

# Parton distributions in the LHC era

*Luigi Del Debbio*<sup>1,\*</sup>

<sup>1</sup>*Higgs Centre for Theoretical Physics, The University of Edinburgh, Edinburgh EH9 3FD, United Kingdom*

**Abstract.** Analyses of LHC (and other!) experiments require robust and statistically accurate determinations of the structure of the proton, encoded in the parton distribution functions (PDFs). The standard description of hadronic processes relies on factorization theorems, which allow a separation of process-dependent short-distance physics from the universal long-distance structure of the proton. Traditionally the PDFs are obtained from fits to experimental data. However, understanding the long-distance properties of hadrons is a nonperturbative problem, and lattice QCD can play a role in providing useful results from first principles. In this talk we compare the different approaches used to determine PDFs, and try to assess the impact of existing, and future, lattice calculations.

## 1 Introduction

The description of physical processes involving nucleons is based on factorization, i.e. on the separation of scales between short-distance (hard) partonic interactions, described in perturbation theory, and the large-distance nonperturbative effects that are responsible for the internal structure of the nucleon, see e.g. Ref. [1]. The nucleon structure is encoded in Parton Distribution Functions (PDFs), which are universal, i.e. they describe properties of the nucleon that do not depend on the physical process under study, e.g. the same PDFs describe the parton content of the nucleon both in Deep Inelastic Scattering (DIS), and in experiments at hadronic colliders like the LHC. A brief introduction to PDFs is presented in Sect. 2.

The lack of evidence for new particles at the LHC experiments suggests that signs of new physics beyond the Standard Model will only show up as small deviations from the Standard Model predictions, which can only be established by precision studies. Increased precision in the experimental results calls for more precise theoretical computations, which severely challenge the techniques currently in use. PDFs are a crucial input in any analysis of collider experiments involving hadrons, and therefore the precision in their determination needs to match the overall precision required in searches for new physics.

PDFs can be extracted from global fits to experimental data. The ever-increasing number of experiments that are included in these fits has led to better fits, with robust methodologies that allow the propagation of the experimental error from data to the fitted functions. Two main concerns arise as we enter the era of precision measurements at hadron colliders. In the kinematical regions that are constrained by data, the reduction in the statistical errors requires a careful assessment of the systematic

---

\* Speaker, e-mail: luigi.del.debbio@ed.ac.uk

errors in global fits. At the same time there are also kinematical regions that are not particularly constrained by data, but are relevant e.g. for searches of new physics, or precision Higgs measurements. In Sect. 3 we summarise the recent results of global fits, trying to assess the statistical and systematic errors, and the target precision that is necessary for lattice calculations to have an impact.

Numerical simulations of QCD allow the computation of PDFs directly from first principles. Recent theoretical developments have paved the way to significant contributions from lattice QCD, which can have an impact on the overall precision of the PDFs determination. We present a review of lattice results in Sect. 4, summarising the current status of each methodology. While we do not aim to present an exhaustive review of results, we hope to address the main theoretical questions that arise in these lattice calculations, and refer to the bibliography for a more exhaustive list of references.

## 2 Parton Distribution Functions

As mentioned above, a generic hadronic observable involves both long- and short-distance contributions, and cannot be obtained simply by a perturbative computation. Factorization theorems allow a separation of the long-distance and short-distance contributions. The latter are obtained by computing matrix elements at the partonic level in perturbation theory, while the distribution of partons in the hadron is encoded in the PDFs. The factorization theorems guarantee that different observables (structure functions, cross sections, ...) can be expressed as functions of the PDFs; inverting these relations allows the PDFs to be determined from experimental data. Moreover PDFs are universal, i.e. the same parton distributions enter in *all* processes involving a given hadron. In this work we will only consider the parton distributions inside the proton, and we will illustrate the relation between PDFs and data in the familiar environment provided by DIS experiments.

### 2.1 Deep Inelastic Scattering

Deep Inelastic Scattering is the scattering of a lepton, with momentum  $k$ , off a nucleon, more precisely a proton in our case, with momentum  $P$ :

$$\ell(k) + N(p) \longrightarrow \ell(k') + X \quad (1)$$

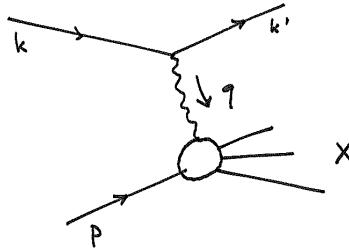
where  $k'$  is the momentum of the outgoing lepton, and  $X$  denotes a generic hadronic final state. We will consider here the simplest case, namely the case where a photon is exchanged between the lepton and the proton, in order to illustrate the main ideas without unnecessary complications. The reaction is depicted in Fig. 1. The computation of the amplitude for this process involves the evaluation of the matrix element of the electromagnetic current between the initial and final hadronic states, which encodes the effects of QCD nonperturbative dynamics. The differential cross section is given by

$$d\sigma = \frac{d^3 k'}{|\mathbf{k}'|} \frac{1}{2s(Q^2 - M_N^2)} L^{\mu\nu}(k, k') W_{\mu\nu}(p, q), \quad (2)$$

where  $M_N$  denotes the mass of the proton,  $Q^2 = -q^2 = -(k - k')^2$ , and  $s = (p + k)^2$ . The leptonic tensor  $L^{\mu\nu}$  is readily computed in perturbation theory, while the hadronic tensor

$$W_{\mu\nu}(p, q) = \frac{1}{4\pi} \sum_X \langle p | j_\mu(0)^\dagger | X \rangle \langle X | j_\nu(0) | p \rangle (2\pi)^4 \delta(p_X - p - q) \quad (3)$$

$$= \frac{1}{4\pi} \int d^4 y e^{iq \cdot y} \langle p | j_\mu(y)^\dagger j_\nu(0) | p \rangle \quad (4)$$



**Figure 1.** Kinematics of DIS: a lepton with momentum  $k$  scatters off a nucleon with momentum  $P$ . The final state contains the lepton with momentum  $k'$  and some hadronic state  $X$ .

involves the nonperturbative matrix elements. Using the transformation properties of  $W_{\mu\nu}$  under Lorentz transformations and parity, and current conservation, the hadronic tensor can be expressed as a function of two independent scalar functions. Introducing the kinematical variables  $v = p \cdot q$ , and  $x = Q^2/(2v)$ , yields

$$W_{\mu\nu}(p, q) = \left( -g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2} \right) F_1(x, Q^2) + \quad (5)$$

$$+ \left( p_\mu - q_\mu \frac{v}{q^2} \right) \left( p_\nu - q_\nu \frac{v}{q^2} \right) \frac{F_2(x, Q^2)}{v}, \quad (6)$$

where  $F_1$  and  $F_2$  are dimensionless structure functions, which depend on  $x$  and  $Q^2$ . The reaction is deeply inelastic when  $(p+q)^2 \gg M_N^2$ . Eqs. 3, and 5 allow to rewrite the cross section in Eq. 2 in terms of  $F_1$  and  $F_2$ , so that a measurement of the cross section yields the structure functions; therefore the structure functions can be considered as physical observables, in particular they are independent of the renormalization scheme, or the renormalization scale.

Parton distributions appear in the factorization theorems for the structure functions. Introducing the convolution

$$f(x) \otimes g(x) = \int_x^1 \frac{dz}{z} f\left(\frac{x}{z}\right) g(z), \quad (7)$$

the factorization theorem yields

$$F_1(x, Q^2) = x \sum_a C_{1a} \left( x, \frac{Q}{\mu}, \alpha_s(\mu) \right) \otimes f_a(x, \mu^2), \quad (8)$$

$$\frac{1}{x} F_2(x, Q^2) = x \sum_a \left[ \frac{1}{x} C_{2a} \left( x, \frac{Q}{\mu}, \alpha_s(\mu) \right) \right] \otimes f_a(x, \mu^2). \quad (9)$$

The sum in both equations is over all partons, labelled by  $a$ , where  $a$  runs over all quarks, antiquarks, and the gluon; <sup>1</sup> for each parton  $a$  there is a PDF  $f_a(x, \mu^2)$ , and the observable structure functions are obtained by taking the convolution of the PDFs with the coefficient functions  $C_{1a}$  and  $C_{2a}$ .

<sup>1</sup>When taking electromagnetic effects into account the photon counts as a parton, with its associated PDF. We will not dwell on this point here; see Ref. [2] and references therein for recent developments.

## 2.2 General properties of PDFs

A discussion of the quantities that enter these two equations provides an effective summary of the important properties of PDFs. First of all it is important to realise that Eqs. 8, and 9 are derived up to terms of order  $O(Q^{-2})$ , i.e. factorization works up to corrections that are suppressed by powers of  $Q^2$ . It is conventional to call this term the leading twist contribution to the structure functions. The coefficient functions  $C_{1a}$  and  $C_{2a}$  describe the perturbative, hard, partonic scattering. They are computed at some given order in perturbation theory, in a given renormalization scheme, and for a given value of the factorization scale  $\mu$ . Two important observations are in order. The *same* PDF appears in the expressions for the two distinct observables. This is the universality of PDFs that was already mentioned above. The structure of the nucleon does not depend on the process that we consider, while the hard scattering does. The second observation is also related to a point we already made above, namely that the structure functions are physical quantities and therefore independent of the details of the renormalization procedure. This means that the PDFs themselves have to be scheme and scale dependent in order to cancel the dependence in the coefficient functions. This is an important point to keep in mind: PDFs are defined in a given renormalization scheme and at some given value of the factorization scale. The dependence of the PDFs on the factorization scale can be computed in perturbation theory, and is encoded in the DGLAP equations:

$$\mu^2 \frac{d}{d\mu^2} f_a(x, \mu^2) = \sum_b P_{ab}(x, \alpha_s(\mu)) \otimes f_b(x, \mu^2). \quad (10)$$

The splitting functions  $P_{ab}$  are known in perturbation theory up to NNLO [3, 4]. The solution of the DGLAP equations yields the evolution of the PDFs with the factorization scale; a consistent treatment requires the evolution kernel to be computed at the same perturbative order, and in the same renormalization scheme, as the coefficient functions. The PDFs at a generic scale  $\mu^2$  are obtained by convoluting the PDFs at a reference scale  $\mu_0^2$  with a perturbative kernel:

$$f_a(x, \mu^2) = \sum_b \Gamma_{ab}(x, \alpha_s(\mu), \alpha_s(\mu_0)) \otimes f_b(x, \mu_0^2). \quad (11)$$

DGLAP evolution has been implemented in numerous codes that are publicly available, see e.g. [5–7].

The PDFs and the coefficient functions satisfy analogous evolution equations, so that the physical structure functions are scale-independent. Clearly when the coefficient functions, and the splitting functions are computed to order  $\alpha_s^N$ , we expect a residual scale dependence of the physical observables, of order  $\alpha_s^{N+1}$ .

Finally, note that Eqs. 8, and 9 can be rewritten in terms of the Mellin moments of the structure functions,

$$\tilde{F}_1(n, Q^2) = \int_0^1 \frac{dx}{x} x^n F_1(x, Q^2), \quad (12)$$

$$\tilde{F}_2(n, Q^2) = \int_0^1 \frac{dx}{x} x^{n-1} F_2(x, Q^2), \quad (13)$$

and the Mellin moments of the PDFs and the coefficient functions, which are defined analogously. The factorization theorems in Mellin space take the simple form

$$\tilde{F}_i(n, Q^2) = \sum_a \tilde{f}_a(n, Q^2) \tilde{C}_{ia}\left(n, \frac{Q}{\mu}, \alpha_s(\mu)\right), \quad i = 1, 2. \quad (14)$$

### 2.3 Operator definition

The PDF for a quark of flavor  $a$  can be defined as the matrix element of a non-local operator

$$f(\xi) = \frac{1}{4\pi} \int dx^- e^{-i\xi P^+ x^-} \langle P | \bar{\psi}(0, x^-, 0_\perp) \gamma^+ \mathcal{G}(x^-, 0) \psi(0) | P \rangle, \quad (15)$$

where  $|P\rangle$  is the hadron state with momentum  $P = (P^+, m^2/(2P^+), 0_\perp)$  in light-cone coordinates,  $\gamma^+ = (\gamma^0 + \gamma^3)/2$ , and

$$\mathcal{G}(x^-, 0) = \mathcal{P} \exp \left\{ -ig \int_0^{x^-} dy^- A^+(0, y^-, 0_\perp) \right\}. \quad (16)$$

A full discussion of this definition can be found in Refs. [1, 8–10].

The operator defined in Eq. 15 requires to be renormalized. Explicit calculations, e.g. in the  $\overline{\text{MS}}$  scheme, show that the renormalization procedure introduces a dependence on the scale  $\mu$ . This dependence is described by the DGLAP equations discussed above. We shall come back to discussing the properties of such operators below.

It is interesting to remark that this definition is not unique [11]. A valid definition of a parton distribution is obtained by convoluting  $f_a$  with any kernel  $D_{ab}(x, Q/\mu)$  that is perturbatively defined:

$$g_a(x, Q^2) = D_{ab}(x, Q/\mu) \otimes f_b(x, \mu^2). \quad (17)$$

A change in the definition of the PDFs entails a change in the coefficient functions, so that the physical observables can still be expressed via a factorization theorem, and remain unchanged. Care must be exercised when comparing PDFs to ensure that the same quantity is actually being computed and compared, especially when matching lattice quantities to quantities that are defined in Minkowski space.

## 3 Global fits from experimental data

Parton Distribution Functions are currently extracted from global fits to available experimental data, using factorization theorems to relate the PDFs to the physical observables. Before discussing the characteristics of these global fits, it is useful to summarise the basic ideas in a simple case, namely the non-singlet structure function for DIS, see e.g. [12].

### 3.1 The non-singlet PDF

The observable in this simple example is

$$F_2^{\text{NS}}(x, Q^2) = F_2^p(x, Q^2) - F_2^d(x, Q^2) \quad (18)$$

$$= C_{\text{NS}}(x, Q^2) \otimes f_{\text{NS}}(x, Q^2), \quad (19)$$

where  $F_2^p$  and  $F_2^d$  are the structure functions for the proton and deuteron respectively. As shown explicitly in Eq. 19, the factorization theorem allows the non-singlet structure function to be expressed as the convolution of the non-singlet PDF

$$f_{\text{NS}}(x, Q^2) = \left[ (u(x, Q^2) + \bar{u}(x, Q^2)) - (d(x, Q^2) + \bar{d}(x, Q^2)) \right], \quad (20)$$

with the coefficient function  $C_{\text{NS}}$ , which is known in perturbation theory [13]. Using DGLAP evolution, the structure function at any value of  $Q^2$  can be written as a function of the PDFs at a single reference scale  $Q_0^2$ :

$$F_2^{\text{NS}}(x, Q^2) = \int_x^1 \frac{dy}{y} K_{\text{NS}}(y, \alpha_s(Q^2), \alpha_s(Q_0^2)) f_{\text{NS}}\left(\frac{x}{y}, Q_0^2\right). \quad (21)$$

Eq. 21 summarises the challenges that global fits are trying to address. Experimental data, which are correlated and have statistical and systematic errors, appear on the left-hand side of the equation, and are used to determine the PDF at the reference scale on the right-hand side. The problem is ill-defined in the sense that the continuous real functions  $f_a(x, Q_0^2)$  cannot be determined from a discrete set of data, no matter how copious this set is. In order to overcome this difficulty, a parametrization for  $f_a(x, Q_0^2)$  needs to be chosen; experimental data are then used in order to constrain the parameters that define the functional form. The parametrizations used for these fits need to be sufficiently flexible, so that they do not introduce a bias in the result of the fit. Moreover the error on the data needs to be propagated into an error on the fitted functions  $f_a(x, Q_0^2)$ .

It is also clear from Eq. 21 that data for  $F_2^{\text{NS}}$  can only constraint the non-singlet PDF, and do not provide information on individual flavors distributions. In order to constrain all PDFs a large variety of processes needs to be included in the analysis.

### 3.2 Global datasets

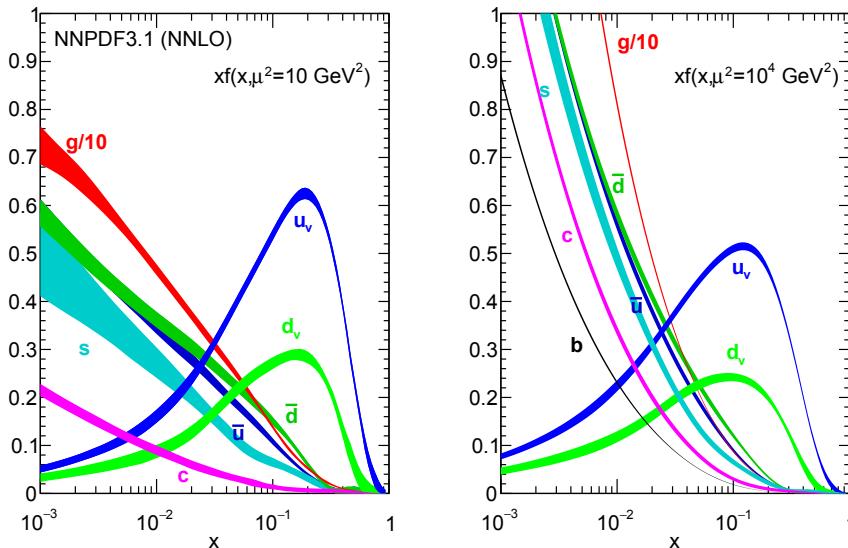
Factorization theorems allow most observables to be written as a convolution of hard partonic cross sections, and PDFs. For example the cross sections for processes at hadron colliders can be written as

$$\sigma(H_1 H_2 \rightarrow X) = \sum_{a,b} \int dx_1 dx_2 f_a(x_1, \mu^2) f_b(x_2, \mu^2) \times \hat{\sigma}_{ab \rightarrow X}(x_1 x_2 s, \mu^2, \mu_R^2), \quad (22)$$

where  $H_1$  and  $H_2$  are the hadrons involved in the collisions, and  $f_a$  and  $f_b$  are the parton distributions in these hadrons. Using DGLAP evolution again, it is clear that Eq. 22 yields an expression for observables as functions of the PDFs at the reference scale. The universality of PDFs allows to combine data from different experiments to constrain the same PDFs. Different experiments will constrain different combinations of PDFs, and different kinematical regions in  $x$ . Being able to combine all the available data, including the rapidly increasing amount of data from the LHC, is crucial to get the best determination of PDFs.

As an example, the result of the latest global fit by the NNPDF Collaboration is shown in Fig. 2. Here we review briefly the new data included in these global fits in going from [14] to [15], trying to identify their impact on the determination of PDFs. The reason for focussing on this specific example is twofold: understanding the current level of precision in global fits, and highlighting the impact of recent LHC data.

Deep-inelastic scattering data are summarised in Tab. 1. The final HERA combination [16] provides stringent bounds on quark distributions at medium values of  $x$ . The bottom [17, 18] and charm [19] structure functions have been considered in order to constrain respectively the determination of the bottom mass, and the charm content of the proton. Tevatron data, reported in Tab. 2, include fixed target Drell-Yan from the E605 [20] and E866 [21–23] experiments, weak boson production from CDF [24] and D0 [25]  $Z$  rapidity distributions, and inclusive jet production [26]. The very precise  $W$  lepton asymmetries in the electron [27] and muon [28], provide important information



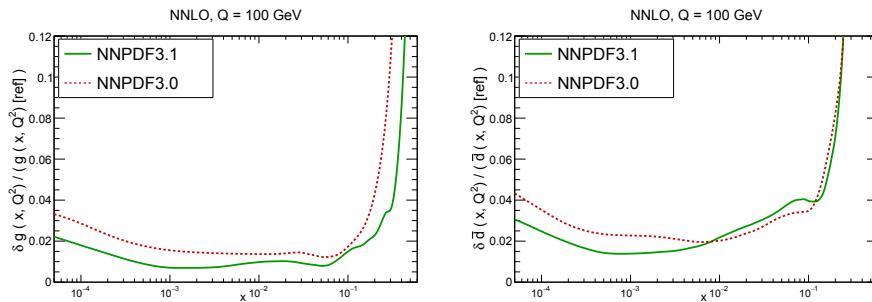
**Figure 2.** Results of the latest global fit by the NNPDF Collaboration. PDFs are shown at factorization scales of  $10 \text{ GeV}^2$  (left) and  $10^4 \text{ GeV}^2$  (right). PDFs at different factorization scales are related by DGLAP evolution. Note the size of the statistical errors for the different partons in different kinematical regions.

on the quark flavor separation at large- $x$ , as demonstrated in [29]. Nowadays an increasing number of LHC results has already been included in global fits. The recent NNPDF3.1 set of PDFs includes data for the  $Z$  boson double-differential distribution [30, 31]; the inclusive  $W^+$ ,  $W^-$ , and  $Z$  rapidity distribution [32, 33]; the top-quark pair production normalized  $y_t$  distribution [34, 35]; the  $t\bar{t}$  total cross section [36–38]; the inclusive jet cross section [39, 40]; the low-mass Drell-Yan [41]; and the inclusive  $W$ ,  $Z$  production [42, 43].

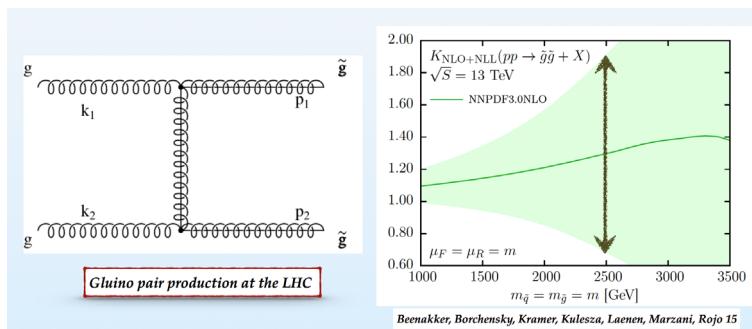
This long list of experimental data should give a feeling for the variety of data used in these fits, and for the LHC contribution to the determination of PDFs, with most PDFs being affected at  $1\sigma$  to  $2\sigma$  level. The impact of these data is discussed in detail e.g. in Ref. [15]. A quantitative estimate of the error reduction due to new data is shown in Fig. 3, where the statistical error for the gluon and the  $\bar{d}$  quark PDFs are shown. An overall reduction of the error is seen for all values of  $x$ , sometimes by a factor of 2, bringing the relative uncertainty at the level of 2%. More detail can be found in the actual publications, but it is useful to keep in mind this order of magnitude as being typical of the uncertainty from global fits, in the regions that are reasonably constrained by the data. Clearly the error blows up at very small values of  $x$ .

The uncertainty on the PDFs is rapidly becoming one the limiting factors in searches for new physics. An example of the impact of the PDF error can be found in Ref. [44], where the relative size of the NLL corrections for gluino pair production was computed. As shown in Fig. 4, the error in the relative size of the NLL corrections grows very quickly as the gluino mass is increased, mostly as a consequence of the large PDF errors at large values of  $x$ .

There are several collaborations that are currently producing global fits, using basically the same datasets, but different methodologies, see Refs. [15, 69–72] for recent updates. Here we would like to summarise what we believe are the important issues to keep in mind when engaging in lattice studies of PDFs. It is interesting to remark that global fits yield consistent results within errors, despite a wide



**Figure 3.** Error reduction in the gluon (left) and the  $\bar{d}$  (right) PDFs due to the inclusion of recent LHC data in the NNPDF global fit. The relative error is shown as a function of  $x$ . The NNPDF31 global fit includes the extra data as discussed in the text above.



**Figure 4.** Impact of the PDF errors on the NLL corrections to gluino pair production at the LHC. The ratio  $\sigma^{\text{NLO+NLL}}/\sigma^{\text{NLO}}$  is shown as a function of the gluino mass. The error band is obtained from the propagation of the PDF error in the numerator. Plot courtesy of J. Rojo.

variety of methodologies, which underscores the robustness of current studies. Based on the current global fits, the uncertainty in the central value of the PDFs is at the % level in the central  $x$  region at  $Q = 100$  GeV. LHC data have already had a significant impact in reducing the error bars, and will continue to do so as Run-2 data are added to the existing fits. The determinations of PDFs in the small  $x$  and large  $x$  region have larger uncertainties, mostly due to the fact that reaching these kinematic regions is difficult in experiments. At the current level of precision the systematic errors in global fits will start to play a significant role, and will need to be assessed carefully. Reliable uncertainties are crucial both for precision Higgs physics, and for potential discoveries. First-principle calculations of PDFs can be performed using lattice QCD, which could thereby yield a significant contribution to the determination of PDFs.

## 4 Lattice computations

Recent work in lattice QCQ has led to significant progresses in the computation of PDFs from first principles. In this report, we will focus on recent developments in the extraction of light-cone PDFs from the so-called quasi-PDFs.

Experiment	Obs.	Ref.	$N_{\text{dat}}$	$x$ range	$Q$ range (GeV)	Theory
NMC	$F_2^d/F_2^p$	[45]	260 (121/121)	$0.012 \leq x \leq 0.68$	$2.1 \leq Q \leq 10$	APFEL
	$\sigma_{\text{NC},p}$	[46]	292 (204/204)	$0.012 \leq x \leq 0.50$	$1.8 \leq Q \leq 7.9$	
SLAC	$F_2^p$	[47]	211 (33/33)	$0.14 \leq x \leq 0.55$	$1.9 \leq Q \leq 4.4$	APFEL
	$F_2^d$	[47]	211 (34/34)	$0.14 \leq x \leq 0.55$	$1.9 \leq Q \leq 4.4$	
BCDMS	$F_2^p$	[48]	351 (333/333)	$0.07 \leq x \leq 0.75$	$2.7 \leq Q \leq 15.1$	APFEL
	$F_2^d$	[49]	254 (248/248)	$0.07 \leq x \leq 0.75$	$3.0 \leq Q \leq 15.1$	
CHORUS	$\sigma_{\text{CC},v}^{\text{CC},v}$	[50]	607 (416/416)	$0.045 \leq x \leq 0.65$	$1.9 \leq Q \leq 9.8$	APFEL
	$\sigma_{\text{CC},\bar{v}}^{\text{CC},\bar{v}}$	[50]	607 (416/416)	$0.045 \leq x \leq 0.65$	$1.9 \leq Q \leq 9.8$	
NuTeV	$\sigma_v^{\text{CC}}$	[51, 52]	45 (39/39)	$0.02 \leq x \leq 0.33$	$2.0 \leq Q \leq 10.8$	APFEL
	$\sigma_{\bar{v}}^{\text{CC}}$	[51, 52]	45 (37/37)	$0.02 \leq x \leq 0.21$	$1.9 \leq Q \leq 8.3$	
HERA	$\sigma_{\text{NC},\text{CC}}^p (*)$	[16]	1306 (1145/1145)	$4 \cdot 10^{-5} \leq x \leq 0.65$	$1.87 \leq Q \leq 223$	APFEL
	$\sigma_{\text{NC}}^c$	[53]	52 (47/37)	$7 \cdot 10^{-5} \leq x \leq 0.05$	$2.2 \leq Q \leq 45$	
	$F_2^b (*)$	[17, 18]	29 (29/29)	$2 \cdot 10^{-4} \leq x \leq 0.5$	$2.2 \leq Q \leq 45$	
EMC	[ $F_2^c$ ] (*)	[19]	21 (16/16)	$0.014 \leq x \leq 0.44$	$2.1 \leq Q \leq 8.8$	APFEL

**Table 1.** Deep-inelastic scattering data included in NNPDF3.1. The EMC  $F_2^c$  data are in brackets because they are only included in a dedicated set but not in the default dataset. New datasets, not included in NNPDF3.0, are denoted (\*). The kinematic range covered in each variable is given after cuts are applied. The total number of DIS data points after cuts is 3102/3092 for the NLO/NNLO PDF determinations (not including the EMC  $F_2^c$  data).

Exp.	Obs.	Ref.	$N_{\text{dat}}$	$\text{Kin}_1$	$\text{Kin}_2$ (GeV)	Theory
E866	$\sigma_{\text{DY}}^d/\sigma_{\text{DY}}^p$	[23]	15 (15/15)	$0.07 \leq y_{ll} \leq 1.53$	$4.6 \leq M_{ll} \leq 12.9$	APFEL+Vrap
	$\sigma_{\text{DY}}^p$	[21, 22]	184 (89/89)	$0 \leq y_{ll} \leq 1.36$	$4.5 \leq M_{ll} \leq 8.5$	APFEL+Vrap
E605	$\sigma_{\text{DY}}^p$	[20]	119 (85/85)	$-0.2 \leq y_{ll} \leq 0.4$	$7.1 \leq M_{ll} \leq 10.9$	APFEL+Vrap
CDF	$d\sigma_Z/dy_Z$	[24]	29 (29/29)	$0 \leq y_{ll} \leq 2.9$	$66 \leq M_{ll} \leq 116$	Sherpa+Vrap
	$k_t$ incl jets	[54]	76 (76/76)	$0 \leq y_{\text{jet}} \leq 1.9$	$58 \leq p_T^{\text{jet}} \leq 613$	NLOjet++
D0	$d\sigma_Z/dy_Z$	[25]	28 (28/28)	$0 \leq y_{ll} \leq 2.8$	$66 \leq M_{ll} \leq 116$	Sherpa+Vrap
	W electron asy (*)	[27]	13 (13/8)	$0 \leq y_e \leq 2.9$	$Q = M_W$	MCFM+FEWZ
	W muon asy (*)	[28]	10 (10/9)	$0 \leq y_\mu \leq 1.9$	$Q = M_W$	MCFM+FEWZ

**Table 2.** Same as Table 1 for the Tevatron fixed-target Drell-Yan and  $W, Z$  and jet collider data. The total number of Tevatron data points after cuts is 345/339 for NLO/NNLO fits.

#### 4.1 Quasi-PDFs

In order to discuss the quasi-PDFs, it is useful to review briefly the field-theoretical definition using bi-local operators

$$O_i(\zeta) = \bar{\psi}(\zeta) \Gamma_i \mathcal{P} \exp \left( -ig \int_0^\zeta d\eta A(\eta) \right) \psi(0), \quad (23)$$

where  $\Gamma_i$  is a matrix in spin space, and  $A_\mu$  is the gauge potential. The path-ordered exponential in Eq. 23 ensures that these operators are gauge invariant. We focus here on the operator  $\mathcal{F}_+$  obtained by

Exp.	Obs.	Ref.	$N_{\text{dat}}$	$\text{Kin}_1$	$\text{Kin}_2$ (GeV)	Theory
ATLAS	$W, Z$ 2010	[55]	30 (30/30)	$0 \leq  \eta_l  \leq 3.2$	$Q = M_W, M_Z$	MCFM+FEWZ
	$W, Z$ 2011 (*)	[32]	34 (34/34)	$0 \leq  \eta_l  \leq 2.3$	$Q = M_W, M_Z$	MCFM+FEWZ
	high-mass DY 2011	[56]	11 (5/5)	$0 \leq  \eta_l  \leq 2.1$	$116 \leq M_{ll} \leq 1500$	MCFM+FEWZ
	low-mass DY 2011 (*)	[41]	6 (4/6)	$0 \leq  \eta_l  \leq 2.1$	$14 \leq M_{ll} \leq 56$	MCFM+FEWZ
	$[Z \text{ } p_T \text{ } 7 \text{ TeV } (p_T^Z, y_Z)]$ (*)	[57]	64 (39/39)	$0 \leq  y_Z  \leq 2.5$	$30 \leq p_T^Z \leq 300$	MCFM+NNLO
	$Z \text{ } p_T \text{ } 8 \text{ TeV } (p_T^Z, M_{ll})$ (*)	[30]	64 (44/44)	$12 \leq M_{ll} \leq 150 \text{ GeV}$	$30 \leq p_T^Z \leq 900$	MCFM+NNLO
	$Z \text{ } p_T \text{ } 8 \text{ TeV } (p_T^Z, y_Z)$ (*)	[30]	120 (48/48)	$0.0 \leq  y_Z  \leq 2.4$	$30 \leq p_T^Z \leq 150$	MCFM+NNLO
	7 TeV jets 2010	[58]	90 (90/90)	$0 \leq  y_{\text{jet}}  \leq 4.4$	$25 \leq p_T^{\text{jet}} \leq 1350$	NLOjet++
	2.76 TeV jets	[59]	59 (59/59)	$0 \leq  y_{\text{jet}}  \leq 4.4$	$20 \leq p_T^{\text{jet}} \leq 200$	NLOjet++
	7 TeV jets 2011 (*)	[39]	140 (31/31)	$0 \leq  y_{\text{jet}}  \leq 0.5$	$108 \leq p_T^{\text{jet}} \leq 1760$	NLOjet++
	$\sigma_{\text{tot}}(t\bar{t})$	[36, 37]	3 (3/3)	-	$Q = m_t$	top++
	$(1/\sigma_{t\bar{t}})d\sigma(t\bar{t})/y_t$ (*)	[34]	10 (10/10)	$0 <  y_t  < 2.5$	$Q = m_t$	Sherpa+NNLO
CMS	$W$ electron asy	[60]	11 (11/11)	$0 \leq  \eta_e  \leq 2.4$	$Q = M_W$	MCFM+FEWZ
	$W$ muon asy	[61]	11 (11/11)	$0 \leq  \eta_\mu  \leq 2.4$	$Q = M_W$	MCFM+FEWZ
	$W + c$ total	[62]	5 (5/0)	$0 \leq  \eta_l  \leq 2.1$	$Q = M_W$	MCFM
	$W + c$ ratio	[62]	5 (5/0)	$0 \leq  \eta_l  \leq 2.1$	$Q = M_W$	MCFM
	2D DY 2011 7 TeV	[63]	124 (88/110)	$0 \leq  \eta_{ll}  \leq 2.2$	$20 \leq M_{ll} \leq 200$	MCFM+FEWZ
	[2D DY 2012 8 TeV]	[64]	124 (108/108)	$0 \leq  \eta_{ll}  \leq 2.4$	$20 \leq M_{ll} \leq 1200$	MCFM+FEWZ
	$W^\pm$ rap 8 TeV (*)	[33]	22 (22/22)	$0 \leq  \eta_l  \leq 2.3$	$Q = M_W$	MCFM+FEWZ
	$Z \text{ } p_T \text{ } 8 \text{ TeV}$ (*)	[31]	50 (28/28)	$0.0 \leq  y_Z  \leq 1.6$	$30 \leq p_T^Z \leq 170$	MCFM+NNLO
	7 TeV jets 2011	[65]	133 (133/133)	$0 \leq  y_{\text{jet}}  \leq 2.5$	$114 \leq p_T^{\text{jet}} \leq 2116$	NLOjet++
	2.76 TeV jets (*)	[40]	81 (81/81)	$0 \leq  y_{\text{jet}}  \leq 2.8$	$80 \leq p_T^{\text{jet}} \leq 570$	NLOjet++
	$\sigma_{\text{tot}}(t\bar{t})$	[38, 66]	3 (3/3)	-	$Q = m_t$	top++
	$(1/\sigma_{t\bar{t}})d\sigma(t\bar{t})/y_{t\bar{t}}$ (*)	[35]	10 (10/10)	$-2.1 < y_{t\bar{t}} < 2.1$	$Q = m_t$	Sherpa+NNLO
LHCb	$Z$ rapidity 940 pb	[67]	9 (9/9)	$2.0 \leq \eta_l \leq 4.5$	$Q = M_Z$	MCFM+FEWZ
	$Z \rightarrow ee$ rapidity 2 fb	[68]	17 (17/17)	$2.0 \leq \eta_l \leq 4.5$	$Q = M_Z$	MCFM+FEWZ
	$W, Z \rightarrow \mu 7 \text{ TeV}$ (*)	[42]	33 (33/29)	$2.0 \leq \eta_l \leq 4.5$	$Q = M_W, M_Z$	MCFM+FEWZ
	$W, Z \rightarrow \mu 8 \text{ TeV}$ (*)	[43]	34 (34/30)	$2.0 \leq \eta_l \leq 4.5$	$Q = M_W, M_Z$	MCFM+FEWZ

**Table 3.** Same as Table 1, for ATLAS, CMS and LHCb data from the LHC Run I at  $\sqrt{s} = 2.76 \text{ TeV}$ ,  $\sqrt{s} = 7 \text{ TeV}$  and  $\sqrt{s} = 8 \text{ TeV}$ . The ATLAS 7 TeV  $Z \text{ } p_T$  and CMS 2D DY 2012 are in brackets because they are only included in a dedicated study but not in the default PDF set. The total number of LHC data points after cuts is 848/854 for NLO/NNLO fits (not including ATLAS 7 TeV  $Z \text{ } p_T$  and CMS 2D DY 2012).

choosing  $\Gamma_i = \gamma^+$ , and integrating over a light-cone direction,  $\zeta = (0, y^-, \vec{0}_\perp)$ ,

$$\mathcal{F}_+(k^+) = \int \frac{dy^-}{4\pi} e^{-ik^+ y^-} O_i(\zeta). \quad (24)$$

$\mathcal{F}_+$  needs to be properly renormalized, e.g.

$$\mathcal{F}_{+,R}(k^+, \mu) = \int_1^\infty \frac{d\xi}{\xi} K(\xi, g_R(\mu)) \mathcal{F}_+(\xi k^+), \quad (25)$$

where we have indicated explicitly the dependence on the renormalization scale  $\mu$  - for a detailed discussion see e.g. Ref. [8].

The PDFs are obtained from the matrix element of  $\mathcal{F}_R$  between hadronic states with momentum  $P$ :

$$f(x, \mu) = \langle P | \mathcal{F}_{+,R}(x P^+, \mu) | P \rangle, \quad (26)$$

where the scale dependence in the PDFs arises from the scale dependence of the renormalized bi-local operator in the right-hand side of Eq. 26.

Clearly the integral along the light-cone direction cannot be performed in Euclidean space. An interesting proposal to overcome this problem was put forward recently in Refs. [73, 74], where so-called quasi-PDFs are introduced by shifting the integration contour in a purely spatial direction, e.g. the  $z$ -direction:

$$q(x, a, M_N, P_z) = \langle P_z | \int \frac{dz}{4\pi} e^{-ixP_z z} \bar{\psi}(z) \gamma^3 \mathcal{P} \exp \left( -ig \int_0^z d\eta A(\eta) \right) \psi(0) | P_z \rangle. \quad (27)$$

Equation 27 defines a bare matrix element in the regularized theory, which can be evaluated by Monte Carlo simulations in Euclidean space. The dependence on the cutoff  $a$  has been made explicit in the expression above as a reminder of the fact that these quantities need to be properly renormalized.

In order to extract PDFs from the quasi-PDFs, the following steps are needed: the renormalization of the lattice operators, including potential power divergencies, the relation between the Euclidean and the Minkowski space results, and a 'factorization theorem' relating the lattice renormalized quantity to the desired distributions. We shall now discuss these issues in turn.

### Renormalization

The renormalization of operators  $O_i(\zeta)$  in Eq. 23 has been studied both in perturbation theory and nonperturbatively.

A one-loop calculation has been presented in Ref. [75, 76]. Although ultimately the operators will need to be renormalized nonperturbatively, a perturbative calculation is helpful in clarifying the renormalization pattern.

The operators considered in Ref. [75, 76] are of the form

$$O_i(z) = \bar{\psi}(x) \Gamma_i \mathcal{P} \exp \left[ ig \int_0^z d\zeta A(\zeta) \right] \psi(x + z\mu), \quad (28)$$

with  $\mu$  set to one, and  $\Gamma_i$  spanning the entire Clifford algebra. These operators can be separated into 8 subgroups:

$$\begin{aligned} S &= O_1, \\ P &= O_{\gamma^5}, \\ V_1 &= O_{\gamma^1}, \\ V_\nu &= O_{\gamma^\nu}, \quad \nu = 2, 3, 4 \\ A_1 &= O_{\gamma^5 \gamma^1}, \\ A_\nu &= O_{\gamma^5 \gamma^\nu}, \quad \nu = 2, 3, 4 \\ T_{1\nu} &= O_{\sigma^{1\nu}}, \quad \nu = 2, 3, 4 \\ T_{\nu\rho} &= O_{\sigma^{\nu\rho}}, \quad \nu, \rho = 2, 3, 4. \end{aligned} \quad (29)$$

The perturbative calculation highlighted the existence of finite mixings between pairs of the operators above under renormalization, and therefore the pattern of renormalization is

$$\begin{pmatrix} O_{1R} \\ O_{2R} \end{pmatrix} = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix}^{-1} \begin{pmatrix} O_1 \\ O_2 \end{pmatrix}, \quad (30)$$

where the dependence of the renormalization constants on the regulator  $X$ , and the renormalization scheme  $Y$  has been omitted to avoid cluttering the equation. We shall reintroduce this dependence below when we discuss conversion factors. The regularizations that we will consider here are dimensional regularization (DR) and the lattice formulation of gauge theories (LR). The renormalization schemes are  $\overline{\text{MS}}$  and variants of the RI-MOM scheme, which we denote RI for simplicity. The elements of the renormalization matrix have been computed at one-loop in perturbation theory:

$$Z_{ij} = \delta_{ij} + g^2 z_{ij} + \mathcal{O}(g^4). \quad (31)$$

The pairs involved in the mixing are:  $\{S, V_1\}$ ,  $\{A_2, T_{34}\}$ ,  $\{A_3, T_{42}\}$ ,  $\{A_4, T_{23}\}$ . The details of the one-loop calculation can be found in Ref. [75]. Besides showing explicitly the mixing pattern, the one-loop calculation provides the conversion factors to relate a nonperturbative scheme like RI-MOM to the  $\overline{\text{MS}}$  scheme used in perturbative QCD, and in the global fits discussed in the section above. These conversion factors are independent of the regularization used, and can be written as

$$C^{\overline{\text{MS}}, \text{RI}} = \left( Z^{\text{LR}, \overline{\text{MS}}} \right)^{-1} \cdot \left( Z^{\text{LR}, \text{RI}} \right) = \left( Z^{\text{DR}, \overline{\text{MS}}} \right)^{-1} \cdot \left( Z^{\text{DR}, \text{RI}} \right). \quad (32)$$

The nonperturbative renormalization of the operators  $O_i(z)$  proceeds along the usual lines. Having identified the renormalization pattern in Eq. 30, the matrix elements are computed by imposing a set of renormalization conditions on amputated correlators  $\Lambda(z, p)$ :

$$Z_q^{-1} \frac{1}{12} \text{Tr} \left[ \Lambda_{i,R}(p, z) \left( \Lambda_j^{\text{tree}}(p, z) \right)^{-1} \right] \Big|_{p^2 = \mu^2} = \delta_{ij}, \quad (33)$$

$Z_q$  is the fermion field renormalization, and the amputated correlators are defined as usual:

$$\Lambda_i(p, z) = S(p)^{-1} \langle O_i(z) \psi(-p) \bar{\psi}(p) \rangle S(p)^{-1}. \quad (34)$$

Note that any divergence coming from the Wilson line is automatically taken into account in this framework.

Nonperturbative renormalization has been implemented recently by two groups, and first results have appeared in Refs. [77–79], where the interested reader can find a detailed discussion, and results.

An alternative approach to the renormalization of these nonlocal operators has been proposed in Ref. [80]. Having different renormalization methodologies, and different lattice discretizations, should ultimately lead to cross-checks and robust results for the continuum extrapolation of the quasi-PDFs.

### Power divergences

It has been suggested in Ref. [81] that the operator defined along a spatial direction in Euclidean space could have power divergencies that do not appear in Minkowski space, where the integral is performed along a light-cone direction. This issue needs to be investigated in more detail, as it could potentially invalidate the lattice approach.

### Euclidean/Minkowski definition

The matrix elements of interest, Eq. 27 are extracted in Monte Carlo simulations from correlators computed in Euclidean time. The latter can be written as a sum of decaying exponentials, where the decay rate is given by the discrete energy levels in the spectrum, and the coefficients of each term in the spectral decomposition yield the finite-volume matrix elements of operators. In the case of the quasi-PDFs it is clear from Eq. 27 that operators of interest only involve fields at time  $t = 0$ , and

should be completely agnostic about the space-time signature. Both the operators, and the eigenstates of the Hamiltonian, are independent of the choice of metric. However the procedure for extracting the matrix elements does depends on it.

For the case of quasi-PDFs the dependence of the matrix elements on the metric has been discussed in Refs. [82, 83]. Here we briefly summarise the argument in Ref. [83] form where we borrow the notation. In Euclidean space the matrix element is extracted from the coefficient of terms that decay exponentially in time. For a three-point function we obtain:

$$\langle N(\tau', \mathbf{P}') O_i N(\tau, \mathbf{P}) \rangle = \langle 0 | N(0, \mathbf{P}') | P' \rangle \langle P' | O_i | P \rangle \langle P | N(0, \mathbf{P}) | 0 \rangle e^{-E_P \tau'} e^{-E_P \tau} + \dots, \quad (35)$$

where  $N(\tau, P)$  is an interpolating operator with the quantum numbers of the proton, and spatial momentum  $\mathbf{P}$ ,  $O_i$  is the operator of interest,  $E_P = \sqrt{\mathbf{P}^2 + m^2}$ , and the ellipses denote terms that decay faster for large values of the Euclidean time  $\tau$ . In Minkowski space the same matrix elements is extracted from the residue of a T-ordered product of the same fields at the pole corresponding to the two protons going on-shell.

A one-loop calculation in a scalar field theory toy model is discussed in full detail in Ref. [83], showing how the two procedures described above yield the same result for the matrix element.

### Factorization/matching

Having performed a proper renormalization of the lattice operator, quasi-PDFs can be defined in the continuum limit:

$$q(x, \mu, M_N, P_z) = \langle P_z | \int \frac{dz}{4\pi} e^{-i(xP_z)z} O_{z,R}(z, \mu, P_z) | P_z \rangle \quad (36)$$

$$= \lim_{a \rightarrow 0} \langle P_z | \int \frac{dz}{4\pi} e^{-i(xP_z)z} Z(z, \mu a) O_z(z, P_z, a) | P_z \rangle. \quad (37)$$

These can be considered like bona fide observables, that are finite when the lattice spacing vanishes, and are independent of the lattice discretization used in their definition. We have emphasized in Eq. 36 the dependence on the renormalization scale  $\mu$  and the momentum of the nucleon state  $P_z$ . Any extrapolation/interpolation to the physical values of the quark masses should also be performed at this stage. The continuum limit of the quasi-PDF obtained using different discretization could, and should be compared. This would yield an idea of the systematics of the lattice calculation before trying to relate the quasi-PDFs to the light-cone ones. It may be desirable at this stage to convert the renormalized quasi-PDF to the  $\overline{\text{MS}}$  scheme, in order to avoid any references to the details of the lattice formulation after this stage.

Having determined the continuum limit of the quasi-PDFs, and having a robust control of the systematics, the final step in the process of computing the light-cone PDFs from first principles is the matching of the quasi-PDFs to the light-cone ones [84–87]. We expect to be able to write this last step in the form of a factorization theorem, relating a physical observable - in this case the continuum limit quasi-PDFs - to the light cone PDFs. The role of the physical scale of the observable is played by the momentum  $P_z$ , and therefore

$$q(x, \mu, M_N, P_z) = C_q \left( x, \frac{P_z}{\mu} \right) \otimes f(x, \mu) + O \left( \frac{\Lambda_{\text{QCD}}^2}{P_z^2}, \frac{M_N^2}{P_z^2} \right), \quad (38)$$

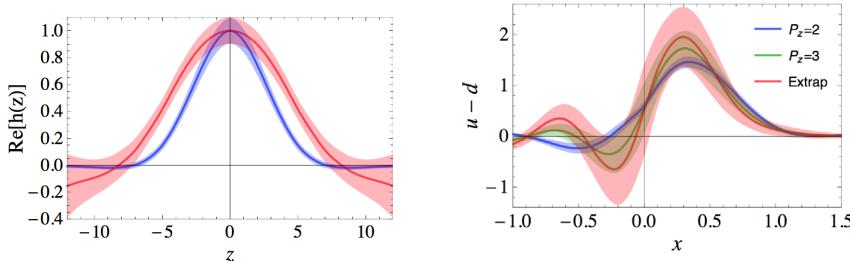
where  $C_q$  is a coefficient function to be computed in perturbation theory, and are the analogues of the coefficient functions that we discussed earlier for the structure functions. Note that the corrections to

the factorization are given by powers of  $\Lambda_{\text{QCD}}^2/P_z^2$ , and  $M_N^2/P_z^2$ ; therefore the matching is expected to hold for large values of  $P_z$ . More specifically the procedure is robust if there is a window such that

$$\Lambda_{\text{QCD}}, M_N \ll P_z, \mu \ll a^{-1}. \quad (39)$$

The left-hand side of the inequality guarantees that the power corrections to the matching in Eq. 38 are indeed small. The right-hand side of the inequality is needed in order to have small lattice artefacts when extrapolating to the continuum limit, and in the RI-MOM renormalization procedure. The perturbative order used in the computation of the coefficient functions  $C_q$  determines the perturbative order at which the PDFs are defined. Therefore a one-loop calculation would yield light-cone PDFs that could be compared with LO PDFs extracted from global fits.

Some preliminary results obtained in Refs. [77, 78] are shown in Fig. 5. While it is premature to try to compare these results with the results obtained from global fits, it is interesting to note that a signal with sensible error bars can be extracted.



**Figure 5.** Light-cone PDFs reconstructed from quasi-PDF. First results presented in Ref. [78] (left), and [77] (right).

## 4.2 Gradient flow

Another interesting approach to the computation of PDFs has been put forward in Ref. [88], where the moments of the quasi-PDFs are defined at some finite flow time  $\tau$  along the gradient flow [89, 90]. The gradient flow acts as a regulator, yielding field correlators that are finite as the lattice spacing  $a \rightarrow 0$ , as long as the flow time  $\tau > 0$ . As it is customary with the gradient flow, the divergences reappear when the flow time vanishes.

Denoting  $q^{(s)}(\xi, \sqrt{\tau}P_z, \sqrt{\tau}\Lambda_{\text{QCD}}, \sqrt{\tau}M_N)$  the smeared quasi-PDF, we can define the moments

$$b_n^{(s)}(\sqrt{\tau}P_z, \sqrt{\tau}\Lambda_{\text{QCD}}, \sqrt{\tau}M_N) = \int d\xi \xi^{n-1} q^{(s)}(\xi, \sqrt{\tau}P_z, \sqrt{\tau}\Lambda_{\text{QCD}}, \sqrt{\tau}M_N). \quad (40)$$

As mentioned above, note that these are finite and are extrapolated to the continuum limit. A small- $\tau$  expansion allows the smeared moments to be related to the Mellin moments of the light-cone PDFs:

$$b_n^{(s)}(\sqrt{\tau}P_z, \sqrt{\tau}\Lambda_{\text{QCD}}, \sqrt{\tau}M_N) = C_n^{(0)}(\sqrt{\tau}\mu, \sqrt{\tau}P_z) a^{(n)}(\mu) + \mathcal{O}\left(\sqrt{\tau}\Lambda_{\text{QCD}}, \frac{\Lambda_{\text{QCD}}^2}{P_z^2}, \frac{M_N^2}{P_z^2}\right), \quad (41)$$

where  $a^{(n)}(\mu)$  are the matrix elements of renormalized twist-two operator. The corrections to this relation come from target mass corrections, and higher-twist effects; details about these effects can be found in the original publication.

Introducing the kernel

$$\left[ C_n^{(0)} \left( \sqrt{\tau}\mu, \sqrt{\tau}P_z \right) \right]^{-1} = \int dx x^{n-1} Z(x, \sqrt{\tau}\mu, \sqrt{\tau}P_z), \quad (42)$$

Eq. 41 can be trivially inverted in Mellin space, and transformed back to  $x$ -space, yielding

$$f(x, \mu) = \int \frac{d\xi}{\xi} Z \left( \frac{x}{\xi}, \sqrt{\tau}\mu, \sqrt{\tau}P_z \right) q^{(s)}(\xi, \sqrt{\tau}P_z, \sqrt{\tau}\Lambda_{\text{QCD}}, \sqrt{\tau}M_N), \quad (43)$$

where we assumed

$$\Lambda_{\text{QCD}}, M_N \ll P_z \ll 1/\sqrt{\tau}, \quad (44)$$

so that the corrections to this formula are negligible. The kernel function can be computed in perturbation theory, the order of the computation defining the order of the PDF. The DGLAP evolution of  $f(x, \mu)$  is obtained from the  $\mu$  dependence of the kernel  $Z$ . It is worthwhile to emphasise that a large momentum  $P_z$  is required to suppress higher-twist contributions. As long as those are negligible, data for different values of  $P_z$  can be combined to extract PDFs.

## 5 Conclusions

There has been a lot of recent progress in the determination of PDFs on the lattice, and it is impossible to give an extensive review of all the results here. Interesting results that we were not able to cover include the momentum smearing technique [91–93], the usage of the Feynman–Hellmann relation to compute the Compton amplitude [94], or the direct computation of the hadronic tensor of the nucleon [95].

While there are theoretical questions that still need to be settled, recent results suggest that lattice determinations of PDFs will become a reality in the near future. While it is theoretically comforting to be able to compute these quantities from first principles, it would be desirable for the lattice to provide some useful input in the study of physical processes at hadron colliders. For this purpose lattice computations need to match, and eventually improve, the accuracy of the determinations obtained from global fits. The latest global fits yield very precise predictions, at the % level. There are however flavor combinations and/or kinematical regions that are poorly constrained by the data. A benchmarking exercise is currently being carried out, which will compare the different determinations, and assess the scope for progress. It will provide an excellent platform for determining future directions in this field. [96]

The LHC is guiding us into an era of precision physics at hadronic colliders. Perturbative and nonperturbative QCD effects need to be known with great accuracy. As the statistical errors decrease, control over systematics becomes mandatory. The challenge for the theory community is set.

## References

- [1] J.C. Collins, D.E. Soper, G.F. Sterman, *Adv. Ser. Direct. High Energy Phys.* **5**, 1 (1989), [hep-ph/0409313](https://arxiv.org/abs/hep-ph/0409313)
- [2] A. Manohar, P. Nason, G.P. Salam, G. Zanderighi, *Phys. Rev. Lett.* **117**, 242002 (2016), [1607.04266](https://arxiv.org/abs/1607.04266)
- [3] S. Moch, J.A.M. Vermaseren, A. Vogt, *Nucl. Phys.* **B688**, 101 (2004), [hep-ph/0403192](https://arxiv.org/abs/hep-ph/0403192)
- [4] A. Vogt, S. Moch, J.A.M. Vermaseren, *Nucl. Phys.* **B691**, 129 (2004), [hep-ph/0404111](https://arxiv.org/abs/hep-ph/0404111)
- [5] A. Vogt, *Comput. Phys. Commun.* **170**, 65 (2005), [hep-ph/0408244](https://arxiv.org/abs/hep-ph/0408244)
- [6] G.P. Salam, J. Rojo, *Comput. Phys. Commun.* **180**, 120 (2009), [0804.3755](https://arxiv.org/abs/0804.3755)
- [7] V. Bertone, S. Carrazza, J. Rojo, *Comput. Phys. Commun.* **185**, 1647 (2014), [1310.1394](https://arxiv.org/abs/1310.1394)
- [8] J.C. Collins, *Phys. Rev.* **D21**, 2962 (1980)
- [9] J.C. Collins, D.E. Soper, *Nucl. Phys.* **B194**, 445 (1982)
- [10] G. Curci, W. Furmanski, R. Petronzio, *Nucl. Phys.* **B175**, 27 (1980)
- [11] G.F. Sterman, *Nucl. Phys.* **B281**, 310 (1987)
- [12] L. Del Debbio, S. Forte, J.I. Latorre, A. Piccione, J. Rojo (NNPDF), *JHEP* **03**, 039 (2007), [hep-ph/0701127](https://arxiv.org/abs/hep-ph/0701127)
- [13] J.A.M. Vermaseren, A. Vogt, S. Moch, *Nucl. Phys.* **B724**, 3 (2005), [hep-ph/0504242](https://arxiv.org/abs/hep-ph/0504242)
- [14] R.D. Ball et al. (NNPDF), *JHEP* **04**, 040 (2015), [1410.8849](https://arxiv.org/abs/1410.8849)
- [15] R.D. Ball et al. (NNPDF), *Eur. Phys. J.* **C77**, 663 (2017), [1706.00428](https://arxiv.org/abs/1706.00428)
- [16] H. Abramowicz et al. (ZEUS, H1), *Eur. Phys. J.* **C75**, 580 (2015), [1506.06042](https://arxiv.org/abs/1506.06042)
- [17] F.D. Aaron et al. (H1), *Eur. Phys. J.* **C65**, 89 (2010), [0907.2643](https://arxiv.org/abs/0907.2643)
- [18] H. Abramowicz et al. (ZEUS), *JHEP* **09**, 127 (2014), [1405.6915](https://arxiv.org/abs/1405.6915)
- [19] J.J. Aubert et al. (European Muon), *Nucl. Phys.* **B213**, 31 (1983)
- [20] G. Moreno et al., *Phys. Rev.* **D43**, 2815 (1991)
- [21] J.C. Webb et al. (NuSea) (2003), [hep-ex/0302019](https://arxiv.org/abs/hep-ex/0302019)
- [22] J.C. Webb, Ph.D. thesis, New Mexico State U. (2003), [hep-ex/0301031](https://arxiv.org/abs/hep-ex/0301031), [http://lss.fnal.gov/cgi-bin/find\\_paper.pl?thesis=2002-56](http://lss.fnal.gov/cgi-bin/find_paper.pl?thesis=2002-56)
- [23] R.S. Towell et al. (NuSea), *Phys. Rev.* **D64**, 052002 (2001), [hep-ex/0103030](https://arxiv.org/abs/hep-ex/0103030)
- [24] T.A. Aaltonen et al. (CDF), *Phys. Lett.* **B692**, 232 (2010), [0908.3914](https://arxiv.org/abs/0908.3914)
- [25] V.M. Abazov et al. (D0), *Phys. Rev.* **D76**, 012003 (2007), [hep-ex/0702025](https://arxiv.org/abs/hep-ex/0702025)
- [26] T. Aaltonen et al. (CDF), *Phys. Rev.* **D78**, 052006 (2008), [Erratum: *Phys. Rev.* D79, 119902 (2009)], [0807.2204](https://arxiv.org/abs/0807.2204)
- [27] V.M. Abazov et al. (D0), *Phys. Rev.* **D91**, 032007 (2015), [Erratum: *Phys. Rev.* D91, no. 7, 079901 (2015)], [1412.2862](https://arxiv.org/abs/1412.2862)
- [28] V.M. Abazov et al. (D0), *Phys. Rev.* **D88**, 091102 (2013), [1309.2591](https://arxiv.org/abs/1309.2591)
- [29] S. Camarda et al. (HERAFitter developers' Team), *Eur. Phys. J.* **C75**, 458 (2015), [1503.05221](https://arxiv.org/abs/1503.05221)
- [30] G. Aad et al. (ATLAS), *Eur. Phys. J.* **C76**, 291 (2016), [1512.02192](https://arxiv.org/abs/1512.02192)
- [31] V. Khachatryan et al. (CMS), *Phys. Lett.* **B749**, 187 (2015), [1504.03511](https://arxiv.org/abs/1504.03511)
- [32] M. Aaboud et al. (ATLAS), *Eur. Phys. J.* **C77**, 367 (2017), [1612.03016](https://arxiv.org/abs/1612.03016)
- [33] V. Khachatryan et al. (CMS), *Eur. Phys. J.* **C76**, 469 (2016), [1603.01803](https://arxiv.org/abs/1603.01803)
- [34] G. Aad et al. (ATLAS), *Eur. Phys. J.* **C76**, 538 (2016), [1511.04716](https://arxiv.org/abs/1511.04716)
- [35] V. Khachatryan et al. (CMS), *Eur. Phys. J.* **C75**, 542 (2015), [1505.04480](https://arxiv.org/abs/1505.04480)
- [36] G. Aad et al. (ATLAS), *Eur. Phys. J.* **C74**, 3109 (2014), [Addendum: *Eur. Phys. J.* C76, no. 11, 642 (2016)], [1406.5375](https://arxiv.org/abs/1406.5375)

- [37] M. Aaboud et al. (ATLAS), Phys. Lett. **B761**, 136 (2016), [Erratum: Phys. Lett.B772,879(2017)], [1606.02699](https://arxiv.org/abs/1606.02699)
- [38] V. Khachatryan et al. (CMS), JHEP **08**, 029 (2016), [1603.02303](https://arxiv.org/abs/1603.02303)
- [39] G. Aad et al. (ATLAS), JHEP **02**, 153 (2015), [Erratum: JHEP09,141(2015)], [1410.8857](https://arxiv.org/abs/1410.8857)
- [40] V. Khachatryan et al. (CMS), Eur. Phys. J. **C76**, 265 (2016), [1512.06212](https://arxiv.org/abs/1512.06212)
- [41] G. Aad et al. (ATLAS), JHEP **06**, 112 (2014), [1404.1212](https://arxiv.org/abs/1404.1212)
- [42] R. Aaij et al. (LHCb), JHEP **08**, 039 (2015), [1505.07024](https://arxiv.org/abs/1505.07024)
- [43] R. Aaij et al. (LHCb), JHEP **01**, 155 (2016), [1511.08039](https://arxiv.org/abs/1511.08039)
- [44] W. Beenakker, C. Borschensky, M. Krämer, A. Kulesza, E. Laenen, S. Marzani, J. Rojo, Eur. Phys. J. **C76**, 53 (2016), [1510.00375](https://arxiv.org/abs/1510.00375)
- [45] M. Arneodo et al. (New Muon), Nucl. Phys. **B487**, 3 (1997), [hep-ex/9611022](https://arxiv.org/abs/hep-ex/9611022)
- [46] M. Arneodo et al. (New Muon), Nucl. Phys. **B483**, 3 (1997), [hep-ph/9610231](https://arxiv.org/abs/hep-ph/9610231)
- [47] L.W. Whitlow, E.M. Riordan, S. Dasu, S. Rock, A. Bodek, Phys. Lett. **B282**, 475 (1992)
- [48] A.C. Benvenuti et al. (BCDMS), Phys. Lett. **B223**, 485 (1989)
- [49] A.C. Benvenuti et al. (BCDMS), Phys. Lett. **B237**, 592 (1990)
- [50] G. Onengut et al. (CHORUS), Phys. Lett. **B632**, 65 (2006)
- [51] M. Goncharov et al. (NuTeV), Phys. Rev. **D64**, 112006 (2001), [hep-ex/0102049](https://arxiv.org/abs/hep-ex/0102049)
- [52] D.A. Mason (2006), fFERMILAB-THESSIS-2006-01
- [53] H. Abramowicz et al. (H1 ,ZEUS), Eur.Phys.J. **C73**, 2311 (2013), [1211.1182](https://arxiv.org/abs/1211.1182)
- [54] A. Abulencia et al. (CDF), Phys. Rev. **D75**, 092006 (2007), [Erratum: Phys. Rev.D75,119901(2007)], [hep-ex/0701051](https://arxiv.org/abs/hep-ex/0701051)
- [55] G. Aad et al. (ATLAS), Phys. Rev. **D85**, 072004 (2012), [1109.5141](https://arxiv.org/abs/1109.5141)
- [56] G. Aad et al. (ATLAS), Phys.Lett. **B725**, 223 (2013), [1305.4192](https://arxiv.org/abs/1305.4192)
- [57] G. Aad et al. (ATLAS), JHEP **09**, 145 (2014), [1406.3660](https://arxiv.org/abs/1406.3660)
- [58] G. Aad et al. (ATLAS), Phys. Rev. **D86**, 014022 (2012), [1112.6297](https://arxiv.org/abs/1112.6297)
- [59] G. Aad et al. (ATLAS), Eur.Phys.J. **C73**, 2509 (2013), [1304.4739](https://arxiv.org/abs/1304.4739)
- [60] S. Chatrchyan et al. (CMS), Phys.Rev.Lett. **109**, 111806 (2012), [1206.2598](https://arxiv.org/abs/1206.2598)
- [61] S. Chatrchyan et al. (CMS), Phys.Rev. **D90**, 032004 (2014), [1312.6283](https://arxiv.org/abs/1312.6283)
- [62] S. Chatrchyan et al. (CMS), JHEP **02**, 013 (2014), [1310.1138](https://arxiv.org/abs/1310.1138)
- [63] S. Chatrchyan et al. (CMS), JHEP **1312**, 030 (2013), [1310.7291](https://arxiv.org/abs/1310.7291)
- [64] V. Khachatryan et al. (CMS), Eur. Phys. J. **C75**, 147 (2015), [1412.1115](https://arxiv.org/abs/1412.1115)
- [65] S. Chatrchyan et al. (CMS), Phys.Rev. **D87**, 112002 (2013), [1212.6660](https://arxiv.org/abs/1212.6660)
- [66] C. Collaboration (CMS) (2016)
- [67] R. Aaij et al. (LHCb), JHEP **1206**, 058 (2012), [1204.1620](https://arxiv.org/abs/1204.1620)
- [68] R. Aaij et al. (LHCb), JHEP **1302**, 106 (2013), [1212.4620](https://arxiv.org/abs/1212.4620)
- [69] L.A. Harland-Lang, A.D. Martin, P. Motylinski, R.S. Thorne, Eur. Phys. J. **C75**, 204 (2015), [1412.3989](https://arxiv.org/abs/1412.3989)
- [70] S. Dulat, T.J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, C.P. Yuan, Phys. Rev. **D93**, 033006 (2016), [1506.07443](https://arxiv.org/abs/1506.07443)
- [71] S. Alekhin, J. Blümlein, S. Moch, R. Placakyte, Phys. Rev. **D96**, 014011 (2017), [1701.05838](https://arxiv.org/abs/1701.05838)
- [72] A. Accardi, L.T. Brady, W. Melnitchouk, J.F. Owens, N. Sato, Phys. Rev. **D93**, 114017 (2016), [1602.03154](https://arxiv.org/abs/1602.03154)
- [73] X. Ji, Phys. Rev. Lett. **110**, 262002 (2013), [1305.1539](https://arxiv.org/abs/1305.1539)
- [74] X. Ji, Sci. China Phys. Mech. Astron. **57**, 1407 (2014), [1404.6680](https://arxiv.org/abs/1404.6680)

- [75] M. Constantinou, H. Panagopoulos, Phys. Rev. **D96**, 054506 (2017), [1705.11193](https://arxiv.org/abs/1705.11193)
- [76] M. Constantinou, H. Panagopoulos, *Perturbative Renormalization of Wilson line operators*, in *35th International Symposium on Lattice Field Theory (Lattice 2017) Granada, Spain, June 18-24, 2017* (2017), [1711.00543](https://arxiv.org/abs/1711.00543), <https://inspirehep.net/record/1634256/files/arXiv:1711.00543.pdf>
- [77] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjyiannakou, K. Jansen, H. Panagopoulos, F. Steffens, Nucl. Phys. **B923**, 394 (2017), [1706.00265](https://arxiv.org/abs/1706.00265)
- [78] J.W. Chen, T. Ishikawa, L. Jin, H.W. Lin, Y.B. Yang, J.H. Zhang, Y. Zhao (2017), [1706.01295](https://arxiv.org/abs/1706.01295)
- [79] J.W. Chen, T. Ishikawa, L. Jin, H.W. Lin, Y.B. Yang, J.H. Zhang, Y. Zhao (2017), [1710.01089](https://arxiv.org/abs/1710.01089)
- [80] J. Green, K. Jansen, F. Steffens (2017), [1707.07152](https://arxiv.org/abs/1707.07152)
- [81] G.C. Rossi, M. Testa, Phys. Rev. **D96**, 014507 (2017), [1706.04428](https://arxiv.org/abs/1706.04428)
- [82] C.E. Carlson, M. Freid, Phys. Rev. **D95**, 094504 (2017), [1702.05775](https://arxiv.org/abs/1702.05775)
- [83] R.A. Briceño, M.T. Hansen, C.J. Monahan, Phys. Rev. **D96**, 014502 (2017), [1703.06072](https://arxiv.org/abs/1703.06072)
- [84] X. Xiong, X. Ji, J.H. Zhang, Y. Zhao, Phys. Rev. **D90**, 014051 (2014), [1310.7471](https://arxiv.org/abs/1310.7471)
- [85] Y.Q. Ma, J.W. Qiu (2014), [1404.6860](https://arxiv.org/abs/1404.6860)
- [86] T. Ishikawa, Y.Q. Ma, J.W. Qiu, S. Yoshida (2017), [1707.03107](https://arxiv.org/abs/1707.03107)
- [87] T. Ishikawa, Y.Q. Ma, J.W. Qiu, S. Yoshida (2016), [1609.02018](https://arxiv.org/abs/1609.02018)
- [88] C. Monahan, K. Orginos, JHEP **03**, 116 (2017), [1612.01584](https://arxiv.org/abs/1612.01584)
- [89] M. Luscher, P. Weisz, JHEP **02**, 051 (2011), [1101.0963](https://arxiv.org/abs/1101.0963)
- [90] M. Luscher, JHEP **04**, 123 (2013), [1302.5246](https://arxiv.org/abs/1302.5246)
- [91] G.S. Bali, V.M. Braun, M. Göckeler, M. Gruber, F. Hutzler, P. Korcyl, B. Lang, A. Schäfer, P. Wein, J.H. Zhang, *Pion distribution amplitude from Euclidean correlation functions*, in *35th International Symposium on Lattice Field Theory (Lattice 2017) Granada, Spain, June 18-24, 2017* (2017), [1709.04325](https://arxiv.org/abs/1709.04325), <https://inspirehep.net/record/1623222/files/arXiv:1709.04325.pdf>
- [92] G.S. Bali, V.M. Braun, M. Göckeler, M. Gruber, F. Hutzler, P. Korcyl, B. Lang, A. Schäfer (RQCD), Phys. Lett. **B774**, 91 (2017), [1705.10236](https://arxiv.org/abs/1705.10236)
- [93] G.S. Bali, B. Lang, B.U. Musch, A. Schäfer, Phys. Rev. **D93**, 094515 (2016), [1602.05525](https://arxiv.org/abs/1602.05525)
- [94] A.J. Chambers, R. Horsley, Y. Nakamura, H. Perlt, P.E.L. Rakow, G. Schierholz, A. Schiller, K. Somfleth, R.D. Young, J.M. Zanotti, Phys. Rev. Lett. **118**, 242001 (2017), [1703.01153](https://arxiv.org/abs/1703.01153)
- [95] J. Liang, K.F. Liu, Y.B. Yang, *Lattice calculation of hadronic tensor of the nucleon*, in *35th International Symposium on Lattice Field Theory (Lattice 2017) Granada, Spain, June 18-24, 2017* (2017), [1710.11145](https://arxiv.org/abs/1710.11145), <https://inspirehep.net/record/1633593/files/arXiv:1710.11145.pdf>
- [96] H. Lin, E. Nocera, F. Olness, K. Orginos, J.e. Rojo