PDF uncertainties at the LHC made easy:

Compression algorithms for the combination of PDF sets

Juan Rojo
STFC Rutherford Fellow
Rudolf Peierls Center for Theoretical Physics
University of Oxford

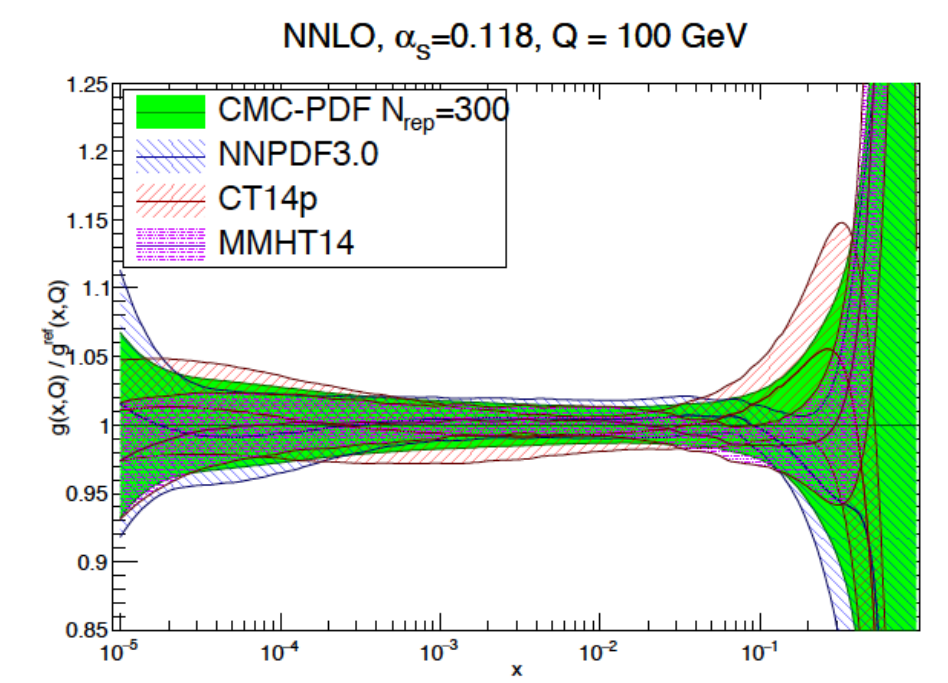
PDF4LHC Meeting
CERN, 13/04/2015

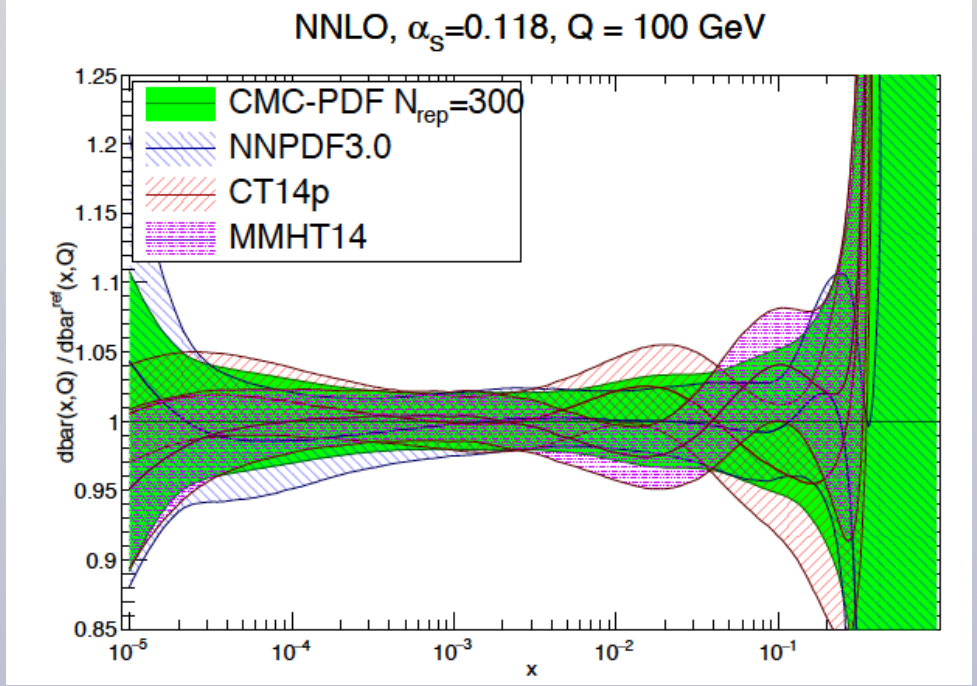
CMC-PDFs Update

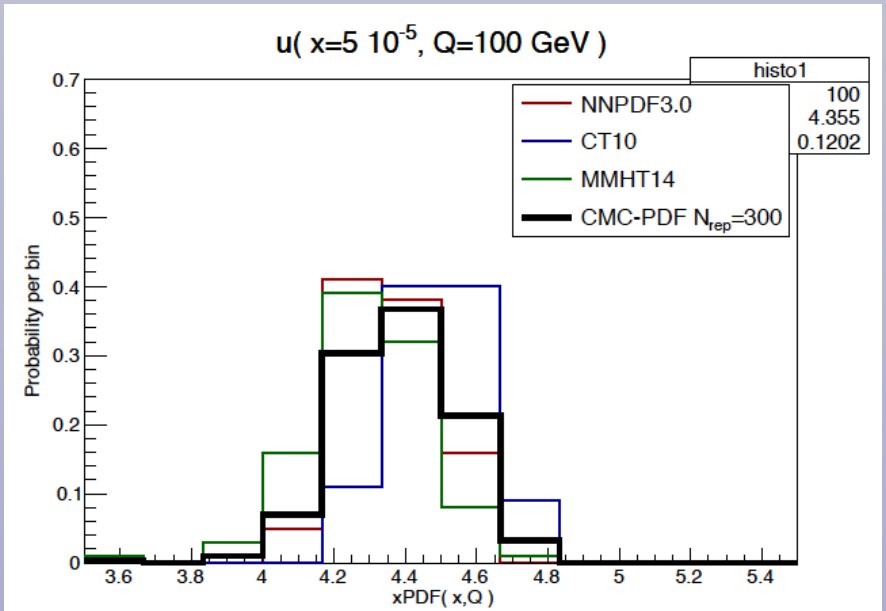
in collaboration with **S. Carrazza**, **J. I. Latorre** and **G. Watt**

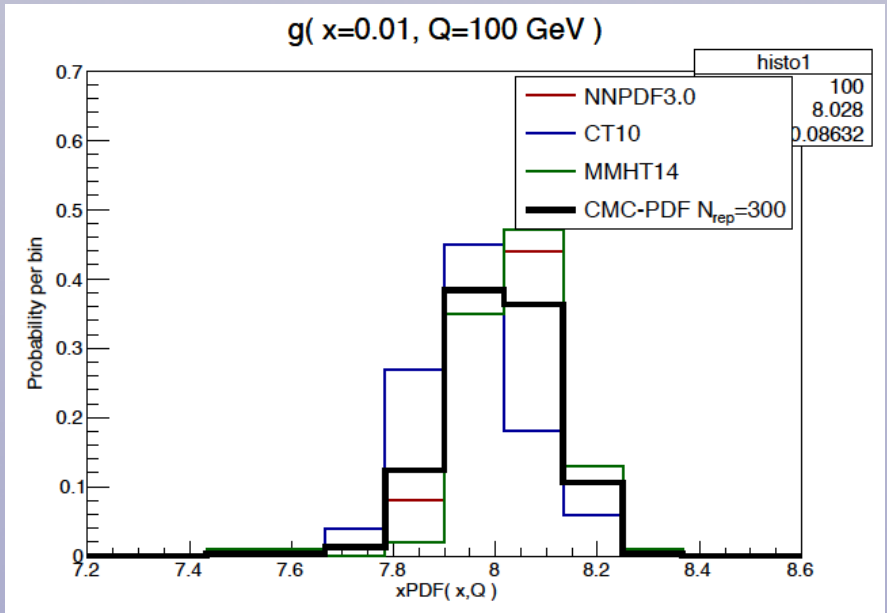
The combined Monte Carlo PDF set

- Updated combination, based on **NNPDF3.0, MMHT14 and CT14p, all at NNLO**



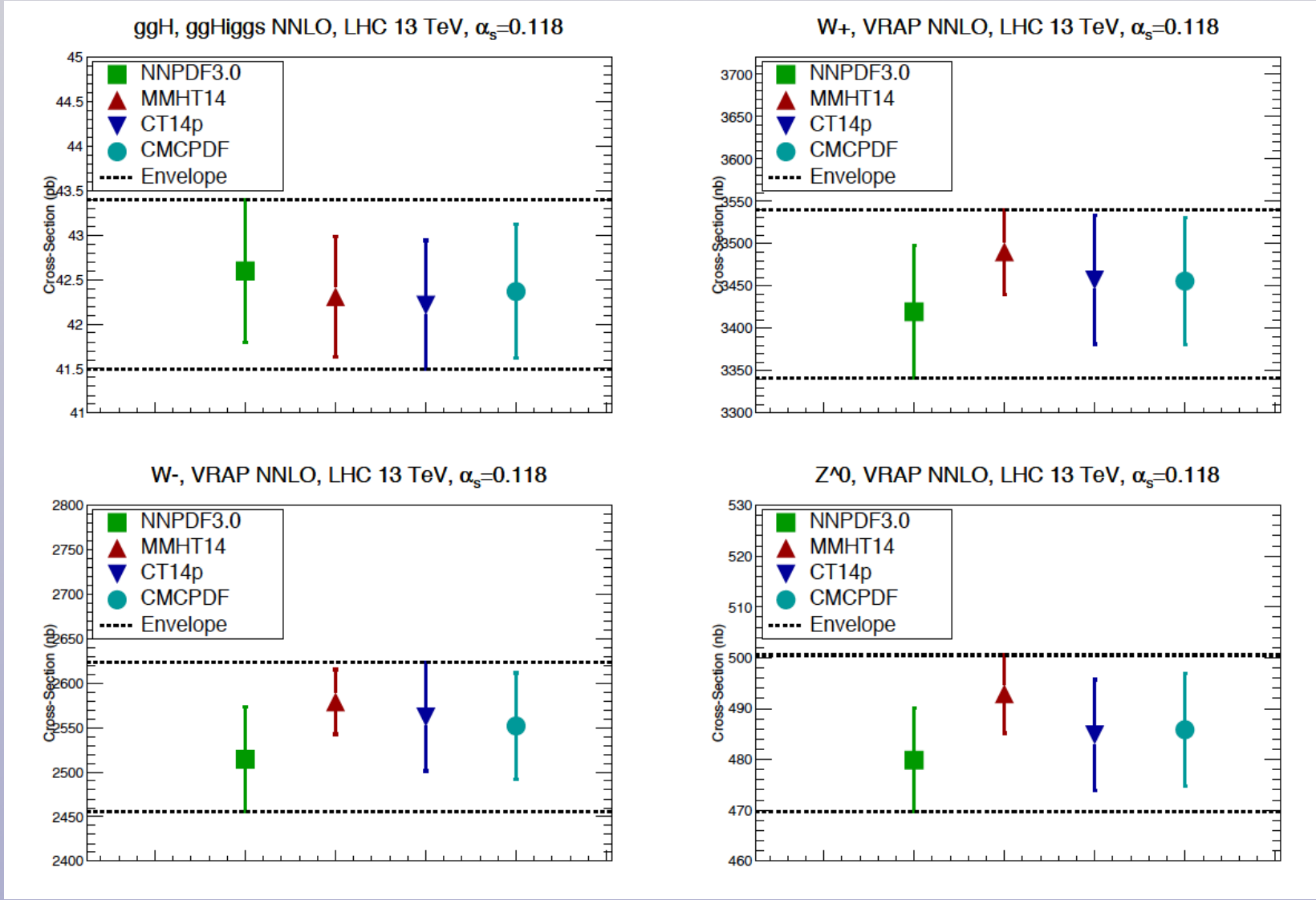






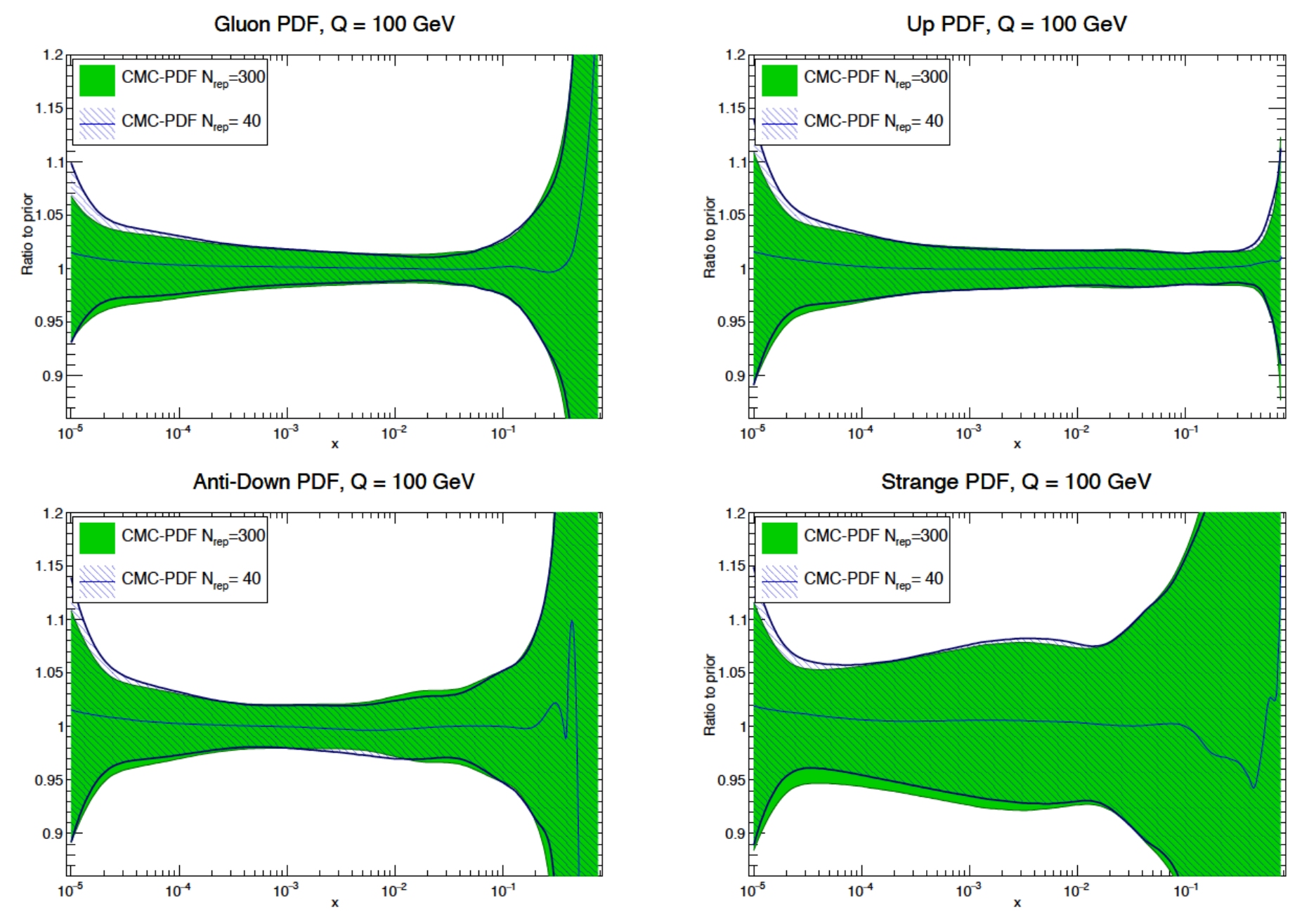
Combined MC set vs PDF4LHC envelope

- As already noted in the **Forte-Watt** study, the **MC combination** leads to somewhat smaller uncertainties than the **PDF4LHC envelope** (same here and in the Meta-PDFs)
- This can be understood because now **each PDF set receives the same weight**, while the PDF4LHC envelope effectively gives more weight to the **outliers**



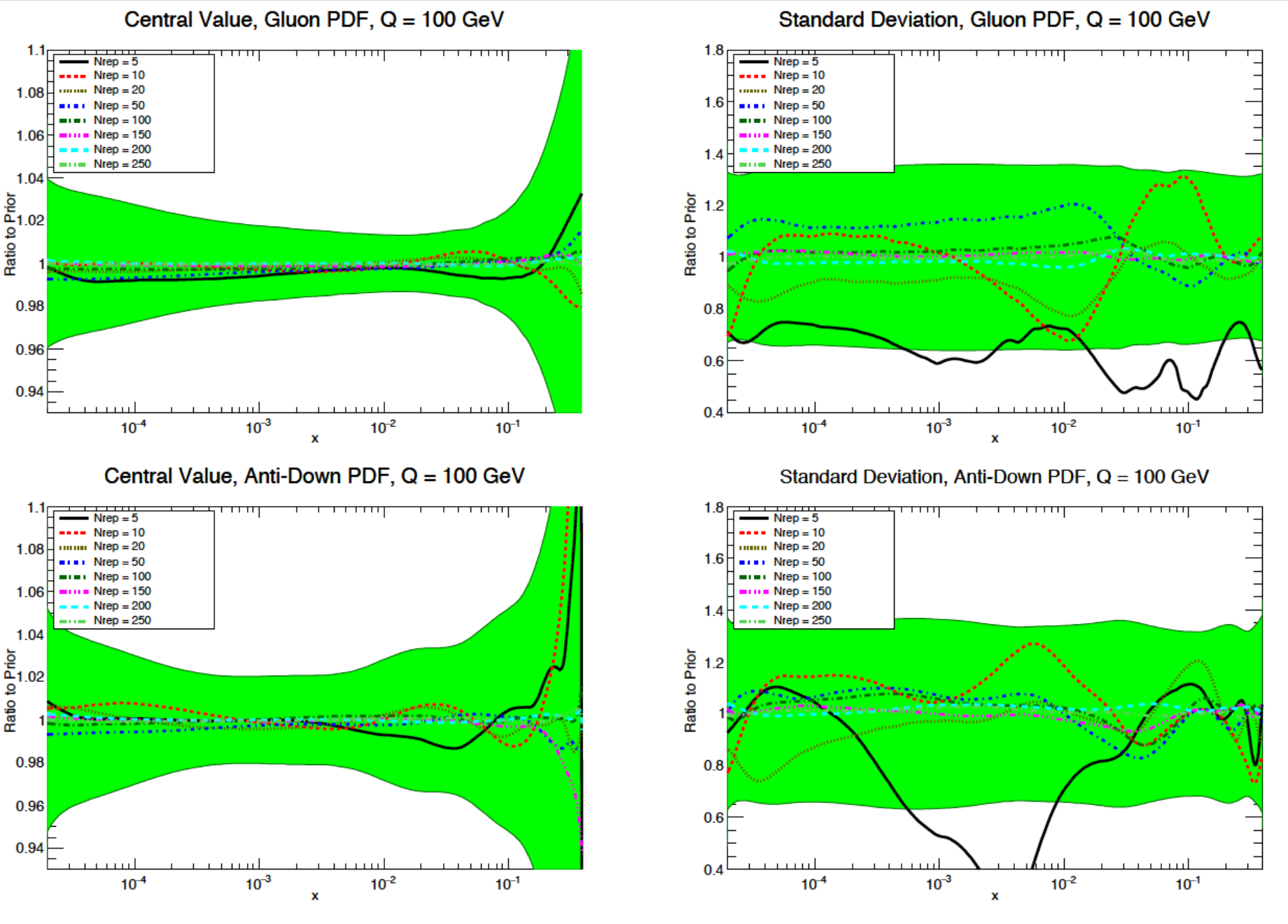
Results of the compression

- For example, for $\mathbf{N_{rep}=40}$ replicas the compressed and the original PDFs are virtually identical



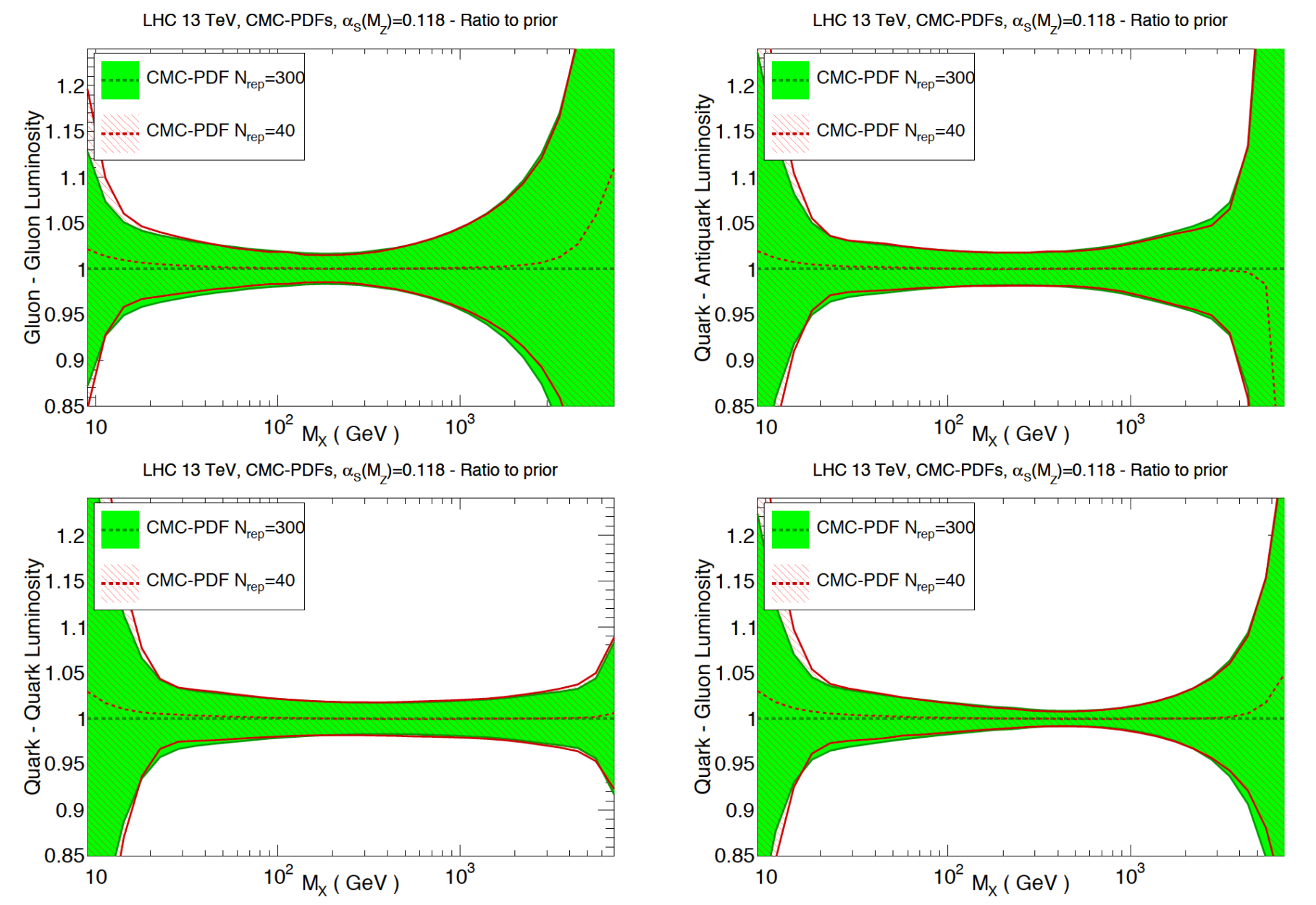
Results of the compression

- For example, for $\mathbf{N_{rep}=40}$ replicas the compressed and the original PDFs are virtually identical



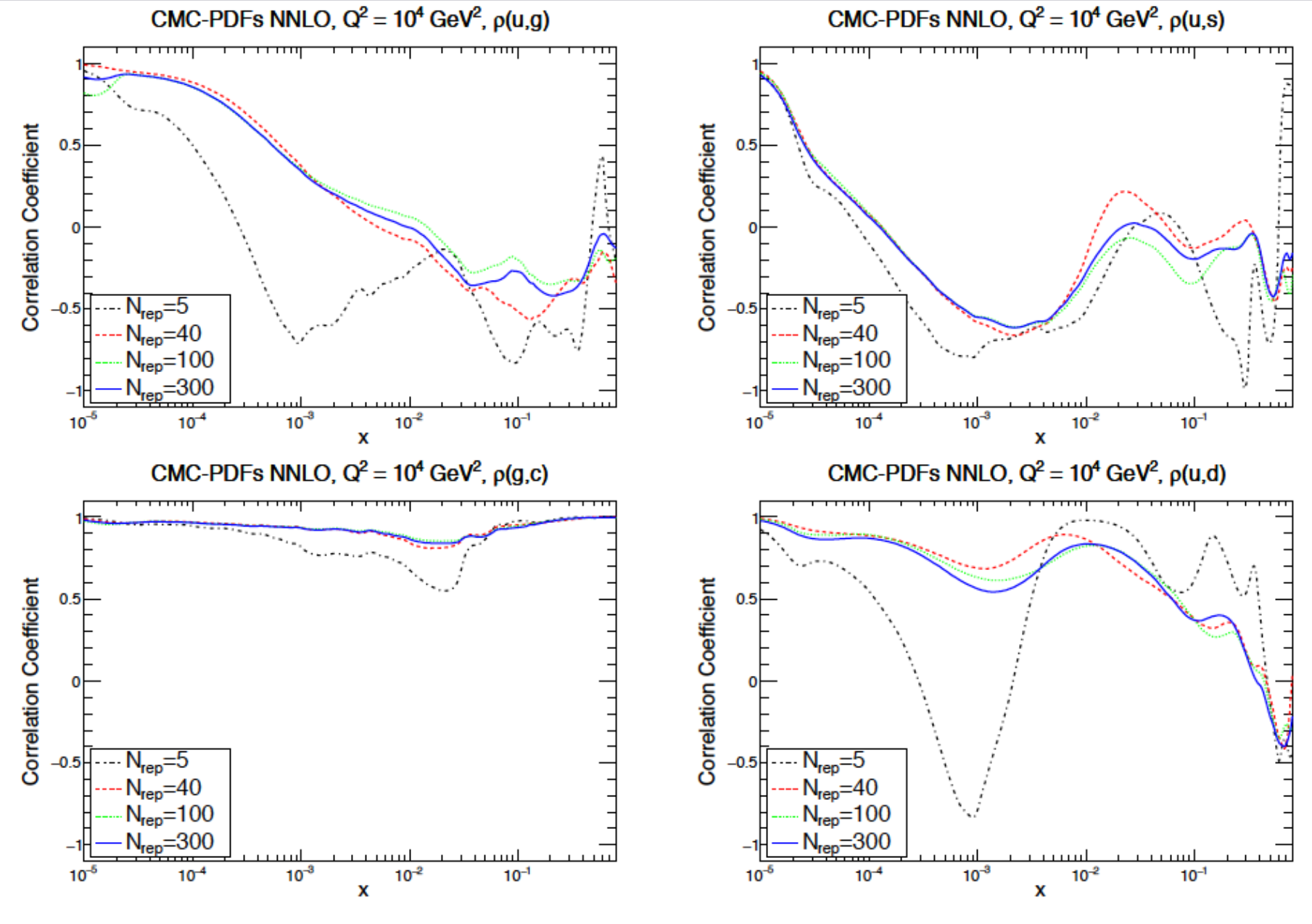
Results of the compression

- For example, for $\mathbf{N_{rep}=40}$ replicas the compressed and the original PDFs are virtually identical



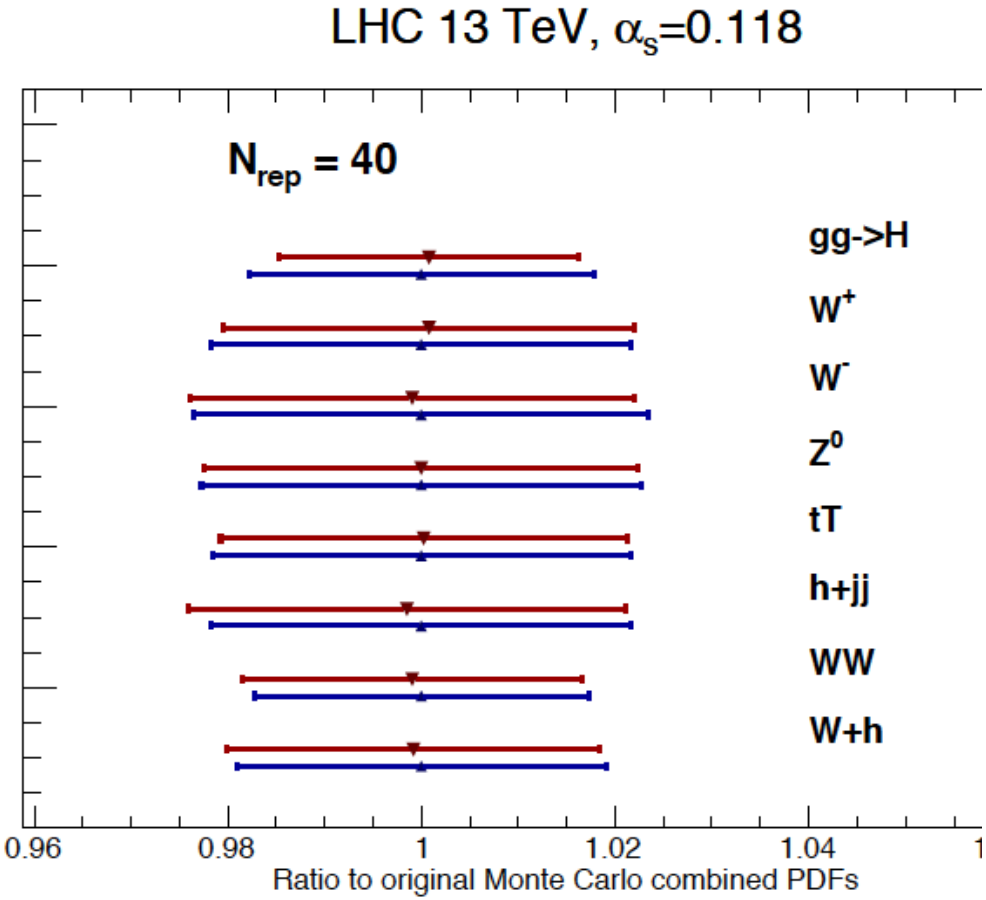
Results of the compression

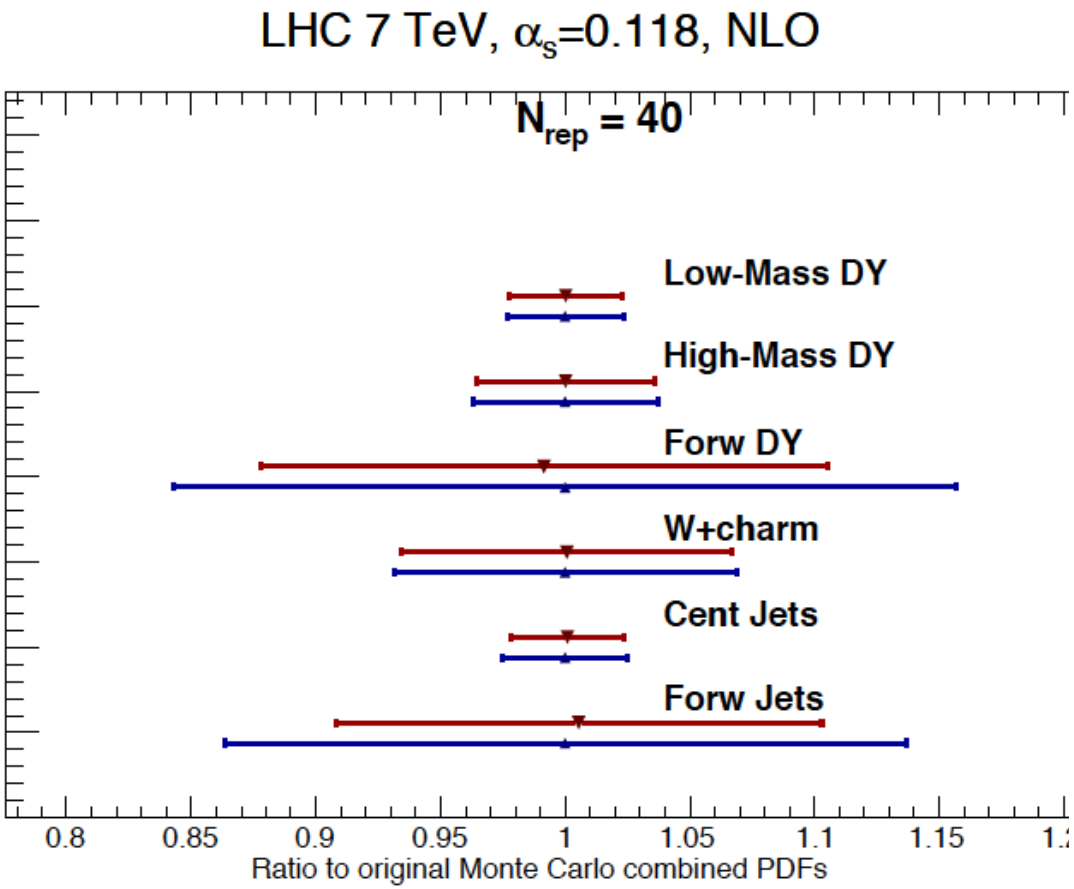
- Correlations between PDFs are also nicely reproduced



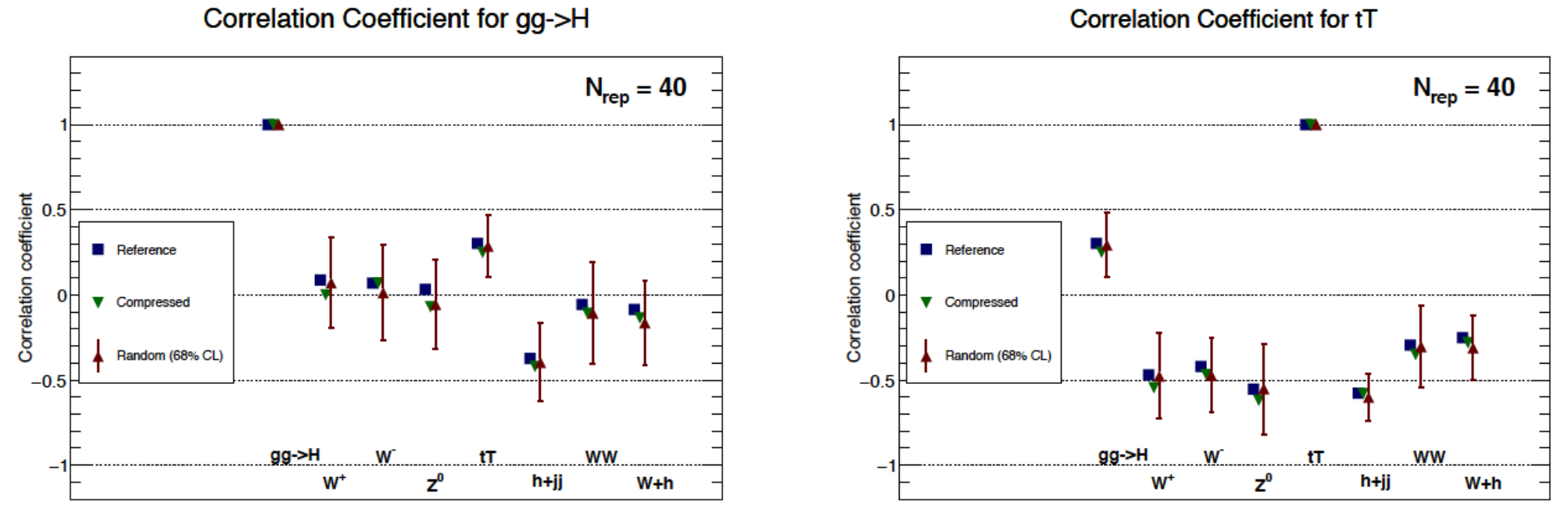
Phenomenology

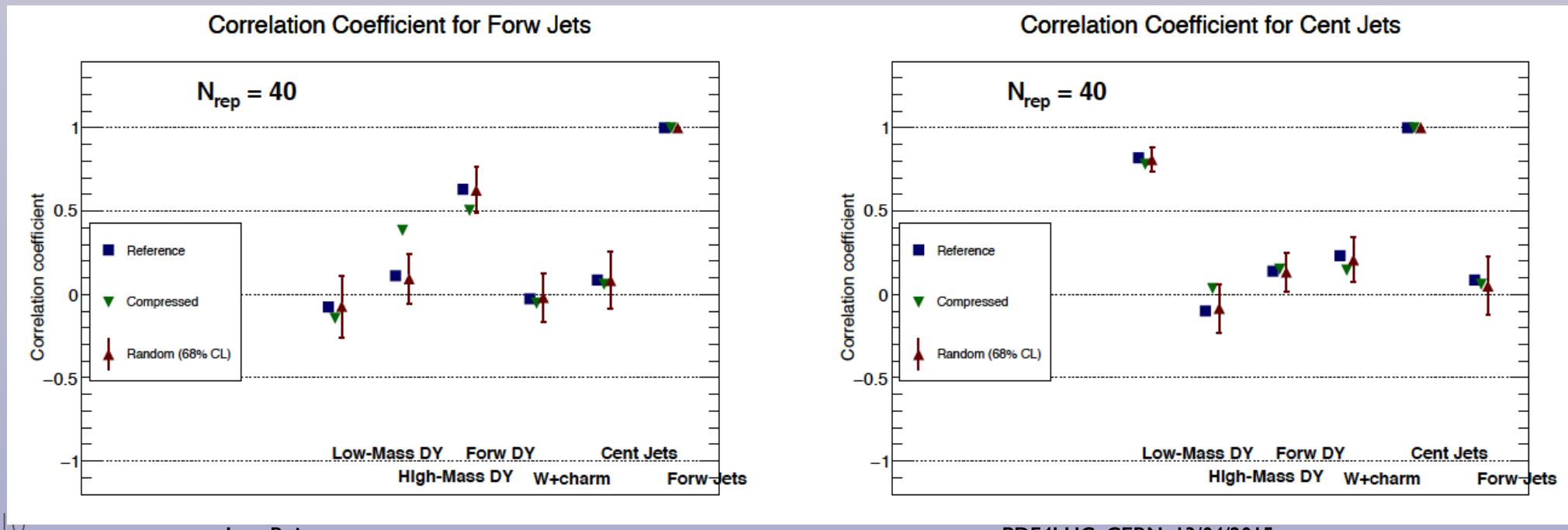
- As expected from the good agreement at the PDF level, CMC-PDFs are also validated at the level of **LHC inclusive cross-sections and differential distributions**





Correlations of LHC xsecs





CMC-PDFs vs META-PDFs Comparisons

S. Carrazza, J. I. Latorre, J.R. and **G. Watt** (CMC-PDFs)
J. Gao, J. Huston, P. Nadolsky (META-PDFs)

Thanks to Pavel and Jun for producing the following comparison plots!

Benchmark comparisons of two combination methods. Work plan (from Benasque workshop)

Input MC ensemble: **NNPDF3.0+CT14+MMHT14 NNLO**, with **alphas(M_Z)=0.118**

Convert to 300 replicas in LHAPDF6 format at $Q_0 = 8$ GeV (above the bottom mass), using two independent codes (JR and JG). Cross-check that results are identical.

Done. The results from two groups agree. Mild differences are due to random variations in the generation of MC replicas.

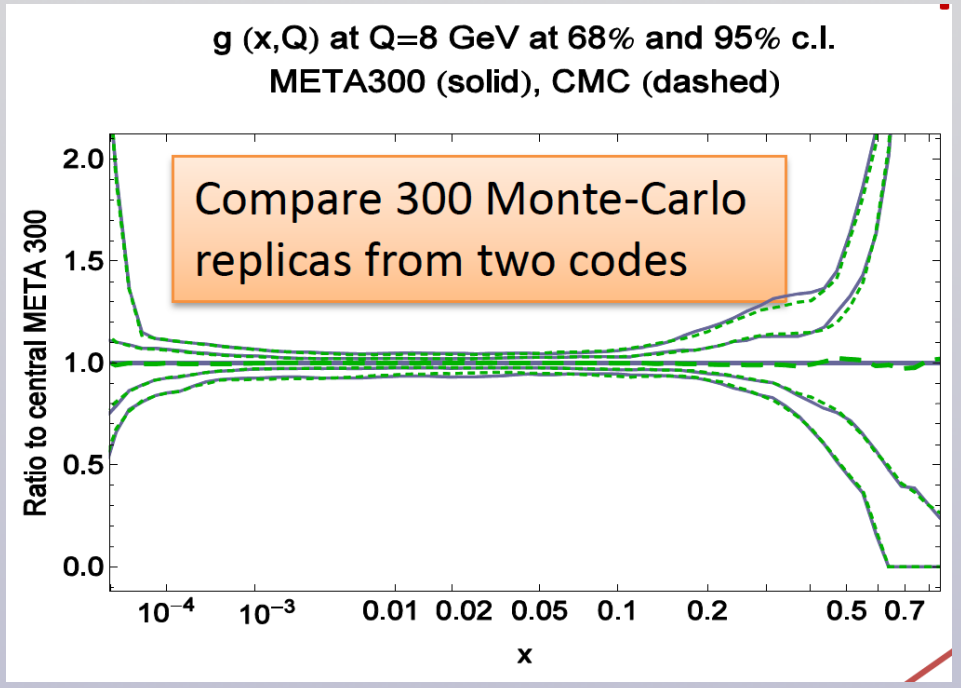
In each approach, reduce the number of replicas to the minimal number that retains 1% or 5% accuracy in reproducing the following properties of the input ensemble:

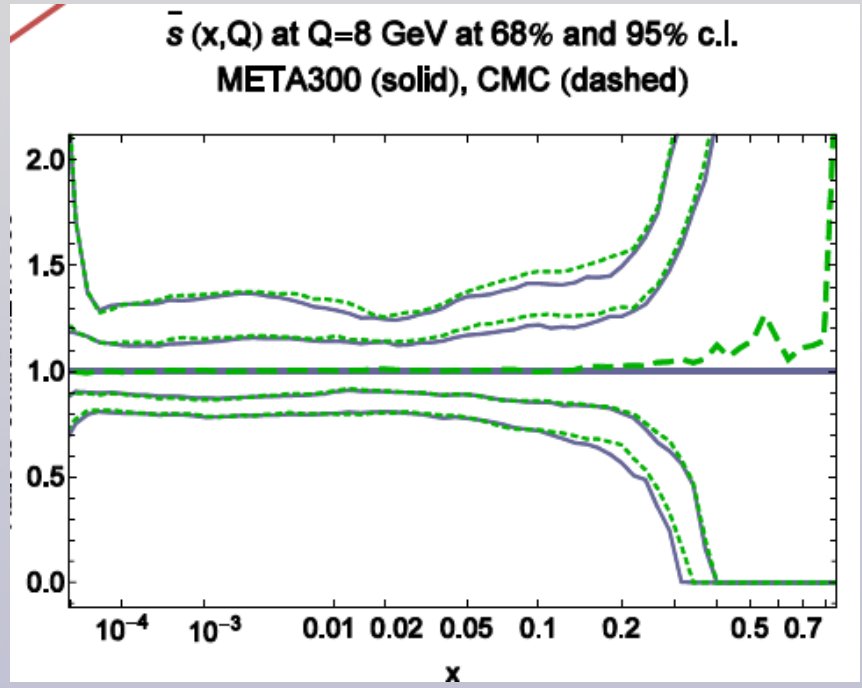
- Means, 68%c.l. PDF uncertainties, higher moments and asymmetry (skewness), PDF-PDF correlations.

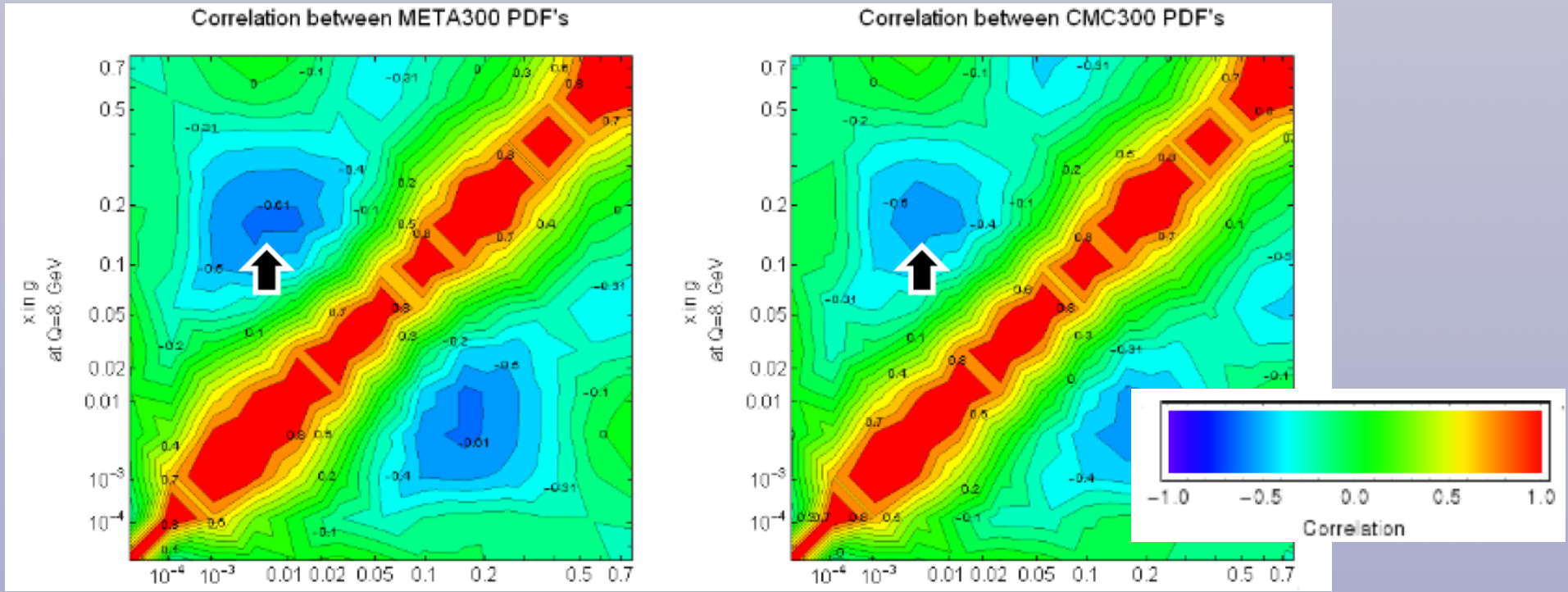
  In progress. Ensembles with 60-100 META PDFs and 40 CMC replicas broadly agree.

- Predictions for the standard candle LHC observables used in the META paper: ggHiggs, ttbar, W,Z [Jun]
- Differential LHC distributions using NNPDF3.0 applgrids, supplemented with some new aMCfast grids [Juan]

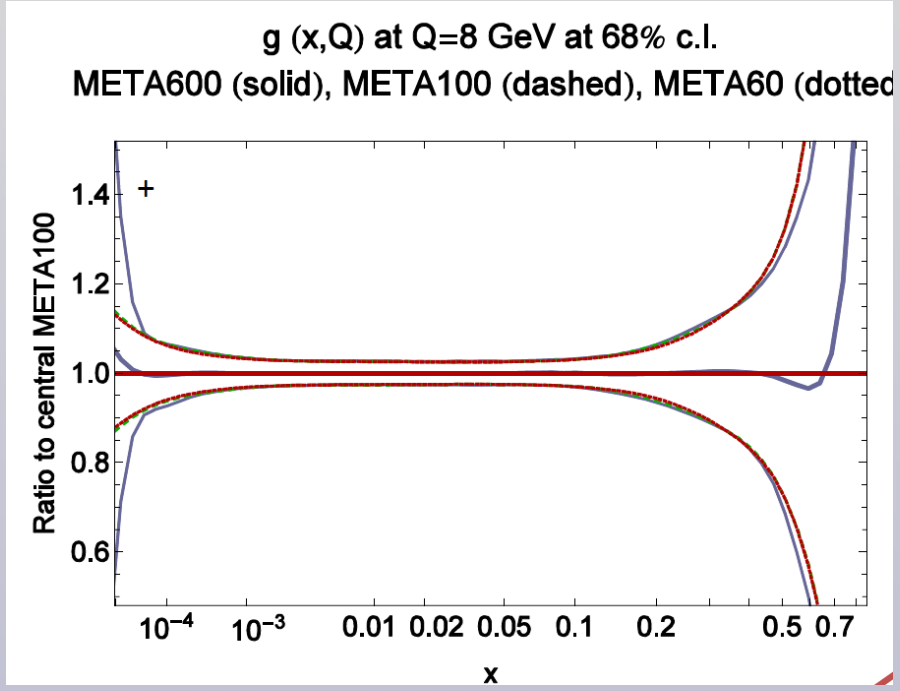
Comparison of prior combined MC sets

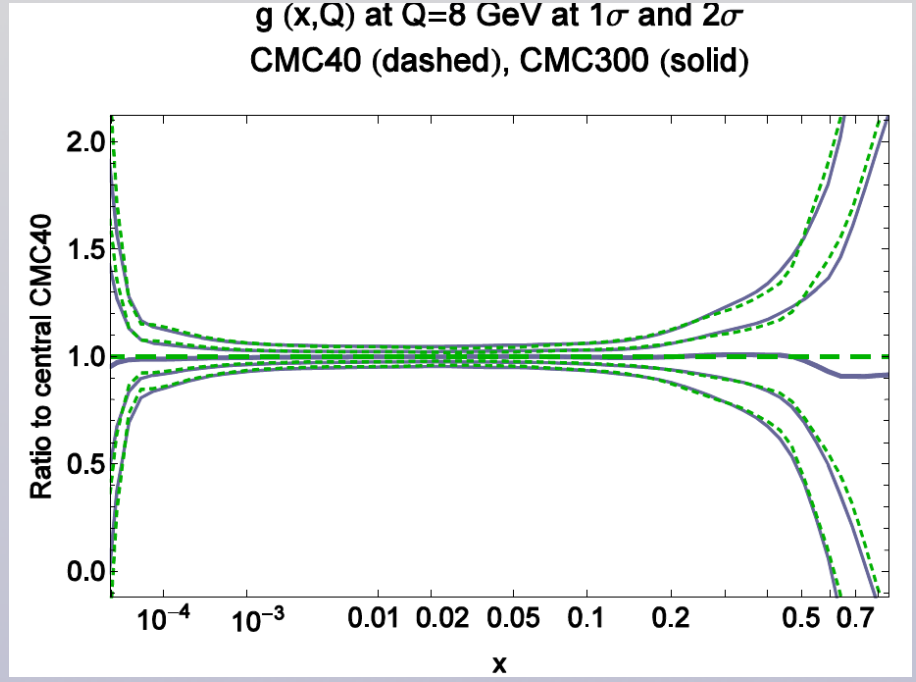


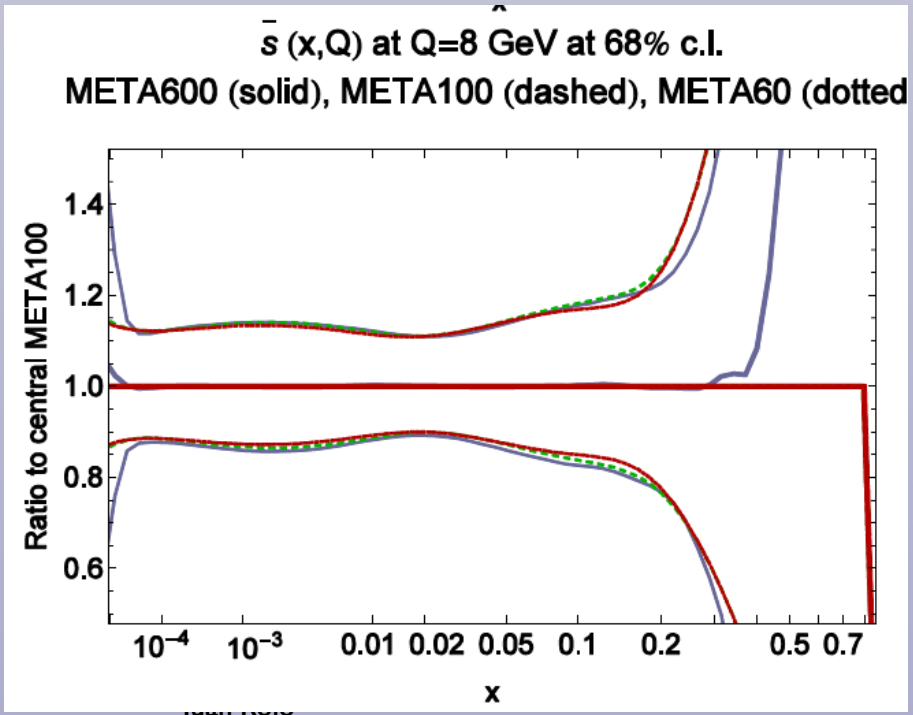


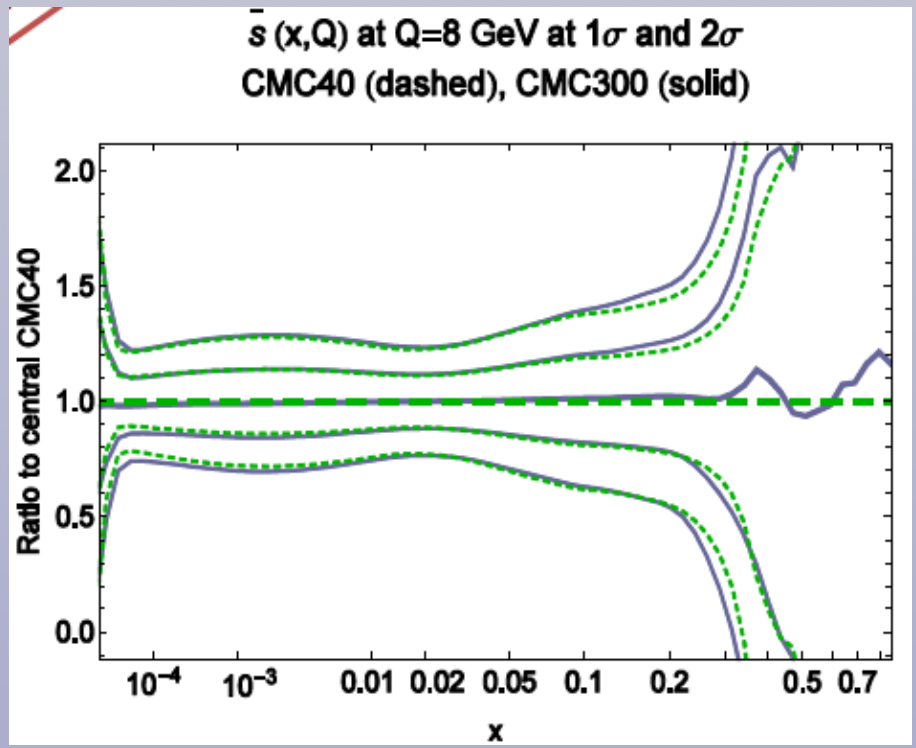


Reduction/Compression - PDFs

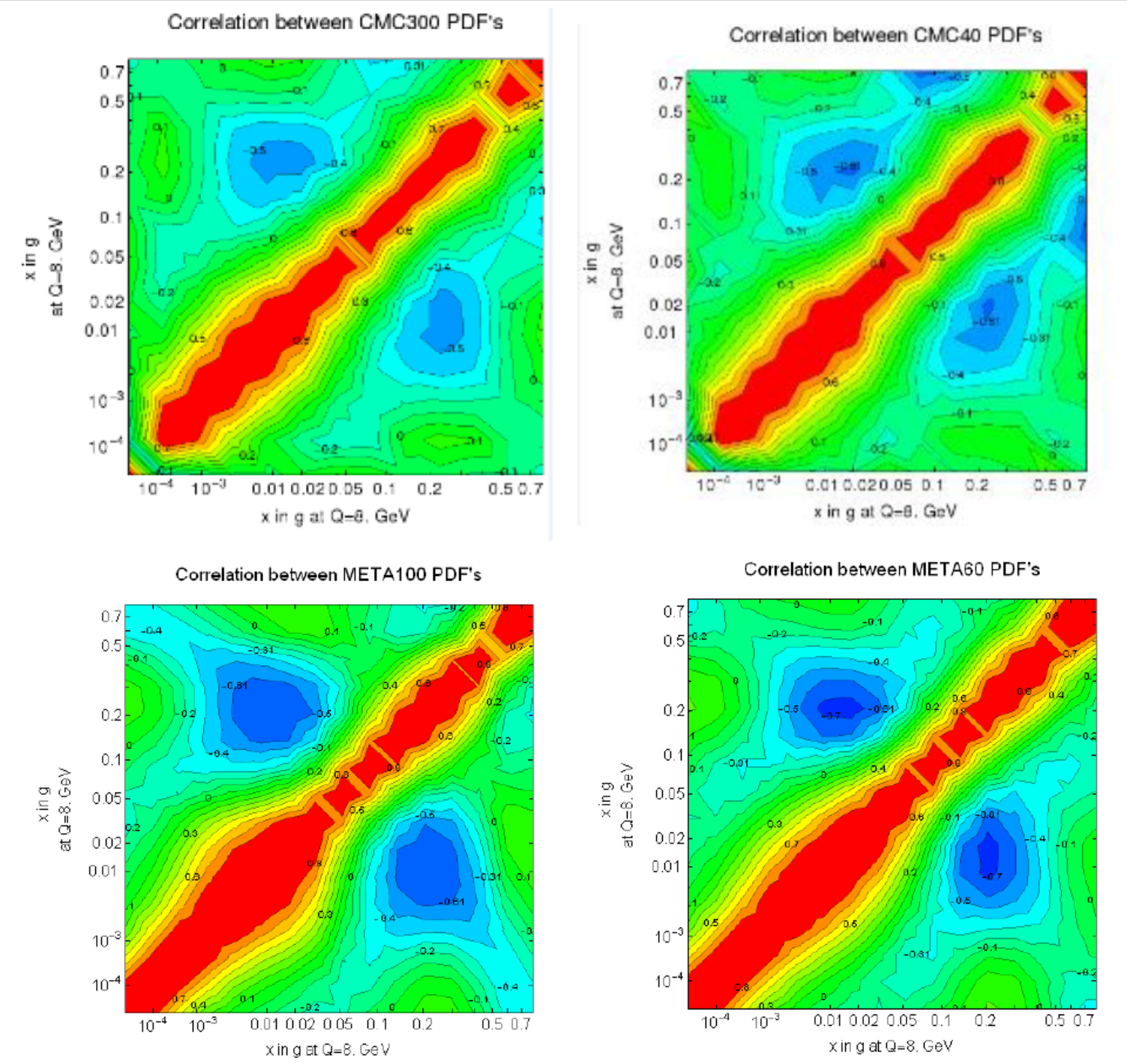








Reduction/Compression - Correlations



- **Correlations are also reasonable maintained** by the reduction algorithms
- In the regions with good data coverage, differences between original and reduced sets are of the order of statistical noise
- Slightly larger differences in PDFs with large uncertainties, like strangeness

Summary and next steps

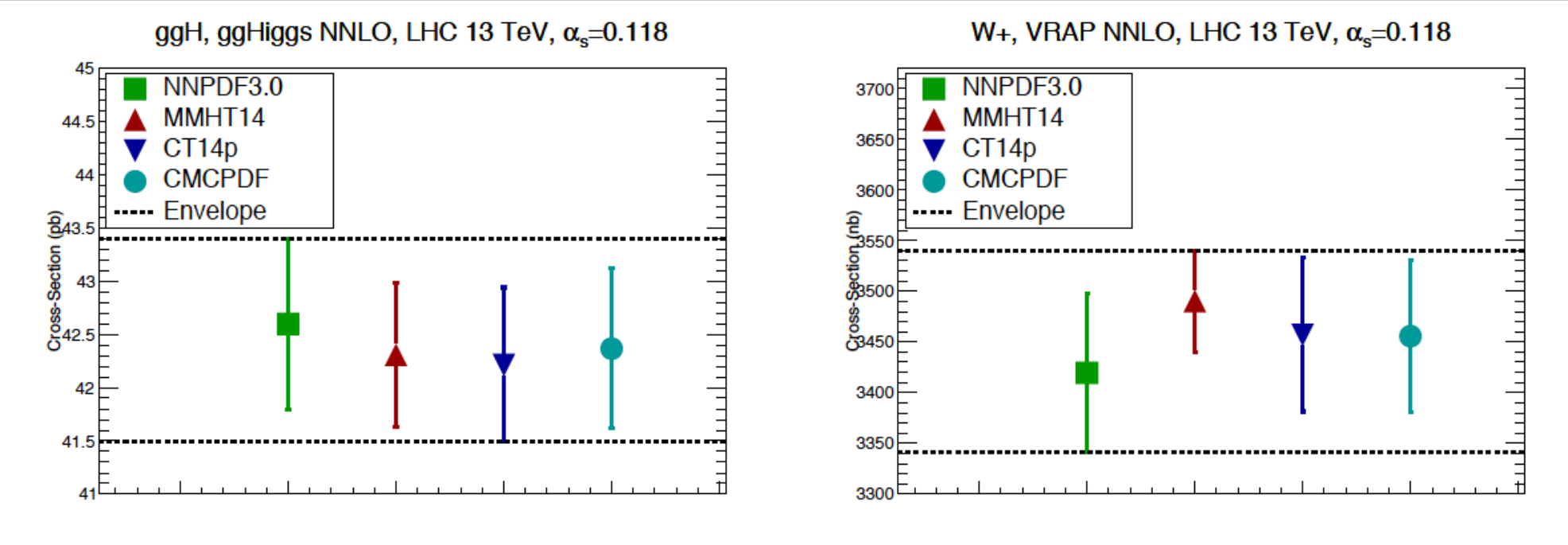
- Two independent codes for the **MC combination of NNPDF3.0, CT14 and MMHT14** have been cross-checked in details: found identical results, up to small statistical differences due to **finite sample size effects**
- For both reduced sets, **CMC40 and META60**, in general good agreement for central values, one-sigma uncertainties and correlations, both among them and with the original prior
- In extrapolation regions, **META-PDF uncertainties narrower than CMC-PDFs** due to the more rigid common interpolation.
- By construction, **non-Gaussian features in the combination** should be reproduced more faithfully by the **CMC-PDFs**. In the data region PDF uncertainties are reasonably Gaussian
- In summary, **META60/CMC40** provide r**easonably accurate implementations of the MC combination of PDF sets, suitable for general-purpose LHC applications**
- CMC-PDFs can also be provided in a **Hessian representation** (S. Carrazza's talk)
- Both CMC-PDFs(Hessian) and META-PDFs can, if needed, be further reduced for specific applications using the **Dataset Diagonalization Method**
- Final step is the **comparison of the two methods at the level of LHC observables**, but the agreement at the PDF level suggests that both methods will work similarly well
- Both approaches provide **ready-to-use LHAPDF6 grid files**

Discussion

some personal thoughts and suggestions ….

Combined MC set vs PDF4LHC envelope

- The **MC combination** leads to somewhat smaller uncertainties than the **PDF4LHC envelope**
- This can be understood because now **each PDF set receives the same weight**, while the PDF4LHC envelope effectively gives more weight to the **outliers**
- **Q: How robust is this combined PDF uncertainty estimate?** What about theory uncertainties, like higher orders? Do we want include some **rescaling factor to be more conservative**?



CMC-PDF / META-PDF usage

- Whilst the two approaches **broadly agree well**, for some PDF combinations in specific kinematic regions (for instance large-x) they **show non-negligible differences**
- e.g. CMC-PDFs uses same parametrisation and theory settings as original sets, while META-PDFs adopts a unified meta-parametrization and common theory settings
- Q: Given the different methodological systematics of the two methods, instead of using **one or the other**, it would be more robust to use **both CMC40 and META60**, so that one cross-checks the other
- The total one-sigma PDF uncertainty band could be then defined as the **envelope of CMC40 and META60?**

CMC-PDF / META-PDF non-usage

- It should be emphasised again that for any analysis where **PDF systematics are really important**, one should never use combined sets, but use individual sets: **only way to understand what drives PDF uncertainties**, and to **learn how to reduce them**
- Crucial to spell out carefully under which conditions CMC-PDF / META-PDF can or cannot be used
- Also, the approach is IMHO not mature enough to **use central CMC-PDF / META-PDF for Monte Carlo event generation**