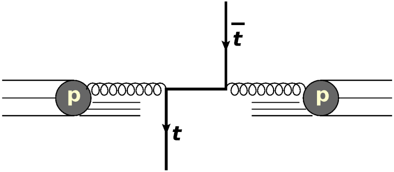
Measurement and QCD analysis of top quark pair double differential cross sections in CMS

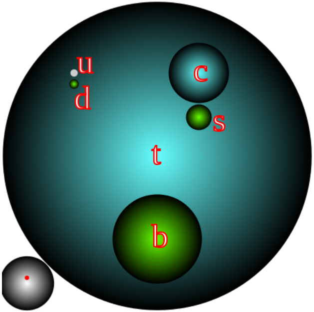
[CMS-PAS-TOP-14-013]

Oleksandr Zenaiev (DESY)

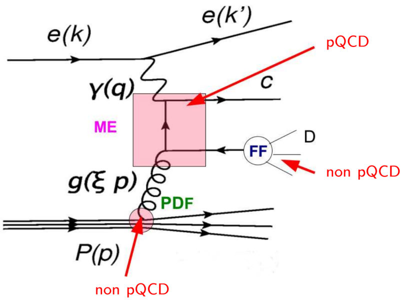
FH Fellow Meeting
29.11.2016

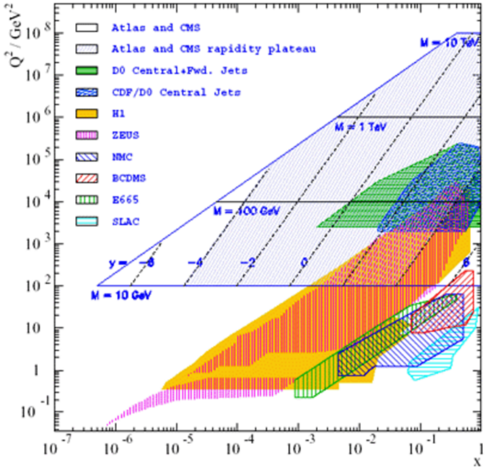
Pedagogical introduction to the physics of HQ hadroproduction: Mangano hep-ph/9711337





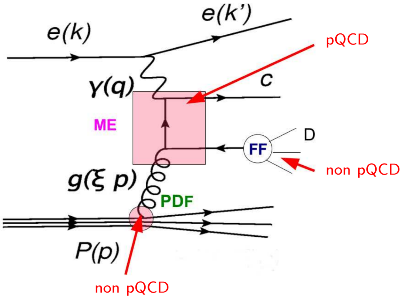
- $m_c, m_b, m_t > \Lambda_{\rm QCD}$ provides a hard scale $\Rightarrow$ pQCD benchmark (LO, NLO, aNNLO, NNLO, ...)
- Different HQ masses ($1.3...170$ GeV) $\Rightarrow$ different impact of radiative corections and non-perturbative effects:
  - e.g. for $m_c \approx 1.3$ GeV fragmentation effects are important
  - while $m_t \approx 172$ GeV is excellent to test pQCD convergence
- Probe of proton structure, e.g. $gg \to t\bar{t}$, $gs \to Wc$ etc.
- Production sensitive to α_s and HQ masses
- May provide insight into possible new physics



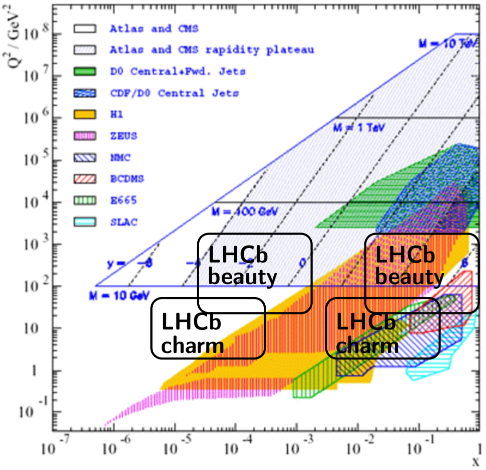


$$\sigma = PDF \otimes ME \otimes FF$$

Measure σ, calculate ME $\Rightarrow$ determine PDF and/or FF



$$\sigma = \color{green}{PDF} \otimes \color{magenta}{ME} \otimes \color{blue}{FF}$$

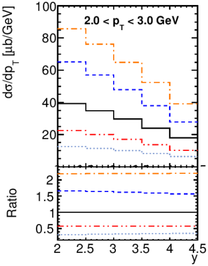


Measure σ, calculate ME $\Rightarrow$ determine PDF and/or FF

Charm production at LHCb $\rightarrow$ gluon at low x $\rightarrow$ atmosphere ν fluxes

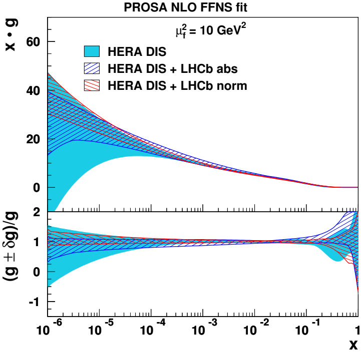
Impact of heavy-flavour production cross sections measured by the LHCb experiment on parton distribution functions at low x

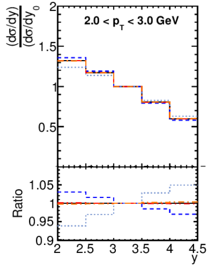
- LHCb measured:
  - charm $0 < p_T < 8$ GeV, $2 < y < 2.5$ [NPB871 (2013) 1]
  - beauty $0 < p_T < 40$ GeV, $2 < y < 2.5$ [JHEP 1308 (2013) 117]
- **PROSA (Proton Structure Analyses in Hadronic Collisions):** QCD analysis of these data
- Reduction of NLO theory uncertainty for y shape
- Improved gluon and sea-quark distributions up to $x \gtrsim 5 \times 10^{-6}$



PROSA,
EPJ C75 (2015) 396

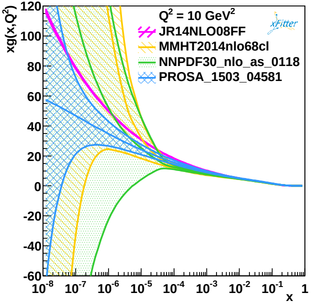
https://prosa.desy.de

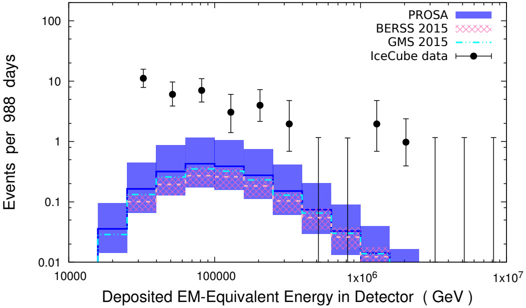




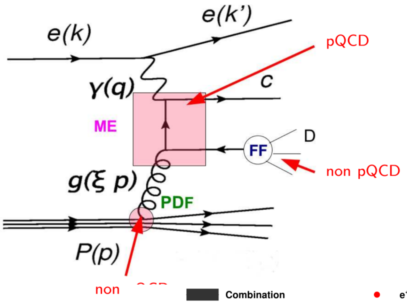
Prompt neutrino fluxes in the atmosphere with PROSA parton distribution functions

- First demonstration how HQ forward hadro-production data can contsrain low-x gluon
  $\Rightarrow$ message for global PDF analyses
- Application: predictions for prompt atmosphere ν fluxes
  $\Rightarrow$ background for very high energy cosmic ν



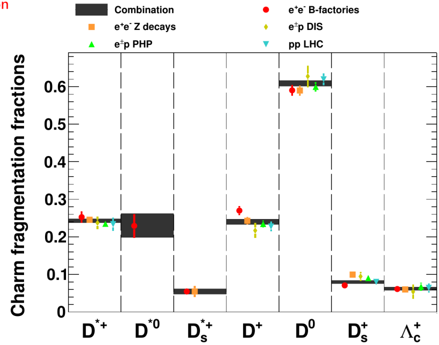


PROSA, arXiv:1611.03815



$$f(c \rightarrow H) = \sigma(H)/\sigma(c)$$

- Fragmentation functions and fragmentation fractions (FF) cannot be calculated perturbatively $\Rightarrow$ have to be extracted from data
- Checked FF universality (e^+e^-, $e^\pm p$, pp)
- FF sum up to 1

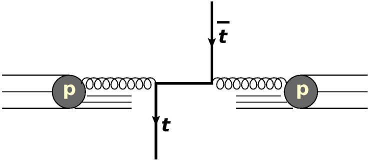


Combined FF have best up-to-date precision

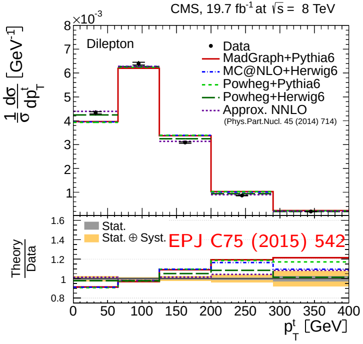
$\Rightarrow$ input for new charm hadroproduction measurements

Lisovyi et al.
EPJ C76 (2016) 397

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

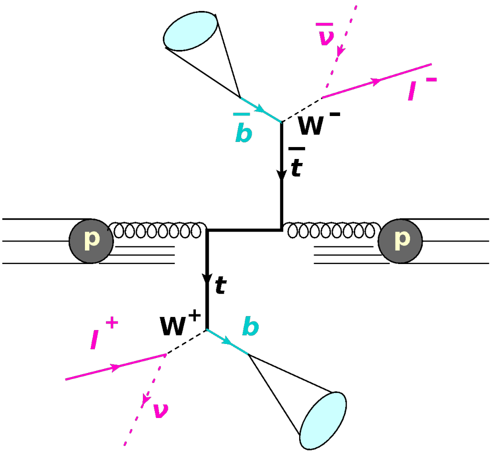


- m_t provides a hard scale
  $\Rightarrow$ ultimate probe of pQCD
  (NLO, aNNLO, NNLO, ...)
- Produced mainly via gg
  $\Rightarrow$ constrain gluon PDF
- Production sensitive to α_s and m_t
- May provide insight into possible new physics



Why measure 2D?

- Previous 1D measurements: overall good agreement, but revealed some trends
- 2D measurement: study production dynamics in more detail
- Especially better PDF sensistivity



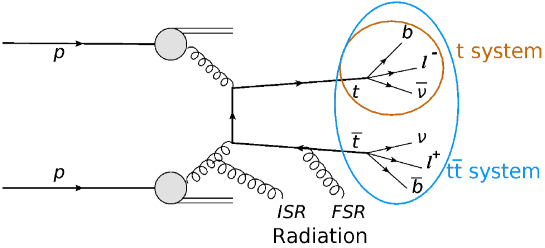
Measurement in $e\mu$ channel:

- **Leptons:**
  - 2 opposite signed e, μ
  - $p_T > 20$ GeV
  - $|\eta| < 2.4$
- **Jets:**
  - at least 2
  - $p_T > 30$ GeV
  - $|\eta| < 2.4$
  - at least 1 b-tagged

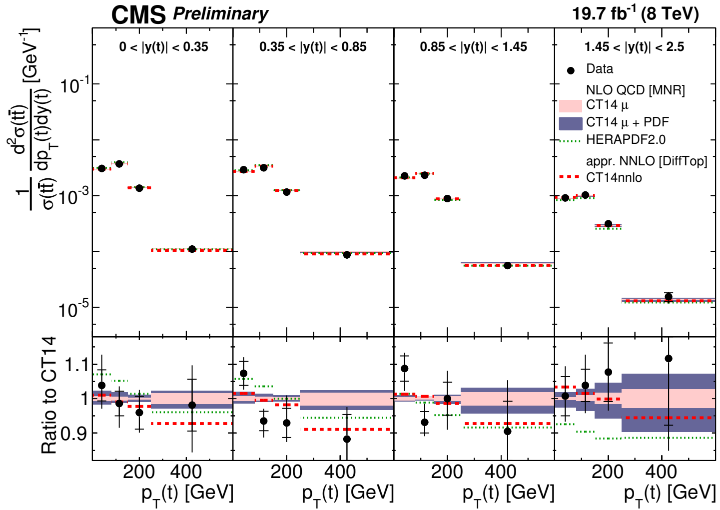
Kinematic reconstruction:

- 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)

(similar to EPJ C75 (2015) 542)



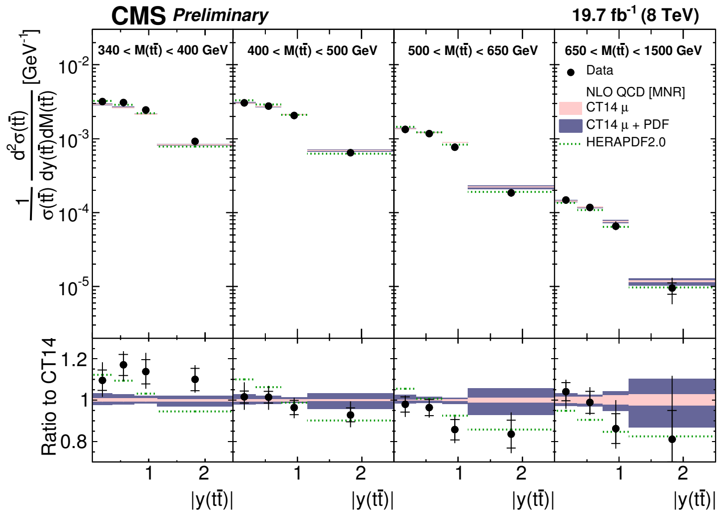
- **t production:**
  - $y(t)$-$p_T(t)$: most simple, aNNLO publicly available
- **$t\bar{t}$ production:**
  - $M(t\bar{t})$-$y(t\bar{t})$: most sensitive to PDFs (at LO $x_{1,2} = \sqrt{\frac{M(t\bar{t})}{s}} e^{\pm y(t\bar{t})}$)
  - $M(t\bar{t})$-$p_T(t\bar{t})$: sensitive to radiation (at LO $p_T(t\bar{t}) \equiv 0$)
- **t, $t\bar{t}$ mixed:**
  - $M(t\bar{t})$-$\Delta\phi(t,\bar{t})$: sensitive to radiation (at LO $\Delta\phi(t\bar{t}) \equiv \pi$)
  - $M(t\bar{t})$-$\Delta\eta(t,\bar{t})$: correlated with $p_T(t)$ as well as sensitive to radiation
  - $M(t\bar{t})$-$y(t)$: sensitive to PDFs (at LO $y(t\bar{t}) = (y(t) + y(\bar{t}))/2$)



- aNNLO provide better data description
- Substantial PDF sensitivity
- Moderate scale uncertainties at NLO (normalised distribution)

	HERA2	CT14	CT14 NNLO
χ^2 (dof = 15)	46	24	13

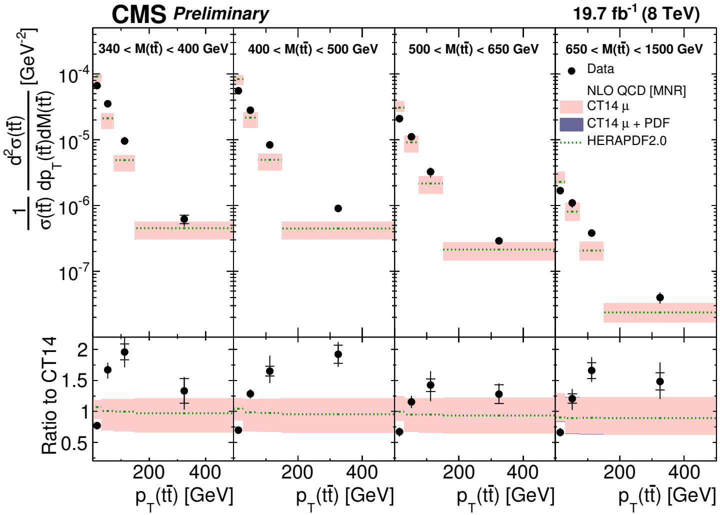
Comparison to fixed-order predictions



- PDF sensitivity exceeds scale uncertainties
- ⇒ can use these data for PDF fits

	HERA2	CT14
χ^2	29	16

(dof = 15)



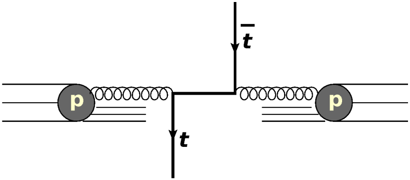
- Large scale uncertainty at NLO (effectively LO for $p_T(t\bar{t})$ shape)
- $\Rightarrow$ useful to test further improvements in pQCD calculations

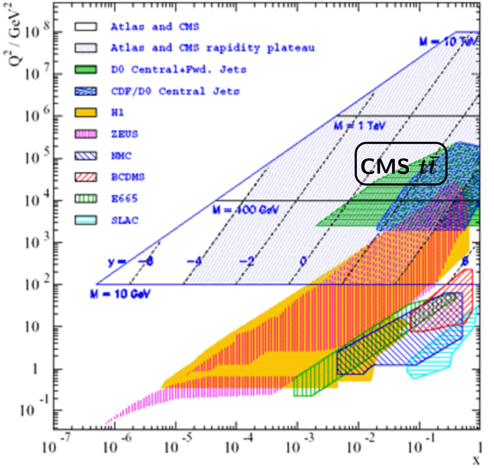
QCD factorisation:

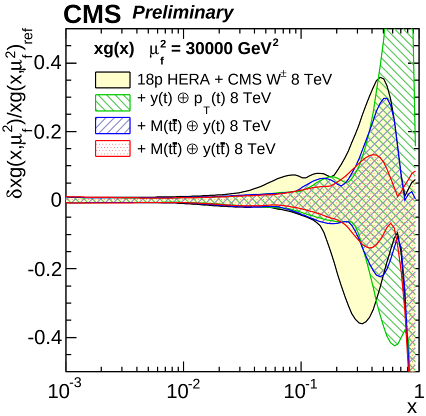
$$\sigma_X = \sum_{a,b} \int dx_1 dx_2 \underline{f_a(x_1,\mu_f^2) f_b(x_2,\mu_f^2)} \hat{\sigma}_{ab\to X}(x_1,x_2,\mu_f^2,\ldots)$$

$\boldsymbol{\mu_f \sim m_t}$

$x_{1,2} = \sqrt{\frac{M(t\bar{t})}{s}} e^{\pm y(t\bar{t})} \quad \Rightarrow \quad \mathbf{0.01 \lesssim x_{1,2} \lesssim 0.25}$





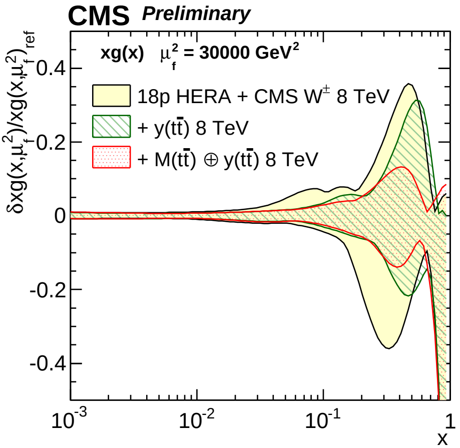


- Using xFitter (former HERAFitter): open-source QCD fit framework [EPJ C75 (2015) 394, xfitter.org]
- Input data:
  - HERA inclusive DIS data [EPJ C75 (2015) 580]
  - CMS W asymmetry [EPJ C76 (2016) 469]
  - these 2D $t\bar{t}$
- NLO calculations for $t\bar{t}$ [NPB373 (1992) 295] using MCFM ⊕ ApplGrid
- Significant improvement of gluon distribution at high x (no other LHC data, e.g. jets)

Best improvement comes from $M(t\bar{t})$–$y(t\bar{t})$:

$$x_{1,2} = \sqrt{\frac{M(t\bar{t})}{s}}\, e^{\pm y(t\bar{t})}$$

- Repeat analysis for 1D distribution $p_T(\mathrm{t})$, $y(\mathrm{t})$, $M(t\bar{t})$, $y(t\bar{t})$
- Check their impact on PDFs



- **2D impact exceeds 1D**
- First study of such kind
- Strongly suggests to use these data in global PDF fits

Summary

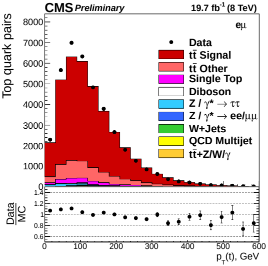
- Measured 2D $t\bar{t}$ normalised cross sections using 8 TeV CMS data
- Quantitative comparison vs MC and fixed-order QCD predictions
- Reasonable overall description, in particular if recent PDFs are used
- PDF fit at NLO using xFitter:
  - 2D $t\bar{t}$ data are well described
  - Best constraints come from $M(t\bar{t}) - y(t\bar{t})$
  - Impact of 2D is better compared to 1D
  - 2D $t\bar{t}$ data are useful as input for global PDF fits
- Interesting to use NNLO calculations when they become available for 2D

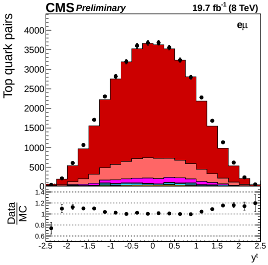
[CMS-PAS-TOP-14-013]

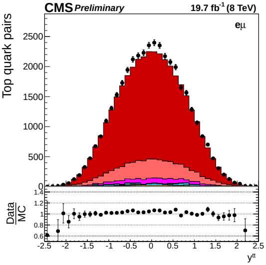
Study of heavy-quark production is an essential ingredient for the understanding of QCD and development of high-energy physics.

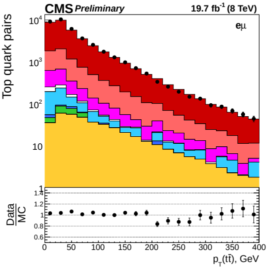
BACKUP

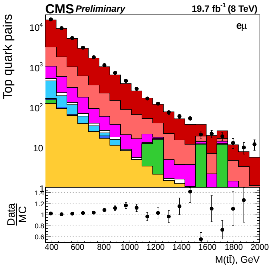
BACKUP. Kinematic distributions





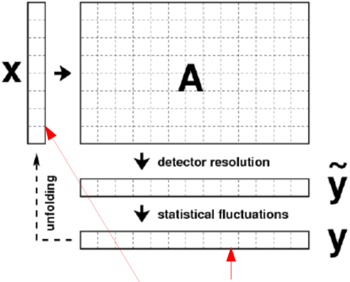






MadGraph + Pythia6 provides overall good description, but:

- $p_T(\mathrm{t})$ harder
- $y(\mathrm{t})$ more central
- $y(t\bar{t})$ less central



`TUnfold` [JINST 7 (2012) T10003]

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

2d distributions are mapped to 1d arrays

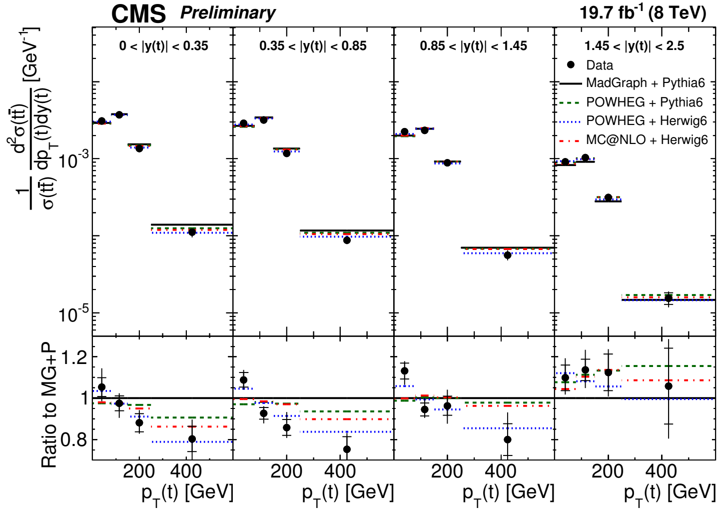
reco. data; unfolded distribution; regularization strength; regularization conditions (second derivative)

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

migration probability matrix; stat. errors of reco.; **gen.** distribution

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

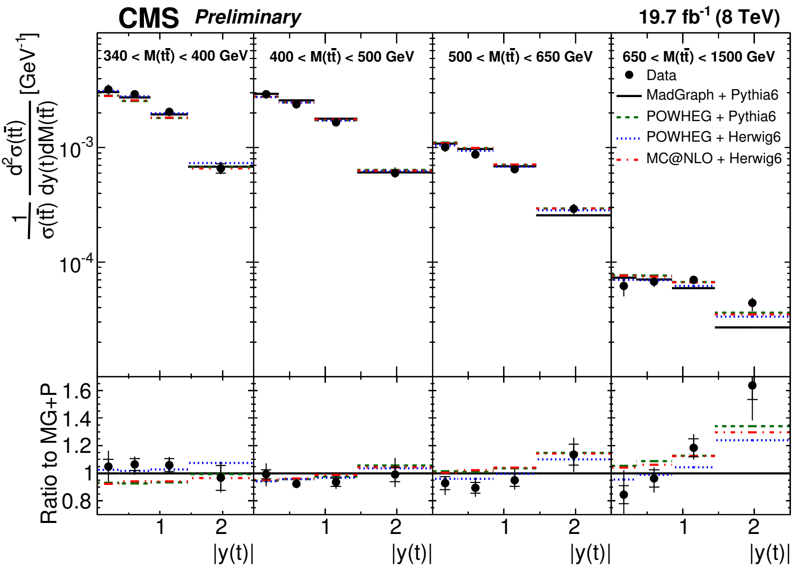
For each Δ**a**^{i}: $\left(\frac{1}{\sigma}\frac{d\sigma}{db}\right)^{ij} = \frac{1}{\sigma} \cdot \frac{X^{ij}}{BR \cdot L \cdot \Delta b^j}$

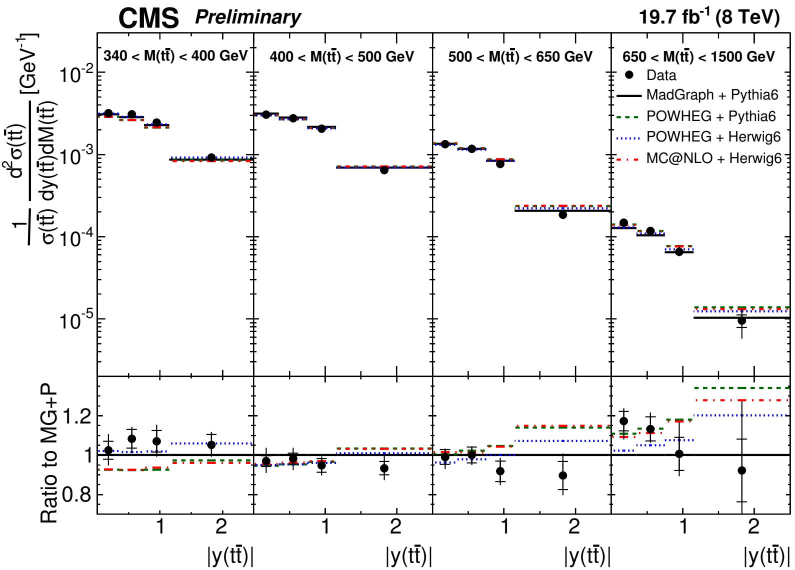


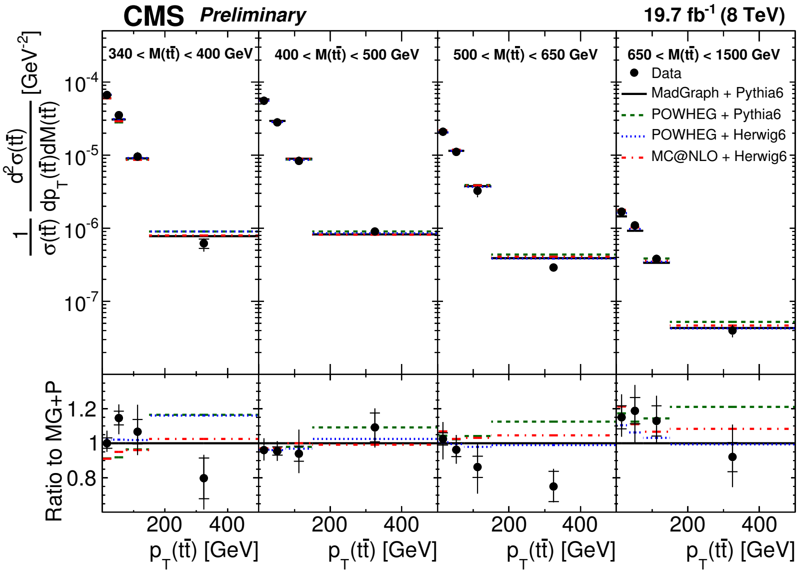
- Harder $p_T(t)$ distribution in all MC
- Trend presents in wide $y(t)$ range

	MP	PP	PH	MH
χ^2	96	58	14	46

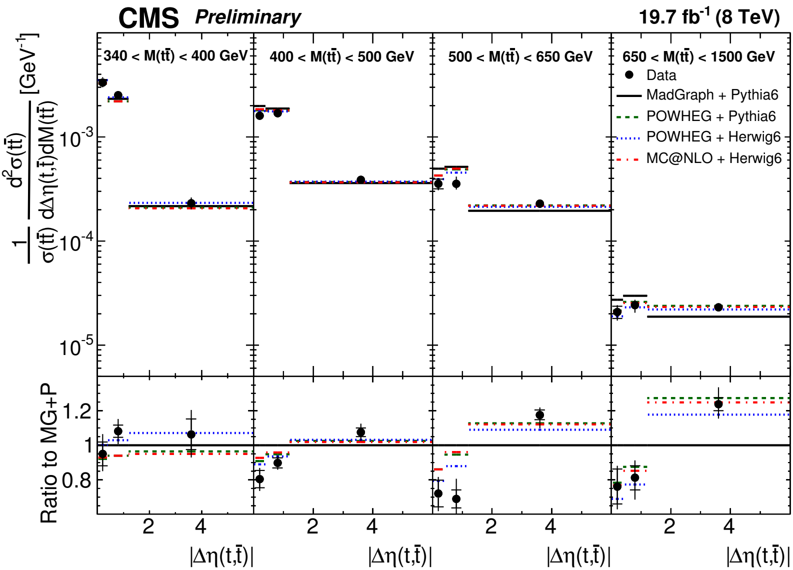
(dof = 15)

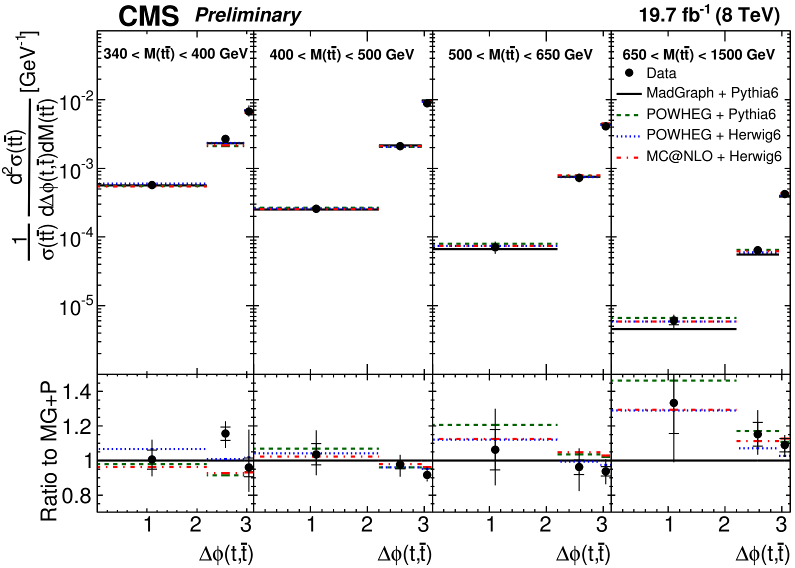

CMS Preliminary
19.7 fb-1 (8 TeV)
340 < M(tt̄) < 400 GeV
400 < M(tt̄) < 500 GeV
500 < M(tt̄) < 650 GeV
650 < M(tt̄) < 1500 GeV
Data
MadGraph + Pythia6
POWHEG + Pythia6
POWHEG + Herwig6
MC@NLO + Herwig6
1/σ(tt̄) d²σ(tt̄)/dy(t)dM(tt̄) [GeV-1]
Ratio to MG+P
|y(t)|

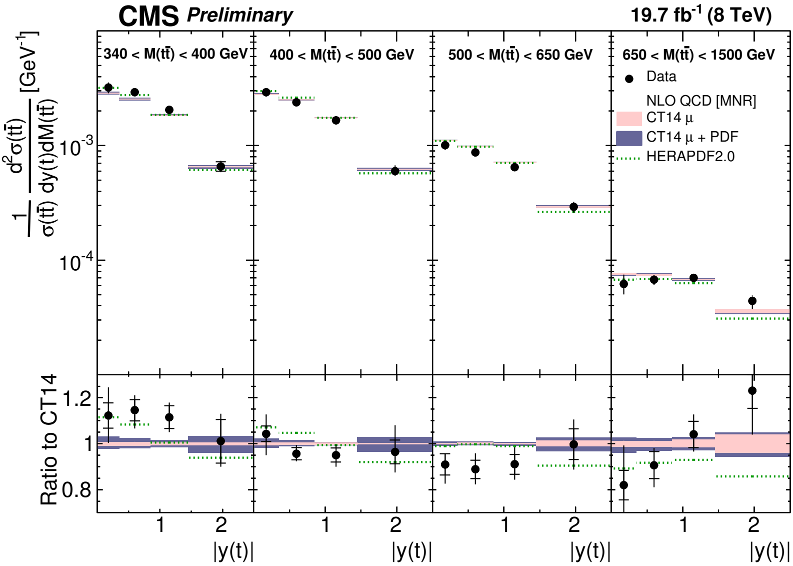

CMS Preliminary
19.7 fb-1 (8 TeV)
340 < M(tt̄) < 400 GeV
400 < M(tt̄) < 500 GeV
500 < M(tt̄) < 650 GeV
650 < M(tt̄) < 1500 GeV
Data
MadGraph + Pythia6
POWHEG + Pythia6
POWHEG + Herwig6
MC@NLO + Herwig6
1/σ(tt̄) d²σ(tt̄)/dy(tt̄)dM(tt̄) [GeV-1]
10-2
10-3
10-4
10-5
Ratio to MG+P
1.2
1
0.8
1
2
|y(tt̄)|

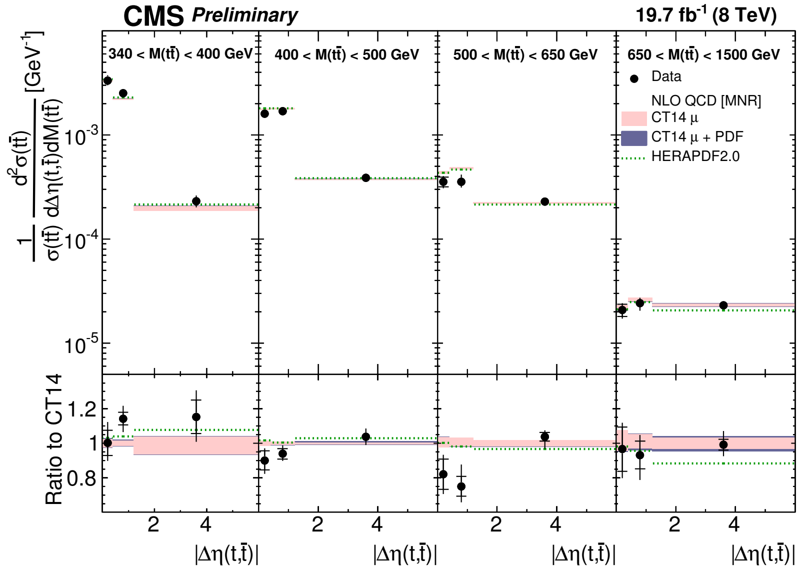

CMS Preliminary
19.7 fb-1 (8 TeV)
340 < M(tt̄) < 400 GeV
400 < M(tt̄) < 500 GeV
500 < M(tt̄) < 650 GeV
650 < M(tt̄) < 1500 GeV
Data
MadGraph + Pythia6
POWHEG + Pythia6
POWHEG + Herwig6
MC@NLO + Herwig6
1/σ(tt̄) d2σ(tt̄)/dpT(tt̄)dM(tt̄) [GeV-2]
10-4
10-5
10-6
10-7
Ratio to MG+P
1.2
1
0.8
200
400
pT(tt̄) [GeV]

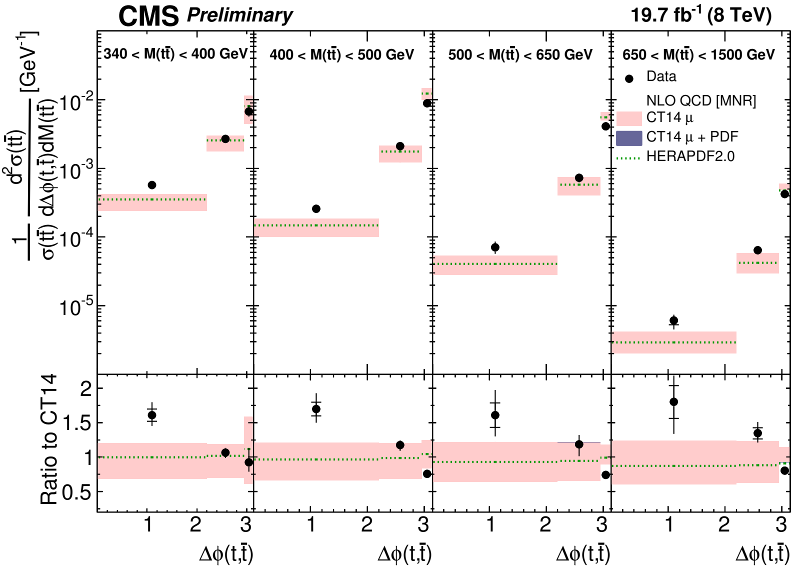
BACKUP. Comparison to MC




CMS Preliminary
19.7 fb^{-1} (8 TeV)
340 < M(tt̄) < 400 GeV
400 < M(tt̄) < 500 GeV
500 < M(tt̄) < 650 GeV
650 < M(tt̄) < 1500 GeV
Data
MadGraph + Pythia6
POWHEG + Pythia6
POWHEG + Herwig6
MC@NLO + Herwig6
$\frac{1}{\sigma(t\bar{t})} \frac{d^2\sigma(t\bar{t})}{d\Delta\phi(t,\bar{t})dM(t\bar{t})}$ [GeV^{-1}]
Ratio to MG+P
$\Delta\phi(t,\bar{t})$


CMS Preliminary
19.7 fb-1 (8 TeV)
340 < M(tt̄) < 400 GeV
400 < M(tt̄) < 500 GeV
500 < M(tt̄) < 650 GeV
650 < M(tt̄) < 1500 GeV
Data
NLO QCD [MNR]
CT14 μ
CT14 μ + PDF
HERAPDF2.0
1/σ(tt̄) d2σ(tt̄)/dy(t)dM(tt̄) [GeV-1]
10-3
10-4
Ratio to CT14
1.2
1
0.8
1
2
|y(t)|


CMS Preliminary
19.7 fb-1 (8 TeV)
340 < M(tt̄) < 400 GeV
400 < M(tt̄) < 500 GeV
500 < M(tt̄) < 650 GeV
650 < M(tt̄) < 1500 GeV
Data
NLO QCD [MNR]
CT14 μ
CT14 μ + PDF
HERAPDF2.0
1/σ(tt̄) d²σ(tt̄)/dΔη(t,t̄)dM(tt̄) [GeV-1]
10-3
10-4
10-5
Ratio to CT14
1.2
1
0.8
2
4
|Δη(t,t̄)|


CMS Preliminary
19.7 fb-1 (8 TeV)
340 < M(tt̄) < 400 GeV
400 < M(tt̄) < 500 GeV
500 < M(tt̄) < 650 GeV
650 < M(tt̄) < 1500 GeV
Data
NLO QCD [MNR]
CT14 μ
CT14 μ + PDF
HERAPDF2.0
Ratio to CT14
Δφ(t,t̄)

BACKUP. Systematics

Following 1D cross section measurement (TOP-12-028):

- pile-up
- lepton selection
- trigger efficiency
- **jet energy scale and resolution ($\lesssim 2\%$)**
- b-tagging efficiency
- kinematic reconstruction efficiency
- background variation:
  - DY varied separately by $\pm 30\%$
  - other backgrounds varied simultaneously by $\pm 30\%$
- model uncertainties ($\lesssim 10\%$):
  - **perturbative scale variation**
  - **matching scale variation**
  - m_t variation
  - PDFs
  - **Hadronization (PowhegHerwig - PowhegPythia)**
  - **Hard scattering (PowhegPythia - MadgraphPythia)**

Distribution	NDoF	MC				NLO nominal (including PDF unc.)						
		MP	PP	PH	MH	HERAPDF2	MMHT14	CT14	NNPDF30	ABM11nf5	JR14	CJ12
$y(t) - p_T(t)$	15	96	58	14	46	46 (40)	26 (24)	24 (21)	28 (25)	62 (51)	47 (47)	24 (23)
$M(t\bar{t}) - y(t)$	15	53	20	13	21	52 (44)	22 (20)	19 (18)	14 (14)	71 (55)	44 (44)	20 (20)
$M(t\bar{t}) - y(t\bar{t})$	15	19	21	15	22	29 (25)	15 (15)	16 (15)	10 (10)	42 (31)	25 (25)	15 (15)
$M(t\bar{t}) - \Delta\eta(t\bar{t})$	11	163	33	20	39	46 (43)	31 (31)	32 (31)	45 (42)	48 (44)	39 (39)	32 (32)
$M(t\bar{t}) - p_T(t\bar{t})$	15	31	83	30	33	485 (429)	377 (310)	379 (264)	251 (212)	553 (426)	428 (415)	382 (378)
$M(t\bar{t}) - \Delta\phi(t\bar{t})$	11	21	21	10	17	354 (336)	293 (272)	296 (259)	148 (143)	386 (335)	329 (324)	297 (295)

Monte Carlo

- MP: MadGraph(CTEQ6) + Pythia6
- PP: Powheg(CT10) + Pythia6
- PH: Powheg(CT10) + Herwig6
- MH: MC@NLO(CTEQ6) + Herwig6

Overall description by MCs:

- best: Powheg + Herwig
- worst: MadGraph + Pythia

NLO

Some comments on PDFs:

- MMHT14, CT14 and NNPDF30: use LHC data
- other PDFs: no LHC data

Overall description by different PDFs:

- best: MMHT14, CT14, NNPDF30, CJ12
- worse: HERAPDF2.0, JR14, ABM11nf5

Particular distributions described by "best" PDFs:

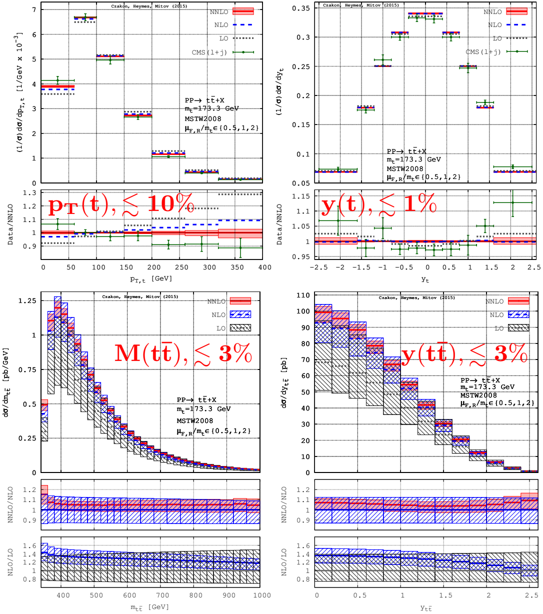
- bad description of $\Delta\phi(t\bar{t})$ and $p_T(t\bar{t})$
- also bad description of $\Delta\eta(t\bar{t})$
- reasonable description of the rest

BACKUP. Framework of QCD analysis

In short: similar to HERAPDF2.0 fit with $t\bar{t}$ data included
(any publicly reproducable PDF fit would serve the purpose)

- Platform: xFitter (former HERAFitter) [www.xfitter.org]
- Input data: HERA $e^{\pm}p$ inclusive data [1506.06042], $Q^2_{\min} > 3.5$ GeV^2 + CMS W asym. 8TeV [1603.01803] + $t\bar{t}$ normalised 2D data
- RT optimal variable-flavour-number scheme with $n_f = 5$, $\alpha_s^{n_f=5}(M_Z) = 0.118$, $M_c = 1.47$ GeV, $M_b = 4.50$ GeV
- Predictions for $t\bar{t}$:
  - MNR calculations [NPB373 (1992) 295] via MCFM $\otimes$ ApplGrid $\otimes$ xFitter
  - pole top mass $m_t = 172.5$ GeV
  - scales $\mu_r^2 = \mu_f^2 = m_t^2 + (p_T^2(t) + p_T^2(\bar{t}))/2$
- PDF parametrisation: 18 free parameters HERAPDF style
- Uncertainties:
  - Experimental: from $\Delta\chi^2 = 1$
  - Model: from theoretical and model parameter variations
  - Parametrisation: from μ^2_{f0} and parameterisation form variation

(pictures taken from arXiv:1511.00549)



Is it OK to use NLO calculations?

Yes, for normalised distributions $\mathbf{p_T(t), y(t), M(t\bar{t}), y(t\bar{t})}$ with current data precision, because:

- These are the only publicly available calculations
- Missing higher order corrections (estimated by scale variation) for NLO are small w.r.t to data:
  - $\simeq 10\%$ fot total x-section
  - $\lesssim 3\%$ for $p_T(t)$, $y(t)$, $M(t\bar{t})$, $y(t\bar{t})$ shapes
  - **Confirmed by exact NNLO calculations** [Czakon et al. arXiv:1511.00549] (slightly larger effect on $p_T(t)$).