A meta analysis of parton distribution functions for LHC applications

Jun Gao
Southern Methodist University
Jun 5, 2013

Preliminary results in collaboration with Pavel Nadolsky

PDF at the LHC era

PDFs describe structures of hadronic matter, exist in several types: **unpolarized proton PDFs**, nuclear PDFs, polarized PDFs, TMD PDFs…

PDFs are essential for BSM searches and other complex LHC measurements. High precision for the PDFs is required to match theoretical efforts in pQCD, e.g., recent top quark pair, jet production, Higgs+jet calculations at the NNLO.

NNLO PDFs are available from 6 groups:

NNLO	CTEQ	MRST/MSTW	NNPDF	ABKM/ABM	HERAPDF	JR
Available since	2012	2001	2011	2009	2011	2008
Latest set	CT10	MSTW2008	NNPDF2.3	ABM11	HERA1.5	JR09

Other groups also provide NLO PDFs, CTEQ-JLab, KT2011, …

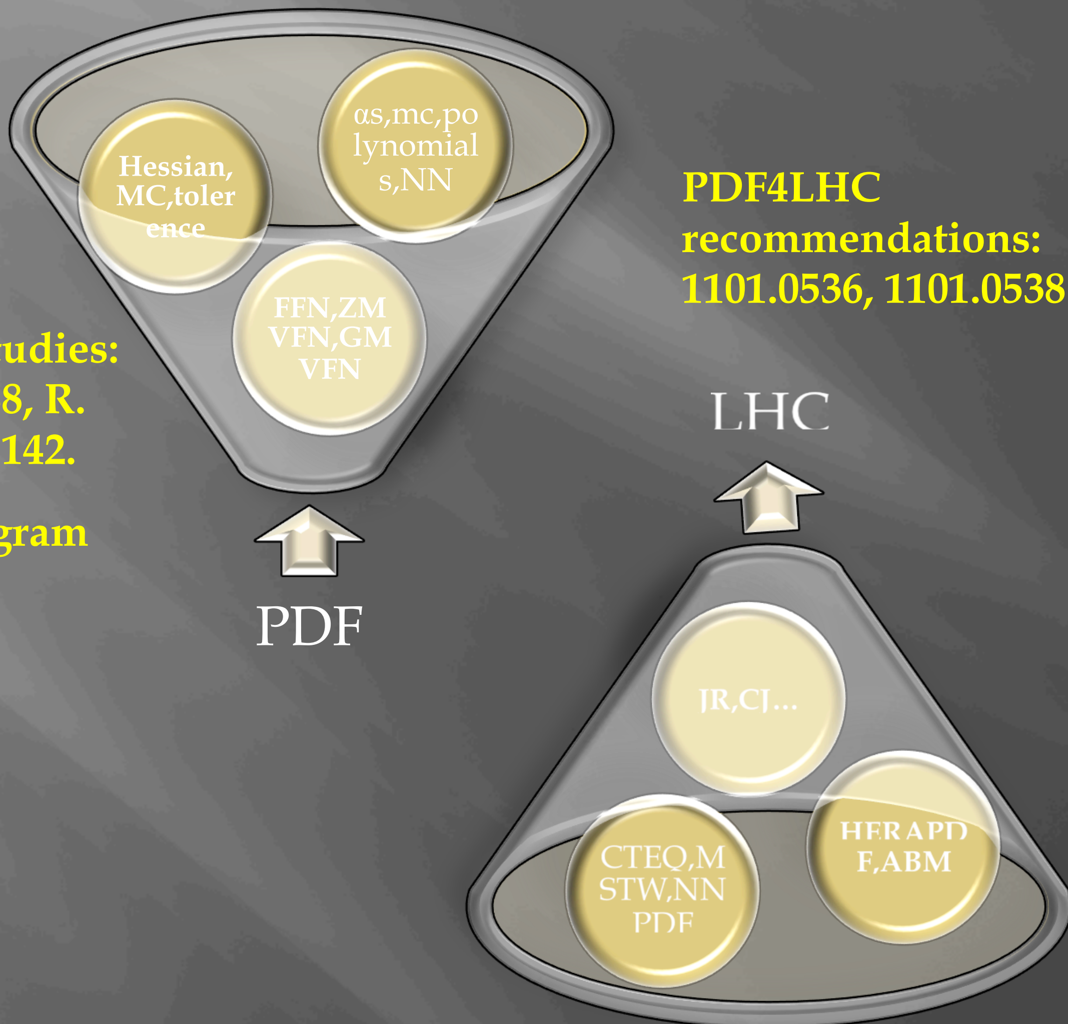
Most constraints on PDFs are derived from non-LHC measurements: HERA, fixed target experiments, Tevatron

Comparing and combining PDFs from different groups

Often, it is necessary to provide a central pQCD prediction and PDF uncertainty using PDF sets from several groups. However, these PDF sets assume distinctly different physics inputs and can not be combined in a simplistic way.

Benchmarking studies: G. Watt, 1106.5788, R. Ball, et al., 1211.5142.

HERA Fitter program



PDF4LHC recommendations: 1101.0536, 1101.0538

Can we try a different path?

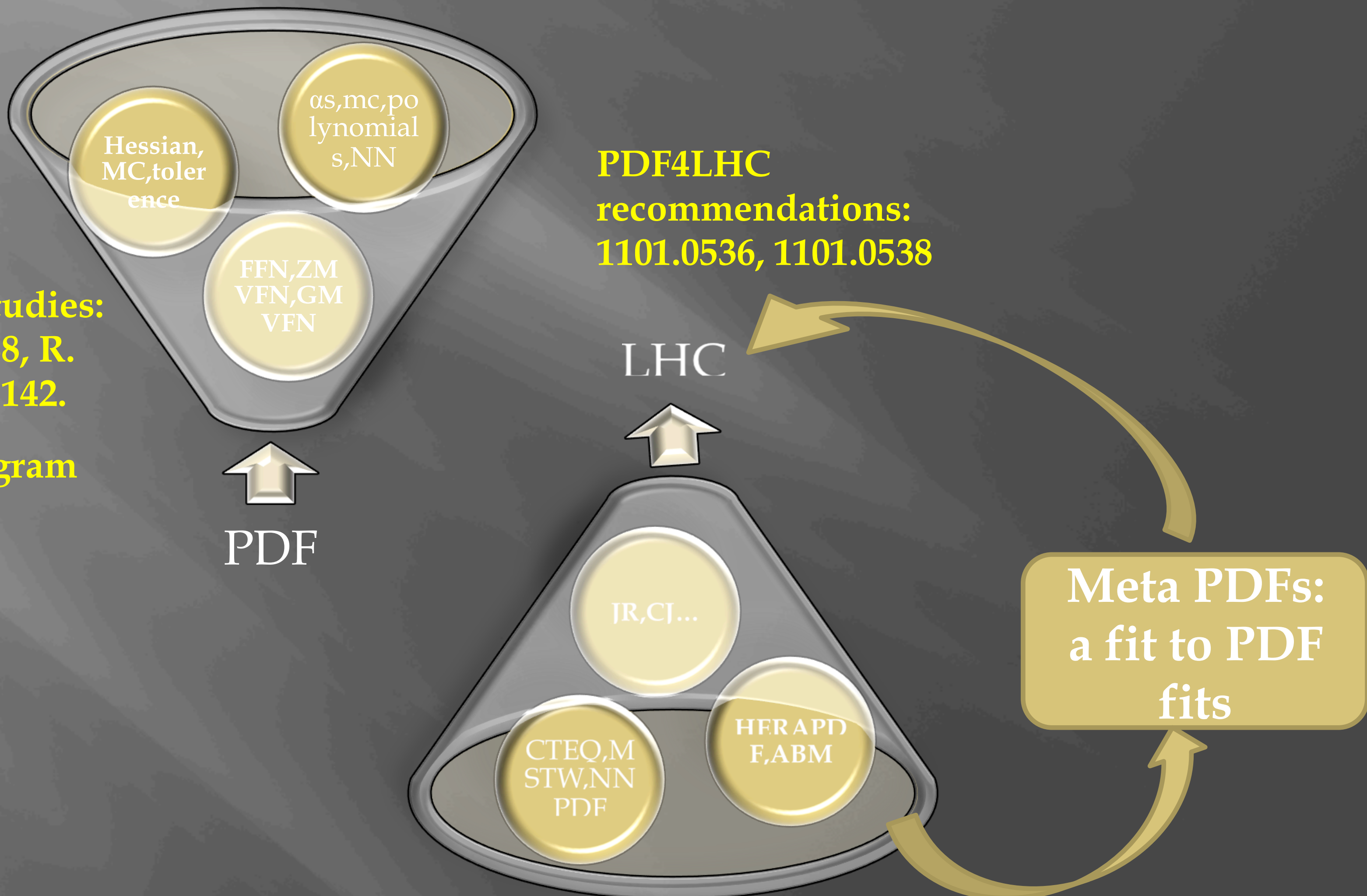
Comparing and combining PDFs from different groups

Often, it is necessary to provide a central pQCD prediction and PDF uncertainty using PDF sets from several groups. However, these PDF sets assume distinctly different physics inputs and can not be combined in a simplistic way.

PDF4LHC recommendations: 1101.0536, 1101.0538

Benchmarking studies: G. Watt, 1106.5788, R. Ball, et al., 1211.5142.

HERA Fitter program



A meta analysis of parton distribution functions for the LHC

For many studies characterized by hard scales $Q>m_{bottom}$~4.5 GeV, all existing PDFs can be represented by a shared parametric form that we call a META parameterization.

This parametric form should be flexible enough to accommodate the full span of variations in the PDF shapes that are allowed by the experimental data.

When PDFs from all groups are converted into the META representation, it is possible to combine them, obtain a central set for the totality of PDFs from all groups, and generate a small number of Hessian or Monte Carlo error sets for estimation of the combined PDF uncertainty.

ALL RESULTS TO BE SHOWN ARE PRELIMINARY, SERVE PRIMARILY TO PROVE THE VIABILITY OF THE META COMBINATION

A meta analysis of parton distribution functions for the LHC

The META PDFs is found by the following procedure:

1, Select several groups of PDFs to be converted into the META form
2, Select a functional form that can describe every selected PDF at an initial scale above the bottom quark mass. Fit this form to each PDF replica and validate the agreement between each original replica and its META representation in the x-Q range of the LHC measurements
3, Compute the central META PDF by averaging all META replicas.

Benefits:

1, A nature way to compare and combine the LHC predictions from different PDF groups in most processes. Works similarly to the PDF4LHC prescription, but PDFs are combined directly in the PDF parameter space
2, Especially desirable for combining a large number of PDF sets, in this case also minimizing numerical computation efforts for massive NNLO calculations.

A functional form for META PDFs

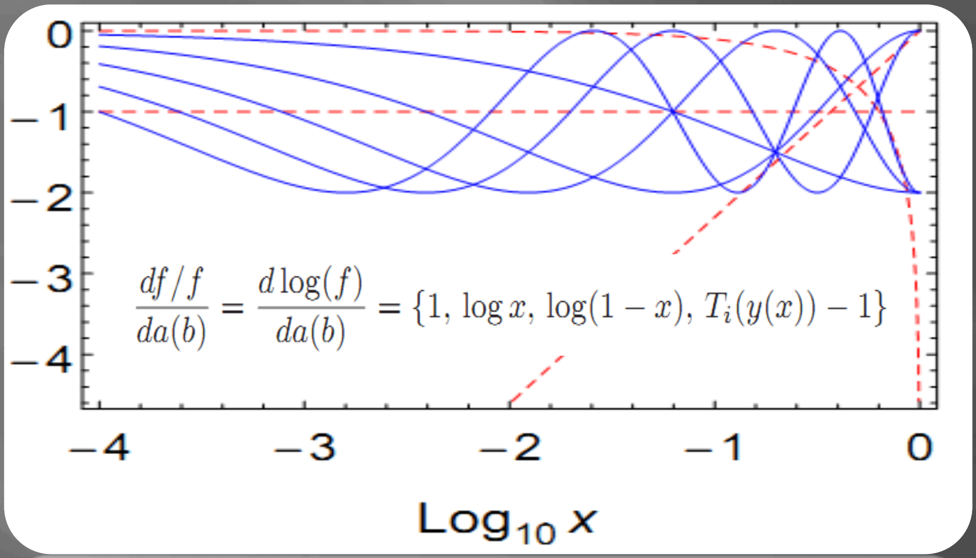
PRELIMINARY

We work with 9 PDF flavors, including strangeness asymmetry, and parameterize each of them as

J. Pumplin, 0909.5176, A. Glazov, et al., 1009.6170, A. Martin, et al., 1211.1215

$$f(x, Q_0) = a_1 x^{a_2} (1-x)^{a_3} e^{\sum_i b_i (T_i(y(x))-1)}$$

The input scale is set to be Q_0=8 GeV. The exponential contains Chebyshev polynomials T(y) with y(x)=cos(πx^β) and β=1/4.

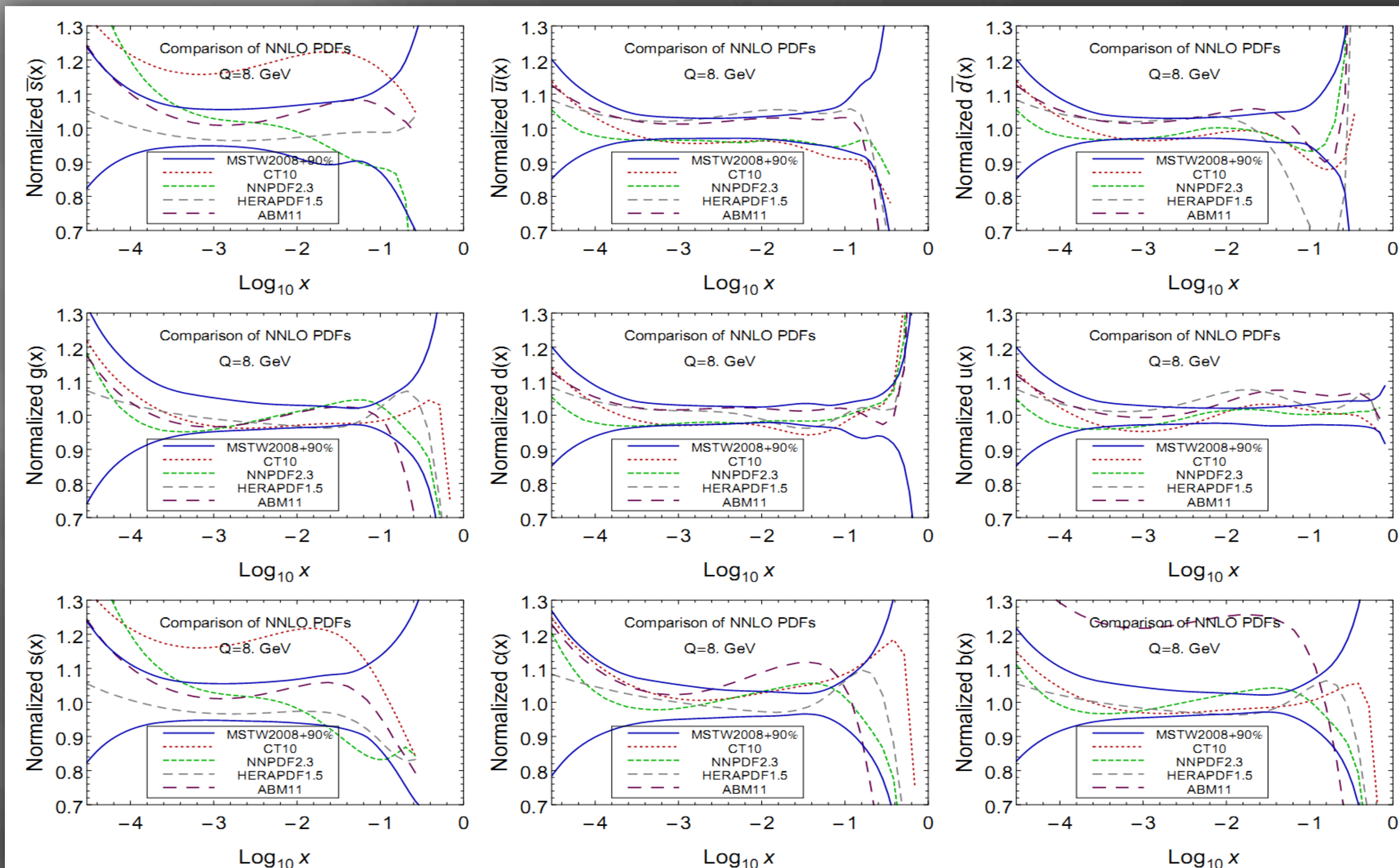


We focus on the x range with the lower limit of 3×10^{-5} for all flavors and upper limits of 0.4 for ubar, dbar; 0.3 for s, sbar; and 0.8 for other flavors. PDFs outside these x regions are determined entirely by extrapolation.

PRELIMINAR

Examined PDF sets

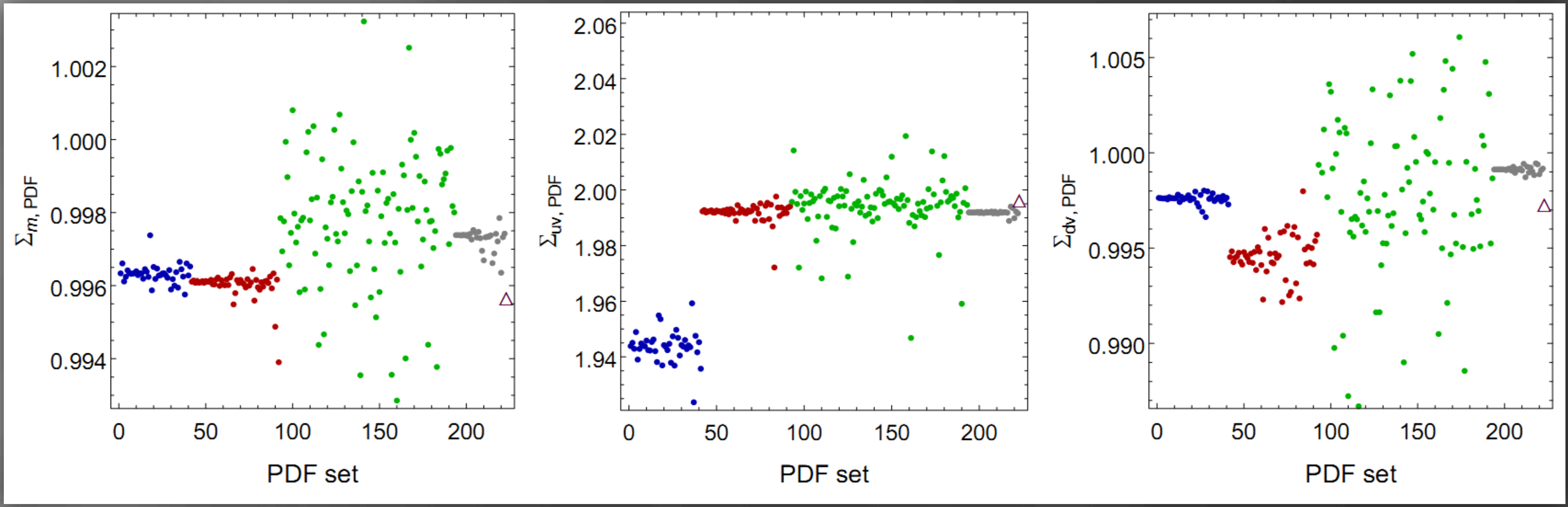
Our sample includes NNLO PDFs from CT10, MSTW2008, NNPDF2.3, HERAPDF1.5 and ABM11 with 5 flavors. The default $\alpha_S(M_Z)$ values are 0.118, 0.11707, 0.118, 0.1176, and 0.118, respectively.



Validation of the META fit: test of QCD sum rules

PRELIMINARY

The momentum , u-valence, and d-valence sums in the specified x range at Q=8 GeV for all the NNLO PDFs, including their error sets.



Left to right: blue(MSTW), red(CT10), green(NNPDF), gray(HERA), magenta(ABM)

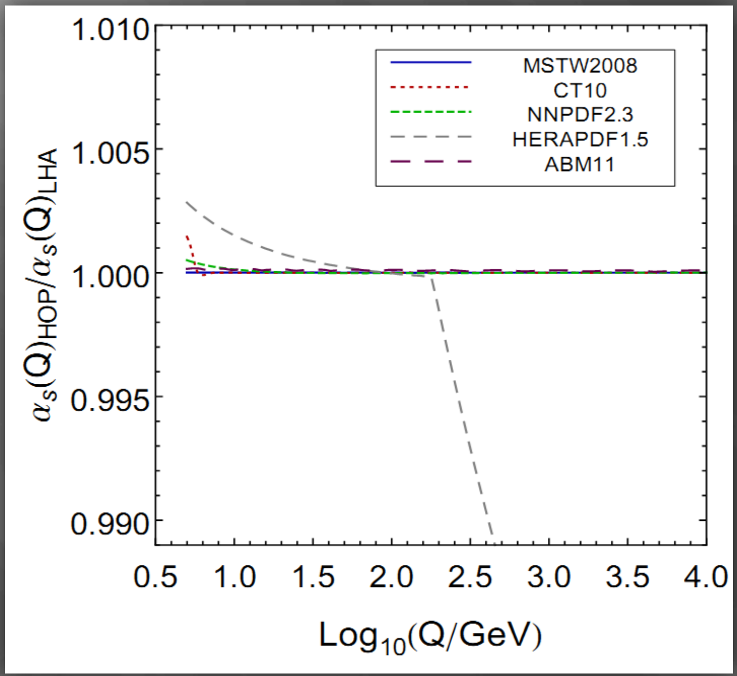
All the NNLO PDFs obey well the sum rules, with an average of 0.996 for the momentum sum, 1.98 for the u-valence, 0.997 for the d-valence. In principle we can apply the sum rule constraints as usual with above averages in our META fit, but not necessary.

PRELIMINARY

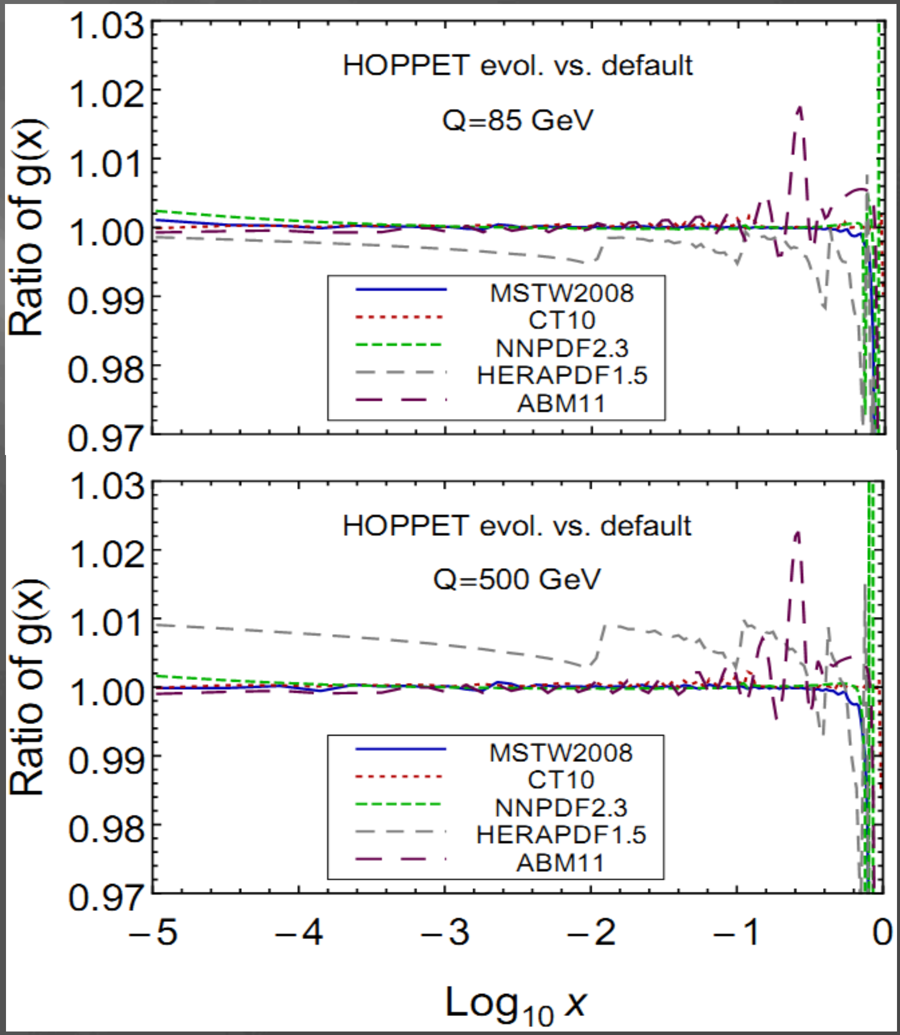
Validation of the numerical DGLAP evolution

Above Q_0=8 GeV, all selected PDFs are expected to follow the same DGLAP evolutions with 3-loop MSbar kernels and N_f=5. Before combining PDFs, it is essential to confirm that their Q dependence tabulated in the grid files is compatible across various sets

They agrees with the explicit numerical evolution from Q_0=8 GeV with HOPPET (Salam, Rojo, 0804.3755) very well



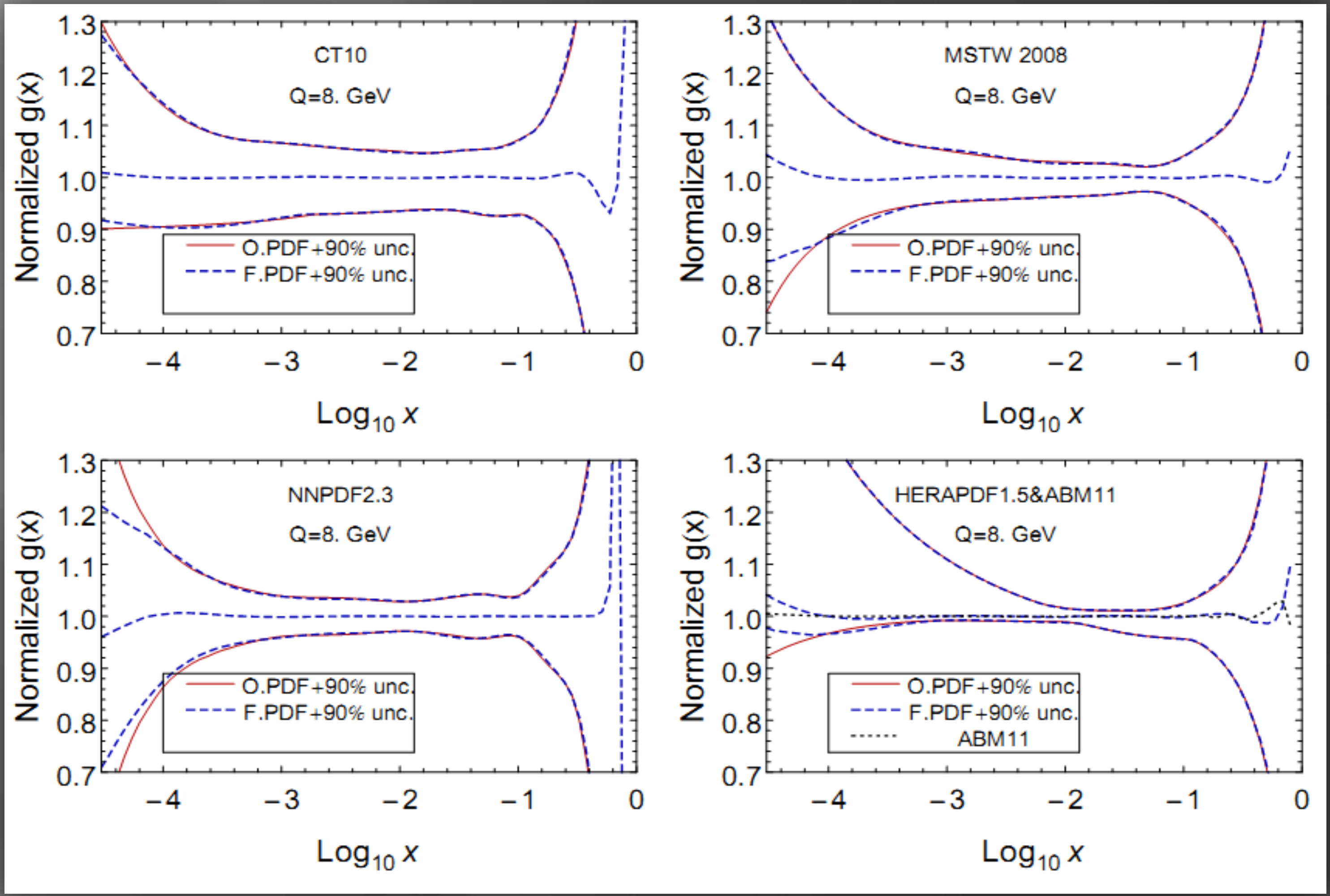
Example: evolution of the gluon PDF



Agreement of the original (O) and fitted (F) PDFs at Q_0

PRELIMINARY

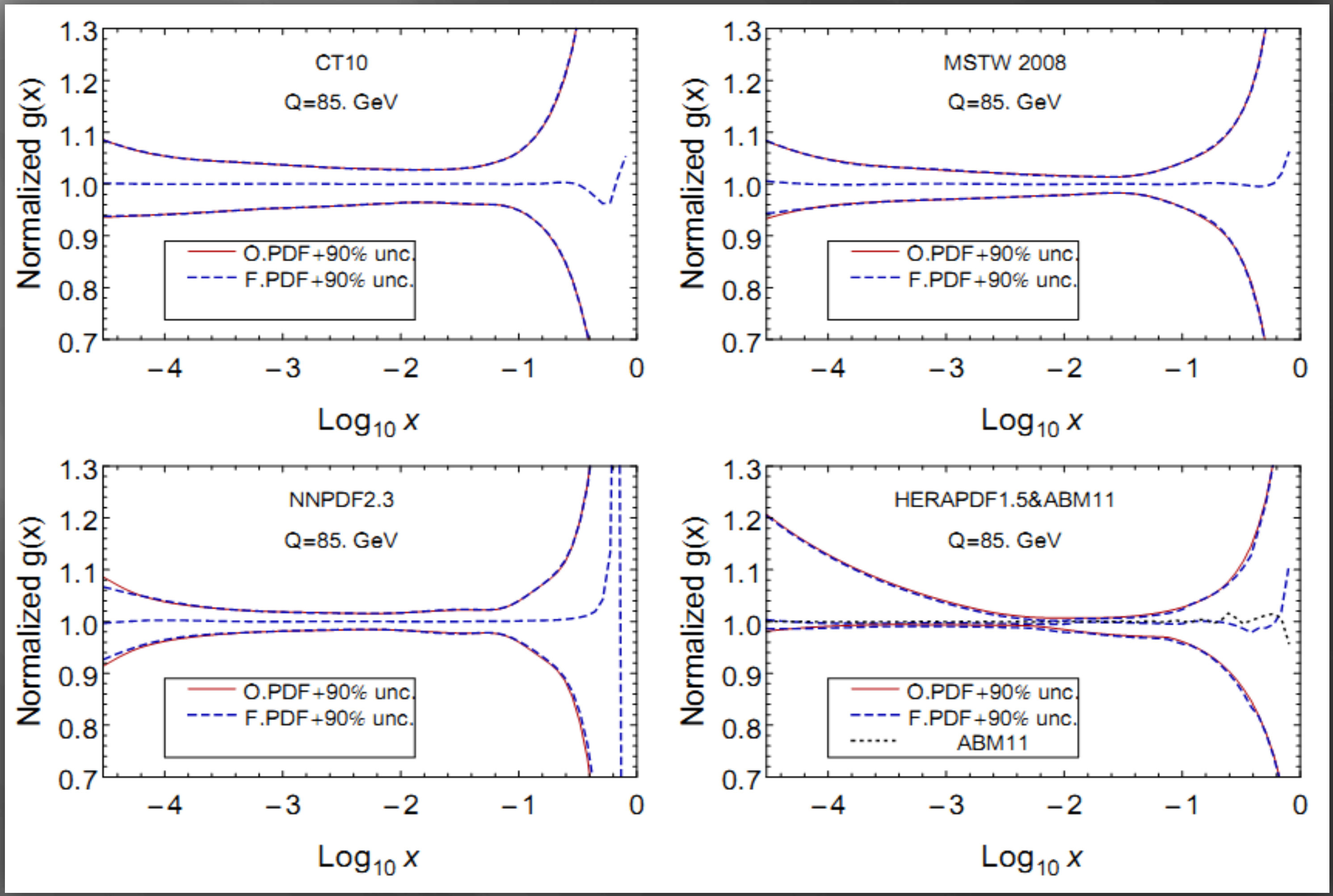
Using the functional form shown before, we include the polynomials up to the 5th order for the u, d quarks and gluon, and 4th order for other flavors. That's 66 PDF parameters in total. We fit PDFs including their error sets, then compare the central set and PDF uncertainties from O.PDF and F.PDF.



Agreement of the original and fitted PDFs at arbitrary Q

PRELIMINARY

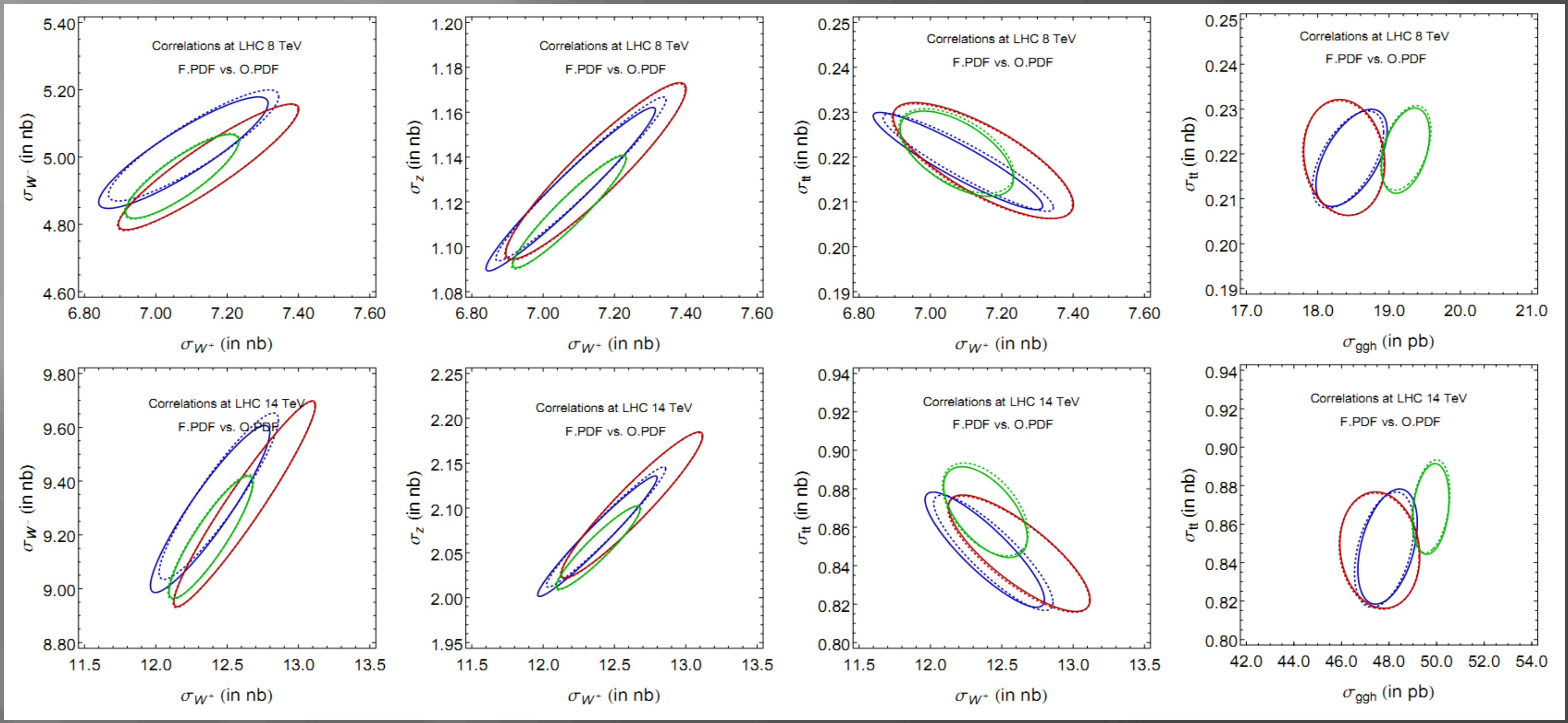
The meta PDFs are fitted at Q=8 GeV and evolved to higher Q using a common numerical program, HOPPET, then compared to the original PDFs at same scales. Excellent agreement, minor discrepancies at small x are further reduced by evolution.



META predictions for benchmark LHC processes

PRELIMINARY

For NNLO inclusive rates of W, Z, Higgs, top quark pair production, NLO jet cross sections in different kinematic bins, at the LHC 8 and 14 TeV, the fitted PDFs can well reproduce predictions of the original PDFs including the PDF induced correlations.



red(CT10), blue(MSTW), green(NNPDF), solid(dotted) for original(fitted) PDFs

Combining PDFs from different groups

Once the original PDF samples are faithfully converted into their META forms, we can combine PDF sets from all groups into one META PDF set

Example: combining CT10, MSTW2008, NNPDF2.3 sets

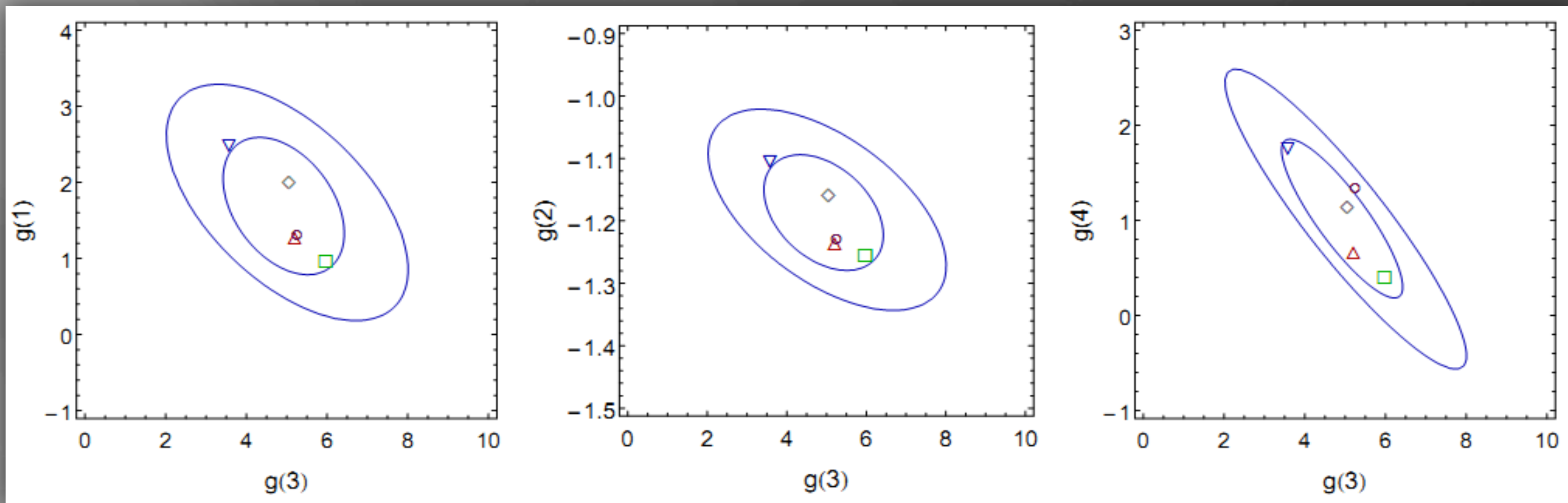
1,Generation of replicas. The PDFs of the three groups at $\alpha_S(M_Z)$=0.118 are generally compatible with each other. Knowing the PDF eigenvectors from each set, we can select 100 MC replicas for each set or generate them for CT10/MSTW using a method similar to the MSTWMC study. Note the differences between the Hessian and MC interpretation of statistical features. We assume the Gaussian distribution in the cases of CT10 and MSTW when generating the replicas.

G. Watt, et al., 1205.4024

2,Averaging all samples. Merge them and get 300 MC replicas. Perform the fit and get the covariance matrix in the PDF parameter space. Reduction of the systematic errors but not experimental errors. Assuming Gaussian distributions in the PDF parameter space, we can find the eigenvector directions, drop ones with small eigenvalue, and arrive at a "Hessian-like" META PDF with 50 eigenvectors (100 error sets).

Example: META PDF parameters for the gluon

PRELIMINARY



up-triangle(CT10), down-triangle(MSTW), square(NNPDF), diamond(HERA), circle(ABM), together with the 90% error ellipse (larger one) of the META PDF

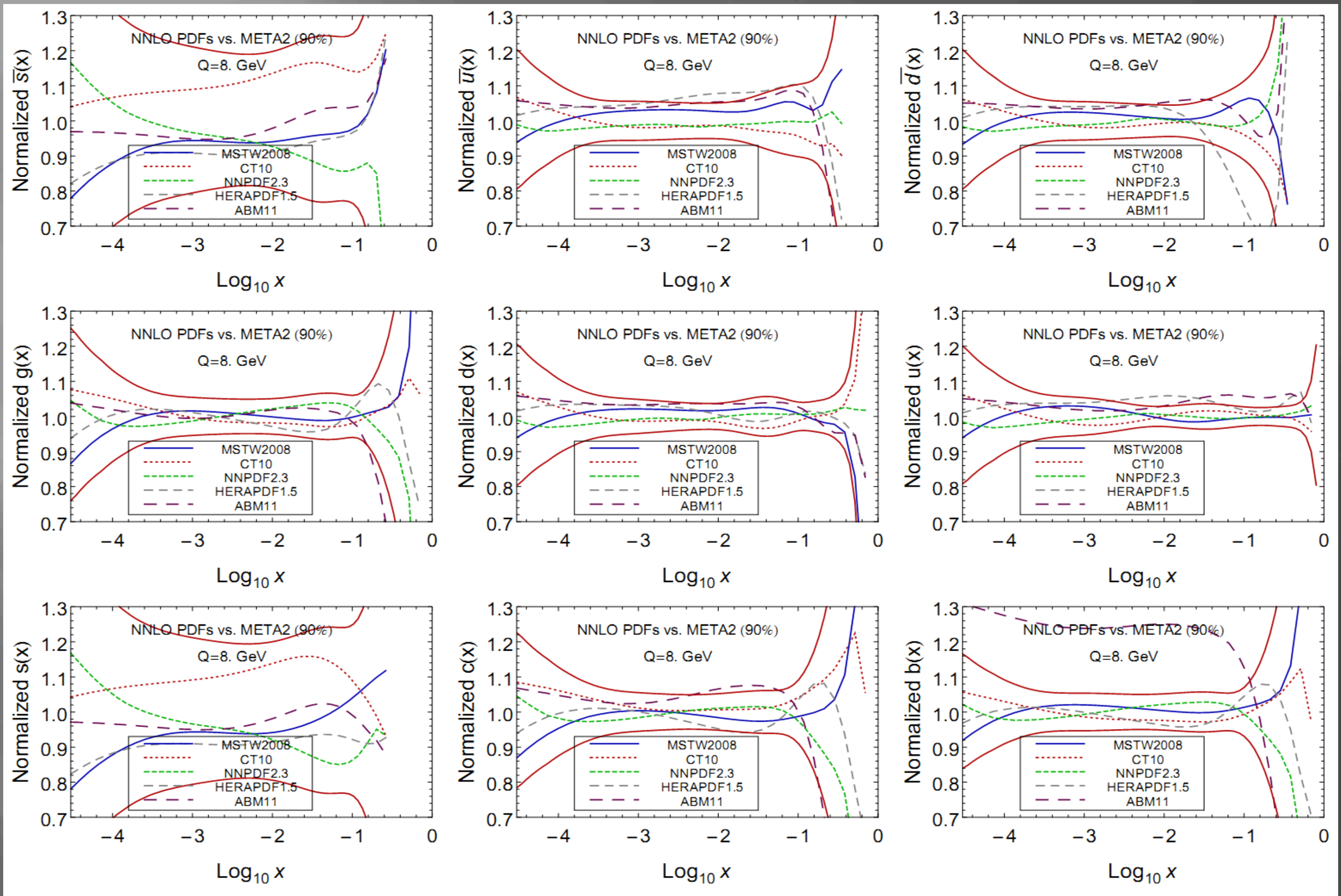
We also provide the PDF α_S series for the META PDF by simply taking the average of the corresponding PDFs from the 3 groups, that can be used to study the α_S uncertainties in the META PDF framework.

We enlarge the α_S variation range to 0.118±0.002 for 68% C.L. and add the α_S and PDF uncertainties linearly to account for the differences between the world average and the preferred value from the global fit of PDFs.

The META PDF

PRELIMINARY

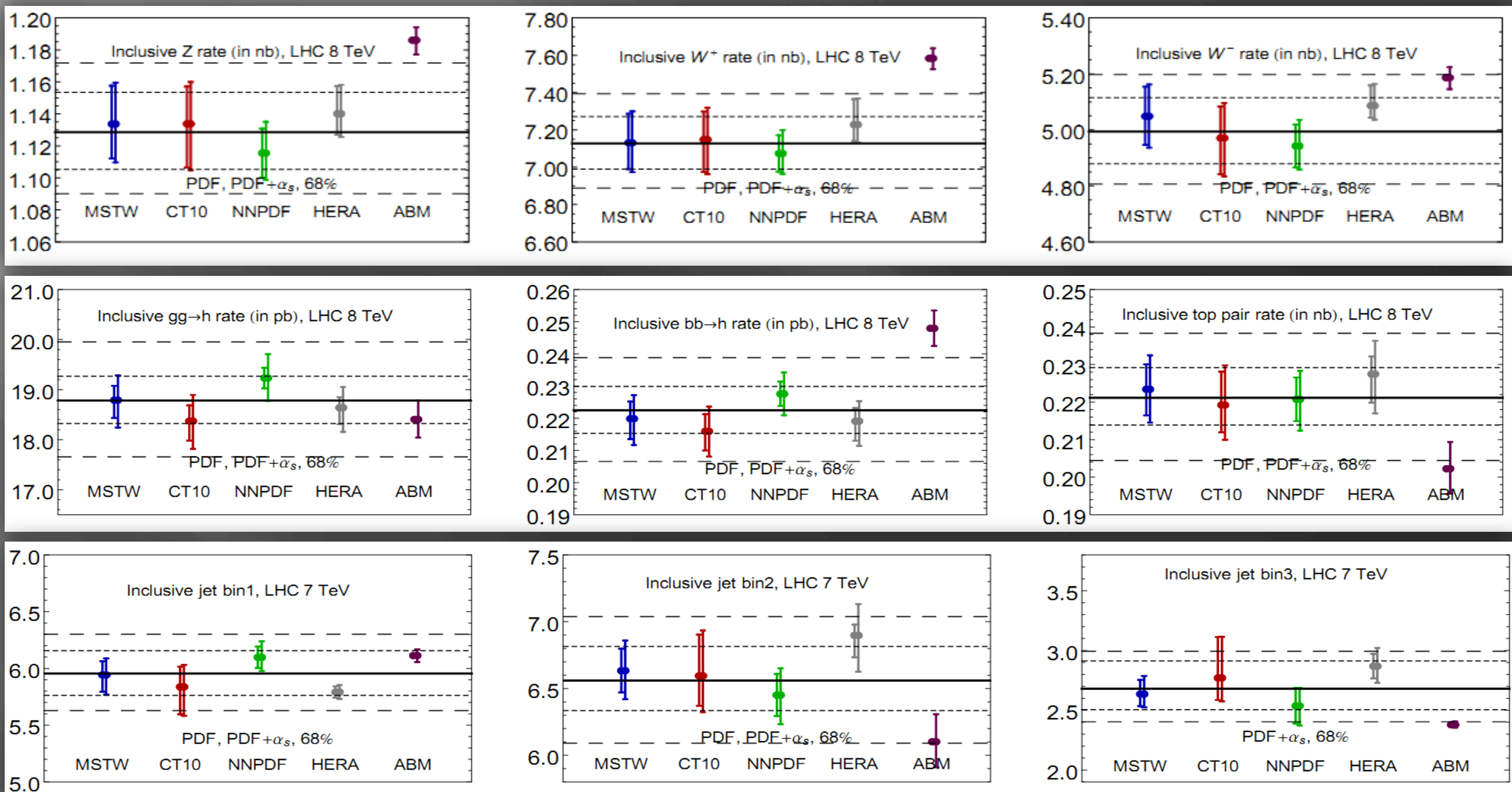
Comparisons of the META PDF with the original central PDF from each groups with $\alpha_S(M_Z)$=0.118 at Q=8 GeV.



PRELIMINAR

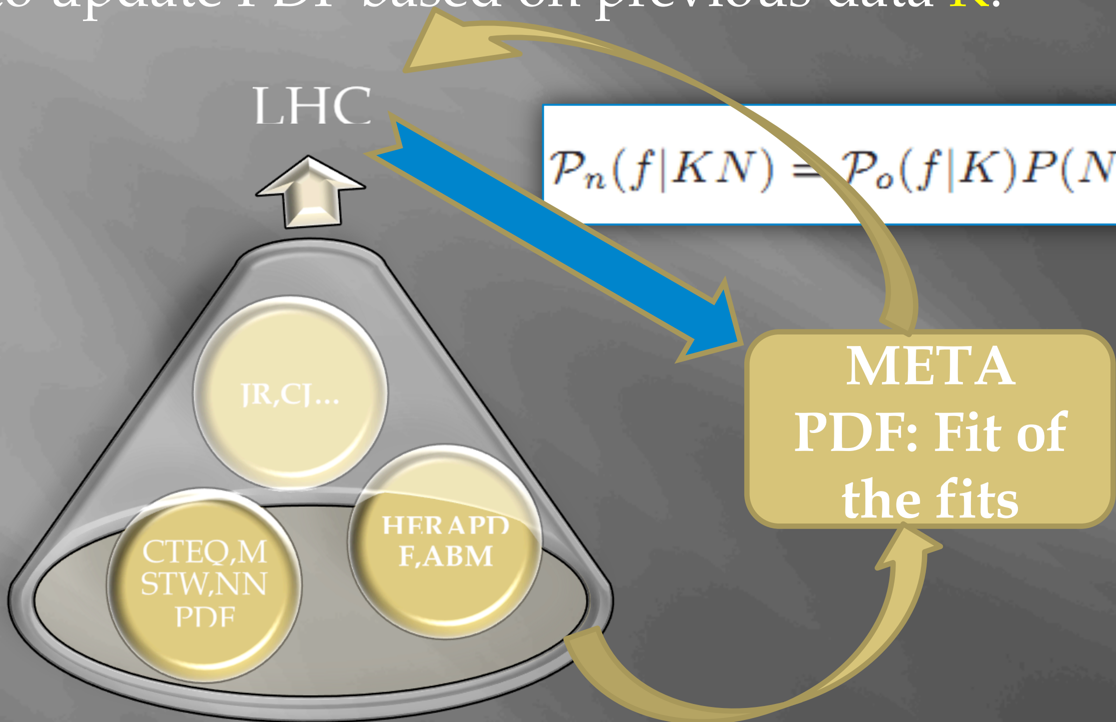
META predictions for the LHC

Comparisons of the LHC predictions, including central prediction, PDF uncertainties, and PDF+alphas uncertainties, at 68% C.L.. Similar results comparing to the envelope prescription in the benchmark study **(R. Ball, et al., 1211.5142)**, e.g., for gg→h, 18.75±1.24 pb there, while 18.78±1.15 pb here.



Further development: constraints from LHC

We can also use the precise LHC data to further constrain the META PDF using PDF reweighting based on the Bayes' theorem. For example, using new data N to update PDF based on previous data K:



$$\mathcal{P}_n(f|KN) = \mathcal{P}_o(f|K)P(N|fK)/\int P(N|fK)\mathcal{P}_o(f|K)Df$$

W. Giele, et al., hep-ph/9803393

Prior pdf of META PDF:

$$\mathcal{P}_o(f|K) \sim e^{-(\sum_{i=1}^{50} x_i^2)/2}$$

Ambiguity in how to take the limit of the infinitesimal volume of the observed data N as pointed out in **(1012.0836, R. Ball, et al.)**, e.g., resulting different weight choices as $\exp[-\chi^2(N|f)/2]$, $(\chi^2)^{n/2-1}\exp[-\chi^2(N|f)/2]$, $(\chi^2)^{(n-1)/2}\exp[-\chi^2(N|f)/2]$, …. We follow a principle that the new data set N should be treated equally as the original data set K. By another meaning, the reweighting should give results close to the refitting with new data N included.

Further development: reweighting schemes

We explore several possible choices for the META PDF

Scheme A: assuming a quadratic dependence of $\chi^2(N|f)$ on PDF parameters x_i , it is straightforward to prove that for the HERA-like fit ($\Delta\chi^2=1$), HERAPDF or ABM, the PDF reweighting with weight $\sim\exp[-\chi^2(N|f)/2]$ is exactly equivalent to the corresponding refitting. Gaussian→ Gaussian.

Scheme D: one variation of scheme A can be motivated by the CTEQ total χ^2 tolerance criterion. $\Delta\chi^2=100$ for 90%, translated to $\Delta\chi^2=h_0=37$ for 68%, and the weight function $\sim\exp[-\chi^2(N|f)/(2h_0)]$.

Scheme B: using the same weight $\sim\exp[-(\chi^2-(n-1)\ln \chi^2)/2]$ as NNPDF, but only keep up to the quadratic terms on x_i in the exponential, so we still get a Gaussian after reweighting.

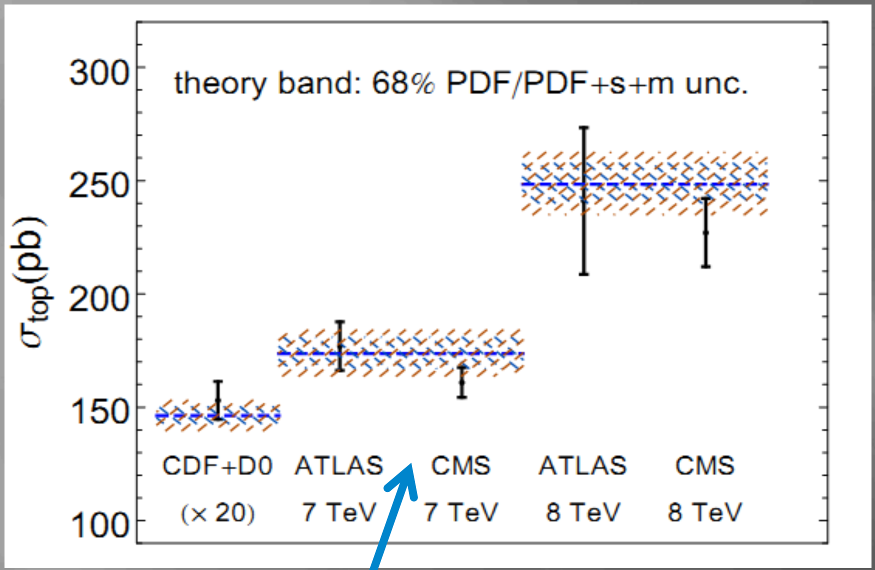
Scheme B*: first generating 50,000 unweighted MC replicas based on the prior of META PDF, then reweight them using the exact NNPDF function form.

Scheme C: MSTW-like, here we fix the best-fit and eigenvector directions. The new PDF uncertainties are determined by the minimum of the original displacements and the newly allowed ones (according to MSTW dynamic tolerance) by data N in each of the directions.

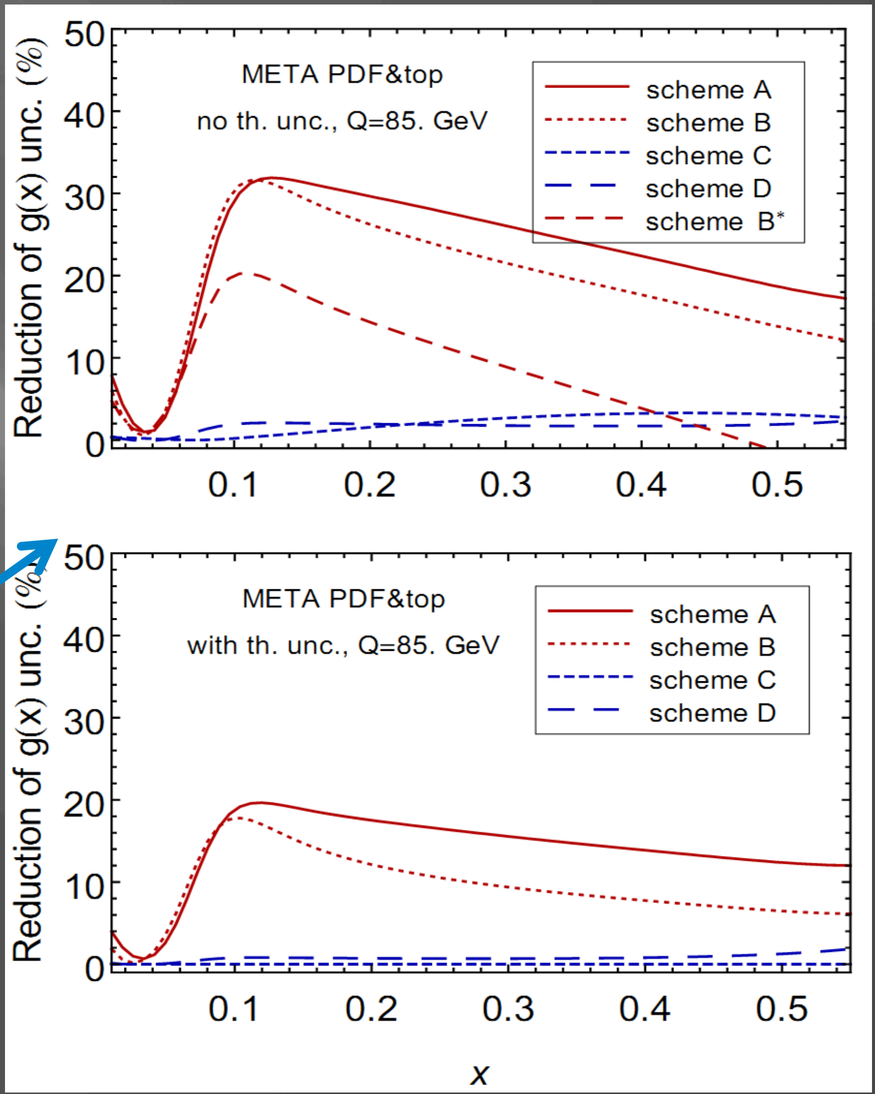
PRELIMINARY

Examples: top quark data

We perform a similar study as in (1303.7215, M. Czakon, et al.) using the measurements of top quark pair inclusive rate to constraint the gluon PDFs.



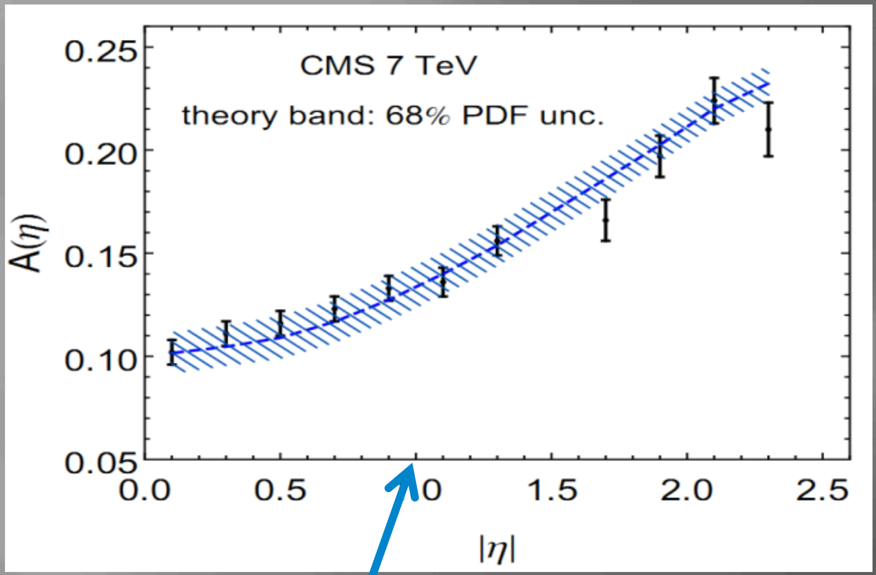
Comparison of META predictions with data before reweighting

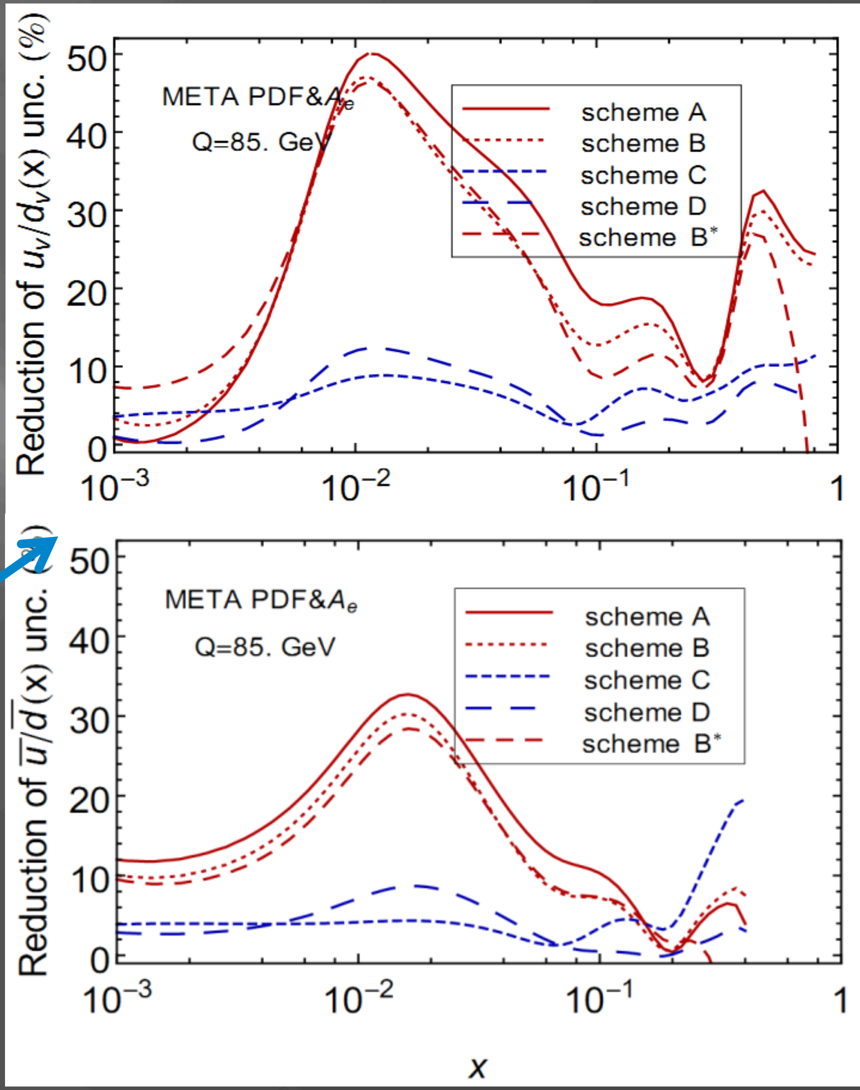


Reduction of the gluon PDF uncertainties under different schemes with and without including theoretical uncertainties.

Examples: CMS electron charge asymmetry

We also study the effects of the CMS measurements of W electron charge asymmetry (840 pb^{-1}) on the u-valence, d-valence, ubar and dbar PDFs.





Comparison of META predictions with data before reweighting

Reduction of the PDF uncertainties of the ratios uv/dv and ubar/dbar under different schemes

Conclusions

We performed a META analysis of the NNLO parton distribution functions focusing on LHC phenomenology studies. The META PDF serves as an average of the chosen PDFs for central predictions and also gives a reasonable estimation of the total PDF uncertainties.

It provides a natural way to compare and combine the LHC predictions from different PDF groups independently of the processes, works similarly to the PDF4LHC prescription, but directly in the PDF parameter space.

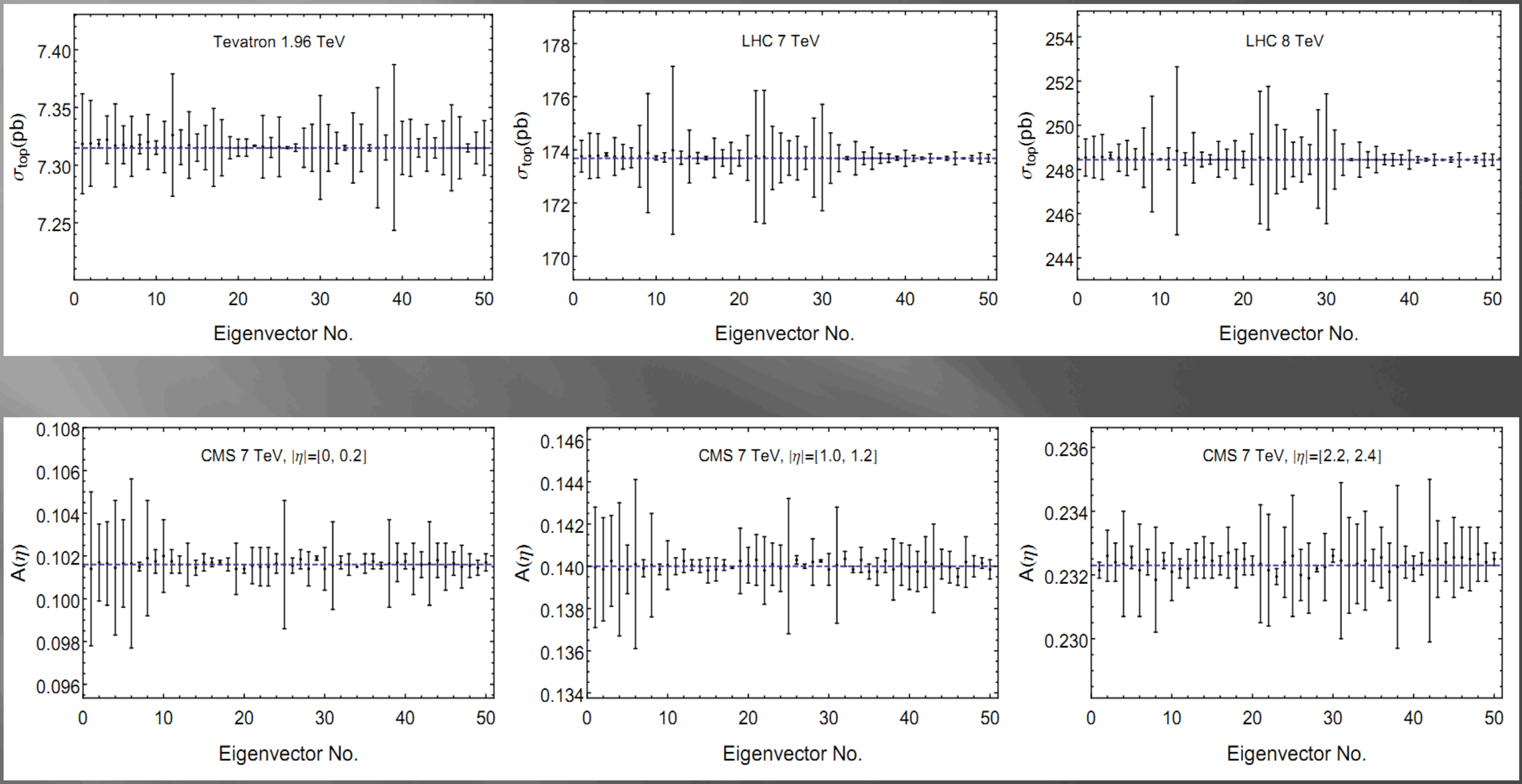
It is suitable for including results from a large number of PDF groups and minimizing numerical computation efforts for massive NNLO calculations.

We explored several possibilities of including constraints from new precise data at the LHC in the META PDF framework based on Bayesian reweighting.

PRELIMINARY

- Examine the LHC predictions from different eigenvectors

It shows the dynamic structure of different eigenvectors and validates the linear approximations.



Examine the sum rule and PDF correlations of META PDF

PRELIMINARY