Quality of Jet Measurements and Impact on a Search for New Physics at CMS

Dissertation

zur Erlangung des Doktorgrades

des Fachbereichs Physik

der Universität Hamburg

vorgelegt von

Matthias Schröder

aus Buxtehude

Hamburg

2012

Gutachter der Dissertation: Prof. Dr. Peter Schleper
Dr. Isabell-Alissandra Melzer-Pellmann
Prof. Dr. Volker Büscher

Gutachter der Disputation: Prof. Dr. Robert Klanner
Prof. Dr. Christian Sander

Datum der Disputation: 13. Juli 2012

Vorsitzender des Prüfungsausschusses: Dr. Georg Steinbrück

Vorsitzender des Promotionsausschusses: Prof. Dr. Peter Hauschildt

Leiterin des Fachbereichs Physik: Prof. Dr. Daniela Pfannkuche

Dekan der Fakultät für Mathematik,
Informatik und Naturwissenschaften: Prof. Dr. Heinrich Graener

Abstract

The search for new physics beyond the Standard Model is one of the primary enterprises in particle physics today. Many models predict the production of new, coloured particles at the LHC. These particles would eventually decay into quarks and gluons, which manifest as particle jets in the detector. Hence, a precise understanding of the properties of jets is essential to study the underlying hard interaction process of hadronic final-states in order to probe the Standard Model and enter the territory beyond.

In this thesis, a measurement of the jet transverse-momentum (p_{T}) response function at the CMS experiment at the CERN LHC is presented, and its impact on a QCD-background estimation method in a search for new physics is described.

The jet-p_{T} response function is determined in two steps from dijet data collected in proton-proton collisions at a centre-of-mass energy of 7 TeV and corresponding to an integrated luminosity of up to $4.90\,\mathrm{fb}^{-1}$. The jets are clustered with the anti-k_T algorithm from individually measured particles that are reconstructed with the Particle-Flow algorithm combining the information from all relevant subdetectors. Firstly, the relative jet-p_{T} resolution σ/p_{T} is measured using an unbinned maximum-likelihood approach that is based on the p_{T} balance in QCD-dijet events and incorporates the jet-p_{T} spectrum. In the central detector region, the relative resolution in data is approximately 10% at $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ and improves to 5.5% at 600 GeV. It is approximately 5% larger on average than in the simulation. Secondly, the tails of the jet-p_{T} response function are measured studying the dijet p_{T}-asymmetry. The tails are caused by semileptonically decaying heavy-flavour quarks as well as instrumental effects and contribute at the percent level to the response function. Their relative size in data is larger than in the simulation by 20 – 60% in the central detector region at medium to high p_{T}.

Furthermore, a search for new physics in a multijet final-state with no leptons and a large p_{T} imbalance is reviewed, which has been conducted by CMS with $4.98\,\mathrm{fb}^{-1}$ of data. The QCD-multijet background is estimated from data in a control region by modelling the jet-p_{T} mismeasurements using the response function determined from dijet events. Taking into account the uncertainties on the response-function measurement, a prediction of the QCD background in the search region with a typical uncertainty of 60 – 70% is achieved. No excess in the number of events has been observed. Derived exclusion limits on the parameters of the Constrained Minimal Supersymmetric Standard Model correspond to a lower limit on equal squark and gluino masses of approximately 1300 GeV.

Kurzfassung

Die Suche nach neuer Physik jenseits des Standardmodells ist eines der Hauptvorhaben der heutigen Teilchenphysik. Viele Modelle sagen die Produktion neuer, farbgeladener Teilchen am LHC voraus. Diese Teilchen würden letztendlich zu Quarks und Gluonen zerfallen, die als Teilchenjets im Detektor sichtbar werden. Aus diesem Grund ist ein präzises Verständnis der Jeteigenschaften unerlässlich, um den harten Wechselwirkungsprozess, der den hadronischen Endzuständen zugrunde liegt, zu studieren und somit das Standardmodell und darüber hinausgehende Bereiche zu untersuchen.

In dieser Arbeit wird eine Messung der Response-Funktion für den Transversalimpuls (p_{T}) von Jets am CMS-Experiment am CERN LHC präsentiert, und deren Bedeutung für eine Methode zur Bestimmung des QCD-Untergrundes bei einer Suche nach neuer Physik beschrieben.

Die Jet-p_{T} Response-Funktion wird in zwei Schritten aus Zweijetereignissen bestimmt, die aus Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von 7 TeV ausgewählt wurden und einer integrierten Luminosität von bis zu $4.90\,\mathrm{fb}^{-1}$ entsprechen. Dazu werden Jets mit dem anti-k_T-Algorithmus aus individuell gemessenen Teilchen rekonstruiert, die zuvor mit dem Particle-Flow-Algorithmus durch Kombination der Informationen aller relevanten Subdetektoren bestimmt wurden. Im ersten Schritt wird die relative Jet-p_{T} Auflösung σ/p_{T} mit Hilfe einer ungebinnten Maximum-Likelihood-Methode gemessen, die auf dem p_{T}-Gleichgewicht in QCD-Zweijetereignissen beruht, und das Jet-p_{T}-Spektrum berücksichtigt. Die relative Auflösung in Daten beträgt im Zentralbereich des Detektors etwa 10% für $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ und verbessert sich zu 5.5% bei 600 GeV. Sie ist im Mittel etwa 5% größer als in der Simulation. Im zweiten Schritt werden die Ausläufer der Jet-p_{T} Response-Funktion durch Untersuchung der Zweijet-p_{T}-Asymmetrie gemessen. Die Ausläufer entstehen durch semileptonische Zerfälle schwerer Quarks sowie durch instrumentelle Effekte und tragen im Prozentbereich zur Response-Funktion bei. Ihr relativer Beitrag in Daten ist etwa 20 – 60% größer als in der Simulation bei mittleren bis hohen p_{T} im Zentralbereich des Detektors.

Desweiteren wird eine Suche nach neuer Physik in Endzuständen mit mehreren Jets, keinen Leptonen und einem starken p_{T}-Ungleichgewicht vorgestellt, die mit einer Datenmenge entsprechend $4.98\,\mathrm{fb}^{-1}$ von CMS durchgeführt wurde. Der QCD-Untergrund wird aus den Daten in einer Kontrollregion gemessen, indem die Jet-p_{T}-Fehlmessungen mit Hilfe der in Zweijetereignissen bestimmten Response-Funktion modelliert werden. Unter Berücksichtigung der Unsicherheit auf die Messung der Response-Funktion wird eine Vorhersage des QCD-Untergrunds in den Suchregionen mit Unsicherheiten von typisch 60 – 70% erreicht. Es wurde kein Überschuss an Ereignissen beobachtet. Deshalb wurden Ausschlussgrenzen für die Parameter des Constrained Minimal Supersymmetric Standard Model bestimmt, die unteren Grenzen von 1300 GeV entsprechen bei gleichen Squark- und Gluinomassen.

Contents

1 Introduction

> *Plötzlich, eines Tages, kam die Reihe an Dich, Dir anzusehen, was der Herr erschaffen hatte.*
>
> Jostein Gaarder

Human's endeavour for knowledge about the surrounding world has led to deep insight into the mechanisms in nature and allowed a glimpse on many astonishing features of the universe. This has only been possible by seeking to understand the fundamental constituents of matter and their interactions. It has been the accomplishment of particle physics to achieve this with unprecedented precision within the framework of the Standard Model (SM), which combines the theory of electroweak interactions with Quantum Chromodynamics (QCD). Numerous measurements at energy scales from a few eV to several TeV are reproduced by the SM, and many of its predictions, e. g. the existence of the W and Z bosons, have been found to be realised in nature. By now, only the source of electroweak symmetry breaking, which in the SM is attributed to the Higgs mechanism, has not been verified experimentally.

Despite its success, the SM is still a very incomplete description of nature for several reasons. First of all, gravity, which governs the large-scale structures in the universe, could so far not be included consistently into the SM framework. Although gravitational interactions are perfectly negligible at the subatomic level, this is certainly unacceptable for a fundamental theory. Even more, Dark Matter and Dark Energy, which together constitute 96% of the total energy in the universe, lack candidates in the SM. These and other observational facts as well as theoretical considerations, such as the apparent need for fine tuning of radiative corrections to the Higgs-boson mass, demonstrate that the SM has to be understood as a low-energy approximation of some more fundamental theory.

Numerous extensions to the SM have been proposed addressing the mentioned shortcomings. One of the most promising of these extensions is the concept of Supersymmetry (SUSY), which relates fermionic and bosonic degrees of freedom. As a consequence, each SM particle obtains a supersymmetric partner the spin of which differs by 1/2. Among the exciting features of SUSY is the fact that several SUSY models naturally contain Dark Matter candidates and that contributions from the new particles can stabilise the Higgs-boson mass with much less fine tuning required than in the SM. Furthermore, the strengths of the fundamental interactions become unified at large energies. Since SUSY particles with the same mass as their SM partners have not been observed, SUSY has to be a broken symmetry.

Several theoretical arguments suggest that the SUSY particles might have masses of the order of 1 TeV. Hence, they might well be in reach of the Large Hadron Collider (LHC), the most powerful and advanced particle accelerator today, which provides proton-

proton collisions at a centre-of-mass energy $\sqrt{s}$ of by now 8 TeV. Two large, multipurpose detectors, ATLAS and CMS, have been constructed to measure the collision products. Their primary mission is to discover the mechanism of electroweak symmetry breaking and to search for signals from possible new physics beyond the SM.

At the LHC, new particles are expected to be produced predominantly in strong reactions leading to quarks and gluons in the final state, which, due to the structure of QCD, hadronise and form particle jets. They are usually accompanied by an imbalance in the measured, total transverse momentum in the event because energy is carried away by weakly-interacting particles which do not deposit signals in the detector, such as possible Dark Matter candidates. However, SM processes resembling this signature occur at a rate larger by orders of magnitude and thus form a huge background to any search for new physics. For example, a large imbalance in the measured transverse momentum can be caused by the neutrinos in $Z \to \nu\bar{\nu}$ + jets, $t\bar{t}$ +jets, or W + jets events. Furthermore, QCD-multijet events where one or more jets are severely mismeasured constitute a background.

A major challenge of the LHC experiments is therefore to measure the final-state objects with high precision in the dense environment of proton-proton collisions. CMS exploits the Particle-Flow (PF) approach which combines the information from all subdetectors to reconstruct individual particles, from which jets are clustered. PF jets are complex objects, which depend on the underlying QCD processes as well as the performance of several different detector systems. A precise understanding of their properties is a necessity in order to probe the SM at the unprecedented energies of the LHC and thus be able to separate interesting events of possible new physics from background processes.

In this thesis, a measurement of the transverse-momentum (p_{T}) response function of PF jets at CMS is presented, which is performed with dijet data collected during the 2011 LHC-run at $\sqrt{s} = 7$ TeV corresponding to an integrated luminosity of up to $4.90\,\mathrm{fb}^{-1}$. Due to the large production cross-section for QCD-multijet events, the measurement has a high reach in p_{T} above 1 TeV. Firstly, an unbinned maximum-likelihood method is developed to fit the Gaussian-like core of the response function based on the assumption of p_{T} balance in QCD-dijet events. The likelihood incorporates an estimate for the jet-p_{T} spectrum. The response function's width, the relative jet-p_{T} resolution, is a fundamental characteristic of measured jets. Hence, many physics analyses with jet final-states, such as QCD or $t\bar{t}$ production cross-section measurements but particularly also new-physics searches, require knowledge of the resolution. Secondly, the non-Gaussian tails of the response function are determined by investigating the tails of the dijet p_{T}-asymmetry distribution. The tails are caused by mismeasurements due to physics effects such as semi-leptonically decaying heavy-flavour quarks inside jets because the neutrino escapes undetected and thus carries away energy as well as by instrumental effects such as inactive channels or miscalibration. Although the tail contributes only at the order of $10^{-3} - 10^{-2}$ to the total distribution, it causes severe mismeasurements which can result in a large p_{T} imbalance in otherwise balanced QCD-multijet events. The results of both the core and the tail measurements are reported as data-to-simulation ratios, which are then used to adjust the simulation to the data. The corrected response functions are provided to the CMS collaboration and employed in numerous physics analyses.

Finally, a search for new physics in events with three or more jets, a large transverse-momentum imbalance, and no leptons in the final state is discussed. It has been performed with data collected by CMS in 2011 at $\sqrt{s} = 7$ TeV corresponding to an integrated luminosity of $4.98\,\mathrm{fb}^{-1}$. The analysis is designed as a model-independent search. A key feature is the prediction of the kinematic properties of the expected SM-background events directly from data. Here, the focus is put on the QCD-background contribution because it is determined from multijet data by modelling the p_T mismeasurements using the aforementioned response function. No significant excess in the number of events is observed above the SM expectations. The result is used to constrain possible SUSY models; specifically, exclusion limits are derived on the parameters of the Constrained Minimal Supersymmetric Standard Model. Furthermore, the results are interpreted within a framework of generic models which assume new massive particles decaying to one or two jets and a stable weakly-interacting particle.

This thesis is organised as follows. The SM and suggested extensions, in particular SUSY, are reviewed in Chapter 2. In Chapter 3, the LHC and the CMS experiment are described. Then, in Chapter 4, the production of jets and their reconstruction at CMS are discussed, and important properties such as their response and resolution are investigated using simulated events. The jet-p_T resolution and response-tail measurements in dijet events are explained in Chapters 5 and 6, respectively. Their results are applied in the QCD-background prediction in the mentioned new-physics search in multijet final-states, which is presented in Chapter 7. Chapter 8 summarises this thesis.

2 The Standard Model and Possible Extensions

This chapter provides an introduction to the theoretical background in particle physics relevant for this thesis. In Section 2.1, the *Standard Model* (SM) of particle physics is presented, which comprises the most advanced description of the fundamental constituents and interactions in nature based on fundamental symmetry principles and only a limited number of parameters. The SM has been remarkably successful and is experimentally verified to an extremely high precision to the order 10^{-3} over a wide range of energies from a few eV to the TeV scale. However, there are a number of observations in tension to its predictions, as discussed in Section 2.2. For example, neither can gravitational interactions be described in the SM, nor does it include candidates for *Dark Matter* or *Dark Energy*, which are assumed to constitute most of the energy density in the universe. Furthermore, radiative corrections to certain processes become divergent at high energies. This and other theoretical considerations clearly point to the limited validity of the SM. Therefore, possible extensions to the SM which can provide solutions to several of the above problems are outlined. In particular, the concept of *Supersymmetry* is introduced and discussed in greater detail in Section 2.3.

2.1 The Standard Model of Particle Physics

Already about 2 400 years ago, Democritus and Leucippus hypothesised that everything in the universe is composed of indivisible units ('átomos') [1], a concept still present in the modern picture of nature. The SM of particle physics describes *quarks* ('q') and *leptons* as fundamental constituents of matter. They can be interpreted as structureless fermions, i. e. particles with spin 1/2, with further quantum numbers such as charge. Their dynamics are completely determined by a small set of fundamental interactions, from which — at least in principle — all other laws of physics[1] can be derived. In the SM, three fundamental interactions are described: the *electromagnetic*, the *weak*, and the *strong interaction*. These interactions, i. e. the transfer of energy and momentum as well as the alteration of quantum numbers, are mediated by the exchange of other elementary particles, bosons, which carry spin 1. The strength of the interactions depends on the fermion charges, which act as *coupling constants* and determine the probability of a fermion to emit or absorb a boson.

Mathematically, the SM is based on quantum field theories, which incorporate special relativity and quantum mechanics. The matter fermions are represented by states of quantised spinor fields, the exchange bosons by states of quantised vector fields in Fock space. All information about a system is encoded in the *Lagrangian (density)* $\mathcal{L}$, a scalar function of the fields. Much in analogy to classical mechanics, the equations of motion of the fields can be obtained from $\mathcal{L}$ when assuming the principle of stationary action $\delta A = 0$, where the action A is defined as

$$A = \int \mathrm{d}t \, \mathcal{L} \, ,$$

---

[1] This does not hold true for gravity-related processes since gravity cannot be incorporated consistently into the SM.

with t denoting time. The resulting equations of motion are the Dirac equation for fermions and the Klein-Gordon equation for bosons. Astonishingly, the interaction terms of the SM are introduced naturally when assuming local invariance of $\mathcal{L}$ under certain unitary transformations (*gauge transformations*).

A phenomenological introduction to the SM is provided in Section 2.1.1, followed by a review of the underlying concept of gauge symmetry in Section 2.1.2. The computation of observables from the SM Lagrangian by means of perturbation theory is discussed in Section 2.1.3. Finally, in Section 2.1.4, the generation of particle masses via gauge-symmetry breaking is described. A detailed introduction to the SM and gauge symmetries is given for instance in [2–4], an introduction to quantum field theory in general in [5]. The properties of the SM particles and related experimental results are reviewed in [6].

2.1.1 Phenomenological Overview

In the SM, twelve different fermions and their corresponding antiparticles are described, which are identical except for the sign of their quantum numbers. They are grouped into three generations and, for reasons that will become clear later, each generation is subdivided further into two leptons and two quarks, cf. Table 2.1.1.

Table 2.1: Fermions and gauge bosons in the SM. Fermion indices 'L' and 'R' denote states of left and right chirality, respectively. Curly brackets indicate to which fermions the particular gauge bosons couple, where the $W^{\pm}$ bosons only couple to the left-chiral states.

	fermions			gauge bosons	
	1. generation	2. generation	3. generation	electroweak	strong
leptons	$\begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}, e_R$	$\begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix}, \mu_R$	$\begin{pmatrix} \nu_{\tau,L} \\ \tau_L \end{pmatrix}, \tau_R$	$\left.\right\} W^{\pm}, Z \left.\right\} \gamma$	
quarks	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}, u_R, d_R$	$\begin{pmatrix} s_L \\ c_L \end{pmatrix}, s_R, c_R$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}, b_R, b_R$		$\left.\right\} g$

The leptons of the first generation are the *electron* ('e') and the *electron neutrino* ('ν_e'), while the corresponding quarks are denoted *up* ('u') and *down* ('d') quark. In a somewhat simplified picture, all ordinary stable matter in the universe consists of electrons and the quarks of the first generation: protons ('p') and neutrons ('n'), the constituents of atomic nuclei, can be considered compounds of u and d quarks and the atomic shell is formed by electrons.

The particles of the other generations have identical properties to their first-generation counterparts except for their masses: the leptons of the second and third generation are the *muon* ('μ') and the *tauon* ('τ') as well as the corresponding neutrinos 'ν_μ' and 'ν_τ', while the quarks are denoted *charm* ('c') and *strange* ('s') as well as *top* ('t') and *bottom* ('b'). While the electron mass[2] amounts to $510.998910 \pm 0.000013\,\mathrm{keV}$ and the u

---

[2] Throughout this thesis, the system of natural units common in particle physics is used, where $\hbar = c = 1$. Hence, masses, momenta, and energies all have the unit of energy.

and d quark masses to a few MeV, the heaviest SM fermion, the t quark, has a mass of $172.0 \pm 0.9 \pm 1.3\,\mathrm{GeV}$. Neutrinos, in contrast, have masses of less than a few eV [6]. The large mass hierarchy is quite intriguing and not understood within the SM.

The electromagnetic interaction can be described by the theory of *Quantum Electrodynamics* (QED), which has been developed to its final form by the 1950s. It is mediated between objects with electric charge by the exchange of *photons* ('γ'), which are themselves electrically neutral. For example, two electrons can scatter via the exchange of a photon, or they can annihilate into a photon (*annihilation*) which in turn again produces an electron and its antiparticle, a *positron* (*pair production*), cf. Fig. 2.1. The electron-type leptons (e, μ, τ) have an electric charge of -1 in units of the elementary charge, while the up-type quarks (u, c, t) have a charge of $+2/3$ and the down-type quarks (d, s, b) of $-1/3$. Since photons are massless and electrically neutral, the electromagnetic interaction is long ranged. The concepts of QED have served as a basis for further quantum field theories describing other interactions.

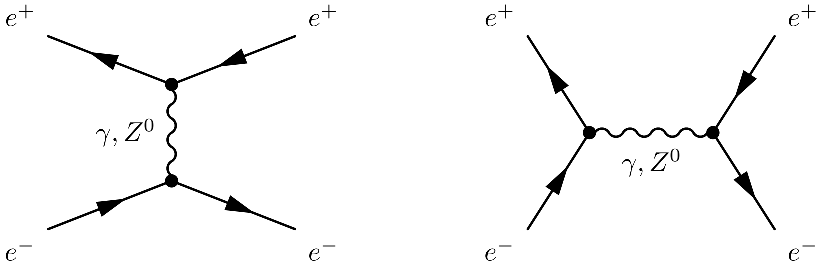


Figure 2.1: Leading-order contributions to the $e^+e^- \to e^+e^-$ process (*Bhabha scattering*): scattering (*left*) and annihilation and pair production (*right*).

It is worthwhile to point out that the electromagnetic interaction in combination with the fermionic nature of the electron, which forms the atomic shell, holds responsible for the entire structure and dynamics of atoms and molecules. A precise understanding of the binding processes has become possible within the framework of QED, like for example the fine-structure splitting of atomic orbitals. In fact, most phenomena acting at the macroscopic scale, like mechanical forces or electric processes, are direct consequences of the electromagnetic interaction.

All fermions in the SM carry a *weak charge* ('g'), the weak interaction is mediated by the exchange of the 'Z' and '$W^\pm$' bosons. In contrast to the photon, the weak bosons are massive: the Z has a mass of about $91.1876 \pm 0.0021\,\mathrm{GeV}$, the $W^\pm$ of $80.399 \pm 0.023\,\mathrm{GeV}$ [6]. As a consequence, the weak interaction is suppressed compared to the electromagnetic interaction below the $Z/W^\pm$-mass scale. Apart from this and the fact that it couples to weak charges, the Z boson behaves in many aspects like the photon and may mediate fermion scattering or pair annihilation and production, cf. Fig. 2.1. The $W^\pm$ bosons, however, have very different features. They carry electric and weak charge, and they alter the type of fermions when coupling to them. An important example is the radioactive β^--decay, where a neutron decays into a proton, an electron, and an anti electron-neutrino. Historically, the investigation of this process lead to Fermi's theory of the β^--decay, an effective theory of the weak interaction [7]. In the modern picture, the β^--decay is attributed to a d quark in the neutron turning into a u quark under emission of a W^-, which decays leptonically as $W^- \to e^-\bar{\nu}_e$. It is via the exchange of $W^\pm$ bosons that the heavier particles of the second and third generation decay into their lighter first-generation counterparts.

It has been found experimentally that the weak interaction violates parity, i. e. does not behave symmetric under spatial point-reflections: fermions from $W^{\pm}$ decays are left-handed by a fraction v/c, where v denotes the velocity of the particle. Here, 'left-' and 'right-handed' are understood as different states of *helicity*, the projection of the spin onto the momentum. The effect has been first observed by Wu in the decay of $^{60}_{27}$Co isotopes [8]. This behaviour can be described in an elegant way in terms of *chirality*: as for helicity, each fermion state can be expanded into two different states of chirality. For massless particles, they correspond to the helicity, which is why they are — somewhat confusingly — also denoted left ('L') and right ('R'). The $W^{\pm}$ bosons couple only to L states (of fermions and R states of antifermions), and the Z boson couples with different strength to the L and R states.

Changes of the flavour eigenstates of the weak interaction (eigenstates of fermions to interactions are denoted as *flavour*) occur to first approximation only within one generation. For example, the electron can transform into an electron neutrino via emission of a W^{-}. However, since the weak flavour-eigenstates of fermions are different to their mass eigenstates, the $W^{\pm}$ can also couple between fermions of different generations. In the quark sector, it is conventional to chose a representation in which the up-type flavour eigenstates correspond to the mass eigenstates and only the down-type states mix. Then, transitions within one generation, i. e. between the u and d mass eigenstates, have a higher probability than between different generations. The relative strength of the various possible transitions are summarised as the coefficients of the Cabbibo-Kobayashi-Maskawa (CKM) matrix [9, 10]. In the leptonic sector, inter-generation coupling is referred to as *neutrino oscillation* [11–13], and its observation implies that neutrinos are in fact massive [14–16].

The weak and the electromagnetic interactions can — and must in fact in order to ensure unitarity conservation — be described by a unified theory as different low-energy manifestations of the underlying *electroweak interaction*, which was described first by Salam, Glashow, and Weinberg [17, 18]. As a consequence, the strength of the electromagnetic and the weak interactions will become of equal size at energies larger than a few 100 GeV, as has been verified experimentally for example in $e^{+}e^{-}$ collisions at the PETRA [19] and LEP [20] experiments at DESY and CERN, respectively. The separation at lower energies is attributed to the spontaneous breaking of electroweak symmetry generated by the *Higgs* mechanism[3]. An essential result of the unified theory was the postulation of the Z boson and the associated *neutral current* reactions, which were discovered shortly afterwards in 1973 with the Gargamelle experiment at CERN [21, 22]. It received further spectacular confirmation through the actual discovery of the $W^{\pm}$ and Z bosons in $p\bar{p}$ collisions at the CERN SPS collider in 1983 [23, 24].

The $W^{\pm}$ bosons are themselves weakly and electrically charged, allowing for various interactions between the γ, Z, and $W^{\pm}$ bosons. In particular WW scattering is interesting because its probability becomes divergent for large momentum transfers without contributions from the exchange of an additional boson like e. g. the Higgs boson.

Finally, the strong interaction is described by the theory of *Quantum Chromodynamics* (QCD). It is mediated via the massless *gluons* between objects that carry *colour charge*. In the SM, these are the quarks and the gluons themselves, collectively called *partons*.

---

[3] Other mechanisms for symmetry breaking are possible, but the Higgs mechanism was assumed in the original formulation of the electroweak theory.

The colour charge is special in the sense that there exist three different linearly-independent states, commonly referred to as 'red', 'blue', and 'green', and their respective anti-states. The concept of three different colour degrees-of-freedom was motivated by the Pauli-principle because particles such as the Δ^{++} and Ω^- could only be described as bound states of three quarks of the same flavour and spin [25, 26]. Their existence was established among others by measurements of the ratio of the $\mu^+\mu^-$ to hadron production-rate in e^+e^- collisions, cf. Fig. 41.6 in [6]. In analogy to chromatics, a neutral or 'white' state consists of the combination of a red, a blue, and a green charge or their respective anticharges. Such colourless states of quarks bound by the strong interaction are termed *hadrons*. In a simplified picture, they can be viewed to consist of either three quarks of different colour (*baryons*) or two quarks of one colour and its anti-colour (*mesons*)[4]. The actual structure of hadrons is more involved, however. The three or two *valence quarks* constantly exchange gluons, which exchange gluons themselves or create virtual quark-antiquark (*sea quark*) pairs that annihilate again.

As a consequence of the number of colour states and the self-interaction of gluons, the dependence of the strong interaction on the distance is opposite to that of the other fundamental interactions: its strength increases with increasing distance. Hence, colour-charged objects cannot exist freely but, when separated, will generate new coloured particles until only colour-neutral states remain, a phenomenon known as *confinement*. This results in the typical dimension of hadrons of about 10^{-15} m in diameter. On the other hand, for small distances coloured particles can be considered free with respect to the strong interaction (*asymptotic freedom*), as Gross, Politzer, and Wilczek proved in 1973 [27, 28].

Importantly for this thesis, confinement leads to the creation of bunches of hadrons (*jets*) by coloured particles originating in high-energy particle collisions, as will be discussed later in Section 4. The observation of three-jet events in 1979 by the PETRA experiments [29–31] was the first experimental evidence for the existence of gluons. QCD has subsequently been probed in great detail for example in *deep-inelastic scattering* experiments at the *ep* collider HERA [32] at DESY.

2.1.2 The Standard Model Lagrangian and Gauge Interactions

The Lagrangian $\mathcal{L}$ of a free, massless fermion spinor ψ is given by

$$\mathcal{L} = \bar{\psi}\left(i\gamma_\mu\partial^\mu\right)\psi \ , \tag{2.1}$$

where the upper and lower Greek indices are understood to run from 0 to 3 and to be summed over according to the Minkowski metric. The γ_μ are the four linear-independent, traceless hermitian 4×4 matrices and ∂^μ denotes the space-time derivative.

In the following, unitary local transformations

$$\psi \to \psi' = U\psi, \quad U = \mathrm{e}^{ig\chi_a(x)T_a}$$

are considered, where the χ_a, $a \in \{0, n\}$, are scalar functions of the space-time coordinate x, and g is a dimensionless number that will be identified with a coupling constant. The T_a are linear-independent, hermitian matrices, which are called the *generators* of the

---

[4] Further combinations which are allowed by QCD, like for example three differently coloured quarks plus two quarks with colour and anti-colour (*pentaquarks*), have not been verified despite of a few claims [6].

transformation, and satisfy the commutation relations

$$[T_a, T_b] = i f_{abc} T_c$$

with the *structure constants* f_{abc}. The action remains unchanged under U if $\mathcal{L}$ is invariant up to a total derivative. This is the case if n new vector fields F_a^μ are introduced that transform like

$$F_a^\mu \to F_a^{\prime\mu} = F_a^\mu - \partial^\mu \chi_a(x) - g f_{abc} \chi_b(x) F_c^\mu$$

under U and if ∂^μ in Eq. (2.1) is replaced by the *covariant derivative*

$$D^\mu = \partial^\mu + i g T_a F_a^\mu \ .$$

The F_a^μ represent massless bosonic particles, *gauge bosons*, and hence, Eq. (2.1) becomes

$$\mathcal{L} = \bar{\psi} i \gamma_\mu \partial^\mu \psi - g \bar{\psi} \gamma_\mu T_a \psi F_a^\mu - \frac{1}{4} F_{a\mu\nu} F_a^{\mu\nu} \ . \tag{2.2}$$

In addition to the kinetic term of the fermion, there is now the second term containing both the fermion and the gauge boson fields. It is interpreted as an interaction between the fermion and the bosons with the coupling strength given by g. The last term, where

$$F_{a\mu\nu} F_a^{\mu\nu} = \left(\partial_\nu F_{a\mu} - \partial_\mu F_{a\nu}\right)\left(\partial^\nu F_a^\mu - \partial^\mu F_a^\nu\right) \ ,$$

represents the kinetic energy of the gauge bosons and is the maximally allowed term conform with the gauge invariance of $\mathcal{L}$. Depending on the structure f_{abc} of the transformation group, $F_{a\mu\nu} F_a^{\mu\nu}$ might also include self interactions of the gauge bosons.

The essence of the SM is that all fundamental interactions are a consequence of gauge invariance. QED follows from invariance under the $U(1)$ symmetry group with the electric charge as generator, while the relevant group for the weak interaction is $SU(2)$ with the three Pauli matrices as generators. However, the properties of both interactions are not fully covered by these separate theories but only by the combined $U(1)_Y \times SU(2)_L$ symmetry group. Invariance of $\mathcal{L}$ requires the introduction of one gauge field B^μ for $U(1)_Y$ and three gauge fields W_i^μ, $i \in \{1, 3\}$, for $SU(2)_L$. Since $SU(2)_L$ has non-trivial structure constants, the elements of the total antisymmetric tensor, there are also interactions between the W_i^μ. Moreover, $SU(2)_L$ transforms only the left-chiral parts of fermion states, illustrated by the index L, which leads to the observed parity-violating nature of the weak interaction. Therefore, the left-chiral fermion states are grouped into doublets in $SU(2)$ flavour-space (cf. Table 2.1.1), for example the leptonic states are

$$\begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}, \quad \nu_{e,R}, \quad e_R \ .$$

The interaction terms in $\mathcal{L}$ can be regrouped accordingly to the physical gauge bosons

$$W^{\mu\pm} = \frac{1}{\sqrt{2}} \left(W_1^\mu \pm W_2^\mu\right) \ ,$$

which act as ladder operators on the doublets. Hence, flavour-changes occur only between the two states in one $SU(2)_L$ doublet (which are different to the mass eigenstates). The physical γ and Z bosons are represented by a superposition of the W_3^μ and the B^μ, corresponding to a rotation in $SU(2)$ flavour-space by the *Weinberg angle* $\sin^2\theta_W = 0.23116 \pm 0.00013$ [6]. As a consequence, the electromagnetic and the weak coupling constants are related by

$$e = g\sin\theta_W \ ,$$

and both the Z and the photon couple to the left- and right-chiral states, but with different strength in case of the Z boson.

QCD finally follows from invariance under $SU(3)_C$ transformations between the three colour states. The generators of the group are the eight Gell-Mann matrices. Hence, there are eight gauge fields corresponding to the gluons. The non-Abelian $SU(3)_C$ structure leads to self interaction of gluons and the discussed phenomenon of confinement, cf. also Section 2.1.3.

In summary, the postulation of local gauge invariance requires the introduction of interaction terms into the Lagrangian. Hence, most remarkably, the dynamics in the SM arise from symmetry under $U(1)_Y \times SU(2)_L \times SU(3)_C$ transformations.

2.1.3 Perturbation Theory and Renormalisation

The equations of motion derived from the SM Lagrangian can in general not be solved analytically for non-trivial systems of interacting particles. It is possible, however, to expand the solution in orders of the coupling constants [2]. For example, the state $|\psi\rangle$ of a system $|\psi_i\rangle$ after some scattering process S may be written as

$$|\psi\rangle = S\,|\psi_i\rangle = \left(\sum_{n=0}^{\infty} g^n C_n\right)|\psi_i\rangle \ . \tag{2.3}$$

For illustration purposes, only one interaction between charges g has been assumed with a coupling constant $\alpha \propto g^2$. Each coefficient C_n in this *perturbation series* can be associated with a distinct physical process contributing to the total interaction. C_0 corresponds to no interaction, i. e. the initial state, and C_1 is zero due to energy and momentum conservation. The first non-trivial (*leading order*, LO) contribution arises from C_2 and is related to the $2 \to 2$ *processes* depicted in Fig. 2.1, which contain two vertices with coupling $\sqrt{\alpha}$. C_3 incorporates *next-to-leading order* (NLO) processes with three vertices such as the diagram shown in Fig. 2.2 (*left*). Physical observables for some final state $|\psi_f\rangle$ are obtained from the matrix element

$$\mathcal{M}_{fi} = \langle\psi_f|\psi\rangle = \langle\psi_f|S|\psi_i\rangle \ ;$$

For example, the cross section σ for obtaining $|\psi_f\rangle$ from $|\psi_i\rangle$ is given by

$$\sigma \propto |\mathcal{M}_{fi}|^2 \ ,$$

where the constant of proportionality contains kinematic phase-space factors.

The C_n can in principle be computed following the *Feynman rules.* Only the first few orders are usually known, though, due to the complex calculations and the rapidly increasing number of diagrams per order. However, depending on the size of the expansion coefficients and on the precision of the experimental result in comparison, it is often sufficient to know only the first few orders of an observable.

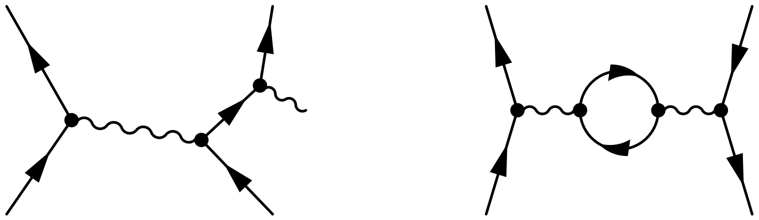


Figure 2.2: Higher-order contributions to the $2 \to 2$ process: final-state radiation (*left*) and virtual loop (*right*).

Conceptual difficulties arise from higher-order diagrams such as Fig. 2.2 (*right*), where the momenta of the virtual particles in the loop are not constraint, and the associated integrals are divergent. If all orders of the perturbation series are taken into account, the divergences cancel with exactly opposite contributions from other diagrams, but this is not the case for a finite number of orders, leading to infinite cross sections.

The situation can be circumvented by *renormalisation*, a technique which allows computation of finite observables to any order [33]. First, divergences are temporarily *regularised* by introducing a cut-off to the loop momenta, the dimensional *renormalisation scale* μ_R. Then, the free parameters of the Lagrangian, i. e. the coupling constant, are redefined (*renormalised*) such that the regularised divergences are included ('absorbed') into the definition of α. Hence, the coupling constant becomes a function of μ_R, i. e. $\alpha \to \alpha(\mu_R^2)$, and is now divergent (*running coupling*)[5]. However, physical observables that depend on α still have to be independent of μ_R since it is an arbitrary parameter. This is assured if α satisfies the *renormalisation group equation*

$$\mu_R^2 \frac{\partial \alpha}{\partial \mu_R^2} = \beta(\alpha) \ .$$

The *beta-function* $\beta(\alpha)$ can be calculated from the structure of the interaction and relates α at μ_R to its value at some reference scale μ. The reference value $\alpha(\mu^2)$ has to be determined experimentally, which reflects the fact that the coupling strength is a free parameter of the Lagrangian and cannot be predicted by the SM. The β function can be written again as expansion in α.

Hence, physical observables, such as the cross section, remain finite after renormalisation to any order in the perturbative expansion. Divergences only occur with the unphysical ('bare') parameters of the Lagrangian. It is also indicative to note that only local gauge-invariant theories are *renormalisable*, i. e. finite results at any order are ensured by introduction of a finite number of renormalised parameters. In case of gravity, for example, the number of renormalised parameters becomes infinite.

In practice, only the first few orders of β can be computed, and therefore, also physical observables become μ_R dependent (but finite). Among typical choices for μ_R is the momentum transfer Q of the investigated process, such that $\alpha(Q^2)$ corresponds to the effective coupling strength in that process.

---

[5] Besides Fig. 2.2 (*right*), there are other types of divergent diagrams that require renormalisation of mass parameters.

At LO, the renormalised electromagnetic coupling of QED with $\mu_R = Q$ is given by

$$\alpha_{\mathrm{em}}\left(Q^2\right) = \frac{\alpha_{\mathrm{em}}\left(\mu^2\right)}{1 - \frac{\alpha_{\mathrm{em}}(\mu^2)}{2\pi}\ln\left(\frac{Q^2}{\mu^2}\right)} \,. \tag{2.4}$$

The running of α_{em} is demonstrated in Fig. 2.3 (*left*). Evidently, the strength of the electromagnetic coupling increases for larger energies. The increase is slow, however, such that perturbation theory Eq. (2.3), which requires $\alpha_{\mathrm{em}} < 1$, is applicable in QED at any relevant energy scale [34][6]. The fine-structure constant $1/\alpha_0 = 137.03599911 \pm 0.00000046$ is a typical reference value for $\mu \to 0$; in comparison, the coupling strength increases to $1/\alpha_{\mathrm{em}}(m_Z^2) = 128.940 \pm 0.048$ at the Z-mass scale [35]. A common semi-classical interpretation of the running is that for example an electron creates a cloud of virtual e^+e^- pairs through loop processes, which screen its bare electric charge. High-Q^2 processes probe small distances around the electron. Hence, more of the bare charge is 'seen' [34].

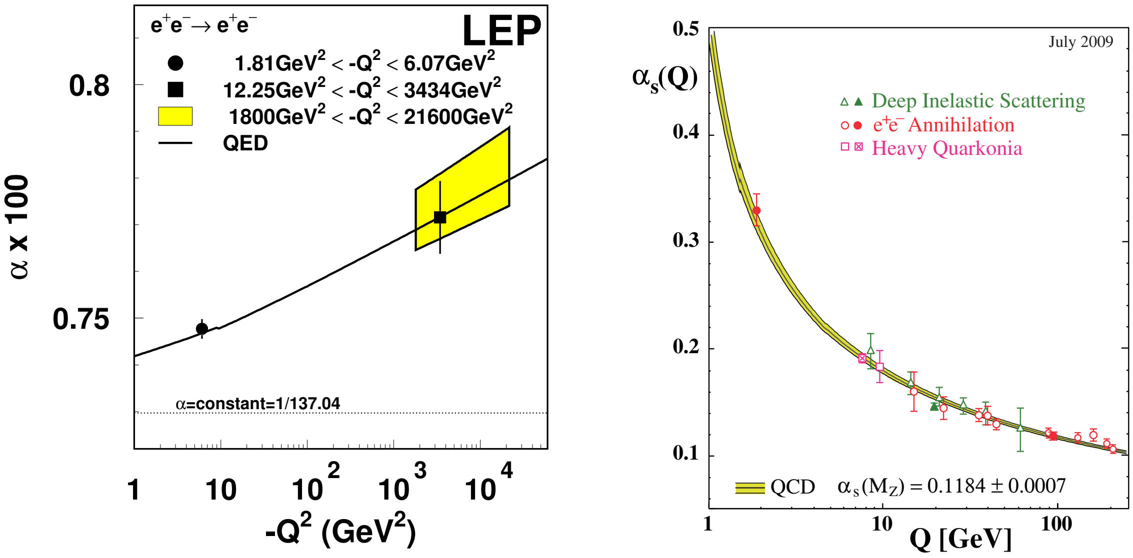


Figure 2.3: Dependence of the electromagnetic (*left*) and the strong (*right*) coupling on the energy scale Q. The predictions (*lines*) from QED and QCD, respectively, are compared to the values extracted from various measurements (*markers*) as indicated in the legends. Taken from [36] and [35].

The renormalised strong coupling of QCD, again for $\mu_R = Q$, is given at LO by

$$\alpha_{\mathrm{s}}\left(Q^2\right) = \frac{12\pi}{(33-2n_f)\ln\left(\frac{Q^2}{\Lambda^2_{\mathrm{QCD}}}\right)} \,, \quad \text{with } \Lambda^2_{\mathrm{QCD}} = \mu^2 \exp\left(\frac{-12\pi}{(33-2n_f)\,\alpha_{\mathrm{s}}(\mu^2)}\right) .$$

where $n_f \leq 6$ denotes the number of quark flavours contributing to the virtual loops, i. e. with $m_q^2 < Q^2$. The positive denominator, a consequence of $n_f \leq 6$ and the gluon self-

---

[6] This is the case also if further loop contributions are considered.

interaction, leads to a scaling behaviour opposite to that in QED, namely a decrease of α_s with increasing Q^2, cf. Fig. 2.3 (*right*). Hence, asymptotic freedom occurs for $Q^2 \to \infty$, such that perturbative expansions as in Eq. (2.3) are valid[7]. In contrast, α_s becomes large for $Q^2 \to \Lambda_\mathrm{QCD}$, which marks the regime where perturbation theory is not applicable anymore. With $\alpha_\mathrm{s}(m_Z^2) = 0.1184 \pm 0.0007$ and $n_f = 5$, Λ_QCD is of the order of 100 MeV [35].

2.1.4 Electroweak Symmetry Breaking

The particles of the SM are in general not massless. However, it is not possible to write explicit mass terms into the Lagrangian that are compatible with gauge invariance. More precisely, mass terms for gauge bosons directly violate the invariance. Fermion mass terms are possible as long as the masses in one multiplet of the symmetry group are the same and the interaction is symmetric under chirality. Both is not the case for the $SU(2)_L$ doublets. Hence, in its purely gauge-symmetric formulation, the electroweak theory fails to correctly describe the experimental facts.

The classical solution to this problem is the *Higgs mechanism* [37–39], which leaves the Lagrangian but not the vacuum state invariant under electroweak transformations, a principle called *spontaneous symmetry breaking*. Reviews of the Higgs mechanism in context of the SM may be found in [2, 4, 40]. In its simplest structure, an $SU(2)$ doublet (the *Higgs field*)

$$\Phi = \begin{pmatrix} \Phi^+ \\ \Phi^0 \end{pmatrix} ,$$

is introduced, which has a charged and a neutral complex scalar component. The corresponding contribution

$$\mathcal{L} = (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi) \tag{2.5}$$

to the Lagrangian is invariant under $U(1)_Y \times SU(2)_L$ transformations. The D^μ are again the covariant derivatives and V is the *Higgs potential*

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda \left(\Phi^\dagger \Phi\right)^2 ,$$

with the parameters λ real and positive and $\mu^2 < 0$. V has degenerated, non-trivial minima Φ_0 defined by $\Phi_0^\dagger \Phi_0 = -\mu^2/2\lambda$, each having the non-vanishing vacuum expectation value

$$\langle 0 | \Phi | 0 \rangle = v = \sqrt{\frac{-\mu^2}{2\lambda}} .$$

The choice of any particular minimum (*vacuum state*) breaks the $U(1)_Y \times SU(2)_L$ symmetry of the system. The vacuum state is chosen electrically neutral as

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} .$$

It is quantised by expansion around the minimum resulting in one massive and three massless bosons, the latter being *Goldstone bosons* [41]. By choice of a suitable gauge, the

---

[7] Again, this result holds also for higher loop contributions.

massless Goldstone boson fields are eliminated and the Higgs field becomes

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.6}$$

with the massive Higgs boson H. Substitution of Eq. (2.6) into Eq. (2.5) results in mass terms[8]

$$\begin{aligned} m_H &= \sqrt{2}\mu \\ m_W &= \frac{1}{2} g v \\ m_Z &= \frac{m_W}{\cos\theta_W} \ . \end{aligned}$$

By virtue of the chosen neutral vacuum state, the photon does not acquire mass. In addition, there are also terms describing interactions between the H and the $W^{\pm}$ and Z bosons with couplings proportional to the vector boson masses, as well as self-coupling terms of the H. The parameter v follows from the known m_W to $v = 246\,\mathrm{GeV}$; the parameter μ, and therefore the Higgs mass, is still unknown.

Also fermion masses can be generated via couplings between the Higgs boson and the fermions (*Yukawa coupling*). The corresponding term in the Lagrangian for the first generation fermions[9] has the form

$$\begin{aligned} \mathcal{L}_{\mathrm{Yukawa}} = &- G_e \left(\bar{e}_L \Phi e_R + \bar{e}_R \Phi^{\dagger} e_L \right) \\ &- G_d \left(\bar{q}_L \Phi d_R + \bar{d}_R \Phi^{\dagger} q_L \right) - G_u \left(\bar{q}_L \Phi_c u_R + \bar{u}_R \Phi_c^{\dagger} q_L \right) \ . \end{aligned} \tag{2.7}$$

Here, Φ_c denotes the charged conjugated Higgs field, e_L and q_L the left-handed lepton and quark $SU(2)$ doublets, and e_R, u_R, and d_R the corresponding right-handed singlet states. The G_i are new coupling constants. Eq. (2.7) includes fermion mass terms

$$m = G \frac{v}{\sqrt{2}} \tag{2.8}$$

as well as fermion-Higgs coupling terms, where the coupling strength is proportional to the fermion mass.

The Higgs boson has not been observed yet, and hence, the mechanism of electroweak symmetry breaking in the SM is not verified. There are, however, theoretical and experimental bounds on the Higgs mass. Given an otherwise divergent high-energy behaviour of the self-coupling parameter λ ('triviality problem'), m_H cannot be larger than a few hundred GeV, depending on the energy scale to which the SM is valid [44]. Besides indirect experimental results, e. g. from electroweak precision measurements at the Z pole [20], there are strong constraints from direct searches in collider experiments at LEP, Tevatron, and the LHC [42,43,45–48]. Currently, a SM Higgs boson with mass below 600 GeV is excluded at 95% confidence level, except in the mass ranges 117.5 – 118.5 GeV and

---

[8] When acquiring mass, the vector bosons obtain a longitudinal polarisation component. These additional degrees-of-freedom correspond to the eliminated Goldstone bosons.

[9] In the original formulation of the SM, neutrinos were considered massless. Therefore, there is no Yukawa-coupling term for neutrinos in $\mathcal{L}_{\mathrm{Yukawa}}$.

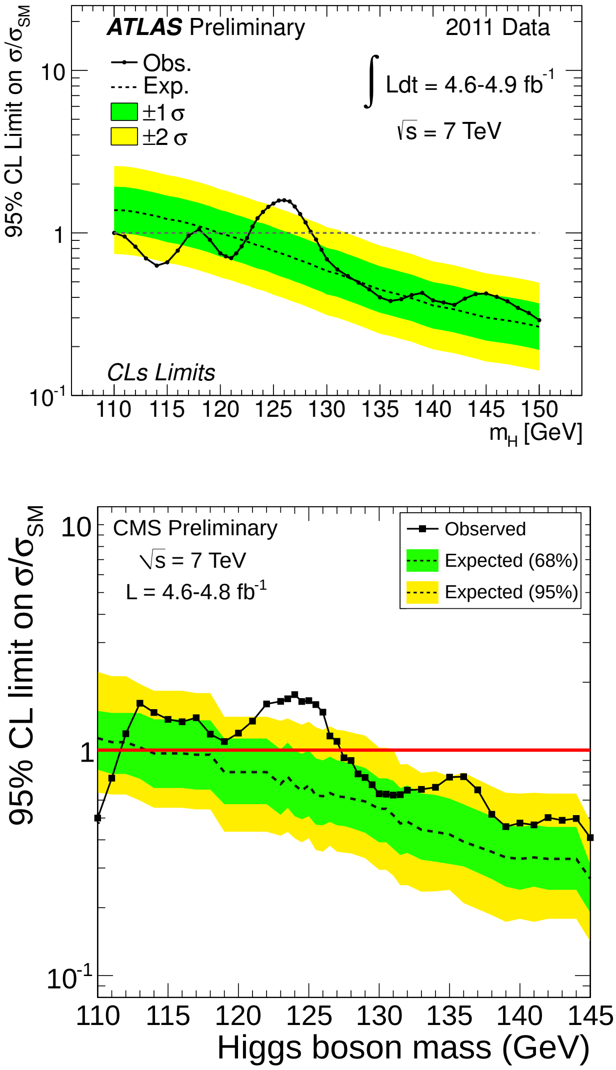


Figure 2.4: Observed (*solid line* and expected (*dashed line*) upper exclusion limit at 95% confidence level on the SM Higgs boson production cross-section relative to the SM expectation as a function of the Higgs mass published by ATLAS (*top*) and CMS (*bottom*). The intervals of 68 (*green band*) and 95% (*yellow band*) confidence level on the expected values are also shown. Taken from [42] and [43].

122.5 – 127.5 GeV. Interestingly, an excess of events above the SM expectations with local significance of about 2.5 σ has been observed by both ATLAS and CMS around 125 – 126 GeV [42, 43], cf. Fig. 2.4.

Considering the experimental sensitivity and foreseen rate of data taking at the LHC, it is likely that a definite answer about the realisation of the SM Higgs mechanism will be found this year. It should be pointed out, however, that the presented results have to be reinterpreted in the context of models beyond the SM, a few of which will be discussed below.

2.2 Beyond the Standard Model

Although remarkably successful, the SM is nevertheless incomplete for a number of reasons, apart from the aspect of the non-established mechanism of electroweak symmetry breaking, some of which are reviewed in Section 2.2.1. Therefore, the SM has to be understood as an effective, low energy model of a more fundamental theory. In Section 2.2.2, a brief overview to the most relevant models beyond the SM is provided. The discussion follows mostly [49].

2.2.1 Shortcomings of the Standard Model

One of the most striking experimental facts demonstrating the incompleteness of the SM is the existence of gravity. Though perfectly negligible compared to the other fundamental interactions at the energy scales described by the SM, it will become relevant at the latest at the Planck scale $\Lambda_\mathrm{P} \approx 10^{19}$ GeV. So far, however, it has not been possible to formulate a renormalisable theory of gravity [50].

Further evidence stems from astrophysical observations. Measurements of galactic rotation curves [51], mass distributions in galaxy clusters [52–54], as well as anisotropies in the cosmic microwave background (CMB) [55,56], imply the existence of Dark Matter (DM), a substance incompatible with the properties of the ordinary SM matter. Moreover, studies of distant type Ia supernovae [57,58] and again of the CMB [55,56] hint to an accelerated expansion of the universe attributed to some exotic form of energy, generally denoted Dark Energy (DE) [59]. Somewhat disturbingly, the DM and DE are inferred to constitute 96% of the total energy and matter content of the universe.

The apparent excess of matter over antimatter in the universe — provided there is no large-scale spatial separation into matter and antimatter regions — requires a mechanism to break the symmetry of SM interactions under charge and parity transformations (CP symmetry). Although some CP-violating contributions are generated by imaginary entries in the CKM matrix, this is not enough to explain the excess expected from current models of baryogenesis [60].

Apart from the listed experimental facts, several conceptual and aesthetic shortcomings of the SM exist.

There are 19 free parameters in the SM implying a lack of understanding of some underlying principles in nature. Moreover, the realised gauge groups and the particle content are completely arbitrary although for example the $Z \to \gamma\gamma$ cross-section remains finite only because there are exactly three generations of fermions with the given distribution of electric charges (*chiral anomaly*) [2].

Another flaw of the SM which receives much theoretical attention arises from radiative corrections to the particle mass-parameters in the Lagrangian due to fermion and boson loop contributions similar to Fig. 2.2 (*right*). The corresponding integrals are divergent and are typically regularised by introduction of a cut-off parameter Λ. Then, the corrections to the fermion and gauge boson mass parameters m become

$$\delta m \propto m \ln \frac{\Lambda}{m} \, .$$

Λ is interpreted as a scale where new physics becomes important and the computations of the SM are not valid anymore, which has to be the case at least at Λ_P or maybe already below. But even if $\Lambda = \Lambda_{\mathrm{P}}$, the corrections are of the order of m, i. e. the observable mass remains close to the mass parameter in the Lagrangian and pose no severe problem. However, in case of a fundamental scalar which does not originate in a gauge symmetry, like the Higgs boson, the corrections turn out to be quadratically divergent with Λ. The observable Higgs mass, corrected to first order (without logarithmic terms) is

$$m_H^2(\mathrm{obs}) \approx m_H^2 + r\Lambda^2 \;, \qquad r \in \mathbb{R} \;.$$

The fact that $m_H^2(\mathrm{obs})$ has to be close to the electroweak scale of about 100 GeV [44] and is not driven to the possibly much larger scale Λ is referred to as the *hierarchy problem* of the SM. The stability of the observable Higgs mass at the electroweak scale can only be assured if the parameter m_H^2 cancels the corrections to an extremely high precision, depending on the size of Λ. Although this is certainly not impossible, it is usually considered not desirable for a fundamental theory and referred to as *fine-tuning problem*. Turning the argument around, this might on the other hand imply new physics at the TeV scale.

2.2.2 Possible Extensions to the Standard Model

A large number of possible extensions to the SM have been proposed which address some of the previously discussed shortcomings. Obviously, any extended (or alternative) version has to reproduce the verified results of the SM.

For example, in *Little Higgs* models [61–63], new particles are added to the SM resulting in cancellation of the quadratic divergences to the scalar masses at lowest order. Of course, this is no fundamental cure and would only soften the hierarchy problem. Moreover, Little Higgs models are strongly constrained by experimental data.

Technicolour models [64–68] in contrast have been designed to avoid elementary scalars. The Higgs is assumed to be a composite state of new heavy fermions, bound by a QCD-like interaction that becomes confining at the TeV scale. However, attempts to generate fermion masses in this framework are difficult and often incompatible with the limits on *flavour-changing neutral currents* (FCNC) as well as electroweak precision measurements.

In a further category of models, generally denoted as *Grand Unified Theories* (GUTs), the SM symmetry groups are embedded into higher-dimensional groups such as $SU(5)$ or $SO(10)$, thus unifying the gauge interactions and relating the different charges and fermion types [69–71]. So far, the additionally introduced gauge bosons (*leptoquarks*) have not been observed in direct searches, e. g. at HERA [72, 73]. GUTs can also be tested indirectly because they predict the existence of magnetic monopoles and proton decays [74, 75]. In fact, the observed limits on the proton lifetime impose severe constraints [76].

In Arkani-Dimopoulos-Dvali (ADD) models on the other hand, the weakness of gravity compared to the other interactions is addressed by prediction of additional, large spatial dimensions [77,78]. While the gauge interactions are confined to ordinary space, gravity is assumed to penetrate the extra dimensions and thus lose flux. Hence, the actual strength of the gravitational coupling could be comparable to that of the gauge couplings, which corresponds to a reduction of the Planck scale to $\bar{\Lambda}_{\mathrm{P}} \ll \Lambda_{\mathrm{P}}$. Consequently, the hierarchy problem can be avoided if $\bar{\Lambda}_{\mathrm{P}}$ is of the order of 1 TeV, which in turn leads to spectacular signatures in high-energy collisions at the TeV scale such as neutral heavy resonances (*Kaluza-Klein resonances*) and the production of microscopic black holes. Recent CMS results, however, appear to disfavour current ADD models [79].

Finally, supersymmetric theories are maybe the most popular category of extensions to the SM. They are discussed below.

2.3 Supersymmetry

In the SM, fermions are postulated a priori while the existence of bosons follows from gauge symmetry[10]. In case of Supersymmetry (SUSY) on the other hand, a symmetry between fermions and bosons is postulated: the laws of physics are invariant under SUSY transformations, which turn fermions into bosons and vice versa,

$$Q\left|\text{fermion}\right\rangle \propto \left|\text{boson}\right\rangle \ , \qquad Q\left|\text{boson}\right\rangle \propto \left|\text{fermion}\right\rangle \ .$$

The SUSY generator[11] Q acts as ladder operator and changes the spin of a state by 1/2. Hence, Q transforms as a spinor itself and satisfies a graded Lie algebra

$$\begin{aligned} \{Q_a, Q_b\} &= \{\bar{Q}_a, \bar{Q}_b\} = 0 \\ \{Q_a, \bar{Q}_b\} &= 2\,(\gamma^\mu)_{ab}\, P_\mu \\ [Q_a, P^\mu] &= [P^\mu, \bar{Q}_a] = 0 \ , \end{aligned} \tag{2.9}$$

where the indices $a, b \in \{1, 4\}$ denote the spinor component and P_μ is the four-momentum operator. It is interesting to note that, given the internal gauge symmetries of the SM, SUSY is the largest possible extension to the Poincaré group, which includes translations, rotations, and boosts, that allows for non-trivial scattering processes [80].

Historically, SUSY has been investigated since the early 1970s, for example in the context of String theories, and the first supersymmetric quantum field theory in four dimensions has been developed by Wess and Zumino in 1974 [81]. However, broad interest began only after it became apparent that the hierarchy problem can be solved in supersymmetric theories. The simplest SUSY model which is not in conflict with phenomenology, the *Minimal Supersymmetric Standard Model* (MSSM), was developed in the early 1980s. Within the MSSM, a new particle (*superpartner*) is assigned to each SM particle with its spin differing by 1/2 but exactly the same properties otherwise. Evidently, this contradicts the experimental facts — if the superpartners had the same mass as the SM particles they would have been observed already — and therefore, SUSY has to be broken.

The MSSM will be discussed in more detail in Section 2.3.2, followed by a brief review of searches for SUSY and the derived constraints in Section 2.3.3. But first, Section 2.3.1 provides a motivation why SUSY is an interesting candidate for physics beyond the SM. An extensive introduction to Supersymmetry can be found for example in [49, 82], which have been used as a basis for this section.

2.3.1 Motivation

As previously outlined, there are divergent loop-corrections to the mass parameters in the SM Lagrangian. In case of SUSY, however, the divergences cancel because the superpartners contribute with equal size but opposite sign as the SM particles due to their different spin nature. Hence, the mass parameters are stabilised against radiative corrections, in

---

[10] Ignoring the fact that the Higgs has to be added rather arbitrarily.

[11] Models with more than one SUSY generator are usually in conflict with phenomenology.

particular also for fundamental scalars, thus solving the hierarchy problem without the necessity of fine tuning[12]. Even if SUSY is broken, quadratically divergent terms cancel exactly, and there are only logarithmic divergences remaining (*little hierarchy problem*). This is acceptable provided the masses of the superpartners are of the order of $\lesssim 1\,\text{TeV}$.

Equation Eq. (2.9) reveals a fundamental relation between SUSY and space-time: successive SUSY transformations correspond to the space-time translation operator. Furthermore, postulating local invariance under SUSY transformations, as it is done in certain models, requires the introduction of gravitational interactions mediated by a spin-2 gauge boson (*graviton*). Although this theory of *supergravity* is not renormalisable, the additional fermionic dimension provided by SUSY might still be a first step towards a consistent quantum field theory of gravity [83, 84]. In fact, realisation of SUSY is assumed in most phenomenologically important String theories, which are usually considered the ultimate framework for a consistent description of nature.

As a consequence of renormalisation, the gauge-coupling parameters of the SM are running. Although their values tend to meet when being evolved to higher scales, this does not happen at the same scale. Surprisingly, this appears to be quite precisely the case for SUSY at about $10^{16}\,\text{GeV}$ due to the additional loop contributions, cf. Fig. 2.5, if the masses of the superpartners are in the range of $100\,\text{GeV} - 10\,\text{TeV}$, thus hinting towards grand unification [85, 86]. In fact, many GUTs assume SUSY to be realised at scales close to the SM. It is also interesting to note that this range corresponds well to the one acceptable in order to provide a solution to the hierarchy problem.

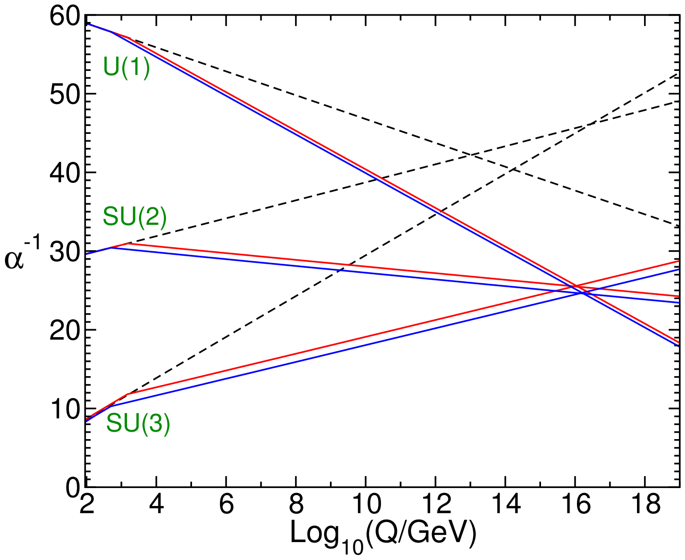


Figure 2.5: Renormalisation-group evolution of the inverse gauge coupling parameters $1/\alpha$ for the SM (*dashed lines*) and the MSSM (*solid lines*), taking into account two-loop effects. Common sparticle mass parameters of 0.5 (*blue lines*) and 1.5 TeV (*red lines*) have been assumed; the kinks in the MSSM lines mark the scale where sparticles start contributing to loops. Taken from [86].

---

[12] While particle masses can be protected against running to a higher scale by the introduction of SUSY, this does not provide an explanation for why the scales are so different in the first place.

In many SUSY models, there is a conserved quantum number *R-parity*

$$R = (-1)^{3(B-L)+2S} \,,$$

where B and L denote the baryon and lepton quantum numbers, respectively, and S the spin. Due to the $(-1)^{2S}$ dependence, the fermionic and bosonic superpartners obtain opposite values of R. This has the important phenomenological implication that SUSY particles can only be produced and annihilated in pairs and in consequence that the *lightest supersymmetric particle* (LSP) is stable. If it is also electrically and colour neutral, the LSP provides an excellent candidate for DM [87, 88].

Furthermore, limits on the mass of the lightest Higgs boson of the MSSM can be derived requiring $m_h \lesssim 130\,\mathrm{GeV}$. This is much more constraining than in case of the SM and compatible with the results of the direct Higgs searches [89, 90].

In conclusion, SUSY is an attractive concept for physics beyond the SM both in theoretical and phenomenological aspects. Several arguments suggest that the predicted new particles have masses of the order of 1 TeV, opening the possibility for their discovery at current collider experiments, in particular the LHC.

2.3.2 The MSSM

In the MSSM, a superpartner (*sparticle*) is assigned to each SM particle. It is denoted in the same way as its SM counterpart but with a tilde on top. The full particle and sparticle content of the MSSM is listed in Table 2.2.

The superpartners to the fermions are spin-zero bosons. They are termed *sfermions*, and likewise the names of the individual superpartners correspond to the SM name with a preceding 's', e. g. the superpartner of the electron would be the *selectron* $\tilde{e}$. In fact, each chiral component of a fermion is assigned a different superpartner. The left-handed fermions and the corresponding sfermions form chiral $SU(2)_L$-doublets, which are grouped into multiplets in *superspace* (*supermultiplets*), and the right-handed fermions and the corresponding sfermions are grouped into superdoublets. It is important to recall that chirality is not actually defined for sfermions because they are bosons; what is referred to is the chirality of the respective fermionic SM partner.

The fermionic superpartners to the SM gauge bosons (*gauginos*) are named with the additional suffix 'ino', e. g. the superpartner to the W boson would be the *Wino* $\widetilde{W}$. Gauge bosons and gauginos are also grouped into superdoublets.

In the MSSM Higgs-sector in contrast, further new particles have to be introduced since the charge conjugated Higgs doublet is not sufficient to generate masses for the up-type quarks. A second Higgs doublet is required. Hence, there are two Higgs doublets together with their superpartners, the *higgsinos*.

2.3.2.1 The Lagrangian

The Lagrangian of the MSSM consists of two parts: a supersymmetric extension of the SM Lagrangian, which includes the SM, plus SUSY breaking terms,

$$\mathcal{L}_{\mathrm{MSSM}} = \mathcal{L}_{\mathrm{SUSY}} + \mathcal{L}_{\mathrm{break}} \,.$$

Table 2.2: Particles and sparticles with spin S in the MSSM. For fermions, the indices 'L' and 'R' denote states of left and right chirality, respectively. For bosons, they refer to the chiral state of the fermionic superpartner. Arrows indicate superpositions of the sparticle's electroweak eigenstates to their mass eigenstates $\tilde{t}_{1,2}$, $\tilde{b}_{1,2}$, $\tilde{\chi}^0_{1,2,3,4}$, and $\tilde{\chi}^\pm_{1,2}$.

type	S	particles / sparticles			
leptons	$\frac{1}{2}$	$\begin{pmatrix}\nu_{e,L}\\ e_L\end{pmatrix}$, e_R	$\begin{pmatrix}\nu_{\mu,L}\\ \mu_L\end{pmatrix}$, μ_R	$\begin{pmatrix}\nu_{\tau,L}\\ \tau_L\end{pmatrix}$, τ_R	
sleptons	0	$\begin{pmatrix}\tilde{\nu}_{e,L}\\ \tilde{e}_L\end{pmatrix}$, $\tilde{e}_R$	$\begin{pmatrix}\tilde{\nu}_{\mu,L}\\ \tilde{\mu}_L\end{pmatrix}$, $\tilde{\mu}_R$	$\begin{pmatrix}\tilde{\nu}_{\tau,L}\\ \tilde{\tau}_L\end{pmatrix}$, $\tilde{\tau}_R$	
quarks	$\frac{1}{2}$	$\begin{pmatrix}u_L\\ d_L\end{pmatrix}$, u_R, d_R	$\begin{pmatrix}s_L\\ c_L\end{pmatrix}$, s_R, c_R	$\begin{pmatrix}t_L\\ b_L\end{pmatrix}$, t_R, b_R	
squarks	0	$\begin{pmatrix}\tilde{u}_L\\ \tilde{d}_L\end{pmatrix}$, $\tilde{u}_R$, $\tilde{d}_R$	$\begin{pmatrix}\tilde{s}_L\\ \tilde{c}_L\end{pmatrix}$, $\tilde{s}_R$, $\tilde{c}_R$	$\begin{pmatrix}\tilde{t}_L\\ \tilde{b}_L\end{pmatrix}$, $\tilde{t}_R$, $\tilde{b}_R$	$\leftrightarrow$ $\tilde{t}_{1,2}$, $\tilde{b}_{1,2}$
gauge bosons	1	$W^\pm$, Z, γ, g			
gauginos	$\frac{1}{2}$	$\tilde{W}^\pm$, $\tilde{Z}^0$, $\tilde{\gamma}$, $\tilde{g}$	$\tilde{\gamma}$, $\tilde{Z}^0$, $\tilde{H}^0_{1,2} \leftrightarrow \tilde{\chi}^0_{1,2,3,4}$; $\tilde{W}^\pm$, $\tilde{H}^\pm \leftrightarrow \tilde{\chi}^\pm_{1,2}$		
higgsinos	$\frac{1}{2}$	$\tilde{H}^0_{1,2}$, $\tilde{H}^\pm$			
Higgs bosons	0	h, H, A, $H^\pm$			

$\mathcal{L}_{\text{MSSM}}$ is constructed to be invariant under the $U(1)_Y \times SU(2)_L \times SU(3)_C$ gauge transformation of the SM and, up to a total derivative[13], under the SUSY transformations Eq. (2.9).

The first part, $\mathcal{L}_{\text{SUSY}}$, contains the kinetic and gauge interaction terms as well as a *superpotential*, which encodes fermion mass terms and further, Yukawa interaction terms. It is in fact the choice of the superpotential which defines a particular supersymmetric model. In its most general form, the superpotential even allows transitions between quark and lepton states. However, given the severe experimental constraints e. g. on the proton life-time [76], such terms must be suppressed. Technically, this can be achieved by requiring invariance of $\mathcal{L}_{\text{SUSY}}$ under discrete transformations generated by the R-parity quantum number.

As mentioned above, SUSY has to be broken for phenomenological reasons. The actual breaking mechanism is unknown, however. Hence, $\mathcal{L}_{\text{break}}$ is generically constructed to comprise all possible breaking mechanisms which do not introduce additional quadratically divergent terms (*soft breaking*). This requires a large number of additional parameters, leading to in total 124 free parameters of the MSSM. Although some of them are related to CP-violating or FCNC processes and are strongly constrained by experimental observations, the number is still too large for the MSSM to be an aesthetically pleasant

---

[13] In fact, it is not possible to construct a Lagrangian that is truly invariant under SUSY transformations but only up to a total derivative, thus still leaving the action invariant.

fundamental theory and practical framework to predict signatures of new physics. Hence, various simplified versions of the MSSM have been developed with specific breaking mechanisms in order to reduce the number of parameters, a few of which are discussed further below. It has to be presumed that, once SUSY signatures are observed, the actual breaking mechanism can be inferred.

2.3.2.2 The Higgs Sector

Analogous to the SM, spontaneous electroweak symmetry breaking in the MSSM is realised via the Higgs mechanism. The scalar potential is a function of the two Higgs fields and has contributions both from the superpotential and the breaking terms[14]. The conditions necessary for the scalar potential to develop non-vanishing vacuum expectation values v_u and v_d for the two Higgs fields are more constraining than in the SM. They are usually parametrised in terms of the phenomenologically important parameter

$$\tan\beta = \frac{v_u}{v_d}$$

and the sign of the Higgs self-coupling strength μ. Its magnitude is fixed, in contrast to the SM, by the Z boson mass. Remarkably, even with universal mass parameters at some large scale, the conditions for electroweak symmetry breaking can be met by renormalisation group evolution of the parameters to the electroweak scale, a phenomenon termed *radiative electroweak symmetry breaking* (REWSB). Since the top quark mass is large, REWSB occurs in a wide range of universal mass parameter values provided $\tan\beta \lesssim 60$.

As in the SM, the mass terms of the $W^\pm$ and Z bosons arise from the kinetic terms of the Higgs fields after symmetry breaking. This absorbs three of the eight degrees of freedom in the two complex Higgs doublets, resulting in five physical Higgs bosons, two charged ($H^\pm$) and three neutral ones (h,H,A). The mass m_h of the lightest one must be below or equal to the Z-boson mass at tree level and $m_h \lesssim 130\,\text{GeV}$ when including radiative corrections.

The matter fermions, i. e. the SM fermions, acquire masses from Yukawa coupling terms with the Higgs fields in the superpotential, leading to mass terms of the form Eq. (2.8), where however v has to be replaced by v_u and v_d for the up and down-type fermions, respectively. The sfermions get an identical Yukawa contribution to their mass since the superpotential is defined in terms of supermultiplets. There are additional contributions to the sfermion masses, of course, stemming from the SUSY breaking terms, which ensures compatibility with the non-observation of sparticles. The additional contributions can be different for the left- and right-handed sfermion states.

Due to the spontaneous $U(1)_Y \times SU(2)_L$ symmetry breaking, the sfermion mass eigenstates are different from the electroweak interaction eigenstates, leading to a mixing of the left- and right-handed sfermion states. The effect is particularly pronounced in case of the third generation where fermion masses are large. For example, the physical superpartners to the t_L and t_R states are $\tilde{t}_{1,2}$, superpositions of $\tilde{t}_L$ and $\tilde{t}_R$. Much alike, physical mass eigenstates of the fermionic superpartners are superpositions of the gaugino and higgsino fields (except for the gluinos since $SU(3)_C$ remains unbroken): the *charginos* $\tilde{\chi}_i^\pm$ are su-

---

[14] In general, the scalar potential is also a function of the sfermion fields. It is constructed in such a way, however, that the deepest minima occur along the 'directions' of the Higgs fields because otherwise charge or colour number conservation might be violated.

perpositions of the charged higgsinos and winos, and the *neutralinos* $\tilde{\chi}_i^0$ are superpositions of the neutral higgsinos, bino, and photino, with $i \in \{1, 4\}$.

2.3.2.3 Specific Breaking Scenarios

A common feature of many soft SUSY breaking scenarios is the existence of a *hidden sector* at a scale much larger than the weak scale. It is named 'hidden' because the involved particles do not (or only very weakly) participate in the SM gauge interactions. The actual SUSY breaking takes place in the hidden sector via some unknown process, and its effects are transmitted to the MSSM. Several concepts of SUSY breaking have been studied, e. g. *gauge-mediated breaking*, where additional gauge interactions are assumed to communicate the effects of SUSY breaking from the hidden to the observable sector, and *gravity-mediated breaking*, where gravitational interactions transmit the breaking. These specific scenarios typically require a much smaller number of free parameters than the more general MSSM.

A popular gravity-mediated model, which also serves as benchmark for the analysis presented in this thesis, is the *Constrained MSSM* (CMSSM) [49,86,91,92]. Following the notion of grand unification, common scalar and gaugino masses m_0 and $m_{1/2}$, respectively, are assumed at $\Lambda \approx 10^{16}$ GeV. Furthermore, the strength of certain coupling parameters in $\mathcal{L}_{\text{break}}$ become unified at that scale. They are parametrised by $\tan\beta$ and the universal trilinear coupling A_0, respectively. Altogether, the number of free parameters additional to the SM is reduced to five within the CMSSM: m_0, $m_{1/2}$, $\tan\beta$, A_0, and $\text{sign}(\mu)$. The coupling and mass parameters at lower scales are obtained by evolution of these parameters using renormalisation group equations.

2.3.3 Experimental and Astrophysical Constraints

A number of direct and indirect searches have been carried out in order to find evidence for SUSY in nature. So far, no clear signals from supersymmetric processes have been observed, but the results were used to constrain the allowed MSSM parameter space.

Various searches for signatures from sparticles directly produced at high-energy colliders have been performed. The most stringent limits of the pre-LHC era were obtained from the experiments at HERA [93,94], LEP [95–99], and Tevatron [100–102] and put a lower bound of a few hundred GeV on the squark and gluino masses, depending on the specific breaking scenario. The limits have been greatly exceeded by the current LHC results, which are a topic of this thesis and will be discussed later in Chapter 7.

The existence of SUSY could also manifest indirectly via higher-order contributions to SM processes. However, combination of various measurements, e. g. Z-pole results from the LEP and SLC experiments as well as measurements of m_W and m_t from the Tevatron experiments [103], proves consistency of the SM expectations only. Interestingly, SM predictions of the anomalous magnetic moment of the muon, '$(g-2)_\mu$', deviate by $3.4\,\sigma$ from precision measurements by the E821 experiment at Brookhaven National Laboratory [104,105]. Furthermore, sizable sources of both FCNC and CP violation can occur in the SUSY breaking sector depending on the specific parameter values. Their occurrence has been investigated extensively in decays involving b quarks. For example, measurements of $b \to s\gamma$ transitions at the CLEO, BaBar, and Belle experiments [106–108] provide no evidence for contributions incompatible with the SM. Moreover, the rate expected from the SM for the rare process $B_s^0 \rightarrow \mu^+\mu^-$ is in good agreement to measurements at Tevatron [109] and LHC [110, 111] although some deviation is reported by CDF [112]. Hence, experiments have verified the SM to very high precision, putting bounds on possible sparticle masses and couplings, and in particular requiring a severe suppression of any SUSY-induced FCNC and CP violation.

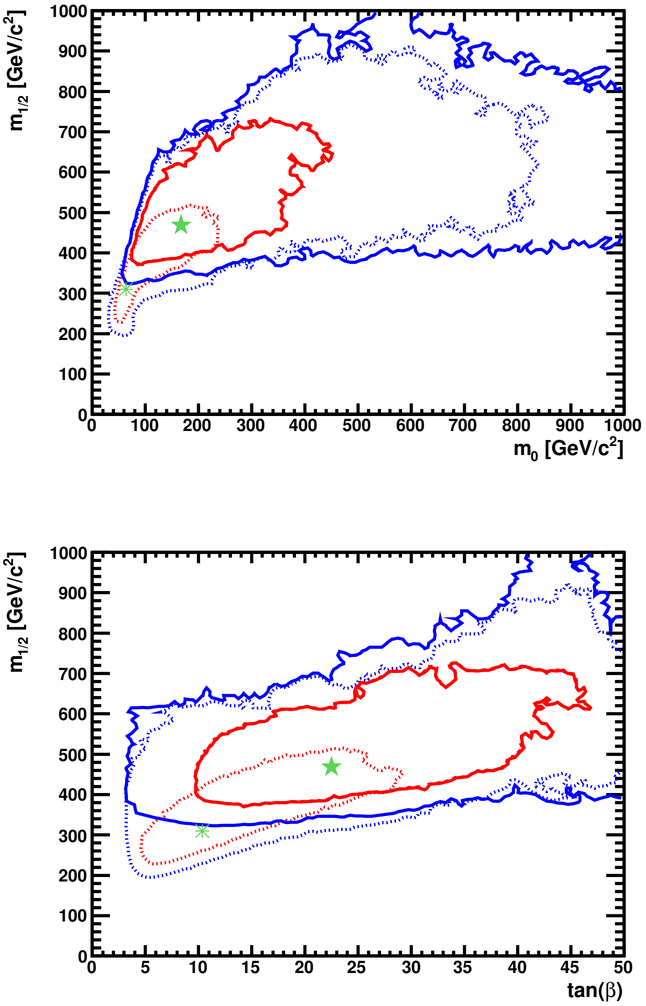


Figure 2.6: Contours of the likelihood function obtained assuming the CMSSM and taking into account data from various direct and indirect SUSY searches. Shown are the regions at 68 (*red lines*) and 95% (*blue lines*) confidence level as well as the best-fit values (*stars*) before (*dashed*) and after (*solid*) inclusion of XENON100 and 2010 LHC results in the $m_0 \times m_{1/2}$ (*top*) and $\tan\beta \times m_{1/2}$ (*bottom*) planes. Taken from [113].

Values of $m_0 = 170^{+330}_{-80}$ GeV, $m_{1/2} = 470^{+140}_{-70}$ GeV, and $\tan\beta = 22^{+27}_{-13}$ are most compatible with the data with a fit probability of 16%. In consequence, gluino masses of $m_{\tilde{g}} \approx 1100$ GeV and LSP masses of $m_{\tilde{\chi}_1^0} \approx 200$ GeV are favoured with uncertainties of approximately 100 – 300 GeV. The impact of recent LHC data on the fit results is discussed later in Section 7.3.5.

So far, astrophysical data suggest that DM corresponds to *weakly interacting massive particles* (WIMPs). After the Big Bang, production and annihilation of WIMPs were in thermal equilibrium, but as the universe expanded, the reaction rate dropped until eventually the number of WIMPs remained constant (*relic density*). Assuming R-parity conservation, the lightest neutralino forms an excellent WIMP candidate provided it is the LSP, which is the case in many MSSM-based scenarios. If indeed neutralinos constitute DM further constraints on the MSSM parameters can hence be derived from the observed relic density [114–116].

Several groups [113, 117–121] have fitted the parameters of the CMSSM and other MSSM-based models to the data from various direct and indirect SUSY searches. For example, the MasterCode collaboration [122, 123] has published an analysis taking into account measurements of $(g-2)_\mu$, electroweak-precision and b-physics observables, and the DM density as well as exclusion limits on the Higgs and sparticle production at colliders including 2010 LHC results [113]. In Fig. 2.6, the best-fit values together with the 68 and 95% confidence-level contours of the likelihood function obtained when assuming the CMSSM are shown in the $m_0 \times m_{1/2}$ and $\tan\beta \times m_{1/2}$ planes.

3 Experimental Setup

Since the development of the first particle accelerators in the 1920s, accelerators became the primary tool in experimental particle physics to produce new particles and probe the structure and interactions of matter [124]. The most powerful particle accelerator and collider to date is the *Large Hadron Collider* (LHC) [125, 126] operated by the *European Organisation for Nuclear Research* (CERN) in Geneva. With the LHC, proton-proton (pp) collisions with unprecedented centre-of-mass energies of 7 TeV are delivered. Furthermore, heavy ions, e. g. lead ions, can be collided with centre-of-mass energies of 2.76 TeV.

The primary goals of the LHC programme are to unravel the mystery of electroweak symmetry breaking and to search for signals from possible new physics beyond the SM. Four large detectors have been installed in underground caverns around the interaction points to measure the collision products: two multipurpose experiments ATLAS [127] and CMS [128, 129], designed for a broad physics programme including in particular the search for the Higgs boson and signs of possible new physics; LHCb [130], specialised to study CP violation and possible indirect signals of new physics in b-quark systems; and ALICE [131], specialised for heavy-ion physics and the study of primordial states of matter.

This chapter is organised as follows. After a review of basic quantities and detector concepts in Sections 3.1 and 3.2, the LHC is described in Section 3.3. In Section 3.4, the design and performance of the CMS detector and its subsystems are discussed. Finally, the particle identification and reconstruction at CMS are explained in Section 3.5.

3.1 Basic Quantities at Hadron Colliders

In the following, the coordinate system at CMS and basic quantities important for the discussion of hadron-collider physics are introduced.

3.1.1 Coordinate System

Throughout this thesis, a right-handed coordinate system is used with its origin at the nominal interaction point inside the CMS detector. In Cartesian coordinates, the x axis points towards the centre of the LHC ring and the y axis upward towards the surface. Thus, the z axis is oriented along the proton-beam direction towards the Jura mountains. In cylindrical coordinates, the radial distance r and azimuthal angle ϕ are defined in the xy plane, where $\phi = 0$ corresponds to the x axis, and the polar angle θ is defined with respect to the z axis.

3.1.2 Kinematic Quantities

At the LHC, the actual high momentum-transfer (*hard*) interactions do not take place between the colliding protons but between two of their partons. Since the partons carry only the fractions x_i, $i = 1, 2$, of the proton's momenta, the centre-of-mass frame of the colliding partons is boosted with respect to the laboratory frame[1] along the direction of the incoming protons, i. e. the z axis. This has implications for the choice of kinematic variables as discussed for example in [132] and reviewed in the following.

If $\sqrt{s}$ denotes the centre-of-mass energy of the pp collision, the parton's four-momenta in the laboratory frame are given by[2]

$$\begin{aligned} p_1 &= \frac{\sqrt{s}}{2}(x_1, 0, 0, x_1) \\ p_2 &= \frac{\sqrt{s}}{2}(x_2, 0, 0, -x_2) \ , \end{aligned}$$

and the centre-of-mass energy $\sqrt{\hat{s}}$ of the parton-parton collision by $\hat{s} = x_1 x_2 s$. Hence, the final state has an invariant mass of $M = \sqrt{x_1 x_2 s}$ and, in the laboratory frame, a rapidity $y = \frac{1}{2}\ln(x_1/x_2)$, and thus

$$\begin{aligned} x_1 &= \frac{M}{\sqrt{s}} e^{y} \\ x_2 &= \frac{M}{\sqrt{s}} e^{-y} \ . \end{aligned}$$

This relationship between the momentum-fractions x_i of the partons and the kinematic variables M and y of the final state is depicted in Fig. 3.1 for $\sqrt{s} = 7\,\text{TeV}$. A high centre-of-mass energy is essential to enable the production of heavy particles and probe interactions at large energy scales.

For technical reasons, it is difficult to measure precisely the four-momenta of the collision products close to the z axis. Hence, the boost of the final-state's centre-of-mass frame cannot be determined accurately. Physics analyses are therefore preferentially performed in terms of the *transverse momentum* p_T, the momentum vector's component in the plane transverse to the z axis, which is the same in both the parton's and the laboratory's centre-of-mass frame.

Particles that interact only weakly, such as neutrinos, do not produce any direct signal in the LHC detectors. They manifest indirectly as an imbalance in the observed total transverse momentum[3] in the event, denoted as *missing transverse momentum.* Commonly, the term *missing transverse energy* ($\not{E}_\mathrm{T}$) is used, implying that the momenta's absolute values are scaled to the energy.

The rapidity is another important quantity to characterise the final state because the number of produced particles is roughly constant per rapidity interval, and rapidity differences are Lorentz invariant. However, the rapidity is again difficult to measure. Therefore, the *pseudorapidity* η, defined as

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \ , \tag{3.1}$$

is commonly used instead. A pseudorapidity $\eta = 0$ corresponds to the direction transverse

---

[1] This is assuming that the centre-of-mass frame of the colliding protons coincides with the laboratory frame, which is the case at CMS.

[2] Natural units are used throughout this thesis, as remarked previously.

[3] Understood as the vectorial sum of all transverse momenta.

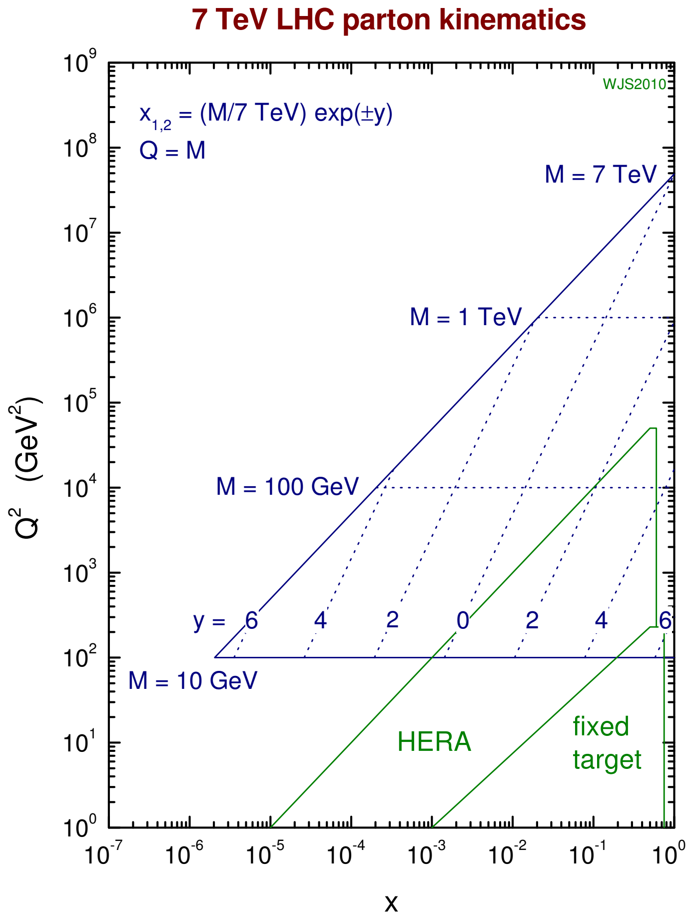


Figure 3.1: Relationship between the momentum transfer Q (which is equal to the final-state invariant mass M), the parton momentum-fraction x, and the final-state rapidity y probed at the LHC with $\sqrt{s} = 7\,\text{TeV}$. The kinematic reach of the HERA collider and of typical fixed-target experiments is also shown. Taken from [133].

to the proton beam and $|\eta| \to \infty$ is directed along the beam. The pseudorapidity becomes equal to the rapidity in the relativistic limit.

3.1.3 Luminosity

The rate of a certain type of process at a particle collider is determined by the product of the process' cross section and the *luminosity* $\mathcal{L}$, which characterises the experimental conditions. At the LHC, the luminosity is given by

$$\mathcal{L} = f \cdot \frac{k_B N_p^2}{a} \,,$$

where k_B denotes the number of bunches per beam, N_p the number of protons per bunch, and f their revolution frequency; the factor a describes the geometric bunch size transverse to the direction of collision [129]. The number of events of this process occurring during the time T is proportional to the *integrated luminosity* $L = \int_0^T \mathrm{d}t\,\mathcal{L}$. Hence, the design of a particle collider typically aims at a large luminosity in order to provide a high rate of the rare, interesting processes expected from new-physics scenarios.

3.2 Detector Concepts in Experimental Particle Physics

Two basic detector concepts are commonly deployed in particle physics to measure the momenta, energies, and directions of particles: *tracking detectors* and *calorimeters*. Their measurement principles and properties are briefly reviewed in the following Sections 3.2.1 and 3.2.2. Modern experiments such as CMS typically combine several different tracking detectors and calorimeters optimised for different types of particles in order to reconstruct the final-state particles' four-momenta with high precision.

3.2.1 Tracking Detectors

With tracking detectors, the positions (*hits*) of traversing particles are measured at several locations in order to reconstruct their trajectories. Various different technologies exist. They are discussed for example in [6] and include two broad classes of detectors, which are sensitive to electrically charged particles:

Gaseous Detectors: Traversing charged particles ionise a gas. The free charge carriers produce a signal for example in wires to which a voltage is applied and that are arranged in parallel thus providing position information.

Semiconductor Detectors: In this case, free charge carriers are produced in the depletion zone of a semiconductor to which a bias voltage is applied such that an electric current occurs.

Typically, tracking detectors are placed in a magnetic field such that the trajectories of charged particles are bent. For a homogeneous field, the momentum p_T perpendicular to the field direction of a relativistic particle with electric charge q in units of e can be determined from

$$p_\mathrm{T} = 0.3\,qB\rho\;,$$

where B denotes the magnetic flux density in units of T and ρ the track's radius of curvature in the transverse plane in units of m [6].

The relative p_T resolution can be approximated by

$$\frac{\sigma(p_\mathrm{T})}{p_\mathrm{T}} = \frac{\sigma_x p_\mathrm{T}}{qBL^2}\sqrt{\frac{720}{N+4}} \oplus \sigma_\mathrm{ms} \tag{3.2}$$

where $\oplus$ denotes addition in quadrature [6]. The first term is due to the hit-position measurement errors σ_x along the trajectory of length L in case of N hits. The second term denotes a contribution due to multiple-scattering and is approximately $\sigma_\mathrm{ms} \propto 1/(p_\mathrm{T}\sqrt{L})$. Hence, the resolution is dominated by multiple-scattering at low p_T and by the hit-position resolution at high p_T. At CMS, multiple-scattering dominates the resolution for $p_\mathrm{T} \lesssim 10\,\mathrm{GeV}$ [128].

3.2.2 Calorimeters

Calorimeters are essentially blocks of dense matter in which incoming high-energy particles induce *particle showers*, cascades of secondary particles with successively less energy. Low energetic, charged shower particles generate photons or free charges in the detector material, the number of which depends on the incident particle's energy. Hence, calorimeters are primarily used to measure the particle's energy and, since they are typically segmented, position. A detailed discussion of particle showers and calorimeters can be found for example in [6, 134–136].

Two basic designs, *sampling* and *homogeneous* calorimeters, exist. The former consists of separated layers of a passive, high-density absorber material, e. g. lead, brass, or uranium, and an active material generating the signal, e. g. a scintillator, an ionising gas, or a semiconductor. The latter are composed of one material, e. g. lead glass, which both absorbs the particles and generates the signal.

Two types of particle showers can be distinguished:

Electromagnetic showers: Above a few 10 MeV, electrons in matter lose energy primarily by bremsstrahlung, and photons in matter convert to e^+e^- pairs. Hence, above approximately 1 GeV, they induce a cascade of photons and electrons[4] with successively lower energies. Below a critical energy, where electrons and photons release their energy by ionisation and the photoeffect, respectively, the shower development stops [6].

Electromagnetic showers develop fairly regularly. Their dimensions scale to first approximation with the matter constants *radiation length* X_0 and *Molière radius* R_M, independently of the details of the absorber material. One X_0 corresponds to both the mean distance after which an electron has lost all but $1/e$ of its energy by bremsstrahlung and 7/9 of a photon's mean free path for pair production. The longitudinal profile $\mathrm{d}E/\mathrm{d}x$ of deposited energy E follows a gamma distribution, where the scale x corresponds to the distance in units of X_0. There is only a weak, approximately logarithmic dependence of the shower maximum on the initial energy. A cylinder with radius R_M contains on average 90% of the deposited energy [6]. Details of the interaction of low-energetic ($\mathcal{O}(1)$ MeV) particles in the shower[5], which generate the measured signal, depend very much on the material, however, and cannot be described properly in terms of X_0 and R_M [135].

Hadronic showers: Hadronic showers are initiated by hadrons, and hence, the strong interaction contributes substantially to the shower development. As a consequence, further hadrons are produced, approximately 90% of which are pions. On average, 30% of these are neutral pions, which decay into two photons which subsequently initiate an electromagnetic shower [135]. The fraction f_em of energy transferred to this electromagnetic component (*electromagnetic fraction*) varies strongly from shower to shower. On average, f_em increases with the incident hadron's energy E since pion production occurs also in subsequent shower steps. It is often[6] parametrised as

---

[4] Here, 'electron' denotes electrons and positrons.

[5] For example, 75% of the energy is deposited by electrons, the rest by positrons [135].

[6] Another common parametrisation, which does not result in $f_\mathrm{em} > 1$ for large E, is $f_\mathrm{em} = 1 - (E/E_0)^{m-1}$ with $E_0 \approx 1$ GeV and m between 0.80 and 0.87 [137].

$f_{\mathrm{em}} = 0.11 \cdot \ln(E/\mathrm{GeV})$ [136]. The charged hadrons decay further, such that the non-electromagnetic energy is on average deposited to 56% by ionising particles (mostly protons with energies between 50 and 100 MeV) and to 10% by soft neutrons with energies of typically 3 MeV. The remaining 34% are lost in releasing nucleons from their nuclei and do not contribute to the calorimeter signal (*invisible energy*) [135].

The average profile of a hadronic shower is similar to that of an electromagnetic shower. Its scale depends on the *nuclear interaction length* λ_I, a matter constant corresponding to the mean distance after which a hadron has lost all but $1/e$ of its energy. However, the dimensions of hadronic showers are typically much larger. For example, the $\mathrm{PbWO_4}$ crystals of CMS's electromagnetic calorimeter have a radiation length of $X_0 = 0.89\,\mathrm{cm}$ but an interaction length of $\lambda_I = 20.28\,\mathrm{cm}$ [6]. Furthermore, the profile of an individual hadronic shower varies strongly due to fluctuations of f_{em}.

A characteristic quantity of calorimeters is the ratio e/h of the individual detection efficiencies e and h for electromagnetic and hadronic energy deposits in a particle shower, respectively [6,135–137]. It is often determined from the measured ratio π/e of calorimeter responses to pions and electrons of the same energy,

$$\frac{\pi}{e} = \frac{f_{\mathrm{em}} e + (1 - f_{\mathrm{em}})h}{e} = \frac{1 + (e/h - 1)\, f_{\mathrm{em}}}{e/h} \ . \tag{3.3}$$

Due to the energy dependence of f_{em}, π/e also depends on the energy: the calorimeter response is said to be *non-linear.* Calorimeters are typically either optimised to measure electrons and photons or hadrons.

The relative calorimetric energy-resolution is usually parametrised as

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

i. e. improves with increasing energy E of the incident particle [6]. Different effects contribute. At low energies, the calorimeter resolution is dominated by electronic noise, described by the *noise term* N. At larger energies, fluctuations of the shower development, e. g. the track length of individual particles in the cascade, dominate, parametrised by the *stochastic term* S. At high energies, the resolution is ultimately limited by miscalibration and non-uniformities of the calorimeters, which is described by the *constant term* C. In case of the electromagnetic calorimeter of CMS, the noise term dominates for $E \lesssim 12\,\mathrm{GeV}$, and the constant term dominates for $E \gtrsim 200\,\mathrm{GeV}$ [129].

3.3 The Large Hadron Collider

The LHC is a 26.7 km long, two-ring hadron accelerator and collider, which consists of eight straight sections and eight arcs. It is installed in a tunnel which lies 45 – 170 m below the surface in the vicinity of Geneva between the lake Léman and the Jura mountains. The protons counter-revolve in two evacuated beam pipes, which cross in the interaction points, enabling collision events. A detailed description of the LHC and its performance can be found in [125,126,138,139].

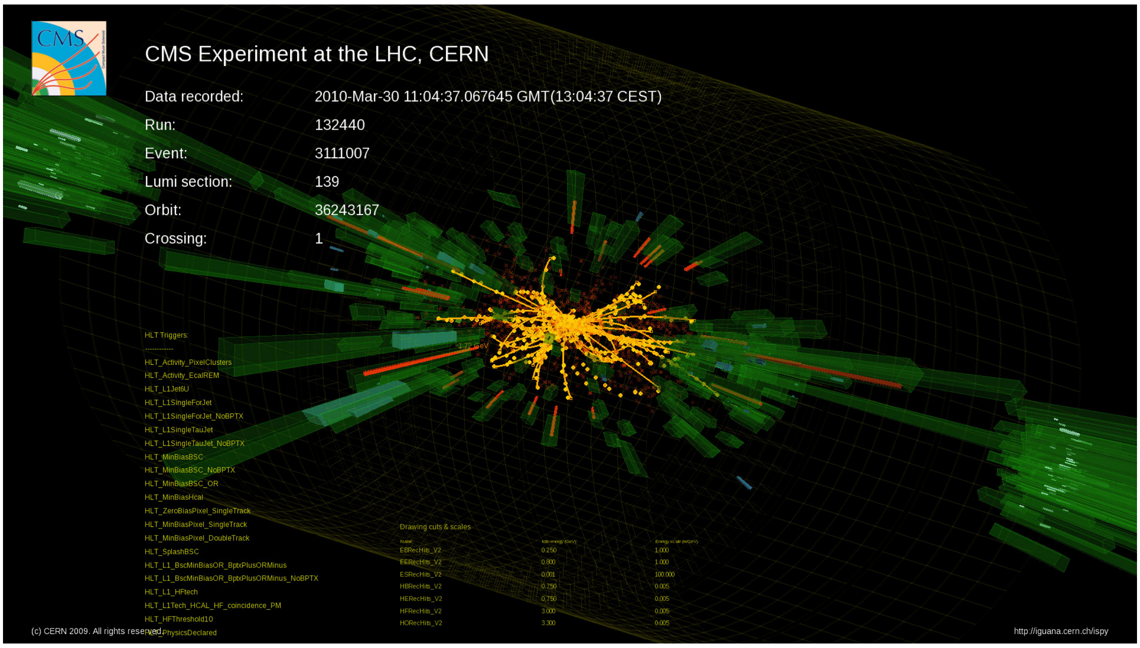
LHC operation started in September 2008. After recovery from a major cooling incident, the first collisions of protons in stable beams at injection energy of 450 GeV were achieved on November 23, 2009. The first ever pp collisions at 7 TeV centre-of-mass energy took place on March 30, 2010, cf. Fig. 3.2. Subsequently, up to $47.03\,\mathrm{pb}^{-1}$ of data were delivered to each experiment in 2010, which were used for detailed studies of the detector performance as well as a number of first searches for new physics. After a shutdown period during winter, almost 150 times more pp collision data were produced by LHC during 2011, corresponding to up to $5.714\,\mathrm{fb}^{-1}$ per experiment as shown in Fig. 3.3.

In the LHC, protons are accelerated and stored using 16 superconducting radio-frequency (RF) cavities operating at 40 MHz with field gradients of up to 5.5 MV/m. The protons are arranged in bunches located at points of equal field strength. As a consequence of the RF pattern, each beam consists of up to 2808 bunches[7] separated by at least 25 ns corresponding to approximately 7.5 m [142]. Nominally, $1.15 \cdot 10^{11}$ protons can be stored per bunch. At the interaction points, the bunches are approximately 8 cm long with an RMS radius of $20\,\mu$m. They are steered and squeezed with a system of nearly 9300 superconducting magnets producing a magnetic conduction of up to 8.33 T. To enable superconductivity, RF cavities and magnets are operated at a temperature of 1.9 K.

Before injection into the LHC, the protons are accelerated to 450 GeV by a chain of pre-

---

[7] Given the 40 MHz RF and the circumference of the LHC, there are in principle 3564 possible bunch positions, but a few larger inter-bunch gaps are required by the proton injection system [142].



Data recorded: 2010-Mar-30 11:04:37.067645 GMT(13:04:37 CEST)
Run: 132440
Event: 3111007
Lumi section: 139
Orbit: 36243167
Crossing: 1

HLT Triggers:

HLT_Activity_PixelClusters
HLT_Activity_EcalREM
HLT_L1Jet6U
HLT_L1SingleForJet
HLT_L1SingleForJet_NoBPTX
HLT_L1SingleTauJet
HLT_L1SingleTauJet_NoBPTX
HLT_MinBiasBSC
HLT_MinBiasBSC_NoBPTX
HLT_MinBiasBSC_OR
HLT_MinBiasHcal
HLT_ZeroBiasPixel_SingleTrack
HLT_MinBiasPixel_SingleTrack
HLT_MinBiasPixel_DoubleTrack
HLT_SplashBSC
HLT_L1_BscMinBiasOR_BptxPlusORMinus
HLT_L1_BscMinBiasOR_BptxPlusORMinus_NoBPTX
HLT_L1_HFtech
HLT_L1Tech_HCAL_HF_coincidence_PM
HLT_HFThreshold10
HLT_PhysicsDeclared

Drawing cuts & scales

Name:	Min energy (GeV):	Energy scale (m/GeV):
EBRecHits_V2	0.250	1.000
EERecHits_V2	0.800	1.000
ESRecHits_V2	0.001	100.000
HBRecHits_V2	0.750	0.005
HERecHits_V2	0.750	0.005
HFRecHits_V2	3.000	0.005
HORecHits_V2	3.300	0.005

(c) CERN 2009. All rights reserved.

http://iguana.cern.ch/ispy

Figure 3.2: Visualisation of the detector signature caused by one of the first pp collision events at 7 TeV centre-of-mass energy recorded by CMS on March 30, 2010. Green cuboids represent cells of the electromagnetic and hadronic calorimeter, respectively, with their height being proportional to the measured transverse energy, and solid yellow lines represent reconstructed particle trajectories. Taken from [140].

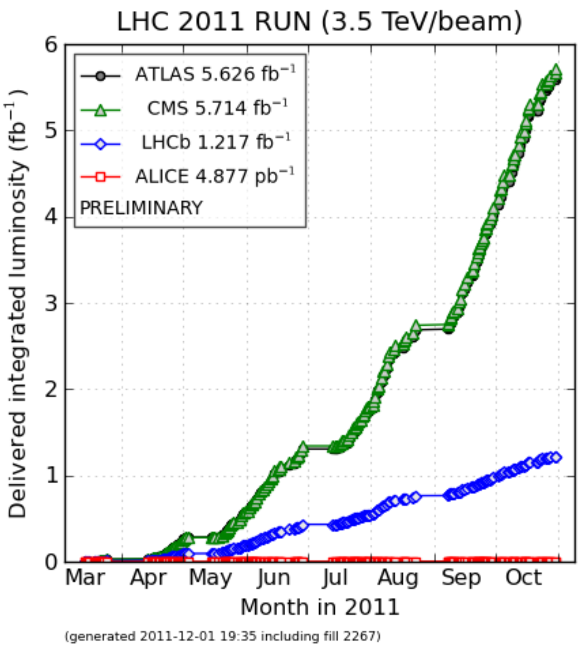


Figure 3.3: Integrated luminosity of the pp collision data delivered to the experiments by LHC in 2011 as a function of time. Taken from [141].

accelerators as depicted in Fig. 3.4. In the LHC, proton energies are increased to 3.5 TeV, which takes approximately 20 minutes, and the bunches are steered to intersect in the interaction points such that protons can collide. Due to the circular beam structure and the huge number of protons per bunch, the bunches can be used for repeated crossings: with one fill, the LHC can operate with sufficient luminosity for up to 30 hours during which the experiments record data (*run*).

During most of the 2011 data-taking period, the LHC was operated with 1380 bunches per beam with a nominal distance of 50 ns. With proton multiplicities of approximately $1.5 \cdot 10^{11}$ per bunch[8], luminosities of a few $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ were achieved [139]. On average, nine pp collisions occurred per bunch crossing, an effect which is referred to as *pile-up*. Pile-up is a byproduct of the desired high luminosity and poses significant challenges on the experiments, which have to deal with a much increased number of collision products per event that have to be assigned to a particular pp interaction. Some relevant LHC machine and performance parameters are listed in Table 3.1.

The total cross section for inelastic pp scattering at $\sqrt{s} = 7$ TeV is approximately 60 mb. Cross sections of important production processes and their rate at the LHC are shown in Fig. 3.5. For example, the $t\bar{t}$ cross-section amounts to approximately 165 pb [144–146], corresponding to approximately one event every 6 seconds at a luminosity of $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The cross section for a SM Higgs boson with a mass of 120 GeV is approximately 10 pb (latest higher-order calculations predict 19 pb [147]), i. e. one Higgs boson is expected to be produced every 1 to 1.5 minutes. Typical SUSY processes have cross sections below the order of 1 pb [148].

---

[8] This is in fact larger than the design value.

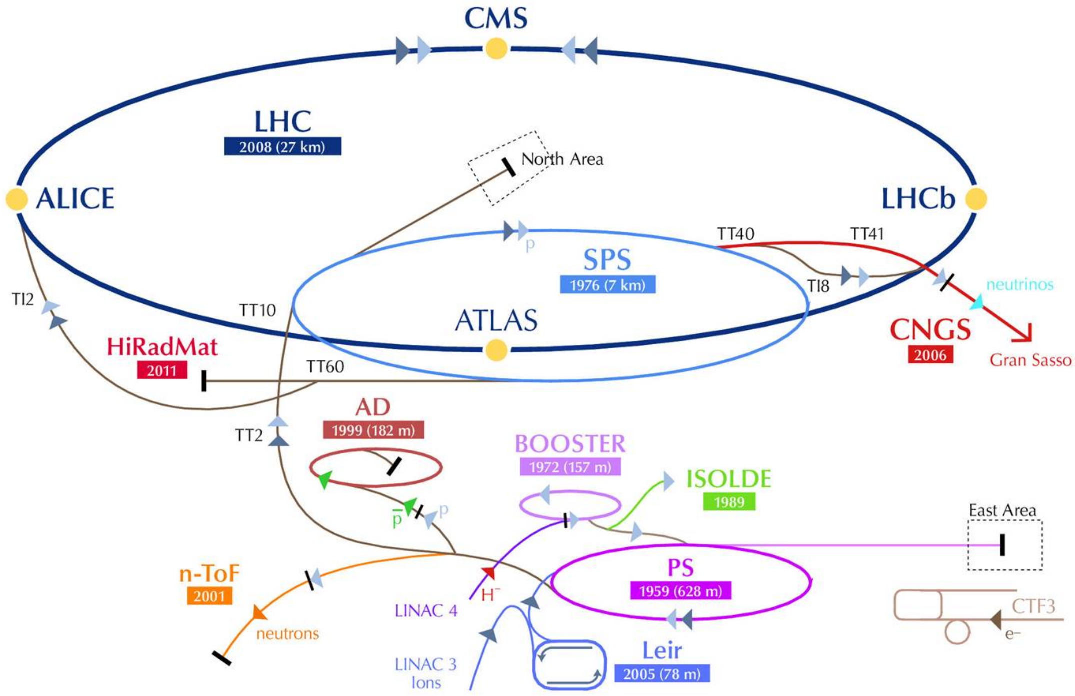


Figure 3.4: Sketch of the CERN accelerator complex. Below the names, the year of first operation and the circumference (in case of ring accelerators) are stated for each machine. Before being injected into the LHC, protons are pre-accelerated by the *Linac4*, the *Proton Synchrotron Booster* (PSB), the *Proton Synchrotron* (PS), and finally the *Super Proton Synchrotron* (SPS). Taken from [143].

Table 3.1: Relevant LHC machine parameters. [129, 139]

	Typical in 2011	Design Value
Centre-of-mass energy per proton (TeV)	7	14
Luminosity ($\mathrm{cm}^{-2}\mathrm{s}^{-1}$)	$2 \cdot 10^{33}$	10^{34}
Bunch separation (ns)	50	25
Number of bunches	1380	2808
Number of protons per bunch	$1.5 \cdot 10^{11}$	$1.15 \cdot 10^{11}$
Number of pp collisions per bunch crossing	≈ 9	≈ 20

For 2012, proton collisions at a centre-of-mass energy of 8 TeV are scheduled[9] with a targeted integrated luminosity of $15\,\mathrm{fb}^{-1}$. Afterwards, a major upgrade of the LHC is intended enabling operation at design centre-of-mass energy of 14 TeV.

---

[9] At the time this thesis is published, LHC has already delivered more than $5.7\,\mathrm{fb}^{-1}$ of 8 TeV data, i. e. succeeded the total integrated luminosity achieved in 2011.

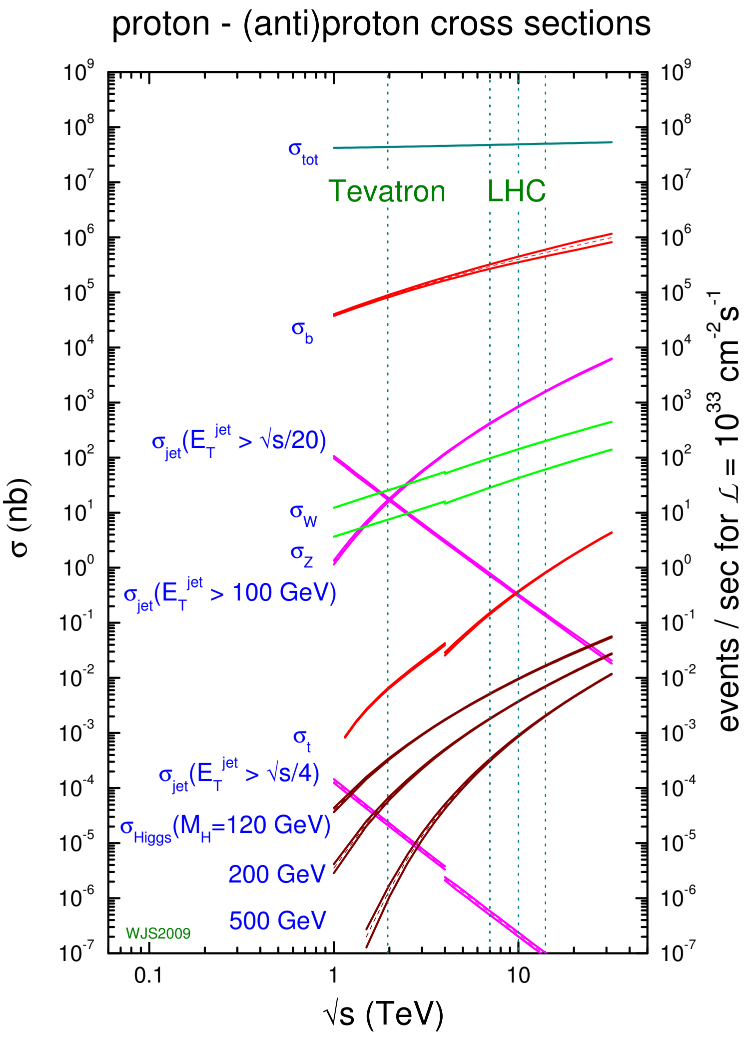


Figure 3.5: Total SM cross-sections for inelastic $p\bar{p}$ and pp scattering and cross sections of important exclusive processes as a function of the centre-of-mass energy $\sqrt{s}$. The corresponding rate for a luminosity of $10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is also shown. Discontinuities in the lines are due to the transition from $p\bar{p}$ scattering at low $\sqrt{s}$ to pp scattering at high $\sqrt{s}$. Taken from [133].

3.4 The Compact Muon Solenoid Experiment

CMS (the *Compact Muon Solenoid*) is one of the two multipurpose detectors at the LHC. An international collaboration of currently over 4000 scientists, students, and engineers from 189 institutes in 41 countries designed and constructed the apparatus over a period of approximately 20 years and are now operating and using it for research. The experiment's primary quests are to explore the mechanism of electroweak symmetry breaking, namely to search for the Higgs boson, and to search for signs of possible new physics, either directly by discovering new particles or indirectly by observing deviations from SM predictions at LHC energies. In consequence, the detector design is aimed in particular at a good photon, electron, and muon identification with excellent momentum resolution to cover the important Higgs decay channels $H \to \gamma\gamma$ and $H \to 4l$. In order to identify τ leptons and b jets, also good vertex reconstruction capabilities are required. Furthermore, R-parity conserving SUSY models predict the production of weakly interacting LSPs, such as neutralinos, which do not produce any direct signal in the detector but only manifest as an imbalance in the observed total transverse momentum in the event. Hence, the expected SUSY signatures motivate the requirement for a good $\not\!\!E_{\mathrm{T}}$ resolution for which large geometric coverage is essential. The layout of the roughly cylinder-shaped CMS detector is depicted in Figs. 3.6 and 3.7. Its primary features are a silicon-based tracking detector at the centre of the apparatus and an electromagnetic calorimeter of lead-tungstate crystals. These and other subdetectors, which are particularly relevant for this thesis, are described below in Sections 3.4.1 to 3.4.4. The tracking volume and most of the calorimetry are contained within a superconducting solenoid with a length of 12.9 m and a diameter of 5.9 m, which provides a uniform axial magnetic field with a flux density of 3.8 T enabling high-resolution measurements of charged particles' momenta. Overall, the CMS detector has a length of 21.6 m and a diameter of 14.6 m with a total weight of 12500 t. Its subsystems are installed in seven lined-up, movable disks, thus enabling access for maintenance also to the innermost systems. Following the typical subdetector design, which consists of a barrel-shaped part plus endcaps, the region at smaller $|\eta|$ is termed 'barrel' or 'central' region, and the regions at larger $|\eta|$ are termed 'endcap' or 'forward' regions.

To cope with large number of particles produced in the pp collisions — approximately 1000 charged particles are expected per event at design luminosity — the subdetectors

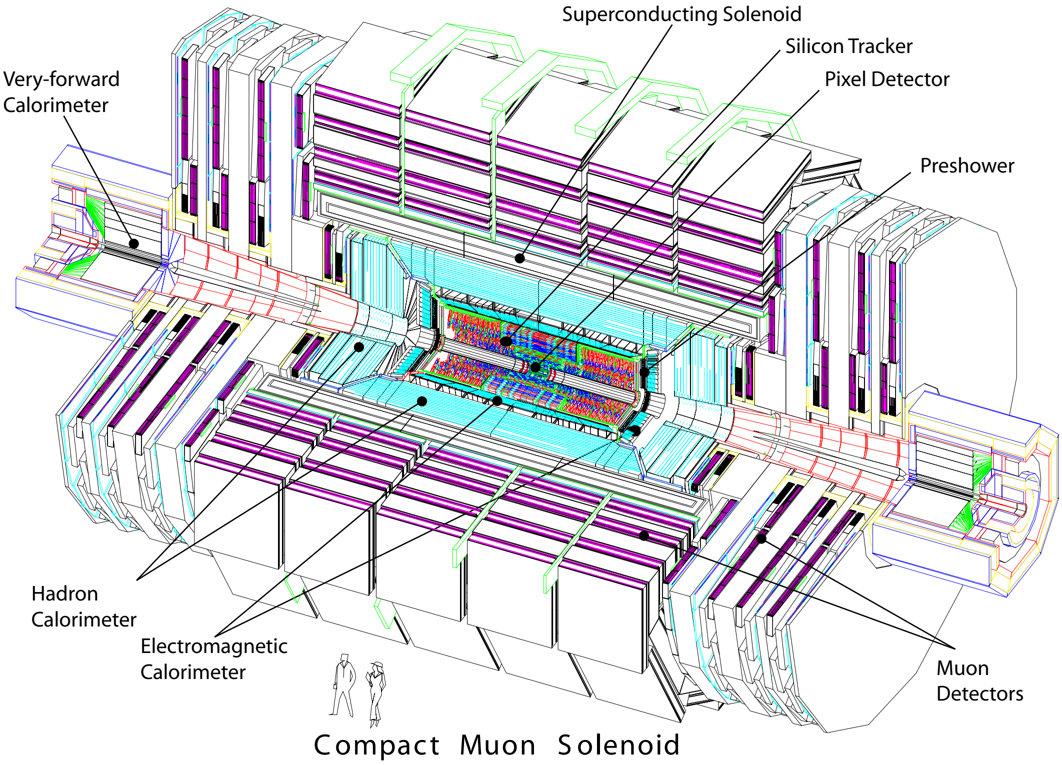


Figure 3.6: Layout of the CMS detector: three-dimensional sketch. Taken from [129].

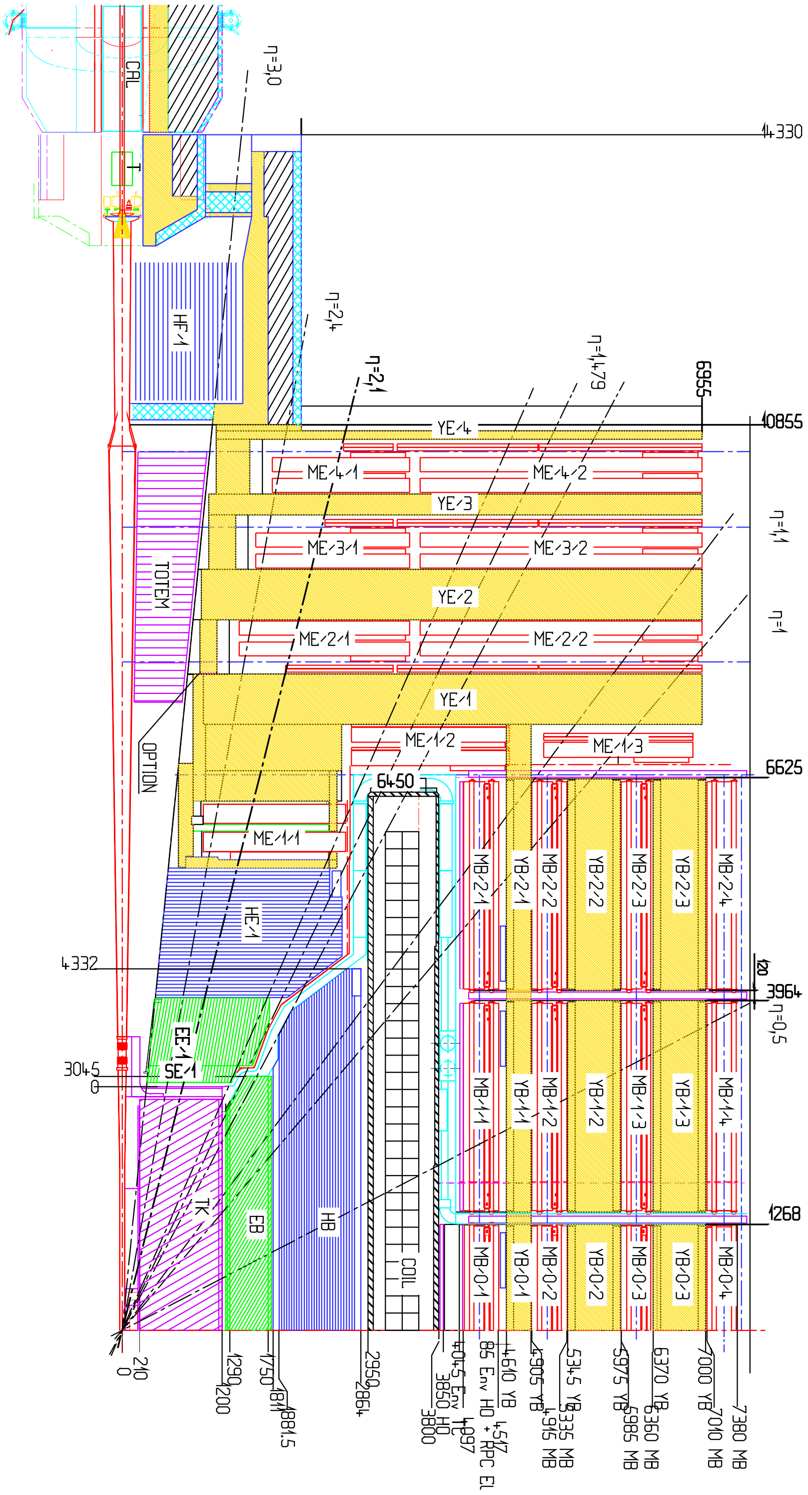


Figure 3.7: Layout of the CMS detector: cross section of one quarter in an rz view. (The interaction point is located at the bottom-left corner.) Dimensions are given in units of mm. Taken from [129].

have a high granularity thus reducing the occupancy. The resulting number of almost 10^8 electronic readout-channels produce information corresponding to approximately 1 MB per event. In combination with the LHC bunch-crossing rate, this leads to a huge data rate of the order of 10 TB per second, which is far too large to be stored or analysed. Hence, an online event selection (*trigger*) system, which identifies potentially interesting events, reduces the rate of recorded events to approximately 300 Hz as described in Section 3.4.5. This still corresponds to a rate of the order of 1 PB per year. The data are stored and analysed on a world-wide distributed network of computing facilities [149, 150].

A detailed description of the CMS detector and all involved subsystems can be found in [128, 129].

3.4.1 Inner Tracking System

The tracking system [128, 129, 151, 152] is the innermost subdetector of CMS. Its sensitive components are silicon-based sensors, which measure the ionisation energy loss of traversing charged particles. A sketch of the tracking system is shown in Fig. 3.8. It is a 5.8 m long cylinder of radius 1.3 m with a sensitive volume up to $|\eta| = 2.5$. The detector consists of two subsystems, which are described in the following.

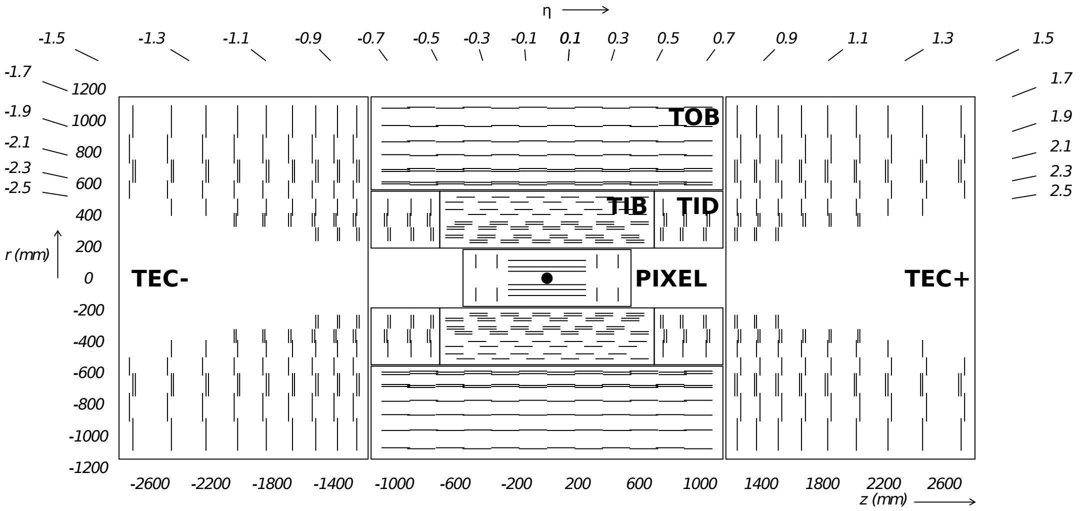


Figure 3.8: Layout of the inner tracking system in an rz view. Lines represent the pixel- and strip-sensor modules. Different parts of the strip tracker are indicated: *Tracker Inner Barrel* and *Disc* (TIB, TID), and *Tracker Outer Barrel* and *End Cap* (TOB, TEC). Taken from [153].

Pixel Tracker: The pixel tracker is located closest to the interaction point around the beampipe with a sensitive volume between an inner radius of 4.4 cm and an outer radius of 10.2 cm. It consists of several layers of modules to which silicon pixel sensors with cell sizes of $100 \times 150\,\mu\mathrm{m}^2$ are mounted. There are 66 million pixels in total, ensuring an occupancy of approximately 10^{-4} per pixel even in the high particle flux close to the interaction point. As a result, a hit-position resolution of

$10-20\,\mu$m is achieved. To support high-quality track reconstruction, the modules are arranged such that at least three hits are present for tracks with $|\eta|$ up to almost 2.5.

Strip Tracker: The strip tracker surrounds the pixel tracker, and its sensitive volume extends to a radius of 116 cm. It is subdivided into an inner and outer system, which each consists again of several module layers. To reduce the number of readout channels, the modules are comprised of silicon strip sensors with pitch sizes between 80 and $180\,\mu$m in the inner and outer part, respectively. Due to the reduced particle flux at larger radii, the occupancy is kept at the percent level. The achieved position resolution varies from 23 to $52\,\mu$m in the $r\phi$ plane and from 230 to $530\,\mu$m in the z direction.

The total track-reconstruction efficiency of the inner tracking system, including effects such as detector geometry, silicon sensor properties, and track-finding performance, has been measured in the environment of 7 TeV collision data [154]. It is above 80% for pions with p_T between 0.5 and 500 GeV, increasing to above 90% between 2 and 50 GeV. For muons, the efficiency is better than 99% for p_T between 0.5 and 500 GeV.

Due to the good hit-position resolution of the tracking system, pp-collision vertices (*primary vertices*) in the event can be reconstructed with high precision using tracks. The spatial resolution of primary vertices with more than 30 tracks has been measured to be between 20 and $25\,\mu$m [155], which is sufficient, for example, to identify pile-up collisions. Furthermore, displaced *secondary vertices* from decays of long-lived particles such as B-mesons can be reconstructed.

The dimensions of the tracking system and the 3.8 T solenoid field enable track-based momentum measurements for charged particles with p_T between 100 MeV and 1 TeV [154]. An excellent relative transverse-momentum resolution of $0.5-1.5\%$ is achieved for particles with p_T of $1-10$ GeV and $|\eta| < 1.5$, which is limited by multiple-scattering processes in the dense detector material. At larger p_T, the relative resolution becomes dominated by the hit-position resolution[10] and degrades approximately proportionally with p_T, cf. Section 3.2.1, but is still only $1.5-2\%$ at 100 GeV. For $|\eta| > 1.5$, the resolution is generally worse due to geometric effects [128, 156].

3.4.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) [128, 129, 157] is a homogeneous calorimeter installed around the inner tracking system. It consists of lead-tungstate ($\mathrm{PbWO_4}$) crystals, which have a short radiation length $X_0 = 0.89$ cm and a small Molière radius $R_\mathrm{M} = 2.2$ cm, enabling a compact detector with good position resolution. The scintillation process of $\mathrm{PbWO_4}$ is fast enough for LHC conditions, 80% of the photons are emitted within 25 ns, but suffers from a low light yield of approximately 30 photons per MeV, requiring high amplification by the photodetectors. Furthermore, both the yield and the performance of the photodetectors feature a temperature dependence of approximately 2%/°C at 18°C. Therefore, the system's temperature has to be stabilised within 0.1°C to preserve the energy resolution.

---

[10] At 100 GeV, multiple scattering accounts for $20-30\%$ of the resolution [128].

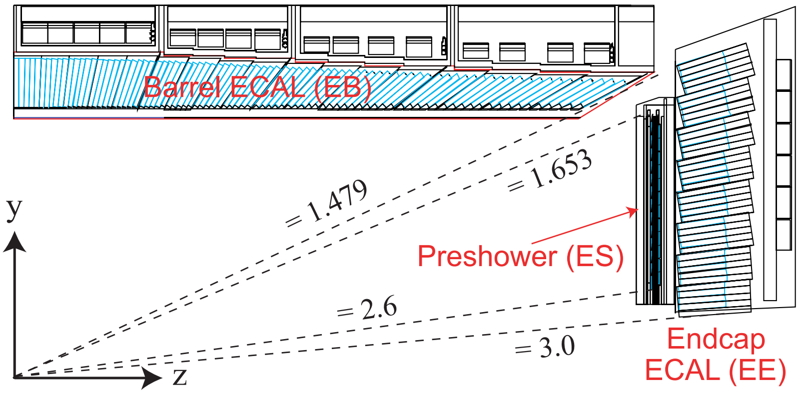


Figure 3.9: Layout of one quarter of the electromagnetic calorimeter in an rz view. Taken from [129].

The ECAL layout is sketched in Fig. 3.9. Since the trajectories of charged particles are bent in the magnetic solenoid field, only particles with momenta above approximately 1.5 GeV actually reach the ECAL's surface. The detector is subdivided into three subsystems described in the following.

Electromagnetic Barrel (EB) Calorimeter: The ECAL barrel (EB) detector is a cylinder with inner radius of approximately 1.3 m covering the region up to $|\eta| = 1.479$. It comprises 61200 truncated-pyramid shaped crystals, which are mounted in a quasi-projective geometry such that their axes are tilted both in η and ϕ by 3° with respect to the direction towards the nominal interaction point. This way, passage of particles along the boundaries of two adjacent crystals is avoided. The crystals have a cross section of 0.0174×0.0174 in $\eta \times \phi$, corresponding to $22 \times 22\,\mathrm{mm}^2$ at the inner side of the detector. Their length of 230 mm equals $25.8\,X_0$. The produced scintillation light is detected by avalanche photodiodes.

Electromagnetic Endcap (EE) Calorimeter: Two disc-shaped ECAL endcaps (EE) perpendicular to the beamline extend the coverage to $|\eta| = 3.0$. Each EE contains 3662 truncated-pyramid shaped crystals with a cross section of $28.6 \times 28.6\,\mathrm{mm}^2$ and a length of 220 mm corresponding to $24.7\,X_0$. The crystals are grouped by 5×5 into clusters, the axes of which point again slightly off the nominal interaction point to avoid passage along the boundaries. Vacuum phototriodes are exploited for readout.

Preshower Detector: In front of each EE calorimeter, a preshower detector is installed covering the region $1.653 < |\eta| < 2.6$. It is a $3\,X_0$ thick sampling calorimeter with absorbers of lead and two layers of silicon strip sensors as active material. With its high granularity, the detector is used to identify neutral pions which decay into two collimated photons.

The ECAL is predominantly utilised in the measurement of photons and electrons but also charged hadrons, cf. Section 3.5. It is of vital importance for example in the search for the Higgs boson via the $H \to \gamma\gamma$ decay channel.

The performance of the ECAL has been measured using test-beam data [129, 158–160]. Its response to pions depends on their energy while it is constant for electrons, resulting in $e/h = 1.6$. The relative resolution to electrons with energy E is parametrised as

$$\frac{\sigma(E)}{E} = \frac{0.124}{E/\mathrm{GeV}} \oplus \frac{0.036}{\sqrt{E/\mathrm{GeV}}} \oplus 0.0026\,, \tag{3.4}$$

i. e. amounts to 0.6% for 50 GeV electrons.

The ECAL calibration is performed with $\pi^0 \to \gamma\gamma$, $W \to e\nu$, and $Z \to ee$ data leading to a crystal-intercalibration precision of 0.6% [161]. Variations of the crystals' transparency are monitored continuously and corrected for using laser-injected light as a reference.

3.4.3 Hadronic Calorimeter

The hadronic calorimeter [128, 129, 162] (HCAL) is a sampling calorimeter enclosing the ECAL. For most parts of the detector, brass with an interaction length of $\lambda_I = 16.42$ cm constitutes the absorber material, interspersed with plastic scintillator tiles. The HCAL layout is shown in Fig. 3.10. It consists of four subsystems described in the following.

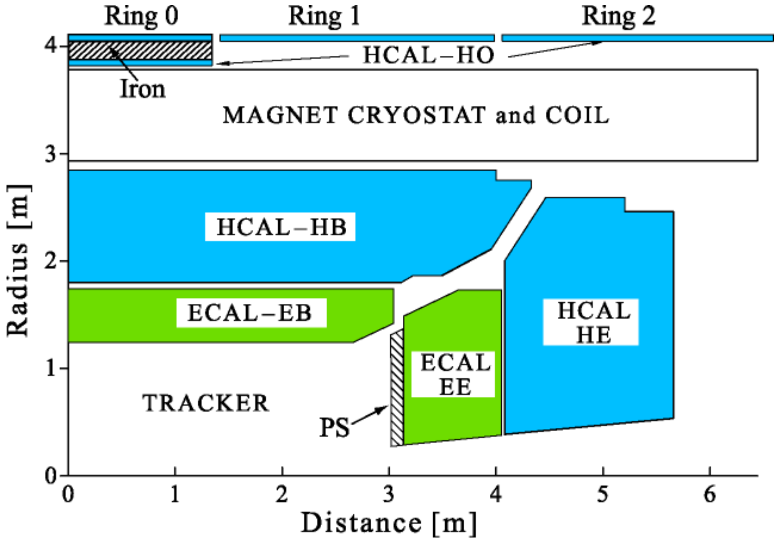


Figure 3.10: Location of one quarter of the electromagnetic and hadronic calorimeters in an rz view. Taken from [160].

Hadronic Barrel (HB) Calorimeter: The hadronic barrel calorimeter is a cylinder covering the volume between the ECAL and the magnet coil up to $|\eta| = 1.392$. It consists of 17 absorber layers between which the scintillator tiles are mounted. The scintillator tiles are segmented into cells with a size of 0.087×0.087 in $\eta \times \phi$, which are arranged in a projective way with respect to the nominal interaction point. All tiles lying upon another in one $\eta \times \phi$ cell are connected by wavelength-shifting fibres and read out as single longitudinal channel, resulting in 2304 truncated-pyramid shaped *HCAL towers*. One tower precisely covers 5×5 ECAL crystals. Since the HB calorimeter is located entirely within the solenoid, its depth is restricted to only $5.8\,\lambda_I$ at $|\eta| = 0$, increasing to $10.6\,\lambda_I$ at $|\eta| = 1.3$; the ECAL amounts to another 1 to $2\,\lambda_I$ in front of the HCAL.

Hadronic Outer (HO) Calorimeter: The hadronic outer calorimeter comprises scintillator tiles mounted outside the magnet coil such that the HB-tower geometry is matched. Due to the additional material of the solenoid, the total absorber depth of the barrel calorimeters is extended to a minimum of $11.8\,\lambda_I$ in the region up to $|\eta| = 1.305$. The HO has not been used for event reconstruction so far.

Hadronic Endcap (HE) Calorimeter: The disc-shaped hadronic endcap calorimeter extends the HB coverage to $|\eta| = 3.0$ and contains 17 layers of scintillator tiles between the brass absorbers. Its tower geometry matches the HB up to $|\eta| = 1.70$, after which the $\eta \times \phi$ cross section increases up to 0.350×0.175 at $|\eta| = 3.0$.

Hadronic Forward (HF) Calorimeter: The disc-shaped hadronic forward calorimeter is installed at $z = \pm 11.2\,\mathrm{m}$ and covers the region $3.0 < |\eta| < 5.0$. Its absorber material is steel with a depth of approximately $10\,\lambda_I$. Radiation-hard quartz fibres are embedded parallel to the z axis as active component, which are suited for operation in the high particle flux in the forward region. Charged shower particles generate Cherenkov light in the fibres, bundles of which are read out by a common photomultiplier, leading to 13 towers with a size of 0.175×0.175 in $\eta \times \phi$. Fibres of two different length are used, such that some collect shower particles over the whole HF depth and some only at larger depth. By this effective longitudinal segmentation, electromagnetic shower-components can be identified, thus compensating for the missing ECAL coverage.

The HCAL is predominantly exploited in the measurement of neutral hadrons, cf. Section 3.5.2. Furthermore, the HF extends the detector's geometric coverage beyond the inner tracking and ECAL systems up to $|\eta| = 5.0$, which is particularly important for a precise determination of $\not{E}_\mathrm{T}$.

The performance of the HCAL has been measured using test-beam data [129, 158–160]. Its pion-energy response has been calibrated for 50 GeV and increases by approximately 50% between 2 and 300 GeV while it is energy independent at the 2% level for electrons, which corresponds to $e/h = 1.4$. The measured relative energy resolution of the combined ECAL and HCAL system is parametrised by

$$\frac{\sigma(E)}{E} = \frac{1.2}{\sqrt{E/\mathrm{GeV}}} \oplus 0.069 \,, \tag{3.5}$$

which corresponds to 18% for pions with an energy of $E = 50\,\mathrm{GeV}$.

3.4.4 Muon System

The muon system [128, 129, 163] is the outermost subdetector of CMS. Located outside the magnet coil, it covers a cylindric volume with inner and outer radii of approximately 4 and 7 m, respectively, with additional, disc-shaped caps such that the geometric acceptance extends up to $|\eta| = 2.4$. The system comprises several layers of gaseous detectors [6] interleaved with iron, which absorbs hadrons and also serves as a return yoke for the magnetic flux. It is employed in the event reconstruction to identify muons and to improve the muon-track and -momentum measurement of the inner tracking system, cf. Section 3.5. Furthermore, it is used by the trigger system for muon identification.

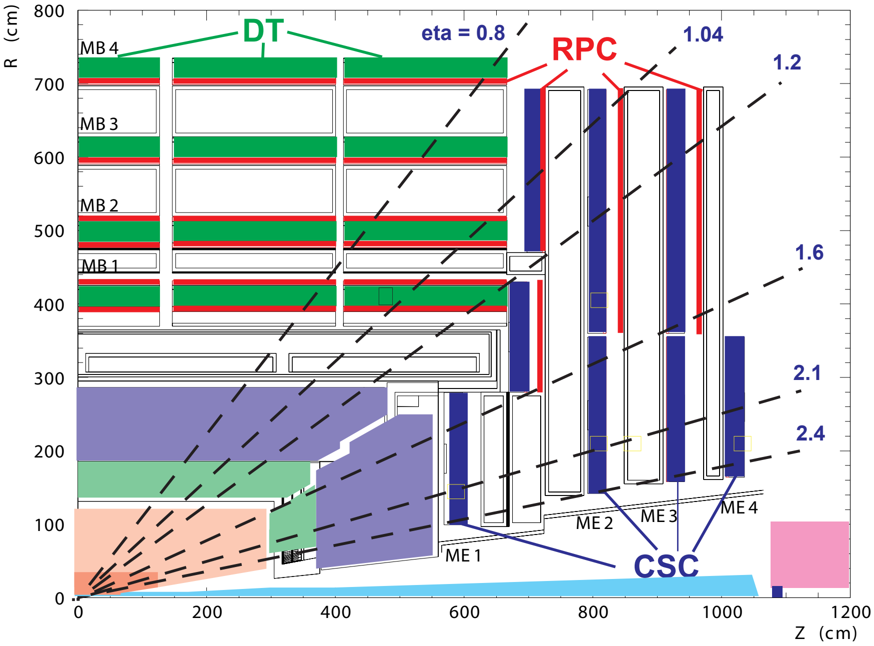


Figure 3.11: Layout of one quarter of the muon system in an rz view. Taken from [129].

A sketch of the muon system is shown in Fig. 3.11. In total, it contains approximately 25000 m^2 of active material with almost 1 million readout channels. Three subsystems with different detector technologies have been chosen since the homogeneity of the magnetic field, the radiation environment, and the timing requirements vary in different pseudorapidity regions [163]. They are described in the following.

Drift Tube Chambers: Drift tube (DT) chambers cover the barrel region with $|\eta| = 1.2$. Here, the rate of muons and neutron-induced background is small, less than 20 $\mathrm{Hz/cm}^2$ at design luminosity, and the magnetic field affecting the detectors is weak. Four layers (*stations*) of up to twelve DT chambers each are installed between the layers of the return yoke, each station providing a directional resolution of 1 mrad in ϕ and better than 100 μm in r.

Cathode Strip Chambers: The forward region up to $|\eta| = 2.4$ is instrumented with cathode strip chambers (CSCs). They enable measurements with high time and good spatial resolution even though the particle flux is high, up to 1 $\mathrm{kHz/cm}^2$, and a strong, non-uniform magnetic field is present. At each side of the CMS detector, 234 CSCs are installed in four discs perpendicular to the beamline such that a muon in the region $1.2 < |\eta| < 2.4$ passes at least three CSCs. The spatial resolution provided by each chamber is approximately 200 μm in r and z and approximately 10 mrad in ϕ.

Resistive Plate Chambers: Both the DT chambers and the CSCs are complemented by resistive plate chambers (RPCs) in the region up to $|\eta| = 1.6$. The RPCs can be operated at high particle rates up to 10 $\mathrm{kHz/cm}^2$ with a good time resolution and a response much faster than 25 ns. Hence, they enable unambiguous assignment of signals to the correct bunch crossing and are also utilised for the trigger system.

The performance of the muon reconstruction, which exploits predominantly the inner tracking and the muon systems, cf. Section 3.5, has been measured in data [164]. Identification and reconstruction efficiencies are above 95% for muons with p_{T} larger than a few GeV with misidentification probabilities below 1%, and the single-muon trigger efficiency is above 90%. The transverse-momentum scale at the Z mass has been calibrated to a precision of 0.2%. For muons with p_{T} below 100 GeV, the relative p_{T} resolution varies between 1 and 2% in the barrel and 6% in the endcaps. It is dominated by the resolution of the inner-tracker because measurements in the muon system are affected by multiple scattering and radiation losses in the material before the muon stations. At higher p_{T}, the degrading precision of the track-curvature measurement limits the resolution, cf. Section 3.2.1. Hence, significant improvements are gained by combination with information from the muon system, which profits from the longer lever arm. This way, the relative p_{T} resolution in the barrel is better than 10% for muons with p_{T} up to 1 TeV [128].

3.4.5 Trigger System

The LHC bunch-crossing rate of 40 MHz and the CMS-event size of approximately 1 MB lead to a massive amount of produced data too large to be stored and analysed. Therefore, a trigger system [165, 166] performs a fast online selection of potentially interesting events, lowering the rate of recorded events to approximately 300 Hz. This reduction by a factor of more than 10^5 does not limit the discovery potential of CMS because the cross sections for processes of interest, such as W/Z-boson or t-quark production and in any case Higgs boson or SUSY production, are even smaller compared to the total pp cross section, cf. Fig. 3.5.

The trigger system operates in two subsequent steps as explained in the following.

Level-1 (L1) Trigger: The level-1 (L1) trigger performs a simplified event reconstruction employing only data from fast detector components such as the calorimeters and muon system at a reduced granularity. Events are selected based on the presence of physics objects such as photons, electrons, muons, or jets above some p_{T} threshold and based on global event characteristics such as the amount of $\not{E}_{\mathrm{T}}$. The time for an L1-trigger decision, including the signal transfer-time from the detector front-end electronics to the remote hardware-based trigger logic and back, is fixed to $3.2\,\mu$s, during which the complete event data is held in buffers. This way, the rate of further-processed events is reduced to a maximum of 100 kHz.

High-Level Trigger (HLT): If an event is accepted by the L1 trigger, further signal processing and compression steps are executed, and the data is placed in front-end readout buffers. Parts are transferred to a computing farm, where the high-level trigger (HLT) software runs in parallel on approximately 5000 CPUs and performs a more sophisticated event reconstruction considering information from all subdetector systems. Interesting events are selected based on the more detailed event properties, achieving a reduction of the event rate to approximately 300 Hz. The decision time for one event is limited to 50 ms. Since the HLT is software-based, reconstruction algorithms and selection requirements can be quickly adjusted during operation.

Several independent trigger requirements (*trigger paths*) are tested in parallel, and an event has to pass only one of them to be accepted. Hence, different data sets can be collected, which are defined for example by the number of jets above some p_{T} threshold in an event or by the amount of $\not{E}_{\mathrm{T}}$, in order to optimise the preselection for different physics analyses.

Since production cross sections are typically decreasing with increasing p_{T}, the trigger thresholds have to be raised with increasing luminosity to keep the final event rate at the desired value. However, trigger paths with lower thresholds can be maintained, for example to collect events containing low-p_{T} jets, by assigning a *prescale* n to the path, which means that only every n-th triggered event is actually stored. Another common approach is to require more exclusive signatures, e. g. two jets above a given p_{T} threshold.

3.5 Particle Identification at CMS

The particles produced in a pp collision emerge from the interaction point, possibly decay further, and finally interact with the detector material in a characteristic way which enables to distinguish different types of particles. In the following Section 3.5.1, the signatures of different particles in the various subdetectors of CMS are briefly discussed. The *Particle-Flow (PF) algorithm* described in Section 3.5.2 combines the information from all subdetectors of CMS in order to identify individual particles and reconstruct their four-momenta.

3.5.1 Signatures of Different Particles

In the following, typical signatures of frequently produced particles with energies above a few 100 MeV in the CMS detector are discussed. They are also illustrated in Fig. 3.12.

Photons: Photons predominantly initiate electromagnetic showers in the ECAL. Due to the high depth of approximately $25\,X_0$, the photon's energy is entirely deposited in the ECAL, distributed over a few crystals. In most cases, more than 94% of the energy is contained within a matrix of 3×3 crystals [168]. Depending on η, between 20 and 60% of the photons already convert to e^+e^- pairs during their passage through the inner tracking detector, which has a depth of $0.4 - 1.8\,X_0$ [128].

Electrons: Electrons deposit most of their energy in the ECAL via electromagnetic showers. However, some energy is lost by bremsstrahlung during the passage through the tracking detector. Due to their electric charge, electrons also produce hits in the inner tracking detector by ionisation of the silicon.

Muons: Due to their much larger mass compared to electrons, muons in matter lose energy predominantly through ionisation [6, 134]. They are minimum ionising for all energies relevant at CMS and hence traverse the entire detector with minimal energy loss. For example, approximately only 2 GeV are deposited in the calorimeters [129]. Hence, muons produce hits in the inner tracking detector and in the muon system.

Hadrons: Charged hadrons in matter lose energy both through ionisation, such as muons, and through strong processes converting the original hadron into a number of new particles [6, 134]. Hence, they produce hits in the inner tracking system and initiate hadronic showers in the calorimeters, mostly the HCAL due to its larger interaction length. Neutral hadrons, in contrast, produce only hadronic showers. Examples for

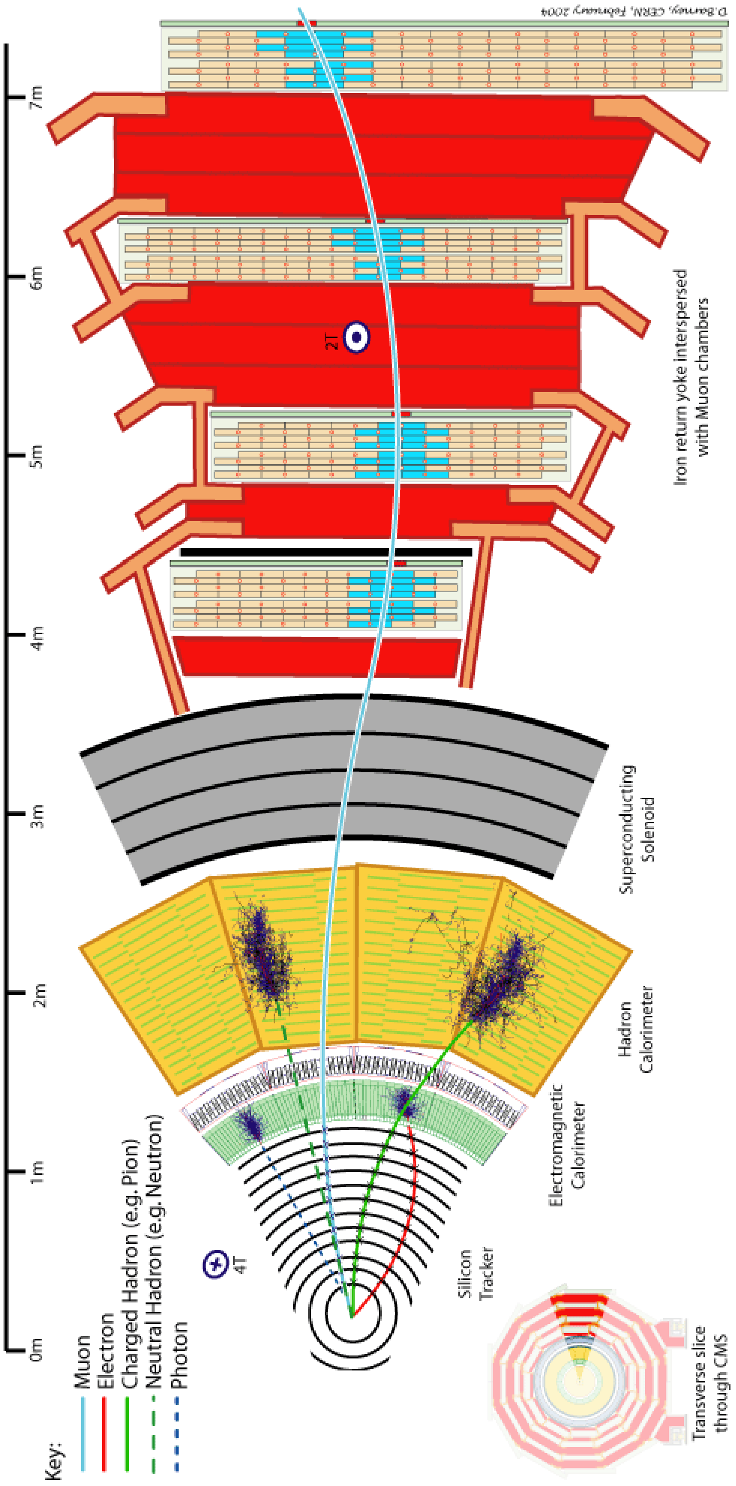


Figure 3.12: Transverse slice through the CMS detector. (The intercation point is located at the bottom.) Signatures of different particles are indicated. Taken from [167].

hadrons which do not already decay within the beampipe and actually reach the sensitive layers of the CMS detector are: charged pions, charged kaons, protons, and neutrons.

3.5.2 The Particle-Flow Algorithm

CMS employs the Particle-Flow (PF) approach to identify the types and reconstruct the four-momenta of all stable particles in an event, such as muons, electrons, photons, as well as charged and neutral hadrons, by combination of the information from all subdetectors. The procedure is explained in detail in [169].

The event reconstruction starts with the identification of charged-particle tracks in the inner tracking and the muon system. Furthermore, clusters are defined from adjacent calorimeter cells with energy deposits. A single particle in general produces several of these elementary signatures. Hence, they are linked to blocks in order to fully reconstruct the particle and avoid double-counting the signatures in different subdetectors. For example, tracks are extrapolated to the calorimeters and possibly matched to clusters, if their positions agree when taking into account deflections by multiple scattering as well as the typical longitudinal shower profiles. Potential Bremsstrahlung photons are added to a track by extrapolating tangents to the track to the ECAL and searching for matching clusters. Tracks in the inner tracking and the muon system are linked depending on the quality of a combined track-fit.

Finally, particles are reconstructed from the blocks by applying dedicated quality criteria. After the identification of muons and removal of their tracks from the blocks, electrons and associated Bremsstrahlung photons are formed from tracks and linked ECAL clusters, which are then also removed. Each of the remaining tracks is considered a charged hadron, and its momentum is determined from a combination of the track momentum and linked cluster energies taking into account the detector resolutions. If the cluster energies exceed the track momentum above the calorimeter resolution, photons and possibly neutral hadrons are formed from the excess energy. Likewise, clusters not linked to any track are interpreted as photons and neutral hadrons. In both cases, the ECAL energy is entirely attributed to the photon because in jets on average 25% of the energy is carried by photons but only 3% is deposited in the ECAL by neutral hadrons.

The calorimetric energy linked to charged hadrons is calibrated to correct for the energy-dependent response of the HCAL to hadrons, the different e/h values of the ECAL and HCAL, as well as instrumental effects such as readout thresholds. The corrections are of the order of 20 − 30%. Since the ECAL has been calibrated for photons and electrons, clusters associated with photons and electrons require only small corrections of the order of 1% to account for residual instrumental effects. In both cases, the calibration constants are obtained from simulation and have been validated with collision data [170]. The single-hadron response has been found to be larger in data than in simulation by up to 5% at p_{T} below 30 GeV. Since the calibration affects mostly neutral hadrons, however, which carry only 15% of the energy in a jet [171], the effect reduces to 2% for jets and is corrected for by the dedicated jet energy calibration described in Section 4.5. The photon calibration has been shown to reproduce the π^0 invariant-mass to better than 1% both in data and simulation.

The high granularity of the different subdetectors, in particular the ECAL, is a key element for the applicability of the PF algorithm at CMS. With it, the performance of jet measurements is greatly improved as demonstrated for example in [169, 172] and in Sections 4.4 and 4.6.

4 Reconstruction and Performance of Jets at CMS

In QCD, the dynamics of coloured objects are described by the exchange of gluons. Due to confinement, however, single quarks and gluons cannot be observed as free particles. Instead, they are bound in colour-neutral states because from a certain distance on it is energetically favourable to create new quarks from the energy of the colour field. Consequently, when quarks or gluons are produced in a high-energy particle collision, they will emit further partons resulting in a collimated spray of hadrons referred to as *jet*. The properties of a jet are closely related to the properties of the initial parton, and thus, jets provide experimental access to the underlying partonic process. However, the exact concept of a jet is ambiguous and has to be established by a *jet definition*, a procedure which particles are grouped into a jet and how their momenta are combined.

Jets are the dominating signature at the LHC, cf. Fig. 3.5. They do not exclusively originate in the initial pp collision, though, but jets are also the product of hadronic decays of heavy particles such as the t quark or some yet undiscovered particle. Hence, the study of jets is not only vital in understanding the detector and, of course, probing the SM at the unprecedented energies of the LHC, but also of great importance in the search for the Higgs boson and physics beyond the SM.

Besides jets, more inclusive event-shape variables such as the *thrust* [173, 174] are also used to characterise the energy flow in an event and probe the underlying QCD processes. Only individually reconstructed jets, however, provide the means to study the hadronic decays of heavy objects and are therefore widely used in searches at CMS.

QCD-multijet production and the reconstruction of jets are reviewed in the following Sections 4.1 and 4.2, respectively. In Section 4.3, the simulation of collision events is briefly described. The last sections are dedicated to properties of jets at CMS. In Section 4.4, the jet transverse-momentum response is investigated using simulated events. The derivation of jet-energy-calibration constants by CMS is reviewed in Section 4.5. Finally, in Section 4.6, the jet transverse-momentum resolution is determined for simulated events. It will serve as a reference in later chapters.

4.1 QCD-Multijet Production

Although the production of jets can in principle be described by QCD, difficulties arise from the running of the strong coupling-constant α_s, which becomes large at low energies. Hence, perturbation theory is not applicable to the low-energy processes of the jet evolution for $Q \to \Lambda_\mathrm{QCD}$, cf. Section 2.1.3. The current understanding of jet physics is based on the *factorisation theorem* [175, 176], according to which a perturbative description of the high-energy parton-parton interaction can be combined with a phenomenological model of the low-energy processes. Jet production in this approach is illustrated in Fig. 4.1.

When two protons collide, two of their constituents can interact in what is called the *hard interaction*, a process with a large momentum transfer Q such that α_s is small. Hence, the cross section of the hard interaction can be computed using perturbation theory. In

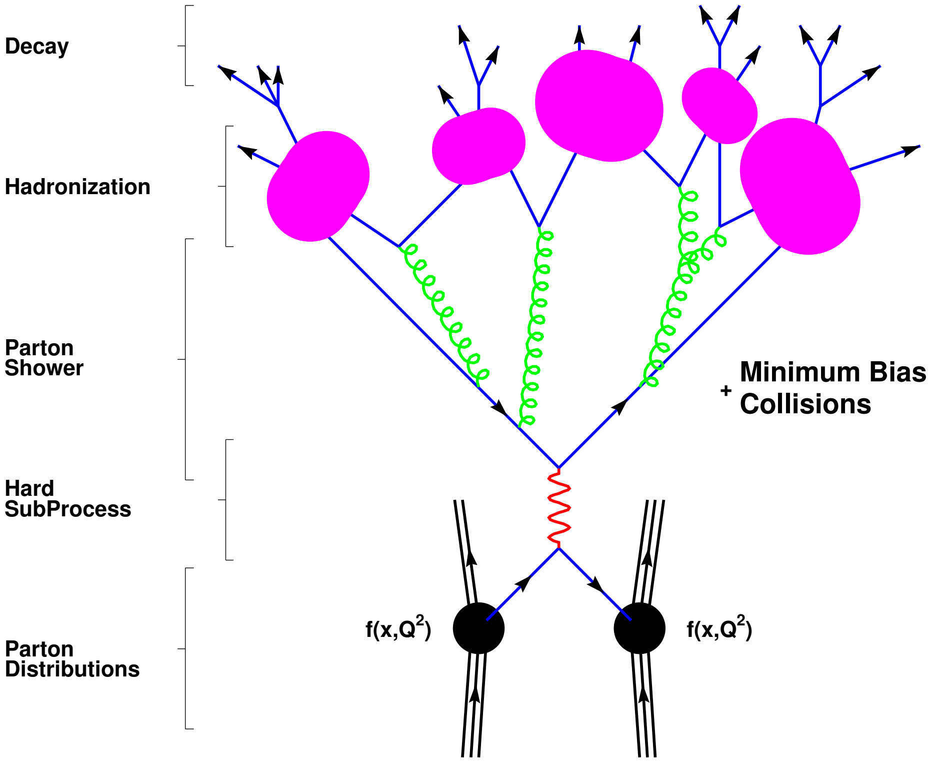


Figure 4.1: Illustration of jet production in a *pp* collision. Taken from [177].

Fig. 3.1, the range of Q accessible at the LHC is shown as a function of x, the fraction of the proton's longitudinal momentum carried by the parton. The probability which type of parton, i, enters the hard interaction depends on the proton structure and is parametrised in the empirically determined *parton-distribution function* (PDF) $f_{i/p}$. A final-state parton will emit further partons, which themselves radiate, a process termed *parton showering*. Eventually, *hadronisation* occurs, and the partons combine to colourless hadrons, which might decay further to stable particles. This collimated bunch of particles, typically mostly hadrons, is referred to as jet. The described steps are discussed further in the following; comprehensive reviews of jet physics can be found for example in [6, 132, 178, 179].

4.1.1 The Initial State: Parton Distribution Functions

Owing to the complex structure of protons, the initial state of a *pp* collision is not exactly determined. Rather, it has to be described by a PDF $f_{i/p}(x, \mu_F)$, the number density of a parton of type i with momentum fraction x inside the proton. Emission of further partons by i after extraction from the proton and before the hard interaction modifies x. Conventionally, this effect is partly included into the proton structure up to the *factorisation scale* μ_F, which thus defines the transition from the non-perturbative to the perturbative regime. The choice of μ_F is arbitrary. The dependence of the PDF on μ_F is given by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [180–182] and does not affect the cross section in the limit that all orders of the perturbative series are considered.

Since the PDFs parametrise soft effects, they cannot be computed in perturbative QCD and are obtained from measurements (although some approaches exist to compute PDFs from lattice QCD [183]). Usually, the results from different experiments and processes, such as deep-inelastic *ep* scattering at the HERA collider [184], are combined to extract the PDFs. A set of PDFs provided by the CTEQ collaboration [185, 186] is shown in Fig. 4.2 for two different values of μ_F. They are gluon dominated except for $x \gtrsim 0.2$, a kinematic regime only probed in case of large $Q \gtrsim 1\,\text{TeV}$, cf. Fig. 3.1, which is why many processes at the LHC are predominantly gluon induced, such as QCD-jet production.

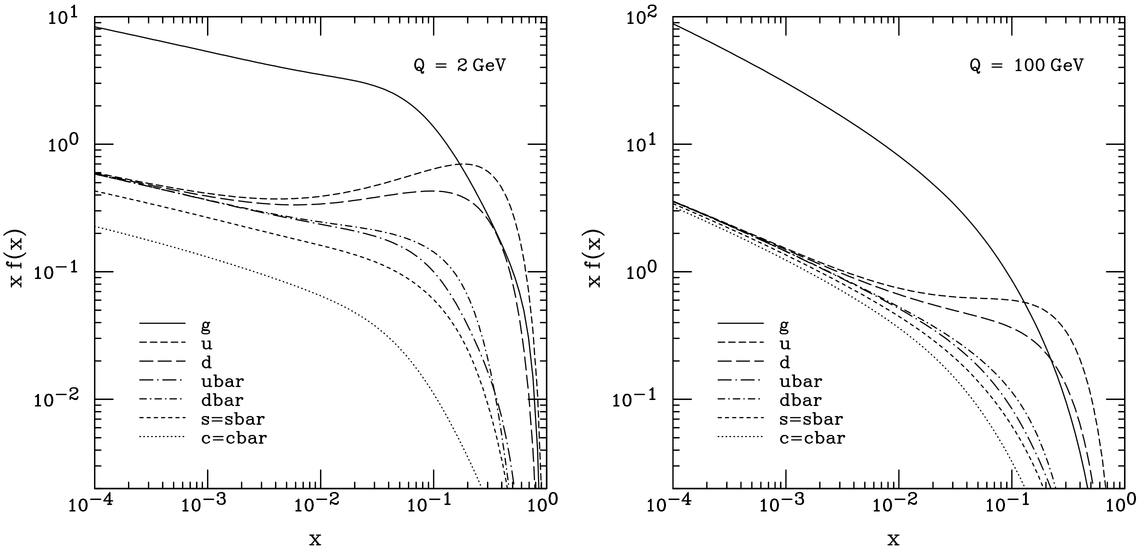


Figure 4.2: CTEQ6M parton-distribution functions, multiplied by x, as a function of x for factorisation scales $\mu_F = Q = 2$ (*left*) and 100 GeV (*right*) for gluons (*solid line*) and quarks (*dashed lines*) as indicated in the legend. Taken from [186].

4.1.2 The Hard Interaction: Inclusive Jet-Production Cross-Section

In LO perturbation theory, jet production in *pp* collisions occurs when two partons interact with sufficient momentum transfer, leading to two final-state partons. In Fig. 4.3, some of the contributing $2 \to 2$ processes are depicted. The inclusive differential jet-production cross-section is given by summation over the cross sections $\mathrm{d}\hat{\sigma}/\mathrm{d}p_\mathrm{T}(ab \to cd, \mu_F, \mu_R)$ of all partonic subprocesses $ab \to cd$, convoluted with the PDF $f_{i/p}(x, \mu_F)$ of the protons 1, 2,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_\mathrm{T}}\left(pp \to \text{jets}\right) = \sum_{abcd} \int\!\!\int \mathrm{d}x_1 \mathrm{d}x_2 \; f_{a/p}\left(x_1, \mu_F\right) f_{b/p}\left(x_2, \mu_F\right) \frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}p_\mathrm{T}}\left(ab \to cd, \mu_F, \mu_R\right) \;, \tag{4.1}$$

where μ_R denotes the renormalisation scale, cf. Section 2.1.3. Both μ_R and μ_F are typically set equal to Q. The inclusive jet cross-section measured by CMS in $4.7\,\mathrm{fb}^{-1}$ of *pp* collision data [187] is shown in Fig. 4.4.

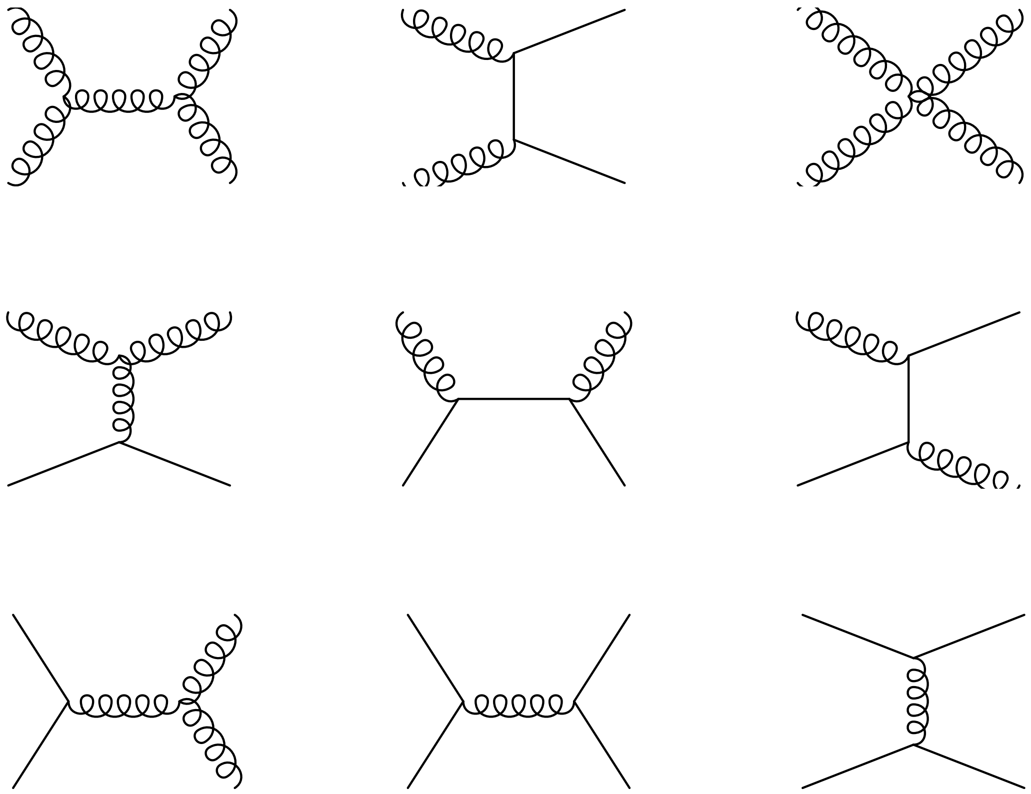


Figure 4.3: Some of the leading-order 2 → 2 diagrams contributing to jet production at the LHC. In case of $qq \to qq$ (*bottom centre and right*), also electroweak bosons can be exchanged.

4.1.3 The Final State: Parton Showering and Hadronisation

After production in the hard interaction, the final-state partons develop a shower, i. e. emit further partons with successively lower energy, which subsequently also radiate. As the energy decreases, α_s becomes larger, and hence, the rate of parton emission increases. When approaching Λ_QCD, the partons eventually hadronise. The time scale of the sketched process is approximately $1/\Lambda_\mathrm{QCD} \approx 10^{-23}\,\mathrm{sec}$ [6].

Since the shower evolution depends to large extents on low-energetic processes which cannot be described perturbatively, they are modelled using empirical *fragmentation functions*, which parametrise the parton showering and hadronisation in analogy to the initial-state description by PDFs. Given the development of a parton shower, the final state typically consists of a few hadrons with large momenta, which are well aligned with the initial parton, and a number of low-momentum hadrons, which can be less aligned, leading to a collimated jet of hadrons. For example, the average p_T of the particles after hadronisation constituting a jet with $p_\mathrm{T} = 500\,\mathrm{GeV}$ is of the order of $10\,\mathrm{GeV}$ [169]. Approximately 65% of the jet's energy is distributed to charged hadrons, mostly pions, 20% to neutral pions, which decay almost instantly into photons[1], and 15% to other neutral hadrons [170–172].

---

[1] The proper mean-lifetime is $(8.4 \pm 0.4) \cdot 10^{-17}\,\mathrm{s}$ for π^0s and $(2.6033 \pm 0.0005) \cdot 10^{-8}\,\mathrm{s}$ for $\pi^\pm$s [6].

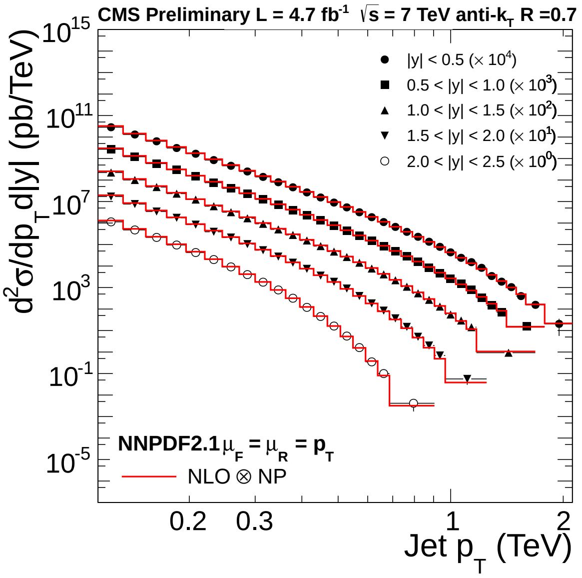


Figure 4.4: Inclusive differential jet-production cross-sections as a function of p_{T} in different intervals of rapidity y, measured by CMS (*markers*) and compared to the next-to-leading-order (NLO) theoretical predictions, corrected for non-perturbative (NP) effects (*solid lines*). The cross sections are scaled by the factors listed in the legend. Taken from [187].

The hadron's momentum components longitudinal to the initial parton scale to first approximation with the parton's energy. Logarithmic corrections occur due to the fact that some energy is required to produce new hadrons and that the multiplicity increases logarithmically with the energy [34]. As a result, high-p_{T} jets are on average narrower: at 20 GeV, approximately 15% of the jet p_{T} is distributed within a radius of 0.1 in $\eta \times \phi$ space around the initial parton while the fraction increases to approximately 90% at 600 GeV [171]. Furthermore, jets are on average more collimated with increasing $|\eta|$, cf. ibid.

The properties of a jet, in particular the multiplicity of its constituents, also depend on the initial parton's type (which defines the *jet flavour*). Due to colour factors, the parton emission rate of gluons is about twice as large as that of quarks. Hence, gluons typically produce more and thus softer particles during the showering than quarks, leading to broader jets [171, 188, 189]. The fraction of gluon jets in QCD events varies with p_{T}. Gluon jets dominate at low p_{T} and quark jets at high p_{T}, cf. Fig. 4.13 (*left*).

Of course, also other particles than partons can be produced during showering, e. g. leptons from heavy-flavour quark decays or photons from electromagnetic interactions of quarks. These processes are rare in comparison, however, due to the size of the strong coupling constant.

According to the above considerations, the process Eq. (4.1) would result in a *dijet event*, a final state with two jets that have opposite transverse momenta due to momentum conservation. However, high-energy partons can be emitted under large angles before and after the hard interaction and produce further, separated jets resulting in a multijet final-state. This is referred to as *initial-state* (ISR) and *final-state radiation* (FSR), respectively.

The occurrence of additional hard jets is suppressed, however, because it requires the presence of additional vertices with coupling α_s [190]. Multijet events can also occur if for example electroweak bosons are produced in association with the final-state partons and decay hadronically or if heavy particles are produced in the hard interaction and decay to partons as in the decay chains of predicted supersymmetric particles.

In the following, dijet events are understood as events with a topology resembling the ideal dijet case but which might still contain further low-p_T jets.

4.1.4 The Underlying Event

All contributions to an event that are not associated to the hard interaction are referred to as *underlying event* (UE). UE activity occurs if further partons of the two colliding protons scatter (*multiple parton interaction*, MPI) [191] and also originates in the coloured proton remnants. The latter produce hadrons predominantly along the direction of the initial protons. They can be — and are — neglected because they have a tiny transverse component and hence hardly affect the observables in physics analysis. In particular through MPI, however, a substantial amount of soft particles is added to the event content, which has to be considered for example in the jet-energy-scale calibration, cf. Section 4.5. The UE activity increases with Q and the centre-of-mass energy $\sqrt{s}$. On average, the UE contributes a transverse momentum of approximately $1 - 2\,\mathrm{GeV}$ per unit of η and ϕ to events with jets up to $100\,\mathrm{GeV}$ [192].

4.2 Jet Reconstruction

Jets can often easily be identified by eye as confined sprays of particles, cf. Fig. 4.5. Their exact properties are ambiguous, however, in particular in events with a lot of activity. Hence, jets are defined by a jet definition, which consists of a *jet algorithm*, a description how to cluster jets out of a list of particles, and a *recombination scheme*, which determines how to compute the jet's four-momentum from its constituents. At CMS, the four-momentum sum of all constituents is assigned to the jet, and this scheme is assumed in the following discussion. Desired features of jet algorithms are discussed in Section 4.2.1, and typical examples are presented. Afterwards, in Section 4.2.2, the measurement of jets at CMS is reviewed.

4.2.1 Jet Algorithms

Jet algorithms define how to combine particles into a jet, usually by some distance measure to decide whether two particles belong to the same jet. They have evolved due to advanced theoretical understanding and requirements as well as increased experimental precision and computing capacities. General requirements on good jet algorithms were collected in the 'Snowmass accord' [193]. Jet algorithms should be simple to implement in experimental and theoretical analysis and applicable to any relevant type of input objects, e. g. final-state hadrons or energy deposits in calorimeter cells. Very importantly, good jet algorithms also have to result in jets that are insensitive to the emission of soft particles (*infrared safety*) and to collinear splitting of particles (*collinear safety*) in the event, termed 'IRC safety'. Otherwise, as soft emission and collinear splittings occur randomly in QCD events, cross sections would depend on the specifics of hadronisation effects. Hence, it would be

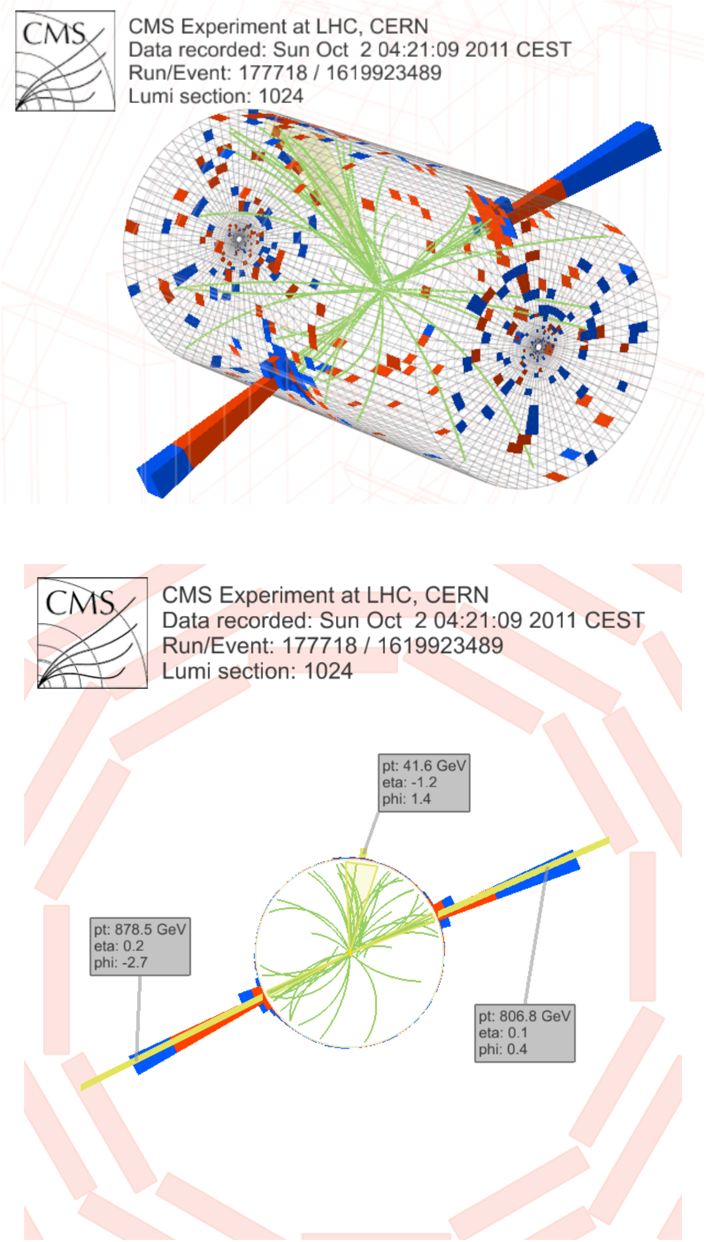


Figure 4.5: Visualisation of the detector signature caused by a typical dijet event at CMS with an average transverse momentum of 842.7 GeV in a three-dimensional (*top*) and a transverse view (*bottom*). (The latter is a projection of the barrel-detector signals onto the transverse plane.) Red and blue cuboids represent towers of the electromagnetic and hadronic calorimeter, respectively, with their height being proportional to the measured transverse energy, and solid green lines represent reconstructed particle tracks. The momenta of clustered Particle-Flow jets are indicated by the solid yellow lines, and their kinematic properties are stated in the grey boxes. Seven primary vertices have been reconstructed.

difficult to compare experimental results with theoretical predictions, in particular also since jets clustered from detector-related quantities such as calorimeter cells provide IRC safety to some extent due to the finite detector resolution. Furthermore, QCD predictions can only be computed to a fixed order in perturbation theory, and calculations rely on the cancellation of divergences from IRC processes and from loop contributions. Cancellations are not guaranteed, however, if jets are sensitive to IRC processes, leading to infinite cross sections.

In the following, the two main categories of jet algorithms, *cone algorithms* and *sequential recombination algorithms*, are described, and specific examples are given. An extensive review of jet algorithms and their properties can be found in [194].

4.2.1.1 Cone Algorithms

Cone algorithms are based on the concept that, due to the dynamics during parton showering and hadronisation, a jet is spatially confined into a cone-shaped region with an axis along the direction of the initial parton. The first jet algorithm, which was developed to identify dijet events in e^+e^- collisions, was in fact a cone algorithm depending on two parameters ϵ and δ [195]: an event was accepted, if more than a fraction $1-\epsilon$ of the energy was contained in two cones of opening angle δ. Most modern variants of the cone algorithm determine the cone's direction in an iterative approach and define a jet as the sum of the four-momenta of all objects inside a stable cone. The number of jets per event is usually not limited, and a procedure has to be chosen which handles possibly overlapping cones.

An example of a common cone algorithm is the *iterative cone* (IC) algorithm. It is applied to a list of all input objects, e. g. calorimeter cells, and the object with highest p_T is taken as seed s. A proto-jet is defined as sum of the four-momenta of all objects k with

$$\Delta\mathrm{R}_{ks} = \sqrt{\Delta\eta_{ks}^2 + \Delta\phi_{ks}^2} < \mathrm{R} \ ,$$

i. e. all objects inside a cone of radius R. The proto-jet is used as new seed, and the procedure is repeated until the proto-jet's direction and p_T change by less than some thresholds. Then, the proto-jet is considered a jet, all constituents k are removed from the input list, and the next jet is built from the remaining objects.

However, the IC algorithm, like other seeded algorithms with progressive removal of input objects, does not meet IRC safety. Nonetheless, in the pre-LHC era cone algorithms were widely in use in particular at hadron colliders because the existing sequential recombination algorithms were computationally too slow for the dense environment in hadron collisions.

There exist also cone algorithms which require no explicit seed but consider all possible stable jet cones instead and thus can be constructed in an IRC-safe way, but these seedless algorithms are typically absurdly slow[2]. A practical implementation, the *seedless infrared-safe cone* (SISCone) algorithm, was presented in 2007 [196], which requires a time of $\mathcal{O}(N^2 \ln N)$ to find a jet out of N particles. Although this is sufficient even for LHC conditions, it is still about 100 times slower than current implementations of sequential

---

[2] The time to cluster jets out of N input objects scales as $N2^N$ for original seedless cone-algorithms. For example, it would take about 10^{17} years to cluster 100 particles [194].

recombination algorithms, cf. e. g. Fig. 11 in [194], and therefore, the SISCone algorithm is not used at CMS.

4.2.1.2 Sequential Recombination Algorithms

Sequential recombination algorithms iteratively cluster the objects which are closest according to some distance measure and thus reflect the process of parton showering. At hadron colliders, commonly the distance d_{ij} between all pairs of objects i, j and the distance d_{iB} between all objects i and the beam are defined as

$$\begin{aligned} d_{ij} &= \min\left(k_{\mathrm{T},i}^{2p}, k_{\mathrm{T},j}^{2p}\right) \frac{\Delta \mathrm{R}_{ij}^2}{\mathrm{R}^2} \\ d_{iB} &= k_{\mathrm{T},i}^{2p} \, . \end{aligned}$$

Here, $\Delta\mathrm{R}_{ij}$ is the distance of i and j in $\eta \times \phi$ space and $k_{\mathrm{T},i}$ the transverse momentum of i. R and p are free parameters of the algorithm. The smallest of all d_{ij} and d_{iB} is determined. If it is d_{ij}, the objects i and j are combined into a new object by addition of their four-momenta, and all distances are recomputed. If d_{iB} is the smallest distance, i is considered a jet and removed from the list of input objects. The procedure is repeated until all objects are clustered.

Three algorithms are distinguished, depending on the value of p: the k_T algorithm with $p = 1$ [197], the *Cambridge/Aachen* (C/A) algorithm with $p = 0$ [198, 199], and the *anti-k_T* algorithm with $p = -1$ [200]. All are IRC safe. Although originally developed for the much cleaner environment of e^+e^--collision events, efficient implementations are provided by the FastJet package [201, 202]. With these, the time to cluster jets out of N particles scales with $\mathcal{O}(N \ln N)$, which allow an application also to hadron-collision events.

In case of the k_T algorithm, which clusters softer objects first, and in case of the C/A algorithm, which clusters objects separated by small angles first, the clustering sequence closely emulates the branching sequence during parton showering. Therefore, both algorithms are suited to investigate the substructure of jets and disentangle for example merged jets from the decay of boosted heavy objects by reversing the last clustering step [203–206]. In contrast to cone algorithms, both might result in irregular shaped jets as illustrated in Fig. 4.6.

The anti-k_T algorithm, on the other hand, starts with clustering harder objects first and resembles a more intuitive clustering order, leading to circular jets of radius R, cf. Fig. 4.6. It is the standard jet algorithm adopted by CMS, and in this thesis, the anti-k_T algorithm with $\mathrm{R} = 0.5$ is used.

4.2.2 Jet Types at CMS

Jets are measured using the four-momenta reconstructed from detector signals as input to a jet algorithm. These jets will be called *detector-level* jets, as opposed to *particle-level* jets one would cluster if the jet algorithm was applied directly to the final-state particles after hadronisation. As mentioned above, the jets' four-momenta are obtained as sum of the input objects' four-momenta.

At CMS, different types of detector-level jets are reconstructed, depending on the utilised subdetectors. In each case, the algorithm implementations of the FastJet package are used

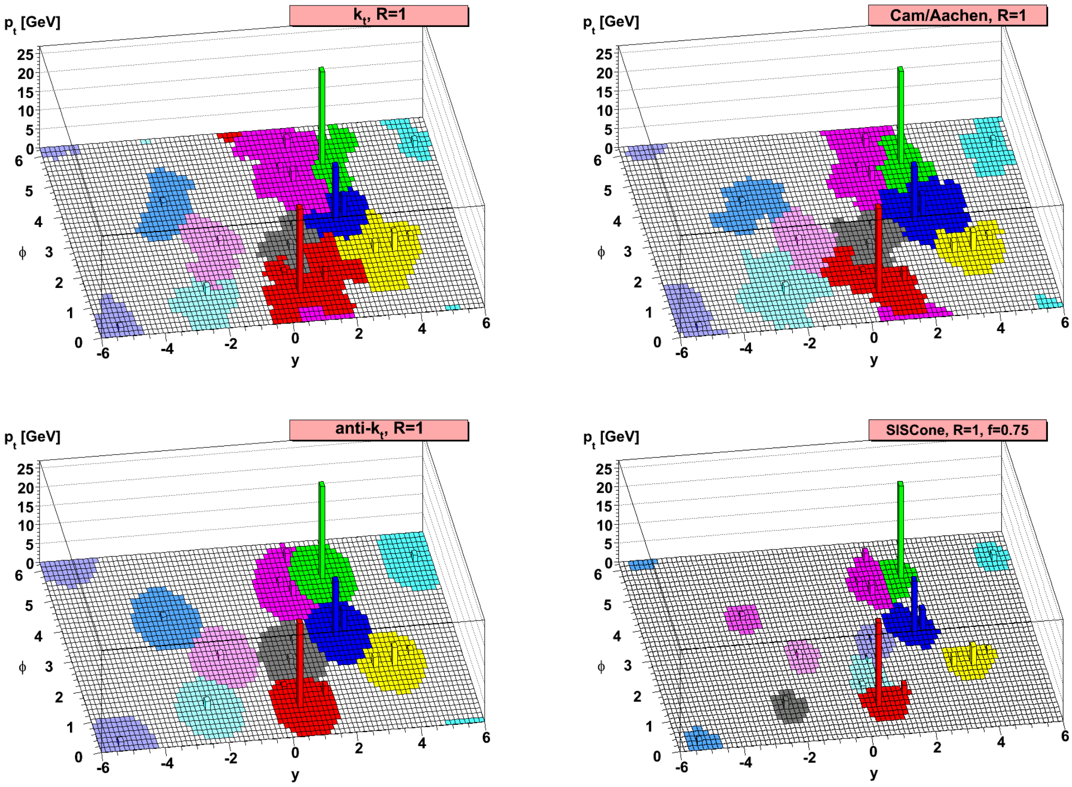


Figure 4.6: Particle-level jets generated with HERWIG++ [207] and clustered with the k_T (*top left*), C/A (*top right*), anti-k_T (*bottom left*), and SISCone (*bottom right*) algorithm. Many random soft particles have been added to the events to illustrate the 'catchment' areas of the jets. Taken from [200].

for jet finding. As a convention, the jets in one event are ordered in p_T, with the first (*leading*) jet being the jet with highest p_T. Two important jet types are described in the following.

Calorimeter (Calo) jets are clustered from the energy deposits in calorimeter towers which comprise one HCAL cell and the covered array of 5 × 5 ECAL cells. The tower's four-momenta are computed taking the direction from the interaction point to the tower centre and assuming zero mass. All towers with a transverse-energy measurement above 300 MeV are considered in the clustering process. Calo jets are relatively simple objects because only calorimeter information is deployed, but they are strongly affected by the non-linearity of the calorimeters. Since the readout of calorimeter measurements is fast, Calo jets are commonly used by the trigger system.

Particle-flow (PF) jets are clustered from the four-momenta of the particle candidates reconstructed by the Particle-Flow algorithm, cf. Section 3.5.2. They have a superior performance compared to calorimeter jets regarding the response and resolution because HCAL contributions to the measurement are limited since only about 15% of the jet energy is distributed among neutral hadrons [169]. Also, since individual candidates are considered, the impact from the calorimeter non-linearity is reduced. Therefore, PF jets are the standard jet type for analysis at CMS. The reconstructed PF jets in a typical dijet event at CMS are visualised in Fig. 4.5.

Furthermore, *track jets* [208], which are clustered from tracks only, and *jet-plus-track jets* [209], which are Calo jets complemented by tracking information, are supported by CMS, but these types are not relevant for this thesis.

4.3 Event Simulation

The simulation of particle collisions and the signatures of the subsequent particles in the detector constitutes an important tool to obtain expectations for physics properties of the collected data and the detector performance. Results from simulated events frequently serve as a guideline in the development of measurement techniques and as a reference to validate their consistency. Sometimes, certain quantities required for the interpretation of a measurement such as the geometric acceptance of the detector or the flavour composition of jets are not accessible in data, e. g. due to insufficient instrumental resolution or lack of statistical precision. In that case, the missing information has to be obtained entirely from simulation, in general at the cost of larger systematic uncertainties though.

Typical simulations involve the modelling of stochastic processes as well as the computation of complicated integrals which are insolvable analytically and which cannot be approximated by numerical methods on a reasonable time scale. Therefore, *Monte Carlo* techniques are exploited which essentially exploit pseudo-random numbers to both simulate individual processes and perform a random sampling of the integrand [179, 210]. In the following, the label 'MC' will frequently serve to denote quantities and results derived from simulation.

4.3.1 Event Generation

The simulation of events is usually separated into several sequential steps, as described extensively for example in [211]. First, the cross section of the hard scattering process of interest is computed for which the matrix elements up to a certain order are considered. For instance, the programme Pythia [212] calculates hard processes with two, and in few cases three, particles in the final state and Madgraph [213] with up to nine. Hence, depending on the definition of the final-state signature, the additional emission of partons can be included to some extent into the simulation at matrix-element level.

The subsequent parton showering cannot be computed from first principles for technical reasons because the number of diagrams increases dramatically with each order of the perturbative series, and, more fundamentally, the energy of the involved partons decreases with increasing order such that perturbation theory is not applicable anymore at some point. Therefore, phenomenological models are used to simulate parton showering and hadronisation as well as the UE activity, following the factorised approach outlined above in Section 4.1. An advanced simulation of these processes is implemented in Pythia, for example, which is commonly interfaced by other generators such as Madgraph. In particular the UE description is difficult, and programmes have to be tuned to the conditions at a certain collider. The simulated particles emerging from the hadronisation are referred to as *generator-level* particles, and associated quantities such as the transverse momentum will be denoted with the superscript ‘gen’.

Finally, the interaction of the generator-level particles with the detector material and the resulting readout signals are simulated with a model of the CMS apparatus based on the GEANT4 programme [214]. This model will be referred to as *full simulation.* In certain situations when particularly large samples are required, a *fast simulation* tool of the CMS detector [215] is used instead, which exploits a number of simplified parameterisations and hence requires less computing resources. The same reconstruction methods can be applied to both data and simulated detector output allowing for a direct comparison of the detector-level objects.

In the following, if results are described to have been obtained from a certain generator programme this will mean, unless stated otherwise, that the full detector simulation is applied to model the detector-level objects.

4.3.2 Definition of a QCD-Multijet Reference Sample

If not stated differently, a sample of 10930800 QCD-multijet events is used throughout this thesis, which have been generated with Pythia 6.4.24 with the CTEQ6L1 PDF set [186] and the tune Z2 described in [192] which also considers CMS soft-p_{T} data. Events have been processed through the full detector simulation.

At generation, the QCD cross-section has been scaled by $\hat{p}_{\mathrm{T}}^{4.5}$, where $\hat{p}_{\mathrm{T}}$ is a scale parameter for the momentum transfer in the hard interaction. That way, the decrease of the event rate is softened and the statistical precision remains comparable over a large range in $\hat{p}_{\mathrm{T}}$. For the shown results, if not stated otherwise, events are weighted inversely such that the realistic QCD spectrum is regained as illustrated in Fig. 4.7. The spectrum simulates the jet cross-section Fig. 4.4.

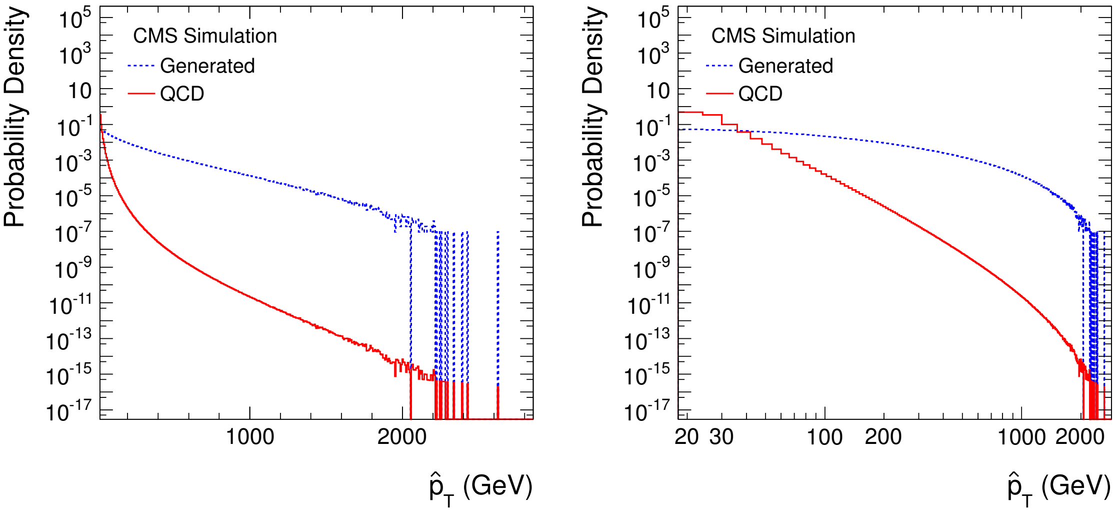


Figure 4.7: $\hat{p}_{\mathrm{T}}$ spectrum of the Pythia QCD-multijet sample used in this analysis on a single- (*left*) and double-logarithmic (*right*) scale. The decrease has been softened at generation by scaling the cross section by $\hat{p}_{\mathrm{T}}^{4.5}$ (*dashed line*). The QCD spectrum is regained by applying the inverse event weight (*solid line*).

In addition to the primary *pp* collision, pile-up collisions have been generated. Their number is distributed uniformly between zero and ten, and decreasing according to a Poisson distribution with mean value 10 for larger multiplicities. Therefore, in order to model the data, an event weight is assigned depending on the number of simulated pile-up interactions [216]. Since this number does not correlate with $\hat{p}_{\mathrm{T}}$, the pile-up weight can be multiplied to the spectral weight. The target pile-up distribution expected for the data is depicted in Fig. 4.8 (*left*). It is derived from the LHC machine conditions during the considered data-taking period together with the total *pp* inelastic cross-section. This approach does not depend on any selection effects.

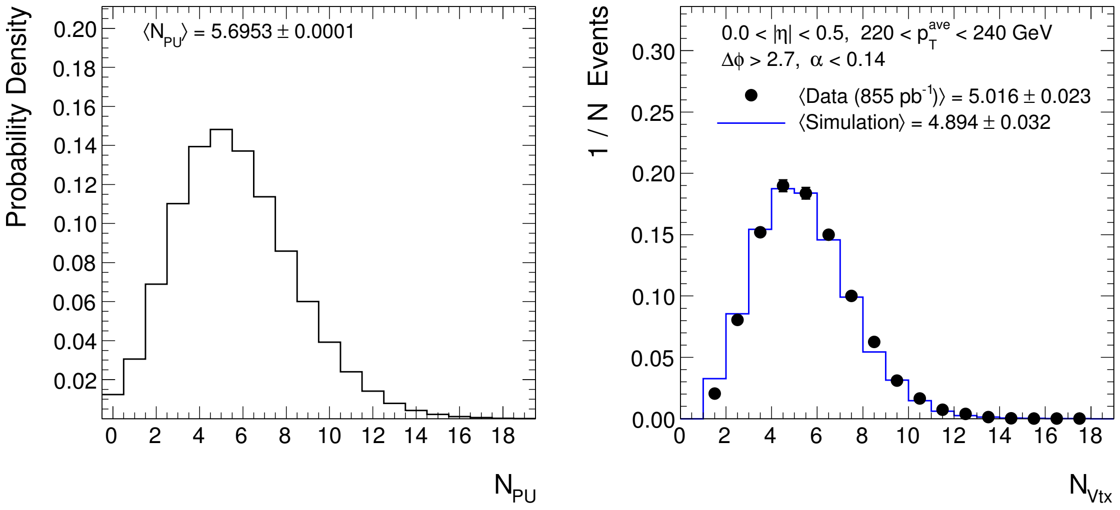


Figure 4.8: Target pile-up distribution expected for the $855\,\mathrm{pb}^{-1}$ of data used in the resolution measurement Chapter 5 (*left*), taken from [216], and the distribution of reconstructed primary vertices (*right*) after a dijet selection. (The stated dijet-selection criteria will be explained later in Section 5.1.5.) The simulation (line), reweighted to the data pile-up conditions, is compared to the data (circles).

A cross-check of how well the pile-up profile is modelled is obtained from the distribution of the number of reconstructed primary vertices, which corresponds to the distribution of pile-up interactions up to the vertex reconstruction efficiency. The data is reasonably well described by the reweighted simulation as demonstrated in Fig. 4.8 (*right*).

The full PF event-reconstruction is performed on the simulated detector output. Detector-level jets are clustered from the reconstructed particles with the anti-k_T algorithm, cf. Section 4.2.1, with size parameter $R = 0.5$. Their four-momenta are corrected on average by applying the full chain of jet energy calibration factors discussed below in Section 4.5. Furthermore, generator-level jets are clustered by applying the same algorithm to all generator-level particles.

In the following, if not stated otherwise, 'PYTHIA sample' or 'simulation' will be used to refer to the described sample.

4.4 Jet Transverse-Momentum Response

In general, the energy measured in the detector does not equal the energy of the original particles. For jets, this effect is quantified by the *jet response*, which in this thesis is defined as the fractional signal measured in the detector for a given amount of energy of the particles emerging from the hadronisation of an initial parton. Given the kinematic constraints at a hadron collider, more specifically the *jet transverse-momentum response*, $\mathcal{R}$, will be used, which is understood as the ratio of the measured transverse momentum, p_{T}, of a jet at detector-level and the transverse momentum, $p_{\mathrm{T}}^{\mathrm{particle}}$, of the underlying particle-level jet,

$$\mathcal{R} = \frac{p_{\mathrm{T}}}{p_{\mathrm{T}}^{\mathrm{particle}}} \,. \tag{4.2}$$

The average response is denoted as *jet energy scale*[3]. After calibration, the jet energy scale should be $\langle\mathcal{R}\rangle = 1$. The width of the response distribution corresponds to the relative *jet transverse-momentum resolution*.

In Section 4.4.1, a description is given to determine $\mathcal{R}$ from simulated events. Fluctuations of $\mathcal{R}$ are discussed in Section 4.4.2, and its dependence on p_{T}, η, and the presence of pile-up collisions is investigated. The jet energy scale and p_{T} resolution are defined more precisely in Section 4.4.3, and the scale is shown in Section 4.4.4. Finally, in Section 4.4.5, techniques to measure the jet-p_{T} response from data are reviewed.

In some cases that are specifically noted, the jet-p_{T} response and related quantities will be shown also after application of the jet energy calibration, which is only described later in Section 4.5. This way, an understanding can be gained of the performance of jets as used in physics analysis at CMS. The focus when discussing general features of the response will be on uncalibrated jets, though.

4.4.1 Determination of the MC-Truth Response

In simulated events, the particle-level jet is represented by the generator-level jet with transverse momentum $p_{\mathrm{T}}^{\mathrm{gen}}$. Thus, $\mathcal{R}$ can be obtained straight-forward as

$$\mathcal{R} = \frac{p_{\mathrm{T}}}{p_{\mathrm{T}}^{\mathrm{gen}}} \,. \tag{4.3}$$

It is referred to as *MC-truth response* and is determined from the Pythia sample as follows.

In each event, the leading two generator-level jets g_i, $i = 1, 2$, are selected, i. e. the two jets with the highest $p_{\mathrm{T}}^{\mathrm{gen}}$. To each selected g_i, the detector level jet j_i closest in $\Delta\mathrm{R}$ is matched, where $\Delta\mathrm{R} = \sqrt{(\Delta\eta(g_i, j_i))^2 + (\Delta\phi(g_i, j_i))^2}$ denotes the distance of g_i and j_i in $\eta \times \phi$ space. The event is discarded, if for one of the two pairs $\Delta\mathrm{R} > \Delta\mathrm{R}_{\mathrm{max}}$. A value of $\Delta\mathrm{R}_{\mathrm{max}} = 0.1$ has been chosen as will be motivated later in Section 4.6.1. For each pair, the response $\mathcal{R} = p_{\mathrm{T}}/p_{\mathrm{T}}^{\mathrm{gen}}$ is recorded in intervals of $p_{\mathrm{T}}^{\mathrm{gen}}$ and $|\eta^{\mathrm{gen}}|$. Example response distributions for two different $p_{\mathrm{T}}^{\mathrm{gen}}$ intervals are shown in Fig. 4.9.

A dependence of $\mathcal{R}$ on the jet's momentum and pseudorapidity is expected because both the response and resolution of the calorimeters and tracking devices, and in this case

---

[3] Often, jet energy scale and response are used as synonyms. Here, response always refers to the relative signal generated by one individual jet.

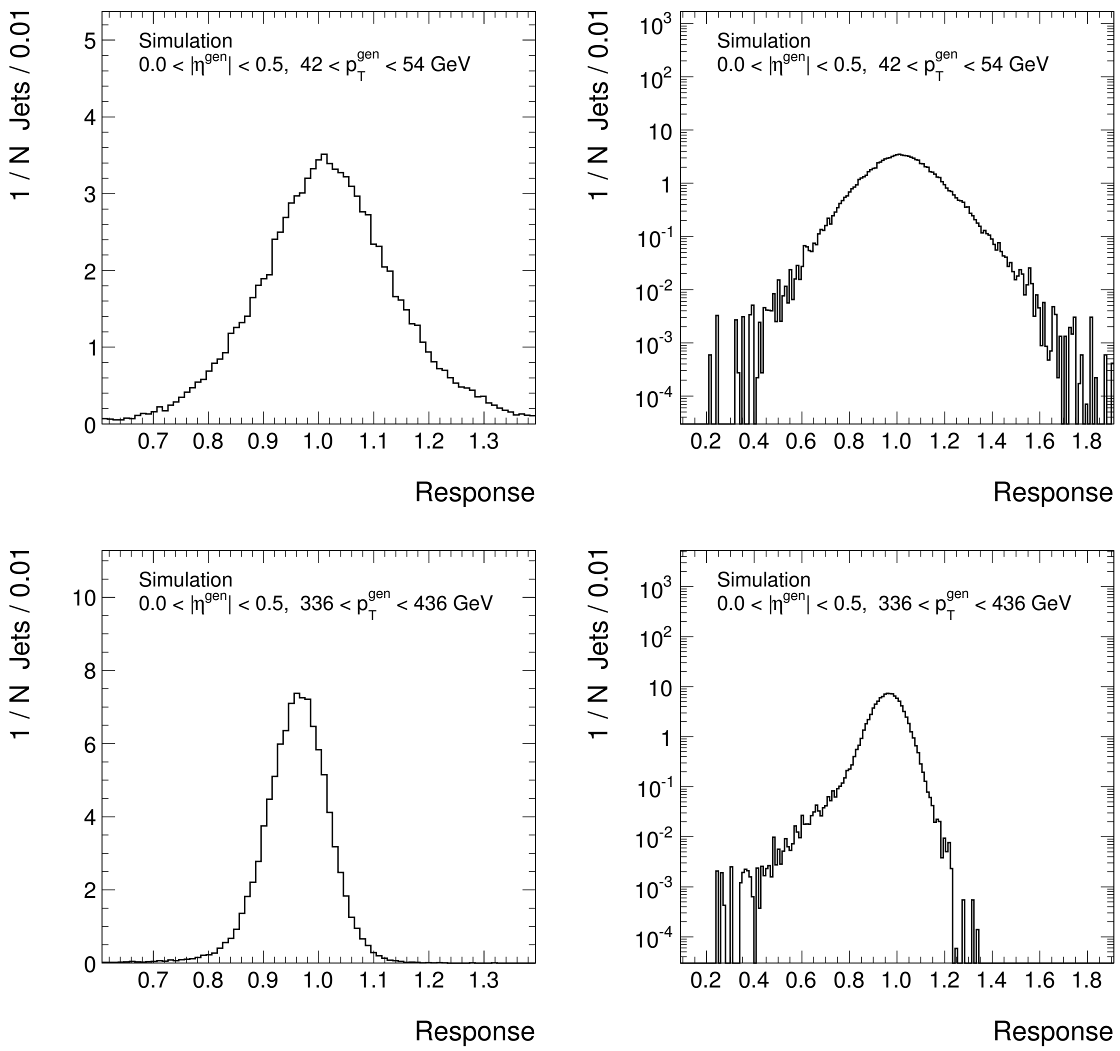


Figure 4.9: MC-truth response distribution before application of the jet energy calibration described in Section 4.5 in a low (*top row*) and a medium (*bottom row*) $p_{\mathrm{T}}^{\mathrm{gen}}$ interval on a linear (*left*) and logarithmic (*right*) scale.

also the track-reconstruction efficiency, depend on the energy. Furthermore, the various subdetectors of different design utilised in the jet reconstruction have different acceptances in η, and likewise different amounts of material that affect the measurement are present in front of the subdetectors.

Here, events are not weighted to the realistic QCD cross-section in order to ensure that the assigned statistical uncertainties reflect the actual number of generated events. Nonetheless, the pile-up weights are applied because the pile-up scenario impacts on the response. However, the induced bias of the statistical uncertainties is much smaller and independent of p_{T}.

$\mathcal{R}$ is recorded in intervals of generator-level quantities in order to avoid biases that occur if detector-level quantities are used instead. Selecting events in intervals of measured

p_{T} leads to migration effects at the boundaries due to the finite jet-p_{T} resolution, i. e. there are jets that fluctuate either into or out of the interval. Additionally, because of the monotonically decreasing jet-p_{T} spectrum, in any interval there are more jets that fluctuated high than jets that fluctuated low in p_{T}. Hence, the selected sample is biased towards jets of lower $p_{\mathrm{T}}^{\mathrm{gen}}$ that fluctuated high in the detector. This is illustrated in Fig. 4.10.

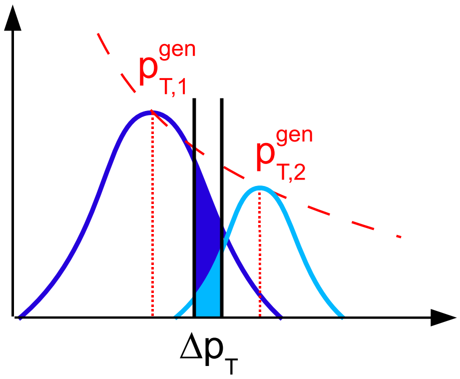


Figure 4.10: Illustration of the bias that occurs if events are selected in intervals Δp_{T} of measured jet-p_{T}. Due to the falling $p_{\mathrm{T}}^{\mathrm{gen}}$ spectrum, more events with low than with high $p_{\mathrm{T}}^{\mathrm{gen}}$ migrate into the Δp_{T} interval.

Migration effects also occur with respect to η but to a smaller extent since the resolution of typically $\delta\eta = 0.005 - 0.03$ [217] is better and the spectrum is more flat for η than for p_{T}. Even more, the impact on the response is much smaller because $\mathcal{R}$ depends only indirectly on η.

4.4.2 Components of the Jet Response

The response distributions in Fig. 4.9 are dominated by a Gaussian-like central region. In particular at low response, additional non-Gaussian components are present, which are of $\mathcal{O}(10^{-3} - 10^{-2})$ of the bulk contribution and are denoted *(non-Gaussian) tails* in the following.

It should be kept in mind for the following discussion that the momenta of the considered detector-level jets are not computed from raw detector signals. Rather, the signals have been precalibrated directly after readout at the simulated hardware-level. Moreover, the calorimeter clusters entering the event reconstruction have been further calibrated as part of the PF algorithm, cf. Section 3.5.2.

For low $p_{\mathrm{T}}^{\mathrm{gen}}$, the average response is a few percent larger than 1 due to the presence of particles originating from pile-up interactions. The average additional energy clustered into jets has been measured to be about 5 GeV in case of eight pile-up collisions [172] as shown in Fig. 4.11. Consequently, limiting the number of pile-up collisions reduces the response: the average response of jets in the same $p_{\mathrm{T}}^{\mathrm{gen}}$ interval from events with less than two pile-up collisions is 96% as demonstrated in Fig. 4.12. Since the pp cross-section is steeply decreasing with p_{T} [218], the occurrence of high-p_{T} pile-up particles is rare, which minimises the absolute offset in the measured jet momentum. Hence, the response at larger $p_{\mathrm{T}}^{\mathrm{gen}}$, with an average of 96%, is hardly affected by the presence of pile-up, cf. Fig. 4.12.

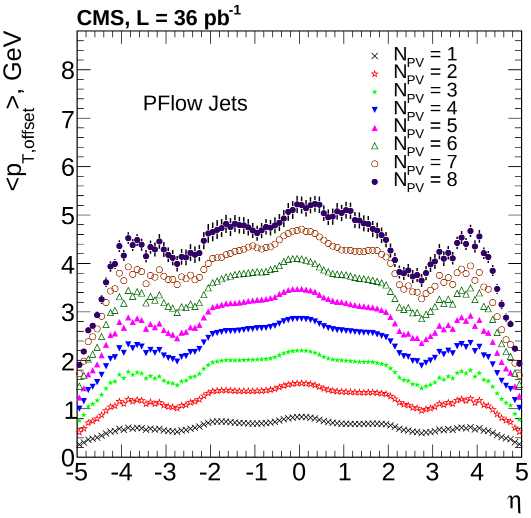


Figure 4.11: Average increase $\langle p_{\mathrm{T,offset}} \rangle$ in the jet's reconstructed transverse momentum due to the additionally deposited energy from instrumental noise and pile-up measured in minimum-bias events, where the pile-up conditions are characterised by the number N_{PV} of reconstructed primary vertices. The noise contribution is quoted to be less than 250 MeV. Taken from [172].

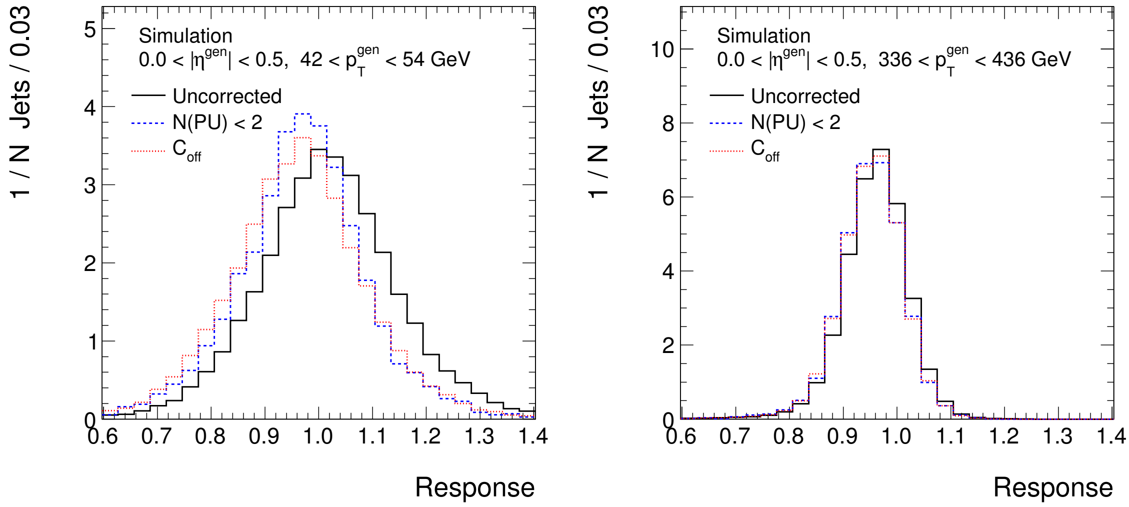


Figure 4.12: MC-truth response in a low (*left*) and a medium (*right*) $p_{\mathrm{T}}^{\mathrm{gen}}$ interval for uncalibrated jets in all events (*solid histogram*) and in events with less than two pile-up collisions (*dashed histogram*), as well as for all events but after application of the offset jet energy correction (*dotted histogram*) described in Section 4.5.

The fluctuations of the response around the mean are due to a combination of the intrinsic resolution of the different subdetectors and the performance of the PF and jet algorithms described in Chapter 3. To a smaller extent, also the flavour composition of the QCD sample contributes. Some aspects of energy and momentum measurements with calorimeters and tracking devices have been summarised in Section 3.2; extensive reviews can be found for example in [134, 135, 219].

The relative p_{T} resolution of the track-based measurements is better than 2% for p_{T} up to 100 GeV, cf. Section 3.4.1. It is caused by the curvature error due to multiple scattering and by the finite hit-position resolution, cf. Eq. (3.2). This latter contribution starts to contribute significantly for $p_{\mathrm{T}} \gtrsim 10\,\mathrm{GeV}$ and leads to a degradation of the relative track-p_{T} resolution proportional to p_{T}. The relative calorimetric energy-resolution evolves in an opposite way and improves with increasing p_{T}, cf. Eq. (3.3). At high p_{T}, it is ultimately limited by miscalibration and non-uniformities of the calorimeters, at medium p_{T}, stochastic fluctuations of the shower development dominate, and at low p_{T}, the calorimeter resolution is dominated by electronic noise in the readout system. The ECAL has an excellent relative p_{T} resolution to electrons and photons of 0.6% at 50 GeV, cf. Eq. (3.4). However, the relative resolution of the calorimetry to hadrons is worse, approximately 18% in case of 50 GeV pions, cf. Eq. (3.5), because additional effects of hadronic showers degrade the resolution. The response of the HCAL to electrons is different than that to hadrons ($e/h = 1.4$), and the electromagnetic fraction f_{em} fluctuates from shower to shower. Furthermore, also the fraction of invisible energy fluctuates. Both effects lead to a variation of the response to hadrons. The impact on the jet response is further complicated because a jet usually contains several hadrons and their showers might overlap.

The impact of the poor HCAL resolution is minimised, however, by the PF event reconstruction. As mentioned above, 85% of the jet energy is carried by charged particles and photons, and their momenta are measured combining tracking and ECAL information. The HCAL contribution is sizable only in the forward region which is not covered by the tracking detector and at very high p_{T}, where the calorimeter resolution is superior and dominates the jet p_{T} measurement.

Another source is the dependence of the response on the jet flavour. Since gluon jets are composed of more and thus softer particles than quark jets, as mentioned previously, they have a smaller response on average due to the non-linearity of the calorimeters. The response of heavy-flavour (b, c) quark jets, on the other hand, is lower on average than that of light-flavour (u, d, s) quark jets because the former frequently decay semi-leptonically into neutrinos which escape undetected and thus carry away energy. Furthermore, the flavour composition of the QCD-multijet sample depends on p_{T}, cf. Fig. 4.13 (*left*). Since PF jets are clustered from individually reconstructed particles, however, the impact of the calorimeter non-linearity is minimised, and the average response differs by less than 4% between the jet flavours, as shown in Fig. 4.13 (*right*).

The performance of the PF algorithm itself also impacts the jet-p_{T} resolution [169]. At low p_{T}, important contributions arise from the efficiency for track identification. It is better than 90% for charged hadrons inside jets; an overall reduction by 5% has been found to reduce the jet response by 2% at low p_{T}. At high p_{T}, uncertainties in the single-hadron response corrections affect the jet response by 2%. Furthermore, studies of the 'Pandora' PF-algorithm designed for detectors at a future linear collider have shown that the limited ability to correctly match calorimetric energy depositions to reconstructed particles constitutes an important contribution to the jet-p_{T} resolution [221].

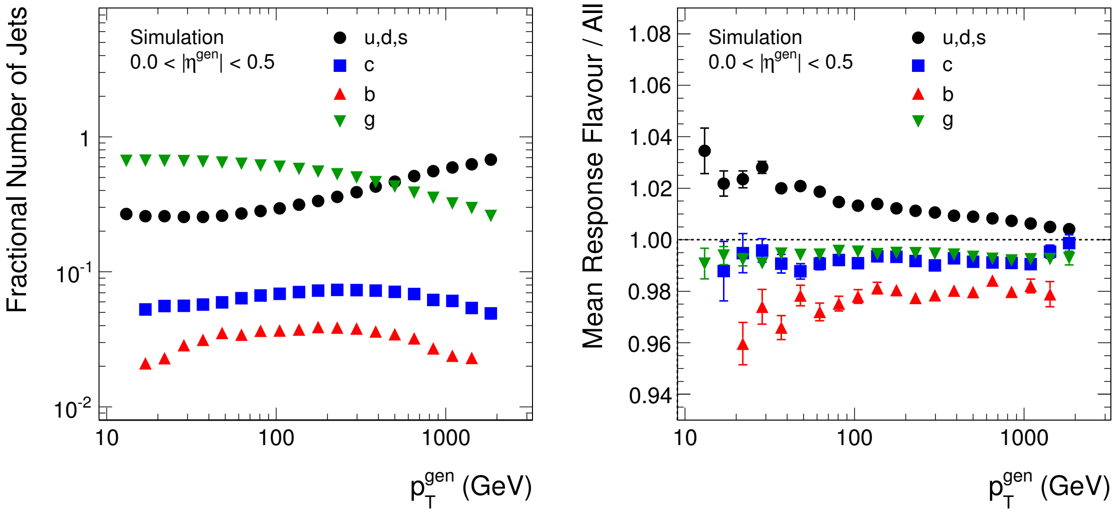


Figure 4.13: Fractional number of jets of a certain flavour (*left*) and their average response relative to the average response of all jets (*right*) as a function of $p_{\mathrm{T}}^{\mathrm{gen}}$ in simulated events. The average response is defined as mean of a Gaussian fit to the MC-truth response of the leading two generator-level jets, as described in Section 4.4.3, in events where both jets meet $|\eta^{\mathrm{gen}}| < 0.5$. The jet flavour has been determined using the 'algorithmic definition' described in [220]. Error bars represent the statistical uncertainties.

Finally, the choice of the jet algorithm and its parameters contribute to the response. Due to the statistical nature of the parton-showering and subsequent hadronisation process, fluctuations occur in the number and momenta of the resulting particles. Thus, some particles might not be clustered into a detector-level jet because they are too soft to be reconstructed or fall out of the geometric reach of the jet. On the other hand, excess energy might be clustered from the underlying event or from overlapping additional jets. Hence, even if a detector-level jet could be perfectly reconstructed, i. e. if the jet resolution due to detector effects could be neglected, its energy would fluctuate around the energy of the underlying parton.

The low-response tail of the distribution is due to severe mismeasurements of the jet p_{T}. This can be due to physics sources such as semi-leptonic decays of heavy-flavour quarks, where energy is transferred to neutrinos that escape undetected. This effect is largest for jets which originate from heavy-flavour quarks in the first place. In Fig. 4.14, the fractional contributions of the different jet flavours to the response are depicted. The fractional contributions of b- and c-quark jets are highest in the low-response tails, where, at higher p_{T}, they even constitute the dominant fraction.

Furthermore, also the detector design contributes to the tails, because shower leakage will occur if not all the energy is deposited within instrumented regions. Moreover, since the calorimeters have a depth of only about 6 nuclear interaction lengths at small $|\eta|$, some of the secondary particles in hadronic showers can even traverse the apparatus (*punch-through*). For example, in QCD-multijet events with a transverse-momentum scale of larger than 500 GeV and $\not{E}_{\mathrm{T}} > 200\,\mathrm{GeV}$, about $10 - 20\%$ of the jets in the barrel are

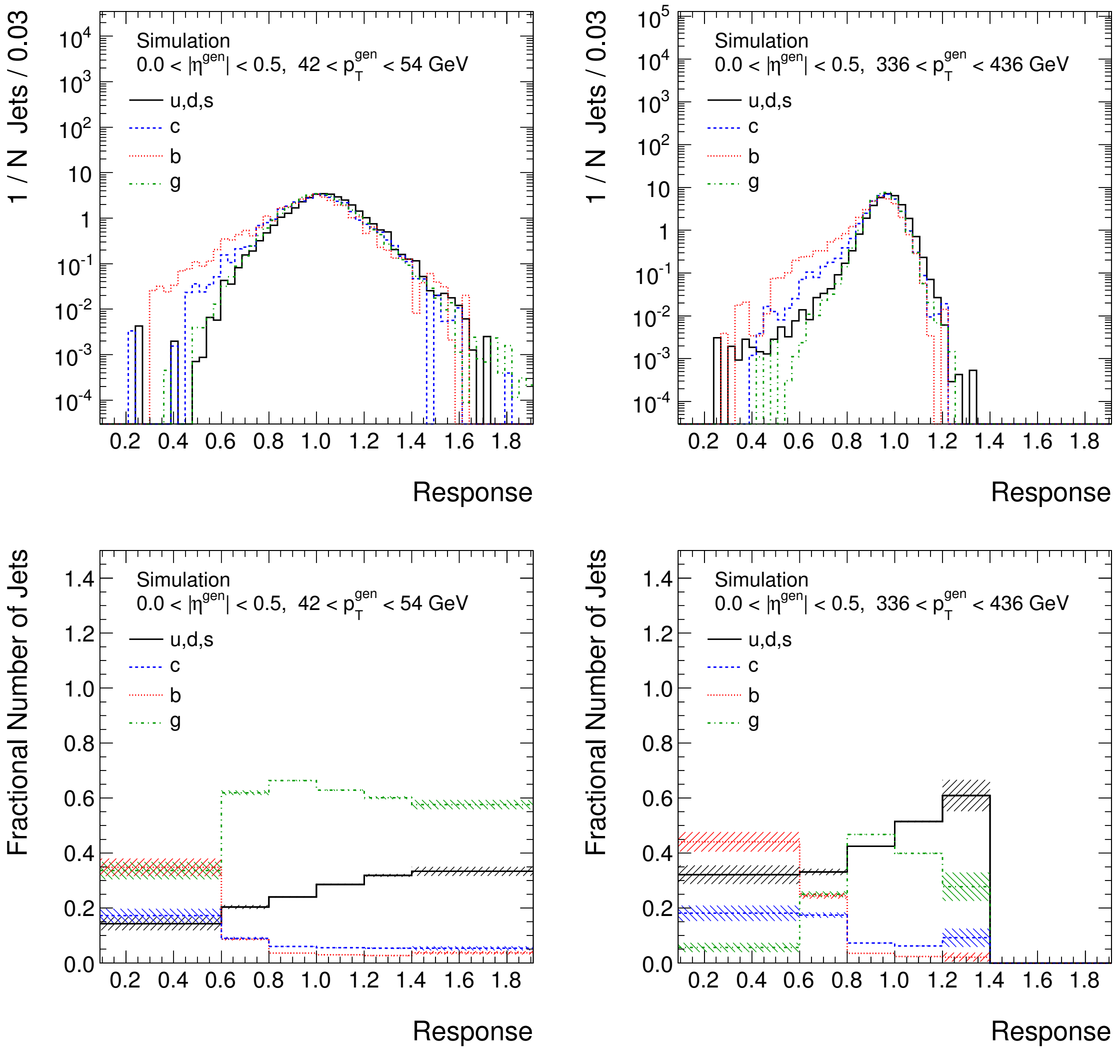


Figure 4.14: MC-truth response for uncalibrated jets of different flavours, where the distributions are normalised to their respective integral (*top*), and the fractional contributions from each jet flavour to the total distribution (*bottom*). Distributions are shown in a low (*left*) and a medium (*right*) $p_{\mathrm{T}}^{\mathrm{gen}}$ interval. The flavour has been determined using the 'algorithmic definition' described in [220]. Hatched bands represent the statistical uncertainties obtained approximating binomial errors.

affected by punch-through [222]. Finally, malfunctioning detector components can also contribute by mismeasurements of the deposited energy or generation of artificial signals, which can lead to both a low- or a high-response tail.

4.4.3 Gaussian Approximation

In order to characterise the response distribution in terms of the mean value and width of the dominating central region, it is approximated by a Gaussian. Its mean value μ and standard deviation $\sigma_{\mathcal{R}}$ define the jet energy scale and relative jet-p_{T} resolution, respectively, in this thesis. This corresponds to approximating the measured $\langle p_{\mathrm{T}} \rangle$ for a given $\langle p_{\mathrm{T}}^{\mathrm{gen}} \rangle$ through the relation

$$\langle \mathcal{R} \rangle = \left\langle \frac{p_{\mathrm{T}}}{p_{\mathrm{T}}^{\mathrm{gen}}} \right\rangle \approx \frac{\langle p_{\mathrm{T}} \rangle}{\langle p_{\mathrm{T}}^{\mathrm{gen}} \rangle} \,, \tag{4.4}$$

and likewise for the resolution.

To avoid biasing the parameter values by the presence of the tails, the Gaussian is fitted in the central interval $\mu \pm 2\,\sigma_{\mathcal{R}}$ which is determined in an iterative procedure. A first approximation μ_1, $\sigma_{\mathcal{R},1}$ of the parameter values is obtained by a Gaussian fit in the interval $\mu_0 \pm 2.5\,\sigma_{\mathcal{R},0}$, where μ_0 and $\sigma_{\mathcal{R},0}$ are the sample mean and standard deviation of the response distribution, respectively. The final values are then fitted in the interval $\mu_1 \pm 2\,\sigma_{\mathcal{R},1}$.

Results of the fitting procedure are shown in Fig. 4.15 for the two example intervals. A χ^2 goodness-of-fit test [223, 224], which results in values of $2-5$ for χ^2 relative to the number of degrees of freedom, clearly indicates some tension between the shape of the distribution and the chosen Gaussian parametrisation, cf. also Fig. 4.16.

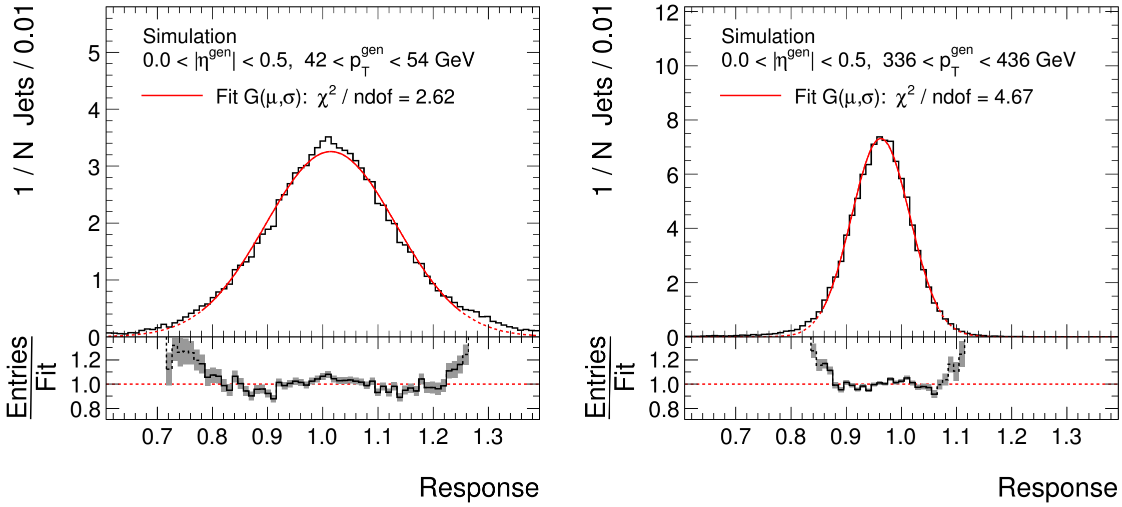


Figure 4.15: MC-truth response (*histogram*) for uncalibrated jets in a low (*left*) and a medium (*right*) $p_{\mathrm{T}}^{\mathrm{gen}}$ interval. The distribution is fitted with a Gaussian G (*line*) with mean μ and standard deviation $\sigma_{\mathcal{R}}$ in the central interval $\mu \pm \sigma_{\mathcal{R}}$ (*solid part of the line*) determined by the iterative procedure described in the text. The bottom part of the pad displays the ratio of the histogram and G together with the statistical uncertainty (*shaded band*).

This observation has a simple physics interpretation. Depending on the type of particle that is reconstructed, measurements from different subdetectors with a different (energy dependent) resolution are combined by the PF algorithm. A jet contains particles of different type and energy, and the composition varies from jet to jet. Hence, the (central part of the) response distribution has to be understood as the sum of several[4] Gaussians

---

[4] So few that the Central Limit Theorem does not apply at the level of required accuracy.

with different width. In the forward region of the detector, the Gaussian model yields a better description of the response, cf. Fig. 4.16. Here, particle energies are measured only by the calorimeter because there is no coverage by the tracking detector.

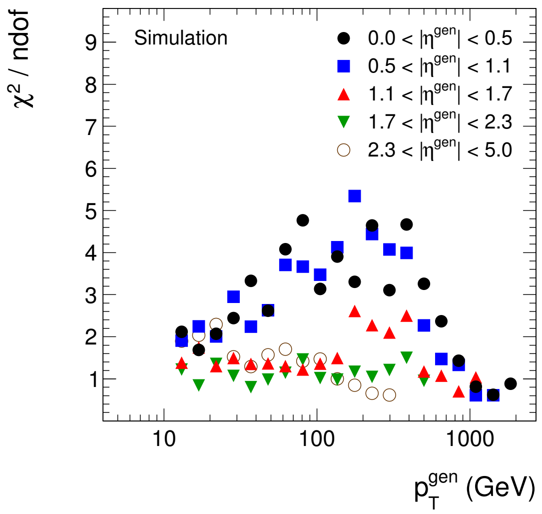


Figure 4.16: Goodness-of-fit obtained from a χ^2 test for a fit of the MC-truth response distributions in different $|\eta^{\mathrm{gen}}| \times p_{\mathrm{T}}^{\mathrm{gen}}$ intervals with Gaussian functions. The fit range is restricted to the central interval $\mu \pm \sigma_{\mathcal{R}}$ determined by the iterative procedure described in the text, where μ and $\sigma_{\mathcal{R}}$ are the mean and width of the Gaussian.

Nevertheless, the Gaussian model is used in the following to parametrise the jet response distribution in order to obtain an unambiguous definition of the average response and resolution that can also be used in the maximum-likelihood technique discussed in Section 5. This has also been done in previous analyses performed by CMS [169]. As will be shown later in Section 4.6.1, the resolution is reasonably stable with respect to the exact choice of the fitting interval.

4.4.4 Discussion of the Jet Energy Scale

The jet energy scale, i. e. the Gaussian mean of $\mathcal{R}$, is shown in Fig. 4.17 as a function of $p_{\mathrm{T}}^{\mathrm{gen}}$ and η^{gen} in two selected intervals. It features little energy dependence since effects due to the non-linear calorimeter response are reduced when clustering jets from individual PF particles. Furthermore, the scale is close to unity even before application of the full jet energy calibration described in Section 4.5 and also in the forward region because PF objects are precalibrated, cf. Section 3.5.2. The impact from pile-up is again clearly visible.

At $|\eta^{\mathrm{gen}}| \approx 1.3$ and $|\eta^{\mathrm{gen}}| \approx 3$, the jet energy scale before calibration features slight dips, which are attributed to the transition between the barrel and the endcap and the transition between the endcap and forward subdetectors, respectively. In these regions, the presence of non-instrumented regions that contain support structures, cooling pipes, and readout cables as well as potential miscalibration between the different subdetector parts are expected to affect the jet measurements.

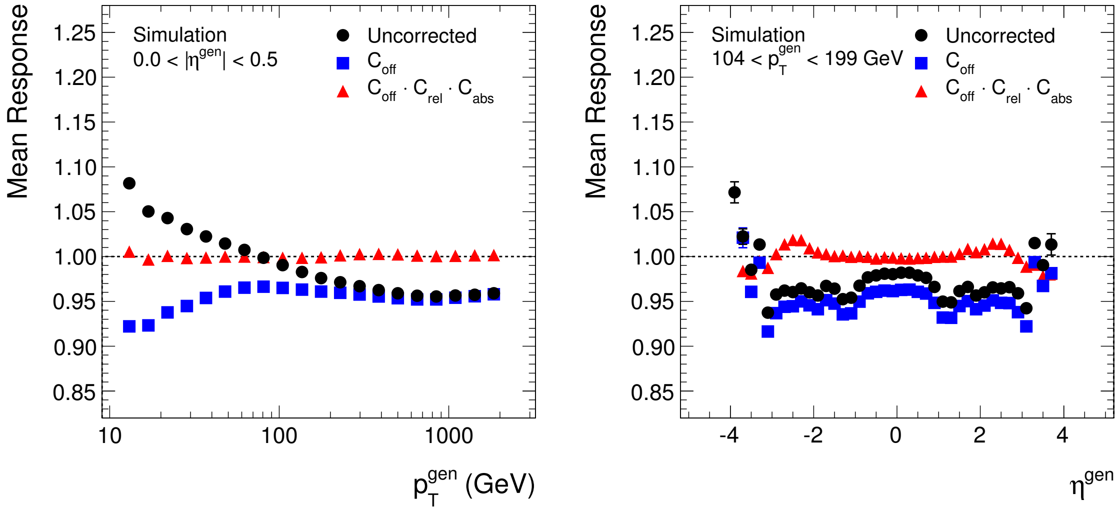


Figure 4.17: Jet energy scale of simulated events as a function of $p_{\mathrm{T}}^{\mathrm{gen}}$ for $|\eta^{\mathrm{gen}}| < 0.5$ (*left*) and as a function of η^{gen} for $104 < p_{\mathrm{T}}^{\mathrm{gen}} < 199\,\mathrm{GeV}$ (*right*) for uncalibrated jets (*circles*), after application of the offset jet energy correction (*squares*), and after application of the full jet energy calibration (*triangles*). The jet energy calibration is described in Section 4.5.

4.4.5 Measurement of the Jet Response in Data

CMS employs several techniques to measure the jet energy scale from data. They are described in [172, 225, 226] and briefly reviewed in the following.

An obvious approach is to determine the response for events where the jet momentum is balanced in the transverse plane against the momentum of a well measured reference object. Two such methods are being used, one using photon + jet events and the other using Z + jet events where the Z boson decays into electrons or muons. Both are based on the superior energy resolution of the reference objects compared to jets due to the excellent performance of the electromagnetic calorimeter as well as the muon and tracking detectors. For example, the relative energy resolution for 100 GeV photons is better than 1% while for 100 GeV jets it is about 10%. Hence, instead of the particle-level jet, the reference object can be used to compute the response Eq. (4.2).

Three principle effects have to be considered, however. Firstly, additional jets due to hard QCD radiation introduce a momentum imbalance between the jet and the reference object. An extrapolation method, similar to the one applied later in Chapter 5, is used to correct for this effect. Secondly, the jet energy scale at CMS is defined relative to the flavour composition of QCD-multijet events, which are dominated by gluons for $p_{\mathrm{T}} \lesssim 400\,\mathrm{GeV}$, cf. Fig. 4.13 (*left*). Jets in photon + jet or Z + jet events, on the other hand, mostly originate from light-flavour quarks. Therefore, the measured jet energy scale will be larger, cf. Fig. 4.13 (*right*). This effect is a major contribution to the uncertainty when measuring the absolute scale from photon + jet or Z + jet events. Finally, the response Eq. (4.2) is defined relative to the particle-level jet p_{T}, which is in general smaller than the p_{T} of the original parton due to hadronisation and jet-clustering effects. In case of the photon or Z boson, on the other hand, these effects are not present. Hence, the response with respect to the parton p_T is probed leading to a smaller jet energy scale. In practice, the impact of the listed effects is much reduced, however, because the methods are employed to measure the relative difference of the jet energy scale in data and simulation as will be explained below in Section 4.5.

While the photon/Z + jet methods are conceptually straight-forward and provide a direct measurement of the jet response, they lack statistical precision due to the relatively low cross sections of the processes[5]. Therefore, another approach, the *p_T-balance method*, which uses QCD-dijet events, is also pursued. The technique is based on the conservation of the momenta of the two jets in the transverse plane and allows for a measurement of their relative response. It is used to relate the transverse-momentum response of jets in arbitrary detector regions to that of jets in a reference region at $|\eta| < 1.3$, where the response is most uniform and the reach in p_T is largest. Again, only events with little extra jet activity are selected to ensure momentum balance at particle level.

The p_T-balance method is affected by an inherent *resolution bias* due to the fact that the resolution of the probed jet and the reference object, the barrel jet, are of similar size (while for example in case of the photon + jet balancing method, the reference object has a much better resolution compared to the probed jet). Hence, the migration effects that occur if events are collected in intervals of measured transverse momentum, cf. Section 4.4.1, are not negligible also for the reference object. If both jets lie in different $|\eta|$ regions, they will have slightly different resolutions and consequently the migration effects will bias the selection differently, resulting in a reduced relative response at large $|\eta|$.

4.5 Jet Energy Calibration

The purpose of the jet energy calibration is to relate the measured energy of a detector-level jet on average to the energy of the particle-level jet. The calibration constants used in this thesis have been centrally provided by the CMS collaboration. In the following, the procedure for their derivation is briefly reviewed; it is described in detail in [172,228,229][6].

The jet-energy-scale calibration at CMS is achieved by application of a correction factor C to the measured[7] four-momentum p_μ^uncorr of the detector-level jet, such that the corrected momentum becomes

$$p_\mu^\mathrm{corr} = C \cdot p_\mu^\mathrm{uncorr} \ .$$

The correction factor C itself factorises into several components which correct for different effects of the energy measurement,

$$C = C_\mathrm{off}(p_\mathrm{T}^\mathrm{uncorr}, \eta) \cdot C_\mathrm{rel}(p'_\mathrm{T}, \eta) \cdot C_\mathrm{abs}(p''_\mathrm{T}) \cdot C_\mathrm{res}(\eta) \ . \tag{4.5}$$

The correction factors are applied sequentially in the stated order; each prime on p_T denotes the transverse momentum after the previous correction steps. Firstly, the *offset correction* C_off compensates for instrumental noise as well as energy contributions from the

---

[5] The cross section for isolated-photon production at the LHC is about $10^3 - 10^4$ times smaller than for QCD-multijet production, cf. [227] and Fig. 4.4.

[6] The cited publications describe the results obtained by CMS with $36\,\mathrm{pb}^{-1}$ of data in 2010. Updated calibration constants have been derived by CMS from a larger data set of $500\,\mathrm{pb}^{-1}$ collected in 2011. They have been provided to the collaboration, and they are in fact used in this thesis, but since they are very similar to the 2010 numbers, no dedicated publication exists.

[7] The measured momentum referred to is again the precalibrated signal mentioned in Section 4.4.2.

UE and pile-up events. Secondly, the *relative correction* C_{rel} ensures a uniform jet energy scale in η, and thirdly, the *absolute correction* C_{abs} shifts the jet energy scale to 1. Finally, the *residual correction* C_{res} removes small differences between data and simulation. Hence, it is applied to data only.

The calibration is determined with respect to the jet-flavour composition of QCD-multijet events. Dedicated further correction steps, which are obtained from simulation, exist to optionally correct the energy scale of jets of a certain flavour for residual differences [229]. This is interesting, for example, in analyses where t-quark decays are reconstructed and the b-quark jet is identified.

For the offset correction, the additional energy is determined on a per-jet basis from the median p_{T} area-density ρ and the jet area A_j. The latter is obtained by adding a large number of infinitely soft particles to the event and clustering them together with the true measurements into jets. Then, A_j is defined as the space occupied by the soft particles. ρ is measured in QCD-multijet events for all reconstructed jets and is insensitive to the presence of hard jets. It is parametrised depending on η. The performance of the offset correction is demonstrated by comparison of the MC-truth response after application of C_{off} to the MC-truth response obtained from events with a low number of pile-up collisions, cf. Fig. 4.12. Evidently, the additional energy is sufficiently removed from the jets.

Both the relative and the absolute corrections are determined entirely from simulation. The MC-truth response $\mathcal{R}$ is obtained from the same PYTHIA sample as described above in intervals of $|\eta^{\text{gen}}|$ and $p_{\text{T}}^{\text{gen}}$, and the correction per interval is defined as the inverse of the average response, $1/\langle\mathcal{R}\rangle$. It is related to the average detector-level jet p_{T} in the same interval, and thus, the correction can be expressed as a function of measured p_{T}. Following this procedure, C_{abs} is determined from events with both leading jets in a reference region at $|\eta| < 1.3$, where $\langle\mathcal{R}\rangle$ has little dependence on η, cf. Fig. 4.17 (*right*). C_{rel} is computed in a slightly modified way relative to $\langle\mathcal{R}\rangle$ in the reference region as a function of p_{T} and in intervals of η from events where one jet falls into the reference region and one jet in the η interval. Since there are more events of this topology than with both jets in the same interval in the forward region, the separation between relative and absolute correction is beneficial because a higher statistical precision is achieved. In total, the combined relative and absolute correction amounts to about 1 – 5% at $p_{\text{T}} > 50\,\text{GeV}$ and $|\eta| < 0.5$ and increases to about 10% at $p_{\text{T}} = 20\,\text{GeV}$ for $|\eta|$ between 2 and 3. The effect of the calibration up to C_{abs} on the jet response and energy scale is demonstrated in Figs. 4.18 and 4.17, respectively. After correction, the scale is uniform in η and at 1 to a precision of $\lesssim 1\%$.

The residual correction is determined similarly in two steps. First, the p_{T}-balance method is employed to measure the average response relative to the $|\eta| < 1.3$ reference region in data and in simulation after the C_{abs} correction-step. The relative jet energy scale in $36\,\text{pb}^{-1}$ of data collected by CMS in 2010 and in simulation is shown in Fig. 4.19 (*left*). In data, the average response increases with respect to the reference region by up to 10% at $|\eta| \approx 2.5$ and decreases again for larger $|\eta|$ due to the discussed resolution bias. In the simulation, on the other hand, it is uniform in pseudorapidity — as it should be after the C_{rel} correction — up to a few percent and decreases only in the very forward regions due to the mentioned bias. The observed differences are attributed to an inaccurate modelling of the jet detection in the simulation. Hence, the data-to-simulation ratio of the relative jet energy scale, Fig. 4.19 (*right*), is utilised to derive an η-dependent residual

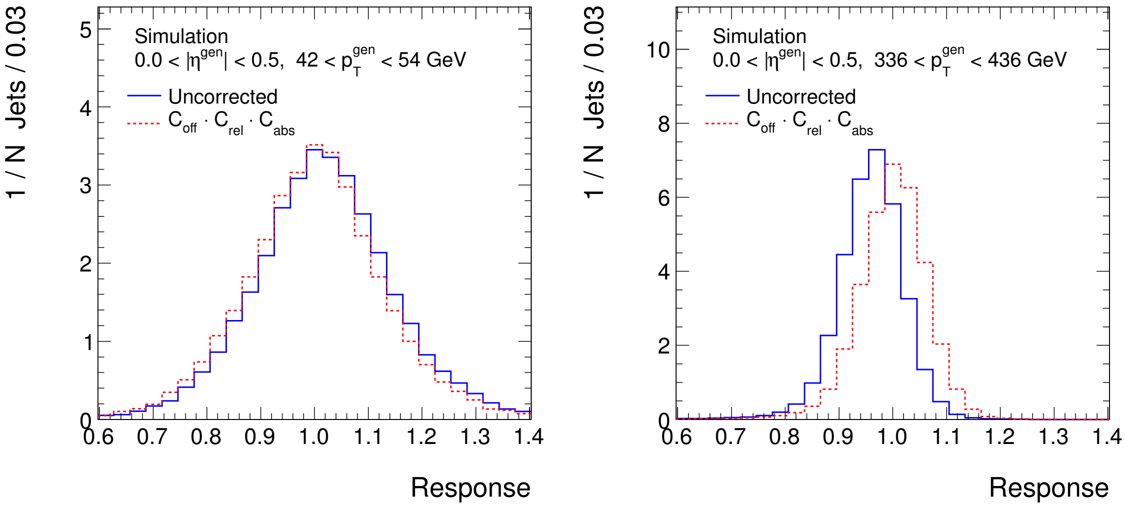


Figure 4.18: MC-truth response in a low (*left*) and a medium (*right*) $p_{\mathrm{T}}^{\mathrm{gen}}$ interval for uncalibrated jets (*solid histogram*) and after application of the offset, relative, and absolute jet energy corrections (*dashed histogram*).

correction, where it is assumed that the resolution bias is the same in both cases[8]. The correction amounts to about $2-3\%$ in general and is as large as 10% for $|\eta| \approx 2.5$ with an uncertainty of up to 5% at large $|\eta|$, dominated by the uncertainty on the jet resolution. After application, the jet energy scale in data is uniform in η within 2% as shown ibid. Finally, the absolute jet energy scale in the reference region is measured from photon + jet events in data and in simulation. The inverse of its ratio serves as a correction, which amounts to about 1.5% constant in p_{T}. Since a data-to-simulation ratio is employed, the correction factor is not affected by the flavour difference between photon + jet and QCD samples. The uncertainty of the result is dominated by a 0.5% uncertainty on the flavour modelling.

The total jet energy correction C Eq. (4.5) adds up to $1-5\%$ with an uncertainty of $1-2\%$ for $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ and $|\eta| < 0.5$. It reaches $10-20\%$ with an uncertainty of $5-7\%$ for $p_{\mathrm{T}} = 20\,\mathrm{GeV}$ in all detector regions. The simulation-based relative and absolute corrections contribute the largest fraction to C, and its uncertainty is dominated by the uncertainty on the relative residual correction.

4.6 Jet Transverse-Momentum Resolution

The relative jet transverse-momentum resolution has been defined previously as the standard deviation of a Gaussian fitted to the response distribution. In the following, in Section 4.6.1, the determination of the resolution from simulation is described and the stability of the method is investigated. Furthermore, its dependence on the number of pile-up collisions is studied. The derived resolution will serve as a reference in validation tests later on in this thesis. Afterwards, in Section 4.6.2, the *dijet asymmetry* is intro-

---

[8] After resolution differences between data and simulation, cf. Section 5.6, have been considered.

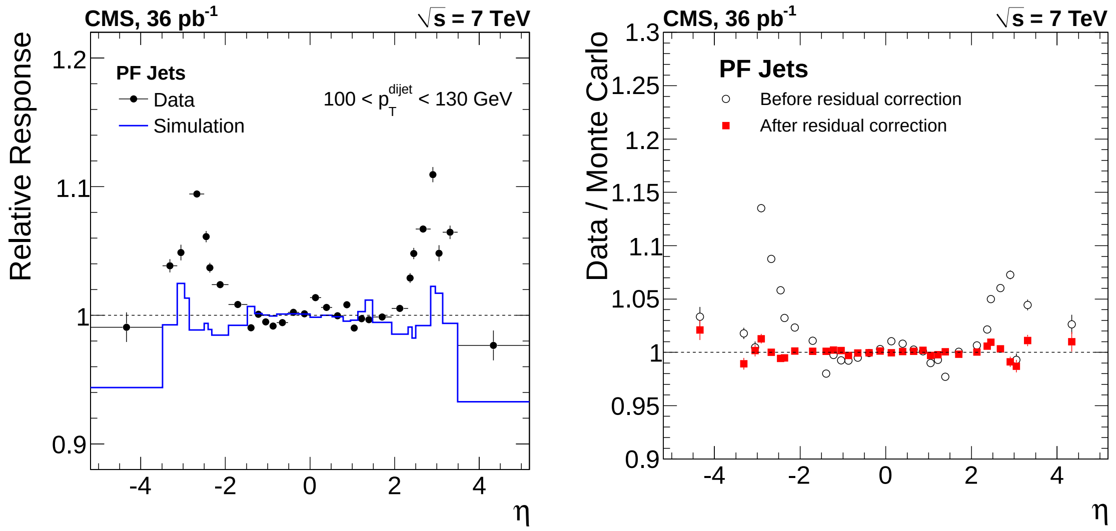


Figure 4.19: Relative jet energy scale after application of the C_{abs} correction step (*left*) measured with the p_{T}-balance method as a function of η in data (*circles*) and simulation (*histogram*), and its ratio (*right*) without (*open circles*) and with (*solid squares*) application of the residual correction. Taken from [172].

duced, which is an important quantity because it can be measured easily in data and it is directly related to the resolution.

4.6.1 MC-Truth Resolution

The jet-p_{T} resolution is determined from the Pythia sample as a function of $p_{\mathrm{T}}^{\mathrm{gen}}$ in intervals of $|\eta^{\mathrm{gen}}|$ after application of the jet energy calibration Section 4.5. It will be referred to as *MC-truth resolution* $\sigma_{\mathrm{MC}}(p_{\mathrm{T}}^{\mathrm{gen}}, \eta^{\mathrm{gen}})$.

The MC-truth response distributions are determined, and the core of each distribution is fitted with a Gaussian using the iterative approach as described above in Sections 4.4.1 and 4.4.3. Following Eq. (4.4), the average relative MC-truth resolution is defined as the standard deviation $\sigma_{\mathcal{R}}$ of the fitted Gaussian[9],

$$\frac{\sigma_{\mathrm{MC}}}{p_{\mathrm{T}}^{\mathrm{gen}}} \equiv \left\langle \frac{\sigma_{\mathrm{MC}}(p_{\mathrm{T}}^{\mathrm{gen}}, \eta^{\mathrm{gen}})}{p_{\mathrm{T}}^{\mathrm{gen}}} \right\rangle = \sigma_{\mathcal{R}} \,. \tag{4.6}$$

Its evolution with $p_{\mathrm{T}}^{\mathrm{gen}}$ is shown in Fig. 4.20 for different $|\eta^{\mathrm{gen}}|$ intervals. The relative resolution in the central detector region improves from 12% at $p_{\mathrm{T}}^{\mathrm{gen}} = 50\,\mathrm{GeV}$ through 9% at 100 GeV and 5% at 600 GeV to 4.5% at 1 TeV. This is similar to the jet-p_{T} resolution achieved by ATLAS [230, 231].

For $p_{\mathrm{T}}^{\mathrm{gen}} > 100\,\mathrm{GeV}$, the resolution in the forward intervals with $|\eta| > 1.7$ is better than in the central intervals. This might be due to kinematic reasons, because jets of the same p_{T} have a higher energy in the forward than in the central region, and the relative resolution of calorimetric measurements, which dominate the jet reconstruction at high

---

[9] Note that σ_{MC} has units of momentum while $\sigma_{\mathcal{R}}$ is dimensionless.

p_{T}, improves as a function of energy rather than p_{T}. The worst resolution occurs in the medium interval $1.1 < |\eta| < 1.7$, which covers the region of transition from the barrel to the endcap parts of the different subdetector systems, cf. Fig. 3.7. As in case of the jet energy scale Fig. 4.17, the presence of non-instrumented regions as well as potential miscalibration between the different detector parts can affect the resolution.

The different measurements in each $|\eta^{\mathrm{gen}}|$ interval are fitted with the function

$$\frac{\sigma_{\mathrm{MC}}(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \sqrt{\mathrm{sgn}(N) \cdot \left(\frac{N}{p_{\mathrm{T}}}\right)^2 + S^2 \cdot p_{\mathrm{T}}^{m-1} + C^2}\ , \tag{4.7}$$

where N, S, C, and m are free parameters. The fitted functions are overlaid in Fig. 4.20 and compared for different $|\eta^{\mathrm{gen}}|$ intervals in Fig. 4.21. The decrease of the fitted curve towards low $p_{\mathrm{T}}^{\mathrm{gen}}$ in the $1.1 < |\eta| < 1.7$ interval is unphysical and caused by the downward fluctuation of the point at 18 GeV, which has a greater statistical precision than the first point and thus drives the fit result. In order to avoid biases for example due to energy thresholds of the calorimeter readout or inefficiencies of the particle reconstruction at low p_{T}, the fit range has been chosen to start at 10 GeV. The fitted parameter values are listed in Table 4.1.

Table 4.1: Fitted parameter values of the MC-truth resolution Eq. (4.7). The quoted uncertainties are the statistical uncertainties returned by the fit.

$\lvert\eta\rvert$	N (GeV)	S (GeV$^{(1-m)/2}$)	C	m
0.0 – 0.5	-1.13 ± 0.14	0.611 ± 0.020	0.03121 ± 0.00074	0.143 ± 0.014
0.5 – 1.1	-0.44 ± 0.40	0.625 ± 0.023	0.03401 ± 0.00088	0.136 ± 0.017
1.1 – 1.7	-3.20 ± 0.18	1.529 ± 0.089	0.04100 ± 0.00097	-0.196 ± 0.024
1.7 – 2.3	-3.37 ± 0.44	2.08 ± 0.30	0.0321 ± 0.0017	-0.416 ± 0.056
2.3 – 5.0	-1.45 ± 0.22	1.034 ± 0.053	0.0000 ± 0.0068	-0.085 ± 0.021

The functional form of Eq. (4.7) was introduced in [172] to describe the resolution of jets reconstructed from tracking and calorimeter measurements[10]. It is a modification of the parametrisation Eq. (3.3) typically employed for calorimeters, for which $m = 0$ and $N > 0$. In case of the PF-based jet reconstruction, the superior momentum resolution of the tracking system at low p_{T} strongly improves the jet-p_{T} resolution, which is considered via the sgn(N) factor. For example, the MC-truth resolution of PF jets at 50 GeV is 30% better than for Calo jets [172]. At medium p_{T}, tracking information still compensate for non-linearities of the calorimeters, which is considered by the parameter m. At high p_{T}, the calorimeters have the superior resolution and dominate the measurement, which is consequently limited by the constant term C due to calorimeter miscalibration and non-uniformities.

The dependence of the MC-truth resolution on the matching criterion for generator-level and detector-level jets has been studied, cf. Fig. 4.22 (*left*). As expected, a larger $\Delta\mathrm{R}_{\mathrm{max}}$

---

[10] In [172], the parameter C is set to 0. Here, it has been also fitted which yields better compatibility with the MC-truth resolution beyond $p_{\mathrm{T}}^{\mathrm{gen}} = 1$ TeV.

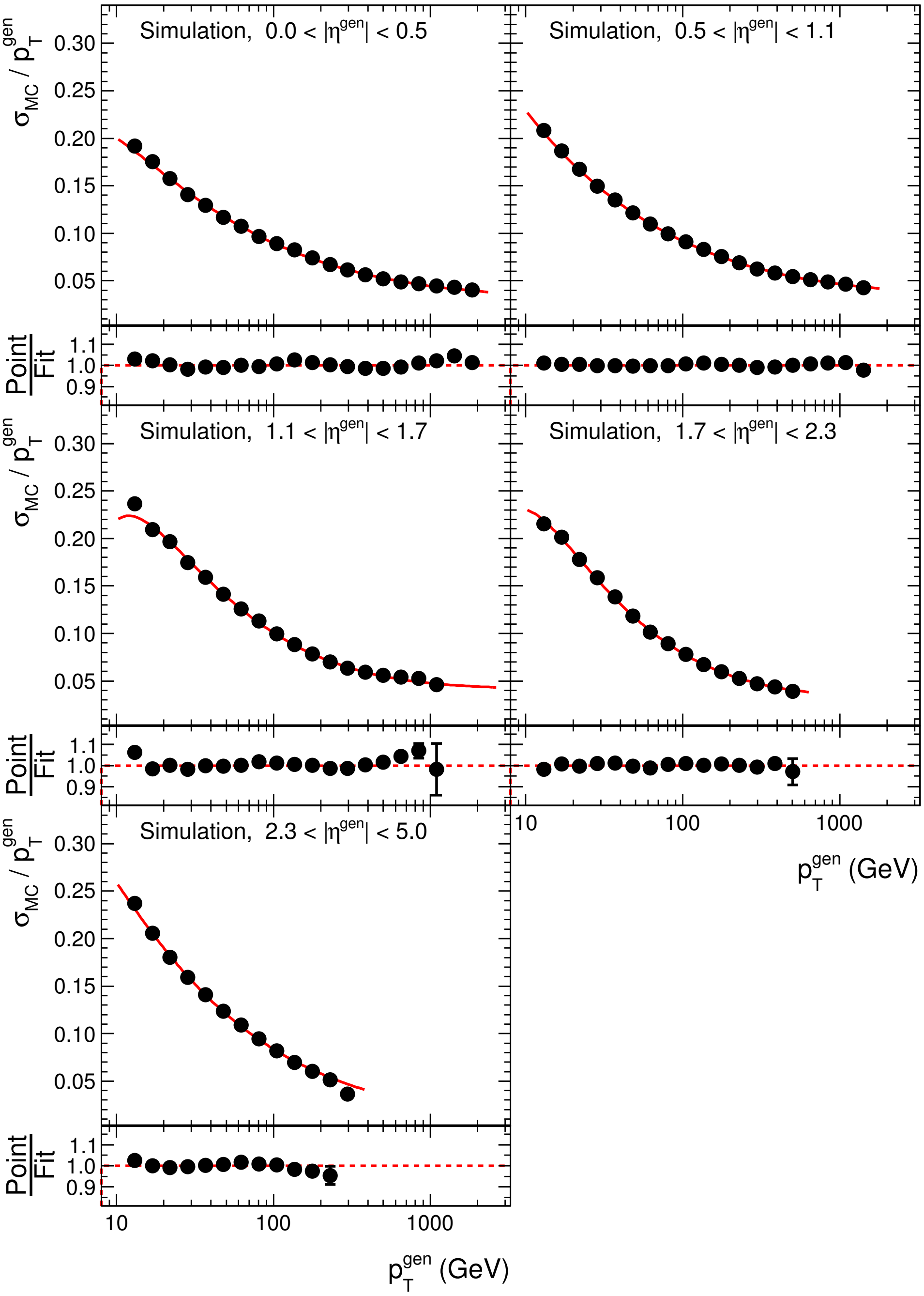


Figure 4.20: Relative MC-truth resolution $\sigma_{\mathrm{MC}}/p_{\mathrm{T}}^{\mathrm{gen}}$ in different $|\eta^{\mathrm{gen}}| \times p_{\mathrm{T}}^{\mathrm{gen}}$ intervals (*circles*). They are fitted with the $p_{\mathrm{T}}^{\mathrm{gen}}$ dependent function Eq. (4.7) in each $|\eta^{\mathrm{gen}}|$ interval (*solid line*). The bottom part of each pad displays the ratio of $\sigma_{\mathrm{MC}}/p_{\mathrm{T}}^{\mathrm{gen}}$ and the fit. Error bars represent the statistical uncertainties.

results in a larger resolution because the impact from jet splitting increases. However, the effect is visible only for $p_{\mathrm{T}}^{\mathrm{gen}} < 60\,\mathrm{GeV}$, and its size is moderate. At 30 GeV, for instance, the resolution degrades by 20% when changing $\Delta\mathrm{R}_{\mathrm{max}}$ from 0.05 to 0.25. As demonstrated

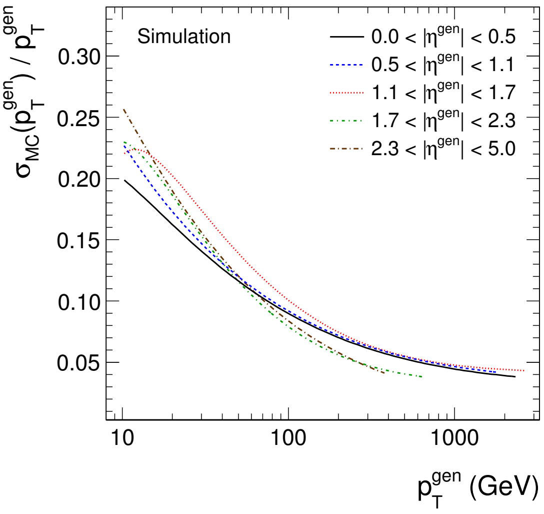


Figure 4.21: Fitted MC-truth resolution functions $\sigma_{\mathrm{MC}}(p_{\mathrm{T}}^{\mathrm{gen}})/p_{\mathrm{T}}^{\mathrm{gen}}$ Eq. (4.7) in different $|\eta^{\mathrm{gen}}|$ intervals.

in Fig. 4.22 (*right*), with the choice of $\Delta\mathrm{R}_{\mathrm{max}} = 0.1$, which is used throughout this thesis[11], a detector-level jet can be matched to the generator-level jets in 99% of the cases.

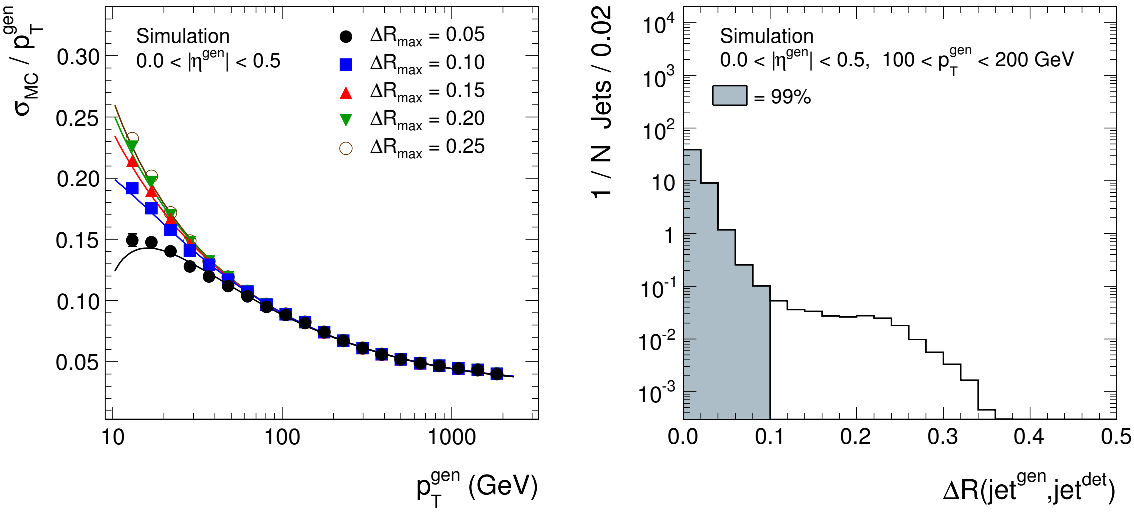


Figure 4.22: Relative MC-truth resolution for different choices of $\Delta\mathrm{R}_{\mathrm{max}}$ as a function of $p_{\mathrm{T}}^{\mathrm{gen}}$ for $|\eta| < 0.5$ (*left*). The $\Delta\mathrm{R}$ resolution is about 0.02 at $p_{\mathrm{T}} = 20\,\mathrm{GeV}$ and improves quickly for larger p_{T} [217]. The distance in $\Delta\mathrm{R}$ between a generator-level jet and the closest detector-level jet is smaller than 0.1 for more than 99% of the jet pairs (*right*).

In Fig. 4.23 (*left*), the dependence of the MC-truth resolution on the chosen response interval for the Gaussian fit is depicted. Choice of a larger interval results in a larger

---

[11] A value of $\Delta\mathrm{R}_{\mathrm{max}} = 0.1$ has also been used by CMS in previous analyses [169].

resolution because the fraction of the aforementioned components with poor resolution increases. The overall effect is small, however. At $p_{\mathrm{T}}^{\mathrm{gen}} = 30\,\mathrm{GeV}$, the resolution degrades by 7% when increasing the fitting interval from 1.5 to $2.5\,\sigma$. Above 100 GeV, no difference is visible.

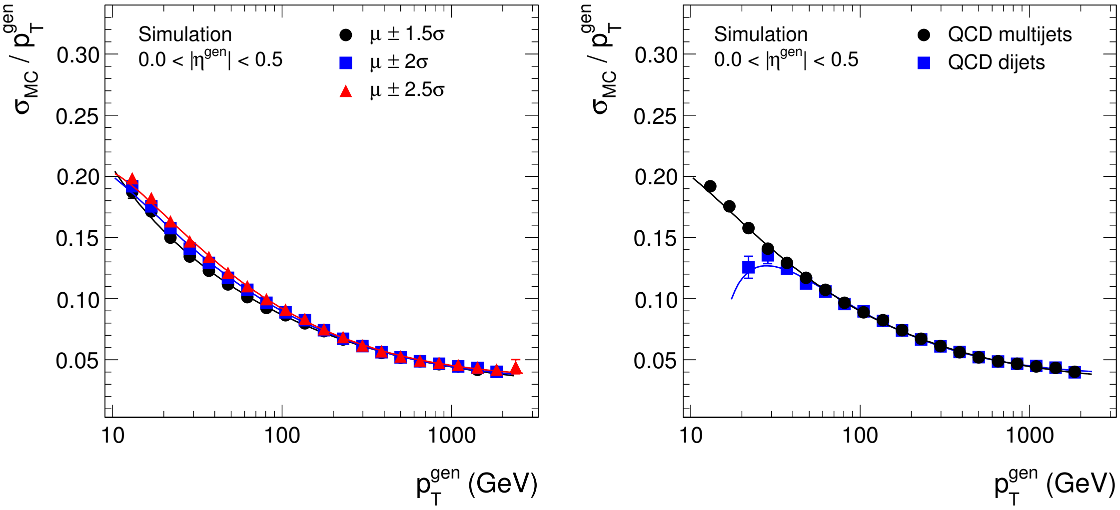


Figure 4.23: Relative MC-truth resolution as a function of $p_{\mathrm{T}}^{\mathrm{gen}}$ for $|\eta| < 0.5$ for different choices of the response interval in which the Gaussian fit is performed (*left*) and compared to the resolution obtained on a dijet sample (*right*).

The MC-truth resolution has been determined from an inclusive QCD-multijet sample, and has thus been averaged over the flavour composition expected in that case. Later, in Section 5.4.1, the MC-truth resolution will serve as reference when validating a method to measure the jet-p_{T} resolution from dijet events. Therefore, the sensitivity of the MC-truth resolution to a dijet selection is investigated. Events are selected by requiring

$$\Delta\phi^{\mathrm{gen}} = \Delta\phi(\vec{p}_{\mathrm{T},1}^{\,\mathrm{gen}}, \vec{p}_{\mathrm{T},2}^{\,\mathrm{gen}}) > 2.7 \tag{4.8}$$

and

$$p_{\mathrm{T},3}^{\mathrm{gen}} < 0.14 \cdot \frac{1}{2}(p_{\mathrm{T},1}^{\mathrm{gen}} + p_{\mathrm{T},2}^{\mathrm{gen}}) \ ,$$

i. e. the leading two generator-level jets in the event have to point into opposite directions in the transverse plane, and the p_{T} of additional generator-level jets has to be less than 14% of the average p_{T} of the leading two jets. These criteria will be motivated in detail in Section 5.1. The MC-truth resolution of the selected events is compatible to the multijet case as apparent from Fig. 4.23 (*right*). The strong decrease of the fitted curve towards low $p_{\mathrm{T}}^{\mathrm{gen}}$ is again an artefact of the downward fluctuation of the point at 20 GeV, which has a greater statistical precision than the first point. In the following, the MC-truth resolution from the QCD-multijet sample will be used in order to profit from the higher statistical precision.

Finally, the presence of pile-up events can impact the resolution because additional energy is distributed in the detector. If these additional energy deposits overlap with a detector-level jet j_i they will be clustered into j_i. The extra energy might not be exactly removed by the jet-energy-scale correction because the offset correction factor C_{off} is determined from the average pile-up energy density ρ in the event. Hence, the computed response of g_i, the generator-level jet j_i is assigned to, gets biased by local fluctuations of the pile-up energy density, and if g_i is one of the leading two generator-level jets the MC-truth resolution will be affected. The relative size of the effect is expected to be reduced at larger p_{T}, however, because the occurrence of high-p_{T} pile-up particles is rare as discussed above. There are further configurations which possibly bias the response, for instance if g_i and j_i originate from different collisions. However, these cases are expected to be highly suppressed because, given the tight ΔR matching criterion and the jet's cone size, it is unlikely that the wrongly assigned jets are reconstructed individually. As observed in Fig. 4.24, at $p_{\mathrm{T}}^{\mathrm{gen}} = 30\,\mathrm{GeV}$ the MC-truth resolution degrades by 25% when increasing the number of pile-up interactions from less than five to more than 14. The difference is less pronounced at higher p_{T}, as expected. Above 100 GeV, no significant effect is observed.

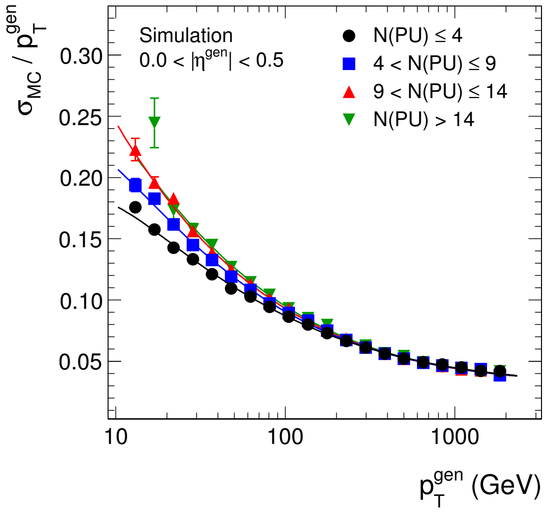


Figure 4.24: Relative MC-truth resolution as a function of $p_{\mathrm{T}}^{\mathrm{gen}}$ for $|\eta| < 0.5$ determined from events with a different number N_{PU} of simulated pile-up collisions.

4.6.2 Dijet Asymmetry

The dijet asymmetry $\mathcal{A}$ is defined for events with at least two jets as

$$\mathcal{A} = \frac{p_{\mathrm{T},1} - p_{\mathrm{T},2}}{p_{\mathrm{T},1} + p_{\mathrm{T},2}} \,, \tag{4.9}$$

where in this case $p_{\mathrm{T},1}$ and $p_{\mathrm{T},2}$ refer to the randomly ordered transverse momenta of the two leading jets. In Fig. 4.25, example asymmetry distributions are shown for QCD-dijet events which have been selected as explained later on in Section 5.1 by requiring a back-to-back topology in the transverse plane and restricting additional jet activity. The average transverse momentum of the two leading jets,

$$p_{\mathrm{T}}^{\mathrm{ave}} = \frac{1}{2}\left(p_{\mathrm{T},1} + p_{\mathrm{T},2}\right) \,, \tag{4.10}$$

has to lie in the interval $240 < p_{\mathrm{T}}^{\mathrm{ave}} < 270\,\mathrm{GeV}$, and they have to be within $|\eta| < 0.5$.

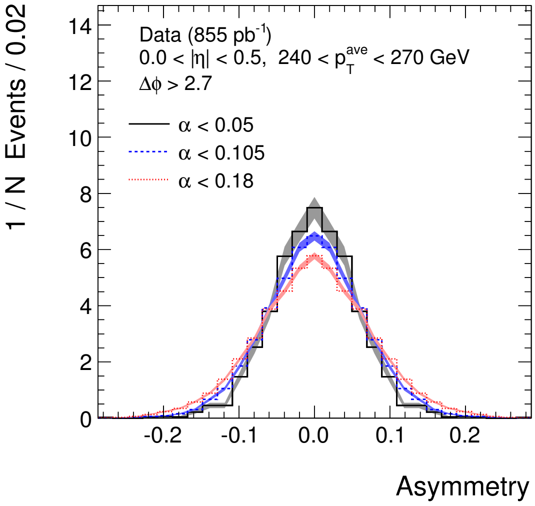


Figure 4.25: Asymmetry distributions in $855\,\mathrm{pb}^{-1}$ of QCD-dijet data. Increased transverse-momentum imbalance, characterised by the fractional transverse momentum of the third jet, $p_{\mathrm{T},3}$, relative to the average transverse momentum of the leading jets, $p_{\mathrm{T}}^{\mathrm{ave}}$, results in a broader asymmetry. Shaded bands represent the statistical uncertainties.

The standard deviation $\sigma_{\mathcal{A}}$ of the asymmetry can be expressed as

$$\sigma_{\mathcal{A}} = \left|\frac{\partial\mathcal{A}}{\partial p_{\mathrm{T},1}}\right| \cdot \sigma(p_{\mathrm{T},1}) \oplus \left|\frac{\partial\mathcal{A}}{\partial p_{\mathrm{T},2}}\right| \cdot \sigma(p_{\mathrm{T},2}) \;.$$

If at particle level the transverse momenta of both jets are balanced and if they are in the same η region, then $\langle p_{\mathrm{T},1}\rangle = \langle p_{\mathrm{T},2}\rangle = \langle p_{\mathrm{T}}\rangle$ and $\sigma(p_{\mathrm{T},1}) = \sigma(p_{\mathrm{T},2}) = \sigma(p_{\mathrm{T}})$. Hence, the jet-$p_{\mathrm{T}}$ resolution $\sigma(p_{\mathrm{T}})$ is related to $\sigma_{\mathcal{A}}$ via

$$\frac{\sigma(p_{\mathrm{T}})}{\langle p_{\mathrm{T}}\rangle} = \sqrt{2}\cdot\sigma_{\mathcal{A}} \;. \tag{4.11}$$

This important relation has previously been exploited at the DØ experiment [232, 233] to measure the jet resolution from dijet data, and the method is also applied by ATLAS [230] and CMS [172, 234].

In realistic collision events, the idealised dijet topology of two jets with exactly balancing transverse momenta at particle level is compromised because momentum is transferred from the original parton to additional jets from ISR/FSR and to soft particles of the UE activity. A transverse-momentum imbalance is induced, and in consequence the recorded asymmetry distribution is broadened, cf. Fig. 4.25. Hence, if the jet resolution is measured using Eq. (4.11) it will be biased. This effect is investigated with simulated events.

Events resembling the dijet topology are selected from the PYTHIA sample as in Eq. (4.8) above by requiring $\Delta\phi^{\mathrm{gen}} > 2.7$ and $|\eta^{\mathrm{gen}}| < 0.5$ for the leading two generator-level jets. From this sample, the asymmetry and the MC-truth response are determined in small intervals of $p_{\mathrm{T}}^{\mathrm{gen}}$. In Fig. 4.26 (*left*), the standard deviation $\sigma_{\mathcal{A}}$, multiplied by $\sqrt{2}$ to account for Eq. (4.11), is shown. Clearly, it is larger than the relative MC-truth resolution[12] $\sigma_{\mathrm{MC}}/\langle p_{\mathrm{T}}^{\mathrm{gen}}\rangle$.

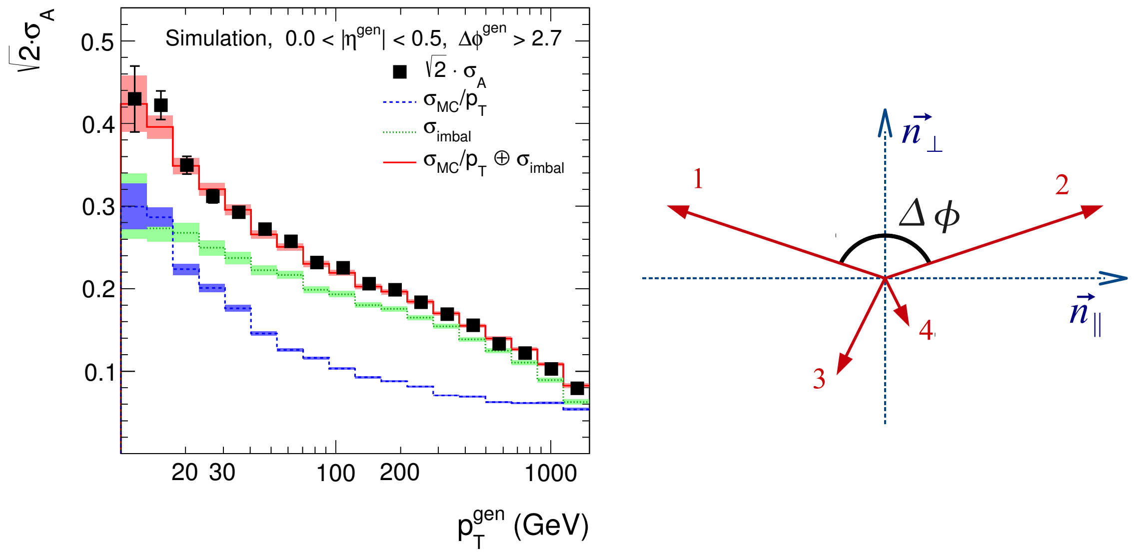


Figure 4.26: Standard deviation $\sigma_{\mathcal{A}}$, multiplied by $\sqrt{2}$, of the dijet asymmetry (*solid squares*) in simulated events as a function of $p_{\mathrm{T}}^{\mathrm{gen}}$ (*left*). The two leading jets lie within $|\eta| < 0.5$ and are separated in azimuth by $\Delta\phi^{\mathrm{gen}} > 2.7$. It is well described by the quadratic sum (*solid line*) of the relative MC-truth resolution $\sigma_{\mathrm{MC}}/\langle p_{\mathrm{T}}^{\mathrm{gen}}\rangle$ (*dashed line*) and the standard deviation of the relative momentum imbalance along the dijet axis, α^{imbal} (*dotted line*). Shaded bands represent the statistical uncertainties. The dijet axis $\vec{n}_{||}$ is defined in the transverse plane as the direction perpendicular to $\vec{n}_{\perp}$, the bisecting line of $\Delta\phi$ (*right*).

The difference is attributed to a transverse-momentum imbalance between the leading two particle-level jets induced by the radiation of hard partons which result in additional jets. Further, small contributions to the imbalance occur due to the aforementioned fluctuating difference between the momenta of the parton and the particle-level jet. They are neglected here but considered later on in Section 5.3.2 for the resolution measurement.

In fact, the imbalance is caused by the components of the additional jets' momenta along the dijet axis, as illustrated in Fig. 4.26 (*right*). The dijet axis $\vec{n}_{||}$ of an event is defined in the transverse plane as the direction perpendicular to the bisecting line of $\Delta\phi$, where $\Delta\phi$ denotes the angle between the two leading jets. Assuming $\Delta\phi = \pi$, the absolute size $p_{\mathrm{T}}^{\mathrm{imbal}}$ of the transverse-momentum imbalance due to additional jets is given by

$$p_{\mathrm{T}}^{\mathrm{imbal}} = \left| \sum_{i>2} p_{\mathrm{T},i}^{\mathrm{gen}} \cdot \cos\left(\phi_i^{\mathrm{gen}} - \phi_{||}\right) \right| ,$$

---

[12] Here, the relative resolution is not obtained from a Gaussian fit to the response distribution but rather determined as the sample standard deviation. This is sufficient for the intended demonstration purpose, even if the presence of tails affect the result, since it is consistent with the definition of $\sigma_{\mathcal{A}}$.

where ϕ_i^{gen} and $\phi_{||}$ denote the azimuthal angle of the i-th jet's momentum and of $\vec{n}_{||}$, respectively. Here, generator-level jet momenta are used in order to be independent of the detector resolution. The broadening of the asymmetry is modelled by recording α^{imbal}, the imbalance relative to the original transverse-momentum scale of the dijet event,

$$\alpha^{\mathrm{imbal}} = \frac{p_{\mathrm{T}}^{\mathrm{imbal}}}{\frac{1}{2}(p_{\mathrm{T},1}^{\mathrm{gen}} + p_{\mathrm{T},2}^{\mathrm{gen}}) + p_{\mathrm{T}}^{\mathrm{imbal}}} \, . \tag{4.12}$$

Its standard deviation σ_{imbal} is also shown in Fig. 4.26 (*left*). As demonstrated, it accounts precisely for the observed difference, and $\sqrt{2} \cdot \sigma_{\mathcal{A}}$ is regained by adding σ_{imbal} in quadrature to the MC-truth resolution,

$$\sqrt{2} \cdot \sigma_{\mathcal{A}} = \frac{\sigma_{\mathrm{MC}}}{\langle p_{\mathrm{T}}^{\mathrm{gen}} \rangle} \oplus \sigma_{\mathrm{imbal}} \, . \tag{4.13}$$

Measurements of the jet-p_{T} resolution which are based on the dijet asymmetry — or on the transverse-momentum balance in dijet events in general — have to be corrected for this systematic effect. Hence, this does also apply to the resolution and response-tail measurements performed in the next Chapters 5 and 6.

5 Measurement of the Jet Transverse-Momentum Resolution

Like the jet energy scale, the jet transverse-momentum resolution defined above is a key quantity characterising the performance of jet measurements. Hence, a precise knowledge of the resolution is required to understand jet final-states. For example, migration effects in the jet-p_{T} spectrum, which are a consequence of the p_{T} resolution, need to be considered when determining the residual jet energy corrections, cf. Section 4.5, and when measuring the QCD or $t\bar{t}$ production cross-sections [235,236]. Furthermore, many searches for physics beyond the SM are carried out in jet final-states, and due to jet mismeasurements, SM processes can fake the expected new-physics signatures such as $\not\!\!E_{\mathrm{T}}$. In the search for new physics [237] discussed in Chapter 7, the jet resolution is employed explicitly to predict the background from QCD-multijet events.

In this chapter, a measurement of the jet transverse-momentum resolution in dijet events with an integrated luminosity of $855\,\mathrm{pb}^{-1}$ collected until summer 2011 is presented. It is performed using an unbinned maximum-likelihood fit based on the assumption of p_{T} balance in QCD-dijet events and a Gaussian parametrisation of the jet p_{T} response function as motivated in Section 4.4.3. The likelihood includes an estimate for the differential jet cross-section. Therefore, biases due to the event selection can be explicitly considered, and an estimate of the average particle-level jet-p_{T} is obtained. To account for the features of realistic QCD-dijet events, the result of the maximum-likelihood fit has to be corrected for the impact from processes that induce a p_{T} imbalance at particle level. The complete measurement procedure is illustrated in Fig. 5.1.

The results of the measurement are also reported as data-to-simulation ratio of the resolution, which can be used to adjust the simulation to the data. Since the cross section for QCD-dijet production is large, the measurement has a high reach in p_{T} up to $1.5\,\mathrm{TeV}$.

The presented results have been published in [238] and are employed in numerous physics analyses performed by CMS. A previous version of the method applied to $36\,\mathrm{pb}^{-1}$ of data collected in 2010 has been published in [172, 239]. The measurement is complemented by a different, dijet-asymmetry-based approach [172, 234, 239] as well as measurements exploiting the transverse-momentum balance in photon + jet events, cf. ibid. The latter feature smaller systematic uncertainties but have a much shorter p_{T} reach[1].

In principle, also other, non-Gaussian response functions can be fitted with the maximum-likelihood method, which is in particular interesting with respect to the aforementioned response tails. Limitations arise, however, from the numerical evaluation of involved integral expressions. Moreover, the likelihood can be extended to incorporate different event types such as photon + jet events.

This chapter is organised as follows: Firstly, in Section 5.1, the selection of QCD-dijet-like events from data and simulation is described. The maximum-likelihood method, including the treatment of biases due to the event selection and the presence of non-

---

[1] As remarked in Section 4.4.5, the cross section for isolated-photon production at the LHC is about 10^3 – 10^4 times smaller than for QCD-multijet production.

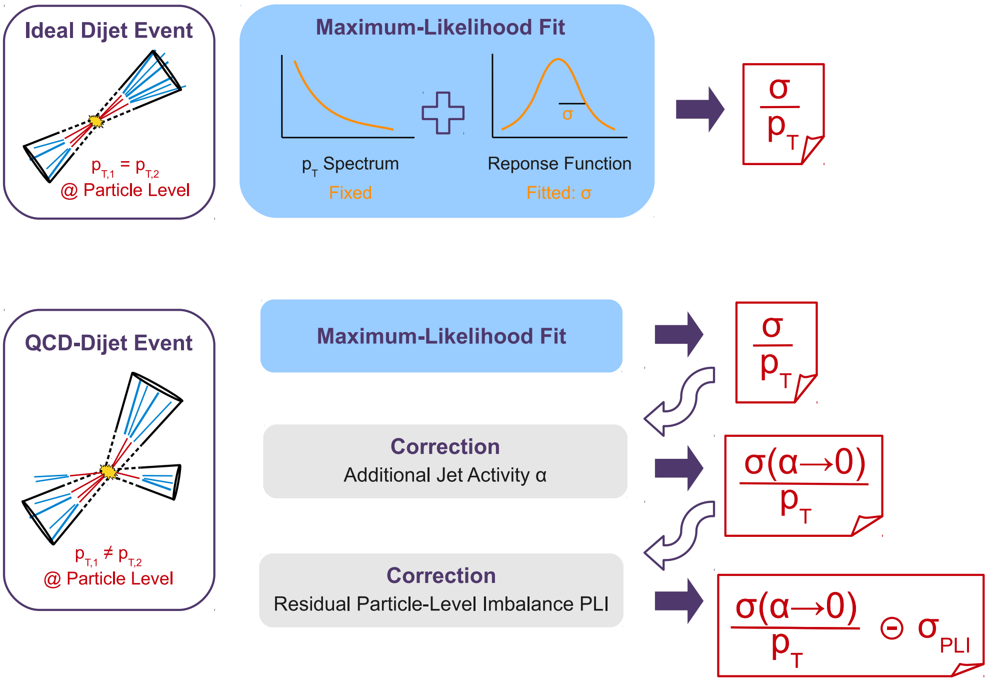


Figure 5.1: Illustration of the resolution measurement. The maximum-likelihood fit is performed assuming p_T balance at particle level (*top*). Subsequent corrections for effects that occur in realistic QCD-dijet events are applied (*bottom*).

Gaussian response tails, is explained in Section 5.2; subsequent corrections for the impact from processes that induce a p_T imbalance at particle level are discussed in Section 5.3. Afterwards, in Section 5.4, the method is validated using simulated events, and the impact from pile-up is investigated. Systematic uncertainties of the measurement are discussed in Section 5.5. The measured transverse-momentum resolution of Particle-Flow (PF) jets as well as the data-to-simulation ratio are presented in Section 5.6, and in Section 5.7, different strategies are reviewed to correct the MC-truth resolution using that ratio. Finally, in Section 5.8, possible extensions of the maximum-likelihood method are discussed.

5.1 Samples and Event Selection

In the following, the selection of events with a dijet-like topology from data and simulation is described. Except for the trigger requirements, which are applied to data only, the selection criteria are the same in both cases.

5.1.1 Data Sample

Multijet events have been collected from *pp*-collision data acquired in 2011 with the CMS detector. With all subdetectors fully functional, the collected data amount to an integrated luminosity of $855\,\mathrm{pb}^{-1}$. The relative uncertainty on the luminosity measurement is 2.2% [240].

Events have been collected with a level-1 trigger that requires the leading jet to have a transverse momentum p_T^{L1} above a certain threshold. At trigger level, calorimeter jets are used. They are clustered with the anti-k_T algorithm with size parameter $R = 0.5$, and the relative and absolute jet-energy-correction factors are applied, cf. Sections 4.2 and 4.5. Afterwards, a high-level trigger selection has been performed requiring a minimum average transverse momentum $p_{\mathrm{T,HLT}}^{\mathrm{ave}}$ as defined in Eq. (4.10) of the two leading trigger-level jets. Since the resolution will be measured in intervals of $p_{\mathrm{T}}^{\mathrm{ave}}$ as motivated below in Section 5.1.5, inefficiencies of the event selection are avoided by the choice of $p_{\mathrm{T,HLT}}^{\mathrm{ave}}$ as trigger quantity. If, for example, a trigger on the leading-jet p_{T} with some threshold X had been chosen, the minimum interval boundary would have to be at least $p_{\mathrm{T,min}}^{\mathrm{ave}} = 2X$. Thus, a sizable amount of data would be rejected. In case of a smaller $p_{\mathrm{T,min}}^{\mathrm{ave}}$, some combinations of $p_{\mathrm{T},1}$ and $p_{\mathrm{T},2}$ are not possible due to the trigger requirement, and this situation would bias the resolution measurement.

Several trigger-paths have been employed with different thresholds on $p_{\mathrm{T}}^{\mathrm{L1}}$ and $p_{\mathrm{T,HLT}}^{\mathrm{ave}}$. The threshold on $p_{\mathrm{T}}^{\mathrm{L1}}$ has been kept sufficiently below the $p_{\mathrm{T,HLT}}^{\mathrm{ave}}$ threshold to avoid the bias discussed above; for example, the HLT path with $p_{\mathrm{T,HLT}}^{\mathrm{ave}} > 370\,\mathrm{GeV}$ is seeded by an L1 trigger requiring $p_{\mathrm{T}}^{\mathrm{L1}} > 128\,\mathrm{GeV}$. Paths with lower thresholds have successively higher prescales to cope with the strong decrease of the QCD-multijet cross-section with p_{T}. Hence, the data rate is kept sufficiently low while at the same time the number of recorded events with high-p_{T} jets is maximised. All HLT paths are listed in Table 5.1 together with the integrated luminosity of the collected data.

Table 5.1: High-level trigger paths employed for the resolution measurement. Stated are the path name which indicates the threshold on the trigger-level $p_{\mathrm{T,HLT}}^{\mathrm{ave}}$, the threshold on the offline $p_{\mathrm{T}}^{\mathrm{ave}}$ where the trigger efficiency is $\geq 99\%$, and the integrated luminosity of the collected data taking into account trigger prescales. The relative uncertainty on the luminosity measurement is 2.2% [240].

Trigger	$p_{\mathrm{T},99}^{\mathrm{ave}}$ (GeV)	Luminosity (pb^{-1})
`HltDiJetAve30`	45	0.01237
`HltDiJetAve60`	75	0.3360
`HltDiJetAve80`	100	1.286
`HltDiJetAve110`	135	6.07
`HltDiJetAve150`	175	29.04
`HltDiJetAve190`	220	92.8
`HltDiJetAve240`	270	392.3
`HltDiJetAve300`	335	855
`HltDiJetAve370`	405	855

For each HLT path, the efficiency has been measured as a function of the $p_{\mathrm{T}}^{\mathrm{ave}}$ of the calibrated PF jets used in the final analysis. The values $p_{\mathrm{T},99}^{\mathrm{ave}}$ are determined where the efficiency reaches more than 99%.

Conventionally, the efficiency $\epsilon_{A|B}$ of a trigger A is defined with respect to an uncorrelated reference trigger B as the fraction of the number $N_{A\wedge B}$ of events passing A and B

out of the number N_B of events passing B,

$$\epsilon_{A|B}(p_{\mathrm{T}}^{\mathrm{ave}}) = \frac{N_{A\wedge B}(p_{\mathrm{T}}^{\mathrm{ave}})}{N_B(p_{\mathrm{T}}^{\mathrm{ave}})} \ .$$

For a specific $p_{\mathrm{T,HLT}}^{\mathrm{ave}}$ path, a path with a lower threshold is more inclusive and can serve as reference trigger. If the respective paths are prescaled, $\epsilon_{A|B}$ has to be scaled accordingly to ensure $0 \leq \epsilon_{A|B}(p_{\mathrm{T}}^{\mathrm{ave}}) \leq 1$. However, since the prescales are large, the amount of available events is drastically reduced preventing a reliable measurement. This is compensated for by measuring

$$\hat{\epsilon}_{A,B}(p_{\mathrm{T}}^{\mathrm{ave}}) = \frac{N_A(p_{\mathrm{T}}^{\mathrm{ave}})}{N_B(p_{\mathrm{T}}^{\mathrm{ave}})} = \frac{\epsilon_A(p_{\mathrm{T}}^{\mathrm{ave}})}{\epsilon_B(p_{\mathrm{T}}^{\mathrm{ave}})} \ ,$$

which is equivalent to the ratio ϵ_A/ϵ_B of the efficiencies of A and B because A and B select events from the same process, namely QCD-multijet production, and N is proportional to the cross section times trigger efficiency. A typical trigger efficiency is monotonically increasing from 0 to 1 in a turn-on region followed by a plateau at 1. Due to the lower threshold, the turn-on region of ϵ_B is shifted towards lower $p_{\mathrm{T}}^{\mathrm{ave}}$ than that of ϵ_A. Hence, as long as B is fully efficient in the considered region[2], $\hat{\epsilon}_{A,B}$ features the same turn on as ϵ_A. Thus, $p_{\mathrm{T,99}}^{\mathrm{ave}}$ can be readily computed from $\hat{\epsilon}_{A,B}(p_{\mathrm{T,99}}^{\mathrm{ave}}) = 0.99 \cdot \hat{\epsilon}_{A,B}(p_{\mathrm{T}}^{\mathrm{ave}} \to \infty)$.

Figure 5.2 depicts $\hat{\epsilon}$ measured in intervals of $p_{\mathrm{T}}^{\mathrm{ave}}$ for the different HLT paths used in this analysis, where in each case the path with the next lower threshold has been used as reference. For technical reasons, the uncertainties are calculated using Gaussian error-propagation. The $\hat{\epsilon}$ are normalised such that the plateau $\hat{\epsilon}(p_{\mathrm{T}}^{\mathrm{ave}} \to \infty)$ lies at 1. They are fitted with the smooth function

$$f(p_{\mathrm{T}}^{\mathrm{ave}}) = \frac{1}{2}\left[\mathrm{erf}\left(p_0 \cdot (p_{\mathrm{T}}^{\mathrm{ave}} - p_1)\right) + 1\right] \ ,$$

from which $p_{\mathrm{T,99}}^{\mathrm{ave}}$ is determined via

$$f(p_{\mathrm{T,99}}^{\mathrm{ave}}) = 0.99\,.$$

The $p_{\mathrm{T,99}}^{\mathrm{ave}}$ of the employed trigger paths are stated in Table 5.1. For the path with the lowest threshold, `HltDiJetAve30`, $p_{\mathrm{T,99}}^{\mathrm{ave}}$ has been deduced from studies of previously implemented trigger paths using uncalibrated jets at trigger level [239].

5.1.2 Simulated Sample

In this analysis, the simulated QCD-multijet events from the PYTHIA sample described in Section 4.3.2 are used. Event weights w_i are applied in order to regain the realistic QCD spectrum and to model the pile-up conditions expected for data, as described ibid. In order to exploit the maximal statistical precision, the absolute normalisation λ is chosen separately for each interval in which the resolution will be measured, cf. Section 5.1.5.

---

[2] This is the case as can be seen from the obtained efficiency curves shown in Fig. 5.2. In any case, ϵ_B saturates well before ϵ_A, and hence, $\hat{\epsilon}_{A,B}$ is at the most affected at the begin of its turn-on region far away from the saturation region.

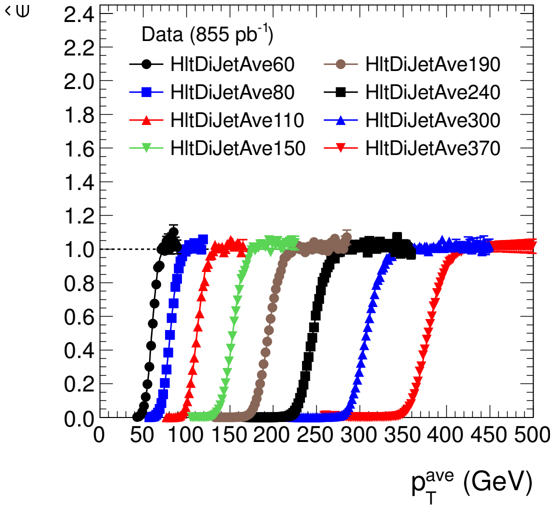


Figure 5.2: Efficiencies $\hat{\epsilon}$ of the employed trigger paths as a function of $p_{\mathrm{T}}^{\mathrm{ave}}$. See the text for a definition of $\hat{\epsilon}$ and an explanation why it exceeds 1 in certain cases.

Therefore, a weight

$$\lambda \cdot w_i = \frac{N}{\sum_{k=1}^{N} w_k} \cdot w_i$$

is assigned to each event i, such that the sum of weights equals the number N of simulated events in that interval.

In the following, simulated events are treated in the same way as data.

5.1.3 Event Reconstruction

All physics objects are reconstructed with the PF algorithm Section 3.5.2. Jets are clustered with the anti-k_T jet algorithm with parameter size $R = 0.5$, and the jet-energy-scale calibration is applied including the residual correction for data. In case of the simulation, generator-level jets are clustered by applying the same jet algorithm to all stable generator-level particles.

5.1.4 Event Cleaning

Several selection criteria (*cleaning filters*) are employed in order to identify events where the reconstruction failed or has been spoiled by instrumental effects or which originated from beam-background processes. Although their fraction is small, the affected events typically feature severely misreconstructed jets. Hence, they populate the extreme tails of the investigated asymmetry Eq. (4.9) and corresponding response distributions and possibly impact the response-tail measurement described in Chapter 6. Therefore, these events are not used for further analysis.

The cleaning filters are mostly standard tools provided by CMS and have partly been developed within the context of the new-physics search [237] described in Chapter 7. Since the results of the resolution measurement are employed by the latter, exactly the same filters are applied here. They are described in the following:

Primary vertex and beam halo: Particles originating from displaced vertices in satellite collisions or from beam-related processes of protons with residual gas molecules in the beam pipe (*beam halo*) possibly prevent a proper event reconstruction. Therefore, each event is required to have at least one high-quality primary vertex that has been reconstructed from more than four tracks and is located within 24 cm in z and 2 cm in xy direction from the nominal interaction point (*good-vertex filter*). Furthermore, the CSC subdetector is utilised to identify and reject events with muons which move parallel to the beam and thus likely stem from beam-halo processes (*beam-halo filter*) [164, 241].

Anomalous calorimeter signals: In some events, particles hitting the readout electronics, scintillation fibres, or photomultipliers cause anomalous signals in the ECAL and HCAL. Dedicated criteria have been defined to identify and reject these events [242, 243]. In rare cases, electronic noise in the ECAL readout-system produces fake large energy deposits at random times which can overlap with a collision event. These anomalies can be identified using timing and pulse-shape information (*HBHE-noise filter*) [244]. Additional rare noise has been identified to affect the ECAL endcaps coherently with the muon system. This has been suppressed by selecting events with less than 2500 energy deposits in the endcaps (*EE-noise filter*) [237, 238, 245].

Tracking failure: Track reconstruction can fail in events with too many seed clusters or in events where a proton in a satellite bunch enters the hard collision. Therefore, the fraction of tracks passing high-quality criteria is required to be greater than 25%, if there are at least ten tracks in the event (*beam-scraping filter*). In addition, the scalar sum of the p_T of tracks associated to the primary vertex is required to be greater than 10% of the scalar sum of the p_T of all jets within the tracker acceptance (*tracking-failure filter*). Moreover, events with seriously misreconstructed muons are rejected (*PF-post-processing filter*) [237, 238, 245].

Dead ECAL-cells: About 1% of the ECAL crystals are not read out because of malfunctioning electronics. The energy deposited in these crystals is lost for the object reconstruction, which affects in particular the jet measurement [246]. The affected events are identified and rejected based on either the information from the parallel readout chain of the trigger system (*TP filter*) or, in case this is not available, on the amount of energy deposited in the crystals surrounding the affected regions (*BE filter*) [244].

In addition, events are rejected that contain isolated electrons or muons with $p_\mathrm{T} > 10\,\mathrm{GeV}$ and $|\eta| < 2.4$ or $|\eta| < 2.5$, respectively (*lepton veto*). This criterion is employed by the new-physics search Chapter 7 (the isolation criterion is explained ibid.) to remove $t\bar{t}$ and $W+\text{jets}$ events where the W bosons decay to leptons. Here, it is applied to ensure a synchronisation to the search analysis, but it has in fact negligible effect on the measured resolution[3]. The efficiencies of all filter requirements as well as the lepton veto are listed in Table A.1 in Appendix A.1.

Furthermore, events are rejected if one of the two leading jets failed dedicated quality criteria (*jet-id*), which identify most jets clustered from detector noise while keeping more than 99% of the real jets [217, 247].

---

[3] In rare cases, e. g. $W+\text{jets}$ events can fake QCD-dijet events with large asymmetries, which is of importance for the response-tail measurement in Chapter 6, though.

In total, approximately 3% of the events are rejected both in data and in simulation.

5.1.5 QCD-Dijet Selection

The emission of additional partons in $2 \to 2$ parton processes can compromise the idealised dijet topology of two jets with exactly balanced transverse momenta as discussed above in Section 4.6.2. The following selection criteria have been applied in order to enhance the number of events which resemble the ideal dijet case.

Events with at least two jets are considered. Dijet-like events are selected based on the azimuthal angle $\Delta\phi$ between the two leading jets,

$$\Delta\phi = \Delta\phi(\vec{p}_{\mathrm{T},1}, \vec{p}_{\mathrm{T},2}) \quad \text{with} \quad 0 < \Delta\phi < \pi \ ,$$

and on the fractional transverse momentum α of the third jet relative to the dijet transverse-momentum scale,

$$\alpha = \frac{p_{\mathrm{T},3}}{p_{\mathrm{T}}^{\mathrm{ave}}} \ , \tag{5.1}$$

The correlation between both variables is shown in Fig. 5.3.

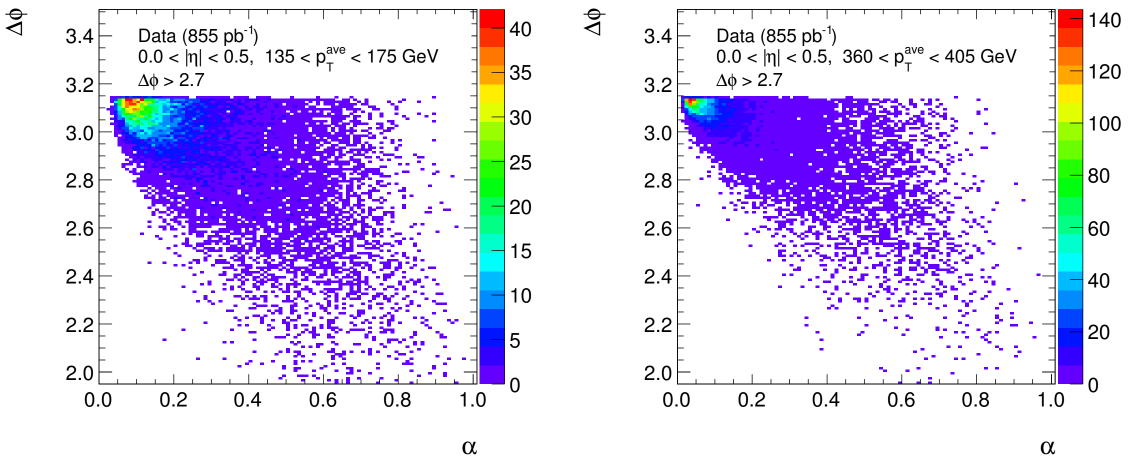


Figure 5.3: Correlation of the dijet-selection variables $\Delta\phi$ and α for data in a low (*left*) and medium (*right*) $p_{\mathrm{T}}^{\mathrm{ave}}$ interval for $|\eta| < 0.5$. (Note the reduced range of the $\Delta\phi$ axis.)

Large values of $\Delta\phi$ correspond to an event topology where the two leading jets point in opposite directions in the transverse plane. This is the case in most events because hard partons are typically emitted under small angles during the parton-showering process, cf. Section 4.1. Throughout this analysis, a threshold of

$$\Delta\phi > 2.7$$

is imposed. This criterion retains most of the events but sufficiently suppresses severely imbalanced ones.

Small values of α indicate transverse-momentum balance between the two leading jets because little transverse momentum is transferred from the original parton to additional ISR/FSR jets. This is further investigated in Section 5.3.1. A significant number of events with large momentum imbalance is present. In the following, when the method will be described, a default threshold of $\alpha < 0.14$ is chosen for illustration purposes, which ensures sufficient momentum balance and an adequate size of the selected data sample. Eventually, the measurement will be performed for different α thresholds as explained in Section 5.3.1. The detector signature of a typical selected dijet event is visualised in Fig. 4.5.

Distributions of $\Delta\phi$ and α after application of the selection criterion on the respective other variable are shown in Fig. 5.4 for data and simulation. The presence of events with large p_{T} imbalance and the predominantly back-to-back topology discussed above are overall well described by the simulation. Some deviations occur for example at low α, though, which reflect the challenges related to the modelling of parton-showering processes.

Several SM processes exist that can fake the QCD-dijet signature, e. g. W + jets and Z + jets events where the W and Z bosons decay hadronically or into a hadronically-decaying τ lepton or, in case of the Z, into two neutrinos[4]. However, the production cross-sections for these processes are lower by orders of magnitude compared to QCD-multijet production, cf. Fig. 3.5, and their rate is further suppressed by the described dijet selection, in particular the $\Delta\phi$ criterion. They are therefore neglected in the following.

Table 5.2: $|\eta|$ and $p_{\mathrm{T}}^{\mathrm{ave}}$ interval boundaries for the resolution measurement.

$\lvert\eta\rvert$	$p_{\mathrm{T}}^{\mathrm{ave}}$ (GeV)
0.0, 0.5	45, 75, 100, 135, 175, 195, 220, 240, 270, 279, 294, 312, 335, 360, 405, 498, 1500
0.5, 1.1	45, 75, 100, 135, 175, 195, 220, 240, 270, 279, 294, 312, 335, 360, 405, 486, 1500
1.1, 1.7	45, 75, 100, 135, 175, 220, 270, 294, 335, 405, 1500
1.7, 2.3	45, 75, 100, 135, 175, 220, 270, 335, 405, 1500
2.3, 5.0	45, 75, 100, 135, 175, 220

In order to account for the transverse momentum and pseudorapidity dependence of the jet resolution, cf. Section 4.4.1, events are additionally distributed into intervals of $p_{\mathrm{T}}^{\mathrm{ave}}$, and both leading jets are required to lie within the same $|\eta|$ interval. By construction, the $p_{\mathrm{T}}^{\mathrm{ave}}$ resolution is better than the jet-p_{T} resolution by approximately a factor $\sqrt{2}$. Hence, the choice of $p_{\mathrm{T}}^{\mathrm{ave}}$ as measure of the event's transverse-momentum scale is preferential to, for example, the p_{T} of the leading jet because in this way the migration effects described in Section 4.4.1 are reduced.

In Table 5.2, the $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ interval boundaries are listed. The $|\eta|$ boundaries have been chosen according to the coverage of the different subdetectors as well as the available number of events. For example, the $1.1 < |\eta| < 1.7$ interval covers the transition region between the barrel and the endcap detectors, where worse resolution is expected, cf. Fig. 4.20. The $p_{\mathrm{T}}^{\mathrm{ave}}$ boundaries have been chosen primarily to coincide with the $p_{\mathrm{T},99}^{\mathrm{ave}}$ trigger-thresholds. Intervals that contain enough events for sufficient statistical precision have been further subdivided to a size which is of the order of the jet-p_{T} resolution. In data, only events collected by a fully efficient trigger path are considered in each interval to avoid biasing

---

[4] Recall that events with isolated leptons are rejected.

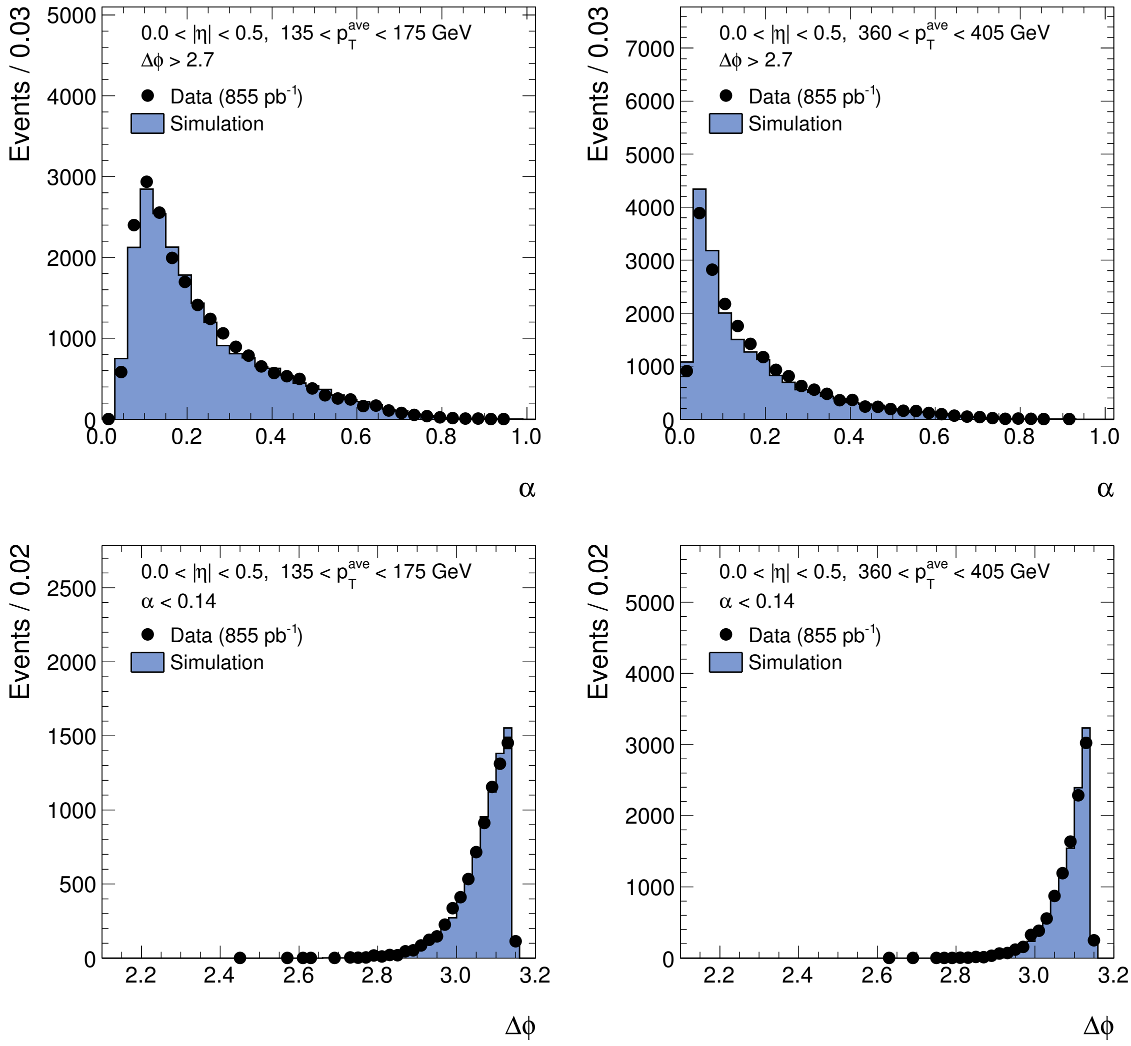


Figure 5.4: α (*top*) and $\Delta\phi$ (*bottom*) distributions for data (*solid circles*) and simulation (*filled histogram*) in a low (*left*) and medium (*right*) $p_{\mathrm{T}}^{\mathrm{ave}}$ interval for $|\eta| < 0.5$. The selection criterion with respect to the other variable has been applied ('n − 1 plots'). The integral of the simulated distribution has been normalised to the integral of the distribution measured in data.

the measurement by selection effects due to trigger inefficiencies.

The distributions of η before and after the $|\eta|$ selection-criterion (the two leading jets in the same $|\eta|$ interval) are shown in Fig. 5.5, and the spectra of $p_{\mathrm{T}}^{\mathrm{ave}}$ and the transverse momenta of the first three jets are depicted in Figs. 5.6 and 5.7 for one example $p_{\mathrm{T}}^{\mathrm{ave}}$ interval and inclusive in $p_{\mathrm{T}}^{\mathrm{ave}}$, respectively. In all cases, the $\Delta\phi$- and α-based dijet-selection criteria have been applied. The distributions reflect the decrease of the jet production cross-section Fig. 4.4 with rapidity and p_{T}. At low p_{T}, the impact of the trigger prescales is visible. The data are well described by the simulation with exception of the aforementioned differences at very low $p_{\mathrm{T},3}$.

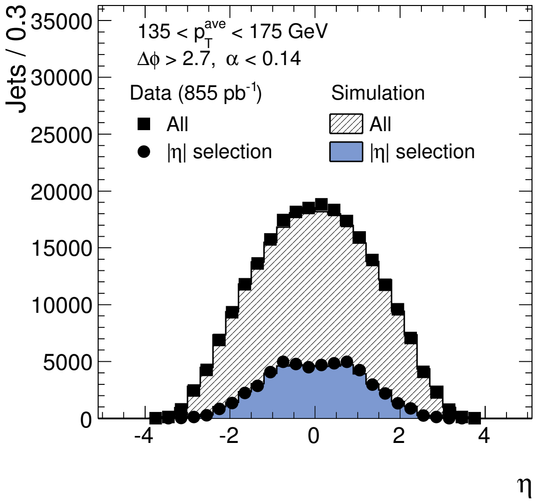


Figure 5.5: η distribution of the leading two jets for data (*markers*) and simulation (*histograms*) in case both jets lie in the same $|\eta|$ interval (*circles and filled histogram*) and in all cases (*squares and hatched histogram*) after the dijet selection $\Delta\phi > 2.7$ and $\alpha < 0.14$. In each case, the integral of the simulated distribution has been normalised to the integral of the distribution measured in data.

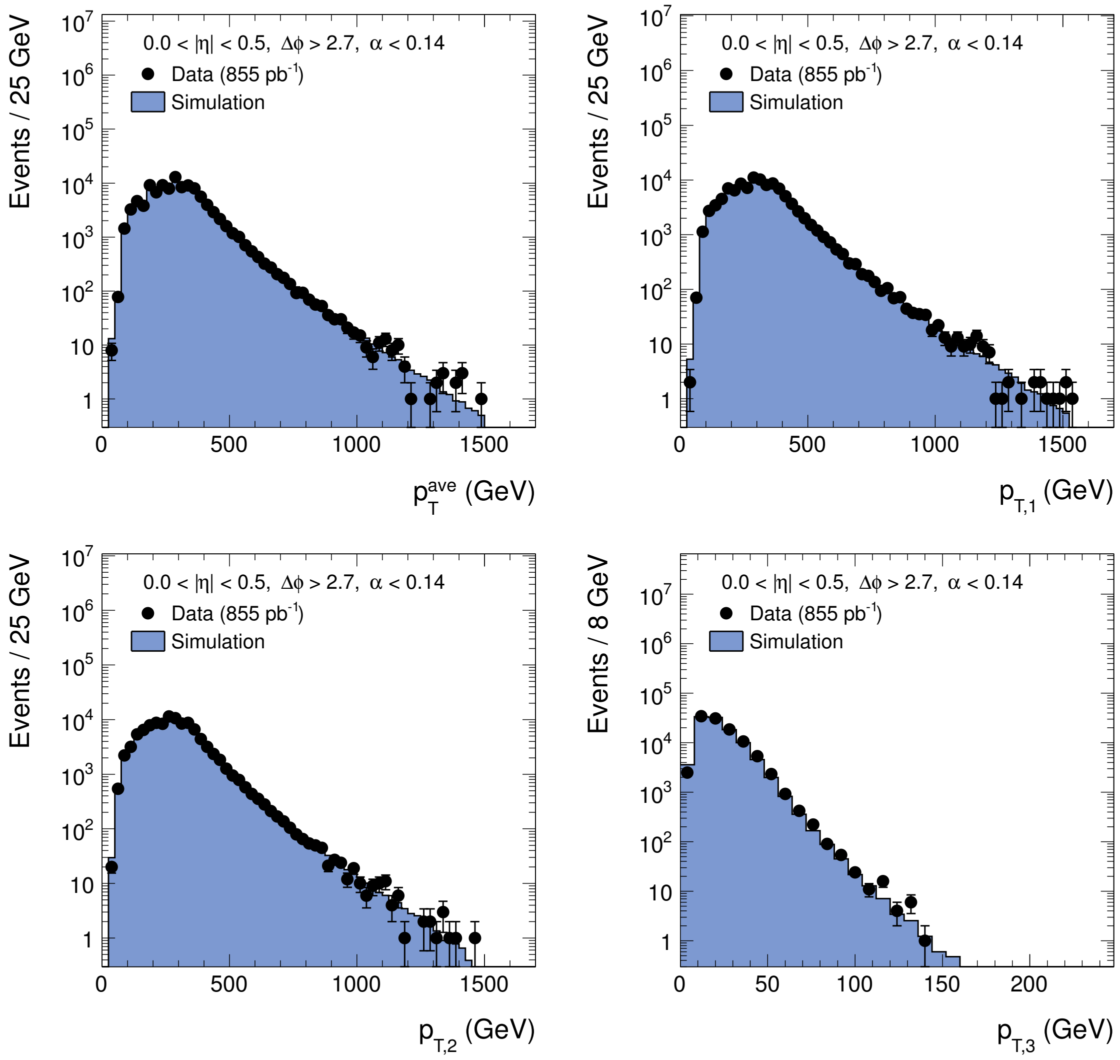


Figure 5.6: Average transverse momentum of the leading two jets (*top-left*) and transverse momenta of the first three jets for data (*solid circles*) and simulation (*filled histogram*) for the $|\eta| < 0.5$ interval after the dijet selection $\Delta\phi > 2.7$ and $\alpha < 0.14$. The impact of the trigger prescales is visible at low p_{T} and has been considered in the weights of the simulated events. In each case, the integral of the simulated distribution has been normalised to the integral of the distribution measured in data.

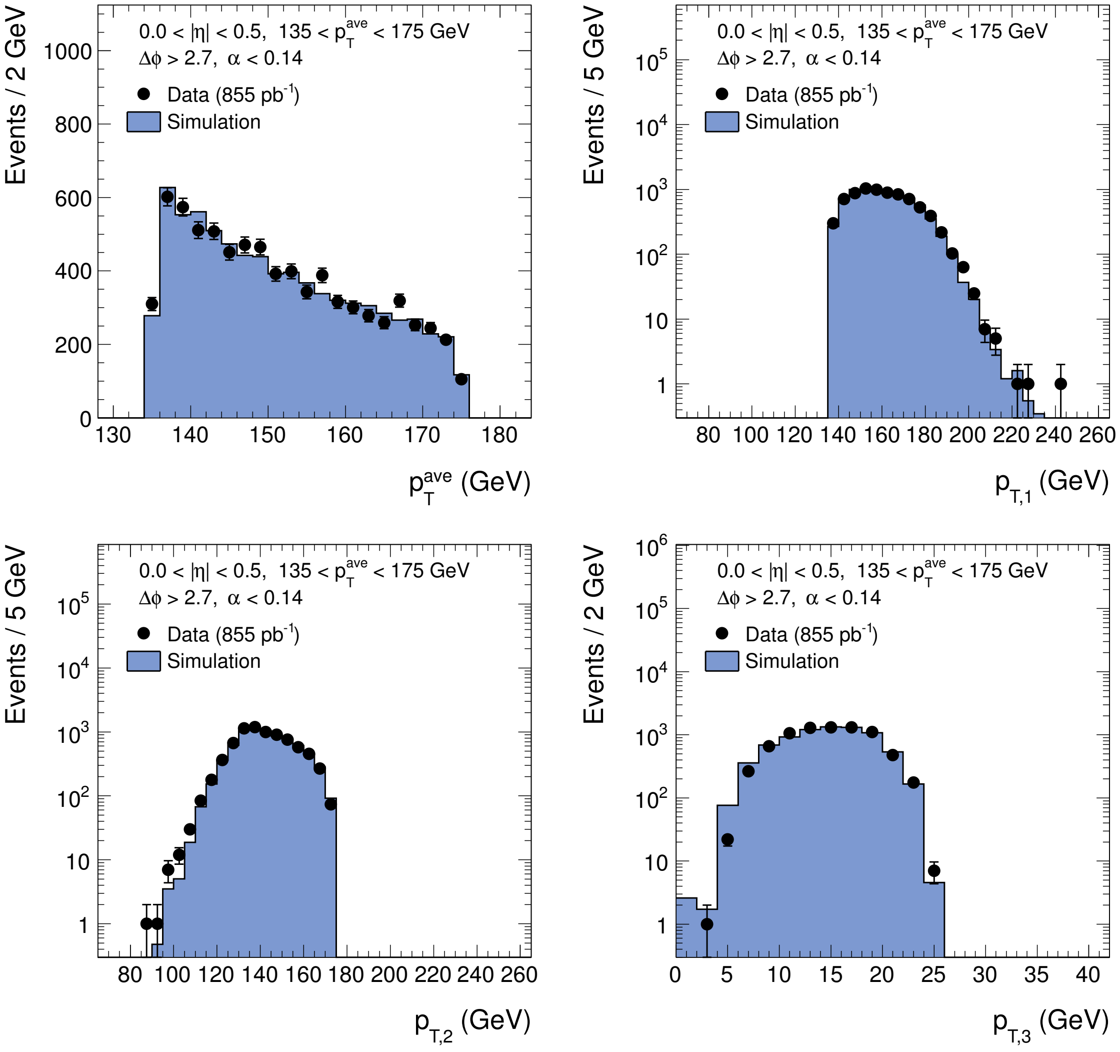


Figure 5.7: Average transverse momentum of the leading two jets (*top-left*) and transverse momenta of the first three jets in the selected dijet events for data (*solid circles*) and simulation (*filled histogram*) for the $135 < p_T^{ave} < 175\,\text{GeV}$ and $|\eta| < 0.5$ interval after the dijet selection $\Delta\phi > 2.7$ and $\alpha < 0.14$. In each case, the integral of the simulated distribution has been normalised to the integral of the distribution measured in data.

5.2 Description of the Method

In the following, the technical details of the unbinned maximum-likelihood method are explained. The concept is established in Section 5.2.1, where a generic likelihood of the measured jet p_T in ideal dijet events is defined, which depends on the jet-p_T response function and the particle-level jet-p_T spectrum. As before in Eq. (4.2), the response is defined with respect to the particle-level jet-p_T. The parameters of the response function can be determined by maximisation of the likelihood. In Sections 5.2.2 to 5.2.4, successive adjustments to the generic likelihood are described which allow an application to the selected dijet data. Firstly, the likelihood expression for a Gaussian response is derived, and a parametrisation with a p_T-independent resolution in a small $p_\mathrm{T}^\mathrm{ave}$ interval is motivated. Secondly, the arising migration effects are incorporated into the likelihood, which leads to an improved estimate of the average particle-level jet-p_T in that interval. Thirdly, a correction for the impact from the presence of non-Gaussian response tails is developed. The resulting dijet likelihood as used in this analysis is summarised in Section 5.2.5, and numerical approximations are discussed. A validation of the likelihood fit with data and simulated events is presented in Section 5.2.6.

5.2.1 An Unbinned Maximum-Likelihood Approach for Ideal Dijet Events

As a start, a likelihood is constructed for ideal dijet events with exactly two jets in the final state. The jets are assumed to be balanced in transverse momentum at particle level[5]. Hence, the probability density function (pdf)[6] $g_{\boldsymbol{\xi}}$ to measure a certain combination of transverse momenta in a dijet event may be defined as

$$g_{\boldsymbol{\xi}}\left(p_{\mathrm{T},1}, p_{\mathrm{T},2}\right) \propto \int \mathrm{d}p_\mathrm{T}^\mathrm{true}\, f\left(p_\mathrm{T}^\mathrm{true}\right) \cdot m_{\boldsymbol{\xi}}\left(p_{\mathrm{T},1}|p_\mathrm{T}^\mathrm{true}\right) \cdot m_{\boldsymbol{\xi}}\left(p_{\mathrm{T},2}|p_\mathrm{T}^\mathrm{true}\right)\ , \tag{5.2}$$

where $p_{\mathrm{T},i}$ and $p_\mathrm{T}^\mathrm{true}$ are the transverse momenta of the i-th jet at detector level, $i = 1, 2$, and at particle level, respectively. The function f denotes the pdf of $p_\mathrm{T}^\mathrm{true}$, i. e. the particle-level differential jet cross-section. The function $m_{\boldsymbol{\xi}}$ denotes the pdf to measure $p_{\mathrm{T},i}$ for the i-th jet given $p_\mathrm{T}^\mathrm{true}$ and depends on the parameters $\boldsymbol{\xi}$ that need to be determined. The two jet-p_T measurements are assumed to be uncorrelated, which is reasonable considering that jets are observed in well separated detector regions. In general, any functional form can be chosen for $m_{\boldsymbol{\xi}}$, in particular also an asymmetric one. In this analysis, a Gaussian parametrisation will be used though as discussed below.

The pdf $r_{\boldsymbol{\xi}}$ of the jet-p_T response $\mathcal{R} = p_\mathrm{T}/p_\mathrm{T}^\mathrm{true}$ is obtained from $m_{\boldsymbol{\xi}}$ by parameter transformation [223]

$$r_{\boldsymbol{\xi}}\left(\mathcal{R}|p_\mathrm{T}^\mathrm{true}\right) = m_{\boldsymbol{\xi}}\left(p_\mathrm{T}\left(\mathcal{R}\right)|p_\mathrm{T}^\mathrm{true}\right) \cdot \left|\frac{\mathrm{d}p_\mathrm{T}}{\mathrm{d}\mathcal{R}}\right|. \tag{5.3}$$

It is important to note that $g_{\boldsymbol{\xi}}$ does not depend on $p_\mathrm{T}^\mathrm{true}$ since the latter is an integration variable. Hence, all possible values of $p_\mathrm{T}^\mathrm{true}$ contribute to $g_{\boldsymbol{\xi}}$ with the individual weights $f(p_\mathrm{T}^\mathrm{true})$. As discussed below, f is taken from the simulation.

---

[5] Biases of the method due to this assumption are discussed in the next Section 5.3.

[6] Not to be confused with the parton-distribution function PDF, although the latter is also a pdf...

For a sample of N dijet events, where the events can be considered statistically independent, a likelihood is defined as

$$\mathcal{L}\left(\boldsymbol{\xi}\right) = \prod_{k=1}^{N} g_{\boldsymbol{\xi},k}\left(p_{\mathrm{T},1}, p_{\mathrm{T},2}\right) \ , \tag{5.4}$$

where $g_{\boldsymbol{\xi},k}(p_{\mathrm{T},1}, p_{\mathrm{T},2})$ denotes the pdf of the k-th dijet event. According to the principle of maximum likelihood [223, 224], the values $\hat{\boldsymbol{\xi}}$ of the parameters $\boldsymbol{\xi}$ which maximise $\mathcal{L}$ correspond to the response $r_{\hat{\boldsymbol{\xi}}}$ that is most compatible with the data.

Since the $\hat{\boldsymbol{\xi}}$ cannot be determined analytically, the programme LVMINI [248, 249] is employed in this analysis to find the maximum. In practice, it is easier to use the *log-likelihood* $\ln \mathcal{L}(\boldsymbol{\xi})$ because it turns the product of pdfs in $\mathcal{L}$ into a sum. In fact, just as similar other programmes, LVMINI does not search for its maximum but for the minimum of the negative log-likelihood function[7]

$$\mathcal{F}(\boldsymbol{\xi}) = -2 \cdot \ln \mathcal{L}(\boldsymbol{\xi}) \ .$$

LVMINI combines the *limited-memory Broyden-Fletcher-Goldfarb-Shanno* (L-BFGS) algorithm [250, 251] for iterative function minimisation with an inexact line-search method based on the Wolfe conditions. A brief description of the strategy behind the L-BFGS algorithm and LVMINI is provided in Appendix A.2; detailed information can be found for example in [223, 249, 251, 252].

LVMINI has been shown to perform equally well to standard minimisation tools such as MINUIT [253, 254] in case of problems with about 10 free parameters, but it is in particular also designed for a large number up to $\mathcal{O}(10^5)$ parameters. Although the dijet likelihood used in this analysis depends on only one free parameter as will be shown below, LVMINI is utilised because the measurement has been implemented into an existing framework for jet-energy calibration based on a maximum-likelihood fit with $\mathcal{O}(10^2)$ parameters, which employs LVMINI for minimisation. Furthermore, possible extensions to the dijet likelihood include the usage of a p_{T} and η dependent response parametrisation, also with a function more complex than a Gaussian. In these cases, the number of free parameters can easily exceed $10 - 15$, which is the maximum for example MINUIT can handle practically [254].

5.2.2 Adjustment 1: Gaussian Response And p_{T}-Independent Resolution

In the following, a Gaussian jet-p_{T} response $r_{\boldsymbol{\xi}}$ is assumed motivated by the discussion above in Section 4.4. The Gaussian has a fixed mean value of 1, i. e. the correct jet energy scale is assumed. Hence, also $m_{\boldsymbol{\xi}}(p_{\mathrm{T},i}|p_{\mathrm{T}}^{\mathrm{true}})$ in Eq. (5.2) is a Gaussian,

$$m_{\sigma}\left(p_{\mathrm{T},i}|p_{\mathrm{T}}^{\mathrm{true}}\right) = \frac{1}{\sqrt{2\pi}\sigma}\,\mathrm{e}^{-\frac{1}{2}\left(\frac{p_{\mathrm{T},i}-p_{\mathrm{T}}^{\mathrm{true}}}{\sigma}\right)^2} \ ,$$

---

[7] Since the logarithm is a strictly monotonic function, $\mathcal{F}$ has the same maximum as $\mathcal{L}$.

which depends on the parameter σ, that is the jet-p_{T} resolution. In that case, a different choice of coordinates than $p_{\mathrm{T},1}$ and $p_{\mathrm{T},2}$ is preferential. With

$$\begin{aligned} p_{\mathrm{T}}^{\mathrm{ave}} &= \frac{1}{2}\left(p_{\mathrm{T},1} + p_{\mathrm{T},2}\right) \\ \Delta p_{\mathrm{T}} &= \frac{1}{2}\left(p_{\mathrm{T},1} - p_{\mathrm{T},2}\right) \ , \end{aligned} \tag{5.5}$$

where Δp_{T} is always positive due to the jet ordering, the dijet pdf Eq. (5.2) becomes [223]

$$\begin{aligned} g_{\sigma'}\left(p_{\mathrm{T}}^{\mathrm{ave}}, \Delta p_{\mathrm{T}}\right) &= g_{\sigma'}\left(p_{\mathrm{T},1}\left(p_{\mathrm{T}}^{\mathrm{ave}}, \Delta p_{\mathrm{T}}\right), p_{\mathrm{T},2}\left(p_{\mathrm{T}}^{\mathrm{ave}}, \Delta p_{\mathrm{T}}\right)\right) \cdot |J| \\ &\propto \underbrace{\frac{1}{\pi\sigma'}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{\Delta p_{\mathrm{T}}}{\sigma'}\right)^2}}_{(A)} \cdot \underbrace{\int \mathrm{d}p_{\mathrm{T}}^{\mathrm{true}}\, f\left(p_{\mathrm{T}}^{\mathrm{true}}\right) \cdot \mathrm{e}^{-\frac{1}{2}\left(\frac{p_{\mathrm{T}}^{\mathrm{ave}} - p_{\mathrm{T}}^{\mathrm{true}}}{\sigma'}\right)^2}}_{(B)} \ , \end{aligned} \tag{5.6}$$

where the Jacobian $|J|$ of the transformation Eq. (5.5) is $|J| = 2$ and

$$\sigma' = \frac{\sigma}{\sqrt{2}} \ .$$

It is evident from Eq. (5.6) that the dijet pdf consists in fact of two components: (A) the pdf of Δp_{T}, i.e. the jet-p_{T} imbalance at detector level, and (B) the pdf of $p_{\mathrm{T}}^{\mathrm{ave}}$, which takes into account the particle-level jet-p_{T} spectrum.

Since the jet-p_{T} resolution σ depends on p_{T} and η, cf. Section 4.6.1, also σ' is a function of p_{T} and η, where the functional form depends on some parameters $\boldsymbol{\xi}'$ that have to be determined by the maximisation. As will be shown later, the result has to be corrected for the impact from additional jet activity using an extrapolation procedure. Since it is not straightforward to apply the procedure to the set of (possibly correlated) parameters $\boldsymbol{\xi}'$, a simplified approach is chosen in the following. Events are selected in small intervals of $p_{\mathrm{T}}^{\mathrm{ave}}$ and $|\eta|$, cf. Section 5.1. In each interval, the dijet pdf Eq. (5.6) is defined to depend solely on one p_{T}-independent parameter

$$\sigma' = \langle \sigma'_{\boldsymbol{\xi}'}(p_{\mathrm{T}}, |\eta|) \rangle \ , \tag{5.7}$$

which is assumed to describe the data adequately. This is justified since the MC-truth resolution determined in Section 4.6.1 varies by less than 10% in the chosen $p_{\mathrm{T}}^{\mathrm{ave}}$ intervals Table 5.2 above 100 GeV and except for the last, high-p_{T} interval.

5.2.3 Adjustment 2: Description of Migration Effects

As described above, events are selected in intervals of $p_{\mathrm{T}}^{\mathrm{ave}}$ and $|\eta|$. Therefore, asymmetric migration occurs at the boundaries of the selected range $p_{\mathrm{T,min}}^{\mathrm{ave}} < p_{\mathrm{T}}^{\mathrm{ave}} < p_{\mathrm{T,max}}^{\mathrm{ave}}$ because of the finite jet-p_{T} resolution in combination with the monotonically decreasing jet-p_{T} spectrum as explained in Fig. 4.10 and Section 4.4.1. As discussed ibid., the impact from migration in η is negligible.

The $p_{\mathrm{T}}^{\mathrm{ave}}$ migration effects are clearly visible in the $p_{\mathrm{T}}^{\mathrm{gen}}$ distributions of simulated events to which the full dijet selection at detector-level has been applied, cf. Fig. 5.8 and also Figs. A.1 to A.5 in Appendix A.3. In general, the shape of the distributions is dominated by the jet-p_T resolution because the size of the $p_\mathrm{T}^\mathrm{ave}$ intervals is of the same order. Merely the last interval with highest $p_\mathrm{T}^\mathrm{ave}$ is much broader, and therefore, the distribution resembles the jet-p_T spectrum.

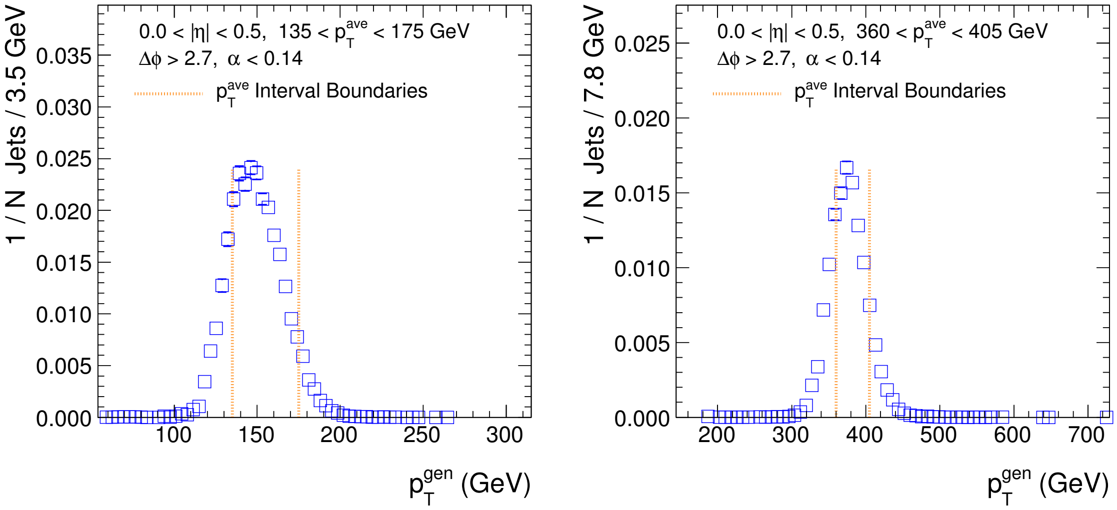


Figure 5.8: $p_\mathrm{T}^\mathrm{gen}$ distributions (*open squares*) of the selected dijet events in a low (*left*) and a medium (*right*) $p_\mathrm{T}^\mathrm{ave}$ interval for $|\eta| < 0.5$ and $\alpha < 0.14$. Selection effects discussed in the text result in migration at the interval boundaries (*dotted lines*).

In order to incorporate the $p_\mathrm{T}^\mathrm{ave}$ selection criteria into the dijet likelihood Eq. (5.6) and thus avoid biases due to migration, the jet-p_T spectrum $f(p_\mathrm{T}^\mathrm{true})$ is modified as

$$f\left(p_\mathrm{T}^\mathrm{true}\right) \to f_{\sigma'}\left(p_\mathrm{T}^\mathrm{true} | p_\mathrm{T,min}^\mathrm{ave}, p_\mathrm{T,max}^\mathrm{ave}\right) ,$$

with

$$f_{\sigma'}\left(p_\mathrm{T}^\mathrm{true} | p_\mathrm{T,min}^\mathrm{ave}, p_\mathrm{T,max}^\mathrm{ave}\right) \propto f\left(p_\mathrm{T}^\mathrm{true}\right) \cdot \int_{p_\mathrm{T,min}^\mathrm{ave}}^{p_\mathrm{T,max}^\mathrm{ave}} \mathrm{d}p_\mathrm{T}^\mathrm{ave} \, \frac{1}{\sqrt{2\pi}\sigma'} \, \mathrm{e}^{-\frac{1}{2}\left(\frac{p_\mathrm{T}^\mathrm{ave} - p_\mathrm{T}^\mathrm{true}}{\sigma'}\right)^2} . \tag{5.8}$$

The Gaussian integrand corresponds to the pdf of $p_\mathrm{T}^\mathrm{ave}$ given $p_\mathrm{T}^\mathrm{true}$. Hence, the integral equals the probability to measure $p_\mathrm{T}^\mathrm{ave}$ between the interval boundaries $p_\mathrm{T,min}^\mathrm{ave}$ and $p_\mathrm{T,max}^\mathrm{ave}$.

In this analysis, the particle-level jet-p_T spectrum f is obtained from simulated events using generator-truth information. The usual dijet selection criteria are applied at generator level, i. e. the leading two generator-level jets in each event have to lie within the same $|\eta^\mathrm{gen}|$ interval and to be back-to-back in the transverse plane by $\Delta\phi(\vec{p}_\mathrm{T,1}^\mathrm{\,gen}, \vec{p}_\mathrm{T,2}^\mathrm{\,gen}) > 2.7$. Additional jet activity is restricted by requiring $\alpha^\mathrm{gen} = p_\mathrm{T,3}^\mathrm{gen} / p_\mathrm{T}^\mathrm{gen,ave} < 0.14$. For each selected event, the $p_\mathrm{T}^\mathrm{gen}$ of both leading jets are filled into a histogram. The resulting $p_\mathrm{T}^\mathrm{gen}$ distributions are depicted in Fig. 5.9 for different $|\eta|$ intervals. They correspond to the measured jet-production cross-section Fig. 4.4 and are used as $f(p_\mathrm{T}^\mathrm{true})$ in Eq. (5.8). Uncertainties arising from the usage of the simulated spectra are evaluated later in Section 5.5.

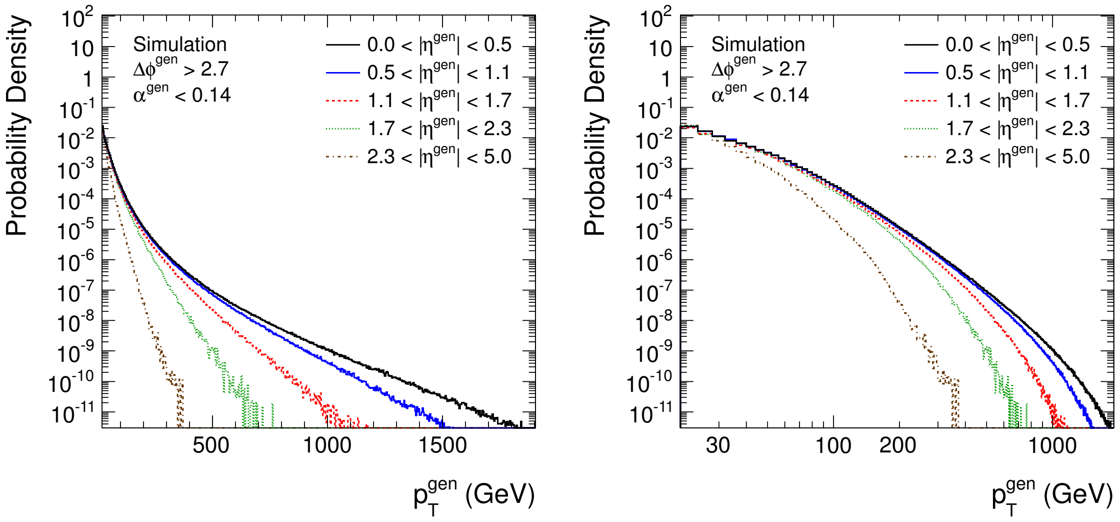


Figure 5.9: $p_{\mathrm{T}}^{\mathrm{gen}}$ spectra of QCD-dijet events simulated with PYTHIA in different $|\eta^{\mathrm{gen}}|$ intervals on a single- (*left*) and double-logarithmic scale (*right*). For each event, the $p_{\mathrm{T}}^{\mathrm{gen}}$ of both leading jets have been filled into the histogram, and a linear interpolation of the bin content has been performed.

The spectrum Eq. (5.8) is also used to obtain the expectation value

$$\langle p_{\mathrm{T}}^{\mathrm{true}} \rangle = \int \mathrm{d}p_{\mathrm{T}}^{\mathrm{true}}\; p_{\mathrm{T}}^{\mathrm{true}} f_{\hat{\sigma}'}(p_{\mathrm{T}}^{\mathrm{true}} | p_{\mathrm{T,min}}^{\mathrm{ave}}, p_{\mathrm{T,max}}^{\mathrm{ave}}) \tag{5.9}$$

for the average particle-level jet-p_{T} in each $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ interval. Its uncertainty is determined by propagating the uncertainty on $\hat{\sigma}'$. Thus, in the following, the relative jet-p_{T} resolution will be given as

$$\text{relative resolution} = \frac{\hat{\sigma}'}{\langle p_{\mathrm{T}}^{\mathrm{true}} \rangle}\;. \tag{5.10}$$

5.2.4 Adjustment 3: Correction for Biases from Non-Gaussian Response Tails

A Gaussian jet-p_{T} response function leads to the dijet pdf Eq. (5.6) and is equivalent to a Gaussian pdf of Δp_{T}. However, the actual jet response of the CMS detector features non-Gaussian tails as discussed previously in Section 4.4.2 and later in Chapter 6. This is also visible in the Δp_{T} distributions shown in Fig. 5.10 for data and simulation. As a consequence, the dijet likelihood will be biased due to the larger tail component. Therefore, only events where the measured transverse momenta lie within the bulk region are considered for the measurement of the Gaussian width.

As observed in Fig. 5.10, non-Gaussian tail components are present for $\Delta p_{\mathrm{T}} \gtrsim 2.5\,\sigma'$. Thus, the likelihood is restricted to events with

$$\Delta p_{\mathrm{T}} < \Delta p_{\mathrm{T}}^{\mathrm{max}} = 2\,\sigma'$$

to suppress the impact of the tails.

Strictly speaking, a dijet-event configuration where both jets fluctuate to extremely

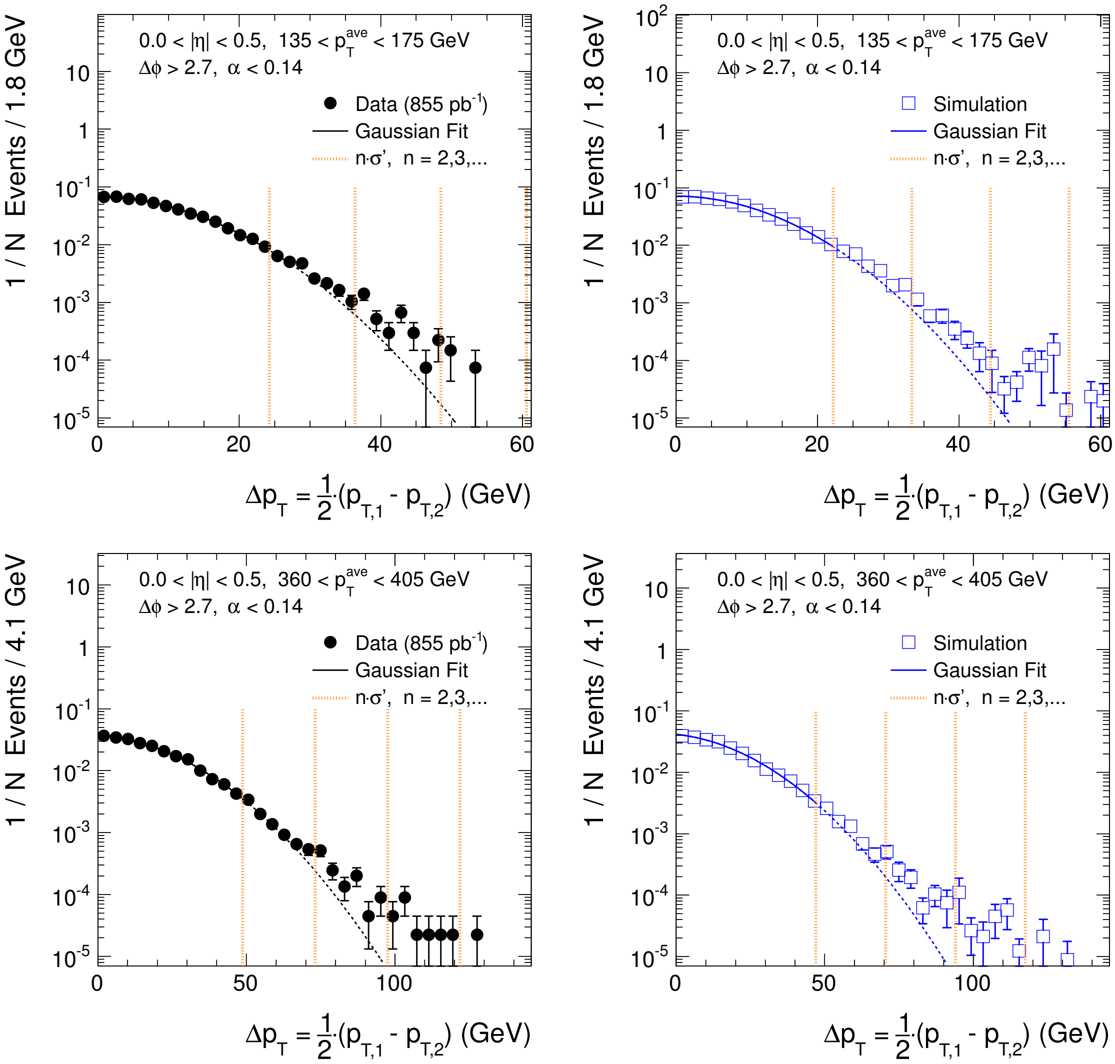


Figure 5.10: Presence of non-Gaussian tails in the Δp_{T} distribution in data (*left*) and simulation (*right*) in a low (*top*) and a medium (*bottom*) $p_{\mathrm{T}}^{\mathrm{ave}}$ interval for $|\eta| < 0.5$. The central part $\Delta p_{\mathrm{T}} < 2\,\sigma'$ of the distribution has been fitted with a Gaussian (*solid line*), which has been extrapolated further (*dashed line*), and $n \cdot \sigma'$ distances (*dotted lines*) are shown. Note that this is not the result of the maximum-likelihood fit but a direct fit to the distribution to guide the eye.

high or extremely low values of p_{T} at the same time can result in small values of Δp_{T} such that the event passes the $\Delta p_{\mathrm{T}}^{\mathrm{max}}$ threshold. However, due to the tiny size of the non-Gaussian jet-p_{T} response tails, which contribute at the percent level or below to the total distribution, this case occurs at a negligible rate compared to events where only one jet fluctuates to extreme values. The latter populate the Δp_{T} distribution's tails and are rejected by the $\Delta p_{\mathrm{T}}^{\mathrm{max}}$ criterion.

The threshold $\Delta p_{\mathrm{T}}^{\mathrm{max}}$ has been chosen proportional to σ', which is a free parameter of the likelihood and is varied accordingly during the maximisation. This, however, biases the result towards smaller values. Hence, an iterative procedure is applied. The threshold is kept fixed and the maximisation is performed, resulting in a first estimate $\hat{\sigma}'$. Then, the threshold is updated with this estimate, and the maximisation is repeated. This procedure is iterated four times, minimising the bias imposed by non-Gaussian tails.

5.2.5 Final Expression for the Dijet Likelihood

Combination of the previously discussed adjustments to the dijet pdf Eq. (5.2) results in

$$g_{\sigma'}\left(p_{\mathrm{T}}^{\mathrm{ave}},\Delta p_{\mathrm{T}}\right)=\frac{1}{\mathcal{N}}\,\mathrm{e}^{-\frac{1}{2}\left(\frac{\Delta p_{\mathrm{T}}}{\sigma'}\right)^2}\int_0^\infty \mathrm{d}p_{\mathrm{T}}^{\mathrm{true}}\; f_{\sigma'}\left(p_{\mathrm{T}}^{\mathrm{true}}|p_{\mathrm{T,min}}^{\mathrm{ave}},p_{\mathrm{T,max}}^{\mathrm{ave}}\right)\,\mathrm{e}^{-\frac{1}{2}\left(\frac{p_{\mathrm{T}}^{\mathrm{ave}}-p_{\mathrm{T}}^{\mathrm{true}}}{\sigma'}\right)^2}$$

with the spectrum

$$\begin{aligned}f_{\sigma'}\left(p_{\mathrm{T}}^{\mathrm{true}}|p_{\mathrm{T,min}}^{\mathrm{ave}},p_{\mathrm{T,max}}^{\mathrm{ave}}\right)&=f\left(p_{\mathrm{T}}^{\mathrm{true}}\right)\int_{p_{\mathrm{T,min}}^{\mathrm{ave}}}^{p_{\mathrm{T,max}}^{\mathrm{ave}}}\mathrm{d}p_{\mathrm{T}}^{\mathrm{ave}}\;\mathrm{e}^{-\frac{1}{2}\left(\frac{p_{\mathrm{T}}^{\mathrm{ave}}-p_{\mathrm{T}}^{\mathrm{true}}}{\sigma'}\right)^2}\\&=f\left(p_{\mathrm{T}}^{\mathrm{true}}\right)\left[\mathrm{erf}\left(\frac{p_{\mathrm{T,max}}^{\mathrm{ave}}-p_{\mathrm{T}}^{\mathrm{true}}}{\sqrt{2}\sigma'}\right)-\mathrm{erf}\left(\frac{p_{\mathrm{T,min}}^{\mathrm{ave}}-p_{\mathrm{T}}^{\mathrm{true}}}{\sqrt{2}\sigma'}\right)\right]\,.\end{aligned}$$

The normalisation factor $\mathcal{N}$ is defined by the condition

$$\int_0^{\Delta p_{\mathrm{T}}^{\mathrm{max}}}\mathrm{d}\Delta p_{\mathrm{T}}\int_{p_{\mathrm{T,min}}^{\mathrm{ave}}}^{p_{\mathrm{T,max}}^{\mathrm{ave}}}\mathrm{d}p_{\mathrm{T}}^{\mathrm{ave}}\;g_{\sigma'}\left(p_{\mathrm{T}}^{\mathrm{ave}},\Delta p_{\mathrm{T}}\right)=1\,.$$

Assuming that the order of integration can be interchanged between $p_{\mathrm{T}}^{\mathrm{ave}}$ and $p_{\mathrm{T}}^{\mathrm{true}}$, $\mathcal{N}$ can be factorised and the dijet pdf can be written as

$$g_{\sigma'}\left(p_{\mathrm{T}}^{\mathrm{ave}},\Delta p_{\mathrm{T}}\right)=\frac{\mathrm{e}^{-\frac{1}{2}\left(\frac{\Delta p_{\mathrm{T}}}{\sigma'}\right)^2}}{\mathcal{N}_{\Delta p_{\mathrm{T}}}}\int_0^\infty \mathrm{d}p_{\mathrm{T}}^{\mathrm{true}}\;\frac{f_{\sigma'}\left(p_{\mathrm{T}}^{\mathrm{true}}|p_{\mathrm{T,min}}^{\mathrm{ave}},p_{\mathrm{T,max}}^{\mathrm{ave}}\right)}{\mathcal{N}_{p_{\mathrm{T}}^{\mathrm{true}}}}\,\frac{\mathrm{e}^{-\frac{1}{2}\left(\frac{p_{\mathrm{T}}^{\mathrm{ave}}-p_{\mathrm{T}}^{\mathrm{true}}}{\sigma'}\right)^2}}{\mathcal{N}_{p_{\mathrm{T}}^{\mathrm{ave}}}}\tag{5.11}$$

with

$$\begin{aligned}\mathcal{N}_{\Delta p_{\mathrm{T}}}&=\sqrt{2\pi}\sigma'\mathrm{erf}\left(\frac{\Delta p_{\mathrm{T}}^{\mathrm{max}}}{\sqrt{2}\sigma'}\right)\\\mathcal{N}_{p_{\mathrm{T}}^{\mathrm{ave}}}&=\sqrt{\frac{\pi}{2}}\sigma'\left[\mathrm{erf}\left(\frac{p_{\mathrm{T,max}}^{\mathrm{ave}}-p_{\mathrm{T}}^{\mathrm{true}}}{\sqrt{2}\sigma'}\right)-\mathrm{erf}\left(\frac{p_{\mathrm{T,min}}^{\mathrm{ave}}-p_{\mathrm{T}}^{\mathrm{true}}}{\sqrt{2}\sigma'}\right)\right]\\\mathcal{N}_{p_{\mathrm{T}}^{\mathrm{true}}}&=\sqrt{\frac{\pi}{2}}\sigma'\int_0^\infty\mathrm{d}p_{\mathrm{T}}^{\mathrm{true}}\;f\left(p_{\mathrm{T}}^{\mathrm{true}}\right)\left[\mathrm{erf}\left(\frac{p_{\mathrm{T,max}}^{\mathrm{ave}}-p_{\mathrm{T}}^{\mathrm{true}}}{\sqrt{2}\sigma'}\right)-\mathrm{erf}\left(\frac{p_{\mathrm{T,min}}^{\mathrm{ave}}-p_{\mathrm{T}}^{\mathrm{true}}}{\sqrt{2}\sigma'}\right)\right]\,.\end{aligned}$$

This is the final form of the dijet pdf as used in this analysis. With it, the final dijet likelihood becomes

$$\mathcal{L}\left(\sigma'\right)=\prod_{k=1}^{N}g_{\sigma',k}\left(p_{\mathrm{T},1},p_{\mathrm{T},2}\right)\,,\tag{5.12}$$

where $g_{\sigma',k}(p_{\mathrm{T},1}, p_{\mathrm{T},2})$ denotes the pdf of the k-th dijet event.

The above dijet pdf contains several integral expressions to which an analytic solution has not been found. Hence, the integration is performed numerically. Since the maximisation of $\mathcal{L}$ with LVMINI requires the repeated evaluation of $\mathcal{L}$ for different σ', the integration has to be fast in order to obtain results on a reasonable time scale. This constraint is lessened by the technical implementation of the analysis, which allows parallel treatment of different dijet events such that the integration can be performed in parallel on several CPUs.

The integral over $p_{\mathrm{T}}^{\mathrm{true}}$ in Eq. (5.11) is computed following *Simpson's 3/8 rule*, an iterative approximation procedure for definite integrals, which is based on cubic interpolation of the integrand [255]. The spectral factor $f_{\sigma'}(p_{\mathrm{T}}^{\mathrm{true}}|p_{\mathrm{T,min}}^{\mathrm{ave}}, p_{\mathrm{T,max}}^{\mathrm{ave}})/\mathcal{N}_{p_{\mathrm{T}}^{\mathrm{true}}}$ has to be determined numerically as well. However, updating this factor after each iteration of LVMINI's maximisation procedure is not feasible even with the available CPU capacity. Hence, the spectral factor is computed before the maximisation with a fixed starting value of σ' for different $p_{\mathrm{T}}^{\mathrm{true}}$, and the values are stored in a look-up table which is accessed during the maximisation. After convergence, the factor is reevaluated, and the maximisation procedure is repeated with the updated look-up table. In total, four iterations are performed. This procedure is possible because the fit result is relatively insensitive to the exact shape of the spectrum as will be shown later in Section 5.5. Together with the look-up table, also the value of $\Delta p_{\mathrm{T}}^{\mathrm{max}}$ is updated. With this approach, the time for one fit takes up to 20 minutes on a current eight-core computer depending on the chosen interval.

For readability, the notational distinction between the parameter σ' and the value $\hat{\sigma}'$ that maximises the likelihood will be omitted later on, and σ' will be used in both cases.

5.2.6 Validation of the Final Dijet Likelihood

Various tests of the presented maximum-likelihood fit have been performed. In Section 5.2.6.1, the validity of the likelihood definition Eq. (5.12) and the performance of the implemented minimisation procedure are verified using a very basic simulation (*toy simulation*) with an ideal dijet-event topology. In Section 5.2.6.2, the fit is applied to PYTHIA-simulated QCD events taking into account realistic detector conditions and to data, and the obtained pdfs for the spectrum and for Δp_{T} are investigated. The complete method, which includes corrections for additional jet activity as explained below, will be validated later in Section 5.4.

5.2.6.1 Validation with a Toy Simulation

A sample of toy dijet-events is generated where each event is characterised by the true transverse momentum $p_{\mathrm{T}}^{\mathrm{true}}$, which follows a simple exponential spectrum

$$f(p_{\mathrm{T}}^{\mathrm{true}}) \propto \exp(-p_{\mathrm{T}}^{\mathrm{true}}/\tau) , \qquad \tau = 100\,\mathrm{GeV} . \tag{5.13}$$

For each event, two independent measurements $\{p_{\mathrm{T},i}\,|\,i = 1, 2\}$ of $p_{\mathrm{T}}^{\mathrm{true}}$ are simulated by weighting $p_{\mathrm{T}}^{\mathrm{true}}$ with a random number according to a Gaussian response pdf[8] with mean value 1 and standard deviation $\sigma(p_{\mathrm{T}}^{\mathrm{true}})/p_{\mathrm{T}}^{\mathrm{true}}$. Then, the maximum-likelihood fit is applied

---

[8] Hence, the limitation $\Delta p_{\mathrm{T}} < 2\sigma'$ to protect from the impact of non-Gaussian tails has no effect.

to determine the p_T resolution σ. Two models with different parametrisations of σ and different event selection criteria are investigated.

Model 1: Constant σ and $p_\mathrm{T}^\mathrm{true}$ Selection As a most fundamental test, the resolution is parametrised as

$$\frac{\sigma(p_\mathrm{T}^\mathrm{true})}{p_\mathrm{T}^\mathrm{true}} = \frac{20\,\mathrm{GeV}}{p_\mathrm{T}^\mathrm{true}} \,,$$

at generation, i. e. $\sigma = 20\,\mathrm{GeV}$ is p_T independent as assumed in the dijet pdf Eq. (5.11). Four-hundred statistically independent samples S_k are generated, each containing 10000 events with $250 < p_\mathrm{T}^\mathrm{true} < 300\,\mathrm{GeV}$. The generated $p_\mathrm{T}^\mathrm{true}$ and simulated Δp_T distributions of the events from one specific sample S_0 are shown in Fig. 5.11 for illustration purposes. Migration effects do not occur since the selection criterion is applied to generator-level quantities. Hence, the spectrum Eq. (5.11) is given directly by the exponential Eq. (5.13). An initial value of $\sigma = 50\,\mathrm{GeV}$ is chosen for the minimisation procedure.

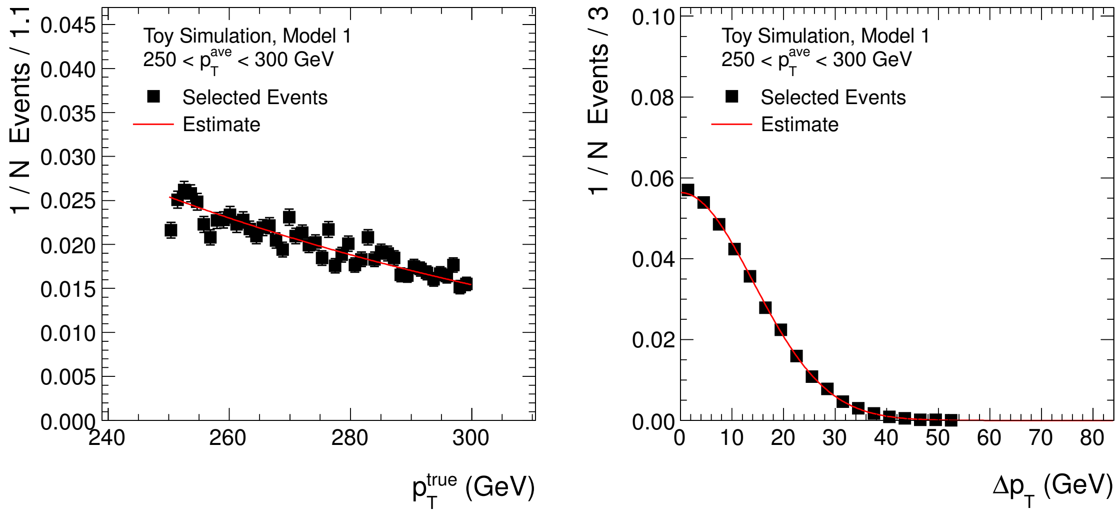


Figure 5.11: $p_\mathrm{T}^\mathrm{true}$ (*left*) and $\Delta p_\mathrm{T} = \frac{1}{2}(p_{\mathrm{T},1} - p_{\mathrm{T},2})$ (*right*) for the sample S_0 of model 1 of the toy simulation. The distributions (*squared markers*) are compared to the assumed pdfs evaluated with $\sigma = \hat{\sigma}_0 = \sqrt{2} \cdot \hat{\sigma}'_0$ (*solid lines*).

In case of S_0, the fit results in $\hat{\sigma}_0 = \sqrt{2} \cdot \hat{\sigma}'_0 = 20.06 \pm 0.13\,\mathrm{GeV}$. The negative log-likelihood function $\mathcal{F}(\sigma)$ in the vicinity of $\hat{\sigma}_0$ is depicted in Fig. 5.12. $\mathcal{F}$ has parabolic shape to good approximation with a well defined minimum at $\hat{\sigma}_0$. The parameter error $\delta\hat{\sigma}_0 = 0.13\,\mathrm{GeV}$ corresponds to the distance from the minimum to the point at which the value of $\mathcal{F}$ has increased by 1, i. e. the interval which covers 68% of the likelihood distribution $\mathcal{L}$. The Δp_T distribution Fig. 5.11 (*right*) is well described by a Gaussian pdf with standard deviation $\hat{\sigma}'_0$ as assumed in Eq. (5.11).

The distributions of the fit results $\hat{\sigma}_k$ of all 400 samples and their deviation $(\hat{\sigma}_k - \langle\sigma\rangle)/\delta\hat{\sigma}_k$ from the expectation $\langle\sigma\rangle = 20\,\mathrm{GeV}$ relative to the fit error $\delta\hat{\sigma}_k$ ('pull' distribution) are shown in Fig. 5.13. Proper statistical behaviour of the likelihood is concluded from the pull distribution's width of 0.94 ± 0.03, which is determined by a Gaussian fit. The likeli-

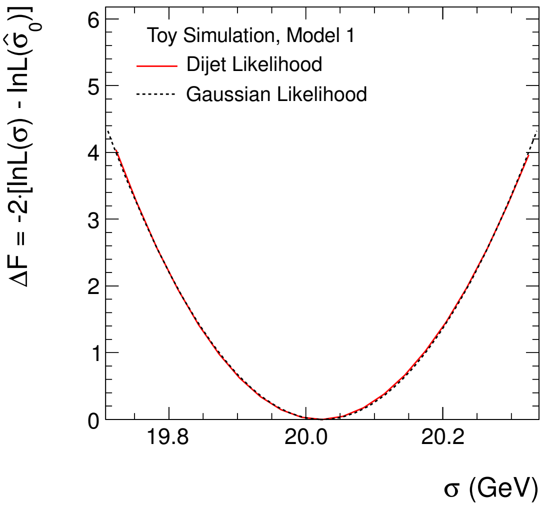


Figure 5.12: Difference $\Delta\mathcal{F} = -2 \cdot [\ln \mathcal{L}(\sigma) - \ln \mathcal{L}(\hat{\sigma}_0)]$ (*solid line*) of the negative log-likelihood function to its minimum at $\hat{\sigma}_0$ as a function of the parameter value σ for the sample S_0 of model 1 of the toy simulation. It is compared to the case of a Gaussian likelihood (for which $\mathcal{F}$ is parabolic) with the same minimum and error (*dashed line*).

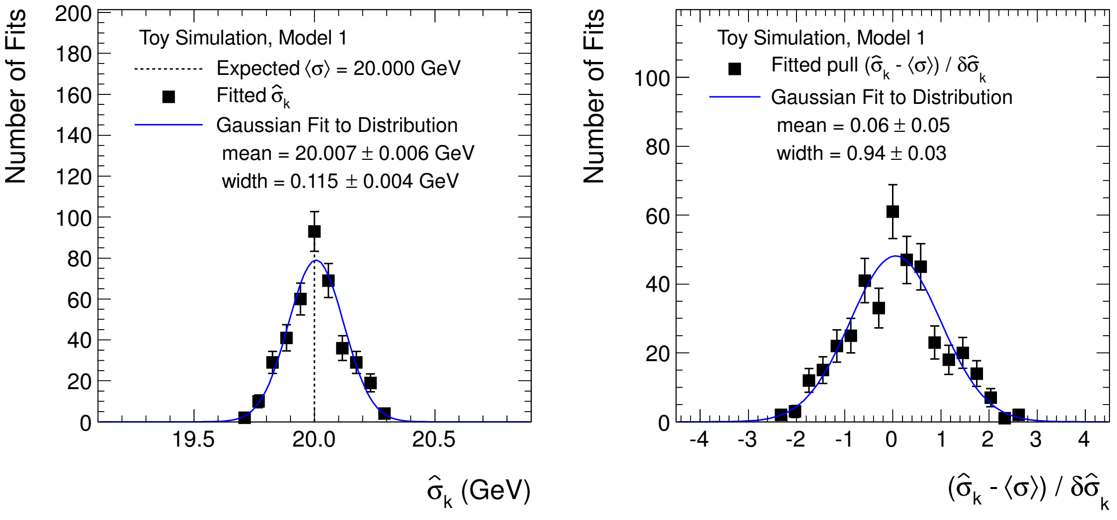


Figure 5.13: Results $\hat{\sigma}_k$ (*left*) of the dijet-likelihood fit and their deviation $(\hat{\sigma}_k - \langle\sigma\rangle)/\delta\hat{\sigma}_k$ (*right*) from the expectation $\langle\sigma\rangle = 20\,\text{GeV}$ relative to the fit error $\delta\hat{\sigma}_k$ ('pull') for all 400 toy samples of model 1. Mean value and width of the distributions (*markers*) are determined by Gaussian fits (*solid lines*).

hood is virtually unbiased: the $\hat{\sigma}_k$ distribution's mean of $20.007 \pm 0.006\,\text{GeV}$ demonstrates that any bias is less than 0.04%.

Model 2: p_T-Dependent σ and $p_\mathrm{T}^\mathrm{ave}$ Selection For a more realistic test, the resolution is parametrised as

$$\frac{\sigma(p_\mathrm{T}^\mathrm{true})}{p_\mathrm{T}^\mathrm{true}} = \frac{4}{p_\mathrm{T}^\mathrm{true}/\mathrm{GeV}} \oplus \frac{1.2}{\sqrt{p_\mathrm{T}^\mathrm{true}/\mathrm{GeV}}} \oplus 0.05\ .$$

Thus, σ depends on $p_\mathrm{T}^\mathrm{true}$ in analogy to the transverse-momentum dependence of the jet resolution at CMS. The event selection is performed using measured quantities such that migration effects have to be considered via Eq. (5.8). Again, 400 samples S_k are generated[9], and 10000 events are selected from each by requiring $250 < p_\mathrm{T}^\mathrm{ave} < 300\,\mathrm{GeV}$.

The $p_\mathrm{T}^\mathrm{true}$ distribution obtained from one specific sample S_0 is shown in Fig. 5.14 (*left*). With its mean $\bar{p}_\mathrm{T}^\mathrm{true}(S_0) = 270.53\,\mathrm{GeV}$, the expectation value $\langle\sigma\rangle_0$ for this sample is approximately

$$\langle\sigma\rangle_0 \approx \sigma(\bar{p}_\mathrm{T}^\mathrm{true}(S_0)) = 24.26\,\mathrm{GeV}\ .$$

The fit result of $\hat{\sigma}_0 = \sqrt{2}\cdot\sigma'_0 = 24.11 \pm 0.17\,\mathrm{GeV}$ agrees well, and $\mathcal{F}(\sigma)$ is again parabolic to good approximation, cf. Fig. 5.15.

The bulk of the $p_\mathrm{T}^\mathrm{true}$ distribution Fig. 5.14 (*left*) is well described by the assumed pdf $f_{\hat{\sigma}'_0}(p_\mathrm{T}^\mathrm{true}|p_\mathrm{T,min}^\mathrm{ave}, p_\mathrm{T,max}^\mathrm{ave})$ Eq. (5.8). Thus, the migration effects have been correctly incorporated into the likelihood. Very small deviations are visible for $p_\mathrm{T}^\mathrm{true} < 220\,\mathrm{GeV}$ and $p_\mathrm{T}^\mathrm{true} > 340\,\mathrm{GeV}$. This is due to the fact that a p_T-independent σ Eq. (5.7) has been assumed within in the $p_\mathrm{T}^\mathrm{ave}$ interval, while in fact it depends on $p_\mathrm{T}^\mathrm{true}$. At small $p_\mathrm{T}^\mathrm{true}$, for example, $\sigma(p_\mathrm{T}^\mathrm{true})$ is smaller than $\langle\sigma\rangle_0$. Hence, less events migrate into the $p_\mathrm{T}^\mathrm{ave}$ interval, which is why the $p_\mathrm{T}^\mathrm{true}$ distribution is overestimated. The Δp_T distribution Fig. 5.14 (*right*) for S_0 again agrees well with its Gaussian pdf.

The expected parameter value $\langle\sigma\rangle$ of all 400 samples is determined as the average of the $\langle\sigma\rangle_k = \sigma((\bar{p}_\mathrm{T}^\mathrm{true}S_k))$ in each sample,

$$\langle\sigma\rangle = \frac{1}{400}\sum_{k=1}^{400}\langle\sigma\rangle_k = 24.309\,\mathrm{GeV}\,. \qquad (5.14)$$

With it, the $\hat{\sigma}_k$ and pull distributions are obtained. They are shown in Fig. 5.16. Again, there is good behaviour of the likelihood. Potential biases are below 0.1%.

5.2.6.2 Validation with Data and with the Full Simulation

The maximum-likelihood fit is applied to the selected data and Pythia-simulated events described in Section 5.1.

In Fig. 5.17 and in Figs. A.6 to A.15 in Appendix A.4, the Δp_T distributions of the events are shown. The central part $\Delta p_\mathrm{T} < 2\cdot\hat{\sigma}'$, which contains the events that contribute to the likelihood, is indicated. It is well described both in data and in the simulation by a Gaussian with standard deviation $\hat{\sigma}'$ as assumed in the dijet pdf Eq. (5.11). Some disagreement is observed merely in the lowest interval $45 < p_\mathrm{T}^\mathrm{ave} < 75\,\mathrm{GeV}$, in particular for $|\eta| > 1.1$, where $\hat{\sigma}'$ is too large. Here, the variation of the jet-p_T resolution over the interval range is relatively large such that the approximation Eq. (5.7) is less adequate.

---

[9] In this case, values of $p_\mathrm{T}^\mathrm{true}$ between 200 and 400 GeV are generated to properly simulate the impact of migration.

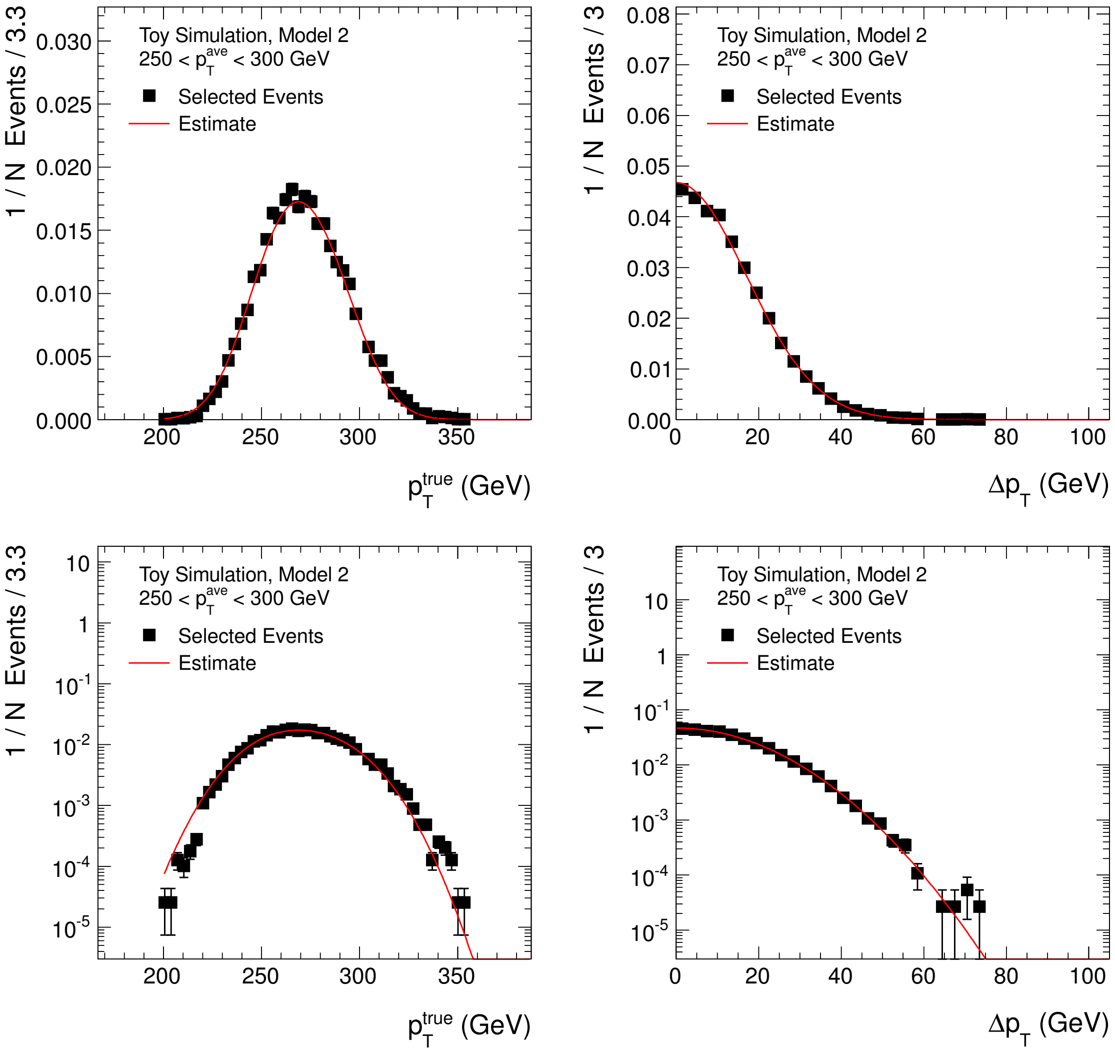


Figure 5.14: $p_{\mathrm{T}}^{\mathrm{true}}$ (*left*) and $\Delta p_{\mathrm{T}} = \frac{1}{2}(p_{\mathrm{T},1} - p_{\mathrm{T},2})$ (*right*) for the sample S_0 of model 2 of the toy simulation on a linear (*top*) and logarithmic (*bottom*) scale. The distributions (*squared markers*) are compared to the assumed pdfs evaluated with $\sigma = \hat{\sigma}_0 = \sqrt{2} \cdot \hat{\sigma}'_0$ (*solid lines*).

Furthermore, the number of events is small such that the statistical precision is limited.

In order to validate the incorporation of migration effects, the $p_{\mathrm{T}}^{\mathrm{gen}}$ distributions of the simulated events are investigated Fig. 5.18 and in Figs. A.1 to A.5 in Appendix A.3. They are well described by the assumed pdf $f_{\hat{\sigma}'}(p_{\mathrm{T}}^{\mathrm{true}} | p_{\mathrm{T,min}}^{\mathrm{ave}}, p_{\mathrm{T,max}}^{\mathrm{ave}})$ Eq. (5.8). In accordance with the above observation, some disagreement occurs in the $45 < p_{\mathrm{T}}^{\mathrm{ave}} < 75\,\mathrm{GeV}$ interval. Further, small deviations arise far away from the interval boundaries as visible in Fig. 5.18 (*bottom*). Here, non-Gaussian components of the jet response distribution, which are not considered in Eq. (5.8), promote the migration. This effect is more pronounced at the high-p_{T} side of the interval since non-Gaussian tails occur predominantly at low response, cf. Fig. 4.9. To a smaller extent, also effects of the assumption Eq. (5.7)

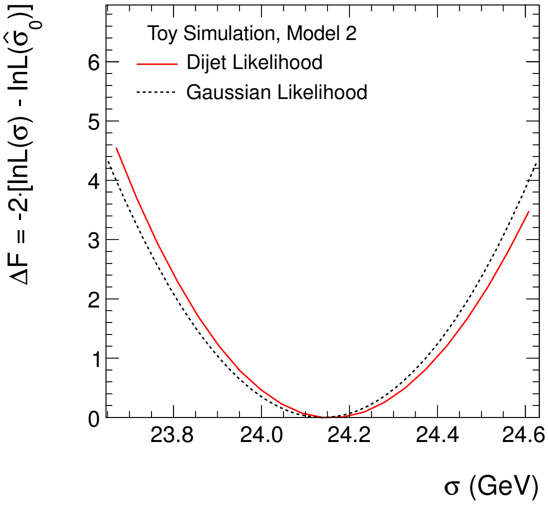


Figure 5.15: Difference $\Delta\mathcal{F} = -2 \cdot [\ln \mathcal{L}(\sigma) - \ln \mathcal{L}(\hat{\sigma}_0)]$ (*solid line*) of the negative log-likelihood function to its minimum at $\hat{\sigma}_0$ as a function of the parameter value σ for the sample S_0 of model 2 of the toy simulation. It is compared to the case of a Gaussian likelihood (for which $\mathcal{F}$ is parabolic) with the same minimum and error (*dashed line*).

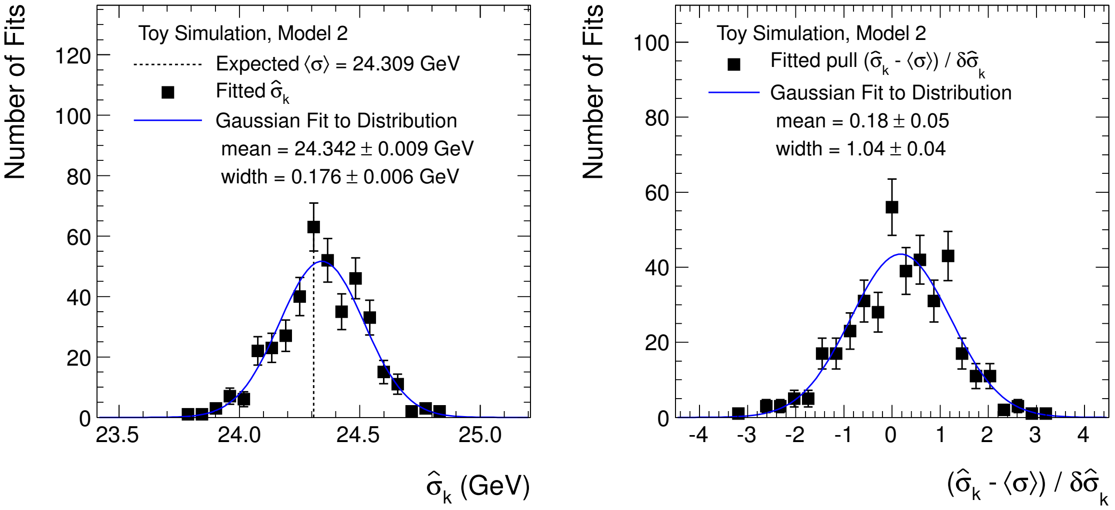


Figure 5.16: Results $\hat{\sigma}_k$ (*left*) of the dijet-likelihood fit and their deviation $(\hat{\sigma}_k - \langle\sigma\rangle)/\delta\hat{\sigma}_k$ (*right*) from the expectation $\langle\sigma\rangle = 24.309\,\text{GeV}$ Eq. (5.14) relative to the fit error $\delta\hat{\sigma}_k$ ('pull') for all 400 toy samples of model 2. Mean value and width of the distributions (*markers*) are determined by Gaussian fits (*solid lines*).

contribute, cf. the above discussion of Fig. 5.14 (*left*).

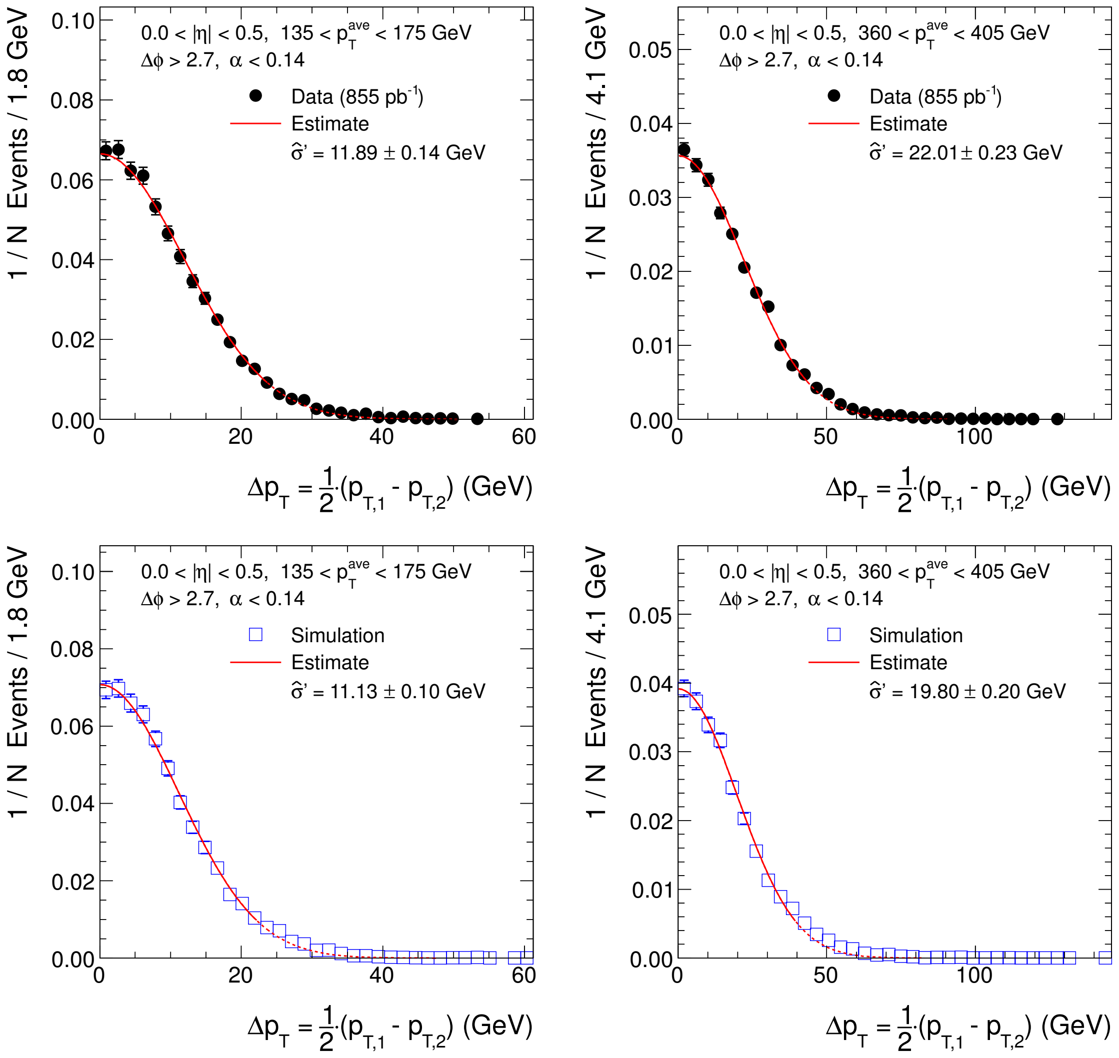


Figure 5.17: Δp_{T} distributions (*markers*) in data (*top*) and for simulated events (*bottom*) in a low (*left*) and a medium (*right*) $p_{\mathrm{T}}^{\mathrm{ave}}$ interval for $|\eta| < 0.5$. The central part is well described by a Gaussian pdf with standard deviation $\hat{\sigma}'$ (*lines*) as assumed in Eq. (5.11). Only events with $\Delta p_{\mathrm{T}} < 2 \cdot \hat{\sigma}'$ (*solid part of the lines*) have been considered for the fit, and all distributions are normalised to this region.

5.3 Corrections to the Resolution

The dijet pdf Eq. (5.11) is defined under the assumption of exactly two jets in the event that are balanced in transverse momentum at particle level. Different effects such as the presence of soft radiation and additional jets which originate in the same hard interaction as well as the statistical nature of the hadronisation process cause a transverse-momentum imbalance between the two particle-level jets and hence broaden the measured resolution, cf. Section 4.6.2. These biases are not directly considered in the likelihood but compensated

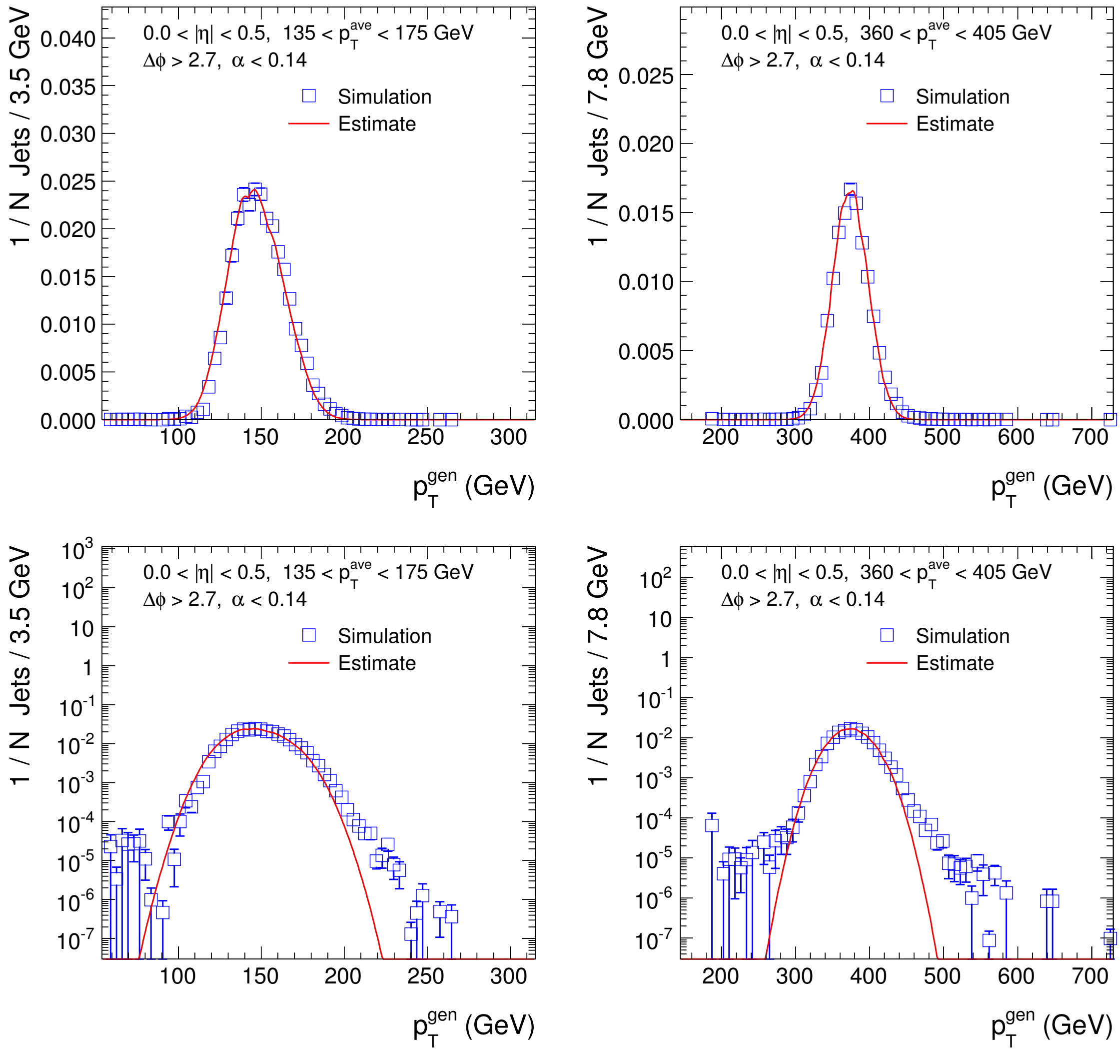


Figure 5.18: $p_{\mathrm{T}}^{\mathrm{gen}}$ distributions (*open squares*) of the simulated dijet events in a low (*left*) and a medium (*right*) $p_{\mathrm{T}}^{\mathrm{ave}}$ interval for $|\eta| < 0.5$ and $\alpha < 0.14$ on a linear (*top*) and logarithmic (*bottom*) scale. All distributions are normalised such that their integral is 1. Selection effects discussed in the text result in migration at the interval boundaries. The dominant part of the distributions is well described by the spectrum $f_{\hat{\sigma}'}(p_{\mathrm{T}}^{\mathrm{true}}|p_{\mathrm{T,min}}^{\mathrm{ave}}, p_{\mathrm{T,max}}^{\mathrm{ave}})$ Eq. (5.8) (*solid line*).

for by subsequent corrections of the result as explained in Sections 5.3.1 and 5.3.2. Similar correction techniques have been used in other analyses [172, 230, 233]. In Section 5.3.3, the resulting, final expression for the resolution as used in the following measurement is summarised.

5.3.1 Correction for Additional Jet Activity

The emission of partons that lead to additional jets induce a transverse-momentum imbalance of the dijet system, and as a consequence, the measured resolution increases. This has been investigated in Section 4.6.2. The measured resolution can be expressed as a quadratic sum of the intrinsic jet-p_{T} resolution and an imbalance term σ_{imbal}, cf. Eq. (4.13), where σ_{imbal} depends on α^{imbal} Eq. (4.12), the fractional p_{T} of all additional jets along the dijet axis.

However, inaccurate energy and position reconstruction of low-p_{T} jets at detector level prevent a precise measurement of α^{imbal} including all further jets and not just the third one. In this analysis, the p_{T} imbalance is therefore measured by α Eq. (5.1), the fractional p_{T} of the third jet relative to the dijet p_{T}-scale. The impact from the fourth and any further jet in the event can be neglected because they typically have successively much smaller p_{T} than the third jet due to the steep decrease of the jet-production cross-section Fig. 4.4 with p_{T}.

The suitability of α is investigated with simulated events using generator-truth information. As demonstrated in Fig. 5.19 (*top*), $\alpha^{\mathrm{gen}} = p_{\mathrm{T},3}^{\mathrm{gen}}/p_{\mathrm{T}}^{\mathrm{gen,ave}}$ correlates strongly with α^{imbal}. α^{gen} is in general larger than α^{imbal} because the case can occur that $p_{\mathrm{T},3}^{\mathrm{gen}}$ has a large magnitude but a small component $p_{||,3}^{\mathrm{gen}}$ along the dijet axis or that its $p_{||,3}^{\mathrm{gen}}$ is compensated by further jets.

The relation between the generator-level α^{gen} and the detector-level α is almost linear as shown in Fig. 5.19 (*bottom*). Merely at low $p_{\mathrm{T}}^{\mathrm{ave}}$ and α^{gen} below 0.04, α is measured systematically too large. In those events, the third-jet p_{T} is below 10 GeV, which lies close to the reconstruction threshold and is affected by noise, UE activity, and large uncertainties of the jet-energy-scale corrections. Furthermore, a pile-up induced jet, which has low p_{T} in general, can be mistaken as the third jet, if it has larger p_{T} than the third jet from the primary interaction. In any case, the measured α is a valid upper bound for the relative p_{T} imbalance α^{imbal}.

In the following, the imbalance contribution to the measured resolution is compensated by an extrapolation procedure. Several sets of dijet events are selected per $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ interval, each with a different threshold α_{max} on α,

$$\alpha < \alpha_{\mathrm{max}} \,. \tag{5.15}$$

In each case, the dijet likelihood Eq. (5.12) is maximised with respect to $\sigma = \sqrt{2} \cdot \sigma'$ while neglecting the presence of additional jets[10]. Assuming that σ_{imbal} is much smaller than the relative intrinsic jet-p_{T} resolution and assuming $\sigma_{\mathrm{imbal}} \propto \alpha$, the measured σ can be shown to depend linearly on α_{max}, independently of the exact shape of the α spectrum Fig. 5.4 (*top*). This is in good agreement to the result of the measurements shown in Fig. 5.20 and in Figs. A.16 to A.20 in Appendix A.5.

Thus, in order to extrapolate the measurement to the case of an ideal dijet event, the σ are fitted with a linear function $\sigma(\alpha_{\mathrm{max}})$, and the y-axis intercept $\sigma(\alpha_{\mathrm{max}} \to 0)$ is used as result for the jet-p_{T} resolution in the further analysis. Hence, the measured relative

---

[10] Recall that the distinction between the parameter σ and the value $\hat{\sigma}$ that maximises the likelihood was dropped for the ease of notation.

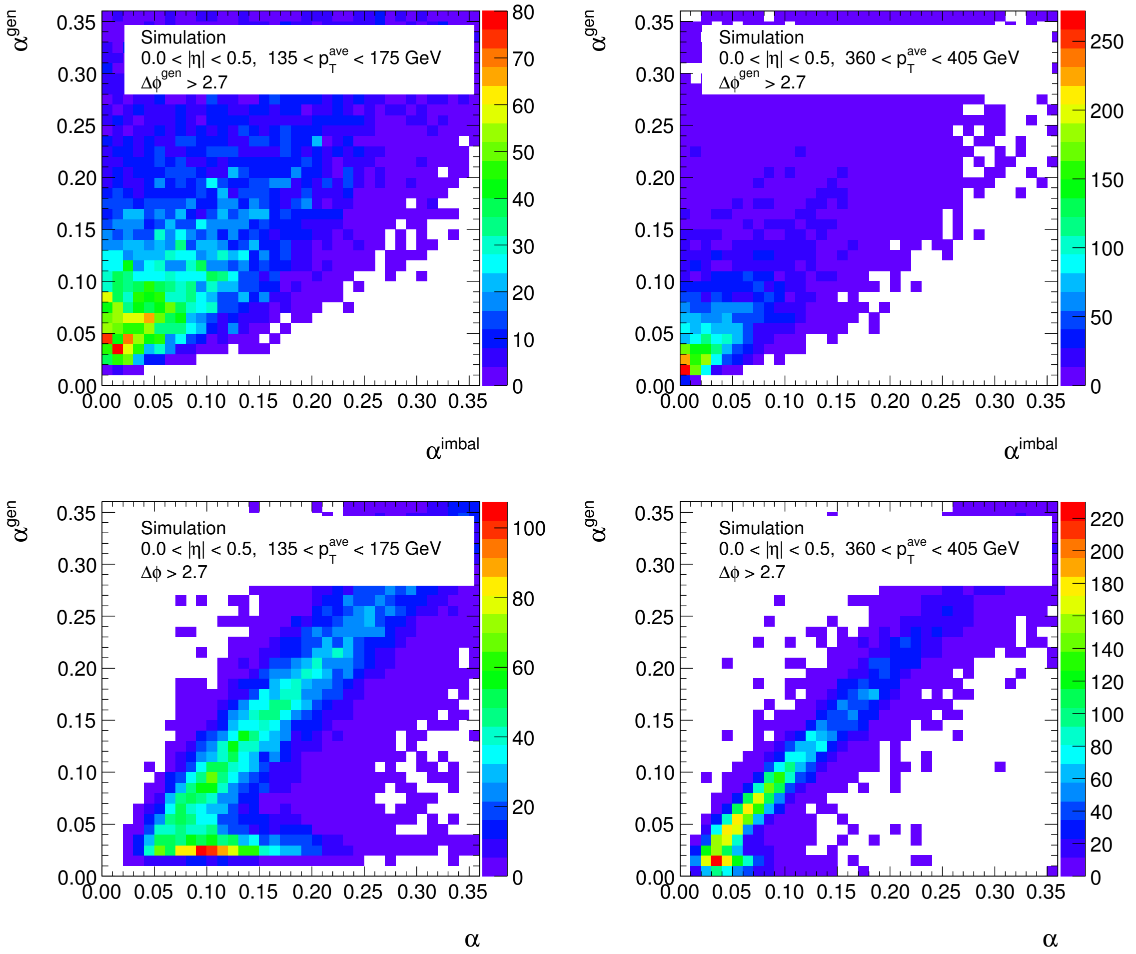


Figure 5.19: Correlation between $\alpha^{\text{gen}} = p_{\text{T},3}^{\text{gen}}/p_{\text{T}}^{\text{gen,ave}}$ and the relative p_{T} imbalance α^{imbal} Eq. (4.12) (*top*), and correlation between α^{gen} and the corresponding detector-level α (*bottom*) for PYTHIA-simulated dijet events in a low (*left*) and medium (*right*) $p_{\text{T}}^{\text{ave}}$ interval for $|\eta| < 0.5$. To obtain the α^{gen} versed α correlation, the generator-level jet closest in ΔR has been matched to each detector-level jet. To select events with a dijet-like topology, only the $\Delta\phi > 2.7$ criterion has been applied. The histograms are normalised to unit volume.

resolution Eq. (5.10) becomes

$$\text{relative resolution} = \frac{\sigma(\alpha_{\max} \to 0)}{\langle p_{\text{T}}^{\text{true}} \rangle} \,, \tag{5.16}$$

where $\langle p_{\text{T}}^{\text{true}} \rangle$ is computed from Eq. (5.9) using $\sigma(\alpha_{\max} \to 0)$.

In the low $p_{\text{T}}^{\text{ave}}$ intervals, the measured σ in data is larger than expected, cf. Figs. A.16 to A.20 in Appendix A.5. This has already been observed and discussed above in Section 5.2.6.2 for the example of $\alpha < 0.14$. The effect is more pronounced for smaller $\alpha_{\max}$, where the correlation between α and α^{gen} is degraded as shown in Fig. 5.19 (*bottom*).

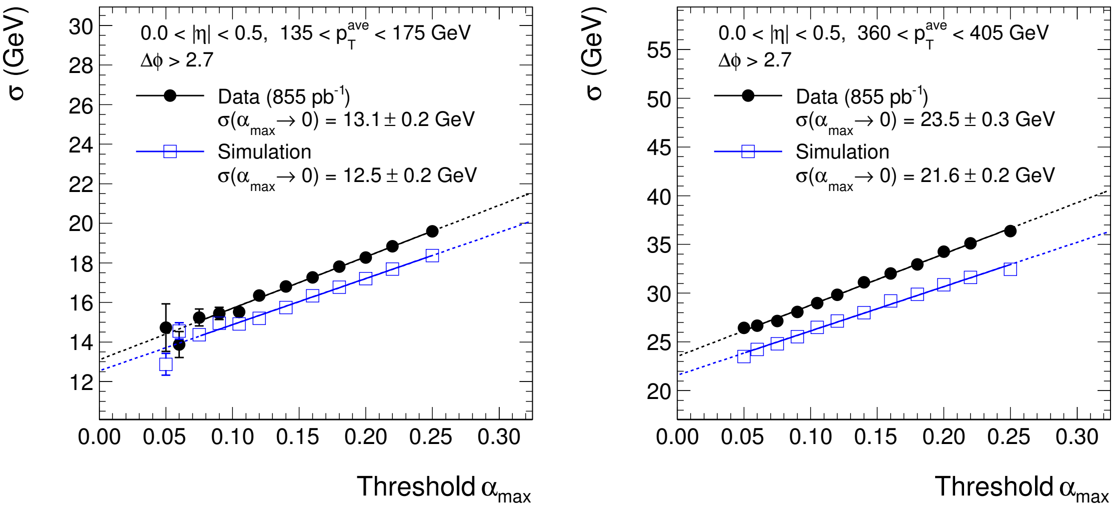


Figure 5.20: Dependence of the fitted σ in data (*solid circles*) and simulation (*open squares*) on the threshold $\alpha_{\max}$ on the fractional third-jet transverse momentum $\alpha = p_{\mathrm{T},3}/p_{\mathrm{T}}^{\mathrm{ave}}$ in a low (*left*) and medium $p_{\mathrm{T}}^{\mathrm{ave}}$ interval (*right*) for $|\eta| < 0.5$. σ is linearly extrapolated (*lines*) to the case of an ideal dijet event $\alpha = 0$. Only points corresponding to an absolute $p_{\mathrm{T},3} > 10\,\mathrm{GeV}$ are considered for the extrapolation fit (*solid part of the lines*).

Therefore, in order to avoid biasing the extrapolation result, only those σ for which $\alpha_{\max}$ corresponds to an absolute $p_{\mathrm{T},3}$ of more than 10 GeV are considered in the linear fit. The uncertainty on the extrapolated value is increased, however, due to the longer lever arm. To ensure good behaviour of the linear fit for $|\eta| < 1.1$, the threshold has to be lowered to $p_{\mathrm{T},3} > 6\,\mathrm{GeV}$ in the intervals with $45 < p_{\mathrm{T}}^{\mathrm{ave}} < 75\,\mathrm{GeV}$. In that case, negative slopes are gained in the three intervals with $|\eta| > 1.1$ and $45 < p_{\mathrm{T}}^{\mathrm{ave}} < 75\,\mathrm{GeV}$, though. The extrapolated values in these three intervals are omitted in the following analysis.

Due to the selection requirement Eq. (5.15), the measured values of σ in one $p_{\mathrm{T}}^{\mathrm{ave}}$ interval for different thresholds $\alpha_{\max}$ are correlated, since all events in a set with a particular value of $\alpha_{\max}$ are also included in the sets with larger $\alpha_{\max}$. A correct treatment of this correlation when propagating the statistical uncertainties on σ to the final result is challenging. Therefore, as in similar analyses [172], a robust, simplified approach is chosen, which neglects the correlations though[11]. In the following, the uncertainty on the value of the linear fit function $\sigma(\alpha_{\max})$ at $\alpha_{\max} = 0$ defines the *extrapolation uncertainty*

$$\delta\sigma_{\mathrm{ex}} = \text{uncertainty on } \sigma(\alpha_{\max} = 0) \tag{5.17}$$

and is taken as uncertainty of the final resolution. The size of $\delta\sigma_{\mathrm{ex}}$ depends on the statistical uncertainty on the measurement, but, due to the selection in inclusive intervals

---

[11] Measuring σ in exclusive intervals ('bins') of α instead in order to avoid the correlation turns out to be impractical because with increasing α, the assumption of a Gaussian Δp_{T} Eq. (5.6) becomes less valid, which biases the measurement at large α. However, in case of the chosen approach with α thresholds, this effect can be neglected since each of the inclusive intervals $0 - \alpha_{\max}$ is dominated by events with small α due to the shape of the α spectrum Fig. 5.4 (*top*).

from $\alpha = 0$ to $\alpha_{\max}$, the statistical uncertainty is expected to be underestimated in the sense that $\delta\sigma_{\mathrm{ex}}$ does not fully cover the fluctuations of the extrapolated value. This will be compensated for by assigning a rather conservative systematic uncertainty on the extrapolation procedure as explained in Section 5.5. In order to obtain an estimate for the correct order of magnitude of the measurement's statistical uncertainty, one might inspect the displayed error bars in Fig. 5.20. They show the statistical uncertainty of σ at a given $\alpha_{\max}$ as returned by the unbinned maximum-likelihood fit before the extrapolation.

5.3.2 Correction for Particle-Level Imbalance

The statistical nature of the parton-showering and the subsequent hadronisation process in connection with the jet algorithm lead to differences between the momenta of the parton and the particle-level jet because some particles might be too soft to be considered in the jet clustering or particles unrelated to the original parton might overlap with the jet, cf. Section 4.4.2. This effect, the *particle-level imbalance* (PLI), is a further source for transverse-momentum imbalance in the dijet system at particle level.

The PLI is determined from the generator-level asymmetry

$$\mathcal{A}^{\mathrm{gen}} = \frac{p_{\mathrm{T},1}^{\mathrm{gen}} - p_{\mathrm{T},2}^{\mathrm{gen}}}{p_{\mathrm{T},1}^{\mathrm{gen}} + p_{\mathrm{T},2}^{\mathrm{gen}}} ,$$

in simulated QCD-multijet events, where the $\{p_{\mathrm{T},i}^{\mathrm{gen}} \,|\, i = 1, 2\}$ refer to the randomly ordered $p_{\mathrm{T}}^{\mathrm{gen}}$ of the leading two generator-level jets. Example distributions after a dijet selection with two different thresholds $\alpha_{\max}^{\mathrm{gen}}$ on α^{gen} are shown in Fig. 5.21 for one $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ interval. Due to additional parton radiation, $p_{\mathrm{T},1}^{\mathrm{gen}}$ is on average not equal to $p_{\mathrm{T},2}^{\mathrm{gen}}$, which leads to a broadening of the distribution for larger $\alpha_{\max}^{\mathrm{gen}}$ (and also results in the dip at $\mathcal{A}^{\mathrm{gen}} = 0$ in Fig. 5.21 (*right*)).

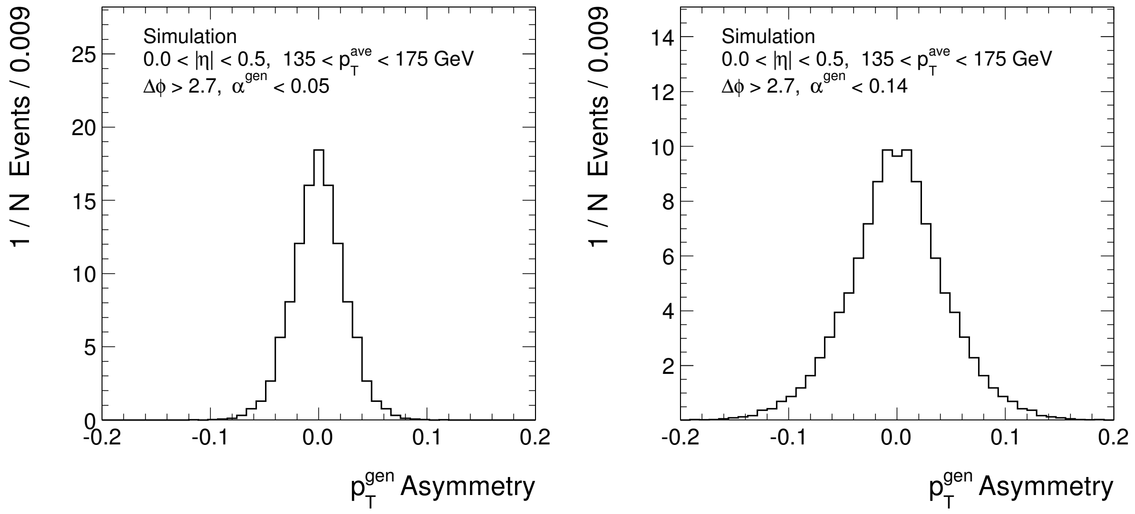


Figure 5.21: Generator-level asymmetry $\mathcal{A}^{\mathrm{gen}}$ for $\alpha^{\mathrm{gen}} < 0.05$ (*left*) and $\alpha^{\mathrm{gen}} < 0.14$ (*right*) in one $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ interval.

The standard deviation $\sigma_{\mathcal{A}^{\text{gen}}}$ of $\mathcal{A}^{\text{gen}}$ is shown in in Fig. 5.22 (*left*) for various thresholds $\alpha^{\text{gen}}_{\text{max}}$. $\sigma_{\mathcal{A}^{\text{gen}}}$ is multiplied by $\sqrt{2}$ to account for the fact that the $p^{\text{gen}}_{\text{T}}$ of both jets fluctuates with respect to the parton p_{T}. As above, a linear extrapolation to $\alpha^{\text{gen}}_{\text{max}} = 0$ is performed. The result,

$$\sigma_{\text{PLI}} = \sqrt{2} \cdot \sigma_{\mathcal{A}^{\text{gen}}}(\alpha^{\text{gen}}_{\text{max}} \to 0) ,$$

is plotted versus $p^{\text{gen}}_{\text{T}}$, as shown in Fig. 5.22 (*right*), and fitted with Eq. (4.7). The fitted parameter values are listed in Table 5.3.

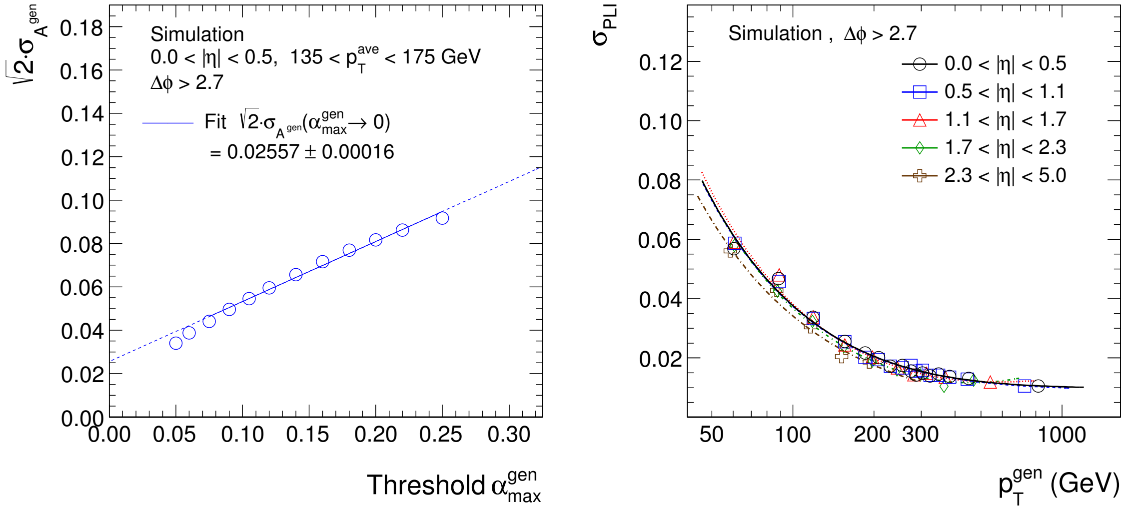


Figure 5.22: Standard deviation $\sigma_{\mathcal{A}^{\text{gen}}}$, multiplied by $\sqrt{2}$, of the generator-level dijet asymmetry (*circles*) for different thresholds on the fractional transverse momentum α^{gen} of the third jet in one $|\eta^{\text{gen}}| \times p^{\text{gen,ave}}_{\text{T}}$ interval (*left*). It is extrapolated to $\alpha^{\text{gen}} = 0$ with a linear fit (*solid line*). The extrapolated values σ_{PLI} are fitted with Eq. (4.7) as a function of $p^{\text{gen}}_{\text{T}}$ in different $|\eta|$ intervals (*right*).

Table 5.3: Fitted parameter values of the particle-level imbalance $\sigma_{\text{PLI}}(p^{\text{gen}}_{\text{T}})$ that is parametrised with Eq. (4.7). (The parameter C has been set to zero to avoid redundancies.) Uncertainties denote the statistical uncertainties returned by the fit.

$\lvert\eta\rvert$	N (GeV)	S (GeV$^{(1-m)/2}$)	m
0.0 – 0.5	3.641 ± 0.023	0.0093 ± 0.0009	1.010 ± 0.028
0.5 – 1.1	3.623 ± 0.023	0.0099 ± 0.0011	0.983 ± 0.033
1.1 – 1.7	3.787 ± 0.028	0.0007 ± 0.0003	1.856 ± 0.133
1.7 – 2.3	3.670 ± 0.023	0.0000 ± 0.0000	2.712 ± 0.227
2.3 – 5.0	-1.703 ± 2.038	3.2783 ± 0.9976	-0.935 ± 0.047

In the following, the fit result $\sigma_{\text{PLI}}(p^{\text{gen}}_{\text{T}})$ is used as an estimator for the PLI. Thus, the measured relative jet-p_{T} resolution Eq. (5.16) is corrected for the effect from non-balanced particle-level jets by subtracting $\sigma_{\mathrm{PLI}}(p_{\mathrm{T}}^{\mathrm{gen}})$ in quadrature. The size of this correction is small, less than 9% above 100 GeV, which justifies the usage of the simulation. A systematic uncertainty is assigned to account for inaccuracies in the PLI description, cf. Section 5.5.

5.3.3 Final Expression for the Measured Resolution

Combination of the above corrections for additional jet activity and for the PLI yields the final expression for the relative jet-p_{T} resolution that will be used in the following:

$$\text{relative resolution} = \frac{\sigma(\alpha_{\max} \to 0)}{\langle p_{\mathrm{T}}^{\mathrm{true}} \rangle} \ominus \sigma_{\mathrm{PLI}}(\langle p_{\mathrm{T}}^{\mathrm{true}} \rangle) \equiv \frac{\sigma}{p_{\mathrm{T}}^{\mathrm{true}}} \,, \tag{5.18}$$

where $\ominus$ denotes subtraction in quadrature and with the estimated particle-level jet p_{T}

$$\langle p_{\mathrm{T}}^{\mathrm{true}} \rangle = \int \mathrm{d}p_{\mathrm{T}}^{\mathrm{true}}\; p_{\mathrm{T}}^{\mathrm{true}} f_{\sigma(\alpha_{\max}\to 0)}(p_{\mathrm{T}}^{\mathrm{true}} | p_{\mathrm{T,min}}^{\mathrm{ave}}, p_{\mathrm{T,max}}^{\mathrm{ave}}) \equiv p_{\mathrm{T}}^{\mathrm{true}} \,, \tag{5.19}$$

where $f_{\sigma}(p_{\mathrm{T}}^{\mathrm{true}} | p_{\mathrm{T,min}}^{\mathrm{ave}}, p_{\mathrm{T,max}}^{\mathrm{ave}})$ is the spectrum Eq. (5.8). The uncertainty on both the resolution Eq. (5.18) and on $p_{\mathrm{T}}^{\mathrm{true}}$ Eq. (5.19) stems from propagation of the extrapolation uncertainty $\delta\sigma_{\mathrm{ex}}$ Eq. (5.17).

5.4 Validation of the Method

The described method to measure the jet-p_{T} resolution, including the imbalance corrections Section 5.3, is validated with the full simulation using the PYTHIA sample defined in Section 5.1.2. In Section 5.4.1, the result is compared to the expectation from the MC-truth resolution, and in Section 5.4.2, the impact from pile-up is studied.

5.4.1 Comparison to the MC-Truth Resolution

In Fig. 5.23, the relative jet-p_{T} resolution measured in the PYTHIA simulation is shown as a function of the estimated particle-level jet $p_{\mathrm{T}}^{\mathrm{true}}$ Eq. (5.19) in the different $|\eta|$ intervals. The corrections for the effects of additional jet activity and the PLI have been applied; the extrapolation fits are shown in Figs. A.16 to A.20 in Appendix A.5. Error bars represent the extrapolation uncertainty $\delta\sigma_{\mathrm{ex}}$ Eq. (5.17) from the extrapolation procedure. As remarked in Section 5.3.1, $\delta\sigma_{\mathrm{ex}}$ underestimates the statistical uncertainty of the measurement, which is why the fluctuations of the points are not fully covered.

The result is compared to the relative MC-truth resolution Fig. 4.20. There is generally good agreement within 10% for $p_{\mathrm{T}} \gtrsim 75\,\mathrm{GeV}$, which is degrading for the intervals with larger $|\eta|$. Also at lower p_{T}, the agreement becomes worse. Here, only points with large $\alpha_{\max}$ are considered in the linear extrapolation-fit, and thus, the extrapolation becomes more sensitive to statistical fluctuation of the remaining points. This effect, which possibly contributes to the increasing disagreement, is not fully covered by the displayed extrapolation uncertainty $\delta\sigma_{\mathrm{ex}}$ as remarked above. It will be included in the systematic uncertainties, the size of which depends on the $\alpha_{\max}$ of the first point considered for the linear extrapolation fit, cf. Section 5.5.

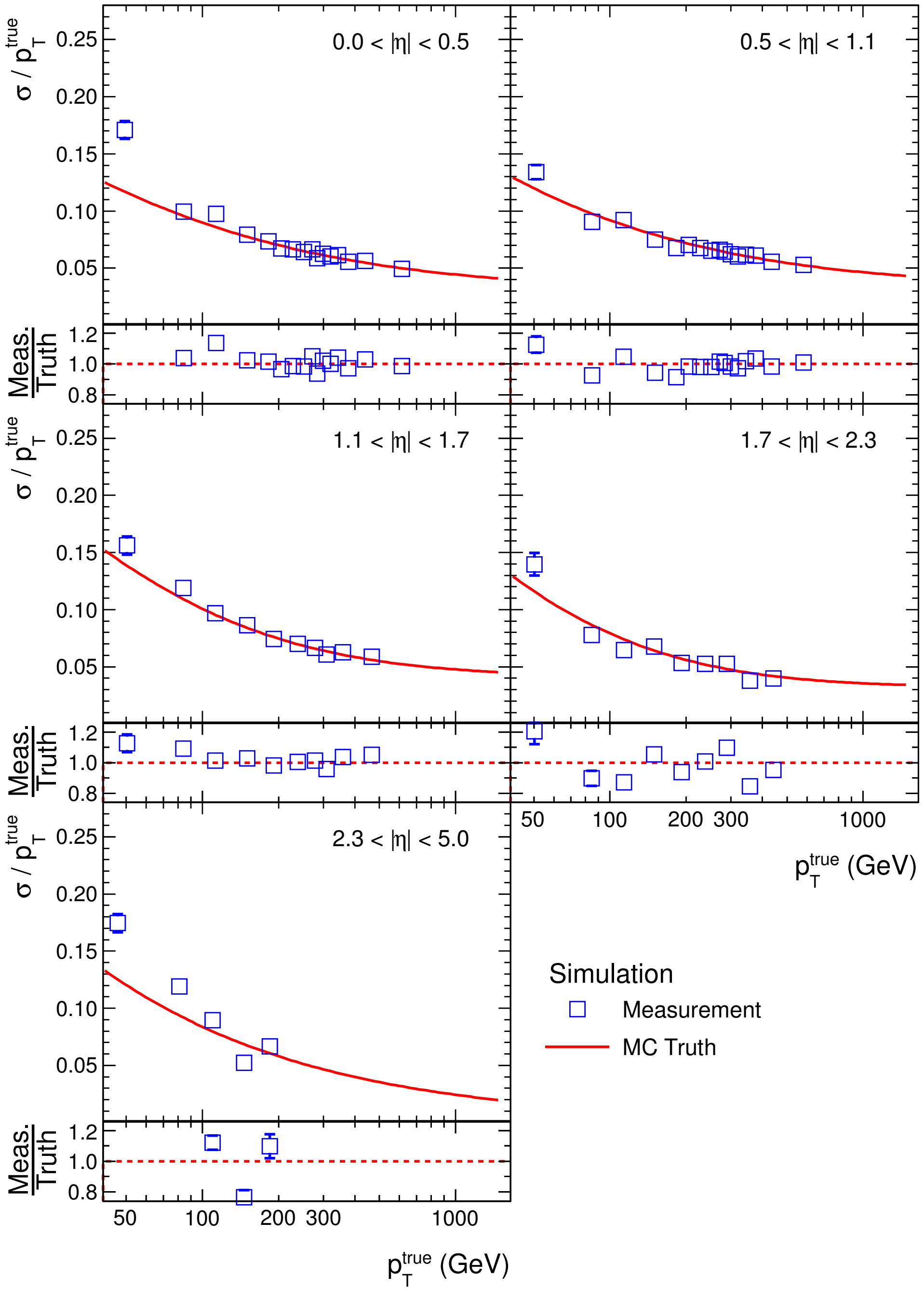


Figure 5.23: Measured relative jet-p_T resolution (*open squares*) in simulated events as a function of the estimated particle-level jet p_T^{true} Eq. (5.19) compared to the relative MC-truth resolution (*solid line*) in different $|\eta|$ intervals. Error bars represent the extrapolation uncertainty $\delta\sigma_{ex}$ Eq. (5.17), which depends on but underestimates the statistical uncertainties. The bottom part of the pads displays the ratio of the measured and the MC-truth resolution.

In the following measurements, the results are not corrected for the residual disagreement, which will be denoted *intrinsic bias*. Rather, the primary result of the analysis will be the ratio of the resolution measured in data and in simulation, and therefore, the intrinsic bias is expected to cancel.

5.4.2 Dependence on Pile-Up

Additional energy deposits by particles from pile-up events bias the resolution at low p_{T} towards larger values, as has been shown for the MC-truth resolution in Section 4.6.1. If the resolution is measured from dijet data with the presented method further effects might occur.

The presence of pile-up events might affect the extrapolation procedure. For example, the additionally distributed energy can increase the measured third-jet p_{T}, or pile-up induced jets can occur with p_{T} larger than that of the third jet from the primary collision. In both cases, α will be too large. Such an effect has been observed in Fig. 5.19 for low $p_{\mathrm{T}}^{\mathrm{ave}}$ and low α. As a result, the slope of the extrapolation might change, and the selected number of events with small α is reduced, which increases the uncertainty.

Furthermore, a jet originating in a pile-up collision can be mistaken as one of the leading two jets from the primary interaction, in which case the assumption of transverse-momentum balance is invalid. However, the rate of this configuration is expected to be very low since high-p_{T} pile-up jets are extremely rare and since the $\Delta\phi$ criterion requires back-to-back jets.

In order to investigate the impact of pile-up on the result of the measurement, the Pythia sample is divided into events with $N_{\mathrm{PU}} \leq 5$ and with $N_{\mathrm{PU}} \geq 6$ pile-up events. Event weights have not been applied for this study. The σ fitted from the two samples are compared in Fig. 5.24 for various thresholds $\alpha_{\max}$.

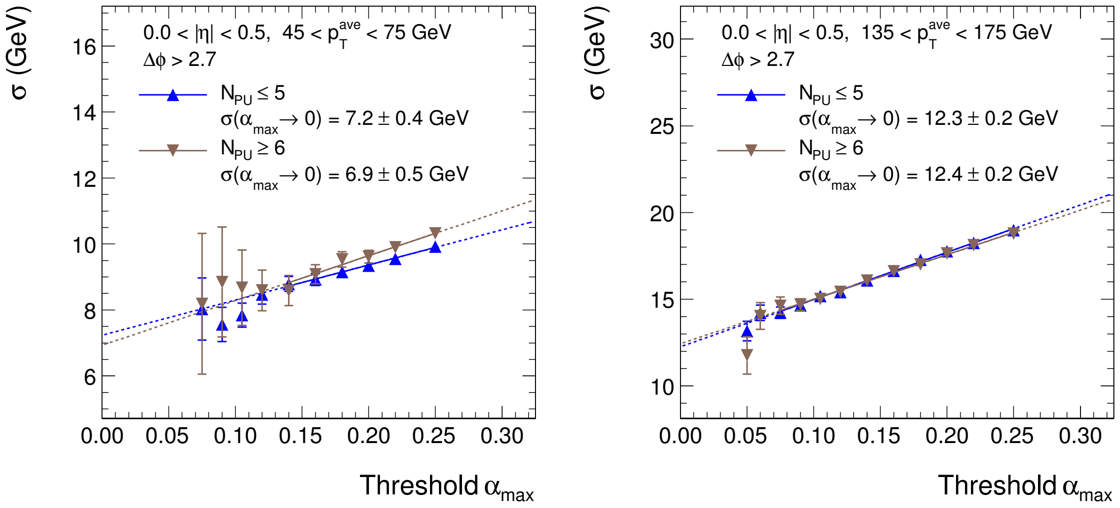


Figure 5.24: Fitted σ as a function of the $\alpha_{\max}$ for simulated events with a different number N_{PU} of pile-up collisions, shown in two different $p_{\mathrm{T}}^{\mathrm{ave}}$ intervals. Only points corresponding to an absolute $p_{\mathrm{T},3} > 10\,\mathrm{GeV}$ are considered for the linear extrapolation-fit.

In the low $p_{\mathrm{T}}^{\mathrm{ave}}$ interval $45 < p_{\mathrm{T}}^{\mathrm{ave}} < 75\,\mathrm{GeV}$, where the largest pile-up impact is expected, the statistical precision is limited. Sufficient precision is obtained only for $\alpha_{\max} \gtrsim 0.14$. There, the σ are generally a few percent larger for the $N_{\mathrm{PU}} \geq 6$ sample. This difference is not reflected by the extrapolated values, though, because the extrapolation fit is dominated by statistical fluctuations of the points. In the larger $p_{\mathrm{T}}^{\mathrm{ave}}$ interval $135 < p_{\mathrm{T}}^{\mathrm{ave}} < 175\,\mathrm{GeV}$, no statistically significant difference is observed between the two pile-up scenarios.

In summary, the presence of pile-up events increases the measured σ by a few percent for $p_{\mathrm{T}}^{\mathrm{ave}} < 75\,\mathrm{GeV}$. The effect is statistically hardly significant, though, and becomes negligible at larger $p_{\mathrm{T}}^{\mathrm{ave}}$. Its impact on the final resolution is reduced because points corresponding to $p_{\mathrm{T},3} < 10\,\mathrm{GeV}$ are not considered in the extrapolation fit as explained in Section 5.3.1. Hence, pile-up is not expected to bias the measurement. In addition, the final result is reported as a data-to-simulation ratio. Since pile-up interactions are modelled in the simulation, biases due to the presence of pile-up are assumed to cancel. Furthermore, the largest impact occurs at low p_{T}, where the statistical precision of the measurement is small due to the limited size of the data sample collected by highly-prescaled triggers. The quoted ratio is p_{T} independent, however, and hence dominated by the medium-p_{T} region[12], which is not affected by pile-up.

5.5 Systematic Uncertainties

The systematic uncertainties of the resolution measurement are evaluated using the Pythia sample defined in Section 5.1.2. The following sources for uncertainties are considered:

Correction for additional jet activity ('Additional Jets'): The transverse-momentum imbalance induced by the presence of additional jets is corrected for by an extrapolation method, where a linear dependence of the measured σ on the α threshold $\alpha_{\max}$ is assumed down to $\alpha = 0$. In order to quantify possible uncertainties introduced by this assumption, 50% of the difference between the extrapolation function's values at $\alpha_{\max} = \alpha_{\max,1}$ and at $\alpha_{\max} = 0$ is quoted as symmetric uncertainty. $\alpha_{\max,1}$ denotes the threshold of the point with smallest $\alpha_{\max}$ still taken into account for the linear fit. The procedure is illustrated in Fig. 5.25 (*left*). Since the assumed linear behaviour is well motivated as discussed in Section 5.3.1, the assigned systematic uncertainty is rather conservative. This is done on purpose to cover the underestimation of the statistical uncertainties by $\delta\sigma_{\mathrm{ex}}$ Eq. (5.17).

PLI correction ('PLI'): The measured resolution is corrected for the estimated bias due to the PLI using the simulation. 25% of the correction is assigned as uncertainty. Since the PLI is subtracted in quadrature from the extrapolated relative resolution, cf. Eq. (5.18), the obtained uncertainty is asymmetric by construction.

Particle-level dijet cross-section ('Spectrum'): The particle-level jet-p_{T} spectrum f in Eq. (5.8) is obtained from a Pythia simulation. Since the cross section is difficult to model, the spectrum obtained from the Herwig++ event generator [207] (with tune 23) is used instead. The difference between the measured resolution in both cases is assigned as symmetric uncertainty. Both spectra are compared in Fig. 5.25 (*right*).

---

[12] At large p_{T}, the amount of data is limited by the QCD cross-section.

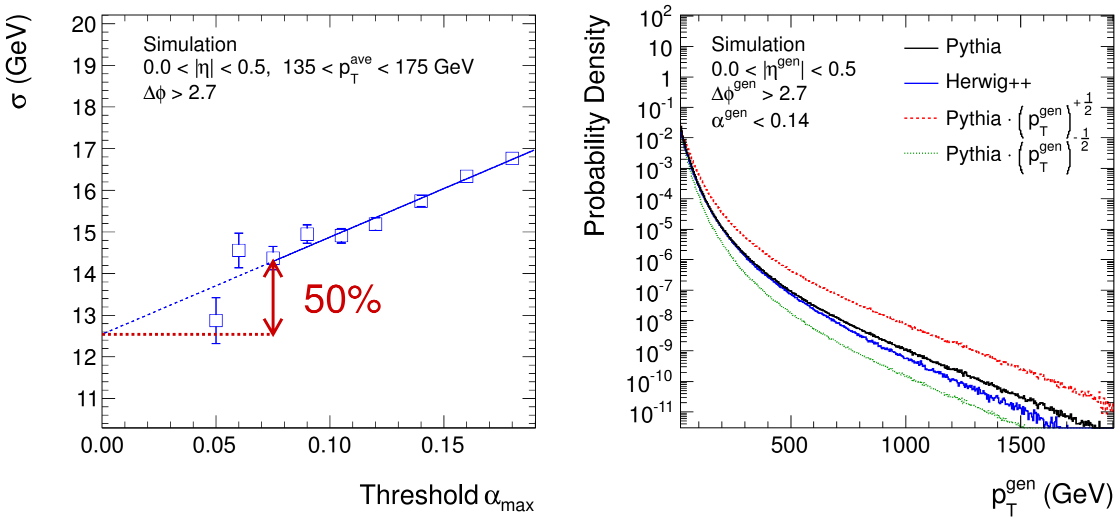


Figure 5.25: σ as a function of $\alpha_{\max}$ (*left*). The systematic uncertainty due to the extrapolation procedure is indicated. It is 50% of the difference between the extrapolation function's values at $\alpha_{\max} = \alpha_{\max,1}$ and at $\alpha_{\max} = 0$, where $\alpha_{\max,1}$ is the position of the point with smallest $\alpha_{\max}$ taken into account for the linear fit. $p_{\mathrm{T}}^{\mathrm{gen}}$ spectra of QCD-dijet events obtained with the PYTHIA and the HERWIG++ generator (*right*). Both are used as $f(p_{\mathrm{T}}^{\mathrm{true}})$ in Eq. (5.8), and the difference of the measured resolutions is taken as systematic uncertainty due to the particle-level dijet cross-section. To test the robustness of the method, also variations of the spectrum by a factor $(p_{\mathrm{T}}^{\mathrm{gen}})^{\pm 1/2}$ have been tested.

In fact, the maximum likelihood method has been found to be quite robust with respect to the exact shape of f. For example, a drastic distortion like scaling f by $(p_{\mathrm{T}}^{\mathrm{gen}})^{\pm 1/2}$ as depicted in Fig. 5.25 (*right*) results in a relative change of the measured resolution by less than 5%. However, the difference between the spectra from the advanced generators PYTHIA and HERWIG++ is considered a realistic estimate of the uncertainty in the simulation. The good description of the measured jet-p_{T} distributions' shape by the PYTHIA generator, cf. Fig. 5.6, supports the assumption of a sufficiently well modelled p_{T} spectrum of the leading two jets.

Jet energy scale ('JES'): In order to evaluate the impact of the jet-energy-scale uncertainty[13], the applied calibration factors are varied up and down according to their uncertainties [172]. The average difference of the measured resolution in the two cases is taken as uncertainty.

Pile-up modelling ('PU'): Biases of the measurement due to the presence of pile-up collisions are small even at low p_{T}, cf. Section 5.4.2. In addition, the final result, which is a data-to-simulation ratio, is independent of the pile-up multiplicity provided it is correctly modelled in the simulation. In order to quantify the impact of imprecise

---

[13] Changing the jet energy scale is equivalent to changing the spectrum; nonetheless, both effects are considered as individual uncertainties since they may both be inaccurate due to different sources.

pile-up modelling, the measurement is repeated with the mean of the pile-up multiplicity distribution Fig. 4.8 (*left*) shifted by ±1. The average of the upper and lower variation of the result is assumed as uncertainty.

The relative size of the systematic uncertainties on the measured resolution in the different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals is shown in Fig. 5.26. The dominant contribution arises from the extrapolation method, which includes parts of the statistical uncertainty on σ.

The components are added in quadrature because their sources are considered uncorrelated. Since the PLI uncertainty is asymmetric, the addition is done separately for upper and lower components[14].

The total uncertainty amounts to $5-10\%$ for $p_{\mathrm{T}} > 100\,\mathrm{GeV}$ in the central detector region. At low p_{T} and at large $|\eta|$, the determined uncertainties are larger. This is partly due to increased statistical fluctuations of the results when performing the systematic variations described above because of the smaller number of selected events. The increase of the JES component reflects the larger jet-energy-calibration uncertainty in these regions [172].

5.6 Results

The relative jet transverse-momentum resolution measured in $855\,\mathrm{pb}^{-1}$ of dijet data and in the Pythia simulation is shown in Fig. 5.27 together with the extrapolation uncertainty[15] $\delta\sigma_{\mathrm{ex}}$ Eq. (5.17) and the systematic uncertainty. The latter is applied only to the data result because the particle-level jet-p_{T} spectrum, the PLI correction, the jet-energy calibration, and the pile-up scenario, which are sources for systematic uncertainties, have been derived from the simulation. Likewise, the systematic uncertainty assigned to the extrapolation procedure does not apply to the simulation since the method works consistently, cf. Section 5.4.1, and any residual intrinsic bias is expected to affect data and simulation in the same way and thus cancel in the ratio Eq. (5.20) defined below.

The large $\delta\sigma_{\mathrm{ex}}$ uncertainties of the measurements in data below $p_{\mathrm{T}}^{\mathrm{ave}} = 100\,\mathrm{GeV}$ are a consequence of the small number of collected dijet events due to the high trigger prescales, cf. Table 5.1. The increase of the uncertainties with $|\eta|$ is due to the rapidity dependence of the jet-production cross-section, cf. Figs. 4.4 and 5.5. For $|\eta| > 1.1$, the data measurements in the $45 < p_{\mathrm{T}}^{\mathrm{ave}} < 75\,\mathrm{GeV}$ interval have been omitted due to negative slopes of the extrapolation fits as discussed in Section 5.3.1. In the $2.3 < |\eta| < 5.0$ and $135 < p_{\mathrm{T}}^{\mathrm{ave}} < 175\,\mathrm{GeV}$ interval, the measurements in data and simulation lie significantly above and below, respectively, the values expected when comparing to the adjacent points. Possible sources such as trigger inefficiencies or effects from jet energy corrections have been carefully investigated but no systematic effect has been found. Hence, a statistical fluctuation is assumed.

The relative jet-p_{T} resolution in data amounts to $\sigma/p_{\mathrm{T}} = (9.80 \pm 0.30\,(\mathrm{ex.})\,^{+0.84}_{-0.81}\,(\mathrm{syst.}))\%$ at $p_{\mathrm{T}} = 113\,\mathrm{GeV}$ and $(5.40 \pm 0.08\,(\mathrm{ex.})\,^{+0.27}_{-0.26}\,(\mathrm{syst.}))\%$ at $611\,\mathrm{GeV}$, where 'ex.' and 'syst.' denote the extrapolation uncertainty $\delta\sigma_{\mathrm{ex}}$ and the total systematic uncertainty, respectively. The p_{T} and $|\eta|$ dependence of the measured resolution generally follows the expec-

---

[14] This is in fact not exactly correct, cf. [256], but the relative size of the difference between the upper and lower component is small, and hence, the procedure is justified.

[15] Recall that the extrapolation uncertainty depends on but underestimates the statistical uncertainty of the extrapolated σ and that the residual contribution is expected to be covered by the systematic uncertainty.

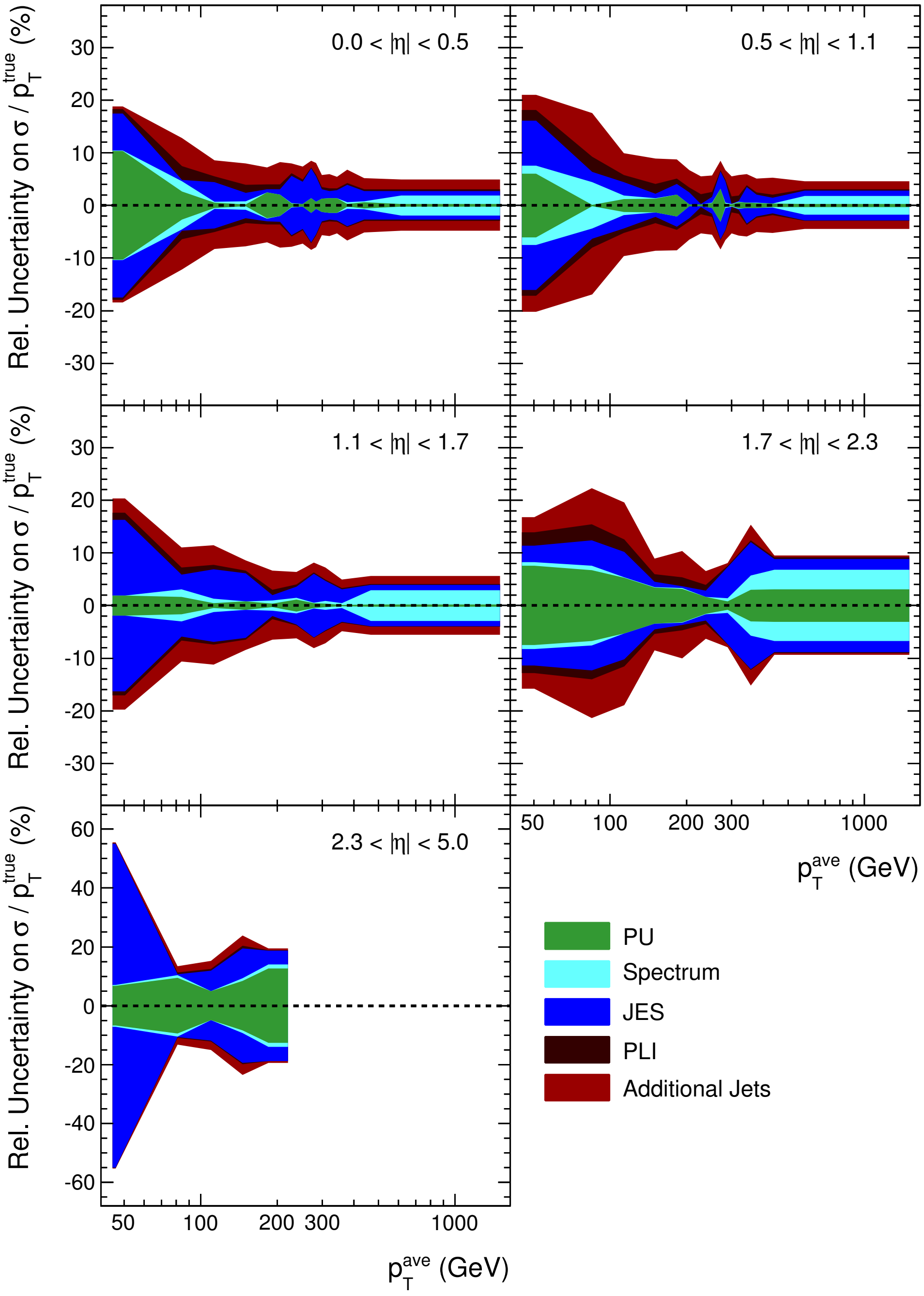


Figure 5.26: Relative systematic uncertainties on the measured resolution as a function of $p_{\mathrm{T}}^{\mathrm{ave}}$ in different $|\eta|$ intervals. The uncertainties are added in quadrature ('stack') and are linearly extrapolated between the $p_{\mathrm{T}}^{\mathrm{ave}}$ intervals. Note the different scale in the bottom-left plot.

tations from the simulation, which have been discussed in detail in Section 4.6.1. For example, an overall worse resolution in the barrel-to-endcap transition region $1.1 < |\eta| < 1.7$ is also observed in data. However, the resolution in data is systematically larger than in

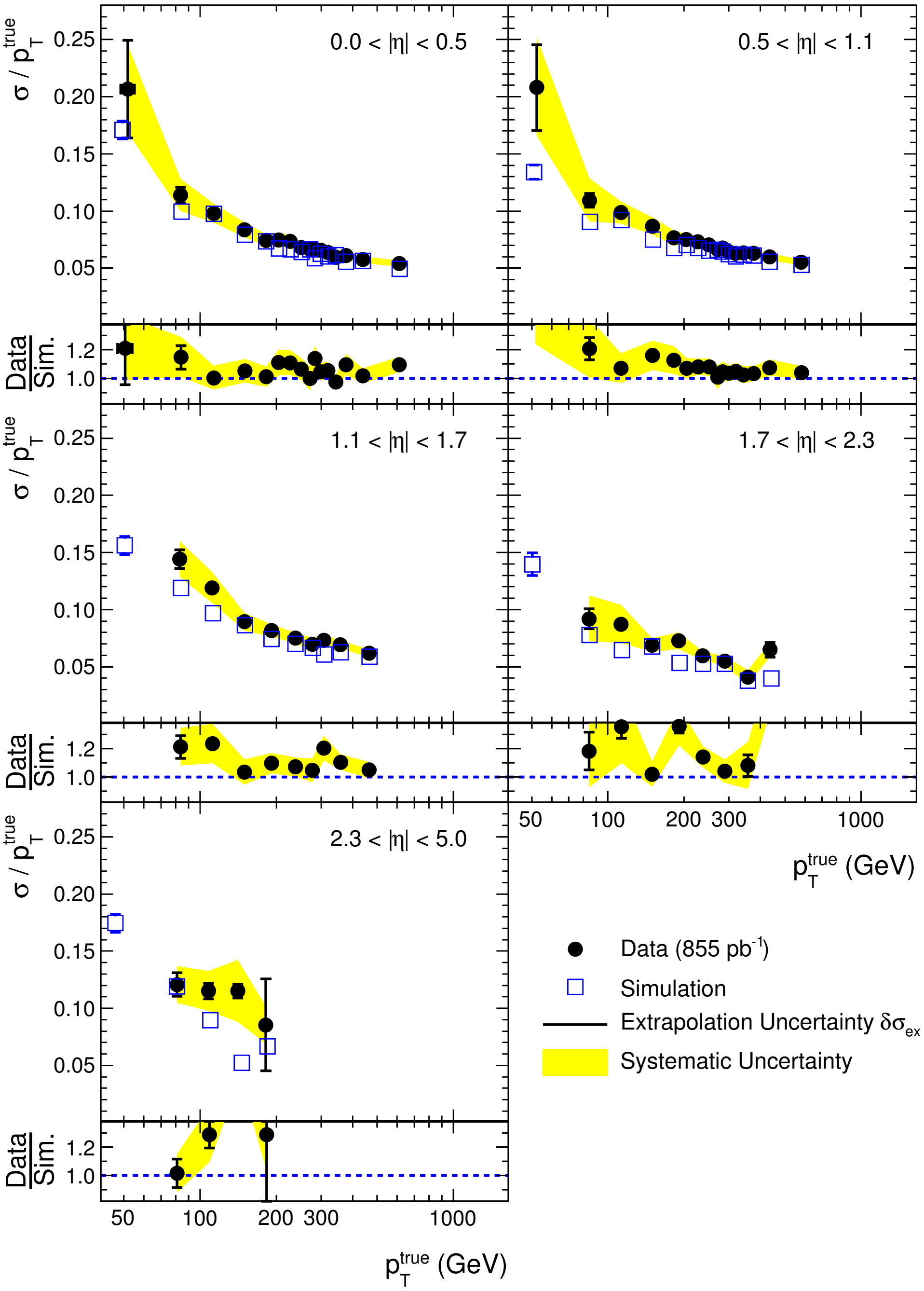


Figure 5.27: Measured relative jet-p_{T} resolution in data (*solid circles*) and simulation (*open squares*) as a function of $p_{\mathrm{T}}^{\mathrm{true}}$ Eq. (5.19) in different $|\eta|$ intervals. Vertical error bars represent the extrapolation uncertainty $\delta\sigma_{\mathrm{ex}}$ Eq. (5.17), horizontal error bars the uncertainty on $p_{\mathrm{T}}^{\mathrm{true}}$, and solid bands represent the total systematic uncertainty. The bottom part of the pads displays the ratio of the resolution in data and the simulation.

the simulation. This effect is clearly visible in Fig. 5.28, where the ratio of the resolution in data and simulation is shown. Since no significant p_T dependence of the data-to-simulation ratio is present, it is parametrised by a constant

$$\rho_\mathrm{res} = \frac{\sigma(\mathrm{data})}{\sigma(\mathrm{simulation})} \; . \tag{5.20}$$

The values of ρ_res are also indicated in Fig. 5.28 and in addition listed in Table 5.4. They increase from $\rho_\mathrm{res} = 1.054 \pm 0.006\,(\mathrm{ex.})\,^{+0.069}_{-0.068}\,(\mathrm{syst.})$ in the central detector region to $1.294 \pm 0.063\,(\mathrm{ex.})\,^{+0.214}_{-0.211}\,(\mathrm{syst.})$ at large $|\eta|$, where 'ex.' and 'syst.' denote the propagated extrapolation uncertainty $\delta\sigma_\mathrm{ex}$ and propagated total systematic uncertainty, respectively.

Table 5.4: Ratio ρ_res of the jet-p_T resolution measured in data and simulation in different $|\eta|$ regions together with the propagated extrapolation uncertainty ('ex.') and the propagated upper and lower systematic uncertainties ('syst.').

$\lvert\eta\rvert$	ρ_res	ex.	syst.
0.0 – 0.5	1.054	$\pm$ 0.006	$^{+0.069}_{-0.068}$
0.5 – 1.1	1.060	$\pm$ 0.006	$^{+0.066}_{-0.064}$
1.1 – 1.7	1.092	$\pm$ 0.008	$^{+0.074}_{-0.072}$
1.7 – 2.3	1.141	$\pm$ 0.020	$^{+0.114}_{-0.111}$
2.3 – 5.0	1.294	$\pm$ 0.063	$^{+0.214}_{-0.211}$

This difference between data and simulation has been observed also for other types than PF jets. The result is consistent with an earlier version of this analysis as well as with dijet-asymmetry- and photon + jet-based measurements, which have been performed with $36\,\mathrm{pb}^{-1}$ of data collected in 2010 [172, 234]. The source of the difference is still under investigation. Possibly, inaccurate descriptions of noise, detector inhomogenities, the e/h differences between the ECAL and HCAL, and shower shapes as well as intercalibration effects, in particular of the HCAL cells [234], might contribute.

5.7 Adjustment of the MC-Truth Resolution to Data

Depending on the application in physics analysis, different methods are suggested how to employ the ρ_res factors to improve the description of the data. If only the jet resolution is required it is sufficient to simply scale the MC-truth resolution Eq. (4.7) by ρ_res. This is useful for example in QCD cross-section measurements where the jet p_T spectrum needs to be unfolded from the resolution. If on the other hand the full response distribution, including the non-Gaussian tails, is required the distribution can be adjusted to the data by convolution with a Gaussian of appropriate width as described in Section 5.7.1. This

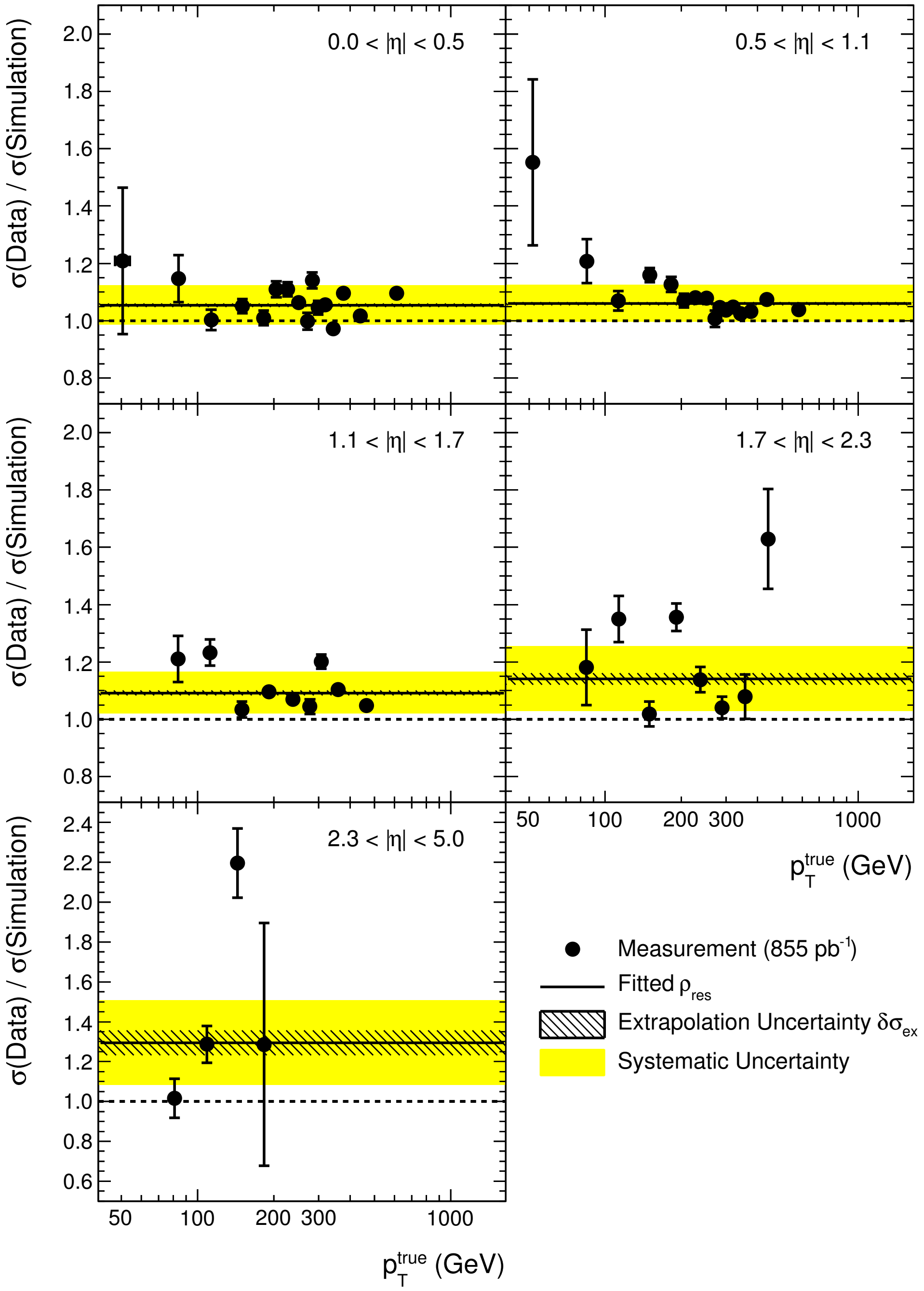


Figure 5.28: Ratios of the jet-p_{T} resolution in data and simulation (*circles*) as a function of $p_{\mathrm{T}}^{\mathrm{true}}$ Eq. (5.19) in different $|\eta|$ intervals. Vertical error bars represent the propagated extrapolation uncertainties $\delta\sigma_{\mathrm{ex}}$ Eq. (5.17), horizontal error bars the uncertainties on $p_{\mathrm{T}}^{\mathrm{true}}$. The ratios are fitted with a constant ρ_{res} (*solid line*); hatched areas indicate the extrapolation, solid areas the systematic uncertainties on ρ_{res}. Note the different scale in the bottom-left plot.

technique is applied later on to the asymmetry distributions in the measurement of the jet response tails in Chapter 6 and also in case of the QCD-background prediction in the new-physics search presented in Chapter 7. Finally, if data-to-simulation correspondence is necessary on a per-jet level a correction can be applied to each simulated detector-level jet depending on the momentum difference to the associated generator-level jet, which is discussed in Section 5.7.2.

5.7.1 Adjustment of the MC-Truth Response Distribution

The full MC-truth response ($\mathcal{R}$) distribution Eq. (4.3) can be adjusted to the data by convolution with a Gaussian of appropriate width σ_c such that afterwards the width $\sigma_{\mathcal{R}}$ of the $\mathcal{R}$ distribution corresponds to the relative resolution in data. Since it is convoluted with a Gaussian, the almost Gaussian shape of the $\mathcal{R}$ distribution will be preserved. This method is interesting, if random numbers are to be sampled according to the jet-p_{T} response.

Assuming $\mathcal{R}$ is Gaussian shaped with standard deviation $\sigma_{\mathcal{R}} = \sigma_{\mathrm{MC}}/p_{\mathrm{T}}$ Eq. (4.6), the modified distribution will have a width

$$\left(\frac{\sigma_{\mathrm{MC}}}{p_{\mathrm{T}}}\right)' = \frac{\sigma_{\mathrm{MC}}}{p_{\mathrm{T}}} \oplus \sigma_c \; .$$

After convolution, the MC-truth resolution should be equal to the resolution in data, i. e. $\sigma_{\mathrm{MC}} = \sigma_{\mathrm{data}}$, and hence

$$\sigma_c = \sqrt{\rho_{\mathrm{res}}^2 - 1} \cdot \left(\frac{\sigma_{\mathrm{MC}}}{p_{\mathrm{T}}}\right) \; . \tag{5.21}$$

Of course, this method is only applicable if $\rho_{\mathrm{res}} > 1$, which is the case however. If not, the technique described in the following has to be used.

5.7.2 Correction of Individual Jet Momenta

Finally, each detector-level jet can be corrected using the difference $\Delta p_{\mathrm{T}} = p_{\mathrm{T}} - p_{\mathrm{T}}^{\mathrm{gen}}$ in transverse momentum to the associated generator-level jet [234]. Since the resolution is proportional to the average Δp_{T}, adjustment of the resolution to data corresponds to scaling Δp_{T} by ρ_{res}. Hence, the smeared transverse momentum p'_{T} becomes

$$p'_{\mathrm{T}} = p_{\mathrm{T}}^{\mathrm{gen}} + \rho_{\mathrm{res}} \cdot \left(p_{\mathrm{T}} - p_{\mathrm{T}}^{\mathrm{gen}}\right) \; .$$

5.8 Outlook

Several extensions of the presented method for the jet-p_{T} resolution measurement appear worthwhile.

A modification of the dijet pdf Eq. (5.11) such that two jets with different resolution can be considered would allow the inclusion of events where the two jets lie in different $|\eta|$ intervals. This would drastically improve the statistical precision of the measurement, which is in particular important in the forward region to enable finer $|\eta|$ intervals and also at low p_{T}.

The likelihood Eq. (5.12) can be extended to incorporate photon + jet or Z + jet events. A pdf based on the p_{T}-balance between the well-measured photon/Z and the jet is simple compared to the dijet pdf Eq. (5.11) because it does not require integration over $p_{\mathrm{T}}^{\mathrm{true}}$, cf. Section 4.4.5. Even though photon/Z + jet events would predominantly contribute at low p_{T} due to their smaller cross section, a reduction of the systematic uncertainties and possibly of the low-p_{T} intrinsic bias is therefore expected. Given the high prescales of the $p_{\mathrm{T}}^{\mathrm{ave}}$ trigger-paths with low thresholds, the rate of collected photon/Z + jet events at low p_{T} becomes in fact comparable to and even higher than that of dijet events. Hence, the statistical precision of the measurement at low p_{T} would also be improved.

This can already be achieved by a simple statistical combination of the results from the dijet-based maximum-likelihood method with the results from independent photon/Z+jet-based measurements. Such a combination is shown in Fig. 5.29 for the results from a previous version of the presented analysis and from a photon + jet analysis that have both been performed with 2010 data [172]. The systematic uncertainty of the photon + jet analysis is approximately 3% in the central detector region, which reduces the uncertainty of the combined ρ_{res}. For the current results presented in this thesis, a combination has not been done because so far the resolution has not been measured by CMS in photon/Z + jet events of the 2011 data.

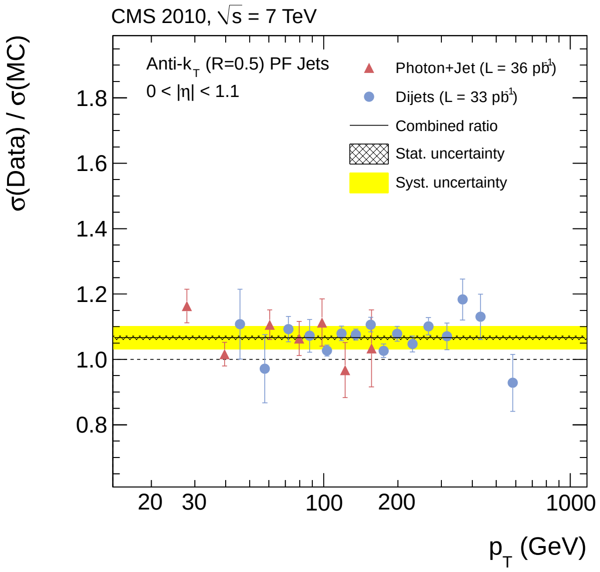


Figure 5.29: Ratio of the resolution measured in data and simulation with data collected in 2010 by CMS. Measurements from a previous version of the presented analysis (*circles*) and from a photon+jet-based analysis (*triangles*) are shown. A combined fit to the results from both analyses is also shown (*solid line*). Details can be found in [172].

Also, a modification of the extrapolation procedure such that the correlations are explicitly taken into account is desirable. Furthermore, a parametrisation of the response with other than a Gaussian function would be beneficial because in that case also the non-Gaussian tails could be accessed. However, it turns out to be challenging to find a parametrisation that is both suitable to describe the response in data accurately over three orders of magnitude and mathematically simple enough to be evaluated many times during the minimisation. Therefore, a different, dijet-asymmetry-based approach to measure the tails is developed in the following Chapter 6. Possible other, technically more challenging extensions to exploit the full potential of the maximum-likelihood approach include the fit of a p_{T}-dependent σ and a simultaneous fit of the particle-level jet-p_{T} spectrum.

6 Measurement of the Jet Transverse-Momentum Response Tails

The jet transverse momentum resolution measured in Chapter 5 characterises the performance of jet measurements at CMS in a compact way sufficient for many analyses. However, a comprehensive understanding of jets requires the knowledge of the full transverse momentum response function including the non-Gaussian tails. The latter are caused by physics effects such as semi-leptonic decays of heavy-flavour quarks and also by instrumental effects such as inactive channels, inhomogeneities of the detector material, electronic noise, or miscalibration, cf. Section 4.4.2. In case of Particle-Flow (PF) jets, cf. Section 4.2.2, artefacts of the complex reconstruction algorithm possibly also contribute. Although the non-Gaussian tails contribute only at the level of a few percent or below to the to response distribution, they cause substantial jet-momentum mismeasurements and thus can significantly affect analyses in extreme phase-space regions, e. g. at high $\not{E}_{\mathrm{T}}$, relevant for new-physics searches as discussed in Chapter 7. Hence, studying the transverse momentum response tails is vital for an evaluation of the performance of detector components and object reconstruction and provides an essential input to searches.

In this chapter, a measurement of the dijet-asymmetry tails, cf. Section 4.6.2, in dijet events with an integrated luminosity of $4.90\,\mathrm{fb}^{-1}$ is presented. Due to the relation between asymmetry and response, non-Gaussian tails in the jet transverse momentum response function manifest as non-Gaussian tails in the asymmetry distribution. However, since the dijet asymmetry results in a symmetric distribution by construction, the assignment of the asymmetry tails to the low- and high-response tails is ambiguous. Therefore, this approach is based on comparing the dijet-asymmetry tails in data to the expectations from simulation. Differences in the fractional number of tail events imply an incorrectly modelled response function. By varying the considered tail region, possible shape differences can also be detected. Results of this measurement have been published in [257]. Previous results of this analysis obtained with $855\,\mathrm{pb}^{-1}$ of 2011 data and $36\,\mathrm{pb}^{-1}$ of 2010 data have been published in [238] and [172, 239], respectively.

Some ambiguities in assigning the observed differences to the low or high-response tails remain, though, and therefore it is worthwhile to complement the results with measurements from for example photon + jet or Z + jet events that allow direct access to the jet-p_{T} response function. In principle, the maximum-likelihood method described in the previous Chapter 5 can also be used to measure the tails but finding a suitable response parametrisation turns out to be challenging.

This chapter is organised as follows. In Section 6.1, the selection of dijet events from data and simulation is described. The selection is analogue to the one for the resolution measurement, but this study is performed on the full data set collected during the 2011 LHC-run. Afterwards, in Section 6.2, the method is explained in detail, and a data-to-simulation ratio of the fractional number of tail events is defined. The dependence of the asymmetry on the presence of pile-up is investigated in Section 6.3. Systematic uncertainties associated with the measurement are discussed in Section 6.4. The measured size of the asymmetry tails of PF jets, absolute and relative to the simulation, is presented in Section 6.5 for different tail regions. A method to correct the simulation using the result of the tail measurement is reviewed in Section 6.6. Finally, in Section 6.7, possible improvements of the analysis are discussed.

6.1 Samples And Event Selection

For this measurement, the dijet data selected for the resolution measurement, cf. Section 5.1, have been updated to an integrated luminosity of $4.90\,\mathrm{fb}^{-1}$. The amount of data collected by the contributing trigger paths[1] is listed in Table 6.1. Also, an updated version of the simulated event sample has been used, which takes into account the increased number of on average 12 pile-up interactions during the later 2011 data-taking period. The selection efficiencies of the cleaning-filter requirements are stated in Table B.1 in Appendix B.1. About 5% of the events are rejected, slightly more than before due to an increased rate of tracking failures in the dense environment of events with many pile-up collisions. In particular the filter against events where a jet covers masked ECAL-channels (TP and BE filters, cf. Section 5.1.4) is of importance for this study because these jets are often severely mismeasured and would appear in the tails of the response distribution [246].

Table 6.1: High-level trigger paths employed for the response-tail measurement. Stated are the path name which indicates the threshold on the trigger-level $p_{\mathrm{T,HLT}}^{\mathrm{ave}}$, the threshold on the offline $p_{\mathrm{T}}^{\mathrm{ave}}$ where the trigger efficiency is $\geq 99\%$, and the integrated luminosity of the collected data taking into account trigger prescales. The relative uncertainty on the luminosity measurement is 2.2% [240].

Trigger	$p_{\mathrm{T},99}^{\mathrm{ave}}$ (GeV)	Luminosity (pb^{-1})
`HltDiJetAve30`	45	0.02458
`HltDiJetAve60`	75	0.733
`HltDiJetAve80`	100	1.323
`HltDiJetAve110`	135	12.01
`HltDiJetAve150`	175	54.6
`HltDiJetAve190`	220	274.1
`HltDiJetAve240`	270	1001
`HltDiJetAve300`	335	4890
`HltDiJetAve370`	405	4900

Since in this study the tails of the dijet asymmetry distributions are measured, the relative rate of the rare SM-background processes mentioned in Section 5.1.5 such as $W+$jets and $Z+$jets events is larger than in case of the resolution measurement. However,

---

[1] The `HltDiJetAve`-trigger definition has been changed shortly after the start of the 2011 data-taking period to incorporate jet energy corrections at trigger level. Therefore, the amount of data employed in this analysis is approximately 1.5% smaller than the total amount of data collected during 2011 with the detector fully functional.

the dijet selection-criteria that require the leading two jets to be back-to-back in azimuth and little relative p_{T} of any further jet as well as the veto of events with isolated leptons are assumed to sufficiently suppress the presence of background events to a level such that any impact on the measurement can be neglected. This is discussed in Section 6.2.2.

In order to acquire enough events in the studied tail regions and to obtain sufficient statistical precision, the asymmetry distributions measured in several of the previously defined $p_{\mathrm{T}}^{\mathrm{ave}}$ intervals are merged. The boundaries of the new, wider intervals, which have been defined to contain roughly the same number of events, are listed in Table 6.2.

Table 6.2: $|\eta|$ and $p_{\mathrm{T}}^{\mathrm{ave}}$ interval boundaries for the response-tail measurement.

$\|\eta\|$	$p_{\mathrm{T}}^{\mathrm{ave}}$ (GeV)
$0.0, 0.5$	$45, 220, 270, 312, 360, 498, 1500$
$0.5, 1.1$	$45, 220, 294, 360, 1500$
$1.1, 1.7$	$45, 220, 335, 1500$
$1.7, 2.3$	$45, 220, 1500$
$2.3, 5.0$	$45, 220$

Since in this study the asymmetry distributions in data and simulation are compared, the trigger prescales have to be considered. For the original selection, only events collected by fully efficient trigger paths are considered in each $p_{\mathrm{T}}^{\mathrm{ave}}$ interval. Hence, the number of available events is reduced by some effective prescale which differs between the intervals. If the distributions from several intervals are merged they will thus contribute with different weights, and this effect has to be considered when combining the simulated distributions because triggers are not included in the simulation. Therefore, the simulated distributions in each of the original intervals are weighted to contain the same number of events as present in data. The weighting factors are determined by comparison of the $p_{\mathrm{T}}^{\mathrm{ave}}$ distributions in data and simulation, as demonstrated in Fig. 6.1.

6.2 Description of the Method

This measurement is based on the comparison of the dijet asymmetry Eq. (4.9) in data, $\mathcal{A}_{\mathrm{data}}$, and in simulation, $\mathcal{A}_{\mathrm{MC}}$. First, in Section 6.2.1, the $\mathcal{A}_{\mathrm{MC}}$ distribution is corrected for the resolution difference between data and simulation. Then, in Section 6.2.2, the asymmetry tails are defined, and in Section 6.2.3, a procedure is described to extrapolate the fractional number of tail events to the case of an ideal dijet event. As discussed in Section 6.2.4, the ratio of the tail size in data and simulation after extrapolation probes simulation differences independently of several systematic effects that affect the absolute measurements.

For technical reasons, the distributions of absolute asymmetry $|\mathcal{A}|$ are considered in the following. Furthermore, the distributions have been normalised to their integral. This is not required for the measurement since relative event numbers are compared, but it enables a by-eye comparison of the superimposed histograms.

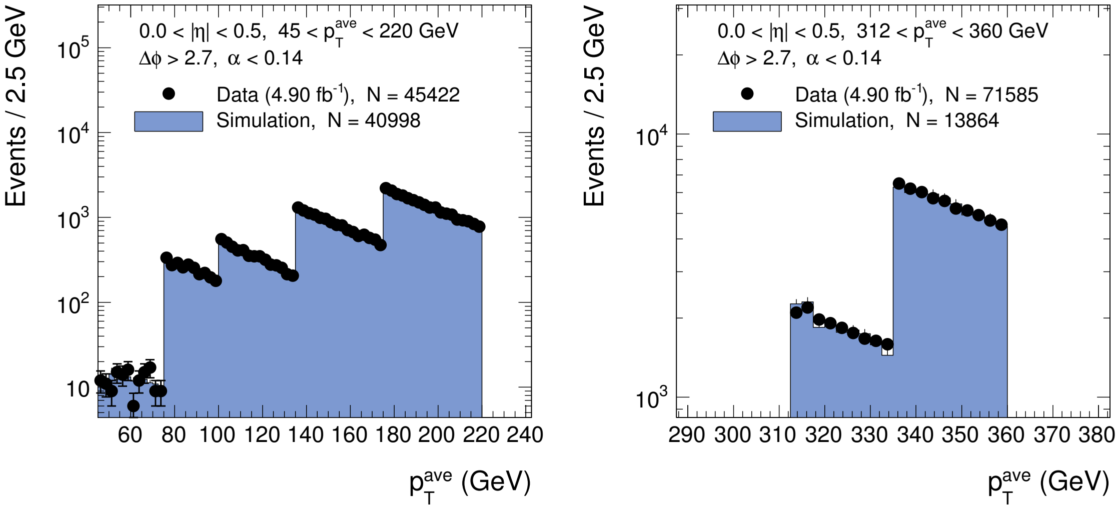


Figure 6.1: $p_{\mathrm{T}}^{\mathrm{ave}}$ distributions in data (*solid circles*) and simulation (*shaded area*) in two $p_{\mathrm{T}}^{\mathrm{ave}}$ intervals for $|\eta| < 0.5$ and $\alpha < 0.14$. The number N of events events is indicated. The data are collected through several trigger paths with different prescales. Edges in the distributions correspond to those $p_{\mathrm{T}}^{\mathrm{ave}}$ where a further path contributes and are equal to the interval boundaries Table 5.2 of the original event selection. In each of these intervals, the simulated histograms are scaled such that the integrated number of events is the same as in data. The scale factors are used to correct the other simulated distributions in this analysis for the effective trigger prescales.

6.2.1 Correction for the Resolution Difference

The jet transverse momentum resolution differs between data and simulation as shown in Chapter 5.6. As a consequence of Eq. (4.11), also the $\mathcal{A}_{\mathrm{data}}$ and $\mathcal{A}_{\mathrm{MC}}$ distributions differ, which is demonstrated in Fig. 6.2 (*left*)[2]. Therefore, in order to be able to compare the tail regions, first the $\mathcal{A}_{\mathrm{MC}}$ distribution is corrected to better describe the data in the core region. The correction is done following the procedure described in Section 5.7.1. Due to the proportionality in Eq. (4.11), the measured resolution ratios ρ_{res} in Table 5.4 are also applicable when correcting the asymmetry distributions. The original and the corrected $\mathcal{A}_{\mathrm{MC}}$ distributions compared to the data are shown in Fig. 6.2 and in Figs. B.1 and B.3 in Appendix B.2.

In the following, the simulated asymmetry distributions, and hence all derived quantities, are always shown after this correction for the larger resolution. They are labelled with 'Corrected $\sigma_{\mathcal{A}}$'.

---

[2] For illustration purposes, distributions of the events with $\alpha < 0.075$ are shown here and in the following. They are closer to the ideal dijet topology than in case of $\alpha < 0.14$ shown so far and contain enough events for sufficient statistical precision when investigating their features.

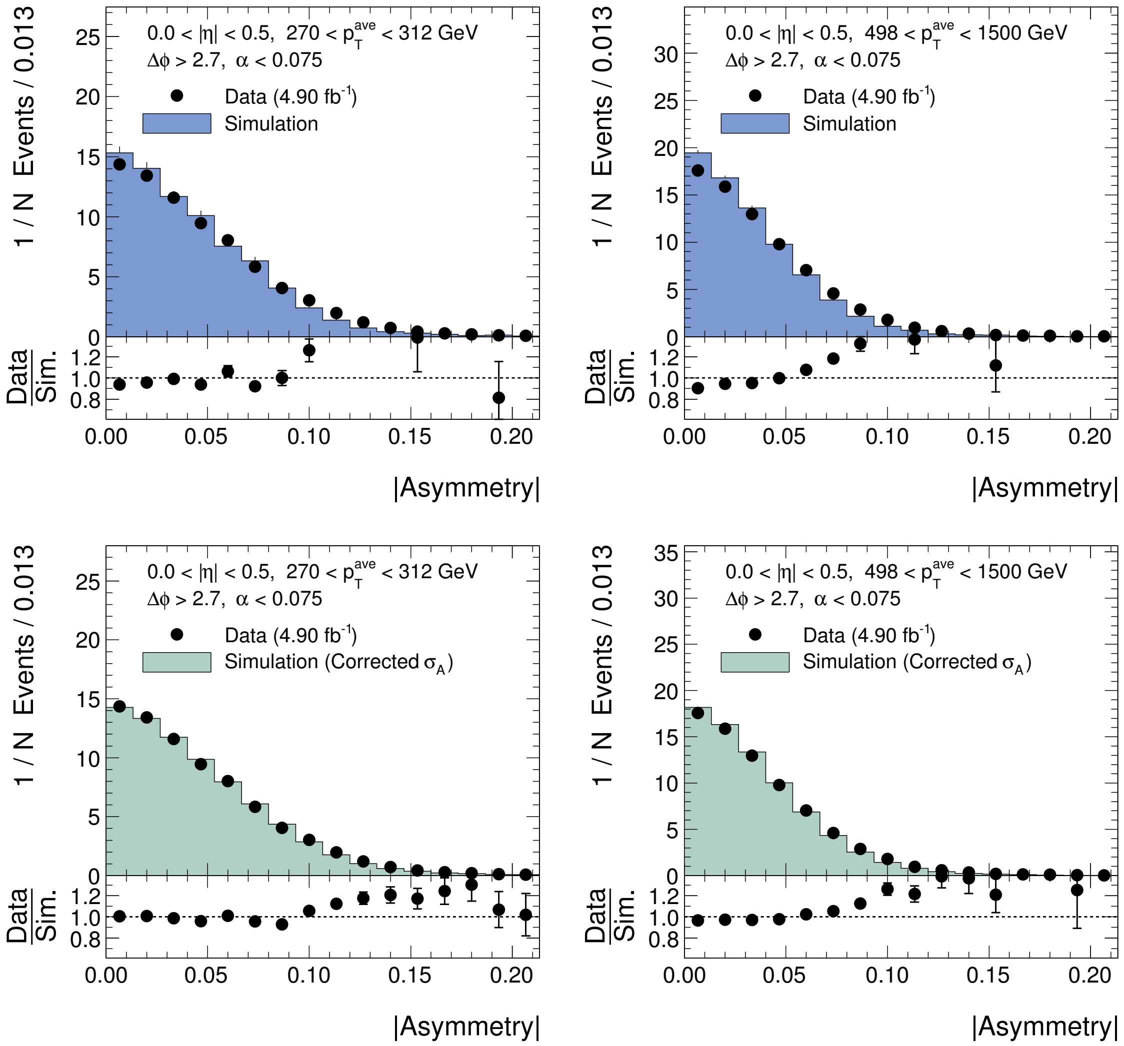


Figure 6.2: Asymmetry distributions in data (*solid circles*) and simulation (*shaded area*) and data-to-simulation ratio (*bottom part of the pads*) in two $p_{\mathrm{T}}^{\mathrm{ave}}$ intervals (*columns*) for $|\eta| < 0.5$ and $\alpha < 0.075$. The simulated distribution is shown before (*top row*) and after (*bottom row*) correction for the larger jet-p_{T} resolution in data ('Corrected $\sigma_{\mathcal{A}}$').

6.2.2 Definition of the Tail Region and the Number of Tail Events

In this analysis, the tail region of an asymmetry distribution is defined in terms of a common standard deviation σ_c. The latter is obtained from a Gaussian fit to the corrected $\mathcal{A}_{\mathrm{MC}}$ distribution of all dijet events with $\alpha < 0.05$, which corresponds to the tightest selection with least distortion of the transverse-momentum balance in the dijet system,

$$\sigma_c = \sigma_{\mathcal{A}_{\mathrm{MC}}}(\alpha < 0.05) = \text{Gaussian standard deviation}\,(\mathcal{A}_{\mathrm{MC}}(\alpha < 0.05))\ . \tag{6.1}$$

The Gaussian is fitted in a core region of $\pm 2\,\sigma_c$ following the iterative procedure described in Section 4.4.3. With $n \in \mathbb{R}^+$, the start of the tail region is defined as $\mathcal{A}_{\text{tail}} = n \cdot \sigma_c$. In this thesis, two different inclusive tail regions are investigated:

- $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$, dominated by the Gaussian part of the distribution, and
- $\mathcal{A}_{\text{tail}} = 3\,\sigma_c$, dominated by the non-Gaussian part of the distribution.

The actual asymmetry distributions, however, are recorded in histograms with finite bin sizes. Therefore, an effective start value $\hat{\mathcal{A}}_{\text{tail}}$ is used, which corresponds to the lower boundary of the bin containing $\mathcal{A}_{\text{tail}}$. For convenience, the notation $\mathcal{A}_{\text{tail}} = n \cdot \sigma_c$ will still be used in the following to denote a specific tail region. The values of $\mathcal{A}_{\text{tail}}$ and $\hat{\mathcal{A}}_{\text{tail}}$ for the region with $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$ are compared in Fig. 6.3 in two $|\eta|$ intervals for illustration purposes. In Table B.2 in Appendix B.3, the values of $\hat{\mathcal{A}}_{\text{tail}}$ are listed in units of σ_c for all $|\eta| \times p_{\text{T}}^{\text{ave}}$ intervals. They differ by up to 10%.

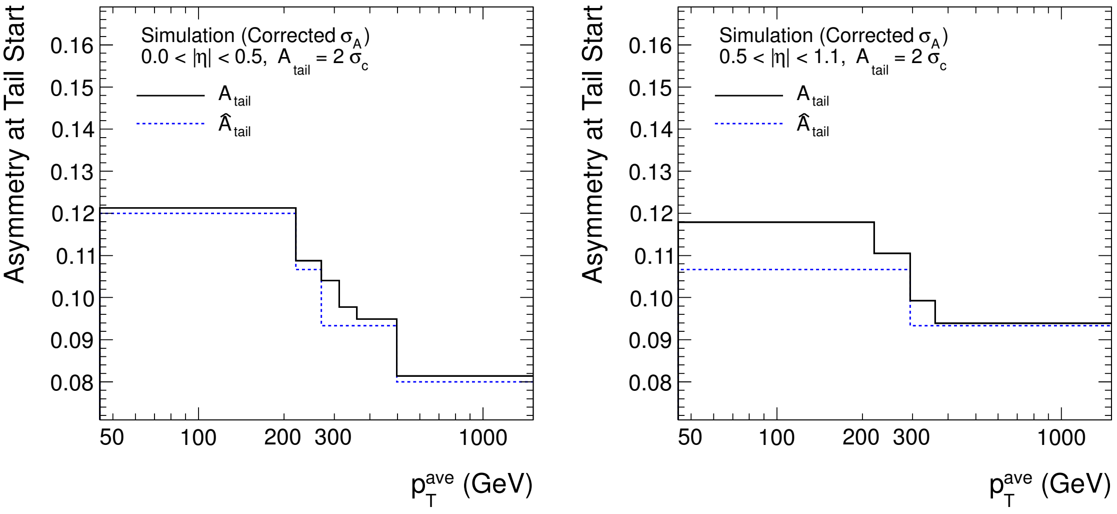


Figure 6.3: Asymmetry $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$ (*solid line*) and effective start value $\hat{\mathcal{A}}_{\text{tail}}$ (*dashed line*) of the corresponding tail region in different $|\eta| \times p_{\text{T}}^{\text{ave}}$ intervals. $\hat{\mathcal{A}}_{\text{tail}}$ is defined as the lower boundary of the histogram bin that contains $\mathcal{A}_{\text{tail}}$.

For the data distribution, the tail region starts at the same $\hat{\mathcal{A}}_{\text{tail}}$, i. e. the distribution is not fitted with a different Gaussian. In Fig. 6.4, examples of asymmetry distributions in data and simulation and their $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$ tail regions are shown. Further distributions can be found in Figs. B.4 and B.5 in Appendix B.3. They are compared to the expectation for a purely Gaussian distribution with the same core width $\sigma_{\mathcal{A}}$ as the simulated asymmetry distribution. Clearly, non-Gaussian components are present both in data and in simulation. The jets in these tail events often contain muons, indicating the occurrence of heavy-flavour quark decays. The non-Gaussian tails in data are slightly more pronounced than in the simulation. This will be discussed in more detail later on.

In data, a few outlier events are present with asymmetries larger than approximately 0.5. They have been individually investigated. No evidence for anomalous event signatures or jets clustered from detector noise that fake a dijet topology has been found. Furthermore, a

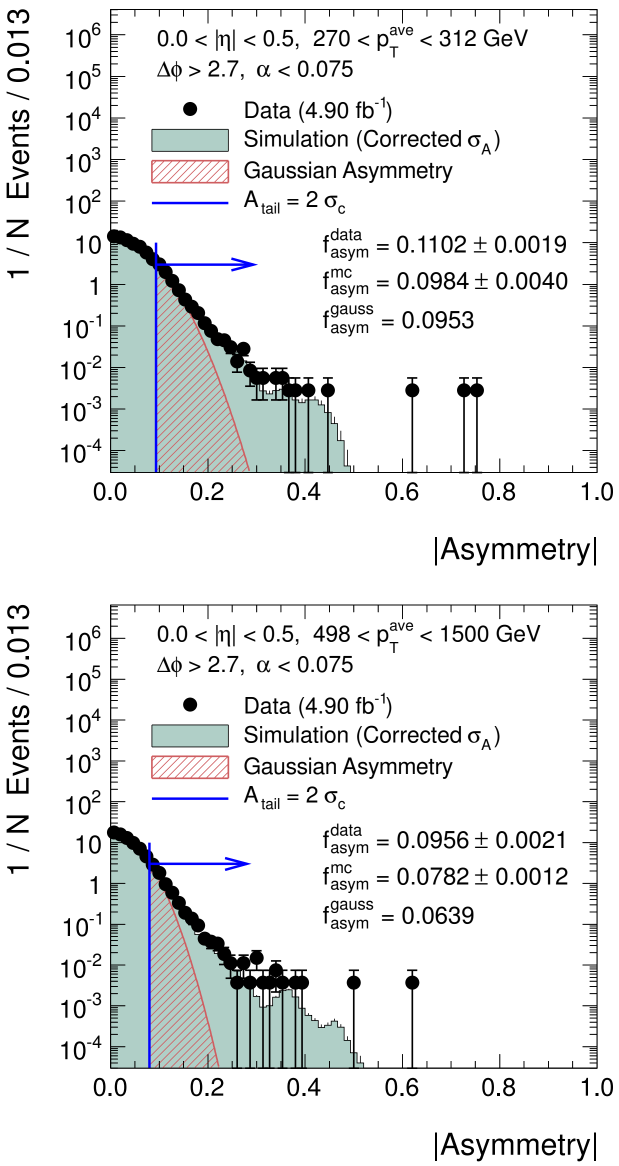


Figure 6.4: Asymmetry distributions in data (*solid circles*) and simulation (*shaded area*) in two different $p_{\mathrm{T}}^{\mathrm{ave}}$ intervals for $|\eta| < 0.5$ and $\alpha < 0.075$. The arrow marks the selected tail region $\mathcal{A}_{\mathrm{tail}} = 2\,\sigma_c$. The fractional numbers f_{asym} Eq. (6.2) of tail events are also indicated. For comparison, the contribution expected from a purely Gaussian asymmetry distribution with the same core width $\sigma_{\mathcal{A}}$ as the simulated asymmetry distribution is also shown (*hatched area*). (Due to the binning effects discussed in the text, $f_{\mathrm{asym}}^{\mathrm{gauss}}$ does not equal exactly the integral over a normal distribution from 2 standard deviations to infinity.)

situation where one of the two leading jets is completely lost due to malfunctioning detector components and a random, low-p_{T} jet is picked-up instead is highly unlikely because PF jets are typically reconstructed using the information from several subdetector systems such that redundancies are present. The events appear as proper QCD-dijet events where one jet fluctuated low and has a p_T of about $50 - 100\,\mathrm{GeV}$. Other contributions might arise from residual SM background processes mentioned in Section 6.1 such as $W \to \tau\nu + \mathrm{jets}$ events where the τ lepton decays hadronically. Candidates for $Z \to \nu\bar{\nu}$ + jets events, which would appear as monojet-like events, have not been observed. In addition, the presence of masked ECAL channels might be a source but this is unlikely, because the masked channels are included in the simulation, and would require the filter efficiency to be different in data than in the simulation. The detector signature of a typical outlier event is visualised in Fig. B.9 in Appendix B.6; it is the event with the largest asymmetry in Fig. 6.4 (*top*).

The size of the asymmetry tails is defined as the fractional number f_asym of events populating the tail region,

$$f_\mathrm{asym} = \frac{\text{Number of events in the tail region}}{\text{Number of events in the total distribution}} \,, \tag{6.2}$$

and determined for data ($f^\mathrm{data}_\mathrm{asym}$) and simulation ($f^\mathrm{mc}_\mathrm{asym}$). The statistical uncertainty δf_asym of f_asym is computed approximating the standard deviation of a binomial probability density function,

$$\delta f_\mathrm{asym} = \sqrt{\frac{f_\mathrm{asym}(1 - f_\mathrm{asym})}{\text{Number of events in the total distribution}}} \,.$$

In Fig. 6.4, the f_asym observed in data and simulation are compared to the expectation $f^\mathrm{gauss}_\mathrm{asym}$ for a Gaussian distribution, which has the same width $\sigma_{\mathcal{A}_\mathrm{MC}}$ as the core of the simulated asymmetry distribution. Naively, one might infer the value of $f^\mathrm{gauss}_\mathrm{asym}$ from the chosen start $\mathcal{A}_\mathrm{tail} = n \cdot \sigma_c$ of the tail region, e. g. $f^\mathrm{gauss}_\mathrm{asym} \approx 0.045$ for $n = 2$. However, it is important to recall

- that the jet-p_T resolution, and hence $\sigma_{\mathcal{A}_\mathrm{MC}}$, depends on α_max, Section 5.3.1,
- that $\mathcal{A}_\mathrm{tail}$ is always defined relative to $\sigma_c = \sigma_{\mathcal{A}_\mathrm{MC}}(\alpha_\mathrm{max} = 0.05)$ Eq. (6.1), and that
- the actual tail starts at $\hat{\mathcal{A}}_\mathrm{tail}$, which by construction is $\hat{\mathcal{A}}_\mathrm{tail} \leq \mathcal{A}_\mathrm{tail}$.

As a consequence, $f^\mathrm{gauss}_\mathrm{asym}$ depends on α_max,

$$\begin{aligned} f^\mathrm{gauss}_\mathrm{asym}(\alpha_\mathrm{max}) &= \frac{1}{\sqrt{\frac{\pi}{2}} \cdot \sigma_{\mathcal{A}_\mathrm{MC}}(\alpha_\mathrm{max})} \int_{\hat{\mathcal{A}}_\mathrm{tail}}^{\infty} \mathrm{d}\mathcal{A}\; \mathrm{e}^{-\frac{1}{2}\left(\frac{\mathcal{A}}{\sigma_{\mathcal{A}_\mathrm{MC}}(\alpha_\mathrm{max})}\right)^2} \\ &= 1 - \mathrm{erf}\left(\frac{\mathcal{A}_\mathrm{tail}}{\sqrt{2} \cdot \sigma_{\mathcal{A}_\mathrm{MC}}(\alpha_\mathrm{max})}\right) \,. \end{aligned} \tag{6.3}$$

Still, $f^\mathrm{gauss}_\mathrm{asym}$ provides the adequate reference to estimate the deviation of $\mathcal{A}_\mathrm{data}$ and $\mathcal{A}_\mathrm{MC}$ from a Gaussian distribution.

Due to the steep decrease of the asymmetry distributions in the tail region, cf. Fig. 6.4, f_asym is dominated by the entries close to the tail start $\mathcal{A}_\mathrm{tail}$. Hence, the defined tail regions with $\mathcal{A}_\mathrm{tail} = 2\,\sigma_c$ and $3\,\sigma_c$ are approximately exclusive with respect to f_asym, thus probing the region of transition from the Gaussian to the non-Gaussian part and the non-Gaussian part, respectively. Furthermore, f_asym is independent of the few outlier events at large asymmetries.

6.2.3 Correction For Additional Jet Activity

The transverse-momentum imbalance induced by the presence of additional jets that originate from the same hard scattering process as the leading two jets affects the measured size f_{asym} of the asymmetry tails as discussed in the previous Section 6.2.2. Thus, it is desirable to also compare $f_{\mathrm{asym}}^{\mathrm{data}}$ and $f_{\mathrm{asym}}^{\mathrm{mc}}$ at $\alpha_{\mathrm{max}} = 0$ in order to be independent of the simulation of radiative effects and achieve closer correspondence to the response-distribution tails. As before in Section 5.3.1, the measurement is therefore corrected by extrapolation to $\alpha_{\mathrm{max}} \to 0$. This is done separately for $f_{\mathrm{asym}}^{\mathrm{data}}$ and $f_{\mathrm{asym}}^{\mathrm{mc}}$.

In Fig. 6.5 and in Figs. B.6 to B.8 in Appendix B.4, f_{asym} is shown for different thresholds α_{max} on α both in data and simulation. As a reference, $f_{\mathrm{asym}}^{\mathrm{gauss}}$ expected for a Gaussian asymmetry distribution is also shown. f_{asym} increases for a looser dijet selection because $\mathcal{A}_{\mathrm{tail}}$ has been kept fixed for all α_{max}. However, the dependence of $f_{\mathrm{asym}}^{\mathrm{data}}$ and $f_{\mathrm{asym}}^{\mathrm{mc}}$ on α_{max} is different than that of $f_{\mathrm{asym}}^{\mathrm{gauss}}$, presumably due to the presence of non-Gaussian tail contributions. Hence, the extrapolation function cannot simply be deduced from Eq. (6.3) and the linear dependence of $\sigma_{\mathcal{A}}$ on α_{max}. Instead, an exponential fit function $f_{\mathrm{asym}}(\alpha_{\mathrm{max}})$ is empirically chosen, which describes the measurements adequately and provides a robust extrapolation behaviour; for example, it does not depend strongly on fluctuations of the first point. As before, the y-axis intercept $f_{\mathrm{asym}}(\alpha_{\mathrm{max}} \to 0)$ is used in the further analysis.

The different measurements of f_{asym} are correlated, as in case of the resolution measurement Section 5.3.1, because events have been selected by applying upper thresholds on α. As before, however, the exponential fit is performed neglecting the correlation. In analogy to Eq. (5.17), the obtained uncertainty on the value of the exponential function at $\alpha_{\mathrm{max}} = 0$ defines the *extrapolation uncertainty*

$$\delta f_{\mathrm{ex}} = \text{uncertainty on } f_{\mathrm{asym}}(\alpha_{\mathrm{max}} = 0) \;, \tag{6.4}$$

which will be cited in the following. Again, δf_{ex} depends the statistical uncertainty of the measurement but is not expected to fully cover the statistical fluctuations of the extrapolated value. This will be compensated for by assigning a rather conservative systematic uncertainty as explained in Section 6.4.

The extrapolation procedure is validated using a toy asymmetry-distribution. In contrast to the previous toy study Section 5.2.6.1, here, the jet-p_{T} response is generated from the MC-truth response distributions Section 4.4.1. The latter are by construction not affected by the presence of additional jets. Hence, the generated toy asymmetry corresponds to the asymmetry expected for ideal dijet events of the full PYTHIA simulation. Therefore, the extrapolated $f_{\mathrm{asym}}^{\mathrm{mc}}(\alpha_{\mathrm{max}} \to 0)$ should correspond to $f_{\mathrm{asym}}^{\mathrm{toy}}$, the fractional number of events in the tail region of the toy asymmetry-distribution.

The MC-truth response distribution is determined as described in Section 4.4.1. In addition, a dijet selection is performed at generator level by requiring $\Delta\phi^{\mathrm{gen}} > 2.7$ and $\alpha^{\mathrm{gen}} < 0.05$ in order to account for potential differences of the response tails between QCD-dijet and -multijet events due to the different flavour composition. Furthermore, the events are weighted according to the QCD spectrum and the effective prescales of the applied trigger paths following the procedures described in Sections 5.1.2 and 6.1,

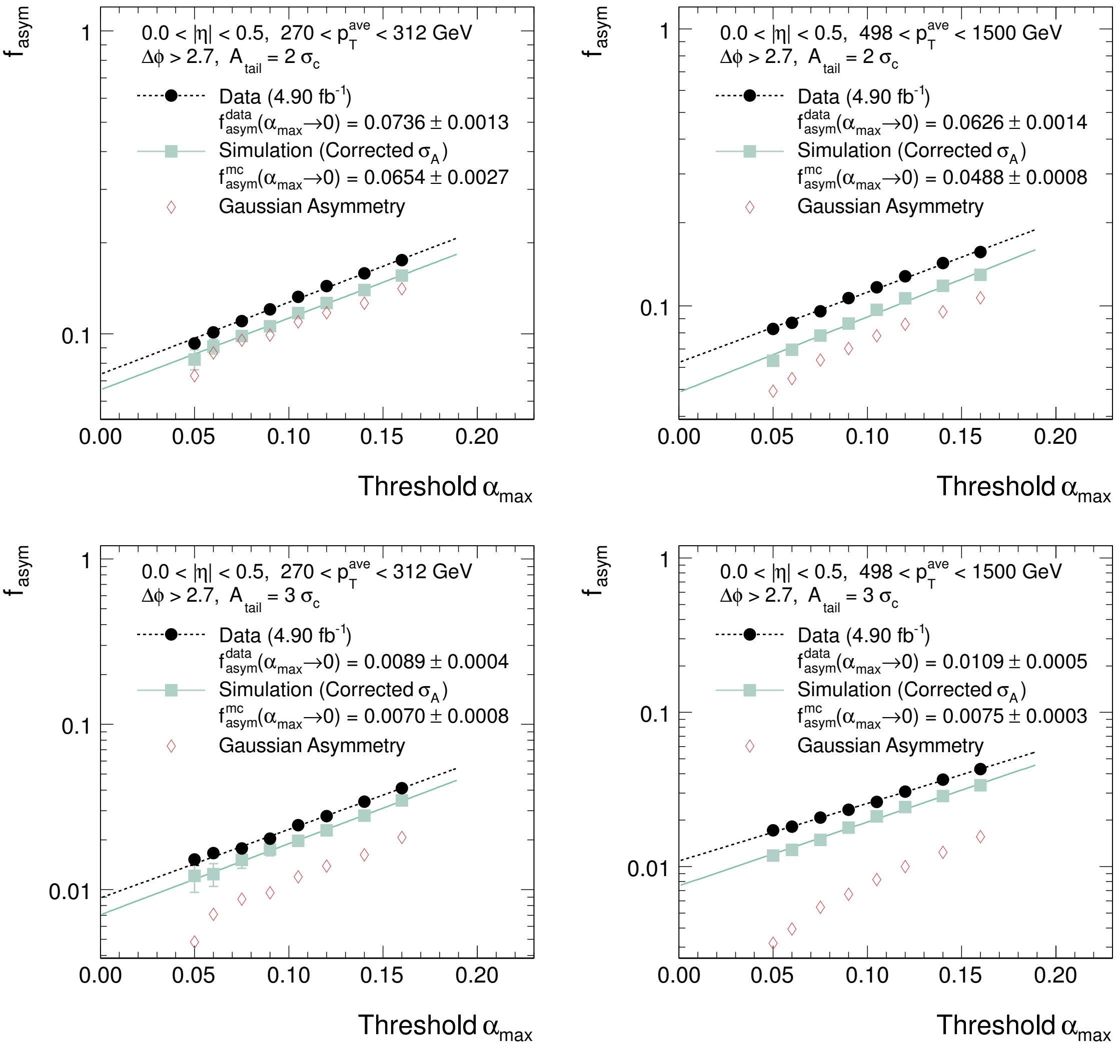


Figure 6.5: Dependence of f_{asym}, the fractional number of tail events, on the threshold α_{max} of the fractional third-jet p_{T} in data (*solid circles*) and simulation (*solid squares*). The measured values are fitted with an exponential function (*lines*) to extrapolate to zero additional jet activity. For comparison, the f_{asym} expected for a Gaussian distribution are overlaid (*open diamonds*). Two different $|\eta| \times p_{\text{T}}^{\text{ave}}$ intervals are shown for the $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$ (*top*) and $3\,\sigma_c$ (*bottom*) tail region.

respectively. Per $|\eta| \times p_{\text{T}}^{\text{ave}}$ interval, 10^6 asymmetry entries

$$\mathcal{A}_{\text{toy}} = \frac{r_1 - r_2}{r_1 + r_2} \tag{6.5}$$

are generated, where the random numbers $\{r_i \,|\, i = 1, 2\}$ are sampled from the MC-truth response distribution, and $f_{\text{asym}}^{\text{toy}}$ is determined.

In Fig. 6.6, the extrapolated $f_{\text{asym}}^{\text{mc}}(\alpha_{\text{max}} \to 0)$ is compared to $f_{\text{asym}}^{\text{toy}}$ in two example intervals. Here and in general, there is agreement within 30%. Residual deviations are expected to equally affect $f_{\mathrm{asym}}^{\mathrm{data}}$ and $f_{\mathrm{asym}}^{\mathrm{mc}}$ and hence to cancel in the final data-to-simulation ratio, cf. Section 6.2.4. Nonetheless, a corresponding systematic uncertainty is assigned as described in Section 6.4.

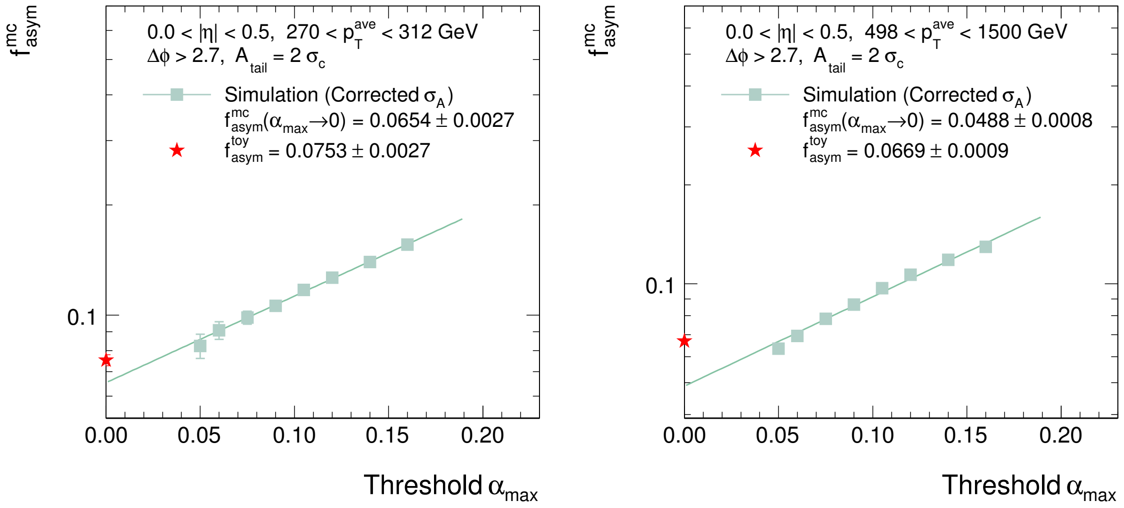


Figure 6.6: Comparison of $f_{\mathrm{asym}}^{\mathrm{mc}}(\alpha_{\mathrm{max}} \to 0)$ obtained from the extrapolation procedure (*line*) and $f_{\mathrm{asym}}^{\mathrm{toy}}$ (*star*), the fractional number of tail events of a toy asymmetry-distribution. The latter has been been generated from the MC-truth response and is by construction not affected by the presence of additional jet activity.

Jet-clustering effects are an additional source of transverse-momentum imbalance at particle level (PLI). As discussed in Section 5.3.2, the resolution measurement must be corrected for this effect. In case of the tail measurement, where the relative difference between $\mathcal{A}_{\mathrm{data}}$ and $\mathcal{A}_{\mathrm{MC}}$ is examined, no correction is applied since the effect is expected to cancel.

6.2.4 Data-to-Simulation Ratio

Events have been selected in intervals of $p_{\mathrm{T}}^{\mathrm{ave}}$. Therefore, migration effects occur because of the finite jet-p_{T} resolution in combination with the falling $p_{\mathrm{T}}^{\mathrm{ave}}$ spectrum as explained in Section 4.4.1. In particular in the tails, where the response of one or both jets fluctuates extremely, events have a high chance of being allocated into a wrong $p_{\mathrm{T}}^{\mathrm{ave}}$ interval[3], and hence, f_{asym} is biased.

The agreement between the asymmetry tails in data and simulation is therefore expressed as ratio ρ_{tail} of $f_{\mathrm{asym}}^{\mathrm{data}}$ and $f_{\mathrm{asym}}^{\mathrm{mc}}$ such that the selection effects, which occur equally in data and simulation, are expected to cancel. Likewise, any residual impact from pile-up on f_{asym} cancels as so far it is correctly modelled in the simulation. The overall sensitivity of the asymmetry in general and f_{asym} in particular on pile-up is small anyway, as will be shown below in Sections 6.3 and 6.5.1, respectively. Furthermore, ρ_{tail} is computed after

---

[3] To be precise, the events are allocated into the wrong $p_{\mathrm{T}}^{\mathrm{gen,ave}}$ interval.

extrapolation of $f_{\mathrm{asym}}^{\mathrm{data}}$ and $f_{\mathrm{asym}}^{\mathrm{mc}}$ to zero additional jet activity in order to minimise the impact from QCD radiation, i. e.

$$\rho_{\mathrm{tail}} = \frac{f_{\mathrm{asym}}^{\mathrm{data}}(\alpha_{\mathrm{max}} \to 0)}{f_{\mathrm{asym}}^{\mathrm{mc}}(\alpha_{\mathrm{max}} \to 0)} \,. \tag{6.6}$$

As argued, ρ_{tail} is expected to be independent of selection biases, pile-up, and radiative effects. Hence, it depends only on the ratio of the fractional number of events in the corresponding tails of the jet-p_{T} response distributions. This relation will be investigated later in Section 6.6. Values of ρ_{tail} different from unity imply mismodelled jet-p_{T} response tails.

6.3 Dependence of the Asymmetry on Pile-Up

The dependence of the asymmetry on the presence of pile-up collisions has been investigated[4] by determining the asymmetry's width $\sigma_{\mathcal{A}}$ as a function of the number of reconstructed primary vertices, N_{Vtx}, which is correlated with the number of pile-up interactions. In Fig. 6.7, $\sigma_{\mathcal{A}}$ is compared for dijet events with $N_{\mathrm{Vtx}} \leq 5$ and with $N_{\mathrm{Vtx}} \geq 9$ in different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals.

A loose dijet selection $\alpha < 0.25$ and broad $|\eta|$ intervals have been chosen in order to increase the statistical precision at very low $p_{\mathrm{T}}^{\mathrm{ave}}$, where the triggers are highly prescaled. The statistical uncertainty of $\sigma_{\mathcal{A}}$ is less than 5% above $p_{\mathrm{T}}^{\mathrm{ave}} = 50\,\mathrm{GeV}$ for $|\eta| < 2.3$. Within these uncertainties, there is agreement between the low and the high pile-up sample.

After application of the offset jet energy correction, which removes the additional energy depositions from pile-up events on average, only local fluctuations of the pile-up energy density can impact the jet measurement, which affects for example the MC-truth resolution, cf. Fig. 4.24. The dijet asymmetry is apparently less sensitive by construction because additional energy depositions cancel to some extent in the ratio.

6.4 Systematic Uncertainties

The following sources of systematic uncertainties of the measurement of the data-to-simulation ratio ρ_{tail} have been considered:

Correction for resolution differences ('$\sigma_{\mathcal{A}}$ Correction'): The simulated asymmetry distributions have been corrected to account for the observed difference of the jet-p_{T} resolution between data and simulation. The measurement is repeated with ρ_{res} in Eq. (5.21) varied by its uncertainty listed in Table 5.4, and the average upper and lower variation of ρ_{tail} is assigned as uncertainty.

Pile-up modelling ('PU'): The asymmetry is only mildly affected by the presence of pile-up collisions because the impact from local density fluctuations of the additionally distributed energy mostly cancels, cf. Section 6.3. Furthermore, any residual impact on f_{asym} is expected to cancel in the ratio ρ_{tail} provided the pile-up is modelled

---

[4] In the previous chapter, it has been shown that the impact of pile-up on the result of the maximum-likelihood fit is small. However, since the dijet pdf Eq. (5.11) is sensitive to the absolute difference $p_{\mathrm{T},1} - p_{\mathrm{T},2}$, here, the impact is explicitly investigated for the asymmetry.

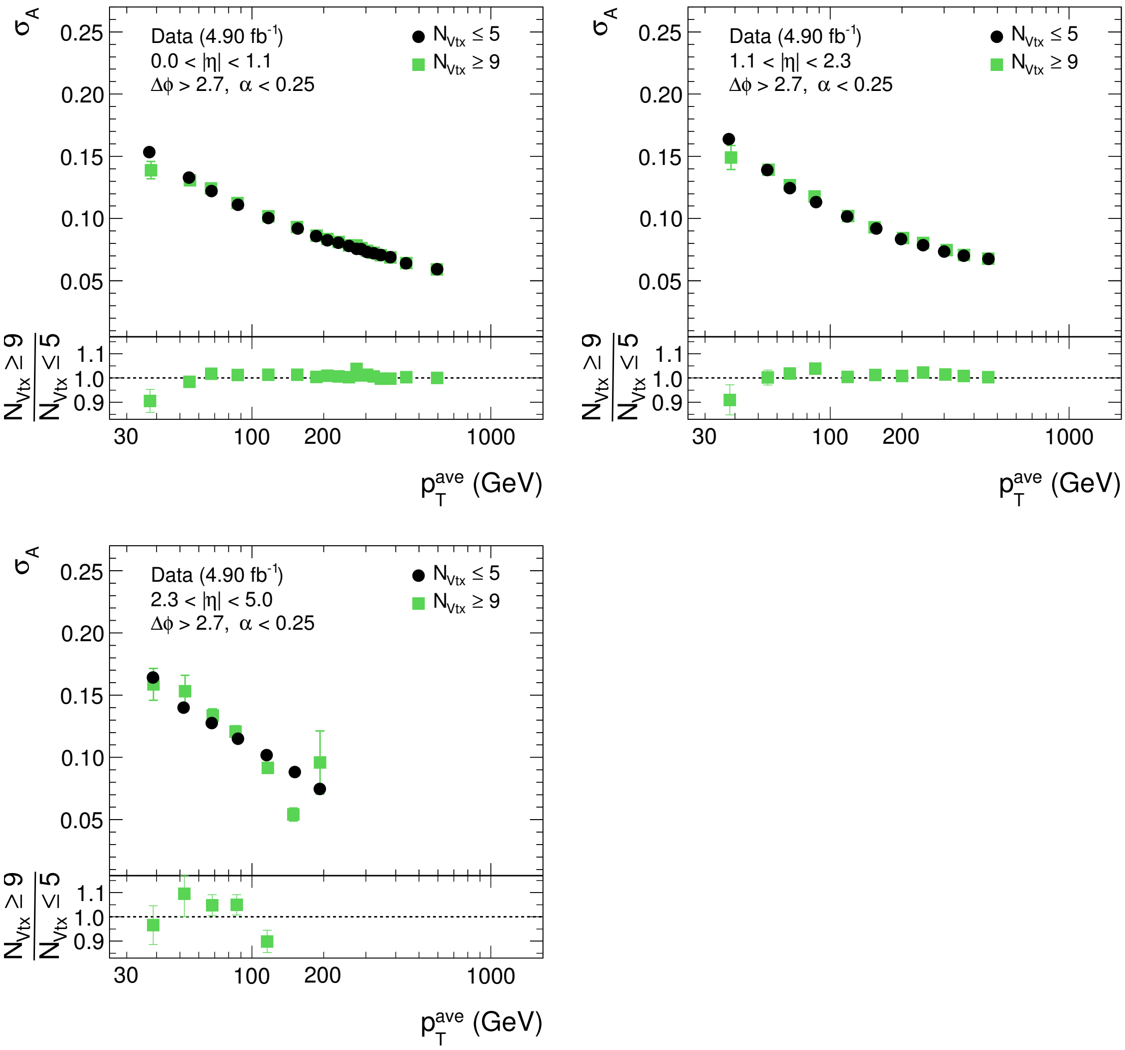


Figure 6.7: Gaussian width $\sigma_{\mathcal{A}}$ of the dijet asymmetry as a function of $p_{\mathrm{T}}^{\mathrm{ave}}$ for $\alpha < 0.25$ in different $|\eta|$ intervals for events with $N_{\mathrm{Vtx}} \leq 5$ (*circles*) and with $N_{\mathrm{Vtx}} \geq 9$ (*squares*). The bottom part of each pad displays the ratio of $\sigma_{\mathcal{A}}$ for the high and the low N_{Vtx} selection.

correctly. In order to quantify the effect of imprecise pile-up modelling, the measurement is repeated with the mean of the pile-up multiplicity distribution shifted by ±1, and the average upper and lower variation of ρ_{tail} is assumed as uncertainty.

Correction for additional jet activity ('Additional Jets'): The transverse-momentum imbalance induced by the presence of additional jets has been corrected for by an extrapolation method. An empirically found, exponential fit-function has been assumed to adequately describe the scaling of f_{asym} with $\alpha_{\max}$ down to $\alpha_{\max} = 0$. In order to quantify possible uncertainties introduced by this choice, the measurement of ρ_{tail} has been repeated using the fit function's value at $\alpha_{\max} = 0.05$ instead of the extrapolated $f_{\mathrm{asym}}(\alpha_{\max} \to 0)$. The full difference between the two results is assigned as symmetric uncertainty. It is assumed to also cover the underestimation of the statistical uncertainties by the extrapolation uncertainty δf_{ex} Eq. (6.4).

Intrinsic bias ('Bias'): The extrapolated $f_{\mathrm{asym}}(\alpha_{\mathrm{max}} \to 0)$ have been compared to $f^{\mathrm{toy}}_{\mathrm{asym}}$, the fractional number of events in a toy asymmetry-distribution generated from the MC-truth response. Although any bias is expected to occur equally in data and simulation and therefore to cancel, an uncertainty is assigned measuring ρ_{tail} using $f^{\mathrm{toy}}_{\mathrm{asym}}$ instead of $f^{\mathrm{mc}}_{\mathrm{asym}}(\alpha_{\mathrm{max}} \to 0)$ in Eq. (6.6) and taking 50% of the difference to the nominal measurement as symmetric uncertainty.

The relative size of the systematic uncertainties is shown in Figs. 6.8 and 6.9 for the tail regions $\mathcal{A}_{\mathrm{tail}} = 2\,\sigma_c$ and $3\,\sigma_c$, respectively. The by far dominant contribution arises from the propagation of the uncertainty on ρ_{res} Table 5.4 due to the correction of the simulation for the larger jet-p_{T} resolution in data. The sources of the different contributions are considered uncorrelated, and thus, the total uncertainty is computed as their quadratic sum. It amounts to approximately 20 – 25% in the central detector region for $\mathcal{A}_{\mathrm{tail}} = 2\,\sigma_c$, and increases to approximately 50 – 60% in the forward region for $\mathcal{A}_{\mathrm{tail}} = 3\,\sigma_c$. In the forward region and also in the fine $p^{\mathrm{ave}}_{\mathrm{T}}$ intervals in the central detector region, the determined uncertainties are partly affected by statistical fluctuations of the results when performing the systematic variations described above because of the small number of selected events.

The jet energy scale is not taken into account as a source of uncertainty for this measurement. Due to the dijet selection, the leading two jets have similar $|\eta|$ and p_{T}. Hence, their calibration factors are similar, and the dijet asymmetry Eq. (4.9) is to first approximation independent of the jet energy scale.

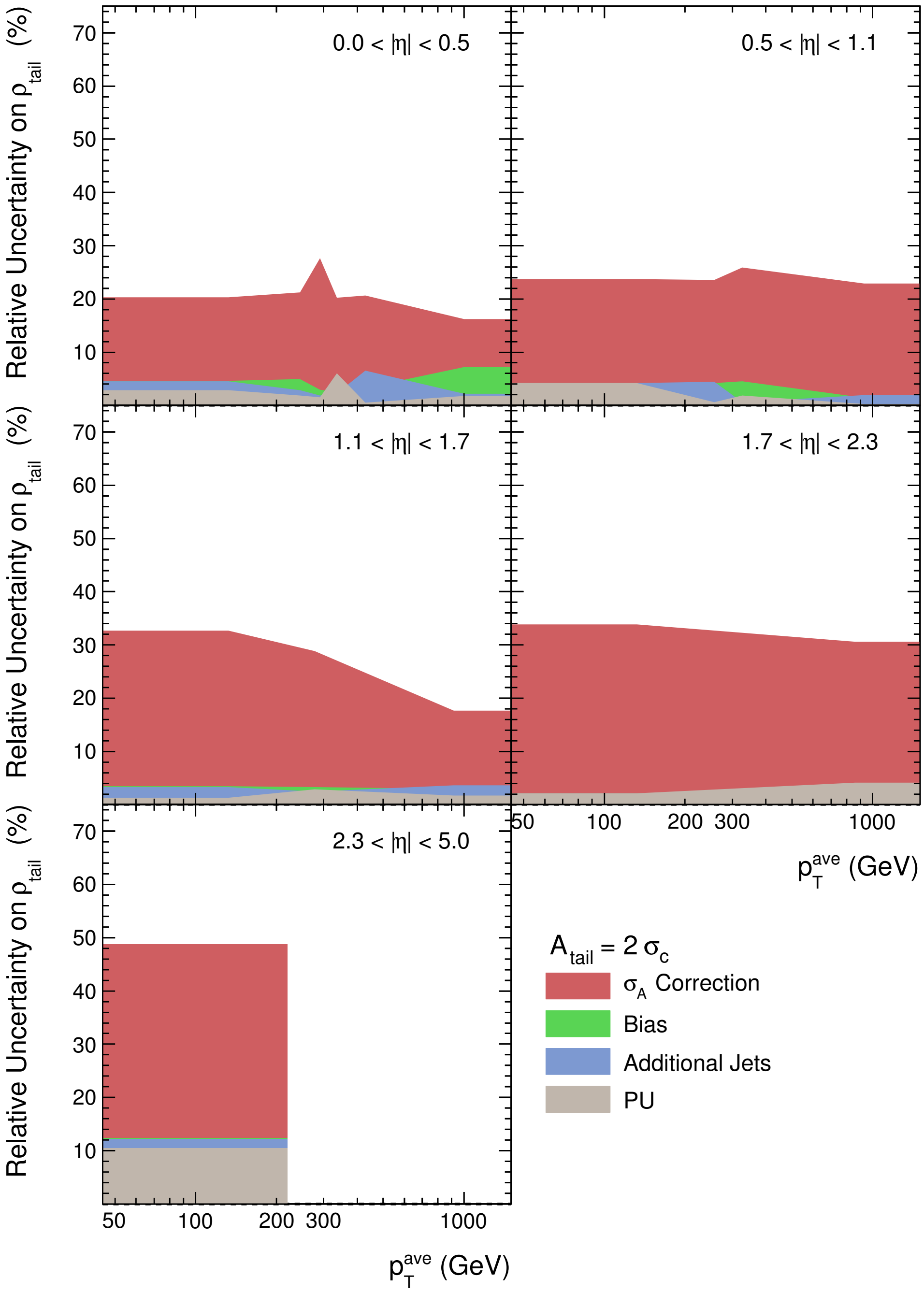


Figure 6.8: Relative systematic uncertainties of the ρ_{tail} measurement as a function of $p_{\mathrm{T}}^{\text{ave}}$ in different $|\eta|$ intervals for the tail region $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$. The uncertainties are added in quadrature ('stack') and are linearly extrapolated between the $p_{\mathrm{T}}^{\text{ave}}$ intervals.

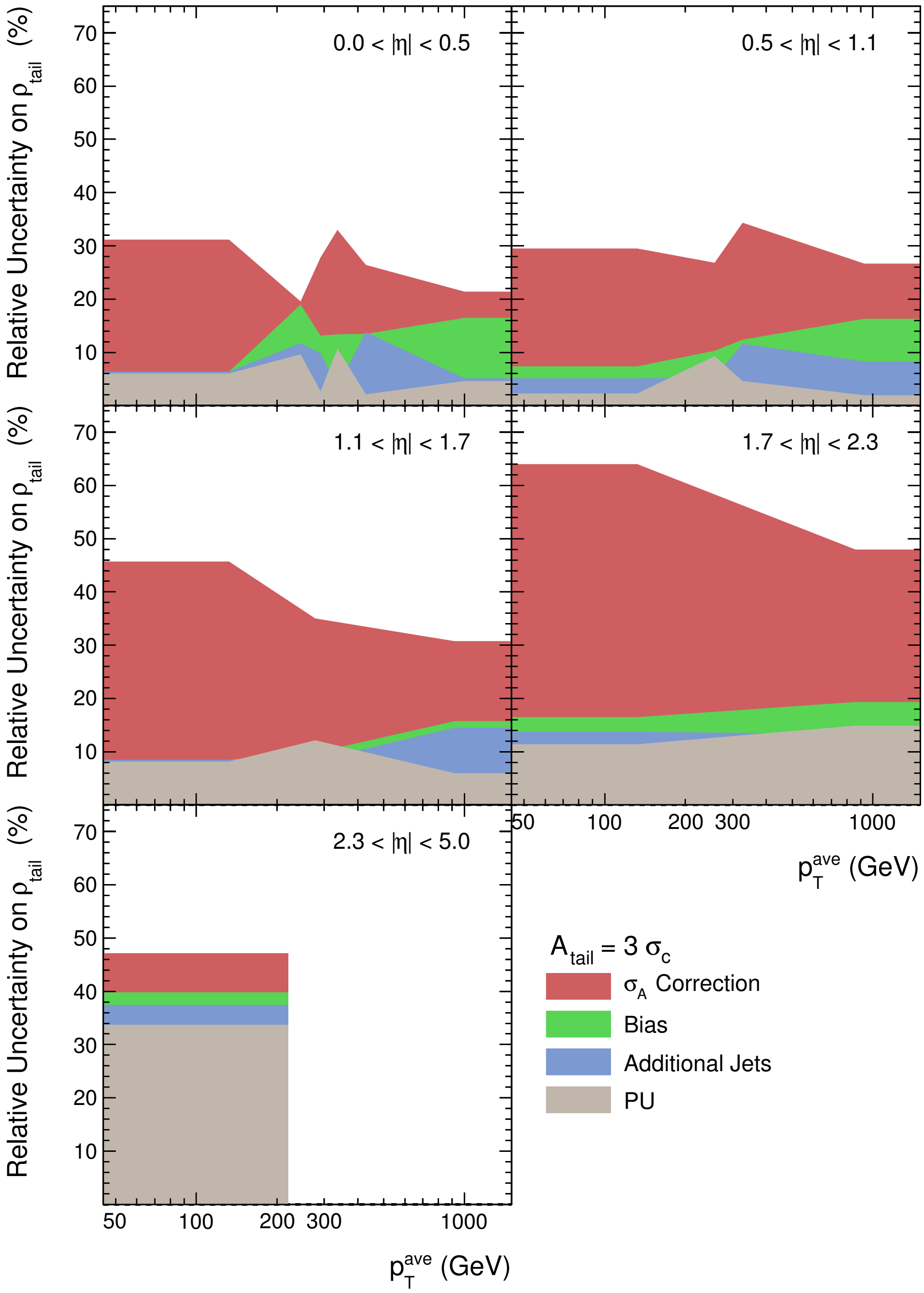


Figure 6.9: Relative systematic uncertainties of the ρ_{tail} measurement as a function of $p_{\mathrm{T}}^{\mathrm{ave}}$ in different $|\eta|$ intervals for the tail region $\mathcal{A}_{\mathrm{tail}} = 3\,\sigma_c$. The uncertainties are added in quadrature ('stack') and are linearly extrapolated between the $p_{\mathrm{T}}^{\mathrm{ave}}$ intervals.

6.5 Results

The fractional size f_{asym} of the dijet-asymmetry tails has been measured in $4.90\,\mathrm{fb}^{-1}$ of dijet data and in the Pythia simulation, where distributions have been corrected for the larger jet-p_{T} resolution in data. In Section 6.5.1, the values of f_{asym} for a particular dijet selection $\alpha < 0.075$ are presented as a function of $p_{\mathrm{T}}^{\mathrm{ave}}$, $|\eta|$, and the tail region $\mathcal{A}_{\mathrm{tail}}$. They are compared to the expectation from a Gaussian asymmetry. The ratios ρ_{tail} of the f_{asym} in data and simulation after extrapolation to zero additional jet activity are shown in Section 6.5.2. In both cases, the dependence of the results on the presence of pile-up is investigated.

6.5.1 Asymmetry Tails

The f_{asym} in data and simulation for $\alpha < 0.075$ are compared in Figs. 6.10 and 6.11 for $\mathcal{A}_{\mathrm{tail}} = 2\,\sigma_c$ and $3\,\sigma_c$, respectively (the underlying asymmetry distributions can be found in Figs. B.4 and B.5 in Appendix B.3). The shown uncertainties are the statistical only. Their increase with $|\eta|$ and $\mathcal{A}_{\mathrm{tail}}$ reflects the rapidity dependence of the jet-production cross-section Fig. 4.4 and the decrease of the number of events with increasing $|\mathcal{A}|$, cf. Fig. 6.3, respectively.

As a reference, the $f_{\mathrm{asym}}^{\mathrm{gauss}}$ of a Gaussian distribution with the same core width as the simulated asymmetry distribution are also depicted. Their size varies due to the binning effects and the related difference between $\mathcal{A}_{\mathrm{tail}}$ and $\hat{\mathcal{A}}_{\mathrm{tail}}$ discussed above in Section 6.2.2. The variations of $f_{\mathrm{asym}}^{\mathrm{mc}}$ and $f_{\mathrm{asym}}^{\mathrm{gauss}}$ are correlated, of course, because σ_c Eq. (6.1) has been determined from a fit to the simulated asymmetry. To be able to compare $f_{\mathrm{asym}}^{\mathrm{data}}$ and $f_{\mathrm{asym}}^{\mathrm{mc}}$ independently of binning effects, their values are also shown relative to $f_{\mathrm{asym}}^{\mathrm{gauss}}$.

Evidently, non-Gaussian tail components are present both in data and simulation. The fractional number of tail events is larger than expected for a Gaussian asymmetry distribution, the difference being roughly $20 - 50\%$ for $\mathcal{A}_{\mathrm{tail}} = 2\,\sigma_c$ and $50 - 300\%$ for $\mathcal{A}_{\mathrm{tail}} = 3\,\sigma_c$. In general, the size of the non-Gaussian component increases for larger p_{T}. Furthermore, the size of the tails in data is larger than simulated, in particular at high p_{T} by approximately $30 - 50\%$, as will be further investigated in the following section.

The dependence of the asymmetry tails on the number of pile-up events is studied by determining f_{asym} as a function of the number N_{Vtx} of reconstructed primary vertices. In Figs. 6.12 to 6.15, f_{asym} in the tail region with $\mathcal{A}_{\mathrm{tail}} = 3\,\sigma_c$ is compared for two exclusive sets of events[5] with $N_{\mathrm{Vtx}} \leq 6$ and with $N_{\mathrm{Vtx}} \geq 7$. No systematic trend of f_{asym} with N_{Vtx} is visible, and the observed differences are consistent with the expected statistical fluctuations.

This is not unexpected from the observations in Section 6.3 because there is only a minor sensitivity of the asymmetry distribution's core to the additionally distributed energy from pile-up collisions. In conclusion, the presence of pile-up also has negligible impact on the measurement of non-Gaussian asymmetry tails.

---

[5] To be independent of binning effects, σ_c has been determined for the inclusive selection, and the same value has been used in both cases. This can be done safely because the core width of the asymmetry distribution does not depend on N_{Vtx} as shown in Fig. 6.7.

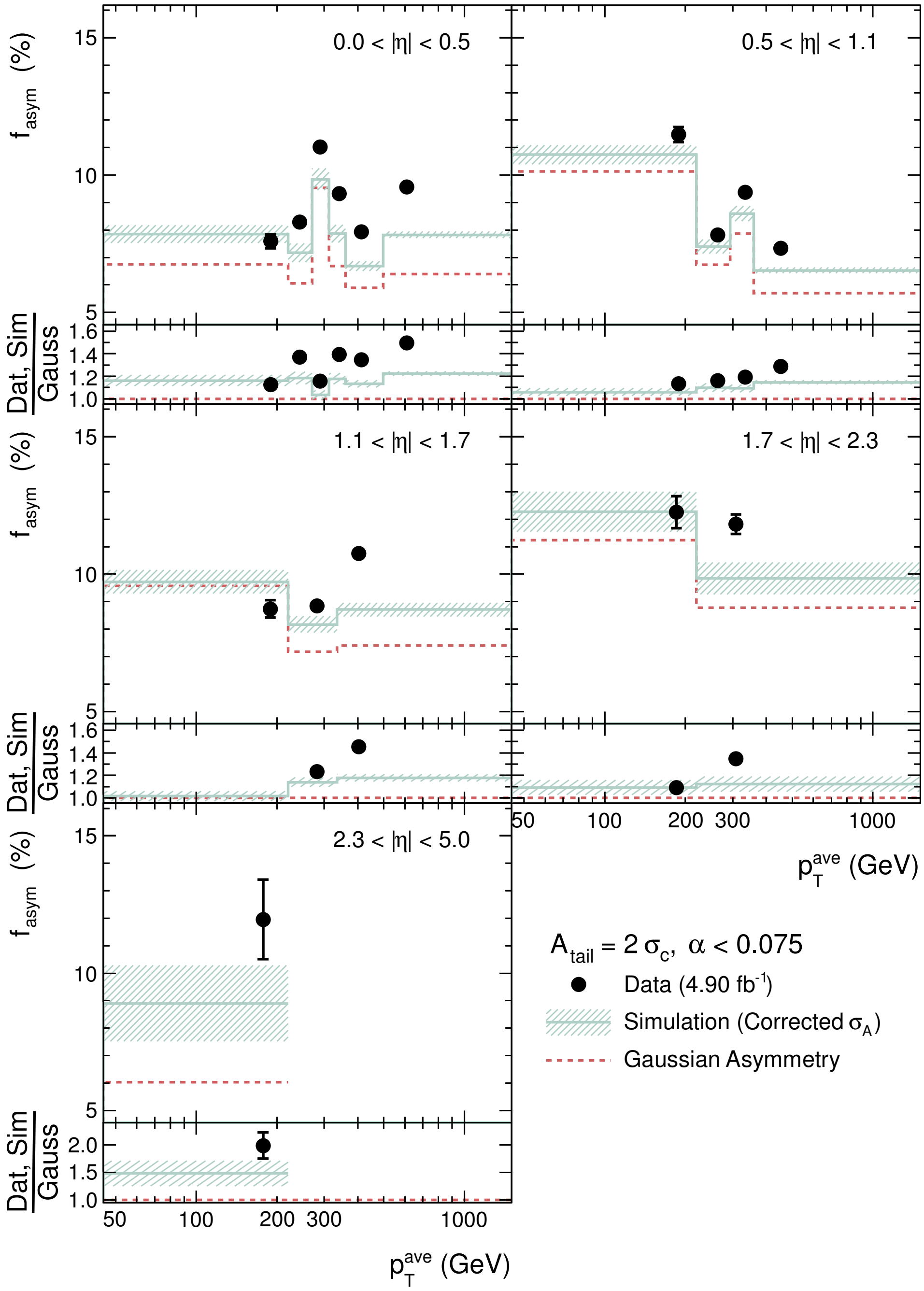


Figure 6.10: Fractional tail-size f_{asym} in data (*circles*), simulation (*solid line*) and for a Gaussian (*dashed line*) in different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals for $\alpha < 0.075$ in the region $\mathcal{A}_{\mathrm{tail}} = 2\,\sigma_c$. The horizontal position of the data points corresponds to the mean $p_{\mathrm{T}}^{\mathrm{ave}}$ in the respective interval. Errors are statistical only. Bottom pads display the ratio of f_{asym} to the Gaussian case. (The absolute size of the Gaussian tails varies due to the discussed binning effects.)

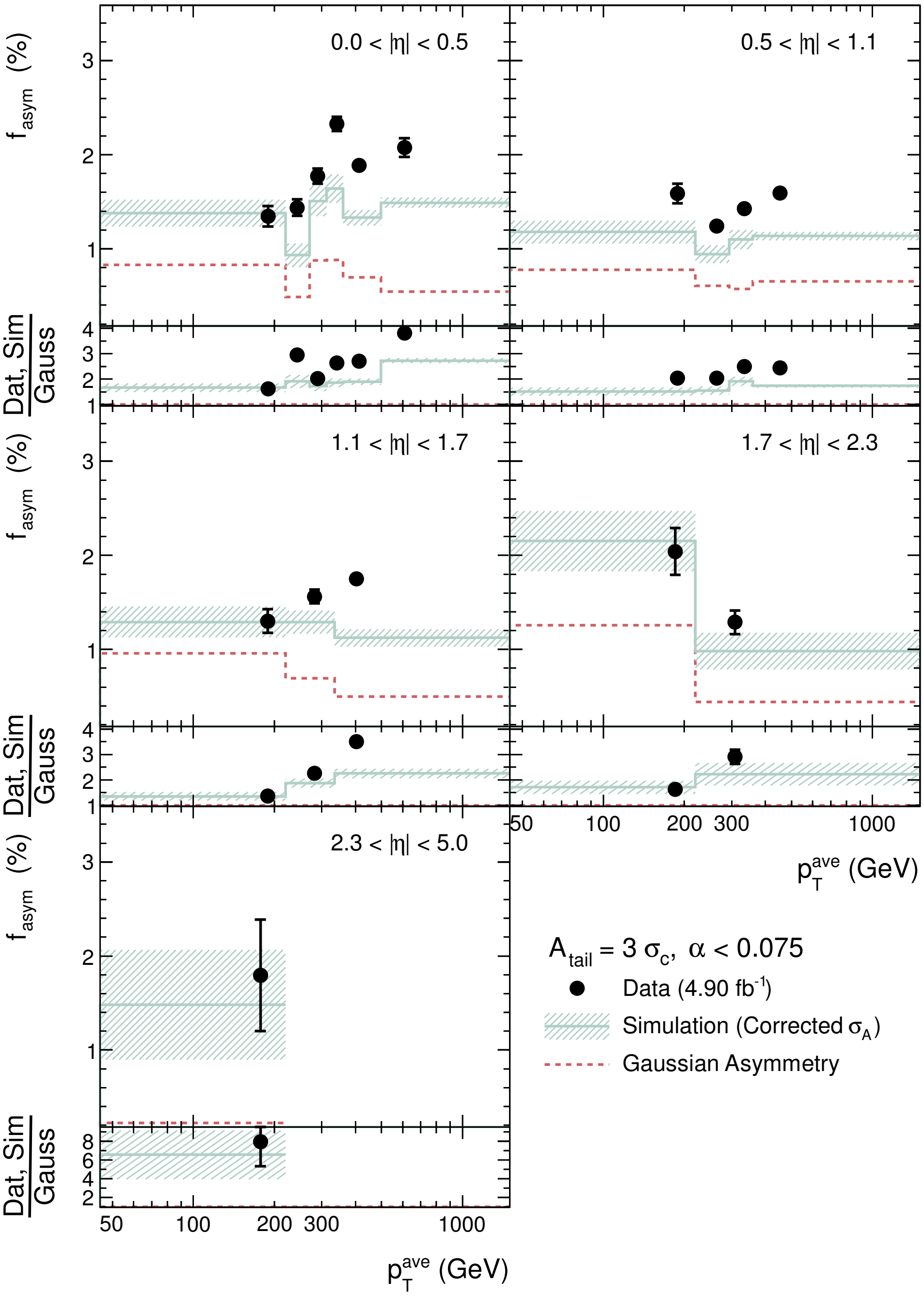


Figure 6.11: Fractional tail-size f_{asym} in data (*circles*), simulation (*solid line*) and for a Gaussian (*dashed line*) in different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals for $\alpha < 0.075$ in the region $\mathcal{A}_{\mathrm{tail}} = 3\,\sigma_c$. The horizontal position of the data points corresponds to the mean $p_{\mathrm{T}}^{\mathrm{ave}}$ in the respective interval. Errors are statistical only. Bottom pads display the ratio of f_{asym} to the Gaussian case. (The absolute size of the Gaussian tails varies due to the discussed binning effects.)

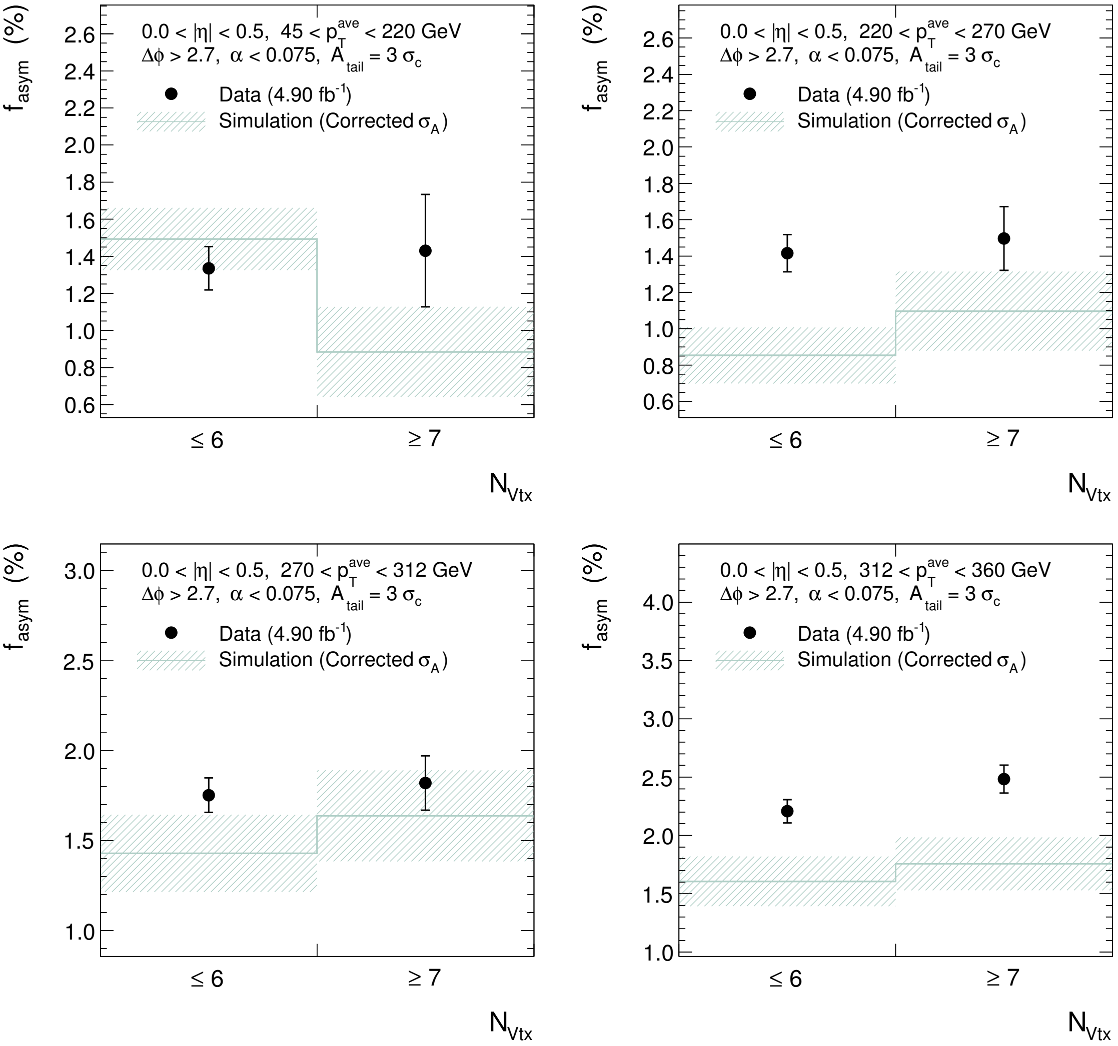


Figure 6.12: Dependence of f_{asym} on the number N_{Vtx} of primary vertices in data (*circles*) and simulation (*solid line*) in different $|\eta| \times p_{\text{T}}^{\text{ave}}$ intervals for $\alpha < 0.075$ and for $\mathcal{A}_{\text{tail}} = 3\,\sigma_c$. The same value of σ_c has been used for both N_{Vtx} selections to avoid fluctuations due to binning effects. Error bars and bands represent the statistical uncertainties. Note that the y-axis scale differs. (Continued in Fig. 6.13.)

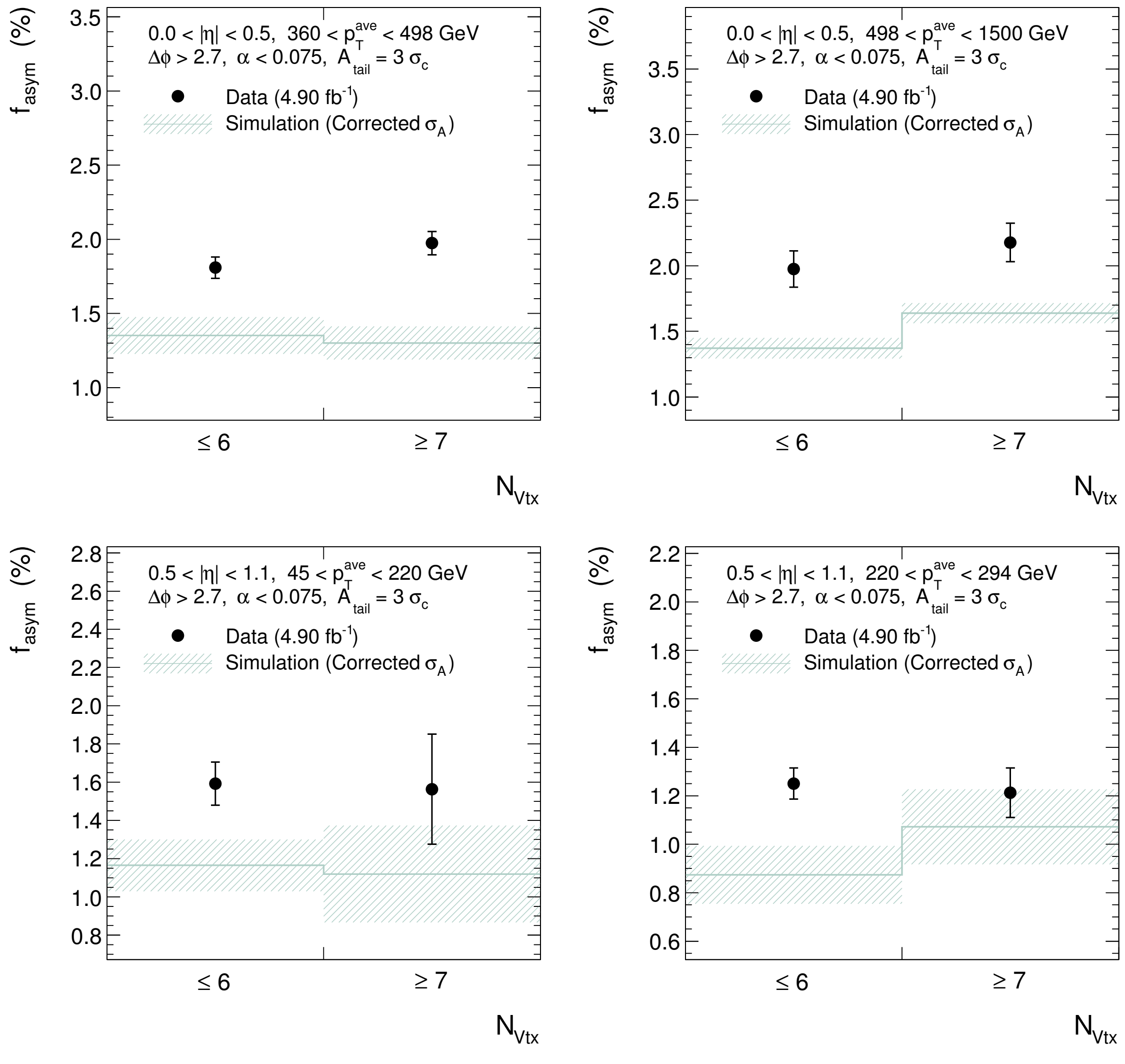


Figure 6.13: Continued from Fig. 6.12 and in Fig. 6.14.

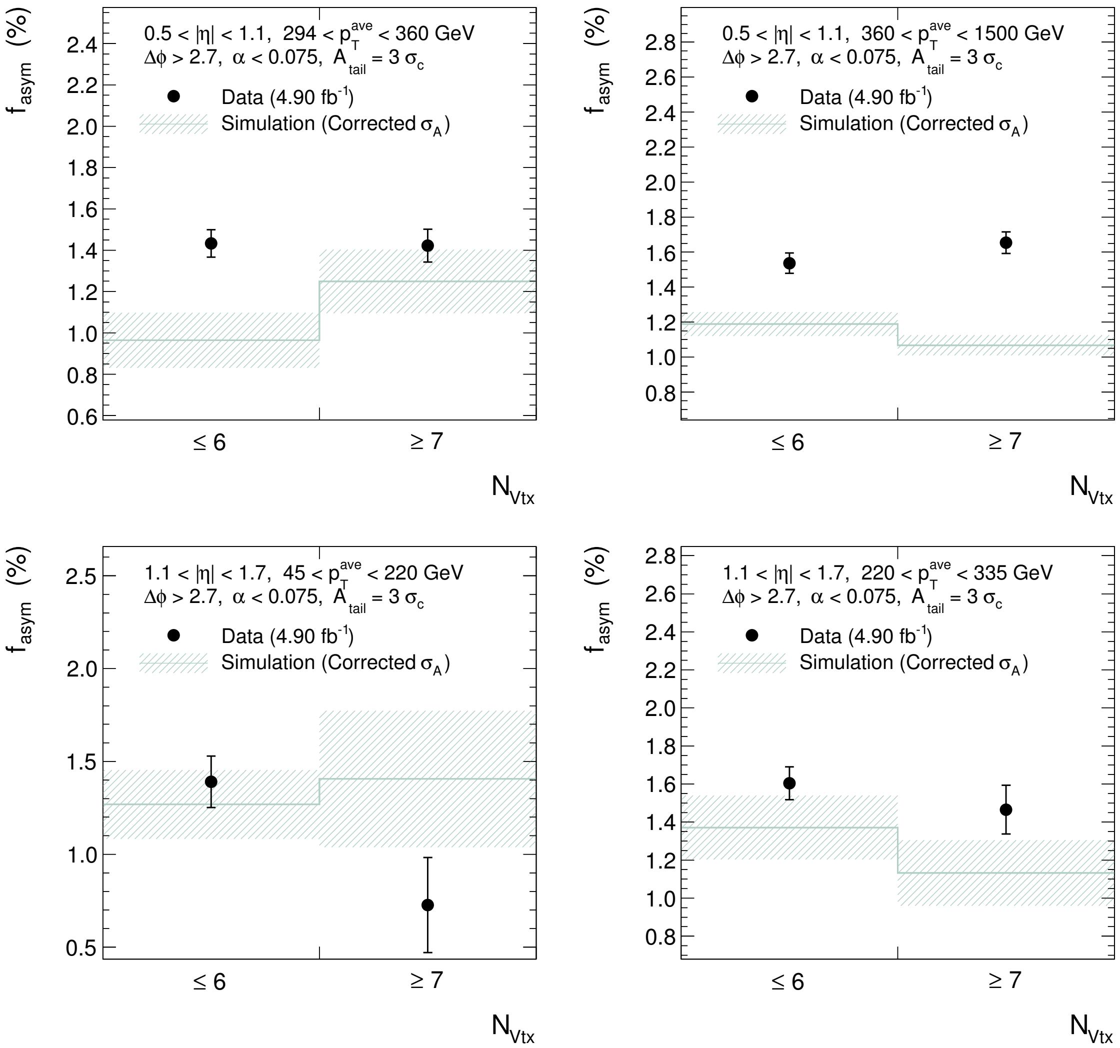


Figure 6.14: Continued from Fig. 6.13 and in Fig. 6.15.

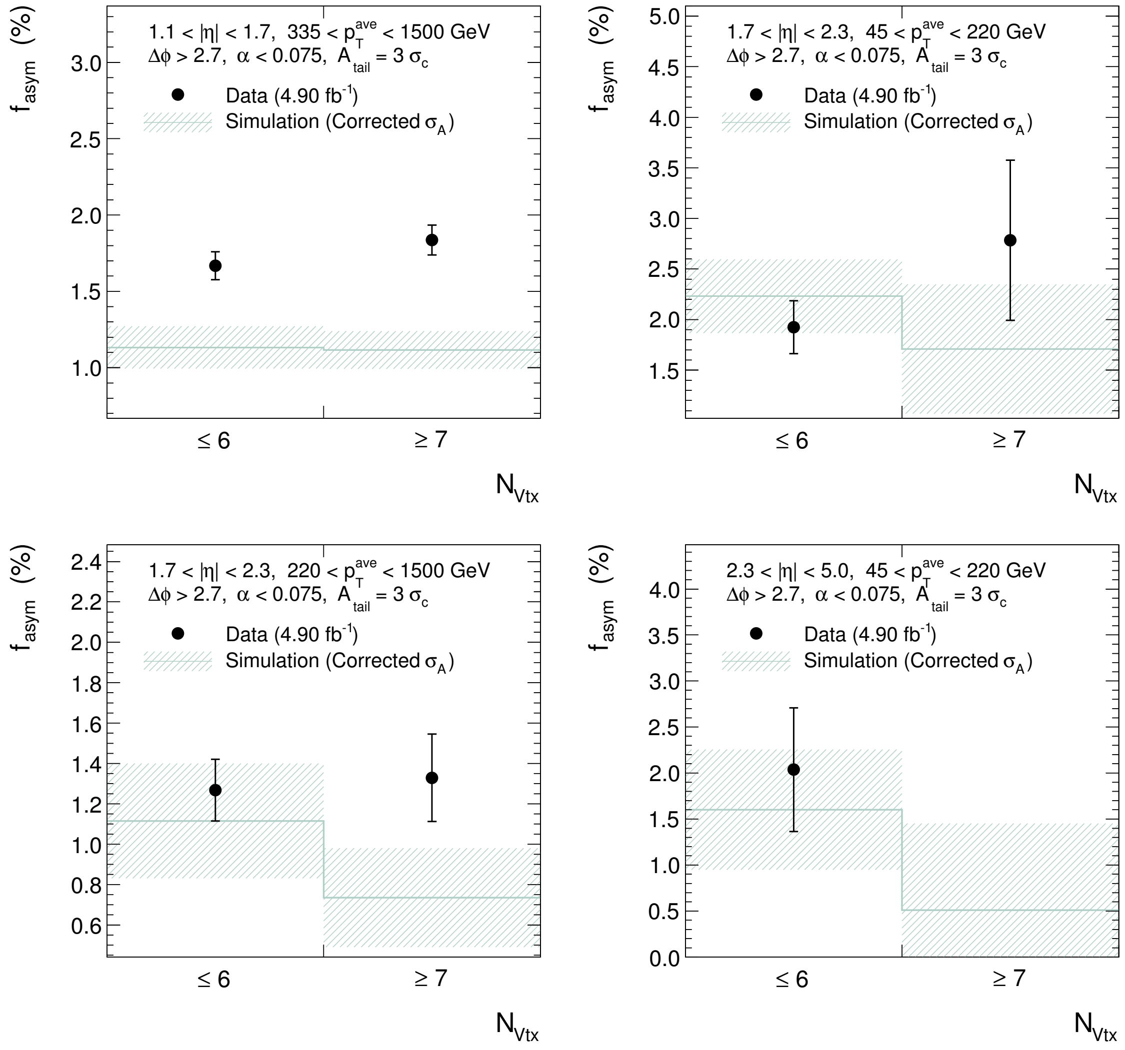


Figure 6.15: Continued from Fig. 6.14. In the last $|\eta|$ interval (*bottom right*), no data events pass the $N_{\mathrm{Vtx}} \geq 7$ selection.

6.5.2 Data-to-Simulation Ratio

The measured data-to-simulation ratios ρ_{tail} Eq. (6.6) of the $f_{\text{asym}}(\alpha_{\max} \to 0)$ in the studied tail regions with $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$ and $3\,\sigma_c$ are shown in Figs. 6.16 and 6.17, respectively, together with the propagated extrapolation uncertainties δf_{ex} Eq. (6.4) and the systematic uncertainties. The extrapolation uncertainty depends on but underestimates the statistical uncertainty of the extrapolated f_{asym}, and the residual contribution is expected to be covered by the systematic uncertainty. However, the statistical uncertainties on the ρ_{tail} can be assumed to be small compared to the systematic uncertainties. This can be concluded from a comparison with Figs. 6.10 and 6.11, which display the f_{asym} for $\alpha < 0.074$ before the extrapolation.

The values of ρ_{tail} are also listed in Table B.3 in Appendix B.5. General features such as the increase of the extrapolation uncertainty with larger $|\eta|$ and $\mathcal{A}_{\text{tail}}$ have been discussed in the previous Section 6.5.1 for the f_{asym} and are propagated to the ρ_{tail}. The tails are generally more pronounced in data than in the simulation. Differences occur in particular in the central detector region at medium to large p_{T}. Here, ρ_{tail} amounts to $1.2 - 1.6$ for $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$ and $3\,\sigma_c$, respectively, with uncertainties between 0.2 and 0.5.

The dependence of ρ_{tail} on the probed tail region $\mathcal{A}_{\text{tail}}$ is investigated in Figs. 6.18 to 6.21[6]. Since ρ_{tail} is dominated by the behaviour close to the tail start, the correlation between measurements with different $\mathcal{A}_{\text{tail}}$ is expected to be small. The systematic uncertainties are partly correlated, though, because the same value of ρ_{res} has been used in Eq. (5.21) when correcting the core of the simulated asymmetry distributions for the larger resolution in data. Overall, the measured values of ρ_{tail} do not differ systematically between the investigated tail regions. Some trend is visible, though, suggesting an increase of ρ_{tail} by approximately $20 - 30\%$ when $\mathcal{A}_{\text{tail}}$ changes from 2 to $3\,\sigma_c$, but this is not conclusive given the uncertainties. Hence, within the uncertainties, the shape of the simulated asymmetry tails is consistent with the data. Based on the results for f_{asym} in the previous Section 6.5.1, the measurement of ρ_{tail} is expected to be insensitive to the presence of pile-up. This is verified in Figs. 6.22 to 6.25, where ρ_{tail} is shown as a function of the number of primary vertices for $\mathcal{A}_{\text{tail}} = 3\,\sigma_c$. As before, the displayed systematic uncertainties are partly correlated. No systematic trend is present.

The presented results are consistent with the results from previous versions of this analysis obtained with $855\,\text{pb}^{-1}$ of 2011 data [238] and $36\,\text{pb}^{-1}$ of 2010 data [172, 239]. The differences between the asymmetry tails in data and simulation are expected to occur similarly for the jet-p_{T} response. Due to the definition of the dijet asymmetry, it is not possible though to determine to what extent the low- or the high-response tail contributes to the observed difference. However, given the potential sources of non-Gaussian contributions discussed in Section 4.4.2 as well as the predictions from simulation, the low-response tail is expected to be clearly larger than the high-response tail. Hence, it is likely that the lower tail also has a more important contribution to the data-to-simulation difference.

Differences between data and simulation occur predominantly in the central detector region in case of high-p_{T} jets. Hence, a possible source might be mismodelled punch-through. Likewise, mismodelled detector inhomogeneities and channel intercalibration might contribute. Furthermore, the simulation of the fraction of heavy-flavour quarks, which contribute an important component to the tails through semi-leptonic decays, are another possible source of the observed difference.

---

[6] An intermediate region $\mathcal{A}_{\text{tail}} = 2.5\,\sigma_c$ has been added for this study.

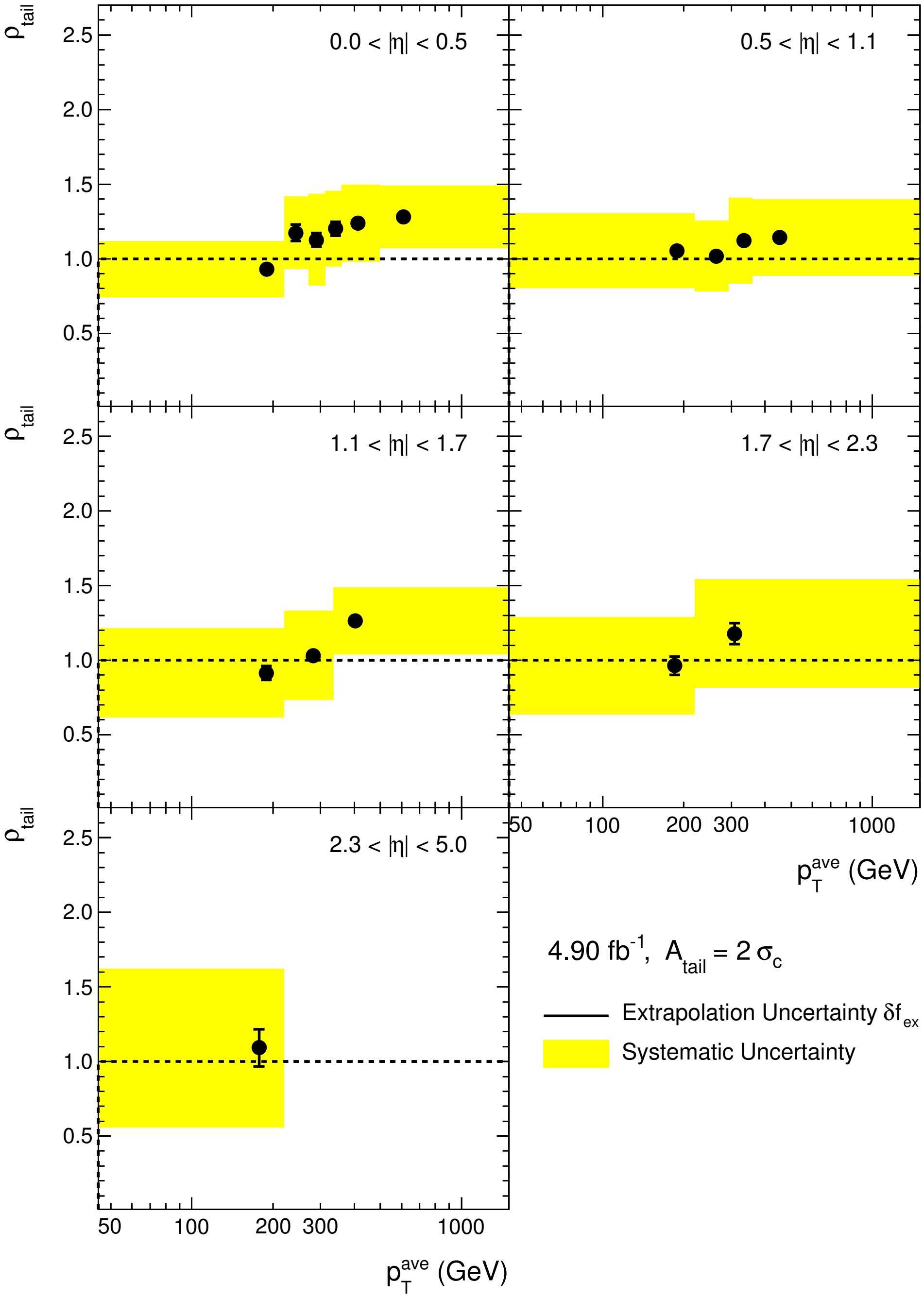


Figure 6.16: Ratio ρ_{tail} of the fractional tail-size f_{asym} in data and simulation in different $|\eta| \times p_{\mathrm{T}}^{\text{ave}}$ intervals in the region $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$. The horizontal position of the points corresponds to the mean $p_{\mathrm{T}}^{\text{ave}}$ in the respective interval. Error bars represent the propagated extrapolation uncertainties δf_{ex}, error bands represent the systematic uncertainties.

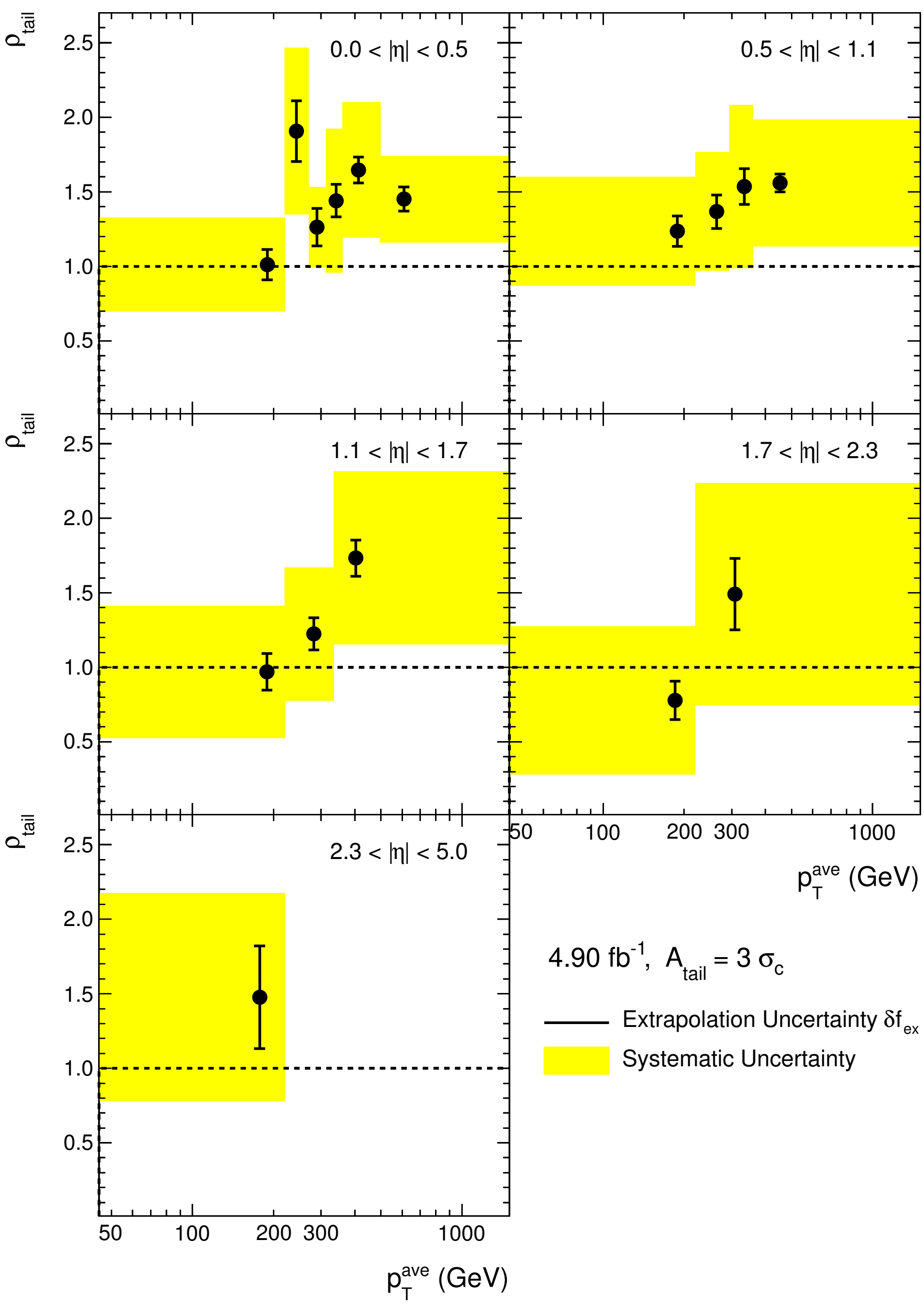


Figure 6.17: Ratio ρ_{tail} of the fractional tail-size f_{asym} in data and simulation in different $|\eta| \times p_{\mathrm{T}}^{\text{ave}}$ intervals in the region $\mathcal{A}_{\text{tail}} = 3\,\sigma_c$. The horizontal position of the points corresponds to the mean $p_{\mathrm{T}}^{\text{ave}}$ in the respective interval. Error bars represent the propagated extrapolation uncertainties δf_{ex}, error bands represent the systematic uncertainties.

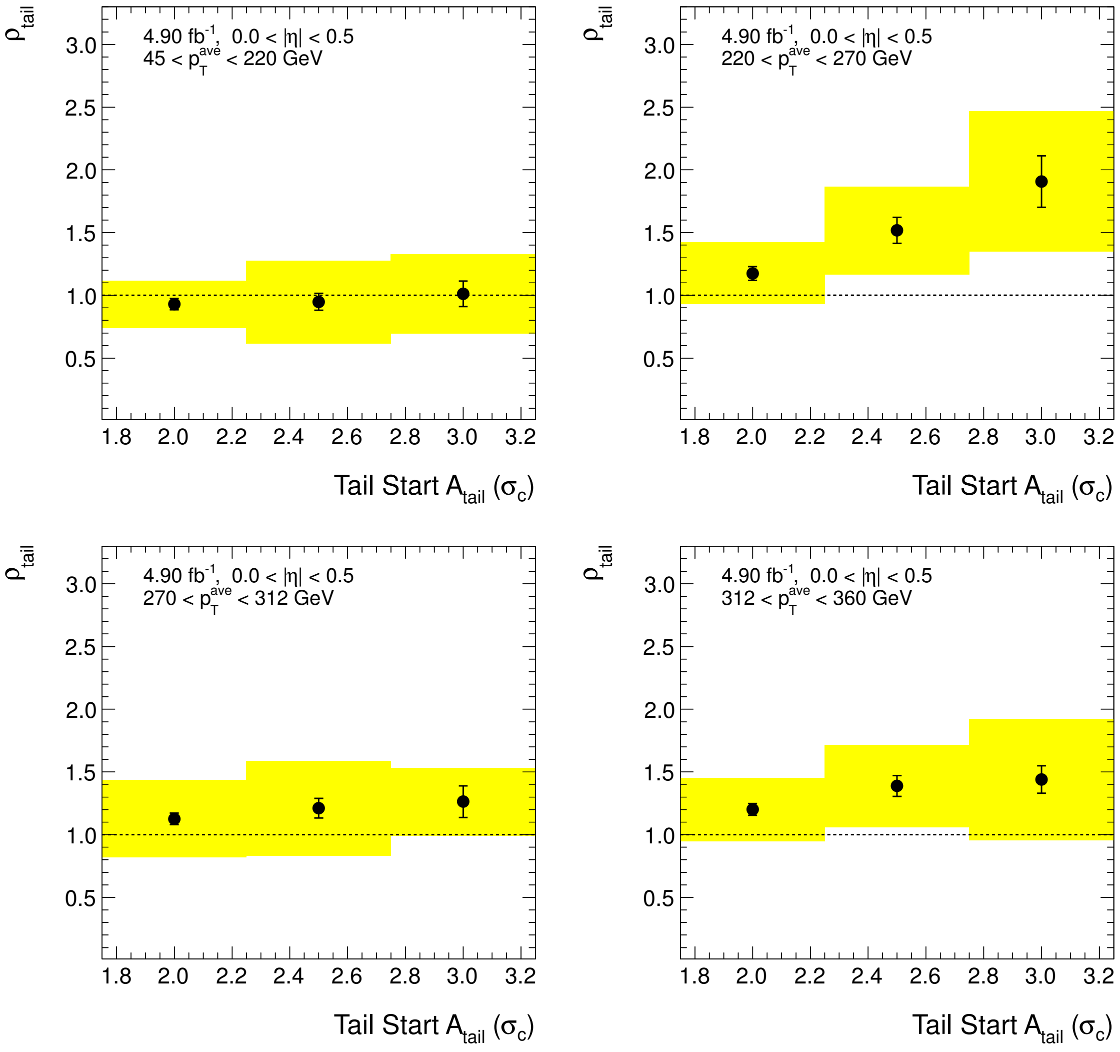


Figure 6.18: Dependence of ρ_{tail} on the selected tail region defined by its start $\mathcal{A}_{\text{tail}} = n \cdot \sigma_c$ in different $|\eta| \times p_{\text{T}}^{\text{ave}}$ intervals. Error bars represent the propagated extrapolation uncertainties δf_{ex}, and error bands represent the systematic uncertainties, which are partly correlated. (Continued in Fig. 6.19.)

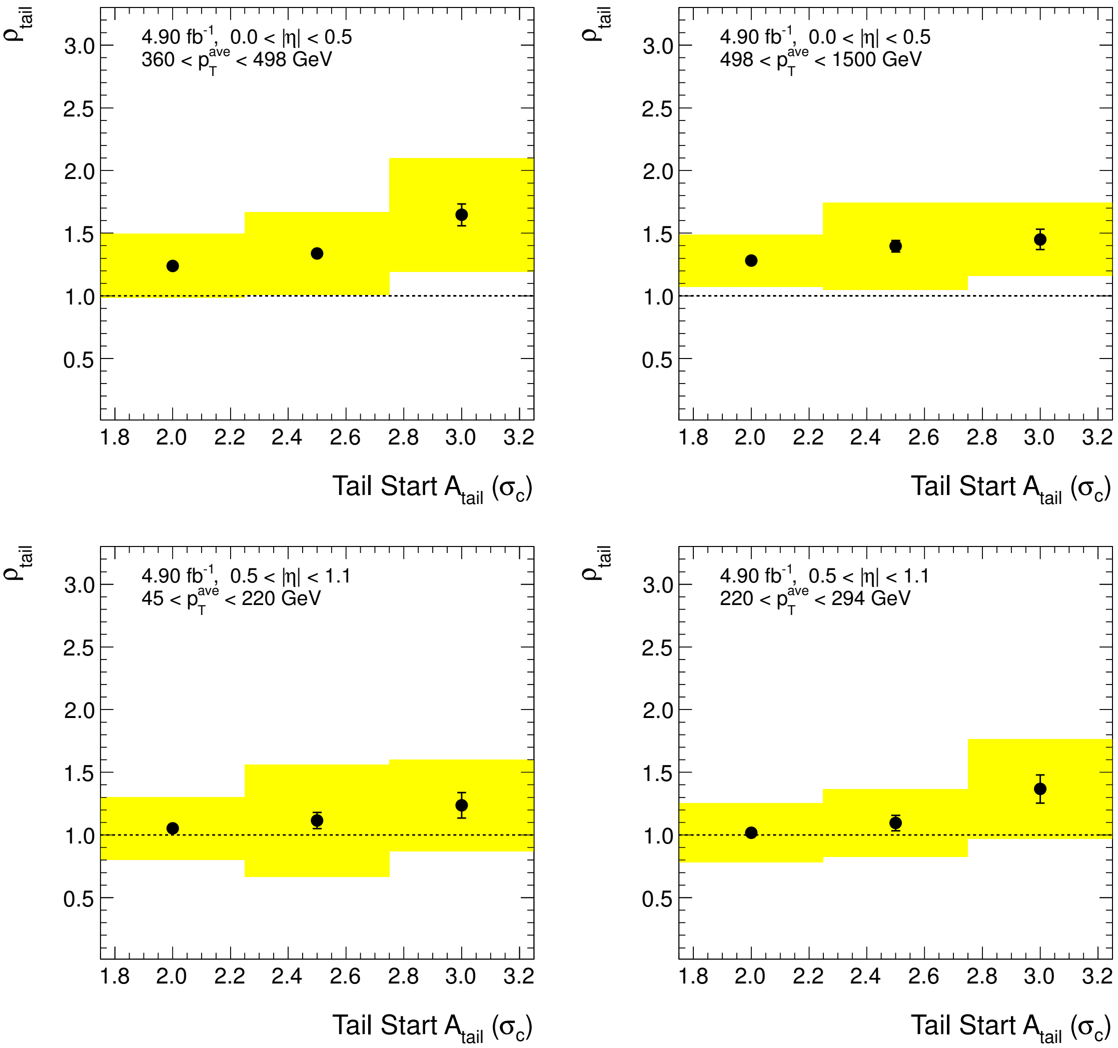


Figure 6.19: Continued from Fig. 6.18 and in Fig. 6.20.

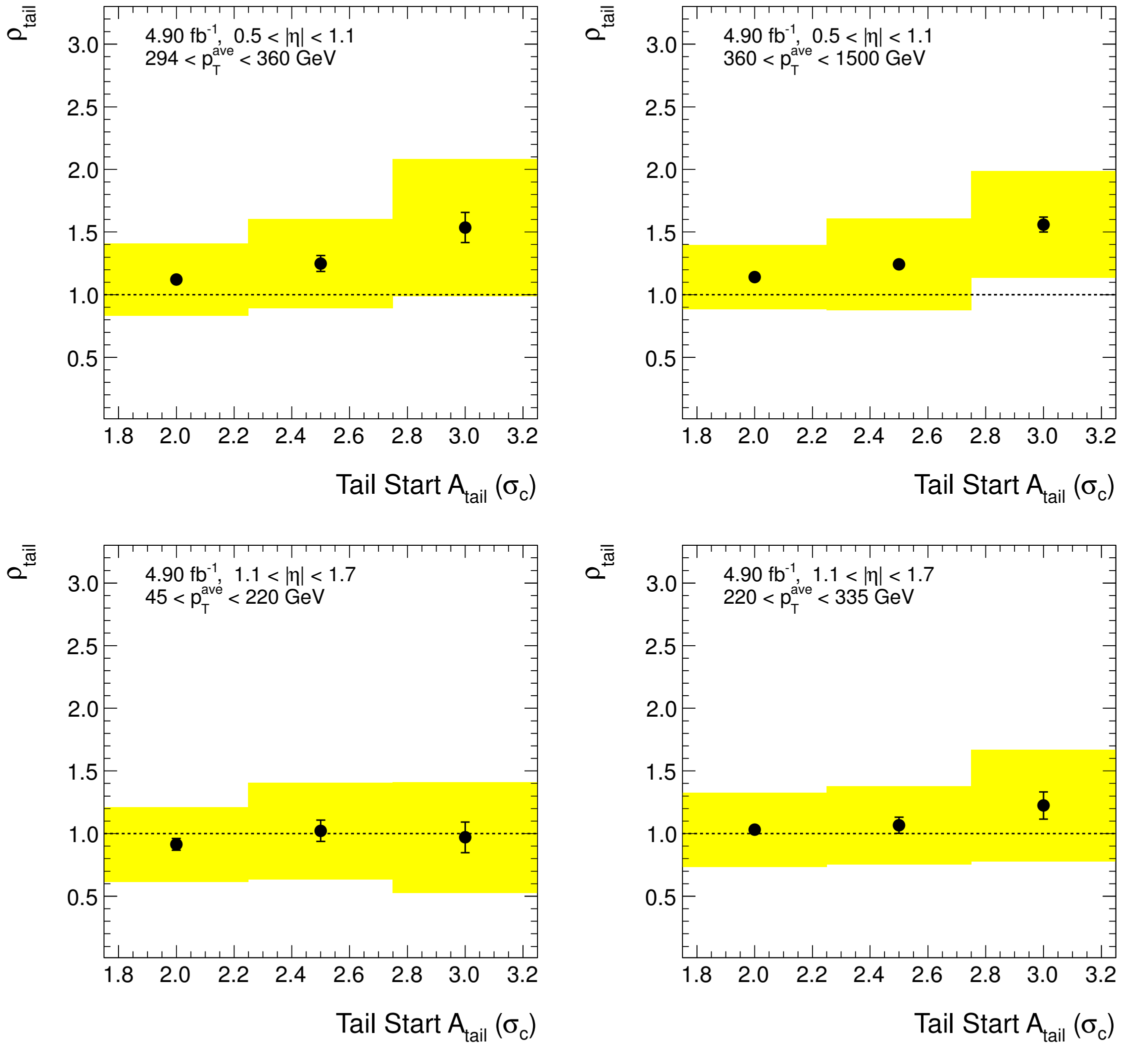


Figure 6.20: Continued from Fig. 6.19 and in Fig. 6.21.

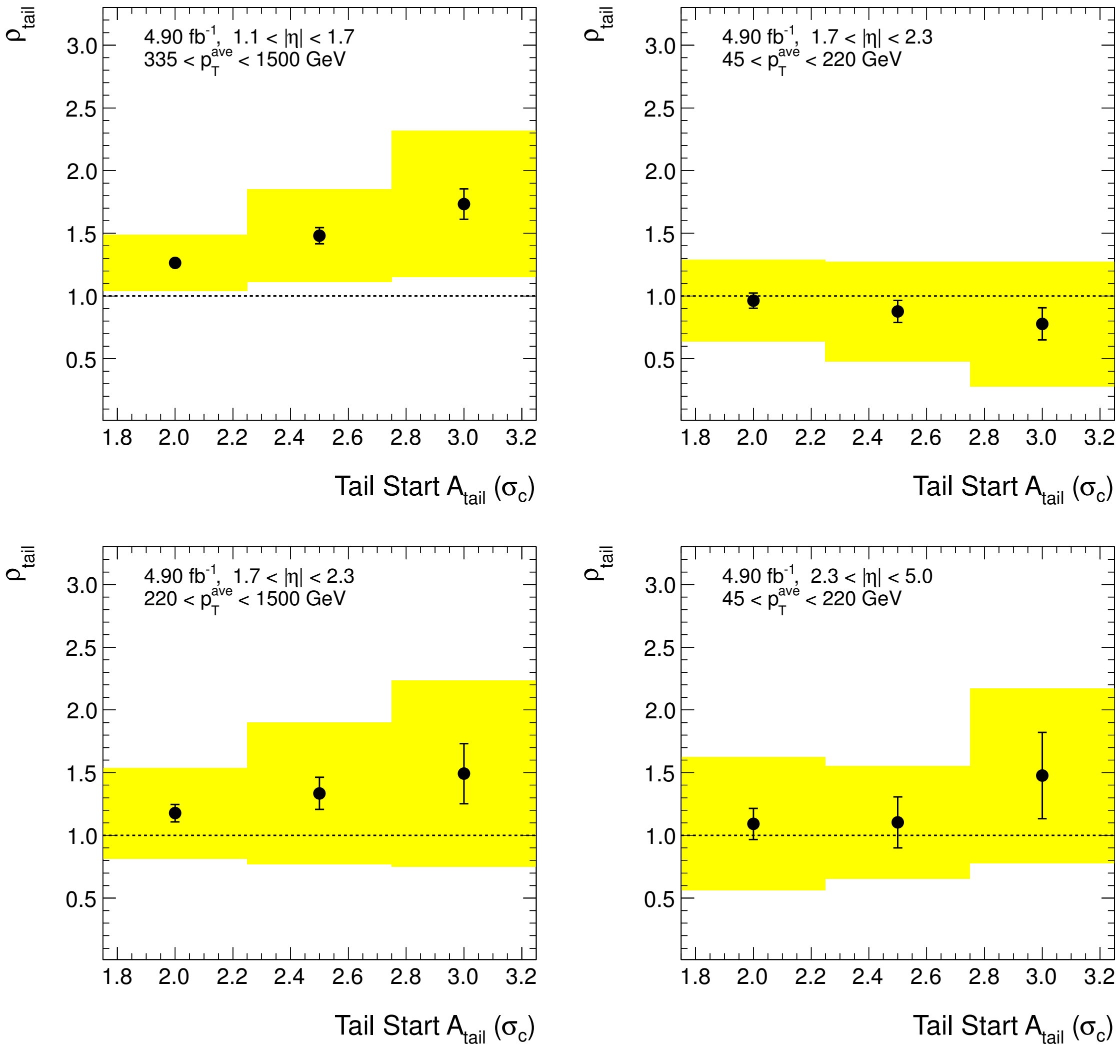


Figure 6.21: Continued from Fig. 6.20.

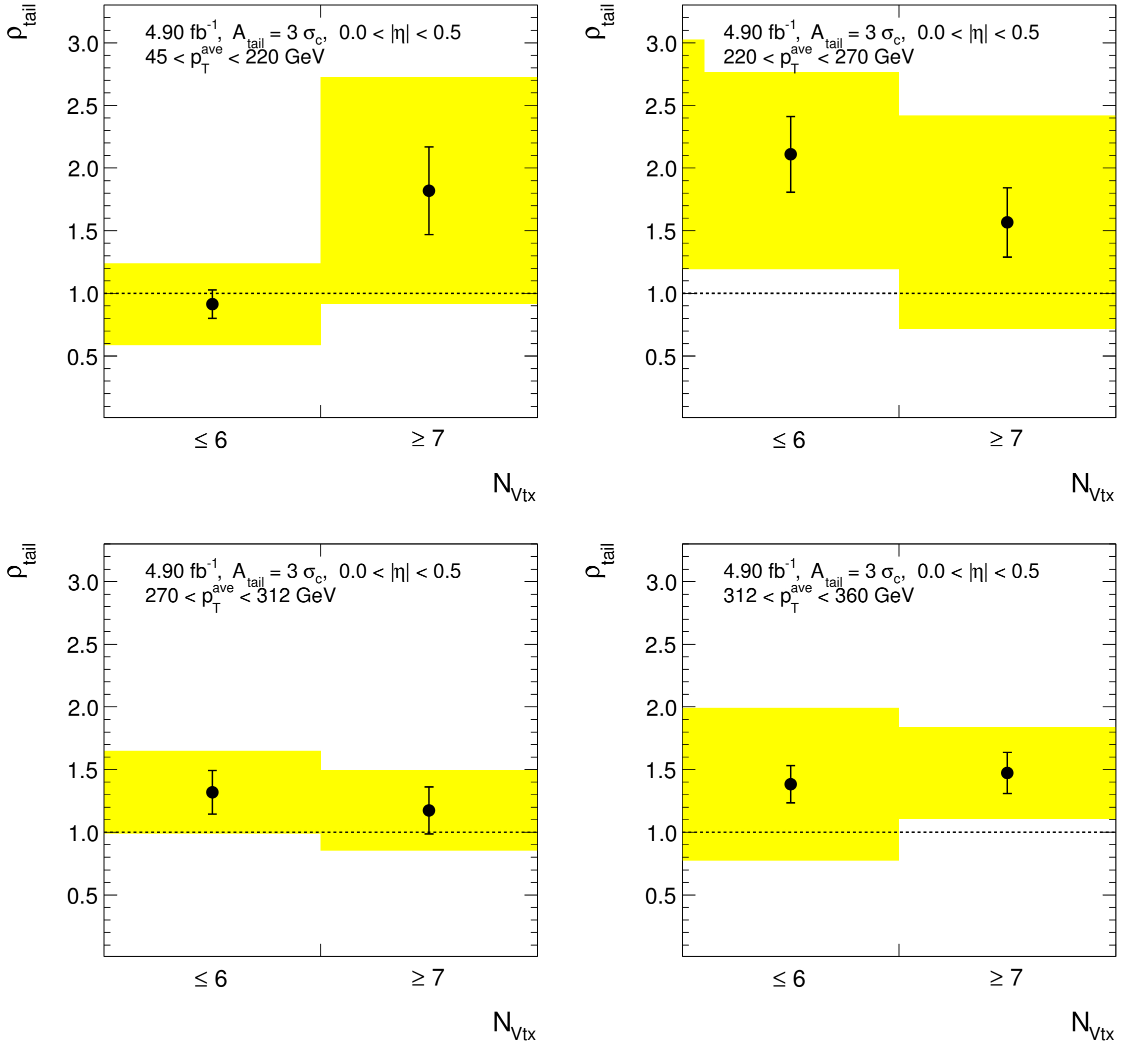


Figure 6.22: Dependence of ρ_{tail} on the number N_{Vtx} of primary vertices in different $|\eta| \times p_{\text{T}}^{\text{ave}}$ intervals for $\mathcal{A}_{\text{tail}} = 3\,\sigma_c$. The same value of σ_c has been used for both N_{Vtx} selections to avoid fluctuations due to binning effects. Error bars represent the propagated extrapolation uncertainties δf_{ex}, and error bands represent the systematic uncertainties, which are partly correlated. (Continued in Fig. 6.13.)

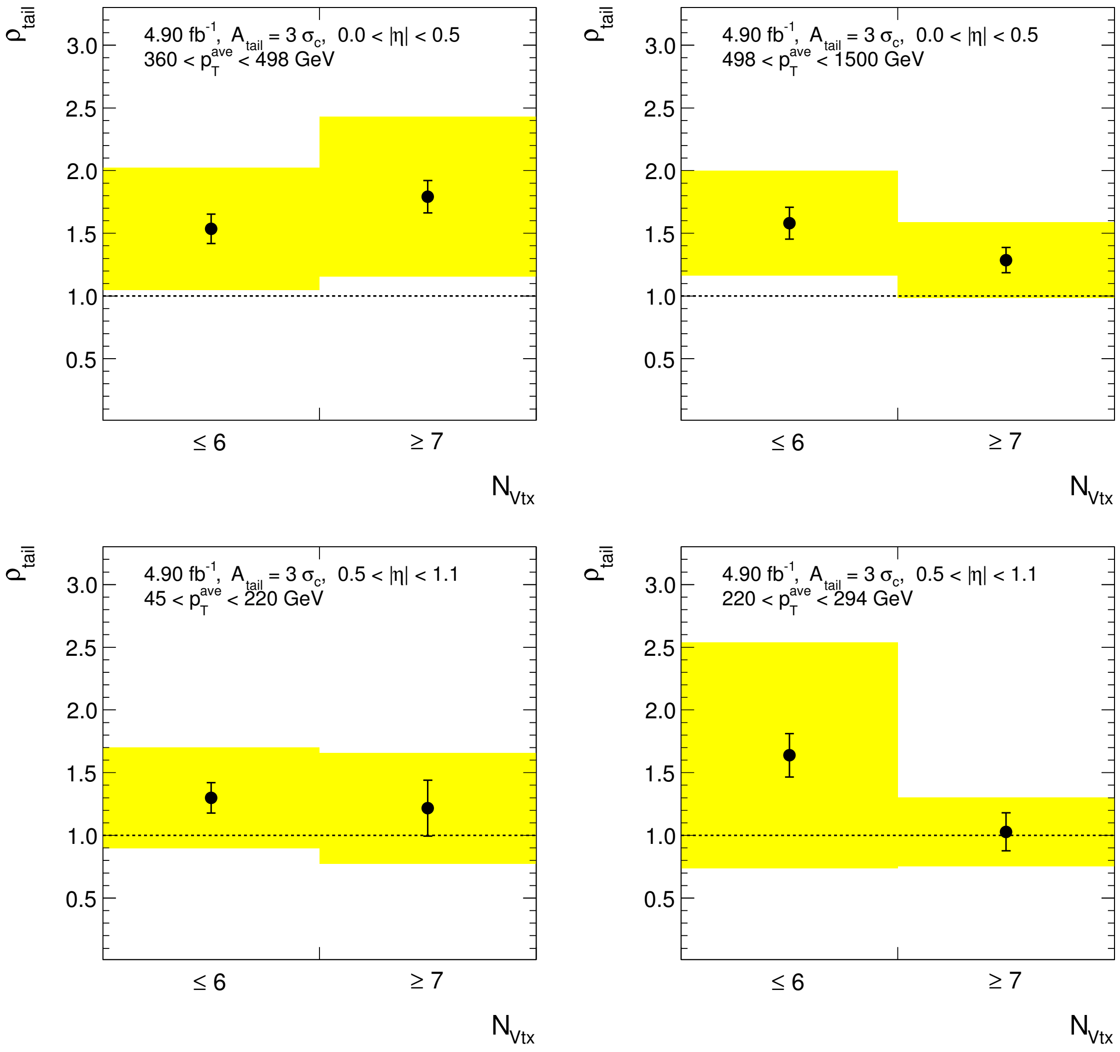


Figure 6.23: Continued from Fig. 6.22 and in Fig. 6.24.

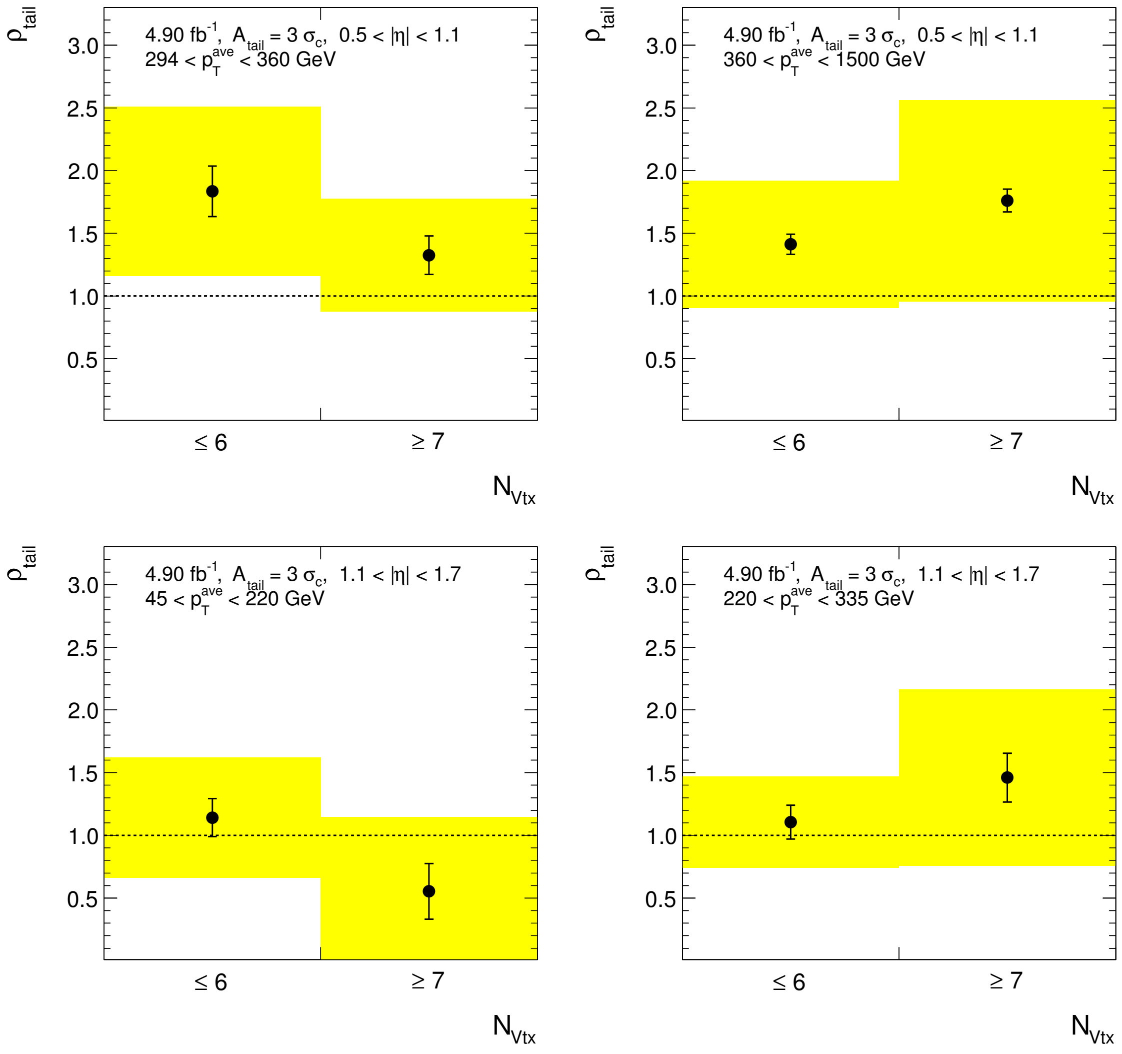


Figure 6.24: Continued from Fig. 6.23 and in Fig. 6.25.

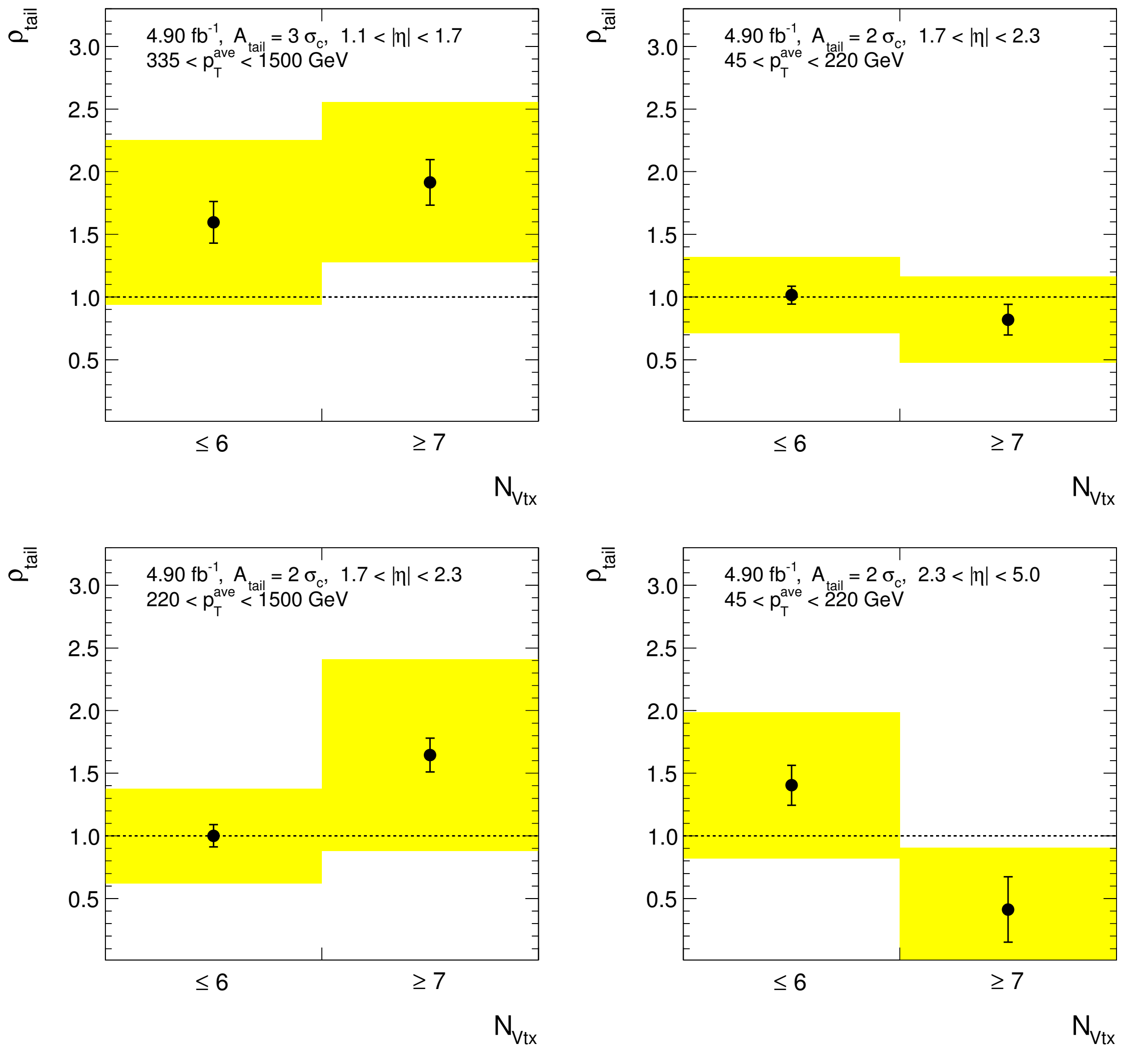


Figure 6.25: Continued from Fig. 6.24. For $|\eta| > 1.7$ (*top right, bottom*), the precision of the measurement becomes insufficient for a comparison due to the small number of events in each subsample. Therefore, results for the $\mathcal{A}_{\text{tail}} = 2\,\sigma_c$ region are shown.

6.6 Adjustment of the MC-Truth Response to Data

Differences between the simulated jet-p_{T} response tails and the response tails in data will induce differences between the asymmetry in simulation and data and will cause measurements of $\rho_{\mathrm{tail}} > 1$. Hence, ρ_{tail} can be used to correct the simulated response. Provided the data-to-simulation ratio of the response tails is the same as the one of the asymmetry tails, the respective parts of the simulated response distribution can simply be scaled by ρ_{tail}. This assumption is validated in Section 6.6.1 using a toy simulation. In Section 6.6.2, a method to correct the MC-truth response based on this assumption is briefly reviewed.

6.6.1 Correspondence Between Response- and Asymmetry-Tail Variations

A set of $5 \cdot 10^5$ toy events is generated. Each event contains two random numbers $\{r_i \,|\, i = 1,2\}$ representing the response of the two jets in a QCD-dijet event. The r_i are sampled from a Gaussian pdf with mean value 1 and standard deviation $\sigma_{\mathrm{toy}} = 0.1$ between 0.75 and 1.25. In 10% of the cases, r_i is sampled from a uniform pdf between 0.3 and 0.75, which corresponds to a tail component with $f_{\mathrm{resp}}^{\mathrm{toy}} = 0.1$, cf. Fig. 6.26 (*left*). Per event, the asymmetry $\mathcal{A}_{\mathrm{toy}}$ is calculated according to Eq. (6.5), and the resulting distribution of absolute values[7] $|\mathcal{A}_{\mathrm{toy}}|$ is shown in Fig. 6.26 (*right*).

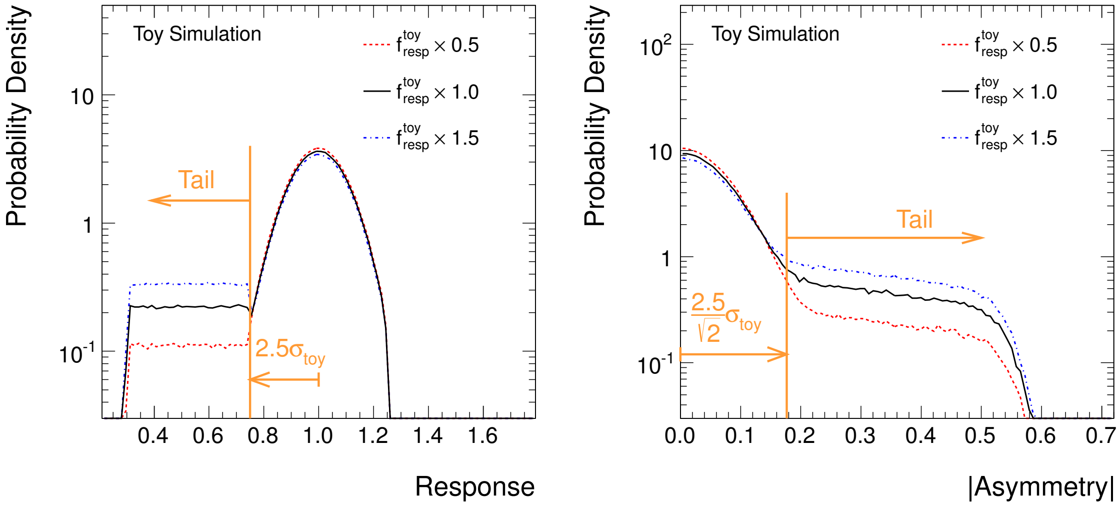


Figure 6.26: Toy simulation to establish the correspondence between tail variations in the response and asymmetry distributions. Jet-p_{T} measurements are sampled from a response pdf (*left*), which consists of a Gaussian core with mean value 1 and a uniform tail below $1 - 2.5\,\sigma_{\mathrm{toy}}$, and are used to compute the dijet asymmetry (*right*). The response-tail component $f_{\mathrm{resp}}^{\mathrm{toy}}$ has been varied from a default contribution of 10% to 5% and 15%, respectively. The corresponding asymmetry-tail components are indicated.

The start of the tail component at a response of 0.75 corresponds to a distance of $2.5\,\sigma_{\mathrm{toy}}$

---

[7] Recall that the asymmetry results in a symmetric distribution by construction.

from the Gaussian mean. Following Eq. (4.11), the tail of the asymmetry distribution is therefore defined as the region with $|\mathcal{A}_{\mathrm{toy}}| > \frac{1}{\sqrt{2}}\,2.5\,\sigma_{\mathrm{toy}}$.

Further sets of events are generated, where the tail component's size $f^{\mathrm{toy}}_{\mathrm{resp}}$ has been varied between 0.05 and 0.15, cf. Fig. 6.26. As can be seen from Fig. 6.27, a variation of $f^{\mathrm{toy}}_{\mathrm{resp}}$ results in a variation of the tail component of the asymmetry, $f^{\mathrm{toy}}_{\mathrm{asym}}$, as expected. The change $f^{\mathrm{toy}}_{\mathrm{asym}}$ is the same within a few percent as the change in $f^{\mathrm{toy}}_{\mathrm{resp}}$. A trend is visible, however, indicating a bias that for larger values of $f^{\mathrm{toy}}_{\mathrm{resp}}$ the corresponding change in $f^{\mathrm{toy}}_{\mathrm{asym}}$ is smaller and vice verse. Since the effect is small compared to the typical uncertainties of ρ_{tail}, it is not considered in the scaling approach described below. As the precision increases however, it appears worthwhile to further study the relation between f_{asym} and f_{resp} for future analyses.

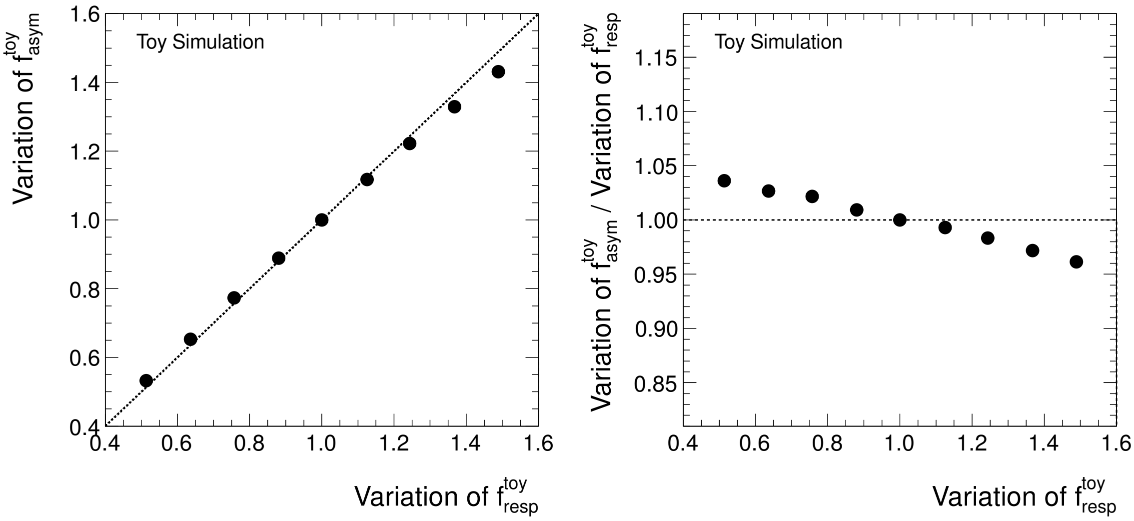


Figure 6.27: Toy simulation to establish the correspondence between tail variations in the response and asymmetry distributions. Shown is the variation in $f^{\mathrm{toy}}_{\mathrm{asym}}$ caused by a variation of $f^{\mathrm{toy}}_{\mathrm{resp}} = 10\%$ by factors between 0.5 and 1.5 (*left*) and the variation in $f^{\mathrm{toy}}_{\mathrm{asym}}$ relative to the variation of $f^{\mathrm{toy}}_{\mathrm{resp}}$ (*right*).

6.6.2 Response Tail Scaling

Based on the correspondance investigated above between variations of the asymmetry and response tails, a method has been developed to correct the MC-truth response tails using the measured values of ρ_{tail} as tail-scale factors [238].

The MC-truth response ('$\mathcal{R}$') distribution is first adjusted to the larger resolution in data as described in Section 5.7.1. Then, a Gaussian is fitted to the central part of the adjusted distribution, where the central part is defined as the range of ±1 standard deviations $\sigma_{\mathcal{R}}$ around the mean $\langle\mathcal{R}\rangle$, and the Gaussian contribution is subtracted form the distribution. A lower response tail is defined as that region of the residual MC-truth response distribution with $\mathcal{R} < \langle\mathcal{R}\rangle - n\,\sigma_{\mathcal{R}}$, $n \in \mathbb{R}^+$ and an upper tail accordingly with $\mathcal{R} > \langle\mathcal{R}\rangle + n\,\sigma_{\mathcal{R}}$. Finally, the tails are scaled whereby the choice of n has to match the definition of $\mathcal{A}_{\mathrm{tail}}$, e. g. 2 or 3, and added back to the Gaussian core distribution[8]. Here, tails contain only the non-Gaussian contributions of a distribution whereas for the ρ_{tail} measurement the total content is considered. Therefore, the actually applied scale factor is a modification of ρ_{tail} to account for this difference. By scaling only the non-Gaussian part, discontinuities of the corrected response distribution are avoided.

Since the dijet asymmetry results in a symmetric distribution by construction, the observed difference between the asymmetry tails in data and simulation cannot be attributed unambigiously to a mismodelled lower or upper response tail, as discussed previously. Therefore, both tails are scaled, and the uncertainty of the corrected MC-truth response distribution is determined as the difference of the distributions obtained when varying ρ_{tail} by its lower and upper uncertainty.

6.7 Outlook

Several extensions of the presented method for the jet-p_{T} response tail measurement appear worthwhile.

The main limitation in precision arises currently from the propagated uncertainty on the measured jet-p_{T} resolution. Hence, increased precision of the latter would also greatly improve the tail measurement.

The assignment of the observed data-to-simulation difference to the low- or high-response tail is ambigious by definition of the asymmetry. Although the low-response tail likely has the predominant contribution, a final conclusion can only be obtained when incorporating information about the true jet-p_{T} scale. This could also help improving the translation of ρ_{tail} into a scale factor for the response tails.

In principle, this could be achieved using the maximum-likelihood method developed in the previous Chapter 5. However, as argued, the fit of non-Gaussian response functions turns out to be difficult. A different possibility is to combine the dijet-asymmetry-based measurement with the results from photon + jet or Z + jet events. As remarked previously, the cross sections of these processes, and thus the reach in p_{T} and statistical precision in the tails, are much lower than for QCD-dijet production, but they allow direct access to the response function as described in Section 4.4.5. Due to the trigger prescales, the rate of collected photon/Z + jet events at low p_{T} is in fact higher than that of dijet events, which additionally improves the statistical precision of the measurement at low p_{T}. A combination has not been done yet because the jet-p_{T} response tails have so far not been investigated by CMS in photon/Z + jet data of sufficient size. Existing results from photon + jet events in $36\,\mathrm{pb}^{-1}$ of data [172, 239] are statistically too limited to study the tails.

---

[8] To be precise, the difference of $\mathcal{A}_{\mathrm{tail}}$ and $\hat{\mathcal{A}}_{\mathrm{tail}}$ as well as the fact that $\sigma_c = \sigma_{\mathcal{A}}(\alpha_{\max} = 0.05)$ have to be considered.

7 Search for New Physics in Events with Jets and Missing Transverse Momentum

As outlined before in Section 2.3, SUSY is a promising and one of the most extensively studied frameworks for physics beyond the SM. Hence, SUSY often serves as a guideline in the design of new-physics searches at collider experiments. At the LHC, sparticles, mostly squarks and gluinos, can be produced through the strong interaction in the initial pp collisions. Due to their high masses, they subsequently decay via cascades into SM particles with large momenta. Since the sparticles are coloured in most cases, they will initially decay into coloured SM particles and hence lead to jets in the detector. In case R-parity is conserved, the sparticles will be created in pairs, and the LSP will be stable. This is illustrated in Fig. 7.1. If only weakly interacting, such as the neutralino in the MSSM, the LSP will not produce any direct signal in the detector but will manifest as $\not\!\!E_\mathrm{T}$.

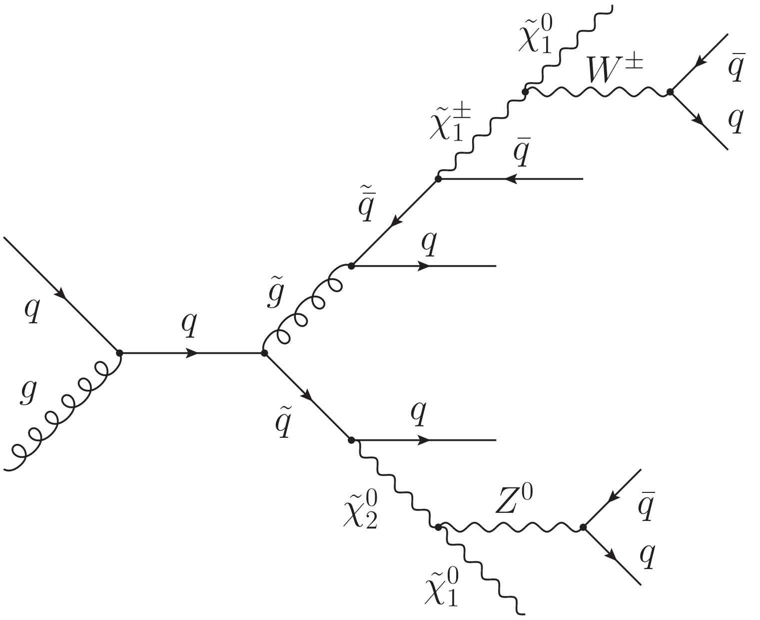


Figure 7.1: Sketch of a possible sparticle-production process at the LHC with a subsequent cascade decay leading to an all-hadronic final-state within the framework of the R-parity conserving MSSM. Taken from [258].

Since, of course, it is not known what kind of new physics — if at all — is realised at the TeV scale, it is not obvious either what kind of signature exactly to expect at the LHC experiments. Therefore, CMS conducts a variety of searches for new physics covering a broad range of possible signatures which are motivated but not limited to supersymmetric models [259]. They can be roughly categorised as searches in lepton, photon, or multijet final-states, each time accompanied by $\not\!\!E_{\mathrm{T}}$. Each channel is covered by several analyses with different techniques and focus, thus allowing for mutual cross checks.

The multijet (*all-hadronic*) final-state is characterised by several high-p_{T} jets, large $\not\!\!E_{\mathrm{T}}$, and no isolated leptons, cf. Fig. 7.1. Given the strong production mechanism and the subsequent decay modes [49], this channel has the largest branching ratio in many SUSY scenarios and can reach highest sensitivity. At the same time, however, it faces a huge background from SM processes with the same signature. Genuine large $\not\!\!E_{\mathrm{T}}$ can be caused by the neutrinos in $Z \to \nu\bar{\nu}$ + jets events as well as in $t\bar{t}$ and W + jets events, where the W boson decays into a neutrino and a light lepton which is not identified or a hadronically decaying τ lepton. A different major background arises from QCD-multijet events where one or more jets are severely mismeasured, for example because of semi-leptonically decaying heavy-flavour quarks inside the jet or because of instrumental effects.

The discussed analysis has been designed as an inclusive search for new physics with minimal kinematic biases induced by the event selection to provide a high acceptance for a wide range of possible new-physics scenarios with all-hadronic final-states. Therefore, the sensitive observables in the analysis are H_{T}, the scalar sum of the jet transverse momenta, and $\not\!\!H_{\mathrm{T}}$ ('missing H_{T}'), the missing transverse-momentum computed from the jet momenta. The aim for broad acceptance does not permit aggressive background suppression and therefore requires a precise understanding of the relatively large contributions from SM processes in the search regions. Hence, the key features of this analysis are robust methods to measure the full background spectra directly from data with minimal reliance on simulation. In the following, the analysis will be termed *jets + $\not\!\!H_T$ analysis.*

The jets + $\not\!\!H_{\mathrm{T}}$ analysis has initially been published based on $36\,\mathrm{pb}^{-1}$ of data acquired by CMS in 2010 [260], and an update with the first $1.1\,\mathrm{fb}^{-1}$ of 2011 data has been published in September 2011 [238]. This thesis focuses on the result obtained with the full data set collected by CMS during 2011 corresponding to an integrated luminosity of $4.98\,\mathrm{fb}^{-1}$ [237].

CMS has conducted further searches in the all-hadronic channel complementing the jets + $\not\!\!H_{\mathrm{T}}$ analysis [261–265]. These analyses are typically based on the assumption of initial pair-production of heavy particles. Hence, different kinematic variables such as α_{T} [266], M_{T2} [267, 268], or the 'razor' variables R and M_R [269] are exploited in order to reduce the SM backgrounds, in particular from QCD-multijet events. The ATLAS collaboration has also published results from searches for new physics in similar all-hadronic final-states [270–272].

In this chapter, the jets + $\not\!\!H_{\mathrm{T}}$ analysis is discussed. Firstly, in Section 7.1, the selection of events with large values of H_{T} and $\not\!\!H_{\mathrm{T}}$ is described. The SM background is much reduced by this *baseline selection*, while the acceptance of possible signal events is retained. Afterwards, more confined, exclusive *search regions* are defined by tightening the criteria on H_{T} and $\not\!\!H_{\mathrm{T}}$ in order to increase the sensitivity to specific new-physics scenarios. Secondly, in Section 7.2, the measurement of the SM backgrounds from data is discussed. Special emphasis is put on the QCD-multijet background because the results of the resolution measurements presented in Chapters 5 and 6 are an essential input to its prediction. Finally, in Section 7.3, the search results are interpreted within the context of the CMSSM and of more generic Simplified Models, and their impact on global MSSM-parameter fits is discussed.

7.1 Sample and Event Selection

The presented analysis is performed on data from *pp* collisions at a centre-of-mass energy of 7 TeV which have been acquired from March until October 2011 by the CMS experiment. With all subdetectors fully functional, an integrated luminosity of $4.98\,\mathrm{fb}^{-1}$ has been recorded.

Events have been collected with a level-1 trigger on $H_{\mathrm{T}}^{\mathrm{L1}}$, which is computed from calorimeter jets with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 5$. The $H_{\mathrm{T}}^{\mathrm{L1}}$ threshold has been adjusted from 100 to 150 GeV during the data-taking period to cope with the increase in luminosity. Afterwards, a high-level trigger selection has been performed by triggering on $H_{\mathrm{T}}^{\mathrm{HLT}}$ and, for most of the data, at the same time on $\not{H}_{\mathrm{T}}^{\mathrm{HLT}}$. All HLT paths are listed in Table 7.1. $H_{\mathrm{T}}^{\mathrm{HLT}}$ and $\not{H}_{\mathrm{T}}^{\mathrm{HLT}}$ are computed from anti-k_T calorimeter jets with size parameter $R = 0.5$, cf. Section 4.2, and with the jet energy scale corrections, cf. Section 4.5, applied. For $H_{\mathrm{T}}^{\mathrm{HLT}}$, jets with $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ and $|\eta| < 3$ which pass dedicated jet-id quality criteria [217, 247], and for $\not{H}_{\mathrm{T}}^{\mathrm{HLT}}$, jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 3$ have been considered. In the latter case, no jet-id requirement is applied because in rare cases the cleaning has been observed to remove real jets and thus to fake $\not{H}_{\mathrm{T}}$. The trigger efficiencies have been measured as a function of the offline H_{T} and $\not{H}_{\mathrm{T}}$ defined below on data collected with more inclusive triggers. In case of the L1 trigger, a single-jet trigger has been used as reference, and the efficiency is greater than 99% for $H_{\mathrm{T}} > 260\,\mathrm{GeV}$; in case of the HLT, more inclusive $H_{\mathrm{T}}^{\mathrm{HLT}}$-$\not{H}_{\mathrm{T}}^{\mathrm{HLT}}$ trigger paths have been used as reference, and the efficiency is greater than 99% for $H_{\mathrm{T}} > 350\,\mathrm{GeV}$ and $\not{H}_{\mathrm{T}} > 200\,\mathrm{GeV}$.

All physics objects in data and simulation are reconstructed with the PF algorithm Section 3.5.2. Jets are clustered from the reconstructed particles with the anti-k_T jet algorithm with parameter size $R = 0.5$. At detector-level, the jet energy scale calibration is applied, including the residual correction for data, cf. Section 4.5.

The event sample for the search is selected by first applying the following baseline-selection criteria that define a validation region which serves as a basis for tighter selection requirements.

- There are at least three jets with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ and $|\eta| < 2.5$ in order to select events with a multijet final-state topology.
- The sample is enhanced with events with a high energy scale of the hard-scattering process by requiring $H_{\mathrm{T}} > 500\,\mathrm{GeV}$, where

  $$H_{\mathrm{T}} = \sum_i |\vec{p}_{\mathrm{T},i}|$$

  and i runs through all jets with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ and $|\eta| < 2.5$.
- In order to suppress contributions from SM processes, in particular QCD-multijet production, $\not{H}_{\mathrm{T}} > 200\,\mathrm{GeV}$ has to be met, where

  $$\not{H}_{\mathrm{T}} = \left| -\sum_i \vec{p}_{\mathrm{T},i} \right|$$

  and i runs through all jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 5$. The different choice of jets in the calculation of H_{T} and $\not{H}_{\mathrm{T}}$ has been introduced in [260] to account for

Table 7.1: Employed high-level trigger paths. Stated are the run range, the path name which indicates the thresholds on the trigger-level $H_{\mathrm{T}}^{\mathrm{HLT}}$ and $\not{H}_{\mathrm{T}}^{\mathrm{HLT}}$, and the integrated luminosity of the collected data taking into account prescales. The relative uncertainty on the luminosity measurement is 2.2% [240].

Run range	Trigger path	Luminosity (pb^{-1})
160431 – 160578	`HLT_HT160_v2`	< 1
160871 – 160943	`HLT_HT240_v2`	< 1
160955 – 160956	`HLT_HT260_MHT60_v2`	< 1
160957 – 161176	`HLT_HT260_v2`	5
161217 – 163261	`HLT_HT250_MHT60_v2`	41
163270 – 163869	`HLT_HT250_MHT60_v3`	169
165088 – 165633	`HLT_HT250_MHT70_v1`	136
165970 – 166345	`HLT_HT250_MHT70_v3`	101
166346 – 166346	`HLT_HT250_MHT70_v4`	4
166347 – 166967	`HLT_HT250_MHT70_v3`	437
167039 – 167913	`HLT_HT250_MHT90_v1`	277
170249 – 172619	`HLT_HT250_MHT90_v2`	390
172620 – 175770	`HLT_HT250_MHT90_v2`	707
175832 – 176023	`HLT_HT250_MHT100_v2`	136
176024 – 176470	`HLT_HT300_MHT90_v2`	251
176545 – 178380	`HLT_HT350_MHT90_v1`	1442
178420 – 178708	`HLT_HT350_MHT100_v3`	192
178712 – 180252	`HLT_HT350_MHT110_v3`	689

trigger-threshold effects.

- Remaining QCD-multijet events in which large $\not{H}_{\mathrm{T}}$ is caused by a single mismeasured jet are vetoed by requiring $\Delta\phi(\vec{p}_{\mathrm{T},i},\vec{\not{H}}_{\mathrm{T}}) > 0.5$, $i = 1, 2$, and $\Delta\phi(\vec{p}_{\mathrm{T},3},\vec{\not{H}}_{\mathrm{T}}) > 0.3$, i. e. $\not{H}_{\mathrm{T}}$ must not be aligned in azimuth with the transverse momenta of one of the leading three jets. The threshold of 0.5 is motivated by the jet size parameter; it is reduced in case of the third jet to keep the signal efficiency high.
- The presence of $t\bar{t}$ and $W+\mathrm{jets}$ events with leptons in the final state is suppressed by rejecting events containing isolated muons or electrons with $p_{\mathrm{T}} > 10\,\mathrm{GeV}$ and $|\eta| < 2.4$ or $|\eta| < 2.5$, respectively. The isolation criterion requires the total p_{T} of all hadrons and photons in a conus in $\eta \times \phi$ with radius $\mathrm{R} = 0.3$ around the lepton to be less than 20% of the lepton p_{T} [241, 273].

Dedicated selection criteria (cleaning filters) have been developed to identify events where the reconstruction failed or has been spoiled by instrumental effects or which originated in beam-background processes. In rare cases, these corrupted events feature large $\not{H}_{\mathrm{T}}$ and can populate the search regions. Although their total rate is small, they might substantially falsify the analysis result since the expected rate of signal processes and also SM-background events after the event selection is likewise small. Therefore, these events are not used for further analysis. The cleaning-filter criteria have been described in detail in Section 5.1.4. Approximately 31% of the events passing the baseline selection criteria are affected. Using simulated events, the filters have been found to introduce an inefficiency of 3% for CMSSM signal-events; this is taken into account in the interpretation of the results in Section 7.3. 1885 events pass the baseline selection, which in the following is understood to include the cleaning-filter requirements.

To increase the sensitivity of the search to different possible signal scenarios, 14 exclusive subsamples (*search regions*) are defined in terms of H_T and $\not\!\!H_\mathrm{T}$ as listed in Table 7.2. For SUSY, regions with large values of H_T are in particular populated by events with high multiplicities as in case of long cascade decays where most of the energy is distributed to the strongly interacting particles. Large values of $\not\!\!H_\mathrm{T}$, on the other hand, reduce the SM-background contributions, especially from QCD-multijet events, and increase the sensitivity to potential signal events where large fractions of the energy are transferred to the LSP. The detector signatures of the selected candidate events with the highest H_T and $\not\!\!H_\mathrm{T}$ are visualised in Figs. 7.2 and 7.3, respectively.

The event yields of different SM-background processes after the baseline selection and in the search regions as expected from simulation are listed in Table 7.2. The yields are compared to the number of events observed in data. In order to illustrate the impact from possible new-physics processes, the expectations from the CMSSM benchmark scenario 'LM5' are also quoted, which is defined by $m_0 = 230\,\mathrm{GeV}$, $m_{1/2} = 360\,\mathrm{GeV}$, $A_0 = 0$, $\tan\beta = 10$, and $\mu > 0$ [148]. The $t\bar{t}$, W, and Z events have been generated using the MADGRAPH generator [213] interfaced with the PYTHIA [212] parton-showering model, and the predictions are normalised to the next-to-leading order (NLO) or, if available, next-to-next-to-leading order (NNLO) precision. The QCD-multijet and CMSSM signal events have been simulated with PYTHIA, and the LM5 signal cross-section has been normalised to 0.41 pb as obtained from the PROSPINO programme [275] at NLO precision. Events have been weighted such that the pile-up scenario matches the distribution expected for data, cf. Section 4.3.2. In Fig. 7.4, additionally the H_T and $\not\!\!H_\mathrm{T}$ distributions in data and simulation are shown after the baseline selection.

The simulated SM-background yields are larger than the number of events observed in data in the high-$\not\!\!H_\mathrm{T}$ search regions[1]. This is unexpected given the better jet-p_T resolution in the simulation, cf. Chapter 5, and is attributed to an overprediction of the yields in the probed, extreme phase-space regions in particular by the W + jets simulation. This overprediction has been observed also in other analyses [276, 277]. The differences strongly motivate a background prediction from data, which is described in the following section. Hence, the simulated predictions are intended for illustration purposes only and not used any further in the analysis.

7.2 Estimation of the Standard-Model Backgrounds From Data

Several SM processes with genuine $\not\!\!E_\mathrm{T}$ caused by neutrinos or instrumental effects constitute large backgrounds to the selected samples. Their contributions can be estimated using simulated events, but these results should not be trusted to the level of accuracy required given that no measurements exist in the probed phase-space regions to verify

---

[1] It should be kept in mind, however, that only the statistical uncertainties of the simulation-based predictions have been determined.

Table 7.2: Number of events in observed in data and expected from simulation after the baseline selection and in the search regions. Individual contributions from different SM-background processes are quoted. ('Other' refers to $WW, WZ, ZZ+$jets events.) They have been obtained using the Pythia generator in case of QCD and the Madgraph generator otherwise and have been normalised to an integrated luminosity of $4.98\,\mathrm{fb}^{-1}$ at NLO ($t\bar{t}$, Other) or at NNLO ($Z \to \nu\bar{\nu} +$ jets, $W+$jets) accuracy. The stated uncertainties are statistical only. For illustration purposes, the expected number of signal events for the CMSSM parameter point LM5 defined in the text are also stated. Their cross section has been computed at NLO precision using the Prospino programme. The predicted event yields are stated for illustration purposes only. Taken from [274].

Selection		Data	Simulation							
H_T (GeV)	$\not{H}_\mathrm{T}$ (GeV)		Sum SM	QCD	$Z \to \nu\bar{\nu}$	$W +$ jets	$t\bar{t}$	Single top	Other	LM5
Baseline		1885	2064	45	744	733	491	40	11	842
500 – 800	200 – 350	1269	1365 ± 36	13.1 ± 3.8	447 ± 6	499 ± 34	369 ± 10	29.2 ± 1.3	7.6 ± 0.6	129
500 – 800	350 – 500	236	251 ± 14	1.2 ± 1.0	131 ± 3	79 ± 14	34.1 ± 3.0	2.9 ± 0.4	2.1 ± 0.3	110
500 – 800	500 – 600	22	45 ± 6.0	—	25.3 ± 1.5	15.8 ± 6	2.9 ± 0.9	0.4 ± 0.1	0.5 ± 0.1	43.2
500 – 800	> 600	6	13 ± 3.9	—	7.1 ± 0.8	5.4 ± 3.8	—	—	—	22.5
800 – 1000	200 – 350	177	172 ± 12	13.6 ± 3.6	56.8 ± 2.2	50 ± 11	47.8 ± 3.6	3.5 ± 0.5	0.3 ± 0.1	93.7
800 – 1000	350 – 500	24	51 ± 8	—	17.8 ± 1.2	23.9 ± 7.8	8.4 ± 1.5	0.9 ± 0.2	0.2 ± 0.1	82.4
800 – 1000	500 – 600	6	16 ± 5	—	5.8 ± 0.7	9.7 ± 4.9	0.3 ± 0.3	0.3 ± 0.1	0.1 ± 0.1	37.1
800 – 1000	> 600	5	12.4 ± 3.6	—	6.5 ± 0.8	4.8 ± 3.5	0.6 ± 0.5	0.2 ± 0.1	0.1 ± 0.1	38.4
1000 – 1200	200 – 350	71	61 ± 8.1	6.8 ± 2.5	17.4 ± 1.2	21.7 ± 7.3	13.6 ± 1.9	1.2 ± 0.3	—	61.3
1000 – 1200	350 – 500	12	20.0 ± 5.3	0.2 ± 0.1	6.2 ± 0.7	10.3 ± 5.2	2.9 ± 0.9	0.4 ± 0.1	—	48.8
1000 – 1200	> 500	4	7.3 ± 2.2	—	4.4 ± 0.7	2.1 ± 2.1	0.5 ± 0.4	0.1 ± 0.1	—	42.4
1200 – 1400	200 – 350	29	22.1 ± 4.0	2.8 ± 0.9	7.1 ± 0.8	5.5 ± 3.6	6.3 ± 1.3	0.4 ± 0.2	—	31.4
1200 – 1400	> 350	8	5.1 ± 0.7	0.1 ± 0.1	4.0 ± 0.6	—	0.9 ± 0.4	0.1 ± 0.1	—	43.3
> 1400	> 200	16	22.3 ± 4.2	7.2 ± 1.8	6.3 ± 1.2	5.1 ± 3.5	3.5 ± 0.9	0.2 ± 0.1	—	57.9

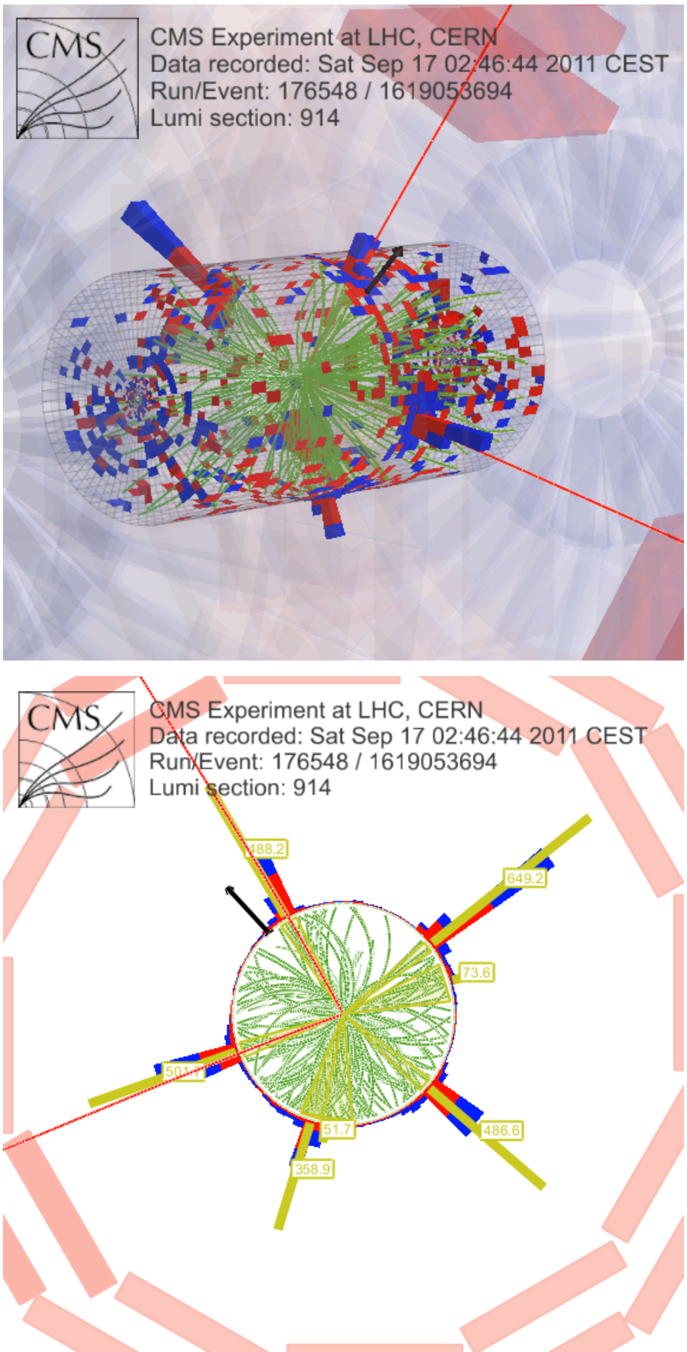


Figure 7.2: Visualisation of the detector signatures caused by the selected candidate event with the highest H_{T}: it has $H_{\mathrm{T}} = 2577\,\mathrm{GeV}$ and $\not{H}_{\mathrm{T}} = 212\,\mathrm{GeV}$. Solid yellow lines represent PF jets with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ with their p_{T} stated in the boxes. (These are the p_{T} before application of the jet energy calibration, which is why their sum of does not amount exactly to the H_{T} cited above.) The $\not{H}_{\mathrm{T}}$ vector is indicated by the black arrow, muon tracks are highlighted in red, other detector-signal representations are as in Fig. 4.5.

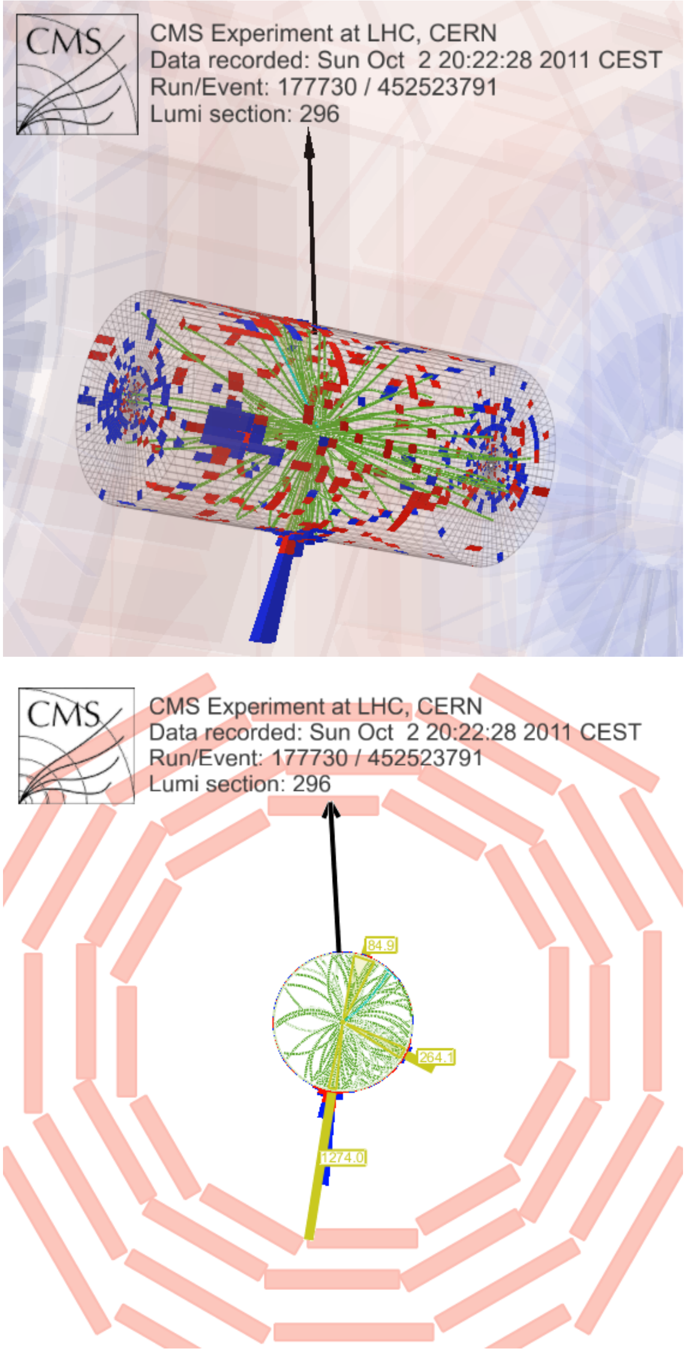


Figure 7.3: Visualisation of the detector signatures caused by the selected candidate event with the highest $\not\!\!H_{\mathrm{T}}$: it has $\not\!\!H_{\mathrm{T}} = 1282\,\mathrm{GeV}$ and $H_{\mathrm{T}} = 1592\,\mathrm{GeV}$. Solid yellow lines represent PF jets with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ with their p_{T} stated in the boxes. (These are the p_{T} before application of the jet energy calibration, which is why their sum of does not amount exactly to the H_{T} cited above.) The $\not\!\!E_{\mathrm{T}}$ vector is indicated by the black arrow, muon tracks are highlighted in red, other detector-signal representations are as in Fig. 4.5.

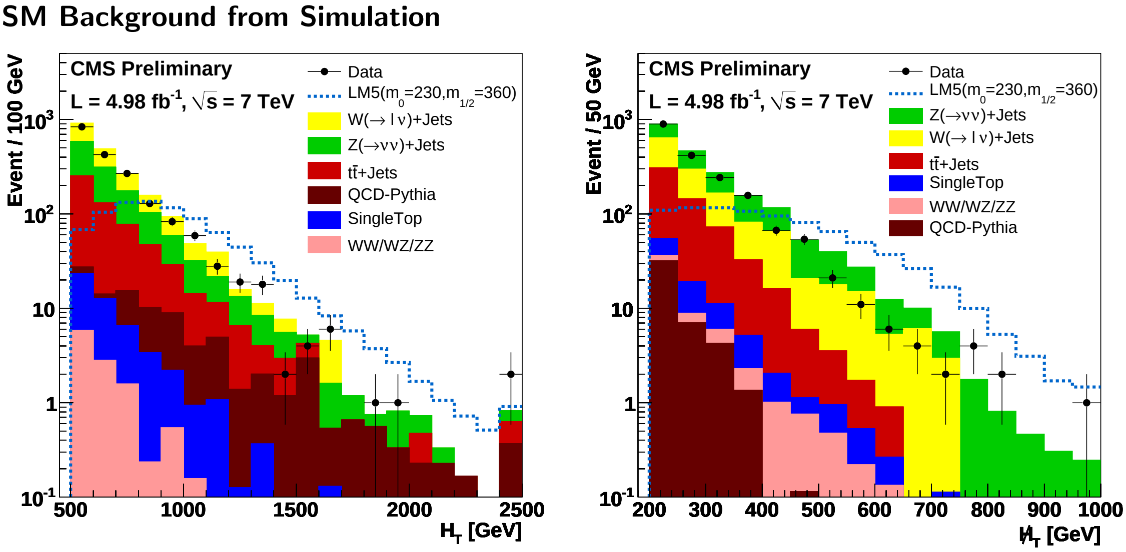


Figure 7.4: H_T (*left*) and $\not\!\!H_\mathrm{T}$ (*right*) distributions after the baseline selection in data (*circles*) compared to the contributions from different SM-background processes predicted from simulation (*stacked histograms*), which have been obtained using the Pythia generator in case of QCD and the Madgraph generator otherwise. The simulated yields are normalised to an integrated luminosity of $4.98\,\mathrm{fb}^{-1}$ at NLO or, if available NNLO, accuracy. For illustration purposes, the expected number of signal events for the CMSSM parameter point LM5 defined in the text are also stated. These events have been generated using the Prospino programme. The last bin contains all events above the maximum value of H_T and $\not\!\!H_\mathrm{T}$ in the figures. The figure is intended for illustration purposes only. Taken from [257].

the simulation. In particular, modelling QCD-multijet events is challenging and thus associated with large uncertainties because the underlying QCD-production processes are theoretically difficult to compute and the simulation of jet measurements requires a precise understanding of detector components and effects. Hence, it is essential to either improve the simulation or provide independent methods to predict the SM-background contributions. Therefore, techniques have been developed to model the full H_T and $\not\!\!H_\mathrm{T}$ spectra of the SM-background processes in the search regions using measured data. These techniques are described in detail in [237, 238, 260]

First, the prediction of the non-QCD backgrounds is briefly reviewed in Section 7.2.1, and then, in Section 7.2.2, the QCD-background estimation is presented in more detail.

7.2.1 Non-QCD Backgrounds

Events with jets and a Z boson produced in association which subsequently decays into two neutrinos form an irreducible SM background to the selected data sample.

An estimation of the background yield is obtained by predicting the kinematic properties of $Z \to \nu\bar{\nu}$ + jets events from a control sample of photon + jets events [237] The neutral electroweak bosons are produced in similar processes, and at p_T larger than the Z-boson mass, the cross-section ratio depends approximately on the different size of the electroweak couplings with asymptotically vanishing mass effects.

Data are collected with a single-photon and, during the latest part of the data-taking period, a photon + H_{T} composite trigger. The single-photon trigger requires at least one photon with $p_{\mathrm{T}} > 70\,\mathrm{GeV}$ during the first and $p_{\mathrm{T}} > 90\,\mathrm{GeV}$ during the later part of data taking; the composite trigger requires at least one photon with $p_{\mathrm{T}} > 70\,\mathrm{GeV}$ and $H_{\mathrm{T}}^{\mathrm{HLT}} > 400\,\mathrm{GeV}$. Photons are distinguished from electrons by requiring that the candidate must not have an associated track segment in the pixel detector. Contributions to the control sample from QCD-multijet events arise from photons that are produced inside jets either during parton showering or through decays of neutral pions. They are suppressed by imposing pile-up subtracted isolation and shower-shape criteria to the photon [168, 242]. Residual contamination with parton-showering photons is estimated with the Jetphox programme using NLO calculations [278, 279]; contamination with pion-decay photons is determined using an isolation-template method in data [280]. The photon selection requirements have also been found to sufficiently suppress a contamination of the control sample by possible CMSSM-signal events.

Finally, the background prediction is obtained by removing the photon from the event reconstruction, applying the jet-related selection requirements to the control sample, and scaling the resulting number of photon + jets events to account for the cross-section ratio $\sigma(Z \to \nu\bar{\nu} + \text{jets})/\sigma(\text{photon} + \text{jets})$ as well as acceptance and reconstruction efficiencies, which are determined from a Madgraph simulation. The dominant uncertainties of the method are an uncertainty of $21 - 42\%$ of the cross-section ratio, an uncertainty of 5% of the detector acceptance, and the statistical uncertainty of $3 - 59\%$ due to the limited size of the control sample.

A second $Z \to \nu\bar{\nu}$ + jets background prediction is obtained using $Z \to \mu^+\mu^-$ + jets events where the muons are removed from the reconstruction in order to mimic the neutrino-induced $\not{H}_{\mathrm{T}}$ in $Z \to \nu\bar{\nu}$ + jets events. After corrections for the acceptance, efficiency, and different branching ratios, the expected $Z \to \nu\bar{\nu}$ + jets background yield is obtained by applying the jet-related selection criteria. Although straightforward, this approach is limited by relatively large uncertainties due to the low production rate of $Z \to \mu^+\mu^-$ + jets events compared to $Z \to \nu\bar{\nu}$ + jets events. Therefore, it is only used as a cross-check, which yields compatible results to the primary method described before.

Events from $t\bar{t}$ and W + jets production can fake the signal topology, if the W bosons decay into leptons and neutrinos thus inducing $\not{E}_{\mathrm{T}}$. They can pass the event selection either if a τ lepton decays hadronically (*hadronic-τ background*) or if the veto of a light lepton fails (*lost-lepton background*) because it is out of geometric or kinematic acceptance, is not identified, or is not isolated. Both backgrounds are predicted using a μ + jets control sample [237]

In case of the lost-lepton background, the control sample is selected by applying the full event selection but without the veto of light leptons. Instead, exactly one well-identified and well-isolated muon is required, defined by identical criteria as for the lepton veto Section 7.1. Moreover, only events with a *transverse mass* $m_{\mathrm{T}} = \sqrt{2\,p_{\mathrm{T}}(\mu)\not{E}_{\mathrm{T}}[1 - \cos(\Delta\phi)]}$ of less than $100\,\mathrm{GeV}$ are selected, with $\Delta\phi$ being the difference in azimuthal angle between the muon momentum and the $\not{E}_{\mathrm{T}}$ vector. This criterion enhances the fraction of $t\bar{t}$ and W + jets events in the control sample and sufficiently suppresses the presence of possible CMSSM-signal events. The lost-lepton contribution to the search regions is modelled by weighting the number of events in the control sample according to the lepton isolation and identification efficiencies as well as acceptance factors. The isolation efficiency is determined separately for electrons and muons from $Z \rightarrow e^+e^-, \mu^+\mu^- +$ jets data, respectively, using a tag-and-probe method [281]. It is parametrised as a function of the lepton p_T and the distance $\Delta\mathrm{R}$ in $\eta \times \phi$ between the lepton and the nearest jet in order to take into account kinematic differences to $t\bar{t}$ and $W +$ jets events. Residual differences have been studied using simulated $t\bar{t}$ and $W +$ jets events generated with MADGRAPH. They have been found to result in an underprediction of the background by about 4% in all search regions, which is applied as additional correction to the final prediction. The identification efficiency and acceptance factor are determined directly from the simulated $t\bar{t}$ and $W +$ jets events. In most search regions, the dominant uncertainty of the prediction is due to the limited size of the $\mu +$ jets control samples and reaches up to 100%. Further major contributions arise from a 10% uncertainty assigned to the isolation-efficiency parametrisation and a 4 – 20% uncertainty based on the limited number of simulated events used to validate the method. The method is described in detail also in [282].

In case of the hadronic-τ background, the control sample is selected by triggering on at least one muon with $p_\mathrm{T} > 17\,\mathrm{GeV}$ and, during the later part of the data-taking period, on additionally two central jets with $p_\mathrm{T} > 30\,\mathrm{GeV}$. Offline, exactly one well-identified and well-isolated muon with $p_\mathrm{T} > 20\,\mathrm{GeV}$ and $|\eta| < 2.1$ is required, applying identical lepton-selection criteria as in Section 7.1. Again, the transverse-mass requirement is imposed to prevent signal contamination of the control sample. Due to the lepton universality of the weak interaction, the kinematic properties of the non-τ objects are the same for the events in the control sample and for the hadronic-τ background events. Hence, the τ jet is modelled by scaling the muon's four-momentum with a random factor obtained from a τ-p_T response template, which has been determined from MADGRAPH-generated $t\bar{t}$ and $W +$ jets events. After recomputation of H_T and $\not{H}_\mathrm{T}$, where the modelled τ jet is taken into account, the event selection is applied to this modified control sample. Given the limited number of events in the control sample, the sampling step is repeated 100 times per event in order to probe the full response template[2]. Correlations between the resulting 100 emulated hadronic-τ events per control-sample event are taken into account by incorporation of a bootstrap technique [283] when computing the statistical uncertainties of the prediction. The hadronic-τ background contribution to the search regions is predicted by scaling the event yield in the control sample with the branching fraction $\mathcal{B}(W \rightarrow \tau\nu(\tau \rightarrow \mathrm{hadrons}))/\mathcal{B}(W \rightarrow \mu\nu) = 0.69 \pm 0.05$ of W bosons decaying into hadronically decaying τ leptons and into muons. Corrections are applied to take into account the muon geometric and kinematic acceptance, which is derived from a MADGRAPH simulation, as well as the trigger, identification, and isolation efficiencies, which are measured from $Z +$ jets data using a tag-and-probe method [281]. Trigger and identification efficiencies are parametrised as a function of lepton p_T and η, the isolation efficiency as above as a function of lepton p_T and $\Delta\mathrm{R}$ to the nearest jet. A 4% overprediction in all search regions has been observed when testing the method with simulated $t\bar{t}$ and $W +$ jets events generated with MADGRAPH, which is considered as an additional correction to the final prediction. The dominant uncertainties of the method arise from the τ-p_T response template (1 – 20%), the acceptance correction (6 – 13%), and the limited number of simulated events used to validate the method (4 – 12%).

---

[2] Prescaled control-sample triggers are taken into account by increasing the number of sampling steps according to the prescale factor.

7.2.2 QCD-Multijet Background

In QCD-multijet events, large values of H_{T} occur naturally while $\not{H}_{\mathrm{T}}$ arises from mismeasurements of the jet momenta. As discussed in previous chapters, the jet-p_{T} response function is dominated by an almost Gaussian core. Severe jet-p_{T} mismeasurements and thus sizable values of $\not{H}_{\mathrm{T}}$ are typically caused predominantly by the additional non-Gaussian tails, which contribute only at the few-percent level or below to the response distribution. Hence, QCD-multijet events mostly populate search regions with high H_{T} but moderate $\not{H}_{\mathrm{T}}$ requirements. Here, they constitute an important background which is expected to be of the order of the other SM backgrounds, cf. Table 7.2. Lack of precision in the QCD-background estimation therefore particularly degrades the sensitivity of the analysis to CMSSM scenarios with large values of m_0. In this case, the masses of squarks are high compared to the LSP mass, and hence, H_{T} is large because a large fraction of energy is transferred to detectable particles.

The QCD background is measured from multijet data using the *rebalance-and-smear* (R+S) method, which has been first developed for the 2010 version of the jets + $\not{H}_{\mathrm{T}}$ analysis [246, 260] and improved for an update in early 2011 [238]; the reviewed version of the method is described in [237] The principles of the R+S method are explained in Section 7.2.2.1, and its validation is reviewed in Section 7.2.2.2. In Section 7.2.2.3, details of the application and the resulting background prediction are described. Systematic uncertainties are discussed in Section 7.2.2.4.

The QCD-multijet background prediction is cross-checked with an alternate method which employs the correlation between H_{T} and an angular variable of $\not{H}_{\mathrm{T}}$ and the leading three jets. Using this correlation, the event yields observed in a control region are extrapolated into the search regions to obtain a prediction. The results are compatible with the R+S prediction. A detailed description of this method can be found in [260, 274, 284].

7.2.2.1 Description of the Rebalance-And-Smear Method

The R+S method can be regarded as a data-based fast simulation: an estimator of $\not{H}_{\mathrm{T}}$ is obtained by first rebalancing the jet momenta in a QCD-multijet event, taking into account the jet-p_{T} resolution, and then weighting the rebalanced momenta according to the full jet-p_{T} response function in order to predict the complete multijet-event kinematics.

The 'rebalance' step: In the rebalance step, a sample of *seed events* is acquired from multijet data. The measured jet momenta are adjusted, resulting in *seed jets*, to regain approximate transverse-momentum balance as expected for QCD-multijet events at particle level. Technically, this is achieved with a *kinematic fit* [285], a least-square fit of the jets' p_{T}, η, and ϕ, where the kinematic constraints for transverse-momentum balance are considered by means of Lagrangian multipliers. The jet-p_{T} measurement uncertainties are approximated by the Gaussian MC-truth resolution since the rebalancing procedure is insensitive to the tiny, non-Gaussian tails. Importantly, contributions to the seed sample from the $t\bar{t}$ or electroweak backgrounds as well as possible signal processes are negligible because also these events will be transformed into QCD-like events with balanced transverse momenta and their production cross-section in the control region is smaller by orders of magnitude than the QCD-multijet cross-section.

The rebalancing procedure is in fact essential for a correct QCD-background prediction. It is not sufficient to simply select a seed sample of events with low $\not{H}_\mathrm{T}$. The jets in the selected events would be biased towards low p_T because $\not{H}_\mathrm{T}$ and H_T are correlated, and the predicted $\not{H}_\mathrm{T}$ would be underestimated [246].

The 'smear' step: In the smear step, the four-momentum of each seed jet is scaled by a random factor sampled from the full MC-truth jet-p_T response distribution, thus modelling effects of the measurement and simulating the expected detector-level jet. The kinematic properties of the original QCD-multijet events are preserved in the obtained sample of events with smeared jets. Hence, the expected distributions of H_T and $\not{H}_\mathrm{T}$ as well as other jet-related variables in the search region can be predicted after applying the full event selection. In order to probe the entire response distribution, in particular also the non-Gaussian tails, and thus improve the stability of the prediction with respect to statistical fluctuations, the smear step is repeated 100 times. As a result, a distribution of 100 event-yield predictions is obtained. Effects due to the correlated seeds are taken into account by incorporation of a bootstrap technique [283]: the mean and the standard deviation of that event-yield distribution are quoted as central prediction and as statistical uncertainty of the R+S method, respectively.

The distributions of the MC-truth jet-p_T response $\mathcal{R}$ Eq. (4.3) are crucial input to the R+S method. They are determined as explained in Section 4.6.1 from a PYTHIA-generated QCD-multijet sample for the three leading generator-level jets in each event in intervals of $p_\mathrm{T}^\mathrm{gen}$ and $|\eta^\mathrm{gen}|$.

Only jets with $p_\mathrm{T} > 15\,\mathrm{GeV}$ are taken into account for the rebalance step in order to reduce the impact from pile-up interactions, which produce a large number of low-p_T particles not originating from the considered hard interaction. If some of these particles are clustered into jets or form extra jets they might spoil the transverse-momentum balance. Possible biases introduced by this requirement have been observed to affect the final prediction by less than 20% and are considered in the uncertainties.

7.2.2.2 Validation of the Rebalance-And-Smear Method

The performance of the R+S method has been studied using a sample of PYTHIA-simulated QCD-multijet events. In Fig. 7.5, the predicted $\not{H}_\mathrm{T}$ distributions are shown after a modified baseline selection where the $\not{H}_\mathrm{T}$ criterion has been dropped and different, tightened H_T criteria are required. The latter correspond to the lower H_T boundaries of the exclusive search regions. The predicted $\not{H}_\mathrm{T}$ distributions agree to the distributions expected from the full simulation within the statistical uncertainties.

Depending on the search region, the central value of the predicted number of events deviates from the simulated one by up to 25%. Although the differences are statistically not significant, they are applied as bias correction factors to the final predictions.

Further cross checks have been performed to verify the robustness of the R+S method. Firstly, the impact of intrinsic biases has been studied by repeated application of the method to simulated QCD-multijet events. The event-yield prediction $N^1_\mathrm{R+S}$ of the first iteration is compared to the prediction $N^2_\mathrm{R+S}$ of the second iteration. Relative differences are below 10% and within the statistical uncertainties in all search regions. Secondly, the method has been validated directly on data. A QCD-dominated sample is selected

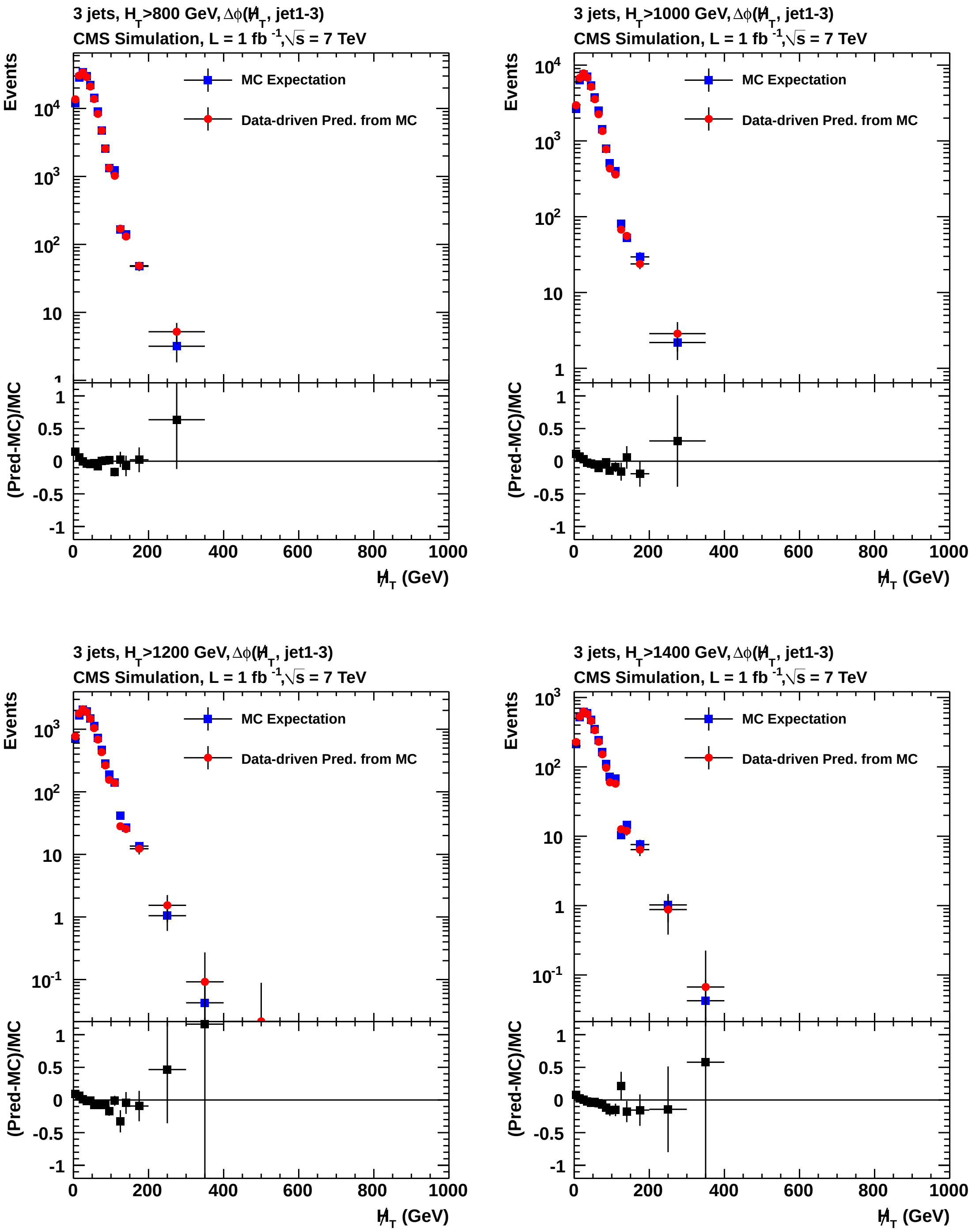


Figure 7.5: Validation of the R+S method with Pythia-simulated events: the predicted $\not{H}_\mathrm{T}$ distributions (*circles*) are compared to the expectation from the full simulation (*squares*). The events pass a modified baseline selection where the $\not{H}_\mathrm{T}$ criterion has been dropped and different H_T criteria are applied as stated in the legends. Taken from [257, 274].

by applying the baseline selection without the H_{T}, $\not{H}_{\mathrm{T}}$, and $\Delta\phi$ criteria. The predicted p_{T} spectra of the leading three jets are in good agreement to the measured distributions, differences are again at the 10% level.

7.2.2.3 QCD-Background Prediction

In contrast to the search sample, the multijet-data input to the R+S method has been collected by triggering on $H_{\mathrm{T}}^{\mathrm{HLT}}$ only with no additional requirement on $\not{H}_{\mathrm{T}}$, which would bias the seed sample as remarked previously. However, the lowest unprescaled trigger path has a threshold of $H_{\mathrm{T}}^{\mathrm{HLT}} = 500\,\mathrm{GeV}$ or higher. Since the baseline selection already requires $H_{\mathrm{T}} > 500\,\mathrm{GeV}$ and since in course of the R+S procedure events can migrate in H_{T}, additional, prescaled trigger paths with lower thresholds are employed. The prescale factor n is taken into account by effectively weighting the event's contribution to the prediction accordingly: in each of the 100 smear steps, an event is smeared n times.

As shown in Chapters 5 and 6, the simulated jet-p_{T} response function differs from the data. Therefore, the MC-truth response $\mathcal{R}$ distributions utilised in the R+S method are adjusted to match the data. First, the $\mathcal{R}$ distributions are broadened by convolution with a Gaussian as described in Section 5.7.1 to account for the larger resolution in data. Afterwards, tails are scaled as described in Section 6.6.2.

The final R+S QCD-background prediction is listed in Table 7.3 together with the assigned uncertainties. As mentioned before, important QCD contributions arise in the search regions with high H_{T} and low $\not{H}_{\mathrm{T}}$ ($< 350\,\mathrm{GeV}$), cf. also Table 7.4: they are of same size as the other SM-background contributions for $H_{\mathrm{T}} > 800\,\mathrm{GeV}$ and dominate for $H_{\mathrm{T}} > 1200\,\mathrm{GeV}$. In search regions with $\not{H}_{\mathrm{T}} > 350\,\mathrm{GeV}$ in contrast, QCD contributions are below 5%.

The R+S predictions are larger by up to 1.5 standard deviations than the predictions in Table 7.2 obtained with the PYTHIA simulation. This is a consequence of the larger jet-p_{T} resolution and response tails in data measured in Chapters 5 and 6. If the jet-p_{T} is measured less accurate than in the simulation large values of $\not{H}_{\mathrm{T}}$ occur at a higher rate, and thus, more QCD events are promoted into the search regions. Furthermore, QCD-multijet production is particularly difficult to model and differences to the data are generally not unexpected.

The systematic uncertainties are discussed in the following Section 7.2.2.4. In the high-$\not{H}_{\mathrm{T}}$ search regions, the statistical uncertainties are larger than 100%. This is an artefact of the employed bootstrap technique. Since the non-Gaussian tails of the jet-p_{T} response are tiny, only a small fraction of the smeared events enters the high-$\not{H}_{\mathrm{T}}$ search regions; in most sampling steps, the predicted event yield is zero. As a consequence, the resulting distributions of event-yield predictions become highly non-Gaussian. Moreover, since the number of seed events is limited, the distributions are also highly discontinuous. Hence, their standard deviations — and therefore the statistical uncertainties of the central predictions — are artificially increased. This is acceptable, though, because the absolute number of expected QCD events is small compared to the size of other SM-background contributions in these regions.

Table 7.3: QCD-background prediction obtained with the R+S method: number of events expected for an integrated luminosity of $4.98\,\mathrm{fb}^{-1}$ in the different search regions together with the statistical ('stat.') and systematic ('syst.') uncertainties. The relative contributions to the latter from the intrinsic R+S bias ('bias') and from the uncertainties due to pile-up ('PU'), the MC-truth response distributions ('$\mathcal{R}$'), and the response-core and -tail correction ('ρ_{res}', 'ρ_{tail}') are also stated. Uncertainties larger than 100% are artefacts of the bootstrap technique as explained in the text. Taken from [274].

Search Region		Prediction			Relative Contribution to Syst.				
H_{T} (GeV)	$\not{H}_{\mathrm{T}}$ (GeV)	yield	stat.	syst.	bias	PU	$\mathcal{R}$	ρ_{res}	ρ_{tail}
500 – 800	200 – 350	118.6	±12	$^{+76}_{-76}$	43%	29%	10%	$^{+14\%}_{-10\%}$	$^{+33\%}_{-34\%}$
500 – 800	350 – 500	2.2	±1.6	$^{+1.8}_{-1.2}$	43%	29%	10%	$^{+40\%}_{-11\%}$	$^{+46\%}_{-3\%}$
500 – 800	500 – 600	0.02	±0.14	$^{+0.02}_{-0.02}$	43%	29%	10%	$^{+50\%}_{-50\%}$	$^{+50\%}_{-100\%}$
500 – 800	> 600	—	—	—	—	—	—	—	—
800 – 1000	200 – 350	34.6	±5.3	$^{+23}_{-24}$	40%	40%	10%	$^{+14\%}_{-12\%}$	$^{+32\%}_{-34\%}$
800 – 1000	350 – 500	1.2	±1.1	$^{+0.7}_{-0.8}$	40%	40%	10%	$^{+14\%}_{-25\%}$	$^{+5\%}_{-34\%}$
800 – 1000	500 – 600	0.03	±0.17	$^{+0.04}_{-0.03}$	40%	40%	10%	$^{+33\%}_{-0\%}$	$^{+100\%}_{-100\%}$
800 – 1000	> 600	0.01	±0.10	$^{+0.01}_{-0.02}$	40%	40%	10%	$^{+100\%}_{-100\%}$	$^{+0\%}_{-100\%}$
1000 – 1200	200 – 350	19.7	±4.4	$^{+13}_{-12}$	31%	40%	10%	$^{+19\%}_{-13\%}$	$^{+37\%}_{-29\%}$
1000 – 1200	350 – 500	0.44	±0.61	$^{+0.30}_{-0.26}$	31%	40%	10%	$^{+23\%}_{-9\%}$	$^{+39\%}_{-30\%}$
1000 – 1200	> 500	0.04	±0.2	$^{+0.03}_{-0.05}$	31%	40%	10%	$^{+0\%}_{-100\%}$	$^{+50\%}_{-25\%}$
1200 – 1400	200 – 350	11.7	±3.4	$^{+7.5}_{-7.6}$	34%	39%	10%	$^{+20\%}_{-24\%}$	$^{+32\%}_{-29\%}$
1200 – 1400	> 350	0.24	±0.53	$^{+0.15}_{-0.20}$	34%	39%	10%	$^{+33\%}_{-33\%}$	$^{+4\%}_{-54\%}$
> 1400	> 200	12.0	±3.8	$^{+8.4}_{-8.1}$	47%	36%	10%	$^{+23\%}_{-17\%}$	$^{+28\%}_{-27\%}$

7.2.2.4 Systematic Uncertainties of the Rebalance-And-Smear Method

The following sources of systematic uncertainties of the R+S method have been investigated:

Intrinsic bias: The R+S method has been validated with simulated events. Relative differences δ between the predicted and the true number of events are below 25% in all search regions, cf. Section 7.2.2.2. They are attributed to the following sources:

- The jet-p_{T} threshold of 15 GeV in the rebalance step biases the prediction. It has been varied between 10 and 20 GeV, resulting in a variation of the final prediction by less than 20%.

- The MC-truth response distributions are recorded in finite intervals of p_{T} and η. The impact of this sampling has been evaluated by smearing generator-level jets, i. e. by omitting the rebalance step and using generator-truth information instead, and the obtained predictions differ by less than 10% from the nominal results.
- Migration effects occur due to the steeply falling jet-p_{T} spectrum in combination with the finite jet-p_{T} resolution. They have been found to impact the final prediction by less than 5% [260].

The differences are statistically compatible with $\delta = 0$ in most cases. Therefore, an overall systematic uncertainty of $\max(\delta, \sigma_\delta)$, where σ_δ denotes the statistical uncertainty of δ, is assigned to account for the intrinsic biases. It amounts to 31 – 47% and covers the aforementioned effects.

Pile-up: A jet-p_{T} threshold of 15 GeV is applied in the rebalance step in order to minimise the impact from pile-up collisions on the R+S method. Since the effect of pile-up is included in the simulation and since the method yields consistent results when being applied to the simulation, the final prediction is considered to be pile-up insensitive to first approximation. To evaluate residual dependences due to incorrectly modelled pile-up effects, the method is performed in data and in simulation for exclusive subsamples containing events with a low ($N_{\mathrm{Vtx}} < 6$), a medium ($6 \leq N_{\mathrm{Vtx}} < 11$), and a high ($N_{\mathrm{Vtx}} \geq 11$) number N_{Vtx} of reconstructed primary vertices. The fraction δ of the predicted yields in data and in simulation observed in the sample with $N_{\mathrm{Vtx}} < 6$ is then used to scale the respective fractions in the other subsamples. Finally, the sum of the scaled predictions in the simulation is compared to the sum of the predictions in data, and the relative difference is considered as systematic uncertainty, which is between 29 and 40%.

Response distributions: The jet-p_{T} response distributions have been determined using the Pythia event generator. Uncertainties related to the assumed parton showering and flavour composition of jets are evaluated by applying the R+S prediction to Pythia-simulated QCD-multijet events and comparing to the results obtained when using response distributions from Madgraph. Differences within 10% have been observed [260]. For the final prediction, the shape of the response distributions has been adjusted to the data using the data-to-simulation ratios ρ_{res} and ρ_{tail} measured in Chapters 5 and 6. Their uncertainties are propagated to the final result by repeating the R+S method with the values of ρ_{res} and ρ_{tail} varied by their uncertainties. In search regions with sizable QCD contributions, i. e. for $200 < \not{H}_{\mathrm{T}} < 350$ GeV, the resulting variations of the predictions are between 10 and 25% in case of the resolution (ρ_{res}) and between 30 and 40% in case of the tails (ρ_{tail}).

Triggers: Since prescales of the seed-sample triggers are explicitly considered in the rebalance step, possible biases of the R+S prediction are avoided. Hence, no additional uncertainty is assigned.

The relative size of each uncertainty component is listed in Table 7.3. They are roughly of the same size, between 10 and 50% in the search regions with low and medium values of $\not\!\!H_{\mathrm{T}}$. The total systematic uncertainty, calculated as the quadratic sum of the individual contributions, varies between 64 and 86%. It increases to 100% and more in the high-$\not\!\!H_{\mathrm{T}}$ search regions. However, these values are dominated by statistical fluctuations of the varied predictions, i. e. the statistical uncertainty on the systematic uncertainty, which are to large extents caused by the applied bootstrap technique as discussed above for the statistical uncertainties. Again, this is acceptable because the QCD contributions are negligible in these regions.

7.3 Results and Interpretation

The number of events observed in $4.98\,\mathrm{fb}^{-1}$ of data in the different search regions and the expected SM-background event yields predicted by the data-based methods described above are listed in Table 7.4. In Fig. 7.6, the H_{T} and $\not\!\!H_{\mathrm{T}}$ distributions observed in data after the baseline selection are compared to the SM-background predictions. The QCD-background prediction has a typical uncertainty between 60 and 70% in search regions with significant QCD contributions.

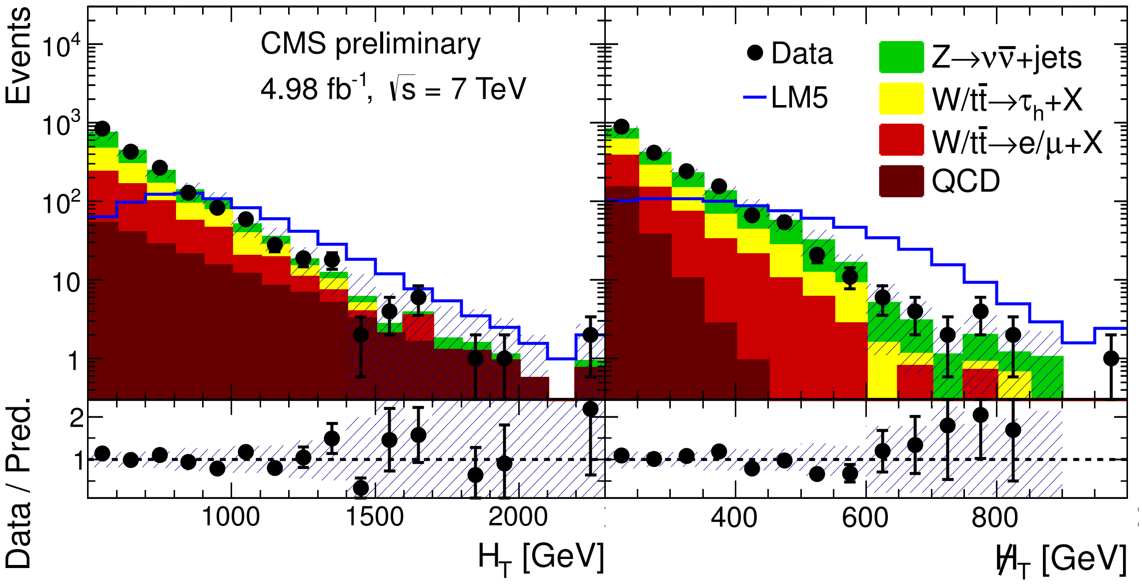


Figure 7.6: H_{T} (*left*) and $\not\!\!H_{\mathrm{T}}$ (*right*) distributions after the baseline selection in data (*circles*) compared to the data-based SM-background (*stacked histograms*) and SUSY LM5 signal predictions. The SM-background distributions are obtained as explained in Section 7.3.1 employing the information in Table 7.4; uncorrelated statistical and fully correlated systematic uncertainties are assumed for their combination. The last bin contains all events above the maxium value of H_{T} and $\not\!\!H_{\mathrm{T}}$ in the figures. The figure has been obtained from the data-based background predictions as explained in Section 7.3.1.

The data are consistent with the SM expectation over the entire range of H_{T} and $\not\!\!H_{\mathrm{T}}$. Systematic overprediction as obtained with the simulation in Table 7.2 is not present, and also the relative size of the background components is different for the data-based

Table 7.4: Number of events predicted for the different SM backgrounds by the data-based methods after the baseline selection and in the search regions for an integrated luminosity of $4.98\,\mathrm{fb}^{-1}$. The total SM-background prediction is obtained assuming uncorrelated statistical and fully correlated systematic uncertainties. It is compared to the number of events observed in data. Taken from [237].

Selection H_T (GeV)	 $\not{H}_\mathrm{T}$ (GeV)	QCD multijets	$Z \rightarrow \nu\bar{\nu} + \mathrm{jets}$	$t\bar{t}/W \rightarrow e, \mu + \mathrm{X}$	$t\bar{t}/W \rightarrow \tau_\mathrm{h} + \mathrm{X}$	Total SM background	Data
Baseline selection		200.7 ± 82.7	596.6 ± 165.0	475.8 ± 76.2	522.5 ± 67.0	1795.4 ± 205.9	1885
500 – 800	200 – 350	118.6 ± 76.9	359.2 ± 82.2	326.5 ± 47.0	348.5 ± 40.1	1152.8 ± 128.4	1269
500 – 800	350 – 500	2.2 ± 2.2	112.3 ± 27.4	47.8 ± 9.2	62.5 ± 8.7	224.8 ± 30.3	236
500 – 800	500 – 600	0.0 ± 0.1	17.6 ± 5.6	5.0 ± 2.2	8.7 ± 2.5	31.3 ± 6.5	22
500 – 800	> 600	0.0 ± 0.0	5.5 ± 3.1	0.8 ± 0.8	2.0 ± 1.8	8.3 ± 3.6	6
800 – 1000	200 – 350	34.6 ± 24.0	48.4 ± 19.1	57.7 ± 15.3	56.3 ± 8.3	197.0 ± 35.3	177
800 – 1000	350 – 500	1.2 ± 1.3	16.0 ± 7.3	5.4 ± 2.3	7.2 ± 2.0	29.8 ± 8.0	24
800 – 1000	500 – 600	0.0 ± 0.2	7.1 ± 4.5	2.4 ± 1.5	1.3 ± 0.6	10.8 ± 4.8	6
800 – 1000	> 600	0.0 ± 0.1	3.3 ± 2.0	0.7 ± 0.7	1.0 ± 0.3	5.0 ± 2.2	5
1000 – 1200	200 – 350	19.7 ± 13.3	10.9 ± 5.5	13.7 ± 3.8	21.9 ± 4.6	66.2 ± 15.5	71
1000 – 1200	350 – 500	0.4 ± 0.7	5.5 ± 3.5	5.0 ± 4.4	2.9 ± 1.3	13.8 ± 5.8	12
1000 – 1200	> 500	0.0 ± 0.2	2.2 ± 2.9	1.6 ± 1.2	2.3 ± 1.0	6.1 ± 3.3	4
1200 – 1400	200 – 350	11.7 ± 8.3	3.1 ± 2.0	4.2 ± 2.1	6.2 ± 1.8	25.2 ± 9.0	29
1200 – 1400	> 350	0.2 ± 0.6	2.3 ± 2.3	2.3 ± 1.4	0.6 ± 0.8	5.4 ± 2.9	8
> 1400	> 200	12.0 ± 9.1	3.2 ± 2.4	2.7 ± 1.6	1.1 ± 0.5	19.0 ± 9.6	16

prediction. The QCD contributions are significantly larger than in the simulation, as discussed above, while the other contributions are lower. These observations emphasise the importance of data-based background prediction methods.

In the following Section 7.3.1, a procedure developed in this thesis to obtain the combined background prediction in Fig. 7.6 is explained. Given the absence of any signal above the SM expectation, the results are used to constrain possible new-physics models. In Section 7.3.2, the derivation of exclusion limits with the CL_s method is described. The results are interpreted specifically in the context of the CMSSM, as discussed in Section 7.3.3, and also in the context of Simplified Models, as briefly presented in Section 7.3.4. Finally, in Section 7.3.5, implications for the MSSM in context with other SUSY searches are reviewed.

7.3.1 Derivation of Combined SM-Background H_T and $\not{H}_\mathrm{T}$ Predictions

The expected H_T and $\not{H}_\mathrm{T}$ distributions Fig. 7.6 of the SM-background processes are obtained using the data-based methods described above. For each background, the integral over each distribution corresponds to the sum of event yields in the 14 exclusive search regions Table 7.4.

The predictions involve corrections such as the $\sigma(Z \to \nu\bar{\nu}$ + jets$)/\sigma($photon + jets$)$ scale factor as well as contributions to the uncertainties such as statistical uncertainties due to control-sample sizes which depend on H_T and $\not{H}_\mathrm{T}$. However, the correction factors and uncertainties are determined on average for each considered search region. Therefore, the evolution of the predictions with H_T and $\not{H}_\mathrm{T}$ might not be accurately represented when solely applying the methods to the baseline selection. Instead, the predictions should be performed in each of the distributions' bins. This is not possible at this stage of the analysis, though, due to lack of statistical precision.

In order to display yet the dependence on H_T and $\not{H}_\mathrm{T}$ with sufficient precision for a granularity as in Fig. 7.6, a procedure has been developed in this thesis to complement the distributions obtained after the baseline selection with information from the search regions Table 7.4. This is explained in the following for the H_T distributions.

1. For each H_T-interval B defined by the search-region boundaries, a prediction N_B with uncertainty σ_B inclusive in $\not{H}_\mathrm{T}$ is determined for each background. Hence, information from the search regions with the same H_T boundaries are combined leading to the five H_T intervals listed below.

H_T interval (GeV)	Contained $\not{H}_\mathrm{T}$ intervals (GeV)
$500 - 800$	$200 - 350$, $350 - 500$, $500 - 600$, $600 - \infty$
$800 - 1000$	$200 - 350$, $350 - 500$, $500 - 600$, $600 - \infty$
$1000 - 1200$	$200 - 350$, $350 - 500$, $500 - \infty$
$1200 - 1400$	$200 - 350$, $350 - \infty$
$1400 - \infty$	$200 - \infty$

The inclusive prediction N_B is calculated as linear sum of the predictions in the corresponding search regions. The uncertainty in each search region has two components, a statistical component due to the control-sample size and a systematic component.

- The statistical components in the search regions are considered to be uncorrelated. Hence, the inclusive statistical component σ_B^{stat} is calculated as their quadratic sum.
- The systematic components in the search regions are considered to be fully correlated. Hence, the inclusive systematic component σ_B^{syst} is calculated as their linear sum.

The total inclusive uncertainty σ_B is calculated as the quadratic sum of σ_B^{stat} and σ_B^{syst}.

2. The prediction N_b and uncertainty σ_b in each bin b of the H_{T} histogram (which has a finer binning than the search regions) are determined from N_B and σ_B. The prediction for each $b \in B$ is set to

$$N_b = \frac{n_b}{\sum_{b \in B} n_b} \cdot N_B \ ,$$

where n_b denotes the number of entries in b as obtained for the baseline prediction only, i. e. with correction factors inclusive in H_{T} and $\not\!\!H_{\mathrm{T}}$. The statistical component of the uncertainty is assumed to evolve as for Poisson-distributed quantities, the relative systematic component is assumed to be constant. Hence, the uncertainty for each $b \in B$ is set to

$$\sigma_b = \sqrt{\left(\sigma_b^{\text{stat}}\right)^2 + \left(\sigma_b^{\text{syst}}\right)^2} \ ,$$

with

$$\begin{aligned} \sigma_b^{\text{stat}} &= s \cdot \sqrt{N_b} \\ \sigma_b^{\text{syst}} &= s \cdot N_b \end{aligned}$$

with the scale factor

$$s = \sqrt{\frac{\left(\sigma_B^{\text{stat}}\right)^2 + \left(\sigma_B^{\text{syst}}\right)^2}{\sum_{b \in B} N_b + N_b^2}} \ .$$

3. The predictions N_b and uncertainties σ_b of all backgrounds are combined assuming uncorrelated uncertainties. Correlations between the statistical components of the lost-lepton and the hadronic-τ uncertainties due to the similar control samples are neglected. Since the size of both is small compared to the total uncertainties, the effect of this simplification is not visible. Hence, the predictions are added linearly, and the uncertainties are added in quadrature.

For the $\not\!\!H_{\mathrm{T}}$ distributions, the procedure works analogue. However, given the boundaries of the chosen search regions, the high-$\not\!\!H_{\mathrm{T}}$ search regions share some H_{T} intervals. To avoid double-counting, the following combinations have been chosen based on the size of the predicted event yield in each region.

$\not\!H_{\mathrm{T}}$ interval (GeV)	Contained H_{T} intervals (GeV)
200 – 350	500 – 800, 800 – 1000, 1000 – 1200, 1200 –1400, 1400 – ∞
350 – 500	500 – 800, 800 – 1000, 1000 – 1200, 1200 –1400, 1400 – ∞
500 – 600	500 – 800, 800 – 1000, 1000 – 1200
600 – ∞	500 – 800, 800 – 1000

The uncertainties calculated in the described way have been estimated to be accurate at the 5% level based on the increase of the uncertainties between the search regions.

7.3.2 Derivation of Exclusion Limits

In this analysis, exclusion limits on the parameters of new-physics models are derived with the CL_s method [286–288], which has been developed for the Higgs searches at LEP to obtain parameter limits in case of small expected signal rates at the limit of the experimental sensitivity. In order to quantify the degree of agreement between the data and a specific hypothesis, a function, the *test statistic*, of the experimental observables and the parameters of the tested new-physics model is defined. Two hypotheses are considered, the *background hypothesis* (b) where the observations are expected to originate in SM processes only, and the *signal + background hypothesis* ($s + b$) where additional contributions arise from the new-physics model. Typically, the test statistic is monotonically increasing where larger function values Q imply less compatibility with the background-only hypothesis. The *confidence level* CL_x of a specific hypothesis x is given by the probability $P_x(Q \leq Q_{\mathrm{obs}})$ to obtain a value Q of the test statistic less than or equal to the value Q_{obs} actually observed in data,

$$\mathrm{CL}_x = P_x(Q \leq Q_{\mathrm{obs}}) \ .$$

If the total number of observed events in data is small and fluctuates below the average expectation for the background-only hypothesis, the confidence level CL_{s+b} might result in unphysical estimates of the model parameters and strong exclusion limits on the signal. To avoid such alleged exclusion of a signal to which the experiment has no sensitivity, the ratio

$$\mathrm{CL}_s = \frac{\mathrm{CL}_{s+b}}{\mathrm{CL}_b}$$

is used instead to derive limits. The signal hypothesis is defined as being excluded at the confidence level α if

$$1 - \mathrm{CL}_s \leq \alpha \ .$$

For the limit calculation, the 14 exclusive search regions are treated as independent channels. In each search region, the different SM-background predictions Table 7.4 are combined assuming Gaussian probability densities with uncorrelated statistical and fully correlated systematic uncertainties. Correlations between the statistical uncertainties of the lost-lepton and the hadronic-τ background due to the similar control samples are taken into account as well. A CL_s likelihood ratio is determined in each search region, and a global test statistic is computed as their product. The procedure follows [289].

7.3.3 Interpretation in the Context of the CMSSM

The search results are interpreted as exclusion limits on the parameters of the Constrained MSSM (CMSSM), cf. Section 2.3.2.3.

The expected number of signal events in the different search regions are determined from simulation assuming $\tan\beta = 10$, $\mu > 0$, and $A_0 = 0$. The values of m_0 and $m_{1/2}$ are varied in steps of 20 GeV, and for each point a sample of 10000 events is generated using the IsaJet programme [290] and the CTEQ6L1 PDF set [186]. Subsequent parton showering and hadronisation processes are simulated with Pythia, and the detector effects are modelled using the CMS fast-simulation tool [215] to save computing time, cf. Section 4.3. Finally, events are weighted to correct the yields to the NLO cross sections calculated with Prospino. The expected event yields correspond to signal acceptances of a few percent up to 8% depending on the search selection and the point in the $m_0 \times m_{1/2}$ parameter plane. Uncertainties on the signal yield due to the following effects are considered:

Statistical uncertainty: The uncertainty due to the limited number of generated signal events amounts to 3 – 30% depending on the search region and the values of m_0 and $m_{1/2}$.

Parton-distribution functions: Uncertainties on the parton-distribution functions are propagated to the signal yields following the recommendations of the PDF4LHC working group [291]. The 68% confidence level of the uncertainties is computed as described by the authors. Half of the maximum difference between positive and negative variations is quoted as uncertainty, which amounts to about 6%.

Jet energy calibration: The jet energy calibration factors are varied by their uncertainties, cf. Section 4.5 and [172], and the average difference of the resulting signal yields is taken as uncertainty, which is about 8%.

Jet transverse-momentum resolution: The jet transverse-momentum resolution in data is larger than simulated, cf. Chapter 5, and the simulation has been adjusted to account for this effect. Propagation of the uncertainties on the measured resolution difference results in a variation of the signal yield of about 2%.

In addition to the listed uncertainties on the signal acceptance, uncertainties due to the luminosity determination (2.2%) [240], trigger inefficiencies (2%), and the event cleaning procedure (3%), as well as the uncertainties of the predicted SM-background yields are taken into account in the limit-setting procedure. Possible background overprediction due to signal events in the control samples used in the data-based estimation methods have been estimated to range from 3% at low m_0 to 20% at high m_0. They are subtracted from the signal + background hypothesis for the limit determination.

The observed and expected CL$_s$ upper limits on the CMSSM cross-section are mapped into lower limits in the $m_0 \times m_{1/2}$ plane where the CMSSM is excluded at the 95% confidence level as shown in Fig. 7.7. The corresponding exclusion limits in the $m_{\tilde{g}} \times m_{\tilde{q}}$ plane are also shown. At a fixed $m_{1/2}$ and low m_0, squark masses are small. Hence, much of the collision energy can be transferred to the LSP, leading to generally large values of $\not{H}_{\mathrm{T}}$. Therefore, the high-$\not{H}_{\mathrm{T}}$ search regions dominate the determined exclusion limit. At larger m_0, on the other hand, more energy has to be transferred to the heavier squarks, resulting in larger H_{T} and a higher sensitivity in the high-H_{T} search regions. At low m_0, values

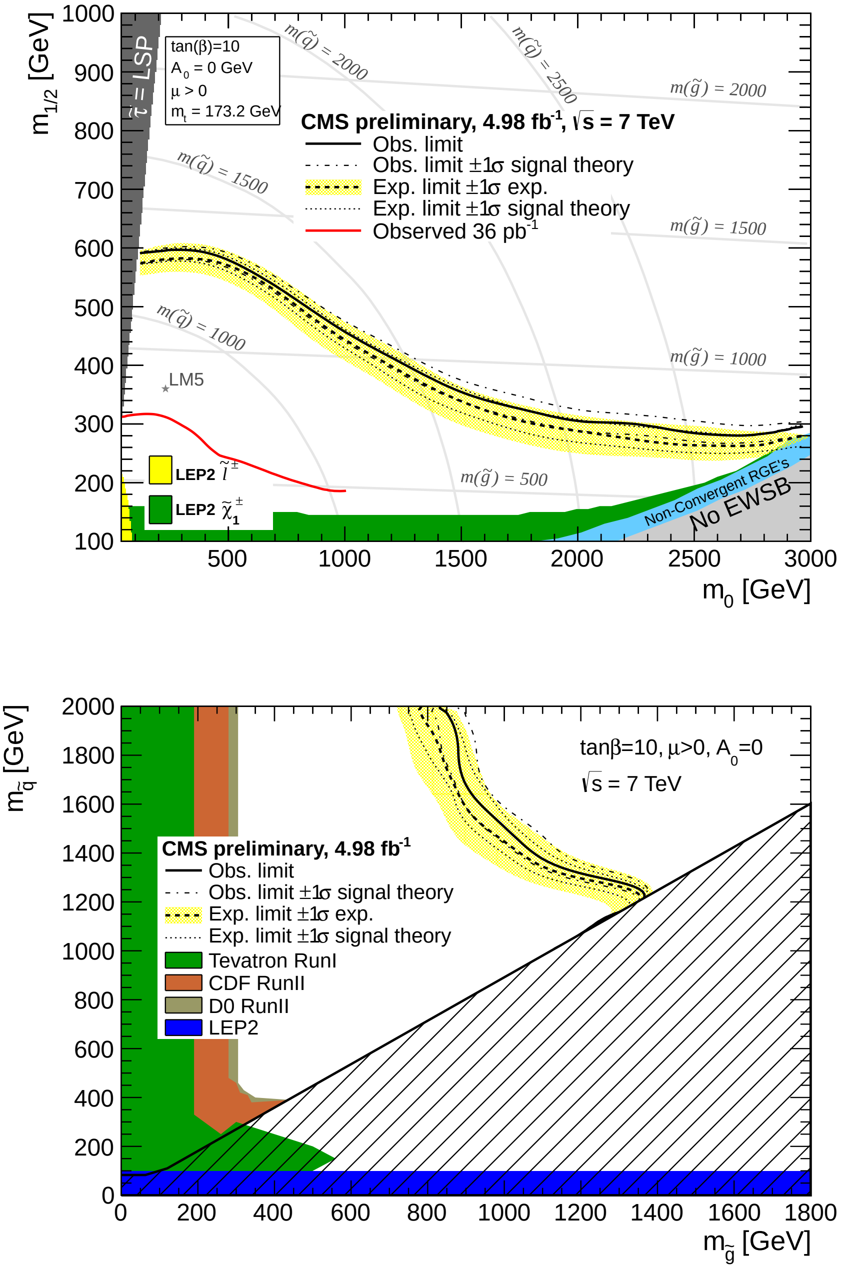


Figure 7.7: Observed (*solid thick black line*) and expected (*dashed thick black line*) 95% confidence-level lower limits on the CMSSM parameters m_0 and $m_{1/2}$ (*top*) and the corresponding $m_{\tilde{g}}$ and $m_{\tilde{q}}$ (*bottom*). Contributions to the limit uncertainty from experimental sources (*yellow band*) and from the signal cross-section (NLO scale and PDF) (*dashed thin lines*) are indicated separately. The remaining CMSSM parameters are $\tan\beta = 10$, $\mu > 0$, and $A_0 = 0$. Observed limits from the 2010 result [260] and from previous collider experiments at Tevatron [100–102] and at LEP [95–99] are also shown. Taken from [257].

of $m_{1/2}$ of up to 600 GeV are excluded, and at $m_0 = 1500$ GeV, values of $m_{1/2}$ of up to 350 GeV are excluded. This corresponds to lower limits between about 800 and 1300 GeV on the gluino mass and between about 1200 and 2500 GeV on the squark mass. Masses of below about 1300 GeV are excluded for $m_{\tilde{g}} = m_{\tilde{q}}$. The results are relatively insensitive to the value of $\tan\beta$.

The exclusion reach of the presented jets + $\not{H}_T$ analysis greatly exceeds the results obtained on the 2010 data-set of 36 pb^{-1} [260] as well as limits from previous collider experiments at Tevatron [100–102] and LEP [95–99]. In comparison to further analyses conducted by ATLAS [292] and CMS [259], the jets + $\not{H}_T$ analysis has a similar sensitivity to the CMSSM parameters as other searches in the all-hadronic final-state and gains better sensitivity than searches in the lepton and photon final-states, cf. Fig. 7.8. This superiority over the analyses in the much cleaner lepton and photon final-states is a result of the achieved precise understanding of the SM backgrounds, in particular also that from QCD-multijet events.

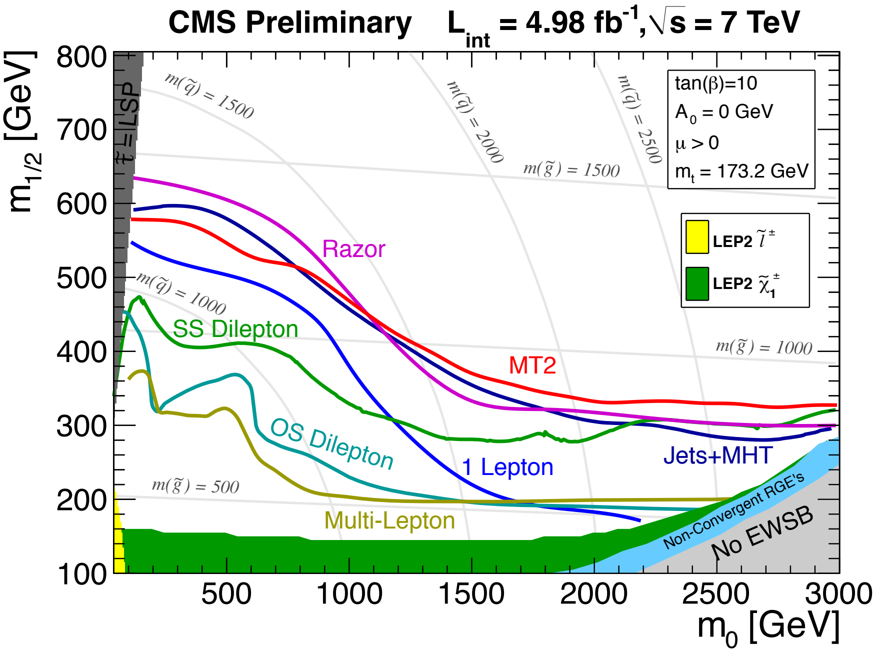


Figure 7.8: Comparison of the observed lower limits on the CMSSM parameters m_0 and $m_{1/2}$ obtained from different searches for new physics conducted by CMS in 2011 with 4.98 fb^{-1} of data. Regions which have been excluded by previous collider experiments at LEP [95–99] are indicated. The remaining CMSSM parameters are $\tan\beta = 10$, $\mu > 0$, and $A_0 = 0$. Taken from [259].

7.3.4 Interpretation in the Context of Simplified Models

The search results are also interpreted in the context of *Simplified Model Spectra* (SMS) [293–296], which are defined by the obtained final state. A limited set of new particles with certain possible decay channels is introduced. Their production cross-sections and the subsequent decay chains are parametrised in terms of the particle masses and their branching ratios. The SMS cross-section limits can be translated into limits on particle production in more complex models.

Two benchmark scenarios with a jets and $\not\!\!H_{\mathrm{T}}$ signature have been considered: gluino pair-production, where each gluino decays into two light quarks and an undetectable weakly interaction particle ($\tilde{\chi}^0$), and squark pair-production, where each squark decays into one light quark and a $\tilde{\chi}^0$. The scenarios are motivated by SUSY-based signatures, which is why the SUSY-terminology has been adopted. In Fig. 7.9, the observed 95% confidence-level exclusion limits on the cross sections of gluino and squark pair-production in the SMS are shown in the $m_{\tilde{g}} \times m_{\tilde{\chi}^0}$ and $m_{\tilde{q}} \times m_{\tilde{\chi}^0}$ planes. They are compared to the contours where the signal cross-sections are excluded. The signal acceptances, which are important for the interpretation of the limits within the context of other models, are documented in [257].

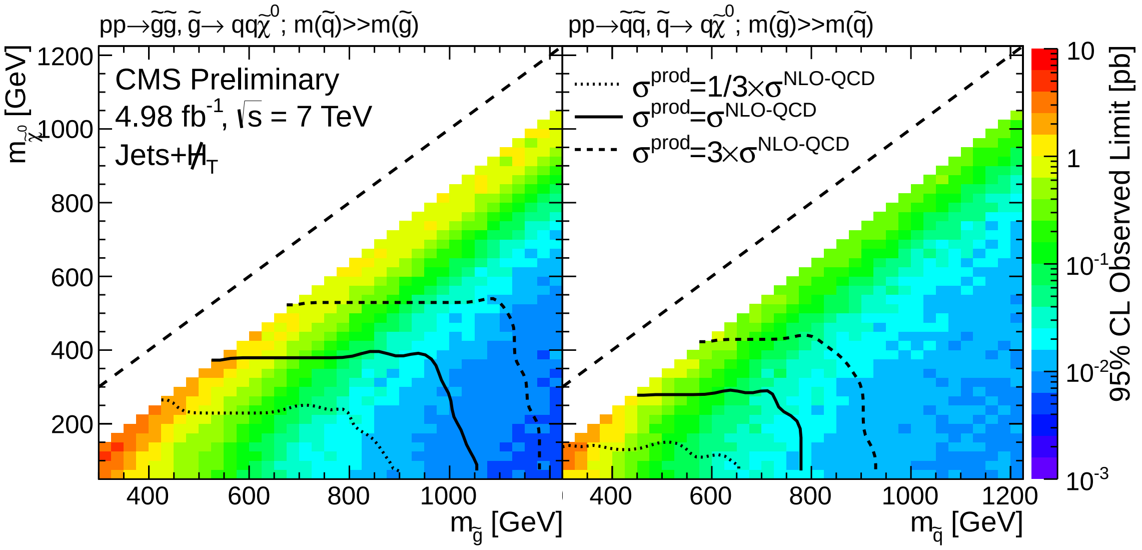


Figure 7.9: Observed 95% confidence-level upper limits on gluino (*left*) and squark (*right*) pair-production cross-sections obtained in the context of two different Simplified Models with a all-hadronic final state. The limits are shown in the $m_{\tilde{g}} \times m_{\tilde{\chi}^0}$ and $m_{\tilde{q}} \times m_{\tilde{\chi}^0}$ planes, respectively. Contours where the signal cross-section and one third and three times this cross section can be excluded are also shown (*lines*). Taken from [257].

7.3.5 Constraints on the MSSM Parameter Space

The LHC SUSY-searches provide direct exclusion limits on the MSSM parameters with a reach unprecedented at collider experiments. Thus, they also greatly impact the global MSSM parameter fits mentioned in Section 2.3.3. An update of the MasterCode analysis discussed ibid. has been published [297], which takes into account various results[3] obtained

---

[3] The $1.1\,\mathrm{fb}^{-1}$ result of the presented jets + $\not\!\!H_{\mathrm{T}}$ analysis [238] has in fact not yet been considered since the global-fit analysis has been published earlier. The most sensitive limits taken into account from CMS are the ones obtained by the α_{T} analysis [263], but the exclusion reach is similar to the one of the jets + $\not\!\!H_{\mathrm{T}}$ analysis, cf. Fig. 7.8.

by ATLAS and CMS with approximately $1\,\mathrm{fb}^{-1}$ of LHC data. Most strikingly, the best-fit values for m_0 and $m_{1/2}$ increase to $450^{+1700}_{-320}\,\mathrm{GeV}$ and $780^{+1350}_{-270}\,\mathrm{GeV}$, respectively, in the context of the CMSSM. The larger uncertainties are a consequence of emerging tension between the results of direct searches at LHC in the hadronic channels and the indirect $(g-2)_\mu$ data, which favour much smaller mass parameters. This tension is also responsible for a larger best-fit value of $\tan\beta = 41^{+16}_{-32}$, but even larger values of $\tan\beta$ conflict with the constraints from heavy-Higgs searches and $B_s^0 \to \mu^+\mu^-$ data. With the updated parameter values, gluino masses of $m_{\tilde{g}} \approx 1600\,\mathrm{GeV}$ are obtained.

This is compatible with the direct exclusion limits presented above. Heavier sparticle masses also affect some of the theoretically appealing features of SUSY. For example, the required fine-tuning factor of the Higgs mass parameter in the CMSSM increases from about 100 to about 300 when including the LHC data into the global fit [297].

Despite the increasing tension between the results from direct and indirect searches, the CMSSM and other gravity-mediated breaking scenarios are still compatible with the data with best-fit probabilities above 10%, which is larger than for the SM only. Their 95% confidence-level regions extend up to $m_0 \approx 2000\,\mathrm{GeV}$ and $m_{1/2} \approx 1500\,\mathrm{GeV}$ [297]. Hence, it is difficult to invalidate these models even if future LHC results significanlty enlarge the excluded $m_0 \times m_{1/2}$ parameter space. On the other hand, tension can be reduced for example in gauge-mediated breaking scenarios, where larger differences are possible between the masses of coloured sparticles, to which the hadronic searches at the LHC are sensitive, and the masses of colourless sparticles, to which the $(g-2)_\mu$ data are most sensitive [113].

Comparison of the results from direct searches for the SM Higgs boson with the expected mass m_h of the lightest MSSM Higgs, which exhibits the same signature, provide further, important information to evaluate the consistency of MSSM-based SUSY models. Values of $m_h \gtrsim 130\,\mathrm{GeV}$ are incompatible. From the discussed global CMSSM fit, a value of $m_h = 119.1^{+3.4}_{-2.9}\,\mathrm{GeV}$ is obtained, where results from the direct Higgs searches have not been considered [297]. It is exciting that candidate events are currently observed for a SM-like Higgs boson with a mass between about 122.5 and 127.5 GeV [42, 43], and that a conclusive answer is to be expected from this year's LHC run.

8 Conclusions

The primary goals of the LHC are the discovery of the mechanism behind electroweak symmetry breaking and the search for new physics beyond the SM. While *pp* collisions are well suited to reach highest centre-of-mass energies, they result in complex final states, which are dominated by jets. Furthermore, the need for highest luminosities leads to the presence of many additional soft pile-up collisions per event. This poses a major challenge to the experiments when reconstructing the final state and measuring the underlying hard-interaction processes.

CMS extensively exploits the Particle-Flow (PF) algorithm for the event reconstruction. The algorithm combines information from various subdetectors to identify and measure individual particles, from which jets can be clustered. PF jets are expected to feature a superior energy scale and resolution compared to jets based only on calorimeter information because the impact from the calorimeter non-linearity is reduced and the usage of information from the hadronic calorimeter, which has a relatively poor resolution, is restricted to the measurement of neutral hadrons. However, PF jets are naturally complex objects as they depend on the properties of very different subdetectors. A precise understanding of their performance is a prerequisite for any analysis exploiting jet final-states.

In this thesis, the transverse-momentum response function of PF jets at CMS has been studied in detail. First, simulated (MC-truth) jet-p_T response distributions have been investigated. They are dominated by an almost Gaussian core region. Its width, the relative jet-p_T resolution, depends on the jet's p_T, η, flavour, and, in particular at low p_T, also on the number of pile-up collisions in the event. At low p_T, the resolution is dominated by the performance of the inner tracking system, and at high p_T by calorimeter measurements. The relative MC-truth jet-p_T resolution amounts to $\sigma_\mathrm{MC}/p_\mathrm{T} = 9\%$ for jets with $p_\mathrm{T} = 100\,\mathrm{GeV}$ and improves to 5% for jets with $p_\mathrm{T} = 600\,\mathrm{GeV}$ in the central detector region. In addition, the response distributions feature non-Gaussian, low-response tails at the percentage level and below. They are caused by semi-leptonically decaying heavy-flavour quarks where energy is carried away by neutrinos as well as instrumental effects such as inactive channels or miscalibration.

Two methods have been developed to measure the jet-p_T resolution and the response tails directly from dijet data. In comparison to e. g. measurements in photon + jet data, these events have a high reach in p_T due to their larger production cross-section.

The jet-p_T resolution is determined using an unbinned maximum-likelihood fit based on the p_T balance in QCD-dijet events. The likelihood includes an estimate of the particle-level differential jet-production cross-section. Thus, migration effects due to the event selection can be explicitly incorporated, and an estimate of the average particle-level jet-p_T is obtained. Biases due to additional jet activity in the events are corrected for by an extrapolation method.

The resolution of PF jets has been measured in $855\,\mathrm{pb}^{-1}$ of *pp*-collision data collected by CMS at $\sqrt{s} = 7\,\mathrm{TeV}$ until summer 2011 and amounts to $\sigma/p_\mathrm{T} = (9.80 \pm 0.30^{+0.84}_{-0.81})\%$ at $p_\mathrm{T} = 113\,\mathrm{GeV}$ and $(5.40 \pm 0.08^{+0.27}_{-0.26})\%$ at $611\,\mathrm{GeV}$. This is systematically larger than in the simulation. The average difference is parametrised by a p_T-independent data-to-simulation ratio ρ_{res}, which is $\rho_{\text{res}} = 1.054 \pm 0.006^{+0.069}_{-0.068}$ in the central detector region and increases to $1.294 \pm 0.063^{+0.214}_{-0.211}$ at large $|\eta|$. The source of this difference is not yet understood, but detector inhomogenities, noise, and miscalibration as well as mismodelled shower shapes might contribute. The uncertainty of ρ_{res} is dominated by the systematic uncertainty due to the extrapolation method. The presence of pile-up collisions does not affect ρ_{res} because pile-up impacts the resolution only below 100 GeV, where the statistical uncertainties of the measurement are large due to low trigger rates.

The size of the non-Gaussian response tails is assessed by measuring the size of the dijet-asymmetry tails. As before, an extrapolation method is used to correct the results for biases due to additional jet acitivity in the events. Differences between the fractional tail-size in data and simulation imply incorrectly modelled response tails because of the relation between the asymmetry and the response. The assignment of such differences to the low- or high-response tail is ambiguous, however, due to the definition of the asymmetry.

A measurement of the asymmetry tails has been performed in $4.90\,\text{fb}^{-1}$ of pp-collision data at $\sqrt{s} = 7\,\text{TeV}$ collected by CMS in 2011. Non-Gaussian components are present that contribute by up to a few percent to the total distribution. In general, the tails are more pronounced in data than in the simulation, even after correction for the resolution difference using ρ_{res}. The tail differences occur in particular in the central detector region at medium to large p_{T}. Here, the fractional size of the tails in data is larger than in the simulation by factors between $\rho_{\text{tail}} = 1.2$ and 1.6. The source of this difference might be due to detector effects such as punch-through or miscalibration but possibly also mismodelling of the heavy-flavour fraction. The uncertainties are of the same order as the difference, though. They are dominated by the propagated uncertainties of ρ_{res}. The presence of pile-up collisions does not affect the result.

Results obtained here have been published in [238] and as supplemental material [257] to [237]. They are used to correct the MC-truth response function, which is employed in numerous physics analyses performed by CMS. The measured differences between data and simulation have already been observed in a previous version of this analysis on $36\,\text{pb}^{-1}$ of data that has been published in [172, 239]. They are consistent with the results from complementary methods based on the dijet asymmetry and on the p_{T} balance in photon + jet events published ibid.

QCD-dijet based methods are expected to remain a cornerstone when measuring the jet response because they provide the highest reach in p_{T} due to the large jet-production cross-section. Further improvements are expected from a combination with the results obtained with photon/Z + jet events, which, in case of the resolution, could be even included directly into the likelihood. This would improve the statistical precision of the measurements at $p_{\text{T}} < 100\,\text{GeV}$, where low trigger rates limit the amount of collected dijet data. In addition, the ambiguities in assigning the observed data-to-simulation differences to the low- or high-response tail could be resolved.

As demonstrated in this thesis, the jet-p_{T} response function of the complex PF jets is understood with good precision by the simulation. Despite of the relatively poor HCAL resolution, the achieved performance of jet measurements at CMS is compatible to that obtained by detector concepts based on high-resolution calorimetry as implemented for example by ATLAS [230, 231]. Residual differences to the data can be measured with reasonable accuracy and are used to correct the simulation. This ability is compulsory to employ PF jets for physics analysis at CMS. It also supports design concepts for detectors at a possible future linear collider, which typically rely on PF event-reconstruction strategies [221] and might profit from the experience gained at CMS.

Finally, a search for new physics in final states with three or more jets, no leptons, and a large transverse-momentum imbalance conducted by CMS has been reviewed, which relies strongly on a precise knowledge of the jet-p_T response function. While the contributions from some SM-background processes are overestimated in the simulation, the simulated QCD-multijet background is significantly underestimated as a consequence also of the too low response tails in the simulation. Therefore, the kinematic properties of the expected SM-background events are predicted directly from data. The QCD-background contribution is determined from multijet data by modelling the momentum mismeasurements using the aforementioned, corrected MC-truth response functions. Taking into account the uncertainties on the response-function measurement, a prediction of the QCD background in the search region with a typical uncertainty of $60 - 70\%$ is achieved.

The analysis has been designed as a model-independent search. When interpreted within the context of the CMSSM, it is presently one of the searches with the highest sensitivity world-wide. This superiority over analyses in the much cleaner lepton and photon final-states is a result of the achieved precise understanding of the SM backgrounds using data-based methods.

A first version of the search has been published using $36\,\mathrm{pb}^{-1}$ of pp-collision data at $\sqrt{s} = 7\,\mathrm{TeV}$ [260]. The latest version [237], which has been reviewed in this thesis, has been performed with $4.98\,\mathrm{fb}^{-1}$ of data collected by CMS in 2011. No significant excess in the number of events has been observed above the SM expectations. The result is used to derive exclusion limits on the CMSSM parameters m_0 and $m_{1/2}$. At low m_0, values of $m_{1/2} < 600\,\mathrm{GeV}$ are excluded, and at $m_0 = 1500\,\mathrm{GeV}$, values of $m_{1/2} < 350\,\mathrm{GeV}$ are excluded. This corresponds to a lower limit of approximately $1300\,\mathrm{GeV}$ on the gluino and squark masses for $m_{\tilde{g}} = m_{\tilde{q}}$.

However, various MSSM-based SUSY scenarios are possible. Hence, a model-independent presentation of the search results is beneficial, such that they can easily be re-interpreted within different models. Therefore, cross-section limits on generic new-physics processes with all-hadronic final-states (Simplified Model Spectra [293–296]) have been derived.

Direct SUSY searches at the LHC, headed by the all-hadronic analyses, have excluded large parts of the CMSSM parameter-space. Hence, the masses of SUSY particles must be above approximately $1.5\,\mathrm{TeV}$, which starts to be in tension with the results from $(g-2)_\mu$ measurements that favour lighter masses. However, the CMSSM and other gravity-mediated breaking scenarios are still compatible with the data with best-fit probabilities larger than for the SM only [297]. Moreover, the currently observed candidate events for a Higgs boson with a mass of approximately $125\,\mathrm{GeV}$ [42, 43] are in good agreement with the SUSY expectations because the MSSM requires a SM-like Higgs boson lighter than approximately $130\,\mathrm{GeV}$. Thus, the direct SM Higgs-boson searches at the LHC, from which conclusive results are expected with the 2012 data, will provide important information whether SUSY may be realised in nature.

Finally, the increase of the LHC centre-of-mass energy from 7 to $8\,\mathrm{TeV}$ in 2012 further enlarges the accessible SUSY parameter space. At the same time, the sensitivity is improved because the production rate for heavier particles increases relatively more than for lighter particles, thus reducing the relative size of the background. For example, the production rate for $100\,\mathrm{GeV}$ particles increases by a factor of 1.3 but for $4\,\mathrm{TeV}$ particles by a factor of 10 [133]. The presented analysis techniques will be applied to the new data.

As argued, significant advancements are expected on the road towards an answer to one of today's primary questions in particle physics: what lies beyond the Standard Model?

A Jet Transverse-Momentum Resolution

A.1 Cleaning-Filter Efficiencies

Table A.1: Number N(passed) of events passing the cleaning filters described in Section 5.1.4 in $855\,\mathrm{pb}^{-1}$ of data and in the PYTHIA simulation. The efficiency is quoted relative to the previous selection step. The total number of data events refers to the number after the trigger requirements.

	Data		Simulation	
Filter	N(passed)	Efficiency (%)	N(passed)	Efficiency (%)
Total	7589851	—	10930800	—
Good vertex	7588500	$99.98220^{+0.00048}_{-0.00049}$	10930718	$99.9992^{+0.0001}_{-0.0001}$
HBHE noise	7533444	$99.27448^{+0.00308}_{-0.00309}$	10899989	$99.7189^{+0.0016}_{-0.0016}$
Beam halo	7515674	$99.76412^{+0.00176}_{-0.00177}$	10899739	$99.9977^{+0.0001}_{-0.0001}$
EE noise	7515672	$99.99997^{+0.00002}_{-0.00002}$	10899739	$100.0000^{+0.0000}_{-0.0000}$
Beam scraping	7515671	$99.99999^{+0.00001}_{-0.00002}$	10899739	$100.0000^{+0.0000}_{-0.0000}$
PF post-processing	7511594	$99.94575^{+0.00085}_{-0.00085}$	10894895	$99.9556^{+0.0006}_{-0.0006}$
Tracking failure	7470144	$99.44819^{+0.00270}_{-0.00271}$	10774349	$98.8936^{+0.0032}_{-0.0032}$
TP	7464547	$99.92508^{+0.00100}_{-0.00101}$	10756505	$99.8344^{+0.0012}_{-0.0012}$
BE	7400881	$99.14709^{+0.00336}_{-0.00337}$	10674908	$99.2414^{+0.0026}_{-0.0027}$
Lepton veto	7380802	$99.72869^{+0.00191}_{-0.00192}$	10662387	$99.8827^{+0.0010}_{-0.0011}$

A.2 Function Minimisation with LVMINI

The *Newton method* is a standard technique to find the minimum of an objective function $F(\vec{x})$ which depends on the n free parameters $\vec{x} \in \mathbb{R}^n$. It is an iterative approach based on a quadratic model[1] $M(\vec{d})$ that describes the change of $F(\vec{x})$ for a change of $\vec{x}$ by $\vec{d}$:

$$M_k(\vec{d}\,) = F(\vec{x}_k\,) + \nabla F^T(\vec{x}_k)\vec{d} + \frac{1}{2}\vec{d}^{\,T} C_k(\vec{x}_k)\vec{d} \approx F(\vec{x}_k + \vec{d}\,) \;,$$

where C_k is the Hessian matrix of second derivatives of F at $\vec{x}_k$ and $k \in \mathbb{N}$ denotes the iteration step. M_k is minimised for

$$\vec{d}_{\min} = -B_k \nabla F^T(\vec{x}_k), \qquad \text{with } B_k = C_k^{-1} \;, \tag{A.1}$$

if C_k is positive definite. If F is not quadratic, the procedure has to be iterated with $\vec{x}_{k+1} = \vec{x}_k + \vec{d}_{\min}$ until convergence with $\nabla F(\vec{x}_k) \to 0$ to find the minimum of F.

In case of an iterative solution, the convergence behaviour can be improved significantly by *line-search algorithms*, where in each iteration the new parameter value is obtained from $\vec{x}_{k+1} = \vec{x}_k + \alpha_{\min}\vec{d}_{\min}$ with $\alpha_{\min}$ minimising the function $\phi(\alpha) = F(\vec{x}_k + \alpha\vec{d}_{\min})$, i. e. $\alpha_{\min}$ optimises the step length in the direction $\vec{d}_{\min}$. Often, *inexact line-search* techniques are applied in order to determine an α at minimal computational expense which achieves adequate reduction of F. The latter can be controlled, for example, by the *Wolfe conditions*, a set of inequalities which relate $F(\vec{x}_k + \alpha\vec{d}_{\min})$ and $F(\vec{x}_k)$ with $\nabla F(\vec{x}_k)$ in order to define sufficient decrease and slope reduction of F.

Since each step of the Newton method involves the evaluation of the Hessian C_k or rather its inverse B_k, the required computing time and memory space increase with at least n^2. In *quasi-Newton methods*, B_k is therefore not computed directly but updated from an initial approximation by the successive evaluation of gradients. The BFGS algorithm belongs to the class of quasi-Newton methods and is defined by a certain updating formula for B_k. Its limited-memory version L-BFGS even avoids explicit construction of B_k at any time by exploiting the fact that B_k appears only in the product Eq. (A.1). Therefore, each update of d_k can be calculated in a sequence of scalar products and vector summations of an initial B_0 with the history of updates $s_k = \vec{x}_{k+1} - \vec{x}_k$ and $y_k = \nabla F_{k+1} - \nabla F_k$ of the position and gradient vectors, respectively. In practice, it is even sufficient to approximate the result from only the last 10 to 20 elements s_k, y_k of the history.

The programme LVMINI implements the L-BFGS algorithm in combination with an inexact line-search method based on the Wolfe conditions. As input, LVMINI requires the values of the objective function and its gradient after each iteration. After minimisation, the covariance matrix can be calculated explicitly to obtain parameter errors.

Detailed information on the reviewed techniques can be found for example in [223, 249, 251, 252]. The programme code and a technical manual of LVMINI are available at [248].

---

[1] This is the Taylor expansion to second order.

A.3 $p_{\mathrm{T}}^{\mathrm{gen}}$ Spectra

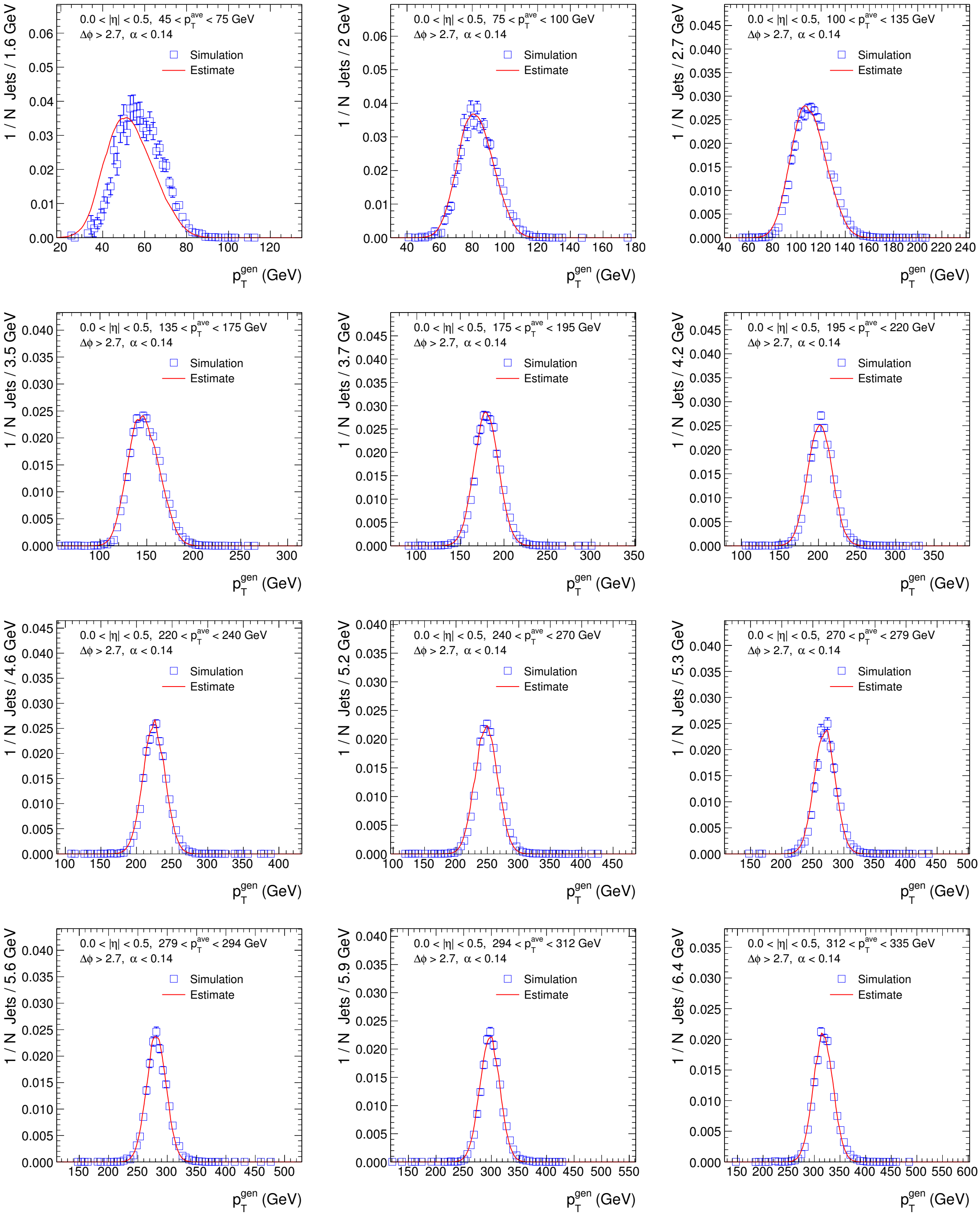


Figure A.1: $p_{\mathrm{T}}^{\mathrm{gen}}$ distributions (*squares*) of simulated events and estimate $f_{\hat{\sigma}'}(p_{\mathrm{T}}^{\mathrm{true}}|p_{\mathrm{T,min}}^{\mathrm{ave}}, p_{\mathrm{T,max}}^{\mathrm{ave}})$ Eq. (5.8) (*line*) in different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals for $\alpha < 0.14$. (Continued in Fig. A.2.)

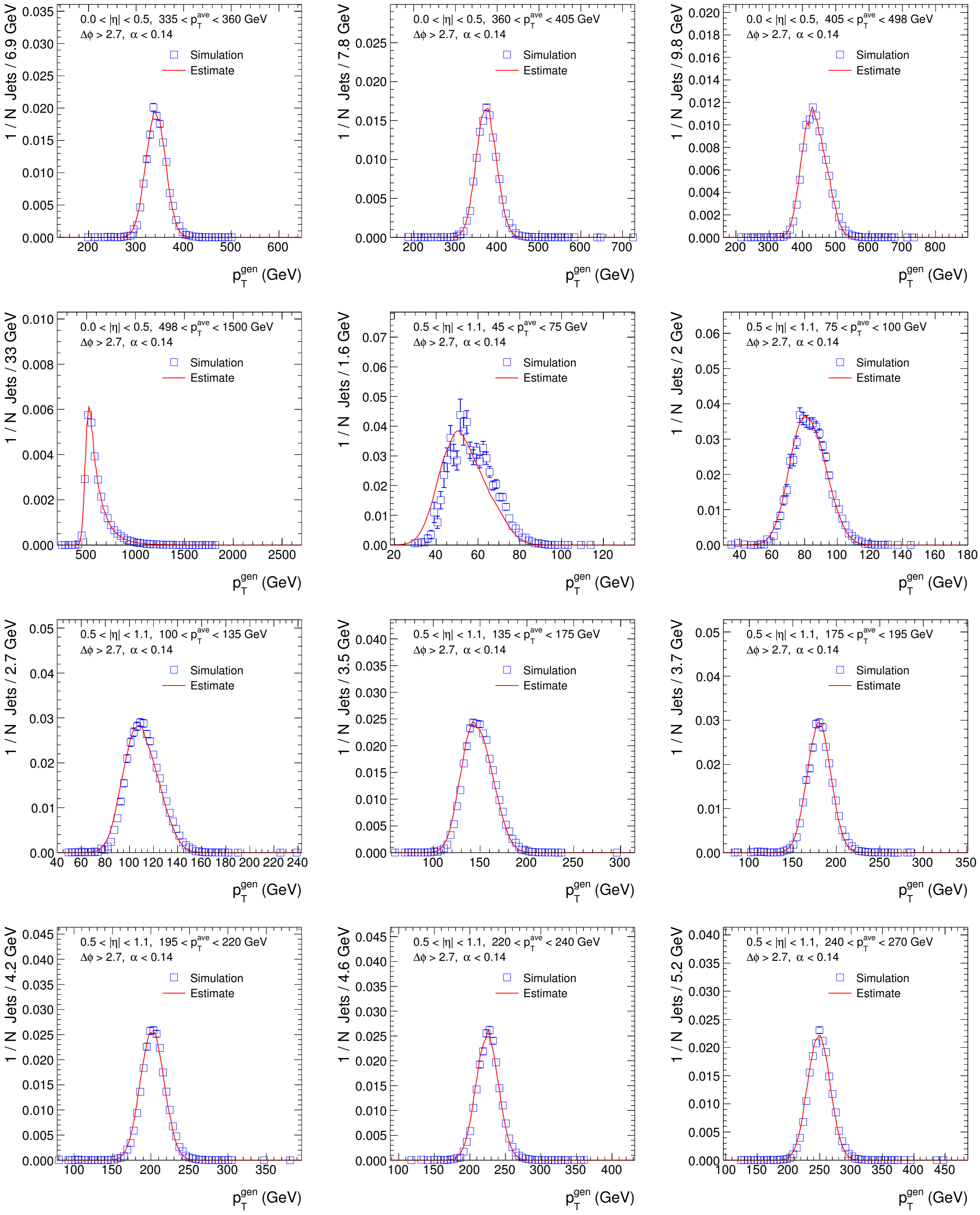


Figure A.2: Continued from Fig. A.1 and in Fig. A.3.

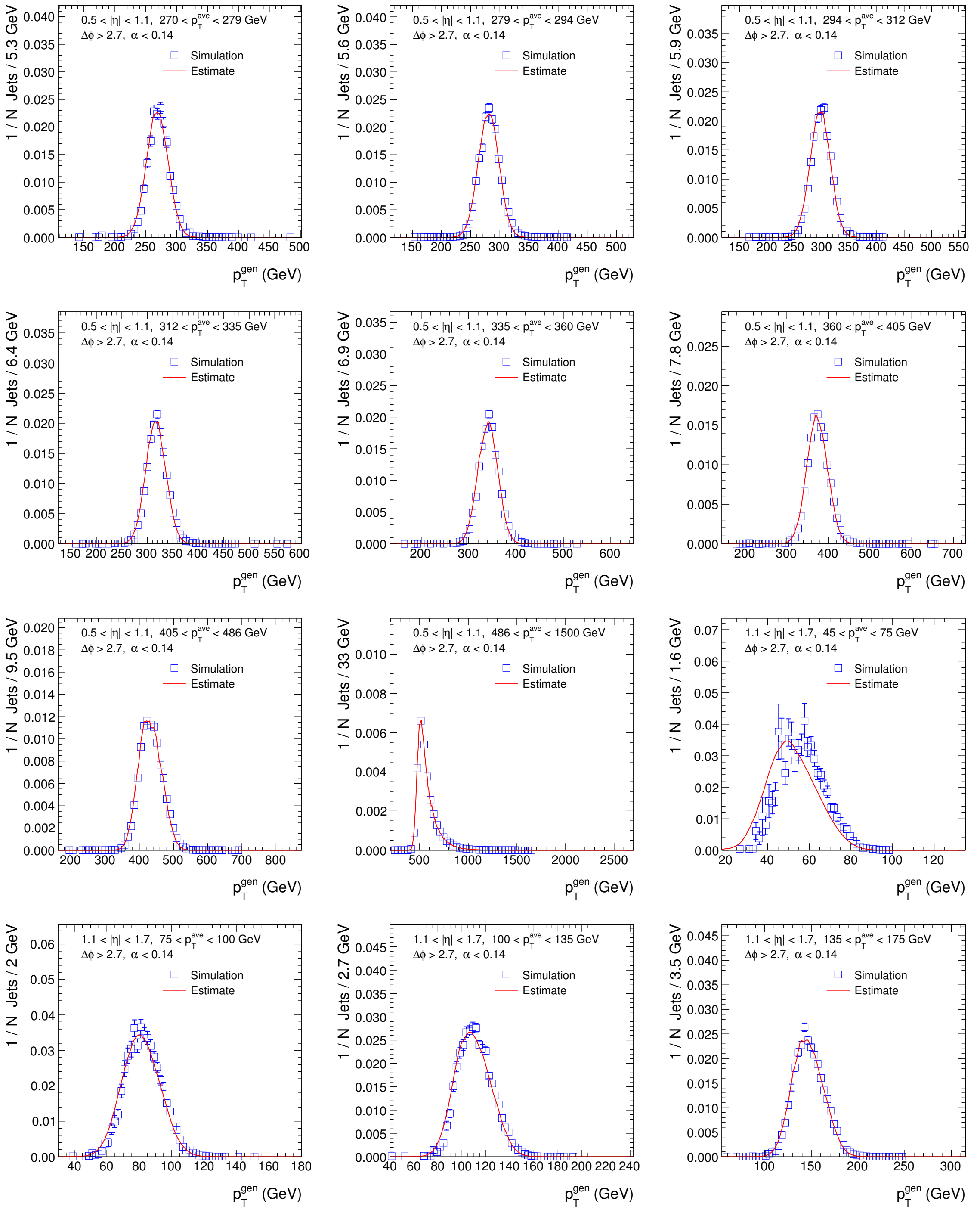


Figure A.3: Continued from Fig. A.2 and in Fig. A.4.

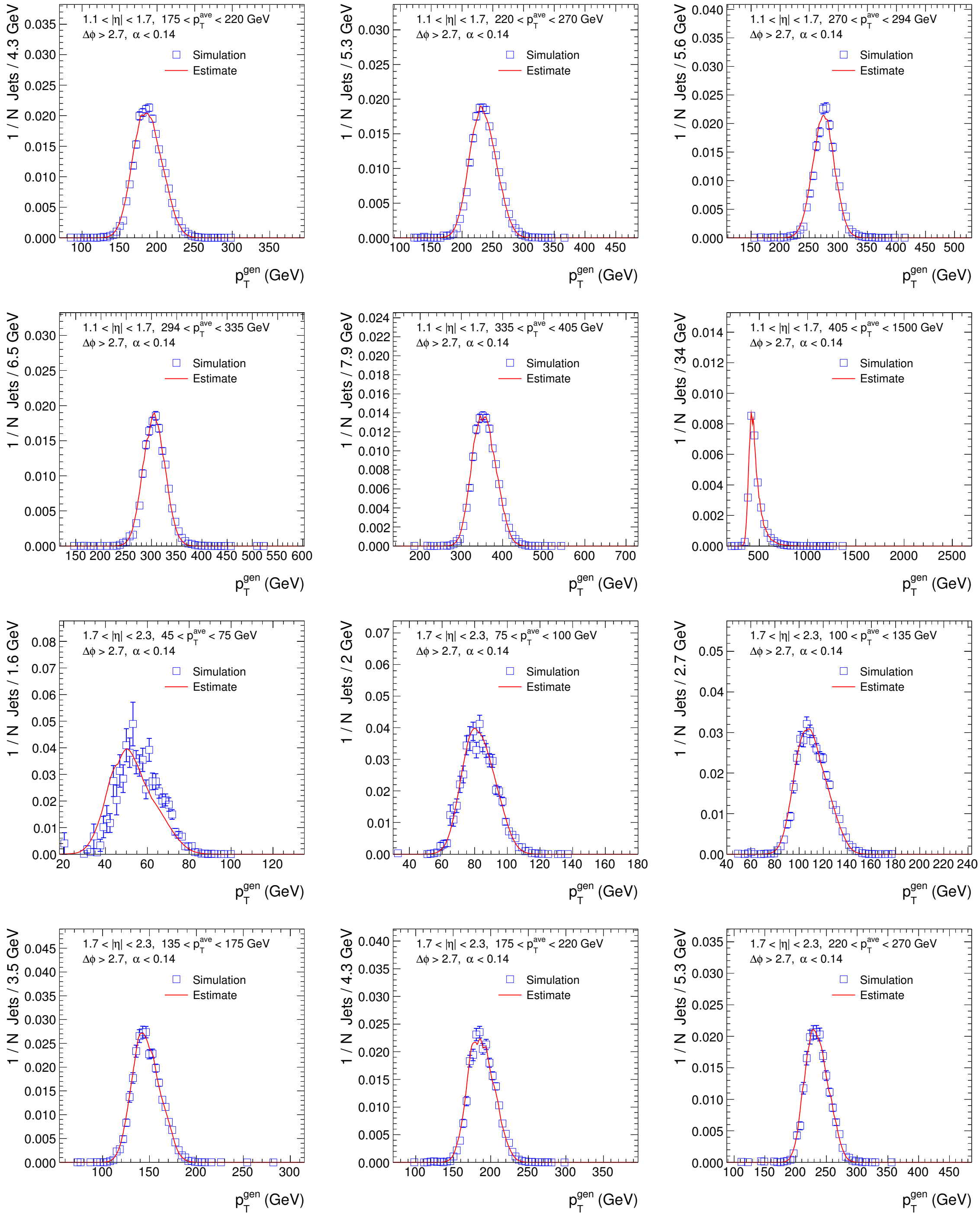


Figure A.4: Continued from Fig. A.3 and in Fig. A.5.

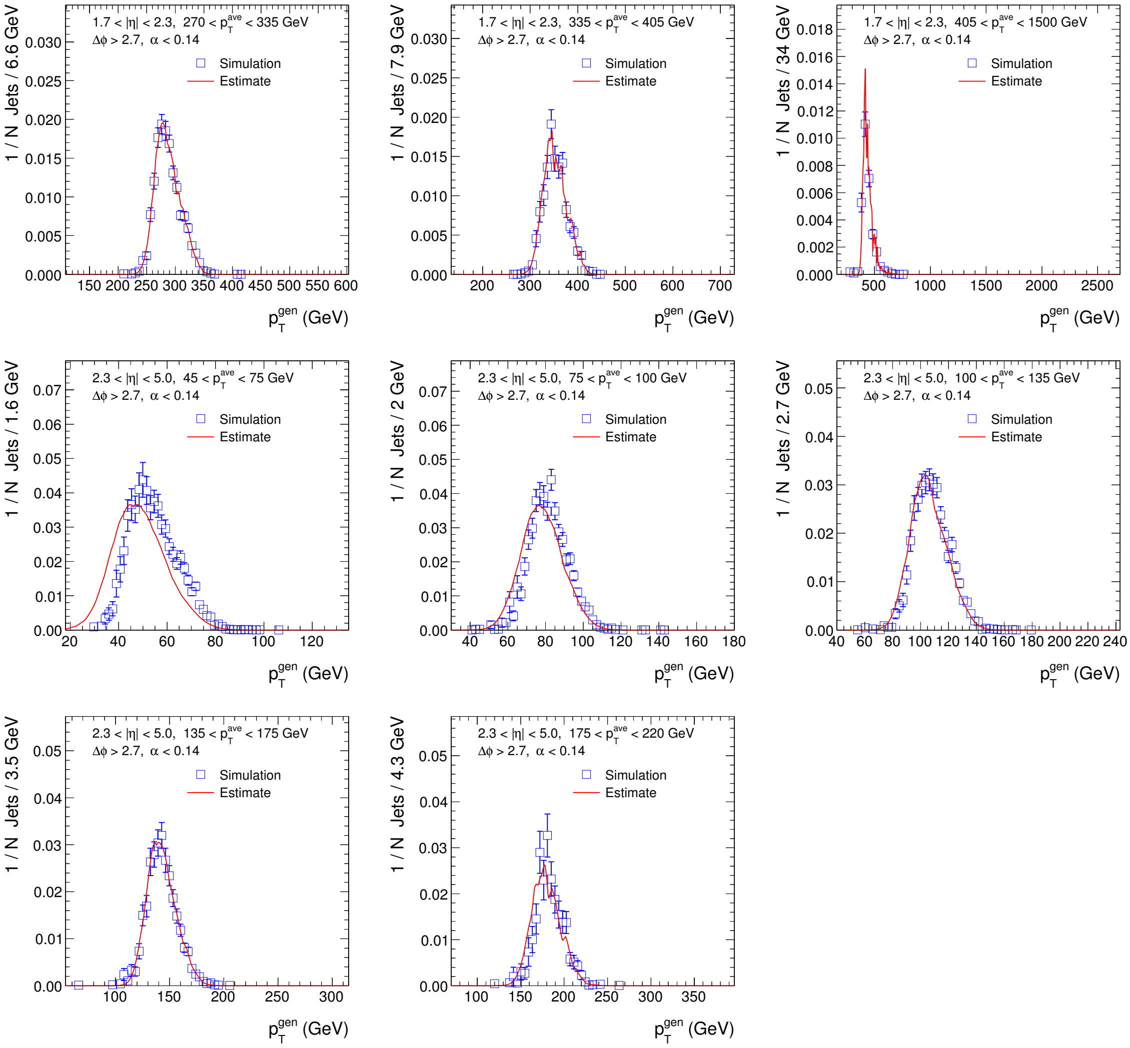


Figure A.5: Continued from Fig. A.4.

A.4 Δp_T Distributions

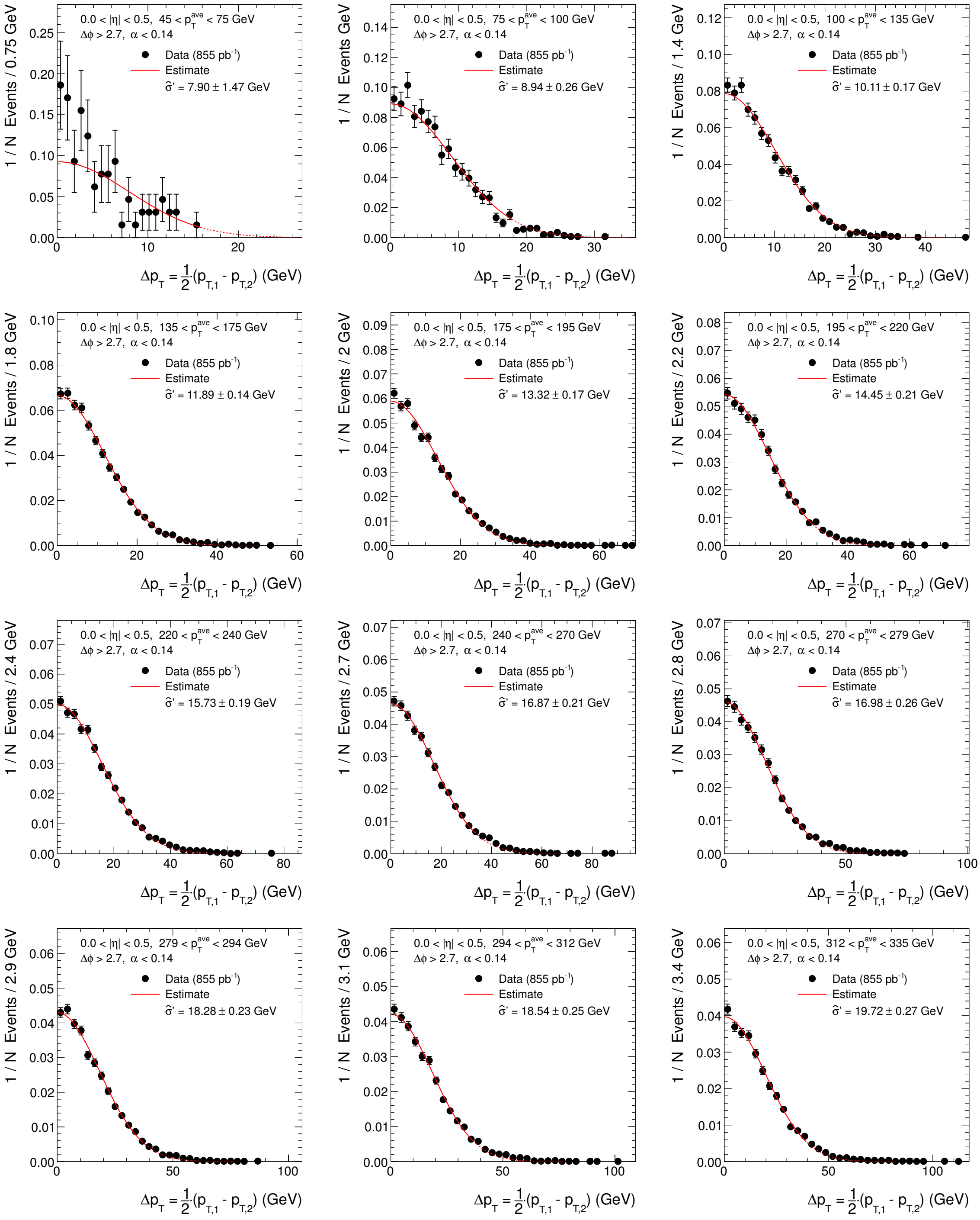


Figure A.6: Δp_T distributions in data (*circles*) and assumed Gaussian pdf (*line*) in different $|\eta| \times p_\mathrm{T}^\mathrm{ave}$ intervals for $\alpha < 0.14$. (Continued in Fig. A.7.)

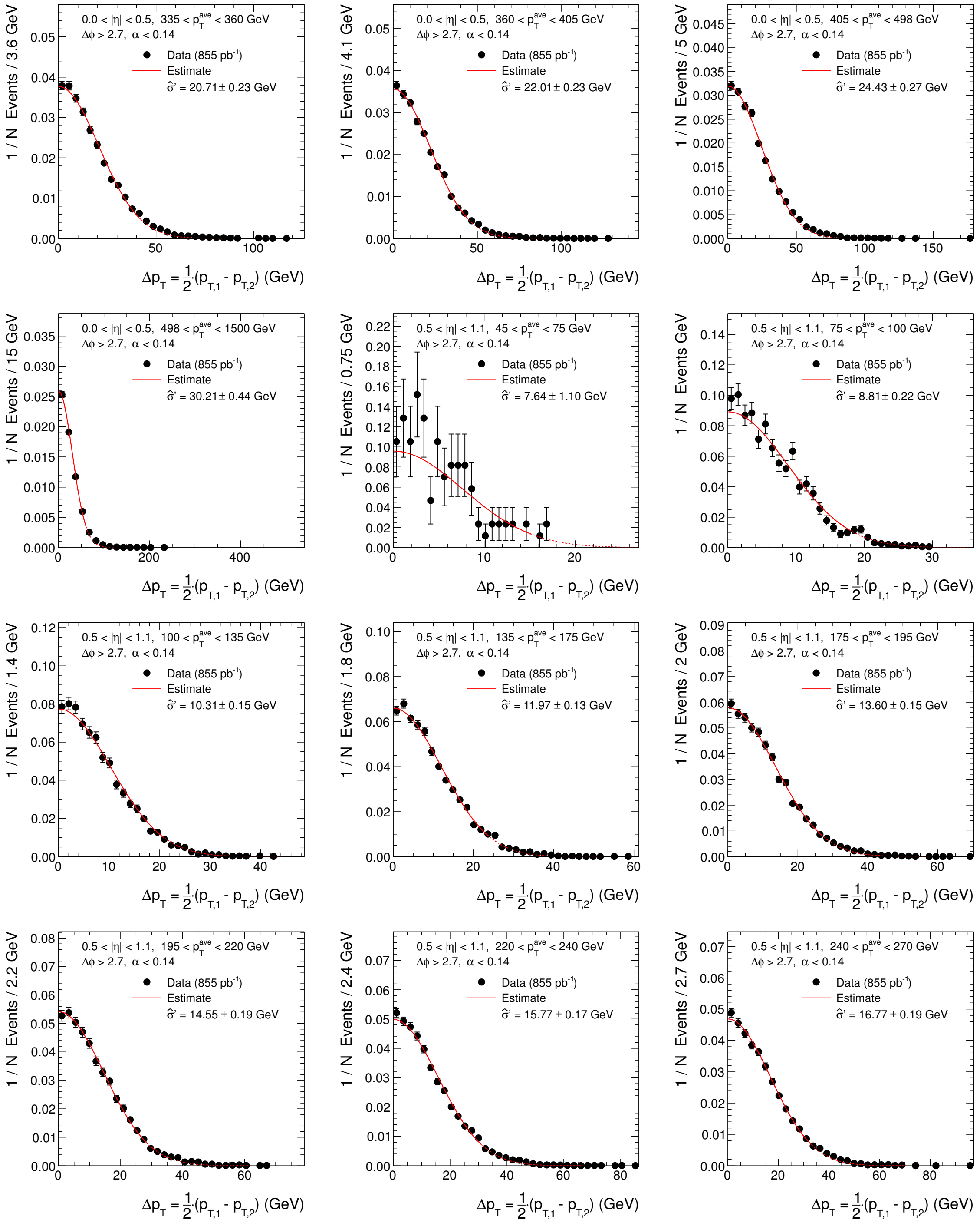


Figure A.7: Continued from Fig. A.6 and in Fig. A.8.

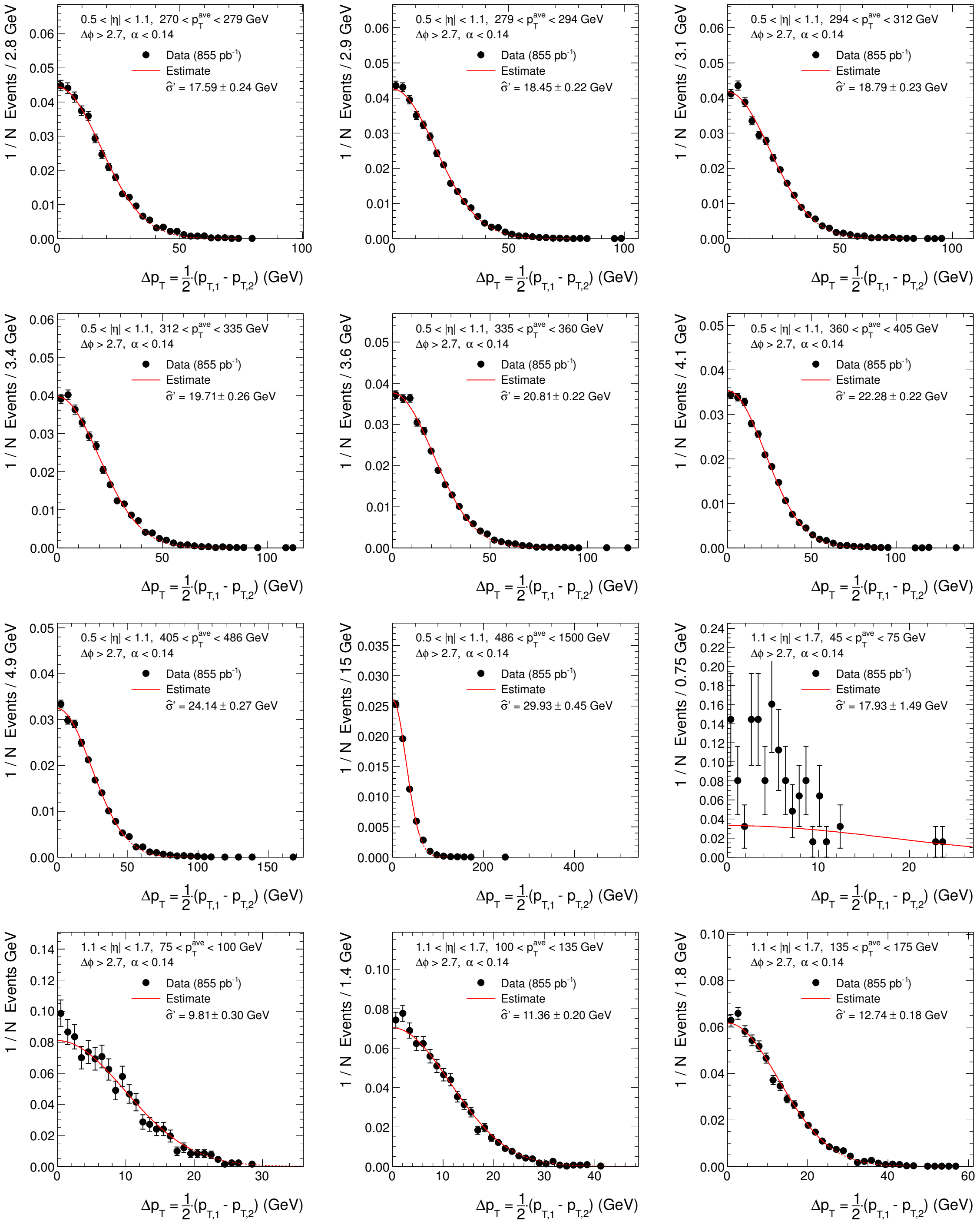


Figure A.8: Continued from Fig. A.7 and in Fig. A.9.

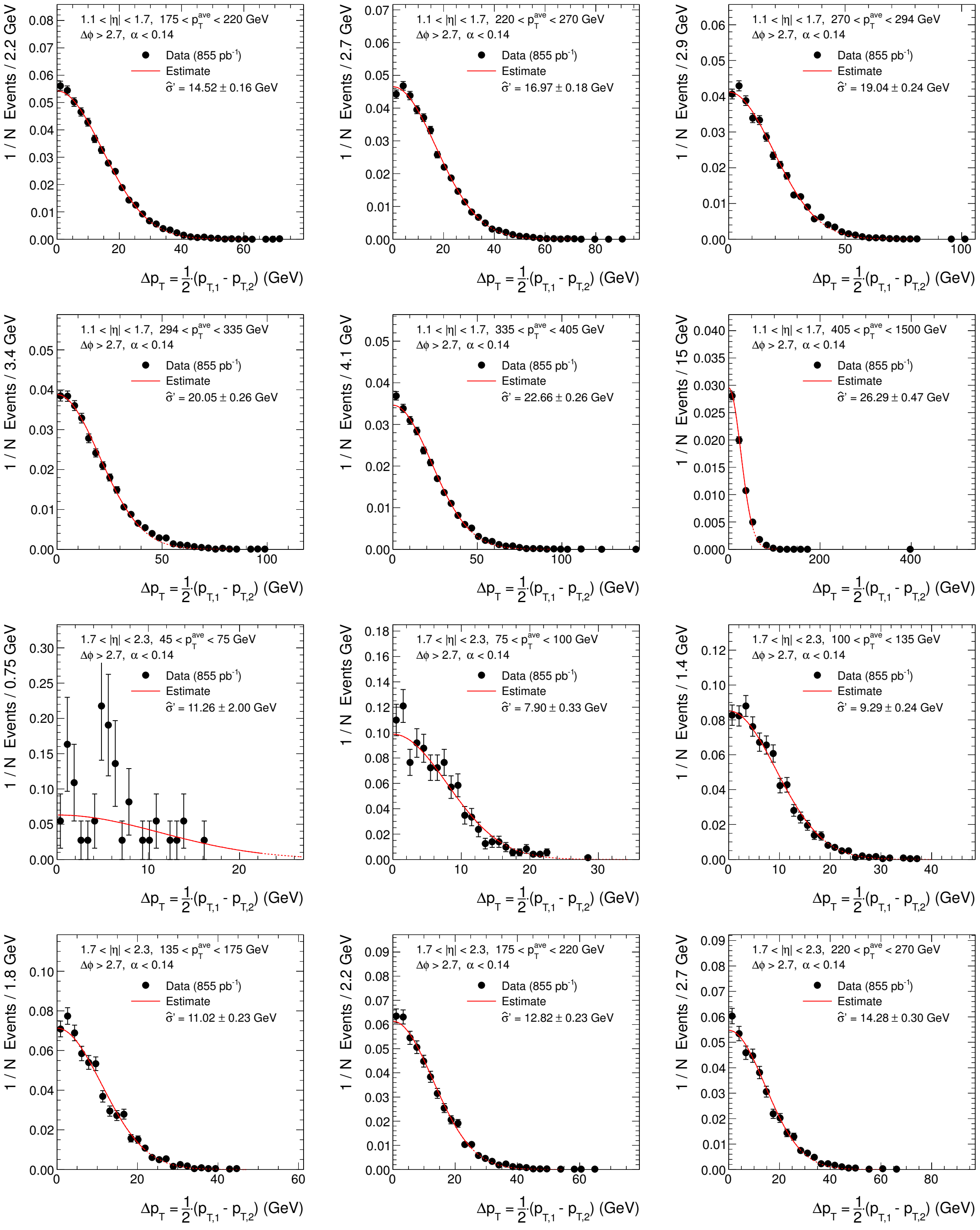


Figure A.9: Continued from Fig. A.7 and in Fig. A.10.

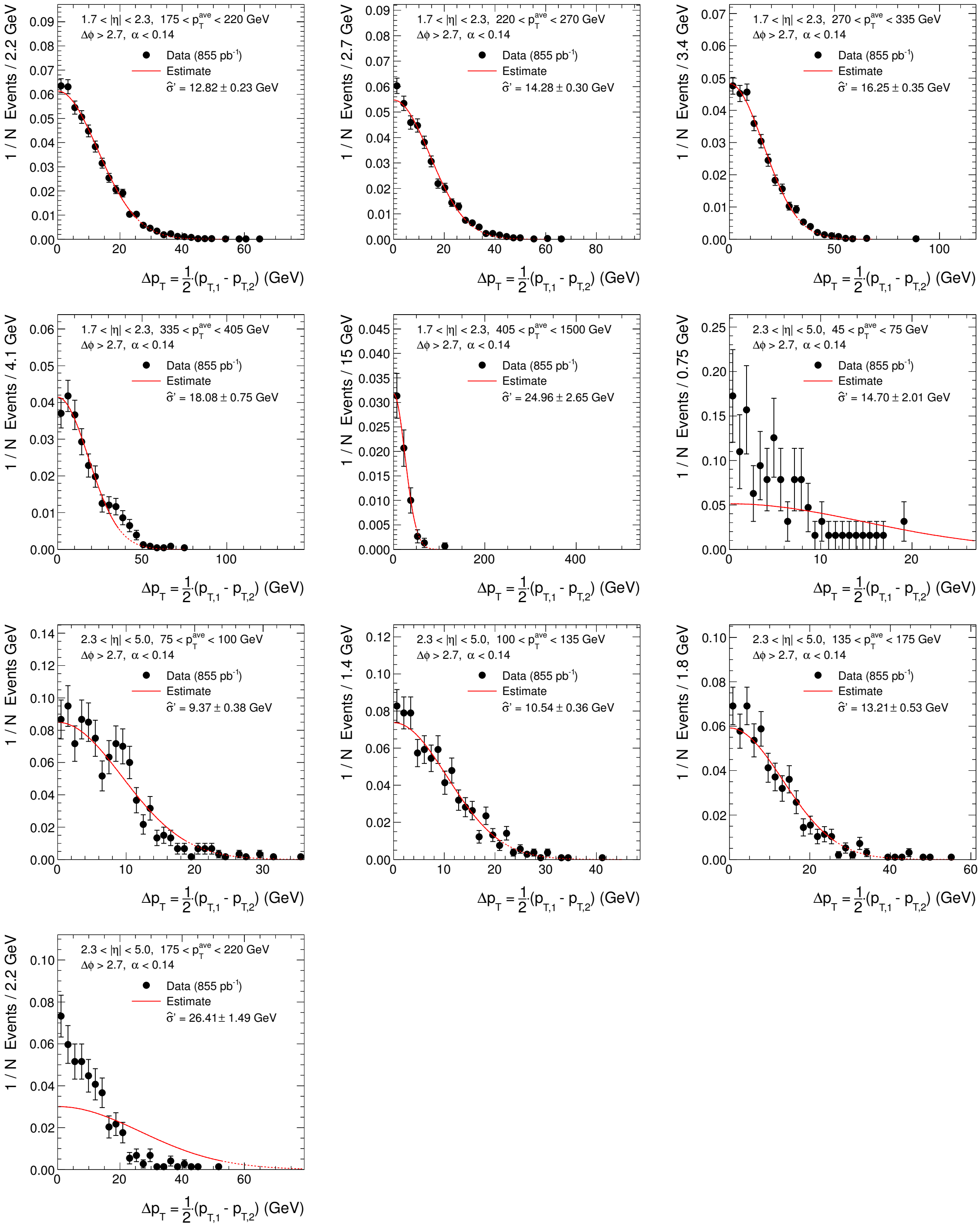


Figure A.10: Continued from Fig. A.9.

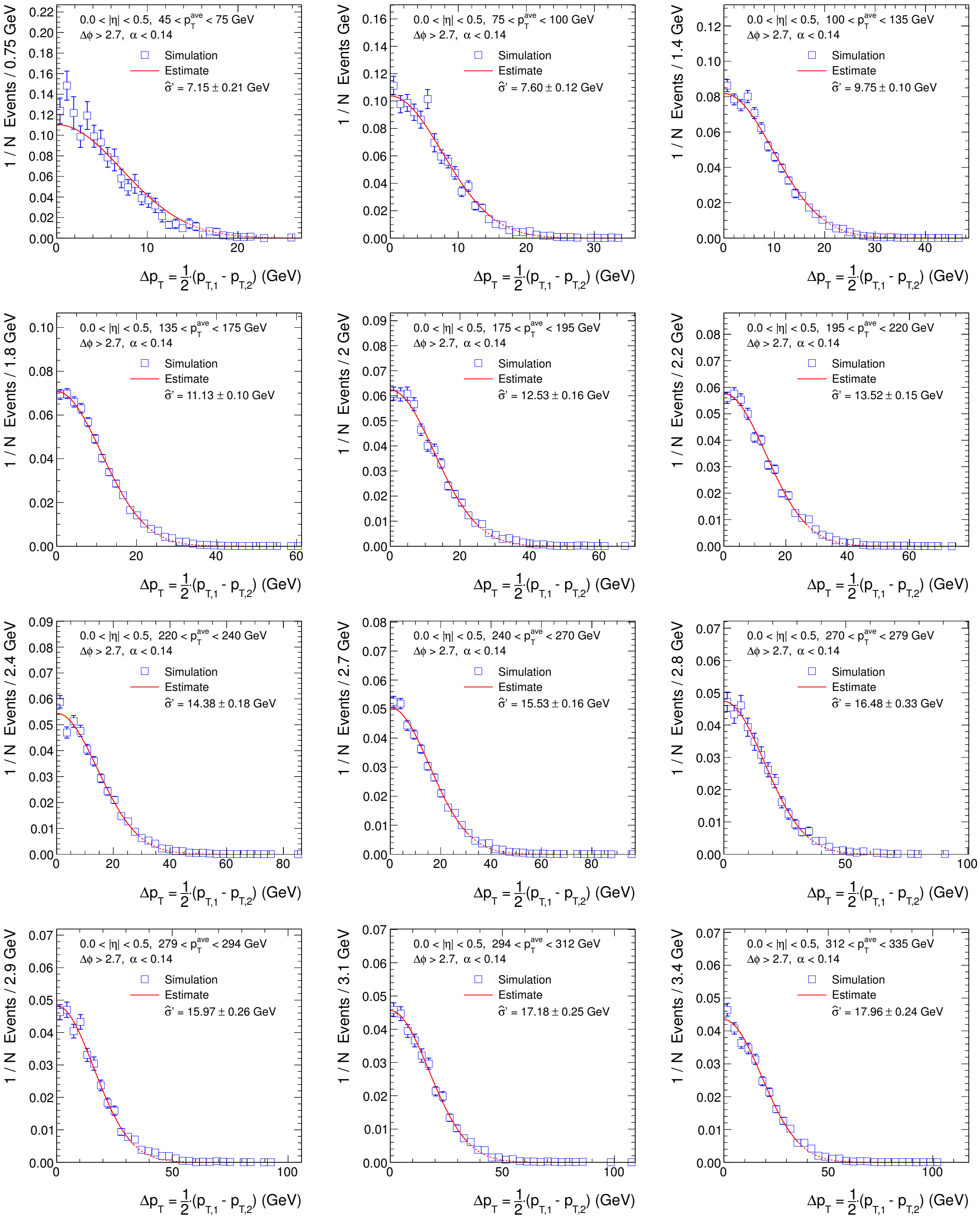


Figure A.11: Δp_{T} distributions of simulated events (*squares*) and assumed Gaussian pdf (*line*) in different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals for $\alpha < 0.14$. (Continued in Fig. A.12.)

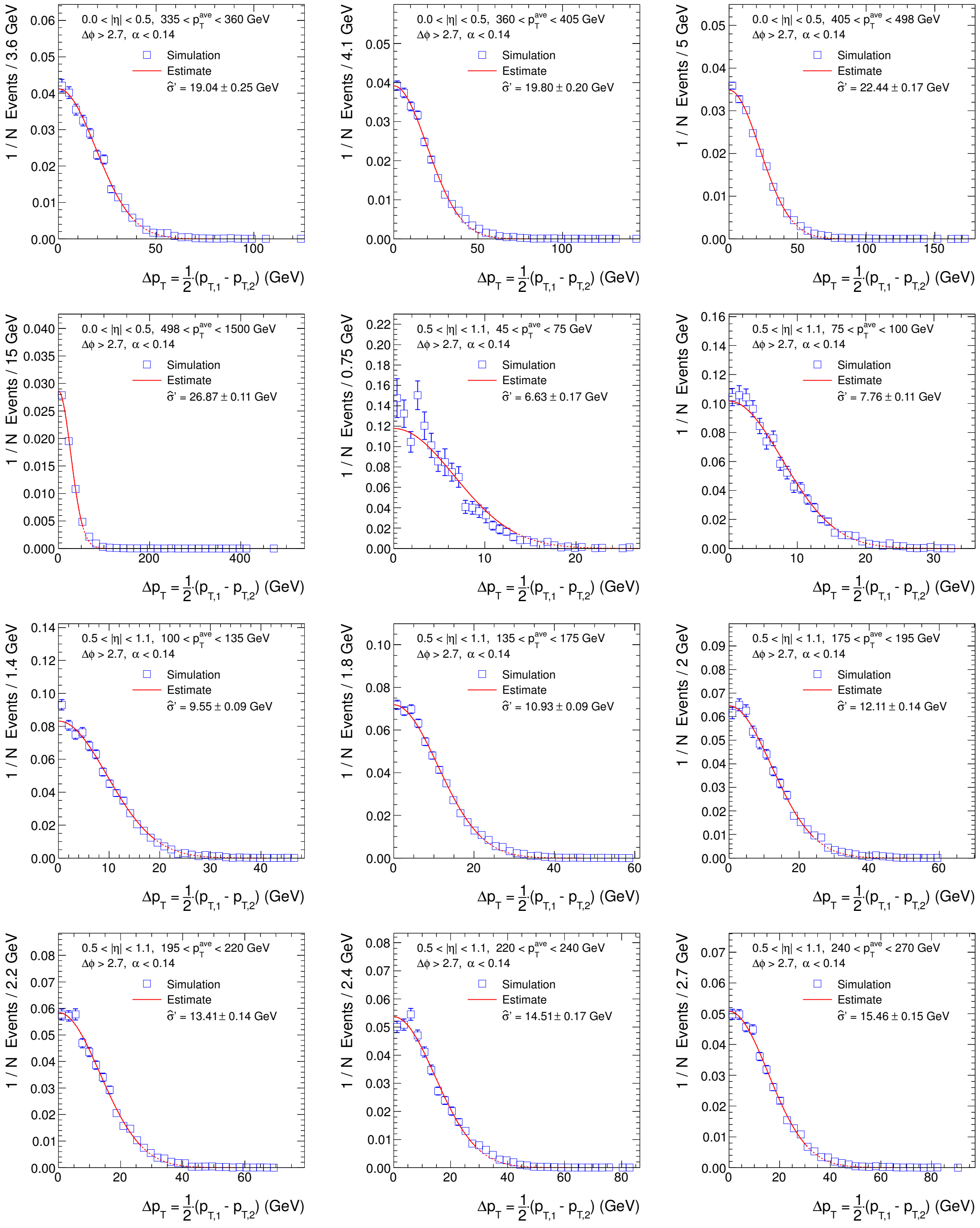


Figure A.12: Continued from Fig. A.11 and in Fig. A.13.

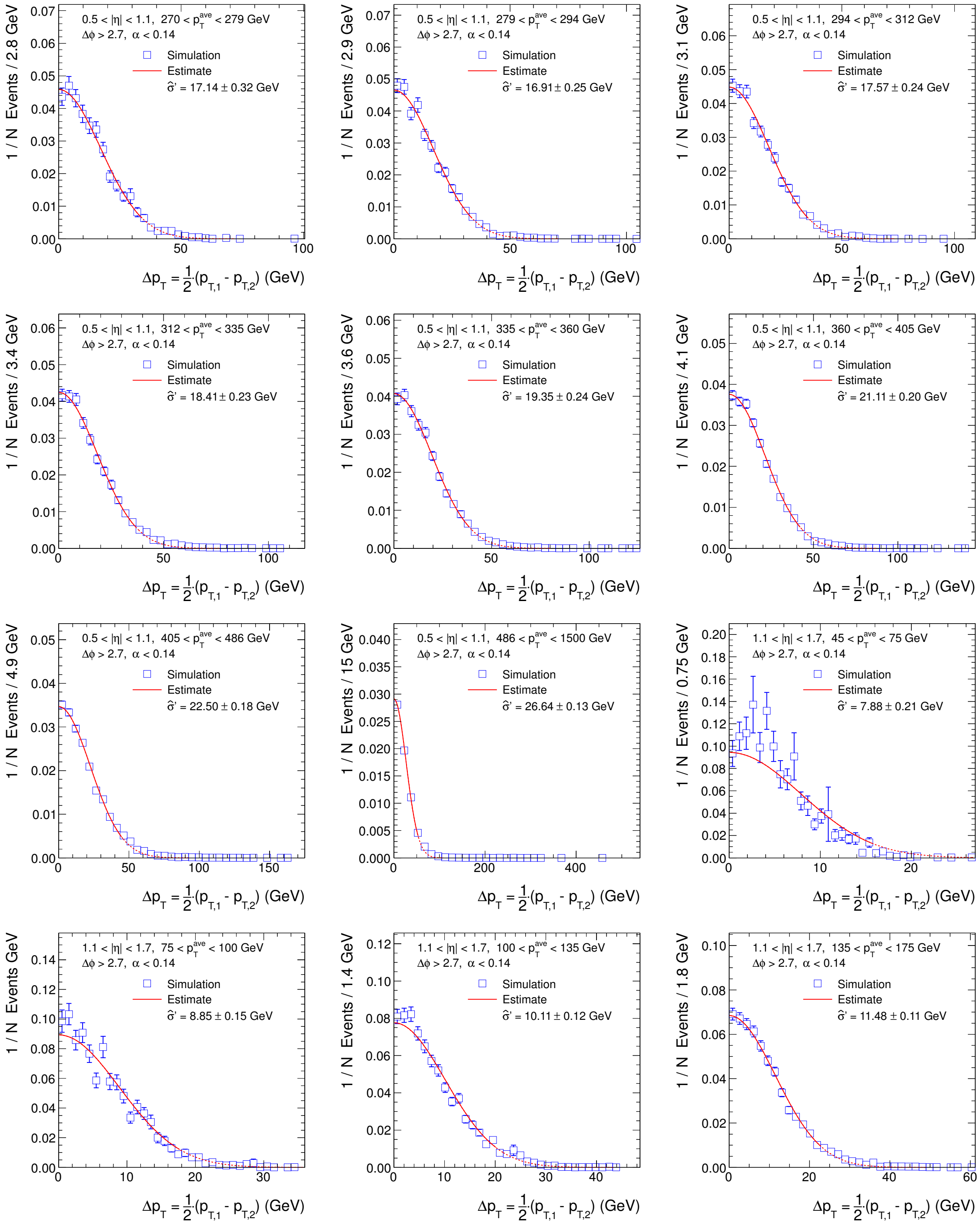


Figure A.13: Continued from Fig. A.12 and in Fig. A.14.

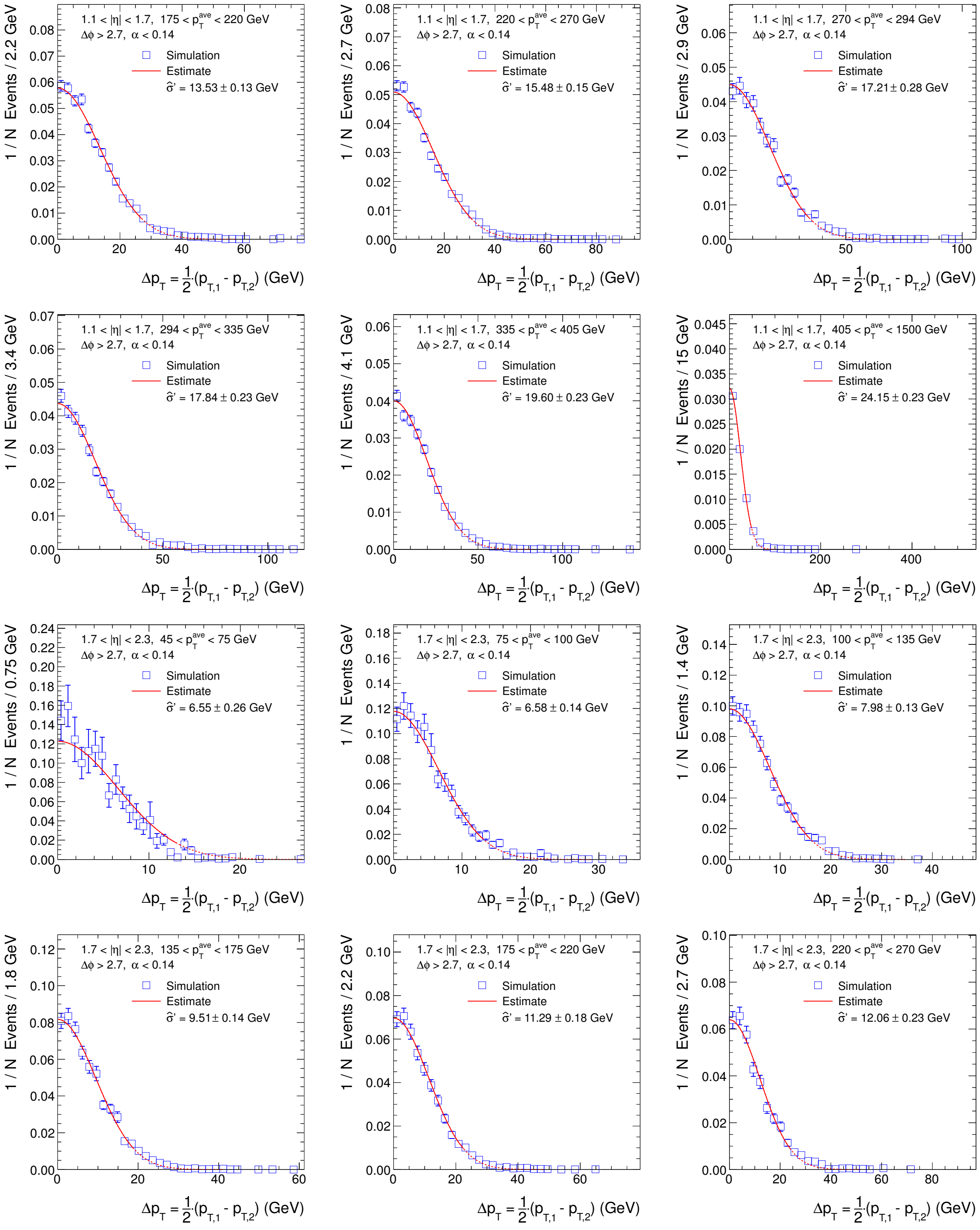


Figure A.14: Continued from Fig. A.12 and in Fig. A.15.

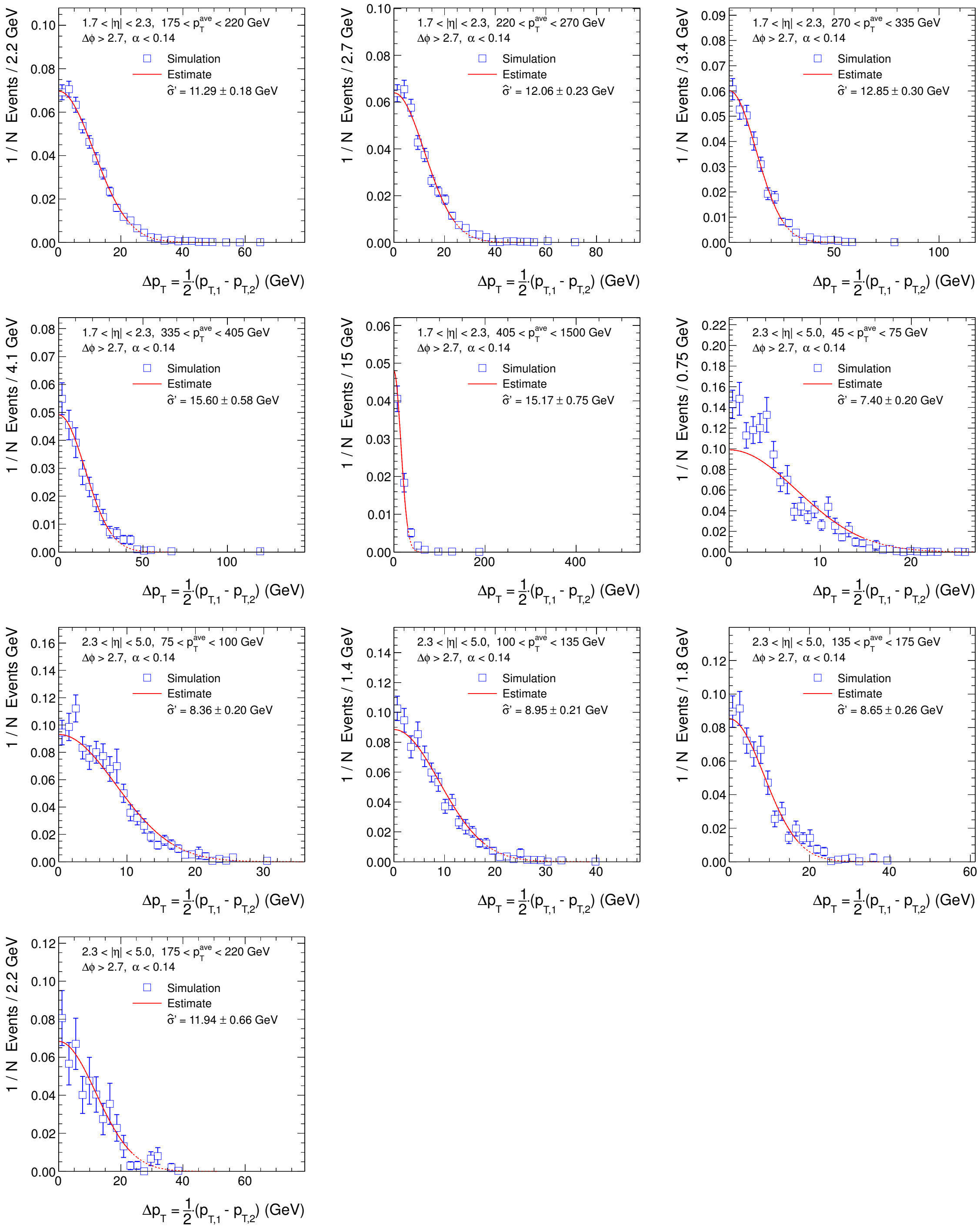


Figure A.15: Continued from Fig. A.14.

A.5 Extrapolation of σ

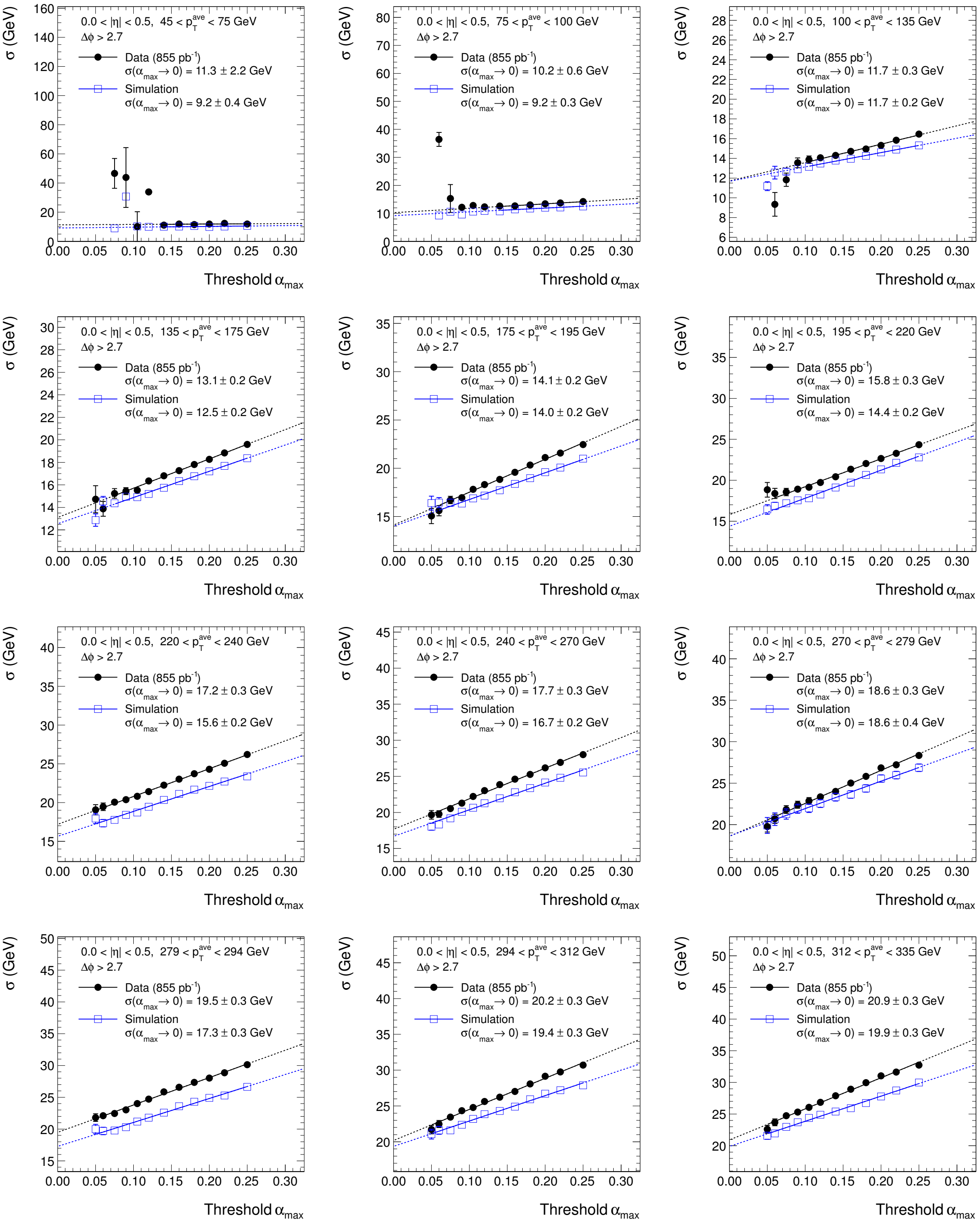


Figure A.16: Dependence of the fitted σ in data (*solid circles*) and simulation (*open squares*) on $\alpha_{\max}$ and linear extrapolation-fits (*lines*) to $\alpha_{\max} = 0$ in different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals. (Continued in Fig. A.17.)

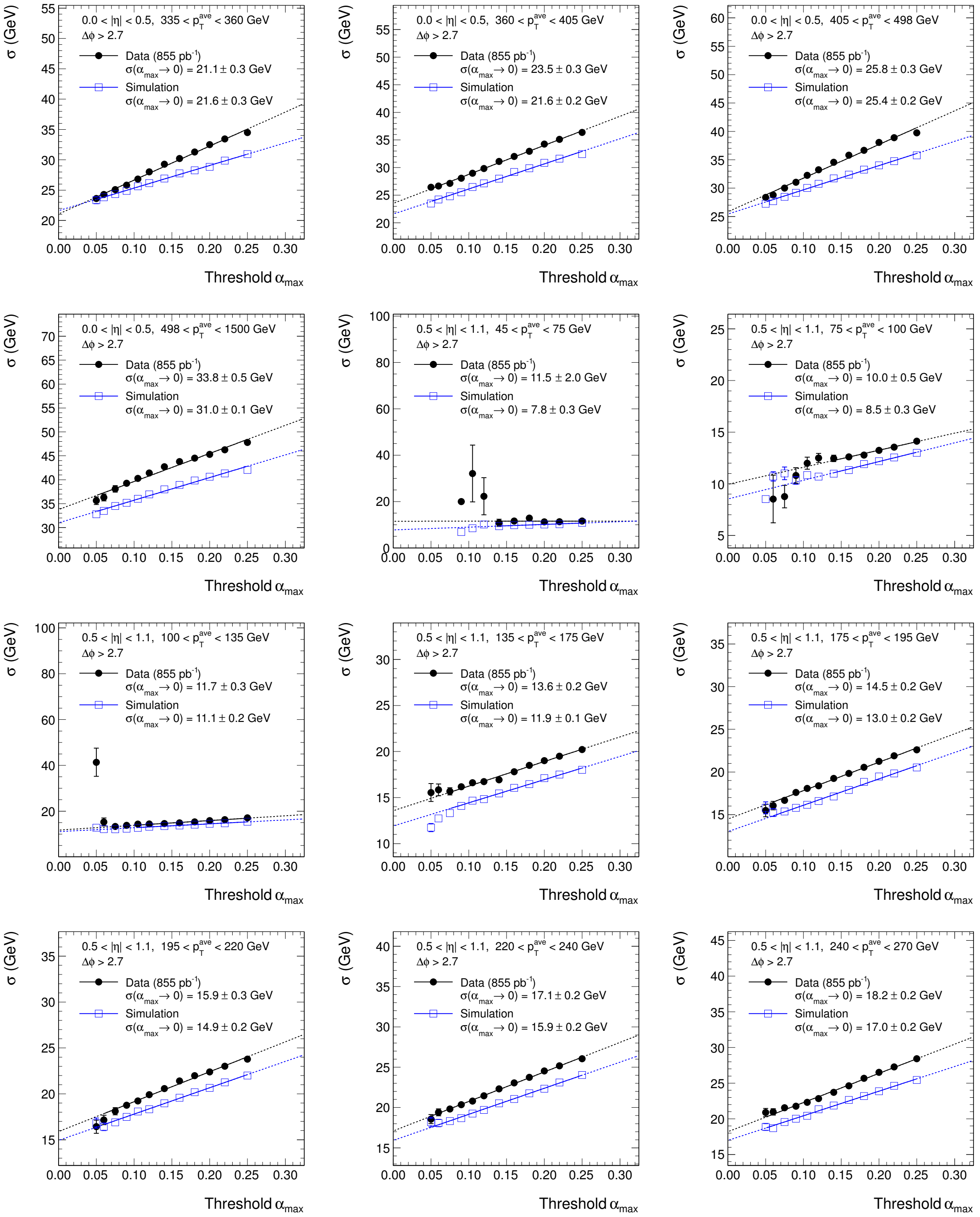


Figure A.17: Continued from Fig. A.16 and in Fig. A.18.

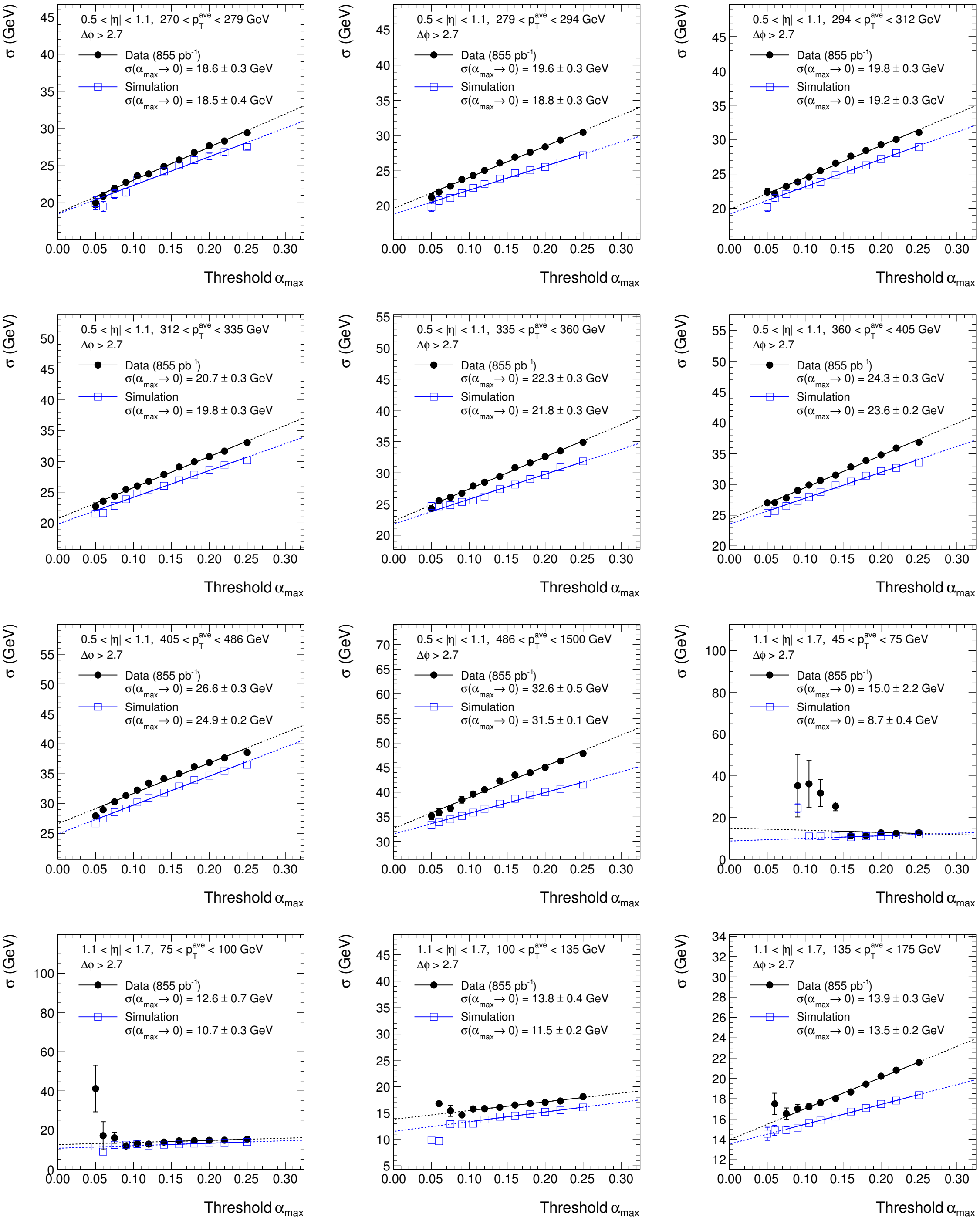


Figure A.18: Continued from Fig. A.17 and in Fig. A.19.

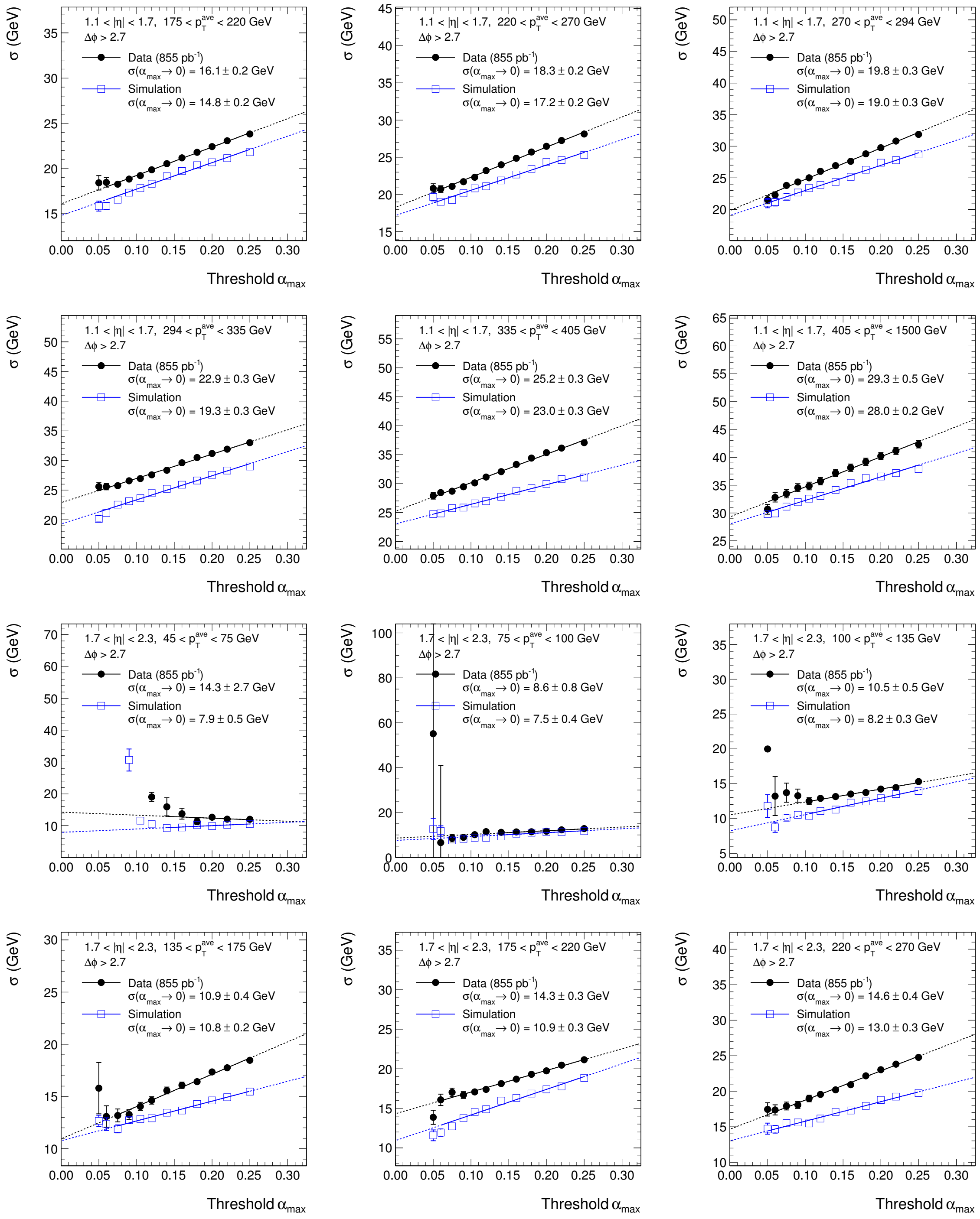


Figure A.19: Continued from Fig. A.18 and in Fig. A.20.

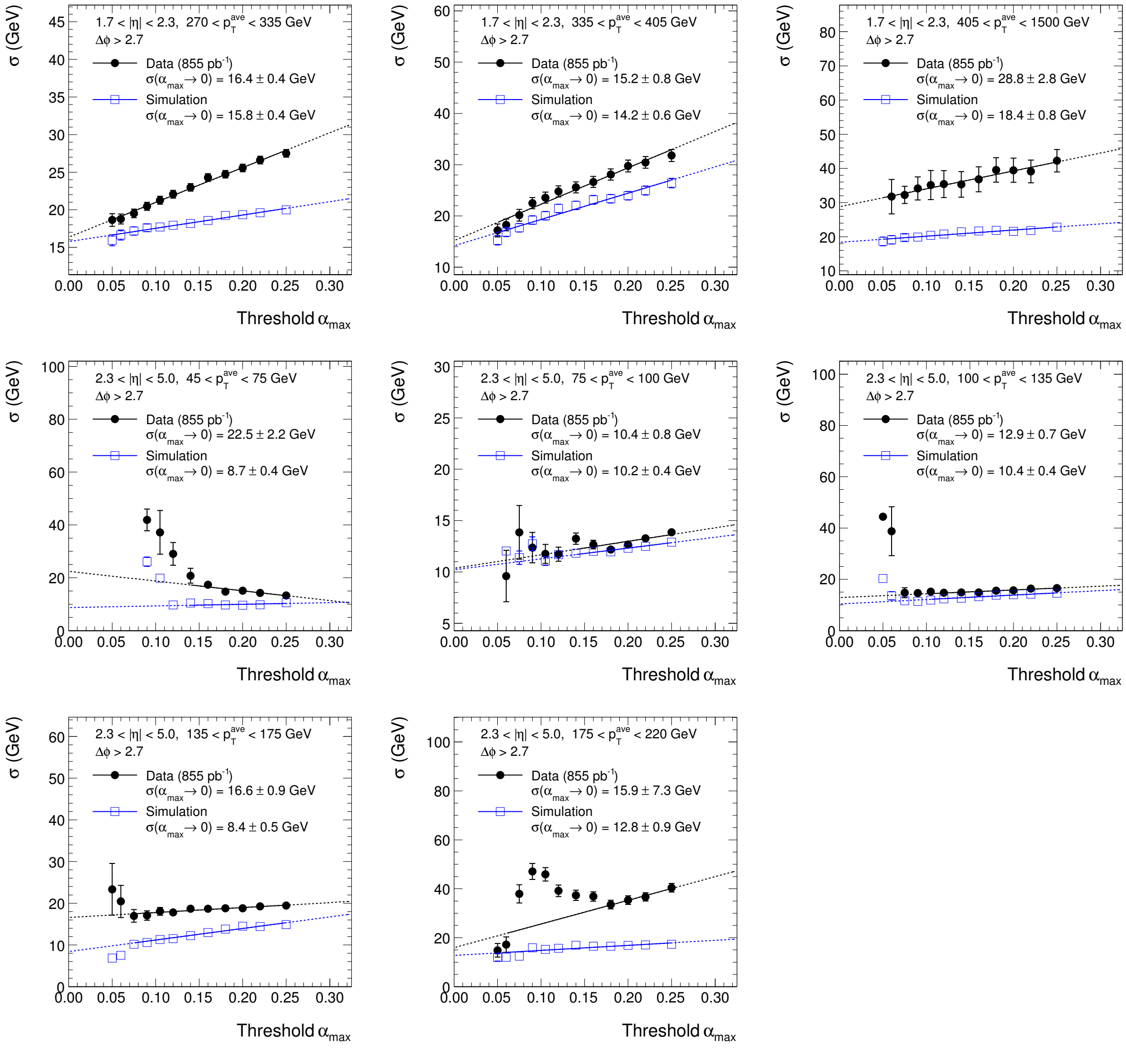


Figure A.20: Continued from Fig. A.19.

B Jet Transverse-Momentum Response Tails

B.1 Cleaning-Filter Efficiencies

Table B.1: Number N(passed) of events passing the cleaning filters described in Section 5.1.4 in $4.90\,\mathrm{fb}^{-1}$ of data and in the PYTHIA simulation. The efficiency is quoted relative to the previous selection step. The total number of data events refers to the number after the trigger requirements.

	Data		Simulation	
Filter	N(passed)	Efficiency (%)	N(passed)	Efficiency (%)
Total	21508173	—	10940800	—
Good vertex	21400293	$99.49842^{+0.00152}_{-0.00153}$	10933408	$99.93244^{+0.00078}_{-0.00079}$
HBHE noise	21213319	$99.12630^{+0.00201}_{-0.00201}$	10897178	$99.66863^{+0.00174}_{-0.00174}$
Beam halo	21161662	$99.75649^{+0.00107}_{-0.00107}$	10896684	$99.99547^{+0.00020}_{-0.00021}$
EE noise	21161658	$99.99998^{+0.00001}_{-0.00001}$	10896485	$99.99817^{+0.00013}_{-0.00013}$
Beam scraping	21161654	$99.99998^{+0.00001}_{-0.00001}$	10896485	$100.00000^{+0.00000}_{-0.00001}$
PF post-processing	21148728	$99.93892^{+0.00054}_{-0.00054}$	10890999	$99.94965^{+0.00068}_{-0.00068}$
Tracking failure	20794068	$98.32302^{+0.00279}_{-0.00280}$	10319083	$94.74873^{+0.00676}_{-0.00677}$
TP	20772589	$99.89671^{+0.00070}_{-0.00071}$	10300804	$99.82286^{+0.00131}_{-0.00131}$
BE	20543691	$98.89808^{+0.00229}_{-0.00229}$	10221909	$99.23409^{+0.00271}_{-0.00272}$
Lepton veto	20484108	$99.70997^{+0.00119}_{-0.00119}$	10210424	$99.88764^{+0.00105}_{-0.00105}$

B.2 Asymmetry Distributions

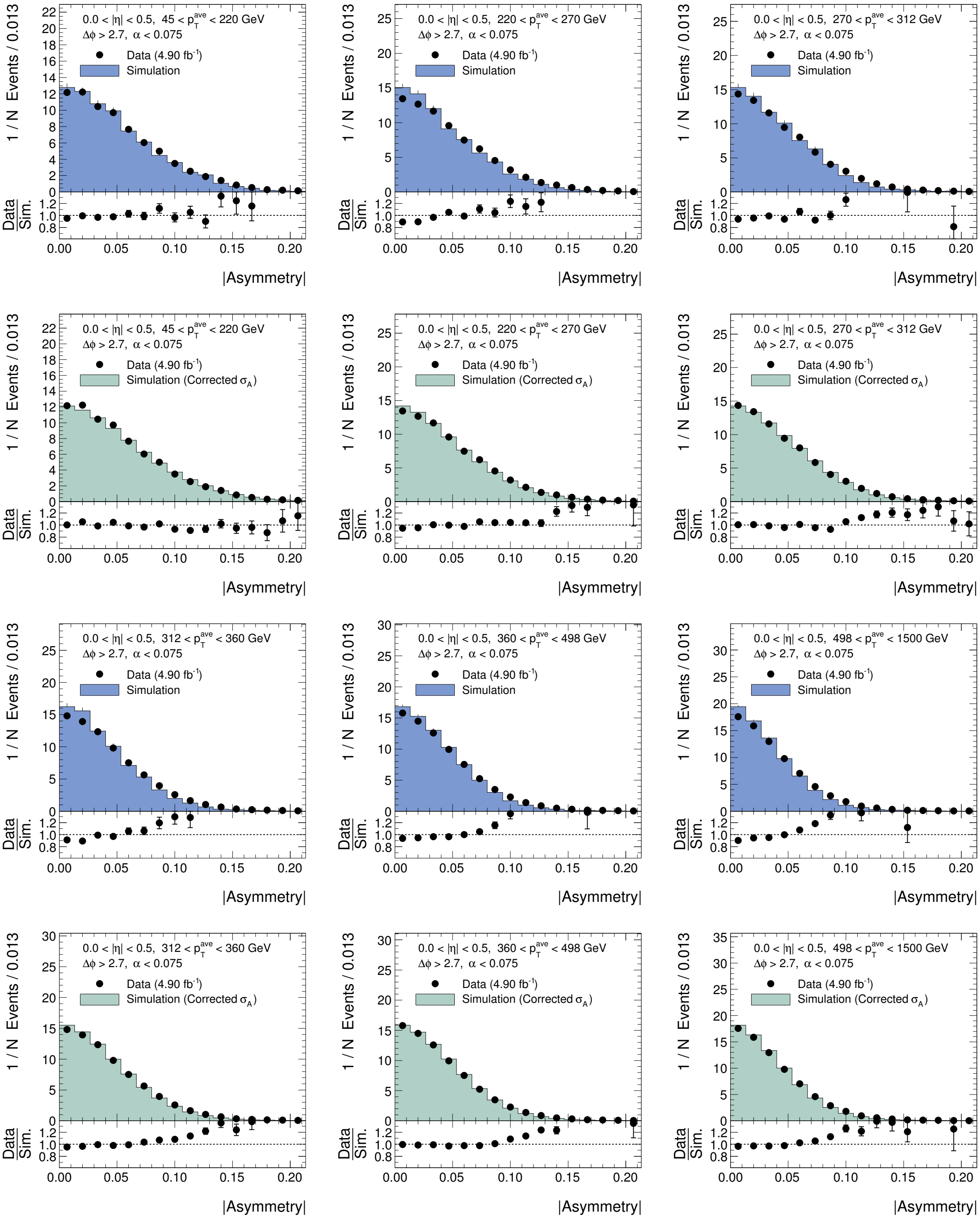


Figure B.1: Asymmetry in data (*solid circles*) and simulation (*shaded area*) before (*1. and 3. row*) and after (*2. and 4. row*) correction for the larger jet-p_{T} resolution in data in different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals for $\alpha < 0.075$. (Continued in Fig. B.2.)

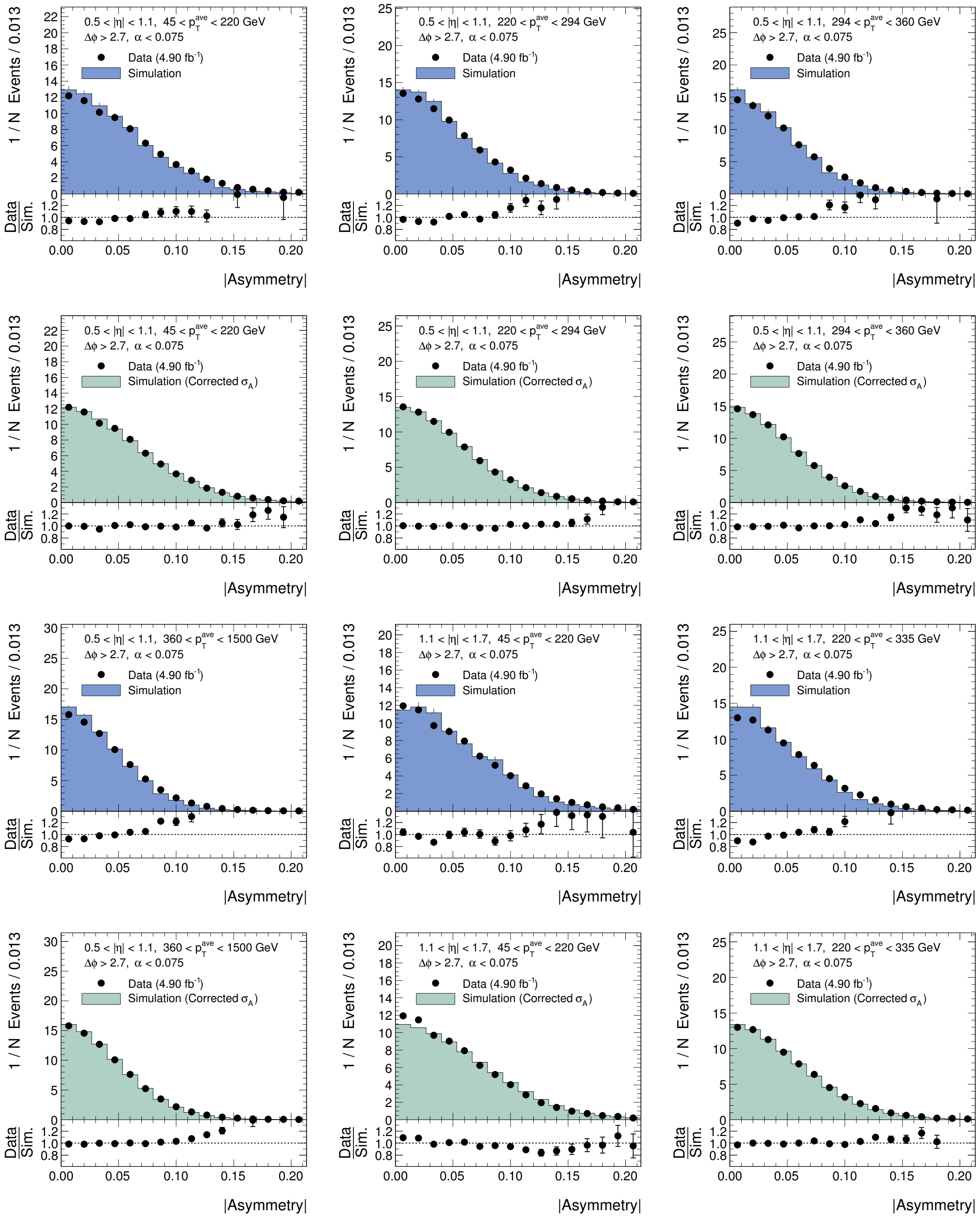


Figure B.2: Continued from Fig. B.1 and in Fig. B.3.

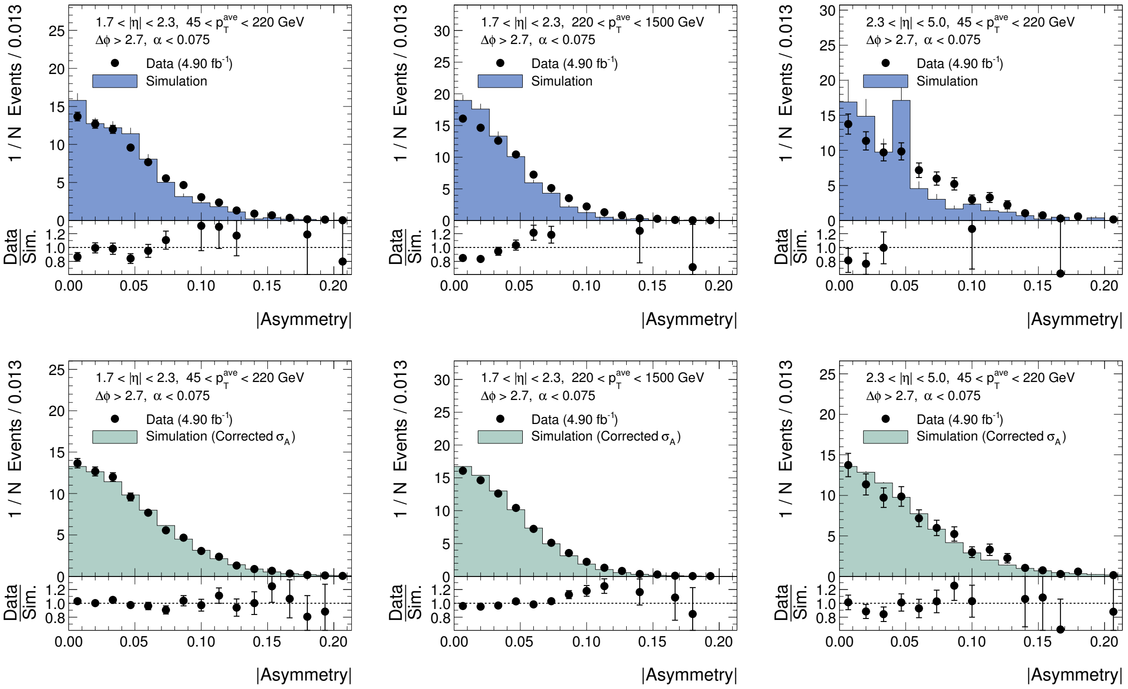


Figure B.3: Continued from Fig. B.2.

B.3 Definition of the Tail Region

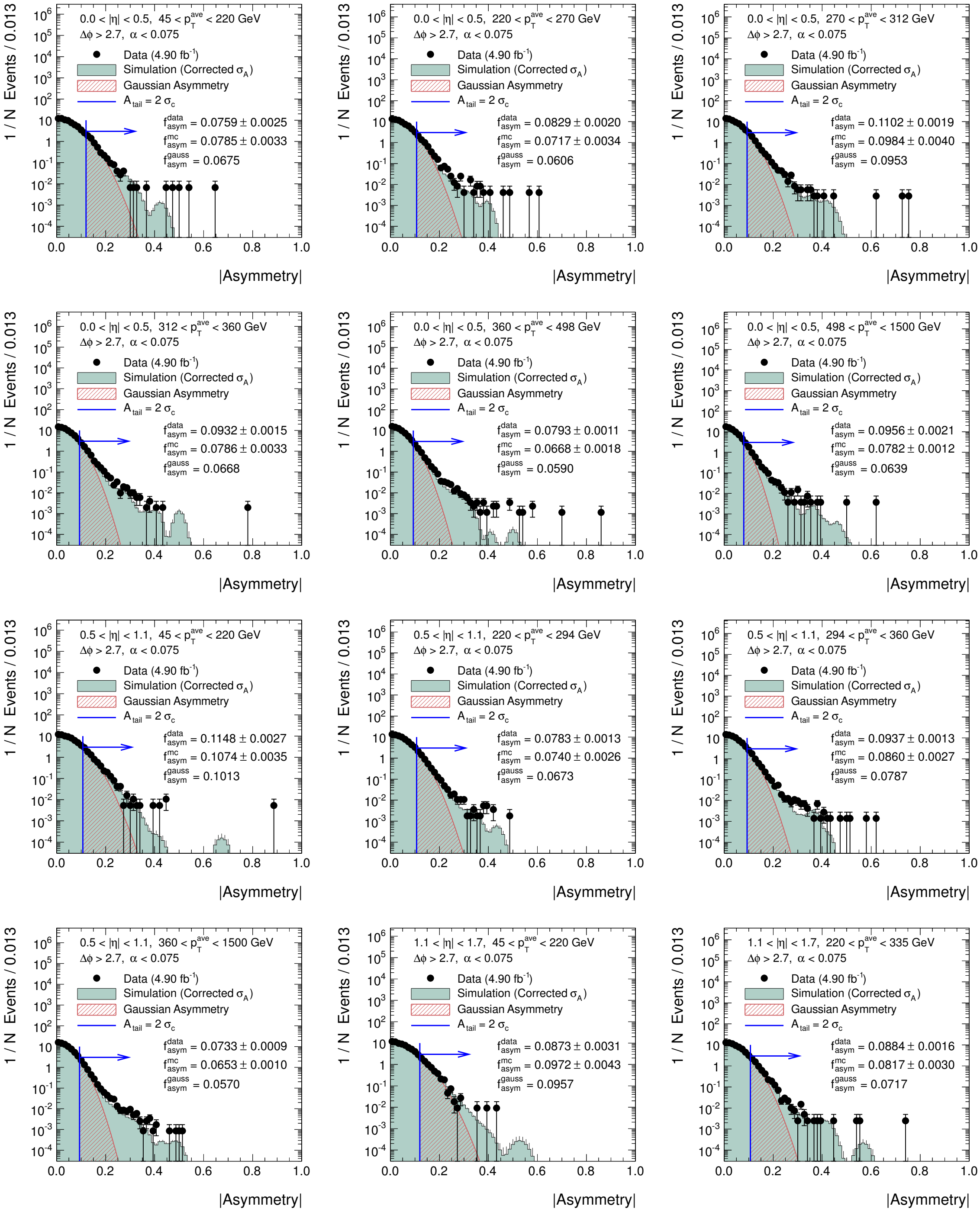


Figure B.4: Asymmetry in data (*solid circles*) and simulation (*shaded area*) in different $|\eta| \times p_T^{ave}$ intervals for $\alpha < 0.075$. The tail region $2\,\sigma_c$ and the expectation for a Gaussian distribution are indicated. (Continued in Fig. B.5.)

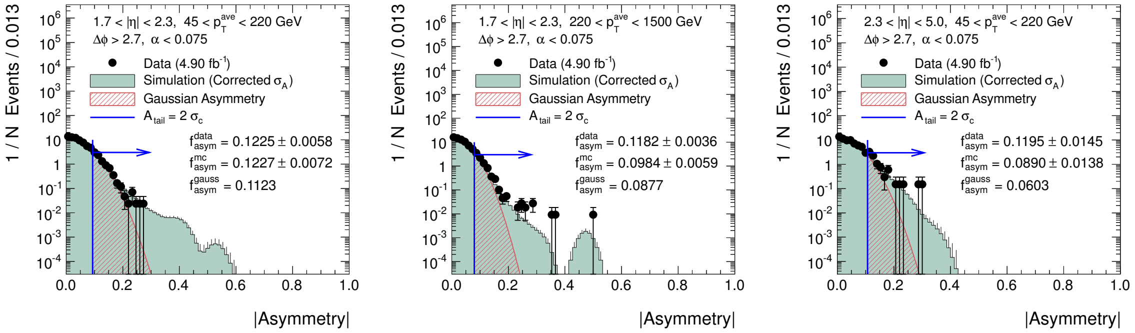


Figure B.5: Continued from Fig. B.4.

Table B.2: Effective start $\hat{\mathcal{A}}_{\text{tail}}$ of the $2\,\sigma_c$ and $3\,\sigma_c$ tail regions in different $|\eta| \times p_{\mathrm{T}}^{\text{ave}}$ intervals.

Interval		$\mathcal{A}_{\text{tail}} = 2\,\sigma_c$		$\mathcal{A}_{\text{tail}} = 3\,\sigma_c$	
$\lvert\eta\rvert$	$p_{\mathrm{T}}^{\text{ave}}$ (GeV)	$\hat{\mathcal{A}}_{\text{tail}}$	$\hat{\mathcal{A}}_{\text{tail}}/\sigma_c$	$\hat{\mathcal{A}}_{\text{tail}}$	$\hat{\mathcal{A}}_{\text{tail}}/\sigma_c$
0.0 – 0.5	45 – 220	0.120	1.979	0.173	2.858
0.0 – 0.5	220 – 270	0.107	1.961	0.160	2.942
0.0 – 0.5	270 – 312	0.093	1.794	0.147	2.819
0.0 – 0.5	312 – 360	0.093	1.909	0.133	2.727
0.0 – 0.5	360 – 498	0.093	1.967	0.133	2.810
0.0 – 0.5	498 – 1500	0.080	1.966	0.120	2.949
0.5 – 1.1	45 – 220	0.107	1.809	0.173	2.940
0.5 – 1.1	220 – 294	0.107	1.931	0.160	2.896
0.5 – 1.1	294 – 360	0.093	1.881	0.147	2.956
0.5 – 1.1	360 – 1500	0.093	1.987	0.133	2.839
1.1 – 1.7	45 – 220	0.120	1.802	0.187	2.803
1.1 – 1.7	220 – 335	0.107	1.868	0.160	2.803
1.1 – 1.7	335 – 1500	0.093	1.877	0.147	2.949
1.7 – 2.3	45 – 220	0.093	1.772	0.147	2.785
1.7 – 2.3	220 – 1500	0.080	1.790	0.133	2.983
2.3 – 5.0	45 – 220	0.107	1.814	0.173	2.948

B.4 Extrapolation of f_{asym}

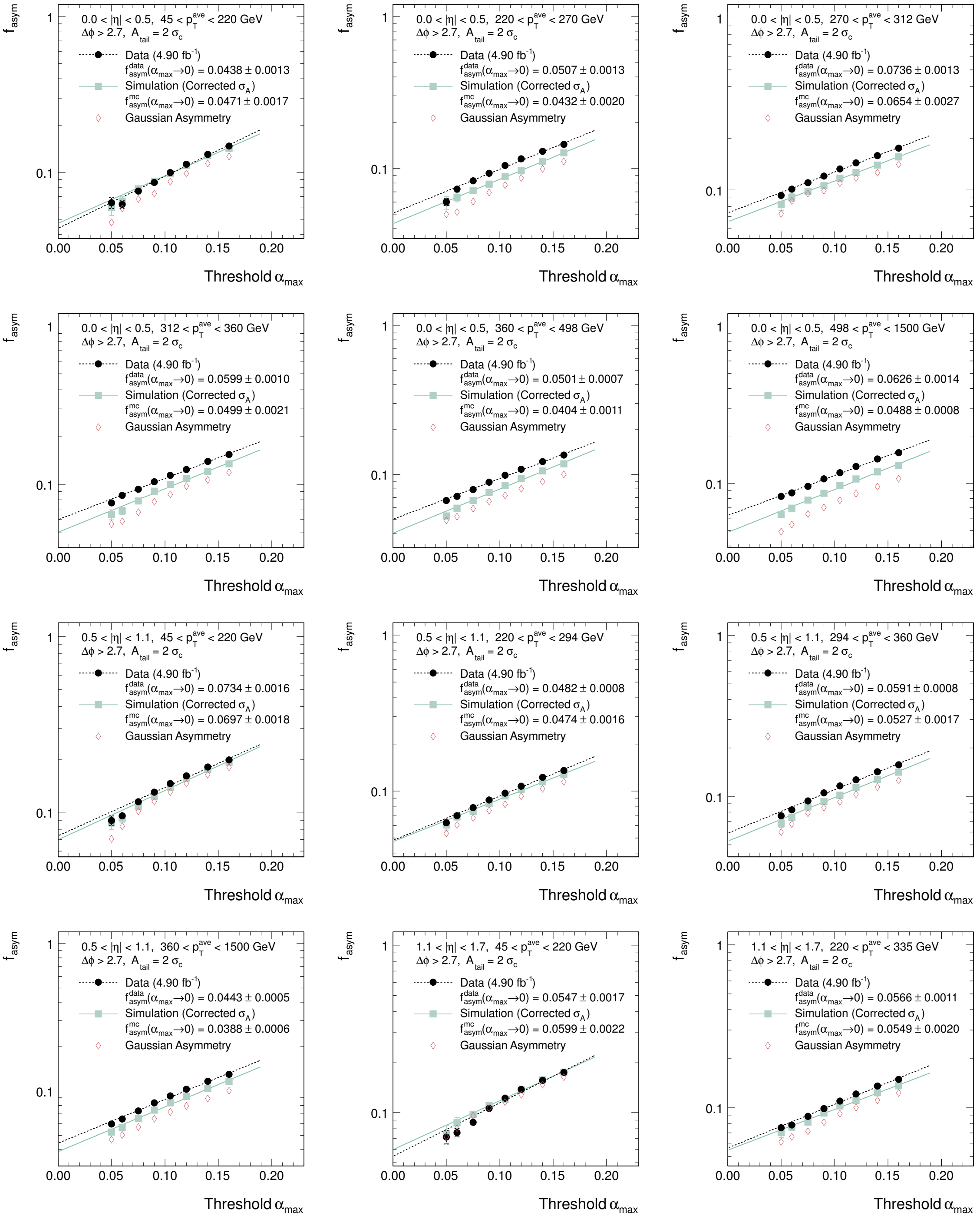


Figure B.6: Dependence of f_{asym} on $\alpha_{\max}$ in data (*circles*), simulation (*squares*), and for a Gaussian (*diamonds*) in different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals for different $\mathcal{A}_{\mathrm{tail}}$. The fit-functions to $\alpha_{\max} = 0$ are also shown. (Continued in Figs. B.7 and B.8.)

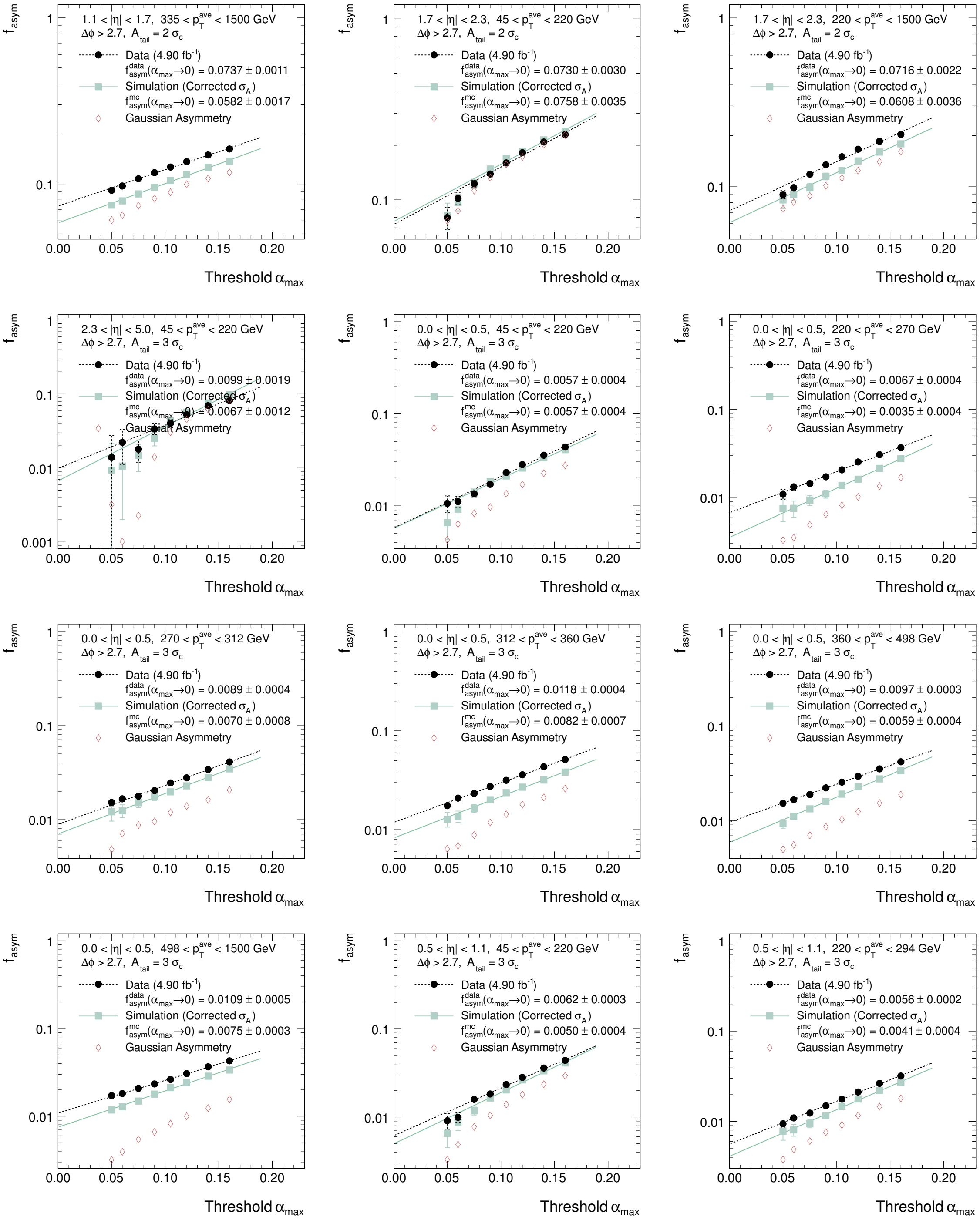


Figure B.7: Continued from Fig. B.6 and in Fig. B.8.

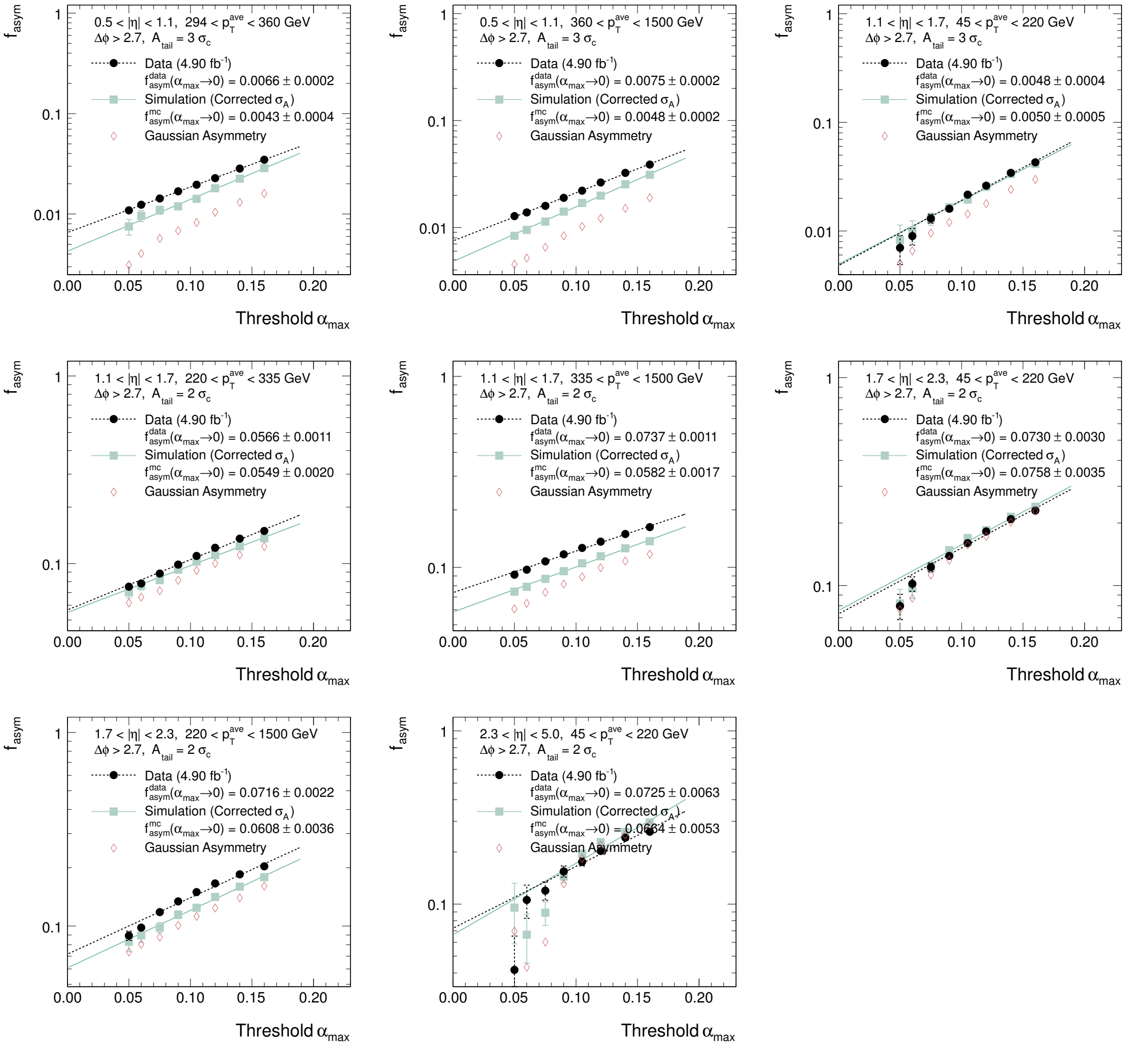


Figure B.8: Continued from Fig. B.7.

B.5 Data-to-Simulation Ratios ρ_{tail}

Table B.3: Mean $p_{\mathrm{T}}^{\mathrm{ave}}$ together with its statistical uncertainty and data-to-simulation ratio ρ_{tail} of the fractional tail-size together with the propagated extrapolation uncertainty δf_{ex} (first quoted error) and the systematic uncertainty (second quoted error) in different $|\eta| \times p_{\mathrm{T}}^{\mathrm{ave}}$ intervals for the $\mathcal{A}_{\mathrm{tail}} = 2\,\sigma_c$ and $3\,\sigma_c$ tail regions.

$\|\eta\|$	$p_{\mathrm{T}}^{\mathrm{ave}}$ (GeV)	$\langle p_{\mathrm{T}}^{\mathrm{ave}} \rangle$ (GeV)	$\rho_{\mathrm{tail}}(\mathcal{A}_{\mathrm{tail}} = 2\,\sigma_c)$	$\rho_{\mathrm{tail}}(\mathcal{A}_{\mathrm{tail}} = 3\,\sigma_c)$
0.0 − 0.5	45 − 220	189.2 ± 0.4	$0.929 \pm 0.045 \pm 0.188$	$1.012 \pm 0.102 \pm 0.315$
0.0 − 0.5	220 − 270	243.3 ± 0.2	$1.174 \pm 0.056 \pm 0.245$	$1.907 \pm 0.205 \pm 0.558$
0.0 − 0.5	270 − 312	289.5 ± 0.1	$1.126 \pm 0.046 \pm 0.307$	$1.263 \pm 0.126 \pm 0.270$
0.0 − 0.5	312 − 360	341.3 ± 0.1	$1.201 \pm 0.047 \pm 0.252$	$1.440 \pm 0.110 \pm 0.482$
0.0 − 0.5	360 − 498	413.4 ± 0.2	$1.239 \pm 0.033 \pm 0.257$	$1.646 \pm 0.086 \pm 0.455$
0.0 − 0.5	498 − 1500	607.9 ± 0.9	$1.282 \pm 0.032 \pm 0.210$	$1.450 \pm 0.081 \pm 0.291$
0.5 − 1.1	45 − 220	188.8 ± 0.4	$1.053 \pm 0.035 \pm 0.250$	$1.236 \pm 0.102 \pm 0.365$
0.5 − 1.1	220 − 294	263.6 ± 0.2	$1.017 \pm 0.038 \pm 0.237$	$1.367 \pm 0.112 \pm 0.399$
0.5 − 1.1	294 − 360	334.0 ± 0.1	$1.121 \pm 0.036 \pm 0.287$	$1.536 \pm 0.120 \pm 0.547$
0.5 − 1.1	360 − 1500	454.1 ± 0.4	$1.141 \pm 0.021 \pm 0.257$	$1.560 \pm 0.059 \pm 0.425$
1.1 − 1.7	45 − 220	189.1 ± 0.6	$0.914 \pm 0.047 \pm 0.299$	$0.970 \pm 0.122 \pm 0.444$
1.1 − 1.7	220 − 335	281.5 ± 0.3	$1.031 \pm 0.041 \pm 0.297$	$1.224 \pm 0.108 \pm 0.446$
1.1 − 1.7	335 − 1500	403.8 ± 0.5	$1.265 \pm 0.036 \pm 0.223$	$1.733 \pm 0.121 \pm 0.582$
1.7 − 2.3	45 − 220	185.2 ± 0.9	$0.963 \pm 0.061 \pm 0.326$	$0.778 \pm 0.128 \pm 0.498$
1.7 − 2.3	220 − 1500	308.8 ± 1.1	$1.177 \pm 0.070 \pm 0.363$	$1.492 \pm 0.240 \pm 0.744$
2.3 − 5.0	45 − 220	177.4 ± 3.2	$1.092 \pm 0.124 \pm 0.532$	$1.477 \pm 0.345 \pm 0.697$

B.6 A Typical Outlier Event With Large Asymmetry

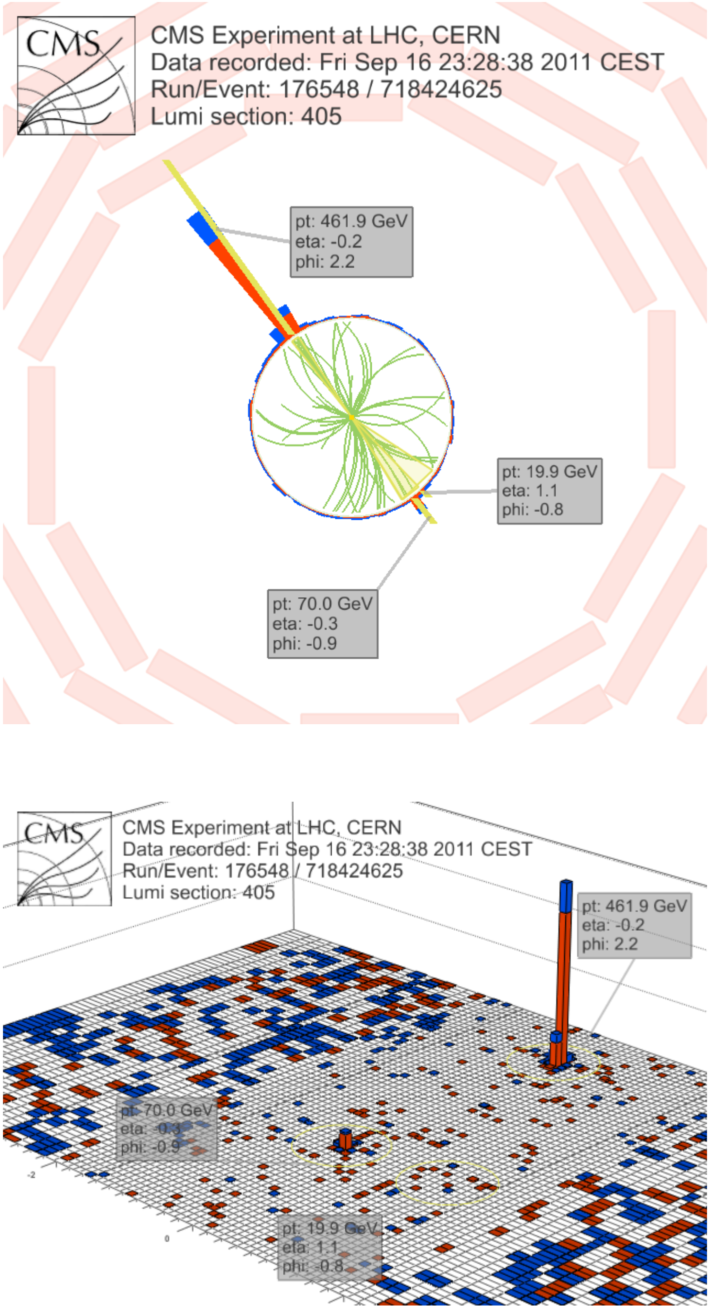


Figure B.9: Visualisation of the detector signature caused by a typical outlier dijet-event with $p_{\mathrm{T}}^{\mathrm{ave}} = 275.2\,\mathrm{GeV}$ and an asymmetry of 0.75. Shown are a transverse projection (*top*) and the $\eta \times \phi$ plane of the calorimeter towers (*bottom*). The representation of the detector signals is explained in Fig. 4.5.

Bibliography

[1] H. J. Störig, “Kleine Weltgeschichte der Philosophie”, pp. 154 – 157. Kohlhammer, Stuttgart, 17. edition, 1999. ISBN 3-17-016070-2.

[2] P. Schmüser, “Feynman-Graphen und Eichtheorien für Experimentalphysiker”. Lecture Notes in Physics. Springer, Berlin, Heidelberg, New York, London, Paris, Tokyo, 1988. ISBN 0-387-18797-9.

[3] I. Aitchison and A. Hey, “Gauge Theories in Particle Physics – Volume 1: From Relativistic Quantum Mechanics to QED”. Graduate Student Series in Physics. Taylor & Francis, third edition, 2002. ISBN 0-750-30864-8.

[4] I. Aitchison and A. Hey, “Gauge Theories in Particle Physics – Volume 2: QCD and the Electroweak Theory”. Graduate Student Series in Physics. Taylor & Francis, third edition, 2003. ISBN 0-750-30950-4.

[5] F. Mandl and G. G. Shaw, “Quantum Field Theory”. Wiley, 1993. ISBN 0-471-94186-7.

[6] K. Nakamura and et al. (Particle Data Group), “Review of particle physics”, *J. Phys.* **G 37** (2010, and 2011 partial update for the 2012 edition) 075021. `doi:10.1088/0954-3899/37/7A/075021`.

[7] E. Fermi, “Versuch einer Theorie der Betastrahlen”, *Zeitschrift für Physik* **88** (1934) 161.

[8] C. S. Wu et al. *Phys. Rev. Lett.* **105** (1957) 1413.

[9] N. Cabibbo, “Unitary Symmetry and Leptonic Decays”, *Phys. Rev. Lett.* **10** (Jun, 1963) 531 – 533. `doi:10.1103/PhysRevLett.10.531`.

[10] M. Kobayashi and T. Maskawa, “CP-Violation in the Renormalizable Theory of Weak Interaction”, *Progress of Theoretical Physics* **49** (1973), no. 2, 652–657. `doi:10.1143/PTP.49.652`.

[11] Z. Maki, M. Nakagawa, and S. Sakata, “Remarks on the unified model of elementary particles”, *Prog. Theor. Phys.* **28** (1962) 870–880. `doi:10.1143/PTP.28.870`.

[12] B. Pontecorvo, “Neutrino Experiments and the Problem of Conservation of Leptonic Charge”, *Soviet Journal of Experimental and Theoretical Physics* **26** (1968) 984.

[13] F. Close, “Neutrino”. Oxford University Press, 2010. ISBN 0-199-57459-6.

[14] B. Pontecorvo, “Neutrino Experiments and the Problem of Conservation of Leptonic Charge”, *Sov.Phys.JETP* **26** (1968) 984–988.

[15] R. Davis, D. S. Harmer, and K. C. Hoffman, "Search for Neutrinos from the Sun", *Phys. Rev. Lett.* **20** (1968) 1205 – 1209. `doi:10.1103/PhysRevLett.20.1205`.

[16] Super-Kamiokande Collaboration, "Evidence for oscillation of atmospheric neutrinos", *Phys. Rev. Lett.* **81** (1998) 1562–1567, `arXiv:hep-ex/9807003`. `doi:10.1103/PhysRevLett.81.1562`.

[17] S. Glashow, "Partial Symmetries of Weak Interactions", *Nucl.Phys.* **22** (1961) 579–588. `doi:10.1016/0029-5582(61)90469-2`.

[18] S. Weinberg, "A Model of Leptons", *Phys.Rev.Lett.* **19** (1967) 1264–1266. `doi:10.1103/PhysRevLett.19.1264`.

[19] B. Naroska, "e^+e^- physics with the jade detector at petra", *Physics Reports* **148** (1987), no. 2-3, 67 – 215. `doi:10.1016/0370-1573(87)90031-7`.

[20] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups, "Precision Electroweak Measurements on the Z Resonance", *Phys. Rept.* **427** (2006) 257, `arXiv:hep-ex/0509008`.

[21] Gargamelle Neutrino Collaboration, "Search for elastic muon neutrino electron scattering", *Phys. Lett.* **B46** (1973) 121–124. `doi:10.1016/0370-2693(73)90494-2`.

[22] Gargamelle Neutrino Collaboration, "Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment", *Phys. Lett.* **B46** (1973) 138–140. `doi:10.1016/0370-2693(73)90499-1`.

[23] UA1 Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540\,\mathrm{GeV}$", *Physics Letters B* **122** (1983), no. 1, 103 – 116. `doi:10.1016/0370-2693(83)91177-2`.

[24] UA1 Collaboration, "Experimental observation of lepton pairs of invariant mass around $95\,\mathrm{GeV/c^2}$ at the CERN SPS collider", *Physics Letters B* **126** (1983), no. 5, 398 – 410. `doi:10.1016/0370-2693(83)90188-0`.

[25] M. Y. Han and Y. Nambu, "Three-Triplet Model with Double $SU(3)$ Symmetry", *Phys. Rev.* **139** (Aug, 1965) B1006 – B1010. `doi:10.1103/PhysRev.139.B1006`.

[26] W. A. Bardeen, H. Fritzsch, and M. Gell-Mann, "Light-cone current algebra, π^0 decay, and e^+e^- annihilation", `arXiv:hep-ph/0211388`.

[27] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Non-Abelian Gauge Theories", *Phys. Rev. Lett.* **30** (Jun, 1973) 1343 – 1346. `doi:10.1103/PhysRevLett.30.1343`.

[28] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?", *Phys. Rev. Lett.* **30** (Jun, 1973) 1346 – 1349. `doi:10.1103/PhysRevLett.30.1346`.

[29] MARK-J Collaboration, "Discovery of Three Jet Events and a Test of Quantum Chromodynamics at PETRA Energies", *Phys. Rev. Lett.* **43** (1979) 830. `doi:10.1103/PhysRevLett.43.830`.

[30] PLUTO Collaboration, “Evidence for Gluon Bremsstrahlung in e^+e^- Annihilations at High-Energies”, *Phys. Lett.* **B86** (1979) 418. `doi:10.1016/0370-2693(79)90869-4`.

[31] TASSO Collaboration, “Evidence for planar events in e^+e^- annihilation at high energies”, *Physics Letters B* **86** (1979), no. 2, 243 – 249. `doi:10.1016/0370-2693(79)90830-X`.

[32] H1 and ZEUS Collaboration, “Combined Measurement and QCD Analysis of the Inclusive $e^\pm p$ Scattering Cross Sections at HERA”, *JHEP* **01** (2010) 109, `arXiv:0911.0884`. `doi:10.1007/JHEP01(2010)109`.

[33] B. Delamotte, “A Hint of renormalization”, *Am.J.Phys.* **72** (2004) 170–184, `arXiv:hep-th/0212049`. `doi:10.1119/1.1624112`.

[34] C. Berger, “Elementarteilchenphysik: Von den Grundlagen zu den modernen Experimenten”. Springer-Lehrbuch. Springer, Berlin, Heidelberg, New York, 2002. ISBN 3-540-41515-7.

[35] S. Bethke, “The 2009 world average of α_s”, *The European Physical Journal C - Particles and Fields* **64** (2009) 689–703, `arXiv:0908.1135`. `doi:10.1140/epjc/s10052-009-1173-1`.

[36] S. Mele, “Measurements of the running of the electromagnetic coupling at LEP”, in *Proceedings of the XXVI Physics in Collision Conference*, J. de Mello Neto and E. Polycarpo, eds. Bùzios, Rio de Janeiro, Brazil, 2006. `arXiv:hep-ex/0610037`. `http://www.slac.stanford.edu/econf/C060706/`.

[37] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, *Phys. Rev. Lett.* **13** (1964) 508 – 509. `doi:10.1103/PhysRevLett.13.508`.

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

[39] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (1964) 585 – 587. `doi:10.1103/PhysRevLett.13.585`.

[40] K. Nakamura and et al. (Particle Data Group), “Higgs bosons: theory and searches” in “Review of particle physics”, *J. Phys.* **G 37** (2010 and 2011 partial update for the 2012 edition) 075021. `doi:10.1088/0954-3899/37/7A/075021`.

[41] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries”, *Phys. Rev.* **127** (1962) 965 – 970. `doi:10.1103/PhysRev.127.965`.

[42] ATLAS Collaboration, “An update to the combined search for the Standard Model Higgs boson with the ATLAS detector at the LHC using up to $4.9\,\mathrm{fb}^{-1}$ of pp collision data at $\sqrt{s} = 7\,\mathrm{TeV}$”, *ATLAS Note* **ATLAS-CONF-2012-019** (2012).

[43] CMS Collaboration, “Combination of SM, SM4, FP Higgs boson searches”, *CMS Physics Analysis Summary* **CMS-PAS-HIG-12-008** (2012).

[44] T. Hambye and K. Riesselmann, "Matching conditions and Higgs mass upper bounds revisited", *Phys. Rev.* **D55** (1997) 7255–7262, `arXiv:hep-ph/9610272`. `doi:10.1103/PhysRevD.55.7255`.

[45] LEP Working Group for Higgs boson searches Collaboration, "Search for the standard model Higgs boson at LEP", *Phys. Lett.* **B565** (2003) 61–75, `arXiv:hep-ex/0306033`. `doi:10.1016/S0370-2693(03)00614-2`.

[46] TEVNPH (Tevatron New Phenomina and Higgs Working Group), CDF and DØ Collaborations, "Combined CDF and DØ Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb^{-1} of Data", `arXiv:1107.5518`. 2011. CDF Note 10606. DØ Note 6226. FERMILAB-CONF-11-354-E. Sumitted to the EPS 2011 Conference.

[47] TEVNPH (Tevatron New Phenomina and Higgs Working Group), CDF and DØ Collaborations, "Combined CDF and DØ Search for Standard Model Higgs Boson Production with up to 10 fb^{-1} of Data", 2012. CDF Note 10806. DØ Note 6303. FERMILAB-CONF-12-065-E.

[48] The ATLAS and CMS Collaborations, "Combined Standard Model Higgs boson searches with up to 2.3 fb^{-1} of pp collision data at $\sqrt{s} = 7\,\text{TeV}$ at the LHC", *CMS Physics Analysis Summary* **CMS-PAS-HIG-11-023** (2011).

[49] H. Baer and X. Tata, "Weak Scale Supersymmetry: From Superfields to Scattering Events". Cambridge University Press, 2006. ISBN 0-521-85786-4.

[50] R. P. Feynman, F. B. Morinigo, W. G. Wagner et al., "Feynman Lectures On Gravitation". Frontiers in Physics. Westview Press, 1995. ISBN 0-201-62734-5.

[51] L. M. J. S. Volders, "Neutral hydrogen in M 33 and M 101", *Bulletin of the Astronomical Institutes of the Netherlands* **14** (1959) 323. `http://adsabs.harvard.edu/abs/1959BAN....14..323V`.

[52] A. Refregier, "Weak gravitational lensing by large scale structure", *Ann.Rev.Astron.Astrophys.* **41** (2003) 645–668, `arXiv:astro-ph/0307212`. `doi:10.1146/annurev.astro.41.111302.102207`.

[53] A. Vikhlinin, A. Kravtsov, W. Forman et al., "Chandra sample of nearby relaxed galaxy clusters: Mass, gas fraction, and mass-temperature relation", *Astrophys.J.* **640** (2006) 691–709, `arXiv:astro-ph/0507092`. `doi:10.1086/500288`.

[54] D. Clowe, M. Bradac, A. H. Gonzalez et al., "A direct empirical proof of the existence of dark matter", *Astrophys.J.* **648** (2006) L109–L113, `arXiv:astro-ph/0608407`. `doi:10.1086/508162`.

[55] WMAP Collaboration, "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results", *Astrophys. J. Suppl.* **192** (2011) 14, `arXiv:1001.4744`. `doi:10.1088/0067-0049/192/2/14`.

[56] WMAP Collaboration, “Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation”, *Astrophys. J. Suppl.* **192** (2011) 18, `arXiv:1001.4538`. `doi:10.1088/0067-0049/192/2/18`.

[57] Supernova Search Team Collaboration, “Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant”, *Astron. J.* **116** (1998) 1009–1038, `arXiv:astro-ph/9805201`. `doi:10.1086/300499`.

[58] Supernova Cosmology Project Collaboration, “Measurements of Ω and Λ from 42 High-Redshift Supernovae”, *Astrophys. J.* **517** (1999) 565–586, `arXiv:astro-ph/9812133`. `doi:10.1086/307221`.

[59] P. Peebles and B. Ratra, “The Cosmological constant and dark energy”, *Rev.Mod.Phys.* **75** (2003) 559–606, `arXiv:astro-ph/0207347`. `doi:10.1103/RevModPhys.75.559`.

[60] M. Beyer, ed., “CP Violation in Particle, Nuclear, and Astrophysics”. Lecture Notes in Physics. Springer, 2010. ISBN 3-642-07830-3.

[61] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, “Electroweak symmetry breaking from dimensional deconstruction”, *Physics Letters B* **513** (2001), no. 1-2, 232 – 240. `doi:10.1016/S0370-2693(01)00741-9`.

[62] N. Arkani-Hamed, T. Gregoire, J. G. Wacker et al., “Phenomenology of Electroweak Symmetry Breaking from Theory Space”, *Journal of High Energy Physics* **2002** (2002), no. 08, 020.

[63] M. Schmaltz and D. Tucker-Smith, “LITTLE HIGGS THEORIES”, *Annual Review of Nuclear and Particle Science* **55** (2005), no. 1, 229–270. `doi:10.1146/annurev.nucl.55.090704.151502`.

[64] S. Weinberg, “Implications of dynamical symmetry breaking”, *Phys. Rev. D* **13** (1976) 974 – 996. `doi:10.1103/PhysRevD.13.974`.

[65] L. Susskind, “Dynamics of spontaneous symmetry breaking in the Weinberg-Salam theory”, *Phys. Rev. D* **20** (1979) 2619 – 2625. `doi:10.1103/PhysRevD.20.2619`.

[66] S. Dimopoulos and L. Susskind, “Mass without scalars”, *Nuclear Physics B* **155** (1979), no. 1, 237 – 252. `doi:10.1016/0550-3213(79)90364-X`.

[67] E. Eichten and K. Lane, “Dynamical breaking of weak interaction symmetries”, *Physics Letters B* **90** (1980), no. 1-2, 125 – 130. `doi:10.1016/0370-2693(80)90065-9`.

[68] R. Foadi, M. T. Frandsen, T. A. Ryttov et al., “Minimal walking technicolor: Setup for collider physics”, *Phys. Rev. D* **76** (Sep, 2007) 055005. `doi:10.1103/PhysRevD.76.055005`.

[69] H. Georgi and S. L. Glashow, “Unity of All Elementary Particle Forces”, *Phys. Rev. Lett.* **32** (1974) 438–441. `doi:10.1103/PhysRevLett.32.438`.

[70] H. Fritzsch and P. Minkowski *Ann. Phys.* **93** (1975) 193.

[71] H. Georgi, H. R. Quinn, and S. Weinberg, "Hierarchy of Interactions in Unified Gauge Theories", *Phys. Rev. Lett.* **33** (Aug, 1974) 451 – 454. `doi:10.1103/PhysRevLett.33.451`.

[72] H1 Collaboration, "Search for leptoquark bosons in *ep* collisions at HERA", *Phys.Lett.* **B629** (2005) 9–19, `arXiv:hep-ex/0506044`. `doi:10.1016/j.physletb.2005.09.048`.

[73] ZEUS Collaboration, "A Search for resonance decays to lepton + jet at HERA and limits on leptoquarks", *Phys.Rev.* **D68** (2003) 052004, `arXiv:hep-ex/0304008`. `doi:10.1103/PhysRevD.68.052004`.

[74] G. 't Hooft, "Magnetic monopoles in unified gauge theories", *Nuclear Physics B* **79** (1974), no. 2, 276 – 284. `doi:10.1016/0550-3213(74)90486-6`.

[75] M. Gell-Mann, P. Ramond, and R. Slansky in *Supergravity. Proceedings of a Workshop at Stony Brook University*, P. V. Nieuwenhuizen and D. Z. Freedman, eds., p. 315. New York, 1979.

[76] Super-Kamiokande Collaboration, "Search for Proton Decay via $p \to e^+\pi^0$ and $p \to \mu^+\pi^0$ in a Large Water Cherenkov Detector", *Phys. Rev. Lett.* **102** (2009) 141801. `doi:10.1103/PhysRevLett.102.141801`.

[77] N. ArkaniHamed, S. Dimopoulos, and G. Dvali, "The hierarchy problem and new dimensions at a millimeter", *Physics Letters B* **429** (1998), no. 3-4, 263 – 272. `doi:10.1016/S0370-2693(98)00466-3`.

[78] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, "Phenomenology, astrophysics, and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity", *Phys. Rev. D* **59** (Mar, 1999) 086004. `doi:10.1103/PhysRevD.59.086004`.

[79] "Search for microscopic black hole signatures at the Large Hadron Collider", *Physics Letters B* **697** (2011), no. 5, 434 – 453. `doi:10.1016/j.physletb.2011.02.032`.

[80] R. Haag, J. T. Łopuszański, and M. Sohnius, "All possible generators of supersymmetries of the S-matrix", *Nuclear Physics B* **88** (1975), no. 2, 257 – 274. `doi:10.1016/0550-3213(75)90279-5`.

[81] J. Wess and B. Zumino, "Supergauge transformations in four dimensions", *Nuclear Physics B* **70** (1974), no. 1, 39 – 50. `doi:10.1016/0550-3213(74)90355-1`.

[82] I. Aitchison, "Supersymmetry in Particle Physics – An Elementary Introduction". Cambridge University Press, 2007. ISBN 978-0-521-88023-7.

[83] A. H. Chamseddine, R. Arnowitt, and P. Nath, "Locally Supersymmetric Grand Unification", *Phys. Rev. Lett.* **49** (1982) 970 – 974. `doi:10.1103/PhysRevLett.49.970`.

[84] B. de Wit, "Supergravity", `arXiv:hep-th/0212245`. 2001 Les Houches Summerschool "Unity from Duality: Gravity, Gauge Theory and Strings".

[85] S. Dimopoulos, S. Raby, and F. Wilczek, "Supersymmetry and the scale of unification", *Phys. Rev. D* **24** (1981) 1681 – 1683. `doi:10.1103/PhysRevD.24.1681`.

[86] S. P. Martin, "A Supersymmetry primer", `arXiv:hep-ph/9709356`.

[87] G. Bertone, D. Hooper, and J. Silk, "Particle dark matter: evidence, candidates and constraints", *Physics Reports* **405** (2005), no. 5-6, 279 – 390. `doi:10.1016/j.physrep.2004.08.031`.

[88] J. L. Feng, "Dark Matter Candidates from Particle Physics and Methods of Detection", *Annual Review of Astronomy and Astrophysics* **48** (2010), no. 1, 495–545. `doi:10.1146/annurev-astro-082708-101659`.

[89] CMS Collaboration, "Search for Neutral MSSM Higgs Bosons Decaying to Tau Pairs in pp Collisions at $\sqrt{s} = 7\,\mathrm{TeV}$", *Phys.Rev.Lett.* **106** (2011) 231801, `arXiv:1104.1619`. `doi:10.1103/PhysRevLett.106.231801`.

[90] K. Nakayama, N. Yokozaki, and K. Yonekura, "Relaxing the Higgs mass bound in singlet extensions of the MSSM", *JHEP* **1111** (2011) 021, `arXiv:1108.4338`. `doi:10.1007/JHEP11(2011)021`.

[91] G. L. Kane, C. F. Kolda, L. Roszkowski et al., "Study of constrained minimal supersymmetry", *Phys.Rev.* **D49** (1994) 6173–6210, `arXiv:hep-ph/9312272`. `doi:10.1103/PhysRevD.49.6173`.

[92] H. Baer, C. Balazs, A. Belyaev et al., "Updated constraints on the minimal supergravity model", *JHEP* **0207** (2002) 050, `arXiv:hep-ph/0205325`.

[93] H1 Collaboration, "A Search for Selectrons and Squarks at HERA", *Phys. Lett.* **B 380** (1996) 461, `arXiv:hep-ex/9605002`. `doi:10.1016/0370-2693(96)00640-5`.

[94] ZEUS Collaboration, "Search for Stop Production in R-Parity-Violating Supersymmetry at HERA", *Eur. Phys. J. C* **50** (2007) 269, `arXiv:hep-ex/0611018`. `doi:10.1140/epjc/s10052-007-0240-8`.

[95] ALEPH Collaboration, "Absolute Mass Lower Limit for the Lightest Neutralino of the MSSM from $\mathrm{e}^+\mathrm{e}^-$ data at $\sqrt{s}$ up to 209 GeV", *Phys. Lett.* **B 583** (2004) 247. See also references therein. `doi:10.1016/j.physletb.2003.12.066`.

[96] DELPHI Collaboration, "Searches for Supersymmetric Particles in $\mathrm{e}^+\mathrm{e}^-$ Collisions up to 208 GeV and Interpretation of the Results within the MSSM", *Eur. Phys. J. C* **31** (2003) 421, `arXiv:hep-ex/0311019`. See also references therein. `doi:10.1140/epjc/s2003-01355-5`.

[97] L3 Collaboration, "Search for Scalar Leptons and Scalar Quarks at LEP", *Phys. Lett.* **B 580** (2004) 37, `arXiv:hep-ex/0310007`. See also references therein. `doi:10.1016/j.physletb.2003.10.010`.

[98] OPAL Collaboration, "Search for Chargino and Neutralino Production at $\sqrt{s} = 192$ GeV to 209 GeV at LEP", *Eur. Phys. J. C* **35** (2004) 1, `arXiv:hep-ex/0401026`. See also references therein. `doi:10.1140/epjc/s2004-01780-x`.

[99] ALEPH, DELPHI, L3 and OPAL Collaborations, “Joint SUSY Working Group”, *LEPSUSYWG/02-06-2* (2004). http://lepsusy.web.cern.ch/lepsusy.

[100] CDF Collaboration, Altonen et al., “The CDF exclusion region in the $m_{1/2}$ vs. m_0 plane”, *Phys. Rev. Lett* **102** (2009) 121801, `arXiv:011.2512`.

[101] DØ Collaboration, Abazov et al., “Search for squarks and gluinos in events with jets and missing transverse energy using 2.1 fb^{-1} of $p\bar{p}$ collision data at $\sqrt{s} = 1.96\,\text{TeV}$”, *Phys. Lett.* **B 660** (2008) 449, `arXiv:0712.3805`.

[102] DØ Collaboration, “Search for associated production of charginos and neutralinos in the trilepton final state using 2.3 fb^{-1} of data”, *Phys. Lett.* **B 680** (2009), no. 1, 34 – 43. `doi:DOI: 10.1016/j.physletb.2009.08.011`.

[103] The ALEPH, CDF, DØ, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD electroweak and heavy flavour groups, “Precision Electroweak Measurements and Constraints on the Standard Model”, `arXiv:1012.2367`. CERN-PH-EP-2010-095, FERMILAB-TM-2480-PPD, LEPEWWG-2010-01, TEVEWWG-2010-01, ALEPH-2010-001-PHYSICS-2010-001, CDF-Note-10338, DELPHI-DELPHI-2010-001-PHYS 952, L3-Note-2837, OPAL-PR-432, SLAC-PUB-14301.

[104] K. Hagiwara, A. D. Martin, D. Nomura et al., “Improved predictions for $g-2$ of the muon and $\alpha_{\text{QED}}(M_Z^2)$”, *Phys. Lett.* **B649** (2007) 173–179, `arXiv:hep-ph/0611102`. `doi:10.1016/j.physletb.2007.04.012`.

[105] Muon G-2 Collaboration, “Final report of the muon E821 anomalous magnetic moment measurement at BNL”, *Phys. Rev.* **D73** (2006) 072003, `arXiv:hep-ex/0602035`. `doi:10.1103/PhysRevD.73.072003`.

[106] CLEO Collaboration, “Branching fraction and photon energy spectrum for $b \to s\gamma$”, *Phys. Rev. Lett.* **87** (2001) 251807, `arXiv:hep-ex/0108032`. `doi:10.1103/PhysRevLett.87.251807`.

[107] BaBar Collaboration, “Measurement of the branching fraction and photon energy moments of $B \to X(s)\gamma$ and $A_{\text{CP}}(B \to X_{s+d}\gamma)$”, *Phys. Rev. Lett.* **97** (2006) 171803, `arXiv:hep-ex/0607071`. `doi:10.1103/PhysRevLett.97.171803`.

[108] Belle Collaboration, “Measurement of Inclusive Radiative B-meson Decays with a Photon Energy Threshold of 1.7 GeV”, *Phys. Rev. Lett.* **103** (2009) 241801, `arXiv:0907.1384`. `doi:10.1103/PhysRevLett.103.241801`.

[109] DØ Collaboration, “Search for the rare decay $B_s^0 \to \mu^+\mu^-$”, *Phys. Lett.* **B693** (2010) 539–544, `arXiv:1006.3469`. `doi:10.1016/j.physletb.2010.09.024`.

[110] J. A. H. Morata, “Rare decays in LHCb”. LHCb-TALK-2012-028, 2012. 47th Rencontres de Moriond on Electroweak Interactions and Unified Theories, La Thuile, Italy.

[111] The CMS and LHCb Collaborations, “Search for the rare decay $B_s^0 \to \mu^+\mu^-$ at the LHC with the CMS and LHCb experiments”,. CMS-PAS-BPH-11-019, LHCb-ANA-2011-039.

[112] CDF Collaboration, “Search for $B_s^0 \to \mu^+\mu^-$ and $B_d^0 \to \mu^+\mu^-$ Decays with CDF II”, *Phys. Rev. Lett.* **107** (2011) 191801, `arXiv:1107.2304`.

[113] O. Buchmueller, R. Cavanaugh, D. Colling et al., “Supersymmetry and Dark Matter in Light of LHC 2010 and Xenon100 Data”, *Eur.Phys.J.* **C71** (2011) 1722, `arXiv:1106.2529`. `doi:10.1140/epjc/s10052-011-1722-2`.

[114] XENON100 Collaboration Collaboration, “Dark Matter Results from 100 Live Days of XENON100 Data”, *Phys.Rev.Lett.* **107** (2011) 131302, `arXiv:1104.2549`. `doi:10.1103/PhysRevLett.107.131302`.

[115] L. Roszkowski, R. Ruiz de Austri, and T. Nihei, “New cosmological and experimental constraints on the CMSSM”, *JHEP* **0108** (2001) 024, `arXiv:hep-ph/0106334`. `doi:10.1088/1126-6708/2001/08/024`.

[116] J. Ellis, J. Hagelin, D. Nanopoulos et al., “Supersymmetric relics from the big bang”, *Nuclear Physics B* **238** (1984), no. 2, 453 – 476. `doi:10.1016/0550-3213(84)90461-9`.

[117] P. Bechtle, K. Desch, M. Uhlenbrock et al., “Constraining SUSY models with Fittino using measurements before, with and beyond the LHC”, *Eur. Phys. J.* **C66** (2010) 215–259, `arXiv:0907.2589`. `doi:10.1140/epjc/s10052-009-1228-3`.

[118] P. Bechtle, T. Bringmann, K. Desch et al., “Constrained Supersymmetry after two years of LHC data: a global view with Fittino”, `arXiv:1204.4199`.

[119] H. Flacher et al., “Revisiting the Global Electroweak Fit of the Standard Model and Beyond with Gfitter”, *Eur. Phys. J.* **C60** (2009) 543–583, `arXiv:0811.0009`. `doi:10.1140/epjc/s10052-009-0966-6`.

[120] R. Trotta, F. Feroz, M. Hobson et al., “The impact of priors and observables on parameter inferences in the constrained MSSM”, *Journal of High Energy Physics* **2008** (2008), no. 12, 024, `arXiv:0809.3792`. `doi:10.1088/1126-6708/2008/12/024`.

[121] S. Sekmen, S. Kraml, J. Lykken et al., “Interpreting LHC SUSY searches in the phenomenological MSSM”, *JHEP* **1202** (2012) 075, `arXiv:1109.5119`. `doi:10.1007/JHEP02(2012)075`.

[122] O. Buchmueller, R. Cavanaugh, A. De Roeck et al., “Prediction for the Lightest Higgs Boson Mass in the CMSSM using Indirect Experimental Constraints”, *Phys.Lett.* **B657** (2007) 87–94, `arXiv:0707.3447`. `doi:10.1016/j.physletb.2007.09.058`.

[123] O. Buchmüller, R. Cavanaugh, D. Colling et al., “The MasterCode project”, 2011. `http://mastercode.web.cern.ch/mastercode`.

[124] P. J. Bryant, “A brief history and review of accelerators”, in *CAS – CERN Accelerator School: 5th General Accelerator Physics Course*, S. Turner, ed., pp. 1 – 16, CERN. CERN, Jyväskylä, Finland, 1992. CERN-94-01-V-1. CERN-94-01-V-2.

[125] L. Evans and P. Bryant, “LHC Machine”, *Journal of Instrumentation* **3** (2008), no. 08, S08001. `doi:10.1088/1748-0221/3/08/S08001`.

[126] O. S. Brüning, P. Collier, P. Lebrun et al., “LHC Design Report”, volume 1: The LHC Main Ring. CERN, Geneva, 2004. CERN-2004-003-V-1.

[127] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* **3** (2008) S08003. `doi:10.1088/1748-0221/3/08/S08003`.

[128] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004. `doi:10.1088/1748-0221/3/08/S08004`.

[129] CMS Collaboration, “CMS Detector Performance and Software”, volume 1 of *CMS Physics Technical Design Report*. CERN, Geneva, 2006. CMS-TDR-008-1. CERN-LHCC-2006-001.

[130] LHCb Collaboration, “The LHCb Detector at the LHC”, *JINST* **3** (2008) S08005. `doi:10.1088/1748-0221/3/08/S08005`.

[131] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST* **3** (2008) S08002. `doi:10.1088/1748-0221/3/08/S08002`.

[132] J. M. Campbell, J. W. Huston, and W. J. Stirling, “Hard interactions of quarks and gluons: a primer for LHC physics”, *Reports on Progress in Physics* **70** (2007), no. 1, 89, `arXiv:hep-ph/0611148`. `doi:10.1088/0034-4885/70/1/R02`.

[133] 2011. W.J. Stirling, private communication, and `http://www.hep.phy.cam.ac.uk/~wjs/plots/plots.html`.

[134] R. Wigmans, “Calorimetry, Energy Measurement in Particle Physics”, volume 107 of *International Series of Monographs on Physics*. Oxford University Press, New York, 2000. ISBN 0-19-850296-6.

[135] R. Wigmans, “Calorimetry”, in *AIP Conference Proceedings*, L. Villasenor and V. Villanueva, eds., volume 674, pp. 144–168. AIP, 2003. `doi:10.1063/1.1604077`.

[136] C. W. Fabjan and T. Ludlam, “Calorimetry in High-Energy Physics”, *Annual Review of Nuclear and Particle Science* **32** (1982), no. 1, 335–389. `doi:10.1146/annurev.ns.32.120182.002003`.

[137] T. Gabriel, D. Groom, P. Job et al., “Energy dependence of hadronic activity”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **338** (1994), no. 23, 336 – 347. `doi:10.1016/0168-9002(94)91317-X`.

[138] “LHC machine outreach”. `http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/`.

[139] M. Lamont, "The LHC from Commissioning to Operation", in *Proceedings of the 2nd International Particle Accelerator Conference IPAC11*, C. Petit-Jean-Genaz, ed., pp. 11 – 15. San Sebastián, Spain, 2011. `http://www.jacow.org/`.

[140] CMS Collaboration, "First high-energy collisions at 7 TeV (30th March 2010)". CMS-PHO-EVENTS-2010-008, 2010.

[141] LHC Programme Coordination, "Luminosity Plots for the 2011 Proton Run", 2011. `http://lpc.web.cern.ch/lpc/lumiplots.htm`.

[142] R. Bailey and P. Collier, "Standard Filling Schemes for Various LHC Operation Modes", Technical Report LHC-PROJECT-NOTE-323, CERN, Geneva, 2003.

[143] C. Lefèvre, "The CERN accelerator complex". CERN-DI-0812015, 2008.

[144] CMS Collaboration, "Combination of top pair production cross section measurements", *CMS Physics Analysis Summary* **CMS-PAS-TOP-11-024** (2011).

[145] M. Aliev et al., "– HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR", *Comput. Phys. Commun.* **182** (2011) 1034–1046, `arXiv:1007.1327`. `doi:10.1016/j.cpc.2010.12.040`.

[146] U. Langenfeld, S. Moch, and P. Uwer, "Measuring the running top-quark mass", *Phys.Rev.* **D80** (2009) 054009, `arXiv:0906.5273`. 22 pages, 10 figures, 5 tables. `doi:10.1103/PhysRevD.80.054009`.

[147] LHC Higgs Cross Section Working Group Collaboration, "Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables", `arXiv:1101.0593`. CERN-2011-002.

[148] CMS Collaboration, "CMS technical design report, volume II: Physics performance", *J. Phys.* **G34** (2007) 995–1579. `doi:10.1088/0954-3899/34/6/S01`.

[149] CMS Collaboration, "CMS computing : Technical Design Report", volume 7 of *Technical Design Report CMS*. CERN, Geneva, 2005. CMS-TDR-007. CERN-LHCC-2005-023.

[150] C. Grandi, D. Stickland, L. Taylor et al., "CMS Computing Model: The CMS Computing Model RTAG", Technical Report CMS-NOTE-2004-031. CERN-LHCC-2004-035. LHCC-G-083, CERN, Geneva, 2004.

[151] CMS Collaboration, "The CMS tracker system project", volume 5 of *Technical Design Report CMS*. CERN, Geneva, 1997. CMS-TDR-005. CERN-LHCC-98-006.

[152] CMS Collaboration, "The CMS tracker: addendum to the Technical Design Report", volume 5-add-1 of *Technical Design Report CMS*. CERN, Geneva, 2000. CMS-TDR-005-add-1. CERN-LHCC-2000-016.

[153] "CMS tracking performance results from early LHC operation", *The European Physical Journal C - Particles and Fields* **70** (2010) 1165–1192, `arXiv:hep-ph/1007.1988v2`. `doi:10.1140/epjc/s10052-010-1491-3`.

[154] CMS Collaboration, “Measurement of Tracking Efficiency”, *CMS Physics Analysis Summary* **CMS-PAS-TRK-10-002** (2010).

[155] CMS Collaboration, “Tracking and Primary Vertex Results in First 7 TeV Collisions”, *CMS Physics Analysis Summary* **CMS-PAS-TRK-10-005** (2010).

[156] CMS Collaboration, “Measurement of Momentum Scale and Resolution of the CMS Detector using Low-mass Resonances and Cosmic Ray Muons”, *CMS Physics Analysis Summary* **CMS-PAS-TRK-10-004** (2010).

[157] CMS Collaboration, “The CMS electromagnetic calorimeter project”, volume 4 of *Technical Design Report CMS*. CERN, Geneva, 1997. CMS-TDR-004. CERN-LHCC-97-033.

[158] D. Green, “Calibration of the CMS calorimeters”, technical report, 2001. FERMILAB-FN-0704.

[159] CMS HCAL Collaboration, “Studies of the response of the prototype CMS hadron calorimeter, including magnetic field effects, to pion, electron, and muon beams”, *Nucl.Instrum.Meth.* **A457** (2001) 75–100, `arXiv:hep-ex/0007045`. `doi:10.1016/S0168-9002(00)00711-7`.

[160] CMS ECAL/HCAL Collaboration, “The CMS barrel calorimeter response to particle beams from 2 to 350 GeV/c”, *The European Physical Journal C - Particles and Fields* **60** (2009) 359–373. `doi:10.1140/epjc/s10052-009-0959-5`.

[161] CMS Collaboration, “Electromagnetic calorimeter calibration with 7 TeV data”, *CMS Physics Analysis Summary* **CMS-PAS-EGM-10-003** (2010).

[162] CMS Collaboration, “The CMS hadron calorimeter project”, volume 2 of *Technical Design Report CMS*. CERN, Geneva, 1997. CMS-TDR-002. CERN-LHCC-97-031.

[163] CMS Collaboration, “The CMS muon project”, volume 3 of *Technical Design Report CMS*. CERN, Geneva, 1997. CMS-TDR-003. CERN-LHCC-97-032.

[164] CMS Collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV”, Technical Report CERN-PH-EP-2012-173. CMS-MUO-10-004-003, CERN, Geneva, Jun, 2012. Submitted to JINST.

[165] CMS Collaboration, “CMS TriDAS project: the trigger system”, volume 6 of *Technical Design Report CMS*. CERN, Geneva, 2000. CMS-TDR-006-1. CERN-LHCC-2000-038.

[166] B. Dahmes, “Triggers for LHC Physics”, 2011. Lecture at CERN Summer Student Programme.

[167] I. P. P. O. Group. http://ippog.web.cern.ch/.

[168] CMS Collaboration, “Photon reconstruction and identification at $\sqrt{s} = 7$ TeV”, *CMS Physics Analysis Summary* **CMS-PAS-EGM-10-005** (2010).

[169] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and E_T^{miss}", *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-001** (2009).

[170] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from *pp* Collisions at 7 TeV", *CMS Physics Analysis Summary* **CMS-PAS-PFT-10-002** (2010).

[171] CMS Collaboration, "Shape, transverse size, and charged hadron multiplicity of jets in *pp* collisions at 7 TeV", *CMS Physics Analysis Summary* **CMS-PAS-QCD-10-029** (2012). Submitted to JHEP.

[172] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *Journal of Instrumentation* **6** (2011) P11002, `arXiv:1107.4277`. `doi:10.1088/1748-0221/6/11/P11002`.

[173] S. Brandt, C. Peyrou, R. Sosnowski et al., "The principal axis of jets an attempt to analyse high-energy collisions as two-body processes", *Physics Letters* **12** (1964), no. 1, 57 – 61. `doi:10.1016/0031-9163(64)91176-X`.

[174] CMS Collaboration Collaboration, "First Measurement of Hadronic Event Shapes in *pp* Collisions at $\sqrt{s} = 7$ TeV", *Phys.Lett.* **B699** (2011) 48–67, `arXiv:1102.0068`. `doi:10.1016/j.physletb.2011.03.060`.

[175] S. Drell and T.-M. Yan, "Partons and their Applications at High-Energies", *Annals Phys.* **66** (1971) 578. `doi:10.1016/0003-4916(71)90071-6`.

[176] J. C. Collins and D. E. Soper, "The Theorems of Perturbative QCD", *Ann.Rev.Nucl.Part.Sci.* **37** (1987) 383–409. `doi:10.1146/annurev.ns.37.120187.002123`.

[177] M. Dobbs and J. Hansen, "HepMC 2 — a C++ Event Record for Monte Carlo Generators: User Manual Version 2.06", *Comput. Phys. Commun.* **134** (2001) 41. `http://lcgapp.cern.ch/project/simu/HepMC`.

[178] G. F. Sterman, "QCD and jets", `arXiv:hep-ph/0412013`. 2004.

[179] V. D. Barger and R. J. N. Phillips, "Collider Physics". Frontiers in Physics. Addison-Wesley, 1987. ISBN 0-201-05876-6.

[180] V. N. Gribov and L. N. Lipatov, "Deep inelastic *ep* scattering in perturbation theory", *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

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

[182] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics", *Sov. Phys. JETP* **46** (1977) 641–653.

[183] D. Brömmel, M. Göckeler, P. Hägler et al., "Hadronic structure from the lattice", *The European Physical Journal - Special Topics* **162** (2008) 63–71. `doi:10.1140/epjst/e2008-00777-8`.

[184] H1 and ZEUS Collaboration, “Combined Measurement and QCD Analysis of the Inclusive $e^{\pm}p$ Scattering Cross Sections at HERA”, *JHEP* **1001** (2010) 109, `arXiv:0911.0884`. `doi:10.1007/JHEP01(2010)109`.

[185] CTEQ Collaboration, “The Coordinated Theoretical-Experimental Project on QCD”, 2011. `http://cteq.org`.

[186] J. Pumplin, D. Stump, J. Huston et al., “New generation of parton distributions with uncertainties from global QCD analysis”, *JHEP* **0207** (2002) 012, `arXiv:hep-ph/0201195`. `doi:10.1088/1126-6708/2002/07/012`.

[187] CMS Collaboration, “Measurement of Differential Jet Cross Sections at $\sqrt{s} = 7\,\mathrm{TeV}$ with the CMS Detector”, *CMS Physics Analysis Summary* **CMS-PAS-QCD-11-004** (2012).

[188] J. Pumplin, “Quark - gluon jet differences at LEP”, *Phys.Rev.* **D48** (1993) 1112–1116, `arXiv:hep-ph/9301215`. `doi:10.1103/PhysRevD.48.1112`.

[189] CDF Collaboration, “Study of jet shapes in inclusive jet production in $p\overline{p}$ collisions at $\sqrt{s} = 1.96\,\mathrm{TeV}$”, *Phys. Rev. D* **71** (2005) 112002, `arXiv:hep-ex/0505013`. `doi:10.1103/PhysRevD.71.112002`.

[190] CMS Collaboration, “Measurement of the ratio of the 3-jet to 2-jet cross sections in pp collisions at $\sqrt{2} = 7\,\mathrm{TeV}$”, *Physics Letters B* **702** (2011), no. 5, 336 – 354, `arXiv:1106.0647`. `doi:10.1016/j.physletb.2011.07.067`.

[191] T. Sjöstrand and M. van Zijl, “A multiple-interaction model for the event structure in hadron collisions”, *Phys. Rev. D* **36** (1987) 2019 – 2041. `doi:10.1103/PhysRevD.36.2019`.

[192] CMS Collaboration, “Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7\,\mathrm{TeV}$ and Comparison with $\sqrt{s} = 0.9\,\mathrm{TeV}$”, *JHEP* **1109** (2011) 109, `arXiv:1107.0330`. `doi:10.1007/JHEP09(2011)109`.

[193] J. E. Huth, N. Wainer, K. Meier et al., “Toward a standardization of jet definitions”, in *1990 DPF Summer Study on High-energy Physics, Research Directions For The Decade, Snowmass, Colorado*, E. L. Berger, ed., number FERMILAB-CONF-90-249-E, pp. 134–136. World Scientific, 1992.

[194] G. P. Salam, “Towards Jetography”, *Eur.Phys.J.* **C67** (2010) 637–686, `arXiv:0906.1833`. `doi:10.1140/epjc/s10052-010-1314-6`.

[195] G. F. Sterman and S. Weinberg, “Jets from Quantum Chromodynamics”, *Phys.Rev.Lett.* **39** (1977) 1436. `doi:10.1103/PhysRevLett.39.1436`.

[196] G. P. Salam and G. Soyez, “A Practical Seedless Infrared-Safe Cone jet algorithm”, *JHEP* **0705** (2007) 086, `arXiv:0704.0292`. `doi:10.1088/1126-6708/2007/05/086`.

[197] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions”, *Phys. Rev. D* **48** (Oct, 1993) 3160 – 3166. `doi:10.1103/PhysRevD.48.3160`.

[198] Y. L. Dokshitzer, G. Leder, S. Moretti et al., “Better jet clustering algorithms”, *JHEP* **9708** (1997) 001, `arXiv:hep-ph/9707323`. `doi:10.1088/1126-6708/1997/08/001`.

[199] M. Wobisch and T. Wengler, “Hadronization corrections to jet cross-sections in deep inelastic scattering”, `arXiv:hep-ph/9907280`. DESY-THESIS-2000-049.

[200] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-k_T jet clustering algorithm”, *Journal of High Energy Physics* **2008** (2008), no. 04, 063, `arXiv:0802.1189`. `doi:10.1088/1126-6708/2008/04/063`.

[201] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual”, `arXiv:1111.6097`.

[202] M. Cacciari and G. P. Salam, “Dispelling the N^3 myth for the k_t jet-finder”, *Phys.Lett.* **B641** (2006) 57–61, `arXiv:hep-ph/0512210`. `doi:10.1016/j.physletb.2006.08.037`.

[203] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, “Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches”, *Phys.Rev.* **D81** (2010) 094023, `arXiv:0912.0033`. `doi:10.1103/PhysRevD.81.094023`.

[204] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, “Techniques for improved heavy particle searches with jet substructure”, *Phys.Rev.* **D80** (2009) 051501, `arXiv:0903.5081`. `doi:10.1103/PhysRevD.80.051501`.

[205] CMS Collaboration, “Search for BSM $t\bar{t}$ Production in the Boosted All-Hadronic Final State”, *CMS Physics Analysis Summary* **CMS-PAS-EXO-11-006** (2011).

[206] D. Rathjens, “Particle Reconstruction at the LHC using Jet Substructure Algorithms”, 2011. Diploma thesis CMS TS-2012-020, Universität Hamburg, Hamburg, Germany.

[207] G. Corcella, I. Knowles, G. Marchesini et al., “HERWIG 6.5 release note”, `arXiv:hep-ph/0210213`.

[208] “Commissioning of TrackJets in pp Collisions at 7 TeV”, *CMS Physics Analysis Summary* **CMS-PAS-JME-10-006** (2010).

[209] CMS Collaboration, “The Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS”, *CMS Physics Analysis Summary* **CMS-PAS-JME-09-002** (2009).

[210] N. Metropolis and S. Ulam, “The Monte Carlo Method”, *Journal of the American Statistical Association* **44** (1949), no. 247, 335–341. `doi:10.2307/2280232`.

[211] M. Dobbs, S. Frixione, E. Laenen et al., “Les Houches guidebook to Monte Carlo generators for hadron collider physics”, `arXiv:hep-ph/0403045`. Compiled by the Working Group on Quantum Chromodynamics and the Standard Model.

[212] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual”, *JHEP* **05** (2006) 026. `doi:10.1088/1126-6708/2006/05/026`.

[213] J. Alwall et al., “MadGraph/MadEvent v4: The New Web Generation”, *JHEP* **09** (2007) 028. `doi:10.1088/1126-6708/2007/09/028`.

[214] S. Agostinelli, J. Allison, K. Amako et al., “G4 – a simulation toolkit”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003), no. 3, 250 – 303. `doi:10.1016/S0168-9002(03)01368-8`.

[215] CMS Collaboration, “Fast Simulation Offline Guide”, *Public CMS Wiki* (2011). https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideFastSimulation, Topic revision: r409.

[216] CMS Collaboration, “Pileup Reweighting”, *Internal CMS Wiki* (2011). https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupReweighting.

[217] CMS Collaboration, “Jet Performance in *pp* Collisions at 7 TeV”, *CMS Physics Analysis Summary* **CMS-PAS-JME-10-003** (2010).

[218] CMS Collaboration, “Charged particle transverse momentum spectra in *pp* collisions at $\sqrt{s} = 0.9$ and 7 TeV”, *JHEP* **1108** (2011) 086, `arXiv:1104.3547`. `doi:10.1007/JHEP08(2011)086`.

[219] R. Glückstern, “Uncertainties in track momentum and direction, due to multiple scattering and measurement errors”, *Nuclear Instruments and Methods* **24** (1963), no. 0, 381 – 389. `doi:10.1016/0029-554X(63)90347-1`.

[220] CMS Collaboration, “Jet Flavour identification (MC truth)”, *Public CMS Wiki* (2011). https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideBTagMCTools, Topic revision: r17.

[221] M. Thomson, “Particle flow calorimetry and the PandoraPFA algorithm”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **611** (2009), no. 1, 25 – 40, `arXiv:0907.3577`. `doi:10.1016/j.nima.2009.09.009`.

[222] U. Gebbert, “Photons, Missing Energy, and the Quest for Supersymmetry at the LHC”. PhD thesis, Universität Hamburg, Hamburg, Germany, 2012.

[223] V. Blobel and E. Lohrmann, “Statistische und numerische Methoden der Datenanalyse”. Teubner Studienbücher Physik. Teubner, Stuttgart, Leipzig, 1998. ISBN 3-519-03243-0.

[224] R. J. Barlow, “Statistics — A Guide to the Use of Statistical Methods in the Physical Sciences”. The Manchester physics series. John Wiley & Sons, 1989. ISBN 0-471-92294-3.

[225] CMS Collaboration, “Jet energy calibration with photon+jet events”, *CMS Physics Analysis Summary* **CMS-PAS-JME-09-004** (2009).

[226] CMS Collaboration, “Calibration of the absolute jet energy scale with $Z \to \mu^+\mu^-$+jet events at CMS”, *CMS Physics Analysis Summary* **CMS-PAS-JME-09-009** (2009).

[227] CMS Collaboration, “Measurement of the Isolated Prompt Photon Production Cross Section in pp Collisions at $\sqrt{s} = 7\,\mathrm{TeV}$”, *Phys.Rev.Lett.* **106** (2011) 082001, `arXiv:1012.0799`. `doi:10.1103/PhysRevLett.106.082001`.

[228] CMS Collaboration, “Determination of the Jet Energy Scale in CMS with pp Collisions at $\sqrt{s} = 7\,\mathrm{TeV}$”, *CMS Physics Analysis Summary* **CMS-PAS-JME-10-010** (2010).

[229] CMS Collaboration, “Plans for Jet Energy Corrections at CMS”, *CMS Physics Analysis Summary* **CMS-PAS-JME-07-002** (2008).

[230] ATLAS Collaboration, “Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy $\sqrt{s} = 7\,\mathrm{TeV}$”, *ATLAS Note* **ATLAS-CONF-2010-054** (2010).

[231] ATLAS Collaboration, “Jet/MissingEt preliminary results on jet resolution from 2011 data”, *ATLAS Wiki* (2011). https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2011JetResolution, Topic revision: r4.

[232] DØ Collaboration, “High-p_{T} jets in $\bar{p}p$ collisions at $\sqrt{s} = 630$ and $1800\,\mathrm{GeV}$”, *Phys. Rev. D* **64** (2001) 032003, `arXiv:hep-ex/0012046`. `doi:10.1103/PhysRevD.64.032003`.

[233] DØ Collaboration, “Measurement of the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\mathrm{TeV}$”, *Phys.Rev.* **D85** (2012) 052006, `arXiv:1110.3771`.

[234] H. Held, “Measurement of the Jet Momentum Resolution and Search for a light Standard Model Higgs Boson in the $H(b\bar{b})W(lv)$ Channel with the CMS Detector at the LHC”. PhD thesis, Karlsruher Instituts für Technologie (KIT), Karlsruhe, Germany, 2012. urn:nbn:de:swb:90-281963.

[235] CMS Collaboration, “Measurement of Top Quark Pair Differential Cross Sections at $\sqrt{s} = 7\,\mathrm{TeV}$”, *CMS Physics Analysis Summary* **CMS-PAS-TOP-11-013** (2011).

[236] CMS Collaboration, “Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7\,\mathrm{TeV}$”, *Phys. Rev. Lett.* **107** (2011) 132001, `arXiv:1106.0208`. `doi:10.1103/PhysRevLett.107.132001`.

[237] CMS Collaboration, “Search for New Physics in the Multijet and Missing Transverse Momentum Final State in Proton-Proton Collisions at $\sqrt{s} = 7\,\mathrm{TeV}$”, *Phys. Rev. Lett.* **109** (2012) 171803, `arXiv:1207.1898`. `doi:10.1103/PhysRevLett.109.171803`.

[238] CMS Collaboration, “Search for supersymmetry in all-hadronic events with missing energy”, *CMS Physics Analysis Summary* **CMS-PAS-SUS-11-004** (2011).

[239] CMS Collaboration, “Jet Energy Resolution in CMS at $\sqrt{s} = 7\,\mathrm{TeV}$”, *CMS Physics Analysis Summary* **CMS-PAS-JME-10-014** (2011).

[240] CMS Collaboration, “Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update”, *CMS Physics Analysis Summary* **CMS-PAS-SMP-12-008** (2012).

[241] CMS Collaboration, “Performance of muon identification in pp collisions at $\sqrt{s} = 7\,\text{TeV}$”, *CMS Physics Analysis Summary* **CMS-PAS-MUO-10-002** (2010).

[242] CMS Collaboration, “Isolated Photon Reconstruction and Identification at $\sqrt{s} = 7\,\text{TeV}$”, *CMS Physics Analysis Summary* **CMS-PAS-EGM-10-006** (2011).

[243] CMS Collaboration, “HCAL performance from first collisions data”, *CMS Detector Performance Summary* **2010-025** (2010).

[244] CMS Collaboration, “Missing transverse energy performance of the CMS detector”, *Journal of Instrumentation* **6** (2011), no. 09, P09001, `arXiv:1106.5048`. `doi:10.1088/1748-0221/6/09/P09001`.

[245] F. Beaudette, R. Bellan, D. Benedetti et al., “Fake missing transverse momentum in the RA2 analysis”, *CMS Analysis Note* **CMS-AN-10-383** (2010). Internal documentation.

[246] S. A. Koay, “A Search for Dark Matter Production with Jets and Missing Momentum Signature in Proton-Proton Collisions at 7 TeV”. PhD thesis, University of California, Santa Barbara, 2011. CMS TS-11-024.

[247] CMS Collaboration, “Calorimeter Jet Quality Criteria for the First CMS Collision Data”, *CMS Physics Analysis Summary* **CMS-PAS-JME-09-008** (2008).

[248] V. Blobel, “Large-scale optimization”. Web site with the programme code `http://www.desy.de/~blobel/largesc.html`.

[249] V. Blobel, “Large-scale optimization with the limited-memory BFGS program LVMINI”, 2007. `http://www.desy.de/~blobel/lvmini.pdf`.

[250] D. C. Liu and J. Nocedal, “On the limited memory BFGS method for large scale optimization”, *Math. Program.* **45** (December, 1989) 503–528. `doi:10.1007/BF01589116`.

[251] J. Nocedal and S. Wright, “Numerical Optimization”. Springer Series in Operations Research. Springer, New York, 2000. ISBN 0-387-98793-2.

[252] F. James, “Function Minimization”, in *Proceedings of the CERN Computing and Data Processing School, 1972, Pertisau, Austria*, pp. 1–52. CERN, Geneva, 1972.

[253] F. James, “MINUIT — Function Minimization and Error Analysis — Reference Manual”. CERN, Geneva, 1994. CERN Program Library Long Writeup D506.

[254] F. James and M. Winkler, “MINUIT User’s Guide (C++ version)”. CERN, Geneva, 2004. `http://seal.cern.ch/documents/minuit/mnusersguide.pdf`.

[255] W. H. Press, S. A. Teukolsky, W. T. Vetterling et al., “Numerical Recipes: The Art of Scientific Computing”. Cambridge University Press, 3. edition, 2007. ISBN 0-52-188068-8.

[256] R. Barlow, "Asymmetric Systematic Errors", `arXiv:0306138`. 2003.

[257] CMS Collaboration, "Search for new physics in the multijets + missing transverse energy final state in 7 TeV proton-proton collisions", *Public CMS Wiki* (2012). https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS12011, Topic revision: r2.

[258] 2012. C. Sander, private communication.

[259] CMS Collaboration, "CMS Supersymmetry Physics Results", *Public CMS Wiki* (2012). https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS, Topic revision: r91.

[260] CMS Collaboration, "Search for new physics with jets and missing transverse momentum in pp collisions at $\sqrt{s} = 7\,\mathrm{TeV}$", *Journal of High Energy Physics* **2011** (2011) 1–46, `arXiv:1106.4503`. `doi:10.1007/JHEP08(2011)155`.

[261] CMS Collaboration, "Search for Supersymmetry in hadronic final states using M_{T2} with the CMS detector at $\sqrt{s} = 7\,\mathrm{TeV}$", *CMS Physics Analysis Summary* **CMS-PAS-SUS-12-002** (2012). Submitted to PRD.

[262] CMS Collaboration, "Search for supersymmetry with the razor variables at CMS", *CMS Physics Analysis Summary* **CMS-PAS-SUS-12-005** (2012).

[263] CMS Collaboration, "Search for Supersymmetry at the LHC in Events with Jets and Missing Transverse Energy", *Phys. Rev. Lett.* **107** (2011) 221804, `arXiv:1109.2352`. `doi:10.1103/PhysRevLett.107.221804`.

[264] CMS Collaboration, "Search for new physics in events with b-quark jets and missing transverse energy in proton-proton collisions at 7 TeV", *CMS Physics Analysis Summary* **CMS-PAS-SUS-11-006** (2011).

[265] CMS Collaboration, "Search for Supersymmetry in Events with b Jets and Missing Transverse Momentum at the LHC", *JHEP* **1107** (2011) 113, `arXiv:1106.3272`. `doi:10.1007/JHEP07(2011)113`.

[266] L. Randall and D. Tucker-Smith, "Dijet Searches for Supersymmetry at the Large Hadron Collider", *Phys. Rev. Lett.* **101** (Nov, 2008) 221803. `doi:10.1103/PhysRevLett.101.221803`.

[267] C. G. Lester and D. J. Summers, "Measuring masses of semi-invisibly decaying particles pair produced at hadron colliders", *Phys. Lett.* **B 463** (1999) 99–103, `arXiv:hep-ph/9906349`. `doi:10.1016/S0370-2693(99)00945-4`.

[268] A. Barr, C. Lester, and P. Stephens, "m_{T2}: the Truth behind the glamour", *J. Phys. G* **29** (2003) 2343–2363, `arXiv:hep-ph/0304226`. `doi:10.1088/0954-3899/29/10/304`.

[269] C. Rogan, "Kinematical variables towards new dynamics at the LHC", `arXiv:1006.2727`.

[270] ATLAS Collaboration, “Search for squarks and gluinos with the ATLAS detector using final states with jets and missing transverse momentum and 4.7 fb^{-1} of $\sqrt{s} = 7$ TeV proton-proton data”, Technical Report ATLAS-CONF-2012-033, CERN, Geneva, 2012.

[271] ATLAS Collaboration, “Hunt for new phenomena using large jet multiplicities and missing transverse momentum with ATLAS in L = 4.7 fb^{-1} of $\sqrt{s} = 7$ TeV proton-proton collisions”, Technical Report ATLAS-CONF-2012-037, CERN, Geneva, 2012.

[272] ATLAS Collaboration, “Search for squarks and gluinos using final states with jets and missing transverse momentum with the ATLAS detector in $\sqrt{s} = 7$ TeV proton-proton collisions”, *Phys.Lett.* **B701** (2011) 186–203, `arXiv:1102.5290`. `doi:10.1016/j.physletb.2011.05.061`.

[273] CMS Collaboration, “Electron reconstruction and identification at $\sqrt{s} = 7$ TeV”, *CMS Physics Analysis Summary* **CMS-PAS-EGM-10-004** (2010).

[274] C. Autermann, A. Bhatti, R. Cavanaugh et al., “Inclusive search for new physics at CMS with the jets and missing momentum signature”, *CMS Analysis Note* **CMS-AN-11-398** (2012). Internal documentation.

[275] W. Beenakker, R. Hopker, and M. Spira, “PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD”, `arXiv:hep-ph/9611232`.

[276] CMS Collaboration, “Measurement of ttbar Pair Production Cross Section at $\sqrt{s} = 7$ TeV using b-quark Jet Identification Techniques in Lepton + Jet Events”, *CMS Physics Analysis Summary* **CMS-PAS-TOP-11-003** (2011).

[277] F. Nowak, “Search for Supersymmetry with Jets, Missing Transverse Momentum, and a Single Tau at CMS”. PhD thesis, Universität Hamburg, Hamburg, Germany, 2012. DESY-THESIS-2012-018.

[278] P. Aurenche, J. P. Guillet, E. Pilon et al., “A new critical study of photon production in hadronic collisions”, *Phys. Rev. D* **73** (2006) 094007. `doi:10.1103/PhysRevD.73.094007`.

[279] CMS Collaboration, “Measurement of the Isolated Prompt Photon Production Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV”, *Phys. Rev. Lett.* **106** (2011) 082001. `doi:10.1103/PhysRevLett.106.082001`.

[280] CMS Collaboration, “Measurement of the differential cross section for isolated prompt photon production in pp collisions at $\sqrt{s} = 7$ TeV”, *Phys. Rev. D* **84** (2011) 052011. `doi:10.1103/PhysRevD.84.052011`.

[281] CMS Collaboration, “Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV”, *JHEP* **1110** (2011) 132, `arXiv:1107.4789`. `doi:10.1007/JHEP10(2011)132`.

[282] J. Thomsen, “Search for Supersymmetry in the Hadronic Channel with Data from the CMS Experiment”. PhD thesis, Universität Hamburg, Hamburg, Germany, 2012. DESY-THESIS-2012-013.

[283] B. Efron, “The Jackknife, the Bootstrap, and Other Resampling Plans”. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial Mathematics, 1987. ISBN 0-898-71179-7.

[284] T. Schum, “QCD Processes and Search for Supersymmetry at the LHC”. PhD thesis, Universität Hamburg, Hamburg, Germany, 2012.

[285] J. D’Hondt, S. Lowette, O. Buchmüller et al., “Fitting of Event Topologies with External Kinematic Constraints in CMS”, *CMS Analysis Note* **CMS-AN-05-025** (2005). Internal documentation.

[286] A. L. Read, “Modified frequentist analysis of search results (the CL_s method)”, in *1st Workshop on Confidence Limits*, L. Lyons, Y. Perrin, and F. E. James, eds. CERN, Geneva, 2000. CERN-OPEN-2000-205. oai:cds.cern.ch:451614.

[287] A. L. Read, “Presentation of search results: the CL_s technique”, *Journal of Physics G: Nuclear and Particle Physics* **28** (2002), no. 10, 2693 – 2704. `doi:10.1088/0954-3899/28/10/313`.

[288] T. Junk, “Confidence level computation for combining searches with small statistics”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **434** (1999), no. 2-3, 435 – 443. `doi:10.1016/S0168-9002(99)00498-2`.

[289] ATLAS Collaboration, CMS Collaboration, LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in Summer 2011”, Technical Report ATL-PHYS-PUB-2011-011. CMS-NOTE-2011-005, CERN, Geneva, 2011.

[290] F. E. Paige, S. D. Protopopescu, H. Baer et al., “ISAJET 7.69: A Monte Carlo event generator for pp, anti-p p, and e+e- reactions”, `arXiv:hep-ph/0312045`. `http://www.nhn.ou.edu/~isajet/`.

[291] The PDF4LHC Working Group, “Recommendation for LHC cross section calculations”, 2011. `http://www.hep.ucl.ac.uk/pdf4lhc/index.html`.

[292] ATLAS Collaboration, “ATLAS Supersymmetry (SUSY) searches”, *ATLAS Wiki* (2012). https://twiki.cern.ch/twiki/bin/view/AtlasPublic/SupersymmetryPublicResults, Topic revision: r204.

[293] N. Arkani-Hamed, P. Schuster, N. Toro et al., “MARMOSET: The Path from LHC Data to the New Standard Model via On-Shell Effective Theories”, `arXiv:hep-ph/0703088`.

[294] J. Alwall, P. Schuster, and N. Toro, “Simplified Models for a First Characterization of New Physics at the LHC”, *Phys.Rev.* **D79** (2009) 075020, `arXiv:0810.3921`. `doi:10.1103/PhysRevD.79.075020`.

[295] J. Alwall, M.-P. Le, M. Lisanti et al., “Model-Independent Jets plus Missing Energy Searches”, *Phys.Rev.* **D79** (2009) 015005, `arXiv:0809.3264`. `doi:10.1103/PhysRevD.79.015005`.

[296] CMS Collaboration, “Interpretation of Searches for Supersymmetry”, *CMS Physics Analysis Summary* **CMS-PAS-SUS-11-016** (2012).

[297] O. Buchmueller, R. Cavanaugh, A. De Roeck et al., “Supersymmetry in Light of 1/fb of LHC Data”, `arXiv:1110.3568`.

Abschließend möchte ich mich bei all denjenigen bedanken, die auf direkte und indirekte Weise an der Entstehung dieser Arbeit beteiligt waren und dazu beigetragen haben, dass meine Doktorandenjahre eine großartige und prägende Zeit waren.

Zuerst gilt mein Dank Prof. Dr. Peter Schleper für die Möglichkeit, in seiner Gruppe promovieren zu können, für seine klugen und hilfreichen Kommentare zu meiner Arbeit sowie die Begutachtung dieser Dissertation und nicht zuletzt für seinen mitreißenden Enthusiasmus und seine hervorragenden Vorlesungen, die wesentlich zu meinem Interesse an Teilchenphysik beigetragen haben. Ich habe eine sehr lehrreiche und inspirierende Doktorandenzeit verbracht!

Des Weiteren gilt mein Dank Dr. Isabell-Alissandra Melzer-Pellmann und Prof. Dr. Volker Büscher für das Erstellen des Zweit- bzw. Drittgutachtens der Dissertation sowie Prof. Dr. Robert Klanner und Prof. Dr. Christian Sander für die Begutachtung der Disputation. Prof. Dr. Christian Sander möchte ich darüber hinaus auch für die lehrreichen, nicht immer ganz stressfreien und erfolgreichen “RA2-Jahre” in Hamburg und Genf danken.

Ein ganz besonderer Dank gilt meinem Betreuer Dr. Hartmut Stadie, der bei Fragen stets die richtige Lösung parat hatte, und von dem ich sehr viel über Physik, Statistik und Programmierung aber auch die Arbeit in einer großen Kollaboration gelernt habe und dessen Scharfsinn manchmal frustrierend war... Danke!

Darüber hinaus möchte ich mich bei den zahlreichen Kollegen, die ich in den vergangenen Jahren in Hamburg, Genf und anderweitig kennengelernt habe, bedanken für ihre Hilfsbereitschaft, die angenehme und inspirierende Arbeitsatmosphäre und nicht zuletzt auch die gemeinsam verbrachte Freizeit. Stellvertretend für diese möchte ich hervorheben: Dr. Holger Enderle, Jörn Lange, Dr. Friederike Nowak und Niklas Pietsch für die großartige Atmosphäre in unserem Büro, die vielen anregenden Dispute über Physik und anderes Relevante sowie die nie versiegenden Süßigkeitenquellen; Matthias Stein für seine geduldige Unterweisung in allen erdenklichen Sportarten; Dr. Ulla Gebbert für einen nie endenden Squashwettkampf; Dr. Thomas Hermanns und Dr. Gordon Kaußen für das gemeinsame Erklimmen sämtlicher Berge rund um Genf; sowie Matthias Edelhoff, Dr. Niklas Mohr und Daniel Sprenger für die “Aachen-Hamburg Coordination” am CERN.

Vor allem aber danke ich meinen Freunden und meiner Familie, insbesondere Stefan, Bente, Bettina, Marion und Horst, für ihre Freundschaft und bedingungslose Unterstützung, auf die ich mich immer verlassen konnte.

Danke!