Klaus Rabbertz

Jet Physics at the LHC

The Strong Force beyond the TeV Scale

Klaus Rabbertz
Institut für Experimentelle Kernphysik
Karlsruhe Institute of Technology (KIT)
Wolfgang-Gaede-Str. 1,
D-76131 Karlsruhe, Germany
e-mail: klaus.rabbertz@kit.edu

To my French and German Family

Preface

The first physics conference I ever attended is the Workshop "QCD — 20 Years Later" at the RWTH Aachen in June 1992. It took place between seminars I had to follow for my studies and the work for my Diploma thesis. In both activities I successfully avoided topics too complicated to understand because of their close relation to the theory of the strong interaction i.e. quantum chromodynamics. Nevertheless, I found the concept of jets, presented by Bryan Webber at this workshop, quite intriguing and also the summary report on the strong coupling constant α_S, given by Guido Altarelli, seemed to be rather interesting. Two years later, encouraged by my thesis advisor Christoph Berger, I found myself venturing on a subject of pure QCD for my Ph.D. thesis: the determination of the strong coupling constant from jet measurements in deep-inelastic scattering with the H1 experiment at the HERA collider in Hamburg. However, I unfortunately could demonstrate that the theoretical calculation available to me at that time had severe shortcomings preventing a timely success of this initial plan A. As a consequence and instigated by another talk from Bryan Webber, I dug even deeper into QCD for Plan B and analysed, this time successfully, the interplay of perturbative and nonperturbative aspects in the context of event shapes and power corrections.

Almost 20 years and two experiments later, my fascination for QCD and its successes has not cessed to increase. Thanks to the CMS experiment and numerous students confiding in me for their thesis work I was able to contribute to the more recent successes of QCD at the LHC. Meanwhile, I also had the great pleasure to lecture on experimental jet physics just after Bryan Webber in the Graduiertenkolleg of KCETA in 2014. Last autumn, I was also selected to present the results of CMS at a workshop dedicated to the strong coupling constant at future colliders — just before the distinguished summary speaker Guido Altarelli. Unfortunately, he passed away just some days before. The proceedings of this workshop are dedicated to his memory. I would have appreciated very much to have his expert opinion on the latest developments and somehow "close the loop".

To share and conserve the knowledge gained during LHC Run 1, I find it timely to summarise my research and to write up the lectures that I have given since 2012 in the form of this book. I hope it proves to be a useful resource on the topic of QCD and jet physics for students and colleagues alike.

Knowing that it is practically impossible to name all who contributed in one or another way to this work, I would like to express my sincere thanks to everybody enabling me to pursuit and accomplish my research goals. I offer my excuses to all those I can not mention explicitly. First of all, I thank my colleagues in CMS for the opportunity to analyse the data of this great experiment. In particular, my thanks go to my long-term collaborators Günther Dissertori, Mikko Voutilainen, Nikos Varelas, Len Apanasevich, Andreas Hinzmann, Panos Kokkas, Manjit Kaur, Katerina Lipka, Ringaile Plačakyte, and Maxime Gouzevitch. Of course, there are also many colleagues in other experiments and projects or from theory, to which I am indebted, and I would like to mention explicitly my friends from FASTNLO Markus Wobisch and Daniel Britzger, my HERAFITTER/XFITTER collaborators, and Sergey Alekhin, Johannes Blümlein, Sven-Olaf Moch, Marco Guzzi, Stefan Dittmaier, Alexander Huss, Peter Uwer, Tancredi Carli, Steffen Schumann, Enrico Bothmann, Simon Plätzer, Juan Rojo, Pavel Nadolsky, and Gavin Salam.

At KIT, I most sincerely thank Günter Quast for, well, everything, and Thomas Müller for his wholehearted and continuous support. I am also very grateful to my colleagues at KIT for all their support and for the possibility to habilitate. I particularly thank my co-lecturer Stefan Gieseke and all the students that have worked or are working with me for their studies at KIT or elsewhere: Andreas Oehler, Oliver Oberst, Michael Heinrich, Volker Büge, Benjamin Klein, Fred Stober, Danilo Piparo, Georg Sieber, Dominik Haitz, Corinna Günth, Frank Polgart, Daniel Săvoiu, Anna Friedel, Pooja Gupta, Manuk Zubin Mehta, Giannis Flouris, and Anterpreet Kaur. Financial support by the Helmholtz Alliance "Physics at the Terascale" and the German Ministry of Education and Research is gratefully acknowledged.

Special thanks go to Dominik Haitz, Georg Sieber, Daniel Săvoiu, and Mikko Voutilainen for proofreading, and to Dominik Haitz, Stefan Gieseke, Danilo Piparo, Rick Field, Joram Berger, Markus Wobisch, Francesco Pandolfi, Fred Stober, Anterpreet Kaur, Christoph Berger, Eckhard Elsen, and Christophe Grojean for their permission to reuse their sketches and figures. Finally, I express very special thanks to Frau Ute Heuser and Springer for their patience and confidence in me despite of missed out deadlines.

Last but not least, my final and most cordial thanks go to Mon Amour, Corinne Béroujon, for all her support and patience.

Collonges-sous-Salève, June 2016

Contents

Chapter 1
Introduction

The restart of the Large Hadron Collider (LHC) at CERN, Geneva, in 2015 at a new record centre-of-mass energy of 13 TeV is a good occasion to collect, discuss, and summarise the experience gained with jet analyses in Run 1. The measurements will become more complex with larger numbers of concurrent proton-proton collisions and with more jets at even higher momenta. New and conceptually refined but also more complicated tools from theory await their application. To meet these challenges a solid understanding of the current state of the art in jet studies at the LHC is crucial. This work aims to convey the required knowledge in experimental techniques, statistical procedures, and theoretical tools in an integrative approach by discussing exemplary analyses in three chapters touching on a multitude of QCD aspects. Building on this experience the conception of new observables and the development of sophisticated, novel measurement tactics for LHC Run 2 should be facilitated.

To demonstrate the achievements from Run 1 as detailed in the following chapters, Figure 1.1 presents for later reference the pre-LHC knowledge on the running of the strong coupling constant as collected in Ref. [1].

This book is organised as follows: Chapters 2 and 3 introduce the theoretical concepts of QCD and the basic experimental technologies, respectively, to perform jet physics at the LHC. Measurements of absolute cross sections, which are subject to all sources of experimental and theoretical uncertainties, are described in Chapter 4. The two strategies to reduce uncertainties by either taking ratios of cross sections or by performing shape comparisons of normalised distributions are each exemplified by a selection of jet analyses in chapters 5 and 6. Perspectives for such measurements at the LHC Run 2 and beyond are presented in Chapter 7. Chapter 8 concludes the book with a summary on jet physics and prospects on the strong coupling constant.

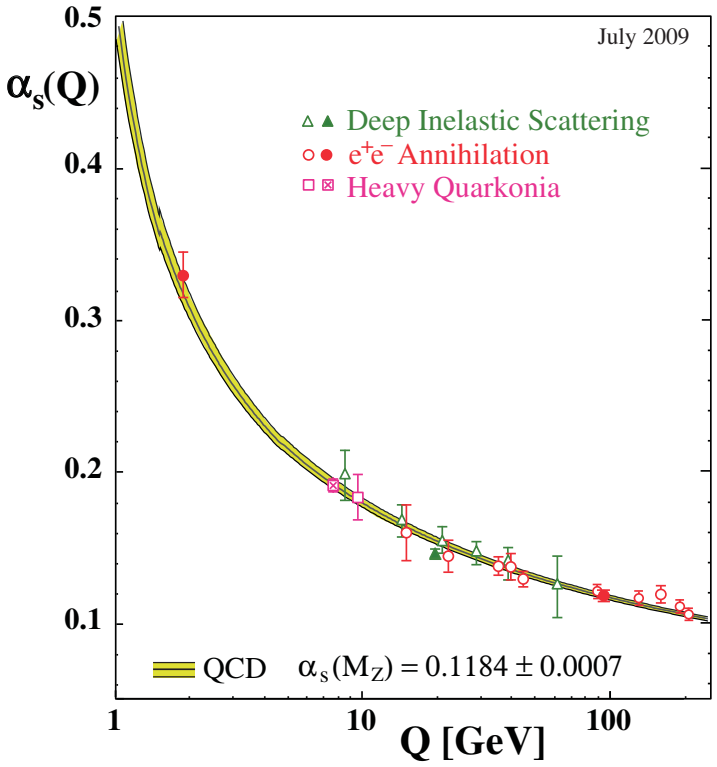

Fig. 1.1 Running of the strong coupling constant as of 2009. Determinations of the strong coupling constant α_S are shown as a function of the relevant energy scale Q of the respective process. The range in Q extends up to 208 GeV from e^+e^- annihilation data taken at the LEP collider. *(Taken from Refs. [1, 2])*

References

1. K. Nakamura and others (Particle Data Group), "Review of Particle Physics", *J. Phys. G* **37** (2010) 075021, doi:10.1088/0954-3899/37/7A/075021.
2. S. Bethke, "The 2009 World Average of $\alpha_s(M_Z)$", *Eur. Phys. J. C* **64** (2009) 689, doi:10.1140/epjc/s10052-009-1173-1, arXiv:0908.1135.

Chapter 2
Theoretical Framework

Jet physics, in particular at a hadron collider such as the LHC, cannot be understood without being thoroughly familiar with the theory of the strong interaction: quantum chromodynamics or short QCD. The material presented in this chapter is intended to provide the required proficiency to comprehend experimental and phenomenological publications on the subject of jet physics, some of which will be discussed in detail in the later chapters of this book. Basic knowledge of other aspects of the Standard Model (SM) of particle physics is implied or expected to be looked up in one of the many relevant textbooks. Hints for further reading will be given at the relevant occasions.

Before presenting a brief outline of the following sections, some notations need to be specified. Natural units, i.e. $\hbar = c = 1$ will be employed throughout so that energy, momentum, and mass all have units of $\mathrm{eV} = e \cdot 1\,\mathrm{V} \approx 1.6 \cdot 10^{-19}\,\mathrm{J}$. In this context, it is particularly useful to recall that $\hbar c = 1 \approx 200\,\mathrm{MeV} \cdot \mathrm{fm}$ can be exploited to translate energy units into units of length and time. Cross sections are given in the customary unit of "barn"[1], $1\,\mathrm{b} = 10^{-24}\mathrm{cm}^2$, with metric prefixes of "pico" or "femto" as appropriate for measurements in particle physics. The coordinate system that will be used is shown in Fig. 2.1, which defines the x, y, and z axes as well as some angular quantities. Symbols written as $\vec{p}$ represent three-vectors, while p normally denotes a four-vector. The notation for matrices is $\mathbf{M}$.

This chapter starts with a historical overview of the development of QCD, followed by a brief reminder of the basics of QCD theory. The next section deals with the central aspects of perturbative QCD. Subsequently, Monte Carlo event generators are introduced, followed by a thorough discussion of jet algorithms. The chapter is completed by a section on theoretical uncertainties and associated techniques for their evaluation.

[1] The use of the unit "barn" goes back to December 1942, when it was introduced during wartime by M.G. Holloway and C.P. Baker. Because of its connection to nuclear physics this information was classified until 1948 [1].

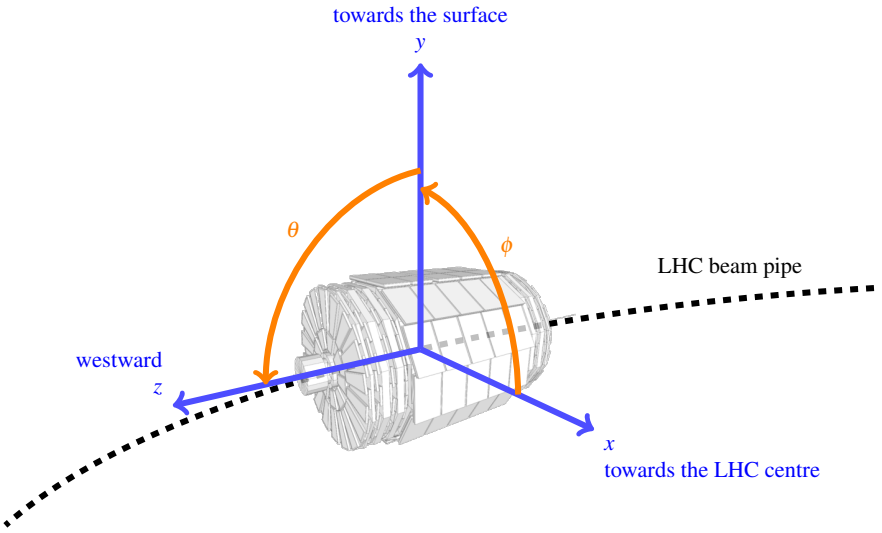

Fig. 2.1 Illustration of the coordinate system used by the LHC experiments at the example of the CMS detector: The experiments define a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the z axis pointing along the direction of the counterclockwise beam. The x axis points from the IP to the centre of the LHC ring, and the y axis points upwards, perpendicular to the plane of the LHC ring. Cylindrical coordinates (r,ϕ) are used in the transverse plane, ϕ being the azimuthal angle around the beam pipe. The pseudorapidity is defined in terms of the polar angle θ as $\eta = -\ln\tan(\theta/2)$; the rapidity, differing from η for massive objects, is defined as $y = 1/2 \cdot \ln[(E+p_z)/(E-p_z)]$. *(Illustration courtesy of D. Haitz [2], background image source: CERN, CMS)*

2.1 Historical Overview

Striving to describe as many phenomena in nature by as few fundamental assumptions as possible, physicists followed in the footsteps of chemists by replacing the roughly 100 chemical "elements" of the periodic table by merely two "elementary particles": the positively charged proton and the negatively charged electron composing the atomic nuclei and shells, respectively. To properly account for atomic weights and to compensate the strongly repellent electrical force between the protons inside a nuclei, the list had to be complemented with electrically uncharged neutrons discovered by J. Chadwick in 1932 [3]. Theoretical developments by P.A.M. Dirac referring to the quantum mechanical description of spin-$\frac{1}{2}$ particles like the electron lead to the prediction of anti-particles [4] and the discovery of the anti-electron, the positron, by C.D. Anderson, also in 1932 [5]. Refining the technologies to observe cosmic rays, the muon was found unexpectedly in 1936 by

C.D. Anderson and S. Neddermeyer [6] and first was mistaken for the so-called mu meson hypothesised one year earlier by H. Yukawa as carrier particle of the strong nuclear force [7]. However, it could be shown that the muon had the wrong properties and in particular did not take part in nuclear reactions. The real mu meson, the pion, was observed only eleven years later in 1947 by C. Lattes, G. Occhialini, and C. Powell, again while studying cosmic radiation [8, 9]. In parallel, sophisticated technologies were developed in the thirties and forties to accelerate charged particles, which later on gave an enormous impetus to the field named *elementary particle physics*. Unfortunately, the huge number of new "elementary" particles discovered in the fifties and sixties, most of them subject to nuclear forces and hence collectively called *hadrons* in contrast to *leptons*, rendered it again difficult to find the ordering principles behind this *particle zoo*.

A clearer picture only started to emerge with the arrangement of the known spin-0 mesons and spin-$\frac{1}{2}$ baryons into octets, cf. Fig. 2.2, according to their electrical charge q and the new quantum property of *strangeness* s that seemed to be conserved in the production process via nuclear collisions, but not in particle decays, which are ascribed to the weak force. M. Gell-Mann coined the term the *Eightfold Way* for this scheme and applying it to the spin-$\frac{3}{2}$ baryons, cf. Fig. 2.3, he and independently Y. Ne'eman declared at the 1962 Rochester conference that a baryon Ω^- with charge $q = -1$ and strangeness $s = -3$ must exist [10], which was discovered two years later in 1964 [11]. The basis for the successful explanation of the observed hadron spectra essentially is their association with an approximate flavour $SU(3)$ symmetry group. For this connection, Gell-Mann and independently G. Zweig invented hadron constituents such that all hadrons known at that time could be composed out of either three such constituents for the baryons, or out of a constituent-anticonstituent pair for the mesons [12–14]. Gell-Mann thought of these constituents as *kworks*, which he later rewrote [15] into *quarks* after reading this expression in James Joyce's "Finnegans Wake". These quarks were postulated to come in the three flavours *up*, *down*, and *strange*, which are conserved in strong (nuclear) and electromagnetic reactions, but not in weak decays.[2] The strangeness of a particle then simply counts the number of antistrange minus the number of strange quarks, where the minus sign is chosen in analogy to the sign of the quark's electrical charge. Peculiarly, however, the postulated quarks had to carry fractional electric charges of $+2/3$, $-1/3$ and $-1/3$, a feature which was never observed in nature.

Support for this quark constituent picture came from measurements of deep-inelastic electron-nucleon scattering (DIS) by the SLAC-MIT experiment at the Stanford Linear Accelerator Center [17,18]. They confirmed a conjecture by J.D. Bjorken from 1969 that the scattering cross section does not depend on the absolute energy or the momentum transfer (squared), Q^2, of the interaction but on dimensionless quantities like energy ratios or angles [19]. This "scaling" behaviour of the measured structure functions leads to a strikingly different prediction than a cross section falling steeply with increasing momentum transfer as expected from the

[2] Supposing there ought to be four constituents in analogy to the four leptons, e, μ, ν_e, and ν_μ, known in 1963, Zweig dubbed them *aces*, but the name did not stick.

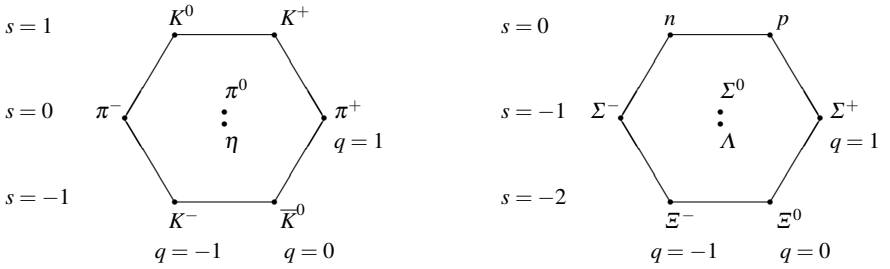

Fig. 2.2 Octets of spin-0 mesons (left) and spin-$\frac{1}{2}$ baryons (right) arranged according to their electrical charge q on the left-leaning diagonal and strangeness s on the horizontal lines. *(Adapted from Source: 2007 Wikipedia , Laurascudder [16])*

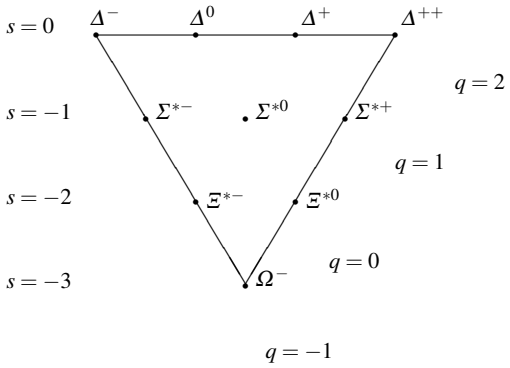

Fig. 2.3 Decuplet of spin-$\frac{3}{2}$ baryons arranged according to their electrical charge q on the left-leaning diagonal and strangeness s on the horizontal lines. *(Adapted from Source: 2007 Wikipedia , Laurascudder [16])*

product of elastic scattering and structure functions representing a finite size of the nucleon charge distribution. Interpreting the momentum transfer of the electron-nucleon scatter in terms of the spatial resolution at which the nucleons are probed, the scaling behaviour translates into an independence of the resolution scale and strongly suggests the presence of point-like scattering centres in a similar way that E. Rutherford concluded on the existence of an atomic nucleus decades before. In the slightly different context of high-energy hadron collisions such point-like constituents had also been proposed by R.D. Feynman who had given them the name *partons*, hence the name *quark-parton model* (QPM).

The so far missing dynamical description of strong interactions was greatly advanced by M. Veltman and G. t'Hooft who proved in 1971 that non-Abelian gauge field theories based on the special unitary group $SU(N)$ are renormalisable [20]. The candidate quantum field theory (QFT) of the strong interaction suggested by Gell-Mann and H. Fritzsch in 1972 was just of this type [21]. In contrast to the previous approximate $SU(3)$ flavour symmetry that is explicitly broken by the different quark

masses, the new threefold quantum number called *colour*, originally suggested by O. Greenberg [22], is associated with an exact $SU(3)$ symmetry of nature and each quark carries one of the three colours *red*, *green*, or *blue*. By attributing a different colour to each quark of a baryon, one can elegantly solve the so-called spin-statistics problem. The Δ^{++} resonance shown in Fig. 2.3 top right has a spin of $\frac{3}{2}$ and consists of three u quarks leading to a fully symmetric wave function, in contradiction to the Pauli principle that demands it to be fully antisymmetric. Using the new colour degree of freedom for quarks, the wave function of the Δ^{++} can be antisymmetrised.

In analogy to quantum electrodynamics (QED), the gauge field theory of the electromagnetic interaction that is based on the unitary group $U(1)$, the theory of the strong interaction is called quantum chromodynamics (QCD). Similarly to QED, the strong force between the colour-charged quarks is mediated via eight massless exchange bosons of spin-1, the *gluons*.[3] However, as a consequence of the non-Abelian character of the $SU(3)$ QCD gauge group, the eight gluons carry colour charges and interact amongst themselves via triple and quartic gluon vertices. This is a striking difference to the electromagnetic force, which is mediated by electrically uncharged photons.

Despite the additional colour degree of freedom, however, there do not seem to be any new hadrons associated with it. For an explanation the dynamics mediated by the self-interacting gluons has to be scrutinised. In QFT the *beta function* encodes the logarithmic dependence of a coupling parameter g on the relevant energy scale μ of a physical process:

$$\beta(g) = \frac{\partial g}{\partial \log(\mu)} \,. \tag{2.1}$$

D.J. Gross, H.D. Politzer, and F. Wilczek calculated in 1973 and 1974 in a perturbative expansion the beta function of QCD to lowest order and found that it has a negative sign in contrast to the beta function of QED [24–27]. As a consequence the strong force increases with distance, inversely to the electromagnetic force, and becomes small at very high energies, i.e. at subnuclear dimensions. This property is called *asymptotic freedom*. Gross, Politzer, and Wilczek also noted that the approximate scaling of structure functions of deep-inelastic electron-nucleon scattering as observed in the SLAC-MIT experiment could now be understood in terms of a QFT. The gross violation of scaling behaviour predicted by any QFT is, in agreement with experiment, reduced to a mild logarithmic scaling violation in asymptotically free theories. Thus, accounting for the gluon degrees of freedom, Feynman's point-like parton constituents of hadrons can be identified with the asymptotically free coloured quarks, antiquarks, and gluons of QCD.

On the other hand the approximately linear growth in strength of the strong force with increasing spatial separation between two colour charges leads to the fact that colour-charged objects are confined to subnuclear dimensions. This property of QCD is called *confinement*. Only entities that are colour singlets, i.e. without any net colour charge, are not subject to strong interactions and can be ob-

[3] The name "gluon" initially was introduced by Gell-Mann in a slightly different context without reference to color [23]

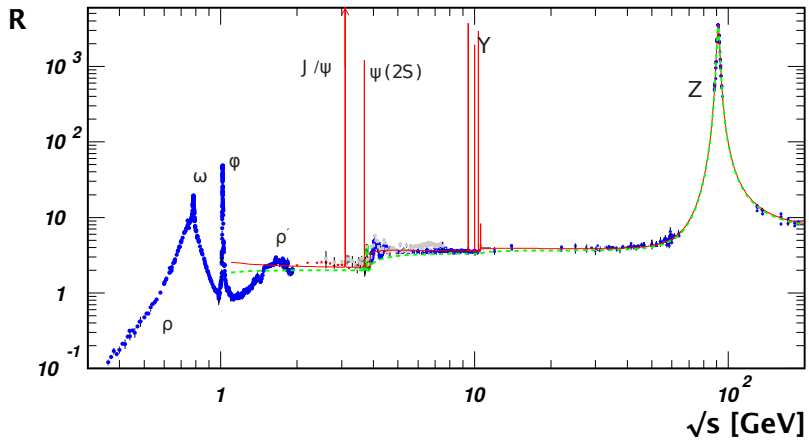

Fig. 2.4 World data on the ratio R of the total cross section of the reaction $e^+e^- \to$ hadrons to the cross section of the reaction $e^+e^- \to \mu^+\mu^-$ as a function of the centre-of-mass energy $\sqrt{s}$. *(Adapted from the 2015 update of Ref. [28])*

served as free particles. This is in complete accordance with experiment where only colourless hadrons are observed. Moreover, all such combinations of three quarks (or antiquarks) in the form of baryons or of a quark-antiquark pair in the form of a meson exhibit integer multiples of the electric charge, which explains why particles with non-integer electrical charges have never been observed in nature. Nevertheless, the fractional charges of the quarks have a direct influence on measurable quantities, notably the ratio R of the total cross section of the reaction $e^+e^- \to$ hadrons compared to $e^+e^- \to \mu^+\mu^-$. At centre-of-mass energies below the production threshold for the pairwise ($q\bar{q}$) production of *charm* or *bottom* quarks, one obtains $R = N_C \cdot \sum_q e_q^2 \cdot (1 + \delta_{m_q,\mathrm{QCD}})$ where $N_C = 3$ is the number of colours and e_q is the charge of the up, down, and strange quarks. $\delta_{m_q,\mathrm{QCD}}$ accounts for finite quark-mass and QCD corrections. When the collision energy surpasses the threshold for charm-anticharm or bottom-antibottom pair production, a corresponding step in this ratio is observed, cf. Fig. 2.4.

Today, QCD has been firmly established as the gauge theory of the strong interaction, one of the four fundamental forces of nature, and represents a cornerstone of the Standard Model of particle physics. It has a broad range of applications from high transverse momentum interactions between coloured quarks and gluons at TeV energies down to the low energy formation of hadrons and mesons. Numerous excellent text books and review articles have appeared in the course of time and provide extensive coverage of QCD and related topics.[4] General principles of QFT are discussed in Refs. [30–34], while Refs. [35–44] are specialised on QCD. The foundations of perturbative QCD are laid out in Ref. [45]. The reviews [46, 47] and the books [48, 49] are recommended specifically in the context of QCD at the LHC. For a historical perspective on QCD Refs. [50] and [51] can be consulted. The experi-

[4] References in this book have been managed with the help of JABREF [29].

mental foundations of particle physics in general are explained in Ref. [10] in their historical context including reprints of numerous original articles. The reference for the current state of research is the Review of Particle Physics by the Particle Data Group (PDG) [28], which is also available online.[5]

2.2 Basic Elements of QCD

As a gauge field theory, QCD is defined via its *Lagrangian density* $\mathcal{L}_{\text{QCD}}$, which is composed of four terms:

$$\mathcal{L}_{\text{QCD}} = \mathcal{L}_{\text{quarks}} + \mathcal{L}_{\text{gluons}} + \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{ghost}}\,, \tag{2.2}$$

where

1. $\mathcal{L}_{\text{quarks}}$ describes the interaction of spin-$\frac{1}{2}$ quark fields q_a of mass m_q with spin-1 gluon fields $\mathcal{A}^A_\mu$,
2. $\mathcal{L}_{\text{gluons}}$ represents the kinetic term of the gluon fields $\mathcal{A}^A_\mu$,
3. $\mathcal{L}_{\text{gauge}}$ defines the chosen gauge,
4. and $\mathcal{L}_{\text{ghost}}$ is the so-called *ghost* term that is a remedy necessary in non-Abelian gauge theories to treat the degeneracy of equivalent gauge field configurations.

Here and in the following, spinor indices are suppressed, Greek letters $\mu,\nu,\ldots \in \{0,1,2,3\}$ represent space-time indices, and $a,b,c \in \{1,\ldots,3\}$ and $A,B,C \in \{1,\ldots,8\}$ are the indices of the triplet and octet representations, respectively, of the colour $SU(3)$ gauge symmetry group. Summation over identical indices is implied. As in QED, the first term can be written with the help of the covariant derivative $(\mathcal{D}_\mu)_{ab}$ as

$$\mathcal{L}_{\text{quarks}} = \sum_{q\in\{u,d,s,c,b,t\}} \overline{q}_a \left(i\gamma^\mu (\mathcal{D}_\mu)_{ab} - m_q\right) q_b\,, \tag{2.3}$$

where the sum runs over all six quark flavours $\{u,d,s,c,b,t\}$ and γ_μ are the *Dirac matrices*. Defining the diagonal metric tensor g as $g^{\mu\nu} = \text{diag}(1,-1,-1,-1)$, the γ matrices satisfy the anticommutation relation

$$\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}\,. \tag{2.4}$$

In contrast to QED, however, the covariant derivative

$$(\mathcal{D}_\mu)_{ab} = \partial_\mu \delta_{ab} + ig_s \mathcal{T}^A_{ab}\mathcal{A}^A_\mu \tag{2.5}$$

not only exhibits colour indices a,b and the gauge coupling g_s of the strong interaction, but also, instead of one photon field for the sole generator of the $U(1)$ group,

[5] PDG: `http://pdg.web.cern.ch/pdg`.
Further useful resources for data, programmes, etc. are:
HEPDATA: `http://durpdg.dur.ac.uk/HepData`,
HEPFORGE: `http://www.hepforge.org/`.

eight gluon fields $\mathcal{A}^A_\mu$ with factors $\mathcal{T}^A_{ab}$ corresponding to the generators of the $SU(3)$ gauge group of QCD. A representation of the generators is given via $\mathcal{T}^A = \lambda^A/2$ by the Hermitian[6] and traceless *Gell-Mann matrices* λ^A:

$$\lambda^1 = \begin{pmatrix} 0 & +1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ +i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^3 = \begin{pmatrix} +1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\lambda^4 = \begin{pmatrix} 0 & 0 & +1 \\ 0 & 0 & 0 \\ +1 & 0 & 0 \end{pmatrix}, \quad \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ +i & 0 & 0 \end{pmatrix}, \quad \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & +1 \\ 0 & +1 & 0 \end{pmatrix},$$
$$\lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & +i & 0 \end{pmatrix}, \quad \lambda^8 = \tfrac{1}{\sqrt{3}} \begin{pmatrix} +1 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \tag{2.6}$$

The (2×2) submatrices of the first three λ_A can be recognised as *Pauli matrices*. The generator matrices $\mathcal{T}^A$ satisfy the commutation relations

$$\left[\mathcal{T}^A, \mathcal{T}^B\right] = i f^{ABC} \mathcal{T}^C, \tag{2.7}$$

where f^{ABC} are the corresponding *structure constants* of $SU(3)$ with values of

$$\begin{aligned} f^{123} &= 1 \\ f^{147} &= -f^{156} = f^{246} = f^{257} = f^{345} = -f^{367} = \frac{1}{2} \\ f^{458} &= f^{678} = \frac{\sqrt{3}}{2}, \end{aligned} \tag{2.8}$$

while all other f^{ABC} not related to these by index permutations are zero. The kinetic term of the gluons then reads

$$\mathcal{L}_{\text{gluons}} = -\frac{1}{4} \mathcal{G}^A_{\mu\nu} \mathcal{G}^{\mu\nu}_A, \quad \text{with} \quad \mathcal{G}^A_{\mu\nu} = \partial_\mu \mathcal{A}^A_\nu - \partial_\nu \mathcal{A}^A_\mu - g_s f^{ABC} \mathcal{A}^B_\mu \mathcal{A}^C_\nu \tag{2.9}$$

being the field strength tensor. In a pictorial representation,[7] these two "classical" parts correspond to the free quark- and gluon-field terms, and the quark-gluon interaction term as depicted in Figs. 2.5 and 2.6. In addition, the non-Abelian group structure of QCD leads to the cubic and quartic gluon self-interaction vertices, which are proportional to g_s and g_s^2, respectively.

Although not obvious from the Gell-Mann matrices or the structure constants, it can be shown that the probabilities for a quark emitting a gluon, or gluons splitting into a quark-antiquark or a gluon pair are identical with respect to each colour. Following the conventional normalisation of the colour $SU(3)$ matrices via the trace $\mathrm{Tr}\left\{\mathcal{T}^A \mathcal{T}^B\right\} = T_F \delta^{AB}$ with $T_F = 1/2$, the relative factors for these processes are

[6] A Hermitian matrix $\mathbf{A}$ is equal to its complex conjugate transpose, i.e. $\mathbf{A} = \mathbf{A}^{*T} = \mathbf{A}^\dagger$.

[7] Feynman diagrams in this book have been drawn with the help of JaxoDraw [52].

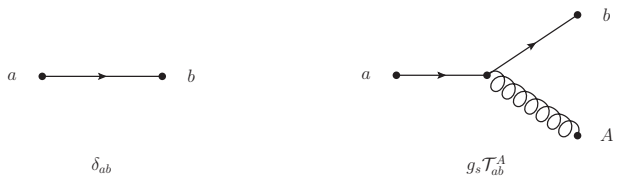

Fig. 2.5 Free quark-field and quark-gluon interaction term.

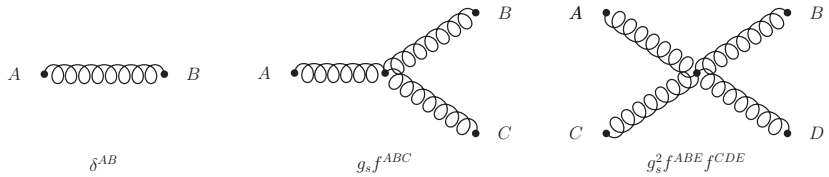

Fig. 2.6 Free gluon-field and cubic and quartic gluon self-interaction terms.

given by $C_F = 4/3$, $T_F = 1/2$, and $C_A = 3$. The ratio of a gluon emission by a gluon relative to the emission by a quark is therefore approximately $C_A/C_F = 9/4 = 2.25$, i.e. gluons radiate stronger than quarks by more than a factor of two. Similarly, gluons split into a gluon pair more often than into a quark-antiquark pair by roughly a factor of $C_A/T_F = 6$.

Of course, this classical QCD Lagrangian exhibits the property of local gauge invariance, i.e. invariance under a simultaneous redefinition of the quark and gluon fields. As a consequence of this internal symmetry, it is impossible to define the gluon field propagator without explicitly specifying a choice of gauge. A popular choice is given as a generalisation of the covariant Lorentz gauge $\partial^\mu \mathcal{A}^A_\mu = 0$ by the class of R_ξ gauges, imposed by adding the term

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{2\xi}\left(\partial^\mu \mathcal{A}^A_\mu\right)^2 \tag{2.10}$$

to the classical Lagrangian. Following L.D. Faddeev and V.N. Popov [53] this must be accompanied by the ghost term

$$\mathcal{L}_{\text{ghost}} = \partial_\mu \eta^{A\dagger}\left(\mathcal{D}^\mu_{AB}\eta^B\right) \tag{2.11}$$

because of the non-Abelian character of the QCD gauge group. The ghosts η^A, with conjugate-transpose $\eta^{A\dagger}$, represent complex scalar fields that nevertheless obey Fermi–Dirac statistics. They do not have a physical meaning, but should be considered as a mathematical trick to cancel nonphysical degrees of freedom otherwise present in calculations with covariant gauges.

This completes the Lagrangian for a consistent QFT of the strong interaction. Further invariant terms to add to the QCD Lagrangian are conceivable in principle. Renormalisability, however, forbids all additions that require coefficients with nega-

tive mass dimensions. Moreover, mass terms for the gluon fields, $m_A^2 \mathcal{A}_\mu^A \mathcal{A}^{A\mu}$, would violate gauge invariance. The only leftover possibility makes use of the *dual field strength tensor* defined as

$$\widetilde{\mathcal{G}}_{\mu\nu}^A = \frac{1}{2} \varepsilon_{\mu\nu}{}^{\rho\sigma} \mathcal{G}_{\rho\sigma}^A \,, \tag{2.12}$$

where $\varepsilon_{\mu\nu}{}^{\rho\sigma}$ is the four-dimensional antisymmetric *Levi-Civita symbol*. With the help of this definition the so-called θ-term is written as

$$\mathcal{L}_\theta = \theta \frac{g_s^2 T_F}{16\pi^2} \mathcal{G}_{\mu\nu}^A \widetilde{\mathcal{G}}^{A\mu\nu} \,. \tag{2.13}$$

This contribution would give rise to violations of the discrete symmetries of parity P and time reversal T by the strong interaction, which have never been observed in nature. Since T-violations are equivalent to a violation of the combined symmetry of charge conjugation C followed by P, experimental upper limits on the CP-violating electric dipole moment of the neutron lead to the conclusion that $|\theta|$ must be smaller than 10^{-10} [28]. Since no obvious reason is known for the smallness of this parameter, this is called the *strong CP problem*. New phenomena like *axions* as suggested by R. Peccei and H. Quinn in Ref. [54, 55] could provide an explanation.

2.3 Perturbative QCD

Given the complete QCD Lagrangian, quantitative predictions can be obtained either by means of lattice gauge theory (LGT), which applies QCD to a world discretised in space and time, or by using perturbative QCD (pQCD), which is valid in the asymptotically free regime, i.e. at high momentum transfers or respectively small distances, where quarks and gluons are weakly coupled. Computations in LGT are extremely complex and time-consuming and for a long time were not possible without severe approximations. For predictions in the context of collisions at the LHC, they are of very limited practical importance at present. On the other hand, such interactions at high-p_T are an ideal testing ground for a perturbative analysis of the strong interaction. In practice, however, the calculations are complicated by the occurrence of singularities that need to be properly addressed.

2.3.1 The Strong Coupling Constant

All Feynman rules required for a perturbative analysis can be deduced from the presented QCD Lagrangian. The relevant parameter in such a perturbative expansion is the gauge coupling g_s or equivalently the *strong coupling constant* α_S, which is defined in analogy to the QED *fine structure constant* α as $\alpha_S = g_s^2/4\pi$. With the knowledge of this fundamental parameter of QCD, in addition to the quark

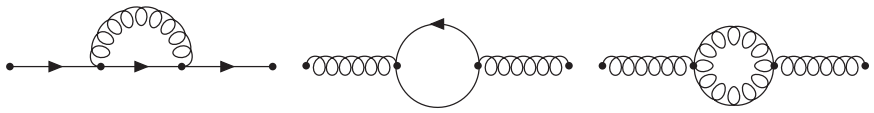

Fig. 2.7 Quark (left) and gluon (middle and right) self-energy corrections.

Fig. 2.8 Quark-gluon vertex corrections.

masses, tree-level amplitudes are calculable. Complications arise when loop diagrams come into play, because the momenta in a loop are not fully constrained by four-momentum conservation. The remaining integral over a loop momentum is logarithmically divergent for arbitrarily large momenta, which corresponds to an infinitely fine resolution. The strategy to overcome such *ultraviolet divergencies* that occur for example in self-energy corrections to quark and gluon lines, cf. Fig. 2.7, or in vertex corrections like in Fig. 2.8, is called *renormalisation*. Since QCD was proven to be a renormalisable QFT [20], the infinities can be absorbed into a finite number of parameters that need to be taken from experiment: the renormalised quark masses and coupling constant. Expressed in renormalised quantities, the Lagrangian of QCD must have exactly the same structure as the "bare" one presented in the previous section, such that counterterms cancel the divergencies to all orders in perturbation theory. As a consequence, precise higher order predictions can be made using the measured renormalised parameters. Details on the theoretical procedures are beyond this work and can be found for example in Ref. [35] specialised on QCD, or in Ref. [30] for QFT in general. For a historical perspective it might be interesting to look up Ref. [56].

The most powerful tool for theoretical predictions, perturbative QCD, can thus be applied and provides meaningful results as long as $\alpha_S \ll 1$. For a given observable $\mathcal{X}$ the expansion can be written as:

$$\mathcal{X} = c_0 + c_1\alpha_S + c_2\alpha_S^2 + \ldots = \sum_{i=0}^{n} c_i \cdot \alpha_S^i \tag{2.14}$$

with the c_i being the perturbative coefficients of this expansion. A price to pay for the renormalisation of the theory, however, is the introduction of a regulator for the infinities, the *renormalisation scale* μ_r. The renormalised parameters, e.g. the strong coupling constant, and hence the predictions in pQCD depend logarithmically on this nonphysical scale unless all orders could be summed up. The exact dependence of $\alpha_S(\mu_r^2)$ on μ_r is given by QCD through the *renormalisation group equation* (RGE), which determines the *running* of the renormalised coupling con-

stant $\alpha_S(\mu_r^2)$

$$\mu_r^2 \frac{\partial \alpha_S(\mu_r^2)}{\partial \mu_r^2} = \beta(\alpha_S(\mu_r^2)), \quad \text{with} \quad \beta(\alpha_S) = -\alpha_S^2 \cdot (b_0 + b_1 \alpha_S + b_2 \alpha_S^2 + \mathcal{O}(\alpha_S^3)). \tag{2.15}$$

The β function of QCD is a prime example of a quantity that can be evaluated in perturbation theory. The first three coefficients of the expansion are

$$b_0 = \frac{33 - 2N_F}{12\pi}, \; b_1 = \frac{153 - 19N_F}{24\pi^2}, \; b_2 = \frac{77139 - 15099N_F + 325N_F^2}{3456\pi^3}, \tag{2.16}$$

where N_F denotes the number of quark flavours with masses m_q smaller than the scale μ_r. Notably, the first term is $-b_0 < 0$ as long as $N_F \leq 16$, in contrast to QED where the corresponding coefficient is $8/3 > 0$. The non-Abelian nature of QCD manifests itself in this negative sign of the β function. In total, the coefficients are known up to four-loop order [57] and are renormalisation-scheme dependent starting with b_2, which is quoted here in the *modified minimal subtraction* ($\overline{\text{MS}}$) scheme [20, 58, 59]. Retaining only the leading term b_0, Eq. (2.15) is solved by

$$\alpha_S(Q^2) = \frac{\alpha_S(\mu_r^2)}{1 + b_0 \ln(Q^2/\mu_r^2)\, \alpha_S(\mu_r^2)}, \tag{2.17}$$

which relates the strength of the coupling at a scale Q to the one at scale μ_r, assuming both scales to be in the perturbative regime. With $b_0 > 0$, the coupling becomes weaker at higher scales Q, or, in other words, the effective colour charge gets smaller when the distance decreases. The consequence is asymptotic freedom, a key property of QCD, caused by the genuine quantum effect of anti-screening of colour charges through gluon self-interactions. The 2012 world average value of the strong coupling constant, quoted at the scale of the Z-boson mass M_Z, is given by

$$\alpha_S(M_Z) = 0.1184 \pm 0.0007, \tag{2.18}$$

derived from hadronic τ-lepton decays, lattice QCD calculations, DIS data, e^+e^- annihilation processes, and electroweak precision fits [60]. To be considered for this average, the perturbative expansion of the theory, to which data are compared, must be known at least to next-to-next-to-leading order. Figure 2.9 shows the respective determinations of α_S as a function of the relevant energy scale Q. In addition, some results from fits at lower theoretical precision are included in the plot to demonstrate the running of the strong coupling constant.

At small momentum transfers Q, i.e. large distances, the perturbatively defined strong coupling constant grows beyond the validity of the perturbative approach. Defining the value where $\alpha_S(Q)$ formally diverges as Λ_{QCD}, an analytic solution to Eq. (2.15) can be given at one-loop level as

$$\alpha_S(Q) = \frac{1}{b_0 \cdot \ln\left(Q^2/\Lambda_{\text{QCD}}^2\right)}, \tag{2.19}$$

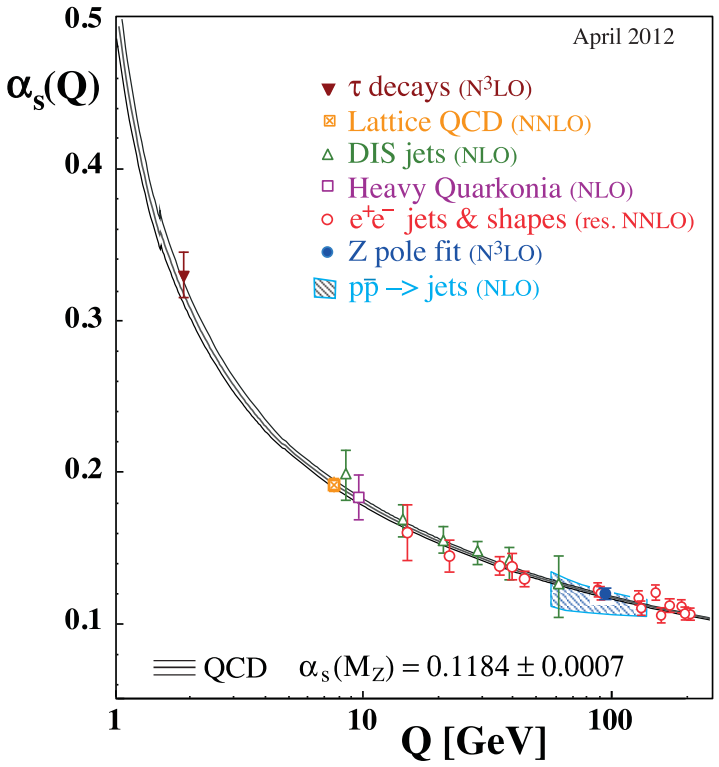

Fig. 2.9 Running of the strong coupling constant as of 2012. Determinations of the strong coupling constant α_S are shown as a function of the relevant energy scale Q of the respective process. The range in Q extends up to 208 GeV from e^+e^- annihilation data taken at the LEP collider. The hatched band indicates recent extractions from $p\bar{p}$ collisions at the Tevatron. *(Taken from Ref. [60])*

where experimentally $\Lambda_{\mathrm{QCD}} \approx 200\,\mathrm{MeV}$ for $N_F = 5$ in the $\overline{\mathrm{MS}}$ scheme. This value, which is close to typical hadron masses and sizes,[8] draws the dividing line between perturbative QCD and the manifestly non-perturbative regime of QCD. This growth of the coupling constant at small scales Q makes QCD the theory of the strong interaction that confines the quarks and gluons into ordinary hadronic matter. The transition from weakly interacting quarks and gluons inside hadrons to the formation of colour-neutral hadrons can not be described by perturbation theory. The alternative approach of LGT is not restricted to expansions around $\alpha_S \ll 1$, but requires vast amounts of computing power and despite severe approximations such an approach originally was considered hopeless. Drastic technical improvements and an exponential growth of computing capacities over the last decades, however, lead to almost "real world" applications nowadays. For example, it has become possible in recent years to calculate hadron masses *ab initio* [61], e.g. of the proton and neutron, provided mass scales for the three light quarks are set via the precisely

[8] Recall that $\hbar c = 1$ in natural units roughly corresponds to 200 MeV/fm.

measured pion, kaon, and Σ baryon masses. Combining pQCD with LGT, the latter can also be used to determine the strong coupling constant at small scales Q with seemingly very small uncertainties as visible in Fig 2.9. The uncertainty of $\alpha_S(M_Z)$ as given in Eq. (2.18) has, in fact, become very small compared to previous determinations because of the estimates of LGT, cf. Ref. [60]. This will further be discussed in Chapter 7 on future perspectives.

2.3.2 Cross-Section Predictions

Perturbation theory so far is applicable to reactions between partons. In nature, however, confinement prohibits to observe free quarks and gluons. Instead only hadrons with a complex internal structure are available for the production of high-p_{T} collisions. The situation is saved by the *factorisation theorem* of QCD [45,62], by which a cross-section computation can be separated into two parts, a short-distance partonic cross section tractable with pQCD, and a manifestly non-perturbative part that parameterises long-distance effects through universal functions extractable from measurements in a process independent way. In this approach, the partonic content of the colliding hadrons is described by parton distribution functions (PDFs), $f_{i/h}(x,\mu_f)$, which quantify the probability to find a parton i with longitudinal momentum fraction x within a hadron h at a resolution characterised by the factorisation scale μ_f. Transverse degrees of freedom in the initial state usually can be neglected compared to the collinear momentum component and are integrated over. Specialising to pp collisions and assuming that collinear factorisation holds, the cross section of a high-p_{T} scattering process can be written in the following form:

$$\mathrm{d}\sigma_{(pp\to X)} = \sum_{i,j} \int \mathrm{d}x\mathrm{d}x' f_{i/p}(x,\mu_f)\cdot f_{j/p}(x',\mu_f) \times \mathrm{d}\hat{\sigma}_{(ij\to X)}(x,x',\mu_f,\mu_r,\alpha_S(\mu_r))\,, \tag{2.20}$$

where i,j are the initial-state parton flavours and $f_{i/p}$, $f_{j/p}$ are the proton PDFs as functions of the fractional momenta x, x' of i and j, respectively. The sum extends over all contributing initial-state partons $i,j \in \{q,\bar{q},g\}$, and the factorisation scale μ_f defines the resolution, below which the physics is absorbed into the non-perturbative PDFs. At higher resolution, any physics process is described by pQCD in the form of a parton-level cross section $\mathrm{d}\hat{\sigma}_{(ij\to X)}$ that depends on the momentum fractions x, x', the factorisation and renormalisation scales μ_f and μ_r, and the strong coupling constant $\alpha_S(\mu_r)$. A pictorial representation is given by Fig. 2.10.

The typical scale Q associated with the partonic process is assumed to be in the perturbative domain, i.e. much larger than Λ_{QCD}. The squared partonic centre-of-mass energy is given by $\hat{s} = xx's$, where $s = (P_1+P_2)^2$ is the squared hadronic centre-of-mass energy and P_1, P_2 are the four-momenta of the incoming hadrons.

The collinear factorisation ansatz is the key element for quantitative predictions in perturbative QCD. Based on the property of asymptotic freedom, the desired cross sections with relevant scale Q can be expanded as a power series in the strong

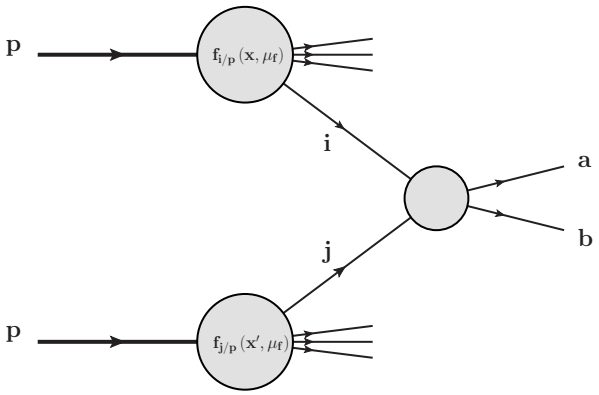

Fig. 2.10 Sketch of one parton i of proton one and one parton j from proton two participating in a high-p_T scattering reaction $A(ij \to ab)$.

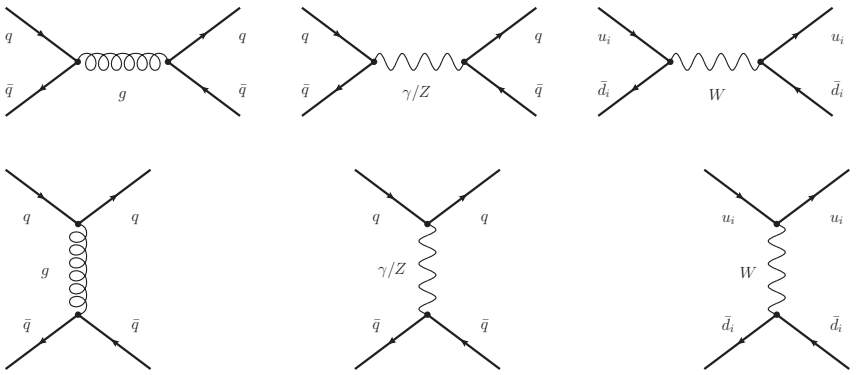

Fig. 2.11 Some LO QCD Feynman diagrams and LO EW corrections for $q\bar{q} \to q\bar{q}$ and $u_i\overline{d_i} \to u_i\overline{d_i}$.

coupling constant $\alpha_S(Q)$. Conventionally, the lowest-order contribution is denoted as leading order (LO), the subsequent ones as next-to-leading order (NLO), next-to-next-to-leading order (NNLO) and so forth. More precisely, these terms should be labelled as e.g. NNLO QCD corrections. For transverse momenta at the TeV scale or under particular kinematic conditions, electroweak (EW) tree-level effects of $\mathcal{O}\left(\alpha\alpha_S, \alpha^2\right)$ and loop effects of $\mathcal{O}\left(\alpha\alpha_S^2\right)$ might become equally or even more important than a QCD NNLO term. Figure 2.11 provides some examples of LO QCD and LO EW Feynman diagrams contributing to the $q\bar{q} \to q\bar{q}$ and $u_i\overline{d_i} \to u_i\overline{d_i}$ processes. For further details Refs. [48, 63, 64] are recommended.

The universal PDFs required in this factorised ansatz must be extracted from data. Their dependence on the factorisation scale μ_f, however, is again prescribed by QCD via the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) [65–68] evolution equations:

$$\mu_f^2 \frac{\partial f_i(x,\mu_f)}{\partial \mu_f^2} = \sum_{j=\{q,\bar{q},g\}} \int_x^1 \frac{dz}{z} \frac{\alpha_S}{2\pi} P_{ij}(z) f_{j/p}(x/z,\mu_f)\,, \tag{2.21}$$

where the P_{ij} are the regularised Altarelli–Parisi (AP) splitting functions

$$\begin{aligned}
P_{qq}(z) &= C_F \left(\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z) \right), \\
P_{qg}(z) &= T_F \left(z^2 + (1-z^2) \right), \qquad P_{gq}(z) = C_F \left(\frac{1+(1-z)^2}{z} \right), \\
P_{gg}(z) &= 2C_A \left(\frac{z}{(1-z)_+} + \frac{1-z}{z} + z(1-z) \right) + \delta(1-z) \frac{11C_A - 4N_F T_F}{6},
\end{aligned} \tag{2.22}$$

with $P_{\bar{q}\bar{q}} = P_{qq}$ and $P_{g\bar{q}} = P_{gq}$.[9] The splitting functions, listed here at one-loop, i.e. LO approximation, are known up to three-loop accuracy [69, 70]. When calculating the LO, NLO, ... estimates of a partonic cross section, the QCD evolution of the PDFs has to be used at the same relative order.

The LO DGLAP evolution allows for an interpretation by means of simple branching processes. A parton i resolved at a scale μ_f may have originated from the branching of a parton j resolved at some higher scale. This transition of parton j to i is accompanied by the emission of an additional QCD parton. When applying the DGLAP equations to solve for the scale evolution, these emissions are ignored by considering inclusive processes only.

First investigated in the context of DIS under kinematic conditions with a large transverse momentum squared Q^2, the task was to sum up all leading terms that give rise to logarithmically enhanced contributions proportional to $(\alpha_S \log Q^2)^n$ to the cross section. With respect to the parton kinematic plane in (x, Q^2) as shown in Fig. 2.12 indicating the accessible phase space for fixed-target, HERA, and LHC experiments, the DGLAP evolution equations connect PDF sensitive measurements at different scales Q^2. On the other hand, for very small momentum fractions x probed at large energies $\sqrt{s}$ in the forward or backward directions with respect to the beams, some PDFs, e.g. the gluon one, rise dramatically. Summing up the leading logarithmic contributions proportional to $(\alpha_S \log(1/x))^n$ the PDF evolution is described by the Balitsky–Fadin–Kuraev–Lipatov (BFKL) [71, 72] equations. For the purpose of the high-p_{T} processes considered here the DGLAP equations are sufficient.

2.3.3 Parton Luminosity

Even without specifying the partonic production process $\mathrm{d}\hat{\sigma}_{(ij\to X)}$ in Eq. (2.20), a lot can already be learned from the parton kinematics and the PDFs. Specifying the in-

[9] The subscript "+" indicates the use of the *plus prescription* defined via the distribution $\int_0^1 \frac{f(x)}{(1-x)_+}\mathrm{d}x = \int_0^1 \frac{(f(x)-f(1))}{(1-x)}\mathrm{d}x$ for any sufficiently smooth function f.

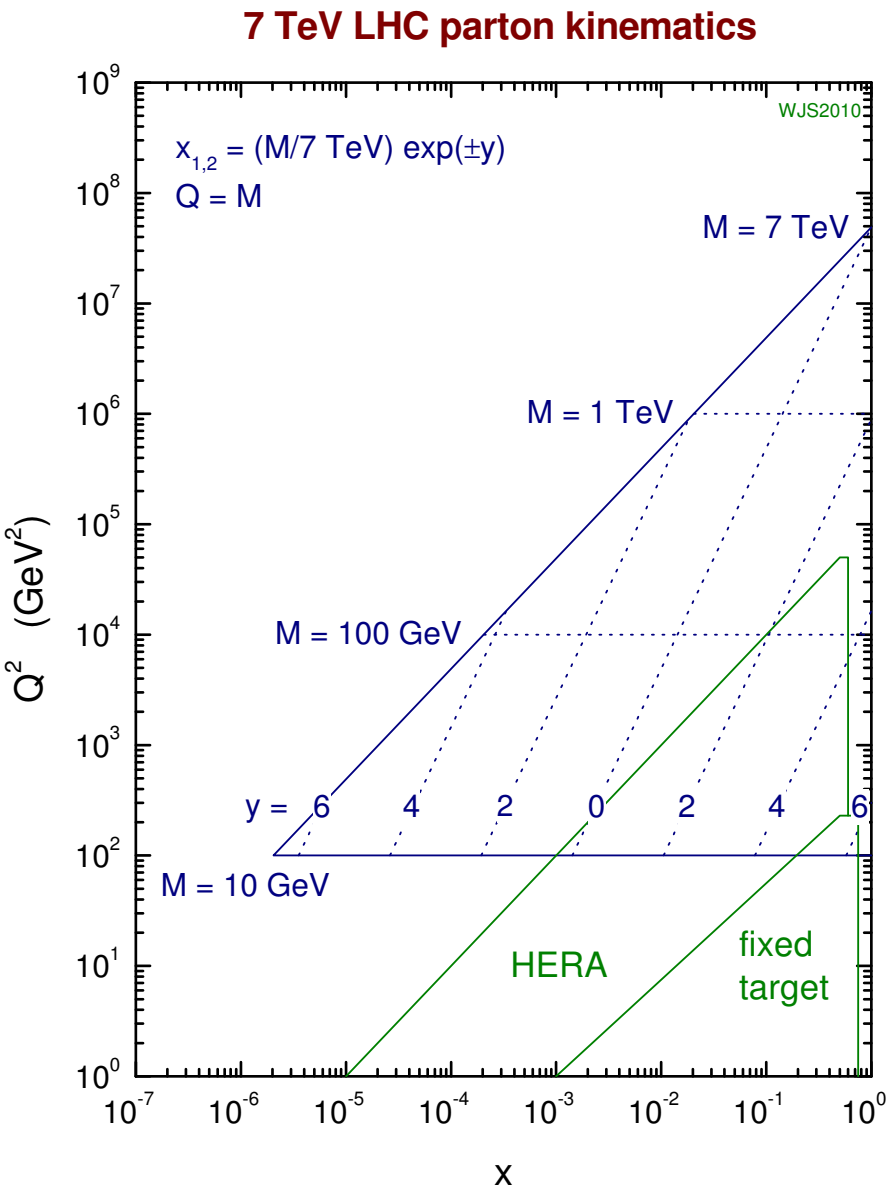

Fig. 2.12 Kinematic plane in parton fractional momentum x and squared energy scale Q^2 accessible to collisions in DIS at fixed-target experiments and at the HERA ep collider, and at the LHC for 7 TeV centre-of-mass energy. *(Taken from Source: 2010 W.J. Stirling [73])*

coming parton four-momenta in the centre-of-mass system as $p_1 = \sqrt{s}/2 \cdot (x,0,0,x)$ and $p_2 = \sqrt{s}/2 \cdot (x',0,0,-x')$, the rapidity y of a final state particle of mass $M = \sqrt{\hat{s}}$ is given by $y = 1/2 \cdot \ln(x/x')$ and the momentum fractions read $x, x' = (M/\sqrt{s}) \cdot e^{\pm y}$ as shown in Fig. 2.12 with $x_1 = x$ and $x_2 = x'$. With these parton momentum fractions the hadron beams effectively are parton beams of variable energy, where the energy profile is given by the PDFs. Assuming that the partonic cross section $\mathrm{d}\hat{\sigma}_{(ij\to X)}$ only depends on $\hat{s}$, it is useful to define the differential *parton luminosity*[10]

$$\frac{\mathrm{d}^2 L_{ij}}{\mathrm{d}\hat{s}\mathrm{d}y} = \frac{1}{s}\frac{1}{1+\delta_{ij}}\left[f_{i/p}(x,\mu_f) f_{j/p}(x',\mu_f) + \left(x \leftrightarrow x'\right)\right] \tag{2.23}$$

and its integral $\mathrm{d}L_{ij}/\mathrm{d}\hat{s}$. The pre-factor with the Kronecker δ corrects for double-counting in case of identical parton flavours i and j. The inclusion of s into the definition allows the comparison between different colliders. Using $\mathrm{d}\hat{s}\,\mathrm{d}y = s\,\mathrm{d}x\,\mathrm{d}x'$, the factorised cross section Eq. (2.20) can be transformed into

$$\mathrm{d}\sigma_{(pp\to X)} = \sum_{i,j} \int \mathrm{d}\hat{s}\mathrm{d}y \left[\frac{\mathrm{d}L_{ij}}{\mathrm{d}\hat{s}\mathrm{d}y}\right] \times \mathrm{d}\hat{\sigma}_{(ij\to X)}(\hat{s},\mu_f,\mu_r,\alpha_S(\mu_r))\,, \tag{2.24}$$

or

$$\sigma_{(pp\to X)} = \sum_{i,j} \int \mathrm{d}\hat{s} \left[\frac{\mathrm{d}L_{ij}}{\mathrm{d}\hat{s}}\right] \times \mathrm{d}\hat{\sigma}_{(ij\to X)}(\hat{s},\mu_f,\mu_r,\alpha_S(\mu_r))\,, \tag{2.25}$$

respectively. The term in square brackets, i.e. the parton luminosity, has units of a cross section. Provided reasonable estimates of the other numerical factors are possible, cf. Refs. [37] or [74], the parton luminosities allow order-of-magnitude estimations for cross sections as a function of the centre-of-mass energy.

2.3.4 Final State

One entity from Eq. (6.8) that has not yet been discussed is the final state X of a collision. The simplest reaction that can be considered is the Drell–Yan process [75], where a quark and an antiquark annihilate to produce a lepton pair: $\hat{\sigma}\,(q\bar{q} \to \ell^+\ell^-)$. Figure 2.13 shows relevant LO and NLO Feynman diagrams. In this case there are no strongly interacting particles in the final state and the theory prediction can directly be compared to the measured leptons. Merely the proton remnants, which fragment into hadrons along the beam lines, have to be described by non-perturbative models. At high transverse momenta, the two leptons are well separated from any such proton debris and high-precision comparisons with theory are possible.

However, in the vast majority of reactions at least some colour-charged partons are produced so that a further step covering the transition from the partonic final state

[10] This should not be confused with the luminosity, which is a characteristic of a collider.

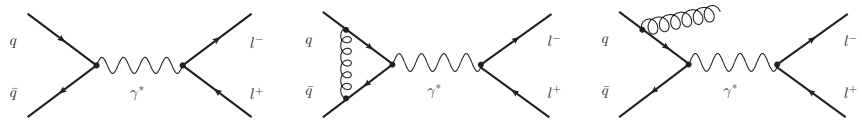

Fig. 2.13 Feynman diagrams for the Drell–Yan process $q\bar{q} \to \ell^+\ell^-$ at LO and NLO: tree-level amplitude (left), virtual correction (middle), and real correction (right).

to measurable particles, the so-called "particle level", is needed. Here, "measurable" refers to colour-neutral particles with mean decay lengths such that $c\tau > 10\,\mathrm{mm}$, where c is the speed of light and τ the mean lifetime of a particle. One possibility to account for this transition is to reuse the concept underlying the PDFs that describe the partonic content of a hadron, only in an inverted sense. The necessary functions $D_{k\to h}(z,\mu_F)$ are called fragmentation functions (FFs) and are the final-state analogues of the PDFs [76]. They parameterise the probability of finding a hadron h within the fragmentation products of parton k, carrying the fraction z of the parton momentum. With this definition, Eq. (2.26) reads

$$\mathrm{d}\sigma_{(pp\to X)} = \sum_{i,j,k} \int \mathrm{d}x\mathrm{d}x'\mathrm{d}z f_{i/p}(x,\mu_f)\cdot f_{j/p}(x',\mu_f) \times \mathrm{d}\hat{\sigma}_{(ij\to k)}(x,x',z,\mu_f,\mu_r,\alpha_S(\mu_r)) \times D_{k\to X}(z,\mu_F)\,, \tag{2.26}$$

where, like for the PDFs, fragmentation functions depend on a non-physical resolution or fragmentation scale μ_F. Again, these functions can currently not be determined by first principles in QCD, but once they have been measured, they are valid universally. Experimentally favourable conditions prevail for example in e^+e^- collisions, where $q\bar{q}$ pairs are created via the inverse of the Drell–Yan process described above.

Other possibilities to account for the transition to measurable particles make use of the concepts of *energy flow* [77] and *particle jets* [78–82]. Instead of scrutinising the detailed production of identified particles—an experimentally very challenging endeavour—for the majority of processes it is sufficient to know how much energy or momentum is carried away by hadrons into a specific direction. Focusing on the normalised spatial distribution of the energy flow, the "shape" of an event (or an ensemble of events) can be compared to QCD radiation patterns. The influence of non-perturbative (NP) effects on such *event shapes* is expected to be power-suppressed with respect to some process-relevant energy scale Q [83]. Event shapes are very popular study subjects in e^+e^- collisions, lead to the discovery of the gluon [84–86] at the e^+e^- colliders of DESY in Hamburg, and are at the basis of one of the $\alpha_S(M_Z)$ determinations entering the world average reported in Ref. [28]. They are further discussed in Section 6.6.

QCD also predicts that the large-distance NP effects are mostly decoupled from the hard reaction so that highly energetic partons fragment into a collimated stream or "jet" of hadrons, which inherits energy and momentum from its parent parton. More precisely, the transverse momenta of individual hadrons with respect to the

jet direction are expected to be typically of the order of Λ_{QCD}. Of course, the term "collimated" requires a mathematical prescription that, given some distance measure, unambiguously decides which objects belong to a jet. Since such particle jets are **THE** primary subject of this work, the whole Section 2.5 is dedicated to their introduction.

If complete knowledge of the final state of a collision event is mandatory, for example in detailed simulations of a complex experimental apparatus, perturbative methods must be complemented with models for the NP effects. This is the domain of *general-purpose Monte Carlo (MC) event generators*, which will be introduced in the next Section 2.4.

2.4 General-purpose Monte Carlo Event Generators

For a first overview it is interesting to compare the cross sections for a set of standard processes as a function of the centre-of-mass energy. Such an overview is given in Fig. 2.14, where typical values can be read off, in decreasing order, for the production of b quarks, jets, W and Z bosons, top quarks, and Higgs bosons separately for the three major production processes. In addition, the total cross section is indicated as well as two diboson production channels and jet production above a minimal jet p_{T} that scales with $\sqrt{s}/20$. The plot spans about a dozen orders of magnitude between the total cross section and Higgs production, i.e. only each trillionth event of any pp collision produces one of the famous Higgs particles discovered in 2012 by the ATLAS and CMS experiments at the LHC [87, 88]. On the other hand, top quarks, discovered with a handful of events at the Tevatron collider in 1995 [89,90], are produced roughly at a rate of one per second at the LHC with $\sqrt{s} = 8\,\mathrm{TeV}$.

The cross sections presented in Fig. 2.14 are sufficient for a rough estimation of event types and rates to expect for a new collider. For the conception of an experiment such as the ones at the LHC, much more comprehensive predictions are indispensable, not only for the precise design with the help of detailed simulations, but also in the later operational phase for an accurate understanding of the recorded data. On the theory side, the most complete description of a collision event is given by general-purpose MC event generators. Based on physically motivated probability distributions, they generate every step in the reaction chain leading from the initial interaction between the two incoming beam particles to the final state that is composed of a list of specific particle types with their respective masses and four-momenta. Everything beyond the final-state particle level, i.e. for flight distances larger than 10 mm, cf. Section 2.3.4, is part of the detector simulation discussed in Section 3.4.1.

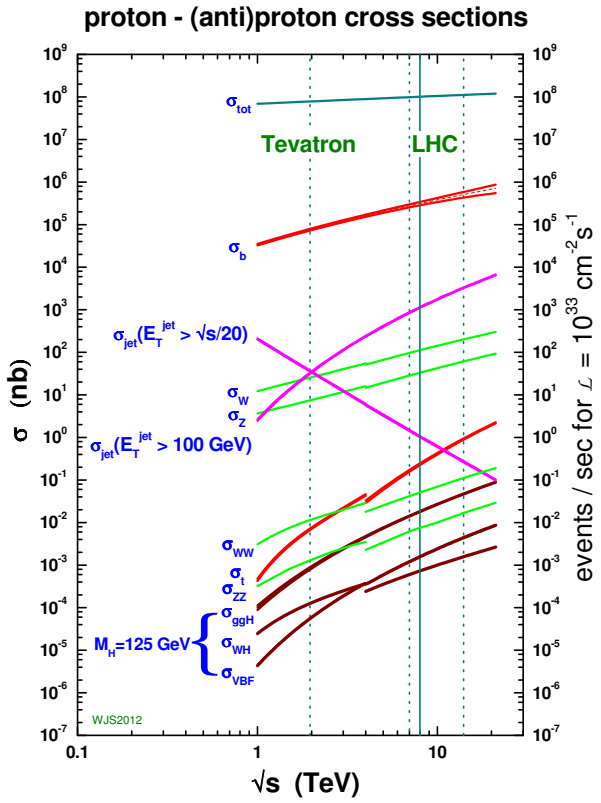

Fig. 2.14 Typical cross sections and event rates for SM processes including Higgs boson production as a function of the centre-of-mass energy in $p\bar{p}$ (below 4 TeV) and pp collisions (above 4 TeV). *(Taken from Source: 2012 W.J. Stirling [73])*

2.4.1 Non-perturbative Modelling

Three general-purpose MC event generators are in widespread use at the LHC: PYTHIA (The Lund Monte Carlo for high-p_{T} Physics), HERWIG (Hadron Emission Reactions With Interfering Gluons), and SHERPA (Simulation of High-Energy Reactions of PArticles). The first two have undergone decades of development from early versions in the FORTRAN programming language, PYTHIA [91–94], HERWIG [95–98], up to the ones in current use based on C++, PYTHIA8 [99, 100] and HERWIG++ [101], which most recently was renamed to HERWIG7 [102]. The new contender, SHERPA, was directly developed in a C++ framework [103].

Within the scope of this work it is hardly possible to do justice to all these developments and even less so to explain the concepts and models in great detail. To

complement the following short recapitulation, it is advised to consult the excellent overview of MC event generators in Ref. [104] or the corresponding chapter of Ref. [105]. References [106] and [107] are recommended as reviews or lectures that include also the latest developments, A pre-LHC summary can be found in Ref. [108].

Coming back to the simple example of Drell–Yan lepton-pair production, a sketch of a MC event generation is shown in Fig. 2.15. The small central blob in red represents a high-p_{T}, short-distance interaction, where in this case the pair of leptons is created. Even though the two leptons do not interact any further, the situation is complicated by the fact that the initial parton lines radiate gluons that evolve into cascades of partons. This process is described perturbatively by the concept of *parton showers* (PS) [109, 110]. If the high-p_{T} scattering products are partons as well, then the initial-state radiation (ISR) must be complemented by final-state radiation (FSR). Considered as a whole, the parton showers represent in their simplest form a perturbative leading-logarithmic approximation (LLA) to the full result. In contrast to the fixed-order partonic cross sections that well describe the high-p_{T} component of a scattering reaction, the LLA works explicitly well for the situation of collinear or soft radiation around a parton. In the basic use case, the MC event generators combine LO predictions with PS (LO+PS). Exploiting the freedom in the choice of shower resolution scale and inclusion of colour-coherence effects, several implementations exist. PYTHIA contains two incarnations with Q^2- and p_{T}-ordered PS, respectively [111–113], while HERWIG employs angular-ordered showers [95, 114]. SHERPA's version is based on the concept of dipole showers [115, 116] first developed in Ref. [117].

Finally, all partons in an event must form colourless compounds, represented by hollow ellipses in Fig. 2.15, that must be transformed into hadrons (yellow circles), which according to their lifetimes might still decay or not. For this manifestly non-perturbative step of *fragmentation*, frequently also called *hadronisation*, only models exist. In PYTHIA this step is performed with the *Lund string fragmentation* [118–120]. The alternative model is *cluster fragmentation*, which is used in HERWIG [121] as well as in SHERPA [122]. The pictorial view of Fig. 2.15 is based on the HERWIG version.

Unfortunately, this picture has to be refined further in order to account for additional, relatively diffuse "soft" particle production observed in hadron-hadron collisions, but absent from the theoretical description so far. Neither ISR and FSR nor the fragmentation of the beam remnants can sufficiently explain the effect. Moreover, the amount of extra particles (and energy) depends not only on the centre-of-mass energy but also on the energy scale of the primary interaction, which is characterised by the leading jet p_{T}. In phenomenological studies everything not associated to the high-p_{T} reaction is investigated and generically denoted as the *underlying event* (UE). Traditionally, this is done by geometrically subdividing an event into different regions of azimuthal angle, "towards", "away", and "transverse", with respect to the jet or particle leading in p_{T} as illustrated in Fig. 2.16. At the same time the p_{T} of the leading object is defined to be the so-called event scale, i.e. the measure of the momentum transfer in the partonic scattering. The transverse region then is

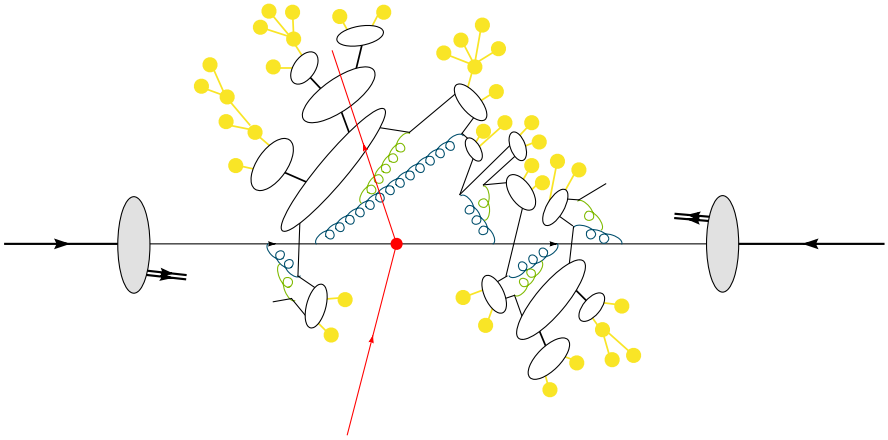

Fig. 2.15 Sketch of a MC event generation: One parton from each colliding hadron (grey blobs) interact and result in two high-p_{T} leptons (red). Gluon radiation off the incoming partons initiate parton showers, whose partons are combined into colourless clusters (hollow ellipses) that finally are fragmented into potentially further decaying hadrons (yellow circles). *(Sketch courtesy of S. Gieseke)*

used to measure the soft particle production that supposedly does not originate from the high-p_{T} scatter. Studies along this line were performed at a series of centre-of-mass energies at the Tevatron [123–126] and at the LHC by the ATLAS [127–130], CMS [131–135], and ALICE experiments [136].

Predictions so far are based on the naive picture of only one interacting parton from each colliding hadron to participate in a high-p_{T} reaction. It is known though that this picture must fail when approaching small transverse momenta, because the respective cross sections grow beyond all limits and finally violate unitarity. The assumption of two simultaneous high-p_{T} interactions is an interesting subject of study on its own and is closer examined in Section 6.4. To simulate the underlying event, the MC event generators accompany the primary reaction by multiple parton interactions (MPI) depending on the overlap in the transverse plane of the percolating hadrons. Collisions with a small impact parameter lead to high transverse momenta and more MPI, while a merely grazing collision exhibits only few soft particles. Partons produced by this extra activity interfere with other partons in the event, including the ones from ISR, FSR, and the beam remnants, and cannot be uniquely separated. A sketch of the resulting hadron-hadron event composition is shown in Fig. 2.17.

The original impact parameter model presented in Ref. [138] is used in PYTHIA, with modifications described in Refs. [113, 139], as well as in SHERPA. The eikonal multiple partonic scattering model applied in HERWIG is explained in Ref. [140]. Together with numerous other parameters steering the parton showers, hadronisation, and decays, the MPI-based phenomenological models employed in the MC event generation must be adapted or *tuned* to data, e.g. from the various experimental results listed above. Two popular tools employed for such tunings are RIVET [141]

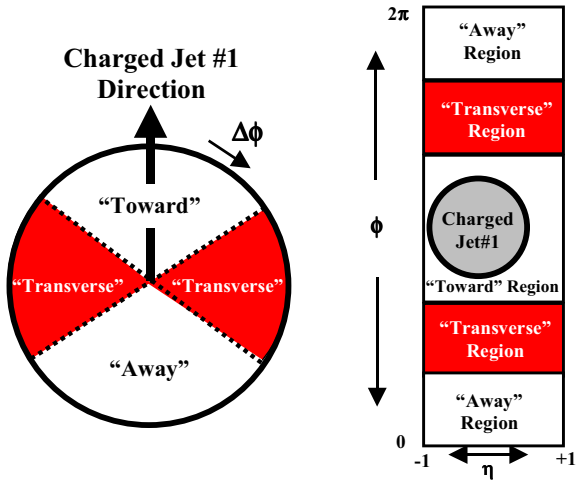

Fig. 2.16 Illustration of the "toward", "away", and "transverse" regions: The highest-p_T jet of an event defines the zero direction of the azimuthal angle in the transverse plane of a hadron-hadron collision. The leading jet as well as the balancing jet in the opposite direction dominate the particle and energy flow in these "toward"- and "away"-side named regions. Perpendicular to the latter is the "transverse" region (red) that is expected to receive predominantly particles from the underlying event. All particles and energies within the respective azimuthal angular intervals are summed up within a pseudorapidity range that usually coincides with the acceptance region of tracking devices, hence "charged" jet. *(Taken from Ref. [123])*

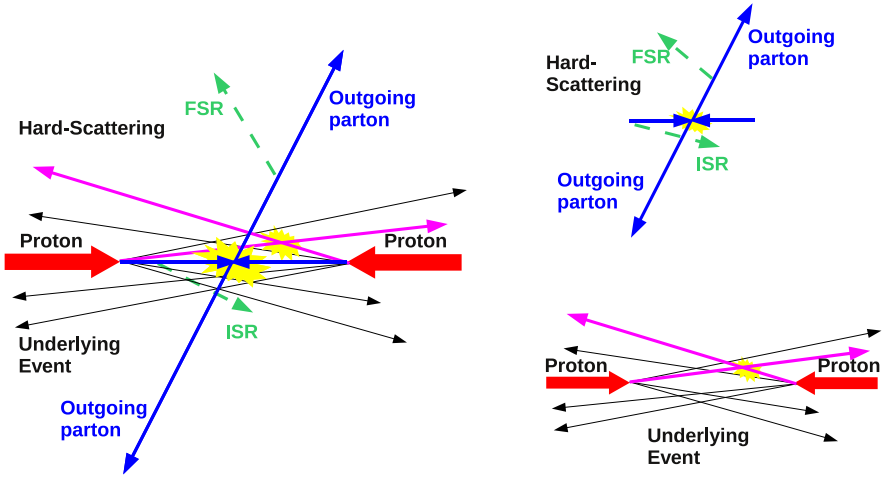

Fig. 2.17 Sketch of the composition (left) of a generated hadron-hadron collision. The "hard" parton-parton scattering, complemented with ISR and FSR, (top right) is accompanied by soft particle production from MPI and hadron remnant fragmentation (bottom right). Because of interference already at parton level it is not possible though to unambiguously attribute any interaction product to a particular effect. *(Adapted from illustration courtesy of D. Piparo [137] following an original design of R. Field)*

and PROFESSOR [142]. RIVET contains numerous MC event analyses that follow as closely as possible the observable definitions and selection criteria of published measurements. Thereby, for each such measurement, it efficiently provides sets of MC predictions for arbitrary choices of parameters to tune. These predictions are then delivered to the PROFESSOR framework, which parameterises the generator response and returns the set of tuned parameters that best fits the input measurements.

Recent developments in MPI and in MC tuning are discussed in a workshop series, cf. Refs. [143–145]. Dedicated MC tuning efforts are reported in Refs. [146–149]; a comprehensive review can be found in Ref. [150].

2.4.2 Perturbative Concepts

The standard perturbative concept applied so far to MC event generation consists in attaching parton showers to a LO prediction, which works well for the situation of collinear or soft radiation around few high-p_{T} partons. At LHC energies, however, high-p_{T} partons are produced abundantly. Therefore, more sophisticated strategies were needed that allow the automated calculation of multi-parton fixed-order results and their combination with parton showers. Several tree-level *merging or matching schemes* were developed that consistently dress varying multi-parton states with parton showers and combine everything into an inclusive event sample: the Catani–Krauss–Kuhn–Webber (CKKW) approach [151,152], the Lönnblad variant CCKW-L [153], and the MLM method [154]. A comparative review is presented in Ref. [155].

For most observables, however, NLO predictions in the strong coupling constant represent the first accurate theoretical estimate that allows an assessment of associated theoretical uncertainties. To improve the fixed-order part of MC event generation correspondingly, two challenges had to be overcome. First, methods had to be developed that deal with the more involved NLO calculations and properly cancel singularities between real-emission graphs and virtual one-loop corrections in a numerically stable way. A number of techniques have been established [156–159]. Secondly, when combining NLO matrix elements with parton showers, the obstacle of double-counting configurations that appear in both approaches needs to be resolved. For example, real-emission corrections as part of the fixed-order calculation have to be properly synchronised with the first, i.e. highest-p_{T}, shower splitting. Two strategies emerged, which are known as MC@NLO [160] and POWHEG [161–163] and which are compared for example in Ref. [164].

Thanks to this enormous progress made over the last years, sometimes dubbed "the NLO revolution", nowadays standard in terms of pQCD ingredients to MC event generation is the combination of NLO with parton showers including multi-parton tree-level corrections or even PS-matched NLO calculations of varying parton multiplicity [165–171]. Here, it should be noted that jet algorithms play an essential part in these developments by providing a safe mapping between m-parton and n-jet ($n \leq m$) final states.

2.5 Jet Algorithms

In common language a jet describes a collimated stream of objects forcefully moving into the same direction, like water molecules in a jet of water. An example of such a "jet d'eau" is the famous water fountain and symbol of the city of Geneva as shown in the photograph Fig. 2.18.

In experimental particle physics, jets are made of measured tracks or energy depositions as illustrated in Fig. 2.19. To decide unambiguously whether a measured object belongs to a jet or not, a mathematical prescription is required: a jet algorithm. Moreover, for comparisons to pQCD, which predicts cross sections in terms of quarks and gluons, an algorithm is needed that is applicable to theoretical calculations as well as to measurements from different experiments. The first prescription of such a jet algorithm was given by G. Sterman and S. Weinberg in 1977 [172] with respect to e^+e^- collisions, where particles (or energy depositions) are grouped

Fig. 2.18 The Jet d'Eau, the famous symbol of Geneva, Switzerland, the hosting city of the CERN laboratory. The water fountain has its origin in a safety valve for a hydraulic power network from 1886. In its nowadays form as a landmark and tourist attraction it reaches a height of 140 m, for which 500 litres of water per second are ejected at a speed of 200 km/h.

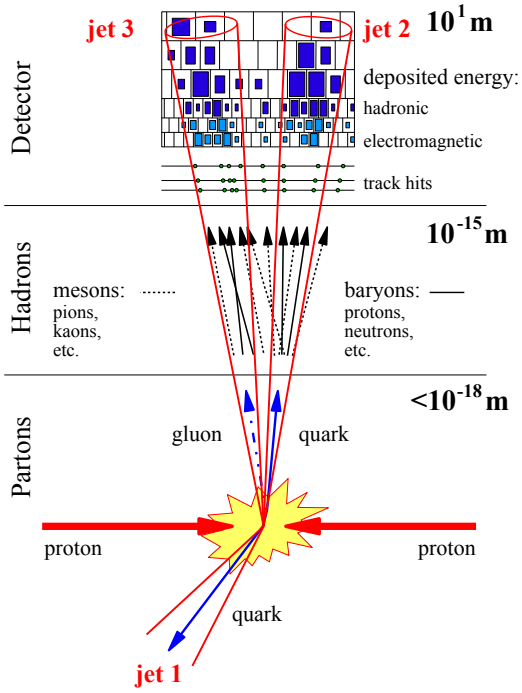

Fig. 2.19 Illustration of a jet to which bundles of partons, hadrons, or detector measurements are grouped together. *(Taken from Ref. [48])*

together depending on their location with respect to an angular cone around a specific direction. In the following decade this prescription was extended by theorists as well as by experimental groups in order to analyse hadron-hadron collisions in terms of a number of cone-shaped jets of a chosen jet size or radius, R, which are localised around the highest concentrations of energy in an event. In the same period a novel type of algorithm based on iterative pairwise clusterings was introduced by the JADE Collaboration for the analysis of e^+e^- events at the PETRA collider [173]. A summary of the jet algorithms in use at that time is presented in Ref. [174]. Two classes of jet algorithms emerged:

1. cone algorithms that assign objects to the leading energy-flow objects in an event based on geometrical criteria;
2. sequential-recombination algorithms that iteratively combine the closest pairs of objects.

It was quickly realised, however, that the comparability of jet quantities between experiment and theory or among different experiments was questionable at best because of serious shortcomings in these early algorithms. For example, many cone algorithms need starting points with a minimum energy or momentum for the cone directions, so-called “seeds”, which spoil their applicability in perturbation theory.

In a series of workshops extending over almost 20 years and starting with the one at Snowmass in 1990 [175–177], the uncovered issues could be addressed and mostly solved leading to a number of requirements, which are listed below from a nowadays perspective. An excellent recapitulation of the encountered problems and the developments can be found in Ref. [178].

2.5.1 General Desiderata

- Order independence: equal applicability to partons, particles, or measured tracks and energy depositions;
- Full specification: disclosure of all necessary details including the required software;
- Ease of implementation: avoidance of complex, proprietary code developments.

These rather general conditions are met by only permitting four-momenta as input objects and by using standardised public code for the jet clustering instead of proprietary implementations as done previously. The software library of reference in use at the LHC is FASTJET [179]. Furthermore, to compare parton-based theory predictions with experiment a clear definition must be given of the final state "truth level" to which measurements and theory are corrected [177], cf. the previous Section 2.3.4. The definitions in use by the ATLAS and CMS experiments are given in Section 3.2.2.

2.5.2 Theoretical Desiderata

- Well-defined finite cross sections at any order of pQCD: collinear and infrared safety;
- Longitudinal boost invariance: independence of jet observable of longitudinal boosts, in particular for hadron-hadron collisions;
- Boundary stability: insensitivity of jet kinematic boundaries from details of the hadronic final state e.g. the number of particles;
- Insensitivity to non-perturbative effects: limitation of impact of hadronisation and additional soft particle production (underlying event).

The by far most pressing issue that needed to be addressed is the *collinear and infrared safety* of a jet algorithm; otherwise the cancellation of collinear and soft singularities appearing in calculations of pQCD is spoiled and the most powerful computational technique for predictions, perturbation theory, is rendered useless. Hence, the outcome of a jet-clustering procedure must neither depend on the splitting or merging of collinear parton four-vectors nor on the addition of arbitrarily soft partons to the set of clustering objects. More formally, to be collinear- and infrared-safe an observable $F_m(p_1, \ldots, p_m)$, defined as a function of m four-momenta p_i, must

satisfy the following conditions:

$$\begin{aligned} F_m(p_1,\ldots,\lambda p_i, p_j=(1-\lambda)p_i,\ldots,p_m) &= F_{m-1}(p_1,\ldots,p_i,p_{j-1},p_{j+1},\ldots,p_m) \\ &\text{and} \\ \lim_{\lambda\to 0} F_m(p_1,\ldots,\lambda p_i,\ldots,p_m) &= F_{m-1}(p_1,\ldots,p_{i-1},p_{i+1},\ldots,p_m) \end{aligned} \tag{2.27}$$

with $0<\lambda<1$. Examples of unsafe behaviour occurring notably for the cone-type jet algorithms, which partially are still employed at the Tevatron [176, 180, 181], are shown in Fig. 2.20. The first row demonstrates an issue in cone algorithms with energy (or p_{T}) thresholds for the seeds of the cone finding procedure. The two in terms of energy flow equivalent situations are distinguished by the number of reconstructed jets. The collinear splitting of a four-momentum leads to the disappearance of a jet (right plot) as compared to the left panel. In the second row, the emission of a soft gluon leads to the merging of two jets into one, spoiling the cancellation of divergences in the virtual corrections (left configuration) against the ones in real corrections (right configuration). A safe cone algorithm, SISCone for *seedless infrared-safe cone*, only exists since 2007 [182].

Longitudinal boost invariance and boundary stability are ensured by using the rapidity y instead of the pseudorapidity for the jet kinematics and by combining two clustering objects through four-vector addition. The combination prescription is called *recombination scheme* (RS), for which other possibilities have been employed in the past [175], but have been abandoned by now. The boundary stability is essential for the applicability of soft gluon resummation to stabilise fixed-order pQCD predictions near exclusive phase space boundaries. As an aside, collinear and infrared safety are necessary, but not sufficient conditions for finite predictions order by order. Unsmooth behaviour of a "safe" observable inside its allowed range, caused for example by a change in phase space limitations when going from an n- to an $n+1$-parton final state, can lead to infinities at these internal borders unless the calculation is complemented with soft gluon resummation [183].

Finally, jet algorithms are designed *per se* to delimit the impact of the non-perturbative hadronisation phase by collecting within one jet hadrons spread out in p_{T} by $\approx \Lambda_{\mathrm{QCD}}$ relative to the jet axis. Obviously, the amount of leakage, the so-called *out-of-cone* effect (OOC), depends on the cone size R. On the other hand, the algorithm itself plays a role as well as can be seen from Fig. 2.21, which shows hadron-jet associations for a 3-jet event in e^+e^- collisions [184]. The same set of four-momenta is differently distributed over the three jets when using the JADE algorithm [173] (left) as compared to the k_{t} algorithm [185] (right). In particular, the JADE algorithm clusters soft particles together into the third jet although they are going into opposite directions, while the k_{t} algorithm is designed to undo perturbative parton splittings and prefers small angle over small mass clusterings. In effect it is found that the k_{t} algorithm is better behaved than JADE in terms of hadronisation corrections and resummability of large logarithms at small values of the resolution.

A detailed comparison of the impact of the jet radius R on the size of NP corrections to jet p_{T} cross sections has been performed in a collinear approximation in

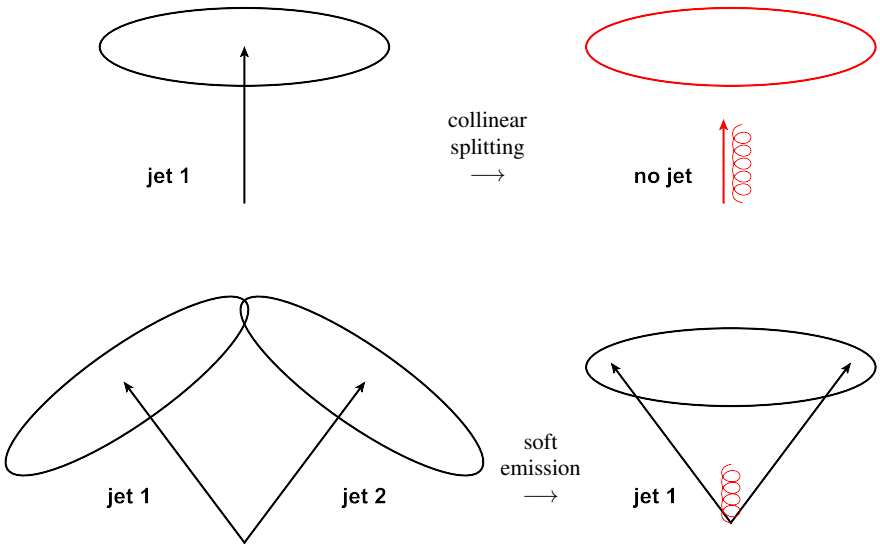

Fig. 2.20 Examples of collinear- and infrared-unsafe behaviour. Upper row: The collinear splitting of a jet seed reduces the seed p_{T}'s below threshold. Lower row: The emission of a soft gluon leads to the merging of two jets into one.

Fig. 2.21 A 3-jet final state in e^+e^- collisions as seen by the JADE (left) and k_{t} (right) jet algorithms. The particle assignments to the first, second, and third jet according to the algorithms are indicated by full, dash-dotted, and dashed lines. *(Redrawn from Ref. [184])*

Ref. [186]. As shown in Fig. 2.22 it was found that perturbative radiation, hadronisation, and the underlying event affect jet transverse momenta for small R roughly proportional to $\ln R$, $-1/R$, and R^2 respectively. As a consequence, the value of the jet radius parameter R determines which aspects of jet formation are emphasised and is a matter of choice.

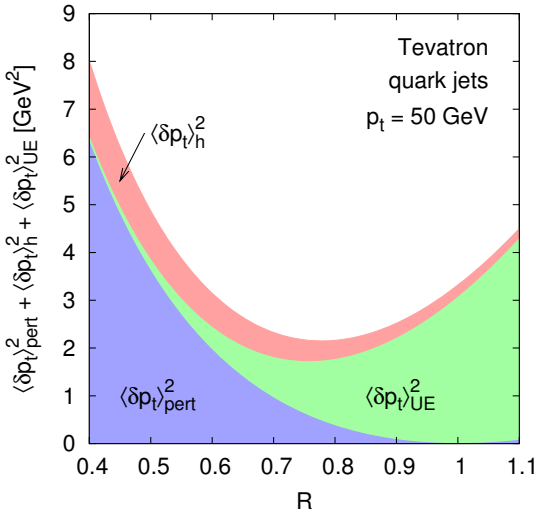

Fig. 2.22 Estimation of the squared average shift in p_T of a jet caused by perturbative radiation (pert), hadronisation (h), and soft particle production from the underlying event (UE) as a function of the jet cone size R. The effects are estimated in a collinear approximation for quark jets at Tevatron energies. *(Taken from Ref. [186])*

2.5.3 Experimental Desiderata

- Detector independence: No algorithmic dependence on detector details;
- Computational efficiency and predictability: predictable computing times that only mildly increase with growing numbers of input objects;
- Maximal reconstruction efficiency: lossless treatment of all input objects (no "dark jets");
- Insensitivity to pile-up collisions: accurate correction for additional energy not coming from the primary interaction;
- Ease of calibration: accurate and straightforward estimation of diverse detector effects on the jet response.
- Minimal resolution smearing and angular biasing: avoidance of algorithmic distortions in addition to detector effects;

Early versions of jet construction by the UA1 and UA2 experiments at the Spp̄S were based on the cell structure of their calorimeters in a way making it difficult to compare consistently to theory or other experiments [187–189]. This could be remedied by using cone algorithms, which as an added bonus are computationally efficient even for a large number N of input objects. On the contrary, sequential-recombination algorithms like k_t, favoured in low multiplicity e^+e^- collisions, were believed to scale with N^3 in terms of computing time. It was discovered in Ref. [190] that much better implementations are possible where the N^3 dependence is drastically reduced to merely N^2 or $N\ln N$. As a consequence k_t type algorithms became

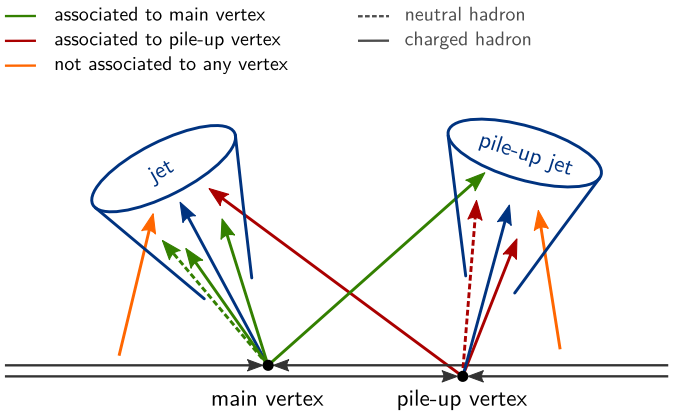

Fig. 2.23 Illustration of a pile-up collision in addition to a primary high-p_T interaction leading to additional particles and energy depositions within the same event. *(Illustration courtesy of J. Berger [192])*

even faster than cone-type ones and it could be envisioned to apply them in the experimental triggering procedure, where fast and predictable reconstruction times are a must.

An undesired feature related to the search for stable cones in some cone-type algorithms is the possibility to have so-called *dark jets*, i.e. measured energy depositions that are not reconstructed as jets [191]. This does not happen with sequential-recombination algorithms.

An experimental complication to jet measurements is caused by the quest for extremely rare processes like Higgs boson or very high-p_T jet production that demand correspondingly high instantaneous luminosities to be provided by a collider. Such luminosities cannot be achieved without piling up multiple proton-proton collisions per bunch crossing. In addition, the finite integration time of detector components in comparison to bunch separations of 50 or 25 ns (from 2015 onwards) leads to crosstalk from adjacent colliding bunches. As a result, each "event" is interspersed with energy depositions from such pile-up collisions (PU), which need to be subtracted. At low p_T even complete jets made of PU energy might be produced as illustrated in Fig. 2.23. An event-by-event identification of PU contributions is possible only for charged particles, which appear as reconstructed tracks that are not associated to the primary vertex of the event.

Jet algorithms that are insensitive to PU do not exist. Instead PU particles and energy are collected roughly proportional to the jet radius R^2 in a similar fashion as for the UE. Since a complete subtraction of PU is impossible, it is advantageous to have a clear idea of the *jet area A*. Although not always exactly correct, for cone-type algorithms the area is estimated to be πR^2, at least for the leading jets. For the k_t algorithm it is much more involved. However, as first proposed in Ref. [193] it is possible to define a jet area for each collinear- and infrared-safe jet algorithm by clustering in addition to the normal list of input objects large numbers of so-called

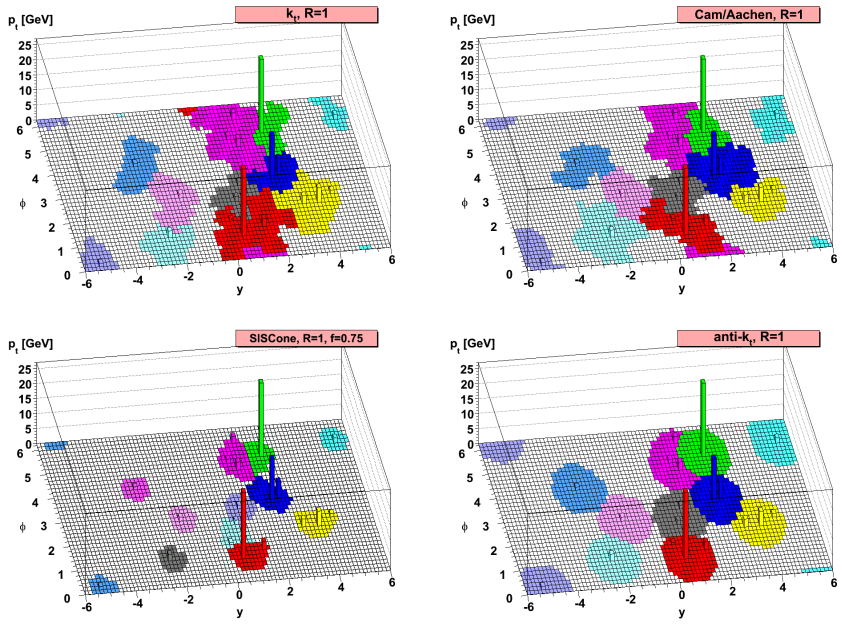

Fig. 2.24 Illustration of jet areas for the k_t, Cambridge/Aachen, SISCone, and anti-k_t jet algorithms. The tiles in rapidity and azimuth are coloured as a function of the number of ghosts clustered into a particular jet. *(Taken from Ref. [193])*

ghost particles that have negligible four-momentum and are uniformly distributed in rapidity and azimuth. Figure 2.24 presents as example for one hadron-hadron collision event the jet areas so defined for the k_t [194], Cambridge/Aachen [195, 196], SISCone [182], and anti-k_t [197] jet algorithms. The exact definition for the three sequential-recombination algorithms is given in the next Section 2.5.4. The k_t algorithm first combines low-p_T objects and leads to larger, irregular-shaped areas, while anti-k_t starts clustering with the highest-p_T objects and produces round-shaped jet areas as if from a cone jet algorithm. The Cambridge/Aachen algorithm solely relies on angular distances and lies somewhere in the middle of the other two.

With respect to PU subtraction and ease of calibration a regular shape of the jet area is favourable, because it simplifies the experimental evaluation of average densities (per area) to implement corrections as described in Chapter 3. In combination with perturbative safety, fast computing speed, and good geometrical resolution characteristics, the cone-like sequential-recombination algorithm anti-k_t was therefore quickly adopted as the standard jet algorithm at the LHC.

2.5.4 Sequential Recombination Algorithms

The sequential-recombination algorithms in use at the LHC can be described in a unified way. First of all two distance measures must be defined that give the relative distance d_{ij} between each pair of N input objects i and j and between each input object i and the beam direction, d_{iB}:

$$d_{ij} = \min\left(p_{\mathrm{T},i}^{2p}, p_{\mathrm{T},j}^{2p}\right)\frac{\Delta R_{ij}^2}{R^2}\,, \tag{2.28}$$

$$d_{iB} = p_{\mathrm{T},i}^{2p}\,. \tag{2.29}$$

Here, ΔR_{ij} is the purely "angular" distance in y and ϕ between i and j,[11]

$$(\Delta R_{ij})^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2\,, \tag{2.30}$$

which is scaled with respect to a parameter R that is equivalent to a cone radius and defines what will be referred to as *jet size* in the rest of this work. The power p decides whether low- or high-p_{T} objects are clustered first and differentiates between the k_{t} ($p = 1$), the Cambridge/Aachen ($p = 0$), and the anti-k_{t} ($p = -1$) jet algorithms.

Then, the minimal distance $d_{\min}$ is found from all pair-wise distances d_{ij} and all beam distances d_{iB} simultaneously. If $d_{\min} \in \{d_{iB}\}$, then the input object is declared a final jet, removed from the list of clustering objects, and added to the list of F jets, which initially is empty. If $d_{\min} \in \{d_{ij}\}$, then the two objects are merged with four-vector addition as recombination scheme and the new object is added to the clustering list while the parent objects i and j are removed. These steps are repeated until there are no more objects left in the clustering list. The list of final jets then is returned as result. An illustrative flowchart of the described algorithms is displayed in Fig. 2.25.

The default jet sizes chosen in LHC Run 1 by the two omni-purpose experiments for the anti-k_{t} algorithm are $R = 0.4$ and 0.6 for ATLAS and $R = 0.5$ and 0.7 for CMS. For Run 2, CMS changed the smaller jet size to 0.4 enabling direct comparisons to ATLAS measurements.

If the substructure of a jet is of interest, for example because the jet is supposed to come from a heavy boosted object whose two- or three-prong decay products all end up in the same jet, then it might be unwise to enforce a round-shaped jet area. For such purposes other jet types are better suited and in particular the Cambridge/Aachen algorithm is widely used. Last but not least, jet areas itself can be exploited for measurements, for example as an alternative to the traditional method to determine the UE [198].

[11] Of course, azimuthal angular separations are delimited to the interval $(-\pi, +\pi]$.

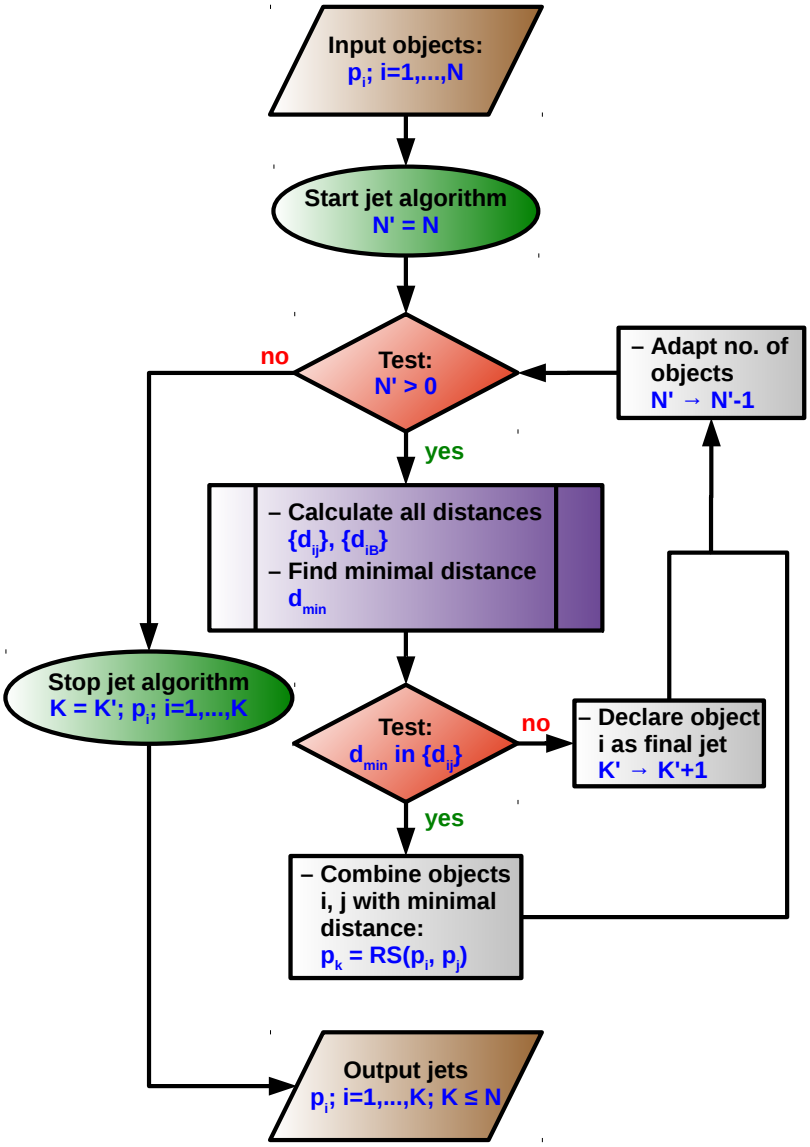

Fig. 2.25 Flowchart of a jet algorithm with sequential recombination. N input objects, usually given in the form of four-momenta p_i, are iteratively clustered following a recombination scheme RS. The final output is a list of K jets where $K \leq N$.

2.6 Theoretical Uncertainties

Given the amount of approximations and assumptions necessary before arriving at a prediction comparable to experimental data, a careful assessment of related uncertainties is mandatory. The most common theoretical uncertainties encountered in jet physics will be addressed in the next sections. They are related to:

1. the truncation of the perturbative series in a fixed-order calculation, colloquially but inadequately referred to as *scale uncertainty*;
2. the limited knowledge of hadron structure, *PDF uncertainty*;
3. the limited knowledge of the strong coupling constant, α_S *uncertainty*;
4. the modelling and tuning of non-perturbative effects, *NP uncertainty*.

Depending on the observable under study and the employed theoretical tools and techniques, further causes such as the fragmentation of heavy quarks, the top mass, decay constants, or limited statistical precision in numerical approximations may give rise to further uncertainties, cf. also Chapter 9 of Ref. [199].

2.6.1 Scale Uncertainties

Because the perturbative expansion in the strong coupling constant α_S needs to be truncated, all fixed-order calculations suffer from the fact that the missing higher orders somehow have to be accounted for in an uncertainty. Apparently, a truncated expansion following Eqs. (2.14)-(2.16) depends on the more or less arbitrary choices of renormalisation scheme and scale μ_r, while the full result of a renormalisable theory does not. More precisely, it can be shown that a quantity expanded in α_S up to terms of power n varies with $\ln \mu_r^2$ proportional to α_S^{n+1}, i.e. one order higher than the expansion itself:

$$\frac{\mathrm{d}\mathcal{X}}{\mathrm{d}\ln \mu_r^2}\left(\alpha_S^n\right) = \mathcal{O}\left(\alpha_S^{n+1}\right) . \tag{2.31}$$

Naturally, this lead to the paradigm of estimating missing higher orders through this residual scale dependence. The differences between the central and varied results when changing the scale μ_r around a central value μ_0 by factors of $1/r$ and r are taken to be the scale uncertainty. However, there are a couple of questions related to this *ad hoc* procedure that, nevertheless, is in widespread use:

- What renormalisation scheme to choose?
- What should the central scale μ_0 be?
- Which factor r should be used for the variation?
- What is the statistical interpretation of this uncertainty?
- What should be done, if multiple choices for μ_0 are involved?
- What should be done with other nonphysical scales like μ_f appearing in calculations of pQCD?

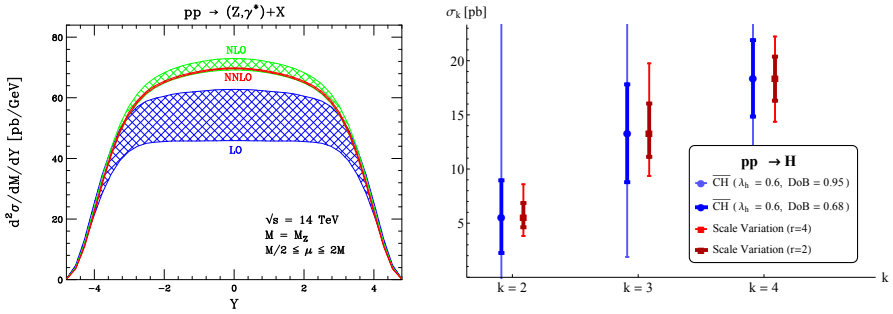

Fig. 2.26 Issues in scale variations: Left: Predictions at LO, NLO, and NNLO including scale uncertainty bands for Z boson production as a function of the Z rapidity Y. Right: Higgs production cross section $pp \to H$ via gluon fusion process at LO ($k = 2$), NLO ($k = 3$), and NNLO ($k = 4$). The light- and dark-blue error bars corresponding to the 1σ and 2σ degrees of belief estimated according to the modified Cacciari–Houdeau approach ($\overline{\text{CH}}$) are compared to the estimates using scale variations by factors of 2 and 4 shown as light- and dark-red error bars. *(Taken from Refs. [201, 203])*

Nowadays adopted standard in terms of renormalisation is the $\overline{\text{MS}}$ scheme, which is chosen for computational simplicity when dealing with one-loop corrections. In this scheme, the central scale μ_0 preferably should be of the order of the "hard" scale of the considered process to improve the accuracy of the perturbative expansion, cf. Ref. [45]. Conventionally, the variation factor r is chosen to be 2, which in many cases leads to reasonable results, see for example Ref. [200] for an examination of single-inclusive jet production. Figure 2.26 left shows the rapidity distribution of Z boson production as predicted to LO, NLO, and NNLO including scale uncertainty bands, where, on the contrary, the NLO prediction lies outside the estimated LO band. If the uncertainty is interpreted in the usual sense of a 1σ or 67% confidence level (CL), one such example might still be fine. However there are much worse cases in the same Ref. [201] and elsewhere. A review and comparison to a more rigorous statistical interpretation in terms of Bayesian degrees of belief (DoB) has been presented in Refs. [202, 203]. Figure 2.26 right presents the example of Higgs boson production via the gluon fusion process. Apparently, the computed higher orders at NLO and NNLO lie beyond the estimated uncertainties for $r = 2$ at lower order, and the uncertainties are largely underestimated even for $r = 4$ in comparison to the DoB as derived in the modified Cacciari–Houdeau approach ($\overline{\text{CH}}$).

The multi-scale problem partially is related to the merging and matching schemes of the NLO+PS event generation mentioned in Section 2.4.2. Discussions with respect to the involved merging/matching scale μ_m and corresponding uncertainties are still ongoing. The factorisation scale μ_f introduced in Section 2.3.2 on the other hand has a totally different physics origin from μ_r and can, in principal, be set independently. For jet production in ep DIS for example, μ_f usually is identified with Q, i.e. the momentum transfer between the scattered electron and the proton, while μ_r is chosen to be the jet p_{T}. At the LHC, μ_r and μ_f are mostly defined identically, but scale variations are performed independently avoiding overly large relative factors

between the two, cf. Refs. [204, 205]:

$$1/r \leq \mu_r/\mu_f \leq r\,. \tag{2.32}$$

With $r = 2$ this leads to the following six variations of μ_r and μ_f from the default choice of $\mu_r = \mu_f = \mu_0$ between $\mu_0/2$ and $2\mu_0$: $(\mu_r/\mu_0, \mu_f/\mu_0) = (1/2, 1/2)$, $(1/2, 1)$, $(1, 1/2)$, $(1, 2)$, $(2, 1)$, and $(2, 2)$. The maximal downwards and upwards deviations from the central result are defined as scale uncertainty, where the dominant effect usually comes from the change in μ_r. Despite the above-mentioned issues the majority of experimental and theoretical particle-physics analyses in the past and still today employ this recipe, or a simplified variant, of scale uncertainties, which although far from optimal at least induce comparability. It is well-known though that scale variations anyway are incapable to account for the appearance of new production channels at the next perturbative order.

Other concepts of scale setting, *fastest apparent convergence* (FAC) [206, 207], *principle of minimal sensitivity* (PMS) [208, 209] , and *Brodsky–Lepage–Mackenzie* (BLM) [210], have been investigated early on. The latter claims to eliminate scale ambiguities in pQCD, but requires to find a proper scale that is different order-by-order for each relevant set of subprocesses and hence is rather impractical. A recent revival of interest in the BLM prescription lead to a reformulation in the form of the *principle of maximum conformality* (PMC) [211–213]. PMC promises to be a systematic method to eliminate the renormalisation scale and scheme ambiguities to all orders in pQCD. A widespread application of PMC is still outstanding, but taken at face value it emphasises once more that scale variations cannot really address the original aim of estimating a *missing higher order uncertainty*. Some progress in that direction has been reported in Refs. [214, 215].

2.6.2 PDF Uncertainties

Initially, the PDFs, necessary for the factorised, long-distance part in Eq. (2.20), were derived from in particular DIS experiments in the form of parameterised functions. The observation of an excess of jet production at high p_{T} by the CDF Collaboration in 1996 [216], cf. Fig. 2.27 left, triggered speculations with respect to new phenomena. However, at that time, the only means to exploit the potential freedom in these PDFs consisted in comparing the predictions of a few candidate functions like the ones used by the CDF Collaboration. A systematic approach to PDF uncertainties simply did not exist. In the end, the excess could be accommodated by adaptations in the gluon PDF [217] as shown in Fig. 2.27 right.

In the following 20 years enormous progress has been made towards statistically well-founded, systematic schemes to determine PDFs including uncertainty estimates. Numerous PDF fitting groups continuously develop and refine their methods, while simultaneously more and more precise data are integrated into the process. Figure 2.28 left exemplifies the MSTW2008 NLO PDFs including uncertainty

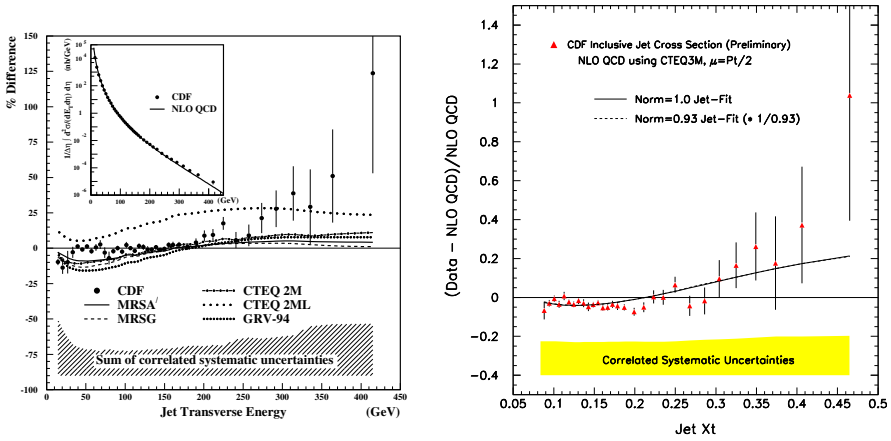

Fig. 2.27 Excess reported by CDF with respect to the predicted inclusive jet p_T spectrum (left) and explanation in terms of a modified gluon PDF (right). *(Taken from Refs. [216, 217])*

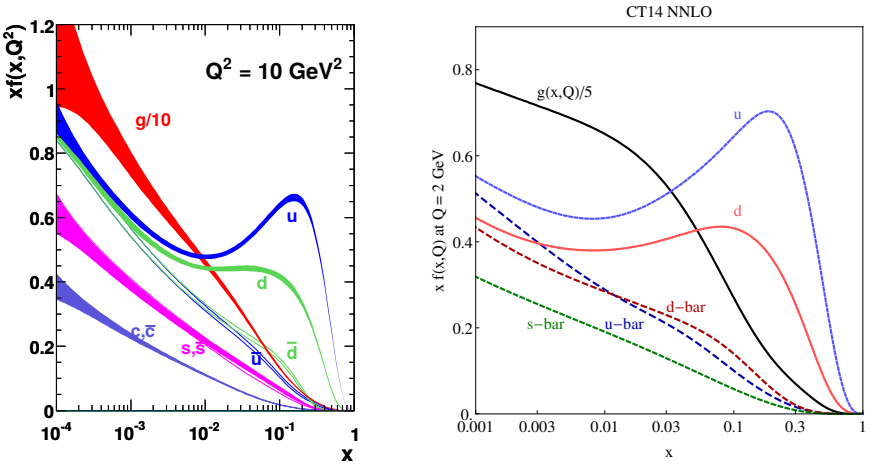

Fig. 2.28 MSTW2008 NLO PDFs including uncertainty bands at a scale of $Q^2 = 10\,\mathrm{GeV}^2$ (left) and shape of the recent CT14 NNLO PDFs at $Q^2 = 4\,\mathrm{GeV}^2$ (right). *(Taken from Refs. [218, 219])*

bands at a scale of $Q^2 = 10\,\mathrm{GeV}^2$. Their shape at higher scales like $10000\,\mathrm{GeV}^2$ as needed at the LHC is completely determined by the DGLAP equations (2.21). An example of the most recent PDFs, displayed without uncertainties, is given in Fig. 2.28 right.

Table 2.1 gives an overview of the primary PDF sets available for comparisons to data at NLO. In most cases, NNLO versions and further variants exist as well. As exhibited in Fig. 2.28 the validity of the DGLAP equations is assumed starting from scales Q even below threshold for the production of b quarks.[12] The inclusion

[12] $m^{\overline{\mathrm{MS}}}_{\mathrm{charm}} = 1.275\,\mathrm{GeV}$, $m^{\overline{\mathrm{MS}}}_{\mathrm{bottom}} = 4.18\,\mathrm{GeV}$ [28].

Table 2.1 The NLO PDF sets used in comparisons to the data in Run 1 (upper rows) and newer sets for Run 2 (lower rows) together with the corresponding number of active flavours N_f, the assumed masses M_t and M_Z of the top quark and the Z boson, respectively, the default values of $\alpha_S(M_Z)$, and the range in $\alpha_S(M_Z)$ variation available for fits. A * behind the $\alpha_S(M_Z)$ values signifies that the parameter was fixed, not fitted. The FFNS (G)JR PDF sets use $N_F = 3$, but allow N_F to rise up to five in the evolution of α_S. VFNS variants of these PDFs exist as well.

Base set	Refs.	N_F	M_t (GeV)	M_Z (GeV)	$\alpha_S(M_Z)$	$\alpha_S(M_Z)$ range
ABM11	[225]	5	180	91.174	0.1180	0.110–0.130
CJ12	[226]	$\leq$5	180	91.1876	0.1180*	—
CT10	[227]	$\leq$5	172	91.188	0.1180*	0.112–0.127
HERAPDF1.5	[228]	$\leq$5	180	91.187	0.1176*	0.114–0.122
GJR08	[229,230]	3(5)	175	91.71	0.1145	—
MSTW2008	[218,231]	$\leq$5	10^{10}	91.1876	0.1202	0.110–0.130
NNPDF2.1	[232]	$\leq$6	175	91.2	0.1190*	0.114–0.124
NNPDF2.3	[233]	$\leq$6	175	91.1876	0.1180*	0.114–0.124
CJ15	[234]	$\leq$5	180	91.1876	0.1180*	—
CT14	[219]	$\leq$5	172	91.1876	0.1180*	0.113–0.123
HERAPDF2.0	[235]	$\leq$5	173	91.1876	0.1180*	0.110–0.130
JR14	[236]	3(5)	173	91.1876	0.1158	—
MMHT2014	[237]	$\leq$5	10^{10}	91.1876	0.1180*	0.108–0.128
NNPDF3.0	[238]	$\leq$5	173	91.2	0.1180*	0.115–0.121

of the heavy quarks, charm and bottom, into pQCD predictions involving PDFs requires special care. Several solutions are proposed. The ABM and (G)JR PDF sets employ a fixed-flavour number scheme (FFNS) [220–222] with three respectively five flavours. All other PDF sets use a general-mass variable-flavour number scheme (GM-VFNS), several variants of which are reviewed in Ref. [223]. The maximum number of flavours is five, $N_{f,\max} = 5$, except for NNPDF2.1 and 2.3, which have $N_{f,\max} = 6$. Even beyond the top quark mass it can be justified to continue using $N_F = 5$ [224] as is the case for the jet analyses presented later in this work.

The PDF sets are supplied for usage in other programs via the unified Les Houches Accord PDF (LHAPDF) interface [239, 240], which in its latest version includes routines to evaluate uncertainties. Two primary strategies evolved to provide these uncertainty estimates from PDF fits:

1. The *Hessian* or *eigenvector method* [241, 242] where the correlated experimental data uncertainties on the fitted parameters for simplicity are provided within an orthonormalised parameter space.
2. The *MC method* [243–245] where the input data are varied according to their precision, including correlations, to produce an ensemble of PDFs that statistically reflects the intrinsic uncertainty.

Since each eigenvector of the Hessian method corresponds to an uncertainty source independent of all other eigenvectors, the total impact on an observable $\mathcal{X}$

can be evaluated by quadratic addition. If one-sided deviations caused by the + and − variations of one eigenvector are to be considered in an asymmetric uncertainty, the formula reads:

$$\Delta\mathcal{X}^{+} = \sqrt{\sum_{i=1}^{N_{\mathrm{EV}}} \left[\max\left(\mathcal{X}_i^{+} - \mathcal{X}_0, \mathcal{X}_i^{-} - \mathcal{X}_0, 0\right)\right]^2},$$
$$\Delta\mathcal{X}^{-} = \sqrt{\sum_{i=1}^{N_{\mathrm{EV}}} \left[\min\left(\mathcal{X}_i^{+} - \mathcal{X}_0, \mathcal{X}_i^{-} - \mathcal{X}_0, 0\right)\right]^2}, \tag{2.33}$$

where N_{EV} is the number of eigenvectors and the index 0 indicates the result for the central PDF stored as zeroth member. The CJ, CT, MSTW, and MMHT sets fall into this category and consist of this central PDF member no. 0 and $2 \cdot N_{\mathrm{EV}}$ members for the uncertainty. The PDF uncertainties customarily are evaluated at 68% confidence level (CL) except for the CT and CJ ones, which provide uncertainties only at 90% CL. For a uniform treatment, the CT/CJ uncertainties must therefore be downscaled by a factor of $\sqrt{2}\,\mathrm{erf}^{-1}(0.9) \approx 1.645$. If symmetric errors are required, for example in covariance matrices, the formula

$$\Delta\mathcal{X}^{\pm} = \pm\sqrt{\sum_{i=1}^{N_{\mathrm{EV}}} \left[\frac{\mathcal{X}_i^{+} - \mathcal{X}_i^{-}}{2}\right]^2} \tag{2.34}$$

can be used instead. In case of the ABM and (G)JR PDF sets a symmetrisation has been performed beforehand such that in their case the Hessian uncertainties are to be derived according to the equation

$$\Delta\mathcal{X}^{\pm} = \pm\sqrt{\sum_{i=1}^{N_{\mathrm{mem}}} \left[\mathcal{X}_i - \mathcal{X}_0\right]^2}, \tag{2.35}$$

where N_{mem} is the total number of members not counting the zeroth one.

For the MC method as employed by the NNPDF sets, the usual formula for a standard deviation is applicable:

$$\Delta\mathcal{X}^{\pm} = \sqrt{\frac{1}{N_{\mathrm{rep}} - 1} \cdot \sum_{i=1}^{N_{\mathrm{rep}}} \left[\mathcal{X}_i - \langle\mathcal{X}\rangle\right]^2}, \tag{2.36}$$

where $N_{\mathrm{rep}} = N_{\mathrm{mem}}$ is the number of replicas, again not counting the zeroth one. $\langle\mathcal{X}\rangle$ corresponds to the average prediction for the observable $\mathcal{X}$, which in general is different from the predicted value for the averaged PDFs present as zeroth member.

The PDF sets listed in Table 2.1 can be further differentiated into “global” ones that evaluate a multitude of different data sets in a global combined fit, ABM, CT, (G)JR, MSTW/MMHT, and NNPDF, and more specialised ones, CJ, HERAPDF.

Out of the global ones, the (G)JR PDF sets are less used, because they lack a series with variations in the value of $\alpha_S(M_Z)$.

The CJ sets follow the general strategy of the CT ones, but include data with nuclear targets, e.g. deuterium, which requires supplementary assumptions or *nuclear correction factors* when extracting proton PDFs. Their advantage consists in larger amounts of data that constrain the high x region, which is particularly important when searching for new phenomena at very high p_T or mass.

The HERAPDF sets are restricted to the exclusive use of DIS measurements from the H1 and ZEUS experiments. Since these data are essential ingredients to all other PDF sets as well, they pose an ideal baseline for comparisons. Moreover, the computer code for performing the fits is publically available in the form of the open source framework HERAFITTER [246].[13] The uncertainty estimation employed in this framework is more involved and is explained in detail in Section 4.5, where such a PDF fit including CMS jet data is described.

Standard model measurements, particularly the ones that promise to be valuable ingredients to fits of SM parameters, clearly should be confronted with predictions for each individual PDF set. In some cases, e.g. limit settings on physics beyond the SM, it might be desirable to provide only one number that in addition accounts for variations caused by the different PDF sets. For this purpose, the PDF4LHC working group has proposed a combination procedure [247] that was updated recently [248].

Effects of polarisation or *transverse momentum dependence* (TMD) in PDFs have been neglected so far. These topics are specially addressed in Refs. [45, 249, 250].

2.6.3 α_S Uncertainties

In addition to the parameters of the PDFs themselves the strong coupling constant enters into the fits. Correlations, specifically between the gluon PDF and the assumed value of $\alpha_S(M_Z)$, are expected and must be addressed when deriving a combined uncertainty. For this purpose, it must be differentiated whether $\alpha_S(M_Z)$ is considered a fixed input parameter, i.e. the "starred" values in Table 2.1, or a fitted output parameter in a PDF fit.
For the first case, three approaches are discussed in Ref. [251]:

- Quadratic addition of the PDF uncertainty and a $\Delta\alpha_S(M_Z)$ uncertainty computed with the same central PDF set.
- Quadratic addition of the PDF uncertainty and a $\Delta\alpha_S(M_Z)$ uncertainty computed with PDF sets fitted using the varied values of $\alpha_S(M_Z)$.
- Correlated propagation of the PDF and $\alpha_S(M_Z)$ uncertainty to the PDF error set.

The first recipe is totally ignorant of correlations, while the last one depends on the method for the estimation of PDF uncertainties, Hessian or MC method. For a

[13] HERAFITTER recently was renamed to xFITTER.

variation of $\alpha_S(M_Z)$ by ± 0.0012 as recommended in Ref. [247] it is observed that all three recipes are rather close to each other with a slight underestimation of the uncertainty by variants one and two. For all practical purposes it is sufficient and recommended to apply the second recipe for the CT and NNPDF sets [252, 253].

In the second case where $\alpha_S(M_Z)$ was fitted together with the PDF parameters, the α_S uncertainty intrinsically is included already in the ABM PDF uncertainties. For the MSTW2008 PDF set, it is recommended to calculate the PDF+$\alpha_S(M_Z)$ uncertainty from the envelope of five PDF sets with uncertainties produced under varying assumptions on $\alpha_S(M_Z)$ [231].

The newer CT, MMHT, and NNPDF sets all employ the same fixed value of $\alpha_S(M_Z) = 0.1180$. Here, the recommended procedure follows recipe two above, but with an enlarged uncertainty of $\Delta\alpha_S(M_Z) = \pm 0.0015$.
The latest 2015 update of the PDG gives the world average as:

$$\alpha_S(M_Z) = 0.1181 \pm 0.0013\,. \tag{2.37}$$

2.6.4 Non-perturbative Uncertainties

Since fixed-order predictions as used in fits of SM parameters are at parton-level only, corrections for non-perturbative effects must be applied. As described in Section 2.4.1 this implies model assumptions and parameter tuning and hence induces a corresponding uncertainty. Customarily, the corrections are estimated from the ratio of distributions for fully hadronised events over the distributions with MPI and hadronisation switched off in the respective LO+PS MC event generators. An envelope is constructed around the predictions by different event generators with various tunes to derive a medium correction and to attribute a systematic uncertainty of half-width of the spread for this factor.

The advent of MC event generators capable of combining NLO+PS with the NP modelling opens up new possibilities. A first step towards NP corrections from NLO+PS event generation has been taken in Ref. [254], cf. Fig. 2.29 left, where the results for LO+PS from HERWIG++ and PYTHIA8 are compared with those for NLO+PS from POWHEG +PYTHIA8. As expected the NP effects are negligible for high transverse momenta, but can become significant for jet p_{T}'s below 300 GeV. Interestingly, they are less pronounced in the NLO+PS case. However, it has to be kept in mind that the p_{T} dependence of the various effects of PS, MPI, and hadronisation differs as detailed in Section 2.4. As a consequence, the NP corrections are sensitive to the choice of jet algorithm and jet size.

To avoid statistical fluctuations in less populated regions of phase space, the NP factors are usually parameterised by a simple polynomial function:

$$f(x) = p_0 + \frac{p_1}{x^{p_2}}\,. \tag{2.38}$$

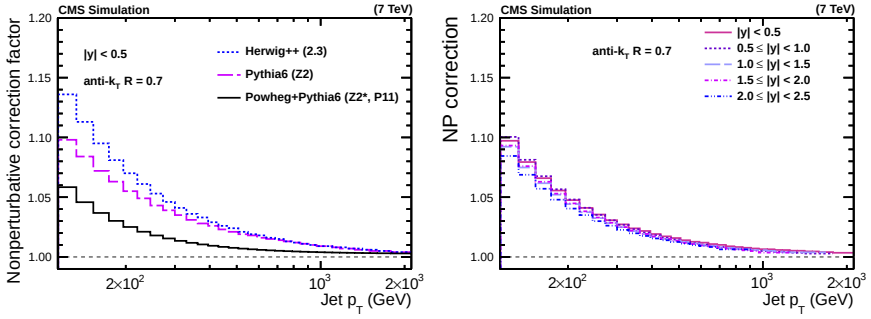

Fig. 2.29 Left: NP corrections at central rapidity $|y| < 0.5$ for inclusive jets as a function of jet p_{T} using two LO+PS and the NLO+PS MC event generator POWHEG +PYTHIA8. Right: Comparison of the combined NP factors in five regions of absolute jet rapidity. *(Taken from Ref. [254])*

Combining the three parameterised predictions for the whole range in rapidity leads to the curves shown in Fig. 2.29 right, exhibiting a similar size and shape versus p_{T}. Details with respect to particular analyses are presented in the relevant sections as appropriate.

2.6.5 Fast Interpolation Techniques

The processes of inclusive jet and dijet production at hadron colliders are known up to NLO since more than two decades from the (M)EKS [255–257], JETRAD [258], and NLOJET++ [259, 260] programs, where the latter also comprises 3-jet production at NLO. A calculation at NNLO is ongoing and at the time of writing is partially completed [261]. 4-jet and 5-jet production at NLO have become available recently [262–264].

All these programs, capable of providing cross sections fully differential in jet observables, are increasingly demanding in terms of CPU time consumption with higher orders and multiplicities. This prevents their direct use in fits of PDF parameters or the strong coupling constant. However, the CPU intensive part entering the calculation according to Eq. (2.20) is localised in the derivation of the partonic cross section

$$\mathrm{d}\hat{\sigma}_{(ij\to X)}(x,x',\mu_f,\mu_r,\alpha_S(\mu_r)) \tag{2.39}$$

that depends on the PDFs and $\alpha_S(\mu_r)$ only indirectly via folding integrals. This fact enables the use of interpolation techniques to separate the time-consuming part from the PDF and $\alpha_S(\mu_r)$ dependence. To factorise for example the PDF dependence from the convolution, the PDFs are expressed as a sum of eigenfunctions (or interpolation kernels) on a grid in the convolution variable x as

$$f_{i/p}(x) \approx \sum_a f_{i/p}(x_a) \cdot E_a(x)\,, \tag{2.40}$$

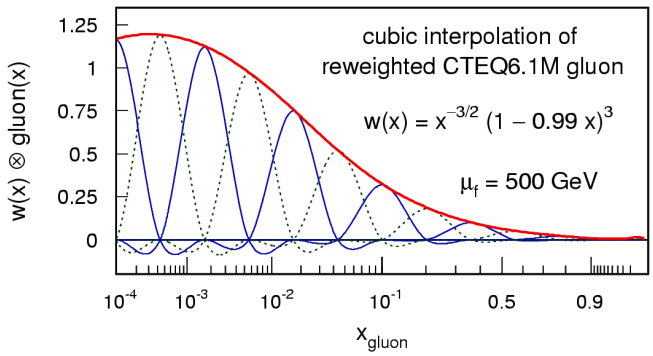

Fig. 2.30 Cubic interpolation of the gluon PDF in FASTNLO. $w(x)$ is an additional weight function to reduce PDF curvature and improve the approximation quality. *(Illustration courtesy of M. Wobisch)*

where a numbers the support nodes x_a, $f_{i/p}(x_a)$ assumes a fixed value, and $E_a(x)$ is the a-th eigenfunction. The set of eigenfunctions $\{E_a(x)\}$ must satisfy the relations

$$\begin{aligned} E_a(x_b) &= \delta_{ab}\,, \quad \forall a,b \\ \sum_a E_a(x) &= 1\,, \quad \forall x\,. \end{aligned} \tag{2.41}$$

A pictorial representation of this procedure employing cubic interpolation kernels is shown in Fig. 2.30.

Replacing $f_{i/p}(x)$ in Eq. (2.20) by the sum of Eq. (2.40) and performing the convolution with the partonic cross section once, the result can be written as a sum of terms, where a change in PDF only requires to adapt the factors $f_{i/p}(x_a)$, which is a very fast operation. For hadron-hadron collisions the same technique can easily be extended to both PDFs $f_{i/p}(x)$ and $f_{j/p}(x')$ giving a two-dimensional grid of support nodes (x_a,x_b) with PDF specific factors $f_{i/p}(x_a)$ and $f_{j/p}(x_b)$.

The quality of this approximation is determined by the number and distribution of support points, and the shape of the interpolation kernels in comparison to the shape of the PDFs and the distribution of the partonic cross section over the phase space. Two possibilities for optimisation are indicated in Fig. 2.30:

1. The support nodes are not evenly distributed in x but proportional to $\sqrt{(}\log_{10}(x))$.
2. An additional weight function $w(x)$ has been applied to reduce in a generic way the curvature of the PDFs, which improves the approximation quality.

Finally, the dependence on μ_r and μ_f can be addressed in the same way as the interpolation in x, where, however, the interpolation kernels and parameters can be optimised separately. In total, the equivalent of Eq. (2.20) then reads

$$\mathrm{d}\sigma_{(pp\to X)} = \sum_{n,i,j,a,b,c} \alpha_s^n(\mu_c) \cdot f_{i/p}(x_a,\mu_c) \cdot f_{j/p}(x_b,\mu_c) \cdot \mathrm{d}\tilde{\sigma}^{(ij\to X)}_{n,i,j,a,b,c}\,, \tag{2.42}$$

where the sum runs over the powers n of α_S, the parton flavours i and j, the x-interpolation nodes a and b, and the μ-interpolation nodes c. The time-consuming ($\mathcal{O}(kh)$) convolution has to be performed only once to compute and tabulate the coefficients

$$\mathrm{d}\tilde{\sigma}^{(ij\to X)}_{n,i,j,a,b,c} = \int \mathrm{d}x\mathrm{d}x' E_a(x)\cdot E_b(x')\cdot E_c(\mu)\times \mathrm{d}\hat{\sigma}^n_{(ij\to X)}(x,x',\mu)\,, \tag{2.43}$$

that fully contain the information on observable definition and phase space restrictions. The evaluation of the sum of products in Eq. (2.42) only requires milliseconds.

This technique of *fast interpolation grids* introduced above hence can be exploited to reduce by many orders of magnitude the amount of CPU time to invest for repeated evaluations of higher-order cross sections with modified assumptions on PDFs, scales μ, or the value and evolution of α_S. Two independent implementations of this technique exist: FASTNLO [265, 266] and APPLGRID [267, 268], both of which are interfaced to various theory programs and utilised in fits with the XFITTER framework. The packages differ in their interpolation and optimisation strategies, but both of them construct tables with grids for each bin of an observable in two steps. In a first step the accessible phase space in momentum fraction x and scale μ is explored and the partonic subprocesses that contribute and need to be differentiated must be identified. This serves to optimise the table size by eliminating empty regions in x and μ as well as reducing the number of subprocesses to store as compared to $11\cdot 11$ linear combinations in a full flavour basis without counting top quarks. Only in the second step the grid is filled for the requested observables and binning. The approximation quality can easily be checked by a comparison to histograms simultaneously filled with the full cross section. Alternatively, the dependence on the number of support nodes could be studied, as for an infinite number the approximation bias must approach zero.

For simplicity, μ_r and μ_f have been set equal in the formulae above. However, it is perfectly possible to differentiate between the two and in an update to the FASTNLO framework [269, 270] it is even possible to separately store terms proportional to $\log(\mu_r^2)^s\cdot\log(\mu_f^2)^t$ as they appear in the coefficients of a perturbative expansion. At NLO, the only combinations possible are $s=t=0$, $s=1,t=0$, and $s=0,t=1$.

References

1. M. G. Holloway and C. P. Baker, "Note on the origin of the term "barn"", technical report, Los Alamos National Laboratory, 1947. Report submitted: 13 September 1944. Report issued: 5 March 1947.
2. D. Haitz, "Precision Measurements of Proton Structure and Jet Energy Scale with the CMS Detector at the LHC". PhD thesis, KIT (Karlsruher Institut für Technologie), Jun, 2016.
3. J. Chadwick, "Possible Existence of a Neutron", *Nature* **129** (1932) 312, doi:10.1038/129312a0.
4. P. A. M. Dirac, "The quantum theory of the electron", *Proc. Roy. Soc. Lond. A* **117** (1928) 610, doi:10.1098/rspa.1928.0023.

5. C. D. Anderson, "The Positive Electron", *Phys. Rev.* **43** (1933) 491, doi:10.1103/PhysRev.43.491.
6. S. H. Neddermeyer and C. D. Anderson, "Note on the Nature of Cosmic Ray Particles", *Phys. Rev.* **51** (1937) 884, doi:10.1103/PhysRev.51.884.
7. H. Yukawa, "On the Interaction of Elementary Particles I", *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48, doi:10.1143/PTPS.1.1.
8. C. M. G. Lattes, G. P. S. Occhialini, and C. F. Powell, "Observations on the Tracks of Slow Mesons in Photographic Emulsions. 1", *Nature* **160** (1947) 453, doi:10.1038/160453a0.
9. C. M. G. Lattes, G. P. S. Occhialini, and C. F. Powell, "Observations on the Tracks of Slow Mesons in Photographic Emulsions. 2", *Nature* **160** (1947) 486, doi:10.1038/160486a0.
10. R. Cahn and G. Goldhaber, "The Experimental Foundations of Particle Physics". Cambridge University Press, Cambridge, 2nd edition, 2009.
11. V. E. Barnes et al., "Observation of a Hyperon with Strangeness -3", *Phys. Rev. Lett.* **12** (1964) 204, doi:10.1103/PhysRevLett.12.204.
12. M. Gell-Mann, "A Schematic Model of Baryons and Mesons", *Phys. Lett.* **8** (1964) 214, doi:10.1016/S0031-9163(64)92001-3.
13. G. Zweig, "An SU_3 model for strong interaction symmetry and its breaking; Version 1", technical report, CERN, Geneva, January, 1964.
14. G. Zweig, "An SU_3 model for strong interaction symmetry and its breaking; Version 2", technical report, CERN, Geneva, February, 1964.
15. M. Gell-Mann, "Quarks, Color, and QCD", in *Proceedings, Workshop on QCD — 20 Years Later*, volume C920609, p. 3. Aachen, Germany, June 9-13, 1992.
16. "Creative Commons License". CC-BY-SA-3.0, http://creativecommons.org/licenses/by-sa/3.0/.
17. E. D. Bloom et al., "High-Energy Inelastic *ep* Scattering at 6-Degrees and 10-Degrees", *Phys. Rev. Lett.* **23** (1969) 930, doi:10.1103/PhysRevLett.23.930.
18. M. Breidenbach et al., "Observed Behavior of Highly Inelastic electron-proton Scattering", *Phys. Rev. Lett.* **23** (1969) 935, doi:10.1103/PhysRevLett.23.935.
19. J. D. Bjorken, "Asymptotic Sum Rules at Infinite Momentum", *Phys. Rev.* **179** (1969) 1547, doi:10.1103/PhysRev.179.1547.
20. G. 't Hooft and M. J. G. Veltman, "Regularization and Renormalization of Gauge Fields", *Nucl. Phys. B* **44** (1972) 189, doi:10.1016/0550-3213(72)90279-9.
21. H. Fritzsch and M. Gell-Mann, "Current algebra: Quarks and what else?", in *Proceedings, 16th International Conference on High-Energy Physics (ICHEP)*, volume C720906V2, p. 135. Batavia, IL, USA, September 6-13, 1972. arXiv:hep-ph/0208010.
22. O. W. Greenberg, "Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons", *Phys. Rev. Lett.* **13** (1964) 598, doi:10.1103/PhysRevLett.13.598.
23. M. Gell-Mann, "Symmetries of baryons and mesons", *Phys. Rev.* **125** (1962) 1067, doi:10.1103/PhysRev.125.1067.
24. D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Nonabelian Gauge Theories", *Phys. Rev. Lett.* **30** (1973) 1343, doi:10.1103/PhysRevLett.30.1343.
25. D. J. Gross and F. Wilczek, "Asymptotically Free Gauge Theories. 1", *Phys. Rev. D* **8** (1973) 3633, doi:10.1103/PhysRevD.8.3633.
26. D. J. Gross and F. Wilczek, "Asymptotically Free Gauge Theories. 2.", *Phys. Rev. D* **9** (1974) 980, doi:10.1103/PhysRevD.9.980.
27. H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?", *Phys. Rev. Lett.* **30** (1973) 1346, doi:10.1103/PhysRevLett.30.1346.
28. K. A. Olive and others (Particle Data Group), "Review of Particle Physics", *Chin. Phys. C* **38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.
29. JabRef Development Team, "JabRef". http://www.jabref.org/, 2003. [Online; accessed 24-May-2016].

30. M. E. Peskin and D. V. Schroeder, "An Introduction To Quantum Field Theory (Frontiers in Physics)". Westview Press, Boulder, CO, October, 1995.
31. I. J. Aitchison and A. J. Hey, "Gauge Theories in Particle Physics: A Practical Introduction, Fourth Edition - 2 Volume set". CRC Press, 4 edition, 12, 2012.
32. G. F. Sterman, "An Introduction to quantum field theory". Cambridge University Press, 1993.
33. C. Itzykson and J.-B. Zuber, "Quantum Field Theory". Dover Publications, February, 2006.
34. J. D. Bjorken and S. D. Drell, "Relativistic Quantum Fields". Mcgraw-Hill College, first edition edition, June, 1965.
35. G. Dissertori, I. G. Knowles, and M. Schmelling, "Quantum Chromodynamics: High Energy Experiments and Theory". Oxford University Press, 2nd edition, May, 2009.
36. F. J. Ynduráin, "The Theory of Quark and Gluon Interactions". Texts and monographs in physics. Springer, Berlin, 4th edition, 2006.
37. R. K. Ellis, W. J. Stirling, and B. R. Webber, "QCD and Collider Physics". Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, Cambridge, 1996.
38. Y. L. Dokshitzer, V. A. Khoze, A. H. Mueller, and S. I. Troian, "Basics of Perturbative QCD". Editions Frontieres, December, 1991.
39. T. Muta, "Foundations Of Quantum Chromodynamics: An Introduction To Perturbative Methods In Gauge Theories". World Scientific, September, 1987.
40. CTEQ Collaboration, "Handbook of perturbative QCD", *Rev. Mod. Phys.* **67** (January, 1995) 157, doi:10.1103/RevModPhys.67.157.
41. D. W. Duke and R. G. Roberts, "Determinations of the QCD Strong Coupling α_s and the Scale $\Lambda_{\rm QCD}$", *Phys. Rept.* **120** (1985) 275, doi:10.1016/0370-1573(85)90112-7.
42. G. Altarelli, "Partons in Quantum Chromodynamics", *Phys. Rept.* **81** (1982) 1, doi:10.1016/0370-1573(82)90127-2.
43. A. H. Mueller, "Perturbative QCD at High-Energies", *Phys. Rept.* **73** (1981) 237, doi:10.1016/0370-1573(81)90030-2.
44. R. P. Feynman, "Photon-hadron Interactions". Westview Press, March, 1998.
45. J. Collins, "Foundations of Perturbative QCD". Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, Cambridge, 2011.
46. S. Moch, "Hard QCD at hadron colliders", *J. Phys. G* **35** (2008) 073001, doi:10.1088/0954-3899/35/7/073001, arXiv:0803.0457.
47. J. M. Campbell, J. W. Huston, and W. J. Stirling, "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics", *Rept. Prog. Phys.* **70** (2007) 89, doi:10.1088/0034-4885/70/1/R02, arXiv:hep-ph/0611148.
48. R. Alemany-Fernandez et al., "The Large Hadron Collider: Harvest of Run 1". Springer, Berlin, Germany, 1st edition, May, 2015.
49. R. M. Barnett et al., "Physics at the Terascale". Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 1st edition, April, 2011.
50. H. Fritzsch and M. Gell-Mann, eds., "50 Years of Quarks". World Scientific Publishing Co. Pte. Ltd., April, 2015.
51. P. M. Zerwas and H. A. Kastrup, eds., "QCD : 20 Years Later : Aachen, June 9-13, 1992". World Scientific Publishing Co. Pte. Ltd., June, 1993.
52. D. Binosi and L. Theussl, "JaxoDraw: A Graphical user interface for drawing Feynman diagrams", *Comput. Phys. Commun.* **161** (2004) 76, doi:10.1016/j.cpc.2004.05.001, arXiv:hep-ph/0309015.
53. L. D. Faddeev and V. N. Popov, "Feynman Diagrams for the Yang-Mills Field", *Phys. Lett. B* **25** (1967) 29, doi:10.1016/0370-2693(67)90067-6.
54. R. D. Peccei and H. R. Quinn, "CP Conservation in the Presence of Pseudoparticles", *Phys. Rev. Lett.* **38** (1977) 1440, doi:10.1103/PhysRevLett.38.1440.
55. R. D. Peccei and H. R. Quinn, "Constraints imposed by CP conservation in the presence of pseudoparticles", *Phys. Rev. D* **16** (1977) 1791, doi:10.1103/PhysRevD.16.1791.

56. G. 't Hooft, "The Glorious days of physics: Renormalization of gauge theories", arXiv:hep-th/9812203.
57. T. van Ritbergen, J. A. M. Vermaseren, and S. A. Larin, "The four-loop β-function in quantum chromodynamics", *Phys. Lett. B* **400** (1997) 379, doi:10.1016/S0370-2693(97)00370-5, arXiv:hep-ph/9701390.
58. G. 't Hooft, "Dimensional regularization and the renormalization group", *Nucl. Phys. B* **61** (1973) 455, doi:10.1016/0550-3213(73)90376-3.
59. W. A. Bardeen, A. J. Buras, D. W. Duke, and T. Muta, "Deep Inelastic Scattering Beyond the Leading Order in Asymptotically Free Gauge Theories", *Phys. Rev. D* **18** (1978) 3998, doi:10.1103/PhysRevD.18.3998.
60. J. Beringer and others (Particle Data Group), "Review of Particle Physics", *Phys. Rev. D* **86** (2012) 010001, doi:10.1103/PhysRevD.86.010001.
61. S. Durr et al., "Ab-Initio Determination of Light Hadron Masses", *Science* **322** (2008) 1224, doi:10.1126/science.1163233, arXiv:0906.3599.
62. J. C. Collins, D. E. Soper, and G. Sterman, "Factorization of Hard Processes in QCD", volume 5 of *Advanced Series on Directions in High Energy Physics*, ch. 1, p. 1. World Scientific Pub Co Inc, 1988. arXiv:hep-ph/0409313. doi:10.1142/9789814503266_0001.
63. J. R. Andersen et al., "Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report", arXiv:1405.1067.
64. S. Dittmaier, A. Huss, and C. Speckner, "Weak radiative corrections to dijet production at hadron colliders", *JHEP* **11** (2012) 095, doi:10.1007/JHEP11(2012)095, arXiv:1210.0438.
65. V. N. Gribov and L. N. Lipatov, "Deep inelastic ep scattering in perturbation theory", *Sov. J. Nucl. Phys.* **15** (1972) 438.
66. L. N. Lipatov, "The parton model and perturbation theory", *Sov. J. Nucl. Phys.* **20** (1975) 94. [Yad. Fiz.20,181(1974)].
67. G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nucl. Phys. B* **126** (1977) 298, doi:10.1016/0550-3213(77)90384-4.
68. 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.
69. S. Moch, J. A. M. Vermaseren, and A. Vogt, "The Three loop splitting functions in QCD: The Nonsinglet case", *Nucl. Phys. B* **688** (2004) 101, doi:10.1016/j.nuclphysb.2004.03.030, arXiv:hep-ph/0403192.
70. A. Vogt, S. Moch, and J. A. M. Vermaseren, "The three-loop splitting functions in QCD: The singlet case", *Nucl. Phys. B* **691** (2004) 129, doi:10.1016/j.nuclphysb.2004.04.024, arXiv:hep-ph/0404111.
71. E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, "The Pomeranchuk Singularity in Nonabelian Gauge Theories", *Sov. Phys. JETP* **45** (1977) 199. [Zh. Eksp. Teor. Fiz.72,377(1977)].
72. I. I. Balitsky and L. N. Lipatov, "The Pomeranchuk Singularity in Quantum Chromodynamics", *Sov. J. Nucl. Phys.* **28** (1978) 822. [Yad. Fiz.28,1597(1978)].
73. W. J. Stirling, "Parton luminosity and cross section plots". Private communication, 2016. http://www.hep.ph.ic.ac.uk/ wstirlin/plots/plots.html.
74. C. Buttar et al., "The QCD, EW, and Higgs Working Group: Summary Report", in *Proceedings, 4th Les Houches Workshop 2005 on Physics at TeV colliders (Les Houches 2005)*. Les Houches, France, May 2-20, 2005. arXiv:hep-ph/0604120.
75. S. D. 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.
76. S. M. Berman, J. D. Bjorken, and J. B. Kogut, "Inclusive Processes at High Transverse Momentum", *Phys. Rev. D* **4** (1971) 3388, doi:10.1103/PhysRevD.4.3388.
77. F. Halzen and D. M. Scott, "Energy Flow: Testing QCD without Structure Functions", in *Proceedings, 11th International Symposium on Multiparticle Dynamics (ISMD)*, p. 0593. Bruges, Belgium, June 22-27, 1980.

78. S. D. Drell, D. J. Levy, and T.-M. Yan, "A Theory of Deep-Inelastic Lepton-Nucleon Scattering and Lepton-Pair Annihilation Processes. I.", *Phys. Rev.* **187** (1969) 2159, doi:10.1103/PhysRev.187.2159.
79. S. D. Drell, D. J. Levy, and T.-M. Yan, "A Theory of Deep-Inelastic Lepton-Nucleon Scattering and Lepton-Pair Annihilation Processes. III. Deep-Inelastic Electron-Positron Annihilation", *Phys. Rev. D* **1** (1970) 1617, doi:10.1103/PhysRevD.1.1617.
80. N. Cabibbo, G. Parisi, and M. Testa, "Hadron Production in e^+e^- Collisions", *Lett. Nuovo Cim.* **4S1** (1970) 35, doi:10.1007/BF02755392. [Lett. Nuovo Cim.4,35(1970)].
81. J. D. Bjorken and S. J. Brodsky, "Statistical Model for electron-Positron Annihilation Into Hadrons", *Phys. Rev. D* **1** (1970) 1416, doi:10.1103/PhysRevD.1.1416.
82. R. P. Feynman, R. D. Field, and G. C. Fox, "Quantum-chromodynamic approach for the large-transverse- momentum production of particles and jets", *Phys. Rev. D* **18** (1978) 3320, doi:10.1103/PhysRevD.18.3320.
83. A. V. Belitsky, G. P. Korchemsky, and G. F. Sterman, "Energy flow in QCD and event shape functions", *Phys. Lett. B* **515** (2001) 297, doi:10.1016/S0370-2693(01)00899-1, arXiv:hep-ph/0106308.
84. A. Ali and G. Kramer, "Jets and QCD: A Historical Review of the Discovery of the Quark and Gluon Jets and its Impact on QCD", *Eur. Phys. J. H* **36** (2011) 245, doi:10.1140/epjh/e2011-10047-1, arXiv:1012.2288.
85. P. Söding, "On the discovery of the gluon", *Eur. Phys. J. H* **35** (2010) 3, doi:10.1140/epjh/e2010-00002-5.
86. B. R. Stella and H.-J. Meyer, "Y(9.46 GeV) and the gluon discovery (a critical recollection of PLUTO results)", *Eur. Phys. J. H* **36** (2011) 203, doi:10.1140/epjh/e2011-10029-3, arXiv:1008.1869.
87. ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
88. CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
89. CDF Collaboration, "Observation of top quark production in $\bar{p}p$ collisions", *Phys. Rev. Lett.* **74** (1995) 2626, doi:10.1103/PhysRevLett.74.2626, arXiv:hep-ex/9503002.
90. D0 Collaboration, "Observation of the top quark", *Phys. Rev. Lett.* **74** (1995) 2632, doi:10.1103/PhysRevLett.74.2632, arXiv:hep-ex/9503003.
91. H. U. Bengtsson, "The Lund Monte Carlo for High p_T Physics", *Comput. Phys. Commun.* **31** (1984) 323, doi:10.1016/0010-4655(84)90018-3.
92. T. Sjostrand, "The Lund Monte Carlo for Jet Fragmentation", *Comput. Phys. Commun.* **27** (1982) 243, doi:10.1016/0010-4655(82)90175-8.
93. T. Sjöstrand, "High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4", *Comput. Phys. Commun.* **82** (1994) 74, doi:10.1016/0010-4655(94)90132-5.
94. T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
95. G. Marchesini and B. R. Webber, "Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation", *Nucl. Phys. B* **310** (1988) 461, doi:10.1016/0550-3213(88)90089-2.
96. G. Marchesini et al., "HERWIG: A Monte Carlo event generator for simulating hadron emission reactions with interfering gluons. Version 5.1 - April 1991", *Comput. Phys. Commun.* **67** (1992) 465, doi:10.1016/0010-4655(92)90055-4.
97. G. Corcella et al., "HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)", *JHEP* **01** (2001) 010, doi:10.1088/1126-6708/2001/01/010, arXiv:hep-ph/0011363.
98. G. Corcella et al., "HERWIG 6.5 release note", arXiv:hep-ph/0210213.

99. T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.
100. T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.
101. M. Bähr et al., "Herwig++ Physics and Manual", *Eur. Phys. J. C* **58** (2008) 639, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.
102. J. Bellm et al., "Herwig 7.0 / Herwig++ 3.0 Release Note", arXiv:1512.01178.
103. T. Gleisberg et al., "Event generation with SHERPA 1.1", *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007, arXiv:0811.4622.
104. A. Buckley et al., "General-purpose event generators for LHC physics", *Phys. Rept.* **504** (2011) 145, doi:10.1016/j.physrep.2011.03.005, arXiv:1101.2599.
105. S. Gieseke and Z. Nagy, "Monte Carlo Generators and Fixed-order Calculations: Predicting the (Un)Expected", ch. 5, p. 97. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 1st edition, April, 2011. doi:10.1002/9783527634965.ch5.
106. S. Gieseke, "Simulation of jets at colliders", *Prog. Part. Nucl. Phys.* **72** (2013) 155, doi:10.1016/j.ppnp.2013.04.001.
107. S. Höche, "Introduction to parton-shower event generators", in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014)*. Boulder, CO, USA, June 2-27, 2014. arXiv:1411.4085.
108. M. A. Dobbs et al., "Les Houches guidebook to Monte Carlo generators for hadron collider physics", in *Proceedings, 3rd Les Houches Workshop 2003 on Physics at TeV colliders (Les Houches 2003)*, p. 411. Les Houches, France, May 26-June 2, 2003. arXiv:hep-ph/0403045.
109. I. G. Knowles, "A linear algorithm for calculating spin correlations in hadronic collisions", *Comput. Phys. Commun.* **58** (1990) 271, doi:10.1016/0010-4655(90)90063-7.
110. I. G. Knowles, "Spin Correlations in Parton - Parton Scattering", *Nucl. Phys. B* **310** (1988) 571, doi:10.1016/0550-3213(88)90092-2.
111. M. Bengtsson and T. Sjöstrand, "A Comparative Study of Coherent and Noncoherent Parton Shower Evolution", *Nucl. Phys. B* **289** (1987) 810, doi:10.1016/0550-3213(87)90407-X.
112. M. Bengtsson and T. Sjöstrand, "Coherent Parton Showers Versus Matrix Elements: Implications of PETRA - PEP Data", *Phys. Lett. B* **185** (1987) 435, doi:10.1016/0370-2693(87)91031-8.
113. T. Sjöstrand and P. Z. Skands, "Transverse-momentum-ordered showers and interleaved multiple interactions", *Eur. Phys. J. C* **39** (2005) 129, doi:10.1140/epjc/s2004-02084-y, arXiv:hep-ph/0408302.
114. S. Gieseke, P. Stephens, and B. Webber, "New formalism for QCD parton showers", *JHEP* **12** (2003) 045, doi:10.1088/1126-6708/2003/12/045, arXiv:hep-ph/0310083.
115. J.-C. Winter and F. Krauss, "Initial-state showering based on colour dipoles connected to incoming parton lines", *JHEP* **07** (2008) 040, doi:10.1088/1126-6708/2008/07/040, arXiv:0712.3913.
116. S. Schumann and F. Krauss, "A parton shower algorithm based on Catani-Seymour dipole factorisation", *JHEP* **03** (2008) 038, doi:10.1088/1126-6708/2008/03/038, arXiv:0709.1027.
117. G. Gustafson and U. Pettersson, "Dipole Formulation of QCD Cascades", *Nucl. Phys. B* **306** (1988) 746, doi:10.1016/0550-3213(88)90441-5.
118. B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton Fragmentation and String Dynamics", *Phys. Rept.* **97** (1983) 31, doi:10.1016/0370-1573(83)90080-7.
119. B. Andersson, G. Gustafson, and B. Söderberg, "A General Model for Jet Fragmentation", *Z. Phys. C* **20** (1983) 317, doi:10.1007/BF01407824.

120. T. Sjöstrand, "The merging of jets", *Phys. Lett. B* **142** (1984) 420, doi:10.1016/0370-2693(84)91354-6.
121. B. R. Webber, "A QCD Model for Jet Fragmentation Including Soft Gluon Interference", *Nucl. Phys. B* **238** (1984) 492, doi:10.1016/0550-3213(84)90333-X.
122. J.-C. Winter, F. Krauss, and G. Soff, "A modified cluster hadronization model", *Eur. Phys. J. C* **36** (2004) 381, doi:10.1140/epjc/s2004-01960-8, arXiv:hep-ph/0311085.
123. CDF Collaboration, "Charged jet evolution and the underlying event in $p\bar{p}$ collisions at 1.8 TeV", *Phys. Rev. D* **65** (2002) 092002, doi:10.1103/PhysRevD.65.092002.
124. CDF Collaboration, "The underlying event in hard interactions at the Tevatron $\bar{p}p$ collider", *Phys. Rev. D* **70** (2004) 072002, doi:10.1103/PhysRevD.70.072002, arXiv:hep-ex/0404004.
125. CDF Collaboration, "Studying the Underlying Event in Drell-Yan and High Transverse Momentum Jet Production at the Tevatron", *Phys. Rev. D* **82** (2010) 034001, doi:10.1103/PhysRevD.82.034001, arXiv:1003.3146.
126. CDF Collaboration, "Study of the energy dependence of the underlying event in proton-antiproton collisions", *Phys. Rev. D* **92** (2015) 092009, doi:10.1103/PhysRevD.92.092009, arXiv:1508.05340.
127. ATLAS Collaboration, "Measurement of underlying event characteristics using charged particles in pp collisions at $\sqrt{s} = 900$ GeV and 7 TeV with the ATLAS detector", *Phys. Rev. D* **83** (2011) 112001, doi:10.1103/PhysRevD.83.112001, arXiv:1012.0791.
128. ATLAS Collaboration, "Measurements of underlying-event properties using neutral and charged particles in pp collisions at 900 GeV and 7 TeV with the ATLAS detector at the LHC", *Eur. Phys. J. C* **71** (2011) 1636, doi:10.1140/epjc/s10052-011-1636-z, arXiv:1103.1816.
129. ATLAS Collaboration, "Measurement of the underlying event in jet events from 7 TeV proton-proton collisions with the ATLAS detector", *Eur. Phys. J. C* **74** (2014) 2965, doi:10.1140/epjc/s10052-014-2965-5, arXiv:1406.0392.
130. ATLAS Collaboration, "Measurement of distributions sensitive to the underlying event in inclusive Z-boson production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *Eur. Phys. J. C* **74** (2014) 3195, doi:10.1140/epjc/s10052-014-3195-6, arXiv:1409.3433.
131. CMS Collaboration, "First Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 0.9$ TeV", *Eur. Phys. J. C* **70** (2010) 555, doi:10.1140/epjc/s10052-010-1453-9, arXiv:1006.2083.
132. CMS Collaboration, "Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and Comparison with $\sqrt{s} = 0.9$ TeV", *JHEP* **09** (2011) 109, doi:10.1007/JHEP09(2011)109, arXiv:1107.0330.
133. CMS Collaboration, "Measurement of the underlying event in the Drell-Yan process in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **72** (2012) 2080, doi:10.1140/epjc/s10052-012-2080-4, arXiv:1204.1411.
134. CMS Collaboration, "Study of the underlying event at forward rapidity in pp collisions at $\sqrt{s} = 0.9$, 2.76, and 7 TeV", *JHEP* **04** (2013) 072, doi:10.1007/JHEP04(2013)072, arXiv:1302.2394.
135. CMS Collaboration, "Measurement of the underlying event activity using charged-particle jets in proton-proton collisions at $\sqrt{s} = 2.76$ TeV", *JHEP* **09** (2015) 137, doi:10.1007/JHEP09(2015)137, arXiv:1507.07229.
136. ALICE Collaboration, "Underlying Event measurements in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV with the ALICE experiment at the LHC", *JHEP* **07** (2012) 116, doi:10.1007/JHEP07(2012)116, arXiv:1112.2082.
137. D. Piparo, "Statistical Combination of Higgs Decay Channels and Determination of the Jet-Energy Scale of the CMS Experiment at the LHC". PhD thesis, KIT (Karlsruher Institut für Technologie), November, 2010.

138. T. Sjöstrand and M. van Zijl, "A Multiple Interaction Model for the Event Structure in Hadron Collisions", *Phys. Rev. D* **36** (1987) 2019, doi:10.1103/PhysRevD.36.2019.
139. T. Sjöstrand and P. Z. Skands, "Multiple interactions and the structure of beam remnants", *JHEP* **03** (2004) 053, doi:10.1088/1126-6708/2004/03/053, arXiv:hep-ph/0402078.
140. M. Bähr, S. Gieseke, and M. H. Seymour, "Simulation of multiple partonic interactions in Herwig++", *JHEP* **07** (2008) 076, doi:10.1088/1126-6708/2008/07/076, arXiv:0803.3633.
141. A. Buckley et al., "Rivet user manual", *Comput. Phys. Commun.* **184** (2013) 2803, doi:10.1016/j.cpc.2013.05.021, arXiv:1003.0694.
142. A. Buckley et al., "Systematic event generator tuning for the LHC", *Eur. Phys. J. C* **65** (2010) 331, doi:10.1140/epjc/s10052-009-1196-7, arXiv:0907.2973.
143. H. Abramowicz et al., "Summary of the Workshop on Multi-Parton Interactions (MPI@LHC 2012)", arXiv:1306.5413.
144. S. Plätzer and M. Diehl, eds., "Proceedings of the 3rd International Workshop on Multiple Partonic Interactions at the LHC (MPI@LHC 2011), Hamburg, Germany, November 21-25", DESY. DESY, Hamburg, Germany, (2011). doi:10.3204/DESY-PROC-2012-03.
145. P. Bartalini and L. Fanó, eds., "Proceedings of the 1st International Workshop on Multiple Partonic Interactions at the LHC (MPI@LHC 2008), Perugia, Italy, October 27-31". (2008). arXiv:1003.4220.
146. CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2015) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
147. CDF Collaboration, R. Field, "The Energy Dependence of the Underlying Event in Hadronic Collisions", in *Proceedings, The 2013 European Physical Society Conference on High Energy Physics (EPS-HEP 2013)*, volume EPS-HEP2013, p. 422. Stockholm, Sweden, July 18-24, 2013.
148. ATLAS Collaboration, "New ATLAS event generator tunes to 2010 data", technical report, CERN, 2011.
149. P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes", *Phys. Rev. D* **82** (2010) 074018, doi:10.1103/PhysRevD.82.074018, arXiv:1005.3457.
150. J. M. Katzy, "QCD Monte-Carlo model tunes for the LHC", *Prog. Part. Nucl. Phys.* **73** (2013) 141, doi:10.1016/j.ppnp.2013.08.002.
151. S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, "QCD matrix elements + parton showers", *JHEP* **11** (2001) 063, doi:10.1088/1126-6708/2001/11/063, arXiv:hep-ph/0109231.
152. F. Krauss, "Matrix elements and parton showers in hadronic interactions", *JHEP* **08** (2002) 015, doi:10.1088/1126-6708/2002/08/015, arXiv:hep-ph/0205283.
153. L. Lönnblad, "Correcting the color dipole cascade model with fixed order matrix elements", *JHEP* **05** (2002) 046, doi:10.1088/1126-6708/2002/05/046, arXiv:hep-ph/0112284.
154. M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-quark production in hadronic collisions", *JHEP* **01** (2007) 013, doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.
155. J. Alwall et al., "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions", *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
156. S. Catani and M. H. Seymour, "A General algorithm for calculating jet cross-sections in NLO QCD", *Nucl. Phys. B* **485** (1997) 291, doi:10.1016/S0550-3213(96)00589-5, arXiv:hep-ph/9605323.
157. S. Frixione, Z. Kunszt, and A. Signer, "Three jet cross-sections to next-to-leading order", *Nucl. Phys. B* **467** (1996) 399, doi:10.1016/0550-3213(96)00110-1, arXiv:hep-ph/9512328.

158. D. A. Kosower, "Antenna factorization of gauge theory amplitudes", *Phys. Rev. D* **57** (1998) 5410, doi:10.1103/PhysRevD.57.5410, arXiv:hep-ph/9710213.
159. Z. Nagy and D. E. Soper, "General subtraction method for numerical calculation of one-loop QCD matrix elements", *JHEP* **09** (2003) 055, doi:10.1088/1126-6708/2003/09/055, arXiv:hep-ph/0308127.
160. S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations", *JHEP* **06** (2002) 029, doi:10.1088/1126-6708/2002/06/029, arXiv:hep-ph/0204244.
161. P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.
162. S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
163. S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
164. S. Höche, F. Krauss, M. Schönherr, and F. Siegert, "A critical appraisal of NLO+PS matching methods", *JHEP* **09** (2012) 049, doi:10.1007/JHEP09(2012)049, arXiv:1111.1220.
165. J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
166. K. Hamilton and P. Nason, "Improving NLO-parton shower matched simulations with higher order matrix elements", *JHEP* **06** (2010) 039, doi:10.1007/JHEP06(2010)039, arXiv:1004.1764.
167. S. Höche, F. Krauss, M. Schönherr, and F. Siegert, "NLO matrix elements and truncated showers", *JHEP* **08** (2011) 123, doi:10.1007/JHEP08(2011)123, arXiv:1009.1127.
168. S. Höche, F. Krauss, M. Schönherr, and F. Siegert, "QCD matrix elements + parton showers: The NLO case", *JHEP* **04** (2013) 027, doi:10.1007/JHEP04(2013)027, arXiv:1207.5030.
169. L. Lönnblad and S. Prestel, "Merging Multi-leg NLO Matrix Elements with Parton Showers", *JHEP* **03** (2013) 166, doi:10.1007/JHEP03(2013)166, arXiv:1211.7278.
170. K. Hamilton, P. Nason, C. Oleari, and G. Zanderighi, "Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching", *JHEP* **05** (2013) 082, doi:10.1007/JHEP05(2013)082, arXiv:1212.4504.
171. S. Höche et al., "Next-to-leading order QCD predictions for top-quark pair production with up to two jets merged with a parton shower", *Phys. Lett. B* **748** (2015) 74, doi:10.1016/j.physletb.2015.06.060, arXiv:1402.6293.
172. G. F. Sterman and S. Weinberg, "Jets from Quantum Chromodynamics", *Phys. Rev. Lett.* **39** (1977) 1436, doi:10.1103/PhysRevLett.39.1436.
173. JADE Collaboration, "Experimental Studies on Multi-Jet Production in e+ e- Annihilation at PETRA Energies", *Z. Phys. C* **33** (1986) 23, doi:10.1007/BF01410449.
174. B. Flaugher and K. Meier, "A Compilation of jet finding algorithms", in *Proceedings, 5th DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90)*, p. 128. Snowmass, CO, USA, Jun 25-Jul 13, 1990.
175. J. E. Huth et al., "Towards a standardization of jet definitions", in *Proceedings, 5th DPF Summer Study on High-energy Physics: Research Directions for the Decade (Snowmass 90)*, p. 134. Snowmass, CO, USA, Jun 25-Jul 13, 1990.
176. G. C. Blazey et al., "Run II jet physics", in *Proceedings, Physics at Run II: QCD and Weak Boson Physics Workshop*, p. 47. Batavia, IL, USA, March 4-6, June 3-4, November 4-6, 1999. arXiv:hep-ex/0005012.

177. C. Buttar et al., "Standard Model Handles and Candles Working Group: Tools and Jets Summary Report", in *Proceedings, 5th Les Houches Workshop 2007 on Physics at TeV colliders (Les Houches 2007)*, p. 121. Les Houches, France, June 11-29, 2007. arXiv:0803.0678.
178. G. P. Salam, "Towards Jetography", *Eur. Phys. J. C* **67** (2010) 637, doi:10.1140/epjc/s10052-010-1314-6, arXiv:0906.1833.
179. M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
180. CDF Collaboration, "Measurement of the Inclusive Jet Cross Section at the Fermilab Tevatron $p\bar{p}$ Collider Using a Cone-Based Jet Algorithm", *Phys. Rev. D* **78** (2008) 052006, doi:10.1103/PhysRevD.79.119902,10.1103/PhysRevD.78.052006, arXiv:0807.2204.
181. D0 Collaboration, "Measurement of the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. D* **85** (2012) 052006, doi:10.1103/PhysRevD.85.052006, arXiv:1110.3771. Long author list - awaiting processing.
182. G. P. Salam and G. Soyez, "A practical Seedless Infrared-Safe Cone jet algorithm", *JHEP* **05** (2007) 086, doi:10.1088/1126-6708/2007/05/086, arXiv:0704.0292.
183. S. Catani and B. R. Webber, "Infrared safe but infinite: Soft gluon divergences inside the physical region", *JHEP* **10** (1997) 005, doi:10.1088/1126-6708/1997/10/005, arXiv:hep-ph/9710333.
184. T. Hebbeker, "Tests of quantum chromodynamics in hadronic decays of Z^0 bosons produced in e^+e^- annihilation", *Phys. Rept.* **217** (1992) 69, doi:10.1016/0370-1573(92)90181-X.
185. S. Catani et al., "New clustering algorithm for multi - jet cross-sections in e+ e- annihilation", *Phys. Lett. B* **269** (1991) 432, doi:10.1016/0370-2693(91)90196-W.
186. M. Dasgupta, L. Magnea, and G. P. Salam, "Non-perturbative QCD effects in jets at hadron colliders", *JHEP* **02** (2008) 055, doi:10.1088/1126-6708/2008/02/055, arXiv:0712.3014.
187. UA1 Collaboration, "Hadronic Jet Production at the CERN Proton-Antiproton Collider", *Phys. Lett. B* **132** (1983) 214, doi:10.1016/0370-2693(83)90254-X.
188. UA2 Collaboration, "Observation of Very Large Transverse Momentum Jets at the CERN $\bar{p}p$ Collider", *Phys. Lett. B* **118** (1982) 203, doi:10.1016/0370-2693(82)90629-3.
189. UA2 Collaboration, "Measurement of Production and Properties of Jets at the CERN anti-p p Collider", *Z. Phys. C* **20** (1983) 117, doi:10.1007/BF01573214.
190. M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett. B* **641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.
191. S. D. Ellis, J. Huston, and M. Tönnesmann, "On building better cone jet algorithms", *eConf* **C010630** (2001) 513, arXiv:hep-ph/0111434.
192. J. Berger, "Search for the Higgs Boson Produced via Vector-Boson Fusion in the Decay Channel H $\to \tau\tau$". PhD thesis, KIT (Karlsruher Institut für Technologie), Jun, 2014.
193. M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188.
194. S. D. Ellis and D. E. Soper, "Successive combination jet algorithm for hadron collisions", *Phys. Rev. D* **48** (1993) 3160, doi:10.1103/PhysRevD.48.3160, arXiv:hep-ph/9305266.
195. Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better Jet Clustering Algorithms", *JHEP* **08** (1997) 001, doi:10.1088/1126-6708/1997/08/001, arXiv:hep-ph/9707323.
196. M. Wobisch and T. Wengler, "Hadronization corrections to jet cross-sections in deep inelastic scattering", in *Proceedings, Monte Carlo Generators for HERA Physics*. Hamburg, Germany, 1998-1999, 1998. arXiv:hep-ph/9907280.

197. M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
198. CMS Collaboration, "Measurement of the underlying event activity in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV with the novel jet-area/median approach", *JHEP* **08** (2012) 130, doi:10.1007/JHEP08(2012)130, arXiv:1207.2392.
199. R. Barlow et al., "Data Analysis in High Energy Physics". Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 1st edition, June, 2013.
200. F. I. Olness and D. E. Soper, "Correlated theoretical uncertainties for the one-jet inclusive cross section", *Phys. Rev. D* **81** (2010) 035018, doi:10.1103/PhysRevD.81.035018, arXiv:0907.5052.
201. C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO", *Phys. Rev. D* **69** (2004) 094008, doi:10.1103/PhysRevD.69.094008, arXiv:hep-ph/0312266.
202. M. Cacciari and N. Houdeau, "Meaningful characterisation of perturbative theoretical uncertainties", *JHEP* **09** (2011) 039, doi:10.1007/JHEP09(2011)039, arXiv:1105.5152.
203. E. Bagnaschi, M. Cacciari, A. Guffanti, and L. Jenniches, "An extensive survey of the estimation of uncertainties from missing higher orders in perturbative calculations", *JHEP* **02** (2015) 133, doi:10.1007/JHEP02(2015)133, arXiv:1409.5036.
204. M. Cacciari et al., "The t anti-t cross-section at 1.8 TeV and 1.96 TeV: A Study of the systematics due to parton densities and scale dependence", *JHEP* **04** (2004) 068, doi:10.1088/1126-6708/2004/04/068, arXiv:hep-ph/0303085.
205. A. Banfi, G. P. Salam, and G. Zanderighi, "Phenomenology of event shapes at hadron colliders", *JHEP* **06** (2010) 038, doi:10.1007/JHEP06(2010)038, arXiv:1001.4082.
206. G. Grunberg, "Renormalization Group Improved Perturbative QCD", *Phys. Lett. B* **95** (1980) 70, doi:10.1016/0370-2693(80)90402-5. [Erratum: Phys. Lett.B110,501(1982)].
207. J. Kubo and S. Sakakibara, "Equivalence of the Fastest Apparent Convergence Criterion and the Principle of Minimal Sensitivity in Perturbative Quantum Chromodynamics", *Phys. Rev. D* **26** (1982) 3656, doi:10.1103/PhysRevD.26.3656.
208. P. M. Stevenson, "Resolution of the Renormalization Scheme Ambiguity in Perturbative QCD", *Phys. Lett. B* **100** (1981) 61, doi:10.1016/0370-2693(81)90287-2.
209. P. M. Stevenson, "Optimized Perturbation Theory", *Phys. Rev. D* **23** (1981) 2916, doi:10.1103/PhysRevD.23.2916.
210. S. J. Brodsky, G. P. Lepage, and P. B. Mackenzie, "On the Elimination of Scale Ambiguities in Perturbative Quantum Chromodynamics", *Phys. Rev. D* **28** (1983) 228, doi:10.1103/PhysRevD.28.228.
211. S. J. Brodsky and L. Di Giustino, "Setting the Renormalization Scale in QCD: The Principle of Maximum Conformality", *Phys. Rev. D* **86** (2012) 085026, doi:10.1103/PhysRevD.86.085026, arXiv:1107.0338.
212. S. J. Brodsky, M. Mojaza, and X.-G. Wu, "Systematic Scale-Setting to All Orders: The Principle of Maximum Conformality and Commensurate Scale Relations", *Phys. Rev. D* **89** (2014) 014027, doi:10.1103/PhysRevD.89.014027, arXiv:1304.4631.
213. H.-Y. Bi et al., "Degeneracy Relations in QCD and the Equivalence of Two Systematic All-Orders Methods for Setting the Renormalization Scale", *Phys. Lett. B* **748** (2015) 13–18, doi:10.1016/j.physletb.2015.06.056, arXiv:1505.04958.
214. M. Rubin, G. P. Salam, and S. Sapeta, "Giant QCD K-factors beyond NLO", *JHEP* **09** (2010) 084, doi:10.1007/JHEP09(2010)084, arXiv:1006.2144.
215. A. David and G. Passarino, "How well can we guess theoretical uncertainties?", *Phys. Lett. B* **726** (2013) 266, doi:10.1016/j.physletb.2013.08.025, arXiv:1307.1843.

216. CDF Collaboration, "Inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **77** (1996) 438, doi:10.1103/PhysRevLett.77.438, arXiv:hep-ex/9601008.
217. J. Huston et al., "Large transverse momentum jet production and the gluon distribution inside the proton", *Phys. Rev. Lett.* **77** (1996) 444, doi:10.1103/PhysRevLett.77.444, arXiv:hep-ph/9511386.
218. A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC", *Eur. Phys. J. C* **63** (2009) 189, doi:10.1140/epjc/s10052-009-1072-5, arXiv:0901.0002.
219. S. Dulat et al., "New parton distribution functions from a global analysis of quantum chromodynamics", *Phys. Rev. D* **93** (2016) 033006, doi:10.1103/PhysRevD.93.033006, arXiv:1506.07443.
220. E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, "On the heavy quark content of the nucleon", *Phys. Lett. B* **291** (1992) 325, doi:10.1016/0370-2693(92)91053-C.
221. E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, "Complete O(α_s) corrections to heavy flavor structure functions in electroproduction", *Nucl. Phys. B* **392** (1993) 162, doi:10.1016/0550-3213(93)90201-Y.
222. S. Riemersma, J. Smith, and W. L. van Neerven, "Rates for inclusive deep inelastic electroproduction of charm quarks at HERA", *Phys. Lett. B* **347** (1995) 143, doi:10.1016/0370-2693(95)00036-K, arXiv:hep-ph/9411431.
223. R. S. Thorne and W. K. Tung, "PQCD Formulations with Heavy Quark Masses and Global Analysis", in *Proceedings, 4th Workshop on the Implications of HERA for LHC Physics (HERA and the LHC)*. Geneva, Switzerland, May 26-30, 2008. arXiv:0809.0714.
224. N. Arkani-Hamed, T. Han, M. Mangano, and L.-T. Wang, "Physics Opportunities of a 100 TeV Proton-Proton Collider", arXiv:1511.06495.
225. S. Alekhin, J. Blümlein, and S. Moch, "Parton Distribution Functions and Benchmark Cross Sections at NNLO", *Phys. Rev. D* **86** (2012) 054009, doi:10.1103/PhysRevD.86.054009, arXiv:1202.2281.
226. J. F. Owens, A. Accardi, and W. Melnitchouk, "Global parton distributions with nuclear and finite-Q^2 corrections", *Phys. Rev. D* **87** (2013) 094012, doi:10.1103/PhysRevD.87.094012, arXiv:1212.1702.
227. H.-L. Lai et al., "New parton distributions for collider physics", *Phys. Rev. D* **82** (2010) 074024, doi:10.1103/PhysRevD.82.074024, arXiv:1007.2241.
228. H1 and ZEUS Collaboration, "Combined Measurement and QCD Analysis of the Inclusive $e^{\pm}p$ Scattering Cross Sections at HERA", *JHEP* **01** (2010) 109, doi:10.1007/JHEP01(2010)109, arXiv:0911.0884.
229. M. Glück, P. Jimenez-Delgado, and E. Reya, "Dynamical parton distributions of the nucleon and very small-x physics", *Eur. Phys. J. C* **53** (2008) 355, doi:10.1140/epjc/s10052-007-0462-9, arXiv:0709.0614.
230. M. Glück, P. Jimenez-Delgado, E. Reya, and C. Schuck, "On the role of heavy flavor parton distributions at high energy colliders", *Phys. Lett. B* **664** (2008) 133, doi:10.1016/j.physletb.2008.04.063, arXiv:0801.3618.
231. A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections", *Eur. Phys. J. C* **64** (2009) 653, doi:10.1140/epjc/s10052-009-1164-2, arXiv:0905.3531.
232. R. D. Ball et al., "Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology", *Nucl. Phys. B* **849** (2011) 296, doi:10.1016/j.nuclphysb.2011.03.021, arXiv:1101.1300.
233. R. D. Ball et al., "Parton distributions with LHC data", *Nucl. Phys. B* **867** (2013) 244, doi:10.1016/j.nuclphysb.2012.10.003, arXiv:1207.1303.
234. A. Accardi et al., "Constraints on large-x parton distributions from new weak boson production and deep-inelastic scattering data", arXiv:1602.03154.

235. H1, ZEUS Collaboration, "Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data", *Eur. Phys. J. C* **75** (2015) 580, doi:10.1140/epjc/s10052-015-3710-4, arXiv:1506.06042.
236. P. Jimenez-Delgado and E. Reya, "Delineating parton distributions and the strong coupling", *Phys. Rev. D* **89** (2014) 074049, doi:10.1103/PhysRevD.89.074049, arXiv:1403.1852.
237. L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs", *Eur. Phys. J. C* **75** (2015) 204, doi:10.1140/epjc/s10052-015-3397-6, arXiv:1412.3989.
238. NNPDF Collaboration, "Parton distributions for the LHC Run II", *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
239. M. R. Whalley, D. Bourilkov, and R. C. Group, "The Les Houches Accord PDFs (LHAPDF) and LHAGLUE", in *Proceedings, HERA and the LHC: A Workshop on the implications of HERA for LHC physics: Vol. B*. Geneva, Switzerland and Hamburg, Germany, March 26-27, October 11-13, March 21-24, 2004-2005. arXiv:hep-ph/0508110.
240. A. Buckley et al., "LHAPDF6: parton density access in the LHC precision era", *Eur. Phys. J.* **C75** (2015) 132, doi:10.1140/epjc/s10052-015-3318-8, arXiv:1412.7420.
241. D. Stump et al., "Uncertainties of predictions from parton distribution functions. 1. The Lagrange multiplier method", *Phys. Rev. D* **65** (2001) 014012, doi:10.1103/PhysRevD.65.014012, arXiv:hep-ph/0101051.
242. J. Pumplin et al., "Uncertainties of predictions from parton distribution functions. 2. The Hessian method", *Phys. Rev. D* **65** (2001) 014013, doi:10.1103/PhysRevD.65.014013, arXiv:hep-ph/0101032.
243. W. T. Giele and S. Keller, "Implications of hadron collider observables on parton distribution function uncertainties", *Phys. Rev. D* **58** (1998) 094023, doi:10.1103/PhysRevD.58.094023, arXiv:hep-ph/9803393.
244. W. T. Giele, S. A. Keller, and D. A. Kosower, "Parton distribution function uncertainties", arXiv:hep-ph/0104052.
245. NNPDF Collaboration, "A determination of parton distributions with faithful uncertainty estimation", *Nucl. Phys. B* **809** (2009) 1, doi:10.1016/j.nuclphysb.2008.09.037, arXiv:0808.1231.
246. S. Alekhin et al., "HERAFitter", *Eur. Phys. J. C* **75** (2015) 304, doi:10.1140/epjc/s10052-015-3480-z, arXiv:1410.4412.
247. M. Botje et al., "The PDF4LHC Working Group Interim Recommendations", arXiv:1101.0538.
248. J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.
249. C. Bourrely, J. Soffer, and F. Buccella, "A Statistical approach for polarized parton distributions", *Eur. Phys. J. C* **23** (2002) 487, doi:10.1007/s100520100855, arXiv:hep-ph/0109160.
250. C. Bourrely and J. Soffer, "New developments in the statistical approach of parton distributions: tests and predictions up to LHC energies", *Nucl. Phys. A* **941** (2015) 307, doi:10.1016/j.nuclphysa.2015.06.018, arXiv:1502.02517.
251. SM and NLO Multileg Working Group Collaboration, T. Binoth et al., "The SM and NLO Multileg Working Group: Summary report", in *Proceedings, 6th Les Houches Workshop 2009 on Physics at TeV colliders (Les Houches 2009; dedicated to Thomas Binoth)*, p. 21. Les Houches, France, June 8-26, 2010. arXiv:1003.1241.
252. H.-L. Lai et al., "Uncertainty induced by the QCD coupling in the CTEQ global analysis of parton distributions", *Phys. Rev. D* **82** (2010) 054021, doi:10.1103/PhysRevD.82.054021, arXiv:1004.4624.
253. F. Demartin et al., "The impact of PDF and alphas uncertainties on Higgs Production in gluon fusion at hadron colliders", *Phys. Rev. D* **82** (2010) 014002, doi:10.1103/PhysRevD.82.014002, arXiv:1004.0962.

254. CMS Collaboration, "Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at $\sqrt{s}$ = 7 TeV", *Eur. Phys. J. C* **75** (2015) 288, doi:10.1140/epjc/s10052-015-3499-1, arXiv:1410.6765.
255. S. D. Ellis, Z. Kunszt, and D. E. Soper, "One-Jet Inclusive Cross Section at Order α_s^3: Quarks and Gluons", *Phys. Rev. Lett.* **64** (1990) 2121, doi:10.1103/PhysRevLett.64.2121.
256. S. D. Ellis, Z. Kunszt, and D. E. Soper, "Two-Jet Production in Hadron Collisions at Order α_s^3 in QCD", *Phys. Rev. Lett.* **69** (1992) 1496, doi:10.1103/PhysRevLett.69.1496.
257. J. Gao et al., "MEKS: a program for computation of inclusive jet cross sections at hadron colliders", *Comput. Phys. Commun.* **184** (2013) 1626, doi:10.1016/j.cpc.2013.01.022, arXiv:1207.0513.
258. W. T. Giele, E. W. N. Glover, and D. A. Kosower, "Higher order corrections to jet cross-sections in hadron colliders", *Nucl. Phys. B* **403** (1993) 633, doi:10.1016/0550-3213(93)90365-V, arXiv:hep-ph/9302225.
259. Z. Nagy, "Three jet cross-sections in hadron hadron collisions at next-to-leading order", *Phys. Rev. Lett.* **88** (2002) 122003, doi:10.1103/PhysRevLett.88.122003, arXiv:hep-ph/0110315.
260. Z. Nagy, "Next-to-leading order calculation of three-jet observables in hadron hadron collisions", *Phys. Rev. D* **68** (2003) 094002, doi:10.1103/PhysRevD.68.094002, arXiv:hep-ph/0307268.
261. J. Currie, A. Gehrmann-De Ridder, E. W. N. Glover, and J. Pires, "NNLO QCD corrections to jet production at hadron colliders from gluon scattering", *JHEP* **01** (2014) 110, doi:10.1007/JHEP01(2014)110, arXiv:1310.3993.
262. Z. Bern et al., "Four-Jet Production at the Large Hadron Collider at Next-to-Leading Order in QCD", *Phys. Rev. Lett.* **109** (2012) 042001, doi:10.1103/PhysRevLett.109.042001, arXiv:1112.3940.
263. S. Badger, B. Biedermann, P. Uwer, and V. Yundin, "NLO QCD corrections to multi-jet production at the LHC with a centre-of-mass energy of $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **718** (2013) 965, doi:10.1016/j.physletb.2012.11.029, arXiv:1209.0098.
264. S. Badger, B. Biedermann, P. Uwer, and V. Yundin, "Next-to-leading order QCD corrections to five jet production at the LHC", *Phys. Rev. D* **89** (2014) 034019, doi:10.1103/PhysRevD.89.034019, arXiv:1309.6585.
265. H1 Collaboration, "Measurement and QCD analysis of jet cross-sections in deep-inelastic positron-proton collisions at $\sqrt{s}$ of 300 GeV", *Eur. Phys. J. C* **19** (2001) 289, doi:10.1007/s100520100621, arXiv:hep-ex/0010054.
266. T. Kluge, K. Rabbertz, and M. Wobisch, "fastNLO: Fast pQCD calculations for PDF fits", in *14th International Workshop on Deep Inelastic Scattering (DIS 2006)*, p. 483. Tsukuba, Japan, April 20-24, 2006. arXiv:hep-ph/0609285. doi:10.1142/9789812706706_0110.
267. T. Carli, G. P. Salam, and F. Siegert, "A Posteriori inclusion of PDFs in NLO QCD final-state calculations", in *HERA and the LHC: A Workshop on the Implications of HERA for LHC Physics (Startup Meeting, CERN, 26-27 March 2004; Working Group Meeting, CERN, 17-21 January 2005; Final Meeting 21-24 Mar 2005) CERN, Geneva, Switzerland, October 11-13, 2004*. 2005. arXiv:hep-ph/0510324.
268. T. Carli et al., "A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project", *Eur. Phys. J. C* **66** (2010) 503, doi:10.1140/epjc/s10052-010-1255-0, arXiv:0911.2985.
269. D. Britzger, K. Rabbertz, F. Stober, and M. Wobisch, "New features in version 2 of the fastNLO project", in *Proceedings, XX. International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012)*, p. 217. Bonn, Germany, March 26-30, 2012. arXiv:1208.3641. doi:10.3204/DESY-PROC-2012-02/165.

270. D. A. Britzger, “Regularized Unfolding of Jet Cross Sections in Deep-Inelastic ep Scattering at HERA and Determination of the Strong Coupling Constant”. PhD thesis, Universität Hamburg, June, 2013.

Chapter 3
Jet Measurement

Besides the theoretical framework for jet physics, an experimental setup is needed. Most prominently, a collider is required to accelerate and collide two particle beams head-on. The two main characteristics of such a collider are its centre-of-mass energy, which delimits the amount of energy available per collision, and its luminosity, which determines the number of collisions per time interval. The centre-of-mass energy constrains the production of new, massive particles, while the luminosity is the limiting factor for the observation of rare processes.

The LHC started Run 1 operations in 2010 with proton-proton collisions at a centre-of-mass energy of 7 TeV, which later was raised to 8 TeV in 2012. Lead-lead beams were accelerated to the equivalent of 2.76 and 5.02 TeV of collision energy per nucleon. A peak instantaneous luminosity of $8 \cdot 10^{33}\text{cm}^{-2}\text{s}^{-1}$ was achieved, exceeding design parameters for $\sqrt{s} = 8\,\text{TeV}$. A figure showing the luminosity accumulated over the years is presented in Chapter 7. More detailed descriptions of the LHC, its history and construction can be found in Ref. [1] and in the Chapters 12 and 1–2 of Refs. [2] and [3], respectively.

This chapter introduces the necessary experimental concepts and tools needed for a generic jet analysis at the LHC. The sections are ordered in a similar way as they would typically appear in an experimental publication and start with a synopsis of the measuring apparatus. This is followed by a description of the jet reconstruction techniques and the corresponding trigger setup initiating the event recording. Subsequently, the important aspect of jet energy calibration is explained in detail. The next section deals with generic jet and event selections, which filter out background processes and remove spurious jet-like signatures originating from detector noise. The chapter closes with a discussion of the jet energy resolution and the unfolding of detector effects.

3.1 The Detectors

The focus is on the two omni-purpose detectors ATLAS and CMS with, the author being a member of the CMS Collaboration, the main emphasis on the latter. While this section presents concise summaries, the two detectors are described in detail in Refs. [4] and [5], respectively. Descriptions of the other two large LHC experiments ALICE and LHCb can be found in Refs. [6, 7]. The general concepts and technologies of particle detectors are delineated in Chapters 12–18 of Ref. [2] and Chapters 1–3 of Ref. [3].

As customary, the two omni-purpose detectors ATLAS and CMS follow the "onion" paradigm: The interaction point is enveloped by layers of detector components, each specialised in measuring specific effects caused by the collision products that traverse the experimental apparatus from the inside outwards. Accounting for the beam pipe in the centre, the subdetectors are arranged in cylindrical layers with increasing radii around the central beam line. The innermost layers consist of silicon pixel detectors to localise ionisation signals caused by charged particles as close as possible to the interaction point and thereby optimise the resolution of multiple primary and of secondary (decay) vertices. They are complemented by multiple layers of silicon strip tracking devices to pursue the trajectories of charged particles. Both are immersed in a strong magnetic field to allow a precise determination of track curvatures and hence the associated momenta.

The surrounding calorimeters might be located in- (CMS) or outside (ATLAS) the solenoidal coil creating the magnetic field. They are subdivided into an inner electromagnetic part absorbing and measuring the energy of electrons, positrons, and photons, and an outer hadronic part responsible for the energy determination of hadronically interacting particles, e.g. protons, neutrons, or charged pions. The outermost layer is made of tracking chambers completing the trajectories of muons escaping the inner parts of the detector. Neutrinos can be recognised only indirectly through a deficit in the transverse momentum balance. Figure 3.1 presents such typical trajectories at the example of a transverse slice through the CMS detector.

Along the beam direction the detectors are completed by endcaps oriented perpendicularly to the beam and subdivided into multiple layers similarly as the cylindrical part, such that the interaction point is enclosed almost hermetically. This general layout is complemented by further instrumentation for triggering and particle identification purposes and approximates in the best possible way a measuring apparatus of 4π geometric coverage. Schematic views of the ATLAS and CMS detectors are shown in Figs. 3.2 and 3.3; concise descriptions of the two have been adapted from Refs. [8] and [9] and are given in the following two sections.

3.1.1 The ATLAS Detector

The ATLAS detector consists of a tracking system, sampling electromagnetic and hadronic calorimeters and muon chambers.

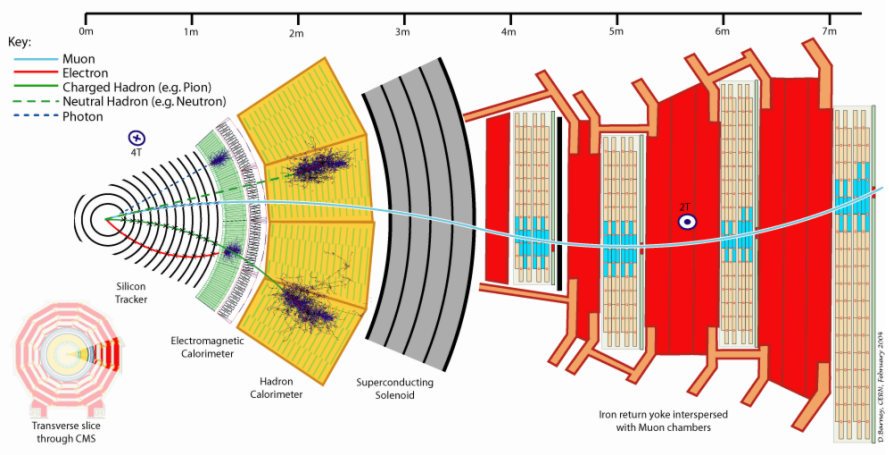

Fig. 3.1 Typical trajectories of different types of particles traversing a transverse slice through the CMS detector. *(Adapted from Source: CERN)*

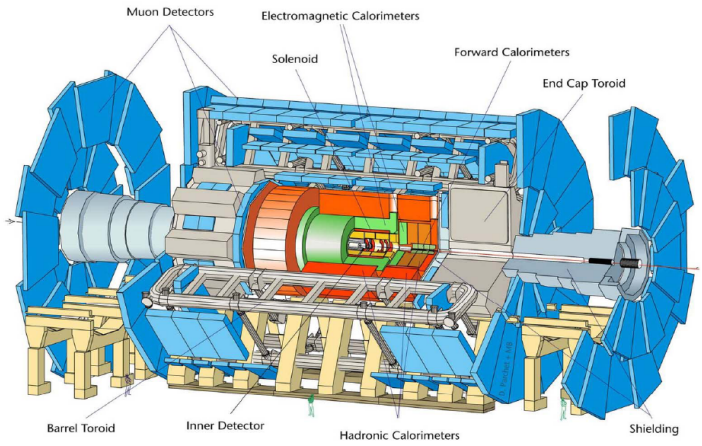

Fig. 3.2 Schematic view of the ATLAS detector with its overall length of 46 m, overall diameter of 22 m and a total weight of 7 000 tons. *(Adapted from Source: CERN)*

The inner Detector has complete azimuthal coverage and spans the pseudorapidity region $|\eta| < 2.5$. It consists of layers of silicon pixel detectors, silicon microstrip detectors and transition radiation tracking detectors, all of which are immersed in a solenoid magnet that provides a uniform magnetic field of 2 T.

Jets are reconstructed using the ATLAS calorimeters, whose granularity and material varies as a function of η. The electromagnetic calorimetry (EM) is provided by high-granularity liquid-argon sampling calorimeters (LAr), using lead as an absorber. It is divided into one barrel ($|\eta| < 1.475$) and two endcap ($1.375 < |\eta| < 3.2$) regions. The hadronic calorimetry is divided into a barrel, endcap, and a forward

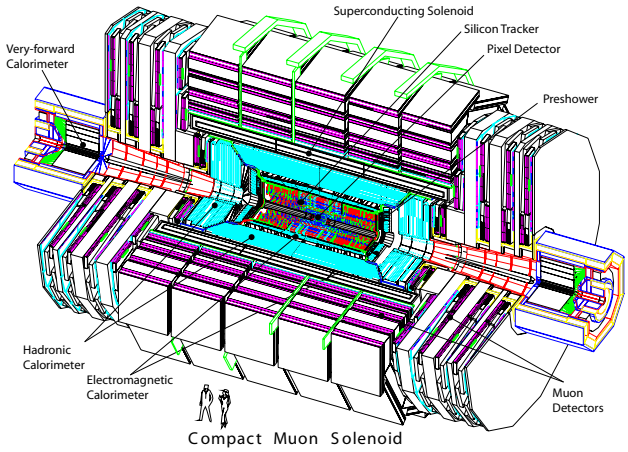

Fig. 3.3 Schematic view of the CMS detector with its overall length of 20 m, overall diameter of 15 m and a total weight of 12 500 tons. *(Adapted from Source: CERN)*

section. The barrel section combines a central region ($|\eta| < 0.8$) with two extended barrel regions ($0.8 < |\eta| < 1.7$). These regions are instrumented with scintillator-tile/steel hadronic calorimeters (Tile). Each barrel region consists of 64 modules with individual ϕ coverages of ~ 0.1 rad. The two hadronic endcap calorimeters (HEC; $1.5 < |\eta| < 3.2$) feature liquid-argon/copper calorimeter modules. The two forward calorimeters (FCal; $3.1 < |\eta| < 4.9$) are instrumented with liquid-argon/copper and liquid-argon/tungsten modules to provide electromagnetic and hadronic energy measurements, respectively.

Up to a pseudorapidity of $|\eta| = 2.7$, the muon spectrometer surrounding the ATLAS calorimeter is immersed in a magnetic field that is generated by a system of three large air-core toroids, a barrel and two endcaps. The muon spectrometer measures muon tracks with three layers of precision tracking chambers and is instrumented with separate trigger chambers.

The trigger system for the ATLAS detector consists of a hardware-based Level 1 (L1) and a software-based High-Level Trigger (HLT). At L1, jets are first built from coarse-granularity calorimeter towers using a sliding window algorithm, and then subjected to early trigger decisions. In the HLT, this is refined using jets reconstructed from calorimeter cells with algorithms similar to the ones applied offline.

3.1.2 The CMS Detector

The central feature of the CMS apparatus is a 3.8 T superconducting solenoid of 6 m internal diameter. Within the field volume are the silicon tracker, the crystal electromagnetic calorimeter (ECAL), and the brass and scintillator hadron calorimeter

(HCAL). The muon system is installed outside the solenoid and embedded in the steel flux-return yoke.

The CMS tracker consists of 1 440 silicon pixel and 15 148 silicon strip detector modules, with full azimuthal coverage within $|\eta| < 2.5$. The ECAL consists of 75 848 lead tungstate crystals, which provide coverage in pseudorapidity $|\eta| < 1.479$ in the central barrel region and $1.479 < |\eta| < 3.000$ in the two forward endcap regions. The HCAL is a sampling calorimeter using alternating layers of brass or steel as absorber and plastic scintillator as active material, it provides a coverage of $|\eta| < 1.3$ in the central region and $1.3 < |\eta| < 3.0$ in the endcap regions. In the forward region ($3.0 < |\eta| < 5.0$), a different calorimeter technology is employed in the hadron forward (HF) detector, which uses the Cherenkov light signals collected by short and long quartz readout fibres to aid the separation of electromagnetic and hadronic signals. The muon system includes barrel drift tubes covering the pseudorapidity range $|\eta| < 1.2$, endcap cathode strip chambers ($0.9 < |\eta| < 2.5$), and resistive-plate chambers ($|\eta| < 1.6$).

Interaction vertices are reconstructed using track information only. The primary interaction or signal vertex is defined as the vertex with the highest sum of the squared transverse momenta of the tracks associated with it.

The L1 of the CMS trigger system, composed of custom hardware processors, uses information from the calorimeters and muon detectors to select the most interesting events in a fixed time interval of less than 4 μs. The HLT processor farm further decreases the event rate from around 100 kHz to less than 1 kHz before data storage.

3.2 Jet Reconstruction

Both experiments employ the collinear- and infrared-safe anti-k_t algorithm [10] to cluster low-level reconstruction objects into final jets. Apart from the diverging choice in default jet size parameters R, 0.4 and 0.6 in the case of ATLAS and 0.5 and 0.7 for CMS, digressions in the construction of the input objects are caused by the differences in detector technologies. The baseline measurements of hadronic jets are performed by the calorimeters, whose performance is characterised by their relative energy resolution:

$$\left(\frac{\Delta E}{E}\right)^2 = \left(\frac{N}{E}\right)^2 + \left(\frac{S}{\sqrt{E}}\right)^2 + C^2, \tag{3.1}$$

where E is to be taken in units of GeV. As usual, the energy dependence is specified in terms of an electronic noise term N independent of the deposited energy, a stochastic term S fluctuating with the number of sampled charged shower particles, and a constant term C absorbing i.a. effects of shower tails, geometric inhomogeneities, and nonlinearities in the signal response. Table 3.1 lists the main

Table 3.1 Main characteristics of the electromagnetic and hadronic calorimeters of the ATLAS and CMS detectors. The relative energy resolution is given in terms of the parameters for the stochastic and constant term according to equation (3.1). *(Adapted from Ref. [3])*

Calorimeter		Material	Number of channels	Angular coverage	Energy resolution S [%]	C [%]
ATLAS elm.	EM barrel	LAr+Pb	109 568	$\lvert\eta\rvert < 1.475$	10	0.7
	EM endcap	LAr+Pb	63 744	$1.375 < \lvert\eta\rvert < 3.2$	10	0.7
	FCal	LAr+Cu	2 016	$3.1 < \lvert\eta\rvert < 4.9$	28.5	3.5
ATLAS had.	Tile	scint.+Pb	9 852	$\lvert\eta\rvert < 1.7$	52	3
	HEC	LAr+Cu	5 632	$1.5 < \lvert\eta\rvert < 3.2$	84	—
	FCal	LAr+W	1 508	$3.1 < \lvert\eta\rvert < 4.9$	94	7.5
CMS elm.	ECAL barrel	$PbWO_4$	61 200	$\lvert\eta\rvert < 1.479$	2.8	0.3
	ECAL endcap	$PbWO_4$	14 648	$1.479 < \lvert\eta\rvert < 3.0$	2.8	0.3
CMS had.	HB	scint.+steel/brass	2 592	$\lvert\eta\rvert < 1.3$	90	9
	HE	scint.+steel/brass	2 592	$1.3 < \lvert\eta\rvert < 3.0$	90	9
	HO	scint.+steel	2 160	$\lvert\eta\rvert < 1.4$	—	—
	HF	quartz fibre+steel	1 728	$3.0 < \lvert\eta\rvert < 5.2$	120	—

characteristics of the electromagnetic and hadronic calorimeters of the ATLAS and CMS detectors.

The liquid-argon sampling calorimeter of ATLAS provides a highly granular and homogeneous coverage of a wide region in pseudorapidity up to $|\eta| = 4.9$ with about 17 000 readout channels in total and a hadronic energy resolution of roughly $50\%/\sqrt{E}$ in the central barrel region. The jet reconstruction in ATLAS therefore is primarily based on so-called *topo-clusters* of energy deposits built from topologically connected calorimeter cells with signals above some noise threshold. The algorithm is described in detail in Ref. [11].

Afterwards, two calibration steps are performed on these topo-clusters. Initially, they are reconstructed at the electromagnetic scale (EM) as appropriate for energy depositions in the calorimeter by particles produced in electromagnetic showers. In a second step, the *local cell weighting* (LCW) method classifies the energy deposits based on their energy density and longitudinal shower profile into electromagnetic and hadronic ones, and subsequently corrects the hadronic signals to account for the non-compensating nature of the ATLAS calorimeter [8]. The LCW method is the primary technique used for hadronic jets. In addition, ATLAS considers *track-jets*, which are made from all reconstructed charged-particle tracks that are associated with the primary interaction vertex.

In contrast, hadronic calorimetry in CMS features in total 9 000 readout channels for a similar range in pseudorapidity with an energy resolution of $\Delta E/E \approx 90\%/\sqrt{E}$ also in the central region. Electromagnetic calorimetry compensates the smaller number of channels, $\approx$ 76 000 in CMS versus $\approx$ 173 000 in ATLAS, through the almost four times better energy resolution in the $PbWO_4$ crystals. However, to com-

pete with ATLAS in terms of hadronic jets it is mandatory to combine the information of all detector components including the tracking devices in the form of a global event reconstruction named *Particle Flow* (PF) described in the next section adapted from Ref. [9].

3.2.1 Particle Flow

In a first step, CMS also employs a topological clustering of calorimeter cells in order to suppress electronic noise and group the cells into energy deposits in the form of "projective towers" pointing to the nominal interaction vertex. Such a calorimeter tower consists of one or more HCAL cells and the geometrically corresponding ECAL crystals. Jets formed from applying a jet algorithm solely to these towers are called *calorimeter jets* (CALO jets). They result from a relatively simplistic yet robust approach, were widely used in the early CMS publications, and are still used for cross-checks.

To profit from a more complete and consistent event interpretation and an improved performance combining all subdetectors, CMS uses the PF technique [12, 13], where the full cell-based information of the calorimeter towers is available. This PF technique reconstructs events as a whole and tries to identify single particles with an optimised combination of all subdetector information. Five categories of particle candidates are differentiated: photons, electrons (including positrons), muons, charged hadrons, and neutral hadrons.

To suppress noise in the calorimeters, only cells with energies above a given threshold are considered, this procedure is referred to as "zero suppression". The energy of photons is obtained directly from the ECAL measurement, corrected for zero-suppression effects. The energy of electrons is determined from a combination of the track momentum at the main interaction vertex, the corresponding ECAL cluster energy, and the energy sum of all bremsstrahlung photons associated with the track. The energy of muons is obtained from the corresponding track momentum. The energy of charged hadrons is determined from a combination of the track momentum and the corresponding ECAL and HCAL energies, corrected for zero-suppression effects, and calibrated for the nonlinear response of the calorimeters. Finally, the energy of neutral hadrons is obtained from the corresponding calibrated ECAL and HCAL energies. In the forward region, energy deposits collected by the HF are considered as electromagnetic or hadronic, depending on the respective energy collected by long and short fibres. The particles reconstructed with the PF algorithm are jointly referred to as PF candidates. An illustration of this algorithm is presented in Fig. 3.4.

PF jets are reconstructed by clustering the four-momentum vectors of PF candidates. The PF jet momentum and spatial resolutions are greatly improved with respect to calorimeter jets, as the use of the tracking detectors and high granularity of the ECAL improves the energy resolution through the independent measurements of charged hadrons and photons inside a jet, which together constitute $\approx$85% of

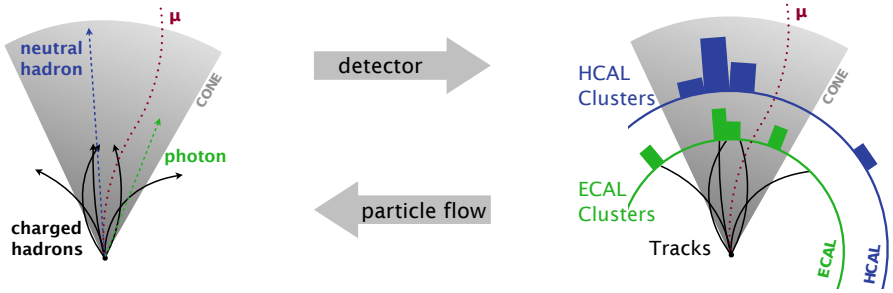

Fig. 3.4 Schematic association of subdetector measurements to physical particle candidates using the PF technique of CMS. *(Illustration courtesy of F. Pandolfi)*

the average jet energy. In reconstructing the PF candidate four-momenta, photons are assumed massless and charged hadrons are assigned the charged pion mass. To demonstrate the percent level agreement of the PF global event interpretation between data and simulation, Fig. 3.5 shows the jet energy fractions attributed to each PF particle category versus p_T within $|\eta| < 1.3$ on the left and versus η for jets with $56 < p_T < 74$ GeV on the right. Outside the tracker coverage beyond $\eta = 2.5$ the differentiation between charged and neutral particles is not possible anymore.

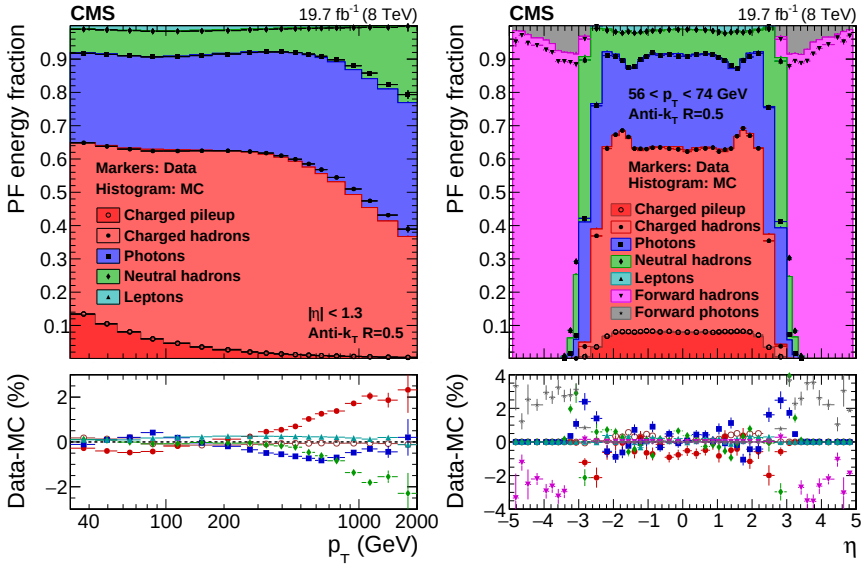

Fig. 3.5 PF jet composition in data and simulation versus p_T at $|\eta| < 1.3$ (left), and versus η at $56 < p_T < 74$ GeV (right). *(Taken from Ref. [9])*

3.2.2 Particle Jets

According to the recommendation on jet definitions in [14] one must clearly specify the final-state "truth level", i.e. the objects of MC generated events up to which detector corrections are performed. At CMS, final-state particles are considered to be stable, if their mean decay lengths exceed $c\tau = 10\,\text{mm}$, compare Section 2.3.4. This is roughly equivalent to $\tau > 30\,\text{ps}$ mentioned in [15] and a factor of three larger than the 10 ps recommended in Ref. [14]. For ATLAS, the situation is somewhat unclear, because in Ref. [8] the condition $c\tau > 10\,\text{mm}$ is given, while Refs. [16] and [17], which compare ATLAS and CMS, explicitly mention $\tau > 10\,\text{ps}$. For all practical purposes concerning simulations for the LHC this difference can presumably be neglected.

By convention, ATLAS and CMS exclude neutrinos from the clustering to obtain the generated particle-level jets, while elsewhere neutrinos are often included at the particle level. This convention allows both experiments to define the jet response in a way that is experimentally accessible and significantly reduces response differences between heavy-flavor (b, c) and light-quark (u, d, s) or gluon jets caused by neutrinos produced in semileptonic decays of heavy-flavor hadrons. Since the jet response is measured from samples with negligible neutrino content, this does not lead to any significant bias for inclusive jet energy calibrations.

To be best adapted to reconstructed PF jets, the particle-level definition of CMS does include muons, while ATLAS only accepts particles detectable in their calorimeters and hence excludes muons. Both, neutrinos and muons carry a significant fraction of the parton energy for jets containing semileptonic decays of heavy quarks, which therefore have to be corrected explicitly and acquire a supplemental systematic uncertainty for the b- and c-jet fragmentation.

3.3 Triggering

The LHC provides colliding bunches separated by time intervals of initially 50 and, since autumn 2015, by 25 ns corresponding to a maximal rate of 40 MHz. Apart from the fact that it is impossible to read out and store events at this rate, most of the recorded data would not contain any useful piece of information. From Fig. 2.14 it is evident that interesting processes like the production of a Higgs boson occur at rates more than ten orders of magnitude smaller than the production of anything, i.e. the total inelastic interaction rate. Therefore, a trigger system must be deployed that recognises rare and interesting events and flags them for further read-out, processing, and storage.

ATLAS and CMS both use a multistage triggering system with an L1 trigger based on custom-made programmable hardware that supports event rates up to 100 kHz with a maximal latency of 2.5 respectively $3.2\,\mu\text{s}$ for a decision. Further event data are buffered meanwhile in a pipelined memory structure. Only a reduced amount of information from the subdetectors with fast signal response, typically

the muon spectrometer and the calorimeters, is available at this stage. Sought-after signatures are high-p_{T} muon, electron/photon, jet, or τ-jet candidates. In addition, global event properties might be used, e.g.:

- the scalar sum of transverse energies $\sum E_{\mathrm{T}}$,
- the scalar sum of jet transverse momenta H_{T},
- the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}} = |\vec{\not{E}}_{\mathrm{T}}|$, where $\vec{\not{E}}_{\mathrm{T}}$ is the negative sum of transverse momentum vectors,
- or a particularly high track multiplicity.

At the analysis stage in CMS, these quantities are derived from all PF candidates in the event.

At a second stage (L2) the ATLAS trigger analyses larger regions of interest of an event and further reduces the output to about 3.5 kHz within 40 ms. Together with the L2, which is missing in the case of CMS, a subsequent filter farm composed of commercial processors constitutes the software-based HLT that fully reconstructs events and reduces the rate of recorded and permanently stored events to a manageable level of several hundred Hertz.

The data samples used for jet measurements are mostly collected with a series of single-jet HLT triggers that require at least one jet in the event to exceed a threshold in calibrated jet p_{T}, increasing with the trigger number. Alternatively, triggers on dijet mass, $\sum E_{\mathrm{T}}$, or H_{T} are in use. Depending on the instantaneous luminosity during a particular data taking period, lower-threshold triggers might be pre-scaled to prioritise the rate budget for high-p_{T} physics.

In the final trigger stage of all LHC detectors, reconstruction algorithms are kept as similar as possible to those used in later offline reconstructions. For various reasons, however, differences occur. In particular, the online jet reconstruction might not have used the latest calibration constants. Together with the fact that the jet trigger objects are derived from fast but simplified algorithms using e.g. only calorimetric information (calorimeter jets), or EM-calibrated jets for one particular jet size R, the necessity arises to check the efficiency of each trigger for the respective analysis. For this purpose, it is extremely useful to have the original trigger objects leading to a decision at disposal at the later analysis stage.

One possibility to determine a trigger efficiency for a set of conditions i, ε_i, consists in comparing against the cross section σ_{ref} measured by a reference trigger with less stringent conditions:

$$\varepsilon_i(O) = \frac{\sigma_i(O)}{\sigma_{\mathrm{ref}}(O)}\,. \tag{3.2}$$

Here, O represents the quantity of interest as a function of which the efficiency is to be determined. For example, Fig. 3.6 left shows for a series of single-jet triggers in ATLAS their respective efficiency as a function of the calibrated jet p_{T}. This method requires to have all relevant luminosity and trigger pre-scale information available in order to evaluate the cross sections and their ratio. Large pre-scales of the lower-threshold reference trigger may lead in this case to an insufficient statistical accuracy.

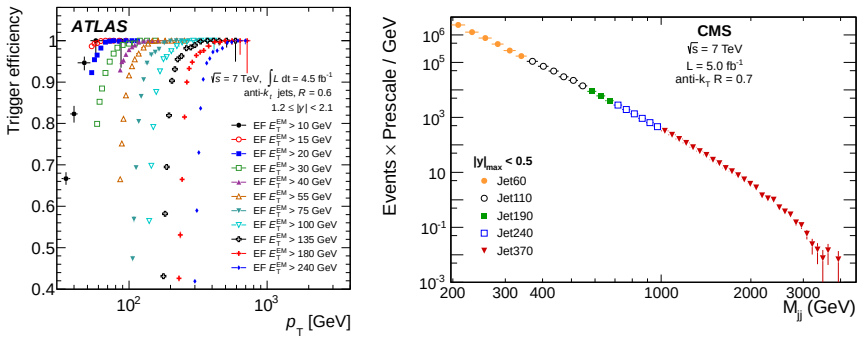

Fig. 3.6 Left: Jet-trigger efficiency as a function of calibrated jet p_{T} at the electromagnetic scale for $R = 0.6$ anti-k_{t} jets in ATLAS within $1.2 \leq |y| < 2.1$ for various trigger thresholds. Right: Dijet-mass spectrum construction for $R = 0.7$ anti-k_{t} jets in CMS within $|y|_{\mathrm{max}} < 0.5$ from various single-jet trigger paths as indicated by different markers. *(Taken from Refs. [19, 20])*

A second method relies on a proper trigger emulation based on the availability of the original trigger objects:

$$\varepsilon_i(O) = \frac{\#\{\mathrm{cond}\,(\mathrm{ref} = \mathrm{true}) \wedge \mathrm{cond}\,(i = \mathrm{true})\}(O)}{\#\{\mathrm{cond}\,(\mathrm{ref} = \mathrm{true})\}(O)}\,. \tag{3.3}$$

Effectively, this means to count the number, #, of events of the set $\{\}$ fulfilling the reference AND, $\wedge$, the additional trigger conditions i in comparison to the number of all events fulfilling the reference requirements. Complications by event weights through large trigger pre-scales or by the luminosity evaluation are avoided, which simplifies fits to determine the $> 99\%$ efficiency threshold. This technique was applied for example in the first publication on inclusive jets by CMS [18].

Both methods, however, only provide the *relative* efficiency with respect to a reference, so that the absolute efficiency must be crosschecked with independent triggers e.g. for muons or for *minimum bias* (MB) events, where only a minimal trigger signal is required like a reconstructed vertex with some tracks pointing to it.

Accepting only events that have passed the highest-threshold, unprescaled trigger above each respective 99% efficiency point, the data samples for each trigger with the respective accumulated integrated luminosity can be combined into a properly weighted spectrum. For example, Fig. 3.6 right presents the trigger composition of a dijet mass spectrum as measured by CMS.

A more complicated procedure involving the combination of multiple trigger paths for an observable bin is necessary, if unprescaled triggers that reach full efficiency as a function of that observable are either not available or would leave gaps in the distribution. This is the case for example in the CMS analysis described in Section 4.3, where single-jet triggers are combined to reproduce the 3-jet mass spectrum. For illustration Fig. 3.7 left shows as before a trigger turn-on curve for a single-jet trigger path as a function of jet p_{T}, where the trigger efficiency errors are

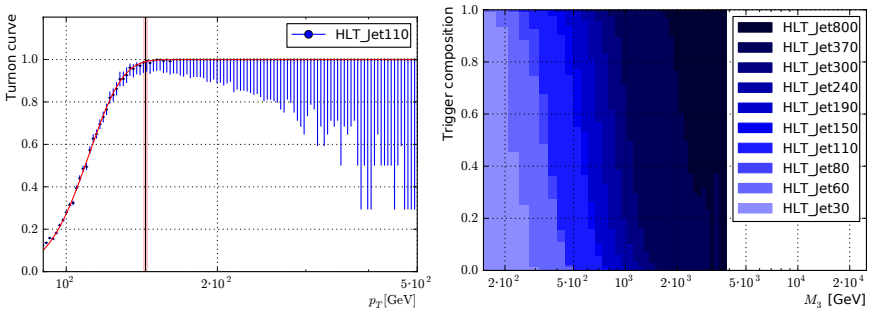

Fig. 3.7 Left: Trigger efficiency as a function of jet p_T for a single-jet trigger path with efficiency errors based on Clopper–Pearson confidence intervals. The trigger turn-on curve is fitted with a Gauss error function (line) to determine the 99% trigger threshold indicated by the vertical line with uncertainty band. Right: Trigger composition of a potential cross-section measurement as a function of 3-jet mass. For each 3-jet mass bin, the fraction of events contributed by a particular single-jet trigger path is shown. *(Adapted from plots courtesy of F. Stober [24])*

based on Clopper–Pearson confidence intervals [21].[1] The point of a minimal efficiency of 99% is determined by fitting a Gauss error function $\mathrm{erf}(x)$ to this curve. The resulting limit with its uncertainty is indicated by the shaded vertical line. As a function of 3-jet mass, however, multiple single-jet trigger paths must be combined as illustrated in Fig. 3.7 right. A more generic discussion on combination schemes of pre-scaled triggers is presented in Ref. [23].

3.4 Jet Energy Calibration

Like all reconstructed objects, jets need to be calibrated such that on average the correct energy can be assigned to them. This jet energy calibration or correction (JEC)[2] procedure is performed in a series of steps, which are similar between ATLAS and CMS:

1. The first step, detailed in Section 3.4.2, accounts for extra energy caused by electronic noise or by PU collisions.
2. In a second stage, cf. Section 3.4.3, the bulk of the correction is performed using detailed detector simulations.
3. Unavoidable imperfections in the detector modelling through simulation are identified and corrected in phase three with data-based methods as a function of jet pseudorapidity and p_T, explained in Section 3.4.4.
4. Finally in Section 3.4.5, MC-based corrections account for differences in the flavour compositions of signal and calibration samples, if necessary.

[1] Alternatively, Wilson score intervals could be used for the binomial proportion confidence intervals [22].

[2] In some figures the JEC uncertainty might also be abbreviated using JES for *jet energy scale*.

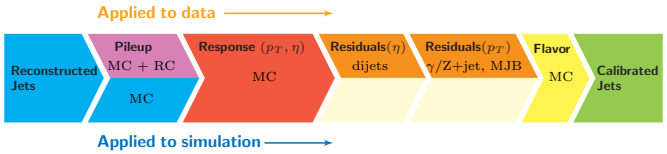

Fig. 3.8 Consecutive stages of the jet energy calibration as performed in CMS [9]. The upper half corresponds to corrections applied to data, while the lower half lists the ones applied to simulation. See text for further explanations. ATLAS proceeds in a similar fashion. *(Taken from Ref. [9])*

The JEC procedure is illustrated in Fig. 3.8 and presented in more detail in the following at the example of the latest CMS study evaluating data that correspond to about $20\,\text{fb}^{-1}$ of integrated luminosity accumulated in 2012 at 8 TeV centre-of-mass energy [9]. This is complemented on occasion with results from ATLAS reported on $5\,\text{fb}^{-1}$ of data taken at $\sqrt{s} = 7\,\text{TeV}$ in 2011 [8].

3.4.1 Simulated Samples

The simulation of modern detectors is based on the GEANT4 package [25] that tracks any initial set of particles with given four-momenta through the detector by accounting for each interaction with its material up to an adjustable level of precision. Thereby, the trajectories of particles are transformed into material hits in the tracking devices. In addition, the evolution of electromagnetic and hadronic showers, initiated by the impact of high-energetic particles, is emulated and energy losses in the sensitive calorimeter material are represented by corresponding energy depositions. Mandatory input to the simulation is a detailed model of the detector geometry and alignment, the calibration of all detector elements, and an emulation of the readout electronics.

The initial set of particles impacting on the detector material is determined with the help of MC event generators as described in the previous chapter, cf. Section 2.4. For the purpose of deriving jet response corrections from simulation, event samples of dijet, multijet, Z+jet, and γ+jet production are required. Additional samples are used for e.g. heavy-quark jet production or for special purposes like single-neutrino generation for events containing exclusively PU and electronic noise. As baseline for the JEC in LHC Run 1, the general-purpose event generators PYTHIA versions 6 and 8, and HERWIG++ version 2 are used. They provide LO predictions for the processes of interest, complemented with parton showers, MPI, and hadronisation. MADGRAPH or ALPGEN are interfaced to PYTHIA or HERWIG++ to complement the tree-level multijet production. Because the emphasis here is not on precision comparisons of cross sections, but on studies of the detector behaviour, more time-consuming higher-order predictions are not required for the purpose of the JEC. The exact composition of JEC simulation samples with respect to MC event generator versions and tunes varies largely between the different experimental collaborations.

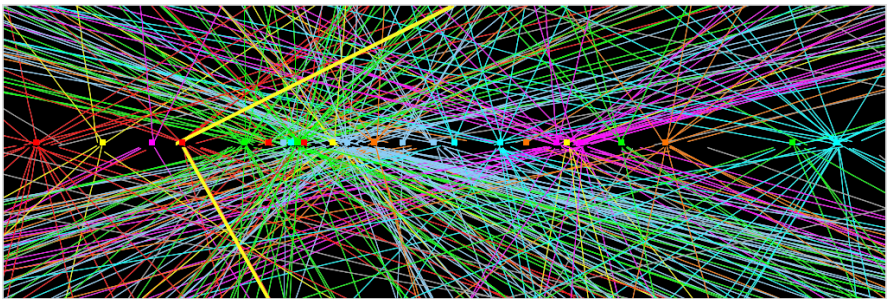

Fig. 3.9 Zoomed display of an ATLAS event with two high-p_T muons (thick yellow lines) and 25 reconstructed primary vertices. *(Source: CERN, ATLAS)*

In the most recent JEC setup of CMS [9], the dijet, γ+jet, and single-neutrino samples are produced with PYTHIA6 tune Z2*. The Z+jet and multijet samples are generated with MADGRAPH matched to the parton showers by PYTHIA6 tune Z2*. Additional samples of dijet and Z+jet processes are provided by HERWIG++ 2.3 with tune EE3C for systematic uncertainty studies.

For the offset corrections in Section 3.4.2, the single-neutrino sample is compared to *zero-bias* (ZB) data. Zero-bias events are collected using a random trigger in the presence of a beam crossing and capture the detector signals from zero collisions up to many PU collisions. If at least a minimal trigger signal is required like a reconstructed vertex with some tracks, then this is again a minimum bias event. They correspond to very low-p_T dijet events and are used to construct the pile-up overlay samples. Dijets at higher p_T are used to simulate the jet response analysed in Section 3.4.3 and to compare to data in the dijet balance analysis described in Section 3.4.4. In the same section, the Z+jet, γ+jet, and multijet simulated samples serve to evaluate residual differences between measured and simulated response.

3.4.2 Offset Corrections

This first step of the JEC accounts for energy collected within an event that does not originate from the principal high-p_T scattering reaction, but from additional pp scatters, i.e. PU, in the same (*in-time*, IT) or neighbouring (*out-of-time*, OOT) bunch crossings or from electronic noise. An event-by-event identification of PU contributions is possible only for charged particles, which appear as reconstructed tracks that are not associated to the primary vertex triggering the readout of the event. For illustration, Fig. 3.9 shows the event display of an ATLAS event, zoomed in on the 25 reconstructed primary vertices, one of which is at the origin of two high-p_T muons represented by two thick yellow lines.

Within the PF concept of CMS, the tracks and four-momenta of the charged hadrons identified as PU can be removed from the jets. When this technique of *charged hadron subtraction* (CHS) has been applied, the corresponding jets are la-

belled as "PF+CHS". The remaining offset energy (or all of it in case of PF jets) is corrected on average by employing jet areas, A, described in the previous chapter. Ignoring the very rare case of two or more simultaneous high-p_{T} scatters, the extra energy from PU collisions can be assumed to be isotropically distributed over the whole event and to correspond to an offset energy density ρ, such that $(\rho \cdot A_j)$ gives a first estimate of the energy to be subtracted from jet j. However, in hadron collisions soft particles are also produced through multiple parton interactions within the primary proton-proton scatter and contribute to the soft energy pedestal via the underlying event. Following Ref. [14] the soft energy dispersed by the UE must not be subtracted from the jets, since a differentiation between particles of the hard subprocess and the UE cannot be made unambiguously. A potential oversubtraction of this component, ρ_{UE}, has to be estimated from events without PU collisions and must be added back. Written in the form of a multiplicative correction to the uncorrected jet p_{T}, $p_{\mathrm{T,uncorr}}$, the factor reads:

$$C_{\mathrm{jetarea}}(p_{\mathrm{T,uncorr}}, A_j, \rho) = 1 - \frac{(\rho - \langle \rho_{\mathrm{UE}} \rangle) \cdot A_j}{p_{\mathrm{T,uncorr}}}\,. \tag{3.4}$$

The input parameters of this *jet area method* are $p_{\mathrm{T,uncorr}}$, the jet area A_j, and the per-event offset density ρ. At Tevatron, where the concept of jet areas initially was not known, the alternative *average offset method* has been developed [26]. Here, the correction depends on the number of primary vertices, N_{PV}, reduced by one for the high-p_{T} scatter, and a pseudorapidity dependent factor $O(\eta)$:

$$C_{\mathrm{offset}}(\eta, p_{\mathrm{T,uncorr}}, N_{\mathrm{PV}}) = 1 - \frac{(N_{\mathrm{PV}} - 1) \cdot O(\eta)}{p_{\mathrm{T,uncorr}}}\,. \tag{3.5}$$

Both, the offset density ρ and N_{PV}, scale almost linearly with the average number of PU interactions per bunch crossing, $\langle \mu \rangle$. Integrating the η dependence of the PU contributions as observed at the Tevatron into the jet area method, both approaches can be unified to give the *hybrid jet area method* as applied by CMS for the data taken at $\sqrt{s} = 7\,\mathrm{TeV}$ [27]:

$$C_{\mathrm{hyb7}}(p_{\mathrm{T,uncorr}}, \eta, A_j, \rho) = 1 - \frac{[(\rho - \langle \rho_{\mathrm{UE}} \rangle) \cdot \beta(\eta)] \cdot A_j}{p_{\mathrm{T,uncorr}}}\,, \tag{3.6}$$

where $(\rho - \langle \rho_{\mathrm{UE}} \rangle)$ and $(N_{\mathrm{PV}} - 1)$, and $O(\eta)$ and $\beta(\eta) \cdot A_j$ play similar roles. The four to five times larger amount of data available at 8 TeV centre-of-mass energy allowed the JEC to be complemented with a logarithmic jet p_{T} dependence for the PU offset. The final correction, where the UE density parameter ρ_{UE} has been absorbed into $\rho_0(\eta)$, reads [9]

$$C_{\mathrm{hyb8}}(p_{\mathrm{T,uncorr}}, \eta, A_j, \rho) = 1 - \frac{[\rho_0(\eta) + \rho \cdot \beta(\eta) \cdot (1 + \gamma(\eta) \cdot \log(p_{\mathrm{T,uncorr}}))] \cdot A_j}{p_{\mathrm{T,uncorr}}} \tag{3.7}$$

and depends on the four input variables jet $p_{\mathrm{T,uncorr}}$, jet η, jet area A_j, and per-event offset density ρ.

The three parameters $\rho_0(\eta)$, $\beta(\eta)$, and $\gamma(\eta)$ provide the necessary flexibility towards the jet η and p_{T} dependent final correction. They are determined from the simulation of samples of QCD dijet events processed with and without overlay of minimum bias events that serve to model PU collisions. Because the average number of PU interactions per bunch crossing in data is not known beforehand, simulations are performed for various hypotheses on $\langle\mu\rangle$ that are combined later to match the actual conditions during data taking. Figure 3.10 left shows the average simulated particle-level offset in jet p_{T}, $\langle p_{\mathrm{T,PU}} - p_{\mathrm{T,noPU}}\rangle$, versus particle-jet p_{T} in the central detector for increasing values of μ. Per PU collision simulation predicts about half a GeV of additional p_{T} captured by a jet. The residual offset after applying the correction remains below that limit.

To account for differences between data and simulation, the method of *random cones* (RC) is invoked. The RC method covers the whole (η,ϕ) space of a ZB event by randomly placed jet cones. As these events are not triggered by any specific detector signal, they generally do not contain any high-p_{T} but PU collision products as well as detector noise. Assuming the noise energy contribution to be negligible with respect to the PU one, the average p_{T} of the jets as measured from the RC method indicates the average energy offset due to PU, for the considered jet algorithm and jet size parameter. Figure 3.10 right presents the RC offset per type of PF candidate as determined in data and from MC simulation, normalised by the average number of PU interactions $\langle\mu\rangle$. The offset fraction labelled "charged hadrons" does not affect PF+CHS jets. The bottom panel shows for PF and PF+CHS jets the ratio of data over simulation, i.e. the scale factor to apply to data to correct for residual differences between the PU offset in data and in simulation. Within the tracker coverage up to $|\eta| = 2.5$ this scale factor, which is larger for PF+CHS jets, is within $\pm 5\%$ around unity, demonstrating the good modelling of PF candidates in simulation for PF or PF+CHS jets. Beyond the reach of the ECAL endcap at $|\eta| = 3.0$ only electromagnetic and hadronic energy deposits in the HF calorimeter can be differentiated leading to scale factors differing by up to 20% from unity.

Assuming on average 20 PU collisions, i.e. $\langle\mu\rangle = 20$, and adding together all sources of systematic uncertainty in the offset correction, the PU uncertainty of the JEC exceeds the level of two per mille only below about 50 GeV in jet p_{T} and is negligible above compared to other sources of systematic uncertainty, cf. Figs. 3.20.

3.4.3 Simulated Response Corrections

The benefits of deriving the bulk of the JEC from simulation are manifold. First of all, the accessible phase space is not limited by data availability and initial corrections can be derived without any data. Secondly, biases present in data-based methods are avoided, and finally the understanding of these methods and inherent correlations are facilitated.

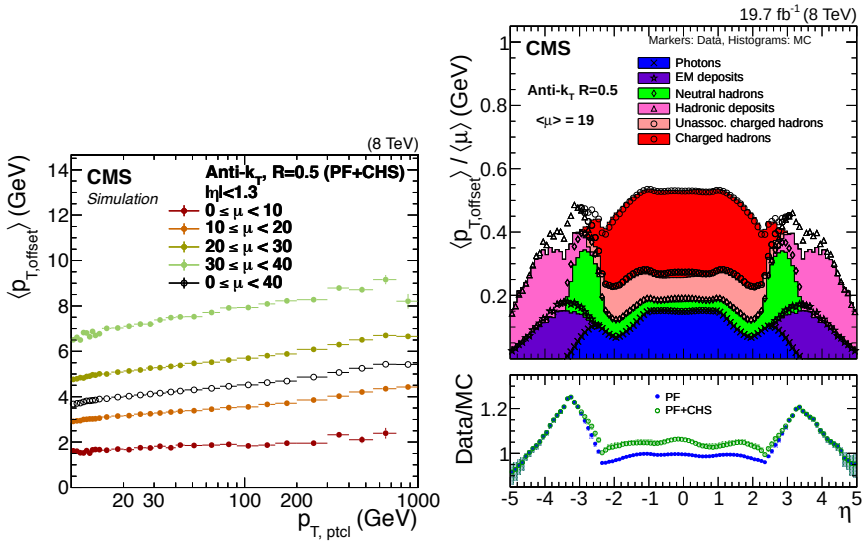

Fig. 3.10 Average simulated particle-level offset in jet p_{T}, $\langle p_{\mathrm{T,PU}} - p_{\mathrm{T,noPU}} \rangle$, (left) for $|\eta| < 1.3$ versus particle-jet p_{T}, $p_{\mathrm{T,ptcl}}$, for different values of average number of PU interactions per bunch crossing, $\langle \mu \rangle$. Random-cone offset per type of PF candidate (right) as determined in data (markers) and from MC simulation (histograms), normalised by the average number of PU interactions $\langle \mu \rangle$. The offset fraction labelled "charged hadrons" does not affect PF+CHS jets. The bottom panel shows for PF and PF+CHS jets the ratio of data over simulation, i.e. the scale factor to apply to data to correct for residual differences between the PU offset in data and in simulation. *(Taken from Ref. [9])*

The simulated jet response is deduced from a large dijet sample generated with PYTHIA6 tune Z2*, where a particle-level jet is matched to the closest reconstructed jet if their distance is smaller than half the jet size parameter R. This method ensures a high matching efficiency and provides a unique match for the anti-k_{t} jets. In the following, jets are assumed to be corrected for the PU offset as described in the previous section. Then the simulated particle response R_{ptcl} is defined as the ratio of arithmetic means of matched reconstructed and particle-level jet transverse momenta,

$$R_{\mathrm{ptcl}}(\langle p_{\mathrm{T}} \rangle, \eta) = \frac{\langle p_{\mathrm{T}} \rangle}{\langle p_{\mathrm{T,ptcl}} \rangle}[p_{\mathrm{T,ptcl}}, \eta], \tag{3.8}$$

in bins of particle-level p_{T} ($p_{\mathrm{T,ptcl}}$) and reconstructed η, where p_{T} is the transverse momentum of the reconstructed jet. The corrections are determined as a function of p_{T} and η for various jet algorithms and sizes and for the QCD dijet flavour mixture. Figure 3.11 left presents as an example the simulated response R_{ptcl} of $R = 0.5$ anti-k_{t} PF+CHS jets versus η for a series of jet p_{T} values. Below pseudorapidities of $|\eta| = 3.0$ and above jet p_{T}'s of 30 GeV more than 90% of the original p_{T} are retained. Beyond $|\eta| = 3.0$ and in particular below 30 GeV a much larger portion of jet p_{T} up to 30% is lost. Figure 3.11 right shows the response after the correction in

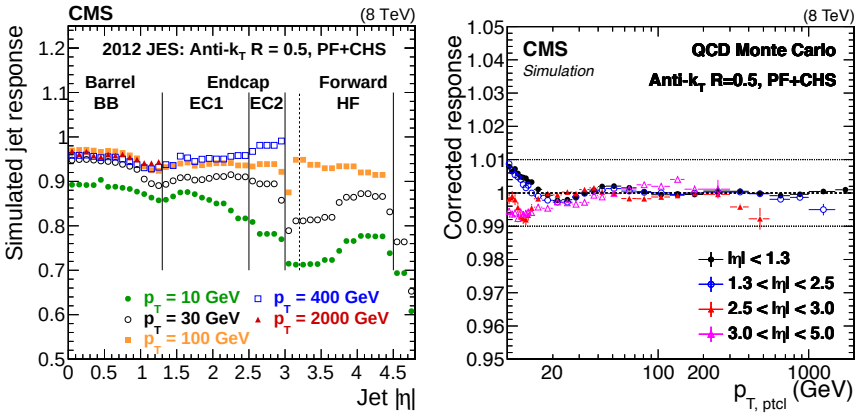

Fig. 3.11 Simulated jet response R_{ptcl} versus $|\eta|$ for $R = 0.5$ (left). Simulated jet response R_{ptcl}, after jet energy corrections have been applied, versus $p_{\mathrm{T,ptcl}}$ for $R = 0.5$ in various η regions (right). *(Taken from Ref. [9])*

the four regions of absolute pseudorapidity defined in the left plot and demonstrates that full response with respect to the particle-jet can be recovered to within $\pm 0.5\%$ for jet p_{T}'s from about 20 GeV up to 2 TeV.

Since PU corrections, which are proportional to the jet area, have been performed already, this significant R^2 dependence of the JEC is removed. Residual differences, visible for the correction factors at a jet p_{T} of 30 GeV in Fig. 3.12 left, are caused primarily by two effects:

- Soft UE contributions to the jet p_{T} have a lower response than the contributions from the hard scattering that are concentrated at the core of a jet. For jets with large jet sizes R the response at low p_{T} is therefore reduced leading to corrections that increase with R. The smallest considered jet size of $R = 0.3$ is an exception because of the detector granularity that becomes important.
- For larger jet sizes, sharp transitions in the detector geometry and hence steps in the response versus η are more averaged out than for narrower jets.

Figure 3.12 right presents the simulated jet responses after the JEC demonstrating agreement with unity within $\pm 1\%$ for jet p_{T}'s larger than 30 GeV in the central detector region of $|\eta| < 1.3$.

The jet response corrected in this step is sensitive to the detector calibrations entering the simulation model. The response of the CMS calorimeters to single pions has been calibrated in test beam studies to within an uncertainty of $\pm 3\%$ in the barrel region, which has subsequently been checked and confirmed in pp collision data with charged pions [28]. Thanks to the inclusion of track measurements within the Particle Flow concept, the PF jet response is much less sensitive to the underlying single-particle calibration within the tracker coverage as compared to relying on the calorimeters alone. Of course, this advantage is most prominent for low-p_{T} PF jets and diminishes for straight, high-p_{T} tracks. The uncertainty of the underlying

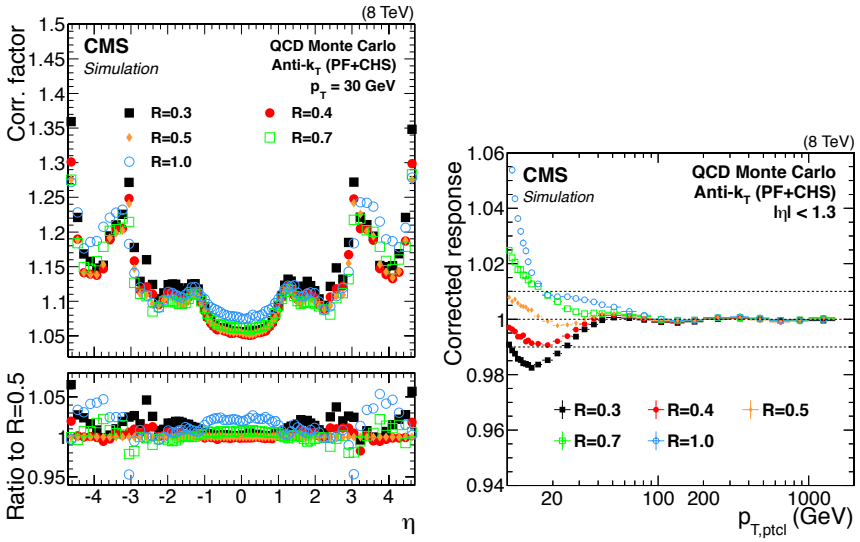

Fig. 3.12 Jet energy correction factors for a jet p_T of 30 GeV as a function of η and for various jet sizes R (left). Simulated jet energy response R_{ptcl} after jet energy corrections as a function of the jet p_T for various jet sizes R in the central detector region of $|\eta| < 1.3$ (right). *(Taken from Ref. [9])*

detector calibration is propagated to the JEC by employing a parameterised fast simulation of the CMS detector (FASTSIM) [29].

Uncertainties arising from the modelling of jet fragmentation are estimated by comparison of the PYTHIA6 samples with HERWIG++ ones.

3.4.4 Residual Corrections for Data

Simulated response corrections suffer from the imperfect modelling of real detectors and therefore must be complemented with data-based corrections. Such residual corrections are measured by exploiting transverse momentum balance to calibrate a jet against more precisely measured reference objects. For reasons of statistical and systematic precision, the residual corrections for data are performed in two steps:

1. In the first step, the detector response is equalised versus pseudorapidity by correcting all jets relative to reference jets in the barrel region $|\eta| < 1.3$. For this purpose, dijet events with at least one central jet are used, which are available in sufficient numbers for jet p_T's up to 1 TeV.
2. In the second step, pairs of unlike-sign high-p_T leptons or a photon corresponding to the topologies of $Z(\rightarrow \mu\mu)$+jet, $Z(\rightarrow ee)$+jet, and γ+jet events are exploited to derive absolute calibration factors. Sufficient amounts of data have been acquired for this purpose within the region of $30 < p_T < 800$ GeV. Be-

yond 800 GeV multijet events are used to constrain one high-p_{T} jet against the recoil of two or more lower p_{T} jets.

Two methods have been developed to assess transverse momentum imbalances in these topologies. In the p_{T}-*balance* method the response of the recoiling jet is determined by a direct comparison to the p_{T} of a reference object. In contrast, the *missing transverse momentum projection fraction* (MPF) method [27] considers the hadronic activity in the whole event to estimate a potential response mismatch. The response for each case is then defined as:

$$R_{\mathrm{jet},p_{\mathrm{T}}} = \frac{p_{\mathrm{T,jet}}}{p_{\mathrm{T,ref}}}\,, \tag{3.9}$$

$$R_{\mathrm{jet,MPF}} = 1 + \frac{\vec{\not{p}}_{\mathrm{T}} \cdot \vec{p}_{\mathrm{T,ref}}}{(p_{\mathrm{T,ref}})^2}\,. \tag{3.10}$$

For the calculation of the missing transverse momentum $\vec{\not{p}}_{\mathrm{T}}$, PF jets up to $|\eta| = 5.0$ and above $p_{\mathrm{T}} = 10\,\mathrm{GeV}$ are considered after applying all preceding corrections. Differences in the two response determinations are used for consistency checks and uncertainty estimates.

Clearly, the existence of jets beyond the ones required by the requested event topology invalidate the assumption of p_{T} balance between the objects under consideration. Therefore, the corrections must be studied as a function of the additional jet activity that is quantified through the observable α, which is defined as the ratio of the p_{T} of the highest-energetic extra jet over the p_{T} relevant for the balancing. So for dijet events, $\alpha = p_{\mathrm{T},3}/p_{\mathrm{T,ave}}$, while for $\mathrm{Z}(\to \mu\mu)$+jet, $\mathrm{Z}(\to ee)$+jet, and γ+jet events $\alpha = p_{\mathrm{T},2}/p_{\mathrm{T},\gamma/\mathrm{Z}}$. In these formulae, $p_{\mathrm{T,n}}$ refers to the n-th highest-energetic jet in the event, $p_{\mathrm{T,ave}}$ is the average p_{T} of the leading two jets, and $p_{\mathrm{T},\gamma/\mathrm{Z}}$ denotes the p_{T} of the photon respectively the Z boson. The measured responses are extrapolated to the limit of zero extra activity, i.e. $\alpha \to 0$, to address only genuine detector effects on the jet energy response.

3.4.4.1 Relative η-dependent corrections

The relative η-dependent corrections to the jet response are obtained from dijet events with the reference or *tag* jet restricted to the central barrel region of $|\eta| < 1.3$, and the *probe* jet unconstrained in pseudorapidity. The bias caused in this method by the rather poor jet energy resolution is reduced by binning in average jet p_{T}, $p_{\mathrm{T,ave}} = 0.5 \cdot (p_{\mathrm{T,tag}} + p_{\mathrm{T,probe}})$, instead of $p_{\mathrm{T,tag}}$. At the same time, biases from ISR+FSR are cancelled to first order.

Because generally the average of a ratio is unequal to the ratio of averages, unless the denominator is almost constant, $p_{\mathrm{T,ave}}$ must appear in the denominator. Hence, the following formulae define the corresponding responses from dijet events for both the p_{T}-balance and MPF method:

$$R_{\text{rel}}^{p_{\text{T}}} = \frac{1+\langle\mathcal{A}\rangle}{1-\langle\mathcal{A}\rangle}, \quad \text{where} \tag{3.11}$$

$$\mathcal{A} = \frac{p_{\text{T,probe}} - p_{\text{T,tag}}}{2p_{\text{T,ave}}}, \quad \text{and} \tag{3.12}$$

$$R_{\text{rel}}^{\text{MPF}} = \frac{1+\langle\mathcal{B}\rangle}{1-\langle\mathcal{B}\rangle}, \quad \text{where} \tag{3.13}$$

$$\mathcal{B} = \frac{\vec{p}_{\text{T}} \cdot (\vec{p}_{\text{T,tag}}/p_{\text{T,tag}})}{2p_{\text{T,ave}}}. \tag{3.14}$$

Extrapolating to zero additional jet activity, i.e. $\alpha \to 0$, and binning sufficiently fine in $p_{\text{T,ave}}$, both variables $R_{\text{rel}}^{p_{\text{T}}}$ and $R_{\text{rel}}^{\text{MPF}}$ reduce to $R_{\text{rel}} = \langle p_{\text{T,probe}}\rangle / \langle p_{\text{T,tag}}\rangle$. Assuming that on average the tag and probe jet p_{T}'s are equal for particle-jets, which holds after correcting for small second-order biases from ISR+FSR and from the jet energy resolution (JER), cf. Section 3.6, this is equivalent to the ratio R_{rel} of jet response $R_{\text{jet,probe}}$ over $R_{\text{jet,tag}}$.

Figure 3.13 left shows the relative η- and p_{T}-dependent factor $R_{\text{rel,MC}}/R_{\text{rel,data}}$ that corrects for residual differences between data and simulation. The curve for $p_{\text{T}} = 480\,\text{GeV}$ stops at $\eta \approx 2.8$ corresponding to a jet energy of $E = p_{\text{T}} \cosh\eta = 4\,\text{TeV}$ thereby restricting the correction to a reasonable range with respect to the collected data. In the barrel reference region of $|\eta| < 1.3$ the factor varies from 0.99 up to 1.01. In the endcap region up to $|\eta| = 2.9$, it rises to 1.06, and for the HF even up to 1.15. A pronounced p_{T} dependence is observed beyond the barrel region with smaller corrections for increasing p_{T}'s as expected from reduced nonlinearities of the calorimeters.

Detailed studies have been performed to correct for biases caused by ISR+FSR and JER effects in the data-based relative calibration. They are reported in Ref. [9]. The total uncertainty on the JEC by this relative correction varies from below 0.5% in the barrel reference region up to about 2.0% for the outer pseudorapidities as shown in Fig. 3.13 right for a jet p_{T} of 100 GeV.

3.4.4.2 Absolute corrections

The data-based absolute calibration of jet energies is performed in the barrel reference region $|\eta| < 1.3$ for jet p_{T}'s between 30 and 1300 GeV. Sufficient amounts of data are at disposal for the $\text{Z}(\to \mu\mu)$+jet and $\text{Z}(\to ee)$+jet topologies between 30 and 400 GeV. In the γ+jet channel the reach is from 40 up to 800 GeV. Beyond 800 GeV the response is constrained using multijet events, where a high-p_{T} jet recoils against a system of two or more lower-p_{T} jets. Within CMS the correction is done in two steps. First, the most precise calibration channel based on $\text{Z}(\to \mu\mu)$+jet events is used to derive a global factor that corrects for the residual mismatch between data and simulation. In a second step, an observed p_{T} dependence of this absolute calibration is taken into account in a simultaneous fit of the individual responses of all channels. To profit from the different sensitivities of the

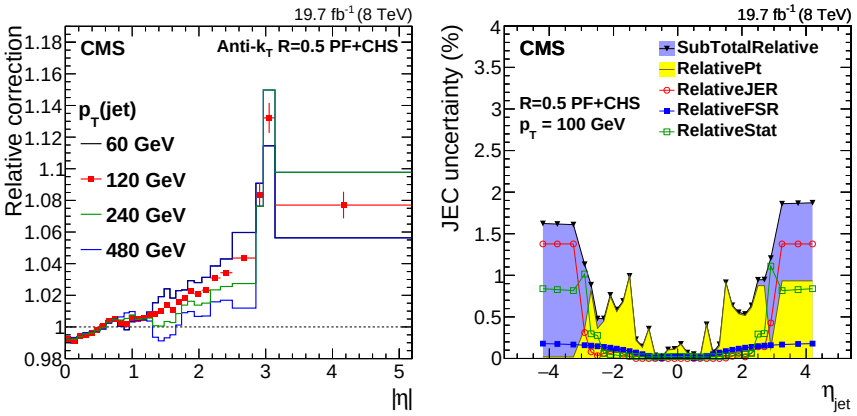

Fig. 3.13 Left: Relative η-dependent energy scale correction $R_{\mathrm{rel,MC}}/R_{\mathrm{rel,data}}$ for $p_{\mathrm{T}} = 60$, 120, 240, and 480 GeV jets. The residual corrections increase towards high η and low p_{T}, where effects from nonlinear calorimeter response become more important. Right: JEC uncertainty of the relative correction versus η. The split-up into uncertainty sources caused by p_{T}-dependence, ISR+FSR, JER, and statistical effects is shown. *(Taken from Ref. [9])*

p_{T}-balance and MPF method to systematic biases, these steps are performed with both approaches.

This absolute calibration is equally affected by a bias through ISR+FSR as the relative one, and subdominantly also by UE and OOC showering effects. In contrast, the energy resolution does not induce a large bias, because the binning is made in terms of the boson's p_{T}, which exhibits a much better resolution than any jet p_{T}. To account for the ISR+FSR bias the jet response is investigated as a function of the suppression of additional jet activity that is quantified in terms of the $\alpha = p_{\mathrm{T},2}/p_{\mathrm{T},\gamma/\mathrm{Z}}$ parameter. Figure 3.14 left shows the corresponding jet responses for data and simulation in both data-based methods, p_{T}-balance and MPF, and their ratios data over simulation in the lower panel. In all cases a linear dependence on α is exhibited that can be extrapolated to $\alpha = 0$ starting from the working point of $\alpha = 0.3$. This extrapolation is indicated in Fig. 3.14 left by lines with shaded bands representing the statistical fit uncertainty. Obviously, the p_{T}-balance method, which ignores event components other than the tag and probe objects, is much more sensitive to additional jet activity than the MPF approach that includes the whole event. As a consequence the relative resolution of the MPF method is superior as compared to the p_{T} balance, particularly so at high p_{T} beyond 100 GeV. On the contrary, additional PU jets at low p_{T} pose a problem for the MPF approach, while these contributions are ignored in the p_{T} balance leading to a superior resolution of the latter method. In summary, both methods give consistent results and complement each other in terms of their performance in the considered p_{T} range.

The simulated response $p_{\mathrm{T,jet}}/p_{\mathrm{T,ptcl}}$, indicated in Fig. 3.14 left with green stars, exceeds unity despite the use of already MC-calibrated jets i.a. because of the lower jet response on average in dijet events as compared to a Z+jet event sample, which

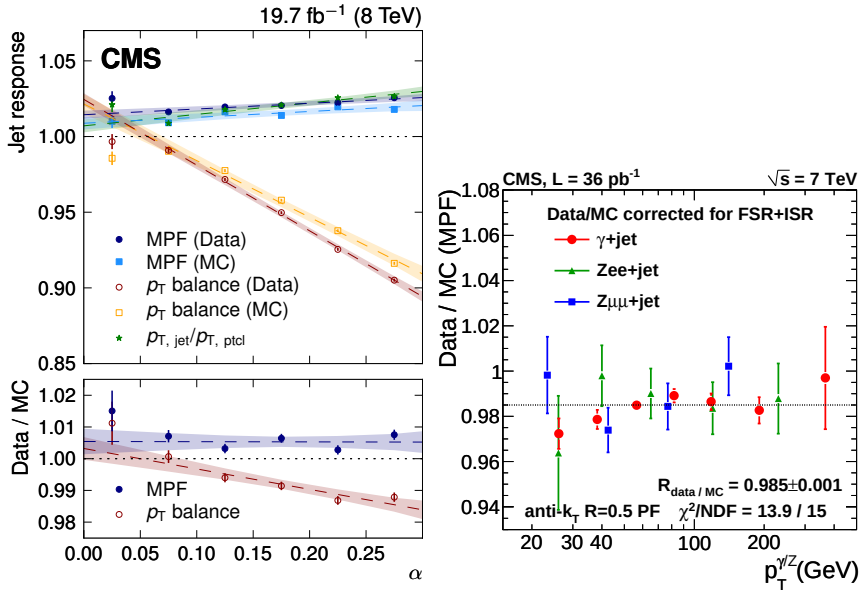

Fig. 3.14 Left: Jet response (upper panel) as a function of the parameter α, delimiting additional jet activity, as obtained in the barrel reference region from simulation (MADGRAPH +PYTHIA6 tune Z2*) and from the MPF and p_T-balance methods in Z($\rightarrow \mu\mu$)+jet events for data and simulation (points). The response in data is scaled by a factor of 1.02, constant as a function of p_T. The simulated jet response $p_{T,jet}/p_{T,ptcl}$ is indicated as well. The dashed lines with shaded bands correspond to linear fits of the α dependence with statistical fit uncertainties. The ratio of the jet response from data and simulation is presented in the bottom panel for both methods. Right: Final data over simulation correction from the MPF method applied to Z+jet and γ+jet event samples of the $\sqrt{s} = 7$ TeV data of 2010. *(Taken from Refs. [9, 27])*

contains a smaller fraction of gluon-initiated jets. This bias is treated in the jet flavour corrections in Section 3.4.5.

After dealing with the remaining effects of UE and OOC, which affect the p_T-balance and MPF methods slightly differently, the response for Z($\rightarrow \mu\mu$)+jet events is found to be lower by about 2% in data than in simulation, in agreement between the two approaches. This global correction factor of 1.02 also coincides with previous findings for the data taken at 7 TeV centre-of-mass energy [27]. The corresponding result for the MPF method is shown in Fig. 3.14 right for Z+jet and γ+jet event samples. In the following, the response in data is presumed to have been scaled by this global correction factor of 1.02, constant as a function of p_T.

The multijet balance (MJB) exploits events where one high-p_T jet recoils against two or more lower p_T jets that have been calibrated already. This technique, introduced by ATLAS in Ref. [30], defines as response

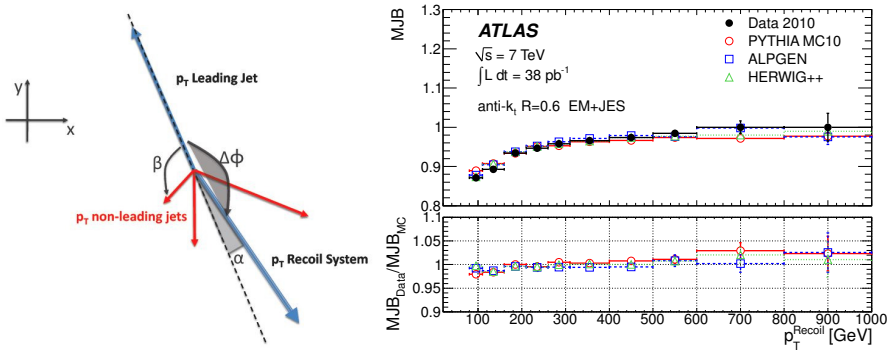

Fig. 3.15 Left: Sketch of the multijet event topology used for calibration purposes. Right: Multijet response MJB from data and simulation as a function of the p_T of the recoil system from the ATLAS analysis. Only statistical uncertainties are provided. The response ratio of data over simulation is shown in the lower panel. *(Taken from Ref. [30])*

$$\mathrm{MJB} = \frac{|\vec{p}_{\mathrm{T},1}|}{|\vec{p}_{\mathrm{T,recoil}}|}, \tag{3.15}$$

where $\vec{p}_{\mathrm{T,recoil}}$ is the vectorial sum of the momenta of all nonleading jets. To ensure a clear one-to-many balancing topology as sketched in Fig. 3.15 left and to confine the recoil jets to a p_T range that has been calibrated already, the following selection criteria must be fulfilled:

1. The azimuthal angular separation $\Delta\phi$ between the leading jet to be calibrated and the recoil system must differ from π by less than 0.3: $|\Delta\phi - \pi| < 0.3$.[3]
2. The azimuthal angular separation β, cf. Fig. 3.15 left, between the leading jet and the closest nonleading jet of the recoil system must exceed unity: $\beta > 1$.
3. The proportion of the next-to-leading jet's p_T, $p_{T,2}$, to the p_T of the recoil system, named asymmetry ratio A in Ref. [30], must not exceed 60%: $A = p_{T,2}/p_{\mathrm{T,recoil}} < 0.6$.

The resulting responses for data and simulation, shown for the analysis by ATLAS in Fig. 3.15 right, upper panel, start at about 90% for 100 GeV of $p_{\mathrm{T,recoil}}$ and approach unity for recoil systems with about 600–1000 GeV of p_T. This behaviour is described to within ±3% by MC simulation using PYTHIA, ALPGEN, or HERWIG++ with a small positive slope of the data-over-MC ratio towards higher p_T's, as visible from the lower panel of Fig. 3.15 right.

CMS extended this technique to include the MPF method with the response defined as [9]:

$$\mathrm{MPF} = 1 + \frac{\left(\vec{p}_{\mathrm{T}} \cdot \vec{p}_{\mathrm{T,recoil}}\right)}{\left|\vec{p}_{\mathrm{T,recoil}}\right|^2}. \tag{3.16}$$

[3] In Fig. 3.15 left, $|\Delta\phi - \pi|$ is defined as α, which is not used here to avoid confusion with the previously defined suppression α of additional jet activity.

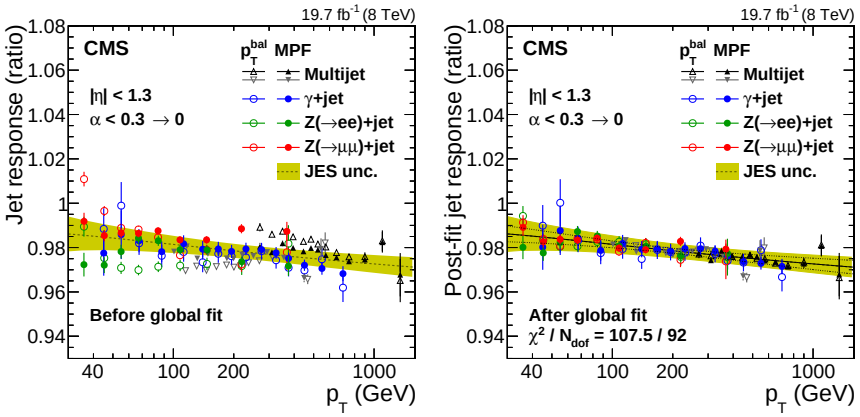

Fig. 3.16 Left: Comparison of the data-to-simulation ratio of the jet response measurements from $Z(\rightarrow \mu\mu)$+jet, $Z(\rightarrow ee)$+jet, γ+jet, and multijet samples after all preceding corrections for the p_T-balance and the MPF method. The uncertainty in the ratio, excluding jet-flavor effects, is shown by the shaded region. Right: Same as left plot, but after application of nuisance parameter shifts as result of the simultaneous fit. The solid line represents the central fit result with the statistical fit uncertainties indicated by dotted lines. *(Taken from Ref. [9])*

Within the kinematic constraints of the CMS JEC, $\vec{\not{p}}_T$ is corrected for all jets with $|\eta| < 5$ and $p_T > 10\,\text{GeV}$, but only jets with $p_T > 30\,\text{GeV}$ are used for the event selection and the construction of $\vec{p}_{T,\text{recoil}}$. The latter condition ensures that the event selection is not biased by PU jets, and that the recoil is composed of jets directly calibrated with data-based methods.

If the MJB constraints are to be used in a simultaneous response fit together with the Z+jet and γ+jet topologies instead of just extending their reach in jet p_T, then a more complicated setup is required. To perform such a fit for the final result on the absolute scale correction for data, including a p_T-dependent term, CMS evaluates the effective p_T of the jets in the recoil system, which is found to lie between 35 to 45% of $p_{T,\text{recoil}}$ and is a measure of the lever arm of the MJB technique towards higher jet p_T [9]. The outcome of the χ^2 fit with nuisance parameters is demonstrated in Fig. 3.16. The left plot shows the data-to-simulation ratio of the jet response measurements from $Z(\rightarrow \mu\mu)$+jet, $Z(\rightarrow ee)$+jet, γ+jet, and multijet samples after all preceding corrections including the constant factor of 1.02 for data, but before performing the simultaneous fit. The right plot demonstrates a much improved agreement amongst the post-fit responses that are shifted according to the fitted nuisance parameters.

Systematic uncertainties beyond differences in the two methods and the bias corrections originate from the corresponding energy/momentum scale uncertainties of the reference objects. The tracker scale uncertainty for muons in $|\eta| < 2.4$ amounts to 0.2% [31], the ECAL scale uncertainty corresponds to 0.5% for electrons within $|\eta| < 2.4$ [32], and to 0.2% for photons within $|\eta| < 1.3$ [33]. The peaks of the invariant-mass distributions of the $Z(\rightarrow \mu\mu)$+jet and $Z(\rightarrow ee)$+jet events is used to

validate the muon (electron) energy scale between data and simulation. These are found to agree within the assumed 0.2 respectively 0.5%.

3.4.5 Jet Flavor Corrections

The amount of quark- or gluon-initiated jets in an event sample has an important impact on the jet response and thereby on the JEC. Based on the larger effective squared colour charge of a gluon, C_A, as compared to that of a quark, C_F, pQCD predicts that gluon-initiated jets are broader in shape with a higher particle multiplicity on average [34]. As a consequence, the wider shape of gluon jets, which entail higher OOC losses for a particular jet of size R, and the higher average multiplicity, which implies the presence of more low-energetic particles, both lead to the expectation of lower jet responses for gluon- than for quark-initiated jets. Therefore such differences have to be examined and controlled for the purpose of JEC.

It must be emphasised, however, that the concept of flavoured jets is not necessarily well-defined beyond LO. As illustrated for example in Fig. 3.17, a $q\bar{q}g$ event can be seen as an NLO correction to the process $q\bar{q} \to q\bar{q}$ via the final-state radiation of a gluon, or as the NLO correction to the process $q\bar{q} \to gg$ with a $g \to q\bar{q}$ splitting in the final state. Since the two final states are indistinguishable by first principles of quantum mechanics, the classification of such an event to either a $q\bar{q}$ or a gg dijet event is ill-defined. For a discussion on an infrared-safe definition of jet flavour Ref. [35] may be consulted. Some observables sensitive to differences in quark- respectively gluon-initiated jets are presented in Section 6.7. In the following the LO-based "physics definition" is used, where each jet is matched to the nearest parton of the LO high-p_T scattering process within $\Delta R = 0.25$. Unassociated jets, e.g. from high-p_T gluon radiation, are labelled as "unmatched". For the purpose of b-tagging sometimes an "algorithmic definition" is employed, where jets with heavy hadrons from gluon splitting are reclassified as b- or c-quark jets. This procedure, however, is in conflict with the parton shower concept of HERWIG++ and is not discussed further here.

For a first look into potential differences with respect to JEC purposes, Fig. 3.18 presents the jet-flavour fractions for the three event classes of dijet, Z+jet, and γ+jet production as derived from PYTHIA6 or MADGRAPH +PYTHIA6 MC predictions. The most striking feature is the more than twice as large gluon-jet fraction in the dijet sample as compared to the Z+jet or γ+jet samples in the whole accessible p_T range up to 800 GeV. But also the two boson+jet samples differ significantly, in particular below p_T's of 100 GeV, where presumably the mass of the Z boson plays a role. In that region, the γ+jet sample exhibits a much higher quark-jet fraction composed of light quarks and about 20% of directly produced charm jets.

Uncertainties arise at two occasions. First, differences might occur between the MC models in the predicted flavour composition of either the calibration sample, the signal sample, or both. Secondly, the jet response for a particular flavour might vary among the consulted theory predictions. Investigations by ATLAS and CMS [8, 9]

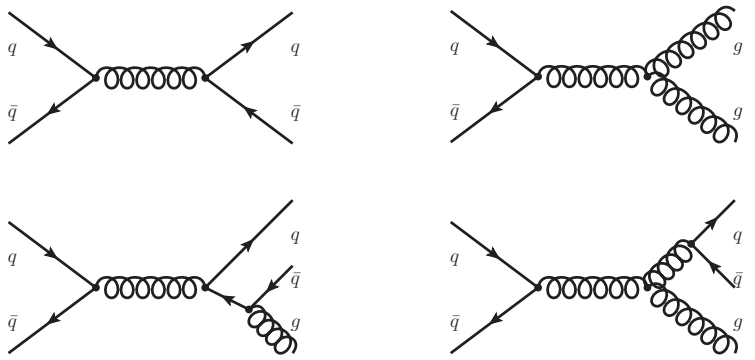

Fig. 3.17 LO Feynman diagrams for the processes $q\bar{q} \to q\bar{q}$ and $q\bar{q} \to gg$ (top row) and two NLO corrections (bottom row) with identical final state of $q\bar{q}g$.

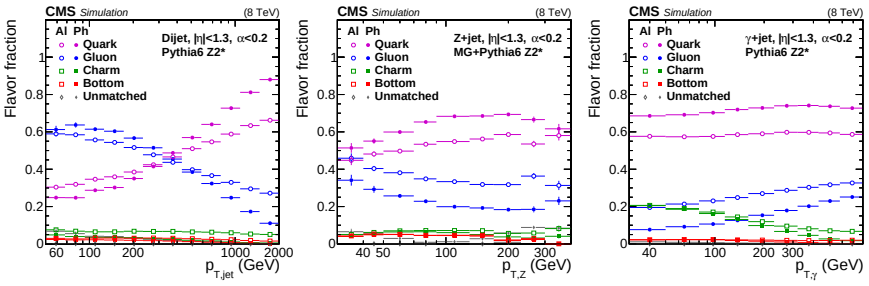

Fig. 3.18 Jet-flavour fractions as a function of reference object p_{T} for central jets from LO PYTHIA6 or MADGRAPH +PYTHIA6 MC predictions according to the two definitions given in the text for the dijet (left), Z+jet (middle), and γ+jet topologies (right) with additional jet activity suppressed via the condition $\alpha < 0.2$. *(Taken from Ref. [9])*

demonstrate that the dominant impact comes from large differences in the gluon-jet response between PYTHIA6 and HERWIG++. As an example, Fig. 3.19 left compares the responses as a function of jet p_{T} for various jet flavours from HERWIG++ tune EE3C and PYTHIA6 tune Z2*. The largest deviations varying from 3% at 30 GeV down to 1% at 1 TeV occur between the gluon-initiated jets of the two MC event generators. Figure 3.19 right presents the uncertainties ensuing from this discrepancy for the JEC as a function of jet p_{T} at $|\eta| = 0$ for light quarks, bottom quarks, gluons, and for typical quark-gluon jet mixtures. These uncertainties are applicable to data versus simulation comparisons regardless of whether or not optional jet-flavor corrections have been utilised with respect to the default corrections provided for dijet production.

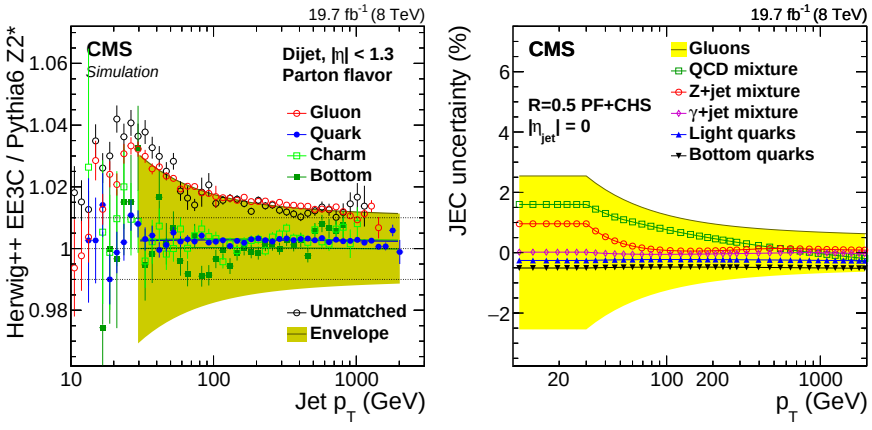

Fig. 3.19 Left: Comparison of jet responses from PYTHIA6 tune Z2* and HERWIG++ tune EE3C versus jet p_{T} for jet flavours according to the "physics definition". The symmetrised yellow band reflects the largest observed discrepancy for gluon jets. Right: JEC systematic uncertainty as a function of jet p_{T} attributed to typical jet-flavour mixtures (dijet, Z+jet, γ+jet) as well as light-quark, bottom-quark, and gluon jets at central pseudorapidity $\eta = 0$. *(Taken from Ref. [9])*

3.4.6 JEC Uncertainties

JEC uncertainties are derived for each calibration step as detailed in the previous sections and are provided by CMS for the purpose of physics analyses in the form of $k = 1,\ldots,K$ mutually uncorrelated sources as functions of jet p_{T} and η. Each such source k is represented through a relative $1\,\sigma$ shift s_{ki} that corresponds to a systematic effect fully correlated in p_{T} and η for any observation bin $i = 1,\ldots,N$. Partial or full decorrelation of a systematic effect within parts of the phase space is implemented by providing multiple sources spanning appropriately limited regions such that each source by itself is fully correlated and their sum in quadrature reproduces the desired degree of correlation for this systematic effect. The quadratic sum of all s_{ki} equals the total relative JEC uncertainty for any point i (or j) in phase space:

$$S_i = \sqrt{\sum_{k=1}^{K} s_{ki}^2}, \quad S_j = \sqrt{\sum_{k=1}^{K} s_{kj}^2}, \quad \rho_{ij} = \frac{\sum_{k=1}^{K} s_{ki} s_{kj}}{S_i S_j}, \tag{3.17}$$

where ρ_{ij} is the correlation coefficient of the uncertainty between the two points i and j. This approach works in complete analogy to the eigenvector method described before in the context of the PDF uncertainties 2.6.2, provided the relative uncertainties are properly scaled with the measurement, D_i, or the expectation from theory, T_i. Testing the goodness-of-fit using a least-squares method, the χ^2 between the N measurements D_i and their theoretical predictions T_i then is defined as:

$$\chi^2 = \sum_{i,j=1}^{N} (D_i - T_i)\, \mathrm{C}_{ij}^{-1}\, (D_j - T_j)\,, \tag{3.18}$$

where the covariance matrix C_{ij} accounts for correlations and is written as:

$$C_{ij} = \sum_{k=1}^{K} s_{ki} \cdot s_{kj} + \hat{\sigma}_{D_i} \hat{\sigma}_{D_j} \delta_{ij} \,. \tag{3.19}$$

Here, the additional term with $\hat{\sigma}_{D_i}$ and $\hat{\sigma}_{D_j}$ corresponds to the estimated statistical uncertainty of the measurement, which is uncorrelated and therefore only appears in the diagonal terms as indicated by Kronecker's delta function δ_{ij}.

As long as the χ^2 fit is unconstrained, it can be proven that the *covariance-matrix* formulation Eq. (3.18) can be written equivalently in a *best-fit* form using *nuisance parameters* [36, 37]:

$$\chi^2 = \sum_{i=1}^{N} \frac{(D_i - T_i \cdot (1 + \sum_{k=1}^{K} \varepsilon_k s_{ki}))^2}{\hat{\sigma}_{D_i}^2} + \sum_{k=1}^{K} \varepsilon_k^2 \,. \tag{3.20}$$

Correlations are taken into account by minimising the χ^2 with respect to the nuisance parameters ε_k, one of each was added per systematic source k. The denominator is composed of only an uncorrelated, statistical uncertainty, while the systematic effects appear with an additional *penalty* term $\sum_{k=1}^{K} \varepsilon_k^2$. The *a priori* expectation at the χ^2 minimum is that systematic shifts via the nuisance parameters ε_k are Gaussian distributed with zero mean and unity root-mean-square. Although the resulting χ^2 at minimum should be identical to the one from the covariance-matrix formulation Eq. (3.18), the latter method has the advantage to provide the nuisance parameters and the best-fit theoretical predictions for sanity checks. Deviations from the expected behaviour of the ε_k would point to problems in the modelling of the systematic uncertainty and the implied assumptions.

As recommended by CMS, it becomes explicit from the $T_i \cdot (1 + \sum_{k=1}^{K} \varepsilon_k s_{ki})$ term in the numerator of Eq. (3.18) that the JEC uncertainties should be propagated as multiplicative factors to the theory expectation T_i. Although some corrections, e.g. the one for the PU offset, could be considered as additive, in total the multiplication of T_i avoids the statistical bias that arises from uncertainty estimates taken from data [38–40].[4]

Figure 3.20 summarises the JEC uncertainty of CMS as a function of jet p_{T} (left) and jet η (right). Going from low to high-p_{T} jets, the dominant uncertainty sources on the experimental side are related to PU and the absolute scale calibration respectively. Below 200 GeV of p_{T} a large fraction of the uncertainty is caused by discrepancies in the theoretical modelling of the gluon jet response. Otherwise JEC uncertainties as small as 0.5% are achieved in the central detector for jet p_{T}'s of 80 GeV and beyond. For outer pseudorapidities beyond tracking and then endcap coverage, the relative η-dependent correction contributes about 1–2% to the total uncertainty. The additional source labelled "time stability" in Fig. 3.20 accounts for time-dependent variations in the ECAL and HCAL responses caused by radiation damage. This effect has been corrected on average for the total sample of un-

[4] Sometimes this effect is called *D'Agostini bias*.

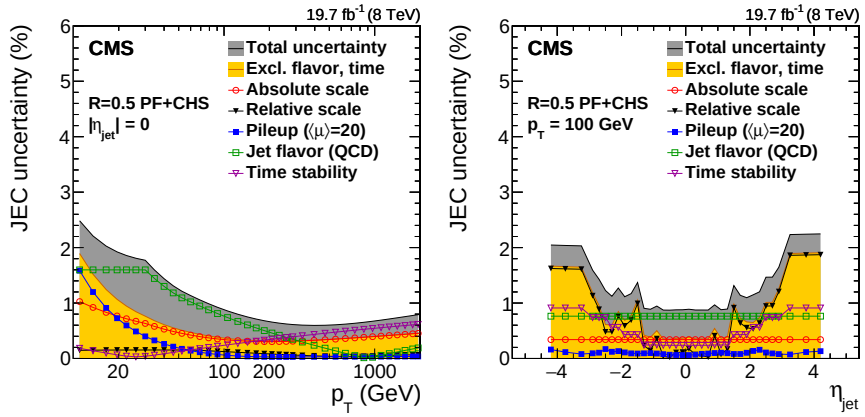

Fig. 3.20 Summary of fractional JEC systematic uncertainties with PF+CHS anti-k_t jets ($R = 0.5$) in CMS as a function of jet p_T at $|\eta| = 0$ (left) and as a function of jet η at $p_T = 100\,\text{GeV}$ (right). The uncertainty contributions from the different JEC steps are shown as markers. The summed up contributions excluding time-dependent and flavour effects, and the total JEC uncertainty are represented as yellow respectively grey bands. *(Taken from Ref. [9])*

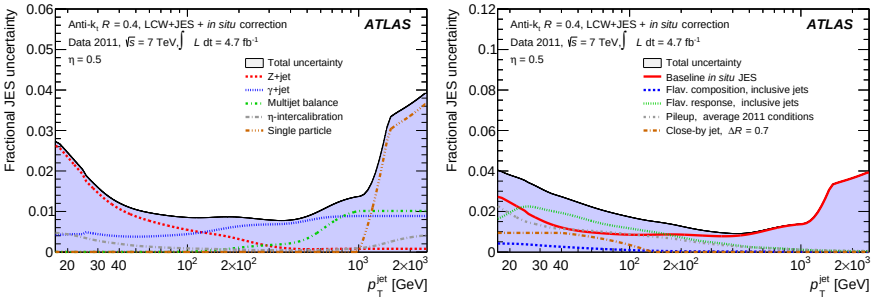

Fig. 3.21 Summary of fractional JEC systematic uncertainties with LCW+JES anti-k_t jets ($R = 0.4$) in ATLAS as a function of jet p_T at $|\eta| = 0.5$. Left: Uncertainty composition of the baseline JEC considering the relative η-intercalibration and the absolute calibration with Z+jet, γ+jet, and multijet event samples, and with single particles. Right: Fractional baseline JEC uncertainty in comparison to additional contributions caused by PU, flavour response and composition as appropriate for inclusive jet samples, and jet overlaps. *(Taken from Ref. [8])*

prescaled events from 2012. The time stability uncertainty therefore is to be applied only if dedicated or pre-scaled subsamples of the total 8 TeV data set are selected.

A similar behaviour and performance of the JEC has been achieved in ATLAS as shown in Fig. 3.21 for data at $\sqrt{s} = 7\,\text{TeV}$. Also here, a significant fraction of the total uncertainty below 200 GeV of jet p_T is caused through flavour effects and their theoretical modelling.

Another interesting aspect arises, when data from multiple experiments are to be used in a common fit. In that case, correlations between the systematic uncertainties of the experiments have to be carefully evaluated. This goes beyond the scope of

results discussed in this work. However, first studies on correlations between the jet measurements in ATLAS and CMS have been reported in Refs. [16] and [17] for 7 and 8 TeV centre-of-mass energy respectively.

3.5 Background Removal

Whenever parts of the detector or the data acquisition and reconstruction chain were not operating properly, the concerned luminosity sections of about 23 seconds duration in CMS are flagged as invalid and are removed from the amount of data available for physics analyses. In some cases the observed problem can be fixed in a later reprocessing and the respective data are recovered. In addition to this upstream validation further careful checks are performed on all events for the chosen set of triggers. The first filtering step consists in the removal of occasional nonphysical reconstruction objects. More precisely, to suppress jet-like signatures resulting from noise in the electromagnetic and/or hadronic calorimeters of CMS, PF jets are required to satisfy the following jet identification criteria. Each PF jet should contain at least two particle candidates, one of which is a charged hadron, and the jet energy fraction carried by neutral hadrons and photons should be less than 90%. These criteria have an efficiency of greater than 99% for physical jets, while the probability for a nonphysical jet to pass the criteria is less than 10^{-6} [20]. In addition, each accepted event must have at least one well-reconstructed proton-proton interaction vertex. Equivalent data-quality, jet-identification, and vertex requirements are imposed by ATLAS, cf. Ref. [41].

In a second step, the event sample is restricted to the desired phase space in jet p_{T} and rapidity y depending on each individual analysis. Because of the huge jet production cross section, backgrounds through other physics processes are usually negligible — a particularity of jet studies. In some cases, CMS requires as an additional safety measure the ratio of the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ over the scalar sum of transverse energy $\sum E_{\mathrm{T}}$ not to exceed 0.3. Figure 3.22 illustrates the effect by means of an inclusive jet (left) and a dijet selection (right). The event excess observed for large $E_{\mathrm{T}}^{\mathrm{miss}}/\sum E_{\mathrm{T}}$ values in the first case is caused by $Z/W+\mathrm{jet}$ production, where the $Z \to \nu\bar{\nu}$ and $W \to l\nu$ decays lead to the missing E_{T}, while for the latter selection the dijet requirement effectively removes this type of events. Since high-p_{T} jet production does not normally generate true $E_{\mathrm{T}}^{\mathrm{miss}}$, the background events can safely be removed by the indicated selection criterion. In special cases, however, like heavy-quark jets with leptonically decaying hadrons, or in particular phase space regions as described in the azimuthal decorrelation study in Section 6.5, additional considerations have to be made.

Finally, in many cases the highest p_{T} (or dijet mass) events are studied for any peculiarities using an event visualisation tool although this is not always explicitly mentioned in the relevant publications.

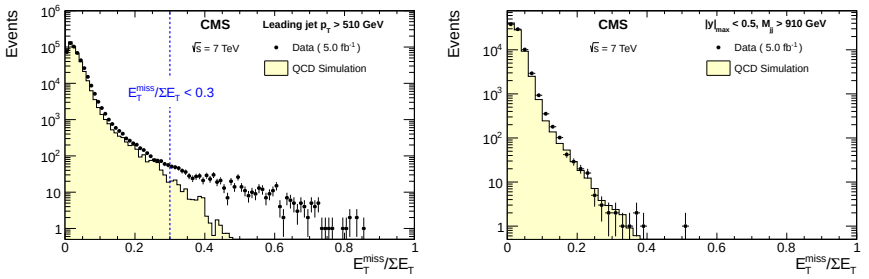

Fig. 3.22 Distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ over $\sum E_{\mathrm{T}}$ for data (black points) and simulated jet events (histograms) in an inclusive jet selection (left) and for a dijet event sample (right). The event excess for large $E_{\mathrm{T}}^{\mathrm{miss}}/\sum E_{\mathrm{T}}$ values in data on the left is caused by $Z/W+$jet production, where the $Z \to \nu\bar{\nu}$ and $W \to l\nu$ decays lead to the missing E_{T}. By requiring a second jet, the dijet selection effectively removes such events. *(Taken from Ref. [20])*

3.6 Jet Energy Resolution and Unfolding

Taking a step back, the achievements so far can be recapitulated as follows: Analysing experimental data means interpreting them in terms of "true" causes while accounting for effects of the experimental setup and the measuring apparatus. With respect to a measuring interval in some observable, undesirable effects can roughly be divided into four parts:

1. **Background:**
   Other physics processes with identical or very similar signatures than the sought for signal might exist. These increase the observation beyond expectation. Similarly, an increase might be provoked by detector artefacts or by different physics processes that, in combination with experimental effects, mimic the desired signature. In careful studies, this background must either be avoided, eliminated, or subtracted. With respect to jet analyses, background is considered to have been dealt with in the previous Section 3.5.
2. **Efficiency:**
   The measuring procedure might suffer from a limited acceptance or from inefficiencies. Both lead to losses that can be corrected by either measuring or estimating the efficiency including its uncertainty. The phase space of the jet studies considered here typically is restricted to the fiducial volume at full efficiency of any detector such that inefficiencies are negligible, cf. Sections 3.2 and 3.3.
3. **Bias:**
   Measured quantities might be displaced with respect to their original values. The prime example here are jet energies, which are subjected to a complicated calibration procedure as described in Section 3.4. Angular measures are, in general, reconstructed quite accurately, but in case of shifted vertices or in problematic detector regions small biases might need to be rectified. At this stage, all known biases of a jet measurement are assumed to be corrected.

4. **Smearing:**
   The last and technically difficult point to tackle concerns the smearing effect of the measuring device because of its finite resolution. As a result, an observation in a measuring interval can be changed both ways, i.e. increased or decreased, through migrations from and to other regions that in addition lead to corresponding correlations. The jet energy resolution and the unsmearing task that is solved by *unfolding* are the topics of the next two sections.

3.6.1 Jet Energy Resolution

As a consequence of the complex and fluctuating formation of hadronic showers, the jet energy or, more precisely, the jet p_T resolution is rather poor compared to the resolution of other physics objects such as muons, photons, or electrons, which are measured from tracks, electromagnetic energy deposits, or both. The considerable impact on jet measurements has to be corrected and the corresponding uncertainty needs to be estimated. As for the JEC, within CMS [9] this is performed through detailed detector simulations with subsequent data-based corrections for processes where a jet is balanced in p_T by either a photon, a Z boson, or a second jet. For e.g. γ+jet events the p_T balance $\mathcal{B}$ can be expanded as:

$$\mathcal{B} = \frac{p_{\mathrm{T,jet}}}{p_{\mathrm{T},\gamma}} = \frac{p_{\mathrm{T,jet}}}{p_{\mathrm{T,jet\ ptcl}}} \cdot \frac{p_{\mathrm{T,jet\ ptcl}}}{p_{\mathrm{T},\gamma\ \mathrm{ptcl}}} \cdot \frac{p_{\mathrm{T},\gamma\ \mathrm{ptcl}}}{p_{\mathrm{T},\gamma}}\,, \tag{3.21}$$

where $p_{\mathrm{T,jet}}$ and $p_{\mathrm{T},\gamma}$ are the reconstructed, and $p_{\mathrm{T,jet\ ptcl}}$ and $p_{\mathrm{T},\gamma\ \mathrm{ptcl}}$ the particle-level jet respectively photon transverse momenta. The quantity of interest, however, is not this balance itself, but the width of its distribution, $\sigma_\mathcal{B}$, for a sample of events. Assuming the three factors in Eq. (3.21) to represent independent and approximately Gaussian distributed random variables, the observed width $\sigma_\mathcal{B}$ can be derived through quadratic summation, indicated by $\oplus$, of the widths of each component:

$$\sigma_\mathcal{B} \cdot k_\mathrm{rad} = \sigma_\mathrm{JER} \oplus \sigma_\mathrm{PLI} \oplus \sigma_\gamma\,. \tag{3.22}$$

Here, $\sigma_\mathrm{JER} = \sigma\left(p_{\mathrm{T,jet}}/p_{\mathrm{T,jet\ ptcl}}\right) \equiv \sigma_{p_\mathrm{T}}/p_\mathrm{T}$ is the sought-after JER of particle-level jets, σ_PLI is the width of the intrinsic, particle-level imbalance (PLI) caused by the UE, OOC effects, and the presence of neutrinos, and σ_γ is the photon p_T resolution. The factor k_rad accounts for ISR and FSR whose impact is eliminated in a similar way as for the JEC by extrapolating to zero additional jet activity. In summary, the observed width $\sigma_\mathcal{B}$ of the $p_{\mathrm{T,jet}}$ versus $p_{\mathrm{T},\gamma}$ balance is the net effect of the JER, an intrinsic, particle-level jet p_T resolution, and the photon p_T resolution. Indicating quadratic subtraction by $\ominus$, Eq. (3.22) can be rearranged to finally give

$$\sigma_\mathrm{JER} = \sigma_\mathcal{B} \cdot k_\mathrm{rad} \ominus \sigma_\mathrm{PLI} \ominus \sigma_\gamma\,. \tag{3.23}$$

For dijet events the asymmetry $\mathcal{A}$, defined as

$$\mathcal{A} = \frac{p_{\mathrm{T,jet1}} - p_{\mathrm{T,jet2}}}{p_{\mathrm{T,jet1}} + p_{\mathrm{T,jet2}}}, \tag{3.24}$$

is used instead of the p_{T} balance. As a complication, jet related resolutions now appear twice, once for the probe and once for the tag jet, for the JER as well as for the intrinsic, particle-level resolution. In the case of both jets within comparable detector regions and sharing the same JER, one obtains:

$$\sigma_{\mathrm{JER}} = \sqrt{2} \cdot \sigma_{\mathcal{A}} \cdot k_{\mathrm{rad}} \ominus \sigma_{\mathrm{PLI}}\,. \tag{3.25}$$

More details on the general case can be found in Ref. [9].

An example of the simulated resolution of the jet response is displayed in Fig. 3.23 left. The core of the distribution up to about two σ is well described by a Gaussian, confirming the assumptions above. Beyond this core region tails are visible on both sides, which approximately follow a power law. At low p_{T}'s as shown in the example plot, the almost symmetric tails are caused through mismatch between jets that have been split either at reconstructed or at particle level but not at both. At high p_{T} the resolution curve becomes increasingly asymmetric with a power-law tail only on the left side because of losses through rare unresponsive detector areas or punch-through of high-p_{T} particles. Experience shows that all these effects can be well modelled by a double-sided Crystal Ball (CB) function that was first used by the Crystal Ball Collaboration in the context of calorimetric resolutions [42]. The piecewise-defined function is illustrated in Fig. 3.23 right at the example of a diphoton mass resolution as used in Ref. [43] and exhibits a Gaussian core with separate power-law tails beyond $\alpha_L \cdot \sigma_{\mathrm{CB}}$ to the left and $\alpha_H \cdot \sigma_{\mathrm{CB}}$ to the right hand side. Defining $x = (m-\mu)/\sigma$ the CB function can be written in a more general form as

$$f_{\mathrm{CB}} = N \cdot \begin{cases} e^{-\frac{1}{2}\alpha_L^2} \cdot \left[\left(\frac{\alpha_L}{n_L}\right)\left(\frac{n_L}{\alpha_L} - [\alpha_L + x]\right)\right]^{-n_L}, & x < -\alpha_L \\ e^{-\frac{1}{2}x^2}, & -\alpha_L \le x \le \alpha_H \\ e^{-\frac{1}{2}\alpha_H^2} \cdot \left[\left(\frac{\alpha_H}{n_H}\right)\left(\frac{n_H}{\alpha_H} - [\alpha_H - x]\right)\right]^{-n_H}, & x > \alpha_H \end{cases}, \tag{3.26}$$

where N is a normalisation factor and α_L and α_H delimit the Gaussian core, which is replaced by a power-law behaviour proportional to $1/n_L$ and $1/n_H$ to the lower respectively higher side. The piecewise function is designed such that the function itself and its first derivative are continuous.

For the purpose of parameterising σ_{JER} for the simulated jet response as a function of $p_{\mathrm{T,jet\ ptcl}}$, only the Gaussian core is considered. To account for PU effects, to which the noise term in PF jet resolutions is very sensitive, the calorimetric *NSC* formula Eq. (3.1) is modified as follows to include a possibly negative noise term:

$$\frac{\Delta p_{\mathrm{T}}}{p_{\mathrm{T}}} = \sqrt{\frac{\mathrm{sign}\,(N) N^2}{p_{\mathrm{T}}^2} + \frac{S^2}{p_{\mathrm{T}}} + C^2}\,, \tag{3.27}$$

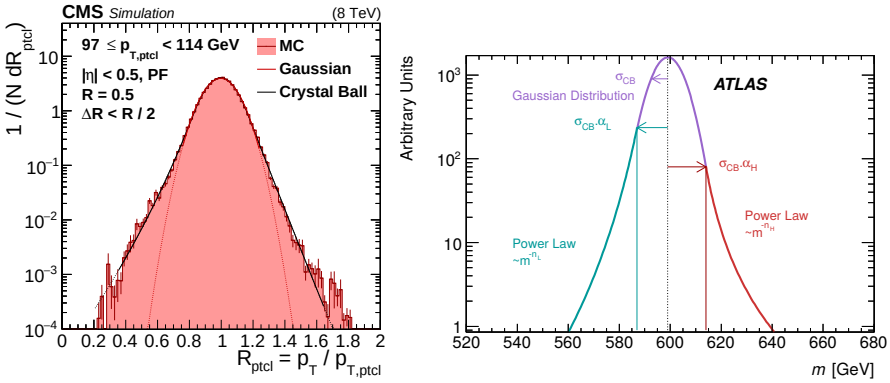

Fig. 3.23 Left: Jet p_{T} resolution in the central detector region, $|\eta| < 0.5$, from particle-flow jets matched to particle jets in the p_{T} range $97 \leq p_{\mathrm{T,ptcl}} < 114\,\mathrm{GeV}$. The distribution is well described by a double-sided Crystal Ball function (solid black line) with a Gaussian core up to $\pm 2\sigma$ (solid red line). Right: Illustration of a double-sided Crystal Ball function that corresponds to a Gaussian core with a lower and upper power-law tail. In this example a diphoton mass resolution is modelled. *(Taken from Refs. [9, 43])*

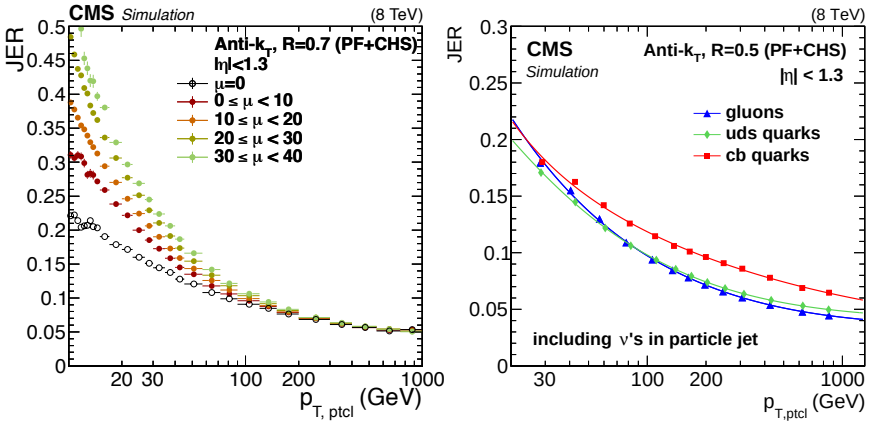

Fig. 3.24 Left: JER as a function of $p_{\mathrm{T,ptcl}}$ for anti-k_{t} jets with $R = 0.7$ without PU and for four ranges in average number of PU interactions, μ, per bunch crossing. Right: Flavour dependent intrinsic JER in simulation for a γ+jet sample and jets with in $|\eta| < 1.3$. The JER is shown for particle-level jets, where neutrinos are exceptionally included in the particle-level jet definition to demonstrate the loss in precision induced for c and b jets. Without neutrinos all curves are rather close to each other. *(Taken from Ref. [9])*

where sign$N$ can be either $+1$ or -1. The sensitivity to PU even for PF+CHS jets is shown in Fig. 3.24 left. The right panel of Fig. 3.24 demonstrates the effect of neutrinos on the JER, if they would be included in the particle-level jet definition of CMS. Without neutrinos, differences in σ_{JER} of the simulated response among the various quark flavours are rather small.

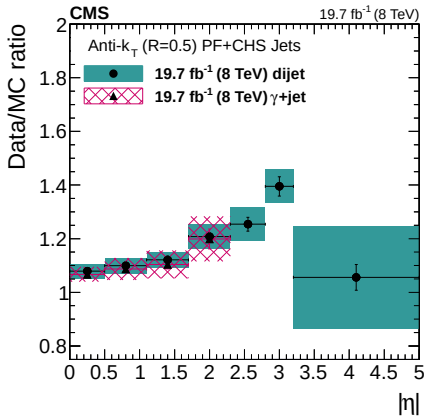

Fig. 3.25 Data-based JER scale factors including uncertainty as a function of $|\eta|$ from the γ+jet p_{T} balance and the dijet asymmetry method to correct the underestimated JER from simulation. *(Taken from Ref. [9])*

The results from simulation are checked in data by using the γ+jet p_{T} balance and the dijet asymmetry method, rewritten in terms of the average p_{T}, $\langle p_{\mathrm{T1,2}} \rangle$, of the two jets as

$$\mathcal{A} = \frac{p_{\mathrm{T,jet1}} - p_{\mathrm{T,jet2}}}{p_{\mathrm{T,jet1}} + p_{\mathrm{T,jet2}}} = \frac{p_{\mathrm{T,jet1}} - p_{\mathrm{T,jet2}}}{2 \cdot \langle p_{\mathrm{T1,2}} \rangle} \,. \tag{3.28}$$

Figure 3.25 presents for both data-based methods the final scale factors including uncertainty that have to be applied to simulation-based JER corrections. Where applicable, both methods agree and indicate an underestimation of the JER in the MC simulation, in particular when approaching the transition region at $|\eta| \approx 3.0$ from the endcap to the HF calorimeter of the CMS detector.

3.6.2 Unfolding

Reflecting on finite detector resolutions, one can draw already two conclusions with respect to the impact of e.g. JER: First, physical causes that would lead to signals with high "frequency" with respect to the typical resolution width are smeared out and become immeasurable. Secondly, variations of the observed data at similarly high "frequencies" are amplified in the inverse process. As a consequence, physical causes can be differentiated into one part that is immeasurable with the present device entailing an unavoidable loss of information, and a second part, on which something might be learned by studying the observations. However, the collected data are subject to statistical fluctuations, which are amplified in the inverse process as well. Both issues together render the unfolding procedure a rather difficult task.

The necessity of unfolding depends on the purpose of a measurement. If the primary goal consists in a comparison to a particular theory prediction, then it is much safer and more efficient to smear the theory by estimating a response function or matrix from detector simulation. This strategy is applied for example in Section 6.1, where a new physics hypothesis is tested against the null hypothesis in order to derive an exclusion limit. If on the other hand observations of different experiments shall be compared or even combined in fits of theory parameters, then there is no alternative to unfolding. Equally, in some cases features may become recognisable only in the unfolded result as for example in a deblurred image. Measurements corrected for detector resolutions are also desirable for future comparisons to not yet developed theories, or, for efficiency reasons, in tunings of MC event generators. Last but not least, culture-wise it is more satisfying at some point to conclude with one "true" figure for posterity than with smeared theory distributions, one for each experiment. Given today's data storage capabilities though, it might be wise to publish both, "raw" data with estimated detector response, and unfolded data with full covariance. In particle physics, this is very rarely done.

As will be seen, drawbacks of unfolding might consist in introducing biases and losses of sensitivity, in particular with respect to small-scale structures in comparison to the typical resolution width, e.g. oscillations or narrow bumps. The following details are tailored to the usage in the discussed particle-physics analyses, where the decision was made to unfold. They cannot do justice to the field of unfolding in general, which is developing dynamically. For an overview with specialisation on particle physics the following Refs. are recommended: [44–47]. In addition, two workshops of the PHYSTAT series that brings together particle physicists, astrophysicists, cosmologists, and statisticians comprise contributions dedicated to unfolding: [48, 49]. To emphasise the importance of proper statistical methods in general, each LHC experimental collaboration has established a Statistics Committee that provides advice and recommendations to the physicists. The following explanations profited on occasion from related discussions with the CMS Statistics Committee.

3.6.2.1 Matrix Inversion

Mathematically, the unfolding problem can be stated in the following form

$$\int_{\Omega} K(y,x) f(x)\, dx = g(y)\,, \tag{3.29}$$

where $g(y)$ is the measured distribution, $f(x)$ is the true distribution, and $K(y,x)$ is a Kernel function describing the measurement process that transforms $f(x)$ into $g(y)$ within some fixed phase space Ω. The solution of such a *Fredholm integral equation of the first kind* counts among the so-called *ill-posed inverse problems*. In the special case, where the Kernel $K(s,t) = K(s-t)$ only depends on the difference of its two arguments, the inverse process to the resulting folding integral is called *deconvolution*.

In practice, the first step towards a numerical solution consists in binning the measurement in the form of a histogram, which effectively smoothes statistical fluctuations, but also signal variations happening at scales comparable to the chosen bin width. Representing this histogram by an m-vector $\vec{y}$ of measured data and the true distribution by an n-vector of unknowns $\vec{x}$, the problem can be rewritten as

$$\mathbf{R}\vec{x} = \vec{y}, \tag{3.30}$$

with $\mathbf{R}$ being a rectangular $(m \times n)$ response matrix. In the special case of a quadratic matrix, $m = n$, the solution reads

$$\vec{x} = \mathbf{R}^{-1}\vec{y} \quad (\mathrm{m} = \mathrm{n}), \tag{3.31}$$

if $\mathbf{R}$ can be inverted, i.e. it has no singular eigenvalue. The physics analyses studied later in more detail have $m = n$. If more information should be extracted than there are measurements, $n > m$, the problem is guaranteed to be ill-posed and should be reconsidered. For $m > n$, the *Moore–Penrose pseudoinverse* $\mathbf{R}^{+}$, with $\mathbf{R}\mathbf{R}^{+}\mathbf{R} = \mathbf{R}^{+}$, $\mathbf{R}^{+}\mathbf{R}\mathbf{R}^{+} = \mathbf{R}$, and both, $\mathbf{R}\mathbf{R}^{+}$ and $\mathbf{R}^{+}\mathbf{R}$ Hermitian, can be used to rewrite the solution in the form

$$\vec{x} = \mathbf{R}^{+}\vec{y} \quad (\mathrm{m} > \mathrm{n}). \tag{3.32}$$

Applying standard methods of linear algebra and statistics, it can be shown that the *matrix inversion method*, Eq. (3.31), provides the *least squares* (LS) and also the *maximum likelihood* (ML) solution with the property of having the smallest variance among all unbiased estimators. In contrast to other fields of application like deblurring in astronomical or medical imaging, it is absolutely essential in particle physics to not only have the unfolded result, but also estimates of its uncertainties and correlations. In other words, it is mandatory to derive the full covariance matrix $\mathbf{C_x}$ in the unfolded space $\vec{x}$:

$$\mathbf{C_x} = \mathbf{R}^{-1}\mathbf{C_y}(\mathbf{R}^{-1})^{T}. \tag{3.33}$$

In favourable situations, for example the measurement of dijet azimuthal decorrelations [50] described in Section 6.5, matrix inversion can be sufficient despite the fact that in general neither $\mathbf{R}$ nor $\vec{y}$ are known exactly. Frequently, however, inevitable statistical fluctuations in the data $\vec{y}$ render this method useless. As explained above, high-frequency signals, suppressed by detector effects, are mimicked by statistical noise and are amplified in the inversion leading to strongly oscillating solutions with large variances. The only means to improve this situation lies in accepting some bias because of additional assumptions in exchange for variances that are better-behaved than for the ML estimator. This procedure is called *regularisation*.

3.6.2.2 Regularisation

Widening the bins introduces an implicit regularisation and helps in reducing bin migrations, in the extreme leading to a simple diagonal response matrix. To avoid that too much information is lost that way, the bin widths are usually chosen to be of the order of the experimental resolution. On the other hand, $\mathbf{R}$ is estimated by MC simulation. In addition to the implied model dependency and statistical fluctuations of the simulated events, this poses a problem explicitly for very steep spectra in combination with wide bins, because the approximation quality of a steep function deteriorates with the bin width. Choosing a binning much narrower than the resolution requires more elaborate regularisations and potentially complicates estimating other systematic uncertainties like the one from JEC, which is not always propagated through the unfolding but implicitly assumes small migrations mostly between neighbouring bins. Given these conflicting boundary conditions, there is no simple recipe for choosing a binning and a case-by-case judgement is advised.

Regularised unfolding techniques used in particle physics can roughly be differentiated into two classes:

1. **Iterative solutions with early stopping:**
   In practice, it might be difficult to calculate the ML estimator via the inversion of the response matrix. However, it can always be computed iteratively through the *expectation-maximisation* (EM) algorithm [51]. In Ref. [52] it has been proven that this procedure converges to an ML estimator for an infinite number of iterations, i.e. it approaches the matrix inversion result including the potential build-up of large oscillations. Regularisation comes into play implicitly by the property of so-called *semi-convergence*. Starting from some initial assumption on the unfolding solution $\vec{x}$, the first iterations account for the dominant structures in the data and show significant improvements, after which the convergence becomes rather slow. The regularisation parameter is the number of iterations and hence the method could be called "regularisation by early stopping of the EM iteration".

2. **Direct solutions with Tikhonov regularisation:**
   In an LS approach the task is to minimise the sum of squared residuals $\|\mathbf{R}\vec{x}-\vec{y}\|^2$, with $\|\,\|$ being the Euclidean or L^2 norm, which leads to the known oscillatory behaviour. By adding a penalty term $\|\mathbf{G}\vec{x}\|^2$ with a properly chosen *Tikhonov matrix* $\mathbf{G}$, preference can be given to desired properties of a solution. This is called *Tikhonov regularisation* [53, 54], where various choices are possible for the penalty term. If $\mathbf{G}$ is chosen to be a multiple of the identity matrix $\mathbf{I}$ as in the original suggestion, then solutions with smaller L^2 norms are preferred.

The iterative technique has been invented a couple of times and is known in optics and astronomy under the name *Lucy–Richardson deconvolution* [55, 56], where it is used for the deblurring of images, e.g. of the Hubble Space Telescope, assuming the response matrix to be a *point spread function*. In a similar way it is applied in medical tomography [57]. Within particle physics it has been re-derived

from Bayesian reasoning without a particular assumption on $\mathbf{R}$ in Ref. [39] and is known as *D'Agostini iterative unfolding*.[5] The D'Agostini iterative approach, which is the method of choice in most of the jet analyses presented in this work, is available within the software framework RooUnfold [58] together with the matrix inversion solution, or bin-by-bin correction factors that should be used only in the simplest case with negligible bin-to-bin correlations. RooUnfold provides two possibilities to propagate the statistical uncertainty from the measurement through the unfolding process, analytic error propagation or toy MC experiments. A term missing from the original description of the analytic method [39] lead to an underestimation of the unfolded uncertainties caused by the fluctuations in data and has been fixed in Ref. [58]. Another iterative method [59] is employed in some of the ATLAS measurements, for example in the dijet mass and 3-jet cross sections, cf. Refs. [19, 60].

Direct numerical methods to solve Eq. (3.32) often make use of the *singular value decomposition* (SVD). Frequently, SVD is employed for crosschecking the iterative approach, but it also figures as primary unfolding method for example in event shape analyses [61, 62], cf. Section 6.6. For $m \geq n$, $\mathbf{R}$ can be written as the product $\mathbf{R} = \mathbf{USV}^T$ of an orthonormal $(m \times n)$-matrix $\mathbf{U}$, a diagonal $(n \times n)$-matrix $\mathbf{S}$, and the transpose of a second orthonormal $(n \times n)$-matrix $\mathbf{V}$.[6] The matrix $\mathbf{S}$ is diagonal with non-negative so-called *singular values* σ_j, $j = 1, \ldots, n$, which conventionally are sorted in decreasing order. As a result the response is rewritten in terms of an orthonormalised vector space with left- and right-singular vectors $\vec{u}_j$, $\vec{v}_j$ that appear as columns of matrices $\mathbf{U}$ and $\mathbf{V}$, respectively, and that exhibit increasing numbers of sign changes in their elements with increasing index number. The singular values σ_j therefore correspond to factors associated to oscillations with frequencies that grow with the index j. Following Ref. [44] the product $\mathbf{R}\vec{x}$ can be expressed via SVD as

$$\mathbf{R}\vec{x} = \mathbf{USV}^T\vec{x} = \sum_{j=1}^{n} \sigma_j \left(\vec{v}_j^T \vec{x}\right) \vec{u}_j = \vec{y}, \tag{3.34}$$

which demonstrates that contributions to $\vec{y}$ are increasingly suppressed by factors σ_j that shrink with the index number j. The *condition number* of a matrix $\mathbf{R}$ is defined as the ratio of the largest to the smallest singular value, i.e. with the current notation

$$\mathrm{cond}(\mathbf{R}) = \frac{\sigma_1}{\sigma_n}, \tag{3.35}$$

and is a measure of the sensitivity of the solution of the inverse problem to small perturbations in the input data. Matrices with condition numbers of the order of unity are said to be *well-conditioned* and their inverse can be computed to good accuracy. Matrices with infinite condition number, i.e. at least one singular value equals zero,

[5] Sometimes the iterative approach, although coinciding with a frequentist technique to compute an ML estimator, is named Bayesian unfolding.

[6] An orthonormal matrix $\mathbf{A}$ satisfies $\mathbf{AA}^T = \mathbf{I}$.

are not invertible and do not lead to stable or unique solutions for the corresponding system of linear equations. In practice, the response matrix $\mathbf{R}$ is estimated from MC simulations with limited precision in each element such that the σ_i are never exactly zero and the condition number ranges from unity to arbitrarily high values associated with ill-posed problems. As a rule of thumb one loses d digits in decimal precision, if the matrix to invert has a condition number of 10^d.

Assuming that all singular values are non-zero and applying the LS method to the inversion problem formulated with the help of SVD, one can derive the LS estimate for $\vec{x}$ and its variance $\mathbf{C_x}$ to [44]:

$$\vec{x} = \mathbf{R}^{+}\vec{y} = \mathbf{V}\mathbf{S}^{-1}\left(\mathbf{U}^T\vec{y}\right) = \sum_{j=1}^{n}\frac{1}{\sigma_j}\left(\vec{u}_j^T\vec{y}\right)\vec{v}_j = \sum_{j=1}^{n}\left(\frac{c_j}{\sigma_j}\right)\vec{v}_j \quad \text{and} \tag{3.36}$$

$$\mathbf{C_x} = \mathbf{R}^{+}\mathbf{C_y}\mathbf{R}^{+T} = \mathbf{V}\mathbf{S}^{-2}\mathbf{V}^T = \sum_{j=1}^{n}\left(\frac{1}{\sigma_j^2}\right)\vec{v}_j\vec{v}_j^T\,, \tag{3.37}$$

where the *Fourier coefficients* $c_j = \vec{u}_j^T\vec{y}$ represent the transformed measurement. In this form, it is demonstrated that small singular values σ_j, associated to immeasurable high-frequency oscillations, can dramatically enhance meaningless random fluctuations of the measurement in the inversion procedure leading to large variances proportional to $1/\sigma_j^2$.

A simple means based on SVD to avoid the build-up of nonphysical oscillations is the truncation of the sum appearing in Eqs. (3.36) and (3.37) to only the first p, $p < n$, terms. The difficulty lies in the choice of the cutoff p requiring a decision, which part of the singular values σ_j to consider and which to ignore. A clear step in size of the σ_j might indicate, where to separate between measurable signal reconstruction and noise amplification. Unfortunately, such a clear step usually is not present, in which case sharp cutoffs between two singular values of similar size are known to lead to the *Gibbs phenomenon* of provoking oscillating components at discontinuities. A solution is to employ more elaborate regularisation prescriptions than the truncation, which effectively lead to a continuous dampening of high-frequency noise. The norm regularisation according to Tikhonov adds the penalty term $\tau\|\mathbf{G}\vec{x}\|^2$ to the quantity to be minimised following the LS method, where $\tau > 0$ is an adjustable parameter setting the regularisation strength, and $\mathbf{G} = \mathbf{I}$. The norm-regularised estimate of $\vec{x}$ then becomes [44]:

$$\vec{x} = \sum_{j=1}^{n}\left(\frac{c_j}{\sigma_j}\right)\varphi_j\vec{v}_j \quad \text{with} \quad \varphi_j = \frac{\sigma_j^2}{\sigma_j^2+\tau}\,, \tag{3.38}$$

with φ_j representing the dampening filter for this case. For $\tau \to 0$ all filter factors φ approach unity and the unregularised solution is recovered. For the index number j, where $\tau \approx \sigma_j^2$, an attenuation by $1/2$ is reached with stronger dampening for larger j. Penalty terms based on derivatives, particularly second ones, that delimit the curvature of the unfolding solution, are popular, too, and lead to other shapes for the filter factors.

The RooUnfold framework [58] comprises two incarnations of Tikhonov regularisation. The first one implements the SVD based method suggested in Ref. [63]. The second one is an interface to the TUnfold package [64], which can also be used separately. Both provide error propagation and numerous additional options and are in wide-spread use in the particle physics community.

3.6.2.3 Forward Smearing

All described methods require the estimated detector response matrix $\mathbf{R}$ as input in addition to the data $\vec{y}$ with variances $\mathbf{C_y}$. However, in particle physics it is very time consuming to derive response matrices through the generation of MC events as input to detailed detector simulations by the Geant4 package [25]. In addition to sparsely populated areas in $\mathbf{R}$ with corresponding large statistical uncertainties, it is non-trivial to account for the residual differences in JER as described in the previous section 3.6.1. A means to overcome these limitations is *forward smearing*, which pursues the following strategy:

1. Generate huge amounts of pseudoevents for the considered observable such that all relevant bins of the response matrix can sufficiently be filled.
2. Weight the generated events so that a theory prediction is reproduced. If the theory prediction is available only in binned form, a smooth approximation, for example by cubic splines, can be performed beforehand.
3. Estimate the resolution in the observable either from data directly, or from detailed MC simulations, which potentially are modified to account for residual differences between data and simulations
4. Smear the generated input events and fill the response matrix accordingly.
5. Perform closure tests by smearing and unfolding the known theory prediction once with identical and once with statistically independent events. If feasible, also compare to the unfolding result when filling the response matrix from a simulation of the input theory. Check on a potential bias by cross-unfolding one theory prediction with another.
6. Unfold the data and compute the respective covariance matrix that includes statistical uncertainties and correlations.
7. Check that the statistical uncertainties after unfolding are mostly larger than before.
8. Estimate systematic uncertainties by varying the JER, parameters of the input theory like PDFs, or the MC generators used in one of the previous steps.

This procedure permits to study and address several problems of unfolding like insufficient amounts of simulated data, binning approximations, dependencies on theory models and parameters, and potential related biases. Figure 3.26 shows for example the corresponding response matrices as they were used for the inclusive jet p_{T} and dijet mass spectra of Ref. [20]. In both cases the matrices are almost symmetric around the principal diagonal, as expected from the procedure, and more than 50% of the jets respectively events stay within the same observable interval

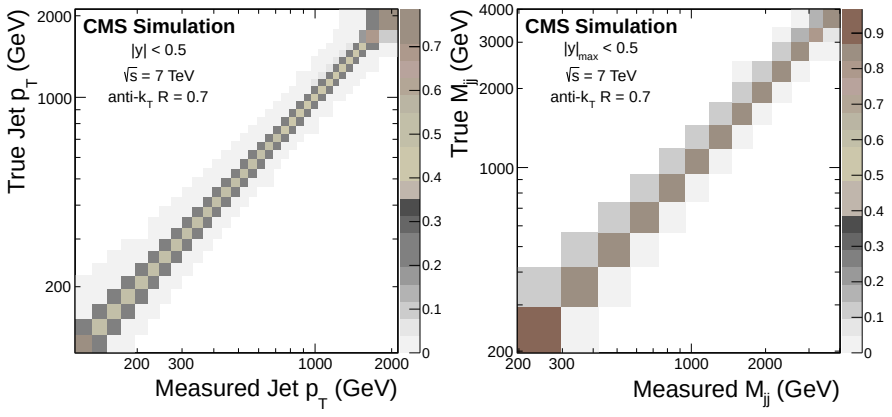

Fig. 3.26 Response matrices as used in the unfolding of the inclusive jet p_{T} spectrum (left) and the dijet mass spectrum (right) for jets within $|y| < 0.5$. The normalised event frequencies are given as colour-coded percentages. *(Taken from Ref. [20])*

with some migrations between neighbouring bins. In combination with steep spectra this leads nevertheless to a significant shift in the measured distribution, which is addressed through the unfolding steps listed above.

Typical closure tests performed in such an unfolding setup are presented in Fig. 3.27. If the same sample of events is used to construct the response matrix and to serve as pseudodata, then the unfolding procedure must, apart from edge effects, reproduce exactly the original distribution irrespective of the statistical precision as shown in the top left panel. In case of statistically independent samples used for the response and the pseudodata, agreement can be expected only within statistical uncertainty as demonstrated by the panels top right for low and bottom left for high accuracy. The last panel, bottom right, illustrates the size of the correction to anticipate for the respective observable. For steeply falling spectra, symmetric smearing effects always exaggerate the measurement towards large values of an observable such that the required resolution correction leads to factors smaller than unity.

As an extension, multidimensional unfolding can be considered.[7] Because of the excellent angular resolutions of the detectors at the LHC, multidimensional unfolding was rarely necessary for jet analyses so far. The impact of smearing in rapidity was found to be negligible for inclusive jet measurements in studies performed before the start-up of the LHC [66]. This was most recently confirmed in Ref. [67] for LHC Run 2. In Ref. [50] on dijet azimuthal decorrelation, a 2-dimensional unfolding in p_{T} and $\Delta\phi_{\mathrm{dijet}}$, performed with the D'Agostini iterative method as implemented in ROOUNFOLD, demonstrated that resolution effects in $\Delta\phi_{\mathrm{dijet}}$ can be neglected. The same method was employed in Ref. [68] for dijet angular distributions binned in dijet mass and $\chi_{\mathrm{dijet}} = \exp(|y_1 - y_2|)$. Some impact was also observed for χ_{dijet},

[7] Depending on the employed software tools this may require additional steps like projecting all n-dimensional bins onto one linear array.

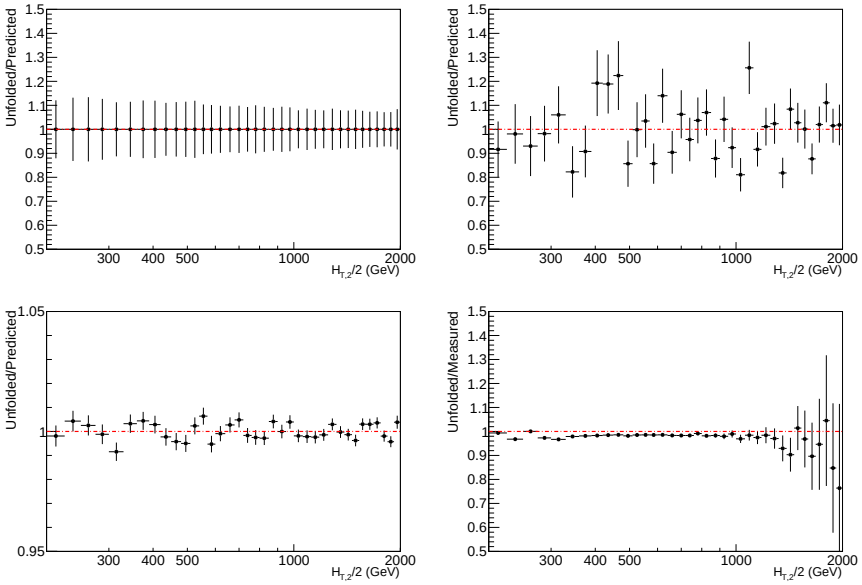

Fig. 3.27 Typical closure tests performed with an unfolding setup at the example of the inclusive dijet cross section as a function of the average p_{T} of the two leading p_{T} jets labelled as $H_{\mathrm{T},2}/2$. Ratio of the unfolding result to the prediction for identical pseudoevents for the response matrix R and pseudodata (top left), and for small (top right) and large (bottom left) samples of statistically independent pseudoevents used for R and the pseudodata. The bottom right plot shows the ratio of the unfolding result to a pseudomeasurement with realistic assumptions on available amounts of data. *(Adapted from plots courtesy of A. Kaur [65])*

but mostly because of a reordering of the leading jets in p_{T} caused by resolution effects instead of migrations in $|\Delta y|$.

3.6.2.4 Regularisation Strength and Bias

A so far unsolved issue is the choice of the regularisation strength, i.e. the number of iterations or the value of τ in the discussed techniques. In some cases it is possible, as described, to clearly separate the singular values of a response matrix into physics-wise relevant and noise-related ones. Generally, however, more elaborate objective criteria are necessary. Two possibilities, the L-curve [69] and correlation minimisation are implemented in TUNFOLD. In the latter, either the average or the maximal correlation as extracted from the covariance $\mathbf{C_x}$ are minimised with respect to τ. Because a certain amount of correlation must be expected from the regularisation, this criterion is not without problems and is strongly dependent on the choice of binning with respect to the experimental resolution. A similar argument, i.e. the observation of a build-up of long-range correlations, can be used to stop the D'Agostini iterative unfolding after a certain number of iterations. As an

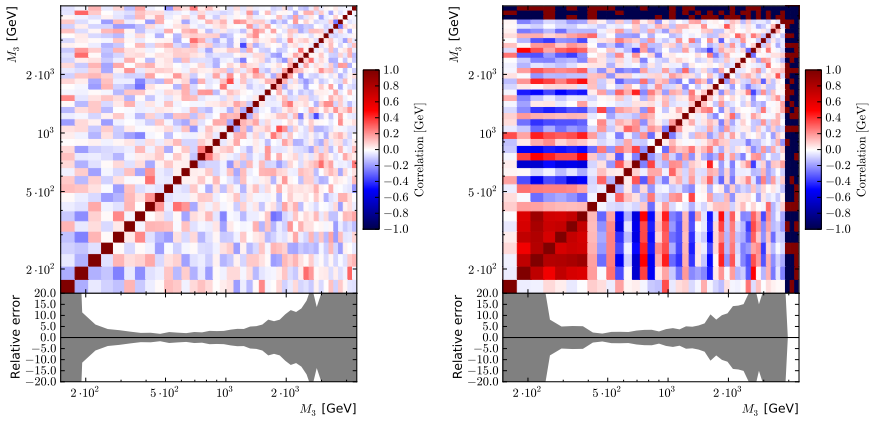

Fig. 3.28 Correlation matrices from the unfolding of a 3-jet mass spectrum at central rapidity with the iterative D'Agostini method after four (left) and after six iterations (right). *(Adapted from plots courtesy of F. Stober [24])*

example favouring four iterations, Figure 3.28 presents correlation matrices from a study of 3-jet mass distributions. The same number of four was found in a χ^2 test in smeared space for inclusive jets at 13 TeV [67] by studying the agreement between uncorrected data and back-folded data as a function of the number of iterations. A concern of this data-based criterion is the unknown influence of the statistical uncertainty of the data. In a previous inclusive jet measurement [20] five iterations were performed to assure that the statistical uncertainty in the unfolded distribution is always larger than in the originally measured one, which is reasonable but not mandatory in all circumstances. Related to the choice of the regularisation strength is the characterisation of the ensuing regularisation bias in terms of an uncertainty. This topic is part of current research in statistics [70–72] with the prospect of interesting new insights to be expected from future developments.

References

1. e. Evans, Lyndon and e. Bryant, Philip, "LHC Machine", *JINST* **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.
2. R. M. Barnett et al., "Physics at the Terascale". Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 1st edition, April, 2011.
3. R. Alemany-Fernandez et al., "The Large Hadron Collider: Harvest of Run 1". Springer, Berlin, Germany, 1st edition, May, 2015.
4. ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.
5. CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

6. ALICE Collaboration, "The ALICE experiment at the CERN LHC", *JINST* **3** (2008) S08002, doi:10.1088/1748-0221/3/08/S08002.
7. LHCb Collaboration, "The LHCb Detector at the LHC", *JINST* **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.
8. ATLAS Collaboration, "Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", *Eur. Phys. J. C* **75** (2015) 17, doi:10.1140/epjc/s10052-014-3190-y, arXiv:1406.0076.
9. CMS Collaboration, "Jet Energy Scale and Resolution in the CMS Experiment in pp Collisions at 8 TeV", (2016). To be submitted to *JINST*.
10. M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
11. W. Lampl et al., "Calorimeter clustering algorithms: Description and performance", technical report, 2008.
12. CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", technical report, CERN, 2009.
13. CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", technical report, CERN, 2010.
14. C. Buttar et al., "Standard Model Handles and Candles Working Group: Tools and Jets Summary Report", in *Proceedings, 5th Les Houches Workshop 2007 on Physics at TeV colliders (Les Houches 2007)*, p. 121. Les Houches, France, June 11-29, 2007. arXiv:0803.0678.
15. A. Buckley et al., "General-purpose event generators for LHC physics", *Phys. Rept.* **504** (2011) 145, doi:10.1016/j.physrep.2011.03.005, arXiv:1101.2599.
16. ATLAS and CMS Collaboration, "Jet energy scale uncertainty correlations between ATLAS and CMS", technical report, CERN, 2014.
17. ATLAS and CMS Collaboration, "Jet energy scale uncertainty correlations between ATLAS and CMS at 8 TeV", technical report, CERN, 2015.
18. CMS Collaboration, "Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **107** (2011) 132001, doi:10.1103/PhysRevLett.107.132001, arXiv:1106.0208.
19. ATLAS Collaboration, "Measurement of dijet cross sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector", *JHEP* **05** (2014) 059, doi:10.1007/JHEP05(2014)059, arXiv:1312.3524.
20. CMS Collaboration, "Measurements of differential jet cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the CMS detector", *Phys. Rev. D* **87** (2013) 112002, doi:10.1103/PhysRevD.87.112002, arXiv:1212.6660.
21. C. J. Clopper and E. S. Pearson, "The Use of Confidence or Fiducial Limits illustrated in the case of the Binomial", *Biometrika* **26** (1934) 404, doi:10.1093/biomet/26.4.404.
22. E. B. Wilson, "Probable Inference, the Law of Succession, and Statistical Inference", *Journal of the American Statistical Association* **22** (1927) 209, doi:10.1080/01621459.1927.10502953.
23. V. Lendermann et al., "Combining Triggers in HEP Data Analysis", *Nucl. Instrum. Meth. A* **604** (2009) 707, doi:10.1016/j.nima.2009.03.173, arXiv:0901.4118.
24. F. Stober, "Measurement of the three-jet mass cross-section at $\sqrt{s} = 7$ TeV". PhD thesis, KIT (Karlsruher Institut für Technologie), October, 2012.
25. S. Agostinelli et al., "GEANT4: A Simulation toolkit", *Nuclear Instruments & Methods in Physics Research A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
26. D0 Collaboration, "Jet energy scale determination in the D0 experiment", *Nucl. Instrum. Meth. A* **763** (2014) 442, doi:10.1016/j.nima.2014.05.044, arXiv:1312.6873.
27. CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
28. CMS Collaboration, "Single-Particle Response in the CMS Calorimeters", technical report, 2010.

29. A. Giammanco, "The Fast Simulation of the CMS Experiment", in *Proceedings, 20th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2013)*, volume 513, p. 022012. Amsterdam, Netherlands, October 14-18, 2014. doi:10.1088/1742-6596/513/2/022012.
30. ATLAS Collaboration, "Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **73** (2013) 2304, doi:10.1140/epjc/s10052-013-2304-2, arXiv:1112.6426.
31. CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", *JINST* **7** (2012) P10002, doi:10.1088/1748-0221/7/10/P10002, arXiv:1206.4071.
32. CMS Collaboration, "Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV", *JINST* **8** (2013) P09009, doi:10.1088/1748-0221/8/09/P09009, arXiv:1306.2016.
33. CMS Collaboration, "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV", *JINST* **10** (2015) P08010, doi:10.1088/1748-0221/10/08/P08010, arXiv:1502.02702.
34. R. K. Ellis, W. J. Stirling, and B. R. Webber, "QCD and Collider Physics". Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, Cambridge, 1996.
35. A. Banfi, G. P. Salam, and G. Zanderighi, "Infrared safe definition of jet flavor", *Eur. Phys. J. C* **47** (2006) 113, doi:10.1140/epjc/s2006-02552-4, arXiv:hep-ph/0601139.
36. L. Demortier, "Equivalence of the best-fit and covariance matrix methods for comparing binned data with a model in the presence of correlated systematic uncertainties", (April, 1999). CDF Note 8661.
37. D. Stump et al., "Uncertainties of predictions from parton distribution functions. 1. The Lagrange multiplier method", *Phys. Rev. D* **65** (2001) 014012, doi:10.1103/PhysRevD.65.014012, arXiv:hep-ph/0101051.
38. L. Lyons, A. J. Martin, and D. H. Saxon, "On the Determination of the B Lifetime by Combining the Results of Different Experiments", *Phys. Rev. D* **41** (1990) 982, doi:10.1103/PhysRevD.41.982.
39. G. D'Agostini, "Bayesian Reasoning in Data Analysis: A Critical Introduction". World Scientific Publishing Co. Pte. Ltd., Singapore, 2003.
40. R. D. Ball et al., "Fitting parton distribution data with multiplicative normalization uncertainties", *JHEP* **05** (2010) 075, doi:10.1007/JHEP05(2010)075, arXiv:0912.2276.
41. ATLAS Collaboration, "Measurement of the inclusive jet cross-section in proton-proton collisions at $\sqrt{s} = 7$ TeV using 4.5 fb^{-1} of data with the ATLAS detector", *JHEP* **02** (2015) 153, doi:10.1007/JHEP02(2015)153, arXiv:1410.8857.
42. M. J. Oreglia, "A Study of the Reactions $\psi' \to \gamma\gamma\psi$". PhD thesis, SLAC (Stanford Linear Accelerator Center), December, 1980.
43. ATLAS Collaboration, "Search for Scalar Diphoton Resonances in the Mass Range $65 - 600$ GeV with the ATLAS Detector in pp Collision Data at $\sqrt{s} = 8$ TeV", *Phys. Rev. Lett.* **113** (2014) 171801, doi:10.1103/PhysRevLett.113.171801, arXiv:1407.6583.
44. R. Barlow et al., "Data Analysis in High Energy Physics". Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 1st edition, June, 2013.
45. G. Bohm and G. Zech, "Introduction to Statistics and Data Analysis for Physicists". Deutsches Elektronen-Synchrotron, Hamburg, Germany, 1st edition, February, 2010.
46. G. Cowan, "Statistical Data Analysis". Clarendon Press, Oxford, 1 edition, June, 1998.
47. G. Zech, "Comparing statistical data to Monte Carlo simulation: Parameter fitting and unfolding", technical report, DESY, Hamburg, June, 1995.
48. G. Cowan, "A survey of unfolding methods for particle physics", in *Proceedings, Conference on Advanced statistical techniques in particle physics*, volume C0203181, p. 248. Durham, UK, March 18-22, 2002.

49. H. B. Prosper and L. Lyons, eds., "Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding", CERN. CERN, Geneva, Switzerland, January 17-20, (2011). doi:10.5170/CERN-2011-006.
50. CMS Collaboration, "Measurement of dijet azimuthal decorrelations in pp collisions at $\sqrt{s} = 8$ TeV", (2015). arXiv:1602.04384. Submitted to *Eur. Phys. J. C*.
51. D. B. R. A. P. Dempster, N. M. Laird, "Maximum Likelihood from Incomplete Data via the EM Algorithm", *Journal of the Royal Statistical Society. Series B (Methodological)* **39** (1977) 1.
52. Y. Vardi, L. A. Shepp, and L. Kaufman, "A Statistical Model for Positron Emission Tomography", *Journal of the American Statistical Association* **80** (1985) 8, doi:10.1080/01621459.1985.10477119.
53. A. Tikhonov, "On the solution of improperly posed problems and the method of regularization", *Soviet Math. Dokl.* **5** (1963) 1035.
54. D. L. Phillips, "A Technique for the Numerical Solution of Certain Integral Equations of the First Kind", *J. ACM* **9** (January, 1962) 84, doi:10.1145/321105.321114.
55. W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration", *J. Opt. Soc. Am.* **62** (January, 1972) 55, doi:10.1364/JOSA.62.000055.
56. L. B. Lucy, "An iterative technique for the rectification of observed distributions", *Astron. J.* **79** (1974) 745–754, doi:10.1086/111605.
57. L. A. Shepp and Y. Vardi, "Maximum Likelihood Reconstruction for Emission Tomography", *Medical Imaging, IEEE Transactions on* **1** (October, 1982) 113, doi:10.1109/TMI.1982.4307558.
58. T. Adye, "Unfolding algorithms and tests using RooUnfold", in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*, p. 313. Geneva, Switzerland, January 17-20, 2011. arXiv:1105.1160. doi:10.5170/CERN-2011-006.
59. B. Malaescu, "An Iterative, dynamically stabilized method of data unfolding", arXiv:0907.3791.
60. ATLAS Collaboration, "Measurement of three-jet production cross-sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector", *Eur. Phys. J. C* **75** (2014) 228, doi:10.1140/epjc/s10052-015-3363-3, arXiv:1411.1855.
61. CMS Collaboration, "First Measurement of Hadronic Event Shapes in pp Collisions at $\sqrt{(s)} = 7$ TeV", *Phys. Lett. B* **699** (2011) 48, doi:10.1016/j.physletb.2011.03.060, arXiv:1102.0068.
62. CMS Collaboration, "Event shapes and azimuthal correlations in Z + jets events in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **722** (2013) 238, doi:10.1016/j.physletb.2013.04.025, arXiv:1301.1646.
63. A. Höcker and V. Kartvelishvili, "SVD Approach to Data Unfolding", *Nucl. Instrum. Meth. A* **372** (1996) 469, doi:10.1016/0168-9002(95)01478-0, arXiv:hep-ph/9509307.
64. S. Schmitt, "TUnfold: an algorithm for correcting migration effects in high energy physics", *JINST* **7** (2012) T10003, doi:10.1088/1748-0221/7/10/T10003, arXiv:1205.6201.
65. A. Kaur. Private communication. To be published in *PhD thesis*.
66. A. Oehler, "Strategy for an initial Measurement of the Inclusive Jet Cross Section with the CMS Detector". PhD thesis, Universität Karlsruhe, December, 2009.
67. CMS Collaboration, "Measurement of the double-differential inclusive jet cross section at $\sqrt{s}$ = 13 TeV", technical report, CERN, 2015.
68. CMS Collaboration, "Search for quark contact interactions and extra spatial dimensions using dijet angular distributions in proton–proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **746** (2015) 79, doi:10.1016/j.physletb.2015.04.042, arXiv:1411.2646.
69. P. C. Hansen, "Analysis of Discrete Ill-Posed Problems by Means of the L-Curve", *SIAM Review* **34** (1992) 561, doi:10.1137/1034115.
70. I. Volobouev, "On the Expectation-Maximization Unfolding with Smoothing", arXiv:1408.6500.

71. M. Kuusela and V. M. Panaretos, "Statistical unfolding of elementary particle spectra: Empirical Bayes estimation and bias-corrected uncertainty quantification", doi:10.1214/15-AOAS857, arXiv:1505.04768.
72. M. Kuusela and P. B. Stark, "Shape-constrained uncertainty quantification in unfolding steeply falling elementary particle spectra", arXiv:1512.00905.

Chapter 4
Absolute Cross Sections

Absolute cross sections are the most fundamental measurements to be made in collision experiments. They represent the proportionality constant between the luminosity characterising the performance of a particle accelerator and event count rates in an experiment. As such they are subject to systematic effects originating from the collider (luminosity) and the detectors (acceptances, efficiencies, calibrations, resolutions). On the theory side, missing higher orders, imprecise knowledge of the proton structure and the strong coupling constant as well as nonperturbative effects have to be taken into account.

After introducing some basic terminology in the next section, Sections 4.2 and 4.3 present measurements of inclusive jet, dijet, and 3-jet cross sections. Subsequently, the measurements are discussed in Sections 4.4 and 4.5 with respect to their use in determining SM parameters like the strong coupling constant $\alpha_S(M_Z)$ and the proton PDFs.

4.1 Cross-Section Terminology

The most basic measurement in collision experiments consists in counting the produced entities of interest, e.g. certain particle types or jets, and to categorise them with respect to their kinematic quantities. Of course, the counts depend on the performance of the respective particle collider, which is expressed in terms of its luminosity $\mathcal{L}$. Measured reaction rates $\dot{N}$ are then connected via

$$\dot{N} = \mathcal{L} \cdot \sigma \tag{4.1}$$

to the underlying physical reaction, which is characterised by its cross section σ that is independent of experimental conditions like the collider performance. Integrating over a period of data taking, this can be written in terms of the integrated luminosity $\mathcal{L}_{\text{int}}$ as

$$\sigma = \frac{N}{\mathcal{L}_{\text{int}}} \,. \tag{4.2}$$

The measured cross section can now be compared to predictions of any adopted theory, e.g. the Standard Model of particle physics and in particular QCD, as detailed in Chapter 2. To make a kinematic dependence explicit, for example on a transverse momentum p_T, the differential form

$$\frac{d\sigma}{dp_T} = \frac{1}{\varepsilon \cdot \mathcal{L}_{int}} \cdot \frac{N}{\Delta p_T} \tag{4.3}$$

can be used, where Δp_T represents the measurement interval in p_T and the additional factor ε accounts for experimental inefficiencies.

Two further distinctions can be mentioned: First, if selected entities are counted irrespective of everything else that is produced within the same event then the cross section is *inclusive*. In contrast, *exclusive* cross sections set e.g. limits on the existence of further jets in an event, so-called *jet vetoes*. As a consequence, exclusive cross sections are sensitive to all energy depositions or particles in an event and are more difficult to treat experiment- and theory-wise. Therefore most cross sections discussed in this book are of the inclusive type.

Secondly, if each entity, for example a jet, satisfying the selection criteria for an analysis, is counted, then this is a jet cross section with two entries per event at LO in pQCD. For comparison, the corresponding inclusive dijet cross section as a function of the average p_T of the two jets leading in p_T receives only one entry per event. As a consequence, at LO the inclusive jet cross section is twice as large as the corresponding dijet event cross section. At the same time, since the two jets within one event are not produced independently from each other, the statistical uncertainty has to account for this correlation in jet cross sections as explained in Section 4.2.

4.2 Inclusive Jet Cross Section

The benchmark observable of jet physics is the inclusive jet production cross section as a function of the jet p_T and, if feasible, of the jet rapidity y. Measured jet yields are transformed into a double-differential cross section via

$$\frac{d^2\sigma}{dp_T\,dy} = \frac{1}{\varepsilon \cdot \mathcal{L}_{int}} \cdot \frac{N_{jets}}{\Delta p_T \Delta y} \tag{4.4}$$

where N_{jets} is the number of jets counted within a bin and corrected for detector distortions, ε is the experimental efficiency, and Δp_T and Δy are the bin widths in jet transverse momentum and rapidity. Since jets in pp or $p\bar{p}$ collisions with equal beam energies are symmetrically distributed in y, the binning is done in terms of the absolute value $|y|$. In that case a factor of two has to be taken into account between bin widths in y and $|y|$: $\Delta y = (2 \cdot \Delta|y|)$.

The first measurement of inclusive jet production has been performed in 1982 by the UA2 Collaboration at the Sp$\bar{p}$S collider for a centre-of-mass energy of 540 GeV [1]. The observed steep decrease of the jet p_T spectrum proportional to

p_{T}^{-n} with $n \approx 4$–6 was correctly predicted by theory at LO [2]. Firm conclusions on the absolute normalisation, however, were not possible because of large experimental and theoretical uncertainties, and lack of a well-defined jet algorithm, cf. Section 2.5. Further measurements have been conducted by experiments at the Sp$\bar{\mathrm{p}}$S, Tevatron, and RHIC colliders at pp or $p\bar{p}$ centre-of-mass energies of 540, 546, and 630 GeV [3–5], at 546 GeV, 630 GeV, 1.8 TeV, and 1.96 TeV [6–12], and at 200 GeV [13], respectively. Figure 4.1 gives an overview of data-theory comparisons for various inclusive jet measurements in $p\bar{p}$ and pp collisions, including first results from the LHC, and in deep-inelastic scattering. It is demonstrated that, within uncertainties, pQCD at NLO adequately describes a huge range of inclusive jet measurements from diverse experiments, processes, and at various centre-of-mass energies.

At the LHC, inclusive jet cross sections have been determined at 2.76 and 7 TeV centre-of-mass energies by the ALICE [16], ATLAS [17–20], and CMS Collaborations [21–25]. Analyses of the data taken in 2012 at $\sqrt{s} = 8$ TeV are still ongoing. All experiments employ the collinear- and infrared-safe anti-k_{t} jet algorithm [26] as implemented in the FastJet package [27], cf. Section 2.5, but with different jet sizes of $R = 0.2$ and 0.4, 0.4 and 0.6, and 0.5, 0.7 for ALICE, ATLAS, and CMS, respectively. Concentrating on the latest results at $\sqrt{s} = 7$ TeV, ATLAS and CMS have published the inclusive jet cross sections for their respective choices of jet sizes double-differentially in jet p_{T} and $|y|$ [20, 23, 24]. The binning in p_{T} approximately follows the jet p_{T} resolution and ranges from 100 GeV (CMS: 60 GeV) up to 2 TeV, while in rapidity equally-sized bins of $\Delta|y| = 0.5$ up to $|y| = 3.0$ have been chosen. Both collaborations have evaluated the full 2011 data set corresponding to 4.5 fb^{-1} (ATLAS) and 5.0 fb^{-1} (CMS) of integrated luminosity. The jet detection efficiencies ε differ by less than 1% from full efficiency and are compatible with 100% within uncertainties—in contrast to first measurements at the Sp$\bar{\mathrm{p}}$S collider with $\varepsilon > 80$–90% only.

Figure 4.2 presents these double-differential cross sections on the left from ATLAS together with predictions at NLO complemented with NP and EW corrections and on the right from CMS compared to particle-level predictions of POWHEG NLO matched to the PS of PYTHIA6 with tune Z2* [28], also multiplied by EW factors according to Eq. 4.6. In both cases, the data are well described by the theory over nine orders of magnitude in cross section and two orders of magnitude in jet p_{T}.

The most complicated experimental tasks consist of calibrating the jet energy scale and unfolding effects of the jet energy resolution and detector inefficiencies as described in Sections 3.4–3.6.2. The dominating source of experimental systematic uncertainty is the JEC which is known to about 2–4% precision depending on jet p_{T} and pseudorapidity η. Because of the steep slope of the p_{T} spectrum $\propto p_{\mathrm{T}}^{-5}$ this translates into an uncertainty on the cross section of 10–20%. When reaching the limit of kinematically accessible phase space, for example at higher rapidity, this effect becomes even worse. The second largest uncertainty is caused by the JER, which smears out the jet p_{T} on top of this steep spectrum and leads to a significant overestimation of the cross section because more jets migrate upwards in p_{T} than downwards. In both experiments this is corrected by iterative unfolding methods

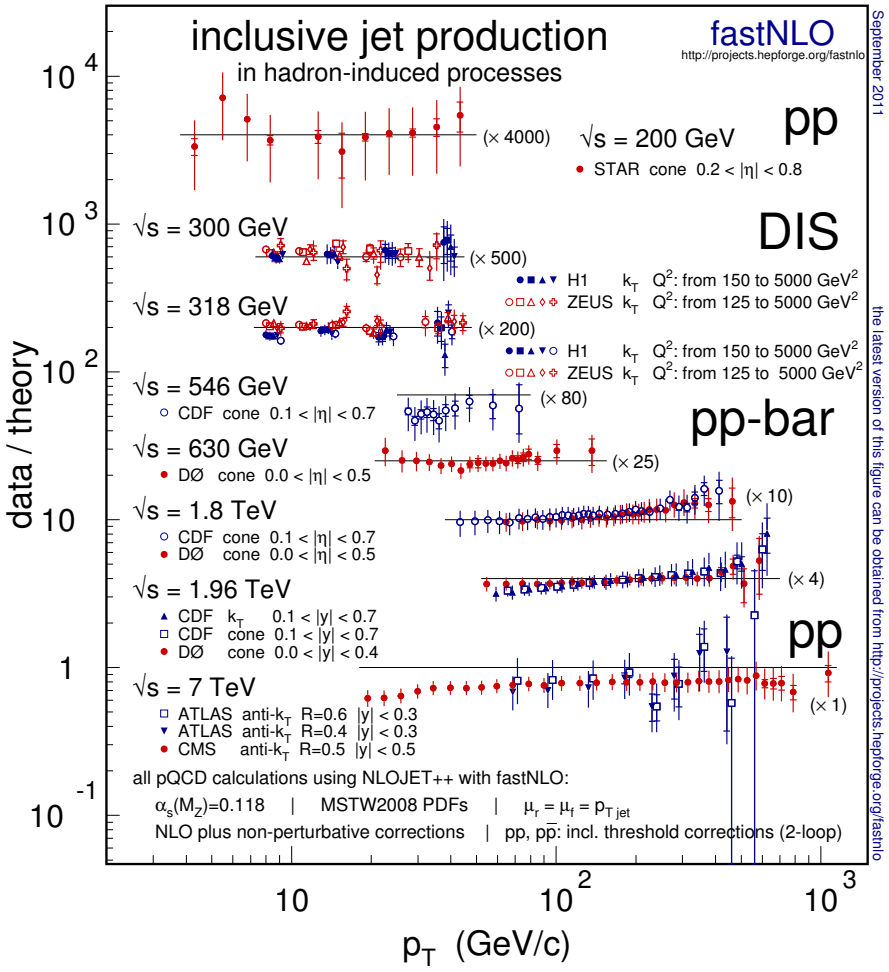

Fig. 4.1 Overview of data over theory comparisons for inclusive jet production as a function of jet p_T in hadron-hadron collisions and in deep-inelastic scattering at various centre-of-mass energies. The NLO predictions have been computed using the MSTW2008 PDFs [14]. *(Taken from Ref. [15])*

based on detector response matrices. Examples of such response matrices, exhibiting limited off-diagonal smearing, have been presented previously in the unfolding Section 3.26. Imprecise knowledge of the exact detector smearing properties, differences between data and simulations, and choices for the physics models in the simulations imply the resulting JER or unfolding uncertainty.

Finally, the limited accuracy of the luminosity determination induces a normalisation uncertainty of 1.8% for ATLAS and 2.2% for CMS. Other sources of systematic effects like trigger and selection efficiencies or angular resolutions contribute only a small fraction to the total uncertainty. Figure 4.3 shows the size of these

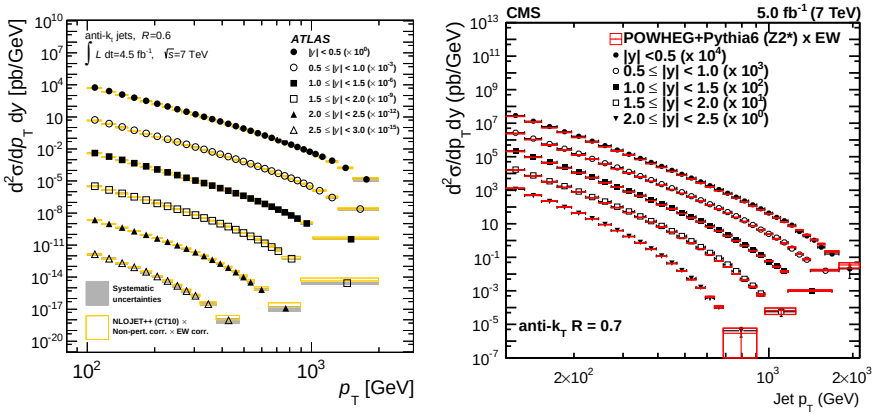

Fig. 4.2 Double-differential inclusive jet cross section at $\sqrt{s} = 7\,\text{TeV}$ from ATLAS in comparison to predictions at NLO complemented with NP and EW corrections (left), and from CMS compared to POWHEG NLO + PYTHIA6 PS and tune Z2*, also multiplied by EW factors. *(Taken from Refs. [20, 29])*

uncertainties for CMS at small rapidity $|y| < 0.5$ (left), and for ATLAS at medium rapidity $1.5 \leq |y| < 2.0$ (right). Evidently, the detector understanding is best at small rapidity, i.e. perpendicular to the beam directions, and slowly deteriorates, especially for the JER, towards smaller angles with respect to the beams. At similar angles ATLAS and CMS perform comparably.

Statistical uncertainties are important at the highest accessible p_T's, but may as well contribute significantly to the total uncertainty at small transverse momenta because of trigger pre-scales, cf. Section 3.3. In addition, for the particular case of the inclusive jet cross section the correlated production of jets within the same event has to be accounted for. This can be achieved by including this correlation into the propagation of uncertainties via pseudo-experiments in the unfolding procedure as done by ATLAS [20], or by means of an explicit correction like in CMS [23] with

$$e_{\text{stat}} = \sqrt{(4-3f)/(2-f)} \cdot \sqrt{N_{\text{jets}}} \tag{4.5}$$

instead of simply $e_{\text{stat}} = \sqrt{N_{\text{jets}}}$. Here, $f = N_1/N_{\text{ev}}$ is the fraction of events that contribute one jet to a given bin. The formula is valid as long as the number of events that contribute more than two jets to a bin is negligible.

On the theory side the process of inclusive jet production, whose cross section is proportional to α_S^2 at LO, is calculated up to NLO using NLOJET++ [30, 31]. For the derivation of the associated theoretical uncertainties, fast re-evaluation techniques [32, 33] for different choices of scales, PDFs, or values of $\alpha_S(M_Z)$ are employed.

The central scale chosen by ATLAS is $p_{\text{T,max}}$, while CMS prefers $p_{\text{T,jet}}$. The scale uncertainties, often found to be asymmetric, are estimated following the six-point variation of the μ_r and μ_f scale factors as described in Section 2.6.1, and range from

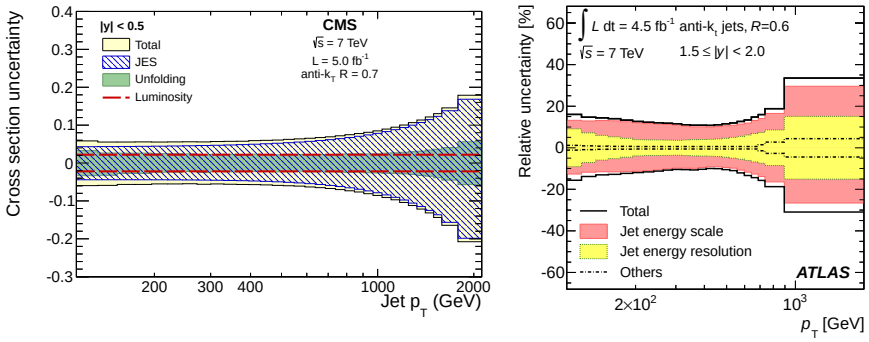

Fig. 4.3 Total experimental systematic uncertainty and most important components at small rapidity $|y| < 0.5$ from CMS (left) and at medium rapidity $1.5 \leq |y| < 2.0$ from ATLAS (right). CMS includes JER effects in the unfolding and a luminosity uncertainty of 2.2% in the total, while the luminosity uncertainty of 1.8% is not contained in the ATLAS plot. *(Taken from Refs. [20, 23])*

5–30%. PDF uncertainties are evaluated according to the respective prescriptions of the PDF groups, cf. Table 2.1 and Section 2.6.2. They become more important, when either very large or very small parton momentum fractions x of the proton are probed, i.e. at large jet p_{T} or at high rapidity $|y|$, where they can reach values of 50% or more. The development of elaborate techniques to determine PDFs including uncertainty estimates over the last 20 years constitutes a substantial progress. The observation of an excess of jet production at high p_{T} by the CDF Collaboration in 1996 [34], when PDF uncertainty estimates did not exist, triggered many speculations with respect to new phenomena. Finally, the excess could be remedied by adaptations in the gluon PDF [35]. To test the compatibility between data and theory, the limited precision of $\alpha_S(M_Z)$ has to be considered leading to a roughly constant additional uncertainty of 2–4%. Figure 4.4 gives an overview of these uncertainties for small rapidity from CMS (left) and ATLAS (right).

Because fixed-order predictions are at parton-level only, one has to apply corrections for the non-perturbative effects of MPI and hadronisation. Habitually, they are estimated from the ratio of distributions for fully hadronised events over the distributions with MPI and hadronisation switched off in the respective LO+PS MC event generators. An envelope is constructed around the predictions by different event generators with various tunes to derive a medium correction and to attribute a systematic uncertainty of half-width of the spread to this factor. Figure 4.5 left displays the NP correction with uncertainty at small rapidity for the ATLAS analysis. As expected the NP effects are negligible for high transverse momenta, but can become significant for $p_{\mathrm{T,jet}} < 300\,\mathrm{GeV}$. Because the p_{T} dependence of the various effects of PS, MPI, and hadronisation differs as detailed in Section 2.4, the NP corrections are sensitive to the choice of jet algorithm and jet size.

For jet p_{T}'s exceeding 1 TeV, EW corrections comprising tree-level effects of $\mathcal{O}\left(\alpha\alpha_S, \alpha^2\right)$ and loop effects of $\mathcal{O}\left(\alpha\alpha_S^2\right)$ get increasingly important, cf. Section2.3.2. Following Refs. [36, 37] the EW corrections have been applied as factors

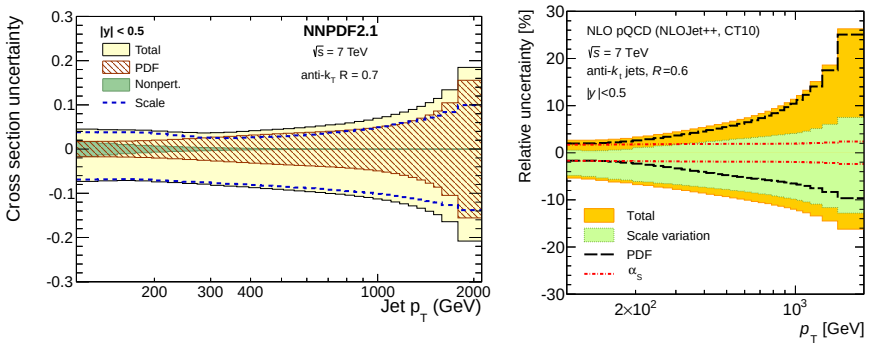

Fig. 4.4 Breakdown of theoretical uncertainties as estimated by CMS (left) and ATLAS (right) at small rapidity. In this plot, ATLAS only includes effects derived from fixed-order calculations and not the NP uncertainty, while CMS presents uncertainties as used later in fits of $\alpha_S(M_Z)$ and consequently omits α_S uncertainties. *(Taken from Refs. [20, 23])*

to the NLO QCD prediction. Hence, a cross section differential in some observable $\mathcal{X}$ is written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathcal{X}} = \frac{\mathrm{d}\sigma}{\mathrm{d}\mathcal{X}_{\mathrm{QCD}}} \cdot \left[1 + \delta_W^{\mathrm{NLO}}(\mathcal{X})\right] , \tag{4.6}$$

where $\delta_W^{\mathrm{NLO}}(\mathcal{X})$ is the relative EW correction and purely photonic corrections $\mathrm{d}\sigma/\mathrm{d}\mathcal{X}_\gamma$ have been neglected. The result shown in Fig. 4.5 right has been calculated by the authors of Ref. [37] for the phase space of the ATLAS measurement. Negligible below p_{T}'s of $\approx 800\,\mathrm{GeV}$, the EW corrections attain values of 13% or more beyond 1 TeV and are much more pronounced at small rapidity. Since fast interpolation techniques for different choices of PDFs or scales as explained in Section 2.6.5 have not yet been interfaced to the EW calculations, the factors are currently applied without attributing appropriate uncertainties, which is acceptable as long as the factors are small.

Putting everything together and including correlations in the uncertainties, one can now test the compatibility between data and theory. For a visual impression one can look at Fig. 4.6 left, where the ratio of theory over data is shown for three different PDF sets entering the fixed-order calculations and for rapidities up to $|y| = 1.5$. All predictions seem to agree with the data within uncertainties. In contrast, this is not anymore the case when using the ABM11-NLO PDF set, which for small rapidities systematically underestimates the cross section as demonstrated in Fig. 11 of Ref. [20]. The final judgement, however, requires a quantitative comparison as performed by ATLAS using a generalised χ^2 test that is described in [38]. As a result, cf. Tables 1 and 2 of Ref. [20], all PDF sets investigated by ATLAS lead to a good description of the measurements except for the ABM11-NLO set, which fails for rapidities below 1.5, and HERAPDF1.5-NLO, which exhibits deviations for $|y| < 1.0$, but only for the larger jet size of $R = 0.6$.

Figure 4.6 right displays the same ratio to ATLAS data, but this time from predictions of POWHEG [39, 40] at NLO matched to the parton showering and hadro-

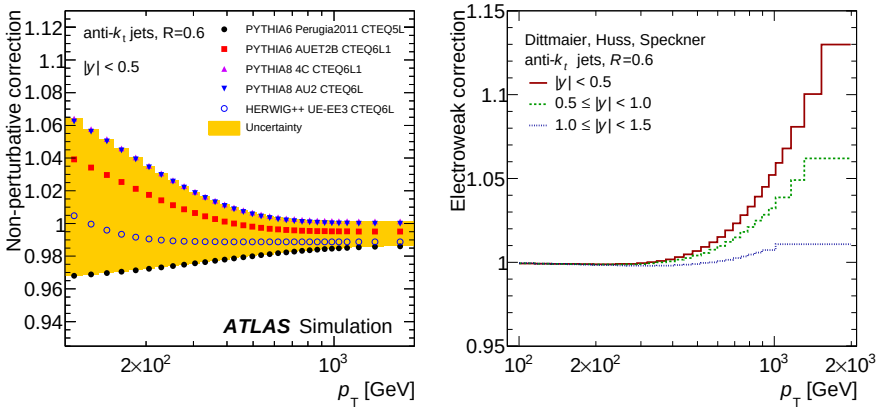

Fig. 4.5 NP (left) and EW (right) corrections as a function of jet p_T for the ATLAS inclusive jet analysis. *(Taken from Ref. [20])*

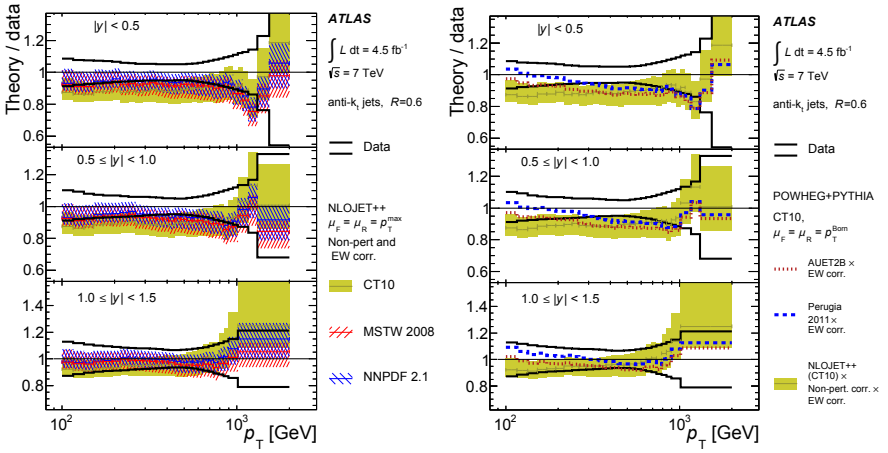

Fig. 4.6 Theory over data ratios from ATLAS including uncertainty bands for NLO predictions using three different PDF sets (left) and for NLO+PS matched predictions from POWHEG +PYTHIA6 (right). *(Taken from Ref. [20])*

nisation model of PYTHIA6 [41], similar as in Fig. 4.2 right for CMS. POWHEG +PYTHIA6 agree with the data for both examined tunes. In principle, the combination of NLO predictions with PS promises to describe a wider range of observables and phase space. However, the crosstalk between the tuning of NP parameters of the LO+PS MC event generators and the NLO+PS matching, and the yet unclear situation how systematic uncertainties are to be determined for the matched predictions, prevents quantitative statements. Establishing NLO+PS matched calculations for future improved pQCD predictions is under active development.

4.3 Inclusive Dijet and 3-Jet Cross Sections

The simplest jet production process corresponds to a $2 \to 2$ reaction with the two outgoing partons fragmenting into a pair of jets. Of course, these two jets are kinematically correlated and in particularly balanced in p_T—a fact that is largely ignored in the inclusive jet cross section. To account for such correlations, the inclusive dijet production cross section is defined as a function of the kinematic properties of the two jets leading in p_T in an event. An obvious choice for one such quantity is the dijet mass, implicitly defined through $M_\mathrm{jj}^2 = (p_1 + p_2)^2$, where p_1 and p_2 are the four-momenta of the two leading p_T jets. The resonant production of new particles that show up as jet pairs in the final state would appear as bumps in the invariant mass distribution on top of the QCD prediction.

Furthermore, assuming massless partons and ignoring any initial parton p_T within the colliding hadrons, the parton fractional momenta x_1 and x_2 can be related at LO to the two emerging jets by means of four-momentum conservation to

$$x_1 = \frac{x_\mathrm{T}}{2}\left(e^{y_1} + e^{y_2}\right) \qquad \text{and} \qquad x_2 = \frac{x_\mathrm{T}}{2}\left(e^{-y_1} + e^{-y_2}\right), \tag{4.7}$$

where $x_\mathrm{T} = 2p_\mathrm{T}/\sqrt{s} = p_\mathrm{T}/E_\mathrm{beam}$ and y_1, y_2 are the rapidities of the two jets.

Two definitions are conventionally used for a second, rapidity-related kinematic quantity that provides a separation of the phase space into exclusive bins. The ATLAS Collaboration characterises the dijet system in terms of half of the rapidity separation of the two jets leading in p_T: $y^* = |y_1 - y_2|/2$. y^* is a longitudinally boost-invariant quantity and corresponds to the rapidity of the leading jet in the two-parton centre-of-mass frame, where it can be written in terms of the polar scattering angle θ^* with respect to the beam axis as:

$$y^* = \frac{1}{2}\ln\left(\frac{1 + |\cos\theta^*|}{1 - |\cos\theta^*|}\right). \tag{4.8}$$

For a $2 \to 2$ process the numbering of the two outgoing partons (or jets) is arbitrary, if parton flavours are ignored. Therefore it is always possible to attribute labels such that $y_1 \geq y_2$. In that case θ^*, defined with respect to parton 1, is restricted to the interval $[0, \pi/2]$ and $\cos\theta^* \in [0, 1]$. The second jet appears at $-y^*$ in the opposite direction. The absolute values in the definition of y^* and in Eq. (4.8) can then be omitted. Expressing the dijet mass in terms of the jet p_T and y^*, $M_\mathrm{jj} = 2p_\mathrm{T}\cosh y^*$, the double-differential cross section reads:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}M_\mathrm{jj}\,\mathrm{d}y^*} = \frac{1}{\varepsilon \cdot \mathcal{L}_\mathrm{int}} \cdot \frac{N}{\Delta M_\mathrm{jj} \Delta y^*}, \tag{4.9}$$

where $\mathcal{L}_\mathrm{int}$ is again the integrated luminosity and ε the experimental efficiency. N is the number of events counted within a bin of M_jj and y^*.

Defining in addition the boost of the centre-of-mass system versus the laboratory system, $y_\mathrm{b} = (y_1 + y_2)/2$, the parton fractional momenta can be rewritten to

$$x_1 = x_\mathrm{T} e^{y_\mathrm{b}} \cosh y^* \,, \qquad x_2 = x_\mathrm{T} e^{-y_\mathrm{b}} \cosh y^* \qquad \text{giving} \qquad y_\mathrm{b} = \frac{1}{2} \ln \frac{x_1}{x_2} \,. \quad (4.10)$$

CMS, in contrast, employs the larger of the two absolute rapidities of the two jets, which is not longitudinally boost-invariant, but has the advantage of maintaining a closer relation to where the jets are localised in the detector. Defining y_max as the signed quantity

$$y_\mathrm{max} = \mathrm{sign}(|\max(y_1, y_2)| - |\min(y_1, y_2)|) \cdot \max(|y_1|, |y_2|) \,, \quad (4.11)$$

the double-differential cross section for the CMS case can be written in a way similar to the inclusive jet cross section, $\mathrm{d}^2\sigma/\mathrm{d}p_\mathrm{T}\,\mathrm{d}y$, including a factor of 2 for rapidity bin widths in terms of $|y_\mathrm{max}|$ instead of $|y|$:

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}M_\mathrm{jj}\,\mathrm{d}y_\mathrm{max}} = \frac{1}{\varepsilon \cdot \mathcal{L}_\mathrm{int}} \cdot \frac{N}{\Delta M_\mathrm{jj} (2\Delta |y|_\mathrm{max})} \,. \quad (4.12)$$

The absolute value of y_max is equal to the maximum $|y|$ of the two leading jets denoted $|y|_\mathrm{max}$.

As in the case of the inclusive jet cross section, the LO for the inclusive dijet production is proportional to α_S^2. Requiring a third jet in the final state, the power of the leading term in the perturbative expansion is incremented by one and the cross section becomes proportional to α_S^3. For CMS, the kinematic definitions of M_jj and y_max can simply be extended to give the 3-jet mass, m_3, and the maximal rapidity y_max respectively $|y|_\mathrm{max}$, which are then derived from the four-momenta p_i and the rapidities y_i of the i^th jet of the three jets leading in p_T. The rapidity separation as used by ATLAS is modified for 3-jet production to:[1]

$$Y^* = |y_1 - y_2| + |y_2 - y_3| + |y_1 - y_3| \,. \quad (4.13)$$

Measurements of dijet cross sections as a function of dijet mass have been performed by the UA2 Collaboration at the Sp$\bar{\mathrm{p}}$S collider [1], by AFS at the ISR [42], and by the CDF and D0 collaborations at the Tevatron [43–46]. Alternatively, dijet events have also been studied triple-differentially in transverse energy, and the pseudorapidities η_1 and η_2 of the two leading jets [47, 48]. This might even better exploit the available information with respect to constraints on the proton PDFs and notably the gluon. However, care has to be taken, because ordering the jets, which are balanced in p_T at LO, and labelling their rapidities accordingly is not infrared-safe as explained in [49, 50]. Only quantities are permitted that remain unaltered when jet 1 and jet 2 are exchanged, for example because of the emission of a soft gluon. As a solution both jet orderings can be accepted as suggested in [49], or one can use y^* and y_b instead as done in [51]. Measurements at the LHC so far have been reported for 7 TeV centre-of-mass energy as a function of dijet mass and either y^* or y_max by the ATLAS and CMS collaborations [17, 18, 23, 38, 52]. The first two of

[1] ATLAS unnecessarily is using the absolute value of Y^*, which will be avoided here.

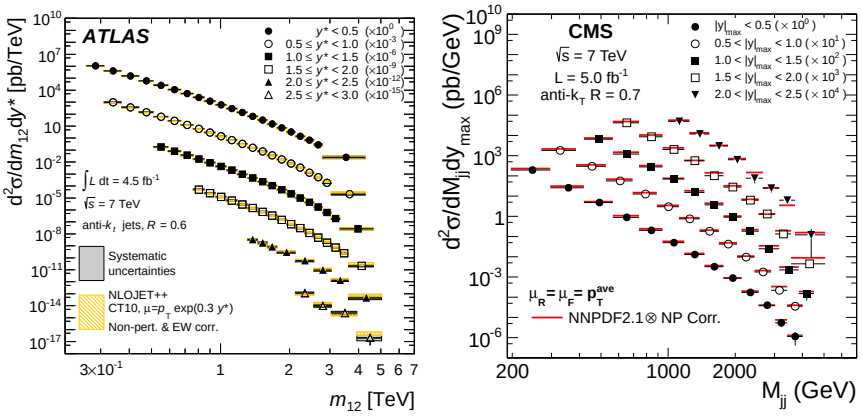

Fig. 4.7 Double-differential inclusive dijet cross section at $\sqrt{s} = 7\,\text{TeV}$ as a function of dijet mass (m_{12}, M_{jj}) and rapidity separation y^* (left) or maximal absolute rapidity y_{max} (right) in comparison to predictions at NLO complemented with NP corrections. The theory for ATLAS in the left plot also accounts for EW effects. *(Taken from Refs. [23, 38])*

these publications evaluate the complete data sets recorded in 2011, which amount to $4.5\,\text{fb}^{-1}$ of integrated luminosity for ATLAS and $5.0\,\text{fb}^{-1}$ for CMS.

Figure 4.7 presents the double-differential inclusive dijet cross sections on the left for the more recent ATLAS publication together with predictions at NLO complemented with NP and EW corrections and on the right for CMS without EW corrections that were not yet available at the time of publication. Again, the data are well described by theory over eight orders of magnitude in cross section, over a large range in dijet mass, and up to jet rapidities of 3. For triggering purposes, minimal jet p_{T} thresholds of 100 GeV (60 GeV) and 50 GeV (30 GeV) are imposed on the leading (subleading) jet by ATLAS and CMS respectively. Both experiments employ single-jet triggers as for the inclusive jet cross section and ensure close to full efficiency within the analysed dijet mass phase space. By using a two-jet trigger strategy combining signals for both leading jets, ATLAS gains about 10% in amount of data compared to single-jet triggering for the leading jet alone and thereby reduces losses caused by trigger pre-scales.

Experimental systematic uncertainties are largely dominated by the JEC as shown for CMS in Fig. 4.8 left and amount to 5–8% up to 2 TeV in dijet mass at small rapidity. In this region, theoretical systematic uncertainties, which are dominated by NP effects for small and by scale as well as PDF uncertainties for large dijet masses, cf. 4.8 right, exceed the experimental ones so that parameters of the theory might be constrained. For masses beyond 2 TeV, measurements become less accurate.

In the CMS analysis the QCD scales μ_r and μ_f have been chosen to be the average p_{T} of the leading two jets, $\langle p_{\text{T1,2}} \rangle$, leading to the scale uncertainties discussed above. For the y^* based phase-space separation, ATLAS tried to employ the same

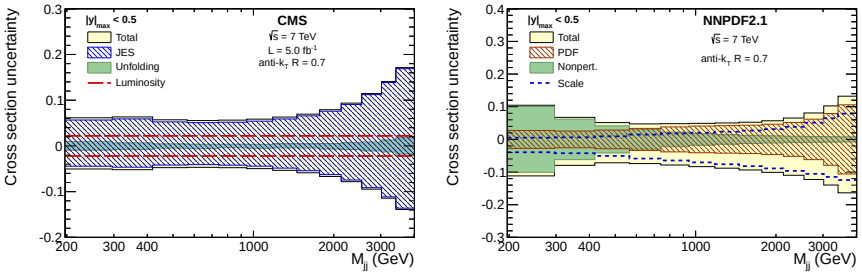

Fig. 4.8 Total experimental systematic uncertainty and most important components of JEC, unfolding, and luminosity uncertainty (left) in comparison to total theoretical systematic uncertainty and its most important components of scale, PDF, and NP uncertainty (right), both at small rapidity $|y|_{\max} < 0.5$ for the CMS measurement. *(Taken from Ref. [23])*

scale as they use for inclusive jets, i.e. $p_{\mathrm{T,max}}$. However, with this choice the cross sections evaluate to negative values for the larger y^* intervals. Motivated by a discussion in [53], ATLAS chooses bins in

$$\mu = \mu_r = \mu_f = p_{\mathrm{T,max}} \cdot e^{(0.3 \cdot y^*)} \tag{4.14}$$

as scale, which follows from a comparison to (twice) the original suggestion in [53]:

$$\mu = \mu_r = \mu_f = \frac{M_{\mathrm{jj}}}{2} \cdot \frac{1}{2 \cdot \cosh\left(0.7 \cdot y^*\right)} . \tag{4.15}$$

With this definition the dijet cross sections remain positive at large y^* even for the usual scale variations by factors of $1/2$ and 2.

Although dijet masses of 200 GeV and beyond are investigated, NP corrections and associated uncertainties can be large, because they scale with jet p_{T}. In Fig. 4.9 left the NP correction as derived by ATLAS reaches 10% at a dijet mass of 600 GeV for a jet size of $R = 0.6$. For CMS with smaller minimal jet p_{T}'s and a larger jet size of $R = 0.7$ this correction even rises up to 20% at $M_{\mathrm{jj}} = 200$ GeV. However, a sizeable correction does not necessarily mean a sizeable uncertainty, as can be seen from a comparison between the NP corrections for dijet mass events with $R = 0.6$ and $R = 0.4$. Although being less susceptible to MPI effects because of the smaller jet size, the NP uncertainty nevertheless is larger in the case of the smaller jet cone.

Finally, Fig. 4.9 right demonstrates that EW effects get increasingly relevant for dijet masses beyond 2 TeV and small rapidity separations y^* between the two leading jets. It is not known yet how they behave exactly for a phase space binning in $y_{\max}$ instead, but since small $|y|_{\max}$ automatically induces a small y^* one can expect them to be of a similar size.

Advancing to 3-jet production, previously measured by the D0 Collaboration [54], the data presented here were recorded during the 2011 data-taking period at 7 TeV centre-of-mass energy and correspond to an integrated luminosity of 4.5 (5.0) fb^{-1} for ATLAS [55] and CMS [56], respectively. For ATLAS, minimal p_{T} thresholds

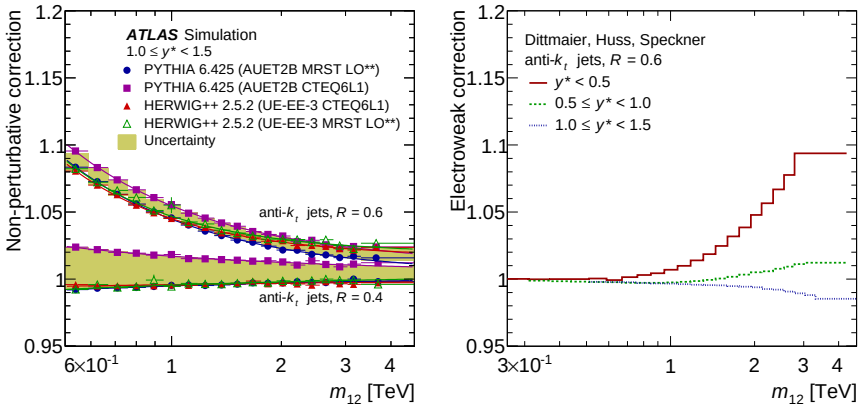

Fig. 4.9 NP and EW corrections as a function of dijet mass m_{12} from ATLAS. The NP factors are shown for the two jet sizes $R = 0.4$ and $R = 0.6$ of the anti-k_t jet algorithm at medium y^* between 1.0 and 1.5. The EW corrections by the authors of Ref. [37] are presented for $R = 0.6$ and three ranges in y^* up to $y^* = 1.5$. *(Taken from Ref. [38])*

of 150, 100, and 50 GeV are imposed on the three leading jets up to $y = 3.0$, while CMS requires the same minimal value of 100 GeV for each of the three jets up to $|y|_{\max} = 2.0$. Figure 4.10 left displays the double-differential cross section as a function of m_{jjj} and Y^* from ATLAS in comparison to NLO predictions computed with NLOJET++ [30, 31] and based on the CT10-NLO PDF set [57]. Within uncertainties agreement is observed similarly as for the CMS measurement with bins in $y_{\max}$, where for better visibility the ratio of data to theory is shown on the right of Fig. 4.10 for various PDF sets. Only the ABM11-NLO PDF set [58] predicts somewhat too small cross sections, potentially caused by its smaller gluon density compared to the alternative PDF sets and visible also for other jet cross sections.

Lacking any significant deviations, in particular the CMS measurement within $|y|_{\max} < 1.0$ can be used to set limits on the mass of gluophilic Z' bosons. The associated Z' production mode investigated in Ref. [59] requires a third jet from ISR such that nonresonant deviations are expected from SM predictions of the 3-jet mass distribution.

The triggering of 3-jet events is more involved. ATLAS determines trigger efficiencies as a function of m_{jjj} in each bin of Y^* using an unbiased sample of events collected with the trigger for a jet p_{T} threshold of 30 GeV. This trigger is fully efficient in events with a leading jet passing the 3-jet analysis requirements. Within each Y^* bin, the range in 3-jet mass is divided into subranges, each of which is associated to one more than 99% efficient single-jet trigger. For CMS, the single-jet triggers define mutually exclusive regions in leading jet p_{T}, in which each respective trigger is more than 99% efficient. The final 3-jet mass spectrum is obtained by summing the spectra measured with the different triggers while taking trigger pre-scale factors into account. In the inner rapidity region, most single-jet triggers contribute up to 50% of the final event yield, with the exception of the two triggers with the lowest

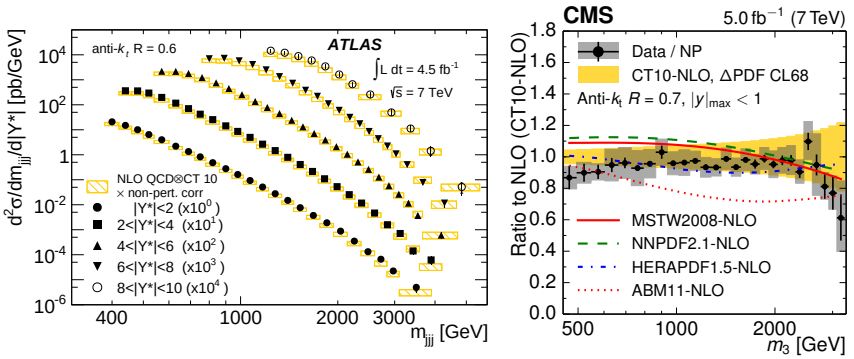

Fig. 4.10 Double-differential inclusive 3-jet cross section at $\sqrt{s} = 7\,\text{TeV}$ as a function of 3-jet mass (m_{jjj}, m_3) and rapidity separation Y^* from ATLAS together with theory at NLO (left) and ratio to NLO theory for a maximal absolute rapidity of $y_{\text{max}} < 1$ from CMS (right). In both cases, pQCD predictions are derived for the CT10-NLO PDF set and complemented with NP corrections. In the right plot, alternative PDF sets are employed as well. *(Taken from Refs. [55, 56])*

and highest threshold, which contribute up to 80 and 100% respectively. In the outer rapidity region, each jet trigger contributes over a large range of 3-jet masses around 25% to the measurement, again except for the lowest and highest threshold triggers.

For the phase space division with Y^*, the JEC in ATLAS is by far the dominant source of experimental uncertainty ranging from 8% up to 26% in the $Y^* < 2.0$ bin as shown in Fig. 4.11 left. In case of the CMS analysis, cf. Fig. 4.11 right, JER and statistical uncertainties, convolved through the unfolding procedure, have a larger share of the total experimental uncertainty and dominate even at the smallest and largest 3-jet masses. The total experimental uncertainties between ATLAS and CMS are of comparable size as are the theoretical ones, which are presented in Fig. 4.12. Here, the renormalisation and factorisation scales have been set to the 3-jet mass by ATLAS and half the 3-jet mass by CMS.

Up to now NP corrections have been derived with the help of the "LO+PS" MC event generators PYTHIA6, PYTHIA8, and HERWIG++. For 3-jet production, however, these programs have not even LO accuracy. The third jet always originates either from the PS or, worse, from MPI. Therefore, CMS evaluates the NP corrections and their uncertainty employing the multi-jet improved MC event generators MADGRAPH5 +PYTHIA6 and SHERPA instead. The result for both y_{max} regions is shown in Fig. 4.13 right. Because NP factors depend on jet p_{T} and only indirectly on 3-jet mass and $|y|_{\text{max}}$, the function used to fit this dependence is different from the usual shape as given by Eq. (2.38). Instead, a logistic function

$$f(x) = 1 + \frac{p_0 - 1}{\left(1 + \log(x)^{\frac{p_2}{p_1}}\right)^{\frac{p_3}{p_2}}} \tag{4.16}$$

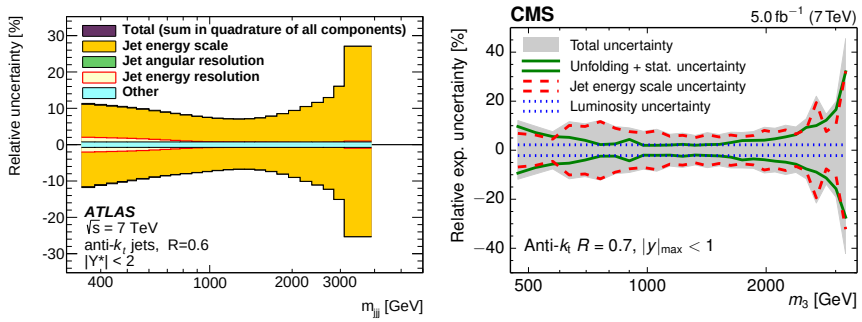

Fig. 4.11 Total experimental systematic uncertainty and most important components for $Y^* < 2$ from ATLAS (left) and for $|y|_{\max} < 1$ from CMS (right). CMS includes JER and statistical effects in the unfolding and a luminosity uncertainty of 2.2% in the total uncertainty, while the luminosity uncertainty of 1.8% or the statistical component is not contained in the ATLAS plot. *(Taken from Refs. [55, 56])*

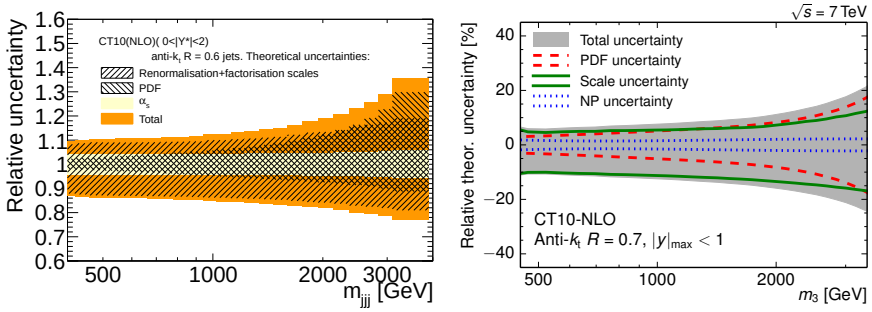

Fig. 4.12 Breakdown of theoretical uncertainties as estimated by ATLAS (left) and CMS (right) at small $Y^* < 2$ respectively $|y|_{\max} < 1$. In this figure, ATLAS only includes effects derived from fixed-order calculations but not the NP uncertainty, while CMS presents uncertainties as used later in fits of $\alpha_S(M_Z)$ and consequently omits α_S uncertainties. *(Taken from Refs. [55, 56])*

is employed that better adapts to a plateau region at small m_3, particularly for the outer $|y|_{\max}$ region. The contribution of events with low p_{T} jets to this phase space region is limited not because of 3-jet mass, but by the minimal p_{T} of 100 GeV required for each of the three leading jets. In contrast, ATLAS derives the NP correction and uncertainty the usual way, which is doubtful in the context of 3-jet production. The ensuing rather small average corrections are displayed in Fig. 4.13 left. EW corrections to 3-jet observables have not yet been calculated.

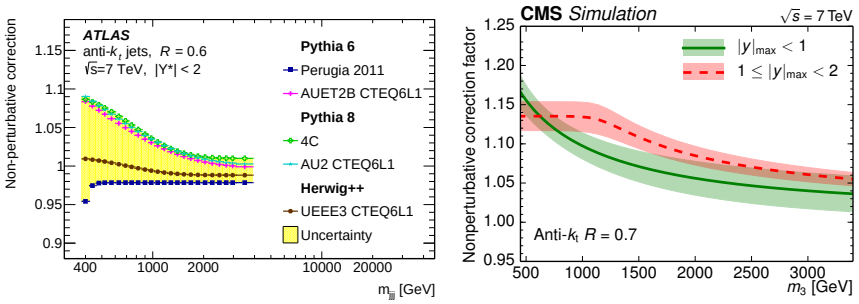

Fig. 4.13 NP corrections as a function of 3-jet mass m_{jjj} (m_3) for ATLAS (left) and for CMS (right). *(Taken from Refs. [55, 56])*

4.4 Determination of the Strong Coupling Constant

Since jet cross sections depend directly on the strong coupling constant, they can be used to determine the parameter $\alpha_S(M_Z)$ from a comparison to data as reported by the CMS Collaboration considering the inclusive jet and the 3-jet cross sections for this purpose [29, 56]. The influence of other theory parameters like PDFs, notably the less known gluon PDF, has to be taken into account. To delimit unknowns, the choice of PDF sets therefore is restricted to global sets that fit data from different experiments, such that only the most precisely known gluon distributions are employed. In a more complete ansatz, combined fits of $\alpha_S(M_Z)$ and the gluon PDF of the proton can be performed. This is explored to some extent in Section 4.5. Here, the correlation between $\alpha_S(M_Z)$ and PDFs is considered by using series of PDF sets, where each set has been fitted to data for a different assumption on the value of $\alpha_S(M_Z)$. Table 2.1 lists i.a. the available PDF sets typically used in LHC Run 1 analyses together with the preferred value of $\alpha_S(M_Z)$ and the available ranges with additional choices for $\alpha_S(M_Z)$, where the step size from lowest to highest value is 0.001. HERAPDF1.5, which is restricted to HERA data as input alone, is not a global PDF set and therefore excluded from the following.

Figure 4.14 left demonstrates the sensitivity of the inclusive jet cross section to $\alpha_S(M_Z)$ by comparing the ratio of the data to theory at NLO for the series of PDFs from CT10-NLO with assumptions on $\alpha_S(M_Z)$ ranging from 0.112 to 0.128. The line at unity corresponds to the ratio for the central $\alpha_S(M_Z)$ of 0.118 for this PDF set. Within uncertainties the data predominantly follow the prediction for one particular value of $\alpha_S(M_Z)$ with some systematic deviations visible at high jet p_{T}, where the measurements become less accurate. A similar picture emerges for the 3-jet cross section or for other regions in rapidity. Other PDF sets like MSTW2008-NLO or NNPDF2.1-NLO also give satisfactory theoretical descriptions of the data with slight variations in sensitivity and shape. The ABM11-NLO set though leads to QCD predictions significantly different in shape to both examined measurements and therefore is excluded from further consideration. This situation might be im-

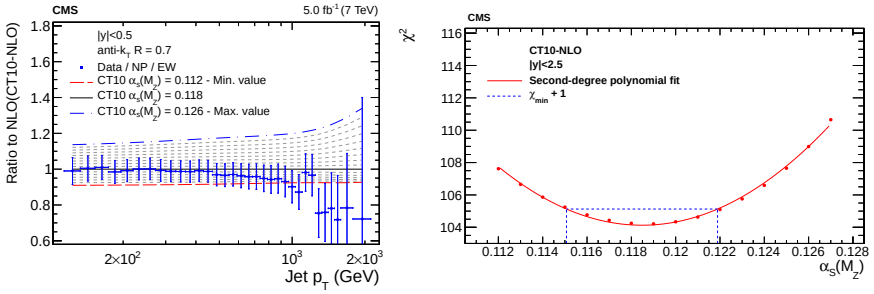

Fig. 4.14 Left: Ratio of the inclusive jet cross section at small rapidity with error bars corresponding to the total experimental uncertainty to predictions using the CT10-NLO PDF set. The value of $\alpha_S(M_Z)$ is varied in the range 0.112–0.126 in steps of 0.001. Right: Example of the χ^2 minimisation with respect to $\alpha_S(M_Z)$ using the CT10-NLO PDF set and data from all rapidity bins. The experimental uncertainty is obtained from the $\alpha_S(M_Z)$ values for which χ^2 is increased by one with respect to the minimum value, indicated by the dashed line. The curve corresponds to a second-degree polynomial fit through the available χ^2 points. *(Taken from Ref. [29])*

proved, once NNLO theory becomes available that could remedy the lower gluon PDF present in ABM11-NLO compared to the other sets.

The fitting procedure to extract the value of $\alpha_S(M_Z)$ minimises the χ^2 between the N measurements D_i and the theoretical predictions T_i. The χ^2 is defined as in Eq. (3.18) to

$$\chi^2 = \sum_{i,j=1}^{N} (D_i - T_i)\, \mathrm{C}_{ij}^{-1}\, (D_j - T_j)\,, \tag{4.17}$$

where the covariance matrix C_{ij} is composed of the following terms:

$$C = \mathrm{cov}_{\mathrm{unf+stat}} + \mathrm{cov}_{\mathrm{uncor}} + \left(\sum_{\mathrm{sources}} \mathrm{cov}_{\mathrm{JEC}} \right) + \mathrm{cov}_{\mathrm{lumi}} + \mathrm{cov}_{\mathrm{PDF}}\,, \tag{4.18}$$

which represent

1. $\mathrm{cov}_{\mathrm{unf+stat}}$: statistical and unfolding uncertainty including correlations induced through unfolding;
2. $\mathrm{cov}_{\mathrm{uncor}}$: uncorrelated systematic uncertainty summing up small residual effects such as trigger and identification inefficiencies, time dependence of the jet p_{T} resolution, or the uncertainty on the trigger pre-scale factor;
3. $\mathrm{cov}_{\mathrm{JEC\,sources}}$: systematic uncertainty for each JEC uncertainty source;
4. $\mathrm{cov}_{\mathrm{lumi}}$: luminosity uncertainty; and
5. $\mathrm{cov}_{\mathrm{PDF}}$: PDF uncertainty.

The first four sources constitute the experimental uncertainty. The JEC and luminosity uncertainty are treated as fully correlated across the whole phase space, where for the JEC uncertainty the procedure recommended in Ref. [29] is applied. To avoid the statistical bias that arises from uncertainty estimations taken from

data [60–62], these fully correlated sources are assumed to be multiplicative. The derivation of PDF uncertainties follows prescriptions for each individual PDF set, cf. Section 2.6.2. As usual, the uncertainty of a result for $\alpha_S(M_Z)$ from a χ^2 fit is obtained from the $\alpha_S(M_Z)$ values for which the χ^2 is increased by one with respect to the minimum value. An example of such a fit is presented in Fig. 4.14, where points in-between the available steps in $\alpha_S(M_Z)$ are interpolated by a second-degree polynomial fit of the χ^2 curve.

The uncertainty in $\alpha_S(M_Z)$ due to the NP uncertainties is evaluated by looking for maximal offsets from a default fit. The theoretical prediction T is varied by the NP uncertainty ΔNP as $T \cdot \mathrm{NP} \to T \cdot (\mathrm{NP} \pm \Delta\mathrm{NP})$. The fitting procedure is repeated for these variations, and the deviation from the central $\alpha_S(M_Z)$ values is considered as the uncertainty in $\alpha_S(M_Z)$.

Finally, the uncertainty due to the variation of the renormalisation and factorisation scales is evaluated by applying the same method as for the NP corrections: μ_r and μ_f are varied from the default choice μ_0 of $\mu_r = \mu_f = p_\mathrm{T}$ (inclusive jets) or $\mu_r = \mu_f = m_3/2$ (3-jet mass) between $\mu_0/2$ and $2\mu_0$ in the following six combinations: $(\mu_r/\mu_0, \mu_f/\mu_0) = (1/2, 1/2)$, $(1/2, 1)$, $(1, 1/2)$, $(1, 2)$, $(2, 1)$, and $(2, 2)$. The two extreme combinations of $(1/2, 2)$ and $(2, 1/2)$ with relative factors of 4 between μ_r and μ_f are not considered following the discussion in Ref. [63]. The χ^2 minimisation with respect to $\alpha_S(M_Z)$ is repeated in each case. The contribution from the μ_r and μ_f scale variations to the uncertainty is evaluated by considering the maximal upwards and downwards deviation of $\alpha_S(M_Z)$ from the central result.

The capability of the FASTNLO framework described in Section 2.6.5 permits to replace the $\alpha_S(M_Z)$ value and the $\alpha_S(Q)$ evolution of a PDF set by alternative choices. Although this technique does not replace a dedicated fit of PDFs for a particular $\alpha_S(M_Z)$ value, because correlations are neglected, it provides a better-founded estimate of a cross section in the vicinity of an available point of $\alpha_S(M_Z)$ than a simple polynomial fit of the χ^2 curve. As long as only interpolations are concerned, differences between the two approaches are found to be negligible. It is observed, however, that offset fits for the evaluation of scale uncertainties might require wider ranges in $\alpha_S(M_Z)$ points than available for a PDF set. Extrapolations based solely on a polynomial fit become doubtful when leaving the region of available $\alpha_S(M_Z)$ points, since they rely on information from inside that region alone. By using FASTNLO, limited extrapolations outside this region at least account for changes in $\alpha_S(M_Z)$ directly and are accepted within the described analyses if necessary for uncertainty evaluations up to a limit of $|\Delta\alpha_S(M_Z)| = 0.003$. Within FASTNLO the $\alpha_S(Q)$ evolution code of the HOPPET toolkit [64] (inclusive jets) or the Glück–Reya–Vogt formula (GRV) [65] (3-jet mass) are used at two-loop order as appropriate for a calculation at NLO. The procedure has been cross-checked using the original $\alpha_S(Q)$ grid of each PDF within LHAPDF and with the evolution code of RUNDEC [66, 67].

Table 4.1 reports the results for $\alpha_S(M_Z)$ from the inclusive jet or the 3-jet production cross sections exploiting the complete measured phase space in rapidity using the CT10-NLO, MSTW2008-NLO, and NNPDF2.1-NLO PDF sets. For fits of $\alpha_S(M_Z)$ from 3-jet production, the very low m_3 region is unsuited because of kine-

Table 4.1 Determination of $\alpha_S(M_Z)$ for the whole measured rapidity range from the inclusive jet cross section (upper three rows) [29] and from the 3-jet mass cross section (lower two rows) [56] using the CT10, MSTW2008, or NNPDF2.1 PDF sets with NLO evolution.

PDF set	$\alpha_S(M_Z)_{\text{def}}$	χ^2/n_{dof}	$\alpha_S(M_Z)$	$\pm$(exp)	$\pm$(NP)	$\pm$(PDF)	$\pm$(scale)
CT10-NLO	0.1180	104.1/132	0.1185	0.0019	0.0004	0.0028	$^{+0.0053}_{-0.0024}$
MSTW2008-NLO	0.1202	107.9/132	0.1159	0.0012	0.0001	0.0014	$^{+0.0024}_{-0.0030}$
NNPDF2.1-NLO	0.1190	103.5/132	0.1150	0.0015	0.0003	0.0024	$^{+0.0025}_{-0.0025}$
CT10-NLO	0.1180	47.2/45	0.1171	0.0013	0.0008	0.0024	$^{+0.0069}_{-0.0040}$
MSTW2008-NLO	0.1202	52.8/45	0.1155	$^{+0.0014}_{-0.0013}$	$^{+0.0008}_{-0.0009}$	$^{+0.0014}_{-0.0015}$	$^{+0.0105}_{-0.0029}$

matic constraints through minimal jet p_{T} requirements. Close to phase space boundaries fixed-order pQCD calculations might be insufficient and resummations might be needed. Therefore, only m_3 bins beyond the maximum of the 3-jet mass cross section in the outer $|y|_{\max}$ bin at around 664 GeV are considered. Also, the range in values available for $\alpha_S(M_Z)$ is not wide enough to reliably estimate the scale uncertainty for the fit of predictions with the NNPDF2.1-NLO PDFs to 3-jet production cross sections. Within experimental uncertainties all results are compatible. Theoretical uncertainties are much larger than experimental ones and are dominated by scale variations emphasising the need for NNLO calculations. The CT10-NLO PDF set is chosen for the main result for two reasons: The range in available $\alpha_S(M_Z)$ values is wide enough to evaluate almost all scale variations within this range, and the value of $\alpha_S(M_Z)$ preferred by the CMS jet data is rather close to the default value of this PDF set making the fitting setup as consistent as possible. The final results then read:

$$\begin{aligned}
\alpha_S(M_Z) &= 0.1185 \pm 0.0019\,(\text{exp}) \pm 0.0004\,(\text{NP}) \pm 0.0028\,(\text{PDF})^{+0.0053}_{-0.0024}\,(\text{scale}) \\
&= 0.1185 \pm 0.0034\,(\text{all except scale})^{+0.0053}_{-0.0024}\,(\text{scale}) \\
&= 0.1185^{+0.0063}_{-0.0042} \qquad (\text{inclusive jets}), \\
\alpha_S(M_Z) &= 0.1171 \pm 0.0013\,(\text{exp}) \pm 0.0008\,(\text{NP}) \pm 0.0024\,(\text{PDF})^{+0.0069}_{-0.0040}\,(\text{scale}) \\
&= 0.1171 \pm 0.0028\,(\text{all except scale})^{+0.0069}_{-0.0040}\,(\text{scale}) \\
&= 0.1171^{+0.0074}_{-0.0049} \qquad (\text{3-jet mass}),
\end{aligned} \tag{4.19}$$

where experimental, NP, PDF, and scale uncertainties have been added quadratically to give the total uncertainty. These results are in agreement with the world average value of $\alpha_S(M_Z) = 0.1185 \pm 0.0006$ [68], with the Tevatron results [69–71], and other results obtained with LHC data [72–74].

Concentrating on the CT10-NLO PDF set, Table 4.2 reports the fit results for each bin in $|y|$ or $|y|_{\max}$ together with the experimental, PDF, NP, and scale uncertainties to investigate potential rapidity dependent biases or incompatibilities. None are found. The fits in separate rapidity regions are consistent with each other and

Table 4.2 Determination of $\alpha_S(M_Z)$ in different regions of jet rapidity from the inclusive jet cross section (upper five rows) [29] and from the 3-jet mass cross section (lower two rows) [56] using the CT10-NLO PDF set [57].

y range	χ^2/n_{dof}	$\alpha_S(M_Z)$	$\pm$(exp)	$\pm$(NP)	$\pm$(PDF)	$\pm$(scale)
$\lvert y\rvert < 0.5$	16.2/32	0.1189	0.0024	0.0008	0.0030	$^{+0.0045}_{-0.0027}$
$0.5 \le \lvert y\rvert < 1.0$	25.4/29	0.1182	0.0024	0.0008	0.0029	$^{+0.0050}_{-0.0025}$
$1.0 \le \lvert y\rvert < 1.5$	9.5/26	0.1165	0.0027	0.0008	0.0024	$^{+0.0043}_{-0.0020}$
$1.5 \le \lvert y\rvert < 2.0$	20.2/23	0.1146	0.0035	0.0013	0.0031	$^{+0.0037}_{-0.0020}$
$2.0 \le \lvert y\rvert < 2.5$	12.6/18	0.1161	0.0045	0.0015	0.0054	$^{+0.0034}_{-0.0032}$
$\lvert y\rvert_{\max} < 1.0$	10.3/22	0.1163	$^{+0.0018}_{-0.0019}$	0.0007	0.0027	$^{+0.0059}_{-0.0025}$
$1.0 \le \lvert y\rvert_{\max} < 2.0$	10.6/22	0.1179	$^{+0.0018}_{-0.0019}$	0.0007	0.0021	$^{+0.0067}_{-0.0037}$

with the outcome of the combined fit, which is nontrivial, since in the combination correlated systematic uncertainties have to be considered. Within the scope of the discussed CMS analyses the original prescription for the JEC uncertainty [75] was found to overestimate the correlation in rapidity and consequently was improved as described in [29].

Finally, to investigate the running of the strong coupling, the fitted region is split into six (seven) bins of p_{T} ($m_3/2$) and the fitting procedure is repeated in each of these bins. The extractions of $\alpha_S(M_Z)$ are reported in Table 4.3, where the same two-loop solutions to the RGE are used, HOPPET and GRV, to evolve the $\alpha_S(M_Z)$ values to the corresponding energy scale Q. The value of Q is calculated as a cross section weighted average in each fit region. These average scale values $\langle Q \rangle$, derived again with the FASTNLO framework, are identical within some GeV for different PDFs. To demonstrate the running of the strong coupling constant, the corresponding $\alpha_S(Q)$ values are listed for each respective $\langle Q \rangle$ scale in the last two columns of Table 4.3. All eleven results are presented in Fig. 4.15 together with the running of the strong coupling $\alpha_S(Q)$ and its total uncertainty as determined in the analysis of the inclusive jets. In the same figure the values of α_S at lower scales determined by the H1 [76–78], ZEUS [79], and D0 [70, 71] collaborations are shown for comparison. Other CMS measurements [73, 74], which are in agreement with the $\alpha_S(M_Z)$ determination of the discussed studies, are displayed as well. The reported results on α_S are consistent with the energy dependence predicted by the RGE and extend the investigated scale values Q beyond the Tevatron limit of 600 GeV into the TeV region.

4.5 Constraining the Proton Structure

The PDFs of the proton are an essential ingredient for precision studies in hadron-induced reactions. They are derived from experimental data involving collider and fixed-target experiments. DIS data from the HERA ep collider experiments cover

Table 4.3 Determination of $\alpha_S(M_Z)$ for different ranges of scale Q from the inclusive jet cross section (upper six rows) [29] and from the 3-jet mass cross section (lower seven rows) [56] using the CT10-NLO PDF set [57]. The last two columns list the corresponding values of the average scales $\langle Q \rangle$ and the strong coupling constant at this scale.

Q range	χ^2/n_{dof}	$\alpha_S(M_Z)$	$\pm$(exp)	$\pm$(NP)	$\pm$(PDF)	$\pm$(scale)	$\langle Q \rangle$/ GeV	$\alpha_S(Q)$
114–196	6.2/19	0.1172	0.0031	0.0007	0.0018	$^{+0.0045}_{-0.0022}$	136	0.1106
196–300	7.6/19	0.1180	0.0034	0.0011	0.0019	$^{+0.0048}_{-0.0025}$	226	0.1038
300–468	8.1/24	0.1194	0.0032	0.0010	0.0023	$^{+0.0049}_{-0.0027}$	345	0.0993
468–638	10.6/19	0.1187	0.0029	0.0006	0.0031	$^{+0.0052}_{-0.0027}$	521	0.0940
638–905	11.2/21	0.1192	0.0034	0.0005	0.0032	$^{+0.0057}_{-0.0030}$	711	0.0909
905–2116	33.6/25	0.1176	0.0047	0.0002	0.0040	$^{+0.0092}_{-0.0020}$	1007	0.0866
664–794	4.5/3	0.1232	$^{+0.0040}_{-0.0042}$	$^{+0.0008}_{-0.0007}$	$^{+0.0019}_{-0.0016}$	$^{+0.0079}_{-0.0044}$	361	0.1013
794–938	7.8/3	0.1143	$^{+0.0019}_{-0.0016}$	0.0008	$^{+0.0034}_{-0.0033}$	$^{+0.0073}_{-0.0042}$	429	0.0933
938–1098	0.6/3	0.1171	$^{+0.0033}_{-0.0034}$	0.0007	0.0022	$^{+0.0068}_{-0.0040}$	504	0.0934
1098–1369	2.6/5	0.1152	0.0026	$^{+0.0008}_{-0.0007}$	$^{+0.0027}_{-0.0026}$	$^{+0.0060}_{-0.0027}$	602	0.0902
1369–2172	8.8/13	0.1168	$^{+0.0018}_{-0.0019}$	$^{+0.0007}_{-0.0006}$	$^{+0.0030}_{-0.0031}$	$^{+0.0068}_{-0.0034}$	785	0.0885
2172–2602	3.6/5	0.1167	$^{+0.0037}_{-0.0044}$	0.0008	$^{+0.0040}_{-0.0044}$	$^{+0.0065}_{-0.0041}$	1164	0.0848
2602–3270	5.5/7	0.1120	$^{+0.0043}_{-0.0041}$	0.0001	$^{+0.0056}_{-0.0040}$	$^{+0.0088}_{-0.0050}$	1402	0.0807

most of the kinematic phase space needed for a reliable PDF extraction. The sensitivity to the gluon PDF, however, is only indirect via scaling violations. In particular in the region of medium to high fractions $x > 0.5 \cdot 10^{-2}$ of the proton momentum, jet cross sections from ep or pp collisions can additionally constrain the gluon PDF and also the valence quark PDFs at very high $x > 0.1$.

The potential impact of LHC jet data can be illustrated by the correlation between the inclusive jet cross section $\sigma_{\mathrm{jet}}(Q)$ and the PDF $xf(x,Q^2)$ for any parton flavour f. The NNPDF Collaboration [80] provides PDF sets in the form of an ensemble of replicas i, which sample variations in the PDF parameter space within allowed uncertainties. The correlation coefficient $\rho_f(x,Q)$ between a cross section and the PDF for flavour f at a point (x,Q) can be computed by evaluating means and standard deviations from an ensemble of N replicas as

$$\rho_f(x,Q) = \frac{N}{(N-1)} \frac{\langle \sigma_{\mathrm{jet}}(Q)_i \cdot xf(x,Q^2)_i \rangle - \langle \sigma_{\mathrm{jet}}(Q)_i \rangle \cdot \langle xf(x,Q^2)_i \rangle}{\Delta_{\sigma_{\mathrm{jet}}(Q)} \Delta_{xf(x,Q^2)}}. \qquad (4.20)$$

Here, the angular brackets denote the averaging over the replica index i, and Δ represents the evaluation of the corresponding standard deviation for either the jet cross section, $\Delta_{\sigma_{\mathrm{jet}}(Q)}$, or a PDF, $\Delta_{xf(x,Q^2)}$. Figure 4.16 presents the correlation coefficient between the inclusive jet cross section and the gluon respectively the u valence quark PDFs in the proton.

The correlation between the gluon PDF and the inclusive jet cross section at small rapidity is large for most jet p_{T}. In contrast, the correlation between the valence quark distributions and the jet cross section is rather small except for very high p_{T}

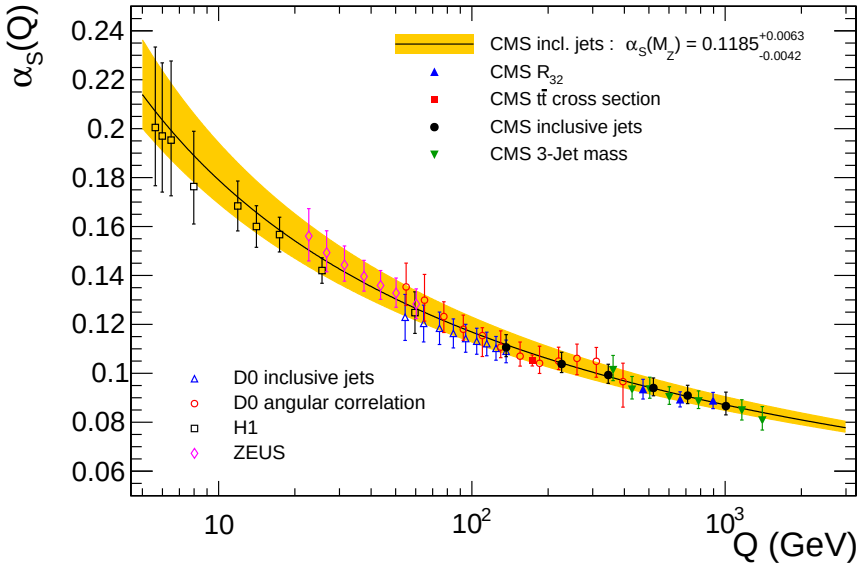

Fig. 4.15 The strong coupling $\alpha_S(Q)$ (full line) and its total uncertainty (band) as a function of the respective choice for the scale Q using a two-loop solution to the RGE. The extractions of $\alpha_S(Q)$ from the inclusive jet and 3-jet mass cross sections in eleven ranges of Q as presented in Table 4.3 are shown together with results from the H1 [77, 78], ZEUS [79], and D0 [70, 71] experiments at the HERA and Tevatron colliders. Other CMS measurements [73, 74] are displayed as well. The uncertainties represented by error bars are subject to correlations. *(Taken from auxiliary material of Ref. [56])*

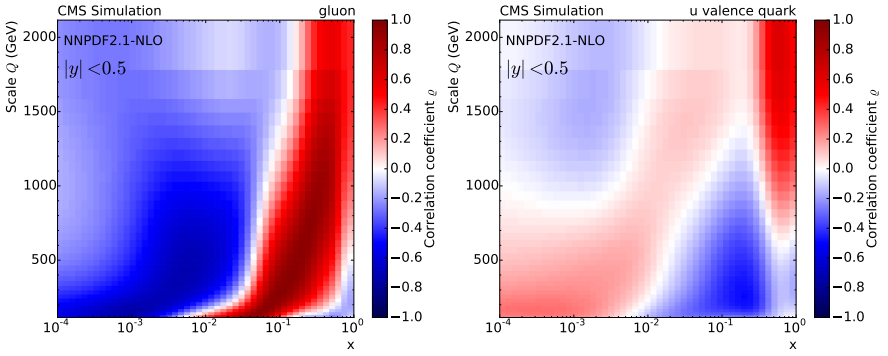

Fig. 4.16 The correlation coefficient between the inclusive jet cross section and the gluon (left), and the u valence quark PDF (right) as a function of the momentum fraction x of the proton and the energy scale Q of the hard process. The correlation is shown for the small rapidity region $|y| < 0.5$. *(Taken from Ref. [29])*

such that some impact can be expected at high x from including these jet data in PDF fits. In the forward region (not shown) the correlation between the valence quark distributions and the jet cross sections is more pronounced at high x and smaller jet p_{T}. In the case of same-side jets with respect to rapidity, the parton fractional momentum of the second proton is rather small because of Eq. (4.10) and lies in a region, where valence quark PDFs are already well known from other processes e.g. DIS. The situation is similar for 3-jet production. Therefore, a significant reduction of PDF uncertainties at medium to high x is expected by including LHC jet data into fits of the proton structure.

These fits can be performed with a software framework such as the open-source HERAFITTER project [81,82], which among other things is designed to fit PDFs to data. It has a modular structure, encompassing a variety of theoretical predictions for different processes and phenomenological approaches for determining the parameters of the PDFs. In the CMS study [29], HERAFITTER version 1.1.1 is employed to estimate the impact of the CMS inclusive jet data on the PDFs and their uncertainties, when combined with DIS data from HERA-I. For both processes the theory is available at NLO, i.e. up to order α_S^2 for DIS and up to order α_S^3 for inclusive jet production in pp collisions.

The impact of the CMS inclusive jet data on proton PDFs is investigated by including the jet cross section measurement in a combined fit at NLO with the HERA-I inclusive DIS cross sections [83], which were the basis for the determination of the HERAPDF1.0 PDF set. The analysis is performed assuming the DGLAP evolution scheme at NLO as implemented in the QCDNUM package [84] and the generalised-mass variable-flavour number Thorne–Roberts scheme [85,86] to treat the massive quark flavours. In contrast to the original HERAPDF fit, the presented results require the DIS data to fulfil $Q^2 > Q^2_{\mathrm{min}} = 7.5\,\mathrm{GeV}^2$ instead of $3.5\,\mathrm{GeV}^2$. The amount of DIS data left out by the increased Q^2_{min} threshold is rather small and concerns a phase space where a perturbative description is less reliable. Crosschecks performed for a cutoff of $Q^2 > Q^2_{\mathrm{min}} = 3.5\,\mathrm{GeV}^2$ did not exhibit any differences beyond an expected reduction of uncertainties at low x.

The following PDFs are parameterised independently in the fit procedure: $xu_v(x)$, $xd_v(x)$, $xg(x)$, and $x\overline{U}(x)$, $x\overline{D}(x)$, where $x\overline{U}(x) = x\overline{u}(x)$, and $x\overline{D}(x) = x\overline{d}(x) + x\overline{s}(x)$. Similar to Ref. [87], a parameterisation with 13 free parameters is used. At the starting scale Q_0 of the QCD evolution, chosen to be $Q_0^2 = 1.9\,\mathrm{GeV}^2$, the PDFs are parameterised as follows:

$$
\begin{aligned}
xg(x) &= A_g x^{B_g}(1-x)^{C_g} - A'_g x^{B'_g}(1-x)^{C'_g},\\
xu_v(x) &= A_{u_v} x^{B_{u_v}}(1-x)^{C_{u_v}}(1+E_{u_v}x^2),\\
xd_v(x) &= A_{d_v} x^{B_{d_v}}(1-x)^{C_{d_v}},\\
x\overline{U}(x) &= A_{\overline{U}} x^{B_{\overline{U}}}(1-x)^{C_{\overline{U}}}, \qquad \text{and}\\
x\overline{D}(x) &= A_{\overline{D}} x^{B_{\overline{D}}}(1-x)^{C_{\overline{D}}}.
\end{aligned}
\tag{4.21}
$$

The normalisation parameters A_g, A_{u_v}, and A_{d_v} are constrained by QCD sum rules. Additional constraints $B_{\overline{U}} = B_{\overline{D}}$ and $A_{\overline{U}} = A_{\overline{D}}(1-f_s)$ are applied to ensure the same

normalisation for the $\overline{u}$ and $\overline{d}$ densities for $x \to 0$. The strangeness fraction is set to $f_s = 0.31$, as obtained from neutrino-induced dimuon production [88]. The parameter C_g' is fixed to 25 [14, 86] and the strong coupling constant to $\alpha_S(M_Z) = 0.1176$.

The agreement between the N data points D_i and the theoretical predictions T_i is quantified via a least-squares method, where

$$\chi^2 = \sum_{i,j=1}^{N} \left(D_i - T_i - \sum_{k=1}^{K} \varepsilon_k \beta_{ik} \right) \mathrm{C}_{ij}^{-1} \left(D_j - T_j - \sum_{k=1}^{K} \varepsilon_k \beta_{jk} \right) + \sum_{k=1}^{K} \varepsilon_k^2 . \qquad (4.22)$$

For unconstrained fits this definition is equivalent to Eq. (4.17) [89, 90]. As a bonus, the systematic shift of the nuisance parameter ε_k for each source in a fit is determined. The *a priori* expectation at the χ^2 minimum is that systematic shifts via the nuisance parameters ε_k are Gaussian distributed with zero mean and unity root-mean-square. Numerous large shifts in either direction indicate a problem as for example observed while fitting $\alpha_S(M_Z)$ with this technique and the uncertainty correlation prescription from Ref. [75].

In the following, the component $\mathrm{cov}_{\mathrm{unf+stat}}$ of the covariance matrix, cf. Eq. (4.18), is separated into an uncorrelated statistical part $\mathrm{cov}_{\mathrm{stat}}$ and a correlated part from unfolding $\mathrm{cov}_{\mathrm{unf}}$. The statistical part together with the uncorrelated systematic component $\mathrm{cov}_{\mathrm{uncor}}$ give the new covariance matrix needed in Eq. (4.22), while the JEC, unfolding, and luminosity determination are treated as fully correlated systematic uncertainties β_{ik} with nuisance parameters ε_k. In comparison to Eq. (3.20) for the JEC uncertainties, the χ^2 here is written with a subtractive term $\sum_{k=1}^{K} \varepsilon_k \beta_{ik}$, where the β_{ik} correspond to absolute systematic shifts of $1\,\sigma$. To rewrite this in a multiplicative form as appropriate for JEC uncertainties, the β_{ik} can be set equal to $T_i \cdot s_{ki}$. Including also the NP uncertainties, treated via the offset method described in Section 4.4, in the form of one nuisance parameter in total K such sources are defined. Of course, PDF uncertainties emerge as results of the fits performed here, in contrast to serving as inputs, as they do in the fits of $\alpha_S(M_Z)$ presented in Section 4.4.

Similarly as for the $\alpha_S(M_Z)$ fit following Eq. (4.18), the fully correlated sources are assumed to be multiplicative. As a consequence, the covariance matrix of the remaining sources has to be re-evaluated in each iteration step. To prevent the compensation of large systematic shifts caused by simultaneously increasing the theoretical prediction and the statistical uncertainties, the systematic shifts of the theory are taken into account before the rescaling of the statistical uncertainty. Otherwise alternative minima in χ^2 can appear that are associated with exaggerated theoretical predictions and corresponding shifts in the nuisance parameters, where the penalty in χ^2 is compensated by uncorrelated uncertainties that are scaled-up with the theory. These alternative minima are clearly undesirable [81].

By employing nuisance parameters, the impact of each systematic source of uncertainty on the fit result can be examined separately. For an adequate estimation of the sizes and the correlations of all uncertainties, the majority of all systematic sources should be shifted by less than one standard deviation from the default in the fitting procedure. Table 4.4 demonstrates that this is the case for the CMS inclusive jet data.

Table 4.4 The 23 independent sources of systematic uncertainty considered in the CMS inclusive jet fits. Out of these, 20 are related to the JEC and are listed first. The shift from the default value in each source of systematic uncertainty is determined by nuisance parameters in the fit and is presented in units of standard deviations. *(Taken from Ref. [29])*

Systematic source		Shift in standard deviations
JEC0	absolute jet energy scale	0.09
JEC1	MC extrapolation	0.00
JEC2a	single-particle response barrel	1.31
JEC2b	single-particle response endcap	−1.46
JEC2c	single-particle decorrelation $\|y\| < 0.5$	0.20
JEC2d	single-particle decorrelation $0.5 \leq \|y\| < 1.0$	0.19
JEC2e	single-particle decorrelation $1.0 \leq \|y\| < 1.5$	0.92
JEC3	jet flavor correction	0.04
JEC4	time-dependent detector effects	−0.15
JEC5	jet p_{T} resolution in endcap 1	0.76
JEC6	jet p_{T} resolution in endcap 2	−0.42
JEC7	jet p_{T} resolution in HF	0.01
JEC8	correction for final-state radiation	0.03
JEC9	statistical uncertainty of η-dependent correction for endcap	−0.42
JEC10	statistical uncertainty of η-dependent correction for HF	0.00
JEC11	data-MC difference in η-dependent PU correction	0.91
JEC12	residual out-of-time PU correction for pre-scaled triggers	−0.17
JEC13	offset dependence in PU correction	−0.03
JEC14	MC PU bias correction	0.39
JEC15	jet rate dependent PU correction	0.29
Unfolding		−0.26
Luminosity		−0.07
NP correction		0.60

Following the HERAPDF prescription [83], the uncertainty in the PDFs is subdivided into experimental, model, and parameterisation uncertainties that are studied separately. In the default setup of the HERAFITTER framework, experimental uncertainties are evaluated following a Hessian method [90] and result from the propagated statistical and systematic uncertainties of the input data.

For the model uncertainties, the offset method [91] is applied considering the following variations of model assumptions:

1. The strangeness fraction f_s, by default equal to 0.31, is varied between 0.23 and 0.38.
2. The b-quark mass is varied by $\pm 0.25\,\mathrm{GeV}$ around the central value of $4.75\,\mathrm{GeV}$.

3. The c-quark mass, with the central value of 1.4 GeV, is varied between 1.35 GeV and 1.65 GeV. For the downwards variation the charm production threshold is avoided by changing the starting scale to $Q_0^2 = 1.8\,\mathrm{GeV}^2$ in this case.
4. The minimum Q^2 value for data used in the fit, $Q^2_{\mathrm{min}} = 7.5\,\mathrm{GeV}^2$, is varied between $5.0\,\mathrm{GeV}^2$ and $10\,\mathrm{GeV}^2$.

The PDF parameterisation uncertainty is estimated as described in Ref. [83]. By employing the more general form of parameterisation

$$\begin{aligned} xg(x) &= A_g x^{B_g}(1-x)^{C_g}(1+D_g x+E_g x^2) - A'_g x^{B'_g}(1-x)^{C'_g}, \\ xf(x) &= A_f x^{B_f}(1-x)^{C_f}(1+D_f x+E_f x^2) \end{aligned} \tag{4.23}$$

for gluons and the other flavours, respectively, it is tested whether the successive inclusion of additional fit parameters leads to a variation in the shape of the fitted results. Furthermore, the starting scale Q_0 is changed to $Q_0^2 = 1.5\,\mathrm{GeV}^2$ and $2.5\,\mathrm{GeV}^2$. The maximal deviations of the resulting PDFs from those obtained in the central fit define the parameterisation uncertainty. The experimental, model, and parameterisation uncertainties are added in quadrature to give the final PDF uncertainty.

Using this fitting setup, the partial χ^2 values per number of data points, n_{data}, are reported in Table 4.5 for each of the neutral current (NC) and charged current (CC) data sets in the HERA-I DIS fit and for the combined fit including the CMS inclusive jet data. The achieved fit qualities demonstrate the compatibility of all data within the presented PDF fitting methods.

Figure 4.17 upper row provides direct comparisons of the two fit results with total uncertainties for the gluon and u valence quark distributions. The uncertainty of the gluon distribution is significantly reduced for almost the whole x range from 10^{-4} up to 0.5. When DIS data below $Q^2_{\mathrm{min}} = 7.5\,\mathrm{GeV}^2$ are included in the fit, the huge effect visible in the low x region $x < 0.01$ is much reduced. Also, for the u valence distribution some reduction in its uncertainty is observed, also at high x ($x \gtrsim 0.1$). At the same time, some wiggly structure can be seen that can be traced back to the parameterisation uncertainties. This might point to a still insufficient flexibility in the parameterisations. Therefore, a comparison is presented using the MC method with the regularisation based on data, which is also implemented within the HERAFITTER framework.

To study more flexible PDF parameterisations, a MC method based on varying the input data within their correlated uncertainties is employed in combination with a data-based regularisation technique. This method was first used by the NNPDF Collaboration and uses a more flexible parameterisation to describe the x dependence of the PDFs [80]. To avoid the fitting of statistical fluctuations present in the input data (over-fitting) a data-based stopping criterion is introduced. The data set is split randomly into a "fit" and a "control" sample. The χ^2 minimisation is performed with the "fit" sample while simultaneously the χ^2 of the "control" sample is calculated using the current PDF parameters. It is observed that the χ^2 of the "control" sample at first decreases and then starts to increase again because of over-fitting. At this point, the fit is stopped. This regularisation technique is used in combination

Table 4.5 Partial χ^2 values, χ^2_p, for each data set in the HERA-I DIS (middle section) or in the combined fit including CMS inclusive jet data (right section). Here, n_{data} is the number of data points available for the determination of the 13 parameters. The bottom two lines show the total χ^2 and χ^2/n_{dof}. The difference between the sum of all χ^2_p and the total χ^2 for the combined fit is attributed to the nuisance parameters. *(Adapted from Ref. [29])*

		HERA-I data		HERA-I & CMS data	
data set	n_{data}	χ^2_p	χ^2_p/n_{data}	χ^2_p	χ^2_p/n_{data}
NC HERA-I H1-ZEUS combined e^-p	145	109	0.75	109	0.75
NC HERA-I H1-ZEUS combined e^+p	337	309	0.91	311	0.92
CC HERA-I H1-ZEUS combined e^-p	34	20	0.59	22	0.65
CC HERA-I H1-ZEUS combined ep	34	29	0.85	35	1.03
CMS inclusive jets	133	—	—	102	0.77
data set(s)	n_{dof}	χ^2	χ^2/n_{dof}	χ^2	χ^2/n_{dof}
HERA-I data	537	468	0.87	—	—
HERA-I & CMS data	670	—	—	591	0.88

with a MC method to estimate the central value and the uncertainties of the fitted PDFs. Before a fit, several hundred replica sets are created by allowing the central values of the measured cross section to fluctuate within their statistical and systematic uncertainties while taking into account all correlations. For each replica, a fit to NLO QCD is performed, which yields an optimum value and uncertainty for each parameter. The collection of all replica fits can then provide an ensemble average and root-mean-square. Moreover, the variations to derive the model dependence of the HERAPDF prescription do not lead to any significant increase of the uncertainty.

For direct comparison, the fit results for the gluon and u valence quark distributions with the MC method are shown in the bottom row of Fig. 4.17. The total uncertainty derived with the MC method is almost always larger than with the HERAPDF technique. In the case of the gluon at low x, it is much larger. In both cases a significant reduction of the uncertainty in the gluon PDF is observed, notably in the x range from 10^{-2} up to 0.5. Both methods also lead to a decrease in the gluon PDF between 10^{-2} and 10^{-1} and an increase for larger x. Although this change is more pronounced when applying the MC method, within the respective uncertainties both results are compatible. Both methods equally agree on a very modest reduction in uncertainty at high $x > 0.05$ in the u valence quark PDF, which is expected from the correlations, studied in Fig. 4.16, where the quark distributions are constrained via the qq contribution to jet production at high $|y|$ and p_T. Finally, Fig. 4.18 shows an overview of the gluon, sea, u valence, and d valence distributions at the starting scale of $Q^2 = 1.9\,\text{GeV}^2$ for the regularised MC method.

Inclusive DIS data alone are not sufficient to disentangle effects on cross section predictions from changes in the gluon distribution or $\alpha_S(M_Z)$ simultaneously. Therefore $\alpha_S(M_Z)$ was always fixed to 0.1176 in the original HERAPDF1.0 derivation. When the CMS inclusive jet data are added, this constraint can be dropped and

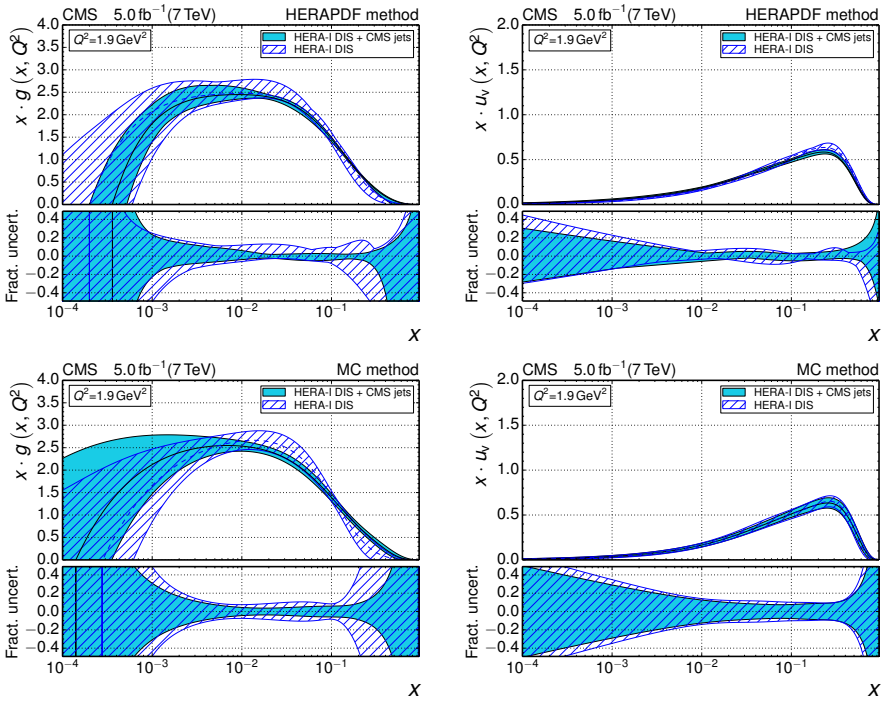

Fig. 4.17 The gluon (left) and the u valence quark PDF (right) as a function of x as derived from HERA-I inclusive DIS data alone (dashed line) and in combination with CMS inclusive jet data (full line). The PDFs are determined employing the HERAPDF method (upper row) and the MC method with data-derived regularisation (lower row). A minimum Q^2 of $Q^2_{\text{min}} = 7.5\,\text{GeV}^2$ was imposed on the DIS data. The PDFs are shown at the starting scale $Q^2 = 1.9\,\text{GeV}^2$. Only the total uncertainty in the PDFs is shown (hatched and solid bands). *(Taken from Ref. [29])*

$\alpha_S(M_Z)$ and its uncertainty (without Q scale variations) is determined to $\alpha_S(M_Z) = 0.1192\,^{+0.0023}_{-0.0019}$ (all except scale). Repeating the fit with the regularised MC method gives $\alpha_S(M_Z) = 0.1188 \pm 0.0041$ (all except scale).

Since a direct correspondence among the different components of the uncertainty can not easily be established, only the quadratic sum of experimental, PDF, and NP uncertainties are presented, which is equivalent to the total uncertainty without scale uncertainty. For example, the HERA-I DIS data contribute to the experimental uncertainty in the combined fits, but contribute only to the PDF uncertainty in separate $\alpha_S(M_Z)$ fits. The HERAPDF prescription for PDF fits tends to small uncertainties, while the uncertainties of the MC method with data-derived regularisation are twice as large. For comparison, the corresponding uncertainty in $\alpha_S(M_Z)$ using more precisely determined PDFs from global fits as in Section 4.4 gives a result between the two: $\alpha_S(M_Z) = 0.1185 \pm 0.0034$ (all except scale).

The evaluation of scale uncertainties is an open issue, which is ignored in all global PDF fits given in Table 2.1. The impact is investigated in Refs. [14, 92–94],

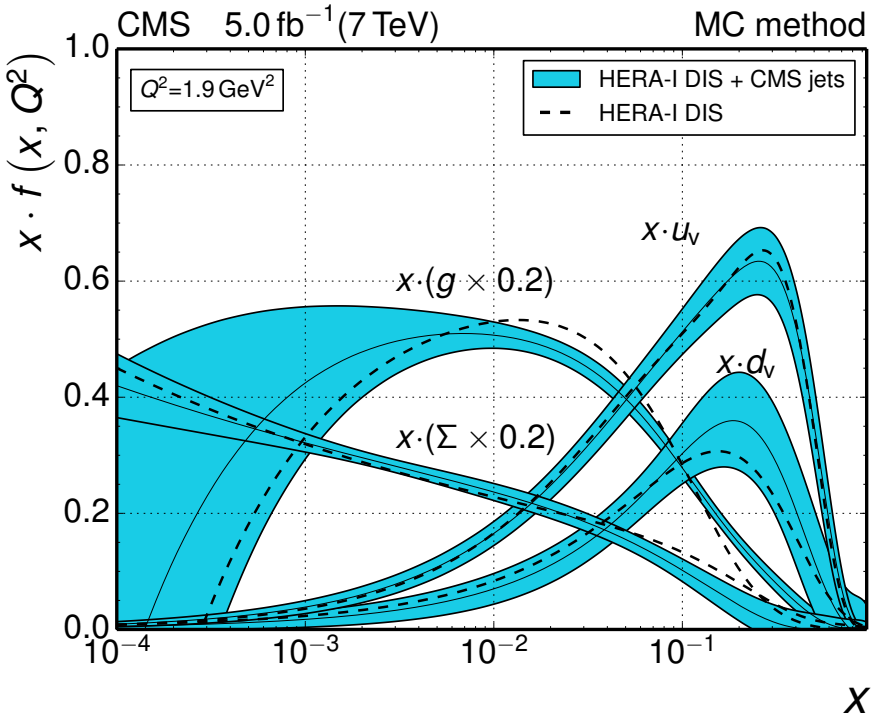

Fig. 4.18 Overview of the gluon, sea, u valence, and d valence PDFs before (dashed line) and after (full line) including the CMS inclusive jet data into the fit. The plots show the PDF fit outcome from the MC method with data-derived regularisation. The PDFs are shown at the starting scale $Q^2 = 1.9\,\mathrm{GeV}^2$. The total uncertainty including the CMS inclusive jet data is shown as a band around the central fit result. *(Taken from Ref. [29])*

where scale definitions and K-factors are varied. Lacking a recommended procedure for the scale uncertainties in combined fits of PDFs and $\alpha_S(M_Z)$, two evaluations are reported for the HERAPDF method. In the first one, the combined fit of PDFs and $\alpha_S(M_Z)$ is repeated for each variation of the scale factors from the default choice of $\mu_r = \mu_f = p_\mathrm{T}$ for the same six combinations as explained in Section 4.4. The scale for the HERA DIS data is not changed. The maximal observed upward and downward changes of $\alpha_S(M_Z)$ with respect to the default scale factors are then taken as scale uncertainty, irrespective of changes in the PDFs: $\Delta\alpha_S(M_Z) = {}^{+0.0022}_{-0.0009}\,(\mathrm{scale})$.

The second procedure is analogous to the method employed to determine $\alpha_S(M_Z)$ in Section 4.4. The PDFs are derived for a series of fixed values of $\alpha_S(M_Z)$ as done for the global PDF sets. Using this series of PDFs with varying values of $\alpha_S(M_Z)$, the combination of PDF and $\alpha_S(M_Z)$ that best fits the HERA-I DIS and CMS inclusive jet data is found. The $\alpha_S(M_Z)$ values determined both ways are consistent with each other. The fits are now repeated for the same scale factor variations, and the maximal

observed upward and downward changes of $\alpha_S(M_Z)$ with respect to the default scale factors are taken as scale uncertainty: $\Delta\alpha_S(M_Z) = {}^{+0.0024}_{-0.0039}\,(\text{scale})$.

In contrast to the scale uncertainty of the first procedure, there is less freedom for compensating effects between different gluon distributions and $\alpha_S(M_Z)$ values in the second procedure. The latter procedure leads to a larger scale uncertainty as expected. In overall size the uncertainty is similar to the final results on $\alpha_S(M_Z)$ reported in the section dedicated to the $\alpha_S(M_Z)$ fit: $\Delta\alpha_S(M_Z) = {}^{+0.0053}_{-0.0024}\,(\text{scale})$. In summary, a combined fit of PDFs and the strong coupling constant becomes possible, when combining HERA DIS and CMS inclusive jet data, and gives compatible results to the separate fit of $\alpha_S(M_Z)$ described in the previous section.

References

1. UA2 Collaboration, "Observation of Very Large Transverse Momentum Jets at the CERN $\bar{p}p$ Collider", *Phys. Lett. B* **118** (1982) 203, doi:10.1016/0370-2693(82)90629-3.
2. R. Horgan and M. Jacob, "Jet Production at Collider Energy", *Nucl. Phys. B* **179** (1981) 441, doi:10.1016/0550-3213(81)90013-4.
3. UA1 Collaboration, "Hadronic Jet Production at the CERN Proton-Antiproton Collider", *Phys. Lett. B* **132** (1983) 214, doi:10.1016/0370-2693(83)90254-X.
4. UA2 Collaboration, "Measurement of the $\sqrt{(s)}$ Dependence of Jet Production at the CERN anti-p p Collider", *Phys. Lett. B* **160** (1985) 349, doi:10.1016/0370-2693(85)91341-3.
5. UA1 Collaboration, "Measurement of the Inclusive Jet Cross-Section at the CERN p anti-p Collider", *Phys. Lett. B* **172** (1986) 461, doi:10.1016/0370-2693(86)90290-X.
6. CDF Collaboration, "Comparison of jet production in $\bar{p}p$ collisions at $\sqrt{s} = 546$ GeV and 1800 GeV", *Phys. Rev. Lett.* **70** (1993) 1376, doi:10.1103/PhysRevLett.70.1376.
7. D0 Collaboration, "High-p_T jets in $\bar{p}p$ collisions at $\sqrt{s} = 630$ GeV and 1800 GeV", *Phys. Rev. D* **64** (2001) 032003, doi:10.1103/PhysRevD.64.032003, arXiv:hep-ex/0012046.
8. D0 Collaboration, "Inclusive jet production in $p\bar{p}$ collisions", *Phys. Rev. Lett.* **86** (2001) 1707, doi:10.1103/PhysRevLett.86.1707, arXiv:hep-ex/0011036.
9. CDF Collaboration, "Measurement of the inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. D* **64** (2001) 032001, doi:10.1103/PhysRevD.65.039903,10.1103/PhysRevD.64.032001, arXiv:hep-ph/0102074.
10. CDF Collaboration, "Measurement of the Inclusive Jet Cross Section using the k_T algorithm in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96 TeV with the CDF II Detector", *Phys. Rev. D* **75** (2007) 092006, doi:10.1103/PhysRevD.75.119901,10.1103/PhysRevD.75.092006, arXiv:hep-ex/0701051.
11. CDF Collaboration, "Measurement of the Inclusive Jet Cross Section at the Fermilab Tevatron $p\bar{p}$ Collider Using a Cone-Based Jet Algorithm", *Phys. Rev. D* **78** (2008) 052006, doi:10.1103/PhysRevD.79.119902,10.1103/PhysRevD.78.052006, arXiv:0807.2204.
12. D0 Collaboration, "Measurement of the inclusive jet cross-section in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV", *Phys. Rev. Lett.* **101** (2008) 062001, doi:10.1103/PhysRevLett.101.062001, arXiv:0802.2400.
13. STAR Collaboration, "Longitudinal double-spin asymmetry and cross section for inclusive jet production in polarized proton collisions at $\sqrt{s} = 200$ GeV", *Phys. Rev. Lett.* **97** (2006) 252001, doi:10.1103/PhysRevLett.97.252001, arXiv:hep-ex/0608030.

14. A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC", *Eur. Phys. J. C* **63** (2009) 189, doi:10.1140/epjc/s10052-009-1072-5, arXiv:0901.0002.
15. fastNLO Collaboration, "Theory-Data Comparisons for Jet Measurements in Hadron-Induced Processes", arXiv:1109.1310.
16. ALICE Collaboration, "Measurement of the inclusive differential jet cross section in pp collisions at $\sqrt{s} = 2.76$ TeV", *Phys. Lett. B* **722** (2013) 262, doi:10.1016/j.physletb.2013.04.026, arXiv:1301.3475.
17. ATLAS Collaboration, "Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector", *Eur. Phys. J. C* **71** (2011) 1512, doi:10.1140/epjc/s10052-010-1512-2, arXiv:1009.5908.
18. ATLAS Collaboration, "Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector", *Phys. Rev. D* **86** (2012) 014022, doi:10.1103/PhysRevD.86.014022, arXiv:1112.6297.
19. ATLAS Collaboration, "Measurement of the inclusive jet cross section in pp collisions at $\sqrt{s} = 2.76$ TeV and comparison to the inclusive jet cross section at $\sqrt{s} = 7$ TeV using the ATLAS detector", *Eur. Phys. J. C* **73** (2013) 2509, doi:10.1140/epjc/s10052-013-2509-4, arXiv:1304.4739.
20. ATLAS Collaboration, "Measurement of the inclusive jet cross-section in proton-proton collisions at $\sqrt{s} = 7$ TeV using 4.5 fb^{-1} of data with the ATLAS detector", *JHEP* **02** (2015) 153, doi:10.1007/JHEP02(2015)153, arXiv:1410.8857.
21. CMS Collaboration, "Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **107** (2011) 132001, doi:10.1103/PhysRevLett.107.132001, arXiv:1106.0208.
22. CMS Collaboration, "Measurement of the inclusive production cross sections for forward jets and for dijet events with one forward and one central jet in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **06** (2012) 036, doi:10.1007/JHEP06(2012)036, arXiv:1202.0704.
23. CMS Collaboration, "Measurements of differential jet cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the CMS detector", *Phys. Rev. D* **87** (2013) 112002, doi:10.1103/PhysRevD.87.112002, arXiv:1212.6660.
24. CMS Collaboration, "Measurement of the ratio of inclusive jet cross sections using the anti-k_T algorithm with radius parameters $R = 0.5$ and 0.7 in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. D* **90** (2014) 072006, doi:10.1103/PhysRevD.90.072006, arXiv:1406.0324.
25. CMS Collaboration, "Measurement of the inclusive jet cross section in pp collisions at $\sqrt{s} = 2.76$ TeV", *Eur. Phys. J. C* **76** (2016) 265, doi:10.1140/epjc/s10052-016-4083-z, arXiv:1512.06212.
26. M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
27. M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.
28. CMS Collaboration, "Study of the underlying event at forward rapidity in pp collisions at $\sqrt{s} = 0.9$, 2.76, and 7 TeV", *JHEP* **04** (2013) 072, doi:10.1007/JHEP04(2013)072, arXiv:1302.2394.
29. CMS Collaboration, "Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **75** (2015) 288, doi:10.1140/epjc/s10052-015-3499-1, arXiv:1410.6765.
30. Z. Nagy, "Three jet cross-sections in hadron hadron collisions at next-to-leading order", *Phys. Rev. Lett.* **88** (2002) 122003, doi:10.1103/PhysRevLett.88.122003, arXiv:hep-ph/0110315.
31. Z. Nagy, "Next-to-leading order calculation of three-jet observables in hadron hadron collisions", *Phys. Rev. D* **68** (2003) 094002, doi:10.1103/PhysRevD.68.094002, arXiv:hep-ph/0307268.

32. D. Britzger, K. Rabbertz, F. Stober, and M. Wobisch, "New features in version 2 of the fastNLO project", in *Proceedings, XX. International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012)*, p. 217. Bonn, Germany, March 26-30, 2012. arXiv:1208.3641. doi:10.3204/DESY-PROC-2012-02/165.
33. T. Carli et al., "A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project", *Eur. Phys. J. C* **66** (2010) 503, doi:10.1140/epjc/s10052-010-1255-0, arXiv:0911.2985.
34. CDF Collaboration, "Inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **77** (1996) 438, doi:10.1103/PhysRevLett.77.438, arXiv:hep-ex/9601008.
35. J. Huston et al., "Large transverse momentum jet production and the gluon distribution inside the proton", *Phys. Rev. Lett.* **77** (1996) 444, doi:10.1103/PhysRevLett.77.444, arXiv:hep-ph/9511386.
36. J. R. Andersen et al., "Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report", arXiv:1405.1067.
37. S. Dittmaier, A. Huss, and C. Speckner, "Weak radiative corrections to dijet production at hadron colliders", *JHEP* **11** (2012) 095, doi:10.1007/JHEP11(2012)095, arXiv:1210.0438.
38. ATLAS Collaboration, "Measurement of dijet cross sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector", *JHEP* **05** (2014) 059, doi:10.1007/JHEP05(2014)059, arXiv:1312.3524.
39. S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
40. S. Alioli et al., "Jet pair production in POWHEG", *JHEP* **04** (2011) 081, doi:10.1007/JHEP04(2011)081, arXiv:1012.3380.
41. T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
42. AFS Collaboration, "Dijet Production Cross-Section and Fragmentation of Jets Produced in pp Collisions at $\sqrt{s} = 63$ GeV", *Z. Phys. C* **30** (1986) 27, doi:10.1007/BF01560675.
43. CDF Collaboration, "A measurement of the differential dijet mass cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. D* **61** (2000) 091101, doi:10.1103/PhysRevD.61.091101, arXiv:hep-ex/9912022.
44. CDF Collaboration, "Measurement of the dijet mass distribution in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. D* **48** (1993) 998, doi:10.1103/PhysRevD.48.998.
45. CDF Collaboration, "Two-jet invariant-mass distribution at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. D* **41** (1990) 1722, doi:10.1103/PhysRevD.41.1722.
46. D0 Collaboration, "The dijet mass spectrum and a search for quark compositeness in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **82** (1999) 2457, doi:10.1103/PhysRevLett.82.2457, arXiv:hep-ex/9807014.
47. CDF Collaboration, "Two-Jet Differential Cross-Section in $\bar{p}p$ Collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **64** (1990) 157, doi:10.1103/PhysRevLett.64.157.
48. CDF Collaboration, "Measurement of the two-jet differential cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1800$ GeV", *Phys. Rev. D* **64** (2001) 012001, doi:10.1103/PhysRevD.65.039902,10.1103/PhysRevD.64.012001, arXiv:hep-ex/0012013.
49. W. T. Giele, E. W. N. Glover, and D. A. Kosower, "The Two-Jet Differential Cross Section at $\mathcal{O}(\alpha_s^3)$ in Hadron Collisions", *Phys. Rev. Lett.* **73** (1994) 2019, doi:10.1103/PhysRevLett.73.2019, arXiv:hep-ph/9403347.
50. W. T. Giele, E. W. N. Glover, and D. A. Kosower, "Inclusive two jet triply differential cross section", *Phys. Rev. D* **52** (1995) 1486, doi:10.1103/PhysRevD.52.1486, arXiv:hep-ph/9412338.

51. A. D. Martin, W. J. Stirling, and R. G. Roberts, "Two jet hadroproduction as a measure of the gluon at small x", *Phys. Lett. B* **318** (1993) 184, doi:10.1016/0370-2693(93)91804-V, arXiv:hep-ph/9309204.
52. CMS Collaboration, "Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **700** (2011) 187, doi:10.1016/j.physletb.2011.05.027, arXiv:1104.1693.
53. S. D. Ellis, Z. Kunszt, and D. E. Soper, "Two-Jet Production in Hadron Collisions at Order α_s^3 in QCD", *Phys. Rev. Lett.* **69** (1992) 1496, doi:10.1103/PhysRevLett.69.1496.
54. D0 Collaboration, "Measurement of three-jet differential cross sections $d\sigma_{3\text{jet}}/dM_{3\text{jet}}$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Lett. B* **704** (2011) 434, doi:10.1016/j.physletb.2011.09.048, arXiv:1104.1986.
55. ATLAS Collaboration, "Measurement of three-jet production cross-sections in pp collisions at 7 TeV centre-of-mass energy using the ATLAS detector", *Eur. Phys. J. C* **75** (2014) 228, doi:10.1140/epjc/s10052-015-3363-3, arXiv:1411.1855.
56. CMS Collaboration, "Measurement of the inclusive 3-jet production differential cross section in proton-proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range", *Eur. Phys. J. C* **75** (2015) 186, doi:10.1140/epjc/s10052-015-3376-y, arXiv:1412.1633.
57. H.-L. Lai et al., "New parton distributions for collider physics", *Phys. Rev. D* **82** (2010) 074024, doi:10.1103/PhysRevD.82.074024, arXiv:1007.2241.
58. S. Alekhin, J. Blümlein, and S. Moch, "Parton Distribution Functions and Benchmark Cross Sections at NNLO", *Phys. Rev. D* **86** (2012) 054009, doi:10.1103/PhysRevD.86.054009, arXiv:1202.2281.
59. O. Ducu, L. Heurtier, and J. Maurer, "LHC signatures of a Z' mediator between dark matter and the SU(3) sector", *JHEP* **03** (2016) 006, doi:10.1007/JHEP03(2016)006, arXiv:1509.05615.
60. L. Lyons, A. J. Martin, and D. H. Saxon, "On the Determination of the *B* Lifetime by Combining the Results of Different Experiments", *Phys. Rev. D* **41** (1990) 982, doi:10.1103/PhysRevD.41.982.
61. G. D'Agostini, "Bayesian Reasoning in Data Analysis: A Critical Introduction". World Scientific Publishing Co. Pte. Ltd., Singapore, 2003.
62. R. D. Ball et al., "Fitting parton distribution data with multiplicative normalization uncertainties", *JHEP* **05** (2010) 075, doi:10.1007/JHEP05(2010)075, arXiv:0912.2276.
63. A. Banfi, G. P. Salam, and G. Zanderighi, "Phenomenology of event shapes at hadron colliders", *JHEP* **06** (2010) 038, doi:10.1007/JHEP06(2010)038, arXiv:1001.4082.
64. G. P. Salam and J. Rojo, "A Higher Order Perturbative Parton Evolution Toolkit (HOPPET)", *Comput. Phys. Commun.* **180** (2009) 120, doi:10.1016/j.cpc.2008.08.010, arXiv:0804.3755.
65. M. Glück, E. Reya, and A. Vogt, "Dynamical parton distributions revisited", *Eur. Phys. J. C* **5** (1998) 461, doi:10.1007/s100520050289, arXiv:hep-ph/9806404.
66. B. Schmidt and M. Steinhauser, "CRunDec: a C++ package for running and decoupling of the strong coupling and quark masses", *Comput. Phys. Commun.* **183** (2012) 1845, doi:10.1016/j.cpc.2012.03.023, arXiv:1201.6149.
67. K. G. Chetyrkin, J. H. Kuhn, and M. Steinhauser, "RunDec: A Mathematica package for running and decoupling of the strong coupling and quark masses", *Comput. Phys. Commun.* **133** (2000) 43, doi:10.1016/S0010-4655(00)00155-7, arXiv:hep-ph/0004189.
68. K. A. Olive and others (Particle Data Group), "Review of Particle Physics", *Chin. Phys. C* **38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.
69. CDF Collaboration, "Measurement of the strong coupling constant from inclusive jet production at the Tevatron $\bar{p}p$ collider", *Phys. Rev. Lett.* **88** (2002) 042001, doi:10.1103/PhysRevLett.88.042001, arXiv:hep-ex/0108034.

70. D0 Collaboration, "Determination of the strong coupling constant from the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. D* **80** (2009) 111107, doi:10.1103/PhysRevD.80.111107, arXiv:0911.2710.
71. D0 Collaboration, "Measurement of angular correlations of jets at $\sqrt{s} = 1.96$ TeV and determination of the strong coupling at high momentum transfers", *Phys. Lett. B* **718** (2012) 56, doi:10.1016/j.physletb.2012.10.003, arXiv:1207.4957.
72. B. Malaescu and P. Starovoitov, "Evaluation of the strong coupling constant α_S using the ATLAS inclusive jet cross-section data", *Eur. Phys. J. C* **72** (2012) 2041, doi:10.1140/epjc/s10052-012-2041-y, arXiv:1203.5416.
73. CMS Collaboration, "Measurement of the ratio of the inclusive 3-jet cross section to the inclusive 2-jet cross section in pp collisions at $\sqrt{s} = 7$ TeV and first determination of the strong coupling constant in the TeV range", *Eur. Phys. J. C* **73** (2013) 2604, doi:10.1140/epjc/s10052-013-2604-6, arXiv:1304.7498.
74. CMS Collaboration, "Determination of the top-quark pole mass and strong coupling constant from the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **728** (2014) 496, doi:10.1016/j.physletb.2013.12.009, arXiv:1307.1907.
75. CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
76. H1 Collaboration, "Jet Production in ep Collisions at High Q^2 and Determination of α_s", *Eur. Phys. J. C* **65** (2010) 363, doi:10.1140/epjc/s10052-009-1208-7, arXiv:0904.3870.
77. H1 Collaboration, "Jet Production in ep Collisions at Low Q^2 and Determination of α_s", *Eur. Phys. J. C* **67** (2010) 1, doi:10.1140/epjc/s10052-010-1282-x, arXiv:0911.5678.
78. H1 Collaboration, "Measurement of multijet production in ep collisions at high Q^2 and determination of the strong coupling α_s", *Eur. Phys. J. C* **75** (2015) 65, doi:10.1140/epjc/s10052-014-3223-6, arXiv:1406.4709.
79. ZEUS Collaboration, "Inclusive-jet photoproduction at HERA and determination of α_s", *Nucl. Phys. B* **864** (2012) 1, doi:10.1016/j.nuclphysb.2012.06.006, arXiv:1205.6153.
80. NNPDF Collaboration, "A determination of parton distributions with faithful uncertainty estimation", *Nucl. Phys. B* **809** (2009) 1, doi:10.1016/j.nuclphysb.2008.09.037, arXiv:0808.1231.
81. S. Alekhin et al., "HERAFitter", *Eur. Phys. J. C* **75** (2015) 304, doi:10.1140/epjc/s10052-015-3480-z, arXiv:1410.4412.
82. HERAFitter web site, http://www.herafitter.org.
83. H1 and ZEUS Collaboration, "Combined Measurement and QCD Analysis of the Inclusive $e^{\pm}p$ Scattering Cross Sections at HERA", *JHEP* **01** (2010) 109, doi:10.1007/JHEP01(2010)109, arXiv:0911.0884.
84. M. Botje, "QCDNUM: Fast QCD Evolution and Convolution", *Comput. Phys. Commun.* **182** (2011) 490, doi:10.1016/j.cpc.2010.10.020, arXiv:1005.1481.
85. R. S. Thorne and R. G. Roberts, "An Ordered analysis of heavy flavor production in deep inelastic scattering", *Phys. Rev. D* **57** (1998) 6871, doi:10.1103/PhysRevD.57.6871, arXiv:hep-ph/9709442.
86. R. S. Thorne, "Variable-flavor number scheme for next-to-next-to-leading order", *Phys. Rev. D* **73** (2006) 054019, doi:10.1103/PhysRevD.73.054019, arXiv:hep-ph/0601245.
87. H1 and ZEUS Collaboration, "Combination and QCD Analysis of Charm Production Cross Section Measurements in Deep-Inelastic ep Scattering at HERA", *Eur. Phys. J. C* **73** (2013) 2311, doi:10.1140/epjc/s10052-013-2311-3, arXiv:1211.1182.
88. NuTeV Collaboration, "Measurement of the Nucleon Strange-Antistrange Asymmetry at Next-to-Leading Order in QCD from NuTeV Dimuon Data", *Phys. Rev. Lett.* **99** (2007) 192001, doi:10.1103/PhysRevLett.99.192001.

89. L. Demortier, "Equivalence of the best-fit and covariance matrix methods for comparing binned data with a model in the presence of correlated systematic uncertainties", (April, 1999). CDF Note 8661.
90. D. Stump et al., "Uncertainties of predictions from parton distribution functions. 1. The Lagrange multiplier method", *Phys. Rev. D* **65** (2001) 014012, doi:10.1103/PhysRevD.65.014012, arXiv:hep-ph/0101051.
91. M. Botje, "Error estimates on parton density distributions", *J. Phys. G* **28** (2002) 779, doi:10.1088/0954-3899/28/5/305, arXiv:hep-ph/0110123.
92. J. Gao et al., "MEKS: a program for computation of inclusive jet cross sections at hadron colliders", *Comput. Phys. Commun.* **184** (2013) 1626, doi:10.1016/j.cpc.2013.01.022, arXiv:1207.0513.
93. R. D. Ball et al., "Parton Distribution Benchmarking with LHC Data", *JHEP* **04** (2013) 125, doi:10.1007/JHEP04(2013)125, arXiv:1211.5142.
94. B. J. A. Watt, P. Motylinski, and R. S. Thorne, "The Effect of LHC Jet Data on MSTW PDFs", arXiv:1311.5703.

Chapter 5
Cross Section Ratios

As demonstrated in the previous chapter, valuable insights can be gained from the measurement of absolute cross sections. However, they are subject to the totality of experimental and theoretical uncertainties. Therefore, it is worthwhile to contemplate alternative observables that are either insensitive or sensitive at a reduced level to the dominant sources of uncertainty. The possibility discussed in this chapter deals with cross section ratios, which, with one exception, are unaffected by imprecise luminosity determinations and which are expected to exhibit at least partial cancellations in JEC and JER effects on the experimental side and a reduced dependence on the QCD scales or on PDFs on the theory side.

The following ratio observables and their interpretation are presented in the next four sections: the ratio of the inclusive 3-jet event to the inclusive 2-jet event cross section, the ratio of inclusive jet cross sections for two different jet sizes or at two different centre-of-mass energies, and the ratio of the dijet cross section at outer to the one at inner pseudorapidity.

5.1 Determination of $\alpha_S(M_Z)$ with n-Jet Ratios

In this first example, the ratio R_{32} of the inclusive 3-jet event to the inclusive 2-jet event cross section is considered as a function of the average p_T, $\langle p_{\mathrm{T}1,2}\rangle$, of the two jets leading in p_T:

$$R_{32}(\langle p_{\mathrm{T}1,2}\rangle) = \frac{\mathrm{d}\sigma_{N_\mathrm{jet}\geq 3}\mathrm{d}\langle p_{\mathrm{T}1,2}\rangle}{\mathrm{d}\sigma_{N_\mathrm{jet}\geq 2}\mathrm{d}\langle p_{\mathrm{T}1,2}\rangle}\,. \tag{5.1}$$

The LO process for the denominator is proportional to α_S^2, while the numerator requires one more jet to be present and therefore is proportional to α_S^3 at LO. This ratio directly depends on α_S, but at the same time profits from uncertainty cancellations. Like in the previous chapter, pQCD predictions are available at NLO for the numerator and denominator from [1, 2]. Other theoretical programs that have become available in the meantime are even capable to compute n-jet cross sections at

NLO up to four or five jets [3–5]. To further decrease the sensitivity to experimental effects, it is attractive to study n-jet cross section ratios leading to proportionalities to higher powers of α_S like $R_{42}, R_{53} \propto \alpha_S^2$ or $R_{52} \propto \alpha_S^3$, although it comes at the expense of having less data and a more restricted phase space. The authors of Ref. [5] also argue that for theoretical reasons the ratios $R_{43}, R_{54} \propto \alpha_S$ might be preferable over R_{32}. Dedicated measurements with the aim of exploiting these new possibilities have not yet been performed.

The UA1 Collaboration published a first investigation of such a ratio in Ref. [6]. The UA2 experiment determined $\alpha_S(M_Z)$ based on the ratio of W($\to e\nu$)+1 jet to W($\to e\nu$)+0 jet production [7]. CMS has measured the ratio R_{32} at the LHC for the data set recorded in 2011 and extracted a value for $\alpha_S(M_Z)$ [8]. The ATLAS experiment so far presented only a preliminary study of R_{32} on data from 2010 [9].[1] Alternative ratios with 3-jet observables in the numerator have been studied at Tevatron [10] and in Ref. [11].

The CMS measurement selects all jets with transverse momenta larger than 150 GeV up to rapidities of five. All events where the two jets leading in p_T are not found in the central detector, i.e. $|y| \leq 2.5$, are rejected. All remaining events contribute at $\langle p_{T1,2} \rangle$ to the inclusive 2-jet event cross section. If a third jet with $p_T > 150$ GeV and $|y| \leq 2.5$ is found, the event also contributes to the numerator, i.e. the inclusive 3-jet event cross section. An example of such a 3-jet event is displayed in Fig. 5.1.

Figure 5.2 presents the CMS data of R_{32} in comparison to NLO predictions, which are compatible with the measurement within uncertainties. The error bars correspond to the total experimental uncertainty and comprise in addition to the statistical uncertainty the systematic components from the JEC determination and from the unfolding of detector effects. The JEC uncertainty, estimated to be 2.0–2.5% for particle-flow jets [12], contributes to about 1.2% to the total systematic uncertainty, which is much smaller than for the absolute cross section measurements discussed before, cf. Figs. 4.3, 4.8, and 4.11. The unfolding uncertainty amounts to less than 1% and accounts for three different mutually uncorrelated uncertainty sources: insufficient knowledge of the $\langle p_{T1,2} \rangle$ resolution, non-Gaussian components in the $\langle p_{T1,2} \rangle$ resolution, and insufficient knowledge of the simulated inclusive 3-jet and 2-jet $\langle p_{T1,2} \rangle$ spectra, which are used to construct the unfolding matrices. Since both experimental uncertainties are very small because of cancellations in the ratio, a potential bias originating from the unfolding technique was studied by comparing the unfolding result of the iterative D'Agostini method with that of the SVD method, cf. Section 3.6.2. The bias was found to be negligible. The by far dominating uncertainty at high $\langle p_{T1,2} \rangle$ is of statistical origin.

NP corrections, which have been applied to the NLO predictions in Fig 5.2, range from only 2% at low $\langle p_{T1,2} \rangle$ to zero at high $\langle p_{T1,2} \rangle$. The corresponding systematic uncertainty is negligible. PDF uncertainties, evaluated according to the respective prescriptions for each set, are shown here for the CT10-NNLO PDFs. At small to medium parton fractional momenta, and hence small p_T's or masses, the PDFs are

[1] ATLAS additionally investigated the quantity N_{32}, which in contrast to R_{32}, receives not only a single entry per event, but one entry per jet in each event.

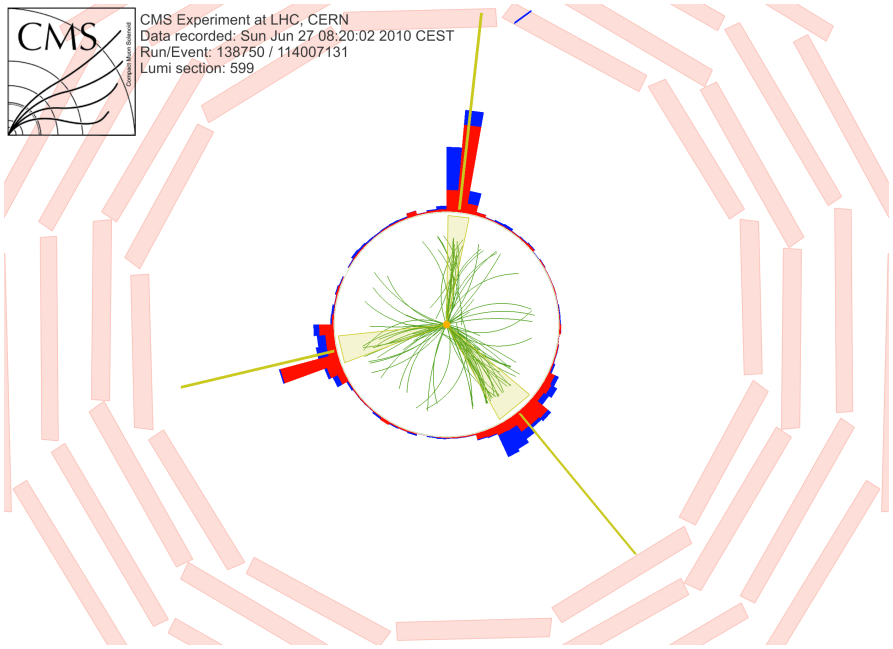

Fig. 5.1 Event display of a 3-jet event recorded with the CMS detector. Starting from the interaction vertex in the centre, tracks of charged particles (green lines), and energy deposits in the electromagnetic and hadronic calorimeters (red and blue columns) are clustered together to form three jets. The jets are represented by yellow, triangular rays and thick yellow lines showing each jet's momentum. The light red rectangles show the muon detection system. *(Source: CERN, CMS)*

well known from previous experiments and cause about 2% uncertainty for jet cross section predictions as well as for the ratio R_{32}, cf. Figs. 4.4, 4.8, and 4.12. At higher jet p_{T}, PDF uncertainties increase considerably up to 20% or more for cross sections, while R_{32} profits from cancellation effects between numerator and denominator limiting the PDF uncertainty to about 3% for $|y| \leq 2.5$ at $\langle p_{\mathrm{T1,2}} \rangle \approx 1.5\,\mathrm{TeV}$. Equally, scale uncertainties indicated by the green bands in Fig 5.2 are reduced to below 5% in the ratio as compared to 5–15% for absolute measurements, if one stays away from phase space regions that are subject to kinematic constraints. Turn-on effects appear below $\approx 420\,\mathrm{GeV}$ of $\langle p_{\mathrm{T1,2}} \rangle$, which has been chosen to be the scale $Q = \mu_r = \mu_f$. If, for example, the leading jet has to balance the minimal p_{T} of 150 GeV of two opposite side jets, it needs at least 300 GeV of p_{T}, which leads to a minimal $\langle p_{\mathrm{T1,2}} \rangle$ of 225 GeV to kinematically permit this configuration. In the vicinity of such thresholds fixed-order pQCD calculations are less reliable and might be subject to large K factors respectively higher-order corrections. In addition, the sensitivity to $\alpha_S(M_Z)$ is reduced as can be concluded from Fig 5.2 right, where the CMS data are compared to NLO predictions for a series of values of $\alpha_S(M_Z)$ available for the CT10-NNLO PDF sets. The picture is similar for the alternative sets NNPDF2.1-NNLO or MSTW2008-NNLO so that like in Section 4.4 a value of $\alpha_S(M_Z)$ can be

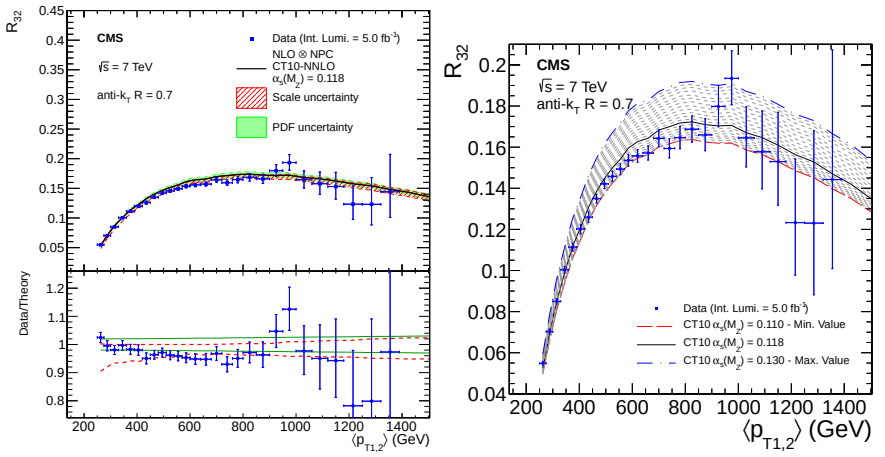

Fig. 5.2 Left: CMS measurement of R_{32} (full circles) in comparison to NLO predictions corrected for NP effects and using the CT10-NNLO PDF set. The error bars correspond to the total uncertainty. The bottom panel shows the ratio of data to the predictions. The bands indicated by lines represent the scale (dashed) and PDF (solid) uncertainties. Right: Sensitivity of R_{32} to $\alpha_S(M_Z)$. The measured R_{32} is presented together with NP corrected NLO predictions using the CT10-NNLO PDF sets for a series of values of $\alpha_S(M_Z)$ from 0.110–0.130 in steps of 0.001. *(Taken from Ref. [8])*

extracted for $\langle p_{\mathrm{T1,2}} \rangle > 420\,\mathrm{GeV}$ from fits to the measured R_{32}:

$$\begin{aligned}\alpha_S(M_Z) &= 0.1148 \pm 0.0014\,(\mathrm{exp}) \pm 0.0000\,(\mathrm{NP}) \pm 0.0018\,(\mathrm{PDF}) \pm 0.0050\,(\mathrm{scale}) \\ &= 0.1148 \pm 0.0023\,(\mathrm{all\ except\ scale}) \pm 0.0050\,(\mathrm{scale}) \\ &= 0.1148 \pm 0.0055 \qquad (R_{32})\,,\end{aligned} \tag{5.2}$$

where the NNPDF2.1-NNLO PDF set was used in combination with a 3-loop solution to the RGE for the evolution of $\alpha_S(Q)$. For comparison, the CT10- and MSTW2008-NNLO PDF sets lead to the compatible results of:

$$\begin{aligned}\text{MSTW2008: } \alpha_S(M_Z) &= 0.1141 \pm 0.0022\,(\mathrm{exp}), \\ \text{CT10: } \alpha_S(M_Z) &= 0.1135 \pm 0.0019\,(\mathrm{exp})\,.\end{aligned} \tag{5.3}$$

Customarily, cross sections are computed with the scale dependence of the perturbative expansion in α_S and of the parton densities at the same order. Here, CMS combined an NLO prediction with PDF sets evolved at NNLO, which sometimes is considered to be inconsistent. Following Ref. [13] this is not necessarily so and depends on the particular case, as long as one is aware that the overall accuracy is always the one of the lowest order involved. CMS argues that R_{32} does not directly depend on the PDFs such that this issue does not pose any problem for this analysis. An explicit cross check by CMS revealed a negligible impact of the usage of NLO versus NNLO PDF sets in fits of the strong coupling constant. This is in contrast

Table 5.1 Determination of $\alpha_S(M_Z)$ for different ranges of scale Q from the inclusive 3-jet over inclusive 2-jet event cross section ratio R_{32} [8] using the NNPDF2.1-NNLO PDF set [15]. The last two columns list the corresponding values of the average scales $\langle Q \rangle$ and the strong coupling constant at this scale.

Q range	χ^2/n_{dof}	$\alpha_S(M_Z)$	$\pm$(exp)	$\pm$(NP)	$\pm$(PDF)	$\pm$(scale)	$\langle Q \rangle$/GeV	$\alpha_S(Q)$
420–600	4.4/5	0.1147	0.0015	—	0.0015	0.0057	474	0.0936
600–800	5.9/4	0.1132	0.0018	—	0.0025	0.0039	664	0.0894
800–1390	5.7/9	0.1170	0.0024	—	0.0021	0.0048	896	0.0889

to fits to cross sections, where NNLO PDF sets usually lead to smaller values of $\alpha_S(M_Z)$ than NLO ones, and confirms, together with the much smaller PDF uncertainty, the reduced dependence of R_{32} on details of the PDF evolution. Hence, fits of $\alpha_S(Q)$ to R_{32} for regions in scale $Q = \langle p_{\text{T1,2}} \rangle$ test the evolution of $\alpha_S(Q)$ predicted by the RGE, which could be modified by the existence of new coloured matter, cf. Section 7.3 and Ref. [14]. Table 5.1 presents the outcome of such fits for three regions in $\langle p_{\text{T1,2}} \rangle$ from 420 up to 1390 GeV, which are shown versus Q together with the results from the last section in Fig. 4.15.

Considering the QCD scale it is certainly natural to choose for dijet production at LO: $Q = p_{\text{T,max}} = p_{\text{T,jet1}} = p_{\text{T,jet2}} = \langle p_{\text{T1,2}} \rangle$. At NLO $p_{\text{T,jet2}}$ may lose energy to a third jet, while $p_{\text{T,max}}$ and $p_{\text{T,jet1}}$ remain unchanged. If a quantity like R_{32} is not expanded itself in terms of powers of α_S, but derived as a ratio of NLO cross sections, it is desirable to have a scale that can be derived the same way for the numerator and denominator, particularly, if one of the two cross sections completely contains the other one. This excludes a scale definition explicitly involving $p_{\text{T,jet3}}$ as suggested in [16], and leaves either $p_{\text{T,max}}$ or $\langle p_{\text{T1,2}} \rangle$. Following good experience in the form of small K factors with $p_{\text{T,jet}}$ or $\langle p_{\text{T1,2}} \rangle$ as scale instead of $p_{\text{T,max}}$, CMS opted for $\langle p_{\text{T1,2}} \rangle$. In the 3-jet case, the typical scale variation by factors of $1/2$ and 2 around $\langle p_{\text{T1,2}} \rangle$ does include the $p_{\text{T,max}}$ values for a symmetric star-like configuration with $\langle p_{\text{T1,2}} \rangle = p_{\text{T,max}}$ and for the extreme case, where one jet balances two opposite-side ones, with $\langle p_{\text{T1,2}} \rangle = 3/4 \cdot p_{\text{T,max}}$. To demonstrate that $\langle p_{\text{T1,2}} \rangle$ is a reasonable scale choice also in case of correlated scale variations in the ratio, Fig. 5.3 shows the scale dependence for the numerator alone and for the ratio R_{32}. The scale dependence of the 2-jet inclusive event cross section is in-between the latter two and leads to the observed reduced scale dependence in Fig. 5.3 right compared to Fig. 5.3 left. The default scale, i.e. $\mu_r/Q = \mu_f/Q = 1$, lies in an uncritical flat region and unreasonable complete cancellations are not exhibited. To extend the scale choice to an arbitrary number of jets, another possibility exists by using the scalar sum of the jet p_{T}'s:

$$H_{\text{T}} = H_{\text{T}}^{(N)} = \sum_{i=1}^{N} p_{\text{T,i}} \,. \tag{5.4}$$

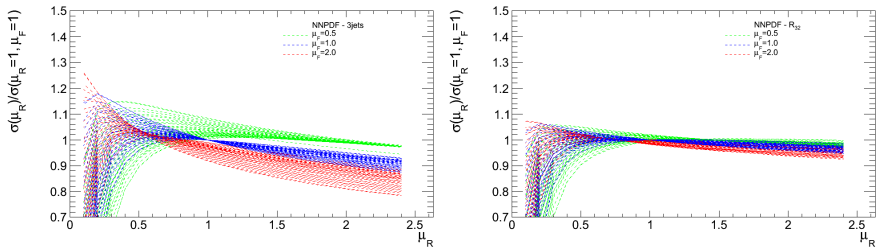

Fig. 5.3 Scale dependence of the NLO predictions for the inclusive 3-jet event cross section (left) and for R_{32} (right). *(Taken from auxiliary material of Ref. [8])*

Limiting the summation to the two leading jets, i.e. $N = 2$, one retrieves $\langle p_{\mathrm{T}1,2} \rangle = H_{\mathrm{T}}^{(2)}/2$. Following Refs. [5, 11] $Q = H_{\mathrm{T}}/2$ could be tried in the future in particular for the case of multi-jet production.

5.2 Parton Showers and Jet Size Ratios

An aspect ignored so far is the choice of the jet size R. ATLAS regularly uses the two jet sizes $R = 0.4$ and 0.6, while CMS employs $R = 0.5$ or 0.7 depending on the analysis. When comparing the measurements separately to pQCD, compatibility within uncertainties is observed. One obvious difference though consists in the NP corrections, which are quite different as shown e.g. for dijet production with $R = 0.4$ and 0.6 in Fig. 4.9. Looking more closely into the results presented by ATLAS for anti-k_{t} jet sizes of $R = 0.4$ and 0.6 [17] or by CMS for anti-k_{t} jet sizes of $R = 0.5$ and 0.7 [18, 19], a tendency becomes visible, namely that the relative normalisation of the measured cross sections to the theoretical predictions exhibits a dependence on R. Theoretically, this R dependence has been examined in Refs. [20, 21], where in a collinear approximation it was found that the impact of perturbative radiation and of the non-perturbative effects of hadronisation and the underlying event on jet transverse momenta scales roughly with $\ln R$, $-1/R$, and R^2 for small R, respectively. By choosing the jet size parameter R, one can therefore vary which aspects of jet formation are emphasised in a jet analysis, cf. Section 2.5.2.

To gain more insight into the interplay of these effects, it is suggested in Ref. [20] to study the relative difference between inclusive jet cross sections that emerge from two different jet definitions labelled here as reference (ref) and alternative (alt):

$$\left(\frac{\mathrm{d}\sigma^{\mathrm{alt}}}{\mathrm{d}p_{\mathrm{T}}} - \frac{\mathrm{d}\sigma^{\mathrm{ref}}}{\mathrm{d}p_{\mathrm{T}}}\right) \Bigg/ \left(\frac{\mathrm{d}\sigma^{\mathrm{ref}}}{\mathrm{d}p_{\mathrm{T}}}\right) = \mathcal{R}(\mathrm{alt}, \mathrm{ref}) - 1\,. \tag{5.5}$$

Provided that partons in opposite hemispheres are not clustered together, a condition that is usually fulfilled, the LO two-parton cross sections are identical for arbitrary jet algorithms. Therefore, only partonic final states with three or more partons lead

to a numerator different from zero. Hence, Equation 5.5 defines a 3-jet observable, $\mathcal{R}$, for which it was shown in Ref. [22] that it is calculable to NLO with terms up to α_S^4 with NLOJET++ [1, 2].

Studies in that direction have been performed at the HERA collider by the ZEUS collaboration for two different jet algorithms [23] and by the ALICE experiment for the two different anti-k_t jet sizes of $R = 0.2$ and 0.4 [24]. The results of a CMS analysis for the two jet sizes of 0.5 and 0.7 is presented in Ref. [25] as a function of jet p_T and rapidity y. It is expected that QCD radiation reduces this jet-radius ratio $\mathcal{R}$ below unity and that this effect disappears with increasing collimation of jets at high p_T. Figure 5.4 confirms this assumption with a comparison of the measured ratio $\mathcal{R}$ as a function of the jet p_T up to the TeV scale to theoretical predictions at small rapidity $|y| < 0.5$. Figure 5.4 left clearly demonstrates that fixed-order calculations up to NLO, even when combined with non-perturbative corrections, are systematically above the data. The LO event generators PYTHIA6 and HERWIG++ lie systematically on either side of the data, as can be seen in Fig. 5.4 right. Presumably, this can be improved by including this observable into the tuning of these generators. The best description is given by POWHEG that matches the dijet production process evaluated at NLO with the parton showers and non-perturbative models of PYTHIA6, emphasising the importance of parton showers for the given choice of jet sizes.

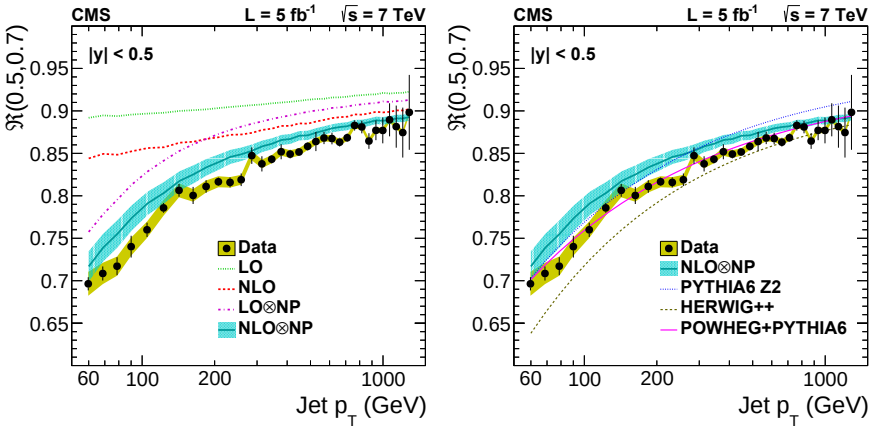

Fig. 5.4 Jet-radius ratio $\mathcal{R}(0.5, 0.7)$ as measured by CMS in comparison to fixed-order predictions with and without non-perturbative corrections (left) and to particle-level predictions of LO and NLO event generators with matched parton showers and modelling of hadronisation and the underlying event (right) at small rapidity $|y| < 0.5$ *(Adapted from auxiliary material of Ref. [25])*

5.3 PDF Constraints from Ratios at different Centre-of-Mass Energies

The LHC has not only taken data at 7 TeV centre-of-mass energy, but also at 8 and at 2.76 TeV.[2] This opens up the possibility to do cross section ratios for two different centre-of-mass energies [26] as studied previously at the Sp$\bar{\text{p}}$S [27, 28] and Tevatron [29, 30]. To remove obvious differences in magnitude and p_{T} range, the invariant cross section $E\mathrm{d}^3\sigma/\mathrm{d}p^3$, where E and p denote energy and momentum of a jet, can be rewritten as a dimensionless quantity [31]

$$p_{\text{T}}^4 E \cdot \frac{\mathrm{d}^3\sigma}{\mathrm{d}p^3} = \frac{p_{\text{T}}^3}{2\pi} \cdot \frac{\mathrm{d}^2\sigma}{\mathrm{d}p_{\text{T}}\mathrm{d}y} = \frac{s}{8\pi} x_{\text{T}}^3 \cdot \frac{\mathrm{d}^2\sigma}{\mathrm{d}x_{\text{T}}\mathrm{d}y} \tag{5.6}$$

that is investigated as a function of scaled momentum $x_{\text{T}} = 2p_{\text{T}}/\sqrt{s}$. The x_{T} *scaling* hypothesis, motivated by the QPM, predicts this re-scaled cross section to be independent of $\sqrt{s}$ [32, 33]. Since PDFs and α_S evolve in QCD theory, deviations from this behaviour are expected. These *scaling violations* were observed at Sp$\bar{\text{p}}$S and Tevatron and were found to be in agreement with QCD predictions. Studying the scale dependence in detail, therefore, provides constraints on the PDFs and $\alpha_S(M_Z)$.

At the LHC, ATLAS studied the ratio of inclusive jet production for 2.76 and 7 TeV centre-of-mass energies as a function of jet rapidity and the scaled momentum x_{T} or the jet p_{T} [34]. Taking the invariant cross section Eq. (5.6) and defining

$$F(y, x_{\text{T}}, s) = \frac{s}{8\pi} x_{\text{T}}^3 \cdot \frac{\mathrm{d}^2\sigma}{\mathrm{d}x_{\text{T}}\mathrm{d}y} \qquad \text{and} \qquad F(y, p_{\text{T}}, s) = \frac{p_{\text{T}}^3}{2\pi} \cdot \frac{\mathrm{d}^2\sigma}{\mathrm{d}p_{\text{T}}\mathrm{d}y} \tag{5.7}$$

the ratios can be written as

$$\rho(y, x_{\text{T}}, s_1/s_2) = \frac{F(y, x_{\text{T}}, s_1)}{F(y, x_{\text{T}}, s_2)} \qquad \text{and} \qquad \rho(y, p_{\text{T}}, s_1/s_2) = \frac{F(y, p_{\text{T}}, s_1)}{F(y, p_{\text{T}}, s_2)} \tag{5.8}$$

such that with $s_1 = 2.76\,\text{TeV}$ and $s_2 = 7\,\text{TeV}$

$$\rho(y, x_{\text{T}}) = \left(\frac{2.76}{7}\right)^3 \frac{\sigma(y, x_{\text{T}}, 2.76\,\text{TeV})}{\sigma(y, x_{\text{T}}, 7\,\text{TeV})} \qquad \text{and} \qquad \rho(y, p_{\text{T}}) = \frac{\sigma(y, p_{\text{T}}, 2.76\,\text{TeV})}{\sigma(y, p_{\text{T}}, 7\,\text{TeV})}\,. \tag{5.9}$$

The inclusive jet cross section at $\sqrt{s} = 7\,\text{TeV}$ with $37\,\text{pb}^{-1}$ of integrated luminosity has been published previously by ATLAS [17], cf. also Section 4.2. The one at 2.76 TeV is measured from $0.20\,\text{pb}^{-1}$ of integrated luminosity collected at the beginning of the 2011 data taking period. The accessible phase space ranges from 20 up to 430 GeV in jet p_{T} and up to 4.4 in rapidity. To perform the ratios to the 7 TeV cross section, the same p_{T} binning has been chosen. In addition, an x_{T} dependent binning has been evaluated whose range in x_{T} is identical to the one derived from

[2] 2.76 TeV is the baseline pp centre-of-mass energy for comparisons with heavy-ion ($PbPb$) collisions at the LHC with 2.76 TeV of energy per nucleon-nucleon pair.

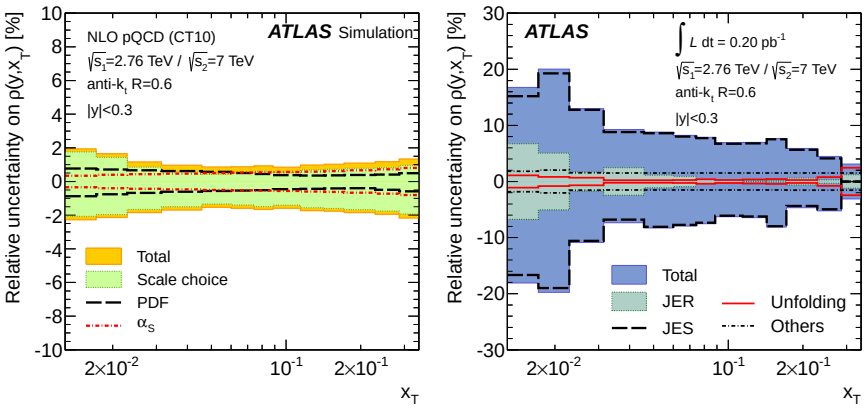

Fig. 5.5 Theoretical (left) and experimental (right) systematic uncertainties derived by ATLAS for the ratio $\rho(y, x_T)$ of invariant jet cross sections at 2.76 and 7 TeV centre-of-mass energies as a function of scaled momentum x_T. *(Taken from Ref. [34])*

the p_T binning at 7 TeV. In the ratio $\rho(y, x_T)$, theoretical uncertainties are reduced in particular the one from PDFs, which are evaluated at similar momentum fractions. However, because jets are measured at different p_T in this ratio, experimental effects profit less from cancellations as shown in Fig. 5.5. Comparing data and theory scaling violations compatible with QCD predictions are clearly visible [34].

To constrain PDFs, the ratio $\rho(y, p_T)$ is better suited than the x_T dependent one, because here the jets in the numerator and denominator are measured in the same y and p_T regions such that experimental uncertainties cancel to a large extent. Theoretical uncertainties are less reduced as visible in Fig. 5.6. Note, however, that in these ratios, $\rho(y, x_T)$ and $\rho(y, p_T)$, luminosity uncertainties do NOT cancel, but add up quadratically to 4.3%, because they are considered to be uncorrelated between the two data taking periods. Luminosity uncertainties are not indicated in Figs. 5.5 and 5.6.

Multi-parton interactions, which model the underlying event, depend on the centre-of-mass energy. Therefore, it can not be expected that NP corrections for the inclusive jet p_T spectrum cancel in the ratio for different centre-of-mass energies. The evaluation by ATLAS for $\rho(y, p_T)$ shown in Fig. 5.7 right is contrasted with the NP corrections for the inclusive jet cross section at 2.76 TeV in Fig. 5.7 left. The larger NP corrections at 7 TeV lead to correction factors for the ratio $\rho(y, p_T)$ of 0.7 at lowest p_T of 20 GeV. For high p_T the NP factors approach unity as expected.

Figure 5.8 presents the ratio of the measured cross section ratio to the one predicted by theory for different regions in rapidity and in comparison to five different PDF sets. For absolute rapidities below $|y| = 2.1$ all PDFs predict slightly too small ratios. The predictions by the ABM11 PDF is particularly low. At higher rapidities the data are better described by theory for small jet p_T's, but start to deviate significantly when approaching the kinematic limit.

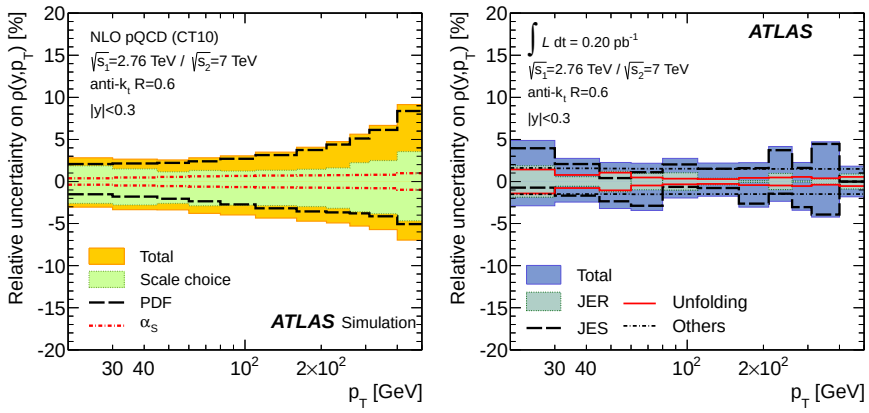

Fig. 5.6 Theoretical (left) and experimental (right) systematic uncertainties derived by ATLAS for the ratio $\rho(y, p_{\mathrm{T}})$ of invariant jet cross sections at 2.76 and 7 TeV centre-of-mass energies as a function of jet p_{T}. *(Taken from Ref. [34])*

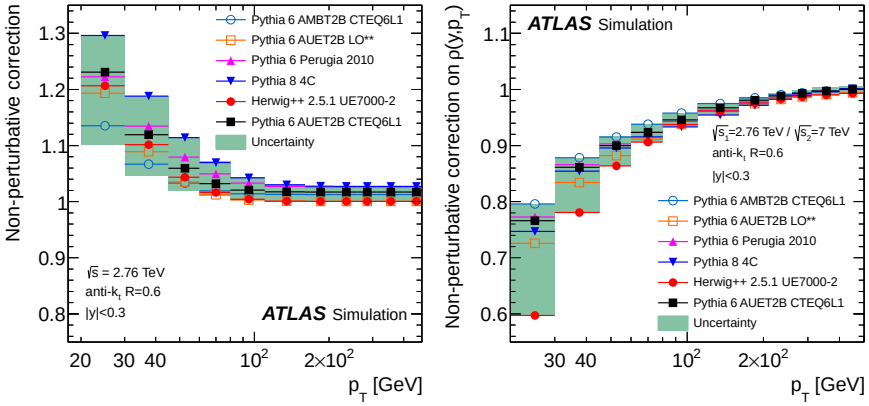

Fig. 5.7 NP corrections as a function of jet p_{T} from ATLAS for the inclusive jet cross section at $\sqrt{s} = 2.76\,\mathrm{TeV}$ (left) and for the ratio $\rho(y, p_{\mathrm{T}})$ of invariant jet cross sections at $\sqrt{s} = 2.76\,\mathrm{TeV}$ and $\sqrt{s} = 7\,\mathrm{TeV}$ (right). *(Taken from Ref. [34])*

To quantify the influence of the measured ratio $\rho(y, p_{\mathrm{T}})$ on PDF determinations, ATLAS performed a very similar study to the one described in Section 4.5 by using the HERAFITTER tool with a combination of HERA-I DIS and the ATLAS data. The impact on the gluon PDF is displayed in Fig. 5.9 left. When the ATLAS jet cross sections are included, either separately or in the ratio $\rho(y, p_{\mathrm{T}})$, the uncertainties on the gluon PDF are somewhat reduced. In particular the ratio also leads to a harder spectrum for the gluon. A similar effect was observed in the analysis of the CMS inclusive jet data presented previously. The resulting improved description of the

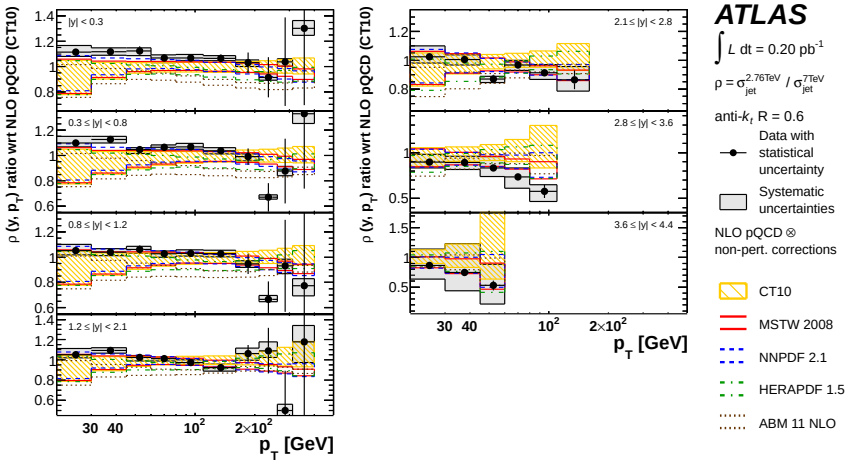

Fig. 5.8 Ratio of data over theory for the ratio $\rho(y, p_{\mathrm{T}})$ of invariant jet cross sections at 2.76 and 7 TeV centre-of-mass energies as a function of jet p_{T} for seven ranges in rapidity y. The predictions of pQCD complemented with NP corrections are derived including uncertainties for the five PDF sets CT10, MSTW2008, NNPDF2.1, HERAPDF1.5, and ABM11 with NLO evolution. *(Taken from Ref. [34])*

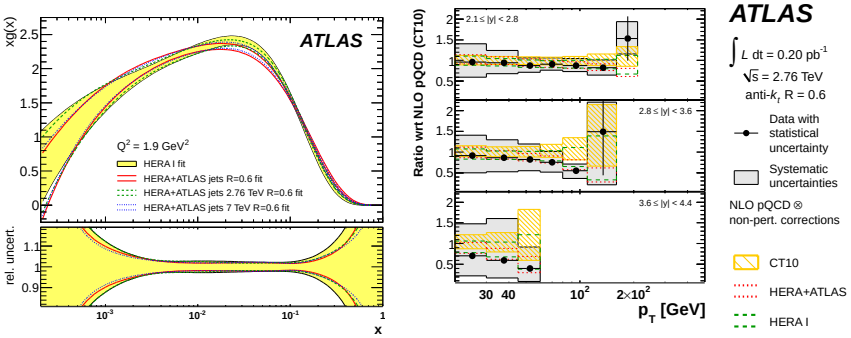

Fig. 5.9 Left: Gluon PDF derived with the HERAFITTER tool from HERA-I DIS data alone and in combination with ATLAS jet data at 2.76 TeV, 7 TeV or both centre-of-mass energies. Right: Improved predictions at outer rapidity for the inclusive jet cross section at $\sqrt{s} = 2.76$ GeV from fits to HERA-I DIS and ATLAS jet data in comparison to HERA-I data alone and to CT10. *(Adapted from Ref. [34])*

jet data in particular at larger rapidity by using the PDF derived with the combined HERA-I DIS and ATLAS jet data is clearly demonstrated in Fig. 5.9 right.

5.4 Search for New Phenomena with the Centrality Ratio

Finally, cross section ratios with reduced sensitivity to systematic effects can be exploited to search for deviations from SM expectations and to set limits on parameters in models of new phenomena beyond the SM [35–41]. Models of composite fermions assume quarks to be made up of even more fundamental constituents. Below a compositeness scale Λ these constituents are strongly bound together to the observed quark states by a new strong gauge interaction. Indications of such a substructure have not been observed and one expects Λ to be at least of the order of the electroweak scale, i.e. the vacuum expectation value of the Higgs boson

$$v = \left(\sqrt{2}G_F\right)^{-1/2} = 246\,\text{GeV}\,, \tag{5.10}$$

where $G_F = 1.1663787 \cdot 10^{-5}\,\text{GeV}^{-2}$ [16] is the Fermi constant. Reusing Fermi's first description of β decay of nuclei or the decay of the muon $\mu \to \nu_\mu e^- \bar{\nu}_e$ as a four-fermion interaction [42, 43], the new gauge interaction can be approximated at energies below Λ as a contact interaction (CI) with an effective coupling proportional to $1/\Lambda^2$ [36, 44]. In case of weak decays, the effective coupling G_F could later be related to the corresponding weak exchange quantum, the W boson, as

$$\frac{G_F}{\sqrt{2}} = \frac{g_w^2}{8M_W^2}\,, \tag{5.11}$$

where g_w denotes the respective weak coupling constant. Similarly, one can add a CI term to the SM Lagrangian like [36, 44]

$$\begin{aligned}\mathcal{L}_{qq} = \frac{g^2}{2\Lambda^2}\Big\{ & \eta_{LL}\left(\bar{q}_L\gamma^\mu q_L\right)\left(\bar{q}_L\gamma_\mu q_L\right) \\ & +2\eta_{LR}\left(\bar{q}_L\gamma^\mu q_L\right)\left(\bar{q}_R\gamma_\mu q_R\right) \\ & +\eta_{RR}\left(\bar{q}_R\gamma^\mu q_R\right)\left(\bar{q}_R\gamma_\mu q_R\right)\Big\}\end{aligned} \tag{5.12}$$

with g the new coupling constant, η_{LL}, η_{LR}, and η_{RR} relative strengths of order unity, and q_L and q_R the left-handed and right-handed chiral components of the quark fields. Conventionally, one defines $g^2/4\pi = 1$ such that the first factor in Eq. (5.12) becomes $2\pi/\Lambda^2$. These terms are flavour-diagonal and avoid any conflict with constraints from non-observation of flavour-changing neutral currents.

In contrast to QCD, where the predominant gluon-exchange channel leads to a cross section that peaks at large pseudorapidity η, cf. Section 6.2, contact interactions produce a more isotropic angular distribution. As a consequence, the ratio of the dijet cross section at outer to the one at inner pseudorapidity, the centrality ratio R_η, is sensitive to an admixture of CI to normal QCD reactions. Depending on the sign of e.g. η_{LL}, the interference between the CI and QCD amplitudes can be constructive ($\eta_{LL} = -1$) or destructive ($\eta_{LL} = +1$) with corresponding compositeness scales Λ^- and Λ^+. The more conservative limit is given by Λ^+.

An investigation of R_η by the D0 Collaboration for both interference modes of the LL term lead to a lower limit of $\Lambda^{\pm} > 2.4\,\text{TeV}$ at 95% confidence level (CL) [45]. CMS repeated this type of analysis, which is typically performed at the start of an experiment because of its reduced sensitivity to an imperfect detector understanding, and defined R_η as the cross section ratio for dijet production with both jets leading in p_T with $|\eta| < 0.7$ over the one with both jets having $0.7 \leq |\eta| < 1.3$ [46]. R_η is measured as a function of the dijet mass so that increasing deviations from QCD predictions are expected when masses approach a Λ-dependent value. A similar study by ATLAS is contained in Ref. [47].

The jets, reconstructed from calorimetric energy depositions alone, are calibrated and used in comparisons to particle-level predictions. Simulation studies show that the impact of detector effects on the dijet mass and on R_η are limited such that for the purpose of shape studies they can be accounted for in an uncertainty.

To estimate a potential lower limit on the compositeness scale from an agreement between data and theory in the centrality ratio R_η, a log-likelihood-ratio statistic, R_LL, is employed. The null hypothesis is represented by the NLO prediction including NP effects, where a constant offset with respect to data is allowed for dijet masses of 490–790 GeV. This limited range avoids the region with NP corrections up to 10% at low M_jj and the potential signal region at high masses. The offset is determined to be -0.05 and is compatible within uncertainties with the SM expectation. Figure 5.10 left shows the measurement from $2.9\,\text{pb}^{-1}$ of integrated luminosity at 7 TeV centre-of-mass energy in comparison to the null hypothesis. No significant discrepancy from the QCD expectation of $R_\eta \approx 0.5$, independent of dijet mass, is observed.

For the signal hypothesis PYTHIA6 is used, which includes a LO version of the CI model. Only the LL term with the more conservative destructive interference is considered. The QCD part of the LO+PS prediction by PYTHIA is corrected for NLO effects, but not the CI part, for which NLO was not available at the time.

The uncertainties entering the limit determination are dominated by the relative JEC between the inner and outer jet pseudorapidities leading to 5–13% of uncertainty on R_η. The theoretical systematic uncertainties associated with the scale and PDF choice are reduced in the ratio to about 3–5% of scale and less than 1% of PDF uncertainty. Larger theoretical uncertainties of about 6% and 4% are caused by the differences in the QCD description between PYTHIA6 and NLO and by the statistical precision of the offset determination between data and NLO.

The total likelihood entering R_LL is the product of individual likelihoods for each M_jj bin. Because the CI model approximation is not valid any more for M_jj very close to the compositeness scale, a Λ dependent upper limit on M_jj is imposed. With this setup, the CL's for signal plus background, CL_{s+b}, and for background alone, CL_b, can be determined. CL_{s+b} and CL_b represent the probabilities for R_LL to have a lower value than that observed, given the alternative and null hypotheses. To protect against setting limits, where the experiment is not sensitive any more, the modified frequentist CL_s method is applied [48–50]. R_LL values, for which $\text{CL}_s = \text{CL}_{s+b}/\text{CL}_b < 0.05$, are then excluded at 95% confidence level. This is shown in

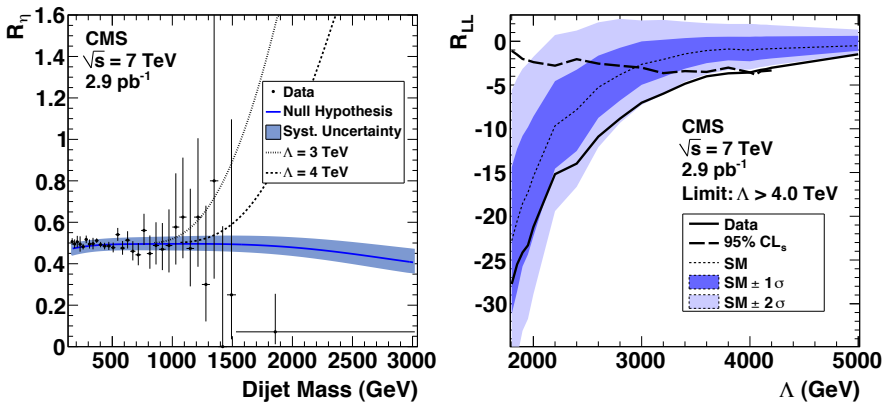

Fig. 5.10 Left: Measured centrality ratio R_η and total uncertainty (points with error bars) as a function of dijet mass in comparison to the QCD prediction at NLO complemented with NP corrections (solid line). The band represents the total theoretical uncertainty including the estimation of an offset allowed between data and theory in the dijet mass range of 490–790 GeV. Predictions comprising the effects of destructive CI interference are shown as dotted and dashed lines for compositeness scales Λ of 3 and 4 TeV. Right: Log-likelihood-ratio R_{LL} versus compositeness scale Λ for data (solid line) and the SM expectation (dotted line) with 1σ and 2σ bands. The dashed line represents the 95% CL_s indicating an exclusion limit of 4 TeV. $\Lambda = \Lambda^+$ here. *(Taken from Ref. [46])*

Fig. 5.10 right, from which a new exclusion limit of 4 TeV on the compositeness scale Λ^+ can be concluded.

References

1. Z. Nagy, "Three jet cross-sections in hadron hadron collisions at next-to-leading order", *Phys. Rev. Lett.* **88** (2002) 122003, doi:10.1103/PhysRevLett.88.122003, arXiv:hep-ph/0110315.
2. Z. Nagy, "Next-to-leading order calculation of three-jet observables in hadron hadron collisions", *Phys. Rev. D* **68** (2003) 094002, doi:10.1103/PhysRevD.68.094002, arXiv:hep-ph/0307268.
3. Z. Bern et al., "Four-Jet Production at the Large Hadron Collider at Next-to-Leading Order in QCD", *Phys. Rev. Lett.* **109** (2012) 042001, doi:10.1103/PhysRevLett.109.042001, arXiv:1112.3940.
4. S. Badger, B. Biedermann, P. Uwer, and V. Yundin, "NLO QCD corrections to multi-jet production at the LHC with a centre-of-mass energy of $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **718** (2013) 965, doi:10.1016/j.physletb.2012.11.029, arXiv:1209.0098.
5. S. Badger, B. Biedermann, P. Uwer, and V. Yundin, "Next-to-leading order QCD corrections to five jet production at the LHC", *Phys. Rev. D* **89** (2014) 034019, doi:10.1103/PhysRevD.89.034019, arXiv:1309.6585.
6. UA1 Collaboration, "Comparison of Three Jet and Two Jet Cross-Sections in p anti-p Collisions at the CERN SPS p anti-p Collider", *Phys. Lett. B* **158** (1985) 494,

doi:10.1016/0370-2693(85)90801-9.
7. UA2 Collaboration, "Measurement of the strong coupling constant α_s from a study of W bosons produced in association with jets", *Phys. Lett. B* **215** (1988) 175, doi:10.1016/0370-2693(88)91093-3.
8. CMS Collaboration, "Measurement of the ratio of the inclusive 3-jet cross section to the inclusive 2-jet cross section in pp collisions at $\sqrt{s}$ = 7 TeV and first determination of the strong coupling constant in the TeV range", *Eur. Phys. J. C* **73** (2013) 2604, doi:10.1140/epjc/s10052-013-2604-6, arXiv:1304.7498.
9. ATLAS Collaboration, "Measurement of multi-jet cross-section ratios and determination of the strong coupling constant in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", technical report, CERN, 2013.
10. D0 Collaboration, "Measurement of angular correlations of jets at $\sqrt{s} = 1.96$ TeV and determination of the strong coupling at high momentum transfers", *Phys. Lett. B* **718** (2012) 56, doi:10.1016/j.physletb.2012.10.003, arXiv:1207.4957.
11. M. Wobisch et al., "A new quantity for studies of dijet azimuthal decorrelations", *JHEP* **01** (2013) 172, doi:10.1007/JHEP01(2013)172, arXiv:1211.6773.
12. CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
13. M. Diehl, "Theory Uncertainties", p. 297. Wiley-VCH Verlag GmbH & Co. KGaA, 2013. doi:10.1002/9783527653416.ch9.
14. D. Becciolini et al., "Constraining new colored matter from the ratio of 3 to 2 jets cross sections at the LHC", *Phys. Rev. D* **91** (2015) 015010, doi:10.1103/PhysRevD.91.015010, arXiv:1403.7411.
15. R. D. Ball et al., "Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology", *Nucl. Phys. B* **849** (2011) 296, doi:10.1016/j.nuclphysb.2011.03.021, arXiv:1101.1300.
16. K. A. Olive and others (Particle Data Group), "Review of Particle Physics", *Chin. Phys. C* **38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.
17. ATLAS Collaboration, "Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector", *Phys. Rev. D* **86** (2012) 014022, doi:10.1103/PhysRevD.86.014022, arXiv:1112.6297.
18. CMS Collaboration, "Measurement of the Inclusive Jet Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **107** (2011) 132001, doi:10.1103/PhysRevLett.107.132001, arXiv:1106.0208.
19. CMS Collaboration, "Measurements of differential jet cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the CMS detector", *Phys. Rev. D* **87** (2013) 112002, doi:10.1103/PhysRevD.87.112002, arXiv:1212.6660.
20. M. Dasgupta, L. Magnea, and G. P. Salam, "Non-perturbative QCD effects in jets at hadron colliders", *JHEP* **02** (2008) 055, doi:10.1088/1126-6708/2008/02/055, arXiv:0712.3014.
21. M. Cacciari, J. Rojo, G. P. Salam, and G. Soyez, "Quantifying the performance of jet definitions for kinematic reconstruction at the LHC", *JHEP* **12** (2008) 032, doi:10.1088/1126-6708/2008/12/032, arXiv:0810.1304.
22. G. Soyez, "A Simple description of jet cross-section ratios", *Phys. Lett. B* **698** (2011) 59, doi:10.1016/j.physletb.2011.02.061, arXiv:1101.2665.
23. ZEUS Collaboration, "Inclusive-jet cross sections in NC DIS at HERA and a comparison of the kT, anti-kT and SIScone jet algorithms", *Phys. Lett. B* **691** (2010) 127, doi:10.1016/j.physletb.2010.06.015, arXiv:1003.2923.
24. ALICE Collaboration, "Measurement of the inclusive differential jet cross section in pp collisions at $\sqrt{s} = 2.76$ TeV", *Phys. Lett. B* **722** (2013) 262, doi:10.1016/j.physletb.2013.04.026, arXiv:1301.3475.
25. CMS Collaboration, "Measurement of the ratio of inclusive jet cross sections using the anti-k_T algorithm with radius parameters $R = 0.5$ and 0.7 in pp collisions at $\sqrt{s}$ = 7 TeV",

Phys. Rev. D **90** (2014) 072006, doi:10.1103/PhysRevD.90.072006, arXiv:1406.0324.
26. M. L. Mangano and J. Rojo, "Cross Section Ratios between different CM energies at the LHC: opportunities for precision measurements and BSM sensitivity", *JHEP* **08** (2012) 010, doi:10.1007/JHEP08(2012)010, arXiv:1206.3557.
27. UA2 Collaboration, "Measurement of the $\sqrt{(s)}$ Dependence of Jet Production at the CERN anti-p p Collider", *Phys. Lett. B* **160** (1985) 349, doi:10.1016/0370-2693(85)91341-3.
28. UA1 Collaboration, "Measurement of the Inclusive Jet Cross-Section at the CERN p anti-p Collider", *Phys. Lett. B* **172** (1986) 461, doi:10.1016/0370-2693(86)90290-X.
29. CDF Collaboration, "Comparison of jet production in $\bar{p}p$ collisions at $\sqrt{s} = 546$ GeV and 1800 GeV", *Phys. Rev. Lett.* **70** (1993) 1376, doi:10.1103/PhysRevLett.70.1376.
30. D0 Collaboration, "High-p_T jets in $\bar{p}p$ collisions at $\sqrt{s} = 630$ GeV and 1800 GeV", *Phys. Rev. D* **64** (2001) 032003, doi:10.1103/PhysRevD.64.032003, arXiv:hep-ex/0012046.
31. J. D. Bjorken, "Can We Measure Parton Parton Cross-Sections?", *Phys. Rev. D* **8** (1973) 4098, doi:10.1103/PhysRevD.8.4098.
32. J. D. Bjorken and E. A. Paschos, "Inelastic Electron Proton and gamma Proton Scattering, and the Structure of the Nucleon", *Phys. Rev.* **185** (1969) 1975, doi:10.1103/PhysRev.185.1975.
33. R. P. Feynman, "Very high-energy collisions of hadrons", *Phys. Rev. Lett.* **23** (1969) 1415, doi:10.1103/PhysRevLett.23.1415.
34. ATLAS Collaboration, "Measurement of the inclusive jet cross section in pp collisions at $\sqrt{s} = 2.76$ TeV and comparison to the inclusive jet cross section at $\sqrt{s} = 7$ TeV using the ATLAS detector", *Eur. Phys. J. C* **73** (2013) 2509, doi:10.1140/epjc/s10052-013-2509-4, arXiv:1304.4739.
35. H. Terazawa, "Subquark Model of Leptons and Quarks", *Phys. Rev. D* **22** (1980) 184, doi:10.1103/PhysRevD.22.184.
36. E. Eichten, K. D. Lane, and M. E. Peskin, "New Tests for Quark and Lepton Substructure", *Phys. Rev. Lett.* **50** (1983) 811, doi:10.1103/PhysRevLett.50.811.
37. U. Baur, I. Hinchliffe, and D. Zeppenfeld, "Excited Quark Production at Hadron Colliders", *Int. J. Mod. Phys. A* **2** (1987) 1285, doi:10.1142/S0217751X87000661.
38. J. L. Hewett and T. G. Rizzo, "Low-Energy Phenomenology of Superstring Inspired E(6) Models", *Phys. Rept.* **183** (1989) 193, doi:10.1016/0370-1573(89)90071-9.
39. P. H. Frampton and S. L. Glashow, "Chiral Color: An Alternative to the Standard Model", *Phys. Lett. B* **190** (1987) 157, doi:10.1016/0370-2693(87)90859-8.
40. E. H. Simmons, "Coloron phenomenology", *Phys. Rev. D* **55** (1997) 1678, doi:10.1103/PhysRevD.55.1678, arXiv:hep-ph/9608269.
41. L. Randall and R. Sundrum, "A Large mass hierarchy from a small extra dimension", *Phys. Rev. Lett.* **83** (1999) 3370, doi:10.1103/PhysRevLett.83.3370, arXiv:hep-ph/9905221.
42. E. Fermi, "Versuch einer Theorie der β-Strahlen. I).", *Z. Phys.* **88** (1934) 161, doi:10.1007/BF01351864.
43. E. Fermi, "Tentativo di una Teoria dei Raggi β", *Nuovo Cim.* **11** (1934) 1, doi:10.1007/BF02959820.
44. E. Eichten, I. Hinchliffe, K. D. Lane, and C. Quigg, "Super Collider Physics", *Rev. Mod. Phys.* **56** (1984) 579, doi:10.1103/RevModPhys.56.579.
45. D0 Collaboration, "The dijet mass spectrum and a search for quark compositeness in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **82** (1999) 2457, doi:10.1103/PhysRevLett.82.2457, arXiv:hep-ex/9807014.
46. CMS Collaboration, "Search for Quark Compositeness with the Dijet Centrality Ratio in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **105** (2010) 262001, doi:10.1103/PhysRevLett.105.262001, arXiv:1010.4439.

47. ATLAS Collaboration, "Search for Quark Contact Interactions in Dijet Angular Distributions in pp Collisions at $\sqrt{s} = 7$ TeV Measured with the ATLAS Detector", *Phys. Lett. B* **694** (2011) 327, doi:10.1016/j.physletb.2010.10.021, arXiv:1009.5069.
48. R. D. Cousins and V. L. Highland, "Incorporating systematic uncertainties into an upper limit", *Nucl. Instrum. Meth. A* **320** (1992) 331, doi:10.1016/0168-9002(92)90794-5.
49. T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl.Instrum.Meth.* **A434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.
50. A. L. Read, "Presentation of search results: The CL(s) technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

Chapter 6
Normalised Cross Sections

A second possibility to reduce uncertainties caused by systematic effects consists in performing shape comparisons only. In this case the measured and the predicted distributions are normalised to the integral over the whole or a part of the investigated phase space. Luminosity uncertainties are irrelevant in this case and JEC, JER, QCD scale, or PDF effects cancel at least partially.

In the following, these advantages of normalised distributions are exploited to search for new phenomena with inclusive jet and dijet events in Sections 6.1 and 6.2, to investigate colour coherence in 3-jet events in Section 6.3, to look for double parton scattering in 4-jet events in Section 6.4, to study multi-jet production through azimuthal decorrelation and event shapes in Sections 6.5 and 6.6, and to examine jet substructure in Section 6.7.

6.1 Search for New Phenomena with Inclusive Jets

As a first example the inclusive jet p_{T} spectrum is examined for potential hints on CI as introduced in Section 5.4. Like with the R_η observable, gradually aggravating deviations from QCD are expected with increasing jet p_{T}. Figure 6.1 illustrates for various compositeness scales Λ the effect of CI versus p_{T} for constructive (left) and destructive interference (right) in the case of left-handed quarks. For the same Λ it is simpler to discover or exclude indications of CI in the case of constructive as compared to destructive interference.

Interestingly, such a study of the inclusive jet p_{T} spectrum performed by CDF at the Tevatron [2] in 1996 did, in contrast to their previous result of $\Lambda_{LL} > 1.4\,\mathrm{TeV}$ at 95%CL [3], not deliver an exclusion limit but a preferred CI scale of $\Lambda_{LL} = 1.6\,\mathrm{TeV}$. Profiting from the fact that deviations are small at low p_{T} and gradually build up with higher p_{T}, both analyses normalised the p_{T} spectrum at NLO QCD to the measurement below 160 GeV. This way a significant source of uncertainty, the relative normalisation of theory to data, is largely eliminated and only the shapes of the distributions are compared. However, systematic effects that potentially exhibit

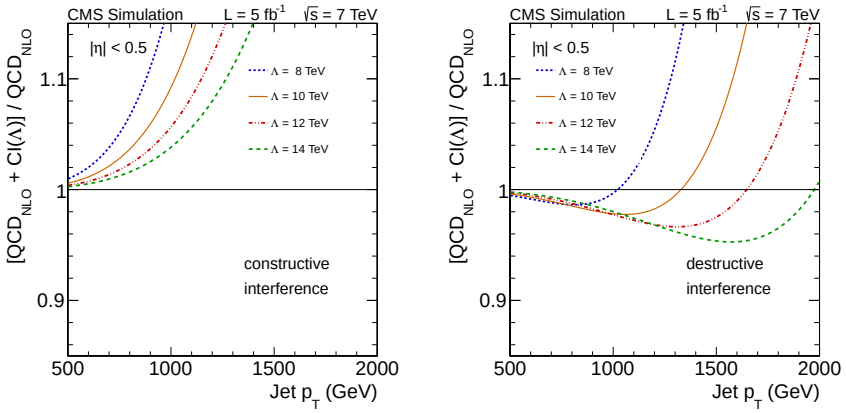

Fig. 6.1 Effect of CI on the inclusive jet p_{T} spectrum for various compositeness scales Λ. The left plot corresponds to constructive ($\eta_{LL} = -1$) and the right one to destructive ($\eta_{LL} = +1$) interference between the QCD and the CI amplitudes. *(Taken from Ref. [1])*

a shape versus p_{T} must carefully be estimated. In the case of the CDF result, the excess could finally be explained by an insufficient knowledge of the gluon PDF and lead to the initiative of improving the estimation of PDF uncertainties as described in Section 2.6.2. A similar measurement by the D0 Collaboration but employing the scalar sum of the jet transverse energies H_{T} instead profited from improved PDFs and did not observe any excess [4].

A CMS study [1] analysed the equivalent of $5.0\,\mathrm{fb}^{-1}$ of data at $\sqrt{s} = 7\,\mathrm{TeV}$. The inclusive jet p_{T} spectrum is measured and calibrated, but not unfolded, for jet $p_{\mathrm{T}} > 500\,\mathrm{GeV}$ at central pseudorapidity $|\eta| < 0.5$. QCD theory is derived at NLO; NP corrections are considered negligible. The signal term $\mathrm{CI}(\Lambda)$ is modelled using the CI implementation in PYTHIA6 by subtracting the LO QCD from the combined prediction of $\mathrm{QCD}_{\mathrm{LO}} + \mathrm{CI}(\Lambda)$. In the absence of any signal, cf. Fig. 6.2 left, exclusion limits can be derived. Since the data have not been corrected for detector effects, however, the theory and its uncertainties must be convolved with the jet response of the CMS detector. Therefore, a parameterised ansatz is used for the Λ-dependent predictions of $(\mathrm{QCD}_{\mathrm{NLO}} + \mathrm{CI}(\Lambda))/\mathrm{QCD}_{\mathrm{NLO}}$. Exploiting the linear dependence on $1/\Lambda^2$, the relative change in cross section can be written as a sum of three terms: unity for the $\mathrm{QCD}_{\mathrm{NLO}}$ prediction, a power-law p_{T} dependent interference term proportional to $1/\Lambda^2$ with two parameters for the relative normalisation and the exponent, and a power-law p_{T} dependent CI term proportional to $(1/\Lambda^2)^2$ again with two parameters for relative normalisation and exponent. Using separate fits of this ansatz for the cases of constructive and destructive interference between QCD and CI amplitudes, theoretical uncertainties from the choice of μ_r, μ_f, and the PDFs as well as the experimental uncertainties from JEC and JER are propagated to the detector level including all correlations. The obtained uncertainty bands for the $\mathrm{QCD}_{\mathrm{NLO}}$ case alone are shown in Fig. 6.2 left demonstrating agreement between

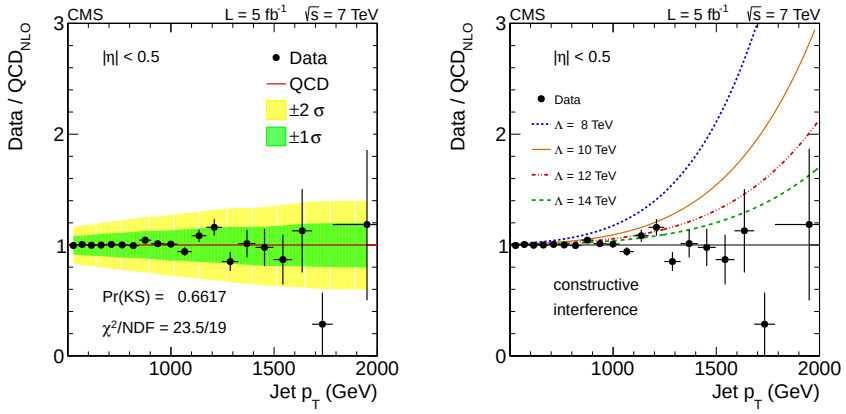

Fig. 6.2 Ratio of data over the $\mathrm{QCD}_{\mathrm{NLO}}$ prediction, convolved with the jet response of the CMS detector, as a function of jet p_{T}. Error bars represent the statistical uncertainty of the data. The shaded areas in the left plot correspond to 1σ and 2σ uncertainty bands of the prediction comprising systematic effects of theoretical (μ_r, μ_f, PDF) and experimental origin (JEC, JER). On the right hand side the same ratio is shown together with the convolved predictions for various assumptions on the compositeness scale Λ for the case of constructive interference ($\eta_{LL} = -1$). *(Taken from Ref. [1])*

data and SM theory. Figure 6.2 right contrasts the measured ratio with expectations assuming a series of four values for the compositeness scale Λ.

As in the Tevatron studies, exclusion limits are very sensitive to the relative normalisation between data and theory, which previously was fixed at low p_{T} below the search region. This is not done in the CMS analysis at hand. Instead the relative normalisation enters as one of the nuisance parameters into a likelihood analysis of the multinomial distribution of event counts versus jet p_{T} bin. Testing the correlated ensembles of convolved predictions with the measurement, lower limits on the compositeness scale of $\Lambda^+ > 9.9\,\mathrm{TeV}$ and $\Lambda^- > 14.3\,\mathrm{TeV}$ are achieved at 95% CL applying the CL_s criterion. Bayesian limits with a reference prior are obtained to be $\Lambda^+ > 10.1\,\mathrm{TeV}$ and $\Lambda^- > 14.1\,\mathrm{TeV}$.

A very similar CMS study [5] using $10.7\,\mathrm{fb}^{-1}$ of data at 8 TeV centre-of-mass energy searches for hints of jet extinction, i.e. a lack of high-p_{T} jet production, as predicted by some models of non-perturbative quantum gravity [6]. Figure 6.3 presents a comparison of the measured inclusive jet p_{T} spectrum to predictions of $\mathrm{QCD}_{\mathrm{NLO}}$ and of three jet extinction scenarios each multiplied by NP corrections and convolved with the jet response of the CMS detector. No significant deviation from the SM prediction becomes apparent. Hence, employing the same general strategy as described above, a lower exclusion limit of 3.3 TeV at 95% CL is set on the extinction mass scale using the CL_s criterion.

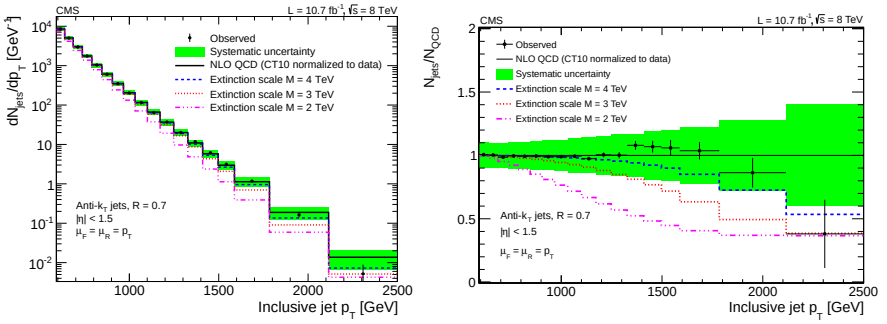

Fig. 6.3 Measured inclusive jet p_{T} spectrum in comparison to predictions of $\mathrm{QCD}_{\mathrm{NLO}}$ and of three jet extinction scenarios each multiplied by NP corrections and convolved with the jet response of the CMS detector. The shaded areas correspond to a 1σ uncertainty band of the prediction comprising systematic effects of theoretical (μ_r, μ_f, PDF) and experimental origin (JEC, JER). The $\mathrm{QCD}_{\mathrm{NLO}}$ prediction has been normalised to the data. The left plot presents the distribution of the number of jets, while the right plot present the ratio of data over prediction. *(Taken from Ref. [5])*

6.2 Search for New Phenomena with Dijet Angular Distributions

Arguing that QCD interactions predominantly are mediated via gluon exchange, the centrality ratio R_η was introduced in Section 5.4 as an observable to differentiate between QCD and CI processes. For substantiation Fig. 6.4 presents the s, t, and u channel graphs characterising a $2 \to 2$ reaction, where s, t, and u are the usual *Mandelstam* invariants. The production of new resonances can only happen via the s channel, while Rutherford scattering for example is mediated via the t-channel exchange of a massless vector particle, the photon. Specialising to a parton-parton scattering process with $\hat{s} = sx_1x_2$, the corresponding quantities $\hat{t}$ and $\hat{u}$ can be written in terms of the scattering angle θ^* in the partonic centre-of-mass system as

$$\hat{t} = -\hat{s}\frac{1-\cos\theta^*}{2} \quad , \quad \hat{u} = -\hat{s}\frac{1+\cos\theta^*}{2} \, . \tag{6.1}$$

As explained in Section 4.3, the numbering of the two final-state partons (or jets) in such a $2 \to 2$ process is arbitrary for massless partons that are not differentiated with respect to their flavour. An exchange of p_3 with p_4 and vice versa merely interchanges $\hat{t}$ with $\hat{u}$ and redefines θ^* with respect to the other parton.

Then it is interesting to have a closer look at the LO QCD $2 \to 2$ partonic processes, which have first been calculated in Ref. [7]. At a pp collider the dominant subprocesses in most of the phase space are $gg \to gg$, $gq \to gq$, and $q_iq_j \to q_iq_j$ ($i \neq j$), cf. Fig. 6.5. Taking into account the $\hat{t} \leftrightarrow \hat{u}$ symmetry, these subprocesses behave to a good approximation like Rutherford scattering with massless vector gluons as exchange particles. Moreover, they scale roughly proportional to the colour charges of the colliding partons, $C_A^2 : C_AC_F : C_F^2$. Hence, relative to $gg \to gg$ the other two subprocesses contribute proportional to $C_F/C_A = 4/9$ and $(C_F/C_A)^2$ [8, 9]. At

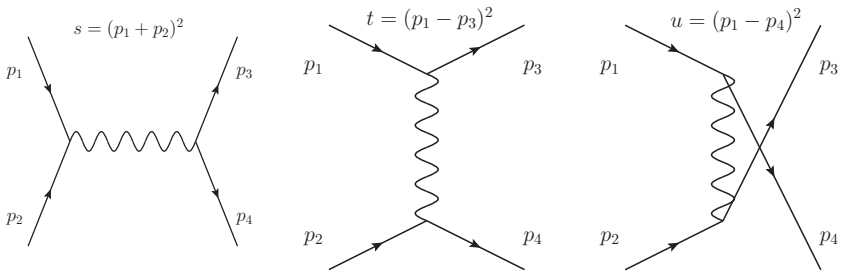

Fig. 6.4 s, t, and u channel graphs characterising a $2 \to 2$ reaction, where s, t, and u are the usual *Mandelstam* variables as defined in the plots.

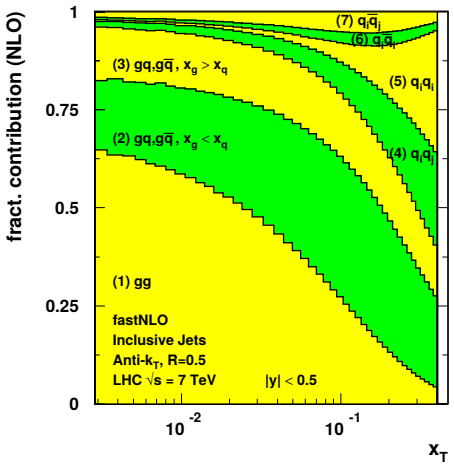

Fig. 6.5 Partonic subprocess decomposition at $\sqrt{s} = 7\,\text{TeV}$ for inclusive jets at the LHC as a function of $x_\text{T} = 2p_\text{T}/\sqrt{s}$.

times when data to determine the proton structure were still sparse, these estimations were exploited in the so-called *single effective subprocess approximation* [10], where only the kinematic shape of the gg channel was retained, but properly rescaled to represent the total cross section. Summarising, QCD behaves to a good approximation similarly to Rutherford scattering, while new phenomena are expected to exhibit a more isotropic angular distribution.

Any difference in angular distribution becomes more apparent when studied in the partonic centre-of-mass instead of the laboratory frame of reference. Rutherford scattering for example reads

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\cos\theta^*} \propto \frac{1}{\sin^4(\theta^*/2)}\,, \tag{6.2}$$

where θ^* is the polar scattering angle in the partonic centre-of-mass frame. To study deviations it is convenient to define the dijet angular variable χ as

$$\chi = \frac{\hat{u}}{\hat{t}} = \frac{1+\cos\theta^*}{1-\cos\theta^*} = \exp(2y^*)\,, \tag{6.3}$$

where for the last transformation Formula (4.8) with $y^* = |y_1 - y_2|$ was used to establish a simple relation to the rapidities of the two jets in the laboratory system. Replacing $\cos\theta^*$ in Eq. (6.2) by χ one obtains

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\chi} \propto \text{constant}\,. \tag{6.4}$$

Interestingly, scalar gluons can be excluded through a measurement of the χ distribution as well, since [8]

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\cos\theta^*} \propto \text{constant} \quad \rightarrow \quad \frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\chi} \propto \frac{1}{(1+\chi)^2}\,. \tag{6.5}$$

ATLAS and CMS, both have repeatedly searched for deviations from QCD expectations in the shape of the dijet angular distributions as a function of χ, unfolded for detector effects and normalised to unit area [11–16]. In addition to standard selection criteria to ensure a good data quality and full efficiency, the following conditions are imposed:

1. the rapidities of the two leading p_T jets are delimited to $|y_1, y_2| < 2.8$ (ATLAS) and $|y_1, y_2| < 2.5$ (CMS),
2. the boost of the dijet system $|y_\mathrm{b}|$ must be smaller than 1.10 (ATLAS) respectively 1.11 (CMS),
3. the maximal allowed value of the χ variable is 30 (ATLAS) and 16 (CMS), which effectively restricts y^* to be smaller than 1.70 and 1.39, respectively.

Depending on the centre-of-mass energy and the size of the investigated data samples, the measurement is performed for a series of bins in dijet mass M_jj, where, lacking any sign of new phenomena, the highest accessible mass ranges provide the most stringent lower exclusion limits on the CI scale. Figure 6.6 presents the final χ distributions from ATLAS at 8 TeV centre-of-mass energy in comparison to SM predictions and to destructive ($\Lambda^+_{LL} = 8\,\mathrm{TeV}$) and constructive CI terms ($\Lambda^-_{LL} = 12\,\mathrm{TeV}$) in the highest mass bin.

Significant refinements that have been introduced in the course of time are the inclusion of CI effects at NLO [17, 18], and of EW corrections [19] into the analyses [13, 15, 16]. The CMS measurement [15] found first indications that EW corrections, shown in Fig. 6.7 left and also included by ATLAS in Fig. 6.6, lead to an improved agreement between data and SM predictions and hence to higher exclusion limits on the CI scales.

CMS also derives limits for other chiral projections of the quark fields of the Lagrangian Eq. (5.12) and investigates the following CI scenarios with colour-singlet couplings between quarks:

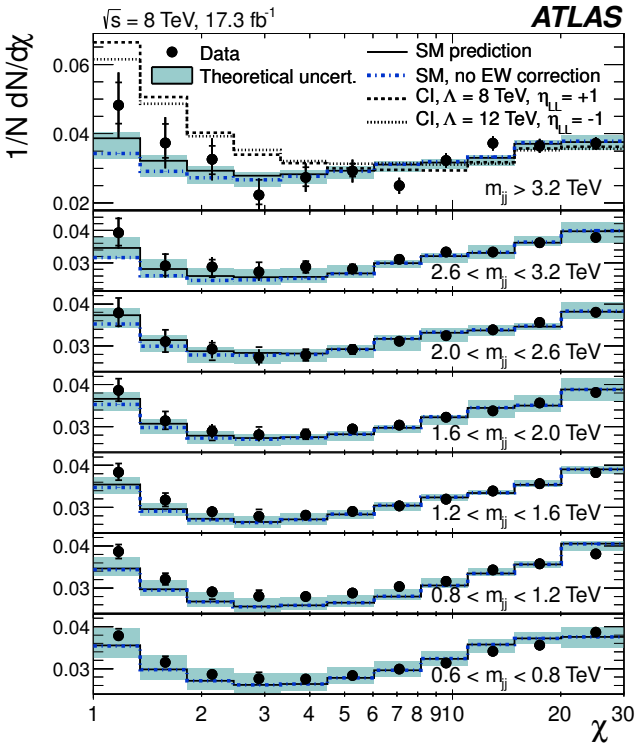

Fig. 6.6 Comparison of dijet angular distributions as measured at $\sqrt{s} = 8\,\text{TeV}$ by ATLAS in various dijet mass regions m_{jj} to predictions of the SM with and without EW corrections (full and dash-dotted lines), and to the *LL* benchmark scenario for destructive (dashed, $\Lambda^{+} = 8\,\text{TeV}$) and constructive (dotted line, $\Lambda^{-} = 12\,\text{TeV}$) interference with the SM process. Error bars represent the quadratically added experimental and statistical uncertainties, where a tick mark indicates the experimental uncertainties alone. Theoretical uncertainties are displayed as shaded bands around the SM predictions. *(Taken from Ref. [16])*

Λ	$(\eta_{LL}, \eta_{RR}, \eta_{RL})$
$\Lambda^{\pm}_{LL}$	$(\pm 1, \quad 0, \quad 0)$
$\Lambda^{\pm}_{RR}$	$(\quad 0, \pm 1, \quad 0)$
$\Lambda^{\pm}_{VV}$	$(\pm 1, \pm 1, \pm 1)$
$\Lambda^{\pm}_{AA}$	$(\pm 1, \pm 1, \mp 1)$
$\Lambda^{\pm}_{(V-A)}$	$(\quad 0, \quad 0, \pm 1)$

Here, it should be noted that for pp collisions the $\Lambda^{\pm}_{LL}$ and $\Lambda^{\pm}_{RR}$ models yield identical tree-level cross sections and NLO corrections, whereas in the cases of $\Lambda^{\pm}_{VV}$, $\Lambda^{\pm}_{AA}$, and $\Lambda^{\pm}_{(V-A)}$, the CI predictions are identical at tree-level, but exhibit different NLO corrections. The expected and observed limits at 95% CL from data taken at $\sqrt{s} = 8\,\text{TeV}$ are shown in Fig. 6.7 right for NLO CI, and for other models with ex-

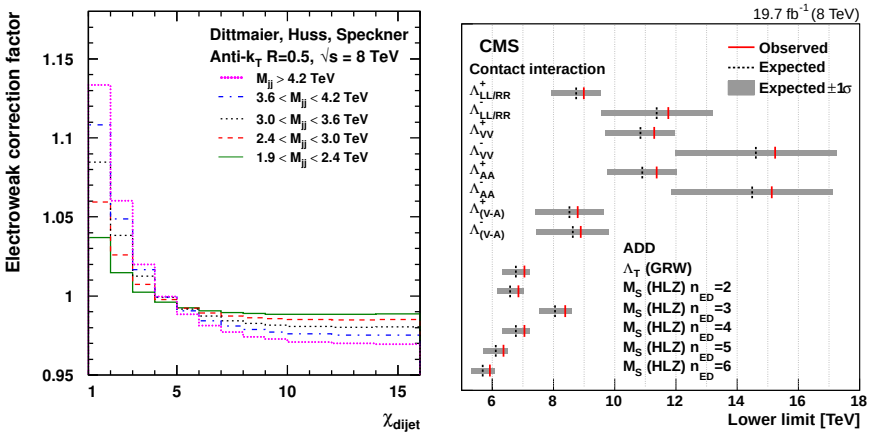

Fig. 6.7 Left: EW corrections to the dijet angular distributions for various dijet mass bins. Right: Observed (solid) and expected lower exclusion limits (dashed lines) at 95% CL on the CI scale Λ at NLO for various combinations of chiral quark field projections. Limits for models with extra spatial dimensions (ADD) are shown as well, cf. Ref. [15] for details. Grey bands indicate the uncertainties in the expected exclusion limits. *(Taken from Ref. [15])*

tra spatial dimensions (ADD), cf. Ref. [15] for details. The limit setting strategy of CMS is unchanged since the beginning and employs the CL_s criterion. In ATLAS, the methods varied with time, but for the final results at 8 TeV centre-of-mass energy the same one as CMS was applied. Profiting from reduced uncertainties through the normalisation and the insensitivity of the χ variable to PDF variations, lower exclusion limits of $\Lambda_{LL}^{+} > 8.1\,\text{TeV}\,(9.0\,\text{TeV})$ and $\Lambda_{LL}^{-} > 12.0\,\text{TeV}\,(11.7\,\text{TeV})$ for ATLAS (CMS) are achieved. The remaining uncertainties in the most sensitive high-mass region are predominantly caused by scale variations in the theory and by the limited amount of data.

6.3 Colour Coherence from 3-Jet Events

An important characteristic of QCD is the fact that partons produced in a hard interaction continue to interfere with each other via the strong force through the hadronisation phase until only colour-neutral particles are left. In particular, self-interacting gluons make a difference with respect to electromagnetic interactions. Multiple collinear gluon emissions can be dealt with using parton showers, cf. Section 2.3, whose effect to a large extent is neutralised by the application of jet algorithms. The k_t algorithm [20] was even designed to revert the parton splittings. Soft gluon emission, however, happens between colour-connected partons up to rather wide angles. This manifests itself by the relative abundance of soft radiation between colour-connected partons and the suppression of such radiation else-

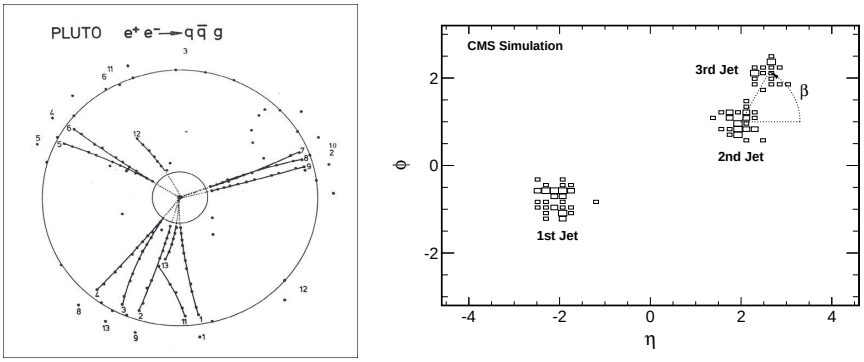

Fig. 6.8 Left: Display of the measured particle tracks of a 3-jet event associated with the gluon bremsstrahlung reaction $e^+e^- \to qqg$. The event has been recorded by the PLUTO Collaboration at DESY. Right: Illustration of the observable β in (η,ϕ) space using a 3-jet event simulated by CMS. The sizes of the rectangular boxes are proportional to the particle energies. *(Taken from Refs. [35–37])*

where. At first glance this phenomenon called *colour coherence* seems to spoil the parton-shower picture of independent subsequent emissions. A solution was found to include colour-coherence effects in the form of angular-ordered emissions, which was first implemented in the HERWIG program [21] and later its successor HERWIG++ [22].

Initially, colour coherence was observed in 3-jet events produced in e^+e^- collisions at the PETRA collider at DESY. Such 3-events, an example by the PLUTO Collaboration is displayed in Fig. 6.8 left, served primarily to establish the existence of the carrier particles of the strong force, the gluon, via the reaction $e^+e^- \to qqg$ [23–26]. Because strongly interacting particles only appear in the final state, effects of colour coherence can be studied as well by analysing the energy or particle flow between the two jets attributed to the $q\bar{q}$-pair on the one hand and between each of these two jets and the third gluonic jet on the other hand. The expected relative suppression of particle production in the region between the quark and antiquark jets has been observed by several experiments at the PETRA [27,28], and later at the PEP [29–32], and LEP colliders [33,34].

In hadron collisions, additional colour connections between the initial- and final-state partons must be taken into account. Both Tevatron experiments have reported evidence for colour coherence effects in measurements of the spatial correlations between neighbouring jets [38,39]. The general strategy is to select events with three jets or more, where two jets stem from the high p_{T} interaction while a third jet with comparably smaller p_{T} is supposed to collect the additional soft radiation. This is in contrast to the previously presented 3-jet analyses, where the emphasis was on high-p_{T} multi-jet production and the same minimal jet p_{T} requirement was applied to all three jets. Ordering the jets in p_{T}, sensitive observables are constructed by studying the separation in pseudorapidity $\Delta\eta_{23} = \mathrm{sign}(\eta_2) \cdot (\eta_3 - \eta_2)$ and azimuthal angle $\Delta\phi_{23} = \phi_3 - \phi_2$ between the second and third jet. CDF and D0 each employ slightly

different variants, namely, $\alpha = \tan^{-1}(\Delta\eta_{23}/|\Delta\phi_{23}|)$ and $\beta = \tan^{-1}(\Delta\phi_{23}/\Delta\eta_{23})$, respectively. CMS [37] adopted, in implicit form, a definition similar to the D0 case

$$\tan\beta = \frac{|\Delta\phi_{23}|}{\Delta\eta_{23}}, \tag{6.6}$$

where $\Delta\phi_{23}$ and β are understood to lie in the ranges $-\pi \leq \Delta\phi_{23} \leq \pi$ and $0 \leq \beta \leq \pi$, respectively. The absolute value of $\Delta\phi_{23}$ in Eq. (6.6) and the sign of the pseudorapidity of the second jet, $\mathrm{sign}(\eta_2)$, in the definition of $\Delta\eta_{23}$ are introduced to map symmetric configurations around $\Delta\phi_{23} = 0$ or $\eta = 0$ onto the same β value. For $\Delta\phi_{23} = 0$, β is defined to be zero or π depending on the sign of $\Delta\eta_{23}$ being positive or negative. In the case of $\Delta\eta_{23} = 0$, which cannot happen simultaneously with $\Delta\phi_{23} = 0$, β is defined to equal $\pi/2$.

In a naive LO model, one of the two partons produced back-to-back in the transverse plane radiates a third parton. In the absence of colour coherence effects there is no preferred direction of emission of this third parton around the radiating parton. In contrast, when colour coherence effects are present, colour connections exist not only between the final-state partons of the high p_T interaction, but also to the initial-state partons and the beam remnants. Therefore, one expects the third jet-initiating parton to preferentially lie in the event plane defined by the emitting parton and the beam axis. This corresponds to configurations, where jet three has the same azimuthal angle as jet two, but a pseudorapidity either between the direction of beam one ($\eta = +\infty$) and η_2 or between beam two ($\eta = -\infty$) and η_2. In case jet three is located between jet two and the closer beam direction, then β becomes zero, otherwise β is equal to π. Out of plane radiation, i.e. $\beta \approx \pi/2$, is suppressed in the presence of colour coherence.

CMS analysed the 2010 data set with an integrated luminosity of $36\,\mathrm{pb}^{-1}$ and selected events, where the two leading jets, reconstructed with the anti-k_t algorithm for a jet size of $R = 0.5$, have pseudorapidities up to 2.5. All jets must have transverse momenta above 30 GeV to be accepted; the p_T of the leading jet must exceed 100 GeV, i.a. for triggering reasons. To ensure a back-to-back topology, the dijet mass M_{12} of the leading two jets is required to be larger than 220 GeV. Finally, the third jet is assured to lie in the vicinity of the second jet by imposing $0.5 < \Delta R_{23} = \sqrt{(\Delta\eta_{23})^2 + (\Delta\phi_{23})^2} < 1.5$, where the lower limit is only a very weak constraint because of the anti-k_t jet size of 0.5. If no differentiation would be made with respect to $\Delta\eta$ and $\Delta\phi$ and only the number of radiated jets would be counted, then this quantity opens another way to determine the strong coupling constant. This has been demonstrated by the D0 Collaboration using an observable named $R_{\Delta R}$ [40].

Because colour coherence effects are expected to become stronger when the angle between the second jet and the remnant beam diminishes, the study of the β variable is performed in two situations: when the second jet is central ($|\eta_2| \leq 0.8$) and when the second jet is more forward ($0.8 < |\eta_2| \leq 2.5$). For these two regions the normalised β distribution is defined as

$$F_{\eta_2,i}(\beta) = \frac{N_{\eta,i}}{N_\eta}\,, \tag{6.7}$$

where $N_{\eta,i}$ is the number of events in a given bin i of the β distribution within an η_2 region, and N_η is the total number of events in that η_2 region. This normalisation significantly reduces the impact of experimental systematic uncertainties and fully eliminates the uncertainty in luminosity. The effect of additional jets through PU collisions was kept at a negligible level by concentrating on the 2010 data set with only two PU events per bunch crossing on average. As a consequence, an additional binning in leading jet p_T could not be applied for reasons of limited amounts of data from the 2010 collider run.

Figure 6.9 presents for both regions in η_2 the CMS measurement in comparison to predictions by the LO MC event generators PYTHIA6, PYTHIA8, HERWIG++, and MADGRAPH + PYTHIA6, all of which by default account for colour coherence effects. Obviously, none of them describes the data satisfactorily. Although it is shown in Ref. [37] for PYTHIA6 that colour coherence favours β values near zero or π, the effect in MC is not strong enough. This is similar for PYTHIA8. MADGRAPH, which includes the LO $2 \to 3$ matrix element, slightly improves the pure parton-shower prediction by PYTHIA6. HERWIG++ finally models the data best but fails in the forward region for β values approaching π. In summary, there is much room for improvement in the description of colour coherence effects by MC event generators.

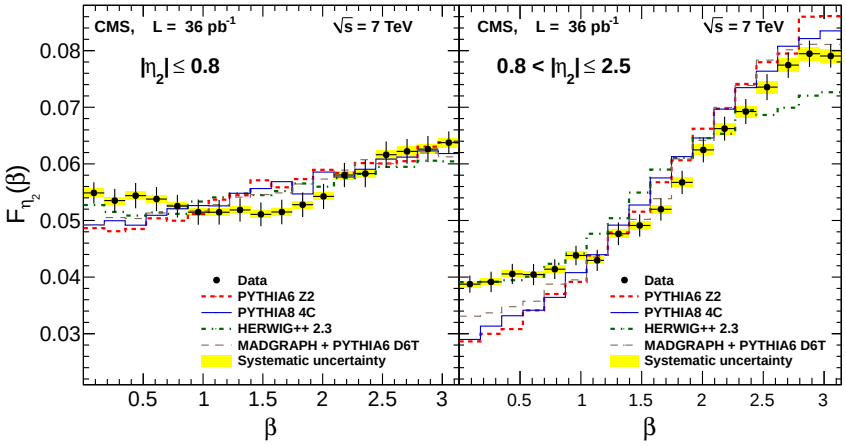

Fig. 6.9 Measured β distributions, corrected for detector effects, in comparison to predictions by the LO MC event generators PYTHIA6, PYTHIA8, HERWIG++, and MADGRAPH + PYTHIA6 in the central ($|\eta_2| \leq 0.8$) and forward ($0.8 < |\eta_2| \leq 2.5$) regions. The error bars show the statistical uncertainties, while the yellow shaded bands correspond to the combined systematic uncertainty. *(Taken from Ref. [37])*

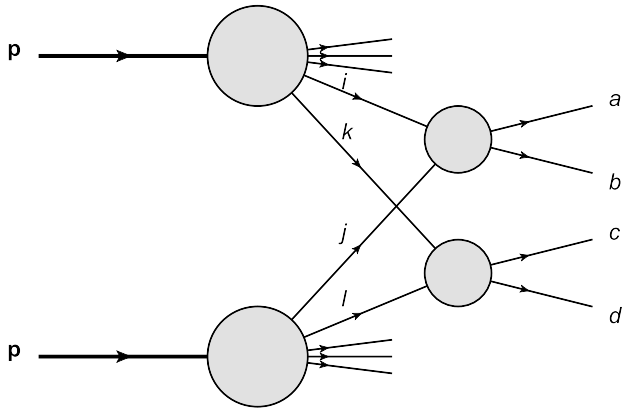

Fig. 6.10 Sketch of two partons i and k from one proton and two partons j and l from a second proton participating in a DPS with the two simultaneous processes $A(ij \to ab)$ and $B(kl \to cd)$.

6.4 Double Parton Scattering in 4-Jet Events

Multiple parton interactions have been introduced in Chapter 2 as a means to model the underlying event [41]. Soft particles are produced in addition to the high-p_{T} process of interest, which is calculated perturbatively following Eq. (6.8) under the assumption that collinear factorisation holds:

$$\mathrm{d}\sigma = \sum_{i,j} \int f_{i/p}(x_1, \mu_f) \cdot f_{j/p}(x_1', \mu_f) \times \mathrm{d}\hat{\sigma}_{(ij\to ab)}(x_1, x_1', \mu_r, \mu_f)\, \mathrm{d}x_1 \mathrm{d}x_1' \,, \qquad (6.8)$$

where i, j are the initial-state parton flavours and $f_{i/p}$, $f_{j/p}$ are the probabilities to find a parton i or j with momentum fraction x_1 respectively x_1' in the proton at factorisation scale μ_f. The parton-level cross section $\mathrm{d}\hat{\sigma}_{(ij\to ab)}$ for process $ij \to ab$ depends on the momentum fractions x_1, x_1' and the factorisation and renormalisation scales μ_f and μ_r. This naive picture of only one interacting parton from each colliding hadron to participate in a high-p_{T} reaction with only longitudinal degrees of freedom has been applied very successfully to an enormous amount of, in particular, inclusive processes. Approaching small transverse momenta though, this picture must fail, because the respective cross sections grow in an unlimited fashion and finally would violate unitarity, i.e. become even larger than the total cross section. Hence, the assumptions need to be modified to regularise the otherwise unlimited growth. The fact that at p_{T} values of a few GeV protons are probed at fractional momenta x of 10^{-3} and below, where the parton densities $f_{i/p}(x, \mu_f)$ increase dramatically, suggests to consider the possibility of two semihard simultaneous parton-parton scatters within the same proton-proton collision. In contrast to the usual *single parton scattering* (SPS) this type of process is called *double parton scattering* (DPS). A sketch of such a reaction is shown in Fig. 6.10.

Accepting DPS as the strategy to stabilise the SPS predictions at low p_{T}, one needs to address the issue of joint probabilities $H_{i,k/p}(x_1,x_2,\mu_{f,A},\mu_{f,B})$ to find a parton i with momentum fraction x_1 and at the same time a second parton k with momentum fraction x_2 within the same proton p at scales $\mu_{f,A}$ and $\mu_{f,B}$ characteristic for the two simultaneous processes $A(ij\to ab)$ and $B(kl\to cd)$, cf. Fig. 6.10. Neglecting again intrinsic transverse momenta $k_{T,i}$ and $k_{T,k}$, and ignoring potentially strong correlations in longitudinal momentum, the poorly known double-parton distribution functions H (DPDFs) are approximated by a product of the usual single PDFs

$$H_{i,k/p}(x_1,x_2,\mu_{f,A},\mu_{f,B}) = f_{i/p}(x_1,\mu_{f,A})\cdot f_{k/p}(x_2,\mu_{f,B})\,. \tag{6.9}$$

Of course, four-momentum conservation $x_1+x_2\leq 1$ is implicitly assumed, so that following Refs. [42–44] the cross section for the DPS process $A+B$ can be written in a simplified form in terms of a process-independent effective cross section σ_{eff}

$$\begin{aligned}\mathrm{d}\sigma^{\mathrm{DPS}}_{A+B} = \frac{m}{2\cdot\sigma_{\mathrm{eff}}}\sum_{i,j,k,l}\int H_{i,k/p}(x_1,x_2,\mu_{f,A},\mu_{f,B})\cdot H_{j,l/p}(x_1',x_2',\mu_{f,A},\mu_{f,B})\times\\ \mathrm{d}\hat{\sigma}^{A}_{ij}(x_1,x_1',\mu_{f,A})\mathrm{d}\hat{\sigma}^{B}_{kl}(x_2,x_2',\mu_{f,B})\,\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_1'\mathrm{d}x_2'\,,\end{aligned} \tag{6.10}$$

where m is a symmetry factor that is equal to unity for indistinguishable final states and two otherwise. Applying Eq. (6.9), integrating over the selected analysis phase space, and solving the equation for σ_{eff} one obtains

$$\sigma_{\mathrm{eff}} = \frac{m}{2}\frac{\sigma_{\mathrm{A}}\cdot\sigma_{\mathrm{B}}}{\sigma^{\mathrm{DPS}}_{\mathrm{A+B}}}\,, \tag{6.11}$$

where σ_{A} and σ_{B} are the cross sections for the independent processes A and B. In this simple picture, σ_{eff} is a measure of the overlap in transverse size of the parton distributions in the colliding hadrons. A naive geometrical interpretation leads to an estimate of $\sigma_{\mathrm{eff}}\approx\pi R_{\mathrm{p}}^2\approx 50$ mb, where R_{p} is the proton radius. The validity of the presented approach relies on a couple of assumptions. For a recent critical overview please cf. Ref. [45].

To maximise the potential production rate of DPS events, one of the two processes A and B is usually identified with inclusive dijet production, i.e. $pp\to jj+X$. For the second process a number of possibilities have been investigated. The largest production rate is expected for a second jet pair leading to 4-jet production, which has to be differentiated kinematically from higher-order SPS reactions with $4+X$ jets. This will be studied closer further below; relevant measurements are found in Refs. [46–49]. A disambiguation of the final state in terms of pairs of high-p_{T} objects can be achieved by requiring two out of the four jets to be b-tagged as proposed in [44]. Alternatively, processes with smaller cross sections but distinguishable final state are chosen for the second reaction, e.g. inclusive $\gamma+\mathrm{jet}+X$ or $W(\to l\nu)+X$ production leading to final states of $\gamma+3\mathrm{jets}+X$ [50–52] and $W(\to l\nu)+2\mathrm{jets}+X$ [53, 54]. Double Drell–Yan production has been discussed in theory [55, 56], but not yet been measured, because the expected cross sections are too small. This topic has recently been re-addressed in the context of discrepancies

between experiment and theory for W^+W^- and ZZ production at the LHC [57]. Another proposal is like-sign WW production, where SPS background processes are suppressed compared to the W^+W^- case [58].

To differentiate DPS from high-multiplicity SPS events, the particular properties of DPS events must be exploited: the jet pair (both jet pairs in 4-jet production) should be in a back-to-back orientation in the transverse plane and the azimuthal angle between the vectors of the final-state objects of the leading scattering and the lower-p_{T} jet pair should be uncorrelated, i.e. randomly oriented. Specialising to 4-jet events, three observables, supposedly sensitive to DPS, can be defined as follows:

(i) the azimuthal angular separation between the jets of the lower-p_{T} (*soft*) jet pair

$$\Delta\phi_{\mathrm{soft}} = \left|\phi\left(\mathrm{j1}_{\mathrm{soft}}\right) - \phi\left(\mathrm{j2}_{\mathrm{soft}}\right)\right| ; \tag{6.12}$$

(ii) the balance in transverse momentum of the two lower-p_{T} jets relative to their scalar jet p_{T} sum

$$\Delta^{\mathrm{rel}}_{\mathrm{soft}} p_{\mathrm{T}} = \frac{\left|\vec{p}_{\mathrm{T}}\left(\mathrm{j1}_{\mathrm{soft}}\right) + \vec{p}_{\mathrm{T}}\left(\mathrm{j2}_{\mathrm{soft}}\right)\right|}{\left|\vec{p}_{\mathrm{T}}\left(\mathrm{j1}_{\mathrm{soft}}\right)\right| + \left|\vec{p}_{\mathrm{T}}\left(\mathrm{j2}_{\mathrm{soft}}\right)\right|} ; \tag{6.13}$$

(iii) the azimuthal angular separation between the high- (*hard*) and low-p_{T} (*soft*) dijet pairs

$$\Delta S = \arccos\left(\frac{\vec{p}_{\mathrm{T}}(\mathrm{j1}_{\mathrm{hard}},\mathrm{j2}_{\mathrm{hard}}) * \vec{p}_{\mathrm{T}}(\mathrm{j1}_{\mathrm{soft}},\mathrm{j2}_{\mathrm{soft}})}{\left|\vec{p}_{\mathrm{T}}(\mathrm{j1}_{\mathrm{hard}},\mathrm{j2}_{\mathrm{hard}})\right| \cdot \left|\vec{p}_{\mathrm{T}}(\mathrm{j1}_{\mathrm{soft}},\mathrm{j2}_{\mathrm{soft}})\right|}\right) . \tag{6.14}$$

Here, $\vec{p}_{\mathrm{T}}$ are the transverse momentum vectors of the soft or hard jets j1, j2, or of a jet pair (j1,j2).

The differential distributions in all three observables are normalised and hence examined for shape differences only to eliminate or reduce significantly experimental and theoretical systematic uncertainties. In addition, ATLAS and CMS restricted so far their analyses to the 2010 data set, where only 1-3 PU collisions occurred per bunch crossing. Although the azimuthal angular separation $\Delta\phi_{\mathrm{soft}}$ between the soft jets significantly favours values close to π in the DPS case, this quantity was nevertheless found to be too sensitive to effects of PU collisions rendering it inefficient in distinguishing DPS from SPS events. $\Delta^{\mathrm{rel}}_{\mathrm{soft}} p_{\mathrm{T}}$, where small p_{T} imbalances are preferred in DPS, and ΔS, where SPS events accumulate at values close to π, proved to be more effective. Figure 6.11 shows, as an example, the normalised differential cross sections as functions of $\Delta^{\mathrm{rel}}_{\mathrm{soft}} p_{\mathrm{T}}$ and ΔS for 4-jet production as measured by CMS [49]. The data are compared to dijet LO+PS (PYTHIA8, HERWIG++), multi-jet LO+PS (SHERPA, MADGRAPH +PYTHIA6), and dijet NLO+PS (POWHEG +PYTHIA6) predictions, all complemented with hadronisation and MPI except for one example using POWHEG +PYTHIA6, where the MPI have been switched off. Without MPI, larger excesses are observed in data for small values of $\Delta^{\mathrm{rel}}_{\mathrm{soft}} p_{\mathrm{T}}$ and ΔS, indicating a lack of DPS events. Including the MPI in the event generation, the data are better described, but particularly for the ΔS observable the MC predictions vary too much among themselves preventing to extract reliably a value for

σ_{eff} from 4-jet final states. The situation is better for $W+2$-jet production and template fits could be performed successfully in the ATLAS [53] or CMS analyses [54]. Figure 6.12 presents the decomposition of the data into a DPS signal and an SPS background component for the corresponding $\Delta^{\text{rel}}_{\text{soft}} p_{\text{T}}$ and ΔS observables of the $W+2$-jet template fit of CMS.

In both analyses, σ_{eff} has been derived as

$$\sigma_{\text{eff}} = \frac{\sigma_{W+0-\text{jet}} \cdot \sigma_{2-\text{jet}}}{\sigma^{\text{DPS}}_{W+2-\text{jet}}} \,. \tag{6.15}$$

Such exclusive event selections, however, are in conflict with Equation (6.11) that implies the processes A and B to be inclusive. Otherwise, σ_{eff} would not be defined independently of the choice for the two scattering processes. This bias of exclusive event selections has correctly been considered in the $W+2$-jet studies of ATLAS and CMS.

A summary of all σ_{eff} determinations from various final states is presented in Table 6.1 and Fig. 6.13. In contrast to the discussed ATLAS and CMS studies of $W+2$-jet final states, the CDF analysis in Ref. [50] involves an exclusive event selection for exactly one photon and three jets, which has been subjected to a correction for this selection bias only later in Refs. [59,60]. Furthermore, for 8 TeV centre-of-mass energy a preliminary study from CMS gives a lower limit on σ_{eff} of 5.9 mb at 95% CL [61].

In conclusion, the existence of double parton scattering has been firmly established. The exact value of the only model parameter, the effective cross section σ_{eff} however, remains to be determined more accurately under consideration of all potential biases.

6.5 Multi-Jet Radiation and Azimuthal Decorrelation

A seemingly simple observable is the dijet azimuthal decorrelation $\Delta\phi_{\text{dijet}}$, which is defined as the azimuthal angular separation between the two jets in an event that define the dijet system:

$$\Delta\phi_{\text{dijet}} = |\phi_{\text{jet1}} - \phi_{\text{jet2}}| \,. \tag{6.16}$$

For a $2 \to 2$ process without initial p_{T}, momentum conservation forces the two jets originating from the final-state partons to be produced back-to-back in the transverse plane. The azimuthal angular separation $\Delta\phi_{\text{dijet}}$ between the two jets hence equals π. Any disturbance of this simple topology leads to deviations from $\Delta\phi_{\text{dijet}} = \pi$ and *decorrelates* the azimuthal angles of the two jets. Following Refs. [62,63], the dijet system initially was defined by D0 in terms of the most forward and most backward jets above some jet p_{T} threshold [64].[1] By studying the azimuthal decorrelation as a function of the rapidity separation it was hoped to find effects of BFKL evolu-

[1] Hadronisation corrections, not described within this publication, can be found in Ref. [65].

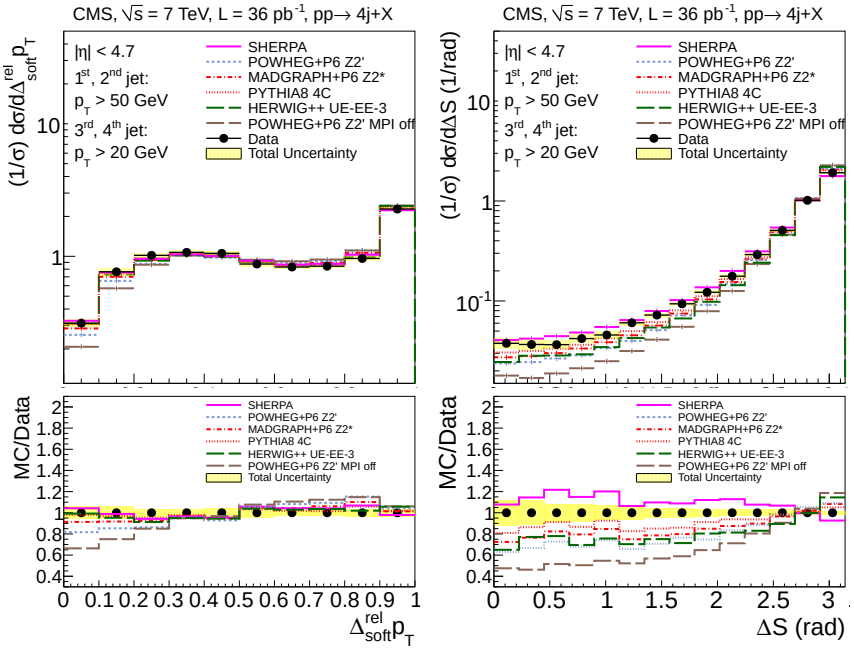

Fig. 6.11 Normalised differential cross section as a function of $\Delta^{\mathrm{rel}}_{\mathrm{soft}} p_{\mathrm{T}}$ (left) and ΔS (right) for 4-jet production in comparison to predictions from various MC event generators. For illustration purposes, the POWHEG +PYTHIA6 prediction is also presented with MPI switched off (POWHEG +P6 Z2' MPI off). *(Taken from Ref. [49])*

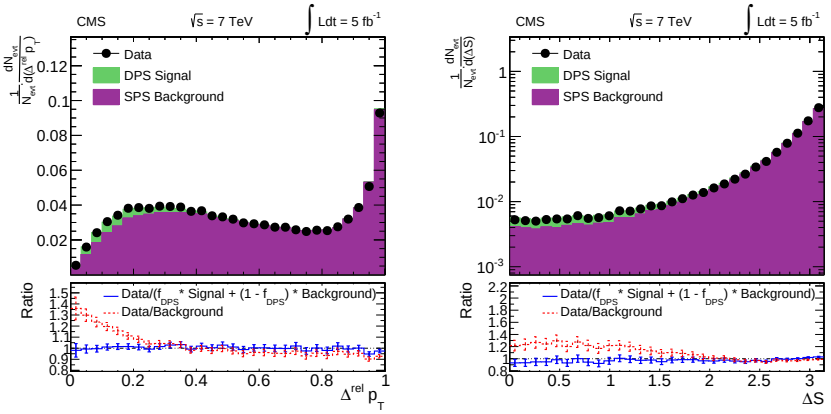

Fig. 6.12 Template fits of SPS background and DPS signal to data for the $\Delta^{\mathrm{rel}}_{\mathrm{soft}} p_{\mathrm{T}}$ (left) and ΔS distributions (right) in W+2-jet final states as measured by CMS. *(Taken from Ref. [54])*

Table 6.1 Summary of σ_{eff} determinations by the AFS, UA2, CDF, D0, ATLAS, and CMS collaborations for various final states. If two uncertainties are given, the first one represents the statistical and the second one the systematic component of the uncertainty.

Experiment	$\sqrt{s}$/GeV	Final state	p_{T} range/GeV	η/y range(s)	σ_{eff}/mb
AFS [46]	63	4 jets	$E_{\text{T}}^{\text{jet}} \geq 4$	$\lvert\eta^{\text{jet}}\rvert < 0.9$	≈ 5
UA2 [47]	630	4 jets	$E_{\text{T}}^{\text{jet}} > 15$	$\lvert\eta^{\text{jet}}\rvert < 2$	$> 8.3\,(95\%\text{C.L.})$
CDF [48]	1800	4 jets	$p_{\text{T}}^{\text{jet}} > 25$	$\lvert\eta^{\text{jet}}\rvert < 3.5$	$12.1^{+10.7}_{-5.4}$
CDF [50]	1800	$\gamma+3$ jets	$E_{\text{T}}^{\text{jet}} > 5$ $E_{\text{T}}^{\text{jet2}} < 7$ $E_{\text{T}}^{\gamma} > 16$	$\lvert\eta^{\text{jet}}\rvert < 4.2$ $\lvert\eta^{\gamma}\rvert < 0.9$	$14.5 \pm 1.7^{+1.7}_{-2.3}$
D0 [51]	1960	$\gamma+3$ jets	$p_{\text{T}}^{\text{jet}} > 15$ $p_{\text{T}}^{\text{jet1}} > 25$ $60 < p_{\text{T}}^{\gamma} < 80$	$\lvert y^{\text{jet}}\rvert < 3.0$ $\lvert y^{\gamma}\rvert < 1.0$ $1.5 < \lvert y^{\gamma}\rvert < 2.5$	$16.4 \pm 0.3 \pm 2.3$
D0 [52]	1960	$\gamma+3$ jets	$p_{\text{T}}^{\text{jet1}} > 15$ $15 < p_{\text{T}}^{\text{jet2,3}} < 35$	$\lvert\eta^{\text{jet}}\rvert < 2.5$ $\lvert y^{\gamma}\rvert < 1.0$	$12.7 \pm 0.2 \pm 1.3$
		$\gamma+c/b$ jet $+2$ jets	$p_{\text{T}}^{\gamma} > 26$	$1.5 < \lvert y^{\gamma}\rvert < 2.5$	$14.6 \pm 0.6 \pm 3.2$
ATLAS [53]	7000	$W(\to l\nu)+2$ jets	$p_{\text{T}}^{\text{jet}} > 20$ $p_{\text{T}}^{l} > 20$ $\lvert\eta^{e}\rvert < 1.37$	$\lvert y^{\text{jet}}\rvert < 2.8$ $\lvert\eta^{\mu}\rvert < 2.4$ $1.52 < \lvert\eta^{e}\rvert < 2.47$	$15 \pm 3^{+5}_{-3}$
CMS [54]	7000	$W(\to \mu\nu)+2$ jets	$p_{\text{T}}^{\text{jet}} > 20$ $p_{\text{T}}^{\mu} > 35$	$\lvert y^{\text{jet}}\rvert < 2.0$ $\lvert\eta^{\mu}\rvert < 2.1$	$20.7 \pm 0.8 \pm 6.6$
CMS [49]	7000	4 jets	$p_{\text{T}}^{\text{jet}} > 20$ $p_{\text{T}}^{\text{jet1,2}} > 50$	$\lvert\eta^{\text{jet}}\rvert < 4.7$	—
CDF result [50] corrected for exclusive selection bias by Bähr et al. [59, 60]					$12.0 \pm 1.4^{+1.3}_{-1.5}$

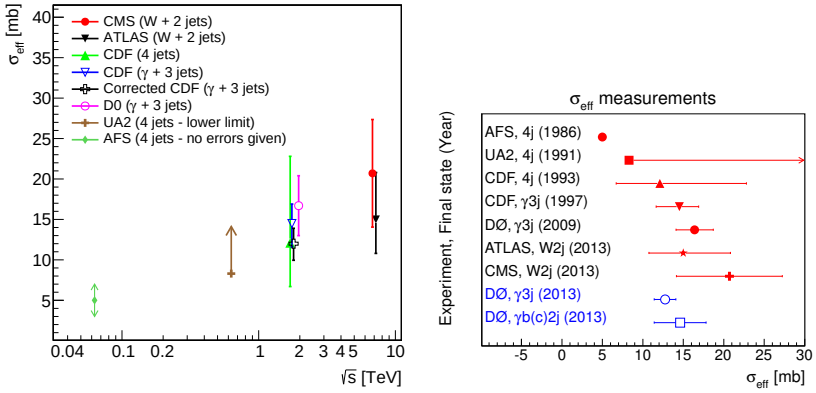

Fig. 6.13 σ_{eff} versus pp ($p\bar{p}$) centre-of-mass energy (left) and comparison of σ_{eff} determinations by the AFS, UA2, CDF, D0, ATLAS, and CMS collaborations (right) both for various final states. The left plot also shows the CDF value corrected in [59, 60] for the exclusive selection bias. The right plot provides two additional points from the more recent measurement by D0. *(Taken from Refs. [52, 54])*

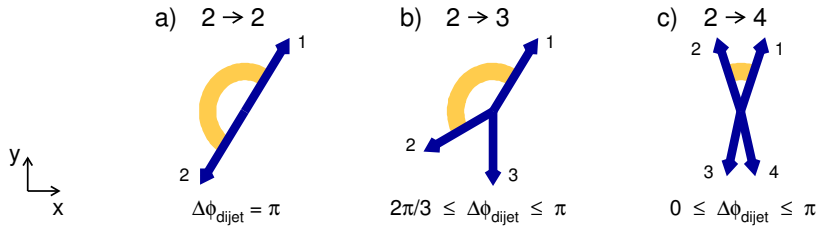

Fig. 6.14 Illustration of different event topologies with up to four jets and the accessible ranges in azimuthal decorrelation $\Delta\phi_{\text{dijet}}$. *(Taken from Ref. [67])*

tion, cf. 2.3. Here, the emphasis will be on the decorrelation in $\Delta\phi_{\text{dijet}}$ of the two jets leading in p_{T} as a consequence of multi-jet production as first investigated by D0 [66] at the Tevatron. Figure 6.14 illustrates various event topologies with up to four jets and the associated ranges in $\Delta\phi_{\text{dijet}}$.

Non-perturbative effects of multiparton interactions or hadronisation perturb the correlation between the two leading jets only mildly and $\Delta\phi_{\text{dijet}} \approx \pi$ still holds. The production of a third high-p_{T} jet, however, leads to a decorrelation in azimuthal angle. The smallest achievable value of $\Delta\phi_{\text{dijet}} = 2\pi/3$ (or 120°) in this case is reached with a symmetric star-shaped 3-jet configuration. Fixed-order calculations for 3-jet production with up to four outgoing partons ($\propto \alpha_S^4$) like from NLOJET++ [68, 69]. provide NLO predictions for this region. If more than three jets are produced, the azimuthal angle between the two leading jets can approach zero, although very small angular separations are suppressed for jets with small rapidity separations because of finite jet sizes for a particular jet algorithm with jet radius R. The measurement of the dijet azimuthal angular decorrelation is therefore a very interesting tool to gain insight into multi-jet production processes without actually measuring jets beyond the leading two.

The actually considered observable is the dijet differential cross section normalised by the inclusive dijet cross section, σ_{dijet}, integrated over $\Delta\phi_{\text{dijet}}$:

$$\frac{1}{\sigma_{\text{dijet}}} \cdot \frac{\mathrm{d}\sigma_{\text{dijet}}}{\mathrm{d}\Delta\phi_{\text{dijet}}}. \tag{6.17}$$

In addition, the phase space is subdivided into bins of the leading jet p_{T}, $p_{\text{T,max}}$. By normalising the $\Delta\phi_{\text{dijet}}$ distribution to the total dijet cross section σ_{dijet} within slices of $p_{\text{T,max}}$, many experimental and theoretical uncertainties are reduced. NP corrections are negligible.

Measurements at the LHC have been reported for 7 TeV centre-of-mass energy by the CMS [70] and ATLAS collaborations [71] and recently at 8 TeV by CMS [72]. Figure 6.15 presents the measurement by CMS at $\sqrt{s} = 8\,\text{TeV}$ with $19.7\,\text{fb}^{-1}$ of integrated luminosity for seven bins in $p_{\text{T,max}}$ ranging from 200 GeV up to 2.2 TeV. Azimuthal angular separations below $\Delta\phi_{\text{dijet}} = \pi/2$ and down to zero are covered for the first time.

The data are compared to fixed-order predictions and here some subtleties arise. As described above, 2-parton final states only contribute to the last bin including $\Delta\phi_{\text{dijet}} = \pi$, while 3-parton final states cover the range $2\pi/3 \leq \Delta\phi_{\text{dijet}} < \pi$, which is well described by 3-jet NLO calculations, if, depending on $p_{\text{T,max}}$, π is not approached too closely. The bin including $\Delta\phi_{\text{dijet}} = \pi$ can be computed as

$$\frac{1}{\sigma_{\text{dijet}}} \cdot \int_{\pi-\Delta\phi_0}^{\pi} \frac{\mathrm{d}\sigma_{\text{dijet}}}{\mathrm{d}\Delta\phi_{\text{dijet}}} \mathrm{d}\Delta\phi_{\text{dijet}}\,, \tag{6.18}$$

where $\Delta\phi_0$ is the bin width. At dijet LO for the integral and the normalization factor the result is unity. At dijet NLO for both, some portion of the real correction corresponding to 3-parton final states lies outside the range $[\pi - \Delta\phi_0, \pi]$ leading to predictions smaller than unity. Excluding the limit $\Delta\phi_{\text{dijet}} = \pi$, a 3-jet NLO prediction becomes negatively singular when approaching π, because collinear and soft gluon radiation leads to enhanced logarithmic terms that spoil the convergence of the perturbative series. In this limit, a resummation becomes necessary, which has been discussed in Ref. [73]. Recent developments including all-order soft-gluon resummation up to next-to-leading logarithmic level are presented in Refs. [74, 75].

In the other direction towards smaller $\Delta\phi_{\text{dijet}}$, the 3-jet NLO also comprises 4-parton configurations and extends the predictions below $\Delta\phi_{\text{dijet}} = 2\pi/3$, but only at 4-jet LO precision with ensuing larger scale uncertainties. In the previous publications by D0, CMS, and ATLAS, Refs. [66, 70, 71], it was claimed that the data in the region $\pi/2 \leq \Delta\phi_{\text{dijet}} < 2\pi/3$ are above the 4-jet LO calculations by up to factors of four. In Ref. [67] it could be shown that this is mostly due to a misconception when forming the ratio of Eq. (6.17). Programs like NLOJET++ deliver differential cross-sections up to a given order in pQCD, but they do not provide directly the expansion in α_S for a ratio R of two cross sections σ_{A} and σ_{B}. By means of expansion in α_S it can be shown though that the ratio $R = \sigma_{\text{A}}/\sigma_{\text{B}}$, where each of the two cross sections is given at the same relative order in α_S, e.g. NLO, is again correct to the same relative order up to terms beyond, as in this example, NLO precision. Of course, this is also a key concept applied in the previous Chapter 5 on cross section ratios. If numerator and denominator are derived in different relative orders, cancellation effects between theoretical uncertainties are compromised, which leads to an artificially increased renormalisation scale dependence as discussed with respect to jet shapes in Ref. [76].[2]

The difference between ratios from Chapter 5 and a normalised quantity like $\Delta\phi_{\text{dijet}}$ is simply the fact that here the denominator represents one normalisation factor for the whole distribution instead of a separate number in each phase-space bin for a ratio. The above-mentioned misconception lies with the fact that in Refs. [66, 70] the 4-jet LO region $\pi/2 \leq \Delta\phi_{\text{dijet}} < 2\pi/3$ was normalised by the same NLO dijet cross section as in the 3-jet NLO region of $2\pi/3 \leq \Delta\phi_{\text{dijet}} < \pi$. Normalising to the LO dijet cross section instead, as done in Fig. 6.15, data are much better described by the consistent LO theory. This can even more explicitly be demon-

[2] See explicitly Sections 3.1 and 4.

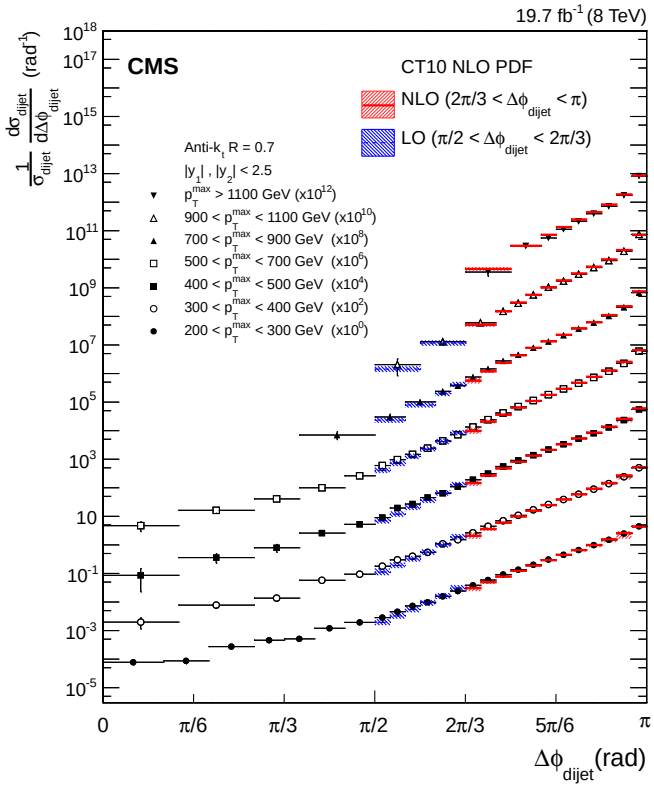

Fig. 6.15 Normalised dijet cross section measured by CMS at $\sqrt{s} = 8\,\text{TeV}$ as a function of $\Delta\phi_{\text{dijet}}$ for seven regions in $p_{\text{T,max}}$, scaled for visibility by the indicated factors. The error bars on the data points comprise statistical and systematic uncertainties. Overlaid on the data are predictions (line) from LO ($\pi/2 \leq \Delta\phi_{\text{dijet}} < 2\pi/3$) and NLO ($2\pi/3 \leq \Delta\phi_{\text{dijet}} \leq \pi$) calculations using the CT10-NLO PDF set. PDF, α_S, and scale uncertainties are added quadratically to give the total theoretical uncertainty that is indicated by the hatched regions. *(Taken from Ref. [72])*

strated by comparing the ratio of data over theory for the two cases as presented in Fig. 6.16 left, adapted from a preliminary study by CMS [77], and right from their final publication [72]. In addition, a reduction of the scale uncertainty from roughly $^{+45}_{-30}\%$ to $^{+30}_{-20}\%$ becomes apparent.

Another issue further discussed in Ref. [67] concerns the selection of the leading jets in an event. In case of the DO and CMS analyses it is guaranteed that only those events are kept, where the two jets leading in p_{T} lie within the $|y|$ range of the final selection. CMS for example accepts initially jets up to $|y| = 5$. For ATLAS [71], the unfavourable situation can occur that at least one of the leading jets is removed in the initial selection of $|y_{\text{jet}}| < 2.8$ leading to very small azimuthal angular separations already for the 3-jet case. The clear separation between the 3-jet and 4-jet region

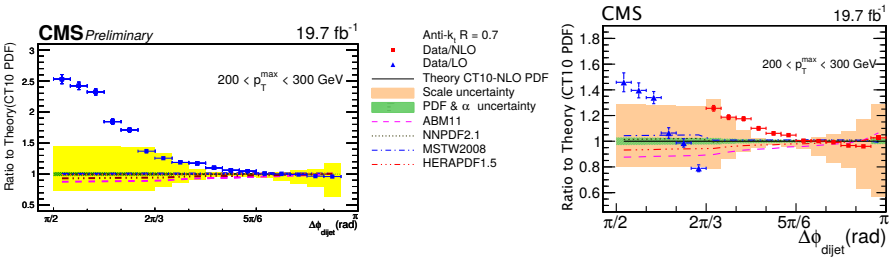

Fig. 6.16 Ratio to theory of the normalised $\Delta\phi_{\text{dijet}}$ distribution for $200\,\text{GeV} < p_{\text{T,max}} < 300\,\text{GeV}$ adapted from a preliminary study by CMS (left), and from their final result (right). The error bars on the data points represent the total experimental uncertainty, while the yellow (left)/orange (right) band represents the scale uncertainty of the theory. In addition to the central theory prediction derived with the CT10-NLO PDF (full line at unity), results for other PDF sets are shown as further lines. The normalisation factor applied in the range of $\pi/2 \leq \Delta\phi_{\text{dijet}} < 2\pi/3$ corresponds to the NLO dijet cross section (left, square symbols without differentiation between the two regions) of the respective $p_{\text{T,max}}$ bin and to the LO one (right). *(Adapted from Refs. [72, 77])*

thus is lost. More appropriate comparisons by ATLAS of up to 4-jet observables to NLO predictions can be found in Ref. [78].

In addition to fixed-order comparisons, dijet azimuthal decorrelations serve to improve the usual LO and multijet improved MC event generators, because they are in particular sensitive to effects of initial state radiation. Further very interesting studies with dijet azimuthal decorrelations, possible already now, comprise the NLO for 4- and 5-jet production [79, 80] and the NLO dijet and 3-jet production matched with parton showers [81, 82]. As a final remark, Ref. [83] points out another observable definition to look into dijet azimuthal decorrelations that is also a promising candidate for a determination of the strong coupling constant $\alpha_S(M_Z)$. A first investigation has been published by the D0 Collaboration [84].

6.6 Multi-Jet Production and Event Shapes

Previous observables always made use of a particular number of jets and their individual kinematic properties or relations between two of them. When studying multijet production, however, it might be even more interesting to characterise events in a global manner without referring to individual jets. Quantities that describe events by one continuous dimensionless number, generically denoted by F in the following, are called *event shapes*. They are calculated from a set of 4-vectors of the final state of a collision, where these 4-vectors might be associated to jets or to individual particles for example. Of course, to compare measured distributions to perturbative QCD, event shapes need to be defined such that they are insensitive to the details of soft non-perturbative effects of QCD, i.e. they have to satisfy the same condition as "good" jet algorithms of being collinear- and infrared-safe, cf. Eqs. (2.27).

Since event shapes can also be calculated from e.g. all final-state particles, they might not even refer to jets at all! A recent example is given by the ATLAS study on event shapes from charged particles [85]. In fact, the first notion of event shapes [86, 87] predates that of jets, although the terms *thrust* and *sphericity* were coined only later [88,89], roughly about the same time as the first proposed jet algorithm [90]. In these early studies of e^+e^- collisions, the main interest was to differentiate between a 2-jet-like structure favoured by QCD and the expectations from alternative models. For a 2-to-2 scattering like $e^+e^- \to q\bar{q} \to$ hadrons, a pencil-like momentum flow is predicted. The first evidence of jet production was established in 1975 making use of sphericity [91], which is, however, not collinear- and infrared safe, because squares of momenta enter in its computation. A modification making use only linearly of the momenta is suggested in [92] under the name of *spherocity*. The thrust T is also a safe event measure. It is defined as the normalised sum of the projections of all momenta onto the event axis, the thrust axis $\vec{n}_T$, that maximises T:

$$T := \max_{\vec{n},\vec{n}^2=1} \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|} = \frac{\sum_i |\vec{p}_i \cdot \vec{n}_T|}{\sum_i |\vec{p}_i|} \,. \tag{6.19}$$

The thrust value varies between $T = 1$ for a strictly linear orientation of all momenta, where the thrust axis coincides with this direction, and $T = T_{\min}$ with $T_{\min} = 0.5$ in case of a completely spherically symmetric distribution of the produced hadrons in e^+e^- annihilation. For practical reasons, event shapes F are defined to equal zero for the simplest possible configuration of final-state momenta as in the respective tree-level Feynman diagrams. Hence, thrust is redefined to $\tau := 1 - T$.

While in e^+e^- collisions the relation between the thrust axis $\vec{n}_T$ and the outgoing back-to-back $q\bar{q}$ pair is straightforward, this is less simple for lepton-hadron or hadron-hadron scattering. In DIS, the strategy is to examine an event in the Breit frame of reference where the incoming parton is back-scattered by the purely space-like photon exchanged with the electron probe. The hemisphere with only the back-scattered quark, the so-called *current hemisphere*, can then be treated as analogue of half a $e^+e^- \to q\bar{q}$ event. Relevant measurements have been reported by the H1 [93–95] and ZEUS collaborations [96, 97]. A topical review of event shapes in e^+e^- and ep scattering is presented in Ref. [98].

For hadron-hadron collisions the centre-of-mass system cannot be determined because of its unknown longitudinal boost. Therefore, events are analysed in the plane perpendicular to the beam directions, where the vectors of all transverse momenta ideally sum up to the null vector. In addition, experimental constraints due to the beams delimit the measurements to pseudorapidities below some maximal value $\eta_{\max}$. This limitation has consequences on the theory side with respect to resummations as discussed in Refs. [99, 100], Correspondingly, such event shapes are called e.g. *central* transverse thrust. Because all event shapes investigated experimentally so far are of the "central" type, this will be omitted in the following. Transverse thrust is then defined as

$$\tau_\perp = 1 - T_\perp = \max_{\vec{n},\vec{n}^2=1} \frac{\sum\limits_{|\eta_i|<\eta_{\max}} \vec{p}_{\perp,i} \cdot \vec{n}}{\sum\limits_{|\eta_i|<\eta_{\max}} \left|\vec{p}_{\perp,i}\right|} = \frac{\sum\limits_{|\eta_i|<\eta_{\max}} \vec{p}_{\perp,i} \cdot \vec{n_T}}{\sum\limits_{|\eta_i|<\eta_{\max}} \left|\vec{p}_{\perp,i}\right|}\,, \tag{6.20}$$

where in this 2-dimensional case the maximal value of $\tau = 1 - T$ is $1 - 2/\pi$. With the help of the thrust axis $\vec{n_T}$ and the plane perpendicular to it, events can again be divided into two hemispheres. Since the orientation into a plus and minus direction of $\vec{n_T}$ is accidental, it is not useful to differentiate the two hemispheres in that aspect alone. However, for each hemisphere one can calculate separately e.g. the *jet broadening*, or the *jet mass* as

$$B_\pm = \frac{1}{2P_T} \sum_{i\in\pm} p_{\mathrm{T},i} \sqrt{(\eta_i - \eta_\pm)^2 + (\phi_i - \phi_\pm)^2}\,, \quad \text{and} \tag{6.21}$$

$$\rho_\pm = \frac{1}{P^2} \left(\sum_{i\in\pm} p_i \right)^2 = \frac{M_\pm^2}{P^2}\,. \tag{6.22}$$

Here, P (P_T) are the sum of all (transverse) momenta of all jets, while the sums with $i \in \pm$ run separately over all jets i with (transverse) momenta ($p_{\mathrm{T},i}$) p_i, pseudorapidities η_i, and azimuthal angles ϕ_i within either the $+$ or $-$ hemisphere. The pseudorapidities $\eta_\pm$ and azimuthal angles $\phi_\pm$ of each hemisphere are defined as

$$\eta_\pm = \frac{\sum\limits_{i\in\pm} p_{\mathrm{T},i} \cdot \eta_i}{\sum\limits_{i\in\pm} p_{\mathrm{T},i}}\,, \quad \text{and} \tag{6.23}$$

$$\phi_\pm = \frac{\sum\limits_{i\in\pm} p_{\mathrm{T},i} \cdot \phi_i}{\sum\limits_{i\in\pm} p_{\mathrm{T},i}}\,. \tag{6.24}$$

With these definitions, the total jet broadening and total jet mass are derived as $B_{\mathrm{tot}} = B_+ + B_-$ respectively $\rho_{\mathrm{tot}} = \rho_+ + \rho_-$. In contrast to somewhat misleading statements in Ref. [101], all three quantities, $\tau_\perp$, B_{tot}, and ρ_{tot}, are identically zero for $2 \to 2$ processes at the level of massless partons of a fixed-order calculation.[3] So values close to zero characterise the 2-jet likeness of an event, while deviations from zero require at least one additional parton to be resolved. It is possible to characterise this without referring to any event axis or hemispheres by using the n-jet resolution parameter. For example to differentiate between 2-jet- and 3-jet-like events, one defines for the chosen jet algorithm the dimensionless distance, where a 3-jet configuration is merged into a 2-jet configuration, as 3-jet resolution parameter Y_{23}. So instead of classifying events into n-jet events by means of a jet algorithm, one characterises the n-jet likeness by a continuous dimensionless quantity, $Y_{n,n+1}$, derived from the jet algorithm. In case of Y_{23}, if only two partons (jets) existed right from the beginning, an infinite resolution corresponding to $Y_{23} \to 0$ would be

[3] Experimental effects can lead to $\tau_\perp$ and ρ_{tot} being non-zero even for dijet events, while B_{tot} and Y_{23} require at least three jets for that to happen.

required. So also Y_{23} is zero in the 2-jet or 2-parton limit. Other event shapes are in use like the C and D parameters [102] or Fox–Wolfram moments [103], some of which go beyond the characterisation of the 2-jet likeness of an event. A quantity related to the C parameter, the energy-energy correlation EEC [104, 105] has been used by ATLAS to determine a value of the strong coupling constant $\alpha_S(M_Z)$ [106]. Measurements on event shapes performed at previous hadron-hadron colliders can be found in Refs. [107, 108] (Sp$\bar{\text{p}}$S) and [109–112] (Tevatron). For further reading on the theory perspective cf. Refs. [99, 100].

Figure 6.17 presents as an example the measurements of transverse thrust $\tau_\perp$, total jet broadening B_{tot}, and 3-jet resolution parameter Y_{23} by CMS [101]. The normalised distributions have been determined from the 2011 data set at $\sqrt{s} = 7\,\text{TeV}$ that corresponds to $5\,\text{fb}^{-1}$ of integrated luminosity. Events are selected, if the two jets leading in p_{T} lie within the central detector with pseudorapidities of $\eta < \eta_{\text{max}} = 2.4$. For the computation of the event shapes, jets within the same η range and a minimal p_{T} of 30 GeV are used. In addition, the phase space is subdivided into five bins of leading jet p_{T} with a minimum of 110 GeV. The example plots of Fig. 6.17 are chosen to be the ones for the largest leading-jet p_{T} bin of $p_{\text{T},1} > 390\,\text{GeV}$.

In the middle panel, ratios of MC over data are shown for the predictions of the 2-jet LO+PS MC event generator PYTHIA6 with three different tunes, Z2, Perugia-P0, and D6T, cf. Chapter 2. Significant deviations are observed for all tunes in all distributions, particularly for high values of the event shapes corresponding to multi-jet production. The best description among these three is given by the Perugia-P0 tune. The bottom panels provide the same ratio for the 2-jet LO+PS event generator programs PYTHIA8, tune 4C, and HERWIG++ with default tune of version 2.3, and for the tree-level multi-jet LO+PS predictions by MADGRAPH +PYTHIA6 with tune Z2. As expected, discrepancies are found in the multi-jet region between the 2-jet LO+PS generators and the data. PYTHIA8 works slightly better than HERWIG++. In contrast, MADGRAPH +PYTHIA6 well describes the measurements over most of the phase space. Astonishingly, almost the inverse behaviour was reported in the very first measurement of event shapes at the LHC, also by CMS [113], for a much smaller data sample. A similar ATLAS study on data from 2010 [114] found a slight advantage in the description of their data by PYTHIA6 with a different tune Perugia2010, and by the multi-jet LO prediction by ALPGEN with HERWIG +JIMMY for the parton shower, hadronisation, and MPI event generation steps.

Although calculable for example with the NLOJET++ program, comparisons to fixed-order predictions at NLO have not been performed yet at the LHC, presumably because past experience from e^+e^- collisions demonstrate the importance of resummation in predicting event shape distributions. Owing to the big progress made by the theory community, in many cases it is possible now to combine NLO with parton showers. In an investigation of event shapes for the Z+jet process, it was found by CMS [115] that the best description of the data is given by NLO+PS events as produced by POWHEG matched to PYTHIA6, tune Z2, for the remaining event generation steps. With the advent of NNLO predictions for jet or W/Z+jet production [116–119], interest in event shapes should increase beyond their current use in MC generator tuning efforts.

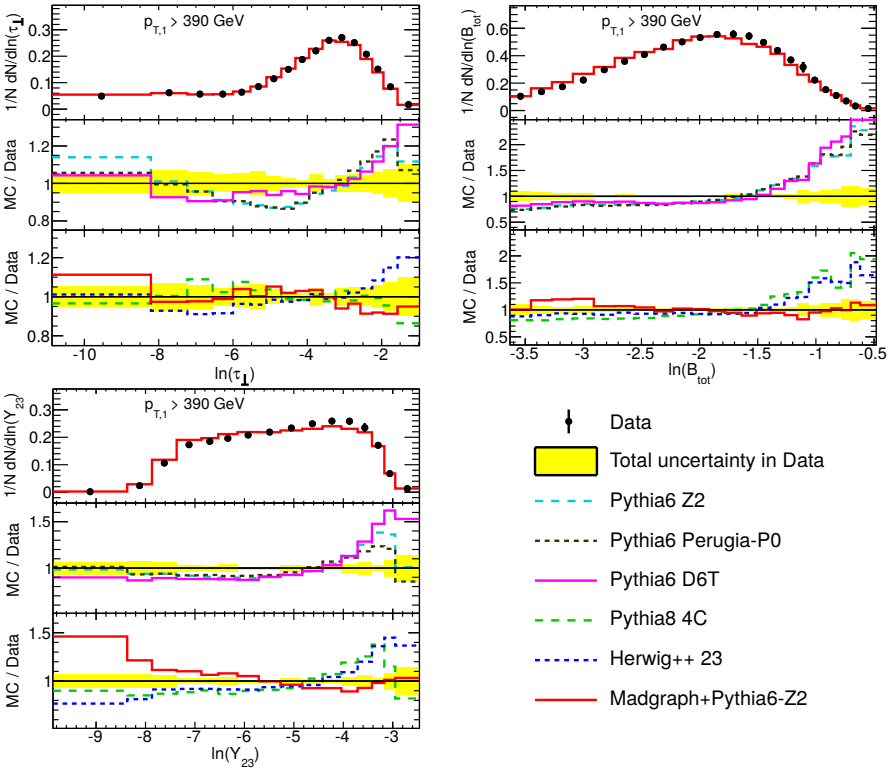

Fig. 6.17 Measured distributions of logarithmic transverse thrust $\tau_\perp$ (top left), total jet broadening B_{tot} (top right), and 3-jet resolution parameter Y_{23} (bottom left) in comparison to predictions from various MC event generators for the bin with largest p_{T} of the leading jet. The error bars represent the statistical uncertainty on the data. The middle and bottom panels show the ratios of the different MC predictions to the data, where the yellow bands correspond to the quadratic sum of statistical and systematic experimental uncertainties. *(Adapted from Ref. [101])*

6.7 Jet Shapes and Jet Substructure

In contrast to jet cross sections and related quantities, the use of event shapes permits to examine the structure of a collision as a whole. In some sense, the jet resolution is decreased such that everything is clustered into only one object, the event. Going the opposite direction, one can ask what we can learn from looking inside the jets. An obvious topic for studies within QCD is, of course, to identify and analyse differences in jets caused by the different flavours of partons initiating a jet, i.e. light quarks, gluons, and the heavier charm and bottom quarks. top quarks decay already at timescales before the onset of the fragmentation process and can not be considered as primary partons *per se*. Instead, the fragmentation of their decay products needs to be studied.

A new situation arises at the LHC, where transverse momenta well beyond the particle masses can be reached. Under such conditions the hadronic decay products are not anymore isolated in different jets, but appear as components of one large (*fat*) jet. With respect to the production and decay of Higgs bosons this has first been investigated in Ref. [120]. Dedicated studies on boosted top quarks have been performed in Refs. [121] and [122], while a more general discussion also including Z and W bosons can be found in Ref. [123]. The possibility to differentiate such jets initiated by heavy boosted objects from "normal" QCD jets initiated by quarks (without top) and gluons has the potential for a big impact not only to identify known, but also new massive particles. Hadronically decaying Z' or t' resonances for example would give rise to jets with two or three, respectively, high-energetic subcomponents that need to be isolated from QCD related background by so-called jet cleaning or *grooming* techniques. Several such techniques have been developed: *jet filtering* [120], *jet trimming* [124], and *jet pruning* [125]. In this context, the investigation of jet substructure has attracted much attention and even lead to the creation of a dedicated series of workshops on "Boosted Object Phenomenology, Reconstruction and Searches in HEP (BOOST)" starting in 2009 at the SLAC National Accelerator Laboratory in Menlo Park, California. Within the scope of this work, it is impossible to do justice to this dynamically developing field that essentially merits a review of its own, cf. e.g. Refs. [126] and [127]. In the following, two observables are discussed with respect to the differentiation of (light) quark- and gluon-initiated jets.

Focusing on the internal structure of jets, the profile of the distribution of transverse momentum within a jet can be examined. For this type of observable the term *jet shape* has been coined [128, 129]. The differential jet shape $\rho(r)$ as a function of the distance $r_i = \sqrt{(\Delta_{i,\text{jet}}y)^2 + (\Delta_{i,\text{jet}}\phi)^2}$ to the jet axis is defined as the average fraction of jet p_T contained inside an annulus of inner radius $r_a = r - \Delta r/2$ and outer radius $r_b = r + \Delta r/2$ for an ensemble of N jets:

$$\rho(r) = \frac{1}{N} \sum_{\text{jets}} \frac{1}{\Delta r} \frac{\sum_{r_a \leq r_i < r_b} p_{\text{T},i}}{\sum_{r_i \leq R} p_{\text{T},i}}, \tag{6.25}$$

where the second sum runs over all jet constituents i. As described in Ref. [76], for fixed-order calculations one has to pay attention to normalise to the same relative order as the numerator to avoid an artificially large μ_r dependence, compare Section 6.5. The integrated jet shape $\Psi(r)$ is then given by the integral of the differential jet shape up to a radius r, cf. the illustrations in Fig. 6.18 right, upper panel. Conventionally, measurements are presented in terms of $1 - \Psi(r = r_\text{core})$, where r_core is taken to be 0.3 [130]. Jet shapes have been measured at the Tevatron by the CDF [129, 130] and D0 [131] experiments and at HERA by the ZEUS [132–134] and H1 [135] collaborations. An analysis specialising to b-jets is reported by CDF in Ref. [136].

Figure 6.18 left shows a CMS measurement [137] of the differential jet shape, for the example of the jet p_T interval $110 < p_\text{T} < 125\,\text{GeV}$, in comparison to several dijet LO MC event generators at particle level. While HERWIG++ and PYTHIA8

predict somewhat broader jet shapes than measured, the PYTHIA6 tunes D6T and Z2 deviate in the opposite direction. The best description of the data is given by PYTHIA6 with the tune Perugia2010, which is similarly observed in an investigation by the ATLAS Collaboration [138]. By employing this tune for the integrated jet shape, shown in the form of $1-\Psi(r=0.3)$ as a function of the jet p_{T} in Fig. 6.18 right, ATLAS demonstrates the sensitivity of this observable with respect to the jet-initiating parton. At small jet p_{T} the examined inclusive jet sample is predominantly composed of gluon jets, while with increasing jet p_{T} the quark-initiated component grows. In the figure, the change in the fraction of quark-initiated to gluon-initiated jets is averaged over the jet rapidity up to $|y|<2.8$. If studied double-differentially, a mild dependence on $|y|$ of this fraction is observed as expected. The sensitivity of jet shapes to non-perturbative effects and modelling uncertainties, however, prevents an extraction of the quark-gluon jet fraction. Instead, they provide valuable input to the tuning of these effects in MC event generators.

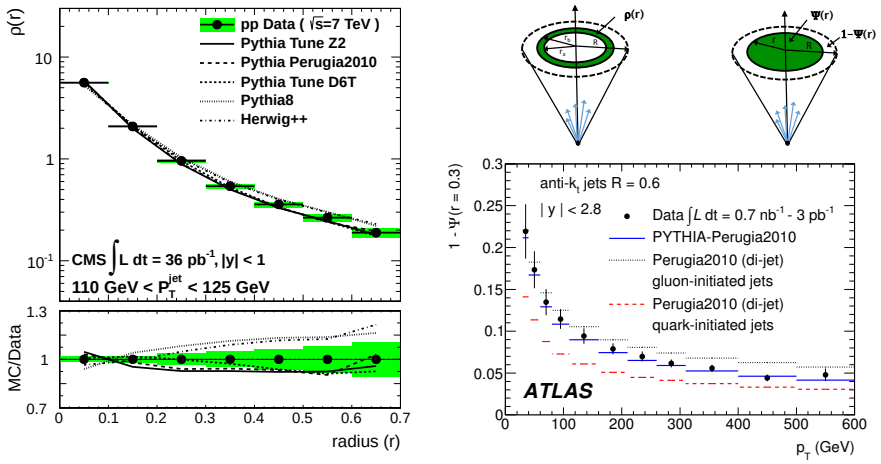

Fig. 6.18 Left: Differential jet shape $\rho(r)$ as a function of the distance r from the jet axis for jets with $110<p_{\mathrm{T}}<125\,\mathrm{GeV}$ from CMS. The data are compared to various tunes of the PYTHIA6 and to the PYTHIA8 and HERWIG++ event generators. Right upper panel: Illustration of the jet shape observables. Right lower panel: Integrated jet shape $1-\Psi(r=0.3)$ as a function of the jet p_{T} from ATLAS. The data are compared to the prediction for gluon-initiated and quark-initiated jets using the PYTHIA6 MC event generator with tune Perugia2010. *(Adapted from Refs. [137, 138])*

Perturbatively, the internal structure of jets is determined by multiple emissions of gluons and depends on the type of initiating parton, i.e. a quark or gluon. Perturbative QCD predicts that gluon-initiated jets are broader in shape with a higher particle multiplicity on average. More precisely, the average multiplicity of any type of object in a gluon jet should roughly be about $C_A/C_F=9/4$ times larger than that in a quark jet [8], in accordance with the larger effective squared colour charge of a gluon, C_A, than that of a quark, C_F. Similarly, the multiplicity distribution of a gluon

jet should be wider by a factor of $\sqrt{C_A/C_F}$ than that of a quark jet [8]. Therefore, it is possible to differentiate statistically between gluon and quark jets by looking into the jet substructure. It should be noted though that strictly speaking the concept of quark and gluon jets is not well-defined beyond LO in pQCD, because higher-order processes cannot be uniquely attributed to a particular LO final-state flavour composition [139]. Nevertheless it is useful to look into alternatives to jet shapes, which are sensitive to non-perturbative effects, for differentiating between jet types. One such proposition [140–142] consists in counting the average number of subjets resolved at a particular jet distance y_{cut} that is smaller than the one used for the initial clustering of the jets. Averaging over a sample of jets, one defines the *average subjet multiplicity* $\langle M \rangle$. An obvious candidate algorithm for undoing again a first jet clustering is the k_{t} jet algorithm as suggested in Ref. [140] and used in the measurements by D0 [143] as well as by the H1 and ZEUS collaborations [135, 144, 145]. In Ref. [144], $\langle M \rangle$ was also used to determine the strong coupling constant.

A challenge of this observable, $\langle M \rangle$, is its dependence on the product of two ingredients that are not well-known: the average multiplicity for quark-initiated jets, $\langle M_q \rangle$, and gluon-initiated ones, $\langle M_g \rangle$, on the one hand, and the relative frequency of quark versus gluon jets, i.e. the *quark and gluon fractions* $1-f$ and f respectively, on the other hand:

$$\langle M \rangle = f \cdot \langle M_g \rangle + (1-f) \cdot \langle M_q \rangle \,. \tag{6.26}$$

Assuming that $\langle M_q \rangle$ and $\langle M_g \rangle$ are independent of the centre-of-mass energy, measurements at two different $\sqrt{s}$ can be exploited to reduce MC modelling dependencies as performed in Ref. [143] at $\sqrt{s} = 630$ and 1800 GeV:

$$\langle M_q \rangle = \frac{f_{s_2} \cdot \langle M \rangle_{s_1} - f_{s_1} \cdot \langle M \rangle_{s_2}}{f_{s_2} - f_{s_1}} \,, \tag{6.27}$$

$$\langle M_g \rangle = \frac{(1-f_{s_1}) \cdot \langle M \rangle_{s_2} - (1-f_{s_2}) \cdot \langle M \rangle_{s_1}}{f_{s_2} - f_{s_1}} \,. \tag{6.28}$$

Of course, the assumption of $\sqrt{s}$ independence has to be either justified e.g. by testing with tagged jets or accounted for by estimating an uncertainty.

At the LHC only a preliminary result on subjet multiplicities at 7 TeV has been reported so far by CMS [146]. Figure 6.19 presents the measured average subjet multiplicity at central rapidity as a function of jet p_{T} in comparison to predictions from PYTHIA6 tunes D6T and Z2 (left) and from HERWIG++ version 2.3 and PYTHIA8 tune 4C (right). Only HERWIG++ well describes the data, while the PYTHIA6 tunes undershoot and PYTHIA8 overshoots them systematically. In addition, the predictions for pure quark and gluon jets are shown for each MC generator, where the quark subjet multiplicities are closer to $\langle M \rangle$ at low p_{T}, while at high p_{T} the gluon subjet multiplicities take over, consistent with an increasing gluon fraction f. This observation is similar to the ATLAS result with jet shapes as shown previously in Fig. 6.18 right lower panel. It would be most interesting to see a study exploiting the by now four different centre-of-mass energies of 2.76, 7, 8, and 13 TeV at the LHC.

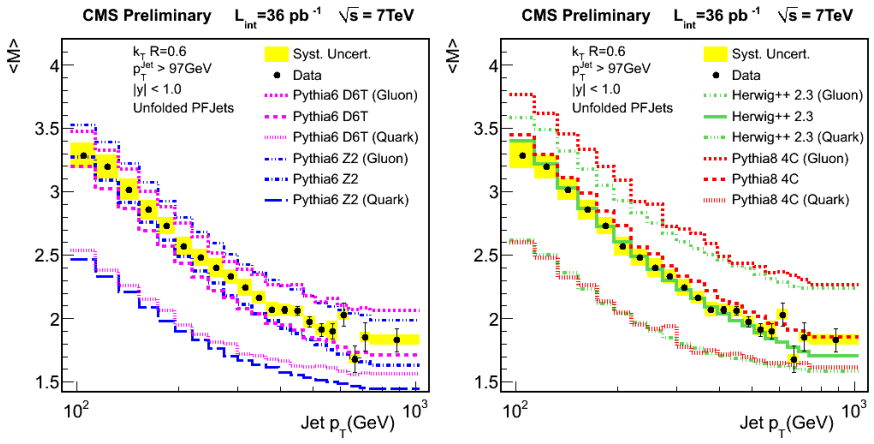

Fig. 6.19 Average subjet multiplicity $\langle M \rangle$ as a function of jet p_T in the central detector region $|y| < 1$ in comparison to predictions from PYTHIA6 tunes D6T and Z2 (left) and from HERWIG++ version 2.3 and PYTHIA8 tune 4C (right). The shaded band (yellow) corresponds to the systematic uncertainty while statistical uncertainties are shown as error bars. In addition, the MC predictions for purely quark- and gluon-initiated jets are shown. *(Adapted from Ref. [146])*

References

1. CMS Collaboration, "Search for contact interactions using the inclusive jet p_T spectrum in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. D* **87** (2013) 052017, doi:10.1103/PhysRevD.87.052017, arXiv:1301.5023.
2. CDF Collaboration, "Inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **77** (1996) 438, doi:10.1103/PhysRevLett.77.438, arXiv:hep-ex/9601008.
3. CDF Collaboration, "Inclusive jet cross-section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **68** (1992) 1104, doi:10.1103/PhysRevLett.68.1104.
4. D0 Collaboration, "Limits on quark compositeness from high-energy jets in $\bar{p}p$ collisions at 1.8 TeV", *Phys. Rev. D* **62** (2000) 031101, doi:10.1103/PhysRevD.62.031101, arXiv:hep-ex/9912023.
5. CMS Collaboration, "Search for jet extinction in the inclusive jet-p_t spectrum from proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **90** (2014) 032005, doi:10.1103/PhysRevD.90.032005, arXiv:1405.7653.
6. C. Kilic et al., "Jet extinction from nonperturbative quantum gravity effects", *Phys. Rev. D* **89** (2014) 016003, doi:10.1103/PhysRevD.89.016003, arXiv:1207.3525.
7. B. L. Combridge, J. Kripfganz, and J. Ranft, "Hadron Production at Large Transverse Momentum and QCD", *Phys. Lett. B* **70** (1977) 234, doi:10.1016/0370-2693(77)90528-7.
8. R. K. Ellis, W. J. Stirling, and B. R. Webber, "QCD and Collider Physics". Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, Cambridge, 1996.
9. G. Dissertori, I. G. Knowles, and M. Schmelling, "Quantum Chromodynamics: High Energy Experiments and Theory". Oxford University Press, 2nd edition, May, 2009.
10. B. L. Combridge and C. J. Maxwell, "Untangling Large p(T) Hadronic Reactions", *Nucl. Phys. B* **239** (1984) 429, doi:10.1016/0550-3213(84)90257-8.

11. CMS Collaboration, "Measurement of Dijet Angular Distributions and Search for Quark Compositeness in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **106** (2011) 201804, doi:10.1103/PhysRevLett.106.201804, arXiv:1102.2020.
12. ATLAS Collaboration, "Search for New Physics in Dijet Mass and Angular Distributions in pp Collisions at $\sqrt{s} = 7$ TeV Measured with the ATLAS Detector", *New J. Phys.* **13** (2011) 053044, doi:10.1088/1367-2630/13/5/053044, arXiv:1103.3864.
13. CMS Collaboration, "Search for quark compositeness in dijet angular distributions from pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **05** (2012) 055, doi:10.1007/JHEP05(2012)055, arXiv:1202.5535.
14. ATLAS Collaboration, "ATLAS search for new phenomena in dijet mass and angular distributions using pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2013) 029, doi:10.1007/JHEP01(2013)029, arXiv:1210.1718.
15. CMS Collaboration, "Search for quark contact interactions and extra spatial dimensions using dijet angular distributions in proton–proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **746** (2015) 79, doi:10.1016/j.physletb.2015.04.042, arXiv:1411.2646.
16. ATLAS Collaboration, "Search for New Phenomena in Dijet Angular Distributions in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV Measured with the ATLAS Detector", *Phys. Rev. Lett.* **114** (2015) 221802, doi:10.1103/PhysRevLett.114.221802, arXiv:1504.00357.
17. J. Gao et al., "Next-to-leading QCD effect to the quark compositeness search at the LHC", *Phys. Rev. Lett.* **106** (2011) 142001, doi:10.1103/PhysRevLett.106.142001, arXiv:1101.4611.
18. J. Gao, C. S. Li, and C. P. Yuan, "NLO QCD Corrections to dijet Production via Quark Contact Interactions", *JHEP* **07** (2012) 037, doi:10.1007/JHEP07(2012)037, arXiv:1204.4773.
19. S. Dittmaier, A. Huss, and C. Speckner, "Weak radiative corrections to dijet production at hadron colliders", *JHEP* **11** (2012) 095, doi:10.1007/JHEP11(2012)095, arXiv:1210.0438.
20. S. Catani et al., "New clustering algorithm for multi - jet cross-sections in e+ e- annihilation", *Phys. Lett. B* **269** (1991) 432, doi:10.1016/0370-2693(91)90196-W.
21. G. Marchesini and B. R. Webber, "Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation", *Nucl. Phys. B* **310** (1988) 461, doi:10.1016/0550-3213(88)90089-2.
22. S. Gieseke, P. Stephens, and B. Webber, "New formalism for QCD parton showers", *JHEP* **12** (2003) 045, doi:10.1088/1126-6708/2003/12/045, arXiv:hep-ph/0310083.
23. MARK-J Collaboration, "Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA", *Phys. Rev. Lett.* **43** (1979) 830, doi:10.1103/PhysRevLett.43.830.
24. TASSO Collaboration, "Evidence for Planar Events in e^+e^- Annihilation at High Energies", *Phys. Lett. B* **86** (1979) 243, doi:10.1016/0370-2693(79)90830-X.
25. PLUTO Collaboration, "Evidence for Gluon Bremsstrahlung in e^+e^- Annihilations at High Energies", *Phys. Lett. B* **86** (1979) 418, doi:10.1016/0370-2693(79)90869-4.
26. JADE Collaboration, "Observation of Planar Three-Jet Events in e^+e^- Annihilation and Evidence for Gluon Bremsstrahlung", *Phys. Lett. B* **91** (1980) 142, doi:10.1016/0370-2693(80)90680-2.
27. JADE Collaboration, "Test of Fragmentation Models by Comparison with Three-Jet Events Produced in $e^+e^- \to$ Hadrons", *Phys. Lett. B* **134** (1984) 275, doi:10.1016/0370-2693(84)90687-7.
28. TASSO Collaboration, "A Study of Three Jet Events in e^+e^- Annihilation Into Hadrons at 34.6 GeV Center-Of-Mass Energy", *Z. Phys. C* **29** (1985) 29, doi:10.1007/BF01571375.

29. TPC/Two Gamma Collaboration, "Tests of models for quark and gluon fragmentation in e^+e^- annihilation at $\sqrt{s} = 29$ GeV", *Z. Phys. C* **28** (1985) 31, doi:10.1007/BF01550246.
30. TPC/Two Gamma Collaboration, "Tests of models for parton fragmentation using three jet events in e^+e^- annihiliation at $\sqrt{s} = 29$ GeV", *Phys. Rev. Lett.* **54** (1985) 270, doi:10.1103/PhysRevLett.54.270. [Erratum: Phys. Rev. Lett.54,1209(1985)].
31. TPC/Two Gamma Collaboration, "Comparison of the particle flow in $q\bar{q}g$ and $q\bar{q}\gamma$ events in e^+e^- annihilation", *Phys. Rev. Lett.* **57** (1986) 945, doi:10.1103/PhysRevLett.57.945.
32. P. D. Sheldon et al., "A Comparison of the Particle Flow in Three Jet and Radiative Two Jet Events From e^+e^- Annihilation at e(CM) = 29 GeV", *Phys. Rev. Lett.* **57** (1986) 1398, doi:10.1103/PhysRevLett.57.1398.
33. OPAL Collaboration, "A Study of coherence of soft gluons in hadron jets", *Phys. Lett. B* **247** (1990) 617, doi:10.1016/0370-2693(90)91911-T.
34. L3 Collaboration, "Evidence for gluon interference in hadronic Z decays", *Phys. Lett. B* **353** (1995) 145, doi:10.1016/0370-2693(95)00552-V.
35. C. Berger, "Results from the PLUTO Experiment on e^+e^- Reactions at High Energies", *eConf* **C790823** (1979) 19.
36. C. Berger, "Elementarteilchenphysik: Von den Grundlagen zu den modernen Experimenten". Springer Spektrum, 3 edition, 9, 2014.
37. CMS Collaboration, "Probing color coherence effects in pp collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **74** (2014) 2901, doi:10.1140/epjc/s10052-014-2901-8, arXiv:1311.5815.
38. CDF Collaboration, "Evidence for color coherence in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. D* **50** (1994) 5562, doi:10.1103/PhysRevD.50.5562.
39. D0 Collaboration, "Color coherent radiation in multijet events from $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Lett. B* **414** (1997) 419, doi:10.1016/S0370-2693(97)01190-8, arXiv:hep-ex/9706012.
40. D0 Collaboration, "Measurement of angular correlations of jets at $\sqrt{s} = 1.96$ TeV and determination of the strong coupling at high momentum transfers", *Phys. Lett. B* **718** (2012) 56, doi:10.1016/j.physletb.2012.10.003, arXiv:1207.4957.
41. T. Sjöstrand and M. van Zijl, "A Multiple Interaction Model for the Event Structure in Hadron Collisions", *Phys. Rev. D* **36** (1987) 2019, doi:10.1103/PhysRevD.36.2019.
42. B. Humpert and R. Odorico, "Multiparton Scattering and QCD Radiation as Sources of Four Jet Events", *Phys. Lett. B* **154** (1985) 211, doi:10.1016/0370-2693(85)90587-8.
43. L. Ametller, N. Paver, and D. Treleani, "Possible Signature of Multiple Parton Interactions in Collider Four Jet Events", *Phys. Lett. B* **169** (1986) 289, doi:10.1016/0370-2693(86)90668-4.
44. E. L. Berger, C. B. Jackson, and G. Shaughnessy, "Characteristics and Estimates of Double Parton Scattering at the Large Hadron Collider", *Phys. Rev. D* **81** (2010) 014014, doi:10.1103/PhysRevD.81.014014, arXiv:0911.5348.
45. M. Diehl, D. Ostermeier, and A. Schäfer, "Elements of a theory for multiparton interactions in QCD", *JHEP* **03** (2012) 089, doi:10.1007/JHEP03(2012)089, arXiv:1111.0910.
46. Axial Field Spectrometer Collaboration, "Double Parton Scattering in pp Collisions at $\sqrt{s} = 63$-GeV", *Z. Phys. C* **34** (1987) 163, doi:10.1007/BF01566757.
47. UA2 Collaboration, "A Study of multi - jet events at the CERN anti-p p collider and a search for double parton scattering", *Phys. Lett. B* **268** (1991) 145, doi:10.1016/0370-2693(91)90937-L.
48. CDF Collaboration, "Study of four jet events and evidence for double parton interactions in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. D* **47** (1993) 4857, doi:10.1103/PhysRevD.47.4857.

49. CMS Collaboration, "Measurement of four-jet production in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. D* **89** (2014) 092010, doi:10.1103/PhysRevD.89.092010, arXiv:1312.6440.
50. CDF Collaboration, "Double parton scattering in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$TeV", *Phys. Rev. D* **56** (1997) 3811, doi:10.1103/PhysRevD.56.3811.
51. D0 Collaboration, "Double parton interactions in photon+3 jet events in pp^- bar collisions $\sqrt{s} = 1.96$ TeV", *Phys. Rev. D* **81** (2010) 052012, doi:10.1103/PhysRevD.81.052012, arXiv:0912.5104.
52. D0 Collaboration, "Double parton interactions in photon + 3 jet and photon + b/c jet + 2 jet events in ppbar collisions at sqrts=1.96 TeV", *Phys. Rev. D* **89** (2014) 072006, doi:10.1103/PhysRevD.89.072006, arXiv:1402.1550.
53. ATLAS Collaboration, "Measurement of hard double-parton interactions in $W(\rightarrow l\nu)$+ 2 jet events at $\sqrt{s}$=7 TeV with the ATLAS detector", *New J. Phys.* **15** (2013) 033038, doi:10.1088/1367-2630/15/3/033038, arXiv:1301.6872.
54. CMS Collaboration, "Study of double parton scattering using W + 2-jet events in proton-proton collisions at $\sqrt{s}$ = 7 TeV", *JHEP* **03** (2014) 032, doi:10.1007/JHEP03(2014)032, arXiv:1312.5729.
55. C. Goebel, F. Halzen, and D. M. Scott, "Double Drell-Yan Annihilations in Hadron Collisions: Novel Tests of the Constituent Picture", *Phys. Rev. D* **22** (1980) 2789, doi:10.1103/PhysRevD.22.2789.
56. M. Mekhfi, "Multiparton processes: An Application to double Drell-Yan", *Phys. Rev. D* **32** (1985) 2371, doi:10.1103/PhysRevD.32.2371.
57. M. W. Krasny and W. Placzek, "On the contribution of the double Drell-Yan process to WW and ZZ production at the LHC", arXiv:1501.04569.
58. A. Kulesza and W. J. Stirling, "Like sign W boson production at the LHC as a probe of double parton scattering", *Phys. Lett. B* **475** (2000) 168, doi:10.1016/S0370-2693(99)01512-9, arXiv:hep-ph/9912232.
59. M. Bähr, M. Myska, M. H. Seymour, and A. Siodmok, "Extracting $\sigma_{\text{effective}}$ from the CDF gamma+3jets measurement", *JHEP* **03** (2013) 129, doi:10.1007/JHEP03(2013)129, arXiv:1302.4325.
60. D. Treleani, "Double parton scattering, diffraction and effective cross section", *Phys. Rev. D* **76** (2007) 076006, doi:10.1103/PhysRevD.76.076006, arXiv:0708.2603.
61. CMS Collaboration, "Double Parton Scattering cross section limit from same-sign W boson pair production in di-muon final state at LHC", technical report, CERN, 2015.
62. V. Del Duca and C. R. Schmidt, "Dijet production at large rapidity intervals", *Phys. Rev. D* **49** (1994) 4510, doi:10.1103/PhysRevD.49.4510, arXiv:hep-ph/9311290.
63. V. Del Duca and C. R. Schmidt, "BFKL versus O(α_s^3) corrections to large rapidity dijet production", *Phys. Rev. D* **51** (1995) 2150, doi:10.1103/PhysRevD.51.2150, arXiv:hep-ph/9407359.
64. D0 Collaboration, "The azimuthal decorrelation of jets widely separated in rapidity", *Phys. Rev. Lett.* **77** (1996) 595, doi:10.1103/PhysRevLett.77.595, arXiv:hep-ex/9603010.
65. D0 Collaboration, "Recent run II QCD results from D0", *AIP Conf. Proc.* **753** (2005) 92, doi:10.1063/1.1896692, arXiv:hep-ex/0411025.
66. D0 Collaboration, "Measurement of dijet azimuthal decorrelations at central rapidities in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. Lett.* **94** (2005) 221801, doi:10.1103/PhysRevLett.94.221801, arXiv:hep-ex/0409040.
67. M. Wobisch and K. Rabbertz, "Dijet azimuthal decorrelations for $\Delta\phi_{\text{dijet}} < 2\pi/3$ in perturbative QCD", *JHEP* **12** (2015) 024, doi:10.1007/JHEP12(2015)024, arXiv:1505.05030.
68. Z. Nagy, "Three jet cross-sections in hadron hadron collisions at next-to-leading order", *Phys. Rev. Lett.* **88** (2002) 122003, doi:10.1103/PhysRevLett.88.122003, arXiv:hep-ph/0110315.

69. Z. Nagy, "Next-to-leading order calculation of three-jet observables in hadron hadron collisions", *Phys. Rev. D* **68** (2003) 094002, doi:10.1103/PhysRevD.68.094002, arXiv:hep-ph/0307268.
70. CMS Collaboration, "Dijet Azimuthal Decorrelations in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **106** (2011) 122003, doi:10.1103/PhysRevLett.106.122003, arXiv:1101.5029.
71. ATLAS Collaboration, "Measurement of Dijet Azimuthal Decorrelations in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **106** (2011) 172002, doi:10.1103/PhysRevLett.106.172002, arXiv:1102.2696.
72. CMS Collaboration, "Measurement of dijet azimuthal decorrelations in pp collisions at $\sqrt{s} = 8$ TeV", (2015). arXiv:1602.04384. Submitted to *Eur. Phys. J. C*.
73. A. Banfi, M. Dasgupta, and Y. Delenda, "Azimuthal decorrelations between QCD jets at all orders", *Phys. Lett. B* **665** (2008) 86, doi:10.1016/j.physletb.2008.05.065, arXiv:0804.3786.
74. P. Sun, C. P. Yuan, and F. Yuan, "Soft Gluon Resummations in Dijet Azimuthal Angular Correlations in Hadronic Collisions", *Phys. Rev. Lett.* **113** (2014) 232001, doi:10.1103/PhysRevLett.113.232001, arXiv:1405.1105.
75. P. Sun, C. P. Yuan, and F. Yuan, "Transverse Momentum Resummation for Dijet Correlation in Hadronic Collisions", *Phys. Rev. D* **92** (2015) 094007, doi:10.1103/PhysRevD.92.094007, arXiv:1506.06170.
76. M. H. Seymour, "Jet shapes in hadron collisions: Higher orders, resummation and hadronization", *Nucl. Phys. B* **513** (1998) 269, doi:10.1016/S0550-3213(97)00711-6, arXiv:hep-ph/9707338.
77. CMS Collaboration, "Measurement of dijet azimuthal decorrelations in pp collisions at $\sqrt{s} = 8$ TeV", technical report, CERN, 2015.
78. ATLAS Collaboration, "Measurement of four-jet differential cross sections in $\sqrt{s} = 8$ TeV proton–proton collisions using the ATLAS detector", arXiv:1509.07335.
79. Z. Bern et al., "Four-Jet Production at the Large Hadron Collider at Next-to-Leading Order in QCD", *Phys. Rev. Lett.* **109** (2012) 042001, doi:10.1103/PhysRevLett.109.042001, arXiv:1112.3940.
80. S. Badger, B. Biedermann, P. Uwer, and V. Yundin, "NLO QCD corrections to multi-jet production at the LHC with a centre-of-mass energy of $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **718** (2013) 965, doi:10.1016/j.physletb.2012.11.029, arXiv:1209.0098.
81. S. Alioli et al., "Jet pair production in POWHEG", *JHEP* **04** (2011) 081, doi:10.1007/JHEP04(2011)081, arXiv:1012.3380.
82. A. Kardos, P. Nason, and C. Oleari, "Three-jet production in POWHEG", *JHEP* **04** (2014) 043, doi:10.1007/JHEP04(2014)043, arXiv:1402.4001.
83. M. Wobisch et al., "A new quantity for studies of dijet azimuthal decorrelations", *JHEP* **01** (2013) 172, doi:10.1007/JHEP01(2013)172, arXiv:1211.6773.
84. D0 Collaboration, "Measurement of the combined rapidity and p_T dependence of dijet azimuthal decorrelations in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Lett. B* **721** (2013) 212, doi:10.1016/j.physletb.2013.03.029, arXiv:1212.1842.
85. ATLAS Collaboration, "Measurement of charged-particle event shape variables in $\sqrt{s} = 7$ TeV proton-proton interactions with the ATLAS detector", *Phys. Rev. D* **88** (2013) 032004, doi:10.1103/PhysRevD.88.032004, arXiv:1207.6915.
86. S. Brandt, C. Peyrou, R. Sosnowski, and A. Wroblewski, "The Principal axis of jets. An Attempt to analyze high-energy collisions as two-body processes", *Phys. Lett.* **12** (1964) 57, doi:10.1016/0031-9163(64)91176-X.
87. J. D. Bjorken and S. J. Brodsky, "Statistical Model for electron-Positron Annihilation Into Hadrons", *Phys. Rev. D* **1** (1970) 1416, doi:10.1103/PhysRevD.1.1416.
88. E. Farhi, "A QCD Test for Jets", *Phys. Rev. Lett.* **39** (1977) 1587, doi:10.1103/PhysRevLett.39.1587.
89. A. De Rujula, J. R. Ellis, E. G. Floratos, and M. K. Gaillard, "QCD Predictions for Hadronic Final States in e^+e^- Annihilation", *Nucl. Phys. B* **138** (1978) 387, doi:10.1016/0550-3213(78)90388-7.

90. G. F. Sterman and S. Weinberg, "Jets from Quantum Chromodynamics", *Phys. Rev. Lett.* **39** (1977) 1436, doi:10.1103/PhysRevLett.39.1436.
91. G. Hanson et al., "Evidence for Jet Structure in Hadron Production by e+ e- Annihilation", *Phys. Rev. Lett.* **35** (1975) 1609, doi:10.1103/PhysRevLett.35.1609.
92. H. Georgi and M. Machacek, "A Simple QCD Prediction of Jet Structure in $^e + e^-$ Annihilation", *Phys. Rev. Lett.* **39** (1977) 1237, doi:10.1103/PhysRevLett.39.1237.
93. H1 Collaboration, "Measurement of event shape variables in deep inelastic e p scattering", *Phys. Lett. B* **406** (1997) 256, doi:10.1016/S0370-2693(97)00754-5, arXiv:hep-ex/9706002.
94. H1 Collaboration, "Investigation of power corrections to event shape variables measured in deep inelastic scattering", *Eur. Phys. J. C* **14** (2000) 255, doi:10.1007/s100520000344, arXiv:hep-ex/9912052.
95. H1 Collaboration, "Measurement of event shape variables in deep-inelastic scattering at HERA", *Eur. Phys. J. C* **46** (2006) 343, doi:10.1140/epjc/s2006-02493-x, arXiv:hep-ex/0512014.
96. ZEUS Collaboration, "Measurement of event shapes in deep inelastic scattering at HERA", *Eur. Phys. J. C* **27** (2003) 531, doi:10.1140/epjc/s2003-01148-x, arXiv:hep-ex/0211040.
97. ZEUS Collaboration, "Event shapes in deep inelastic scattering at HERA", *Nucl. Phys.* **B767** (2007) 1–28, doi:10.1016/j.nuclphysb.2006.05.016, arXiv:hep-ex/0604032.
98. M. Dasgupta and G. P. Salam, "Event shapes in e^+e^- annihilation and deep inelastic scattering", *J. Phys. G* **30** (2004) R143, doi:10.1088/0954-3899/30/5/R01, arXiv:hep-ph/0312283.
99. A. Banfi, G. P. Salam, and G. Zanderighi, "Resummed event shapes at hadron - hadron colliders", *JHEP* **08** (2004) 062, doi:10.1088/1126-6708/2004/08/062, arXiv:hep-ph/0407287.
100. A. Banfi, G. P. Salam, and G. Zanderighi, "Phenomenology of event shapes at hadron colliders", *JHEP* **06** (2010) 038, doi:10.1007/JHEP06(2010)038, arXiv:1001.4082.
101. CMS Collaboration, "Study of hadronic event-shape variables in multijet final states in pp collisions at $\sqrt{s}$ = 7 TeV", *JHEP* **10** (2014) 87, doi:10.1007/JHEP10(2014)087, arXiv:1407.2856.
102. R. K. Ellis, D. A. Ross, and A. E. Terrano, "The Perturbative Calculation of Jet Structure in e+ e- Annihilation", *Nucl. Phys. B* **178** (1981) 421, doi:10.1016/0550-3213(81)90165-6.
103. G. C. Fox and S. Wolfram, "Observables for the Analysis of Event Shapes in e+ e- Annihilation and Other Processes", *Phys. Rev. Lett.* **41** (1978) 1581, doi:10.1103/PhysRevLett.41.1581.
104. C. L. Basham, L. S. Brown, S. D. Ellis, and S. T. Love, "Energy Correlations in electron - positron Annihilation: Testing QCD", *Phys. Rev. Lett.* **41** (1978) 1585, doi:10.1103/PhysRevLett.41.1585.
105. C. L. Basham, L. S. Brown, S. D. Ellis, and S. T. Love, "Energy Correlations in electron-positron Annihilation in Quantum Chromodynamics: Asymptotically Free Perturbation Theory", *Phys. Rev. D* **19** (1979) 2018, doi:10.1103/PhysRevD.19.2018.
106. ATLAS Collaboration, "Measurement of transverse energy-energy correlations in multi-jet events in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector and determination of the strong coupling constant $\alpha_s(m_Z)$", *Phys. Lett. B* **750** (2015) 427, doi:10.1016/j.physletb.2015.09.050, arXiv:1508.01579.
107. UA1 Collaboration, "Hadronic Jet Production at the CERN Proton-Antiproton Collider", *Phys. Lett. B* **132** (1983) 214, doi:10.1016/0370-2693(83)90254-X.
108. UA2 Collaboration, "Jet measures and hadronic event shapes at the CERN $\bar{p}p$ collider", *Z. Phys. C* **36** (1987) 175, doi:10.1007/BF01579132.

109. CDF Collaboration, "Measurement of QCD jet broadening in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. D* **44** (1991) 601, doi:10.1103/PhysRevD.44.601.
110. D0 Collaboration, I. A. Bertram, "Jet results at the D0 experiment", in *Proceedings, 10th International Workshop on Deep-Inelastic Scattering (DIS 2002)*, volume 33, p. 3141. Krakow, Poland, April 30-May 4, 2002.
111. D0 Collaboration, V. D. Elvira, "Jet measurements at D0 using a k_T algorithm", in *Proceedings, 9th High-Energy Physics International Conference on Quantum chromodynamics (QCD'02)*, volume 121, p. 21. Montpellier, France, July 2-9, 2003. arXiv:hep-ex/0209073. doi:10.1016/S0920-5632(03)01805-X.
112. CDF Collaboration, "Measurement of Event Shapes in Proton-Antiproton Collisions at Center-of-Mass Energy 1.96 TeV", *Phys. Rev. D* **83** (2011) 112007, doi:10.1103/PhysRevD.83.112007, arXiv:1103.5143.
113. CMS Collaboration, "First Measurement of Hadronic Event Shapes in pp Collisions at $\sqrt(s) = 7$ TeV", *Phys. Lett. B* **699** (2011) 48, doi:10.1016/j.physletb.2011.03.060, arXiv:1102.0068.
114. ATLAS Collaboration, "Measurement of event shapes at large momentum transfer with the ATLAS detector in pp collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **72** (2012) 2211, doi:10.1140/epjc/s10052-012-2211-y, arXiv:1206.2135.
115. CMS Collaboration, "Event shapes and azimuthal correlations in Z + jets events in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **722** (2013) 238, doi:10.1016/j.physletb.2013.04.025, arXiv:1301.1646.
116. J. Currie, A. Gehrmann-De Ridder, E. W. N. Glover, and J. Pires, "NNLO QCD corrections to jet production at hadron colliders from gluon scattering", *JHEP* **01** (2014) 110, doi:10.1007/JHEP01(2014)110, arXiv:1310.3993.
117. R. Boughezal, C. Focke, X. Liu, and F. Petriello, "W-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD", *Phys. Rev. Lett.* **115** (2015) 062002, doi:10.1103/PhysRevLett.115.062002, arXiv:1504.02131.
118. A. G.-D. Ridder et al., "Precise QCD predictions for the production of a Z boson in association with a hadronic jet", arXiv:1507.02850.
119. R. Boughezal et al., "Z-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD", arXiv:1512.01291.
120. J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, "Jet substructure as a new Higgs search channel at the LHC", *Phys. Rev. Lett.* **100** (2008) 242001, doi:10.1103/PhysRevLett.100.242001, arXiv:0802.2470.
121. J. Thaler and L.-T. Wang, "Strategies to Identify Boosted Tops", *JHEP* **07** (2008) 092, doi:10.1088/1126-6708/2008/07/092, arXiv:0806.0023.
122. D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, "Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks", *Phys. Rev. Lett.* **101** (2008) 142001, doi:10.1103/PhysRevLett.101.142001, arXiv:0806.0848.
123. L. G. Almeida et al., "Substructure of high-pT Jets at the LHC", *Phys. Rev. D* **79** (2009) 074017, doi:10.1103/PhysRevD.79.074017, arXiv:0807.0234.
124. D. Krohn, J. Thaler, and L.-T. Wang, "Jet Trimming", *JHEP* **02** (2010) 084, doi:10.1007/JHEP02(2010)084, arXiv:0912.1342.
125. 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. D* **81** (2010) 094023, doi:10.1103/PhysRevD.81.094023, arXiv:0912.0033.
126. A. Altheimer et al., "Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks", *J. Phys. G* **39** (2012) 063001, doi:10.1088/0954-3899/39/6/063001, arXiv:1201.0008.
127. A. Altheimer et al., "Boosted objects and jet substructure at the LHC. Report of BOOST2012, held at IFIC Valencia, 23rd-27th of July 2012", *Eur. Phys. J. C* **74** (2014) 2792, doi:10.1140/epjc/s10052-014-2792-8, arXiv:1311.2708.
128. S. D. Ellis, Z. Kunszt, and D. E. Soper, "Jets at hadron colliders at order $\alpha - s^{3:}$ A Look inside", *Phys. Rev. Lett.* **69** (1992) 3615, doi:10.1103/PhysRevLett.69.3615, arXiv:hep-ph/9208249.

129. CDF Collaboration, "A Measurement of jet shapes in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **70** (1993) 713, doi:10.1103/PhysRevLett.70.713.
130. CDF Collaboration, "Study of jet shapes in inclusive jet production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. D* **71** (2005) 112002, doi:10.1103/PhysRevD.71.112002, arXiv:hep-ex/0505013.
131. D0 Collaboration, "Transverse energy distributions within jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Lett. B* **357** (1995) 500, doi:10.1016/0370-2693(95)00889-S.
132. ZEUS Collaboration, "Measurement of jet shapes in photoproduction at HERA", *Eur. Phys. J. C* **C2** (1998) 61, doi:10.1007/s100520050124, arXiv:hep-ex/9710002. DOI not accessible!
133. ZEUS Collaboration, "Measurement of jet shapes in high Q^2 deep inelastic scattering at HERA", *Eur. Phys. J. C* **C8** (1999) 367, doi:10.1007/s100520050471, arXiv:hep-ex/9804001.
134. ZEUS Collaboration, "Substructure dependence of jet cross sections at HERA and determination of α_s", *Nucl. Phys. B* **700** (2004) 3, doi:10.1016/j.nuclphysb.2004.08.049, arXiv:hep-ex/0405065.
135. H1 Collaboration, "Measurement of internal jet structure in dijet production in deep inelastic scattering at HERA", *Nucl. Phys. B* **545** (1999) 3, doi:10.1016/S0550-3213(99)00118-2, arXiv:hep-ex/9901010.
136. CDF Collaboration, "Measurement of b-jet Shapes in Inclusive Jet Production in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.96-TeV", *Phys. Rev. D* **78** (2008) 072005, doi:10.1103/PhysRevD.78.072005, arXiv:0806.1699.
137. CMS Collaboration, "Shape, Transverse Size, and Charged Hadron Multiplicity of Jets in pp Collisions at 7 TeV", *JHEP* **06** (2012) 160, doi:10.1007/JHEP06(2012)160, arXiv:1204.3170.
138. ATLAS Collaboration, "Study of Jet Shapes in Inclusive Jet Production in pp Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector", *Phys. Rev. D* **83** (2011) 052003, doi:10.1103/PhysRevD.83.052003, arXiv:1101.0070.
139. A. Banfi, G. P. Salam, and G. Zanderighi, "Infrared safe definition of jet flavor", *Eur. Phys. J. C* **47** (2006) 113, doi:10.1140/epjc/s2006-02552-4, arXiv:hep-ph/0601139.
140. M. H. Seymour, "The Average number of subjets in a hadron collider jet", *Nucl. Phys. B* **421** (1994) 545, doi:10.1016/0550-3213(94)90516-9.
141. M. H. Seymour, "The Subjet Multiplicity in Quark and Gluon Jets", *Phys. Lett. B* **378** (1996) 279, doi:10.1016/0370-2693(96)00399-1, arXiv:hep-ph/9603281.
142. J. R. Forshaw and M. H. Seymour, "Subjet rates in hadron collider jets", *JHEP* **09** (1999) 009, doi:10.1088/1126-6708/1999/09/009, arXiv:hep-ph/9908307.
143. D0 Collaboration, "Subjet multiplicity of gluon and quark jets reconstructed with the k_T algorithm in $p\bar{p}$ collisions", *Phys. Rev. D* **65** (2002) 052008, doi:10.1103/PhysRevD.65.052008, arXiv:hep-ex/0108054.
144. ZEUS Collaboration, "Measurement of subjet multiplicities in neutral current deep inelastic scattering at HERA and determination of α_s", *Phys. Lett. B* **558** (2003) 41, doi:10.1016/S0370-2693(03)00216-8, arXiv:hep-ex/0212030.
145. ZEUS Collaboration, "Subjet distributions in deep inelastic scattering at HERA", *Eur. Phys. J. C* **63** (2009) 527, doi:10.1140/epjc/s10052-009-1090-3, arXiv:0812.2864.
146. CMS Collaboration, "Measurement of the Subjet Multiplicity in Dijet Events from proton-proton Collisions at $\sqrt{s} = 7$ TeV", technical report, CERN, 2010.

Chapter 7
Future Perspectives

On June 3rd, 2015, the LHC resumed operations after a 2-years break for consolidation and upgrades. Most notably, the centre-of-mass energy was raised from 8 to 13 TeV. The second goal of reducing the time between two bunch crossings from 50 to 25 ns was achieved in the course of 2015. The original aim of providing $10\,\mathrm{fb}^{-1}$ already in the start-up year of Run 2 proved to be somewhat too ambitious. Nevertheless, $4.2\,\mathrm{fb}^{-1}$ could be accumulated, out of which more than $1\,\mathrm{fb}^{-1}$ were gathered in the last week of operation alone, cf. Fig. 7.1. For the first production year, 2016, $25\,\mathrm{fb}^{-1}$ are foreseen with the goal of delivering in total $100\,\mathrm{fb}^{-1}$ to the experiments during LHC Run 2 from 2015 until end of 2018.

In the following two sections, first results will be presented on new phenomena from dijets and on inclusive jet measurements that were released by the LHC experiments at or before the so-called end-of-year jamboree on December 15th, 2015. This is complemented with an outlook on determinations of the strong coupling constant.

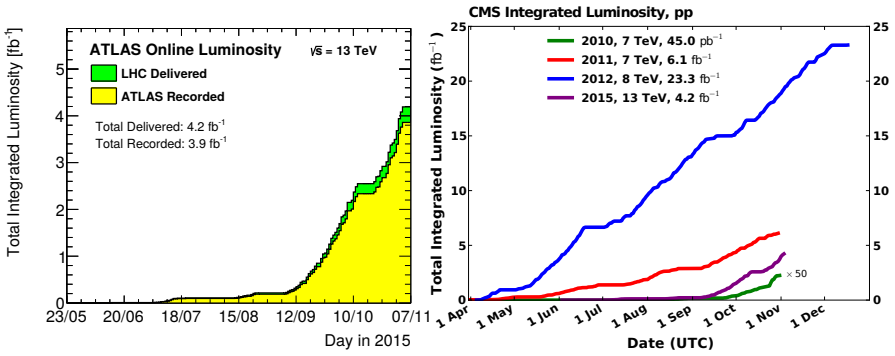

Fig. 7.1 Integrated pp luminosity recorded by ATLAS in 2015 and accumulated by CMS over all operational years of the LHC. *(Adapted from Source: CERN, ATLAS and CMS)*

7.1 Reach for New Phenomena

Among the first publications of ATLAS and CMS at the new energy frontier of $\sqrt{s} = 13\,\text{TeV}$ are searches for a resonant production of new particles with dijet final states [1, 2]. In this type of analysis, the dijet mass spectra are scrutinised for the appearance of peak-like structures on top of the SM background. These so-called "bump hunts" will not be further discussed here. Instead the repetition of the dijet angular analysis, cf. Section 6.2, as published by ATLAS in the same Ref. [2] for $3.6\,\text{fb}^{-1}$ of integrated luminosity at $\sqrt{s} = 13\,\text{TeV}$ is addressed. Figure 7.2 presents the corresponding χ distributions in comparison to SM predictions and to destructive ($\Lambda_{LL}^{+} = 12\,\text{TeV}$) and constructive CI terms ($\Lambda_{LL}^{-} = 17\,\text{TeV}$) in the four highest-mass ranges. Lacking any significant deviation from SM predictions including EW effects, the observed lower exclusion limits are raised by more than 45% from 8.1 to 12.0 TeV for destructive, and from 12.0 to 17.5 TeV for constructive interference, respectively. Preliminary results presented by CMS [3] are available for LO CI only and indicate a smaller increase of about 20–25% in the respective limits, alas for only a tenth of the amount of data recorded at $\sqrt{s} = 8\,\text{TeV}$ as compared to a fifth in the case of ATLAS.

For a historical perspective Table 7.1 presents the lower exclusion limits on CI for destructive and constructive interference with QCD as determined at hadron-hadron colliders starting with the UA2 Collaboration at the Sp$\bar{\text{p}}$S in 1984. The tested observables are listed in column four and comprise:

- the inclusive jet p_{T} spectrum, cf. Section 6.1,
- the ratio R_χ of the dijet cross section with a small separation in rapidity of the two leading jets over the one with a wide separation,
- the centrality ratio R_C, cf. Section 5.4,
- the scalar sum of jet transverse momenta H_{T},
- the dijet angular distribution χ, cf. Section 6.2,
- and the ratio F_χ of the dijet cross section with a small separation in rapidity of the two leading jets over the cross section integrated over χ.

The combination of NLO QCD with LO CI model predictions was subject to a couple of changes. The label "NLO($\times K_{\text{CI}}$)" in column five means that the NLO QCD prediction is multiplied by the ratio of a LO+CI over LO calculation. Implicitly, such way the SM K factor of NLO/LO QCD is applied to CI as well, which is not necessarily correct. The same happens explicitly for the case of "LO$\times K_{\text{QCD}}$" in column seven. The later ATLAS and the CMS studies ensure to apply QCD K factors to the SM part only or make use of the CI at NLO, which became available in 2011 [4, 5]. As visible from the CMS results, the exclusion limits are reduced by about 10% for CI at NLO as compared to LO only.

The initially employed χ^2 tests for the limit determinations were later replaced by a Bayesian method. Two priors were considered, one flat in the CI Lagrangian, i.e. $\propto \Lambda^{-2}$, and one flat in the CI cross section $\propto \Lambda^{-4}$, where the latter results in smaller exclusion limits. If both are provided, the more conservative limit is quoted in Table 7.1. Although Bayesian limits proved to be well-founded, the ambiguity in

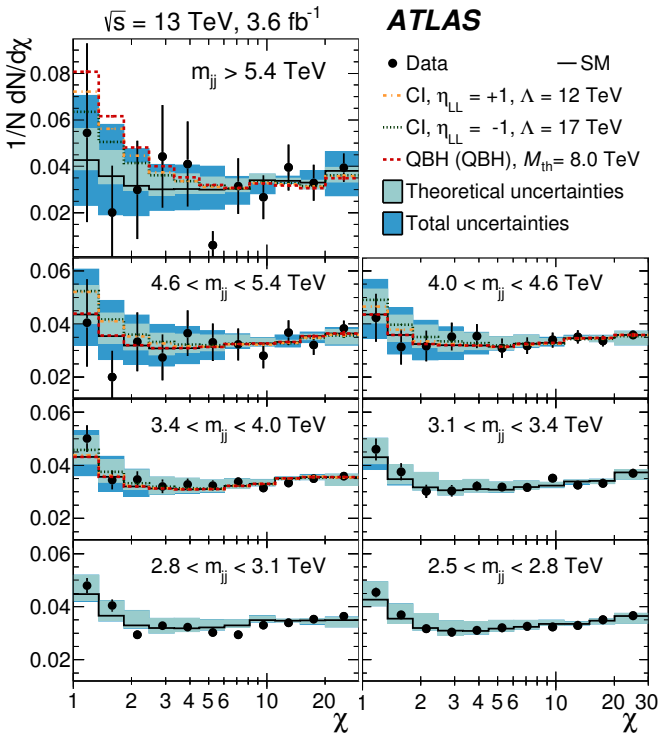

Fig. 7.2 Comparison of dijet angular distributions as measured at $\sqrt{s} = 13\,\text{TeV}$ by ATLAS in various dijet mass regions m_{jj} to predictions of the SM including EW corrections, and to the *LL* benchmark scenario for destructive (dash-dotted, $\Lambda^+ = 12\,\text{TeV}$) and constructive (dotted line, $\Lambda^- = 17\,\text{TeV}$) interference with the SM process. A model calculation involving quantum black holes (QBH, dashed line) is shown as well. Error bars represent the statistical uncertainties. The theoretical and the total, i.e. quadratically summed-up experimental and theoretical uncertainty, are displayed as shaded bands around the SM predictions. *(Taken from Ref. [2])*

the choice of prior assumptions lead to the adoption of the modified frequentist CL_s method that, as described in Section 5.4, avoids setting limits, where an experiment is not sensitive. For that reason, the CL_{s+b} method is not recommended.

For the early searches only one number is given for the two limits Λ_{LL}^+ and Λ_{LL}^-, because it is not clearly specified in the relevant publications. The CDF analysis from 1996 did not derive an exclusion limit but a preferred value for Λ_{LL}, since at that time a proper estimation of PDF uncertainties was not yet possible, cf. Section 6.1. In later studies usually both limits were given except for some cases, where it is claimed, based on previous results at the Tevatron, that they should differ by around 1% only, which in general is not correct.

Irrespective of such differences in details of the analyses listed in Table 7.1, it appears that a step-up in centre-of-mass energy has a much more significant impact on lower exclusion limits than an increase in integrated luminosity. The num-

Table 7.1 Historical development of lower exclusion limits on *LL* contact interaction scales for destructive (Λ_{LL}^{+}) and constructive interference (Λ_{LL}^{-}) with the QCD Lagrangian.

Experiment	$\sqrt{s}$ [TeV]	$\mathcal{L}_{\text{int}}$ [fb^{-1}]	Observable	QCD	EW	CI	stat. Method	Λ_{LL}^{+} [TeV]	Λ_{LL}^{-} [TeV]	Ref.
UA2	0.54	0.0001	$p_{\text{T,jet}}$	LO	—	LO	χ^2	0.275		[6]
UA2	0.63	0.0003	$p_{\text{T,jet}}$	LO	—	LO	χ^2	0.370		[7]
UA1	0.63	0.0003	R_χ	LO	—	LO	χ^2	0.415		[8]
CDF	1.8	$7\cdot10^{-6}$	$p_{\text{T,jet}}$	LO	—	LO	χ^2	0.700		[9]
UA2	0.63	0.0075	$p_{\text{T,jet}}$	LO	—	LO	χ^2	0.825		[10]
CDF	1.8	0.0042	$p_{\text{T,jet}}$	NLO	—	LO	χ^2	1.4		[11]
CDF	1.8	0.0042	R_χ	NLO	—	LO	χ^2	1.0		[12]
CDF	1.8	0.0195	$p_{\text{T,jet}}$	NLO	—	LO	χ^2	—		[13]
CDF	1.8	0.106	R_χ	NLO	—	LO×K_{QCD}	χ^2	1.8	1.6	[14]
D0	1.8	0.094	R_χ	NLO(×K_{CI})	—	LO	Bayesian	2.1	2.2	[15]
D0	1.8	0.0919	R_C	NLO(×K_{CI})	—	LO	Bayesian	2.7	2.4	[16]
D0	1.8	0.0919	H_{T}	NLO(×K_{CI})	—	LO	Bayesian	2.0	2.0	[17]
D0	1.8	0.0957	R_χ R_C	NLO(×K_{CI})	—	LO	Bayesian	2.0 2.4	1.9 2.4	[18]
D0	1.96	0.7	χ	NLO	—	LO×K_{QCD}	Bayesian χ^2	2.84 2.92	2.82 2.96	[19]
ATLAS	7	0.0031	R_C F_χ F_χ	NLO	—	LO×K_{QCD}	Bayesian Bayesian Neyman	2.0 3.2 3.4	—	[20]
CMS	7	0.0029	R_C	NLO	—	LO	CL$_s$	4.0	—	[21]
CMS	7	0.036	χ	NLO	—	LO	CL$_s$	5.6	6.7	[22]
ATLAS	7	0.031	$F_\chi(M_{\text{jj}})$ $F_\chi(M_{\text{jj}})$ F_χ χ	NLO	—	LO	CL$_{s+b}$ Bayesian Neyman CL$_{s+b}$	9.5 6.7 6.8 6.6	—	[23]
CMS	7	2.2	χ	NLO	—	LO NLO	CL$_s$	8.4 7.5	11.7 10.5	[24]
ATLAS	7	4.8	χ $F_\chi(M_{\text{jj}})$	NLO	—	LO	Bayesian	7.6 7.6	—	[25]
CMS	7	5	$p_{\text{T,jet}}$	NLO	—	LO	CL$_s$ Bayesian	9.9 10.1	14.3 14.1	[26]
CMS	8	19.7	χ	NLO	✓	LO NLO	CL$_s$	10.3 9.0	12.9 11.7	[27]
ATLAS	8	17.3	χ	NLO	✓	NLO	CL$_s$	8.1	12.0	[28]
ATLAS	13	3.6	χ	NLO	✓	NLO	CL$_s$	12.0	17.5	[2]
CMS prel.	13	2.6	χ	NLO	✓	LO	CL$_s$	12.1	16.3	[3]

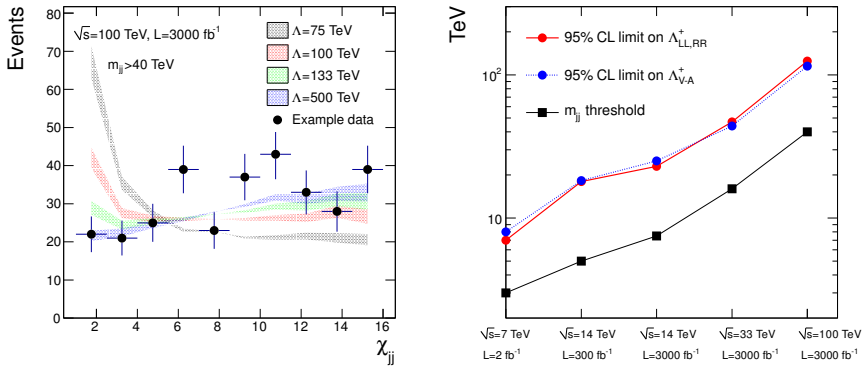

Fig. 7.3 Left: Dijet angular distribution at $\sqrt{s} = 100\,\text{TeV}$ for $3000\,\text{fb}^{-1}$ of integrated luminosity in comparison to a series of CI predictions for Λ^+_{LL} at LO. Right: Projected dijet mass thresholds and 95% CL exclusion limits for $\Lambda^+_{LL,RR}$ and Λ^+_{V-A} at LO for various combinations of centre-of-mass energy and integrated luminosity at a hadron-hadron collider. *(Taken from Ref. [29])*

ber of expected events for a scenario beyond the SM e.g. leading to CI is given as usual by $N = \mathcal{L}_{\text{int}} \cdot \sigma_{\text{CI}}(s)$ and scales linearly with the accumulated luminosity. When approaching a new mass scale, however, the production cross section can rise dramatically. Figure 7.3 presents projections from the 2013 Snowmass Study [29] for measurements of dijet angular distributions at a 100 TeV hadron-hadron collider and corresponding dijet mass thresholds and exclusion limits for $\Lambda^+_{LL,RR}$ at LO for various combinations of centre-of-mass energy and integrated luminosity. For $\sqrt{s} = 14\,\text{TeV}$ and $\mathcal{L}_{\text{int}} = 300\,\text{fb}^{-1}$ a lower exclusion limit of $\approx 19\,\text{TeV}$ can be read off Fig. 7.3 right. Comparing the last four entries of Table 7.1, the improvement in the LO and NLO limits for Λ^+_{LL} when going from 8 to 13 TeV centre-of-mass energy can be derived from the CMS and ATLAS results to $+1.8\,\text{TeV}$ and $+3.9\,\text{TeV}$. Considering 19 TeV as maximal reach for the CI limit, one concludes that with merely $\approx 1\%$ of $\mathcal{L}_{\text{int}}$ already 20-40% of the projected exclusion potential has been covered.

7.2 Jets at 13 TeV Centre-of-Mass Energy

Preliminary results on inclusive jets at 13 TeV centre-of-mass energy have been presented by ATLAS at central rapidity, $|y| < 0.5$, for an anti-k_t jet size $R = 0.4$ and by CMS up to a rapidity of $|y| = 4.7$ for the two jet sizes $R = 0.4$ and 0.7 from 78 respectively $72\,\text{pb}^{-1}$ of integrated luminosity [30,31]. This is the first time that both experiments employ at least one common jet size. The ratios of the CMS data to various theoretical calculations are presented in Fig. 7.4 for $|y| < 0.5$. The reach in transverse momentum does not yet exceed previous measurements at $\sqrt{s} = 8\,\text{TeV}$. Since PDFs are rather well known for the smaller fractional momenta x probed at

$\sqrt{s} = 13\,\text{TeV}$ as compared to 8 TeV for the same jet p_T, a very reasonable description of data by fixed-order predictions including NP corrections (top row of plots in Fig. 7.4) is recovered. For the smaller jet size $R = 0.4$, differences between data and theory are observed to be somewhat larger than for $R = 0.7$. The predictions from dijet NLO with matched PS and complemented with MPI and hadronisation from POWHEG +PYTHIA8 (bottom row) perform equally well for both jet sizes. This confirms at $\sqrt{s} = 13\,\text{TeV}$ the previous study on jet size dependence, discussed in Section 5.2 and reported in [32], where POWHEG +PYTHIA8 correctly described the R dependence of the inclusive jet cross section, while fixed-order predictions at NLO were insufficient in that respect. With the expected substantial increase in integrated luminosity up to $100\,\text{fb}^{-1}$, jet production can be tested in the regime of multi-TeV jet transverse momenta, which at the same time provides valuable information about the gluon PDF at large parton fractional momenta x.

7.3 Prospects for the Strong Coupling Constant

Numerous extractions of the strong coupling constant $\alpha_S(M_Z)$ have been performed at hadron colliders, in particular from jet cross sections. The latest results by the H1 and ZEUS experiments at the ep collider HERA and by the CDF and D0 experiments at the $p\bar{p}$ collider Tevatron are reported in Refs. [33–35] and [36–38], respectively. The range in scale Q for the running coupling constant $\alpha_S(Q)$ covered by these measurements runs from about 5 GeV up to 400 GeV. Evolving the fit results to the reference scale of the Z boson mass, $M_Z = 91.1876\,\text{GeV}$, $\alpha_S(M_Z)$ values between 0.1160 and 0.1206 are obtained. Within their rather large uncertainties up to 10%, they are compatible with each other. The dominant contributions to the uncertainty are of theoretical nature, namely scale uncertainties accounting for missing higher orders in the NLO predictions, and uncertainties caused by the limited knowledge of the proton PDFs.

With the advent of the LHC, unprecedented scales Q in the multi-TeV range become accessible. In addition, the modern particle detectors ATLAS and CMS in combination with more elaborate jet calibration and detector simulation techniques easily outperform the much older Tevatron experiments leading to jet calibration uncertainties even below percent level, cf. Chapter 3. $\alpha_S(M_Z)$ extractions using ATLAS data have been reported in Refs. [39] and [40]. The CMS Collaboration has determined $\alpha_S(Q)$ from scales Q close to the Z boson mass up to 1.4 TeV, which surpasses the pre-LHC upper limit of 400 GeV more than threefold. The fits employ jet cross sections as functions of the inclusive jet p_T [41] and the 3-jet mass m_3 [42], and the 3- to 2-jet cross section ratio R_{32} [43]. Again, scale and PDF uncertainties, amounting to roughly 4% and 2%, are the limiting factors preventing a better accuracy.

The potential gain through NNLO theory is demonstrated in another analysis by CMS, where for the first time the $t\bar{t}$ cross section is used to extract the strong coupling constant [44]. Theoretical calculations, available at NNLO+NNLL (next-to-

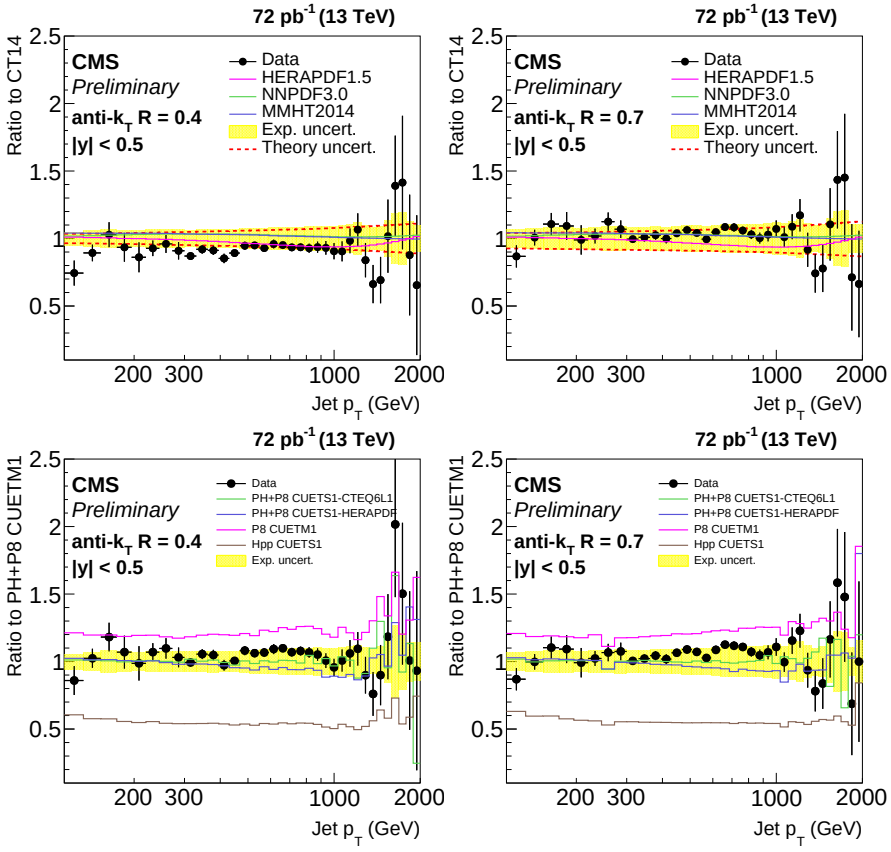

Fig. 7.4 Ratio of data to NLO predictions from NLOJET++ for the CT14-NLO PDF (top row) and to NLO with matched PS, MPI, and hadronisation from POWHEG +PYTHIA8 tune CUETM1 (bottom row). The inclusive jet cross section at central rapidity $|y| < 0.5$ is compared for the two anti-k_t jet sizes $R = 0.4$ (left column) and $R = 0.7$ (right column). The error bars correspond to the statistical uncertainty of the data and the shaded band to their total systematic uncertainty. Additional predictions are shown from NLO for three alternative PDF sets (top), and from POWHEG +PYTHIA8, PYTHIA8, and HERWIG++ for various tunes. *(Taken from Ref. [31])*

next-to-leading-logarithmic) [45], reduce the scale uncertainty in this case to 0.7%. This is the first hadron-hadron collider observable that has been included into the derivation of the world average of $\alpha_S(M_Z)$ at this level of accuracy. Significant further progress can therefore be expected for the near future considering that NNLO predictions have been completed recently for W+jet and Z+jet final states and are partially complete for inclusive jet and dijet production [46–49]. EW corrections that become similarly important at the TeV scale, which is easily within reach of the inclusive jet and dijet final states, are known as well [50].

Figure 7.5 left presents the timeline of $\alpha_S(M_Z)$ determinations over the last decades with indicators of the main developments leading to an improved precision.

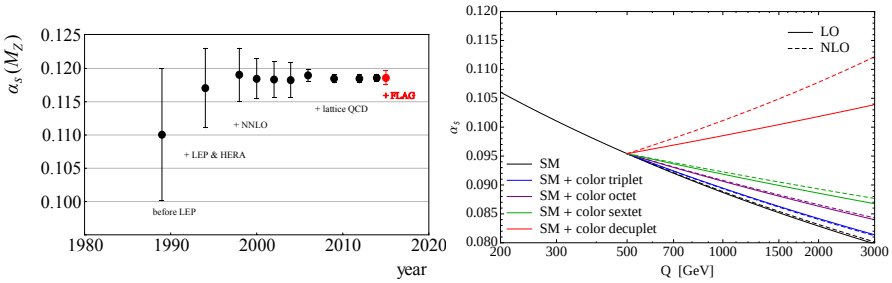

Fig. 7.5 Left: Timeline of $\alpha_S(M_Z)$ determinations over the last decades with indicators of the main developments leading to an improved precision. The plot from Ref. [54] has been complemented with the latest 2015 update to Ref. [52]. Right: Example of a deviation from the SM $\alpha_S(Q)$ evolution as induced by a new fermion of mass 500 GeV in various representations of the colour gauge group. The evolution of $\alpha_S(Q)$ is performed at NLO with LO shown for comparison. *(Left: Adapted from Ref. [54]; Right: taken from Ref. [55])*

The oldest entry from 1989 corresponds to the first "Summary and Conclusion on α_S" by G. Altarelli in Ref. [51]. Lacking NNLO theory, most of the hadron-hadron collider data are excluded from the world average since almost twenty years. This changed only with the inclusion of the CMS result obtained from $t\bar{t}$ production into the newest derivation of the world average of $\alpha_s(M_Z) = 0.1181 \pm 0.0013$ in the 2015 update of Ref. [52]. Remarkably, the uncertainty on $\alpha_S(M_Z)$ has doubled compared to the previous world average of Ref. [52], because the rather aggressive uncertainty estimations by individual groups working with lattice gauge theory were revised by the FLAG working group [53]. Their new uncertainty estimation amounts to about 1%.

Table 7.2 summarises the uncertainties for the hadron collider determinations of $\alpha_S(M_Z)$. Except for scale uncertainties where always the larger deviation is considered (scale), asymmetric uncertainties have been symmetrised for simpler comparison and are presented as a $\pm$ percental uncertainty on the respective central value. Separately detailed uncertainties of experimental origin, including statistical ones, are quadratically added together (exp). The same is done for multiple uncertainties caused by similar sources like a PDF uncertainty quoted for one PDF set and an additional uncertainty quoted for using different PDF sets (PDF). The entry of "scl" in the column for the non-perturbative effects from multiple-parton interactions and hadronisation indicates their inclusion within the scale uncertainty. The extra column "other" lists additional uncertainties that have been attributed to the choice of an infrared-unsafe(!) jet algorithm (CDF), to the choice of the jet size $R = 0.4$ or 0.6 (ATLAS incl. jets), and to the top pole-mass M_t^{pole} as input parameter (CMS $\sigma(t\bar{t})$).

Within CMS the focus was on testing the running of $\alpha_S(Q)$ up to the highest scales possible. As shown previously in Fig. 4.15 of Section 4.4, no indication of a significant deviation was found up to 1.4 TeV. A preliminary result from inclusive jets at 8 TeV increases the investigated range in scale Q even to 1.5 TeV [56]. With more data at $\sqrt{s} = 13$ TeV it is expected to extend these tests into the multi-TeV

Table 7.2 Summary of latest $\alpha_S(M_Z)$ determinations at hadron colliders. For each process the power in α_S of the leading order (LO), the centre-of-mass energy, the accessed range of scale Q, and the number of fitted data points is given. Theory is used at NLO accuracy except for the D0 inclusive jets, where additional threshold corrections are considered, and for CMS $\sigma(t\bar{t})$, where theory is known to NNLO+NNLL precision. The experimental, PDF, scale, NP, and additional uncertainties are presented as percental uncertainties on the extracted $\alpha_S(M_Z)$ value. All numbers are derived from the given references.

Process	LO	$\sqrt{s}$	Q	N_p	$\alpha_s(m_Z)$	$\Delta\alpha_s(m_Z)/\alpha_s(m_Z)$ [%]					Ref.
ep, $p\bar{p}$, pp	α_s^n	[TeV]	[GeV]			exp	PDF	scale	NP	other	
H1 jets low Q^2	1	0.32	5-57	62	0.1160	1.2	1.4	8.0	scl	-	[33]
ZEUS γp jets	1	0.32	21-71	18	0.1206	1.9	1.9	2.5	0.4	-	[34]
H1 jets high Q^2	1	0.32	10-94	64	0.1165	0.7	0.8	3.1	0.7	-	[35]
CDF incl. jets	2	1.8	40-250	27	0.1178	7.5	5.0	5.0	-	2.5	[36]
D0 incl. jets	2	1.96	50-145	22	0.1161	2.9	1.0	2.5	1.1	-	[37]
D0 ang. corr.	1	1.96	50-450	102	0.1191	0.7	1.2	5.5	0.1	-	[38]
ATLAS incl. jets	2	7	45-600	42	0.1151	4.3	1.8	3.8	1.9	5.2	[39]
ATLAS EEC	1	7	250-1300	22	0.1173	0.9	1.4	5.4	0.2	-	[40]
CMS incl. jets	2	7	114-2116	133	0.1185	1.6	2.4	4.5	0.3	-	[41]
CMS 3-jet mass	3	7	332-1635	46	0.1171	1.1	2.0	5.9	0.7	-	[42]
CMS $R_{3/2}$	1	7	420-1390	21	0.1148	1.2	1.6	4.4	scl	-	[43]
CMS $\sigma(t\bar{t})$	2	7	M_t^{pole}	1	0.1151	2.2	1.5	0.7	-	1.1	[44]

range, where new coloured matter potentially changes the running of α_S. Figure 7.5 right gives examples for such deviations from the SM evolution of $\alpha_S(Q)$ as induced by a new fermion of mass 500 GeV in various representations of the colour gauge group [55]. As not only the running of $\alpha_S(Q)$ but also the DGLAP evolution of the PDFs is modified by the existence of new coloured matter, the CMS jet cross section ratio R_{32} was found to be particularly useful in that respect because of its reduced dependence on PDFs and consequently has been employed in Ref. [55] to set limits on such models of new phenomena. The best observables and strategies to detect such deviations are subject to future developments.

Moreover, there is still a significant optimisation potential for a precision determination of $\alpha_S(M_Z)$ within CMS. Concentrating on the best understood detector parts at central rapidity and within a jet p_{T} range restricted to the most precisely calibrated jets, experimental uncertainties can be further reduced. At the same time PDFs are more precisely known and the uncomfortable increase in uncertainty when evolving downwards to $Q = M_Z$ is smaller for medium jet p_{T}. With respect to PDFs, one must assure though that the same data are not used twice, first for constraining the proton structure and then to extract the strong coupling constant with these PDFs. The latter effect of a change in relative precision roughly proportional to $\alpha_S(M_Z)/\alpha_S(Q)$ [57] is also known under the name "incredibly shrinking error", which refers to the fact that an $\alpha_S(Q)$ determination to 4.2% accuracy at the τ mass of 1.777 GeV translates into a relative precision of 1.3% at the Z boson mass.

The technique of cross section ratios to cancel at least partially some systematic effects might be helpful as well. Corresponding cross sections are known to NLO up to jet multiplicities of five [58, 59]. For ratios like R_{32} at NNLO the 3-jet cross sections, however, are not yet available at this level of accuracy.

The extraction of the strong coupling constant $\alpha_S(M_Z)$ at hadron colliders, in particular from jet cross sections, has a long tradition, but was hampered by the complex and strongly interacting initial states and theoretical predictions that were limited in accuracy to NLO. This will be cured in the near future advancing the precision to the one percent level! Further improvements can be obtained by addressing all heretofore subdominant uncertainties including non-perturbative effects. For the ultimate precision possible with hadron collisions further refinements of combining multiple observables to disentangle e.g. correlations between the top mass M_t, the gluon PDF $g(x)$, and the strong coupling constant $\alpha_S(M_Z)$, and combining multiple experiments, i.e. ATLAS, CMS, and the HERA and Tevatron experiments, will be necessary.

References

1. CMS Collaboration, "Search for narrow resonances decaying to dijets in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Phys. Rev. Lett.* **116** (2016) 071801, doi:10.1103/PhysRevLett.116.071801, arXiv:1512.01224.
2. ATLAS Collaboration, "Search for New Phenomena in Dijet Mass and Angular Distributions from pp Collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS Detector", *Phys. Lett. B* **754** (2016) 302, doi:10.1016/j.physletb.2016.01.032, arXiv:1512.01530.
3. CMS Collaboration, "Searches for quark contact interactions and extra spatial dimensions with dijet angular distributions in proton proton collisions at 13 TeV", technical report, CERN, 2015.
4. J. Gao et al., "Next-to-leading QCD effect to the quark compositeness search at the LHC", *Phys. Rev. Lett.* **106** (2011) 142001, doi:10.1103/PhysRevLett.106.142001, arXiv:1101.4611.
5. J. Gao, C. S. Li, and C. P. Yuan, "NLO QCD Corrections to dijet Production via Quark Contact Interactions", *JHEP* **07** (2012) 037, doi:10.1007/JHEP07(2012)037, arXiv:1204.4773.
6. UA2 Collaboration, "Measurement of Very Large Transverse Momentum Jet Production at the CERN anti-p p Collider", *Phys. Lett. B* **138** (1984) 430, doi:10.1016/0370-2693(84)91935-X.
7. UA2 Collaboration, "Measurement of the $\sqrt{(s)}$ Dependence of Jet Production at the CERN anti-p p Collider", *Phys. Lett. B* **160** (1985) 349, doi:10.1016/0370-2693(85)91341-3.
8. UA1 Collaboration, "Angular Distributions for High Mass Jet Pairs and a Limit on the Energy Scale of Compositeness for Quarks from the CERN p anti-p Collider", *Phys. Lett. B* **177** (1986) 244, doi:10.1016/0370-2693(86)91065-8.
9. CDF Collaboration, "Measurement of the inclusive jet cross-section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **62** (1989) 613, doi:10.1103/PhysRevLett.62.613.
10. UA2 Collaboration, "Inclusive jet cross-section and a search for quark compositeness at the CERN $\bar{p}p$ collider", *Phys. Lett. B* **257** (1991) 232, doi:10.1016/0370-2693(91)90887-V.

11. CDF Collaboration, "Inclusive jet cross-section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **68** (1992) 1104, doi:10.1103/PhysRevLett.68.1104.
12. CDF Collaboration, "The Dijet angular distribution in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **69** (1992) 2896, doi:10.1103/PhysRevLett.69.2896.
13. CDF Collaboration, "Inclusive jet cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **77** (1996) 438, doi:10.1103/PhysRevLett.77.438, arXiv:hep-ex/9601008.
14. CDF Collaboration, "Measurement of Dijet Angular Distributions by the Collider Detector at Fermilab", *Phys. Rev. Lett.* **77** (1996) 5336, doi:10.1103/PhysRevLett.77.5336, arXiv:hep-ex/9609011. [Erratum-ibid. **78**, 4307 (1997)], author list incomplete.
15. D0 Collaboration, "Measurement of dijet angular distributions and search for quark compositeness", *Phys. Rev. Lett.* **80** (1998) 666, doi:10.1103/PhysRevLett.80.666, arXiv:hep-ex/9707016.
16. D0 Collaboration, "The dijet mass spectrum and a search for quark compositeness in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV", *Phys. Rev. Lett.* **82** (1999) 2457, doi:10.1103/PhysRevLett.82.2457, arXiv:hep-ex/9807014.
17. D0 Collaboration, "Limits on quark compositeness from high-energy jets in $\bar{p}p$ collisions at 1.8 TeV", *Phys. Rev. D* **62** (2000) 031101, doi:10.1103/PhysRevD.62.031101, arXiv:hep-ex/9912023.
18. D0 Collaboration, "High-p_T jets in $\bar{p}p$ collisions at $\sqrt{s} = 630$ GeV and 1800 GeV", *Phys. Rev. D* **64** (2001) 032003, doi:10.1103/PhysRevD.64.032003, arXiv:hep-ex/0012046.
19. D0 Collaboration, "Measurement of dijet angular distributions at $\sqrt{s} = 1.96$ TeV and searches for quark compositeness and extra spatial dimensions", *Phys. Rev. Lett.* **103** (2009) 191803, doi:10.1103/PhysRevLett.103.191803, arXiv:0906.4819.
20. ATLAS Collaboration, "Search for Quark Contact Interactions in Dijet Angular Distributions in pp Collisions at $\sqrt{s} = 7$ TeV Measured with the ATLAS Detector", *Phys. Lett. B* **694** (2011) 327, doi:10.1016/j.physletb.2010.10.021, arXiv:1009.5069.
21. CMS Collaboration, "Search for Quark Compositeness with the Dijet Centrality Ratio in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **105** (2010) 262001, doi:10.1103/PhysRevLett.105.262001, arXiv:1010.4439.
22. CMS Collaboration, "Measurement of Dijet Angular Distributions and Search for Quark Compositeness in pp Collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **106** (2011) 201804, doi:10.1103/PhysRevLett.106.201804, arXiv:1102.2020.
23. ATLAS Collaboration, "Search for New Physics in Dijet Mass and Angular Distributions in pp Collisions at $\sqrt{s} = 7$ TeV Measured with the ATLAS Detector", *New J. Phys.* **13** (2011) 053044, doi:10.1088/1367-2630/13/5/053044, arXiv:1103.3864.
24. CMS Collaboration, "Search for quark compositeness in dijet angular distributions from pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **05** (2012) 055, doi:10.1007/JHEP05(2012)055, arXiv:1202.5535.
25. ATLAS Collaboration, "ATLAS search for new phenomena in dijet mass and angular distributions using pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2013) 029, doi:10.1007/JHEP01(2013)029, arXiv:1210.1718.
26. CMS Collaboration, "Search for contact interactions using the inclusive jet p_T spectrum in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. D* **87** (2013) 052017, doi:10.1103/PhysRevD.87.052017, arXiv:1301.5023.
27. CMS Collaboration, "Search for quark contact interactions and extra spatial dimensions using dijet angular distributions in proton–proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **746** (2015) 79, doi:10.1016/j.physletb.2015.04.042, arXiv:1411.2646.
28. ATLAS Collaboration, "Search for New Phenomena in Dijet Angular Distributions in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV Measured with the ATLAS Detector", *Phys. Rev. Lett.* **114** (2015) 221802, doi:10.1103/PhysRevLett.114.221802, arXiv:1504.00357.

29. Y. Gershtein et al., "Working Group Report: New Particles, Forces, and Dimensions", in *Proceedings, Community Summer Study 2013: Snowmass on the Mississippi (CSS2013)*. Minneapolis, MN, USA, July 29-August 6, 2013. arXiv:1311.0299.
30. ATLAS Collaboration, "Measurement of the inclusive-jet cross section in proton-proton collisions at 13 TeV centre-of-mass energy with the ATLAS detector", technical report, CERN, 2015.
31. CMS Collaboration, "Measurement of the double-differential inclusive jet cross section at $\sqrt{s}$ = 13 TeV", technical report, CERN, 2015.
32. CMS Collaboration, "Measurement of the ratio of inclusive jet cross sections using the anti-k_T algorithm with radius parameters $R = 0.5$ and 0.7 in pp collisions at $\sqrt{s}$ = 7 TeV", *Phys. Rev. D* **90** (2014) 072006, doi:10.1103/PhysRevD.90.072006, arXiv:1406.0324.
33. H1 Collaboration, "Jet Production in ep Collisions at Low Q^2 and Determination of α_s", *Eur. Phys. J. C* **67** (2010) 1, doi:10.1140/epjc/s10052-010-1282-x, arXiv:0911.5678.
34. ZEUS Collaboration, "Inclusive-jet photoproduction at HERA and determination of α_s", *Nucl. Phys. B* **864** (2012) 1, doi:10.1016/j.nuclphysb.2012.06.006, arXiv:1205.6153.
35. H1 Collaboration, "Measurement of multijet production in ep collisions at high Q^2 and determination of the strong coupling α_s", *Eur. Phys. J. C* **75** (2015) 65, doi:10.1140/epjc/s10052-014-3223-6, arXiv:1406.4709.
36. CDF Collaboration, "Measurement of the strong coupling constant from inclusive jet production at the Tevatron $\bar{p}p$ collider", *Phys. Rev. Lett.* **88** (2002) 042001, doi:10.1103/PhysRevLett.88.042001, arXiv:hep-ex/0108034.
37. D0 Collaboration, "Determination of the strong coupling constant from the inclusive jet cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. D* **80** (2009) 111107, doi:10.1103/PhysRevD.80.111107, arXiv:0911.2710.
38. D0 Collaboration, "Measurement of angular correlations of jets at $\sqrt{s} = 1.96$ TeV and determination of the strong coupling at high momentum transfers", *Phys. Lett. B* **718** (2012) 56, doi:10.1016/j.physletb.2012.10.003, arXiv:1207.4957.
39. B. Malaescu and P. Starovoitov, "Evaluation of the strong coupling constant α_S using the ATLAS inclusive jet cross-section data", *Eur. Phys. J. C* **72** (2012) 2041, doi:10.1140/epjc/s10052-012-2041-y, arXiv:1203.5416.
40. ATLAS Collaboration, "Measurement of transverse energy-energy correlations in multi-jet events in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector and determination of the strong coupling constant $\alpha_s(m_Z)$", *Phys. Lett. B* **750** (2015) 427, doi:10.1016/j.physletb.2015.09.050, arXiv:1508.01579.
41. CMS Collaboration, "Constraints on parton distribution functions and extraction of the strong coupling constant from the inclusive jet cross section in pp collisions at $\sqrt{s}$ = 7 TeV", *Eur. Phys. J. C* **75** (2015) 288, doi:10.1140/epjc/s10052-015-3499-1, arXiv:1410.6765.
42. CMS Collaboration, "Measurement of the inclusive 3-jet production differential cross section in proton-proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range", *Eur. Phys. J. C* **75** (2015) 186, doi:10.1140/epjc/s10052-015-3376-y, arXiv:1412.1633.
43. CMS Collaboration, "Measurement of the ratio of the inclusive 3-jet cross section to the inclusive 2-jet cross section in pp collisions at $\sqrt{s}$ = 7 TeV and first determination of the strong coupling constant in the TeV range", *Eur. Phys. J. C* **73** (2013) 2604, doi:10.1140/epjc/s10052-013-2604-6, arXiv:1304.7498.
44. CMS Collaboration, "Determination of the top-quark pole mass and strong coupling constant from the $\mathrm{t\bar{t}}$ production cross section in pp collisions at $\sqrt{s}$ = 7 TeV", *Phys. Lett. B* **728** (2014) 496, doi:10.1016/j.physletb.2013.12.009, arXiv:1307.1907.
45. M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_s^4)$", *Phys. Rev. Lett.* **110** (2013) 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

46. J. Currie, A. Gehrmann-De Ridder, E. W. N. Glover, and J. Pires, "NNLO QCD corrections to jet production at hadron colliders from gluon scattering", *JHEP* **01** (2014) 110, doi:10.1007/JHEP01(2014)110, arXiv:1310.3993.
47. R. Boughezal, C. Focke, X. Liu, and F. Petriello, "W-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD", *Phys. Rev. Lett.* **115** (2015) 062002, doi:10.1103/PhysRevLett.115.062002, arXiv:1504.02131.
48. A. G.-D. Ridder et al., "Precise QCD predictions for the production of a Z boson in association with a hadronic jet", arXiv:1507.02850.
49. R. Boughezal et al., "Z-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD", arXiv:1512.01291.
50. S. Dittmaier, A. Huss, and C. Speckner, "Weak radiative corrections to dijet production at hadron colliders", *JHEP* **11** (2012) 095, doi:10.1007/JHEP11(2012)095, arXiv:1210.0438.
51. G. Altarelli, "Experimental Tests of Perturbative QCD", *Ann. Rev. Nucl. Part. Sci.* **39** (1989) 357, doi:10.1146/annurev.ns.39.120189.002041.
52. K. A. Olive and others (Particle Data Group), "Review of Particle Physics", *Chin. Phys. C* **38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.
53. S. Aoki et al., "Review of lattice results concerning low-energy particle physics", *Eur. Phys. J. C* **74** (2014) 2890, doi:10.1140/epjc/s10052-014-2890-7, arXiv:1310.8555.
54. R. Alemany-Fernandez et al., "The Large Hadron Collider: Harvest of Run 1". Springer, Berlin, Germany, 1st edition, May, 2015.
55. D. Becciolini et al., "Constraining new colored matter from the ratio of 3 to 2 jets cross sections at the LHC", *Phys. Rev. D* **91** (2015) 015010, doi:10.1103/PhysRevD.91.015010, arXiv:1403.7411.
56. CMS Collaboration, "Measurement of the double-differential inclusive jet cross section at $\sqrt{s} = 8$ TeV", technical report, CERN, 2015.
57. B. R. Webber, "QCD and jet physics", arXiv:hep-ph/9410268.
58. Z. Bern et al., "Four-Jet Production at the Large Hadron Collider at Next-to-Leading Order in QCD", *Phys. Rev. Lett.* **109** (2012) 042001, doi:10.1103/PhysRevLett.109.042001, arXiv:1112.3940.
59. S. Badger, B. Biedermann, P. Uwer, and V. Yundin, "Next-to-leading order QCD corrections to five jet production at the LHC", *Phys. Rev. D* **89** (2014) 034019, doi:10.1103/PhysRevD.89.034019, arXiv:1309.6585.

Chapter 8
Summary and Outlook

Jets are versatile tools that are produced abundantly in high-p_{T} hadron-hadron collisions. When energy frontiers are probed they enable searches for new phenomena at the highest scales achievable. Signatures in the form of contact interactions could be detected in jet final states up to around 19 TeV in Run 2 of the LHC.

Owing to the unprecedented experimental precision reached with the new detectors at the LHC, jet measurements have evolved into precision tests of QCD. Such tests must be accompanied by equally accurate theoretical predictions, a fact that has sparked rapid progress in the field of perturbative calculations using both analytic methods and modern Monte Carlo event generators. The potential of some of these new possibilities like cross sections for multi-jet topologies at NLO still needs to be fully exploited, while other developments like the NNLO calculation for jet production are eagerly awaited.

From the interplay between experiment and theory in a multitude of subjects ranging from nonperturbative effects over parton showers and parton distribution functions to the strong coupling constant a more refined and detailed picture of QCD has emerged than ever before. Last but not least the much improved understanding of the QCD dynamics including the ability to precisely predict even complicated high-multiplicity final states helps estimating the background in many searches for new phenomena.

Of particular interest is the unique potential of jet measurements to better characterise the gluon parton distribution function in the proton and to determine the strong coupling constant, both of which represent significant uncertainties in predictions for the Higgs boson production via the gluon fusion channel. The current knowledge on $\alpha_S(M_Z)$ is dominated by derivations from lattice gauge theory with an estimated precision of around 1%. Considering the recent experimental and theoretical developments, a similar accuracy should be achievable within Run 2 of the LHC from hadron collider measurements alone, which currently are limited to 3–5% of precision mostly because of lacking theory ingredients. An accuracy below 1% will be difficult (but not impossible) to reach, since multiple sources including the modelling of nonperturbative effects contribute at this level to the total uncertainty. Global projections discussed in Refs. [1–4] conclude on a prospect of further

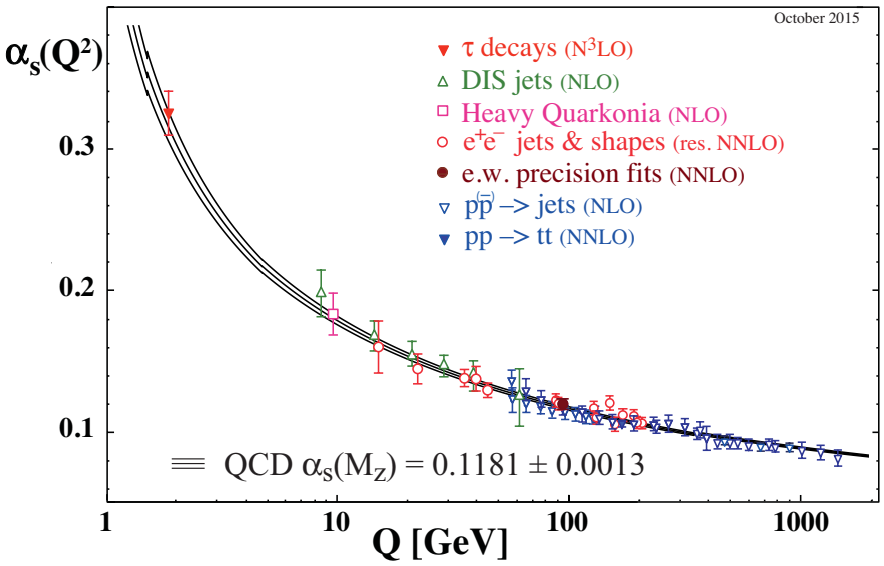

Fig. 8.1 Running of the strong coupling constant as of 2015. Determinations of the strong coupling constant α_S are shown as a function of the relevant energy scale Q of the respective process. By including LHC data, the range in Q could be extended beyond the TeV scale. The small uncertainty of a first result from $t\bar{t}$ production at the LHC with theory at NNLO demonstrates the future potential. *(Taken from the 2015 update of Ref. [5])*

reductions to $\approx 0.3\%$ of uncertainty from lattice gauge theory within the next five years and on a timescale of 10–20 years to $\approx 0.1\%$ from a Giga-Z program at a future e^+e^- collider, or from an ep collider like LHeC.

Precision studies of the running of $\alpha_S(Q)$ are equally important, since indications on new phenomena might well become visible through a modified evolution of the strong coupling and the PDFs. The latest status is shown in Figure 8.1, where thanks to the LHC jet data the range in Q has been extended beyond the TeV scale.

References

1. D. d'Enterria and P. Skands, eds., "Proceedings of the Workshop on High-precision α_s measurements: From LHC to FCC-ee, October 12-13, 2015". Geneva, Switzerland, (2015). arXiv:1512.05194.
2. J. M. Campbell et al., "Working Group Report: Quantum Chromodynamics", in *Proceedings, Community Summer Study 2013: Snowmass on the Mississippi (CSS2013)*. Minneapolis, MN, USA, July 29-August 6, 2013. arXiv:1310.5189.
3. M. Bardeen et al., "Planning the Future of U.S. Particle Physics: Report of the 2013 Community Summer Study of the APS Division of Particles and Fields",.
4. ESPPPG Collaboration, "Physics Briefing Book: Input for the Strategy Group to draft the update of the European Strategy for Particle Physics",.

5. K. A. Olive and others (Particle Data Group), "Review of Particle Physics", *Chin. Phys. C* **38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.

List of Figures

List of Tables

About this book

Particle jets are produced abundantly in hadron-hadron collisions. At the LHC, they probe the highest energies for new phenomena. Likewise, owing to the unprecedented accuracy attained by the LHC experiments, jet measurements have evolved into precision tests of QCD, the theory of the strong interaction. A key prediction, the running of the strong coupling constant, has been established up to energy scales three times larger than previously possible.

This book reviews the latest experimental results on jet physics at the LHC by means of a multitude of observables illuminating different aspects of QCD like the interplay with predictions involving new coloured matter, the coupling strength, or the gluon content of the proton, which constitutes a significant source of uncertainty in predictions of Higgs boson production. A novel approach used in this book is the classification of these observables into the three categories of absolute, ratio, and shape measurements; the advantages of each category and the differences between them are clearly pointed out. The significance of each observable is elaborated with the help of numerous illustrations.

About the author

Klaus Rabbertz obtained his Ph.D. in 1998 at the RWTH Aachen for research performed within the H1 experiment at the electron-proton collider HERA. As a CERN research fellow he worked within the OPAL experiment at the e^+e^- collider LEP. Since 2002, he is a member of the CMS Collaboration at the LHC as senior scientist for the University of Karlsruhe, now Karlsruhe Institute of Technology (KIT). Initially, he spearheaded the worldwide deployment of the CMS software and was the senior production manager for CMS at the German Tier 1 computing site GridKa. Since 2007, he lead physics working groups for CMS and lectured on jet physics. He is recognised as one of the most prominent researchers on QCD and jets at the LHC.