PDF fits including top and learning from PDF fits for EFT fit

C.-P. Yuan
Michigan State University
Wu-Ki Tung Endowed Professor

November 24, 2020

LHC top WG meeting

For the CTEQ-TEA Collaboration

CTEQ – Tung et al. (TEA)
in memory of Prof. Wu-Ki Tung

Outline

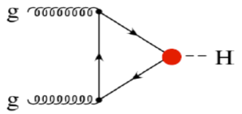
- Motivation
- Probing top quark Yukawa coupling in EFT
- “New Physics Found” vs. “refining PDFs”
- CT18 family PDFs
- Impact of LHC 8 TeV top quark pair data to CT18
- Fits to CMS 13 TeV top quark pair data
- Lessons learned from PDF fits for EFT

Probing top quark Yukawa coupling in EFT

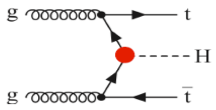
Li, Xu, Yan, CPY, arXiv: 1904.12006

Measuring Top Yukawa coupling at the LHC

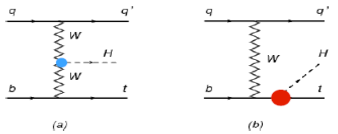
- Indirect probe: gluon fusion



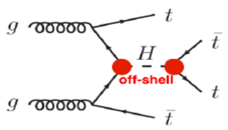
- Htt associated production



- Htj associated production



- Multi-top production



- In Unitarity gauge, two dim-6 effective operators can modify top quark Yukawa coupling:

$$c_t = \frac{2+3c_y}{\sqrt{1+c_H(2+c_y)}}$$

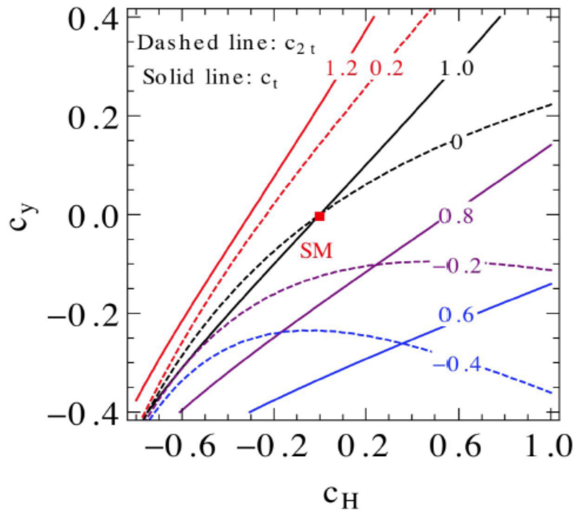
$$\mathcal{O}_H = \frac{1}{2v^2}\partial^\mu(H^\dagger H)\partial_\mu(H^\dagger H)$$

$$\mathcal{O}_y = -\frac{y_t^{\rm SM}}{v^2}H^\dagger H\bar{Q}_L\tilde{H}t_R$$

- SM corresponds to $c_t = 1$
- Degeneracy exist for two effective operators!
  - If $c_t \neq 1$, new physics is discovered
  - If $c_t = 1$, there is still possible new physics

Can we distinguish the effects of these effective operators?

Breaking the degeneracy!



- Htt and HHtt couplings:

$$c_t = \frac{2+3c_y}{\sqrt{1+c_H(2+c_y)}}$$

$$c_{2t} = \frac{c_H(3c_y-2)+6c_y}{2(c_H+1)^2(c_y+2)}$$

- The dependence on Wilson coefficients are different
- They can be measured in Higgs pair production

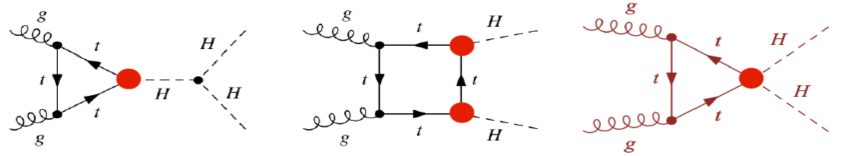
Breaking the degeneracy with Higgs Pair Production!

$\mathcal{O}_H = \frac{1}{2v^2}\partial^\mu(H^\dagger H)\partial_\mu(H^\dagger H)$ $\quad$ $\mathcal{O}_y = -\frac{y_t^{\rm SM}}{v^2}H^\dagger H\bar{Q}_L\tilde{H}t_R$

⇩

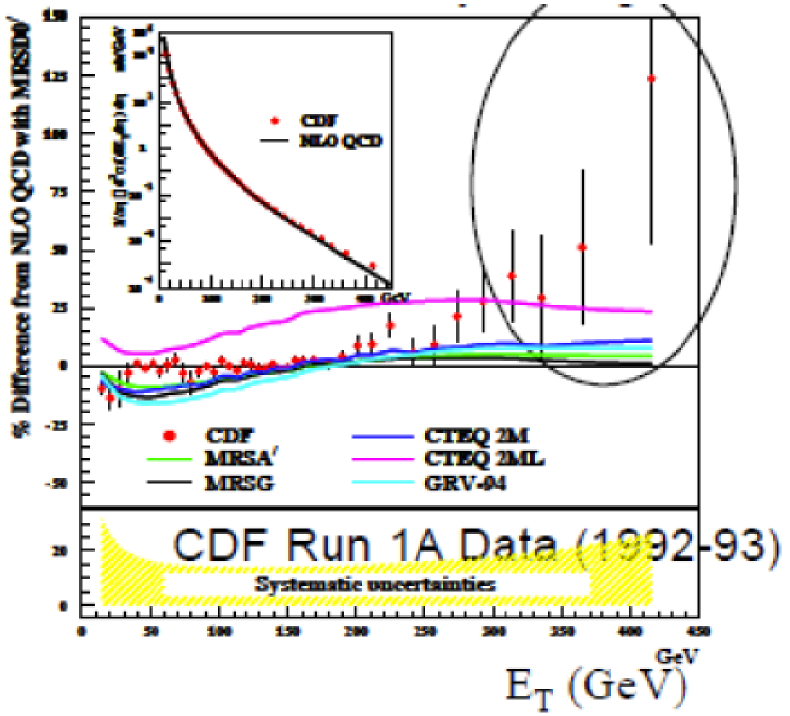
Higgs effective couplings: c_t $\quad$ c_{2t}

⇩



New Physics Found (in 1996) ?

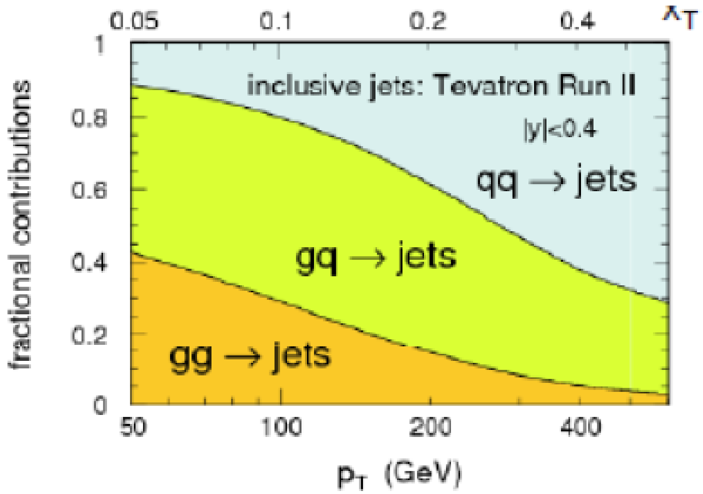
C T E Q

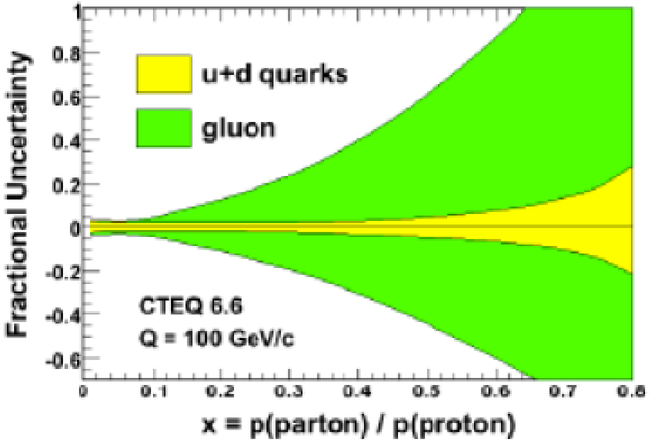


Phys. Rev. Lett. 77, 438 (1996)

High-x gluon not well known

...can be accommodated in the Standard Model





Explained by having better determined PDFs from global analysis

J. Huston, E. Kovacs, S. Kuhlmann, J.L. Lai, J.F. Owens, D. Soper, W.K. Tung , Phys. Rev. Lett. 77 (1996) 444.

CT18 family PDFs

Hou, et al., arXiv: 1912.10053

- Start with CT14-HERAII (HERAII combined data released after publication of CT14)
- Examine a wide range of non-perturbative 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 (single and double differential) data ⇒ compared to NNLO theory predictions.
- 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 PDFSense, ePump results and global fit
- Lagrange Multiplier studies to examine constraints of specific data sets on PDF distributions, or on $\alpha_s(m_Z)$ and (in some cases) the tensions (useful information)

LHC data sets included in CT18

C T E Q

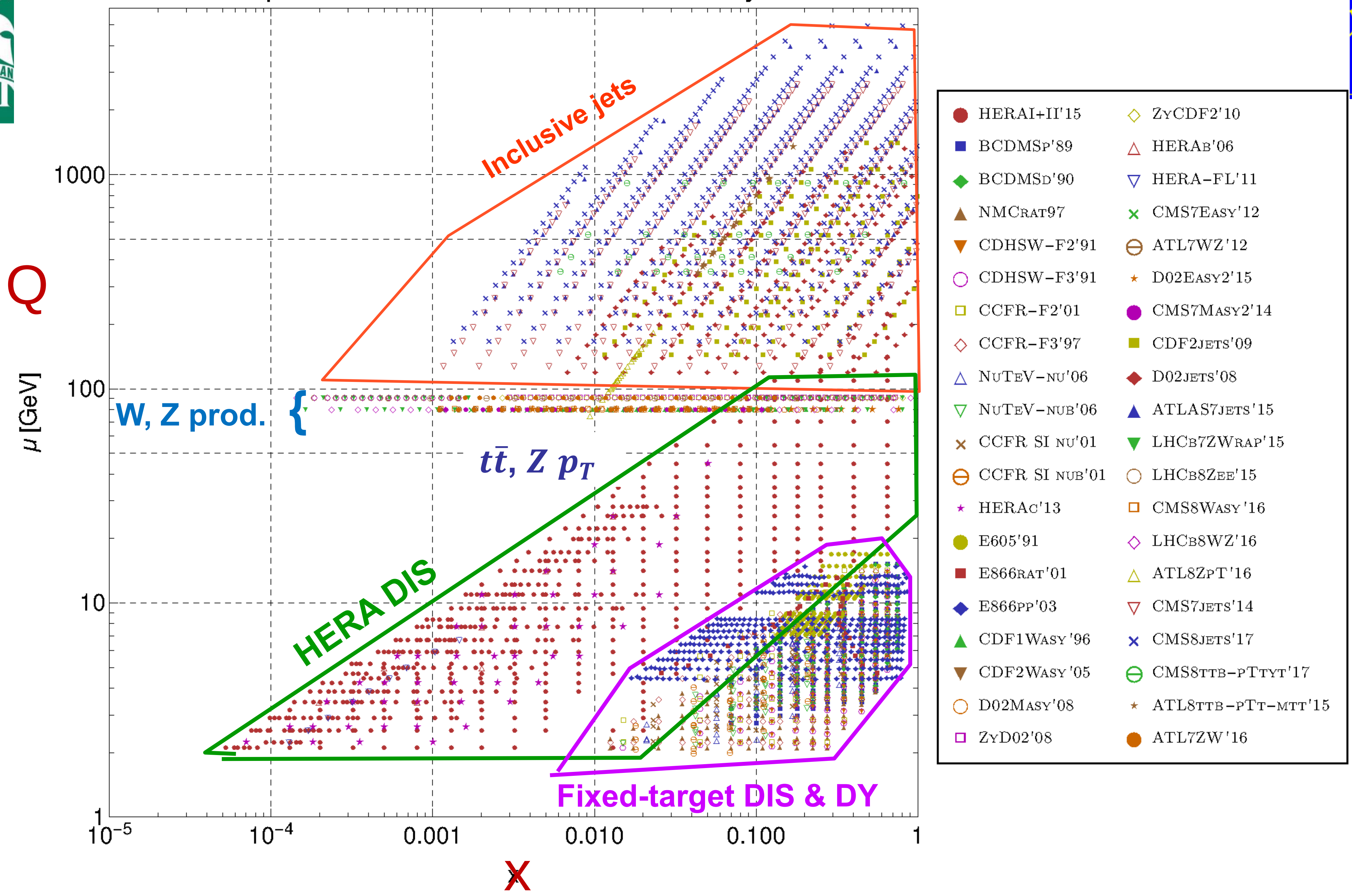
245 1505.07024 LHCb Z (W) muon rapidity at 7 TeV(applgrid)
246 1503.00963 LHCb 8 TeV Z rapidity (applgrid);
249 1603.01803 CMS W lepton asymmetry at 8 TeV (applgrid)
250 1511.08039 LHCb Z (W) muon rapidity at 8 TeV(applgrid)
253 1512.02192 ATLAS 7 TeV Z pT (applgrid)

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)

580 1511.04716 ATLAS 8 TeV tT pT diff. distributions (fastNNLO)
+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 PDFs

Experimental data in CT18 PDF analysis



CT18 LHC data treatment

- CT18 analysis includes new LHC experiments on W/Z, high-p_T Z, jet, $t\bar{t}$ production; up to 30 candidate LHC data sets available
- The challenge is to select and implement relevant and consistent experiments
- We include as large a rapidity interval for the ATLAS jet data as we can, using the ATLAS de-correlation model, rather than using a single rapidity interval. Using a single rapidity interval may result in selection bias.
- We use two $t\bar{t}$ single differential observables from ATLAS (using statistical correlations) and double differential measurement from CMS in order to include as much information as possible. Again, there is a risk of bias, as some of the observables are in tension with each other.
- Previous data continue having an impact on global fits and tend to dilute the impact of new data

CT18 family PDFs

- CT18 (main PDF)
- CT18A (including ATLAS 7 TeV precision W and Z data)
- CT18X (special scale for DIS; mimicking small-x resummation)
- CT18Z (including both ATLAS 7 TeV W/Z data and special scale for DIS)

➡

- CT18A differ from CT18 mainly in s-PDF.
- CT18X differ from CT18 mainly in g-PDF.
- CT18Z represents the maximal difference from CT18, particularly on g-PDF and s-PDF.

Findings

- The CT18 PDF uncertainty is mildly reduced at NNLO compared to the CT14 PDF uncertainty.
- 700+ data points from 12 new LHC data sets. The LHC constraints on the CT18 PDFs are weaken by some inconsistencies between the LHC data sets and the pre-LHC data sets.
- HERA DIS and fixed-target experiments deliver key constraints on CT18 PDFs.
- We observe some impact on PDFs from ATLAS and CMS incl. jet data, ATLAS, CMS, LHCb W/Z data and ATLAS 8 TeV Z pT data.
- LHC top quark pair data provides a similar impact to g-PDF as incl. jet data, but cannot reduce g-PDF errors as strong as incl. jet data due to its much smaller number of data points.
- ATLAS 7 TeV W and Z rapidity data is included in the CT18Z PDF analysis.

Selected Top Quark Pair Observables from ATLAS and CMS

- Modest effect observed if t-tbar data are included together with the Tevatron and LHC jet production data.
- Its impact on gluon PDF is consistent with jet data, though jet data provide stronger constraint.
- For ATLAS 8 TeV, select the pT and mtT distributions that directly probes large-x region; statistical correlations are included in order to fit pT and mtT simultaneously; fully correlated for experimental systematics except for decorrelation of PS sys. error.

	χ^2/N_{pt} (with CT18 PDFs)	nominal	w/o PS decorrelation	w/o statistical correlation
ATLAS 8 TeV 1511.04716	ATLAS 8 TeV abs. dσ/d pT & dσ/d mtt (Npts=15)	0.62	3.55	0.51
CMS 8 TeV 1703.01630	CMS 8 TeV nor. $d^2\sigma/(d\ pT\ d\ yt)$ (Npts=16)	1.18	—-	—-

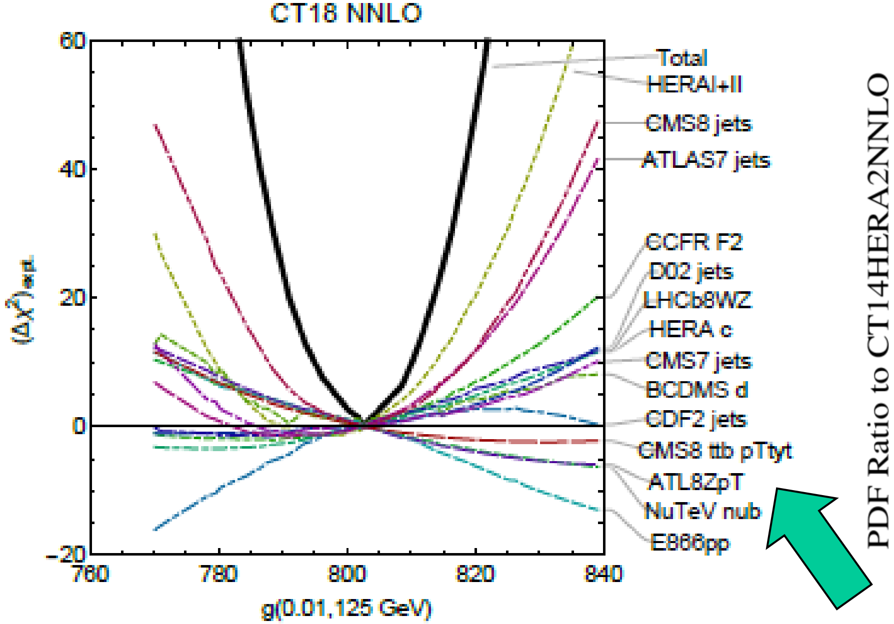
Theory calculations @NNLO

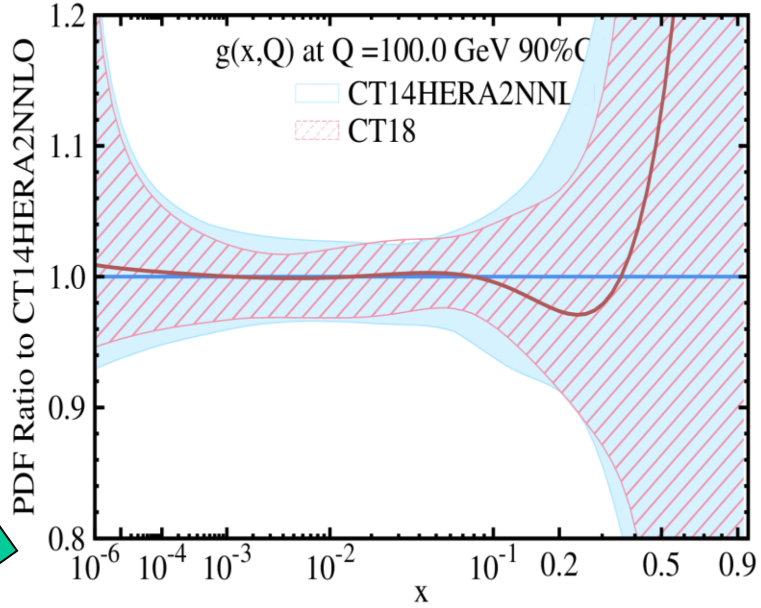
- Jet pT, (W,Z) rapidity, Z pT, t-tbar

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			$\frac{H_T}{4}, \frac{m_T}{2}$

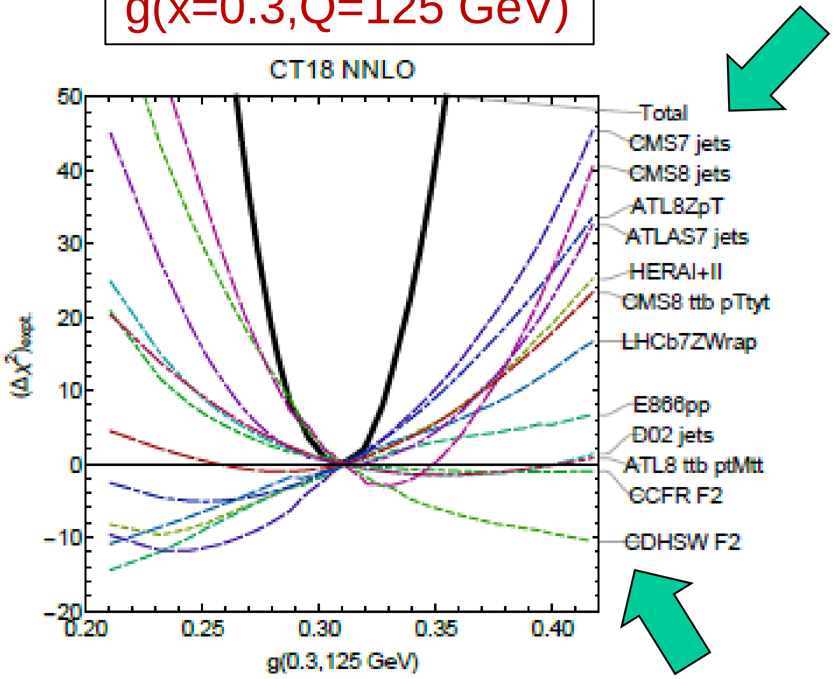
CT18 NNLO PDFs (g-PDF)

g(x=0.01,Q=125 GeV)





g(x=0.3,Q=125 GeV)

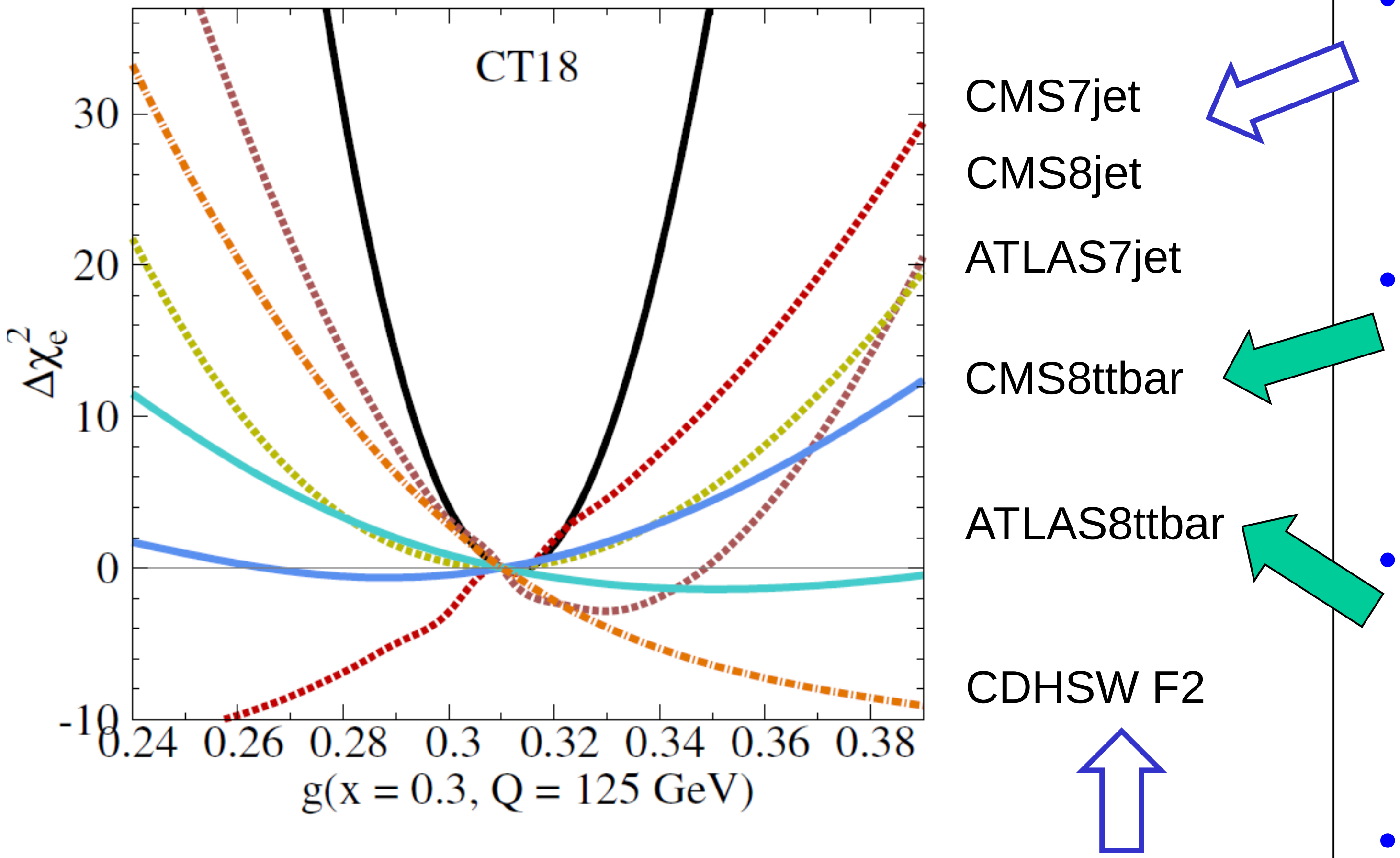


Lagrange Multiplier Scans

- At x around 0.01, ATLAS8 Z pT data prefer a slightly larger gluon PDF.
- At x around 0.3, competing with the CDHSW F2 and Tevatron jet data, which prefer larger gluon, the ATLAS7 jet, CMS7 jet and ATLAS8 Z pT data prefer a smaller gluon; some tension found in CMS7 and CMS8 jet data.
- The gluon PDF as $x \to 1$ is parametrization form dependent.

CT18 NNLO PDFs (g-PDF)

g(x=0.3; Q=125 GeV)



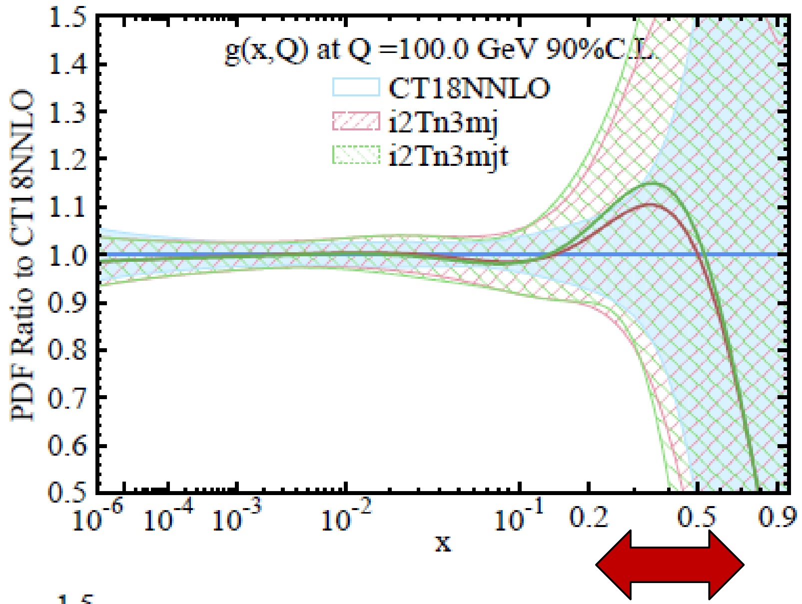
Lagrange Multiplier Scans

- At large x=0.3, Q=125 GeV, CMS + ATLAS jet data and CDHSW F2 data dominate the constraint on g-PDF.
- CMS 8 TeV t-tbar (pT_t,y_t) data provides similar constraint as HERA I+II data on g-PDF, favoring softer gluon.
- ATLAS 8 TeV t-tbar (pT_t ,M_tt) data provides similar constraint as D0 Run 2 jet data on g-PDF, favoring harder gluon.
- Some tension found in CMS7 (favoring softer gluon) and CMS8 (favoring harder gluon) jet data.

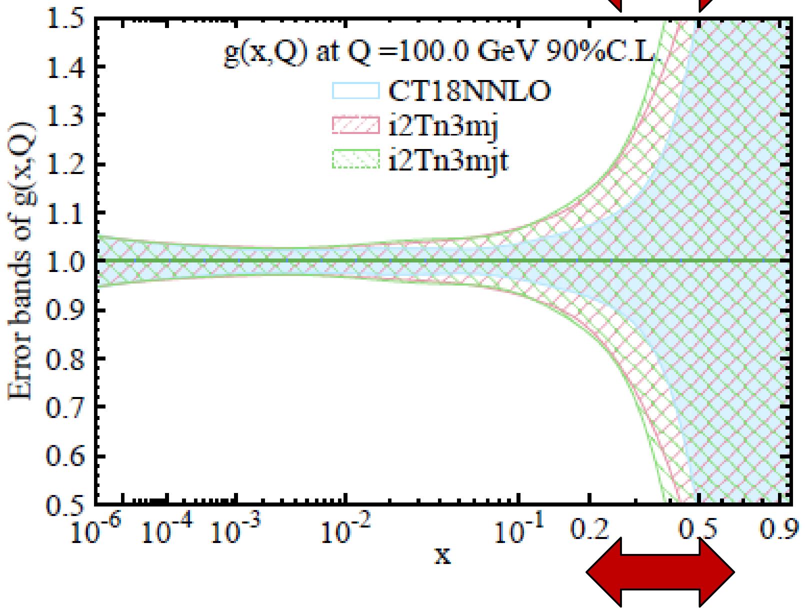
Impact of LHC 8 TeV t-tbar data to CT18

C T E Q

Compare g-PDFs in various global fits at 90% CL



Compare g-PDF error bands



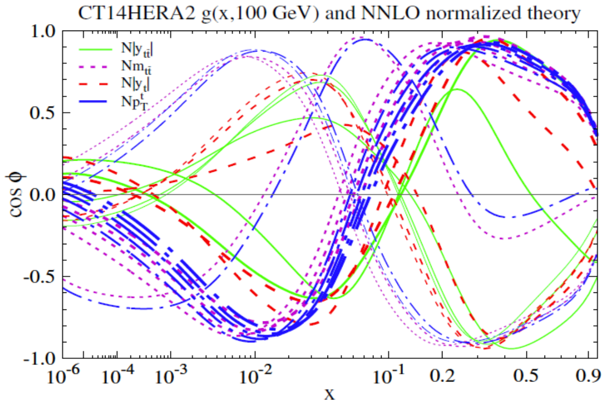
- g-PDF ratios to CT18NNLO PDFs (in blue)
- Red curve (and its PDF error band) is a global fit with ALL the jet data removed from the CT18 fit.
- Green curve is a global fit with ALL the jet data and LHC 8 TeV t-tbar data (ID=573 and 580) removed from the CT18 fit.
- g-PDF at large x is mainly constrained by the jet data, especially CMS and ATLAS jet data.
- The impact of LHC 8 TeV t-tbar data is to reduce g-PDF and shrink its error band at x around 0.2-0.5.

Which exp. obs. to use in the global fit?

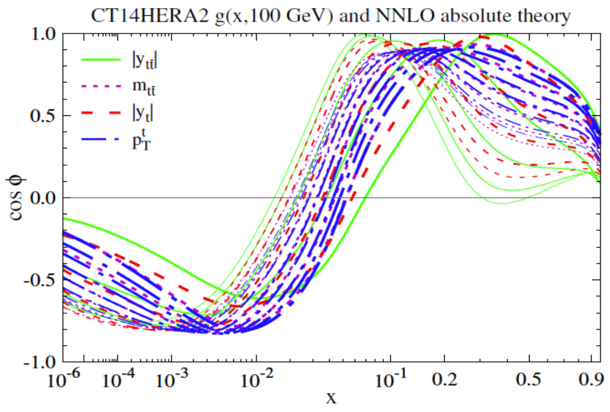
Kadir, et al, arXiv: 2003.13740

TABLE II: Number of data points and χ^2/N_{pts} for inclusive jet and top-quark pair data, after ePump updating from the CT14HERA2 and CT14HERA2mJ PDFs.

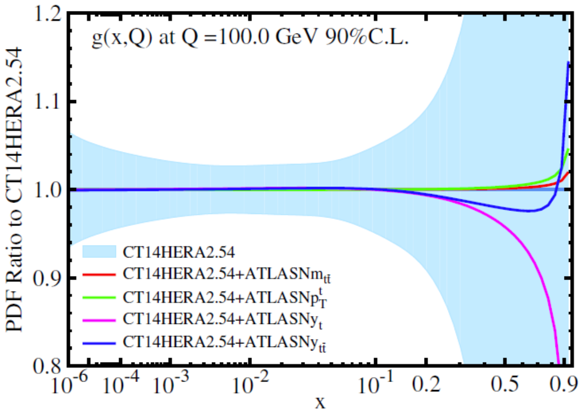
Observable	Detector		N_{pts}	χ^2/N_{pts}(CT14HERA2) weight=1.0	χ^2/N_{pts}(CT14HERA2) weight=9.0	χ^2/N_{pts}(CT14HERA2mJ)
inclusive jet	CDF	[8]	72	1.46		1.50
inclusive jet	D0	[9]	110	1.03		1.03
inclusive jet	ATLAS	[10]	90	0.57		0.57
inclusive jet	CMS	[11]	133	0.89		0.93
$\frac{1}{\sigma}\frac{d\sigma}{dp_T^t}$, $\frac{1}{\sigma}\frac{d\sigma}{dp_T^t}$	ATLAS, CMS	[6, 7]	8,8	0.39, 3.55	0.38, 2.20	0.38, 4.82
$\frac{1}{\sigma}\frac{d\sigma}{d\|y_t\|}$, $\frac{1}{\sigma}\frac{d\sigma}{dy_t}$	ATLAS, CMS	[6, 7]	5,10	2.40, 2.52	1.45, 2.50	5.34, 3.32
$\frac{1}{\sigma}\frac{d\sigma}{dm_{t\bar{t}}}$, $\frac{1}{\sigma}\frac{d\sigma}{dm_{t\bar{t}}}$	ATLAS, CMS	[6, 7]	7,7	0.25, 7.69	0.25, 3.96	0.35, 9.30
$\frac{1}{\sigma}\frac{d\sigma}{d\|y_{t\bar{t}}\|}$, $\frac{1}{\sigma}\frac{d\sigma}{dy_{t\bar{t}}}$	ATLAS, CMS	[6, 7]	5,10	2.21, 2.31	1.18, 1.07	5.21, 3.34
$\frac{d\sigma}{dp_T^t}$	ATLAS	[6]	8	0.34	0.33	0.32
$\frac{d\sigma}{dy_t}$	ATLAS	[6]	5	2.83	1.62	5.79
$\frac{d\sigma}{dm_{t\bar{t}}}$	ATLAS	[6]	7	0.45	0.42	0.40
$\frac{d\sigma}{dy_{t\bar{t}}}$	ATLAS	[6]	5	3.83	1.48	7.29



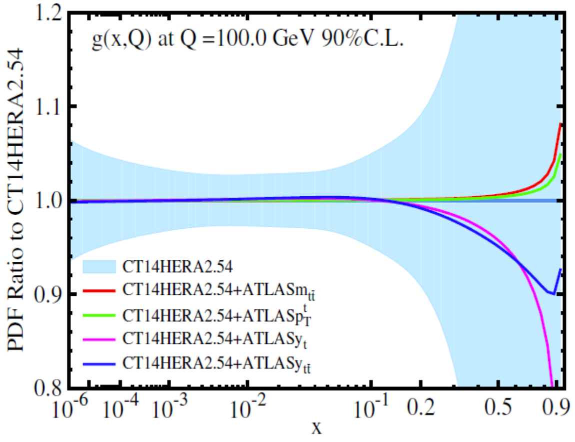
ATLAS norm. data



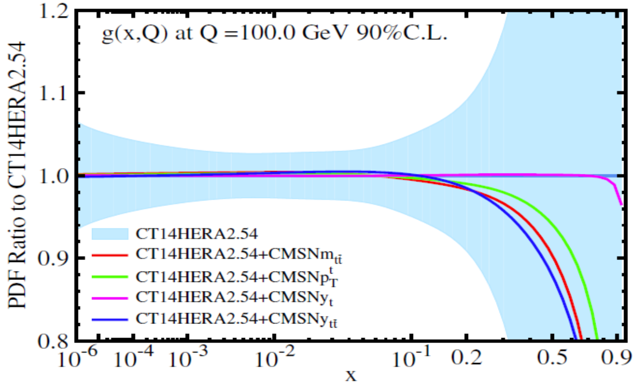
ATLAS abs. data



ATLAS norm. data



ATLAS abs. data



CMS norm. data

Impact of t-tbar data

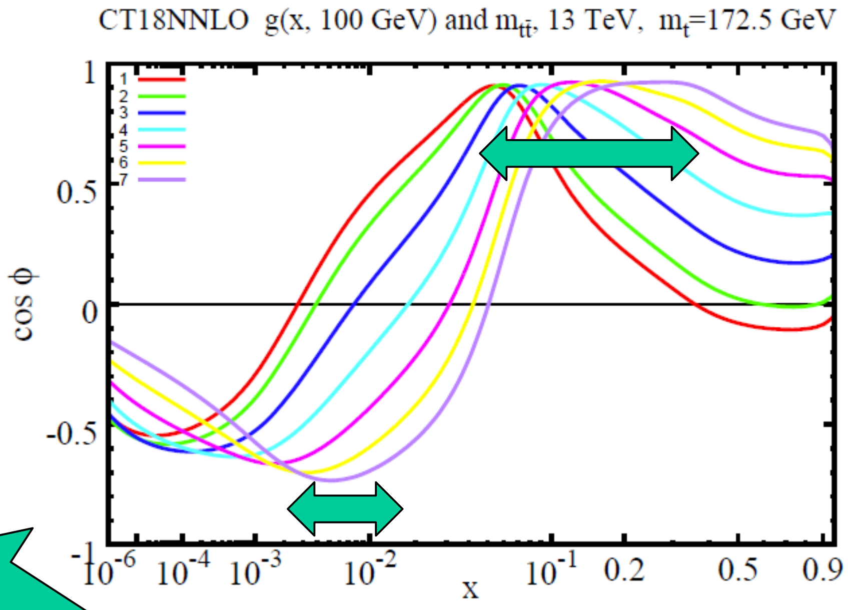
- g-PDF will be modified in the large x region, consistent with inclusive jet data constraint, but depending on the chosen (absolute or normalized) exp. observables (m_tt, pT_t, y_t, y_tt, etc.)
- Need to carefully include statistical and (correlated) systematic errors.
- Use double-differential top distributions to include more data information.

Czakon et al, arXiv:1912.08801

Fits to CMS 13 TeV top quark pair data

χ^2/Npt table for CMS 13 t-tbar dis., CT18NNLO (mt=172.5 GeV)

χ^2/Npt	Absolute obs. 1811.06625		Normalized obs. 1811.06625	
obs.	CT18	ePump updated	CT18	ePump updated
m_tt	3.57	3.42	5.38	5.26
pT_t	3.59	3.46	3.96	3.88
pT_tb	3.52	3.38	3.68	3.63
pT_tt	3.24	3.17	5.31	5.30
y_t	0.76	0.74	0.90	0.89
y_tb	0.58	0.56	0.56	0.55
y_tt	0.86	0.79	0.88	0.81
Δy_tt	1.03	1.00	1.23	1.19



$d\sigma/dm_{t\bar{t}}$

- Sensitive to g-PDF at x in [0.05-0.5] and around 0.01

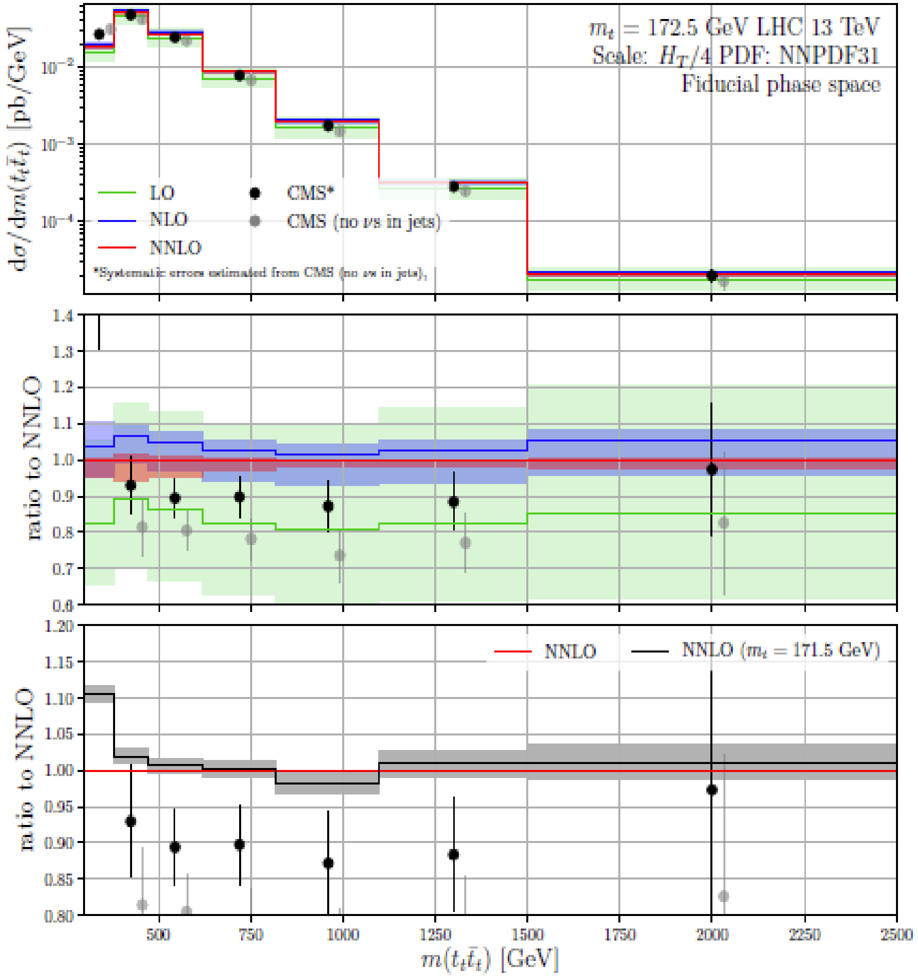
Not a good fit, why?

- Revised Data Analysis
- Improved Theory calculations

- Removing the first M_{tt} data point, which is important for determining top quark mass, does not change the fitted g-PDF.

Revised Data Analysis

$d\sigma/dm_{t\bar{t}}$ (abs. xsec; true top; fiducial phase space)



CMS Data: Sirunyan et al, arXiv: 1181.06625

- Black data points; CMS*
- Gray data points; (no neutrino in jets); not include the neutrinos from semileptonic B-meson decays; breaking of partonic-particle jet equivalence

Theory calculation:

Czakon, Mitov, Poncelet; arXiv: 2008.11133

Improved Theory Calculations (beyond NNLO in QCD)

- NNLO QCD + NLO QED

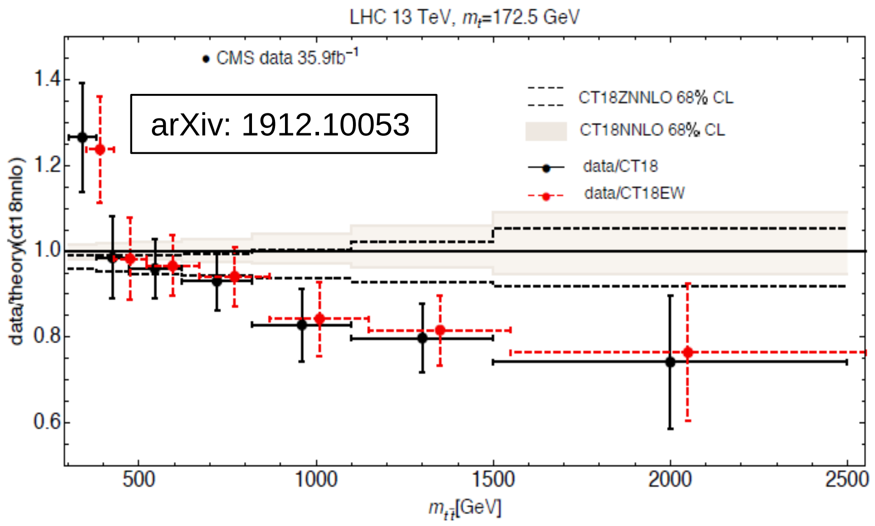
  arXiv: 1705.04105

- Resummation:
  - Soft gluon resummation in the threshold and in the boosted-soft limit

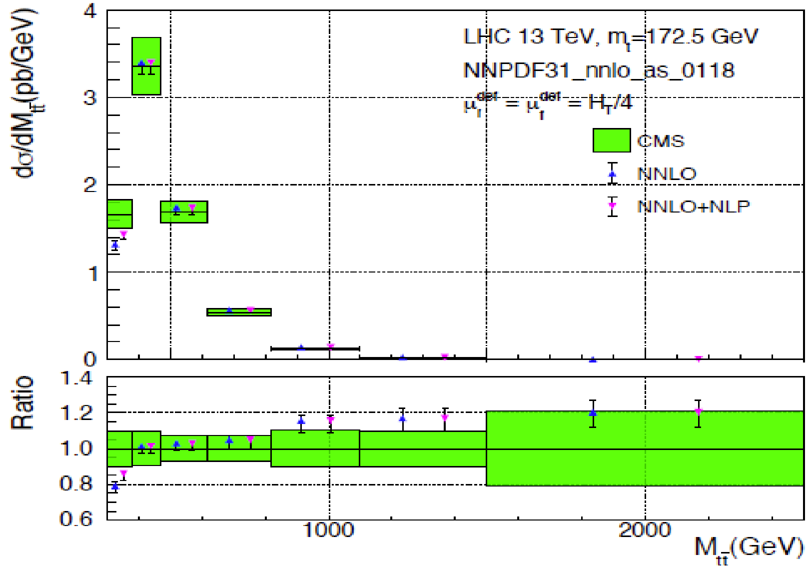
    arXiv: 1803.07623

  - Resum Coulomb corrections near the threshold region

    arXiv: 1908.02179, 2004.03088



$$d\sigma/dm_{t\bar{t}}$$



Lessons learned from PDF fits for EFT fit

Experimentalists + Theorists

2017 Featured Story #1: Million-dollar gift establishes endowed professorship in honor of the late Dr. Wu-Ki Tung

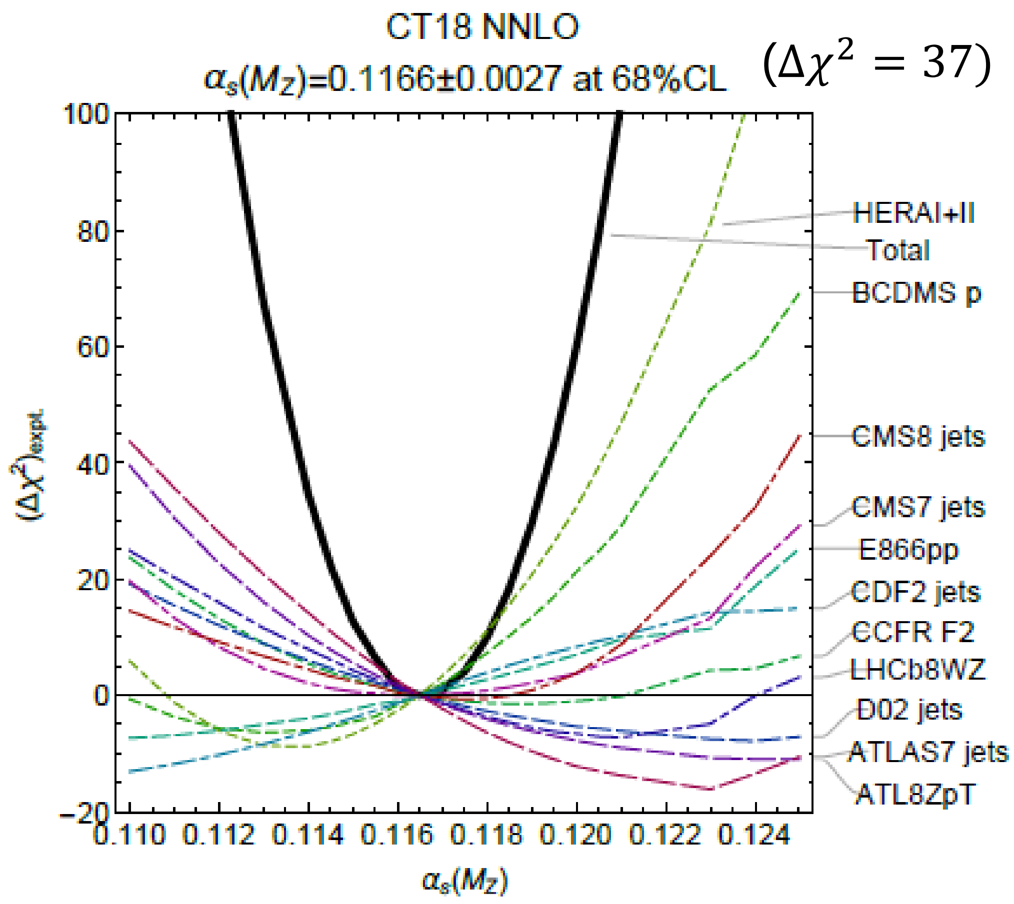
Michigan State University
(1992-2009)

http://www.pa.msu.edu/node/5921

- Co-founder of CTEQ (**The Coordinated Theoretical-Experimental Project on QCD)** in 1989 – present
- Nowadays, many, like this one -- **LHC Top WG**, are doing precisely that.

Backup slides

α_s (Mz) for CT18



HERA I+II

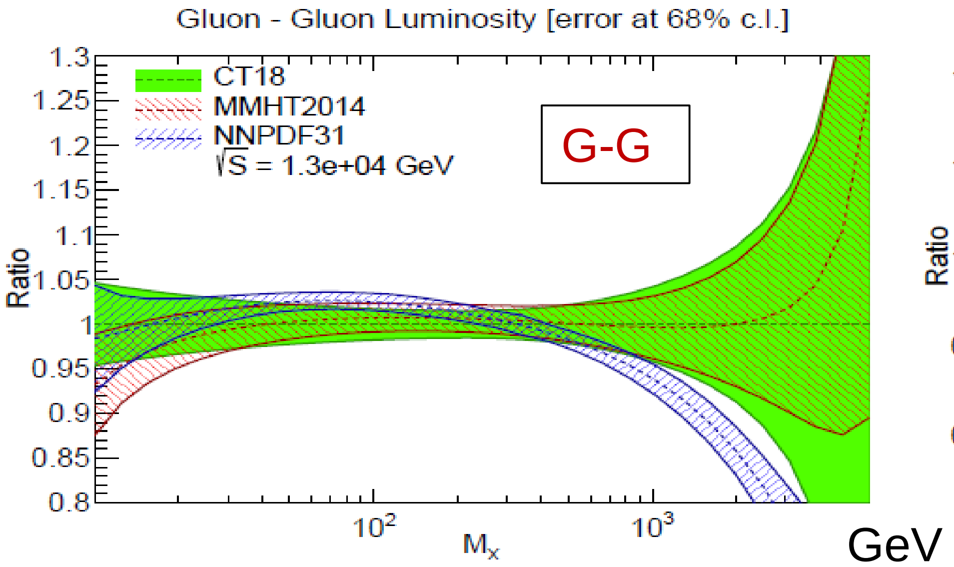
Lagrange Multiplier scan

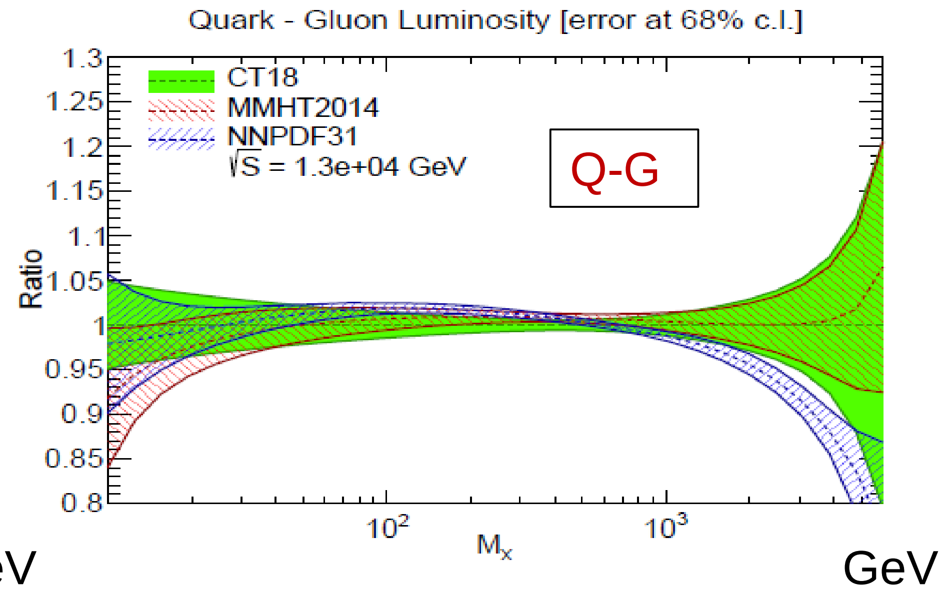
ATLAS 7 jet

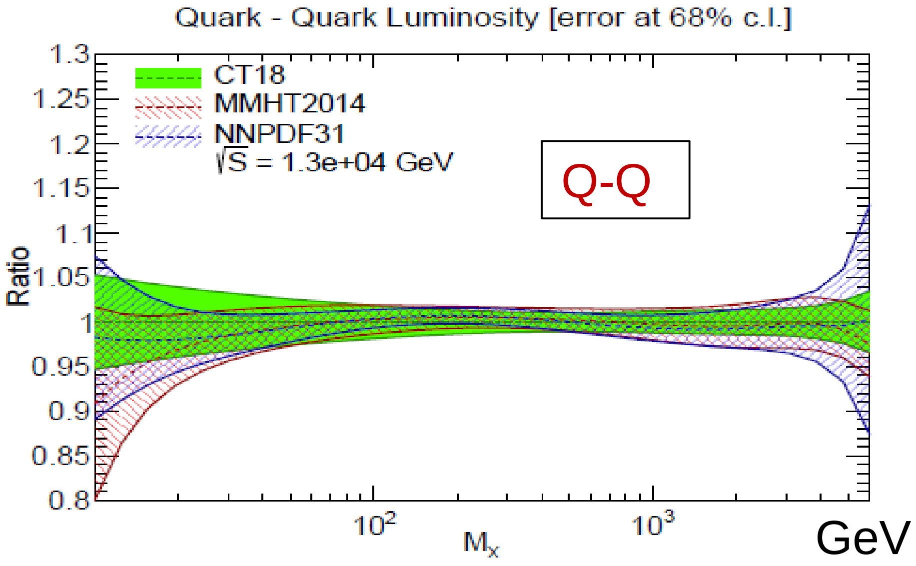
ATLAS 8 Z pT

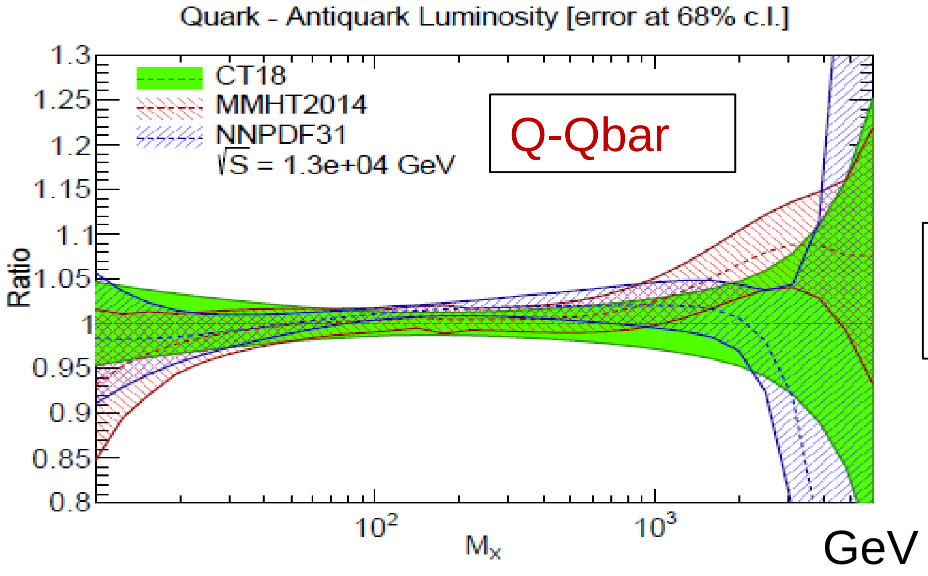
- The fixed target F2 data and HERA DIS data prefer smaller αs value.
- The ATLAS 8TeV Z pT and ATLAS 7 TeV incl. jet data, bring the central value of αs (Mz) from $0.115^{+0.006}_{-0.004}$ (CT14) to 0.1166 ± 0.0027 (CT18).

PDF Luminosities at 13 TeV LHC
CT18, MMHT14 and NNPDF3.1









PDF errors at 68% CL

CT18Z
LHC data treatment

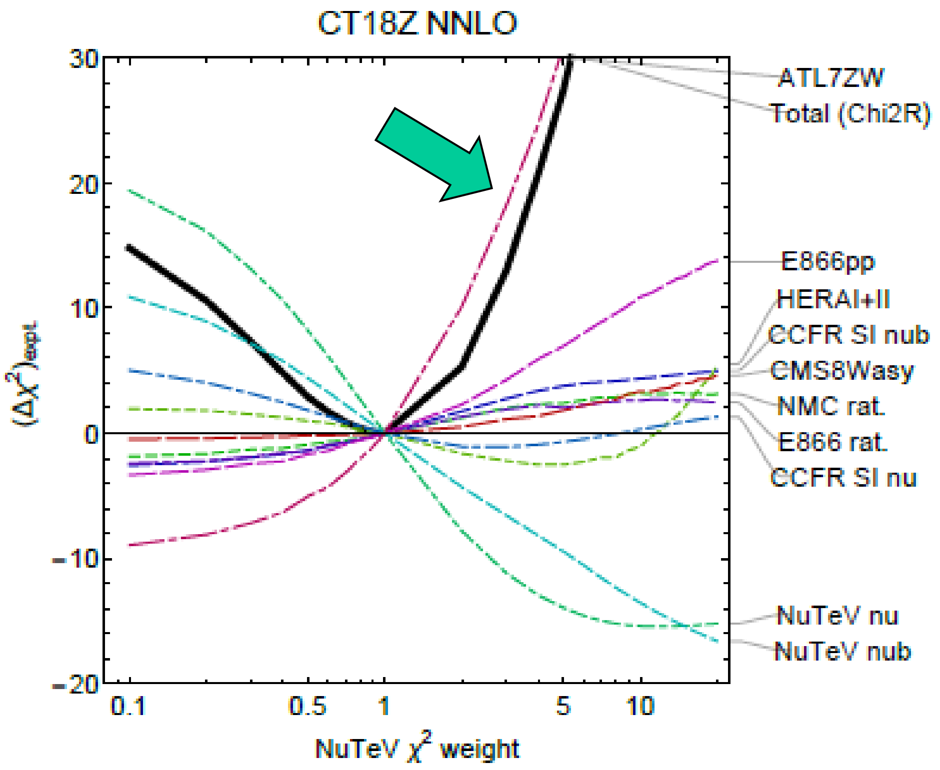
CTEQ

- Start with CT18 data set
- Add in ATLAS 7 TeV *W* and *Z* rapidity data (arXiv:1612.03016; 4.6 1/fb); large chi^2/d.o.f ~ 2.1
- Remove CDHSW data
- Use a special x-dependent factorization scale $\mu_{DIS,x}$ at NNLO calculation
- CT18Z uses a combination of $\mu_{DIS,x}$ (preferred by DIS) and an increased $m_c^{pole} = 1.4$ GeV (preferred by LHC vector boson production, disfavored by DIS)

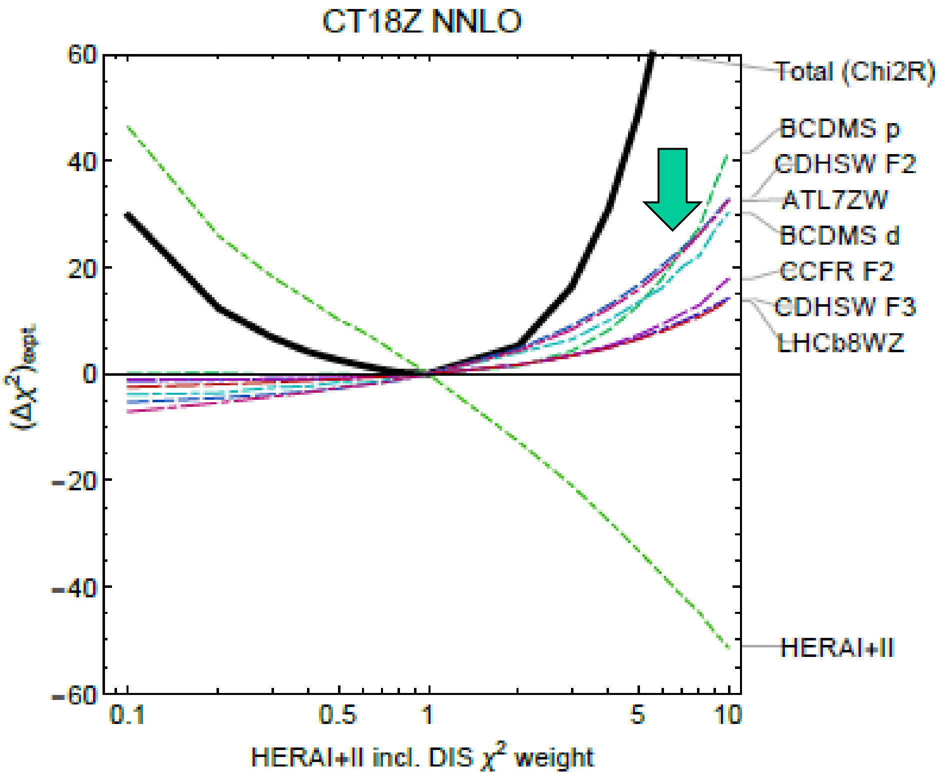
→ CT18Z PDFs

CT18Z fit

- ATLAS 7 TeV W and Z rapidity data have obvious tensions with NuTeV di-muon data; and some tension with HERA I+II data.



NuTeV di-muon data



HERA I+II data