Benchmarking lattice determinations of x-dependent hadron structure

Chris Monahan (he/him/his)
Colorado College

A report on (and context for) PRL 135 (2025) 191901

QUARK-GLUON TOMOGRAPHY COLLABORATION

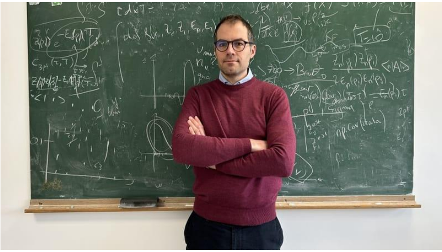
HadStruc

The 43rd International Symposium on Lattice Field Theory
August 1, 2026

Co-authors

Kostas Orginos, W&M

Joe Karpie, JLab



Savvas Zafeiropoulos
CNRS Marseille

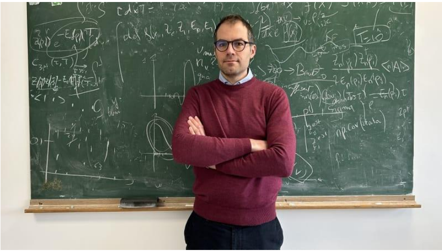
“*Window observables for benchmarking parton distribution functions*”
PRL 135 (2025) 191901

Chosen as a Jefferson Lab Significant Accomplishment
Read more at “Widening windows to peer into the proton”

Co-authors and collaborators

Kostas Orginos, W&M

Joe Karpie, JLab



Savvas Zafeiropoulos
CNRS Marseille

HadStruc Collaboration
Hervé Dutrieux
Robert Edwards
Danieal Kovner
Yamil Medrano
Anatoly Radyushkin
David Richards
Eloy Romero
Jianwei Qiu
Ben Slimmer
Alex Sturzu

Physics friends and recent collaborators
Òscar Crosas
Anna Hasenfratz
Axel Klomparens
Ruby Kofron
Lorenzo Maio
Cédric Mezrag
Kay Quintana
Matthew Rizik
Andrea Shindler
Peter Stoffer
Oliver Witzel

Claim 1:
We do not understand protons and neutrons.

Claim 2:
But! We can, with dedicated community effort.

Claim 1:
We do not understand protons and neutrons.

We have not yet calculated nucleon properties with precision.

Claim 2:
But! We can, with dedicated community effort.

In particular, through quantified, complete error budgets and appropriate benchmarks.

We propose:
Window observables serve as one set of appropriate benchmarks

Can be defined in regions of mutual validity, without extrapolation, for both lattice and experiment

Provide precise benchmarks by taking advantage of correlations in momentum fraction

Outline

Part 1: claim 1 in context

Collinear x-dependent hadron structure on the lattice
The need for benchmarking

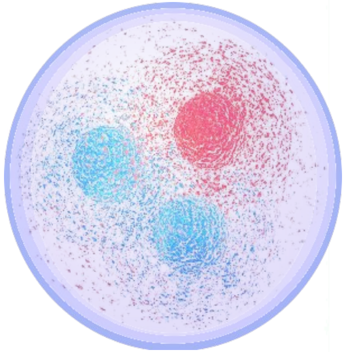
Part 2: our proposal

Window observables
Some preliminary studies

Plenaries
G. Bali, Fri. 11:00
A. Shindler, Sat. 09:45

Hadron structure at a glance

Global properties → Form factors → Parton distribution functions (PDFs) → Generalised parton distribution functions (GPDs) → Wigner functions



How is the spin of the hadron split amongst its constituents?

What are the internal correlations of the proton?

How does the proton interact with the other fundamental forces?

How is the momentum of the constituents correlated with their transverse structure?

How is the momentum of a fast moving hadron spread amongst its constituent quarks and gluons?

Hadron structure at a glance

Global properties → Form factors

$\boxed{q^2 = 0}$ $\boxed{q^2 \neq 0}$

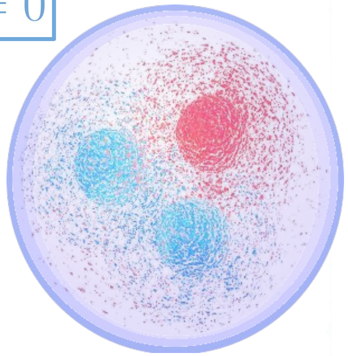
How is the spin of the
hadron split amongst

Extracted from matrix elements of **local** operators

interact with the other
fundamental forces?

Parton distribution functions (PDFs) → Generalised parton distribution functions (GPDs)

$\boxed{q^2 = 0}$ $\boxed{q^2 \neq 0}$



How is the mon
fast moving ha
amongst its c
quarks and

Extracted from matrix elements of **nonlocal** operators

Wigner functions

???

What are the
internal correlations
of the proton?

How is the momentum
of the constituents
correlated with their
structure?

Plenary
A. Shindler, Sat. 09:45

Hadron structure at a glance

Global properties

Form factors

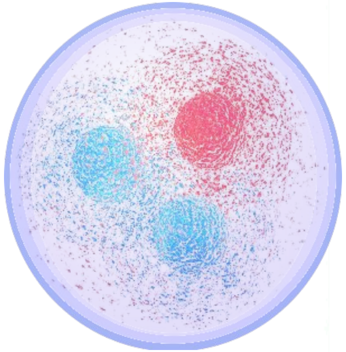
Parton distribution functions (PDFs)

Generalised parton distribution functions (GPDs)

Wigner functions

How is the spin of the hadron split amongst its constituents?

How does the proton interact with the other fundamental forces?

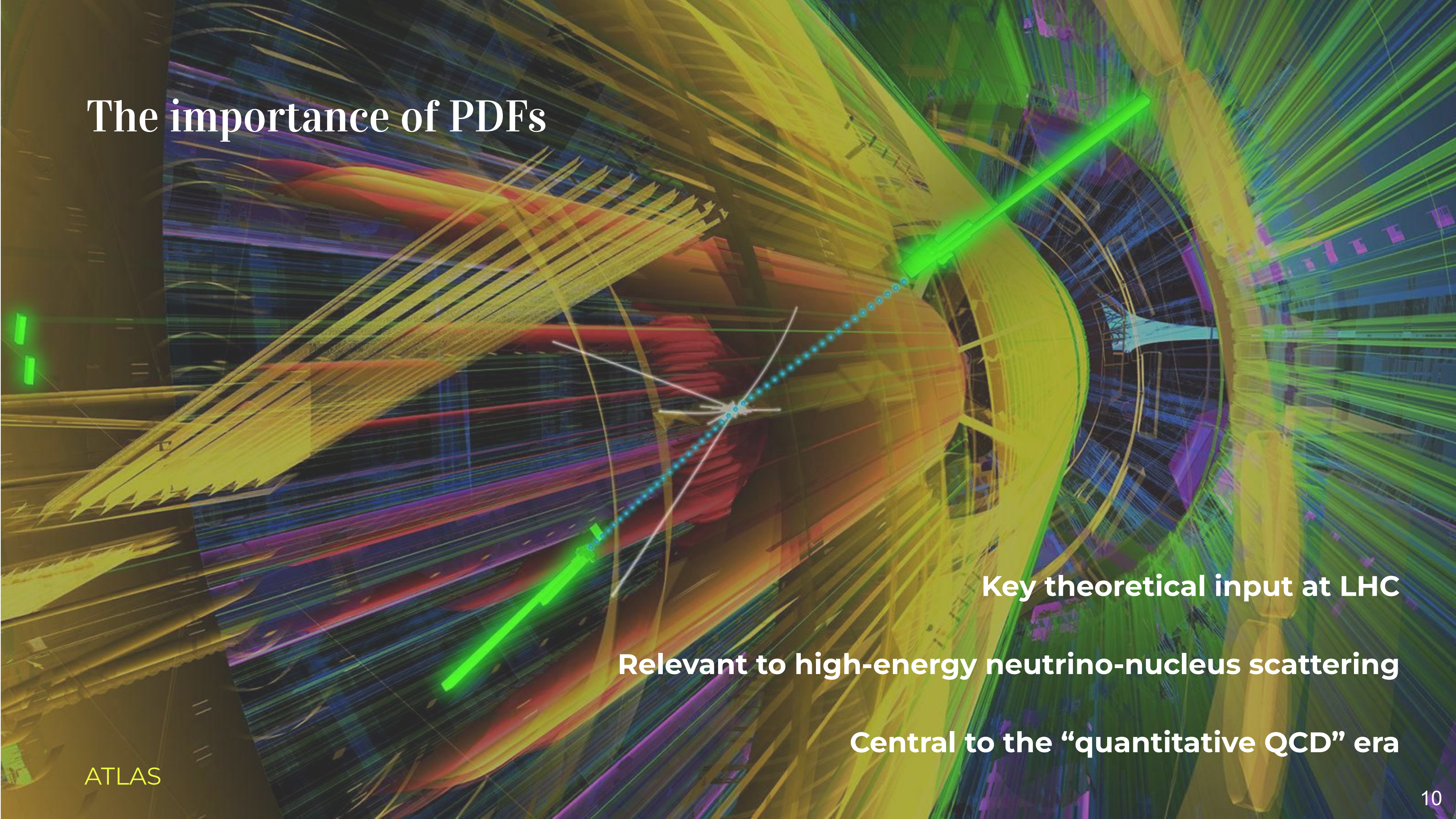


How is the momentum of a fast moving hadron spread amongst its constituent quarks and gluons?

What are the internal correlations of the proton?

How is the momentum of the constituents correlated with their transverse structure?

The importance of PDFs



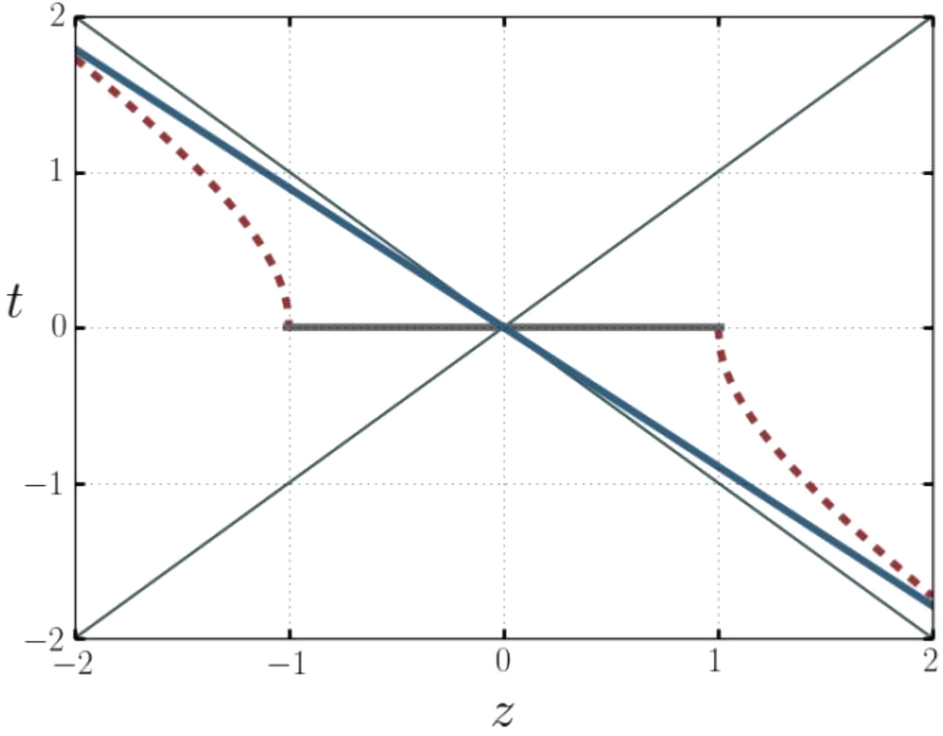
ATLAS

Key theoretical input at LHC

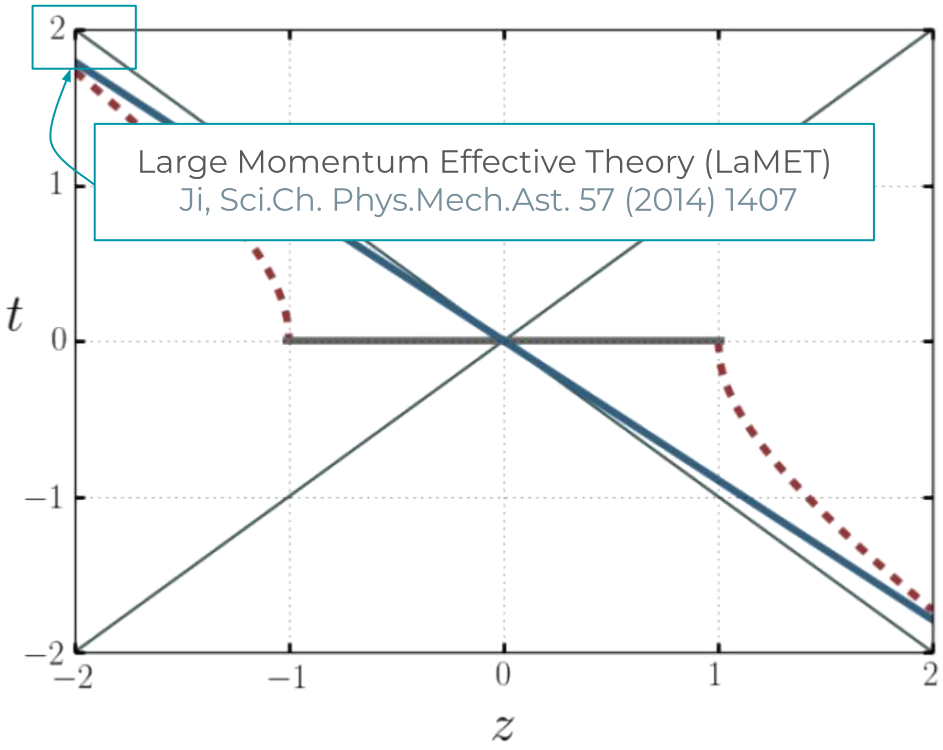
Relevant to high-energy neutrino-nucleus scattering

Central to the “quantitative QCD” era

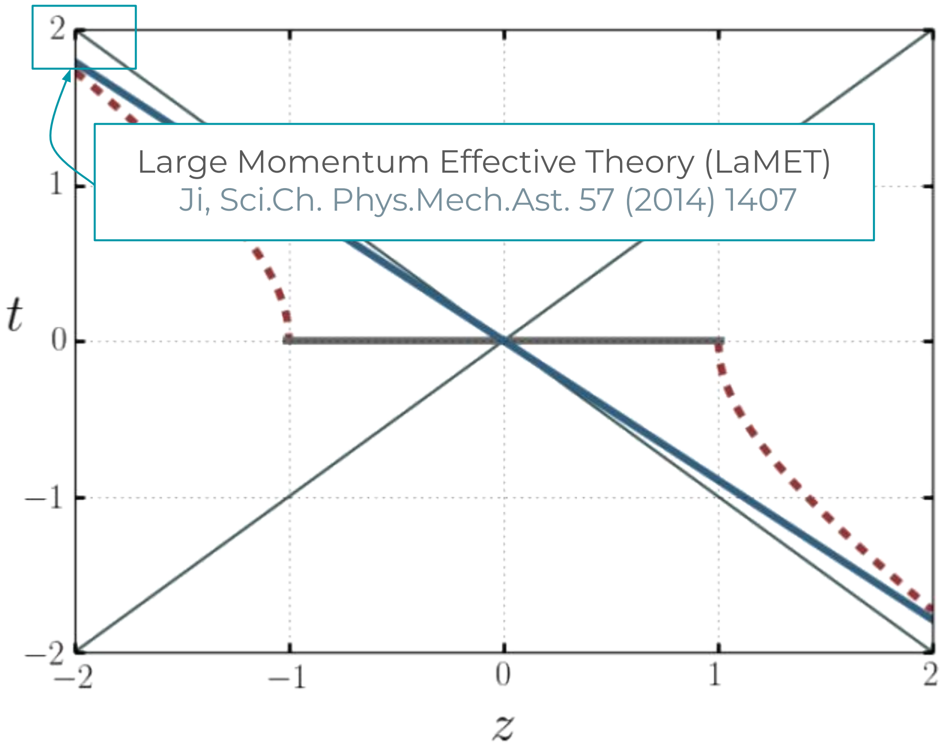
Towards light-cone physics



Towards light-cone physics



Towards light-cone physics



Ji, PRL 110 (2013) 262002
Radyushkin, PRD 96 (2017) 034025

Note history of different approaches:
Davoudi & Savage, PRD 86 (2012) 054505
Musch et al., PRD 83 (2011) 094507
Braun & Müller, EPJC 55 (2008) 349
Detmold & Lin, PRD 73 (2006) 014501
Liu & Dong, PRL 72 (1994) 1790

The distribution zoo

$$\langle P|G^{\mu\nu}(z)W(z,0)G^{\rho\sigma}(0)|P\rangle \longrightarrow M_{pp}(\nu = z\cdot P, z^2)$$

The distribution zoo

$$f_g(x) \sim \int_{-\infty}^{\infty} \mathrm{d}z^- e^{-iz^- xP^+} M_{pp}(\nu, 0)$$

PDFs

$$z^2 = 0 \quad \Big\uparrow \quad z^-$$

$$\langle P|G^{\mu\nu}(z)W(z,0)G^{\rho\sigma}(0)|P\rangle \longrightarrow M_{pp}(\nu = z\cdot P, z^2)$$

The distribution zoo

$$f_g(x) \sim \int_{-\infty}^{\infty} \mathrm{d}z^- e^{-iz^- x P^+} M_{pp}(\nu, 0)$$

PDFs

$$z^2 = 0 \quad z^-$$

$$\langle P | G^{\mu\nu}(z) W(z,0) G^{\rho\sigma}(0) | P \rangle \longrightarrow M_{pp}(\nu = z \cdot P, z^2)$$

$$z \qquad z^2 < 0 \qquad \nu$$

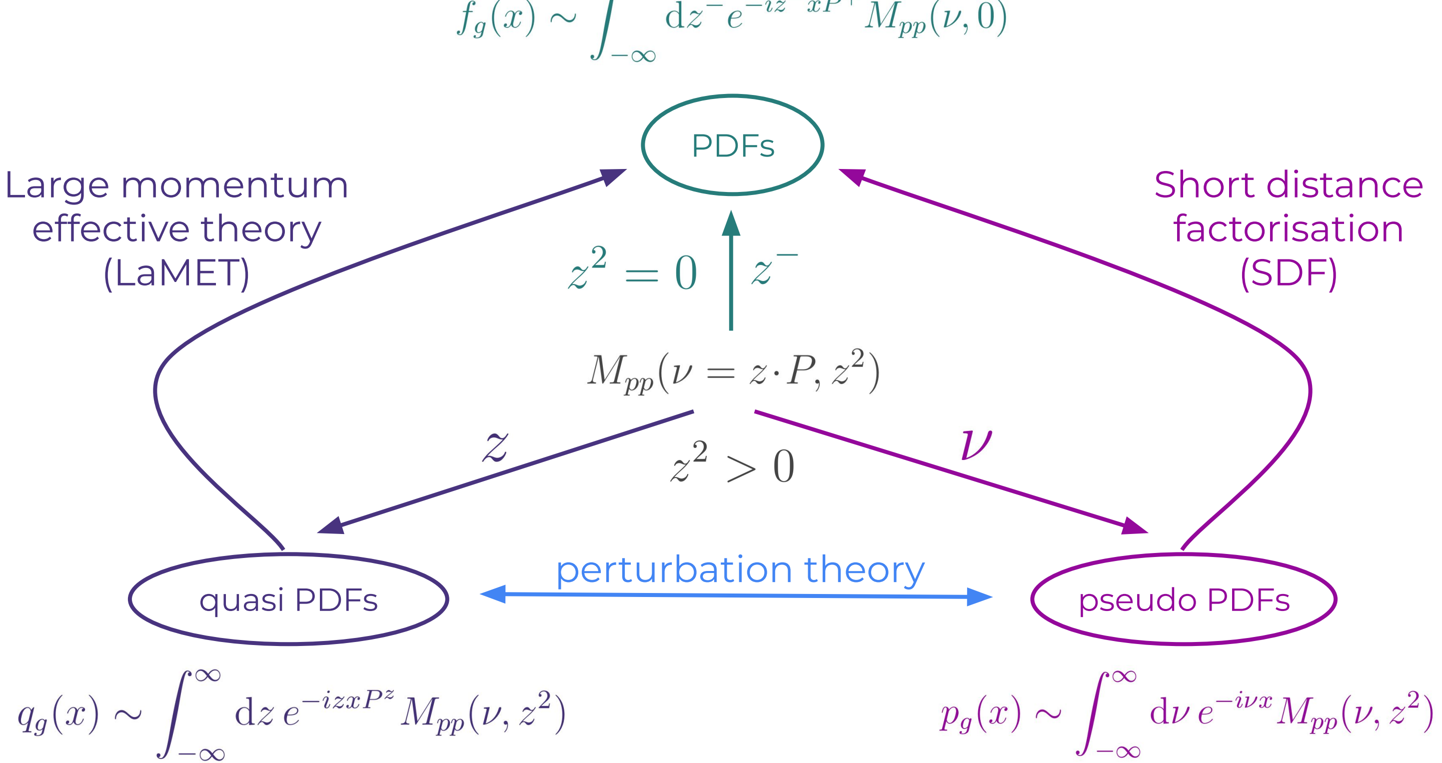
quasi PDFs

pseudo PDFs

$$q_g(x) \sim \int_{-\infty}^{\infty} \mathrm{d}z\, e^{-izxP^z} M_{pp}(\nu, z^2)$$

$$p_g(x) \sim \int_{-\infty}^{\infty} \mathrm{d}\nu\, e^{-i\nu x} M_{pp}(\nu, z^2)$$

The distribution zoo



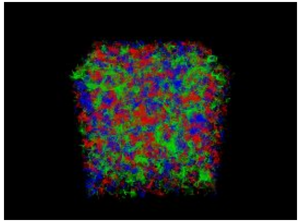
Complementary windows into hadron structure

"Validation from variety"
Wilson Award
N. Klco, Thurs. 11:30

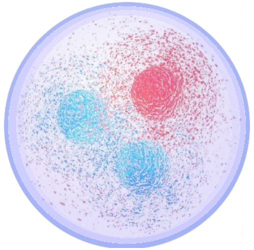
Plenary
G. Bali, Fri. 11:00

Ioffe-time space

$$I_{q/H}(\nu, \mu^2)$$



CSSM, Adelaide



LHC

$$a_n(\mu^2)$$

$$f_{q/H}(x, \mu^2)$$

Mellin space

momentum-fraction space

Plenary
A. Shindler, Sat. 09:45

Plenary
C. Aidala, Fri. 10:00

Examining uncertainties

$$\langle P|G^{\mu\nu}(z)W(z,0)G^{\rho\sigma}(0)|P\rangle \longrightarrow M_{pp}(\nu = z\cdot P, z^2)$$

Examining uncertainties

$$\langle P|G^{\mu\nu}(z)W(z,0)G^{\rho\sigma}(0)|P\rangle \longrightarrow M_{pp}(\nu = z\cdot P, z^2)$$

Typical lattice sources of uncertainty:

- statistics
- excited state contamination
- discretisation effects
- chiral effects
- finite volume effects

Examining uncertainties

perturbative truncation effects

power corrections

PDFs

Short distance factorisation (SDF)

renormalisation of nonlocal operators

$$\langle P|G^{\mu\nu}(z)W(z,0)G^{\rho\sigma}(0)|P\rangle \longrightarrow M_{pp}(\nu = z\cdot P, z^2)$$

ν

Typical lattice sources of uncertainty:

- statistics
- excited state contamination
- discretisation effects
- chiral effects
- finite volume effects

Ill-posed inverse problem:

Fourier transform

pseudo PDFs

Examining uncertainties

perturbative truncation effects

power corrections

PDFs

Short distance factorisation (SDF)

renormalisation of nonlocal operators

$$\langle P|G^{\mu\nu}(z)W(z,0)G^{\rho\sigma}(0)|P\rangle \longrightarrow M_{pp}(\nu = z\cdot P, z^2)$$

ν

Typical lattice sources of uncertainty:

- statistics
- excited state contamination
- discretisation effects
- chiral effects
- finite volume effects

Ill-posed inverse problem:

Fourier transform

pseudo PDFs

Ill-posed inverse problems

Determination of PDFs via Fourier transform is an ill-posed inverse problem

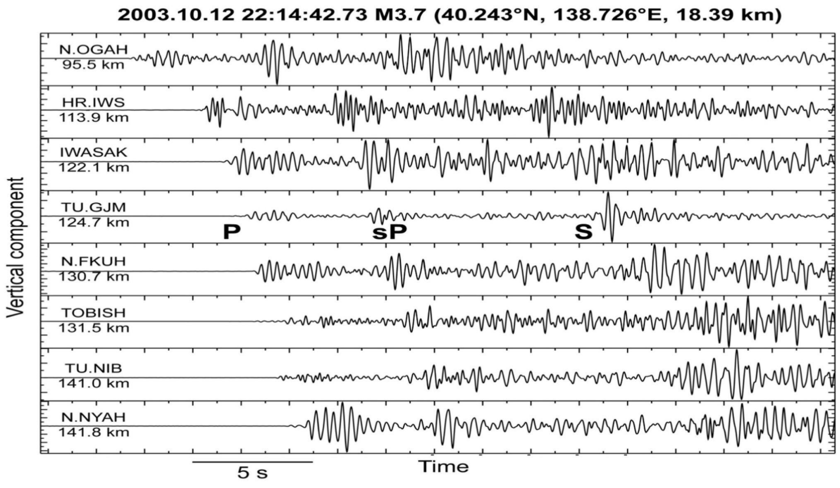
"Starting with the effects to discover the causes"

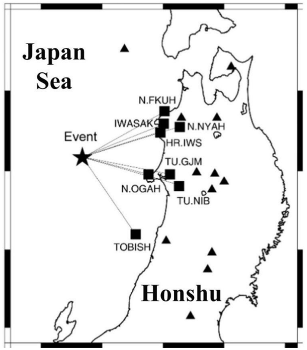
Plenary
S. Hashimoto, Mon. 11:00

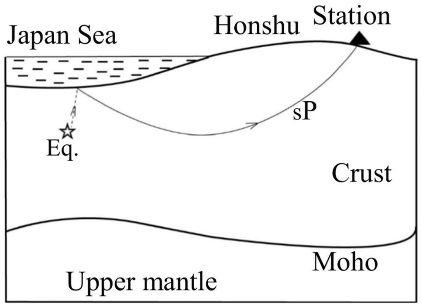
Ill-posed inverse problems

Determination of PDFs via Fourier transform is an ill-posed inverse problem

“Starting with the effects to discover the causes”







Zhao, Phys. EPI 296 (2019) 106314

Ill-posed inverse problems

Determination of PDFs via Fourier transform is an ill-posed inverse problem

> "Starting with the effects to discover the causes"

Given map *M(p)*, we want model parameters {*p*} that match observed data

$$d_{\text{obs.}} = M(p) \qquad \Rightarrow \qquad p = M^{-1}(d_{\text{obs.}})$$

Well-posed (according to Hadamard) if solution:

1. Exists
2. Is unique
3. And is stable

Fourier transforms of discrete, noisy data = ill-posed

Determination of PDFs via Fourier transform is an ill-posed inverse problem

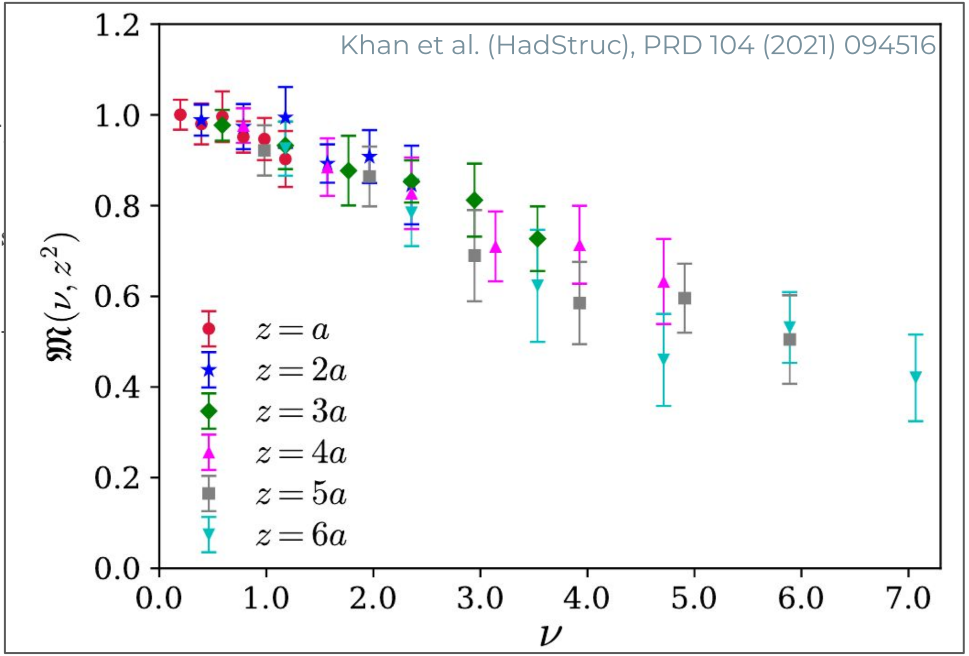
"Starting with the effects to discover the

Given map *M(p)*, we want model parameter

$$d_{\text{obs.}} = M(p) \qquad \Rightarrow \qquad p = M^{-1}(d_{\text{obs}}$$

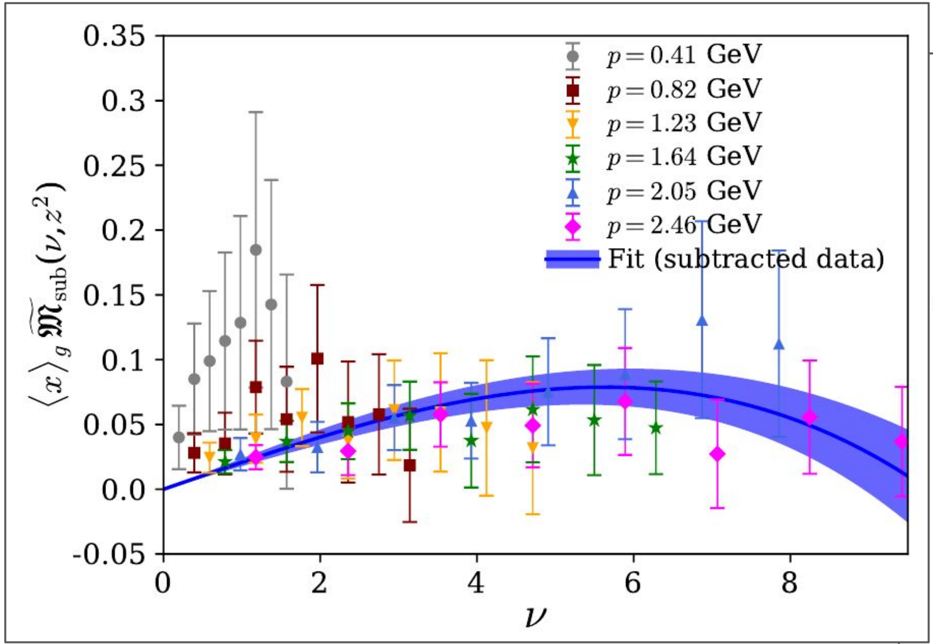
Well-posed (according to Hadamard) if solu

1. Exists
2. Is unique
3. And is stable

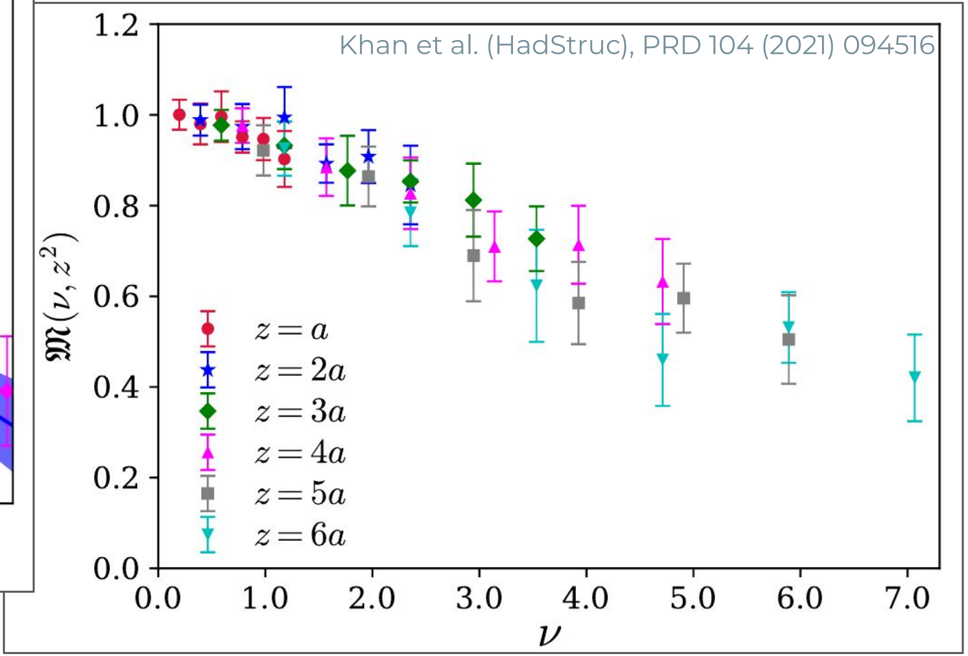


Fourier transforms of discrete, noisy data = ill-posed

Determination of PDFs via Fourier transform is an ill-posed inverse problem



Egerer et al. (HadStruc), PRD 106 (2022) 094511



Khan et al. (HadStruc), PRD 104 (2021) 094516

Power corrections

Perturbative matching of light-cone and quasi PDFs via LaMET valid up to power corrections

$$f_g(x,\mu) \sim C\left(x,\mu\right) \otimes q_g(x,\mu) + \mathcal{O}\left(\frac{\Lambda_{\mathrm{QCD}}^2}{x^2 P_z^2}, \frac{\Lambda_{\mathrm{QCD}}^2}{(1-x)^2 P_z^2}\right)$$

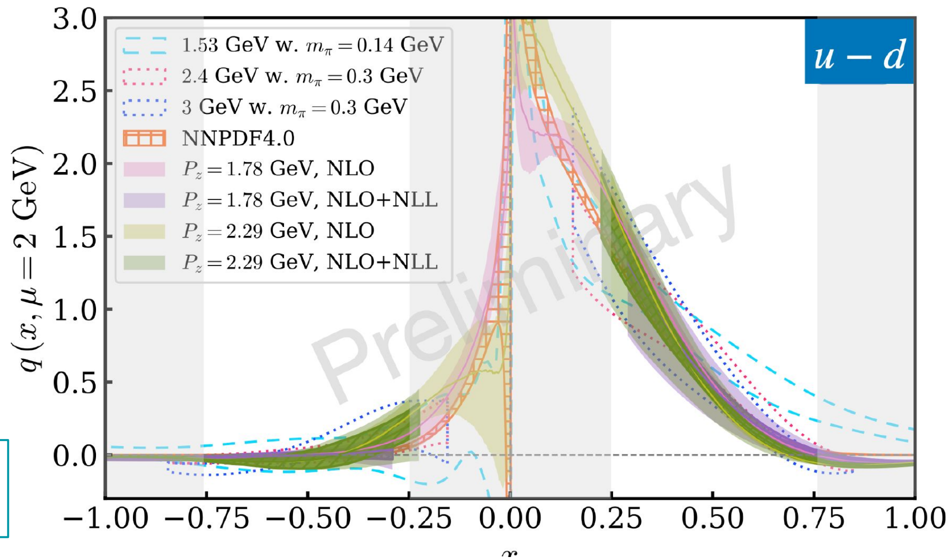
Power corrections

Perturbative matching of light-cone and quasi PDFs via LaMET valid up to power corrections

$$f_g(x,\mu) \sim C(x,\mu) \otimes q_g(x,\mu) + \mathcal{O}\left(\frac{\Lambda^2_{\mathrm{QCD}}}{x^2P_z^2}, \frac{\Lambda^2_{\mathrm{QCD}}}{(1-x)^2P_z^2}\right)$$

large for x ~ 0

large for x ~ 1



Thanks to Q. Shi
Wed. 09:20

Outline

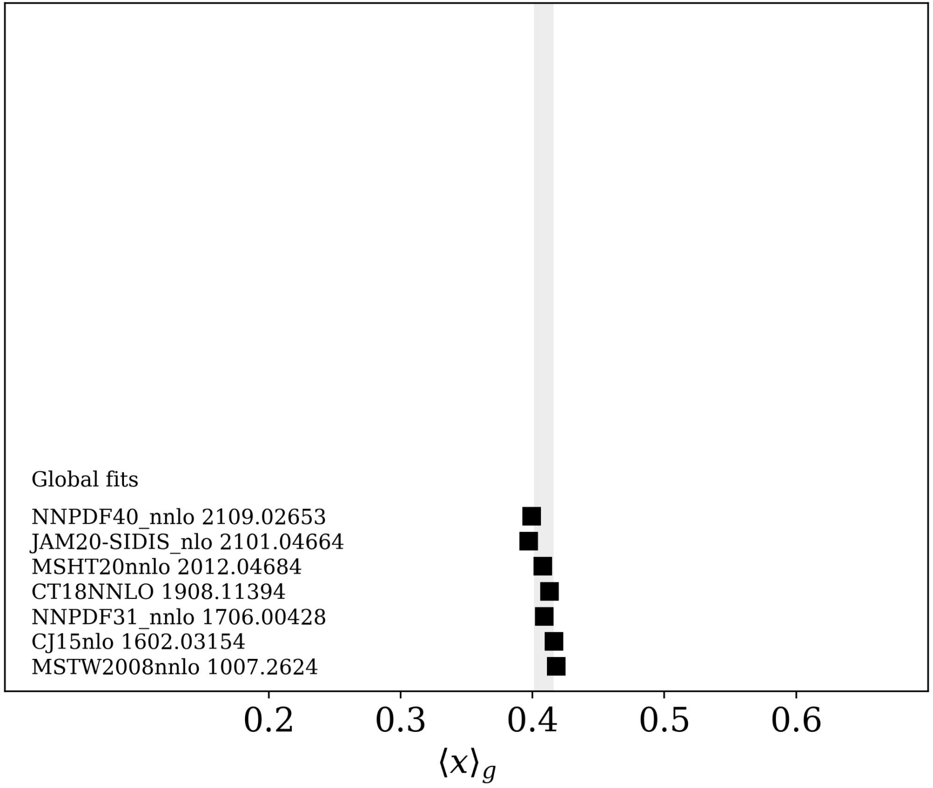
Part 1: claim 1 in context

Collinear x-dependent hadron structure on the lattice
The need for benchmarking

Part 2: our proposal

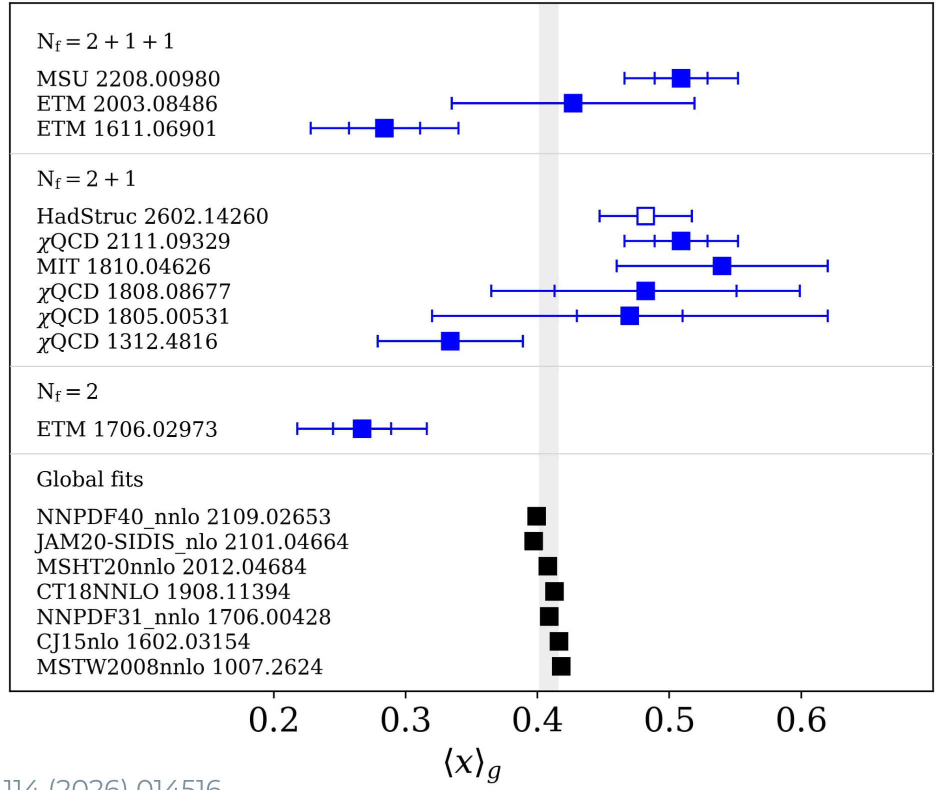
Window observables
Some preliminary studies

Gluon momentum fraction of the nucleon



Gluon momentum fraction of the nucleon

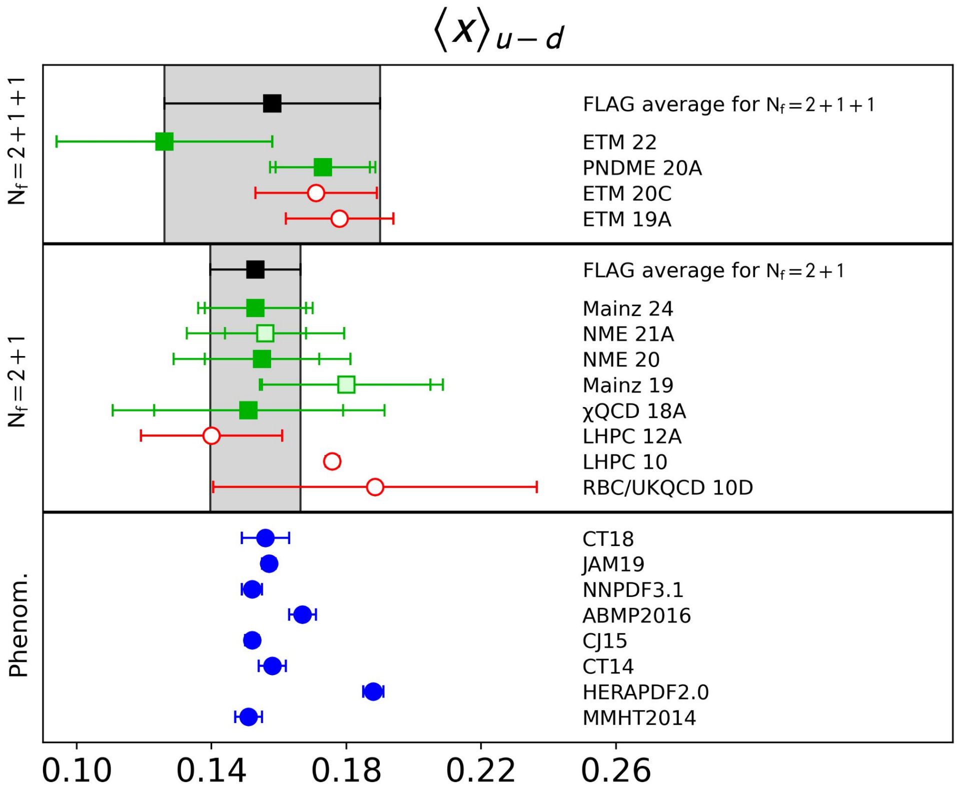
A. Sturzu, Tu. 16:10



Edwards et al. [HadStruc], PRD 114 (2026) 014516

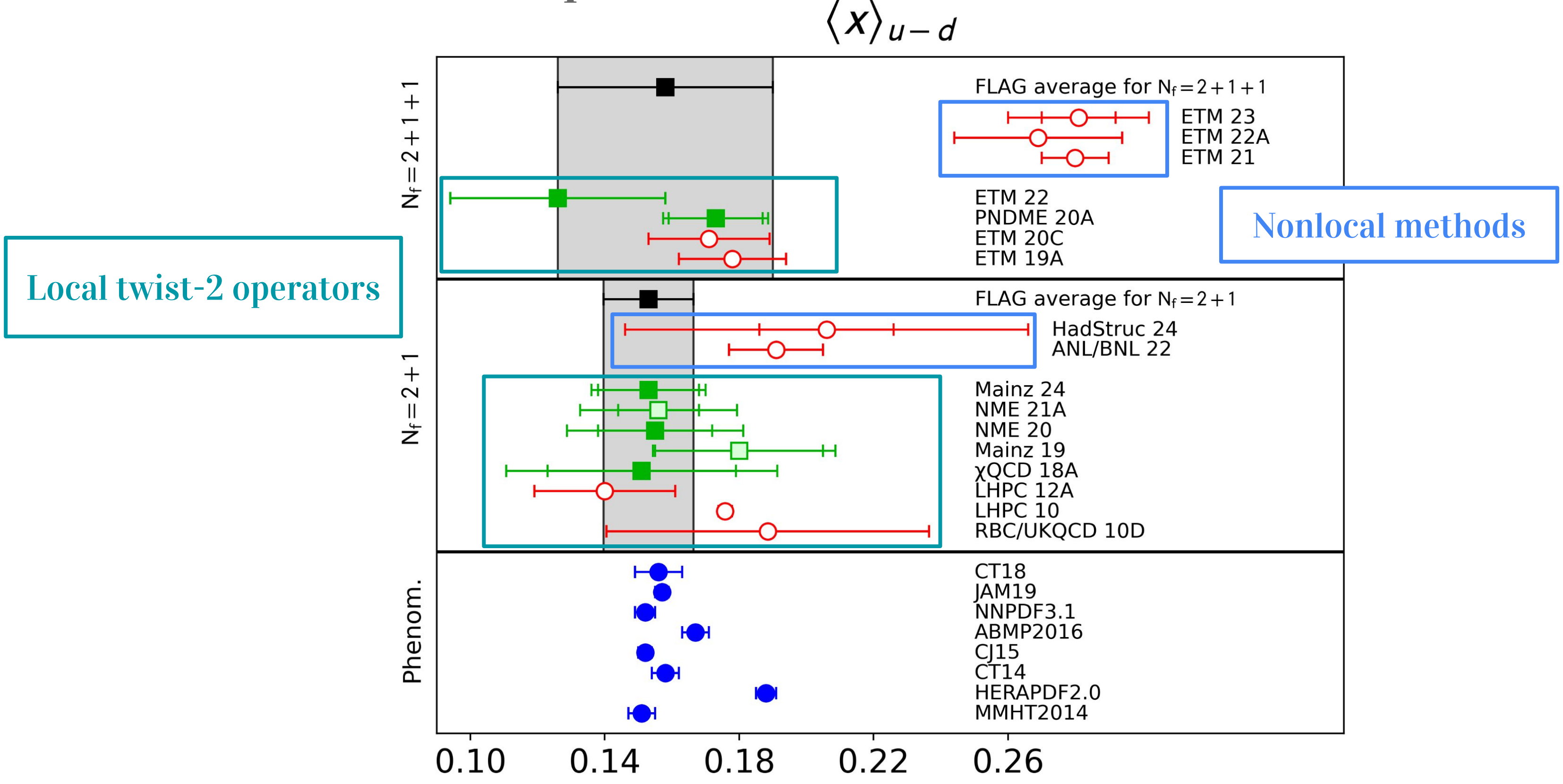
Local operators

With apologies to FLAG Nucleon Matrix Element Working Group
Based on FLAG, PRD 113 (2026) 014508



Local and nonlocal operators

With apologies to FLAG Nucleon Matrix Element Working Group
Based on FLAG, PRD 113 (2026) 014508



Outline

Part 1: claim 1 in context

Collinear x-dependent hadron structure on the lattice
The need for benchmarking

Part 2: our proposal

Window observables
Some preliminary studies

Our proposal: window observables

Karpie et al., PRL 135 (2025) 191901

Propose **window observables** for benchmarking lattice calculations and global fits

$$W^{(f)}(x_\pm) = \int_{-1}^{1} \mathrm{d}x\, w(x; x_\pm) f(x)$$

For example, window moments:

$$a_n(x_-, x_+) = \int_{x_-}^{x_+} \mathrm{d}x\, x^n f(x)$$

Motivated by analogy with (and enormous success of) windows for muon g-2

Blum et al., RBC/UKQCD PRL 121 (2018) 022003

Plenary
S. Hashimoto, Mon. 11:00

In the spirit of Shoji Hashimoto's take-home message
"Don't ask if the ~~spectral functions~~ PDFs can be fully reconstructed;
Ask if phenomenologically useful observables can be constructed."

Our proposal: window observables

Karpie et al., PRL 135 (2025) 191901

Propose **window observables** for benchmarking lattice calculations and global fits

$$W^{(f)}(x_\pm) = \int_{-1}^{1} \mathrm{d}x \, w(x; x_\pm) f(x)$$

For example, window moments:

$$a_n(x_-, x_+) = \int_{x_-}^{x_+} \mathrm{d}x \, x^n f(x)$$

Goal is "validation from variety" and precision

Wilson Award
N. Klco, Thurs. 11:30

Two key advantages as benchmarks that help achieve this goal:

1. Can be defined in regions of mutual validity for both lattice and experiment
2. Provide precise benchmarks by taking advantage of correlations in x

Our proposal: window observables

Karpie et al., PRL 135 (2025) 191901

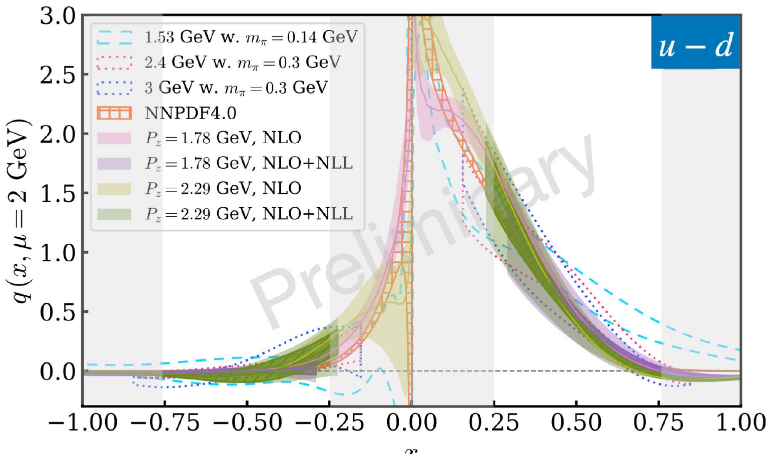
Propose **window observables** for benchmarking lattice calculations and global fits

$$W^{(f)}(x_\pm) = \int_{-1}^{1} \mathrm{d}x\, w(x; x_\pm) f(x)$$

For example, window moments:

$$a_n(x_-, x_+) = \int_{x_-}^{x_+} \mathrm{d}x\, x^n f(x)$$

Goal is "validation from variety" and precision



Thanks to Q. Shi
Wed. 09:20

Two key advantages as benchmarks that help achieve this goal:

1. Can be defined in regions of mutual validity for both lattice and experiment
2. Provide precise benchmarks by taking advantage of correlations in x

Proof-of-principle results

Egerer et al., PRD 105 (2022) 034507

Extracted window moments from lattice data

- 2+1 clover-improved Wilson fermions with tree-level tadpole-improved Symanzik glue
- configurations generated by the JLab-LANL-MIT-WM collaboration

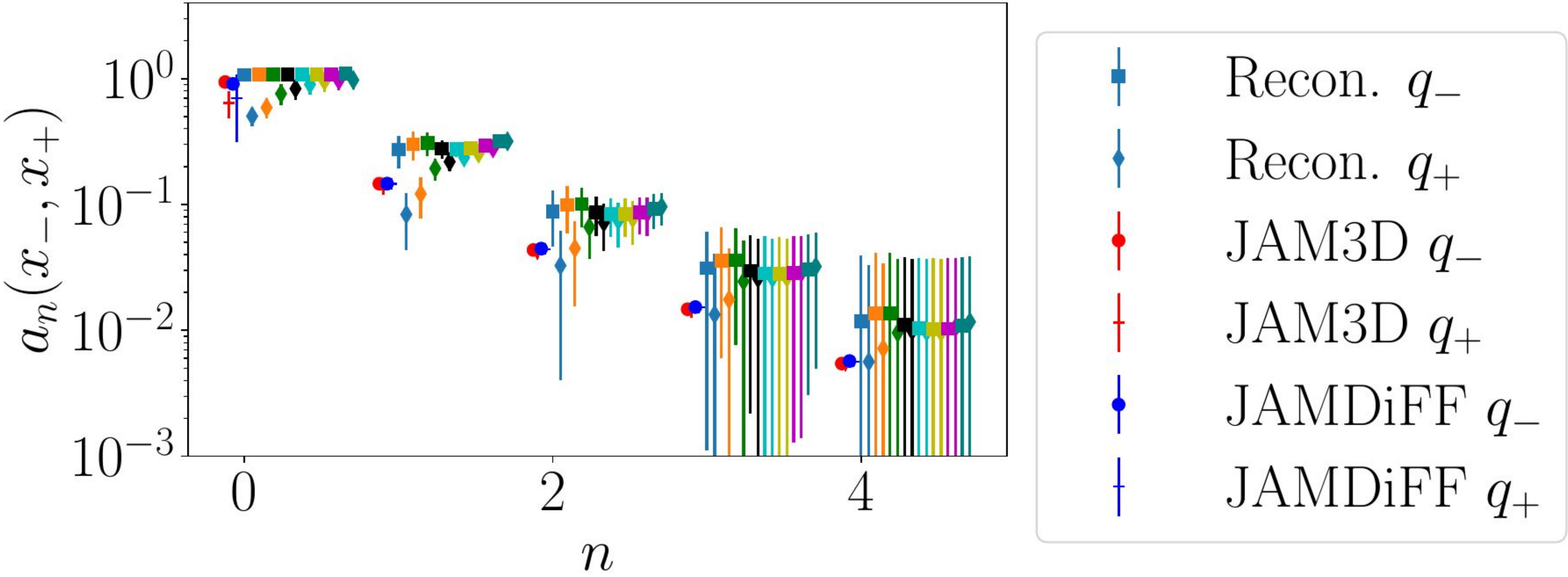
ID	a (fm)	m_π (MeV)	$L^3 \times T$	N_{cfg}	N_{srcs}	N_{D}
$a094m358$	0.094(1)	358(3)	$32^3 \times 64$	1121	64	64

Compared with two datasets from the JAM collaboration

- JAM3D
- JAMDiFF

Gamberg et al., PRD 106 (2022) 034014
Cocuzza et al., PRL 132 (2024) 091901
Gamberg et al., PRD 109 (2024) 034024

Proof-of-principle comparison

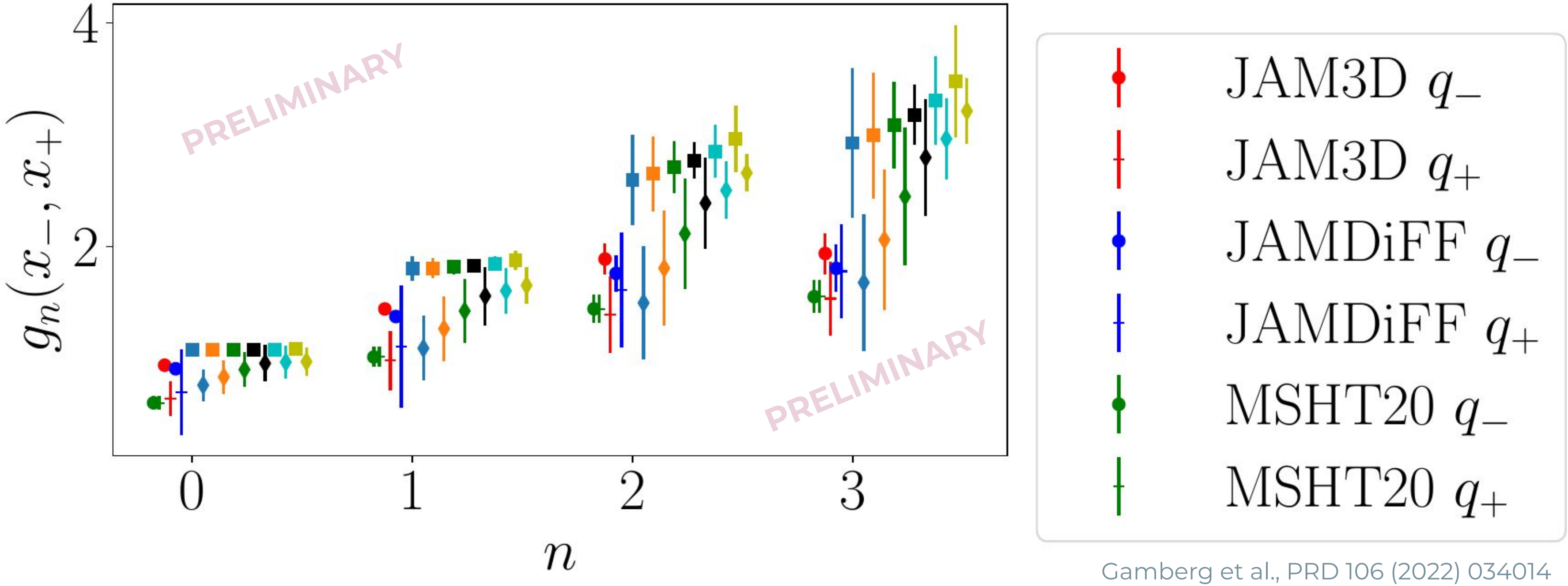


Egerer et al., PRD 105 (2022) 034507

Gamberg et al., PRD 106 (2022) 034014
Cocuzza et al., PRL 132 (2024) 091901
Gamberg et al., PRD 109 (2024) 034024

Preliminary work with C. Flore

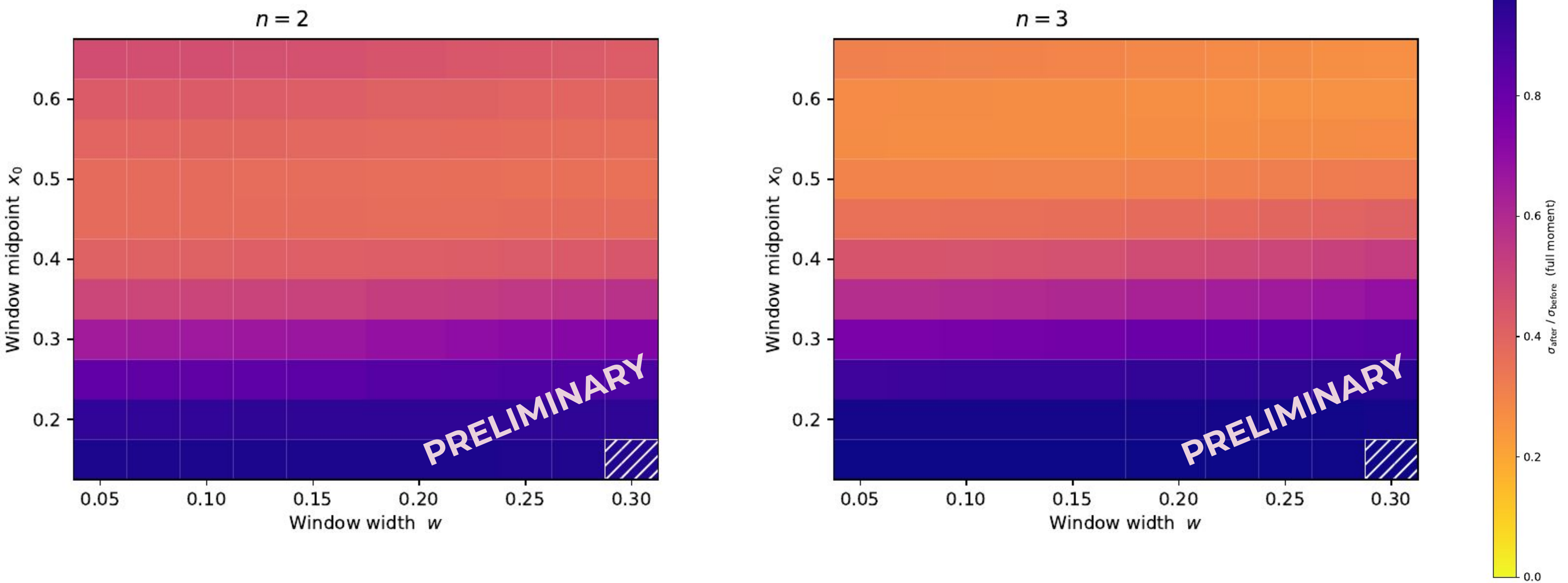
Proof-of-principle comparisons



Egerer et al., PRD 105 (2022) 034507

Gamberg et al., PRD 106 (2022) 034014
Cocuzza et al., PRL 132 (2024) 091901
Gamberg et al., PRD 109 (2024) 034024
Bailey et al., EPJC 81 (2021) 341
C. Flore, arXiv:2409.18751

Future impact?



Tests via the epump framework

Schmidt et al., PRD 98 (2018) 094005

(Reformulated) Claim 1:
We are entering the "quantitative QCD" era for protons and neutrons on the lattice.

We have not (yet) calculated nucleon properties with precision.

(Reformulated) Claim 2:
Our community history shows that dedicated community effort can achieve great things.

In particular, through quantified, complete error budgets and appropriate benchmarks.

We propose:
Window observables serve as one set of appropriate benchmarks

Can be defined in regions of mutual validity, without extrapolation, for both lattice and experiment

Provide precise benchmarks by taking advantage of correlations

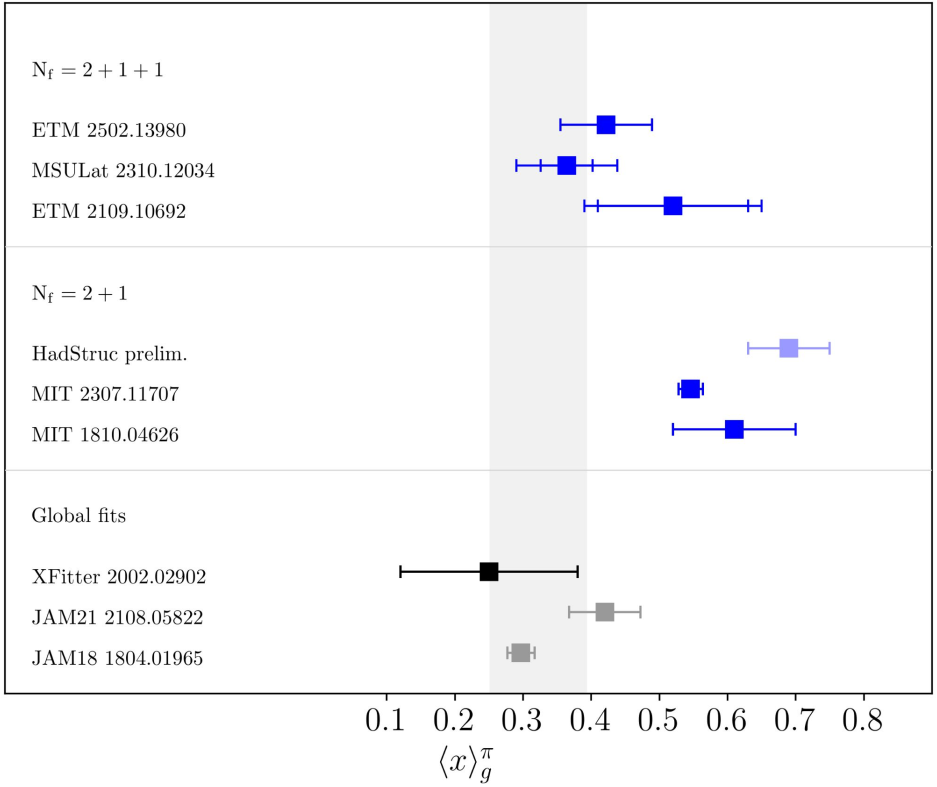
Thank you!

cjmonahan@coloradocollege.edu

QUARK-GLUON TOMOGRAPHY COLLABORATION

HadStruc

Gluon momentum fraction of the pion



PDFs from and for experiment

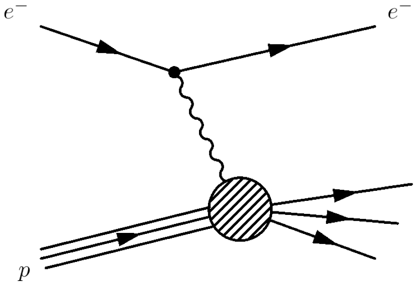
Extracted from and input into inclusive high-energy lepton-hadron scattering processes

$$e^- \, (\gamma^*) \, H \to e^- X$$

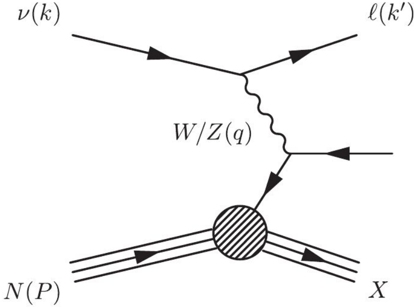
$$\nu_\ell \, (W^{\pm\,*}) \, H \to \ell^\pm X$$

$$\nu_\ell \, (Z^*) \, H \to \nu_\ell \, X$$

Cross-sections factorise: perturbative scattering coefficient x nonperturbative PDFs



Universität Zürich



Xie et al. (CTEQ-TEA), PRD 109 (2024) 113001

Additional details for inverse problems

Overcoming noise

$$d_{\text{obs.}} = M(p) \qquad \Rightarrow \qquad p = M^{-1}(d_{\text{obs.}})$$

Three obvious strategies to overcome noise in inverse problems

1. Acquire more accurate data to reduce $|p_1 - p_2|$ arising from noise
2. Change the map *M(p)* and acquire different data
3. Restrict the space in which parameters are sought

In other words, introduce a prior model:

1. Penalisation procedure (e.g. a regularisation)
2. Bayesian framework

Note that some authors consider the first option a subset of the latter, with Bayesian methods having the advantage of making the impact of prior information explicit.

Noise and ill-posed inverse problems

In the presence of noisy data, we can capture stability through stability estimates

$$\boxed{|p_1 - p_2| \leq f(|M(p_1) - M(p_2)|)}$$

Provides a numerical, but subjective, approach to ill-posedness:

- when noise is significantly amplified, the problem is ill-posed
- noise can also lead to data falling outside the range of *M(p)*

Natural to introduce noise through a modified forward map

$$\boxed{d_{\text{obs.}} = M(p) + \eta}$$

- statistical properties of noise become important
- Monte Carlo statistical uncertainty well-modelled by Gaussian noise
- less clear how to analytically model lattice systematic uncertainties

Some historical strategies

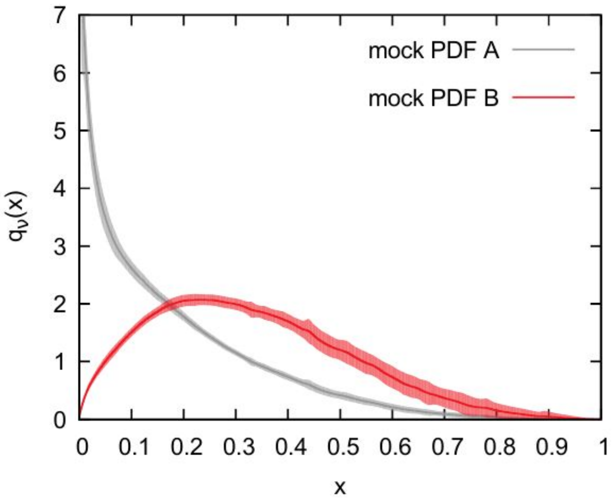
Three obvious strategies to overcome noise in inverse problems

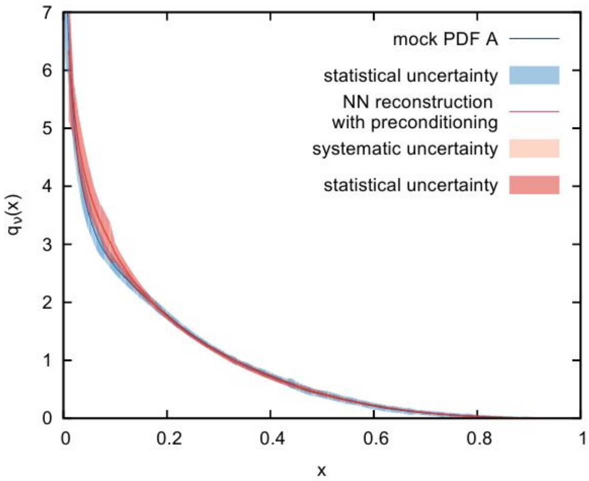
1. Acquire more accurate data to reduce $|p_1 - p_2|$ arising from $|\eta|$
2. Change the map $M(p)$ and acquire different data
3. Restrict the space in which parameters are sought

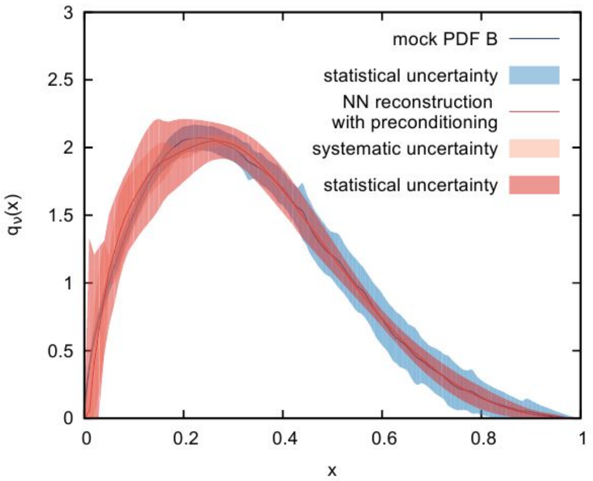
Lin et al. (LP[3]), PRD 91 (2014) 054510	Direct Fourier transform
Lin et al. (LP[3]), PRD 98 (2018) 054504	Low-pass filtering
Ishikawa et al., SCPMA 62 (2019) 991021	"Derivative method"
	Gaussian weighting
	Backus-Gilbert
Karpie et al., JHEP 04 (2019) 057	Maximum entropy method
	Bayesian reconstruction
Dutrieux et al., PRD 111 (2025) 034515	Neural nets
Del Debbio et al., EPJC 85 (2025) 2	Gaussian processes
Izubuchi et al., PRD 100 (2019) 034516 Cichy, Del Debbio & Giani, JHEP 2019 (2019) 137	Global fits

Model data and closure tests

Reconstruction methods first systematically studied in Karpie et al., JHEP 04 (2019) 057
Generated two PDF models and analysed advanced reconstruction methods







Gaussian process regression

Applied to x-dependent distributions in
Alexandrou et al., PRD 102 (2020) 094508
Dutrieux et al., PRD 111 (2024) 034515
Cahuana Medrano et al., JHEP 04 (2026) 182

Non-parametric approach to generating flexible models that implement desired constraints

- can be interpreted as a neural network with a single wide hidden layer

Gaussian process: set of random variables that follows a joint multivariate Gaussian distribution

Goal of regression is to find mean and covariance of the distribution such that the mean is an estimate of data and covariance describes deviations

Typically implemented via Bayesian inference

- requires choice of prior covariance function and corresponding hyperparameters

Gaussian process regression

Applied to x-dependent distributions in
Alexandrou et al., PRD 102 (2020) 094508
Dutrieux et al., PRD 111 (2024) 034515
Cahuana Medrano et al.,JHEP 04 (2026) 182

Non-parametric approach to generating flexible models that implement desired constraints

Can be interpreted as an implementation of neural networks with a single wide hidden layer

Gaussian process:

- collection of random variables $f(z)$, with continuous real index z, that follows a joint multivariate Gaussian distribution with mean function $\mu(z)$ and covariance function $k(z, z_0)$
- for points $z_1 \ldots z_n$, the distribution of $f(z)$ is given by

$$p(f_1, \cdots, f_n) = \det(2\pi K)^{-1/2} e^{-1/2 \sum_{ij} (f_i - \mu_i) K^{-1}_{ij} (f_j - \mu_j)}$$

where

$$f_i = f(z_i) \qquad \mu_i = \mu(z_i) \qquad K_{ij} = k(z_i, z_j)$$

Goal: find $\mu(z)$ and $k(z, z_0)$ such that $\mu(z)$ is an estimate of matrix elements $h(z)$ and $k(z, z_0)$ describes deviations of $\mu(z)$ from $h(z)$

Bayesian inference

Applied to x-dependent distributions in
Alexandrou et al., PRD 102 (2020) 094508
Dutrieux et al., PRD 111 (2024) 034515
Cahuana Medrano et al.,JHEP 04 (2026) 182

Regression (determination of $\mu(z)$) typically undertaken with Bayesian inference

- requires specification of a prior Gaussian process, specified by $\mu_P(z)$ and $k_P(z, z_0)$
- prior Gaussian process adapted to data using Bayes' theorem
- posterior Gaussian process can be calculated analytically

$$\mu(z) = \mu_P(z) + \sum_{ij} k_P(z, z_i) \left[K_{ij} + (\Delta h_i)^2 \delta_{ij}\right]^{-1} (h_j - \mu_P(z_j))$$

Choice of prior covariance and hyperparameters is critical to avoid over/underfitting

GPR as Bayesian prior

data likelihood given by correlated likelihood with data

$$P(M|f_i^q) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\left[-\frac{1}{2}(M_k - B\cdot f^q)^T \Sigma^{-1}(M_k - B\cdot f^q)\right]$$

In the case relevant to GPDs (or PDFs), data take the form

$\mathcal{M}(\nu,\overline{\nu},t,z^2) = B(\nu,\overline{\nu}) \otimes f^q(t,z^2)$

where $B(\nu,\overline{\nu})$ is a linear operator and we wish to infer $f^q(t,z^2)$

vector index *i* labels vertices of fine mesh for reconstruction

Assuming a prior given by a multivariate normal distribution

$I = \mathcal{N}(g_i, K)$

Bayes' theorem gives the posterior probability distribution

Prior distribution density

$$P(f_i^q|I) = \frac{1}{\sqrt{\det(2\pi K)}} \exp\left[-\frac{1}{2}(f_i - g_i)^T K^{-1}(f_i - g_i)\right]$$

$$P(f_i^q|M,I) = \frac{\boxed{P(M|\boxed{f_i^q})}\boxed{P(f_i^q|I)}}{\boxed{P(M|I)}}$$

evidence term

given the data (also a multivariate normal distribution)

$$M = \mathcal{N}(\boxed{M_k = \mathcal{M}(\nu_k,\overline{\nu}_k,t_k,z_k^2)}, \boxed{\Sigma})$$

covariance of lattice data

discrete set of lattice data at specific kinematic points (labelled by k)