Global Fits of PDFs

Tommaso Giani

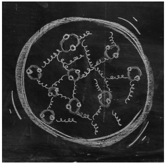
Wednesday, October 21

QCD Real-Time Dynamics and Inverse Problems

Proton structure and PDFs

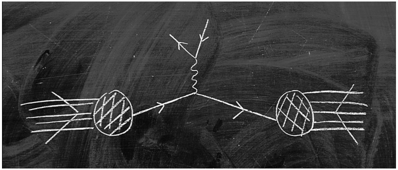
Proton energy divided among its constituents: quarks and gluons.
Proton structure encoded in Parton Distribution Functions (PDFs):

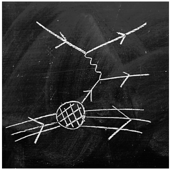
> *probability distributions of sampling a parton (quarks, gluon) with given momentum from the proton.*



$$f_q\left(x, Q^2\right) \qquad x = p_{\text{parton}}/p_{\text{proton}}$$

$Q =$ characteristic scale of the process





- important input to make predictions and computations in collider physics
- dominant source of uncertainty in many analyses: determination of standard model parameters, Higgs boson characterisation and searches for New Physics
- non perturbative and universal objects, extracted from experimental data

→ https://lhapdf.hepforge.org/

Overview

1. PDFs from experimental data: the NNPDF framework
2. PDFs from the lattice
3. Summary and ideas for future work

PDFs from experimental data: the NNPDF framework

PDFs from global QCD analysis

1. Factorization theorem relating any sufficiently inclusive observable to the collinear PDFs

$$\mathcal{O}_I\left(x, Q^2\right) = \sum_i C_i^I\left(x, \alpha_s\left(Q^2\right)\right) \otimes f_i\left(x, Q^2\right) + \text{power corrections}$$

   with

$$C_{I,i} \otimes f_i \equiv \int_x^1 \frac{dy}{y} C_i^I\left(\frac{x}{y}\right) f_i\left(y\right)$$

2. Perturbative input: coefficient functions and DGLAP evolution

$$C_{I,i} = \sum_{k=0} C_{I,i}^{(k)} \alpha_s^k$$

$$f_i\left(x, Q^2\right) = \sum_j \Gamma_{ij}\left(Q^2, Q_0^2, \alpha_s\right) \otimes f_j\left(y, Q_0^2\right)$$

   with

$$\Gamma_{ij} = \sum_{k=0} \Gamma_{ij}^{(k)} \alpha_s^{k+1}$$

3. Experimental data

PDF parametrization
$f\left(x, Q_0^2\right)$

evolve with DGLAP
to the data scale Q^2
$f_i\left(x, Q^2\right) = \Gamma_{ij} \otimes f_j\left(x, Q_0^2\right)$

compute the cross section
$\sigma_{\text{theo}} = f_i \otimes C_i$

compare to data
$\sigma_{\text{theo}} - \sigma_{\text{data}}$

update model to minimize χ^2

PDFs set

8 independent neural net parametrizations for:

g

$\Sigma = \left(u + \bar{u} + d + \bar{d} + s + \bar{s}\right)$

$T_3 = \left(u + \bar{u} - d - \bar{d}\right)$

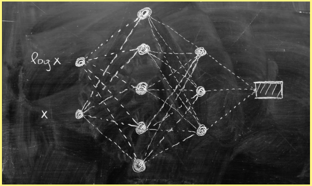
$T_8 = \left(u + \bar{u} + d + \bar{d} + -2s - 2\bar{s}\right)$

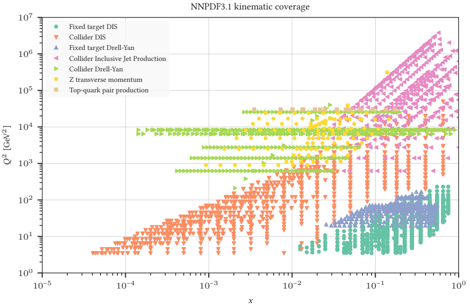
$V = \left(u - \bar{u} + d - \bar{d} + s - \bar{s}\right)$

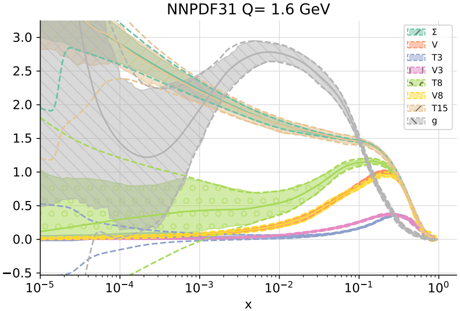
$V_3 = \left(u - \bar{u} - d + \bar{d}\right)$

$V_8 = \left(u - \bar{u} + d - \bar{d} - 2s + 2\bar{s}\right)$

$T_{15} = (u + \bar{u} + d + \bar{d} + s + \bar{s} - 3c - 3\bar{c})$







NNPDF collaboration, "Parton distributions from high-precision collider data", Eur. Phys. J. C77 (2017)

Towards NNPDF4.0

- new experimental data (for example dijet cross sections)
- based on improved machine learning algorithms
- NNLO QCD and NLO EWK corrections
- new theoretical improvements, like PDFs positivity, integrability...
- ...

Work of many people! $\rightarrow$

http://nnpdf.mi.infn.it/

PDFs from the lattice

What about lattice data?

As before we need 3 things:

1. Factorization theorem relating some lattice observable to the collinear PDFs

$$\mathcal{O}_I\left(z, Q^2\right) = \sum_i \mathcal{C}_i^I\left(z, \alpha_s\left(Q^2\right)\right) \circledast f_i\left(Q^2\right) + \text{power corrections}$$

   where the specific expression of $\mathcal{C}_i^I$ and the definition of the operation $\circledast$ will depend on the specific observable we are interested in.

2. Perturbative input: coefficient functions and DGLAP evolution

$$\mathcal{C}_{I,i} = \sum_{k=0} \mathcal{C}_{I,i}^{(k)}\,\alpha_s^k$$

$$f_i\left(x, Q^2\right) = \sum_j \Gamma_{ij}\left(Q^2, Q_0^2\right) \otimes f_j\left(Q_0^2\right) \quad \text{with} \quad \Gamma_{ij} = \sum_{k=0} \Gamma_{ij}^{(k)}\,\alpha_s^{k+1}$$

3. Lattice data

Having these ingredients we can run on lattice data *the exact same framework* used for experimental data. Let's see two examples...

quasi-PDFs (large-P_3 limit)

[X. Ji, Phys. Rev. Lett. 110 (2013)]

$$q\left(y,\mu^2,P_3^2\right)=\int_{-1}^{1}\frac{d\xi}{|\xi|}f\left(\xi,\mu^2\right)\mathcal{C}\left(\frac{y}{\xi},\frac{\mu^2}{\xi^2P_3^2}\right)+\mathcal{O}\left(\frac{M^2}{P_3^2},\frac{\Lambda_{QCD}^2}{P_3^2}\right)$$

$$\mathcal{C}\left(\xi,\eta\right)=\delta\left(1-\xi\right)+\frac{\alpha_s}{2\pi}C_F\begin{cases}\left[\frac{1+\xi^2}{1-\xi}\log\frac{\xi}{\xi-1}+1+\frac{3}{2\xi}\right]_{+(1)} & \xi>1\\ \left[\frac{1+\xi^2}{1-\xi}\log\left[\frac{1}{\eta^2}\left(4\xi\left(1-\xi\right)\right)\right]-\frac{\xi(1+\xi)}{1-\xi}\right]_{+(1)} & 0<\xi<1\\ \left[-\frac{1+\xi^2}{1-\xi}\log\frac{\xi}{\xi-1}-1+\frac{3}{2(1-\xi)}\right]_{+(1)} & \xi<0\end{cases}.$$

$$\mathcal{M}\left(\nu,-z_3^2;\,\mu^2\right)=\int_{-\infty}^{\infty}dx\,e^{-i(xP_z)z}\int_{-1}^{+1}\frac{dy}{|y|}C_3\left(\frac{x}{y},\frac{\mu}{|y|P_z}\right)f_3\left(y,\mu^2\right)$$

$$\text{Re}\left[\mathcal{M}\right]\left(\nu,-z_3^2;\,\mu^2\right)=\int_0^1 dx\,\tilde{C}^{\,\text{Re}}\left(x\nu,\mu^2z_3^2\right)V_3\left(x,\mu^2\right)$$

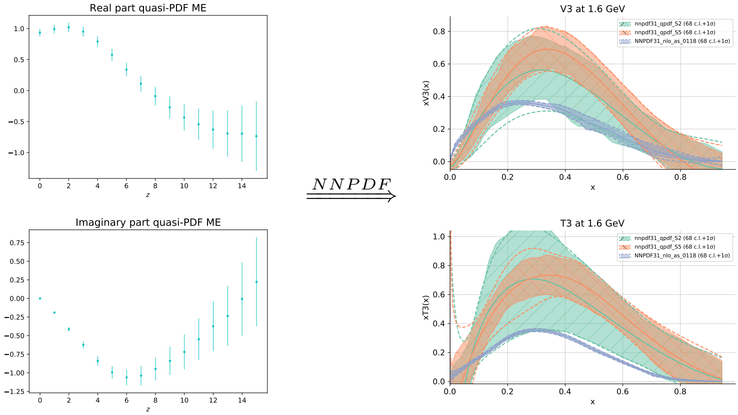
$$\text{Im}\left[\mathcal{M}\right]\left(\nu,-z_3^2;\,\mu^2\right)=\int_0^1 dx\,\tilde{C}^{\,\text{Im}}\left(x\nu,\mu^2z_3^2\right)T_3\left(x,\mu^2\right)$$

$$V_3\left(x\right)=u\left(x\right)-\bar{u}\left(x\right)-\left[d\left(x\right)-\bar{d}\left(x\right)\right],$$

$$T_3\left(x\right)=u\left(x\right)-\bar{u}\left(x\right)+\left[d\left(x\right)-\bar{d}\left(x\right)\right]$$

[JHEP 137(2019), K. Cichy , L. Del Debbio and TG]

Quasi-PDFs data from ETMC collaboration [Phys. Rev. Lett. 121 (2018)]



- 16 points for quasi-PDFs ME (high momenta only)
- different possible scenarios for systematic uncertainties (backup slides)

Reduced pseudo-ITD (small-z_3 limit)

[Phys. Rev. D96 (2017), Phys. Lett. B781 (2018), Phys. Rev. D100 (2019), ..]

$$\mathfrak{M}\left(\nu, -z_3^2\right) = \frac{\mathcal{M}\left(\nu, -z_3^2; a\right)}{\mathcal{M}\left(0, -z_3^2; a\right)}.$$

At 1-loop we have

$$\mathfrak{M}\left(\nu, -z_3^2\right) = \int_{-1}^{1} dx\, C\left(x\nu, \mu^2 z_3^2\right) f\left(x, \mu^2\right) + \mathcal{O}\left(z_3^2 \Lambda^2\right),$$

$$C\left(\xi, \mu^2 z_3^2\right) = e^{i\xi} - \frac{\alpha_s}{2\pi} C_F \int_0^1 dw \left[\frac{1+w^2}{1-w} \log\left(z_3^2 \mu^2 \frac{e^{2\gamma_E+1}}{4}\right) + 4\frac{\log\left(1-w\right)}{1-w} - 2\left(1-w\right)\right]_+ e^{i\xi w} + \mathcal{O}\left(\alpha_s^2\right).$$

$$\text{Re}\left[\mathfrak{M}\right]\left(\nu, -z_3^2\right) = \int_0^1 dx\, C^{\text{Re}}\left(x\nu, \mu^2 z_3^2\right) V_3\left(x, \mu^2\right),$$

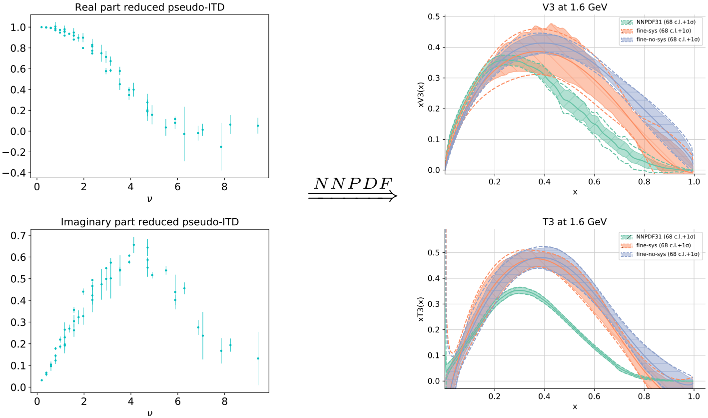
$$\text{Im}\left[\mathfrak{M}\right]\left(\nu, -z_3^2\right) = \int_0^1 dx\, C^{\text{Im}}\left(x\nu, \mu^2 z_3^2\right) T_3\left(x, \mu^2\right)$$

$$V_3\left(x\right) = u\left(x\right) - \bar{u}\left(x\right) - \left[d\left(x\right) - \bar{d}\left(x\right)\right],$$

$$T_3\left(x\right) = u\left(x\right) - \bar{u}\left(x\right) + \left[d\left(x\right) - \bar{d}\left(x\right)\right]$$

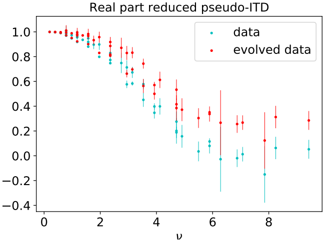
[arXiv:2010.03996, L.Del Debbio, TG, J. Karpie , K. Orginos, A. Radyushkin and S. Zafeiropoulos]

Reduced pseudo-ITD data from HadStruc Collaboration [JHEP 12 (2019) 081, 2004.01687]



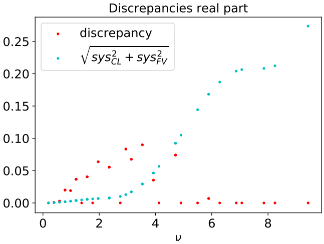
- 48 points for pseudoITD ME (all values of momentum can be used)
- no need to invert the factorization formula: we can use directly position space equations
- systematic uncertainties are accounted for in the analysis

$$\text{Re}\left[\mathcal{M}\right]\left(\nu, z_0^2\right) = \text{Re}\left[\mathcal{M}\right]\left(\nu, z^2\right) - C_F \frac{\alpha_s}{2\pi} \log \frac{z_0^2}{z^2} \int_0^1 dx \left[\int_0^1 dw\, B\left(w\right) \cos\left(x\nu w\right)\right] V_3\left(x, \mu^2\right)$$



- After evolution to a common scale z_0^2 points having same Ioffe-time should be compatible within error
  $\rightarrow$ tension between some datapoints
- need to account for systematic errors

- Finite Volume and Continuum Limit systematics estimate by lattice data
- Not enough to account for observed discrepancies



Including an unknown systematic effects

Examples from the literature: [Eur.Phys.J.C 79 (2019), JCAP 07 (2018)]

Figure of merit minimized in a fit

$$\mathcal{P}\left(D|\theta\right) = e^{-\frac{1}{2}(D-T(\theta))^T\,\Sigma^{-1}\,(D-T(\theta))}\,. \tag{1}$$

If we assume the presence of unknown systematic Δ

$$\rightarrow \mathcal{P}\left(D,\Delta|\theta\right) = e^{-\frac{1}{2}(D+\Delta-T(\theta))^T\,\Sigma^{-1}\,(D+\Delta-T(\theta))}\,. \tag{2}$$

Assuming a gaussian prior $\mathcal{P}\left(\Delta\right) = \exp\left[-\frac{1}{2}\Delta^T\,\hat{\Sigma}^{-1}\,\Delta\right]$, marginalize over Δ

$$\int d\Delta\,\mathcal{P}\left(\Delta\right)\mathcal{P}\left(D,\Delta|\theta\right) \propto e^{-\frac{1}{2}(D-T(\theta))\left(\Sigma+\hat{\Sigma}\right)^{-1}(D-T(\theta))}, \tag{3}$$

which defines the relevant likelihood to be minimized.

unknown systematic effects can be accounted for by introducing in the likelihood an additional contribution to the covariance matrix, denoted by $\hat{\Sigma}$, which defines the prior probability distribution of these systematics.

Our recipe for this exercise:

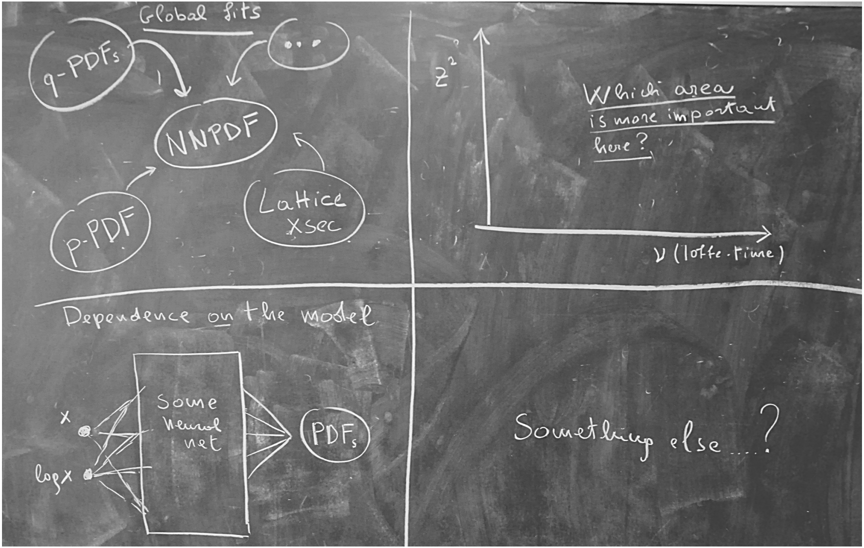
"...for each couple of points having a given Ioffe-time value, we will define the two corresponding diagonal components of $\hat{\Sigma}$ *as half of the distance between evolved points, setting the off diagonal elements to zero. Each point sharing the same Ioffe time value with at least another one will therefore be affected by an additional, uncorrelated systematic such that, after evolution, datapoints having the same Ioffe-time will be compatible between each other."*

Summary and ideas for future work

Summary

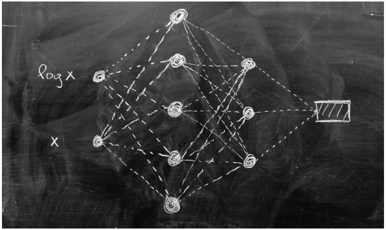
- PDFs usually extracted from experimental data through global analysis $\rightarrow$ NNPDF framework
- we can compute in lattice QCD some quantity related to PDFs through a factorization theorem. Data for such lattice observables can be treated on the same footing as experimental data.
- we have extracted PDFs from a selection of the available lattice data, studying the possible impact of systematics over the resulting distributions
  1. q-PDFs matrix elements
  2. p-PDFs matrix elements
- full information regarding systematic uncertainties and their correlation is necessary to get statistically meaningful numbers

Some ideas for future work

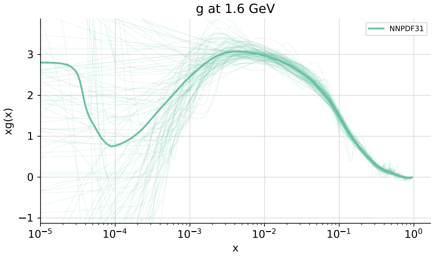


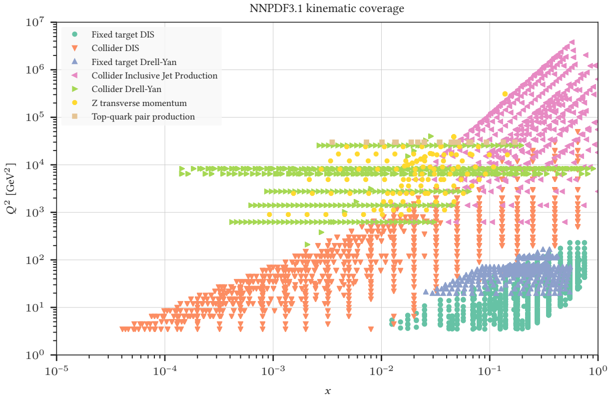

THANKS!

$$f_q\left(x, Q_0^2\right) = x^{\alpha_q}\left(1-x\right)^{\beta_q} \mathsf{NN}_q\left(x\right)$$



$$\chi^2\left(\theta\right) = \frac{1}{N_{\text{dat}}}\sum_{i,j}\left(\mathcal{O}_i - \mathcal{O}_i^{\text{th}}\left(\theta\right)\right)\left[\text{Cov}^{-1}\right]_{ij}\left(\mathcal{O}_j - \mathcal{O}_j^{\text{th}}\left(\theta\right)\right), \qquad \theta = \text{free parameters}$$





	N_{dat} (NNLO/NLO)	χ^2/N_{dat} (NNLO)	χ^2/N_{dat} (NLO)
FT DIS	1881/1881	1.15	1.20
HERA DIS	1211/1221	1.11	1.14
FT DY	189/189	1.25	0.96
Tevatron	150/156	1.08	1.06
ATLAS	360/358	1.09	1.37
CMS	409/397	1.06	1.20
LHCb	85/93	1.47	1.61
Total	4285/4295	1.148	1.168

Lattice data systematics: q-PDFs case

Different sources of systematics from the lattice simulation

- cut-off effects (finite value of lattice spacing a, UV regulator)
- finite volume effects (finite size of the box L, IR reguator)
- excited states contaminations
- truncation effects (coefficients to go from RI'-MOM to minimal subtraction)

We came up with different possible scenarios

Scenario	Cut-off	FVE	Excited states	Truncation
S1	10%	2.5%	5%	10%
S2	20%	5%	10%	20%
S3	30%	$e^{-3+0.062z/a}\%$	15%	30%
S4	0.1	0.025	0.05	0.1
S5	0.2	0.05	0.1	0.2
S6	0.3	$e^{-3+0.062z/a}$	0.15	0.3

q-PDFs and p-PDFs results