CT update

J. Huston
MSU
for the CTEQ-TEA (Tung et al) collaboration

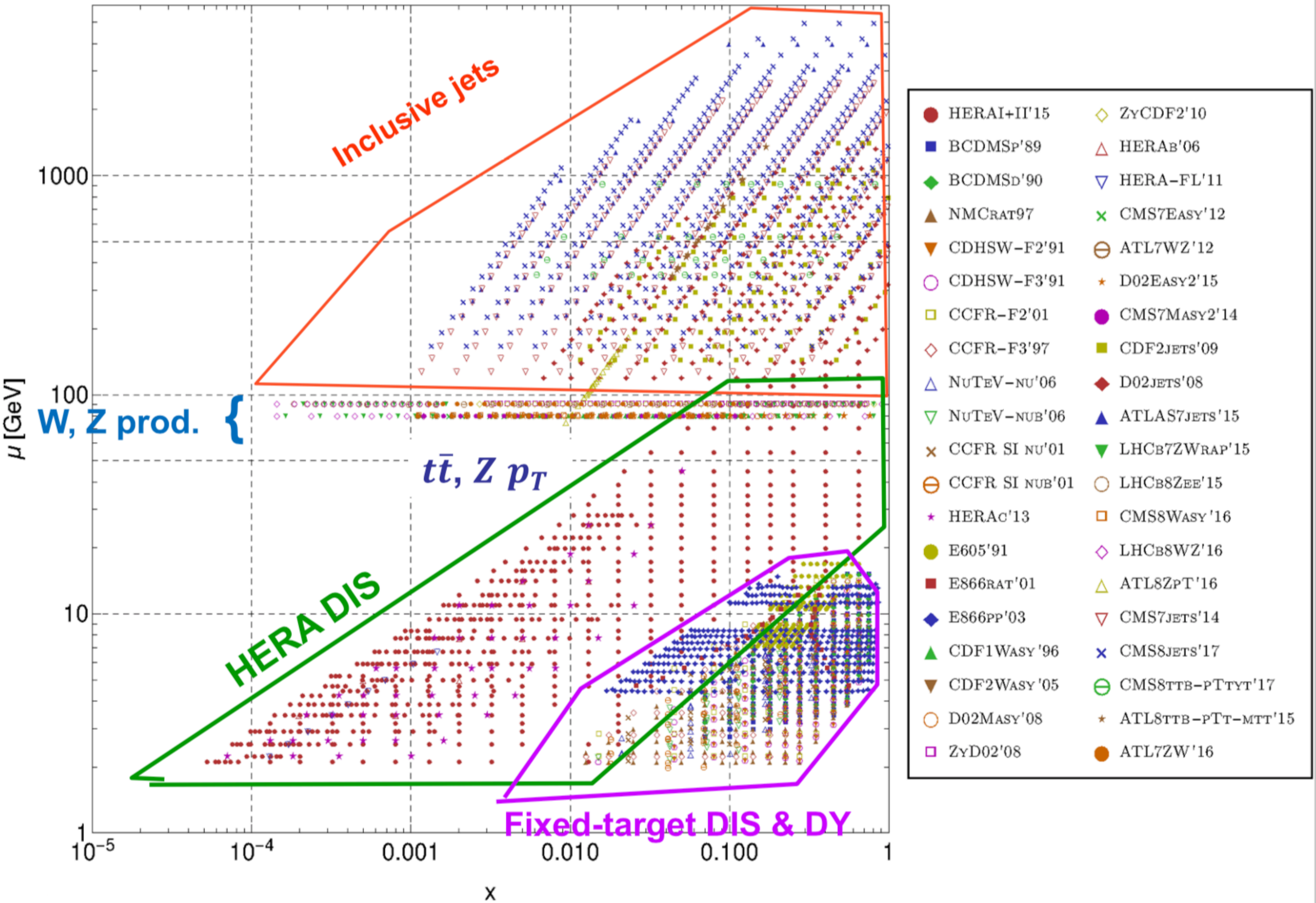
PDF4LHC meeting Durham Sept 2019

Thanks to Pavel Nadolsky from whom I've borrowed a number of slides

CT18 in a nutshell

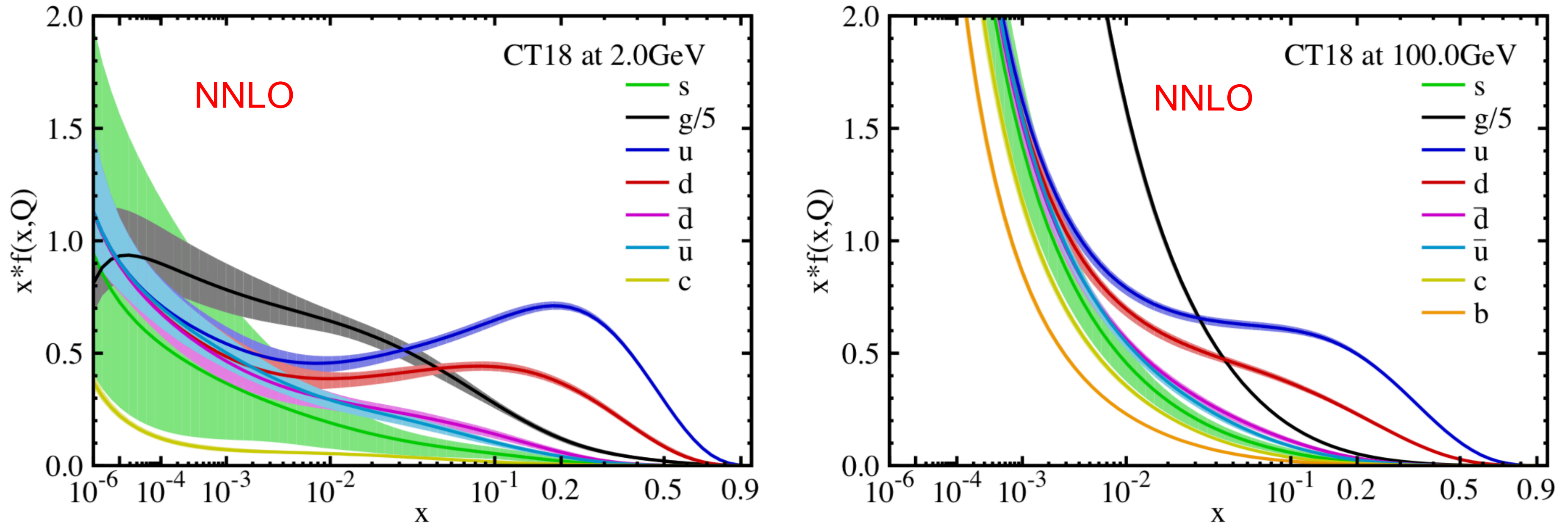
- Start with CT14-HERAII (HERAII combined data released after publication of CT14)
- Examine a wide range of PDF parameterizations
- Use as much relevant LHC data as possible using applgrid/fastNLO interfaces to data sets, with NNLO/NLO K-factors, or fastNNLO tables in the case of top pair production
- PDFSense (arXiv:1803.02777), L2 sensitivity 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
- 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)

Experimental data in CT18 PDF analysis



CT18...

- Main product is CT18 (NNLO, NLO,LO)



- Including full data set except for ATLAS 7 TeV W/Z, which has a sizeable impact on the global fit (strange quark)

...and family

PDF ensemble	Factorization scale in DIS	ATLAS 7 Z/W data included?	CDHSW $F_2^{p,d}$ data included?	Pole charm mass, GeV
CT18	$\mu^2_{F,DIS} = Q^2$	No	Yes	1.3
CT18X	$\mu^2_{F,DIS} = 0.8^2 \left(Q^2 + \frac{0.3\ GeV^2}{x^{0.3}} \right)$	No	Yes	1.3
CT18A	$\mu^2_{F,DIS} = Q^2$	Yes	Yes	1.3
CT18Z	$\mu^2_{F,DIS} = 0.8^2 \left(Q^2 + \frac{0.3\ GeV^2}{x^{0.3}} \right)$	Yes	No	1.4

CT18 PDFs available from https://tinyurl.com/ct18pdfs-1

ID#	**Experimental data set**		$N_{pt,n}$	χ^2_n	$\chi^2_n/N_{pt,n}$	S_n
160	HERAI+II 1 fb^{-1}, H1 and ZEUS NC and CC $e^{\pm}p$ reduced cross sec. comb.	[10]	1120	1408.7(1377.8)	1.3(1.2)	5.7(5.1)
101	BCDMS F_2^p	[22]	337	373.7(383.8)	1.1(1.1)	1.4(1.8)
102	BCDMS F_2^d	[23]	250	280.4(287.0)	1.1(1.1)	1.3(1.6)
104	NMC F_2^d/F_2^p	[24]	123	125.7(116.2)	1.0(0.9)	0.2(-0.4)
108	CDHSW† F_2^p	[25]	85	85.6(86.8)	1.0(1.0)	0.1(0.2)
109	CDHSW† F_3^p	[25]	96	86.5(85.6)	0.9(0.9)	-0.7(-0.7)
110	CCFR F_2^p	[26]	69	78.8(76.0)	1.1(1.1)	0.9(0.6)
111	CCFR xF_3^p	[27]	86	33.8(31.4)	0.4(0.4)	-5.2(-5.6)
124	NuTeV $\nu\mu\mu$ SIDIS	[28]	38	18.5(30.3)	0.5(0.8)	-2.7(-0.9)
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS	[28]	33	38.5(56.7)	1.2(1.7)	0.7(2.5)
126	CCFR $\nu\mu\mu$ SIDIS	[29]	40	29.9(35.0)	0.7(0.9)	-1.1(-0.5)
127	CCFR $\bar{\nu}\mu\mu$ SIDIS	[29]	38	19.8(18.7)	0.5(0.5)	-2.5(-2.7)
145	H1 σ_r^b	[30]	10	6.8(7.0)	0.7(0.7)	-0.6(-0.6)
147	Combined HERA charm production	[31]	47	58.3(56.4)	1.2(1.2)	1.1(1.0)
169	H1 F_L	[16]	9	17.0(15.4)	1.9(1.7)	1.7(1.4)
201	E605 Drell-Yan process	[32]	119	103.4(102.4)	0.9(0.9)	-1.0(-1.1)
203	E866 Drell-Yan process $\sigma_{pd}/(2\sigma_{pp})$	[33]	15	16.1(17.9)	1.1(1.2)	0.3(0.6)
204	E866 Drell-Yan process $Q^3 d^2\sigma_{pp}/(dQdx_F)$	[34]	184	244.4(239.7)	1.3(1.3)	2.9(2.7)
225	CDF Run-1 electron A_{ch}, $p_{T\ell} > 25$ GeV	[35]	11	9.0(9.3)	0.8(0.8)	-0.3(-0.2)
227	CDF Run-2 electron A_{ch}, $p_{T\ell} > 25$ GeV	[36]	11	13.5(13.4)	1.2(1.2)	0.6(0.6)
234	DØ Run-2 muon A_{ch}, $p_{T\ell} > 20$ GeV	[37]	9	9.1(9.0)	1.0(1.0)	0.2(0.1)
260	DØ Run-2 Z rapidity	[38]	28	16.9(18.7)	0.6(0.7)	-1.7(-1.3)
261	CDF Run-2 Z rapidity	[39]	29	48.7(61.1)	1.7(2.1)	2.2(3.3)
266	CMS 7 TeV 4.7 fb^{-1}, muon A_{ch}, $p_{T\ell} > 35$ GeV	[40]	11	7.9(12.2)	0.7(1.1)	-0.6(0.4)
267	CMS 7 TeV 840 pb^{-1}, electron A_{ch}, $p_{T\ell} > 35$ GeV	[41]	11	11.8(16.1)	1.1(1.5)	0.3(1.1)
268	ATLAS 7 TeV 35 pb^{-1} W/Z cross sec., A_{ch}	[42]	41	44.4 50.6)	1.1(1.2)	0.4(1.1)
281	DØ Run-2 9.7 fb^{-1} electron A_{ch}, $p_{T\ell} > 25$ GeV	[43]	13	22.8(20.5)	1.8(1.6)	1.7(1.4)
504	CDF Run-2 inclusive jet production	[44]	72	122.4(117.0)	1.7(1.6)	3.5(3.2)
514	DØ Run-2 inclusive jet production	[45]	110	113.8(115.2)	1.0(1.0)	0.3(0.4)

TABLE I. Data sets included in the CT18(Z) global analysis. The numbers in round brackets are for the CT18Z fit.

ID#	Experimental data set		$N_{pt,n}$	χ^2_n	$\chi^2_n/N_{pt,n}$	S_n
245	LHCb 7 TeV 1.0 fb^{-1} W/Z forward rapidity cross sec.	[46]	33	53.8 (39.9)	1.6 (1.2)	2.2 (0.9)
246	LHCb 8 TeV 2.0 fb^{-1} $Z \to e^-e^+$ forward rapidity cross. sec.	[47]	17	25.8 (23.0)	1.5 (1.4)	1.4 (1.0)
248	ATLAS‡ 7 TeV 4.6 fb^{-1}, W/Z combined cross sec.	[19]	34	287.3 (88.7)	8.4 (2.6)	13.7 (4.8)
249	CMS 8 TeV 18.8 fb^{-1} W cross sec. and A_{ch}	[48]	11	11.4 (12.1)	1.0 (1.1)	0.2 (0.4)
250	LHCb 8 TeV 2.0fb^{-1} W/Z cross sec.	[49]	34	73.7 (59.4)	2.1 (1.7)	3.7 (2.6)
251	ATLAS 8 TeV 20.3 fb^{-1} single diff. high-mass cross sec.	[50]	12	20.4 (25.6)	1.7 (2.1)	1.6 (2.3)
253	ATLAS 8 TeV 20.3 fb^{-1}, Z p_T cross sec.	[51]	27	30.2 (28.3)	1.1 (1.0)	0.5 (0.3)
542	CMS 7 TeV 5 fb^{-1}, single incl. jet cross sec., $R = 0.7$ (extended in y)	[52]	158	194.7 (188.6)	1.2 (1.2)	2.0 (1.7)
544	ATLAS 7 TeV 4.5 fb^{-1}, single incl. jet cross sec., $R = 0.6$	[53]	140	202.7 (203.0)	1.4 (1.5)	3.3 (3.4)
545	CMS 8 TeV 19.7 fb^{-1}, single incl. jet cross sec., $R = 0.7$, (extended in y)	[54]	185	210.3 (207.6)	1.1 (1.1)	1.3 (1.2)
573	CMS 8 TeV 19.7 fb^{-1}, $t\bar{t}$ norm. double-diff. top p_T & y cross sec.	[55]	16	18.9 (19.1)	1.2 (1.2)	0.6 (0.6)
580	ATLAS 8 TeV 20.3^{-1}, $t\bar{t}$ p_T^t and $m_{t\bar{t}}$ abs. spectrum	[56]	15	9.4 (10.7)	0.6 (0.7)	-1.1 (-0.8)

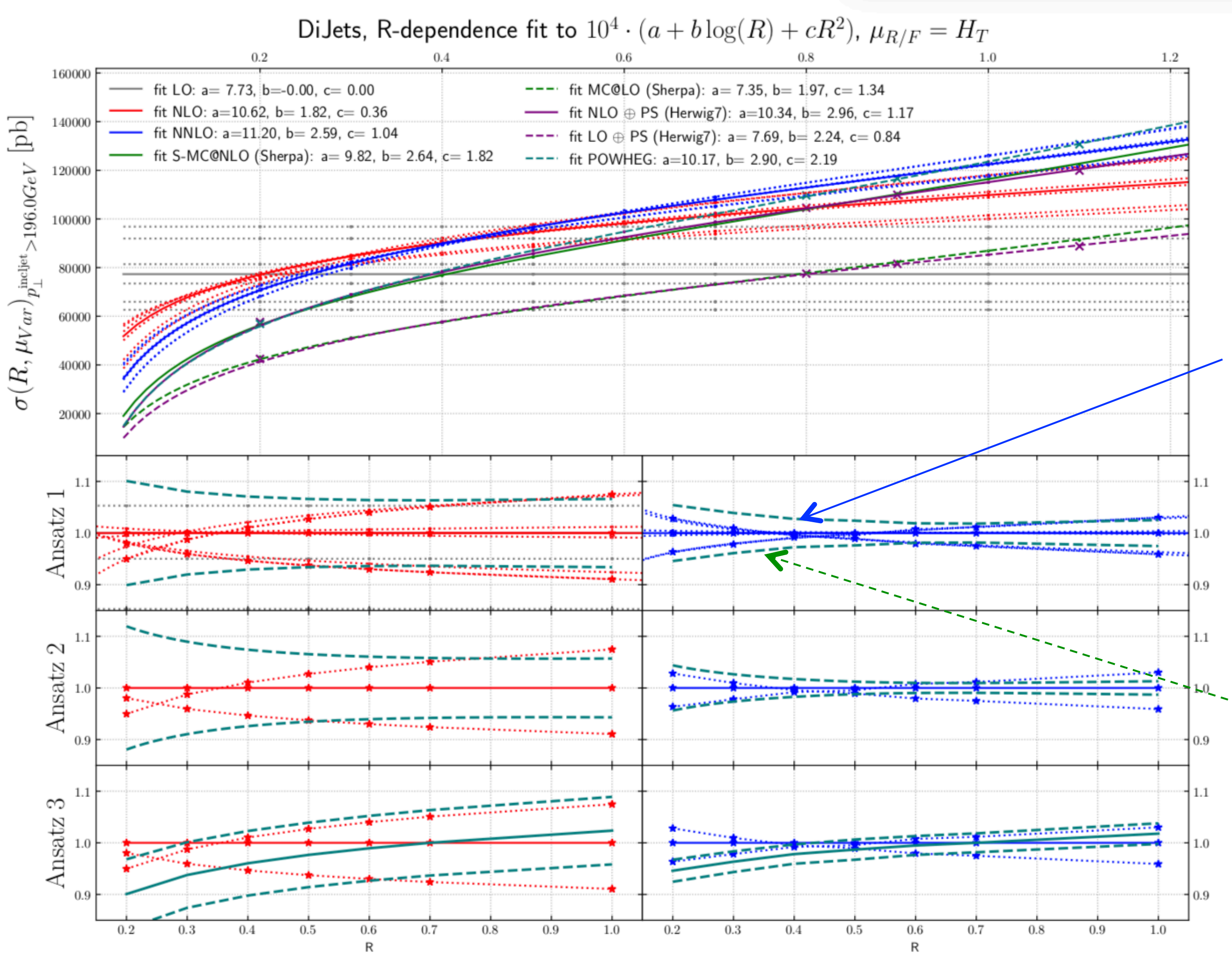
TABLE II. High precision LHC measurements employed in the CT18(Z) analysis. The numbers in round brackets are for the CT18Z fit. $N_{pt,n}$, χ^2_n are the number of points and value of χ^2 for the n-th experiment at the global minimum. S_n is the effective Gaussian parameter [18, 57, 58] quantifying agreement with each experiment. The ATLAS data, labelled by ‡, are included in the CT18Z global fit, but not in CT18.

Treatment of new LHC data

- Include processes that have a sensitivity for the PDFs of interest, and for which NNLO predictions are available.
- Include as large a rapidity interval for the jet data as possible
  - for ATLAS this involves using the ATLAS de-correlation model, rather than using a single rapidity interval. Using a single rapidity interval may result in selection bias. The result is a worse χ^2 due to the remaining tensions in the ATLAS jet data, and a reduced sensitivity compared to the CMS jet data.
  - the use of only a single jet rapidity interval provides incomplete information
- Use multiple t-tbar observables, possible using experimentally provided statistical correlations.
  - and for CMS, using the double differential calculation from Mitov et al
  - again, some of the observables are in tension with each other.
- NB: previous data (including CMS 7 TeV W,Z data) continue having an impact on global fits and tend to dilute the impact of new data

Treatment of new LHC data

- For jet data, always use the largest jet size available; use of smaller R sizes can result in inaccurate scale uncertainties



arXiv:1903.12563

- At NNLO, there are accidental cancellations, that lead to an artificially low scale uncertainty for processes with small R (0.4) jets
- Prescription for restoring reasonable uncertainty estimate
- Original idea from Salam et al ->Les Houches Accord

Theory input

Obs.	Expt.	fast table	NLO code	K-factors	R,F scales
Inclusive jet	ATL 7 CMS 7/8	APPLgrid fastNLO	NLOJet++	NNLOJet	p_T, p_T^1
p_T^Z	ATL 8	APPLgrid	MCFM	NNLOJet	$\sqrt{Q^2 + p_{T,Z}^2}$
W/Z rapidity W asymmetry	LHCb 7/8 ATL 7 CMS 8	APPLgrid	MCFM/aMCfast	FEWZ/MCFM	$M_{W,Z}$
DY (low,high mass)	ATL 7/8 CMS 8	APPLgrid	MCFM/aMCfast	FEWZ/MCFM	Q_{ll}
$t\bar{t}$	ATL 8 CMS 8	fastNNLO			$\frac{H_T}{4}, \frac{m_T}{2}$

when justified, a small Monte-Carlo error (typically 0.5%) added for NNLO/NLO K-factors

Theoretical calculations for vector boson production

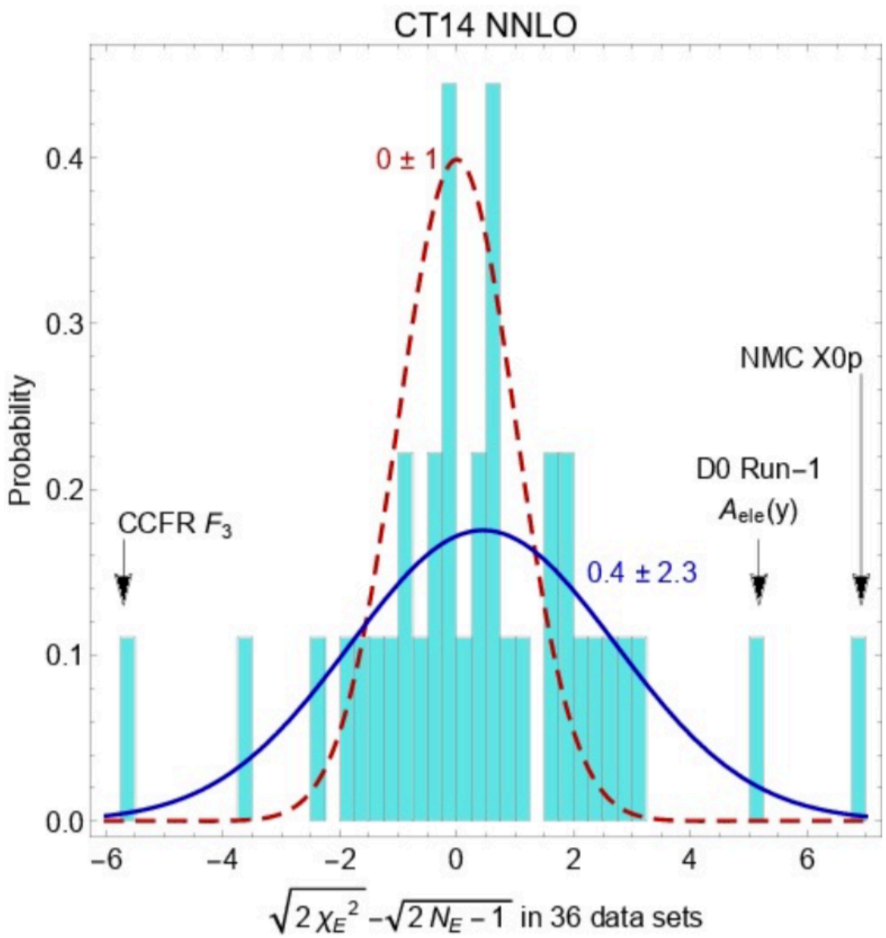
ID	Obs.	Expt.	fast table	NLO code	K-factors	$\mu_{R,F}$
245 246 250 249	$y_{\mu\mu}, \eta_\mu$ y_{ee} $y_{\mu\mu}, \eta_\mu$ $A(\mu)$	LHCb7ZW LHCb8Z LHC8ZW CMS8W	APPLgrid	MCFM/aMCfast	MCFM/FEWZ	$M_{Z,W}$
253	p_T^{ll}	ATL8Z	APPLgrid	MCFM	NNLOJet	M_T^{ll}
201 203 204	$\sqrt{\tau}, y$ $\sigma_{pd}/\sigma_{pp}, x_F$ 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	$M_{Z,W}$

What experiments provide the greatest constraints?

- ...and how well do the LHC experiments agree among themselves?

- We use
  - Effective Gaussian variables
  - PDFSense and L2 sensitivity
  - Lagrange Multiplier scans
  - ePump

Effective Gaussian variables



Define $S_n(\chi^2, N_{pt})$ for experiment n so that, in a perfect fit, it would approximately obey the standard normal distribution $\mathcal{N}(0,1)$ (mean=0, half-width=1) independently of $N_{pt,n}$

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

[H.-L. Lai et al., arXiv:1007.2241; S.Dulat et al., arXiv:1309.0025; K. Kovarik, P.N., D. Soper, arXiv:1905.06957]

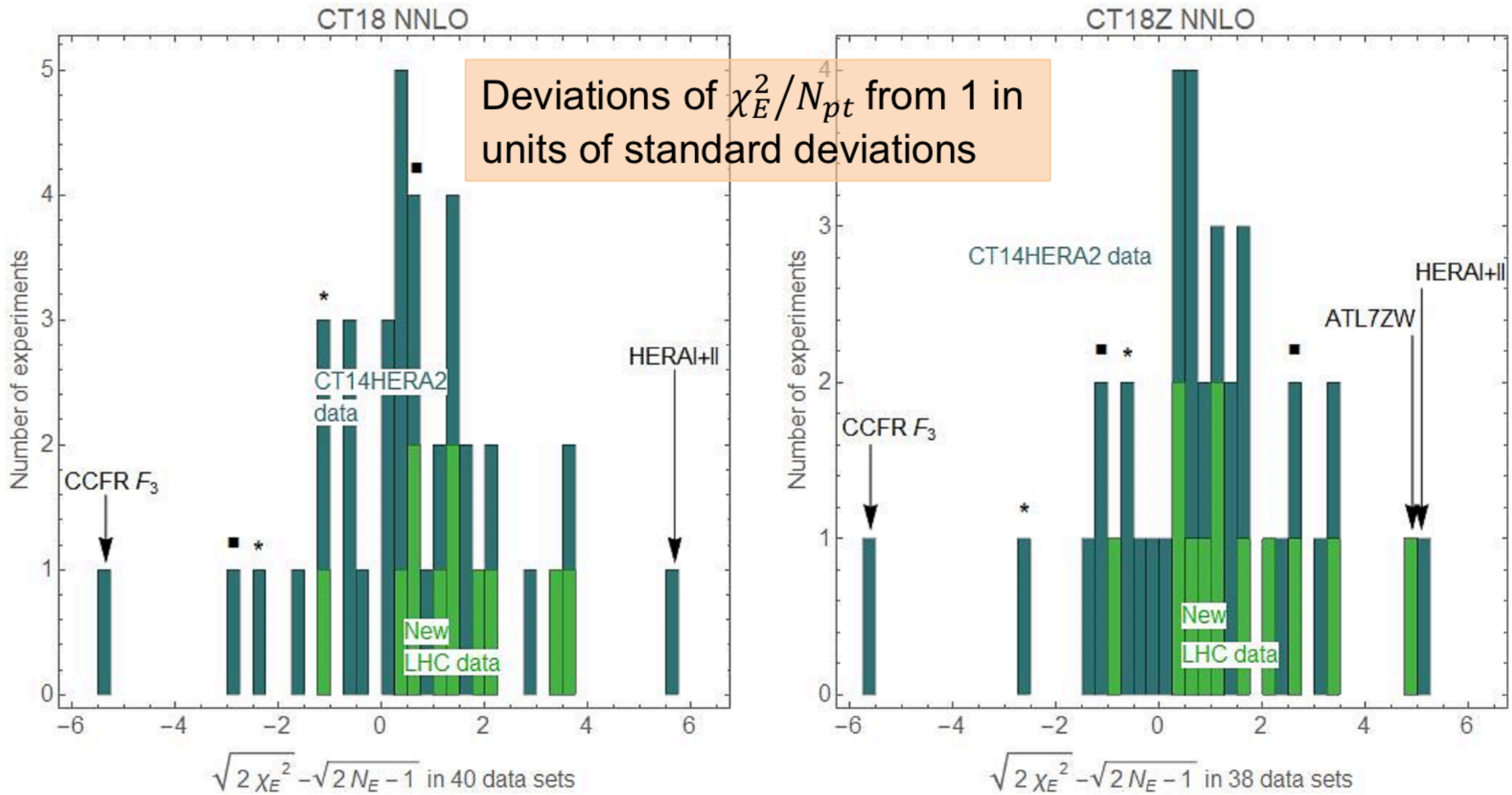
An empirical S_n distribution can be compared to $\mathcal{N}(0,1)$ visually or using a statistical (Anderson-Darling, Kolmogorov-Smirnov, …) test

CT18 (CT18Z) NNLO

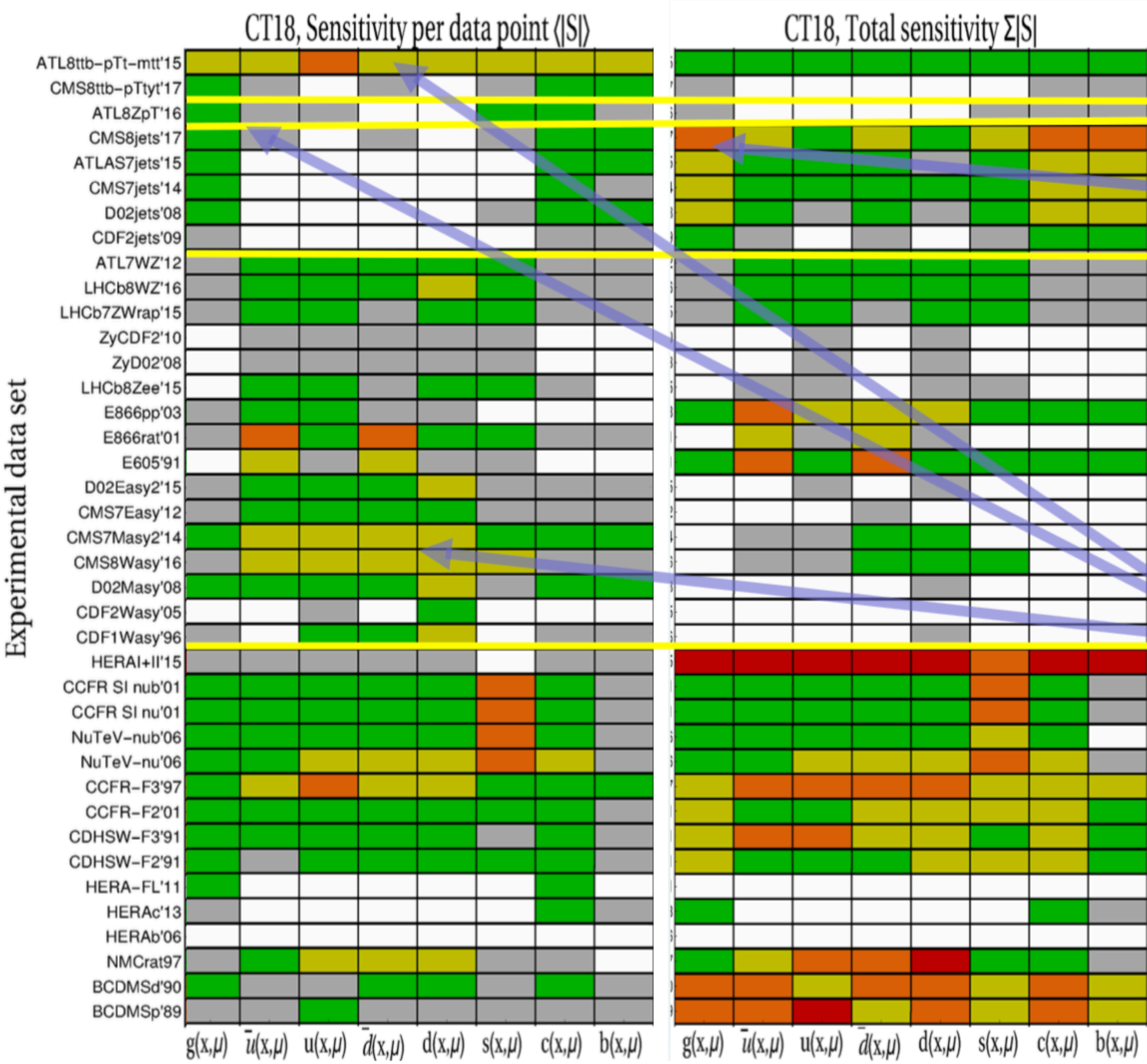
13 (14) new LHC experiments with 665 (711) data points

LHC experiments, especially ATLAS 7 TeV Z, W production (only in the CT18A and Z fits) tend to have elevated χ_n^2/N_{pt} in global fits

Deviations of χ_E^2/N_{pt} from 1 in units of standard deviations



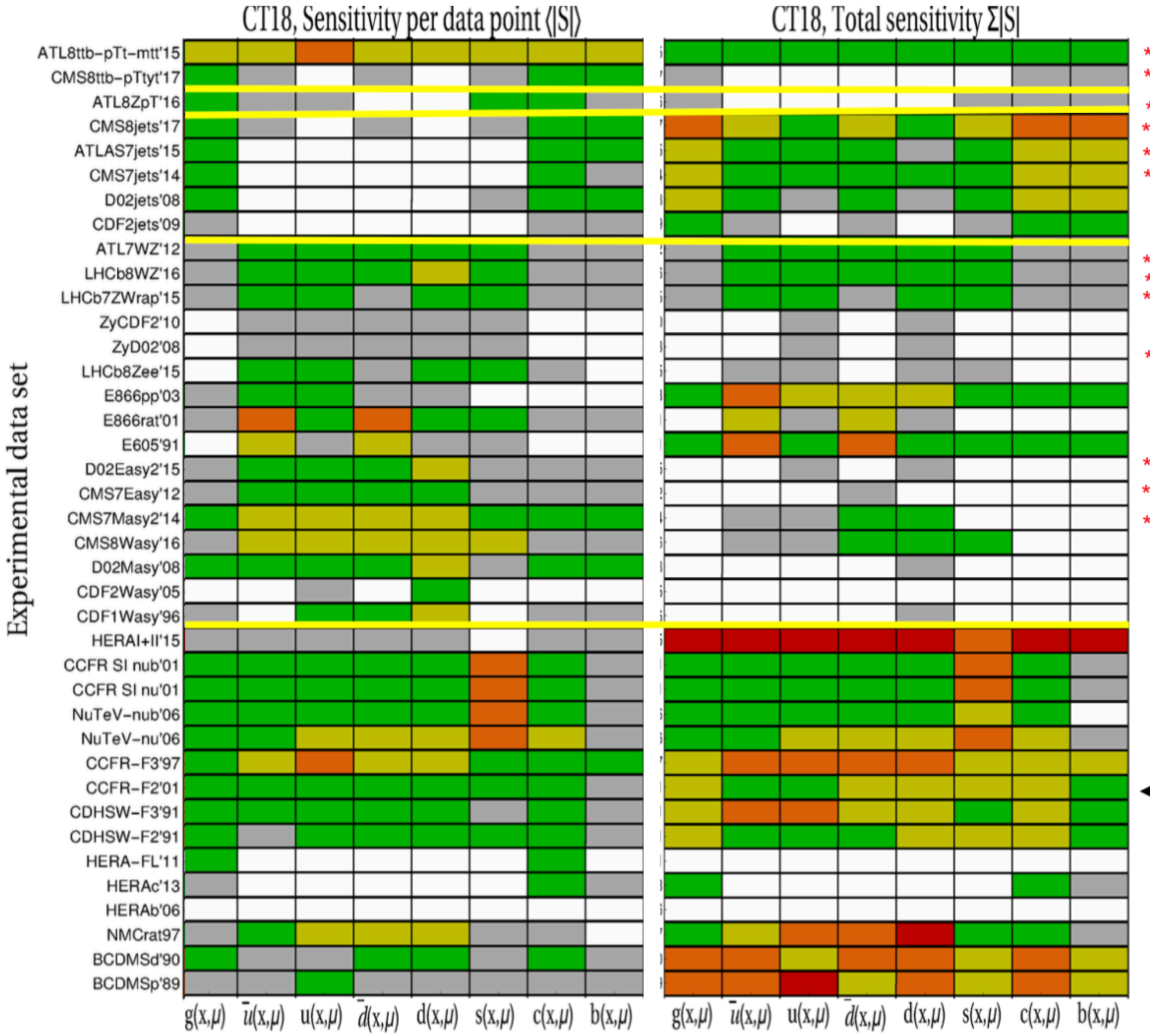
Sensitivity per point and total sensitivity



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 ($N_{pt} \sim 10 - 20$)

Sensitivity per point and total sensitivity

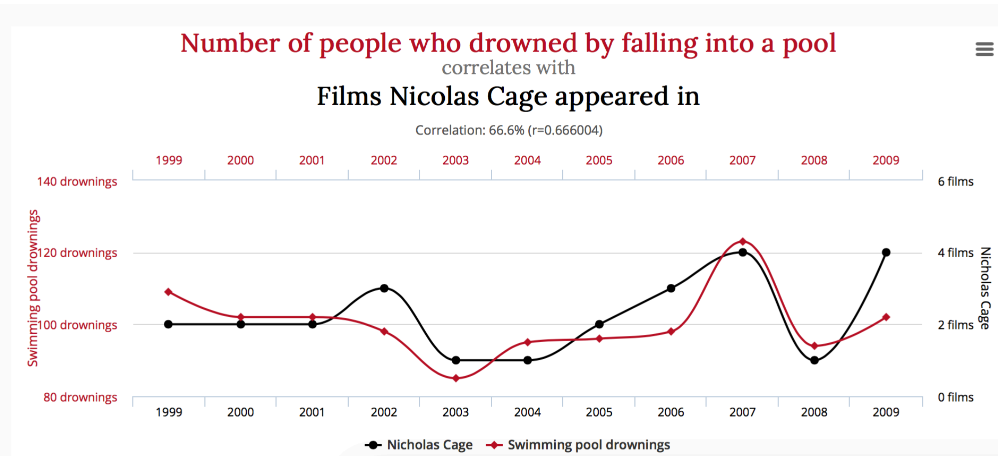


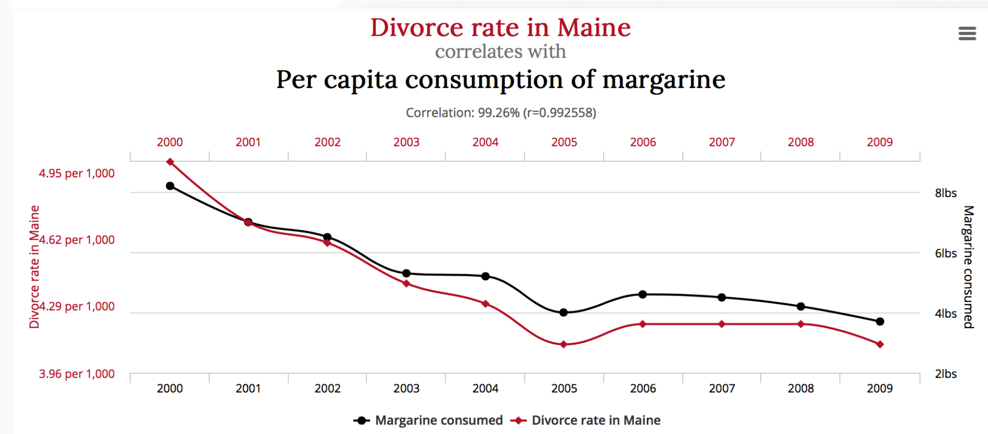
The LHC data sets (*) hold a great promise – if they agree

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

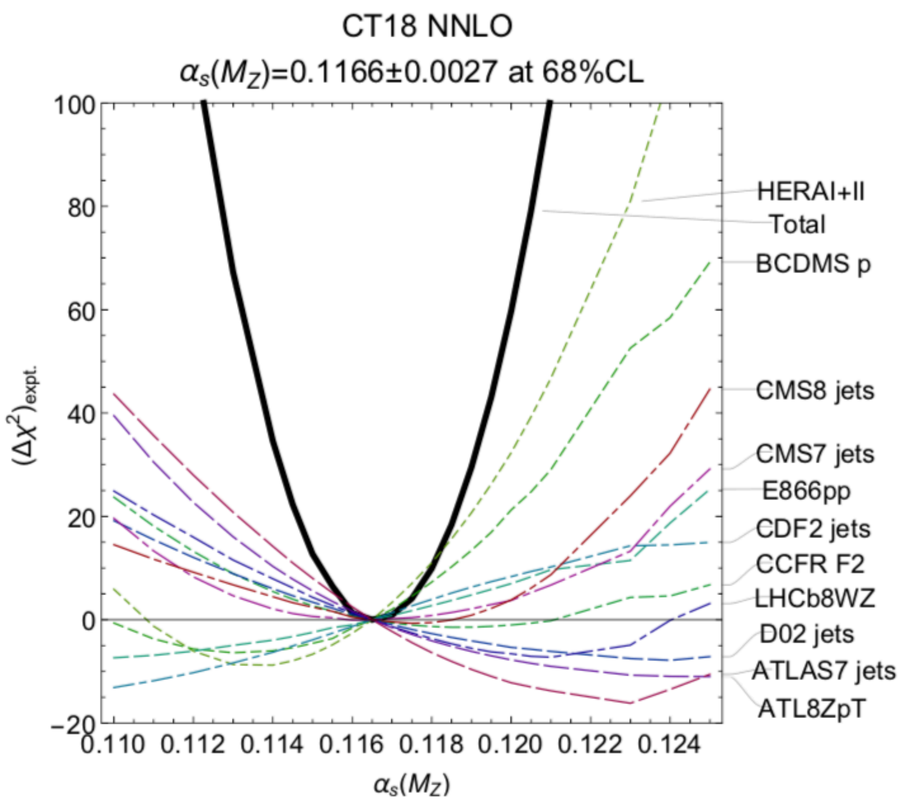
Correlations

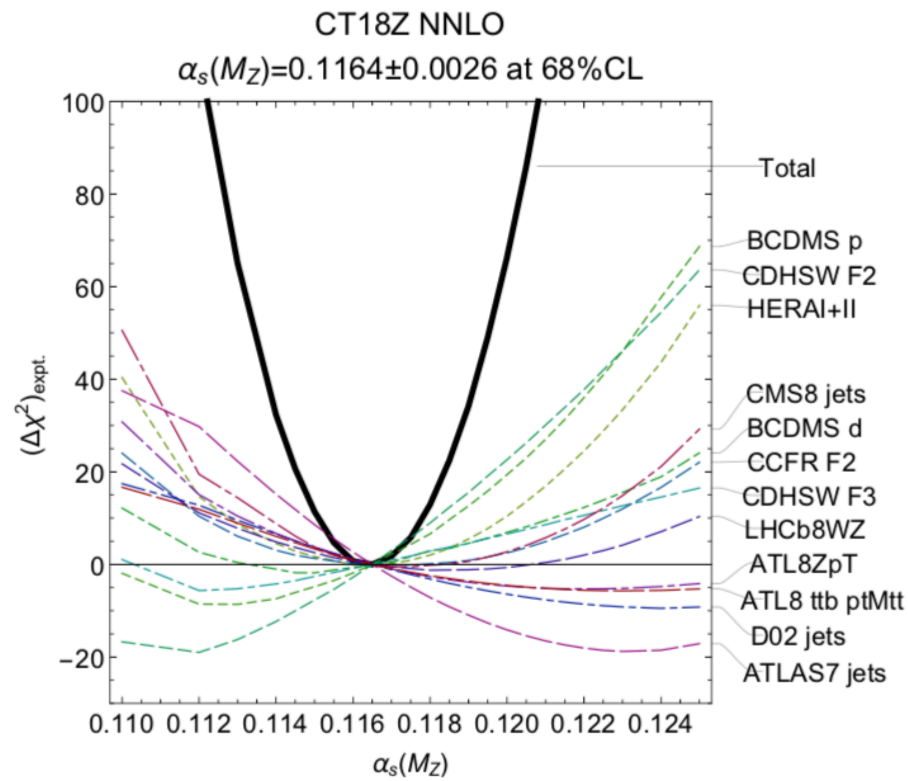
Correlations are important, but not sufficient. The statistical power of the data set also has to be there. The most effective data sets may have low correlation, but high sensitivity. Tensions within the data set also may reduce the ultimate sensitivity.





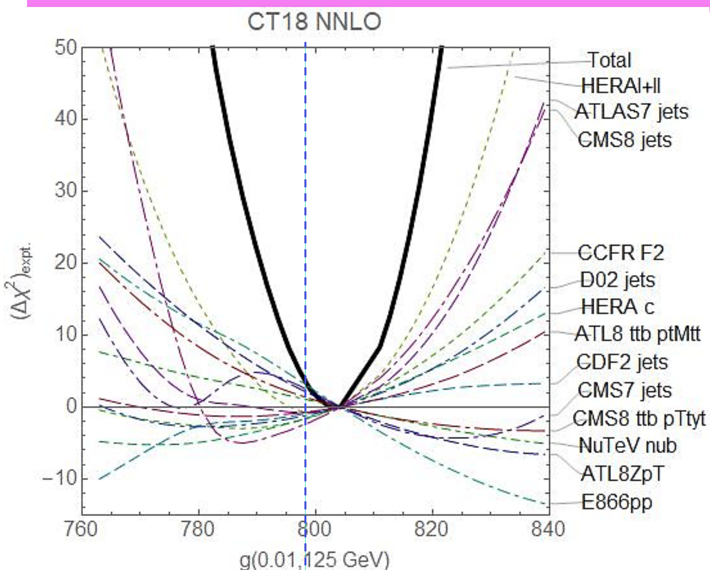
Lagrange Multiplier Scans

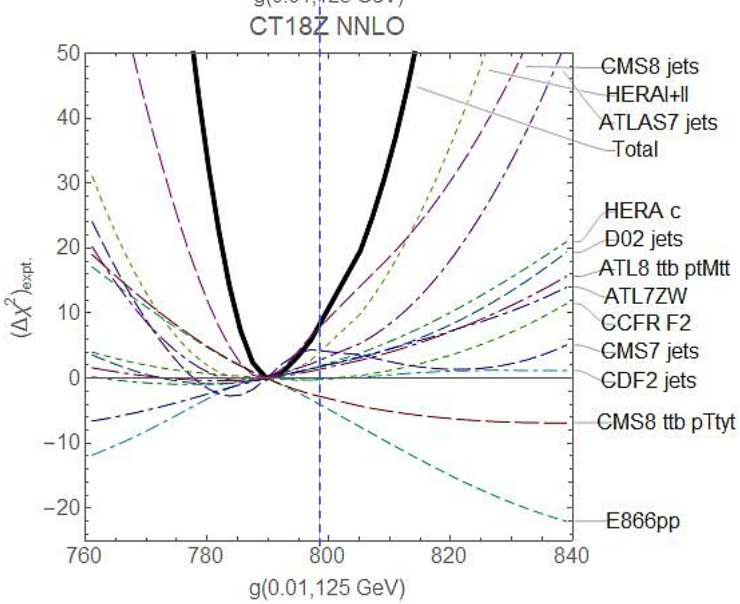




New PDG $\alpha_s(m_Z)$ and uncertainty coming out soon, including input from PDF fits; divide inputs into lattice/not-lattice

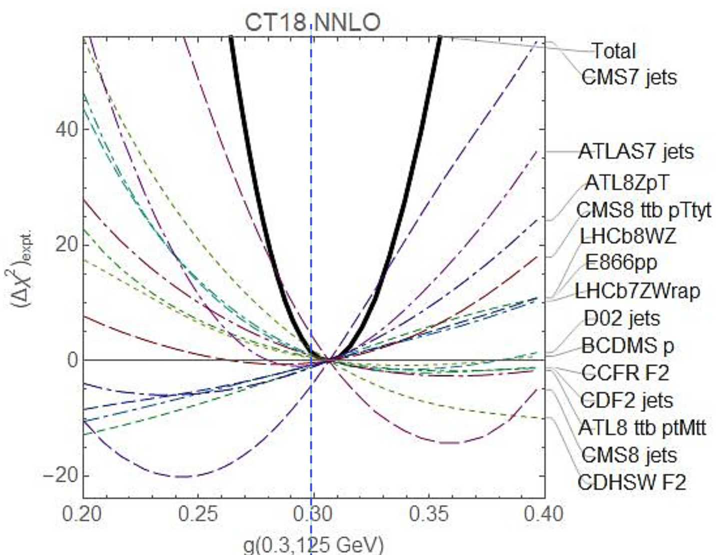
Lagrange Multiplier scan (gluon: x=0.01, Q=125 GeV)

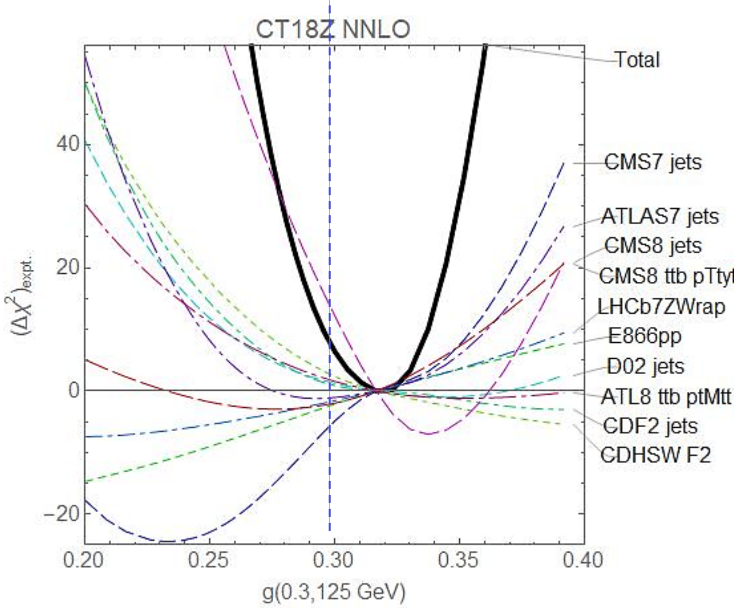




- Top: CT18
  - HERA1+II data set provides the dominant constraint, followed by ATLAS, CDF2,D02 jet production, HERA charm...
  - tt double differential cross sections provide weaker constraints
- Lower: CT18Z
  - a 1% lower NNLO gluon in the Higgs production region than for CT14/ CT18 as a result of
    - higher charm mass, m_c^{pole}=1.4 GeV
    - including ATLAS7 W/Z production
    - a special factorization scale in DIS that mildly improves χ^2 and approximates the effect of small-x resummation

Lagrange Multiplier scan (gluon: x=0.3, Q=125 GeV)





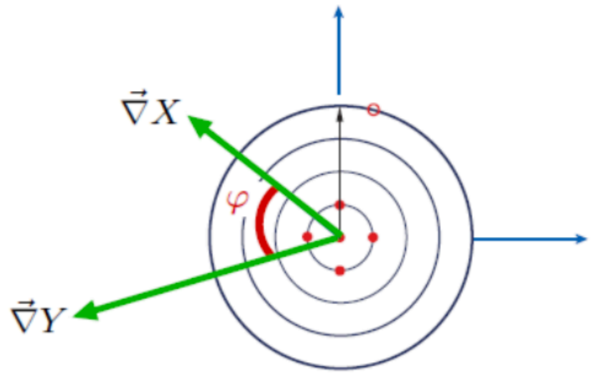
- Upper: CT18
- Lower: CT18Z
- Opposite pulls from ATLAS7/ CMS7 jet production on one hand, and CMS8 jet production on the other hand
- Similarly, ATLAS tt distributions (dm_{tt}, dp_T^t) and CMS double tt distributions ($dp_T^t dy_t$) at 8 TeV impose weak opposite pulls
  - NB: tt data has relatively large impact if jet data removed, or if statistical precision increased to match that of the jet data
- Constraints from ATLAS8 Z p_T production are moderate

L_2 sensitivity, definition

Tolerance hypersphere in the PDF space

2-dim (i,j) rendition of N-dim (22) PDF parameter space

Faster than Lagrange Multiplier; works in the Hessian approximation; See for example slide 27.



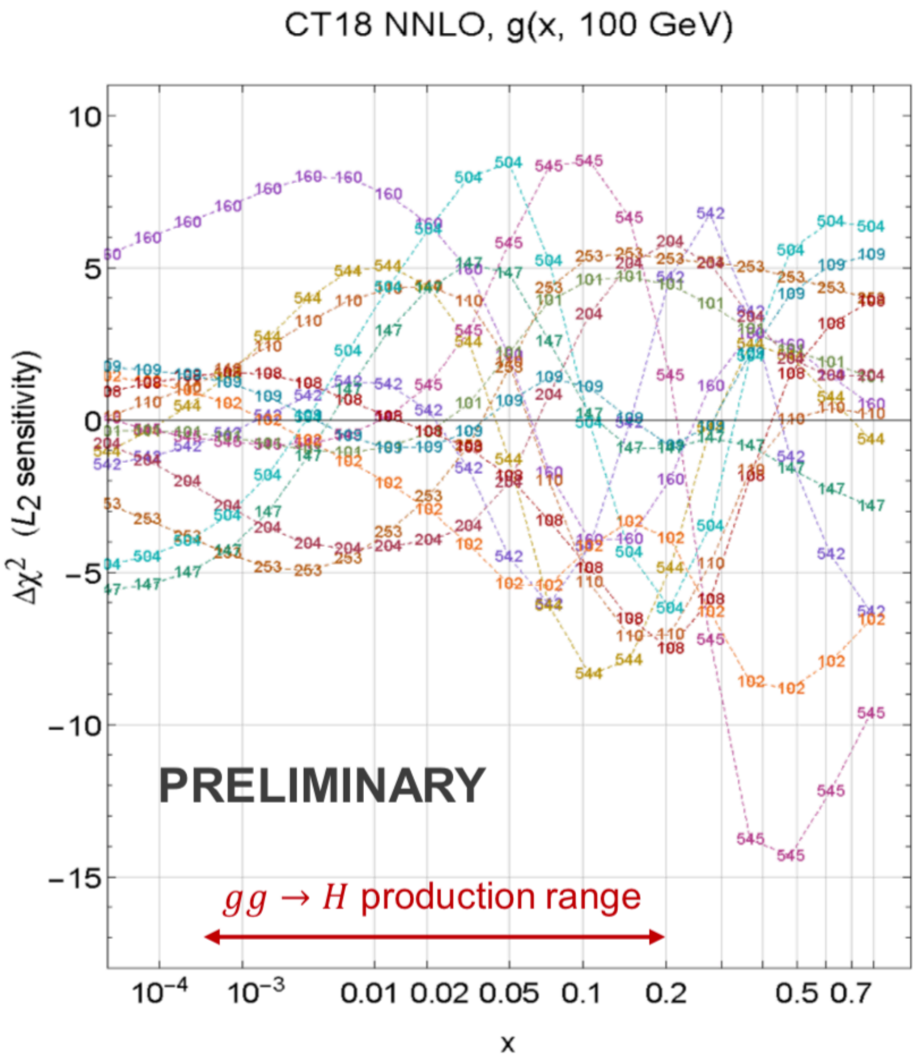
(b)
Orthonormal eigenvector basis

L_2 sensitivity. Take $X = f_a(x_i, Q_i)$ or $\sigma(f)$; $Y = \chi^2_E$ for experiment E. Find $\Delta Y(\vec{z}_{m,X})$ for the displacement $|\vec{z}_{m,X}| = 1$ along the direction $\vec{\nabla}X / \left|\vec{\nabla}X\right|$ (corresponding to $\Delta\chi^2_{tot} = T^2$ and $X(\vec{z}) = X(0) + \Delta X$):

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

Estimated χ^2 pulls from experiments

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



CT18 NNLO, gluon at Q=100 GeV

15 core-minutes

Most sensitive experiments

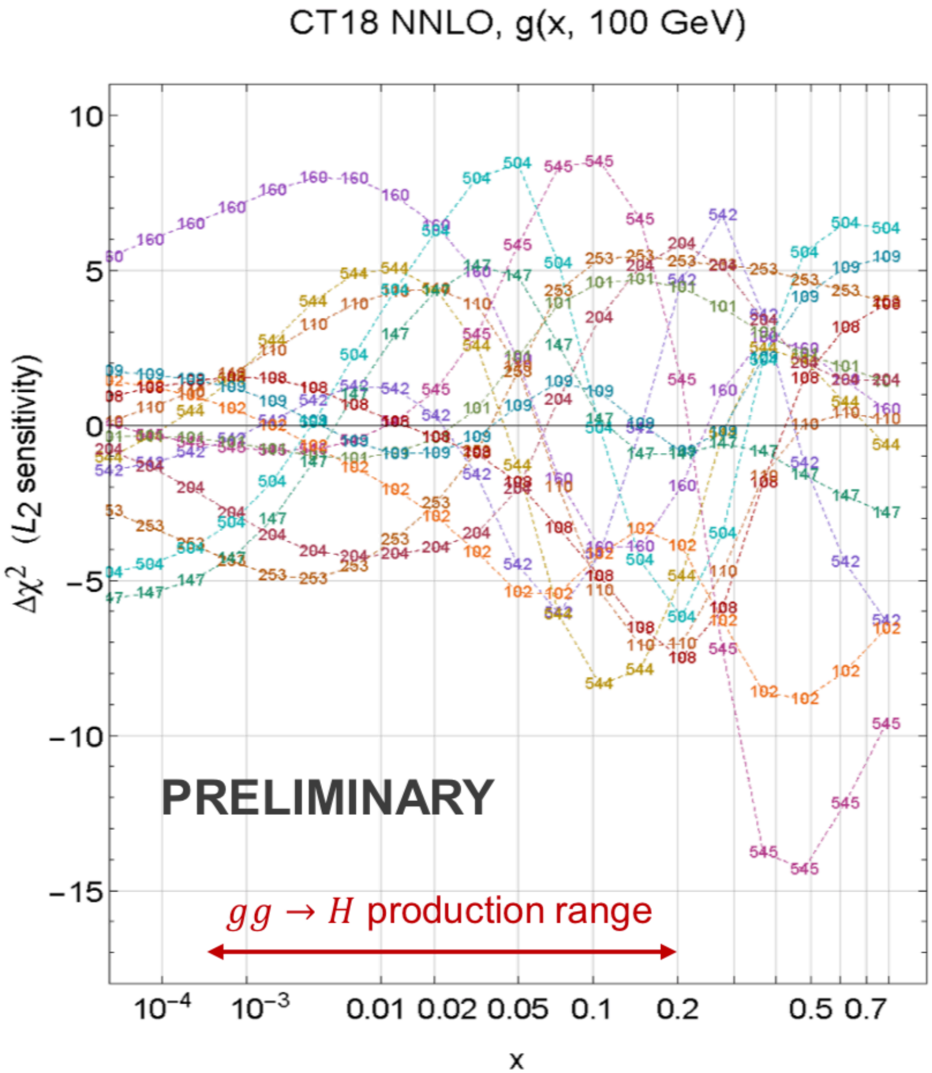
- 253 ATL8ZpTbT
- 542 CMS7jtR7y6T
- 544 ATL7jtR6uT
- 545 CMS8jtR7T
- 160 HERAIpII
- 101 BcdF2pCor
- 102 BcdF2dCor
- 108 cdhswf2
- 109 cdhswf3
- 110 ccfrf2.mi
- 147 Hn1X0c
- 204 e866ppxf
- 504 cdf2jtCor2

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 using CT18 Hessian PDFs

Estimated χ^2 pulls from experiments

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



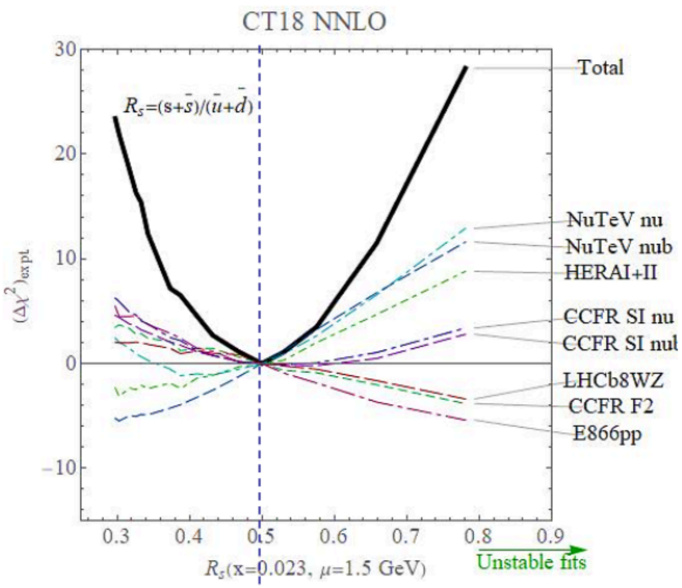
CT18 NNLO, gluon at Q=100 GeV

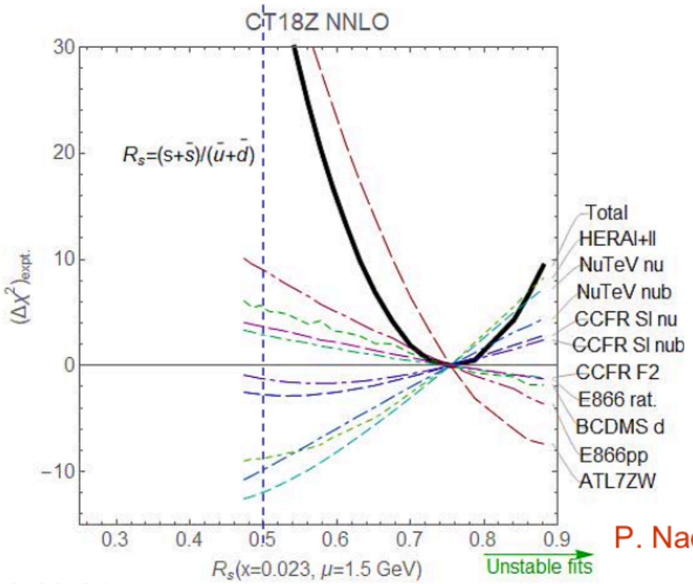
Most sensitive experiments

- 253 ATL8ZpTbT
- 542 CMS7jtR7y6T
- 544 ATL7jtR6uT
- 545 CMS8jtR7T
- 160 HERAIpII
- 101 BcdF2pCor
- 102 BcdF2dCor
- 108 cdhswf2
- 109 cdhswf3
- 110 ccfrf2.mi
- 147 Hn1X0c
- 204 e866ppxf
- 504 cdf2jtCor2

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

Lagrange Multiplier scan: $\mathrm{R}_s(x = 0.023, \mu = 1.5\ GeV)$

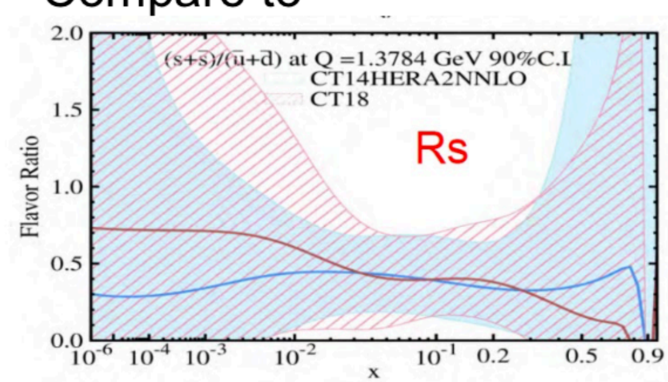




The CT18Z strangeness is increased primarily as a result of including the ATLAS 7 TeV W/Z production data (not in CT18), as well as because of using the DIS saturation scale and $m_c^{pole} = 1.4$ GeV

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

Compare to



Look at residuals and nuisance parameters

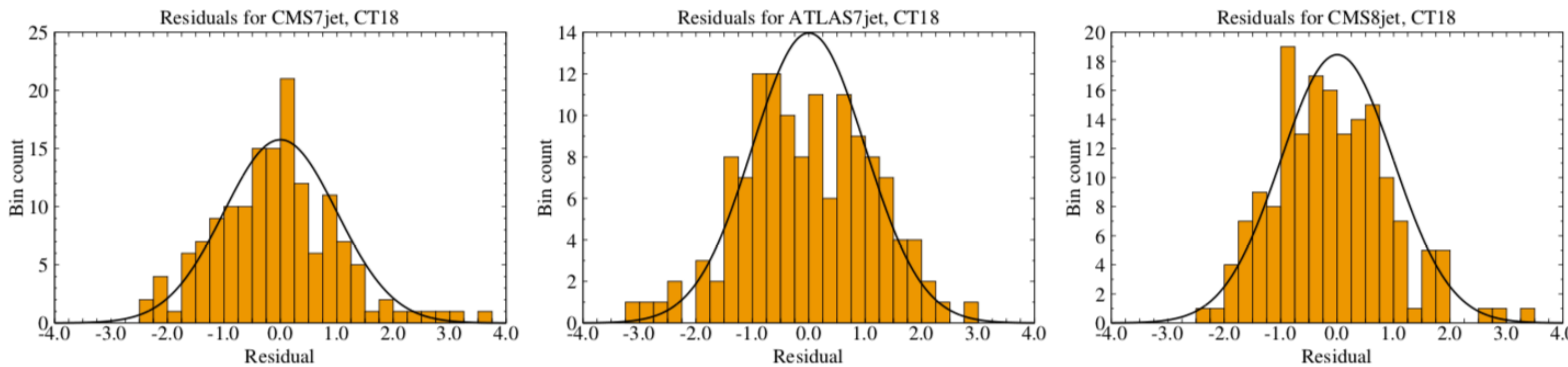


FIG. 41. Distribution of residuals for the CMS (ID=542 and 545) and ATLAS (ID=544) jet data.

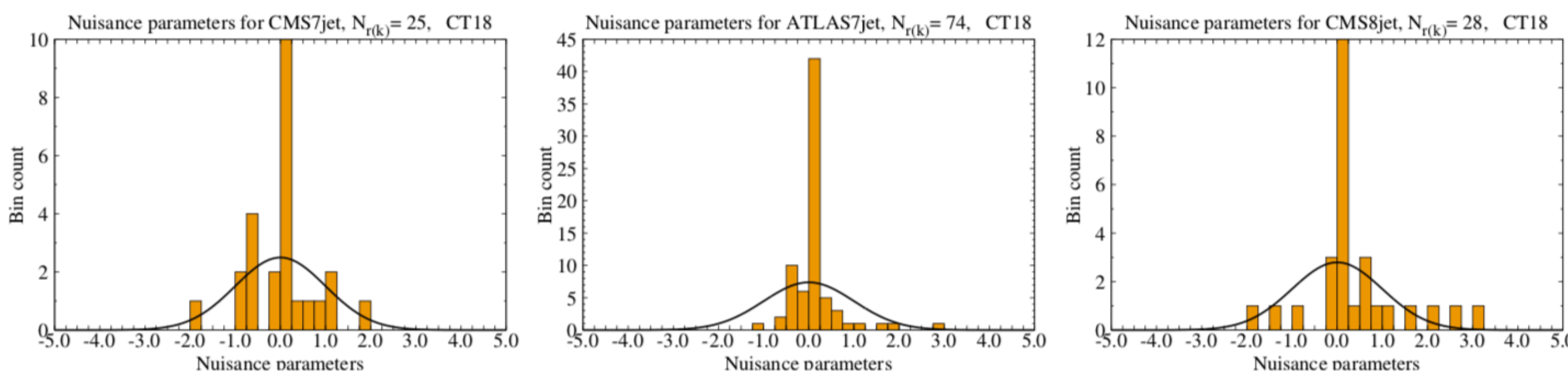
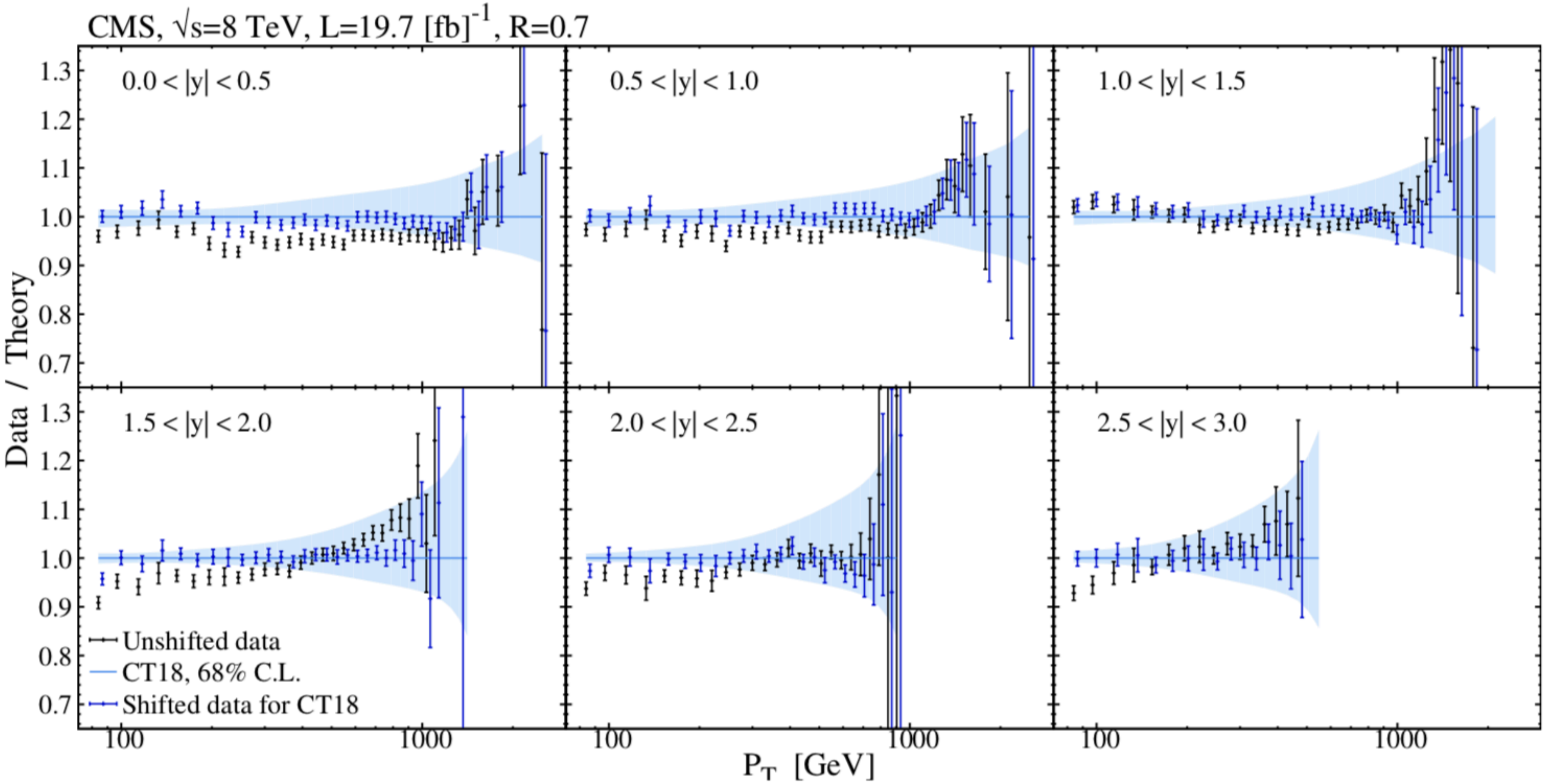
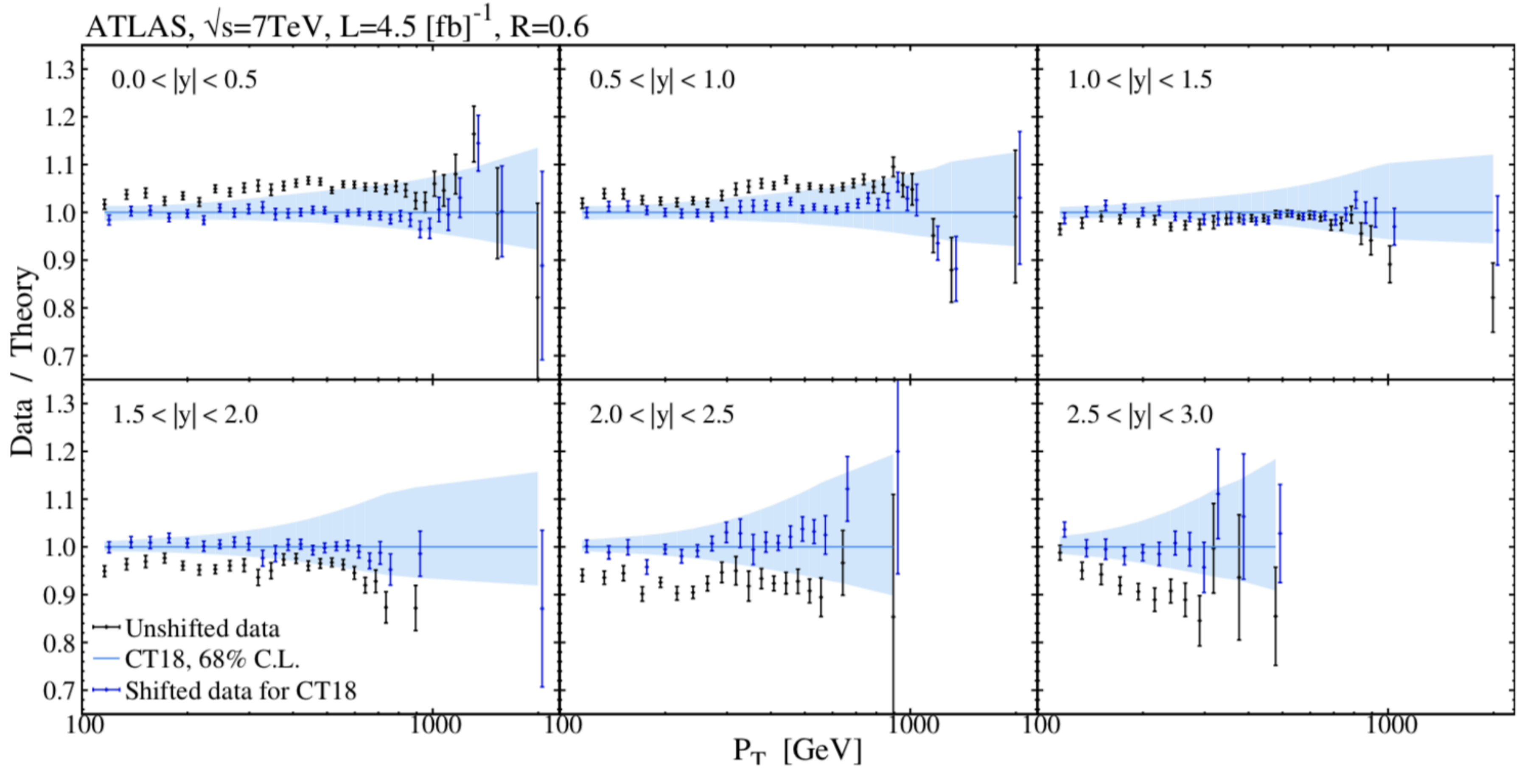


FIG. 42. Distribution of nuisance parameters for the CMS (ID=545) and ATLAS (ID=544) jet data.

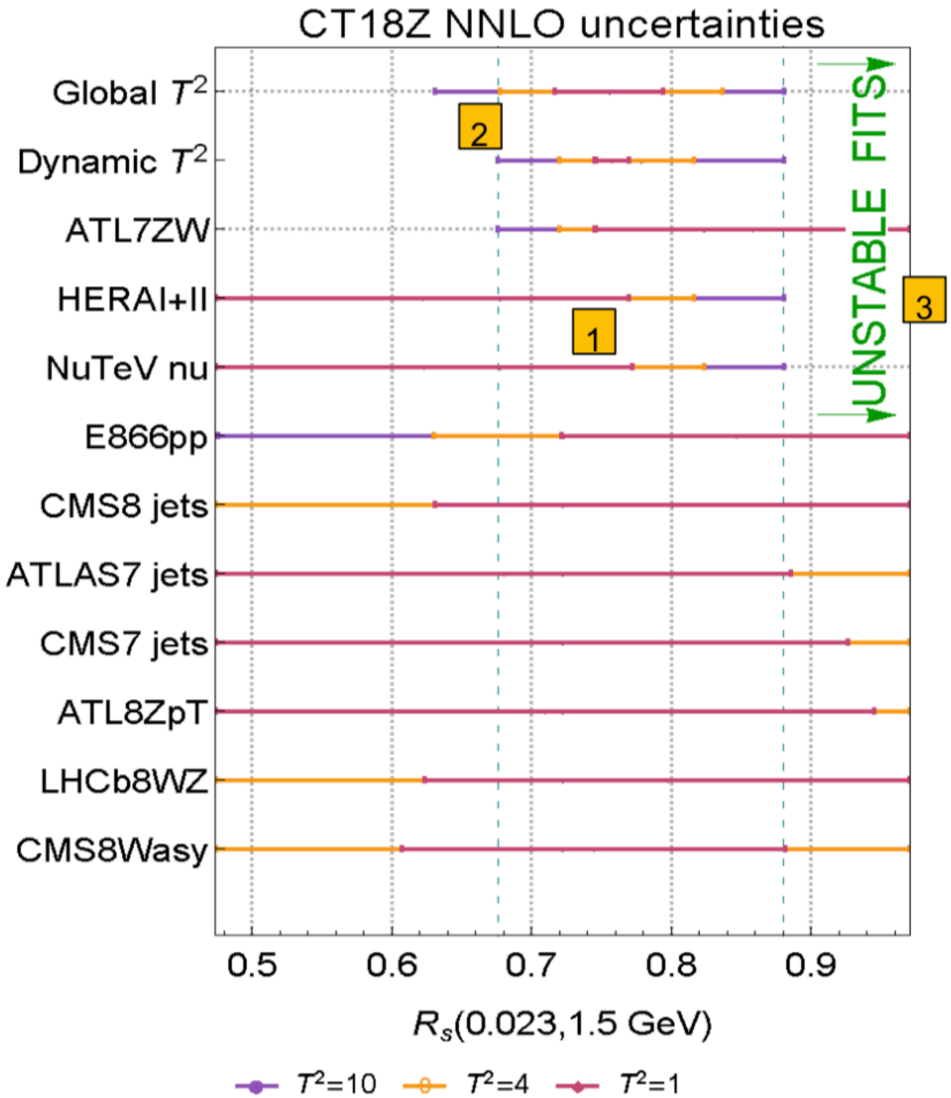
Shifted vs unshifted: CMS 8 TeV jets



Shifted vs un-shifted: ATLAS 7 TeV jets



Effect on PDF uncertainties



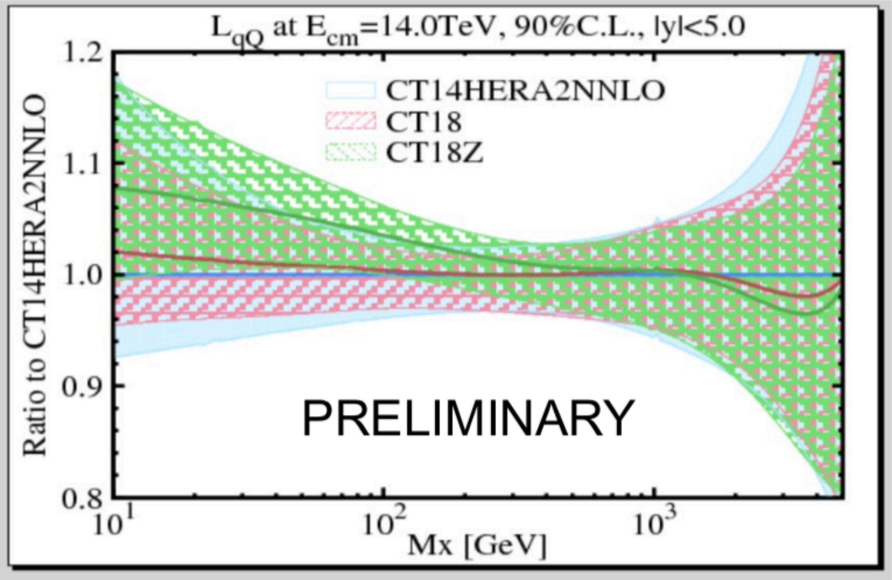
The LM scan reveals details not captured by other methods

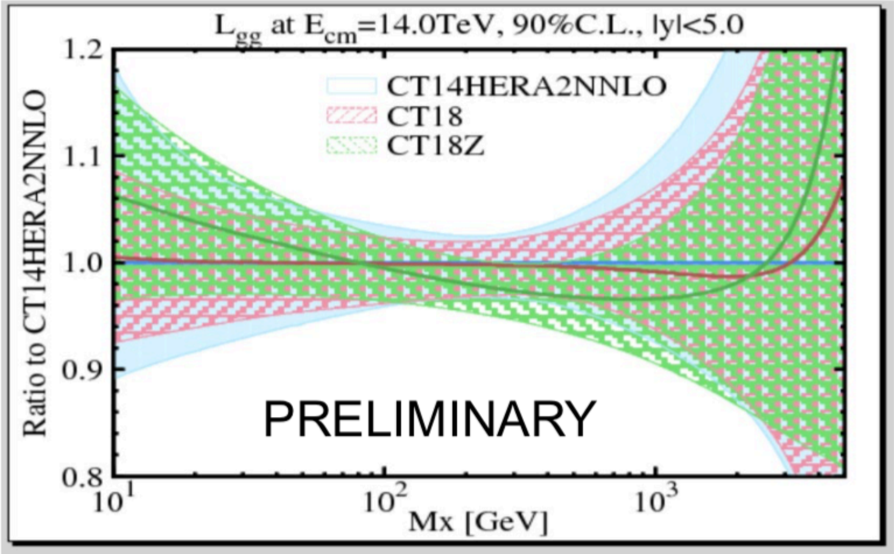
1 **Nonlinearities:** the error bands for tolerance $T^2 = 1, 4, 10$ may not scale according to the Gaussian distribution

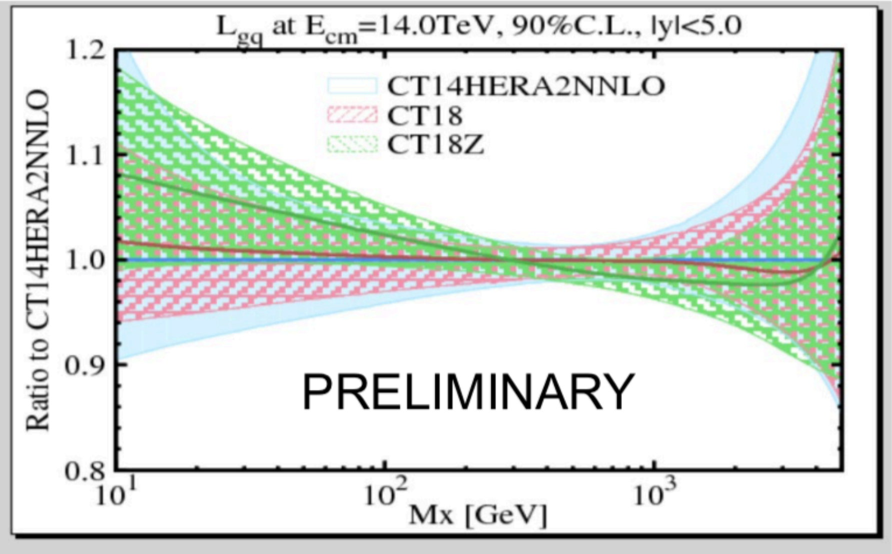
2 **Tensions:** in the affected direction(s), the global tolerance and **especially dynamic tolerance** may underestimate the true PDF error.

3 χ^2 **instability:** Neither the "global T^2" nor "dynamic T^2" reflect instability of fits at $R_s > 0.9$

PDF luminosities



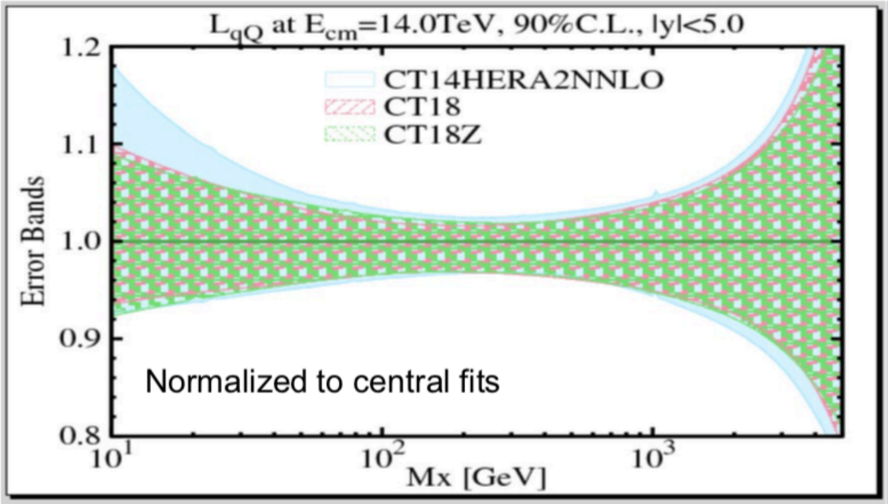


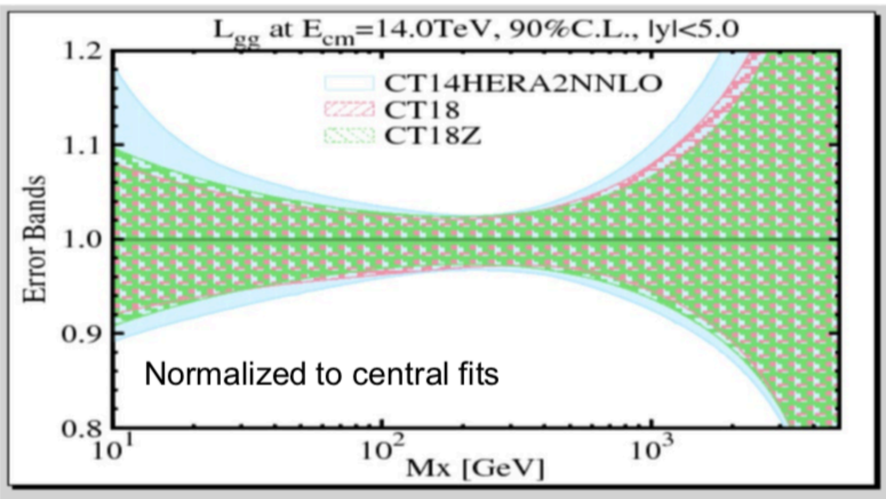


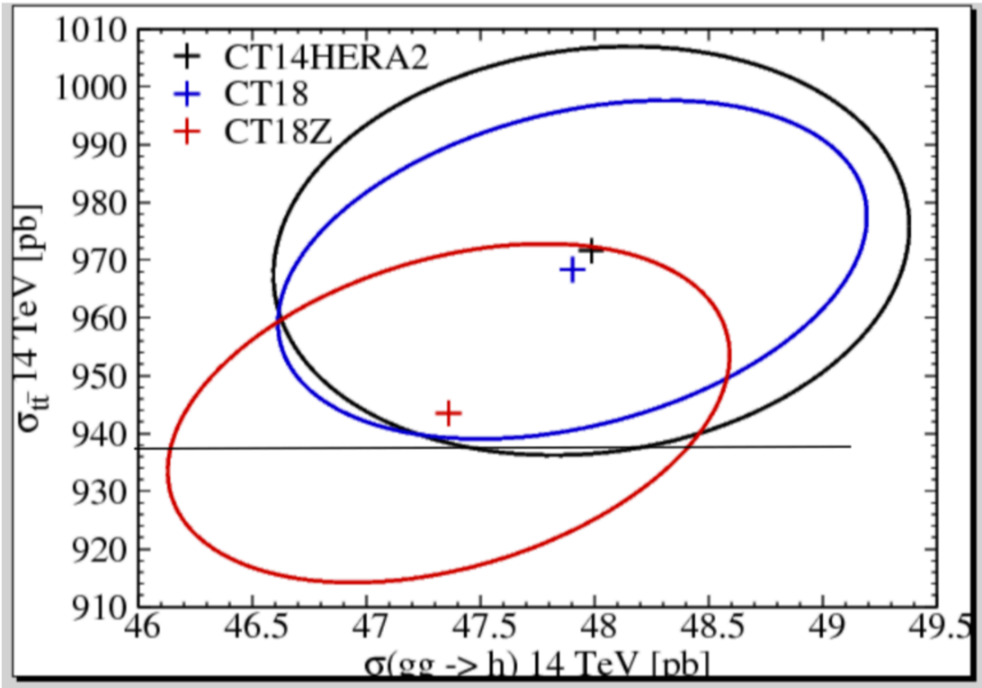
CT18 consistent with CT14

CT18Z has a somewhat different shape, especially at low invariant masses M_X

PDF luminosities



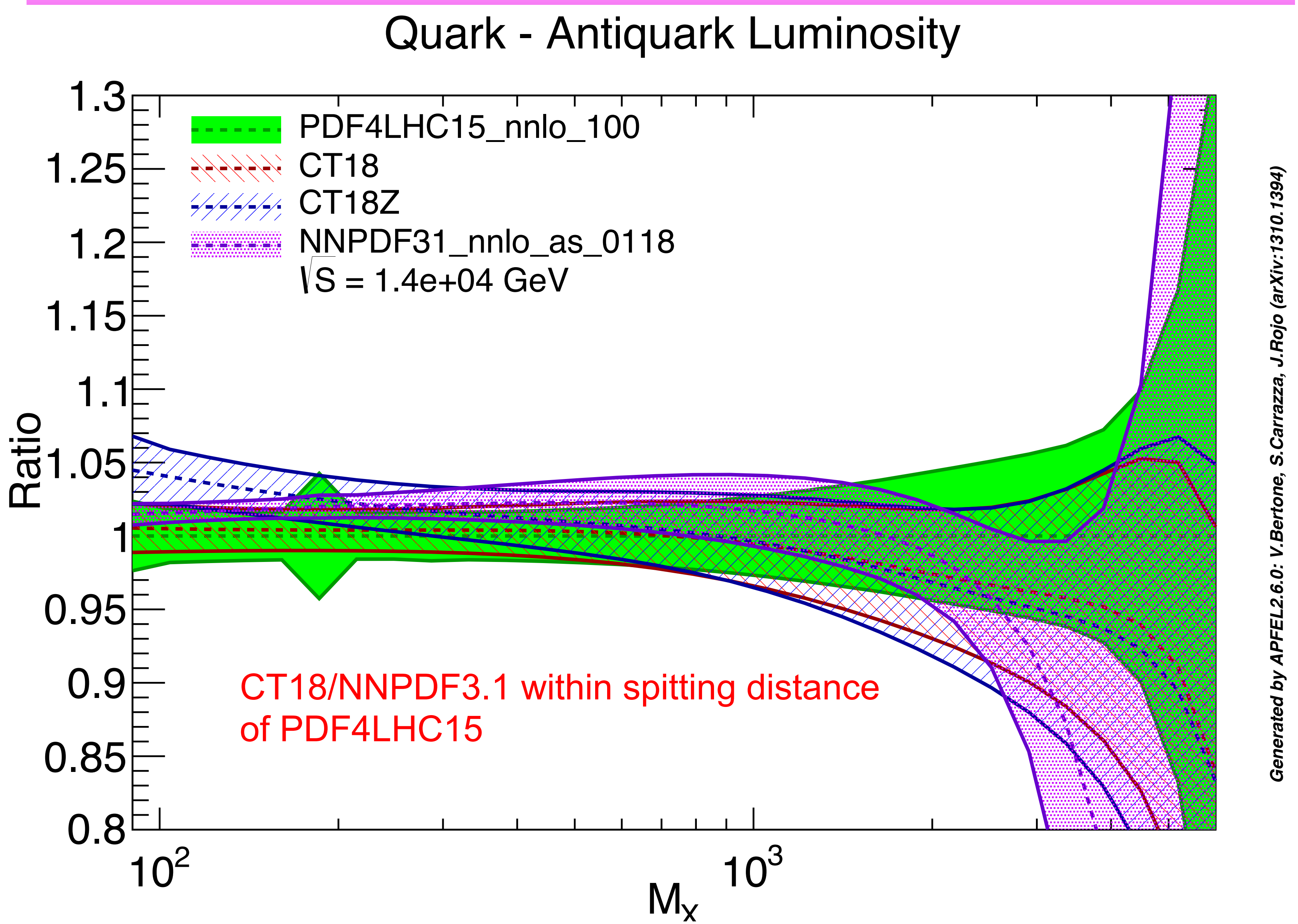




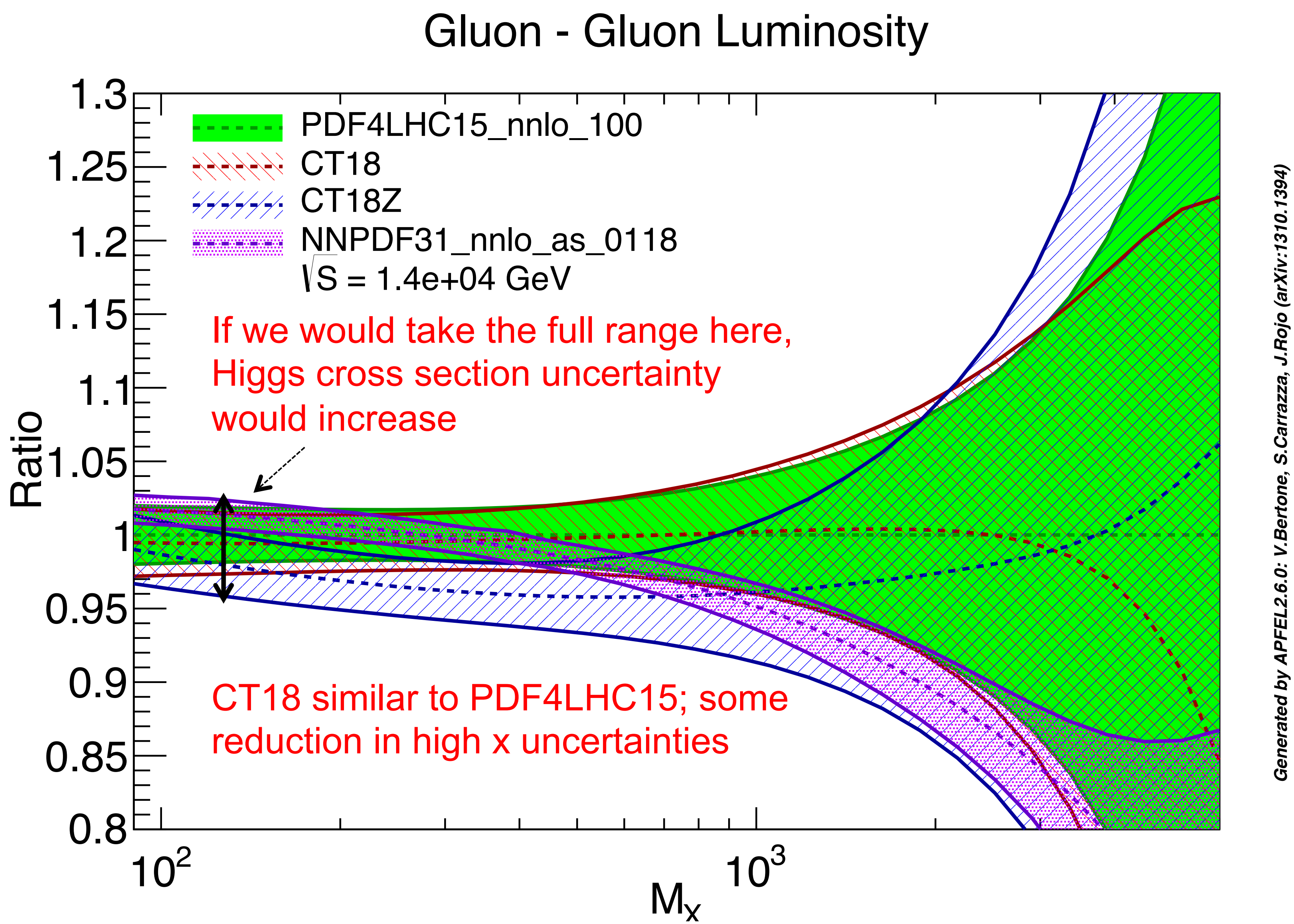
CT18Z $gg \to H$ and $t\bar{t}$ production cross sections lower by about 1 and 2.5% compared to CT14HERA2

PRELIMINARY

Comparisons to PDF4LHC15



Comparison to PDF4LHC15



Towards PDF4LHC20

- We are gearing up for the full 300 fb^{-1} expected from Runs 1-3
- New generations of PDF sets are/will be coming out
- It's been 5 years since PDF4LHC15
- It may be time for a PDF4LHC20 combination
- ...which will of necessity include some benchmarking before-hand
- First benchmark Drell-Yan codes/vector boson production; see Sergey's presentation at the EWWG meeting
- There are a large number of tools that have been developed to examine the impact of data pre- and post-fit; let's make use of them to understand the impact of the LHC datasets on each PDF fit, both in terms of the central fit and the uncertainty
- Go back to HERA1+II data? Confirm what the central values and uncertainties look like with new formalisms.
- Add top data from ATLAS; each single distribution, several distributions simultaneously using the statistical correlations; do we see similar effects/ influences for the different fitting programs?
- Add ATLAS/CMS jet data, each jet y bin individually and then together
- Add jet and top data together

Extras

ePump (error PDF updating package)

- ePump (Error PDF Updating Method Package) is a set of classes, functions, etc. for analyzing the impact of new data on the PDF predictions and uncertainties, in the Hessian method.
- It assumes quadratic dependence of the global χ^2 function on the parameters, and linear dependence of the observables on the parameters.
- It allows for the inclusion of a dynamical tolerance in each of the original Hessian eigenvector directions.
- Extensively cross-checked against actual global fitting

It contains two main executables

1) UpdatePDFs

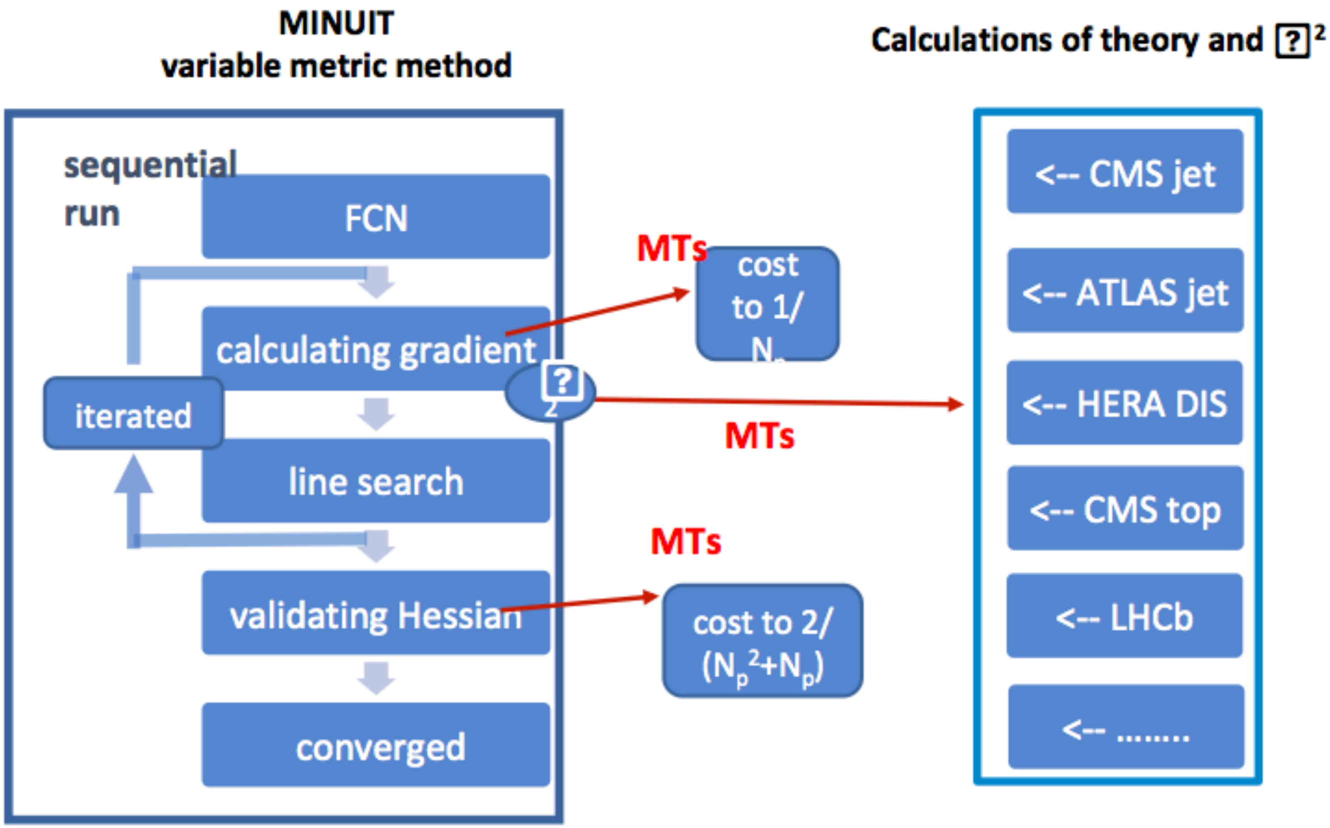
- Given the original theory predictions for a set of observables, and the experimental data for some subset of the observables, it computes the updated predictions and uncertainties for all of the observables, incorporating the effects of the new data.
- If the original best-fit and Hessian error PDFs are supplied, it also computes updated best-fit and Hessian error PDFs that incorporate the effects of the new data.

2) OptimizePDFs

- Given the original best-fit and Hessian error PDFs, along with theory predictions for a set of observables, it computes a new set of Hessian error PDFs that are optimized for the particular set of observables.
- The new set of error PDFs produce equivalent results to the original set of error PDFs (at least in the linear/ quadratic approximations assumed in the Hessian method). However, each new Hessian eigenvector PDF has an associated eigenvalue that gives the sum of the relative contributions of that eigenvector direction to the variance of each of the observables in the given set. The eigenvalues can be used to choose a reduced set of Hessian error PDFs by discarding those that are irrelevant to the given set of observables.

Fitting code parallelization with multi-threads

upgrade to a parallelized version of the fitting code, two-layer parallelization: 1. through rearrangement of the minimization algorithm; 2. via redistribution of the data sets



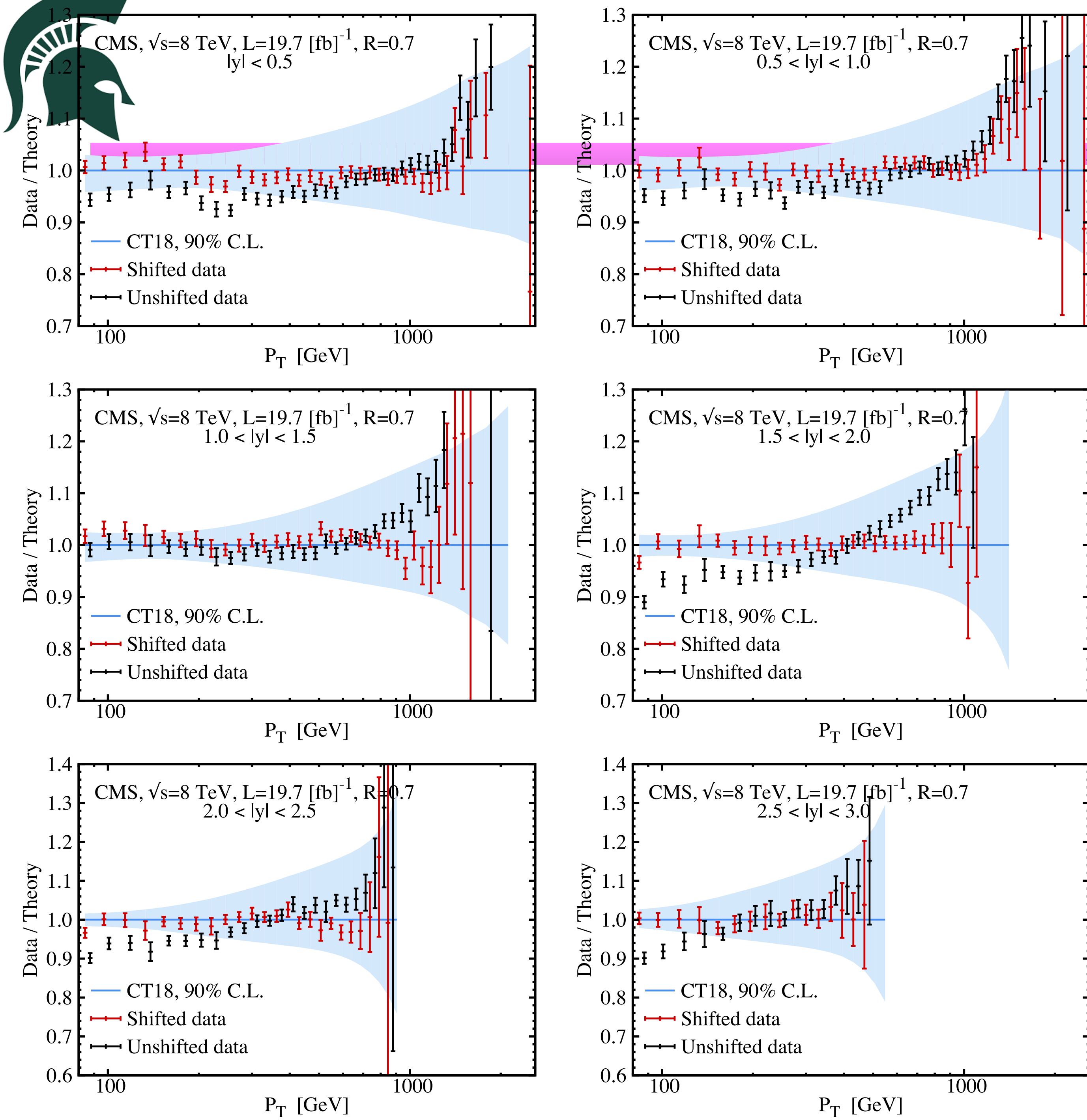
Layer 1: after all a factor of 4~5 improvement on speed is achieved!

Layer 2: further speed up by a factor of 2, depending on data sets included

Results of the fit (LHC observables)

ID	Observable	Npts	χ^2/N(CT18)	χ^2/N(CT18Z)
245	LHCb 7 TeV ZW	33	1.5	1.2
246	LHCb 8 TeV Z	17	1.4	1.2
248	ATLAS 7 TeV ZW			2.4
249	CMS 8 TeV W asym.	11	0.5	0.5
250	LHCb 8 TeV WZ	34	2.1	1.7
253	ATLAS 8 TeV DY $d^2\sigma/dydM$	27	1.6	1.4
542	CMS 7 TeV jet	158	1.3	1.3
544	ATLAS 7 TeV jet	140	1.5	1.5
545	CMS 8 TeV jet	185	1.3	1.2
573	CMS 8 TeV $d^2\sigma_{t\bar{t}}/dp_T^t dy^t$	16	1.9	1.9
580	ATLAS 8 TeV $d\sigma_{t\bar{t}}/dp_T^t$ and $d\sigma_{t\bar{t}}/dm_{t\bar{t}}$	15	1.1	1.4
Total		3493/3681	1.2	1.2

CMS 8 TeV

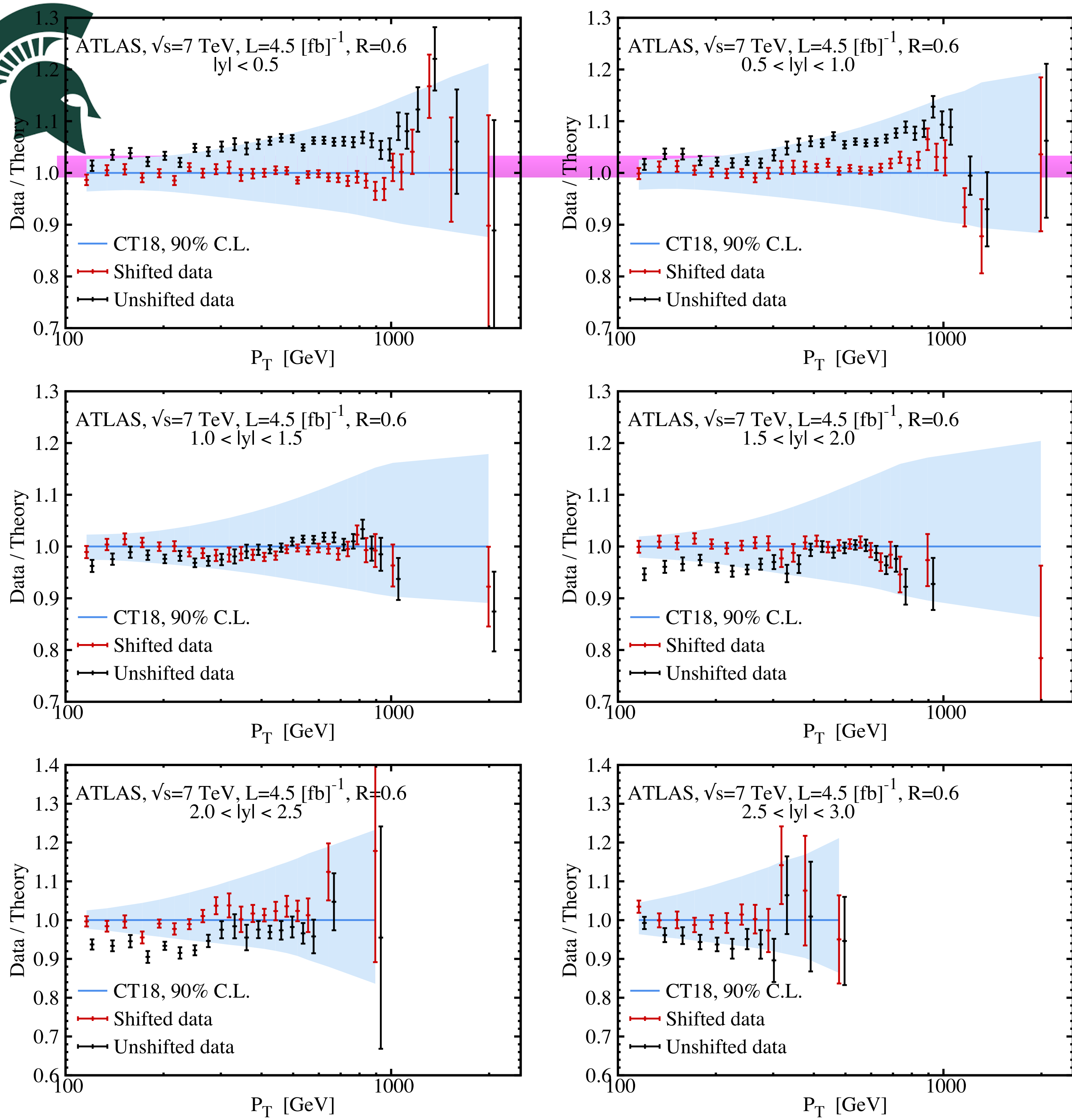


drop the low pileup data from 21-74 GeV

χ^2=168 (for 185 points) before fitting; 132 after fitting

PDFSense predicted this would be the highest impact LHC data set

results in some reduction in gluon uncertainty at moderate and high x



140 data points

χ^2=1.5/dof
after fitting

moderate impact
on gluon
uncertainty