Parton distribution with scale uncertainties: A Monte Carlo sampling approach

DIS, Santiago de Compostela

Zahari Kassabov,

in collaboration with with Cameron Voisey and Maria Ubiali

May 3, 2022

DAMTP, University of Cambridge

A new perspective

What problem are we solving

PDFs with “theory uncertainties” are often requested.

Scale variations $\subset$ Missing higher orders (in QCD) $\subset$ Theory uncertainties

Methods

- Assume first perturbative coefficients determine later ones (Cacciari-Houdeau, Passarino-David, Bonvini).
- Scale variations
  - Factor of two (?)
  - Relation to missing higher orders

Scale uncertainties are uncertainties in their own right

MHOU $\longrightarrow$ Scale uncertainties

Scales in PDF fits

- PDFs produce theory predictions given other theory predictions and experimental data in the PDF fit.
- Each theory prediction in the fit requires a factorization and a renormalization scale.
- When making predictions using PDFs we also need to set scales.
- We get PDFs based on fits to the data.
    - PDFs with bad fit quality do not make sense!

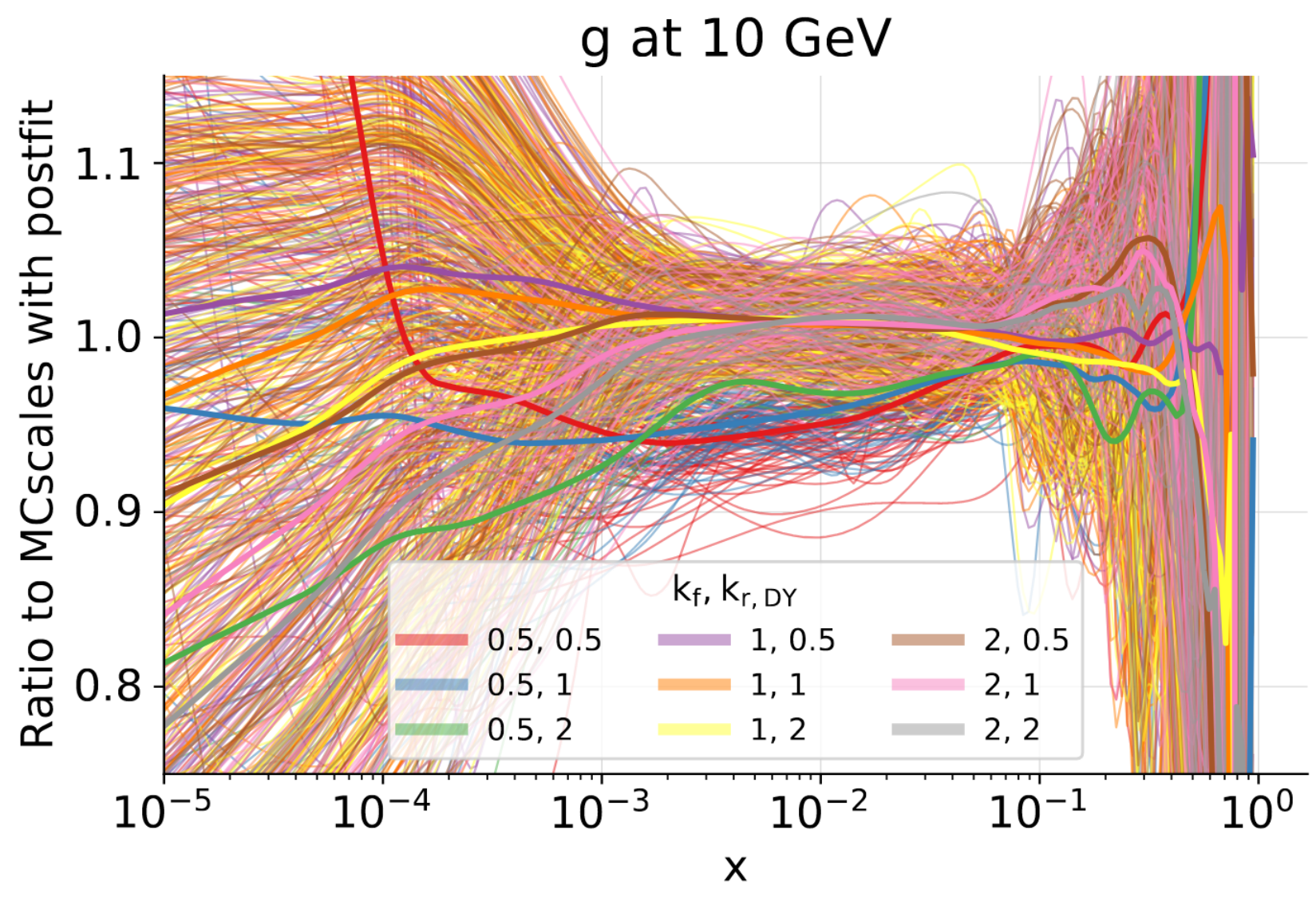
Questions

- How many scales (i.e. possible variables)?
- What is their distribution?
  - What is the range of variations?
  - How are they correlated?
- Implementation in a PDF fit?
- Interpretation within a PDF fit?
- Predictions with matched scales between PDFs and hard cross sections?

MCscales

- NNPDF:
  - Produce samples from the experimental data distribution and fit a PDF replica for each.
- MCscales:
  - Produce samples from the experimental data and theory distribution and fit a PDF replica for each.
- Borrow some assumptions from our earlier attempt and including scale variations in PDF fits [NNPDF 2019, arxiv:1906.10698].

What we'll get



`MCscales` replicas for the Drell Yan Process

Initial distribution of scales

- Assume scale variations by a factor 2 around a central scale.
- Factorization scale same for all processes
- Renormalization scale correlated by process
  - Split data in `mcscales_processes: ["DIS CC", "DIS NC", "DY", "JETS", TOP]`

With that

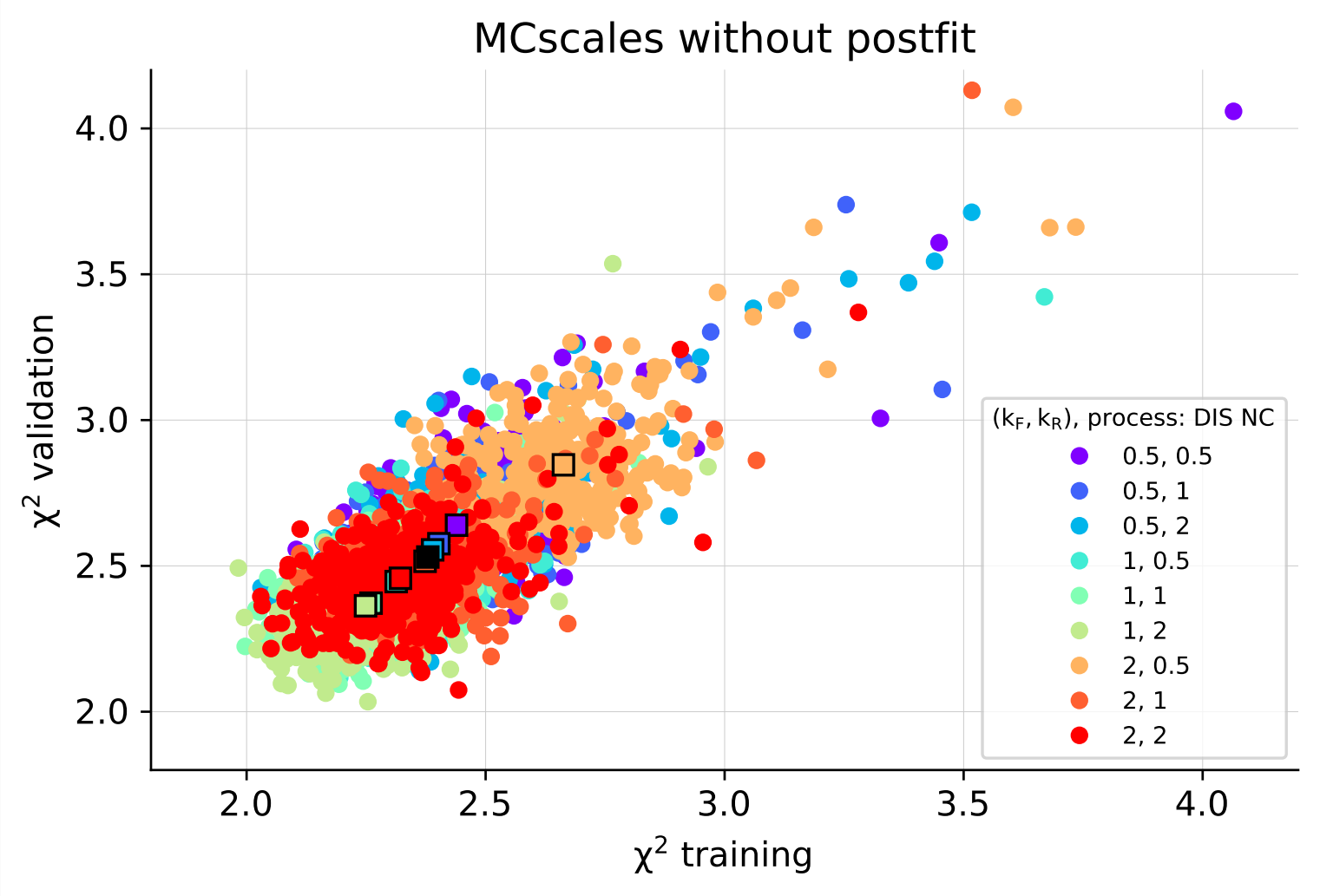
$$\Xi = \left\{\frac{1}{2}, 1, 2\right\}$$

$$\Omega = \left\{(\xi_f, \xi_1 \cdots \xi_{N_p}) \; \forall \; \xi_f, \xi_1, \cdots, \xi_{N_p} \in \Xi\right\},$$

and for each replica we sample $\omega \in \Omega$.

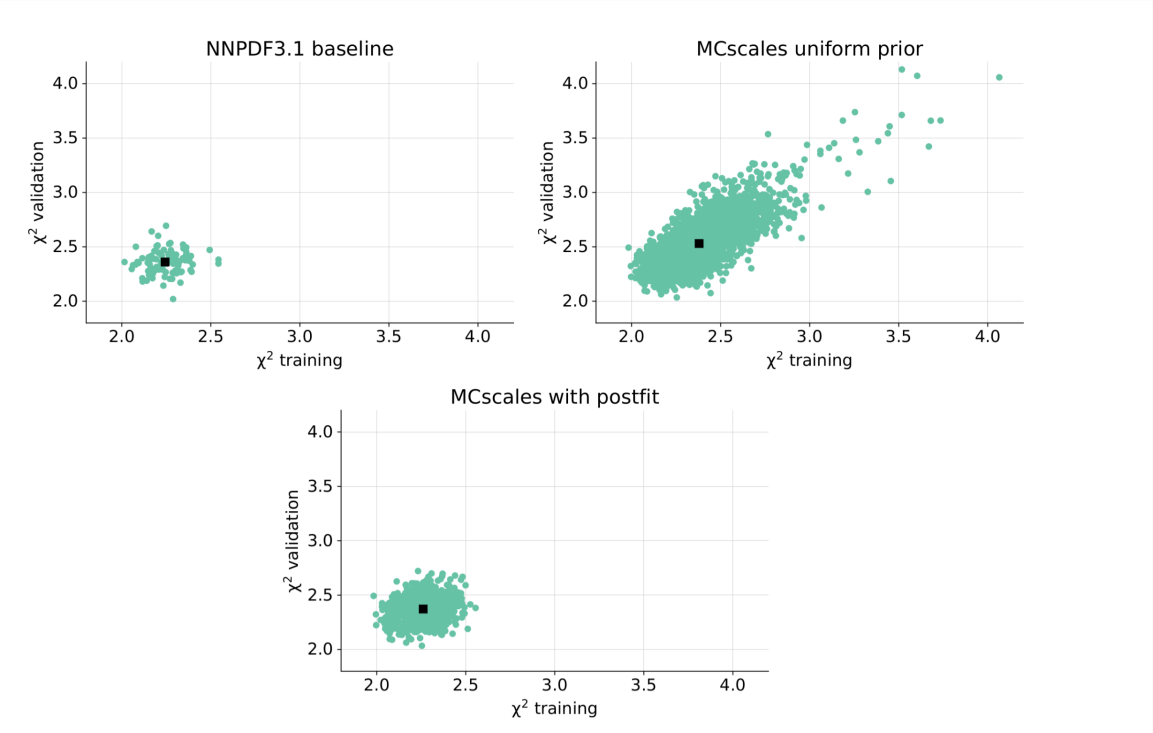
- Assume uniform sampling in this space.

Fit quality when we fit



Different scale choices result in very different fit quality

Postfit selection



- Use the same selection criterion as the normal NNPDF fit, assuming central scales only

$$\chi^2_n > \langle \chi^2 \rangle_{\omega=\{1\ldots1\}} + 4\,\mathrm{std}(\chi^2)_{\omega=\{1\ldots1\}}$$

- This is a feature not a bug!

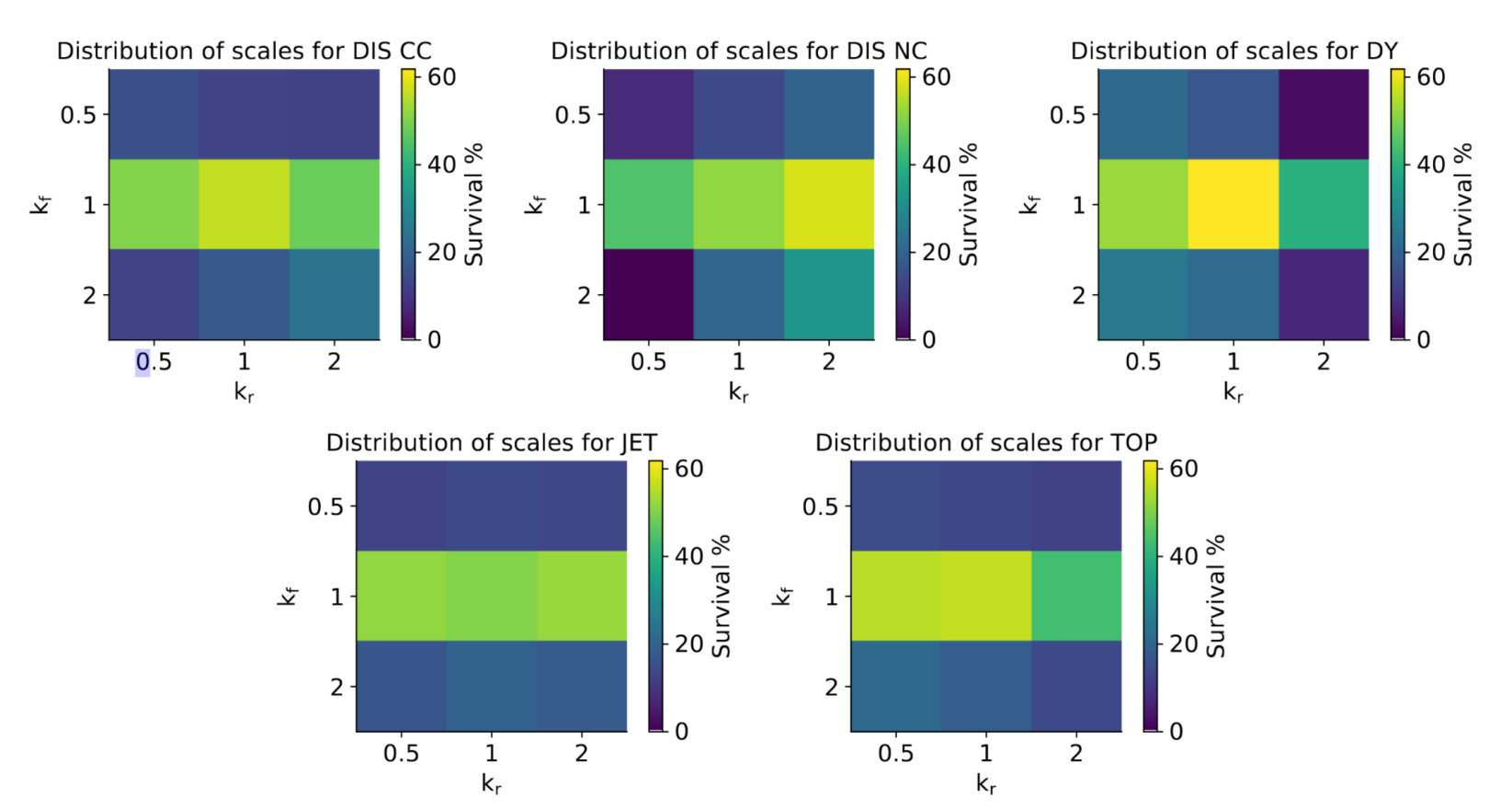
Postfit replicas interpretation

- So far we are talking about a *prior* scale distribution
- Surviving replicas define a *postrior* (on scales+PDF) with the likelihood

$$P(D|k,\omega) \propto \begin{cases} 1 & \text{Postfit passes} \\ 0 & \text{Otherwise} \end{cases}$$

- Can make statements about scale variations consistent with data
  - We are already fitting PDFs assuming theory matched data.

Survival fraction



- Renormalisation variation much more uniform than factorisation variation
- Big differences between processes

Matched scales convolution

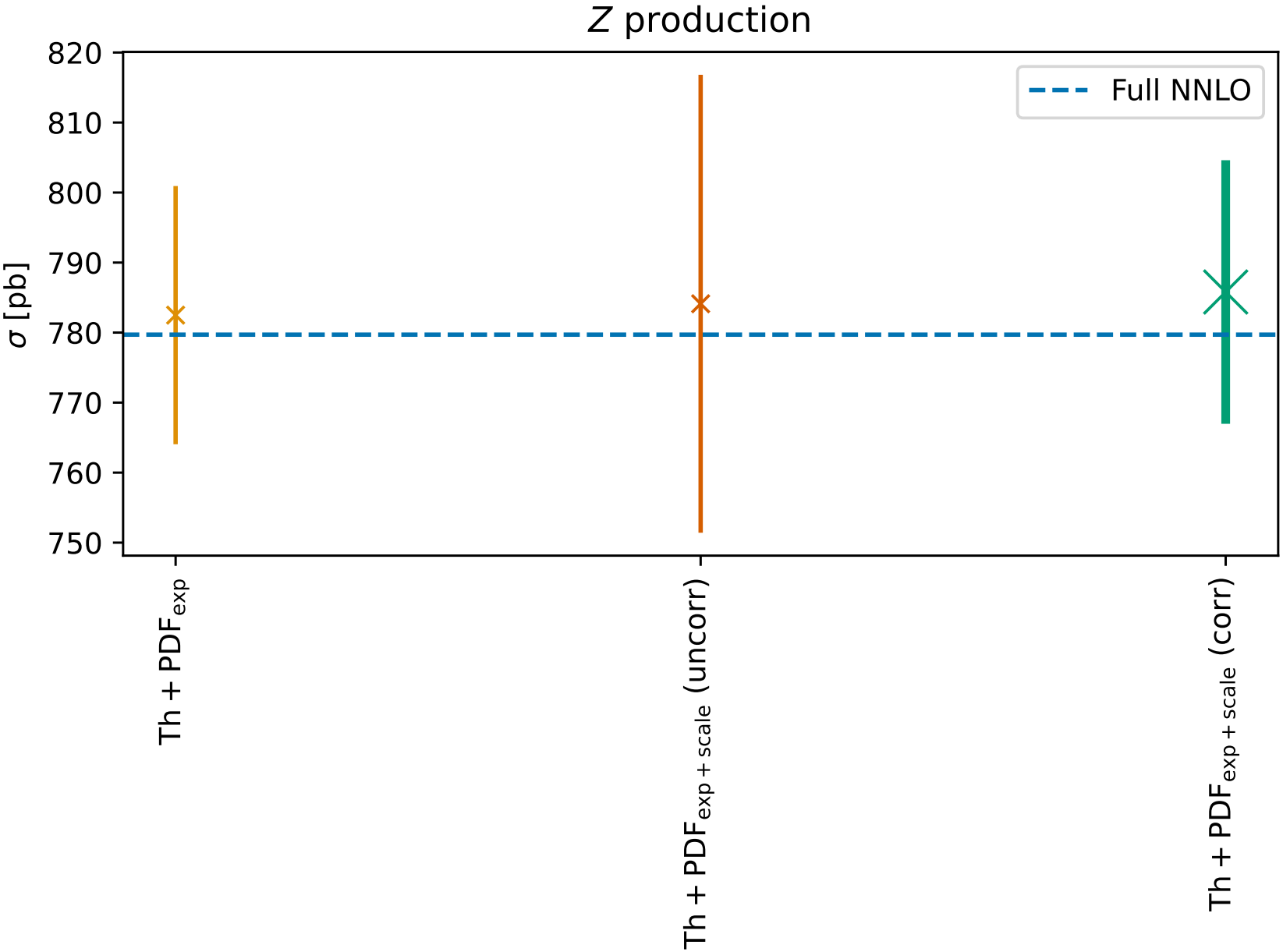
We record the scale choices for each fitted replicas. This allows matching the partonic cross section with the scale choices within each replica

- Monte Carlo sample of N_{rep} `MCscales` prediction including correlated PDF and scale uncertainty

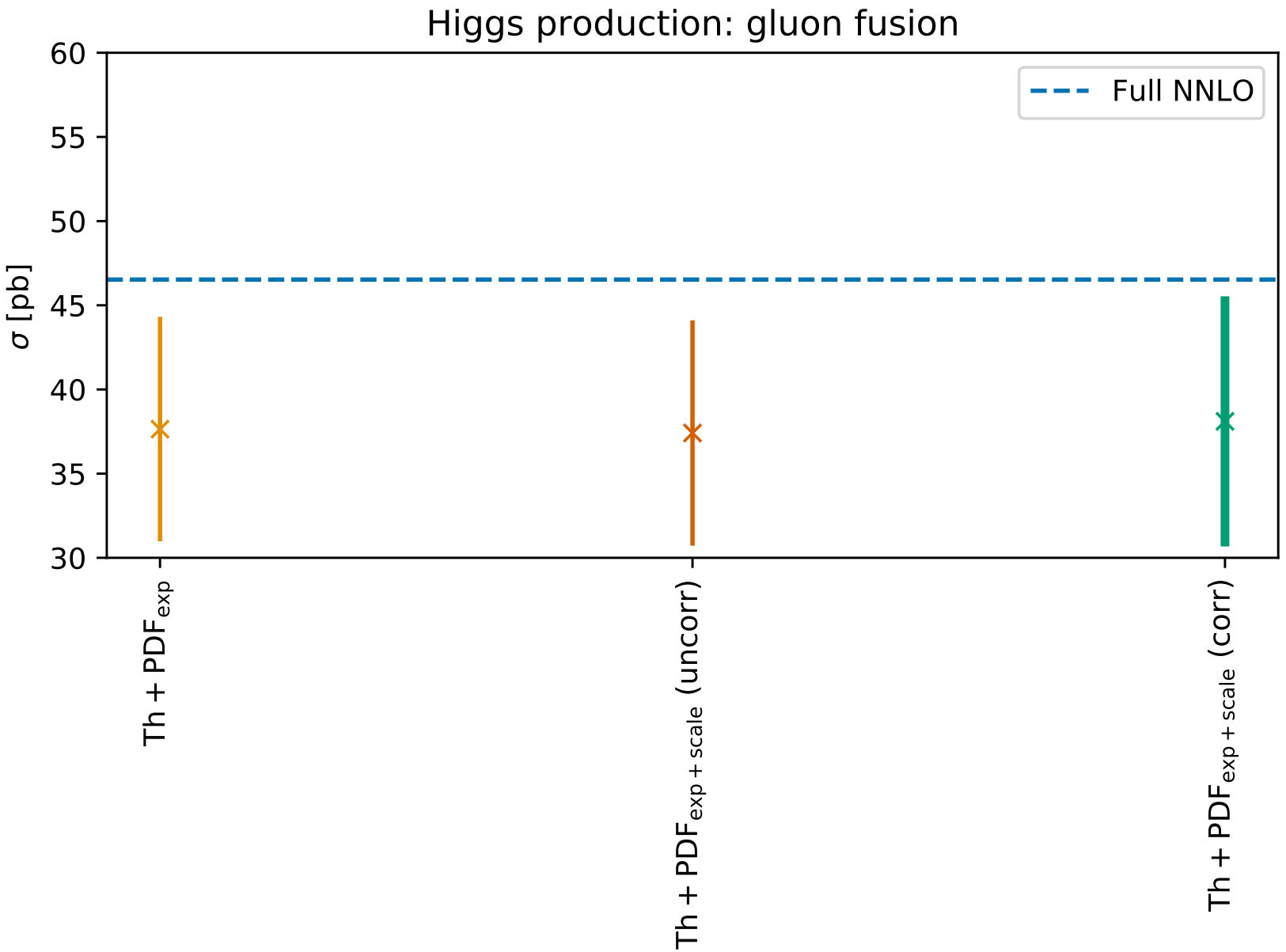
$$\left\{\sigma_k = \hat{\sigma}_p(k_f^{(k)}, k_{r_p}^{(k)}) \otimes f_k(k_f^{(k)}, k_{r_p}^{(k)}) \;\; \forall\, k \in 1 \dots N_{\text{rep}}\right\}$$

- Could be integrated on a Monte Carlo but no need to (tools provided)
- Easy to adapt for processes not in the PDF fit (e.g. Higgs)

Matched predictions: Z production



Matched predictions: Higgs production



Achievements

- LHAPDF set with replica scale variations based on NNPDF 3.1 NLO
  - https://data.nnpdf.science/pdfs/mcscales_v1.tar.gz
- Tools to operate on it
  - https://github.com/Zaharid/mcscales_tools
- First comprehensive benchmark of scale variation fit quality
  - Uniform treatment per process not optimal
  - Factor of 2 for ren variations likely small
  - Factor of 2 for fac ok after postfit selection
- First method to match scale variations in PDFs and hard cross section correctly.