

# Isvector parton distribution functions of the proton on a superfine lattice

Zhouyou Fan,<sup>1</sup> Xiang Gao<sup>2,3,\*</sup>, Ruizi Li,<sup>1</sup> Huey-Wen Lin,<sup>1,4</sup> Nikhil Karthik,<sup>2</sup> Swagato Mukherjee,<sup>2</sup> Peter Petreczky,<sup>2</sup> Sergey Syritsyn,<sup>5,6</sup> Yi-Bo Yang<sup>7,1</sup> and Rui Zhang<sup>1,4</sup>

<sup>1</sup>*Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA*

<sup>2</sup>*Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA*

<sup>3</sup>*Physics Department, Tsinghua University, Beijing 100084, China*

<sup>4</sup>*Department of Computational Mathematics, Michigan State University, East Lansing, Michigan 48824, USA*

<sup>5</sup>*RIKEN-BNL Research Center, Brookhaven National Laboratory, Upton, New York 11973, USA*

<sup>6</sup>*Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, USA*

<sup>7</sup>*CAS Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*



(Received 29 May 2020; accepted 22 September 2020; published 12 October 2020)

We study isovector unpolarized and helicity parton distribution functions (PDF) of the proton within the framework of large-momentum effective theory. We use a gauge ensemble, generated by the MILC Collaboration, with a superfine lattice spacing of 0.042 fm and a pion mass of 310 MeV, enabling us to simultaneously reach sub-Fermi spatial separations and larger nucleon momenta. We compare the spatial dependence of quasi-PDF matrix elements in different renormalization schemes with the corresponding results of the global fits, obtained using one-loop perturbative matching. We present determinations of the first four moments of the unpolarized and helicity PDFs of proton from the Ioffe-time dependence of the isovector matrix elements, obtained by employing a ratio-based renormalization scheme.

DOI: [10.1103/PhysRevD.102.074504](https://doi.org/10.1103/PhysRevD.102.074504)

## I. INTRODUCTION

Decades of deep inelastic scattering (DIS) and semi-inclusive DIS (SIDIS) data over wide kinematic ranges have provided us insight into the structure of nucleons. Significant progress also has been made in recent years—for example, the determination of the polarized gluon distribution at small  $x$  [1] based on the inclusive jet and pion production data from polarized  $p$ - $p$  collisions at the Relativistic Heavy-Ion Collider (RHIC) [2–4], double spin asymmetries from open-charm muon production at COMPASS [5], and the constraints on the polarization of sea quarks and antiquarks with longitudinal single-spin asymmetries in  $W^\pm$ -boson production [6,7]. In the future, the kinematic coverage of nucleon parton distribution functions (PDFs) will be greatly extended by the data from the Jefferson Lab 12-GeV program [8] and the Electron-Ion Collider (EIC) [9]. On the energy frontier, nucleon PDFs not only were a critical input for the discovery of the Higgs boson at the Large Hadron

Collider (LHC) [10,11], but also are expected to play critical roles in determining the Standard Model backgrounds during LHC’s search for physics beyond the Standard Model in future Runs 3–5.

Despite great progress on the experimental and phenomenological sides, nonperturbative determinations of the PDFs starting from the microscopic theory of quantum chromodynamics (QCD) remains a challenge. To obtain the quark PDF, one has to calculate the matrix element with the quark fields separated along the light cone between the hadronic states. Due to the light cone separation, straightforward calculation of the PDF is not possible using lattice QCD, a technique based on Euclidean-time formulation. One can bypass this obstacle by calculating a similar matrix element with spatially separated quark fields at equal time within highly boosted hadron states, which defines the so-called quasi-PDF (qPDF) [12,13]. For large hadron momenta, this matrix element can be related to the PDF [12,13]. The large-momentum effective theory (LaMET) provides a systematic way to relate the qPDF at large, but finite, hadron momentum to the PDF order by order in perturbation theory [13]. Related approaches to connect PDF to matrix elements of boosted hadrons calculable in the Euclidean-time lattice computations, such as “the good lattice cross section” [14,15] and the pseudo-PDF [16,17], have also been proposed. Renormalization of the underlying boosted hadron matrix elements, usually referred to

\*[xgao@bnl.gov](mailto:xgao@bnl.gov)

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

as the Ioffe-time distributions (ITDs), involves the Wilson line. The multiplicative renormalizability of the ITDs to all orders of perturbation theory has been proven [18,19]. Practical ways to implement renormalization on the lattice, such as the use of the regularization-invariant momentum-subtraction scheme (RI-MOM) [20–24] and reduced Ioffe-time distributions [17], have been established. The relations between different theoretical approaches are also now understood [25]. Based on these theoretical developments, unpolarized and polarized nucleon PDFs have been calculated on the lattice [24,26–33]. Furthermore, lattice calculations of the valence pion PDF have also appeared [34–38]. The status of this field is well summarized in recent review papers [39–42]. All of these calculations for the nucleon, so far, have been carried out with lattice spacing  $a > 0.08$  fm.

Having small lattice spacing plays a crucial role in calculation of the PDF within the LaMET framework. To suppress the target mass and higher twist corrections, the hadron momentum  $P_z$  should be large. But to avoid large discretization effects, one must ensure  $aP_z \ll 1$ . Furthermore, to obtain the light cone PDF from a qPDF, one needs perturbative matching, which presently is known only up to one-loop order. Applicability of one-loop perturbative matching can be guaranteed only for spatial separations  $z\Lambda_{\text{QCD}} \ll 1$ , and therefore demands the use of fine lattices. The main goal of the present work is to study systematic effects of the PDF calculations within the LaMET framework by going to the extreme limit with the use of a superfine lattice having  $a = 0.042$  fm. The lattice spacing used in this study is at least twice smaller than that used in any previous lattice calculations of the nucleon PDF. The unpolarized and helicity PDFs of the nucleon are well constrained through global fits to experimental results. Thus, we study the systematic effects of our calculations by comparing the  $P_z$  and  $z$  dependence of renormalized qPDF matrix elements with the same reconstructed from the well-known phenomenological PDFs using the LaMET framework.

The rest of the paper is organized as follows: In Sec. II, we discuss the general features of LaMET and our lattice setup. In Sec. III, we discuss the nucleon two-point functions for large values of  $P_z$  and the determination of the energy levels of a fast-moving nucleon. Section IV is dedicated to the analysis of the nucleon three-point functions and the calculations of bare qPDF. Section V describes the nonperturbative RI-MOM renormalization. Comparisons of the lattice results on qPDF with the results of global analysis of unpolarized and helicity PDFs are discussed in Secs. VI and VII, respectively. Different from RI-MOM renormalization, we discuss the analysis of ratios of nucleon matrix elements in Sec. VIII. Finally, Sec. IX contains our conclusions.

## II. LATTICE SETUP AND LaMET

In this paper, we report the results of a lattice QCD calculation using clover valence fermions on an ensemble

of  $N_f = 2 + 1 + 1$  gauge configurations with lattice spacing  $a = 0.042$  fm, with space-time dimensions of  $64^3 \times 192$  and pion mass  $M_\pi \approx 310$  MeV in the continuum limit. The gauge configurations have been generated using highly improved staggered quarks (HISQ) [43] by the MILC Collaboration [44]. The gauge links entering the clover Wilson-Dirac operator have been smeared using hypercubic (HYP) smearing [45]. We used the tree-level tadpole improved result for the coefficient of the clover term, and the bare quark mass has been tuned to recover the lowest pion mass of the staggered quarks in the sea [46–49]. We use only one step of HYP smearing to improve the discretization effects, since it is possible that multiple applications of smearing could alter the ultraviolet results for the PDF. We use a multigrid algorithm [50,51] in the Chroma software package [52] to perform the inversion of the clover fermion matrix, allowing us to collect a relatively high statistics sample. We collected a total of 3258 measurements using six sources per configuration and 543 gauge configurations. In the following, we elaborate on the steps of our computation.

### A. Nucleon two-point correlators

The two crucial components of the lattice computation are the two-point function and the three-point function involving the boosted nucleon and the qPDF operator. The two-point function for the nucleon boosted to spatial momentum  $\mathbf{P}$  is the standard operator

$$C_{2\text{pt}}(t_s) = \langle \hat{N}_s(\mathbf{P}, t_s) \hat{N}_s^\dagger(\mathbf{P}, 0) \rangle, \\ \hat{N}_s(\mathbf{P}, t) = \sum_{\mathbf{x}} \epsilon_{abc} u_a^{(s)}(\tilde{\mathbf{x}}) (u_b^{(s)}(\tilde{\mathbf{x}})^T C \gamma_5 d_c^{(s)}(\tilde{\mathbf{x}})) e^{-i\mathbf{P}\cdot\mathbf{x}}, \quad (1)$$

where  $\tilde{\mathbf{x}} = (\mathbf{x}, t)$  and  $t_s$  is the source-sink separation along the Euclidean-time direction. The index  $s$  refers to the kind of quark smearing that is applied to improve the signal-to-noise ratio of the boosted nucleon states. We either use point quark operators  $\psi(x)$ , or we use the Gaussian momentum smeared [53] for the quark fields  $\psi^{(s)}(x)$  that enters  $\hat{N}_s$ :

$$\psi^{(s)}(\tilde{\mathbf{x}}) = S_{\text{mom}} \psi(\tilde{\mathbf{x}}) \\ = \frac{1}{1+6\alpha} \left( \psi(\tilde{\mathbf{x}}) + \alpha \sum_j U_j(\tilde{\mathbf{x}}) e^{i\mathbf{k}\cdot\hat{j}} \psi(\tilde{\mathbf{x}} + \hat{j}) \right), \quad (2)$$

where  $\mathbf{k}$  is the momentum of the quark field,  $U_j(\tilde{\mathbf{x}})$  are the gauge links in the  $\hat{j}$  direction, and  $\alpha$  is a tunable parameter as in traditional Gaussian smearing. The quark momentum should be chosen such that the signal-to-noise ratio is optimal for the given nucleon momentum. Naively, one would expect that  $|\mathbf{k}|$  should be one third of the nucleon momenta [53]. For this particular study, we use  $j = z$  and  $k_z = 4\pi/L$ , and a large Gaussian-smearing parameter

$\alpha = 10$ . Such a momentum source is designed to align the overlap with nucleons of the desired boost momentum, and we are able to reach higher boost momentum for the nucleon states with reasonable signals. In the nucleon two-point correlators, we can study multiple values of the nucleon momentum,  $\mathbf{P} = \{0, 0, P_z\}$ , with

$$P_z = n_z \frac{2\pi}{L}, \quad n_z \in [0, 6], \quad (3)$$

without a significant increase in computational needs. These values of  $n_z$  from 1 to 6 correspond to  $P_z = 0.46, 0.92, 1.38, 1.84, 2.31$ , and  $2.77$  GeV in physical units, respectively. We either use smeared fields for both the source and sink, which we refer to as SS, or we use smeared fields only for the source and point fields for the sink, which we refer to as SP in the rest of the paper.

### B. Nucleon three-point function

The three-point function we compute is of the form

$$C_{3\text{pt}}(t_s, \tau) = \mathbb{P} \langle \hat{N}_s(\mathbf{P}, t_s) O_\Gamma(z; \tau) \hat{N}_s^\dagger(\mathbf{P}, 0) \rangle, \quad (4)$$

where  $O_\Gamma(z; \tau)$  is the  $u$ - $d$  isovector qPDF operator

$$O_\Gamma(z; \tau) = \sum_{\mathbf{x}} \bar{u}(\tilde{\mathbf{x}} + z) \Gamma W_z(\tilde{\mathbf{x}} + z, \tilde{\mathbf{x}}) u(\tilde{\mathbf{x}}) - \sum_{\mathbf{x}} \bar{d}(\tilde{\mathbf{x}} + z) \Gamma W_z(\tilde{\mathbf{x}} + z, \tilde{\mathbf{x}}) d(\tilde{\mathbf{x}}), \quad (5)$$

where  $\tilde{\mathbf{x}} = (\mathbf{x}, \tau)$ , and  $W_z$  is the straight Wilson line along the spatial  $z$  direction, connecting lattice sites  $\tilde{\mathbf{x}}$  and  $\tilde{\mathbf{x}} + z$ . The Dirac  $\Gamma$  used will determine the quantum numbers of the PDF:  $\Gamma = \gamma_t$  for the unpolarized case, and  $\Gamma = \gamma_z \gamma_5$  for the longitudinally polarized case. The projector operator,  $\mathbb{P}$ , is given by  $\mathbb{P} = \frac{1+\gamma_t}{2}$  for the unpolarized case and  $\mathbb{P} = i\gamma_z \gamma_5 \frac{1+\gamma_t}{2}$  for the longitudinally polarized case. We only use smeared quark sources for the computation of  $C_{3\text{pt}}$ . In order to reduce the computational cost, we only compute the  $C_{3\text{pt}}$  for two large values,  $P_z = 1.84$  and  $2.31$  GeV, and for source-sink separations  $t_s = 16a, 18a, 20a$ .

### C. Extraction of nucleon matrix element and perturbative matching to PDF

Using the three-point and two-point functions whose calculations are described above, we can extract the bare matrix element

$$h(z, P_z, \Gamma) = \langle P_z | O_\Gamma(z) | P_z \rangle \quad (6)$$

formally in the infinite source-sink separation  $t_s$  limit of their ratio,

$$R(z, P_z, \Gamma; \tau, t_s) = \frac{C_{3\text{pt}}(\tau, t_s)}{C_{2\text{pt}}(t_s)}. \quad (7)$$

To obtain the matrix element  $h(z, P_z)$  from the above ratio, we calculate the nucleon three-point function with the insertion of the  $O_\Gamma(z)$  operator at three nucleon three-point source-sink separations, approximately  $t_s = 0.67, 0.76$ , and  $0.84$  fm, and we describe its  $t_s$  and  $\tau$  dependence through two- and three-state *Ansätze*. In Sec. IV, we describe our extraction of the bare matrix element from various extrapolations in detail.

The next step of the computation is the renormalization of the bare matrix element  $h$ . One possible choice for  $O_\Gamma$  is  $O_{\gamma_z}$ . However, for this case of  $\Gamma = \gamma_z$ , there is a mixing with the quark bilinear operator containing the unit matrix,  $\Gamma = 1$ , if Wilson fermions are used [20,21,54]. This mixing is absent if we use  $\Gamma = \gamma_t$ , and we will use this choice for the unpolarized PDF in this study. One way to perform the renormalization procedure on the lattice is to use the RI-MOM scheme [20,22], where the renormalized matrix element is defined as

$$h^R(z, P_z, \mu_R, p_z^R) = Z(z, \mu^R, p_z^R) h(z, P_z, \Gamma). \quad (8)$$

The non-perturbatively determined RI-MOM renormalization constant  $Z(z, \mu^R, p_z^R)$  depends on the separation  $z$ , the norm of the renormalization point  $\mu^R = (p^R)^2$  and the  $z$  component of renormalization point  $p_z^R$ . The dependence on  $p_z^R$  arises because the  $z$  component of the momentum now plays a special role. We will discuss the details of the RI-MOM renormalization in Sec. V. We will also consider an alternate ratio scheme that has a well-defined continuum limit in Sec. VIII. Here, the multiplicative renormalization factor  $Z_{\text{ratio}}(z)$  can be taken as the hadron matrix element at a different fixed momentum  $P'_z$ —i.e.,  $Z_{\text{ratio}}(z) = (h(z, P'_z, \Gamma))^{-1}$ .

After the RI-MOM renormalization, one obtains the renormalized matrix element  $h_R(z, P_z, \mu_R, p_z^R)$ , from which we can define the qPDF as a function of Bjorken- $x$ :

$$\tilde{q}(x, P_z, \mu_R, p_z^R) \equiv \int_{-\infty}^{\infty} \frac{dz}{4\pi} e^{ixP_z z} h_R(z, P_z, \mu_R, p_z^R). \quad (9)$$

From this formula, it is clear that  $h_R(z, P_z, \mu_R, p_z^R)$  can be considered as the coordinate space qPDF. For finite momentum  $P_z$ ,  $\tilde{q}(x, P_z, \mu_R, p_z^R)$  has support in  $-\infty < x < \infty$ . Unlike the physical PDF, which is frame independent, the qPDF has a nontrivial dependence on the nucleon momentum  $P_z$ . When the nucleon momentum  $P_z \gg \{M, \Lambda_{\text{QCD}}\}$ , with  $M$  being the nucleon mass, the qPDF in the RI-MOM scheme can be matched to the PDF defined in the  $\overline{\text{MS}}$  scheme,  $q(x, \mu)$ , through the factorization theorem [12,13,25]

$$\tilde{q}(x, P_z, p_z^R, \mu_R) = \int_{-1}^1 \frac{dy}{|y|} C\left(\frac{x}{y}, r, \frac{yP_z}{\mu}, \frac{yP_z}{p_z^R}\right) q(y, \mu) + \mathcal{O}\left(\frac{M^2}{P_z^2}\right) + \mathcal{O}\left(\frac{\Lambda_{\text{QCD}}^2}{P_z^2}\right), \quad (10)$$

where  $r = (\mu_R/p_z^R)^2$ ,  $C$  is the perturbative matching coefficient,  $\mathcal{O}(M^2/P_z^2)$  is the target-mass correction due to the nonzero nucleon mass, and  $\mathcal{O}(\Lambda_{\text{QCD}}^2/P_z^2)$  stands for higher-twist contributions. The flavor indices of  $q$ ,  $\tilde{q}$ , and  $C$  are implied. In what follows, we will discuss the nonsinglet case, and therefore, mixing with gluon and sea-quark PDFs is absent in the above formula. We use the one-loop expression of the kernel  $C$ . (The one-loop matching, including for the singlet case, has also been worked out in Refs. [55,56]).

The matching kernel  $C(x, r, P_z/\mu, P_z/p_z^R)$  for  $\Gamma = \gamma_t$  was derived in Ref. [24] and depends on details of the RI-MOM scheme. It can be written in the following form:

$$C\left(x, r, \frac{P_z}{\mu}, \frac{P_z}{p_z^R}\right) = \delta(1-x) + \left[ f_{1,\Gamma}\left(x, \frac{P_z}{\mu}\right) - \left| \frac{P_z}{p_z^R} \right| f_{2,\Gamma,\mathcal{P}}\left(1 + \frac{P_z}{p_z^R}(x-1), r\right) \right]_+. \quad (11)$$

The subscript “+” stands for the plus-prescription. Both the functions,  $f_{1,\Gamma}$  and  $f_{2,\Gamma,\mathcal{P}}$ , depend on the choice of the  $\Gamma$  in the operator insertion [24]. On the other hand,  $f_{1,\Gamma}$  is independent of the projection operator ( $\mathcal{P}$ ) used in defining the RI-MOM renormalization condition, but  $f_{2,\Gamma,\mathcal{P}}$  is different for different choices of the RI-MOM renormalization condition [24]. We also note that it is also possible to convert  $h_R(z, P_z, \mu_R, p_z^R)$  to the  $\overline{\text{MS}}$  scheme and define

the corresponding qPDF  $\tilde{q}(x, P_z, \mu)$ , that then can be directly matched to the  $\overline{\text{MS}}$  PDF [22].

To study the longitudinally polarized quark PDF, one can use  $\Gamma = \gamma_z \gamma_5$  or  $\Gamma = \gamma_t \gamma_5$ . In the case  $\Gamma = \gamma_z \gamma_5$ , there is no mixing with quark bilinear operators with  $\Gamma = 1$  [22]. Therefore, we will use this choice to study the longitudinally polarized quark PDF and qPDF. The bare matrix element of  $O_{\gamma_z \gamma_5}$  can be renormalized using the RI-MOM scheme and then match to PDF in the same manner as was done for the unpolarized case. The RI-MOM renormalization for the longitudinally polarized case will be discussed in Sec. V, while details of the matching procedure, including the formulas for  $f_1$  and  $f_2$  functions, will be given in Sec. VII.

### III. ANALYSIS OF THE NUCLEON TWO-POINT FUNCTION

For the extraction of the qPDF matrix element of the nucleon at large momenta, it is important to understand the contribution of different energy states to the nucleon two-point correlation function. We calculated the nucleon two-point function using a smeared source and smeared sink (SS correlator), as well as a smeared source and point sink (SP correlator), for seven values of the momenta  $aP_z = 2\pi/L \cdot n_z$ ,  $n_z = 0, 1, 2, 3, 4, 5$ , and 6. From the two-point correlators,  $C_{2\text{pt}}^i(t_s, P_z)$ ,  $i = \text{SS or SP}$ , we define the effective mass

$$aE_{\text{eff}}(t_s, P_z) = \ln\left(\frac{C_{2\text{pt}}^i(t_s/a, P_z)}{C_{2\text{pt}}^i(t_s/a + 1, P_z)}\right). \quad (12)$$

Our results for the effective masses are shown in Fig. 1 for the SP and SS correlators.

The effective mass should approach a constant corresponding to the ground-state energy  $E_0(P_z)$  at sufficiently large  $t_s$ . The momentum dependence of the ground-state

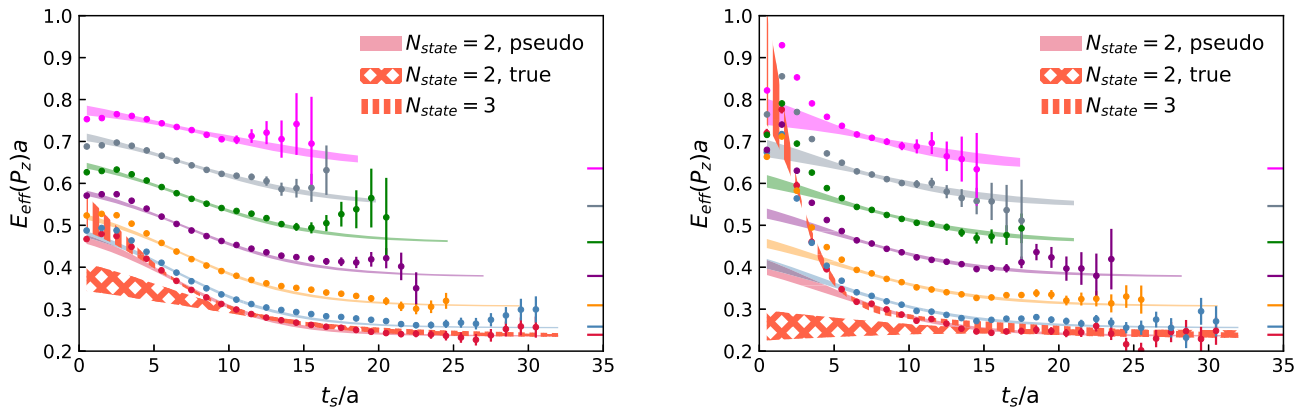


FIG. 1. The effective masses obtained from SP (left) and SS (right) correlators for different momenta shown in Eq. (3). The bands come from the results of two-state ( $N_{\text{state}} = 2$ ) and three-state ( $N_{\text{state}} = 3$ ) fits. For  $N_{\text{state}} = 2$ , “pseudo” indicates that the effective pseudoplateau in the range  $5a < t_{\min} < 10a$  for the first excited state  $E_1$  has been used, and “true” indicates that the true plateau value of  $E_1$  in the range  $t_{\min} > 11a$  has been used (see text for details).

energy is expected to be described by the dispersion relation  $E_0(P_z) = \sqrt{P_z^2 + M^2}$ , with  $M$  being the nucleon mass. Therefore, in Fig. 1, we show the expected ground-state energy at different  $P_z$  values obtained from the dispersion relation as horizontal lines at the right for comparison. Along with the expected asymptotic values at large  $t_s$ , we also show the  $t_s$  dependence of the effective mass based on an effective two-state fit to the two-point function, as we will explain shortly. Indeed, we see that the effective masses approach the corresponding values. The effective masses corresponding to the SP correlator reach a plateau at a slightly larger  $t_s$  than the SS correlators. On the other hand, at small  $t_s$ , the effective masses for the SP correlators are smaller than those for SS correlators. This implies that the contribution of the excited states is smaller for the SP correlator, for which a plausible reason could be that the different excited states contribute with different signs to the correlator. Thus, even though the ground and the excited-state energies are the same in the SP and SS correlators, the two are affected differently by the higher excited states, which we can take advantage of to obtain the excited state spectrum reliably.

In order to determine the energy levels, we fit the spectral decomposition of  $C_{2pt}(t_s)$ ,

$$C_{2pt}(t_s) = \sum_{n=0}^{N_{\text{state}}-1} A_n e^{-E_n t_s}, \quad (13)$$

truncated at  $N_{\text{state}}$ , to the two-point function data over a range of values of  $t_s$  in the range  $[t_{\min}, 32a]$ . Since the lattice extent in the time direction is 192, we do not find any

effect of lattice periodicity in this range of  $t_s$  to be important. We perform this fitting with the one-state ( $N_{\text{state}} = 1$ ), two-state ( $N_{\text{state}} = 2$ ), and three-state ( $N_{\text{state}} = 3$ ) *Ansätze*. The ground-state energies  $E_0$  from the fits of SS correlators for  $n_z = 3$  and 4 are shown in the left panels of Fig. 2 as functions of  $t_{\min}$ , where  $t_{\min}$  indicates that only  $C_{2pt}(t_s > t_{\min})$  has been fitted. Similar results were obtained at the other values of the momenta. The horizontal lines in the figures correspond to the results from the dispersion relation for  $E_0$ . The single exponential fits give a good description of the SS correlator for  $t_{\min} > 11a$ , while two exponential fits give stable results for the ground-state energy already for  $t_{\min} > 5a$ .

We find the determination of the excited-state energies from the SS correlators to be more problematic than that from SP correlators. The excited-state energy for SS is not well constrained by simple two-exponential fits, and it is also not very stable with respect to the variation of  $t_{\min}$ . Since the SP and SS correlators receive different contributions from excited states, we perform a combined analysis of them to obtain more reliable results for the excited-state energies. Since we are able to obtain the ground-state energy  $E_0$  reliably from one or two exponential fits to both the SS and SP correlators, and they agree with the expectation from the dispersion relation well, we use  $E_0$  as a prior to perform more stable two-exponential fits. The results from the two-state exponential fits, with  $E_0$  as prior, for  $n_z = 3$  and  $n_z = 4$  are shown in the middle and right panels of Fig. 2 for the SP and SS correlators, respectively. For the SP correlators, the excited-state energy  $E_1$  seems to approach a plateau smoothly for  $t_{\min} > 13a$ .

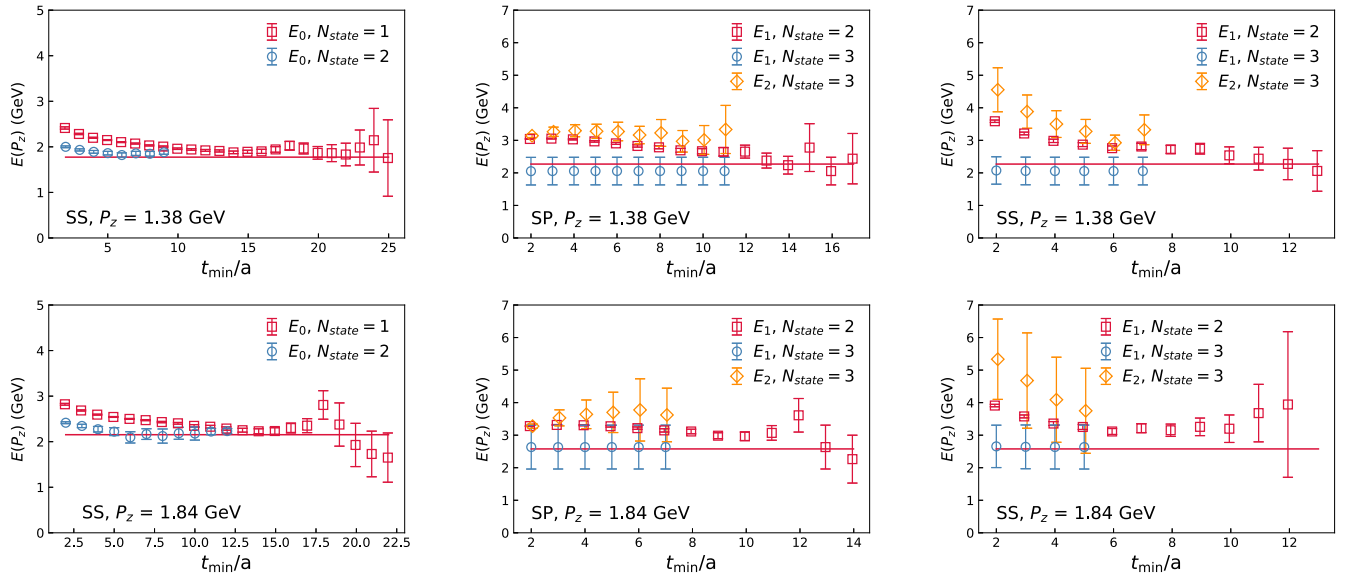


FIG. 2. Fit results for  $n_z = 3$  (upper panels) and  $n_z = 4$  (lower panels) for the nucleon two-point function. The left panels are for the ground state ( $E_0$ ) from one-state and two-state fits. The middle and right panels are for the first ( $E_1$ ) and second ( $E_2$ ) excited states, determined by two-state and three-state prior-based fits (see text for details). The horizontal lines are the values calculated from the dispersion relation.

It is interesting to note that, empirically, we observe that the values of the plateaus agree with the dispersion relation  $E_1(P_z) = \sqrt{P_z^2 + E_1(P_z=0)^2}$ , which is shown with the horizontal lines. While being an interesting observation, such a stringent identification of this state is not important to our analysis and requires further studies to rigorously establish this. For the SS correlator,  $E_1$  develops a pseudoplateau for  $5a < t_{\min} < 10a$ , and it relaxes to the true plateau (i.e., as identified from the SP case) for  $t_{\min} > 11a$ . For  $n_z = 4$ , it is actually difficult to identify the true plateau. To model the excited-state contributions to  $R(\tau, t_s)$  in the range  $0 < \tau < t_s/2$ , with  $t_s = 16a, 18a, 20a$ , one might consider using the well-determined values of  $E_0$  and  $E_1$  from the SP correlator at large  $t_s$ . However, as we will demonstrate now, such choices provide a less accurate description of the SS two-point function in the range  $5a < t_s < 10a$ . A better description of the excited-state contributions to the  $C_{2pt}(5a < t_s < 10a)$  can be obtained by using the effective pseudoplateau value of  $E_1$  in the range  $5a < t_{\min} < 10a$ .

Since we observe  $E_1$  to be well described by a particle-like dispersion relation for sufficiently large  $t_s$ , we perform three-state fits for both SP and SS correlators by imposing a prior on  $E_1$  as well, using its best estimate from a two-state fit of SP correlators with the corresponding jackknife errors [35]. The results are shown in the middle and right panels in Fig. 2. We see that with the prior-based three-state exponential fits, we can obtain stable results for the first excited-state energy  $E_1(P_z)$  already for relatively small  $t_{\min}$ , which agrees with the dispersion relation value that we input via the prior. The value of the second excited state is also shown in Fig. 2, and it roughly agrees with the values of  $E_1$  from the two-exponential fit (with a prior only on  $E_0$ )

at smaller  $t_{\min}$ . Since the value of  $E_2$  is quite large, the third exponential probably corresponds to a combination of several excited states. In Fig. 1, we show the  $1\sigma$  bands for the effective mass corresponding to (1) a two-state fit that uses values of  $E_0$  and the true value of  $E_1$ ; (2) a two-state fit obtained by setting  $E_1$  to be the effective value in the range  $5a < t_{\min} < 10a$ ; and (3) the three-state fit that we described above. We find that the curves (2) and (3) agree quite well with each other in the range of  $5a < t_s < 10a$ , and they extrapolate in a similar fashion to the asymptotic value  $E_0$ . However, the curve (1) fails in capturing the data in the range  $5a < t_s < 10a$ . Since for our three-point calculations the source-sink separations were chosen to be  $t_s = 16a, 18a, 20a$ , we must model the effective excited-state contributions to the three-point functions in the range  $0 < \tau < t_s/2$ . Thus, through this analysis on SP and SS correlators, we numerically demonstrate that the usage of an effective value of  $E_1$  in the range  $5a < t_s < 10a$  that is higher than the true value of  $E_1$  is justified, and that it is the best extrapolation one could perform for the extraction of bare matrix elements in the absence of enough data to perform a three-state fit.

Let us now summarize the analysis of the nucleon two-point function. Using boosted Gaussian sources, we were able to extract ground-state energy levels up to momenta 2.7 GeV from SP and SS correlators. The ground-state energy dependence on  $P_z$  seem to follow the continuum dispersion relation. Using this fact, we performed prior-based fits using the energy from the dispersion relation as a prior and extracted the excited-state energies as functions of  $P_z$ . For SP correlators, the extracted value of  $E_1$  agrees well with the one from the dispersion relation. We show this in the left panel of Fig. 3. Furthermore, we were able to extract an effective third energy level. These results are shown as

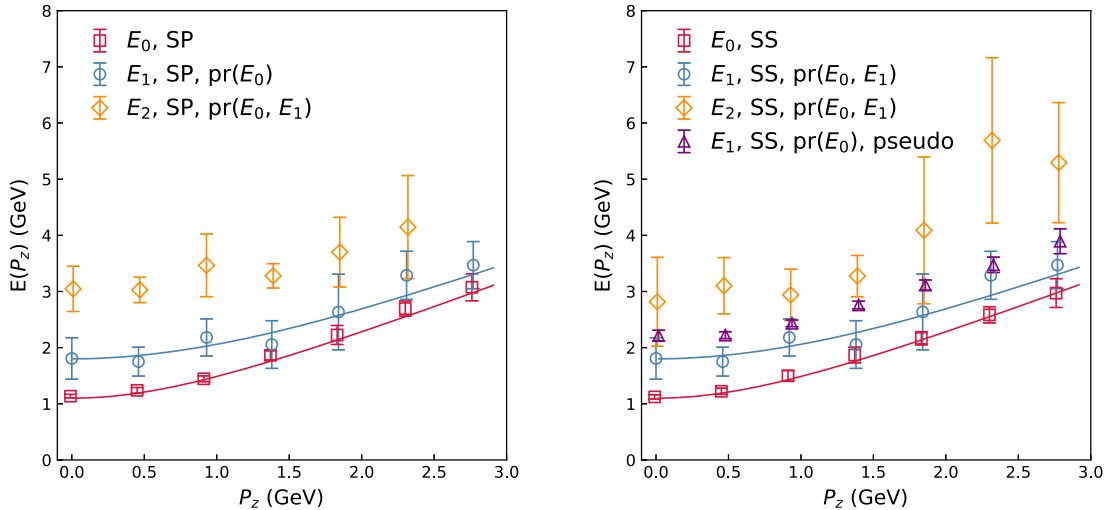


FIG. 3. The energies of different states as functions of  $P_z$ . In the left panel, the  $P_z$  dependence of  $E_0$ ,  $E_1$ , and  $E_2$  for SP correlators are shown. The values of  $E_1$  were obtained from a two-state fit with a prior only on  $E_0$ , and those of  $E_2$  from a three-state fit with priors on both  $E_0$  and  $E_1$ . In the right panel, the  $P_z$  dependence of true values (blue points) and effective values of  $E_1$  (purple points) for SS correlators are both shown (see text for details). The lines indicate the corresponding continuum dispersion relations.

blue points ( $E_1$ ) and orange points ( $E_2$ ) in the left panel of Fig. 3. Our results for the energy levels obtained from SS as functions of  $P_z$  are summarized in the right panel of Fig. 3. Here, the effective values of  $E_1$  from the pseudoplateau and the true values are both shown. The main point of the elaborate analysis is that even though a third excited state contributes in the relatively shorter range of  $t_s$  we use in the paper, it is possible to describe the SS correlator very well by a two-state form with an effective value of  $E_1$ , which is larger than the energy of the physical excited state. Further details on the analysis of the two-point functions are provided in Appendix A.

#### IV. NUCLEON THREE-POINT CORRELATORS

In order to obtain the nucleon qPDF matrix element, we consider the ratio of the three-point function to the two-point function,  $R(z, P_z; t_s, \tau)$ , at different source sink separations,  $t_s$ , and operator insertion,  $\tau$ . At fixed  $(z, P_z)$ , we are interested in fitting the  $(t_s, \tau)$  dependence as expected from the spectral decomposition of  $R$ . If only two states contribute to the correlation functions, the dependence of this ratio on  $\tau$  and  $t_s$  is given by the following form:

$$R_3^{\text{fit}}(t_s, \tau) = \frac{B_0 + e^{-\Delta E t_s/2} (B_1 e^{-\Delta E(t_s/2-\tau)} + B_2 e^{\Delta E(t_s/2-\tau)}) + B_3 e^{-\Delta E t_s}}{1 + \frac{A_1}{A_0} e^{-\Delta E t_s}}. \quad (14)$$

Here  $B_0$  is the desired matrix element  $h$ , and  $\Delta E = E_1 - E_0$ . Generically,  $B_1$  and  $B_2$  are independent fit parameters, except at  $z = 0$ , where  $B_1 = B_2$ . If we assume that the terms proportional to  $A_1$  are small, the denominator can be expanded to leading order to obtain a simpler form:

$$R_2^{\text{fit}}(t_s, \tau) = B_0 + e^{-\Delta E t_s/2} (B_1 e^{-\Delta E(t_s/2-\tau)} + B_2 e^{\Delta E(t_s/2-\tau)}) + B_3 e^{-\Delta E t_s}. \quad (15)$$

Finally, if the term proportional to  $B_3$  is also small compared to other terms, we get an even simpler expression that depends only on three parameters,  $B_0$ ,  $B_1$ , and  $B_2$ :

$$R_1^{\text{fit}}(t_s, \tau) = B_0 + e^{-\Delta E t_s/2} (B_1 e^{-\Delta E(t_s/2-\tau)} + B_2 e^{\Delta E(t_s/2-\tau)}). \quad (16)$$

For each  $(z, P_z)$ , we fit the  $(t_s, \tau)$  dependence of  $R(z, P_z; t_s, \tau)$  to Eqs. (14), (15), and (16) and determine

$B_0$  in each case. In all these fits, we use a fixed value of  $\Delta E(P_z) = E_1(P_z) - E_0(P_z)$ , with the pseudoplateau values of  $E_1(P_z)$  and the ground-state energies  $E_0(P_z)$  determined from the two-point SS correlation function, as shown in the right panel of Fig. 3.

In the following, we discuss the ratio of the three-point function to the two-point function,  $R(z, P_z; t_s, \tau)$ , and the corresponding fits for  $\Gamma = \gamma_t$  and  $n_z = 4$ . In Fig. 4, we show the lattice data on this ratio, together with the fit results for two representative values of  $z$ , namely  $z = 0$  and  $z = 8a$ . The  $t_s$  dependence of the lattice results is small compared to the statistical errors. In particular, the difference between  $t_s = 12a$  and  $t_s = 16a$  data is quite small. This means that the contribution of the excited states is not large even though the source-sink separation is below 1 fm. Given that the  $t_s$  dependence of the ratio is small, it is natural to set  $B_3 = 0$ , since the term is suppressed by  $e^{-t_s \Delta E}$ , and to perform fits using  $R_3^{\text{fit}}(t_s, \tau)$ . We perform fits of the lattice results using the three different fit forms

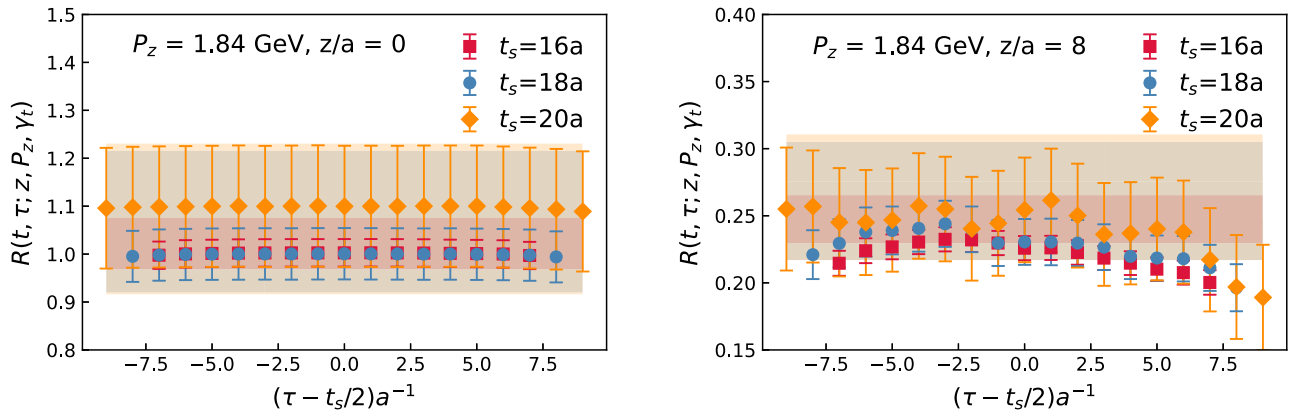


FIG. 4. The real-part ratio of the three-point function to the two-point function at  $z = 0$  (left) and  $z = 8a$  (right) as a function of  $\tau$  and for different  $t_s$  values. The red, orange, and gray bands correspond to the bare matrix elements  $B_0$  extracted from fits to  $R_1^{\text{fit}}$ ,  $R_2^{\text{fit}}$ , and  $R_3^{\text{fit}}$ , respectively.

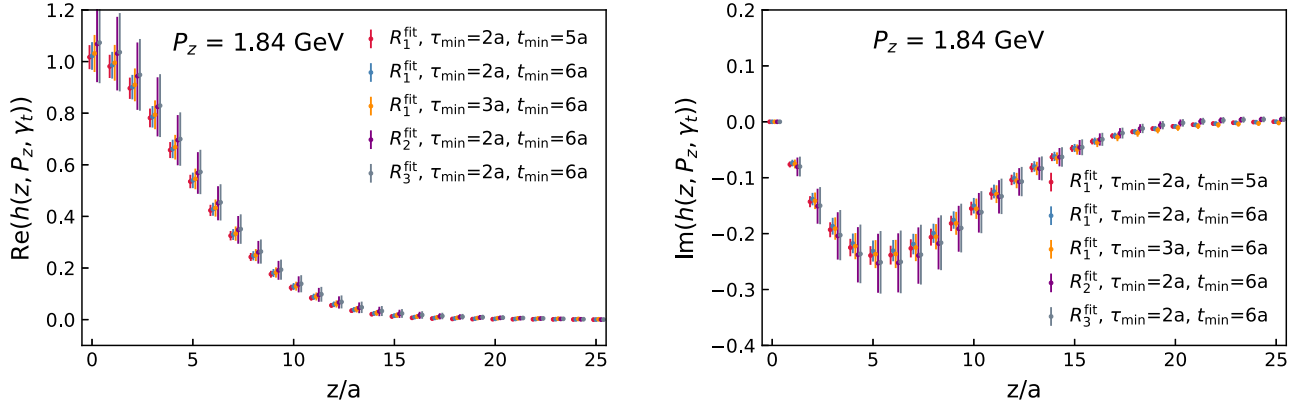


FIG. 5. The  $z$  dependence of the bare matrix element for  $n_z = 4$ . The left panel corresponds to the real part, and the right panel corresponds to the imaginary part. The different colors are the matrix elements obtained by various extrapolation methods (denoted by  $R_1^{\text{fit}}$ ,  $R_2^{\text{fit}}$ , and  $R_3^{\text{fit}}$ ), depending on the number of operator insertion points skipped near the source and sink (denoted by  $\tau_{\min}$ ), and the  $t_{\min}$  value in the two-point function fit from which the excited state  $E_1$  was obtained.

above, and the value of  $\Delta E$  obtained from two-state fits of the two-point functions with  $t_{\min} = 6a$ . We use the operator insertion time  $\tau \geq \tau_{\min} = 2a$  in the fits. The matrix elements,  $B_0$ , obtained using the three fit functions agree within errors. The real and imaginary parts of the ratio  $R$  should be symmetric and antisymmetric with respect to  $z = 0$  at any fixed  $t_s$  and  $\tau$ . In general, our lattice data are compatible with this expectation. Hence, we symmetrize and antisymmetrize the real and imaginary parts of the ratio with respect to  $z = 0$ . The  $z$  dependence of the bare matrix elements is shown in Fig. 5 for all three types of fits. We see again that all three fits give consistent results within errors. Since  $B_0$  obtained from  $R_3^{\text{fit}}$  and  $R_2^{\text{fit}}$  is consistent with that obtained from  $R_1^{\text{fit}}$ , but with larger errors, in the following we will focus on the results obtained from  $R_1^{\text{fit}}$ . We also carry out additional checks for any systematic effects, as discussed below.

We perform several checks to understand the systematic effects in  $R_1^{\text{fit}}$ . First, we study the dependence of the extracted matrix element on  $\tau_{\min}$  and find no significant dependence on it. Second, we perform the fits using only a single source-sink separation  $t_s$  and compare the corresponding results from the three values of  $t_s$ . Interestingly, the matrix elements calculated for source-sink separations  $t_s = 16a, 18a$ , and  $20a$  agree within errors, though the  $t_s = 20a$  results have very large errors. We also study the variation of the extracting matrix element on  $\Delta E$  by using  $E_1$  obtained using different values of  $t_{\min}$ . We find no significant variation. Finally, we use the summation method to obtain the matrix element. This determination has very large statistical errors, but it is still compatible with all other determinations. The above checks of systematic effects are discussed further in Appendix B. We perform a similar analysis for the three-point functions corresponding to the helicity qPDF—i.e., for  $\Gamma = \gamma_z \gamma_5$ . Details of those analyses are also discussed in Appendix B.

## V. NONPERTURBATIVE RENORMALIZATION

We calculate the nonperturbative renormalization of the qPDF operator in the RI-MOM scheme using off-shell quark states in the Landau gauge [20,22]. The matrix element of  $O_\Gamma(z)$  in an off-shell quark state  $|p\rangle$  is

$$\Lambda(p, z, \Gamma) = \langle S(p) \rangle^{-1} \left\langle \sum_w \gamma_5 S^\dagger(p, w + zn) \gamma_5 \Gamma W_z(w + zn, w) S(p, w) \right\rangle \langle S(p) \rangle^{-1}, \quad (17)$$

where  $n^\mu = (0, 0, 0, 1)$  is the unit vector along the  $z$  direction, and the summation is over all lattice sites  $w$ . The quark propagators are defined as

$$S(p, x) = \sum_y e^{ipy} \langle \bar{\psi}(x) \psi(y) \rangle, \quad S(p) = \sum_x e^{-ipx} S(p, x), \quad (18)$$

and  $\gamma_5$  is inserted on both sides of  $S^\dagger(p, w + zn)$  in Eq. (17) to get the necessary propagator  $\sum_y e^{-ipy} \langle \bar{\psi}(y) \psi(w + zn) \rangle$ .

For the unpolarized qPDF, we use the RI-MOM renormalization constant defined via

$$Z_{\text{mp}}(z, p_z^R, a^{-1}, \mu_R) = \frac{\text{Tr}[\mathcal{P} \Lambda_{\text{tree}}(p, z, \gamma_t)]}{\text{Tr}[\mathcal{P} \Lambda(p, z, \gamma_t)]} \Big|_{p^2 = \mu_R^2, p_z = p_z^R}, \quad (19)$$

where  $\Lambda_{\text{tree}}(p, z, \gamma_t) = \gamma_t e^{-izp_z}$  is the tree-level matrix element in the momentum space. Furthermore,  $\mathcal{P} = \gamma_t - (p_t/p_x) \gamma_x$  is the projection operator corresponding to the so-called minimal projection, where only the term with the Dirac structure proportional to  $\gamma_t$  is kept [23,24]. Hence, we use the subscript “mp” for the renormalization constant. The renormalization constant  $Z_{\text{mp}}(z, p_z^R, a^{-1}, \mu_R)$  depends

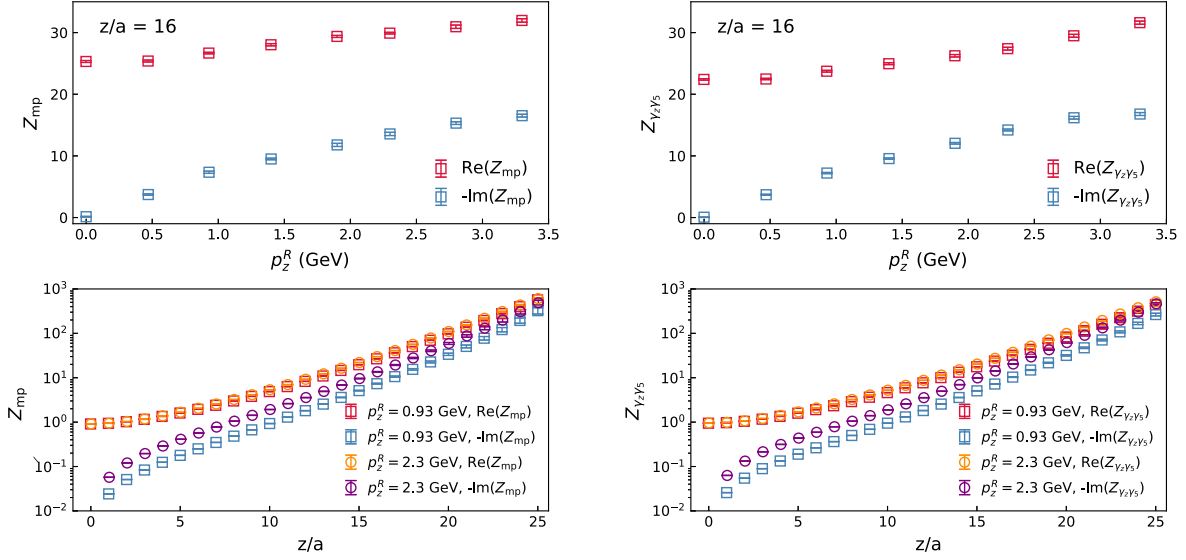


FIG. 6. The renormalization constant  $Z_{\text{mp}}$  and  $Z_{\gamma_z\gamma_5}$  at  $\mu_R = 4$  GeV. The upper panels show  $Z_{\text{mp}}$  and  $Z_{\gamma_z\gamma_5}$  as functions of  $p_z^R$  at  $z = 16a \approx 0.67$  fm. The lower panels show the  $z/a$  dependence with  $p_z^R = 0.93$  GeV and 2.3 GeV.

on the lattice spacing  $a$ , as well as the other two scales  $p_z^R$  and  $\mu_R$ .

We follow a similar procedure for the longitudinally polarized case, where the RI-MOM renormalization constant is defined as

$$Z_{\gamma_z\gamma_5}(z, p_z^R, a^{-1}, \mu_R) = \frac{\text{Tr}[\mathcal{P}\Lambda_{\text{tree}}(p, z, \gamma_z\gamma_5)]}{\text{Tr}[\mathcal{P}\Lambda(p, z, \gamma_z\gamma_5)]} \Big|_{p^2=\mu_R^2, p_z=p_z^R}, \quad (20)$$

with  $\Lambda_{\text{tree}}(z, p_z, \gamma_z\gamma_5) = \gamma_z\gamma_5 e^{-ip_z z}$ . The projection operator  $\mathcal{P}$  in this case is chosen to be  $\mathcal{P} = \gamma_5\gamma_z/4$ .

We calculate the nonperturbative RI-MOM renormalization constants in the Landau gauge. The calculations are performed using 14 gauge configurations. The relative uncertainties of the renormalization constants for  $z = 0, 16$ , and 32 are 0.02%, 1%, and 10%, respectively. Such precision is much better than that of our nucleon matrix elements with the same  $z$ , so it is enough at the present stage. We use the following values of the momenta for the off-shell quark state:  $ap = \frac{2\pi}{L}(5, 5, 5, 0)$ ,  $\frac{2\pi}{L}(6, 2, 1, 17/3)$ , and  $\frac{2\pi}{L}(7, 4, 3, 1/3)$ , with  $L = 64$  being the spatial size of the lattice. These momenta correspond to  $\mu_R = |p| = 3.99$  GeV, 3.94 GeV, and 3.97 GeV—i.e., to  $\mu_R \sim 4$  GeV within 1.5%. Since all the spatial directions are equivalent, each of them could be considered as the  $z$  direction, and therefore, with the above choice of the three momenta we have  $p_z^R = 0.46 \times \{0, 1, 2, \dots, 7\}$  GeV.

The renormalization constant is plotted in Fig. 6. Due to the linear divergence, the renormalization constant can be far from 1 at a large  $z \approx 0.67$  fm, making the nonperturbative renormalization unavoidable. Figure 6 also shows that the renormalization constant will be sensitive to the value of

$p_z^R$ , while such a dependence should be canceled by the matching in the continuum if we have the matching formula up to all orders, because the PDFs or the Mellin moments in the  $\overline{\text{MS}}$  scheme have no dependence on  $p_z^R$ . We will consider the residual  $p_z^R$  dependence in the final PDF prediction as a systematic uncertainty.

Having determined the renormalization constants  $Z_{\text{mp}}$  and  $Z_{\gamma_z\gamma_5}$  we obtain the renormalized matrix elements—i.e., coordinate space qPDF. For the unpolarized case,

$$h_R(z, P_z, \mu_R, p_z^R) = Z_q Z_{\text{mp}}(z, p_z^R, a^{-1}, \mu_R) h(z, P_z, \gamma_t), \quad (21)$$

and for the longitudinally polarized case,

$$\Delta h_R(z, P_z, \mu_R, p_z^R) = Z_q Z_{\gamma_z\gamma_5}(z, p_z^R, a^{-1}, \mu_R) h(z, P_z, \gamma_z\gamma_5). \quad (22)$$

In the above equations,  $Z_q$  is the quark wave function renormalization factor.

In Fig. 7, we show the renormalized matrix elements, modulo the factor  $Z_q$ , in the RI-MOM scheme at  $p_z^R = 0$ ,  $\mu_R = 4$  GeV. We find that the errors are large. We can achieve substantial error reductions at  $z \neq 0$  by redefining the renormalized matrix elements as

$$h_R(z, P_z, \mu_R, p_z^R) \equiv \frac{h_R(z, P_z, \mu_R, p_z^R)}{h_R(z=0, P_z, \mu_R, p_z^R)} \quad \text{and} \quad \Delta h_R(z, P_z, \mu_R, p_z^R) \equiv \frac{\Delta h_R(z, P_z, \mu_R, p_z^R)}{\Delta h_R(z=0, P_z, \mu_R, p_z^R)}. \quad (23)$$

The errors of the matrix elements for  $z \neq 0$  are reduced due to the strong correlations between  $z \neq 0$  (particularly for

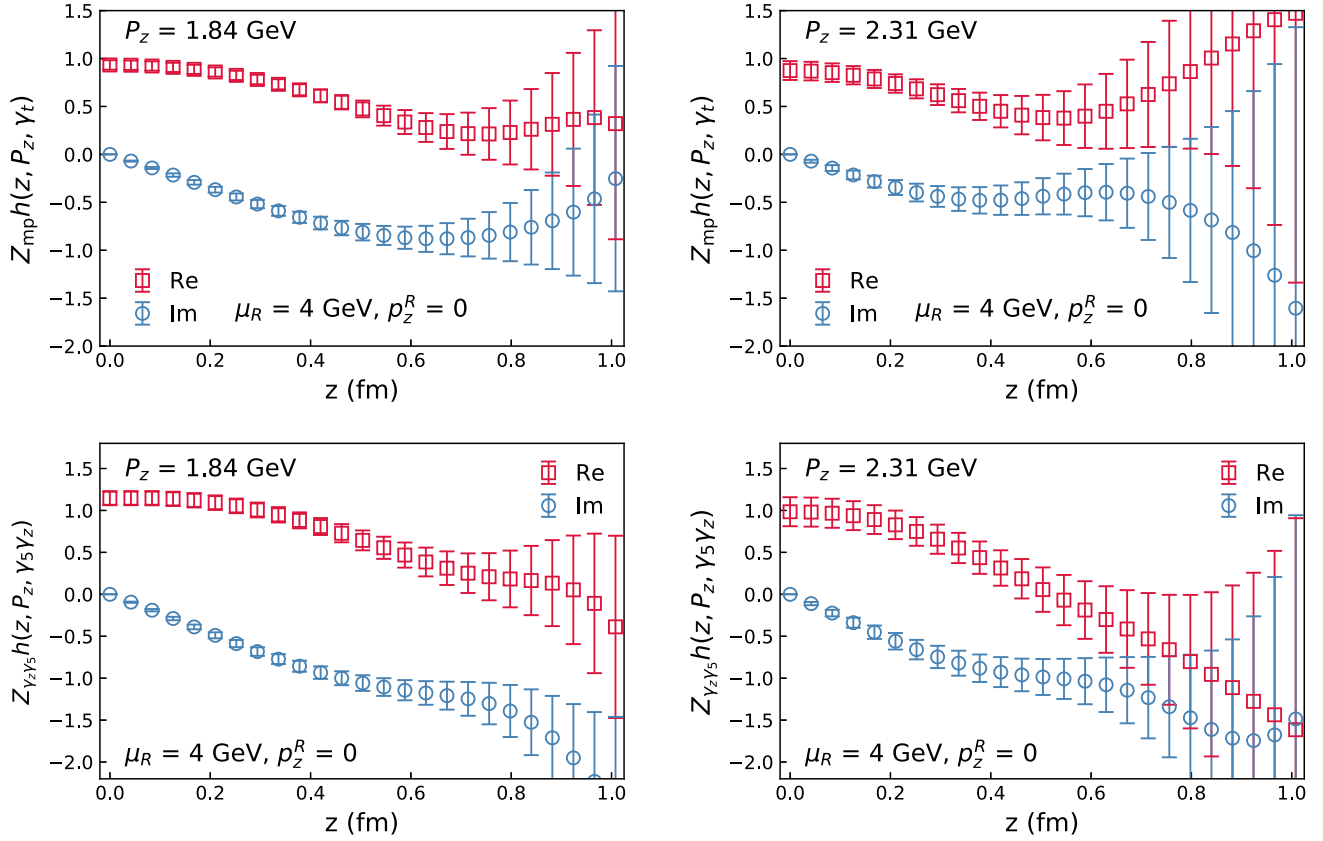


FIG. 7. Top panels: the  $z$  dependence of the real and imaginary parts of the RI-MOM renormalized (modulo the wave function renormalization,  $Z_q$ ), unpolarized qPDF matrix element for  $P_z = 1.84$  GeV (left) and  $2.31$  GeV (right). Bottom panels: similar results for the real and imaginary parts of the helicity matrix element.

small  $z$  close to  $z = 0$ ) and  $z = 0$  matrix elements for each gauge configuration. The effectiveness of this procedure in can be seen from Figs. 12 and 16. As one can see, the error reduction due to this division is very significant. In fact, with this method, the errors are reduced enough that the  $z$  dependence of the matrix element is well constrained also for  $n_z = 5$ . Since for the extraction of the qPDF we are only interested in the  $z$  dependence of the matrix element, and we know that the unpolarized isovector nucleon matrix element at  $z = 0$  is the isospin of the nucleon, which is unity [in our convention, cf. Eq. (5)] after renormalization, we can consider the above improved ratio of renormalized matrix elements. However, the effect of taking this ratio is not trivial in the case of the matrix element of the helicity qPDF—the value of the renormalized matrix element at  $z = 0$  should be  $g_A \approx 1.3$ ; this procedure is equivalent to studying a helicity PDF with the first moment normalized to unity, i.e., in a normalization where  $g_A = 1$ .

## VI. UNPOLARIZED PDF: PERTURBATIVE MATCHING AND COMPARISONS WITH $h_R(z, P_z)$

In this section, we will discuss how the renormalized coordinate space qPDF,  $h_R(z, P_z, \mu_R, p_z^R)$ , can be related and compared with phenomenological unpolarized nucleon

PDFs, such as the CT18 [57] and NNPDF3.1 [58], extracted from the global analysis of experimental data. The unpolarized quark PDF in the valence region is well constrained through global analysis. Therefore, it is natural to start from these phenomenological PDFs as a function of Bjorken- $x$ , use the perturbative matching to reconstruct the corresponding coordinate space qPDF as a function of  $z$  for different  $P_z$  values, and compare with our results for  $h_R(z, P_z, \mu_R, p_z^R)$ . The reason for comparing in the  $z$ -space, rather than constructing the  $x$ -dependent PDF from our  $h_R(z, P_z, \mu_R, p_z^R)$  and then comparing with the phenomenological PDFs, is the following: As can be seen from Figs. 12 and 16,  $h_R(z, P_z, \mu_R, p_z^R)$  is quite noisy for  $z \geq 0.5$  fm. Thus, the Fourier transformation which is needed to calculate the qPDF in  $x$  space is difficult to perform. A similar approach had also been used for the pion PDF [35].

Even at the leading  $\alpha_s^0$  order, the qPDF and the PDF differ due to the trace term in the small- $z$  expansion [12, 25]. This difference was explicitly calculated in Ref. [59]. In the context of DIS, such corrections have been studied long ago [60], and are known as target-mass corrections. Following Ref. [59], we introduce the target-mass-corrected PDF

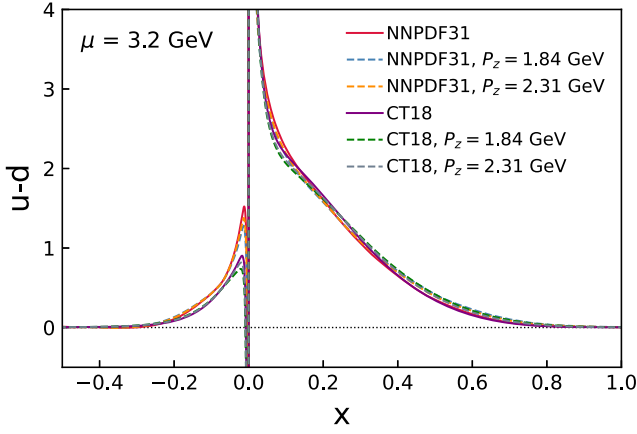


FIG. 8. The NNLO isovector nucleon PDFs CT18 [57] and NNPDF3.1 [58] (solid lines), and the corresponding target-mass-corrected ones (dashed lines), at a scale  $\mu = 3.2$  GeV. See text for details.

$$q'(x, P_z) = \frac{1}{\sqrt{1+4c}} \left[ \frac{f_+}{2} q\left(\frac{2x}{f_+}\right) - \frac{f_-}{2} q\left(\frac{-2x}{f_-}\right) \right], \quad (24)$$

where  $c = M^2/(4P_z)^2$ ,  $f_{\pm} = \sqrt{1+4c} \pm 1$ , and  $q(x)$  is the usual PDF that corresponds to  $P_z \rightarrow \infty$ . In our analysis, we use two sets of NNLO PDF for the  $u$  and  $d$  quark and antiquark distributions, the CT18 [57] and NNPDF3.1 [58], evaluated at the scale  $\mu = 3.2$  GeV. If the matching were known to all orders of perturbation theory, the prediction for real-space qPDF should be independent of the value of  $\mu$  at which the PDF was evaluated. Since the matching is only known to one-loop order, we choose a scale  $\mu = 3.2$  GeV that is of the same order as the other momentum scales used in our computations, and thereby avoided corrections due to large logarithms. The light-cone quark PDF for  $u$  quarks is calculated as  $q_u(x) = u(x)$ ,  $x > 0$  and

$q_u(x) = -\bar{u}(-x)$ ,  $x < 0$ . The isovector nucleon PDF,  $q'_u(x) - q'_d(x)$  is shown in Fig. 8.

In Fig. 8, we also show the target-mass-corrected isovector nucleon PDF for the two momenta used in our study, namely 1.84 GeV and 2.31 GeV. We see from the figure that target-mass correction is small for the values of  $P_z$  used in this study. Using the target-mass-corrected NNPDF3.1 isovector nucleon PDF obtained from Eq. (24) and the one-loop matching to RI-MOM, we obtain the corresponding qPDF for  $P_z = 1.84$  GeV and  $P_z = 2.31$  GeV,  $\mu_R = 4$  GeV, and  $p_z^R = 0, 0.93, 1.9$  GeV. The functions  $f_{1,\gamma_1}$  and  $f_{2,\gamma_1,mp}$  in Eq. (11) for the one-loop matching to the RI-MOM scheme with minimal projection were taken from Eqs. (28) and (31) of Ref. [24]. Figure 9 shows comparisons of the NNPDF3.1 with the corresponding qPDFs. In these comparisons,  $\alpha_s$  is evaluated at the scale  $\mu = 3.2$  GeV, which results in the value  $\alpha_s = 0.25$ . We see significant differences between the PDF and qPDF. For large positive  $x$ , the qPDF is larger than the PDF, while for negative  $x$  the qPDF can turn negative for some  $P_z^R$ . The qPDF strongly depends on the choice of the RI-MOM scales. It is possible to choose the RI-MOM scale such that the qPDF is negative for  $x < -0.2$ , even though the PDF is positive.

By Fourier-transforming the CT18 and NNPDF3.1 target-mass-corrected qPDFs with respect to  $x$ , we obtain the corresponding distributions as a function of the so-called Ioffe time,  $zP_z$ —i.e., the corresponding ITDs [61]. Since the matching is only up to one-loop order, the scale entering  $\alpha_s$  is not fixed. We consider three choices of the scale for  $\alpha_s$ , namely  $\mu/2$ ,  $\mu$ , and  $2\mu$ . The corresponding variations in the ITDs can be considered as estimates of the perturbative uncertainties, and are shown as bands in Fig. 10. In the same figure, we also compare with the lattice results for the ITDs in RI-MOM renormalization, at the renormalization scales of  $\mu_R = 4$  GeV and  $p_z^R = 0$  GeV. Notwithstanding the large errors, for both values of  $P_z$  the real parts of the ITDs compare well at least up to  $zP_z \lesssim 5$ . However, lattice results

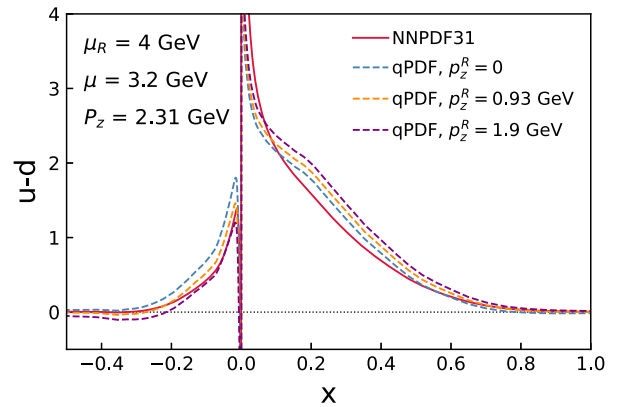
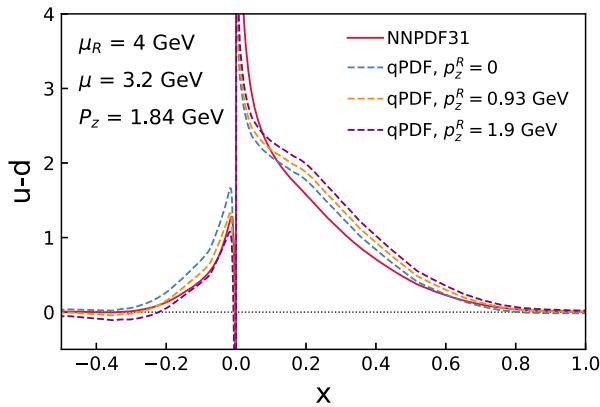


FIG. 9. qPDF corresponding to NNPDF3.1 for  $P_z = 1.84$  GeV (left) and  $P_z = 2.3$  GeV (right) with  $\alpha_s = 0.25$  and three different RI-MOM renormalization conditions.

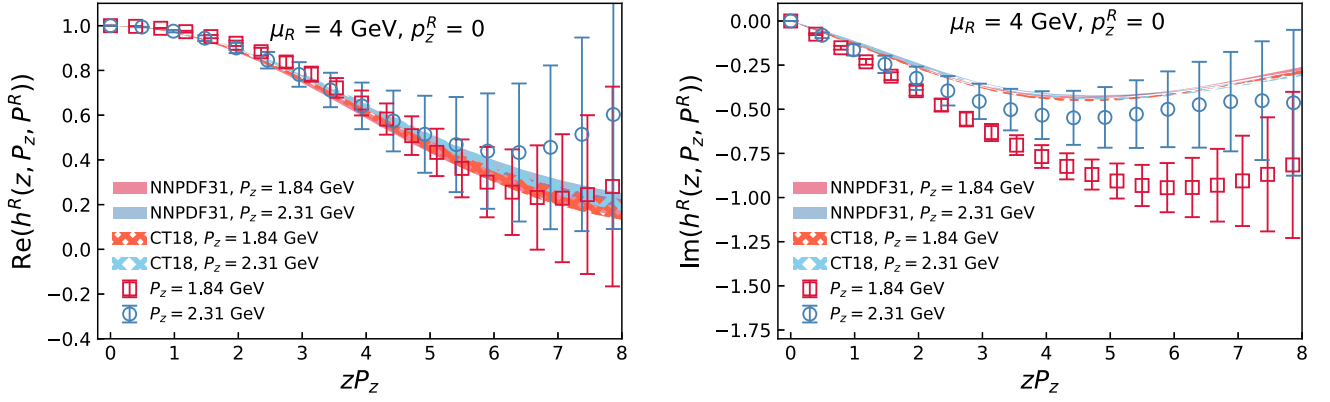


FIG. 10. Comparisons of the real (left panel) and imaginary (right panel) parts of the ITDs, in the RI-MOM renormalization at the scales  $p_z^R = 0$  and  $\mu_R = 4$  GeV, with that obtained from the CT18 and NNPDF3.1 unpolarized isovector nucleon PDFs.

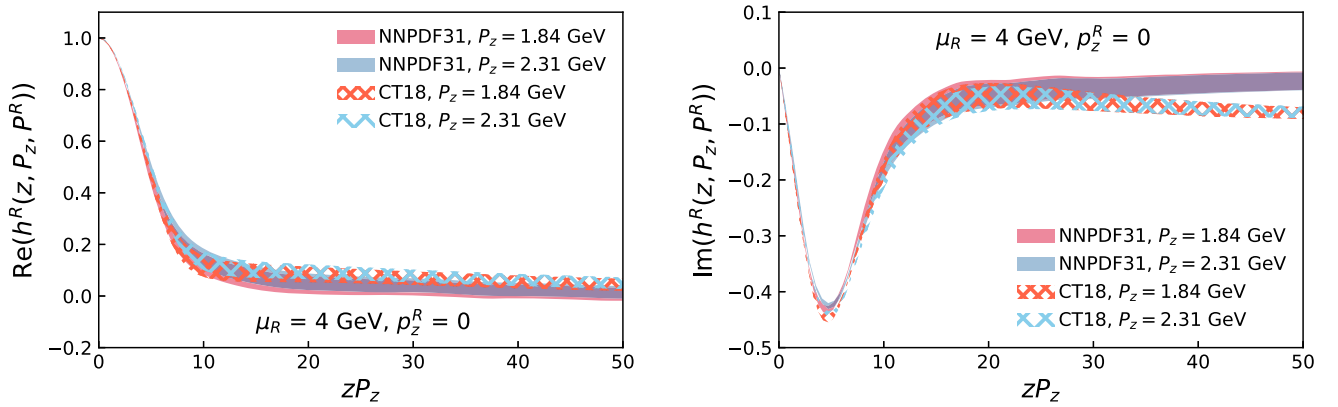


FIG. 11. Real (left) and imaginary (right) parts of the ITDs corresponding to target-mass-corrected CT18 and NNPDF3.1 PDFs, for  $p_z^R = 0$ , in an extended range of the Ioffe time.

for the imaginary parts of ITDs undershoot the phenomenological ITDs even for  $zP_z \gtrsim 2$ .

Although there is a significant difference between the CT18 and NNPDF3.1 PDFs in the small- $x$  region, Fig. 10 does not show any visible difference in their corresponding ITDs. To understand this better, in Fig. 11 we explore these ITDs in an extended range of Ioffe time. The difference between the PDFs in the negative- $x$  region is only reflected in a  $< 10\%$  difference in the imaginary part of the ITDs for  $zP_z > 25$ , essentially showing no difference in the real-part ITDs even up to  $zP_z = 50$ .

To explore the dependence of the lattice results on the choice of RI-MOM scale  $p_z^R$  and the range of validity of the one-loop matching, in Fig. 12 we show comparisons between the qPDFs as a function of  $z$  obtained in the lattice calculations and from the global analysis of the PDF for two different choices of renormalization scale, namely  $p_z^R = 0$  and  $0.93$  GeV. Very little dependence on  $p_z^R$  was observed. While the real part of the qPDF obtained from the

global analysis agrees with the lattice results up to  $z \sim 1$  fm within relative large errors, the agreement is limited only for  $z \lesssim 0.2$  fm. For  $P_z = 2.31$  GeV, the agreements seem to extend to larger values of  $z$ , partly because of larger errors. However, it is encouraging that the central value seems to shift towards the global analysis results as  $P_z$  is increased from  $1.84$  to  $2.31$  GeV. In any case, at large  $z$ , we see clear tension between the imaginary part of the lattice qPDF and the results of global analysis. This suggests that the range of applicability of one-loop matching is perhaps limited to  $z \lesssim 0.2$  fm in the case of the nucleon. It remains to be seen if this agreement gets better with the addition of higher-loop corrections, or if this observed discrepancy arises because of contamination of higher-twist effects at larger  $z$ . This observation has an important implication for our ability to describe the  $x$  dependence of PDF within the LaMET framework. For example, if the one-loop perturbative matching works only for  $z \simeq 0.2$  fm, reliable calculations of nucleon PDF down to  $x \simeq 0.1$  will need  $P_z \gtrsim 10$  GeV.

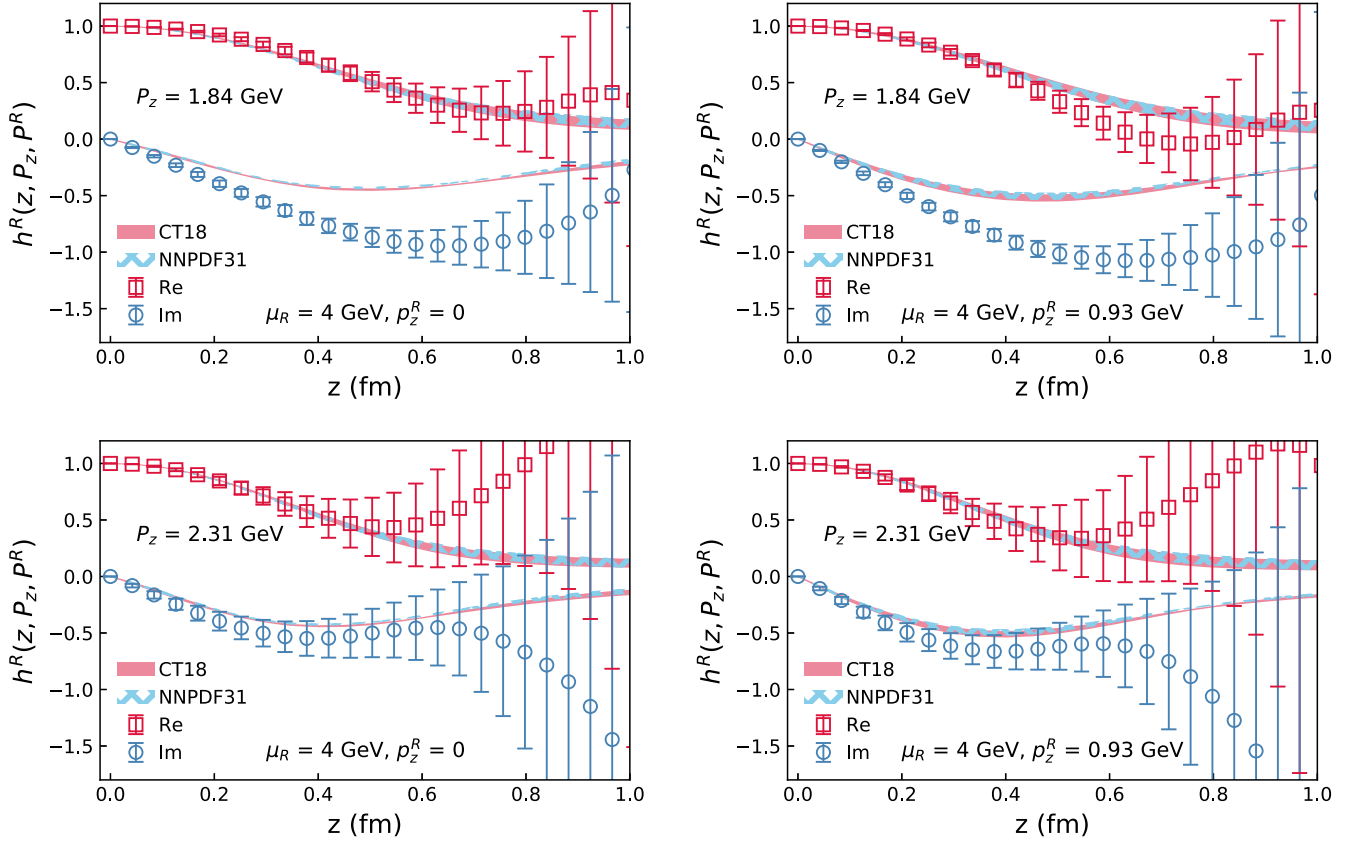


FIG. 12. Comparisons of the qPDF with those obtained from the global analysis for two values of the RI-MOM renormalization scale,  $p_z^R = 0$  GeV (left row) and  $p_z^R = 0.93$  GeV (right row), and for two values of the nucleon boost momenta,  $P_z = 1.84$  GeV (upper column) and  $P_z = 2.31$  GeV (lower column).

## VII. HELICITY PDF: PERTURBATIVE MATCHING AND COMPARISONS WITH $\Delta h_R(z, P_z)$

Our analysis of helicity qPDF closely follows the analysis performed in the unpolarized case; namely, we start from the helicity PDF obtained in global analyses, reconstruct the corresponding target-mass-corrected qPDF, and then compare with the lattice results. The helicity PDFs have been extracted from the global analysis by the NNPDF Collaboration using DIS, inclusive  $W^\pm$  and jet production data from RHIC, as well as the open charm data

from COMPAS, resulting in NNPDFpol1.1 [62]. The JAM Collaboration used the DIS and SIDIS data in their global analysis, combined with  $e^+e^-$  data to constrain the fragmentation functions at NLO [63]. The resulting PDF parametrization is called JAM17. In Fig. 13, we show the isovector helicity PDF  $\Delta q_u - \Delta q_d$ . The positive- $x$  region corresponds to the quark contribution, while the negative- $x$  region corresponds to the antiquark region. The target-mass-corrected helicity PDF,  $\Delta q'(x, P_z)$ , is obtained from helicity PDF,  $\Delta q(x)$ , following Ref. [59]:

$$\Delta q'(x, P_z) = \frac{1}{1+4c} \left[ \frac{f_+}{2} \Delta q\left(\frac{2x}{f_+}\right) + \frac{f_-}{2} \Delta q\left(\frac{-2x}{f_-}\right) \right] - \int_{\pm\infty}^x \frac{2c}{(1+4c)^{3/2}} \left[ \Delta q\left(\frac{2y}{f_+}\right) + \Delta q\left(\frac{-2y}{f_-}\right) \right], \quad (25)$$

where  $c = M^2/4P_z^2$ ,  $f_\pm = \sqrt{1+4c} \pm 1$ , and for the integration limits  $+\infty$  ( $-\infty$ ) corresponds to  $x > 0$  ( $x < 0$ ).

Although the matching for helicity qPDF has not been explicitly presented in the literature before, it is straightforwardly deduced from the results presented in Ref. [24]. The key observation here is the fact that, owing to the chiral symmetry, for massless quarks in one-loop perturbation

theory,  $\text{Tr}[\gamma_5 \gamma_z \Lambda(p, z, \gamma_z \gamma_5)] = \text{Tr}[\gamma_z \Lambda(p, z, \gamma_z)]$ . Thus, the one-loop matching of the helicity qPDF in the RI-MOM scheme with minimal projection is same as that for the unpolarized qPDF with  $\Gamma = \gamma_z$  (instead of the  $\Gamma = \gamma_t$  used before), and with the RI-MOM renormalization condition corresponding to the projection operator  $\mathcal{P} = \gamma_z$  (instead of the minimal projection). The one-loop matching for the

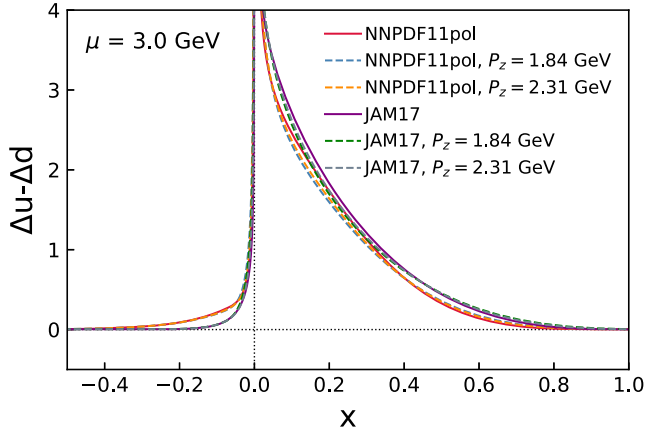


FIG. 13. NNPDF11pol and JAM17 isovector helicity PDFs at a scale  $\mu = 3$  GeV. Also shown are the corresponding target-mass-corrected isovector helicity PDFs (dashed lines) for  $P_z = 1.84$  GeV and  $P_z = 2.31$  GeV.

$\Gamma = \gamma_z$  operator is known for two different RI-MOM projections, the minimal projection and the  $\not{p}$  projection, corresponding to  $\mathcal{P} = \gamma_z - (p_z/p_x)\gamma_x$  and  $\mathcal{P} = \not{p}/(4p_z)$ , respectively [24]. The function that depends on the RI-MOM projection operator—i.e.,  $f_{2,\gamma_z,\gamma_z}$ —entering the matching coefficient in Eq. (11) was simply deduced from these known results. The Lorentz structure of  $\Lambda(p, z, \gamma^\alpha)$  for a general  $\gamma_\alpha$ ,  $\alpha = x, y, z, t$  is given by

$$\begin{aligned} \Lambda^{(1)}(p, x, \gamma_\alpha) = & \gamma_\alpha [\tilde{f}_t(x, \rho)]_+ \\ & + \gamma_z \frac{p_\alpha}{p_z} [\tilde{f}_z(x, \rho)]_+ + \frac{\not{p} p_\alpha}{p^2} [\tilde{f}_p(x, \rho)]_+, \end{aligned} \quad (26)$$

and  $f_{2,\gamma_z,\text{mp}} = \tilde{f}_t + \tilde{f}_z$  and  $f_{2,\gamma_z,\not{p}} = \tilde{f}_t + \tilde{f}_z + \tilde{f}_p$  [24]. Here, the subscript “+” refers to the standard plus-prescription and  $\rho = -p^2/p_z^2$ . The functions  $\tilde{f}_t$ ,  $\tilde{f}_z$ , and  $\tilde{f}_p$  have been calculated in Ref. [24], and we use the same

notations here. Therefore, for the case of  $\mathcal{P} = \gamma_z$  the RI-MOM projection-dependent function is given by

$$\begin{aligned} f_{2,\gamma_z,\gamma_z} &= \tilde{f}_t + \tilde{f}_z + (p_z^2/p^2)\tilde{f}_p \\ &= f_{2,\gamma_z,\text{mp}} + (f_{2,\gamma_z,\not{p}} - f_{2,\text{mp}})/r. \end{aligned} \quad (27)$$

Thus, for the helicity qPDF, the one-loop matching RI-MOM function in the minimal projection scheme is the same as in Eq. (11), but with  $f_{2,\gamma_z,\text{mp}}$  given by Eq. (27), and with  $f_{1,\gamma_z}$ ,  $f_{2,\gamma_z,\text{mp}}$ , and  $f_{2,\gamma_z,\not{p}}$  given by Eqs. (A6-A8) of Ref. [24].

Using the matching discussed above, we can obtain the isovector helicity qPDF from the target-mass-corrected NNPDF11pol and JAM17. As before, for the one-loop matching we use  $\alpha_s$  evaluated at the scale  $\mu = 3.0$  GeV, and the scale is varied between  $\mu/2$  and  $2\mu$  to estimate the scale uncertainty. We find a noticeable difference between the isovector helicity PDFs and the corresponding qPDFs in Fig. 14. By Fourier-transforming the qPDFs, we obtain the isovector helicity ITDs and compare them with our lattice results in Fig. 15. Since we normalized our lattice results by the value of matrix element at  $z = 0$ , we normalize the phenomenological ITDs by dividing with  $g_A = 1.25$ . Within the large statistical errors, we do not find a significant  $P_z$  dependence of the lattice results. While the real parts of the lattice results agree with those obtained from the phenomenological PDFs up to  $zP_z \lesssim 3$ , the imaginary parts do not agree quantitatively but also have larger errors. We also explore the dependence of our result on the choice of RI-MOM scales. In Fig. 16, we compare the qPDFs for  $\mu_R = 4$  GeV, and for  $p_z^R = 0$  GeV and  $p_z^R = 0.93$  GeV. From the figure, we see that the comparison between the results of lattice calculation, as well as the fact that the global analyses are not sensitive to  $P_z$ , the agreement between the lattice and the global analyses extends to values of  $|z|$  of about 0.3 fm for the real parts, but not for the imaginary parts. In the next section, we will discuss how these disagreements show up in the moments of the PDFs.

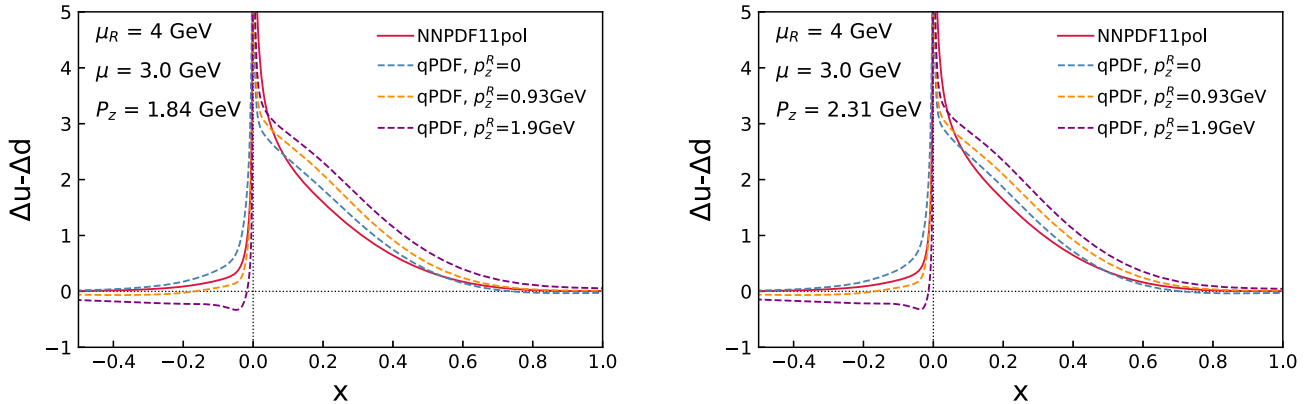


FIG. 14. qPDF corresponding to NNPDF1.1pol for  $P_z = 1.84$  GeV (left) and  $P_z = 2.3$  GeV (right) with  $\alpha_s = 0.25$  and three different RI-MOM renormalization conditions.

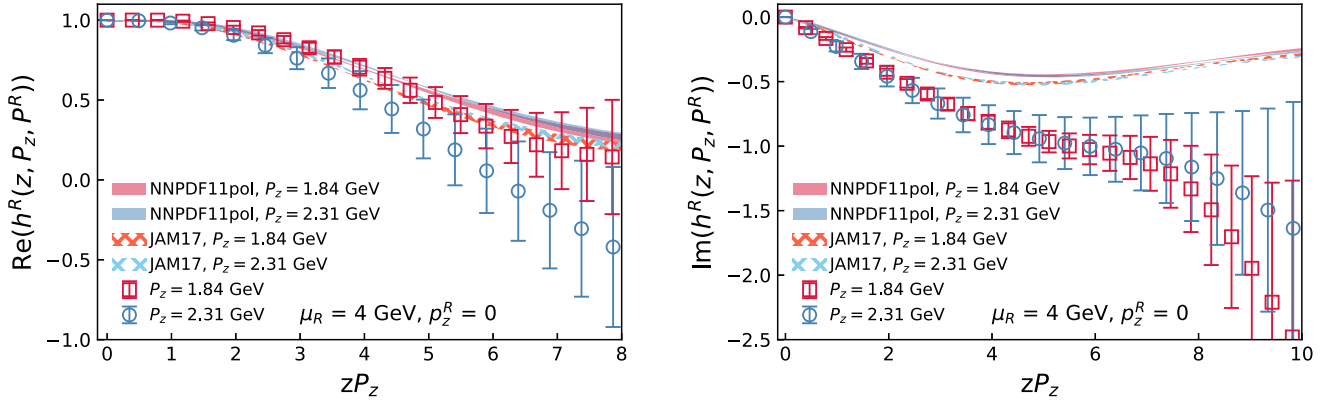


FIG. 15. Comparisons of the real (left panel) and imaginary (right panel) parts of the isovector helicity ITDs, in the RI-MOM renormalization at the scales  $p_z^R = 0$  and  $\mu_R = 4$  GeV, with that obtained from NNPDF1.1pol and JAM17.

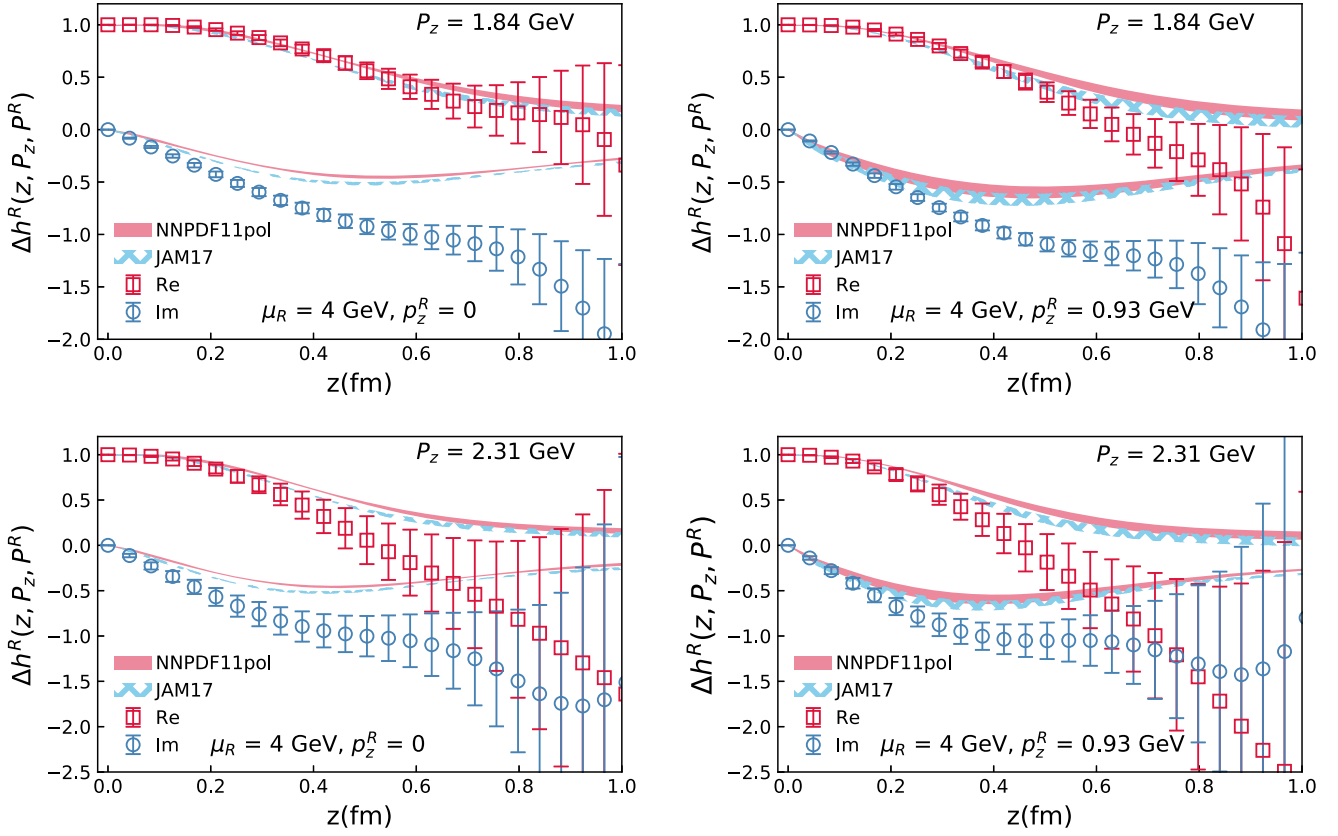


FIG. 16. Comparisons of the isovector helicity qPDF with those obtained from the global analysis for two values of the RI-MOM renormalization scale,  $p_z^R = 0$  GeV (left row) and  $p_z^R = 0.93$  GeV (right row), and for two values of the nucleon boost momenta,  $P_z = 1.84$  GeV (upper column) and  $P_z = 2.31$  GeV (lower column).

### VIII. MOMENTS OF PDF FROM RATIO OF IOFFE-TIME DISTRIBUTIONS

In the previous sections, we analyzed the boosted nucleon matrix elements renormalized in the RI-MOM scheme and matched them to the PDFs in the  $\overline{\text{MS}}$  scheme. Due to the multiplicative renormalizability of  $h(z, P_z, \gamma_t)$

and  $h(z, P_z, \gamma_z \gamma_5)$ , we can form well-defined renormalized quantities by taking the ratios of matrix elements at two different momenta  $P_z$  and  $P'_z$  as

$$\mathcal{M}(z, P_z, P'_z, \Gamma) = \frac{h(z, P_z, \Gamma) h(0, P'_z, \Gamma)}{h(z, P'_z, \Gamma) h(0, P_z, \Gamma)}. \quad (28)$$

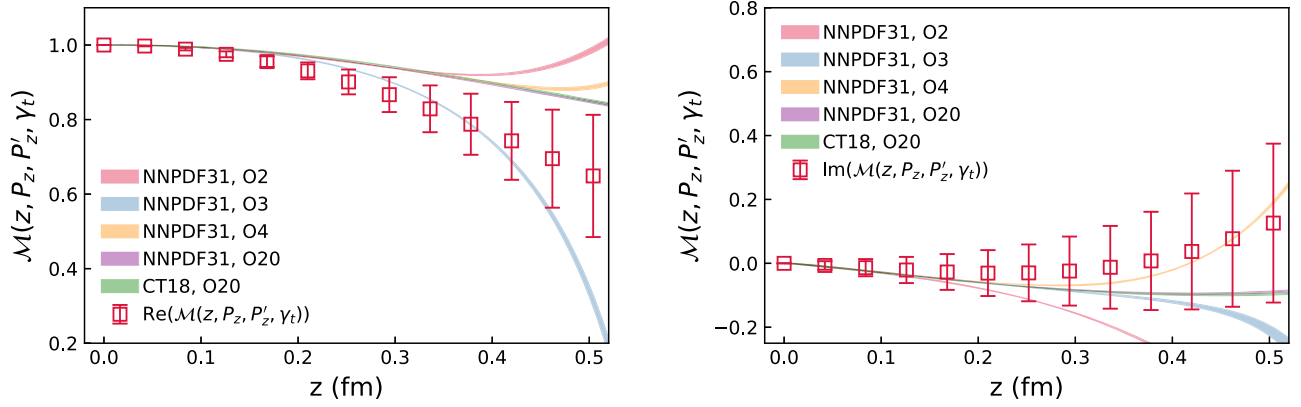


FIG. 17. The real (left) and imaginary (right) parts of  $\mathcal{M}(z, P_z, P'_z, \gamma_t)$  are shown for  $P_z = 2.31$  GeV and  $P'_z = 1.84$  GeV. The data points are from our lattice calculations, whereas the various colored bands are the corresponding results from the isovector unpolarized PDFs from NNPDF3.1 and CT18. The bands in these phenomenological expectations arise due to variations of  $\alpha_s(\mu)$  within the scale  $\mu/2$  to  $2\mu$ . For NNPDF3.1, we also show results by truncating the expansion in Eq. (29) at various orders,  $n = n_{\max}$ , in the PDF moments; these results are denoted by  $On_{\max}$ .

The second factor on the right-hand side of the above definition normalizes the  $z = 0$  matrix element to unity, as we did in the case of the RI-MOM scheme. The choice  $P'_z = 0$  in the ratio is usually referred to as the reduced Ioffe-time distribution [16], and one should think of  $P'_z \neq 0$  as a generalization of this choice. Here, we take  $P_z = 2.31$  GeV and  $P'_z = 1.84$  GeV, respectively. Since both  $P_z, P'_z > \Lambda_{\text{QCD}}$  and the nucleon mass, we expect this ratio to be simply described by the leading twist expression [25],

$$\mathcal{M}(z, P_z, P'_z, \Gamma) = \frac{\sum_{n=0} \frac{c_n(\mu z)}{c_0(\mu z)} \frac{(-izP_z)^n}{n!} \langle x^n \rangle_{P_z}(\mu)}{\sum_{n=0} \frac{c_n(\mu z)}{c_0(\mu z)} \frac{(-izP'_z)^n}{n!} \langle x^n \rangle_{P'_z}(\mu)}. \quad (29)$$

Following Ref. [59], the target-mass-corrected unpolarized PDF moments  $\langle x^n \rangle$  can be obtained by relation:

$$\frac{\langle x^n \rangle_{P_z}}{\langle x^n \rangle} = \sum_{i=0}^{\lfloor (n+1)/2 \rfloor} C_{n-i+1}^i c^i, \quad (30)$$

and for the helicity case,

$$\frac{\langle x^n \rangle_{P_z}}{\langle x^n \rangle} = \sum_{i=0}^{\lfloor n/2 \rfloor} \left( \frac{n-i+1}{n+1} \right) C_{n-i}^i c^i, \quad (31)$$

where  $C_n^i$  is the binomial function,  $c = M^2/4P_z^2$ . In Eq. (29),  $c_n(\mu z)$  is the one-loop-order Wilson coefficient in the  $\overline{\text{MS}}$  scheme. The Wilson coefficient describes the  $z$  dependence of the twist-2 local operator associated with the  $n$ th moment of the PDF,  $\langle x^n \rangle(\mu)$ , in the  $\overline{\text{MS}}$  scheme and at a factorization scale  $\mu$ . As in our RI-MOM analysis, we will use  $\mu = 3.2$  GeV for the unpolarized case and  $\mu = 3$  GeV for the helicity case in the following analysis.

Now, we can perform an independent analysis that avoids the usage of RI-MOM procedure completely and

compare the outcome to the prediction for  $\mathcal{M}(z, P_z, P'_z, \Gamma)$  from the knowledge of NNPDF and CTEQ PDF moments. We perform such a comparison in Fig. 17. For this, we use the values of  $\langle x^n \rangle(\mu)$  up to an order  $n = n_{\max}$  for NNPDF3.1 in Eq. (29) and the complete result for CT18 to obtain the phenomenological expectation for the ratio  $\mathcal{M}(z, P_z, P'_z, \gamma_t)$ . The results obtained by using the truncation order  $n_{\max} = 2, 3, 4, 20$  using the NNPDF3.1 values for  $\langle x^n \rangle$  are shown as different colored bands in Fig. 17. It is clear that inclusion of up to  $n_{\max} = 20$  moments is sufficient for convergence to the correct PDF within  $z \leq 0.5$  fm. For  $z < 0.3$  fm, which is where the lattice data has a good signal-to-noise ratio, we find that  $N = 4$  is sufficient to describe the lattice results. This gives us an idea of which moments are being probed by our lattice data at different  $z$ . We observe some discernible differences between the phenomenological expectations and our lattice  $\mathcal{M}(z, P_z, P'_z, \gamma_t)$  for  $z > 0.2$  fm, as we also observed in the case of RI-MOM scheme in Fig. 10. To understand this, we estimate the values of the moments  $\langle x^n \rangle$  that best describe our lattice data. To avoid overfitting the data, we truncate the expansion in Eq. (29) at most by  $n = 4$ . In order to avoid lattice artifacts that might be present for  $z$  of the order of lattice spacing, we fit the data only from  $z = 2a$  to a value  $z_{\max}$ . The variation of the best-fit values of  $\langle x^n \rangle$  with  $z_{\max}$  is a source of systematic error. In Fig. 18, we show the  $z_{\max}$  dependence of our estimates for  $\langle x^1 \rangle$ ,  $\langle x^2 \rangle$ ,  $\langle x^3 \rangle$ , and  $\langle x^4 \rangle$ . From Fig. 17, we note the noisy determination of the imaginary part of  $\mathcal{M}$ . As a consequence, we find our estimates of  $\langle x^1 \rangle$  and  $\langle x^3 \rangle$  to be noisy as well. On the contrary, we are able to determine  $\langle x^2 \rangle$  and  $\langle x^4 \rangle$  reasonably well. In addition to  $z_{\max}$  dependence, we also study whether our determination of the moments is affected by the order of truncation used in Eq. (29). We observe no significant variations with truncation. For comparison, the NNPDF and CT18 values of these moments are shown by the horizontal lines. Further, when we fix the values of  $\langle x^1 \rangle$

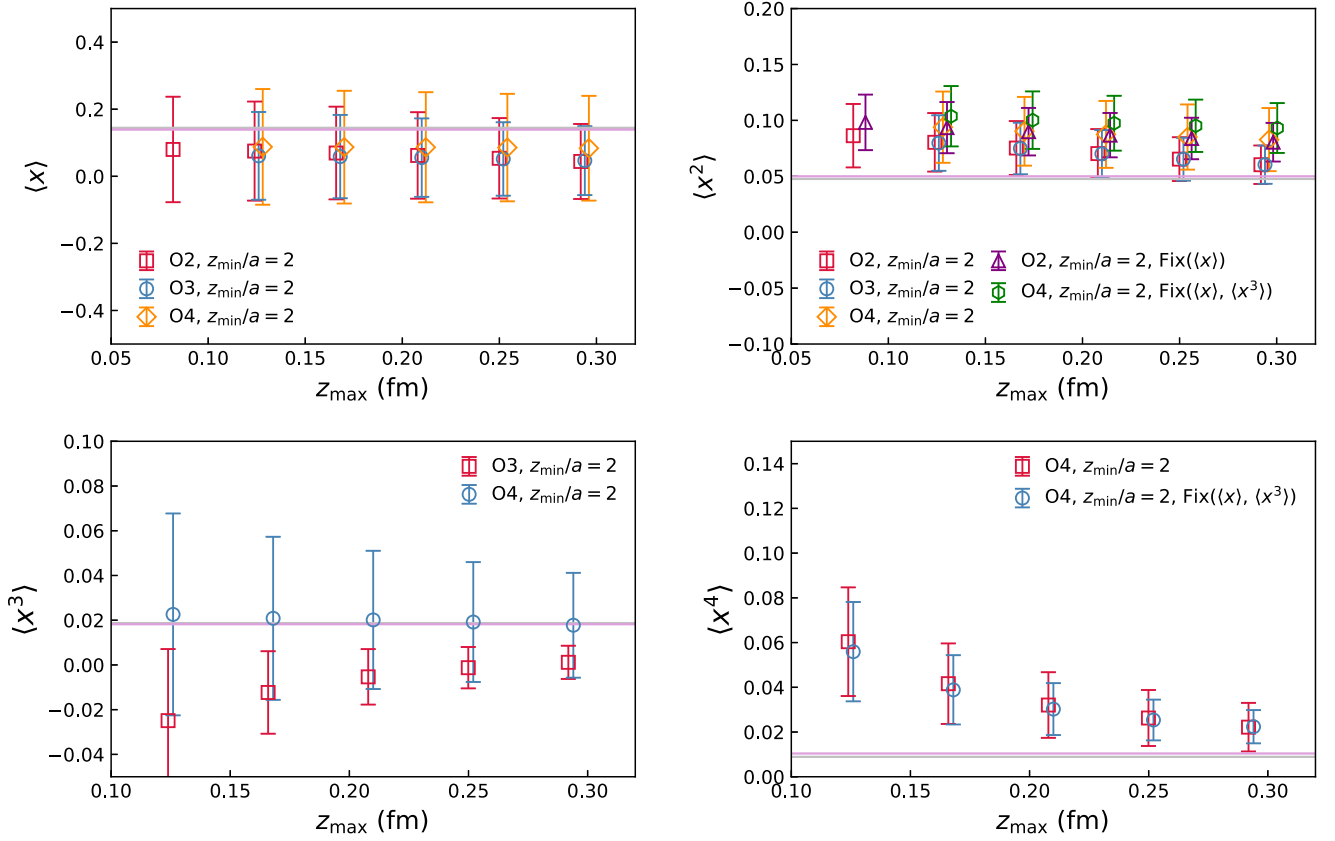


FIG. 18. The moments of the isovector unpolarized PDF,  $\langle x^1 \rangle$  (top left),  $\langle x^2 \rangle$  (top right),  $\langle x^3 \rangle$  (bottom left), and  $\langle x^4 \rangle$  (bottom right) that best describe the ratio  $\mathcal{M}(z, P_z, P'_z, \gamma_t)$  with  $P_z = 2.31$  GeV and  $P'_z = 1.84$  GeV. In each of the panels, the moment  $\langle x^n \rangle$  is shown as a function of  $z_{\max}$  of the fit using the functional form in Eq. (29) over a range  $[2a, z_{\max}]$  of the data. The results from fits using only moments up to  $n = 2$  as free parameters in Eq. (29) are labeled “O2,” and those up to  $n = 4$  are labeled “O4.” The results from fits that fix the moment  $\langle x^1 \rangle$ , or  $\langle x^1 \rangle$  and  $\langle x^3 \rangle$ , to their global fit values are also shown. For comparisons, results from CT18 and NNPDF3.1 are shown as horizontal lines.

and  $\langle x^3 \rangle$  from NNPDF to reduce the number of fit parameters, we find the estimates for  $\langle x^2 \rangle$  to be slightly elevated in value and in the direction away from the NNPDF or CT18

value. To a small extent, this is seen in  $\langle x^4 \rangle$  as well. Thus, the observed difference between our lattice result and the NNPDF and CT18 results could be attributed to this tendency

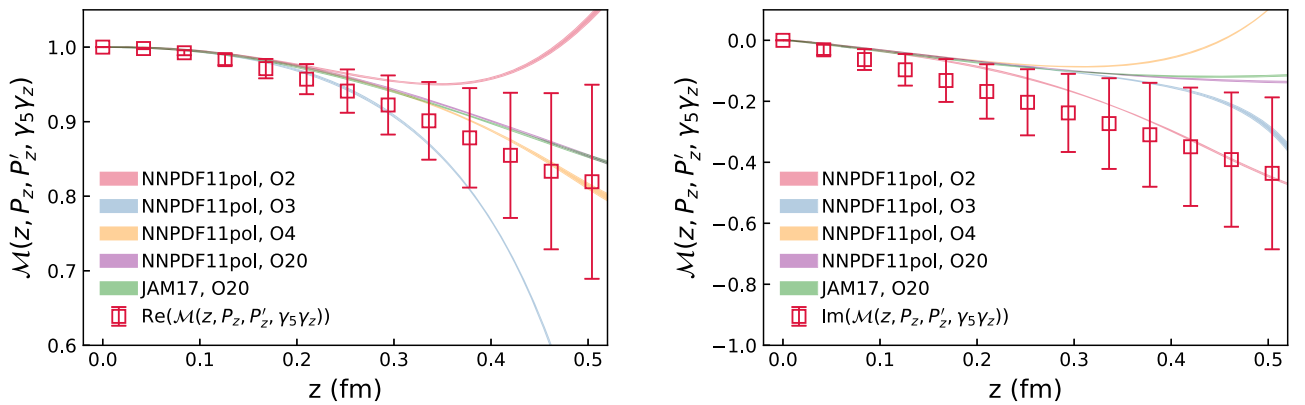


FIG. 19. The real (left) and imaginary (right) parts of  $\mathcal{M}(z, P_z, P'_z, \gamma_5 \gamma_z)$  are shown for  $P_z = 2.31$  GeV and  $P'_z = 1.84$  GeV. The data points are from our lattice calculations, whereas the various colored bands are the corresponding results from the isovector helicity PDFs from NNPDF1.1pol and JAM17. For NNPDF1.1pol, we also show results by truncating the expansion in Eq. (29) at various orders,  $n = n_{\max}$ , in the PDF moments; these results are denoted by “O $_{n_{\max}}$ ”.

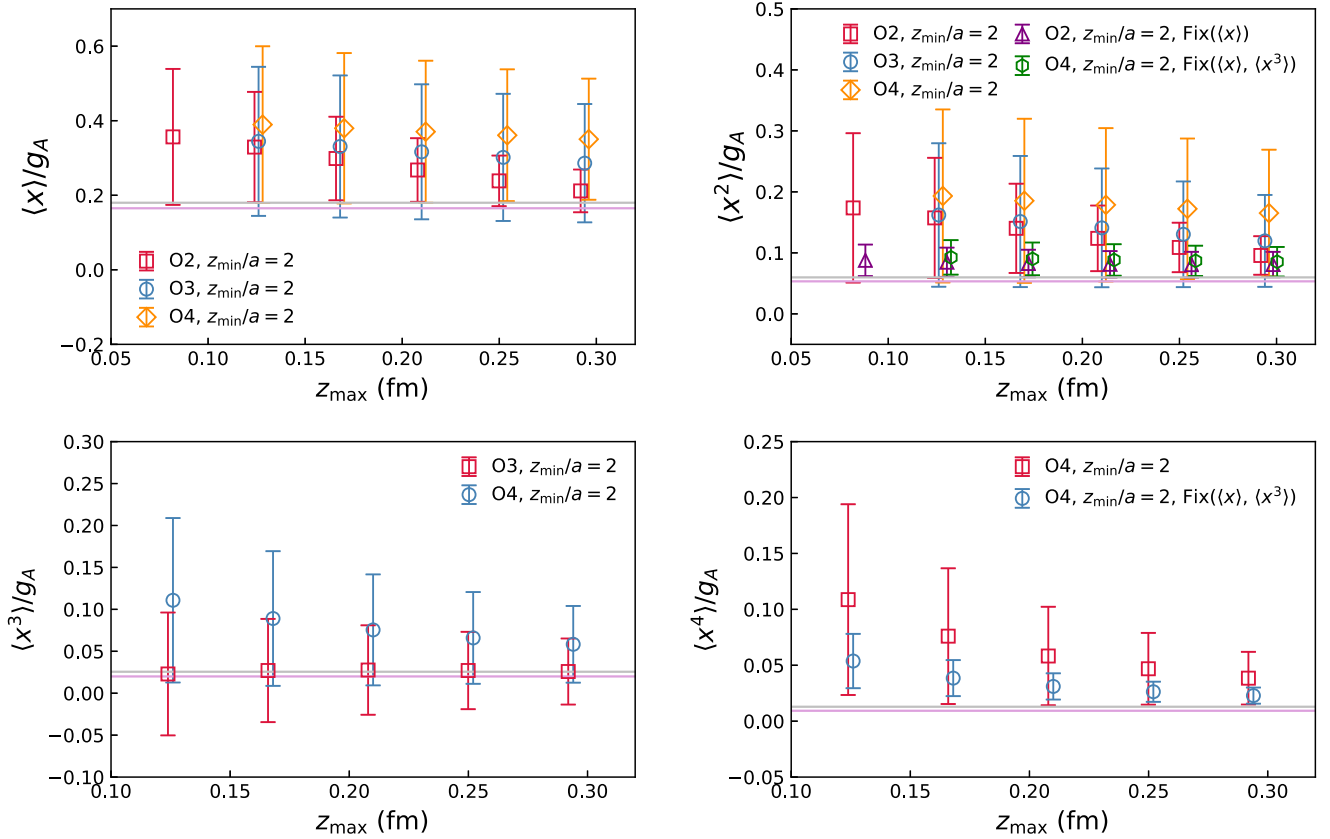


FIG. 20. The moments of the helicity PDF,  $\langle x^1 \rangle$  (top left),  $\langle x^2 \rangle$  (top right),  $\langle x^3 \rangle$  (bottom left), and  $\langle x^4 \rangle$  (bottom right), that best describe the ratio  $\mathcal{M}(z, P_z, P'_z, \gamma_5 \gamma_z)$  with  $P_z = 2.31$  GeV and  $P'_z = 1.84$  GeV. In each of the panels, the moment  $\langle x^n \rangle$  is shown as a function of  $z_{\max}$  of the fit using the functional form in Eq. (29) over a range  $[2a, z_{\max}]$  of the data. The results from fits using only moments up to  $n = 2$  as free parameters in Eq. (29) are labeled “O2,” and those up to  $n = 4$  are labeled “O4.” The results from fits that fix the moment  $\langle x^1 \rangle$ , or  $\langle x^1 \rangle$  and  $\langle x^3 \rangle$ , to their global fit values are also shown. For comparisons, results from JAM17 and NNPDF1.1pol are shown as horizontal lines.

for our lattice values of  $\langle x^2 \rangle$ ,  $\langle x^4 \rangle$  to be slightly higher than the corresponding phenomenological values.

We repeat a similar analysis for the helicity matrix element,  $\Gamma = \gamma_5 \gamma_z$ . In this case, the Wilson coefficients  $c_n(\mu z)$  are the same as in the case of the unpolarized case with  $\Gamma = \gamma_z$ . Since we are setting the value of the matrix elements at  $z = 0$  to be 1 through the ratio, we only obtain the values of  $\langle x^n \rangle / \langle x^0 \rangle$  in the expansion in Eq. (29), with  $\langle x^0 \rangle = g_A$ . In Fig. 19, we compare the results corresponding to the NNPDF1.1pol and JAM17 values with the lattice result for the ratio. As in the case of the unpolarized matrix element, we also test the dependence of this comparison on the truncation order  $n_{\max}$ . The sensitivity to higher moments is a bit more than that for the unpolarized case, and we find convergence at only  $n_{\max} = 6$  at  $z < 0.3$  fm. Surprisingly, the global fit expectation agrees quite well with our lattice result even though there is a little tension in the imaginary parts. As explained above, we also obtain the best-fit values of  $\langle x^1 \rangle / g_A$ ,  $\langle x^2 \rangle / g_A$ ,  $\langle x^3 \rangle / g_A$ , and  $\langle x^4 \rangle / g_A$  that describe our lattice data via Eq. (29) truncated at most by the fourth order. In Fig. 20, we show the results as a

function of the largest  $z$  used in the fits,  $z_{\max}$ . Like the unpolarized PDF case,  $\langle x^1 \rangle / g_A$  is noisy, but it seems to agree with the global fit results. The more precisely determined value of  $\langle x^2 \rangle / g_A$  is quite robust to various ways of fitting the data and agrees nicely with the global fit values. To compare with other lattice calculations, we truncate the expansion in Eq. (29) at  $n = 2$ , and estimate  $\langle x^1 \rangle / g_A$  at  $\mu = 2$  GeV with the  $z$  in the range  $[2a, 0.3]$  fm. Our result  $\langle x \rangle / g_A = 0.219(56)$  is compatible with the ETMC result [64]  $0.229(30)/1.242(57)$  within the error.

## IX. SUMMARY AND CONCLUSIONS

In this paper, we studied isovector unpolarized and helicity PDFs of the proton using the LaMET approach. The lattice calculations have been performed for an unphysically large pion mass of 310 MeV. On the other hand, our lattice study was carried out using lattice spacing  $a = 0.042$  fm, which is the smallest lattice spacing used in such studies. We argued that such small lattice spacing is essential for the validity of one-loop perturbative matching

between the PDF and qPDF, which is a key ingredient of LaMET.

Extracting the nucleon matrix elements for such large momenta and small lattice spacing is challenging because of poor signal-to-noise ratio. To deal with this problem, we performed a detailed study of the nucleon two-point function with a momentum-smeared source and sink, as well as with a momentum-smeared source and point sink to better control the excited-state contributions. We demonstrated that the ground state can be reliably isolated up to the highest momenta used in this study. Furthermore, for the Euclidean-time separations used that are relevant for our lattice analysis, the two-point function is very well described by the ground state and an “effective” excited-state contribution, with an energy that is larger than the true excited-state energy. Therefore, we argued that the two-state *Ansätze* are sufficient to describe the dependence of the three-point function on the source-sink separation and on the operator insertion time. We showed that the qPDF matrix elements can be extracted in this way, and the results do not depend on the choices of the fit interval used in our study, demonstrating the robustness of our analysis procedure.

After nonperturbative RI-MOM renormalizations, we compared the lattice calculations of the spatial  $z$  dependence of qPDFs with those from the phenomenological PDFs, obtained from the global pQCD-based analyses of pertinent experimental data performed by different collaborations. Working in  $z$ -space allowed us to test the LaMET approach. The comparisons showed that there is a rough agreement between the lattice results and the results of global analysis, but only at quite small distances. Even for the very small lattice spacing used in this study, there was not enough data points to constrain the  $x$  dependence of the PDFs. Instead, to translate our  $z$ -space comparisons to  $x$  dependence, we introduced a new ratio-based renormalization scheme for the Ioffe-time distributions. Using our

lattice calculations for Ioffe-time distributions, renormalized via this new ratio-based scheme, we determined the first moments of the isovector unpolarized and helicity PDFs of the proton, and compared these moments with those from the corresponding phenomenological PDFs.

## ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics through Contract No. DE-SC0012704; by the U.S. Department of Energy, Office of Science, Office of Nuclear Physics and Office of Advanced Scientific Computing Research within the framework of Scientific Discovery through Advance Computing (SciDAC) award Computing the Properties of Matter with Leadership Computing Resources; and by the Brookhaven National Laboratory’s Laboratory-Directed Research and Development (LDRD) project No. 16-37. The work of Z. F., R. L., H. L., Y. Y., and R. Z. are supported by the U.S. National Science Foundation under Grant No. PHY 1653405 CAREER: Constraining Parton Distribution Functions for New-Physics Searches. The work of X. G. is partially supported by NSFC Grant No. 11890712. S. S. also acknowledges support by the RHIC Physics Fellow Program of the RIKEN BNL Research Center and by the National Science Foundation under CAREER Grant No. PHY-1847893. This research used awards of computer time provided by the INCITE program at the Oak Ridge Leadership Computing Facility, a DOE Office of Science User Facility operated under Contract No. DE-AC05-00OR22725. We are grateful to Yong Zhao for advice on perturbative matching of the helicity qPDF.

## APPENDIX A: ANALYSIS OF THE NUCLEON TWO-POINT FUNCTION

In this appendix, we discuss some details of the analysis of the SP and SS two-point correlators. In Fig. 21, we show

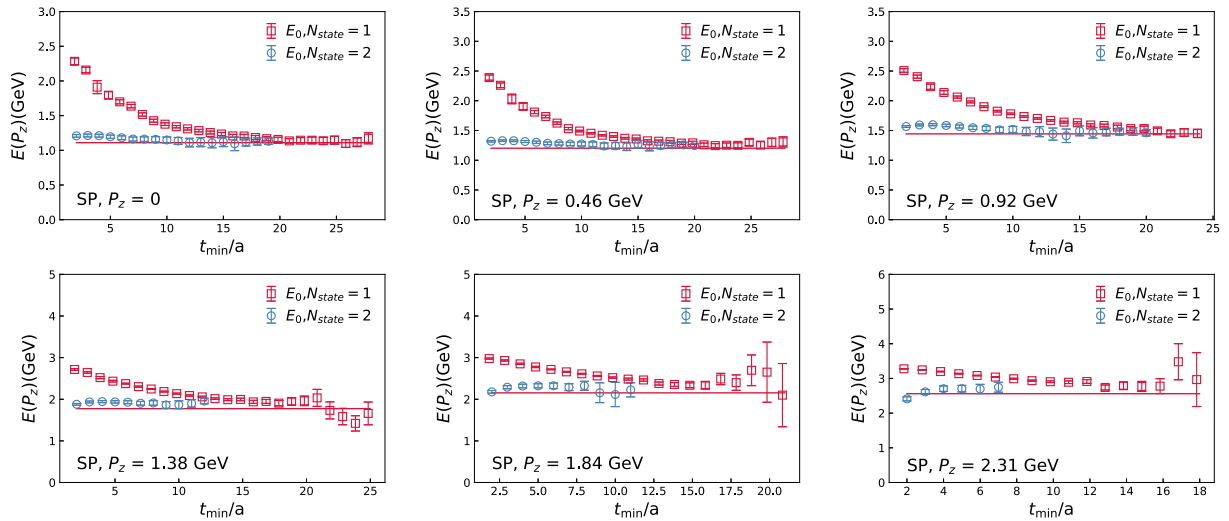


FIG. 21. Ground-state energy from unconstrained one-state and two-state fits of the SP correlators.

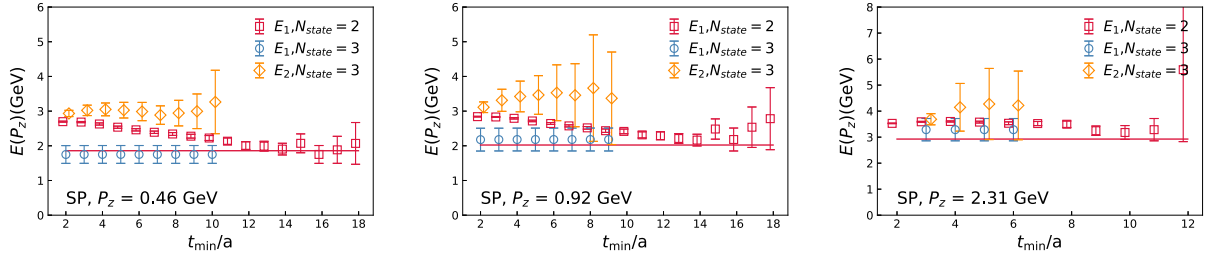


FIG. 22. The energies of the first ( $E_1$ ) and second ( $E_2$ ) excited states from constrained two-state and three-state fits of the SP correlator for  $n_z = 1$  (left),  $n_z = 2$  (middle), and  $n_z = 5$  (right). The horizontal lines denote the values calculated from the dispersion relation.

the ground-state energy from one- and two-exponential fits of SP correlators as a function of  $t_{\min}$ . Contrary to the fits of the SS correlators' stable result for the ground-state energy,  $E_0$  is only obtained for  $t_{\min} \geq 20$ .

As discussed in the main text, we performed prior-based fits of SP and SS correlators for all values of  $p_z$ . In Fig. 22, we show the results on  $E_1(p_z)$  for  $n_z = 1, 2$ , and 5 for prior-based fits of the SP correlator.

We see clearly that  $E_1$  approaches the value expected from the dispersion relation for  $t_{\min} > 11$  if the two-

exponential fit is used. For constrained three-exponential fits, the same value is approached for  $t_{\min} = 2$ . In Fig. 23, we show the amplitudes,  $A_i$ ,  $i = 1, 2, 3, \dots$ , of different states normalized by the value of the two-point correlator at  $t = 0$ , which by definition is equal to  $\sum_i A_i$ . We see that  $A_1$  is slightly higher than  $A_0$ , while  $A_2$  is significantly larger than either  $A_0$  or  $A_1$ .

A similar analysis was performed for SS correlators, and the results for the excited state energies and amplitudes are shown in Figs. 24 and 25, respectively. From these figures,

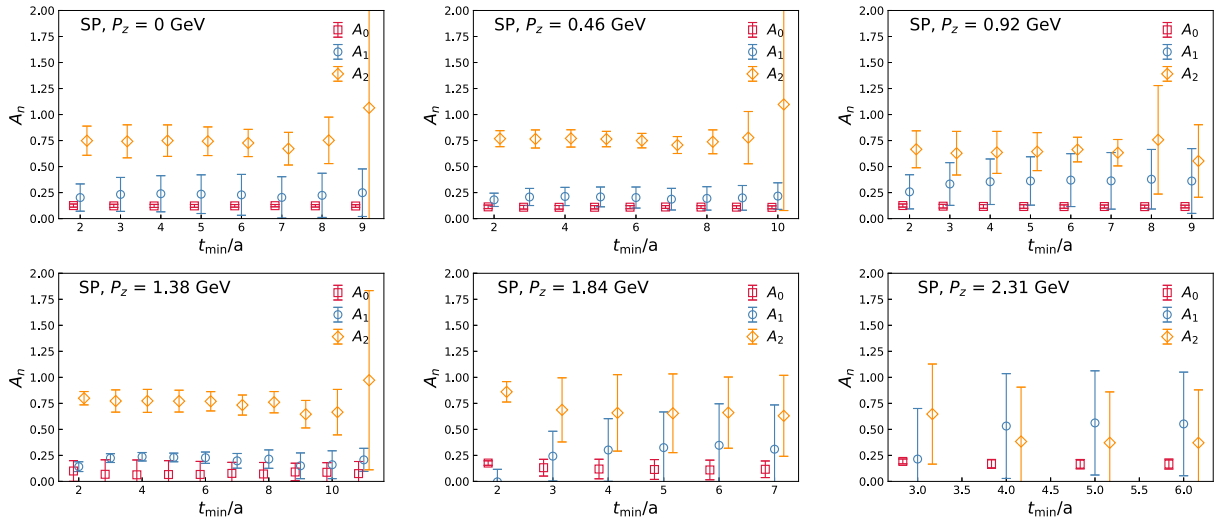


FIG. 23. The amplitudes of different states obtained from the constrained three-state fit of the SP correlator and normalized by  $C_{2pt}^{SP}(t=0)$  as a function of temperature.

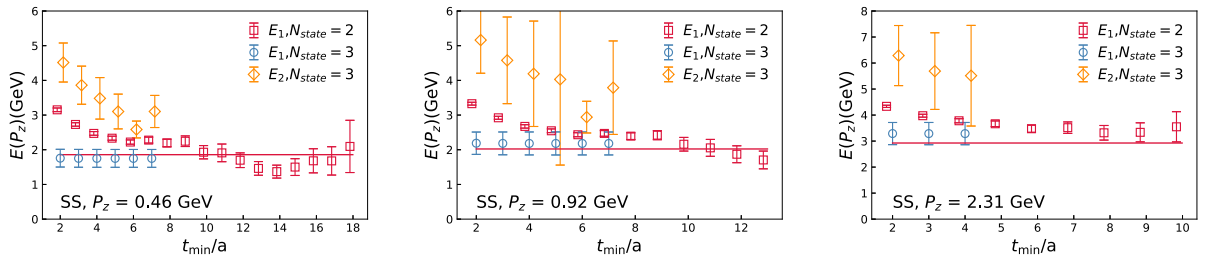


FIG. 24. The energies of the first ( $E_1$ ) and second ( $E_2$ ) excited states from constrained two-state and three-state fits of the SS correlator for  $n_z = 1$  (left),  $n_z = 2$  (middle), and  $n_z = 5$  (right). The horizontal lines denote the values calculated from the dispersion relation.

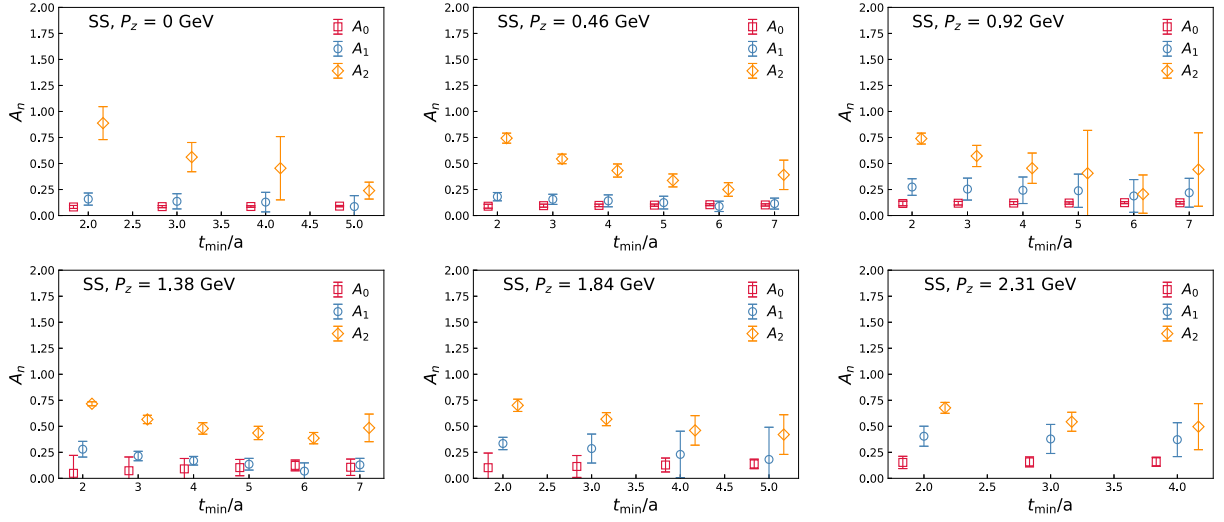


FIG. 25. The amplitudes of different states obtained from the constrained three-state fit of the SS correlator and normalized by  $C_{2pt}^{SS}(t=0)$  as a function of temperature.

we see that a pseudoplateau develops for the first excited states for  $5 < t_{\min} < 10$  of the two-state fit. We see that  $A_0$  and  $A_1$  are similar in this case, and  $A_2$  decreases as  $t_{\min}$  increases.

## APPENDIX B: ANALYSIS OF THE THREE-POINT FUNCTION

In this appendix, we discuss further details of the extraction of the bare matrix element of the qPDF operator. First, we show our results for the ratio of the three-point function to the two-point function for different source-sink separation and different values of  $z$  as a function of the operator insertion time  $\tau$  in Fig. 26 for  $n_z = 4$ . In this figure, we also show the results for  $R_1^{\text{fit}}$ . As one can see from

the figure,  $R_1^{\text{fit}}$  can describe the data well for all values of  $t$ . In Fig. 27, we show the same analysis but for  $n_z = 5$ .

As discussed in the main text, we perform  $R_1^{\text{fit}}$  using a single value of source-sink separation. The results are shown in Fig. 28 for the real part of the matrix element. As one can see from the figure, the results obtained from this fit for  $t = 16, 18$  and  $20$  agree within errors. We performed fits using the form  $fit_1$  with  $\tau > \tau_{\min}$  and taking the value of  $E_1$  from the two-point function fit with  $t > t_{\min}$ . The results are shown in Fig. 29. We see no significant dependence on  $\tau_{\min}$  and  $t_{\min}$ .

Another way to obtain the matrix element is to use the summation method. The summation method is illustrated in Fig. 30 for  $n_z = 4$ . The results obtained from the summation method agree with those from  $R_1^{\text{fit}}$  but have much larger

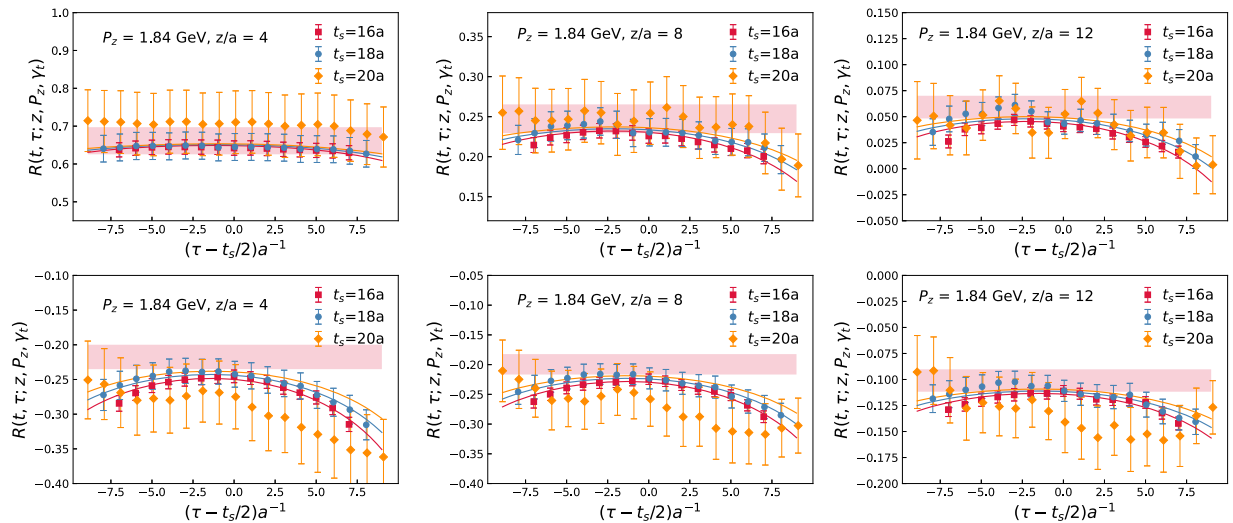


FIG. 26. The ratio of the three-point function to the two-point function for  $z = 4, 8, 12$  and  $n_z = 4$ . The upper panels show the real part, while the imaginary part is shown in the lower panels. The results of  $R_1^{\text{fit}}$  are shown as lines.

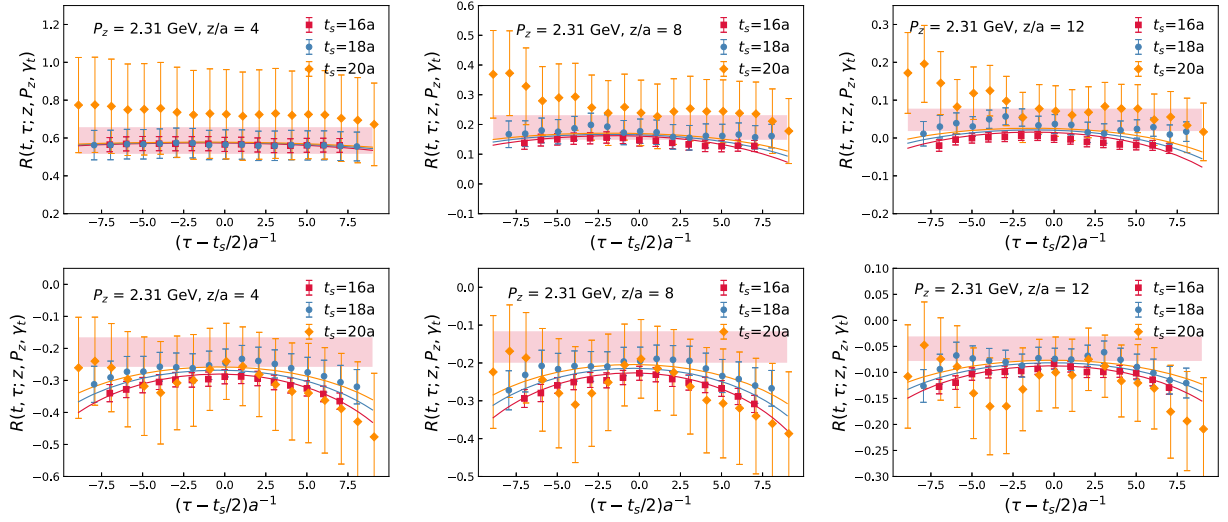


FIG. 27. The ratio of the three-point function to the two-point function for  $z = 4, 8, 12$  and  $n_z = 5$ . The upper panels show the real part, while the imaginary part is shown in the lower panels. The results of  $R_1^{\text{fit}}$  are shown as lines.

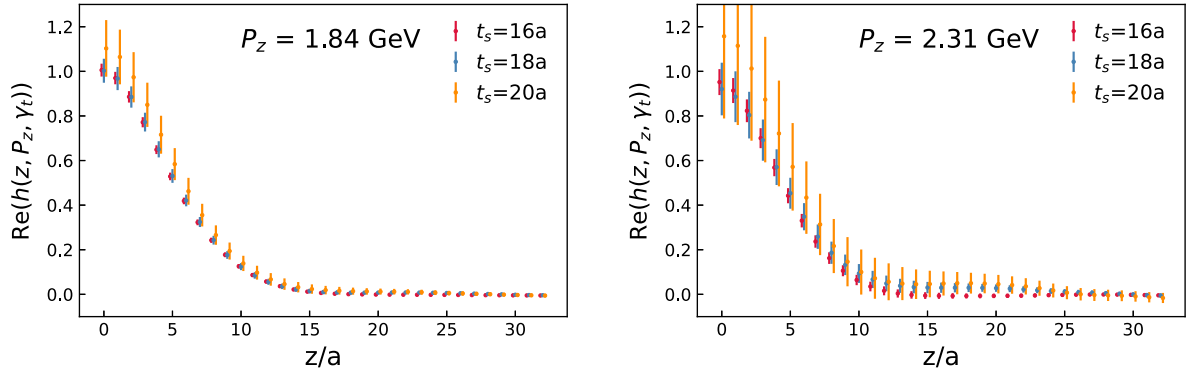


FIG. 28. The  $z$  dependence of the qPDF matrix element obtained using  $R_1^{\text{fit}}$  with a single value of the source-sink separation for  $n_z = 4, 5$ .

errors. The statistical errors of the  $n_z = 5$  data are too large to use the summation method. Furthermore, we could also reduce the error in the summation method by dividing by the matrix element at  $z = 0$ , as can be seen in Fig. 31.

Similar analysis of the ratio of the three-point function to the two-point function is carried out for longitudinally

polarized qPDF operator. The results are summarized in Figs. 32, 33, and 34.

To take advantage of correlation between different  $z$ 's and cancel the field renormalization factor, we divided the bare matrix elements by the matrix element at  $z = 0$ . The errors are much smaller after this division, as discussed in the main text.

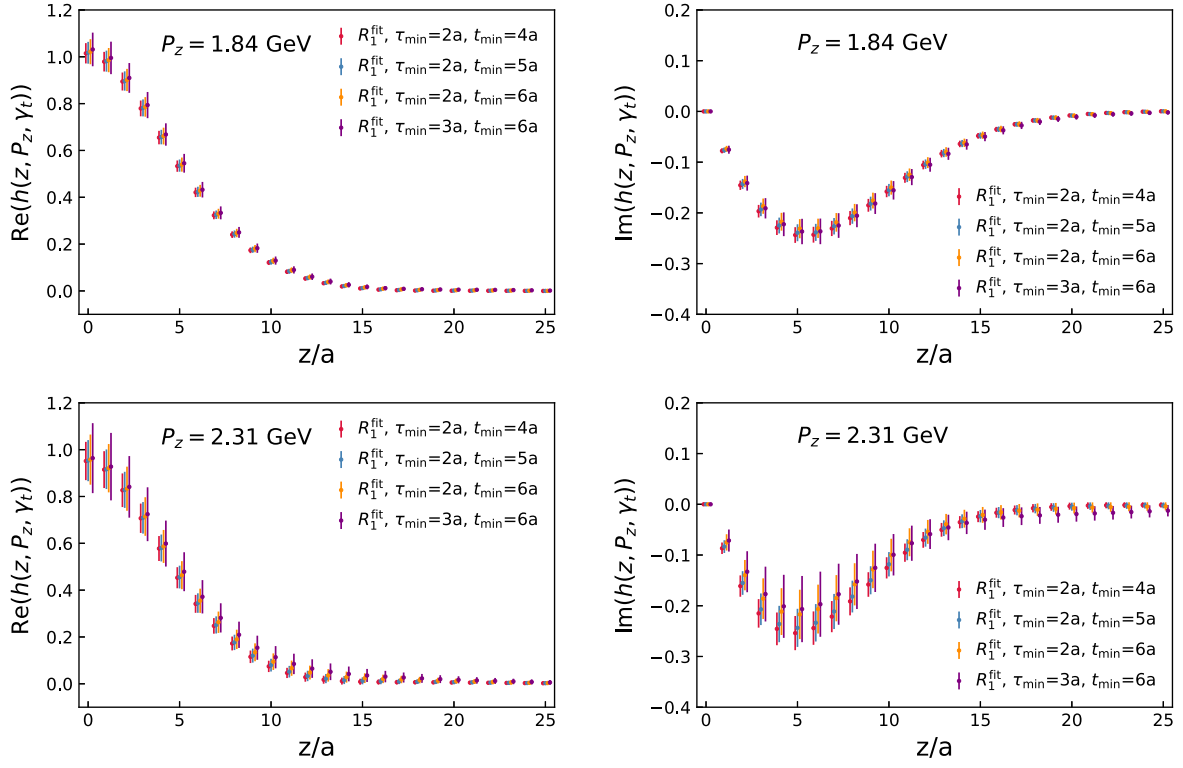


FIG. 29. Real (left) and imaginary (right) parts of the bare matrix as a function of  $z$ . The top panel shows the results for  $n_z = 4$ , and the bottom panel shows the results for  $n_z = 5$ . The results for different choices of  $\tau_{\min}$  and  $t_{\min}$  in the two-point function fits are shown.

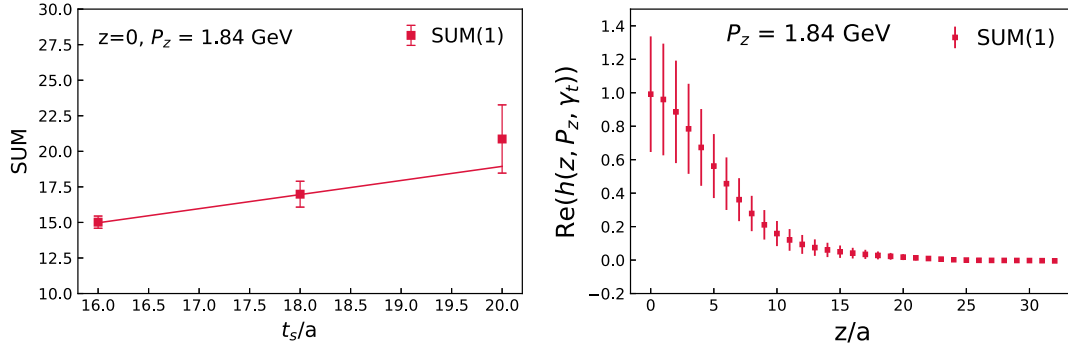


FIG. 30. The  $\tau$  dependence of the sum of the ratio of the three-point function to the two-point function (left) and the  $z$  dependence of the matrix element extracted from the summation method (right). SUM( $n$ ) means the summation fit with  $n$  skipped time insertion.

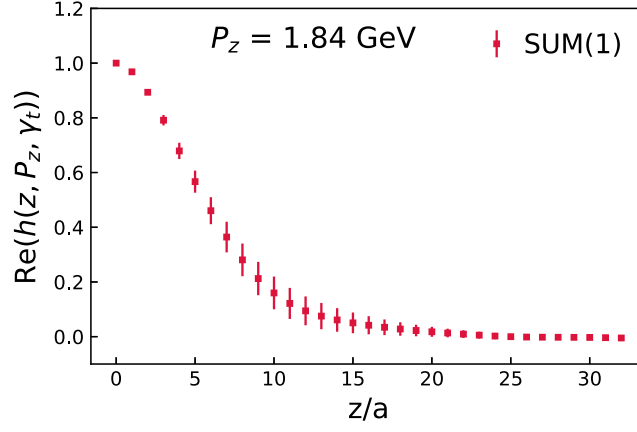


FIG. 31. The  $z$  dependence of the real part of the bare qPDF matrix element obtained by the summation method after division by the matrix element for  $z = 0$  at  $n_z = 4$ . SUM( $n$ ) means the summation fit with  $n$  skipped time insertion.

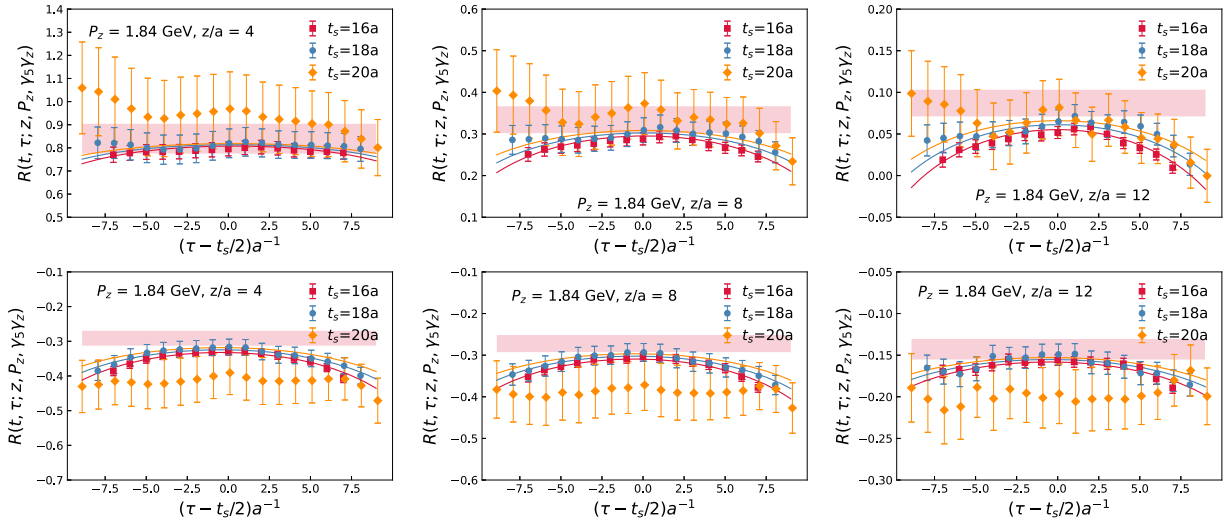


FIG. 32. The ratio of the three-point function to the two-point function corresponding to the helicity qPDF for  $z = 4, 8, 12$  and  $n_z = 4$ . The upper panels show the real part, while the imaginary part is shown in the lower panels. The results of  $R_1^{\text{fit}}$  are shown as lines.

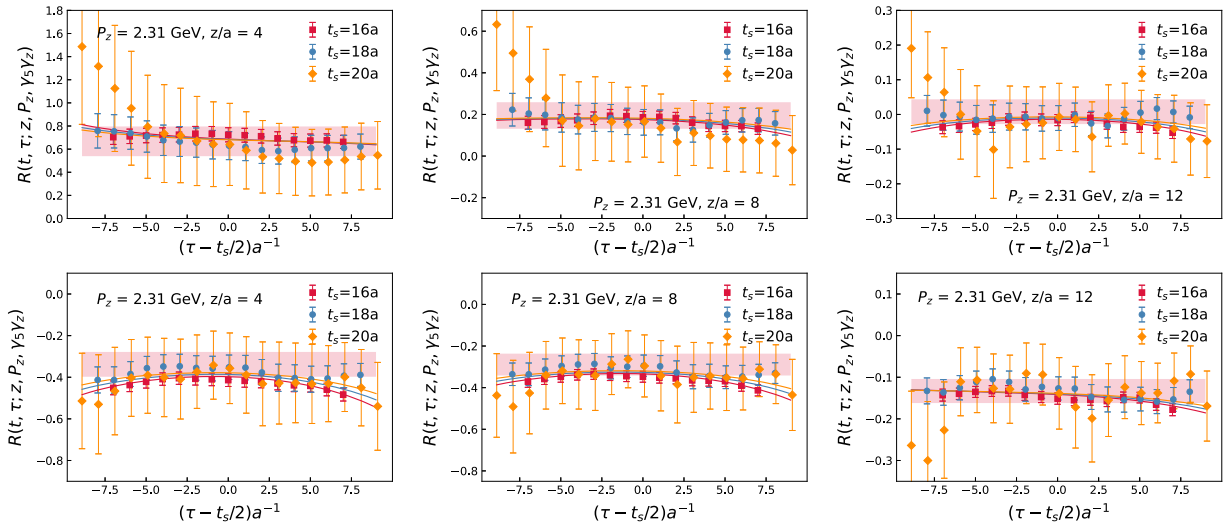


FIG. 33. The ratio of the three-point function to the two-point function corresponding to the helicity qPDF for  $z = 4, 8, 12$  and  $n_z = 5$ . The upper panels show the real part, while the imaginary part is shown in the lower panels. The results of  $R_1^{\text{fit}}$  are shown as lines.

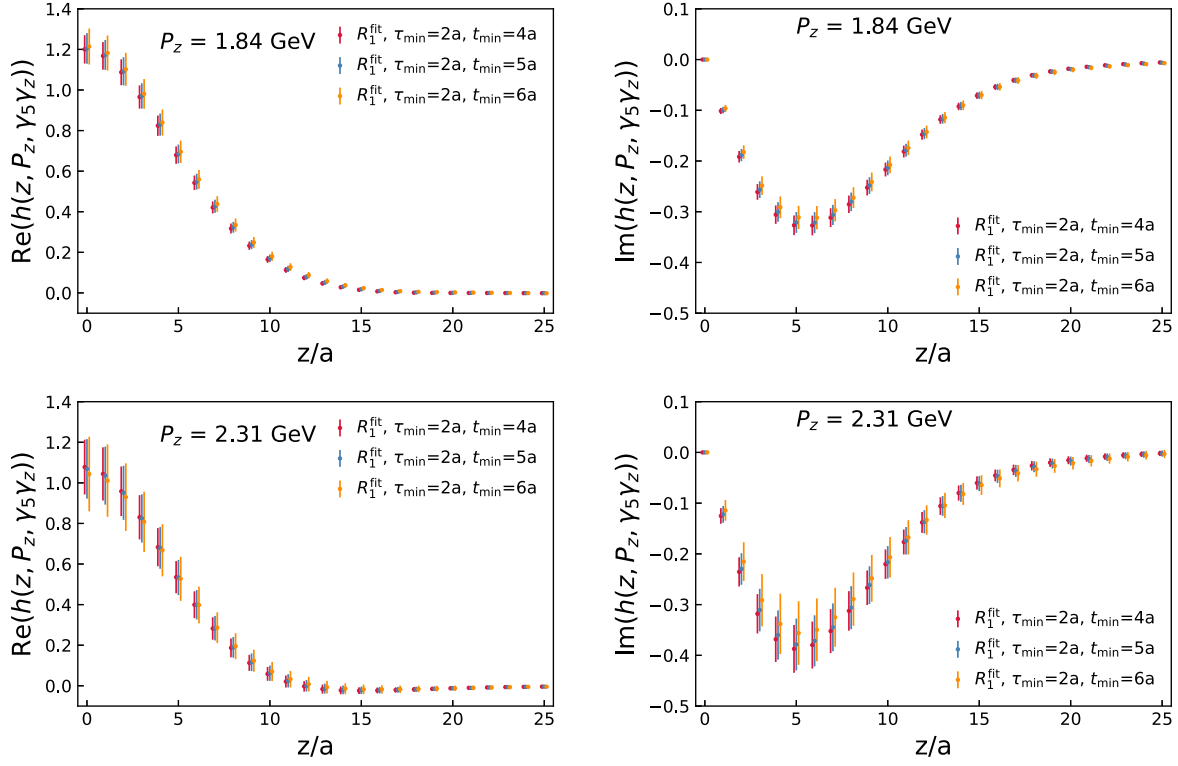


FIG. 34. The real (left) and the imaginary (right) parts of the bare matrix corresponding to the helicity qPDF. The upper panels correspond to  $n_z = 4$ , while the lower panels correspond to  $n_z = 5$ . See text for further details.

- 
- [1] D. de Florian, R. Sassot, M. Stratmann, and W. Vogelsang, *Phys. Rev. Lett.* **113**, 012001 (2014).
- [2] L. Adamczyk *et al.* (STAR Collaboration), *Phys. Rev. Lett.* **115**, 092002 (2015).
- [3] A. Adare *et al.* (PHENIX Collaboration), *Phys. Rev. D* **90**, 012007 (2014).
- [4] A. Adare *et al.* (PHENIX Collaboration), *Phys. Rev. D* **93**, 011501 (2016).
- [5] C. Adolph *et al.* (COMPASS Collaboration), *Phys. Rev. D* **87**, 052018 (2013).
- [6] L. Adamczyk *et al.* (STAR Collaboration), *Phys. Rev. Lett.* **113**, 072301 (2014).
- [7] A. Adare *et al.* (PHENIX Collaboration), *Phys. Rev. D* **93**, 051103 (2016).
- [8] J. Dudek *et al.*, *Eur. Phys. J. A* **48**, 187 (2012).
- [9] A. Accardi *et al.*, *Eur. Phys. J. A* **52**, 268 (2016).
- [10] S. Chatrchyan *et al.* (CMS Collaboration), *Science* **338**, 1569 (2012).
- [11] G. Aad *et al.* (ATLAS Collaboration), *Science* **338**, 1576 (2012).
- [12] X. Ji, *Phys. Rev. Lett.* **110**, 262002 (2013).
- [13] X. Ji, *Sci. China Phys. Mech. Astron.* **57**, 1407 (2014).
- [14] Y.-Q. Ma and J.-W. Qiu, *Phys. Rev. D* **98**, 074021 (2018).
- [15] Y.-Q. Ma and J.-W. Qiu, *Phys. Rev. Lett.* **120**, 022003 (2018).
- [16] A. V. Radyushkin, *Phys. Rev. D* **96**, 034025 (2017).
- [17] K. Orginos, A. Radyushkin, J. Karpie, and S. Zafeiropoulos, *Phys. Rev. D* **96**, 094503 (2017).
- [18] X. Ji, J.-H. Zhang, and Y. Zhao, *Phys. Rev. Lett.* **120**, 112001 (2018).
- [19] T. Ishikawa, Y.-Q. Ma, J.-W. Qiu, and S. Yoshida, *Phys. Rev. D* **96**, 094019 (2017).
- [20] J.-W. Chen, T. Ishikawa, L. Jin, H.-W. Lin, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, *Phys. Rev. D* **97**, 014505 (2018).
- [21] M. Constantinou and H. Panagopoulos, *Phys. Rev. D* **96**, 054506 (2017).
- [22] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, H. Panagopoulos, and F. Steffens, *Nucl. Phys. B* **923**, 394 (2017).
- [23] I. W. Stewart and Y. Zhao, *Phys. Rev. D* **97**, 054512 (2018).
- [24] Y.-S. Liu *et al.* (Lattice Parton Collaboration), *Phys. Rev. D* **101**, 034020 (2020).
- [25] T. Izubuchi, X. Ji, L. Jin, I. W. Stewart, and Y. Zhao, *Phys. Rev. D* **98**, 056004 (2018).
- [26] Y.-S. Liu, J.-W. Chen, L. Jin, R. Li, H.-W. Lin, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, *arXiv:1810.05043*.

- [27] H.-W. Lin, J.-W. Chen, X. Ji, L. Jin, R. Li, Y.-S. Liu, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, *Phys. Rev. Lett.* **121**, 242003 (2018).
- [28] J.-W. Chen, L. Jin, H.-W. Lin, Y.-S. Liu, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, [arXiv:1803.04393](#).
- [29] C. Alexandrou, K. Cichy, M. Constantinou, K. Hadjiyiannakou, K. Jansen, A. Scapellato, and F. Steffens, *Phys. Rev. D* **99**, 114504 (2019).
- [30] C. Alexandrou, K. Cichy, M. Constantinou, K. Jansen, A. Scapellato, and F. Steffens, *Phys. Rev. D* **98**, 091503 (2018).
- [31] C. Alexandrou, K. Cichy, M. Constantinou, K. Jansen, A. Scapellato, and F. Steffens, *Phys. Rev. Lett.* **121**, 112001 (2018).
- [32] B. Joó, J. Karpie, K. Orginos, A. Radyushkin, D. Richards, and S. Zafeiropoulos, *J. High Energy Phys.* **12** (2019) 081.
- [33] B. Jo, J. Karpie, K. Orginos, A. V. Radyushkin, D. G. Richards, and S. Zafeiropoulos, [arXiv:2004.01687](#).
- [34] J.-H. Zhang, J.-W. Chen, L. Jin, H.-W. Lin, A. Schr, and Y. Zhao, *Phys. Rev. D* **100**, 034505 (2019).
- [35] T. Izubuchi, L. Jin, C. Kallidonis, N. Karthik, S. Mukherjee, P. Petreczky, C. Shugert, and S. Syritsyn, *Phys. Rev. D* **100**, 034516 (2019).
- [36] R. S. Sufian, J. Karpie, C. Egerer, K. Orginos, J.-W. Qiu, and D. G. Richards, *Phys. Rev. D* **99**, 074507 (2019).
- [37] B. Joó, J. Karpie, K. Orginos, A. V. Radyushkin, D. G. Richards, R. S. Sufian, and S. Zafeiropoulos, *Phys. Rev. D* **100**, 114512 (2019).
- [38] R. S. Sufian, C. Egerer, J. Karpie, R. G. Edwards, B. Joó, Y.-Q. Ma, K. Orginos, J.-W. Qiu, and D. G. Richards, *Phys. Rev. D* **102**, 054508 (2020).
- [39] Y. Zhao, *Int. J. Mod. Phys. A* **33**, 1830033 (2018).
- [40] K. Cichy and M. Constantinou, *Adv. High Energy Phys.* **2019**, 3036904 (2019).
- [41] C. Monahan, *Proc. Sci.*, LATTICE2018 (2018) 018 [[arXiv:1811.00678](#)].
- [42] X. Ji, Y.-S. Liu, Y. Liu, J.-H. Zhang, and Y. Zhao, [arXiv:2004.03543](#).
- [43] E. Follana, Q. Mason, C. Davies, K. Hornbostel, G. P. Lepage, J. Shigemitsu, H. Trottier, and K. Wong (HPQCD, UKQCD Collaborations), *Phys. Rev. D* **75**, 054502 (2007).
- [44] A. Bazavov *et al.* (MILC Collaboration), *Phys. Rev. D* **87**, 054505 (2013).
- [45] A. Hasenfratz and F. Knechtli, *Phys. Rev. D* **64**, 034504 (2001).
- [46] R. Gupta, Y.-C. Jang, H.-W. Lin, B. Yoon, and T. Bhattacharya, *Phys. Rev. D* **96**, 114503 (2017).
- [47] T. Bhattacharya, V. Cirigliano, S. Cohen, R. Gupta, A. Joseph, H.-W. Lin, and B. Yoon (PNDME Collaboration), *Phys. Rev. D* **92**, 094511 (2015).
- [48] T. Bhattacharya, V. Cirigliano, R. Gupta, H.-W. Lin, and B. Yoon, *Phys. Rev. Lett.* **115**, 212002 (2015).
- [49] T. Bhattacharya, S. D. Cohen, R. Gupta, A. Joseph, H.-W. Lin, and B. Yoon, *Phys. Rev. D* **89**, 094502 (2014).
- [50] R. Babich, J. Brannick, R. C. Brower, M. A. Clark, T. A. Manteuffel, S. F. McCormick, J. C. Osborn, and C. Rebbi, *Phys. Rev. Lett.* **105**, 201602 (2010).
- [51] J. C. Osborn, R. Babich, J. Brannick, R. C. Brower, M. A. Clark, S. D. Cohen, and C. Rebbi, *Proc. Sci.*, LATTICE2010 (2010) 037 [[arXiv:1011.2775](#)].
- [52] R. G. Edwards and B. Joo (SciDAC LHPC, UKQCD Collaborations), *Nucl. Phys. B, Proc. Suppl.* **140**, 832 (2005).
- [53] G. S. Bali, B. Lang, B. U. Musch, and A. Schaefer, *Phys. Rev. D* **93**, 094515 (2016).
- [54] J.-W. Chen, T. Ishikawa, L. Jin, H.-W. Lin, Y.-B. Yang, J.-H. Zhang, and Y. Zhao, *Chin. Phys. C* **43**, 103101 (2019).
- [55] W. Wang, J.-H. Zhang, S. Zhao, and R. Zhu, *Phys. Rev. D* **100**, 074509 (2019).
- [56] J.-H. Zhang, X. Ji, A. Schr, W. Wang, and S. Zhao, *Phys. Rev. Lett.* **122**, 142001 (2019).
- [57] T.-J. Hou *et al.*, [arXiv:1912.10053](#).
- [58] R. D. Ball *et al.* (NNPDF Collaboration), *Eur. Phys. J. C* **77**, 663 (2017).
- [59] J.-W. Chen, S. D. Cohen, X. Ji, H.-W. Lin, and J.-H. Zhang, *Nucl. Phys. B* **911**, 246 (2016).
- [60] O. Nachtmann, *Nucl. Phys.* **B63**, 237 (1973).
- [61] B. L. Ioffe, *Phys. Lett.* **30B**, 123 (1969).
- [62] E. R. Nocera, R. D. Ball, S. Forte, G. Ridolfi, and J. Rojo (NNPDF Collaboration), *Nucl. Phys.* **B887**, 276 (2014).
- [63] J. J. Ethier, N. Sato, and W. Melnitchouk, *Phys. Rev. Lett.* **119**, 132001 (2017).
- [64] A. Abdel-Rehim *et al.*, *Phys. Rev. D* **92**, 114513 (2015); **93**, 039904(E) (2016).