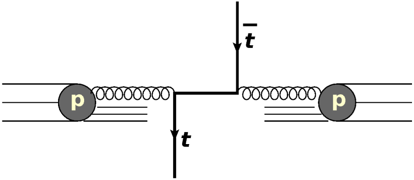
CMS normalised multi-differential $t\bar{t}$ cross-sections and simultaneous determination of α_s, m_t^{pole} and PDFs
[CMS-PAS-TOP-18-004]

Oleksandr Zenaiev (DESY)
on behalf of the CMS Collaboration

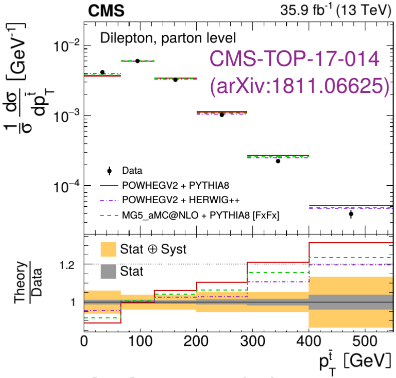
LHCTopWG Meeting, CERN
20.11.2018

Introduction

Why measure $t\bar{t}$ 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

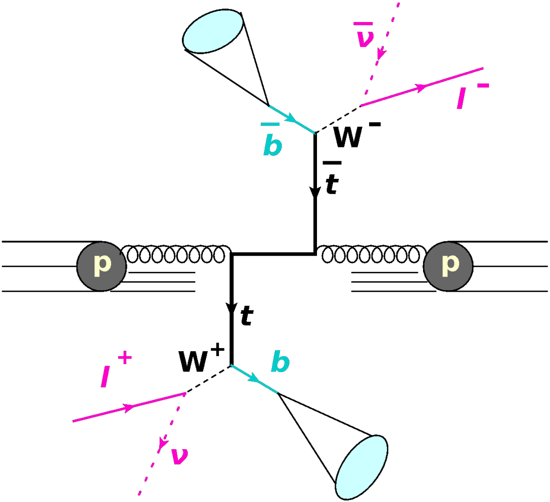


CMS-TOP-17-014
(arXiv:1811.06625)

Why measure 2D/3D?

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

Event selection

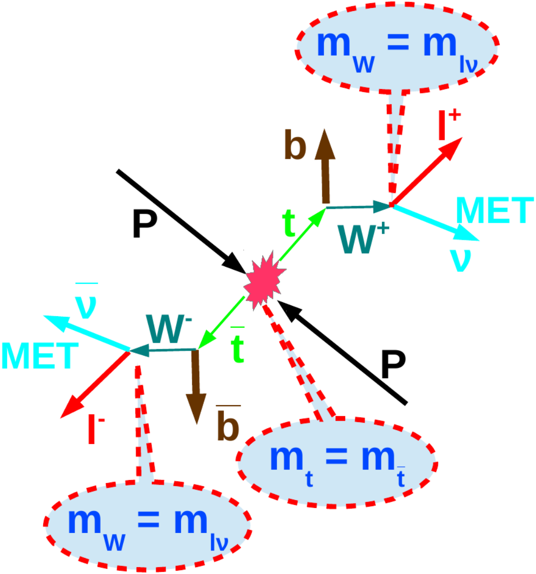


Follows 1D measurement:

CMS-TOP-17-014 (arXiv:1811.06625)

- **Leptons:**
  - 2 isolated $l^{\pm} l^{\mp}$
  - $p_T > 20(25)$ GeV
  - $|\eta| < 2.4$
- **Jets:**
  - at least 2 jets
  - $p_T > 30$ GeV
  - $|\eta| < 2.4$
  - at least 1 *b*-tagged

Kinematic reconstruction



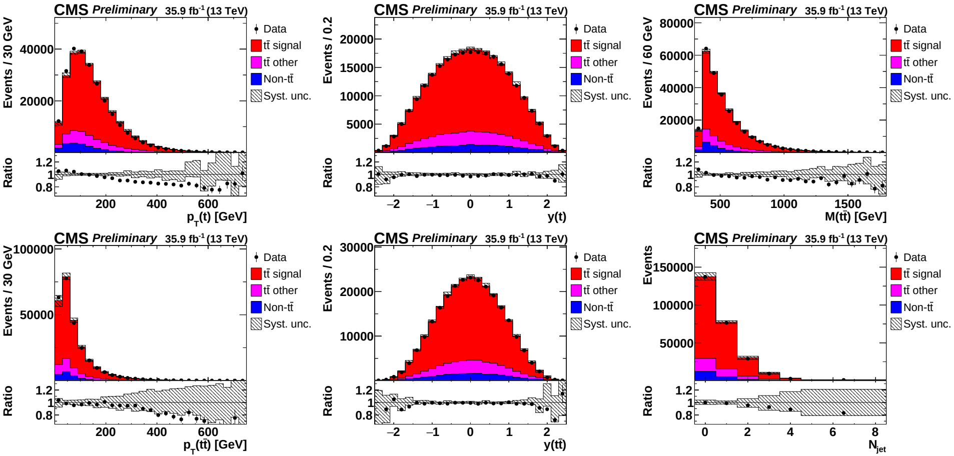
- Measured input: leptons, jets, MET
- Unknowns: $\bar{p}_\nu$, $\bar{p}_{\bar{\nu}}$ (6)
- Constraints:
  - m_t, $m_{\bar{t}}$ (2)
  - m_{W^+}, m_{W^-} (2)
  - $(\bar{p}_\nu + \bar{p}_{\bar{\nu}})_T = \mathrm{MET}$ (2)

Two variants:

(1) Full reconstruction:
  - recover t, $\bar{t}$
  - use all constraints

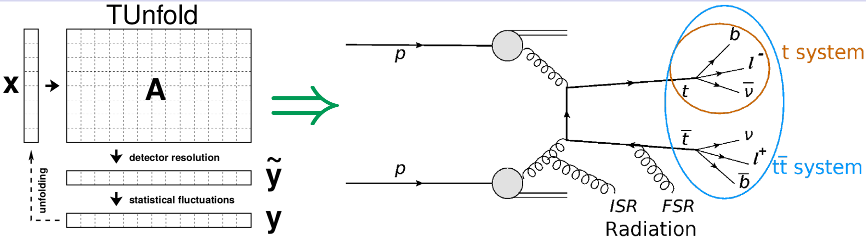
(2) Loose reconstruction:
  - recover $t\bar{t}$
  - m_t constraints not used → reliable to extract m_t^{pole}

Kinematic distributions



- $t\bar{t}$ signal MC: POWHEGV2 + PYTHIA8 (details in BACKUP)
- Overall good description of data within uncertainties
- Central MC predictions for $p_T(t)$, $p_T(t\bar{t})$, $M(t\bar{t})$, $N_{\rm jet}$ are softer than data

Overview of measured cross sections

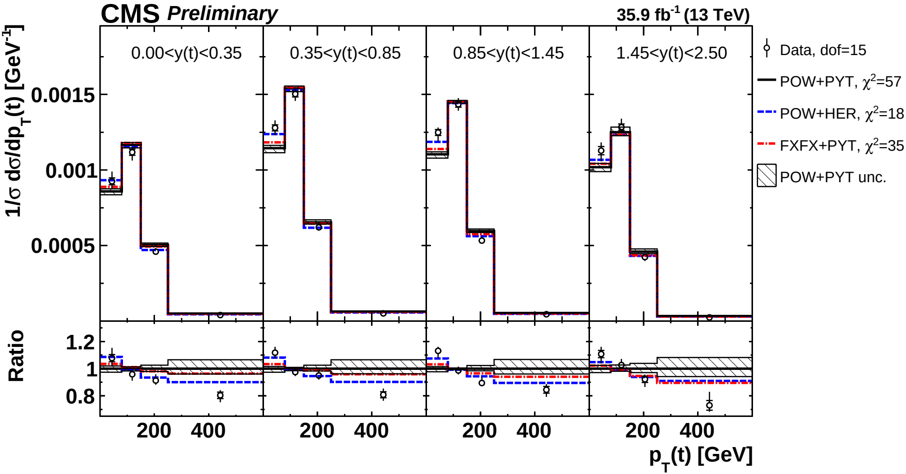


- ***t* production:**
  - [$y(t)$,$p_T(t)$]: most simple
- ***$t\bar{t}$* production:**
  - [$M(t\bar{t})$,$y(t\bar{t})$]: most sensitive to PDFs (at LO $x_{1,2} = \sqrt{\frac{M(t\bar{t})}{s}}e^{\pm y(t\bar{t})}$)
  - [$M(t\bar{t})$,$p_T(t\bar{t})$]: sensitive to radiation (at LO $p_T(t\bar{t}) \equiv 0$)
- ***t*, *$t\bar{t}$* mixed:**
  - [$M(t\bar{t})$,$y(t)$]: sensitive to PDFs (at LO $y(t\bar{t}) = (y(t) + y(\bar{t}))/2$)
  - [$M(t\bar{t})$,$\Delta\phi(t,\bar{t})$]: sensitive to radiation (at LO $\Delta\phi(t\bar{t}) \equiv \pi$)
  - [$M(t\bar{t})$,$\Delta\eta(t,\bar{t})$]: correlated with $p_T(t)$ and may shade light on $p_T(t)$ problem
  - [$M(t\bar{t})$,$p_T(t)$]: may shade further light on $p_T(t)$ problem
- **NEW *$t\bar{t}$* production with extra jets:**
  - [$N_{\text{jet}}^{0,1+}$,$M(t\bar{t})$,$y(t\bar{t})$]: sensitive to α_s, $m_t^{\text{pole}}$and PDFs (nominal extraction)
  - [$N_{\text{jet}}^{0,1,2+}$,$M(t\bar{t})$,$y(t\bar{t})$]: sensitive to α_s, $m_t^{\text{pole}}$and PDFs (cross check)

Results

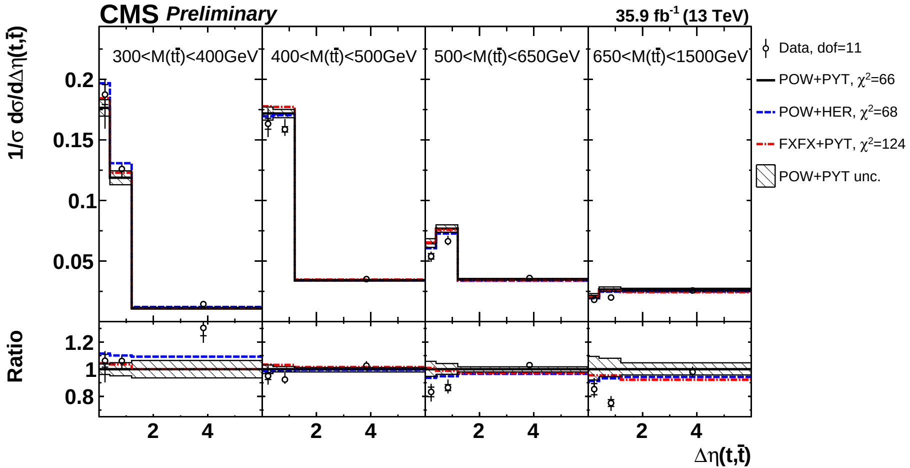
- Measured seven 2D and two 3D cross sections
- All cross sections are provided at parton level for $t\bar{t}$ (before t decay) but particle level for jets
  - corrections particle→ parton level (e.g. for FO predictions) derived from MC ($\lesssim$ 5%)
- 2D and 3D cross sections are compared to MC predictions (details in BACKUP)
  - POWHEGV2 + PYTHIA8, CUETP8M2T4 ('POW-PYT')
  - POWHEGV2 + HERWIG++, EE5C ('POW-HER')
  - MG5_aMC@NLO + PYTHIA8 [FxFx], CUETP8M2T4 ('FXFX-PYT')
- Each comparison is quantified by χ^2 which takes into account data statistical and systematical unc. (list in BACKUP), their correlation, and cross section normalisation
  - resulted χ^2 are translated into *p*-values and compared on one plot (caveat: no theory uncertainties → p-value have limited value)
- Further, 3D cross sections are exploited for α_s +m_t +PDF extraction using NLO (highest order available for $t\bar{t}$ + jets) calculations
  - sensitivity to PDFs from $M(t\bar{t})$, $y(t\bar{t})$ ($x_{1,2} = (M(t\bar{t})/\sqrt{s}) \exp\left[\pm y(t\bar{t})\right]$)
  - sensitivity to α_s from N_{jet} and $M(t\bar{t})$, $y(t\bar{t})$ (PDFs)
  - sensitivity to m_t from $M(t\bar{t})$ via threshold and cone effects

Results: 2D x-sections $[y(t), p_T(t)]$



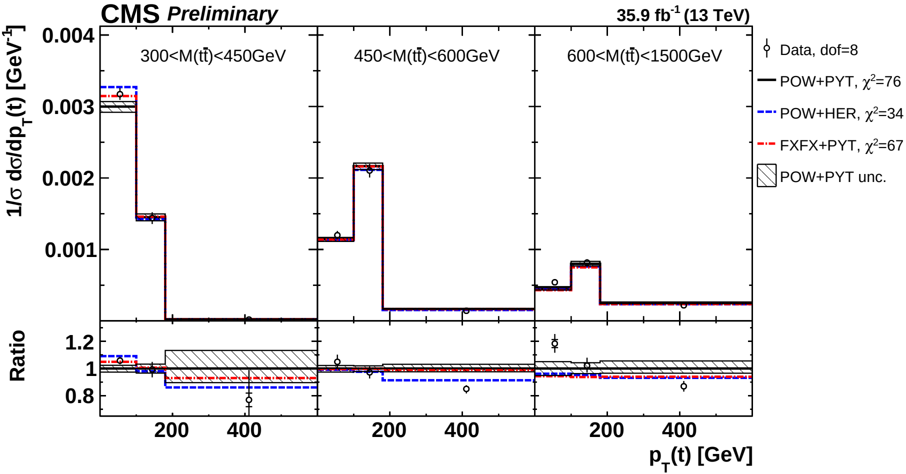
- ‘POW-PYT’ and ‘FXFX-PYT’ predict softer $p_T(t)$ in entire $y(t)$ range
- better description by ‘POW-HER’

Results: 2D cross sections $[M(t\bar{t}),\Delta\eta(t,\bar{t})]$



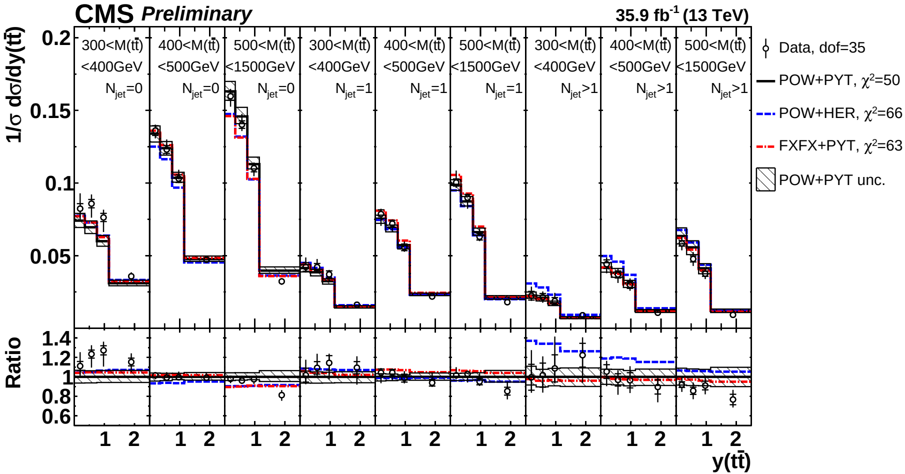
- predicted $\Delta\eta(t,\bar{t})$ are too low at medium and high $M(t\bar{t})$
- at large $M(t\bar{t})$, t and $\bar{t}$ have a larger η separation than in MC: correlated with a lower $p_T(t)$
- bad description by all MC central predictions, strongest disagreement for 'FXFX-PYT'

Results: 2D cross sections $[M(t\bar{t}), p_T(t)]$



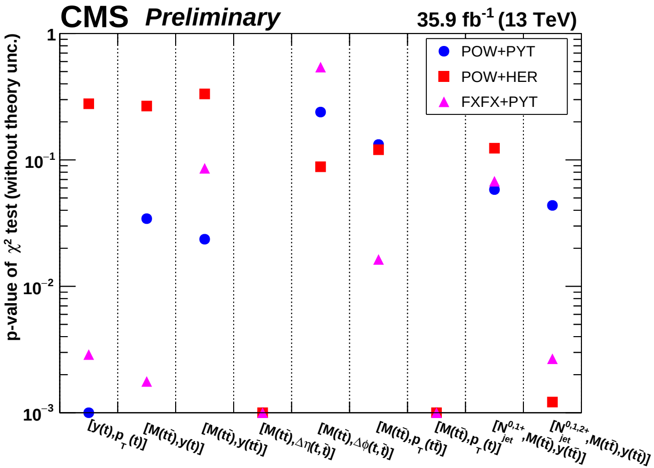
- bad description by all MC, strongest disagreement for ‘POW-PYT’
- notice: ‘POW-HER’ describes $p_T(t)$ in entire $y(t)$ range well, but predicts too hard $p_T(t)$ at high $M(t\bar{t})$

Results: 3D cross sections $[N_{\mathbf{jet}}^{0,1,2+}, M(t\bar{t}), y(t\bar{t})]$



- only 'POW-PYT' is in satisfactory agreement with data
- 'POW-HER' predicts too high cross section at $N_{\text{jet}} > 1$
- 'FXFX-PYT' describes worse $M(t\bar{t})$ at $N_{\text{jet}} = 1$
- . . . more plots in BACKUP

Results: summary of comparison to MC models



- none of central MC predictions is able to describe all distributions, in particular $[M(t\bar{t}), \Delta\eta(t,\bar{t})]$, $[M(t\bar{t}), p_T(t)]$
- overall, best description is provided by ‘POW-PYT’ and ‘POW-HER’:
  - ‘POW-HER’ describes better distributions probing $p_T(t)$
  - ‘POW-PYT’ describes better distributions probing N_{jet} and radiation

Data interpretation consists of two parts:

(1) comparison theory vs data using external PDF sets:

- extracting α_s keeping m_t^{pole} fixed
- extracting m_t^{pole} keeping α_s fixed

→ this presents α_s, m_t^{pole} extraction from $t\bar{t}$ data only

(2) simultaneous fit of PDFs, α_s and m_t^{pole} using $t\bar{t}$ and HERA DIS:

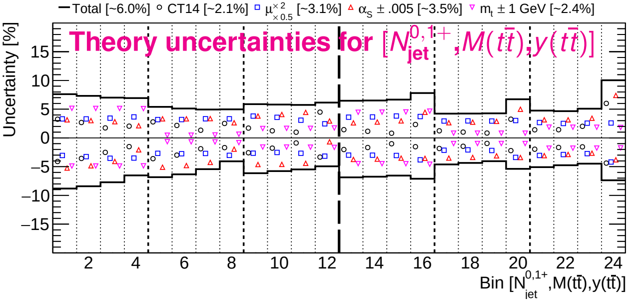
→ this presents fully consistent extraction of α_s, m_t^{pole} and PDFs, but using also HERA data

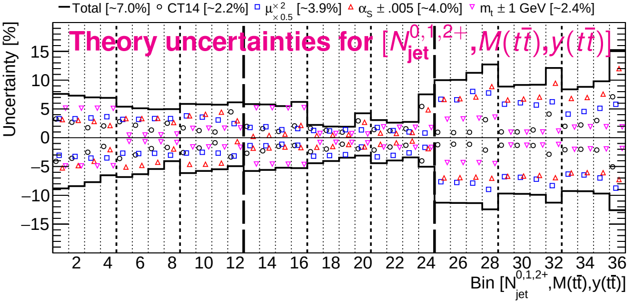
→ important as exercise to understand new $t\bar{t}$ data, providing baseline for future global fits

NLO calculations

- NLO predictions for inclusive $t\bar{t}$, $t\bar{t}$ + 1 jet and $t\bar{t}$ + 2 jets are computed and compared to data using MadGraph5_aMC@NLO + aMCfast + ApplGrid + xFitter:
    - $[N_{\mathrm{jet}}^{0,1+}, M(t\bar{t}), y(t\bar{t})]$ with 2 N_{jet} bins:
        - $\sigma^{\mathrm{NLO}}(N_{\mathrm{jet}} = 0) = \sigma^{\mathrm{NLO}}(t\bar{t}) - \sigma^{\mathrm{NLO}}(t\bar{t} + 1\mathrm{jet})$
        - $\sigma^{\mathrm{NLO}}(N_{\mathrm{jet}} > 0) = \sigma^{\mathrm{NLO}}(t\bar{t} + 1\mathrm{jet})$
    - $[N_{\mathrm{jet}}^{0,1,2+}, M(t\bar{t}), y(t\bar{t})]$ with 3 N_{jet} bins:
        - $\sigma^{\mathrm{NLO}}(N_{\mathrm{jet}} = 0) = \sigma^{\mathrm{NLO}}(t\bar{t}) - \sigma^{\mathrm{NLO}}(t\bar{t} + 1\mathrm{jet})$
        - $\sigma^{\mathrm{NLO}}(N_{\mathrm{jet}} = 1) = \sigma^{\mathrm{NLO}}(t\bar{t} + 1\mathrm{jet}) - \sigma^{\mathrm{NLO}}(t\bar{t} + 2\mathrm{jets})$
        - $\sigma^{\mathrm{NLO}}(N_{\mathrm{jet}} > 1) = \sigma^{\mathrm{NLO}}(t\bar{t} + 2\mathrm{jets})$
- $\mu_{\mathrm{r}} = \mu_{\mathrm{f}} = H'/2$, $H' = \sum_i m_{T,i}$ where the sum runs over all final-state partons (t, $\bar{t}$ and up to three light partons in the $t\bar{t} + 2$ jets calculations) and $m_T = \sqrt{m^2 + p_T^2}$. Uncertainties:
    - $\mu_{\mathrm{r}}, \mu_{\mathrm{f}}$ are varied by factor 2 (6 variations in total)
    - alternative functional form $\mu_{\mathrm{r}} = \mu_{\mathrm{f}} = H/2$, $H = \sum_i m_{T,i}$ with the sum runs over t and $\bar{t}$
- $m_t^{\mathrm{pole}} = 172.5 \pm 1$ GeV (sometimes ± 5 GeV for presentation purposes)
- PDFs and α_s from several groups via LHAPDF, $\alpha_s \pm 0.001$ for uncertainties (sometimes ± 0.005 for presentation purposes)
- multiplied with non-perturbative corrections ($< 5\%$) from parton to particle jet level (BACKUP)

Data and theory uncertainties $[N_{\text{jet}}^{0,1+}, M(t\bar{t}), y(t\bar{t})]$



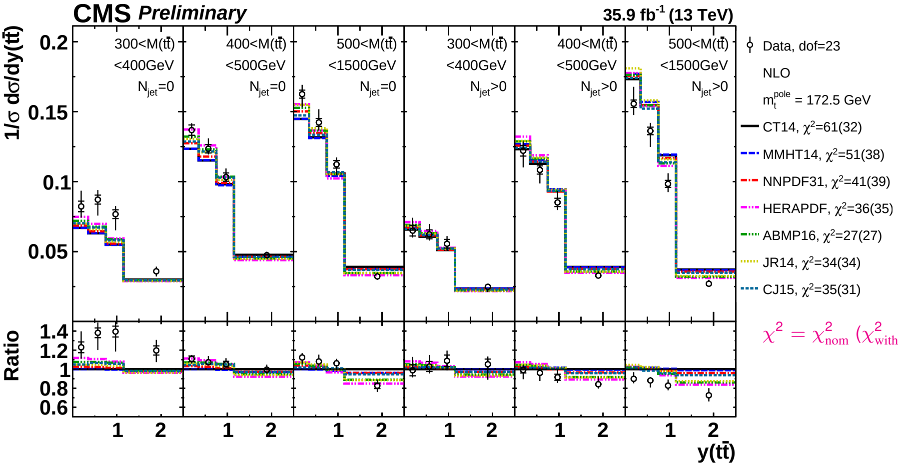


- Bins are grouped for $y(t\bar{t})$, $M(t\bar{t})$ and N_{jet} (separated by different vertical lines)
- NLO scale uncertainties are comparable to PDF, α_s and m_t uncertainties

  → data can constrain PDF, α_s and m_t
- Scale uncertainties are considerably smaller for $[N_{\text{jet}}^{0,1+}, M(t\bar{t}), y(t\bar{t})]$

  → $[N_{\text{jet}}^{0,1,2+}, M(t\bar{t}), y(t\bar{t})]$ is used for cross check only

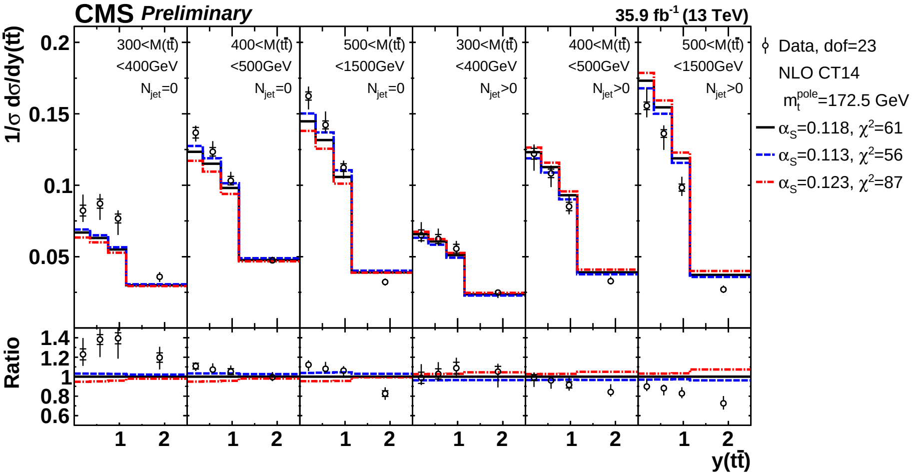
Results: $[N_{\text{jet}}^{0,1+}, M(t\bar{t}), y(t\bar{t})]$ **compared to NLO pred. with diff. PDFs**



$\chi^2 = \chi^2_{\text{nom}}\ (\chi^2_{\text{with PDF unc.}})$

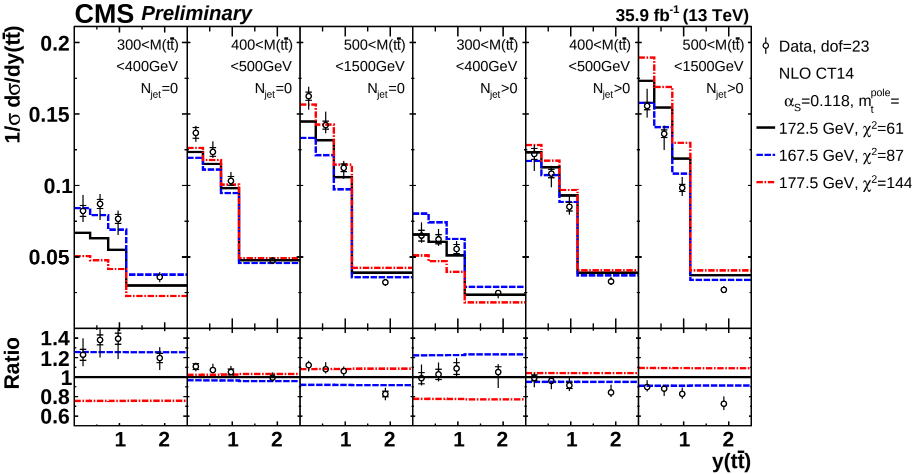
- description depends on PDFs → data are sensitive to PDFs
- All modern PDF sets considered:
  - MMHT2014, ABMP16: total $\sigma(t\bar{t})$ data
  - NNPDF3.1: total and differential (Run-I) $\sigma(t\bar{t})$ data
  - other PDFs: no $t\bar{t}$ data

Results: $[N_{\text{jet}}^{0,1+}, M(t\bar{t}), y(t\bar{t})]$ compared to NLO pred. with diff. α_s



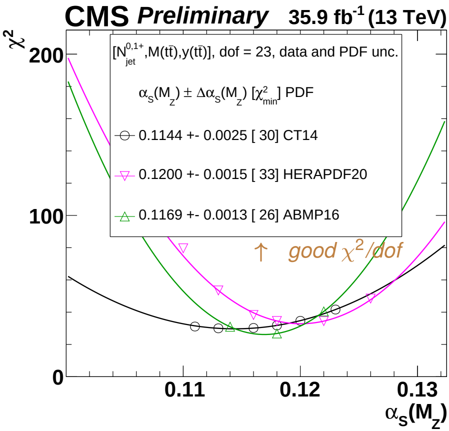
- α_s sensitivity comes from different N_{jet} bins
- also (indirect) sensitivity comes from $[M(t\bar{t}), y(t\bar{t})]$ via sensitivity to PDFs

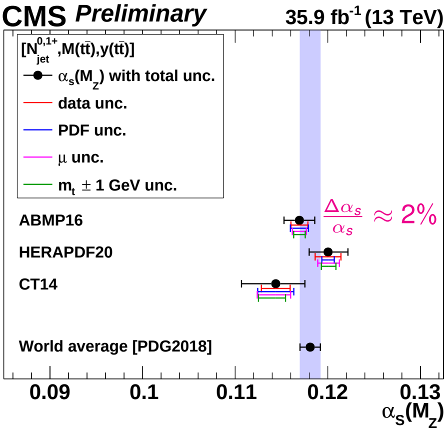
Results: $[N_{\mathbf{jet}}^{0,1+}, M(t\bar{t}), y(t\bar{t})]$ compared to NLO pred. with diff. $m_t^{\mathbf{pole}}$



- m_t sensitivity comes from $M(t\bar{t})$, mainly 1st bin
- this method differs from extracting m_t^{pole} from total $t\bar{t}$ x-section, and is similar to extracting m_t^{pole} from $t\bar{t}j$ diff. x-section [EPJ C73 (2013) 2438, CMS-PAS-TOP-13-006,u JHEP 1510 (2015) 121]
- previous determination using this method: prelim. D0 results [FERMILAB-CONF-16-383-PPD]

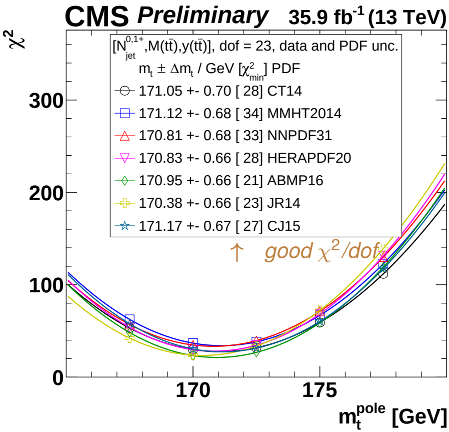
Results: extraction of α_s from $[N_{\text{jet}}^{0,1+}, M(t\bar{t}), y(t\bar{t})]$

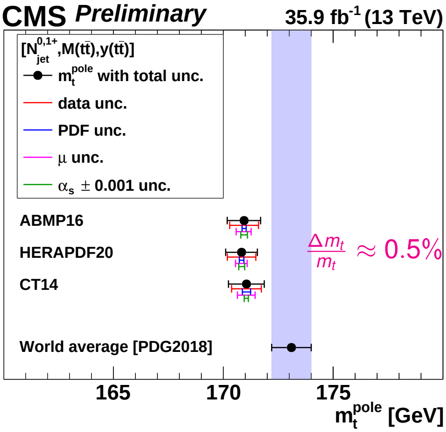




- used $m_t^{\text{pole}} = 172.5$ GeV in ME for all PDF sets (ABMP16 fitted $m_t^{\text{pole}} = 171.44$ GeV)
- precise determination of α_s is possible using these data
- significant dependence on PDF set observed (correlation between g and α_s)

Results: extraction of m_t^{pole} from $[N_{\text{jet}}^{0,1+}, M(t\bar{t}), y(t\bar{t})]$





- used α_s from each PDF set ($\alpha_s = 0.118$ in CT and HERAPDF, $\alpha_s = 0.119$ in ABMP)
- precise determination of m_t^{pole} is possible using these data
- no significant dependence on PDF set

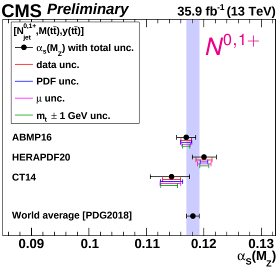
Cross checks

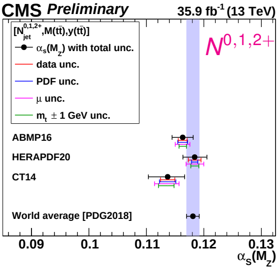
Cross checks of α_s and m_t^{pole} extraction (all results in backup):

- using $[N_{\text{jet}}^{0,1,2+}, M(t\bar{t}), y(t\bar{t})]$
- using single-differential N_{jet}, $M(t\bar{t})$ or $y(t\bar{t})$ cross sections
- using $[p_T(t\bar{t}), M(t\bar{t}), y(t\bar{t})]$ cross sections with 2 $p_T(t\bar{t})$ bins
- using unnormalised cross sections

→ consistent results obtained in all cross checks

→ in this analysis, observables $(\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\ldots})$ have been chosen to have **maximum sensitivity to QCD paramaters and minimum experimental and scale uncertainties**

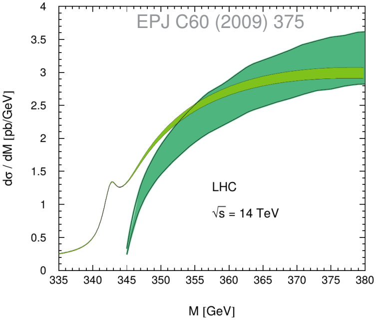


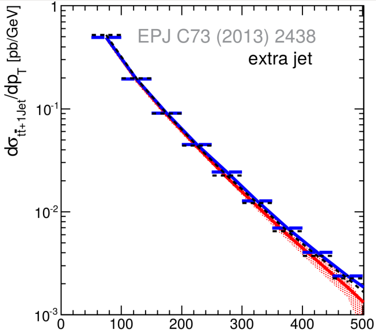


Remarks on limitations in theory calculations

NLO is the only available theory publicly available today, but there are limitations:

- impact of missing threshold resummation is $\Delta m_t \sim 0.7$ GeV [Eur.Phys.J. C60 (2009) 375]
- impact of missing FSR resummation is $\Delta m_t \sim 0.5$ GeV [Eur. Phys. J. C73 (2013) 2438]
  - in general, good agreement between NLO and NLO+PS [Fig. 1 in Eur. Phys. J. C73 (2013) 2438]
- EW corrections could be a few % near threshold [Phys. Rev. D91 (2015) 014020] [JHEP10 (2017) 186]
- **most wanted is NNLO QCD**

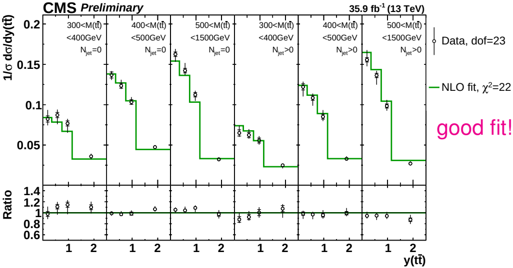




Simultaneous PDF + α_s + m_t^{pole} fit: results

- followed standard approach: using HERA DIS data only, or HERA + $t\bar{t}$ data to demonstrate added value from $t\bar{t}$ on PDF and α_s determination
- settings follow HERAPDF2.0 fit (very similar to TOP-14-013), use xFitter-2.0.0
- input data: combined HERA DIS [1506.06042] + $t\bar{t}$ (further details in BACKUP)

Data sets	χ^2/dof	
	Nominal fit	$+[N_{\text{jet}}, y(\text{t}\bar{\text{t}}), \text{M}(\text{t}\bar{\text{t}})]$
CMS $\text{t}\bar{\text{t}}$		10/23
HERA CC e^-p, $E_\text{p} = 920\,\text{GeV}$	55/42	55/42
HERA CC e^+p, $E_\text{p} = 920\,\text{GeV}$	38/39	39/39
HERA NC e^-p, $E_\text{p} = 920\,\text{GeV}$	218/159	217/159
HERA NC e^+p, $E_\text{p} = 920\,\text{GeV}$	438/377	448/377
HERA NC e^+p, $E_\text{p} = 820\,\text{GeV}$	70/70	71/70
HERA NC e^+p, $E_\text{p} = 575\,\text{GeV}$	220/254	222/254
HERA NC e^+p, $E_\text{p} = 460\,\text{GeV}$	219/204	220/204
Correlated χ^2	82	90
Log-penalty χ^2	+2	−7
Total χ^2/dof	1341/1130	1364/1151



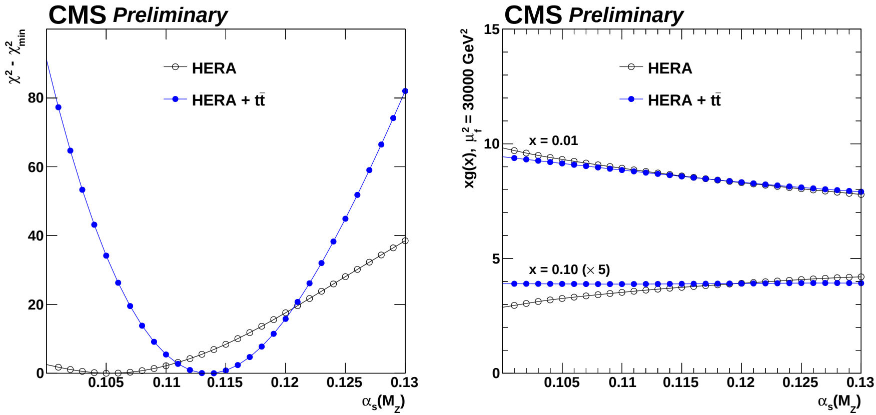
good fit!

$$\alpha_s(M_Z) = 0.1135 \pm 0.0016(\text{fit})^{+0.0002}_{-0.0004}(\text{mod})^{+0.0008}_{-0.0001}(\text{par})^{+0.0011}_{-0.0005}(\text{scale}) = 0.1135^{+0.0021}_{-0.0017}(\text{total})$$

$$m_t^{\text{pole}} = 170.5 \pm 0.7(\text{fit})^{+0.1}_{-0.1}(\text{mod})^{+0.0}_{-0.1}(\text{par})^{+0.3}_{-0.3}(\text{scale})\text{ GeV} = 170.5 \pm 0.8(\text{total})\text{ GeV}$$

→ two SM parameters are simultaneously determined from these data to high precision with only weak correlation between them ($\rho = 0.3$) + constraints on PDFs (next slides)

Simultaneous PDF + α_s + m_t^{pole} fit: correlation between α_s and gluon

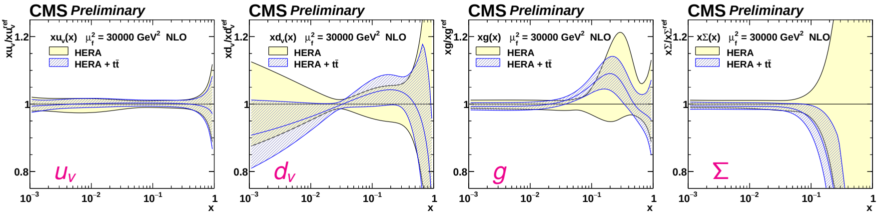


Adding $t\bar{t}$ data:

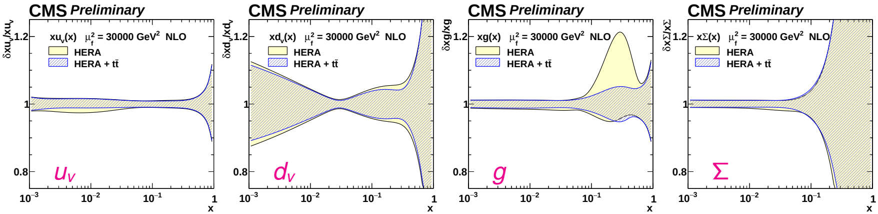
- constrain α_s (left)
- reduce correlation between α_s and gluon (g) (right)
  - weak correlation (α_s,m_t) → weak correlation (g,m_t)

Simultaneous PDF + α_s + m_t^{pole} fit: Impact on PDFs

PDFs (α_s in HERA-only fit set to $\alpha_s = 0.1135 \pm 0.0016$)



Relative PDF uncertainties

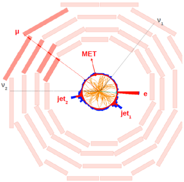


→ reduced g uncertainty at high x

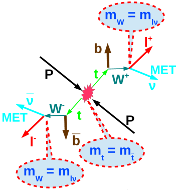
→ smaller impact on other distributions via correlations in the fit

Summary

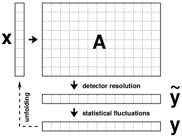
Event selection:
as in 1D



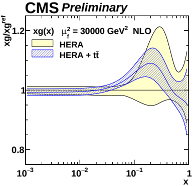
Kinematic reconstruction:
as in 1D + loose for m_t^{pole}

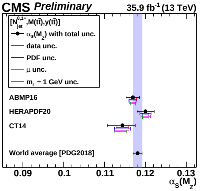


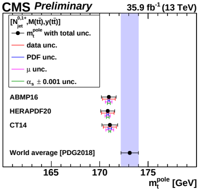
Unfolding: TUnfold,
no nuisance par. fit



Interpretation:
new approach



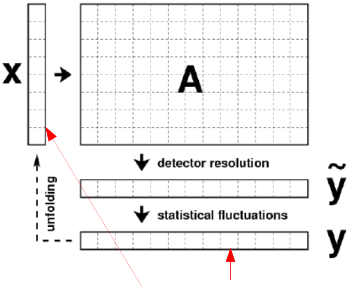




- Measured 2D and 3D $t\bar{t}$ cross section in dilepton channel using 2016 data
- Quantitative comparison to several MC predictions:
  - data distinguish between predictions and reveal trends
  - NNLO from theorists is not yet available
- Used measured 3D cross sections to constrain α_s, m_t^{pole}, PDFs at NLO
  - first extraction of such kind using differential $t\bar{t}$ cross sections
  - most precise result on m_t^{pole} from single analysis to date:
    - ★ uncertainty on m_t^{pole} is comparable to PDG 2018
  - α_s and m_t^{pole} are extracted simultaneously
  - need 3D NNLO prediction, especially for future analyses

BACKUP

Unfolding



`TUnfold` [JINST 7 (2012) T10003]

χ^2 minimisation with regularisation ($\approx 1\%$)

2d distributions are mapped to 1d arrays

reco. data

unfolded distribution

regularization strength

regularization conditions (second derivative)

$$\chi^2 = (Y - AX)^T V_Y^{-1} (Y - AX) + \tau^2 (X - X_0)^T L^T L (X - X_0)$$

migration probability matrix

stat. errors of reco.

gen. distribution

$$Y = N_{measured} - N_{Background}$$

For each Δ**a**^{i}:

$$\boxed{\left(\frac{1}{\sigma}\frac{d\sigma}{db}\right)^{ij} = \frac{1}{\sigma} \cdot \frac{X^{ij}}{BR \cdot L \cdot \Delta b^j}}$$

Systematic uncertainties

Experimental uncertainties:

- JES (splitted in sources, also propagated to MET)
- JER
- b-tagging SFs
- lepton ID/ISO SFs
- triggers SFs
- pileup reweighting
- non-$t\bar{t}$ background normalisation varied by 30%
- lumi and branching ratios cancel for normalised cross section

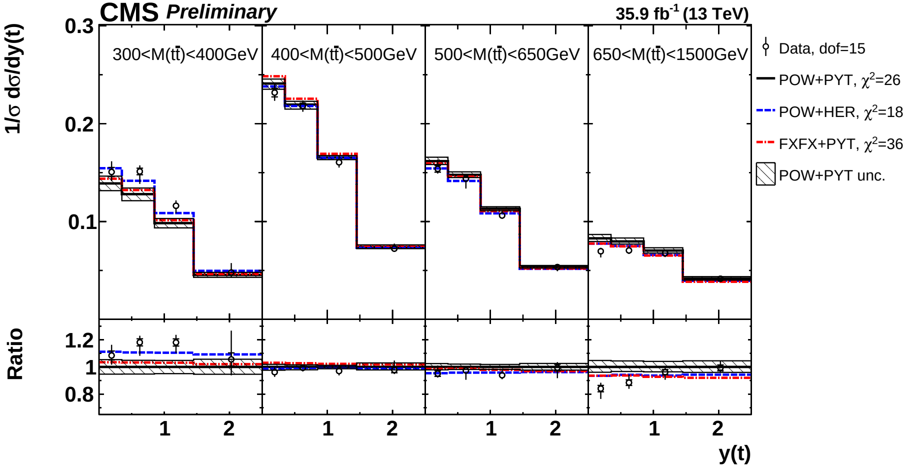
Model uncertainties:

- based on weights:
  - ME scales (envelope of 6 variations dominated by simultaneous μ_r,μ_f var.)
  - PDFs and α_s (CT14 eigenvectors)
  - b-quark fragmentation (envelope of varied Bowler-Lund and Peterson funct.)
  - b-hadron branching ratios
- based on independent samples:
  - $m_t \pm 1$ GeV (using samples with ± 3 GeV $\rightarrow$ rescaled by $1/3$)
  - $0.996 m_\mathrm{t} < h_\mathrm{damp} < 2.239 m_\mathrm{t}$
  - ISR μ, FSR μ variations (latter rescaled by $1/\sqrt{2}$)
  - color reconnection: envelope of 3 samples with different tunes
  - underlying event tune variation

MC predictions

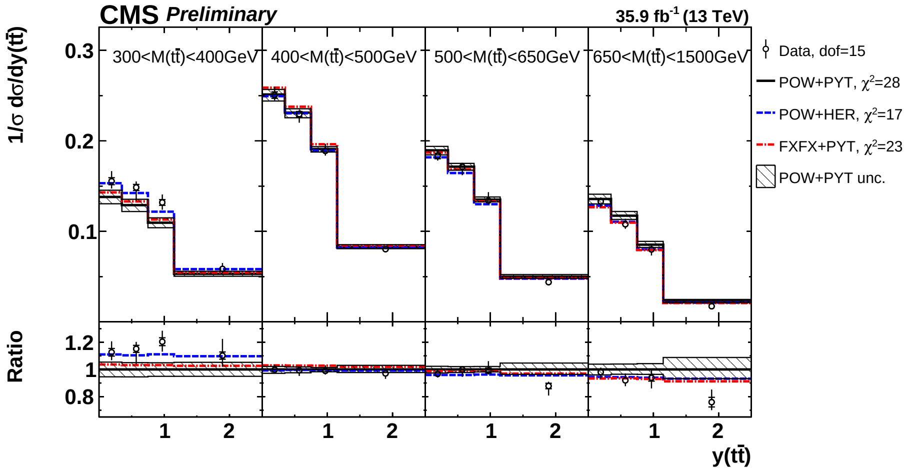
- POWHEGV2 + PYTHIA8
  - $h_{\mathrm{damp}} = 1.581\,m_t$
  - $m_t = 172.5$ GeV
  - CUETP8M2T4 tune [CMS-PAS-TOP-16-021]
- POWHEGV2 + HERWIG++
  - $h_{\mathrm{damp}} = 1.581\,m_t$
  - $m_t = 172.5$ GeV
  - EE5C tune [JHEP10 (2013) 113]
- MG5_aMC@NLO + PYTHIA8
  - FxFx prescription for $t\bar{t}$, $t\bar{t}$ +1 jet, $t\bar{t}$ +2 jets @ NLO [JHEP12 (2012) 061]
  - $m_t = 172.5$ GeV
  - CUETP8M2T4 tune [CMS-PAS-TOP-16-021]

Results: 2D x-sections $[M(t\bar{t}), y(t)]$



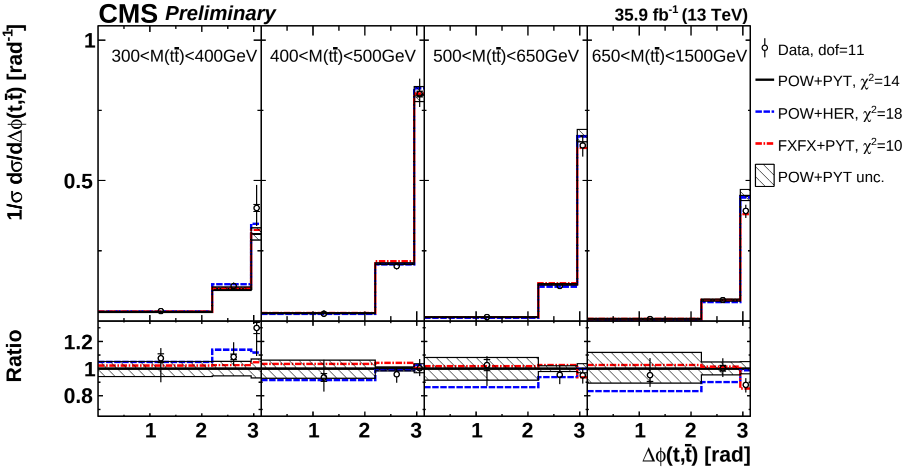
- → MC is more central than data at largest $M(t\bar{t})$
- → best description by 'POW-HER' (mainly $M(t\bar{t})$ slope)

Results: 2D cross sections $[M(t\bar{t}),y(t\bar{t})]$



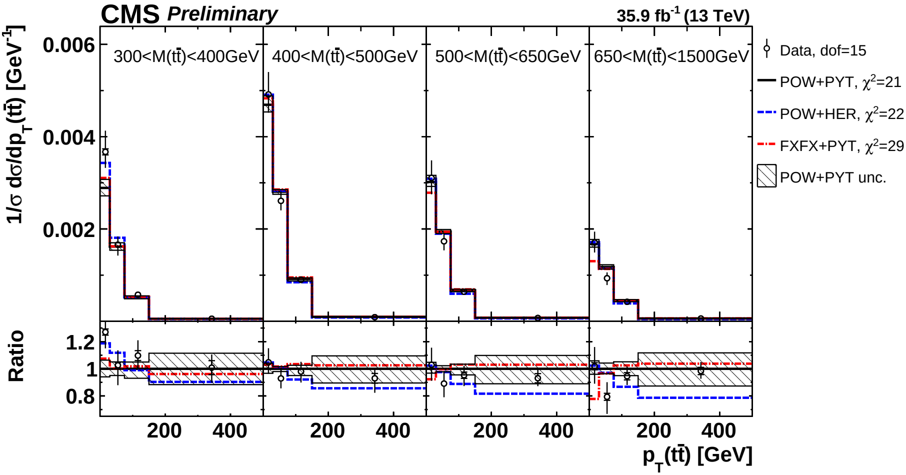
- MC is (somewhat) less central than data at largest $M(t\bar{t})$
- best description by ‘POW-HER’ (mainly $M(t\bar{t})$ slope)

Results: 2D x-sections $[M(t\bar{t}),\Delta\phi(t,\bar{t})]$



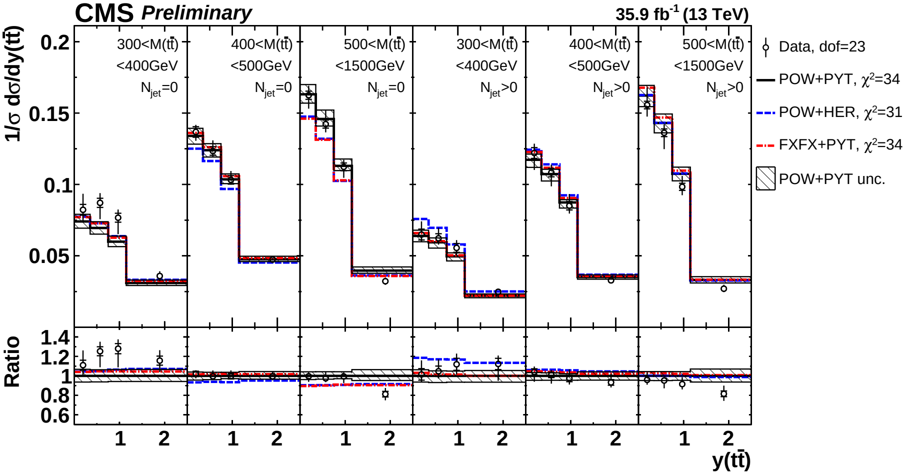
→ all MC describe data well

Results: 2D x-sections $[M(t\bar{t}), p_T(t\bar{t})]$



→ all MC describe data well, but 'FXFX-PYT' predicts too hard $p_T(t\bar{t})$ at highest $M(t\bar{t})$

Results: 3D x-sections $[N_{\mathbf{jet}}^{0,1+}, M(t\bar{t}), y(t\bar{t})]$



→ all MC describe data well

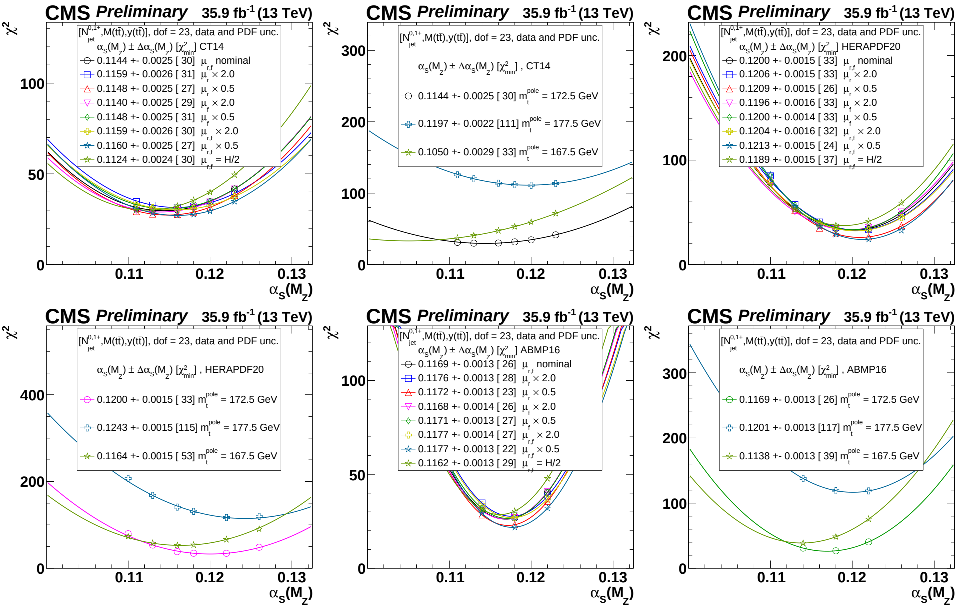
List of discovered features/bugs in MadGraph5_aMC@NLO

- https://bugs.launchpad.net/mg5amcnlo/+bug/1737367
- https://bugs.launchpad.net/mg5amcnlo/+bug/1737368
- https://bugs.launchpad.net/mg5amcnlo/+bug/1752981
- https://bugs.launchpad.net/mg5amcnlo/+bug/1758683
- + many more features and improvements were just implemented locally to provide smooth running on a cluster

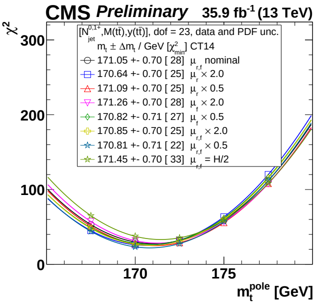
Definition of extra jets (not from top decay)

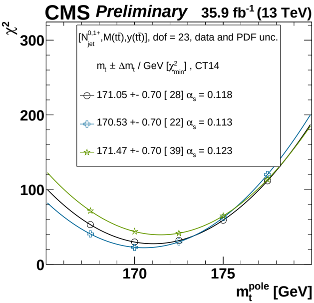
- NLO predictions for inclusive $t\bar{t}$, $t\bar{t}$ + 1 jet and $t\bar{t}$ + 2 jets computed and compared to data using MadGraph5_aMC@NLO + aMCfast + ApplGrid + xFitter
- particle-level jet definition used in measurement, further corrected to parton level using separate MC POWHEGV2 + PYTHIA8 simulations
  - $p_T(j) > 30$ GeV, $|\eta(j)| < 2.4$
  - 'Particle level': particle jets (no ν) required to be isolated within $\Delta R > 0.4$ from l and b from $t\bar{t}$
  - Parton level: standalone POWHEGV2 + PYTHIA8 generated without
    - **(1)** top decays: $C_{\mathrm{def}} = \sigma_{\text{no l,b from } t\bar{t}} \,/\, \sigma_{\text{no } t\bar{t}}$
    - **(2)** hadronisation: $C_{\mathrm{had}} = \sigma_{\text{with had.}} \,/\, \sigma_{\text{no had.}}$
    - **(3)** MPI: $C_{\mathrm{MPI}} = \sigma_{\mathrm{with\ MPI}} \,/\, \sigma_{\mathrm{no\ MPI}}$
- $C_{\mathrm{NP}} = \sigma_{\text{no l,b from } t\bar{t}} / \sigma_{\text{no } t\bar{t}\text{,had.,MPI}}$ $[C_{\mathrm{NP}} \approx C_{\mathrm{def}} \times C_{\mathrm{had}} \times C_{\mathrm{MPI}}]$
- theoretical predictions = $\mathrm{NLO} \times \mathrm{C_{NP}}$
- similar procedure used in jet measurements (although without excluding decay products)

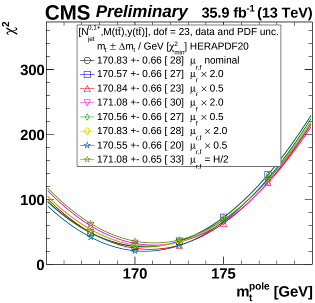
Impact of scale variations on α_s extraction

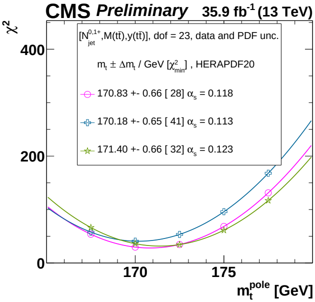


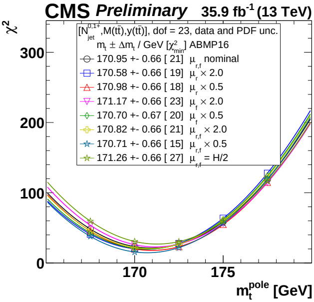
Impact of scale variations on m_t^{pole} extraction

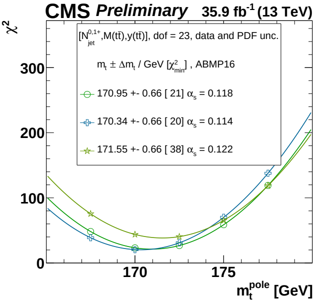




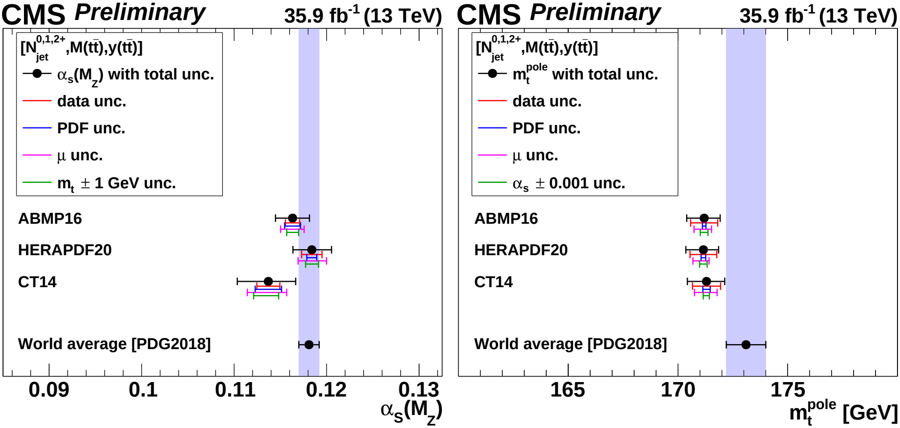




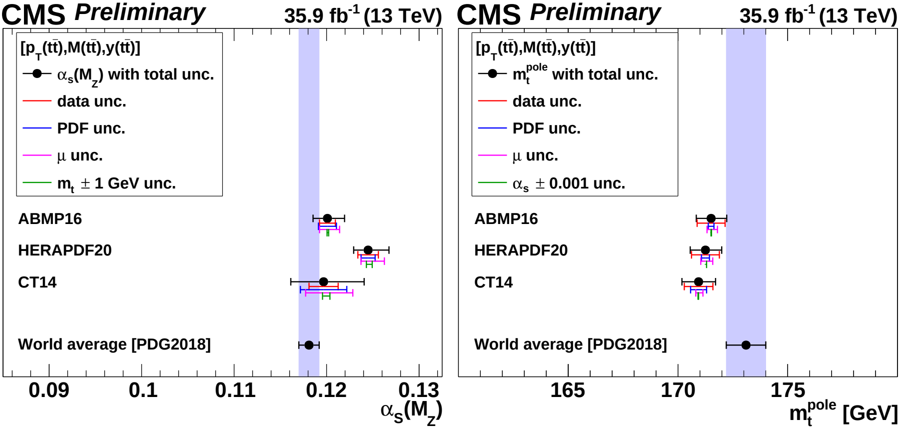




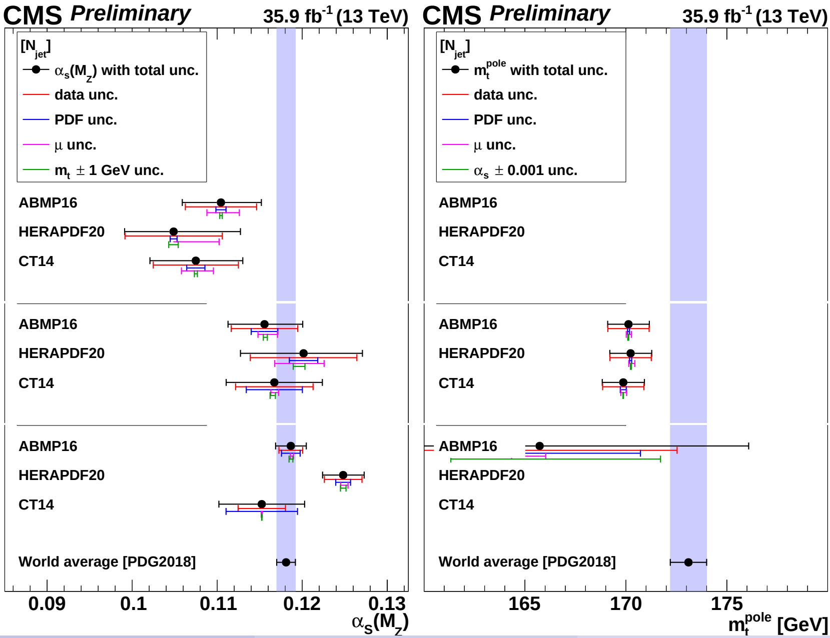
α_s and m_t^{pole} from $[N_{\text{jet}}, mtt, y(t\bar{t})]$ with 3 N_{jet} bins



α_s and m_t^{pole} from $[p_T(t\bar{t}), mtt, y(t\bar{t})]$ with 2 $p_T(t\bar{t})$ bins



α_s and m_t^{pole} from single-differential cross sections



Simultaneous PDF$+\alpha_s + m_t^{\text{pole}}$ fit: settings

- followed standard approach: using HERA DIS data only, or HERA + $t\bar{t}$ data to demonstrate added value from $t\bar{t}$ on PDF and α_s determination
- settings follow HERAPDF2.0 fit (very similar to TOP-14-013), use xFitter-2.0.0
- input data: combined HERA DIS [1506.06042] + $t\bar{t}$
- RTOPT, $M_c = 1.47$ GeV, $M_b = 4.5$ GeV, $Q^2_{\text{min}} = 3.5^{+1.5}_{-1.0}$ GeV^2
- predictions fot $t\bar{t}$ data via MadGraph5_aMC@NLO + aMCfast + ApplGrid, $\mu_r = \mu_f = H_t/4, H_t = \sqrt{m_t^2 + (p_T(t))^2} + \sqrt{m_t^2 + (p_T(\bar{t}))^2}$ varied by factor 2
  - dependence on α_s and scales written in ApplGrid tables
  - dependence on m_t^{pole}derived by linear interpolation between tables generated with different values of m_t^{pole}(new feature for xFitter)
  - kinematic range probed by $t\bar{t}$: $x = (M(t\bar{t})/\sqrt{s}) \exp\left[\pm y(t\bar{t})\right] \Rightarrow 0.01 \lesssim x \lesssim 0.1$
- 15-parameter form (backup) determined using parametrisation scan (one extra g parameter required by $t\bar{t}$ data) at $Q_0^2 = 1.9$ GeV^2, $f_s = 0.4 \pm 0.1$
- DGLAP NLO PDF evolution via QCDNUM-17.01.14
- PDF uncertainties: fit ($\Delta\chi^2 = 1$ via HESSE, cross checked with MC replica method), model and parametrisation; in addition for α_s and m_t^{pole}scale uncertainties for $t\bar{t}$ are considered

Simultaneous PDF, α_s and m_t^{pole} fit: PDF parametrisation

Determined using parametrisation scan:

$$
\begin{aligned}
x g(x) &= A_g x^{B_g} (1-x)^{C_g} (1 + E_g x^2) - A'_g x^{B'_g} (1-x)^{C'_g}, \\
x u_v(x) &= A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + D_{u_v} x), \\
x d_v(x) &= A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}, \\
x\overline{U}(x) &= A_{\overline{U}} x^{B_{\overline{U}}} (1-x)^{C_{\overline{U}}} (1 + D_{\overline{U}} x), \\
x\overline{D}(x) &= A_{\overline{D}} x^{B_{\overline{D}}} (1-x)^{C_{\overline{D}}},
\end{aligned}
$$

- additional gluon parameyter (E_g) required by new $t\bar{t}$ data
- PDF parametrisation uncertainties given by $A'_g = 0$ (13p) and $E_g = 0$ (14p), and $Q_0^2 = 1.9 \pm 0.3$ GeV^2 variation