Overview of proton PDFs
and small-x resummation

Marco Bonvini

INFN, Rome 1 unit

Initial Stages 2019, Columbia University, NY, June 24, 2019
Theoretical predictions with hadrons in the initial state

Collinear factorization theorem in QCD:

\[
\frac{d\sigma}{dQ^2 dY dp_t...} \sim \sum_{i,j=g,q} \int_0^1 dx_1 \, dx_2 \, f_i(x_1, Q^2) \, f_j(x_2, Q^2) \, C_{ij} \left( \frac{Q^2}{x_1 x_2 s}, y, p_t, ..., \alpha_s(Q^2) \right)
\]

- partonic cross sections \( C_{ij}(x, ..., \alpha_s) \) (observable-dependent, perturbative)
- parton distribution functions (PDFs) \( f_i(x, Q^2) \) (universal, non-perturbative)
Scale dependence of the PDFs

DGLAP evolution:

$$\mu^2 \frac{d}{d\mu^2} f_i(x, \mu^2) = \sum_{j=g,q} \int_x^1 \frac{dz}{z} P_{ij}(z, \alpha_s(\mu^2)) f_j \left( \frac{x}{z}, \mu^2 \right)$$

- splitting functions $P_{ij}(x, \alpha_s)$ (universal, perturbative)

PDFs at a given scale $\mu_0 +$ DGLAP evolution $\rightarrow$ PDFs at any scale $\mu$
PDF determination from first principles

Non-perturbative problem $\rightarrow$ numerical simulations on a discretized spacetime. However, the field-theoretic definition of PDFs involves light-cone distances:

$$f_q(x, \mu^2) = \int \frac{d\xi}{4\pi} e^{-ix\xi P^+} \langle P|\bar{\psi}_q(\xi)\gamma \eta U_n(\xi, 0)\psi_q(0)|P\rangle$$

$$\xi^2 = 0$$

but in lattice QCD simulation the spacetime is Euclidean, where light-cone separations are only trivial ($\xi = 0$) $\rightarrow$ PDFs cannot be computed in lattice QCD!

Some possible alternatives:

- compute properties of PDFs (Mellin moments) on the lattice
- compute on the lattice a different object (quasi-PDFs, pseudo-PDFs) which tends to the light-cone PDFs in some limit (some issues though)
- compute scattering amplitudes on the lattice and extract PDFs
- use non-perturbative techniques in the continuum (using e.g. the Bethe-Salpeter equation)

None of these approaches provides sufficient precision for phenomenology today.
Strategy: fit $f_i(x, \mu_0^2)$ by comparison with (many) data

Such fitted PDFs depend unavoidably on the accuracy on the perturbative ingredients $P_{ij}(x, \alpha_s)$, $C_{ij}(x, ..., \alpha_s)$
Various PDF fitting groups (more on LHAPDF):

- CTEQ (CT and CJ)
- MRST/MSTW/MMHT
- NNPDF
- ABM/ABKM/ABMP
- HERAPDF
- xFitter → ATLAS, CMS, ...
- ....

Differences:

- datasets
- theory inputs/assumptions
- PDF parametrization
- fitting methodology
- ....
Today most PDF sets are based on the following consolidated ingredients:

- **NNLO accuracy in DGLAP evolution and in cross section computation**
- **variable flavour number scheme (VFNS) with correct mass effects**
- **flexible PDF parametrizations (with some caveats)**
- **careful treatment of experimental uncertainties**
- **a large dataset:**
  - **HERA** (inclusive DIS and heavy flavour production)
  - **fixed-target DIS** (BCDMS, NMC, NuTeV, ...)
  - **fixed-target Drell-Yan** (E866, E605)
  - **Tevatron Drell-Yan**
  - **LHC** (mostly ATLAS+CMS):
    - Drell-Yan
    - jets
    - $t\bar{t}$ production
    - $Z$ $p_t$-distribution
    - ....

Kinematic coverage

**Collider Drell-Yan**

**Collider DIS**

**Fixed-Target Drell-Yan**

**Fixed-Target DIS**

**Jets**

**top production**

**Z differential**
Variable flavour number scheme

The number $n_f$ of “active” flavours changes during the evolution (factorization scheme choice to resum large collinear logarithms from heavy quark pair production).

Matching relation between PDFs in schemes with different $n_f$

$$f^{[n_f+1]}_i(\mu^2) = \sum_{j=\text{light}} A_{ij} \left( \frac{m^2}{\mu^2} \right) \otimes f^{[n_f]}_j(\mu^2)$$

$A_{ij}$ = perturbative matching coefficients

PDFs formally independent of the matching scale; perturbative dependence remain

The dependence is reduced by including higher orders in the matching conditions (and in DGLAP)

Marco Bonvini New insights on the proton's structure

Overview of proton PDFs and small-x resummation

8
Matching conditions at the charm threshold

\[ \kappa_c = \frac{\mu_c}{m_c}, \quad \mu_c = \text{charm matching scale (threshold)} \]

The perturbatively generated charm PDF, since the scale is low (thus \( \alpha_s \) is large), is affected by large higher order corrections, somehow probed by \( \mu_c \) variations.
recent developments
Improving the charm PDF

Option 1: use $\mu_c$ dependence to improve agreement with data

![Graph showing $\chi^2/\chi_0^2$ vs $\mu_c$ (GeV) for charm PDFs at NLO and NNLO.](image)

at NNLO $\mu_c \sim 3.5\text{GeV} \sim 2.5m_c$ gives the best agreement, but it’s not perturbatively stable. One should also vary $\mu_c$ to get an uncertainty.

Option 2: directly fit the charm PDF from data

![Graphs showing charm PDF fits at NLO and NNLO for different $m_c$ values.](image)

fitted charm has larger uncertainty but is much less dependent on the charm mass than the perturbatively generated charm.

fitted charm may also contain some “intrinsic” component. See also [CT 1707.00657]
Impact of fitted charm PDF

Moderate effect for LHC phenomenology

[NPDF 1605.06515]

Inclusive Z production @ NLO, LHC 13 TeV

Z+Charm production, LHC 13 TeV

processes strongly influenced by the charm PDF in the initial state (like Z + c) are more sensitive
The proton contains also photons, with probability suppressed by $\alpha_{em}$.

However, for percent precision knowing the photon PDF is important.

Past: fitting it from data along with the other PDFs (large uncertainties)

A breakthrough: the LUXqed approach

$$\sigma = C_\gamma \otimes f_\gamma + \alpha_{em} \sum_q C_q \otimes f_q + \ldots$$

= exact formula in terms of measured structure functions

The cross section for the process $l^+ + ! L$ in the QCD improved parton model.

At this point a comment is in order. We can systematically compute the cross section

Due to differences in quarks –

The cross section for the process $l^+ + ! L$ in the QCD improved parton model.

At this point a comment is in order. We can systematically compute the cross section
Neutral Drell-Yan production: sizeable effect at low mass

Higgs + $W$ production: sizeable effect up to very high $p_t^H$
**Theory uncertainties**

\[ \chi^2 = \sum_{i,j=1}^{N_{\text{data}}} (T_i - D_i)(\text{cov}^{-1})_{ij}(T_j - D_j) \]

the covariance matrix \( \text{cov} \) usually contains ONLY experimental uncertainties

Theoretical predictions are perturbative \( \rightarrow \) **uncertainty from missing higher orders**

Usually estimated through (unphysical) scale variation: \( \mu_R, \mu_F \)

Include such theory uncertainties in the covariance matrix  

[NNPDF 1905.04311]

![Experimental Correlation Matrix](image1)

![Experimental + Theory Correlation Matrix (9 pt)](image2)
Including theory uncertainties in the fit leads to (slightly) larger PDF uncertainties.

The fitted PDF at next order is contained within the band: good!

However, varying again the scale when making a prediction using these PDFs can lead to a double counting of the scale variation effect [Harland-Lang, Thorne 1811.08434]
small-\(x\) resummation
Low $x$ at HERA: importance of resummation in PDF fits

Deep-inelastic scattering (DIS) data from HERA extend down to $x \sim 3 \times 10^{-5}$

Tension between HERA data at low $Q^2$ and low $x$ with fixed-order theory

Also leads to a deterioration of the $\chi^2$ when including low-$Q^2$ data

Attempts to explain this deviation with higher twists, phenomenological models, ...

Successful description of this region when including small-$x$ resummation!

[Ball,Bertone,MB,Marzani,Rojo,Rottoli 1710.05935] [xFitter+MB 1802.00064] [MB,Giuli 1902.11125]
Logarithmic enhancements → all-order resummation

Structure of logarithmically enhanced contributions

\[ \text{pert. coeff. } (P, A, C) = a_0 L + \alpha_s [a_1 L + b_1] + \alpha_s^2 [a_2 L^2 + b_2 L + c_2] + \alpha_s^3 [a_3 L^3 + b_3 L^2 + c_3 L + d_3] + \alpha_s^4 [a_4 L^4 + b_4 L^3 + c_4 L^2 + d_4 L + e_4] + \ldots \]

If/when \( \alpha_s L \sim 1 \) the fixed-order expansion is no longer predictive!

Resum the logs, and convert to a “logarithmic-order” expansion:

\[ g_{\text{LL}}(\alpha_s L) + \alpha_s g_{\text{NLL}}(\alpha_s L) + \alpha_s^2 g_{\text{NNLL}}(\alpha_s L) + \ldots \]

Leading log (LLx), next-to-leading log (NLLx), next-to-next-to-leading log (NNLLx)...

Small-x resummation formalism developed in the 90s-00s
Known at LLx for partonic cross sections and NLLx for DGLAP evolution

[Catani,Ciafaloni,Colferai,Hautmann,Salam,Stasto] [Thorne,White] [Altarelli,Ball,Forte]
$P_{gg}(x, \alpha_s)$ splitting function at fixed order

Logarithms start to grow for $x \lesssim 10^{-2} \rightarrow \text{perturbative instability}$
Example: small-$x$ logarithms in gluon-gluon splitting function

\[ P_{gg}(x, \alpha_s) \] splitting function at fixed order

Logarithms start to grow for $x \lesssim 10^{-2} \rightarrow \text{perturbative instability}$

Resummation obtained with the HELL public code

[MB,Marzani,Peraro 1607.02153] [MB,Marzani,Muselli 1708.07510] [MB,Marzani 1805.06460]
Another example: matching conditions at the charm threshold

\[ \kappa_c = \frac{\mu_c}{m_c}, \quad \mu_c = \text{charm matching scale (threshold)} \]

The perturbatively generated charm PDF is much less dependent on the (unphysical) matching scale when small-\(x\) resummation is included!
The first two PDF fits with small-\(x\) resummation

**HELL** → makes possible a PDF fit with small-\(x\) resummation

<table>
<thead>
<tr>
<th><strong>NNPDF3.1sx [1710.05935]</strong></th>
<th><strong>xFitter [1802.00064]</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>NeuralNet parametrization of PDFs</td>
<td>polynomial parametrization</td>
</tr>
<tr>
<td>MonteCarlo uncertainty</td>
<td>Hessian uncertainty</td>
</tr>
<tr>
<td>charm PDF is fitted</td>
<td>charm PDF perturbatively generated</td>
</tr>
<tr>
<td>DIS+tevatron+LHC ((\sim 4000) datapoints)</td>
<td>only HERA data ((\sim 1200) datapoints)</td>
</tr>
<tr>
<td>NLO, NLO+NLLx, NNLO, NNLO+NLLx</td>
<td>NNLO, NNLO+NLLx</td>
</tr>
</tbody>
</table>

The quality of the fit improves substantially including small-\(x\) resummation

\[
\chi^2/N_{\text{dat}} \quad \text{NNLO} \quad \text{NNLO+NLLx}
\]

| **xFitter** | 1.23 | 1.17 |
| **NNPDF3.1sx** | 1.130 | 1.100 |

**smaller!**

Stable upon inclusion of low-\(x\) data →

![Graph showing \(\chi^2/N_{\text{dat}}\) vs. -log(\(x_{\text{min}}\)) for different PDF fits]

Ball et al 17, xFitter 18
Significantly improved description of the HERA data

**Q^2 = 3.5 GeV^2**

The better description mostly comes from a larger resummed $F_L$

\[
\sigma_{r,NC} = F_2(x, Q^2) - \frac{y^2}{1 + (1 - y)^2} F_L(x, Q^2)
\]

\[
y = \frac{Q^2}{x s}
\]
Significantly improved description of the HERA data

\[ Q^2 = 3.5 \text{ GeV}^2 \]

The better description mostly comes from a larger resummed \( F_L(x, Q^2) \).

\[ \sigma_{r,\text{NC}} \]

The tension between HERA data at low \( x \) and DIS data from HERA extend down to \( x \sim 10^{-3} \).

**Figure 34**: The combined low-\( x \) at HERA: importance of resummation in PDFs

**Marco Bonvini**

-  A new frontier of precision and accuracy at LHC and future colliders
-  Significant improved description of the HERA data
-  The better description mostly comes from a larger resummed \( F_L(x, Q^2) \)
Small-$x$ resummation mostly affects the gluon PDF (and the total quark singlet).

Note: future higher energy colliders will probe smaller values of $x$ → small-$x$ resummation will be even more important in future! \( (x_{\text{min}} \sim Q^2/s) \)
**Impact of small-$x$ resummation at LHC and future colliders**

$gg \rightarrow H$ inclusive cross section:

\[ \text{ggH production cross section} \quad \text{--- \ effect of small-x resummation} \]

\[ \begin{array}{cccccccc}
  \text{f.o. PDFs: NNPDF31sx_nnlo_as_0118} & \text{res PDFs: NNPDF31sx_nnlonllx_as_0118} \\
  \text{N}^3\text{LO}, \text{f.o. PDFs} & \text{N}^3\text{LO}, \text{res PDFs} \\
  \text{N}^3\text{LO}+\text{LLx}, \text{res PDFs} \\
\end{array} \]

\[ m_H = 125 \text{ GeV} \quad \mu_F = \mu_R = m_H/2 \]

\[ \frac{\text{ratio to N}^3\text{LO}}{} \]

\[ \sqrt{s} [\text{TeV}] \]

$ggH$ cross section at FCC-hh $\sim 10\%$ larger than expected!

At LHC $+1\%$ effect; larger effect expected at differential level

Other processes (Drell-Yan, $c\bar{c}$, ...): work in progress (at multi-differential level)
Conclusions

Our study of the structure of the proton keeps progressing and producing very interesting new results

Stay tuned
Backup slides
A digression on the theory of small-\(x\) resummation

Small-\(x\) resummation formalism based on \(k_t\)-factorization and BFKL [Altarelli,Ball,Forte] Developed in the 90s-00s [Catani,Ciafaloni,Colferai,Hautmann,Salam,Stasto] [Thorne,White] Known at LL\(x\) and NLL\(x\) since many years, but very limited number of applications until very recently, because small-\(x\) resummation is a hell!

Recent developments: [MB,Marzani,Peraro 1607.02153][MB,Marzani,Muselli 1708.07510]

- improved ABF [Altarelli,Ball,Forte 1995,...,2008] procedure to resum splitting functions and new formalism for coefficient functions
- all the ingredients for describing DIS process at small \(x\), including mass effects and heavy flavour matching conditions in DGLAP evolution
- match resummation to NNLO, allowing NNLO+NLL\(x\) phenomenology
- we released (and keep developing) a public code HELL: High-Energy Large Logarithms www.ge.infn.it/~bonvini/hell which delivers resummed splitting functions and coefficient functions
- HELL interfaced to APFEL (apfel.hepforge.org) \(\rightarrow\) PDF fits
- matching to \(N^3\)LO also available [MB,Marzani 1805.06460]
- resummation of LHC observables (Higgs in gluon fusion) [MB 1805.08785]
Towards $N^3$LO evolution

Recent impressive progress towards $N^3$LO splitting functions
[Davies, Vogt, Ruijl, Ueda, Vermaseren 1610.07477] [Moch, Ruijl, Ueda, Vermaseren, Vogt 1707.08315]

At small $x$, approximate predictions from NLL$x$ resummation [MB, Marzani 1805.06460]

Large uncertainties from subleading logs

$N^3$LO splitting functions are much more unstable at small $x$ → need resummation!
The role of FCC-eh (and LHeC)

Prediction in the LHeC and FCC-eh kinematic regions for \( F_2 \) and \( F_L \)

Pseudo data show a small errors: significant constraining power!

Fit to pseudo data shows a significantly reduced uncertainty, and a huge effect of small-\( x \) resummation
Higgs production: parton-level results

\[ \sigma(m_H^2, s) = \sigma_0(m_H^2) \sum_{i,j=g,q} \int_\tau^1 \frac{dz}{z} C_{ij}(z, \alpha_s(m_H^2)) L_{ij} \left( \frac{\tau}{z}, m_H^2 \right) \]

\[ \tau = \frac{m_H^2}{s} \]

Marco Bonvini

Overview of proton PDFs and small-x resummation
Higgs production: parton luminosities

\[
\sigma(m_H^2, s) = \sigma_0(m_H^2) \sum_{i,j=g,q} \int_\tau \frac{dz}{z} C_{ij}(z, \alpha_s(m_H^2)) \mathcal{L}_{ij} \left( \frac{\tau}{z}, m_H^2 \right) \tau = \frac{m_H^2}{s}
\]
The large effect of the resummation is due to the NNLO being perturbatively unstable at small $x$. 

Marco Bonvini
Overview of proton PDFs and small-$x$ resummation
35
New parametrization in xFitter

Default xFitter parametrization \( x f(x) = A x^B (1 - x)^C \left[ 1 + D x + E x^2 \right] \)
Not flexible enough at small \( x \)! May lead to bias.

Newly proposed parametrization \[ MB, Giuli 1902.11125 \]
\( x f(x) = A x^B (1 - x)^C \left[ 1 + D x + E x^2 + F \log x + G \log^2 x + H \log^3 x \right] \)
much more flexible at small \( x \)!

Reduction of the \( \chi^2 \), both at fixed order and with resummation
Slightly different PDF shapes, well compatible with the (more flexible) NNPDF fit