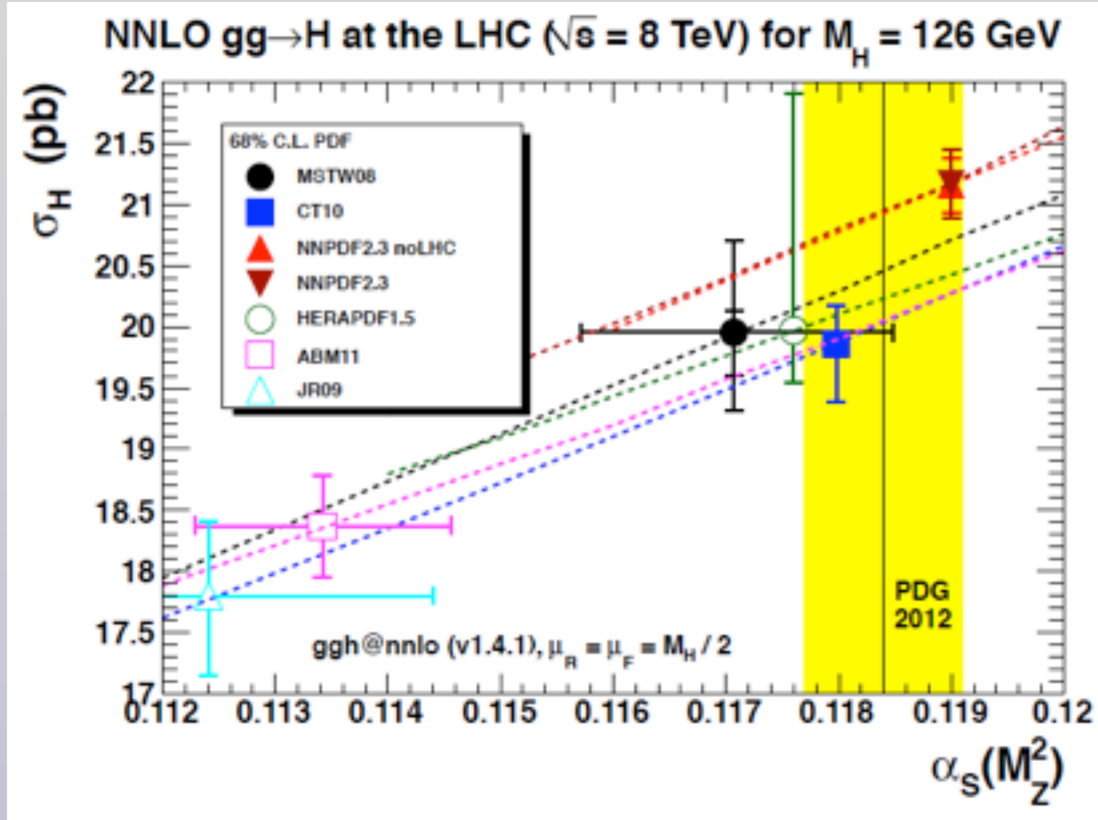
Neural Network PDFs and the W mass at the LHC

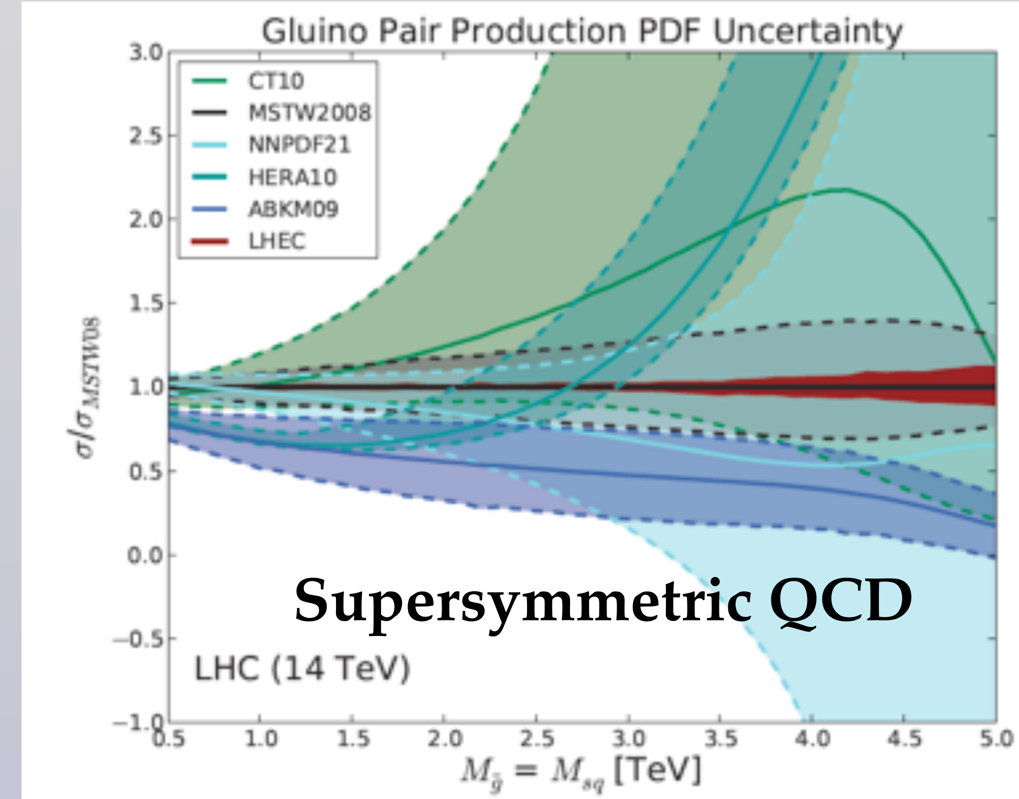
Juan Rojo
CERN, PH Division, TH Unit

Snowmass Electroweak Workshop
Duke University
19/02/2013

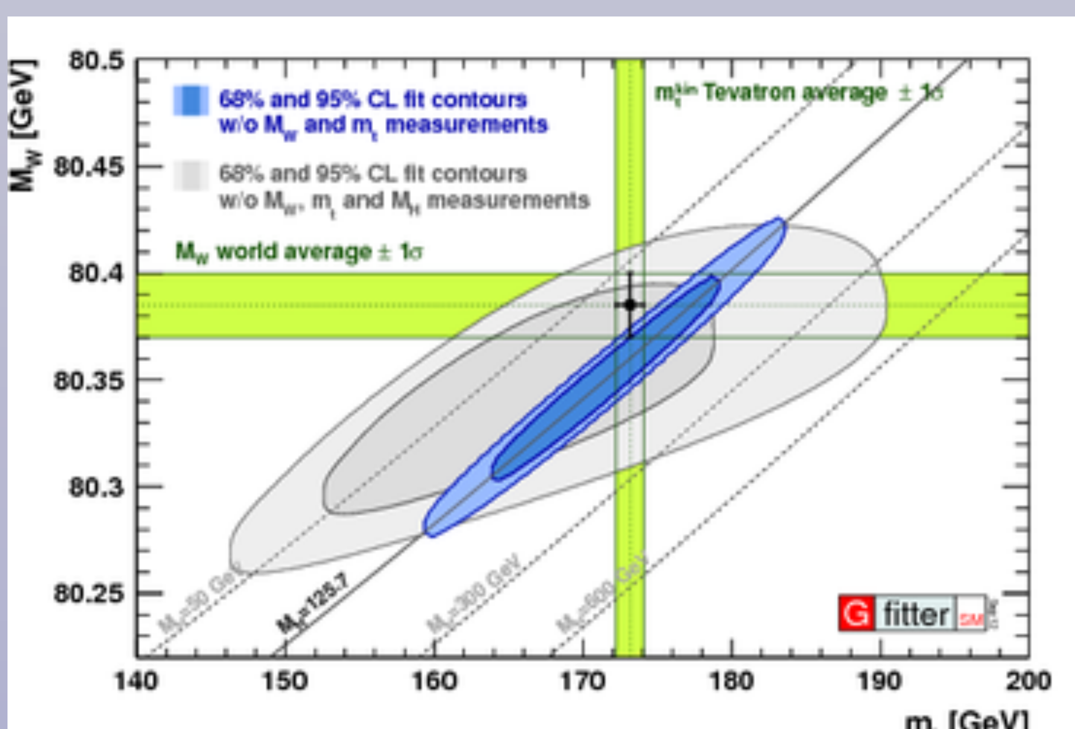
Parton Distributions and LHC phenomenology



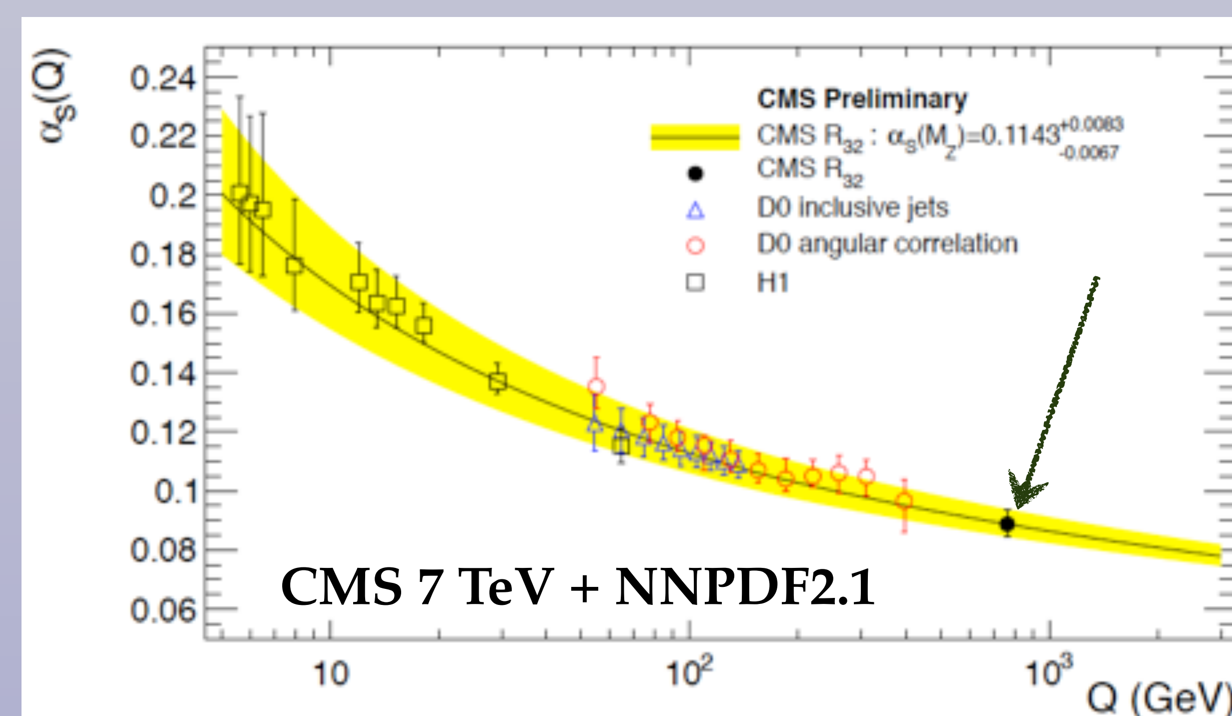
1) PDFs fundamental limit for Higgs production characterization in terms of cross sections and couplings



2) Very large PDF uncertainties for any New Physics model at large masses: supersymmetry, extra dimensions,...



3) PDFs dominant systematic for precision measurements, like W boson mass, that test consistency of the Standard Model



4) Accurate PDFs required for indirect searches of New Physics in the TeV region, like deviations from QCD evolution

Global QCD analysis

PDF determination is based on a **global analysis of hard scattering data** to extract, thanks to the factorization theorem, **universal PDFs for LHC predictions**

Experimental data → **Lepton-proton structure functions**, heavy quark production, jet production, Drell-Yan pair production, isolated photons, **LHC data itself**, ...

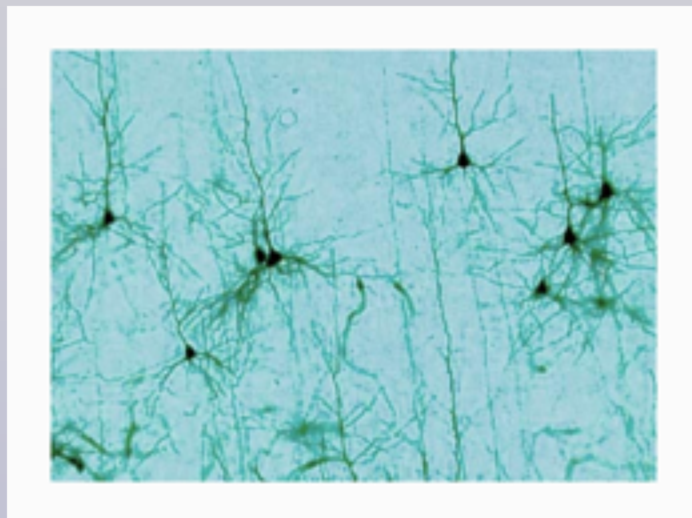
QCD Theory → **DGLAP evolution, hard scattering cross sections**, heavy quark treatment, strong coupling, **electroweak and QED effects**, NNLO calculations, ...

Methodology → *Ansatze* **for the x-dependence of PDFs $q(x,Q_0)$**, propagation of **experimental errors from data to PDFs**, model uncertainties, statistical analysis, ...

Parton Distributions → **LHC phenomenology**, predictions for LHC processes, PDF and Higgs production, **improving New Physics searches**, ...

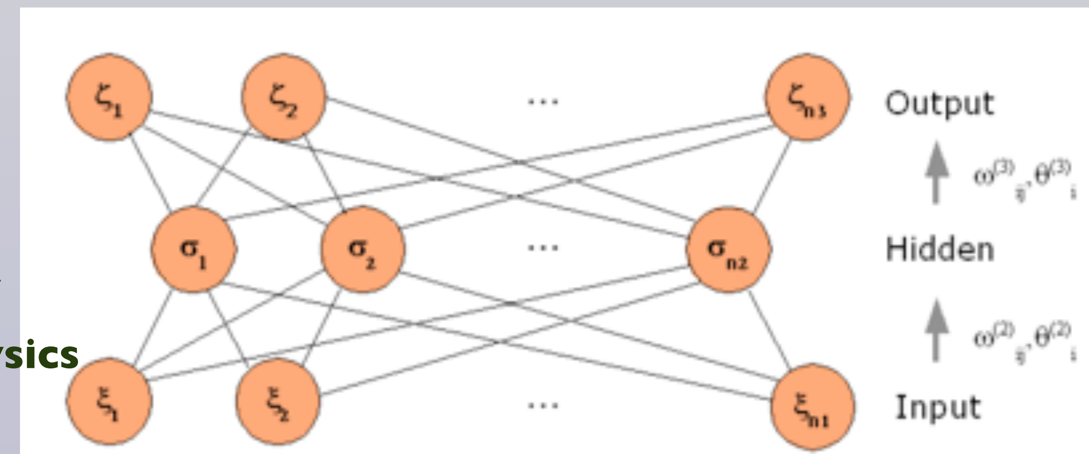
(Artificial) Neural Network PDFs

Inspired by **biological brain models**, **Artificial Neural Networks** (ANNs) are **mathematical algorithms** widely used in a wide range of applications, from **high energy physics** to **targeted marketing** and **finance forecasting. ANNs** excel in same domains as their biological counterparts: **pattern recognition, forecasting, classification**, ... where our **evolution-driven biology** outperforms traditional algorithms



from biology...

... to high energy physics



- In the NNPDF approach, **ANNs** provide **universal unbiased interpolants** to parametrize non-perturbative PDF dynamics, and **learn the underlying physical laws** from experimental data
- **No theory bias** introduced in the PDF determination by the choice of *ad-hoc* functional forms
- **Faithful extrapolation**: PDF errors **blow up** in regions with scarce data, crucial for LHC searches

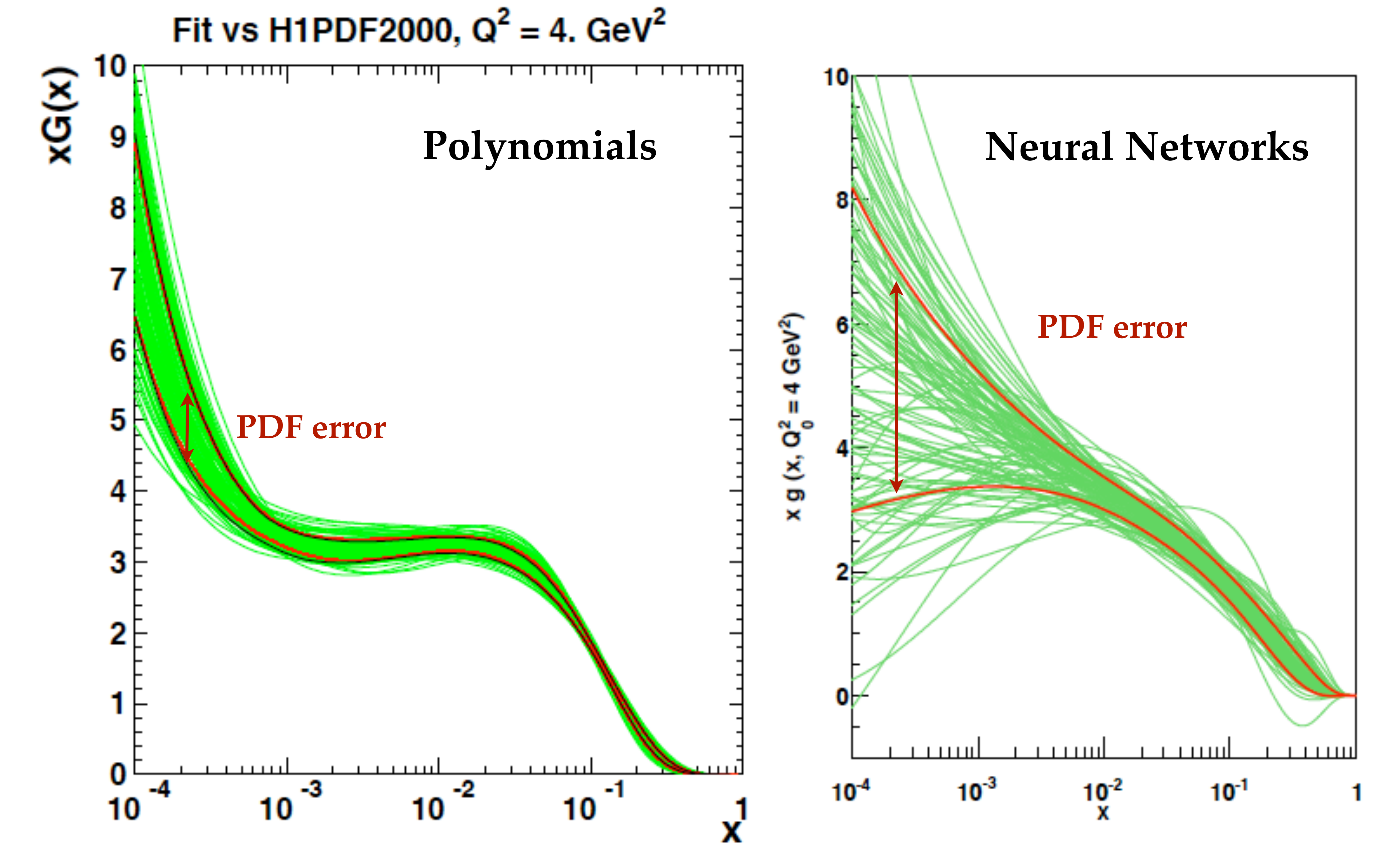
NNPDF approach

$$\begin{aligned}\Sigma(x,Q_0^2) &= (1-x)^{m_\Sigma}x^{-n_\Sigma}\mathrm{NN}_\Sigma(x)\\ g(x,Q_0^2) &= A_g(1-x)^{m_\Sigma}x^{-n_\Sigma}\mathrm{NN}_g(x)\end{aligned}$$

Traditional approach

$$\begin{aligned}\Sigma(x,Q_0^2) &= (1-x)^{m_\Sigma}x^{-n_\Sigma}\left(1+a_\Sigma\sqrt{x}+b_\Sigma x+\ldots\right)\ ,\\ g(x,Q_0^2) &= A_g(1-x)^{m_\Sigma}x^{-n_\Sigma}\left(1+a_g\sqrt{x}+b_g x+\ldots\right)\end{aligned}$$

Neural Network PDFs Learning from Data



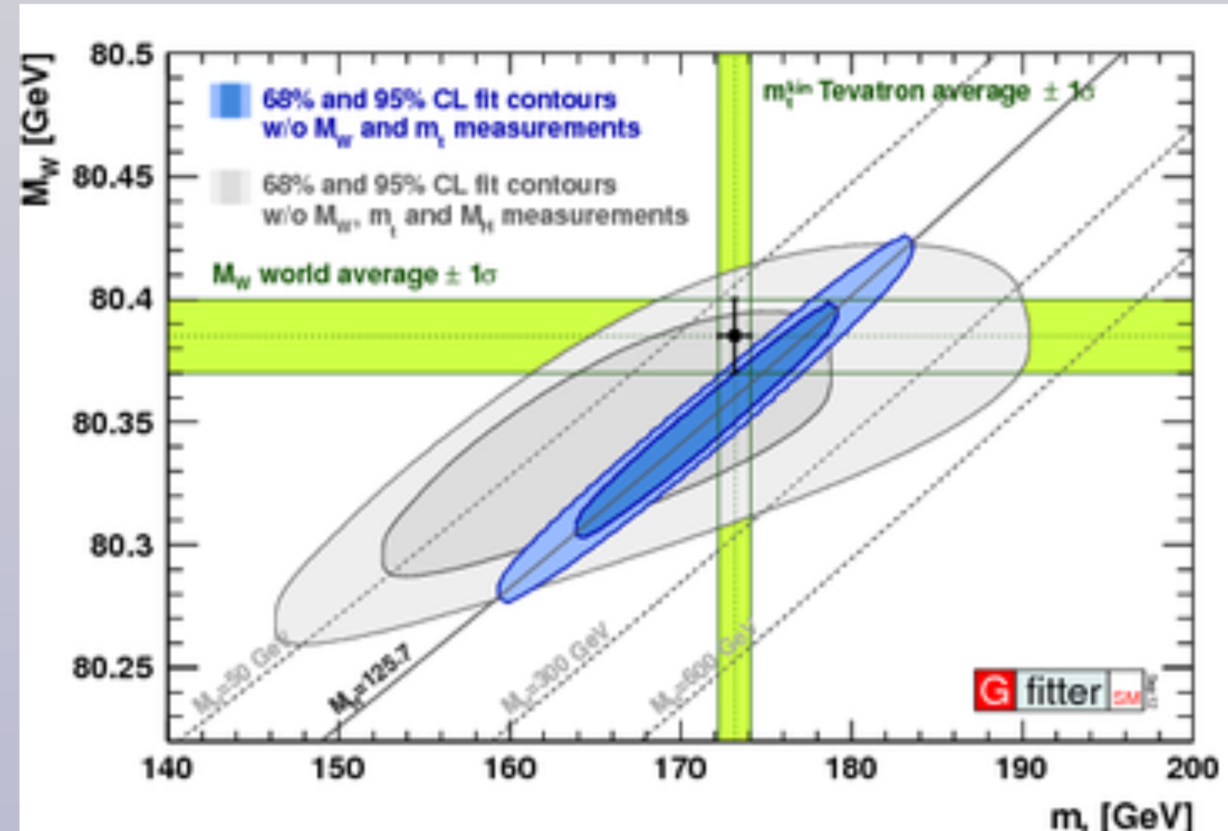
Enhanced flexibility crucial for accurate extrapolations at small and large-x

W mass at hadron colliders

- **The W mass** measurement is one of the legacy analysis of the Tevatron
- **LEP-like precision**, dominated by **theory systematics,** in particular **Parton Distributions**
- Stringent tests on the **consistency of the Standard Model**, indirect bounds on the Higgs boson mass

New CDF Result (2.2 fb^{-1})
Transverse Mass Fit Uncertainties

	electrons	muons
W statistics	19	16
Lepton energy scale	10	7
Lepton resolution	4	1
Recoil energy scale	5	5
Recoil energy resolution	7	7
Selection bias	0	0
Lepton removal	3	2
Backgrounds	4	3
pT(W) model	3	3
Parton dist. Functions	10	10
QED rad. Corrections	4	4
Total systematic	18	16

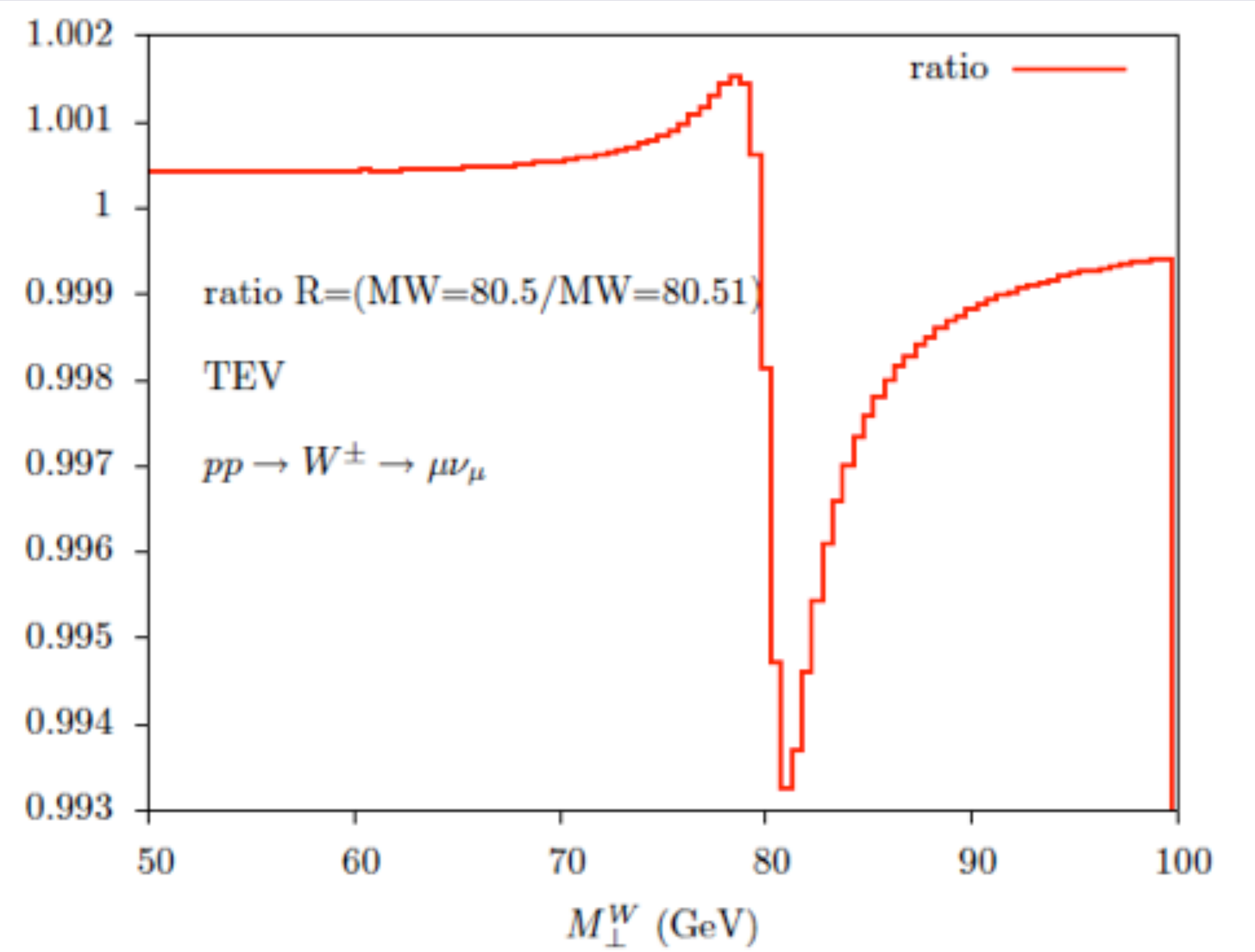


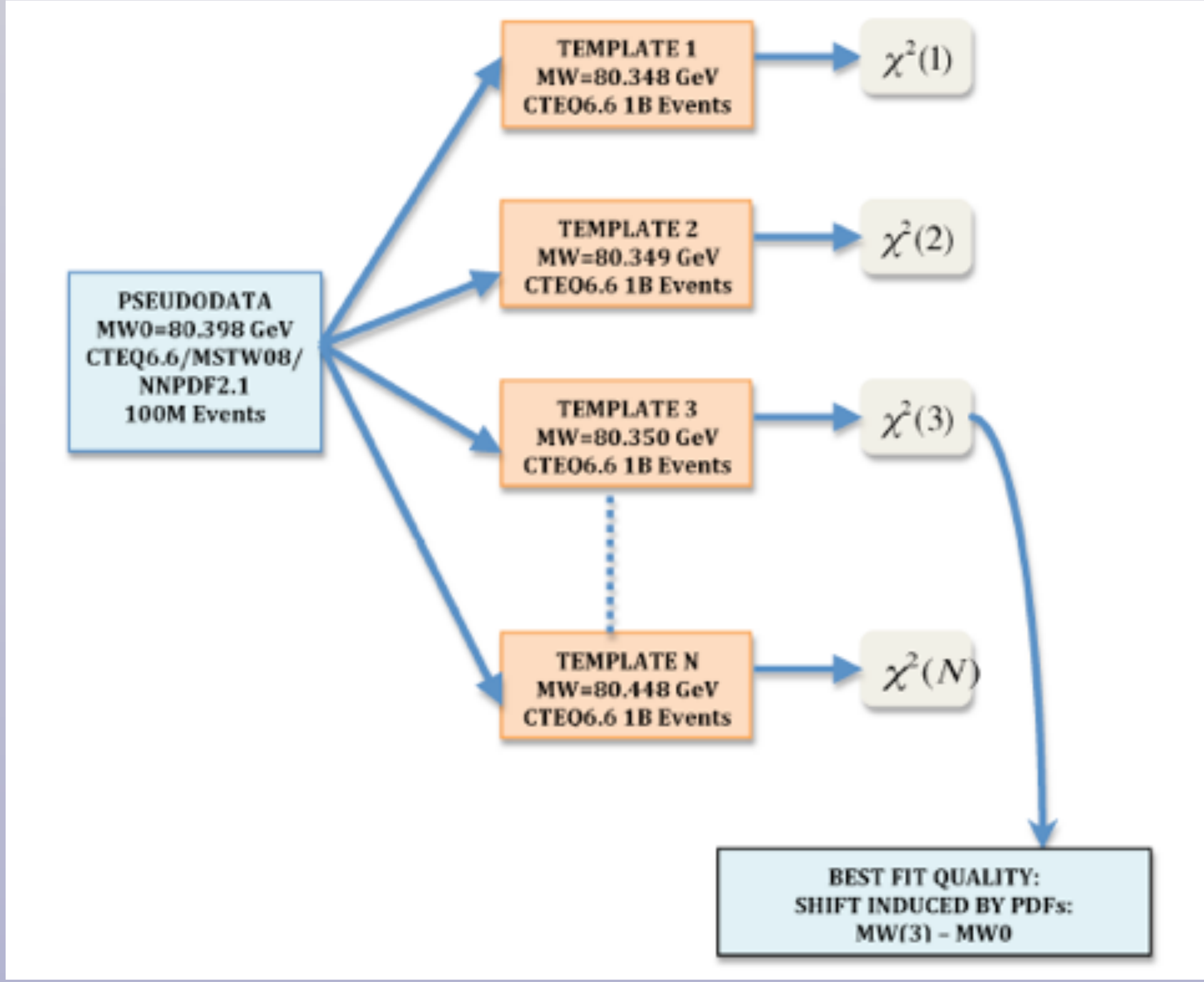
Is it feasible to achieve the **same precision/improve at the LHC**? Not everyone agrees, see for example **Krasny et al, arXiv:1004.2597:**

$\Delta M_W \leq 10$ MeV/c^2 at the LHC: a forlorn hope? †

W determination at the LHC

- The shape of various distributions in W production (**$p_{W,T}$, $M_{W,T}$**) are very sensitive to value of **M_W**
- At the Tevatron, a variation of **10 MeV** in the mass leads to a variation of **7 permille** in the distribution: Need to control **experimental and theory systematics** at similar degree of accuracy to obtain a competitive determination
- **Statistics** not an issue at the LHC: one can use only muons or electrons
- At the LHC **W+ different from W-**: provide two independent and complementary determinations, then **average**

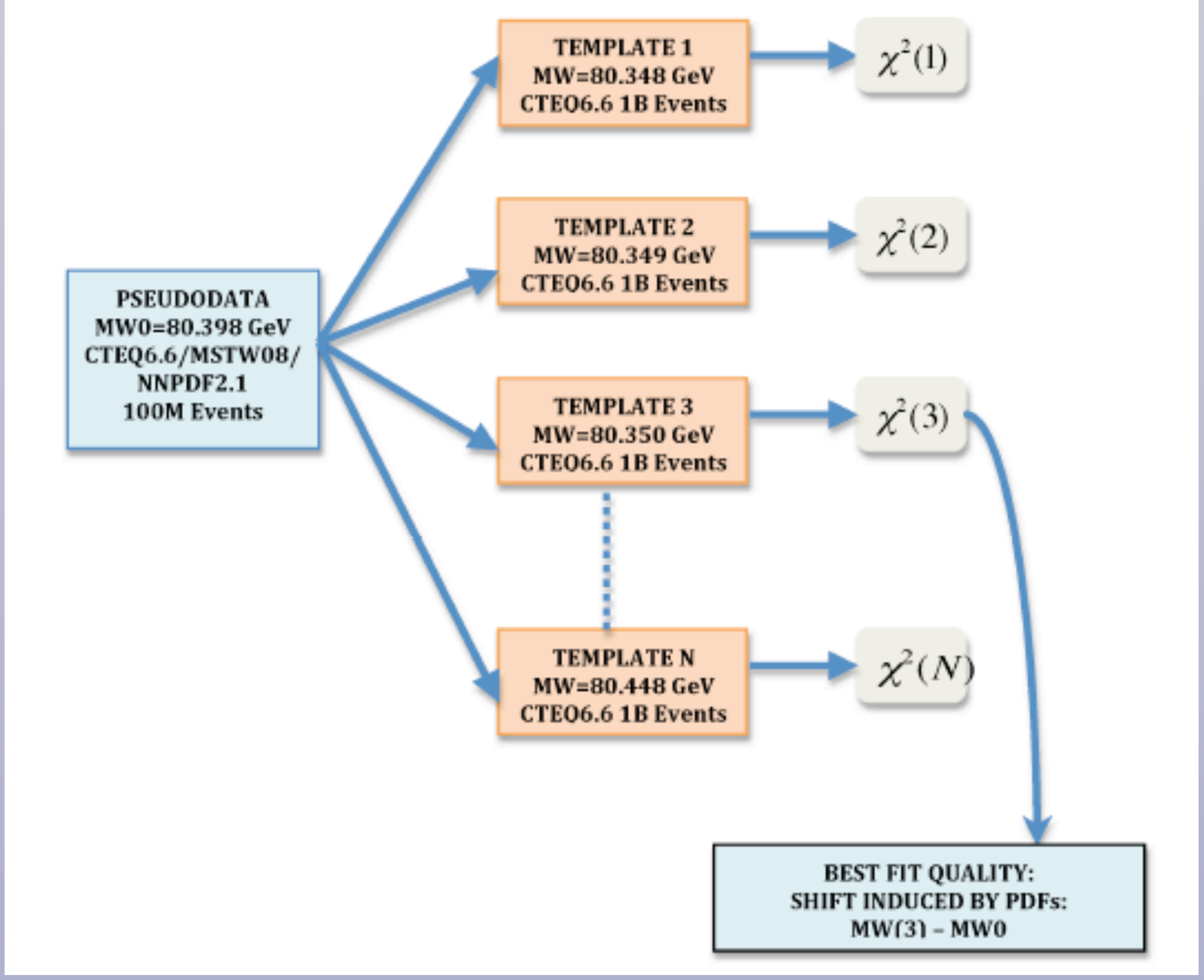




G. Bozzi, J. Rojo and A. Vicini, ``The Impact of PDF uncertainties on the measurement of the W boson mass at the Tevatron and the LHC," Phys. Rev. D83, 113008 (2011)

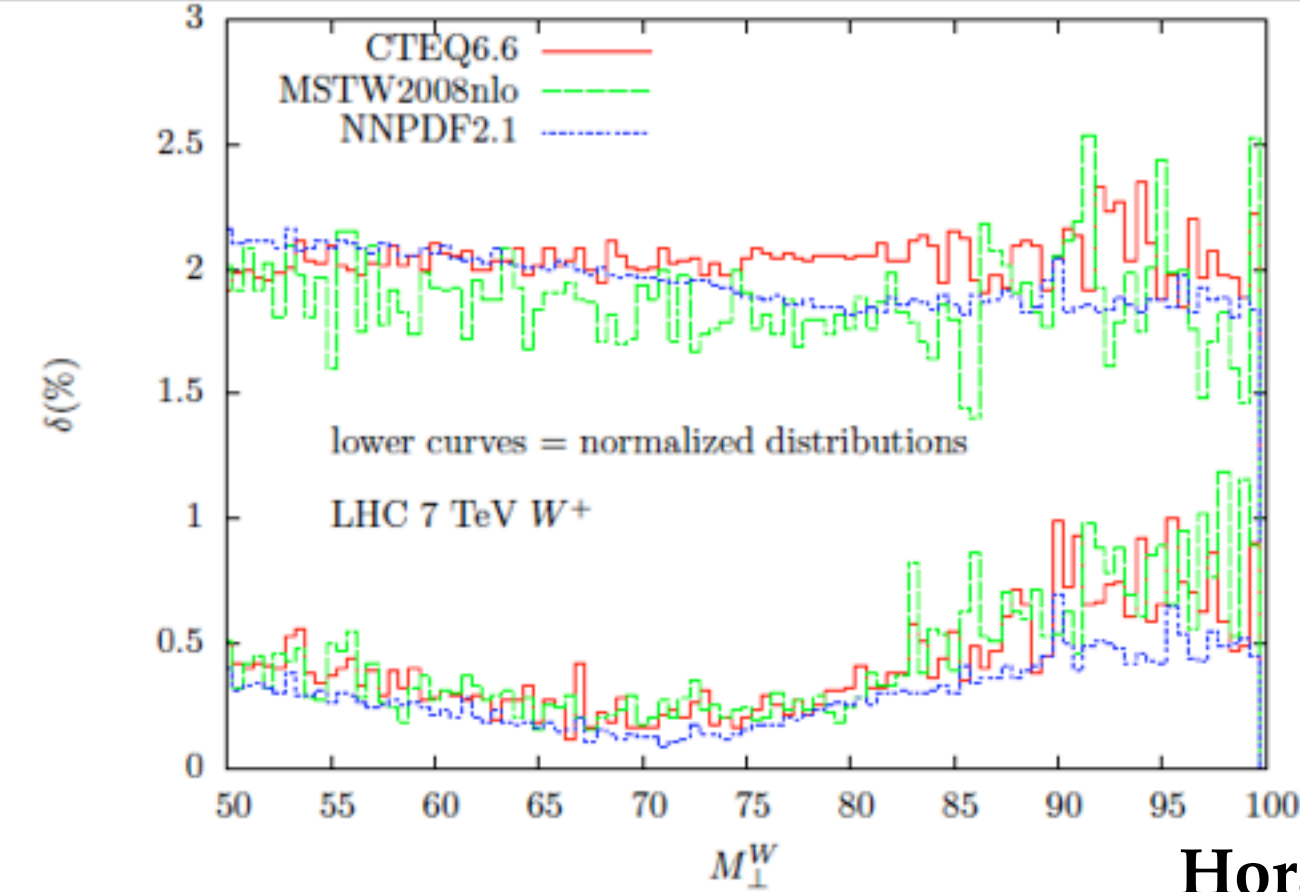
W determination at the LHC

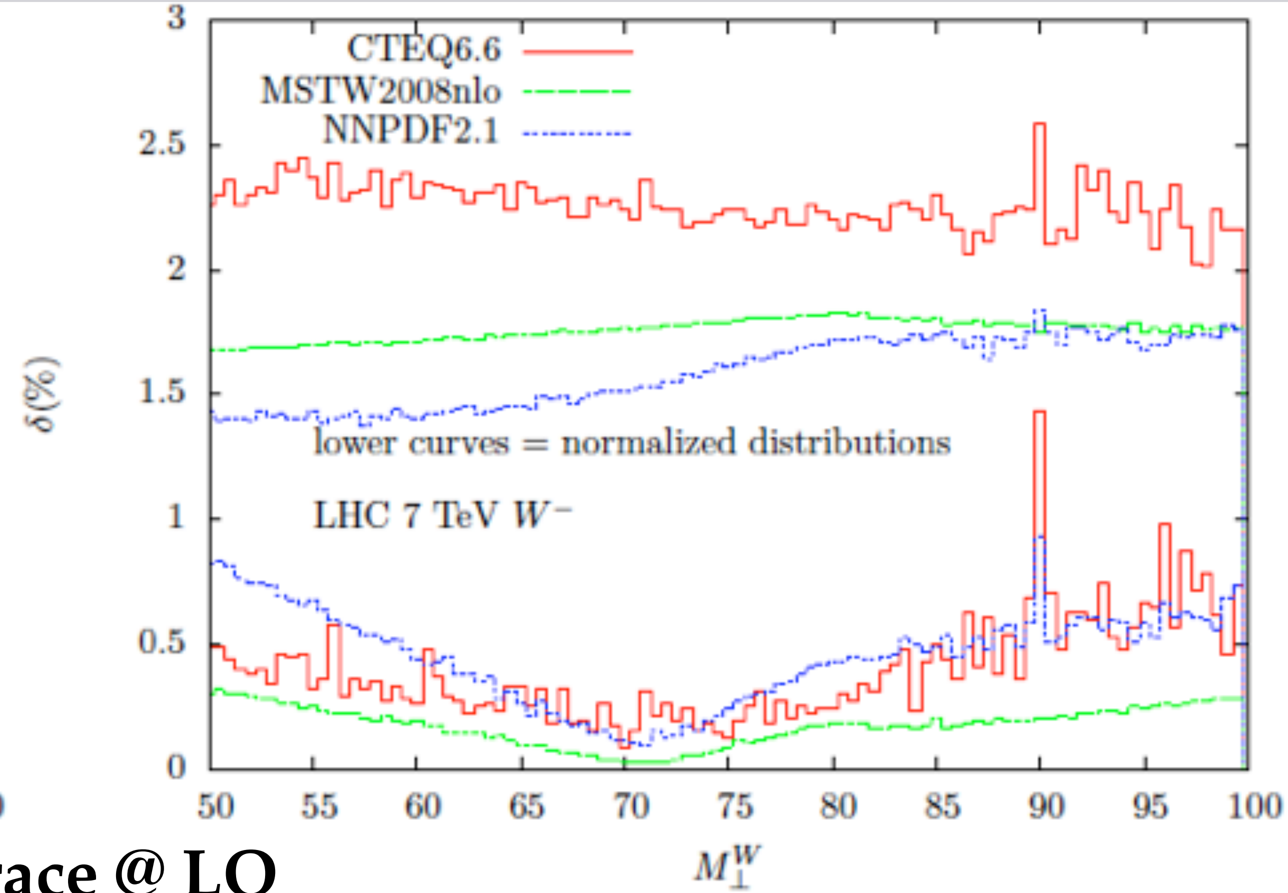
- They key ingredient of the W mass determination is the construction of **templates** for various differential distributions, using the most accurate theory, for a large number of MW values
- Then we **fit experimental data and find the template** that leads to the same agreement
- The procedure is repeated for all **theoretical uncertainties**: PDF variations, scale uncertainties, model parameters....
- An alternative possibility is take the **Z distributions from data** and provide templates for the **W/Z ratios**



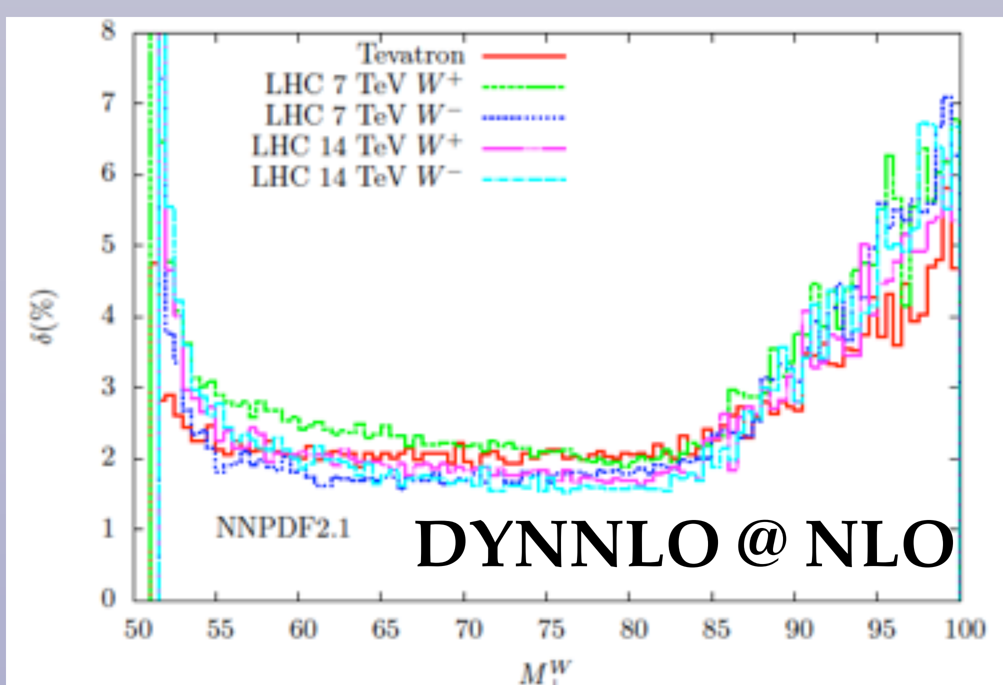
W determination at the LHC

The templates for the **differential W distributions** should be normalized to reduce PDF uncertainties



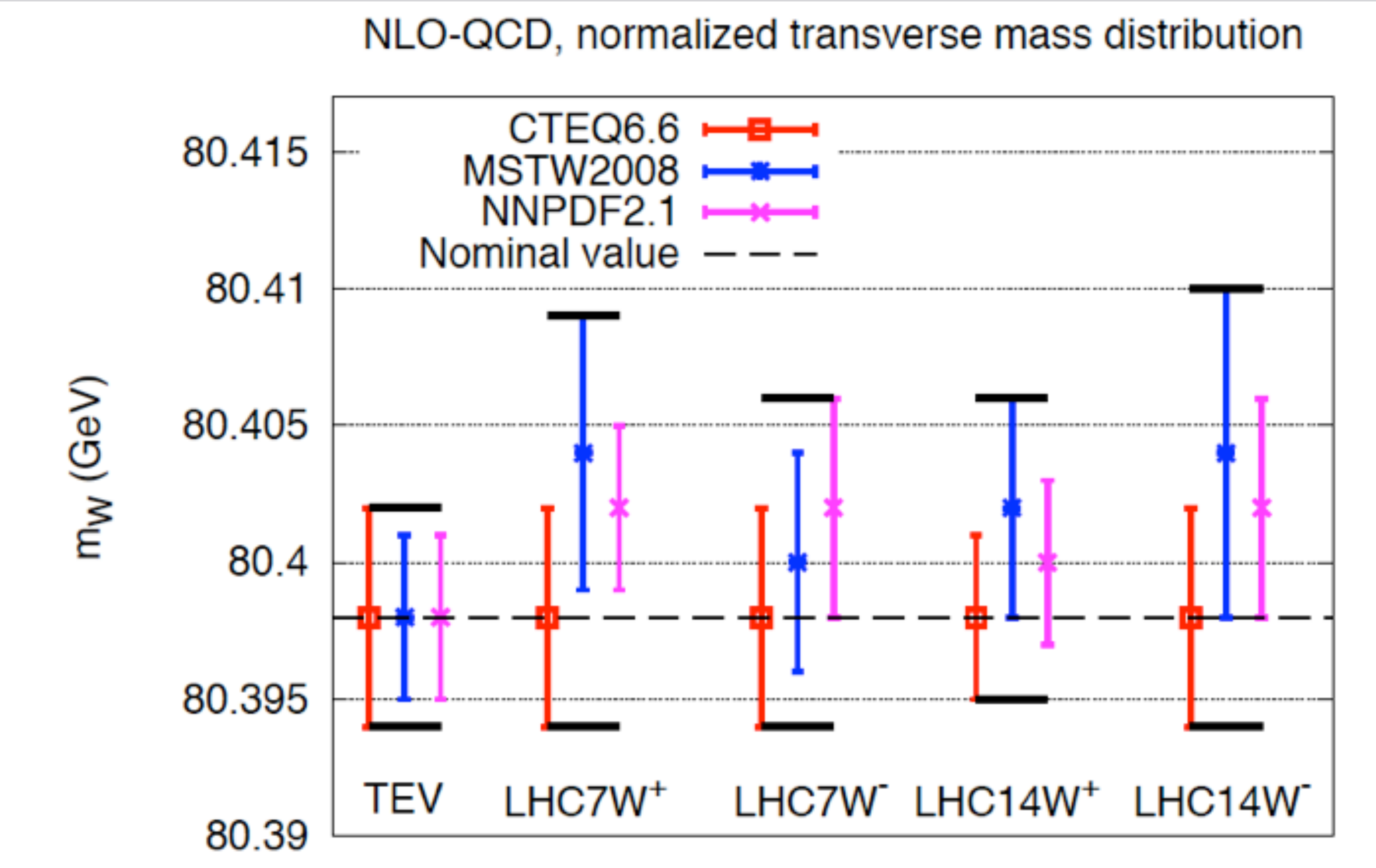


Horace @ LO



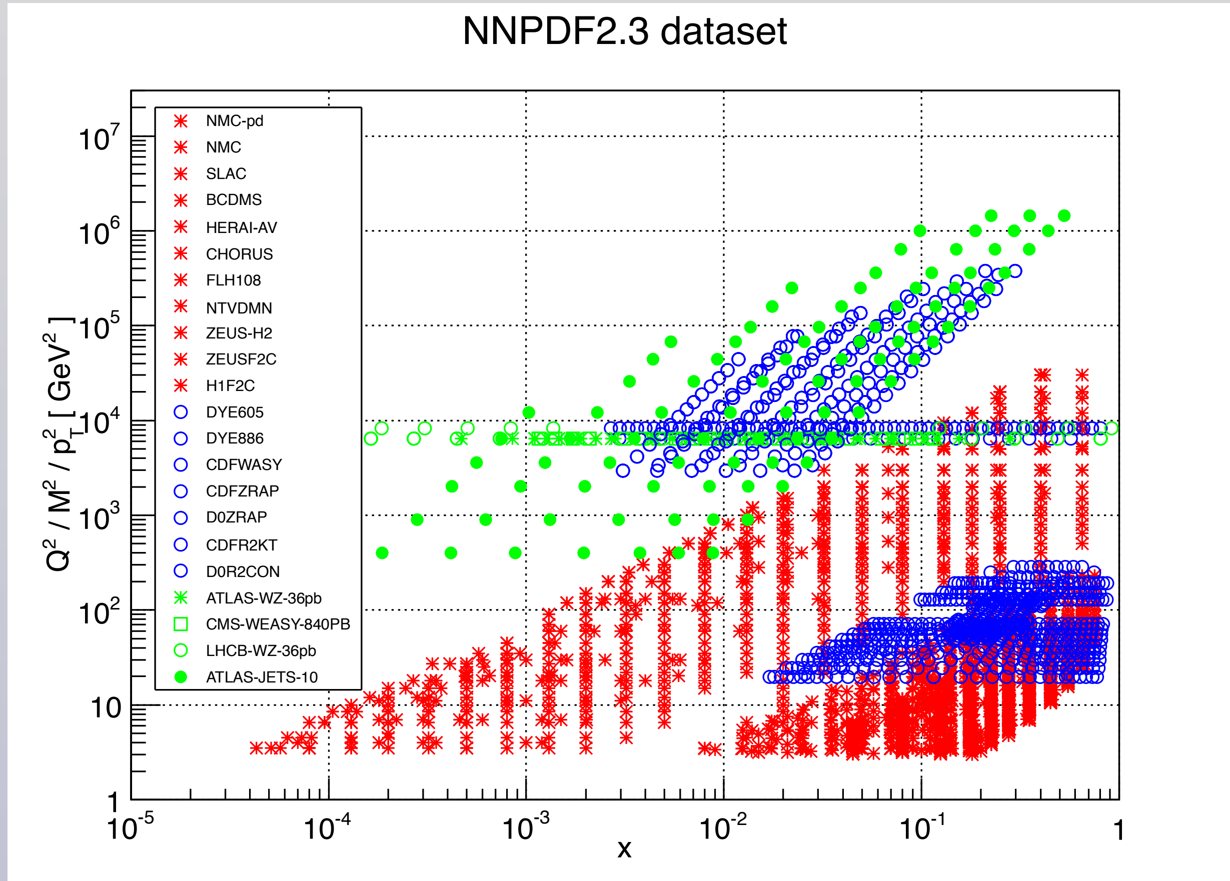
- The dependence on MW in the templates arises from the **shape of the distribution**, not on the **absolute normalization**
- Using normalized distributions in the template fits **reduces substantially PDF uncertainties**, wo removing MW sensitivity
- PDF uncertainties very **similar at Tevatron and at the LHC** at various collider energies

W determination at the LHC



- To provide a **conservative estimate of PDF errors**, we use the **PDF4LHC prescription**: combine in envelope **NNPDF, CT and MSTW**
- We found that a **20 MeV** uncertainty at the LHC was a reasonable estimate. No huge increase of PDF errors from Tevatron to LHC as claimed in the literature
- Our study was based on **parton level templates**, but we checked that a **simple detector-like smearing** did not modify our results qualitatively. Need to check with full simulation
- Variations in α**s** and in the **heavy quark masses** explicitly shown to be negligible

Improving PDFs with LHC data

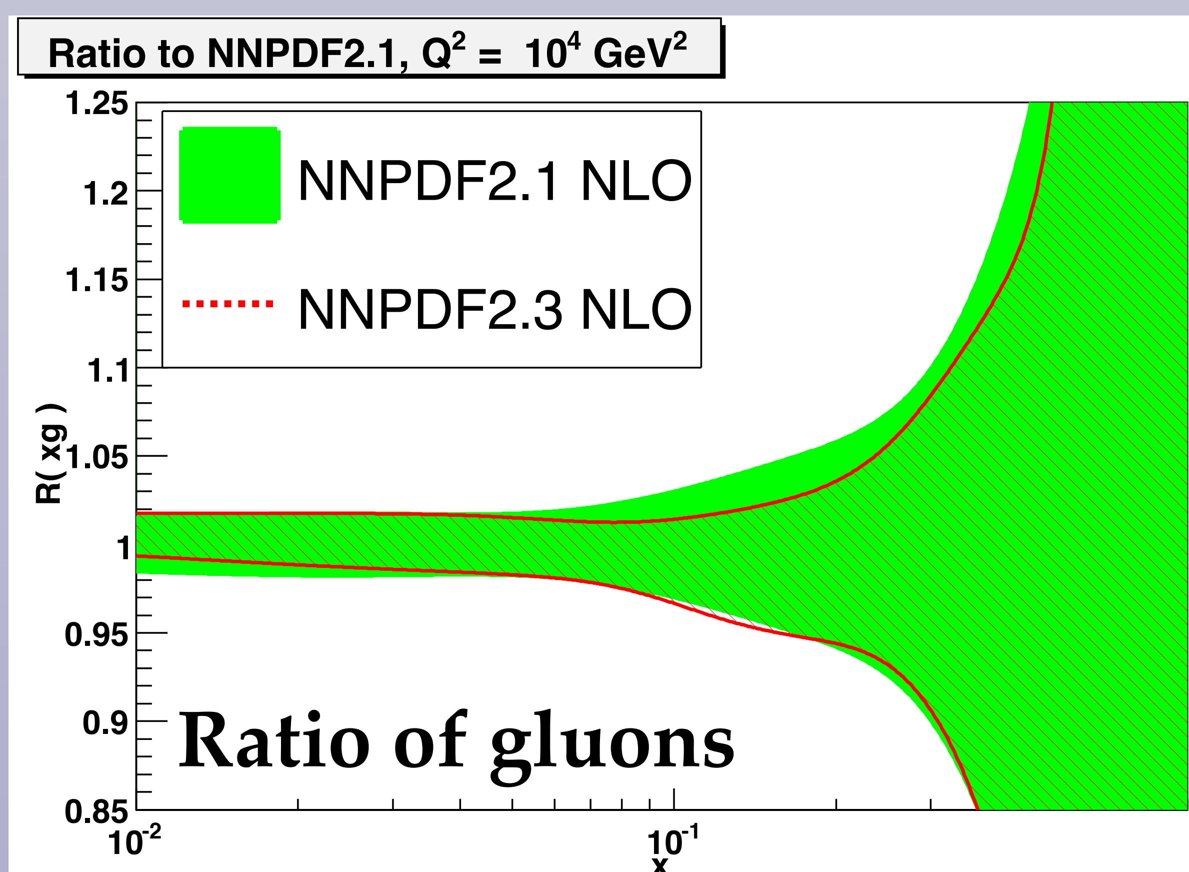


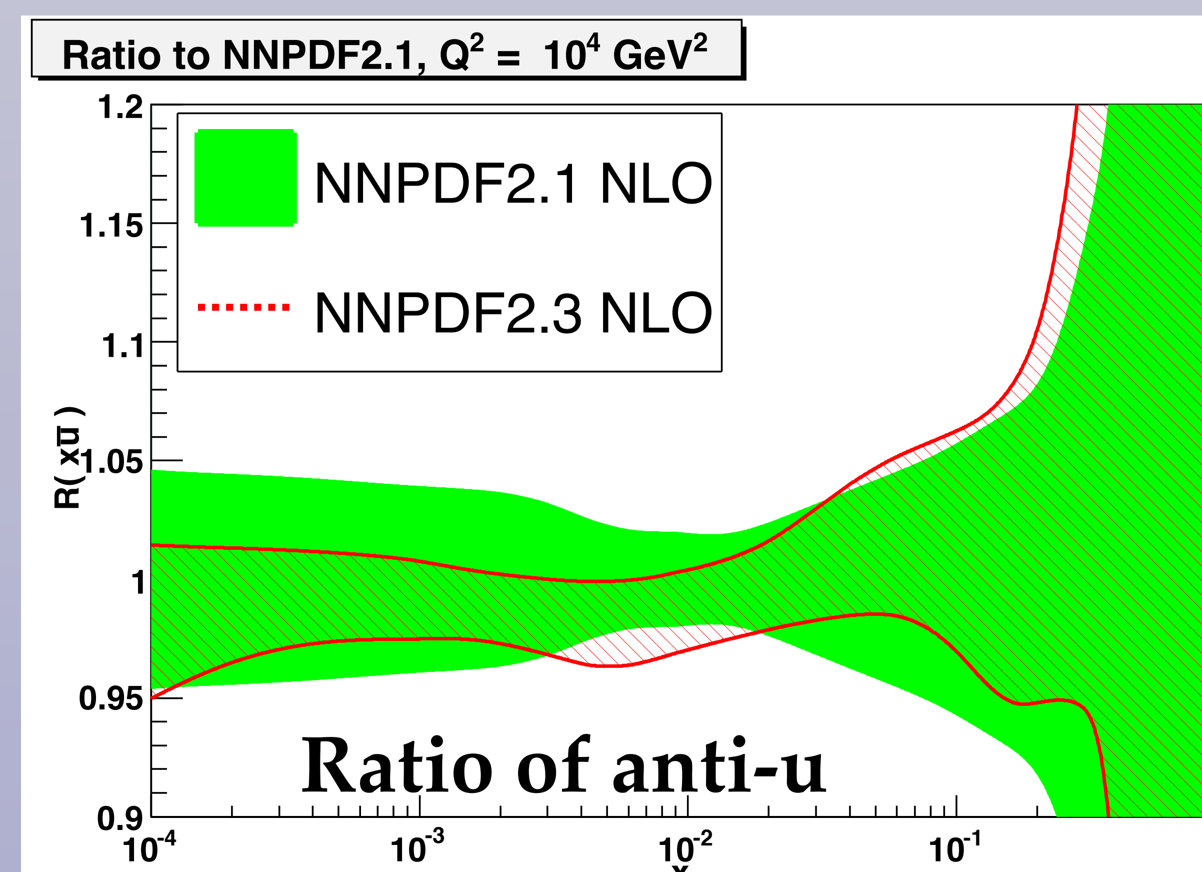
LHC data already part of global PDF analysis, *ie*. the recent **NNPDF2.3** sets

The **inclusive jet data** constrains large-x gluon

The **W and Z production** data from **CMS, ATLAS and LHCb** constrain medium-x antiquarks

PDFs with LHC data are more reliable for determination of MW because of **reduced uncertainties for W production kinematics**

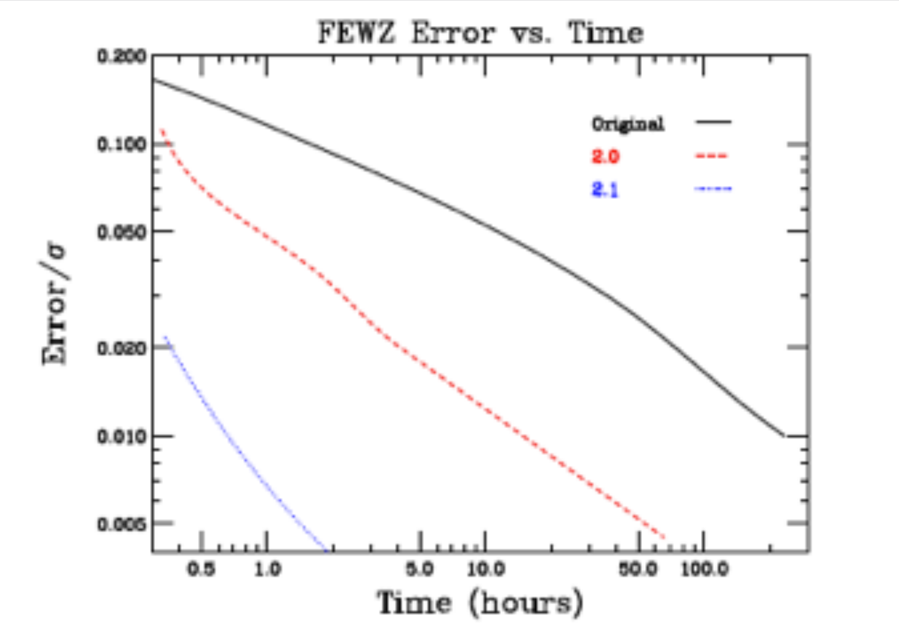




Improving PDFs with LHC data

- From the **experimental data point of view,** all the **current and future needs of the LHC in terms of PDFs,** including the **measurement of MW** could be addressed by a **specific PDF program at the LHC**
- There is a long list of measurements to be pursued, that will provide all required information on PDFs:
  - Inclusive jets and dijets, central and forward: **large-x quarks and gluons**
  - Isolated photons: **medium-x gluons**
  - Inclusive W and Z production and asymmetries: **quark flavor separation, strangeness**
  - W production with charm quarks: **direct handle on strangeness**
  - W production with jets: **medium small-x gluon**
  - Off resonance DY and W production at small and high mass: **quarks at very small and very large-x**
  - Top quark distributions: **large-x gluon**
  - Z+charm: **intrinsic charm PDF**
  - Single top production: **gluon and bottom PDFs**
  - Charmonium production: **small-x gluon**
- To maximize the **LHC data impact on PDFs**, it is crucial to **coordinate a detailed PDF program** between the LHC experiments and the Theory community
- Important input to improve **W mass measurements at the LHC**

MC Generators



Several **MC generators** available for precision W/Z physics

- **FEWZ**: Fully Exclusive W and Z production up to NNLO
  - FEWZ2.1 with improved integration (speed)
  - FEWZ3.0 also includes **NLO QED+EWK** corrections
  - Allows to compute **PDF uncertainties within a single run**
  - **Poor description or low pt region** (needed to model W recoil)
- **DYNNLO**
  - Same accuracy as FEWZ, but **different method to subtract NNLO singularities**, need different runs to compute PDF and scale errors
- **ResBos**
  - NLO accuracy supplemented with NNLL **p_T resummation**, but dependence of **modeling of NP corrections** (fitted to Tevatron data)
  - **Good description of the low pt region**
  - **Reduced flexibility** from the user point of view, for instance small subset of PDFs available

All parton level generators can be interfaced to **QED montecarlos for multiple photon emission**, like HORACE or PHOTOS

MC Generators

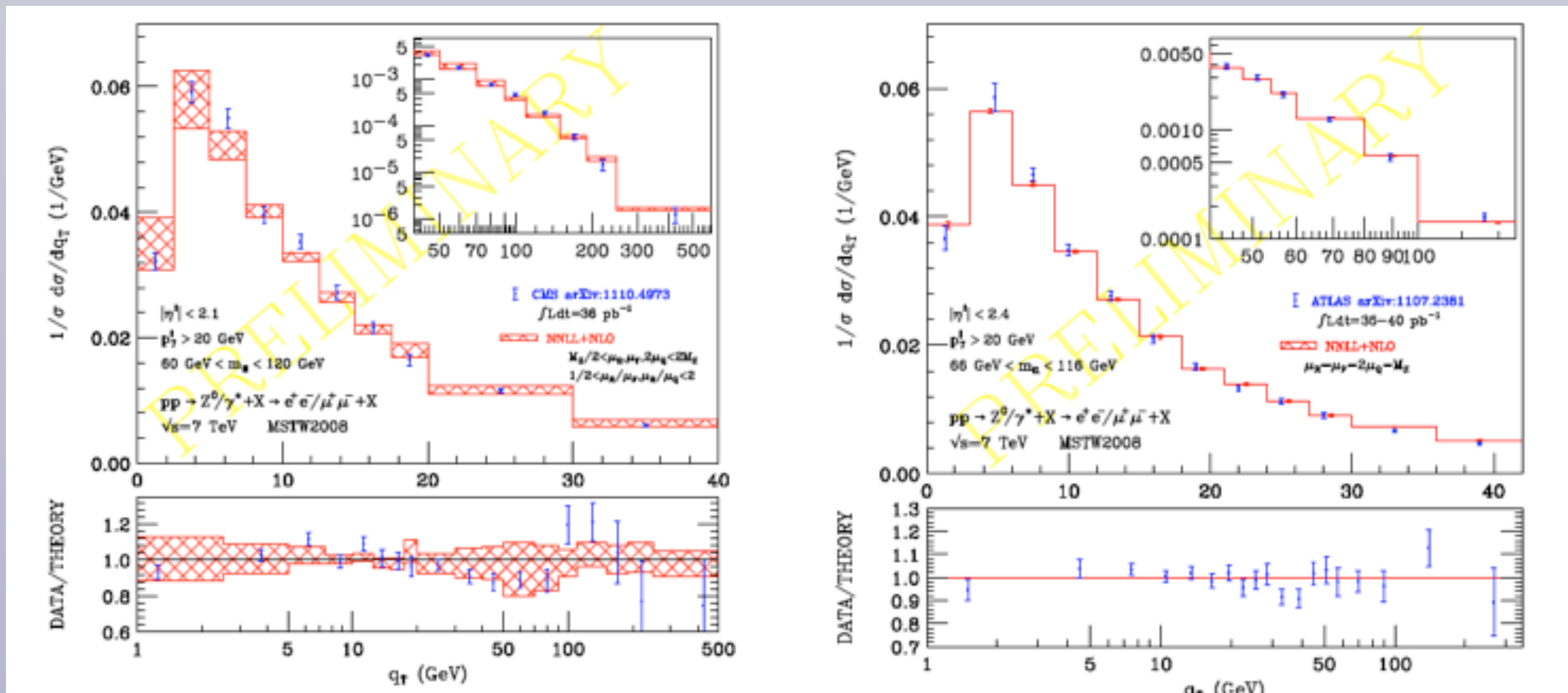
Several **MC generators** available for precision W/Z physics

- **aMCatNLO**: NLO+PS
  - NLO accuracy matched to various parton showers: **Pythia6, Herwig, Herwig++**
  - Allows to compute PDF uncertainties and scale errors within a single run
- **POWHEG:** NLO+PS
  - Same accuracy as **aMCatNLO,** exclusive description of event
  - Both aMCatNLO and POWHEG provde a correct description of the low p_T region

- For a realistic event generation at the LHC, we need NLO+PS exclusive events
- First step: **aMCatNLO/POWHEG** generation, also to estimate the full detector simulation corrections
- **Reweight** the low p_T region with the most precise **analytical resummed** calculations available
- Use at least three different PDF sets (**NNPDF2.3, CT10, MSTW08**) to estimate PDF errors
- **Huge CPU requirements**: templates for many different M_W values, PDFs, scales, full simulation: need to carefully design an optimal strategy

Predicting the W/Z bosons pt

- To model the **Higgs transverse momentum at the LHC**, one can **reweight the exclusive hadron NLO+PS distributions** from say POWHEG with the most **reliable perturbative calculations, NLO+NNLL**
- The same approach is possible **for W/Z production**. The **DYres code**, based on DYNNLO, allows to compute the W/Z pt with NLO+NNLL precision
- An **excellent description of the LHC data** obtained even at very small pt **without the need of any non-perturbative modeling**



CMS data for the Z q_T spectrum compared with NNLL+NLO result.
Scale variation:
$/2 \leq \{\mu_F/m_Z, \mu_R/m_Z, \mu_F/\mu_R, 2Q/m_Z, Q/\mu_R\} \leq 2$

ATLAS data for the Z q_T spectrum compared with NNLL+NLO result.

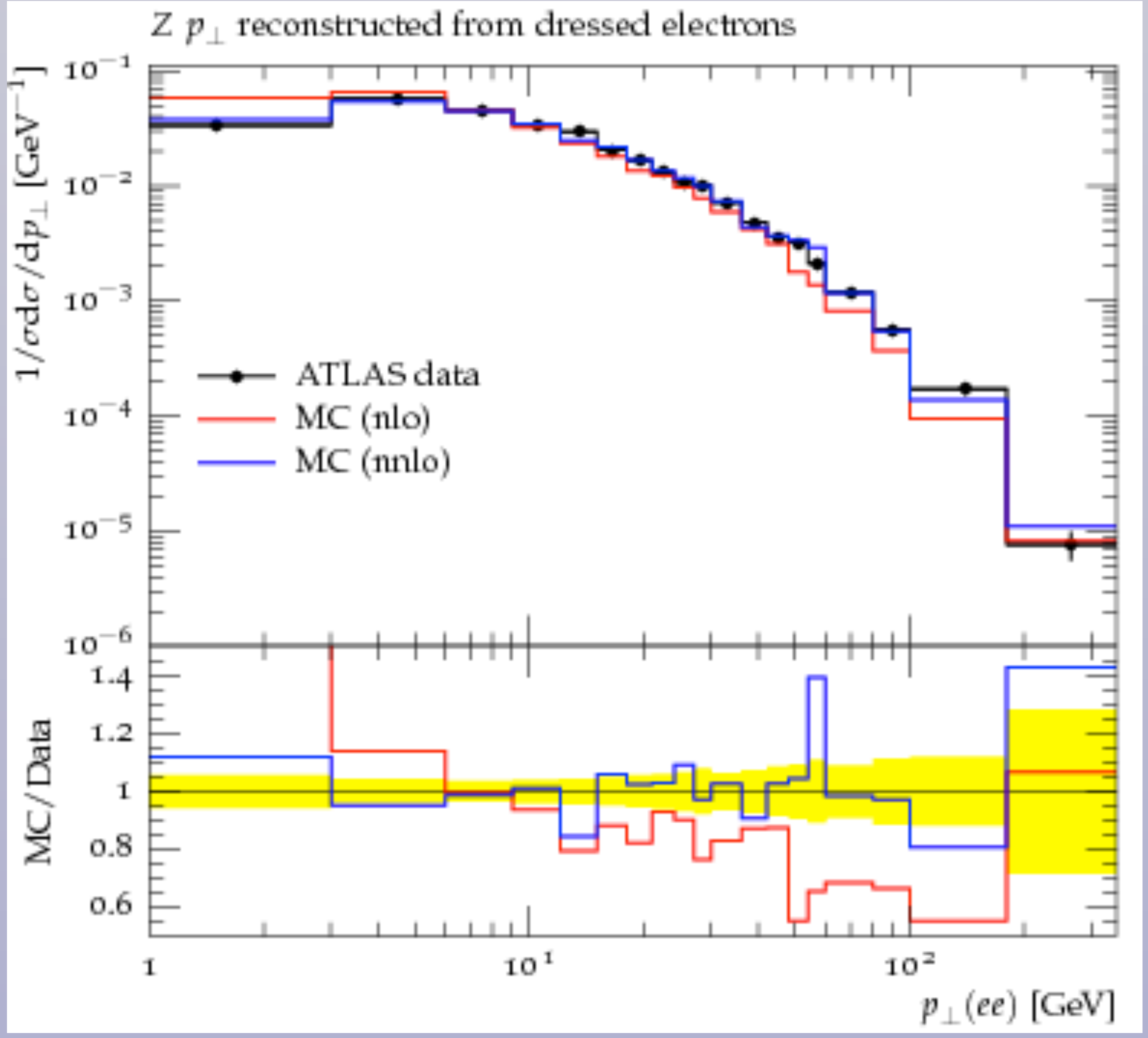
G. Ferrera, Moriond 2012

Predicting the W/Z bosons pt

- To model the **Higgs transverse momentum at the LHC**, one can **reweight the exclusive hadron NLO+PS distributions** from say POWHEG with the most **reliable perturbative calculations, NLO+NNLL**

- The same approach is possible **for W/Z production**. The **DYres code**, based on DYNNLO, allows to compute the W/Z pt with NLO+NNLL precision

- An **excellent description of the LHC data** obtained even at very small pt **without the need of any non-perturbative modeling**



- Preliminary studies (P. Lenzi) show that a **NLO event sample** can be **reweighted** with **DYres** to achieve excellent data-theory description

- Important ingredient for the **W mass determination at the LHC**

NNPDFs with QED corrections

- **Photon-initiated diagrams** are required for electroweak precision calculations, including W production
- The **DGLAP QCD evolution equations** can be modified to account for **QED corrections,** introducing a **photon PDF**
- Then the global PDF analysis is repeated, and we study
    - The **modifications on the proton and neutron PDFs due to QED corrections**
    - The size of the **photon PDF** allowed by experimental data
    - Violation of **isospin symmetry** (connection with nuclear models)

$$Q^2 \frac{\partial}{\partial Q^2}\gamma(x,Q^2) = \frac{\alpha(Q^2)}{2\pi}\int_x^1 \frac{d\xi}{\xi}\left\{P_{\gamma\gamma}(\xi)e_\Sigma^2\gamma\left(\frac{x}{\xi},Q^2\right) + P_{\gamma q}(\xi)\sum_j e_j^2 q_j\left(\frac{x}{\xi},Q^2\right)\right\}$$

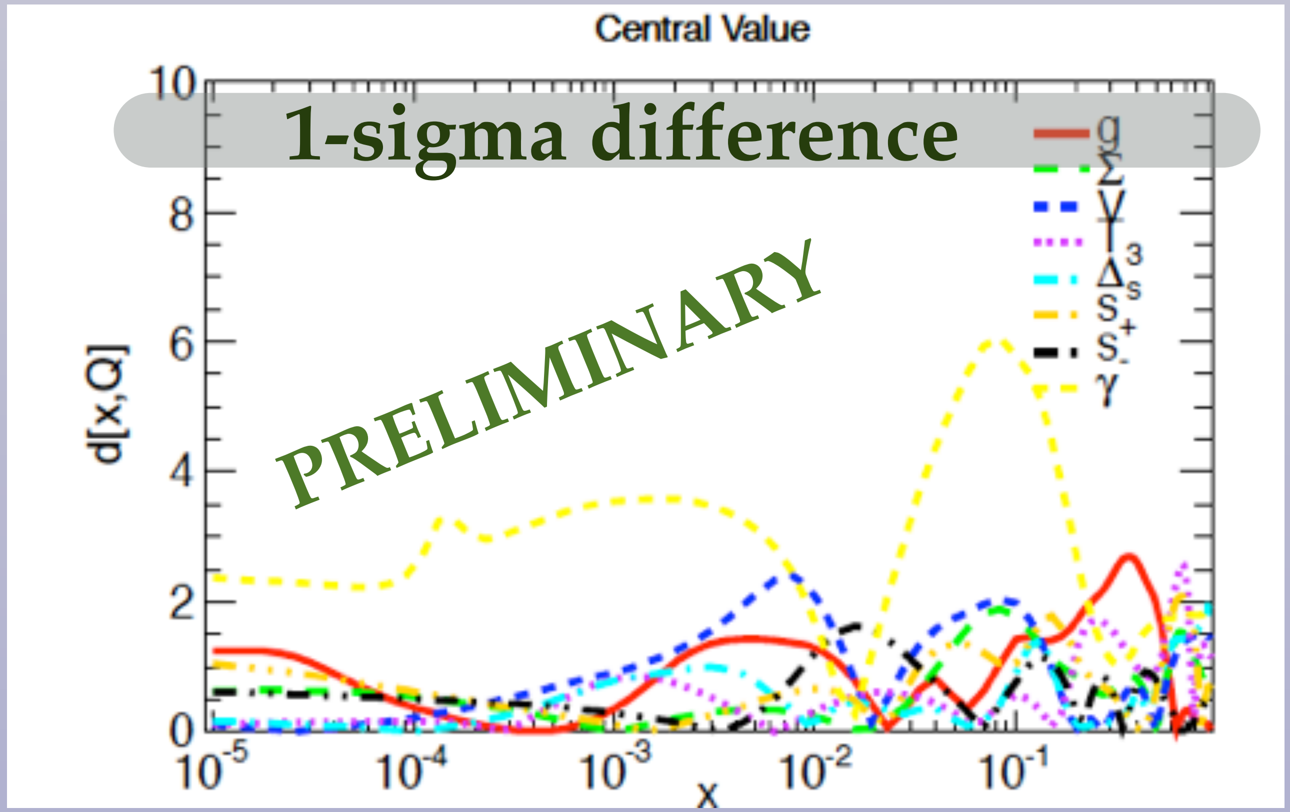
$$Q^2 \frac{\partial}{\partial Q^2}q_i(x,Q^2) = \frac{\alpha(Q^2)}{2\pi}\int_x^1 \frac{d\xi}{\xi}\left\{P_{q\gamma}(\xi)e_i^2\gamma\left(\frac{x}{\xi},Q^2\right) + P_{qq}(\xi)e_i^2 q_i\left(\frac{x}{\xi},Q^2\right)\right\}$$

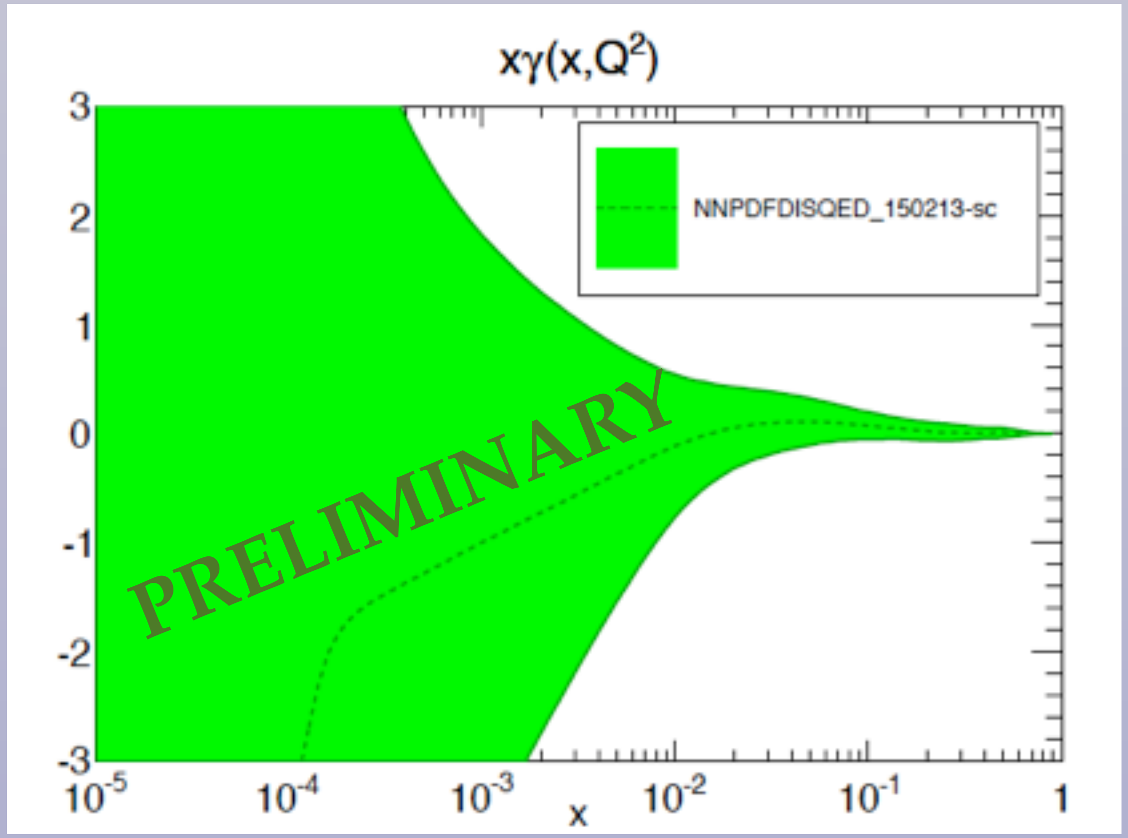
- As opposed to MRST2004QED, where a model was used for the photon PDF, in NNPDF we parametrize the photon with **a neural network with 37 parameters**, to be **extracted from data**

NNPDFs with QED corrections

- Very preliminary results based on **NNPDF2.3**, suggest that
  - 1) **Impact on QED** on **quarks and gluons small**, well below PDF uncertainties
  - 2) Experimental data leave still some room for a substantial **photon PDF:** large uncertainties
- Work in progress to study the **phenomenological implications of the photon NNPDF**, in particular for Drell-Yan production at the LHC

NNPDF2.3 QED





Summary

- **The W mass** is a very important measurement at hadron colliders: provide **stringent test of SM consistency** and **sets bounds on many BSM scenarios**
- **Theoretical uncertainties** now dominant for W mass at the Tevatron, will be crucial to carefully assess them at the LHC, but **no huge difference expected**
- PDF uncertainties reduce substantially if **normalized templates** are used in the W mass fit, **without affecting the sensitivity to M_W**
- PDF errors to **M_W** @ LHC not larger than **20 MeV** level, to be improved soon with LHC data
- Crucial ingredient is theory modeling of the **templates.** **Theoretically most reliable option** is to NLO+PS exclusive events and reweight the W/Z pt spectrum with the analytical NLO+NNLL calculation
- Both the **W mass measurement**, as many **other important LHC analysis**, will benefit from **improved PDFs** thanks to LHC data
- **NNPDF2.3 with QED corrections** will soon allow to consistently include **QED effects in NC and CC Drell-Yan** with an up-to-date PDF set