

VU Research Portal

Linear vs non-linear QCD evolution

Albacete, J. L.; Milhano, J. G.; Quiroga-Arias, P.; Rojo, J.

published in

European Physical Journal C. Particles and Fields
2012

DOI (link to publisher)

[10.1140/epjc/s10052-012-2131-x](https://doi.org/10.1140/epjc/s10052-012-2131-x)

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Albacete, J. L., Milhano, J. G., Quiroga-Arias, P., & Rojo, J. (2012). Linear vs non-linear QCD evolution: from HERA data to LHC phenomenology. *European Physical Journal C. Particles and Fields*, *C72*, 2131. <https://doi.org/10.1140/epjc/s10052-012-2131-x>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Linear vs non-linear QCD evolution: from HERA data to LHC phenomenology

J. L. Albacete^{a,b}, J. G. Milhano^{c,d}, P. Quiroga-Arias^e, J. Rojo^d

^a*IPNO, Université Paris-Sud, CNRS/IN2P3, F-91406, Orsay, France.*

^b*IPhT, CEA/Saclay, 91191 Gif-sur-Yvette cedex, France.*

^c*CENTRA, Instituto Superior Técnico, Universidade Técnica de Lisboa, Av. Rovisco Pais, P-1049-001
Lisboa, Portugal.*

^d*Physics Department, Theory Unit, CERN, CH-1211 Genève 23, Switzerland.*

^e*LPTHE, UPMC Univ. Paris 6 and CNRS UMR7589, Paris, France*

Abstract

The very precise combined HERA data provides a testing ground in which the relevance of novel QCD regimes, other than the successful linear DGLAP evolution, in small- x inclusive DIS data can be ascertained. We present a study of the dependence of the AAMQS fits, based on the running coupling BK non-linear evolution equations (rcBK), on the fitted dataset. This allows for the identification of the kinematical region where rcBK accurately describes the data, and thus for the determination of its applicability boundary. We compare the rcBK results with NNLO DGLAP fits, obtained with the NNPDF methodology with analogous kinematical cuts. Further, we explore the impact on LHC phenomenology of applying stringent kinematical cuts to the low- x HERA data in a DGLAP fit.

Contents

1	Introduction	2
2	rcBK and DGLAP evolution	4
3	NNLO DGLAP analysis with kinematical cuts	8
4	rcBK analysis with kinematical cuts	10
5	Implications for LHC phenomenology	13
6	Conclusions and outlook	17

*Preprint number: CERN-PH-TH/2012-060

Email addresses: albacete@ipno.in2p3.fr (J. L. Albacete), guilherme.milhano@ist.utl.pt (J. G. Milhano), pquiroga@lpthe.jussieu.fr (P. Quiroga-Arias), juan.rojo@cern.ch (J. Rojo)

1. Introduction

The knowledge of the partonic structure of the proton at all relevant observation scales plays a crucial role in the analysis of data from present high-energy hadronic colliders, most notably at the LHC. In practice such information is provided by phenomenological parton fits to previously existing data based on the use of perturbative QCD renormalization group equations and, in the framework of factorization theorems, it is used as input for establishing the theoretical expectations and uncertainties for the production rates of any process of interest.

The different QCD approaches for the description of the scale dependence of parton distribution functions – analogously, for gauge invariant operators encoding the parton flux into the collision – share a similar strategy in resumming to all orders radiative terms enhanced by large logarithms. The most widely used framework are the DGLAP equations [1–3], that is, the renormalization group equations that describe the scale dependence of parton distribution functions through a resummation of large logarithms $\sim \alpha_s \ln Q^2/Q_0^2$ with Q_0 some initial scale. The DGLAP equations have been successfully and intensively tested against experimental data and, together with asymptotic freedom and factorization theorems, provide a fundamental tool for establishing controlled predictions at the LHC. Successful as they are, the DGLAP equations are also expected to break down in some kinematic regimes. In particular, at small values of Bjorken- x , large logarithms $\sim \alpha_s \ln(x_0/x)$ emerge and need to be resummed to all orders.

In turn, analogous resummation schemes aimed at describing the small- x evolution of hadron structure have also been developed. In this direction in the kinematic plane, orthogonal to DGLAP evolution, the relevant logarithms are $\sim \alpha_s \ln(x_0/x)$, resummed to all orders in the BFKL approach [4, 5]. Let us recall that although a consistent formulation of DGLAP evolution which also accounts for small- x resummation has been formulated [6–9] in the recent years, its phenomenological consequences have not yet been fully explored. Additionally, the enhancement of gluon emission at small- x naturally leads to the – empirically observed – presence of large gluon densities and to the need of non-linear recombination terms in order to stabilize the diffusion towards the infrared characteristic of BFKL evolution. Most importantly, the presence of non-linear terms is ultimately related to the preservation of unitarity of the theory.

Both the resummation of small- x logarithms and the inclusion of non-linear density dependent corrections are consistently accounted for by the B-JIMWLK [10–14] equations. The presence of non-linear terms in the small- x evolution equations limits the growth rate of gluon number densities for modes of transverse momentum smaller than the saturation scale Q_s (see section 2 for a precise definition). This novel, semi-hard dynamical scale marks the onset of non-linear corrections in QCD evolution and leads to distinctive dynamical effects such as the generation of geometric scaling [15].

An important improvement in devising a phenomenological tool for the empiric study of the saturation phenomenon based on first principles QCD was brought by the recent determination of higher order corrections to the B-JIMWLK equations [16–19]. In particular, it was shown in [20] that the BK equation [10, 11] – the large- N_c limit of the full B-JIMWLK hierarchy – including running coupling corrections as calculated in [16–18] was compatible with experimental data from different collision systems. This expectation was confirmed by

the AAMQS collaboration in [21], where global fits to e+p data based on the use of the running coupling BK equation (referred to as rcBK equation henceforth) were performed for the first time, providing a good description of data. This result was then updated in [22] to include in the fitted data set the combined H1+ZEUS data [23] on the reduced cross section and also include the contribution of heavy quarks. The AAMQS data set covers the region $x < 0.01$ and $Q^2 < 50 \text{ GeV}^2$, including data in the photoproduction region where $Q^2 \ll 1 \text{ GeV}^2$. A similar strategy was followed in [24], where data on the diffractive cross section was also fitted consistently. Overall, a very good description of data is provided in those three studies. Although the rcBK fitting technology has not yet reached the level of sophistication and accuracy of DGLAP based approaches, it shares a similar strategy and methodology, thus allowing for systematic comparison to DGLAP studies, as we intend to do in this work.

Theoretical arguments alone can only strictly establish the applicability of either DGLAP or BK in their asymptotic limits of very large Q^2 (for DGLAP) or very small x (for BK). On the phenomenological side, where intermediate (x, Q^2) kinematics is probed, the situation remains unclear. On one hand, DGLAP based approaches have continuously reported good fits to all available data above some initial scale $Q_0^2 \sim 1 \div 4 \text{ GeV}^2$ and for x -values as small as $\sim 10^{-5}$. On the other hand, similarly good fits to data in the same kinematic region are achieved using the rcBK evolution equation, including data points up to $Q^2 = 50 \text{ GeV}^2$. Moreover, some features identified in data such as geometric scaling or the ratio of total over diffractive cross sections are naturally explained in a non-linear evolution framework, while they are difficult to reconcile with the strictly linear approach. Although geometric scaling has been shown to be fully compatible with the linear DGLAP equations [25], the dynamical scale generation at its origin in the non-linear context is not viable in a linear framework.

Clearly, the reliability of one or the other approach in the region of moderate (x, Q^2) cannot be determined on a priori theoretical arguments. It is also clear that claims in favor of one particular approach should not be done solely on the basis of agreement with experimental data: it is well known that one can obtain an excellent fit to the HERA low- x data with a very reduced number of free parameters [26]. Such is a necessary but not sufficient condition. Further, beyond describing existing data, the usefulness of a given approach rests on its predictive power towards kinematic regions experimentally unexplored so far. This latter condition rules out phenomenological data descriptions such as the structure function parametrizations of Refs. [27, 28]. The predictivity requirement is clearly satisfied by both the BK and DGLAP approaches, both endowed with a well defined QCD dynamical input that we briefly review in Sect. 2. Their predictive power is, however, oriented towards different directions in the kinematic plane: small- x and high- Q^2 respectively.

A pertinent question to ask is whether corrections to the limit in which both formalisms are well defined, large- Q^2 and small- x respectively, are important in intermediate kinematic regions and, if so, to what extent this missing dynamical effects could be accommodated in the boundary conditions for the respective rcBK or DGLAP evolution. Thus, it may be that the flexibility in the initial conditions for DGLAP evolution is hiding some interesting QCD dynamics, namely the presence of non-linear behavior or the onset of the regime dominated by high energy QCD (small- x resummation) not included in DGLAP fits at their present degree of accuracy. A first attempt to answer this question was carried out in [29, 30]. Also, it is conceivable that higher twist corrections not included in the collinear factorization framework may be sizable at moderate and small Q^2 . Analogous concerns arise in relation to

the applicability of the rcBK approach as it is conceivable that the maximal value $x_0 = 0.01$ considered in the AAMQS fits might not be small enough for the dipole formalism they rely upon to be fully applicable. In particular, it remains to be understood why data at $Q^2 = 50 \text{ GeV}^2$ can be described in a rcBK fit, far beyond the naive range of application of the nonlinear description. In that case, sizable energy conservation corrections, negligible at very small- x , may influence the values of the fit parameters.

A natural step towards elucidating whether interesting dynamics is hidden or absorbed in the boundary conditions in either approach is to systematically displace those boundaries and assess the fit stability under such changes or, alternatively, look for correlations between fit parameters and the position of the boundary. These boundary conditions are the Parton Distribution Functions (PDFs) for DGLAP evolution and the Unintegrated Gluon Distribution (UGD) for BK evolution. Such sensitivity would indicate that the resulting PDFs (in the case of DGLAP) or UGDs (in the case of rcBK) extracted from the fit are contaminated with physics effects beyond the dynamical content of the respective evolution equations. In Sect. 2 we describe the strategy pursued in [29, 30] to search for systematic deviations of DGLAP fits by imposing selected kinematic cuts to the data set. We shall extend it in an analogous fashion to the framework of rcBK fits. The results for the updated DGLAP and rcBK fits to HERA data with kinematic cuts are presented and discussed in Sects. 3 and 4 respectively.

The presence of any systematic deviations from fixed order DGLAP evolution will have direct implications for LHC phenomenology: indeed, parton distributions extracted from a DGLAP analysis which includes HERA data in the small- (x, Q^2) region are then evolved upwards in Q^2 to predict any LHC processes. In Sect. 5 we shall quantify and explore how this potential source of theoretical uncertainty propagates towards high Q^2 scales by computing benchmark LHC cross sections with PDF sets both with and without the small- x kinematical cuts. Conclusions and outlook are presented in Sect. 6.

2. rcBK and DGLAP evolution

In order to frame the discussion and facilitate the interpretation of the results presented in the next sections, and before describing the strategy adopted to chart the small- (x, Q^2) kinematic territory, we first review very succinctly the main features of both BK and DGLAP approaches. Using rather compact notation¹ the BK (for the dipole scattering amplitude $\mathcal{N}(x, r)$) and DGLAP evolution equations (for vector PDFs $f(x, Q^2)$) can be written as follows:

$$\text{BK: } \frac{\partial \mathcal{N}(r, x)}{\partial \ln(x_0/x)} = \int d^2 r_1 \mathcal{K}(r, r_1, r_2) [\mathcal{N}(r_1, x) + \mathcal{N}(r_2, x) - \mathcal{N}(r, x) - \mathcal{N}(r_1, x)\mathcal{N}(r_2, x)] ; \quad (1)$$

$$\text{DGLAP: } \frac{\partial f(x, Q^2)}{\partial \ln(Q^2/Q_0^2)} = \int_x^1 \frac{dy}{y} P(\alpha_s(Q^2), x/y) f(y, Q^2). \quad (2)$$

The object evolved by the BK equation is the (imaginary part of the) quark-antiquark color dipole amplitude to scatter off a hadronic target, $\mathcal{N}(r, x)$, where r is the transverse

¹The BK equation written here assumes translational invariance of the dipole amplitude \mathcal{N} . Also, some constant factors have been absorbed in the evolution kernel \mathcal{K} , see e.g. [22] for precise definitions.

coordinate and x is the usual Bjorken scaling variable in deep-inelastic scattering. The onset of the black-disk limit is given by the condition $\mathcal{N}(r_s = 1/Q_s(x), x) = \kappa \sim 1$, which can also be used to define the saturation scale, $Q_s(x)$. The amplitude $\mathcal{N}(r, x)$ is the main dynamical ingredient of the dipole formulation of deep inelastic scattering, extensively used in the phenomenological searches of non-linear dynamics in data, including the AAMQS global fits. In this framework the virtual photon-proton cross section for transverse (T) and longitudinal (L) polarization of the virtual photon reads

$$\sigma_{T,L}(x, Q^2) = 2\sigma_0 \sum_f \int_0^1 dz \int d\mathbf{r} |\Psi_{T,L}^f(e_f, m_f, z, Q^2, \mathbf{r})|^2 \mathcal{N}(r, x), \quad (3)$$

where $\Psi_{T,L}^f$ is the light-cone wave function for a virtual photon to fluctuate into a quark-antiquark dipole of quark flavor f . Other observables of interest, such as structure functions F_2 , F_L , are straightforwardly related to $\sigma_{T,L}$ in Eq. (3). A clearer physical interpretation of the dipole amplitude is obtained by recalling that its Fourier transform yields the unintegrated gluon distribution (UGD) of the target: $\phi(x, k_t) = \int d^2r e^{i\vec{r}\cdot\vec{k}_t} \mathcal{N}(r, x)$. In turn, to LO accuracy the UGD is related to the standard integrated gluon distribution: $xg(x, Q^2) = \int^{Q^2} d^2k_t \phi(x, k_T)$. Thus, the dipole amplitude provides direct information on the gluon content of the hadron. The kernel \mathcal{K} of the BK equation is obtained via perturbative QCD calculations for small- x gluon emission, in a fashion analogous to that employed for the calculation of the DGLAP splitting functions, $P(x, \alpha_s)$ in Eq. (2). The evolution kernel has been calculated up to NLO accuracy in the resummation variable $\alpha_s \ln(x_0/x)$ [19]. Alternatively, a partial subset of NLO corrections can be resummed to all orders to obtain the running coupling BK equation used later in this work.

Three main distinctive features of the BK equation deserve to be highlighted. First, it is an evolution equation in Bjorken- x , i.e. it provides the change of hadron structure when smaller values of x are probed. As such, it has no predictive power in the orthogonal Q^2 -direction, with all the Q^2 dependence of structure functions computed with rcBK fixed by the wave function in Eq. (3). Second, it is a non-linear equation. The presence of non-linear terms is required by preservation of unitarity², and can be interpreted, in the appropriate gauge and frame, as due to gluon recombination processes. Third, it resums all twists or multiple scatterings, and thus is applicable also for small values of Q^2 .

In contrast, the DGLAP equation, Eq. (2), provides the Q^2 evolution of proton structure. The matrix of splitting functions that govern the evolution are known up to $\mathcal{O}(\alpha_s^2)$ in pQCD [31, 32]. It has no predictive power in the orthogonal x -direction: in a DGLAP analysis, for values of $x \leq x_{\min}$, with x_{\min} the smallest value of x of the data included, the DGLAP predictions become unreliable. Also, it is a linear, leading twist equation, two conditions that are expected to break down for sufficiently small values of Q^2 , where gluon densities are higher and the contribution from higher twists may be important.

Contact between these two orthogonal approaches can be made in the diagonal limit of $x \rightarrow 0$ and $Q^2 \rightarrow \infty$, where both BK and DGLAP equations converge to the Double Logarithmic Approximation (see e.g. [11]). However, no smooth interpolation between the

²It is indeed straightforward to check that the non-linear term in Eq. (1) prevents the dipole scattering amplitude to grow above unity, provided the initial condition is unitary itself, i.e $\mathcal{N} \leq 1$.

two is known to date in the more interesting phenomenological region of moderate x and Q^2 .

Both the DGLAP or the BK equations are initial value problems, i.e. they are well defined only after initial conditions at the initial evolution scale have been provided. This introduces free parameters, ultimately of non-perturbative origin, to be fitted to data. In the case of BK, initial conditions for the dipole amplitude (the UGD) should be specified at the initial evolution scale $x = x_0$ and for all values of the the dipole size r (equivalently, for all values of³ Q^2). In the AAMQS rcBK fits to HERA data the initial conditions are taken in the form

$$\mathcal{N}(r, x = x_0) = 1 - \exp\left(-\frac{(r^2 Q_{s0}^2)^\gamma}{4} \ln\left(\frac{1}{r\Lambda_{\text{QCD}}} + e\right)\right), \quad (4)$$

where Q_{s0} and γ are free parameters to be determined from the data. For DGLAP, initial conditions for the PDFs have to be provided at some low initial scale Q_0^2 and for all values of x : $xf(Q^2 = Q_0^2, x)$. Within the NNPDF approach, these initial conditions are parametrized with artificial neural networks, which provide universal interpolants and avoid any theoretical bias arising from the choice of a particular functional form for the input PDFs. Note that, in principle, this technology could also be applied for the initial conditions of the rcBK analysis.

Lacking any better quantitative criterion, the choice of initial scale – x_0 for BK or Q_0^2 for DGLAP – that limits the fitting the data set in global fits is commonly determined a posteriori: if good quality fits to data can be obtained with a given choice of the initial scale, then the validity of the applied evolution scheme down to such scale is accepted. However, this empirical, self-consistent criterion can be a misleading one. It is perfectly conceivable that the adjustable parameters entering the initial conditions in either evolution scheme may absorb some relevant dynamics not properly described by the physics content comprised in the equations themselves.

Thus, one needs to define some suitable strategy to identify the regimes of validity of each formalism and quantify the potential deviation from these. Such a strategy to search for statistically significant deviations from DGLAP evolution was laid down in Refs. [29, 30]. There, subsets of data on the reduced lepton–proton DIS cross section $\sigma_r(x, Q^2)$ measured at HERA [23], below some given kinematic cuts,

$$Q^2 \leq Q_{\text{cut}}^2 \equiv A_{\text{cut}} x^{-\lambda} \quad (5)$$

with $\lambda \sim 0.3$ and different values of A_{cut} , were excluded from the data set, so that the PDFs were fitted excluding the potentially dangerous region. The cuts were motivated by the generic expectation that possible deviations from fixed order DGLAP are larger at small- x and Q^2 . Then, backwards DGLAP evolution was used to compare with the excluded data, thus providing a direct test of fit stability under changes in the boundary conditions.

The analysis of [29, 30], carried at NLO in the massless scheme for structure functions, found a systematic discrepancy, albeit with not large enough statistical significance for a decisive statement to be made. The observed discrepancy indicated that the parton evolution in the unfitted kinematic region may not be fully accounted for by the physics encoded in NLO DGLAP equations. Rather, additional dynamics should be responsible for it. Such

³The photon wave function $\Psi_{T,L}^f$ in Eq. (3) peaks at $r \sim 2/Q$.

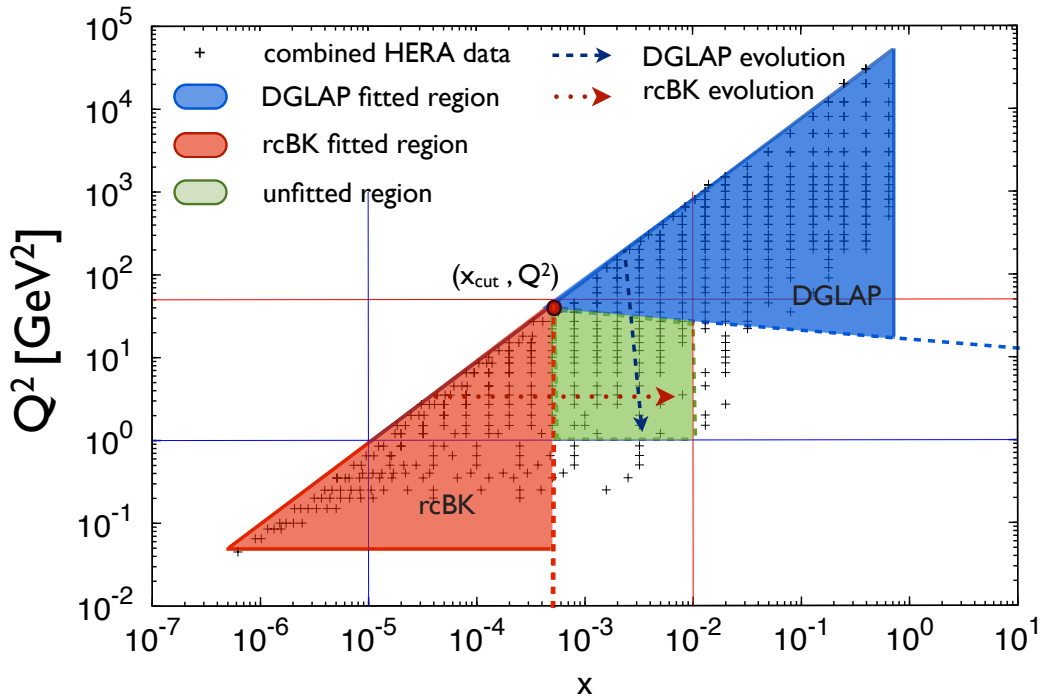


Figure 1: Sketch of the kinematic plane with cuts for DGLAP and rcBK fits. The arrows indicate backwards evolution in either formalism to the unfitted *test* region.

deviations are qualitatively consistent with the behavior predicted by small- x perturbative resummation [7], but incompatible with next-to-next-to-leading order corrections. Also, it was suggested that an improved treatment of the heavy quark masses may have a sizable impact for the relatively low Q^2 values in the region excluded by the saturation inspired kinematical cuts mentioned above.

In this work, we shall study the stability of the AAMQS fits with respect to the choice of dataset following an analogous procedure. The AAMQS fits to HERA data, based on the rcBK non-linear evolution equations included data in the region $x < x_0 = 0.01$ and $Q^2 < 50$ GeV². We shall perform fits to data using the AAMQS set up for the case of only three active flavors (the lightest ones) with 4 free parameters (Q_0^2 , γ , σ_0 and C) as described in detail in [22]. We shall systematically reduce the largest experimental value of x included in the fit, which we denote by x_{cut} , and then use the parametrization for the dipole scattering amplitude resulting from the fit to predict the value of the reduced cross section in the unfitted region $x_{\text{cut}} < x < x_0$.

The rcBK equation at the basis of the AAMQS approach is a non-linear equation and, therefore, extremely unstable under backwards evolution. Equivalently, as it is well known, the solutions of the BK equation at asymptotically small- x are universal, i.e independent of the initial conditions. This immediately implies that backward BK evolution is not well defined. Thus, even for fits with a cut in the Bjorken variable at x_{cut} , we shall start the evolution, and hence determine the boundary conditions, at $0.01 = x_0 > x_{\text{cut}}$. This will allow

us to know the dipole scattering amplitude corresponding to fits with cuts in all the unfitted region, i.e. $x_{\text{cut}} < x < x_0$. The fitting strategy for the DGLAP and rcBK analysis with kinematical cuts is summarized in Fig. 1.

3. NNLO DGLAP analysis with kinematical cuts

Before presenting the results for the rcBK fits with cuts, we present an updated DGLAP fit with kinematic cuts, along the lines of those presented in Refs. [29, 30], but now based on the NNPDF2.1 NNLO set [33]. NNPDF2.1 NNLO is based on the FONLL-C GM-VFN scheme [34] for an accurate treatment of heavy quarks in DIS, while the original studies [29, 30] were based on PDFs sets extracted in the ZM-VFN scheme [35–37]. Although we will refrain from a detailed discussion of the NNPDF2.1 NNLO cut fits that were produced for comparison with the rcBK fit with cuts presented below, we briefly explain our modus operandi and highlight some of its features.

The reference DGLAP fit is NNPDF2.1, a NNLO fit to all data ($A_{\text{cut}} = 0$) above $Q^2 > 3 \text{ GeV}^2$, including $\mathcal{O}(3500)$ experimental points from DIS, DY, W/Z production and inclusive jet data. Then, based on that fit, the χ^2 per number of data points⁴ is calculated for different subsets of interest. Of relevance to this work is the neutral current positron subset, HERA-I NCp, since these are the only data included in the rcBK analysis, and also the only data affected by the kinematic cuts. We have then performed a new NNLO DGLAP fit including only those data points which survive the cut $A_{\text{cut}} = 1.5$. The NNLO PDFs with $A_{\text{cut}} = 1.5$, see Eq. (5), are compared to the reference NNPDF2.1 NNLO PDFs (with $A_{\text{cut}} = 0$) in Fig. 2, where the two PDFs that are mostly affected by the applied kinematical cuts – the singlet and the gluon – are shown, both at the initial scale and at $Q^2 = 10^4 \text{ GeV}^2$, a typical value for LHC phenomenology. In the latter case we show the ratios with respect to the reference uncut fit. This comparison shows that while PDF uncertainties increase substantially when the low- x , low- Q^2 HERA data are removed, the two PDF sets are always consistent at the one-sigma level. The differences remain when DGLAP evolution is used to determine the PDFs at a typical LHC scale of $Q^2 = 10^4 \text{ GeV}^2$ where both the gluon and the singlet are larger at small- x as compared to the uncut fit.

It is known that in the small- x and large Q^2 limit, the trajectories dictated by DGLAP evolution become independent of the initial condition. To quantify if this asymptotic regime is reached for realistic kinematics, we show the distances between PDFs⁵ as the scale is raised in Fig. 3. This comparison allows us to quantify how DGLAP evolution modifies the differences in the initial scale PDFs as the scale is raised. Let us recall that once the initial conditions are fixed at some low scale Q_0^2 , the PDFs at any other scale $Q^2 \geq Q_0^2$ are completely determined by the DGLAP evolution equations. As one can see, DGLAP evolution modifies the PDFs but does not wash up the initial differences, even for scales as large as $Q^2 = 10^8$

⁴In the NNPDF analysis the full information on the correlated systematics is taken into account into the definition of the χ^2 , and the normalization uncertainties are included following the t_0 prescription [38]. Let us recall that non negligible differences are expected if the systematic uncertainties are added in quadrature to the statistical error, and that the χ^2 will be artificially about 15% lower in this latter approximation [23].

⁵The distances defined to compare PDFs in the NNPDF framework are defined in the Appendix B of Ref. [37]. They should not be confused with the relative and statistical distances, eqs. (6) and (7), introduced later in this work.

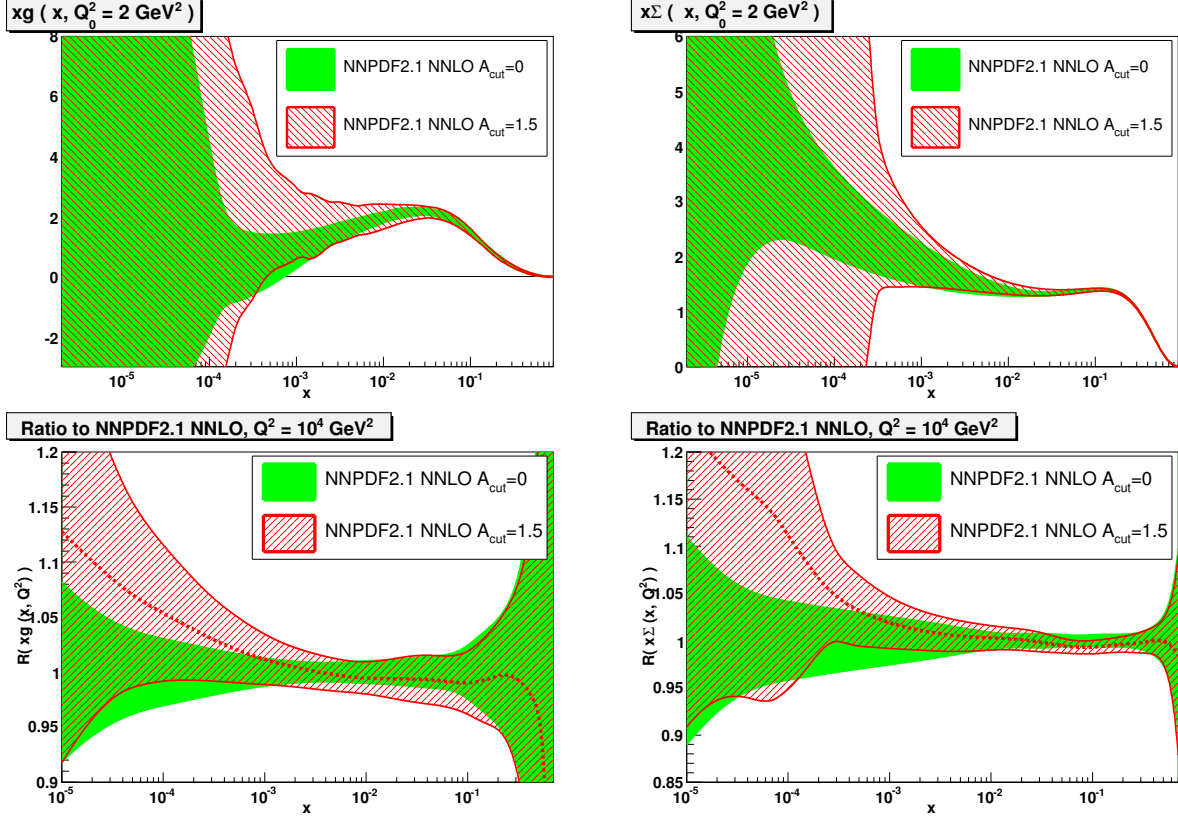


Figure 2: Comparison between the NNPDF2.1 NNLO PDFs with $A_{\text{cut}} = 1.5$ and $A_{\text{cut}} = 0$ for the gluon (left plots) and the quark singlet (right plots). We perform the comparison of the absolute PDFs at the initial scale $Q_0^2 = 2 \text{ GeV}^2$ in the upper plots, and we compare the ratio with respect the reference set at a typical LHC scale $Q^2 = 10^4 \text{ GeV}^2$ in the lower plots.

GeV^2 , shifting these differences towards smaller- x values, as clearly seen for example in the gluon PDF.

Other PDF fitting groups have investigated the same topic, with inconclusive results so far. In the CT collaboration study [39], based on the CT10 NLO analysis, the dependence of the fit quality to the HERA-I data with the same kinematical cuts as in Refs. [29, 30] was studied. It was found that some deviation for the smallest x and Q^2 was obtained only for given choices of PDF parametrizations, while for other parametrizations the results were consistent with NLO QCD. This is not in contradiction with the NNPDF analysis, which instead probe simultaneously a wide range of possible initial conditions. A full clarification would require, within the CT framework, the exploration of a larger variety of initial PDF conditions in order to statistically quantify the significance of the possible deviations with respect to fixed order DGLAP evolution.

The HERAPDF collaboration has also studied the impact of raising the kinematical cuts and assessed the variation of fit quality when going from NLO to NNLO. In Ref. [40] a NLO QCD analysis was performed on the combined HERA-I data supplemented by HERA-II taken at lower beam energies. When the kinematical cut was raised to $Q^2 \geq 5 \text{ GeV}^2$ the resulting small- x singlet and gluon were found to be outside the reference PDF uncertainty

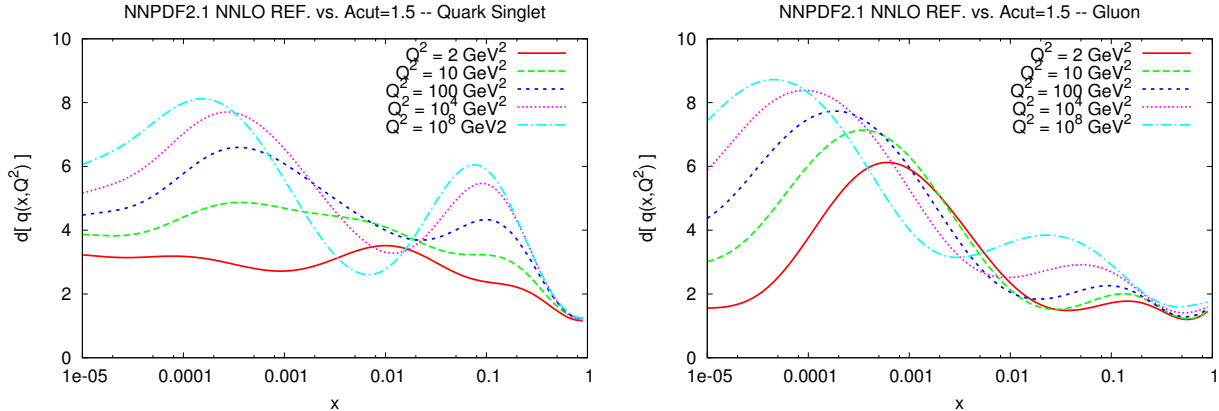


Figure 3: Distances between PDF central values as a function of the scale Q^2 of the PDFs. We show the results for the quark singlet PDF (left plot) and for the gluon PDF (right plot).

band. Also, HERAPDF1.0 NNLO fits lead to a worse fit quality (by more than 50 units in χ^2 for about 500 data points) than the corresponding NLO results. Similar conclusions are found in the updated HERAPDF1.5 NLO and NNLO analysis [41]. There a larger PDF uncertainty on the gluon distribution is found at NNLO as compared to NLO.

4. rcBK analysis with kinematical cuts

We now present the rcBK results with various kinematical cuts as described in Sect. 2, and compare them with the DGLAP cut fits presented in Sect. 3. Fig. 4 shows the comparison of the theoretical results stemming from fits to data with different x -cuts ($x_{\text{cut}} = 10^{-2}$, $3 \cdot 10^{-3}$, 10^{-3} , $3 \cdot 10^{-4}$ and 10^{-4}) to experimental data on the reduced cross section for $x < 10^{-2}$ and $Q^2 < 50 \text{ GeV}^2$. Several comments are in order. First, fits to data with different cuts yield comparably good $\chi^2/\text{d.o.f.}$ ⁶, despite the decreasing number of points with decreasing x_{cut} . Also, the extrapolations of the results for the reduced cross section from fits with cuts to the unfitted region, i.e. to $x > x_{\text{cut}}$, yield a good description of the data. This illustrates the stability of the AAMQS fits under changes in the boundary conditions and lends support to the idea that the non-linear small- x dynamics comprised in the running coupling BK equation describes well the scale dependence of the proton structure in this test region of moderate values of $x \lesssim 0.01$ and $Q^2 \lesssim 50 \text{ GeV}^2$. This result is in contrast with the DGLAP cut fits presented in the previous section, where systematic deviations among fits with and without cuts were observed.

Fig. 5 shows our results corresponding to the rcBK fit with the most stringent cut $x_{\text{cut}} = 10^{-4}$ together with experimental data and the analogous results from the DGLAP fit with cut $A_{\text{cut}} = 1.5$. While the DGLAP extrapolations to the unfitted, *test* region are compatible with data within the uncertainty bands, the central values of the predictions show significant

⁶In the AAMQS approach the χ^2 is calculated as $\chi^2 = \sum_i \frac{(\sigma_{r,\text{th}} - \sigma_{r,\text{exp}})^2}{\Delta\sigma_{r,\text{exp}}^2}$, where $\Delta\sigma_{r,\text{exp}}^2$ is the total experimental error obtained adding in quadrature all experimental uncertainties, and thus neglects the effects of correlations.

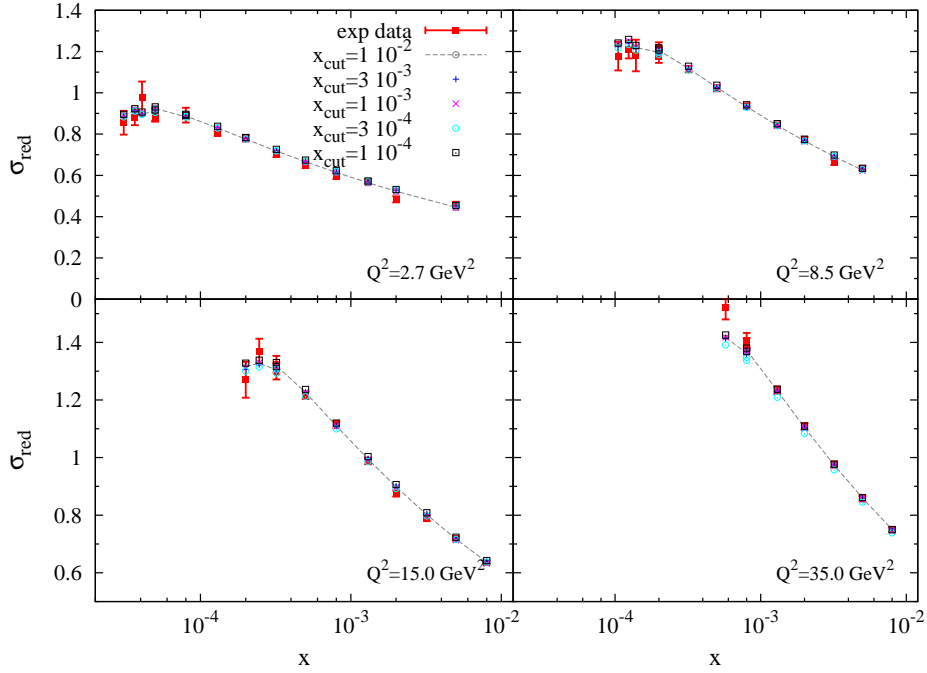


Figure 4: Comparison of the result for reduced cross section obtained with rcBK fits with different cuts: $x_{\text{cut}} = 10^{-2}$, $3 \cdot 10^{-3}$, 10^{-3} , $3 \cdot 10^{-4}$ and 10^{-4} and HERA data for four different bins in $Q^2 = 2.7, 8.5, 15$ and 35 GeV^2 .

deviations from data in the region of small- x . Also, the PDF error bands blow up at small- x . This can be understood as a consequence of the fact that there is no data in the DGLAP fitting set causally connected to small values of x or, equivalently, the DGLAP predictions at small values of x are not empirically constrained by data. This explains the large error bars in this region.

In order to measure these deviations, we plot in Fig. 6 the relative distance between the theoretical results and experimental data both for the rcBK and DGLAP cut fits, the cut values being $x_{\text{cut}} = 10^{-4}$ and $A_{\text{cut}} = 1.5$ respectively. The relative distance is defined as

$$d_{\text{rel}}(x, Q^2) = \frac{\sigma_{\text{r,th}} - \sigma_{\text{r,exp}}}{(\sigma_{\text{r,th}} + \sigma_{\text{r,exp}})/2}, \quad (6)$$

where $\sigma_{\text{r,th}}$ are the theoretical predictions for the HERA reduced cross sections and $\sigma_{\text{r,exp}}$ the corresponding experimental data. As shown in Fig. 6, $d_{\text{rel}}(x, Q^2)$ is on average much smaller for the rcBK fits than it is for the DGLAP one, the latter also showing a systematic trend to underestimate data at small- x and to overshoot them at larger x . In turn, the rcBK values for $d_{\text{rel}}(x, Q^2)$ alternate in sign in all the unfitted region.

Measurements of the deviations of theoretical predictions from data can be endowed with a more statistically meaningful measure, the statistical distance defined as

$$d_{\text{stat}}(x, Q^2) = \frac{\sigma_{\text{r,th}} - \sigma_{\text{r,exp}}}{\sqrt{\Delta\sigma_{\text{r,th}}^2 + \Delta\sigma_{\text{r,exp}}^2}}. \quad (7)$$

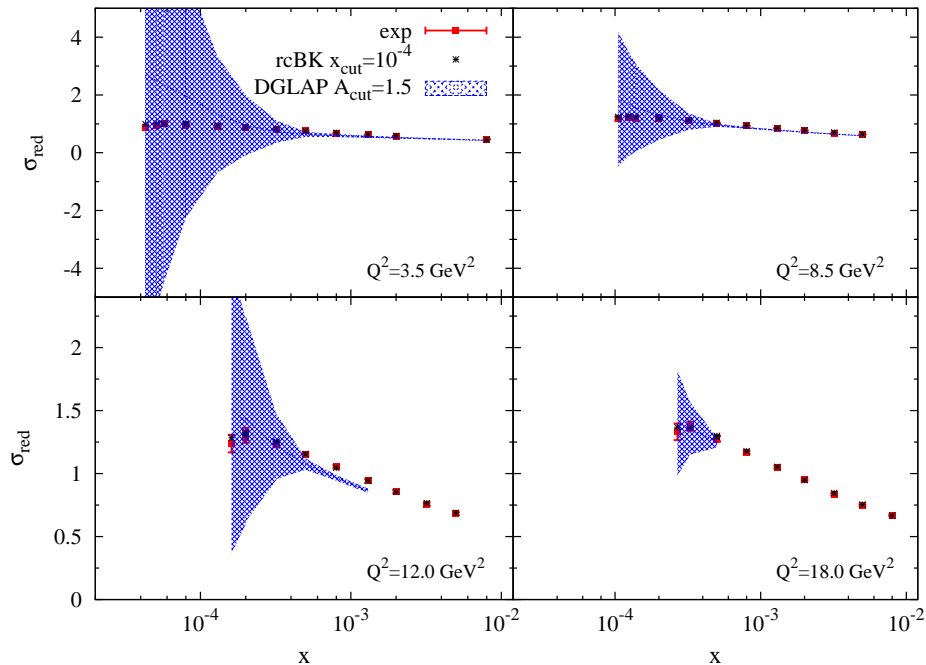


Figure 5: Reduced cross section obtained with the rcBK cut fit with $x_{\text{cut}} = 10^{-4}$ and the DGLAP fit with $A_{\text{cut}} = 1.5$, compared to the experimental HERA-I data. The comparison is shown in four different bins in $Q^2 = 3.5, 8.5, 12$ and 18 GeV^2 . In the DGLAP case the band corresponds to the PDF uncertainties.

While the relative distance Eq. (6) measures in absolute terms the deviation, Eq. (7) gives the statistical significance of that deviation in units of the standard deviation. The values of $d_{\text{stat}}(x, Q^2)$ for the rcBK fit and the DGLAP ones are shown in Fig. 7. In the case of the rcBK fits the theoretical error has been estimated as the maximal difference among the theoretical predictions corresponding to fits with different cuts. The average distances for the rcBK fits with the most stringent cut, $x_{\text{cut}} = 10^{-4}$, are $\langle d_{\text{rel}} \rangle = (5 \pm 41) \cdot 10^{-3}$, and $\langle d_{\text{stat}} \rangle = 0.3 \pm 9$. For the DGLAP cut fit, the average relative distance in the cut region is $\langle d_{\text{rel}} \rangle = 0.1 \pm 0.3$, while the statistical distance in the same region is $\langle d_{\text{stat}} \rangle = -0.8 \pm 1.1$. Both $\langle d_{\text{rel}} \rangle$ and $\langle d_{\text{stat}} \rangle$ are considerably smaller for rcBK than they are for DGLAP, despite the fact that theoretical errors are probably underestimated in the rcBK approach. Note however that the initial conditions in the rcBK analysis are more restrictive than in the DGLAP fit, and that adopting a more flexible input in the rcBK might affect the above results.

In order to explore the predictive power of the rcBK approach and the sensitivity to boundary effects encoded in the different initial conditions for the evolution under the inclusion/exclusion of subsets of data we extrapolate our results for the total $F_2(x, Q^2)$ and longitudinal $F_L(x, Q^2)$ structure functions to values of x smaller than those currently available experimentally. The results of such extrapolation are presented in Figs. 8 and 9. We find that the predictions stemming from different fits converge, within approximately one percent accuracy, at values of $x \sim 10^{-4}$. The fit which extrapolation deviates the most, $\sim 2 \div 3\%$, is the one corresponding to $x_{\text{cut}} = 3 \cdot 10^{-4}$, which also yields a larger $\chi^2/\text{d.o.f.}$ to all data.

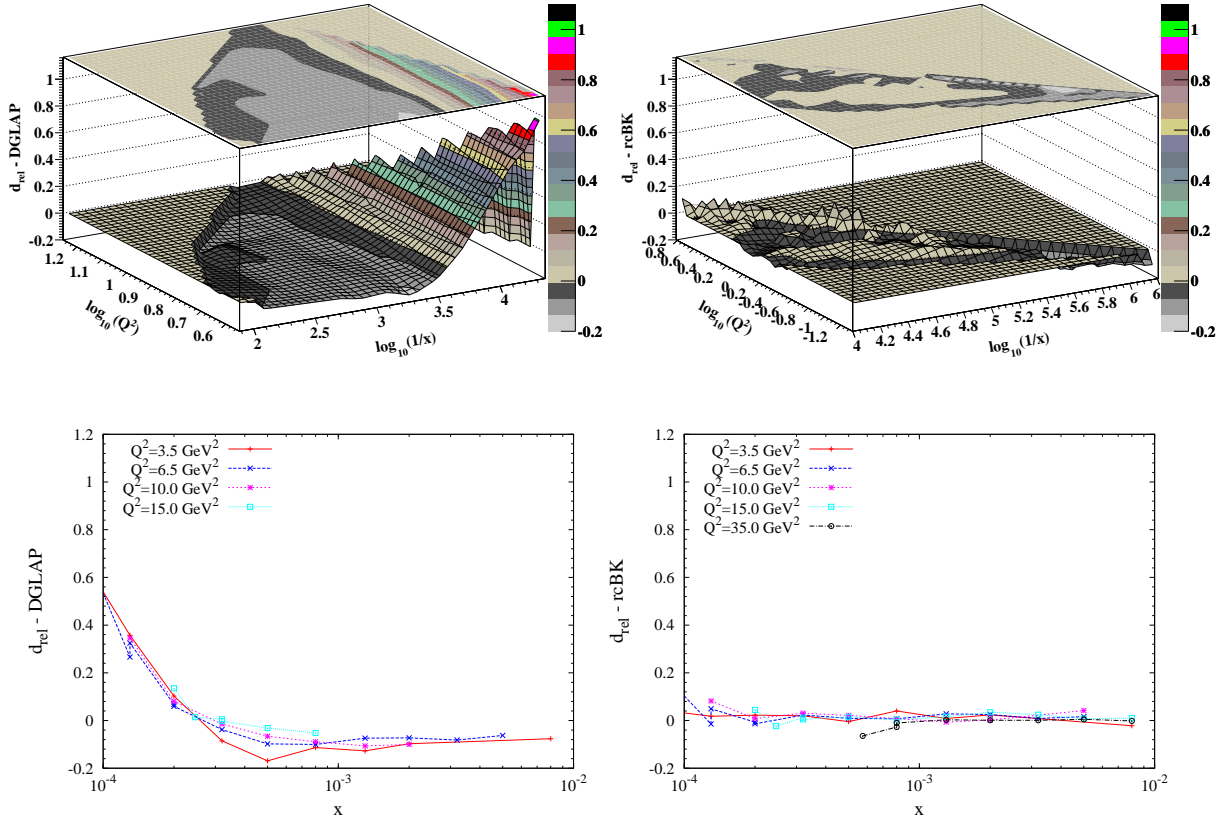


Figure 6: The relative distance, $d_{\text{rel}}(x, Q^2)$ Eq. (6), for DGLAP (left) and rcBK (right) cut fits.

This convergence is due to the fact that rcBK admit asymptotic solutions independent of initial conditions, as already known in the literature. Importantly, these predictions could be verified in planned facilities as the Large electron-Hadron Collider (LHeC) [42] or the Electron Ion Collider [43], where a much extended kinematic reach in x would be available⁷.

5. Implications for LHC phenomenology

It follows from the previous discussion that if deviations from fixed order DGLAP were conclusively found in HERA data, resulting from either the presence of small- x resummation or non-linear effects, one should either exclude this data from the DGLAP analysis or perform a new analysis combining for instance fixed order DGLAP with small- x resummation.⁸ To quantify the worst-case scenario, in this final section we estimate the theoretical uncertainty stemming from these potential deviations in DGLAP fits. To do so we proceed along the lines of Refs. [29, 30] and compute benchmark LHC cross sections with the PDF sets both with and without the small- x kinematical cuts, that is using the same settings as in Ref. [33].

⁷In the context of future DIS facilities, the physics that can be probed at low x within the DGLAP framework has been studied in [44, 45].

⁸Of course such deviations from fixed order DGLAP might also contaminate extractions of the strong coupling constant from global PDF analyses, in which the HERA data play an important role [46].

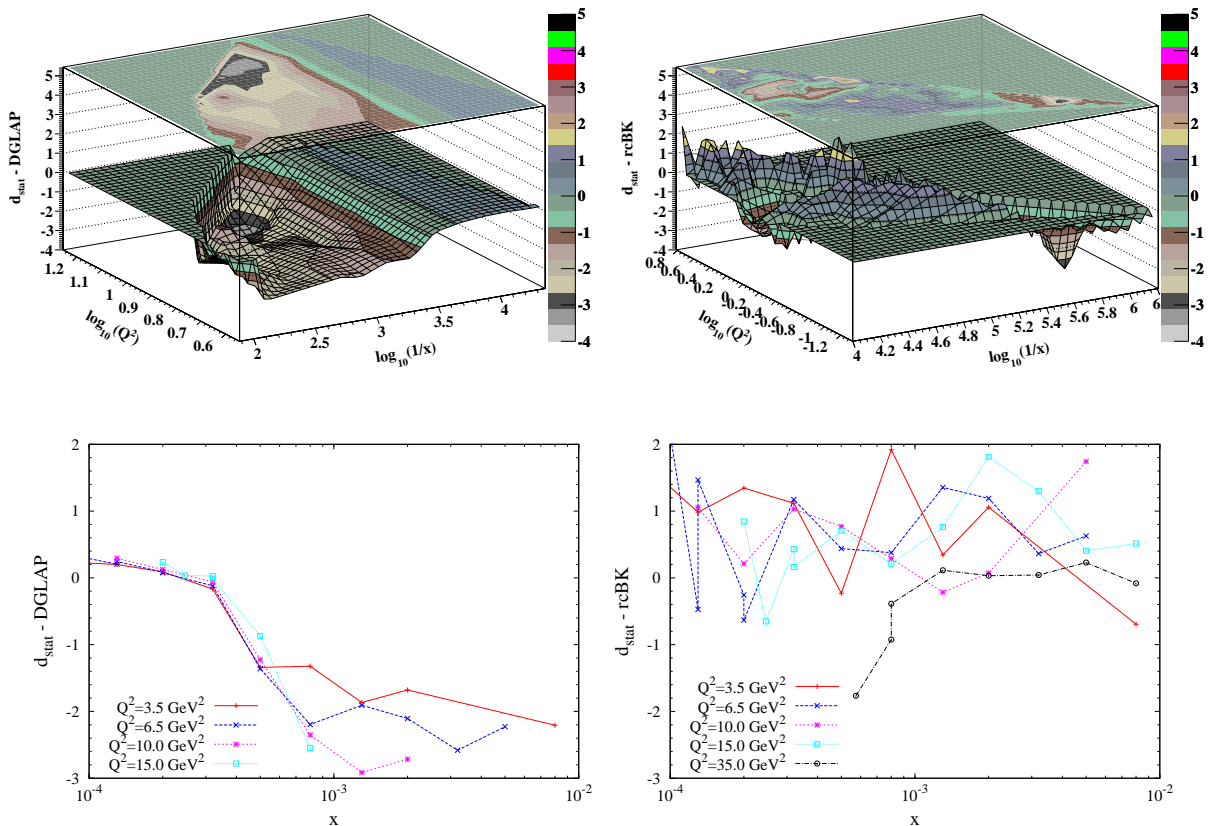


Figure 7: The statistical distance, $d_{\text{stat}}(x, Q^2)$ Eq. (7), for DGLAP (left) and rcBK (right) cut fits.

We have computed, with the NNPDF2.1 NNLO set, the cross sections in NNLO QCD for electroweak gauge boson production with VRAP [47], top quark production with HATHOR [48] (based on an approximate NNLO calculation) and Higgs production in gluon fusion with the code of Ref. [49]. We consider only PDF uncertainties, the strong coupling is kept fixed at its reference value of $\alpha_s(M_Z) = 0.119$ and the renormalization and factorization scales are not varied.

The results are summarized in Table 1 and represented graphically in Fig. 10. As can be seen, the impact of cutting the small- x and small- Q^2 HERA data from the fit is rather moderate at LHC 7 TeV: the uncut results are always within the reference 1-sigma band and the PDF uncertainties increase less than a factor two. On the other hand, the impact is much larger at LHC 14 TeV, since the larger center of mass energy leads to smaller values of x in the PDFs being probed, and these are more affected by the kinematical cut. Indeed, for the electroweak boson production cross sections the PDF uncertainties increase by up to a factor five, though the central predictions are still in agreement with the original uncut results. On the other hand, the cross section for Higgs boson production in gluon fusion is very stable against the kinematical cuts.

These results suggest that understanding what is the correct dynamics that drives the small- x and Q^2 HERA data is important for precision physics at the LHC specially once the machine energy is raised up to its design value of 14 TeV.

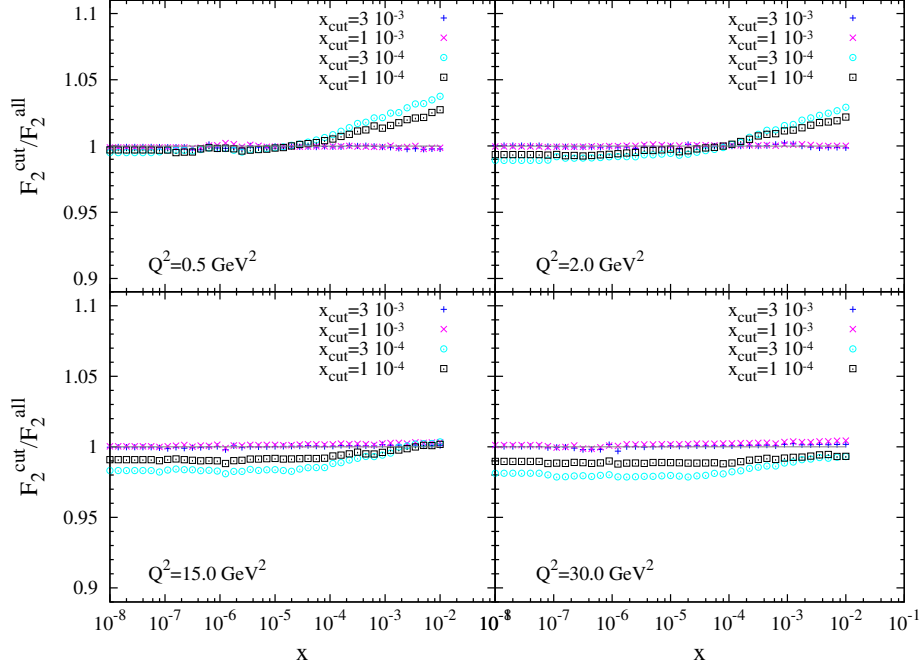


Figure 8: Extrapolation to the low- x region from the rcBK cut fits presented in Fig. 4. The total structure function, $F_2(x, Q^2)$, is calculated down to $x = 10^{-8}$. The results are presented as a ratio of the prediction for the different cut fits to the prediction for the uncut fit, i.e. a fit to all data with $x < x_0 = 10^{-2}$ and $Q^2 < 50 \text{ GeV}^2$.

LHC 7 TeV					
	$\sigma(W^+)B_{l\nu}$ (nb)	$\sigma(W^-)B_{l\nu}$ (nb)	$\sigma(Z^0)B_{ll}$ (nb)	$\sigma(tt)$ (pb)	$\sigma(ggH)$ (pb)
NNPDF2.1 $A_{\text{cut}} = 0$	6.20 ± 0.10	4.21 ± 0.07	0.972 ± 0.013	167 ± 5	13.3 ± 0.3
NNPDF2.1 $A_{\text{cut}} = 1.5$	6.13 ± 0.17	4.17 ± 0.10	0.962 ± 0.021	171 ± 7	13.3 ± 0.3
LHC 14 TeV					
	$\sigma(W^+)B_{l\nu}$ (nb)	$\sigma(W^-)B_{l\nu}$ (nb)	$\sigma(Z^0)B_{ll}$ (nb)	$\sigma(tt)$ (pb)	$\sigma(ggH)$ (pb)
NNPDF2.1 $A_{\text{cut}} = 0$	12.45 ± 0.22	9.14 ± 0.15	2.08 ± 0.03	935 ± 17	44.1 ± 0.5
NNPDF2.1 $A_{\text{cut}} = 1.5$	12.69 ± 1.19	9.24 ± 0.50	2.10 ± 0.11	920 ± 30	44.0 ± 0.7

Table 1: Predictions for LHC cross sections, computed at NNLO QCD with two different PDF sets: the reference NNPDF2.1 NNLO fit with $A_{\text{cut}} = 0$ and with the NNPDF2.1 NNLO fit with $A_{\text{cut}} = 1.5$. Upper table: LHC 7 TeV. Lower table: LHC 14 TeV. For Higgs boson production a mass of $m_H = 125 \text{ GeV}$ has been used, while for top quark production we have used $m_t = 172.5 \text{ GeV}$.

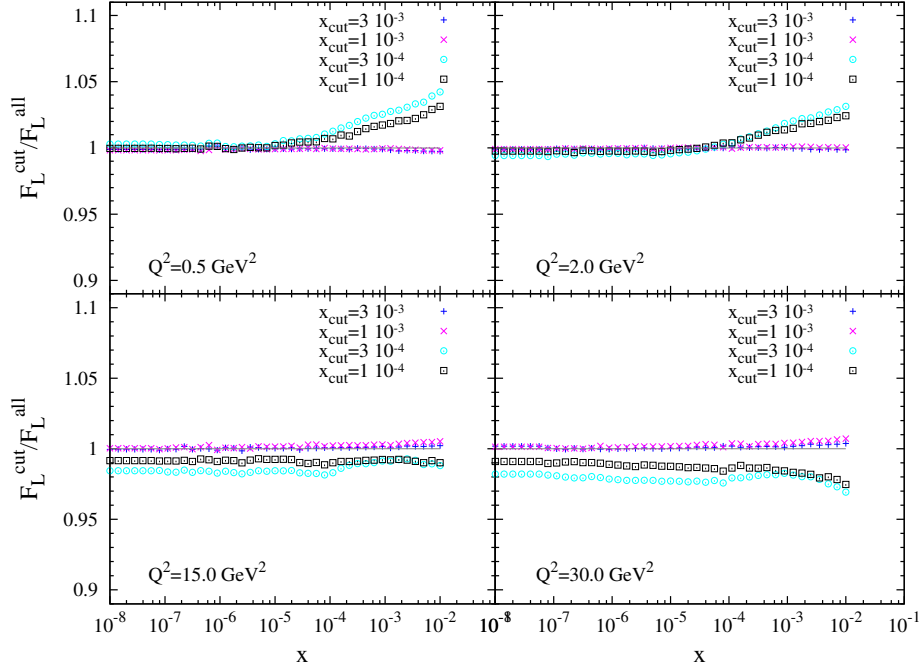


Figure 9: Low- x extrapolation analogous to Fig. 8 for the longitudinal structure function $F_L(x, Q^2)$.

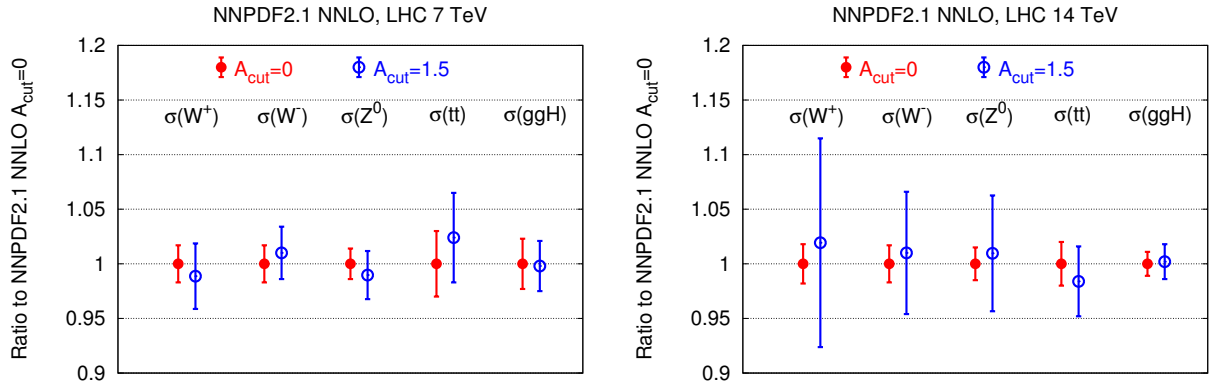


Figure 10: Graphical summary of the results of Table 1: comparison of the predictions for LHC NNLO cross sections for the reference NNPDF2.1 NNLO fit with $A_{\text{cut}} = 0$ and with the NNPDF2.1 NNLO fit with $A_{\text{cut}} = 1.5$. Cross sections are shown as ratios to the uncut $A_{\text{cut}} = 0$ predictions. We show results both for LHC 7 TeV (left plot) and for LHC 14 TeV (right plot).

6. Conclusions and outlook

In summary, we have presented a precision study on the suitability of the rcBK and DGLAP approaches to describe HERA data in the kinematic region of moderate x and Q^2 . Our strategy consists in setting a common *test* ground through selected kinematic cuts to the rcBK and DGLAP fitting procedures and perform systematic comparisons between theory extrapolations and the unfitted data set. Our main findings are:

- i) DGLAP fits display sensitivity to the exclusion of small- x data from the fitted sample. While the deviations found are not statistically significant (below the 1-sigma level), this might give support to the general idea that novel interesting physics is being obscured by its encoding in the freedom of choice of initial conditions.
- ii) rcBK fits are robust against the exclusion of data for values of x down to $x = 10^{-4}$ and for data with $Q^2 < 50 \text{ GeV}^2$.
- iii) The exclusion of small- x data from DGLAP fits has consequences for LHC phenomenology, in particular in a significant increase on the theoretical uncertainty for standard production cross sections at a collision energy of 14 TeV.
- iv) rcBK evolution yields robust predictions at small x and thus can be decisively confronted with data from proposed facilities like the LHeC or the EIC. Within a DGLAP analysis uncertainties grow very fast for low x outside the data region, and that renders the framework non predictive in that domain.

The results of this paper beg for the development of a framework in which both Q^2 (DGLAP) evolution and x (rcBK) non-linear evolution are jointly accounted for, just as the common framework for the DGLAP and BFKL equations has been consistently obtained by the small- x resummation formalism. Ultimately, one would wish for a resummation of the non-linear kernel into DGLAP or alternatively to the encoding of Q^2 evolution into the B-JIMWLK framework. Unfortunately, at present no clear pathway in this direction exists.

At a more pragmatic level, the non-linear corrections implied by rcBK evolution could be encoded in an effective manner as a correction factor or further constrain to the results obtained within DGLAP. For instance, the paucity of data at small- x could be made up for by including rcBK generated pseudodata in DGLAP fits. Also, a thorough study of the impact of the small- x resummed formalism in DGLAP fits would be very useful to delimit to what extent non-linear corrections embodied in rcBK evolution are actually necessary to describe present data.

Acknowledgments

The research of J.L.A. is supported by a fellowship from the Théorie LHC France initiative funded by the IN2P3. J. R. is grateful to S. Forte for discussions. The research of J. R. has been supported by a Marie Curie Intra-European Fellowship of the European Community's 7th Framework Programme under contract number PIEF-GA-2010-272515. JGM acknowledge the support of Fundação para a Ciência e a Tecnologia (Portugal) under project CERN/FP/116379/2010. The work of PQA is funded by the French ANR under contract ANR-09-BLAN-0060.

References

- [1] Y.L. Dokshitzer, Sov. Phys. JETP 46 (1977) 641.
- [2] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. 15 (1972) 438.
- [3] G. Altarelli and G. Parisi, Nucl. Phys. B126 (1977) 298.
- [4] E.A. Kuraev, L.N. Lipatov and V.S. Fadin, Sov. Phys. JETP 45 (1977) 199.
- [5] I. Balitsky and L. Lipatov, Sov.J.Nucl.Phys. 28 (1978) 822.
- [6] M. Ciafaloni et al., Phys. Rev. D68 (2003) 114003, hep-ph/0307188.
- [7] G. Altarelli, R.D. Ball and S. Forte, Nucl. Phys. B799 (2008) 199, arXiv:0802.0032.
- [8] S. Forte, G. Altarelli and R.D. Ball, Nucl.Phys.Proc.Suppl. 191 (2009) 64, arXiv:0901.1294.
- [9] C. White and R. Thorne, Phys.Rev. D75 (2007) 034005, hep-ph/0611204.
- [10] I. Balitsky, Nucl. Phys. B463 (1996) 99, hep-ph/9509348.
- [11] Y.V. Kovchegov, Phys. Rev. D60 (1999) 034008, hep-ph/9901281.
- [12] J. Jalilian-Marian et al., Phys. Rev. D59 (1998) 014014, hep-ph/9706377.
- [13] A. Kovner, J.G. Milhano and H. Weigert, Phys. Rev. D62 (2000) 114005, hep-ph/0004014.
- [14] H. Weigert, Nucl. Phys. A703 (2002) 823, hep-ph/0004044.
- [15] A.M. Stasto, K.J. Golec-Biernat and J. Kwiecinski, Phys. Rev. Lett. 86 (2001) 596, hep-ph/0007192.
- [16] E. Gardi et al., Nucl. Phys. A784 (2007) 282, hep-ph/0609087.
- [17] Y. Kovchegov and H. Weigert, Nucl. Phys. A 784 (2007) 188, hep-ph/0609090.
- [18] I.I. Balitsky, Phys. Rev. D 75 (2007) 014001, hep-ph/0609105.
- [19] I. Balitsky and G.A. Chirilli, Phys. Rev. D77 (2008) 014019, arXiv:0710.4330.
- [20] J.L. Albacete and Y.V. Kovchegov, Phys. Rev. D75 (2007) 125021, arXiv:0704.0612 [hep-ph].
- [21] J.L. Albacete et al., Phys. Rev. D80 (2009) 034031, arXiv:0902.1112.
- [22] J.L. Albacete et al., Eur.Phys.J. C71 (2011) 1705, arXiv:1012.4408.
- [23] H1, F.D. Aaron et al., JHEP 01 (2010) 109, arXiv:0911.0884.

- [24] J. Kuokkanen, K. Rummukainen and H. Weigert, Nucl.Phys. A875 (2012) 29, arXiv:1108.1867.
- [25] F. Caola and S. Forte, Phys.Rev.Lett. 101 (2008) 022001, arXiv:0802.1878.
- [26] S. Forte and R.D. Ball, Acta Phys.Polon. B26 (1995) 2097, hep-ph/9512208.
- [27] T. Lastovicka, Eur.Phys.J. C24 (2002) 529, hep-ph/0203260.
- [28] The NNPDF collaboration, L. Del Debbio et al., JHEP 03 (2005) 080, hep-ph/0501067.
- [29] F. Caola, S. Forte and J. Rojo, Phys. Lett. B686 (2010) 127, arXiv:0910.3143.
- [30] F. Caola, S. Forte and J. Rojo, Nucl.Phys. A854 (2011) 32, arXiv:1007.5405.
- [31] S. Moch, J. Vermaseren and A. Vogt, Nucl.Phys. B688 (2004) 101.
- [32] S. Moch, J. Vermaseren and A. Vogt, Phys. Lett. B691 (2004) 129.
- [33] The NNPDF Collaboration, R.D. Ball et al., Nucl.Phys. B855 (2012) 153, arXiv:1107.2652.
- [34] S. Forte et al., Nucl. Phys. B834 (2010) 116, arXiv:1001.2312.
- [35] The NNPDF Collaboration, R.D. Ball et al., Nucl. Phys. B809 (2009) 1, arXiv:0808.1231.
- [36] The NNPDF collaboration, R.D. Ball et al., Nucl. Phys. B823 (2009) 195, arXiv:0906.1958.
- [37] The NNPDF collaboration, R.D. Ball et al., Nucl. Phys. B838 (2010) 136, arXiv:1002.4407.
- [38] The NNPDF collaboration, R.D. Ball et al., JHEP 05 (2010) 075, arXiv:0912.2276.
- [39] H.L. Lai et al., Phys. Rev. D82 (2010) 074024, 1007.2241.
- [40] H1 and ZEUS Collaborations, (2010), H1prelim-10-044, ZEUS-prel-10-008.
- [41] H1 and ZEUS Collaborations, (2011), H1prelim-11-042, ZEUS-prel-11-00.
- [42] M. Klein *et. al.*, EPAC'08, 11th European Particle Accelerator Conference (2008).
- [43] *The Electron Ion Collider: A white paper*, BNL Report BNL-68933-02/07-REV, Eds. A. Deshpande, R. Milner and R. Venugopalan. .
- [44] A. Accardi, V. Guzey and J. Rojo, (2011), arXiv:1106.3839.
- [45] J. Rojo and F. Caola, (2009), arXiv:0906.2079.
- [46] R.D. Ball et al., Phys.Lett. B707 (2012) 66, arXiv:1110.2483.

- [47] C. Anastasiou et al., Phys. Rev. D69 (2004) 094008, hep-ph/0312266.
- [48] M. Aliev et al., Comput. Phys. Commun. 182 (2011) 1034, 1007.1327.
- [49] R. Bonciani, G. Degrossi and A. Vicini, JHEP 11 (2007) 095, arXiv:0709.4227.