UPDATE FROM NNPDF

Nathan Hartland Nikhef / VU Amsterdam

DIS 2018 Kobe, 17/04/18





CURRENT GLOBAL FIT STATUS



Last release: NNPDF3.1 [1706.00428]

- Broad dataset inc. LHC measurements
- Statistically validated methodology

Theoretical developments for NNPDF3.1

- ► NNLO Results
 - $t\bar{t}$ Czakon, Heymes, Mitov [1511.00549], [1606.03350]
- $W/Z \ pT$ Boughezal *et al*, Gehrmann et al[1504.02131], [1507.02850]

Inc. Jets Currie et al [1310.3993] [1611.01460]

Fitted/intrinsic charm Ball et al [1510.02491], [1605.06515]

NEW UAIA IN NNPUF3.1 Measurement Measurement	Data taking Data Taking	Motivation Target PDF
Combined HERA inclusive data	Run I+II	quark singlet and gluon
D0 legacy W asymmetries	Run II	quark flavor separation
ATLAS inclusive W, Z rap 7 TeV	2011	strangeness
ATLAS inclusive jets 7 TeV	2011	large- <i>x</i> gluon
ATLAS low-mass Drell-Yan 7 TeV	2010+2011	small- <i>x</i> quarks
ATLAS Z pT 7,8 TeV	2011+2012	medium- <i>x</i> gluon and quarks
ATLAS and CMS tt differential 8 TeV	2012	large- <i>x</i> gluon
CMS Z (pT,y) 2D xsecs 8 TeV	2012	medium- <i>x</i> gluon and quarks
CMS Drell-Yan low+high mass 8 TeV	2012	small- <i>x</i> and large- <i>x</i> quarks
CMS W asymmetry 8 TeV	2012	quark flavor separation
CMS 2.76 TeV jets	2012	medium and large-x gluon
LHCb W,Z rapidity dists 7 TeV	2011	large-x quarks
LHCb W,Z rapidity dists 8 TeV	2012	large-x quarks

(Table thanks to J. Rojo)

NNPDF3.1 GLOBAL – PHENOMENOLOGY (GG)

Significant reduction in uncertainties across the kinematic range



NNPDF3.1 GLOBAL – PHENOMENOLOGY (QQ)

QQ Uncertainties decrease despite greater parametrisation freedom



Uncertainties often towards the percent level There are other sources of uncertainty to worry about

[1802.03398]

DETERMINING ALPHAS FROM A GLOBAL FIT



 σ_{θ}

 θ

 $\widehat{\theta}$

 σ_{α}

 σ_{θ}

Previous NNPDF determination Based on a scan of NNPDF2.1 [1110.2483]

Measure the χ^2 of best fit PDF parameters as a function of α_S

Neglects correlations with PDF fit parameters (θ) - important when experimental uncertainties are small

Ideally one should minimise PDFs and α_S simultaneously

NNPDF FITS ARE EXPENSIVE

Monte Carlo uncertainties

- PDFs are formed by *ensembles*:
 Fits to independent, equally likely samples (replicas) of the input dataset
- Each result requires 100/1000 statistically independent analysis runs



Fitting a large dataset only possible making use of pre-computed tables

$$\sigma_{pp \to X} = \sum_{i,j} \int_0^1 dx_1 dx_2 \ f_i(x_1, Q^2) \ f_j(x_2, Q^2) \ \sigma_{ij \to X} \left(x_1, x_2, Q^2 \right)$$
$$\sigma = \sum_{i,j}^{n_f} \sum_{\alpha,\beta}^{n_x} W_{ij\alpha\beta} \ f_i(x_\alpha, Q_0^2) f_j(x_\beta, Q_0^2)$$

THE CORRELATED REPLICA METHOD

How can we take into account PDF/ α_S correlations in a 'MC' way?



- For each data sample (replica) perform a scan in α_S
- Each replica has a preferred value of α_S (the minimum of each parabola)
- These preferred values form a MC distribution





 $\alpha_s^{\text{NNLO}}(m_Z) = 0.1185 \pm 0.0005^{\text{exp}} \pm 0.0001^{\text{meth}} \pm 0.0011^{\text{th}} = 0.1185 \pm 0.0012 \ (1\%)$



THEORETICAL UNCERTAINTIES IN PDF FITS

PDF uncertainties often represent only experimental and procedural factors Parametric uncertainties due to e.g strong coupling straightforward to handle

A more difficult question

How can we estimate uncertainties due to missing higher orders?



At NLO - measure difference between fits at different perturbative orders

NNPDF3.1 DIS-only NLO, Q = 100 GeV



THEORETICAL UNCERTAINTIES IN PDF FITS

Can we build a theoretical 'covariance matrix'?

For a set of predictions for an observable at a central scale

 $\mathcal{O}_i(\mu_R, \mu_F), \quad (0 < i < N_{\text{dat}})$

Consider shifts due to three-point scale variations

$$\Delta_i^+ = \mathcal{O}_i(\mu_R, \mu_F) - \mathcal{O}_i(2\mu_R, 2\mu_F),$$

$$\Delta_i^- = \mathcal{O}_i(\mu_R, \mu_F) - \mathcal{O}_i(\mu_R/2, \mu_F/2),$$

One can then construct a theory 'covariance'

$$\operatorname{Cov}_{\operatorname{Th}}[\mathcal{O}_i, \mathcal{O}_j] = \Delta_i^+ \Delta_j^+ - \Delta_i^- \Delta_j^-$$

To be added to the experimental matrix

 $\operatorname{Cov}_{\operatorname{Total}} = \operatorname{Cov}_{\operatorname{Th}}[\mathcal{O}_i, \mathcal{O}_j] + \operatorname{Cov}_{\operatorname{Ex}}[\mathcal{O}_i, \mathcal{O}_j]$

THEORY UNCERTAINTIFS



Very preliminary Combined covariance for DIS sets

In progress How do these matrices influence the fit? How do they compare to scale-varied fits?

Theory covariance matrix



Experiment + theory covariance matrix



PATH TO NNPDF 4.0 - NEW DATA

Datasets/processes under consideration for NNPDF 4.0

DIS	13TeV Standard Candles	New processes
HERA c. F_2^c, F_2^b	Inclusive jets	2/3-jets
NOMAD $\mu\mu$	W/Z production	Prompt photon
CHORUS $\mu\mu$	High-Mass DY	Single top
JLAB S.Fs	top-pair production	D-meson

Impact of ATLAS prompt photon studied in J. Campbell et al [1802.03021]



PATH TO NNPDF 4.0 - METHODOLOGY

As our dataset gets larger, it's important to assess our methodology

A large dataset with a flexible parametrisation presents a complex optimisation problem

Our parametrisation is flexible to minimise bias, but can it be made more *efficient*?





Fits are computationally expensive *Can modern optimisation tools help?*

- Evolutionary Strategies
- Analytical gradients

SUMMARY

NNPDF3.1: released June 2017

- ► NNPDF3.1sx October 2017
- ► NNPDF3.1luxQED December 2017
- ► NNPDF α_S January 2018



The path to NNPDF4.0

► New data

Almost 40 new datasets due to be investigated

► Theory uncertainties

How can we best represent PDF uncertainties due to MHO corrections?

Methodology

How can our fitting procedure keep pace with the dataset?

BACKUPS

INTRINSIC CHARM

The charm PDF is a borderline perturbative object
 Most PDF fits assume that charm is generated perturbatively by evolution
 Such an assumption can lead to a disproportionate influence of the charm mass

Relaxing this assumption by fitting charm can stabilise results



NNPDF FITS ARE EXPENSIVE

Several procedural factors conspire to make NNPDF fits particularly demanding

Monte Carlo uncertainties

PDFs are formed by ensembles: Each result requires 100/1000 statistically independent analysis runs

Neural Network parametrisation

 Standard gradient descent is difficult: Minimisation by Genetic Algorithm
 typically 50,000 generations



Fitting a large dataset only possible making use of pre-computed tables

$$\sigma_{pp \to X} = \sum_{i,j} \int_0^1 dx_1 dx_2 \ f_i(x_1, Q^2) \ f_j(x_2, Q^2) \ \sigma_{ij \to X} \left(x_1, x_2, Q^2 \right)$$
$$\longrightarrow \sigma = \sum_{i,j}^{n_f} \sum_{\alpha,\beta}^{n_x} W_{ij\alpha\beta} \ f_i(x_\alpha, Q_0^2) f_j(x_\beta, Q_0^2)$$

NNPDF FITS ARE EXPENSIVE

Several procedural factors conspire to make NNPDF fits particularly demanding

Monte Carlo uncertainties

PDFs are formed by ensembles: Each result requires 100/1000 statistically independent analysis runs

Neural Network parametrisation

 Standard gradient descent is difficult: Minimisation by Genetic Algorithm
 typically 50,000 generations





NNPDF3.1 GLOBAL FIT RESULTS

LHC 13 TeV, NNLO



Fit Quality χ^2 3.1 NNLO 3.0 NLO **HERA** 1.16 1.14 **ATLAS** 1.09 1.37 CMS 1.06 1.20 LHCb 1.47 1.61 TOTAL 1.148 1.168 (FC) TOTAL 1.187 1.197 (PC)

NNPDF3.1 GLOBAL FIT RESULTS – DATA VS METHODOLOGY

NNLO, Q = 100 GeV



NNPDF3.1 GLOBAL FIT RESULTS – DATA VS METHODOLOGY

NNLO, Q = 100 GeV



THE STRANGENESS PUZZLE





Tension in strangeness between global fits and xFitter persists in NN3.1

	NNPDF3.1 Global	NNPDF3.1 Collider
ATLAS 2011 W/Z	2.14	1.55
ATLAS 2010 W/Z	0.96	0.92
NuTeV dimuon	0.82	26.5

Driven by disagreement between collider data and neutrino DIS

THE STRANGENESS PUZZLE

 $R_s(x,Q^2) = \left[s(x,Q^2) + \bar{s}(x,Q^2)\right] / \left[\bar{u}(x,Q^2) + \bar{d}(x,Q^2)\right]$

NNLO, Q=100 GeV



INCLUSIVE JET DATA AT NNLO

While the full NNLO calculation for inclusive jet production has been finalised, exact K-factors for several of the jet datasets were not available at time of publication

Therefore for NNLO fits, the jet data was included at NLO accuracy, but with an additional uncertainty determined by NLO scale variation



Reliability verified by comparison against fit with available NNLO corrections