LHC differential top-quark pair production cross sections in the ABMP16 PDF fit

Sergey Alekhin[1], Maria Vittoria Garzelli[1], Sven-Olaf Moch[1], Sasha Zenaiev[1]
with input from Javier Mazzitelli[2]

[1] Hamburg University
[2] PSI

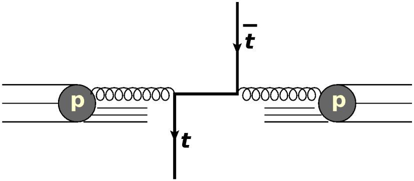
arxiv:2311.05509

arxiv:24XX.XXXXX in preparation

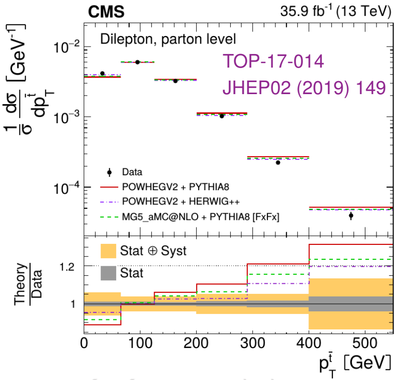
DIS2024 (Grenoble, France)
9 Apr 2024

Introduction

Why study $t\bar{t}$ production?



- m_t provides a hard scale
  $\Rightarrow$ **ultimate probe of pQCD**
  (NLO, aNNLO, NNLO, ...)
- Produced mainly via gg
  $\Rightarrow$ constrain gluon PDF at high x
- Production sensitive to α_s and m_t
- May provide insight into possible new physics

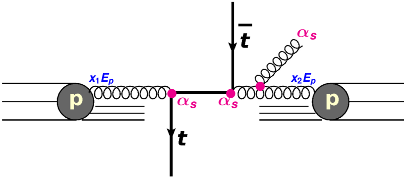


Why study 2D/3D?

- 1D measurements: overall good agreement, but reveal some trends
- 2D [EPJ C77 (2017) 459, PRD97 (2018) 112003]: study production dynamics in more detail
- 3D [EPJ C80 (2020) 658]: simultaneously constrain α_s (extra jets), m_t^{pole}, PDFs

Introduction

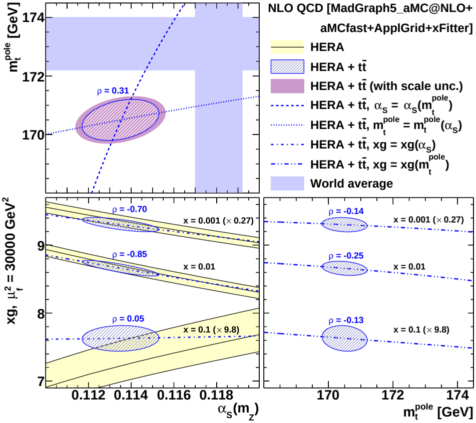
Why study $t\bar{t}$ production?



- m_t provides a hard scale
  $\Rightarrow$ ultimate probe of pQCD
  (NLO, aNNLO, NNLO, ...)
- Produced mainly via *gg*
  $\Rightarrow$ constrain gluon PDF at high x
- Production sensitive to α_s and m_t^{pole}
- May provide insight into possible new physics

Example:

CMS TOP-18-004 [EPJ C80 (2020) 7]

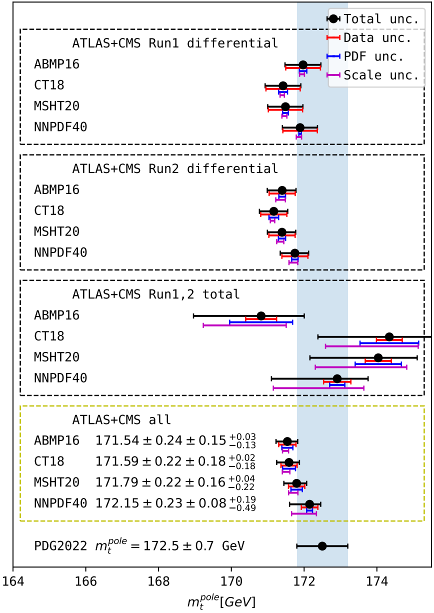


- Simultaneous extraction of PDFs, α_s, m_t^{pole} using normalised triple-differential cross sections at NLO
- Extended to $\overline{\text{MS}}$, MSR schemes in JHEP 04 (2021) 043 [Garzelli, Kemmler, Moch, Zenaiev]

Scope of this work

- NNLO calculations for total and fully differential $t\bar{t}$ (q_T subtraction) are now publicly available with MATRIX framework [Catani, Devoto, Grazzini, Kallweit, Mazzitelli Phys.Rev.D 99 (2019) 5, 051501; JHEP 07 (2019) 100]
  - fully differential NNLO calculations were also published in JHEP 04 (2017) 071 [Czakon, Heymes, Mitov], but no public code available
- Multi-differential $t\bar{t}$ measurements enable extraction of PDFs, α_s, m_t^{pole}
  - for 3D x-sections $M(t\bar{t})$, $y(t\bar{t})$, N_{jet} [CMS TOP-18-004] NNLO calculations are not available for $t\bar{t}$ + jets
  - therefore for the time being we focus on double-differential $M(t\bar{t})$, $y(t\bar{t})$ x-sections
    - ⋆ $M(t\bar{t})$ provides sensitivity to m_t
    - ⋆ $y(t\bar{t})$ provides sensitivity to PDFs via relation to partonic momentum fraction x: at LO $x_{1,2} = (M(t\bar{t})/\sqrt{s})\exp[\pm y(t\bar{t})]$
- Recently we have extracted m_t^{pole} from the total and differential ATLAS and CMS measurements of $t\bar{t}$ production using different PDF sets [Garzelli, Mazzitelli, Moch, Zenaiev arXiv:2311.05509]
  - overall uncertainty on m_t^{pole} is 0.3 GeV (experimental, PDF, scale variations)
  - possible extra uncertainty (not accounted for): renormalon ambiguity, m_t in MC used to unfold the data to parton level
- **Now we do a global PDF fit with total and differential $t\bar{t}$ within the ABMP16 framework**
  - using collider and fixed-target DIS, DY, updated single top and $t\bar{t}$ data (see BACKUP)

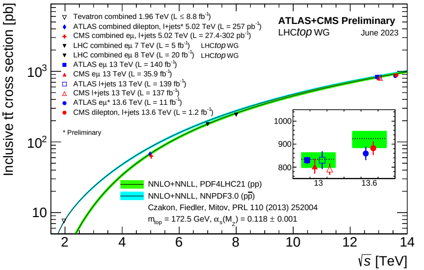
Previous work: extraction of m_t^{pole} [Garzelli, Mazzitelli, Moch, SZ arXiv:2311.05509]



Theoretical calculations with MATRIX framework

- Using “private” (custom) version of MATRIX provided by Javier Mazzitelli
- Interfaced to PineAPPL [Carrazza at al., JHEP 12 (2020) 108] to produce PDF interpolation grids which are further used in xFitter `https://gitlab.com/fitters/xfitter`
  - reproduce NNLO calculations using any PDF set and/or varied μ_r, μ_f in $\sim$ seconds
  - interface implemented privately and only for $t\bar{t}$ production
  - **no NNLO/NLO K-factors etc.**
- Further modifications to MATRIX to make possible runs with $\Delta\sigma_{t\bar{t}} < 1‰$
  - skip calculation of identical things (tailored for $t\bar{t}$)
  - adapted to DESY Bird Condor cluster and local multicore machines
  - technical fixes related to memory and disk space usage etc.
- We did 6 runs with $m_t^{\text{pole}} \in$ [165,177.5] GeV with step of 2.5 GeV and $\Delta\sigma_{t\bar{t}} = 0.2‰$
  - $\approx$ 350K $\times$ 6 $\approx$ 2M CPU hours (200 years on single CPU, 3 months of real time)
  - for differential distributions, integration uncertainties in bins are $\lesssim$ 0.5%
- Differential distributions obtained with fixed $r_{cut} = 0.0015$ (q_T subtraction)
  - checked that extrapolation to $r_{cut} = 0$ for total cross section produces difference $< 1\%$ (see also S. Catani et al., JHEP 07 (2019) 100)
- $\mu_r = \mu_f = H_T/4$, $H_T = \sqrt{m_t^2 + p_T^2(t)} + \sqrt{m_{\bar{t}}^2 + p_T^2(\bar{t})}$, varied up and down by factor 2 with $0.5 \le \mu_r/\mu_f \le 2$ (7-point method)
- For PDF and top quark mass fit, m_t^{pole} is converted to $m_t(m_t)$, as customary in ABMP16

Data used in this analysis



Selection of data:

- all measurements of total $\sigma(t\bar{t})$:
  - 10 data points, including recently combined CMS+ATLAS cross section at 7 and 8 TeV
- differential measurements $\frac{1}{\sigma(t\bar{t})}\frac{d\sigma(t\bar{t})}{dO}$ which satisfy following criteria:
  - as function of $M(t\bar{t})$ (if available, 2D $M(t\bar{t})$ and $y(t\bar{t})$)
  - unfolded to parton level (no cuts on p_T, y of leptons or jets): no LHCb data
  - normalized cross sections (to avoid unknown correlation with total $\sigma(t\bar{t})$ and to reduce unknown correlations between different data sets)
  - bin-by-bin correlations are available
  - for the moment only Run-2 2D data included in the PDF fit (besides the total $t\bar{t}$ x-section data)

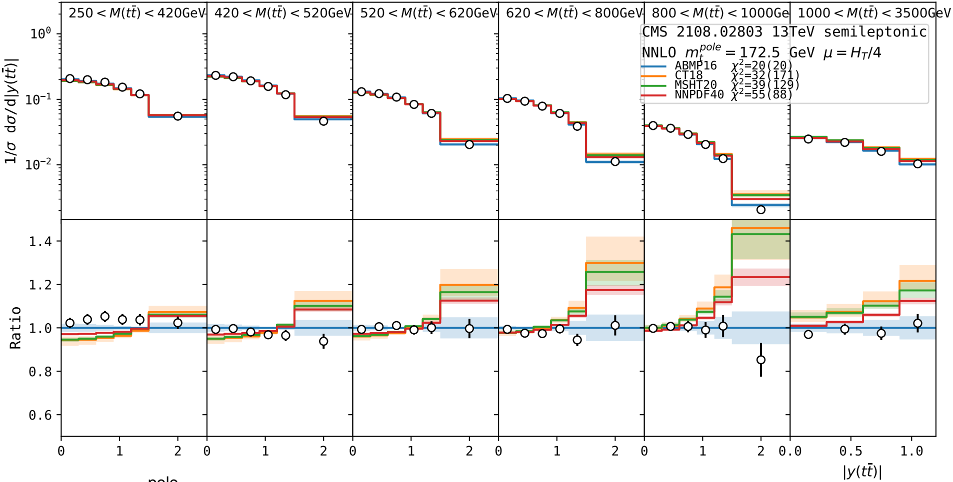
$$\sigma(t\bar{t})$$

Experiment	decay channel	dataset	luminosity	$\sqrt{s}$
ATLAS & CMS	combined	2011	5 fb^{-1}	7 TeV
ATLAS & CMS	combined	2012	20 fb^{-1}	8 TeV
ATLAS	dileptonic, semileptonic	2011	257 pb^{-1}	5.02 TeV
CMS	dileptonic	2011	302 pb^{-1}	5.02 TeV
ATLAS	dileptonic	2015-2018	140 fb^{-1}	13 TeV
ATLAS	semileptonic	2015-2018	139 fb^{-1}	13 TeV
CMS	dileptonic	2016	35.9 fb^{-1}	13 TeV
CMS	semileptonic	2016-2018	137 fb^{-1}	13 TeV
ATLAS	dileptonic	2022	11.3 fb^{-1}	13.6 TeV
CMS	dileptonic, semileptonic	2022	1.21 fb^{-1}	13.6 TeV

$$\frac{1}{\sigma(t\bar{t})}\frac{d\sigma(t\bar{t})}{dO}$$

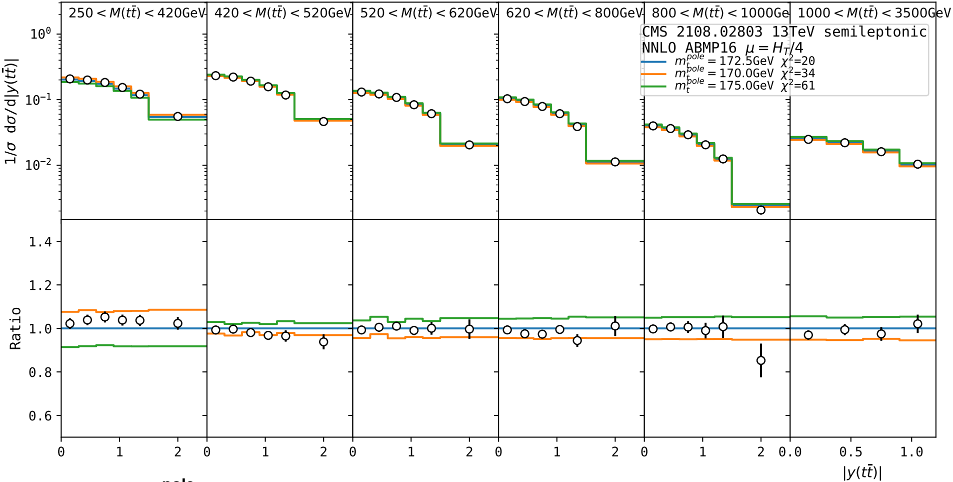
Experiment	decay channel	dataset	luminosity	$\sqrt{s}$	observable(s)	n
CMS	semileptonic	2016–2018	137 fb^{-1}	13 TeV	$M(t\bar{t})$, $\lvert y(t\bar{t})\rvert$	34
CMS	dileptonic	2016	35.9 fb^{-1}	13 TeV	$M(t\bar{t})$, $\lvert y(t\bar{t})\rvert$	15
ATLAS	semileptonic	2015–2016	36 fb^{-1}	13 TeV	$M(t\bar{t})$, $\lvert y(t\bar{t})\rvert$	19
ATLAS	all-hadronic	2015–2016	36.1 fb^{-1}	13 TeV	$M(t\bar{t})$, $\lvert y(t\bar{t})\rvert$	10
CMS	dileptonic	2012	19.7 fb^{-1}	8 TeV	$M(t\bar{t})$, $\lvert y(t\bar{t})\rvert$	15
ATLAS	semileptonic	2012	20.3 fb^{-1}	8 TeV	$M(t\bar{t})$	6
ATLAS	dileptonic	2012	20.2 fb^{-1}	8 TeV	$M(t\bar{t})$	5
ATLAS	dileptonic	2011	4.6 fb^{-1}	7 TeV	$M(t\bar{t})$	4
ATLAS	semileptonic	2011	4.6 fb^{-1}	7 TeV	$M(t\bar{t})$	4

CMS 2108.02803 vs NNLO predictions using different PDFs



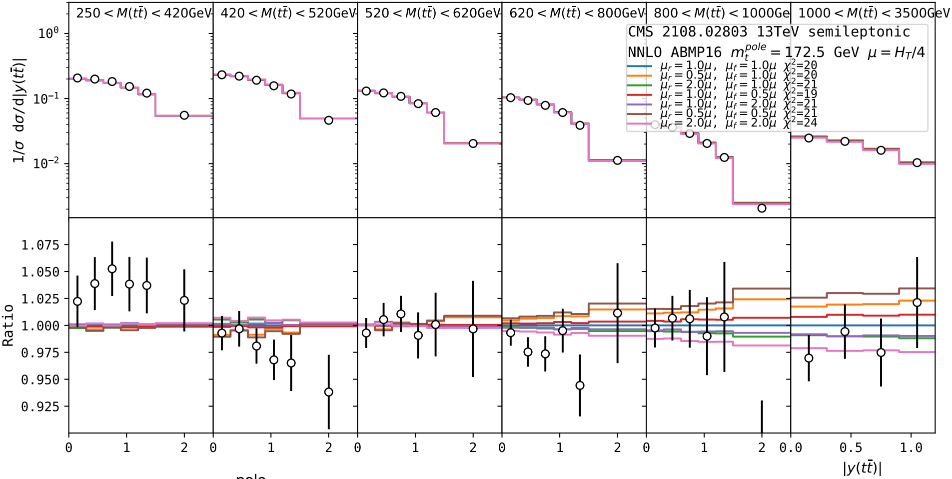
- Fixed $m_t^{\text{pole}} = 172.5$ GeV (not always consistent with PDF sets), $\mu_r = \mu_f = H_T/4$
- Reported χ^2 values with (and without) PDF uncertainties
- All PDF sets describe data reasonably well, with best description by ABMP16
  - CT18, MSHT20 and NNPDF40 show clear trend w.r.t data at high $y(t\bar{t})$ (large x)

CMS 2108.02803 vs NNLO predictions using different PDFs



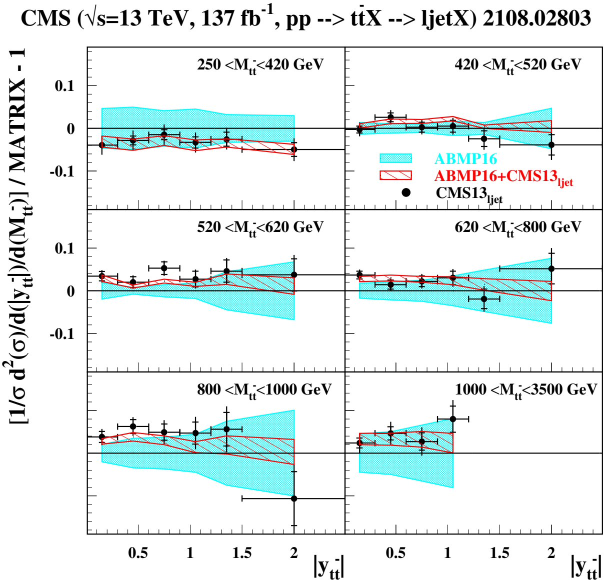
- **NOTE: m_t^{pole} values on this plot are not the same as the ones obtained in ABMP16 fit**
- Low $M(t\bar{t})$: strong dependence on m_t^{pole} via threshold effects
- High $M(t\bar{t})$: opposite dependence due to cross section normalization

CMS 2108.02803 vs NNLO predictions using different PDFs



- ABMP16, fixed $m_t^{\text{pole}} = 172.5$ GeV
- Scale variations $< 1\%$ at low $M(t\bar{t})$ (largest cancellation), reach $\approx 4\%$ at high $M(t\bar{t})$
  - → **these data are useful to provide constraints on m_t and PDFs**

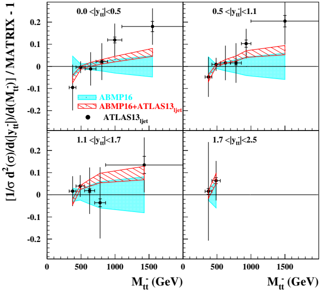
CMS 2108.02803 in ABMP16 PDF fit



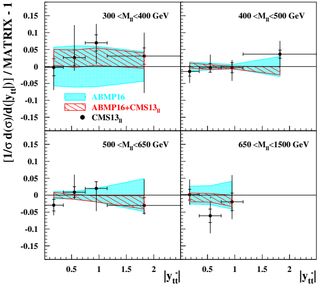
- Good description of the data
- **Significantly constrained PDF uncertainty band**

Other $t\bar{t}$ differential data in ABMP16 PDF fit

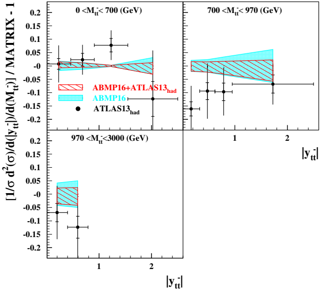
ATLAS (√s=13 TeV, 36 fb⁻¹, pp --> tt̄X --> ljetX) 1908.07305



CMS (√s=13 TeV, 36 fb⁻¹, pp --> tt̄X --> l⁺l⁻X) 1904.05237

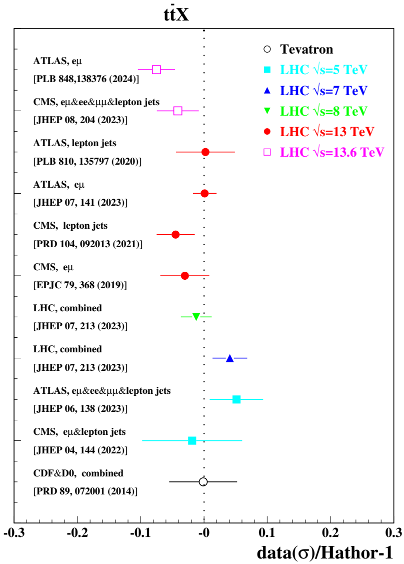


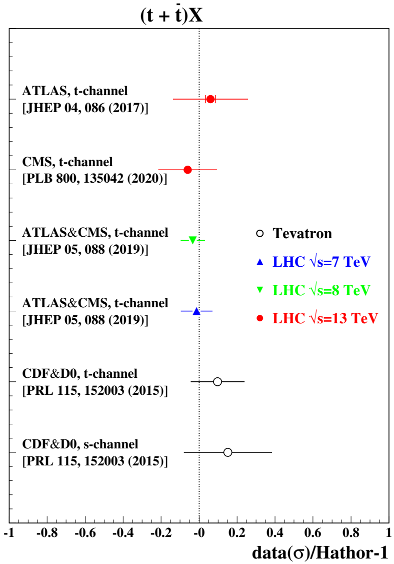
ATLAS (√s=13 TeV, 36 fb⁻¹, pp --> tt̄X --> hadronsX) 2006.09274



- ATLAS ljet data tend to be above theory predictions at $M(t\bar{t}) \gtrsim 1000$ GeV (but still ok within large data uncertainties)
- ATLAS hadronic data tend to be below theory predictions at $M(t\bar{t}) \gtrsim 700$ GeV (but still ok within large data uncertainties)
- **In summary, all data are in good agreement with NNLO theoretical predictions and put significant constraints on the PDFs**

$\sigma(t\bar{t})$ nd single t in ABMP16 PDF fit





χ^2

Experiment	Data set	$\sqrt{s}$ (TeV)	Reference	NDP	χ^2		
					I	II	III
ATLAS	$ATLAS\,13_{ljet}$	13	[25]	19	34.2	27.2	–
	$ATLAS\,13_{had}$	13	[26]	10	11.8	12.1	–
CMS	$CMS\,13_{ll}$	13	[24]	15	21.1	–	19.9
	$CMS\,13_{ljet}$	13	[22]	34	42.2	–	40.8

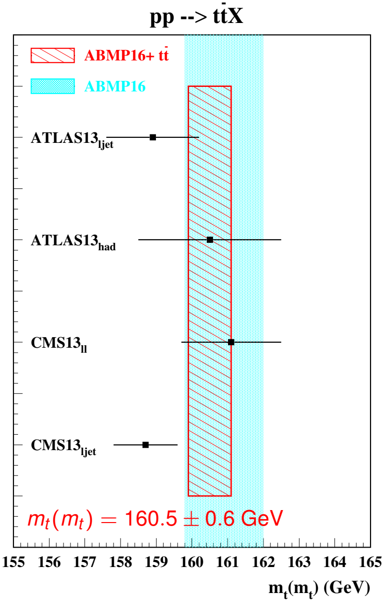
I: both ATLAS and CMS

II: only ATLAS

III: only CMS

→ **Overall good description of data by NNLO theoretical predictions, but some tension between ATLAS and CMS differential $t\bar{t}$ data is noticeable**

$m_t(m_t)$ and α_s in ABMP16 PDF fit



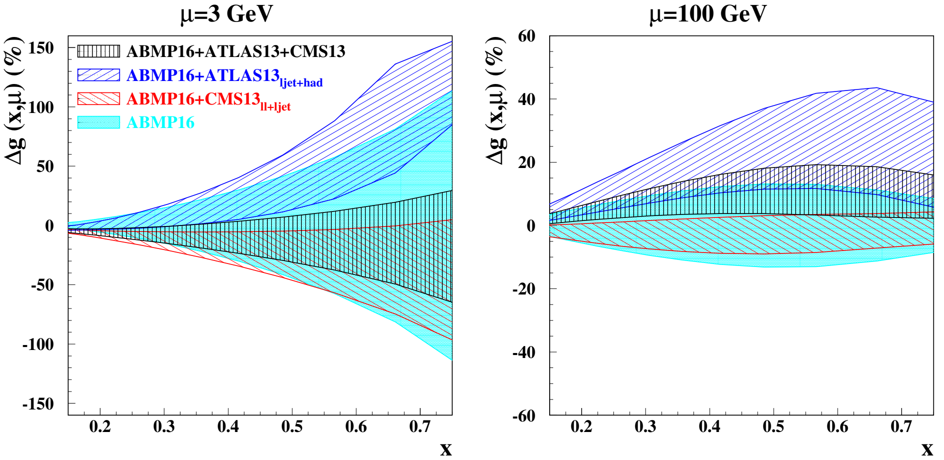
Data set	$m_t(m_t)$ (GeV)	χ^2/NDP
$ATLAS\,13_{ljet}$	158.9 ± 1.3	25.2/19
$ATLAS\,13_{had}$	160.5 ± 2.0	11.3/10
$CMS\,13_{ll}$	161.1 ± 1.4	13.9/15
$CMS\,13_{ljet}$	158.7 ± 0.9	37.4/34

- Overall, good agreement between $m_t(m_t)$ extracted from different data sets
- Good agreement with ABMP16 fit and $\sim$ 50% **reduced uncertainty** on $m_t(m_t)$

$\alpha_s(M_Z, N_f = 5)$		$m_t(m_t)$ (GeV)
Fitted	0.1148(9)	160.5(6)
Fixed	0.114	160.1(4)
	0.116	161.0(4)
	0.118	161.9(4)
	0.120	162.7(4)
	0.122	163.5(4)

- Positive correlation between α_s and $m_t(m_t)$ **reduced once $t\bar{t}$ differential data included in the fit**

Fitted gluon PDF (work in progress)



- Significant reduction of the gluon PDF uncertainty once differential $t\bar{t}$ data are included
- The fitted gluon PDF is consistent with ABMP16
- Some tension between ATLAS and CMS is noticeable
- Work in progess on producing this fit in the xFitter framework (more details in my next talk)

Summary & Outlook

Summary

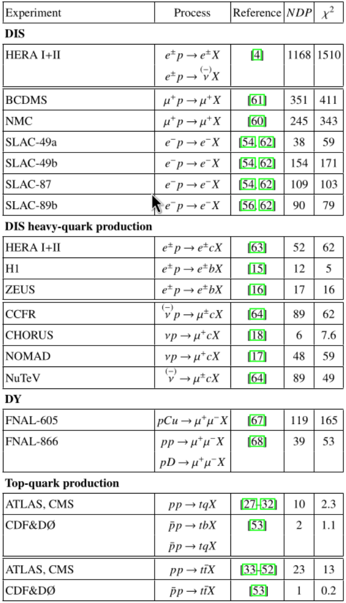
- Measurements of $t\bar{t}$ production at LHC have very rich potential for phenomenology
  - provide information on m_t, α_S and gluon PDF
- We have extracted m_t^{pole} using most recent ATLAS and CMS Run-1,2,3 measurements of $t\bar{t}$ production with a few hundred MeV precision [arXiv:2311.05509]
  - however, pushing this to higher precision might require better understanding of the top quark mass definition used in MC to unfold the data to parton level (+ renormalon)
- Included most recent Run-2 measurements in ABMP16 fit [to appear arXiv:24XX.XXXXX]
  - the new data are consistent with ABMP16
  - significant ($\sim$ of factor 2) constraints on the gluon PDF at $x \gtrsim 0.1$
  - improved ($\sim$ of factor 2) determination of $m_t(m_t)$
  - reduced correlation between α_S and $m_t(m_t)$

Outlook

- Current experimental accuracy is already at % level for normalized $t\bar{t}$ cross sections
  - theory tools require a lot of resources to compute predictions: need to improve
- Differential measurements of $t\bar{t}$ production seem to show some tension
  - ATLAS+CMS effort on combining differential $t\bar{t}$ data will be useful
- LHCb covers a complementary kinematic region sensitive to higher x, but no measurements unfolded to parton level (without cuts p_T, y of decay products) were done
  - would be useful to have such measurements from LHCb
- It would be very nice to have NNLO differential calculations for $t\bar{t}$+jets
  - would allow direct constraints on α_S

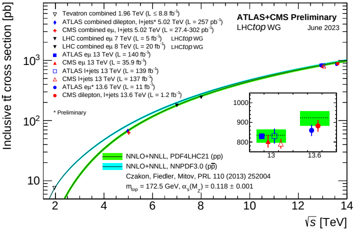
BACKUP

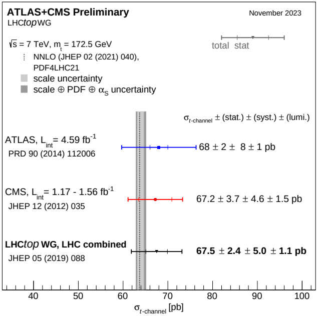
Data in ABMP16 fit [Alekhin et al., arXiv:1701:05838] (1)



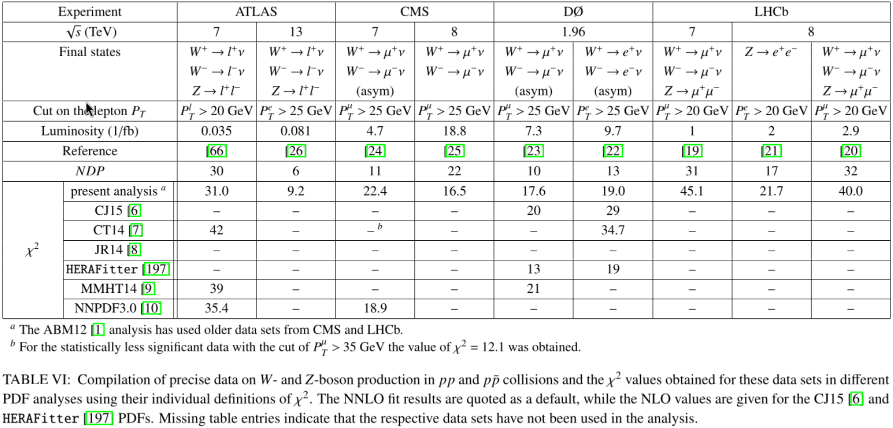
Experiment	Process	Reference	NDP	χ^2
DIS				
HERA I+II	$e^\pm p \to e^\pm X$ $e^\pm p \to \overset{(-)}{\nu}X$	[4]	1168	1510
BCDMS	$\mu^+ p \to \mu^+ X$	[61]	351	411
NMC	$\mu^+ p \to \mu^+ X$	[60]	245	343
SLAC-49a	$e^- p \to e^- X$	[54, 62]	38	59
SLAC-49b	$e^- p \to e^- X$	[54, 62]	154	171
SLAC-87	$e^- p \to e^- X$	[54, 62]	109	103
SLAC-89b	$e^- p \to e^- X$	[56, 62]	90	79
DIS heavy-quark production				
HERA I+II	$e^\pm p \to e^\pm cX$	[63]	52	62
H1	$e^\pm p \to e^\pm bX$	[15]	12	5
ZEUS	$e^\pm p \to e^\pm bX$	[16]	17	16
CCFR	$\overset{(-)}{\nu} p \to \mu^\pm cX$	[64]	89	62
CHORUS	$\nu p \to \mu^+ cX$	[18]	6	7.6
NOMAD	$\nu p \to \mu^+ cX$	[17]	48	59
NuTeV	$\overset{(-)}{\nu} \to \mu^\pm cX$	[64]	89	49
DY				
FNAL-605	$pCu \to \mu^+\mu^- X$	[67]	119	165
FNAL-866	$pp \to \mu^+\mu^- X$ $pD \to \mu^+\mu^- X$	[68]	39	53
Top-quark production				
ATLAS, CMS	$pp \to tqX$	[27–32]	10	2.3
CDF&DØ	$\bar{p}p \to tbX$ $\bar{p}p \to tqX$	[53]	2	1.1
ATLAS, CMS	$pp \to t\bar{t}X$	[33–52]	23	13
CDF&DØ	$\bar{p}p \to t\bar{t}X$	[53]	1	0.2

$t\bar{t}$ and single top LHC data have been updated to latest LHCTOPWG:





Data in ABMP16 fit [Alekhin et al., arXiv:1701:05838] (2)



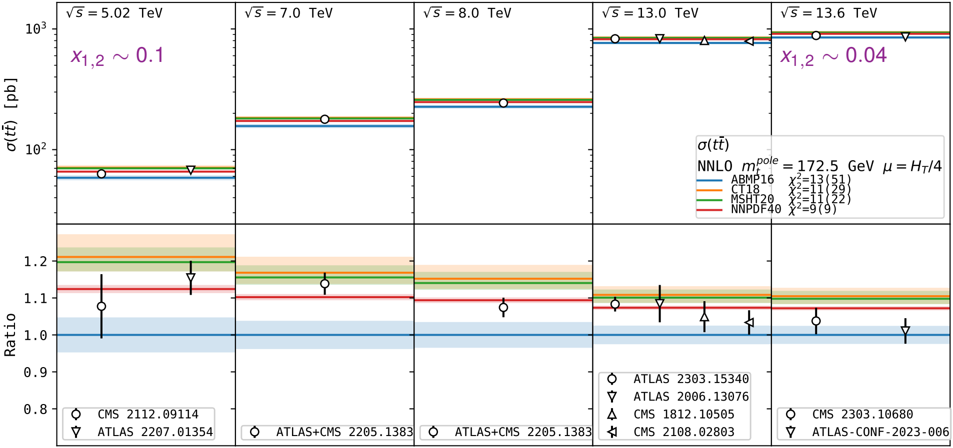
Experiment		ATLAS		CMS		DØ		LHCb		
$\sqrt{s}$ (TeV)		7	13	7	8	1.96		7	8	
Final states		$W^+ \to l^+\nu$ $W^- \to l^-\nu$ $Z \to l^+l^-$	$W^+ \to l^+\nu$ $W^- \to l^-\nu$ $Z \to l^+l^-$	$W^+ \to \mu^+\nu$ $W^- \to \mu^-\nu$ (asym)	$W^+ \to \mu^+\nu$ $W^- \to \mu^-\nu$	$W^+ \to \mu^+\nu$ $W^- \to \mu^-\nu$ (asym)	$W^+ \to e^+\nu$ $W^- \to e^-\nu$ (asym)	$W^+ \to \mu^+\nu$ $W^- \to \mu^-\nu$ $Z \to \mu^+\mu^-$	$Z \to e^+e^-$	$W^+ \to \mu^+\nu$ $W^- \to \mu^-\nu$ $Z \to \mu^+\mu^-$
Cut on the lepton P_T		$P_T^l > 20$ GeV	$P_T^e > 25$ GeV	$P_T^\mu > 25$ GeV	$P_T^\mu > 25$ GeV	$P_T^\mu > 25$ GeV	$P_T^e > 25$ GeV	$P_T^\mu > 20$ GeV	$P_T^e > 20$ GeV	$P_T^\mu > 20$ GeV
Luminosity (1/fb)		0.035	0.081	4.7	18.8	7.3	9.7	1	2	2.9
Reference		[66]	[26]	[24]	[25]	[23]	[22]	[19]	[21]	[20]
NDP		30	6	11	22	10	13	31	17	32
χ^2	present analysis [a]	31.0	9.2	22.4	16.5	17.6	19.0	45.1	21.7	40.0
	CJ15 [6]	–	–	–	–	20	29	–	–	–
	CT14 [7]	42	–	– [b]	–	–	34.7	–	–	–
	JR14 [8]	–	–	–	–	–	–	–	–	–
	`HERAFitter` [197]	–	–	–	–	13	19	–	–	–
	MMHT14 [9]	39	–	–	–	21	–	–	–	–
	NNPDF3.0 [10]	35.4	–	18.9	–	–	–	–	–	–

[a] The ABM12 [1] analysis has used older data sets from CMS and LHCb.

[b] For the statistically less significant data with the cut of $P_T^\mu > 35$ GeV the value of $\chi^2 = 12.1$ was obtained.

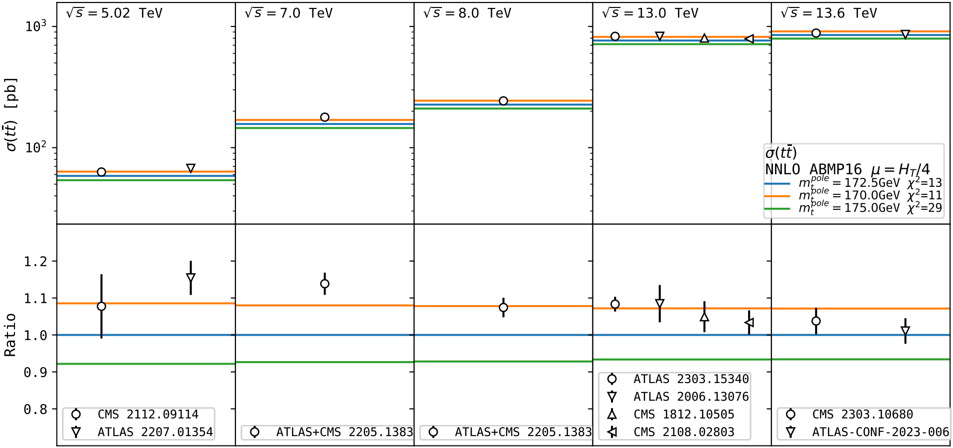
TABLE VI: Compilation of precise data on W- and Z-boson production in pp and $p\bar{p}$ collisions and the χ^2 values obtained for these data sets in different PDF analyses using their individual definitions of χ^2. The NNLO fit results are quoted as a default, while the NLO values are given for the CJ15 [6] and `HERAFitter` [197] PDFs. Missing table entries indicate that the respective data sets have not been used in the analysis.

$\sigma(t\bar{t})$ vs NNLO predictions using different PDFs



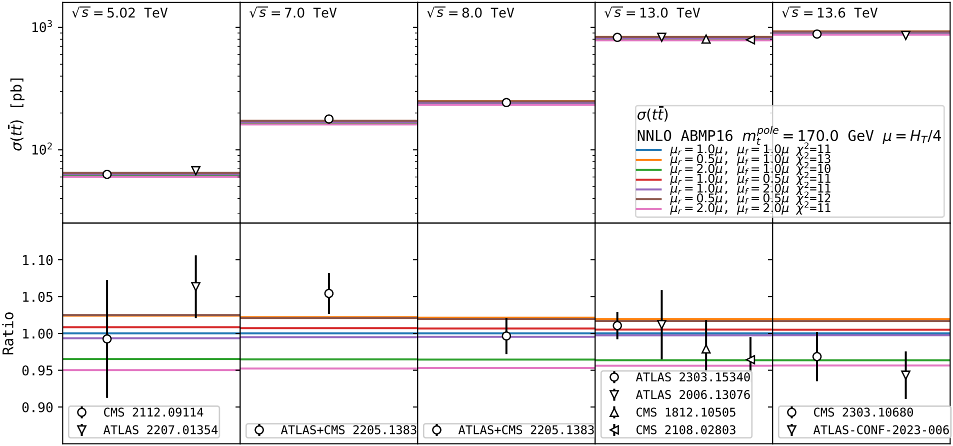
- Fixed $m_t^{\text{pole}} = 172.5$ GeV, $\mu_r = \mu_f = H_T/4$
- Reported χ^2 values with (and without) PDF uncertainties
- All PDF sets describe data reasonably well (depends on m_t^{pole}, α_S)
- Sensitivity to PDFs reduces with increasing $\sqrt{s}$ (lower x probed)

$\sigma(t\bar{t})$ vs NNLO predictions using different m_t^{pole}



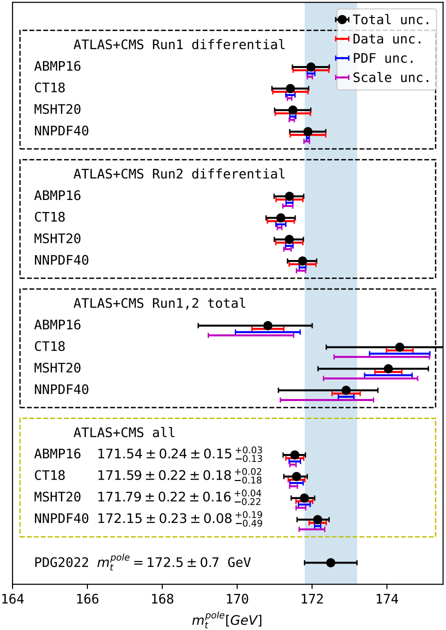
- ABMP16, fixed $\mu_r = \mu_f = H_T/4$
- Change of m_t^{pole} by 1 GeV → change of $\sigma(t\bar{t})$ by $\approx 3\%$
- Preferable $m_t^{\text{pole}} \sim 170$–$172.5$ GeV (depends on PDF and α_S)

$\sigma(t\bar{t})$ vs NNLO predictions with scale variations



- ABMP16, fixed $m_t^{\text{pole}} = 172.5$ GeV
- Scale variations $^{+3\%}_{-5\%}$:
  - larger than data uncertainty (best data uncertainty ±1.9%)
  - limit precision of m_t^{pole} extraction to 1 GeV
  - can be reduced by using e.g. $\overline{\text{MS}}$ mass $m_t(m_t)$ EPJ C74 (2014) 3167, JHEP04 (2021) 043

Extraction of m_t^{pole}: summary

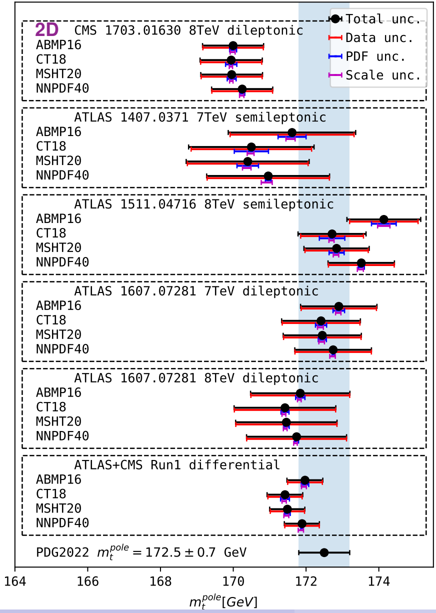


- Extracted m_t^{pole} values with precision ± 0.3 GeV are consistent with PDG value 172.5 ± 0.7 GeV
  - data uncertainty ~ 0.2 GeV
  - PDF uncertainty ~ 0.1 GeV
  - NNLO scale uncertainty ~ 0.2 GeV
- Significant dependence on PDFs (~ 0.5 GeV):
  - different m_t^{pole} used in different PDFs
  - PDFs, m_t^{pole}, α_s should be determined simultaneously
- For CMS 1904.05237, NNLO results are consistent with published results obtained at NLO
  - good convergence of perturbative series
- Larger sensitivity comes from differential data
  - 2D differential x-sections in $M(t\bar{t})$, $y(t\bar{t})$ constrain m_t^{pole}, PDFs and (indirectly) α_S
  - ideally, 3D cross section in $M(t\bar{t})$, $y(t\bar{t})$ and number of extra jets constrain α_S directly, but NNLO not yet available for $t\bar{t}$ +jets
- Possible effects from Coulomb and soft-gluon resummation near the $t\bar{t}$ production threshold are neglected: might be ~ 1 GeV?

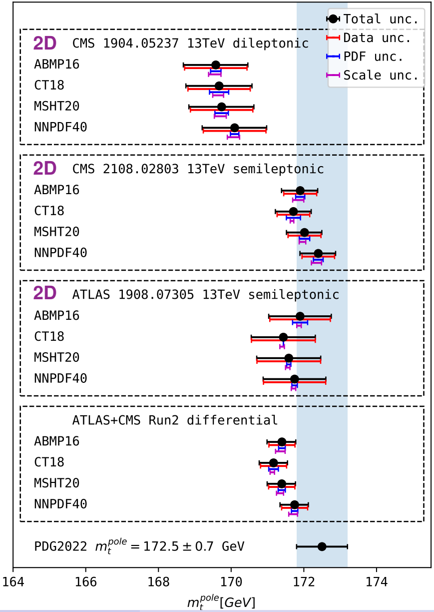
[CMS Coll. EPJ C80 (2020) 658; Kiyo, Kuhn, Moch, Steinhauser, Uwer EPJ C60 (2009) 375; Mäkelä, Hoang, Lipka, Moch 2301.03546]

Extraction of m_t^{pole}: differential Run 1, Run 2

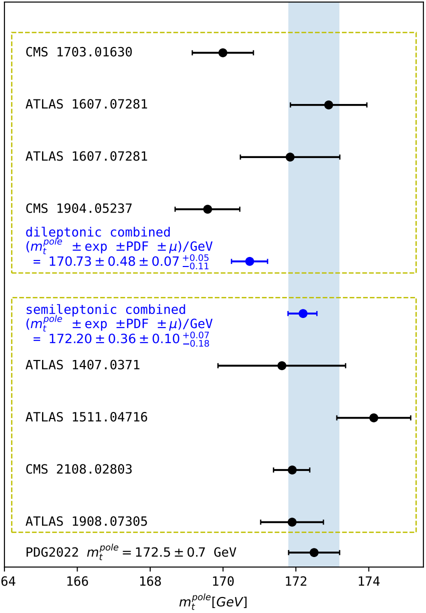
Run 1 differential

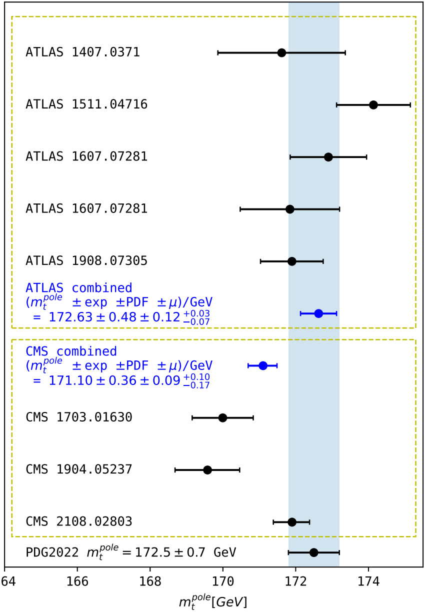


Run 2 differential

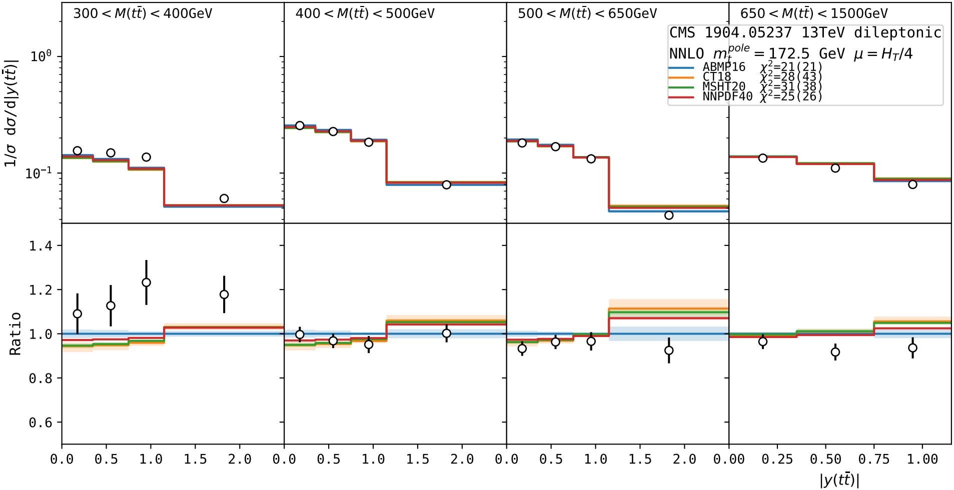


Extraction of m_t^{pole}: dilepton vs semileptonic, ATLAS vs CMS

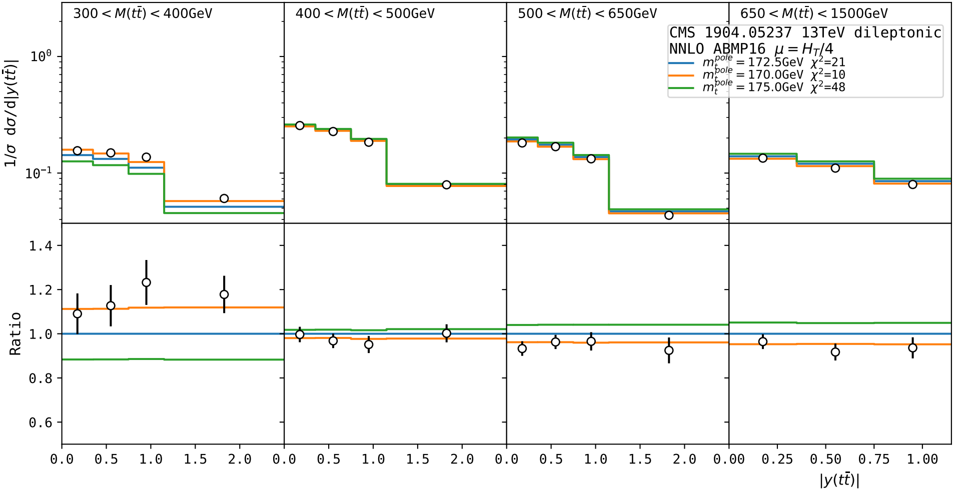




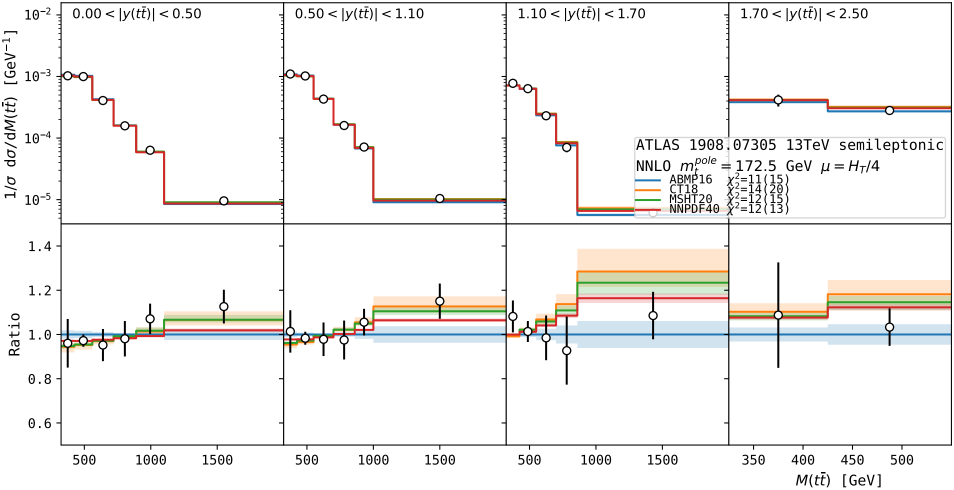
Data vs NNLO predictions using different PDFs



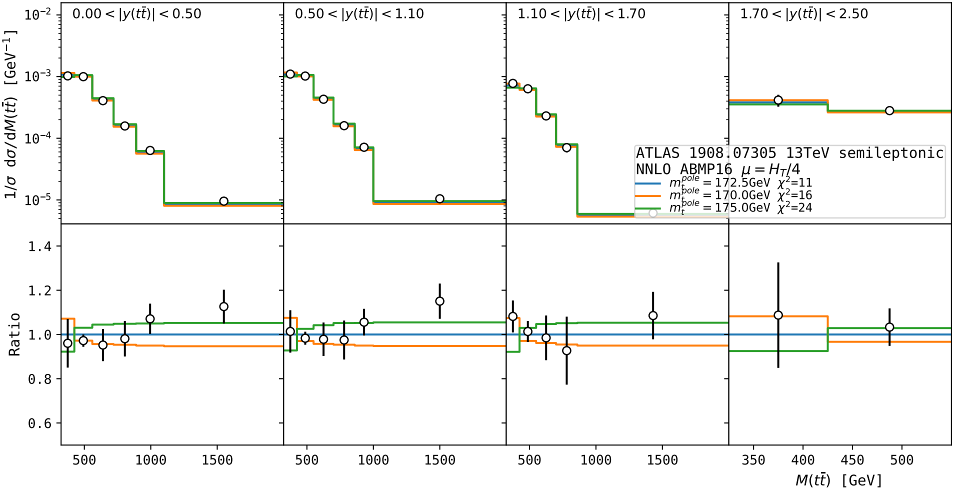
Data vs NNLO predictions using different m_t^{pole}



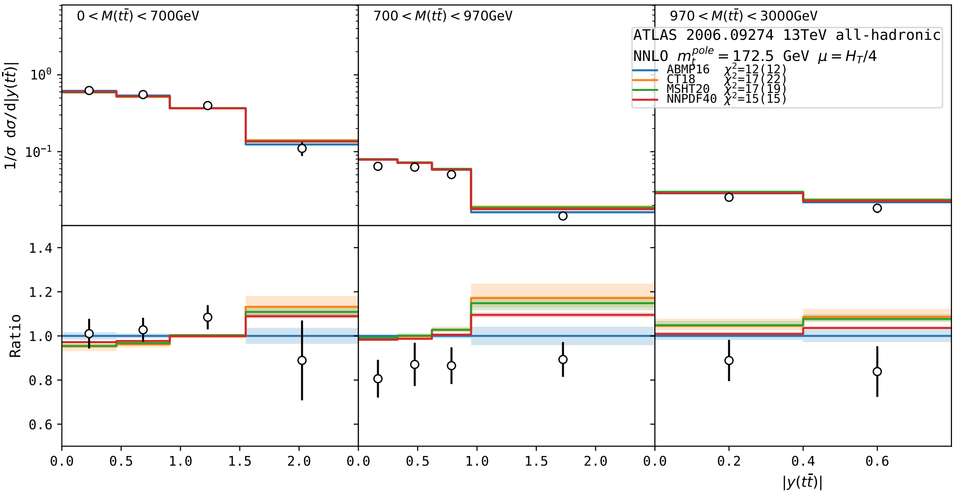
Data vs NNLO predictions using different PDFs



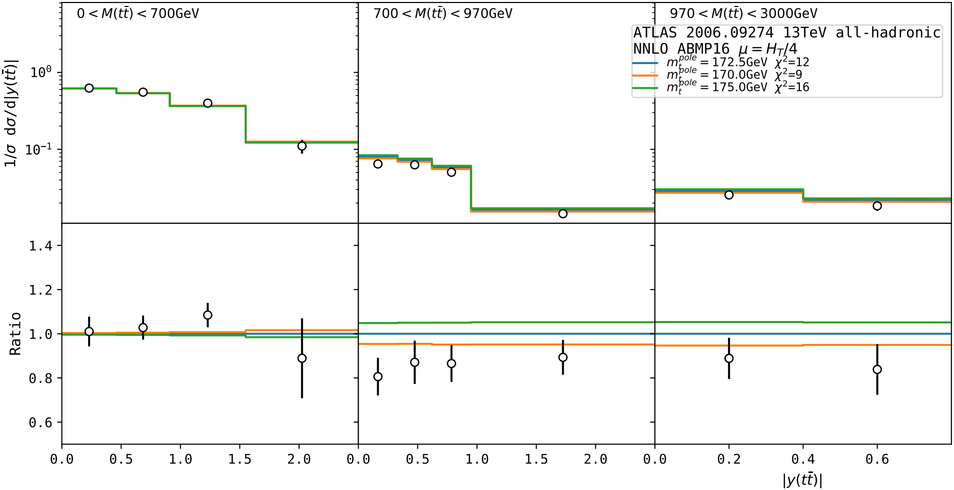
Data vs NNLO predictions using different m_t^{pole}



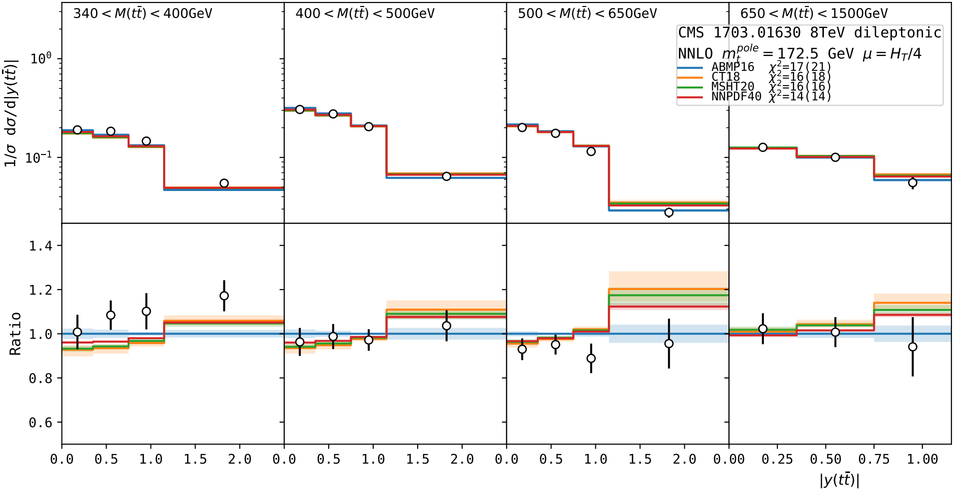
Data vs NNLO predictions using different PDFs



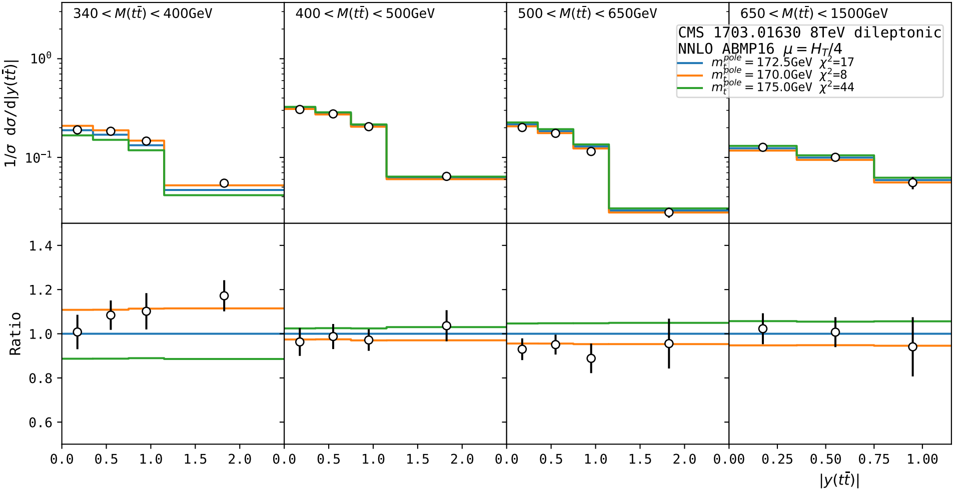
Data vs NNLO predictions using different m_t^{pole}



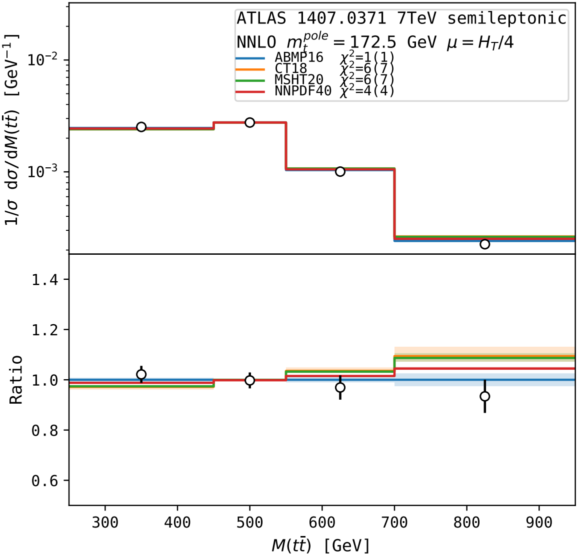
Data vs NNLO predictions using different PDFs



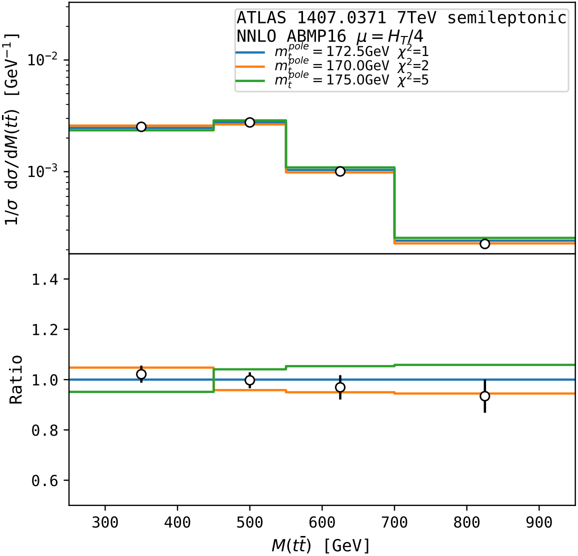
Data vs NNLO predictions using different m_t^{pole}



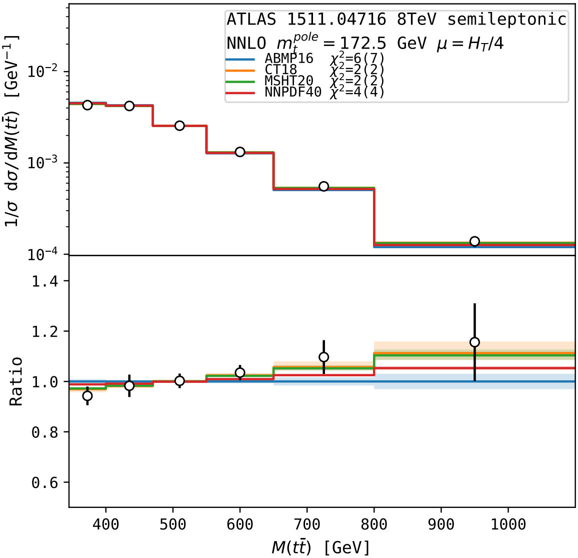
Data vs NNLO predictions using different PDFs



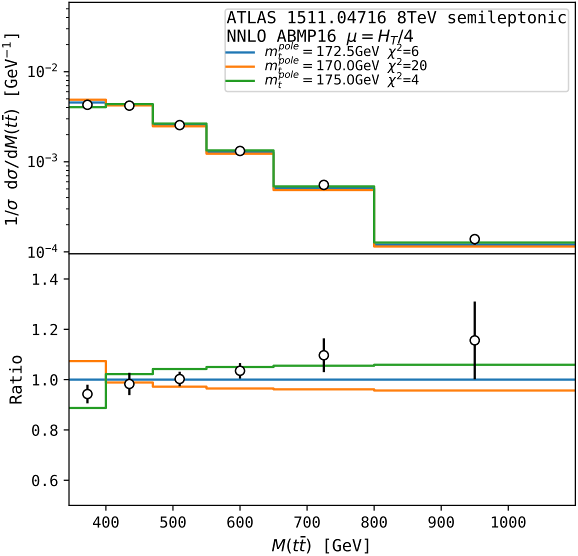
Data vs NNLO predictions using different m_t^{pole}



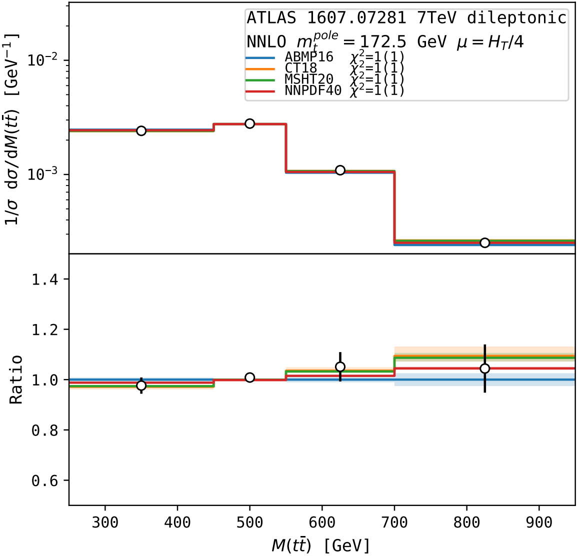
Data vs NNLO predictions using different PDFs



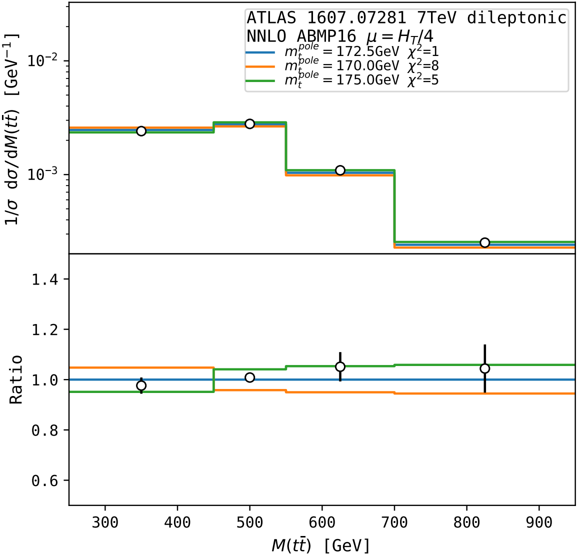
Data vs NNLO predictions using different m_t^{pole}



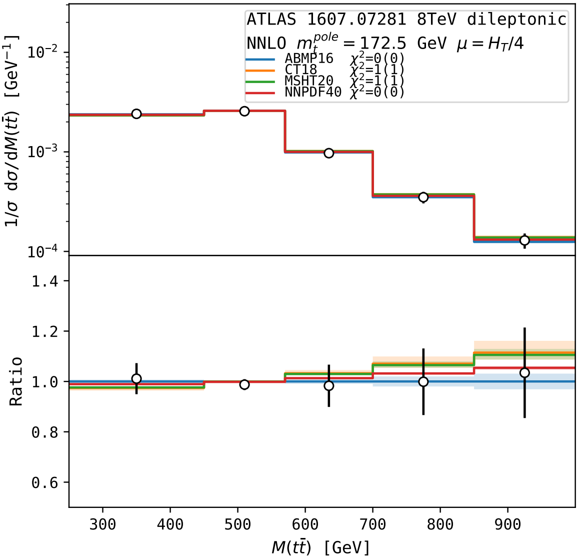
Data vs NNLO predictions using different PDFs



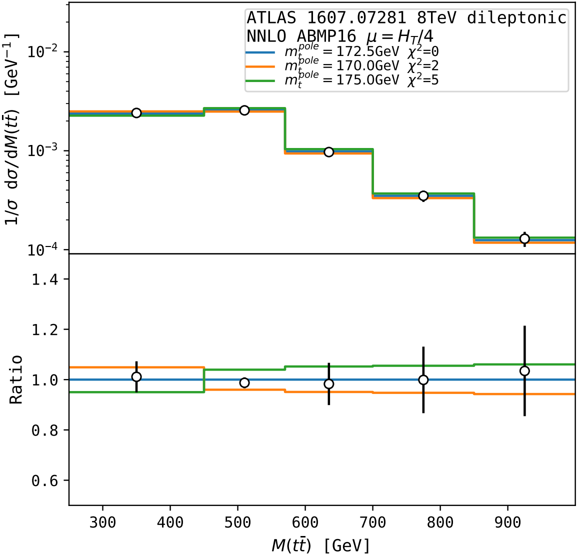
Data vs NNLO predictions using different m_t^{pole}



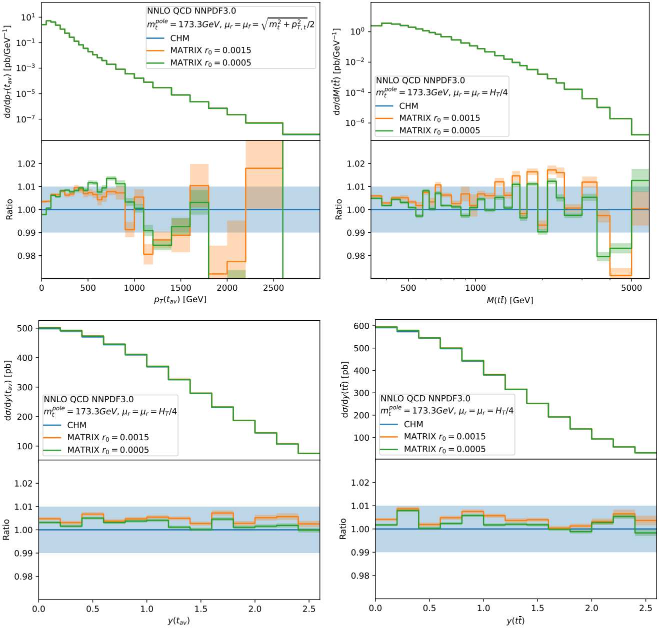
Data vs NNLO predictions using different PDFs



Data vs NNLO predictions using different m_t^{pole}

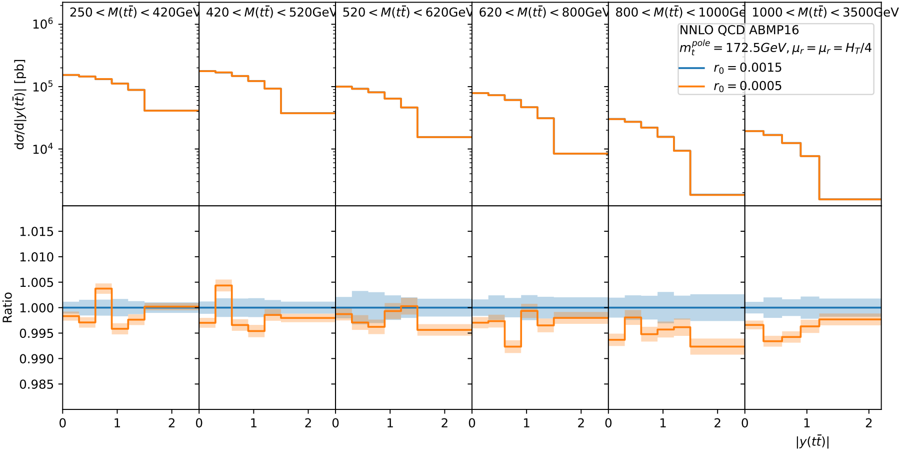


Variation of r cut, validation vs JHEP 04 (2017) 071 by Czakon et al. [CHM]

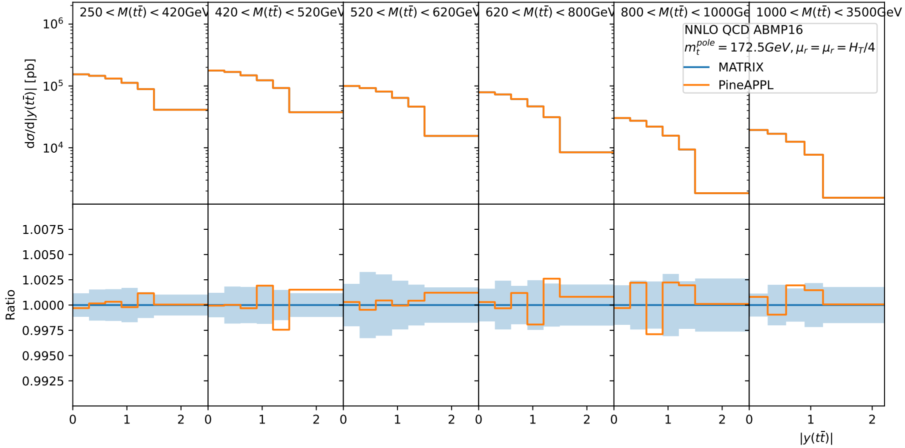


Good agreement $< 1\%$

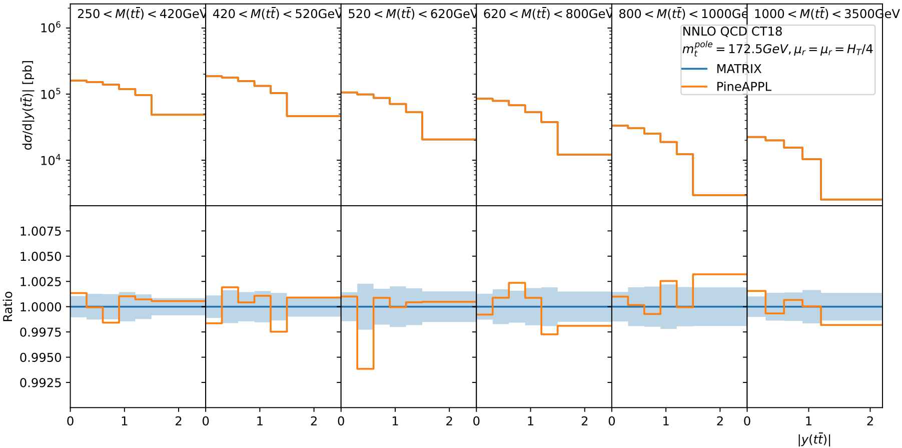
r cut variation in bins of TOP-20-001



PineAppl vs MATRIX in bins of TOP-20-001 [ABMP16]

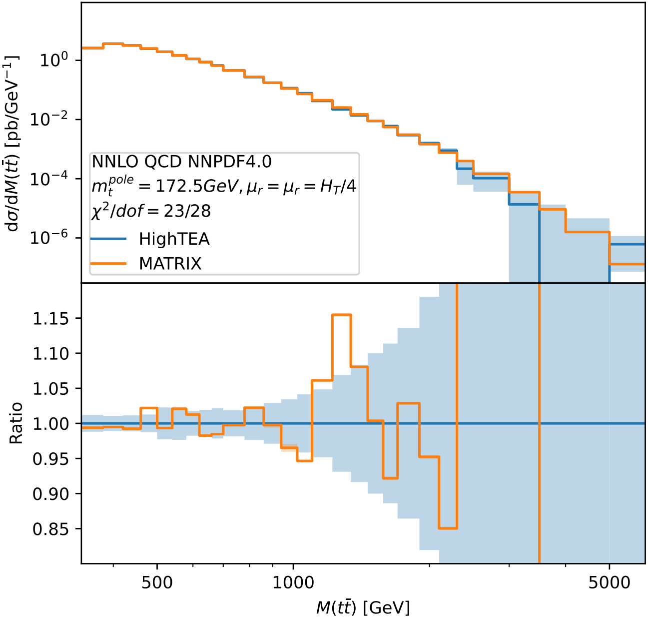


PineAppl vs MATRIX in bins of TOP-20-001 [CT18]

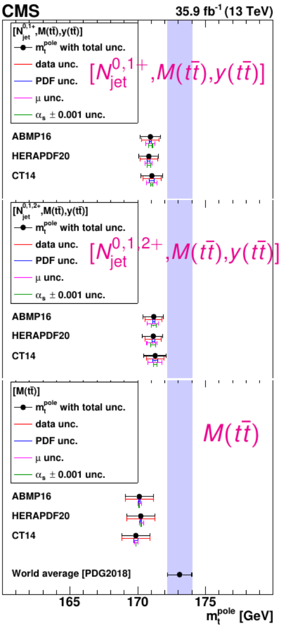


- grids were produced with ABMP16

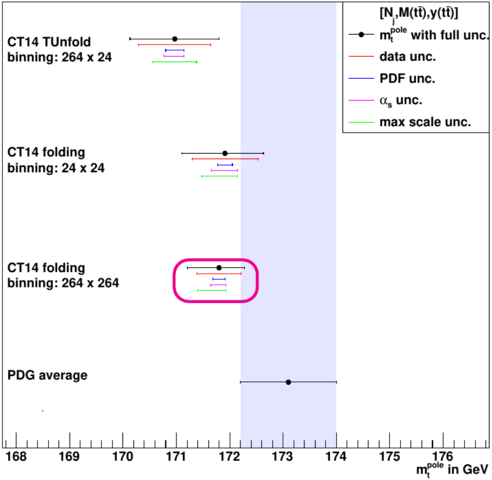
MATRIX vs HighTEA



CMS TOP-18-004 checks



DESY 2018 summer school, L. Materne, bachelor thesis “Differential Top-Pair Production Cross Section with the CMS Detector - Optimization of Measurement Information”, Karlsruher Institut für Technologie (KIT), Bachelorarbeit, 2018 [ETP-Bachelor-KA/2018-11]



m_t dependence of measured cross sections [CMS TOP-18-004]