# Summary of PDFLattice 2024

JLab
November 18-20, 2024

*A. Courtoy, on behalf of the organizers*

*C. Keppel, A.Kronfeld, H.-W. Lin, E. R. Nocera, F. Olness, J. Qiu*

# Scope of the workshop

The 2024 edition of the workshop will focus on **uncertainty quantification** in **PDF determinations** from global analyses and lattice computations.

6 key talks on PDF and lattice — state-of-the-art and uncertainty quantification

16 focused talks on PDF and lattice — GPDs, TMDs, FFs, methodology, evolution

10 amazing posters

discussions — always too short

# Inverse problem

An inverse problem entails determination of causal factors from the effects or observations they produced.

Antonym: *forward problem*

- well-posed problems — Hadamard conditions:
  1. The problem has a solution
  2. The solution is unique
  3. The solution's behavior changes continuously with the initial conditions

- creasons for complexity of inverse problems
  direct products
  convolutional problems
  Hausdorff moment problem

Inverse problems arise at many steps of our analyses …
baseline of this workshop.

# Regression

Entails the definition of an optimization framework with

- a loss / objective function

  (log-) likelihood and priors or penalties, treatment of systematic uncertainties
- parametric form(s)

  model sampling, first principle constraints
- criteria for goodness-of-fit

  metric, closure tests
- criteria for uncertainty quantification (UQ)

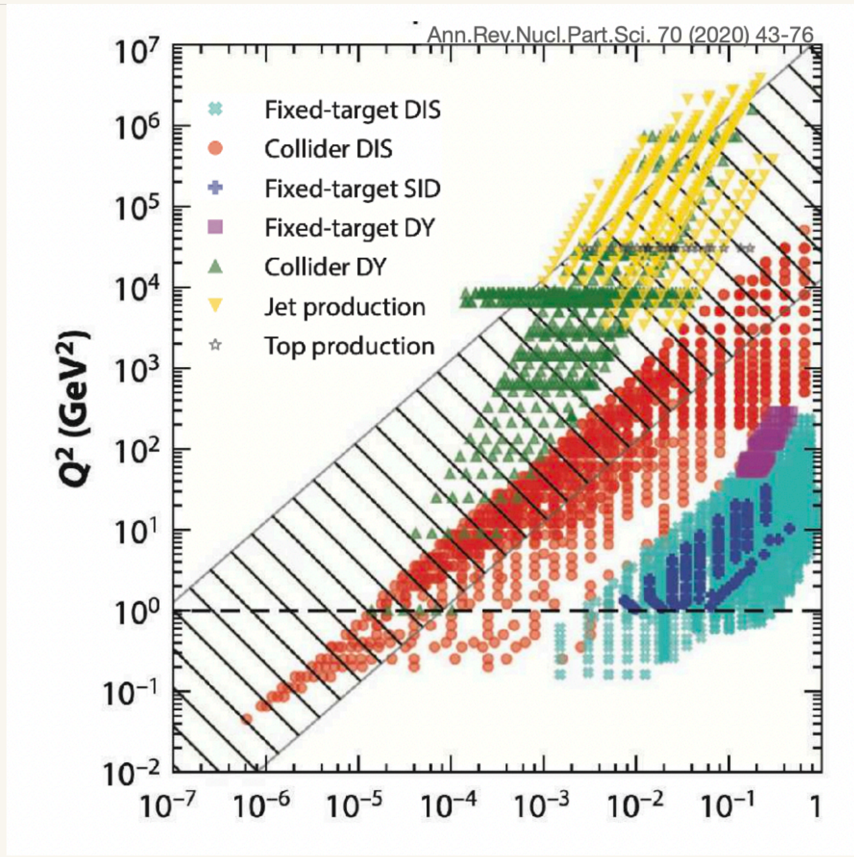
  spelling out the error budget

**This list apply for both the global analyses and lattice.**

# Data for PDF determination

- **experimental data**
  for fixed target, collider DIS, Tevatron, LHC (with a variety of processes)

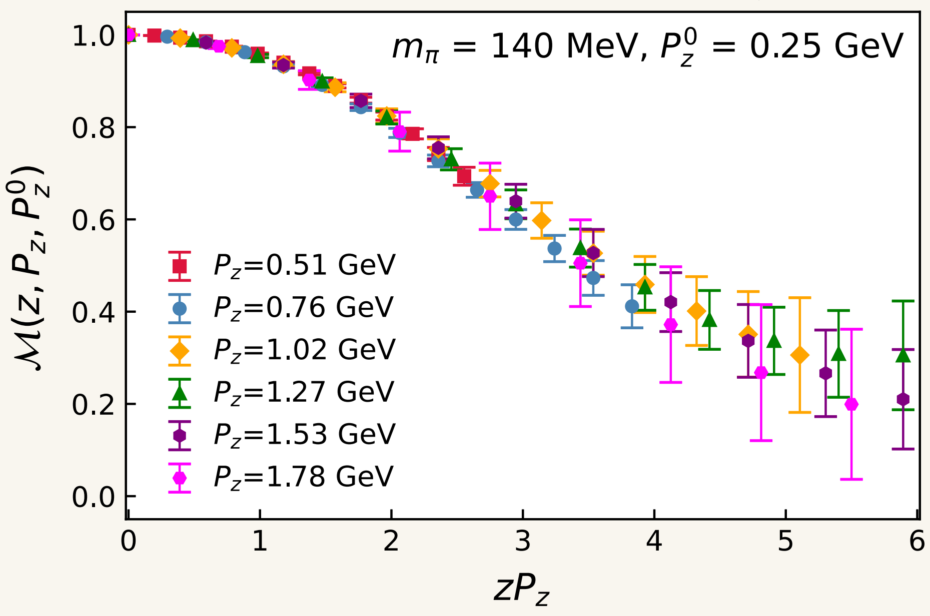
  ~4000 data points for unpolarized PDF




- **lattice data**
  for 3-pt correlation function (in necessary ratios…) from various collaborations

  ~30-50 data points per observable per collaboration

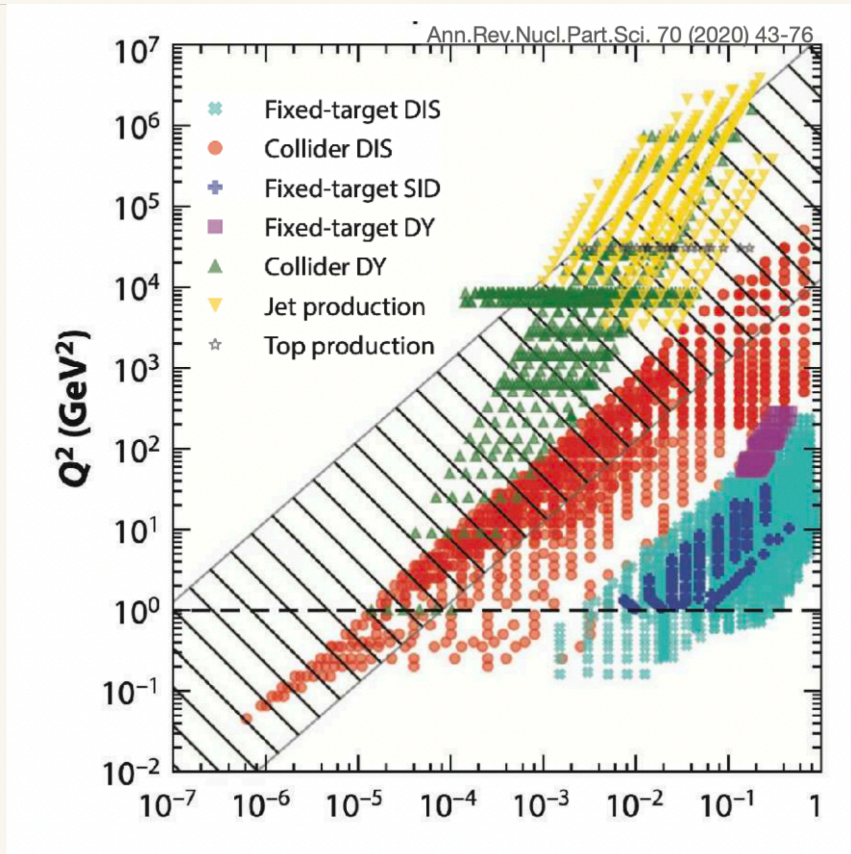
Gao et al, PRD106



# Data for PDF determination

- **experimental data**
  for fixed target, collider DIS, Tevatron, LHC (with a variety of processes)

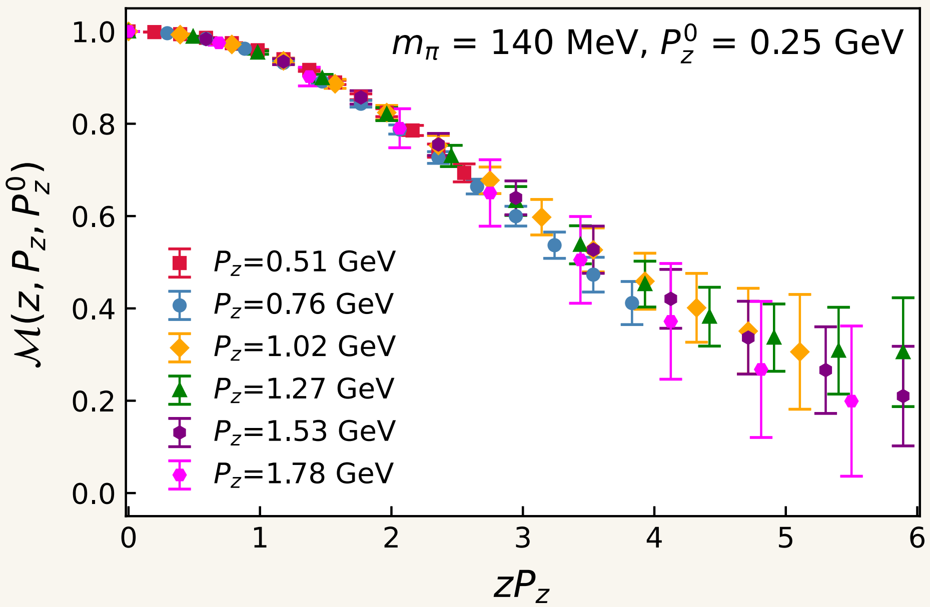
  Kinematics constraints, statistical and systematic uncertainties
  Correlation among data



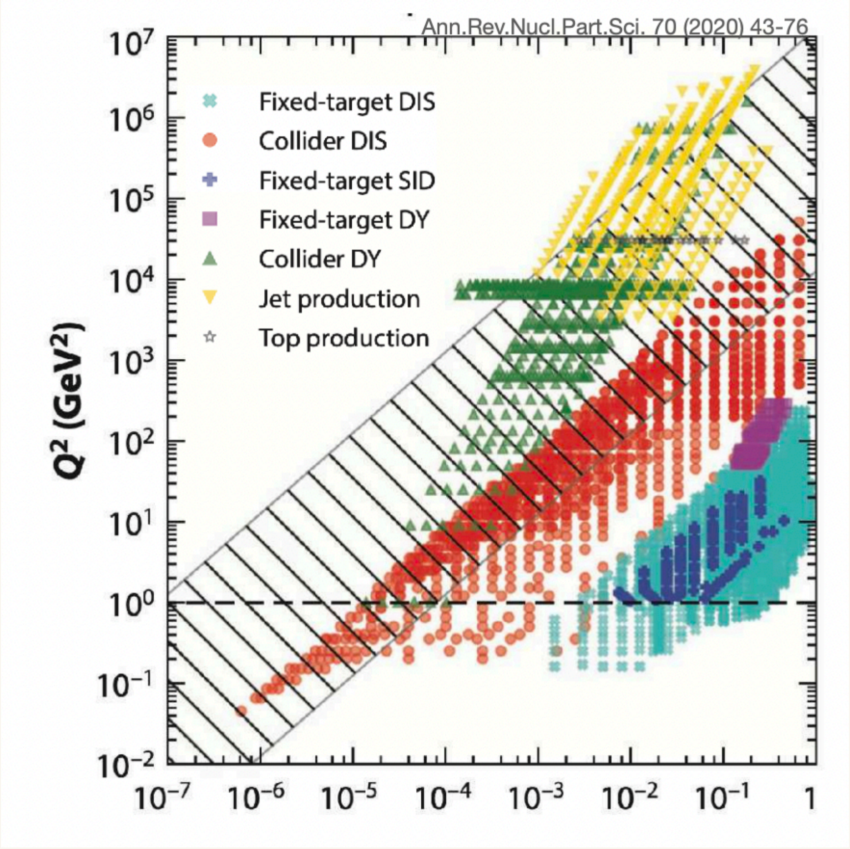

- **lattice data**
  for 3-pt correlation function (in necessary ratios…) from various collaborations

  Lattice configuration, statistical and systematic uncertainties
  Correlation among data

Gao et al, PRD106



# Data for PDF determination



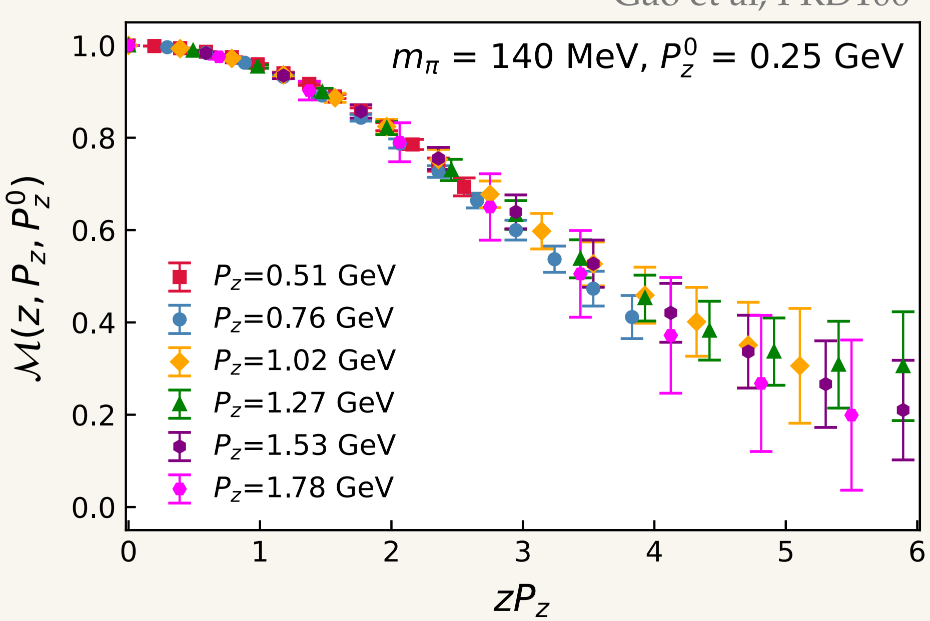

- **experimental data**
  for fixed target, collider DIS, Tevatron, LHC (with a variety of processes)

  Perturbative QCD framework with factorization theorems.

- **lattice data**
  for 3-pt correlation function (in necessary ratios…) from various collaborations

  Short-distance factorization, LaMET formalism, Compton amplitudes, good lattice cross section …

Gao et al, PRD106



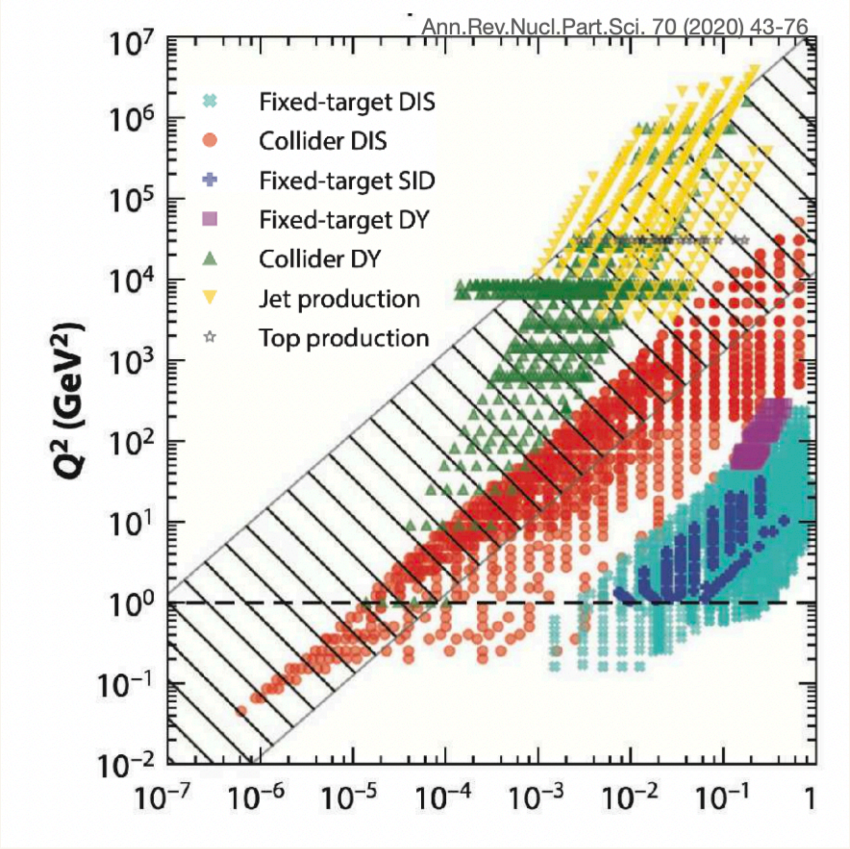
# Data for PDF determination

- **experimental data**
  for fixed target, collider DIS, Tevatron, LHC (with a variety of processes)

  Perturbative QCD framework with factorization theorems.

*Inverse problem analyzed by global QCD analyses practitioners*

***historically started by experimentalists*

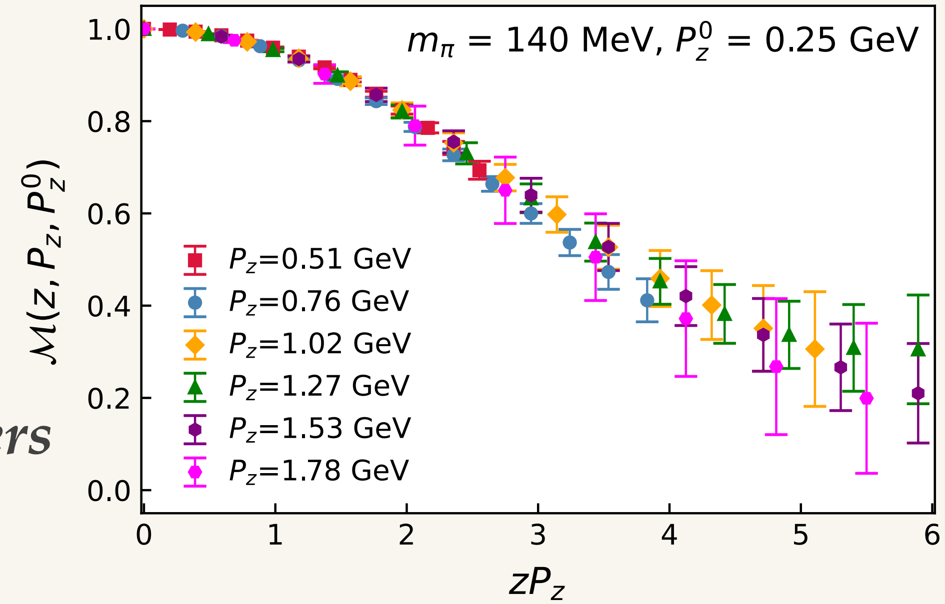



- **lattice data**
  for 3-pt correlation function (in necessary ratios…) from various collaborations

  Short-distance factorization, LaMET formalism, Compton amplitudes, good lattice cross section …

*Inverse problem analyzed by lattice practitioners*

Gao et al, PRD106



# Distribution functions

OPE of the hadronic tensor involves

$$j(z)j(0) = \sum_{i,n} C_n^{(i)}(z^2)\, z^{\mu_1} \ldots z^{\mu_n}\, \mathcal{O}_{\mu_1 \ldots \mu_n}(0)$$

with z on the LC and with $\mathcal{O}$ a 4-field operator

OPE of the spatial correlator, $\tilde{\mathcal{O}}_{\gamma^z}(z,\mu) \propto \bar{\psi}(z)\Gamma\, U(z,0)\psi(0)$, involves

$$\tilde{\mathcal{O}}_{\gamma^z}(z,\mu) \to \sum_n C_n(\mu^2 z^2) \frac{(-iz)^n}{n!} e^{\mu_1} e^{\mu_2} \cdots e^{\mu_n}\, \mathcal{O}_{\mu_0 \mu_1 \cdots \mu_n}(\mu)$$

with $z^\mu = (0,0,0,z)$ and with $\mathcal{O}$ a 2-field operator

In both cases, the Mellin moments can be found

$$\langle P | O_1^{\mu_0 \mu_1 \cdots \mu_n} | P \rangle = 2a_{n+1}(\mu)(P^{\mu_0} P^{\mu_1} \ldots P^{\mu_n} - \text{trace}),$$

Ioffe time, $z \cdot P$, in both OPEs.

# Convolutions

To access the $x$ dependence of PDFs, we must address convolution problems:

- Convolution in structure functions — Wilson coefficients, etc.
- Convolution in lattice observables — Fourier transform and matching conditions, Wilson coefficients etc.

▶ Could they both be treated in a unique framework of global analyses?

*Alternatively:*
*to access the $x$ dependence of PDFs, we must address the Hausdorff moment problem.*

# Methodologies to address inverse problems

*Buzz words to be defined in a glossary*

- Hessian formalism — central value+ covariance matrices
- Neural networks — bootstrap ("Monte Carlo")
- Markov Chain Monte Carlo
- Iterative Monte Carlo — with functional form
- Further AI/ML tools — Variational Auto Encoder, Deep Neural Network, pixelation,…

*Bayesian or frequentist?*

# Uncertainty budget for lattice

- starts with importance sampling (TBC)
- truncation and model averaging
- agreed upon standards for lattice configurations and validation — though not applied in hadron structure
- display of correlation?

# Uncertainty budget for global analyses of exp. data

*Partial opinion of the speaker:*

- Experimental
- Theoretical
- Epistemic:
    - Methodological
    - Parametrization

includes consideration of all sampling sources, of treatment of tensions…

# QCD precision

- Unpolarized PDF up to (a)N3LO, polarized PDFs and TMDs at NLO, GPDs at LO,…
  *Scale given by the kinematics of the physical process.*
- Lattice available with matching coefficients up to NNLO
  *Scale depends on $z$ — limited to ranges where pQCD is valid.*
  $P_z \to \infty$ *necessary to demonstrate convergence.*
- Collins-Soper kernel studies.
- Resummation available for lattice, very few global analyses include it.
- Higher-twist ($\Lambda^2/Q^2$) corrections in both formalisms— treated differently

# Synergy for the combined analysis of experimental & lattice data

*Two main focuses:*

- Mellin moments as integral constraints
- Lattice data (~3-pts correlation function) on the same footing as experimental data
- *less popular idea: use lattice-extracted LC PDF as direct constraints*

Complementarity of data in extrapolation regions or when data scarce (e.g., transversity, GPDs, …)

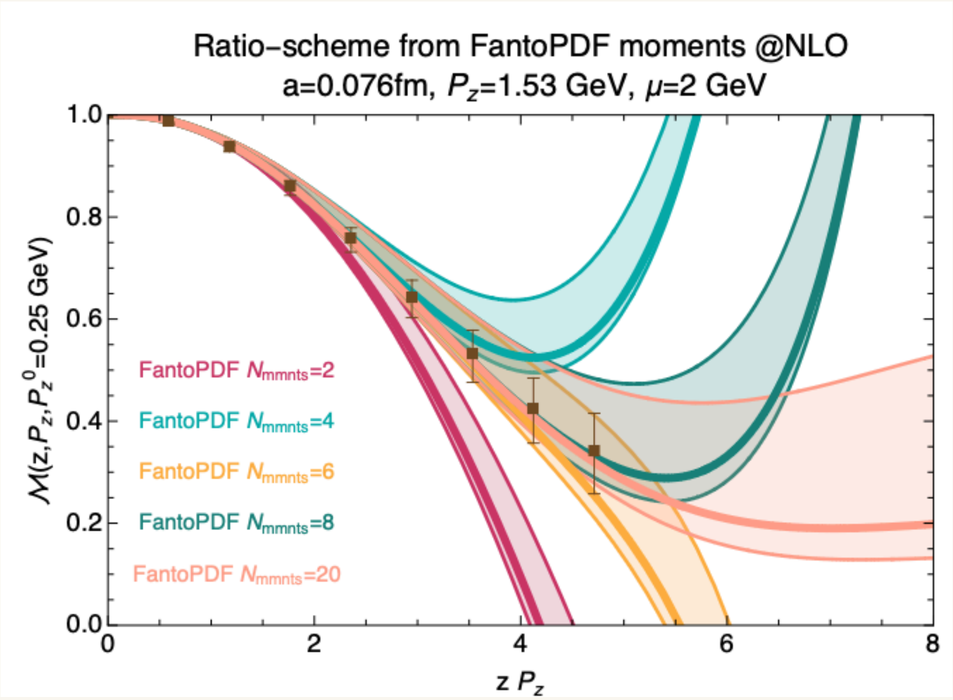
# Benchmarking?

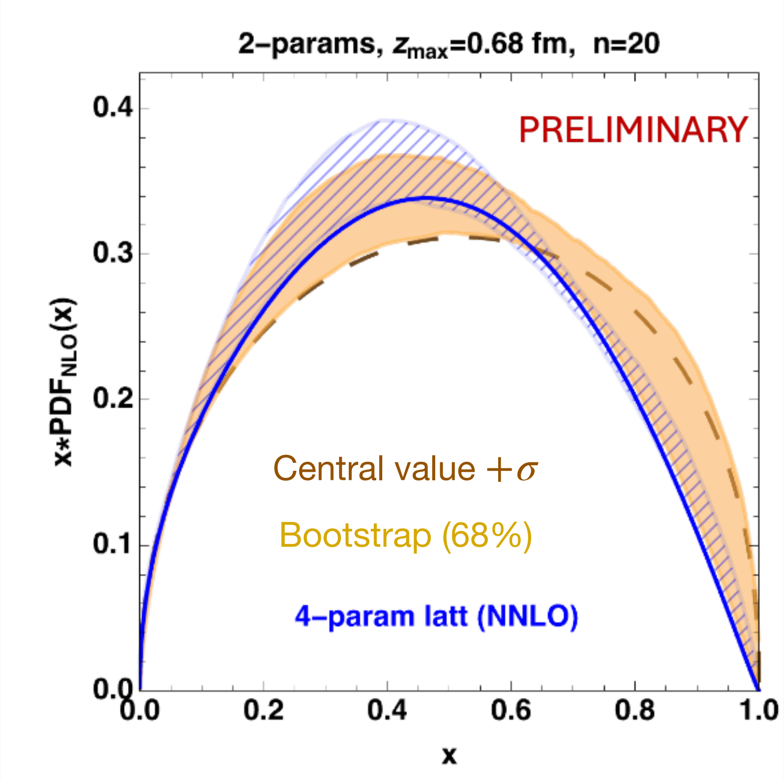
- Lattice: the axial charge as a benchmark example
- Pheno analyses benchmarked unpolarized PDFs — PDF4LHC
- *Benchmark or combined exercise to solve lattice inverse problem with global-fitter tools? And vice versa?*

➡ Would help both sides to get acquainted with limitations, approximations, etc.

# Benchmarking?

## Reproducing [Gao et al, PRD106] from pheno point of view




[Courtoy, Gao, Nadolsky, Zhao and students, in progress]

Other examples that could be explored?

# Original key questions

1. Accessing PDFs: global analyses and lattice computations

   $\longrightarrow$ How does PDF determination work in global analyses and lattice QCD?

2. Global QCD analyses: inverse problem and objective functions

   $\longrightarrow$ How is the inverse problem entailed by PDF determination addresed?

3. Lattice QCD: considerations on the validity of the perturbative matching

   $\longrightarrow$ How is the equivalence between $zP_z$ and $\xi^- P^+$ defined?

4. Setting up a common language: definitions and benchmarks

   $\longrightarrow$ How to benchmark lattice moments and quasi-/pseudo-PDFs with global analyses?

5. Combining lattice and experimental data to determine PDFs

   $\longrightarrow$ What are the efforts/limitations to incorporate lattice data in PDF determinations?

6. Uncertainty quantification and bias/variance trade-off

   $\longrightarrow$ How are aleatoric and epistemic uncertainties combined? How is a model chosen?

# White Paper

*Thanks for the very productive two days of presentations!*

*Now, let's work on the White Paper. No structure defined yet. We can use tools such as:*

- Initial list of questions
- Idea of a glossary
- Basics of statistics
- Exemplary observables
- Common exercise (benchmark)
- ….