

# **Heavy-flavor production in global QCD analyses**

**Marco Guzzi**

**Kennesaw State University**

**with**

**A. Ablat, S. Dulat, T.-J. Hou, P. Nadolsky**

**I. Sitiwaldi, K. Xie, and C.-P. Yuan**

# Summary of CTEQ activities on HF

- Impact of Heavy-Flavor (HF) production at hadron colliders on the structure of the proton:
  - $t\bar{t}$  @LHC13, c/b production @HERA, (this talk)
  - c/b production @LHCb13, Z+c/b @LHC13 (Keping Xie)
- Nonperturbative charm in the proton (Tim Hobbs)
- HF treatments in precision theory predictions at the LHC
- Keping Xie and I, will show results of recent studies **beyond** CT18

# Motivations I (what are data telling us?)

## Goals:

- Assess the impact of heavy-flavor (HF) production on unpolarized collinear PDFs
- Improve PDF uncertainties in global QCD analyses (in particular, constrain HF PDFs)
- Probe QCD dynamics at small and large  $x$

Important for: precision BSM searches, precise and accurate theory predictions (pQCD),...

## Heavy-flavor treatment in theory predictions (GMVFN schemes):

- DIS: S-ACOT- $\chi$  (default in CTEQ analyses) based on Collins' HF factorization in DIS
- Extend S-ACOT to PP collisions: S-ACOT-MPS. Now at NLO. NNLO needed.
- Implemented for inclusive charm [[FPF, 2109.10905](#), [2203.05090](#)] and bottom [[2203.06207](#)] production.
- S-ACOT-MPS can be extended to other processes (See Keping Xie's talk).

# Motivations II

- Heavy-quark production at the LHC at small  $p_T$  and large rapidity  $y$  of the heavy quark: sensitive to PDFs at both small and large  $x$  (especially true for c/b production)

$$x_{1,2} \approx \frac{\sqrt{p_T^2 + m_Q^2}}{\sqrt{S}} e^{\pm y}$$

- In this kinematic region PDFs are poorly constrained by other experiments in global PDF fits.
- In particular, c/b production in the  $4 < |y| < 4.5$  rapidity range in pp collisions at the LHC 13 TeV can probe  $x \leq 10^{-5}$ . When  $p_T \geq 40$  GeV, it can probe  $x \geq 0.2$
- Top-quark pair production @LHC can probe the gluon PDF already at  $x \gtrsim 0.01$
- c/b prod. @HERA: sensitive to intermediate and small  $x$

# Motivations II

- Probing this regime (and beyond, at future facilities) will further shed light on interesting problems: **(nonperturbative) heavy-flavor content** of the proton (**See Tim Hobbs' slides**), and on **small-x dynamics**.
- LHC delivered precise measurements for these observables, (e.g.,  $t\bar{t}$  prod. at **ATLAS** and **CMS**; D-meson prod. at **LHCb**).
- In addition, we have the **HERA** legacy: most recent c/b production combination in DIS

# The CT18 analysis

## New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC

Tie-Jiun Hou,<sup>1,†</sup> Jun Gao,<sup>2</sup> T. J. Hobbs,<sup>3,4</sup> Keping Xie,<sup>3,5</sup> Sayipjamal Dulat,<sup>6,‡</sup> Marco Guzzi,<sup>7</sup> Joey Huston,<sup>8</sup> Pavel Nadolsky,<sup>9,§</sup> Jon Pumplin,<sup>8,\*</sup> Carl Schmidt,<sup>8</sup> Ibrahim Sitiwaldi,<sup>6</sup> Daniel Stump,<sup>8</sup> and C.-P. Yuan<sup>8,||</sup>

TABLE I. Datasets included in the CT18(Z) NNLO global analyses. Here we directly compare the quality of fit found for CT18 NNLO vs CT18Z NNLO on the basis of  $\chi_E^2$ ,  $\chi_E^2/N_{pt,E}$ , and  $S_E$ , in which  $N_{pt,E}$ ,  $\chi_E^2$  are the number of points and value of  $\chi^2$  for experiment  $E$  at the global minimum.  $S_E$  is the effective Gaussian parameter [38,42,56] quantifying agreement with each experiment. The ATLAS 7 TeV 35 pb<sup>-1</sup>  $W/Z$  dataset, marked by ‡, is replaced by the updated one (4.6 fb<sup>-1</sup>) in the CT18A and CT18Z fits. The CDHSW data, labeled by †, are not included in the CT18Z fit. The numbers in parentheses are for the CT18Z NNLO fit.

Exp. ID#	Experimental dataset		$N_{pt,E}$	$\chi_E^2$	$\chi_E^2/N_{pt,E}$	$S_E$
160	HERA1 + II 1 fb <sup>-1</sup> , H1 and ZEUS NC and CC $e^\pm p$ reduced cross sec. comb.	[30]	1120	1408 (1378)	1.3 (1.2)	5.7 (5.1)
101	BCDMS $F_2^p$	[57]	337	374 (384)	1.1 (1.1)	1.4 (1.8)
102	BCDMS $F_2^d$	[58]	250	280 (287)	1.1 (1.1)	1.3 (1.6)
104	NMC $F_2^d/F_2^p$	[59]	123	126 (116)	1.0 (0.9)	0.2 (-0.4)
108†	CDHSW $F_2^p$	[60]	85	85.6 (86.8)	1.0 (1.0)	0.1 (0.2)
109†	CDHSW $x_B F_3^p$	[60]	96	86.5 (85.6)	0.9 (0.9)	-0.7 (-0.7)
110	CCFR $F_2^p$	[61]	69	78.8 (76.0)	1.1 (1.1)	0.9 (0.6)
111	CCFR $x_B F_3^p$	[62]	86	33.8 (31.4)	0.4 (0.4)	-5.2 (-5.6)
124	NuTeV $\nu\mu\mu$ SIDIS	[63]	38	18.5 (30.3)	0.5 (0.8)	-2.7 (-0.9)
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS	[63]	33	38.5 (56.7)	1.2 (1.7)	0.7 (2.5)
126	CCFR $\nu\mu\mu$ SIDIS	[64]	40	29.9 (35.0)	0.7 (0.9)	-1.1 (-0.5)
127	CCFR $\bar{\nu}\mu\mu$ SIDIS	[64]	38	19.8 (18.7)	0.5 (0.5)	-2.5 (-2.7)
145	H1 $\sigma_r^p$	[65]	10	6.8 (7.0)	0.7 (0.7)	-0.6 (-0.6)
147	Combined HERA charm production	[66]	47	58.3 (56.4)	1.2 (1.2)	1.1 (1.0)
169	H1 $F_L$	[33]	9	17.0 (15.4)	1.9 (1.7)	1.7 (1.4)
201	E605 Drell-Yan process	[67]	119	103.4 (102.4)	0.9 (0.9)	-1.0 (-1.1)
203	E866 Drell-Yan process $\sigma_{pd}/(2\sigma_{pp})$	[68]	15	16.1 (17.9)	1.1 (1.2)	0.3 (0.6)
204	E866 Drell-Yan process $Q^3 d^2\sigma_{pp}/(dQ dx_F)$	[69]	184	244 (240)	1.3 (1.3)	2.9 (2.7)
225	CDF run-1 lepton $A_{ch}$ , $p_{T\ell} > 25$ GeV	[70]	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	[71]	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	[72]	9	9.1 (9.0)	1.0 (1.0)	0.2 (0.1)
260	DØ run-2 $Z$ rapidity	[73]	28	16.9 (18.7)	0.6 (0.7)	-1.7 (-1.3)
261	CDF run-2 $Z$ rapidity	[74]	29	48.7 (61.1)	1.7 (2.1)	2.2 (3.3)
266	CMS 7 TeV 4.7 fb <sup>-1</sup> , muon $A_{ch}$ , $p_{T\ell} > 35$ GeV	[75]	11	7.9 (12.2)	0.7 (1.1)	-0.6 (0.4)
267	CMS 7 TeV 840 pb <sup>-1</sup> , electron $A_{ch}$ , $p_{T\ell} > 35$ GeV	[76]	11	4.6 (5.5)	0.4 (0.5)	-1.6 (-1.3)
268‡	ATLAS 7 TeV 35 pb <sup>-1</sup> $W/Z$ cross sec., $A_{ch}$	[77]	41	44.4 (50.6)	1.1 (1.2)	0.4 (1.1)
281	DØ run-2 9.7 fb <sup>-1</sup> electron $A_{ch}$ , $p_{T\ell} > 25$ GeV	[78]	13	22.8 (20.5)	1.8 (1.6)	1.7 (1.4)
504	CDF run-2 inclusive jet production	[79]	72	122 (117)	1.7 (1.6)	3.5 (3.2)
514	DØ run-2 inclusive jet production	[80]	110	113.8 (115.2)	1.0 (1.0)	0.3 (0.4)

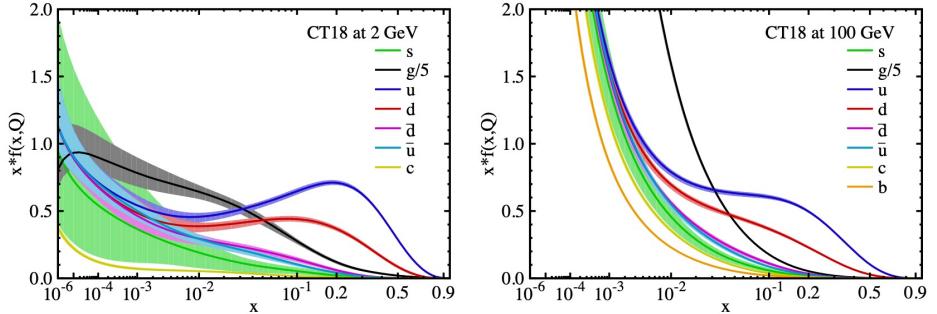


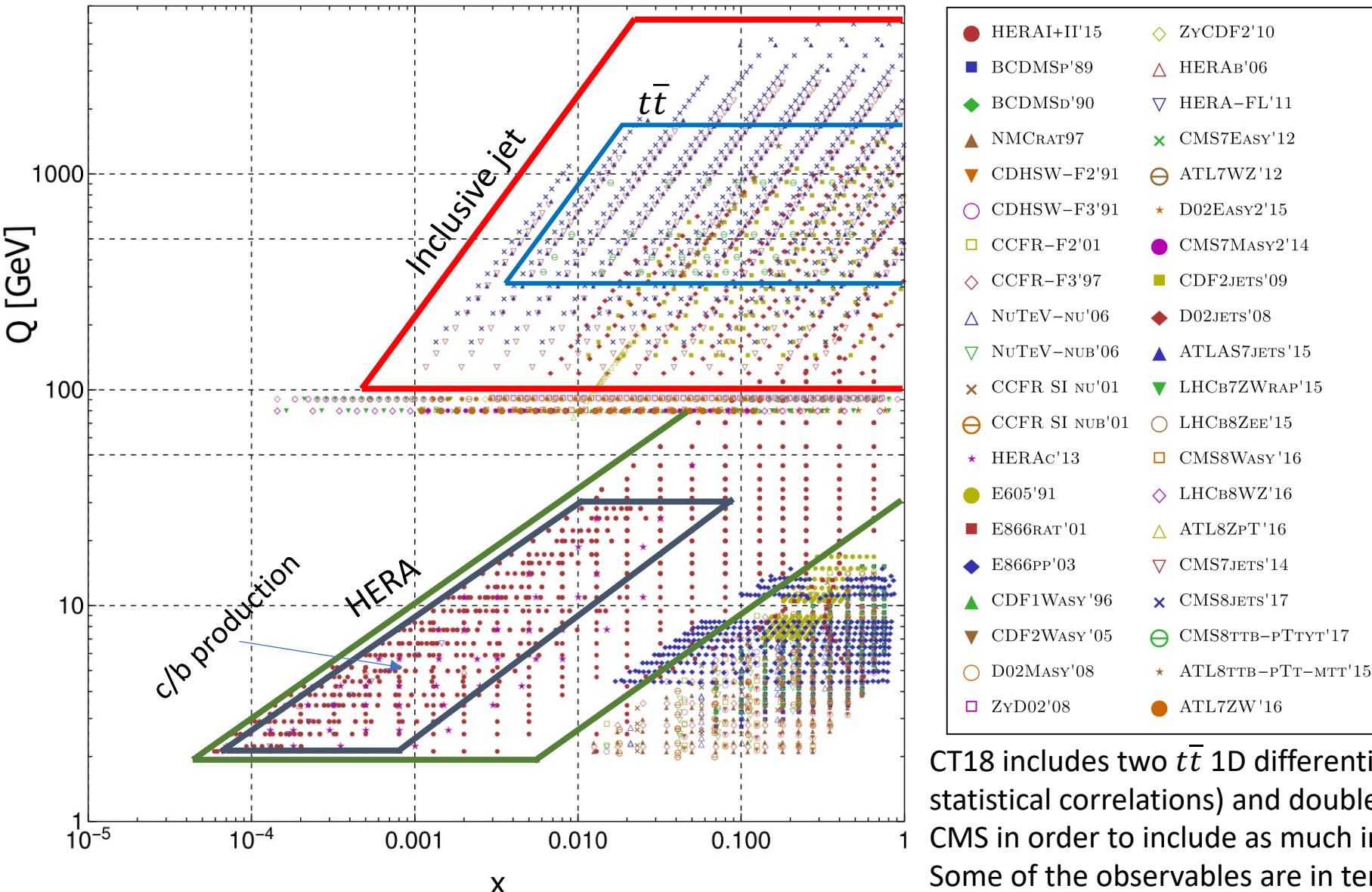
TABLE II. Like Table I, for newly included LHC measurements. The ATLAS 7 TeV  $W/Z$  data (4.6 fb<sup>-1</sup>), labeled by ‡, are included in the CT18A and CT18Z global fits, but not in CT18 and CT18X. The ATLAS 7 TeV  $W/Z$  data (4.6 fb<sup>-1</sup>), marked by ‡, is replaced by the updated one (4.6 fb<sup>-1</sup>) in the CT18A and CT18Z fits. The CDHSW data, labeled by †, are not included in the CT18Z fit. The numbers in parentheses are for the CT18Z NNLO fit.

Exp. ID#	Experimental dataset		$N_{pt,E}$	$\chi_E^2$	$\chi_E^2/N_{pt,E}$	$S_E$
245	LHCb 7 TeV 1.0 fb <sup>-1</sup> $W/Z$ forward rapidity cross sec.	[81]	33	53.8 (39.9)	1.6 (1.2)	2.2 (0.9)
246	LHCb 8 TeV 2.0 fb <sup>-1</sup> $Z \rightarrow e^-e^+$ forward rapidity cross sec.	[82]	17	17.7 (18.0)	1.0 (1.1)	0.2 (0.3)
248‡	ATLAS 7 TeV 4.6 fb <sup>-1</sup> , $W/Z$ combined cross sec.	[39]	34	287.3 (88.7)	8.4 (2.6)	13.7 (4.8)
249	CMS 8 TeV 18.8 fb <sup>-1</sup> muon charge asymmetry $A_{ch}$	[83]	11	11.4 (12.1)	1.0 (1.1)	0.2 (0.4)
250	LHCb 8 TeV 2.0 fb <sup>-1</sup> $W/Z$ cross sec.	[84]	34	73.7 (59.4)	2.1 (1.7)	3.7 (2.6)
253	ATLAS 8 TeV 20.3 fb <sup>-1</sup> , $Z$ $p_T$ cross sec.	[85]	27	30.2 (28.3)	1.1 (1.0)	0.5 (0.3)
542	CMS 7 TeV 5 fb <sup>-1</sup> , single incl. jet cross sec., $R = 0.7$ (extended in $y$ )	[86]	158	194.7 (188.6)	1.2 (1.2)	2.0 (1.7)
544	ATLAS 7 TeV 4.5 fb <sup>-1</sup> , single incl. jet cross sec., $R = 0.6$	[9]	140	202.7 (203.0)	1.4 (1.5)	3.3 (3.4)
545	CMS 8 TeV 19.7 fb <sup>-1</sup> , single incl. jet cross sec., $R = 0.7$ , (extended in $y$ )	[87]	185	210.3 (207.6)	1.1 (1.1)	1.3 (1.2)
573	CMS 8 TeV 19.7 fb <sup>-1</sup> , $t\bar{t}$ norm. double-diff. top $p_T$ and $y$ cross sec.	[88]	16	18.9 (19.1)	1.2 (1.2)	0.6 (0.6)
580	ATLAS 8 TeV 20.3 fb <sup>-1</sup> , $t\bar{t}$ $p_T^l$ and $m_{t\bar{t}}$ abs. spectrum	[89]	15	9.4 (10.7)	0.6 (0.7)	-1.1 (-0.8)

Heavy-flavor production measurements at HERA and LHC included in the CT18 NNLO QCD global analysis.

# PDF Kinematics: the $Q$ - $x$ plane

Experimental data in CT18 PDF analysis



Jet and  $t\bar{t}$  complement each other in the kinematic plane. They impact the gluon PDF at large  $x$ . Important to disentangle the effect due to jet production and top-quark data.

## Top and jet Data in CT18

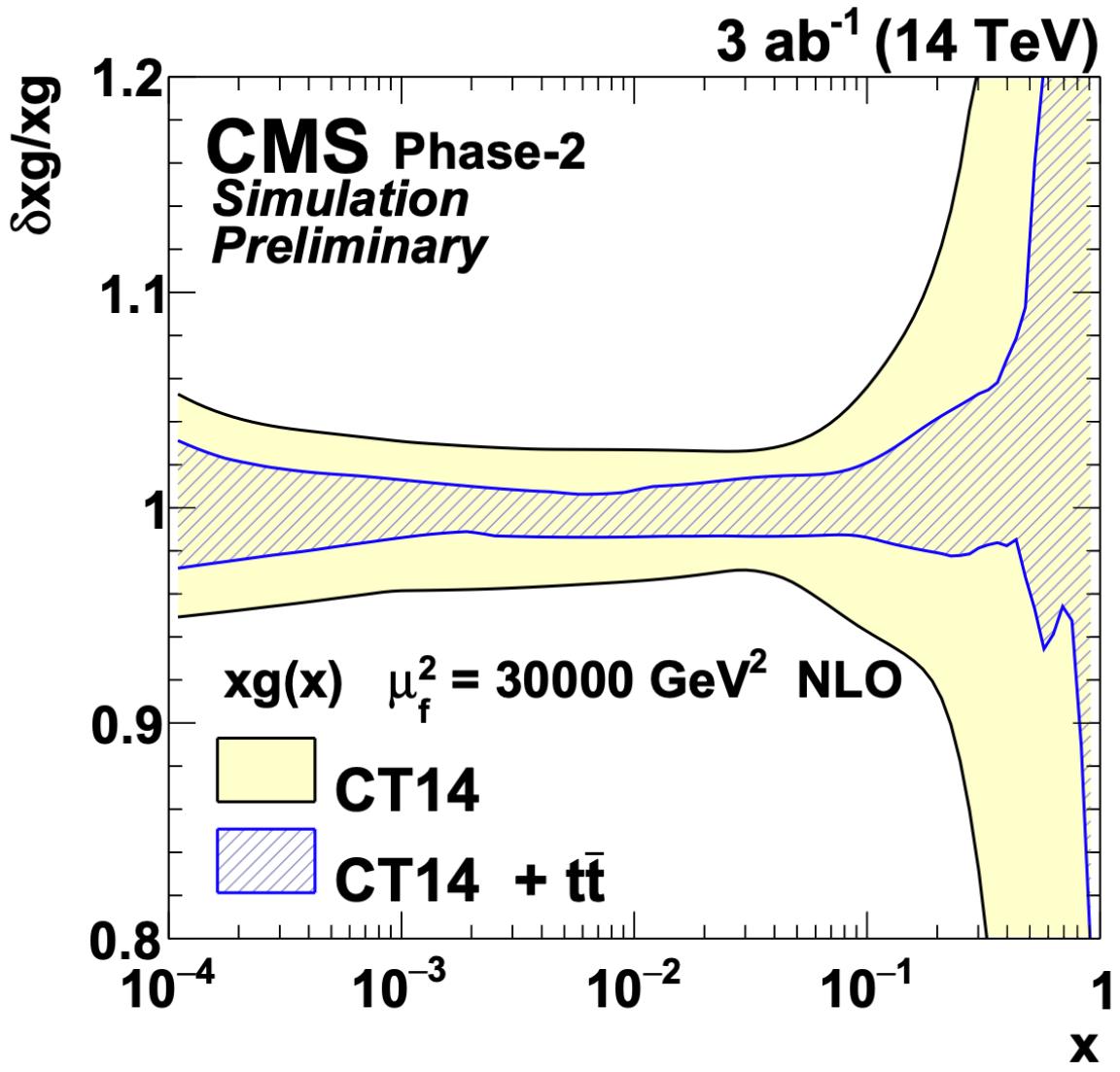
Top-quark

- 1511.04716 ATLAS 8 TeV tt<sub>b</sub> ptT diff. distributions
- 1511.04716 ATLAS 8 TeV tt<sub>b</sub> mtt diff. distributions
- 1703.01630 CMS 8 TeV tt<sub>b</sub> (pT, yt) 2d diff. distrib.

Jet production

- 1406.0324 CMS incl. jet at 7 TeV with R=0.7
- 1410.8857 ATLAS incl. jet at 7 TeV with R=0.6
- 1609.05331 CMS incl. jet at 8 TeV with R=0.7

# How constraining are $t\bar{t}$ data @LHC run III (and beyond)?



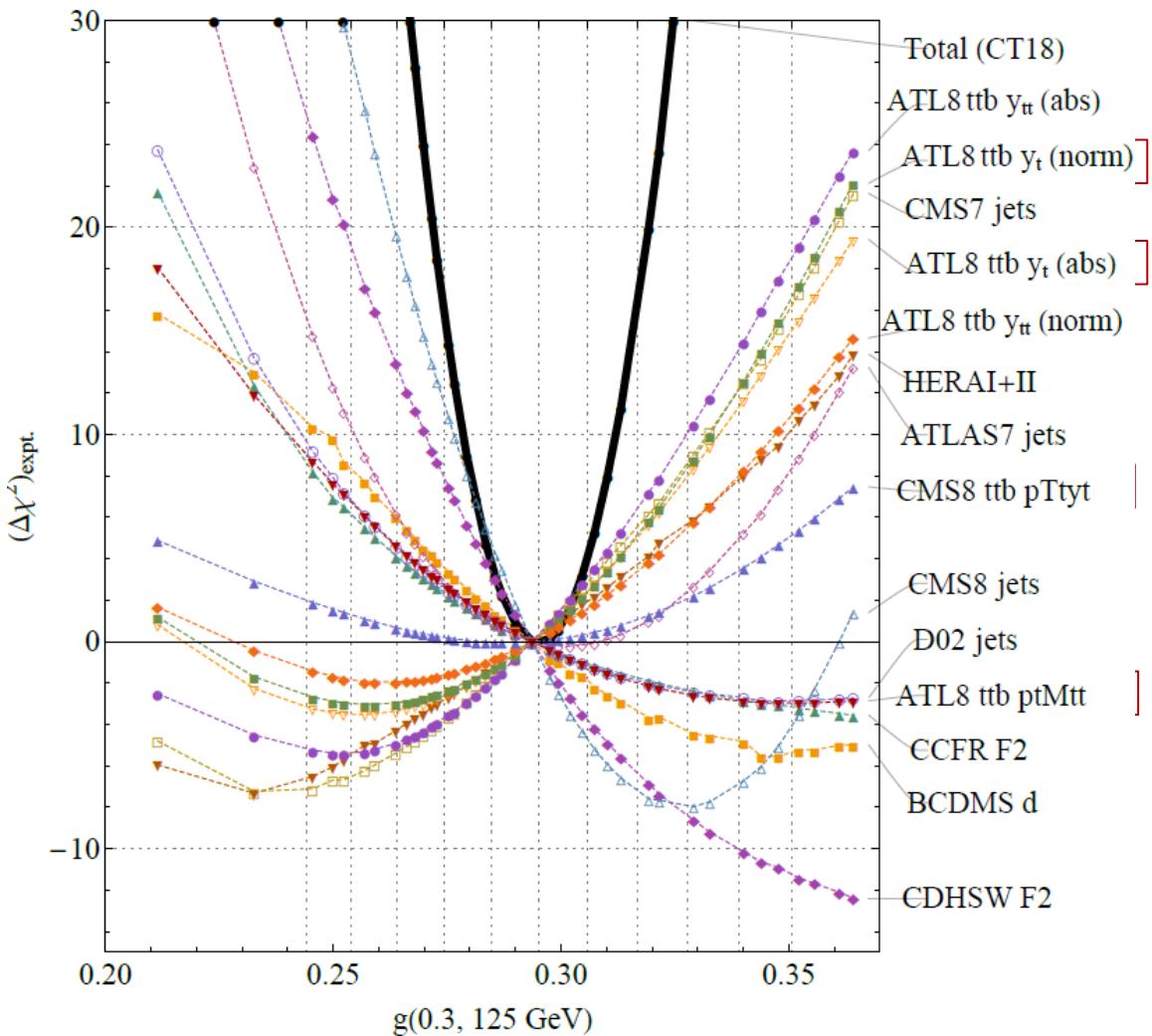
Projections with  $\Delta\chi^2 = 1$  tolerance predict strong constraints on the gluon and other flavors.

Such projections effectively emphasize a given measurement over the other experiments.

Figure: An estimated reduction of the relative uncertainty on the gluon PDF by profiling CT14 PDFs using simulated  $t\bar{t}$  measurements at the HL-LHC [CMS-PAS-FTR-18-015].

# How constraining are $t\bar{t}$ data: CT18 analysis

CT18 NNLO + unfitted ATLAS 8 TeV top single-diff. data



More realistic estimates account for multiple PDF functional forms and some disagreements between the measurements.

They predict milder impact from  $t\bar{t}$  data

In the figure, pulls on the gluon from ATLAS8  $y_{t\bar{t}}$  and  $y_t$  distributions (absolute or normalized) agree with HERA DIS, oppose ATLAS8  $d^2\sigma/(dp_{T,t} dm_{t\bar{t}})$  and CMS8  $d^2\sigma/(dp_{T,t} dy_{t,\text{ave}})$

# Impact of LHC 13 TeV $t\bar{t}$ production beyond CT18

Exp	Obs	Npt	ePump updated Chi2/Npt			HT/2	Global fit Chi2/Npt	HT/4
			HT	HT/2	HT/4			
ATLAS hadronic channel 36.1 fb <sup>-1</sup>	mtt	9	1.749	1.574	1.601		1.532026	1.4691
	HTtt	11	1.982	1.769	1.585		1.499361	1.74098
	ytt	12	1.279	1.15	0.938		1.051071	1.07351
	pTt1	10	1.301	1.185	1.118		1.196207	1.33326
	pTt2	8	1.132	0.843	1.047		0.84058	1.59056
CMS dilepton channel 35.9 fb <sup>-1</sup>	mtt	7	3.457	3.068	3.142		3.121005	3.22675
	ytt	10	1.66	0.969	0.679		0.938607	0.67252
	pTt	6	3.598	3.701	3.679		3.558017	3.04841
	yt	10	1.334	0.944	0.867		1.002635	0.68848
ATLAS lepton + jet 36 fb <sup>-1</sup>	mtt	7	2.395	1.165	0.681		0.826805	0.65684
	ytt	10	0.909	0.69	0.621		0.740418	0.74866
	pTt	6	2.337	2.012	2.469		1.353523	1.43062
	yt	10	1.298	1.073	1.095		1.161363	0.68198
CMS lepton + jet 137 fb <sup>-1</sup>	mtt	15	1.485	1.383	1.808		1.203901	1.66676
	ytt	10	6.469	6.238	6.424		6.005668	5.87508

ATLAS all hadronic, JHEP 01 (2021) 033, arXiv:2006.09274

ATLAS lepton + jets, EPJC 79 (2019) 1028, arXiv:1908.07305

CMS dilepton, JHEP 1902 (2019) 149, arXiv:1811.06625

CMS lepton + jets, PRD 104 092013 (2021), arXiv:2108.02803

Correlated Systematic Uncertainties:

ATLAS -> nuisance parameters

CMS -> Covariance matrix representation (converted to nuisance param. )

Statistical correlations not provided → data added one at a time

# Theory predictions: setup

- CMS: FastNNLO grids – ([Czakon et al. 1704.08551](#))
- ATLAS: bin-by-bin NNLO/NLO K-factors generated by MATRIX ([Catani, Grazzini, et al. PRD2019](#))

The NLO QCD calculation is obtained using our in-house APPLGrid fast tables ([Carli et al. EPJC 2010](#)) for the public MCFM calculation ([Campbell, Ellis JPG 2015](#))

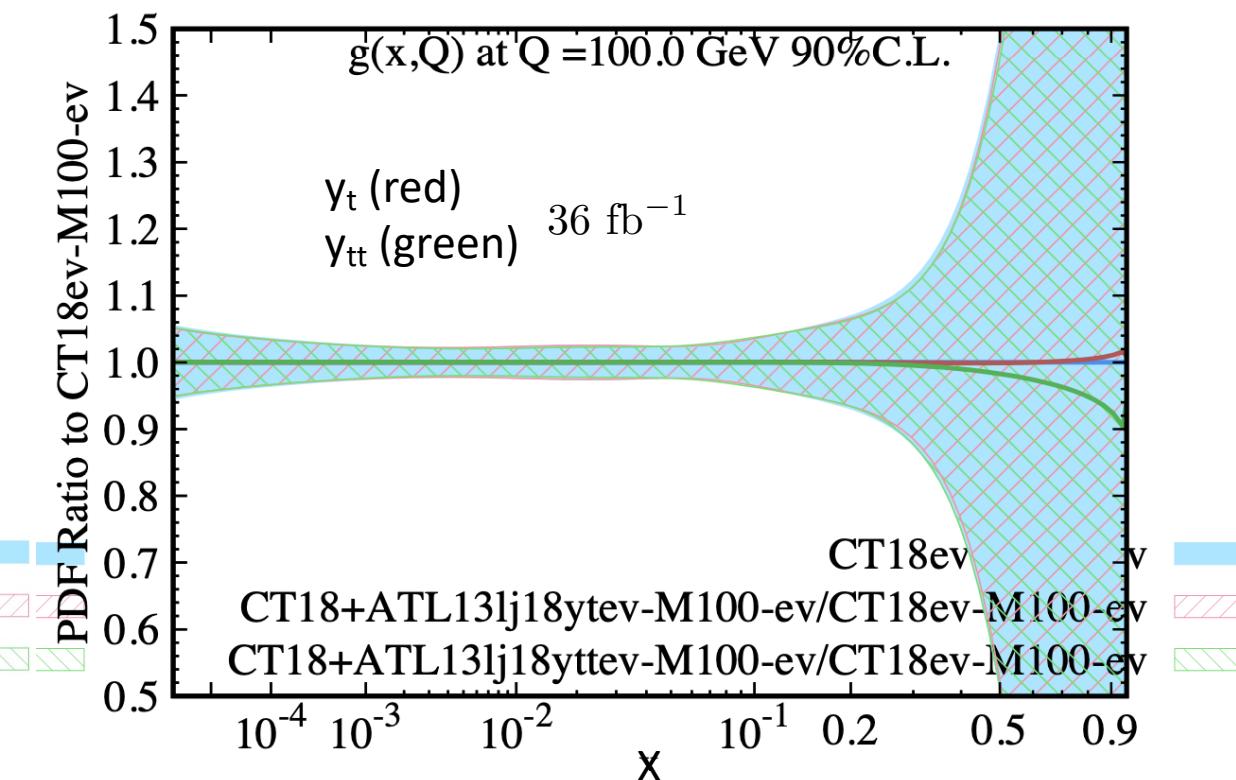
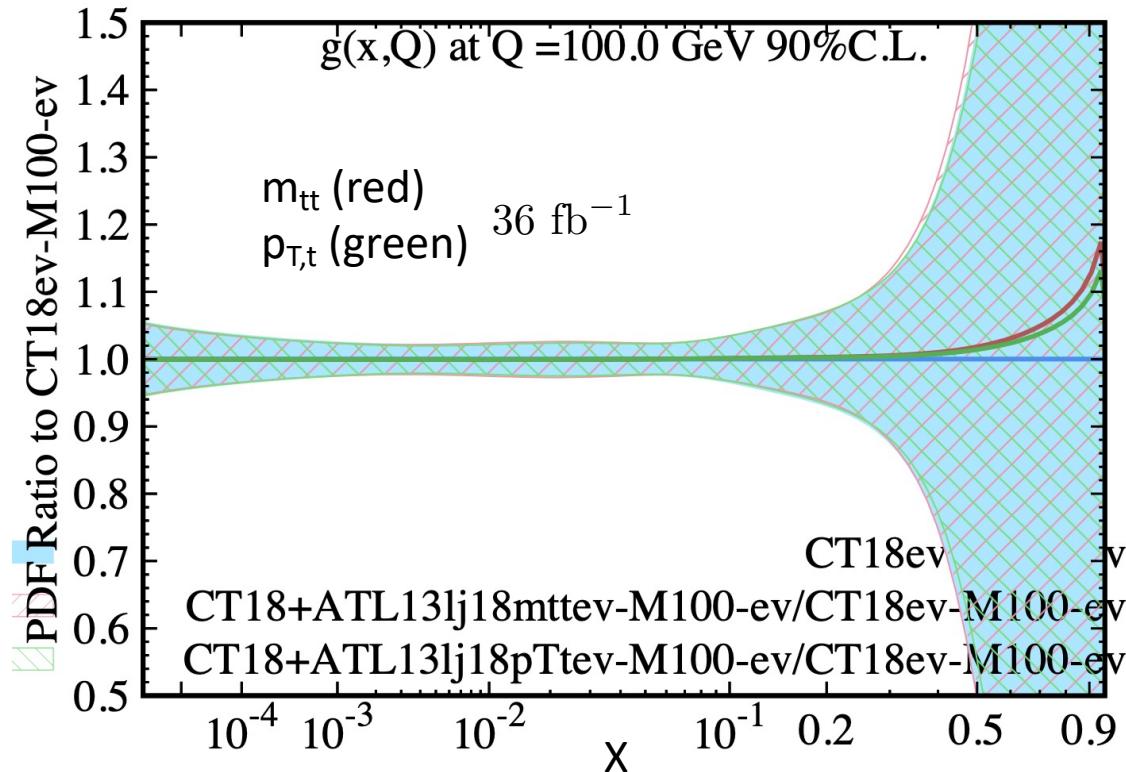
- $m_t(\text{pole}) = 172.5 \text{ GeV}$
- Fact/Ren scale choice:

$m_{tt}, p_{T,tt}, y_{tt}, y_t$  use:  $H_T/4$  (or  $H_T/2$ );  $p_{T,t}$ , use  $M_T$ ;  $p_{T,t} \text{ avg}$  use  $M_T/2$  ([Czakon et al. JHEP 2017](#))

$$\mu_F = \mu_R = H_T/4 = \left( \sqrt{m_t^2 + p_{T,t}^2} + \sqrt{m_t^2 + p_{T,\bar{t}}^2} \right) / 4 \quad \mu_{F,R} = M_T^t/2 = \sqrt{m_t^2 + p_T^2}/2$$

- **EW corrections considered:** negligible impact on our fits.

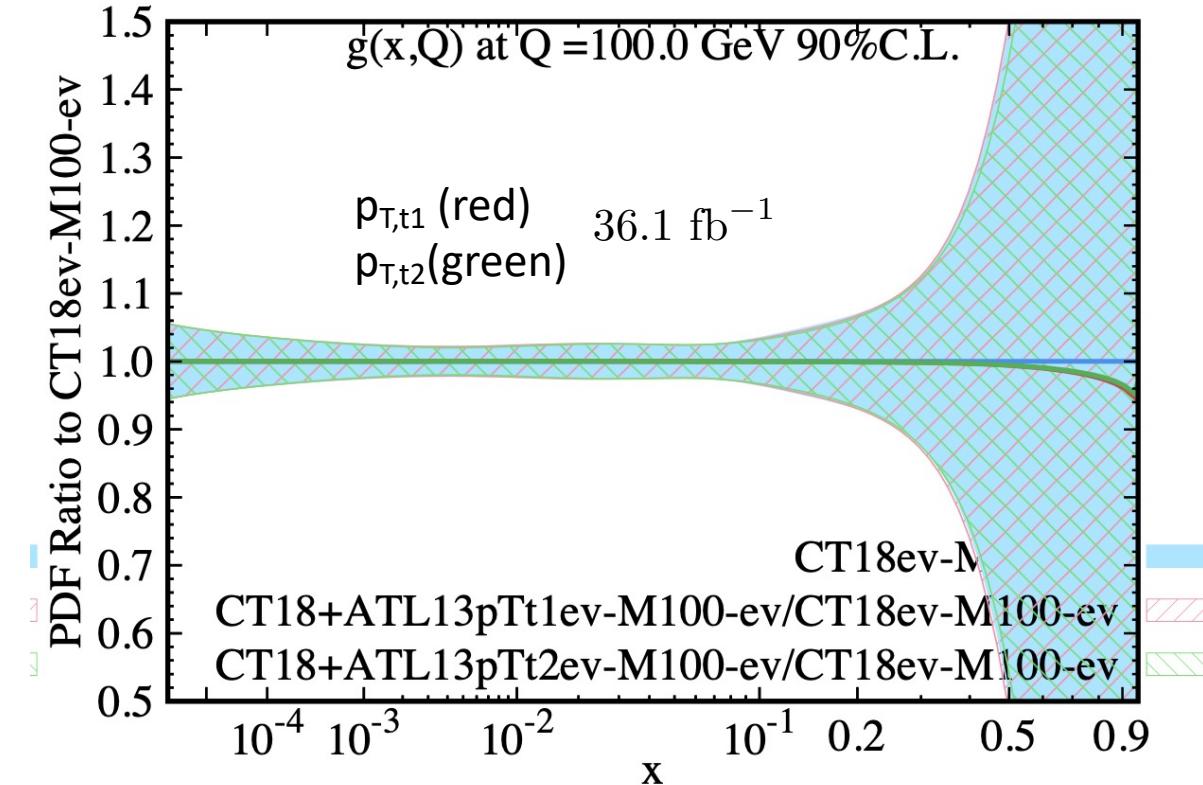
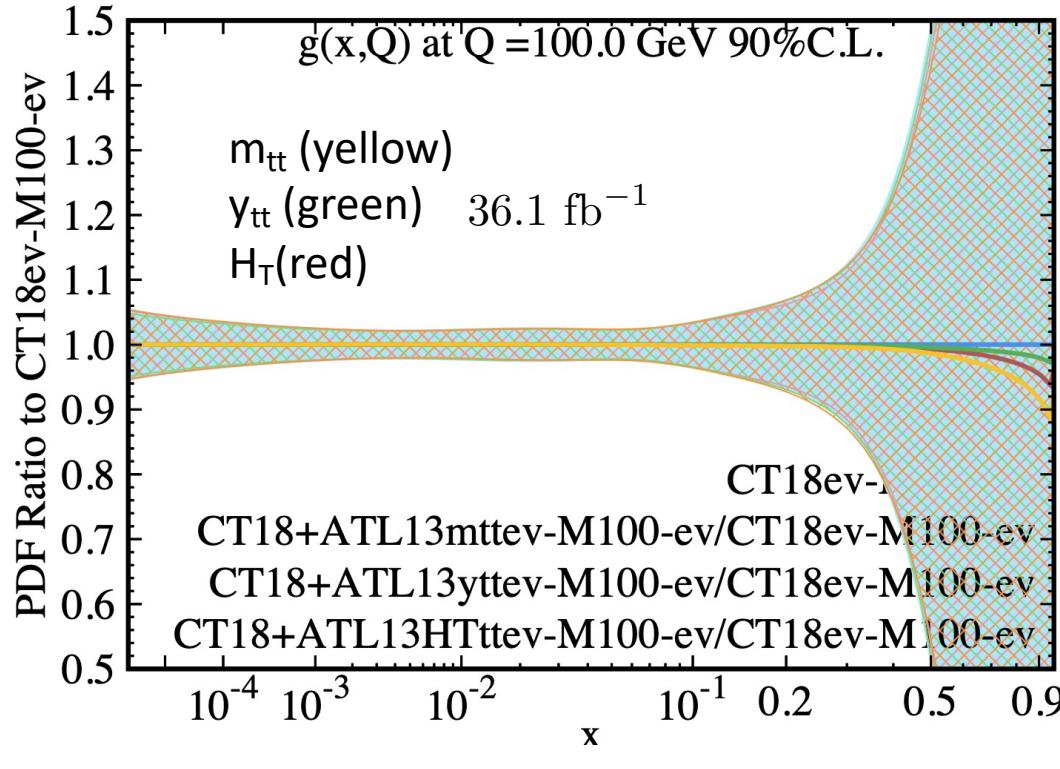
# Global fit: Impact on $g(x,Q)$ from ATLAS lep+jets



$m_{tt}$ ,  $p_{T,t}$ ,  $y_t$ , and  $y_{tt}$  1D absolute distributions added one by one in the global fit

Pulls are not in the same direction

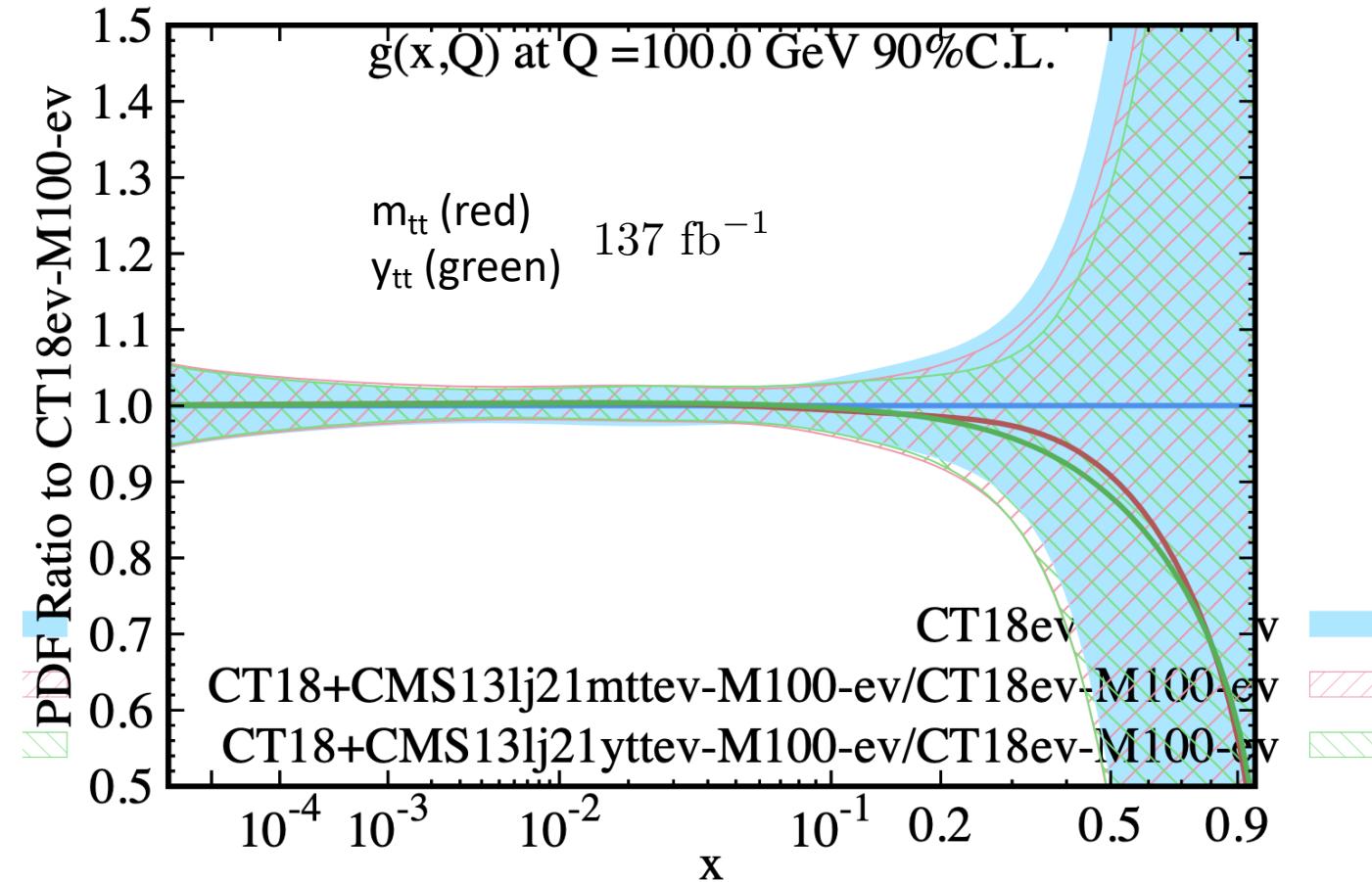
# Global fit: Impact on $g(x,Q)$ from ATLAS all hadronic



$m_{tt}$ ,  $p_{T,t1}$ ,  $p_{T,t2}$ ,  $H_T$ , and  $\gamma_{tt}$  1D absolute distributions added one by one in the global fit

Pulls are in the same direction here

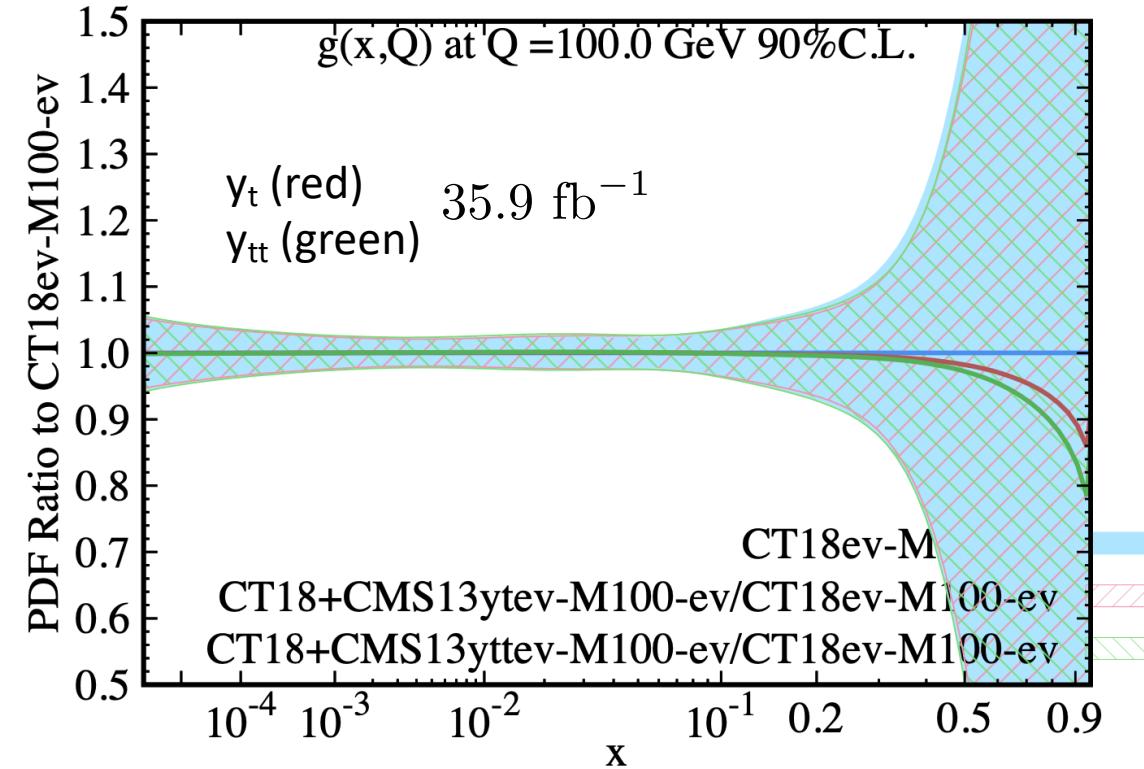
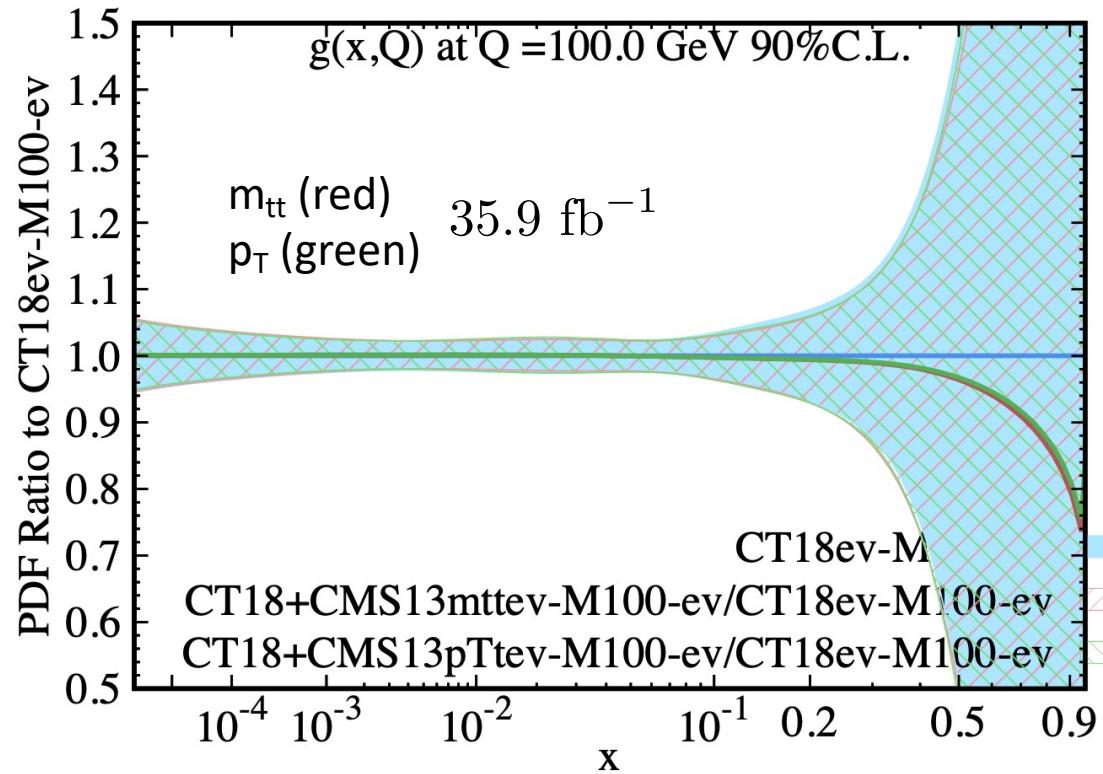
# Global fit: Impact on $g(x,Q)$ from CMS lep+jets



$m_{tt}$ , and  $y_{tt}$  1D absolute distributions added one by one in the global fit

Pulls are in the same direction here: stronger impact due to higher precision of these data

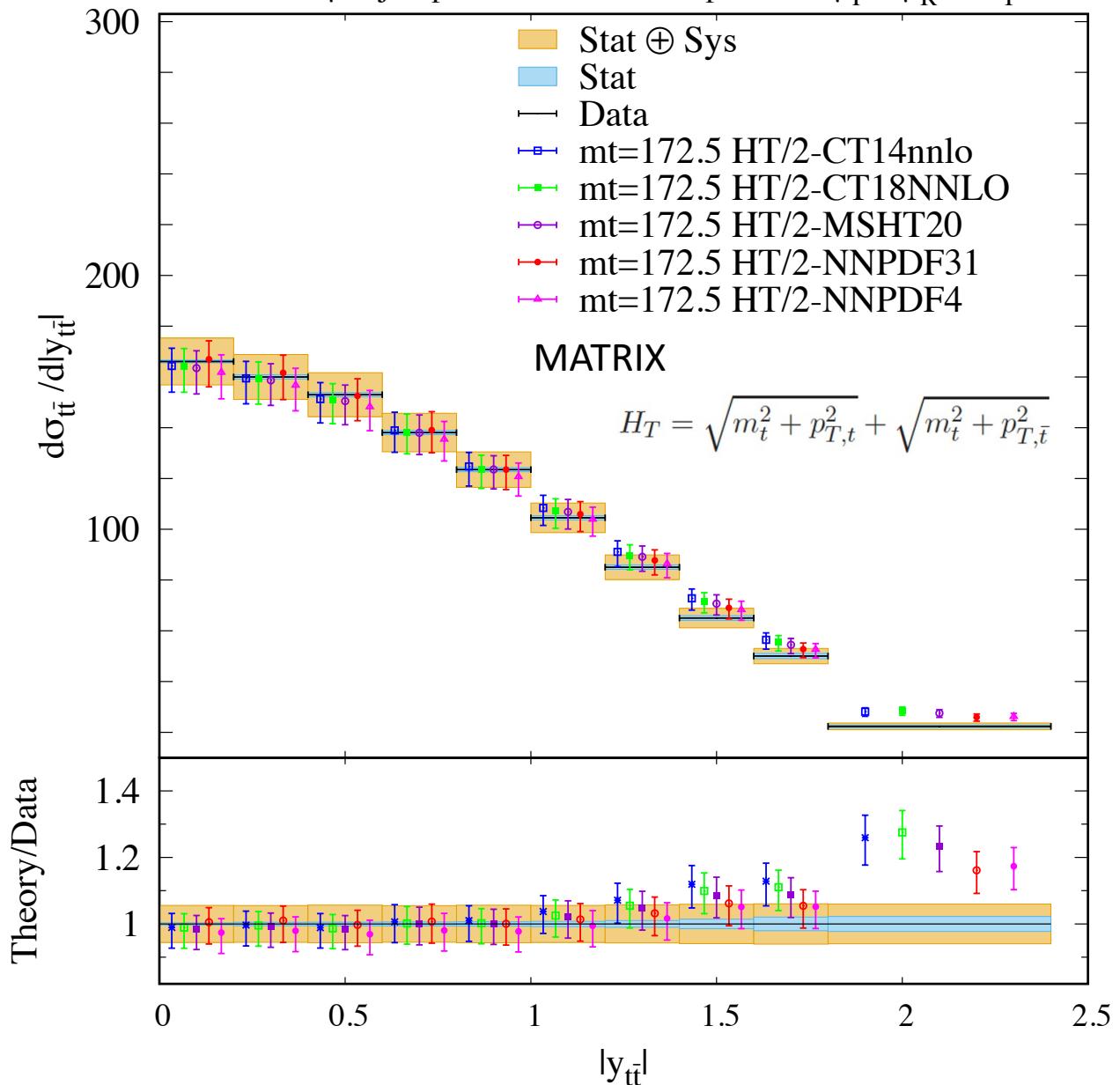
# Global fit: Impact on $g(x,Q)$ from CMS 2-lep



$m_{tt}$  ,  $p_T$  ,  $y_{tt}$  , and  $y_t$  1D absolute distributions added one by one in the global fit

Pulls are in the same direction here

CMS13 e/ $\mu$  + jets parton level. Scale dependence  $\mu_F = \mu_R = H_T / 2$



SCALE DEPENDENCE: HT/2 varied up and down by a factor of 2

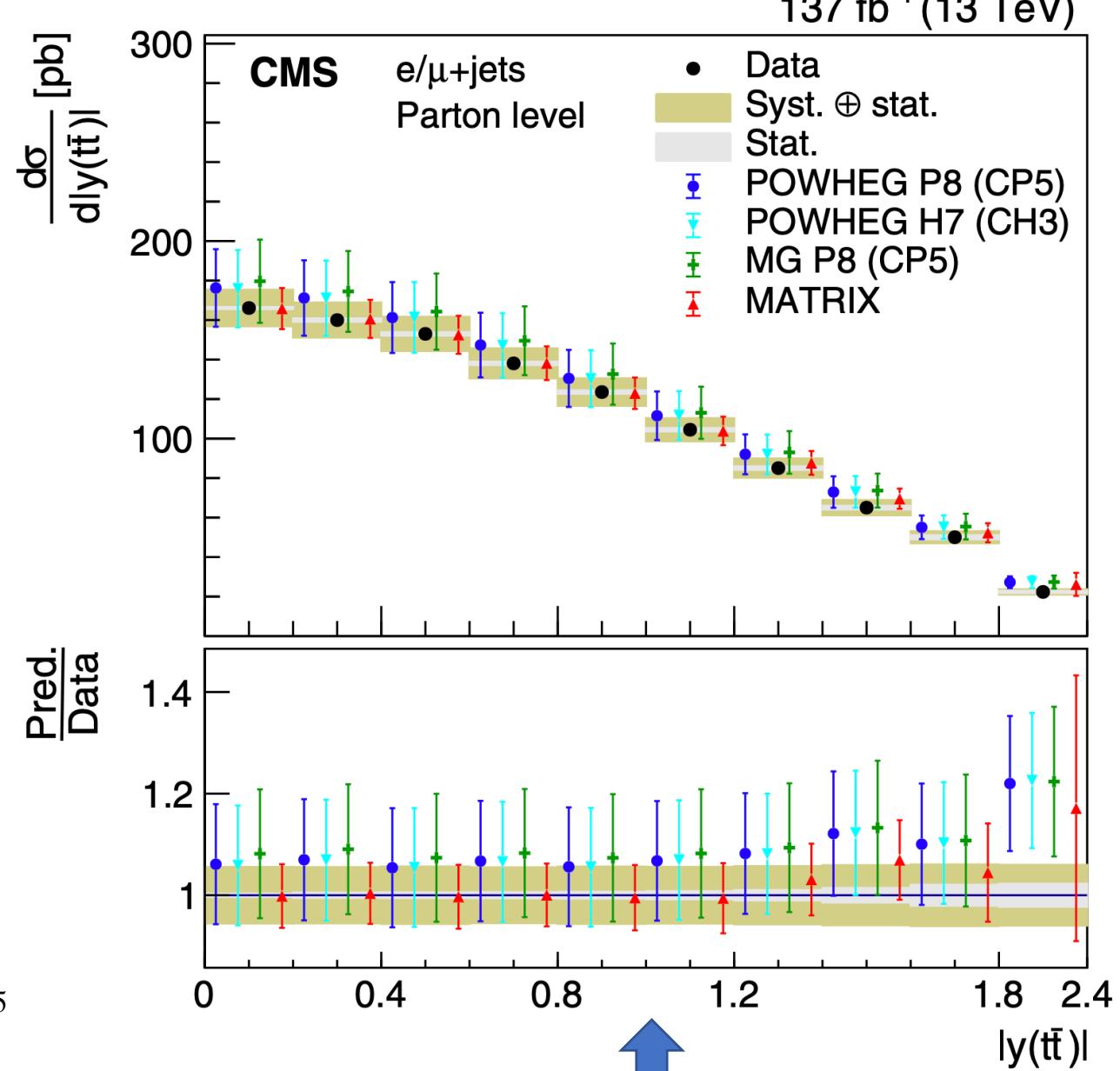
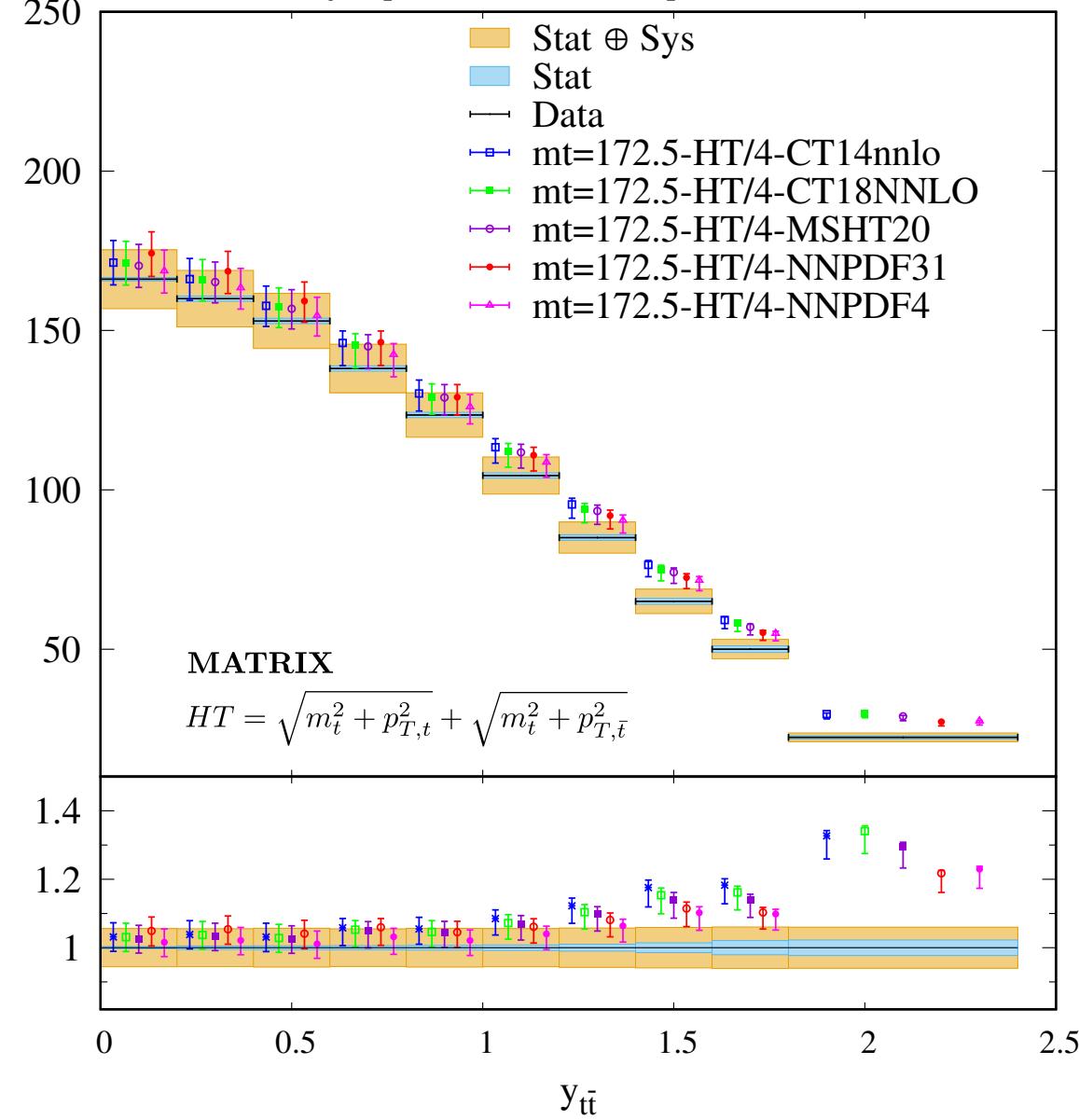


Figure from CMS publication, arXiv:2108.02803  
Error: PDF + scale + ?

CMS 13 e/ $\mu$ +jets parton level. Scale dependence  $\mu_R = \mu_F = HT/4$



SCALE DEPENDENCE: HT/4 varied up and down by a factor of 2

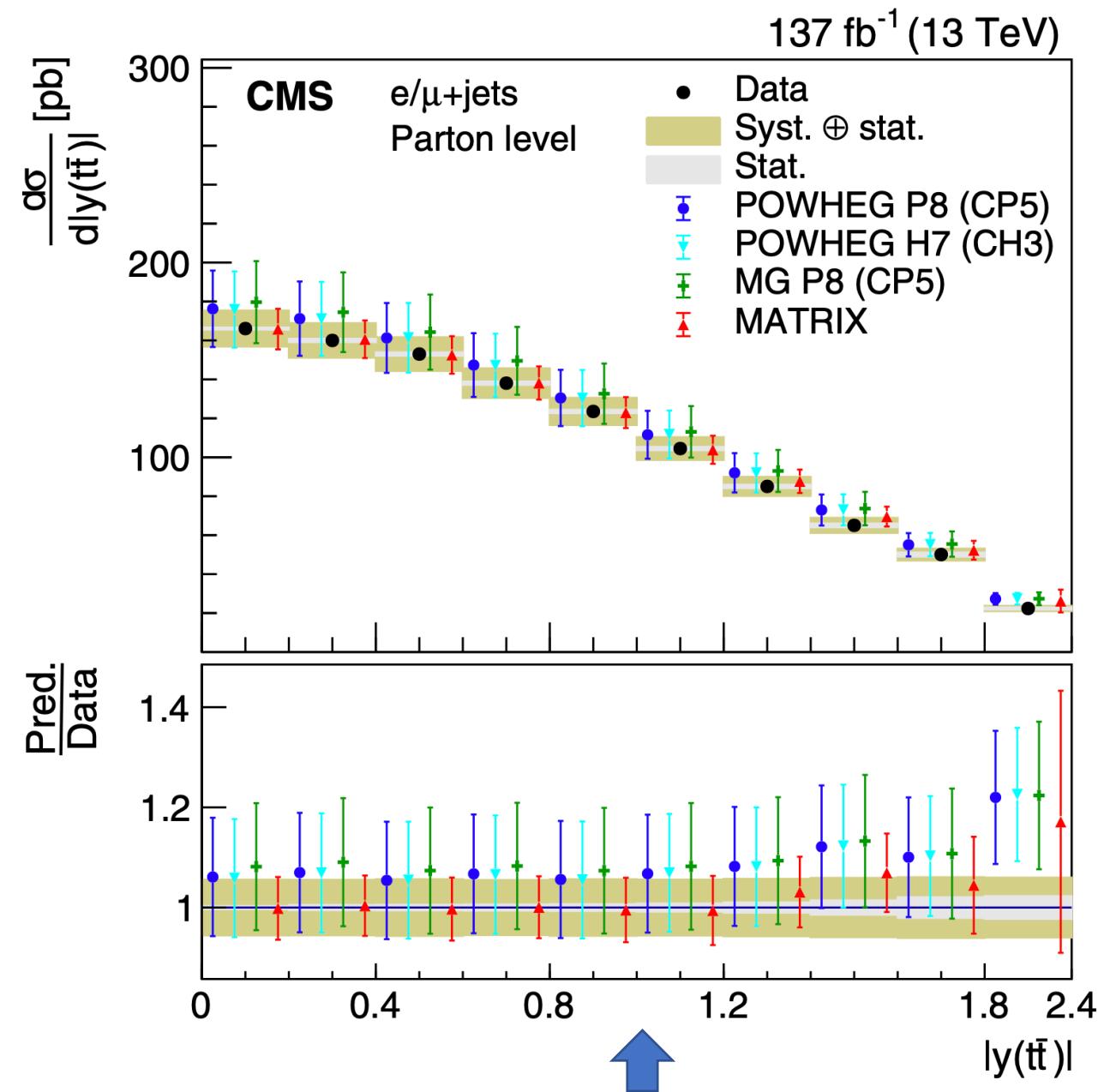


Figure from CMS publication, arXiv:2108.02803  
Error: PDF + scale + ?

# Messages from the ttbar 13 TeV analysis

- First comprehensive exploration of LHC 13 TeV ttbar impact on PDFs
- Most of the impact: high-precision data from CMS
- Pulls on the gluon not in the same direction for some distr. in ATLAS
- Scale uncertainty: the recommended scale choice is not always the best
- Interplay between jets and ttbar (we have partial results, work is in progress)
- Optimal baseline: ATLAS hadronic: ytt absolute, CMS dilepton: ytt absolute, ATLAS lep+jet: ytt, CMS lep+jet: mtt absolute.
- ttbar 13 TeV data seem to prefer a softer gluon at large x.
- Interplay with the most recent DIS charm and bottom production measurements from HERA

# Impact from new combined c- and b-quark production at HERA

(H1 and ZEUS Coll. 1804.01019)

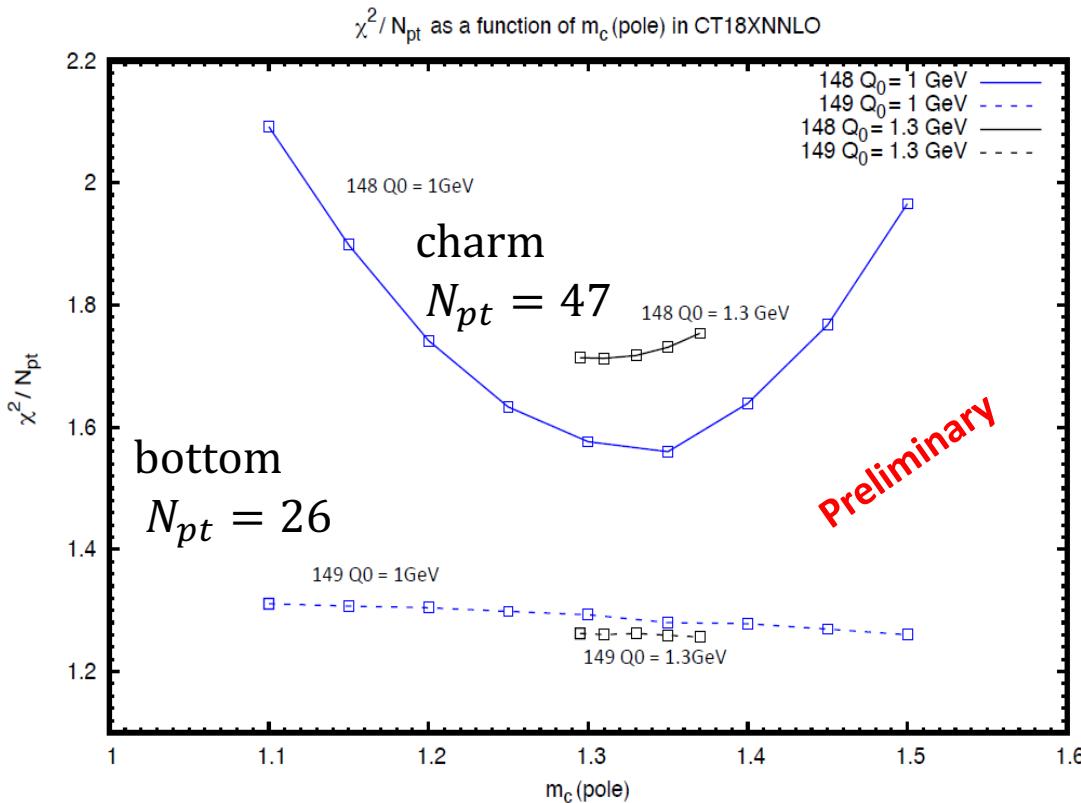


We fit these data using the SACOT- $\chi$  heavy-quark scheme at NNLO.

In all tried scenarios, we get  $\chi^2/N_{pt}$  no less than 1.5, reached when the combined HERA HQ SIDIS data is included (even with large statistical weight )

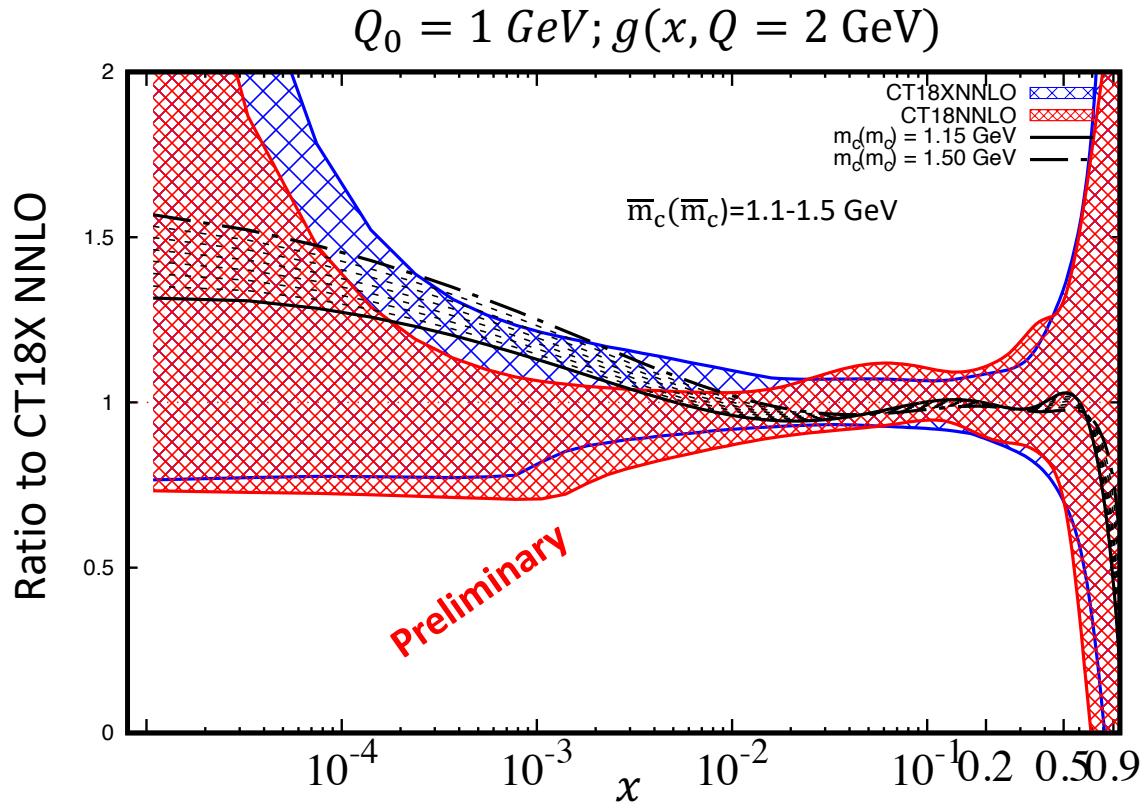
These data prefer a harder gluon at intermediate and small  $x$ .

Our  $\chi^2$  values are similar to those found by MSHT20 and to predictions from other groups reported in Table 4 of the HERA publication.

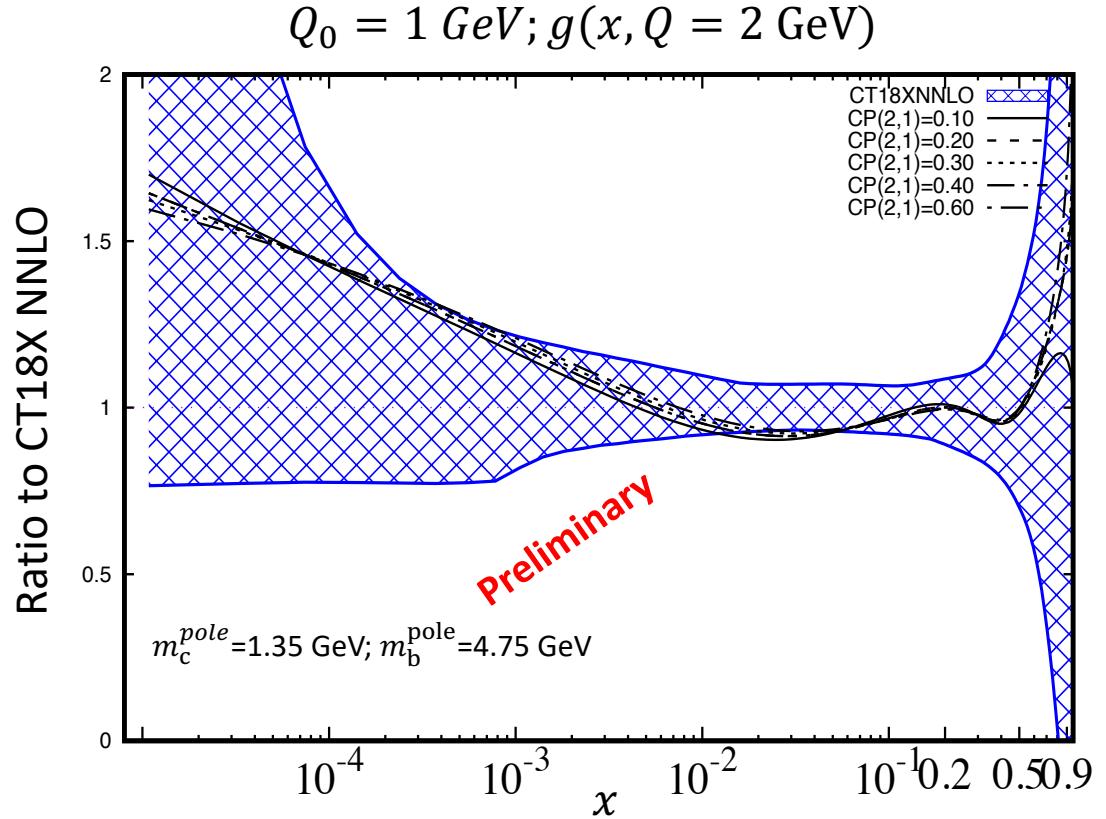


# CT18XNNLO + combined HERA $c/b$ DIS data set

Fits with varied  $\overline{m}_c(\overline{m}_c)$



Fits with varied small-x scale



This data set mildly prefers CT18XNNLO to CT18NNLO.

But  $\chi^2/N_{pt}$  is never lower than 1.5 for all explored combinations

$$\mu_{DIS}(x) = A \sqrt{m_Q^2 + B^2/x^C}$$

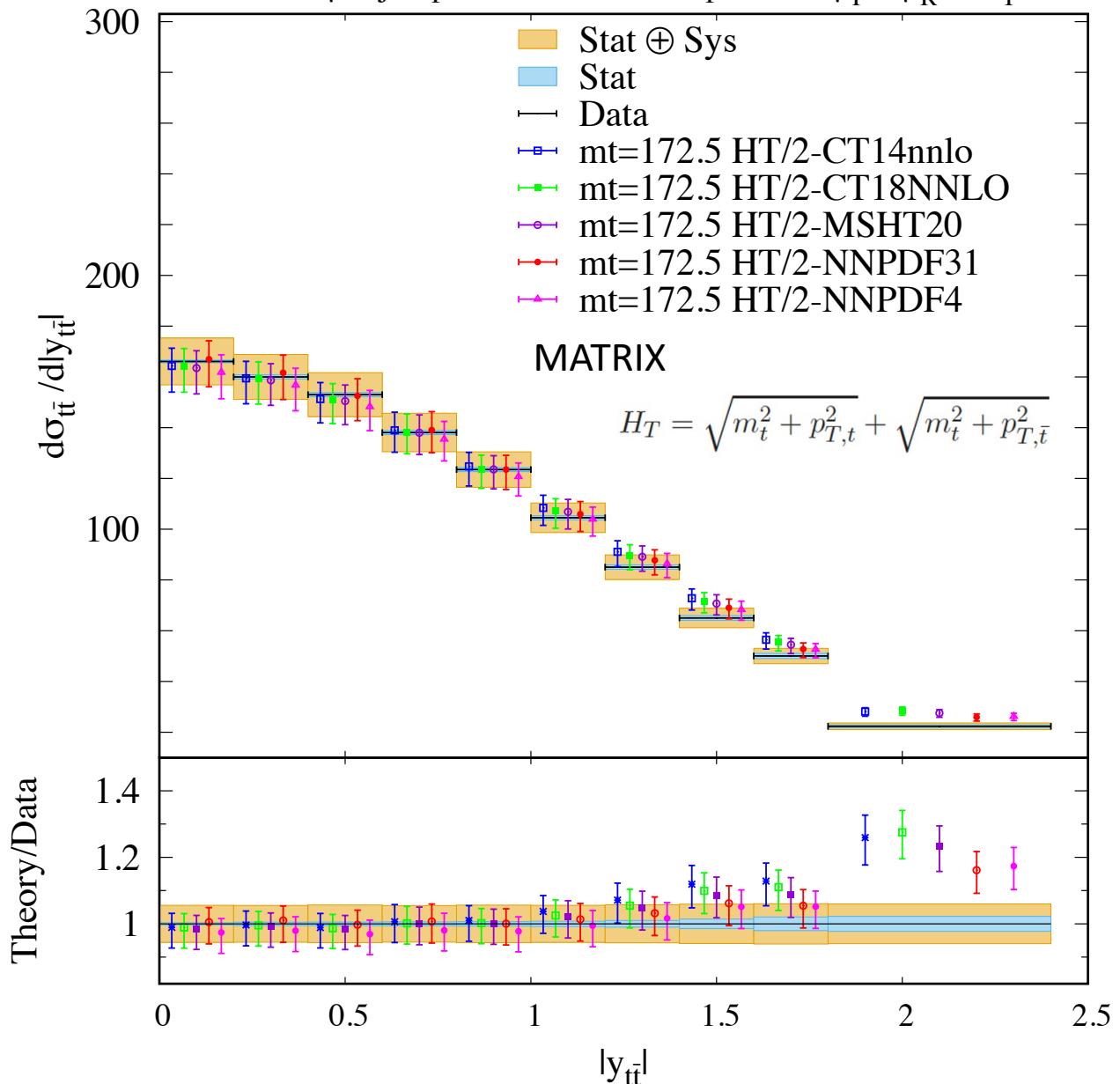
Vary  $B=CP(2,1)$ , while keeping  $A=0.5$  and  $C=0.33$  fixed

# Points of discussion

- We explored the impact of 13 TeV  $t\bar{t}$  LHC 1D measurements on CT18 PDFs
- MATRIX theory @NNLO implemented via K-factors with NLO Applgrids from MCFM:
  - We gained flexibility for the binning, but K-factors may introduce extra uncertainty.
- ttbar 13 TeV fastNNLO grids available but not flexible (mt, bins, scale, ...).
- We studied 1D distributions in full phase space.
- Impact from data in fiducial PS? Again, we need even more flexible fast tables.
- How CPU/time consuming is to generate flexible fast tables for theory predictions?

# BACK UP

CMS13 e/ $\mu$  + jets parton level. Scale dependence  $\mu_F = \mu_R = H_T / 2$



SCALE DEPENDENCE: HT/2 varied up and down by a factor of 2

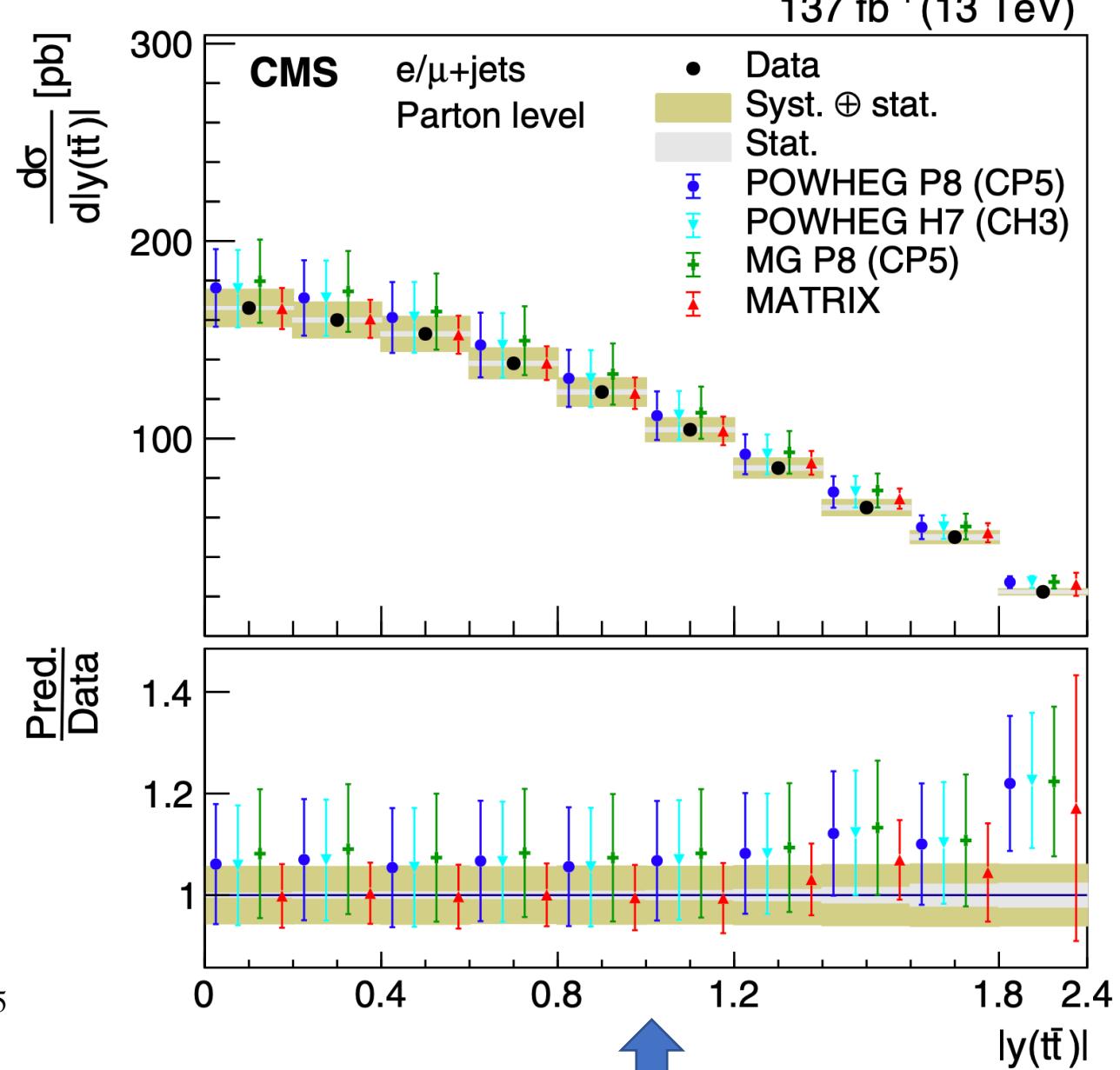
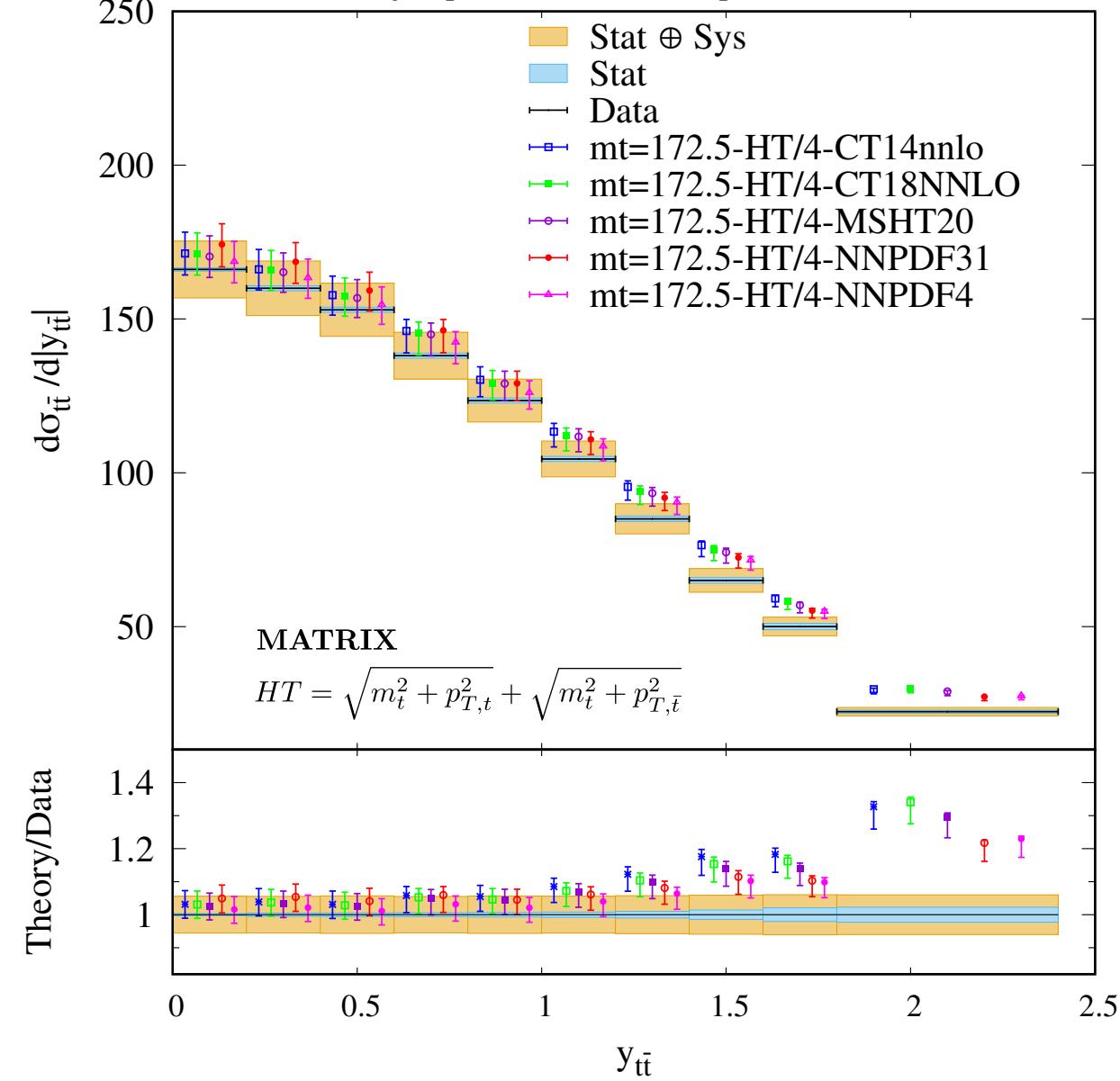


Figure from CMS publication, arXiv:2108.02803  
Error: PDF + scale + ?

CMS 13 e/ $\mu$ +jets parton level. Scale dependence  $\mu_R = \mu_F = HT/4$



SCALE DEPENDENCE: HT/4 varied up and down by a factor of 2

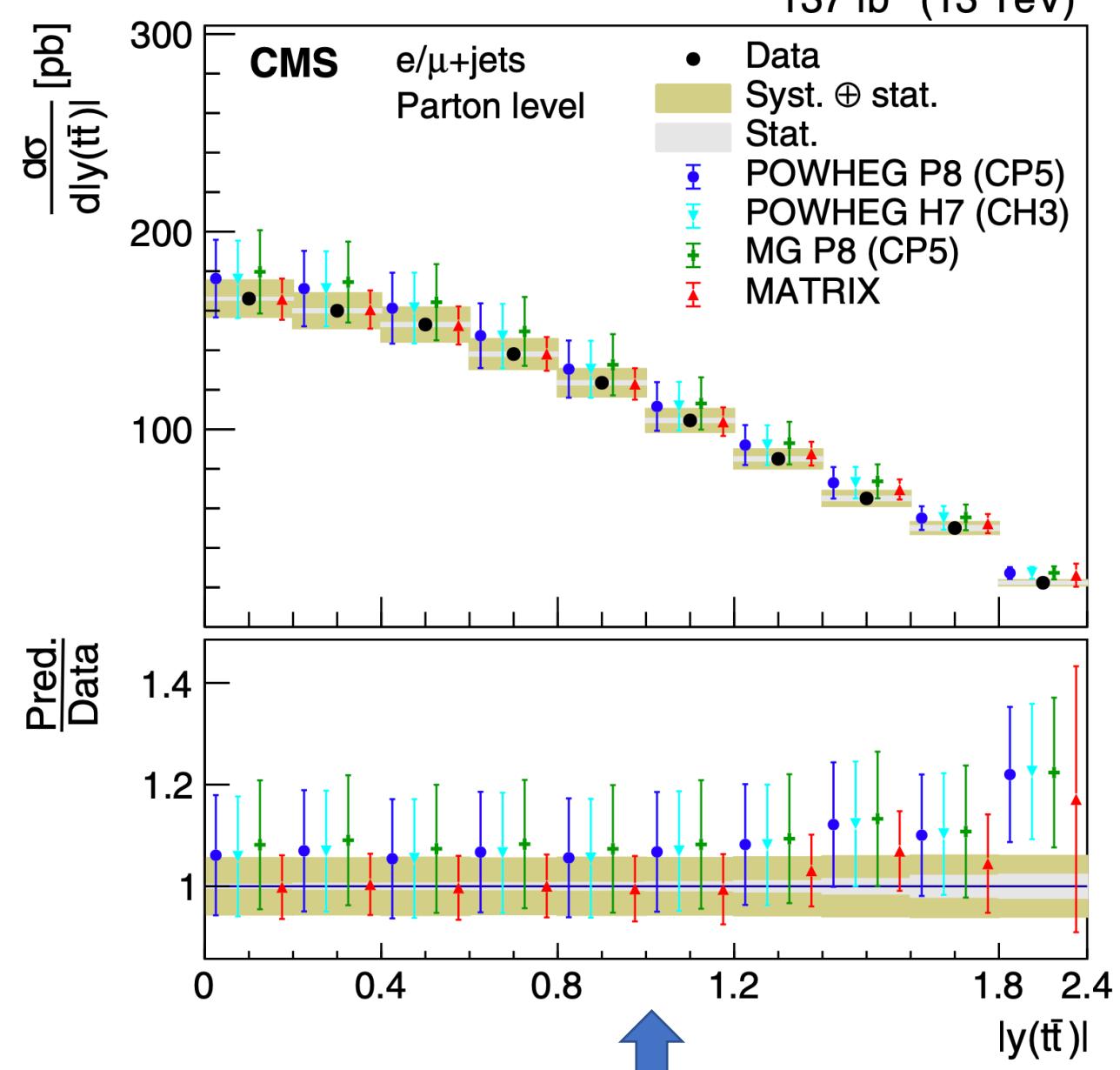


Figure from CMS publication, arXiv:2108.02803  
Error: PDF + scale + ?

