\sigma_{\text{inel.}}$ and the average frequency of bunch crossings in the LHC $N_{\text{bunch}} \times f_{\text{LHC}}$.
Fixing N_{PV} while measuring $\langle\mu\rangle$ gives information about the out-of-time pile-up. The correction for pile-up is divided into two steps. At first a jet-by-jet correction based on jet areas [50] is applied, followed by a residual correction parametrized in N_{PV} and $\langle\mu\rangle$ [51].

Jet area based pile-up subtraction

For the jet area based pile-up subtraction, the additional energy from pile-up is modelled as a uniform underlying p_T density in the event. If the p_T density as well as the jet area are known, the additional p_T contribution from pile-up to jets can be computed and subtracted. Calculating this p_T density for each event as well as the jet areas allows a jet-by-jet correction.

THE JET AREA In order to calculate the jet area, infinitesimally soft *ghost particles* ($p_T = \mathcal{O}(10^{-100}\ \text{GeV})$), are uniformly added to the event before the jet finding. Ghost particles do not change the result of the jet finding for IRC-safe jet clustering algorithms, as the anti-k_t algorithm used here, but the number of ghosts associated to a jet is a measure of the jet area.
Exploiting the fact that every ghost particle has a four-momentum g_i, the jet area four-momentum A_{jet} is defined as the sum of the ghost four-momenta, normalized by the transverse momentum density of the ghosts $\nu_g \langle g_t \rangle$:

$$A_{\text{jet}} = \frac{1}{\nu_g \langle g_t \rangle} \sum_{g_i \in \text{jet}} g_i \tag{8.5}$$

In practice, the ghost particle four momentum is constant. For the limit of small jets, this area is equivalent to the *scalar area*: the number of ghosts associated to the jet divided by the ghost number density.

THE MEDIAN p_T DENSITY ρ The median p_T density measured separately in every event is defined as

$$\rho = \text{median} \left\{ \frac{p_{T,i}^{\text{jet}}}{A_i^{\text{jet}}} \right\} \tag{8.6}$$

where i enumerates all jets found by the anti-k_t jet clustering algorithm with a distance parameter $R = 0.4$, using topological clusters with positive energy in the central region ($|\eta| < 2.0$) as input. The jets emerging from the hard scattering could be part of the list. As ρ should be a measure of the pile-up activity, the median is used instead of the mean value to suppress the influence of these large p_T jets. As shown in Fig. 8.6, the median p_T density ρ drops towards 0 for the outer regions of the detector ($|\eta| > 2.5$). The main reason for this behaviour is the broader granularity in this region. A broader granularity leads to a decrease in noise-only cluster (Fig. 8.3) and therefore a decreased pile-up sensitivity (better noise suppression). The small ρ values are therefore not caused by the absence of pile-up, but by the incapability of measuring the contribution consistently. This shows that the pile-up contribution is uniform in η. The results in Sec. 13.1 provide evidence for a mainly spherical p_T distribution in ϕ in minimum bias events, what in-time pile-up mostly is made of. That proves the assumption, that ρ is constant in η and ϕ. In order to have a good estimator for pile-up in the forward region, ρ gets extrapolated from the central region to the forward region up to $|\eta| < 4.9$. What could also be seen from Fig. 8.6 is that for events with nearly no pile-up (red solid line with $\langle \mu \rangle = 1$) ρ is not zero, as expected for a pile-up only sensitivity of ρ. That could be explained by the fact that ρ is sensitive to the underlying event activity. Therefore, the presented method does not only correct for pile-up, but also for underlying event

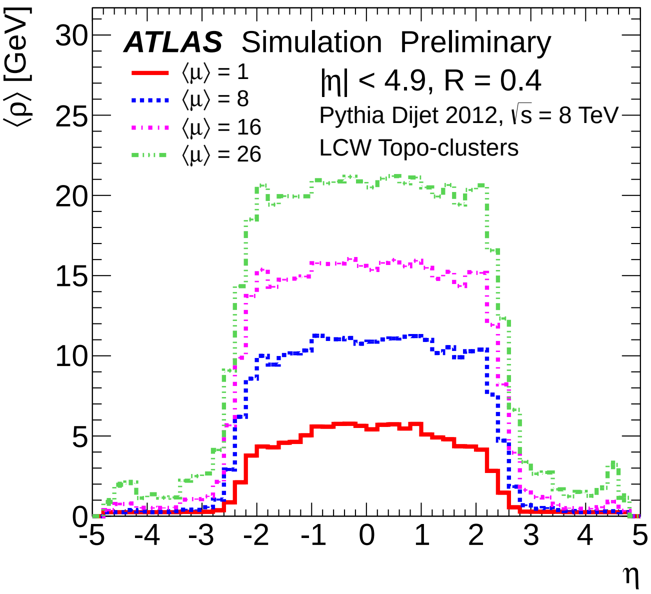


Figure 8.6: The median p_T density, ρ, as a function of η using a sliding η interval of width 0.7 in the range $|\eta| < 4.9$ for four different $\langle\mu\rangle$ values [52].

contributions.

To obtain the corrected jet transverse momentum, the p_T density is multiplied by the jet area and the product gets subtracted of the transverse momentum of the jet:

$$p_T^{\text{jet,corr}} = p_T^{\text{jet}} - \rho \cdot \mathcal{A}, \tag{8.7}$$

where $\mathcal{A}$ is the transverse component of A_{jet}.

Residual correction

The jet area method is not able to correct for all pile-up contributions. Since jets used to compute ρ are only built from topological clusters with positive energy, the jet area method cannot correct for negative energy contributions from out-of-time pile-up and noise. Fig. 8.5 shows the bipolar pulse shape of a detector signal. The negative overshoot can produce negative energy contributions if not compensated by pile-up contributions. In addition, the extrapolation of ρ from $|\eta| < 2.0$ to the full detector range ($|\eta| < 4.9$) can introduce an over- or undercorrection in the forward region of the detector.

The residual correction is obtained for a reference pile-up condition ($N_{PV} = N_{PV}^{\text{ref}}$, $\langle\mu\rangle = \langle\mu\rangle^{\text{ref}}$) and all other pile-up conditions are first extrapolated back to the reference condition and then, they are corrected using the known correction for the reference point.

The extrapolation to the reference point is derived using Monte Carlo, studying the dependencies of jet p_T over N_{PV} in bins of η, $\langle\mu\rangle$, and $p_{T,\text{truth}}$. The correction $\mathcal{O}$ is

$$\mathcal{O}(N_{PV}, \langle\mu\rangle, \eta_{\text{det}}) = \alpha(\eta_{\text{det}}) \cdot \left(N_{PV} - N_{PV}^{\text{ref}}\right) + \beta(\eta_{\text{det}}) \cdot \left(\langle\mu\rangle - \langle\mu\rangle^{\text{ref}}\right) \tag{8.8}$$

where α and β are taken from Monte Carlo studies (such as in Fig. 8.7). α and β are η-dependent as the effect of pile-up in different detector regions is different due to e.g.

the granularity (as discussed before) and different pulse shapes in different detector regions. The choice of reference points is arbitrary. For 2012, mean values of $N_{PV}^{ref} = 11.8$ and $\langle\mu\rangle^{ref} = 20.7$ were used.

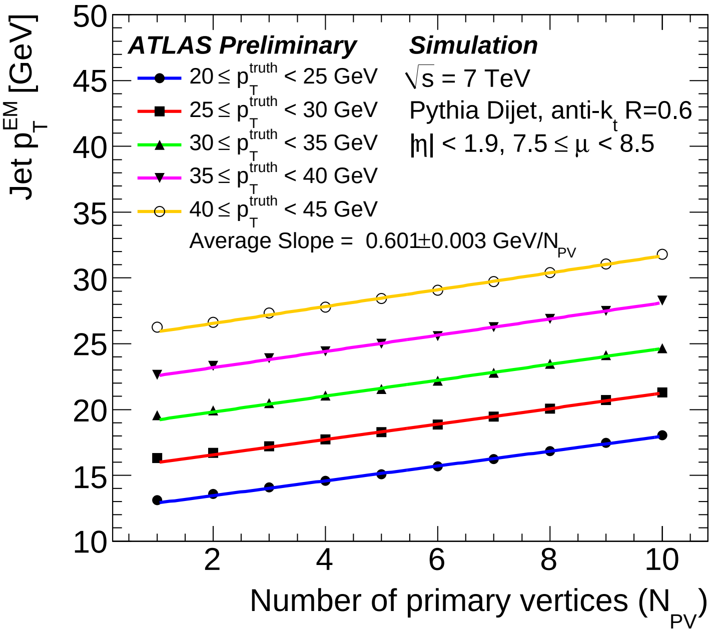


Figure 8.7: The average reconstructed transverse momentum (p_T) on EM scale for jets in MC simulation, as function of the number of reconstructed primary vertices (N_{PV}) and $7.5 < \langle\mu\rangle < 8.5$, in various bins of truth jet p_T, for jets with R = 0.6[51].

The dependency of the reconstructed jet p_T on N_{PV} and $\langle\mu\rangle$ as a function of $|\eta|$ is shown in Fig. 8.8. It is visible that the jet area correction is not able to correct for out-of-time pile-up, as the dependency even increases after applying the correction. In the forward region, the effect of out-of-time pile-up is more pronounced compared to the central region. Due to the broader granularity of the detector in that region, the probability that a jet signal and a negative signal from out-of-time pile-up overlap is increased and therefore, the effect of out-of-time pile-up is increased. For in-time pile-up, the result shows nearly no dependency on $|\eta|$.
After applying both corrections, the reconstructed jet p_T is mostly independent from pile-up, which shows the success of the correction method.

Systematic uncertainties

Four sources of systematic uncertainties were taken into account. The residual correction has a dependency on jet p_T^{true} for both, the N_{PV} and the $\langle\mu\rangle$ term. The corrections were derived at a reference point of $p_T^{true} = 25$ GeV. The dependency follows a logarithmic function, therefore the uncertainty is constructed as

$$\Delta_{p_T}^{p_T} = \pm\left(\alpha_{N_{PV}} \cdot \left(N_{PV} - {N_{PV}}^{ref}\right) + \alpha_{\langle\mu\rangle} \cdot \left(\langle\mu\rangle - \langle\mu\rangle^{ref}\right)\right) \cdot \log\left(\frac{p_T}{25\text{ GeV}}\right) \tag{8.9}$$

where $\alpha_{N_{PV}}$ and $\alpha_{\langle\mu\rangle}$ were extracted from the logarithmic fits on the remaining dependencies on $\langle\mu\rangle$ and N_{PV} as a function of p_T^{true} shown in Fig. 8.9.

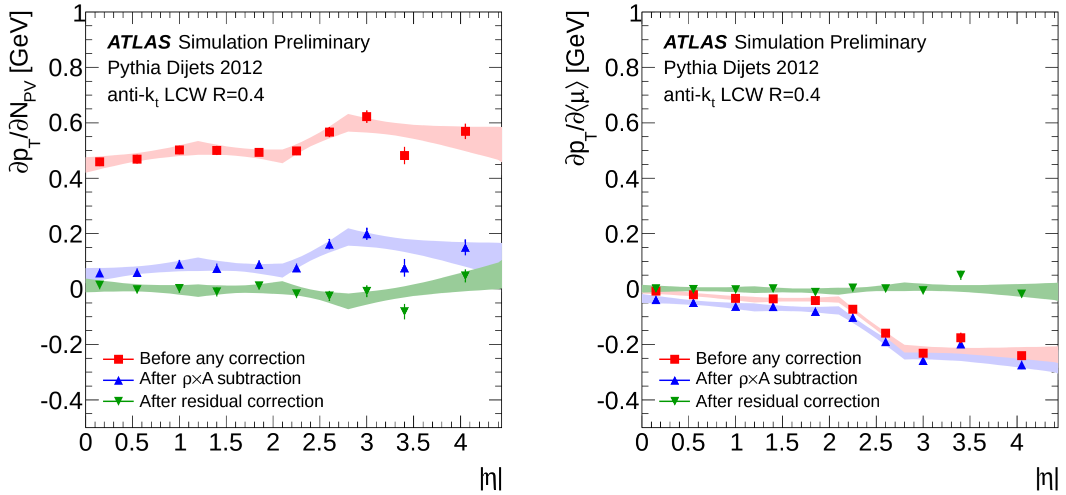


Figure 8.8: Reconstructed jet p_T (anti-k_t, R=0.4, LCW scale) dependence on in-time pile-up (N_{PV}, left) and out-of-time pile-up ($\langle\mu\rangle$, right) before any correction, after $\rho \cdot$area subtraction, and after the residual correction. The dependence is shown in bins of jet $|\eta|$ and fitted as in the residual correction itself. The error bands show the 1σ confidence intervals of the fits. The dependence was obtained by comparison to truth particle jets in simulated dijet events, and corresponds to a truth-jet p_T of 25 GeV [52].

Closure tests using dijet, γ+jet and Z+jet events in data and Monte Carlo generated events were performed to assign uncertainties to $\frac{\partial p_T}{\partial N_{PV}}$ and $\frac{\partial p_T}{\partial \langle\mu\rangle}$. The values of $\Delta\left(\frac{\partial p_T}{\partial N_{PV}}\right)$ and $\Delta\left(\frac{\partial p_T}{\partial \langle\mu\rangle}\right)$ are taken as the maximum of the non-closure, the difference between data and Monte Carlo, and the statistical uncertainty of the latter. These uncertainties propagate through the offset correction, resulting in the two uncertainty components $\Delta_{p_T}^{\langle\mu\rangle}$ and $\Delta_{p_T}^{N_{PV}}$, which are defined as:

$$\Delta_{p_T}^{\langle\mu\rangle} = \pm\Delta\left(\frac{\partial p_T}{\partial\langle\mu\rangle}\right) \cdot \left(\langle\mu\rangle - \langle\mu\rangle^{\text{ref}}\right) \tag{8.10}$$

$$\Delta_{p_T}^{N_{PV}} = \pm\Delta\left(\frac{\partial p_T}{\partial N_{PV}}\right) \cdot \left(N_{PV} - N_{PV}^{\text{ref}}\right) \tag{8.11}$$

Potential mis-modeling of the sample dependence introduced by the use of ρ for pile-up subtraction was addressed by a systematic uncertainty. The difference between data and Monte Carlo was studied using dijet, Z + jet and γ + jet events. It was found that the dependence of ρ on $\langle\mu\rangle$ follows a linear function as shown in Fig. 8.10.

While the slope is consistent among the different samples, the intercept varies. This was evaluated using

$$\Delta_I = \left(\rho(\langle\mu\rangle^{\text{ref}})^{\text{Ch1}}_{\text{data}} - \rho(\langle\mu\rangle^{\text{ref}})^{\text{Ch2}}_{\text{data}}\right) - \left(\rho(\langle\mu\rangle^{\text{ref}})^{\text{Ch1}}_{\text{MC}} - \rho(\langle\mu\rangle^{\text{ref}})^{\text{Ch2}}_{\text{MC}}\right) \tag{8.12}$$

where $\rho(\langle\mu\rangle^{\text{ref}})$ is the value of ρ for $\langle\mu\rangle = \langle\mu\rangle^{\text{ref}} = 20.7$ taken from a linear fit and Ch1 and Ch2 represent all possible combinations of the channels specified above. The systematic shift due to Δ_I then is

$$\Delta_{p_T}^{I} = \pm\max(\Delta_I) \cdot C^{\text{JES}} \cdot \pi R^2. \tag{8.13}$$

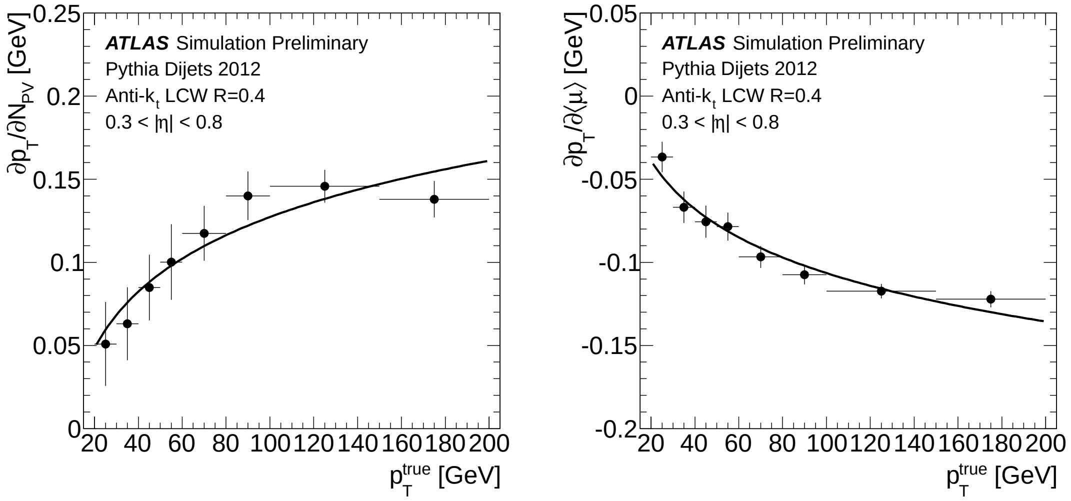


Figure 8.9: Dependence of the coefficients of the residual pile-up correction on the p_T of matched truth particle jets in simulated dijet events, for anti-k_t $R = 0.4$ jets at the LCW scale and within the pseudorapidity range $0.3 < |\eta| < 0.8$. The points are consistent with a logarithmic functional form. The first point in both plots represents the nominal correction [52].

The average jet size for anti-k_t jets is πR^2 and C^{JES} is the average jet energy calibration factor for jets at $p_T = 25$ GeV.
The combined systematic uncertainty is typically below 2% in the central region and increases up to 5.6% at higher η values.

8.2.2 *Origin correction*

4-vectors of the topological clusters, making them point back to the primary hard-interaction vertex. All kinematic variables of the topological cluster are recomputed and the jet properties reconstructed again. While the jet energy does not change, the angular resolution and the p_T response are slightly improved ($< 1\%$). The origin correction has not been implemented yet at the time writing the thesis and therefore is not applied. The impact on the final results can be estimated to be negligible.

8.2.3 *Monte Carlo based corrections*

The jet response is calculated in Monte Carlo. The response $\mathcal{R}$ is defined as the ratio of the measured energy over the true energy.

$$\mathcal{R} = \frac{E_{\text{jet,reco}}}{E_{\text{jet,truth}}} \tag{8.14}$$

This response was measured as a function of η and p_T and is applied to the reconstructed jets after the pile-up subtraction. Fig. 8.11 shows an example of the response for a Monte Carlo Simulation with a center of mass energy of $\sqrt{s} = 7$ TeV. The higher the jet energy, the better the response. The energy loss due to out-of-cluster energy deposition is more pronounced for low energetic jets. A dead material distribution and detector

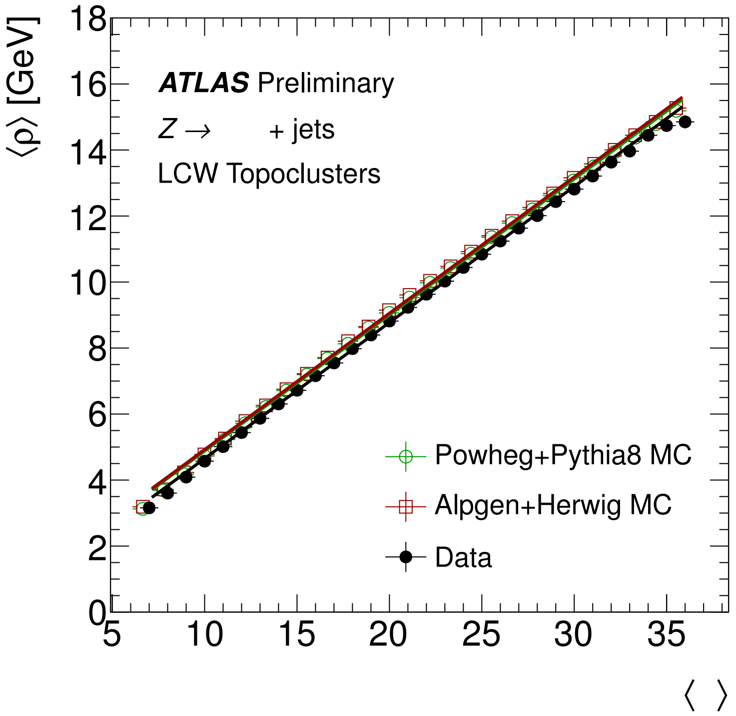


Figure 8.10: Dependence of ρ on ⟨μ⟩ in Z + jet events for data and two different Monte Carlo predictions using LCW topoclusters [52].

structure is clearly visible in the $|\eta_{\mathrm{det}}|$ dependency. Regions with high amount of dead material and/or crack regions in the detector can be observed as a drop in the jet response. The consequences of the LCW calibration is a highly improved jet response in dependency of E_{truth}, due to the hadronic energy and out-of-cluster corrections, as well as in $|\eta_{\mathrm{det}}|$, mainly due to the dead material correction.

8.2.4 *In situ calibration*

In addition to the Monte Carlo based correction, an in-situ calibration is applied to data only. The momentum conservation in the transverse plane is used to derive these corrections. In a first step, jets in the central region are calibrated using events with well calibrated objects like photons and Z bosons recoiling against a single jet. Due to the low statistics of highly boosted Z and γ, multijet balance method is used for jets with a high transverse momentum: Events with one high energetic jet recoiling to many smaller jets are used.

In a second step, the now well calibrated central jets are used to calibrate jets in the more forward part of the detector. Events with a dijet topology are selected and forward jets get balanced to central jets.

To account for the difference of data and Mote Carlo, the correction factor of each in-situ calibration step is constructed as a double ratio:

$$\frac{\langle p_T^{\mathrm{jet}}/p_T^{\mathrm{ref}}\rangle_{\mathrm{data}}}{\langle p_T^{\mathrm{jet}}/p_T^{\mathrm{ref}}\rangle_{\mathrm{MC}}} \tag{8.15}$$

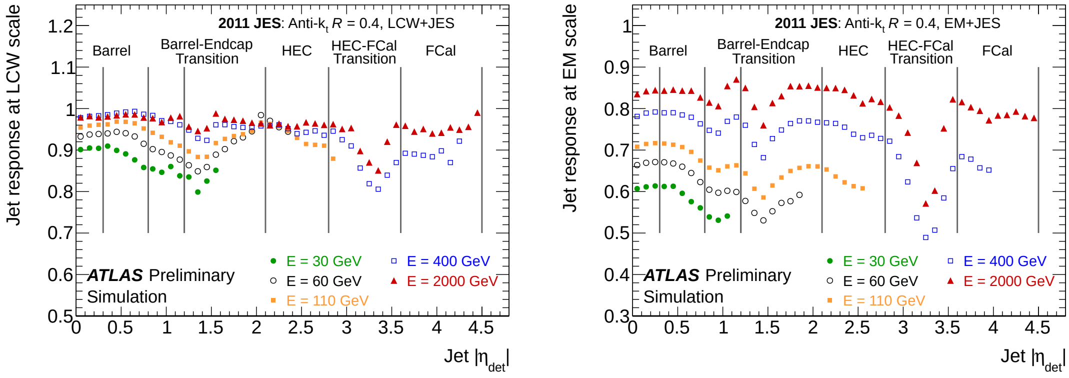


Figure 8.11: Average jet response as a function of η for different E_{truth} bins for the LCW (left) and EM (right) calibration scheme and anti-k_t jets with a radius parameter of $R = 0.4$ for a center of mass energy of $\sqrt{s} = 7$ TeV[48].

Z + JET DIRECT BALANCE Events in which one Z boson, decaying into two electrons①, recoils against exactly one isolated jet are used to calibrate the recoiling jet. To reconstruct the Z boson, electrons with a transverse energy above 20 GeV within the region covered by the inner detector ($|\eta| < 2.47$, excluding a small transition region in the detector) and an invariant mass $M_{e^+e^-}$ between 80 GeV and 116 GeV are selected. To suppress imbalances due to radiations, events with a second jet with a transverse momentum above 20% of p_T^Z, but at least 10 GeV, are rejected. If the jet is in the central region covered by the inner detector and has a $p_T < 50$ GeV, the jet vertex fraction *JVF* is requested to be above 0.25 to exclude jets from pile-up. The jet vertex fraction is defined as the percentage of track transverse momentum associated with the jet_i and the primary vertex PV_j:

$$\text{JVF}(\text{jet}_i, \text{PV}_j) = \frac{\sum_k p_T(\text{track}_k^{\text{jet}_i}, \text{PV}_j)}{\sum_n \sum_l p_T(\text{track}_l^{\text{jet}_i}, \text{PV}_n)} \tag{8.16}$$

where k counts all tracks originating from PV_j matched to the jet_i. n runs over all primary vertices in the event and l runs over the tracks originating from PV_n matched to the jet_i. A summary of the event selection is given in Table 8.1

The $p_T^{\text{jet}}/p_T^{\text{ref}}$ ratio is computed in bins of p_T^{jet} and $\Delta\phi(\text{jet}, Z)$. The mean value for each p_T^{jet} bin is obtained using a χ^2 fit of a Poisson distribution, extended to non-integer values and multiplied by a turn-on curve to account for the p_T^{jet} threshold. All events then get extrapolated to a back-to-back configuration ($\Delta\varphi = \pi$) with a constant function to determine the mean balance in every p_T^{jet} bin independent from $\Delta\phi(\text{jet}, Z)$.
The systematic uncertainties on the Z + jet direct balance are described in the following:

Extrapolation to $\Delta\varphi = \pi$ Instead of a constant function, a linear fit was used. The difference between the two extrapolations is taken as a systematic uncertainty.

Second jet veto/radiation suppression The cut on the second jet veto was varied up and down, and the average of the absolute value of the deviations is taken as a systematic uncertainty.

---

① Electrons denotes for electrons and positrons.

Variable	Selection	Description
e_1, e_2	$E_T^{e_{1,2}} > 20$ GeV	electron
	$80 \text{ GeV} < M_{e^+e^-} < 116$ GeV	pre-selection
	$\lvert\eta^{e_{1,2}}\rvert < 2.47$ excl. $1.37 < \lvert\eta^e\rvert < 1.52$	
leading jet	$\lvert\eta^{\text{jet}}\rvert < 0.8, \lvert\text{JVF}\rvert > 0.25$ if $p_T < 50$ GeV	jet pre-selection
$\Delta R_{j,e} = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}$	> 0.35 for anti-k_t $R = 0.4$ jets	isolation/topology
	> 0.5 for anti-k_t $R = 0.6$ jets	
$p_T^{\text{jet}_2}$	$< \max(0.2 \cdot p_T^Z, 10 \text{ GeV})$	radiation/topology
JVF(jet$_2$)	> 0.25 if $\lvert\eta^{\text{jet}_2}\rvert < 2.4$ and $p_T < 50$ GeV	pile-up removal
		for sub-leading jet

Table 8.1: Summary of event selection criteria for the Z + jet direct balance. $\Delta R_{j,e}$ is the distance between the jet and both electrons from the selection. jet$_2$ is defined as the highest p_T jet besides the leading jet.

Pile-up jet rejection The impact of additional proton–proton interactions is studied by varying the pile-up suppression cut on the JVF on the leading and the sub-leading jet.

Electron energy scale The measurement assumes a well reconstructed Z boson as a reference object to derive the JES. The Z p_T is reconstructed using two electrons, hence the uncertainty on the electron energy scale was propagated to the Z + jet measurement. The biggest effect comes from the material description of the EM calorimeter and presampler.

Monte Carlo generator The corrections are derived using Monte Carlo predictions generated with POWHEG and SHERPA. These generators use different theoretical models for all steps of the event generation and are therefore well suited to study the Monte Carlo generator dependency on the correction. This uncertainty component is the biggest one with an uncertainty of 1.5%.

Out-of-cone effects Even perfect Z + jet events do not have a perfect balancing between the jet and the Z boson on detector level. Two main effects are causing this: The Z boson is recoiling against a whole system of particles and the jet will not include all these particles in most of the cases. Contributions from the underlying event are added primarily to the jet, as electrons are rather narrow objects. The energy uncertainty due to out-of-cone effects and underlying event contributions were added as a systematic uncertainty.

The uncertainties described above are presumed to be uncorrelated and the components are therefore added in quadrature to calculate the combined uncertainty. The combined uncertainty is $1.5\% - 2\%$ for $35 \text{ GeV} < p_T^{\text{ref}} < 210$ GeV and increases up to $3 - 3.5\%$ for lower transverse momenta.

PHOTON + JET DIRECT BALANCE The same technique as for the Z + jet direct balance is used for the γ + jet direct balance. Differences are mainly in the event selection. As reference, photons identified by strict identification criteria, with a transverse momentum of $p_T^\gamma > 85$ GeV in the central region $\lvert\eta^\gamma\rvert < 1.37$ are selected. Additional cuts

Variable	Selection	Description
leading γ	$p_T^\gamma > 85$ GeV and $\lvert\eta^\gamma\rvert < 1.37$	photon pre-selection
$E_T^{\gamma \mathrm{Iso}}$	< 3 GeV	γ isolation
$E_T^{\gamma \mathrm{Cluster}}/(\sum p_T^{\mathrm{tracks}})$	$\in [0, 2]$ (single-track conversion)	jets faking photons
	$\in [0.5, 1.5]$ (double-track conversion)	
leading jet	$p_T^{\mathrm{jet1}} > 12$ GeV and $\lvert\eta^{\mathrm{jet1}}\rvert < 0.8$	jet pre-selection
$\Delta\Phi_{\mathrm{jet}-\gamma}$	> 2.9 radian	radiation suppression
p_T^{jet2}	$< 0.2 \cdot p_T^\gamma$	radiation suppression
JVF(jet2)	> 0.25 if $\lvert\eta^{\mathrm{jet2}}\rvert < 2.4$	pile-up removal

Table 8.2: Event selection for the γ + jet in-situ calibration. $E_T^{\gamma \mathrm{Iso}}$ is the pile-up and photon energy leakage corrected transverse energy contained in a cone of size $R = 0.4$ around the photon. jet2 is defined as the highest p_T jet besides the leading jet. $E_T^{\gamma \mathrm{Cluster}}/(\sum p_T^{\mathrm{tracks}})$ is the ratio of the photon candidate cluster transverse energy and the sum of transverse momenta of tracks associated with the cluster. $\Delta\Phi_{\mathrm{jet}-\gamma}$ is the azimuthal angle between the photon and the leading jet.

to reduce fake photons are applied, based on the photon transverse energy isolation and cluster–track energy matching. At most 3 GeV of transverse energy depositions in a cone with a radius of $R = 0.4$ around the photon is allowed after pile-up and photon energy leakage corrections. Photons reconstructed using an electron–positron pair are rejected if no tracks are found pointing from the reconstructed conversion vertex to the photon candidate cluster. To increase the photon reconstruction efficiency, converted photons with only one reconstructed track (single-track conversion) are used as well as converted photons with two tracks (double-track conversion). To reduce background from jets faking converted photons, a cut on the ratio of the photon candidate cluster transverse energy and the sum of transverse momenta of tracks associated with the cluster is applied.

The recoiling jet is requested to be in the central region $|\eta| < 0.8$ and must have a minimum transverse momentum of $p_T^\gamma > 12$ GeV. In order to minimize imbalances due to radiation, events containing any jet with $p_T^{\mathrm{jet}} > 0.2 \cdot p_T^\gamma$ or $|\Delta\phi|_{\mathrm{jet}-\gamma} > 2.9$ radians are rejected. All cuts are summarized in Table 8.2. The ratio $p_T^{\mathrm{jet}}/p_T^\gamma$ is calculated in bins of p_T^γ. The mean value in each bin is obtained with the same Poisson fit as for the Z + jet balance. The systematic uncertainties are listed below:

Soft radiation suppression The cuts on the sub-leading jet transverse momentum and on $\Delta\Phi_{\mathrm{jet}-\gamma}$ are varied independently. The maximum deviation from the nominal value is used as the uncertainty.

Background from jets Signal and background enhanced samples were studied. The uncertainty due to jets faking photons can be estimated as

$$(1 - P) \cdot (\mathcal{R}_{\mathrm{dijet}} - \mathcal{R}_{\gamma-\mathrm{jet}})/\mathcal{R}_{\gamma-\mathrm{jet}},$$

with the purity P of the $\gamma-$jet sample and the $\gamma-$jet balance in background events $\mathcal{R}_{\mathrm{dijet}}$ and in signal events $\mathcal{R}_{\gamma-\mathrm{jet}}$. The differences of the balance in signal and background were studied by loosening and tightening the photon identification criteria. The purity was measured in data with a sideband technique [53].

Variable	Selection	Description
N_{jets}	$\geqslant 3$	jet pre-selection
leading jet	$\lvert\eta^{\mathrm{jet1}}\rvert < 1.2$ and $p_T^{\mathrm{jet1}} > 25$ GeV	jet pre-selection
jet2, jet3	$\lvert\eta^{\mathrm{jet2}}\rvert, \lvert\eta^{\mathrm{jet3}}\rvert < 2.8$ and $p_T^{\mathrm{jet2}}, p_T^{\mathrm{jet3}} > 25$ GeV	jet pre-selection
β	> 2.14 radian	topology
α	< 0.3 radian	topology
p_T^{jeti}, jeti $\in$ recoil system	< 800 GeV	topology
p_T^{jet2}	$< 0.6 \cdot p_T^{\mathrm{recoil}}$	topology

Table 8.3: Event selection for the multi jet balance in-situ calibration. jet2 and jet3 are the sub-leading and second-sub-leading jet in the event. jeti is any jet of the multi-jet recoil system. $\alpha = \lvert\Delta\phi - \pi\rvert$, where $\Delta\phi$ is the azimuthal angle between the leading jet and the vector-sum of all recoiling jets. β is the azimuthal angle between the leading jet and the closest jet of the recoil system.

Photon energy scale The photon energy scale is derived from the electron energy scale. Each component of the photon/electron energy scale uncertainty had been varied up and down by 1σ and were then added in quadrature to get the full uncertainty. Again, the material description is the biggest uncertainty component.

Monte Carlo generator dependency The full γ – jet balancing is derived using PYTHIA and HERWIG. PYTHIA is used as the nominal value, since it is used to derive the JES corrections, and the difference to HERWIG is used as the systematic uncertainty.

Out-of-cone effects The description of out-of-cone effects is equivalent for Z + jet and γ + jet, as they only contribute to the jets, but not to narrow reference objects like photons. Therefore, the same systematic uncertainty as for the Z + jet balance is applied to the γ + jet balance.

The combined relative uncertainty is $\sim 1\%$ at intermediate p_T^{γ} and increases up to 1.3% at very low and very high p_T^{γ} values.

MULTI JET BALANCE MJB The Z/γ + jet calibration suffer from low statistics at high jet transverse momentum. To be able to calibrate jets at high p_T, events where many lower p_T jets, which are already calibrated by the Z/γ + jet calibration, recoil against one high p_T jet are used. The high-p_T jet is supposed to be isolated. To suppress dijet events with a soft radiation, the second-leading jet is only allowed to have 60% of the transverse momentum of the leading jet. The event selection is summarized in Table 8.3

To reach even the highest jet transverse momenta in data, this method is applied iteratively to increase the p_T range of the recoiling jets above the limits of the previous in-situ calibration steps. Jets calibrated using the MJB in a first step are used again in the MJB to balance a higher p_T jet. This is repeated recursively.
The systematic uncertainties can be divided into three components:

Energy calibration of the recoil system jets The uncertainties of all already applied calibration steps, being the γ/Z + jet direct balance and previous iterations of the MJB, were propagated through the multi jet balance as a systematic uncertainty.

Variable	Selection	Description
N_{jets}	$\geqslant 2$	jet pre-selection
$\Delta\phi(\text{jet1}, \text{jet2})$	> 2.5 radian	topology
$p_T^{\text{jet3,ctr}}$	$\max(0.4 \cdot p_T^{\text{avg}}, 12\text{ GeV})$	radiation/topology
$p_T^{\text{jet3,fwd}}$	$\max(0.4 \cdot p_T^{\text{avg}}, 10\text{ GeV})$	radiation/topology
$\text{JVF}(\text{jet3}^{\text{ctr}})$	> 0.25	pile-up suppression

Table 8.4: Event selection for the η intercalibration using dijet topologies. jet3^{ctr} is the second-sub-leading jet within $|\eta| < 2.4$ and jet3^{fwd} the second-sub-leading jet outside $|\eta| > 2.4$.

Monte Carlo generator dependency The Monte Carlo part of the MJB calibration was performed using PYTHIA and HERWIG. The difference of the two results was taken as the systematic uncertainty.

Event selection The event selection introduces a bias to the MJB calibration. Several sets of alternative cuts were applied and the differences added to the systematic uncertainty of the MJB.

The relative uncertainty of the multi jet balance is around 1% increasing to 2% below $p_T^{\text{recoil}} < 400 GeV$.

η INTERCALIBRATION USING DIJET TOPOLOGIES After calibrating the central jets using the Z/γ-jet and multijet balances, jets in the forward region are calibrated with the η intercalibration technique. Events with a dijet topology are used. The momentum conservation in the transverse plane is used once again to balance a dijet event topology, in which one of the jets is falling into the central region $|\eta| < 0.8$, called reference jet or jet_{ref} from here on. A summary of the event and jet selection is summarized in Table 8.4

The p_T balance is characterized by the asymmetry of the two leading jets $\mathcal{A}$:

$$\mathcal{A} = \frac{p_T^{\text{probe}} - p_T^{\text{ref}}}{p_T^{\text{avg}}} \tag{8.17}$$

where p_T^{avg} is defined as $p_T^{\text{avg}} = (p_T^{\text{probe}} + p_T^{\text{ref}})/2$. In the case of both jets falling into the reference region, each jet is used to probe the other jet and the resulting mean asymmetry is zero by construction.
The asymmetry is then used to calculate the calibration factor c:

$$\frac{p_T^{\text{probe}}}{p_T^{\text{ref}}} = \frac{2+\mathcal{A}}{2-\mathcal{A}} = 1/c \tag{8.18}$$

The correction gets calculated in bins of p_T^{avg} and η_{det}. In each bin, the average asymmetry is used to calculate the correction factor:

$$c_{ik} = \frac{2 - \langle\mathcal{A}\rangle_{ik}}{2 + \langle\mathcal{A}\rangle_{ik}} \tag{8.19}$$

where i indicates the η_{det} and k the p_T^{avg} bin. The downside of this method, called *central reference method* hereafter, is the decreasing statistical significance the further the

jets are separated in η.
To avoid that drawback, the method described above is extended by replacing *probe* and *reference* with *left* and *right* where left and right are defined as $\eta^{\mathrm{left}} < \eta^{\mathrm{right}}$. That way, the reference jet is not restricted to the central region and the η range can be extended iteratively, similar to the concept of the MJB. The asymmetry is consistently translated into

$$\mathcal{A} = \frac{p_T^{\mathrm{left}} - p_T^{\mathrm{right}}}{p_T^{\mathrm{avg}}} \tag{8.20}$$

In addition, a ratio of the *central reference method* correction factors of the *left*, c^{left}, and *right*, c^{right}, jet is defined.

$$\mathcal{R} = \frac{p_T^{\mathrm{left}}}{p_T^{\mathrm{right}}} = \frac{c^{\mathrm{right}}}{c^{\mathrm{left}}} = \frac{2+\mathcal{A}}{2-\mathcal{A}} \tag{8.21}$$

For each η^{left}-bin i, η^{right}-bin j, and p_T^{avg}-bin k, a response density distribution $\mathcal{R}_{ijk}$ and its average $\langle \mathcal{R}_{ifk} \rangle$ is computed. In order to obtain the relative correction factor c_{ik} for a given jet η-bin i and a fixed p_T^{avg}-bin k, the matrix of linear equations gets minimized:

$$S(c_{1k}, \ldots, c_{Nk}) = \sum_{j=1}^{N} \sum_{i=1}^{j-1} \left(\frac{1}{\Delta\langle \mathcal{R}_{ijk} \rangle} \left(c_{ik} \langle \mathcal{R}_{ijk} \rangle - c_{jk} \right) \right)^2 + X(c_{1k}, \ldots, c_{Nk}) \tag{8.22}$$

where N is the number of η-bins, $\Delta\langle \mathcal{R} \rangle$ the statistical uncertainty of $\langle \mathcal{R} \rangle$, and the function $X(c_{1k}, \ldots, c_{Nk})$ is chosen in a way that the trivial solution of $c_{ik} = 0$ is suppressed, while not changing the result otherwise. Performing the minimization for every p_T-bin k leads to the correction factors c_i, which are scaled in a way that the average correction factor in the reference region is equals unity. The results of each bin then is plotted as a function of $p_T^{J3}/p_T^{\mathrm{avg}}$ with the transverse momentum p_T^{J3} of the jet with the third highest p_T and the average transverse momentum of the two leading jets p_T^{avg}. An interpolation to $p_T^{J3}/p_T^{\mathrm{avg}} \to 0$ is performed with a linear fit to extrapolate to a zero-radiation configuration. These extrapolated values are used as the final correction factors the η-intercalibration. The systematic uncertainties were determined to be:

Monte Carlo generator dependency As before, the calibration step was performed using PYTHIA and HERWIG and the full difference between the two results is used as a systematic uncertainty.

Extrapolation to zero radiation Pseudo experiments are used to assess the uncertainty due to the extrapolation method. 2000 sets of pseudo experiments were generated on data, PYTHIA and HERWIG by smearing the response, using a Gaussian distribution with a mean equal to the nominal response and standard deviation equals to the systematic uncertainty on the response. For each pseudo experiment, the extrapolation was derived again. The RMS of the distribution of the gained correction factors is taken as the systematic uncertainties and all three values (for data, PYTHIA, and HERWIG) were added in quadrature for the full systematic uncertainty.

Radiation suppression The cuts to reduce the impact of radiations, the p_T cut on the second leading jet and the angular separation of the two leading jets $\Delta\phi(j1, j2)$

were varied separately and the shifts added in quadrature to obtain the combined uncertainty.

Jet energy resolution The systematic uncertainty of the jet energy resolution, described in Sec. 15.6.2, was propagated through the calibration method.

The combined uncertainty on the η-intercalibration is 6% in its maximum for $p_T = 35$ GeV and $|\eta_{det}| = 4$, decreasing to $\sim 1\%$ at high p_T and low $|\eta_{det}|$.

A more detailed description of the jet energy calibration and its uncertainties for $\sqrt{s} = 7$ TeV could be found in [48, 53]. The publication for the jet energy scale used in this analysis has not been released yet at the time of publishing this thesis. Most calibration methods are unchanged and the previous cited publications provide a more detailed description. The main difference is the pile-up removal. The new approach is described in detail in [52].

8.3 JET CLEANING

Jet reconstruction only includes minor information about the jet itself. Therefore, everything looking similar to a jet is reconstructed as a jet. To remove jets which have emerged from beam background, detector effects or cosmic rays (called *bad* jets), a jet cleaning is applied. In addition, jets falling into problematic (e.g. inactive or disabled) detector regions are rejected as a correct energy reconstruction is not possible in these cases. These jets will be referred to as *ugly* jets.

8.3.1 *Bad jet identification*

Different sets of variables remove fake jets from three sources. The cuts are presented in the following:

HEC spikes Sporadic noise-bursts in the end-caps of the hadronic calorimeter (HEC) can fake jets. In those cases, most of the jet energy is deposited in the HEC and a small number of cells hold a majority of the jet energy. To identify these jets, the following variables are used as discriminators:

HECf Fraction of jet energy stored in the HEC.

HECQ (HEC quality) Fraction of energy in HEC cells with big differences between measured and expected pulse shapes.

HECQmean Normalized jet quality computed as the energy squared cells mean HECQ.

negative E Negative energy in the jet as a negative energy is an indicator for noise fluctuations.

A jet is identified as emerging from a HEC spike if

$$(\text{HECf} > 0.5 \;\&\&\; |\text{HECQ}| > 0.5 \;\&\&\; \text{LArQmean} > 0.8) \parallel (\text{negative E} > 60\text{ GeV})$$

Coherent noise in the EM calorimeter Similar as noise-bursts in the HEC could lead to fake jets, noise in the EM calorimeter could also fake jets. The checked variables are to the ones for HEC spikes, but translated into the electromagnetic calorimeter:

EMf Fraction of energy stored in the EM calorimeter.

LArQ (LAr quality) Fraction of the jet energy in LAr cells with big differences between measured and expected pulse shapes.

LArQmean Normalized jet quality computed as the energy squared cells mean LArQ.

Pseudorapidity η Pseudorapidity of the jet.

To identify a jet as coherent noise, the following needs to be true

$$\text{EMf} > 0.95 \;\&\&\; |\text{LArQ}| > 0.8 \;\&\&\; \text{LArQmean} > 0.8 \;\&\&\; |\eta| < 2.8$$

Non-collision background Jets with a significant low energy deposition in the electromagnetic calorimeter and a small fraction of charged particles inside the jet (jets are estimated to have a mean charged fraction of about 2/3) are identified as non-collision background as well as jets where most of the energy is deposited in one calorimeter layer. The cleaning variables are therefore:

EMf as described above.

Chf The charged fraction *Chf* is the fraction of energy measured from tracks associated with the jet. The Chf can only be measured in the central region (coverage of the inner detector).

fmax The maximum energy fraction in one calorimeter layer.

The cuts were found to be optimized at

$$(\text{EMf} < 0.05 \;\&\&\; (\text{Chf} < 0.05 \;||\; |\eta| \geqslant 2.0)) \;||\; (\text{fmax} > 0.99 \;\&\&\; |\eta| < 2.0)$$

If a jet is identified as either of the above fake scenarios, it is marked as a *bad* jet.

8.3.2 *Ugly jet identification*

Modules which are known to be dysfunctional are masked and the energy in the cell is recovered dependent on the surrounding cell's energy. If the measured energy needs to be corrected by more than a factor of 2 or if a large amount of energy is deposited inside the gap in the detector between the TileCal barrel and end-cap, the jet is marked as *ugly*.

9

TRACK RECONSTRUCTION

Charged particles cannot be directly detected and identified in the inner detector. In order to reconstruct particle trajectories, the signals induced by the particles are used. The reconstruction happens in three steps:

Space point reconstruction and seed finding Hits from the pixel and SCT detectors are getting transfered into space points. Hits in the pixel detector directly give information on the coordinates of the hits (position of the module and 2-dimensional measurement of the hit position on the module), while space point reconstruction for SCT hits are not that straight forward as they only allow a 1-dimensional location measurement. The position of the module is used to determine the second dimension. To get the remaining coordinate, the combination of two SCT hits in a double-layer is used. The SCT detector is built in a sandwich module structure, in which two modules are rotated by 40 mrad compared to each other. If a particle produces hits in two modules of one double-layer in the SCT, the combination of the two signals gives information about the location of the hit in the (η, ϕ) direction. The advantage of this space point construction is an intrinsic noise suppression due to the requirement of two module hits.
After all space points have been reconstructed, all combinations of three space points from different detector layers are transfered into track seeds and the impact parameter d_0, the smallest distance to the beam spot, as well as the radius ρ of the circular trajectory are computed. Assuming a homogeneous magnetic field B parallel to the z-axis results, the transverse momentum of the track seed p_T can be calculated via

$$p_T[\text{GeV}] = \rho[\text{mm}] \cdot 3 \times 10^{-4} \cdot q[e] \cdot B[\text{T}]$$

with the charge of the particle q. An example how to reconstruct the mentioned parameter is shown in Fig. 9.1. Fig. 9.2 shows all track seeds in a generated $t\bar{t}$-event. The large number of seeds reflects the complexity of the track reconstruction and the need for a high space-point resolution. All seeds below a given transverse momentum threshold and above a maximum distance parameter d_0 are rejected. The remaining track seeds are used as input for the track candidate finding.

Constructing track candidates and track assessment The track seeds get extrapolated along the estimated trajectories. Hits along that extrapolation, plus a small window to take uncertainties into account, are added to the seed, now called track candidates. Hits can be assigned to several track candidates. To solve ambiguities, each track candidate receives a score, based on various properties. An overview

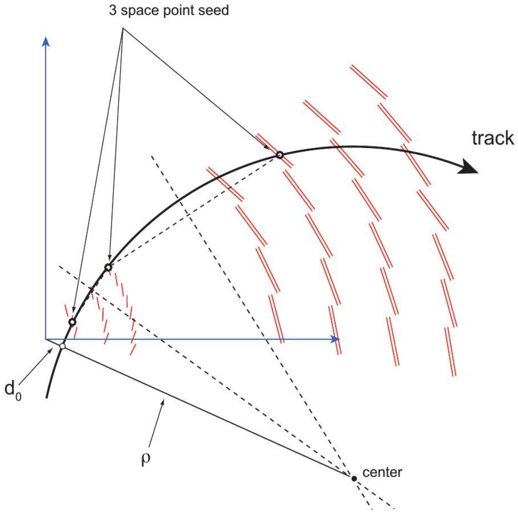


Figure 9.1: Possible track seed and extraction of the track seed parameter ρ, radius of the circular trajectory, and d_0, shortest distance to the beam spot, represented by the point of origin (0,0,0) [54].

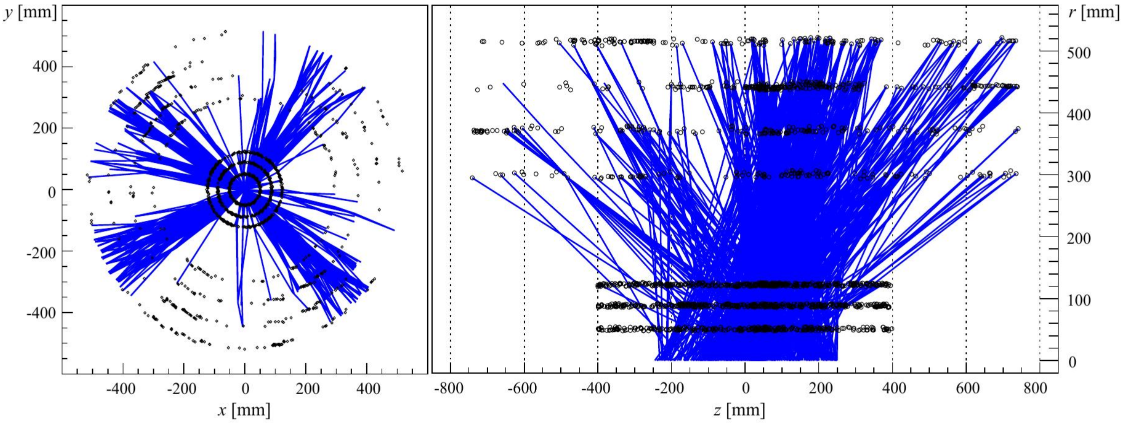


Figure 9.2: All track seeds in a simulated $t\bar{t}$-event [55].

of track characteristics changing the track score is given in Table 9.1. Overall, each hit along the trajectory increases the score, while missing hits in modules or sensors are penalized. The χ^2/n_{dof} of a helix-fit to the trajectory is taken into account to determine the final track score. If a hit is shared between several track candidates, it is assigned to the track candidate with the highest track score and removed from the others. After solving the ambiguities, basic track selections (min. number of hits in detector parts, missing hits in both subdetectors, ...) are applied and all tracks with a track score of 0 or below are removed. Only about 38% of all track seeds represent real tracks and are converted into track candidates. Fig. 9.3 shows the number of track candidates as function of pseudorapidity as well as the number of removed track candidates, split into reasons for rejection. Most track candidates are removed due to the track score and only above $|\eta| > 2$, a noticeable fraction of tracks is removed due to quality requirements.

Extension to the TRT After finding all track candidates in the silicon detector, they are extended into the TRT. The hits in the pixel and SCT detector remain unchanged.

Track characteristics	Detector	Effect on the track score
B-layer hole	pixel	strong penalty
Layer hole	pixel	penalty
Overlap hits	pixel, SCT	strong benefit
Sensor hole	SCT	weak penalty
Layer hole (module)	SCT	strong penalty

Table 9.1: Track characteristics that lead to benefits or penalties in the ATLAS silicon detector track score [55].

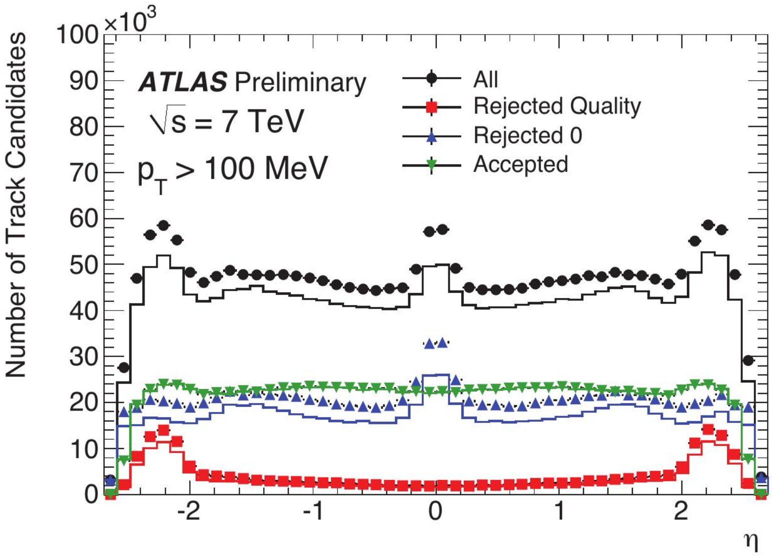


Figure 9.3: Number of track candidates as a function if η entering the ambiguity solver (black), rejected due to a track score of 0 or below (blue), rejected because of quality cuts (red) and accepted as resolved tracks (green). Data is shown in markers, Monte Carlo is drawn as a histogram. The p_T distribution of the accepted tracks candidates in the simulation have been reweighted and normalized to the accepted track candidates in data [54].

> Only hits in the TRT are added, where available, to improve the fit of the trajectory. The track parameters described above are reconstructed again using the new fit with increased precision.

The knowledge about the position of the single modules is crucial for the track finding. Therefore, they are frequently remeasured to ensure a high track resolution. Cosmic rays and tracks in pp collisions with the highest possible track scores are used for the alignment. The squared sum of the distances between hits and reconstructed trajectories is minimized over all used tracks. This results in a new measurement of the position of the detector modules. Assuming 6 degrees of freedom (3 translational and 3 rotational) for every module, about 35000 degrees of freedom have to be determined in the silicon detector.

Part IV

MEASUREMENT OF CHARGED-PARTICLE EVENT SHAPE VARIABLES IN INCLUSIVE $\sqrt{s} = 7$ TeV PROTON-PROTON INTERACTIONS

10

MOTIVATION

Event shape variables measure the geometrical distribution of the energy flow in collider experiments. They were used first in $e^+e^-$and ep deep inelastic scattering (DIS) experiments, e. g. for precise measurements of the strong coupling constant α_S [56–60], and later at the ISR [61] and Sp$\bar{\text{p}}$S [62, 63] at CERN to examine the emergence of jets, and at Tevatron [64] to study the underlying event. In addition, first studies have shown that some event shape observables are suitable to measure the CP eigenvalues of the Higgs boson [65] and are candidates to analyze supersymmetric particle decays as well as black-hole decays.

Recent adoptions to hadron colliders were made by Banfi, Salam and Zanderighi [66] to construct event shape variables which are invariant under Lorentz boosts along the beam axis. First studies have indicated that these newly constructed event shape variables are sensitive to the underlying event.

As mentioned before, the "soft" part of pp collisions, where soft refers to interactions with low momentum transfer between the scattering particles, are not yet completely understood and cannot be calculated within perturbative QCD. Measurements like the one presented below help to increase the understanding of these soft interactions at LHC center-of-mass energies.

Phenomenological models are used in Monte Carlo generators to describe the non-perturbative contribution. These models contain many parameters, which are initially unknown. Measurements sensitive to the underlying event are eligible to be used to tune the phenomenological models and thus constrain the free parameters of the models. ATLAS [67–70], ALICE [71–73], and CMS [74, 75] published several measurements sensitive to soft-QCD processes, and first Monte Carlo tunes using LHC data were published. The analysis presented in this thesis has been published [76] in order to provide additional information about non-perturbative QCD and to be available as input for Monte Carlo tuning outside the ATLAS collaboration. Further tuning of the Monte Carlo generators can improve the quality of existing predictions.

As Monte Carlo generators are widely used in many other analyses, including the search for new physics and precision measurements of the standard model, this analysis can improve indirectly many other ongoing and future measurements.

Three of the hadronic event shape observables proposed in [66] are used in this measurement: The transverse thrust, thrust minor and transverse sphericity. In addition to the inclusive distributions, the evolution of the three event shape variables as a function of the leading charged particle transverse momentum p_T^{lead}, charged particle multiplicity N_{ch} and summed transverse momentum $\sum p_T$ are presented and compared to different Monte Carlo predictions.

The event shapes were calculated from charged particle tracks. Studies [77] have shown that adding neutral particles, by using calorimeter information instead of tracks, does not change the distributions of the event shape variables used. However, the correction to particle level appears to be more complicated and would introduce large uncertainties. Therefore, tracks were preferred in this analysis. New techniques in unfolding or more sophisticated methods to match energy deposition and single particles can allow to reconsider measuring event shapes for charged and uncharged particles.

11

THE EVENT SHAPE OBSERVABLES

Event shape observables measure the geometric properties of energy flow in the final state. They are constructed to describe the events from a configuration, where all particles fly along a single axis continuously to a configuration, where the energy is homogeneously distributed in ϕ. Three event shape variables were chosen to be studied in this analysis. The decision was made by selecting the event shape variables most sensitive to the underling event according to studies performed by Banfi et al. [66].

Event shapes in hadron collisions have to be Lorentz invariant under boosts along the beam axis, as the center-of-mass frame of the hard scattering process is usually boosted along this axis. In order to achieve this, only transverse momenta are used in the definitions. As the transverse components of the momentum are invariant under these boosts, the event shape variables itself are invariant by construction. Event shape variables also share the properties of being infrared and collinear safe (Eq. (**??**)) and calculable up to NLL. The last statement is only true for *directly global* and *exponentially suppressed* event shapes [66, 78], where global means that all particles in an event are used. This is not possible in experiments and therefore, the *exponentially suppressed* event shapes were introduced, where particles in the forward direction are suppressed. In this analysis, a third version was used: *Central* event shapes. These event shapes only use a subset of the particles in an event from a central region. Analyses based on tracks are restricted to this class of event shape variables as the inner detector only covers pseudo-rapidities up to $|\eta| < 2.5$. In addition, using only charged particles destroys the IRC-safety anyway, as collinear splitting of a charged particle into a charged and a neutral particle is not collinear safe. These central event shapes do not lose the sensitivity to non-perturbative effects at low momentum transfer and therefore are still experimentally useful observables.

All event shape variables used here were constructed to reach values of zero for an event, where all particles fly along one axis in the $\phi - r$-plane, later called *dijet-like* or *pencil-like*, to their maximum for the perfectly homogeneous event, later referred to as *spherical configuration*.

To minimize uncertainties, all event shape variables used are ratios of measured quantities, i. e. are normalized to the scalar sum of charged particle transverse momenta $\sum p_T$.

11.1 TRANSVERSE THRUST

The thrust is probably the most widely known and used event shape variable:

$$T = \max_{\hat{n}} \frac{\sum_i |\vec{p}_i \cdot \hat{n}|}{\sum_i |\vec{p}_i\,|}, \tag{11.1}$$

where the sum is performed over the momenta of all particles within the kinematic selection.

The transverse thrust is an adaption to provide the Lorentz-invariance needed. It is defined as:

$$T_\perp = \max_{\hat{n}} \frac{\sum_i |\vec{p}_{T,i} \cdot \hat{n}|}{\sum_i |\vec{p}_{T,i}\,|}, \tag{11.2}$$

The thrust axis $\hat{n}_T$ is the unit vector that maximizes the ratio in equation **??**. The range of $T_\perp$ values reaches from 1, for a perfectly balanced back-to-back configuration, to $T_\perp = \langle|\cos\psi|\rangle = 2/\pi$ for a circularly symmetric distribution of particles in the transverse plane, where ψ is the azimuthal angle between the thrust axis and each respective particle.

As said before, all event shape variables used in this analysis have values of zero for the pencil-like topology and reach their maximum for the isotropic distribution in the transverse plane. To achieve this, the new quantity $\tau_\perp = 1 - T_\perp$ is defined, which shares this property. From here on, transverse thrust will refer to this quantity $\tau_\perp$. Examples for a dijet-like and a spherical event together with the thrust axis are shown in Fig. 11.1.

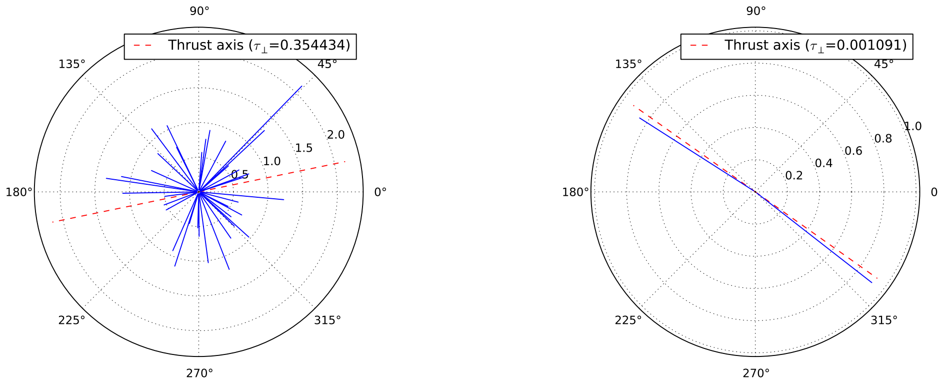


Figure 11.1: Example configurations for a spherical (left) and a pencil-like (right) event configuration.

11.2 THRUST MINOR

The thrust minor defines the out-of-event-plane energy flow, where the event plane is the plane spanned by the thrust axis and the beam axis $\hat{z}$. The hard scattering outgoing particles are in this event plane, and the energy flow perpendicular to that plane is dominated by non-perturbative effects.

$$T_M = \frac{\sum_i |\vec{p}_{T,i} \cdot \hat{n}_m|}{\sum_i |\vec{p}_{T,i}|}\,, \quad \hat{n}_m = \hat{n}_T \times \hat{z}\,. \tag{11.3}$$

The sum again is performed over the transverse momenta of all particles within the kinematic selection. The thrust minor is zero for all events without any momentum flow outside the event plane and $2/\pi$ for the spherical configuration.

11.3 SPHERICITY

The third event shape observable used in this analysis is the transverse sphericity $S_\perp$, which is derived from the widely used Sphericity S [79].

The sphericity describes the energy flow based on the momentum tensor,

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2}\,, \tag{11.4}$$

where the Greek indices represent the x, y, and z components of the momentum of the particle i. The sphericity of the event is defined in terms of the two smallest eigenvalues of this tensor, λ_2 and λ_3:

$$S = \frac{3}{2}(\lambda_2 + \lambda_3). \tag{11.5}$$

The sphericity could be interpreted as a measure of $\sum p_T^2$ with respect to the event axis [79, 80], where the event axis is defined as the line passing through the interaction point and oriented along the eigenvector associated with the largest eigenvalue, λ_1.

The transverse sphericity was constructed by restricting the definition of the sphericity to the transverse components:

$$S^{xy} = \sum_i \frac{1}{|\vec{p}_{T,i}|^2} \begin{bmatrix} p_{x,i}^2 & p_{x,i}\, p_{y,i} \\ p_{x,i}\, p_{y,i} & p_{y,i}^2 \end{bmatrix} \tag{11.6}$$

and

$$S_\perp = \frac{2\lambda_2^{xy}}{\lambda_1^{xy} + \lambda_2^{xy}}\,, \tag{11.7}$$

where $\lambda_2^{xy} < \lambda_1^{xy}$ are the two eigenvalues of S^{xy}. It has the usual range from zero for dijet-like events to 1 for isotropic events.

The event shape obervables were calculated using CAESAR [81].

12

ANALYSIS

The following chapter describes the analysis of the event shape variables. Most control distributions are only shown for the most inclusive event shape distributions. The conclusions drawn here were also verified for all distributions entering the results. If other distributions behave differently than the shown ones, this is explicitly mentioned. The single steps described in the following sections were performed with all phase space configurations shown in Sec. 13.1.

12.1 EVENT AND TRACK SELECTION

The data used in this analysis was taken during the first two periods of $\sqrt{s} = 7$ TeV LHC data taking in March and April 2010, corresponding to an integrated luminosity of $\sim 168\ \mu\text{b}^{-1}$. During that time, the LHC ran with a low instantaneous luminosity ($\mathcal{L}_{\text{inst}} < 2 \times 10^{30}\ \text{cm}^{-2}\text{s}^{-1}$), which results in a low number of simultaneous pp interactions. The average number of simultaneous interactions per bunch crossing $\langle\mu\rangle$ was always below 0.1 (Fig. 12.1). This condition is needed to perform this analysis, as it is very sensitive to pile-up. To even further reduce pile-up, all events containing a vertex with more than three tracks, excluding the primary interaction vertex, were rejected. Vertices are ordered by $\sum p_T^2$ of the tracks assigned to that vertex. The vertex with the highest $\sum p_T^2$ is named the primary interaction vertex of the event. In addition to the vertex cut, the event must have been triggered by the MBTS trigger, all inner detector subsystems needed to be in nominal operating condition and the event must have more than six

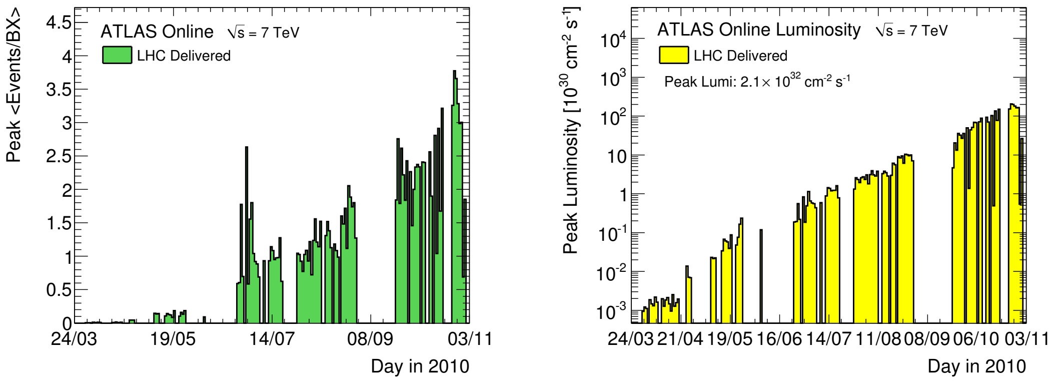


Figure 12.1: The average number of events per beam crossing (BX) at the start of an LHC fill (left) and the maximum instantaneous luminosity (right) versus day [82]. For this analysis, only data taken in April 2010 was used.

tracks which pass the track selection below. The last cut suppresses single- and double-diffractive events (see Sec. 12.3).

Event selection

- MBTS trigger
- $\leqslant 1$ primary vertex with $n_{\text{Tracks}}^{\text{vtx}} > 3$
- Inner detector subsystems in good condition
- At least 6 reconstructed tracks ($N_{ch} \geqslant 6$)

The track selection is chosen to secure a good balance between the most inclusive selection and track reconstruction efficiency/purity [68]. All tracks used in this analysis have:

Track selection

1. $p_T > 0.5$ GeV.
2. at least one hit in the pixel detector and $\geqslant 6$ hits in the SCT.
3. a hit in the innermost pixel layer, if the corresponding module was active.
4. a track-fit probability $\chi^2 > 0.01$ for tracks with $p_T > 10$ GeV in order to remove mis-measured tracks.
5. transverse and longitudinal impact parameters with respect to the primary vertex, $|d_0| < 1.5$ mm and $|z_0| \sin\theta < 1.5$ mm.
6. $|\eta| < 2.5$.

Tracks with a transverse momentum below 0.5 GeV suffer from a decreased track reconstruction efficiency due to interactions with inactive material. The track selection criteria 2.- 4. remove tracks with bad fits or insufficient fit input. The limitation of the impact parameters in 5. removes non-beam background like particles from cosmic showers crossing the detector simultaneous to the collision. The last cut in pseudorapidity is given by the detector geometry. The inner detector ends at $|\eta| < 2.5$ and no tracks can be reconstructed outside this region. To reject any mis-measured or badly fitted tracks, this cut was applied in the analysis.
After all cuts, the analysis is based on 17 million events containing 300 million tracks in total.

12.2 MONTE CARLO TUNES

Seven tunes from three different Monte Carlo generators are used in the analysis. The main differences (besides the differences of the generators discussed in Sec. 5.1.1) are mainly the data used for tuning. The impact of using different PDF sets is estimated to be small compared to the influence of which data was used for tuning. An overview is given in Table 12.1.

For the study of most systematic uncertainties and to compute correction factors, the Monte Carlo generated needs to be passed through the full detector simulation provided by GEANT4. This was only the case for PYTHIA 6 AMBT1, PYTHIA 6 DW and Herwig++ 2.5.0. Therefore, the reference tune throughout the analysis was PYTHIA 6 AMBT1, as it provided the highest number of generated events on particle and detector level, and Herwig++ 2.5.0 was used to study the model dependent unfolding uncertainty (see Sec. 12.5.2). The other tunes (PYTHIA 6 AMBT2B, PYTHIA 6 DW, PYTHIA 8 A2 and Herwig++ UE7-2) have only been used as a comparison to unfolded data. PYTHIA 6 AMBT2B is the successor of PYTHIA 6 AMBT1. Additional LHC measurements were used to further improve the PYTHIA 6 AMBT1 tune.

Generator	Version	Tune	PDF	Focus	Data	FS
PYTHIA 6	6.425	AMBT1 [83]	MRST LO** [84]	MB	Early LHC	Yes
PYTHIA 6	6.425	AMBT2B [85]	CTEQ6L1 [86]	MB	LHC	No
PYTHIA 6	6.421	DW [87]	CTEQ5L [88]	UE	Tevatron	Yes
PYTHIA 6	6.425	Z1 [89]	CTEQ5L	UE	LHC	No
PYTHIA 8	8.157	A2 [90]	MSTW2008LO [91]	MB	LHC	No
Herwig++	2.5.1	UE7-2 [92]	MRST LO**	UE	LHC	No
Herwig++	2.5.0	Default	MRST LO**	UE	LHC	Yes

Table 12.1: Monte Carlo models used in the analysis. Tunes use data from different experiments for constraining different processes, but for brevity only the *data* which had the most weight on each specific tune are shown. Here *LHC* indicates data taken at $\sqrt{s} = 7$ TeV, although $\sqrt{s} = 900$ GeV data were also included in ATLAS tunes, with much smaller weight. Some tunes are focused on describing the minimum-bias (MB) distributions better, while the rest are tuned to describe the underlying event (UE) distributions, as indicated. Not all Monte Carlo samples were passed through the full detector simulation (FS) and information was available only on particle level.

12.3 BACKGROUND CONTRIBUTIONS

Multiple backgrounds can contribute to the measurement. Single- and double-diffractive contributions are suppressed by removing events with less than six tracks which pass the track selection. Fig. 12.2 shows that the contribution from single- and double-diffractive events is negligible after applying the cut.

A second possible background is pile-up. The number of average interactions per bunch crossing $\langle\mu\rangle$ for the data used in this analysis is always below 0.1 (Fig. 12.1). Remaining pile-up events are removed by the vertex cut described in Sec. 12.1. Comparing Monte Carlo with pile up contributions and applied cuts to Monte Carlo without pile-up contributions shows that no measurable influence of remaining pile-up events can be observed after applying the cuts.

The amount of additional beam (e. g. beam–gas and beam–material interactions) and non-beam (e. g. cosmic rays and detector noise) backgrounds were estimated using hits in the pixel detector not associated to any reconstructed tracks [68]. The amount of background after all selections was found to be negligible and no dedicated background subtraction is performed. The full analysis has been performed with $N_{ch} \geqslant 2$

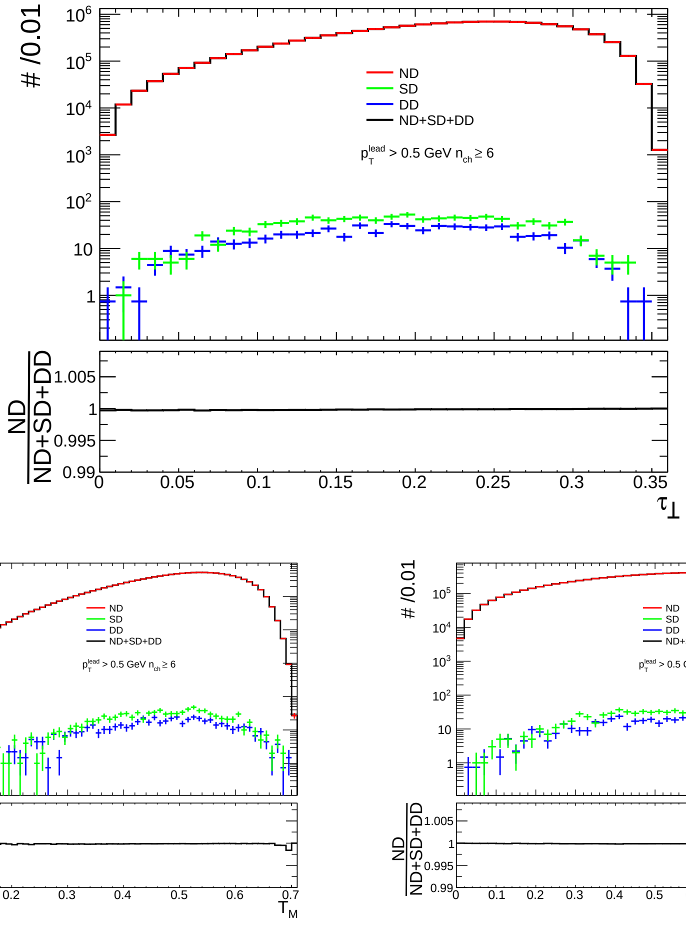


Figure 12.2: Inclusive distribution of transverse thrust (top), thrust minor (left) and transverse sphericity (right) from PYTHIA 6 AMBT1 sample, stacking non-, single- and double-diffractive events.

instead of 6. As the results barely add any additional information, but have a larger uncertainty due to the single- and double-diffractive background, the results will not be discussed here.

12.4 CORRECTION TO PARTICLE LEVEL

In order to make the results comparable to theory predictions and results from other experiments and to provide a suitable input for MC tuning, the event shape distributions need to be corrected to particle level. To do so, a two-step correction was chosen. In a first step, corrections for trigger and selection efficiencies are applied. The corrected distributions are then used to perform a bin-by-bin unfolding to correct for the remaining detector effects, e. g. detector resolution.

12.4.1 *Event-level corrections*

The vertex reconstruction and trigger efficiency was measured in [68] as a function of number of tracks passing all selection criteria except the primary vertex cut, $N_{\text{sel}}^{\text{BS}}$.

The MBTS trigger efficiency, $\epsilon_{\text{trig}}\left(N_{\text{sel}}^{\text{BS}}\right)$, was determined using an orthogonal trigger, which only required to coincide with the beam collision and 4 pixel as well as 4 SCT hits on the second trigger level L2. The efficiency was found to be around 97% for low multiplicities and close to 100% otherwise. No dependencies on p_T and η of the selected tracks were observed.

The vertex reconstruction efficiency $\epsilon_{\text{vtx}}\left(N_{\text{sel}}^{\text{BS}}, \langle\eta\rangle\right)$ was determined by taking the ratio of the number of events triggered and the number of events with a reconstructed vertex. For an event selection criterion of at least six tracks with a transverse momentum above 500 GeV and $|\eta| < 2.5$, this efficiency is close to unity.

The total efficiency correction applied to data is:

$$w_{\text{ev}} = \frac{1}{\epsilon_{\text{trig}}\left(N_{\text{sel}}^{\text{BS}}\right)} \cdot \frac{1}{\epsilon_{\text{vtx}}\left(N_{\text{sel}}^{\text{BS}}, \langle\eta\rangle\right)} \tag{12.1}$$

accounting for the two efficiencies described above. For the event selection used in this analysis, the correction $(1 - w_{\text{ev}})$ is below 1% [68].

12.4.2 *Bin-by-bin corrections*

To correct for additional, mostly unknown detector effects, a bin-by-bin unfolding was performed on top of the event-level correction described above.

For bin-by-bin unfolding, a correction factor for each bin in each distribution is calculated from Monte Carlo simulation:

$$U_{\text{bin}} = \frac{V_{\text{bin}}^{\text{Gen}}}{V_{\text{bin}}^{\text{Reco,eff corr}}} \tag{12.2}$$

$V_{\text{bin}}^{\text{Gen}}$ is the number of entries on Monte Carlo generator level in this bin, $V_{\text{bin}}^{\text{Reco,eff corr}}$ the number of entries in this bin after reconstruction with event-level corrections applied. Each bin in the data distribution is multiplied by the corresponding correction factor U_{bin} to unfold the data.

Other unfolding techniques such as iterative Bayesian unfolding and a data-driven unfolding method were tested. The bin-by-bin correction established itself as the most robust and effective correction method and therefore is used throughout this analysis and is applied to all measured event shape distributions.

Some prerequisites need to be fulfilled for the bin-by-bin unfolding to work properly:

- A clearly visible correlation between generator and reconstruction level.
- Bin sizes need to be chosen according to the resolution. Large bin migrations decrease the performance of the unfolding.
- The Monte Carlo should describe the data on reconstruction level.

The three prerequisites are tested in the following sections:

Correlation of generator and reconstruction level

The correlation between generator and detector level are shown in Fig. 12.3. While bin migration is clearly visible (non-diagonal entries), a clean linear correlation between generator and reconstruction level is visible. A slight asymmetry to lower values on reconstruction level can be observed. This is caused by a combination of N_{ch}–dependence of event shape variables and the fact that more tracks fail to be reconstructed than fake tracks are produced. This causes a reduction in N_{ch} which then causes a decrease in the values of event shape variables. Looking into the projections in Appendix A shows that the asymmetry is very small.

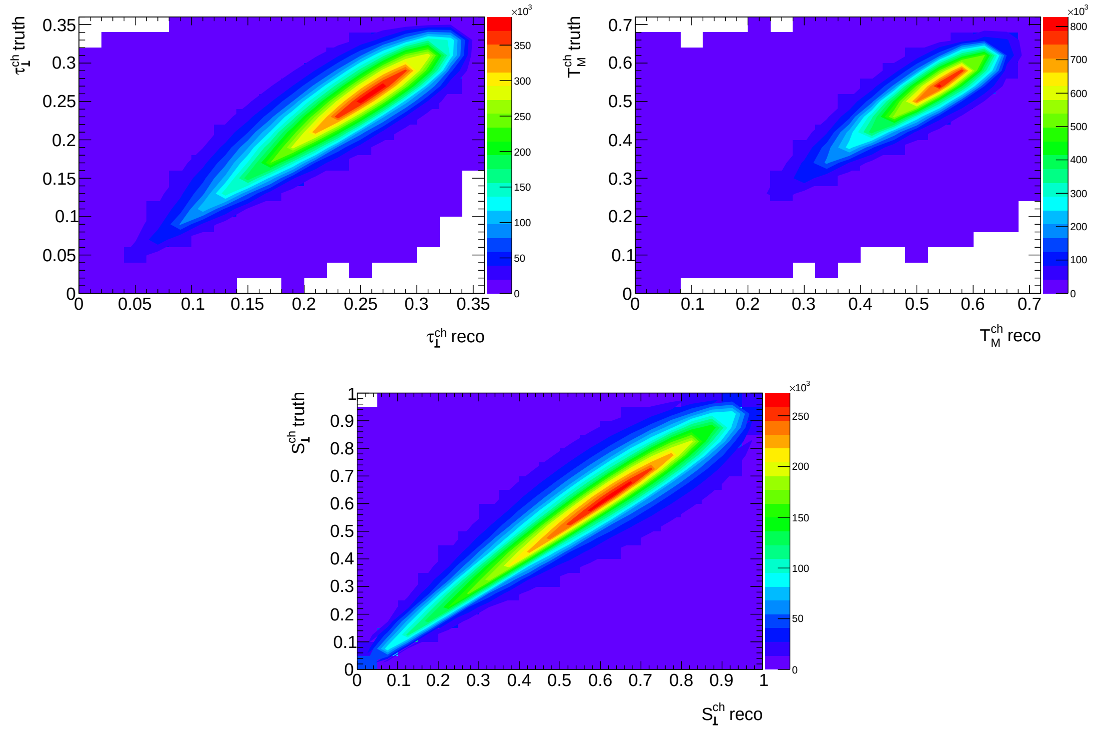


Figure 12.3: Correlation between generator and reconstruction level for transverse thrust (top left), thrust minor (top right) and sphericity (bottom).

Resolution and Purity

The resolution for every bin is calculated and fitted with a Gaussian (Appendix A). The fits are not optimized as the results are only used as an estimator. The RMS for every bin is shown in Fig. 12.4. In the worst case, the resolution is about 150% of the bin size. Given that the fits mostly underestimate the resolution, the bin sizes are big enough to prevent large bin migrations. This is shown in the right plots of Fig. 12.4. The purity $\mathcal{P}$ is defined as

$$\mathcal{P} = \frac{N_{\mathrm{Ev}}^{\mathrm{same\,bin}}}{N_{\mathrm{Ev}}} , \qquad (12.3)$$

where $N_{\mathrm{Ev}}^{\mathrm{same\,bin}}$ is the number of events which are in the same bin for generator and reconstruction level, and N_{Ev} the number of events in the bin on generator level. For transverse thrust, $\mathcal{P}$ is at 70% in the first bins and drops down to 30%. Only the last bin is significantly below 15%. A similar behaviour is observed for thrust minor. Starting at close to 90% in the first bin, the purity decreases to 40% before dropping down to 20% in the last bins. The purity in the sphericity distributions starts with above 80% in the first bin, but thereafter drops to below 30%. Usually, bin widths bigger than the resolution are desired. The low purity and resolutions broader than the bin width might introduce biases during the unfolding due to high bin migrations. To test if problems occur, a closure test is performed. The Monte Carlo sample is split into two statistically independent halves (using even and odd event numbers). One half is used to calculate the correction factors with which the other half is corrected. Comparing the corrected spectrum with the generator level, a good agreement proves a good performance of the unfolding. Fig. 12.5 shows, that for all three event shapes observables, the deviation between corrected reconstruction and generator level is less than 2%. This shows, that the unfolding method is able to handle the bin migrations.

Data - MC agreement on detector level

The unfolding is based on the assumption that data and Monte Carlo agree on detector level and the detector is well described/modeled in Monte Carlo. Therefore, data and two Monte Carlo tunes (PYTHIA6 AMBT1 and HERWIG) are compared in Fig. 12.6. None of the Monte Carlo tunes used is able to describe the data, but the discrepancies are small. The correction factors using PYTHIA 6 AMBT1 and HERWIG are different and lead to different unfolding results. A systematic uncertainty is applied to cover the difference (Sec. 12.5).

12.4.3 *Alternative correction methods*

The bin-by-bin unfolding strongly depends on the Monte Carlo simulation used to compute the correction factors. Therefore, two additional unfolding methods were tested and compared to the bin-by-bin unfolding.

Iterative Bayesian unfolding

The first cross check is performed with an unfolding technique proposed by D'Agostini [94] using Bayes' theorem. The goal is to maximize the probability

$$P(x_C|x_E, \Lambda, I)\,, \tag{12.4}$$

that a ("true") distribution (x_C) causes the measured spectrum (x_E), given a smearing matrix Λ (calculated from Fig. 12.3) and the state of information I under which the analysis was performed (advance information). Bayes' theorem allows to rewrite Eq. (**??**) to

$$P(x_C|x_E, \Lambda, I) \propto P(x_E|x_C, \Lambda, I) \cdot P(x_C|I), \tag{12.5}$$

where $P(x_E|x_C, \Lambda, I)$ is the likelihood that x_C leads to x_E. $P(x_C|I)$ is the *prior*, the probability of x_C given an advance information I. Translating this from distributions to bins (to unfold binned results as in this analysis) leads to

$$P(C_i|E_j, I) \propto P(E_j|C_i, I) \cdot P(C_i|I), \tag{12.6}$$

in which C_i and E_j are the bins of the "true" and measured distributions and $P(E_j|C_i, I)$, for instance, is the probability that an event from C_i gets measured in E_j and can be directly taken from Fig. 12.3. $P(C_i|I)$ is the prior again.
Starting from the true distribution of the Monte Carlo prediction as a prior and iteratively using the result as new prior, [94] states that the results from different Monte Carlo generators will converge and no model dependency is left.
The results, side by side with bin-by-bin results, are shown in Fig. 12.7. The central values of the unfolded data are similar throughout all bins and distributions. The systematic uncertainties, however, are even a bit larger. Therefore, the iterative bayesian unfolding endorses the bin-by-bin correction, but will not be used as the default unfolding method for this analysis.

Data-driven unfolding method

A data-driven approach to correct for detector effects in analyses which are dominated by track reconstruction and resolution effects was used in [95]. The correction does not use Monte Carlo but starts with the data spectrum on detector level. Starting from there, for every track a random number between 0 and 1 is generated. If this number is above the track reconstruction efficiency (as function of p_T and η), the track is removed. This results in a new distribution containing fewer tracks, and the procedure is repeated 4 times. For every bin in the distribution to unfold, the change of value is plotted and interpolated by a fit. Counting the original distribution (data on detector level) as iteration 1, "iteration 0" is the corrected particle level distribution. To achieve that, each bin value is interpolated by the fit.
This method would have removed the Monte Carlo uncertainty of the unfolding, but was not applicable to this analysis. The main challenge is to find the correct function to interpolate to "iteration 0". The change in shape and skewness with each iteration leads to an unstable and imprecise result.

12.5 SYSTEMATIC UNCERTAINTIES

Various systematic uncertainties were studied. Only uncertainties which are not negligible are presented in the following: Track reconstruction, unfolding model dependency and unfolding statistical uncertainty. All other systematic uncertainties combined are below 0.1%.

12.5.1 *Track and vertex reconstruction*

Effects of the track reconstruction efficiency are corrected for, based on the the results provided in [68]. The efficiencies are shown in Fig. 12.8. The reconstruction efficiency is above 80% for tracks with $p_T > 500$ MeV. The efficiency as a function of η is flat around 80% in the very central region $|\eta| < 1.0$ and falls below 50% for the most forward tracks. Using only tracks above $p_T > 500$ MeV, leads to an higher efficiency than shown in Fig. 12.8. To estimate the systematic uncertainty on this correction, tracks are removed from Monte Carlo truth level, according to the provided efficiency $\epsilon(p_T, \eta)$ and $\epsilon(p_T, \eta) \pm 1\sigma$. To do so, for every charged particle a random number between 0 and 1 is generated and compared to the track reconstruction efficiency $\epsilon(p_T, \eta)$. If the

Trigger and vertex efficiency	< 0.1
Track reconstruction	$0.1 - 0.5$
Correction model difference	$1 - 10$
PYTHIA correction stat. uncertainty	$0.1 - 2$
Total systematic uncertainty	1-5

Table 12.2: Summary of systematic uncertainties in %.

random number is above $\epsilon(p_T, \eta)$, the charged particle is removed. This is repeated with an $\epsilon(p_T, \eta)$ increased and decreased by 1σ to estimate the uncertainty. The difference between the up and the down shift is taken as a systematic uncertainty. The uncertainties for the inclusive distributions are shown in Fig. 12.9. Applying the track reconstruction efficiency to truth level shifts the distribution towards the detector level. The lines are almost fully covered by the detector level distribution. This shows that the detector effects are dominated by the track reconstruction efficiency. Shifting the applied efficiency up and down does not change the distributions much. The ratio in the bottom part of the plots shows that the resulting systematic uncertainty is below 1%. The same applies for the different p_T^{lead} slices and the 2-dimensional distributions over $\sum p_T$ and even charged multiplicity N_{ch}. The uncertainty due to the vertex reconstruction efficiency is found to be negligible.

12.5.2 *Unfolding: Model dependency*

The result of the bin-by-bin unfolding is highly dependent on the Monte Carlo model used to calculate the unfolding factor. A series of Monte Carlo generator and tunes were used to unfold the data. From all Monte Carlo predictions that are able to reproduce the data distribution on detector level, the two most different results are used to evaluate the uncertainty (PYTHIA 6 AMBT1 and HERWIG). The ratio of the unfolded data with these two Monte Carlo predictions is used as the systematic uncertainty of the model dependency. The unfolding factors for PYTHIA 6 AMBT1 and HERWIG are shown in Fig. 12.6. Fig. 12.10 shows the resulting systematic uncertainty for the three most inclusive p_T^{lead} bins. The resulting systematic uncertainty is $1 - 10\%$ with few exceptions at the edges of the distributions with 20% and above.

12.5.3 *Unfolding: Statistical uncertainty*

In addition to the model dependency, a statistical uncertainty of the unfolding due to the finite size of the Monte Carlo samples needs to be evaluated. This is done using a bootstrapping method. Every entry in the correlation matrix (Fig. 12.3) is replaced by a random number drawn from a Poisson distribution with a mean of the original bin content and the result is used to unfold the data. This is repeated 10000 times and the spread of all results is used as the statistical uncertainty on the unfolding procedure. The uncertainties are shown in Fig. 12.10. The uncertainty is below 1% except for some of the outermost bins. All uncertainties discussed are considered as uncorrelated and added in quadrature. A summary of all uncertainties is shown in Table 12.2.

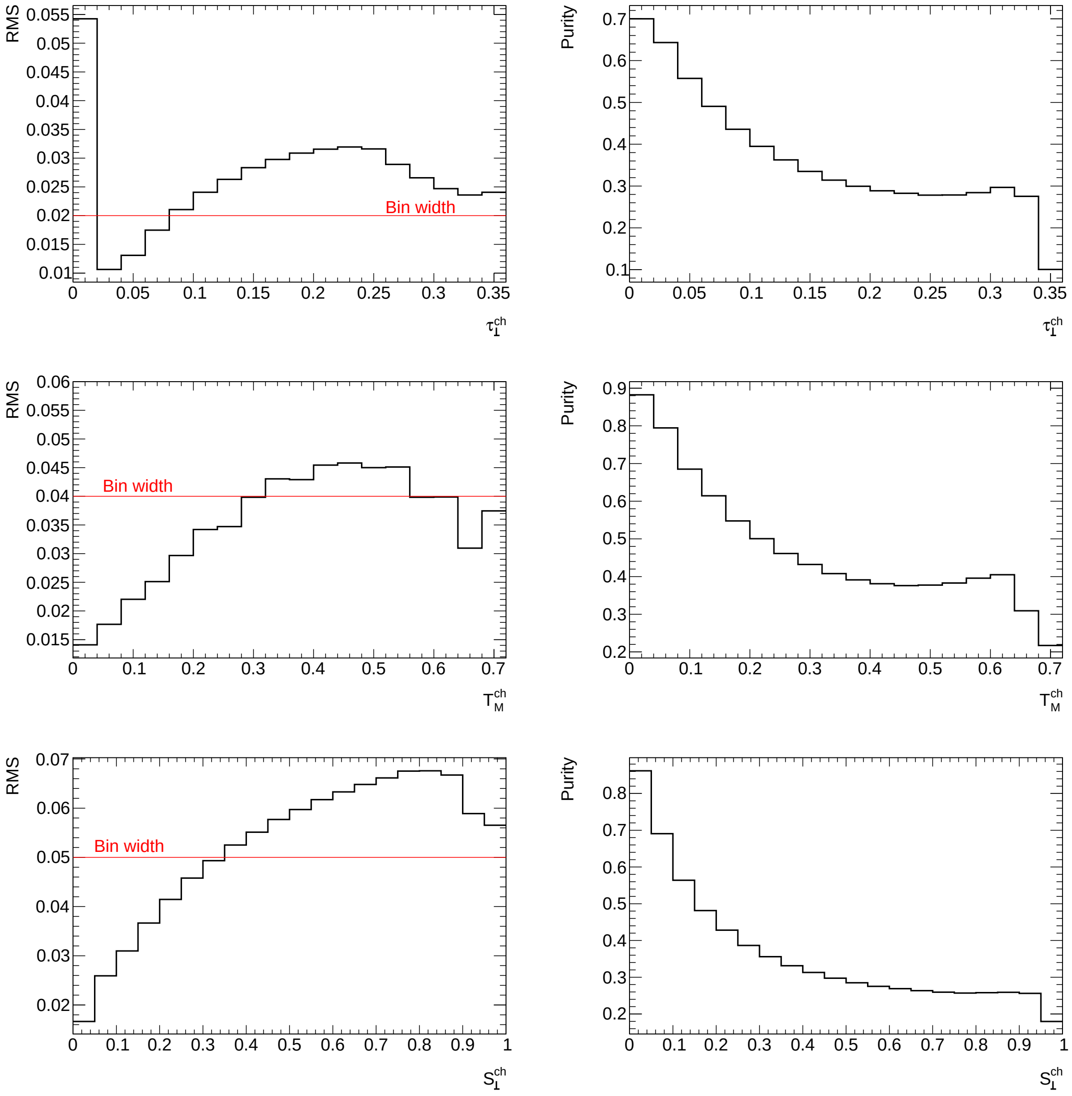


Figure 12.4: Left: Approximate resolution for each bin and each event shape observable. RMS taken from automated gauss-fits (Appendix A). Spikes are caused by non-converging fits. The red lines represents the bin width. Right: Purity for each bin and each event shape observable (Eq. (**??**)).

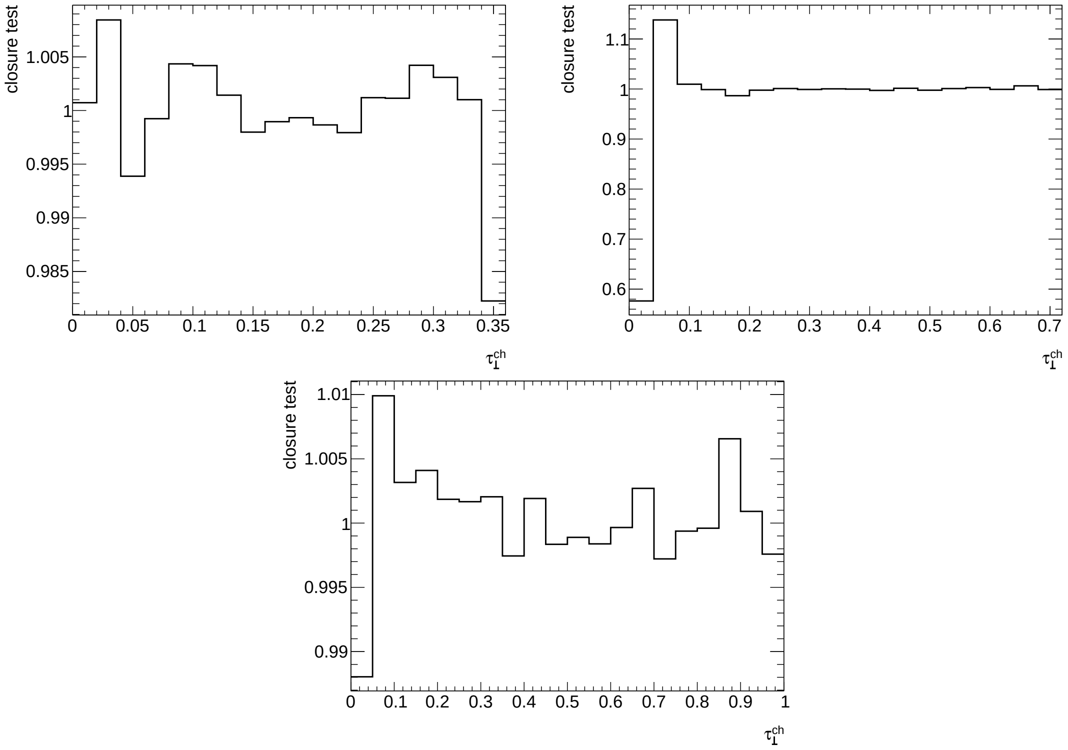


Figure 12.5: Closure test for the three most inclusive event shape distributions. The deviations to one are of statistical nature and added as a systematic uncertainty to the final result.

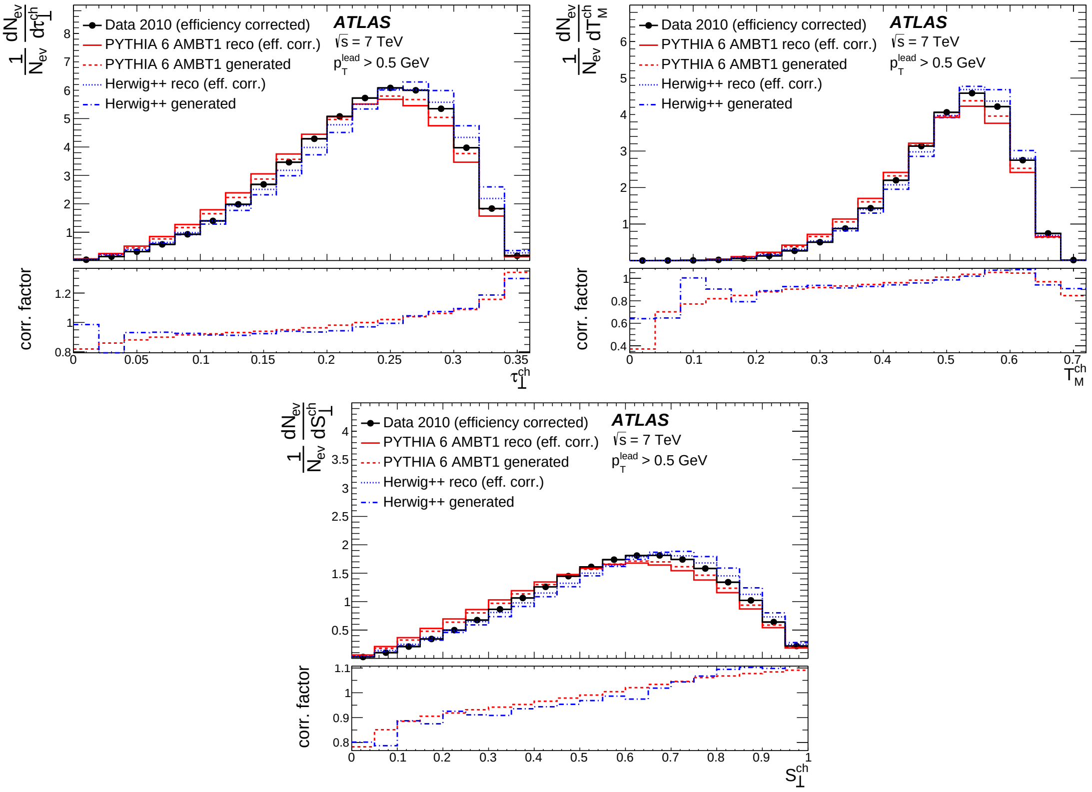


Figure 12.6: Comparison between efficiency corrected data, detector level Monte Carlo and truth level Monte Carlo for PYTHIA 6 AMBT1 and HERWIG for the most inclusive distributions. Top left shows the transverse thrust, top right thrust minor and bottom plot transverse sphericity. The bottom part of the plots show the correction factors for the two Monte Carlo tunes [93].

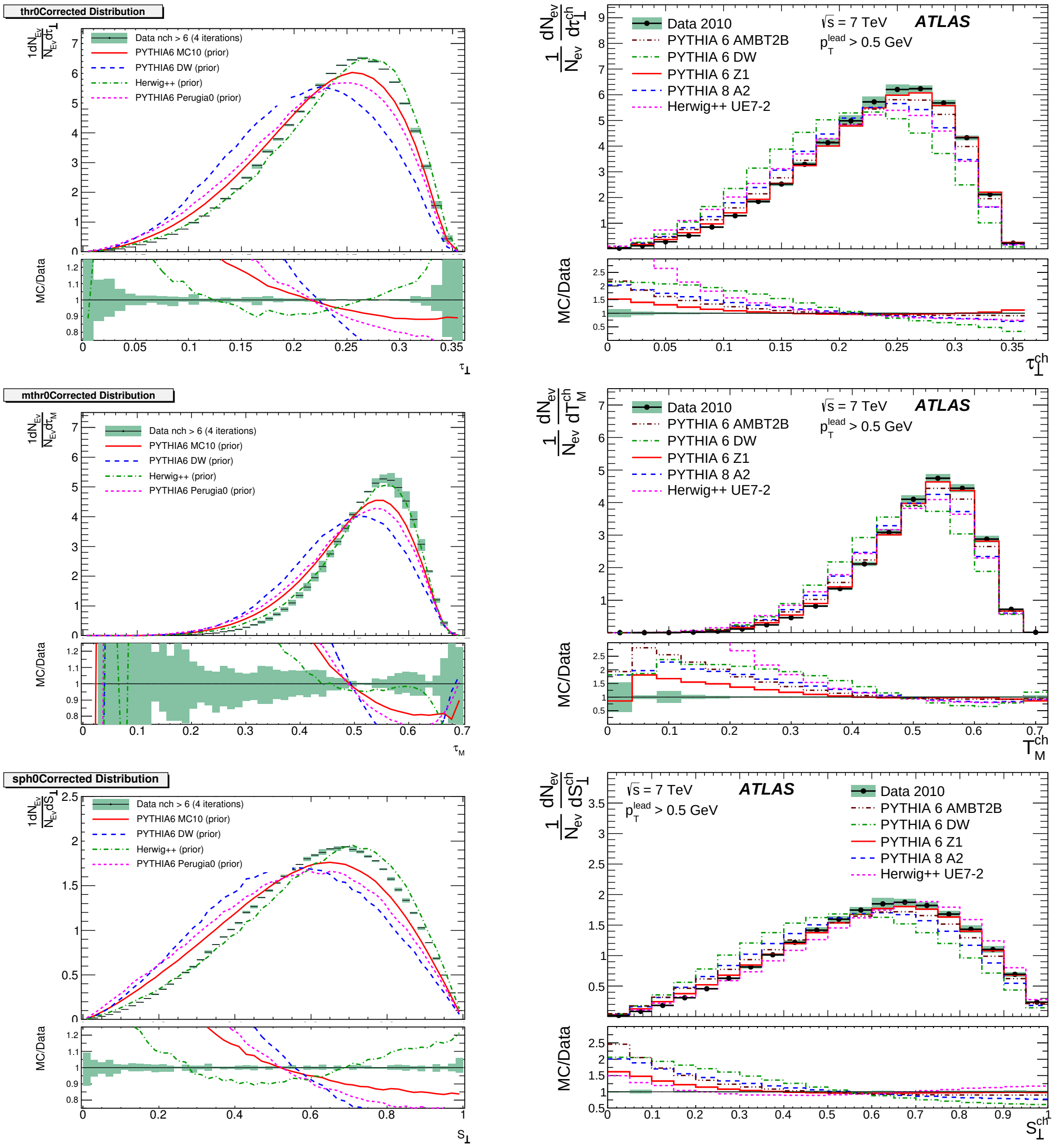


Figure 12.7: Results from bayesian (left) and bin-by-bin unfolding (right) compared to each other. Top plots show transverse thrust, middle ones thrust minor and bottom plots sphericity [93].

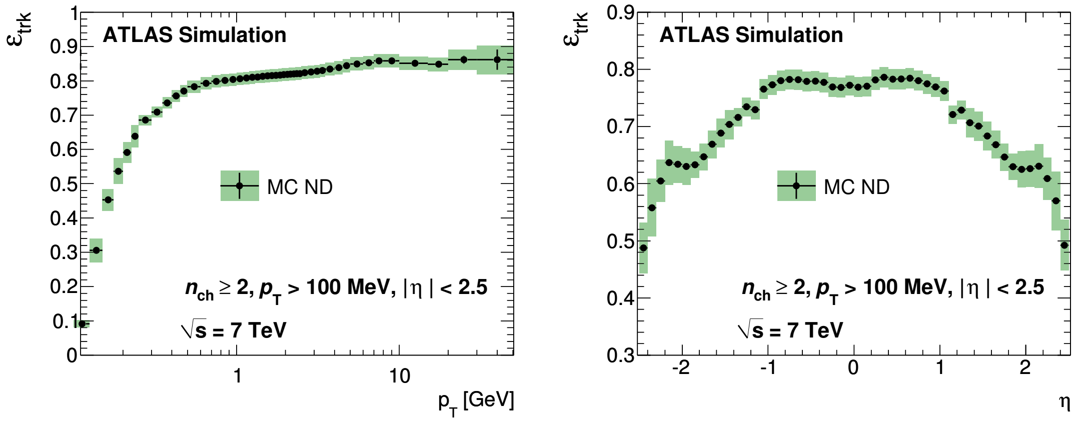


Figure 12.8: Track reconstruction efficiency as a function of p_T (left) and η (right) [68].

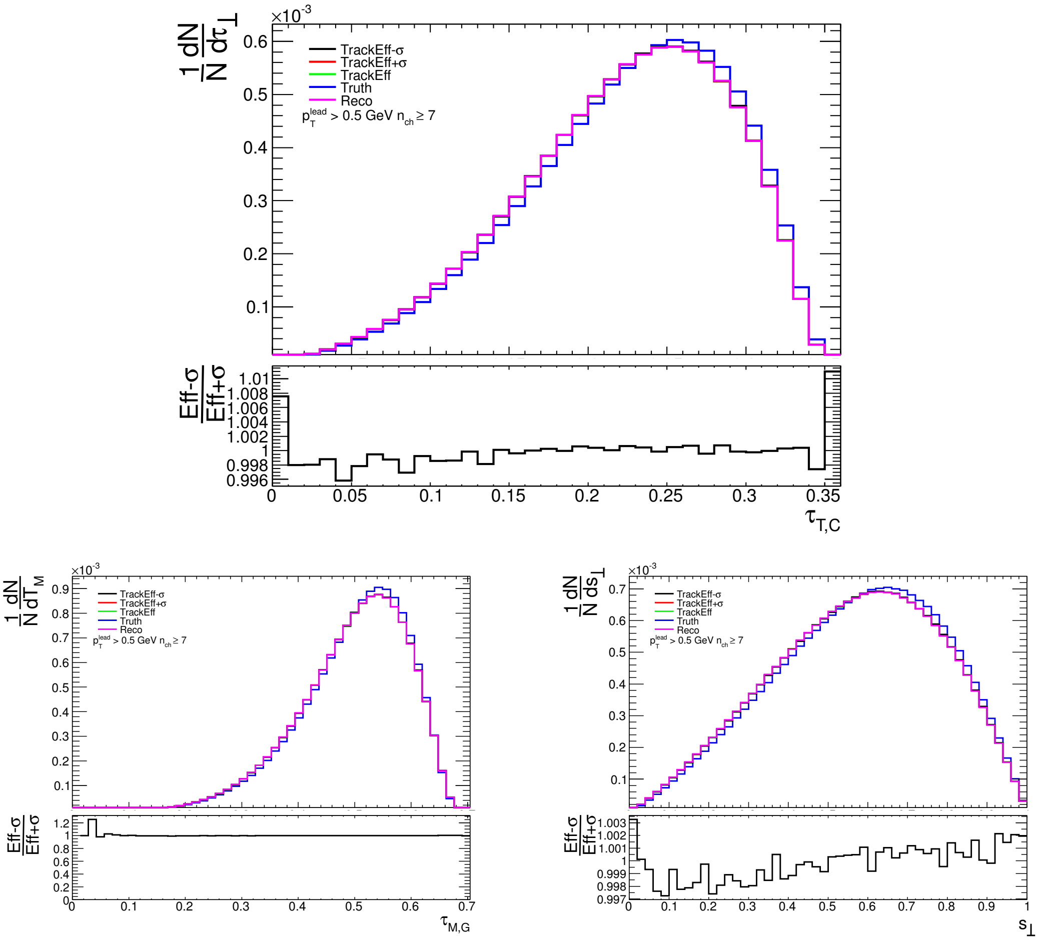


Figure 12.9: Calculation of the systematic uncertainty due to the track reconstruction efficiency. Top left plot shows the transverse thrust (donated as $\tau_{T,C}$), top right the thrust minor (named $\tau_{M,G}$) and sphericity (labeled as $s_\perp$) is shown at the bottom. The magenta line represents the detector-level, blue the truth-level distribution. The red line results from applying the track reconstruction efficiency to truth level, the black line applying the track reconstruction efficiency shifted by one sigma down, the green line shifted by one sigma up. These three lines with applied track reconstruction efficiencies are covered by the magenta line. The bottom part of each plot shows the ratio of efficiencies shifted down and up.

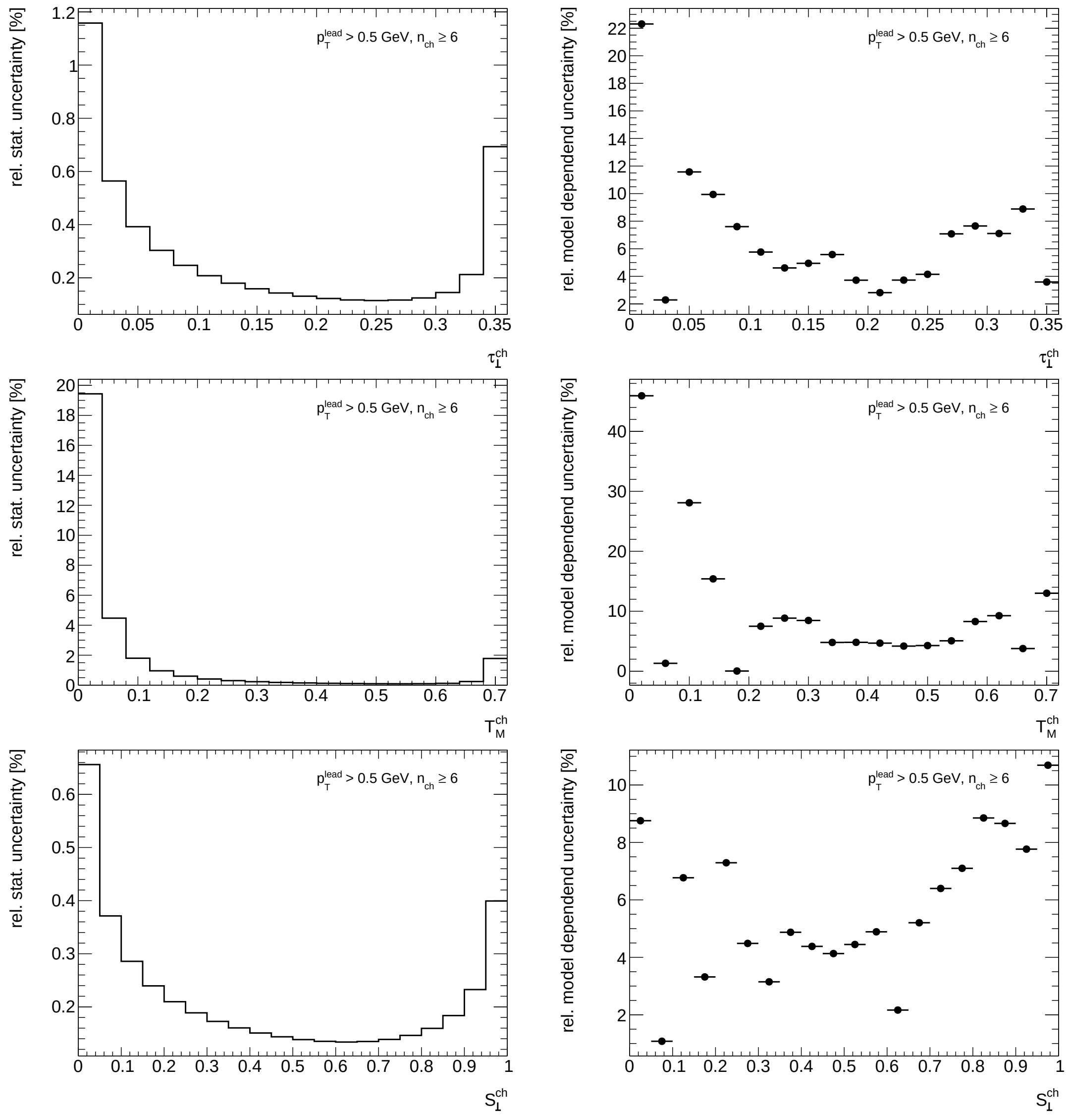


Figure 12.10: Statistical uncertainty on the bin-by-bin unfolding calculated using a bootstrapping method shown in the left row. Right plots show the relative systematic uncertainty due to the model dependency on the unfolding. All plots are for the most inclusive p_T^{lead} bin with $p_T^{lead} > 500$ MeV and $N_{ch} \geqslant 6$.

13

RESULTS

13.1 1D EVENT SHAPE DISTRIBUTIONS

The 1-dimensional event shape distributions are produced differentially in p_T^{lead} as well as in steps of minimum p_T^{lead}. While the differential plots are statistical independent, the inclusive distributions are preferred for Monte Carlo tuning. The physical interpretations are redundant and one set does not add much information to the other one. Only the differential plots are discussed in detail and the physical conclusions drawn from the discussion of the differential distributions are valid for the integrated plots as well. The integrated distributions can be found in Appendix B.
To estimate the number of events in each p_T^{lead} slice, the leading track transverse momentum as well as the inclusive track transverse momentum is plotted in Fig. 13.1. This shows for the integrated plots, that distributions with low p_T^{lead} are not completely dominated by the lowest p_T^{lead} contributions, but also have significant contributions from higher p_T^{lead} slices. An overview of the contributions to every p_T^{lead} bin is shown in Table 13.1.

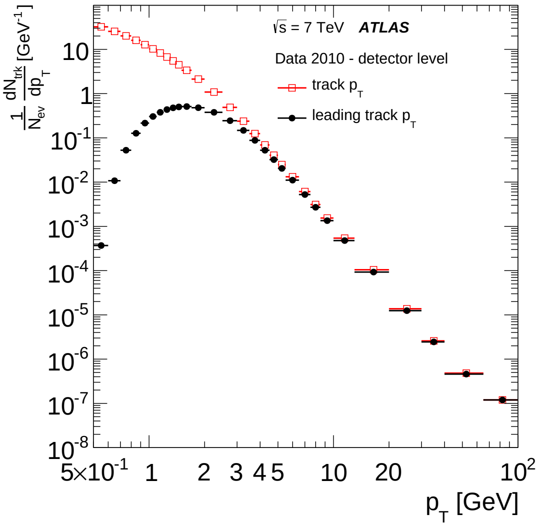


Figure 13.1: Track transverse momentum and leading track transverse momentum for the selected events and tracks. Errors are statistical only. Where not shown, the statistical error is smaller than the marker size[93].

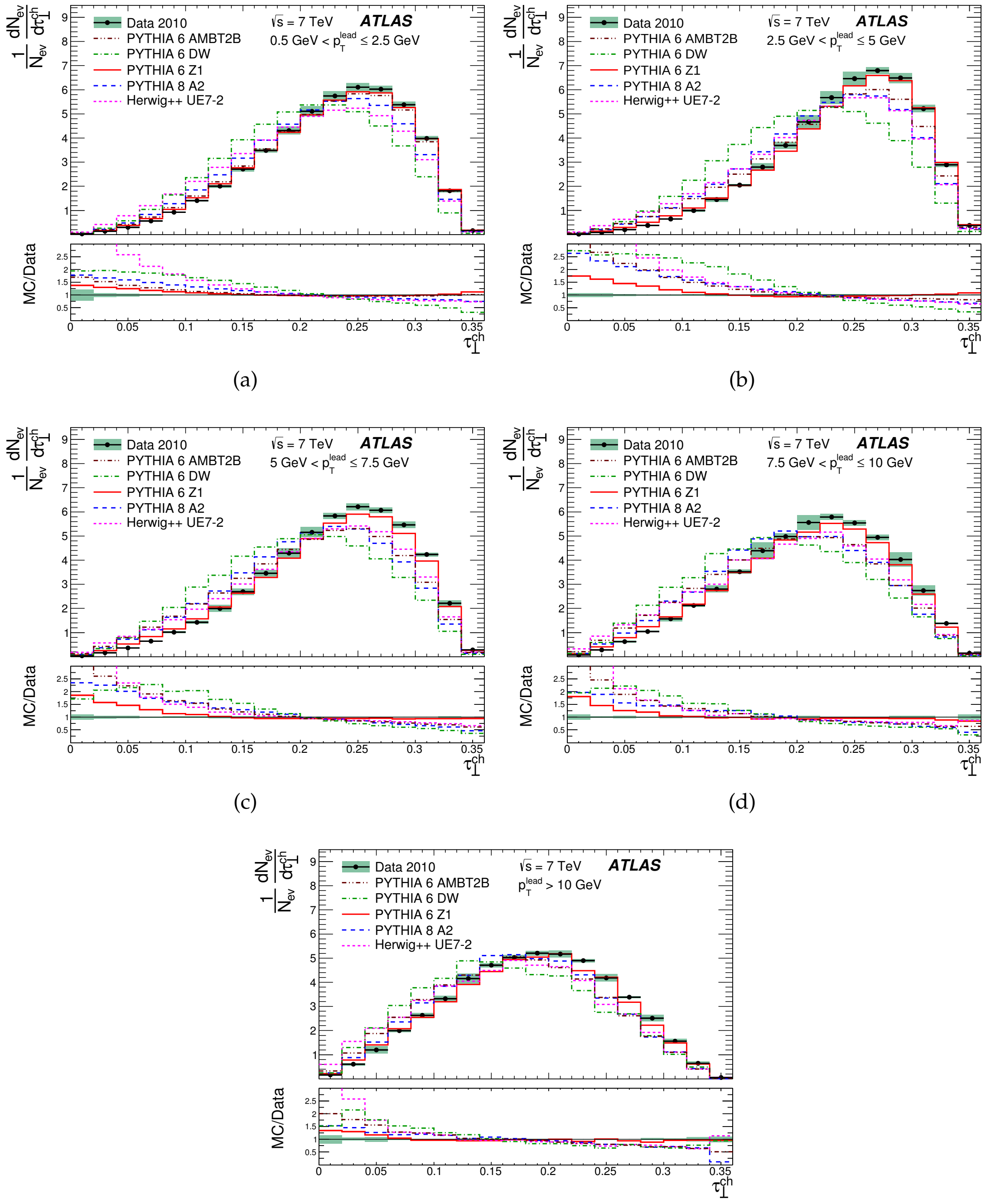
As expected, the higher the p_T^{lead} slice, the smaller the number of events contributing to it.
The differential 1-dimensional results are shown in Fig. 13.2, 13.3 and 13.4 for the three

Table 13.1: Percentage of events in each p_T^{lead} bin.

p_T^{lead} bin [GeV]	Percentage of events
0.5–2.5	68.45
2.5–5.0	28.20
5.0–7.5	2.65
7.5–10.0	0.47
> 10.0	0.23

event shapes. Fig. B.1, B.2 and B.3 show the event shapes integrated over the transverse momentum of the leading charged particle, p_T^{lead}, for different minimums of p_T^{lead}. The 2-dimensional distributions, mean values of event shape variables over charged particle multiplicity and sum of charged particle transverse momenta $\sum p_T$ respectively, are presented in Fig. 13.6. The distributions in the first three p_T^{lead} bins for the transverse thrust in Fig. 13.2a – 13.2c are fairly similar. The distributions peak at high values, indicating mostly spherical events, with a long tail to lower values, vanishing at the lowest bin. A shift towards lower values can be observed in the two highest p_T^{lead} bins (Fig. 13.2d – 13.2e). At the same time, the distributions broaden with increasing p_T^{lead}. The shift indicates that events with a higher p_T^{lead} are less spherical than those with a smaller p_T^{lead}. As the leading charged particle transverse momentum can be understood as a measure of the hardness of an event or as an estimator for the momentum transfer, this would indicate that harder events (events with a higher Q^2) tend to have a more dijet-like structure, i. e. a privileged direction. The broadening of the distribution indicates that the effect described before is not dominant and that a transition region is observed where events start to become more dijet-like . An overview over the mean value and RMS for the results, as measure of the transition to lower values and broadening, is shown in Table 13.2 and Fig. 13.5.

None of the Monte Carlo predictions shown are able to describe the data within the uncertainties, and all tunes are shifted to lower values of event shapes, i. e. predict more dijet-like events. The PYTHIA 6 Z1 tune provides the overall best description while the PYTHIA 6 DW tune is consistently furthest from data. The overall bad performance of PYTHIA 6 DW was already observed in [68], and DW is the only tune which is not tuned to LHC data. For the lower p_T^{lead} bins, PYTHIA 6 AMBT2B gives a better description than the HERWIG UE7-2 and PYTHIA 8 A2 tune, but vice versa for the intermediate and highest p_T^{lead} bins. The good performance of HERWIG UE7-2 at higher p_T^{lead} results from its tuning to underlying event data, like the PYTHIA 6 Z1 tune, which prefers events with a hard-scattering process, resulting in higher p_T^{lead} values. PYTHIA 6 AMBT2B and PYTHIA 8 A2 are both tuned to minimum bias measurements, which explains the accurate description at lower p_T^{lead}. The improved MPI modeling in PYTHIA 8 compared to PYTHIA 6 can explain the better performance of PYTHIA 8 A2 at higher p_T^{lead}. The thrust minor shows the same behaviour as the transverse thrust. The spectra for lower p_T^{lead} in Fig. 13.3a – 13.3c peak at high values and get shifted downwards with increasing p_T^{lead} (Fig. 13.3d – 13.3e). No Monte Carlo prediction is able to describe the data within systematic and statistical uncertainties. The observations made for the transverse thrust



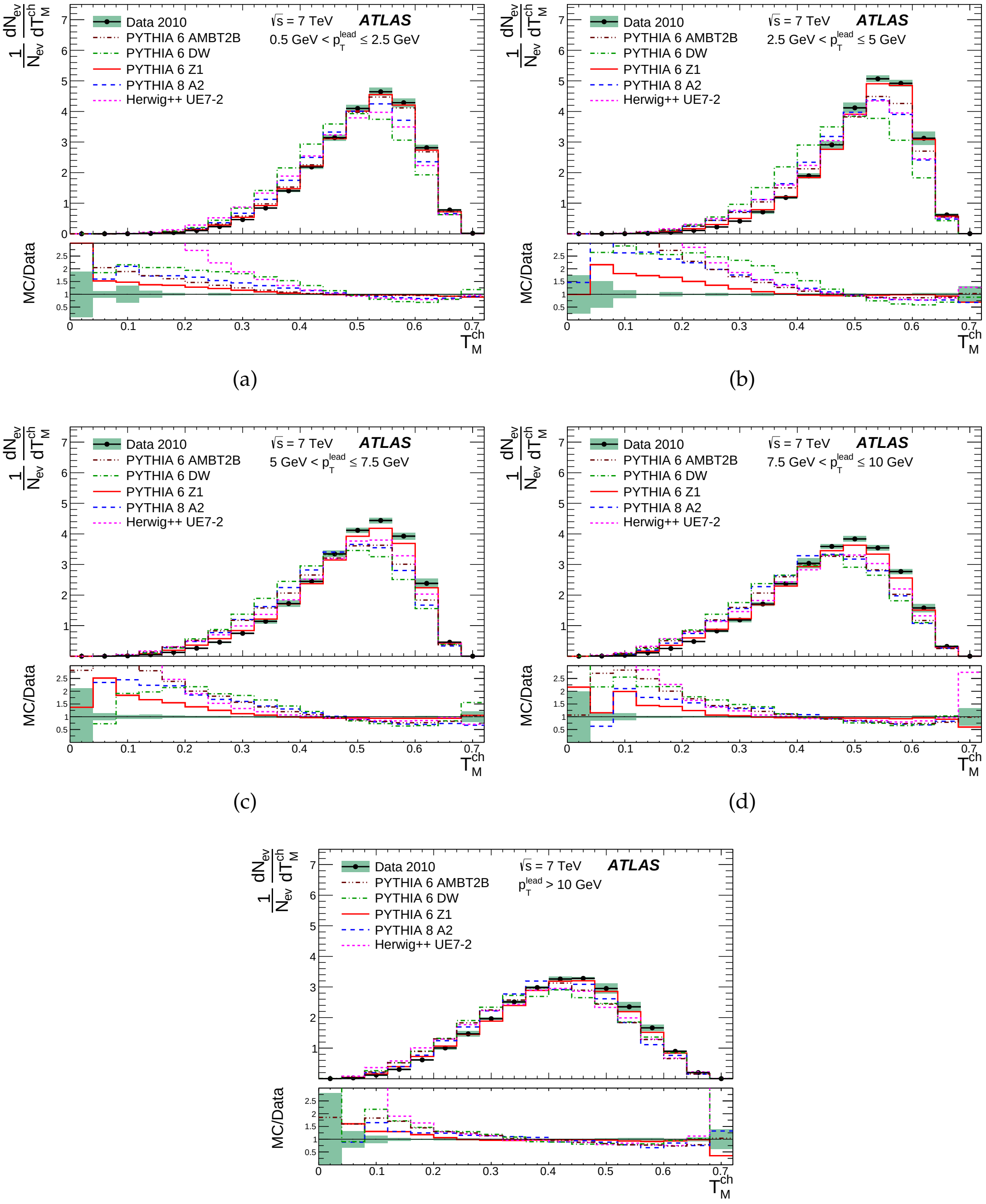
(a) (b)

(c) (d)

(e)

Figure 13.2: Normalized distributions of the transverse thrust using at least six tracks with $|\eta| < 2.5$ and $p_T > 0.5$ GeV for different bins in transverse momentum of the leading charged particle, p_T^{lead}. The error bars drawn indicate statistical uncertainties only. The shaded area represents the combined statistical and systematic uncertainties. Where not visible, the statistical uncertainty is smaller than the marker size [93].

can be confirmed in the thrust minor. The sphericity is more sensitive to the transition in p_T^{lead} and the shift to lower values is more dominant compared to the two other event shapes. After the spectrum is broadened during the transition to the intermediate p_T^{lead} bins (Fig. 13.4a → Fig. 13.4c), it narrows down again for the highest p_T^{lead} bins (Fig. 13.4c



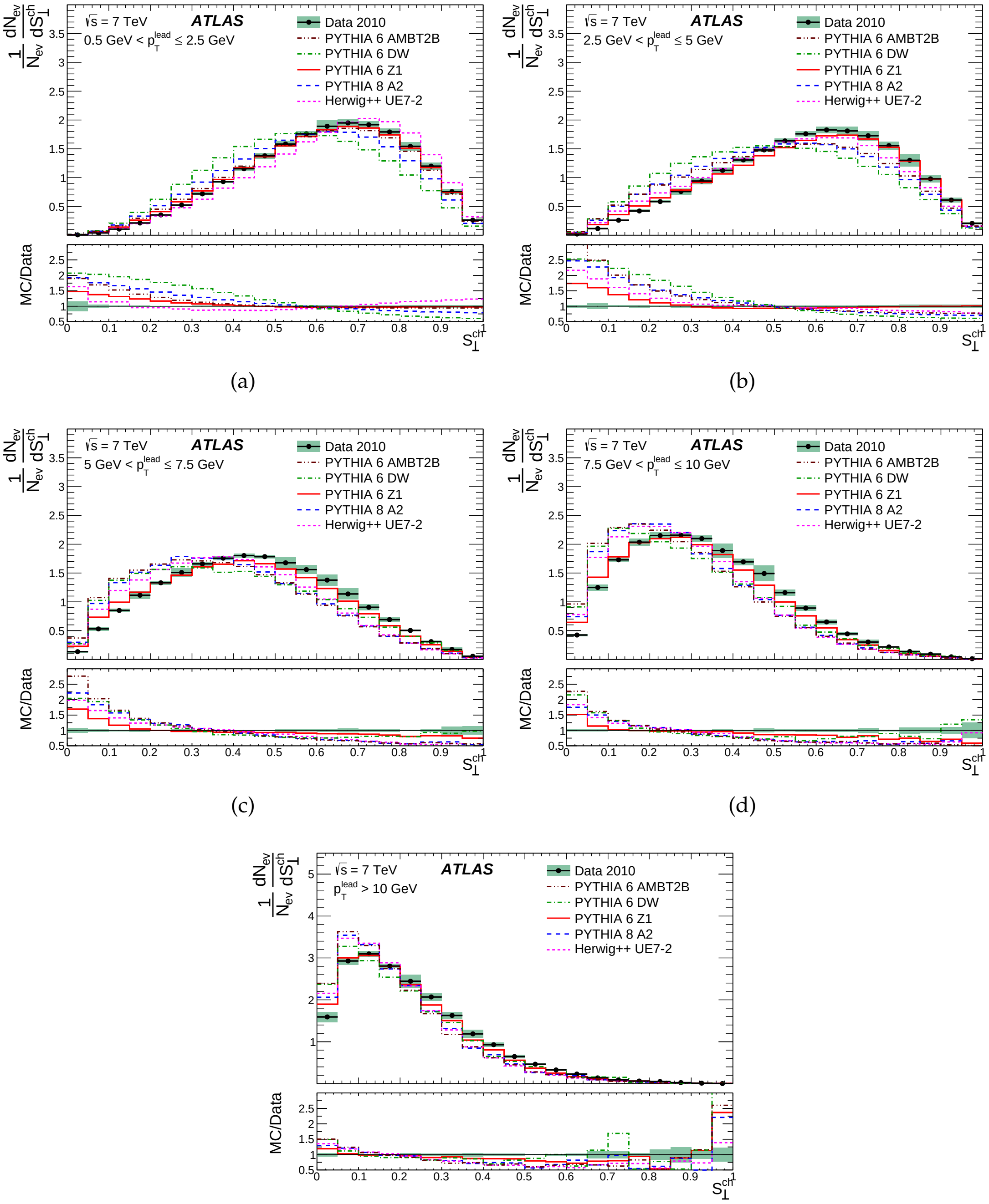
(a) (b) (c) (d) (e)

Figure 13.3: Normalized distributions of the thrust minor using at least six tracks with $|\eta| < 2.5$ and $p_T > 0.5$ GeV for different bins in transverse momentum of the leading charged particle, p_T^{lead}. The error bars drawn indicate statistical uncertainties only. The shaded area represents the combined statistical and systematic uncertainties. Where not visible, the statistical uncertainty is smaller than the marker size [93].

→ Fig. 13.4e). Fig. 13.2 - Fig. 13.4 show that the different Monte Carlo tunes follow the same trend, but the transition is less pronounced.

The ability of the different Monte Carlo tunes to describe the data is unchanged when considering the thrust minor and transverse thrust. PYTHIA 6 Z1 provides the best description throughout all distributions and PYTHIA 6 DW is furthest away. The tunes tuned to underlying event data show their strength in the higher p_T^{lead} bins, while Monte Carlo predictions tuned to minimum bias data are better at the lower bins. Again, PYTHIA 8 A2 is able to describe the data in a more accurate way than PYTHIA 6 AMBT2B, probably due to the more advanced MPI treatment.

An overview of all mean and RMS values is provided in Table 13.2 and Fig. 13.5.



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

Figure 13.4: Normalized distributions of the transverse sphericity using at least six tracks with $|\eta| < 2.5$ and $p_T > 0.5$ GeV for different bins in transverse momentum of the leading charged particle, p_T^{lead}. The error bars drawn indicate statistical uncertainties only. The shaded area represents the combined statistical and systematic uncertainties. Where not visible, the statistical uncertainty is smaller than the marker size [93].

Transverse Thrust

p_T^{lead} range	Mean	RMS
$0.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 2.5\ \text{GeV}$	0.227 ± 0.002	0.064 ± 0.008
$2.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 5.0\ \text{GeV}$	0.240 ± 0.006	0.062 ± 0.001
$5.0\ \text{GeV} < p_T^{\text{lead}} \leqslant 7.5\ \text{GeV}$	0.227 ± 0.007	0.065 ± 0.003
$7.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 10\ \text{GeV}$	0.210 ± 0.010	0.068 ± 0.005
$p_T^{\text{lead}} > 10\ \text{GeV}$	0.185 ± 0.011	0.070 ± 0.006

Thrust Minor

p_T^{lead} range	Mean	RMS
$0.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 2.5\ \text{GeV}$	0.508 ± 0.002	0.090 ± 0.010
$2.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 5.0\ \text{GeV}$	0.514 ± 0.005	0.087 ± 0.012
$5.0\ \text{GeV} < p_T^{\text{lead}} \leqslant 7.5\ \text{GeV}$	0.490 ± 0.006	0.099 ± 0.010
$7.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 10\ \text{GeV}$	0.459 ± 0.007	0.107 ± 0.009
$p_T^{\text{lead}} > 10\ \text{GeV}$	0.415 ± 0.010	0.117 ± 0.011

Transverse Sphericity

p_T^{lead} range	Mean	RMS
$0.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 2.5\ \text{GeV}$	0.618 ± 0.005	0.190 ± 0.006
$2.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 5.0\ \text{GeV}$	0.579 ± 0.013	0.204 ± 0.003
$5.0\ \text{GeV} < p_T^{\text{lead}} \leqslant 7.5\ \text{GeV}$	0.449 ± 0.019	0.206 ± 0.002
$7.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 10\ \text{GeV}$	0.337 ± 0.017	0.183 ± 0.004
$p_T^{\text{lead}} > 10\ \text{GeV}$	0.230 ± 0.024	0.157 ± 0.007

Table 13.2: The mean and RMS for each event shape distribution in different intervals of p_T^{lead}. The combined statistical and systematic uncertainties are also given, where the systematic uncertainty is obtained from the difference of unfolded results using PYTHIA 6 and HERWIG++ MC predictions.

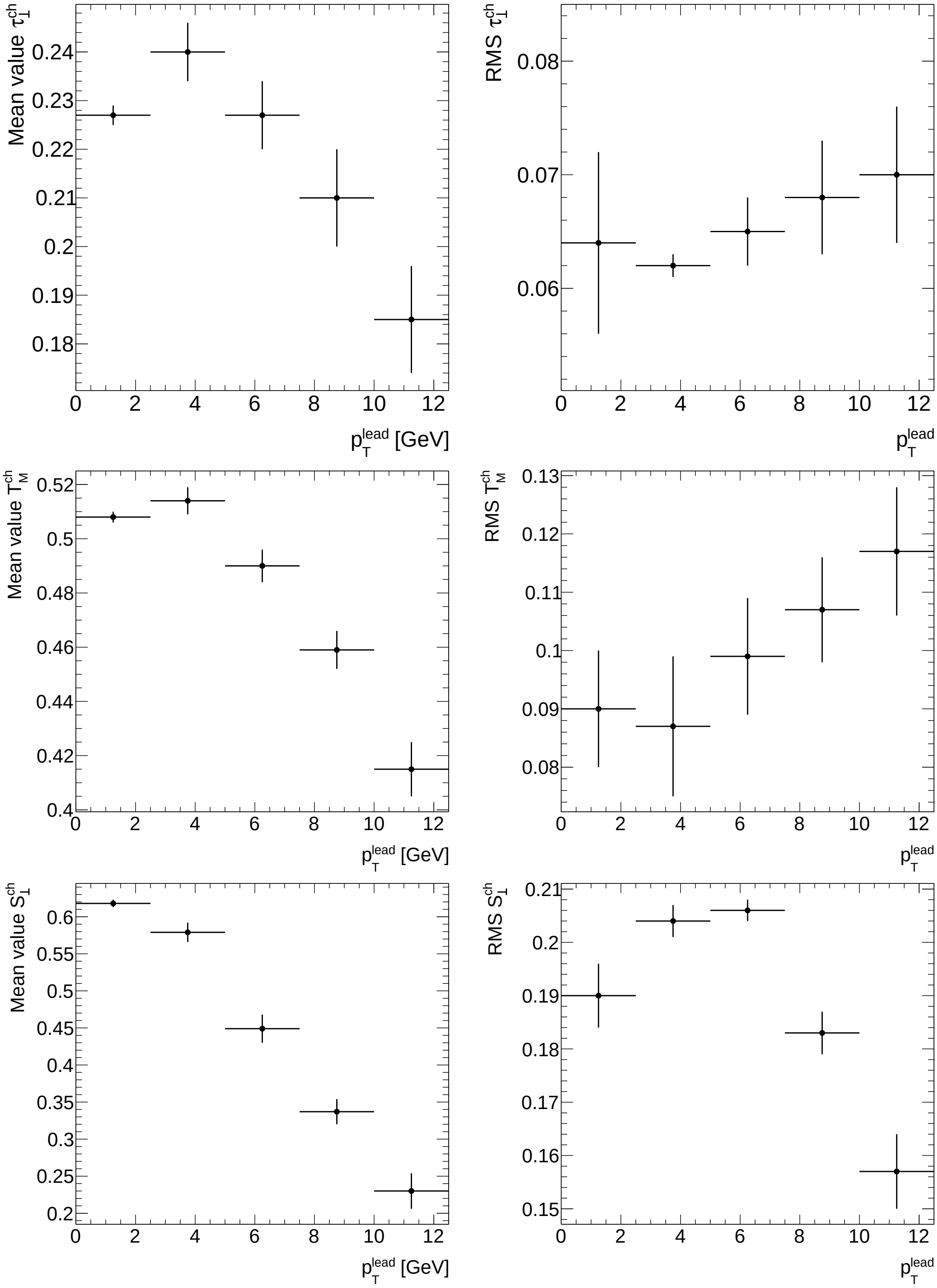


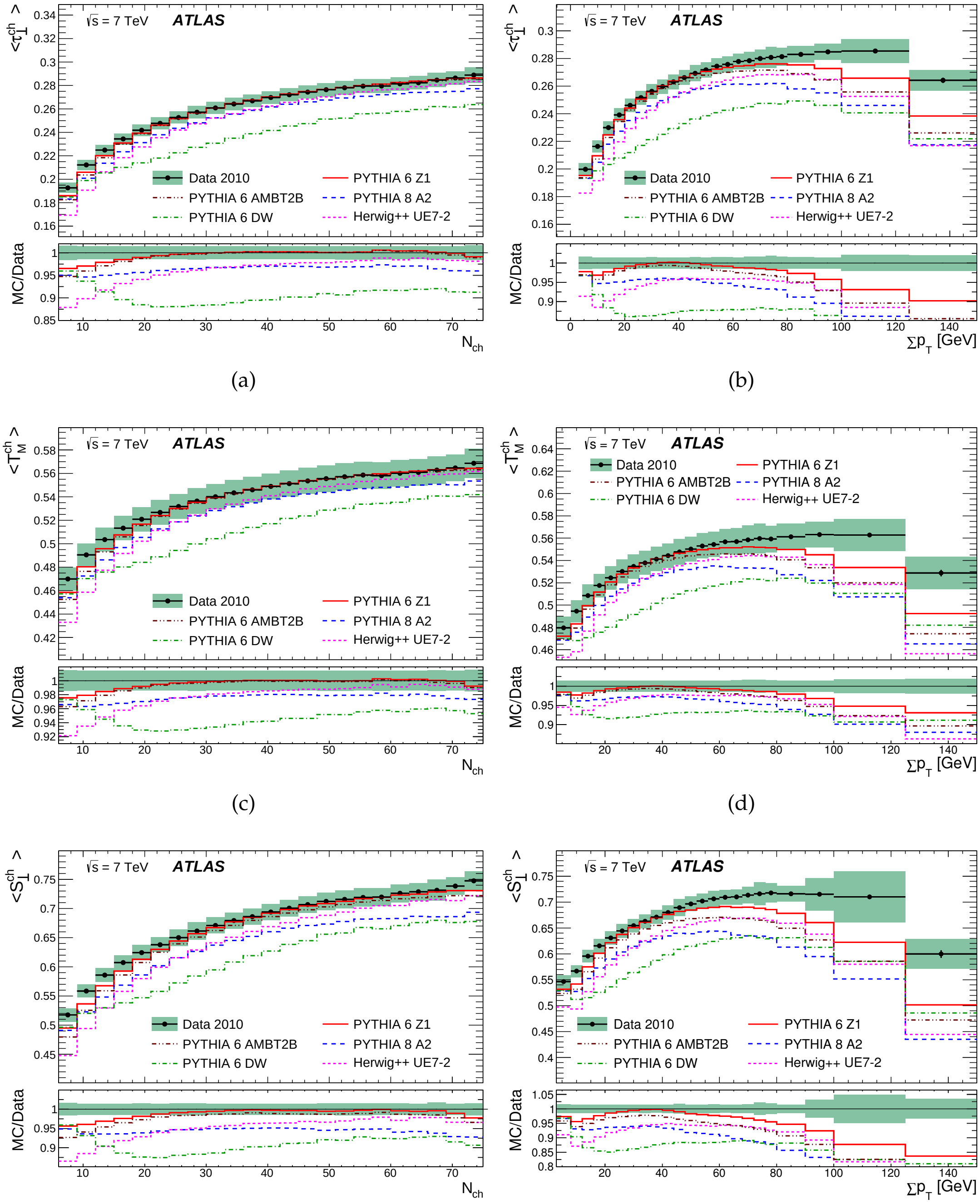
Figure 13.5: Graphical presentation of results from Table 13.2. Top row shows transverse thrust, middle one thrust minor and in the bottom row, values for sphericity are shown. Left column shows the mean values, the RMS is plotted on the right side. Last bin expands to $+\infty$.

In addition to the 1-dimensional spectra, two 2-dimensional distributions were measured. The mean values of event shapes were plotted versus the charged particle multiplicity, N_{ch}, and versus the scalar sum of the charged particle transverse momentum, $\sum p_T$, in the event. N_{ch} was chosen, because it is known to be very sensitive to the underlying event. $\sum p_T$ is another measure of the hardness of an event, like the p_T^{lead} in the 1-dimensional distributions. The results are shown in Fig. 13.6.

The mean values rise with increasing charged multiplicity, but the slope decreases at about $N_{ch} = 30$. For lower values of N_{ch}, the mean values of the event shape variables correspond to less spherical events, and at higher N_{ch}, the events are more spherical. Again, all Monte Carlo tunes predict less spherical events over the full shown N_{ch} range. However, PYTHIA 6 Z1 and AMBT2B are able to describe the data for $N_{ch} > 20$ and are very close for smaller multiplicities. The three remaining Monte Carlo predictions fail to describe the data. PYTHIA 6 DW is again furthest away, while PYTHIA 8 A2 and HERWIG UE7-2 are compatible above $N_{ch} \sim 25$. Below that, PYTHIA 8 A2 is able to provide a better agreement than HERWIG UE7-2. A measurement provided by the ALICE collaboration [96] is compatible to this result.

For the event shape variables as function of the $\sum p_T$, the distributions start to rise quickly for low $\sum p_T$ and flatten until they drop down at the highest values of $\sum p_T$. As $\sum p_T$ is a measure of the "hardness" of the event, based on the previous observations, a different behaviour was expected. At low values of $\sum p_T$, i. e. "soft" events, the mean values indicate more dijet-like events than in harder events. That is the exact opposite to the observations made before. However, the strong correlation between N_{ch} and $\sum p_T$ can explain the behaviour. The drop in mean values at $\sum p_T \sim 120$ GeV indicates that the events become more dijet-like again if going to higher $\sum p_T$, as expected.

While the Monte Carlo was able to describe the shape of the data for the event shape–N_{ch} distributions over a broad range, it fails to describe the event shape variables as a function of $\sum p_T$. The drop at higher $\sum p_T$ starts earlier than in data. In addition, and as observed before, Monte Carlo predicts consistently less spherical events. No tune is able to describe the data. PYTHIA 6 Z1 continues to give the best description, followed by PYTHIA 6 AMBT2B.



(a) (b) (c) (d) (e) (f)

Figure 13.6: Mean values of the transverse thrust, thrust minor and transverse sphericity (top to bottom), using at least six tracks with $|\eta| < 2.5$ and $p_T > 0.5$ GeV, versus charged particle multiplicity (left) and versus scalar sum of charged particle transverse momentum (right) of the event. The shaded area represents the combined statistical and systematic uncertainties. Where not visible, the statistical uncertainty is smaller than the marker size [93].

The three event shape observables transverse thrust, thrust minor and sphericity are presented in dependency of the leading charged particle transverse momentum, the charged particle multiplicity and the scalar sum of the transverse momenta of charged particles in the event. Events with at least six tracks in $|\eta| < 2.5$ and $p_T > 500$ MeV in $\sqrt{s} = 7$ TeV proton–proton collisions are used. The data is compared to predictions of five different Monte Carlo tunes.

An overall conclusion can be drawn: With increasing p_T^{lead}, as well as with increasing N_{ch}, the distributions shift towards less spherical events. An exception is the distribution of mean values of event shapes against $\sum p_T$. Here, the correlation between N_{ch} and $\sum p_T$ causes events with low $\sum p_T$ to appear more dijet-like. A special feature of this distribution is that, in contrast to all other results, the Monte Carlo is not only shifted but also fails to describe the shape of the measured data. The drop to less spherical events at higher $\sum p_T$ starts earlier in Monte Carlo compared to data, and all Monte Carlo tunes predict less spherical events than observed in data. While PYTHIA 6 Z1 shows the best performance, PYTHIA 6 DW is always furthest away from the data. This shows that tuning to LHC data is substantial for providing accurate predictions for LHC analyses.

For softer events, Monte Carlo tuned to minimum bias analyses give a better description than those tuned to underlying event data. For harder events, it is the other way around and the underlying event tunes provide a better description. But as no Monte Carlo prediction is able to describe the data, this analysis provides complementary input to previous measurements used for Monte Carlo tuning. The results are published in the HEPdata format [97], in order to be available for future Monte Carlo tuning.

Part V

MEASUREMENT OF THE DOUBLE-DIFFERENTIAL DIJET CROSS SECTION AT $\sqrt{s} = 8$ TeV

14

MOTIVATION

Jet production is one of the most dominant processes at hadron colliders and covers a wide range in the available phase space. Cross sections for proton–(anti)proton collisions as a function of center-of-mass energy for different channels are shown in Fig. 14.1. The jet production is for example about two orders of magnitude above the Z-boson production. Therefore, very high statistics can be expected. Jets can be measured from very low transverse momenta, starting below $p_{T,jet} = 20$ GeV, up to $p_{T,jet} > 2$ TeV. The results of the cross section measurement can be used to test perturbative QCD predictions up to very high scales and therefore improves the understanding of QCD and its coupling constant α_S. In addition, it can be used to learn more about the proton structure, especially the parton distribution functions of the gluon at high momentum

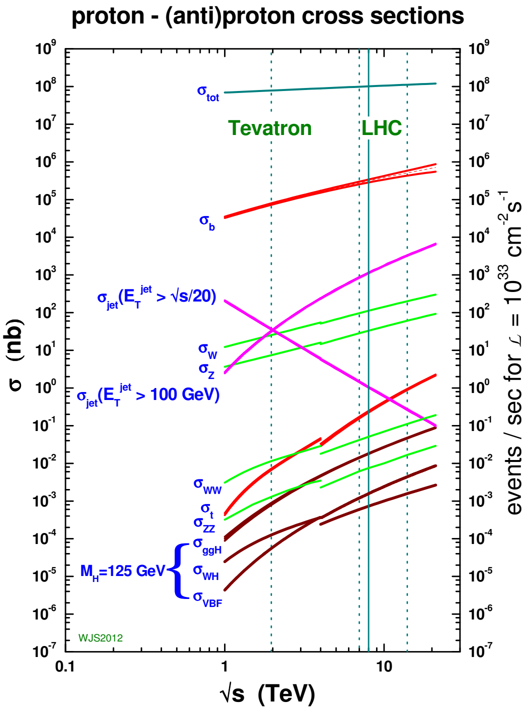


Figure 14.1: Cross sections for different production processes as a function of center-of-mass energy in proton–(anti)proton collisions. The solid vertical line represents $\sqrt{s} = 8$ TeV. The dashed lines corresponds to earlier LHC energies ($\sqrt{s} = 7$ TeV), the LHC design (and possible future) energy $\sqrt{s} = 14$ TeV as well as the center-of-mass energy of the Tevatron $\sqrt{s} = 1.96$ TeV [24].

scale Q^2 and high Bjorken-x. Fig. 14.2 shows the phase space covered by $\sqrt{s} = 7$ TeV pp collisions. For $\sqrt{s} = 8$ TeV, even higher Q^2 can be reached. Two variables with one

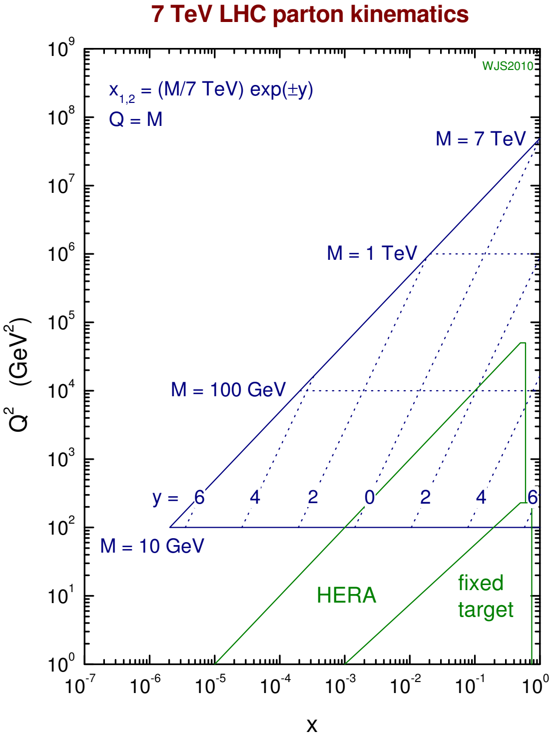


Figure 14.2: Phase space in x, Q^2 covered for $\sqrt{s} = 7$ TeV pp collisions drawn in blue. Phase space covered by HERA and fixed target experiments indicated in green [24].

of the highest sensitivity to PDFs are the dijet mass and y^*. The dijet mass is defined as

$$m_{12} = \sqrt{2p_{T,1}p_{T,2}\left[\cosh(\Delta y) - \cos(\Delta\Phi)\right]} \tag{14.1}$$

and y^* is half the separation in rapidity of the two leading jets:

$$y^* = |y_1 - y_2|/2. \tag{14.2}$$

In a simplified $2 \to 2$ process with two massless outgoing partons, y^* describes the rapidities of the partons in the dijet center-of-mass frame. The double-differential cross section as a function of dijet mass m_{12} and rapidity separation $y^* = |y_1 - y_2|/2$ of the two leading jets is defined as

$$\frac{d^2\sigma}{dm_{12}dy^*} = \frac{1}{\epsilon\mathcal{L}}\frac{1}{\Delta m_{12}\Delta y^*}. \tag{14.3}$$

The collected luminosity is denoted as $\mathcal{L}$, being 20.3 fb^{-1} for this measurement. N is the number of events in the dijet mass range Δm_{12} and y^* range Δy^*. To correct for detector resolution and efficiency effects, the correction factor ϵ is applied. Earlier studies at lower center-of-mass energies and with less statistics [98, 99] have shown agreement over a wide range between data and theory predictions, but first hints of a disagreement are observed at very high dijet masses. The higher center-of-mass energy as well as the increased amount of data available for this analysis can give more information about the agreement or disagreement at highest dijet masses.

15

ANALYSIS

This section descirbes the analysis steps to obtain the results. At first, the event and jet selection is presented in Sec. 15.1. The binning optimization in dijet mass bins is described in Sec. 15.2 followed by the pile-up reweighting in Monte Carlo in Sec. 15.3. Some basic jet distributions are shown and discussed in Sec. 15.4. The correction to particle level is presented in Sec. 15.5 and Sec. 15.6 evaluates the treatment of statistical and systematic uncertainties.

15.1 EVENT AND JET SELECTION

A set of *data quality*, DQ, cleaning and kinematic cuts are applied in the analysis. The details are explained in the following sections.

Used data and Monte Carlo samples

The analysis used the full dataset of the 2012 LHC data taking. This corresponds to an integrated Luminosity of $\mathcal{L} = 20.3\ \text{fb}^{-1}$. Therefore, the luminosity is increased by a factor of five compared to the dijet cross-section measurement from ATLAS in 2011. This and the higher center-of-mass energy of $\sqrt{s} = 8$ TeV contributes significantly to decreased statistical uncertainties compared to earlier measurements. The cumulative luminosity is shown in Fig. 15.1. In order to achieve a higher integrated luminosity in a shorter time period, compared to the $\sqrt{s} = 7$ TeV data taking, the instantaneous luminosity was increased by a factor of two. This leads to a maximum instantaneous luminosity of $\sim 8 \cdot 10^{33}\ \text{cm}^{-2}\text{s}^{-1}$. The peak luminosity delivered at ATLAS as function of time from 2010 to 2012 is shown in Fig. 15.2. The high instantaneous luminosity leads to a higher number of simultaneous proton–proton interactions. In 2012, the average number of interactions per bunch crossing peaks at around 17 and reaches more than 40 simultaneous interactions per bunch-crossing, as shown in Fig. 15.3. To ensure a good correction for the pile-up conditions, a new per-jet pile-up correction based on jet areas has been developed and applied (Sec. 8.2.1).

As an input for the unfolding procedure, to calculate the jet energy scale uncertainties and for some first comparisons with data, Monte Carlo produced with the PYTHIA 8 generator and the AUET2B tune using the CT10 PDF was used. The Monte Carlo was produced in eight slices, named JZ0W to JZ7W, of leading jet transverse momentum on parton level ($\hat{p}_T$). Table 15.1 gives an overview over the covered $\hat{p}_T$ range and the available number of events per slice. Slices two to six are produced *flat* in $\hat{p}_T$, which results in constant statistics over $\hat{p}_T$, in contrast to decreasing statistics over $\hat{p}_T$ accord-

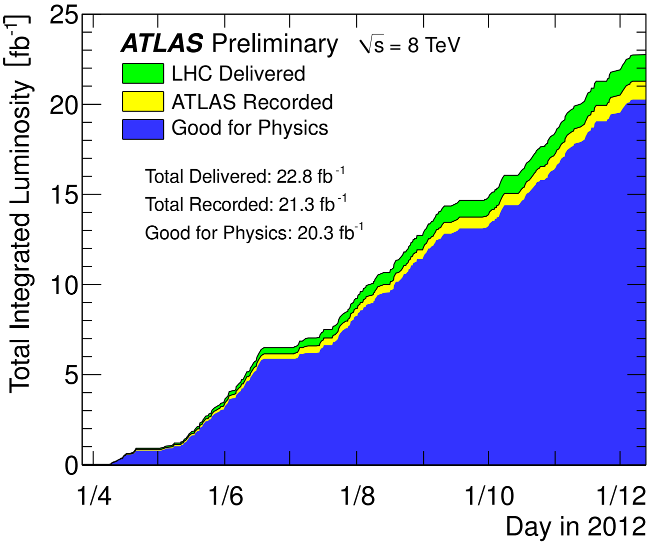


Figure 15.1: Cumulative luminosity versus time delivered to (green), recorded by ATLAS (yellow), and certified to be good quality data (blue) during stable beams and for pp collisions at 8 TeV center-of-mass energy in 2012 [82].

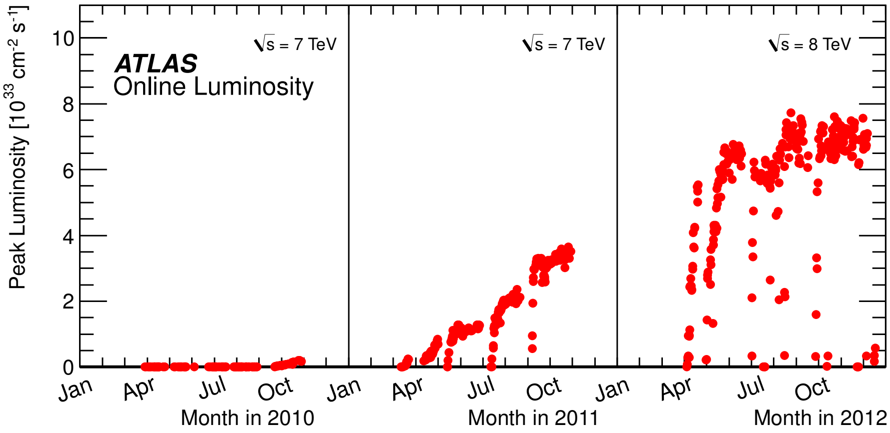


Figure 15.2: Peak instantaneous luminosity delivered to ATLAS per day versus time during the p-p runs of 2010,2011 and 2012. The online luminosity measurement is used for this plot [82].

ing to the physical cross section. These flat samples decrease statistical uncertainties at the upper edge of the slices and avoids *spikes* in the uncertainties due to high weights in the resulting spectrum. Fig. 15.4 shows the combination of the eight Monte Carlo samples. Requiring a minimum jet transverse momentum of 40 GeV removes any contributions from the JZ0W sample. The top plot shows the unweighted Monte Carlo as it is produced initially. The production with constant statistics as function of $\hat{p}_T$ is visible. Only the JZ7W sample is produced as a correctly falling spectrum. Entries above the $\hat{p}_T$ limits given in Table 15.1 are caused by contributions from the underlying event and pileup to the jet energy. The bottom plot in Fig. 15.4 shows the weighted spectrum, reproducing the expected falling spectrum in p_T.

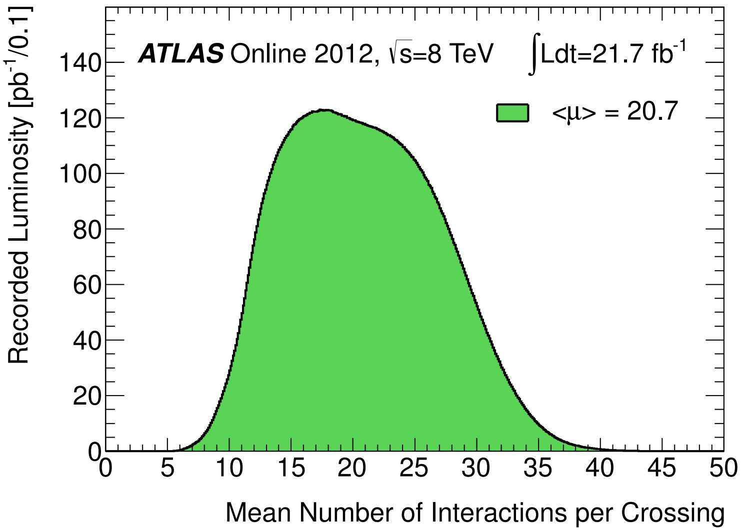


Figure 15.3: Luminosity-weighted distribution of the mean number of interactions per crossing for 2012 [82].

Monte Carlo slice	covered $\hat{p}_T$ range	generated events
JZ0W	0-20 GeV	1350000
JZ1W	20-80 GeV	1499994
JZ2W	80-200 GeV	5999034
JZ3W	200-500 GeV	5997214
JZ4W	500-1000 GeV	5837256
JZ5W	1000-1500 GeV	2996082
JZ6W	1500-2000 GeV	2993651
JZ7W	> 2000 GeV	2991955

Table 15.1: Splitting of Monte Carlo in $\hat{p}_T$ slices and available statistics.

Data quality requirements

In order to provide a well-defined experimental environment, events in which parts of the detector were dysfunctional or not yet fully functional were removed. A list of luminosity blocks① in which the detector was fully operational is provided by ATLAS DQ. To reduce non-collision background, such as cosmic particles, at least two tracks need to be associated with the primary interaction vertex.

Trigger selection

As described earlier, the event rate at the LHC is too high to be able to record every event. Dedicated jet triggers, selecting events with at least one jet above a certain threshold in transverse energy, are used in this analysis. After selecting jet events in

---

① A luminosity block is a short interval of usually 1 minute with approximative constant beam and detector conditions

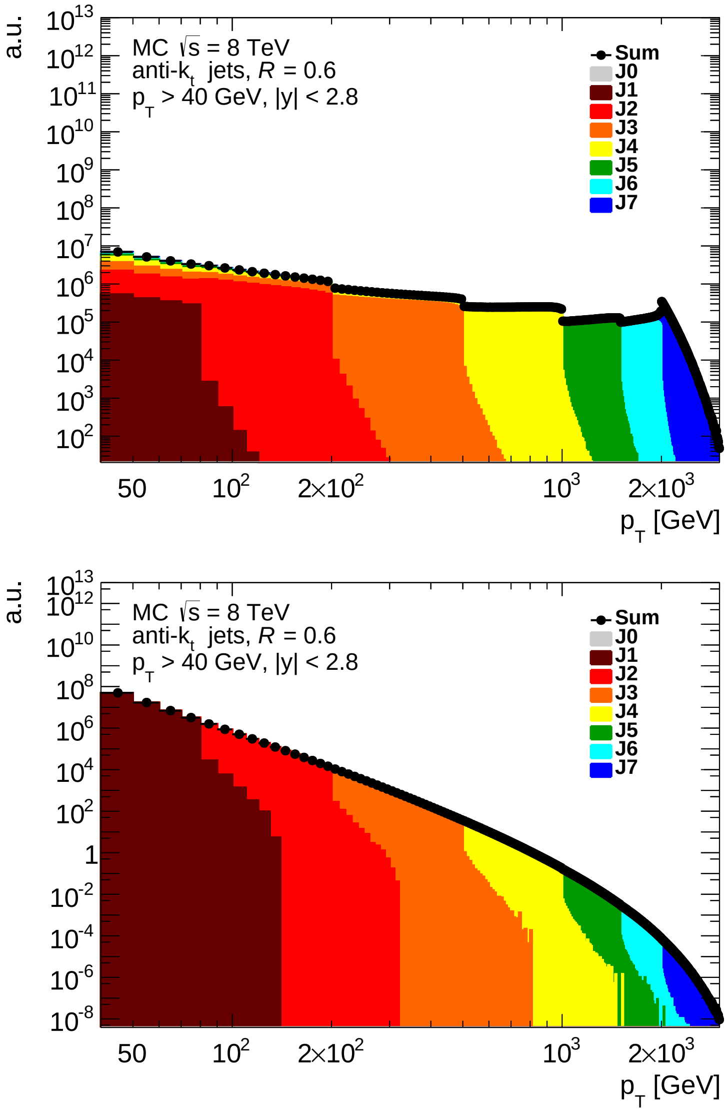


Figure 15.4: Combination of the eight Monte Carlo samples as a function of the jet transverse momentum at particle level. The top plots shows the unweighted spectrum, as it is originally produced, and the bottom plot shows the weighted p_T spectrum.

the trigger system, the event rates are still too high. Therefore, so called prescales were introduced. Not every event fulfilling the trigger requirements is stored, but only every x-th event, where x is called prescale factor. Only for triggers with a high transverse momentum threshold the prescale is one, which means that every event is recorded. The lower the threshold, the higher the prescale. The prescales are chosen in a way that the number of recorded events at the trigger threshold is approximately equal for all triggers.. As the jet spectrum is steeply falling, the jet transverse momentum distribution is mainly populated at lower values close to the threshold. Statistics at very high transverse momenta are therefore low. To have the highest possible statistical significance, the highest trigger which is fully efficient was used. The *EF_j360_a4tchad* trigger is the lowest unprescaled trigger. Therefore, no higher triggers are in use. The trigger name can be decoded as follows: *EF* stands for event filter, the highest trigger level. This is followed by a j for jet trigger and the threshold on transverse energy in GeV. a4 describes the used jet algorithm to reconstruct the jets: **a**nti-k_t with a radius parameter of $R = 0.$**4**. tchad means that jets are build from **t**opological **c**luster including cells in the **had**ronic calorimeter. For the dijet analysis, trigger requirements are applied to the offline leading jet.

The efficiencies for the different jet triggers derived for anti-k_t jets with a radius parameter of $R = 0.6$ are shown in Fig. 15.5. The thresholds are used for both jet Radii, $R = 0.6$ and $R = 0.4$. In order to obtain the trigger efficiency ε for each trigger tested, a reference trigger needs to be defined. This reference trigger has to be fully efficient significantly below the threshold of the test trigger. To ensure that, a random trigger is used to determine the efficiencies of the lowest jet triggers. From there on, a bootstrapping method is applied, using the highest possible trigger as a reference to test the next jet trigger. Table 15.2 shows the reference triggers for each trigger tested. To determine the efficiency, an unbiased set of events needs to be selected. This is done by selecting events which passed the reference trigger. The efficiency then is defined as the percentage of events out of the unbiased sample which pass the test trigger.

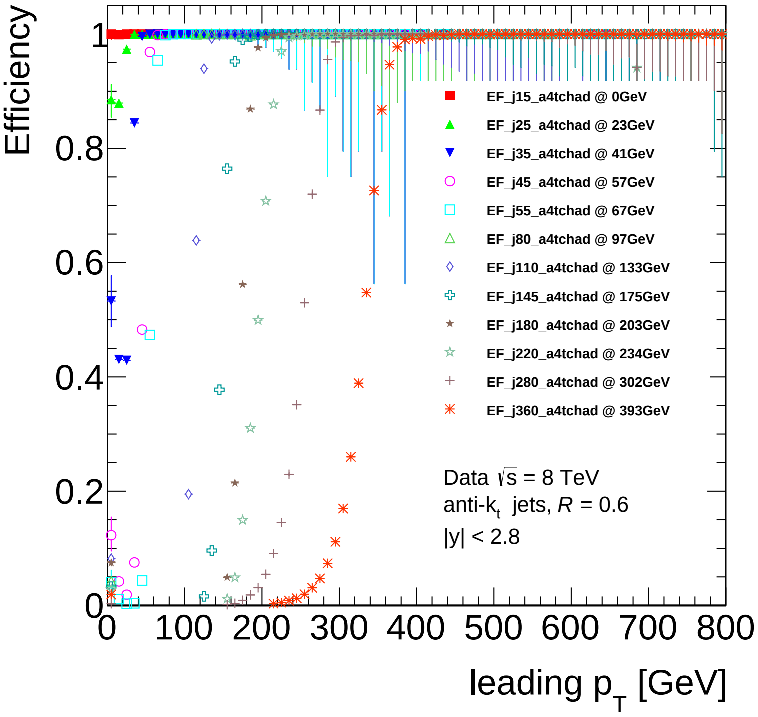


Figure 15.5: Trigger efficiencies for single jet trigger. Efficiency points are defined at a efficiency of 99.5%. Additional 5% of the value found was added to account for differences in the methods for determining the threshold. Only jets within $|y| < 2.8$ were considered.

$$\varepsilon = \frac{\text{\#events passing reference and test trigger}}{\text{\#events passing reference trigger}} \qquad (15.1)$$

As the trigger decisions are based on jet transverse momenta and only the leading jet is required to be triggered efficiently in the analysis, the efficiency is plotted as a function of p_T^{lead}. The transverse momentum from which on the trigger is used is determined as the point at which the trigger is 99.5% efficient. There is an ambiguity in determining this 99.5% efficiency point. Three possible and tested methods are:

- Using the lowest leading jet p_T with an efficiency above 99.5%.
- Fitting the efficiency curve and evaluating the fitted function at $\varepsilon = 99.5\%$.
- Starting from $p_{T,\text{min}} = 0$, calculating $\epsilon = \frac{\int_{p_{T,\text{min}}}^{\infty} dp_{T,\text{test}}}{\int_{p_{T,\text{min}}}^{\infty} dp_{T,\text{test}}}$ and increasing $p_{T,\text{min}}$ until $\epsilon > 99.5\%$.

While the first method is only useful for monotonically rising efficiency distributions and therefore not used very often, the last two methods are both commonly used. The last method is more sensitive to small drops in the spectra and is also a more conservative approach because it ignores statistical uncertainties completely. The fit method itself is not consistent, as different fit functions can be used, leading to different results. The last method is used to determine the 99.5% efficiency point. To cover the differences between the methods, an additional 5% of the leading jet p_T, at which the method predicts a 99.5% efficiency, is added. The resulting p_T ranges, in which the single triggers are used, are summarized in Table 15.3, together with the luminosity collected by the trigger. All triggers used are central jet triggers, meaning that only jets with $|\eta| < 3.2$ are considered. As the trigger uses jets with a radius parameter of $R = 0.4$, some energy of jets above $|\eta| < 2.8$ might not be considered during the triggering process. Therefore, the efficiency drops down to about 50% at $|\eta| = 3.2$. The efficiency as a function of rapidity, calculated for the 2011 data taking, is shown in Fig. 15.6.

The applied prescales cause a distortion in the spectra. To extrapolate back to the original number of events (before prescales) a luminosity weight is applied. The luminosity weight is proportional to the prescale and the reciprocal collected luminosity in

Tested trigger		reference trigger
EF_j15_a4tchad	→	EF_rdo_filled_NoAlg_Eff
EF_j25_a4tchad	→	EF_rdo_filled_NoAlg_Eff
EF_j35_a4tchad	→	EF_rdo_filled_NoAlg_Eff
EF_j45_a4tchad	→	EF_rdo_filled_NoAlg_Eff
EF_j55_a4tchad	→	EF_rdo_filled_NoAlg_Eff
EF_j80_a4tchad	→	EF_j35_a4tchad
EF_j110_a4tchad	→	EF_j55_a4tchad
EF_j145_a4tchad	→	EF_j80_a4tchad
EF_j180_a4tchad	→	EF_j110_a4tchad
EF_j220_a4tchad	→	EF_j110_a4tchad
EF_j280_a4tchad	→	EF_j110_a4tchad
EF_j360_a4tchad	→	EF_j180_a4tchad

Table 15.2: Reference triggers used to determine the trigger efficiencies of the tested triggers. EF_rdo_filled_NoAlg_Eff is a random trigger, only requiring the buffers in the trigger system to be filled.

Trigger name	p_T range [GeV]	Luminosity [μb^{-1}]
EF_25_a4tchad	100 - 136	0.441786
EF_j80_a4tchad	136 - 190	2.3164e+06
EF_j110_a4tchad	190 - 204	9.81141e06
EF_j145_a4tchad	204 - 252	36.2647e06
EF_j180_a4tchad	252 - 304	78.7753e+06
EF_j220_a4tchad	304 - 404	261.379e+06
EF_j280_a4tchad	404 - 522	1164.69e+06
EF_j360_a4tchad	>522	20.277e+09

Table 15.3: Triggers used for the dijet analysis, covered p_T range and the collected integrated luminosity for each trigger are listed. Lower limits in p_T are inclusive, upper exclusive.

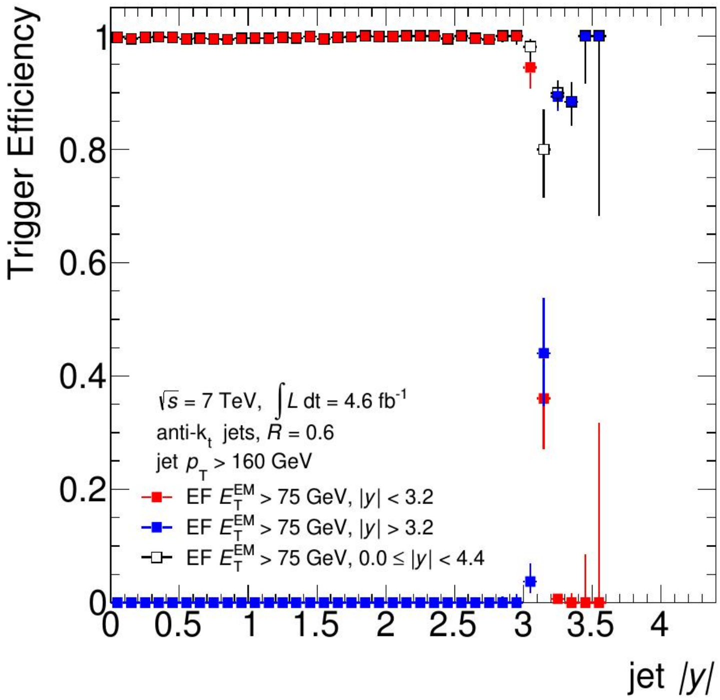


Figure 15.6: Trigger efficiency as a function of $|y|$ for the EF_j75_a4tc_EFFS trigger from the 2011 data taking period [100].

pb while the trigger was active. This corrects the prescales and normalizes the distribution to 1 pb^{-1} as shown in Fig. 15.7.

Kinematic and jet quality cuts

The fiducial area of the analysis is defined by kinematic cuts on the two leading jets:

- $nJets \geqslant 2$
- $|y_{1,2}| \leqslant 2.8$

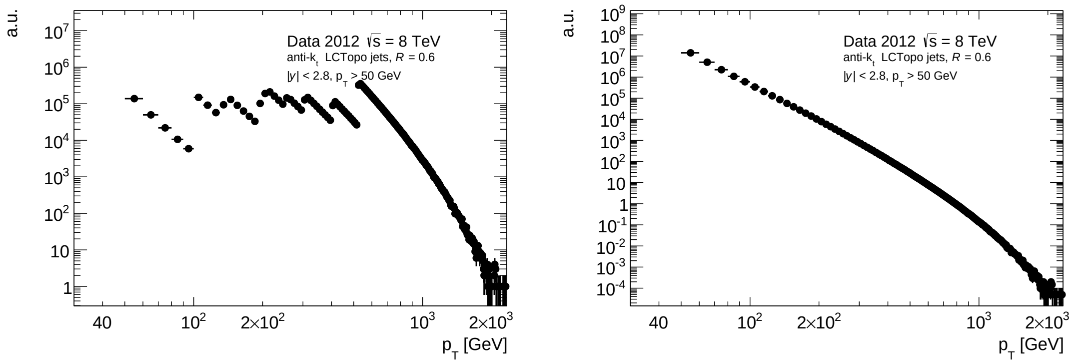


Figure 15.7: Inclusive jet transverse momentum distribution in data. Left side shows distribution before prescale correction, right side luminosity corrected results.

- $p_{T,1} \geqslant 100$ GeV
- $p_{T,2} \geqslant 50$ GeV
- $y^* = \frac{|y_1 - y_2|}{2} \leqslant 2.8$

where index 1 indicates the leading and index 2 the subleading jet. Jets below a transverse momentum of 50 GeV cannot be consistently well measured, mostly due to the high amount of pile-up and the low energy resolution. Choosing different thresholds for the leading and the subleading jet in p_T originates from instabilities, while generating NLO predictions with NLOjet++ when using symmetric cuts. As mentioned before, the trigger starts to become inefficient at around $|y| > 3.0$. In order to operate with a fully efficient trigger without lowering the trigger thresholds significantly, the two leading jets are required to be within $|y_{1,2}| \leqslant 2.8$. This directly results in an implicit cut of $y^* \leqslant 2.8$.
In addition to these kinematic cuts, events with poorly measured or fake jets are being removed:

- Both leading jets neither bad nor ugly
- No other ugly jet with $p_T^{\text{other}} > 0.3 \cdot p_{T,2}$

The two leading jets should not be flagged as *bad* or *ugly* to avoid fake jets or poorly measured jets. Jets which are tagged as ugly are not well measured and the true energy can be much higher than the measured one. To ensure that the leading and subleading jets are identified correctly, events which contain any other ugly jet with a transverse momentum $p_T^{\text{other}} > 0.3 \cdot p_{T,2}$ are being removed.

15.2 BINNING OPTIMIZATION

The choice of the binning highly influences the results of the analysis. Using broad bins results in better statistics and less bin migrations. In the case of very high bin migrations, the unfolding procedure is not able to perform satisfactorily. On the other hand, physics information is lost when using broad bins. Using for example only one bin in dijet mass will result in the loss of information about the dijet dependency. Using only few bins will include the dependency, but the finer the bins, the higher the information about this dependency, as long as enough statistics are accessible. Therefore a compromise between good statistics/low bin migrations and accessible physics information needs to be found. Therefore, all bins should be as narrow as possible, while fulfilling the two following requirements:

- The purity needs to be above 60%. The purity is defined as the number of events in one dijet mass bin on truth level, divided by the number of events in the same dijet mass bin on reconstruction level (Eq. (**??**)). Only events which pass reconstruction and truth level are considered.
- The statistical uncertainty should always be smaller than the systematic uncertainty.

The purity for the lowest y^* bin is shown in Fig. 15.8. The vertical red line indicates the lowest bin used in the analysis, while the horizontal red line is drawn at the aimed value of 60%. The higher the jet energy, the better the detector resolution. The influence of pile-up and underlying event activity is reduced and, as jets with a high energy are more collimated, out-of-cone effects are reduced as well. This leads to a rise in purity with increasing dijet mass. In addition, bins at higher dijet masses are broader. The limiting factor for the bin size are the bin-migrations at lower dijet masses and the lack of statistics at higher dijet masses.

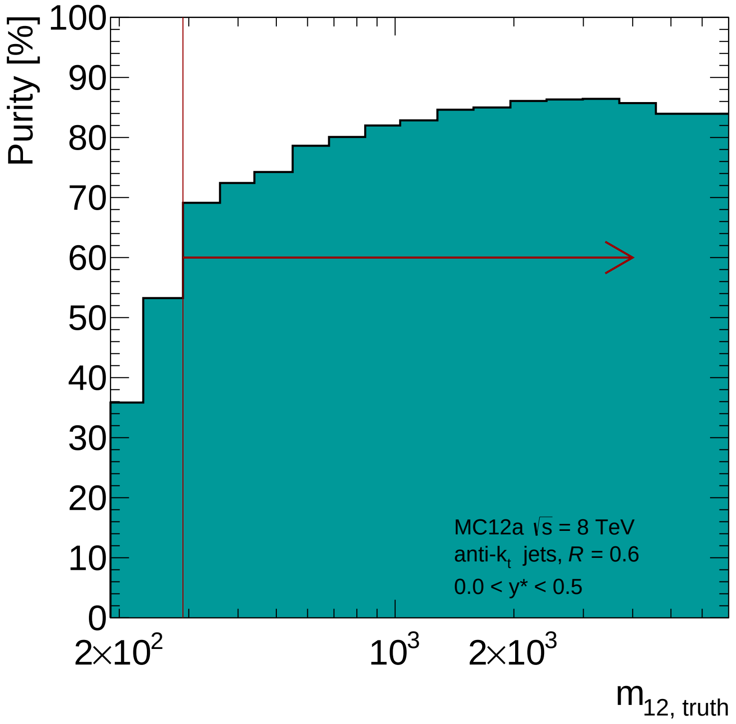


Figure 15.8: Purity vs. dijet mass on truth level. The vertical line represents the lowest m_{12}-bin used in the analysis and the vertical line marks 60% purity.

The instantaneous luminosity during the 2012 data taking was unknown when the Monte Carlo was produced. Therefore, the distribution of average interactions per bunch-crossing was simulated flat over most of the expected range with a short rise and fall at the edges. To correct for the different amount of pile-up in data and Monte Carlo, the Monte Carlo gets reweighted to fit the data's μ distribution.

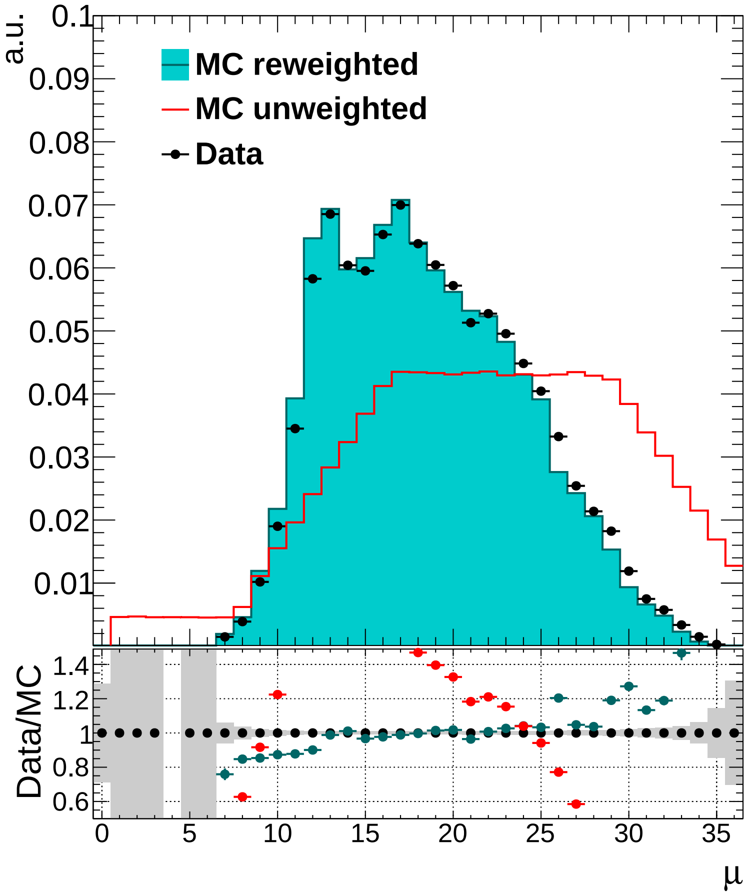


Figure 15.9: Distribution of the average number of interactions per bunch-crossing in Monte Carlo (red triangles), Monte Carlo after applied reweighting (blue) and data (black). The bottom plot shows the ratio of data over Monte Carlo. The grey band represents statistical uncertainties in data.

A full agreement between data and MC cannot be achieved due to the data selection: Reweighting factors were derived from the full dataset and Monte Carlo while Fig. 15.9 shows the distribution after analysis cuts. But the agreement has been improved and the difference between data and Monte Carlo is within a few percent over most of the available range. As the pile-up correction of the applied jet energy scale removes the dependency of pile-up, the effect of the pile-up reweighting is small and the non-agreement therefore negligible in the results.

A few basic jet/dijet distributions are presented in this section. The distributions on detector level from data are compared to LO Monte Carlo predictions. Only jets passing the event and jet selection described above are included into the plots.
Fig. 15.10 shows the transverse momentum distribution of the leading jet for two different normalizations. Normalizing the plots to luminosity and bin width enables a direct comparison of the event yield on detector level, while normalizing to the area makes it easier to compare shapes, without knowing the absolute scale of both plots.

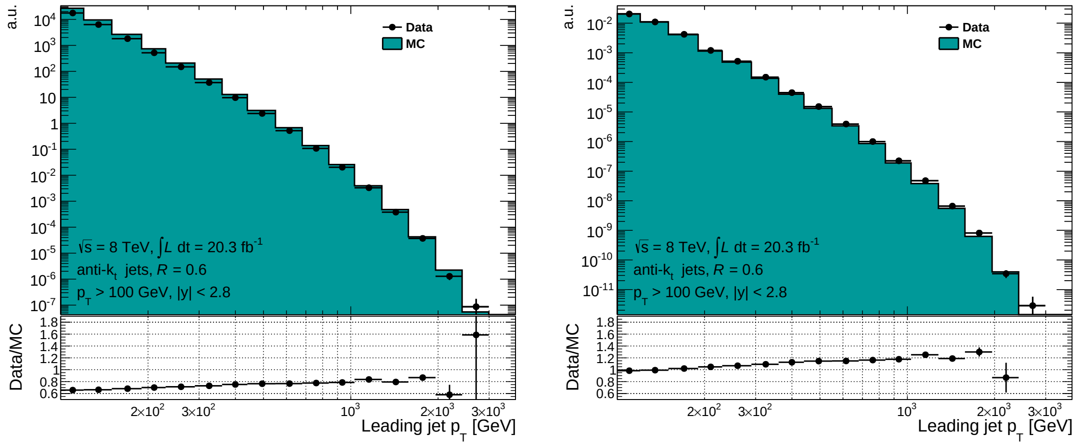


Figure 15.10: Transverse momentum spectra for the leading jet. Normalized by luminosity and bin width (left) and normalized to the area (right). Data is shown in black points, leading-order Monte Carlo predictions drawn in blue.

The leading jet transverse momentum is a steeply falling spectrum as expected. The leading-order Monte Carlo is not able to describe the data. Missing contributions from higher orders result in an increased cross section and lead to an overall shift of about 20 - 30% as seen in the left plot of Fig. 15.10. Normalizing the histogram to its area (right side of Fig. 15.10) allows a better shape comparison. However, it is not possible to specify whether the shape description is better at lower or higher transverse momenta as the global shift is dominated by highly populated bins, low p_T in this case. Besides the global offset, the shape is not described well either. Differences up to 30% can be observed. Fig. 15.11 shows the pseudorapidity and the azimuthal angle ϕ of the leading jet. While the azimuthal angle is well described, the pseudorapidity shows disagreements of up to 25%. At $|\eta| \approx 1.7$, deviations between data and Monte Carlo can be observed. This is the transition region between the tile calorimeter and the hadronic end cap calorimeter. A not well modeled detector response in this region might cause the difference.
Overall, the comparison between data and Monte Carlo shows that an unfolding method which is not sensitive to the shape of the Monte Carlo is needed.
Figures 15.12, C.1 and C.2 show the leading jet and subleading jet p_T distributions, comparing LO Monte Carlo with data. NLO contributions are not negligible and the comparisons are bound to fail. However, an approximative agreement is needed to use the Monte Carlo for unfolding. As the distributions are normalized to their integrals, an absolute evaluation is not possible. For both, the leading and the subleading jet, the

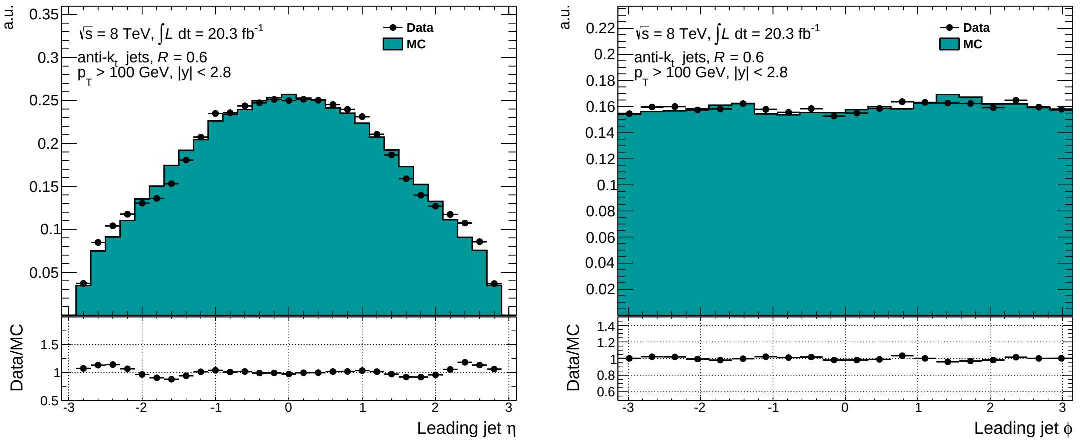


Figure 15.11: Leading jet pseudorapidity η (left) and azimuthal angle ϕ (right) of the leading jet. Data is shown in black points, leading-order Monte Carlo predictions drawn in blue. The distributions are normalized to their integrals.

Monte Carlo predicts a steeper drop towards high transverse momentum. The higher the rapidity separation, the earlier this trend changes, and the data distribution starts getting steeper. For the lowest y^* bin, the drop off is not clearly visible, only the last p_T bin is lower compared to the previous ones. With increasing y^*, the turning point of the ratios move towards lower p_T. This results in a better description of the shape at high y^* values. A special region for the subleading jet p_T distribution is between 50 GeV and 100 GeV. In a perfect $2 \to 2$ process on parton level, as used in the LO Monte Carlo, the leading and subleading jets are balanced in p_T and no contributions below 100 GeV are expected. Final state radiation/parton shower, jet energy resolution effects, pile-up and hadronization can contribute to an imbalance in p_T, and subleading jets populate the bins below 100 GeV. The Monte Carlo fails to describe the data, but NLO and non-perturbative effects play a role in these bins, therefore no accurate description is expected.

Fig. 15.13 - 15.15 show the η distribution of the leading and subleading jet normalized to their inegrals. For the lowest y^* bin, the leading and subleading jet η distributions are similar to each other, peaking at around zero and falling off towards more forward regions. The more forward a jet is, the higher the energy needed to get the same transverse momentum. Therefore, the p_T cut determines the shape of the distribution. With increasing y^*, the effects of the restrictions of the rapidity separation become visible. The central area is more pronounced and the shoulders get flatter. Leading jets above $|y| > 1.8$② need to have a subleading jet in the central region. Otherwise, the subleading jet would have been outside $|y| < 2.8$ and the event is rejected. The leading jet is always preferred to be in the central region, due to its higher transverse momentum. At $1.5 < y^* < 2.0$, all central jets are rejected. A central jet at $y = 0$ needs a subleading jet at $3.0 < |y| < 4.0$. This configuration is again rejected by the rapidity cut. Leading jets again tend to be more central. Therefore, the leading jet η is tilted towards the central region, while the subleading jet η is tilted towards higher η. The $1.0 < y^* < 1.5$ slice is in the transition region. Central jets are not completely rejected, but their *dijet partner* needs to be at high η, which is suppressed due to the high energy needed to

---

② The mass of the jet is usually small compared to its energy. Therefore η and y are approximately equal.

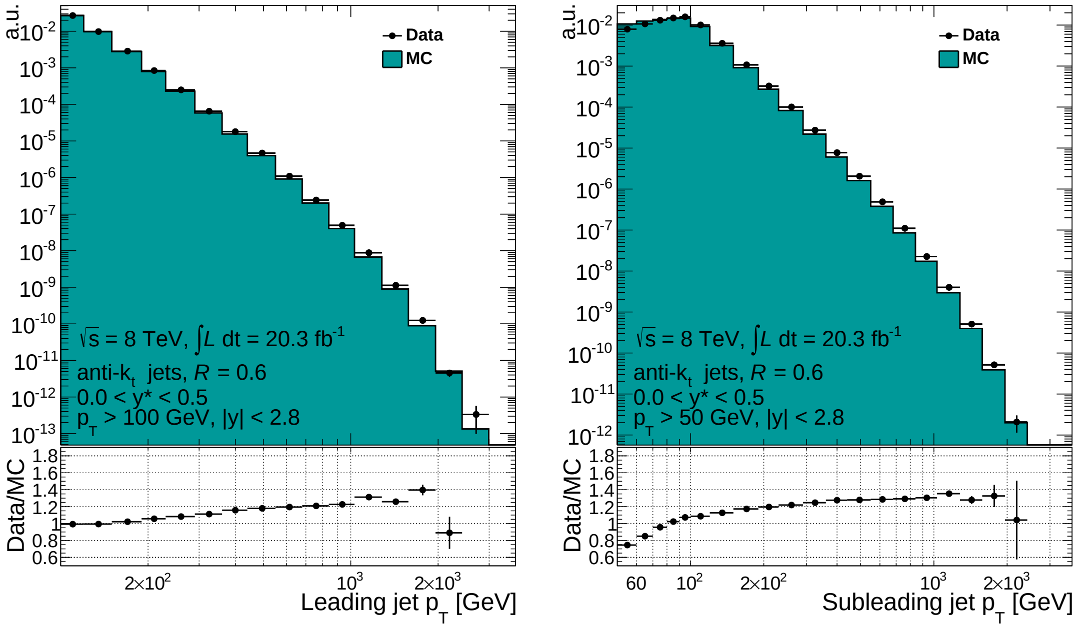


Figure 15.12: Comparison of the leading (left) and subleading (right) p_T distributions in Monte Carlo and data for $0 < y^* < 0.5$. The distributions are normalized to their integrals.

pass the p_T cut. At the highest y^* bin, the *forbidden zone* in the central region spreads from $2.8 - 4 = -1.2$ to $-2.8 + 4 = 1.2$ with a minimum rapidity separation of 4.

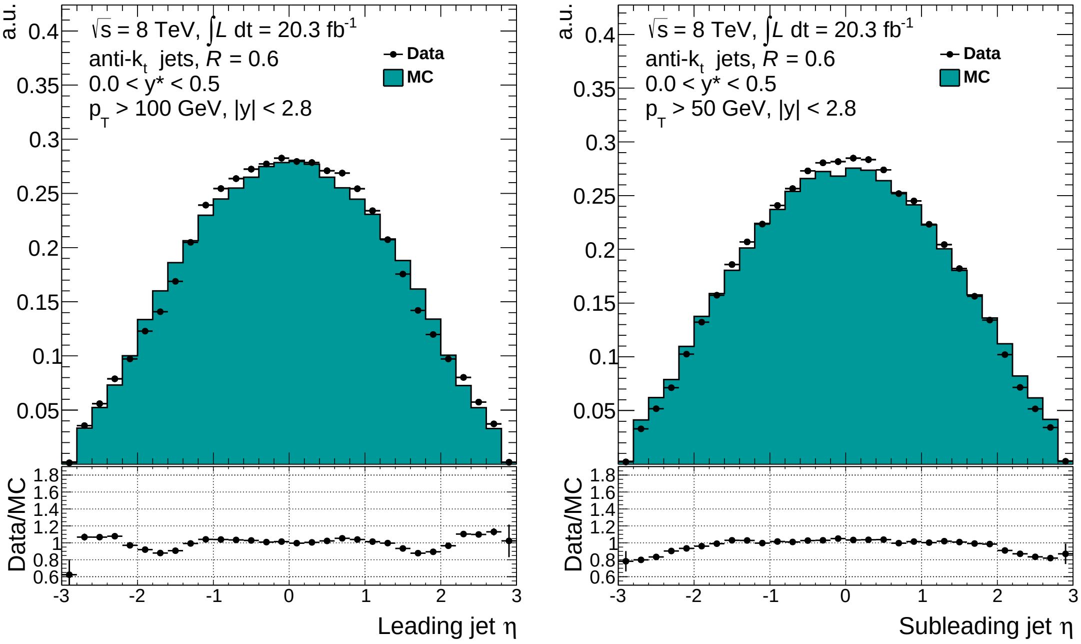


Figure 15.13: Comparison of the leading (left) and subleading (right) η distributions in Monte Carlo and data for $0 < y^* < 0.5$. The distributions are normalized to their integrals.

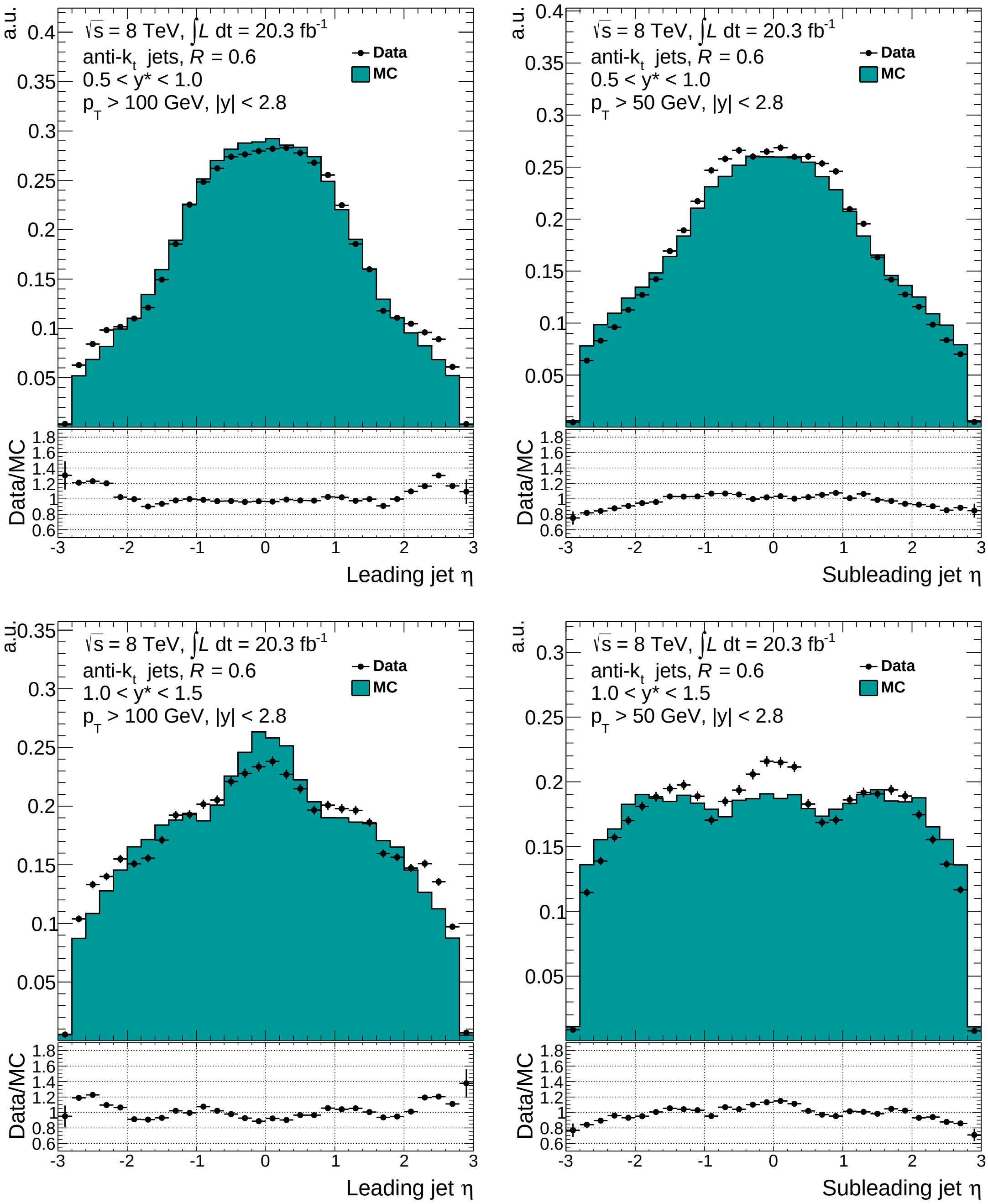


Figure 15.14: Comparison of the leading (left) and subleading (right) η distributions in Monte Carlo and data for $0.5 < y^* < 1.0$ (top) and $1.0 < y^* < 1.5$ (bottom). The distributions are normalized to their integrals.

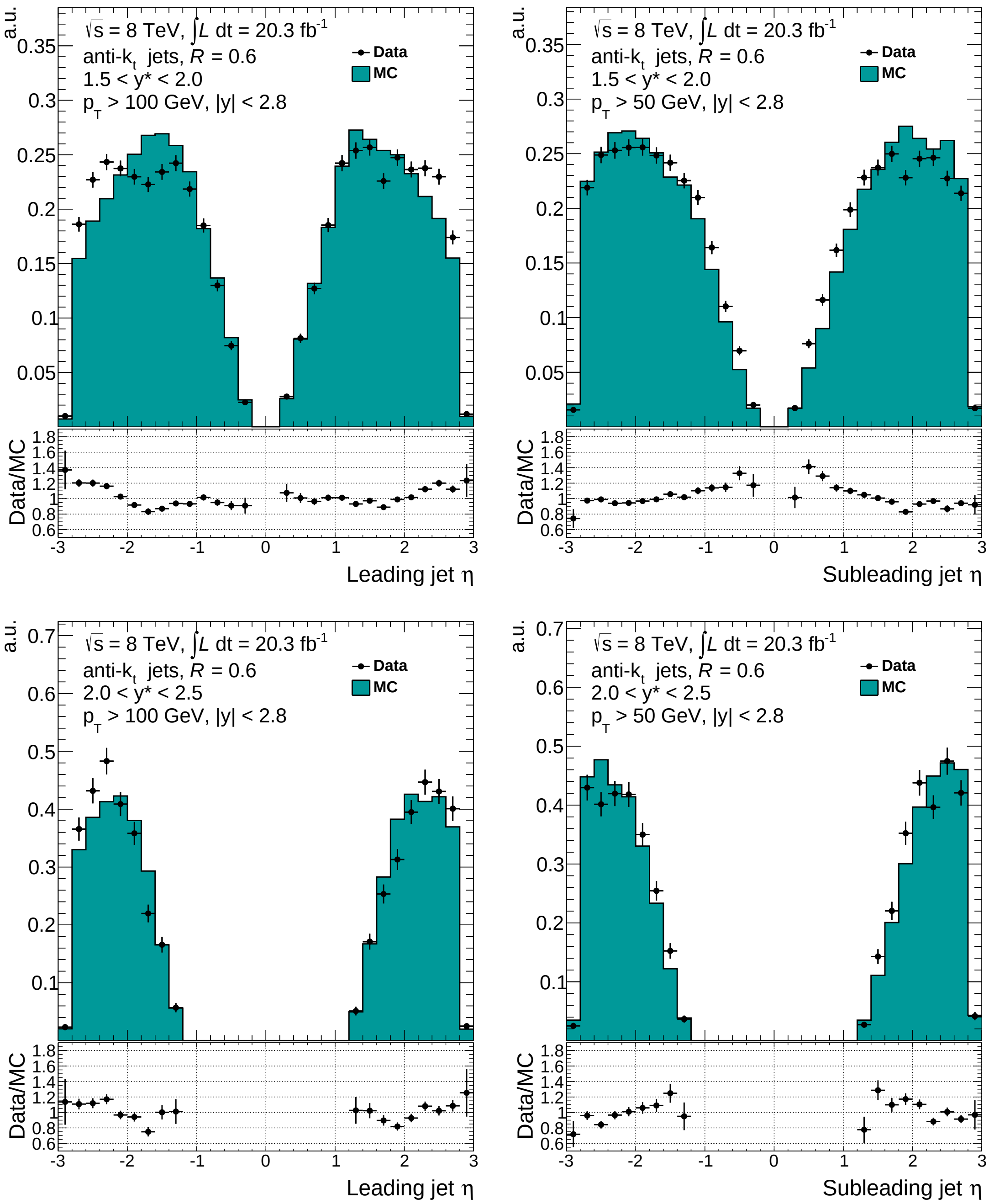


Figure 15.15: Comparison of the leading (left) and subleading (right) η distributions in Monte Carlo and data for $1.5 < y^* < 2.0$ (top) and $2.0 < y^* < 2.5$ (bottom). The distributions are normalized to their integrals.

In order to be able to compare the results to theory predictions or results from other experiments, the detector level distributions need to be corrected back to particle level. All information about detector effects like jet smearing used in the unfolding are derived from MC. Detailed information about bin migrations are drawn from so called transfer matrices, as shown in Fig. 15.16. They show the correlation between the dijet mass on reconstruction and particle level. Only *matched* events, i. e. events which pass all cuts on reconstruction and truth level, are considered. Bin migrations decrease with increasing dijet mass, as already shown in Fig. 15.8. A clear correlation between particle and reconstruction level is visible, which is a prerequisite for the correction to work. The bin migrations and correlations in y^* are much smaller compared to those in m_{12}. For that reason, the unfolding is performed in $y*$ bins and the fluctuations between different y^* bins is corrected for in an separated correction step. Unmatched events, i. e. events which are rejected on either truth or reconstruction level, are stored in separate histograms and used for efficiency corrections.

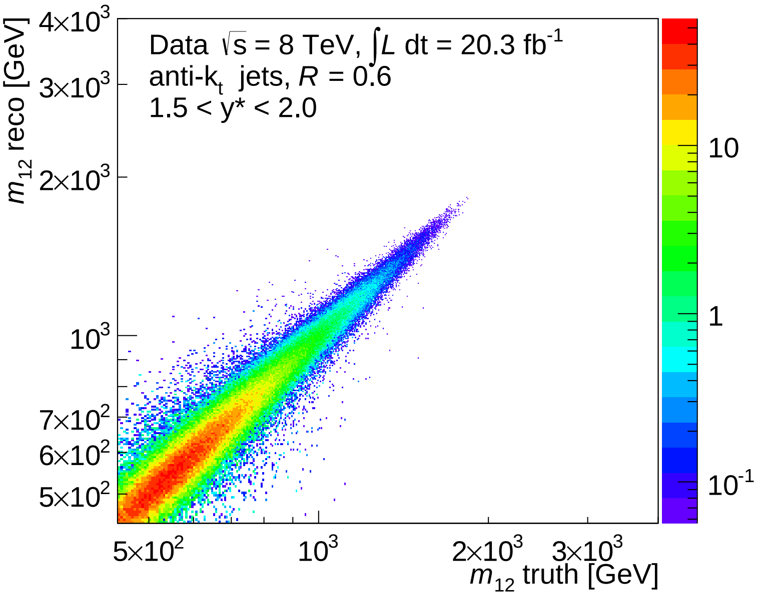


Figure 15.16: Transfer matrix for $1.0 < y^* < 1.5$ and anti-k_t jets with a radius parameter of $R = 0.6$. Entries at higher dijet masses are suppressed by the color scale.

The correction is performed in three steps. In a first step, an efficiency correction is applied to correct for fake dijets events, i. e. events where the dijet system does not pass the cuts on truth level, but on reconstruction level. The efficiency is calculated by dividing the dijet spectra of matched events on detector level by the spectrum of all events passing reconstruction level. The efficiency corrected data spectrum is called *matched level*.

This is followed by an iterative, dynamically stabilized unfolding method, described in [101], which is an extension to the D'Agostini unfolding, described in Sec. 12.4.3. The extension iteratively reweights the transfer matrix that the Monte Carlo truth spectrum matches the unfolded data. To ensure that statistical uncertainties are not amplified, the significance of the difference between the two spectra is used for the reweighting and not the nominal difference. After a fixed number of iterations, the modified transfer matrix is used to unfold the matched-level data, using the unfolding described in Sec. 12.4.3. This results in an improved description of the dijet mass spectrum in data by the Monte Carlo and reduces biases due to shape differences. The comparison between data and Monte Carlo in Sec. 15.4 shows that such a measure is needed. The number of iterations is a trade-off between systematic and statistical uncertainties. Each iteration reduces the shape bias, but enhances statistical uncertainties. One iteration was found to result in the lowest combined uncertainty.

In a final step, the matching efficiency is applied to the unfolded matched-level data distribution. The correction factors are derived correspondingly to the first correction step, but on particle level: The dijet spectra of matched events on particle level is divided by the dijet spectra of all events on particle level. An overview of all modification steps is shown in Fig. 15.17. The left plot shows the ratio of the acceptance corrected data (black), modified matched-level detector level Monte Carlo (red), non-matched true (blue) and reconstructed (green) Monte Carlo to the matched detector-level Monte Carlo for the lowest y^*-bin ($y^* < 0.5$). The agreement between the acceptance corrected data and modified matched-level Monte Carlo shows the quality of the Monte Carlo reweighting. The efficiency corrections represented by the blue and green line in Fig. 15.17 are very small and only have negligible impact. The right plot of Fig. 15.17 shows the equivalent correction factors for $y^* < 0.5$, which are below 5% over the full range. Same is true for all other y^*-bins. For all unfolding steps, one bin is added below

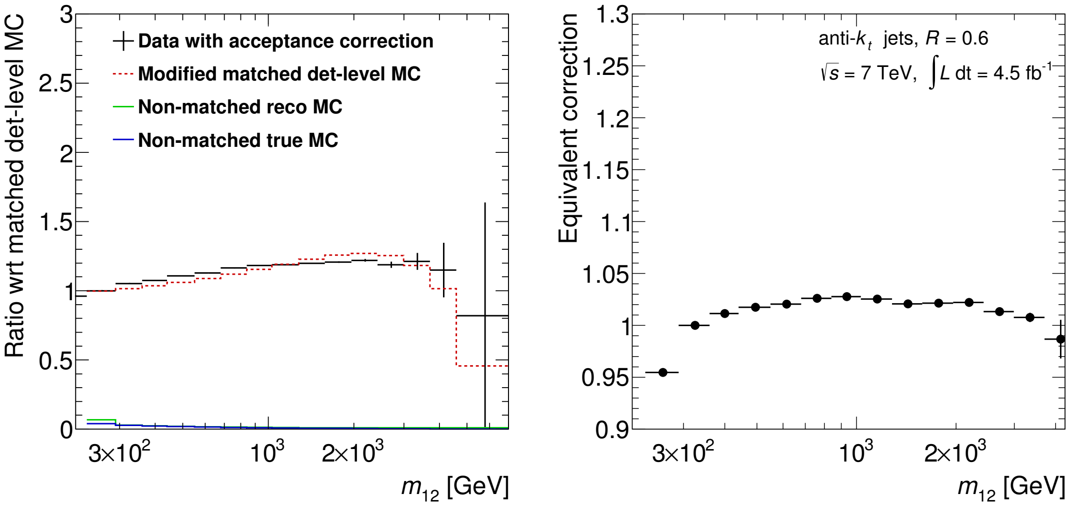


Figure 15.17: Ratio of acceptance corrected data (black), modified matched-level Monte Carlo (red), non-matched true (blue) and reco (green) Monte Carlo with matched detector-level Monte Carlo for $y^* < 0.5$ (left). The equivalent correction factors for $y^* < 0.5$ are shown on the right.

the lowest dijet mass bin of the measurement. If this would not have been done, bin migrations out of the lowest bin would be corrected, but migrations from lower dijet masses into that bin would not, which would result in an underestimation of the cross section in the first bin.

15.6 UNCERTAINTIES

This section will present the uncertainties on the presented measurement. In a first part, the evaluation of the statistical uncertainties are presented before describing the treatment of the systematic uncertainties.

15.6.1 *Statistical uncertainties*

The statistical uncertainties on data and Monte Carlo are derived using a bootstrapping method. To get the statistical uncertainty on data due to fluctuations, but also include bin correlations, a series of pseudo experiments are performed. Every event is fluctuated with a Poisson function with $\lambda = 1$, and the resulting fluctuated spectra are then unfolded. A covariance matrix is constructed by using the difference of each pseudo experiment to its average. The covariance matrix contains all information about the statistical uncertainties and the bin correlations introduced by the unfolding.
In order to estimate the systematic uncertainty on Monte Carlo, the Monte Carlo events are fluctuated in the same way as the data, producing a set of replicas of the transfer matrix. The data then gets unfolded with the fluctuated transfer matrices. A separate covariance matrix for Monte Carlo is constructed.
As the Monte Carlo and data uncertainties are uncorrelated, the two covariance matrices are added together and the square-root of the diagonal entries, the standard deviation, is used as the combined statistical uncertainty. The uncertainties for the $y^* < 0.5$ are shown in Fig. 15.18. The systematic uncertainty is dominated by data, and the Monte Carlo is about a factor of 10 below. The combined uncertainty is in the sub-percentage range below 2 TeV and rises above to 10% and more with increasing m_{12}. For higher y^* bins, the uncertainty on data is still dominating, but the uncertainty rises to 5% below 2 TeV and rises steeply to above 50% in the highest dijet mass bins.

15.6.2 *Systematic uncertainties*

All systematic uncertainties below are propagated through the unfolding procedure. For each component, all jet transverse momenta are shifted up and down by 1σ of the uncertainty. The shifted spectra as well as the unshifted spectrum are then unfolded using the nominal transfer matrix. The difference between the unfolded nominal value and shifted values are taken as the uncertainty on the measurement. As the components are considered to be uncorrelated, the quadratic sum of all deviations to the nominal result after unfolding is used as the combined systematic uncertainty. If the up and the down variation shift the spectra in the same direction, the bigger shift is taken as the uncertainty and the uncertainty on the opposite direction is set to zero. To reduce statistical fluctuations while deriving the systematic uncertainties on the dijet cross-section measurement, Monte Carlo is used. The result is independent of using data or Monte Carlo while Monte Carlo provides better statistics at high dijet masses.

Systematic uncertainty on unfolding

The systematic uncertainty on the unfolding process is derived by unfolding the modified reconstruction-level Monte Carlo and comparing it to the modified particle level distribution. To decrease the effect of statistical fluctuations, the transfer matrix is fluctuated 1000 times and the unfolding procedure is repeated for every replica. The average difference between the particle and unfolded detector-level Monte Carlo is then used as a systematic uncertainty, shown in the left plot of Fig. 15.18 for SVD, bin-by-bin and 0 and 1 iterations of IDS unfolding. The improved performance of the IDS unfolding after only one iteration, with >1% uncertainty over the full range, approves the choice of IDS unfolding over the other methods.

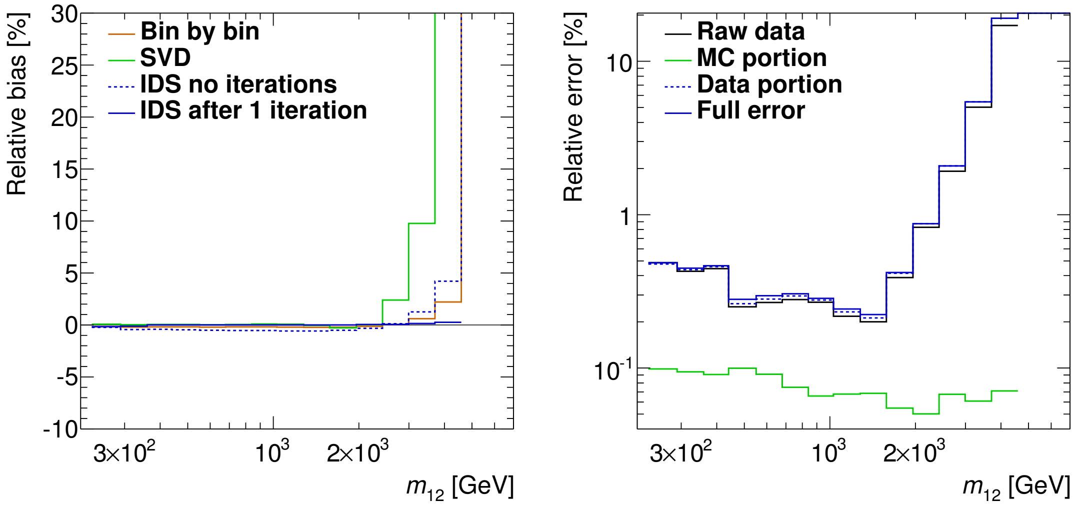


Figure 15.18: Relative bias for SVD, bin-by-bin and 0 and 1 iteration of IDS unfolding (left) and statistical uncertainty split into Monte Carlo (green) and data (black dotted) portions. The solid blue line shows the combined uncertainty. The solid black line shows the statistical uncertainty on data before unfolding. Both plots show the lowest y^*-bin ($y^* < 0.5$).

Jet energy scale uncertainties

Most of the jet energy scale uncertainties are described in Sec. 8.2. In addition to these, the statistical uncertainties of each calibration step are added as an extra component. The following uncertainties are also considered:

Flavour composition

The Monte Carlo used to derive the jet energy calibration has a fixed quark/gluon fraction (Ratio of jets induced by quarks/gluons). The γ and Z+jet calibration merely use jets from gluons. The differences in jet p_T responses between quark and gluon jets is evaluated in Monte Carlo. The uncertainty is defined as:

$$\Delta\mathcal{R}_s = \Delta f_g \times (\mathcal{R}_q - \mathcal{R}_g) \oplus f_g \times \Delta\mathcal{R}_g \qquad (15.2)$$

with the jet responses for quark and gluon initiated jets $\mathcal{R}_q$ and $\mathcal{R}_g \pm \Delta\mathcal{R}_g$, measured in Monte Carlo, the gluon jet fraction f_g and its uncertainty Δf_g. Fig. 15.19 shows the difference in the jet response for light quark and gluon induced jets. For this analysis, the default values of $f_g = 0.5$ and $\Delta f_g = 10\%$ are used. The resulting uncertainty is between 0.5% and 2%.

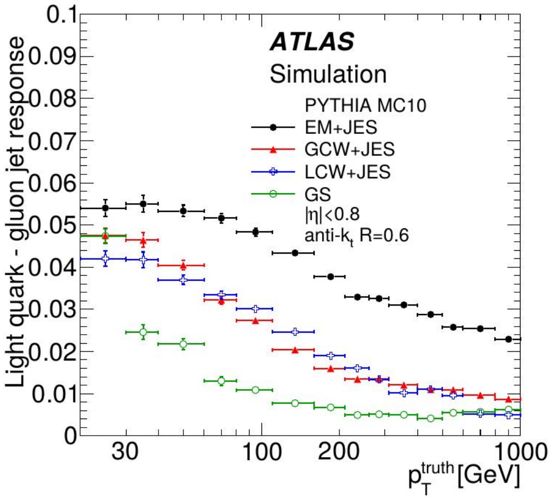


Figure 15.19: Difference in average response of gluon and light quark jets as a function of the truth jet p_T for anti-k_t jets with $R = 0.6$ in the central region. Various calibration schemes are shown [48].

Single particle response

The single particle response was derived first for the 2010 data and was updated with new Monte Carlo for the 2011/12 measurements [102]. The jet energy response is estimated from single particle responses. Jets are treated as a superposition of energy deposits of single particles. The energy responses were derived from test-beam measurements and minimum bias events in early 2010. For hadrons above 400 GeV, a flat uncertainty of 10% is assumed, to account for possible detector non-linearities and leakage. For jets above 1 TeV, the in-situ uncertainty components are assumed constant, using the value at 1 TeV because they cannot be reliable evaluated at the highest jet transverse momenta. The single particle response is used to derive the uncertainties above $p_T = 1$ TeV. To avoid overestimating the uncertainty, the other uncertainty components are subtracted in quadrature from the single particle response uncertainty.

Jet energy resolution

The jet energy resolution (JER) reflects the mis-modeled detector resolution in Monte Carlo simulations. As the energy resolution influences the bin migrations and therefore the unfolding, the mis-modeled resolution might have a large effect on the resulting cross section. The jet resolution is measured in data and Monte Carlo using two different in-situ methods. The first one is a dijet balance technique, in which the transverse momenta of the two leading jets are balanced, in order to derive a jet energy resolution. The second one is the bisector method using the balance of the projection of the two leading jets on an orthogonal coordination system in the transverse plane which bisects the azimuthal angle between the two leading jets as shown in Fig. 15.20 and discussed in [103].

Uncertainties on the JER are derived from different sources. To estimate the uncertainty due to soft radiation, the cuts to suppress those, such as the angular separation in ϕ of the two leading jets and a cut on the third jet transverse momentum, are varied. The *true*

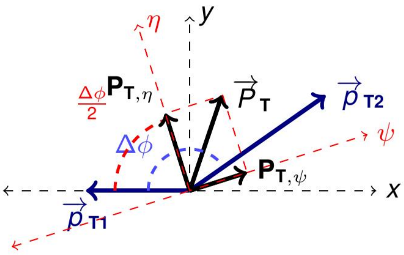


Figure 15.20: Coordination system used for the bisection method. $p_{T,1}$ and $p_{T,2}$ are the two leading jets and $\vec{P}_T$ is the imbalance transverse momentum vector [103].

resolution was derived with various Monte Carlo generators and tunes to predict the uncertainty due to the mis-modeling of the Monte Carlo simulation. The discrepancy between the two methods as well as between data and Monte Carlo is another source of uncertainty. The uncertainties on the fits used to compute the resolution were taken into account as a final contribution to the uncertainty. To propagate the uncertainty to the dijet analysis, each jet is fluctuated with the quadratic difference between the truth resolution and the truth resolution plus the uncertainty σ_{smear}:

$$\sigma_{\text{smear}} = \sqrt{(\sigma_{\text{nominal}} + \Delta\sigma)^2 - \sigma^2_{\text{nominal}}} \tag{15.3}$$

where σ_{nominal} is the jet energy resolution and $\Delta\sigma$ its uncertainty. The data is unfolded with the smeared unfolding matrix. The difference between the smeared result and the nominal result is taken as the jet energy resolution uncertainty on the dijet measurement.

Jet angular resolution

In analogy to the jet energy resolution, a jet angular resolution (JAR) can be assigned to the measurement. The resolution is derived from matching jets on reconstruction level to particle level in Monte Carlo. In previous measurements, the uncertainty on the dijet cross section was found to be below 0.5% with some exceptions at very high dijet masses (up to 3%). Due to the very small magnitude of the uncertainty compared to the JER and JES uncertainties, the uncertainty is neglected in this analysis.

The combined uncertaintiy for the lowest y^* bin is shown in Fig. 15.21 and in Fig. D.1 for the other four y^* bins. The single components are discussed in detail in the following:

Special The *special* component combines the uncertainties on the pile-up subtraction, flavour composition and single particle response. It is the biggest source of uncertainties over a broad dijet mass range. At the lowest dijet masses, the uncertainty is 10%, decreasing to a few percent at around 3 TeV and rapidly increasing to 50% thereafter. At low energies, the pile-up and flavour response uncertainties are dominating, while the single particle response dominates at higher masses.

- The uncertainty on the pile-up calibration is approximately constant over p_T. However, the lower the jet energy, the higher the relative uncertainty.

- The difference in the energy response of quark and gluon induced jets increases with decreasing jet energies (Fig. 15.19). Therefore, the flavour composition uncertainty increases as well.
- Due to a lack of measurements of the single particle responses at high p_T, a realistic, yet conservative uncertainty is applied and dominates for high p_T jets.

MJB The multi jet balance MJB is not applied to low p_T jets, so the uncertainty at low dijet masses is close to zero. It rises up to $\sim$ 15% at higher masses, at which the method is mainly applied. To balance a high p_T jet, the multi jet balance needs to be applied iteratively. With each iteration, the systematic uncertainties of all previous iterations on the recoil system jets are propagated to the MJB.

γ+Gjet The uncertainty on the γ-jet direct balance is dominated by the propagation of the photon energy scale uncertainty, which increases itself with increasing photon transverse momentum. At a dijet mass of around 200 GeV, it is the dominating uncertainty with 6%-7%. It is even increasing further to 16% at the highest dijet masses, only exceeded by the MJB and single particle response.

Z+jet, JER and η-intercalibration The other uncertainties are usually below 5%. The JER increases from close to no contribution at low masses up to 3% at 6 TeV.

The combined uncertainty is between 7% and 12% for dijet masses up to 2 TeV, increasing to 20% in the highest dijet mass bin populated with data (second last bin). For higher y^* values, the combined uncertainty increases. The combined pile-up, flavour response and single particle response uncertainty is dominating even more, the higher the y^*. Also, the relative trends between the single contributions are similar to the lowest y^* bin. The limited statistics to evaluate the uncertainties are clearly visible as statistical fluctuations in the higher y^* bins.

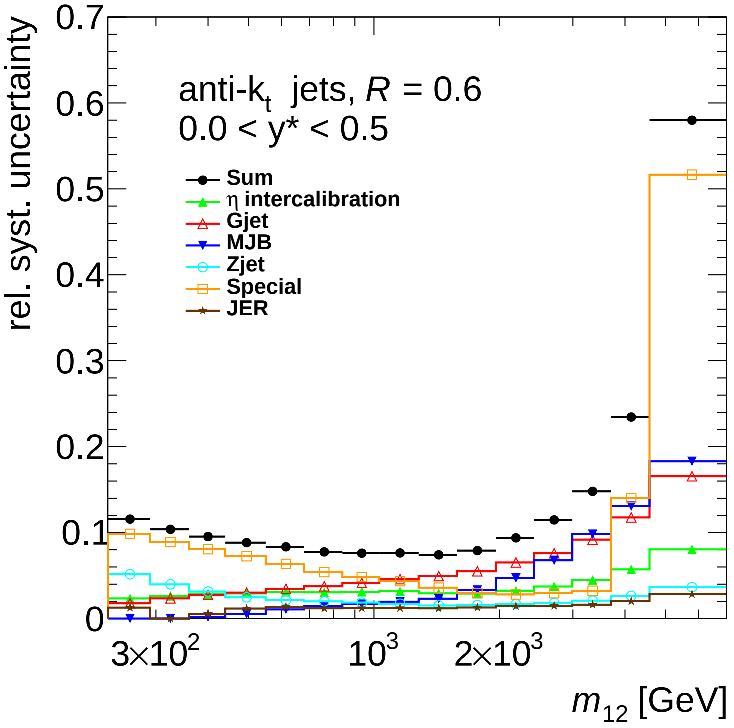


Figure 15.21: Relative systematic uncertainties on the dijet mass propagated through the unfolding. Uncertainties from the single jet energy scale correction methods are combined. Special combines the remaining uncertainties on the pile-up subtraction, flavour composition and single particle response. Black points represents the combined uncertainties.

16

RESULTS

In this chapter, the results of the double differential dijet cross section measurement is presented and compared to several NLO predictions. The unfolded cross sections on particle level with the statistical and systematic uncertainties are compared to NLO predictions produced with NLOjet++ folded with four different PDF sets. The NLO predictions are corrected for non-perturbative effects.

16.1 DOUBLE DIFFERENTIAL DIJET CROSS SECTION RESULTS

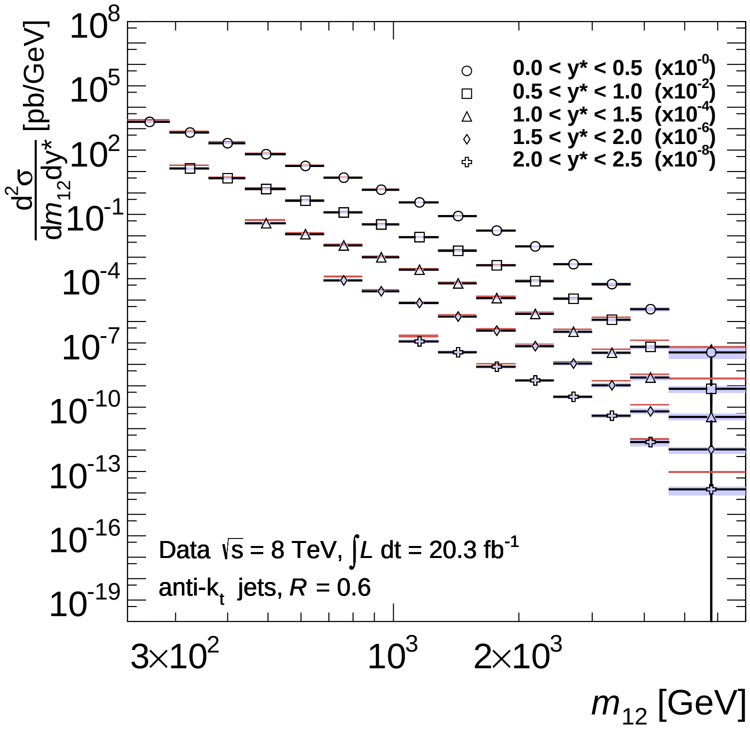


Figure 16.1: Double differential dijet cross section for anti-k_t jets with a radius parameter of $R = 0.6$ as a function of dijet mass and y^*. For a better readability, y^*-bins are scaled as described in the legend. Error bars represent statistical uncertainties while colored areas indicate the combined systematic uncertainties. For comparison, an NLOjet++ prediction, using the MSTW 2008 PDF set, corrected for non-perturbative effects, is shown in red.

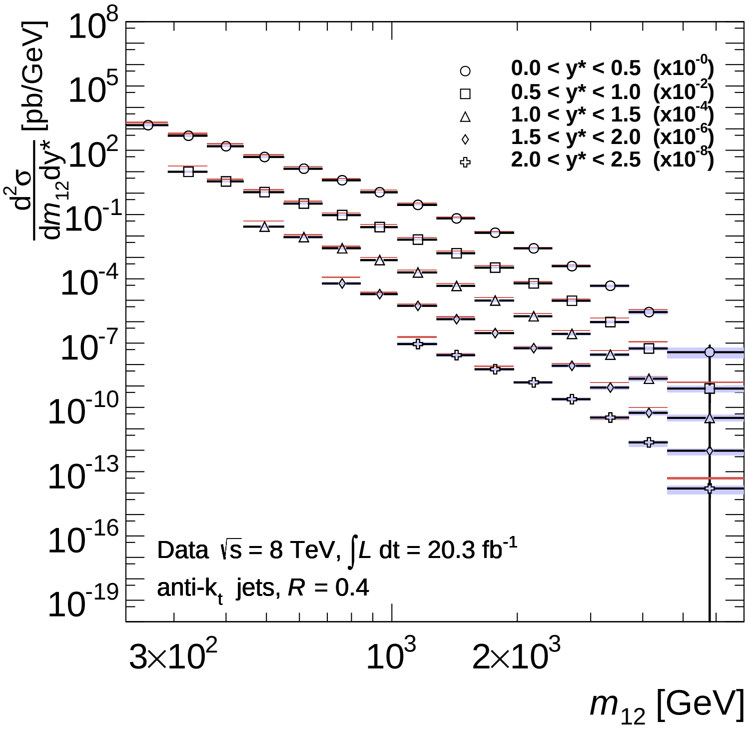


Figure 16.2: Double differential dijet cross section for anti-k_t jets with a radius parameter of $R = 0.4$ as a function of dijet mass and y^*. For a better readability, y^*-bins are scaled as described in the legend. Error bars represent statistical uncertainties while colored areas indicate the combined systematic uncertainties. For comparison, an NLOjet++ prediction, using the MSTW PDF set, corrected for non-perturbative effects, is shown in red.

Fig. 16.1 and Fig. 16.2 show the double differential cross section for anti-k_t jets with a radius parameter of $R = 0.6$ and $R = 0.4$ respectively as a function of dijet mass m_{12} and y^*. The data is compared to an NLO prediction, produced by NLOjet++, using the MSTW 2008 PDF set. The NLO prediction is corrected for non-perturbative effects. The systematic uncertainty on data is the quadratic sum of all uncertainty components discussed in Sec. 15.6. For the NLOjet++ calculation, the uncertainties on scale choices, α_S, PDF sets and non-perturbative corrections are added in quadrature for the final, combined uncertainty. The theory prediction is able to describe the general behaviour of the data over many orders of magnitude. The dijet masses reach values of above 6 TeV. However, discrepancies at these high dijet masses can be observed.

16.2 COMPARISON WITH NLOJET++ PREDICTIONS

For a better comparison, the ratio of theory over data for different PDF sets are shown in Fig. 16.3 - Fig. 16.6. NLOjet++ agrees with the measurement within the uncertainties up to a dijet mass of 3 TeV for all PDF sets and all y^* bins. Beyond that, single bins show a disagreement and NLOjet++ predicts a higher cross section. The higher the dijet mass, the bigger the discrepancy. Even though the theory predictions using the NNPDF 2.3 and MSTW2008 PDF sets agree within the uncertainties, they system-

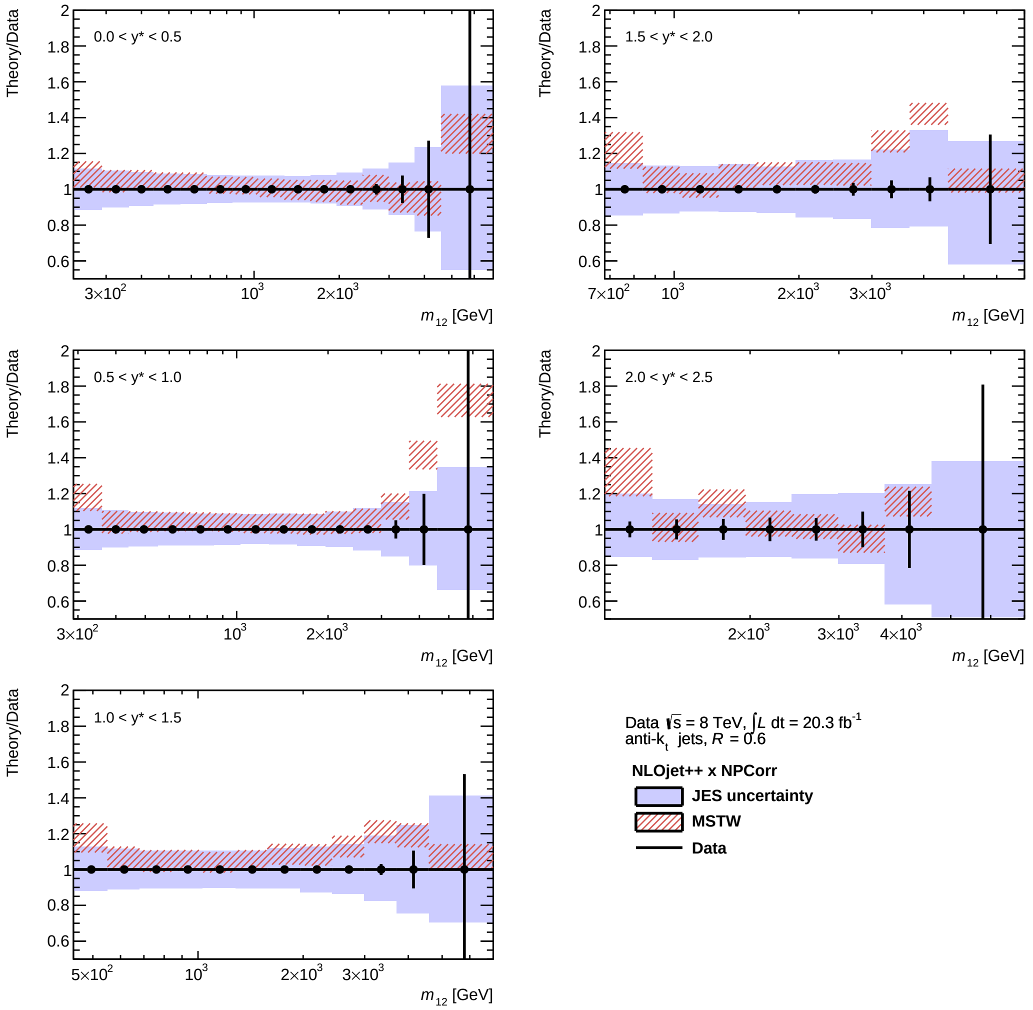


Figure 16.3: Ratio of the NLOjet++ theory prediction, using the MSTW2008 PDF set and corrected for non-perturbative corrections, and the measurement of the double differential cross section for anti-k_t jets with a radius parameter of $R = 0.6$. Error bars represent statistical uncertainties of the data. The combined systematic uncertainty is indicated by the blue area. Systematic uncertainties on the NLOjet++ predictions are drawn as red shaded area.

atically predict a higher cross section than measured. Using HERAPDF results in the opposite observation. Below 2 TeV, the measured cross section is above the predicted one. From 2 TeV on, this changes and the NLOjet++ predictions are above the data. The CT10 prediction provides a fairly accurate prediction at lower dijet masses, but the trend of the prediction to overshoot at high m_{12} is also present. An up-fluctuation in the first dijet mass bin can be observed in all comparisons. This might be caused by not-considered bin migrations into the first bin of the measurement during the unfolding. The combined systematic uncertainties on the measured data are at around 10% and rise up to 50% and above for the highest dijet mass bins with a minimum at around 1 TeV of slightly below 10% for the lowest y^* bin. The systematic uncertainties on NLOjet++ paired with CT10 and HERAPDF 1.5 are about the same magnitude as the systematic uncertainty on data. Using MSTW2008 and NNPDF 2.3 results in systematic uncertainties roughly half the magnitude as for CT10 and HERAPDF. At high

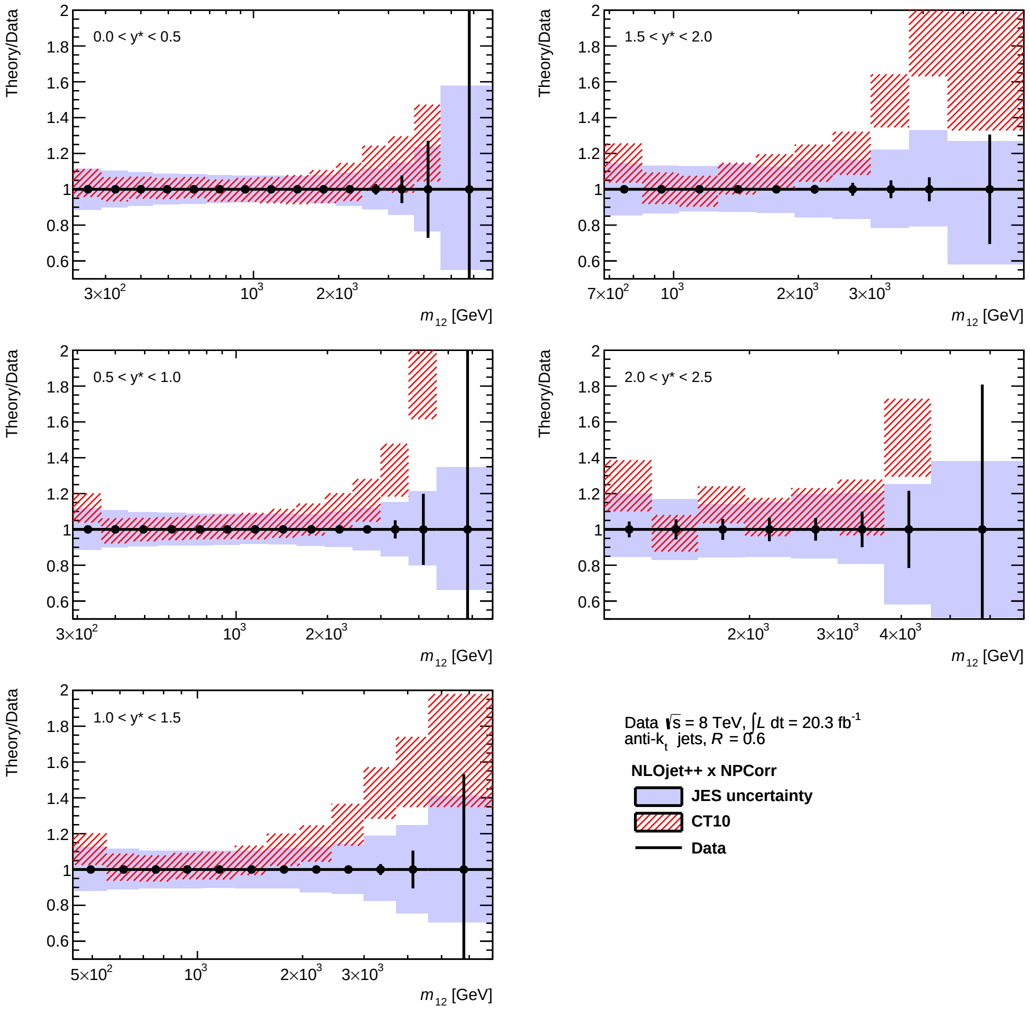


Figure 16.4: Ratio of the NLOjet++ theory prediction, using the CT10 PDF set and corrected for non-perturbative effects, and the measurement of the double differential cross section for anti-k_t jets with a radius parameter of $R = 0.6$. Error bars represent statistical uncertainties of the data. The combined systematic uncertainty is indicated by the blue area. Systematic uncertainties on the NLOjet++ predictions are drawn as red shaded area.

y^* values, the NLOjet++ prediction is limited by the statistical uncertainties, which leads to increased systematic uncertainties and fluctuating bins. The same is true for higher dijet mass bins. For the two y^*-bins below 1.0, the discrepancy increases with increasing dijet mass. For $1.0 < y^* < 1.5$, however, the maximal discrepancy is reached at approximately 4 TeV, before decreasing again. The ratio plots for the second radius parameter $R = 0.4$ are shown in Appendix E. They yield the same conclusions and share the same behaviour as the results for the bigger radius parameter $R = 0.6$. However, it can be observed that the theory predictions tend to overestimate the cross section even more can be observed.

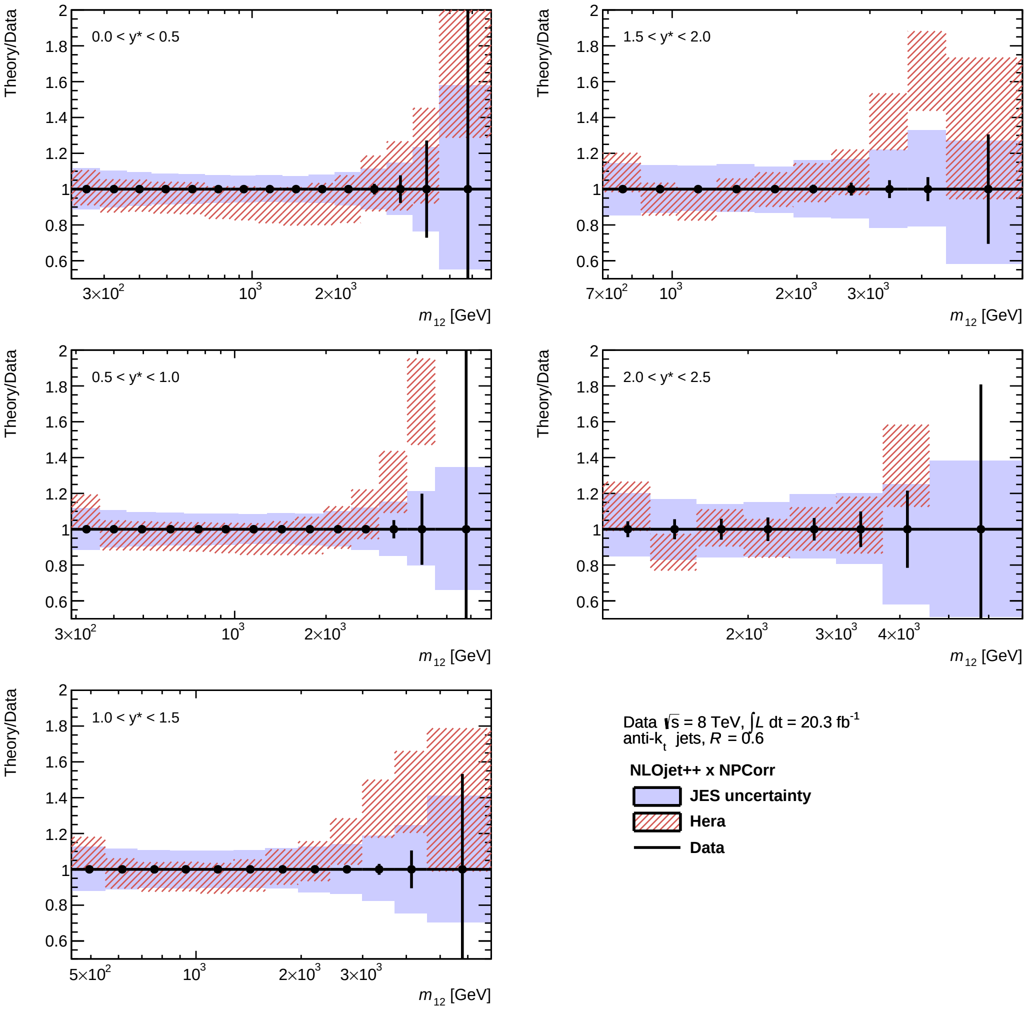


Figure 16.5: Ratio of the NLOjet++ theory prediction, using the HERAPDF 1.5 PDF set and corrected for non-perturbative corrections, and the measurement of the double differential cross section for anti-k_t jets with a radius parameter of $R = 0.6$. Error bars represent statistical uncertainties of the data. The combined systematic uncertainty is indicated by the blue area. Systematic uncertainties on the NLOjet++ predictions are drawn as red shaded area.

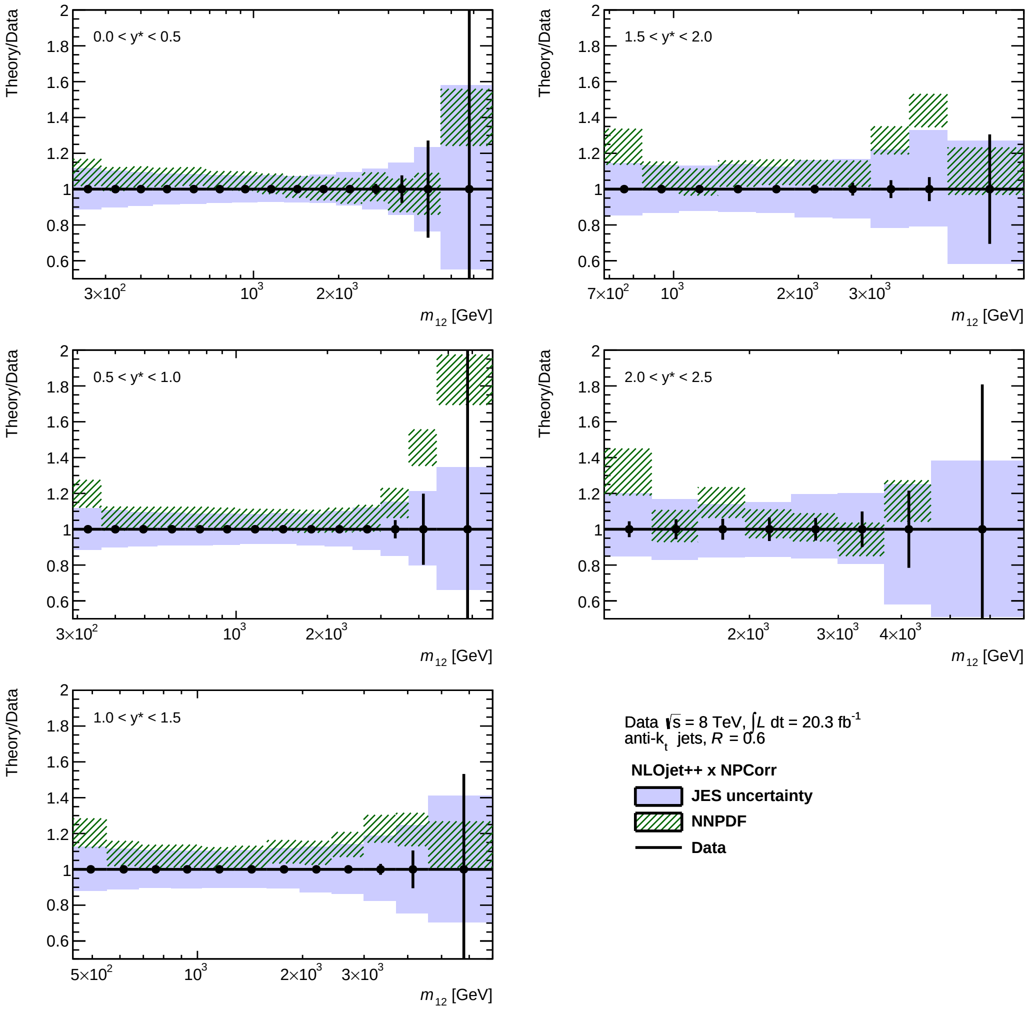


Figure 16.6: Ratio of the NLOjet++ theory prediction, using the NNPDF 2.3 PDF set and corrected for non-perturbative corrections, and the measurement of the double differential cross section for anti-k_t jets with a radius parameter of $R = 0.6$. Error bars represent statistical uncertainties of the data. The combined systematic uncertainty is indicated by the blue area. Systematic uncertainties on the NLOjet++ predictions are drawn as green shaded area.

Ongoing studies have shown that corrections applied to compensate dead modules in the calorimeter, so called bad channel, *BCH*, corrections, can lead to an overcalibration of jets. This mainly effects jets with a very high transverse momentum. The more collimated the jet is, the bigger the possible overcalibration. To estimate the energy deposited in a nonfunctional module, the mean value of the two neighboring cells are used, as indicated in Fig. 16.7. The smaller the jet energy, the broader the jet. High-p_T-jets tend to be more collimated and the possibility of a mis-modeling is increased. Four modules in the tile calorimeter have been broken over larger parts of the 2012 data taking and only one of them is modeled in the available Monte Carlo. Additional random power trips lead to short-time dysfunctional modules, which are not modeled as well. Therefore, a Monte Carlo based correction for the mis-calibration of the jets was not possible at the time of writing this thesis. Removing all events in which at least one jet is close to any dead module, called *tight* BCH cleaning, results in a loss of about 1/3 of all data. The position of all removed jets for the tight cleaning is shown in Fig. 16.8. The four modules are clearly visible as red spots. Modules effected by short dead times are randomly distributed in the tile calorimeter. 36% of all events would have been removed with this cleaning. Fig. 16.9 compares the uncleaned dijet mass spectrum for the lowest y^* bin with the dijet mass spectrum after removing all events, which are potentially affected by the mis-calibration. The shape difference between the two spec-

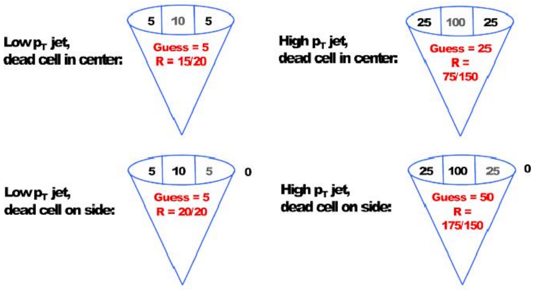


Figure 16.7: Schematic view of bad channel corrections [104].

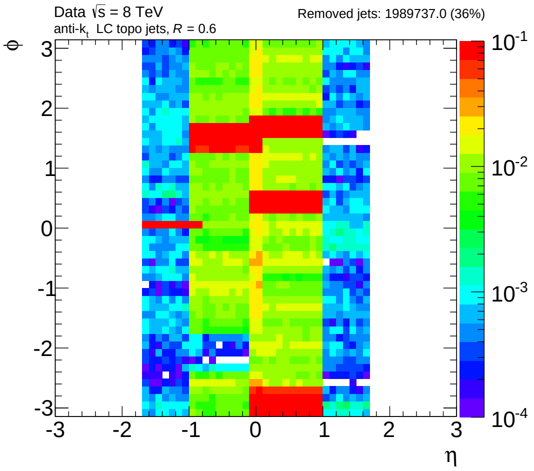


Figure 16.8: Leading jets masked as possible mis-calibrated due to BCH corrections.

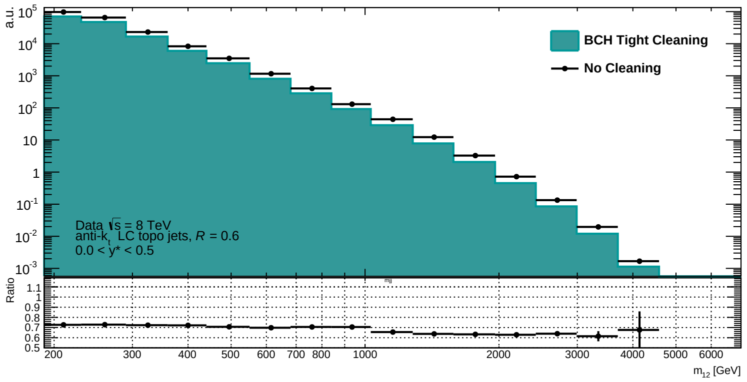


Figure 16.9: Comparison of the dijet mass spectrum on detector level before and after a tight BCH cleaning for two y^* bins, using anti-k_t jets with a radius parameter of $R = 0.6$.

tra shows that the effect depends on the dijet mass. First tests predict that the cross section is overestimated at high dijet masses. Fig. 16.10 shows the ratios of the fully unfolded cross sections with and without BCH cleaning applied to the Monte Carlo detector level and data, using a radius parameter of $R = 0.6$. In both y^* bins shown, a lower cross section after the cleaning is observed. The two y^* bins show different trends. For $y^* < 0.5$ the ratio rises up to 20% while for $1.0 < y^* < 1.5$, the effect is smaller and decreases with increasing dijet masses. This could be caused by the BCH overcorrection, but the capability of the Monte Carlo to correct for the over calibration cannot be proven. Therefore, the true impact on the unfolded cross section cannot be predicted. The studies on this issue were ongoing while this thesis was written and no

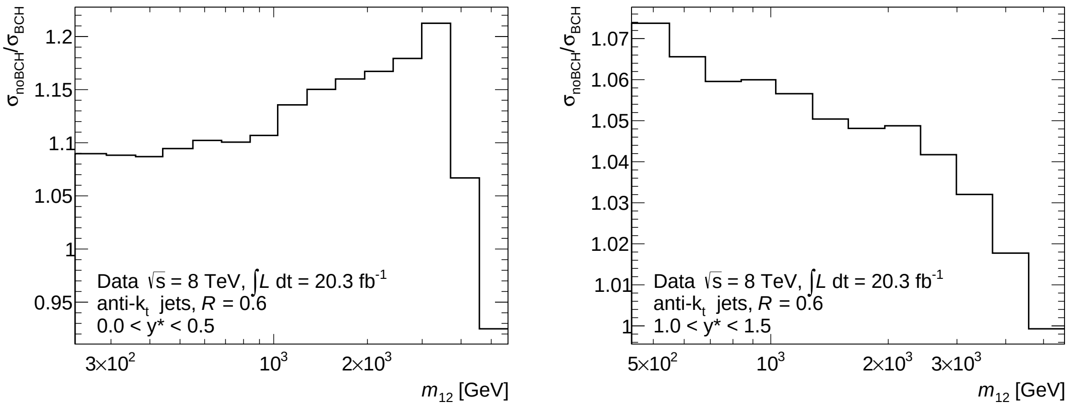


Figure 16.10: Ratio of the nominal dijet cross section (σ_{noBCH}) and a BCH corrected cross section (σ_{BCH}) for anti-k_t jets with a radius parameter of $R = 0.6$ for two different y^* bins.

suitable solution could be found. A well motivated estimation of an uncertainty due to the BCH correction cannot be derived without additional Monte Carlo studies.

16.4 OUTLOOK

To improve the presented measurement, a solid solution for the BCH corrections explained above is needed. The impact is most probably of non-negligible size. In addition, the final jet energy scale had not been released when writing the thesis. The origin correction was not be applied and the flavour response uncertainties should decrease due to additional Monte Carlo studies. The impact of the improvements of the JES are expected to be small and should not change the outcome of the measurement to a noticeable extent. The analysis is prepared to include the mentioned changes as soon as they are available.

Calculations of electroweak corrections [105], including electroweak loops to the QCD NLO prediction provided by NLOjet++ have not been finished before writing this thesis. In general, they are predicted to be negligible for larger values of y^*, but increasing the predicted cross section even further at low y^* and high dijet mass values. Correction factors for dijet production at $\sqrt{s} = 7$ TeV are shown in [100].

An extension to the forward region, loosen the $|y|$ and y^* cuts, would give access to a broader range of Bjorken-x, as shown in Fig. 14.2. In a simplified picture, imagine two solid scattering partners with impulse x, two partons with different x values are needed to boost the dijet system into the forward region. PDF extractions can profit from these measurements if the uncertainties are not too large. The analysis presented was designed to be easily expanded to the forward region. The main reasons for restricting the measurement to the central region are the transition region between the central and forward jet trigger (as shown in Fig. 15.6) and the more complicated jet energy calibration. Correction factors to compensate the less effective jet triggers need to be derived. The coarser granularity and worse energy resolution① of the forward calorimeter, as well as the increased pile-up and underlying event activity, combined with the missing tracking information, makes it difficult to measure jets in the forward region. Reduced statistics due to the minimum p_T cut impede the measurement even further.

Another way to improve the sensitivity to parton distribution functions is to add a third dimension and measure the triple-differential cross section. A possible choice as a third variable is $y^{\text{boost}} = |y_1 + y2|/2$, together with the dijet mass and y^*. y^{boost} is a measure of the boost of the dijet system. The separation in y^{boost} helps to select events with high or very low Bjorken-x, similar to the expansion to the forward region. However, splitting the data in even more bins reduces the statistical power of every single bin. First distributions on detector level were measured to test the basic performance and the statistical uncertainties. Fig. 16.11 shows the lowest y^{boost} bin of the uncorrected triple-differential cross section on detector level for jets reconstructed using the anti-k_t jet algorithm with a radius parameter of $R = 0.6$. The expected falling spectrum is observed. The statistical uncertainties on detector level are still small enough for a sensible measurement. Fig. 16.12 shows a higher bin in y^{boost}, covering $2.0 < y^{\text{boost}} < 2.5$. Dijet events cannot have high values for y^{boost} and y^* at the same time (a high separation in rapidity always results in a small y^{boost} value). The statistical uncertainties are still tolerable. The analysis framework is set up to perform the analysis steps presented for the double-differential cross section also for the triple differential cross section. All uncertainty variations were already performed.

To improve the impact on PDF extractions even further, the combination of different jet measurements are aimed at. Therefore, the different analyses endeavor to agree on as many similarities as possible, such as used jet radii and calibration methods. Consider-

---

① $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$ compared to $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$ in the barrel and end-cap of the hadronic calorimeter, with E being in units of GeV.

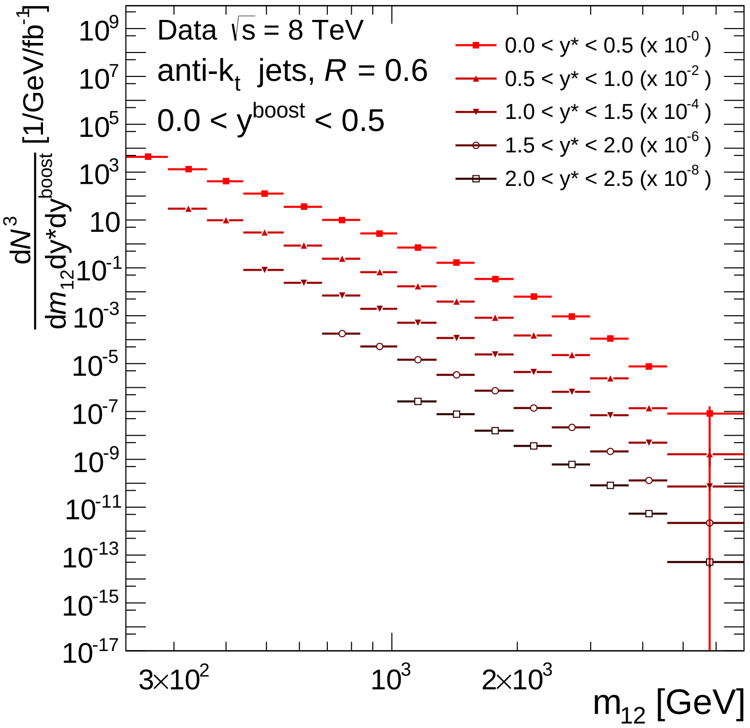


Figure 16.11: Triple-differential cross section as a function of the dijet mass, y^* and y^{boost}. The lowest y^{boost} bin covering $0.0 < y^{boost} < 0.5$ is shown. The single y^* slices are scaled to increase readability.

ing the uncertainty correlations among the different analyses reduces the propagated uncertainties to the PDFs. Candidates for jet analyses to be combined with the dijet cross section measurement are the three-jet cross section and the inclusive jet cross section. The uncorrected inclusive jet cross section on detector level, using jets reconstructed with the anti-k_t jet algorithm and a radius parameter of $R = 0.6$ is shown in Fig. 16.13. Jets with a transverse momentum of up to 3 TeV are observed. All shown jet distributions, however, are impacted by the impairment of the BCH correction. The fundament for future analyses of jet cross sections at $\sqrt{s} = 8$ TeV with the ATLAS detector is established by the performed analysis and the developed analysis framework. Looking even further ahead, new jet cross section measurements during Run II of the LHC at increased center-of-mass energies increase the accessible phase space even further. The observed discrepancies at higher dijet masses can be analyzed with increased statistical power. The increased integrated luminosity, which will presumably be recorded during Run II, can also reduce the uncertainties on the in situ jet calibration.

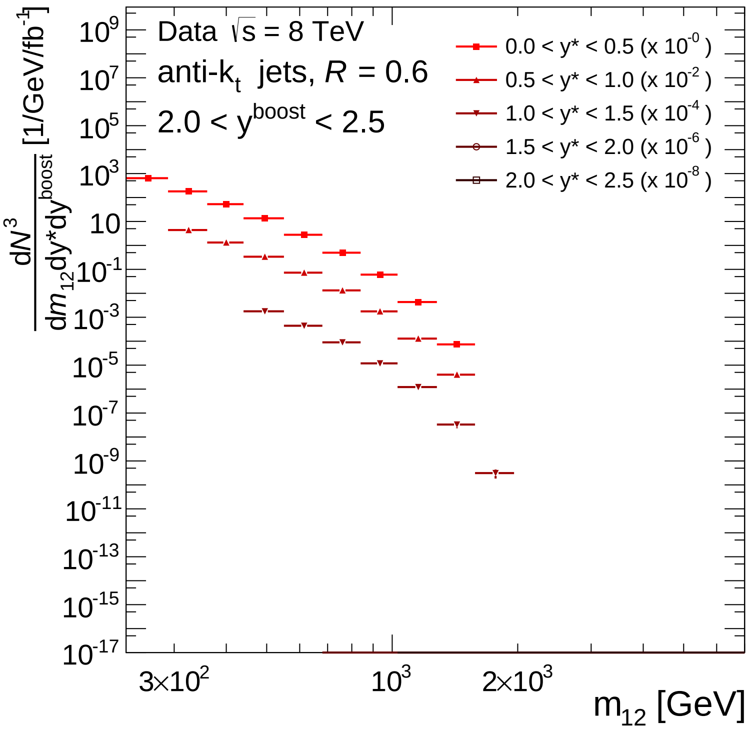


Figure 16.12: Triple-differential cross section as a function of the dijet mass, y^* and y^{boost}. The y^{boost} bin covering $2.0 < y^{\text{boost}} < 2.5$ is shown. The single y^* slices are scaled to increase readability.

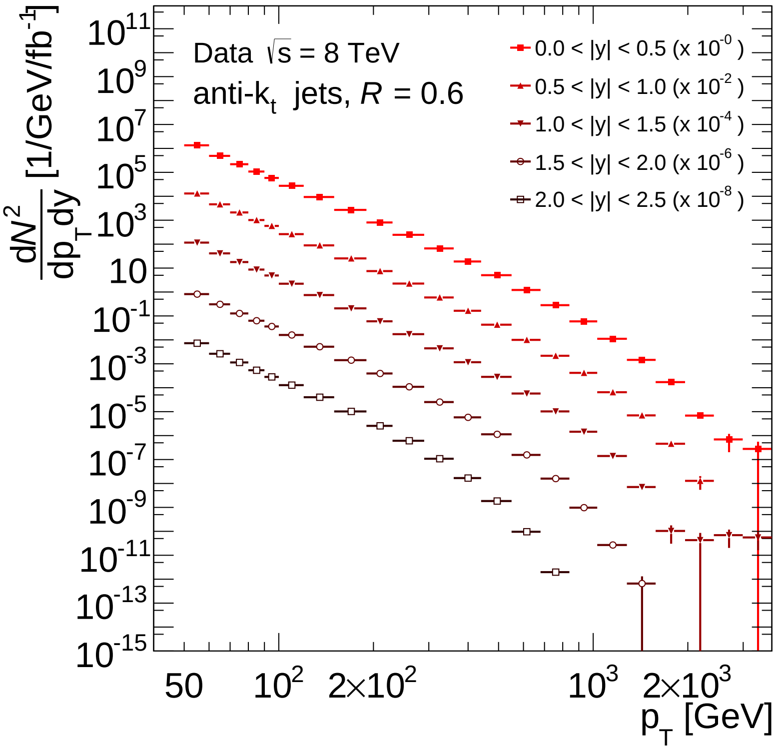


Figure 16.13: Inclusive jet cross section as a function of the jet transverse momentum and absolute values of rapidity on detector level. Jets reconstructed with the anti-k_t jet algorithm with a radius parameter of $R = 0.6$ are used. Single $|y|$ slices are scaled to increase the readability.

Part VI

CONCLUSION AND OUTLOOK

17

CONCLUSION AND OUTLOOK

Two measurements using data recorded by the ATLAS experiment have been presented in this thesis. Both are capable to test QCD predictions, however at different momentum scales. Proton–proton collisions with a small momentum transfer are described by phenomenological models since perturbative calculations are not applicable at low momentum scales. Measuring properties of these *soft* collisions can improve the phenomenological models and the understanding of this non-perturbative regime of QCD. The high center-of-mass energy at the LHC allows tests of perturbative theory predictions at momentum scales never reached before. The dijet cross section is able to test theory predictions over a broad range of the accessible physe space of the hard scattering.

The measurement of charged-particle event shape variables at $\sqrt{s} = 7$ TeV in inelastic proton–proton collisions uses charged-particle tracks to measure the momentum flow transverse to the beam axis. The transverse thrust, the thrust minor and the transverse sphericity are sensitive to the so called underlying event, and can be used to test phenomenological models used in Monte Carlo generators to simulate soft processes with a low momentum transfer. Due to the sensitivity to the underlying event and other soft contributions, only data taken in the beginning of 2010 was used, as the LHC provided a low instantaneous luminosity to ATLAS, resulting in a more or less pile-up free data taking. The collected data corresponds to an integrated luminosity of 168 μb. The data was cleaned from the remaining pile-up events by rejecting events with more than one reconstructed primary vertex. To suppress background from single and double diffractive scattering, only events with at least six tracks were used. The minimum track p_T was chosen to be 500 MeV. The event shape distributions were analyzed in bins of the transverse momentum of the leading track, p_T^{lead}. The ATLAS tracking system provides tracking information down to 100 MeV, but the track reconstruction efficiency significantly drops below 500 MeV. The measured data was corrected for detector effects and compared to five different Monte Carlo simulations: PYTHIA 6 AMBT2B, PYTHIA 6 DW, PYTHIA 6 Z1, PYTHIA 8 A2 and Herwig++ UE7-2. The discrepancy between unfolded results, using different Monte Carlo simulations for the unfolding is the dominant systematic uncertainty of the measurement. The uncertainty is between 1% to 10%, with the exception of a few bins with uncertainties up to 40%. Data predicts the events to be mainly spherical, with a shift towards more dijet-like events with increasing p_T^{lead}, N_{chg} or $\sum p_T$. All Monte Carlo simulations predict less spherical events compared to the data. PYTHIA 6 Z1 provides the overall best description while PYTHIA 6 DW is consistently furthest away from the data. The PYTHIA 6 DW Monte Carlo tune is the only simulation not tuned to LHC data. The bad description of the data by PYTHIA 6

DW shows that tuning Monte Carlo to LHC data is substantial for providing accurate predictions for LHC measurements. While Monte Carlo tunes focused on minimum bias data provide a more precise description in low p_T^{lead} bins, underlying event Monte Carlo tunes improve with increasing p_T^{lead} and give the better description in the highest p_T^{lead} bins. However, none of the Monte Carlo tunes tested was able to describe the data within the systematic uncertainties.

To analyze the continuous transition to harder events, the mean values of event shape variables were analyzed as a function of the charged particle multiplicity, N_{ch}, and the scalar sum of charged particle transverse momentum, $\sum p_T$. For the event shape variables as a function of the N_{ch}, PYTHIA 6 Z1 and PYTHIA 6 AMBT2B are able to describe the data within the uncertainties over a broad range, but they fail to describe the data at low N_{ch}. The description of the event shape variables as a function of $\sum p_T$ is bad for all Monte Carlo tunes tested and no Monte Carlo predicts the shape correctly. The results are published in the HEPdata format, in order to be available for future Monte Carlo tuning. Future Monte Carlo simulations can make use of the measurement by tuning its parameters to the results provided.

Increased statistics in Monte Carlo predictions and more sophisticated unfolding methods compensating the model dependency could reduce the systematic uncertainties in future measurements of event shape variables. Repeating the analysis at higher center-of-mass energies, for example in the upcoming data taking period at the LHC with $\sqrt{s} = 13$ TeV, can be used to extract information about the underlying event at higher energies. However, low pile-up conditions need to be provided. Adding neutrals particles to the measurement, by using calorimeter information instead of tracks, allows for measuring infrared and collinear safe versions of the event shape variables. With increasing integrated luminosities, event shapes can also be used to measure the CP eigenvalues of the Higgs boson and are candidates to analyze supersymmetric particle decays as well as black-hole decays.

As a second analysis, the measurement of the double differential cross section as a function of the dijet mass and rapidity separation of the two leading jets was presented. The jet production is one of the most dominant processes at the LHC and can be used to probe theoretical predictions up to high scales. The sensitivity of the dijet cross section to the parton density functions qualifies the measurement to contribute to future PDF extractions.

Proton–proton collisions recorded in 2012 at a center-of-mass energy of $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of 20.3 fb^{-1}, were analyzed. Jets were identified using the anti-k_t jet algorithm with two different jet radius parameters, $R = 0.4$ and $R = 0.6$. The two jets with the highest transverse momenta, called the leading jets, were required to be in the central region $|y_{1,2}| < 2.8$ and to have transverse momenta above $p_T > 100$ GeV and $p_T > 50$ GeV respectively. The results are presented in bins of the dijet mass and $y^* = |y_1 - y_2|/2$. Dijet masses between 280 GeV and 6 TeV are covered by the analysis. The results were corrected to particle level and compared to theory predictions, calculated using the NLOjet++ package, and corrected for non-perturbative effects such as hadronization and underlying event contributions. The main uncertainty contribution is the uncertainty on the jet energy scale. The uncertainty is around 10% at ~ 1 TeV after propagating the uncertainty through the unfolding process. Only a small increase towards lower dijet masses, predominantly due to uncertainties of the pile-up correction, can be observed. Towards higher dijet masses, the uncertainty rises to 20% and above. The lack of statistics at high jet transverse momenta does not allow for deriving and applying in-situ calibration methods with small uncertainties, as achieved in the lower p_T regime. Instead, the uncertainty is derived from using single particle responses, and combining them to a jet energy response.

The dijet cross section is consistently overestimated by the NLO prediction, except in the prediction using the HERAPDF 1.5 PDF set and jets with a radius parameter of $R = 0.6$. However, a agreement between theory predictions and data within the given uncertainties up to dijet masses of 2 TeV can be observed. For the two y^*-bins below 1.0, the discrepancy increases with increasing dijet mass. For $1.0 < y^* < 2.0$, the maximal discrepancy is reached at approximately 4 TeV before decreasing again. In some bins, the difference is not covered by the combined uncertainties anymore. At high y^* values, the NLOjet++ prediction is limited by the statistical uncertainties which leads to increased systematic uncertainties and fluctuating bins. Calculations of electroweak corrections, including electroweak loops to the QCD NLO prediction provided by NLOjet++ have not been finished before writing this thesis. In general, they are predicted to be negligible for larger values of y^*, but increasing the predicted cross section even further at low y^* and high dijet mass values.

A possible overcalibration of jets, affecting around 35% of the collected data, was discovered in mid 2014. At the time of writing this thesis, no suitable correction method was available, making a final evaluation of the results difficult, as the effect on the data is unknown. A main task of all jet measurements in ATLAS is to find a correction for the miscalibrated jets. Including a correction for the overcalibrated jets and applying electroweak corrections to the NLO predictions would round up the analysis. Once the analysis is finalized, the results can be used as additional input for PDF fits. The analysis was designed to be easily expanded to the more forward region. The extension to the forward region is of specific interest for PDF extractions. Extending the measurement to a triple differential cross section also increases the sensitivity to PDFs. Event selection, calibration methods, and used jet radii are matched to other jet cross section measurements within ATLAS for an easy combination of the results and for determination of uncertainty correlations across the analyses. With an increased center-of-mass energy for Run II at the LHC, starting in 2015, the accessible phase space is increased. The deviations observed in this measurement at higher dijet masses can be analyzed with a higher precision in future measurements.

Part VII

APPENDIX

A

EVENT SHAPE DISTRIBUTIONS - RESOLUTION

A.1 TRANSVERSE THRUST

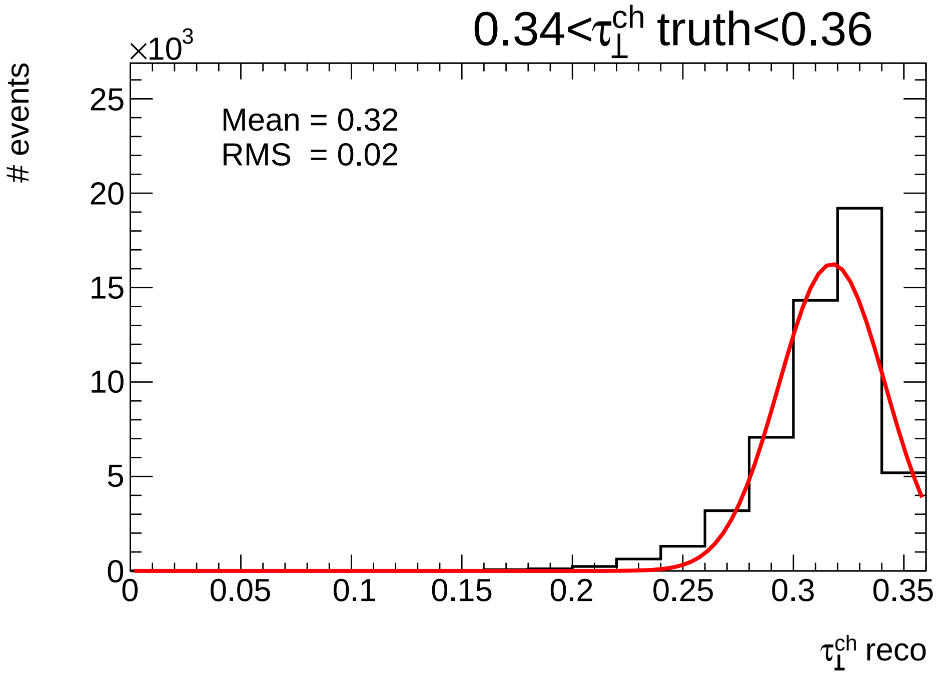


Figure A.1: Distribution of reconstructed values of the transverse thrust for events with a transverse thrust on truth level between $0.17 < \tau_{\perp}^{ch} < 0.18$. The 1σ-width of a Gaussian fit is used to determine the resolution.

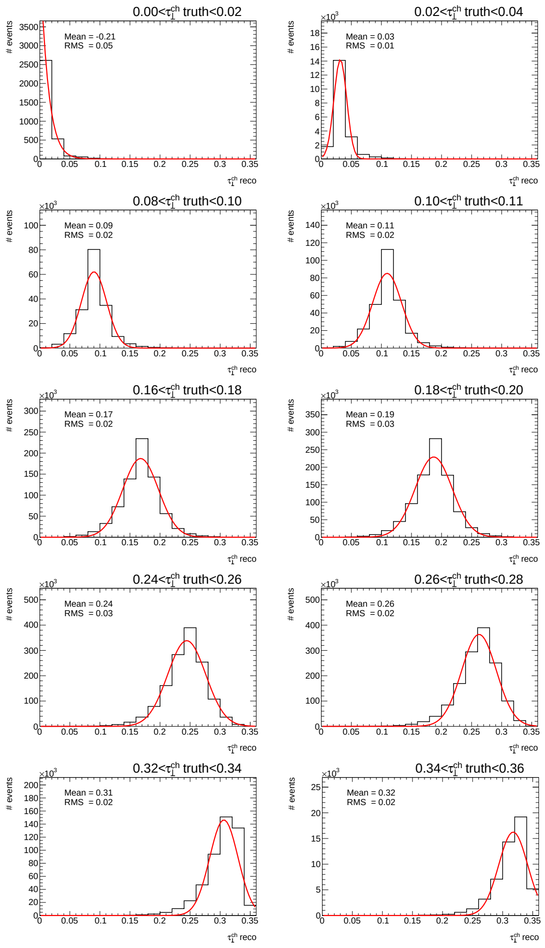


Figure A.2: Distributions of reconstructed values of the thrust minor for events for each bin in thrust minor on truth level. The 1σ-width of a Gaussian fit is used to determine the resolution.

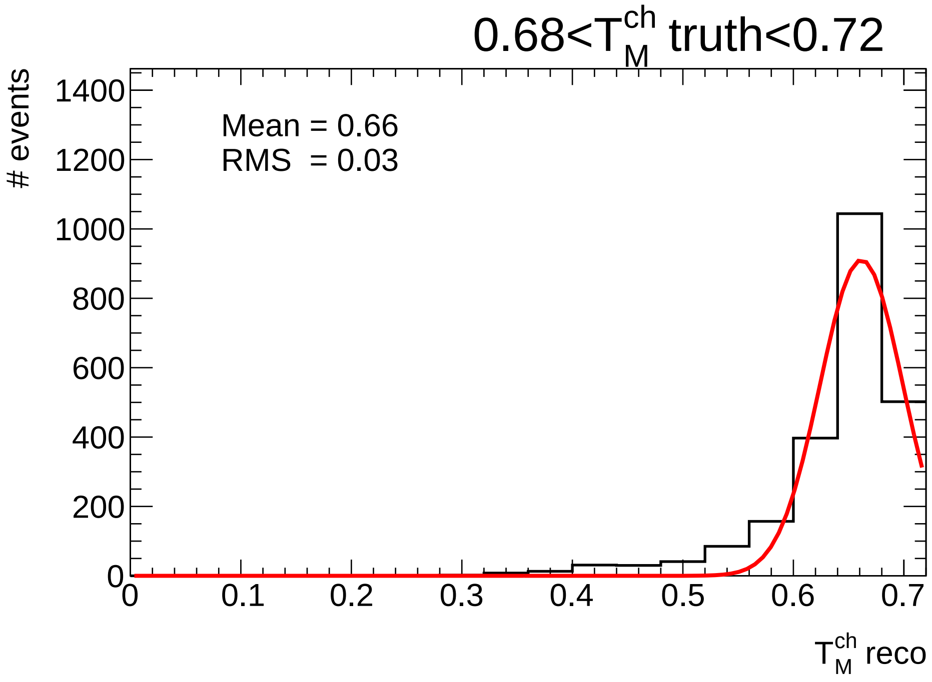


Figure A.3: Distribution of reconstructed values of the thrust minor for events with a thrust minor on truth level between $0.17 < \tau_{\perp}^{ch} < 0.18$. The 1σ-width of a Gaussian fit is used to determine the resolution.

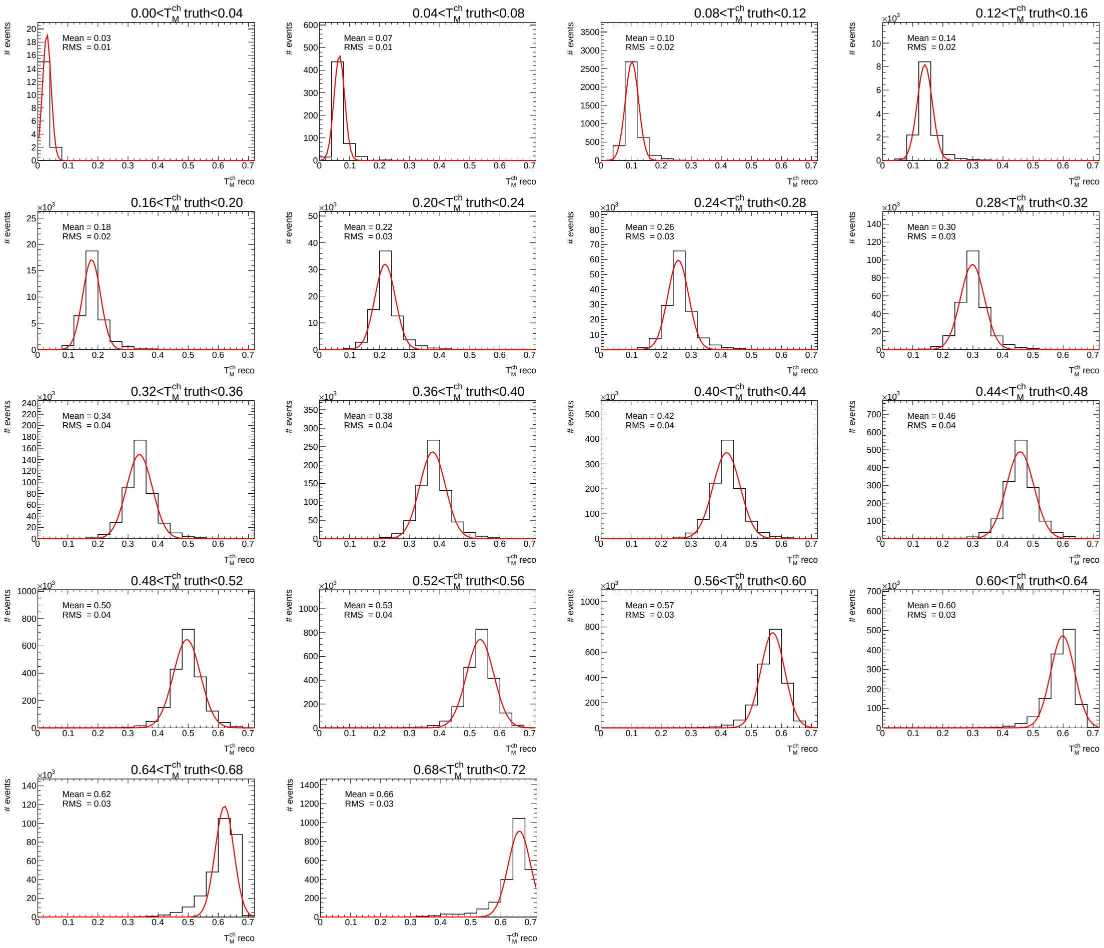


Figure A.4: Distributions of reconstructed values of the thrust minor for events for each bin in thrust minor on truth level. The 1σ-width of a Gaussian fit is used to determine the resolution.

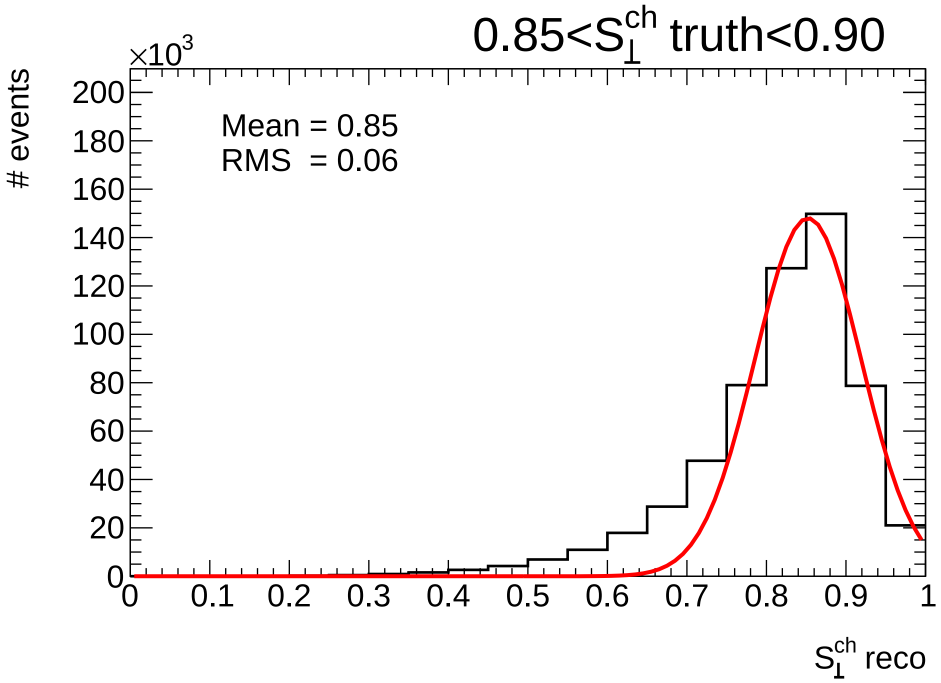
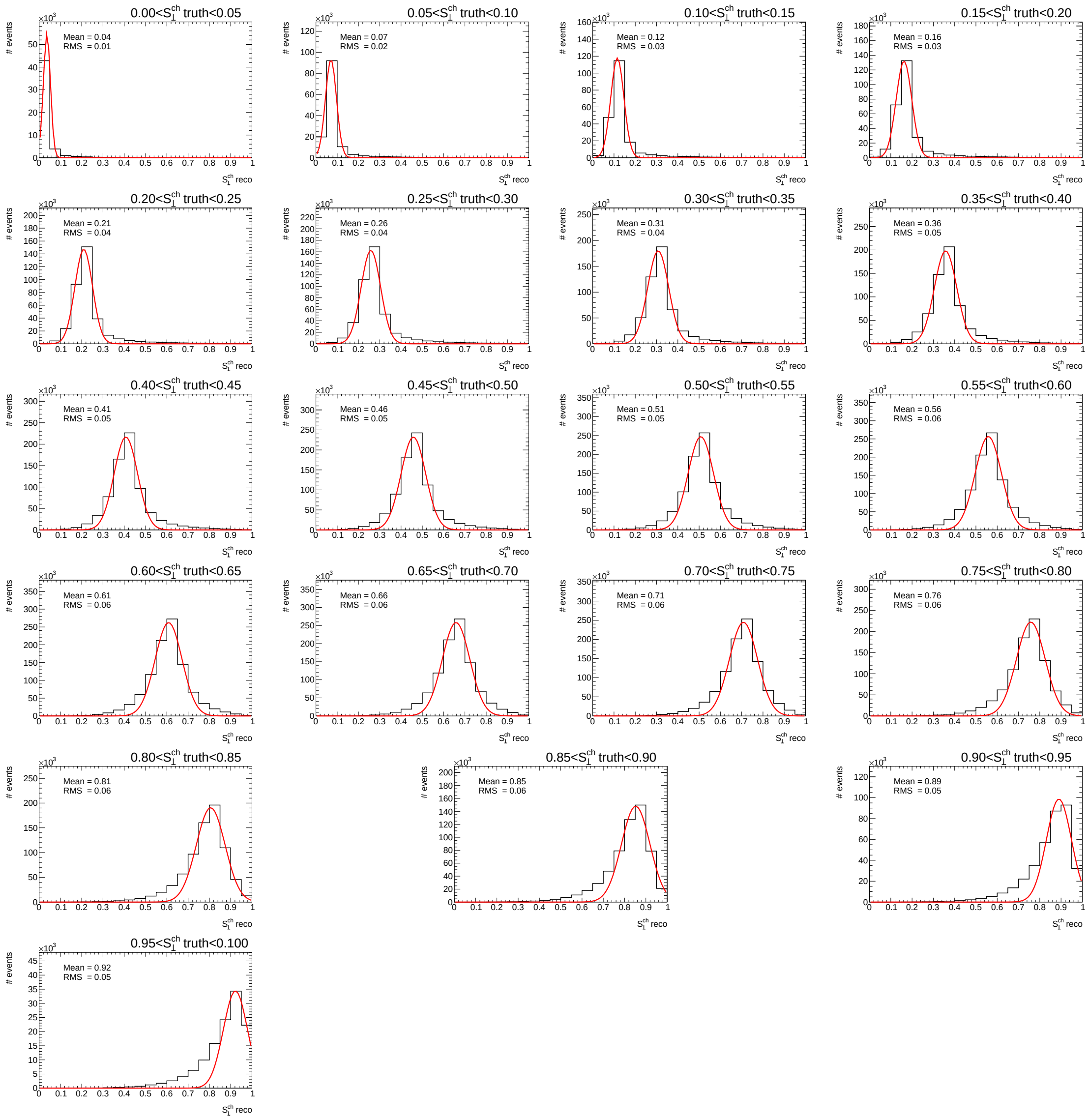


Figure A.5: Distribution of reconstructed values of the thrust minor for events with a thrust minor on truth level between $0.17 < \tau_\perp^{ch} < 0.18$. The 1σ-width of a Gaussian fit is used to determine the resolution.


$0.00<S_{\perp}^{ch}$ truth<0.05

$0.05<S_{\perp}^{ch}$ truth<0.10

$0.10<S_{\perp}^{ch}$ truth<0.15

$0.15<S_{\perp}^{ch}$ truth<0.20

$0.20<S_{\perp}^{ch}$ truth<0.25

$0.25<S_{\perp}^{ch}$ truth<0.30

$0.30<S_{\perp}^{ch}$ truth<0.35

$0.35<S_{\perp}^{ch}$ truth<0.40

$0.40<S_{\perp}^{ch}$ truth<0.45

$0.45<S_{\perp}^{ch}$ truth<0.50

$0.50<S_{\perp}^{ch}$ truth<0.55

$0.55<S_{\perp}^{ch}$ truth<0.60

$0.60<S_{\perp}^{ch}$ truth<0.65

$0.65<S_{\perp}^{ch}$ truth<0.70

$0.70<S_{\perp}^{ch}$ truth<0.75

$0.75<S_{\perp}^{ch}$ truth<0.80

$0.80<S_{\perp}^{ch}$ truth<0.85

$0.85<S_{\perp}^{ch}$ truth<0.90

$0.90<S_{\perp}^{ch}$ truth<0.95

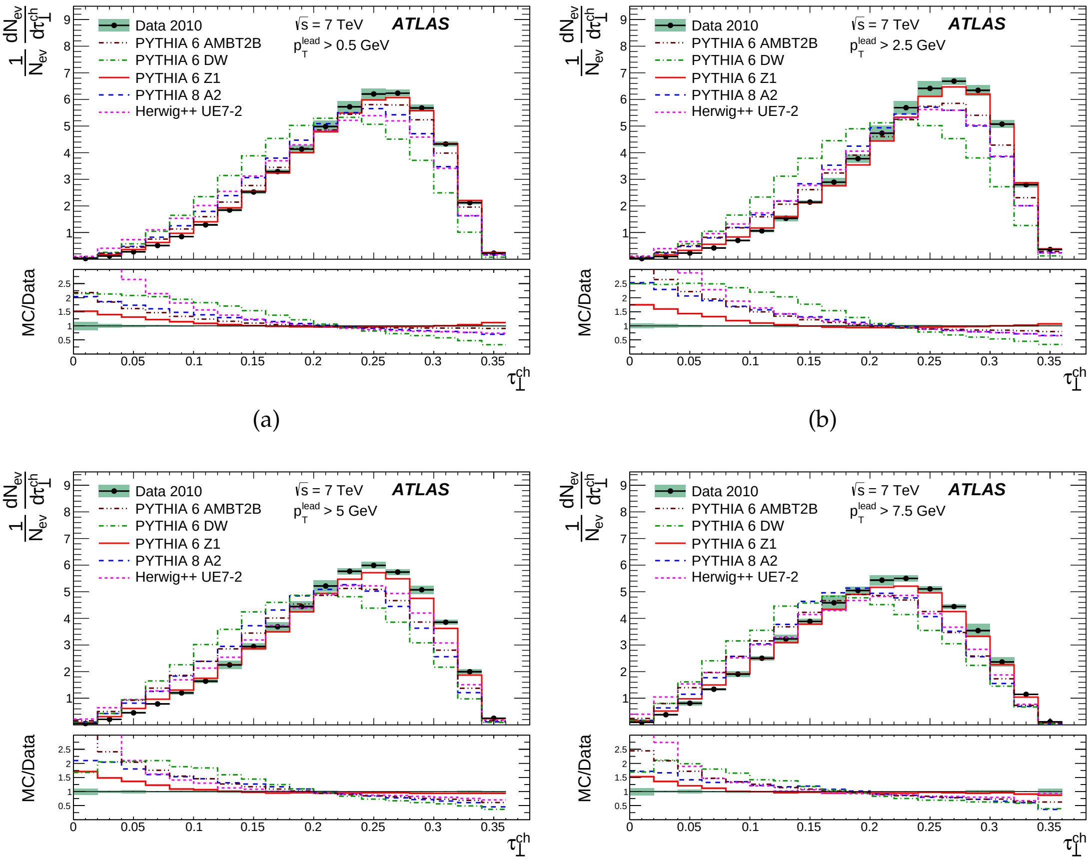
$0.95<S_{\perp}^{ch}$ truth<0.100

Figure A.6: Distributions of reconstructed values of the transverse sphericity for events for each bin in transverse sphericity on truth level. The 1σ-width of a Gaussian fit is used to determine the resolution.

B

INTEGRATED EVENT SHAPE DISTRIBUTIONS

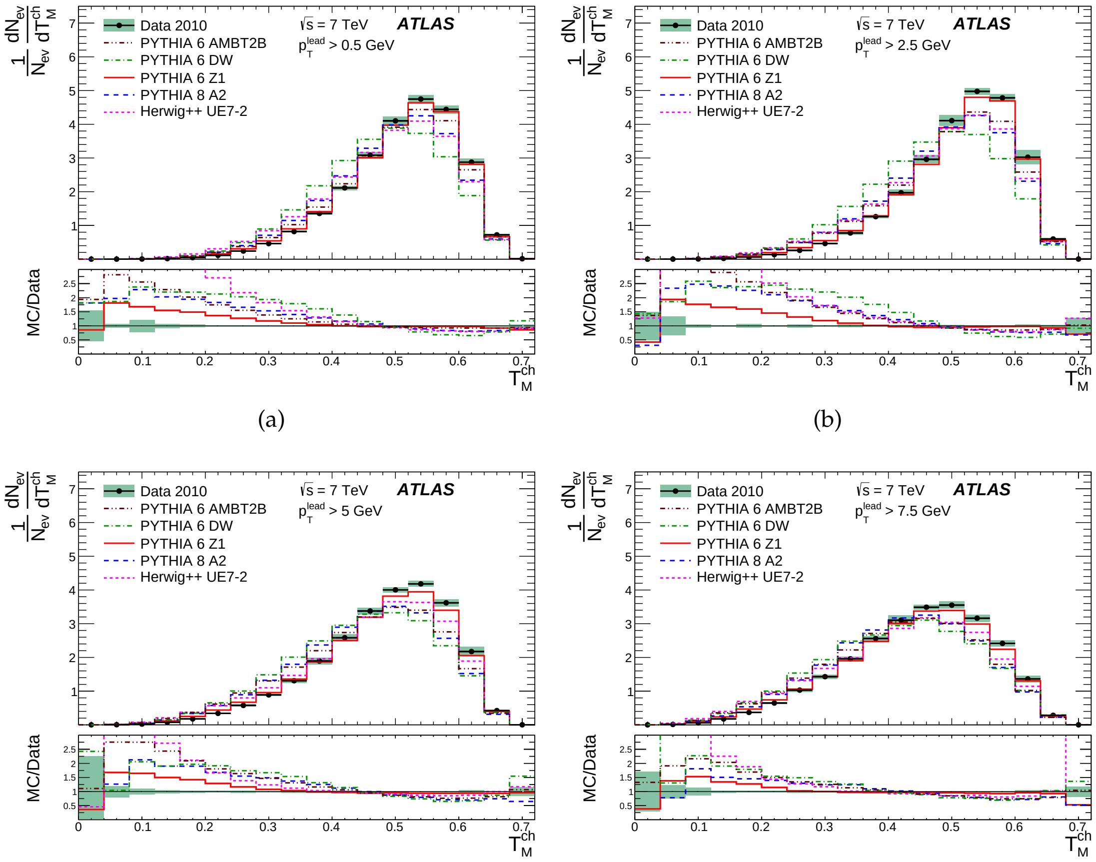
As discussed in Sec. 13.1, the integrated distributions show the exact same behaviour as the differential ones. Table 13.1 shows the reason for the similarity: The integrated distributions are dominated by the events with the lowest included p_T^{lead}, due to the steeply falling spectrum of the leading track transverse momentum. While in the first plot ($p_T^{\text{lead}} \leqslant 2.5\ \text{GeV}$), 68% of all events are within the the lowest p_T^{lead} bin ($0.5\ \text{GeV} < p_T^{\text{lead}} \leqslant 2.5\ \text{GeV}$), the second minimum leading track transverse momentum ($p_T^{\text{lead}} \leqslant 5\ \text{GeV}$) 90% of all events are in the lowest allowed p_T^{lead} bin. This continues in the following bins in the same way. Therefore, all conclusions drawn in Sec. 13.1 are valid for the following spectra as well and no sever difference can be observed.



(a) (b)

(c) (d)

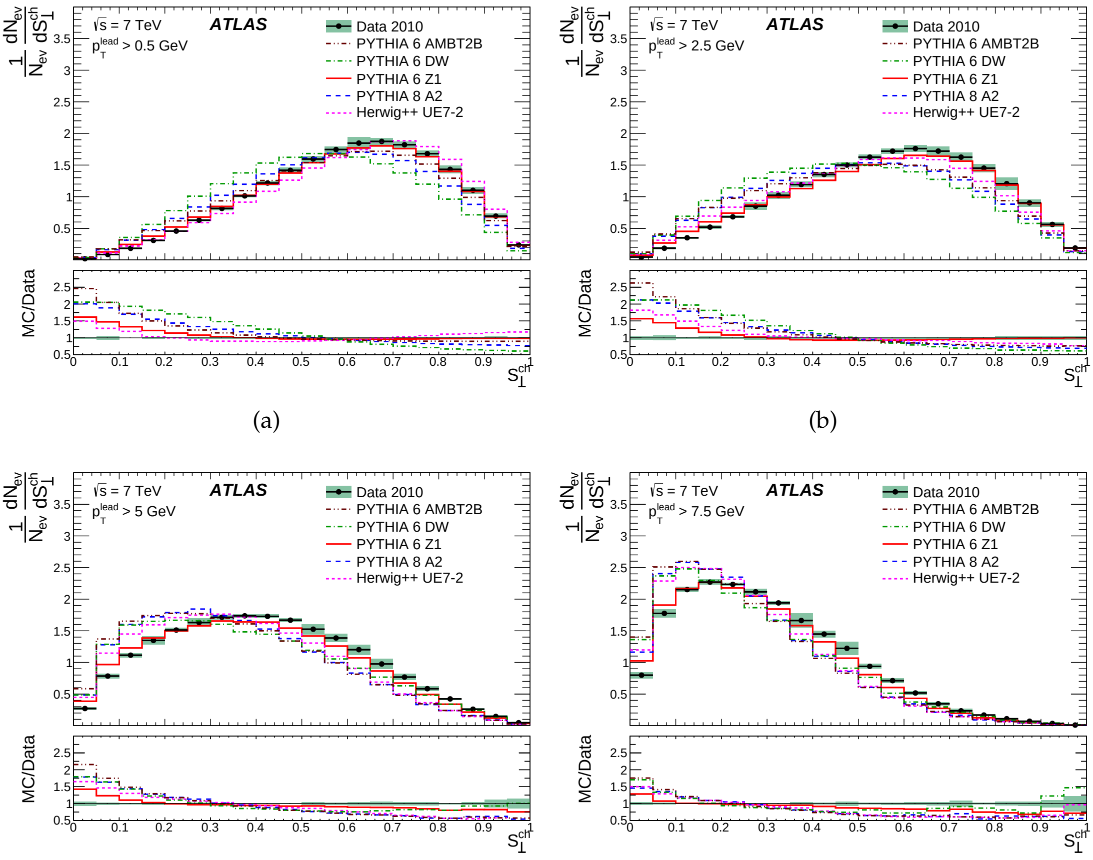
Figure B.1: Normalized distributions of the transverse thrust using at least six tracks with $|\eta| < 2.5$ and $p_T > 0.5$ GeV for different minimum transverse momentum of the leading charged particle, p_T^{lead}. The error bars drawn indicate statistical uncertainties only. The shaded area represent the combined statistical and systematic uncertainties. Where not visible, the statistical uncertainty is smaller than the marker size[93].



(a) (b)

(c) (d)

Figure B.2: Normalized distributions of the thrust minor using at least tracks with $|\eta| < 2.5$ and $p_T > 0.5$ GeV for different minimum transverse momentum of the leading charged particle, p_T^{lead}. The error bars drawn indicate statistical uncertainties only. The shaded area represent the combined statistical and systematic uncertainties. Where not visible, the statistical uncertainty is smaller than the marker size[93].



(a) (b)

(c) (d)

Figure B.3: Normalized distributions of the transverse sphericity using at least six tracks with $|\eta| < 2.5$ and $p_T > 0.5$ GeV for different minimum transverse momentum of the leading charged particle, p_T^{lead}. The error bars drawn indicate statistical uncertainties only. The shaded area represent the combined statistical and systematic uncertainties. Where not visible, the statistical uncertainty is smaller than the marker size[93].

C

BASIC KINEMATIC DISTRIBUTIONS IN BINS OF RAPIDITY SEPARATION

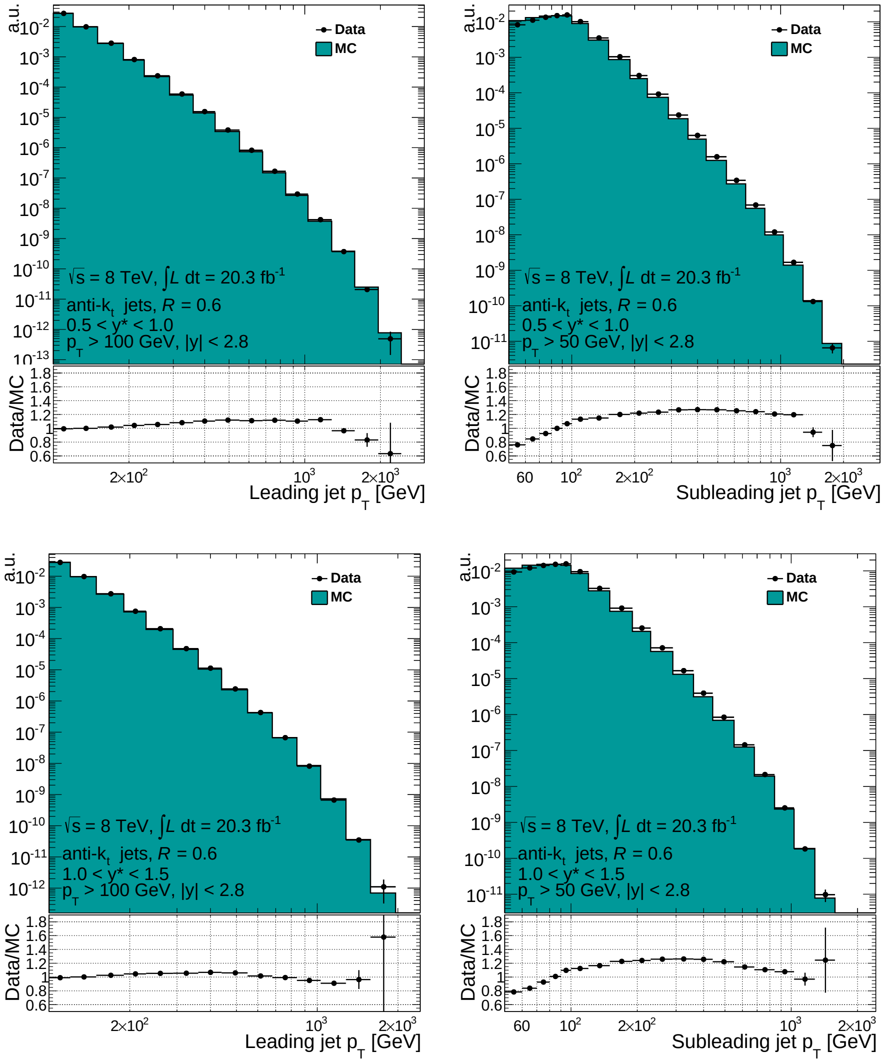


Figure C.1: Comparison of the leading (left) and subleading (right) p_T distributions in Monte Carlo and data in five different y^* bins. The distributions are normalized to their integrals.

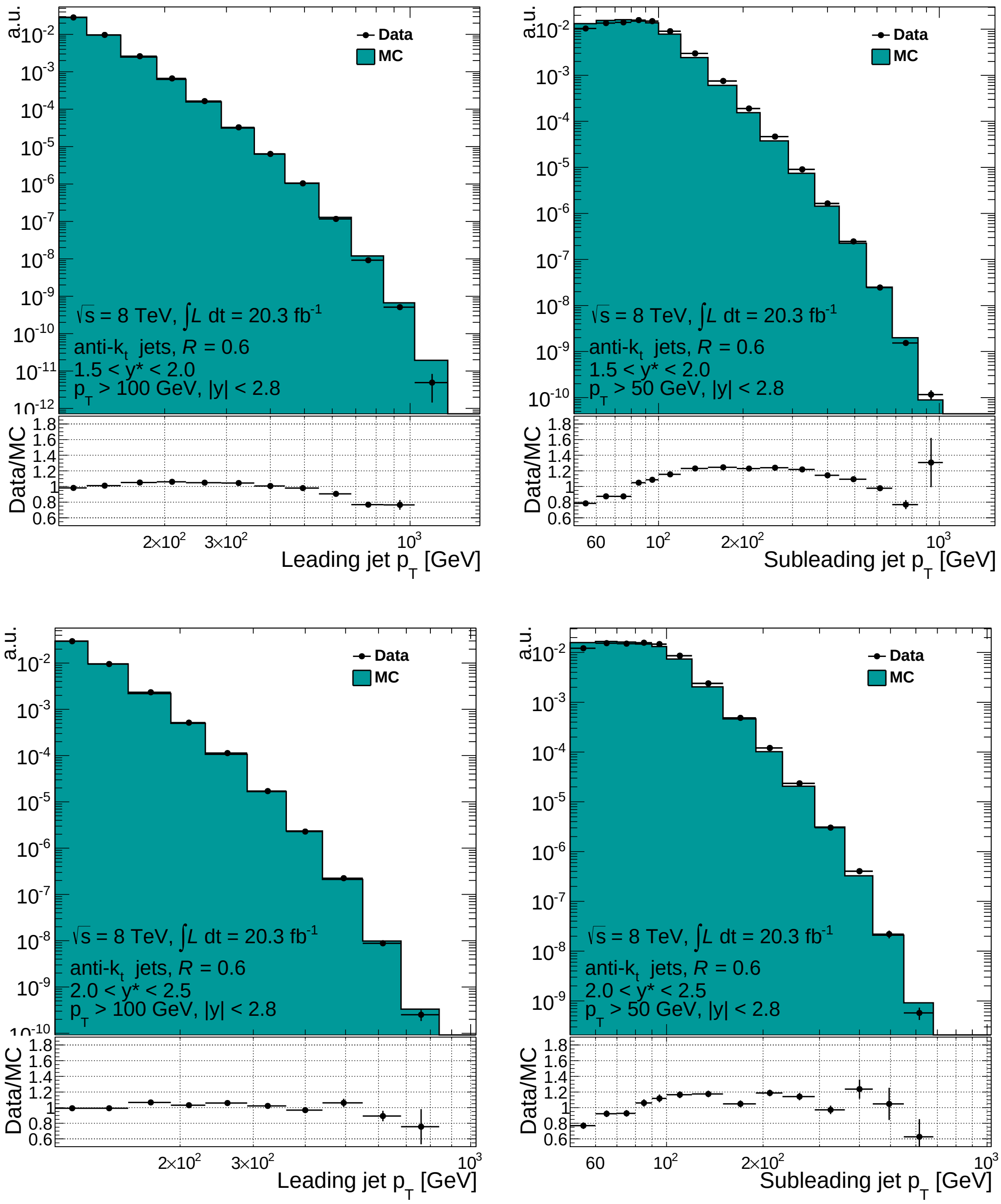


Figure C.2: Comparison of the leading (left) and subleading (right) p_T distributions in Monte Carlo and data in five different y^* bins. The distributions are normalized to their integrals.

D

SYSTEMATIC UNCERTAINTY ON THE MEASURED DIJET CROSS SECTION

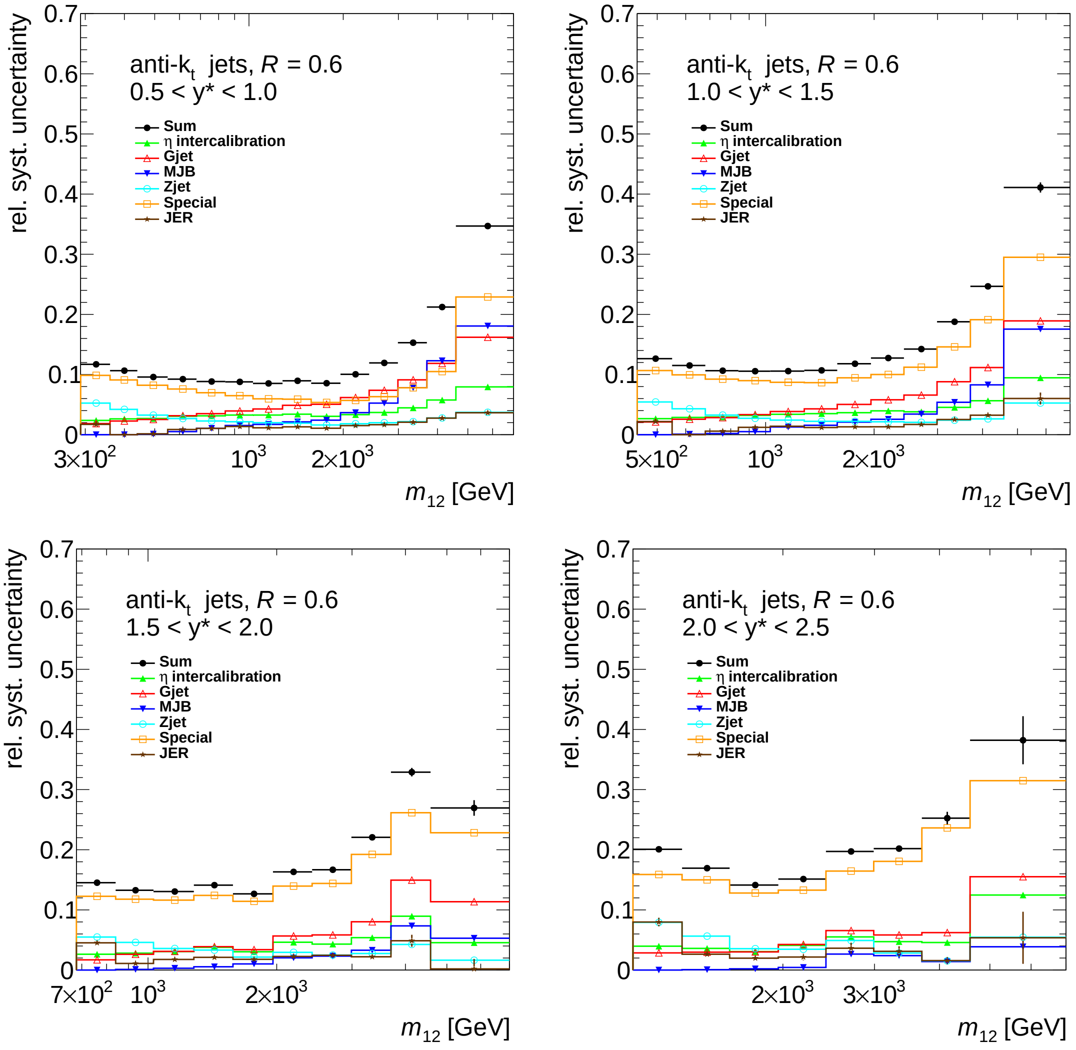


Figure D.1: Relative systematic uncertainties on the dijet mass propagated through the unfolding. Uncertainties from the single jet energy scale correction methods are combined. Special combines the remaining uncertainties on the pile-up subtraction, flavour composition and single particle response. *Insitu*, green, denotes as the η-intercalibration method. Black points represents the combined uncertainties.

E

DIJET CROSS SECTION COMPARISON WITH THEORY - ANTI-k_t $R = 0.4$

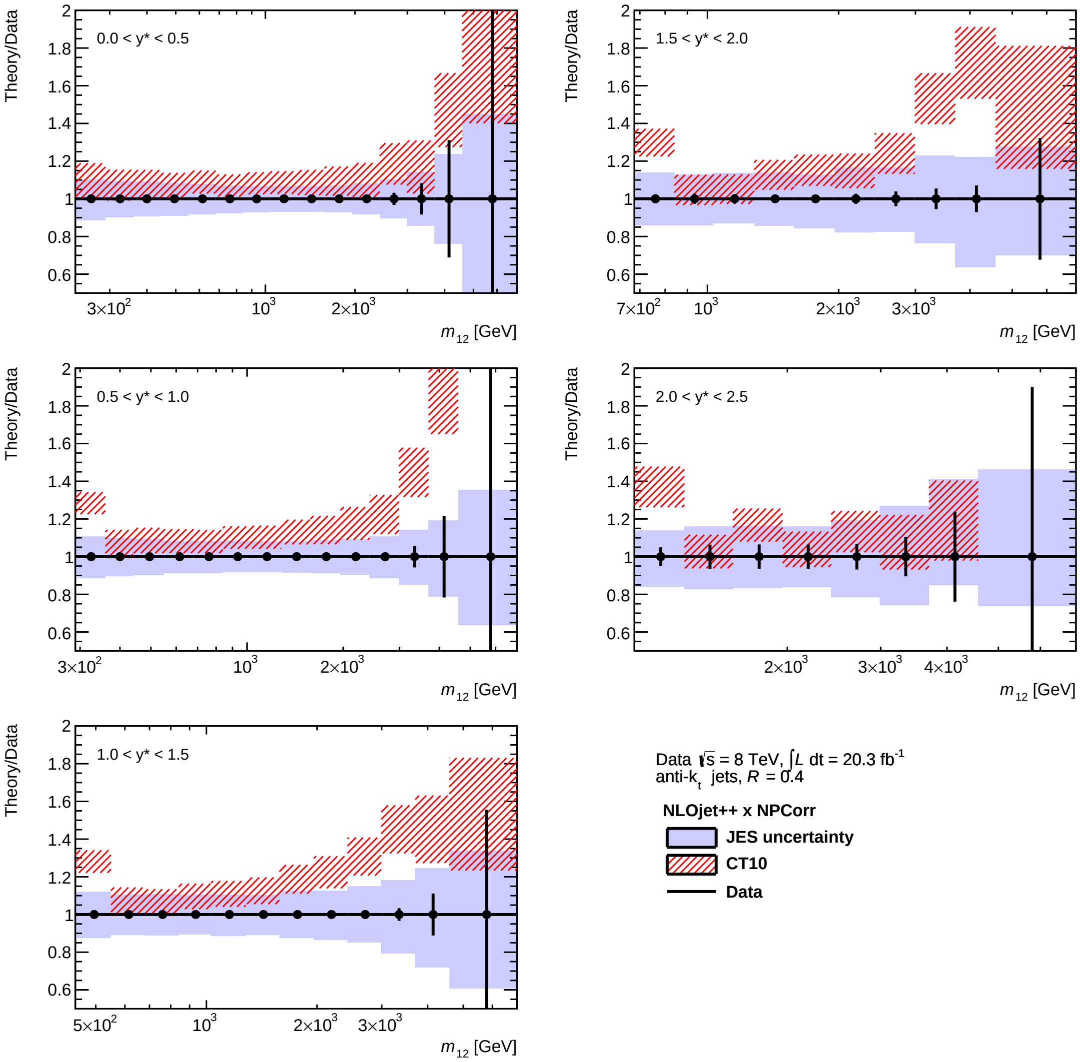


Figure E.1: Ratio of the NLOjet++ theory prediction, using the CT10 PDF set and corrected for non-perturbative corrections, and the measurement of the double differential cross section for anti-k_t jets with a radius parameter of $R = 0.4$. Error bars represent statistical uncertainties of the data. The combined systematic uncertainty is indicated by the blue area. Systematic uncertainties on the NLOjet++ predictions are drawn as red shaded area.

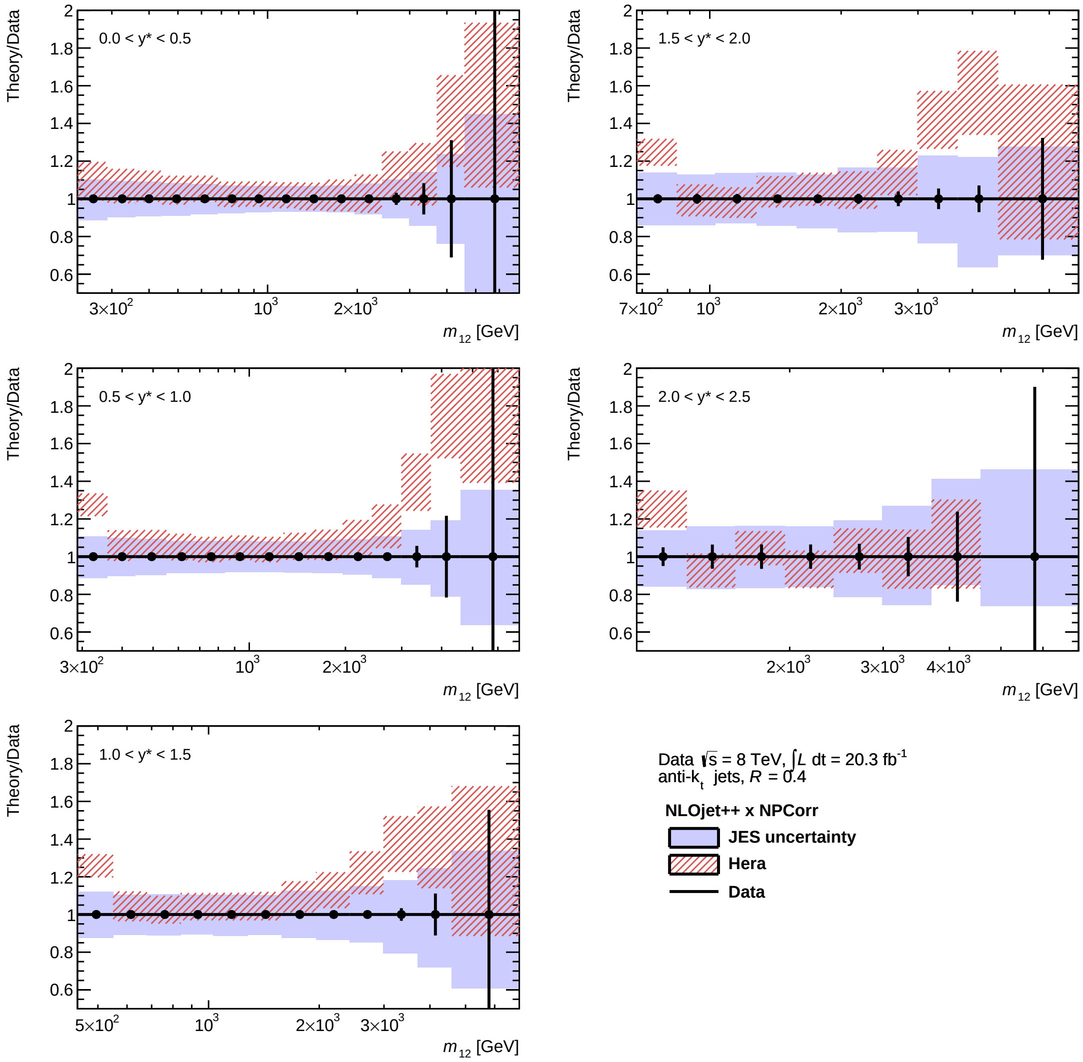


Figure E.2: Ratio of the NLOjet++ theory prediction, using the HERAPDF 1.5 PDF set and corrected for non-perturbative corrections, and the measurement of the double differential cross section for anti-k_t jets with a radius parameter of $R = 0.4$. Error bars represent statistical uncertainties of the data. The combined systematic uncertainty is indicated by the blue area. Systematic uncertainties on the NLOjet++ predictions are drawn as red shaded area.

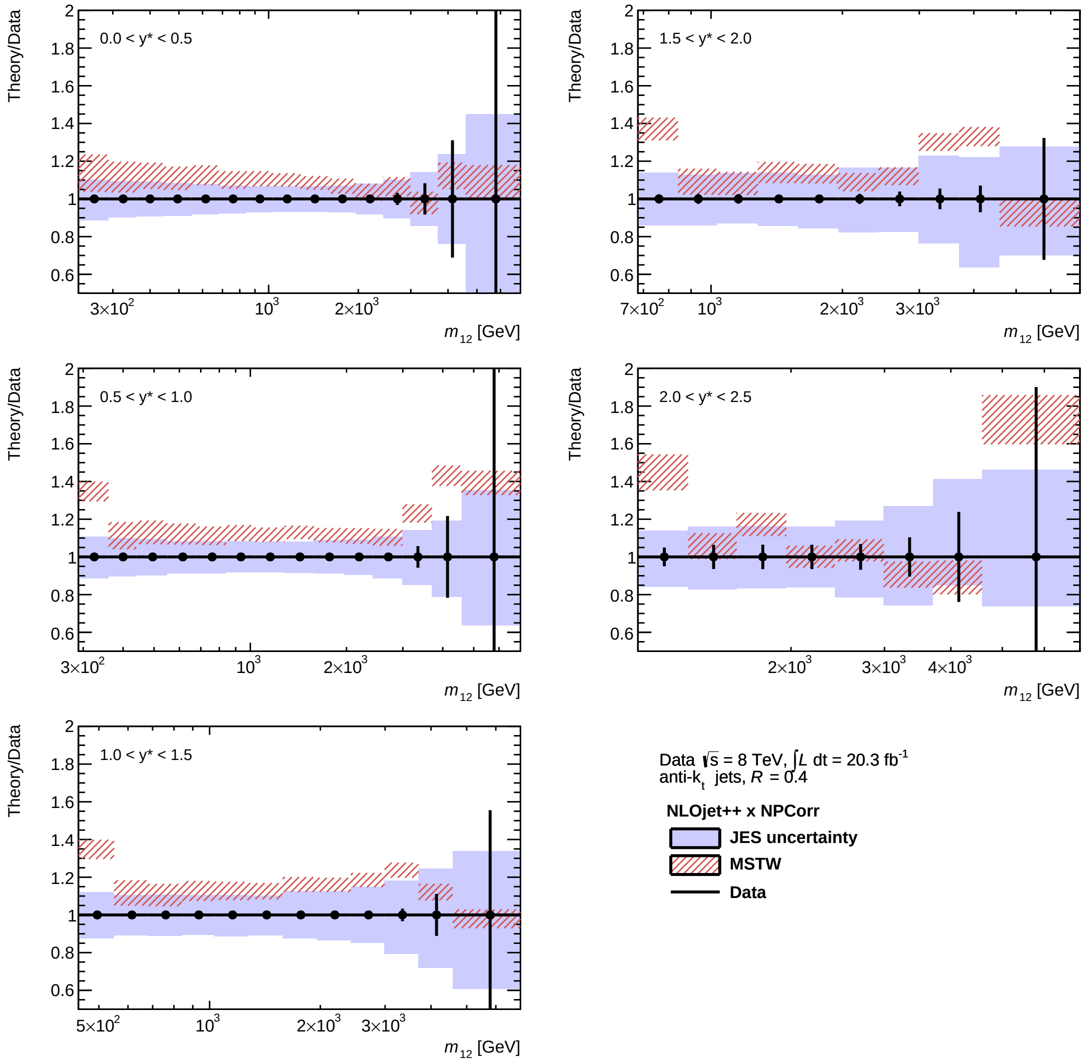


Figure E.3: Ratio of the NLOjet++ theory prediction, using the MSTW2008 PDF set and corrected for non-perturbative corrections, and the measurement of the double differential cross section for anti-k_t jets with a radius parameter of $R = 0.4$. Error bars represent statistical uncertainties of the data. The combined systematic uncertainty is indicated by the blue area. Systematic uncertainties on the NLOjet++ predictions are drawn as red shaded area.

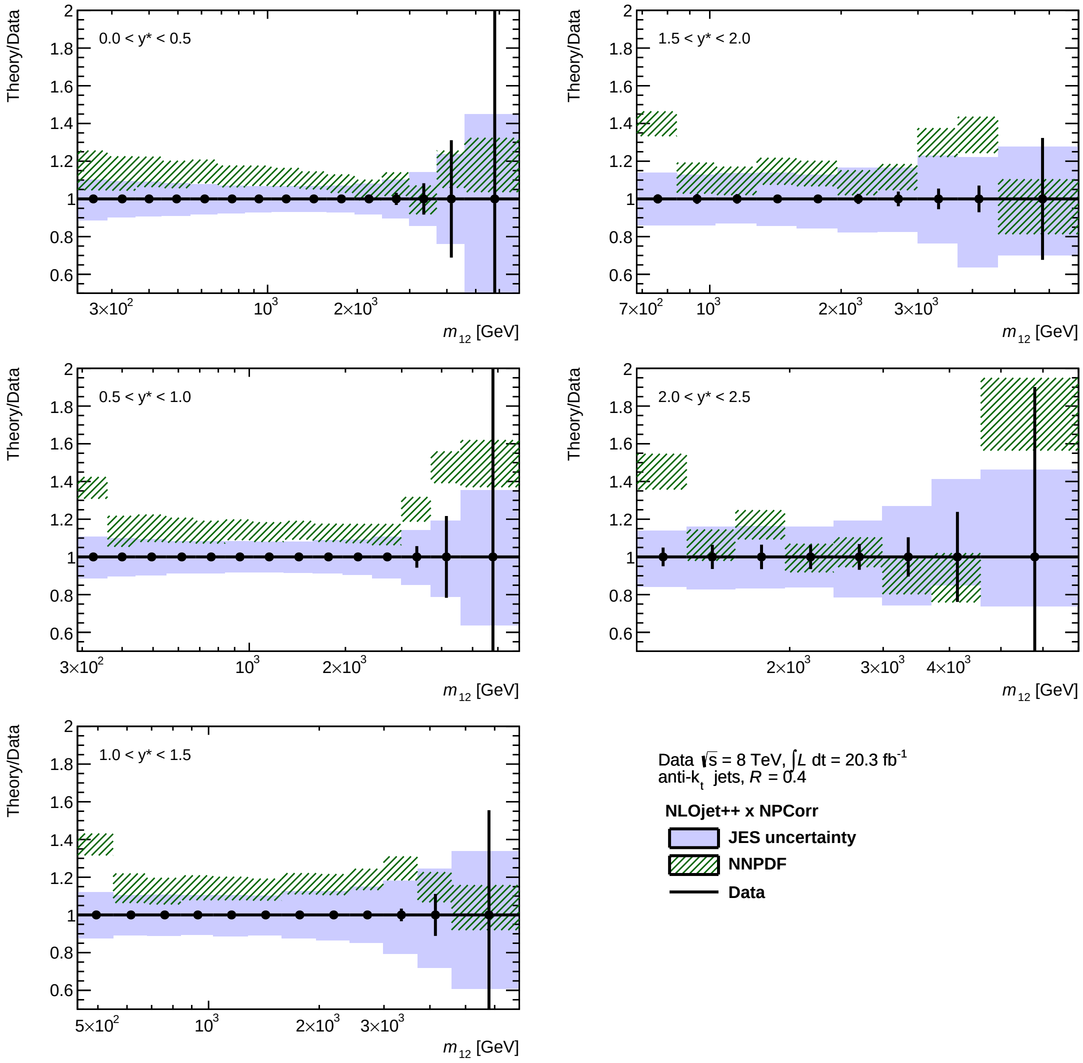


Figure E.4: Ratio of the NLOjet++ theory prediction, using the NNPDF 2.3 PDF set and corrected for non-perturbative corrections, and the measurement of the double differential cross section for anti-k_t jets with a radius parameter of $R = 0.4$. Error bars represent statistical uncertainties of the data. The combined systematic uncertainty is indicated by the blue area. Systematic uncertainties on the NLOjet++ predictions are drawn as green shaded area.

ACKNOWLEDGMENTS

Aus Datenschutzgründen entfernt.

LIST OF FIGURES

LIST OF TABLES

BIBLIOGRAPHY

[1] Murray Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys.Lett.*, 8: 214–215, 1964. doi: 10.1016/S0031-9163(64)92001-3.

[2] G. Zweig. An SU(3) model for strong interaction symmetry and its breaking. Version 1. 1964.

[3] S.L. Glashow. Towards a Unified Theory: Threads in a Tapestry. *Rev.Mod.Phys.*, 52:539–543, 1980. doi: 10.1103/RevModPhys.52.539.

[4] Abdus Salam. Gauge Unification of Fundamental Forces. *Rev.Mod.Phys.*, 52:525–538, 1980. doi: 10.1103/RevModPhys.52.525.

[5] Steven Weinberg. Conceptual Foundations of the Unified Theory of Weak and Electromagnetic Interactions. *Rev.Mod.Phys.*, 52:515–523, 1980. doi: 10.1103/RevModPhys.52.515.

[6] Georges 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.*, B716:1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.

[7] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.

[8] Wikimedia Commons. Standard model of elementary particles. `http://upload.wikimedia.org/wikipedia/commons/thumb/0/00/Standard_Model_of_Elementary_Particles.svg/819px-Standard_Model_of_Elementary_Particles.svg.png`, . Date Accessed: July 22, 2014.

[9] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys.Rev.Lett.*, 13:321–323, 1964. doi: 10.1103/PhysRevLett.13.321.

[10] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys.Lett.*, 12:132–133, 1964. doi: 10.1016/0031-9163(64)91136-9.

[11] T. Kinoshita. Mass singularities of feynman amplitudes. *J. Math. Phys*, 3:650–67, 1962.

[12] T. D. Lee and M. Nauenberg. Degenerate systems and mass singularities. *Phys. Rev.*, 133:B1549–B1562, 1964.

[13] J. Beringer et al. Review of particle physics. *Phys. Rev. D*, 86:010001, 2012.

[14] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton Fragmentation and String Dynamics. *Phys.Rept.*, 97:31–145, 1983. doi: 10.1016/0370-1573(83)90080-7.

[15] M. Bahr et al. Herwig++ Physics and Manual. *Eur. Phys. J. C*, 58:639–707, 2008. doi: 10.1140/epjc/s10052-008-0798-9.

[16] T. Sjostrand, S. Mrenna, and P. Skands. PYTHIA 6.4 physics and manual. *J. High Energy Phys.*, 05:026, 2006.

[17] Siegert, Frank. *Monte-Carlo event generation for the LHC*. PhD thesis, Durham University, 2010. Available at Durham E-Theses Online: `http://etheses.dur.ac.uk/484/`.

[18] Guido Altarelli and G. Parisi. Asymptotic Freedom in Parton Language. *Nucl.Phys.*, B126:298, 1977. doi: 10.1016/0550-3213(77)90384-4.

[19] C. Adloff et al. Measurement of neutral and charged current cross-sections in positron proton collisions at large momentum transfer. *Eur.Phys.J.*, C13:609–639, 2000. doi: 10.1007/s100520050721.

[20] S. Chekanov et al. Measurement of the neutral current cross-section and F(2) structure function for deep inelastic e + p scattering at HERA. *Eur.Phys.J.*, C21: 443–471, 2001. doi: 10.1007/s100520100749.

[21] L.N. Lipatov. The parton model and perturbation theory. *Sov.J.Nucl.Phys.*, 20: 94–102, 1975.

[22] V.N. Gribov and L.N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov.J.Nucl.Phys.*, 15:438–450, 1972.

[23] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov.Phys.JETP*, 46:641–653, 1977.

[24] J. Stirling. MSTW PDFs : key plots. URL `https://mstwpdf.hepforge.org/plots/plots.html`.

[25] A.D. Martin, W.J. Stirling, R.S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur.Phys.J.*, C63:189–285, 2009. doi: 10.1140/epjc/s10052-009-1072-5.

[26] Gavin P. Salam and Gregory Soyez. A Practical Seedless Infrared-Safe Cone jet algorithm. *JHEP*, 0705:086, 2007. doi: 10.1088/1126-6708/2007/05/086.

[27] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Phys.Rev.*, D48:3160–3166, 1993. doi: 10.1103/PhysRevD.48.3160.

[28] Gavin P. Salam. Towards Jetography. *Eur.Phys.J.*, C67:637–686, 2010. doi: 10.1140/epjc/s10052-010-1314-6.

[29] Mrinal Dasgupta, Lorenzo Magnea, and Gavin P. Salam. Non-perturbative QCD effects in jets at hadron colliders. *JHEP*, 0802:055, 2008. doi: 10.1088/1126-6708/2008/02/055.

[30] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *J. High Energy Phys.*, 11:040, 2004. doi: 10.1088/1126-6708/2004/11/040.

[31] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *J. High Energy Phys.*, 11:070, 2007. doi: 10.1088/1126-6708/2007/11/070.

[32] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *J. High Energy Phys.*, 06:043, 2010. doi: 10.1007/JHEP06(2010)043.

[33] S. Agostinelli et al. Geant4: a simulation toolkit. *Nucl. Instr. Meth. A*, 506:250–303, 2003. doi: 10.1016/S0168-9002(03)01368-8.

[34] S. Catani and M.H. Seymour. A General algorithm for calculating jet cross-sections in NLO QCD. *Nucl.Phys.*, B485:291–419, 1997. doi: 10.1016/S0550-3213(96)00589-5.

[35] Zoltan Nagy. Next-to-leading order calculation of three jet observables in hadron hadron collision. *Phys.Rev.*, D68:094002, 2003. doi: 10.1103/PhysRevD.68.094002.

[36] Tancredi Carli, Dan Clements, Amanda Cooper-Sarkar, Claire Gwenlan, Gavin P. Salam, et al. A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project. *Eur.Phys.J.*, C66:503–524, 2010. doi: 10.1140/epjc/s10052-010-1255-0.

[37] Lyndon Evans, (ed.) and Philip Bryant, (ed.). LHC Machine. *JINST*, 3:S08001, 2008. doi: 10.1088/1748-0221/3/08/S08001.

[38] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi: 10.1088/1748-0221/3/08/S08003.

[39] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008. doi: 10.1088/1748-0221/3/08/S08004.

[40] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008. doi: 10.1088/1748-0221/3/08/S08002.

[41] Jr. Alves, A. Augusto et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi: 10.1088/1748-0221/3/08/S08005.

[42] G. Anelli et al. The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008. doi: 10.1088/1748-0221/3/08/S08007.

[43] O. Adriani et al. The LHCf detector at the CERN Large Hadron Collider. *JINST*, 3:S08006, 2008. doi: 10.1088/1748-0221/3/08/S08006.

[44] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi: 10.1088/1748-0221/3/08/S08003.

[45] Wikimedia Commons. Pseudorapidity. `http://commons.wikimedia.org/wiki/File:Pseudorapidity_plot.svg`, . Date Accessed: July 22, 2014.

[46] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl.Instrum.Meth.*, A389:81–86, 1997. doi: 10.1016/S0168-9002(97)00048-X.

[47] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, et al. Calorimeter clustering algorithms: Description and performance. 2008.

[48] ATLAS Collaboration. Jet energy scale and its systematic uncertainty in proton-proton collisions at sqrt(s)=7 TeV with ATLAS 2011 data. Technical Report ATLAS-CONF-2013-004, CERN, Geneva, Jan 2013.

[49] ATLAS Collaboration. Drift time measurement in the atlas liquid argon electromagnetic calorimeter using cosmic muons. *arXiv*, 1002.4189, 2010.

[50] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Physics Letters B*, 659(1–2):119 – 126, 2008. ISSN 0370-2693. doi: http://dx.doi.org/10.1016/j.physletb.2007.09.077. URL `http://www.sciencedirect.com/science/article/pii/S0370269307011094`.

[51] ATLAS Collaboration. Pile-up corrections for jets from proton-proton collisions at $\sqrt{s} = 7$ TeV in ATLAS in 2011. Technical Report ATLAS-CONF-2012-064, CERN, Geneva, Jul 2012.

[52] ATLAS Collaboration. Pile-up subtraction and suppression for jets in ATLAS. Technical Report ATLAS-CONF-2013-083, CERN, Geneva, Aug 2013.

[53] ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur.Phys.J.*, C73:2304, 2013. doi: 10.1140/epjc/s10052-013-2304-2.

[54] Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-072, CERN, Geneva, Jul 2010.

[55] T Cornelissen, M Elsing, S Fleischmann, W Liebig, E Moyse, and A Salzburger. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT). Technical Report ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar 2007.

[56] Siegfried Bethke. α_s at Zinnowitz 2004. *Nucl. Phys. Proc. Suppl.*, 135:345–352, 2004. doi: 10.1016/j.nuclphysbps.2004.09.020.

[57] OPAL Collaboration. A simultaneous measurement of the QCD colour factors and the strong coupling. *Eur. Phys. J. C*, 20:601–615, 2001. doi: 10.1007/s100520100699.

[58] DELPHI Collaboration. Tuning and test of fragmentation models based on identified particles and precision event shape data. *Z. Phys. C*, 73:11–60, 1996. doi: 10.1007/s002880050295.

[59] Thomas Kluge. Review of power corrections in DIS. 2006.

[60] P. A. Movilla Fernandez, S. Bethke, O. Biebel, and S. Kluth. Tests of power corrections for event shapes in e+ e- annihilation. *Eur. Phys. J. C*, 22:1–15, 2001. doi: 10.1007/s100520100750.

[61] T. Akesson et al. The dominance of jets at large transverse energy in a full-azimuth hadron calorimeter at ISR energies. *Physics Letters B*, 128(5):354 – 360, 1983. doi: 10.1016/0370-2693(83)90274-5.

[62] C. Albajar et al. Analysis of the highest transverse energy events seen in the ua 1 detector at the spps collider. *Zeitschrift für Physik C Particles and Fields*, 36(1), 1987. ISSN 0170-9739. doi: 10.1007/BF01556162.

[63] R. Ansari et al. JET MEASURES AND HADRONIC EVENT SHAPES AT THE CERN anti-p p COLLIDER. *Z. Phys.*, C36:175, 1987. doi: 10.1007/BF01579132.

[64] CDF Collaboration. Measurement of Event Shapes in Proton-Antiproton Collisions at Center-of-Mass Energy 1.96 TeV. *Phys. Rev. D*, 83:112007, 2011. doi: 10.1103/PhysRevD.83.112007.

[65] Christoph Englert, Michael Spannowsky, and Michihisa Takeuchi. Measuring Higgs CP and couplings with hadronic event shapes. *JHEP*, 1206:108, 2012. doi: 10.1007/JHEP06(2012)108.

[66] Andrea Banfi, Gavin P. Salam, and Giulia Zanderighi. Phenomenology of event shapes at hadron colliders. *J. High Energy Phys.*, 06:038, 2010. doi: 10.1007/JHEP06(2010)038.

[67] ATLAS Collaboration. Charged-particle multiplicities in pp interactions at $\sqrt{s} = 900$ GeV measured with the ATLAS detector at the LHC. *Physics Letters B*, 688: 21–32, 2010. doi: 10.1016/j.physletb.2010.03.064.

[68] ATLAS Collaboration. Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC. *New J. Phys.*, 13:053033, 2011. doi: 10.1088/1367-2630/13/5/053033.

[69] ATLAS Collaboration. Measurement of underlying event characteristics using charged particles in pp collisions at $\sqrt{s} = 900$ GeV and 7 TeV with the ATLAS detector. *Phys. Rev. D*, 83:112001, 2011. doi: 10.1103/PhysRevD.83.112001.

[70] ATLAS Collaboration. Measurements of underlying-event properties using neutral and charged particles in pp collisions at 900 GeV and 7 TeV with the ATLAS detector at the LHC. *Eur. Phys. J. C*, 71:1636, 2011. doi: 10.1140/epjc/s10052-011-1636-z.

[71] ALICE Collaboration. Charged-particle multiplicity measurement in proton–proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC. *Eur. Phys. J. C*, 68:345–354, 2010. doi: 10.1140/epjc/s10052-010-1350-2.

[72] ALICE Collaboration. Charged-particle multiplicity measurement in proton–proton collisions at $\sqrt{s} = 0.9$ and 2.36 TeV with ALICE at LHC. *Eur. Phys. J. C*, 68:89–108, 2010. doi: 10.1140/epjc/s10052-010-1339-x.

[73] ALICE Collaboration. Underlying Event measurements in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV with the ALICE experiment at the LHC. *J. High Energy Phys*, 07: 116, 2012. doi: 10.1007/JHEP07(2012)116.

[74] CMS Collaboration. Transverse momentum and pseudorapidity distributions of charged hadrons in pp collisions at $\sqrt{s} = 0.9$ and 2.36 TeV. *J. High Energy Phys.*, 02:041, 2010. doi: 10.1007/JHEP02(2010)041.

[75] CMS Collaboration. Measurement of the underlying event activity in proton–proton Collisions at 0.9 TeV. 2010.

[76] ATLAS Collaboration. Measurement of charged-particle event shape variables in $\sqrt{s} = 7$ TeV proton-proton interactions with the ATLAS detector. *Phys.Rev.*, D88: 032004, 2013. doi: 10.1103/PhysRevD.88.032004.

[77] Tobias Hülsing. *Messung von Ereignisformvariablen bei inelastischen* pp*-Ereignissen mit dem ATLAS-Detektor am LHC bei* $\sqrt{s} = 7\,\text{TeV}$. Diploma thesis, Johannes Gutenberg-Universität Mainz, `http://butler.physik.uni-mainz.de/~huelsing/Diplomarbeit.pdf`, 01 2011. URL `http://butler.physik.uni-mainz.de/~huelsing/Diplomarbeit.pdf`.

[78] Andrea Banfi, Gavin P. Salam, and Giulia Zanderighi. Resummed event shapes at hadron - hadron colliders. *J. High Energy Phys.*, 08:062, 2004. doi: 10.1088/1126-6708/2004/08/062.

[79] G. Hanson et al. Evidence for jet structure in hadron production by e^+e^- annihilation. *Phys. Rev. Lett.*, 35:1609, 1975.

[80] J.D. Bjorken and Stanley J. Brodsky. Statistical Model for electron-Positron Annihilation Into Hadrons. *Phys. Rev. D*, 1:1416–1420, 1970. doi: 10.1103/PhysRevD.1.1416.

[81] G. Salam. Caesar: Computer automated resummations. Technical Report LPTHE-P04-05, LPTHE, 6 2004.

[82] ATLAS Collaboration. Luminosity Public Results. `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`. Accessed:2014/02/11.

[83] ATLAS Collaboration. New ATLAS event generator tunes to 2010 data. ATL-PHYS-PUB-2011-008, 2011. URL `http://cdsweb.cern.ch/record/1345343`.

[84] A. Sherstnev and R.S. Thorne. Parton Distributions for LO Generators. *Eur. Phys. J. C*, 55:553–575, 2008. doi: 10.1140/epjc/s10052-008-0610-x.

[85] ATLAS Collaboration. ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11. ATL-PHYS-PUB-2011-009, 2011. URL `http://cdsweb.cern.ch/record/1363300`.

[86] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *J. High Energy Phys.*, 07:012, 2002.

[87] R. Field. Min-bias and the underlying event at the Tevatron and the LHC. A talk presented at the Fermilab MC Tuning Workshop, Fermilab, Oct 2002.

[88] H.L. Lai et al. Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. *Eur. Phys. J. C*, 12:375–392, 2000. doi: 10.1007/s100529900196.

[89] Rick Field. Min-Bias and the Underlying Event at the LHC. *Acta Phys. Polon. B*, 42:2631–2656, 2011. doi: 10.5506/APhysPolB.42.2631.

[90] ATLAS Collaboration. Further ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11. ATL-PHYS-PUB-2011-014, 2011. URL `http://cdsweb.cern.ch/record/1400677`.

[91] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J. C*, 63:189–285, 2009. doi: 10.1140/epjc/s10052-009-1072-5.

[92] S. Gieseke, C.A. Rohr, and A. Siodmok. Multiple Partonic Interaction Developments in Herwig++. 2011.

[93] ATLAS Collaboration. Measurement of charged-particle event shape variables in $\sqrt{s} = 7$ TeV proton-proton interactions with the ATLAS detector. *Phys.Rev.*, D88: 032004, 2013. doi: 10.1103/PhysRevD.88.032004.

[94] G. D'Agostini. A Multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Meth. A*, 362:487–498, 1995. doi: 10.1016/0168-9002(95)00274-X.

[95] ATLAS Collaboration. Measurement of inclusive two-particle angular correlations in pp collisions with the ATLAS detector at the LHC. *JHEP*, 1205:157, 2012. doi: 10.1007/JHEP05(2012)157.

[96] ALICE Collaboration. Transverse sphericity of primary charged particles in minimum bias proton-proton collisions at $\sqrt{s} = 0.9$, 2.76 and 7 TeV. *Eur. Phys. J. C*, 72:2124, 2012. doi: 10.1140/epjc/s10052-012-2124-9.

[97] ATLAS Collaboration. Measurement of charged-particle event shape variables in sqrt(s) = 7 tev proton-proton interactions with the atlas detector - reaction database full record display. URL `http://hepdata.cedar.ac.uk/view/ins1124167`.

[98] ATLAS Collaboration. Measurement of dijet cross sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector. 2013.

[99] ATLAS Collaboration. Measurement of inclusive jet and dijet cross sections in proton-proton collision data at 7 TeV centre-of-mass energy using the ATLAS detector. 2011.

[100] Christopher Meyer and Mark Oreglia. *Measurement of Dijet Cross Sections in Proton-Proton Collisions at 7 TeV Center-of-Mass Energy Using the ATLAS Detector*. PhD thesis, Chicago U., 2013. Presented 2013.

[101] Bogdan Malaescu. An Iterative, Dynamically Stabilized (IDS) Method of Data Unfolding. (arXiv:1106.3107. CERN-PH-EP-2011-111):5 p, Jul 2011. Contribution to the proceedings of PHYSTAT2011. To be published in a CERN yellow report.

[102] Georges Aad et al. Single hadron response measurement and calorimeter jet energy scale uncertainty with the ATLAS detector at the LHC. *Eur.Phys.J.*, C73: 2305, 2013. doi: 10.1140/epjc/s10052-013-2305-1.

[103] ATLAS Collaboration. Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector. *Eur.Phys.J.*, C73:2306, 2013. doi: 10.1140/epjc/s10052-013-2306-0.

[104] Caterina Doglioni. High-pt jets in masked tile calorimeter modules. 09 2013. URL https://indico.cern.ch/event/252704/session/0/contribution/39.

[105] Stefan Dittmaier, Alexander Huss, and Christian Speckner. Weak radiative corrections to dijet production at hadron colliders. *JHEP*, 1211:095, 2012. doi: 10.1007/JHEP11(2012)095.