Parton Distribution Functions - lecture 2 -

Alberto Accardi Hampton U. and Jefferson Lab

CTEQ summer school 2013





Lecture 1 - recap

data theory • DIS: p, d pQCD at NLO • p+p(pbar) \rightarrow l⁺l⁻, W[±], Z Factorization & universality • Large-*x*, low-*Q*², nuclear corr. • p+p(pbar) \rightarrow jets, γ +jet fits • Parametrize PDF at Q_0 , evolve to Q• Minimize χ^2 **PDFs F**₂(n) W, Z / W', Z', Higgs (or any other "hard" observable) accardi@jlab.org 2 cieu 2013 – Leclure 2

Lecture 2 -PDF uncertainties and Applications

PDF uncertainties

- Statistical, in detail
- Theoretical, by examples

Comparison of PDFs

- PDFs, parton luminosities

🗆 LHC

- Standard candles, Higgs and t-tbar
- PDF constraints from LHC data

Large-x connections

- Nuclear uncertainty, a new parametrization bias
- d/u ratio, large mass searches at LHC
- Using proton targets to constrain nuclear phsyics

PDF uncertainties



PDF uncertainties

Experimental ("PDF errors"):

- uncertainties in measured data propagate into the fitted PDFs
- can be quantified adapting statistical methods: "PDF error bands"
- Need to interpreted with care

Theoretical

- Several sources, cannot be quantified easily
 - Choice of data sets, kinematic cuts
 - Choice of χ^2 function
 - Parametrization bias
 - Truncation of pQCD series, scale choice, alpha strong
 - Heavy-quark scheme
 - Higher-twist, target mass effects
 - Nuclear corrections

(a few examples will be discussed later)

(We will see this in some detail)



Hessian method

- PDF parameters denoted by $\{a_{\mu}\}, \mu = 1, \dots, d$
- As a byproduct of the fitting process, one obtains the Hessian $H_{\mu\nu}$

$$H_{\mu\nu} \equiv \frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_\mu \, \partial a_\nu}$$

which is evaluated at the minimum of χ^2 .

• To estimate the error on some observable X(a), taking into account only the experimental errors which entered into the calculation of χ^2 one uses the "Master Formula"

$$(\Delta X)^2 = \prod_{\mu,\nu} \sum_{\mu,\nu} \frac{\partial X}{\partial a_{\mu}} (H^{-1})_{\mu\nu} \frac{\partial X}{\partial a_{\nu}}$$

Representation of PDF uncertainties

Error PDFs

 $f^{(0)}$

- Diagonalize the Hessian in parameter space
- Parameter eigenvectors define orthogonal directions in the functional PDF space
- A number of PDF error sets is provided:

Central value (best fit)

 $f^{(i)}$ $i = 1, N_{par}$ Error PDFs (one for each orthogonal direction)

Observables and their uncertainty are then calculated as *

$$X_{0} = X[f^{(0)}]$$

$$\Delta X^{2} = T \sum_{i}^{N_{par}} \left(X[f^{(i)}] - X_{0} \right)^{2}$$

* NOTE: in CT10, tolerance is absorbed inside the error PDFs CTEQ 2013 – Lecture 2 7



Representation of PDF uncertainties

Monte-Carlo error set

- Construct a number N of PDF replicas by generating a multi-Gaussian distribution of parameter values:
 - One Gaussian per orthogonal direction
 - Centered on best fit parameter, width provided by the Hessian matrix
- Observables and their uncertainty are then calculated as

$$X_{0} = \frac{1}{N} \sum_{i=1}^{N} X[f^{(i)}]$$
$$\Delta X^{2} = \frac{1}{N^{2}} \sum_{i=1}^{N} \left(X[f(i)] - X_{0}] \right)^{2}$$

– Previous results recovered as $N \rightarrow \infty$; N=50-100 suitable

accardi@jlab.org

Tolerance

 \Box Open a textbook, *T*= $\Delta\chi$ =1 means 67% confidence level

But Hessian method works only if

- all data sets are statistically compatible
- Exp. errors are Gaussian...
- ...and have not been underestimated
 (e.g., by neglect of a source of systematics)

Correct this by a larger tolerance factor so that most data (90%, 67% of them) fall inside the PDF error band

- Fixed tolerance (T=10 for CTEQ6.1 and CJ12, T=7 for MRST)
- "Dynamical tolerance" for each fit (MSTW08, CT10)

Monte Carlo method

 \Box Generate N_{rep} replicas of the chosen data set

- In each replica, randomize central data point within quoted errors
- Make a fit for each replica

Obtain PDF errors from statistical analysis of all fit results

$$X_{0} = \frac{1}{N} \sum_{1}^{N_{rep}} X[f^{(i)}]$$
$$\Delta X^{2} = \frac{1}{N_{rep}^{2}} \sum_{1}^{N_{rep}} \left(X[f(i)] - X_{0}] \right)^{2}$$

This method originally adopted by the NNPDF collaboration (Hessian not available in their case) but not limited to neural network based fits

Advantages of a Monte-Carlo representation of errors

"Monte-Carlo" error PDF sets (either by NNPDF or conventional Hessian methods) can be used in several ways

Bayseian reweighting:

- New data sets can be included by "reweighting" according to the previously obtained Monte Carlo PDFs
- No need to perform new fits if N large enough

Combination of PDFs from different groups

- E.g., for equal a-priori weighting, take same number of Monte-Carlo PDFs from each group
- Calculate X and ΔX as before

Speed up uncertainty calculations in Monte-Carlo generators

Example: combination of CT10, NNPDF2.3, MSTW08



Open markers: usual best-fit and 68% C.L. Hessian uncertainty. Closed markers: average and s.d. over random predictions.

Open markers: usual best-fit and 68% C.L. Hessian uncertainty. Closed markers: average and s.d. over random predictions.

NNPDF2.3

 $a_{e}(M^{2}) = 0.1190$

NNPDF2.3

 $a_{a}(M^{2}) = 0.1190$

accardi@jlab.org

Lagrange multiplier method

Given an observable X, minimize a new function for fixed values of Lagrange multiplier λ

 $\Psi(\lambda, \{a\}) = \chi_g^2(\{a\}) + \lambda (X(\{a\}) - X_0)$

– Obtain a new set of parameters, A_{min} , and the pair $\{\chi_g^2(\lambda), X(\lambda)\}$





Comparison of PDFs



Some differences between PDF groups

	MSTW08	CT10	NNPDF2.3	HERAPDF1.5
PDFs	7	6	7	5
Params.	20 + 8	25	259	14
Heavy Q.	TR	ACOT	FONLL	TR
Statistics	Hess.+DT	Hess.+DT	MC	Hess.+par+modl
$W_{\mathrm{m}in}^2$	high	high	high	high
HT, TMC				
Nucl. corr.	explored		explored	
large-x d				

	ABM11	JR09	CJ12
PDFs	6	5	5
Params.	24	12	19 + 3
Heavy Q.	FFN	FFN	ZM
Statistics	Hess.	Hess.+T	Hess.+T
W^2_{min}	low	high	low
HT, TMC	\checkmark		\checkmark
Nucl. corr.	\checkmark		\checkmark
large-x d			\checkmark

accardi@jlab.org

Non-singlet PDFs

u, well determined, some

 \Box Uncertainties in d_v at medium-large x

- See later for details

 $\Box \overline{d} - \overline{u}$ well determined by DY data at x = 0.05 - 0.25

- otherwise unconstrained

- □ s − s partially determined by EMC v+A dimuon production
 - But large experimental, nuclear uncertainties
 - CJ12 omits these data altogether



Singlet, gluon PDFs

 $\Box u + \overline{u}$ and $d + \overline{d}$ dominant at large x

 $\Box s + \overline{s}$ becomes comparable at small x

Gluons dominate by far at small x



accardi@jlab.org

Parton luminosities: q - qbar

- Fairly good agreement between MSTW08, CT10, NNPDF2.3
 - "standard candles" under control
- But at variance with other sets... Due to choice of data?
 - ABM and JR have DIS+DY only
 - HERAPDF have HERA DIS only
- Uncertainties and spread of central values grow at large x
 - Limitation for large mass, forward rapidity observables



 $L^{ij} \equiv q_{i/A}(x_1)\bar{q}_{j/B}(x_2) + \{A \leftrightarrow B\}$

Parton luminosities: g - g

- More variation
 - In particular in the region
 relevant to Higgs production
- HERAPDF has larger PDF errors
 - Due to use of HERA-only data
- \square Note: α_s varies with each PDF
 - Strongly correlated with gluons









W and Z cross sections - a standard candle at the LHC

Partonic cross section for inclusive W,Z production well-known

use to monitor parton luminosity reduce systematic in other cross sects.

$$\sigma_X(s, M_X^2) = \int_{\tau}^1 \frac{dx}{x} \mathcal{L}(x) C\left(\frac{\tau}{x}, \alpha_s(M_X^2)\right)$$
$$\tau \equiv \frac{M_X^2}{z}$$
$$\mathcal{L}(x) \equiv \int_x^1 \frac{dz}{z} f_{a/h_1}(z, M_X^2) f_{b/h_2}\left(\frac{x}{z}, M_X^2\right)$$

NNLO W and Z cross sections at the LHC (\sqrt{s} = 7 TeV)

NNLO W⁺ and W⁻ cross sections at the LHC ($\sqrt{s} = 7$ TeV)



Higgs and t-tbar

Sensitivity to gluons, alpha strong



Both g-g and q-qbar lumi underestimated (also CJ12 for q-qbar) – What are we missing?

PDF sets with LHC data

J.Rojo, DIS 2013 major improvement since DIS12 is use of LHC data on jets and W,Z production by the PDF groups

NNPDF2.3 is only publicly available PDF set that includes constrains from LHC jet and W,Z data, and other groups have presented preliminary updates quantifying the impact of LHC measurements



accardi@jlab.org

Experimental constraints

J.Rojo, DIS 2013 On top of traditional processes, like jets and W, Z production, a wide range of new processes that provide PDF information is now available at the LHC (see Voica's talk)

Top quarks: constrain large-x gluon





W+charm: sensitivity to strangeness



Isolated photons: complementary probe of the gluon, same x-range as for gg Higgs production



24

Strange quark from W, Z measurements at ATLAS

Strange guark is not so well constrained:

Radescu

DIS 2013

- Neutrino dimuon data favours suppressed strange
- At LHC, Z cross sections together with y₇ shape may provide a constraint on s-quark density and it can be cross checked by W+charm data.
 - The results for NNLO fits to inclusive W.Z differential data with free and fixed s:
 - \Rightarrow For W+ and W- there is little difference, helps to fix the normalisation.
 - \rightarrow For Z, the cross section is increased and the shape is modified.

epWZ free s

ATLAS

rs

1.2



Nuclear suppression of charm?

accardi@jlab.org

Q² = 1.9 GeV², x=0.023

experimental uncertainty

0.2

0.6

 $r_s = 1.00 \pm 0.20 \exp \pm 0.07 \mod_{-0.15}^{+0.10} \Pr_{-0.07}^{+0.06} \alpha_s \pm 0.08 \text{th.}$

0.4

0.8

ABKM09

NNPDF2.1 MSTW08

CT10 (NLO)

-0.2

total uncertainty





Why large x ?

- Large (experimental) uncertainties in Parton Distribution Functions (PDFs)
- Precise PDFs at large x are needed, e.g.,
 - Non-perturbative nucleon structure:
 - d/u, $\Delta u/u$, $\Delta d/d$ at $x \rightarrow 1$
 - at LHC, Tevatron
 - New physics as large p_{τ} excess
 - High mass searches
 - Forward physics
 - At RHIC:
 - Polarized gluons at the smallest x
 - Neutrino oscillations, ...



Large-x, small-Q² corrections



Deuteron corrections

No free neutron! Best proxy: Deuteron

- Parton distributions (to be fitted)
- nuclear wave function (AV18, CD-Bonn, WJC1, ...)
- Off-shell nucleon modification (model dependent)

Theoretical uncertainty



Effect of theory corrections in a nutshell

PDFs stable with respect to low W cut

If TMC included and residual power correction ("HT") fitted

Accardi et al., PRD 81 (2010)

New *d*-quark parametrization

 $d'(x) = d(x) + \alpha x^{\beta} u(x)$

- Allows d/u to be non-zero at x = 1
 (as required in non-perturbative models)
- Produces *dramatic increase in d PDF in x* \rightarrow 1 *limit*

Sensitivity to nuclear corrections

- *d*-quark at *x* > 0.5 almost fully correlated to nuclear model model:
 Very large theoretical uncertainty at large x
- Modest, non negligible impact also at 0.2 < x < 0.5

Accardi et al. PRD81 (2010)

Ball et al. ArXiv:1303.1189 (2013)



CJ12 fits: nuclear and PDF uncertainty



Large overall reduction in uncertainty with relaxed cuts

```
accardi@jlab.org
```

Applications: *d/u* ratio



Applications: new physics at LHC

Accardi et al., PRD 84 (2011) 014008

New physics signal require accurate determination of QCD background

Uncertainties in large-x PDFs could affect interpretation of experiments searching for new particles



Differential parton luminosities

accardi@jlab.org

Applications: large mass searches at LHC

Brady, Accardi, Melnitchouk, Owens, JHEP 1206 (2012) 019

- New physics signal require accurate determination of QCD background
- Uncertainties in large-x PDFs could affect interpretation of experiments searching for new particles



Example: W' and Z' total cross sections

Large-x parametrization bias



Dramatic increase in d quark with more flexible parametrization

 \Box Standard (old) d-quark: either $d/u \rightarrow 0$ or $d/u \rightarrow \infty$

- Large bias, neglected in all other fits

accardi@jlab.org

Large-x parametrization bias



accardi@jlab.org

CJ12: old vs. new d quark



Standard d-quark too stiff at x > 0.6

- Underestimates central value *and nuclear uncertainty*
- Full unbiasing could be obtained in a NN analysis with low W cuts

Constraining the nuclear uncertainty

DIS data minimally sensitive to nuclear corrections

- DIS with slow spectator proton (BONUS)
 - Quasi-free neutrons
- DIS with fast spectator (DeepX)
 - Off-shell neutrons
- ³He/³H ratios

Data on free (anti)protons, sensitive to d

- e+p: parity-violating DIS HERA (e^+ vs. e^-), EIC, LHeC
- -v+p, v+p (no experiment in sight)
- *p+p, p+p*-at large positive rapidity
 - W charge asymmetry, Z rapidity distribution

Cross-check data

- *p+d* at large <u>negative</u> rapidity dileptons; *W, Z*
 - Sensitive to nuclear corrections, cross-checks *e*+*d*

Jlab12, EIC

Tevatron: CDF, D0(?) LHCb(?) RHIC AFTER@LHC

RHIC??

AFTER@LHC

accardi@jlab.org

Use protons to study nuclei (!)

Brady, Accardi, Melnitchouk, Owens, JHEP 1206 (2012) 019



Needs to be corroborated:

See also MMSTWW, EPJ C73 (2013)

- -W, Z at RHIC, Z (and W?) at LHC, W at DØ (??)
- PVDIS at JLab 12, CC @ EIC

accardi@jlab.org

Use protons to study nuclei (!)

Brady, Accardi, Melnitchouk, Owens, JHEP 1206 (2012) 019



accardi@jlab.org

Use protons to study nuclei (!)

Brady, Accardi, Melnitchouk, Owens, JHEP 1206 (2012) 019



W charge asymmetry at Tevatron

Brady, Accardi, Melnitchouk, Owens, JHEP 1206 (2012) 019



Too little large-x sensitivity in lepton asymmetry:

– need reconstructed W

W charge asymmetry at LHC

Brady, Accardi, Melnitchouk, Owens, JHEP 1206 (2012) 019



Would be nice to reconstruct W at LHCb

- Does not seem feasible (too many holes in detector)
- What about RHIC, AFTER@LHC?

Z rapidity distribution

Brady, Accardi, Melnitchouk, Owens, JHEP 1206 (2012) 019



Direct Z reconstruction is unambiguous in principle, but:

- Needs better than 5-10% precision at large rapidity
- Experimentally achievable?
 - At LHCb? **RHIC?** AFTER@LHC?
 - Was full data set used at Tevatron?

At RHIC: p+p collisions

W reconstruction: an almost unique RHIC measurement

- Cross checks CDF measurements (tension with lepton asymmetry)
- − Energy systematics: $500 \rightarrow 1960 \text{ GeV}$
- Large x > 0.5 needs 1.1 < y < 1.9</p>
- Hard/impossible at LHCb (large x)
- Not a priority at ATLAS, CMS (small x)
 - Lepton asymmetry good substitute in this case

Z reconstruction much easier

- Similar motivations, kinematics
- − Energy systematics: $500 \rightarrow 1960 \rightarrow 7000$

Possibly the hardest

experimental requirement

Final thoughts

- **Data** extended kinematic range, precision needed
 - LHC, JLab12, E906/J-PARC, RHIC, EIC/LHeC
- Large-x: use nuclei to find new particles; use protons to study nuclear physics

Experimental PDF uncertainties

- use MC error representation, compare to Hessian + tolerance
- Theoretical uncertainties will be crucial
 - Need to be estimated, provided to users not explored enough so far

Parametrization biases

- Use extended parametrizations; how to determine an optimal fit?

Perturbative order

- NNLO to be calculated for all relevant processes where to stop?
- Do we want/need PDFs with resummation?
- Heavy quarks schemes to be extended beyond DIS
 - Intrinsic charm (hadron structure, new Higgs channels at large y)
- Electroweak corrections to PDFs as large as NNLO need to include

Lecture 2 - recap

Global PDF fits as a tool for particle physics

- integrate across hadronic physics from JLab to the LHC
- connect with rest of subatomic physics



Appendices

A1 Impact of a new accelerator



Questions

- What are the requirements in terms of energy, luminosity?
- What physics do we expect to learn?
- "Is it worthwhile building that accelerator?"

For example:

 Is a DIS cross section measurement at the EIC going to improve the PDF measurements?

This we can answer with a global fit:

- Generate pseudo-data
- Include them in a global fit
- Compare with old result

e+p collisions – using CTEQ-JLab fits [Accardi, Ent, Keppel]

- Pseudo data:



accardi@jlab.org

e+A collisions – using NNPDF2.0 fits [Accardi, Guzey, Rojo]

QCD fit to EIC pseudo-data for <u>Pb only</u>



- e+A collisions using NNPDF2.0 fits [Accardi, Guzey, Rojo]
 - With only 1 nucleus target, impact comparable to present day world data:



A2 Other large x observables



At RHIC: p+d collisions

W,Z, dileptons at <u>negative</u> rapidity (or d+p collisions)

- Large x_A in nucleus
- Large Q2 ==> no power corrections
- Cross checks nuclear corrections in DIS assuming universality of "nuclear PDFs":



$$q_d(x,Q^2) = \int_{x_B}^A dy \, \mathcal{S}_A(y,\gamma=1) q(x,Q^2) \left(1 + \frac{\delta^{off} q(x)}{q(x)}\right)$$

(see also Kamano, Lee, PRD86 (2012) 094037 for a DY convolution formula analogous to the DIS one)

At RHIC: ...and more...

Sea asymmetry at mid-rapidity

Bourrelly, Soffer, NPB 423 (1994) 329 Peng, Jensen, PLB 354 (1995) 460

- Measures \bar{u}/\bar{d} at x = 0.06-0.45 (for -1 < y < 1)
- Useful to compare p+p vs. p+n
- Play W vs. Z production

Forward physics

- What to do with what you will measure forward?
 - Photons
 - Pions
 - ...

Spectator tagging at Jlab: quasi-free neutrons



γ. **N** X **D** P

accardi@jlab.org

Spectator tagging at JLab12

- Neutron off-shellness depends on on spectator momentum:
 - Slow: nearly on-shell (BONUS12)
 - Fast: more and more off-shell (LAD)





Spectator tagging at EIC: even better!

- measure neutron F, in D target
 - flavor separation

- measure proton F₂ in D target
 - Unique at colliders
 - Compare off-shell to free proton

proton, neutron in light nuclei

accardi@jlab.org

embedding in nuclear matter (a piece of the EMC puzzle)

CTEQ 2013 – Lecture 2

⁶*He*

 \mathcal{N}

³H e

 ^{7}Li

n

⁴H e

⁴H e

60

A3 CJ12 vs. others



CJ12 vs. others

Owens, Accardi, Melnitchouk, arXiv:1212.1702



accardi@jlab.org

CTEQ 2013 – Lecture 2

62

CJ12 vs. others

Owens, Accardi, Melnitchouk, arXiv:1212.1702

