Constraints on the gluon PDF from top quark production at hadron colliders

Juan Rojo
Rudolf Peierls Center for Theoretical Physics
University of Oxford

Topical Workshop on Top Quark Differential Distributions
Cannes, 27/09/2014

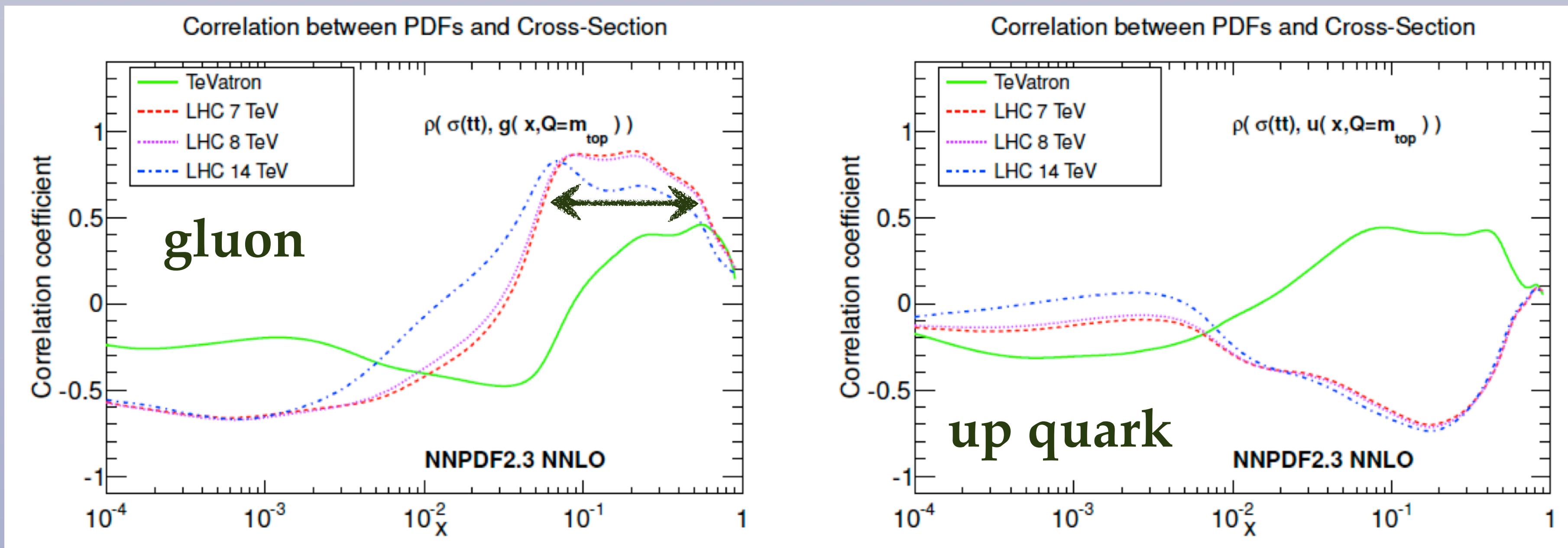
Top quarks as gluon luminometers

- Top quark pair production at the LHC is **directly sensitive to the gluon luminosity**, thus provides a new observable to constrain gluons in **global PDF analysis**

	TeVatron	LHC 7 TeV	LHC 8 TeV	LHC 14 TeV
gg	15.4%	84.8%	86.2%	90.2%
$qg+\bar{q}g$	-1.7%	-1.6%	-1.1%	0.5%
qq	86.3%	16.8%	14.9%	9.3%

Contribution to the NNLO+NNLL cross section from different subprocesses

- Top production probes the **gluon PDF** in the range between **0.1 ≲ x ≲ 0.5**



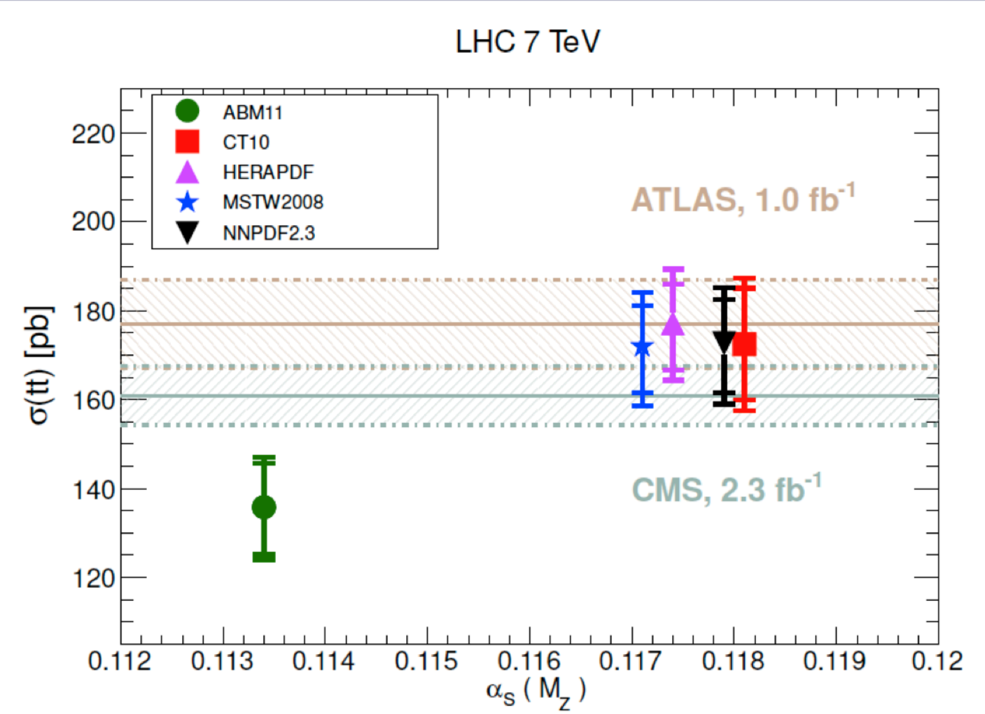
PDF constraints using the inclusive ttbar cross section

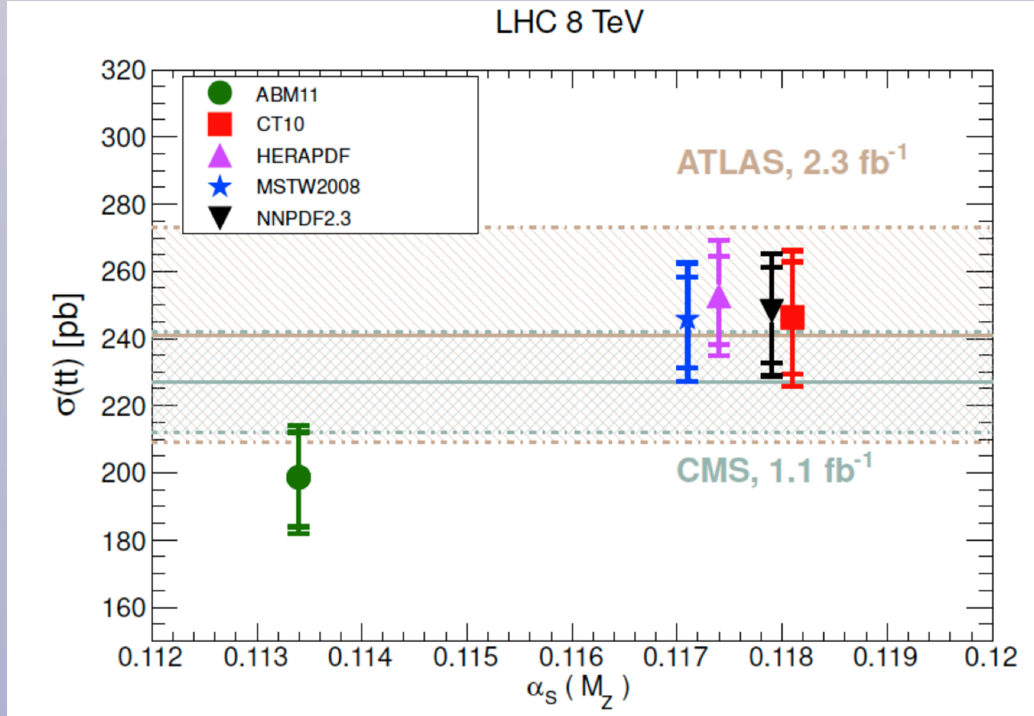
M. Czakon, M. Mangano, A. Mitov and J. Rojo, arXiv:1303.7215

PDF dependence of the top cross section

- Compute predictions at NNLO+NNLL with **top++2.0** for different PDF sets with the associated theoretical uncertainties: **PDFs, m_t ,α_S** and **missing higher orders**
- Top mass fixed to **m_t = 173.3 GeV**. Assume **δm_t = 1 GeV**, and **$\delta\alpha_S$ = 0.007** (*PDG values*)
- Parametric uncertainties (PDFs, m_t ,α_S) added in quadrature, then linearly to scale uncertainty
- Compare to the most precise **ATLAS** and **CMS** 7 and 8 TeV data (as of March 2013)

When available, experimental data corrected to nominal m_t = 173.3 GeV





PDF dependence of the top cross section

- Compute predictions at NNLO+NNLL with **top++2.0** for different PDF sets with the associated theoretical uncertainties: **PDFs, m_t ,α_S** and **missing higher orders**
- The contributions from the different sources of theory uncertainty are similar

$$\delta_{scale} \approx 2.5 - 3.5\% \;,\quad \delta_{PDF} \approx 3.0 - 4.5\% \;,\quad \delta_{\alpha s} \approx 1.5 - 2.2\% \;,\quad \delta_{mt} \approx 3.0\%$$

$$\delta_{tot} \approx 7.0 - 8.5\%$$

LHC 7 TeV						
PDF set	σ_{tt} (pb)	$\delta_{\rm scale}$ (pb)	$\delta_{\rm PDF}$ (pb)	δ_{α_s} (pb)	δ_{m_t} (pb)	$\delta_{\rm tot}$ (pb)
ABM11	135.8	$^{+3.5}_{-4.2}$ $(^{+2.6\%}_{-3.1\%})$	$^{+6.4}_{-6.4}$ $(^{+4.7\%}_{-4.7\%})$	$^{+0.0}_{-0.0}$ $(^{+0.0\%}_{-0.0\%})$	$^{+4.3}_{-4.2}$ $(^{+3.2\%}_{-3.1\%})$	$^{+11.2}_{-11.8}$ $(^{+8.2\%}_{-8.7\%})$
CT10	172.5	$^{+4.6}_{-6.0}$ $(^{+2.7\%}_{-3.5\%})$	$^{+8.0}_{-6.5}$ $(^{+4.6\%}_{-3.8\%})$	$^{+3.7}_{-3.7}$ $(^{+2.2\%}_{-2.2\%})$	$^{+5.3}_{-5.1}$ $(^{+3.1\%}_{-3.0\%})$	$^{+14.9}_{-15.0}$ $(^{+8.6\%}_{-8.7\%})$
HERA1.5	177.2	$^{+4.8}_{-4.2}$ $(^{+2.7\%}_{-2.3\%})$	$^{+4.0}_{-6.4}$ $(^{+2.3\%}_{-3.6\%})$	$^{+3.0}_{-3.0}$ $(^{+1.7\%}_{-1.7\%})$	$^{+5.4}_{-5.2}$ $(^{+3.1\%}_{-2.9\%})$	$^{+12.2}_{-12.9}$ $(^{+6.9\%}_{-7.3\%})$
MSTW08	172.0	$^{+4.4}_{-5.8}$ $(^{+2.6\%}_{-3.4\%})$	$^{+4.7}_{-4.7}$ $(^{+2.7\%}_{-2.7\%})$	$^{+2.9}_{-2.9}$ $(^{+1.7\%}_{-1.7\%})$	$^{+5.3}_{-5.1}$ $(^{+3.1\%}_{-3.0\%})$	$^{+12.1}_{-13.4}$ $(^{+7.0\%}_{-7.8\%})$
NNPDF2.3	172.7	$^{+4.6}_{-6.0}$ $(^{+2.7\%}_{-3.5\%})$	$^{+5.2}_{-5.2}$ $(^{+3.0\%}_{-3.0\%})$	$^{+2.7}_{-2.7}$ $(^{+1.6\%}_{-1.6\%})$	$^{+5.3}_{-5.2}$ $(^{+3.1\%}_{-3.0\%})$	$^{+12.5}_{-13.7}$ $(^{+7.2\%}_{-8.0\%})$

Not a single factor limits the accuracy of the theory prediction

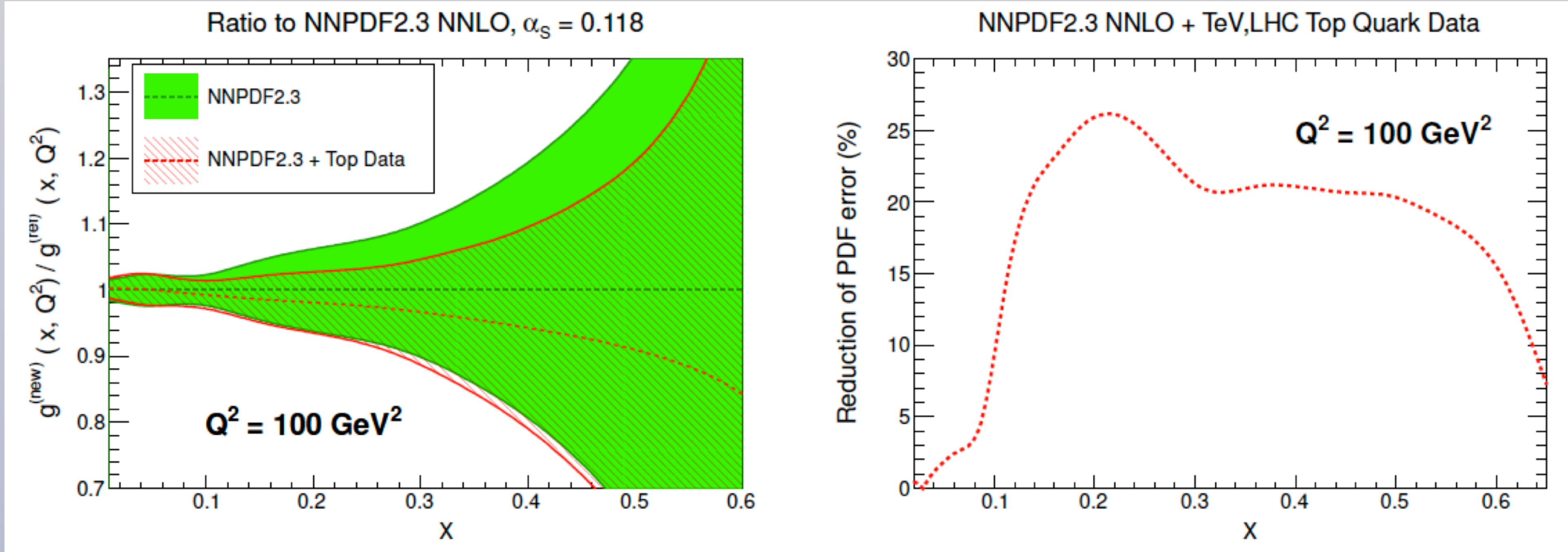
***Scale*, *PDF* and *top mass* uncertainties all of the similar order**

Similar conclusions for other LHC energies

(*) For ABM11, $\delta_{\alpha s}$ included in δ_{PDF}

Pinning down the gluon with top data

- Top quark cross-section data **discriminates between PDF sets**
- In addition, it can also be used to **reduce the PDF uncertainties** within a single PDF set
- We included the most precise top quark data into the **NNPDF2.3** global PDF analysis



- Top quark cross-section data **reduces the PDF uncertainty** in the **large-x gluon** by up to **20%**
- The impact is restricted to the region between **0.1<x<0.5**, where the correlation between the gluon and the top cross section is most significant

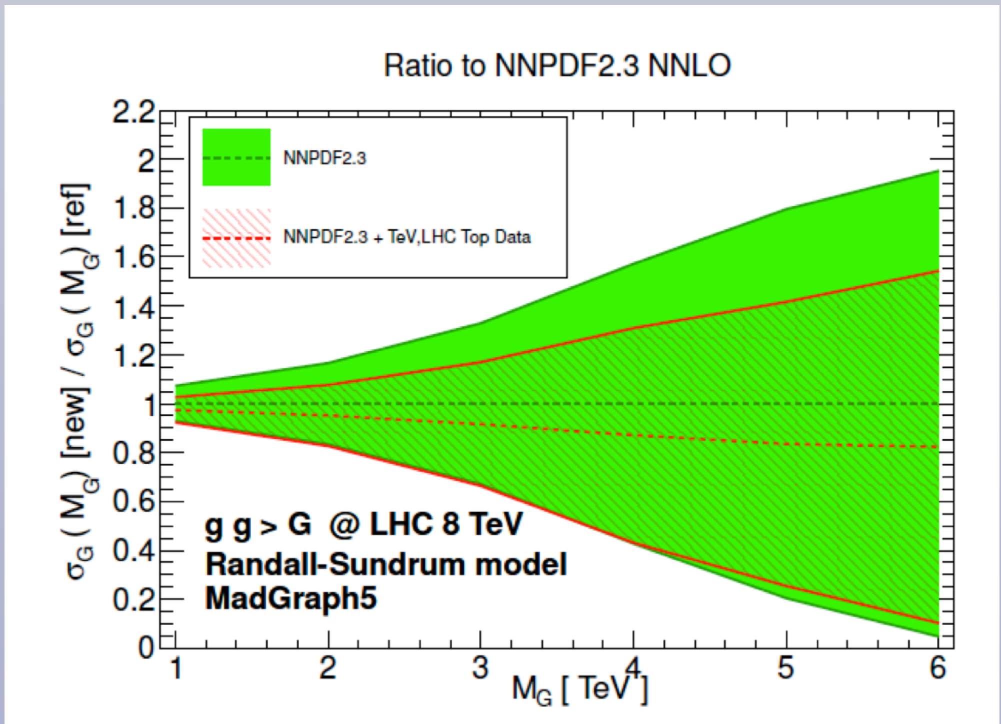
Impact in BSM searches

- Thanks to the top quark data, the **smaller large-x gluon PDF uncertainties** improve the theory predictions for **gluon-initiated BSM processes**

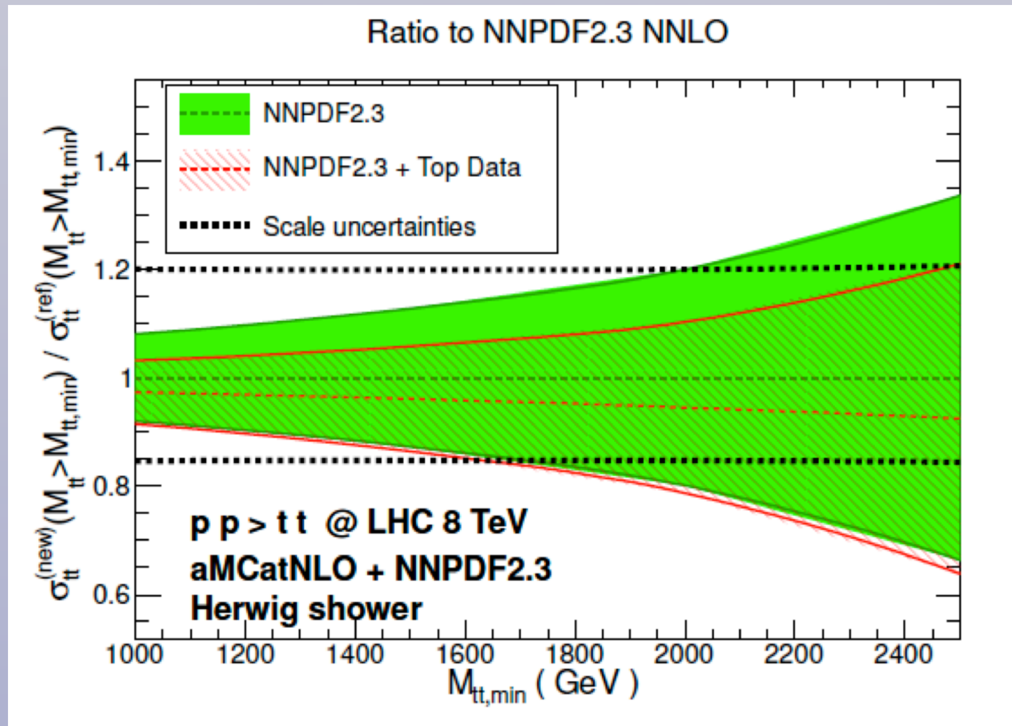
- *Example 1*: reduction of PDF errors for **high mass Graviton production** in warped extra dimensions scenarios

- *Example 2*: reduction of PDF errors in the **high mass tail of the M_{tt} distribution**, used for searches of **resonances decaying into top pairs**

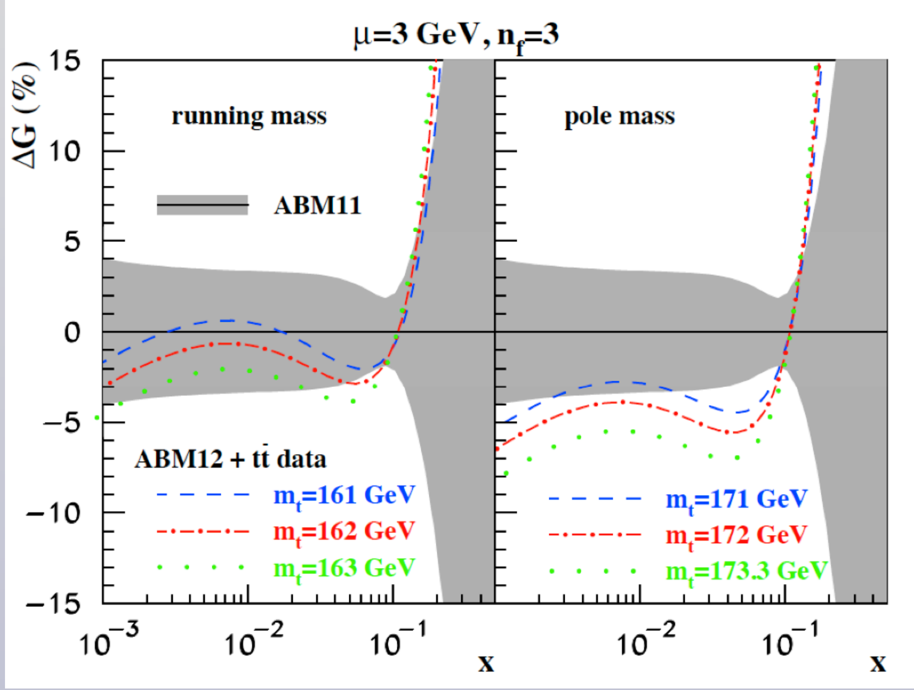
High mass KK graviton production



High mass tail of the M_{tt} distribution

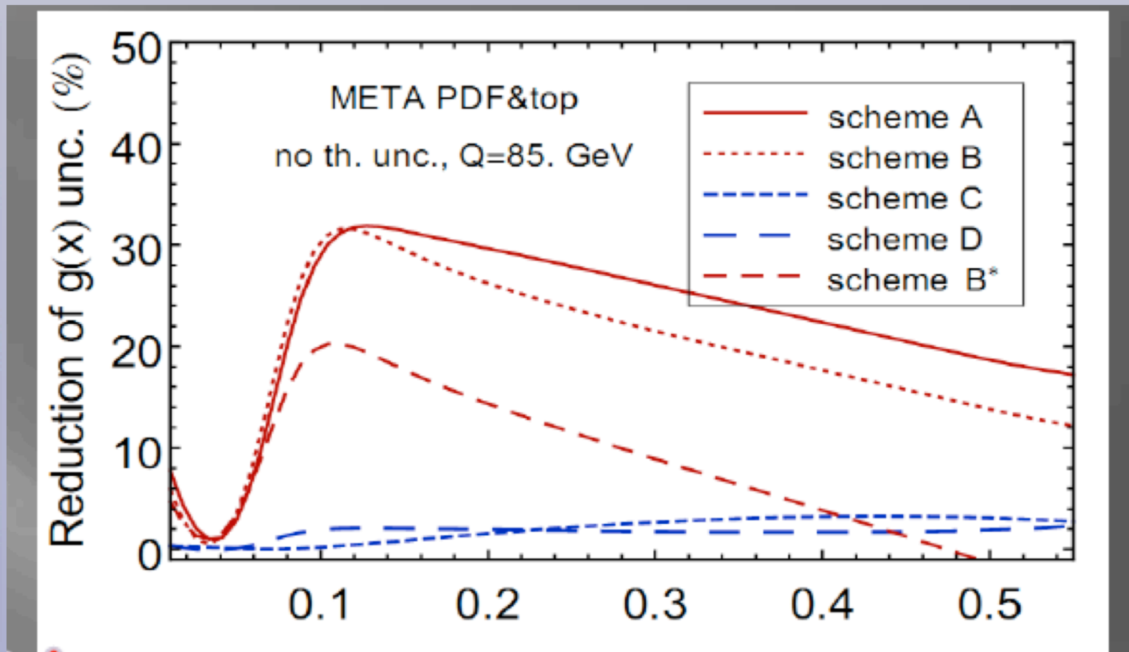


Top quark data in other PDF set



- **ABM13:** substantial impact, for x>0.2 gluon with ttbar **outside the one-sigma PDF uncertainty band**

Jun Gao, PDF4LHC, Dec13



- **CT with MetaPDFs:** Similar quantitative conclusions for reduction of gluon PDF errors as in Czakon et al. (provided same statistical treatment)
- The impact of ttbar total cross-section data has thus been **validated independently** by different groups
- LHC ttbar data included also in **NNPDF3.0** and **MMHT14**

High-precision tests of QCD and BSM with cross-section ratios

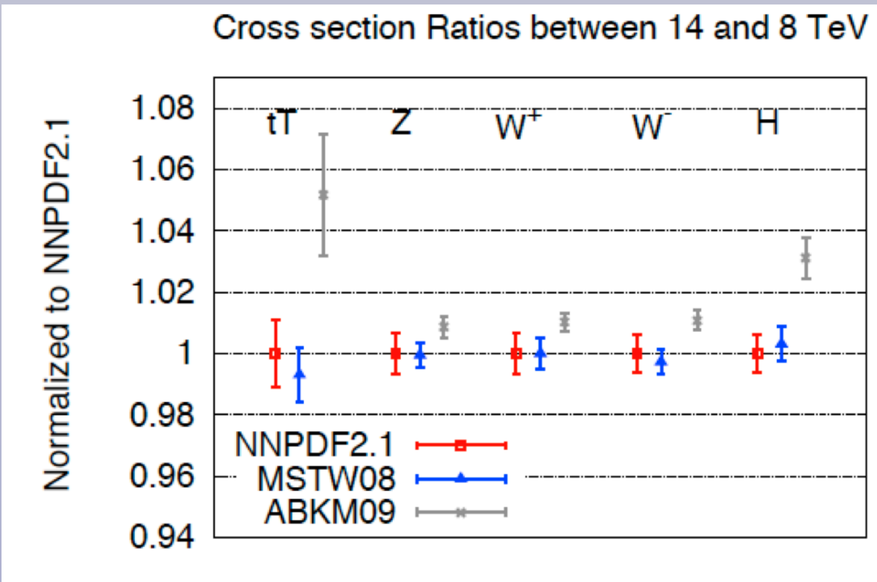
M. Mangano and J. Rojo, arXiv:1206.3557

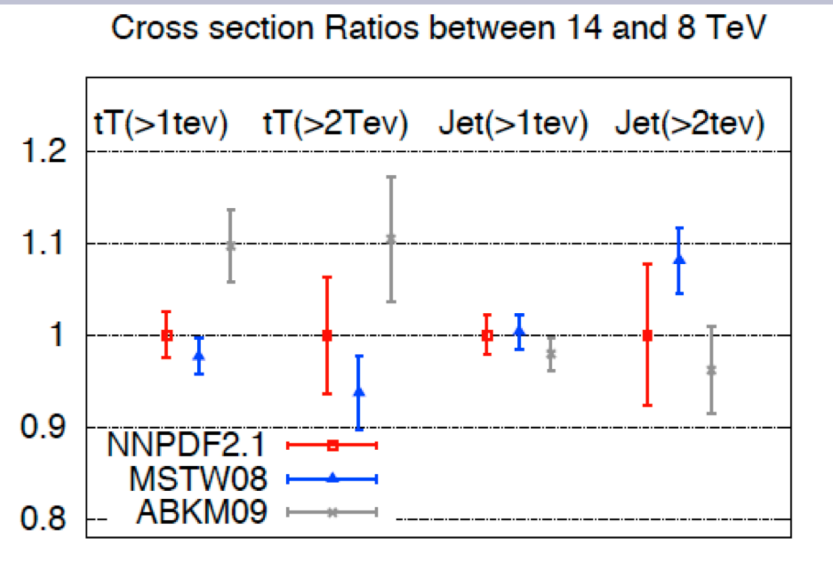
Cross section Ratios between 7, 8 and 14 TeV

- The **staged increase of the LHC beam energy** provides a new class of interesting observables: **cross section ratios** for different beam energies

$$R_{E_2/E_1}(X) \equiv \frac{\sigma(X,E_2)}{\sigma(X,E_1)} \qquad R_{E_2/E_1}(X,Y) \equiv \frac{\sigma(X,E_2)/\sigma(Y,E_2)}{\sigma(X,E_1)/\sigma(Y,E_1)}$$

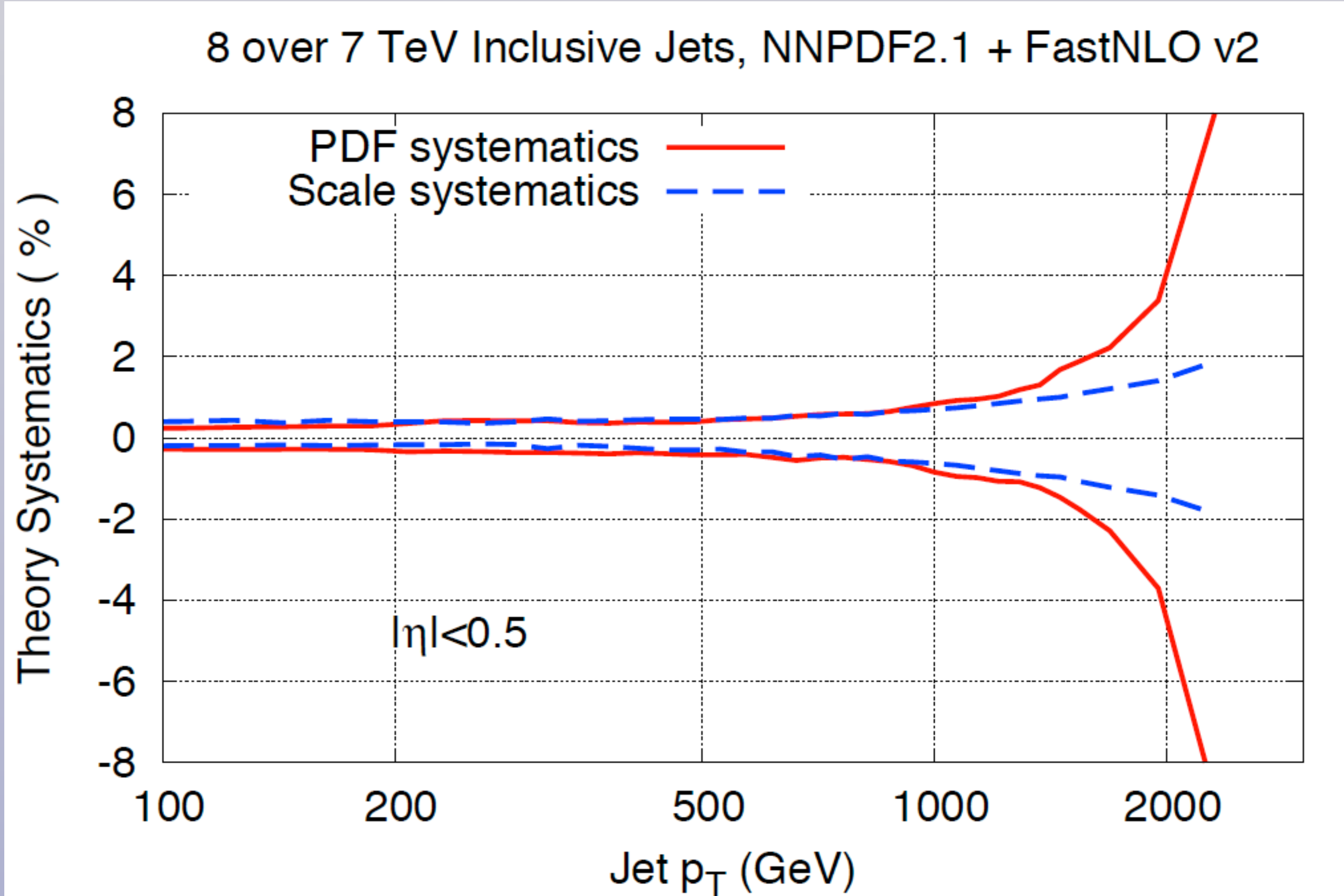
- These ratios can be computed with **very high precision** due to the large degree of **correlation of theoretical uncertainties** at different energies
- **Experimentally** these ratios can also be measured accurately since many systematics, like luminosity or jet energy scale, **cancel partially in the ratios**
- These ratios allow **stringent precision tests of the SM**, like **PDF discrimination**





Cross section Ratios between 7, 8 and 14 TeV

- For inclusive jets, the 8/7 ratio brings **theory systematics below the percent level**
- Above 1 TeV, PDF systematics larger than scale systematics: **opportunity for PDF constraints**
- Greatly extended lever arm in the **14 TeV / 8 TeV ratios**

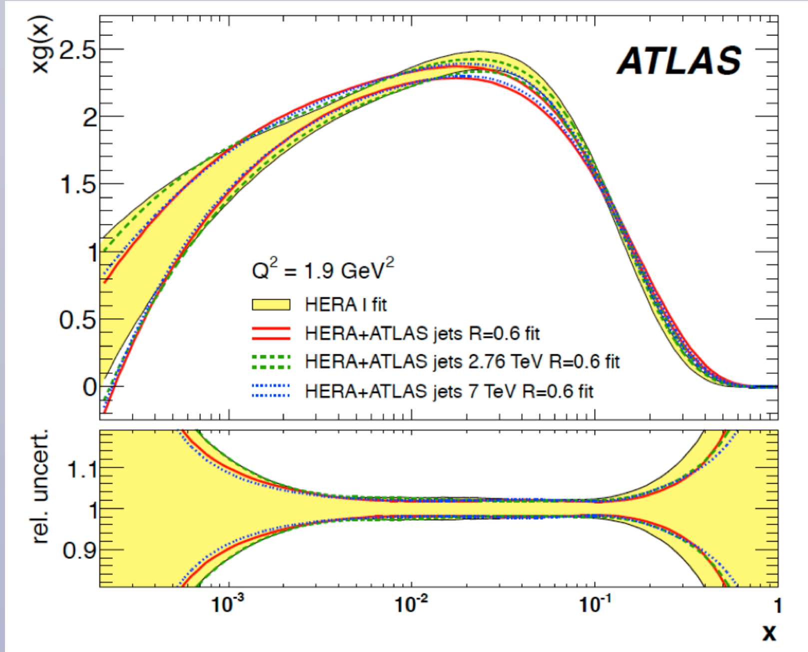


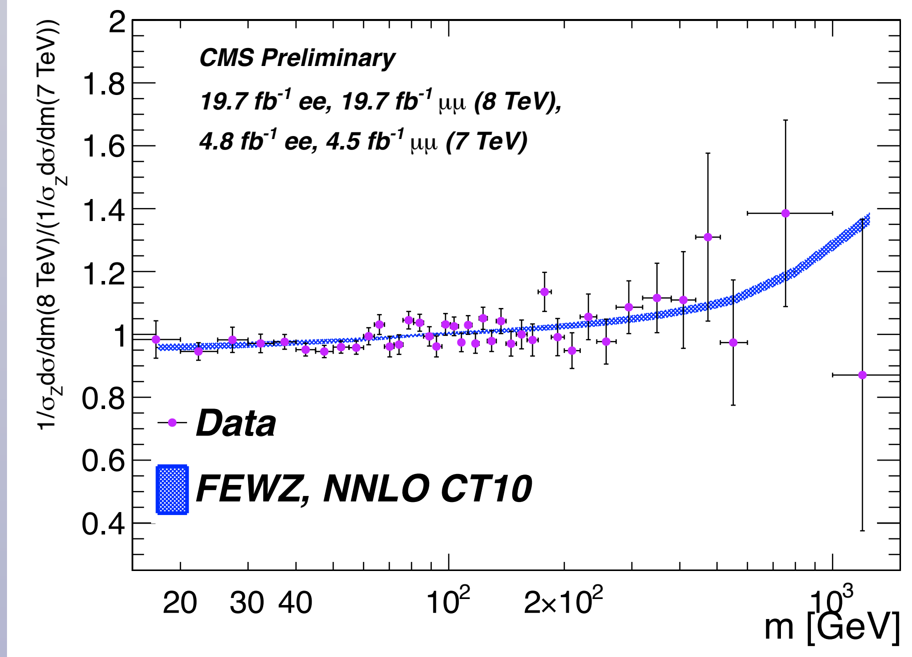
Proof-of-concept measurements at the LHC

ATLAS: measurement of the ratio of **inclusive jet cross-sections between 2.76 TeV and 7 TeV**

$$\rho(y, p_T) = \frac{\sigma(y, p_T, 2.76\ \text{TeV})}{\sigma(y, p_T, 7\ \text{TeV})},$$

CMS: measurement of the ratio of **Drell-Yan distributions between 8 TeV and 7 TeV**





Used in a HERAfitter analysis to study sensitivity to the **gluon PDF**

Could have sensitivity to **large-x antiquarks**

Cross-section ratios of top cross sections

LHC 8 TeV / 7 TeV ratio						
PDF set	σ_{tt}	$\delta_{\rm scale}$	$\delta_{\rm PDF}$	δ_{α_s} (pb)	δ_{m_t}	$\delta_{\rm tot}$
ABM11	1.463	$^{+0.001}_{-0.002}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.006}_{-0.006}$ $(^{+0.4\%}_{-0.4\%})$	$^{+0.000}_{-0.000}$ $(^{+0.0\%}_{-0.0\%})$	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.007}_{-0.008}$ $(^{+0.5\%}_{-0.5\%})$
CT10	1.428	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.008}_{-0.010}$ $(^{+0.5\%}_{-0.7\%})$	$^{+0.002}_{-0.002}$ $(^{+0.2\%}_{-0.2\%})$	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.009}_{-0.011}$ $(^{+0.6\%}_{-0.8\%})$
HERA1.5	1.426	$^{+0.001}_{-0.002}$ $(^{+0.0\%}_{-0.1\%})$	$^{+0.003}_{-0.003}$ $(^{+0.2\%}_{-0.2\%})$	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.004}_{-0.005}$ $(^{+0.3\%}_{-0.4\%})$
MSTW08	1.429	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.004}_{-0.004}$ $(^{+0.2\%}_{-0.2\%})$	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.005}_{-0.005}$ $(^{+0.3\%}_{-0.3\%})$
NNPDF2.3	1.437	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.006}_{-0.006}$ $(^{+0.4\%}_{-0.4\%})$	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.001}_{-0.001}$ $(^{+0.1\%}_{-0.1\%})$	$^{+0.007}_{-0.007}$ $(^{+0.5\%}_{-0.5\%})$
ATLAS	1.36					$\pm$ 0.11 (8%)
CMS	1.40					$\pm$ 0.08 (6%)

8 TeV / 7 TeV: **permille TH accuracy**

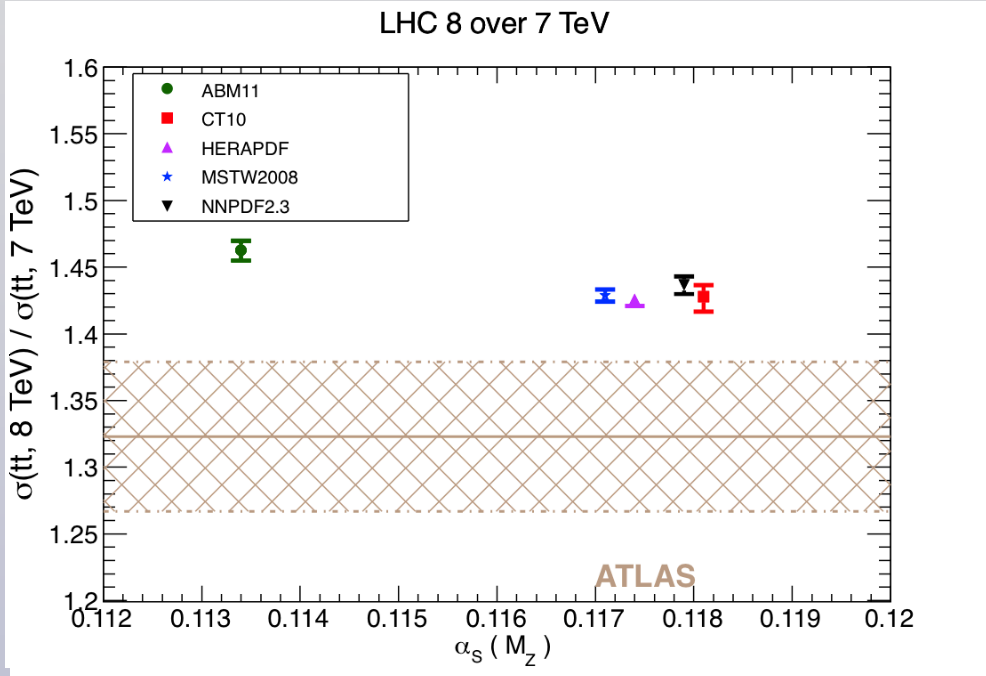
14 TeV / 8 TeV: **1-2% TH accuracy**

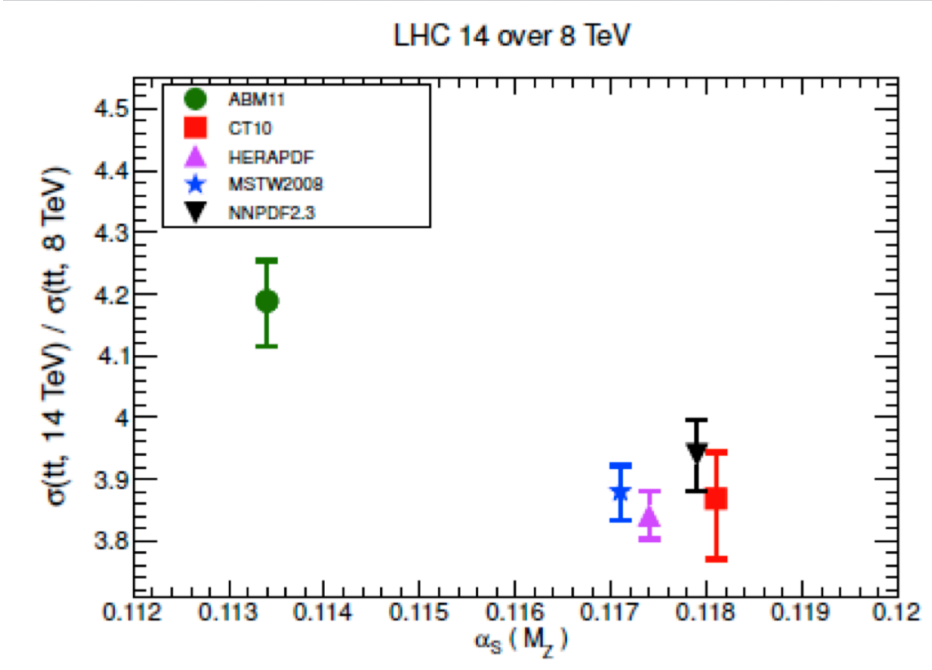
Compare to **7-8%** of absolute xsec

LHC 14 TeV / 8 TeV ratio						
PDF set	σ_{tt}	$\delta_{\rm scale}$	$\delta_{\rm PDF}$	δ_{α_s} (pb)	δ_{m_t}	$\delta_{\rm tot}$
ABM11	4.189	$^{+0.008}_{-0.016}$ $(^{+0.2\%}_{-0.4\%})$	$^{+0.057}_{-0.057}$ $(^{+1.4\%}_{-1.4\%})$	$^{+0.000}_{-0.000}$ $(^{+0.0\%}_{-0.0\%})$	$^{+0.012}_{-0.012}$ $(^{+0.3\%}_{-0.3\%})$	$^{+0.067}_{-0.074}$ $(^{+1.6\%}_{-1.8\%})$
CT10	3.869	$^{+0.006}_{-0.009}$ $(^{+0.2\%}_{-0.2\%})$	$^{+0.068}_{-0.088}$ $(^{+1.8\%}_{-2.3\%})$	$^{+0.020}_{-0.020}$ $(^{+0.5\%}_{-0.5\%})$	$^{+0.010}_{-0.010}$ $(^{+0.2\%}_{-0.2\%})$	$^{+0.077}_{-0.100}$ $(^{+2.0\%}_{-2.6\%})$
HERA1.5	3.841	$^{+0.005}_{-0.012}$ $(^{+0.1\%}_{-0.3\%})$	$^{+0.033}_{-0.025}$ $(^{+0.9\%}_{-0.7\%})$	$^{+0.010}_{-0.010}$ $(^{+0.3\%}_{-0.3\%})$	$^{+0.009}_{-0.010}$ $(^{+0.2\%}_{-0.2\%})$	$^{+0.041}_{-0.041}$ $(^{+1.1\%}_{-1.1\%})$
MSTW08	3.880	$^{+0.006}_{-0.009}$ $(^{+0.2\%}_{-0.2\%})$	$^{+0.036}_{-0.036}$ $(^{+0.9\%}_{-0.9\%})$	$^{+0.011}_{-0.011}$ $(^{+0.3\%}_{-0.3\%})$	$^{+0.010}_{-0.010}$ $(^{+0.2\%}_{-0.2\%})$	$^{+0.045}_{-0.048}$ $(^{+1.2\%}_{-1.2\%})$
NNPDF2.3	3.940	$^{+0.006}_{-0.010}$ $(^{+0.2\%}_{-0.3\%})$	$^{+0.048}_{-0.048}$ $(^{+1.2\%}_{-1.2\%})$	$^{+0.009}_{-0.009}$ $(^{+0.2\%}_{-0.2\%})$	$^{+0.010}_{-0.010}$ $(^{+0.3\%}_{-0.3\%})$	$^{+0.056}_{-0.060}$ $(^{+1.4\%}_{-1.5\%})$

Ratios at NNLO+NNLL within uncertainty band of NLO+NNLL: validation of estimate of scale error in ratios

Cross section ratios of top cross sections





The cross-section ratios are essentially independent of the value of the top quark mass used

- Compare theory predictions for the **8 TeV / 7 TeV ratio** with the recent **ATLAS measurement**:

$$R_{t\bar{t}} = 1.326 \pm 0.024 \pm 0.015 \pm 0.049 \pm 0.001$$

- Interestingly, the data seem to **undershoot the theory prediction by 2-sigma**, and the tension with **AMB11 is enhanced** (3 sigma). To be understood ...

- For the **14 TeV / 8 TeV** ratio, 10% spread between different PDF sets: **clear discrimination power**, but needs dedicated measurements

Precision QCD to enhance BSM searches

- Consider a **final state X** that receives contributions **both from SM and BSM processes**

$$\sigma(pp \to X) = \sigma^{SM}(pp \to X) + \sigma^{BSM}(pp \to X)$$

- Then one can write, assuming the **BSM contribution is small wrt SM one**

$$R^X_{E_1/E_2} \sim \frac{\sigma_X^{SM}(E_1)}{\sigma_X^{SM}(E_2)} \times \left\{ 1 + \frac{\sigma_X^{BSM}(E_1)}{\sigma_X^{SM}(E_1)} \Delta_{E_1/E_2}\left[\frac{\sigma_X^{BSM}}{\sigma_X^{SM}}\right]\right\}$$

$$\Delta_{E_1/E_2}(A) = 1 - \frac{A(E_2)}{A(E_1)}$$

- The **visibility of a BSM contribution** in the **evolution with energy** of the cross section requires that it evolves **differently** from the SM contribution
- The **threshold** to be **sensitive to BSM contributions** is given by the **precision of the SM prediction**, taken into account all theory systematics

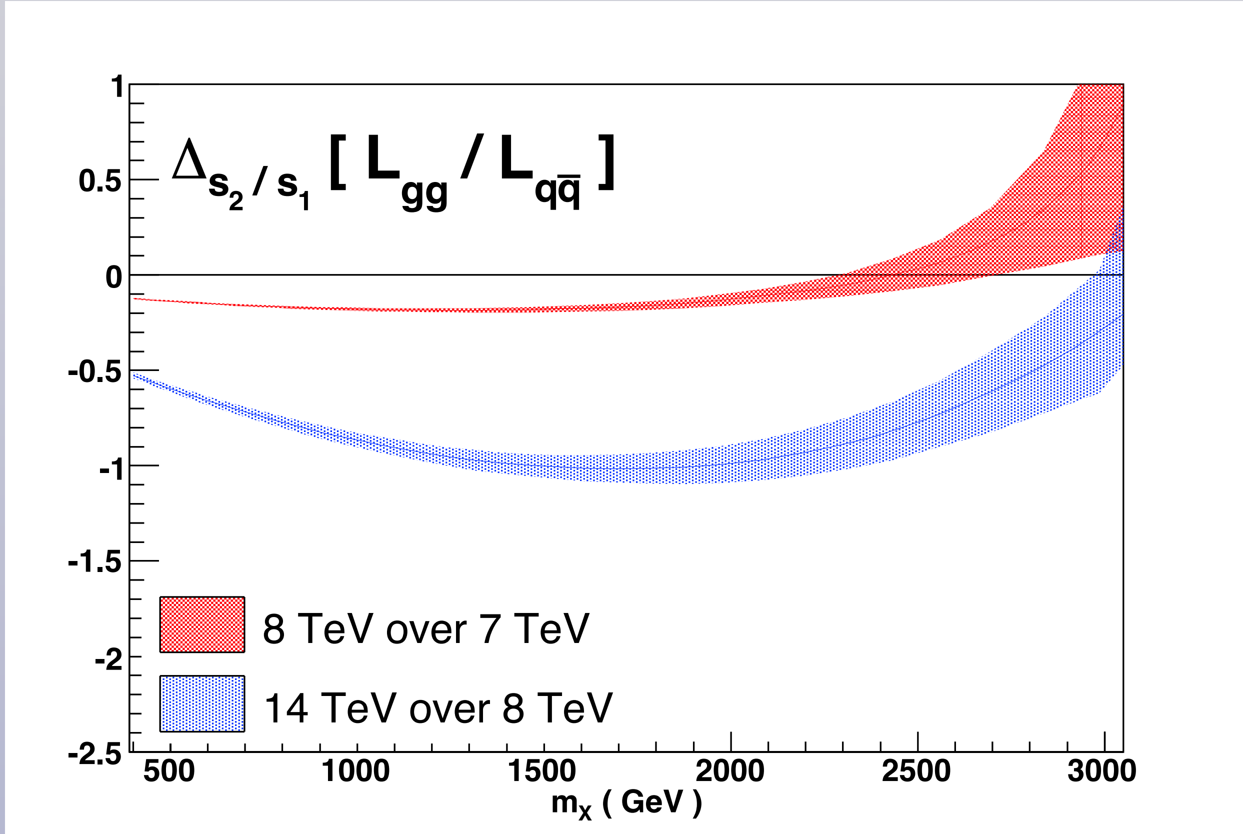
$$\frac{\sigma_X^{BSM}(E_1)}{\sigma_X^{SM}(E_1)} \times \Delta_{E_1/E_2}\left[\frac{\sigma_X^{BSM}}{\sigma_X^{SM}}\right] > \delta_{TH} \equiv \frac{\delta R^{SM}_{E_1/E_2}}{R^{SM}_{E_1/E_2}}$$

- This scaling with energy is driven by the **partonic luminosities**

$$\Delta_{E_1/E_2}\left[\frac{\sigma_X^{BSM}}{\sigma_X^{SM}}\right] \sim \Delta_{E_1/E_2}\left[\frac{\mathcal{L}^{ij}(M)}{\mathcal{L}^{ab}(M)}\right]$$

Precision QCD to enhance BSM searches

- Example: **High mass Z production** dominated by **qqbar luminosity**
- Very different scaling with energy of **gg luminosity:** sensitivity to **BSM gluon initiated contributions** that lead to the same final state as of **Z production**



- Xsec ratios can be **measured very precisely**
- Can be predicted from **theory** with small uncertainties
- In a wide range of scenarios, the **BSM contribution** in these ratios is **enhanced** as compared to absolute cross-sections
- Great opportunities for measurements at Run II

Relative deviation wrt SM scaling

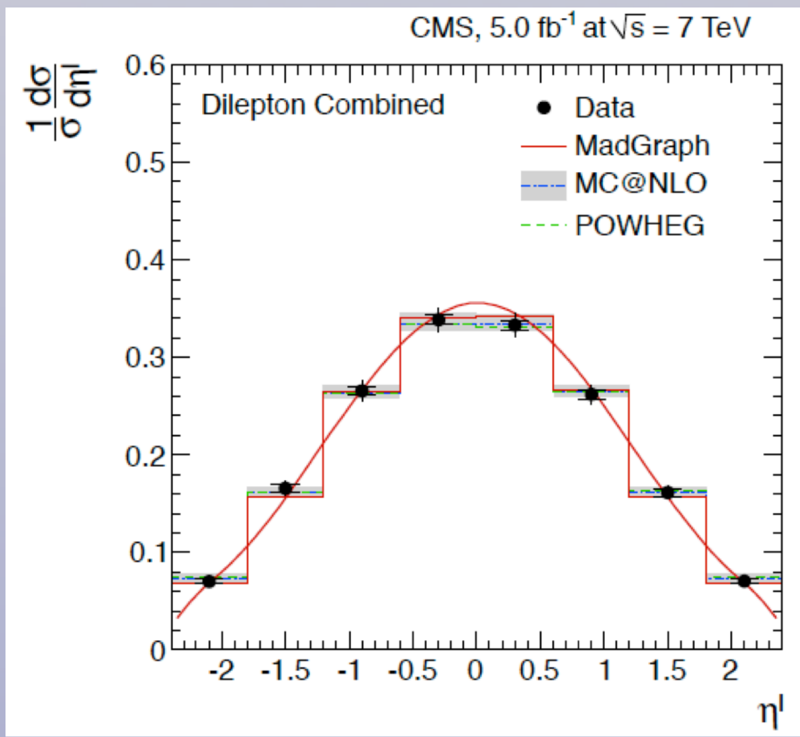
$$\frac{\sigma_Z^{\rm BSM}(m_X)}{\sigma_Z^{\rm SM}(m_X)} \Delta_{E_1/E_2}\left[\frac{\mathcal{L}_{gg}(m_X)}{\mathcal{L}_{q\bar{q}}(m_X)}\right]$$

PDF constraints using the inclusive ttbar cross section

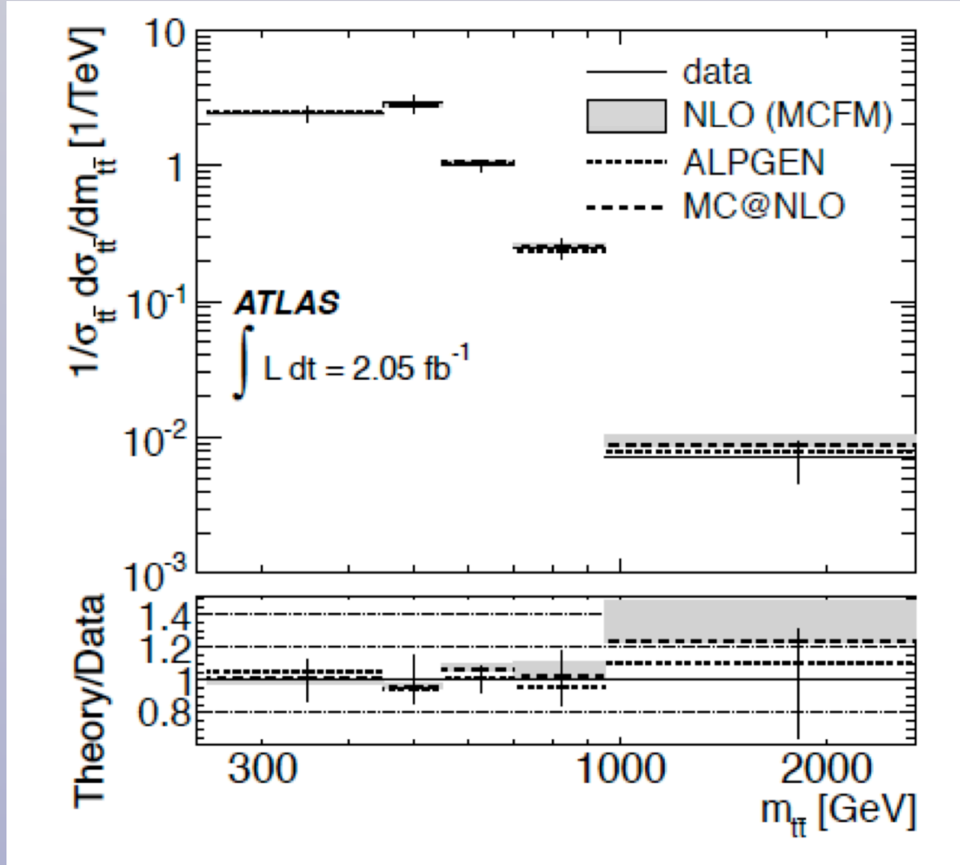
Top differential distributions

- The availability of the fully differential top quark pair calculations at **NNLO** suggest that one should also include **ttbar differential distribution in the PDF fit**
- Measurements available for **ATLAS 7 TeV** and **CMS 7 TeV** and **8 TeV** (the latter preliminary)
- In all cases the **full covariance matrix** is available: ready to be included in PDF fits!

CMS, arxiv:1211.2220

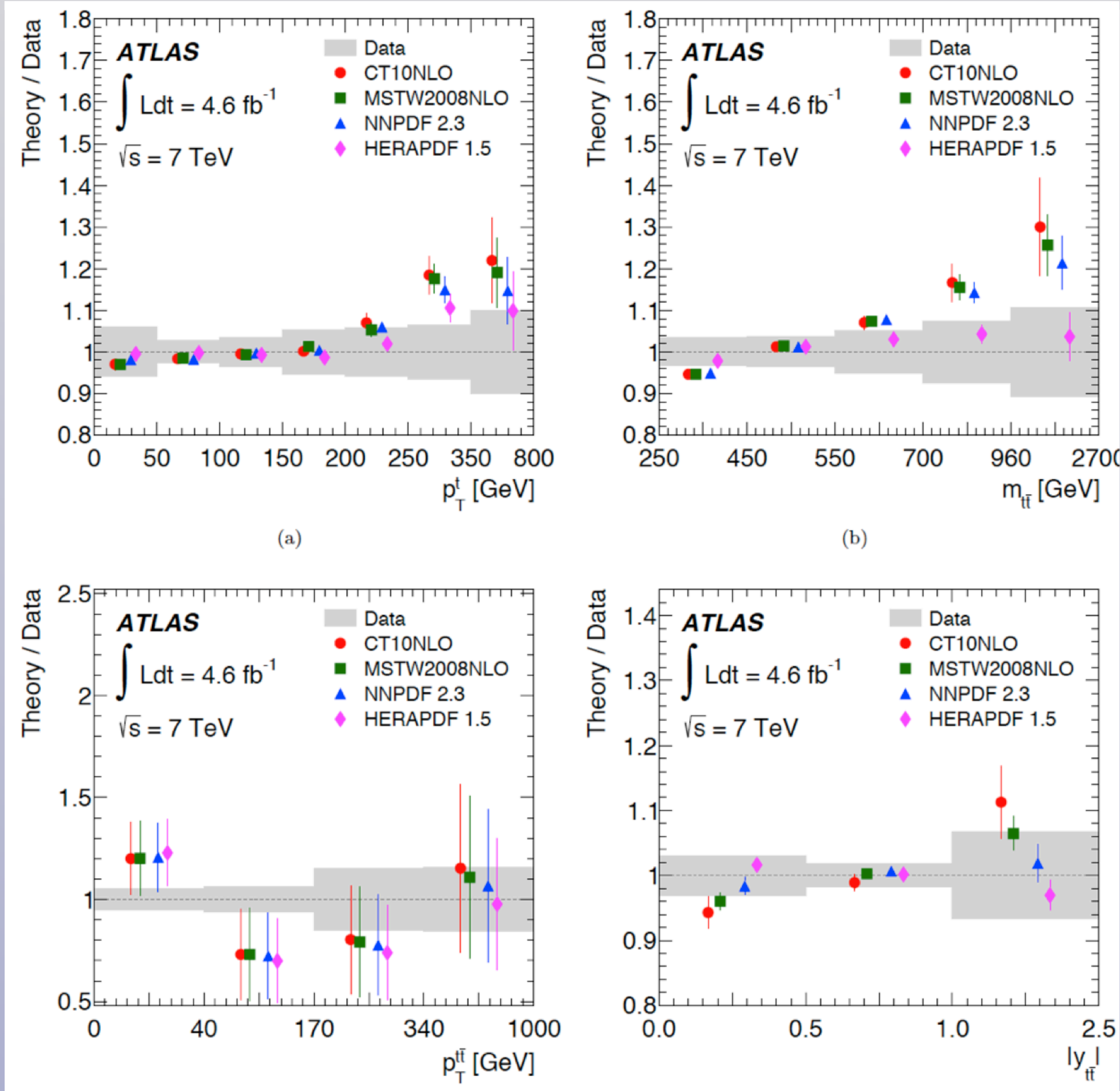


ATLAS, arxiv:1207.5644



Top differential distributions

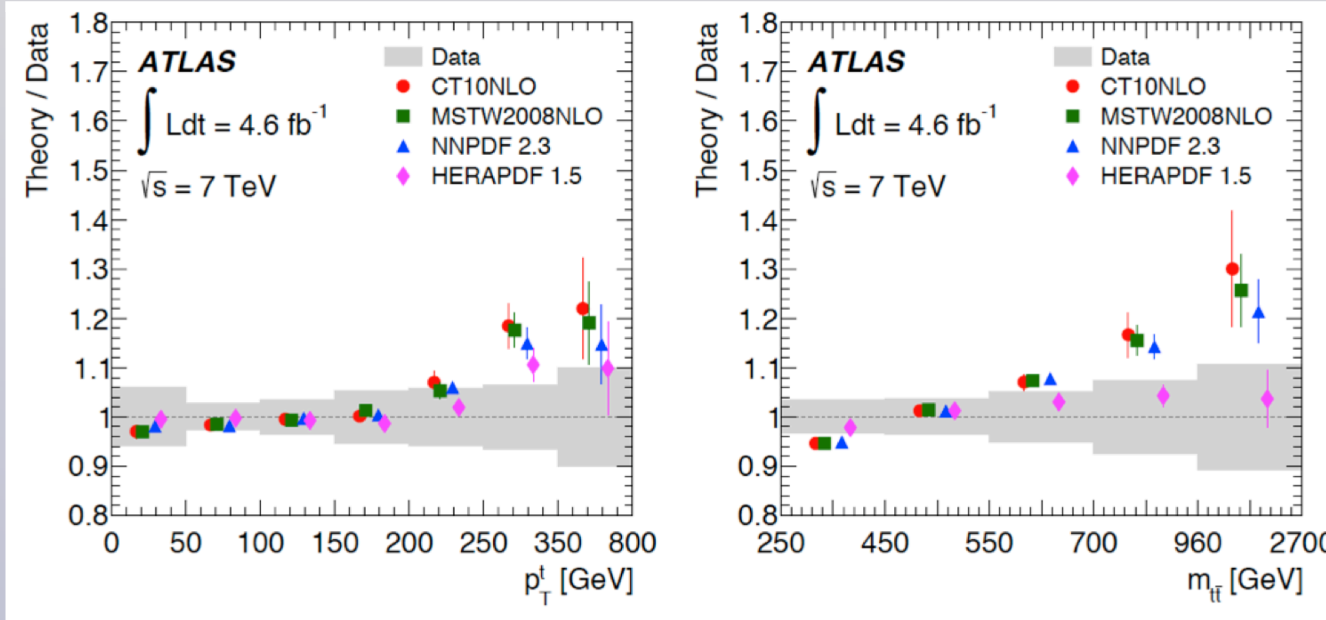
- Recent ATLAS 7 TeV analysis compares with **different PDF sets:**



- Calculation based on MCFM NLO. What are the effects of **NNLO** for these distributions?
- Data/theory compared for **normalized distributions:** absolute should be better
- χ^2 analysis shows **similar level of agreement of the 4 sets,** with the best description provided by HERAPDF1.5 and NNPDF2.3

Top differential distributions

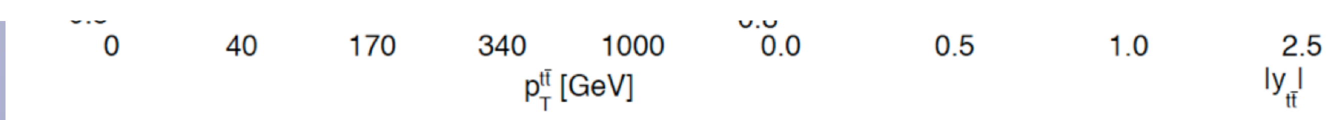
- Recent ATLAS 7 TeV analysis compares with **different PDF sets:**



(a) (b)

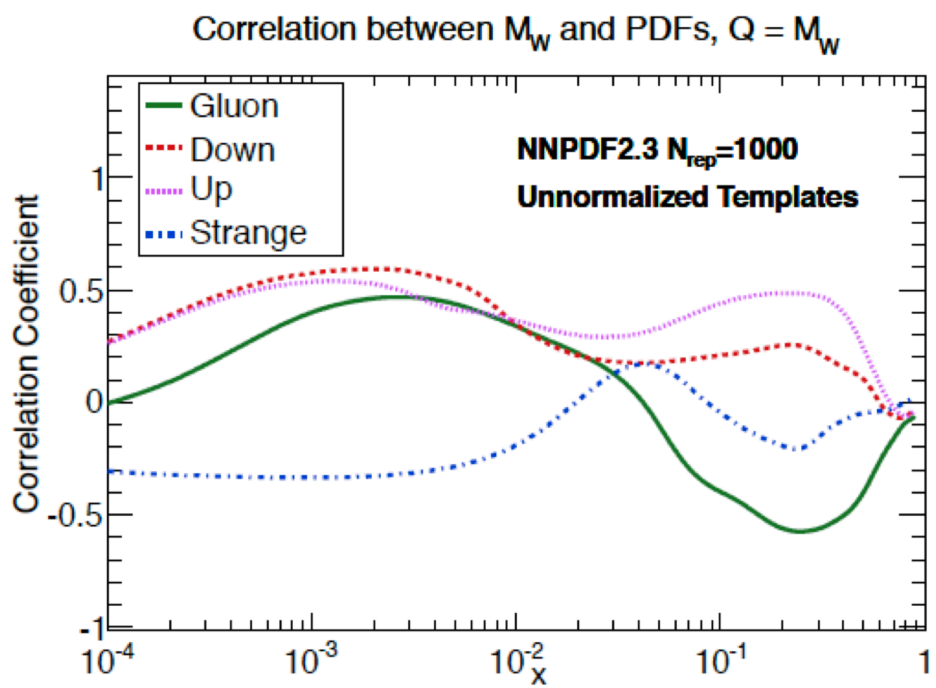
- Calculation based on MCFM NLO. What are the effects of **NNLO** for these distributions?
- Data/theory compared for **normalized distributions:** absolute should be better
- χ^2 analysis shows **similar level of agreement of the 4 sets,** with the best description provided by HERAPDF1.5 and NNPDF2.3

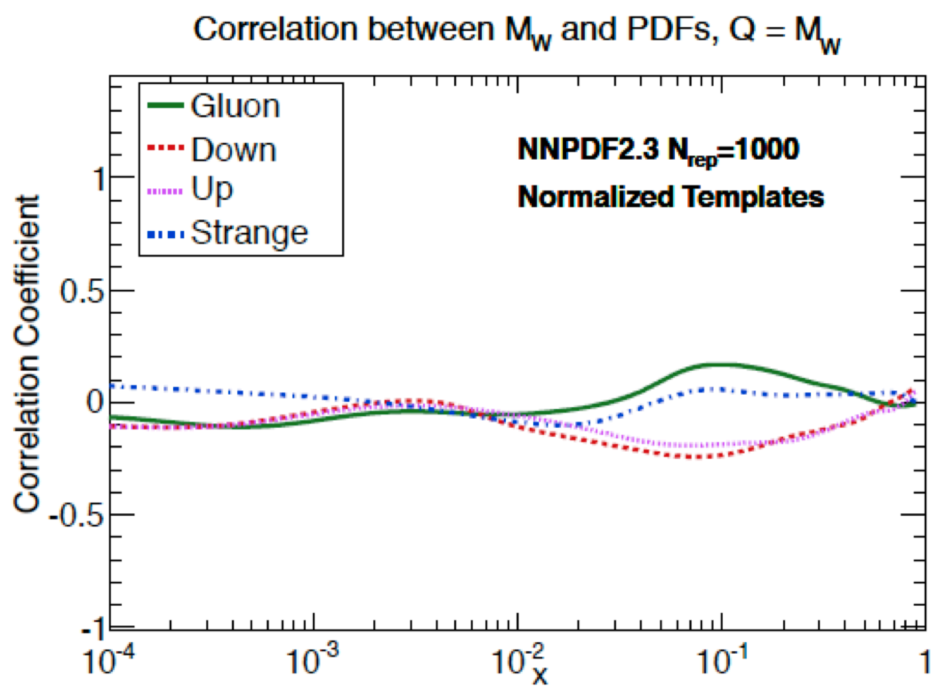
Variable	CT10		MSTW2008NLO		NNPDF 2.3		HERAPDF 1.5	
	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
p_T^t	9.5/6	0.15	9.8/6	0.14	8.2/6	0.22	5.5/6	0.49
$m_{t\bar{t}}$	5.5/4	0.24	6.0/4	0.20	5.2/4	0.27	0.63/4	0.96
$p_T^{t\bar{t}}$	14./3	0.00	13./3	0.01	12./3	0.01	9.1/3	0.03
$\|y_{t\bar{t}}\|$	3.7/2	0.16	4.0/2	0.13	1.3/2	0.52	0.44/2	0.80



Absolute vs Normalized Distributions

- Normalized distributions are nicer because **many experimental systematics cancel**, but the risk is that the **PDF sensitivity is washed out**
- Suggestion: to release **both absolute and normalized distributions,** since each might be more useful in specific physical applications
- To illustrate the point: **correlation** between **PDFs** and the value of the **W mass** extracted from the template fit method in either **absolute or normalized distributions**: PDF sensitivity is substantially reduced in the latter, which is great to extract $\mathbf{M_W}$, but not to fit PDFs!





J. Rojo and A. Vicini, arXiv:1309.1311

NNLO Top differential in PDF fits: K-factors

- There are basically two methods that can be used to include the **NNLO differential top distributions** into a global PDF analysis
  - **Interpolation of partonic matrix elements** and convolution with PDFs and scales a **posteriori**: **APPLgrid, FastNLO, aMCfast** (see later)
  - NNLO/NLO **K-factor approximation**
- The idea for the K-factors is very simple, and d**oes not require any additional computation** provided the suitable distributions, with fine enough binning, are available:
  - Compute, for the same NNLO PDF, the relevant distributions **at NLO and NNLO**
  - Define the **K-factor, bin-by-bin**, as the ratio of the two calculations

$$C_{\rm fact} = \frac{\hat{\sigma}^{\rm NNLO} \otimes \mathcal{L}^{\rm NNLO}}{\hat{\sigma}^{\rm NLO} \otimes \mathcal{L}^{\rm NNLO}}$$

  - Now compute a **fast NLO grid**, for example with **aMC@NLO/aMCfast,** suitable to be included in the PDF fit, and then supplement it with the **NNLO/NLO K-factors**
  - This can also be used to compute **predictions for the PDF error sets**, and for different PDF sets, without having to redo the lengthy NNLO calculations
  - Main drawback: the **PDF sensitivity of the C-factor** should be small but **non negligible**, specially between **different PDF sets** (rather than between PDF error sets)

Including NNLO calculations into global PDF analysis

Ideas inspired by the **aMCfast interface to MadGraph5_aMC@NLO**

V. Bertone, R. Frederix, S. Frixione, J. Rojo, and M. Sutton arXiv:1406.7693

NLO and NNLO Calculations in PDF analyses

- **NLO and NNLO QCD calculations** are too CPU-time intensive to be used directly into PDF analysis, which require iterating the theory predictions a very large number of times

- The traditional solution, **LO supplemented by iterative bin-by-bin K-factors**, is not suitable to match the **precision of LHC data**

- In the recent years, various approaches have been proposed to provide **fast interfaces to NLO calculations**, that can be used directly in PDF analysis, the main ones being:

    - ☑ **APPLgrid**: interfaced to **MCFM** and **NLOJet++**
    - ☑ **FastNLO**: interfaced to **NLOJet++**
    - ☑ **aMCfast:** fast interface to **NLO** and **NLO+PS** calculations in **MadGraph5_aMC@NLO**

- Basic strategy: **interpolate PDFs** in a suitable basis, and **precompute the partonic cross-section** into a set of grids, reconstructing the final distributions via a **fast convolution**. The same ideas underlie most **x-space PDF evolution codes**: HOPPET, QCDNUM, APFEL, ...

Thanks to the **aMCfast** development, the problem of **fast interfaces for NLO and NLO+PS calculations** at the LHC is **completely solved:**

- ☑ **All LHC processes** that can be used to constrain PDFs, now or in the future, can be computed in the form of a **fast NLO interface for PDF fits**

- ☑ Allows to perform **PDF fits with NLO+PS accuracy**, study the stability of PDF fits wrt higher order corrections, increase the number of observables that can be used in PDF fits, and eventually provide **specific PDF sets for NLO event generators**

- ☑ Include consistently **electroweak corrections in PDF fits** at the matrix element level, and include QED effects that allow to perform a **precision determination of the photon PDF** from LHC data

Interpolation Strategy

- The fast interface to **MadGraph5_aMC@NLO, aMCfast,** is constructed using the routines provided by the **APPLgrid** library
- The key idea is to use a generic **higher-order Lagrange interpolation**

$$F(z) = \sum_{i=0}^{s} F\left(\left(\left\lfloor \frac{z}{\delta} - \frac{s-1}{2} \right\rfloor + i\right)\delta\right) I_i^{(s)}\left(\frac{z}{\delta} - \left\lfloor \frac{z}{\delta} - \frac{s-1}{2} \right\rfloor\right)$$

$$I_i^{(s)}(u) = \frac{(-1)^{s-i}}{i!(s-i)!} \prod_{k=0,k\neq i}^{s} (u-k)$$

$$\lfloor u \rfloor \in \mathbb{Z}\,, \qquad u-1 < \lfloor u \rfloor \leq u\,, \qquad u \in \mathbb{R}$$

- We want to use this expansion to **compute a generic integral**, defined by the convolution of two terms

$$J = \int_a^b dz\, S(z)\, F(z) = \sum_{k=1}^{M} \Phi_k\, S(z_k)\, F(z_k)$$

"Slow" function, CPU-time intensive, to be precomputed only one (ie NLO cmatrix element)

"Fast" function, quick evaluation
To be expanded in Lagrange polynomials (ie PDFs)

- The derivation which follows now is completely generic, and applies to **NLO** as well as to **NNLO** calculations: could be used for **NNLO jets, NNLO ttbar** etc

Interpolation Strategy

- Now we can show that the integral J can be written as

$$
\begin{aligned}
J &= \sum_{k=1}^{M} \Phi_k \, S(z_k) \sum_{j=p_\delta(z_k)}^{s+p_\delta(z_k)} F(j\delta) I^{(s)}_{j-p_\delta(z_k)} \left(\frac{z_k}{\delta} - p_\delta(z_k) \right) \\
&= \sum_{j=-\infty}^{\infty} F(j\delta) \, G_j \, ,
\end{aligned}
$$

- That is, in terms of the function **F(z)** evaluated only at a finite subset of **interpolation grid nodes,** weighted by a **factor** that depends on the **node position,** and is by

$$
G_j = \sum_{k=1}^{M} \Phi_k \, S(z_k) \, I^{(s)}_{j-p_\delta(z_k)} \left(\frac{z_k}{\delta} - p_\delta(z_k) \right) \Theta \left(p_\delta(z_k) \leq j \leq s + p_\delta(z_k) \right)
$$

- The idea of **aMCfast** is to extract the information on the hard-scattering matrix elements from **MadGraph5_aMC@NLO,** use this to fill the **APPLgrid** interpolating grids, and then **reconstruct the original distributions a posteriori** with arbitrary PDFs and scales.

- Note that in principle it is possible to change a posterior not only the **numerical values of the scales,** but also its **functional form**: extremely powerful to investigate scale variations with a minimum of CPU effort

NLO FKS short-distance cross-sections

- In the **FKS subtraction formalism**, a generic 2 -> n short distance cross-section is given by

$$
\begin{aligned}
d\sigma^{(\mathrm{NLO})} &\longleftrightarrow \left\{ d\sigma^{(\mathrm{NLO},\alpha)} \right\}_{\alpha=E,S,C,SC}, \\
d\sigma^{(\mathrm{NLO},\alpha)} &= f_1(x_1^{(\alpha)}, \mu_F^{(\alpha)}) f_2(x_2^{(\alpha)}, \mu_F^{(\alpha)}) W^{(\alpha)} d\chi_{Bj} d\chi_{n+1}
\end{aligned}
$$

with the contribution of the **fully resolved event** (E) and of the **soft** (S), **collinear** (C) and **soft-collinear** (SC) counterevents

The information on **partonic matrix-elements**, including scale dependence, is encoded in the functions:

$$
\begin{aligned}
W^{(\alpha)} &= g_S^{2b+2}(\mu_R^{(\alpha)}) \left[\widehat{W}_0^{(\alpha)} + \widehat{W}_F^{(\alpha)} \log\left(\frac{\mu_F^{(\alpha)}}{Q}\right)^2 + \widehat{W}_R^{(\alpha)} \log\left(\frac{\mu_R^{(\alpha)}}{Q}\right)^2 \right] \\
&+ g_S^{2b}(\mu_R^{(\alpha)}) \widehat{W}_B \delta_{\alpha S},
\end{aligned}
$$

where the various pieces are the **Born**, the **scale independent NLO term**, and the terms which encode the **renormalization and factorization scale dependences** of the NLO calculation

In **aMCfast,** the information on the **event-by-event *W* weights** is extracted and used to fill the **APPLgrid** interpolation grids

It should be able to express **any NLO or NNLO calculation in a similar form**, and then the same methods use in aMCfast could be used there

Short-distance cross-sections

- One can finally write the **four components of the FKS cross-section** in terms of **PDFs and strong coupling** evaluated only **at the grid nodes**, and grid weight factors which encode **all dependence on matrix elements**, and that need to be computed only once

$$
\begin{aligned}
\sigma_{O,0}^{(h)} &= \sum_{j_1,j_2,j_3,j_4} f_1(j_1\delta_1, j_3\delta_3)\, f_2(j_2\delta_2, j_3\delta_3)\, g_S^{2b+2}(j_4\delta_4)\, G_{j_1j_2j_3j_4}^{(h,0)}\,, \\
\sigma_{O,F}^{(h)} &= \sum_{j_1,j_2,j_3,j_4} f_1(j_1\delta_1, j_3\delta_3)\, f_2(j_2\delta_2, j_3\delta_3)\, g_S^{2b+2}(j_4\delta_4) \log\left(\frac{j_3\delta_3}{Q}\right)^2 G_{j_1j_2j_3j_4}^{(h,F)}\,, \\
\sigma_{O,R}^{(h)} &= \sum_{j_1,j_2,j_3,j_4} f_1(j_1\delta_1, j_3\delta_3)\, f_2(j_2\delta_2, j_3\delta_3)\, g_S^{2b+2}(j_4\delta_4) \log\left(\frac{j_4\delta_4}{Q}\right)^2 G_{j_1j_2j_3j_4}^{(h,R)}\,, \\
\sigma_{O,B}^{(h)} &= \sum_{j_1,j_2,j_3,j_4} f_1(j_1\delta_1, j_3\delta_3)\, f_2(j_2\delta_2, j_3\delta_3)\, g_S^{2b}(j_4\delta_4)\, G_{j_1j_2j_3j_4}^{(h,B)}\,,
\end{aligned}
$$

- As opposed to other NLO calculations, in the FKS formalism we require **four independent interpolation grids** rather than two, to account for the most general possible scale variations
- The final distributions are reconstructed **a posteriori** and require only the **knowledge of PDFs and scales in a pre-determined (x,Q) grid**: extremely fast and efficient

Scale variations

- Once the APPLgrid interpolating grids have been produced with **aMCfast**, we can recompute the **original kinematical distributions** for arbitrary **PDFs, scales and value of α_S**
- We assume that scales vary wrt the central value by a **constant factor** in all phase space

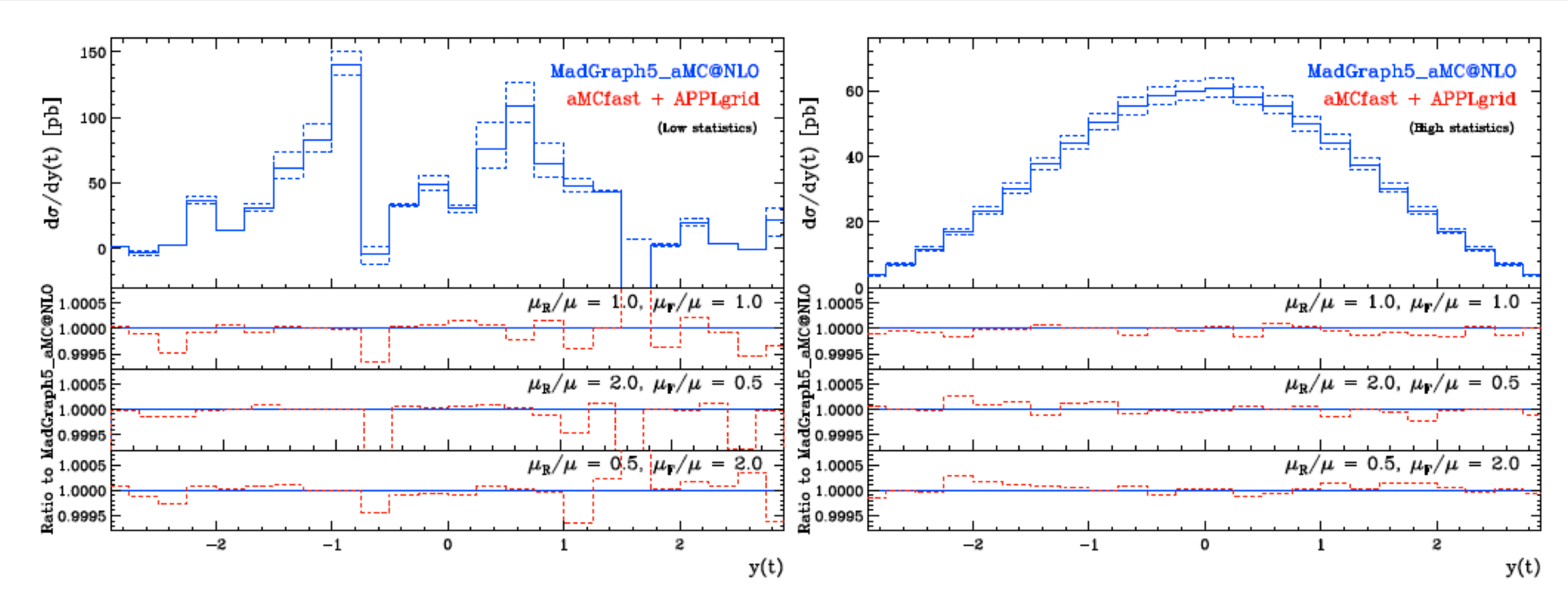
$$\mu_F = \xi_F \, \mu \,, \qquad \mu_R = \xi_R \, \mu$$

- We end up with the following final expressions in terms of PDFs in the grid and node weight factors:

$$\begin{aligned}
\sigma^{(h)}_{O,0} &= \sum_{j_1,j_2,j_3} \sum_{l=1}^{n_l} \widehat{\mathcal{F}}^{(l)}(j_1\delta_y, j_2\delta_y, j_3\delta_\tau)\hat{g}_S^{2b+2}(j_3\delta_\tau)\, G^{(h,0,l)}_{j_1j_2j_3}\,, \\
\sigma^{(h)}_{O,F} &= \sum_{j_1,j_2,j_3} \sum_{l=1}^{n_l} \widehat{\mathcal{F}}^{(l)}(j_1\delta_y, j_2\delta_y, j_3\delta_\tau)\,\hat{g}_S^{2b+2}(j_3\delta_\tau)\ \log\xi_F^2\, G^{(h,F,l)}_{j_1j_2j_3}\,, \\
\sigma^{(h)}_{O,R} &= \sum_{j_1,j_2,j_3} \sum_{l=1}^{n_l} \widehat{\mathcal{F}}^{(l)}(j_1\delta_y, j_2\delta_y, j_3\delta_\tau)\,\hat{g}_S^{2b+2}(j_3\delta_\tau)\ \log\xi_R^2\, G^{(h,R,l)}_{j_1j_2j_3}\,, \\
\sigma^{(h)}_{O,B} &= \sum_{j_1,j_2,j_3} \sum_{l=1}^{n_l} \widehat{\mathcal{F}}^{(l)}(j_1\delta_y, j_2\delta_y, j_3\delta_\tau)\,\hat{g}_S^{2b}(j_3\delta_\tau)\, G^{(h,B,l)}_{j_1j_2j_3}\,,
\end{aligned}$$

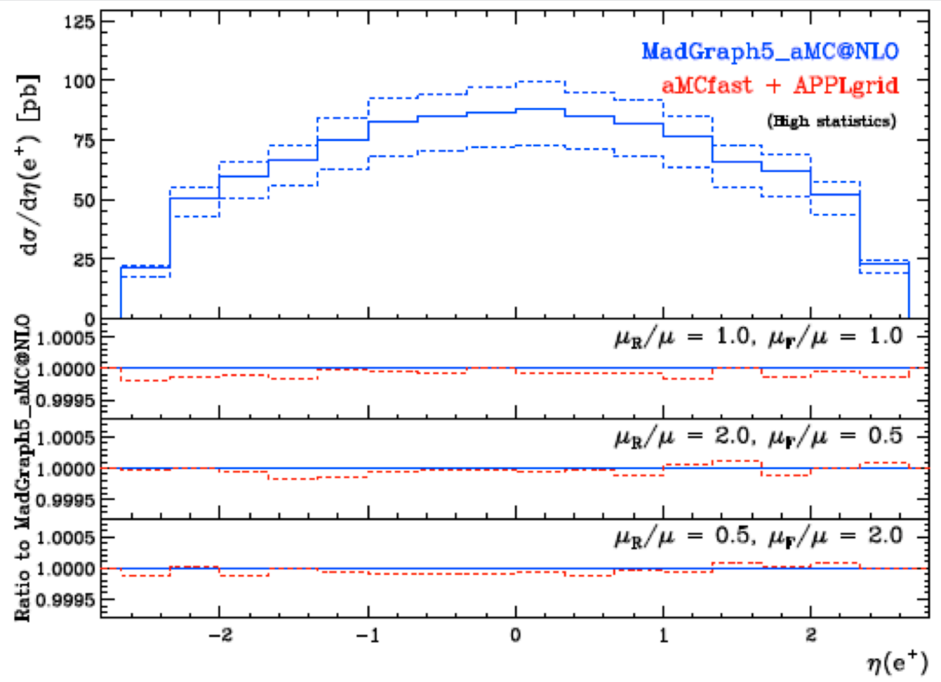
The **aMCfast method is fully generic**, valid on a **event-by-event basis** and completely stand-alone. Also, if needed, it can be generalized to more complex choices of scale variations

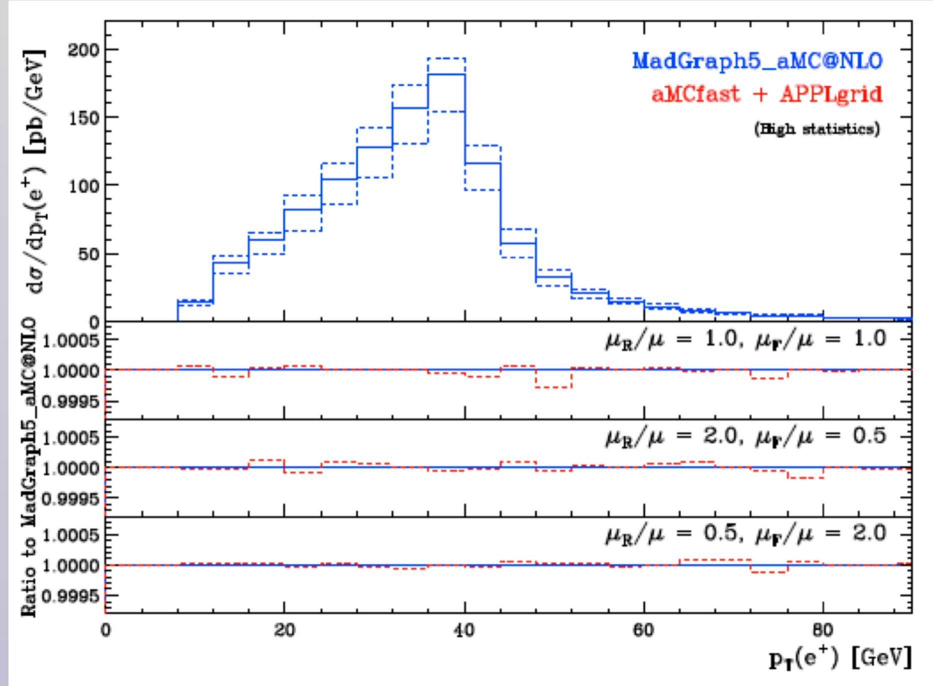
Top-quark pair production



- Useful to constrain the **gluon PDF**
- For this process, we have the contribution from 7 independent PDF luminosities.
- The PDF Luminosities for each new processes are **determined automatically at run time**
- Excellent agreement with original distributions both at low and at high statistics

W production in association with charm quarks





- Crucial experimental data to constrain the **strange PDF**
- Illustrates that **resonance decays** can the trivially included in aMCfast
- Any **theoretical refinement of the original MadGraph5_aMC@NLO calculation**, such as the use of complex mass scheme to account for finite widths of resonances, **translates automatically into aMCfast**
- Can be used to include the hadron level ATLAS W+D data into a PDF fit

Towards a standard fast interface to NNLO?

- In aMCfast, the accuracy of the interpolation can be made **arbitrarily high** by using a finer grid in (x,Q), and O(10^{-5}) accuracy is achieved with only **30-50 points in x and 20 in Q**
- The **numerical cancellation between weights** that have large difference in size is not a problem: in aMCfast we have filled interpolating grids by weights that can differ by 15 or 20 orders of magnitude, and the interpolation always works beautifully. This should be important for NNLO!
- It would be a massive benefit for the LHC community if **different NNLO calculations** could be expressed in a way that allowed to then to perform a **fast interpolation**, using a similar strategy as in aMCfast
- Ideally, one would like to provide a **"Les-Houches"-like convention** for how, event by event, information on NNLO calculations should be available to **plug-in the fast interface**, and therefore enormously facilitate the **availability of all these great NNLO results for global PDF analysis**
- Perhaps document such an agreement in the **Whitepaper of this workshop**?

Summary

- **Top quark pair production** at the LHC is full of unique opportunities for **high precision tests of the SM**, as well as a testing ground for many **scenarios of BSM dynamic**s

- The availability of the **fully differential NNLO calculation** marks the beginning of a dedicated program of phenomenological applications, one of the main ones being **PDF studies**

- Total cross-sections already provide **constraints on the large-x gluon PDF**, validated by various groups, and more information should be provided by the **differential distributions** from **ATLAS and CMS at 7 and 8 TeV**, and then Run II data, which will extend their kinematic reach

- The cross-section ratios between **8 and 7 TeV** and **14 and 8 TeV** provide a **new avenue for high-precision QCD checks** and for **BSM searches**. In ttbar production, it provides **complementary constraints on PDFs** that are highly stable against higher orders and top quark mass choice

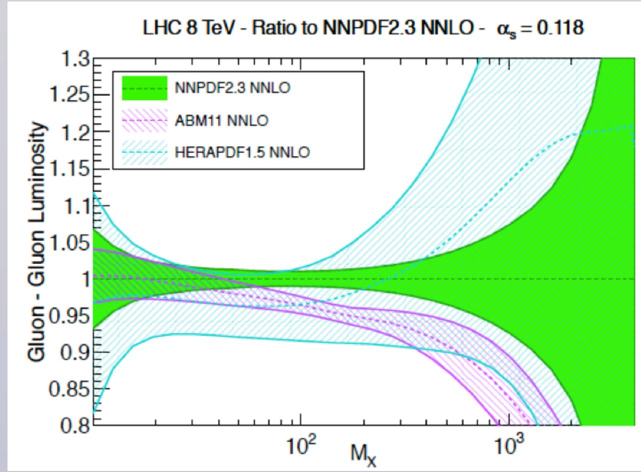
- **NLO and NNLO calculations** of LHC processes are an essential ingredients of **modern global PDF fits**. With the development of the **aMCfast interface to MadGraph5_aMC@NLO**, the problem of fast interfaces to NLO processes is **solved once and for all**

- The same strategy can and should be applied to **NNLO calculations**, but programs should be able to provide the right information on hard cross-sections, PDFs and scales event-by event. This could be the key to extend the **NNLO revolution directly into PDF fits**!

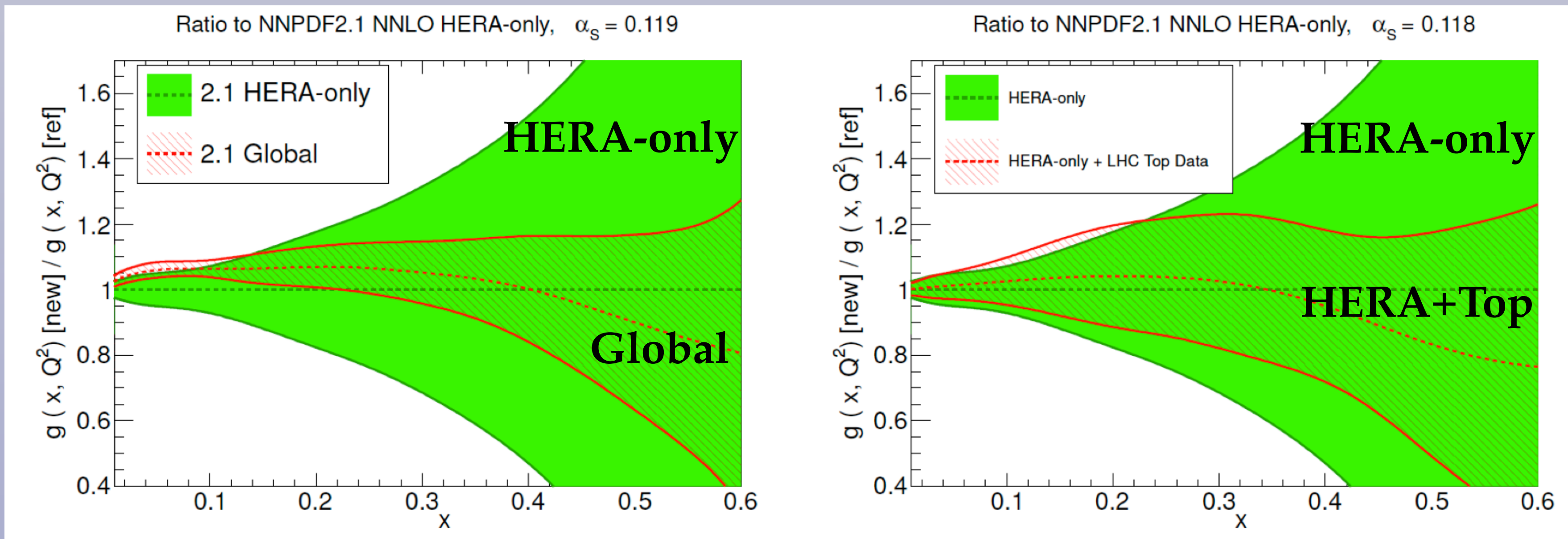
Extra Material

Impact in DIS-only fits

- PDF fits based on **reduced datasets**, such as HERAPDF, display large PDF uncertainties for the gluon due to the lack of direct constraints

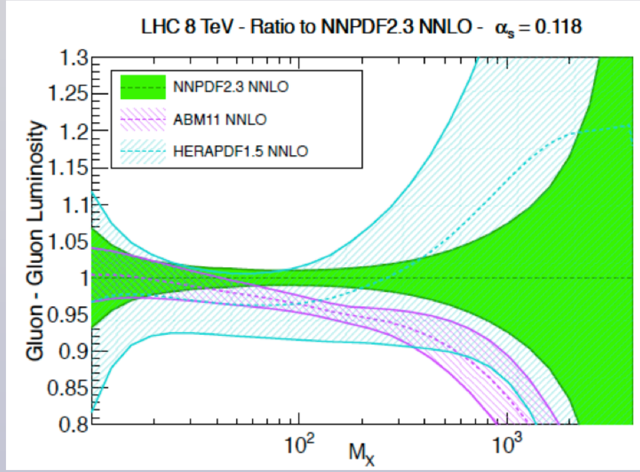


- Top quark data can be included in a **NNLO fit based on HERA data**
  - *Substantial reduction of PDF uncertainties*
  - *The HERA+Top gluon PDF is close to the gluon from the global PDF fit*



Impact in DIS-only fits

- PDF fits based on **reduced datasets**, such as HERAPDF, display large PDF uncertainties for the gluon due to the lack of direct constraints

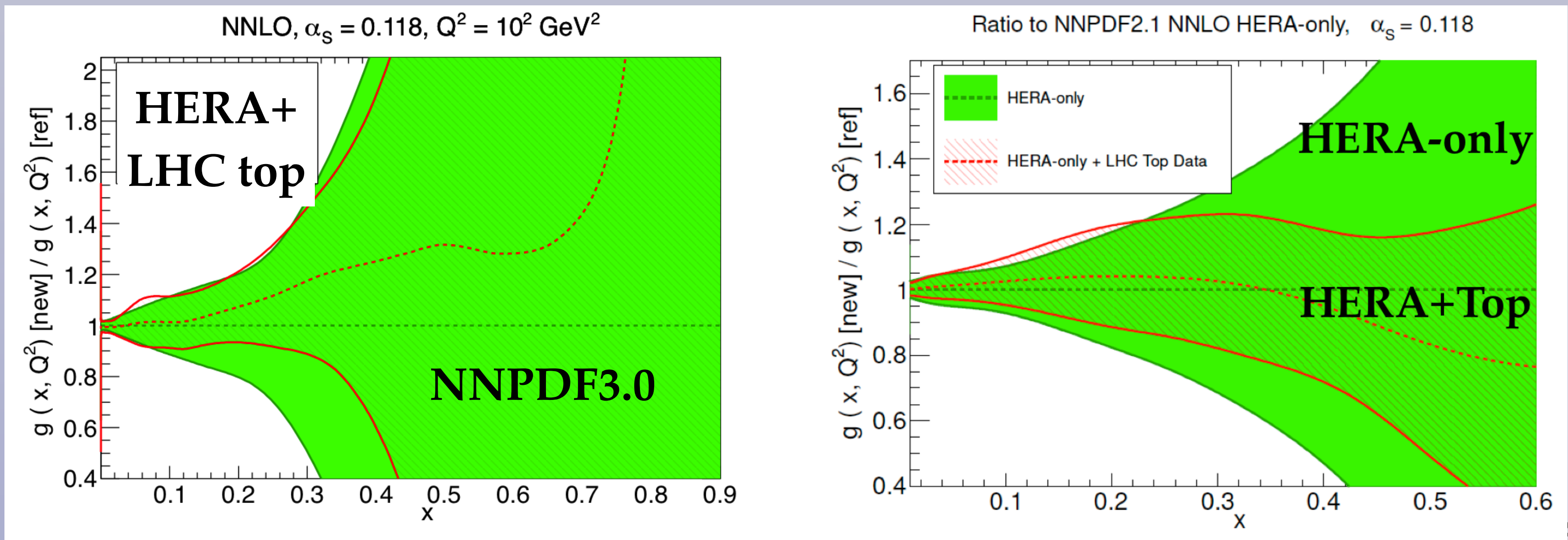


gg luminosity

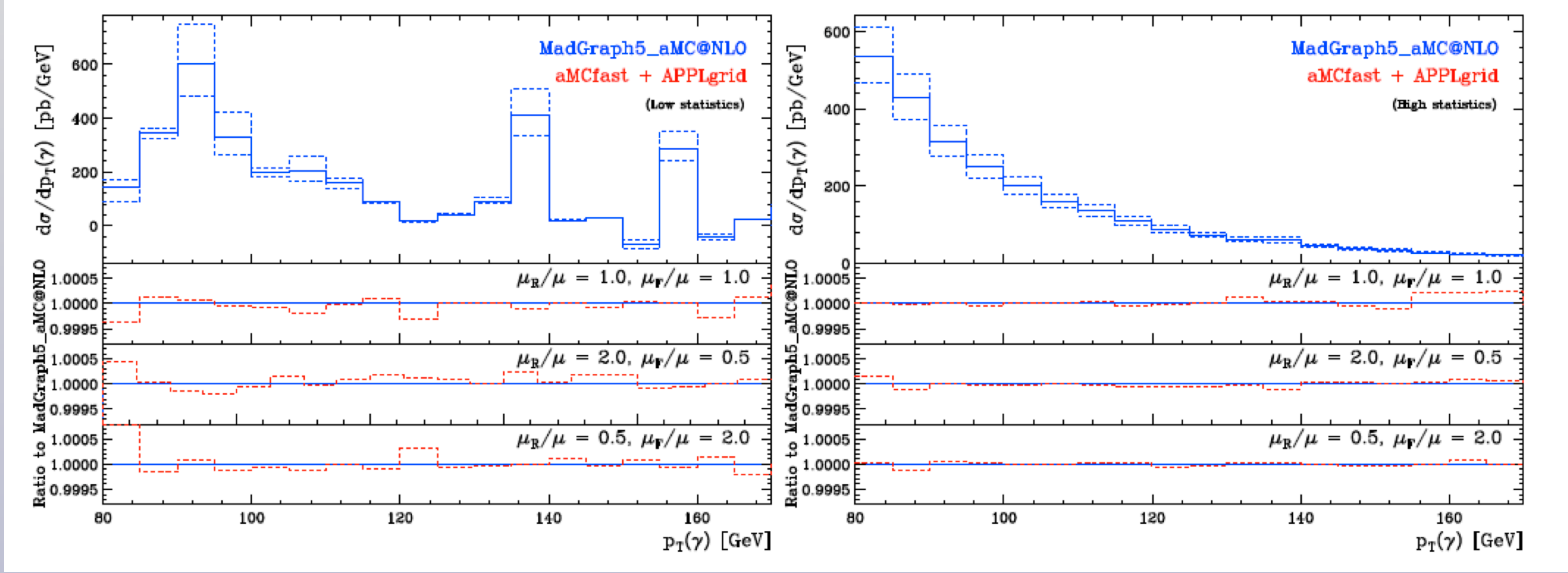
- Top quark data can be included in a **NNLO fit based on HERA data**

  Substantial reduction of PDF uncertainties

  The HERA+Top gluon PDF is close to the gluon from the global PDF fit



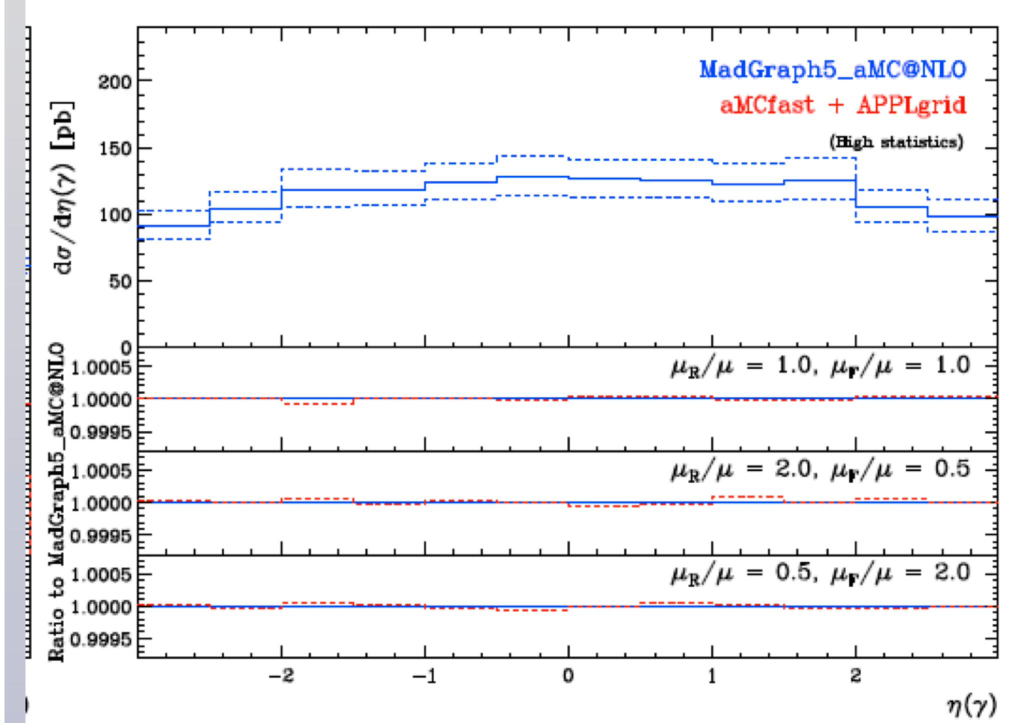
Photon+jet production




```
MG5_aMC> define j = g u d s c b u~ d~ s~ c~ b~
MG5_aMC> generate p p > a j [QCD]
```

- Could have PDF discrimination power if experimental measurements are precise enough
- Same generation commands as before, adding the correct jet definition for a 5-flavor scheme
- Again, **excellent agreement** for all distributions and choices of scales
- Uses the **Frixione isolation criterion** to remove the fragmentation component
- Generation-level cuts and analysis parameters (including isolation) **accessible from the run card**

Photon+jet production

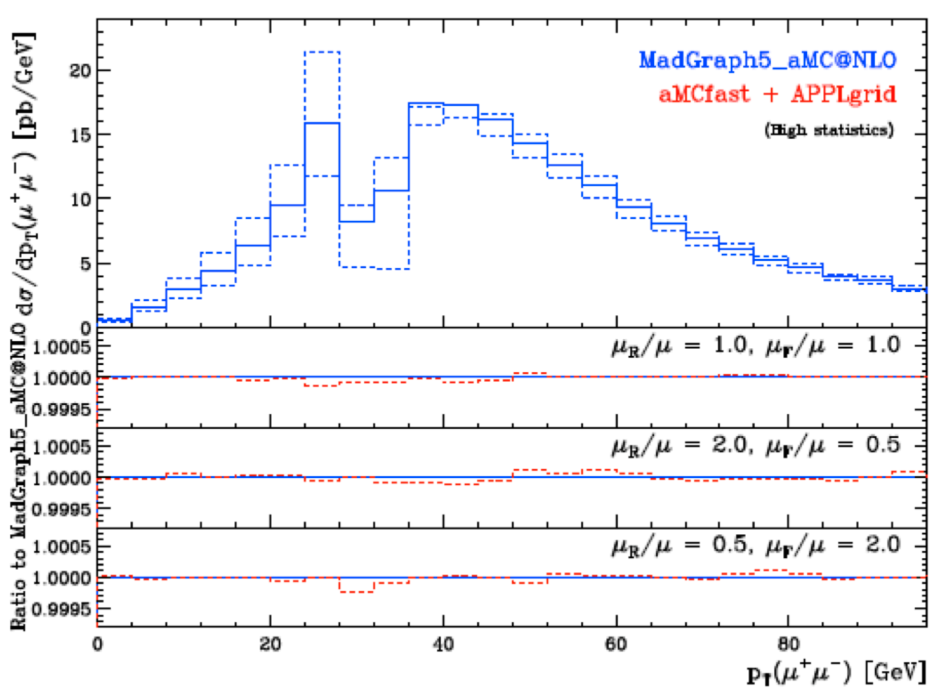


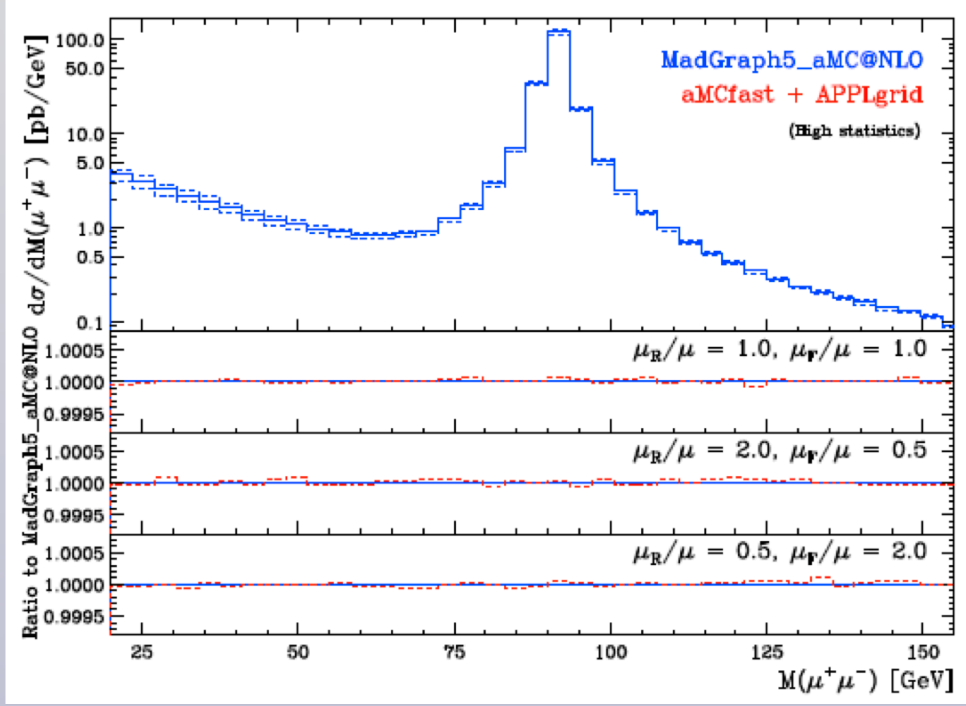
- Due to the complexity of this process, here we have **33 independent PDF luminosities**
- Determined at run time, this information is **completely transparent to the user**

l	n_{rs}	(r,s)					
1	2	(g,u)	(g,c)				
2	6	$(\bar{b},u)$	$(\bar{b},c)$	$(\bar{s},u)$	$(\bar{s},c)$	$(\bar{d},u)$	$(\bar{d},c)$
3	6	(d,u)	(d,c)	(s,u)	(s,c)	(b,u)	(b,c)
4	2	$(\bar{c},c)$	$(\bar{u},u)$				
5	2	(u,u)	(c,c)				
6	2	$(\bar{c},u)$	$(\bar{u},c)$				
7	2	(u,c)	(c,u)				
8	1	(g,g)					
9	3	(g,d)	(g,s)	(g,b)			
10	3	$(\bar{b},b)$	$(\bar{s},s)$	$(\bar{d},d)$			
11	3	(d,d)	(s,s)	(b,b)			
12	6	$(\bar{c},d)$	$(\bar{c},s)$	$(\bar{c},b)$	$(\bar{u},d)$	$(\bar{u},s)$	$(\bar{u},b)$
13	6	(u,d)	(u,s)	(u,b)	(c,d)	(c,s)	(c,b)
14	6	$(\bar{b},d)$	$(\bar{b},s)$	$(\bar{s},d)$	$(\bar{s},b)$	$(\bar{d},s)$	$(\bar{d},b)$
15	6	(d,s)	(d,b)	(s,d)	(s,b)	(b,d)	(b,s)
16	2	$(g,\bar{c})$	$(g,\bar{u})$				
17	6	$(\bar{b},\bar{c})$	$(\bar{b},\bar{u})$	$(\bar{s},\bar{c})$	$(\bar{s},\bar{u})$	$(\bar{d},\bar{c})$	$(\bar{d},\bar{u})$
18	6	$(d,\bar{c})$	$(d,\bar{u})$	$(s,\bar{c})$	$(s,\bar{u})$	$(b,\bar{c})$	$(b,\bar{u})$
19	2	$(\bar{c},\bar{c})$	$(\bar{u},\bar{u})$				
20	2	$(u,\bar{u})$	$(c,\bar{c})$				
21	2	$(\bar{c},\bar{u})$	$(\bar{u},\bar{c})$				
22	2	$(u,\bar{c})$	$(c,\bar{u})$				
23	3	$(g,\bar{b})$	$(g,\bar{s})$	$(g,\bar{d})$			
24	3	$(\bar{b},\bar{b})$	$(\bar{s},\bar{s})$	$(\bar{d},\bar{d})$			
25	3	$(d,\bar{d})$	$(s,\bar{s})$	$(b,\bar{b})$			
26	6	$(\bar{c},\bar{b})$	$(\bar{c},\bar{s})$	$(\bar{c},\bar{d})$	$(\bar{u},\bar{b})$	$(\bar{u},\bar{s})$	$(\bar{u},\bar{d})$
27	6	$(u,\bar{b})$	$(u,\bar{s})$	$(u,\bar{d})$	$(c,\bar{b})$	$(c,\bar{s})$	$(c,\bar{d})$
28	6	$(\bar{b},\bar{s})$	$(\bar{b},\bar{d})$	$(\bar{s},\bar{b})$	$(\bar{s},\bar{d})$	$(\bar{d},\bar{b})$	$(\bar{d},\bar{s})$
29	6	$(d,\bar{b})$	$(d,\bar{s})$	$(s,\bar{b})$	$(s,\bar{d})$	$(b,\bar{s})$	$(b,\bar{d})$
30	2	(u,g)	(c,g)				
31	3	(d,g)	(s,g)	(b,g)			
32	2	$(\bar{c},g)$	$(\bar{u},g)$				
33	3	$(\bar{b},g)$	$(\bar{s},g)$	$(\bar{d},g)$			

Juan Rojo

dilepton+jet production

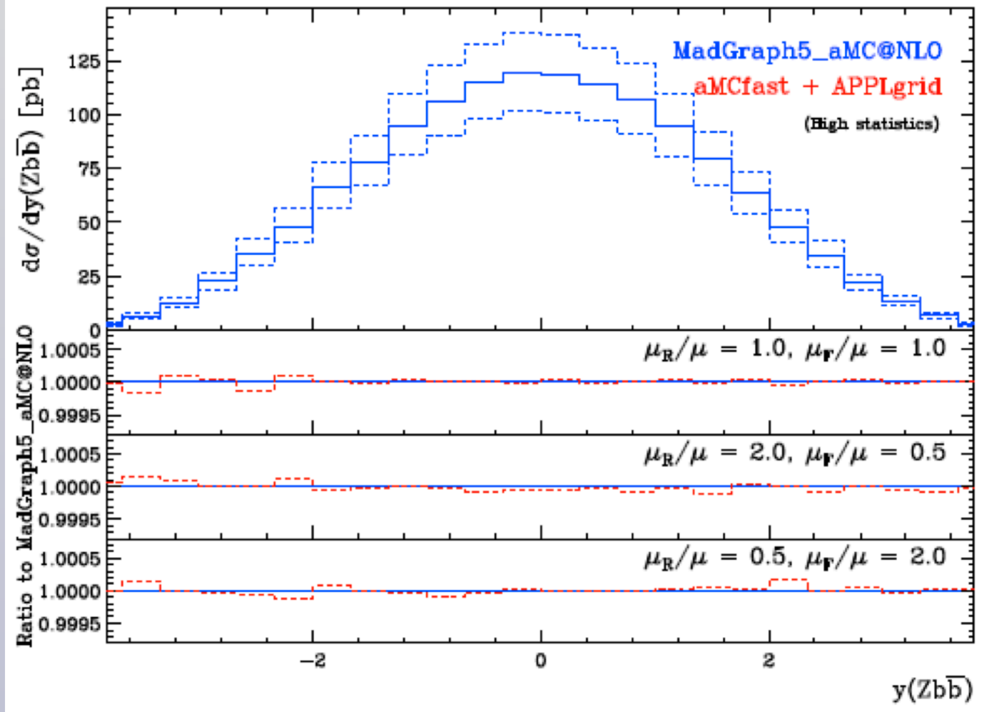


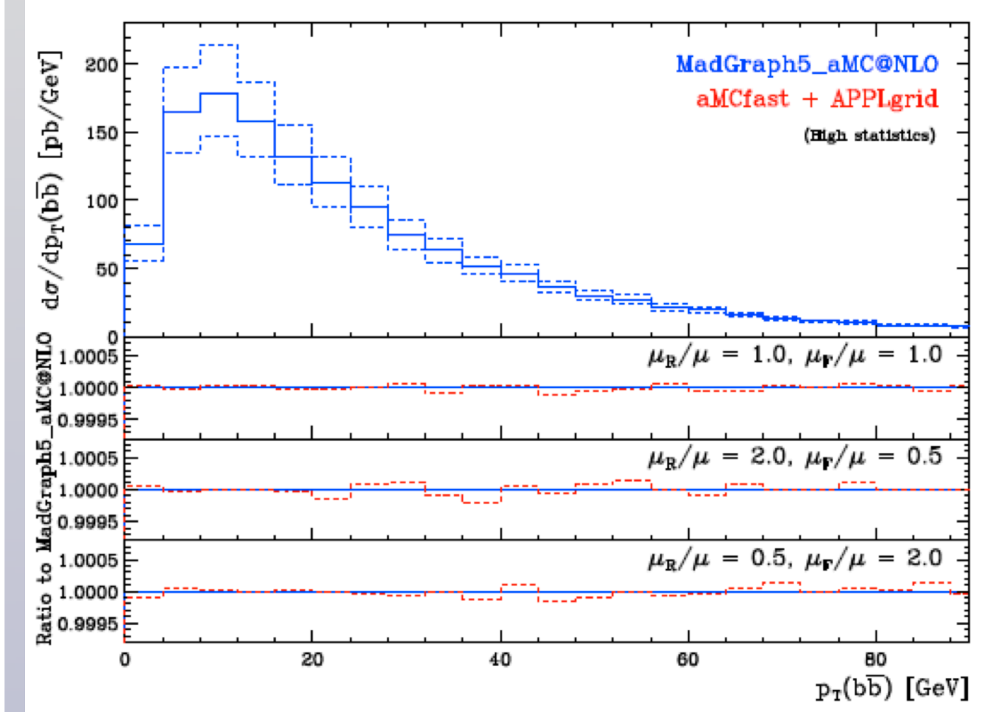



```
MG5_aMC>  generate p p > mu+ mu- j [QCD]
```

- Useful to constrain the **gluon** and the **antiquarks**
- Feature around pT = 30 GeV understood from perturbative instability in fNLO calculation, requires resummation
- Ratios of this process with W+jets could provide important information on **quark flavor separation**

Zbb production



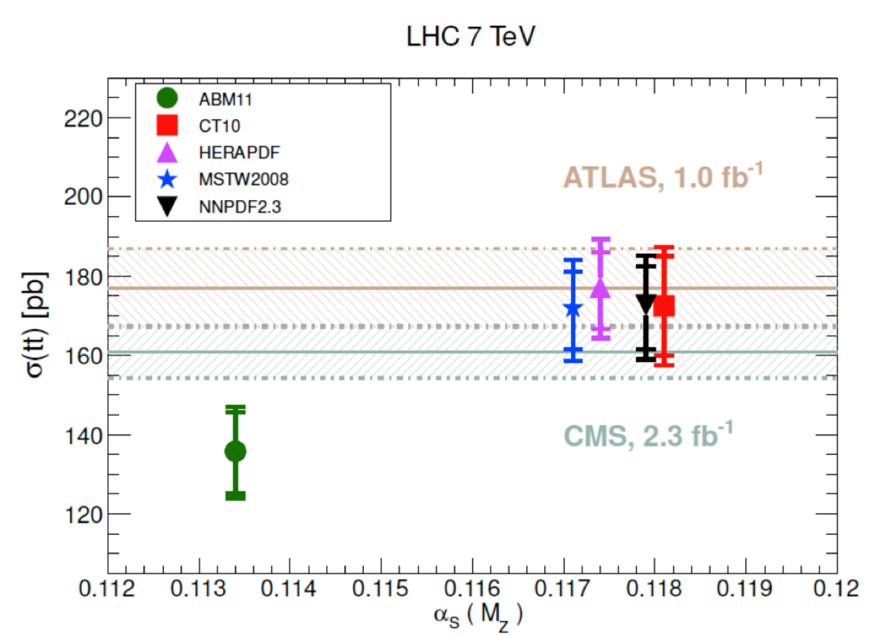


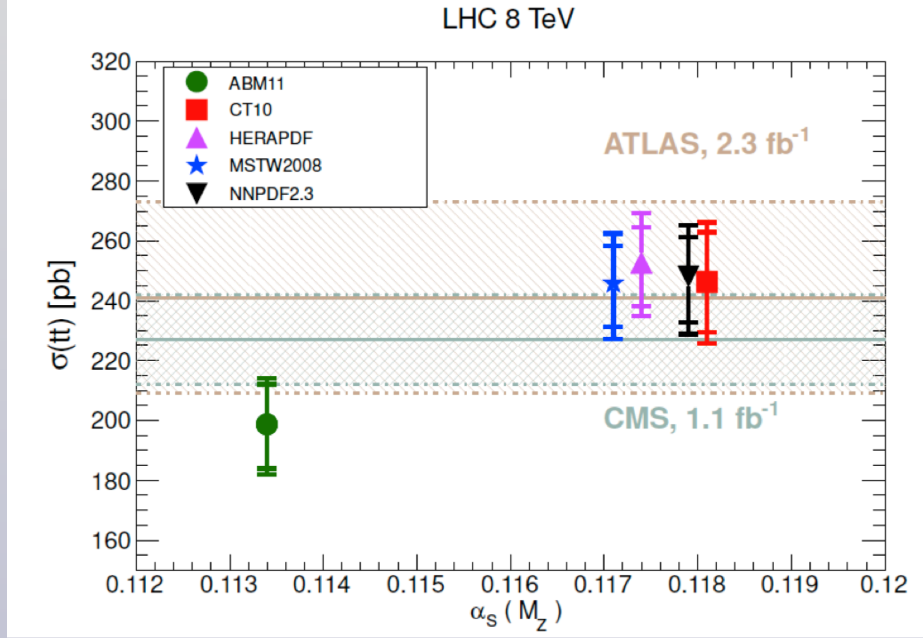

```
MG5_aMC>  generate p p > z b b~ [QCD]
```

- Example of a quite complicated final state that benefits from the complete automation of fast NLO calculations in **aMCfast**
- Calculation needs to be performed in a scheme with massive b quark
- In this case we have 13 independent PDF lumis

l	n_{rs}	(r,s)	
1	1	(g,g)	
2	2	$(\bar{s},g)$	$(\bar{d},g)$
3	2	(d,g)	(s,g)
4	2	$(\bar{c},g)$	$(\bar{u},g)$
5	2	(u,g)	(c,g)
6	2	$(g,\bar{s})$	$(g,\bar{d})$
7	2	(g,d)	(g,s)
8	2	$(g,\bar{c})$	$(g,\bar{u})$
9	2	(g,u)	(g,c)
10	2	$(u,\bar{u})$	$(c,\bar{c})$
11	2	$(d,\bar{d})$	$(s,\bar{s})$
12	2	$(\bar{c},c)$	$(\bar{u},u)$
13	2	$(\bar{s},s)$	$(\bar{d},d)$

PDF dependence of the top cross section





- Most PDF sets provide a **good quantitative description** of Tevatron and LHC top data
- **ABM11** is systematically lower than **other PDF sets** and than **experimental data**

χ^2: EXP errors only

$$\chi^2 = \sum_{i=1}^{N_{\rm dat}} \frac{\left(\sigma_{t\bar{t}}^{(\rm exp)} - \sigma_{t\bar{t}}^{(\rm th)}\right)^2}{\delta_{\rm tot}^{(\rm exp)2}}$$

Pull: EXP + TH errors

$$P = \frac{1}{N_{\rm dat}} \sum_{i=1}^{N_{\rm dat}} \frac{\left(\sigma_{t\bar{t}}^{(\rm exp)} - \sigma_{t\bar{t}}^{(\rm th)}\right)^2}{\delta_{\rm tot}^{(\rm exp)2} + \delta_{\rm tot}^{(\rm th)2}}$$

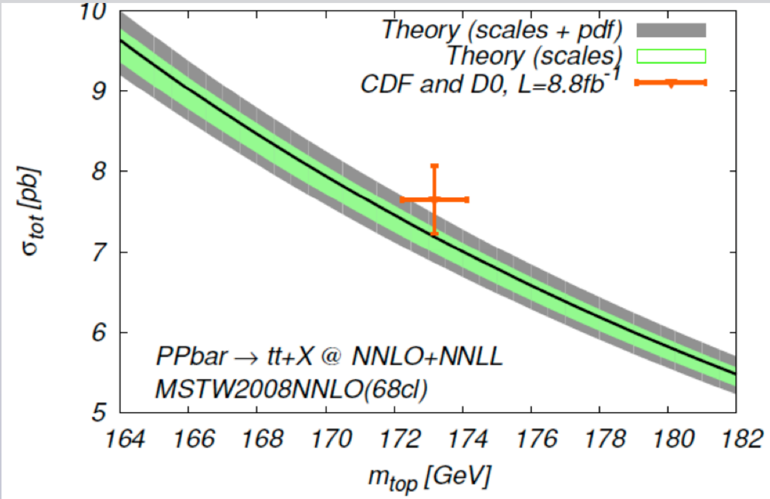
	$\chi^2_{\rm tev}$	$\chi^2_{\rm lhc7}$	$\chi^2_{\rm lhc8}$	$\chi^2_{\rm tot}$	$\chi^2_{\rm tot}/N_{\rm dat}$	P
AMB11	3.5	31.4	5.3	40.2	8.0	3.2
CT10	0.4	3.3	1.7	5.3	1.1	0.3
HERAPDF15	0.0	6.1	3.1	9.2	1.8	0.5
MSTW08	1.3	3.1	1.6	6.0	1.2	0.4
NNPDF2.3	0.9	3.4	2.0	6.3	1.3	0.4

LHC top data already discriminates between PDF sets!

Dependence on the top quark mass

- Compare total theory uncertainty **with** and **without top quark mass uncertainty**

- Thanks to the improvement of the NNLO calculation, now all theory uncertainties of similar size, only **mild reduction (< 1.5%)** in the total theory errors if one assumes that $\boldsymbol{\delta_{mt} \approx 0}$



Collider	σ_{tt} (pb)	$\delta_{\mathrm{PDF+scales}+\alpha_s}$ (pb)	δ_{tot} (pb)
Tevatron	7.258	$+0.267$ ($+3.7\%$) -0.352 (-4.9%)	$+0.390$ ($+5.4\%$) -0.469 (-6.5%)
LHC 7 TeV	172.7	$+10.4$ ($+6.0\%$) -11.8 (-6.8%)	$+12.5$ ($+7.2\%$) -13.7 (-8.0%)
LHC 8 TeV	248.1	$+14.0$ ($+5.6\%$) -16.2 (-6.5%)	$+17.1$ ($+6.9\%$) -19.1 (-7.7%)
LHC 14 TeV	977.5	$+44.1$ ($+4.5\%$) -55.8 (-5.7%)	$+57.4$ ($+5.9\%$) -68.5 (-7.0%)

Dependence on the strong coupling

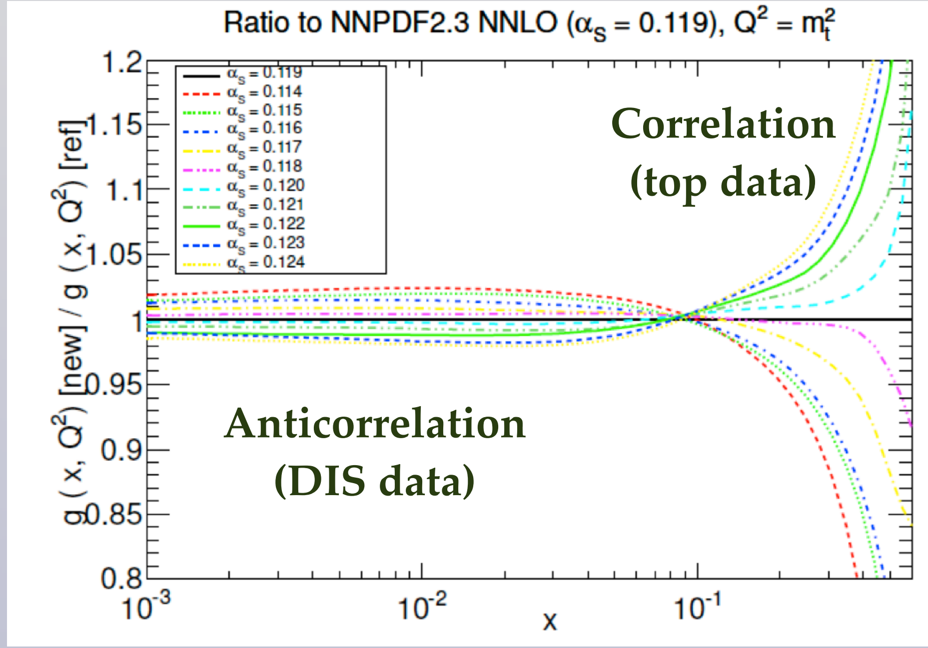
- The same reasons that suggest the use of top data for PDFs motivate the **extraction of the strong coupling** $\alpha_S(M_Z)$ from the total cross-section

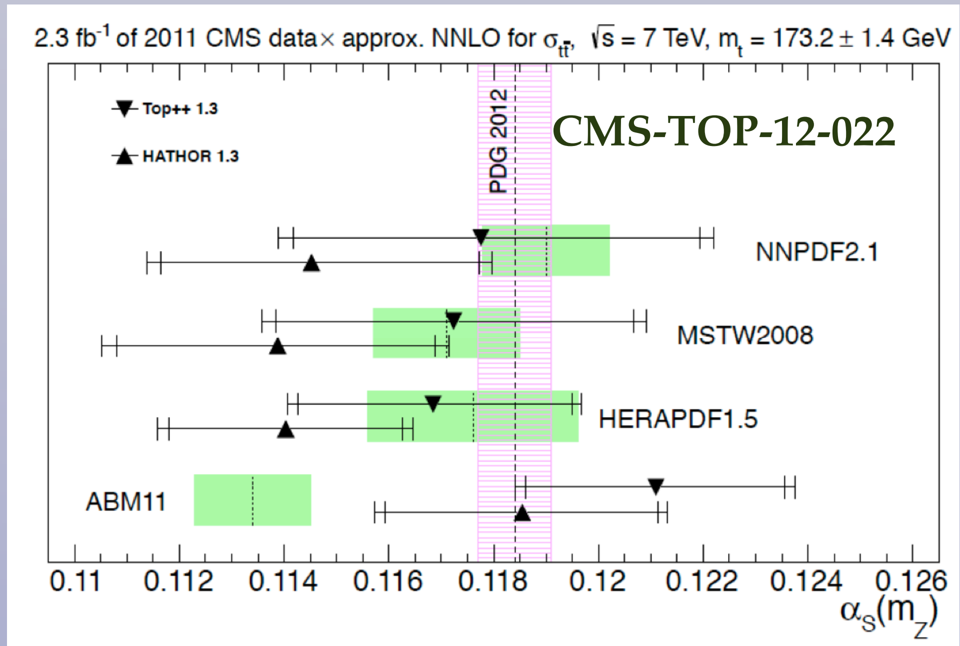
    Small scale uncertainties at NNLO

    Reduced non-perturbative corrections as compared to jets

- The dependence between σ(top) and $\alpha_S(M_Z)$ is enhanced as compared to the naive power counting because of **correlation between** $\alpha_S(M_Z)$ **and the gluon at large-x**

- First determination by **CMS** based on approximate NNLO, results based on full NNLO in preparation





Pinning down the gluon with top data

- Adding data from lower energy colliders: reduced theory uncertainties at higher energies
- Adding **TeV+LHC7** data to NNPDF2.3, we obtain the **best possible theory prediction for LHC8**
- Not only PDF uncertainty reduced, also central value **shifts** to increase agreement with data

Collider	Ref	Ref+TeV	Ref +TeV+LHC7	Ref+TeV+LHC7+8
Tevatron	7.26 ± 0.12	-	-	-
LHC 7 TeV	172.5 ± 5.2	172.7 ± 5.1	-	-
LHC 8 TeV	247.8 ± 6.6	248.0 ± 6.5	245.0 ± 4.6	-
LHC 14 TeV	976.5 ± 16.4	976.2 ± 16.3	969.8 ± 12.0	969.6 ± 11.6

PDF uncertainty only

- Using TeV+LHC7 data, optimal fit description for LHC8
- The **precise LHC7** data carry most of the information, but full 8 TeV analysis still missing

Collider	χ^2 (Total, $N_{\rm dat} = 5$)	χ^2 (LHC 8 TeV, $N_{\rm dat} = 2$)
NNPDF2.3	6.28	1.64
NNPDF2.3 + TeV,LHC data	4.88	1.24
NNPDF2.3 + TeV,LHC7 data	4.87	1.24