Top quark pair differential distributions with aMC@NLO: constraints on the gluon and PDFs for NLO event generators

Valerio Bertone

CERN and Albert-Ludwigs-Universität Freiburg

PDF4LHC

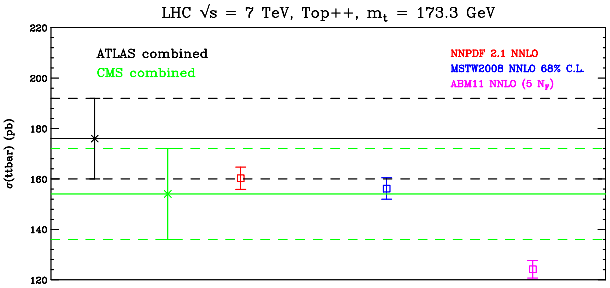
May 23$^{\mathrm{th}}$ 2012, CERN

In collaboration with:
S. Frixione, J. Rojo and P. Torrielli

Outline

Motivation: why $t\bar{t}$?

- Potentially **discriminating** between different PDF sets:



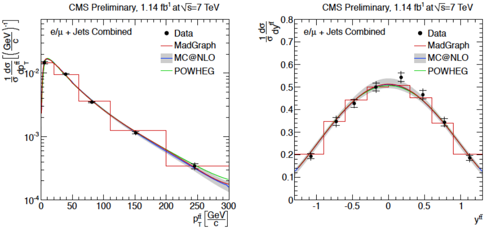
[Reference: Francesco Spanó: arXiv:1112.3906]

- Provide a strong constraining on the **Gluon PDF** at large values of x:

$$\frac{\sigma_{gg}(t\bar{t})}{\sigma_{\rm tot}(t\bar{t})} \sim 80\% \quad \text{at LHC 7 TeV}$$

Motivation: why $t\bar{t}$?

- $t\bar{t}$ **inclusive** and **differential** cross sections measured at LHC:
  - CMS:



[Reference: CMS PAS TOP-11-013]

  - ATLAS in progress,
  - Sizable amount of data to use to perform PDF fits.
- Exact fully **NNLO calculation** soon available:
  - $q\bar{q} \rightarrow t\bar{t} + X$ already present,

    [Reference: Bärnreuther, Czakon and Mitov, arXiv:1204.5201]
  - $gg, gq, qq' \rightarrow t\bar{t} + X$ in preparation.

aMC@NLO in a nutshell

aMC@NLO:
automatic tool that is flexible and allows to compute any process at **NLO accuracy**, including matching to the **parton shower**.

How it works:

- **MadFKS** computes all the NLO contributions, except the finite part of the virtual amplitude.
- **MadLoop** computes the virtual corrections using CutTools:
  - if the analytical virtual correction is available (like $t\bar{t}$ production itself), possibility to skip MadLoop.
- **Combine MadFKS and MadLoop** to get any observable at NLO accuracy.
- Add terms to **remove double counting** when matching to the parton shower *á la* MC@NLO.
- **Shower** the generated events using HERWIG or PYTHIA to get **fully exclusive** predictions at NLO accuracy.

Why aMC@NLO

- It allows to use **identical analysis procedures** as the ones applied to data:
  - application of kinematical cuts,
  - jet reconstruction,
  - ...
  - **direct comparison with data**.

- Consistent computation of correction factors from hadron to parton level:
  - a **common framework** for all processes.

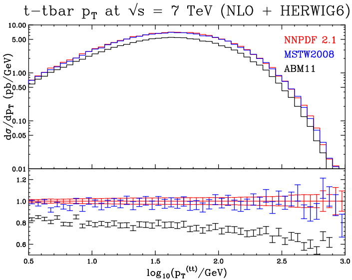
- Possibility to compute **scale variation and PDF uncertainties** by means of the reweighting method:
  - no extra cost,
  - available at both NLO and NLO+PS level.

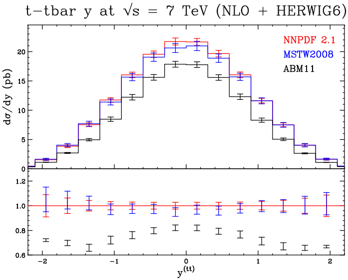
Differential Distributions with aMC@NLO: the setup

- Generation of $t\bar{t}$ **events at LHC** $\sqrt{s} = 7$ TeV:
  - pure NLO,
  - NLO + PS with **HERWIG6**,
  - NLO + PS with **PYTHIA6 virtuality ordered**.
- Study of all the relevant **differential distributions**:
  - **top** and **bottom** distributions,
  - **lepton** distributions.
- Comparison of the results for **different PDF sets** with respective uncertainties:
  - NNPDF 2.1
  - MSTW2008
  - AMB11
- Assessment of which distributions are **more sensitive to the shower effects**.

Differential Distributions with aMC@NLO: the impact of PDFs

- The LHC $t\bar{t}$ **differential distributions** could be potentially **discriminating** for the PDF sets.



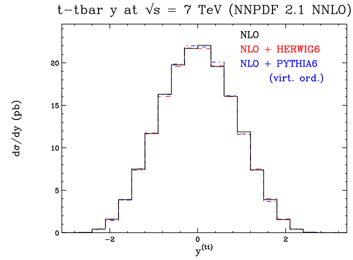


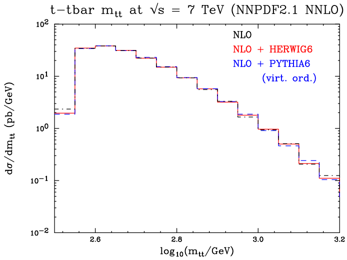
- Differences between NNPDF/MSTW and ABM of **order 20%**:
  - same trend as in the inclusive cross section.

Differential Distributions with aMC@NLO: the effect of the Parton Shower

- Impact of the **parton shower** on the **differential distributions**,
  - **inclusive** cross section **unchanged** (unitarity of the parton shower),
  - difference in the shape of the **differential distributions**.

- Comparison of NLO and NLO+PS predictions:

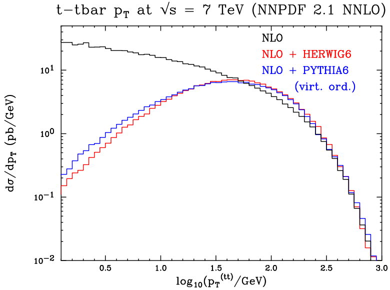




Generally moderate impact of the PS ...

Differential Distributions with aMC@NLO: the effect of the Parton Shower

- Some distributions present a **stronger dependence** on the PS:



$\Rightarrow$ the inclusion of these data in a PDF fit **requires NLO+PS** accuracy.

- Presently available PDF sets do not include PS effects:
  - conceptual **inconsistency** when used with NLO MC generators (MC@NLO, POWHEG).
- aMC@NLO could be used to provide a **PDF set including PS effects**:
  - to be used with **NLO MC generators**.

How to estimate the impact on PDFs

Question:

Can we **estimate the impact** of the inclusion of the $t\bar{t}$ differential distributions on PDFs?

Exercise:

- Generate **pseudo data** of $t\bar{t}$ differential cross sections (NLO+PS) using **NNPDF2.1** and **ABM11** central values and use the statistical uncertainty of 5 fb^{-1} of data as error.
- **Reweight** the NNPDF 2.1 NNLO set with those pseudo data:
  - NNPDF2.1: assess the impact on a **given PDF set**,
  - ABM11: assess how **discriminating** these distributions are with respect to different PDF sets.

Reweighting of PDFs in a nutshell

It allows to **incorporate a new set of data** points $\mathcal{O}_i$ (and covariance matrix σ_{ij}), with $i = 1, \ldots, n$, in a set of PDFs $\{f_k\}$, with $k = 1, \ldots, N_{rep}$, **without need of refitting**.

- Valid for a Monte Carlo representation of PDFs like the NNPDF sets,
  [References: The NNPDF Collaboration, arXiv:1108.1758 and arXiv:1012.0836]
- recently made available also for MSTW2008.
  [Reference: Watt and Thorne, arXiv:1205.4024]

Procedure:

- Evaluate the χ^2 of the dataset for each PDF replica:

$$\chi^2_k = \frac{1}{n} \sum_{i,j}^{n} (\mathcal{O}_i - \mathcal{O}_i[f_k]) \sigma^{-1}_{ij} (\mathcal{O}_j - \mathcal{O}_j[f_k])$$

- Evaluate the weights:

$$w_k \propto \chi_k^{n-1} \exp\left[-\frac{1}{2}\chi^2_k\right] \quad \text{with} \quad \sum_{k=1}^{N_{rep}} w_k = N_{rep}$$

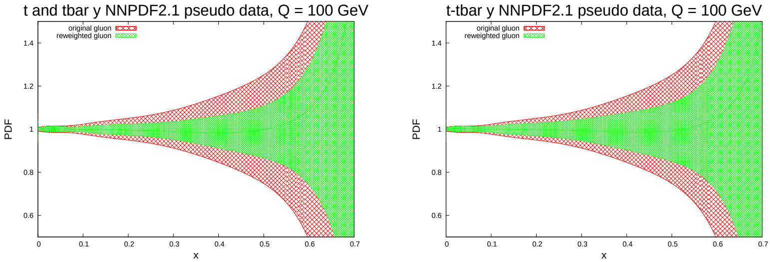
- Compute the expectation value and variance of any observable $\mathcal{P}$ like:

$$\langle \mathcal{P} \rangle = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} w_k \mathcal{P}[f_k] \quad \text{and} \quad \sigma^2_{\mathcal{P}} = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} (\langle \mathcal{P} \rangle - w_k \mathcal{P}[f_k])^2$$

NNPDF2.1 pseudo data

Reweighting with NNPDF2.1 NNLO pseudo data:

- only statistical and **no systematic error** (best case).



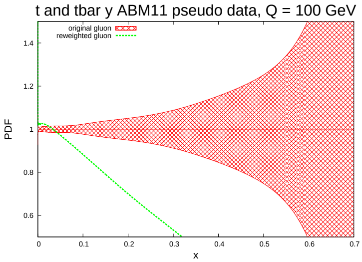
Sizable impact on the **gluon distribution**:

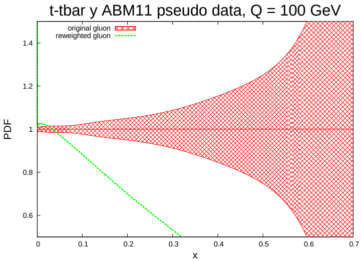
- central value **stable** (self-consistency of the procedure),
- shrinking of the error at large **values of** x.

ABM11 pseudo data

Reweighting with ABM11 NNLO pseudo data:

- only statistical and **no systematic error** (best case).





Strong incompatibility of ABM11 with NNPDF2.1:

- only **one replica** survived the reweighting procedure,
- $t\bar{t}$ differential distributions **strongly discriminanting**.

Conclusions

- aMC@NLO extremely useful toolbox for the LHC phenomenology.
- Impact of the PS on the $t\bar{t}$ distributions:
  - generally moderate $\Rightarrow$ present PDF fits are a good approximation,
  - some differential distributions (e.g. the p_T of the pair $t\bar{t}$) are strongly dependent on the PS $\Rightarrow$ need for the PS to be included in fits.
- Impact of the $t\bar{t}$ distributions on the PDFs:
  - constraint on the gluon PDF at large x,
  - LHC data could be able to discriminate between PDF sets.