An overview

Pavel Nadolsky

Southern Methodist University

with CTEQ-TEA working group, Karol Kovarik, Fred Olness, and Dave Soper

2nd workshop on Parton distributions and lattice calculations

MSU KBS, Gull Lake, Michigan

2019-09-28

Q: what should the “Overview talk” be about?

Q: what should the “Overview talk” be about?

1. Show some pictures from the 1st PDFLattice workshop

PDFLattice workshop

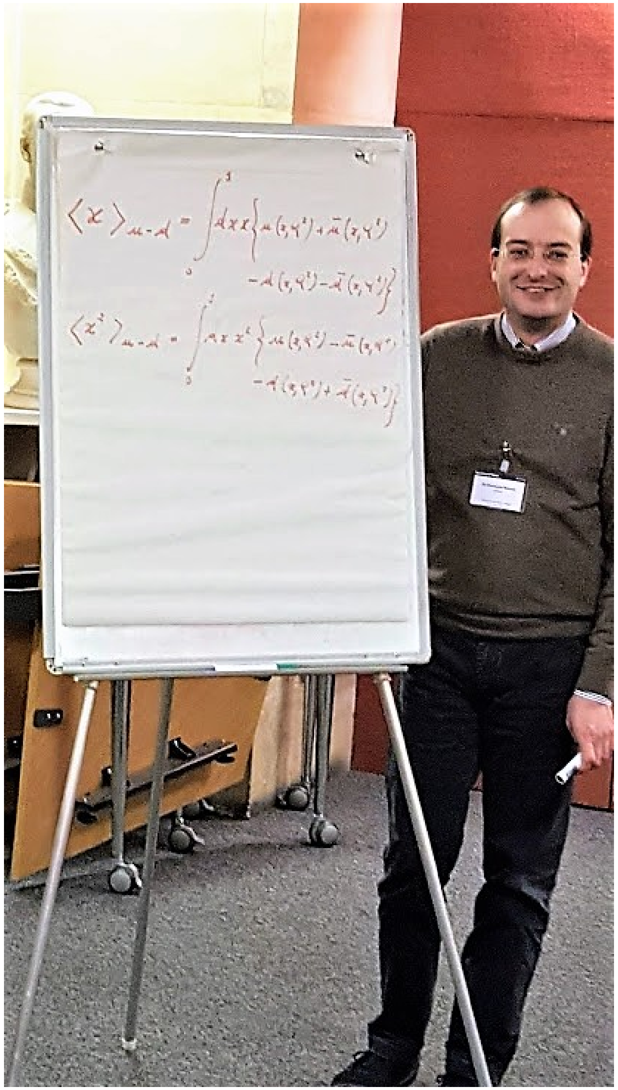
Balliol College, Oxford, March 2017

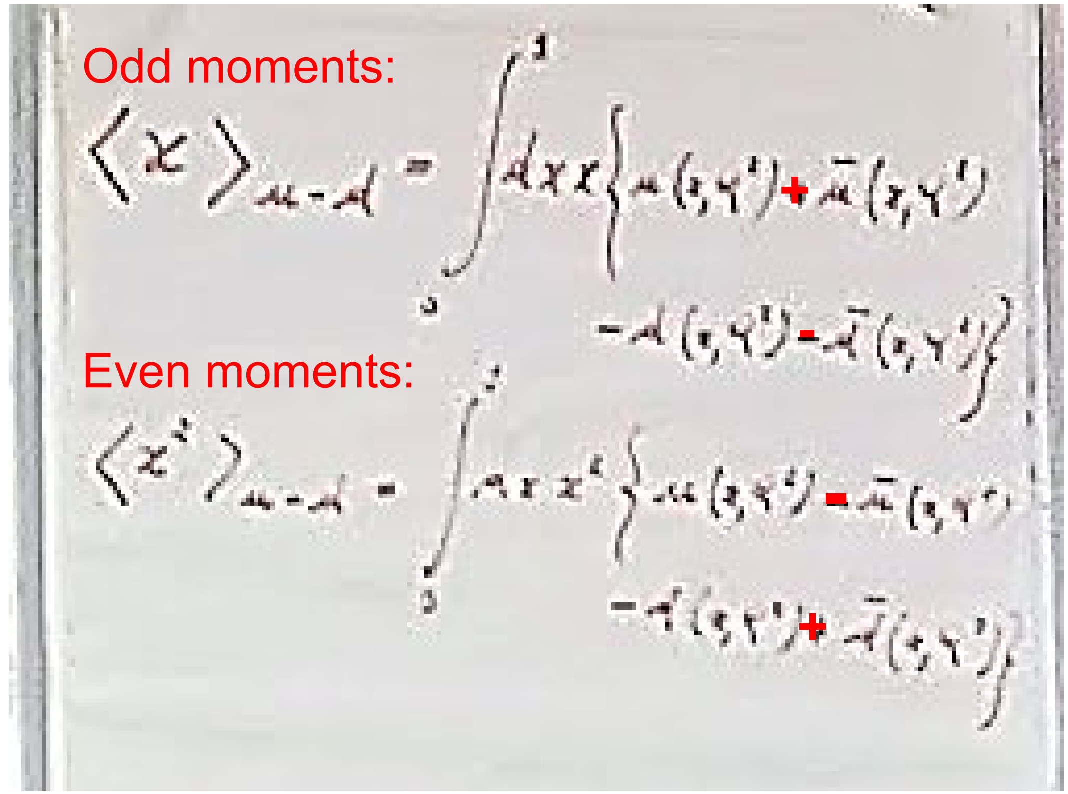
PDFLattice workshop

Balliol College, Oxford, March 2017

PDFLattice workshop

Balliol College, Oxford, March 2017





Q: what should the "Overview talk" be about?

2. Rethinking the global QCD analysis and lattice QCD calculations for the next decade

- common physics goals
- shared resources
- agreed-upon practices
- benchmarking studies

⇒ **A concrete proposal**

Frequently asked questions

1. What is the meaning (definition) of my PDF?
2. Do I measure what I claim to be measuring?
3. How large are higher-order contributions?
4. How should the PDF uncertainties be estimated?
5. How much am I biased by my parametrization of PDFs?
6. Are there hidden PDF uncertainties?
7. Are there **reliable** shortcuts to get the complete answer? [Error PDF reweighting, vector techniques, …]

These questions are best understood for unpolarized collinear PDFs

They must be addressed when trying to measure or compute the nuclear, polarized, meson, TMD, GPD, quasi-, pseudo- PDFs

$$f_{a/h}(x, Q)$$

The status of determination and usage of unpolarized PDFs is addressed in several recent reviews,

1. S. Forte, G. Watt, Ann. Rev. Nucl. Part. Sci. 63, 291 (2013)
2. J. Gao, L. Harland-Lang, J. Rojo, Phys. Rept. 742, 1 (2018)

3. **K. Kovarik, PN, D. Soper, Hadron structure in high-energy collisions, arXiv:1905.06957**

and group studies:

4. PDF4LHC working group, J. Phys. G 43, 023001 (2016)
5. A. Accardi et al., Eur. Phys. J. C 76, 471 (2016).

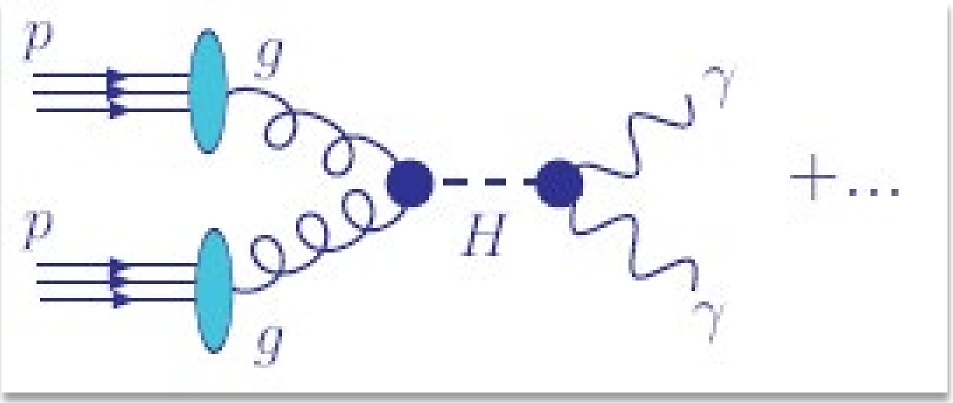
Frontiers of the PDF analysis

Significant advances on all frontiers will be necessary to meet the targets of the HL-LHC program by 2030

Exceptional opportunities to learn about the 3D hadron structure at a future EIC/LHeC

- precision (N)NNLO studies of (un)polarized PDFs at an EIC will also be of a great value for the high-lumi LHC physics program

Parton distributions describe long-distance dynamics in high-energy collisions



$$\sigma_{pp\to H\to\gamma\gamma X}(Q) = \sum_{a,b=g,q,\bar{q}} \int_0^1 d\xi_a \int_0^1 d\xi_b \hat{\sigma}_{ab\to H\to\gamma\gamma}\left(\frac{x_a}{\xi_a}, \frac{x_b}{\xi_b}, \frac{Q}{\mu_R}, \frac{Q}{\mu_F}; \alpha_s(\mu_R)\right)$$

$$\times f_a(\xi_a, \mu_F) f_b(\xi_b, \mu_F) + O\left(\frac{\Lambda_{QCD}^2}{Q^2}\right)$$

$\hat{\sigma}$ is the hard cross section

$f_a(x, \mu_F)$ is the distribution for parton a with momentum fraction x, at scale μ_F

Operator definition of PDFs; evolution equations

To all orders in α_s, PDFs are **defined** as matrix elements of certain correlator functions:

$$f_{q/p}(x,\mu) = \frac{1}{4\pi}\int_{-\infty}^{\infty} dy^- e^{iy^-p^+} \langle p \left| \overline{\psi}_q(0,y^-,\vec{0}_T)\gamma^+\psi_q(0,0,\vec{0}_T) \right| p \rangle, \text{ etc.}$$

The exact form of $f_{a/p}$ is not known; but its μ dependence is described by **D**okshitzer-**G**ribov-**L**ipatov-**A**ltarelli-**P**arisi (**DGLAP**) equations

$$\mu \frac{df_{i/p}(x,\mu)}{d\mu} = \sum_{j=g,u,\bar{u},d,\bar{d},\dots} \int_x^1 \frac{dy}{y} P_{i/j}\left(\frac{x}{y}, \alpha_s(\mu)\right) f_{j/p}(y,\mu)$$

$P_{i/j}(x,\mu)$ are known up to N3LO

⇒ Starting from parametrizations of $f_{i/p}(x,\mu_0)$ at $\mu_0 \approx 1\ GeV$, DGLAP equations predict $f_{i/p}(x,\mu)$ at $\mu \geq \mu_0$

Perturbative QCD loop revolution

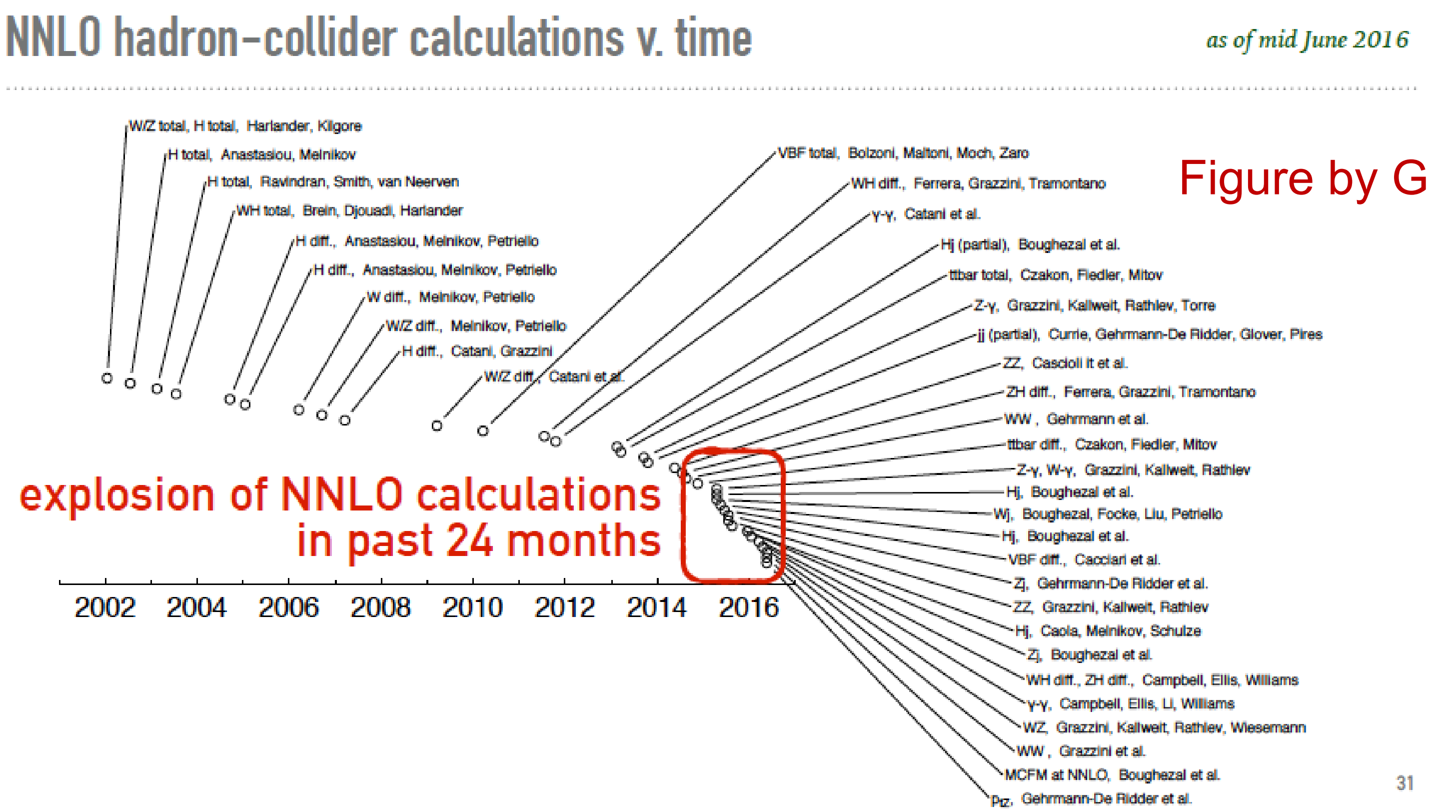
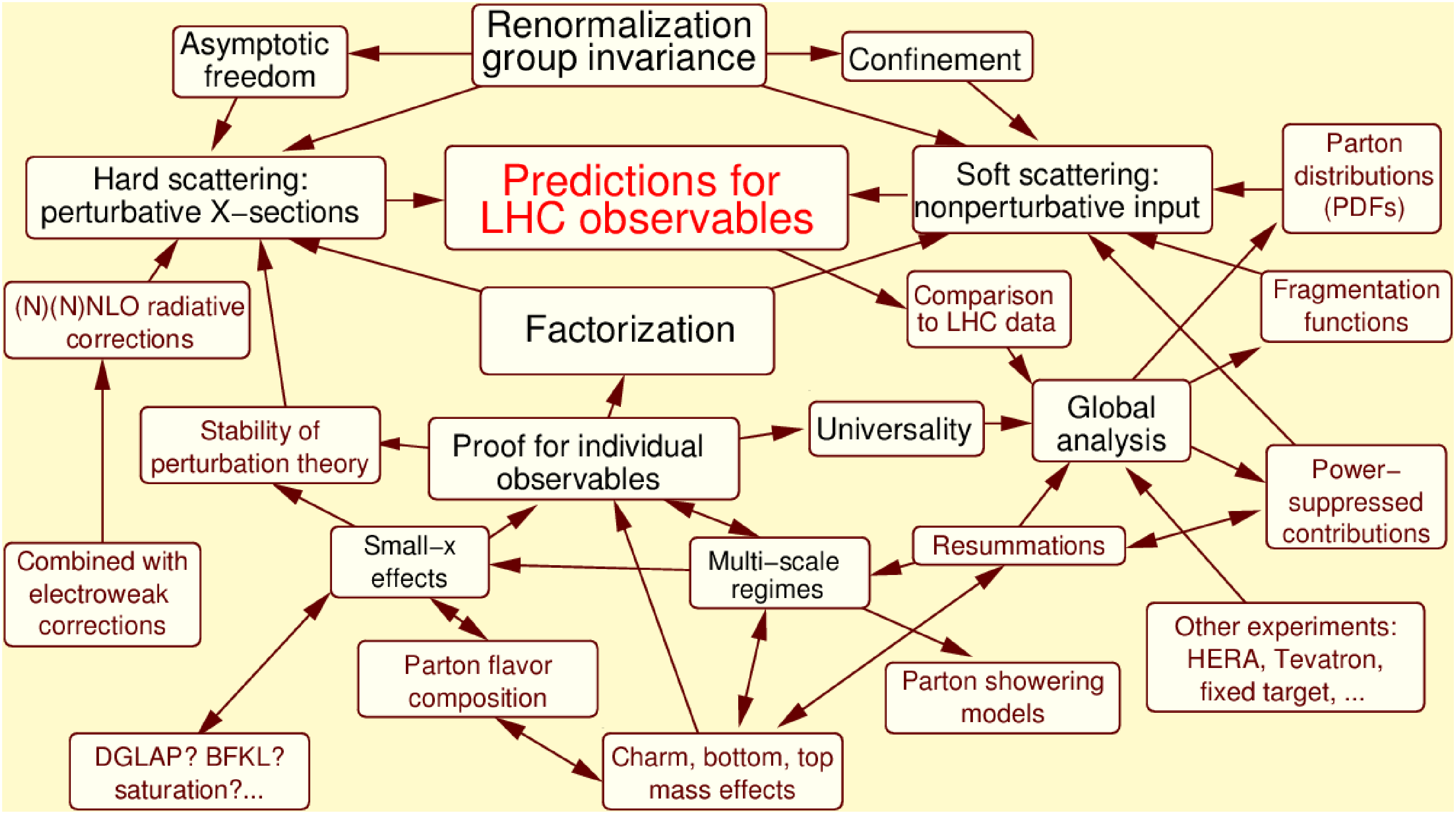


Figure by G. Salam

Since 2005, generalized unitarity and related methods dramatically advanced the computations of **perturbative** NLO/NNLO/N3LO hard cross sections $\hat{\sigma}$.

To make use of it, accuracy of PDFs $f_{a/p}(x,\mu)$ must keep up

At the (N)NNLO accuracy level, multiple aspects affect the PDF behavior



Classes of PDFs

General-purpose

For (N)NLO calculations with $N_f \leq 5$ active quark flavors

From several groups:
ABMP'16
CTEQ-Jlab (CJ'2015)
HERA2.0
CT14 (→ CT18)
MMHT'14 (→ 19)
NNPDF3.1 (→4.0)

Specialized

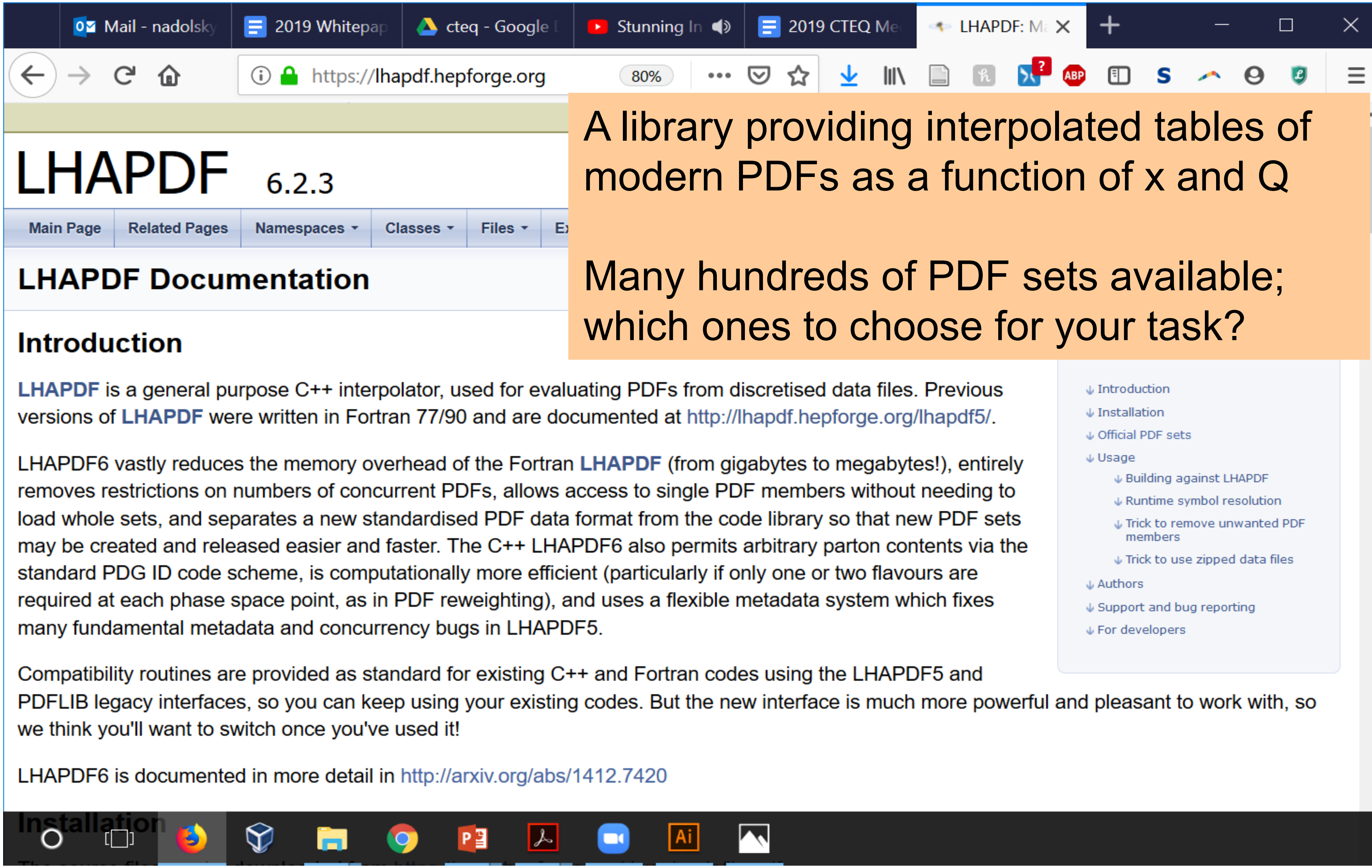
For instance, for CT14:
CT14 LO
CT14 $N_f = 3$, 4, 6
CT14 HERA2 **[arXiv:1609.07968]**
CT14 Intrinsic charm **[1707.00065]**
CT14 QCD+QED **[1509.02905]**
CT14 Monte-Carlo **[1607.06066]**

ATLAS & CMS exploratory

Combined **[1509.03865]**

PDF4LHC'15=CT14+MMHT'14+NNPDF3.0

Les Houches Accord PDF repository



A library providing interpolated tables of modern PDFs as a function of x and Q

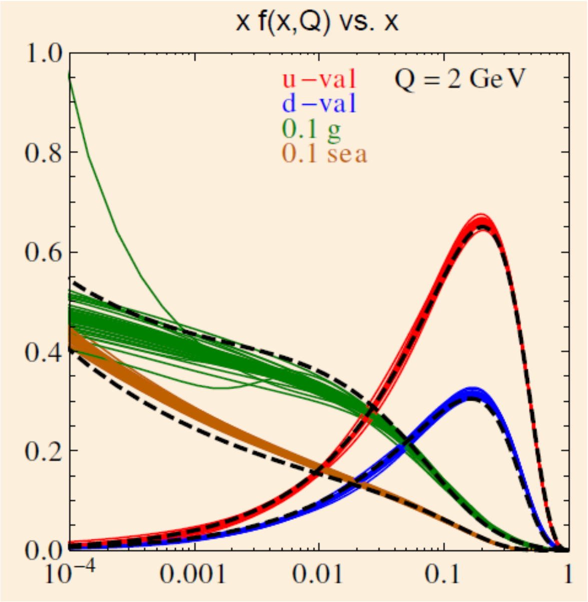
Many hundreds of PDF sets available; which ones to choose for your task?

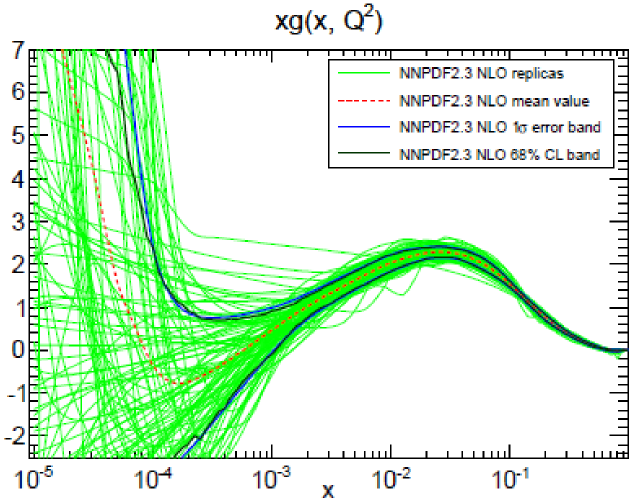
“Error sets” for computing PDF uncertainties

1. Based on diagonalization of the Hessian matrix
   - singular value decomposition of the covariance matrix in the Gaussian approximation
   - Default representation by CTEQ, MMHT, ABM, HERAPDF

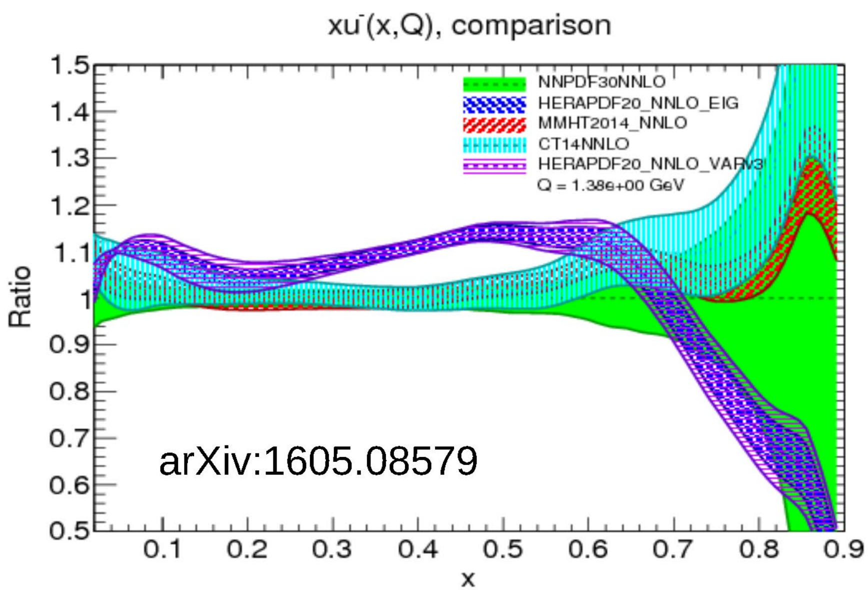
2. Based on Monte-Carlo sampling of probability
   - default representation by Neural Network PDF (NNPDF) collaboration

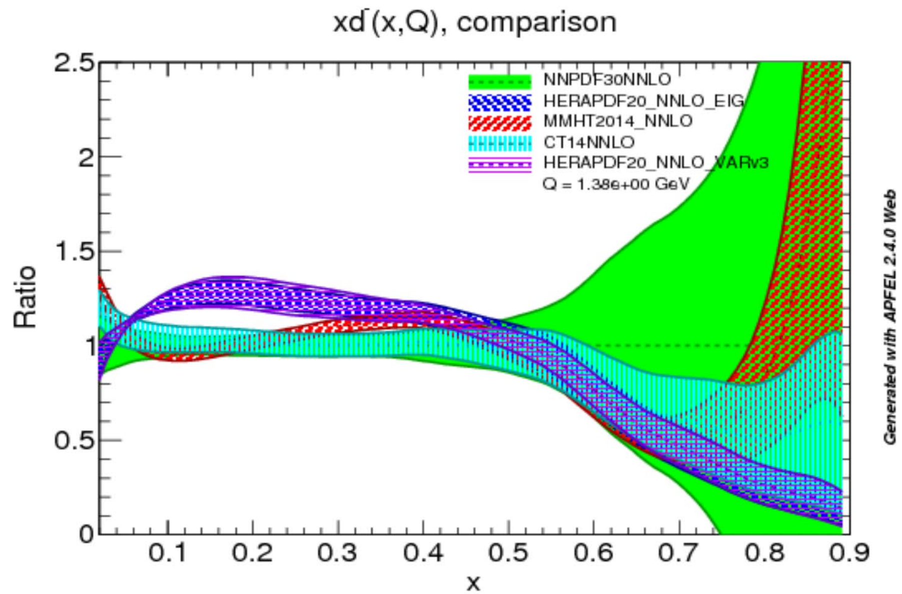
Hessian error sets can be converted into Monte Carlo ones, and back

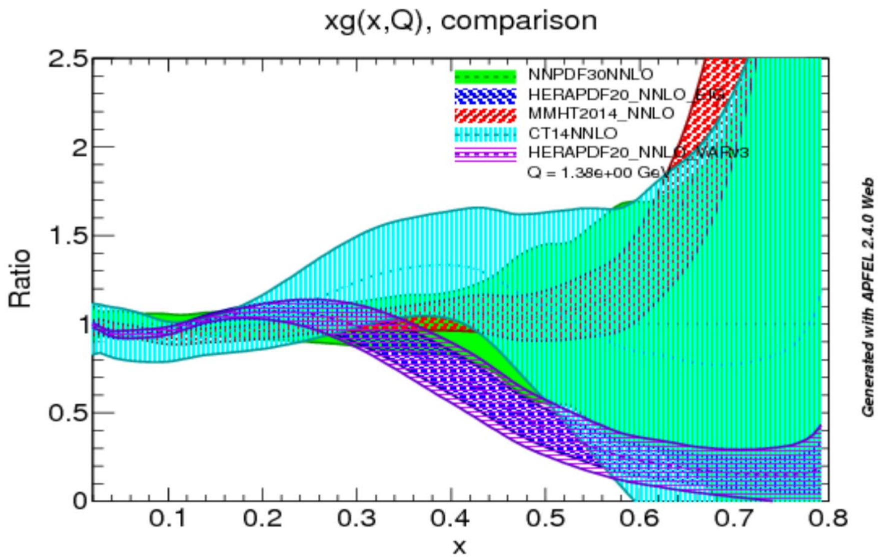




HERAPDF2.0 vs. CT14, MMHT14, NNPDF3.0







Easy online plots of PDFs using **APFEL WEB** (V. Bertone, S. Carrazza, J. Rojo, et al.)

The plots reveal broad agreement between three global PDF ensembles

More pronounced differences with PDF sets based on alternative methodologies, such as HERAPDF2.0

DON'T PANIC

OUTP-15-17P
SMU-HEP-15-12
TIF-UNIMI-2015-14
LCTS/2015-27
CERN-PH-TH-2015-249

arXiv:1510.03865v2 [hep-ph] 12 Nov 2015

PDF4LHC recommendations for LHC Run II

Jon Butterworth[1], Stefano Carrazza[2,4], Amanda Cooper-Sarkar[3], Albert De Roeck[4,5], Joël Feltesse[6], Stefano Forte[2], Jun Gao[7], Sasha Glazov[8], Joey Huston[9], Zahari Kassabov[2,10], Ronan McNulty[11], Andreas Morsch[4], Pavel Nadolsky[12], Voica Radescu[13], Juan Rojo[14] and Robert Thorne[1].

[1] *Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK.*
[2] *TIF Lab, Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy*
[3] *Particle Physics, Department of Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK.*
[4] *PH Department, CERN, CH-1211 Geneva 23, Switzerland*
[5] *Antwerp University, B2610 Wilrijk, Belgium*
[6] *CEA, DSM/IRFU, CE-Saclay, Gif-sur-Yvette, France*
[7] *High Energy Physics Division, Argonne National Laboratory, Argonne, Illinois 60439, U.S.A.*
[8] *Deutsches Elektronen-Synchrotron (DESY), Notkestrasse 85, D-22607 Hamburg, Germany.*
[9] *Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824 U.S.A.*
[10] *Dipartimento di Fisica, Università di Torino and INFN, Sezione di Torino, Via Pietro Giuria 1, I-10125 Torino, Italy*
[11] *School of Physics, University College Dublin Science Centre North, UCD Belfeld, Dublin 4, Ireland*
[12] *Department of Physics, Southern Methodist University, Dallas, TX 75275-0181, U.S.A.*
[13] *Physikalisches Institut, Universität Heidelberg, Heidelberg, Germany.*
[14] *Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, University of Oxford, OX1 3NP Oxford, UK*

A Hitchhiker's Guide to choosing the right PDF set

"A common mistake that people make when trying to design something completely foolproof is to underestimate the ingenuity of complete fools."
–D. Adams

Weak and strong goodness-of-fit criteria

Kovarik, P. N., Soper, **arXiv:1905.06957**

Weak (common) goodness-of-fit (GOF) criterion

Based on the global χ^2

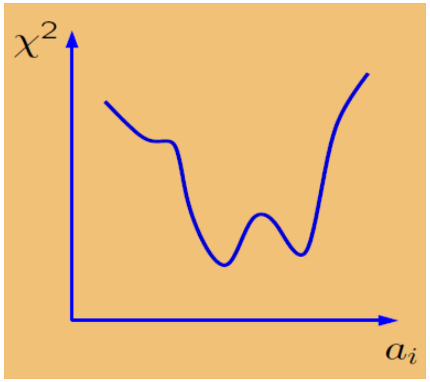
A fit of a PDF model to N_{exp} experiments with N_{pt} points ($N_{pt} \gg 1$) is good at the probability level p if $\chi^2_{global} \equiv \sum_{n=1}^{N_{exp}} \chi^2_n$ satisfies

$$P\left(\chi^2 \geq \chi^2_{global}, N_{pt}\right) \geq p; \qquad e.g.$$

$$\left|\chi^2_{global} - N_{pt}\right| \lesssim \sqrt{2N_{pt}} \text{ for } p = 0.68$$

Even when the weak GOF criterion is satisfied, parts of data can be poorly fitted

Then, **tensions between experiments** may lead to **multiple solutions** or **local χ^2 minima** for some PDF combinations



An excellent fit requires more than a good global χ^2

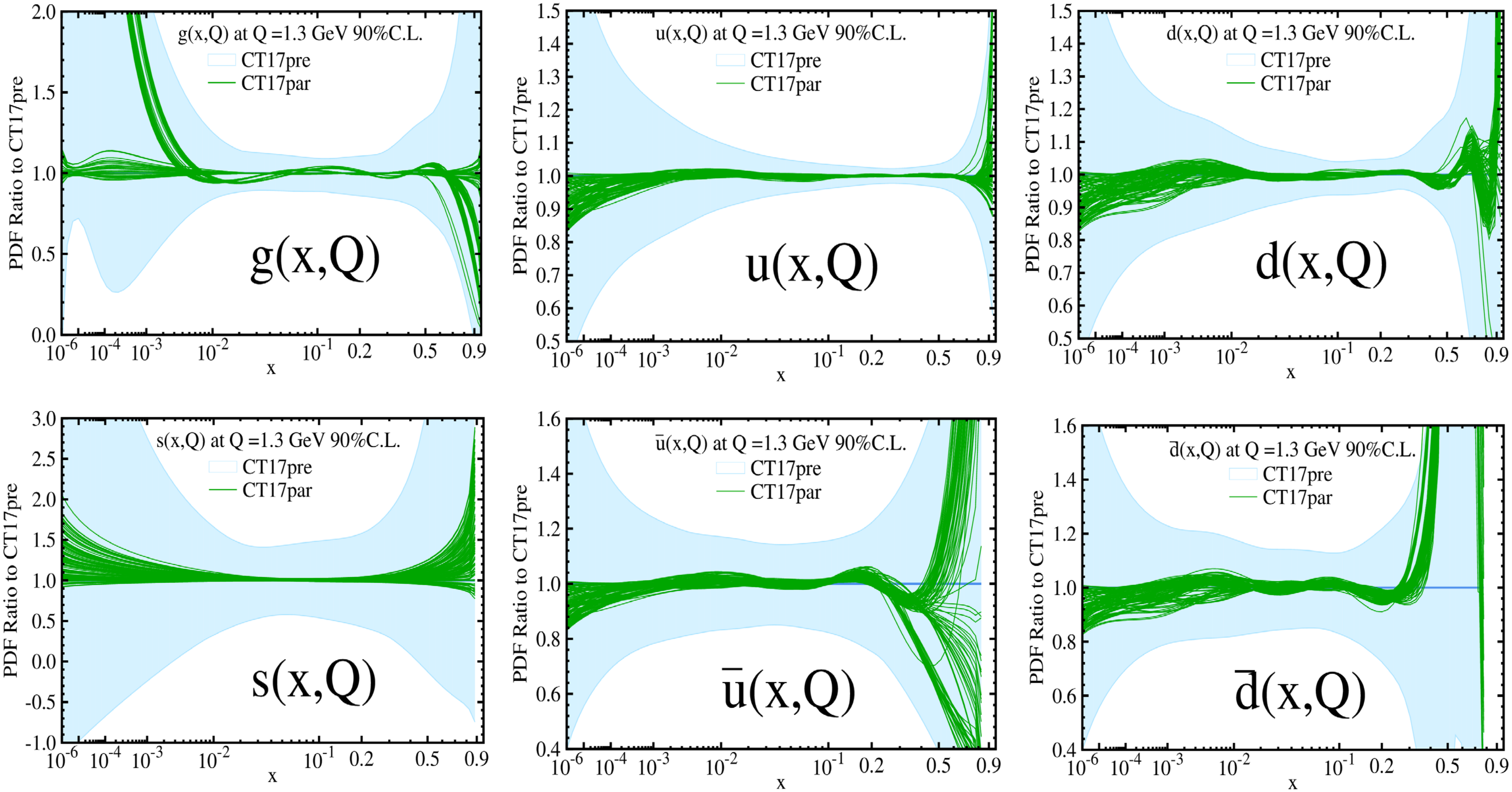
It passes a number of quality tests, called together the **strong set of goodness-of-fit criteria**

1. Each possible partition n of the global data set has a good χ^2
   – differences between theory and data for this partition are indistinguishable from random fluctuations
   – $P(\{\chi_n^2\}) \geq 0.68$ for the distribution of χ_n^2 over N_{part} partitions
2. Best-fit nuisance parameters obey the expected probability distribution
3. **Resampling test:** the data are neither underfitted nor overfitted
4. A closure test is passed, such as the one used in NNPDF 3.x
5. …

Functional forms of PDFs and resampling test

The uncertainty due to the PDF functional form contributes as much as 50% of the total PDF uncertainty in CT fits. The CT18 analysis estimates this uncertainty using 100 trial functional forms.

Explore various non-perturbative parametrization forms of PDFs



- CT17par – sample result of using various non-perturbative parametrization forms.
- No data constrain very large x or very small x regions.

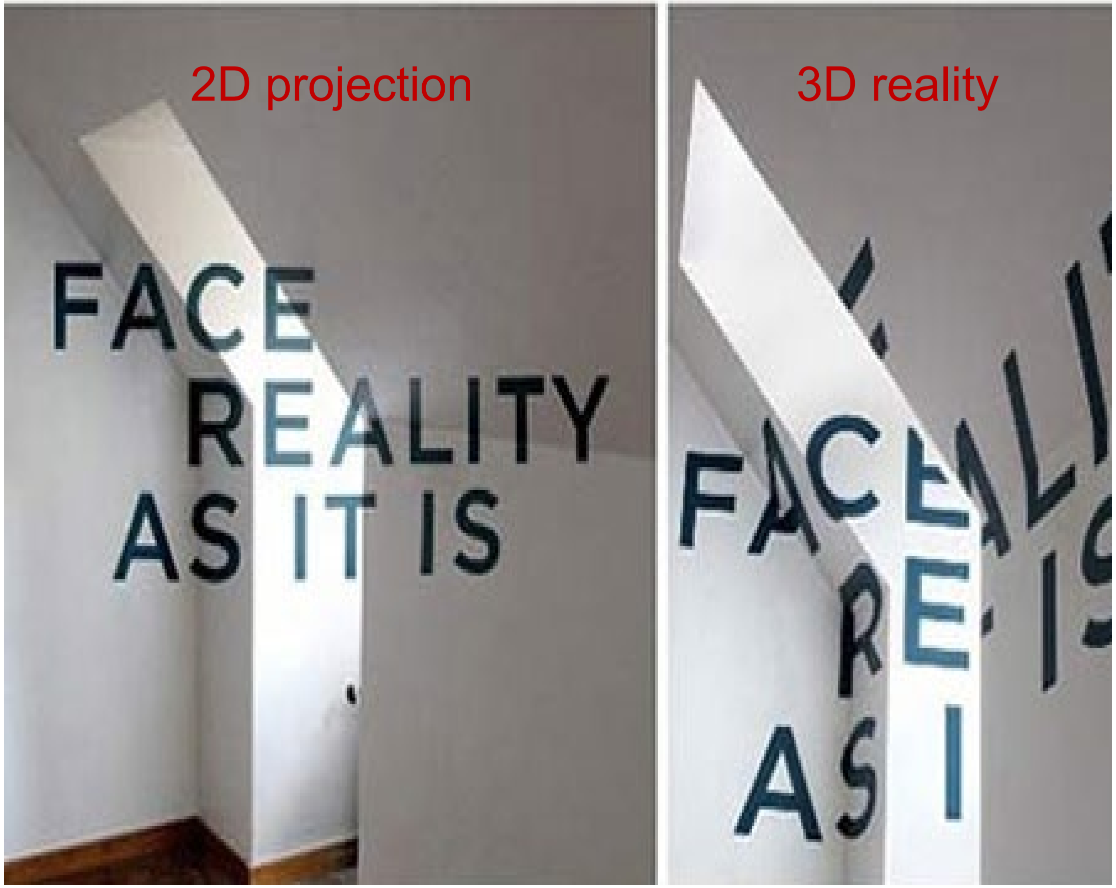
CT14: parametrization forms

- CT14 relaxes restrictions on several PDF combinations that were enforced in CT10. [These combinations were not constrained by the pre-LHC data.]
  - The assumptions $\frac{\bar{d}(x,Q_0)}{\bar{u}(x,Q_0)} \to 1$, $u_v(x,Q_0) \sim d_v(x,Q_0) \propto x^{A_{1v}}$ with $A_{1v} \approx -\frac{1}{2}$ at $x < 10^{-3}$ are relaxed once LHC W/Z data are included
  - CT14 parametrization for $s(x,Q)$ includes extra parameters
- Candidate CT14 fits have 30-35 free parameters
- In general, $\mathrm{f_a}\,(\mathrm{x}, \mathrm{Q_0}) = \mathrm{A} x^{a_1}(1-\mathrm{x})^{\mathrm{a_2}} \mathrm{P_a}(\mathrm{x})$
- CT10 assumed $P_a(x) = \exp(a_0 + a_3\sqrt{x} + a_4 x + a_5\, x^2)$
  - exponential form conveniently enforces positive definite behavior
  - but power law behaviors from a_1 and a_2 may not dominate
- In CT14, $\mathrm{P_a}(\mathrm{x}) = \ {}_{\mathrm{a}}(\mathrm{x})\mathrm{F_a}(\mathrm{z})$, where $\ {}_{\mathrm{a}}(\mathrm{x})$ is a smooth factor
  - $\mathrm{z} = 1 - 1(1-\sqrt{\mathrm{x}})^{\mathrm{a_3}}$ preserves desired Regge-like behavior at low x and high x (with $\mathrm{a_3}$>0)
- Express $F_a(z)$ as a linear combination of Bernstein polynomials:

$$z^4, 4z^3(1-z), 6z^2(1-z)^2\,, 4z(1-z)^3, (1-z)^4$$

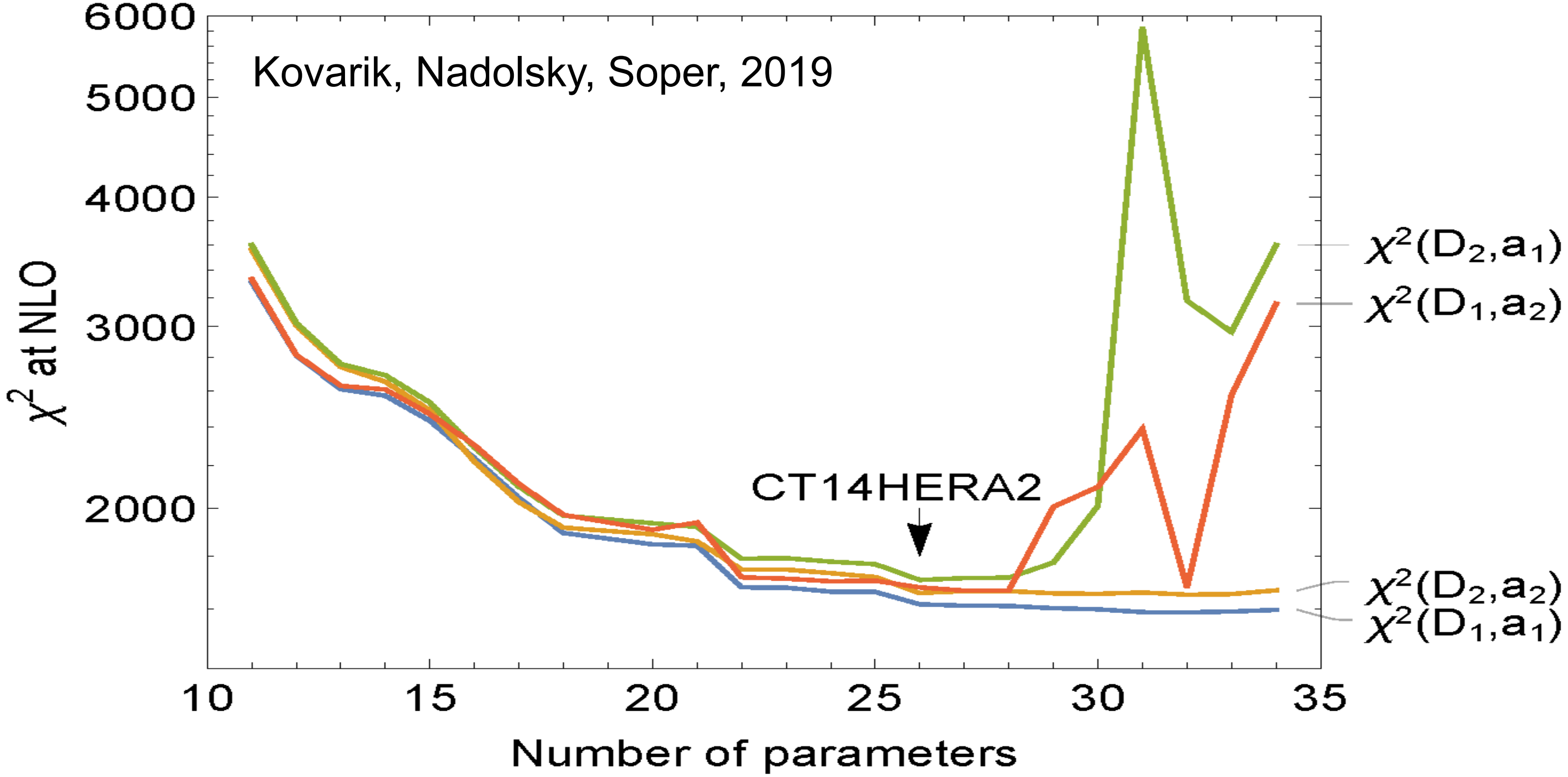
  - each basis polynomial has a single peak, with peaks at different values of z; reduces correlations among parameters

If too few parameters



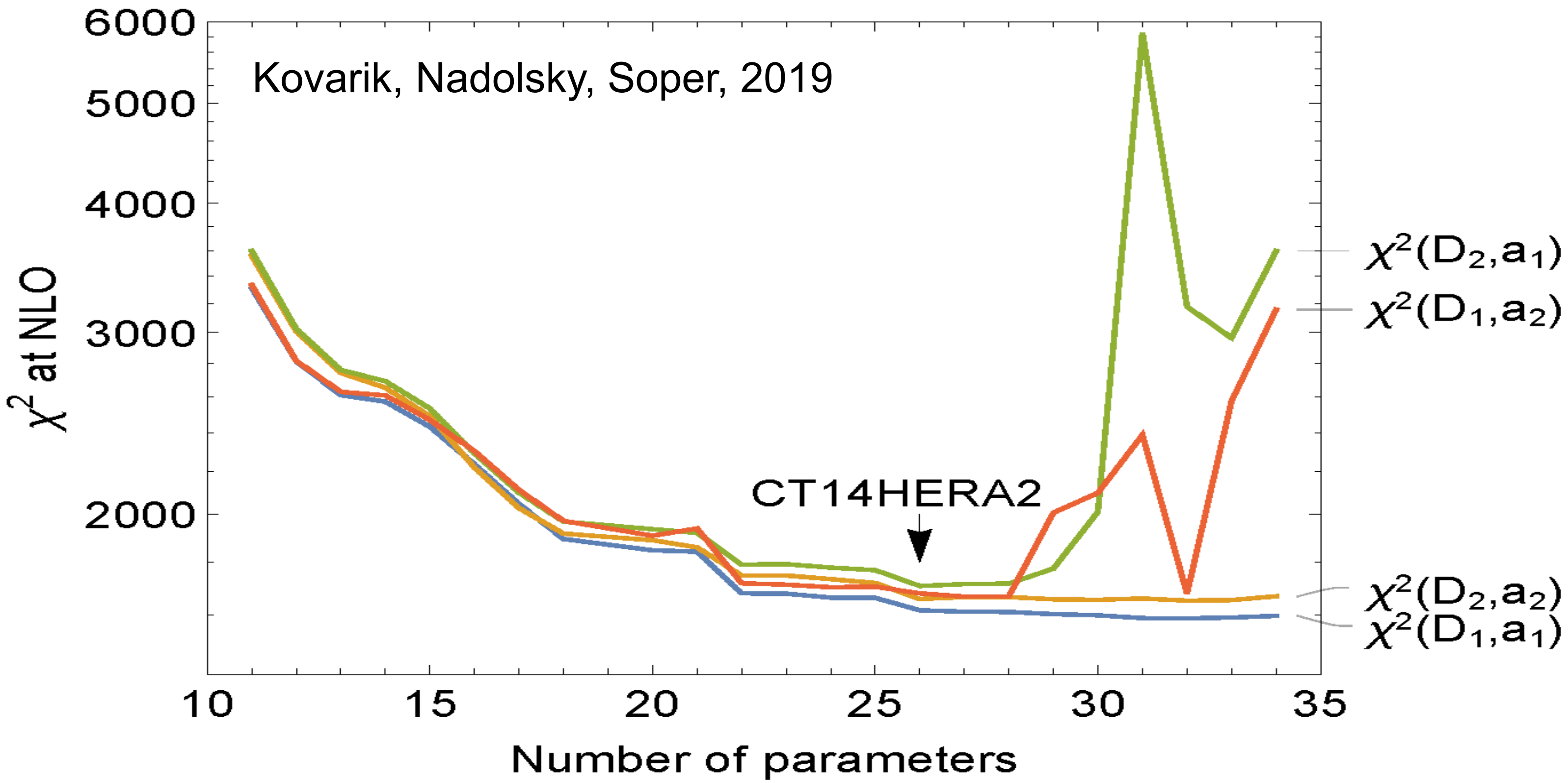
The solution can be consistent and false

If too many parameters



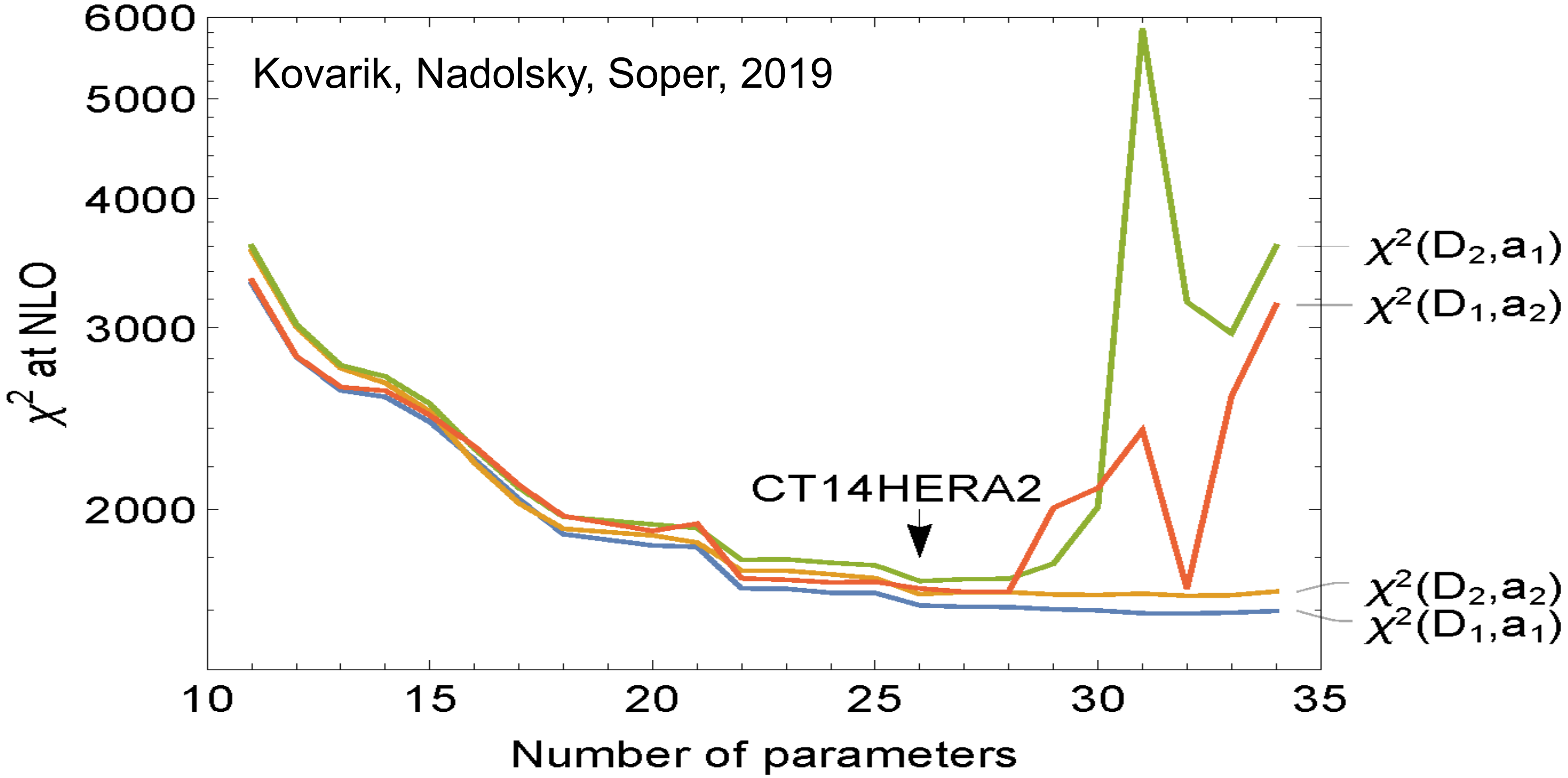
- Randomly split the CT14HERA data set into two halves, D_1 and D_2
- Find parameter vectors a_1 and a_2 from the best fits for D_1 and D_2, respectively

If too many parameters



- **Fitted samples:** $\chi^2(D_1, a_1)$ and $\chi^2(D_2, a_2)$ uniformly decrease with the number of parameters; eventually the fits become unstable ("fitting noise")
- **Control samples:** $\chi^2(D_2, a_1)$ and $\chi^2(D_1, a_2)$ fluctuate when the number of parameters is larger than about 30

If too many parameters

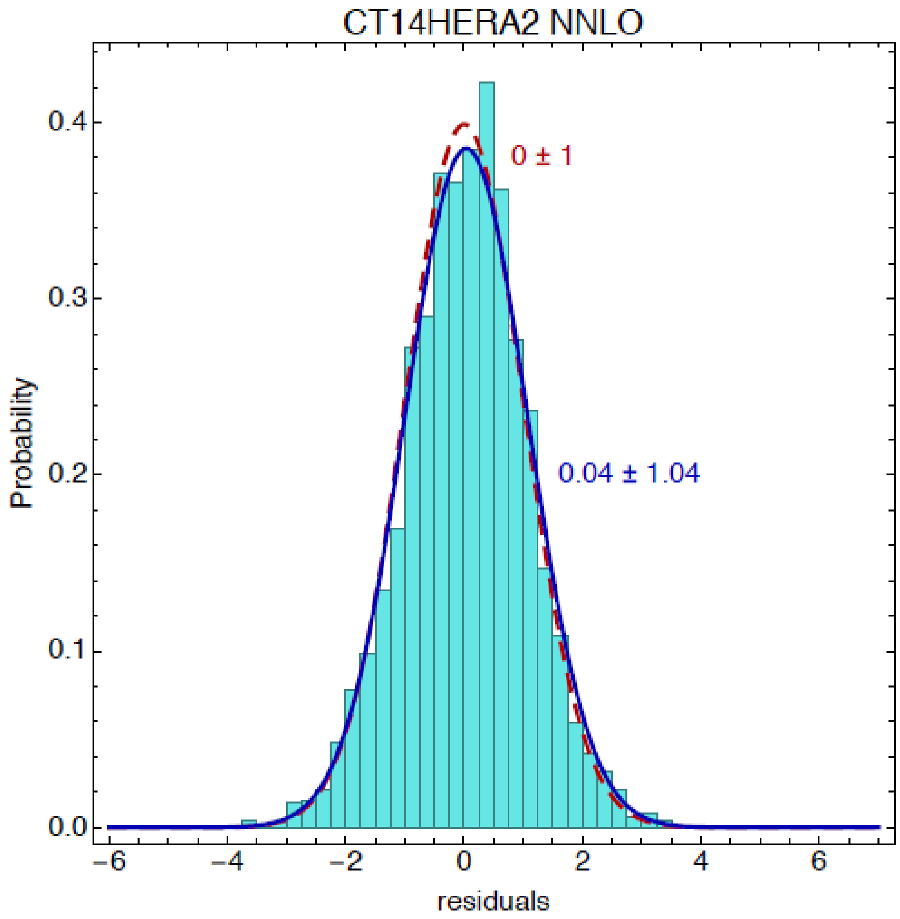


$\lesssim 30$ parameters (26 in CT14HERA2) is optimal for describing the CT14HERA2 data set. 15 parameters or less is optimal for nuclear PDFs

How well are the data described?

Note: It is convenient to define $S_n(\chi^2, N_{pt})$ that approximately obeys the standard normal distribution (mean=0, width=1) independently of N_{pt}

Example: data residuals r_n

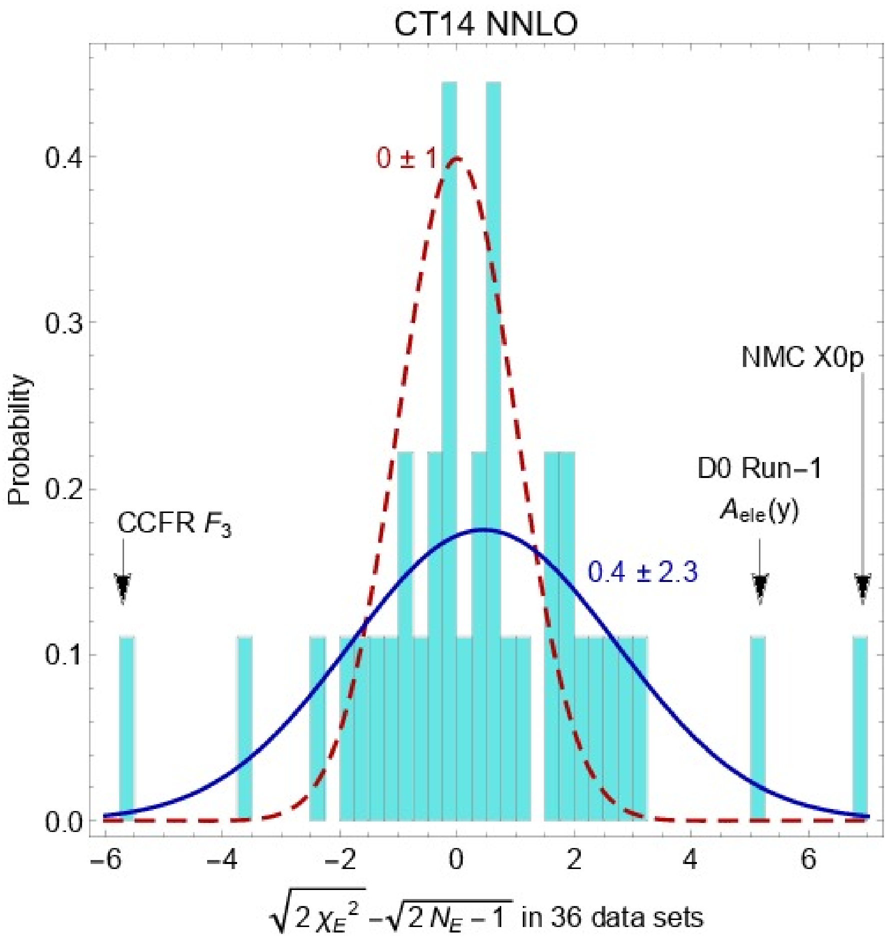


$$r_n \equiv \frac{T_n(\{a\}) - D_n^{shifted}(\{a\})}{\sigma_n^{uncorrelated}}$$

The distribution of residuals is consistent with the standard normal distribution

Full definition of r_n in the backup slides

Example: individual experiments



Define

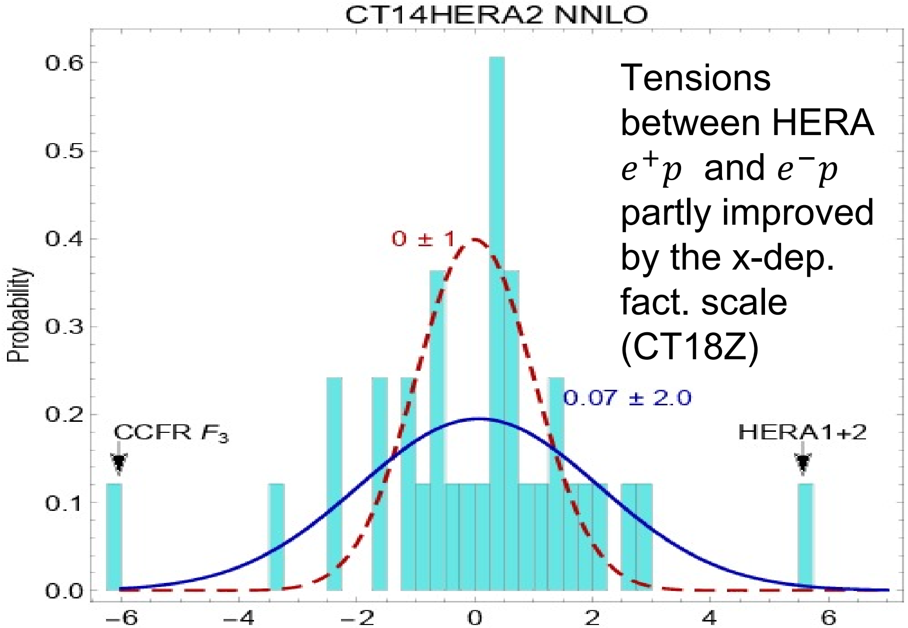
$$S_n(\chi^2, N_{pt}) \equiv \sqrt{2\chi^2} - \sqrt{2N_{pt} - 1}$$

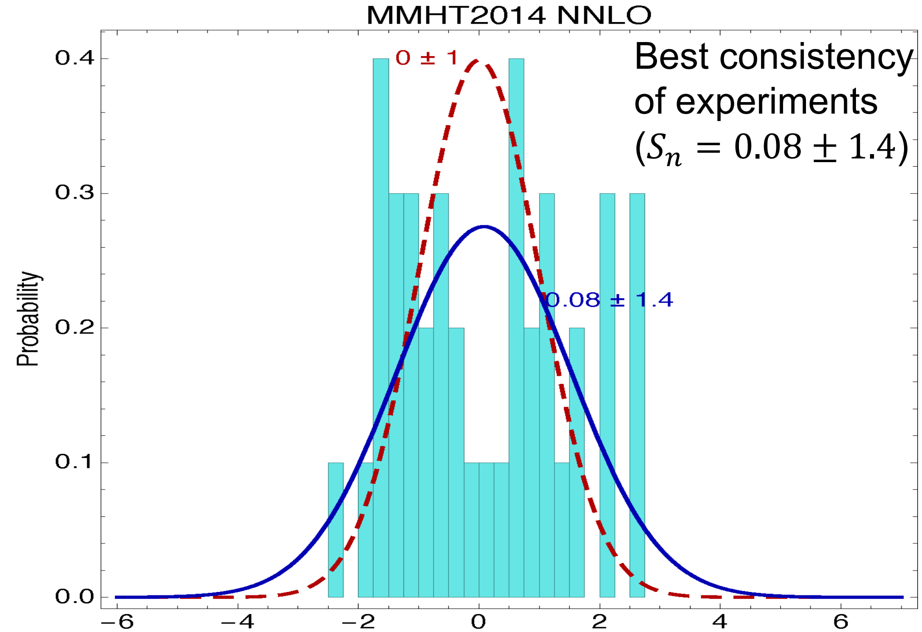
$S_n(\chi_n^2, N_{pt,n})$ are Gaussian distributed with mean 0 and variance 1 for $N_{pt,n} \geq 10$ **[R.A.Fisher, 1925]**

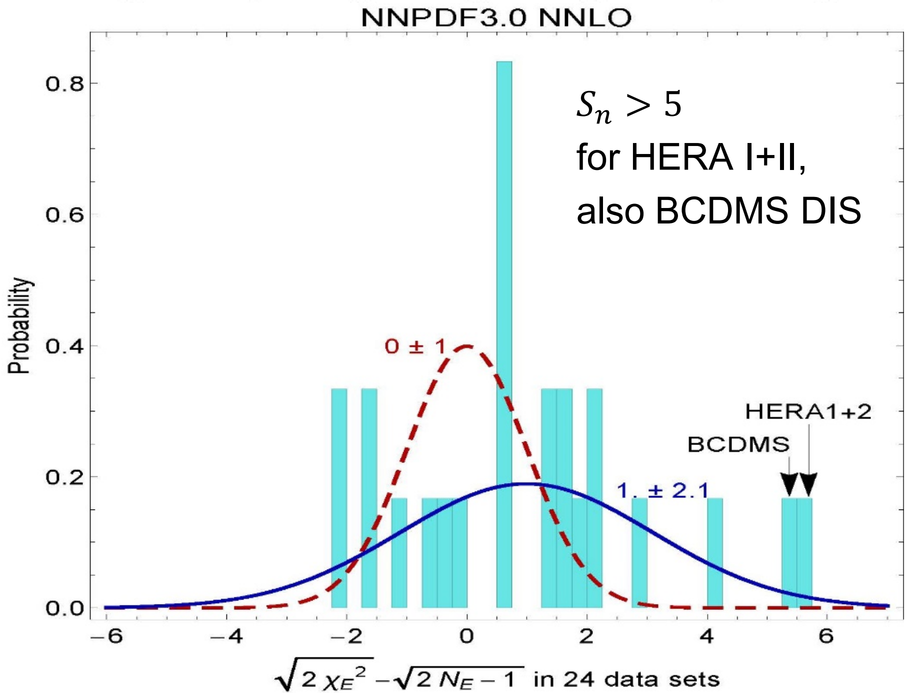
Even more accurate (χ^2, N_{pt}): **T.Lewis, 1988**

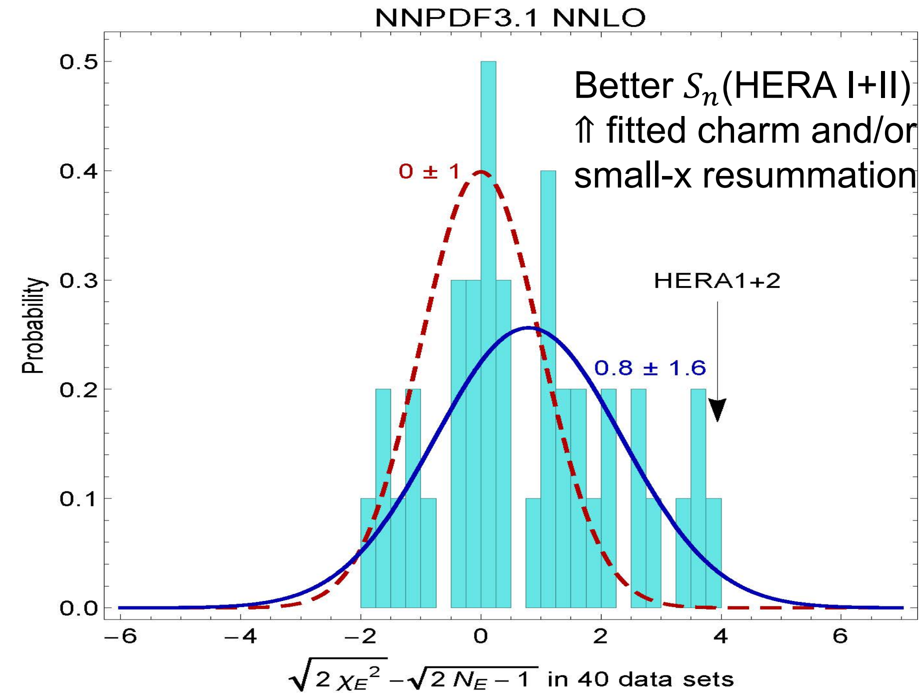
An empirical S_n distribution can be compared to $\mathcal{N}(0,1)$ visually or using a statistical (KS or related) test

Effective Gaussian variables









A fast test of experimental constraints using L_2 sensitivity

T.J. Hobbs et al., 1904.00022

- Requires:
  - Hessian or Monte Carlo error PDFs
  - χ^2 values for fitted or envisioned experiments for each error PDF

- Quantifies:
  - strengths of constraints from individual experiments on given PDFs
  - agreement among the experiments (universality of the best-fit PDFs)
  - sensitivity of processes not included in the global fit

If data point residuals for each error PDF set are also provided, a related $\boldsymbol{L_1}$ **sensitivity** (B. T. Wang et al., 1803.02777, see backup) can be computed to visualize kinematic distributions of experimental constraints in the x-Q plane

Tolerance hypersphere in the PDF space

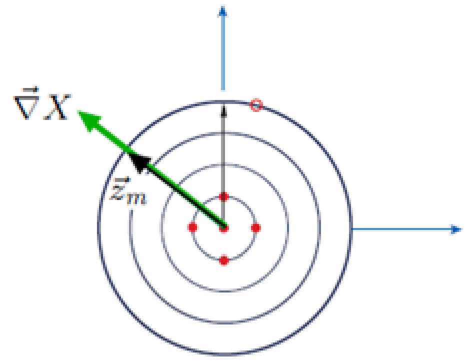
2-dim (i,j) rendition of N-dim (26) PDF parameter space

Hessian method: Pumplin et al., 2001

A symmetric PDF error for a physical observable X is given by

$$\Delta X = \vec{\nabla} X \cdot \vec{z}_m = \left|\vec{\nabla} X\right|$$

$$= \frac{1}{2}\sqrt{\sum_{i=1}^{N}\left(X_i^{(+)} - X_i^{(-)}\right)^2}$$



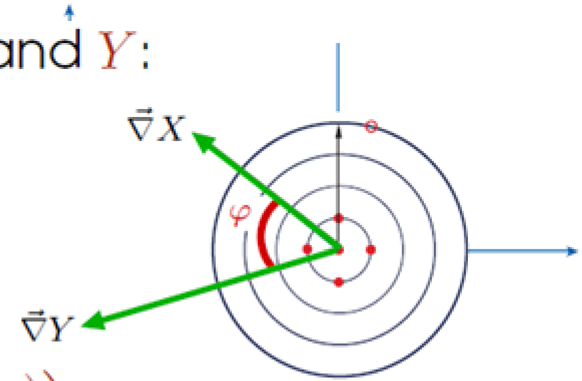
(b)
Orthonormal eigenvector basis

Correlation cosine for observables X and Y:

hep-ph/0110378

$$\cos\varphi = \frac{\vec{\nabla} X \cdot \vec{\nabla} Y}{\Delta X \Delta Y} =$$

$$\frac{1}{4\Delta X\,\Delta Y}\sum_{i=1}^{N}\left(X_i^{(+)} - X_i^{(-)}\right)\left(Y_i^{(+)} - Y_i^{(-)}\right)$$

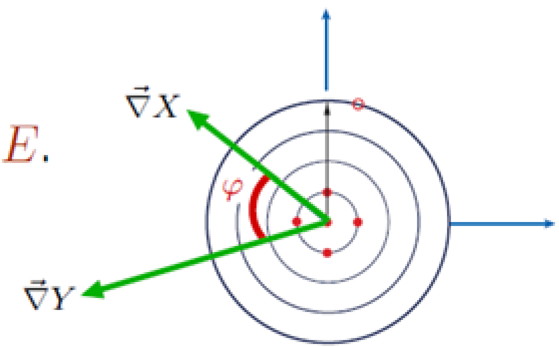


(b)
Orthonormal eigenvector basis

L_2 sensitivity, definition

$S_{f,L_2}(E)$ for experiment E is the estimated $\Delta\chi^2_E$ for this experiment when a PDF $f_a(x_i, Q_i)$ increases by the +68% c.l. Hessian PDF uncertainty

Take $X = f_a(x_i, Q_i)$ or $\sigma(f)$; $Y = \chi^2_E$ for experiment E.

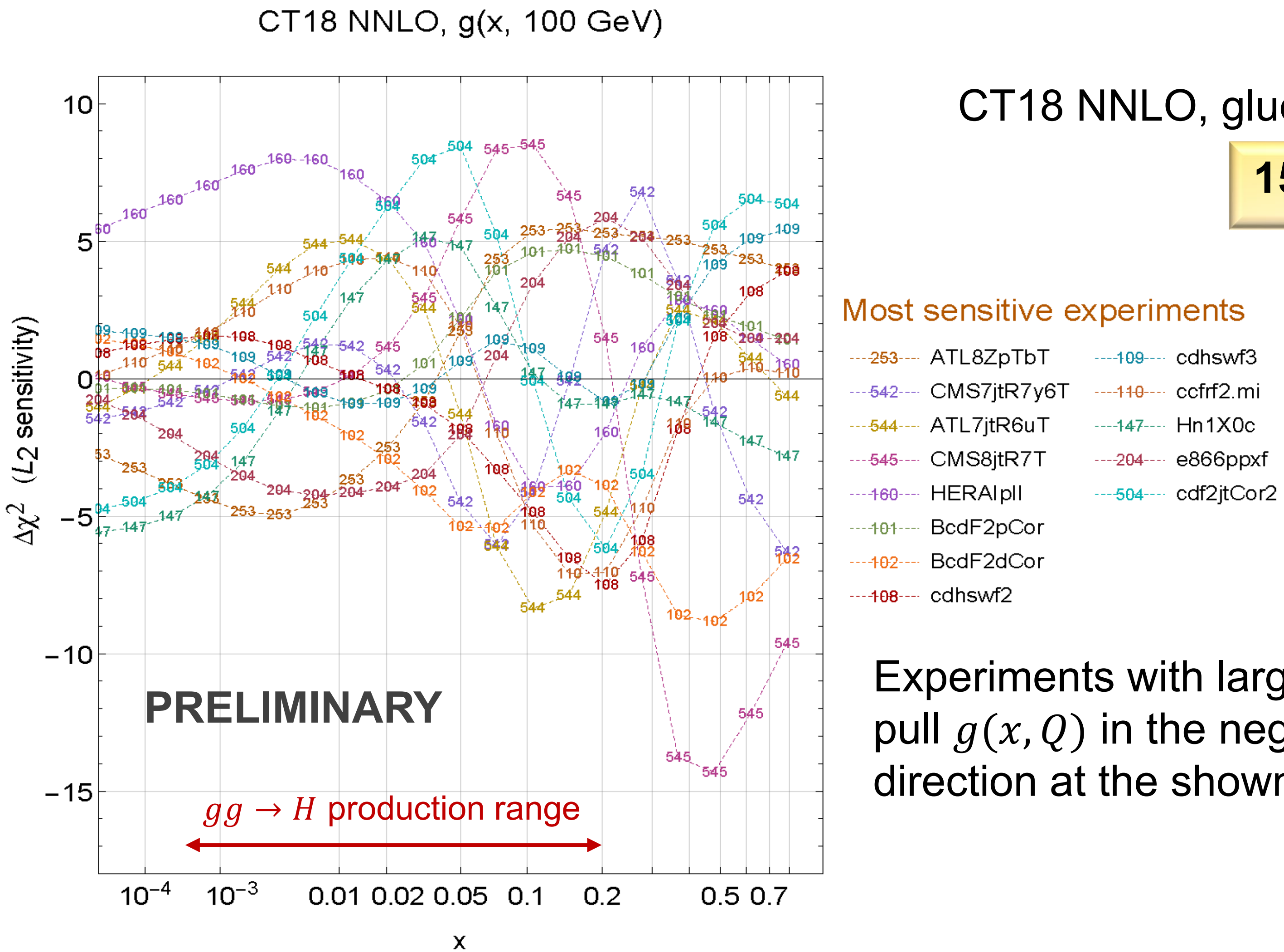


$$S_{f,L_2} \equiv \Delta Y(\vec{z}_{m,X}) = \vec{\nabla} Y \cdot \vec{z}_{m,X} = \vec{\nabla} Y \cdot \frac{\vec{\nabla} X}{|\vec{\nabla} X|} = \Delta Y \cos\varphi.$$

A fast version of the Lagrange Multiplier scan of χ^2_E along the direction of $f_a(x_i, Q_i)$!

Estimated χ^2 pulls from experiments

(L_2 sensitivity, arXiv:1904.00222, v. 2)



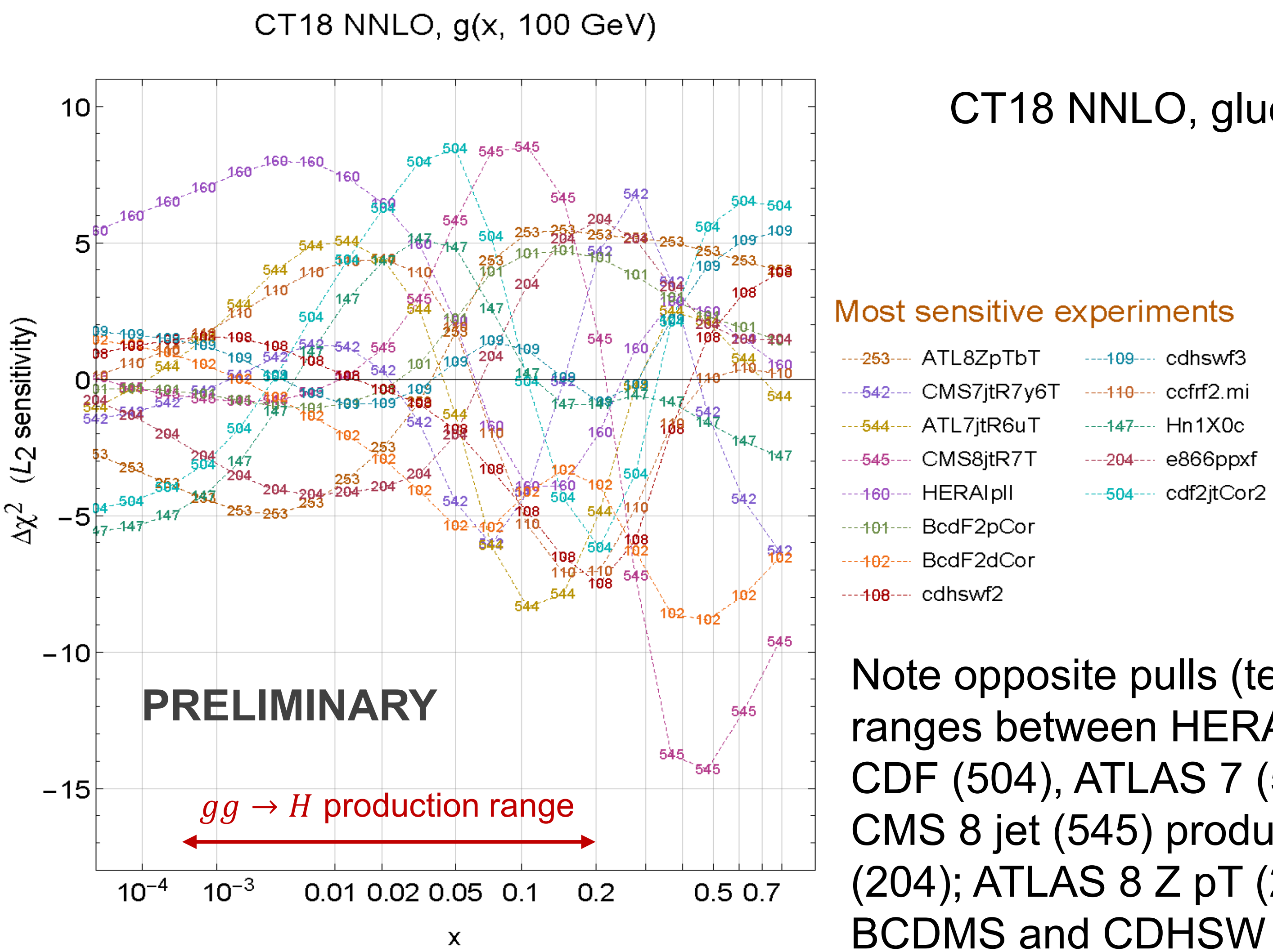
CT18 NNLO, gluon at Q=100 GeV

15 core-minutes

Experiments with large $\Delta\chi^2 > 0$ [$\Delta\chi^2 < 0$] pull $g(x, Q)$ in the negative [positive] direction at the shown x

Estimated χ^2 pulls from experiments

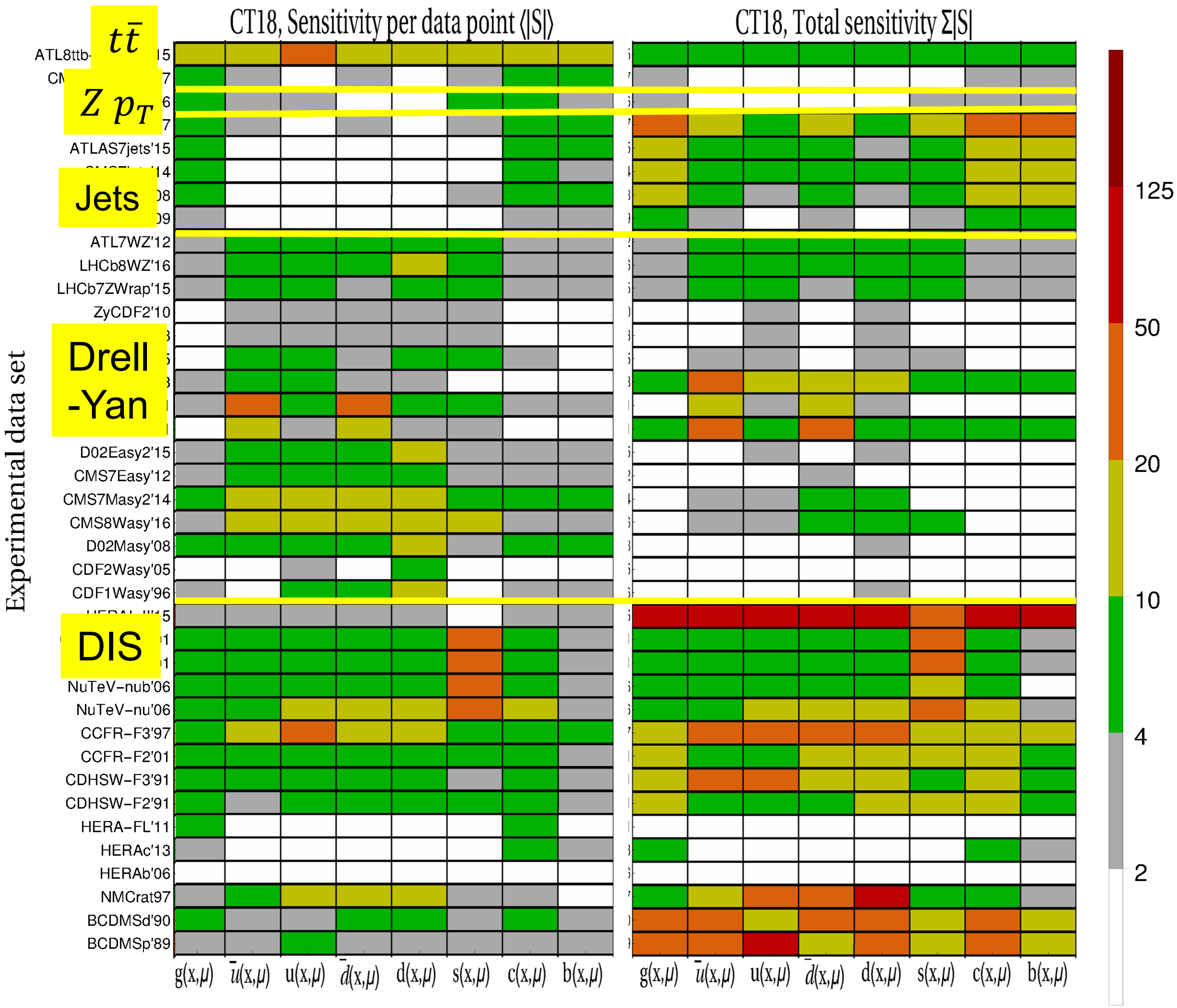
(L_2 sensitivity, arXiv:1904.00222, v. 2)



CT18 NNLO, gluon at Q=100 GeV

Note opposite pulls (tensions) in some x ranges between HERA I+II DIS (ID=160); CDF (504), ATLAS 7 (544), CMS 7 (542), CMS 8 jet (545) production; E866pp DY (204); ATLAS 8 Z pT (253) production; BCDMS and CDHSW DIS

L_1 sensitivity of hadronic experiments to CT18 PDFs

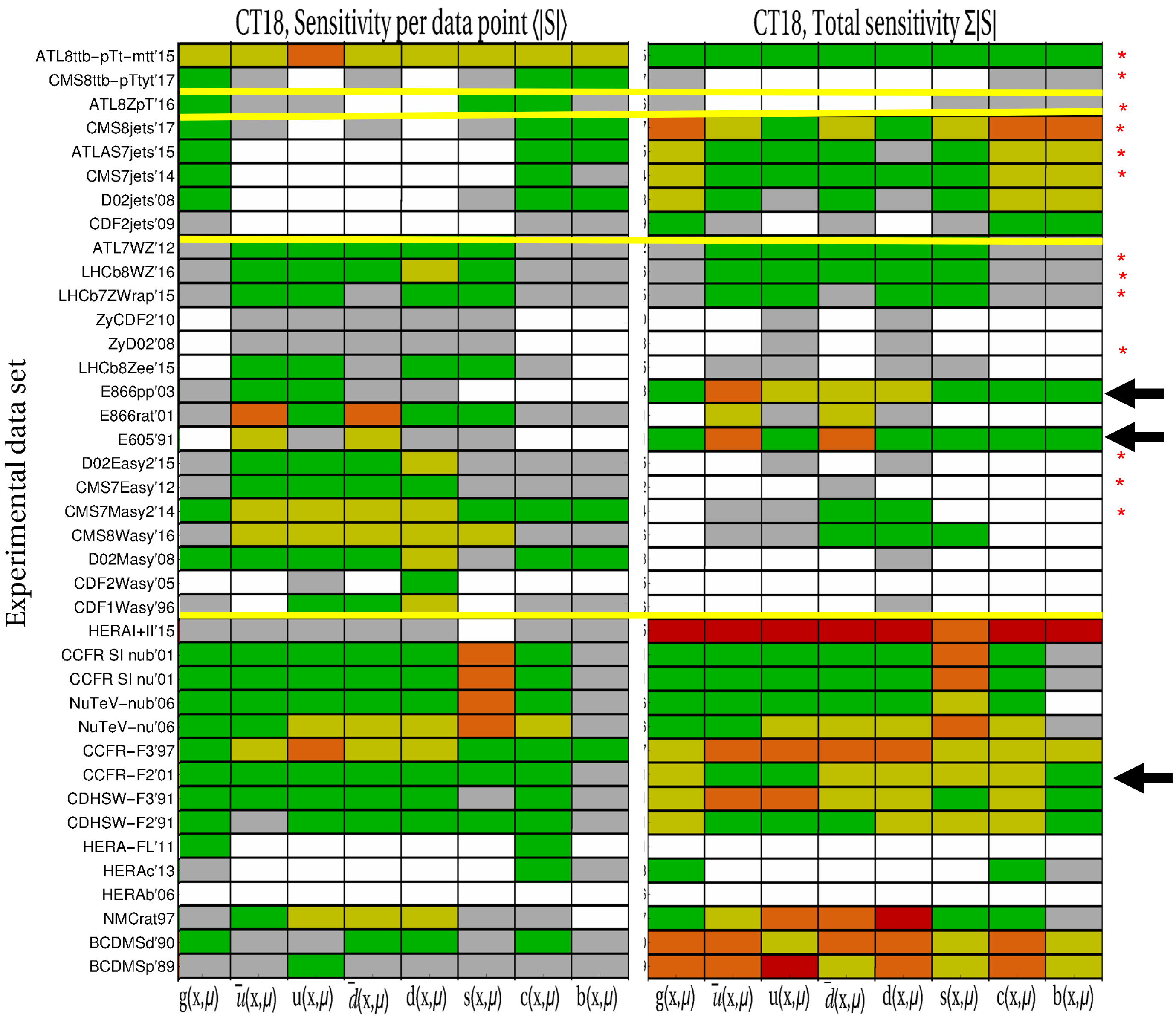


Per-data point and total sensitivity of experimental data sets in the CT18 NNLO analysis to indicated PDF flavors

Computed using the **PDFSense** code
[arXiv:0803.02777]

Red bars = most sensitive experiments

L_1 sensitivity of hadronic experiments to CT18 PDFs

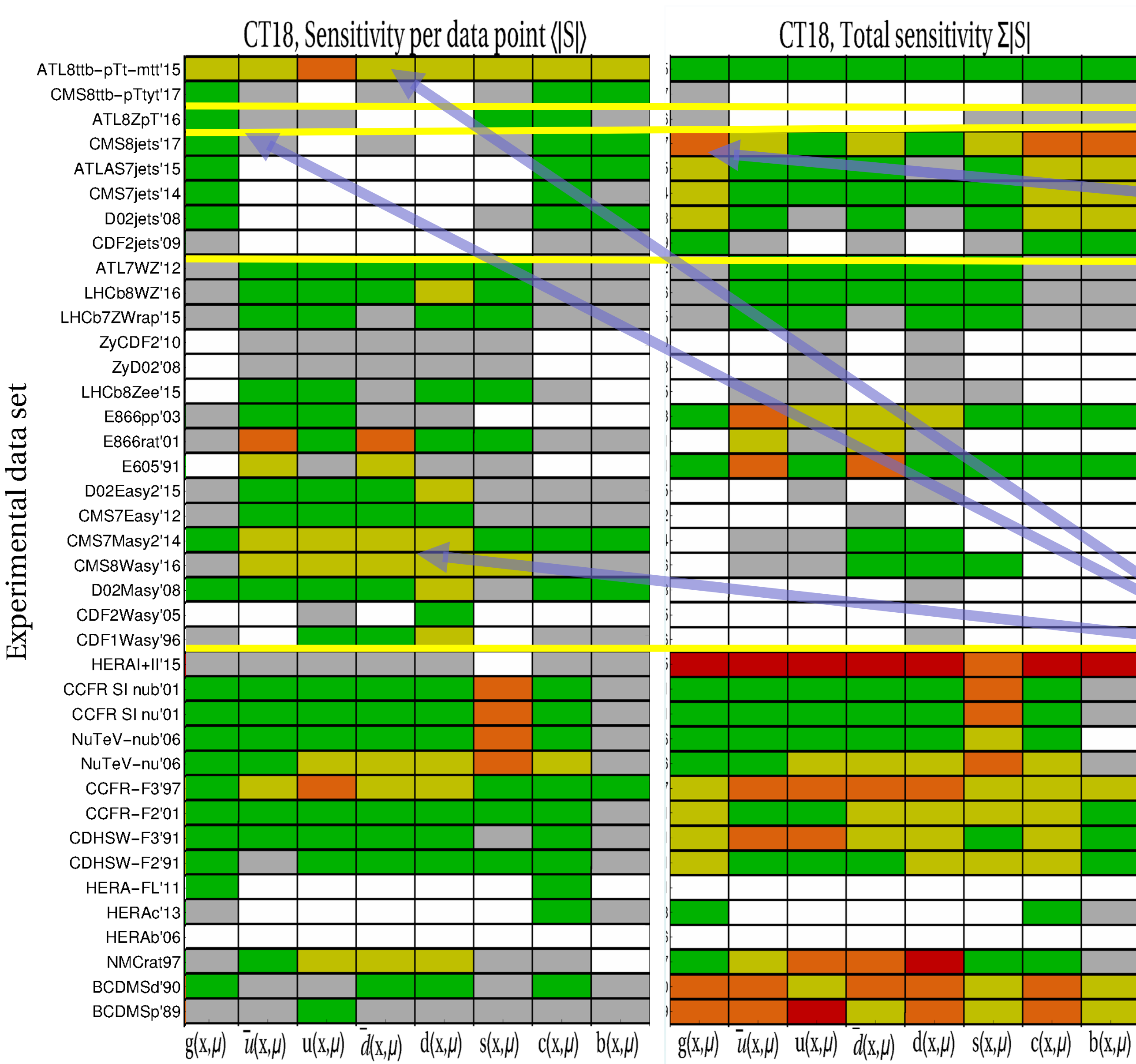


The LHC data sets (*) hold a great promise – if they agree

"Midsize" Drell-Yan experiments (E605, E866, SeaQuest, STAR…) continue to provide competitive constraints

HERA I+II, BCDMS, NMC, DIS data sets dominate present experimental constraints. Large numbers of data points matter!

Sensitivity of hadronic experiments to PDFs



CMS 7 & 8 TeV single-inclusive jet production has highest total sensitivity ($N_{pt} > 100$), modest sensitivity per data point

$t\bar{t}$, CMS W asy, high-p_T Z production have high sensitivity per data point, smaller total sensitivity ($N_{pt} \sim 10 - 20$)

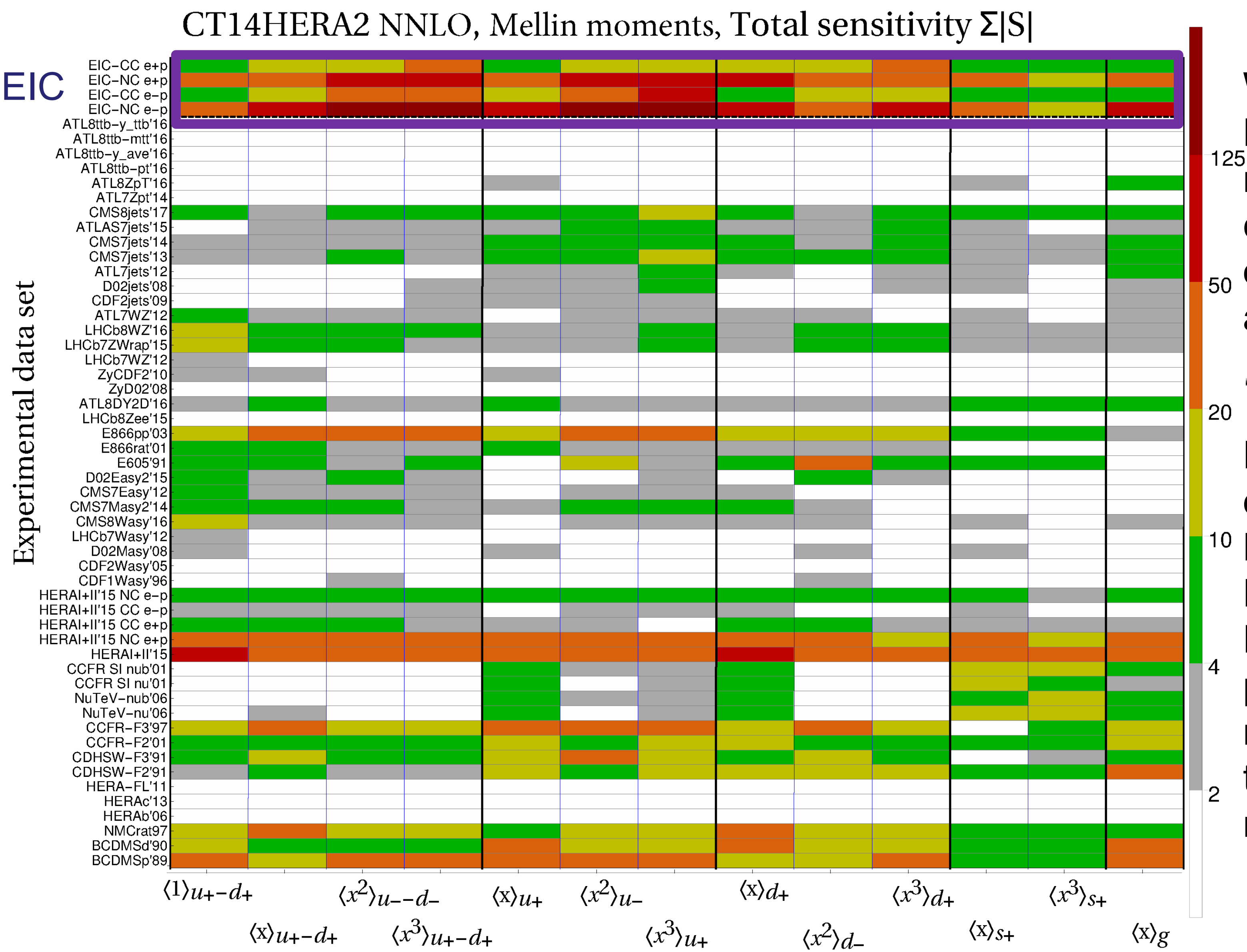
What about *future experiments*

…like the EIC or LHeC?

especially, in the context of other measurements at HL-LHC

- **EIC and LHeC** PDFSense projections by Hobbs and Wang **[arXiv:1907.00988]**
- Compared to **HL-LHC** projections by Abdul Khalek, Bailey, Gao, Harland-Lang, Rojo **[arXiv:1810.03039]**

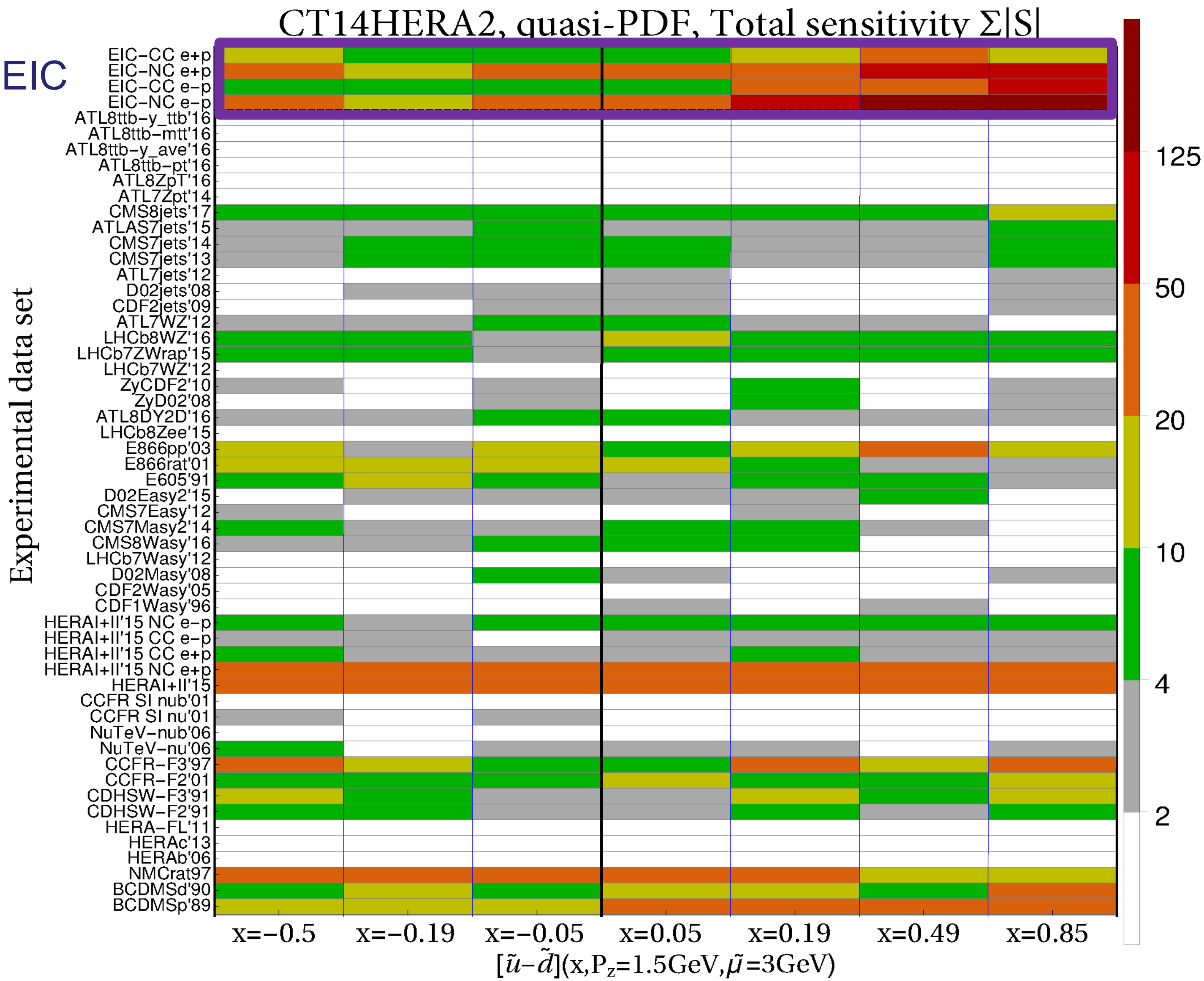
Total sensitivity to Mellin moments



We show Mellin moments computable on the lattice at the scale $\mu_F = 2\,\mathrm{GeV}$

Pre-2019 data: HERA, BCDMS, NMC, E866 DY pair production are most sensitive to the moments

Total sensitivity to lattice quasi-PDFs



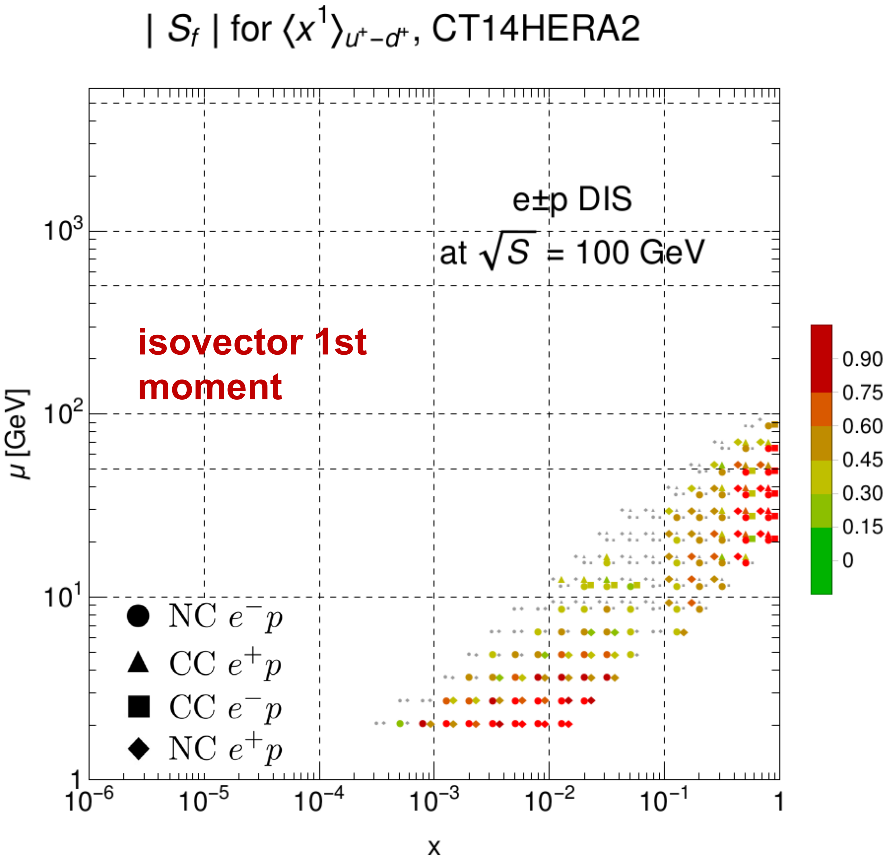
$$\tilde{q}(x) \equiv \begin{cases} q(x), \\ x > 0; \\ -\bar{q}(|x|), \\ x < 0 \end{cases}$$

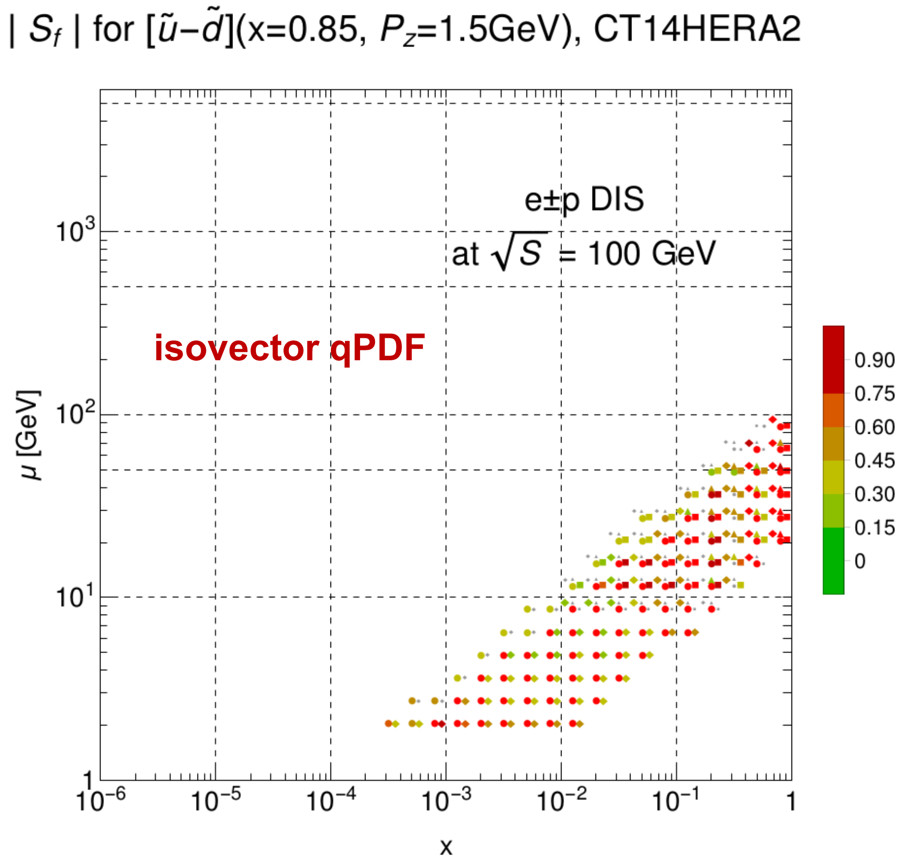
$Q = 3$ GeV,
$P_z = 1.5$ GeV

An EIC would drive lattice phenomenology

- A high-luminosity lepton-hadron collider will impose very tight constraints on many lattice observables; below, the isovector first moment and qPDF

- Many of the experiments most sensitive to PDF Mellin moments and qPDFs involve nuclear targets ⟶ **eA data from EIC would sharpen knowledge of nuclear corrections**





Rethinking the PDFs for the next decade

Some ideas

- common physics goals
  ⇒ learn about the 3D hadron structure!
- shared resources
  - LHAPDF-like repository for interpolations of polarized/TMD/GPD PDFs? For χ^2 values for error PDFs? Other outputs of the fits?
  - Coordinated software development for global fits?
- agreed-upon practices
  - presentation of data and theory predictions? RIVET for the EIC?
  - common definitions of PDF uncertainties?
  - a common standard for PDF validation tests?
- benchmarking studies
  - **explore experimental constraints on various types of PDFs and from various available and future processes at (N)(N)LO using the L_2 sensitivity technique ← SMU can lead**

Backup

A shifted residual r_i

$r_i(\vec{a}) = \frac{T_i(\vec{a}) - D_i^{sh}(\vec{a})}{s_i}$ are N_{pt} **shifted residuals** for point i, PDF parameters $\vec{a}$

$\bar{\lambda}_\alpha(\vec{a})$ are N_λ **optimized nuisance parameters** (dependent on $\vec{a}$)

The $\chi^2(\vec{a})$ for experiment E is

$$\chi^2(\vec{a}) = \sum_{i=1}^{N_{pt}} r_i^2(\vec{a}) + \sum_{\alpha=1}^{N_\lambda} \bar{\lambda}_\alpha^2(\vec{a}) \approx \sum_{i=1}^{N_{pt}} r_i^2(\vec{a})$$

$T_i(\vec{a})$ is the theory prediction for PDF parameters $\vec{a}$

D_i^{sh} is the data value **including the optimal systematic shift**

$$D_i^{sh}(\vec{a}) = D_i - \sum_{\alpha=1}^{N_\lambda} \beta_{i\alpha} \bar{\lambda}_\alpha(\vec{a})$$

s_i is the uncorrelated error

$r_i(\vec{a})$ and $\bar{\lambda}_\alpha(\vec{a})$ are tabulated or extracted from the cov. matrix

Finding shifted residuals r_i from the covariance matrix

The CTEQ-TEA fit returns tables of $r_i(\vec{a})$ and $\bar{\lambda}_\alpha(\vec{a})$ for every i and α

Alternatively, they can be found from the covariance matrix:

$$r_i(\vec{a}) = s_i \sum_{j=1}^{N_{pt}} (\mathrm{cov}^{-1})_{ij} \, (T_j(\vec{a}) - D_j), \qquad \bar{\lambda}_\alpha(\vec{a}) = \sum_{i,j=1}^{N_{pt}} (\mathrm{cov}^{-1})_{ij} \frac{\beta_{i\alpha}}{s_i} \frac{(T_j(\vec{a}) - D_j)}{s_j}$$

Vectors of data residuals

For every data point i, construct a vector of residuals $r_i(\vec{a}_k^{\pm})$ for 2N Hessian eigenvectors. $\mathrm{k} = 1, \ldots, N$, with $N = 28$ for CT14 NNLO:

$$\vec{\delta}_i = \{\delta_{i,1}^{+}, \delta_{i,1}^{-}, \ldots, \delta_{i,N}^{+}, \delta_{i,N}^{-}\} \qquad [\mathrm{N} = 28]$$

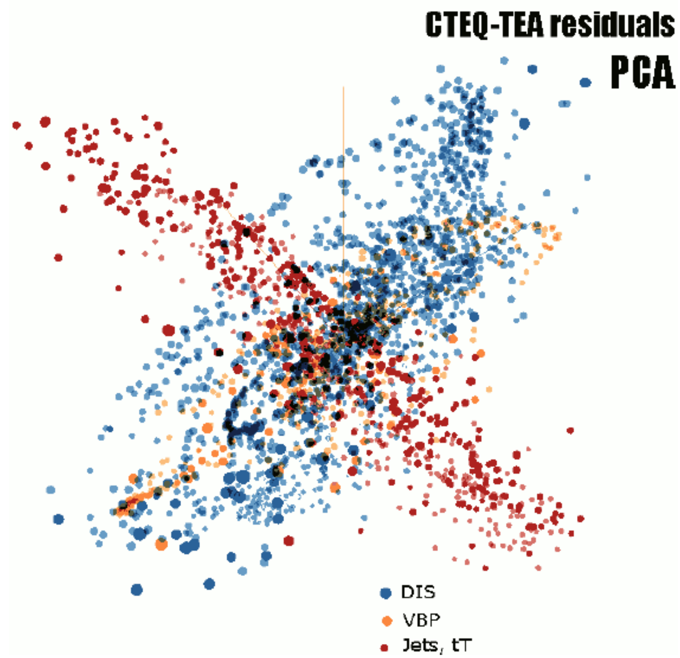
$$\delta_{i,k}^{\pm} \equiv \left(r_i\left(\vec{a}_k^{\pm}\right) - r_i(\vec{a}_0)\right) / \langle r_0 \rangle_E$$

-- a 56-dim vector normalized to $\langle r_0 \rangle_E$, the root-mean-squared residual for the experiment E for the central fit $\vec{a}_0$

$$\langle r_0 \rangle_E \equiv \sqrt{\frac{1}{N_{pt}} \sum_{i=1}^{N_{pt}} r_i^2(\vec{a}_0)} \approx \sqrt{\frac{\chi_E^2\,(\vec{a}_0)}{N_{pt}}}$$

$\langle r_0 \rangle_E \approx 1$ in a good fit to E

r_i is defined in the backup



The TensorFlow Embedding Projector (http://projector.tensorflow.org) represents CT14HERA2 $\vec{\delta}_i$ vectors by their 10 principal components indicated by scatter points. A sample 3-dim. projection of the 56-dim. manifold is shown above. A symmetric 28-dim. representation can be alternatively used.

PDFSense maps, kinematical matchings

- **residual-PDF correlations and sensitivities are evaluated at parton-level kinematics determined according to leading-order matchings with physical scales in measurements**

x_i	: parton mom. fraction
μ_i	: factorization scale

deeply-inelastic scattering:

$$\mu_i \approx Q\big|_i \,,\ x_i \approx x_B\big|_i$$

hadron-hadron collisions:
$AB \to CX$

$$\mu_i \approx Q\big|_i \,,\ x_i^{\pm} \approx \frac{Q}{\sqrt{s}}\,\exp(\pm y_C)\Big|_i$$

single-inclusive jet production: $Q = 2p_{Tj},\ y_C = y_j$

$t\bar{t}$ pair production: $Q = m_{t\bar{t}},\ y_C = y_{t\bar{t}}$

$d\sigma/dp_T^Z$ measurements: $Q = \sqrt{(p_T^Z)^2 + (M_Z)^2},\ y_C = y_Z$

etc...