Parton distribution functions for beyond-standard-model searches

Intersections of BSM phenomenology and QCD for new physics searches

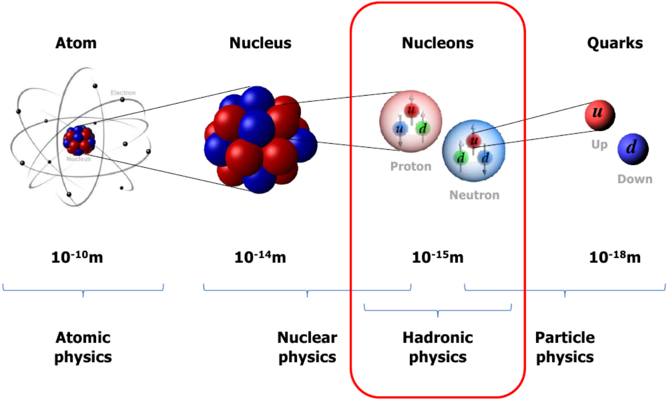
Emanuele R. Nocera

Rudolf Peierls Centre for Theoretical Physics
University of Oxford

Institute for Nuclear Theory - October 19, 2015

Hadron physics, or the quest for the nucleon structure

Nucleons make up all nuclei, and hence most of the visible matter in the Universe
They are bound states with internal structure and dynamics
Such a structure is encoded in Parton Distribution Functions
Parton Distribution Functions are essential tools in high-energy particle physics



Outline

1. The QCD structure of nucleons
    - Theory: factorization, evolution
2. A global analysis of parton distributions
    - Practice: methodology, experimental data
3. PDFs and new physics at the LHC and beyond
    - QED corrections, resummation, polarization (with phenomenology)
4. Conclusions

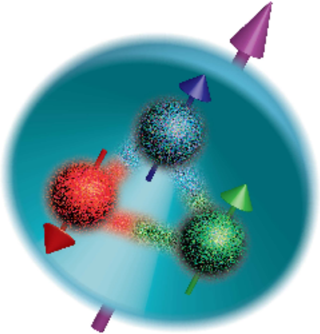
DISCLAIMER

Not a comprehensive review of recent developments in PDF analyses.
Rather, a partial and subjective view on parton distributions
mostly based on results obtained by the NNPDF Collaboration recently.
Apologies in advance for not mentioning your favorite subject.

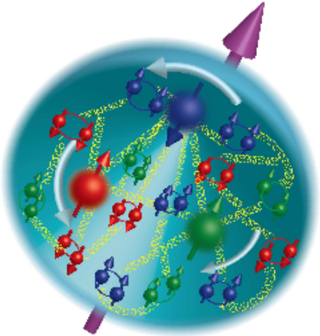
1. The QCD structure of nucleons

The QCD picture of the nucleon

naive picture



realistic picture



three non-relativistic quarks

$\overset{QCD}{\underset{factorization, evolution}{\longleftrightarrow}}$

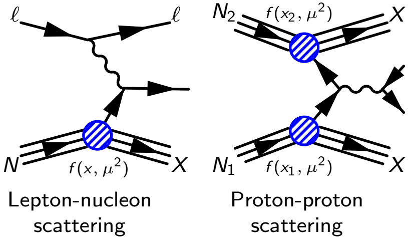
indefinite number of relativistic quarks and gluons

Factorization of physical observables [Adv.Ser.Direct.High Energy Phys. 5 (1988) 1]

1. A variety of sufficiently inclusive processes allow for a factorized description

short-distance part	factorization ⟷ scheme & scale μ	long-distance part
hard interaction of partons		nucleon structure
process-dependent kernels		universal parton distributions

2. Physical observables are written as a convolution of coefficient functions and PDFs



$$\mathcal{O}_I = \sum_{f=q,\bar{q},g} C_{If}(y,\alpha_s(\mu^2)) \otimes f(y,\mu^2)$$

$$\sigma = x\sum_{a,b} \frac{1}{y}\hat{\sigma}_{ab}(y,\alpha_s(\mu^2)\otimes \mathcal{L}_{ab}(y,\mu^2)$$

$$\mathcal{L} = f_1(y_1,\mu^2)\otimes f_2(y_2,\mu^2)$$

$$f\otimes g = \int_x^1 \frac{dy}{y} f\left(\frac{x}{y}\right) g(y)$$

3. Coefficient functions/partonic cross sections allow for a perturbative expansion

$$C_{If}(y,\alpha_s) = \sum_{k=0} a_s^k C_{If}^{(k)}(y) \qquad \hat{\sigma}_{ab}(y,\alpha_s(\mu^2)) = \sum_{k=0} a_s^k \hat{\sigma}_{ab}^{(k)}(y), \qquad a_s = \alpha_s/(4\pi)$$

4. Incredible progress in higher-order computations of $\hat{\sigma}_{ab}^{(k)}$ recently

Scale-dependence of PDFs: DGLAP equations [NP B126 (1977) 298]

1. A set of $(2n_f + 1)$ integro-differential equations, n_f is the number of active flavors

$$\frac{\partial}{\partial \ln \mu^2} f_i(x,\mu^2) = \sum_j^{n_f} \int_x^1 \frac{dz}{z} P_{ji}\left(z, \alpha_s(\mu^2)\right) f_j\left(\frac{x}{z}, \mu^2\right)$$

2. Often written in a convenient basis of PDFs

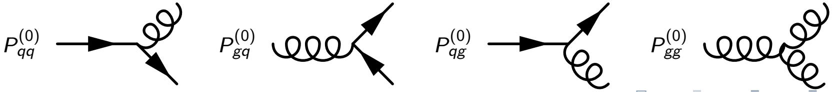
$$q_{\mathrm{NS};\pm} = (q_i \pm \bar{q}_i) - (q_j \pm \bar{q}_j) \qquad q_{\mathrm{NS;v}} = \sum_i^{n_f} (q_i - \bar{q}_j) \qquad \Sigma = \sum_i^{n_f} (q_i + \bar{q}_j)$$

$$\frac{\partial}{\partial \ln \mu^2} q_{\mathrm{NS};\pm,v}(x,\mu^2) = P^{\pm,\mathrm{v}}(x,\mu_F^2) \otimes q_{\mathrm{NS};\pm,v}(x,\mu^2)$$

$$\frac{\partial}{\partial \ln \mu^2} \begin{pmatrix} \Sigma(x,\mu^2) \\ g(x,\mu^2) \end{pmatrix} = \begin{pmatrix} P^{qq} & P^{gq} \\ P^{qg} & P^{gg} \end{pmatrix} \otimes \begin{pmatrix} \Sigma(x,\mu^2) \\ g(x,\mu^2) \end{pmatrix}$$

3. With perturbative computable splitting functions

$$P_{ji}(z,\alpha_s) = \sum_{k=0} a_s^{k+1} P_{ji}^{(k)}(z), \qquad a_s = \alpha_s/(4\pi)$$

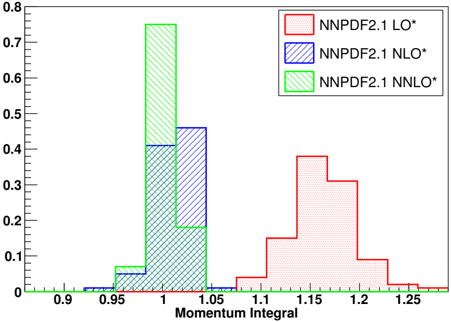


Theoretical constraints

1. Momentum & valence sum rules

$$M \equiv \int_0^1 dx\, x\, [\Sigma + g] = 1 \qquad \int_0^1 dx\, [u - \bar{u}] = 2 \qquad \int_0^1 dx\, \left[d - \bar{d}\right] = 1$$

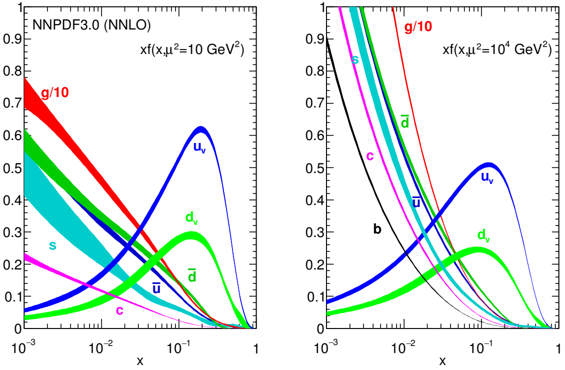
2. Positivity of cross sections
3. Can sum rules be satisfied automatically without being imposed? [NP B855 (2012) 608]



$M^{\mathrm{LO}} = 1.161 \pm 0.032 \qquad M^{\mathrm{NLO}} = 1.011 \pm 0.018 \qquad M^{\mathrm{NNLO}} = 1.002 \pm 0.014$

PDFs: some general remarks

1. There is one independent PDF for each parton in the proton
2. Heavy quark PDFs are generated radiatively
3. Beyond LO PDFs become scheme-dependent
4. The shape and the normalization of PDFs are very different for each flavor

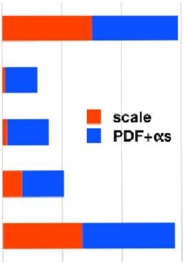


PDFs (and their uncertainties): why should we bother?

First:

PDFs are a fundamental limit for Higgs boson characterization

		σ (8 TeV)	uncertainty
NNLL QCD +NLO EW	gg→H	19.5 pb	14.7%
NNLO QCD +NLO EW	VBF	1.56 pb	2.9%
	WH	0.70 pb	3.9%
	ZH	0.39 pb	5.1%
NLO QCD	ttH	0.13 pb	14.4%

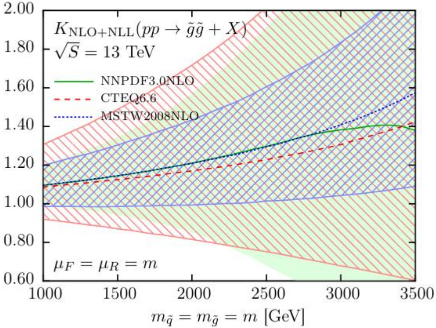


1. All production cross sections require an accurate knowledge of PDFs
   → *gg* fusion, *ttH* (gluon luminosity)
   → vector-boson fusion (quark-quark luminosity)
   → associated production with W/Z: quark-antiquark luminosity
2. PDF uncertainties are now dominant for a number of crucial LHC processes
   → example: Higgs production in *gg* fusion, known up to N^{3}LO [PRL 114 (2015) 212001]

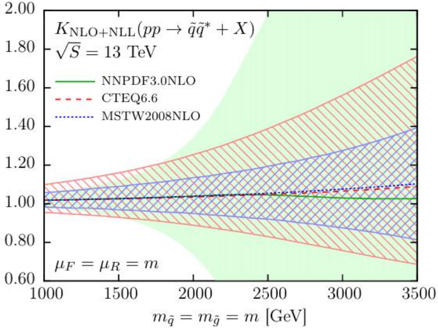
PDFs (and their uncertainties): why should we bother?

Second:
PDF uncertainties are huge ($> 100\%$) for BSM heavy particle production

gluino pair production



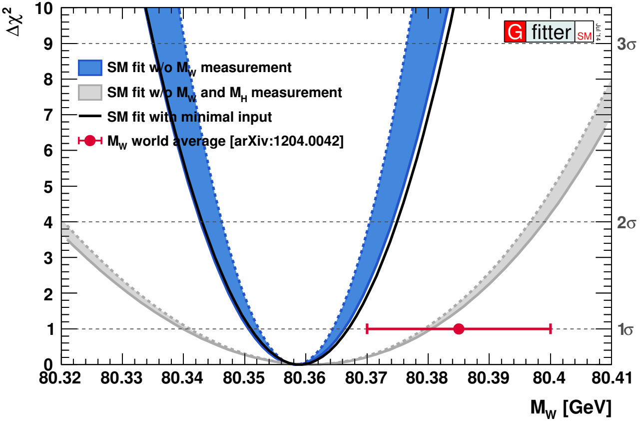
squark-antisquark pair production



$$K_{\mathrm{NLO+NLL}} = \frac{\sigma^{\mathrm{NLO+NLL}}}{\sigma^{\mathrm{NLO}}}$$

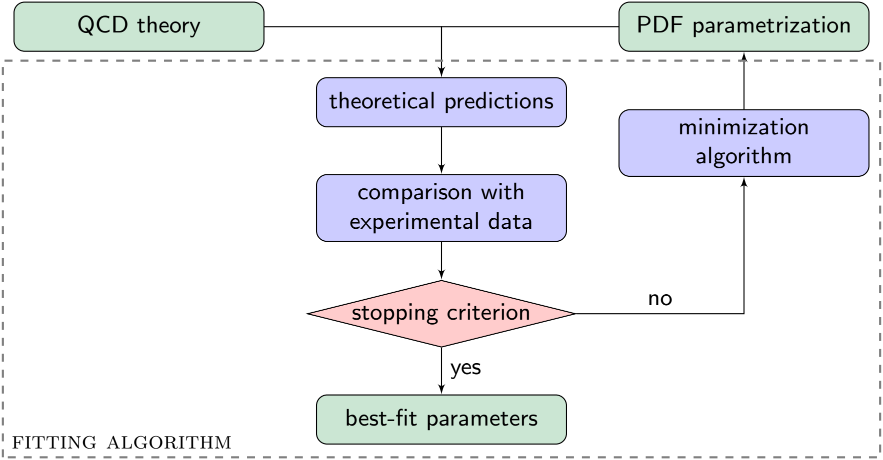
PDFs (and their uncertainties): why should we bother?

Third:
PDFs are the dominant source of systematics for precision measurements, like the W bson mass, that provide consistency stress-tests of the Standard Model



2. A global analysis of parton distributions

A global PDF determination: the underlying strategy



Assume a reasonable PDF parametrization

Obtain theoretical predictions for various processes and compare predictions to data

Determine the best-fit parameters via minimization of a proper figure of merit (*e.g.* χ^2)

A global PDF determination: the ingredients we need

theory	methodology	data
DGLAP evolution	parametrization	fixed-target data
partonic cross sections	uncertainty estimation	(DIS, DY)
heavy quark treatment	error propagation	collider data
QED/EW corrections	minimization strategy	(*ep pp*)

Need for a choice of

1. **theory**, or the theoretical details of the QCD analysis
   (perturbative order, treatment of heavy quarks, treatment of α_s, theoretical constraints)
2. **methodology**, or a prescription to determine PDFs and their uncertainties
   (uncertainty estimates are crucial to make reliable predictions based on PDFs)
3. **data**, or the set of observables to be included in the analysis
   (constrain all possible PDFs in the widest range of Bjorken-x)

Each of these ingredients is a source of uncertainty on the PDF determination

Methodology: the standard route

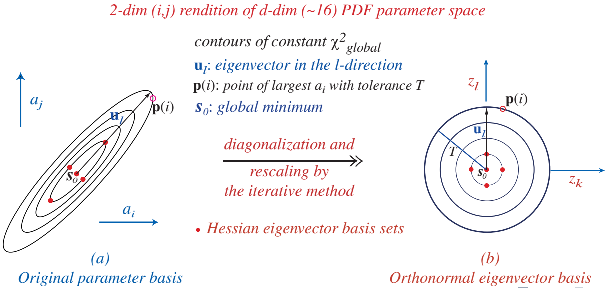
1. Simple analytical parametrization of PDFs, *e.g.*

$$xf(x,\mu_0^2) = \eta_f x^{a_f}(1-x)^{b_f}\left(1+\rho_f x^{\frac{1}{2}}+\gamma_f x\right) \qquad \{\mathbf{a}\} = \{a,b,\eta,\rho,\gamma\}$$

$\Rightarrow$ **potential bias if the parametrization is too rigid**

2. Hessian propagation of errors
   - expand the χ^2 about its global minimum at first order, $\chi^2\{\mathbf{a}\} \approx \chi^2\{\mathbf{a_0}\} + \delta a^i H_{ij}\delta a^j$
   - diagonalize the Hessian matrix and take the hypersphere of radius $\sqrt{\chi^2} = 1$

$\Rightarrow$ **is linear approximation adequate? do we need a tolerance $\mathbf{T} = \sqrt{\chi^2} > 1$?**



Methodology: the NNPDF route

1. Neural network parametrization of PDFs
   - redundant and flexible parametrization, $\mathcal{O}(200)$ parameters
   - requires a proper minimization algorithm and stopping criterion

   $\Rightarrow$ **reduce the theoretical bias due to the parametrization**

2. Monte Carlo propagation of errors
   - generate experimental data replicas assuming multi-Gaussian probability distribution
   - validate against experimental data to determine the sample size ($N_{\rm rep} \sim 100$)

   $\Rightarrow$ **no need to rely on linear error propagation, no tolerance needed**

PDF replicas are equally probable members of a statistical ensemble which samples the probability density $\mathcal{P}[f_i]$ in the space of PDFs

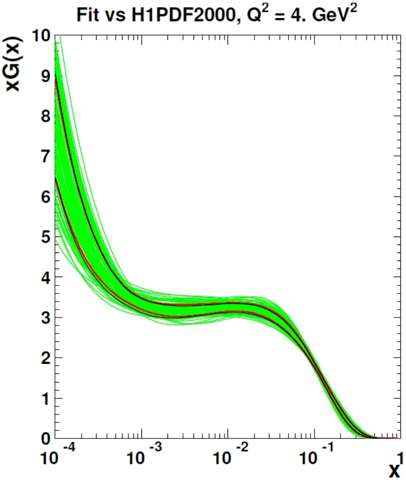
$$\langle \mathcal{O} \rangle = \int \mathcal{D}f_i \mathcal{P}[f_i] \mathcal{O}[f_i]$$

Expectation values for observables are Monte Carlo integrals

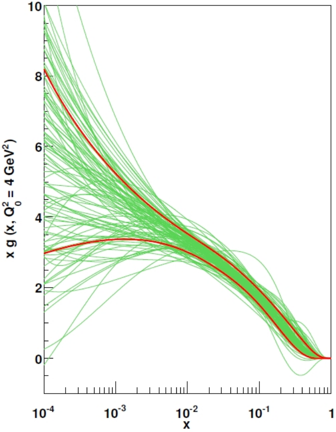
$$\langle \mathcal{O}[f_i(x,Q^2)] \rangle = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \mathcal{O}[f_i^{(k)}(x,Q^2)]$$

Methodology: standard *vs* neural network parametrization

HERA-LHC 2009 PDF benchmark

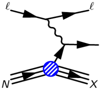
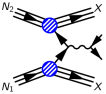


simple functional forms



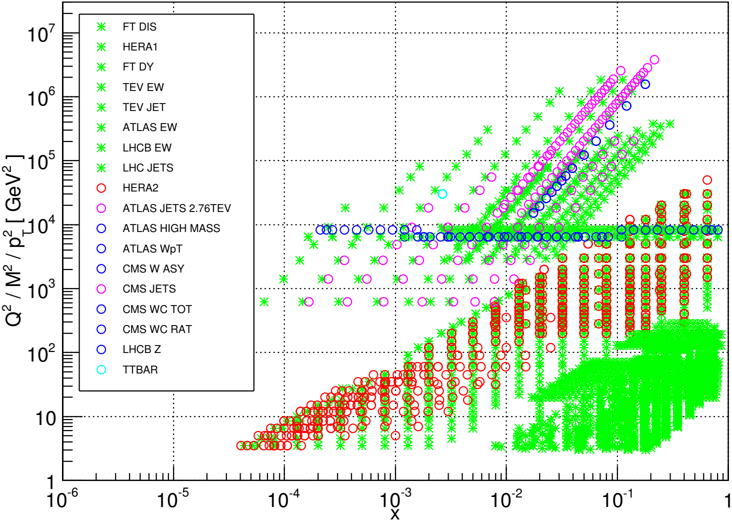
neural networks

Experimental data in a global PDF determination

Process	Reaction	Subprocess	PDFs probed	x
 DIS	$\ell^{\pm}\{p,n\} \to \ell^{\pm} + X$	$\gamma^* q \to q$	$q, \bar{q}, g$	$x \gtrsim 0.01$
	$\ell^{\pm} n/p \to \ell^{\pm} + X$	$\gamma^* d/u \to d/u$	d/u	$x \gtrsim 0.01$
	$\nu(\bar{\nu})N \to \mu^-(\mu^+) + X$	$W^* q \to q'$	$q, \bar{q}$	$0.01 \lesssim x \lesssim 0.5$
	$\nu N \to \mu^- \mu^+ + X$	$W^* s \to c$	s	$0.01 \lesssim x \lesssim 0.2$
	$\bar{\nu} N \to \mu^+ \mu^- + X$	$W^* \bar{s} \to \bar{c}$	$\bar{s}$	$0.01 \lesssim x \lesssim 0.2$
	$e^{\pm} p \to e^{\pm} + X$	$\gamma^* q \to q$	$g, q, \bar{q}$	$0.0001 \lesssim x \lesssim 0.1$
	$e^+ p \to \bar{\nu} + X$	$W^+\{d,s\} \to \{u,c\}$	d, s	$x \gtrsim 0.01$
	$e^{\pm} p \to e^{\pm} c\bar{c} + X$	$\gamma^* c \to c, \gamma^* g \to c\bar{c}$	c, g	$0.0001 \lesssim x \lesssim 0.1$
	$e^{\pm} p \to jet(s) + X$	$\gamma^* g \to q\bar{q}$	g	$0.01 \lesssim x \lesssim 0.1$
 pp	$pp \to \mu^+ \mu^- + X$	$u\bar{u}, d\bar{d} \to \gamma^*$	$\bar{q}$	$0.015 \lesssim x \lesssim 0.35$
	$pn/pp \to \mu^+ \mu^- + X$	$(u\bar{d})/(u\bar{u}) \to \gamma^*$	$\bar{d}/\bar{u}$	$0.015 \lesssim x \lesssim 0.35$
	$p\bar{p}(pp) \to jet(s) + X$	$gg, qg, qq \to 2jets$	g, q	$0.005 \lesssim x \lesssim 0.5$
	$p\bar{p} \to (W^{\pm} \to \ell^{\pm}\nu) + X$	$ud \to W^+, \bar{u}\bar{d} \to W^-$	$u, d, \bar{u}, \bar{d}$	$x \gtrsim 0.05$
	$pp \to (W^{\pm} \to \ell^{\pm}\nu) + X$	$u\bar{d} \to W^+, d\bar{u} \to W^-$	$u, d, \bar{u}, \bar{d}, (g)$	$x \gtrsim 0.001$
	$p\bar{p}(pp) \to (Z \to \ell^+\ell^-) + X$	$uu, dd(u\bar{u}, d\bar{d}) \to Z$	$u, d(g)$	$x \gtrsim 0.001$
	$pp \to (W + c) + X$	$gs \to W^- c, g\bar{s} \to W^+\bar{c}$	$s, \bar{s}$	$x \sim 0.01$
	$pp \to t\bar{t} + X$	$gg \to t\bar{t}$	g	$x \sim 0.01$

CERN	CERN	SLAC	DESY	FERMILAB
NMC, BCDMS, CHORUS	ATLAS, CMS, LHCb	E142, E143, E154, E155	HERA, ZEUS, H1	NuTev, E605, E866, CDF, D0

Experimental data in a global PDF determination



$+$ kinematic cuts in order to remove the sensitivity to higher-twist effects

$$N_{\rm dat}^{\rm unp} \sim \mathcal{O}(4000)$$

Recent determinations of PDF sets

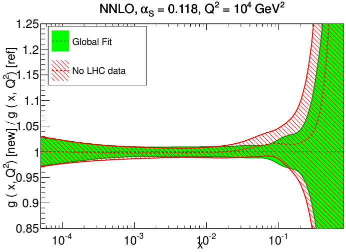
	CT14	MMHT14	NNPDF3.0	ABM12	HERAPDF1.5
fixed-target DIS	✓	✓	✓	✓	✗
HERA	✓	✓	✓	✓	✓
fixed-target DY	✓	✓	✓	✓	✗
Tevatron (W, Z)	✓	✓	✓	✗	✗
Tevatron (jets)	✓	✓	✓	✗	✗
LHC	✓	✓	✓	✓ (W, Z)	✗
statistical treatment	Hessian $\Delta\chi^2 = 100$	Hessian $\Delta\chi^2$ dynamical	Monte Carlo	Hessian $\Delta\chi^2 = 1$	Hessian $\Delta\chi^2 = 100$
parametrization	Bernstein pol. (28 pars)	Čebyčëv pol. (25 pars)	neural network (259 pars)	polynomial (14 pars)	polynomial (14 pars)
HQ scheme	ACOT-χ	TR′	FONLL	FFN	TR′
α_s	varied	fitted+varied	varied	fitted	varied
latest update	arXiv:1506.07443	EPJ C75 (2005) 204	JHEP 1504 (2015) 040	PR D89 (2014) 054028	PoS EPS-HEP2011 (2011) 320

All PDF sets listed above are available through the LHAPDF interface

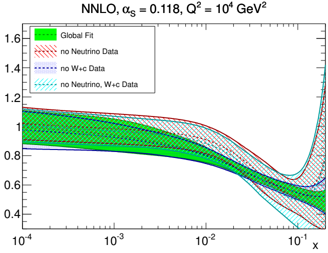
`https://lhapdf.hepforge.org/`

Parton distributions with LHC data [JHEP 1504 (2015) 040]

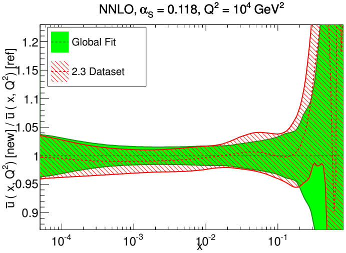
inclusive jets and $t\bar{t}$ production: large-x gluon

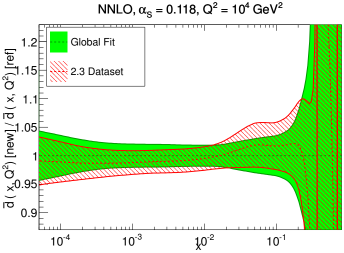


$W + c$: direct handle on strangeness



inclusive W/Z production: $q/\bar{q}$

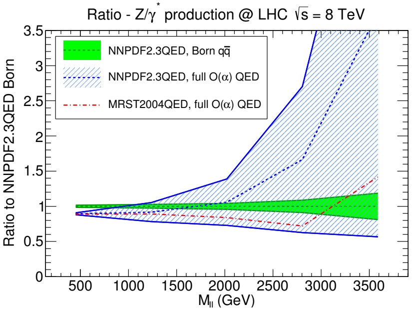




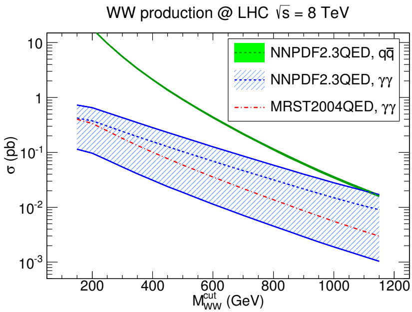
3. PDFs and new physics at the LHC and beyond

PDFs with QED corrections [NP B877 (2013) 290]

1. QED and electroweak corrections are essential for precision LHC phenomenology
   - *W* and *Z* production, *W*-mass determination, *WW* boson pair production
   - TeV scale jet and top quark pair production, searches for new W', Z' bosons
   - neglecting photon-initiated contributions underestimates theory errors in crucial BSM search channels
2. The inclusion of EW effects requires PDFs with QED corrections and a γ PDF
3. First determination of a γ PDF from LHC data: `NNPDF2.3QED` [NP B877 (2013) 290]



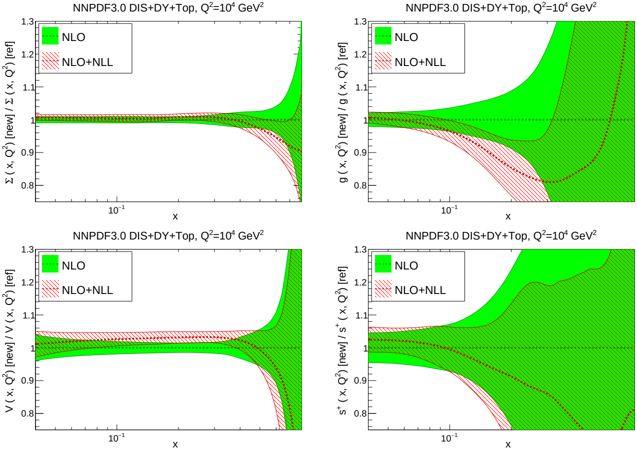
$\gamma\gamma \to \ell^+\ell^-$



$\gamma\gamma \to WW$

PDFs with threshold resummation [JHEP 1509 (2015) 191]

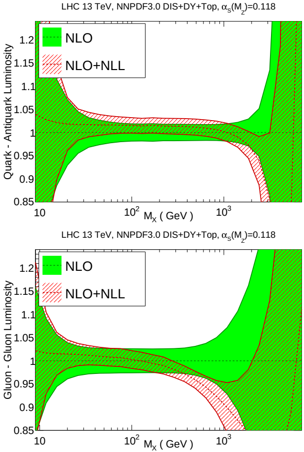
1. Resummation of threshold logs performed consistently at (N)NLO + (N)NLL
2. Only a subset of the data included in `NNPDF3.0` are included (DIS+DY+$t\bar{t}$)



The main effect of threshold resummation is to suppress all PDF flavors for $x \gtrsim 0.1$; at $0.01 \lesssim x \lesssim 0.1$ PDFs are enhanced by sum rules; the effect is negligible at $x \lesssim 0.1$

PDFs with threshold resummation [JHEP 1509 (2015) 191]

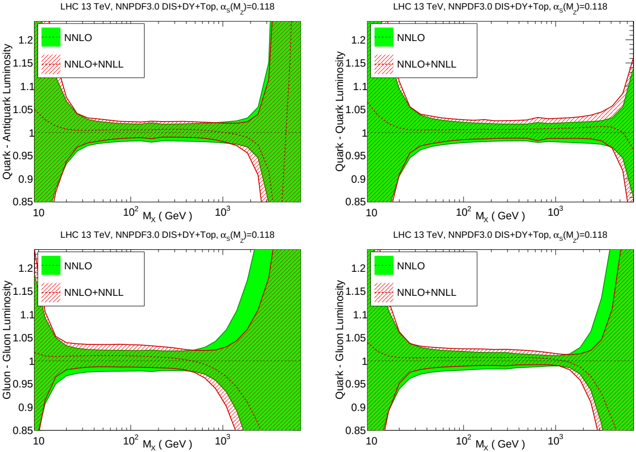
1. Resummation of threshold logs performed consistently at (N)NLO + (N)NLL
2. Only a subset of the data included in `NNPDF3.0` are included (DIS+DY+$t\bar{t}$)



The suppression at the level of PDFs becomes important for luminosities at:
$M_X \gtrsim 400$ GeV (gg); $M_X \gtrsim 1$ TeV ($q\bar{q}$ and qg); $M_X \gtrsim 5$ TeV (qq)

PDFs with threshold resummation [JHEP 1509 (2015) 191]

1. Resummation of threshold logs performed consistently at (N)NLO + (N)NLL
2. Only a subset of the data included in NNPDF3.0 are included (DIS+DY+$t\bar{t}$)



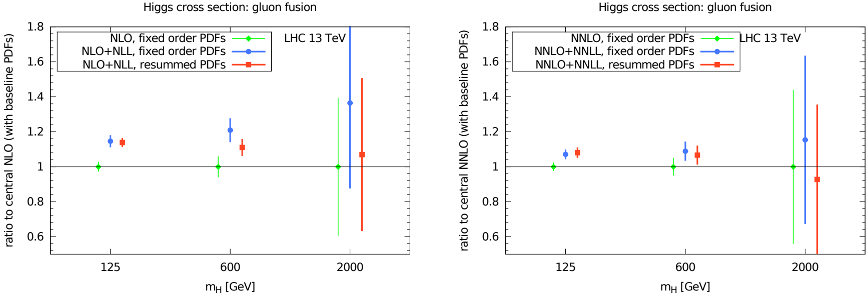
The trend is similar at NNLO,

but in this case differences between fixed-order and resummed PDFs are much smaller

PDFs with threshold resummation: phenomenology

First: SM and BSM Higgs production in gg fusion (searches for heavy Higgs)

1. Resummation in the PDFs can cancel out resummation in the matrix element
2. at NNLO this effect is less prominent
3. PDF uncertainties are large at large Higgs masses (lack of jet data in the resummed fits)
4. Using resummed PDFs for SM/BSM production at the LHC has negligible effect

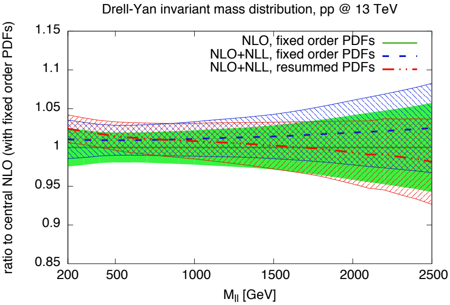


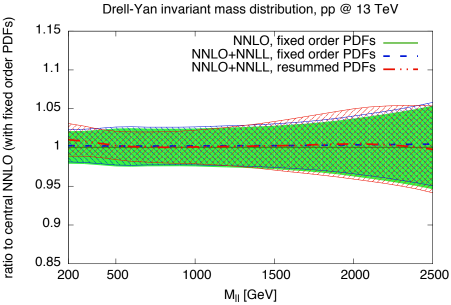
(results are normalized to the central value of the fixed (N)NLO calculation)

PDFs with threshold resummation: phenomenology

Second: High-mass DY dilepton mass distributions (searches for Z')

1. At NLO and at large invariant masses the effect of threshold resummation is moderate (the NLL correction amounts to about 4% at $M_{\ell\ell} = 2.5$ TeV, within PDF uncertainties
2. For $M_{\ell\ell} \in [1.5, 2.5]$ TeV, NLO+NLL agrees with fixed-order NLO by less than 1%
3. *A fortiori*, the effect of resummed PDFs is completely negligible at NNLO+NNLL



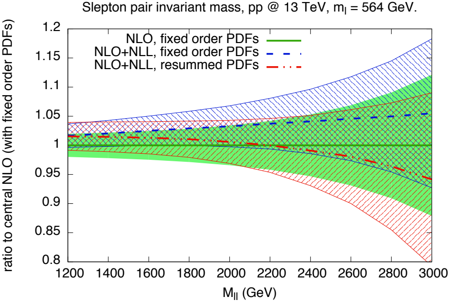


(results are obtained with `Vrap` supplemented with threshold resummation provided by `TROLL`)

PDFs with threshold resummation: phenomenology

Third: Supersymmetric sparticle (slepton pair) production (signatures of SUSY)

1. Resummation in the matrix element only:
   cross section enhancement from 1% at $M_{\ell\bar{\ell}} \sim 1.2$ TeV to 5% at $M_{\ell\bar{\ell}} \sim 3$ TeV
2. Resummation included in both PDFs and matrix element:
   cross section enhancement limited to $M_{\ell\bar{\ell}} \sim 1.2$ TeV
3. Resummation of PDFs compensate resummation of the matrix element

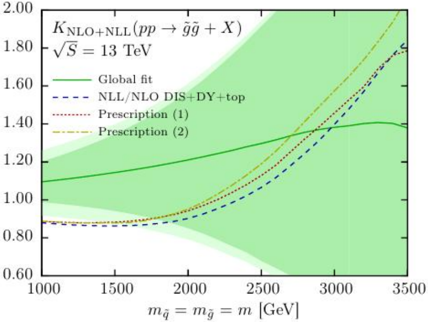


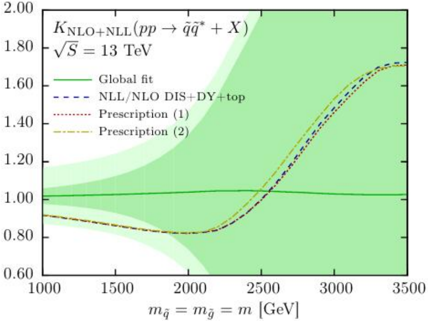
(results are obtained with `Resummino` and are normalized to the fixed NLO calculation)

PDFs with threshold resummation: phenomenology

Fourth: Squark and gluino production at the LHC run II [arXiv:1510.00375]

1. NLO+NLL cross sections are significantly shifted
2. This shift is within the total theory band, so current exclusions limits are unaffected
3. Will become crucial if we ever need to characterize SUSY particles from LHC data

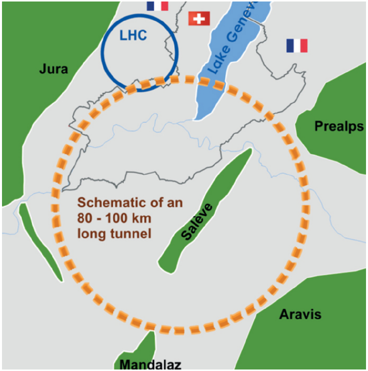


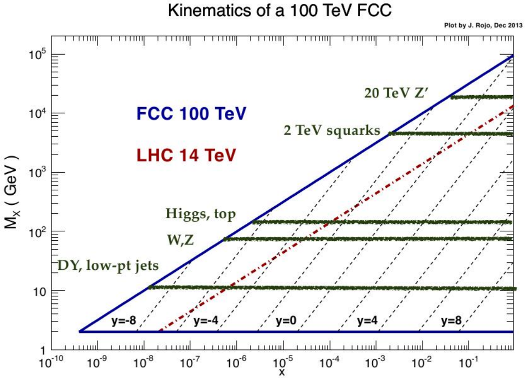


$$K \equiv \frac{\sigma^{\rm NLO+NLL}_{\rm NLO,global}}{\sigma^{\rm NLO}_{\rm NLO,global}} \qquad K^{(1)} \equiv \frac{\sigma^{\rm NLO+NLL}_{\rm NLO,global}}{\sigma^{\rm NLO}_{\rm NLO,global}} \times \frac{\sigma^{\rm NLO+NLL}_{\rm NLL,DIS+DY+top}}{\sigma^{\rm NLO+NLL}_{\rm NLO,DIS+DY+top}}$$

PDFs at a Future Circular Collider

1. A 100 TeV hadron collider, possibly with e^+e^- and *ep* operation modes
2. Rich phenomenology of PDFs at such extreme energies: top quark PDFs, EW effects on PDFs and W/Z boson PDFs, ultra-low-x physics, BFKL dynamics, ...

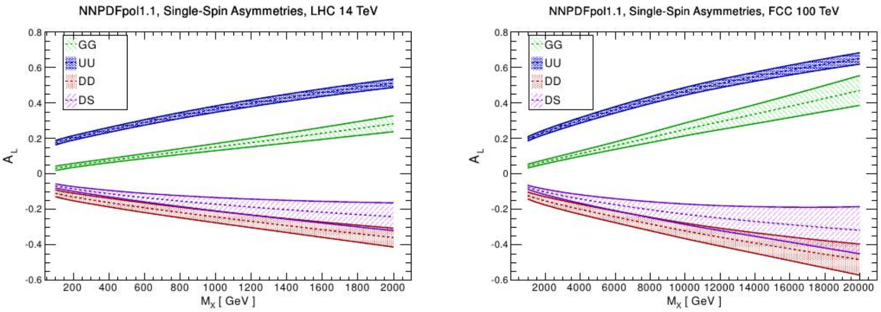




First studies are now being performed by the CERN FCC working group

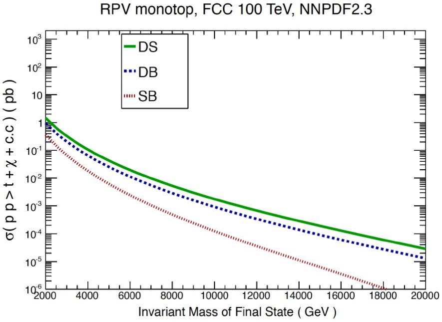
Polarized PDFs at a Future Circular Collider [JHEP 1405 (2014) 045]

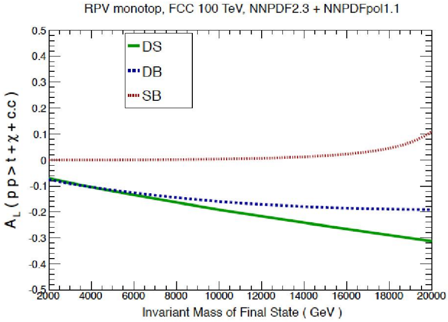
1. Polarized PDFs are rather different from their unpolarized counterparts
2. If BSM physics is discovered at high values of the final-state invariant mass, spin-asymmetry measurements could be used to characterize the structure of BSM physics
3. Provided enough statistics, BSM cross-sections will look very different if they are initiated by up quarks (large positive asymmetry), gluons (moderate positive asymmetry) or down quarks (negative asymmetry)
4. The idea is to pin down BSM couplings within a given model, and possibly to discriminate among different models that lead to the same signature



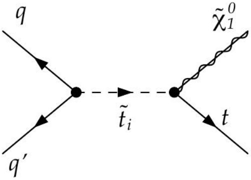
Polarized PDFs at a Future Circular Collider [JHEP 1405 (2014) 045]

Example: Monotop production in R-parity violation MSSM





1. Unpolarized case: all partonic production channels show the same qualitative behavior
2. Polarized case: asymmetries will vary between -30% and +10% depending on the dominant coupling
3. Provided a signal in unpolarised collisions, polarized data would help in understanding the nature of BSM physics



4. Conclusions

Summary

Parton Distribution Functions are an essential ingredient for LHC phenomenology

Accurate PDFs are required for precision SM measurements,
Higgs characterisation and New Physics searches

The accuracy of a PDF determination closely depends on
the experimental data, the fitting methodology and the theoretical details

A plethora of new, precise data will be available at the LHC run II
These may be supplemented with the data from RHIC and JLAB

The NNPDF methodology allows for a determination of minimally biased PDF sets

An increasing effort is being devoted to determine PDFs
including the best theory options available on the market:
QED/EW corrections, resummation, dynamic higher-twist corrections,
potential interplay with non-perturbative models, . . .

Summary

Parton Distribution Functions are an essential ingredient for LHC phenomenology

Accurate PDFs are required for precision SM measurements,
Higgs characterisation and New Physics searches

The accuracy of a PDF determination closely depends on
the experimental data, the fitting methodology and the theoretical details

A plethora of new, precise data will be available at the LHC run II
These may be supplemented with the data from RHIC and JLAB

The NNPDF methodology allows for a determination of minimally biased PDF sets

An increasing effort is being devoted to determine PDFs
including the best theory options available on the market:
QED/EW corrections, resummation, dynamic higher-twist corrections,
potential interplay with non-perturbative models, ...

Thank you for your attention