

NNPDF Parton Distribution Functions

Maria Ubiali

(RWTH Aachen)

In collaboration with

S.Forte (U Milano), J. Rojo (CERN) R.Ball, L.Del Debbio, N. Hartland (U Edinburgh) V.Bertone (U Freiburg), A.Guffanti(NBI) F.Cerutti, J.I.Latorre (U Barcelona) [NNPDF collaboration]

17th SFB/TR9 meeting, PDFs mini workshop, KIT Karlsruhe

19.03.2012

Parton Distributions for the LHC



- Can we trust PDF uncertainties?
- How do we interpret the differences between predictions obtained by using different parton sets?

 PDF4LHC: huge effort in understanding differences & improving theoretical and statistical treatments in PDF analyses

$PDFs \rightarrow LHC$

PDF uncertainties are crucial input for LHC: both for standard candle processes and for exclusion and discovery



Outline

⇒ The NNPDF parton sets

- NNPDF in a nutshell
- The NNPDF2.1 parton sets
- LHC phenomenology

⇒ Inclusion of LHC data

- Collider-only fit
- Inclusion of the LHC data by refitting
- Bayesian Reweighting in the NNPDF approach

Conclusions and outlook

The NNPDF parton sets

- NNPDF in a nutshell
- The NNPDF2.1 parton sets
- LHC phenomenology

NNPDF partons Monte Carlo & Neural Networks



NNPDF partons Monte Carlo & Neural Networks

- 1 Monte Carlo replicas of experimental data
 - Generation through MC sampling of data
 - Validation against experimental data
 - No need of relying on linear propagation of errors
 - Possibility to test for non Gaussian behavior in fitted PDFs
- 2 Fit PDFs with a set of Neural Networks for each rep
 - Redundant parametrization:
 - 7 independent PDFs, 259 free parameters
 - Dynamical stopping criterion: cross- validation method
 - NN provide an unbiased parametrization
- Sexpectation values for observables are MC integrals (same for errors, confidence levels, correlations...)



$$\begin{split} \langle \mathcal{F} \rangle &= \frac{1}{N_{\text{set}}} \sum_{i=1}^{N_{\text{set}}} \mathcal{F}[q^{(i)}] \\ \sigma_{\mathcal{F}}^{\text{MC}} &= \left(\frac{1}{N_{\text{N}_{\text{set}}}} \sum_{k=1}^{N_{\text{set}}} \left(\mathcal{F}[\{q^{(k)}\}] - \langle \mathcal{F}[\{q\}] \rangle \right)^2 \right)^{1/2} \end{split}$$

NNPDF partons Timeline

	2008 2009		2010	2010 2011		•				
							•			
	NNPDF1.0	NNPDF1.2	NNPDF2.0	NNPDF2.1 NLO	NNPDF2.1 LO and NNLO	NNPDF2.2	MSTW08	CT10	HERAPDF1.5	ABM11
DIS	~	v	v	v	~	~	V	v	×	
Drell-Yan data	×	×	>	~	~	v	~	>	×	× .
Jet data	×	×	v	~	~	~	~	 ✓ 	×	×
LHC data	×	×	×	×	×	 	×	×	×	×
Independent strange and anti-strange	×	•	~	~	v	~	~	~	~	×
Heavy Quark masses	×	×	×	~	~	~	~	~	~	v
NNLO	×	×	×	×	~	~	~	×	~	~

v only (s+sbar), s = sbar

The NNPDF2.1 parton set Experimental data

NNPDF2.1 dataset



- All systematic correlated uncertainties are included (when available)
- Kinematical cuts $Q^2 > 3 \text{ GeV}^2$ $w^2 = Q^2 (1-x)/x > 12.5 \text{ GeV}^2$
- In total 3333 (LO), 3338 (NLO), 3357 (NNLO) experimental points included

The NNPDF2.1 parton set Features of the fit

- NNPDF2.1 is an ensemble of PDF sets presently available at LO, NLO and NNLO The NNPDF collaboration, Ball et al, Nucl.Phys. B849 (2011) The NNPDF collaboration, Ball et al, Nucl.Phys. B855 (2012)
- Global fit: DIS + DY + JET data but available also for data subsets (NNPDF2.2 includes also first LHC data)
- Heavy quark mass effects included using the General-Mass FONLL scheme up to NNLO Forte, Laenen et al, Nucl.Phys. B834 (2010)
- Fast Kernel method for the inclusion of exact higher order corrections The NNPDF collaboration, Ball et al, Nucl.Phys. B838 (2010)
 - DIS up to NNLO
 - DY and JET up to NLO
- NNLO correction to DY included by means of K-factors

• NNLO correction to inclusive JET observables implemented using FastNLO with approximated NNLO corrections based on threshold resummation

> Kluge, Rabbertz, Wobisch, hep-ph/0609285 (2006) Kidonakis, Owens, Phys. Rev. D (2001)

• Available for a range of $\alpha_s = 0.114, 0.115, ..., 0.124$

$$m_c = 1.4, \dots, 1.7$$

 $m_h = 4.25, 4.5, 4.75, 5.0$

The NNPDF2.1 parton set Results





The NNPDF2.1 parton set Perturbative stability

xΣ/2 Sxg/5 ZxT₃

OxV

 Δ_{s}

]xs⁺

xs

10⁻¹

At the starting scale (2 GeV²)

NNPDF2.1 LO, $Q^2 = 2 \text{ GeV}^2$

2.5

1.5

0.5

10*



NNPDF collaboration, 1107.2652



At the EW scale (100^2 GeV^2)

10⁻³

10-2



The NNPDF2.1 parton set Perturbative stability



Momentum sum rule

Even without imposing momentum sum rules in the fit, the fitted PDFs tend to respect momentum sum rules and this tendency increases as perturbative order increases

 $[M]_{\text{LO}} = 1.161 \pm 0.032$ $[M]_{\text{NLO}} = 1.011 \pm 0.018$ $[M]_{\text{NNLO}} = 1.002 \pm 0.014$

NNPDF collaboration, 1107.2652

The NNPDF2.1 parton set Implications for LHC

At hadron colliders, observables depend on PDFs through parton luminosities

$$\Phi_{ij} = \frac{1}{S_{\text{had}}} \int_{\tau}^{1} \frac{dy}{y} f_i(y, M_X^2) f_j\left(\frac{\tau}{y}, M_X^2\right) \qquad \tau = \frac{M_X^2}{S_{\text{had}}}$$



- GG luminosity particularly stable in the standard Higgs regions
- QQ luminosity significantly larger at NNLO in the W, Z region

The NNPDF2.1 parton set Implications for LHC



- Gluon fusion Higgs production: noticeable agreement among global sets (MSTW and NNPDF at NNLO plus CT10 at NLO)
- Agreement increases when same value of $\alpha_s(M_z)$
- Sizeable difference with the ABKM09 predictions, partially accounted for different value of $\alpha_s(M_z)$

The NNPDF2.1 parton set Implications for LHC





• W and Z production: weaker dependence on $\alpha_s(M_z)$ but non negligible higher order corrections

Smaller differences among PDF sets

$$\frac{\sigma_{W^+}}{\sigma_{W^-}} \sim \frac{u(x_1)\,\bar{d}(x_2)}{d(x_1)\,\bar{u}(x_2)} \sim \frac{u(x_1)}{d(x_1)}$$

G. Watt, JHEP 1109 (2011) 069

Inclusion of LHC data

- Collider-only fit?
- Inclusion of the LHC data by refitting
- Bayesian Reweighting in the NNPDF approach

Collider-only fit



 No fixed target data
 No low energy troubles (nuclear corrections, higher twists...)

PDFs from HERA + Tevatron?

Collider-only fit







- No fixed target data
 No low energy troubles (nuclear corrections, higher twists...)
- HERA + Tevatron:
 - ✓ Good accuracy for gluon
 - * Loss of accuracy for flavor separation and strange

What about

HERA + Tevatron + LHC ?

Constraints from LHC data



- → Medium and large x gluon
 - Prompt photon
 - Precision jets data
 - Top pairs

 \rightarrow Light flavors at medium and small x

- Low-mass Drell-Yan
- Z rapidity distributions
- W asymmetries
- Polarized W

Strangeness and heavy flavors

- Wc for strangeness
- Zc and γc for charm
- Zb for bottom

Inclusion of LHC data by refitting

• To include LHC data in parton fits, TOOLS to interface slow NLO/NNLO codes to the fast computation needed in a fit are essential

APPLGRID[T. Carli et al, Eur.Phys.J. C66 (2010)]FASTNLO[T. Kluge et al, hep-ph/0609285]FASTDY[NNPDF, Nucl.Phys. B838 (2010)]

These codes are based on

- Pre-computation of hard part of the process on a grid
- Polynomial Interpolation

$$d\sigma = \sum_{p} \sum_{m=1}^{N} w_m^{(p)} \left(\frac{\alpha_s(Q_m^2)}{2\pi}\right)^p f_1(x_{1m}, Q_m^2) f_2(x_{2m}, Q_m^2)$$

Medium-long term project [work in progress... NNPDF2.3 summer 2012]

Inclusion of LHC data by refitting

N. Hartland, Moriond QCD 2012

$$W = \sum_{p} \sum_{l=0}^{n_{\rm sub}} \sum_{i_{y_1}} \sum_{i_{y_2}} \sum_{i_{\tau}} W_{i_{y_1}, i_{y_2}, i_{\tau}}^{(p)(l)} \left(\frac{\alpha_s \left(Q^{2^{(i_{\tau})}} \right)}{2\pi} \right)^p F^{(l)} \left(x_1^{(i_{y_1})}, x_2^{(i_{y_1})}, Q^{2^{(i_{\tau})}} \right)$$

Fast ... but can we get faster? \rightarrow combine weight tables with FastKernel evolution:

$$E_{\alpha\beta jk}^{\tau} = \int_{x_{\alpha}}^{1} \frac{dy}{y} \Gamma_{ij} \left(\frac{x_{\beta}}{y}, Q_{0}^{2}, Q_{\tau}^{2}\right) \mathcal{I}^{(\beta)}(y).$$
$$f_{i}(x_{\alpha}, Q_{\tau}^{2}) = \sum_{j}^{N_{\text{pdf}}} R_{ij} N_{j}(x_{\alpha}, Q_{\tau}^{2}) = \sum_{\beta}^{N_{x}} \sum_{j,k}^{N_{\text{pdf}}} R_{ij} E_{\alpha\beta jk}^{\tau} N_{k}^{0}(x_{\beta})$$

Combined Weight-Evolution tables

- More of the calculation is precomputed
- Smaller flavour basis at initial scale

$$W = \sum_{lpha,eta}^{N_{\mathrm{x}}} \sum_{i,j}^{N_{\mathrm{pdf}}} \sigma_{lphaeta_{ij}} N_i^0(x_lpha) N_j^0(x_eta)$$

In the meantime, is it possible to have a tool to estimate quickly the impact of new data without refitting?

Inclusion of LHC data by reweighting

R.D.Ball et al. ArXiv:1012.0836

- Bayesian reweighting method inspired by Giele and Keller [hep-ph/9803393]
- The IDEA:
 - \checkmark N_{rep} of a Monte Carlo fit give the probability density in the space of PDFs
 - Expectation values are MC integrals. Same for errors, correlations...
 - ✓ One can assess the impact of including new data in the fit by **updating** the probability density distribution!

Refitting:	1 1 1
Whenever add new data, need to do full refitting, tune parametrization and statistic treatment	
 Can be done only by PDF fitting collaborations themselves. 	

Reweighting:	ì
 Immediate: no need to refit 	į
 Anybody can do it just evaluating weights with each replica of a PDF set and producing a new PDF set through unweighting 	

Reweighting versus refitting

- Start from NNPDF2.0 DIS+DY only fit (BLUE)
- Add CDF and D0 jet data by refitting (as a benchmark)
- Add first CDF data by reweighting, unweight, then reweight D0 data on unweighted set
- Add first D0 data by reweighting, unweight, then reweight CDF data on unweighted set
- Add CDF+D0 at the same time by reweighting
- Obtain the same results!



The (old) W lepton asymmetry data



- D0 electron and muon asymmetry distributions [ArXiv:0709.4254] [ArXiv:0807.3367]
- ATLAS muon charge asymmetry 36 pb⁻¹ [ArXiv: 1103.2929]
- CMS electron and muon asymmetry 36 pb⁻¹ [ArXiv: 1103.3470]

Tevatron parton kinematics

LHCb muon asymmetry [LHCb-CONF-2011-039]





The NNPDF2.2 set



 NNPDF2.2: Added ATLAS and CMS W lepton asymmetry data and Tevatron D0 electron and muon asymmetry data at the same time

- These data did not include full covariance matrix
- Their inclusion reduces uncertainty and moves central values
- Total uncertainty reduction is already pretty significant
- d PDFs: tension between Tevatron and low-energy data?

Updated LHC data

- LHC data in NNPDF2.2 now superseded
 - Full covariance matrix is available for ATLAS W lepton and Z rapidity distributions
 - Higher luminosity 234 pb⁻¹ for CMS muon asymmetry
- Additional LHC data
 - 36 pb⁻¹ inclusive jet measurements (full covariance matrix ATLAS)
 - 36 pb⁻¹ LHCb Z rapidity distribution, W asymmetries at high rapidity
 - 840 pb⁻¹ CMS W electron asymmetry with full covariance matrix
 - 4.67 fb⁻¹ CMS inclusive jet measurements

χ^2 to electroweak vector boson production data

Dataset , χ^2	NNPDF2.1	MSTW08	ABKM09	JR09	HERAPDF1.5
ATLAS W/Z Rapidity	2.7	3.6	3.6	5.0	2.0
CMS μ asym + Z Rap	2.0	3.0	2.8	3.6	2.8
LHCb W asym + Z Rap	0.8	0.7	1.2	0.4	0.6

χ^2 to inclusive jet data

Dataset, χ^2	NNPDF2.1	MSTW08	ABKM09	JR09	HERAPDF1.5
ATLAS Incl. Jets $R = 0.4$	0.93	1.18	1.41	1.63	1.21
ATLAS Incl. Jets $R = 0.6$	1.38	1.31	1.46	1.88	1.43

NNLO predictions obtained with DYNNLO

Impact of the LHC EW data (ATLAS)

* Full covariance matrix provided

Dataset	χ^2	$\chi^2_{ m rw}$	N _{eff}
ATLAS	2.7	1.2	16
ATLAS W^+ 36 pb ⁻¹	5.7	1.5	17
ATLAS W^- 36 pb ⁻¹	2.5	1.0	205
ATLAS Z 36 pb^{-1}	1.8	1.1	581



Impact of the LHC EW data (CMS)

* Covariance matrix not yet available

Dataset	χ^2	$\chi^2_{ m rw}$	N _{eff}
CMS	2.0	1.2	56
CMS Z rapidity 36 pb ⁻¹	1.9	1.4	223
CMS muon asymmetry 234 pb^{-1}	2.0	0.4	200



Impact of the LHC EW data (LHCb)

* Covariance matrix not yet available

Dataset	χ^2	$\chi^2_{ m rw}$	N _{eff}
LHCb	0.8	0.8	972
LHCb Z rapidity 36 pb ⁻¹	1.1	1.0	962
LHCb W lepton asymmetry 36 pb ⁻¹	0.8	0.5	961



Impact of the LHC EW data on PDFs



Conclusions and outlook

Reliable PDF uncertainties are crucial for the LHC analyses

• NNPDF partons are produced with a method which overcomes some drawbacks of standard approaches: restricted parametrization, tolerances and tuning of statistical treatment to data included in the fits

[ArXiv:1103.2369,1102.4407,1101.1300,1012.0836,1005.0397...]

 Next steps: inclusion of LHC data, MS mass for charm and bottom quarks, exploring intrinsic charm contribution, EW corrections, study of theoretical uncertainty, inclusion of resummations

 NNPDF2.2 has the same features and it includes already the effect of early LHC data via reweighting

- Inclusion of new LHC data is crucial for obtaining up-to-date parton sets
- Next NNPDF releases will use as much as possible up-to-date LHC data

Thank you for your attention!

BACK-UP

Impact of the LHC inclusive jet data

Dataset	χ^2	$\chi^2_{ m rw}$	$N_{\rm eff}$
NNPDF2.1 NNLO + ATLAS Incl. Jets $R = 0.4$	0.93	0.91	904
NNPDF2.1 NNLO + ATLAS Incl. Jets $R = 0.6$	1.42	1.24	610



The NNPDF2.1 parton set Results

In order to quantify the impact on the NNLO corrections, one defines the distance:

$$d^{2} = \sum_{i=1}^{N_{rep}} \frac{(\mathcal{O}_{\text{NLO}}^{(i)} - \mathcal{O}_{\text{NNLO}}^{(i)})^{2}}{\sigma_{\text{NLO}}^{2} + \sigma_{\text{NNLO}}^{2}} \implies \text{for } N_{rep} = 100, \ d \sim \left\{ \begin{array}{cc} 1 & \text{Statistical equivalence} \\ \sqrt{50} \simeq 7 & 1 - \sigma \text{ shift} \end{array} \right.$$



- NLO and NNLO PDFs very similar at small-x,
- Largest distances for quarks at x ∼ 0.1 − 0.2,
- very small distances for PDF uncertainties:
 - same (and only) experimental uncertainties for both NLO and NNLO fit.

Description of SLAC data

NNPDF21_FFN_NF3 NLO predictions versus SLAC data (F^p₂ and F^d₂ overlapped)



Good description of the data

 $\chi^2/dof = 1.28$ (p) $\chi^2/dof = 1.05$ (d)

Description of HERA NC data



Strangeness of the proton



10⁻³

х

10⁻⁴

10⁻²

10⁻¹

0^上 10⁻⁵



R_s determination





No discrepancy observed with Rs for NNPDF collider +W/Z only fit

The NNPDF2.1 parton set Comparison to MSTW08





- Apart from NNPDF MSTW is presently the only NNLO PDF set available for different values of α_s
- Comparison performed at $\alpha_s(M_z) = 0.119$
- Reasonable agreement for most PDF combinations
- NNPDF singlet has larger uncertainty in extrapolation region

The NNPDF2.1 parton set Comparison to MSTW08





- Apart from NNPDF MSTW is presently the only NNLO PDF set available for different values of α_s
- Comparison performed at $\alpha_s(M_z) = 0.119$
- Reasonable agreement for most PDF combinations
- NNPDF singlet has larger uncertainty in extrapolation region
- NNPDF exhibits a more stable gluon at small-x
- Flexible param. of strangeness

The NNPDF2.1 parton set Comparison to ABKM09





- ABKM09 provide only one value of α_s determined from the fit

 $\alpha_s(M_Z)_{\rm ABKM}=0.1135$

- Harder gluon due to smaller α_s
- Worse agreement than with MSTW
- ABKM fit exhibits a smaller uncertainty in extrapolation region
- To be compared with ABM11

The NNPDF α_s determination



Remarkable perturbative stability
Somewhat lower but compatible results when fit is performed on DIS data only

 $\alpha_s^{\rm NNLO} \left(M_Z \right) = 0.1166 \pm 0.0008^{\rm stat} \pm 0.0001^{\rm proc}$

NNPDF coll, ArXiv 1110.2483 [PLB in press]

- Performed NNPDF2.1 NNLO fits at different $\alpha_s(M_z)$ with large statistics
- Determine bootstrap errors of each point

 Result in agreement with PDG value, small statistical uncertainties

 $\alpha_s^{\text{NLO}}(M_Z) = 0.1191 \pm 0.0006^{\text{stat}} \pm 0.0001^{\text{proc}}$ $\alpha_s^{\text{NNLO}}(M_Z) = 0.1173 \pm 0.0007^{\text{stat}} \pm 0.0001^{\text{proc}}$





PDF + α_s uncertainty

Exact physical parameters+PDFs uncertainty obtained as average over the joined Monte Carlo ensemble [ArXiv:1005.0397]

$$\langle \mathcal{F}
angle_{\mathrm{rep}} = \sum_{j=1}^{N^{lpha_s}} \sum_{k_j=1}^{N^{lpha_s^{(j)}}} \mathcal{F}\left(PDF^{(k_j,j)}, lpha_s^{(j)}
ight)$$





$$\langle \mathcal{F} \rangle_{\rm rep} = \frac{1}{N_{\rm rep}} \sum_{i=1}^{N_{m_c}} \sum_{j=1}^{N_{m_b}} \sum_{k_{ij}=1}^{N_{\rm rep}^{(i,j)}} \mathcal{F} \left(\text{PDF}^{(k_{ij},i,j)}, m_c^{(i)}, m_b^{(j)} \right)$$
$$N_{\rm rep}^{(i,j)} \propto \exp\left(-\frac{\left(m_c^{(i)} - m_c^{(0)} \right)^2}{2\delta_{m_c}^2} - \frac{\left(m_b^{(j)} - m_b^{(0)} \right)^2}{2\delta_{m_b}^2} \right)$$

The NMC issue

- ABKM report a 3(1)-σ shift at NNLO (NLO) on the Higgs production cross section in gluon fusion at the LHC (and Tevatron) (arXiv:1101.5261)
- Claim is different treatment of fixed target DIS NMC data: used as structure functions (MSTW, NNPDF, CT) or cross sections (ABKM) → Origin of ABKM/MSTW discrepancy?

$$\widetilde{\sigma}(x, y, Q^2) = F_2(x, Q^2) \left(2 - 2y + y^2 / \left[1 + R\left(x, Q^2\right) \right] \right) + TMCs$$

 NNPDF finds negligible impact of the treatment of NMC data for Higgs production, both at NLO (arXiv:1102.3182) and at NNLO – even removing NMC altogether has moderate effect



The NMC issue

- ABKM report a 3(1)-σ shift at NNLO (NLO) on the Higgs production cross section in gluon fusion at the LHC (and Tevatron) (arXiv:1101.5261)
- Claim is different treatment of fixed target DIS NMC data: used as structure functions (MSTW, NNPDF, CT) or cross sections (ABKM) → Origin of ABKM/MSTW discrepancy?

$$\widetilde{\sigma}(x, y, Q^2) = F_2(x, Q^2) \left(2 - 2y + y^2 / \left[1 + R\left(x, Q^2\right) \right] \right) + TMCs$$

 NNPDF finds negligible impact of the treatment of NMC data for Higgs production, both at NLO (arXiv:1102.3182) and at NNLO – even removing NMC altogether has moderate effect



Statistical distances between NNPDF2.1 NMC-F2 and NNPDF2.1 NMC-XSEC

The NMC issue

- ABKM report a 3(1)-σ shift at NNLO (NLO) on the Higgs production cross section in gluon fusion at the LHC (and Tevatron) (arXiv:1101.5261)
- Claim is different treatment of fixed target DIS NMC data: used as structure functions (MSTW, NNPDF, CT) or cross sections (ABKM) → Origin of ABKM/MSTW discrepancy?

$$\widetilde{\sigma}(x, y, Q^2) = F_2(x, Q^2) \left(2 - 2y + y^2 / \left[1 + R\left(x, Q^2\right) \right] \right) + TMCs$$

 NNPDF finds negligible impact of the treatment of NMC data for Higgs production, both at NLO (arXiv:1102.3182) and at NNLO – even removing NMC altogether has moderate effect



The treatment of NMC data has negligible impact on collider Higgs production Also at NNLO

Heavy quark schemes



• All GM-VFNS include terms included in ZM-VFN scheme and FFNS at a given order in PT, but sub-leading terms are different

The latter may have a not negligible impact

• See LH benchmark [Andersen et al, 1003.1241] and HERA studies [Placakyte's talk, QCD@LHC]

