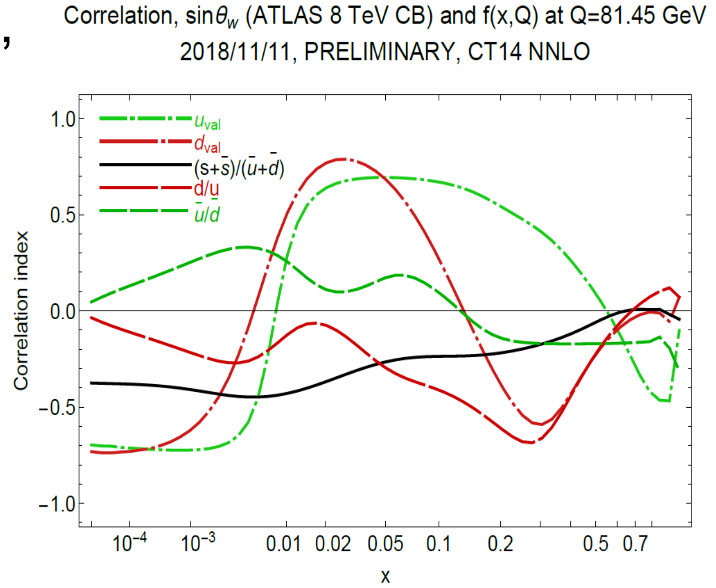
A global likelihood is not enough

CTEQ-TEA PDFs in the LHC electroweak precision studies

Pavel Nadolsky

Southern Methodist University

In collaboration with the CTEQ-TEA group, Karol Kovarik, and Dave Soper



Evolving PDF models

- EW precision fits and PDF fits are fundamentally different.
  - In an EW fit ("ZFITTER"), the Standard Model parameters are found by fitting a **fixed** theoretical model.
  - In a PDF fit ("XFITTER"), the theoretical model (PDF parametrization) **evolves** when more data are added.

⇒ A PDF model can change its functional form within some limits to evade falsification by a new data set

- 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. This part of analysis requires significant human intervention.
  - In the EW benchmarking exercise, as many as 10,000*100 toys may be necessary to estimate the CT parametrization uncertainty

Discriminating between two PDF fits based on the Bayes theorem

Given a new data set D ["$A_\ell(\mathrm{y}_\ell)$"], we can determine the ratio ρ of posterior likelihoods $P(T_i|D)$ of two fits T_1 and T_2:

$$\rho = \frac{P(T_2|D)}{P(T_1|D)} = \frac{P(D|T_1)\cdot P(T_1)}{P(D|T_2)\cdot P(T_2)}$$

- $P(D|T) \propto e^{-\chi^2(D,T)/2}$ is determined from the fit to D
- The prior $P(T)$ is determined by many theoretical considerations and past experimental measurements

Discriminating between two PDF fits based on the Bayes theorem

Given a new data set D ["$A_\ell(y_\ell)$"], we can determine the ratio ρ of posterior likelihoods $P(T_i|D)$ of two fits T_1 and T_2:

$$\rho = \frac{P(T_2|D)}{P(T_1|D)} = \frac{P(D|T_1)\cdot P(T_1)}{P(D|T_2)\cdot P(T_2)}$$

- T_1 is very unlikely compared to T_2 iff $\rho \ll 1$
- ρ can be used to establish the PDF uncertainty, with the caveats that
  - In the PDF fits, the functional form and other assumptions affect both $P(D|T)$ and $P(T)$
  - $P(D|T)$ need not be Gaussian for 1 experiment (examples to follow)

To do the "LHC2" EW fit, the groups must agree on the treatment of the prior $P(T)$ [!!!]

Origin of differences between PDF sets

1. **Updates** on exptl. data sets, **corrections** of outdated theory

2. **PDF uncertainty**

- ≡ a range of allowed PDF shapes for plausible input assumptions, reflected by the PDF error band
- is associated with
  - the choice and treatment of fitted experiments
  - experimental errors propagated into PDF's
  - handling of inconsistencies between experiments
  - choice of $\alpha_S(M_Z)$, m_c, m_b, renormalization/factorization schemes and scales, parametrizations for PDF's, higher-twist terms, nuclear effects,...
  - statistical methodology
- leads to non-negligible differences between the newest central PDF sets

How good are our PDF fits?

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

Based on the global χ^2

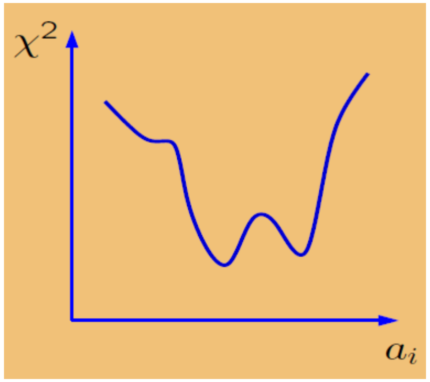
A fit of a PDF model to N_{exp} experiments with N_{pt} points ($N_{pt} \gg 1$) is good at the probability level p if $\chi^2_{global} \equiv \sum_{n=1}^{N_{exp}} \chi^2_n$ satisfies

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

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

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



Strong (new?) GOF criterion

(From Kovarik, P.N., Soper, paper in preparation)

Shatter the global data set by dividing it into N_{part} partitions with $N_{pt,n}$ points each

$$1 \leq N_{part} \leq N_{pt}$$

$$\sum_{n=1}^{N_{part}} N_{pt,n} = N_{pt}$$

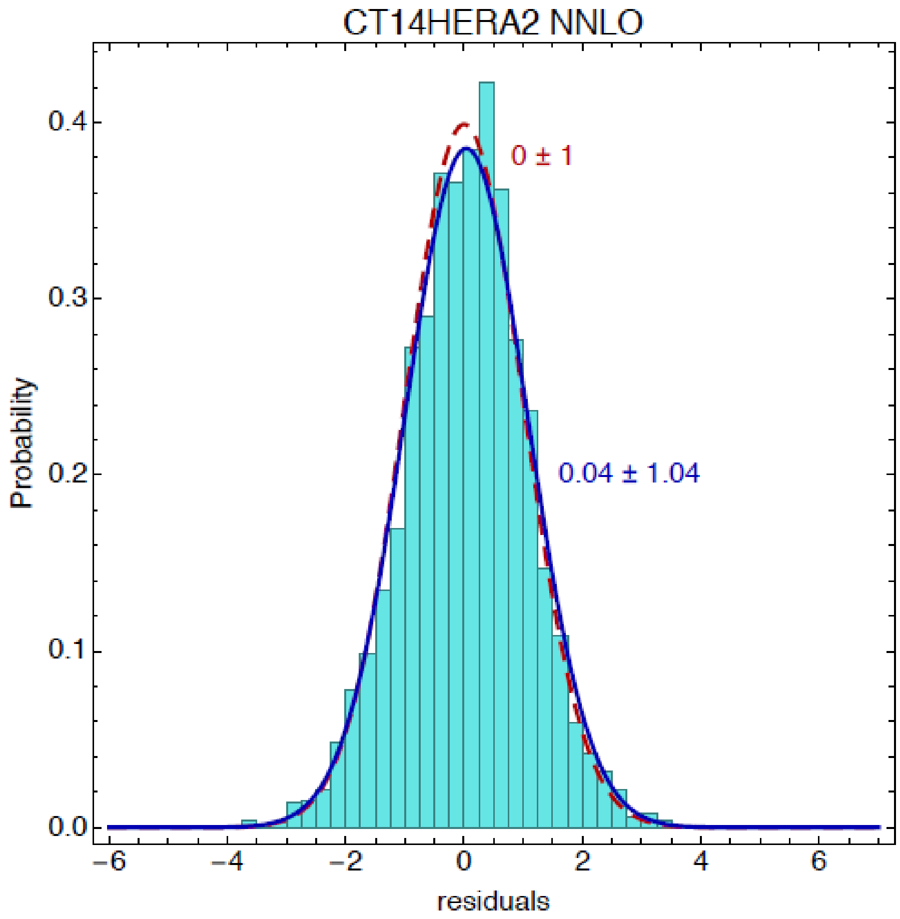
A fit is good iff the weak GOF criterion is satisfied for every partition. That is, for each possible partition n:

- differences between theory and data are indistinguishable from random fluctuations
- $P\left(\chi^2 \geq \chi_n^2, N_{pt,n}\right) \geq 0.68$ for $\mathrm{n} = 1, \dots, N_{part}$

How good are our PDF fits?

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: $N_{part} = N_{pt}$, data residuals r_n

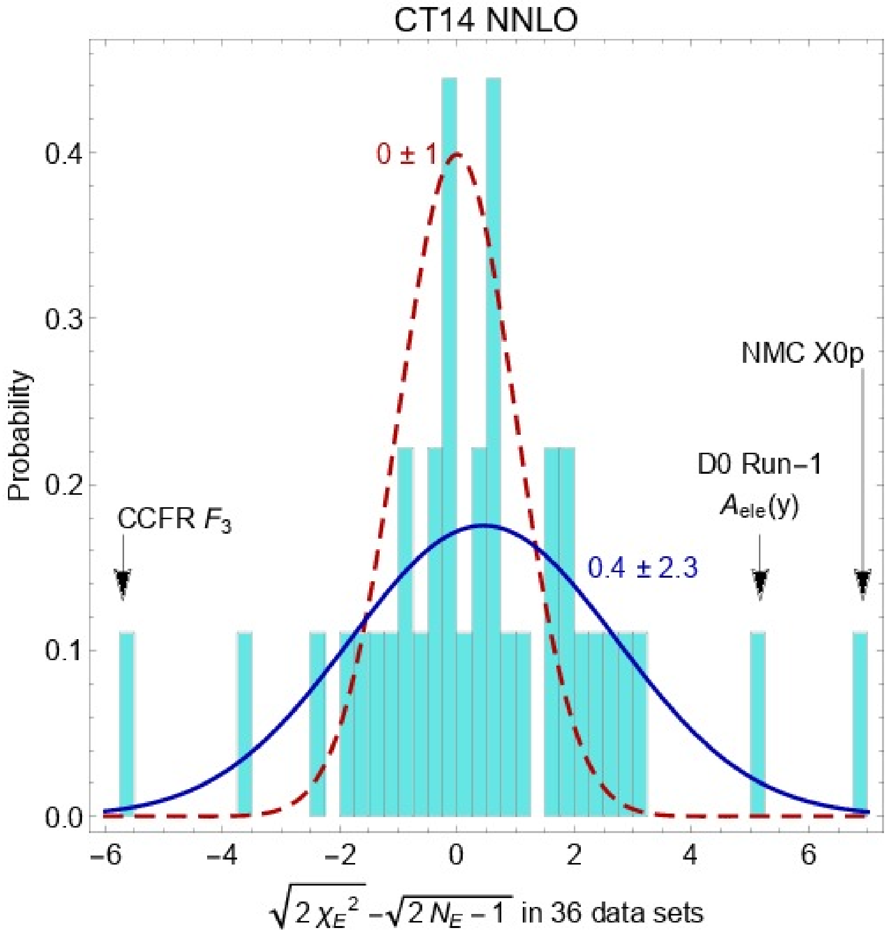


$$r_n \equiv \frac{T_n(\{a\}) - D_n^{shifted}(\{a\})}{\sigma_n^{uncorrelated}}$$

The distribution of residuals is consistent with the standard normal distribution

Full definition of r_n in the backup slides

Example: $N_{part} = N_{exp}$, individual experiments



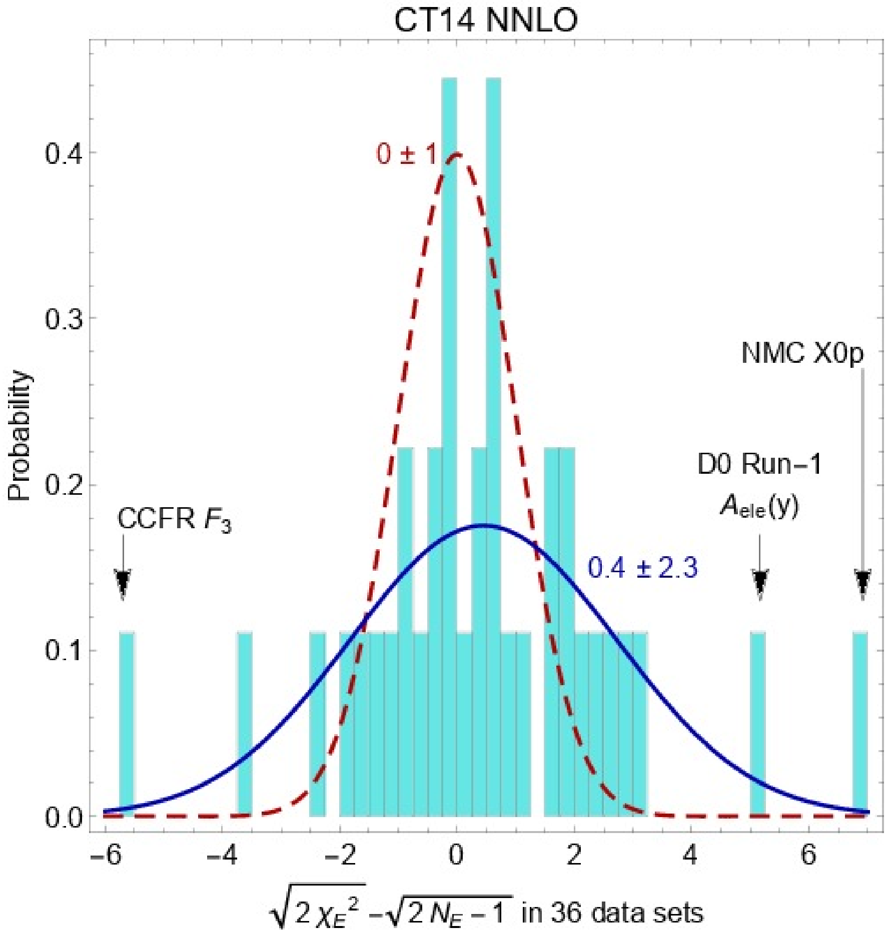
Define

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

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

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

Example: $N_{part} = N_{exp}$, individual experiments

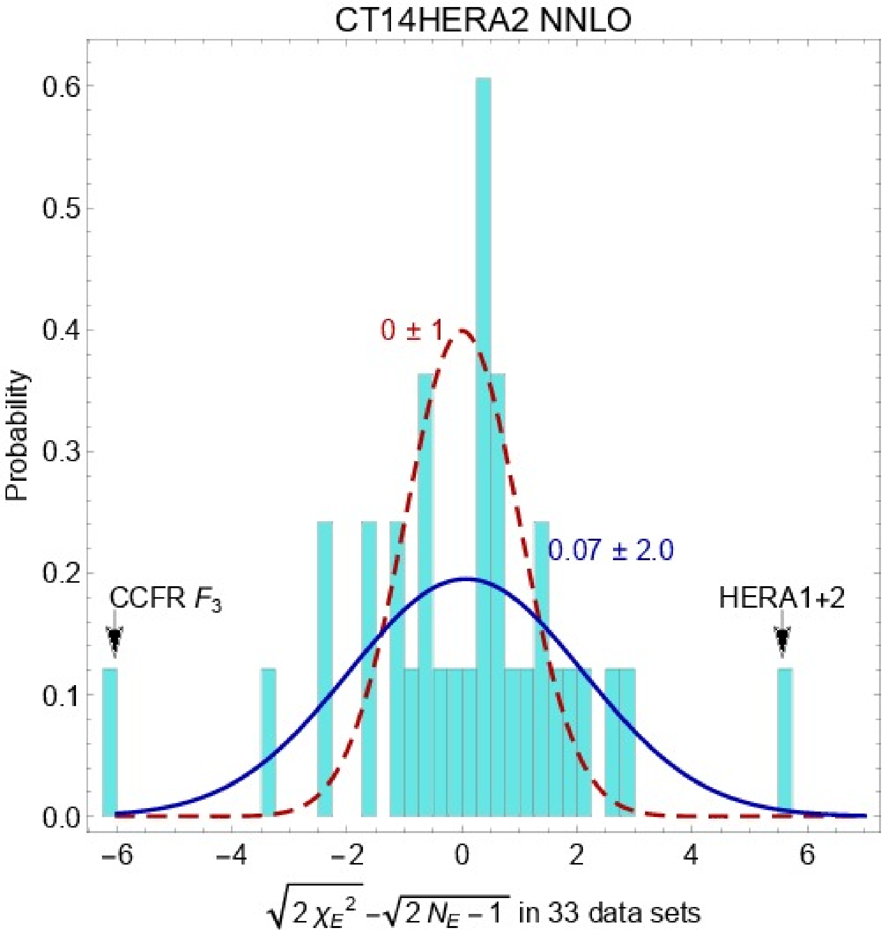


Some S_n are too big or too small in a global fit

CT14 NNLO:

- $S_n > 4$ for NMC DIS ep cross section and D0 Run-1 electron charge asymmetry
- These data sets are eliminated in CT14HERA2/CT18 fits
- The rest of CT14 experiments are reasonably consistent; $S_n \sim \mathcal{N}(0.3, 1.6)$

Example: $N_{part} = N_{exp}$, individual experiments



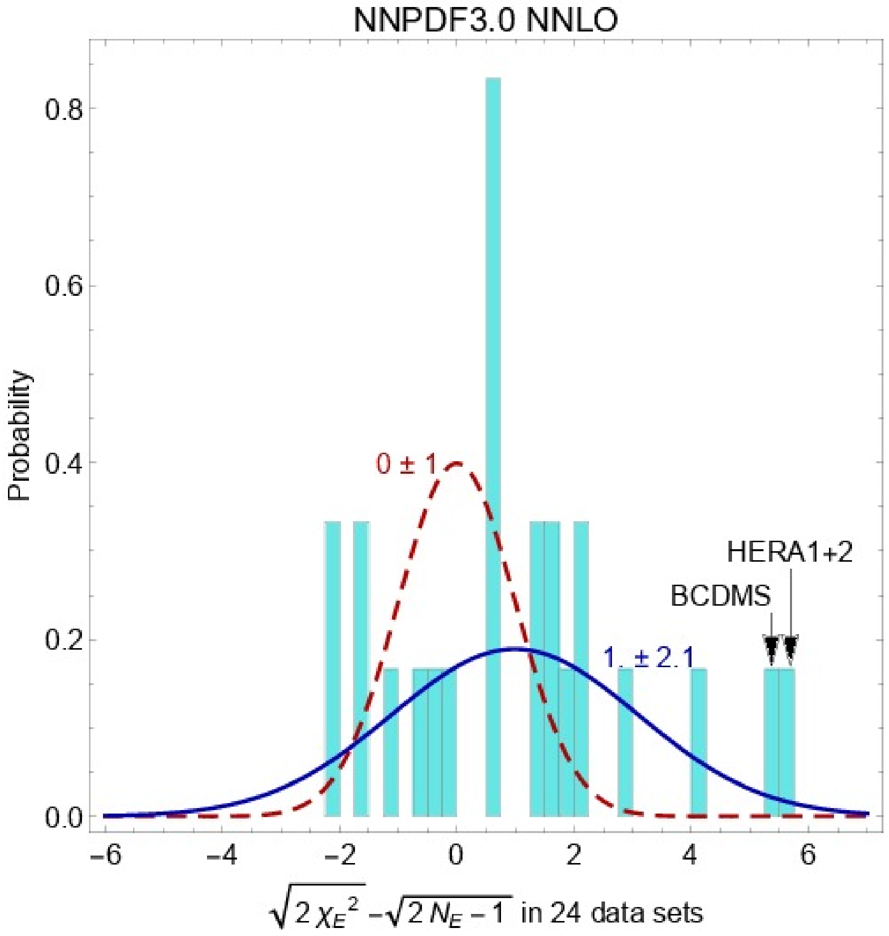
CT14 HERA2 NNLO:

For HERA 1+2 inclusive DIS data

- $\frac{\chi^2}{N_{pt,n}} > 1.15$: not good for $N_{pt,n} = 1120$

  $\Rightarrow S_n^{HERA\ I+II} = 5.89$
- Tensions between e^+p and e^-p DIS channels
- Partly improved by the x-dependent factorization scale (CT18Z) or small-x resummation

Example: $N_{part} = N_{exp}$, individual experiments



Similar tensions observed in other global fits

NNPDF3.0 NNLO:

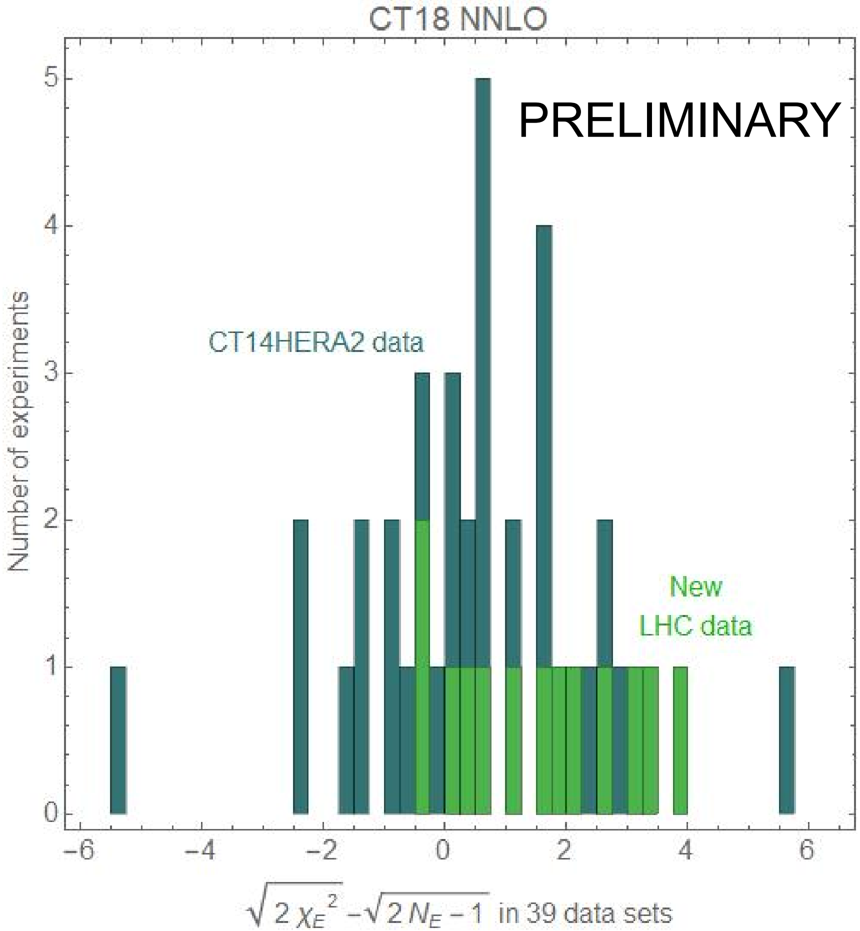
$S_n > 5$ for HERA I+II, also BCDMS DIS

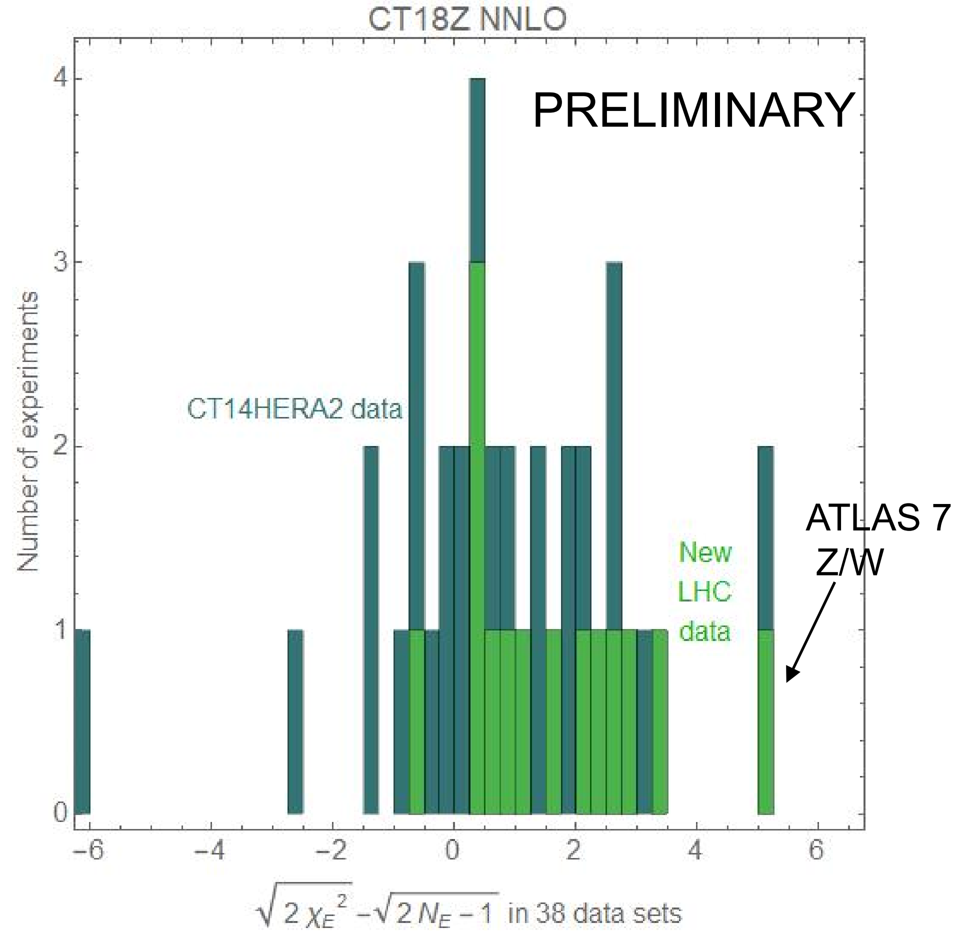
In NNPDF3.1, S_n(HERA I+II) is improved to ≈ 3 by using fitted charm and/or small-x resummation

CT18 (CT18Z) NNLO

⇒ J. Huston's talk, PDF4LHC mtg. 13 (14) LHC experiments with 665 (711) data points

- New LHC experiments tend to have larger S_n
- ATLAS 7 TeV Z, W production has $S_n \approx 5.2,$ included in CT18Z fit only





CT18: theoretical calculations for vector boson production

ID	Obs.	Expt.	fast table	NLO code	K-factors	$\mu_{R,F}$
245 246 250 249	$y_{\mu\mu}, \eta_\mu$ y_{ee} $y_{\mu\mu}, \eta_\mu$ $A(\mu)$	LHCb7ZW LHCb8Z LHC8ZW CMS8W	APPLgrid	MCFM/aMCfast	MCFM/FEWZ	$M_{Z,W}$
253	p_T^{ll}	ATL8Z	APPLgrid	MCFM	NNLOJet	M_T^{ll}
201 203 204	$\sqrt{\tau}, y$ $\sigma_{pd}/\sigma_{pp}, x_F$ Q, x_F	E605 E866 E866	CTEQ		FEWZ	Q_{ll}
225 227 234 281	$A(e)$ $A(e)$ $A(\mu)$ $A(e)$	CDF1Z CDF2W D02W D02W	CTEQ		ResBos	Q_{ll} M_W
260 261	y_{ll} y_{ll}	D02 CDF2	CTEQ		VRAP	Q_{ll}
266 267 268	$A(\mu)$ $A(e)$ $y_{ll}, \eta_l, A(l)$	CMS7W CMS7W ATL7ZW(2012)	CTEQ		ResBos	M_W $M_{Z,W}$
248	y_{ll}, η_l	ATL7ZW(2016)	APPLgrid	MCFM/aMCfast	MCFM/FEWZ	$M_{Z,W}$

Benchmarking of NNLO theory predictions for vector boson production is a high priority!

Q_T resummation may be important when fitting lepton data with $p_{T,\ell}$ cuts

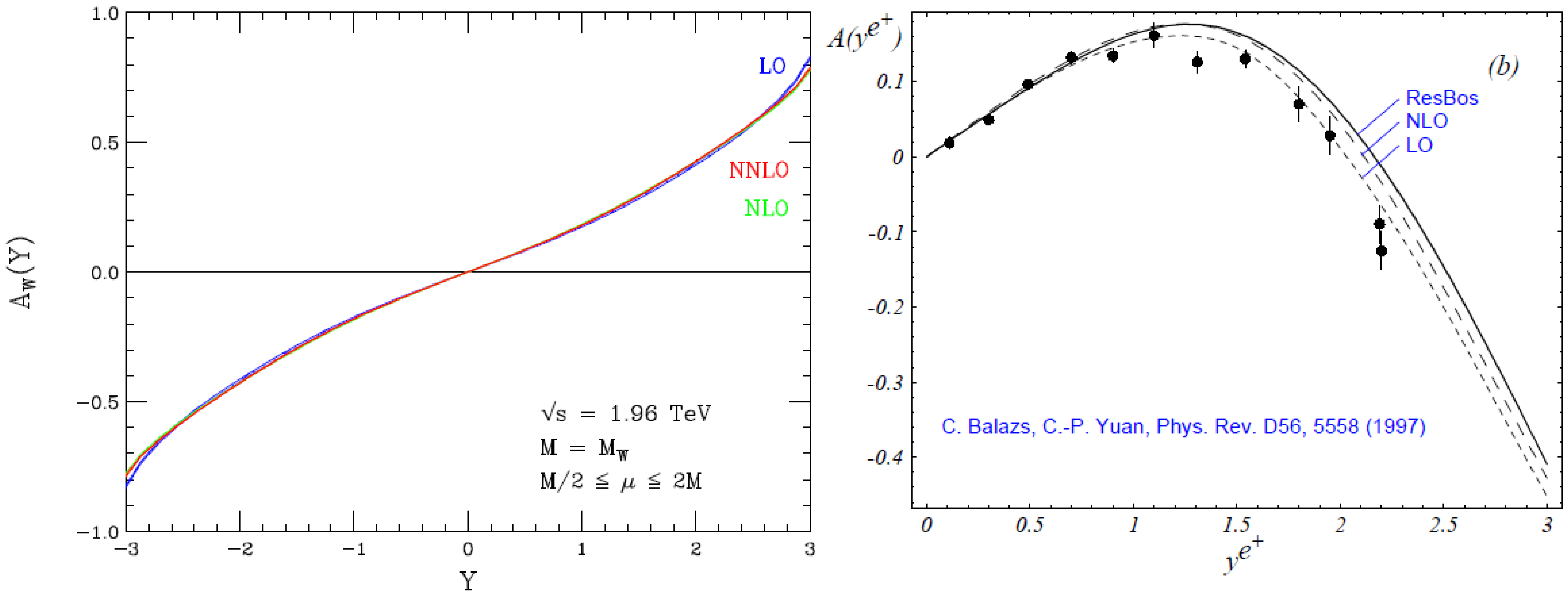


Figure 5. (a) Charge asymmetry $A_{ch}(y)$ of W boson rapidity distributions, $d\sigma/dy$, at various orders of α_s (no acceptance cuts are imposed) [21]; charge asymmetry $A_{ch}(y_\ell)$ in the Tevatron Run-1 for $q_T < 30$ GeV, $p_T^{e^+,\nu} > 25$ GeV in the leading-order, next-to-leading order, and resummed calculations [23].

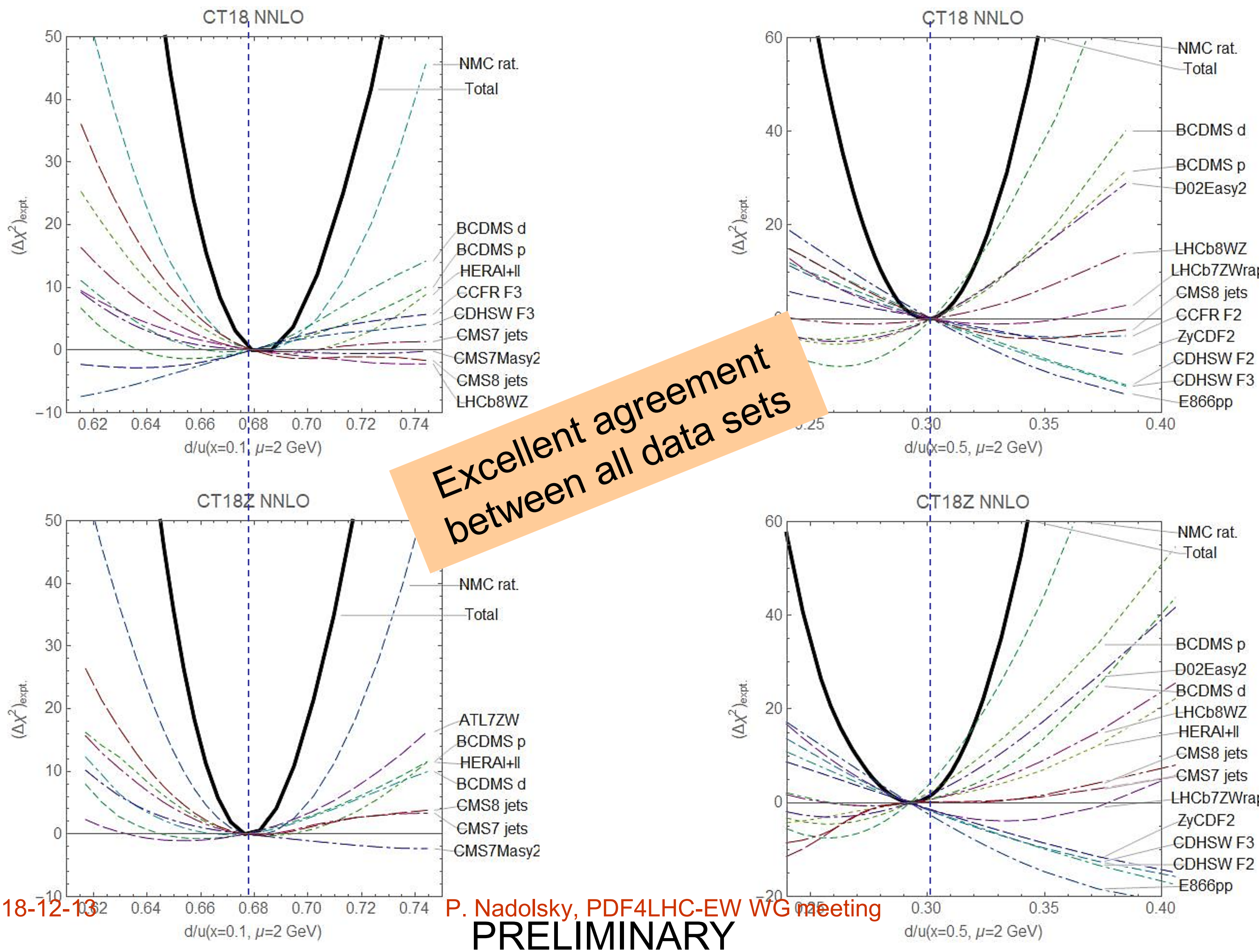
From P. Nadolsky, Theory of W and Z production, hep-ph/0412146

Recommendations for EW precision fits

A realistic precise fit must

1. include the majority of pre-LHC data from DIS, Drell-Yan pair production, and jet production
2. perform a benchmark comparison of (fast) NNLO theory cross sections for DY and other relevant processes
3. resolve tensions affecting EW observables within the HERA I+II data set, and various tensions involving non-LHC, ATLAS, CMS, and LHCb experiments
4. provide a trustworthy estimate of the parametrization uncertainty (10,000 x ? toys for $\sim 10 - 100$ parametrization models)

Lagrange Multiplier scan: d/u at x=0.1 and 0.5

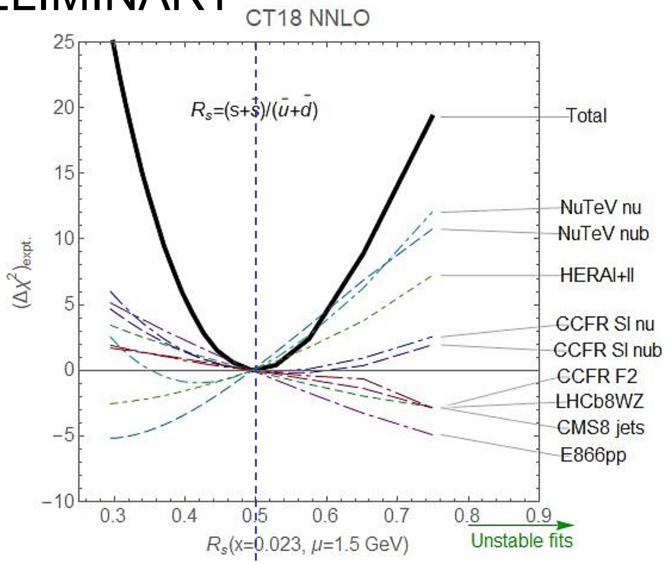


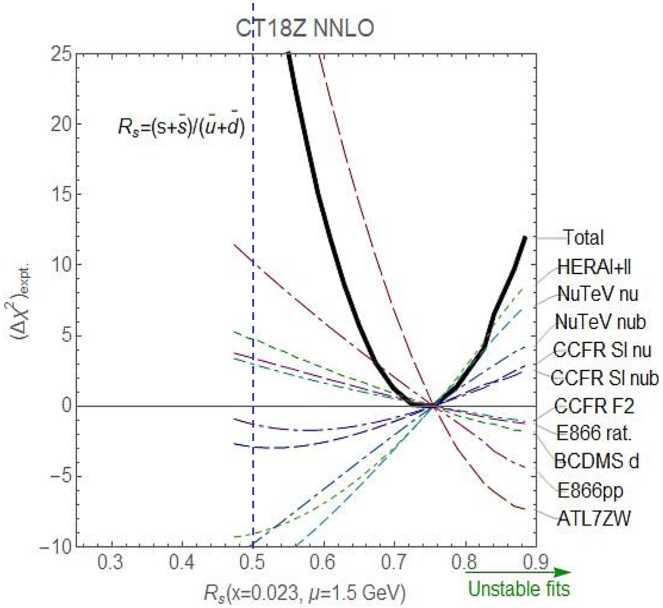
Excellent agreement between all data sets

PRELIMINARY

Lagrange Multiplier scan: $R_s(x = 0.023, \mu = 1.5\ GeV)$

PRELIMINARY



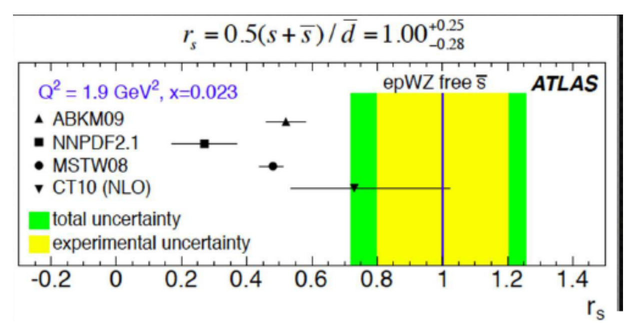


$$R_s(x,\mu) \equiv \frac{s(x,\mu)+\bar{s}(x,\mu)}{\bar{u}(x,\mu)+\bar{d}(x,\mu)}$$

Upper/lower rows: CT18/CT18Z

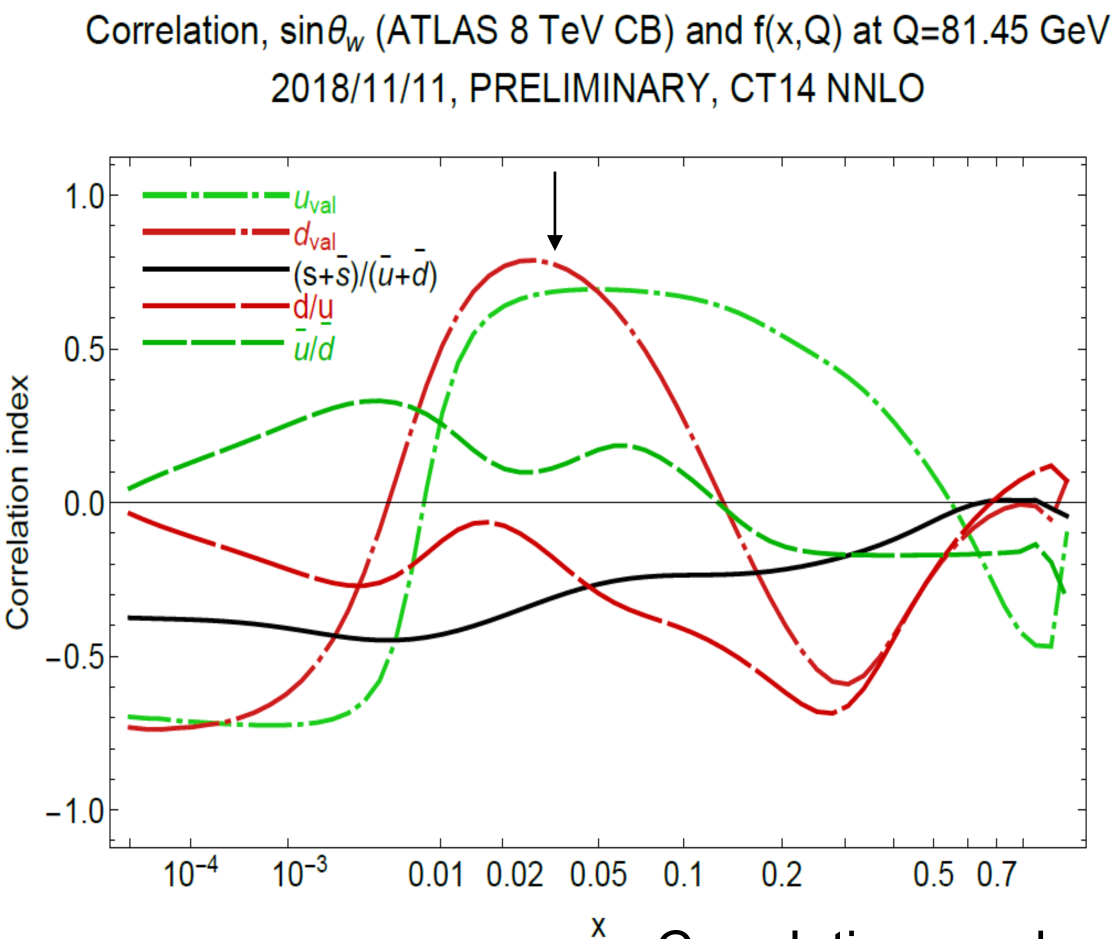
The CT18Z strangeness is increased primarily as a result of including the ATLAS 7 TeV W/Z production data (not in CT18), as well as because of using the DIS saturation scale in $m_c^{pole} = 1.4$ GeV

In either CT18 or CT18Z fit, observe instability in the fits for $R_s > 1$ at $x = 0.01 - 0.1$



Hessian correlation for $\sin^2\theta_w$ at 8 TeV

Presented at the EW precision subgroup meeting, Nov. 13, 2018



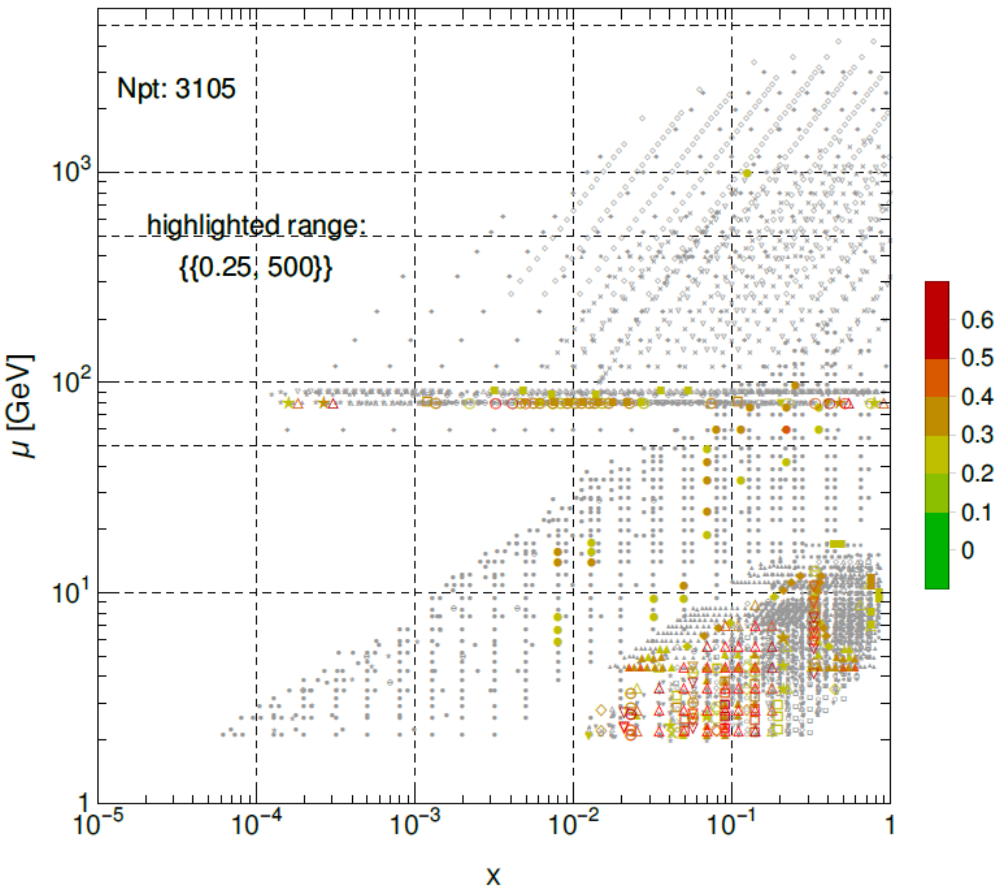
Strongest correlations of s2w with u_{val}, d_{val} at $x = 0.01 - 0.2$

weak correlations with $\bar{u}$, $\bar{\mathrm{d}}$, $\bar{s}$, g

Correlations and sensitivities (next slide) elucidate PDF properties using **published** error PDF sets

Sensitivity of CT14 experiments to s2w

| S_f | for s2w, ATLAS 8 TeV (prel., CB), CT14NNLO

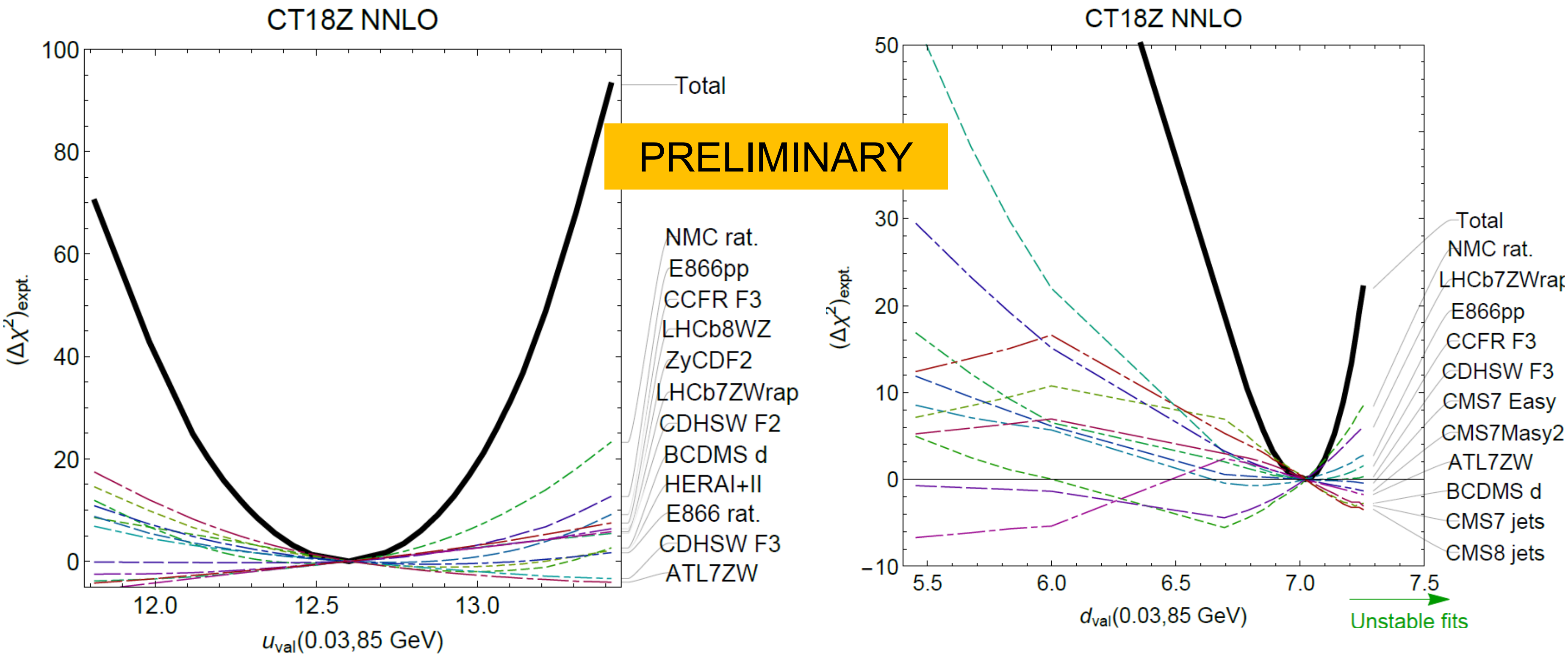


No significant difference between CB and CBF s2w samples

Based on the **PDFSense** [arXiv:1803.02777] analysis, the most sensitive CT14 data sets to s2w are

- **combined HERA1 DIS [most sensitive]**
- CCFR νp DIS $F_{3,2}$
- BCDMS $F_2^{p,d}$
- NMC ep, ed DIS
- CDHSW νA DIS
- NuTeV $\nu A \to \mu\mu X$
- CCFR $\nu A \to \mu\mu X$
- E866 $pp \to \ell^+\ell^- X$
- ATLAS 7 TeV W/Z (35 pb^{-1})
- …

CT18Z Lagrange Multiplier scan on $u_{val}(0.03, 85\ GeV)$ and $d_{val}(0.03, 85\ GeV)$



In the central rapidity region at 8 TeV, u_{val} and d_{val} **refitted** with the new LHC data are constrained by the NMC ed/ep ratio, a combination of DY, DIS experiments, and even jet experiments

CTEQ-TEA recommendations for LHC DY measurements DRAFT, 2018-12-13, page 1

1. CT18 NNLO or CT14HERA2 NNLO
2. CT18 fits find contradictory preferences for strangeness $x \geq 10^{-3}$ between fitted (SI)DIS experiments, on one hand, and some LHC experiments, especially ATLAS W/Z production measurements and to some extent LHCb W/Z measurements. Benchmarking of LHC measurements and theoretical predictions, as well as new (SI)DIS experiments can be highly effective for resolving these tensions.
3. Theoretical programs for DY processes used in CT18 NNLO are summarized in a next slide. NLO cross sections The NNLO cross sections for DY are obtained by multiplying fast NLO cross sections by tabulated point-by-point NNLO/NLO ratios (close to 1 in DY processes) computed for a recent CTXX PDF set. Parton shower effects are very limited, especially when NNLO predictions are used.

NEW

4. Alternative candidate fits of the CT18 NNLO analysis estimate the QCD scale and numerical uncertainties in high-p_T Z production. In our opinion, NNLO theoretical uncertainties are under good control in the fitted region $50 < p_{TZ} < 150$ GeV of the high-p_T Z production data in the CT18 NNLO analysis.
5. The photon PDFs do not significantly affect the inclusive QCD observables included in the CT18 NNLO analysis.

CTEQ-TEA recommendations for LHC DY measurements DRAFT, 2018-12-13, page 2

6. When it is relevant, QCD predictions using CT18/CT14 PDFs must use the SACOT-chi scheme and the same charm and bottom mass values as those used to fit the CT18 PDFs. For the LHC observables with all scales much larger than the c, b masses, the S-ACOT-chi hard cross section coincides with the zero-mass MSbar hard cross section. On the other hand, the mass effects may be relevant in W/Z p_T distributions in c, b channels at $p_T^2 \lesssim m_{c,b}^2$. A comprehensive study of the power-suppressed/intrinsic/fitted charm distribution is published in **JHEP 1802 (2018) 059 / arXiv:1707.00657.** CTEQ-TEA does not see it mandatory to use the fitted charm parametrizations throughout. The PDFs with fitted charm such as CT14 IC or NNPDF3.1 do not provide a better theoretical framework than the standard CT14 PDFs. A large part of the fitted charm PDF may arise from twist-4 contributions that are unique to low-Q DIS.
7. The TMD effects are negligible in the recent CTEQ-TEA analyses.
8. No, various kinds of parametrization and methodological uncertainties are accounted for in the CTEQ-TEA PDF errors and are studied regularly as a part of the CTEQ-TEA analysis.

CTEQ-TEA recommendations for LHC DY measurements DRAFT, 2018-12-13, page 3

9. As of 2018, we do not recommend to fit the PDFs only to the LHC or DY data. The most significant constraints arise from other experiments, such as fixed-target DIS. It is ok to perform this type of study with a reduced number of data sets as a benchmarking exercise among the PDF groups, but the resulting PDFs will be less accurate/precise than the global PDF fits.
10. To a great degree, the important uncertainties, those due to the experimental errors of the datasets included in the fit, are already completely correlated. Correlation of other issues, such as parameterizations/scale choices can be studied.
11. If the PDF sets include the data, but do not agree with the data, and the other PDF sets do, then it is crucial to understand the source of the disagreement.
12. If the measurements do not have clearly defined systematic errors (in the modern sense), then it is justified to not use them in a global PDF fit. If the data sets are in strong tension with the other data sets used in a global fit, then they can be excluded. Of course, this happens on a case-by-case basis.

CTEQ-TEA recommendations for LHC DY measurements DRAFT, 2018-12-13, page 4

13. The Hessian and MC approaches are complementary. In recent years, the PDF groups have gained a great deal of experience in converting between Hessian and MC replica PDFs, strengthening the understanding of both. The Hessian PDFs are sufficient for the majority of estimates of PDF uncertainty in the case of sufficient experimental constraints. The MC error PDFs are useful in the case of weak experimental constraints or persistent non-Gaussian effects.
14. Conceptual foundations of PDF reweighting have not been explored sufficiently, which may result in its spurious applications. This area needs additional exploration before PDF reweighting can be safely used in high-stake situations such as in item 11.

Thank you!

CT18 in a nutshell

- Start with CT14-HERAII (HERAII combined data released after publication of CT14)
- Examine a wide range of PDF parameterizations
- Use as much relevant LHC data as possible using applgrid/fastNLO interfaces to data sets, with NNLO/NLO K-factors, or fastNNLO tables in the case of top pair production
- PDFSense (arXiv:1803.02777) to determine quantitatively which data will have impact on global PDF fit
- ePump (arXiv:1806.07950) on quickly exploring the impact of data prior to global fit within the Hessian approximation
  - good agreement between ePump results and global fit (see talk in March PDF4LHC meeting)
- Implement a parallelization of the global PDF fitting to allow for faster turn-around time
- Lagrange Multiplier studies to examine constraints of specific data sets on PDF distributions, and (in some cases) the tensions (useful information; will spend some time on this)

Experiments in the CT14 HERA2 fit

ID#	Experimental dataset		N_d
101	BCDMS F_2^p	[47]	337
102	BCDMS F_2^d	[48]	250
104	NMC F_2^d/F_2^p	[49]	123
108	CDHSW F_2^p	[50]	85
109	CDHSW F_3^p	[50]	96
110	CCFR F_2^p	[51]	69
111	CCFR xF_3^p	[52]	86
124	NuTeV $\nu\mu\mu$ SIDIS	[40]	38
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS	[40]	33
126	CCFR $\nu\mu\mu$ SIDIS	[41]	40
127	CCFR $\bar{\nu}\mu\mu$ SIDIS	[41]	38
145	H1 σ_r^b (57.4 pb^{-1})	[53][54]	10
147	Combined HERA charm production (1.504 fb^{-1})	[39]	47
160	HERA1+2 Combined NC and CC DIS (1 fb^{-1})	[6]	1120
169	H1 F_L (121.6 pb^{-1})	[55]	9

ID#	Experimental dataset		N_d
201	E605 DY	[56]	119
203	E866 DY, $\sigma_{pd}/(2\sigma_{pp})$	[57]	15
204	E866 DY, $Q^3 d^2\sigma_{pp}/(dQdx_F)$	[58]	184
225	CDF Run-1 $A_e(\eta^e)$ (110 pb^{-1})	[59]	11
227	CDF Run-2 $A_e(\eta^e)$ (170 pb^{-1})	[60]	11
234	DØ Run-2 $A_\mu(\eta^\mu)$ (0.3 fb^{-1})	[61]	9
240	LHCb 7 TeV W/Z muon forward-η Xsec (35 pb^{-1})	[62]	14
241	LHCb 7 TeV W $A_\mu(\eta^\mu)$ (35 pb^{-1})	[62]	5
260	DØ Run-2 Z $d\sigma/dy_Z$ (0.4 fb^{-1})	[63]	28
266	CMS 7 TeV $A_\mu(\eta)$ (4.7 fb^{-1})	[64]	11
267	CMS 7 TeV $A_e(\eta)$ (0.840 fb^{-1})	[65]	11
268	ATLAS 7 TeV W/Z Xsec, $A_\mu(\eta)$ (35 pb^{-1})	[66]	41
281	DØ Run-2 $A_e(\eta)$ (9.7 fb^{-1})	[67]	13
504	CDF Run-2 incl. jet $(d^2\sigma/dp_T^j dy_j)$ (1.13 fb^{-1})	[36]	72
514	DØ Run-2 incl. jet $(d^2\sigma/dp_T^j dy_j)$ (0.7 fb^{-1})	[37]	110
535	ATLAS 7 TeV incl. jet $(d^2\sigma/dp_T^j dy_j)$ (35 pb^{-1})	[68]	90
538	CMS 7 TeV incl. jet $(d^2\sigma/dp_T^j dy_j)$ (5 fb^{-1})	[69]	133

Candidate experiments in the CTEQ-TEA fit

ID#	Experimental dataset		N_d
245	LHCb 7 TeV Z/W muon forward-η Xsec (1.0 fb^{-1})	[70]	33
246	LHCb 8 TeV Z electron forward-η $d\sigma/dy_Z$ (2.0 fb^{-1})	[71]	17
247	ATLAS 7 TeV $d\sigma/dp_T^Z$ (4.7 fb^{-1})	[72]	8
249	CMS 8 TeV W muon, Xsec, $A_\mu(\eta^\mu)$ (18.8 fb^{-1})	[73]	33
250	LHCb 8 TeV W/Z muon, Xsec, $A_\mu(\eta^\mu)$ (2.0 fb^{-1})	[74]	42
252	ATLAS 8 TeV Z $(d^2\sigma/d\|y\|_{ll}dm_{ll})$ (20.3 fb^{-1})	[75]	48
253	ATLAS 8 TeV $(d^2\sigma/dp_T^Z dm_{ll})$ (20.3 fb^{-1})	[76]	45
542	CMS 7 TeV incl. jet, R=0.7, $(d^2\sigma/dp_T^j dy_j)$ (5 fb^{-1})	[34]	158
544	ATLAS 7 TeV incl. jet, R=0.6, $(d^2\sigma/dp_T^j dy_j)$ (4.5 fb^{-1})	[33]	140
545	CMS 8 TeV incl. jet, R=0.7, $(d^2\sigma/dp_T^j dy_j)$ (19.7 fb^{-1})	[35]	185
565	ATLAS 8 TeV $t\bar{t}$ $d\sigma/dp_T^t$ (20.3 fb^{-1})	[38]	8
566	ATLAS 8 TeV $t\bar{t}$ $d\sigma/dy_{<t/\bar{t}>}$ (20.3 fb^{-1})	[38]	5
567	ATLAS 8 TeV $t\bar{t}$ $d\sigma/dm_{t\bar{t}}$ (20.3 fb^{-1})	[38]	7
568	ATLAS 8 TeV $t\bar{t}$ $d\sigma/dy_{t\bar{t}}$ (20.3 fb^{-1})	[38]	5

N_d is the number of data points

New LHC datasets for CT18

- 245 1505.07024 LHCb Z (W) muon rapidity at 7 TeV(applgrid)
- 246 1503.00963 LHCb 8 TeV Z rapidity (applgrid);
- 249 1603.01803 CMS W lepton asymmetry at 8 TeV (applgrid)
- 250 1511.08039 LHCb Z (W) muon rapidity at 8 TeV(applgrid)
- 253 1512.02192 ATLAS 7 TeV Z pT (applgrid)
- 542 1406.0324 CMS incl. jet at 7 TeV with R=0.7 (fastNLO)
- 544 1410.8857 ATLAS incl. jet at 7 TeV with R=0.6 (applgrid)
- 545 1609.05331 CMS incl. jet at 8 TeV with R=0.7 (fastNLO)
- 565 1511.04716 ATLAS 8 TeV tT pT diff. distributions (fastNNLO)
- 567 1511.04716 ATLAS 8 TeV tT mtT diff. distributions (fastNNLO)
- 573 1703.01630 CMS 8 TeV tT (pT , yt) double diff. distributions (fastNNLO)

- 248 1612.03016 ATLAS 7 TeV Z and W rapidity (applgrid)->CT18Z
  - also uses special small x factorization scale, m_c=1.4 GeV
  - serious changes in PDFs, so warrants a separate PDF

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 $\vec{a}$

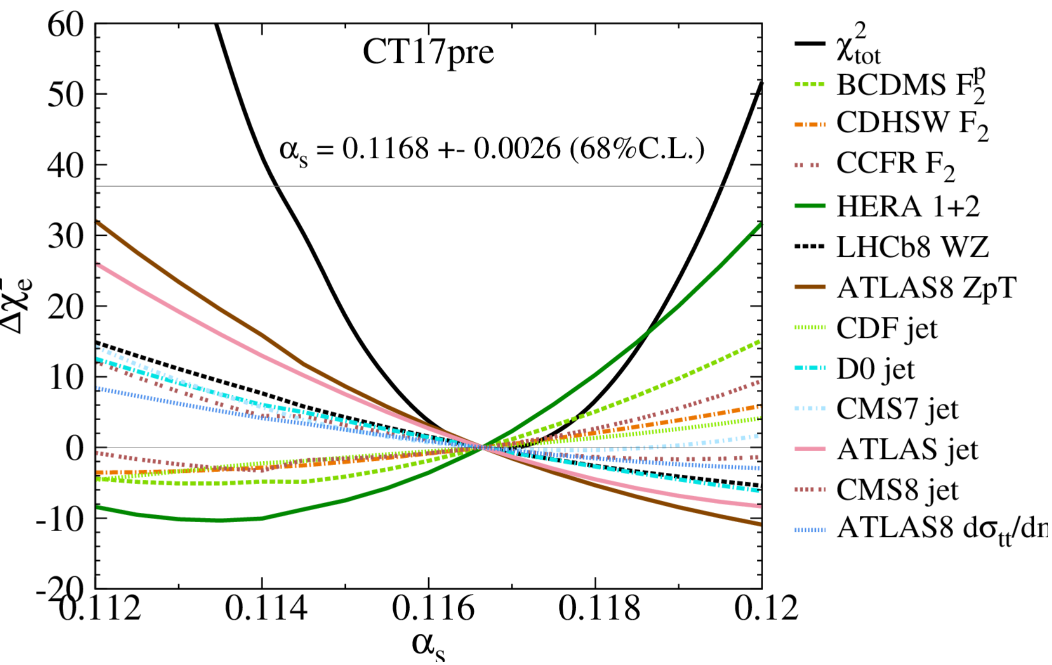
D_i^{sh} is the data value **including the optimal systematic shift**

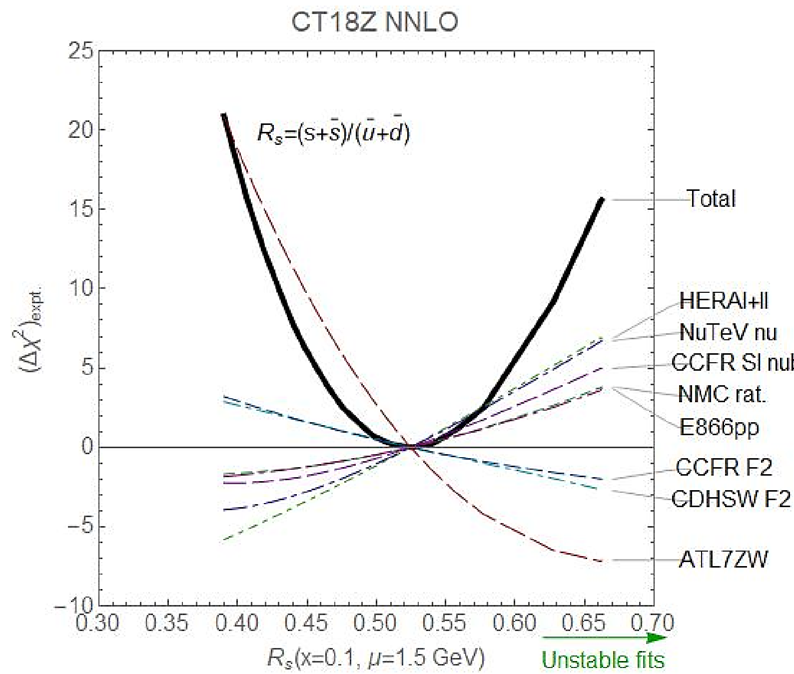
$$D_i^{sh}(\vec{a}) = D_i - \sum_{\alpha=1}^{N_\lambda} \beta_{i\alpha}\bar{\lambda}_\alpha(\vec{a})$$

s_i is the uncorrelated error

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

Lagrange Multiplier Scans



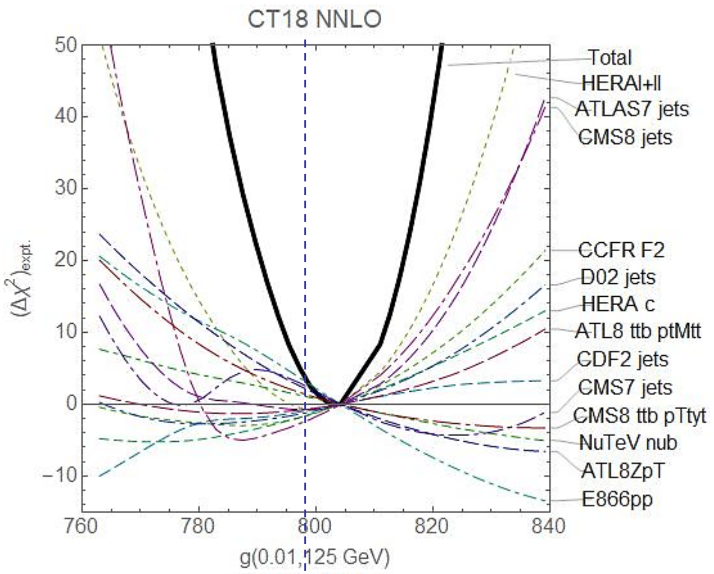


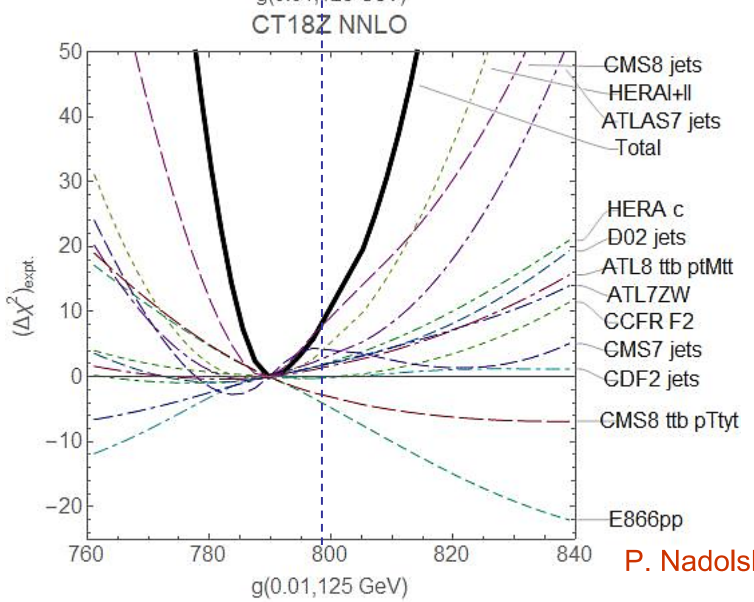
$\alpha_s(m_Z)$ from global fit closer to 0.117 than to 0.118

- LHCb W and Z (7,8 TeV) data prefer a larger strange in the small x region
- ATLAS 8 TeV Z pT data prefer a slightly larger strange
- NuTeV dimuon data strongly prefer smaller strange
- ATLAS 7 TeV precision W/Z strongly prefer larger strangeness

There is an alternative version of CT18 (CT18Z) including these ATLAS data ←

Lagrange Multiplier scan: $g(0.01, 125\ GeV)$





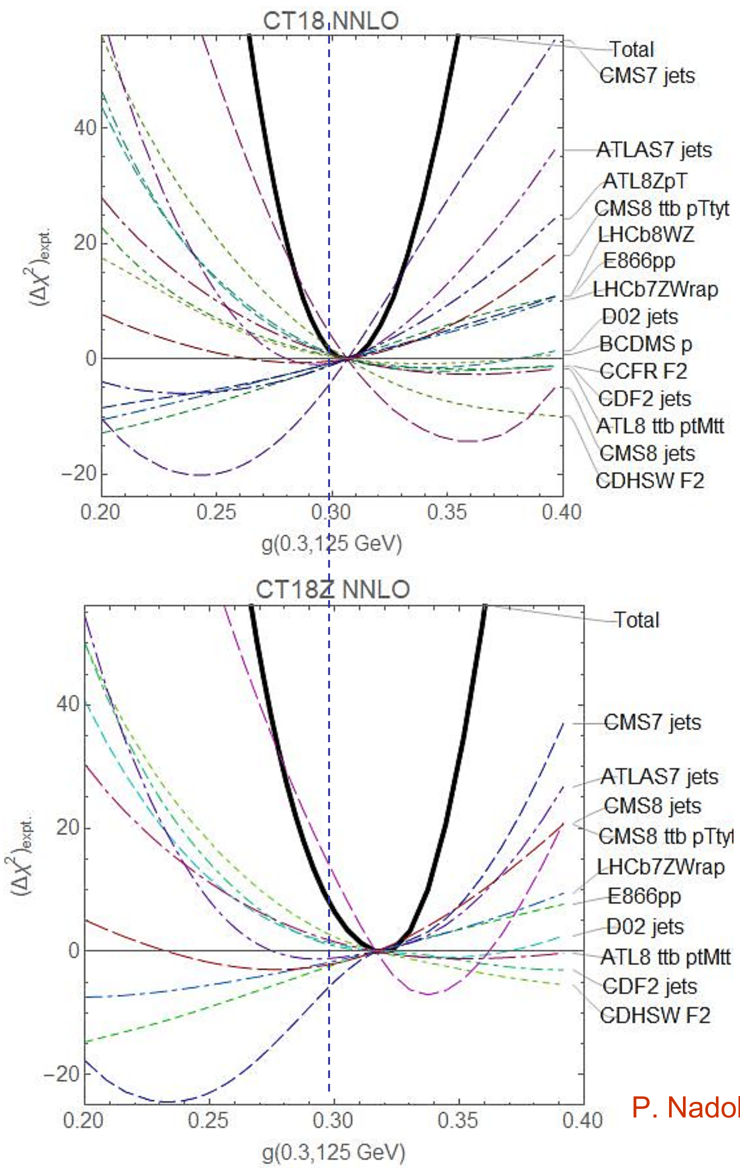
Upper row: CT18

- HERAI+II data set provides the dominant constraint, followed by ATLAS, CDF2, CMS, D02 jet production, HERA charm,…
- $t\bar{t}$ double-diff. cross sections provide weaker constraints

Lower row: CT18Z

- CT18Z: a 1% lower NNLO gluon in the Higgs production region than for CT14/CT18 as a result of
  - higher charm mass, $m_c^{\rm pole} = 1.4$ GeV
  - Including ATLAS7 W/Z production
  - a special factorization scale in DIS that mildly improves χ^2 and approximates effect of small-x resummation

Lagrange Multiplier scan: $g(0.3, 125\ GeV)$



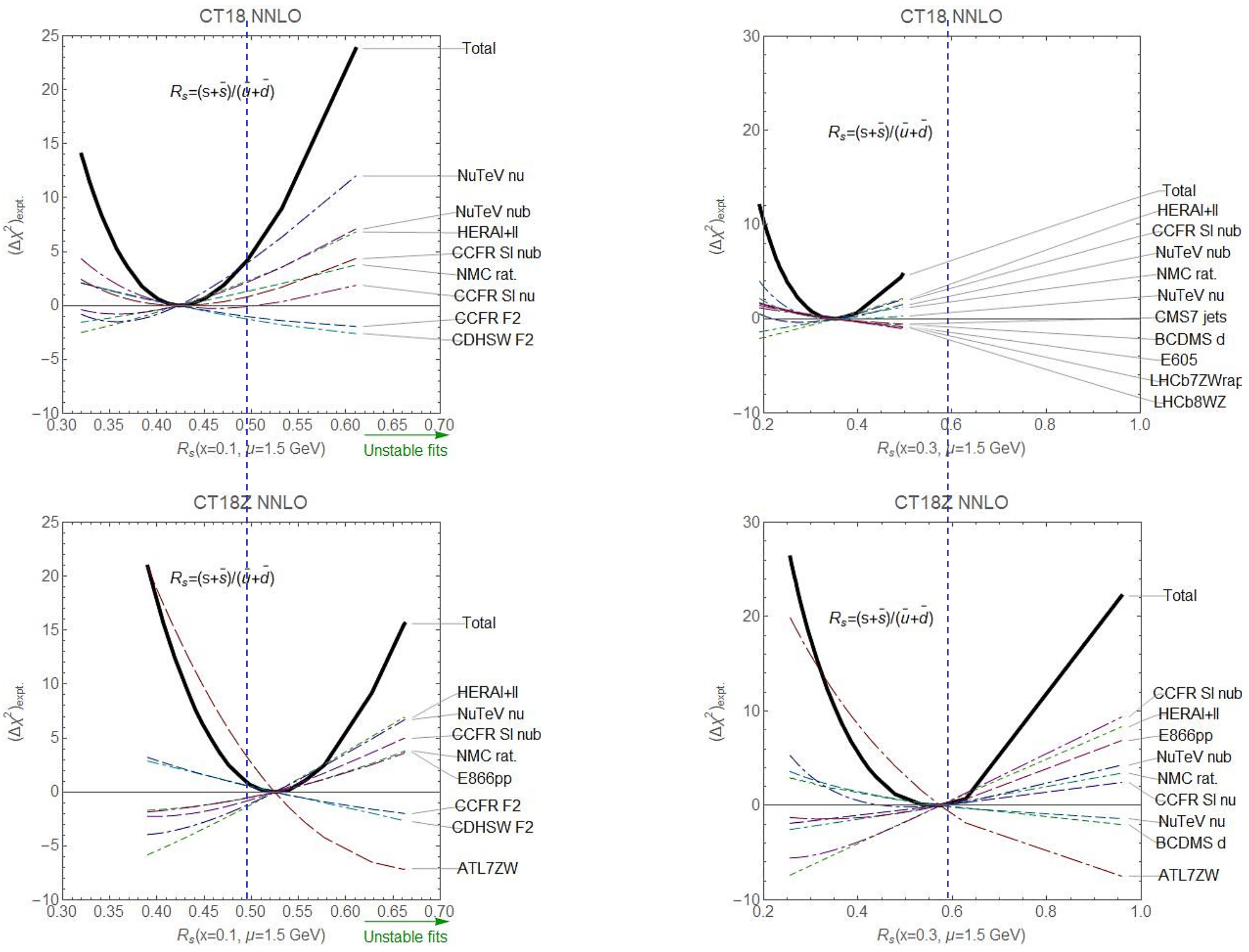
Upper/lower rows: CT18/CT18Z

Observe opposite pulls from ATLAS7/CMS7 jet production, on one hand, and CMS8 jet production, on the other hand

Similarly, ATLAS $t\bar{t}$ distributions $\mathrm{d}^2\sigma/(dp_{T,t}dm_{t\bar{t}})$ and CMS $t\bar{t}$ distributions $\mathrm{d}^2\sigma/(dp_{T,t}dy_{t,ave})$ at 8 TeV impose weak opposite pulls

Constraints from ATLAS 8 $Z\ p_T$ production data are moderate and still affected by NNLO scale uncertainty

Lagrange Multiplier scan: R_s at x=0.1 and 0.3



Theoretical uncertainty in DIS

Mild NNLO theoretical uncertainties in **large** DIS data sets have a non-negligible overall effect on the global χ^2

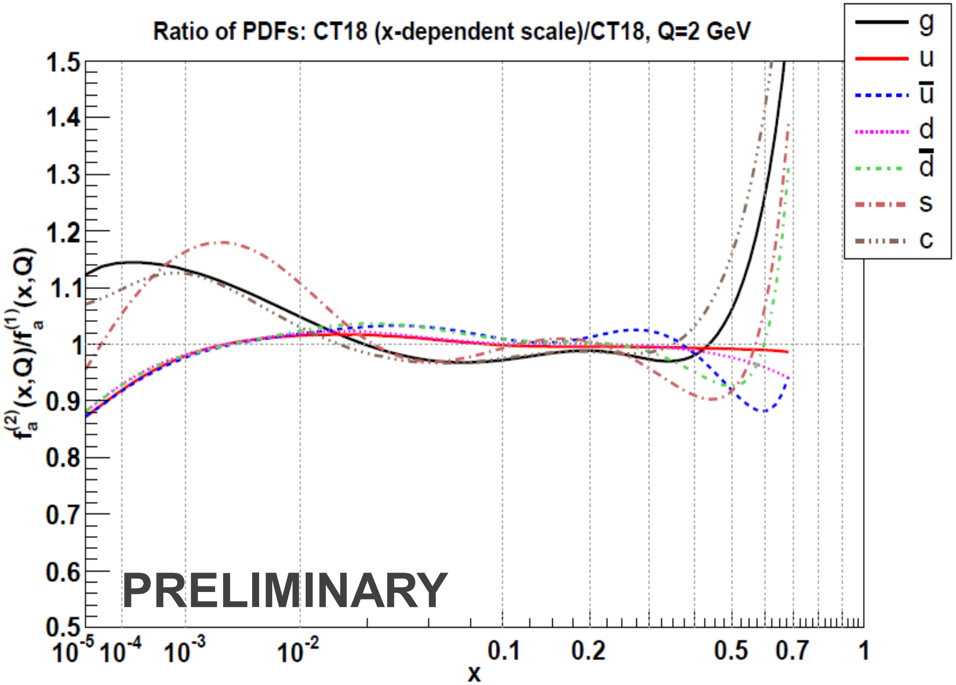
The following **x-dependent** factorization scale at NNLO improves description of CTEQ-TEA DIS data sets by mimicking

- missing N3LO terms at x>0.001
- small-x/saturation terms at x<0.001

$$\mu_{DIS,X}^2 = 0.8^2 \left(Q^2 + \frac{0.3\ GeV^2}{x^{0.3}}\right)$$

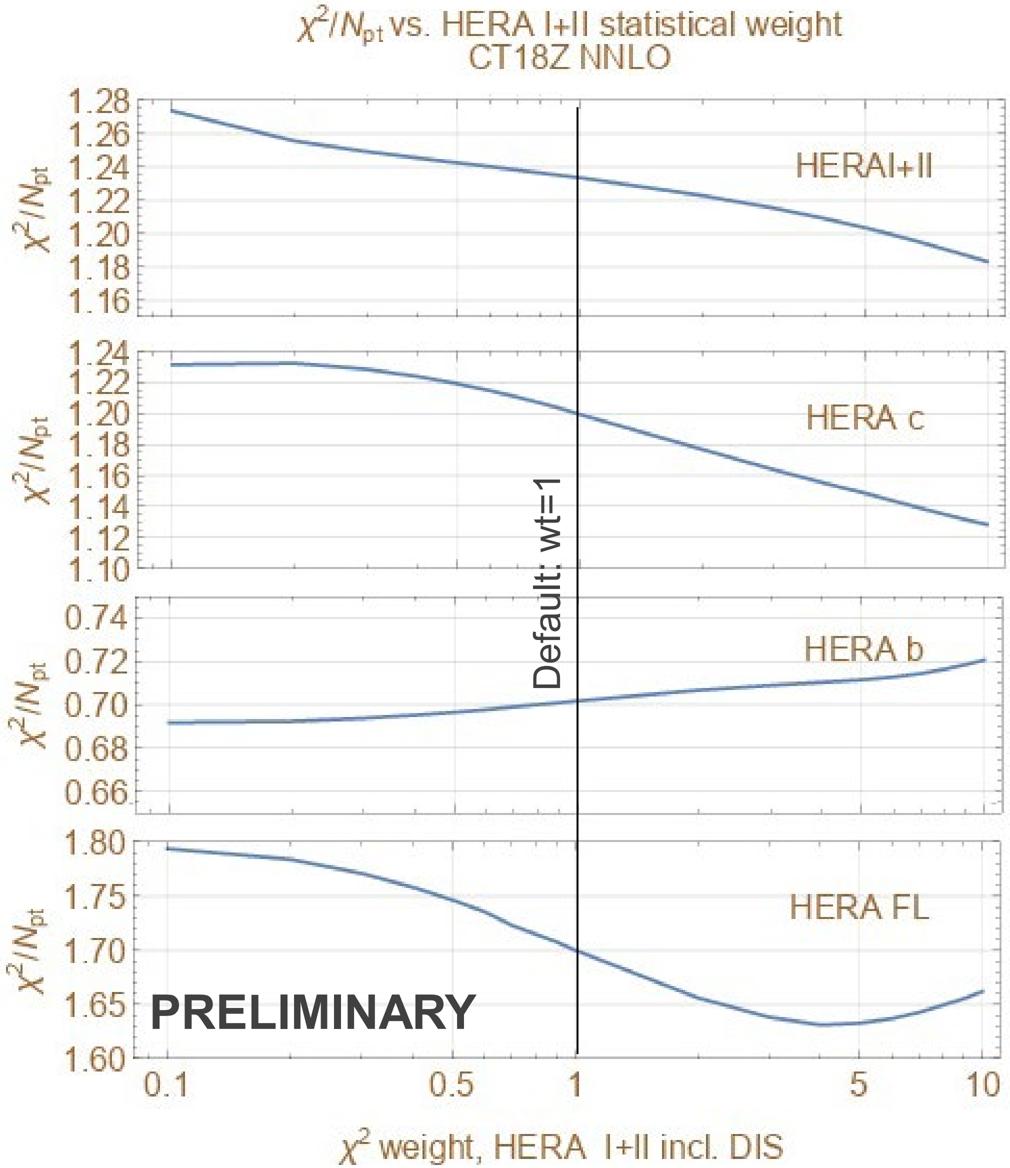
CT18Z uses a combination of $\mu_{DIS,X}$ (preferred by DIS) and increased $m_c^{pole} = 1.4$ GeV (preferred by LHC vector boson production, disfavored by DIS)

X-dependent DIS scale, effect on PDFs



Using $\mu_{DIS,X}$ in a **fixed-order** NNLO cross section bears similar effect to small-x resummation/saturation. In particular, the gluon and strange PDF are enhanced at $x < 10^{-2}$

Varied statistical weight for HERA I+II DIS set



The CT18Z fits using the $\mu_{DIS,X}$ scale reproduce many features of NNLO-NNLx fits with $\ln(1/x)$ resummation by the NNPDF [arXiv:1710.05935] and xFitter [1802.0064] groups.

Left: when the statistical weight for the HERA I+II data set is increased to $wt = 10$ to suppress pulls from the other experiments, χ^2_{CT18Z}/N_{pt} for HERA I+II DIS **and** HERA heavy-quark production decreases to about the same levels as in NNLO+NNLx fits to HERA DIS only by NNPDF and xFitter.