

Proton's isovector PDF with updated analysis of large-momentum lattice data

Xiangdong Ji, Yushan Su

[arXiv:2606.22224](https://arxiv.org/abs/2606.22224)

Outline

Benchmark of LaMET

Proton unpolarized isovector PDF as a benchmark to evaluate LaMET's x -dependence calculation;
Previous works: lattice artifacts and less-advanced theoretical analysis lead to inconsistent results among lattice groups and many deviate from phenomenological PDFs.

Reanalysis of lattice datasets

Incorporating all the latest theoretical tools;
Mitigating lattice artifacts empirically.

LaMET results compared with global fit

The variations among lattice groups are dramatically reduced;
The physical-limit result is consistent with global fit PDF within $\sim 1\sigma$.

Outline

Benchmark of LaMET

Proton unpolarized isovector PDF as a benchmark to evaluate LaMET's x -dependence calculation;
Previous works: lattice artifacts and less-advanced theoretical analysis lead to inconsistent results among lattice groups and many deviate from phenomenological PDFs.

Reanalysis of lattice datasets

Incorporate advanced theoretical tools;
Mitigation of lattice artifacts empirically.

LaMET results compared with global fit

The variations among lattice groups are dramatically reduced;
The physical-limit result is consistent with global fit PDF within $\sim 1\sigma$.

EFT and high-energy collision

- An EFT: retains specific dofs from the full theory, while integrating out the others.
- QCD factorization viewed as EFT

- A field can be decomposed

$$\text{e.g. } \psi = \psi_{\text{high}} + \psi_{\text{low}} \quad \text{Ji:2024oka}$$

- Cross sections are factorized into UV (high energy) and IR (low energy) parts:

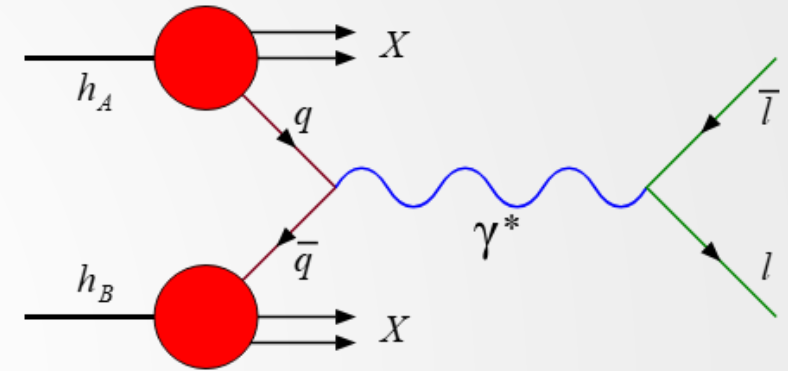
$$\sigma = \text{UV} * \text{IR} + \dots$$

UV: experimental probes&final state interactions, perturbatively calculable

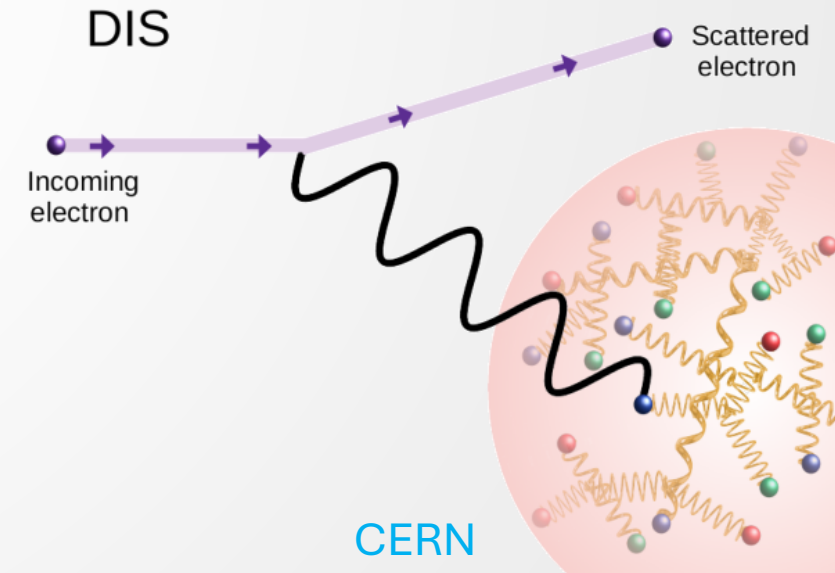
IR: internal hadronic structures, non-perturbative physics

...: power corrections suppressed in the high energy limit

- How to formulate the low energy EFT?



Wikipedia: Drell-Yan process



LaMET for x -dependence calculation of PDF

5

- Collinear modes along the light-cone

([Collins_2023](#), [Stewart:2013SCETnotes](#), [Becher:2014oda](#)): parton distribution function (PDF)

$$f(x, \mu) = \int \frac{d\xi^-}{4\pi} e^{-i\xi^- P^+ x} \langle P | \bar{\psi}(\xi^-) U(\xi^-, 0) \gamma^+ \psi(0) | P \rangle$$

Beauty: momentum density interpretation

Issue: cannot be simulated on a Euclidean lattice

- Collinear modes selected by the hadron velocity

([Ji:2013dva](#), [Ji:2014gla](#), [Ji:2020byp](#), [Ji:2022ezo](#), [Ji:2024oka](#)): longitudinal momentum distribution (LMD)

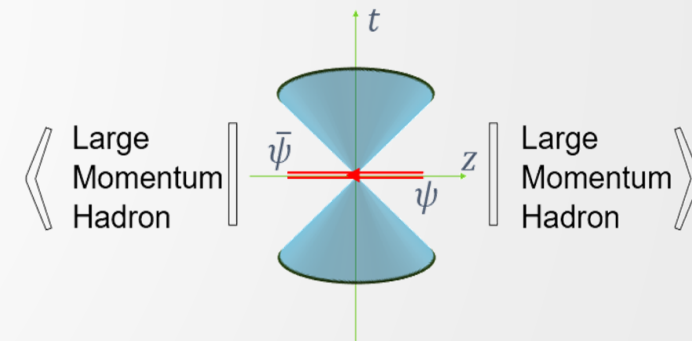
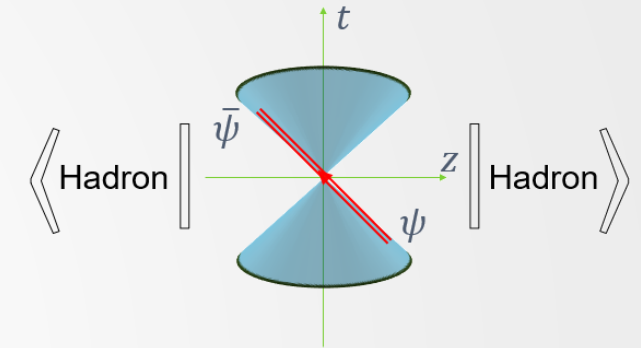
$$\tilde{f}(y, P^z) = \int \frac{dz}{4\pi} e^{i z P^z y} \langle P | \bar{\psi}(z) \gamma^t U(z, 0) \psi(0) | P \rangle$$

measurable in lattice QCD

- Large momentum expansion ([Ji:2024oka](#))

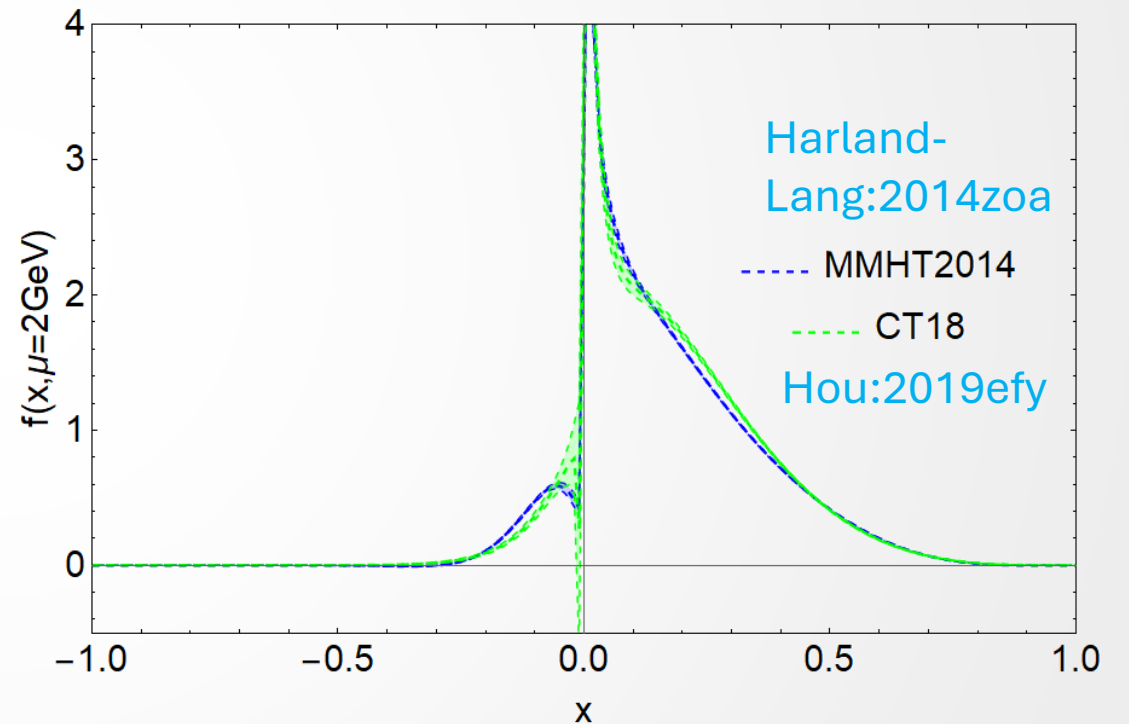
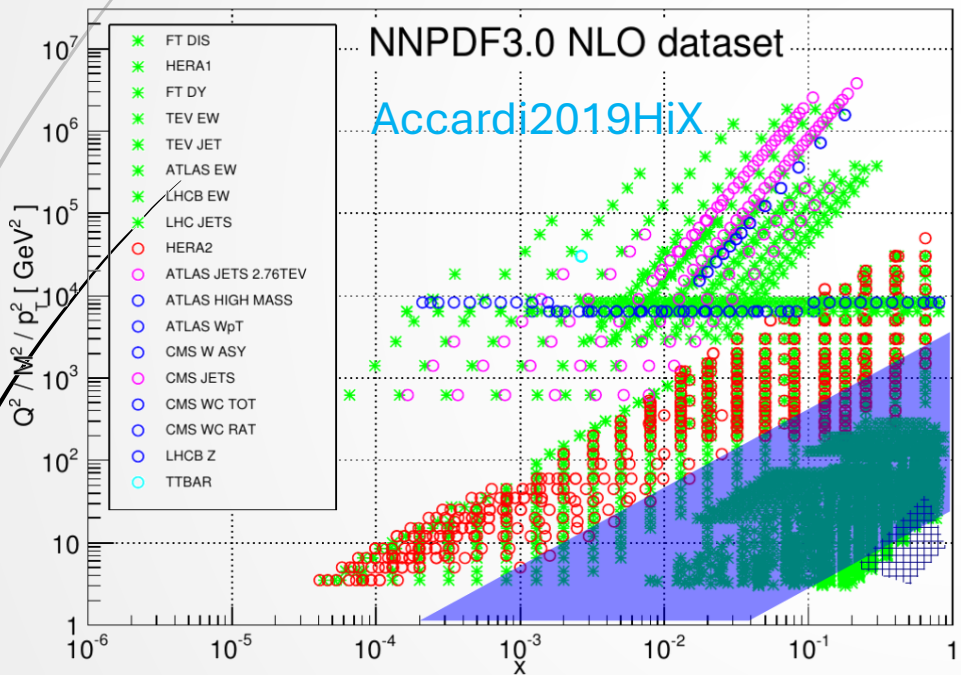
$$f(x, \mu) = \int \frac{dy}{|y|} C\left(\frac{x}{y}, \frac{\mu}{|y|P^z}\right) \tilde{f}(y, P^z) + O\left(\frac{\Lambda_{\text{QCD}}^2}{x^2 P_z^2}, \frac{\Lambda_{\text{QCD}}^2}{(1-x)^2 P_z^2}\right)$$

which allows a **direct x -dependence calculation of PDF with controlled precision**



Benchmark: proton unpolarized isovector PDF

- Measured in high precision from experiments: numerous datasets; no mixing with gluon channel



- Convenient for lattice simulation: no quark-line disconnected diagrams
- A natural benchmark to evaluate LaMET's performance on x -dependence

Previous lattice datasets

LaMET, gauge link (focus of this talk)

Works	ETMC18(Alexandrou:2018pbm) & ETMC19 (Alexandrou:2019lfo)	ETMC20 (Alexandrou:2020qtt)	LP3 (Chen:2018xof)	LPC18(LatticeParton:2018gjr)	CLQCD24(Chen:2024rgi)
Ensemble	ETMC	ETMC	MILC (a09m130)	CLS	CLQCD
a (fm)	0.0938	0.0934, 0.0820, 0.0644	0.09	0.085, 0.064	0.105
m_π (MeV)	130.4	~370	135	~350	~290, 940
P^z (GeV)	0.83, 1.11, 1.38	1.66, 1.89, 1.80	2.2, 2.6, 3.0	0~3.2	1.5~2.5
Works	JLab20(Joo:2020spy)	BNL/MSU20(Fan:2020nzz)	MSULat20(Lin:2020fsj)	ANL/BNL22(Gao:2022uhg)	ANL/BNL26(Gao:2026hix)
Ensemble	JLab/W&M	MILC (a04)	MILC	HotQCD	HotQCD
a (fm)	~0.09	0.042	0.12, 0.09, 0.06	0.076	0.0601
m_π (MeV)	358, 278, 172	310	~310, 230, 138	140	300
P^z (GeV)	Up to ~2	1.84, 2.31	up to 2.2, 2.6, 3.1	0, 0.25, 1.02, 1.53	2.43, 3.04

Pseudo PDF

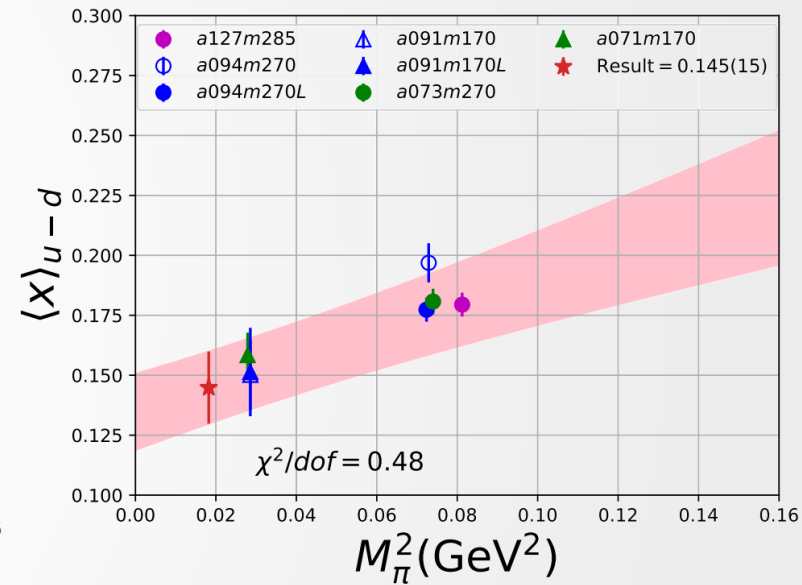
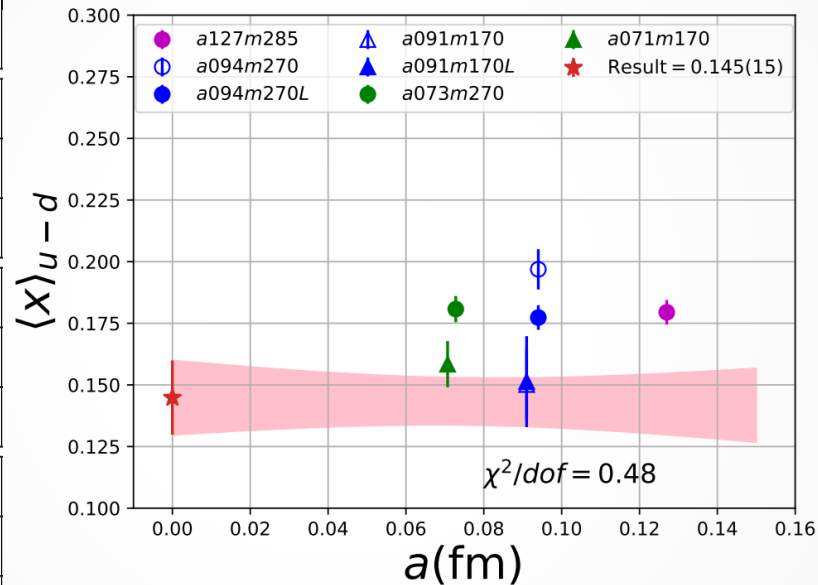
LaMET, Coulomb gauge. Efforts in this direction in Qi Shi's and Joshua Lin's talks

Lattice artifacts

8

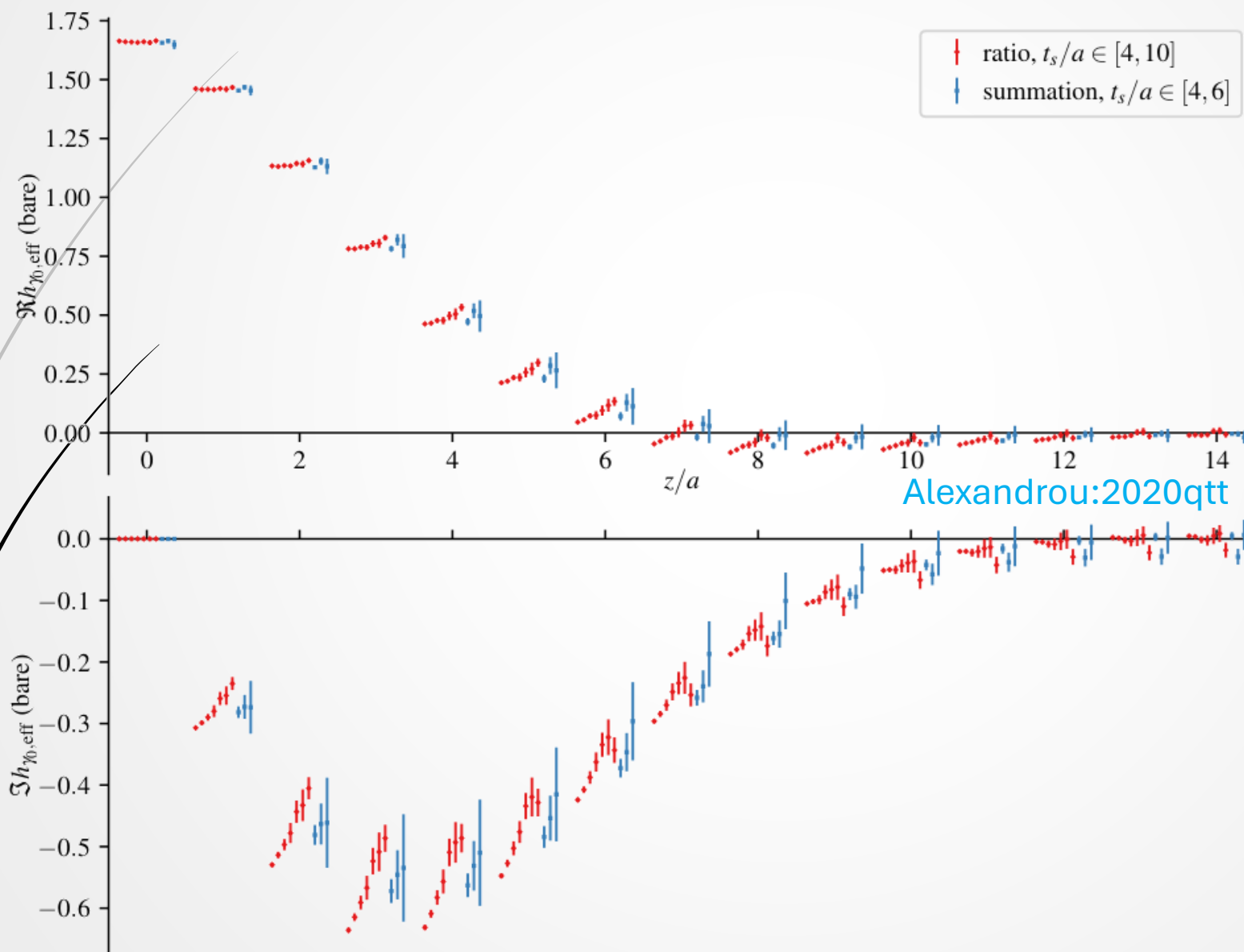
- The first moment calculation (zero momentum) is sensitive to excited state contaminations, discretization effects, and unphysical pion mass [Mondal:2020ela](#)

strategy	Renorm Method	$\langle x \rangle_{u-d}$	
		CC	CCFV
$\{4^{N\pi}, 3^*\}$	A	0.160(13)	0.145(15)
$\{4^{N\pi}, 3^*\}$	B	0.160(12)	0.145(15)
$\{4^{N\pi}, 3^*\}$	Final	0.160(13)	0.145(15)
$\{4, 3^*\}$	A	0.170(15)	0.161(16)
$\{4, 3^*\}$	B	0.168(14)	0.159(16)
$\{4, 3^*\}$	Final	0.169(15)	0.160(16)
$\{4, 2^{\text{free}}\}$	A	0.209(11)	0.193(12)
$\{4, 2^{\text{free}}\}$	B	0.206(11)	0.190(12)
$\{4, 2^{\text{free}}\}$	Final	0.208(11)	0.192(12)



- Compared to zero momentum, these effects in large momentum matrix elements are expected to be more difficult to control.

Excited state contaminations in quasi-PDF matrix elements



$$h_{\Gamma, \text{eff}}^{\text{ratio}}(z; t_s) \equiv \frac{C_{3\text{pt}}^{\Gamma, z}(\frac{t_s}{2}, t_s)}{C_{2\text{pt}}(t_s)} = h_{\Gamma}(z) + O(e^{-\Delta E t_s/2}),$$

$$h_{\Gamma, \text{eff}}^{\text{summ}}(z; t_s) \equiv \frac{S_{\Gamma, z}(t_s + a) - S_{\Gamma, z}(t_s)}{a} = h_{\Gamma}(z) + O(e^{-\Delta E t_s}),$$

where $S_{\Gamma, z}(t_s) \equiv a \sum_{\tau/a=1}^{t_s/a-1} \frac{C_{3\text{pt}}^{\Gamma, z}(\tau, t_s)}{C_{2\text{pt}}(t_s)}$

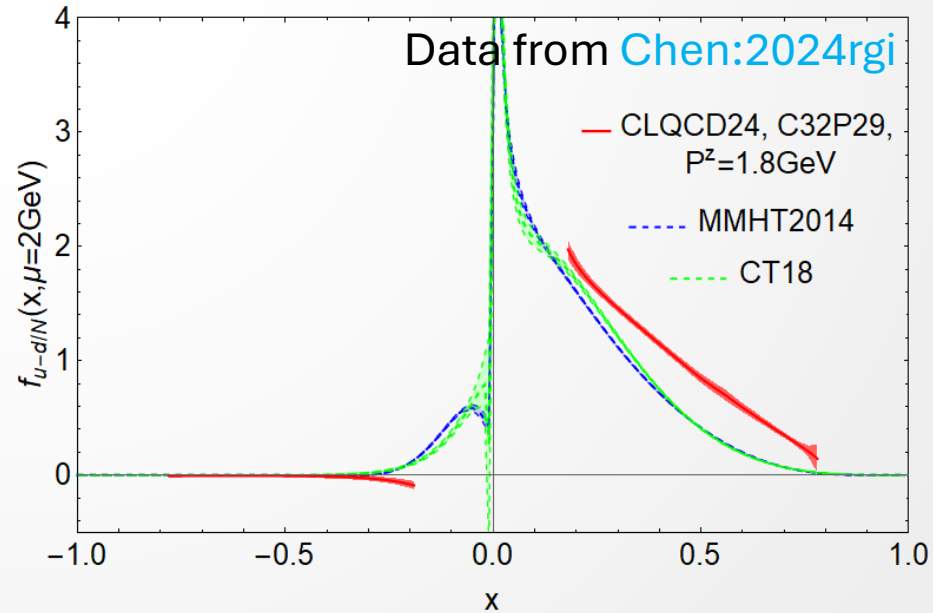
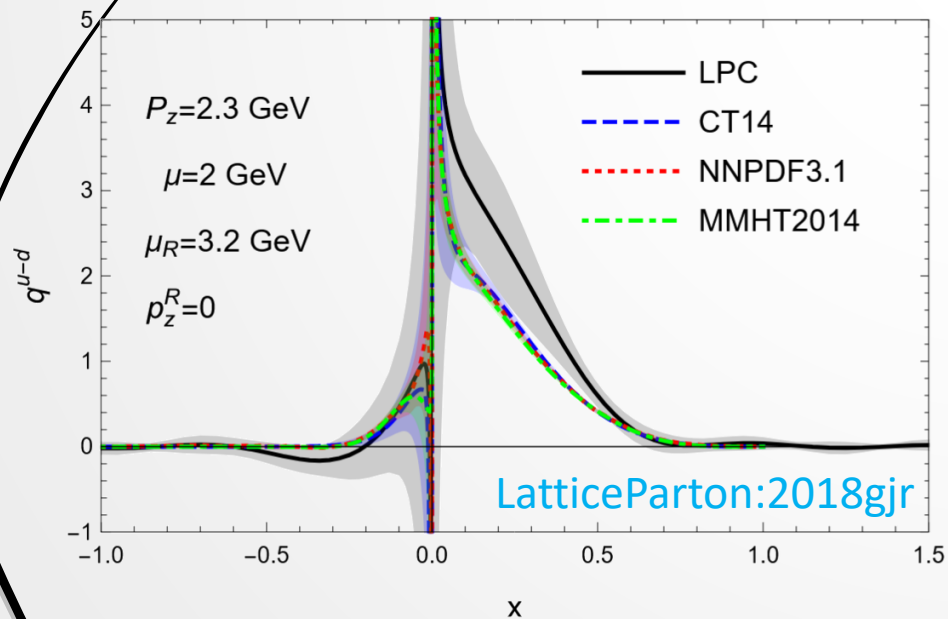
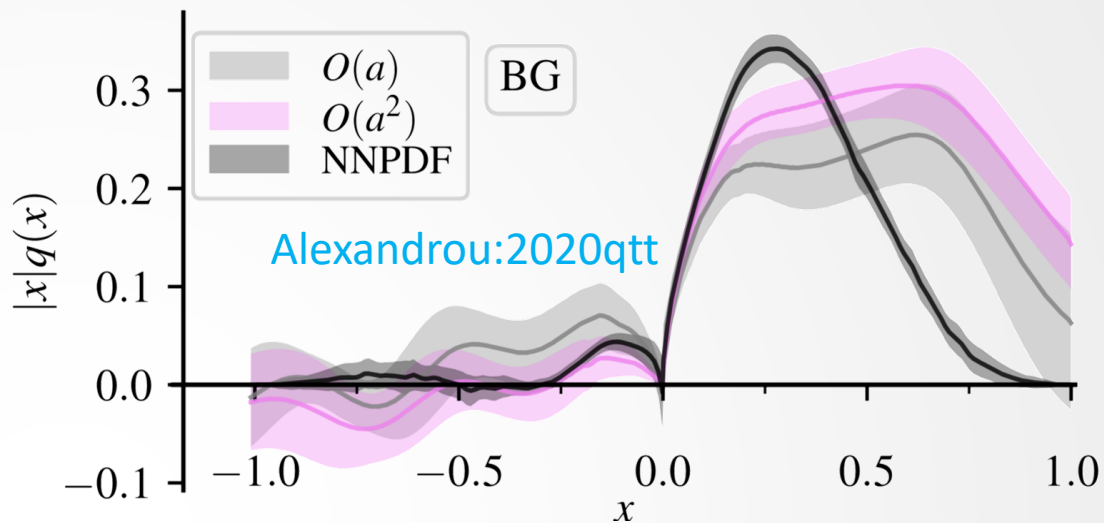
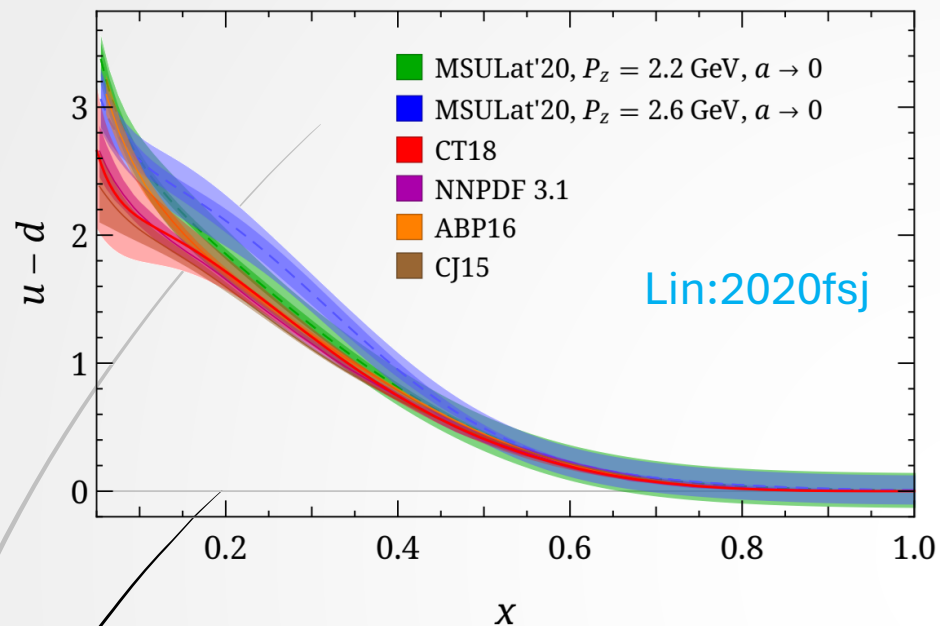
Less advanced theoretical analysis

10

	ETMC18,19	ETMC20	LP3	LPC18
Hybrid renormalization (Ji:2020brr)	No	No	No	No
Latest large distance formulas (Ji:2026vir)	No	No	No	No
Perturbative resummations (Gao:2021hxl , Su:2022f iu , Zhang:2023bxs , Ji:20 23pba , Liu:2023onm , Ji: 2024hit)	No	No	No	No
	CLQCD24	BNL/MSU20	MSULat20	ANL/BNL22
Hybrid renormalization	Yes	No	No	Yes
Latest large distance formulas	No	No	No	No
Perturbative resummations	RGR+LRR No TR	No	No	No

Various shapes in x -space

11



Outline

Benchmark of LaMET

Proton unpolarized isovector PDF as a benchmark to evaluate LaMET's x -dependence calculation;
Previous works: lattice artifacts and less-advanced theoretical analysis lead to inconsistent results among lattice groups and many deviated from the experimental PDFs.

Reanalysis of lattice datasets

Incorporating all the latest theoretical tools;
Mitigating lattice artifacts empirically.

LaMET results compared with global fit

The variations among lattice groups are dramatically reduced;
The physical-limit result is consistent with global fit PDF within $\sim 1\sigma$.

Selected datasets

Works	MSULat20(Lin:2020 fsj)	ETMC20 (Alexandrou:2020q tt)	LPC18(LatticeParton:2018gjr)	CLQCD24(Chen:2024rgi)
Ensemble	MILC	ETMC (B55)	CLS	CLQCD
a (fm)	0.12, 0.09, 0.06	0.0820	0.085	0.105
m_π (MeV)	~310, 230, 138	~370	350	~290
P^z (GeV)	2.2	1.89	2.3	1.85

They are selected under the following considerations:

- large momentum (at least $\sim 2\text{GeV}$, corresponding to $v \sim 0.9c$);
- statistical precision;
- data accessibility.

Hybrid renormalization

Ji:2020brr

14

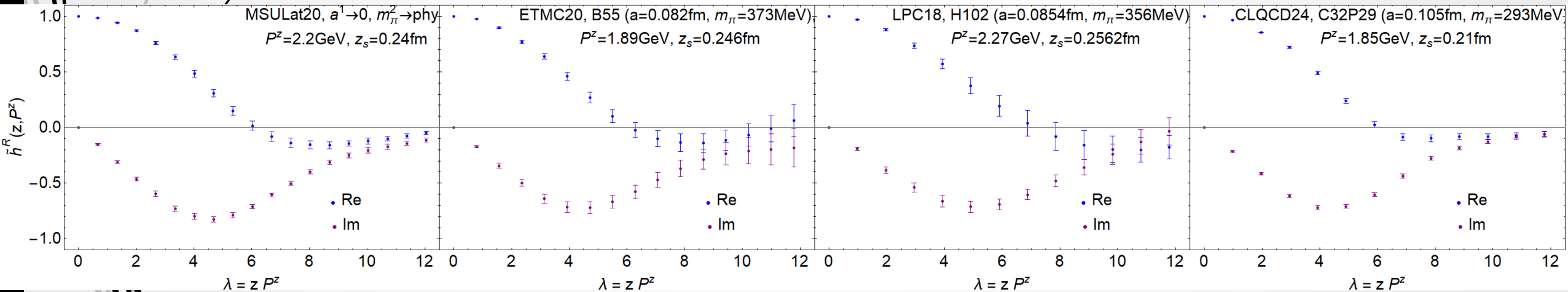
➤ Basic idea: an almost minimal UV subtraction while mitigating short-distance discrepancies between lattice and continuum theories.

➤ Formulation:

$$\tilde{h}^R(z, P^z) = \frac{\tilde{h}^B(z, P^z, a)}{\tilde{h}^B(z, 0, a)} \theta[z_s - |z|] + \frac{\tilde{h}^B(z, P^z, a) Z_R(z_s, \mu, a)}{Z_R(z, \mu, a) \tilde{h}^B(z_s, 0, a)} \theta[|z| - z_s]$$

where $Z_R(z, \mu, a)$ is determined using the methods in [LatticePartonLPC:2021gpi](#) and [Zhang:2023bxs](#) using NNLO+RGR+LRR pert.

➤ Hybrid renormalized matrix elements:



Mitigating lattice artifacts

15

- Idea: enforce the first moment $\langle x \rangle$ of lattice data to match the phenomenological result.
- A mitigation formula should impact short distance slope of imaginary part: $1 + i z P^Z \delta y$. To avoid unrealistically large imaginary part at large z , we unitarize it to an exponential:

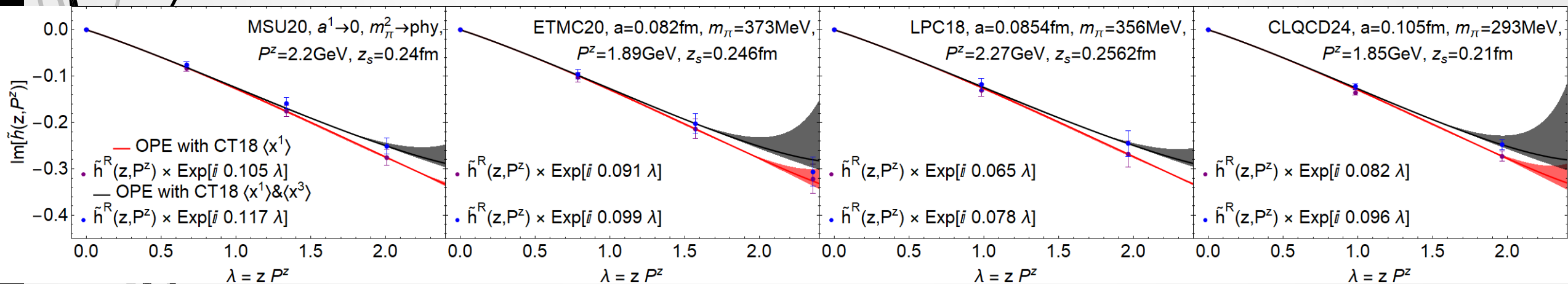
$$\tilde{h}(z, P^Z) = \tilde{h}^R(z, P^Z) e^{i z P^Z \delta y}$$

where δy is determined by fitting the imaginary part of $\tilde{h}(z, P^Z)$ to the short distance OPE

$$1 - i z P^Z \langle x \rangle(\mu) \frac{C_1(z^2 \mu^2)}{C_0(z^2 \mu^2)} + \dots$$

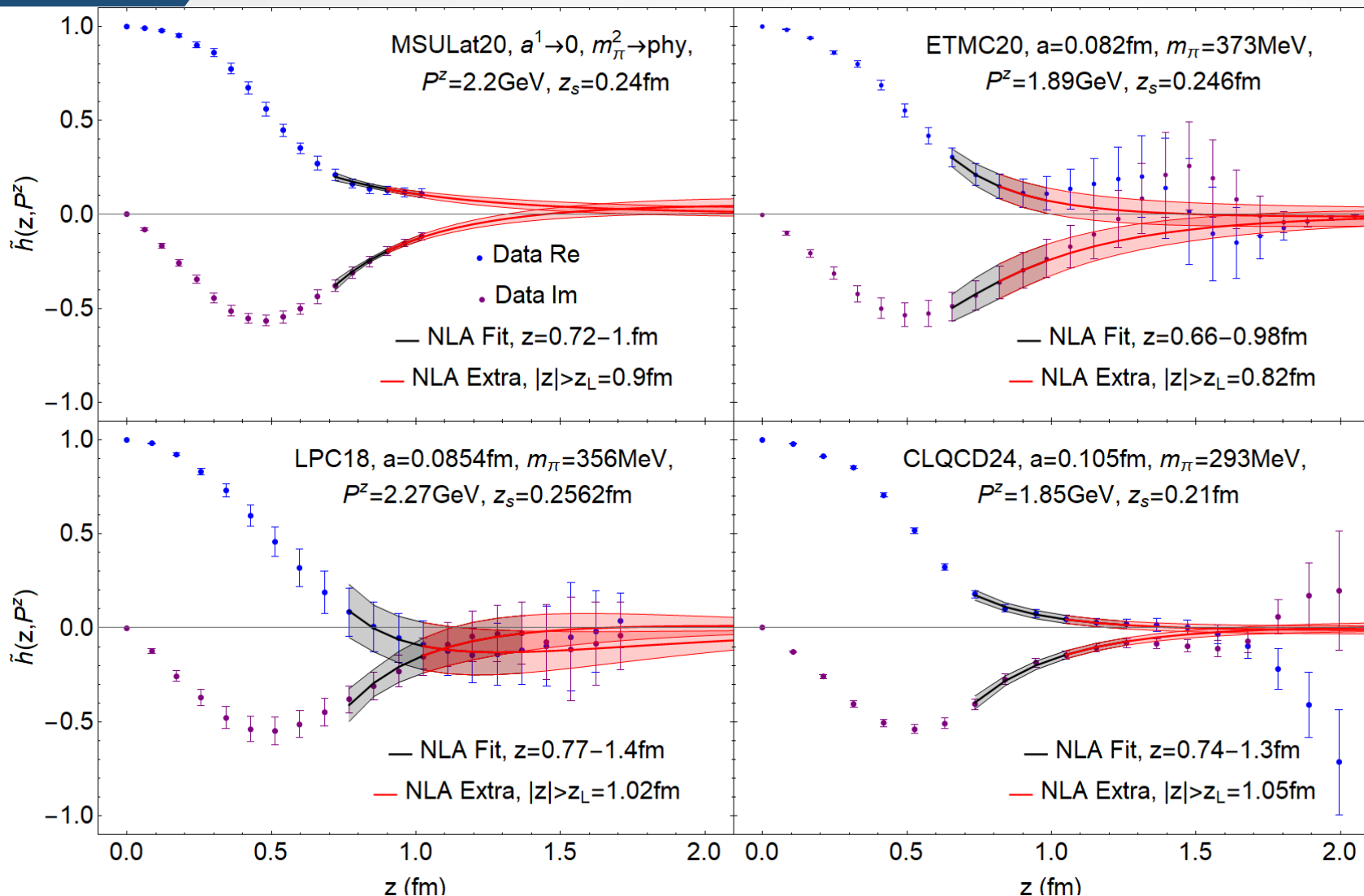
where the first moment is fixed to CT18 value $\langle x \rangle(\mu = 2\text{GeV}) = 0.1562(24)$, and the Wilson coefficients C_i are NNLO+RGR+LRR accuracy.

Results



Large distance extrapolation and Fourier transform

16



- Use large distance asymptotic formulas (Ji:2026vir) as physical constraints:

$$\tilde{h}^{\text{asym}}(z, P^z) = \left[A e^{i \phi \text{sgn}(z)} + \frac{A' e^{i \phi' \text{sgn}(z)}}{|z|} \right] e^{-\Lambda |z|}$$

- Fourier transform

$$\tilde{f}(y, P^z) = P^z \int \frac{dz}{2\pi} e^{i z P^z y} \left[\tilde{h}(z, P^z) \theta(z_L - |z|) + \tilde{h}^{\text{asym}}(z, P^z) \theta(|z| - z_L) \right]$$

Perturbative matching

- Purpose: deal with the UV difference between PDF and LMD

([Xiong:2013bka](#),[Ma:2014jla](#),[Ma:2017pxb](#),[Izubuchi:2018srq](#),[Wang:2019tgg](#),[Ji:2020ect](#)):

$$f(x, \mu) = \int \frac{dy}{|y|} C\left(\frac{x}{y}, \frac{\mu}{|y|P^Z}\right) \tilde{f}(y, P^Z) + O\left(\frac{\Lambda_{\text{QCD}}^2}{x^2 P_Z^2}, \frac{\Lambda_{\text{QCD}}^2}{(1-x)^2 P_Z^2}\right)$$

- The matching kernel $C\left(\frac{x}{y}, \frac{\mu}{|y|P^Z}\right)$ is at NNLO+RGR+LRR+TR accuracy:
 - Next-to-next-to-leading order (NNLO) ([Li:2020xml](#),[Chen:2020ody](#)): perturbative precision
 - DGLAP log resummation (RGR) ([Su:2022fiu](#)): precision control at small x
 - Threshold log resummation (TR) ([Gao:2021hxl](#),[Ji:2023pba](#),[Liu:2023onm](#),[Ji:2024hit](#)): precision control at large x
 - Leading renormalon resummation (LRR) ([Zhang:2023bxs](#)): perturbative convergence, linear power accuracy

Outline

Benchmark of LaMET

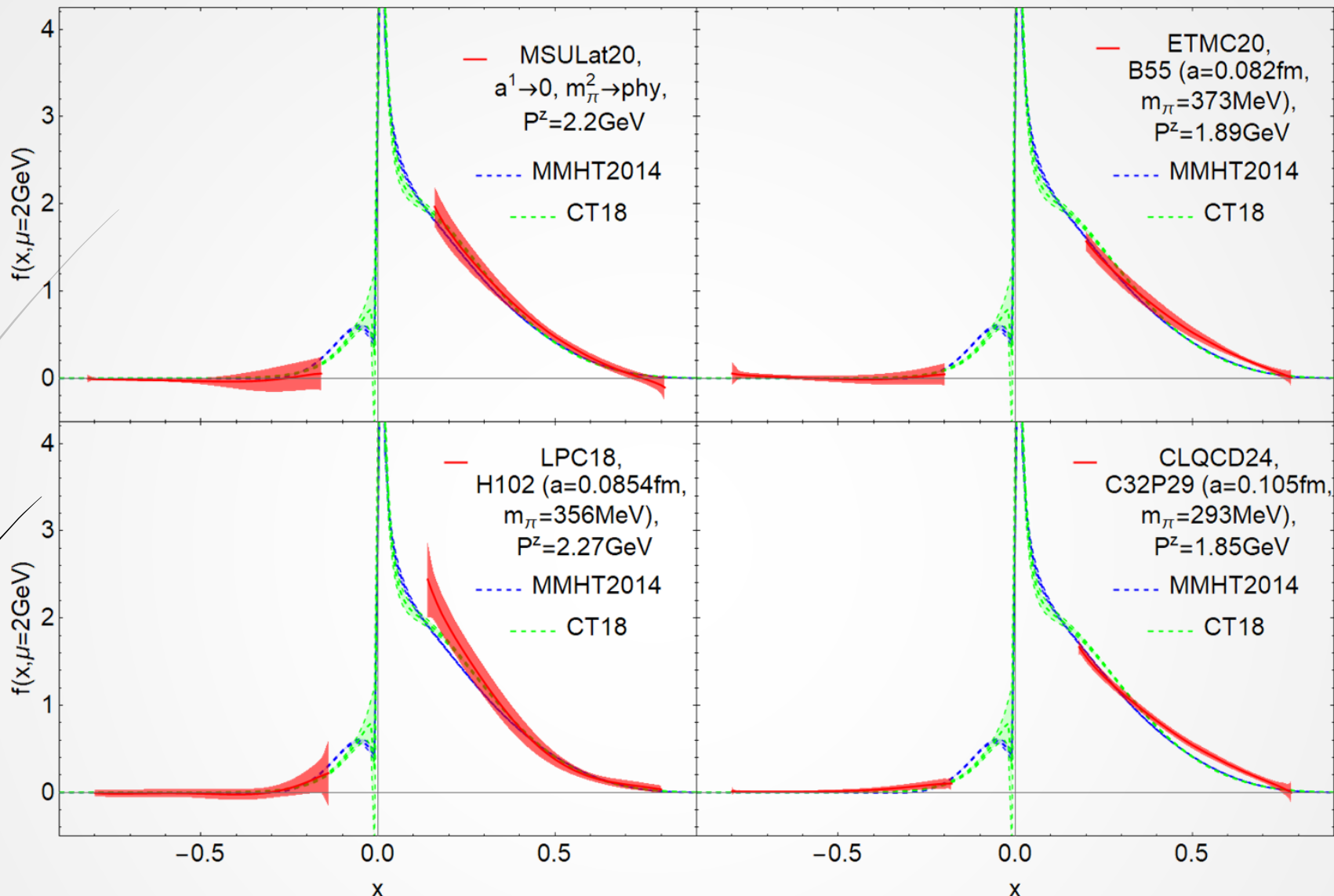
Proton unpolarized isovector PDF as a benchmark to evaluate LaMET's x -dependence calculation;
Previous works: lattice artifacts and less-advanced theoretical analysis lead to inconsistent results among lattice groups and many deviate from experimental PDFs.

Reanalysis of lattice datasets

Incorporation of advanced theoretical tools;
Mitigation of lattice artifacts empirically.

LaMET results compared with global fit

The variations among lattice groups are dramatically reduced;
The physical-limit result is consistent with global fit PDF within $\sim 1\sigma$.



- After implementing a unified analysis framework, the spread among the resulting PDFs is substantially reduced compared with the original analyses.
- The physical-limit lattice PDF is consistent with phenomenological PDFs within $\sim 1\sigma$. The remaining datasets exhibit broadly compatible x -dependence.

Conclusions

- ▶ We present a LaMET calculation of the proton isovector unpolarized PDF using existing lattice QCD datasets, along with theoretical precision-control techniques and an empirical method to remove a large part of lattice artifacts.
- ▶ The disagreement among earlier lattice PDF calculations appears to be largely attributable to analysis methodology and lattice artifacts.
- ▶ The resulting PDFs align with phenomenological determinations, supporting LaMET's potential to achieve quantitatively reliable determinations of the x -dependence of parton distributions.

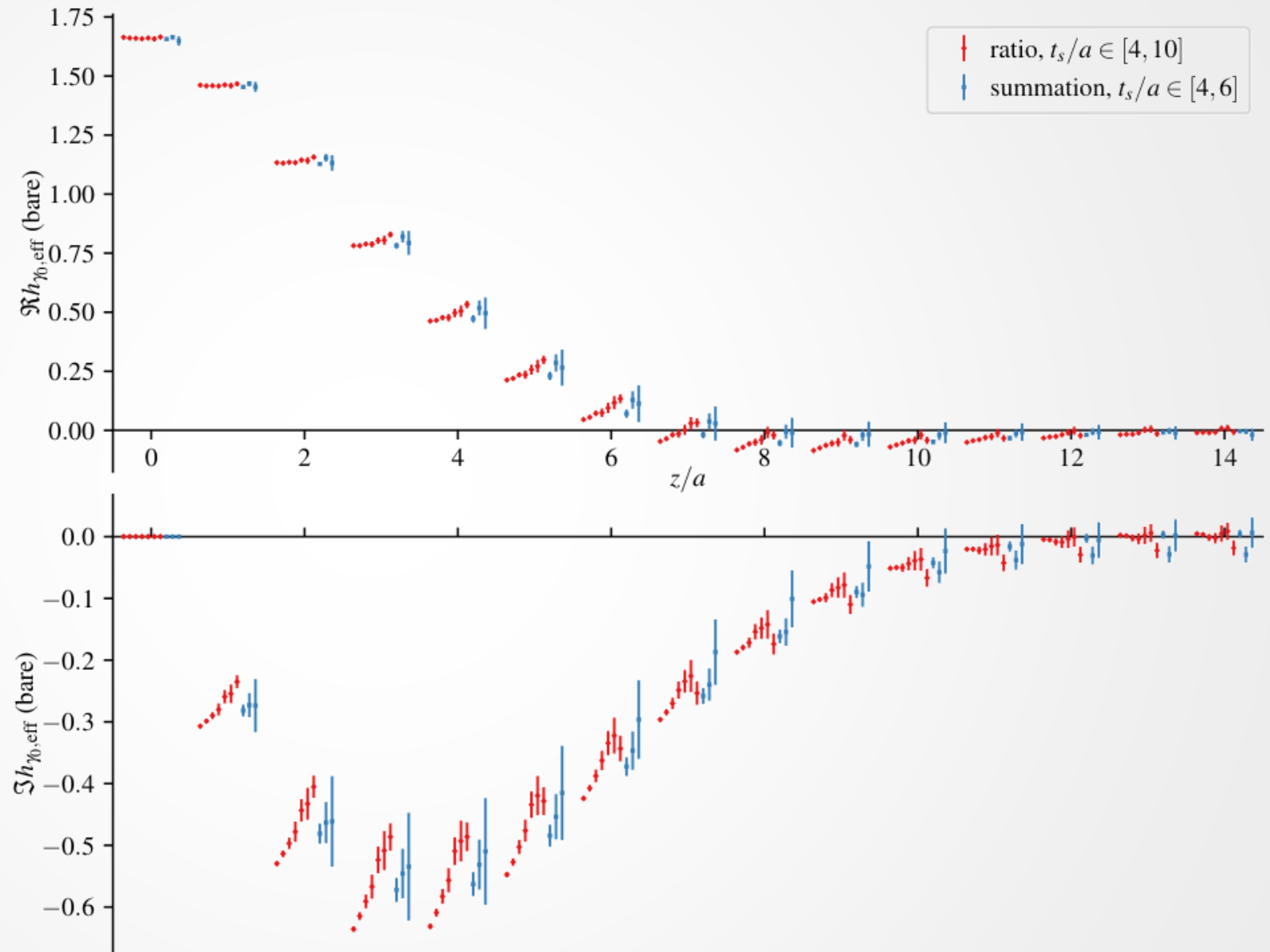
Appendix

Evidence of excited state contamination

$$h_{\Gamma,\text{eff}}^{\text{ratio}}(z; t_s) \equiv \frac{C_{3\text{pt}}^{\Gamma,z}(\frac{t_s}{2}, t_s)}{C_{2\text{pt}}(t_s)}$$
$$= h_{\Gamma}(z) + O(e^{-\Delta E t_s/2}),$$

$$h_{\Gamma,\text{eff}}^{\text{summ}}(z; t_s) \equiv \frac{S_{\Gamma,z}(t_s + a) - S_{\Gamma,z}(t_s)}{a}$$
$$= h_{\Gamma}(z) + O(e^{-\Delta E t_s}),$$

where $S_{\Gamma,z}(t_s) \equiv a \sum_{\tau/a=1}^{t_s/a-1} \frac{C_{3\text{pt}}^{\Gamma,z}(\tau, t_s)}{C_{2\text{pt}}(t_s)}$



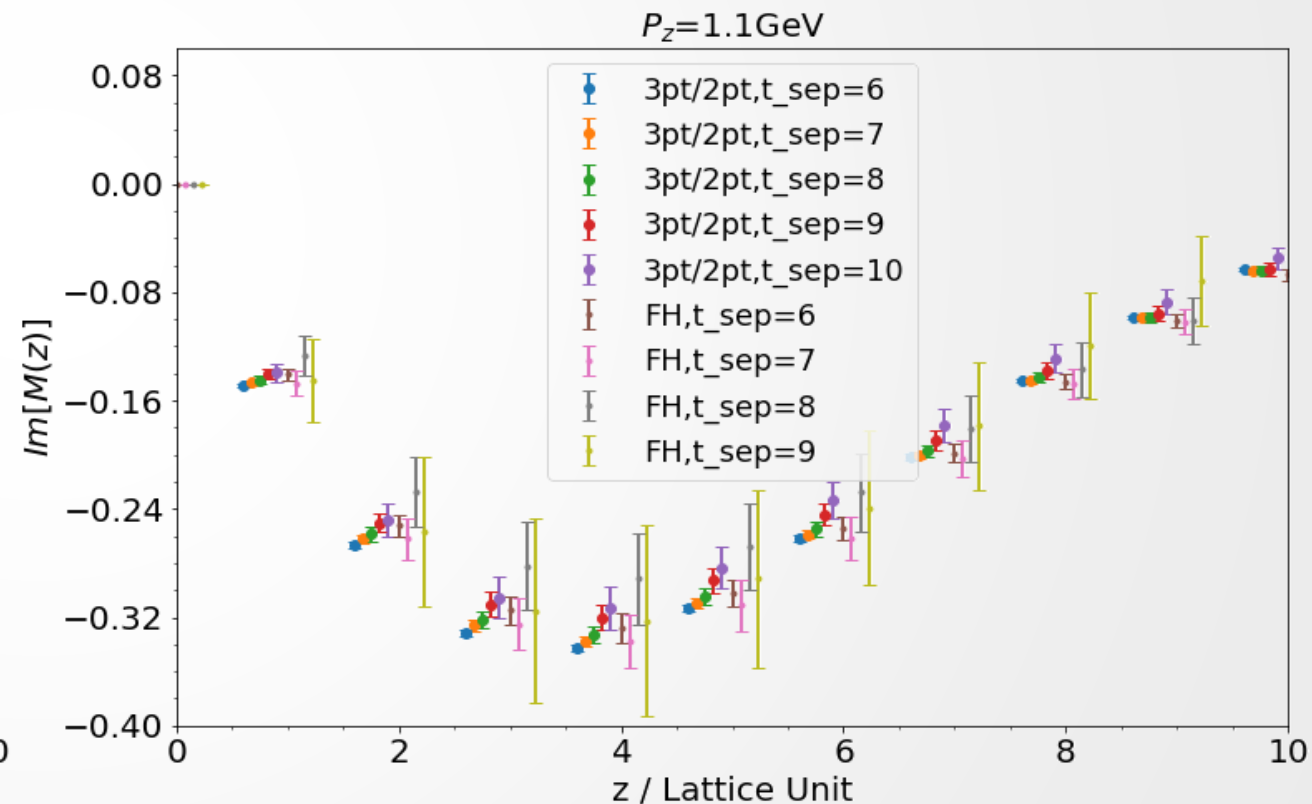
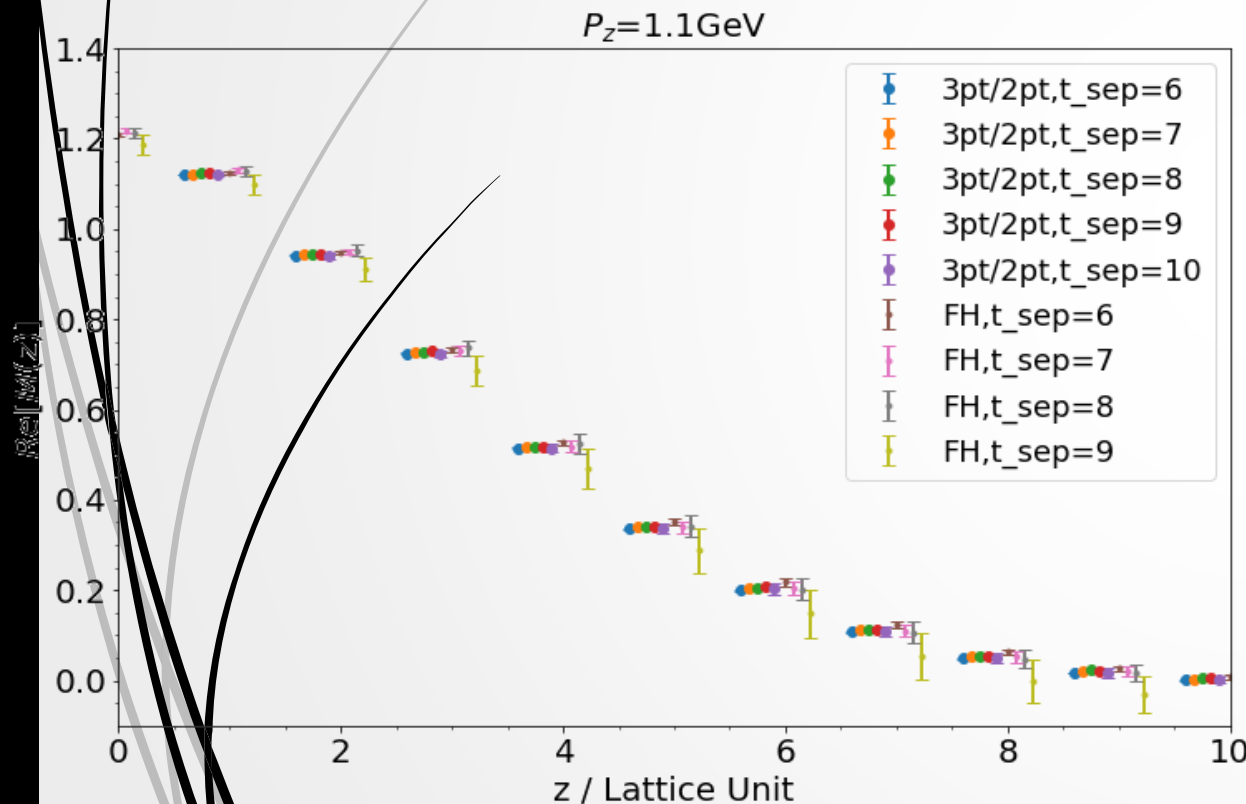
Evidence of excited state contamination

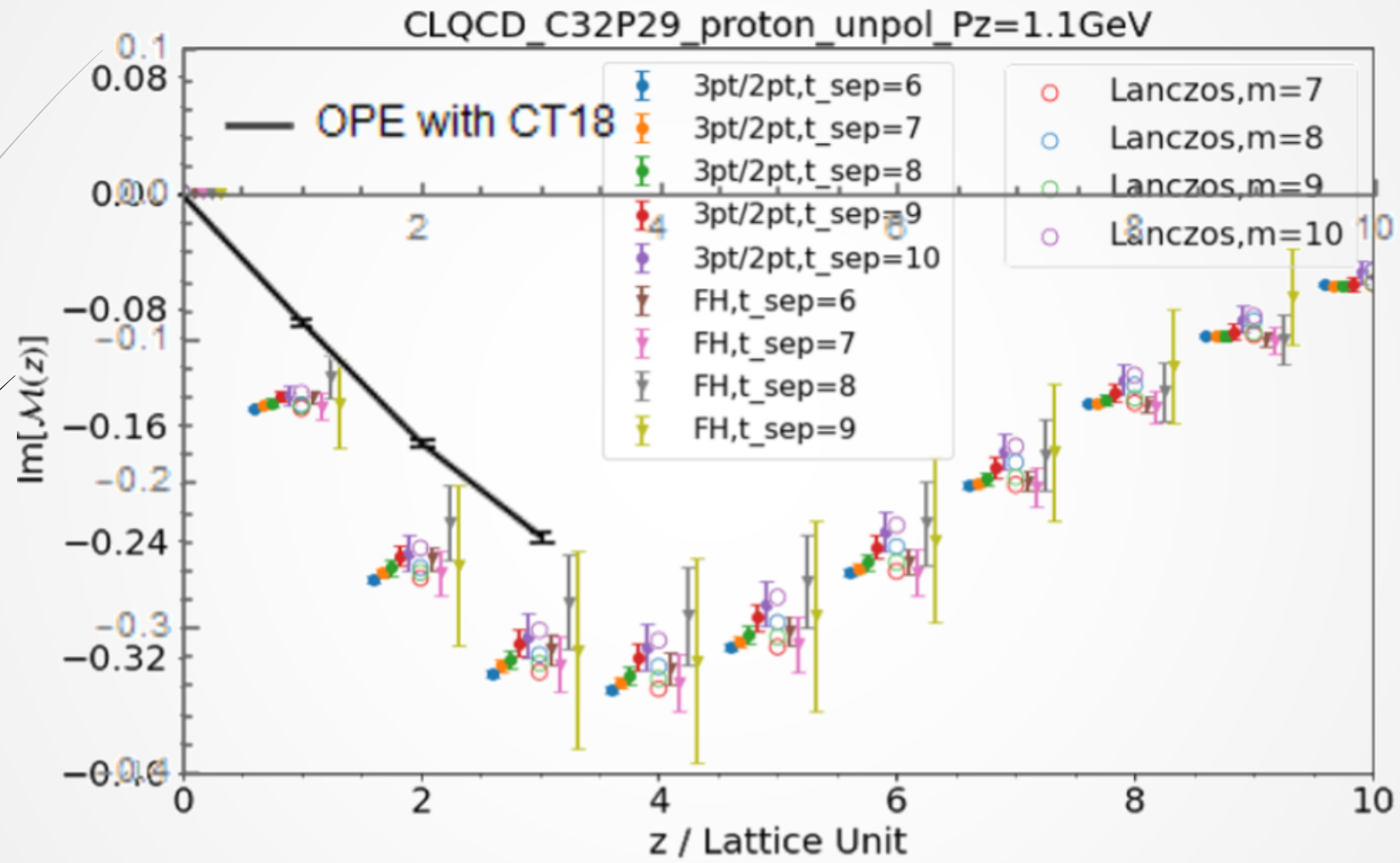
$$R\left[t_{\text{sep}}, t = \frac{t_{\text{sep}}}{2}\right] = \frac{C_3\left[t_{\text{sep}}, t = \frac{t_{\text{sep}}}{2}, z, P_z\right]}{C_2\left[t_{\text{sep}}, P_z\right]}$$

$$S[t_{\text{sep}}, \tau_c] = \sum_{t=\tau_c}^{t_{\text{sep}}-\tau_c} R[t_{\text{sep}}, t]$$

$$FH[t_{\text{sep}}, \tau_c] = S[t_{\text{sep}} + 1, \tau_c] - S[t_{\text{sep}}, \tau_c]$$

$$\tau_c = 2$$





Reanalysis of selected datasets

Works	ETMC18(Alexandrou:2018pbm) & ETMC19 (Alexandrou:2019lfo)	ETMC20 (Alexandrou:2020qtt)	LP3 (Chen:2018xof)	LPC18(LatticeParton:2018gjr)	CLQCD24(Chen:2024rgi)
Ensemble	ETMC	ETMC (B55)	MILC (a09m130)	CLS	CLQCD
a (fm)	0.0938	0.0820	0.09	0.085	0.105
m_π (MeV)	130.4	~370	135	350	~290
P^Z (GeV)	0.83, 1.11, 1.38	1.89	2.2, 2.6, 3.0	2.3	1.85

Works	BNL/MSU20(Fan:2020nzz)	MSU20(Lin:2020fsj)	ANL/BNL22(Gao:2022uhg)
Ensemble	MILC (a04)	MILC	HotQCD
a (fm)	0.042	0.12, 0.09, 0.06	0.076
m_π (MeV)	310	~310, 230, 138	140
P^Z (GeV)	1.84, 2.31	2.2	0, 0.25, 1.02, 1.53

LaMET analysis procedure

- ▶ Lattice data generation of $(u - d)/N$ quasi-PDF matrix element

$$\tilde{h}^B(z, P^z, a) = \langle N(P^z) | [\bar{u}(z)\gamma^t U(z, 0)u(0) - \bar{d}(z)\gamma^t U(z, 0)d(0)] | N(P^z) \rangle$$

- ▶ Hybrid renormalization

$$\tilde{h}^R(z, P^z) = \frac{\tilde{h}^B(z, P^z, a)}{\tilde{h}^B(z, 0, a)} \theta[z_s - |z|] + \frac{\tilde{h}^B(z, P^z, a) Z_R(z_s, \mu, a)}{Z_R(z, \mu, a) \tilde{h}^B(z_s, 0, a)} \theta[|z| - z_s]$$

where $Z_R(z, \mu, a)$ is determined using the methods in [LatticePartonLPC:2021gpi](#) and [Zhang:2023bxs](#) using NNLO+RGR+LRR pert.

- ▶ Calibration of first moment

$$\tilde{h}(z, P^z) = \tilde{h}^R(z, P^z) e^{i z P^z \Gamma}$$

where Γ is determined by fitting imaginary part of $\tilde{h}(z, P^z)$ at $|z| < z_s$ to the short distance OPE with $\langle x \rangle$:

$$\tilde{h}(z, P^z) = 1 - i z P^z \langle x \rangle(\mu) \frac{C_1(z^2 \mu^2)}{C_0(z^2 \mu^2)} + \dots$$

where $C_{\#}$ are NNLO+RGR+LRR Wilson coefficients.

- ▶ Large distance asymptotic analysis. Fit and extrapolate $\tilde{h}(z, P^z)$ to large distance based on [Ji:2026vir](#). Then perform the Fourier transformation:

$$\tilde{f}(y, P^z) = P^z \int \frac{dz}{2\pi} e^{i z P^z y} \tilde{h}(z, P^z)$$

- ▶ Perturbative matching

$$f(x, \mu) = \int \frac{dy}{|y|} C\left(\frac{x}{y}, \frac{\mu}{y P^z}\right) \tilde{f}(y, P^z)$$

where $C\left(\frac{x}{y}, \frac{\mu}{y P^z}\right)$ is the NNLO+RGR+LRR+TR matching kernel in [Ji:2024hit](#). All the LRRs mentioned here are in the same renormalon scheme.

MSU20(Lin:2020fsj) – Conversion to hybrid scheme

Fig.~4 of [Lin:2020fsj](#) shows the RIMOM renormalized matrix element

$$\tilde{h}^{\text{RM}}(z, P^z) = \frac{\tilde{h}^B(z, P^z, a)}{Z^{\text{RM}}(z, a)}$$

to implement LaMET analysis, we need to convert it to hybrid scheme

- According to [LatticePartonLPC:2021gpi](#), the RM factor can be factorized into the renormalization factor and intrinsic nonperturbative physics

$$Z^{\text{RM}}(z, a) = Z_R(z, \mu, a) Z^{\text{RM-inp}}(z, \mu)$$

So does the zero momentum matrix element

$$\tilde{h}^B(z, 0, a) = Z_R(z, \mu, a) \tilde{h}^{B\text{-inp}}(z, \mu)$$

Therefore, the conversion can be implemented as follows

$$\begin{aligned} \tilde{h}^R(z, P^z) &= \frac{\tilde{h}^{\text{RM}}(z, P^z) Z^{\text{RM}}(z, a)}{\tilde{h}^B(z, 0, a)} \theta[z_s - |z|] + \frac{\tilde{h}^{\text{RM}}(z, P^z) Z^{\text{RM}}(z, a) Z_R(z_s, \mu, a)}{Z_R(z, \mu, a) \tilde{h}^B(z_s, 0, a)} \theta[|z| - z_s] \\ &= \frac{\tilde{h}^{\text{RM}}(z, P^z) Z^{\text{RM-inp}}(z, \mu)}{\tilde{h}^{B\text{-inp}}(z, \mu)} \theta[z_s - |z|] + \frac{\tilde{h}^{\text{RM}}(z, P^z) Z^{\text{RM-inp}}(z, \mu)}{\tilde{h}^{B\text{-inp}}(z_s, \mu)} \theta[|z| - z_s] \end{aligned}$$

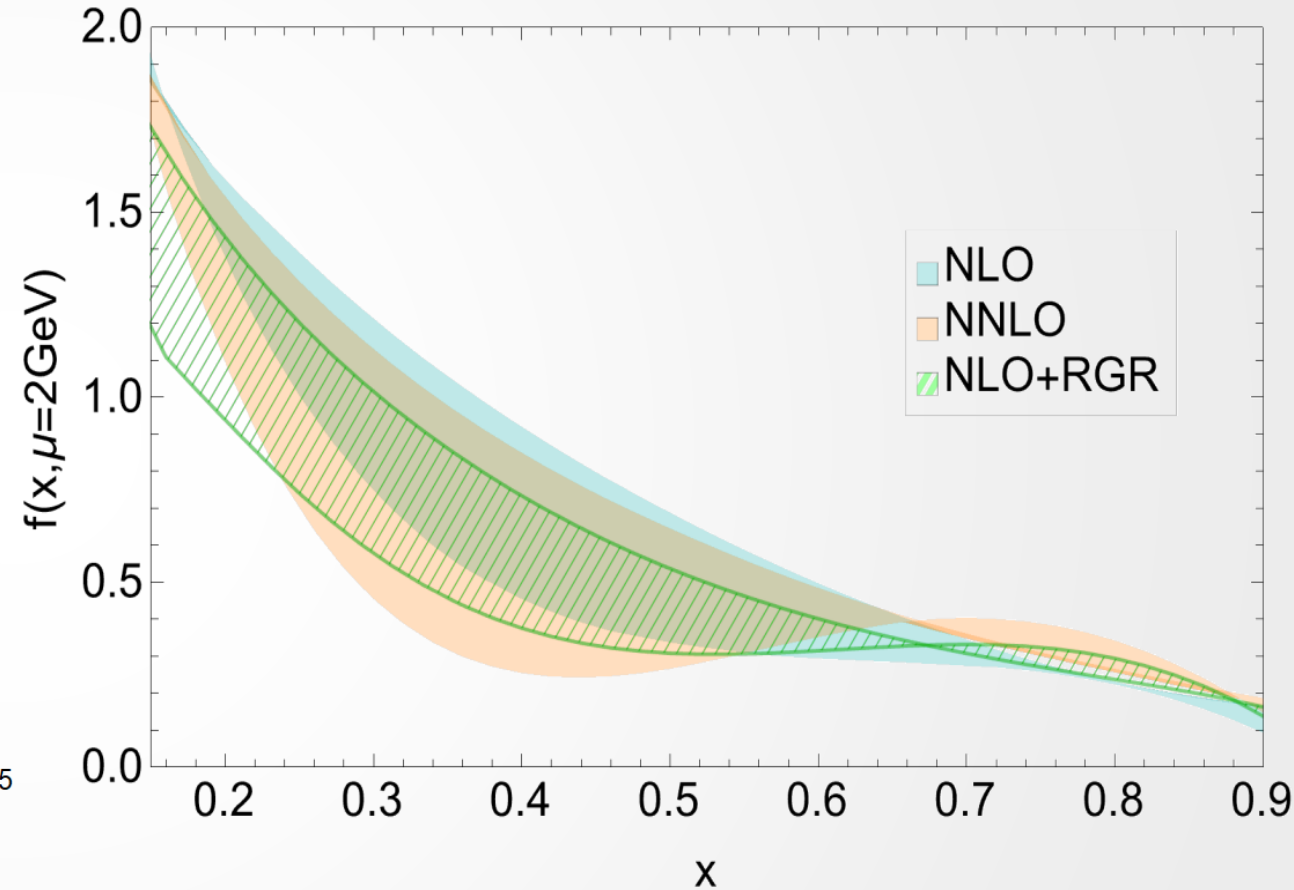
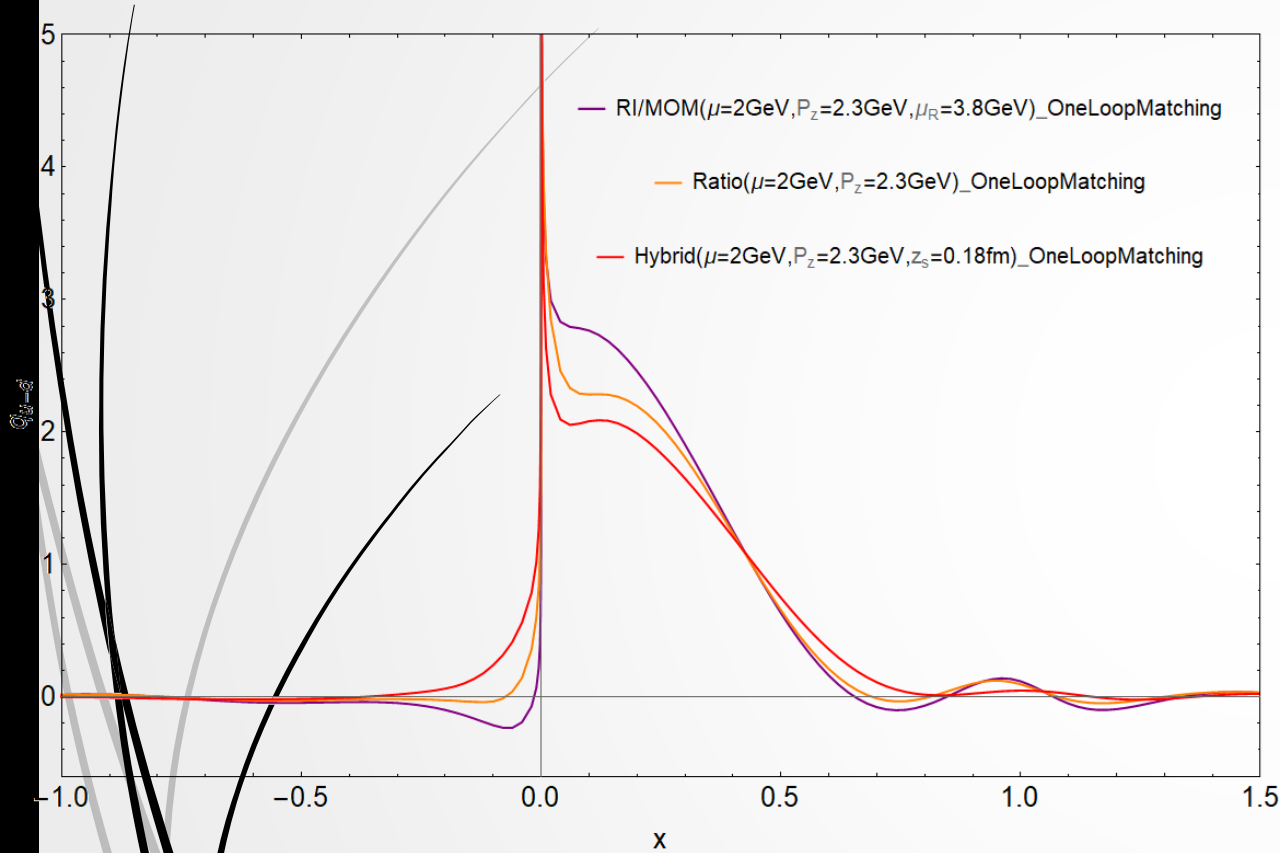
- In practice, $\tilde{h}^{\text{RM}}(z, P^z)$ is digitized from Fig.~4 of [Lin:2020fsj](#), $Z^{\text{RM}}(z, a)$ from Fig.~2 of [Lin:2020fsj](#), and $\tilde{h}^B(z, 0, a)$ from [LatticePartonLPC:2021gpi](#). The $Z^{\text{RM-inp}}(z, \mu)$ and $\tilde{h}^{B\text{-inp}}(z, \mu)$ are extracted from $Z^{\text{RM}}(z, a)$ and $\tilde{h}^B(z, 0, a)$ following the methods in [LatticePartonLPC:2021gpi](#) and [Zhang:2023bxs](#).

Impacts of renormalization methods

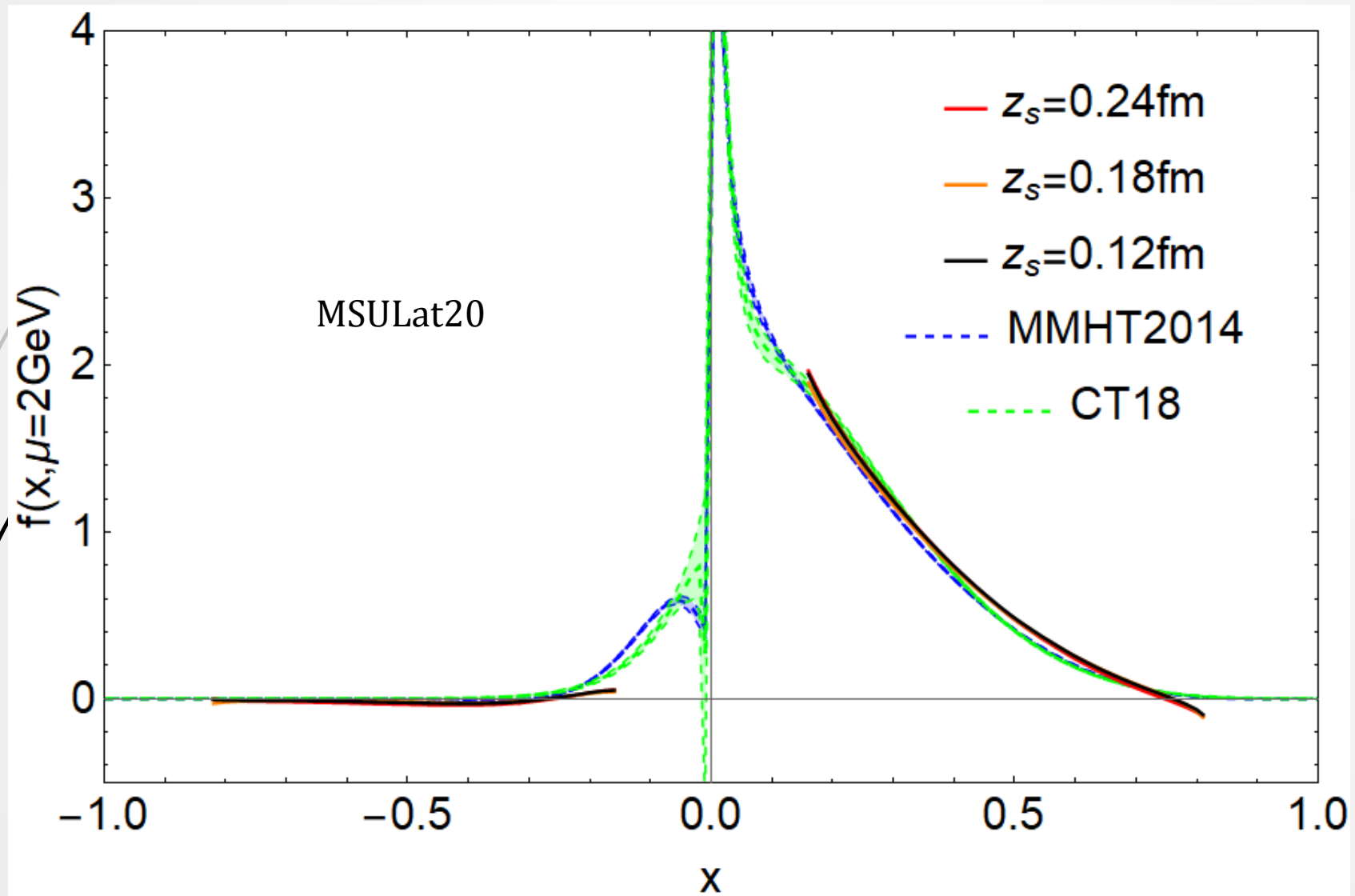
28

➤ Different renormalization methods using LPC18 dataset

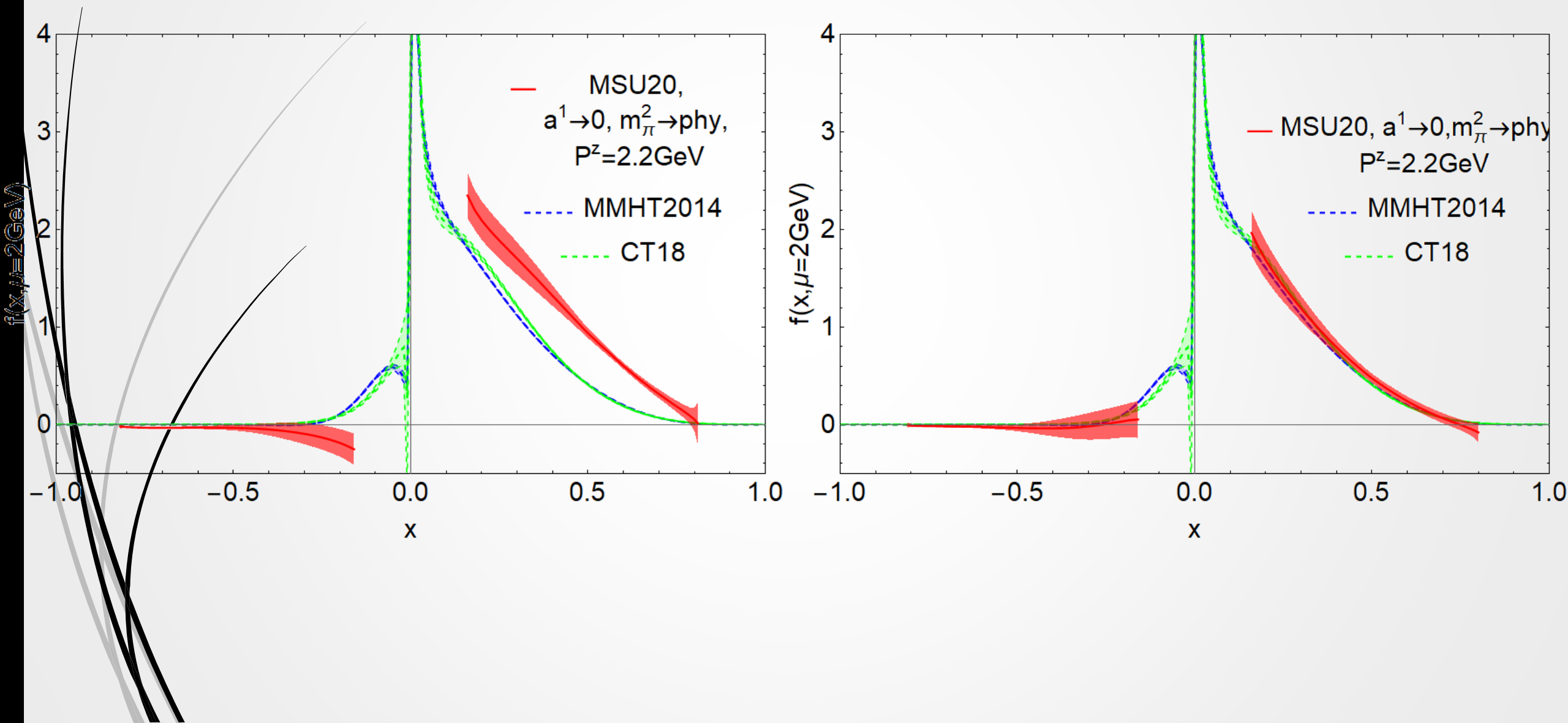
➤ Influence from m_0



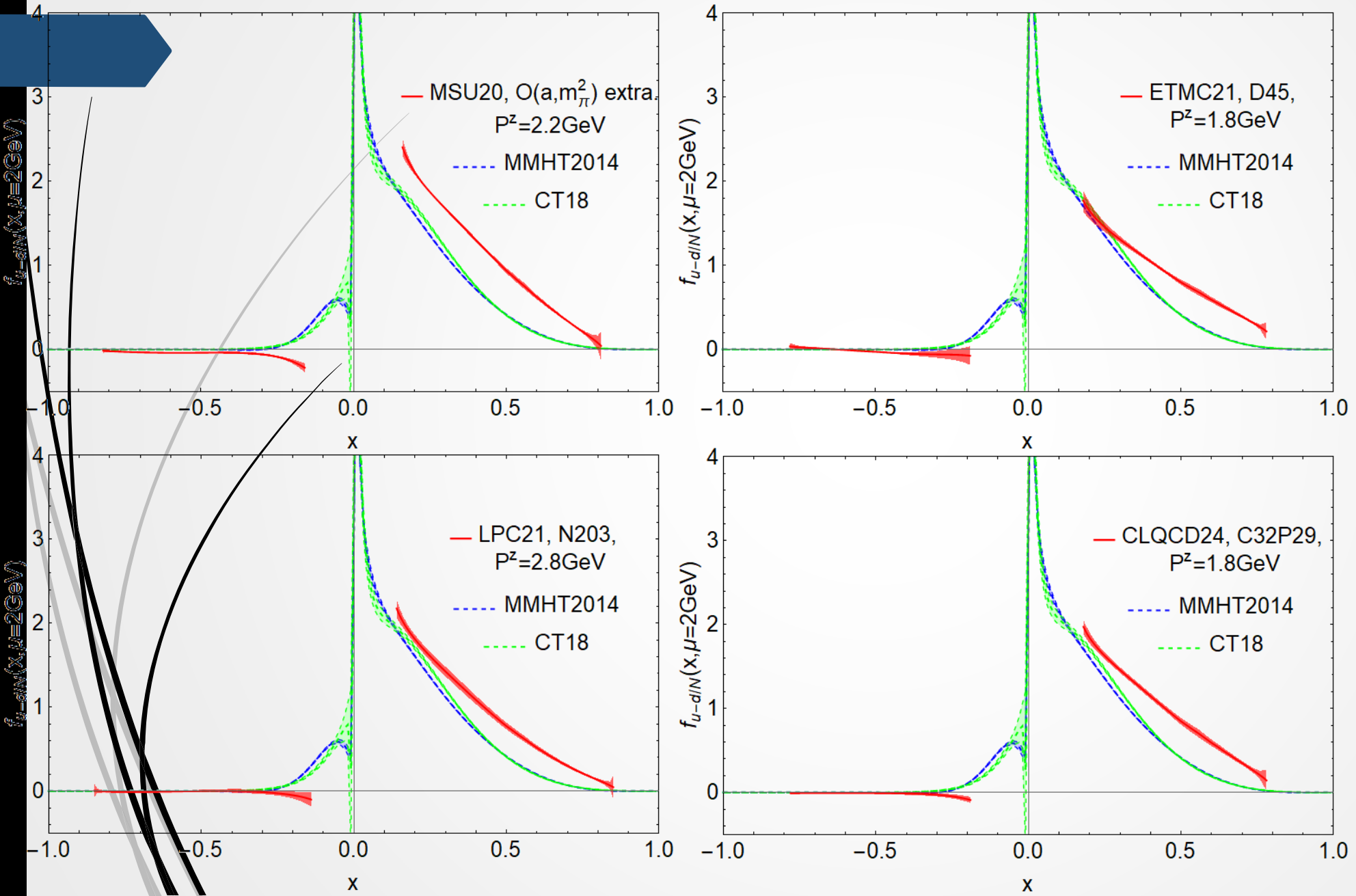
Zhang:2023bxs

Test on z_S 

Without and with the calibration of $\langle x \rangle$



Without the calibration of $\langle x \rangle$



- Without any modifications, the lattice PDFs deviate a lot from global fit ones.
- CLQCD and LPC contain stat err. MSU20 and ETMC21 do not.

Comparing calibration ansatz

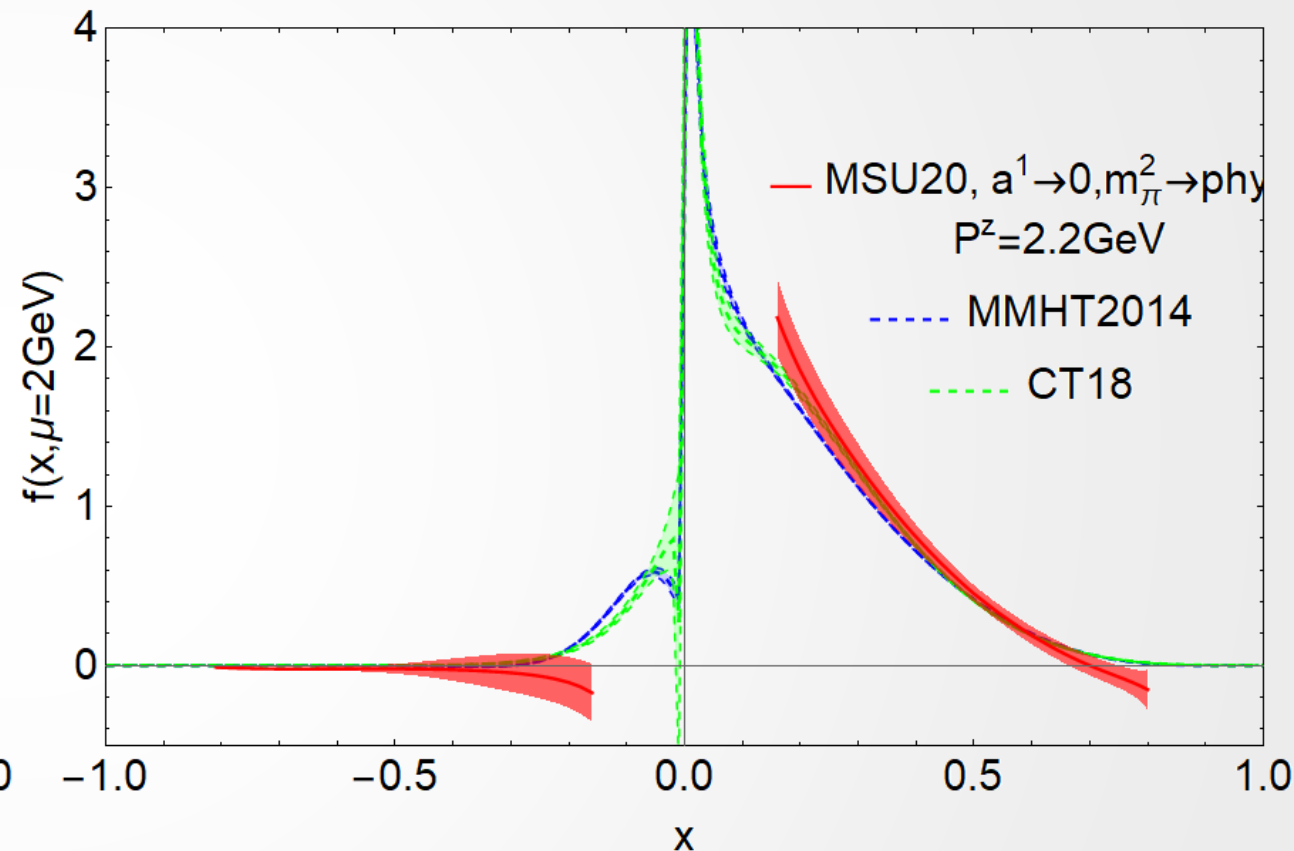
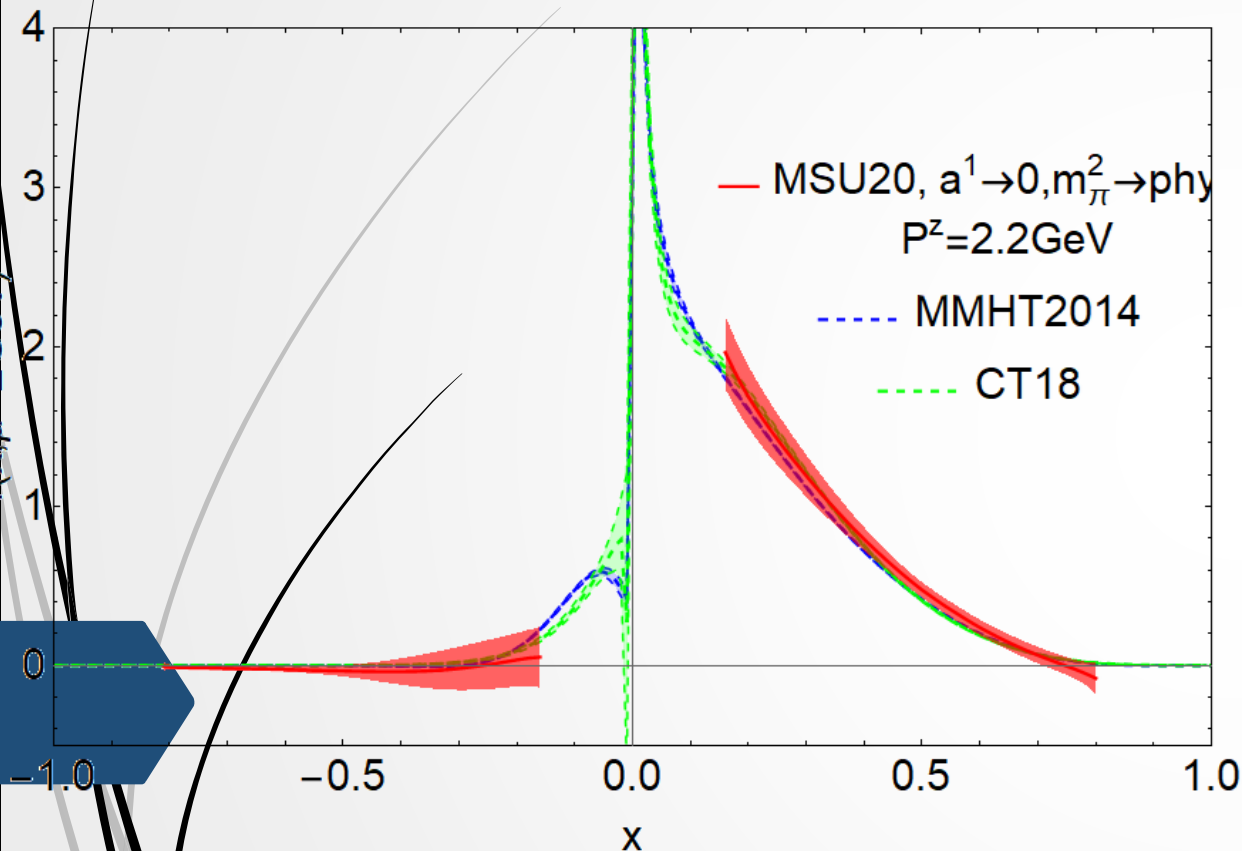
$$\tilde{h}(z, P^Z) = \tilde{h}^R(z, P^Z) e^{i z P^Z \delta y}$$

$$\tilde{h}(z, P^Z) = \tilde{h}^R(z, P^Z) (1 + i z P^Z \delta y)$$

f(x, μ=2 GeV)

0

-1.0



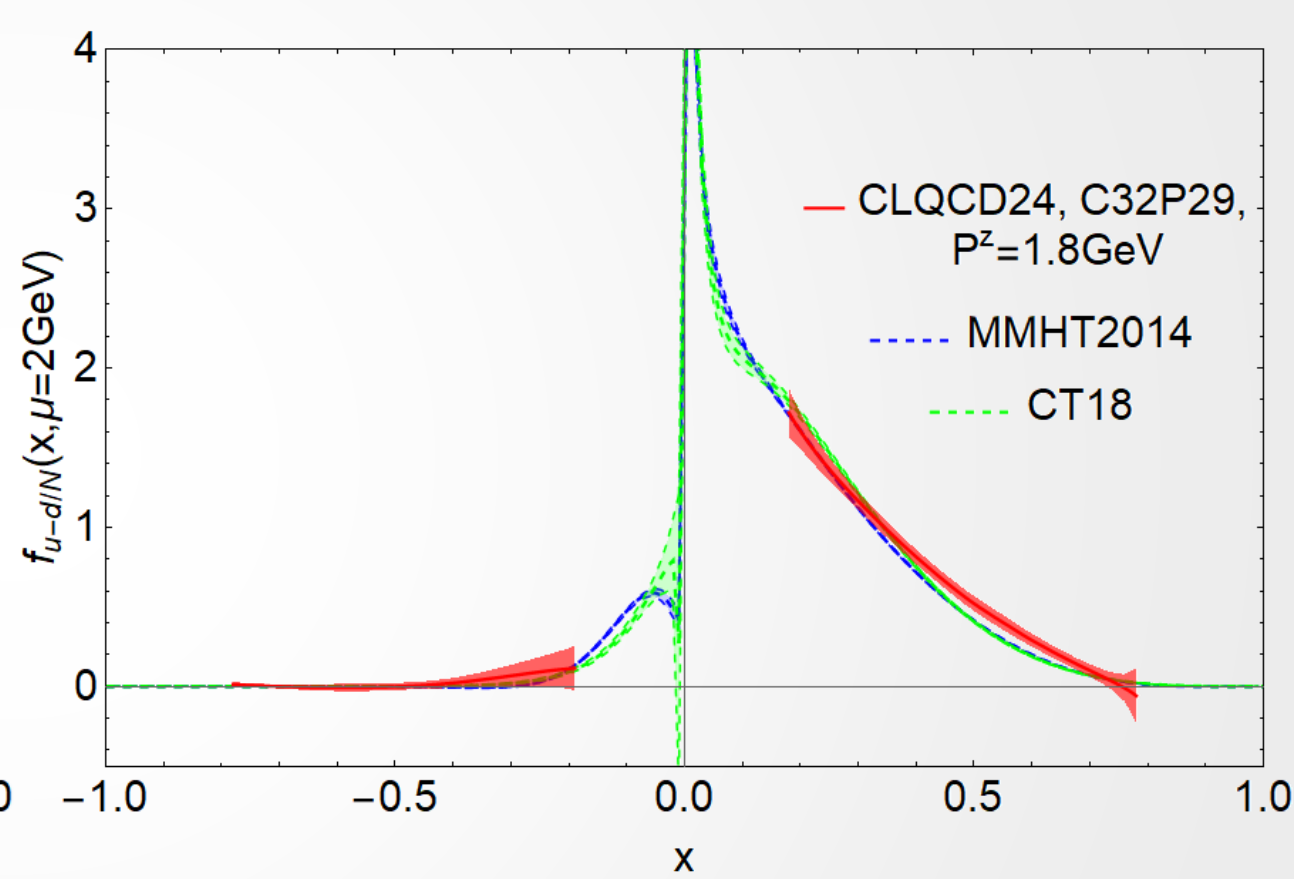
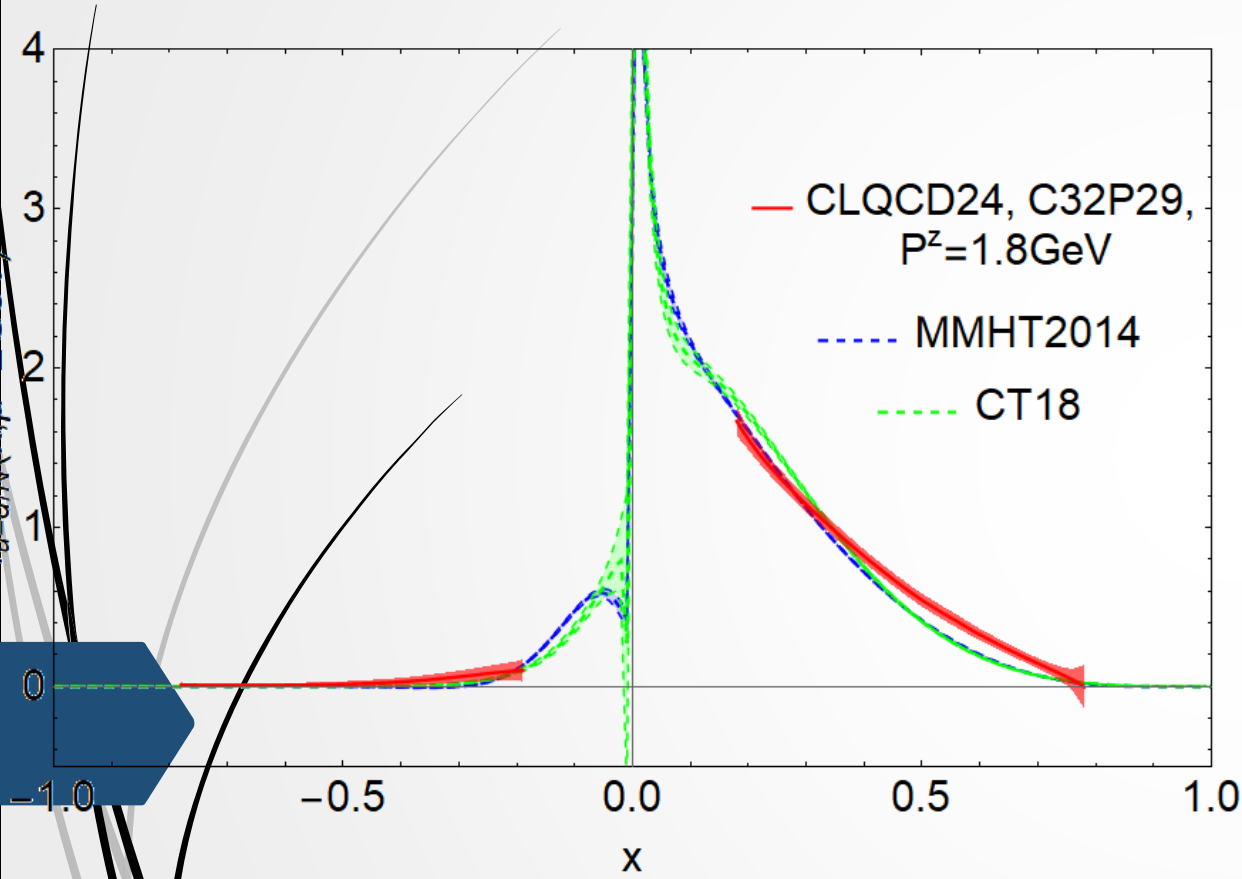
Comparing calibration ansatz

$$\tilde{h}(z, P^Z) = \tilde{h}^R(z, P^Z) e^{i z P^Z \delta y}$$

$$\tilde{h}(z, P^Z) = \tilde{h}^R(z, P^Z) (1 + i z P^Z \delta y)$$

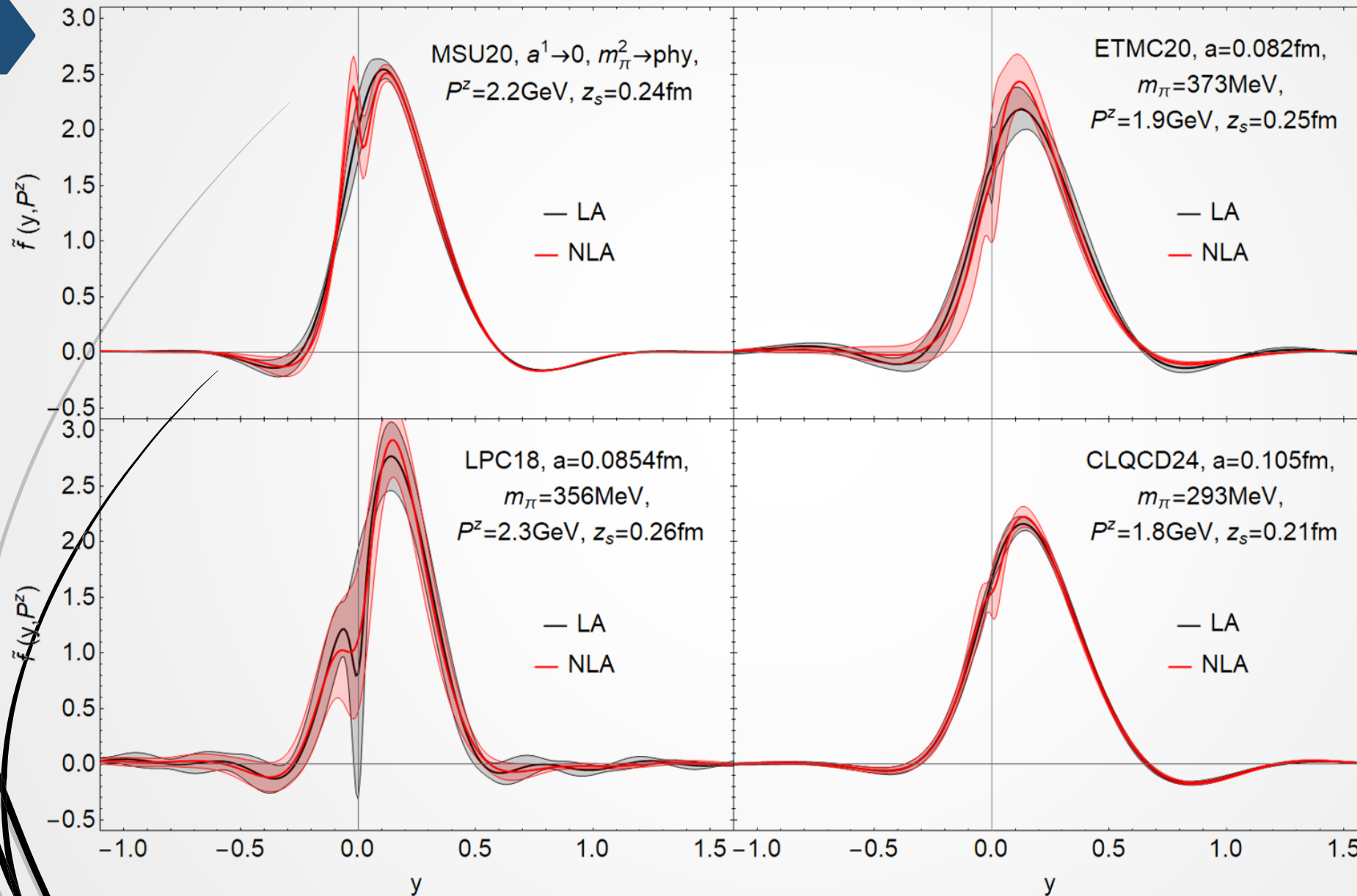
$f_{u-d/N}(x, \mu=2\text{GeV})$

0

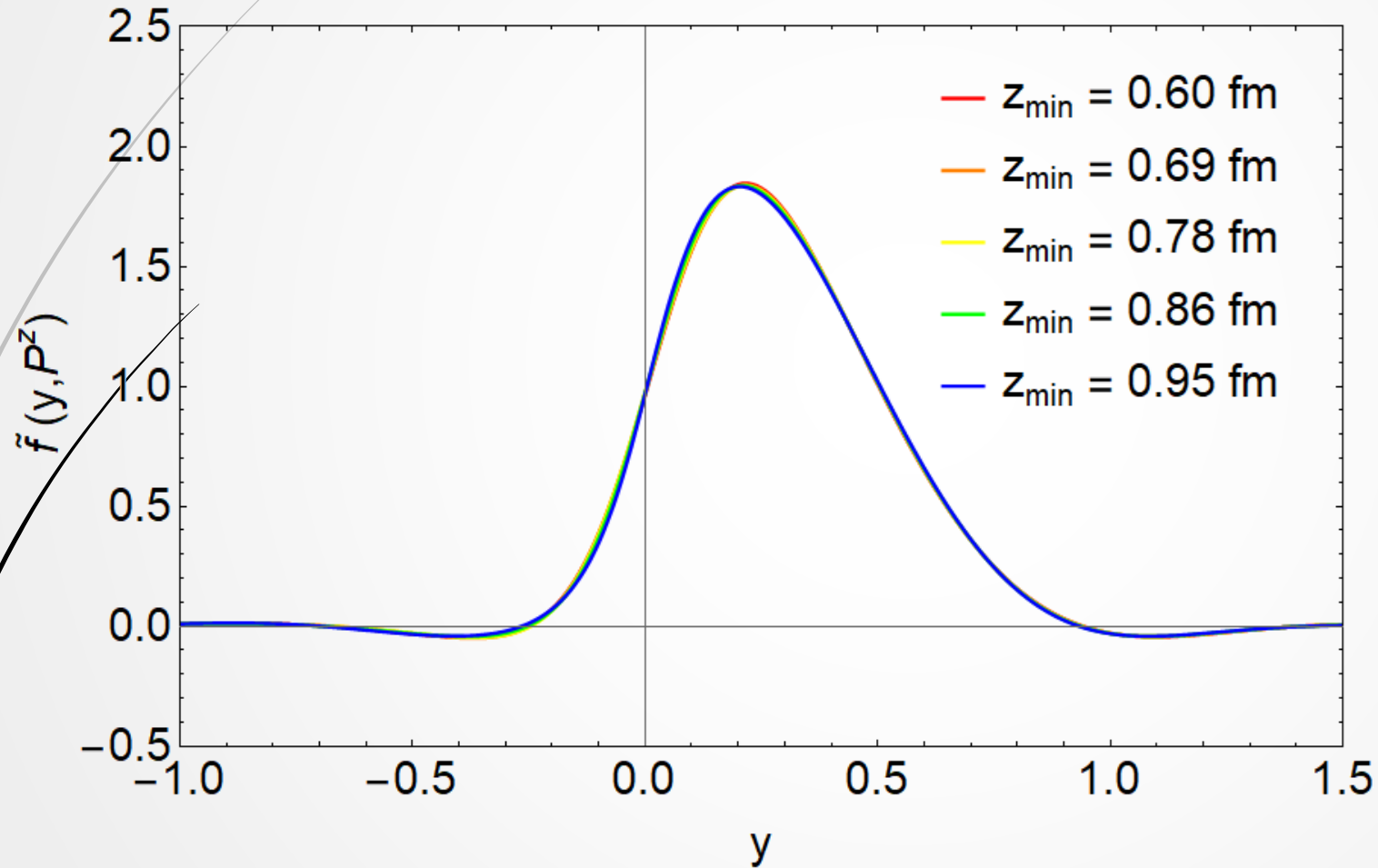


qPDFs obtained from different fit formulas

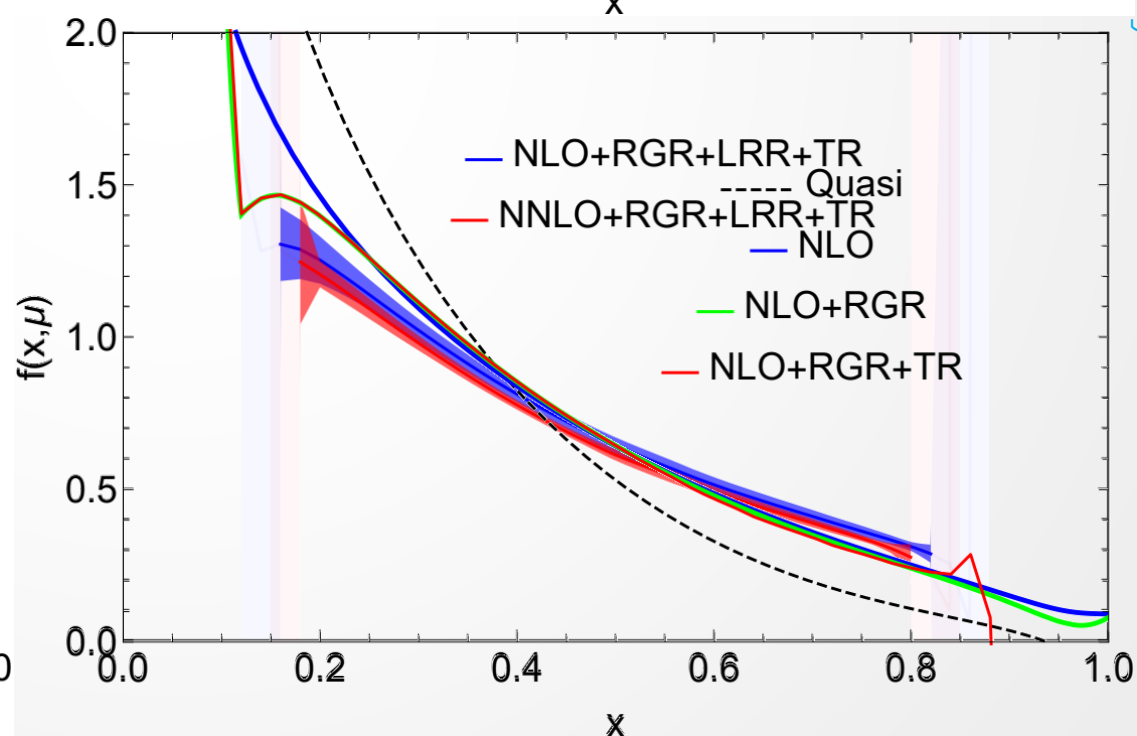
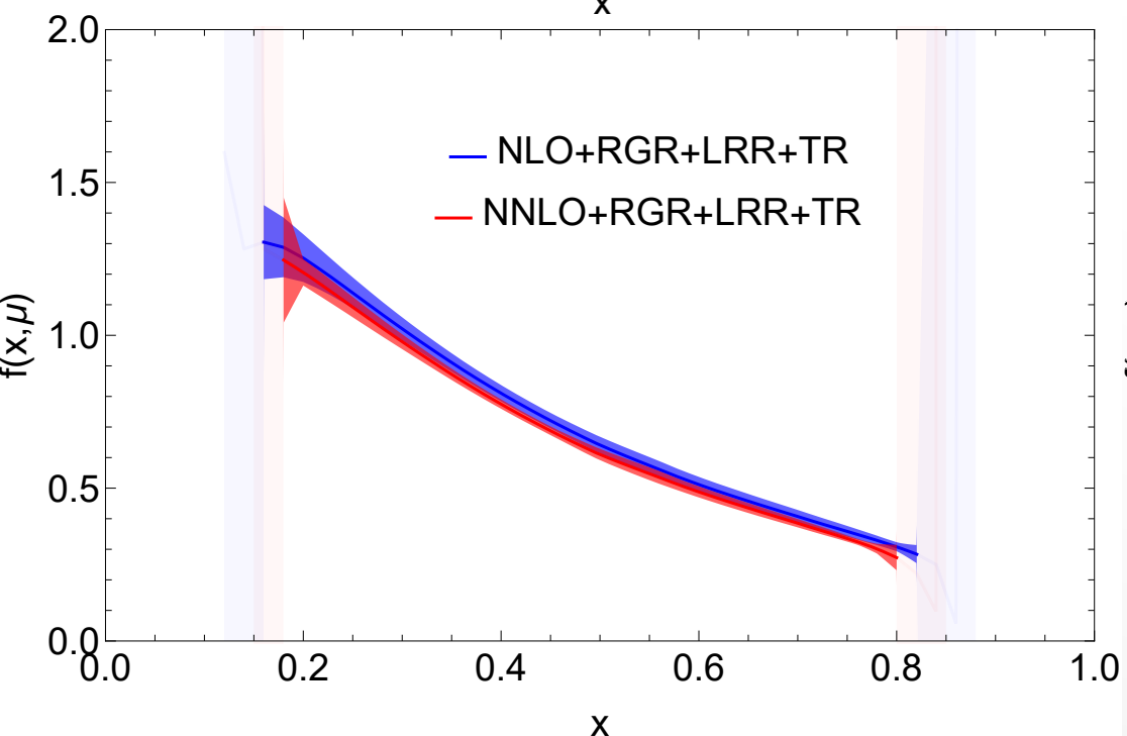
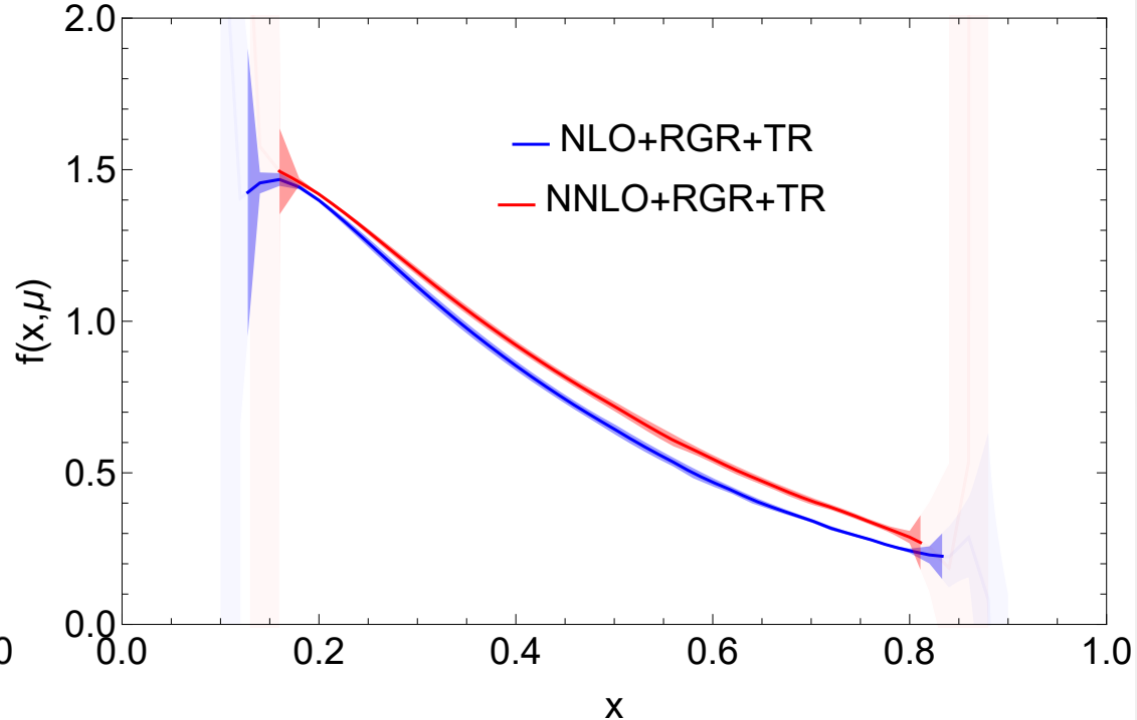
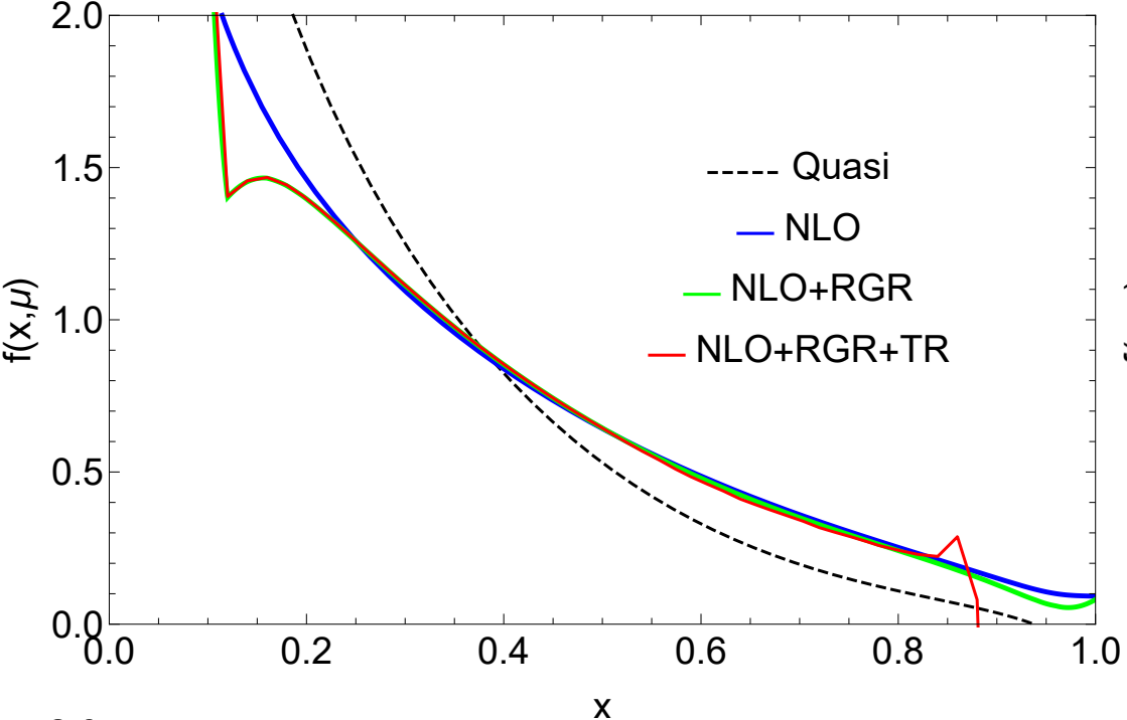
34



Fit range test

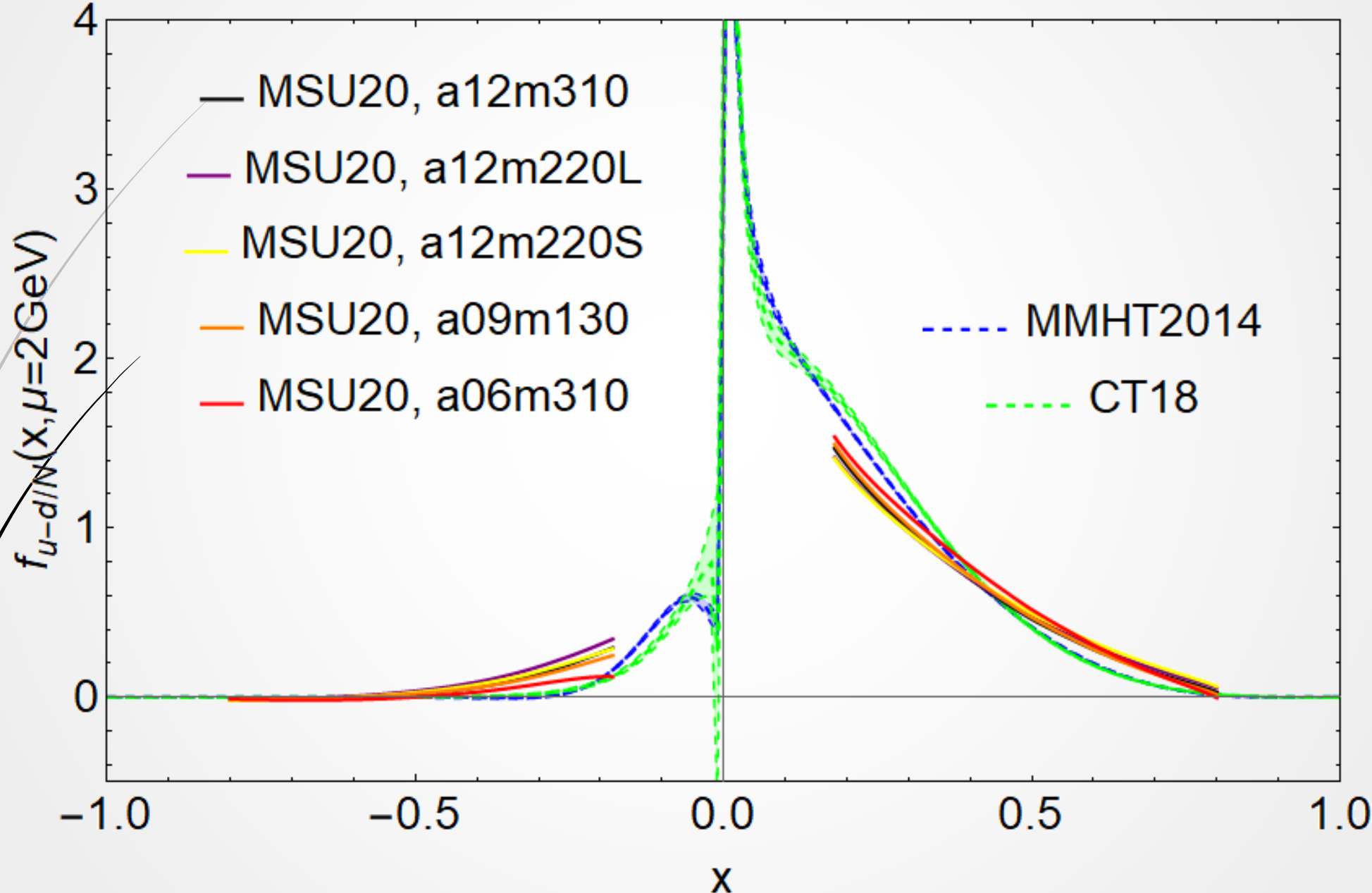


$$z_{\max} = 1.2 \text{ fm}$$



Ji:2024hit

MSU20(Lin:2020fsj): result of each ensemble



MSU20(Lin:2020fsj): coor extra (red) v.s. mom extra (black)

