A meta analysis of parton distribution functions for LHC applications

Jun Gao
Southern Methodist University
Dec 13, 2013

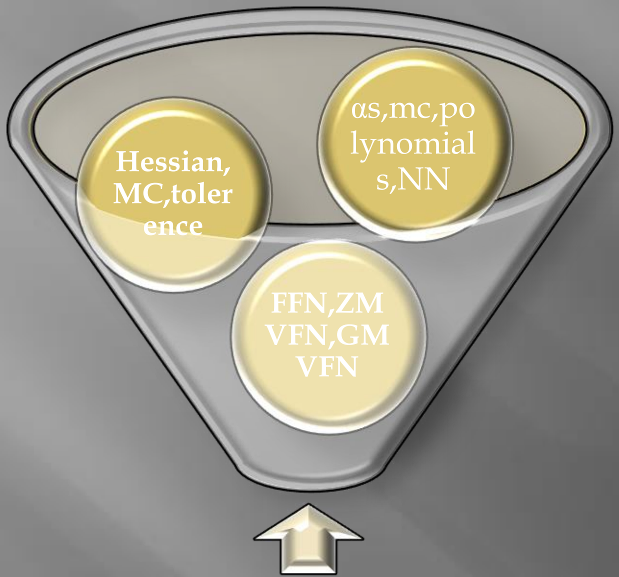
In collaboration with Pavel Nadolsky, arXiv:1312.xxxx

http://metapdf.hepforge.org

Introduction

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

PDFs at the LHC era: aiming at higher accuracies, great diversities



PDF

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

HERA Fitter program

PDF4LHC recommendations: 1101.0536, 1101.0538, or envelope prescription,1211.5142

LHC

Requiring fast calculation programs, especially at NNLO, for LHC applications, e.g., FastNLO, ApplGrid

Done process by process

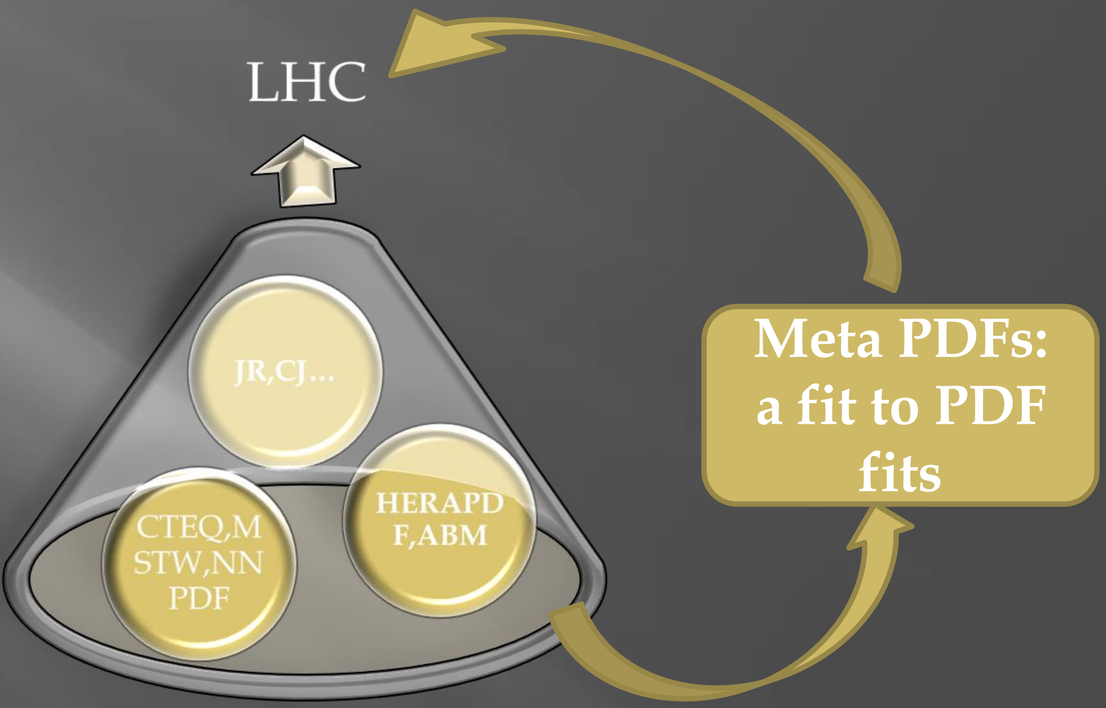
Large meta data bases through LHAPDF

The idea of META PDFs

A meta-analysis of PDFs can be first performed to provide an estimation of central prediction and uncertainty using PDF sets from several groups. Then the new PDF set can be applied to LHC predictions.

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 calculations.



However, different PDF sets always assume distinctly different physics inputs and can not be combined in a simplistic way.

Steps of the meta analysis

1. Select the input PDF ensembles
2. Fit each PDF sample in the input ensembles by a common functional form (“**a meta-parametrization**”)
3. Combine MC replicas of meta-parametrizations from all input ensembles into one Hessian/MC PDF set

Selection of input PDF ensembles

We select most recent NNLO PDFs with 5 active flavors from 5 groups, i.e., CT10, MSTW2008, NNPDF2.3, HERAPDF1.5, and ABM11, focusing on LHC applications.

NNLO	CTEQ	MSTW	NNPDF	HERA	ABM
Initial scale (GeV)	1.3	1.0	1.414	1.378	3.0
Error sets	50	40	100	28	Only best-fit used
α_S for error sets	0.118	0.1171	0.118	0.1176	0.118
Error type	Heesian	Hessian	MC	Hessian	Hessian

For many studies characterized by hard scales $Q>m_{bottom}\sim 4.8$ GeV, all existing PDFs can be represented by a shared parametric form that we call a META parameterization.

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.

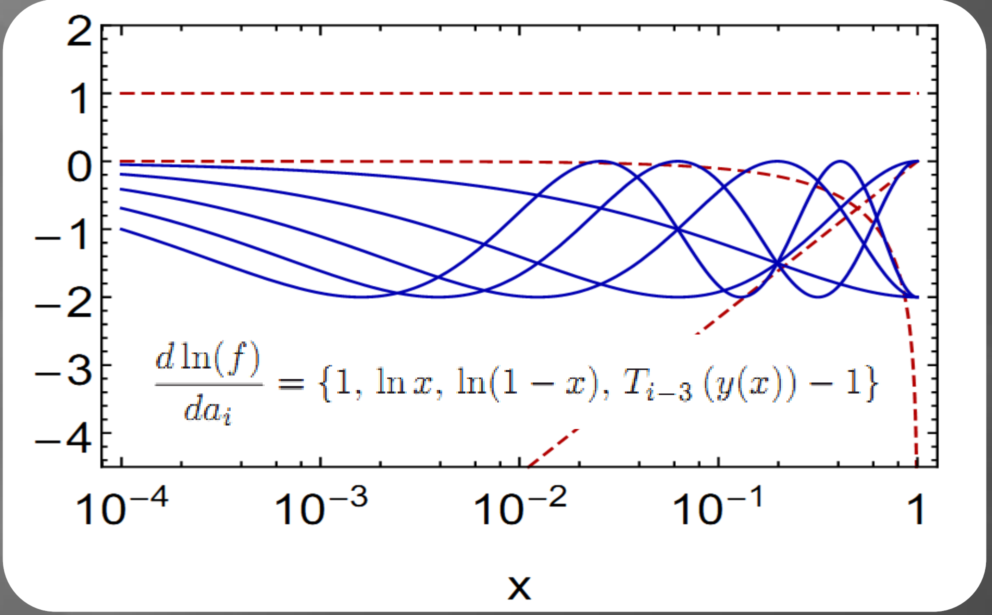
META parametrization

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

$$f(x, Q_0; \{a\}) = e^{a_1} x^{a_2} (1-x)^{a_3} e^{\sum_{i\geq 4} a_i \, (T_{i-3}(y(x))-1)}$$

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

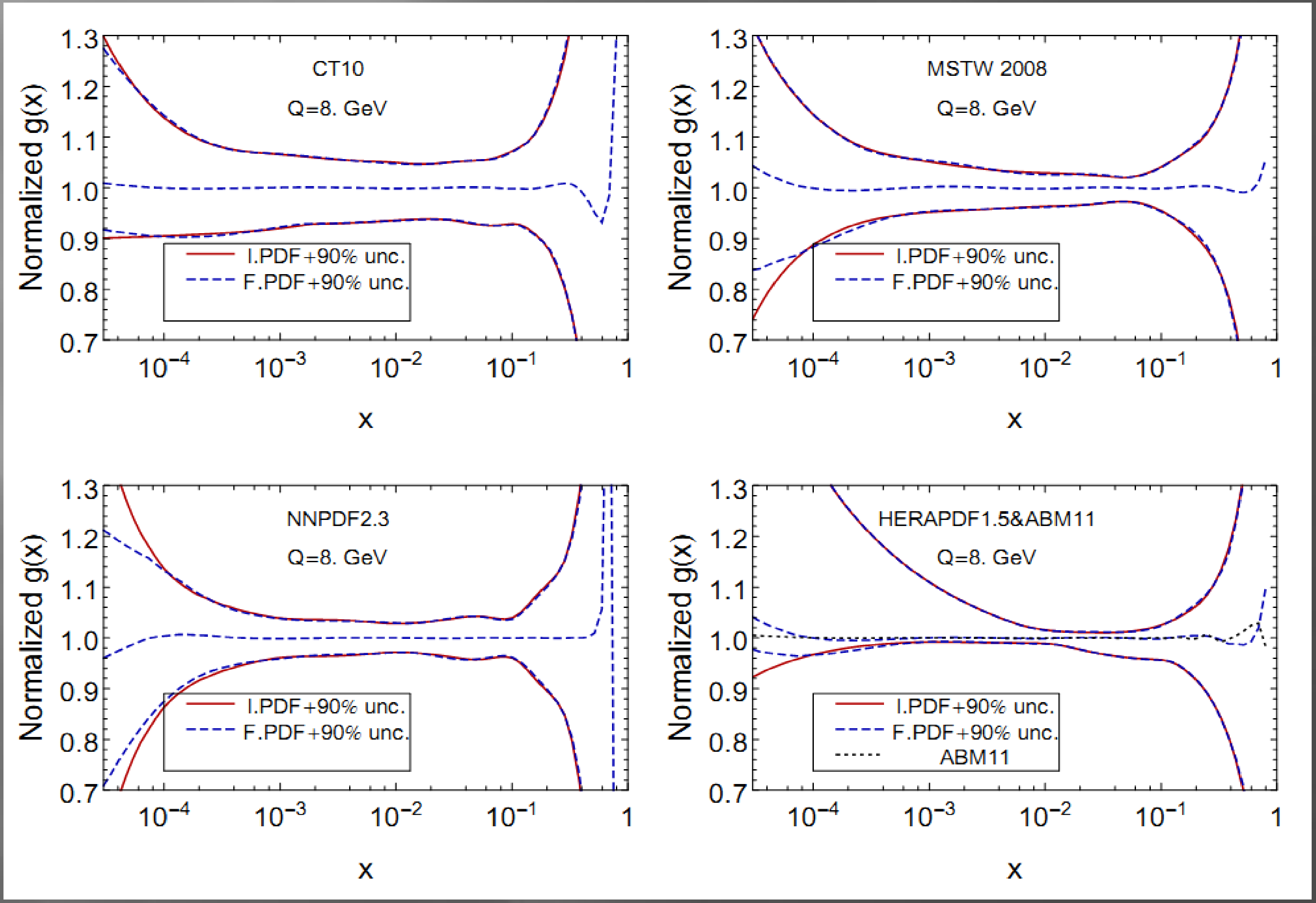
The input scale is set to be Q_0=8 GeV. The exponential contains Chebyshev polynomials T(y) with $y(x)=\cos(\pi x^{\beta})$ and β=1/4.



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.

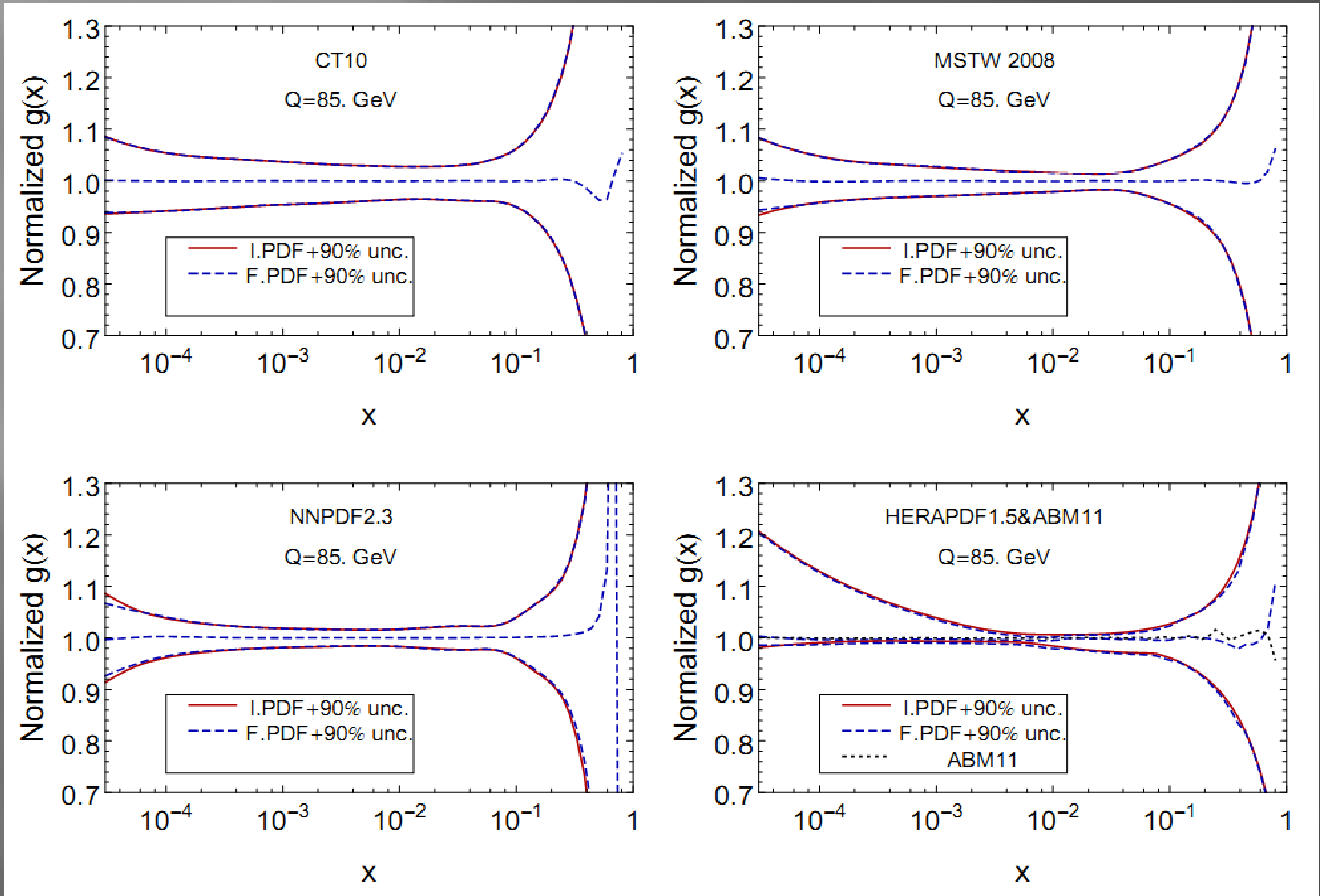
Agreement of the input (I) and fitted (F) PDFs at Q_0

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 I.PDF and F.PDF.



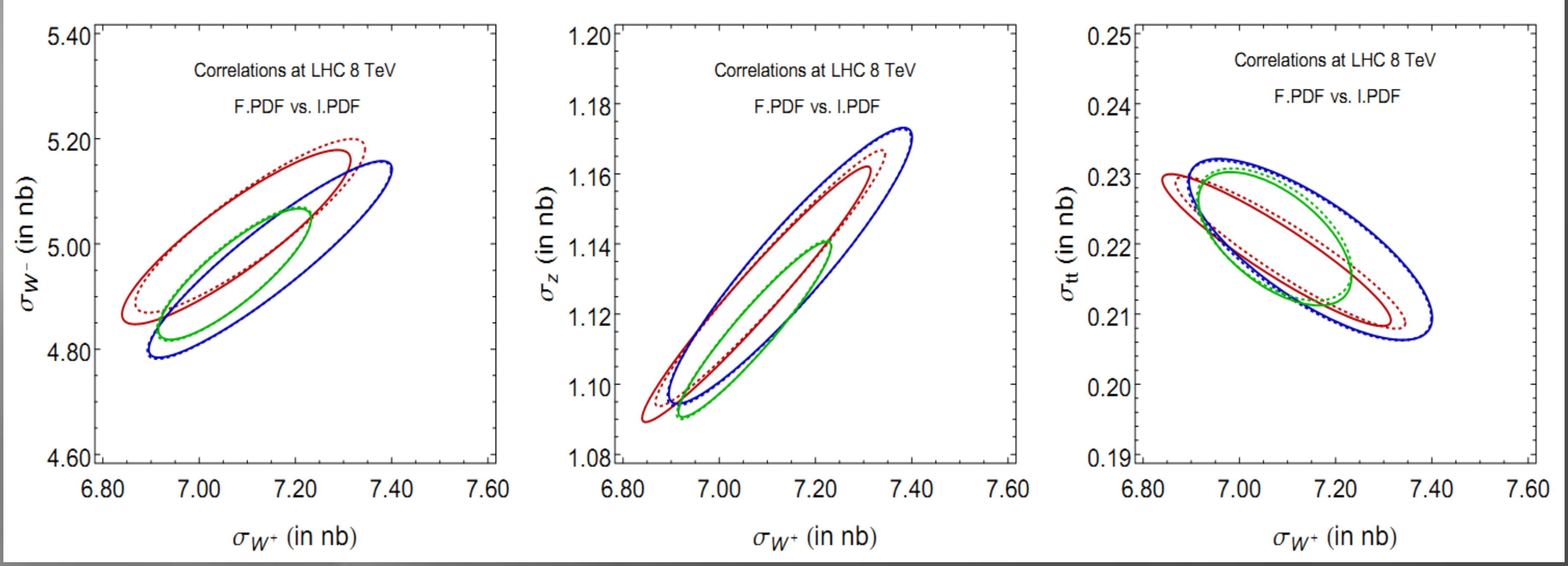
Agreement of the input and fitted PDFs at arbitrary Q

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.



Agreement on predictions for benchmark LHC processes

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.



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

Combining PDFs from different groups

Once the input 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

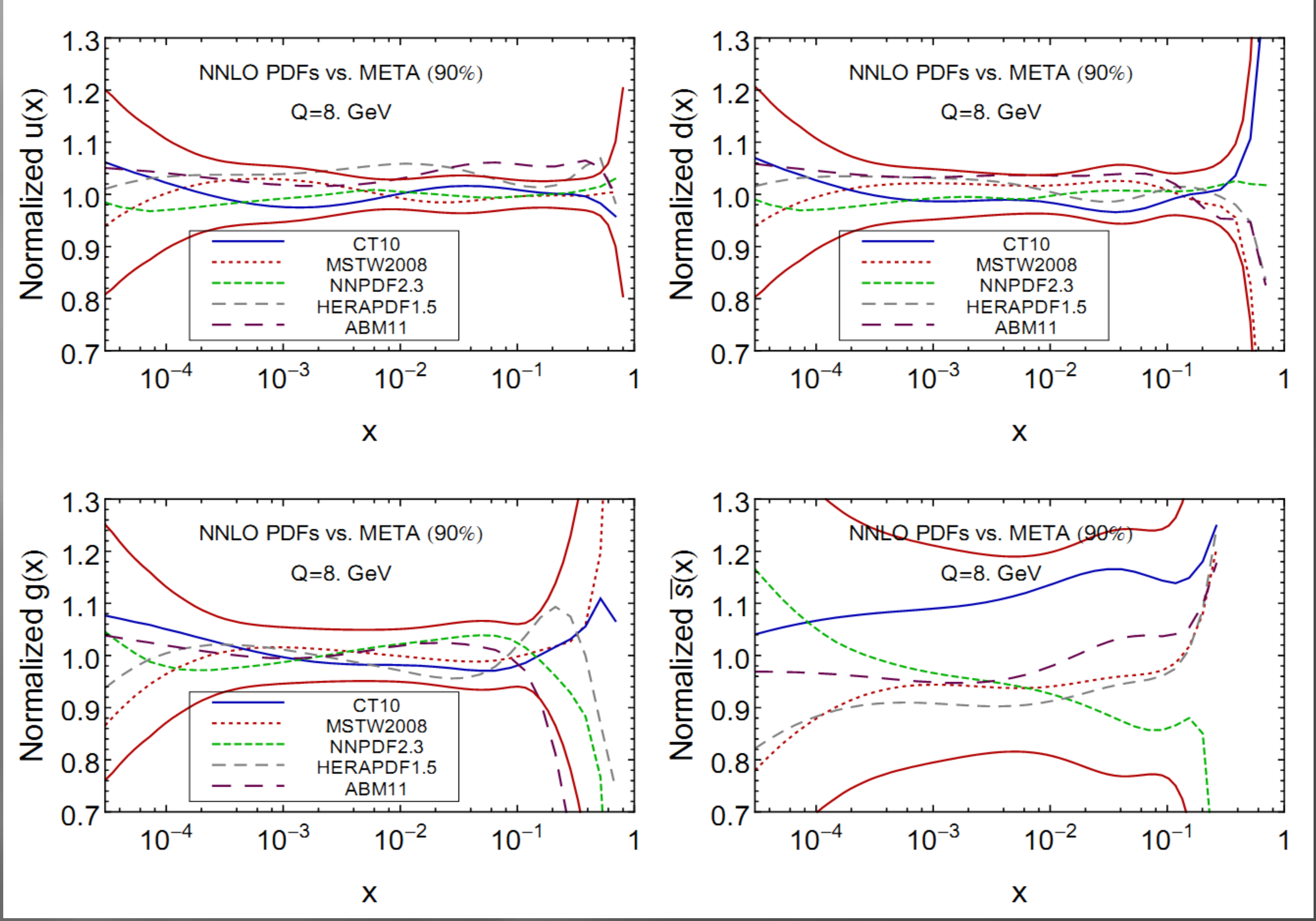
1,Generation of replicas. The PDFs of the three groups at $\alpha_S(M_Z)$=0.118 are generally compatible with each other. Knowing the PDF eigenvectors from each set, we can select 100 MC replicas for each set or generate them for CT10/MSTW using a method similar to the MSTWMC study. Note the differences 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

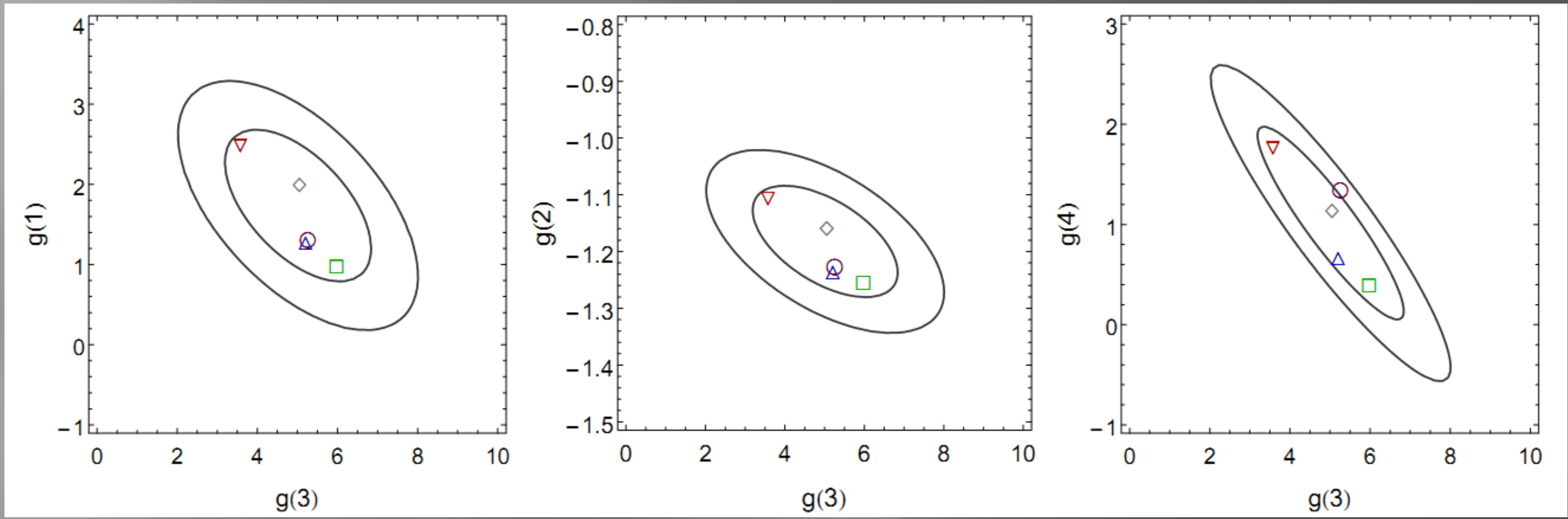
2,Averaging all samples. Merge them and get 300 MC replicas. Perform the fit and get the covariance matrix in the PDF parameter space. Assuming Gaussian distributions in the PDF parameter space, we can find the eigenvector directions, drop ones with small eigenvalue, and arrive at a Hessian META PDF with 50 eigenvectors (100 error sets).

The META PDFs

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



- Example: META PDF parameters for the gluon



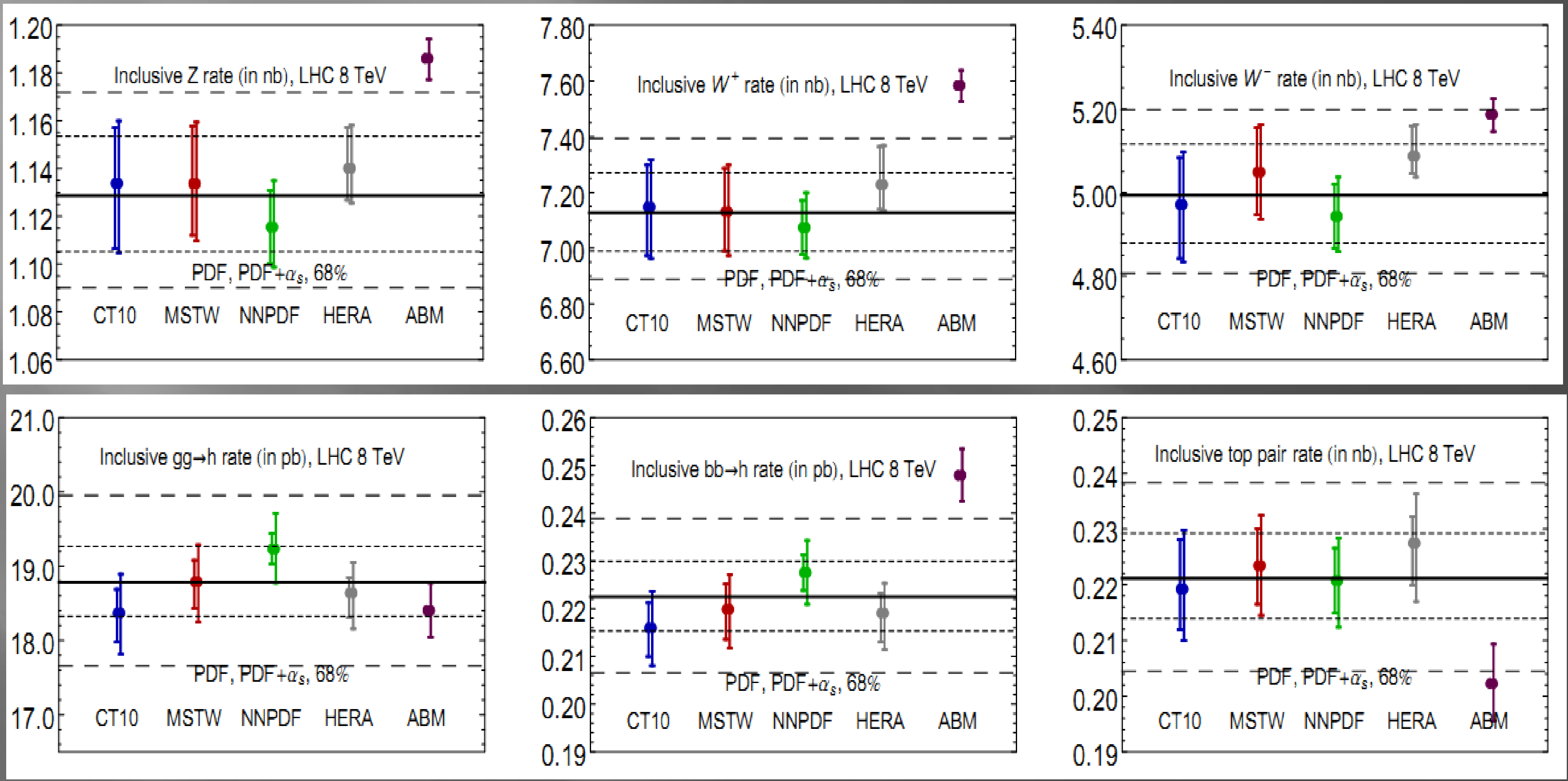
up-triangle(CT10), down-triangle(MSTW), square(NNPDF), diamond(HERA), circle(ABM), together with the 90(68)% error ellipse of the META PDF

We also provide the PDF α_S series for the META PDF that can be used to study the α_S uncertainties in the META PDF framework.

Possibilities of combining PDF and α_S uncertainties: 1, ± 0.0007 for 68% c.l., adding in quadrature; 2, ± 0.0012 for 68% c.l.; 3, ± 0.002 for 68% c.l. and adding 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.

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: data set diagonalization

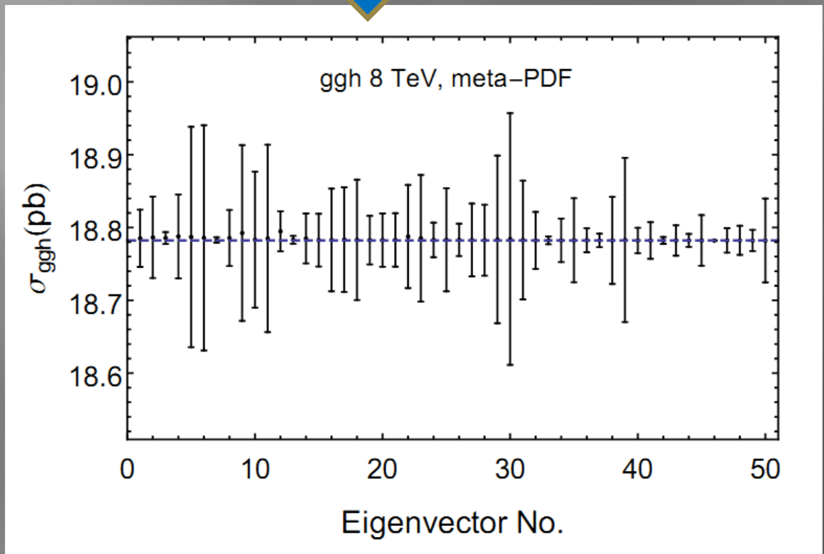
In Hessian method there are additional freedoms on choosing the eigenvector basis. A basis 2 can be produced by an arbitrary orthogonal transformation of basis 1. They represent identical measurements in PDF parameter space and predict same results for physical observables with linear dependence.

Basis 1

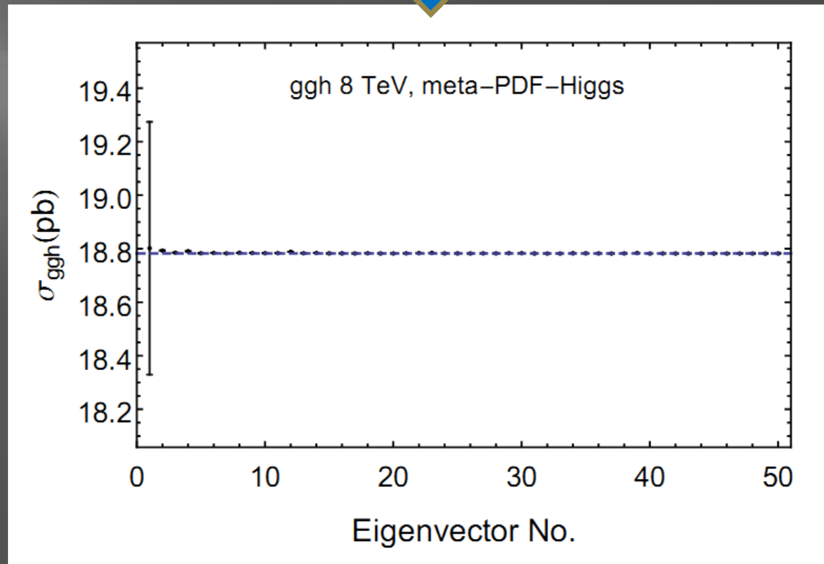
By rotation specified for some physical observables

Basis 2

J. Pumplin, et al., 2009



=

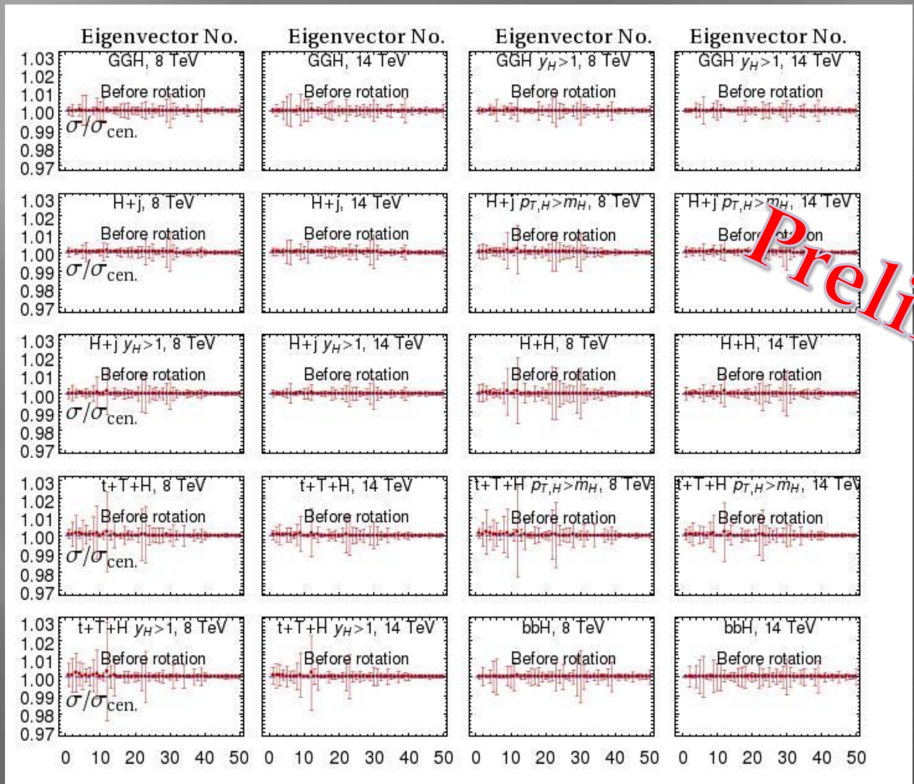


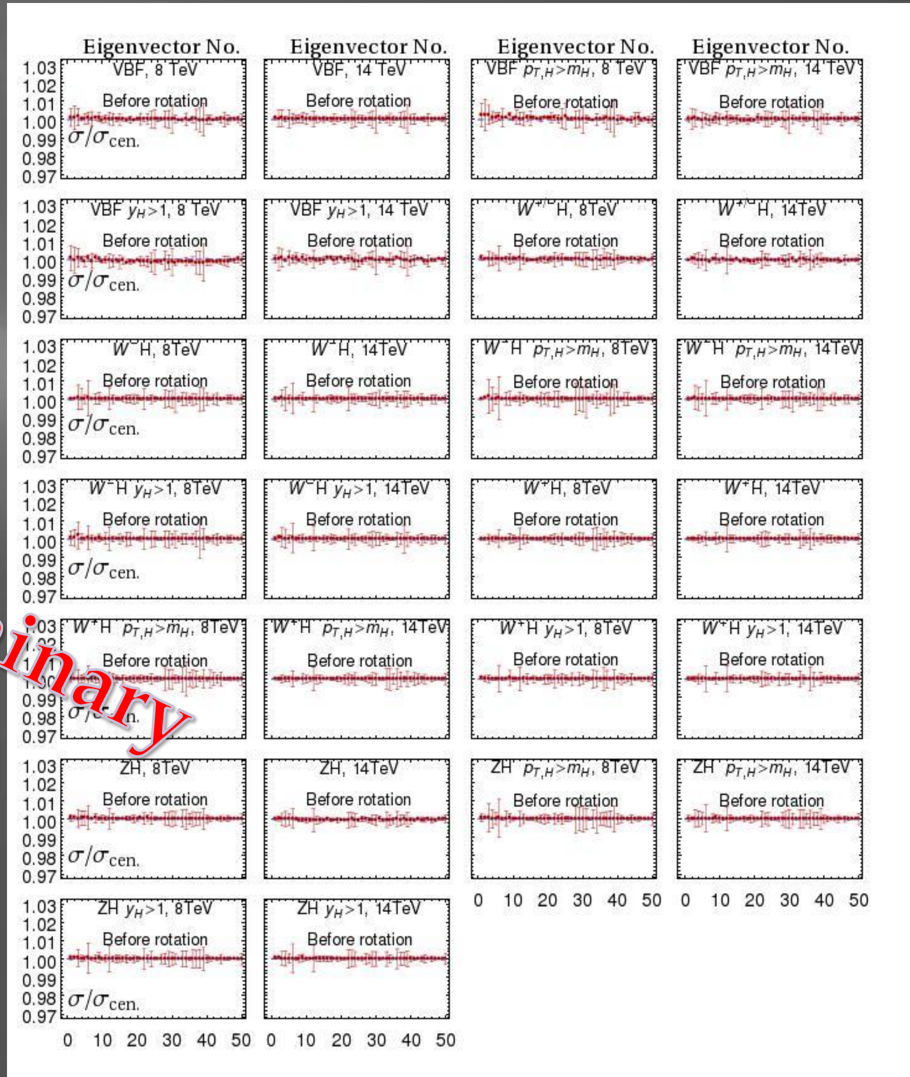
Benefit: less eigenvectors are needed for full descriptions, including uncertainties and correlations, of a certain group of physical observables

Examples: LHCH PDF set for SM Higgs boson study

We select a group of 46 physical observables, including SM Higgs production cross sections at LHC 8 and 14 TeV, in gg->h, bb->h, VBF, Wh, Zh, hh, tth, h+jet channels, for inclusive rate and rate in high pt or large rapidity region of the Higgs boson.

Predictions of regular META PDF set

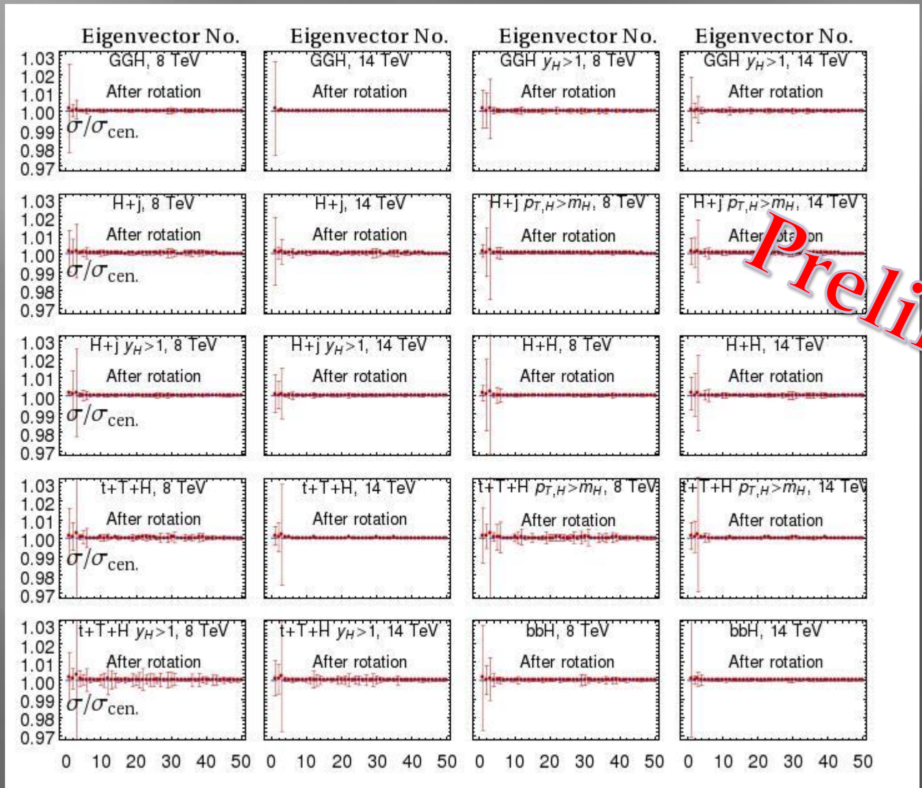


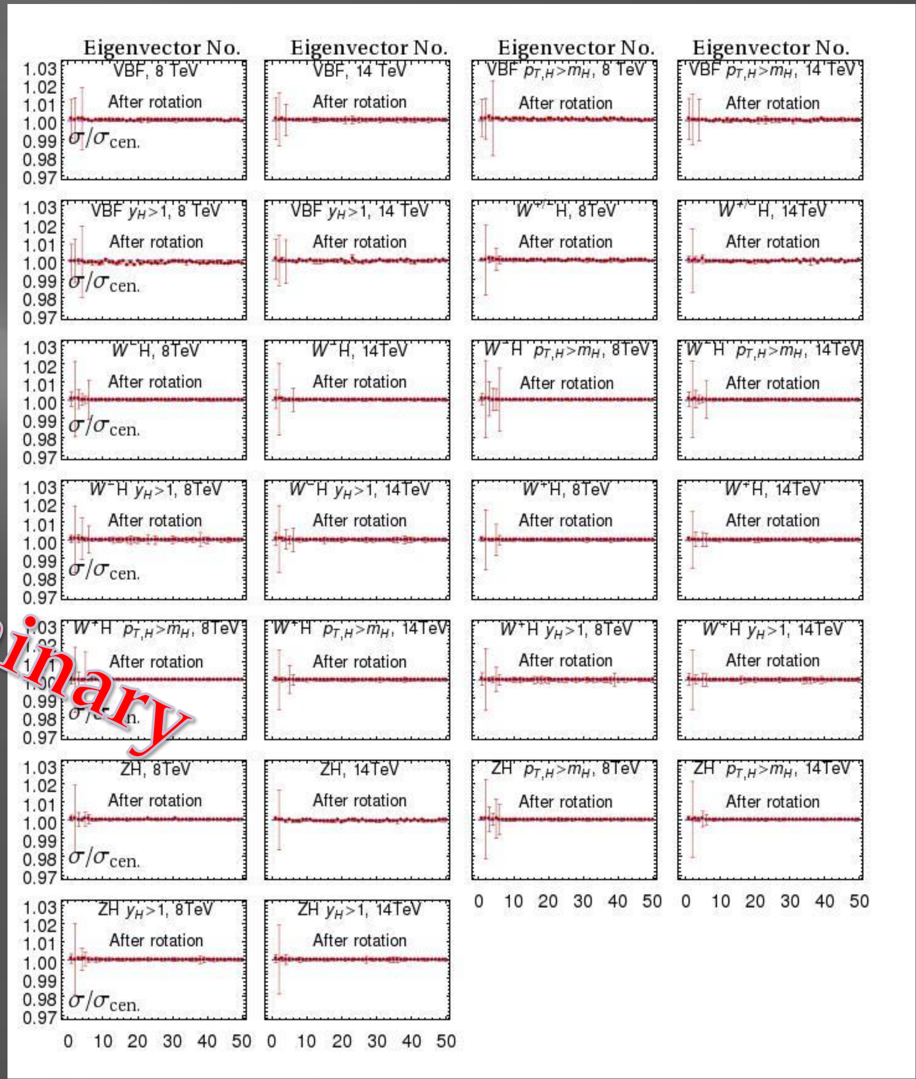


Examples: LHCH PDF set for SM Higgs boson study

We select a group of 46 physical observables, including SM Higgs production cross sections at LHC 8 and 14 TeV, in gg->h, bb->h, VBF, Wh, Zh, hh, tth, h+jet channels, for inclusive rate and rate in high pt or large rapidity region of the Higgs boson.

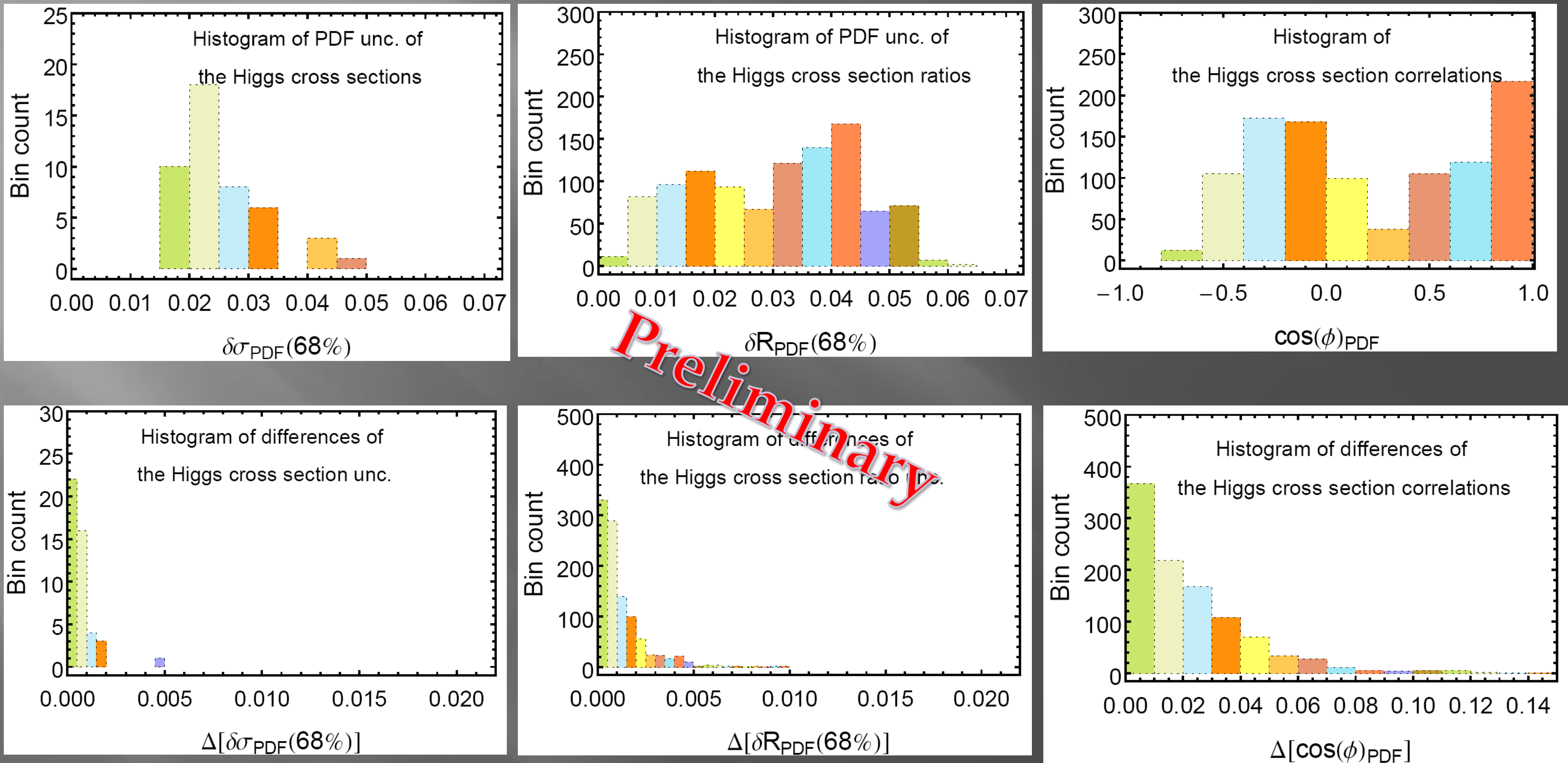
Predictions of the LHCH set, fully described by 6 eigenvectors





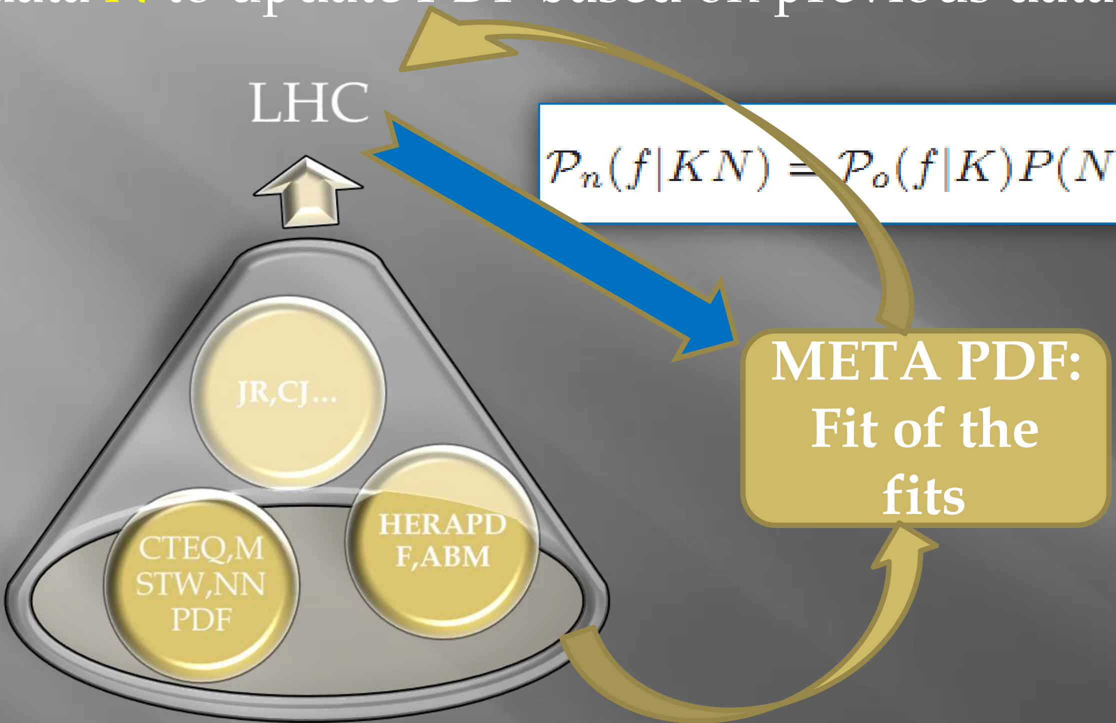
Examples: LHCH PDF set for SM Higgs boson study

Distribution of the PDF uncertainties of the 46 Higgs observables from the regular set, and comparisons of the first 6 eigenvectors of LHCH with the regular set



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., 1310.1089, N. Sato et al.)**, e.g., resulting different weight choices as $\exp[-\chi^2(N|f)/2]$, $(\chi^2)^{n/2-1}\exp[-\chi^2(N|f)/2]$, $(\chi^2)^{(n-1)/2}\exp[-\chi^2(N|f)/2]$, ….

The reweighting can be performed using MC sampling or even analytically assuming a quadratic shape of χ^2 on PDF parameters.

Further development: reweighting schemes

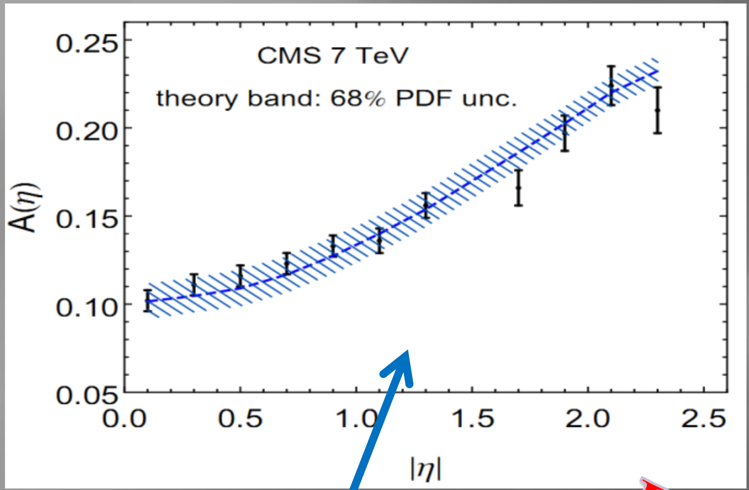
We follow a principle that the new data set N should be treated equally as the original data set K. By another meaning, the reweighting should give results close to the refitting with new data N included.

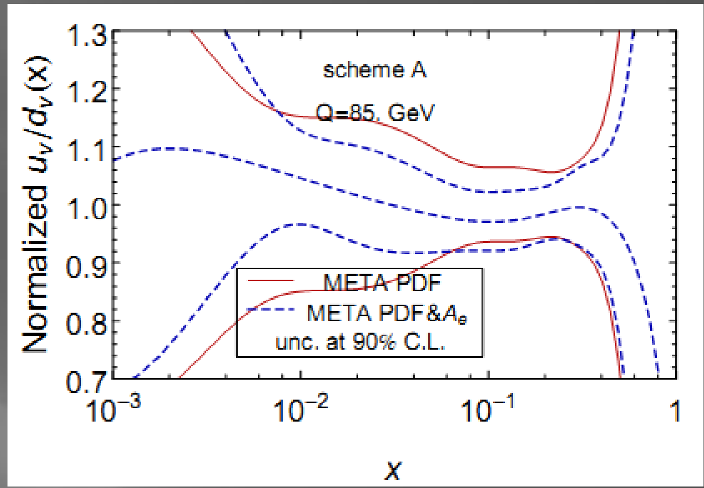
Scheme A: assuming a quadratic dependence of $\chi^2(N|f)$ on PDF parameters x_i, it is straightforward to prove that for the HERA-like fit ($\Delta\chi^2=1$), HERAPDF or ABM, the PDF reweighting with weight $\sim\exp[-\chi^2(N|f)/2]$ is exactly equivalent to the corresponding refitting. Gaussian→ Gaussian.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

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.

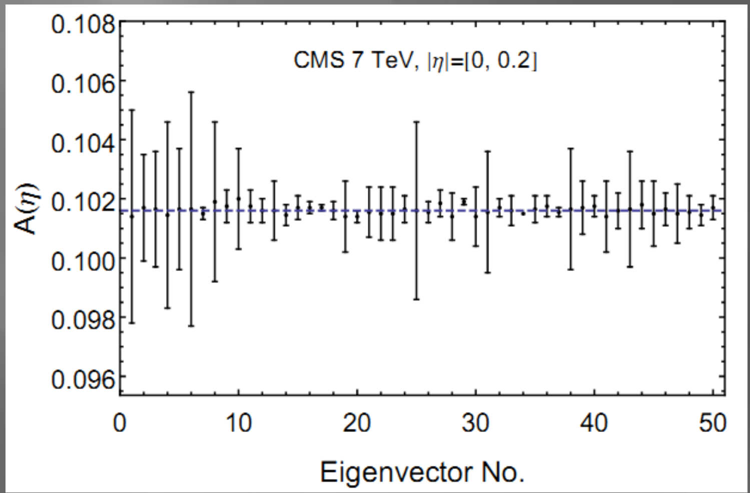


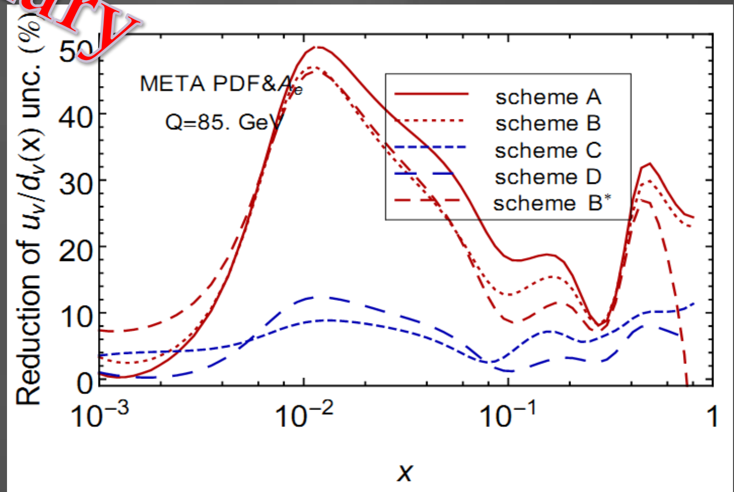


Comparison of META predictions with data before reweighting

Preliminary

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 process, and 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 in that case minimizing numerical computation efforts for massive NNLO calculations.

It can be applied with Bayesian reweighting to include constraints from new experiments or data set diagonalization to simplify the PDF analyses for certain observables.

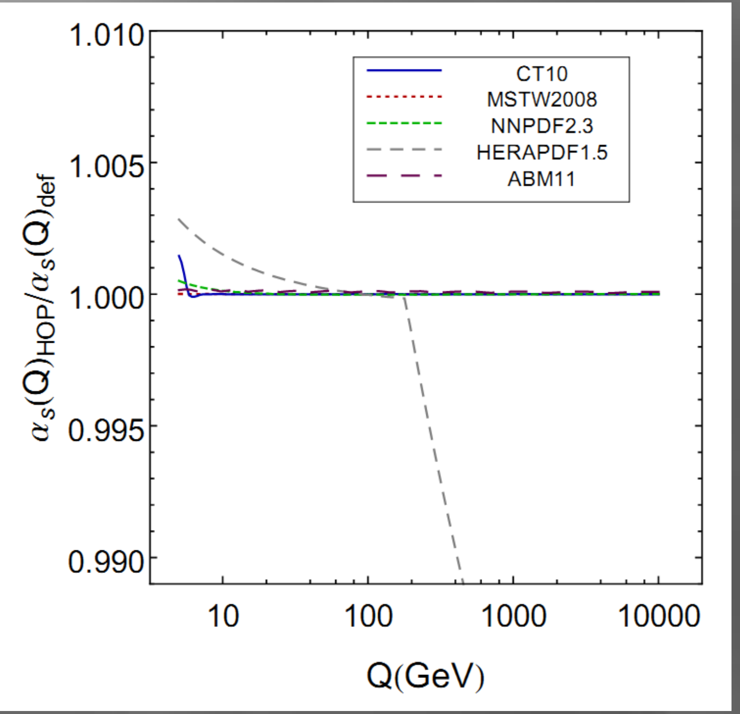
The LHAPDF grids(v5 & v6), table files, and interfaces (Mathematica, Fortran, C++) are publically available at **http://metapdf.hepforge.org**

Validation of the numerical DGLAP evolution

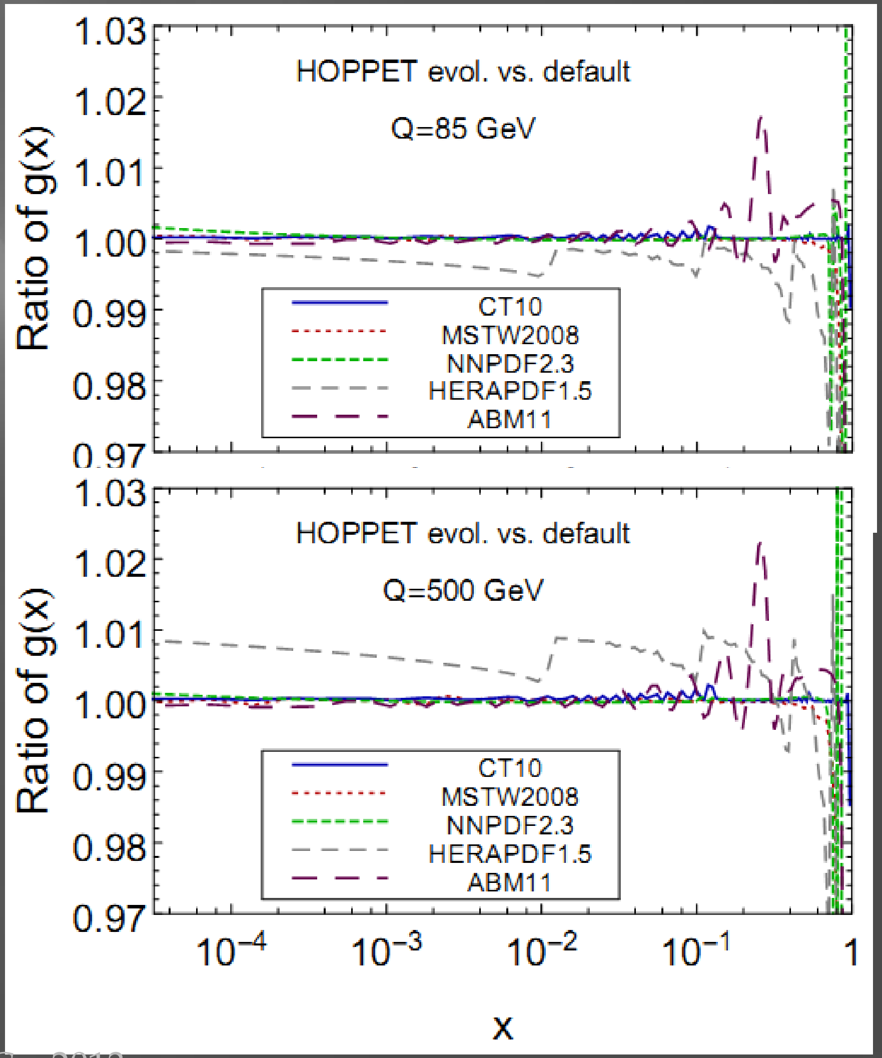
Backups

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



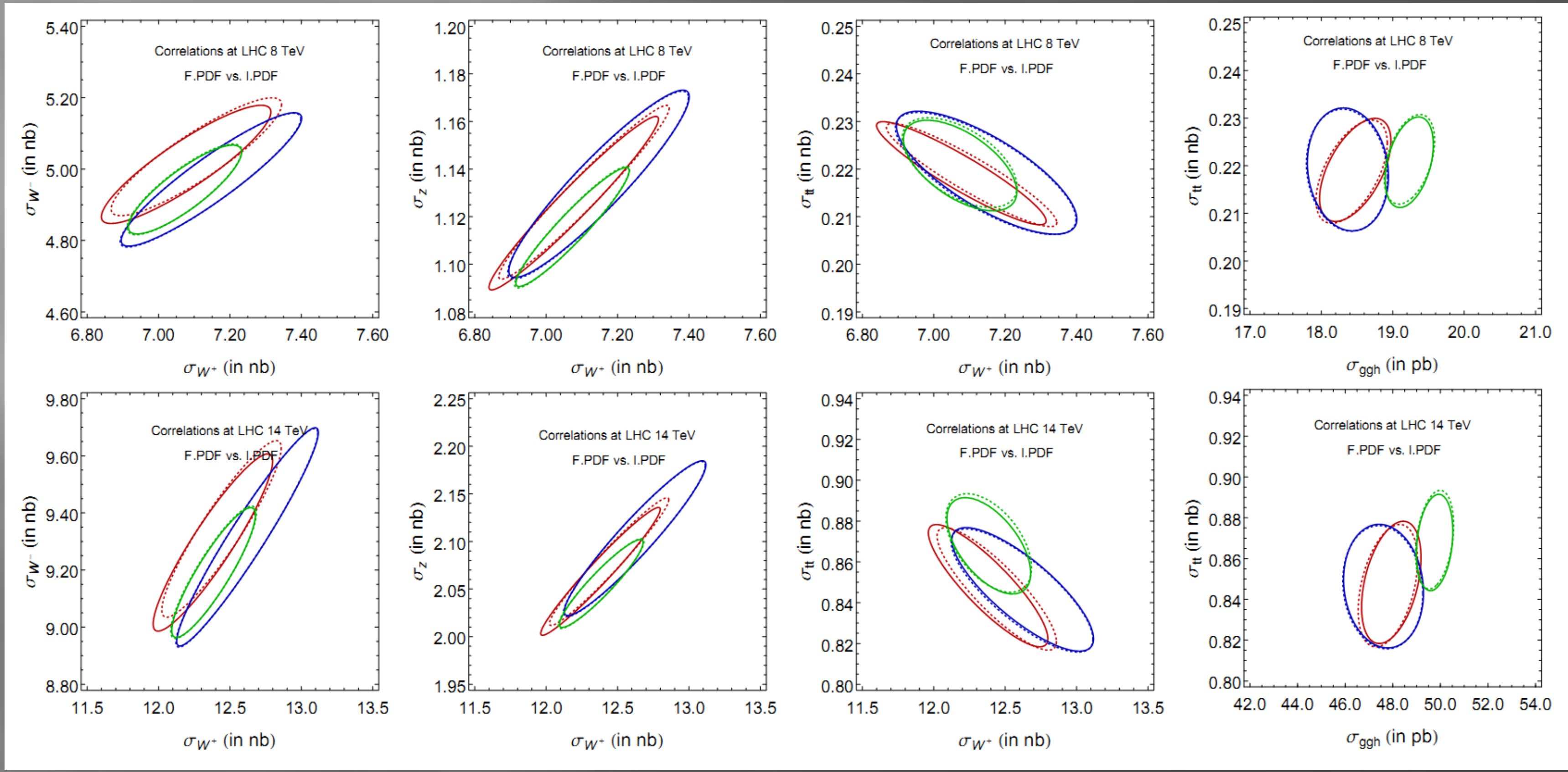
Example: evolution of the gluon PDF



Agreement on predictions for benchmark LHC processes

Backups

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.

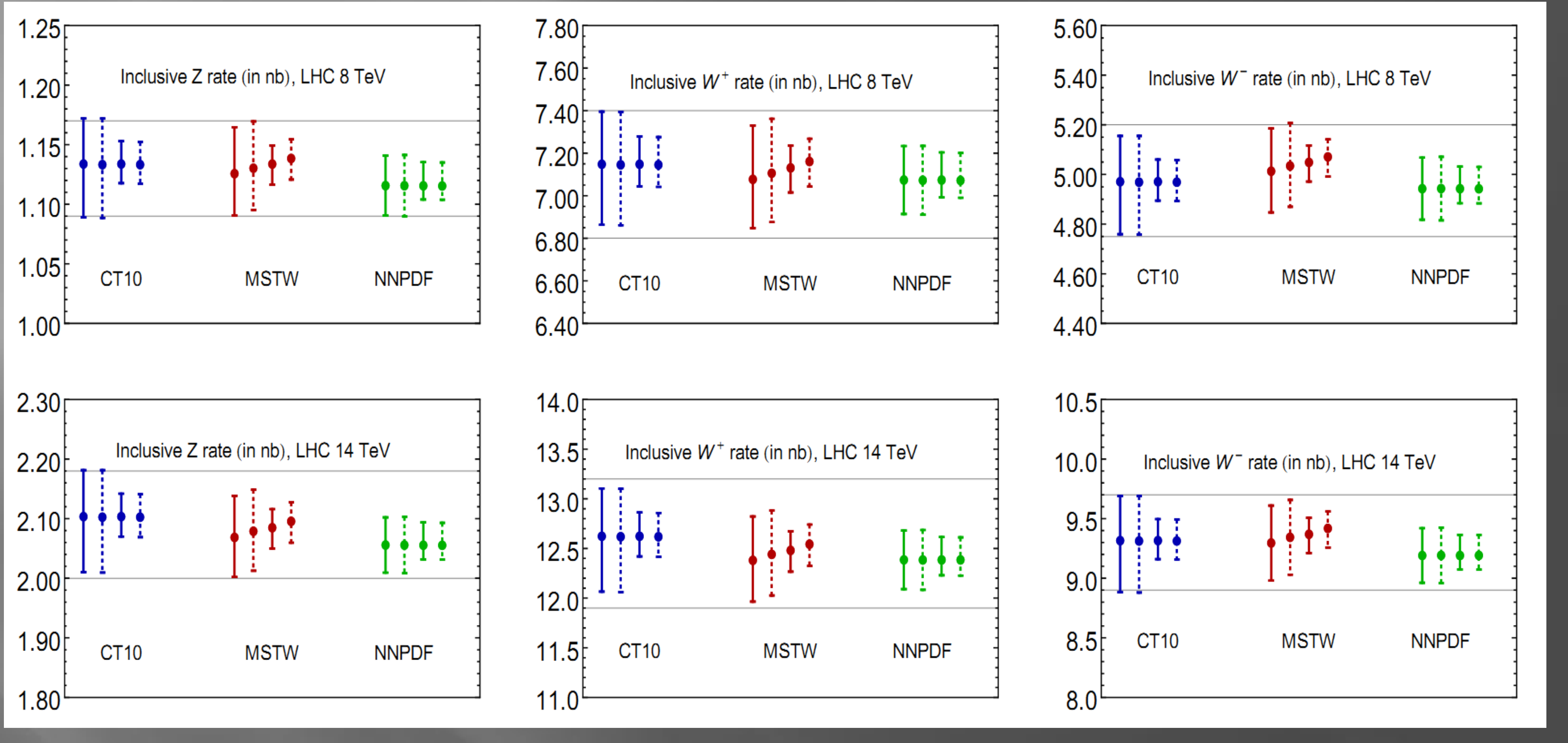


blue(CT10), red(MSTW), green(NNPDF), solid(dotted) for input (fitted) PDFs

Backups

- Agreement on predictions for benchmark LHC processes

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.

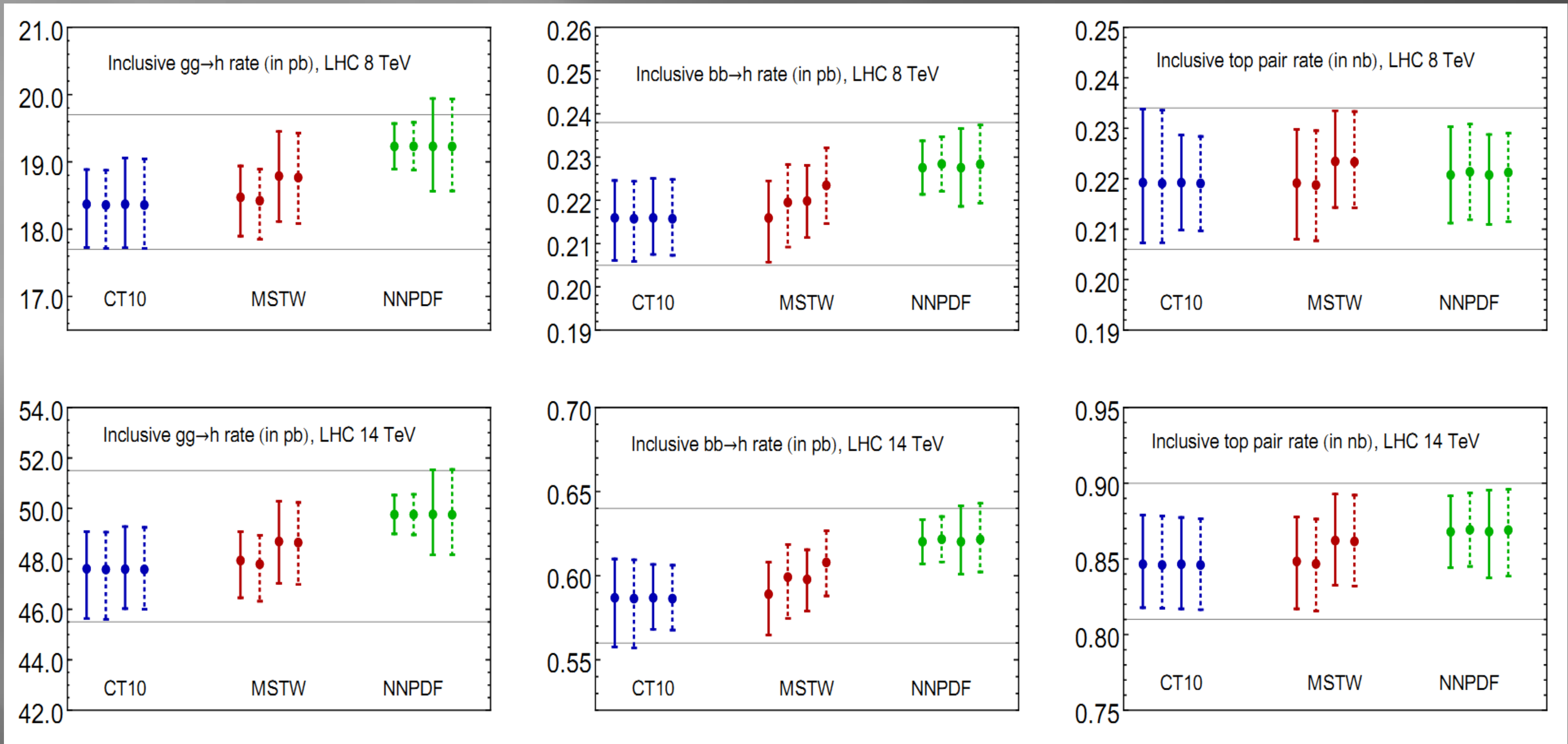


blue(CT10), red(MSTW), green(NNPDF), solid(dotted) for input (fitted) PDFs

Agreement on predictions for benchmark LHC processes

Backups

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.

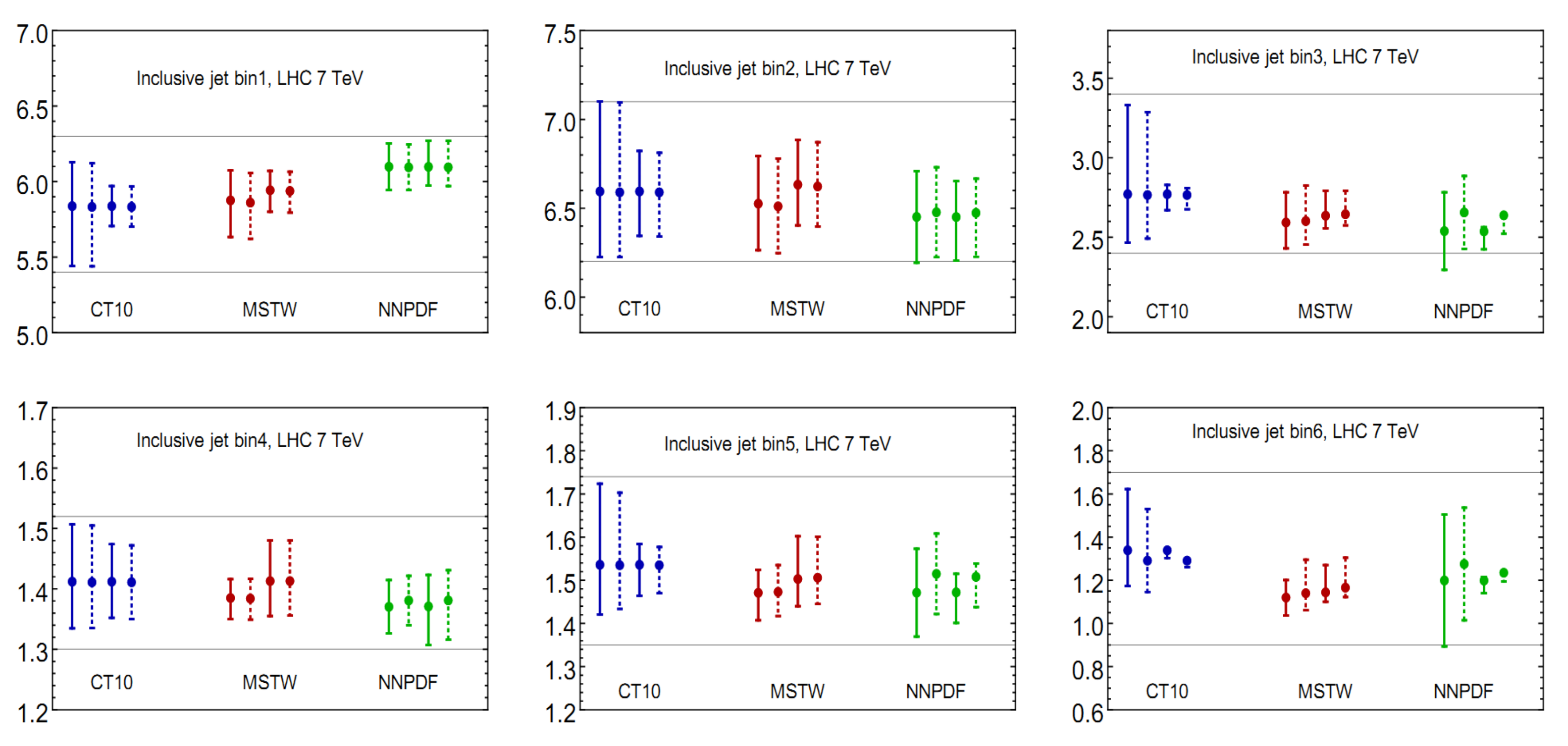


blue(CT10), red(MSTW), green(NNPDF), solid(dotted) for input (fitted) PDFs

- Agreement on predictions for benchmark LHC processes

Backups

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.

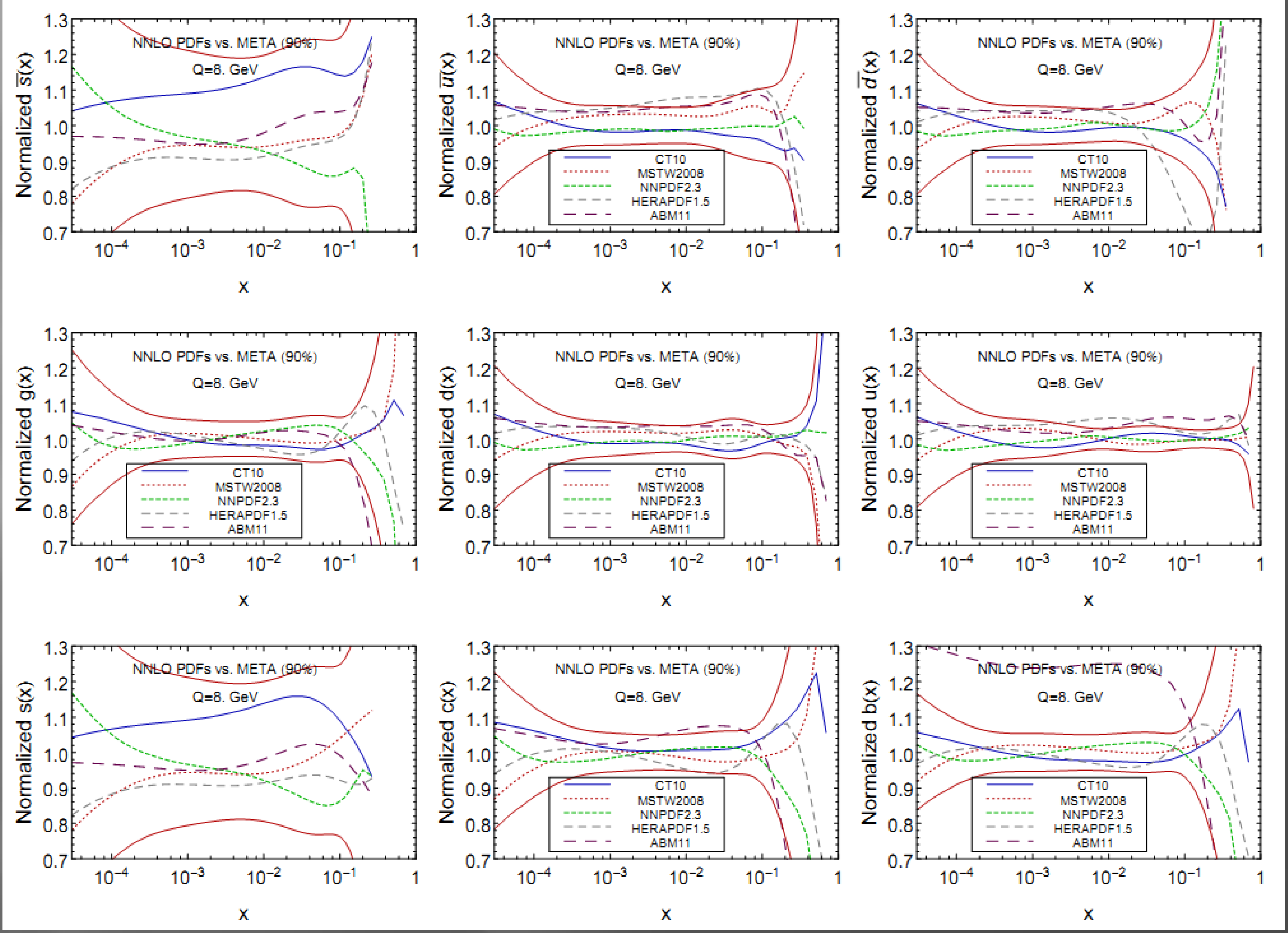


blue(CT10), red(MSTW), green(NNPDF), solid(dotted) for input (fitted) PDFs

The META PDFs

Backups

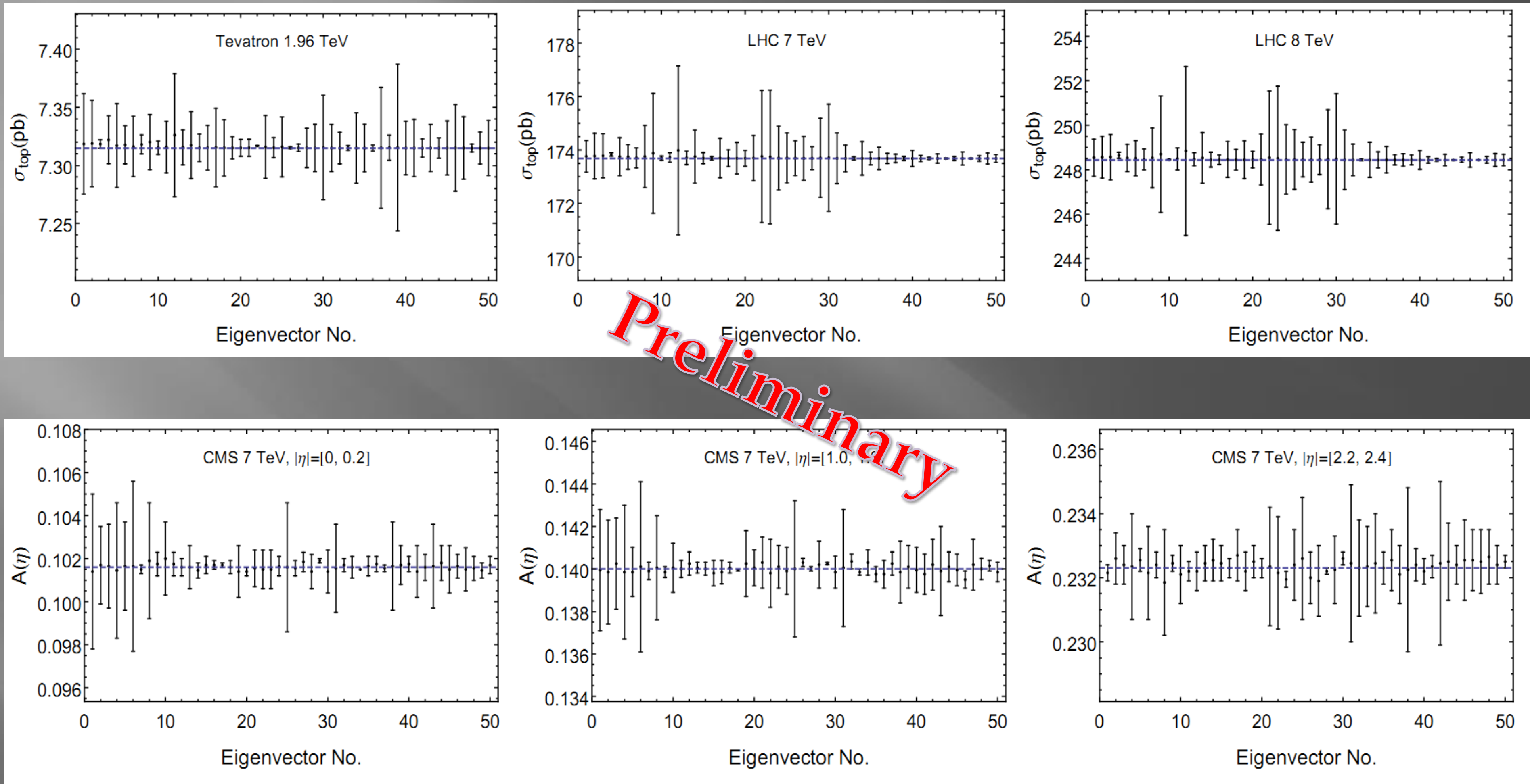
Comparisons of the META PDF with the original central PDF from each groups with $\alpha_S(M_Z)$=0.118 at Q=8 GeV.



- Examine the LHC predictions from different eigenvectors

Backups

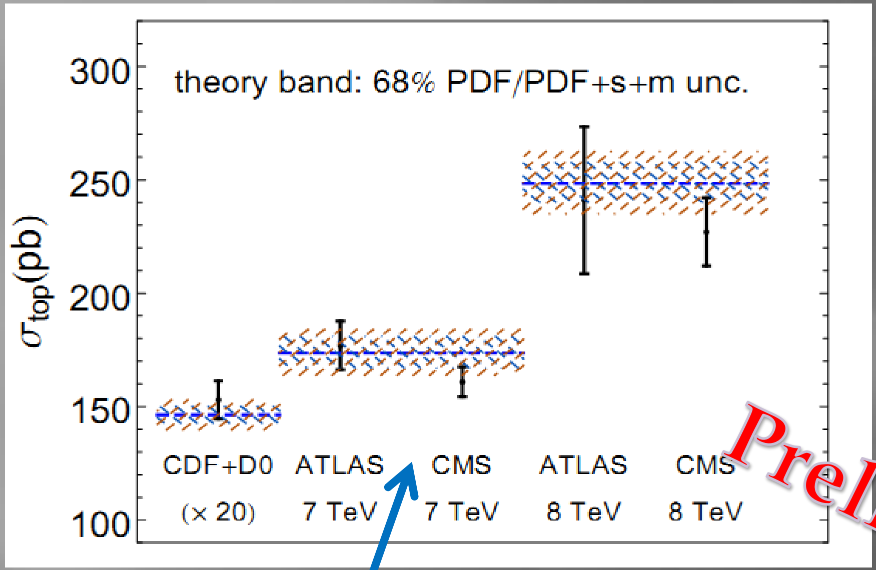
It shows the dynamic structure of different eigenvectors and validates the linear approximations.

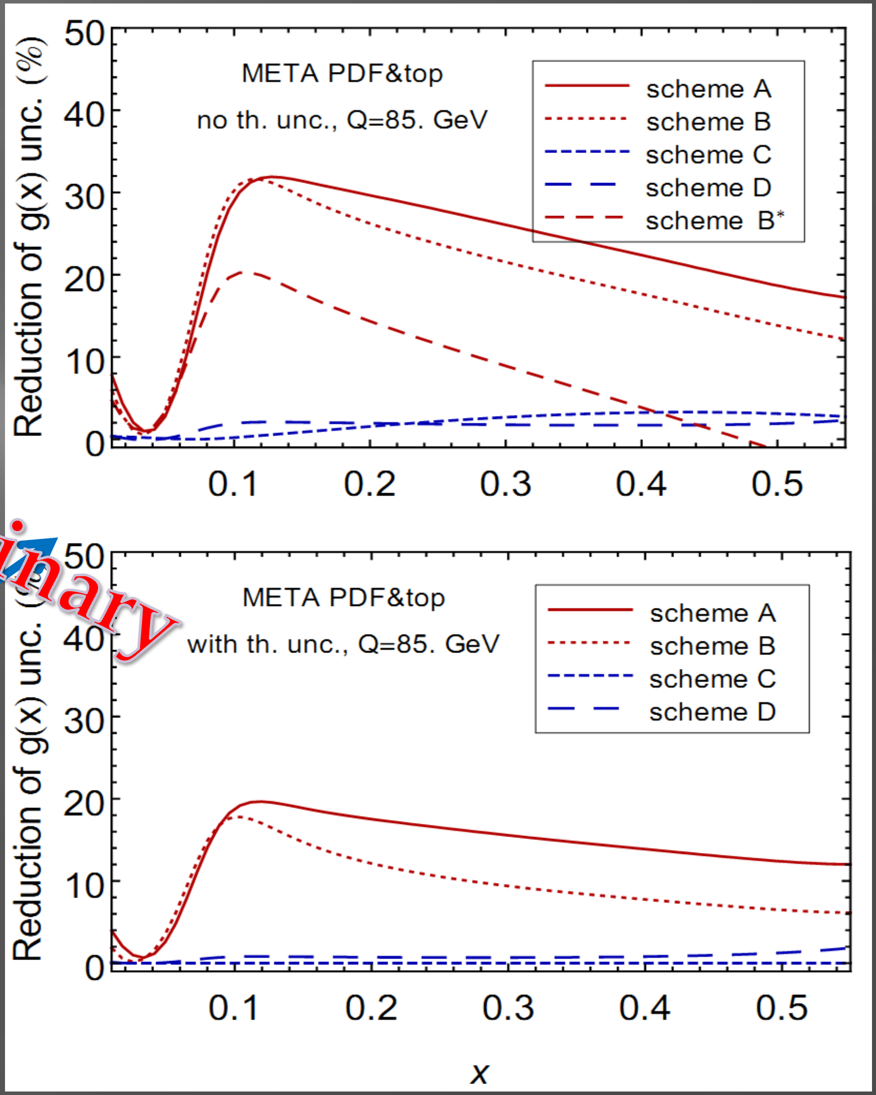


Examples: top quark data

Backups

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.





Preliminary

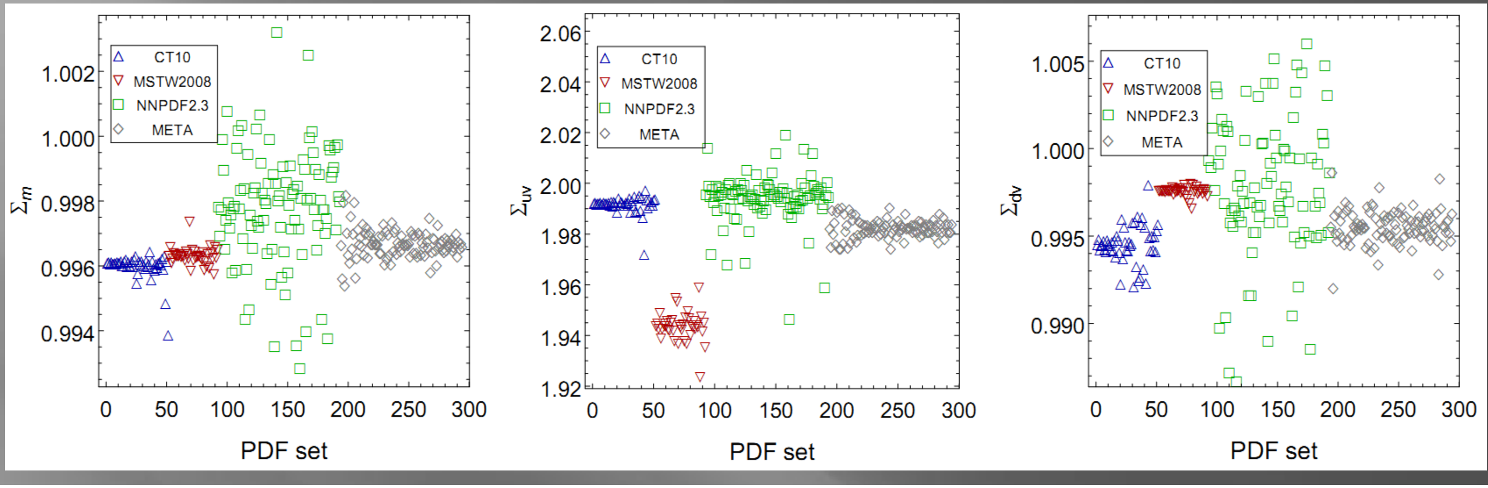
Comparison of META predictions with data before reweighting

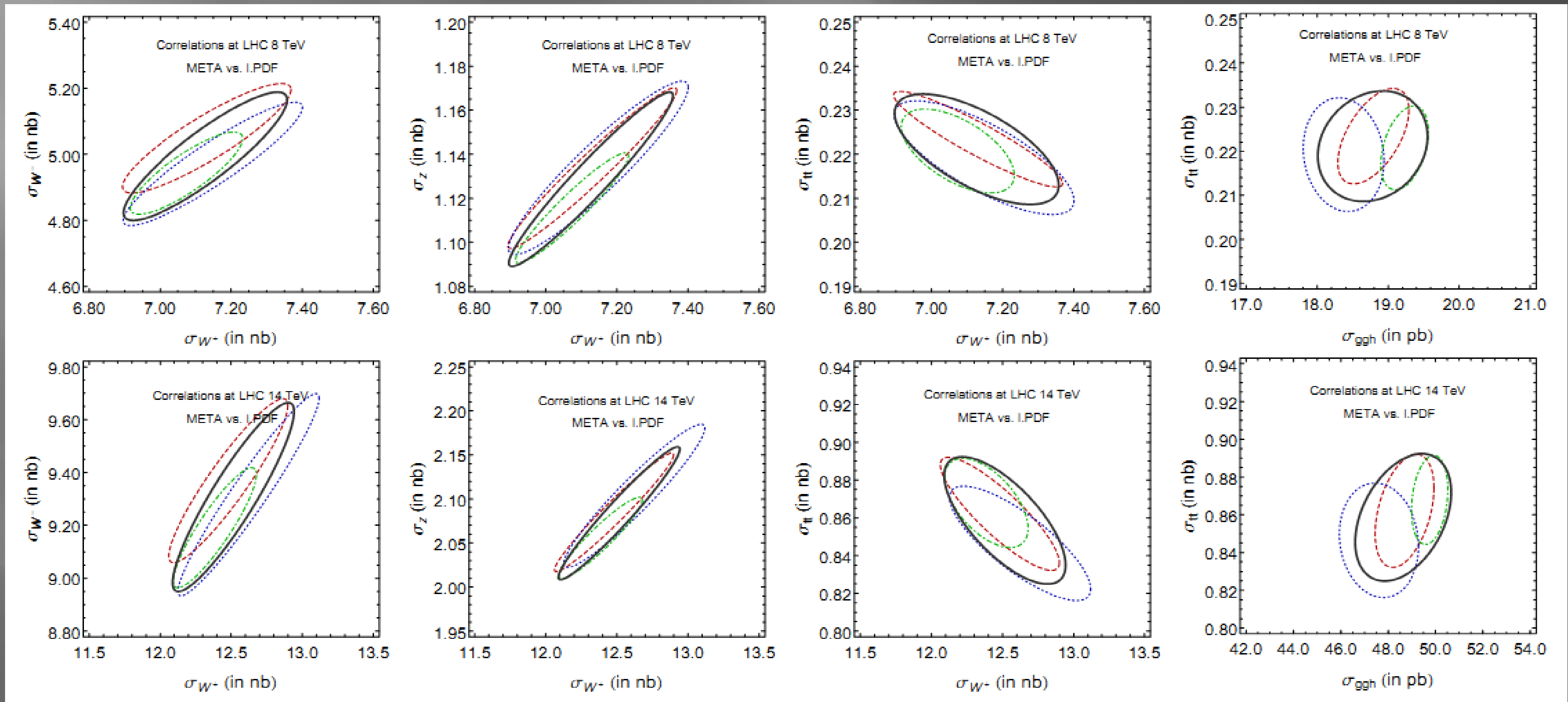
Reduction of the gluon PDF uncertainties under different schemes with and without including theoretical uncertainties.

Examine the sum rule and PDF correlations of META PDFs

Backups

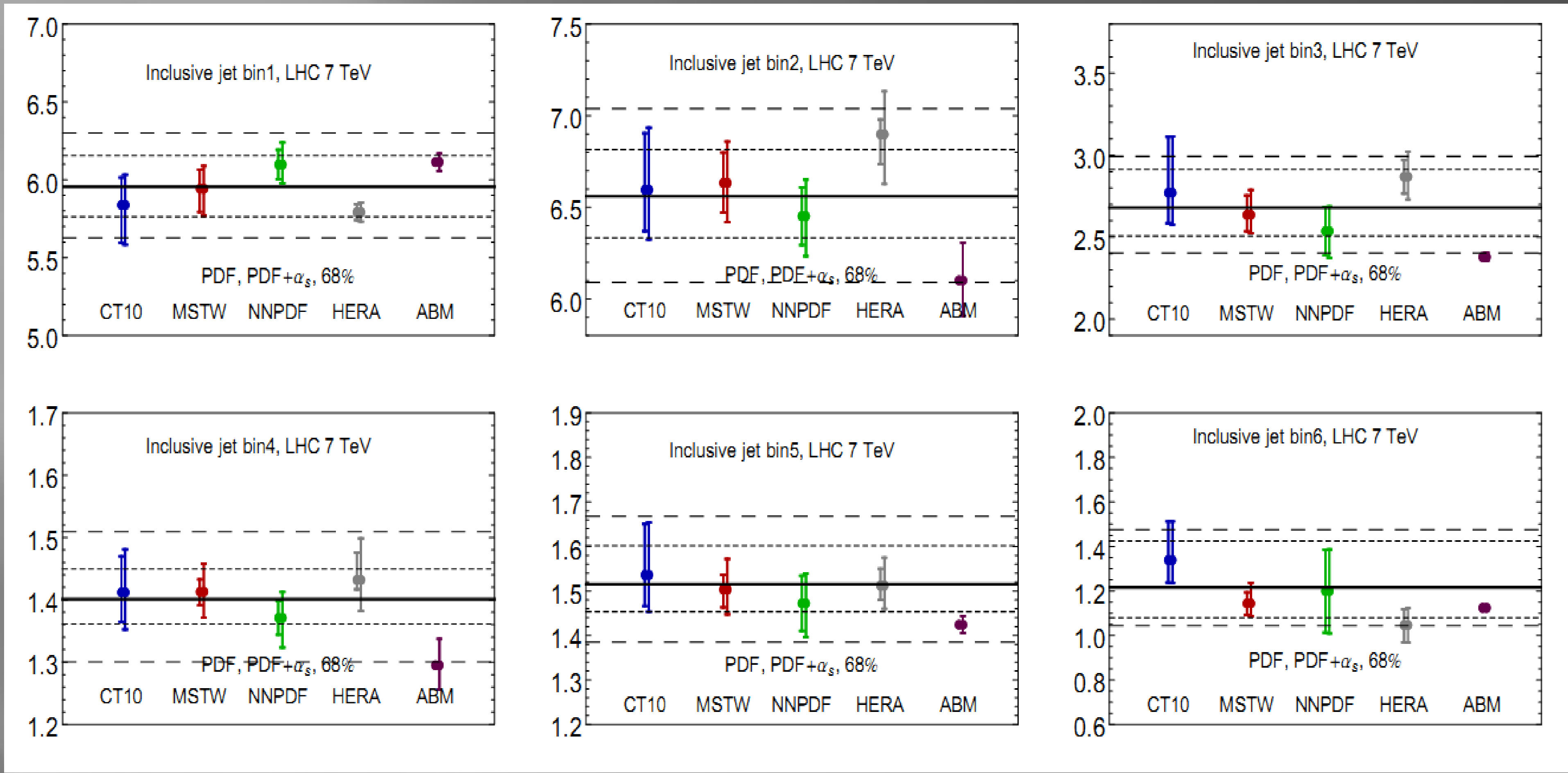
META PDFs vs. input PDFs, sum rules and correlations





Examine of jet cross sections from META PDFs

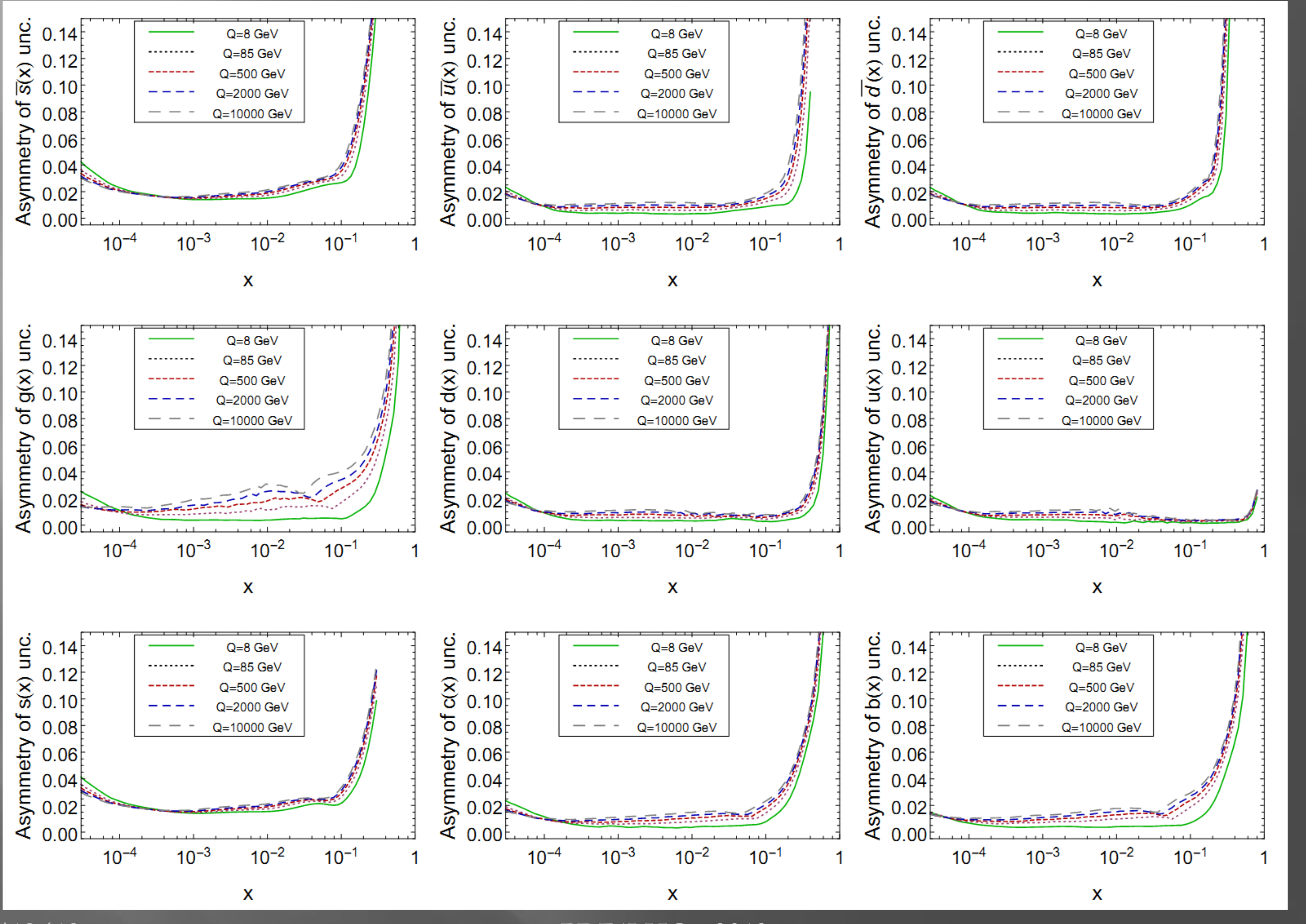
Backups



Examine of the linear dependence of META PDFs

Backups

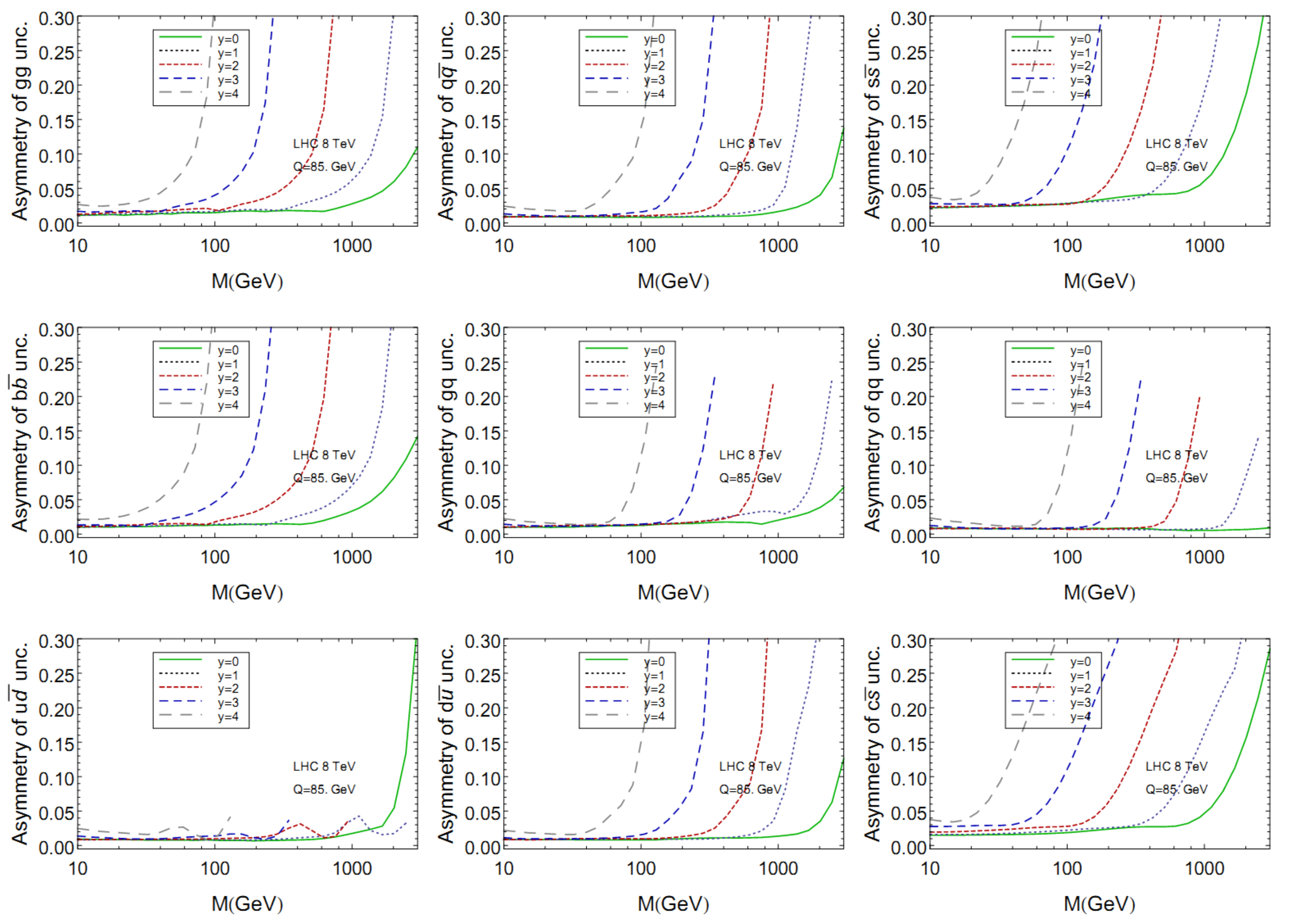
Asymmetry of PDF uncertainties of PDFs with different flavors



Examine of the linear dependence of META PDFs

Backups

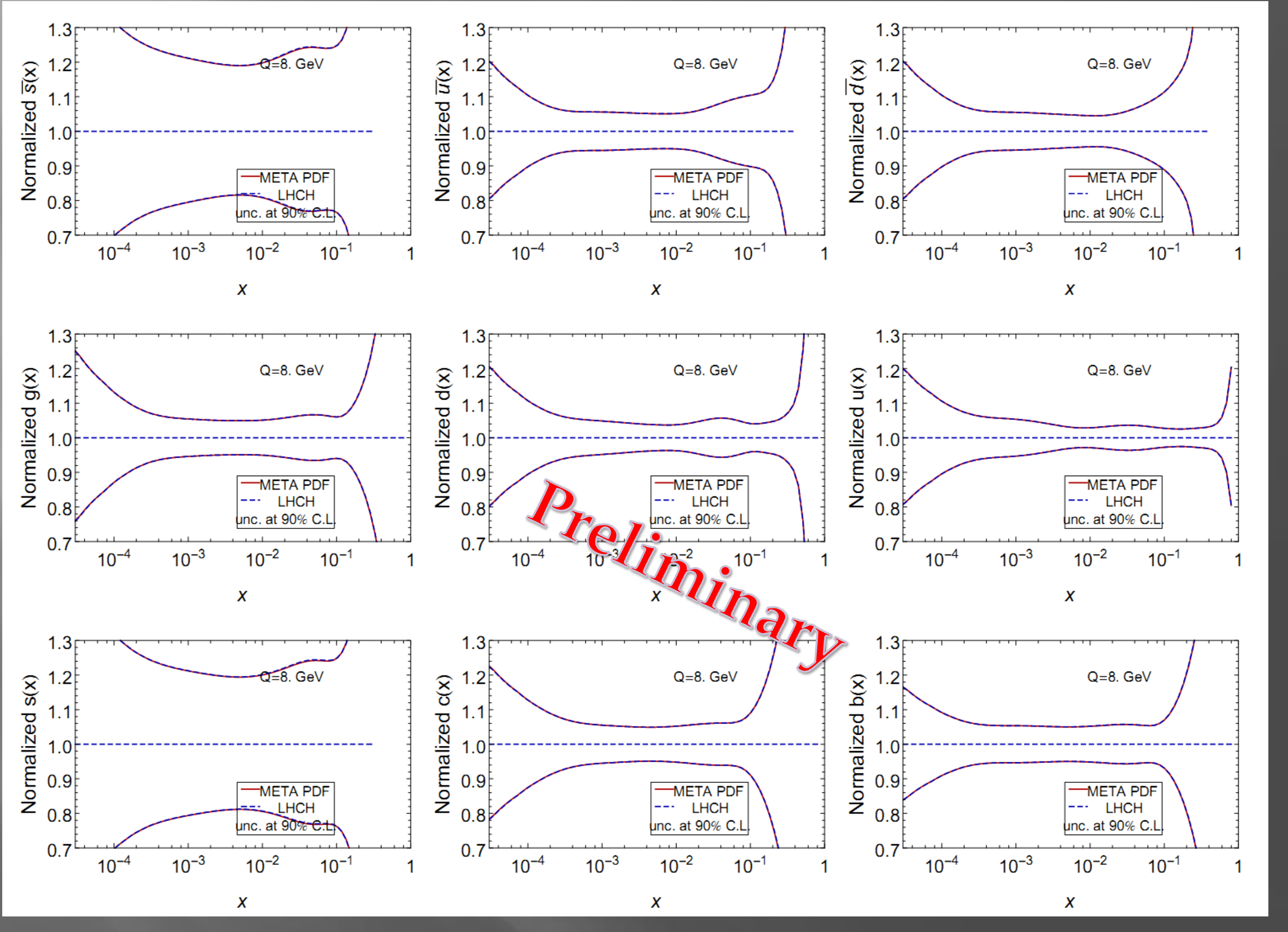
Asymmetry of PDF uncertainties of parton luminosities



The LHCH set

Backups

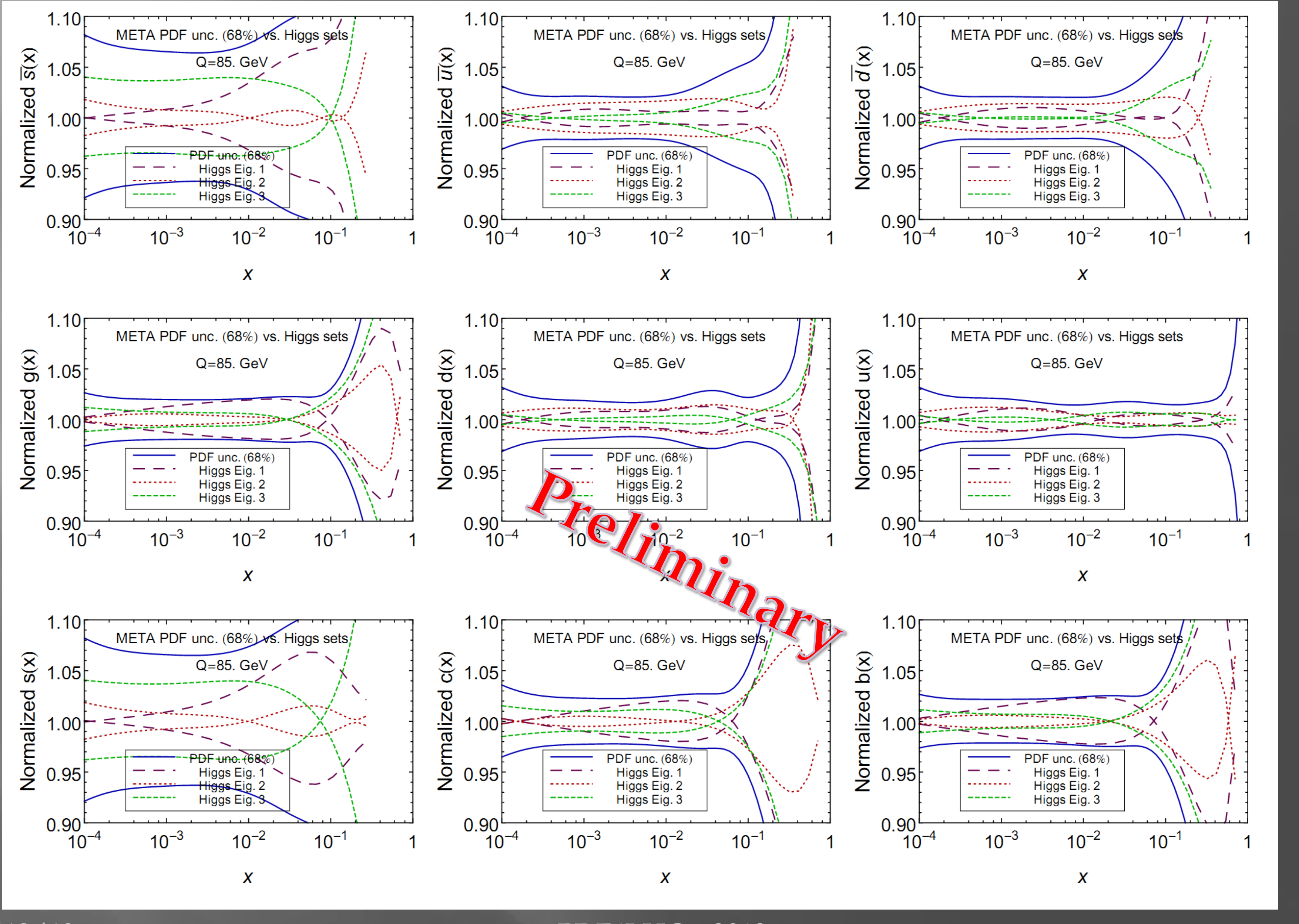
Equivalence of the PDFs before and after rotation



- The LHCH set

PDFs from the first 6 eigenvectors

Backups



The LHCH set

Backups

PDFs from the first 6 eigenvectors